A Comprehensive Scan for Heterotic SU(5) GUT models

Lara B. Anderson,[a] **Andrei Constantin,**[b] **James Gray,**[c] **Andre Lukas**[b] **and Eran Palti**[d,e]

[a] *Center for the Fundamental Laws of Nature, Jefferson Laboratory, Harvard University,*
17 Oxford Street, Cambridge, MA 02138, U.S.A.

[b] *Rudolf Peierls Centre for Theoretical Physics, Oxford University,*
1 Keble Road, Oxford, OX1 3NP, U.K.

[c] *Arnold-Sommerfeld-Center for Theoretical Physics, Department für Physik,*
Ludwig-Maximilians-Universität München, Theresienstraße 37, 80333 München, Germany

[d] *Centre de Physique Theorique, Ecole Polytechnique, CNRS,*
91128 Palaiseau, France

[e] *Institut für Theoretische Physik, Universitat Heidelberg,*
Philosophenweg 19, D-69120 Heidelberg, Germany

E-mail: lara@physics.harvard.edu, a.constantin1@physics.ox.ac.uk,
james.gray@physik.uni-muenchen.de, lukas@physics.ox.ac.uk,
Eran.Palti@cpht.polytechnique.fr

Abstract: Compactifications of heterotic theories on smooth Calabi-Yau manifolds remain one of the most promising approaches to string phenomenology. In two previous papers, arXiv:1106.4804 and arXiv:1202.1757, large classes of such vacua were constructed, using sums of line bundles over complete intersection Calabi-Yau manifolds in products of projective spaces that admit smooth quotients by finite groups. A total of 10^{12} different vector bundles were investigated which led to 202 SU(5) Grand Unified Theory (GUT) models. With the addition of Wilson lines, these in turn led, by a conservative counting, to 2122 heterotic standard models. In the present paper, we extend the scope of this programme and perform an exhaustive scan over the same class of models. A total of 10^{40} vector bundles are analysed leading to $35,000$ SU(5) GUT models. All of these compactifications have the right field content to induce low-energy models with the matter spectrum of the supersymmetric standard model, with no exotics of any kind. The detailed analysis of the resulting vast number of heterotic standard models is a substantial and ongoing task in computational algebraic geometry.

Contents

1 Introduction and summary

Heterotic string compactifications [1–4] on Calabi-Yau threefolds have provided one of the most promising approaches to string phenomenology for almost three decades. Several approaches have been proposed and used over the years: smooth Calabi-Yau compactifications based on the standard embedding [4–8], non-standard embedding models [9–20], models based on orbifolds [21–30], free fermionic strings [31–36] and Gepner models [37–39]. In the present paper, we display the latest results of a large scale model building

programme in the context of smooth Calabi-Yau compactifications of the heterotic string. This programme, aimed at achieving more detailed phenomenology than has to date be possible in this context, was initiated in the publications [40, 41] (for an early example of a systematic scan over heterotic models in the free fermionic context, see [31–33]). The history of string phenomenology suggests that it is difficult to fine tune any particular construction in order to simultaneously meet all the properties of the Standard Model. Instead, the approach we take is that of a 'blind' automated scan over a huge number of models; for the present scan this number is of order 10^{40}.

Following this approach, what lies in front of the heterotic string model builder is a set of highly non-trivial challenges that can be summarised in the following checklist:

1. Construct a geometrical set-up, such that the 4-dimensional compactification of the $\mathcal{N} = 1$ supergravity limit of the heterotic string contains the symmetry SU(3) $\times$ SU(2) $\times$ U(1) of the Standard Model of particle physics. This step is usually realised in two stages, by firstly breaking the E_8 heterotic symmetry to a Grand Unified Theory (GUT) group and then breaking the latter to the Standard Model gauge group (plus possibly U(1) factors). This requires a VEV of the gauge connection on the internal (compact) 6-dimensional space X, or, equivalently, one needs to construct a vector bundle $V \to X$.

2. Derive the matter spectrum of the 4-dimensional theory. At low energy, the fermion fields transforming under the broken gauge group must be massless modes of the Dirac operator on the internal space X. The number of massless modes for a given representation is given by the dimension of certain bundle-valued cohomology groups on X. Such cohomology computations are generically difficult to perform. At this stage, one would like to identify models with the matter spectrum of the minimally supersymmetric Standard Model (MSSM), typically a very small sub-set of all consistent models constructed in this way.

3. Constrain the resulting Lagrangian, in order to avoid well-known problems of supersymmetric GUT models, such as fast proton decay. For this purpose, additional discrete or continuous symmetries derived from the compactification set-up may be helpful.

4. Derive information about the detailed properties of the model, such as the superpotential, the holomorphic Yukawa couplings, fermion mass-terms and μ-terms. Such holomorphic quantities can usually be understood using techniques from algebraic geometry.

5. Compute the physical Yukawa couplings. The physical Yukawa couplings consist of holomorphic superpotential terms times a non-holomorphic prefactor, whose computation requires the explicit knowledge of the metric on X and the gauge connection on the vector bundle V. For the case when X is a Calabi-Yau manifold, Yau's proof [42] guarantees the existence of a Ricci-flat metric, while for poly-stable vector bundles on Calabi-Yau manifolds, the Donaldson-Uhlenbeck-Yau theorem [43, 44] guarantees

the existence of a Hermitian Yang-Mills connection. However, except in very special cases, these quantities are not known analytically. So far, one can approach this differential geometric problem only numerically [45–47].

6. Stabilize the moduli and break supersymmetry. Recently, some progress has been made by including the effect of the $E_8 \times E_8$ bundle flux [48–50].

7. Compute soft-breaking parameters.

Every phenomenological requirement in this list will lead to a substantial reduction in the number of viable models. It is, therefore, crucial to start with a large number of models at the initial stages, if one hopes to retain a realistic model in the end. In this paper, we will concentrate on precisely this task and obtain, within a certain class of constructions, the largest possible set of models after the first two steps.

The history of this field can be largely understood by looking at the types of poly-stable holomorphic vector bundles that have been the focus of study at any given time. In the early days of the subject researchers largely concentrated on small deviations from the "standard embedding", where the gauge bundle was taken to be a holomorphic deformation of the tangent bundle [5, 6]. Such work has been continued to the current day with the first exact MSSM being produced from such an approach relatively recently [8]. In the 1990's and later more general poly-stable holomorphic vector bundles, or "non-standard embeddings", began to be considered in earnest [9–20, 51]. These gauge bundles were typically taken to have structure groups SU(3), SU(4) or SU(5) leading to an E_6, SO(10) or SU(5) GUT group, respectively.

Recently, a new approach to building heterotic models on smooth Calabi-Yau three-folds has been advanced [40, 41]. In this approach, the vector bundles in consideration were chosen to be simple sums of line bundles. This construction leads to a GUT group which naively includes additional U(1) factors in the GUT group beyond the gauge groups mentioned above. However, these extra U(1)'s are frequently broken, in addition to other effects, by the Green-Schwarz mechanism [52, 53]. As such, these models are just as capable of leading to acceptable particle physics phenomenology as their non-abelian cousins.

There are several advantages to working with sums of line bundles, as opposed to irreducible vector bundles. Firstly, such configurations are relatively simple to deal with from a computational point of view, and as a result, vastly greater numbers of models can be considered as compared to other approaches, such as [18–20]. Secondly, although broken at a high scale, the additional U(1) symmetries that are present in these models can greatly constrain the Lagrangian of these models giving more information about the superpotential, and in particular the Kähler potential, than is usually available in smooth heterotic constructions. Finally, although line bundle sums often represent special loci in the moduli space of vector bundles of a given topology, one can move away from the 'split' locus by turning on VEVs for certain bundle moduli, thus reaching non-abelian bundles. As such, these simple configurations provide a computationally accessible window into an even bigger moduli space of heterotic compactifications.

The two previous papers in which the line bundle construction was employed, [40, 41], achieved a number of goals. These publications presented the results of a scan over some 10^{12} line bundle sums in a search for heterotic standard models. The number of models investigated was in some sense arbitrary. The range of integer values defining the first Chern class of the line bundles scanned over was relatively restricted. These values were chosen simply to give a large number of models which was, nevertheless, manageable with a relatively simply implemented algorithm. From these 10^{12} vector bundles, the authors extracted 202 SU(5) GUT models which had precisely three generations of GUT families, no anti-families, at least one $5 - \bar{5}$ pair of Higgs fields and no other charged matter of any kind whatsoever. Each of these models was constructed such that there was at least one Wilson line which could be added to the configuration which led to exactly the spectrum of the MSSM with the Higgs triplets being projected out. In fact, each of these GUT models led to many different standard models due to choices, closely related to the Wilson line breaking, which were available in the construction. The results of including Wilson lines were presented explicitly and the 202 GUT models led to, by a conservative counting, 2122 heterotic standard models. The constraints on the effective field theory description of these models coming from the broken U(1) gauge factors was also explicitly computed. The complete data set including higher dimensional construction and tabulation of the resulting four-dimensional effective field theories was presented in a data base which can be found here [54].

In this paper, we extend the scope of the scan to make it comprehensive within the class of heterotic compactifications being studied. Instead of restricting the range of integers defining the line bundles in an arbitrary manner, we developed an algorithm which allows for an exhaustive scan. This leads to a considerable increase in the size of the data set being considered. Instead of examining 10^{12} configurations as before, in this paper we present the results of a scan over 10^{40} different compactifications of heterotic string theory. Even with improvements to our methodology from the previously published work, this scan, which is described in section 5 ran on a computer cluster over a period of seven months. The data set of GUT models we have obtained is, unsurprisingly, much larger than the 202 models discussed above. A conservative counting results in $34,989$ such GUT models which we expect will lead to one order of magnitude more heterotic standard models when Wilson lines are added. Given the size and extra technical complications resulting from dealing with such huge numbers of heterotic compactifications, we will present the detailed analysis of incorporating the effects of the Wilson line breaking in a separate paper. This represents in and of itself a huge task in computational algebraic geometry, which will take several months to complete.

The table below presents a statistics on the total number of consistent GUT models which have resulted from the search detailed in section 5. The first column counts SU(5) GUT models having the correct chiral asymmetry, which can, however, suffer from the presence of $\overline{\mathbf{10}}$ multiplets or the absence of $\mathbf{5} - \overline{\mathbf{5}}$ pairs to accomodate the Higgs content of the standard model. In the second column we eliminate those models that contain $\overline{\mathbf{10}}$ anti-family matter. This step relies on computations of line bundle cohomology groups, which we are able to perform in 94% of all cases. The number in parentheses indicates the

GUT models	no $\overline{\mathbf{10}}$ multiplets	no $\overline{\mathbf{10}}$s and at least one $\mathbf{5}-\overline{\mathbf{5}}$ pair
63325	44343 (3606)	34989 (5291)

Table 1. Statistics on the number of models.

GUT models for which we could not decide upon the presence of $\overline{\mathbf{10}}$ multiplets. Similarly, in the third column we select from the 44343 models that definitely have no anti-families, those which contain at least one $\mathbf{5} - \overline{\mathbf{5}}$ pair to contain MSSM Higgs fields.

The rest of this paper is structured as follows. In the next section we review the line bundle construction, and in addition provide a discussion of the structure group of holomorphic sums of line bundles, which is crucial in determining correctly the GUT gauge group. The latter discussion is new and, for the purpose of preserving the fluidity of the text, we defer a full presentation to appendix A. In the following two sections we define the class of manifolds under consideration and present the constraints imposed on the vector bundles. In section 5 we outline the algorithm used in the automated scan, while in section 6 we list the number of viable models obtained over each manifold, noting that in all cases we reach a limit beyond which no realistic line bundle vacua exist. We conclude with an example and final remarks.

2 Overview of the construction

The structure of $E_8 \times E_8$ heterotic compactifications on smooth Calabi-Yau three-folds with Abelian vector bundles, as well as the class of $\mathcal{N} = 1$ four dimensional supergravities to which they lead, have been thoroughly discussed in two previous publications [40, 41]. As such, we limit the scope of this section to merely summarising the central features of heterotic line bundle standard models. Additionally, we provide a discussion on the possible structure groups of vector bundles constructed as direct sums of holomorphic line bundles.

2.1 Heterotic line bundle compactifications

Schematically, the construction and analysis of heterotic string line bundle standard models can be broken up into three steps.

1. In the first step, a solution to the 10-dimensional supergravity limit of the $E_8 \times E_8$ heterotic string is obtained by specifying several geometrical elements. Firstly, we compactify 10-dimensional space-time on a smooth Calabi-Yau threefold X. Over this manifold we specify a poly-stable holomorphic vector bundle V with structure group $H \subset E_8 \times E_8$ which describes the gauge field expectation values in the supergravity solution. The possible choices of V over a given X are restricted by several consistency requirements as described in sections 2.2 and 4. In the line bundle construction, the vector bundle is taken to be a direct sum of five holomorphic line bundles

$$V = \bigoplus_a L_a \, .$$

As we will discuss in section 2.2, we choose the five line bundles L_a such that the structure group $H \subset E_8$ is Abelian and of the form $H = S\left(\mathrm{U}(1)^5\right) \cong \mathrm{U}(1)^4$.

If the background derived in this first step was used to dimensionally reduce the heterotic string theory to obtain an $\mathcal{N} = 1$ four dimensional supergravity without further modification, then the result would be a supersymmetric GUT. The gauge group seen in four dimensions G would, naively, be the commutant of H inside E_8. For the line bundle models mentioned above, this leads to a GUT group

$$G = \mathrm{SU}(5) \times S\left(\mathrm{U}(1)^5\right)$$

However, the additional U(1) factors are generically Green-Schwarz anomalous and thus the associated gauge bosons often obtain Stückelberg masses which are close to the compactification scale in magnitude.

2. In the second step, Wilson lines are added on the Calabi-Yau in such a way as to break the GUT group described above, down to that of the Standard Model. Adding such structure to the compactification is only possible if X is not simply connected. Most standard constructions of Calabi-Yau threefolds lead to manifolds for which $\pi_1(X) = 0$. Fortunately this situation can be resolved by quotienting a monifold X obtained from one of the usual constructions by a freely acting discrete symmetry Γ. The fundamental group of the resulting smooth quotient manifold $\hat{X} = X/\Gamma$ is non-trivial, and in fact is isomorphic to Γ.

The vector bundle V constructed in step 1 must be consistent with this quotienting procedure. We must ensure that our bundle $V \to X$ descends to a well defined vector bundle $\hat{V} \to \hat{X}$. This is only the case if V admits an equivariant structure under the symmetry Γ. Indeed, the set of vector bundles on $\hat{X}$ is in one-to one correspondence with the set of equivariant vector bundles on X.

The heterotic theory is then compactified to four dimensions on this new quotiented configuration including a non-trivial Wilson line. The gauge group obtained in four dimensions is then the commutant of the structure group of the flat bundle associated to the Wilson line inside G. This result is corrected as described in the first step by the Green-Schwarz mechanism. If the configuration is chosen correctly this can lead to the standard model gauge group G_{SM} in four dimensions. The matter content must be computed by the usual techniques of dimensional reduction — including the effects of the Wilson line. One wishes to obtain examples where the resulting four dimensional standard model charged matter is exactly that of the MSSM.

3. As a final step in analysing a heterotic line bundle standard model, one can use the additional U(1) four dimensional gauge symmetries to constrain the operators present in the four dimensional Lagrangian even though they are spontaneously broken by the Green-Schwarz mechanism. This allows a degree of analytical control over the low energy theory associated to these models which is unusual in the context of smooth Calabi-Yau reductions — in particular with regard to the Kähler potential

for matter fields. In specific models, these symmetries can help forbid operators in the four dimensional theory whose presence can be problematic for issues such as proton stability.

In this paper we present the results obtained after pursuing the avenue described in the first step above. We construct a large class of GUT models, postponing the remaining analysis for a future publication. However, we stress that the full analysis is feasible and has already been carried out for the more restricted set of models described in ref. [40, 41]. In the rest of this section we describe the GUT gauge group and particle spectrum that is obtained in such constructions in more detail.

2.2 The GUT gauge group

As discussed in refs. [40, 41], the gauge group of the GUT models that we hope to construct using line bundles is $\mathrm{SU}(5) \times S\left(\mathrm{U}(1)^5\right)$, the maximal subgroup of E_8 which commutes with $S\left(\mathrm{U}(1)^5\right)$. We reserve this section for determining when the structure group of a direct sum of five holomorphic line bundles with vanishing first Chern class is indeed $S\left(\mathrm{U}(1)^5\right)$, leading to the desired GUT gauge group. We will outline below the possible structure groups (and obstructions) for $V = \bigoplus_{a=1}^5 L_a$.

It is a long standing problem in vector bundle geometry that in general, the structure group, H, of a vector bundle cannot be determined without explicit knowledge of the H-valued connection (satisfying the equations of motion, here the Hermitian Yang-Mills equations [40]). However, for holomorphic bundles in certain cases, knowledge of the topology of the bundle and other facts may be enough to fully specify H.

As discussed more fully in appendix A, for the present scans of rank 5, reducible bundles, built as a sum of holomorphic line bundles $\bigoplus_{a=1}^5 L_a$, there are only a few possibilities for H. We demand that $\sum_a c_1(L_1) = 0$. Thus, it can be argued that H must be a subgroup of $\mathrm{SU}(5)$, $\mathrm{SO}(5)$ or $\mathrm{Sp}(4)$ (see appendix A). The latter two structure groups are possible for a rank 5 vector bundle only if V_5 admits either a real or symplectic fiber structure (see [55, 56] and (A.7)), in the form of a vector bundle isomorphism, $\phi : V \to V^*$. Since $\bigoplus_a L_a$ is an odd sum of 5 line bundles, such an isomorphism is possible if and only if $L_a = \mathcal{O}_X$ for at least one a. To avoid this case we impose for all $a = 1, \ldots, 5$ that

$$c_1^r(L_a) \neq 0 \quad \text{for at least one value of} \quad r = 1, \ldots h^{1,1}(X) . \tag{2.1}$$

This constraint, combined with the vanishing of the first Chern class, means that H must be a sub-group of $\mathrm{SU}(5)$. It only remains to determine whether $H = S\left(\mathrm{U}(1)^5\right)$ or a proper sub-group thereof. If H is not equal $S\left(\mathrm{U}(1)^5\right)$, it is possible that its commutant in E_8 (the $4d$ GUT group, G) is not of the form $\mathrm{SU}(5) \times \mathrm{U}(1)^4$, but rather another group less suitable for realistic model-building. For example, the structure group $H = S\left(\mathrm{U}(1)^2\right) \times S\left(\mathrm{U}(1)^3\right)$ has commutant, $G = \mathrm{SU}(6) \times \mathrm{U}(1)^3$. To eliminate such phenomenologically unviable possibilities, the following condition is imposed on the Chern classes of the line bundle sum:

$$\sum_{a \in S} c_1(L_a) \neq 0 \text{ for all proper subsets } S \subset \{1, \ldots, 5\} . \tag{2.2}$$

repr.	cohomology	total number	required for MSSM
$\mathbf{1}_{\mathbf{e}_a - \mathbf{e}_b}$	$H^1(L_a \otimes L_b^*)$	$\sum_{a,b} h^1(L_a \otimes L_b^*) = h^1(V \otimes V^*)$	-
$\mathbf{5}_{-\mathbf{e}_a - \mathbf{e}_b}$	$H^1(L_a^* \otimes L_b^*)$	$\sum_{a<b} h^1(L_a^* \otimes L_b^*) = h^1(\wedge^2 V^*) = h^1(\wedge^2 V)$	n_h
$\overline{\mathbf{5}}_{\mathbf{e}_a + \mathbf{e}_b}$	$H^1(L_a \otimes L_b)$	$\sum_{a<b} h^1(L_a \otimes L_b) = h^1(\wedge^2 V)$	$3|\Gamma| + n_h$
$\mathbf{10}_{\mathbf{e}_a}$	$H^1(L_a)$	$\sum_a h^1(L_a) = h^1(V)$	$3|\Gamma|$
$\overline{\mathbf{10}}_{-\mathbf{e}_a}$	$H^1(L_a^*)$	$\sum_a h^1(L_a^*) = h^1(V^*)$	0

Table 2. The possible SU(5) matter representations which may be obtained in four dimensions and their U(1) charges. The dimensions of the cohomology groups indicated in the second column determine the multiplicity of each representation in the four dimensional spectrum. The third column gives the total number of each SU(5) representation present in the four dimensional effective theory, of any U(1) charge. The final column gives the number of each SU(5) multiplet that we require in the GUT theory in order to obtain the standard model spectrum (with at most n_h pairs of Higgs doublets) after the addition of suitable Wilson lines.

With these conditions in hand, it is guaranteed that the heterotic line bundle construction leads to a $4d$ GUT symmetry of the form $\mathrm{SU}(5) \times \mathrm{U}(1)^4$ (with the abelian factors generically Green-Schwarz massive). We are now ready to turn to the more detailed question of the charged matter particle spectrum of the low-energy theory.

2.3 The GUT spectrum

If the conditions of the previous section are satisfied, a sum of five line bundles breaks the E_8 heterotic symmetry to $\mathrm{SU}(5) \times S\left(\mathrm{U}(1)^5\right)$. In such a case, the computation of the spectrum of the heterotic line bundle model has been explained in detail in [40, 41]. Here we simply state the results for the convenience of the reader.

We represent $S\left(\mathrm{U}(1)^5\right)$ representations by vectors $\mathbf{q} = (q_1, \ldots, q_5)$ of five integer charges. Due to the determinant condition associated to the "S" of $S\left(\mathrm{U}(1)^5\right)$, two such vectors $\mathbf{q}$ and $\mathbf{q}'$, represent the same representation and, hence, have to be identified if $\mathbf{q} - \mathbf{q}' \in \mathbb{Z}\mathbf{n}$, where $\mathbf{n} = (1,1,1,1,1)$. We also introduce the standard basis $\{\mathbf{e}_a\}_{a=1,\ldots,5}$ in five dimensions. Charges for GUT multiplets are indicated by adding the charge vector as a subscript so that, for example, $\mathbf{10}_{\mathbf{e}_1}$ represents a $\mathbf{10}$ multiplet of SU(5) with charge 1 under the first U(1) and uncharged under the others. A list of the relevant multiplets and their properties is provided in table 2, below. Different SU(5) representations are associated with different patterns of U(1) charges. For example, the $\mathbf{10}$ multiplets carry charge one under precisely one of the five U(1) symmetries, while the $\overline{\mathbf{5}}$ multiplets carry charge one with respect to two U(1) symmetries. Apart from such rules, the precise assignment of charges across the spectrum (including bundle moduli) is model-dependent. This is of major phenomenological importance: invariance under the (global remnant of the) $S\left(\mathrm{U}(1)^5\right)$ symmetry constrains the allowed operators in the low-energy theory. Indeed, one can easily envisage situations in which the pattern of charges is such that, for example proton decay operators are forbidden.

The final column of the table shows the number of each SU(5) representation that we require in the GUT model such that after quotienting the Calabi-Yau manifold and adding

suitable Wilson lines, we obtain the spectrum of the MSSM. As before, Γ denotes a freely acting finite group by which we quotient the Calabi-Yau manifold and $|\Gamma|$ is its order.

We have now completed our overview of what is required for a successful heterotic line bundle standard model construction. In the next sections we move on to describe the particular class of Calabi-Yau threefolds which we will study, as well as some details of the line bundles over them.

3   The manifolds

Historically the first class of Calabi-Yau three-folds explicitly constructed [57], complete intersections in products of projective spaces (CICYs) have often served as the starting point in heterotic model building [5–8, 12–15, 18–20]. The present systematic computer-based scan for standard models will continue this tradition.

The choice is based on two crucial features of the CICY class of manifolds. In the first place, there exists a systematic classification of all linearly realised freely acting discrete symmetries on the CICY manifolds in the database [58, 59]. More accurately, Braun's classification [59] provides a list of all such symmetries which descend from a linearly acting symmetry on the ambient space. Furthermore, a given Calabi-Yau manifold can frequently be embedded in many different products of projective spaces. The symmetry classification of [59] is carried out for a limited selection of possible ambient spaces for each Calabi-Yau. As discussed in the previous section, a knowledge of such symmetries is an essential ingredient in breaking the GUT group to the Standard Model gauge group. We note that in constructing the GUT models presented here, it is in fact only knowledge of the possible orders of the available groups, and thus the possible values of $|\Gamma|$, which is required.

The second feature of the CICY's which makes them particularly suitable for the current work is related to their relative simplicity. The embedding of Calabi-Yau manifolds in such simple ambient spaces means that computations of the cohomology of line bundles over these manifolds can be effectively automatised on a computer.[1] This is particularly true when there is a strong connection between line bundles on the ambient space and line bundles on the Calabi-Yau threefold. As such we restrict our attention to CICYs presented in a 'favourable' embedding.

Favourable embeddings can be described in many equivalent ways. For example, on manifolds such as those we are considering, isomorphism classes of line bundles are completely classified by their first Chern class. This is an element of the second cohomology group $H^2(X, \mathbb{Z})$ of the base space X. Favourable CICYs can be defined to be those whose second cohomology descends entirely from the second cohomology of the embedding space. In such a case, all of the line bundles on X are restrictions of line bundles over the ambient product of projective spaces. For this property to hold for a given description of a CICY, certain requirements have to be satisfied, as discussed in appendix B.

---

[1]Using the "CICY Package" by L.B. Anderson, J. Gray, Y.-H. He, S.-J. Lee and A. Lukas based on methods described in in [60–64].

4 The bundles

Recall that our bundles $V \to X$ are taken to be sums of five line bundles

$$V = \bigoplus_{a=1}^{5} L_a \,. \tag{4.1}$$

A single line bundle is specified by its first Chern class and, hence, by a set of $h^{1,1}(X)$ integers. Given this, a sum of 5 line bundles is specified by a matrix of integers with $h^{1,1}(X)$ rows and 5 columns. In our systematic investigation, we have scanned over $\sim 10^{40}$ such matrices and selected approximately $35,000$ bundles which lead to phenomenologically consistent SU(5) GUTs. In the following subsections we list the criteria that these $35,000$ models satisfy.

4.1 Topological constraints

We require that,

$$c_1(V) = 0 \,, \tag{4.2}$$

as well we eq. (2.2), such that the structure group of V is $S(\mathrm{U}(1)^5)$ which leads to a GUT group $G = \mathrm{SU}(5) \times S\left(\mathrm{U}(1)^5\right)$. Apart from this group theoretical advantage, imposing $c_1(V) = 0$ guarantees the existence of a spin structure on V.

In addition, the integrability condition on the Bianchi Identity for the Neveu-Schwarz two form leads to the following constraint on the vector bundle V.

$$\mathrm{ch}_2(TX) - \mathrm{ch}_2(V) - \mathrm{ch}_2(\tilde{V}) + [C] = 0 \tag{4.3}$$

Here $[C] \in H^4(X)$ is the Poincaré dual to the effective holomorphic curve class wrapped by a five-brane and $\tilde{V}$ is the hidden-sector bundle (which we will take to be trivial). The simplest way to guarantee that this condition can be satisfied is to require that $c_2(TX) - c_2(V) \in$ Mori cone of X, where we have used that $c_1(V) = 0$. In this case, an effective curve class which saturates the condition (4.3) for a trivial hidden bundle $\tilde{V}$ exists (although, typically, solutions with non-trivial $\tilde{V}$ can be found as well). For favourable CICYs we have a basis $\{J_r\}$ of (1,1)-forms on X, obtained from hyperplane classes of the embedding projective spaces, such that the Kähler forms $J = t^r J_r$ correspond to positive values, $t^r > 0$, of the Kähler parameters t^r. The above Mori cone condition can then be written as

$$\int_X (c_2(TX) - c_2(V)) \wedge J_r \geq 0, \text{ for all } r \in \{1, \dots, h^{1,1}(X)\} \,. \tag{4.4}$$

4.2 Constraints from stability

Demanding an $\mathcal{N} = 1$ supersymmetric vacuum in four dimensions leads to the requirement that the gauge connection on V satisfies the hermitian Yang-Mills equations at zero slope. By the Donaldon-Uhlenbeck-Yau theorem this is possible if and only if V is holomorphic, has vanishing slope and is polystable.

The slope of a vector bundle V is defined as

$$\mu(V) = \frac{1}{\mathrm{rk}(V)} \int_X c_1(V) \wedge J \wedge J = \frac{1}{\mathrm{rk}(V)} \sum_{r,s,t=1}^{h^{1,1}(X)} d_{\mathrm{rst}}\, c_1^r(V) t^s t^t \,,$$

where $d_{\mathrm{rst}} = \int_X J_r \wedge J_s \wedge J_t$ are the triple intersections on X.

For the case of interest, V is a direct sum of line bundles and $c_1(V) = 0$. The vanishing slope condition $\mu(V) = 0$ is therefore automatically satisfied. In addition, these sums of line bundles are automatically holomorphic. On the other hand, poly-stability reduces to the requirement that,

$$\exists\, t^r \quad \text{such that} \quad \mu(L_a)|_{t^r} = 0 \ \ \forall a \tag{4.5}$$

somewhere in the interior of the Kähler cone ($t^r > 0 \ \forall r$).

Finally, we note that for slope(poly)-stable bundles on a Calabi-Yau threefold there is a positivity condition on the second Chern class, given by the so-called Bogomolov bound [65]. For $SU(n)$ bundles this takes the simple form

$$\int_X c_2(V) \wedge J \geq 0 \tag{4.6}$$

and J is any Kähler form for which V is poly-stable.

4.3 Constraints from the GUT spectrum

The $SU(5) \times S\left(U(1)^5\right)$ GUT spectrum has already been discussed in section 2.3. In order to secure a chiral asymmetry of 3 after taking the quotient of X by Γ, we must require that $h^1(X, V) - h^2(X, V) = h^1(X, \wedge^2 V) - h^2(X, \wedge^2 V) = 3|\Gamma|$. Since for a poly-stable bundle the zeroth and the top cohomologies vanish, the chiral asymmetry conditions can be formulated in terms of the indices

$$\mathrm{ind}(V) = \mathrm{ind}(\wedge^2 V) = -3|\Gamma| \,. \tag{4.7}$$

In fact, for an $SU(5)$ bundle, $\mathrm{ind}(V) = \mathrm{ind}(\wedge^2 V)$, so one needs to check only one chiral asymmetry. Furthermore, in order to exclude anti-families, we require the absence of $\overline{\mathbf{10}}$ multiplets, and hence that

$$h^2(X, V) = 0 \tag{4.8}$$

Finally, in order to safeguard the presence of at least one Higgs doublet, it is necessary to demand the existence of at least one $\mathbf{5} - \overline{\mathbf{5}}$ pair which is expressed by the requirement

$$h^2(X, \wedge^2 V) > 0 \,. \tag{4.9}$$

4.4 Equivariance and the doublet-triplet splitting problem

In addition to the above constraints, we demand that, for each $a < b$

$$\mathrm{ind}(L_a \otimes L_b) \leq 0 \tag{4.10}$$

In the case where each of the line bundles composing V are individually equivariant, this condition is necessary for it to be possible to project out all Higgs triplets upon the addition of a Wilson line. If the line bundles composing V are individually equivariant, then V descends to a simple sum of line bundles on the quotient space $\hat{X}$. In such a case there is a relation between the index of the line bundle products on X and $\hat{X}$.

$$\text{ind}(\widetilde{L}_a \otimes \widetilde{L}_b) = \frac{1}{|\Gamma|} \text{ind}(L_a \otimes L_b)$$

All indices involved must, of course, be integers and in the case where the line bundles are individually equivariant the size of $\text{ind}(L_a \otimes L_b)$ will be such as to ensure that this is true. Given this relationship, if $\text{ind}(L_a \otimes L_b) > 0$ then so is $\text{ind}(\widetilde{L}_a \otimes \widetilde{L}_b)$. This ensures that in such cases there is at least one complete set of **5** degrees of freedom of SU(5) in the four dimensional effective theory — leading to the presence of Higgs triplets. This result is unaffected by the presence of Wilson lines as the undesirable particle content is protected by an index which such gauge configurations do not affect.

For line bundle sums with non-trivial equivariant blocks the situation is more complicated since divisibility of the index only applies to each equivariant block rather than to individual line bundles. In the simplest such case, an equivariant block is formed by two or more same line bundles which are permuted by the equivariant structure. More complicated equivariant blocks can consist of different line bundles which are mapped into each other, typically subject to an additional permutation of their integer entries. At any rate, bundle isomorphisms between the relevant line bundles must exist in this case so they must have the same index. In conclusion, line bundles within equivariant blocks must have the same index and, hence, if this index is positive so is the index of the equivariant block. Then, a generalization of the above index argument to the entire block leads to the same conclusion, namely the inevitable presence of Higgs triplets. Hence, the condition (4.10) should be imposed in all cases.

5   The scanning algorithm

If a sum of five line bundles passes the criteria (4.2)–(4.10), it leads to a consistent four dimensional GUT theory which, with appropriate Wilson line breaking, will lead to heterotic standard models. As mentioned above, a sum of five line bundles is specified by $5 \cdot h^{1,1}(X)$ integers. An examination of the CICY database reveals that all manifolds which are favourable and admit known linear free actions of discrete groups have $h^{1,1}(X)$ in the range $2 \leq h^{1,1}(X) \leq 6$.[2] Thus, we are interested in investigating bundles described by matrices of between 10 and 30 integers, and deciding when they obey the criteria we have described.

One could envisage a scan over all line bundle sums with entries between, say, -10 and 10, for the manifolds with $h^{1,1}(X) = 6$. This would require us to check $\sim 10^{30}$ matrices representing sums of line bundles. For comparison, a year has $\sim 3 \cdot 10^7$ seconds. It is rather

---

[2]Manifolds with $h^{1,1}(X) = 1$ cannot lead to consistent line bundle models since the slope zero condition (4.5) cannot be satisfied.

clear that such an attempt would be impossible if one desired to explicitly construct each such line bundle matrix and then check the necessary criteria. A better approach is based on the observation that the criteria (4.2)–(4.10) impose certain conditions on individual line bundles, as well as on partial sums of line bundles. These restrictions are of four kinds: stability related, index related, conditions stemming from the integrability of the heterotic Bianchi identity and restrictions on cohomology.

The constraint from stability imposes that each collection of up to five line bundles $\{L_a\}$ can only describe a heterotic vacuum if there exists a point in the interior of the Kähler cone, $t^r > 0$, such that simultaneously for all a,

$$\sum_{r,s,t=1}^{h^{1,1}(X)} d_{\text{rst}}\, c_1^r(L_a)\, t^s t^t = 0.$$

This is condition (4.5) of the previous section.

Deciding whether a quadratic equation in several variables has positive solutions can be a computationally intensive question. Establishing the existence of common positive solutions for a collection of such equations is an even more formidable problem. On the other hand, given that the Kähler cone for our manifolds is given by $t^r > 0$ for all $r = 1, \ldots, h^{1,1}(X)$, a fairly strong, and obviously necessary, condition for the existence of positive solutions to the slope-zero equation for each line bundle L_a is that the matrix $(M_a)_{st} = d_{\text{rst}}\, c_1^r(L_a)$ has both positive and negative entries. For any subset $\{L_{a_1}, \ldots, L_{a_n}\} \subset \{L_1, \ldots, L_5\}$ of our five line bundles to have common positive solutions, it is necessary that any linear combination of the matrices $\{M_{a_1}, \ldots, M_{a_n}\}$ has both positive and negative entries. In practice, we consider linear combinations with integer coefficients between -5 and 5. This turns out to be a remarkably effective way of eliminating line bundle sums that are not poly-stable.

For the line bundle sums that pass this necessary criterion we explicitly find common solutions to the slope-zero equations in the interior of the Kähler cone. In a great majority of the cases we are able to find exact solutions, while in the remaining cases (most of which appear for the $h^{1,1} = 6$ manifolds), we have to resort to numerical methods.

The index-related criteria impose, for any subset of the five line bundles, that the sum of their indices is negative and greater than or equal to $-3|\Gamma|$. This criterion simply follows from equation (4.7) — we do not want more than three generations of standard model particles. We must also check the index based criteria given by equation (4.10). Indices of line bundles can be computed very rapidly in terms of their defining integers using the following standard formula:

$$\text{ind}(L) = \frac{1}{12} \left(2\, c_1(L)^3 + c_1(L)\, c_2(TX) \right)$$
$$= \sum_{r,s,t=1}^{h^{1,1}(X)} d_{\text{rst}} \left(\frac{1}{6}\, c_1^r(L)\, c_1^s(L)\, c_1^t(L) + \frac{1}{12} c_1^r(L)\, c_2^{s\,t}(TX) \right)$$

The condition (4.4) stemming from the integrability of the heterotic Bianchi Identity constrains the full sum of five line bundles. This can be rewritten in terms of the integers

describing the line bundles as follows.

$$\int_X c_2(TX) \wedge J_r \geq \int_X c_2(V) \wedge J_r = \frac{1}{2} d_{\mathrm{rst}} \sum_{a=1}^{5} c_1^s(L_a) \, c_1^t(L_a)$$

The restrictions on the cohomology of the line bundle sums (4.8)–(4.9) require a larger amount of computational resources to implement than do the simple checks already described. Therefore, at first, we only take into account the remaining constraints (4.2)–(4.7) and (4.10). This stage of the scan leads to GUT models with the correct chiral asymmetry, but does not exclude the possibility of having $\overline{\mathbf{10}}$ multiplets or no Higgs doublets at all. In the second stage of the scan, we attempt to eliminate the models containing anti-families. In 94% of the cases we are able to compute the required cohomology and thus decide upon the fate of the corresponding models. Finally, we eliminate the models that have no Higgs doublets, with a rate of decidability of 88%. The code we use to evaluate the relevant cohomolgy groups can only compute the ranks of a subset of the maps involved in the calculation. The 6% and 12% of cohomologies respectively which could not be determined here require these ranks to be computed. The computation of the cohomology of line bundles over CICYs is reviewed, for example, in ref. [50].

Below, we schematically present the algorithm used in this automated search. The input parameters are the Calabi-Yau data (configuration matrix, intersection numbers, $c_2(TX)$, a list of row permutations that leave the configuration matrix unchanged); the order of a freely acting discrete group Γ and the maximal value for a line bundle entry, $k_{\max}$. The list of permutations present in the Calabi-Yau data is used in order to eliminate redundant line bundle sums, that is, line bundle sums that can be related to one another by a trivial re-labeling of the ambient space projective factors. The algorithm outputs a list of *Models* represented as matrices of integers whose columns stand for the first Chern classes of the 5 line bundles.

1. assemble *List_1* containing line bundles satisfying:

    i) $-3|\Gamma| \leq \text{ind}\,(L) \leq 0$ and

    ii) $\mu(L) = 0$, somewhere in the interior of the Kähler cone

2. obtain *List_1r* $\subset$ *List_1* by removing all redundant line bundles;

3. for each $L_{i_1} \in$ *List_1r* assemble *List_2*$(L_{i_1}) \subset$ *List_1* containing line bundles such that, for every $L_{i_2} \in$ *List_2*(L_{i_1}) the following relations hold:

    i) $-3|\Gamma| \leq \text{ind}\,(L_{i_1}) + \text{ind}\,(L_{i_2})$;

    ii) $-3|\Gamma| \leq \text{ind}\,\left(\wedge^2\left(L_{i_1} \oplus L_{i_2}\right)\right) = \text{ind}\,(L_{i_1} \otimes L_{i_2}) \leq 0$;

    iii) $\mu(L_{i_1}) = \mu(L_{i_2}) = 0$, somewhere in the interior of the Kähler cone

4. given $L_{i_1} \in$ *List_1r*, for each $L_{i_2} \in$ *List_2*(L_{i_1}) assemble *List_3*$(L_{i_1}, L_{i_2}) \subset$ *List_2*(L_{i_1}), such that any $L_{i_3} \in$ *List_3*(L_{i_1}, L_{i_2}) satisfies:

$i)$ $-2\,k_{\max} \leq c_1^r(L_{i_1}) + c_1^r(L_{i_2}) + c_1^r(L_{i_3}) \leq 2\,k_{\max}$, for all $r \in \{1,\ldots,h^{1,1}(X)\}$

$ii)$ $-3|\Gamma| \leq \mathrm{ind}\,(L_{i_1}) + \mathrm{ind}\,(L_{i_2}) + \mathrm{ind}\,(L_{i_3})$;

$iii)$ $-3|\Gamma| \leq \mathrm{ind}(L_{i_1} \otimes L_{i_3}) \leq 0$; $\quad -3|\Gamma| \leq \mathrm{ind}(L_{i_2} \otimes L_{i_3}) \leq 0$;

$iv)$ $-3|\Gamma| \leq \mathrm{ind}\left(\wedge^2\left(L_{i_1} \oplus L_{i_2} \oplus L_{i_3}\right)\right) = \mathrm{ind}\,(L_{i_1}\otimes L_{i_2}) + \mathrm{ind}\,(L_{i_1}\otimes L_{i_3}) + \mathrm{ind}(L_{i_2}\otimes L_{i_3}) \leq 0$;

$v)$ $\mu(L_{i_1}) = \mu(L_{i_2}) = \mu(L_{i_3}) = 0$, somewhere in the interior of the Kähler cone

5. given L_{i_1}, L_{i_2} and L_{i_3} as above, select from $List\_3(L_{i_1}, L_{i_2})$ those line bundles L_{i_4}, such that the line bundle L_{i_5} defined by $c_1(L_{i_1} \oplus L_{i_2} \oplus L_{i_3} \oplus L_{i_4} \oplus L_{i_5}) = 0$ satisfies:

   $i)$ $-k_{\max} \leq c_1^r(L_{i_5}) \leq k_{\max}$

   $ii)$ $-3|\Gamma| \leq \mathrm{ind}\,(L_{i_5}) \leq 0$;

   $iii)$ $\mu(L_{i_5}) = 0$ somewhere in the interior of the Kähler cone

6. given $L_{i_1}, L_{i_2}, L_{i_3}, L_{i_4}$ and L_{i_5} as above, check:

   $i)$ $-3|\Gamma| = \mathrm{ind}\,(L_{i_1}) + \mathrm{ind}\,(L_{i_2}) + \mathrm{ind}\,(L_{i_3}) + \mathrm{ind}\,(L_{i_4}) + \mathrm{ind}\,(L_{i_5})$;

   $ii)$ $-3|\Gamma| = \mathrm{ind}\left(\wedge^2\left(L_{i_1} \oplus L_{i_2} \oplus L_{i_3} \oplus L_{i_4} \oplus L_{i_5}\right)\right)$;

   $iii)$ $\mathrm{ind}(L_{i_a} \otimes L_{i_b}) \leq 0$ for all pairs $a < b$ that have not been checked so far;

   $iv)$ $\mu(L_{i_1}) = \mu(L_{i_2}) = \mu(L_{i_3}) = \mu(L_{i_4}) = \mu(L_{i_5}) = 0$, in the interior of the Kähler cone

   $v)$ $2\,d_{\mathrm{rst}}\,c_2^{s\,t}(TX) \geq d_{\mathrm{rst}} \sum_{a=1}^{5} c_1^s(L_{i_a})\,c_1^t(L_{i_a})$

   if these requirements are satisfied, append $L_{i_1} \oplus L_{i_2} \oplus L_{i_3} \oplus L_{i_4} \oplus L_{i_5}$ to $Models$

7. remove all redundant line bundle sums from $Models$.

8. for the remaining $Models$, check stability by explicitly finding points in the Kähler cone where all line bundles have slope zero.

9. eliminate models with $\overline{\mathbf{10}}$ multiplets

10. eliminate models with no $\mathbf{5} - \overline{\mathbf{5}}$ pairs

In the next section we discuss the results of running this algorithm and will describe in what sense the list of models obtained is comprehensive in nature.

6 Results and finiteness

It is expected that the number of slope(poly)-stable vector bundles relevant for a smooth heterotic compactification is in fact finite. To begin, it is possible to see that for any bundle which is stable somewhere in the Kähler cone, the possible values of its topology, at least, are finite. For example, for the bundles under consideration here, the first Chern class is constrained to vanish (the condition for spinors) and the second Chern class is bounded from above by the anomaly cancellation condition (4.3) and from below by the Bogomolov bound (4.6), thus yielding a finite range of possible values for $c_2(V)$. Moreover, for an $SU(n)$ vector bundle which is semistable somewhere in the Kähler cone, it is known that there can be only finitely many values possible for the third Chern class [66, 67]. Moreover for fixed topology (that is, fixed total Chern class) the moduli space of semi-stable sheaves on a Calabi-Yau threefold is known (by algebraicity of the family [65, 66]) to have only finitely many components.

An important subtlety arises here for the problem at hand — namely that the mathematical definition of this moduli space proceeds by first defining an explicit choice of Kähler form (that is, a ray in Kähler moduli space) with respect to which the sheaves are semistable. However, for the purposes of heterotic model building, we aim to build all Standard Model bundles which are stable for some (not necessarily all the same) Kähler form. We wish to know, then, if there are finitely many components to the collection of all moduli spaces, allowing the Kähler moduli to vary over *any* relevant values in the Kähler cone. For this harder problem, some boundedness results are still known (with the most detailed bounds possible in the case of complex surfaces) (see [65, 67] for a review) for the families of moduli spaces relevant for present scans.

Thus, viewing our poly-stable sums of line bundles as special points in the moduli space of semistable sheaves, it is expected that for a given Calabi-Yau threefold, X, there are a bounded number of line bundle Standard Models that can be constructed. However, for the most part, the bounds described above are non-constructive. Thus, in this work, we will empirically bound the number of such models by algorithmic scanning and explicitly constructing the poly-stable vector bundles.

The number of models over a certain manifold admitting discrete symmetries of a fixed order is an increasing and saturating function of the maximal integer appearing in the line bundles (ignoring sign). This can be observed for all the pairs $(X, |\Gamma|)$ that we have considered, as shown in the tables below. However, we believe that our results reflect a more general phenomenon. In practice, we have applied the above algorithm to all pairs $(X, |\Gamma|)$ of favourable CICYs with the orders $|\Gamma|$ of a freely-acting symmetries and all line bundle sums (4.1) with $|c_1^r(L_a)| \leq k_{\mathrm{m}}$, for a fixed upper bound k_{m}. In each case, the number of viable models has then been determined for increasing values of k_{m} until saturation occurred. As a practical criterion for the onset of saturation we have required the number of models to remain unchanged for three consecutive values of k_{m}. This criterion has been checked by going to large integer entries in the line bundles in simple cases where the computation can be completed in a modest amount of time. However, although it seems that the data set is indeed comprehensive, there is no formal proof that this criterion

$X, |\Gamma|$	$k_{\mathrm{m}} = 1$	$k_{\mathrm{m}} = 2$	$k_{\mathrm{m}} = 3$	$k_{\mathrm{m}} = 4$	$k_{\mathrm{m}} = 5$	$k_{\mathrm{m}} = 6$	$k_{\mathrm{m}} = 7$
7484, 4	0	0	0	1	1	1	
7669, 3	0	0	2	2	2		
7669, 9	0	0	1	1	1		
7735, 8	0	0	0	0	1	1	1
7745, 8	0	0	0	0	1	1	1

Table 3. Number of models as a function of k_{m} on CICYs with $h^{1,1}(X) = 3$. Total number of models: 6.

is enough to ensure that all relevant sums of line bundles have been found. The results, before imposing the absence of $\overline{\mathbf{10}}$ multiplets, eq. (4.8), and the presence of Higgs multiplets, eq. (4.9), can be found in the subsequent tables. The Calabi-Yau manifolds, X, are specified by a number, given in the first column of the tables below, which represents their position in the standard list of CICYs compiled in refs. [57, 68] and explicitly accessible here [69]. As is evident from the tables all viable models consist of line bundles satisfying

$$|c_1^r(L_a)| \leq 10 \ . \tag{6.1}$$

As one would expect, their number increases dramatically with $h^{1,1}(X)$, the number of Kähler parameters. For $h^{1,1}(X) = 1, 2, 3, 4, 5, 6$ we find $0, 0, 6, 552, 21731, 41036$ models, respectively, for a total of 63325 models, the number already quoted in the introduction. When the two further constraints (4.8) and (4.9) are imposed this number reduces to 44343 and 34989, as already indicated in table 1. The number of models at each stage, for all pairs $(X, |\Gamma|)$ is tabulated in appendix C. Numbers in parentheses in the tables of this appendix indicate how many cases a given quantity could not be computed for, given the limitations in the code used to compute cohomology which we have already discussed. The complete list of these models can be accessed here [70].

We note that the average number of viable models per pair $(X, |\Gamma|)$ as a function of $h^{1,1}(X)$ is approximately given by $0.3, 20, 530, 4560$ for $h^{1,1} = 3, 4, 5, 6$, respectively. Very roughly, this corresponds to an increase of one order of magnitude per additional Kähler parameter. At this point it is tempting to speculate about the total number of standard models, that is, models with the MSSM spectrum, in string theory. Known Calabi-Yau three-folds have Hodge numbers in the range of up to $h^{1,1}(X) \leq 500$. If the increase by an order of magnitude observed at small $h^{1,1}(X)$ continues to such large values the number of string standard models is enormous. However, a line bundle sum is determined by $5 \cdot h^{1,1}(X)$ integers and it seems likely that the three-family constraint becomes more difficult to satisfy for a large number of these integers. We would, therefore, expect the increase to slow down at larger $h^{1,1}(X)$. Currently, we do not see any way of checking this by explicit scanning since $h^{1,1}(X) = 6$ marks out the reach of present computational power.

7   An example

For illustration, we would like to present a model from our database, which is accessible here [70]. The example is based on the CICY with number 7447, defined by the configura-

X, $|\Gamma|$	$k_{\mathrm{m}} = 1$	$k_{\mathrm{m}} = 2$	$k_{\mathrm{m}} = 3$	$k_{\mathrm{m}} = 4$	$k_{\mathrm{m}} = 5$	$k_{\mathrm{m}} = 6$	$k_{\mathrm{m}} = 7$	$k_{\mathrm{m}} = 8$	$k_{\mathrm{m}} = 9$
6784, 2	0	0	2	10	12	12	12		
6784, 4	0	6	38	50	62	70	70	70	
6828, 2	0	0	1	5	6	6	6		
6828, 4	0	3	19	25	31	35	35	35	
6831, 2	0	1	2	2	2				
7204, 2	0	2	14	22	22	22			
7218, 2	0	1	7	11	11	11			
7241, 2	0	1	7	11	11	11			
7245, 2	0	1	4	4	4				
7247, 3	0	19	57	59	59	59			
7270, 2	0	2	14	22	22	22			
7403, 2	0	3	6	6	6				
7435, 2	0	0	0	2	2	2			
7435, 4	0	0	5	8	9	10	10	10	
7462, 2	0	0	0	6	6	6			
7462, 4	0	0	15	24	27	30	30	30	
7468, 2	0	5	7	7	7				
7491, 2	0	0	0	2	2	2			
7491, 4	0	0	5	8	9	10	10	10	
7522, 2	0	0	0	6	6	6			
7522, 4	0	0	15	24	27	30	30	30	
7719, 2	0	4	14	26	26	26			
7736, 2	0	2	7	13	13	13			
7742, 2	0	2	7	13	13	13			
7862, 2	0	5	7	10	10	10			
7862, 4	0	9	46	54	58	58	58		
7862, 8	0	3	40	53	58	62	64	64	64
7862, 16	0	0	0	1	4	5	5	5	

Table 4. Number of models as a function of k_{m} on CICYs with $h^{1,1}(X) = 4$. Total number of models: 552.

tion matrix and line bundle sum

$$
X = \begin{array}{c} \mathbb{P}^1 \\ \mathbb{P}^1 \\ \mathbb{P}^1 \\ \mathbb{P}^1 \\ \mathbb{P}^1 \end{array} \left[\begin{array}{cc} \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{1} \end{array} \right]^{5,45}_{-80} \, , \qquad V = \begin{bmatrix} -1 & -2 & 1 & 1 & 1 \\ 0 & -2 & -1 & 1 & 2 \\ 0 & 2 & -1 & 1 & -2 \\ 0 & 2 & 0 & 0 & -2 \\ 1 & 0 & 0 & -2 & 1 \end{bmatrix} \, .
$$

According to ref. [59], the manifold X can be smoothly quotiented by a group of order 4. The columns of the second matrix correspond to the first Chern classes of the five line bundles composing V. The dimension $h^\bullet(X, V) = \big(h^0(X, V), h^1(X, V), h^2(X, V), h^3(X, V) \big)$ of the bundle cohomologies for V are explicitly given by

$$
\begin{aligned}
h^\bullet(X, V) &= (0, 12, 0, 0) \\
h^\bullet(X, \wedge^2 V) &= (0, 15, 3, 0)
\end{aligned}
$$

$X, |\Gamma|$	$k_{\mathrm{m}} = 1$	$k_{\mathrm{m}} = 2$	$k_{\mathrm{m}} = 3$	$k_{\mathrm{m}} = 4$	$k_{\mathrm{m}} = 5$	$k_{\mathrm{m}} = 6$	$k_{\mathrm{m}} = 7$	$k_{\mathrm{m}} = 8$	$k_{\mathrm{m}} = 9$
5256, 2	0	575	727	775	779	779	779		
5256, 4	0	672	1857	2085	2173	2180	2180	2180	
5301, 2	0	144	182	194	195	195	195		
5301, 4	0	169	466	523	545	547	547	547	
5452, 2	0	574	726	774	778	778	778		
5452, 4	0	672	1854	2083	2171	2177	2177	2177	
6024, 3	0	303	510	513	513	513			
6204, 2	0	62	116	122	125	125	125		
6225, 2	0	147	221	231	232	232	232		
6715, 2	0	96	148	184	184	184			
6715, 4	0	165	690	812	844	848	848	848	
6724, 2	0	19	34	36	39	39	39		
6732, 2	0	434	778	880	880	880			
6770, 2	0	216	307	329	331	331	331		
6777, 2	0	434	778	880	880	880			
6788, 2	0	96	148	184	184	184			
6788, 4	0	165	690	812	844	848	848	848	
6802, 2	0	432	775	877	877	877			
6804, 2	0	59	154	169	173	173	173		
6834, 2	0	218	390	441	441	441			
6836, 2	0	24	37	46	46	46			
6836, 4	0	43	175	206	214	215	215	215	
6836, 8	0	6	94	120	131	133	137	137	137
6836, 16	0	0	0	0	2	3	3	3	
6890, 2	0	860	1546	1750	1750	1750			
6896, 2	0	218	390	441	441	441			
6927, 2	0	144	222	276	276	276			
6927, 4	0	244	1030	1212	1260	1266	1266	1266	
6927, 8	0	34	554	706	770	782	806	806	806
6947, 2	0	24	37	46	46	46			
6947, 4	0	43	175	206	214	215	215	215	
6947, 8	0	6	94	120	131	133	137	137	137
6947, 16	0	0	0	0	2	3	3	3	
7279, 2	0	128	204	212	218	218	218		
7447, 2	0	56	87	93	93	93			
7447, 4	0	214	377	419	428	430	432	432	432
7447, 10	0	6	58	72	81	82	83	83	83
7487, 2	0	277	430	459	459	459			
7487, 4	0	1052	1851	2058	2101	2111	2121	2121	2121

Table 5. Number of models as a function of k_{m} on CICYs with $h^{1,1}(X) = 5$. Total number of models: 21731.

The model has a chiral asymmetry of 12, which, after quotienting, is reduced to 3. It contains a number of $\mathbf{5} - \overline{\mathbf{5}}$ pairs, which after introducing Wilson lines lead to one (or possibly more than one) pair of Higgs doublets.

The above example is interesting as it satisfies the anomaly cancellation condition without the addition of any 5-branes. In this case,

$$c_2(TX).J_i = c_2(V).J_i = (24, 24, 24, 24, 24)$$

$X, |\Gamma|$	$k_\mathrm{m} = 1$	$k_\mathrm{m} = 2$	$k_\mathrm{m} = 3$	$k_\mathrm{m} = 4$	$k_\mathrm{m} = 5$	$k_\mathrm{m} = 6$	$k_\mathrm{m} = 7$	$k_\mathrm{m} = 8$	$k_\mathrm{m} = 9$	$k_\mathrm{m} = 10,$ 11, 12, 13
3413, 3	0	2278	2897	2906	2906	2906				
4190, 2	11	766	1175	1243	1246	1247	1249	1249	1249	
5273, 2	29	4895	7149	7738	7799	7810	7810	7810		
5302, 2	0	4314	5978	6360	6369	6369	6369			
5302, 4	0	11705	16988	17687	17793	17838	17868	17868	17868	
5425, 2	0	2381	3083	3305	3337	3337	3337			
5958, 2	0	148	224	240	253	253	253			
6655, 5	0	92	178	189	194	194	198	201	202	203
6738, 2	1	2733	4116	4346	4386	4393	4399	4399	4399	

Table 6. Number of models as a function of k_m on CICYs with $h^{1,1}(X) = 6$. Total number of models: 41036.

As the ranks of V and TX are the same, and their second Chern classes match, one could study the interesting problem[3] of deforming V to TX, which would bring us back to the standard embedding. Our database contains 348 such models which saturate the inequality (4.4).

8   Final comments and outlook

In this paper, we have presented the results of a comprehensive scan over heterotic line bundle models on favourable complete intersection Calabi-Yau manifolds (CICYs) with freely-acting symmetries. There are 68 such manifolds with $h^{1,1}(X) = 2, \ldots, 6$ contained in the standard list of CICY three-folds [57, 68] available at [69]. We have focused on rank five line bundle sums, leading to SU(5) GUT models, and a scan over about 10^{40} configurations has produced 63325 consistent and physically viable such models, available here [70]. Furthermore, we have shown computationally that this exhausts the set of physically viable line bundle models on the aforementioned class of CICYs. More precisely, by a consistent and physically viable model we mean a model with a poly-stable line bundle sum which allows for a global completion and whose chiral asymmetries have the correct values to produce a standard model upon taking the quotient by the freely-acting symmetry and including the Wilson line. When we require, in addition, the absence of $\overline{\mathbf{10}}$ multiplets and the presence of at least one $\mathbf{5}$–$\overline{\mathbf{5}}$ pair to account for the Higgs the number of viable models is reduced to about 35000.

The task ahead involves constructing the standard models associated to these GUT models. From prior experience with a smaller data set [40, 41] we expect this will lead to a larger number of standard models compared to the number of GUT models. A number of technical hurdles have to be overcome in order to complete this task, notably devising and implementing a complete algorithm for computing (equivariant) line bundle cohomology on CICYs. This work is currently in progress. The resulting models will provide by far the largest data set of standard models in any type of string construction and they will

---

[3]This idea was suggested to one of us by S.-T. Yau in a private communication.

provide a starting point for a systematic study of phenomenological questions beyond the spectrum, such as proton decay, μ-problem and the structure of Yukawa-couplings.

We hope to report on the results of this ongoing work in the near future.

Acknowledgments

A   Bundle structure groups

In this appendix, we review some useful results regarding principal and vector bundle geometry in heterotic compactifications. In particular, we address the problem of how to determine the structure group, $H \subset E_8$, of a vector bundle without knowing an explicit form for the connection.

A.1   Principal bundles vs. vector bundles

In compactifications of the heterotic string, for each E_8 factor, the gauge fields over the Calabi-Yau threefold are specified by a *principal H-bundle*, $\mathcal{V}$, with $H \subset E_8$. Given an explicit embedding of H into E_8, $\mathcal{V}$ determines a collection of associated vector bundles, V_α, carrying specific representations of H, as determined by the decomposition of the **248** representation of E_8. For example, if a principal SU(3) bundle, $\mathcal{V}_{\mathrm{SU}(3)}$, is embedded into E_8 via the direct product $(E_6 \times \mathrm{SU}(3))/\mathbb{Z}_3$, then the decomposition of the adjoint representation of E_8 yields the following representations, carried by the corresponding trio of vector bundles with appropriate rank (fiber dimension):

$$
\begin{array}{c|c|c}
\mathbf{3} & \overline{\mathbf{3}} & \mathbf{8} \\
\hline
V_3 & V_3^* & \mathrm{End}_0(V_3)
\end{array}
\tag{A.1}
$$

Given the rank 3 vector bundle, V_3, in the fundamental representation of SU(3), we can straightforwardly build those bundles corresponding to the $\overline{\mathbf{3}}$ and the $\mathbf{8}$ by taking the dual or tensor products.

In practice, however, in building the background geometry for a heterotic compactification we do not explicitly construct the principal bundle $\mathcal{V}$, but rather, first, the vector bundle in the fundamental representation, and from it the full collection of vector bundles, V_α, in the relevant representations. Moreover, as an added difficulty, except in very special

cases, there are no tools available to explicitly construct the H-valued connections, ∇_α of the relevant vector bundles.[4]

Instead, our starting point is an explicit formal construction of a rank n holomorphic vector bundle (for example a sum of line bundles, or a bundle constructed via a monad [18], or by extension [71]). The question now becomes, can we be sure that the given collection of vector bundles really arose from an H-valued principle bundle? Suppose, for example, that we consider a holomorphic rank 3 vector bundle, V_3, with structure group $H \subset U(3)$ and $c_1(V_3) = 0$. We may be tempted to declare this an SU(3) vector bundle from this data alone. However, suppose further that the bundle satisfies the non-trivial condition that

$$V \simeq V^* \tag{A.2}$$

Now, from this new information, it is clear that the previous conclusion was too hasty. Since the **3** of SU(3) is not a real representation, it follows that no SU(3) vector bundle can satisfy the self-duality condition in (A.2). Instead, the given V_3 could actually be carrying the symmetric, **3**-representation of SU(2) (more precisely, it could correspond to $S^2 V_2$ for some fundamental, rank 2, SU(2)-bundle, V_2); or similarly, the **3** of an SO(3)-bundle. An obstruction of this type could occur for any vector bundle in the collection V_α, and we must make sure that no such topological obstacles exist in building a bundle with the desired structure group.

In this work, we focus on SU(5) principal bundles breaking E_8 to an SU(5) GUT symmetry in 4-dimensions via,

$$\mathbf{248}_{E_8} \to [(\mathbf{1}, \mathbf{24}) \oplus (\mathbf{5}, \overline{\mathbf{10}}) \oplus (\overline{\mathbf{5}}, \mathbf{10}) \oplus (\mathbf{10}, \mathbf{5}) \oplus (\overline{\mathbf{10}}, \overline{\mathbf{5}}) \oplus (\mathbf{24}, \mathbf{1})]_{\mathrm{SU}(\mathbf{5}) \times \mathrm{SU}(\mathbf{5})} \tag{A.3}$$

Thus, we must construct the associated vector bundles with fiber-dimensions corresponding the $\mathbf{5}, \overline{\mathbf{5}}, \mathbf{10}, \overline{\mathbf{10}}, \mathbf{24}$ representations (see table 2).

Beginning with the fundamental **5**-representation, for the vector bundles constructed in this work, we will check here that there are no obstructions, such as the one described above, which would prevent the sum of five line bundles, $\bigoplus_a L_a$, from having structure group $S\left(U(1)^{\otimes 5}\right)$.

We will outline in the following paragraphs a set of tools for determining the structure groups of rank n holomorphic vector bundles with structure group $H \subset U(n)$ and $c_1(V) = 0$. We will focus on distinguishing the groups $SU(n), Sp(2n)$ and $SO(n)$. The exceptional sub-groups of E_8 will not arise in the dimensions of representation in consideration here and we will omit them from this discussion.

A.2 Chern classes and structure groups

The first and most important ingredient we have in determining the structure group of a vector bundle is its topology. As a simple example, consider the following direct sums of two line bundles on a threefold X

V	$L_1 \oplus L_2$	$L \oplus L$	$L \oplus L^*$
H	$U(1) \times U(1)$	$U(1) \times U(1)$ or $U(1)$	$S[U(1) \times U(1)] = U(1)$

$$\tag{A.4}$$

---

[4]For numeric approaches to this problem see [45–47].

For the first sum of line bundles, $c_1(L_1) \neq c_1(L_2)$ implies that for all possible connections on this sum, the structure group is $U(1) \times U(1)$. However, for the sum of two identical line bundles with the same first Chern class, $L \oplus L$, there is some flexibility in the choice of connection. For generic, independent, U(1)-valued connections, the structure group likewise is generic, that is, $U(1) \times U(1)$. For this topology, however, a non-generic choice is also available, and by choosing the two connections $\nabla_1 = \nabla_2$, the structure group is simply U(1). Finally, in the last example, the sum of a line bundle and its dual, the only structure group compatible with the reducible connection and vanishing trace condition is the diagonal $U(1) \subset SU(2)$.

For phenomenology we require that the low energy GUT symmetry in 4-dimensions is SU(5) times possible U(1) factors. So long as the commutant of H is of this form, $SU(5) \times S(U(1)^5) \subset E_8$, the Green-Schwarz Mechanism will guarantee that the U(1) symmetries are generically massive (see [40, 41, 52, 53]). Just as in the case of two line bundles described above, here we must guarantee that the topology of our sum does not force a smaller sub-group than $S\left(U(1)^{\otimes 5}\right)$ in such a way that the commutant contains other non-Abelian factors beyond SU(5). For example, if the sum of 5 line bundles satisfies

$$c_1(L_1) + c_1(L_2) + c_1(L_3) = 0 \quad , \quad c_1(L_4) + c_1(L_5) = 0 \tag{A.5}$$

then structure group is $H = S\left(U(1)^{\otimes 3}\right) \times S\left(U(1)^{\otimes 2}\right) \simeq U(1)^{\otimes 3}$, but its commutant in E_8 is $SU(6) \times U(1)^{\otimes 3}$ which would not be suitable for model-building. Thus, in the scans outlined in the main body of the text, we have, in addition to $\sum_a c_1(L_a) = 0$, imposed that

$$\sum_{a \in S} c_1(L_a) \neq 0 \text{ for all proper subsets } S \subset \{1, \dots, 5\} . \tag{A.6}$$

Finally, having eliminated the possibility of undesirable sub-groups of $S\left(U(1)^{\otimes 5}\right)$ we must still worry about accidental isomorphisms of the form described in (A.2), which could force the structure group to be non-unitary and perhaps even larger than $S\left(U(1)^{\otimes 5}\right)$.

A.3 Hermitian, real and symplectic fiber structures

To guarantee that the fibers of V carry the SU(5) representations given in (A.3), we must check that no other topological obstructions, beyond the Chern class conditions described above, exist which could force a different structure group. To begin, we note that we can distinguish between the classical simple groups: $SU(n), SO(n)$ and $Sp(2n)$ by determining whether the vector bundles carry more than the standard Hermitian fiber metric (characteristic of U(n) bundles [55]), but also symplectic or real fiber structures (see [56] for a review). For example, an $Sp(2n)$-bundle can be represented by a rank $2n$ holomorphic vector bundle (with trivial determinant) equipped with a skew-symmetric, holomorphic pairing, $V \otimes V \rightarrow \mathbb{C}$. The pairing can be viewed as an isomorphism $\varphi : V \rightarrow V^*$ which is skew-symmetric and non-degenerate on each fiber. The morphism φ is referred to as a "symplectic fiber structure". The case of an $SO(n)$-bundle is identical for rank n holomorphic bundles, where in this case the morphism φ is symmetric and forms a "real fiber

structure". These conditions are summarized in (A.7).

$$
\begin{array}{c|c|c|c}
\mathrm{Sp}(2n) & \varphi : V \hookrightarrow V^* & \varphi^* = -\varphi & \varphi \in H^0(X, \wedge^2 V) \\
\hline
\mathrm{SO}(n) & \varphi : V \hookrightarrow V^* & \varphi^* = \varphi & \varphi \in H^0(X, S^2 V)
\end{array}
\tag{A.7}
$$

Returning to the case of $V = \bigoplus_{a=1}^5 L_a$, no exceptional sub-group of E_8 carries **5**-dimensional representations, so to guarantee that $H \subset S\left(\mathrm{U}(1)^{\otimes 5}\right)$ we have only to eliminate the possibility that V_5 corresponds to the **5** of either $\mathrm{SO}(5)$ or $\mathrm{Sp}(4)$. However, from the conditions above in (A.7) it is clear that this is only possible if $V \simeq V^*$.

Happily, for the sum of 5 line bundles considered here such an isomorphism is only possible if $L_a = \mathcal{O}_X$ for at least one $a \in \{1, \dots 5\}$. This possibility can be explicitly excluded in scans by demanding that for all $a \in \{1, \dots, 5\}$

$$
c_1^r(L_a) \neq 0 \quad \text{for at least one value of } r .
\tag{A.8}
$$

Thus, for a sum of five holomorphic line bundles, satisfying $\sum_a c_1(L_a) = 0$ in which each summand is a non-trivial line bundle, both real and symplectic fiber structures are not possible.

Having eliminated the possibility of an $\mathrm{SO}(n)$ or $\mathrm{Sp}(n)$ structure group, via the condition (A.8), and $V \neq V^*$, and with no exceptional group representations of the appropriate dimension, by process of elimination we have determined that the structure group of $V = \bigoplus_{a=1}^5 L_a$ satisfies $H \subset \mathrm{SU}(5)$. Combining this with the condition (A.6) to exclude undesirable subgroups which might lead to non-Abelian commutants in E_8, we have a necessary set of conditions to guarantee a $4d$ GUT symmetry of $\mathrm{SU}(5) \times \mathrm{U}(1)^4$.

B Favourable embeddings

In section 3 we noted that the line bundle scan has been carried out over the class of favourable CICYs, that is, CICYs for which the entire second cohomology descends from the ambient space. We would now briefly like to discuss the precise meaning of this property as well as some criteria which can be used to decide whether a given CICY is favourable. We begin with a CICY, X, defined in the ambient space $\mathcal{A} = \bigotimes_{r=1}^m \mathbb{P}^{n_r}$, as the common zero locus of certain polynomials which can be thought of as sections of the line bundle sum $\mathcal{N}$ on $\mathcal{A}$. We denote the restriction of $\mathcal{N}$ to X by $N = \mathcal{N}|_X$ and also introduce the bundle $S = \bigoplus_{r=1}^m \mathcal{O}_X(\mathbf{e}_r)^{\oplus(n_r+1)}$, where $\mathbf{e}_r$ are the standard unit vectors in m dimensions.

The tangent bundle TX of the CICY X can be obtained from the two short exact sequences

$$
0 \to TX \to T\mathcal{A}|_X \to N \to 0 , \quad 0 \to \mathcal{O}_X^{\oplus m} \to S \to T\mathcal{A}|_X \to 0 .
\tag{B.1}
$$

Noting that $H^{1,1}(X) \cong H^2(X, TX)$ and $H^3(X, TX) \cong H^{0,1}(X) = 0$ the two associated long exact sequences lead to the following relations for the second cohomology of X

$$
H^{1,1}(X) \cong \mathrm{Coker}\left(H^1(X, S) \to H^1(X, N)\right) \oplus \mathrm{Ker}\left(H^2(X, T\mathcal{A}|_X) \to H^2(X, N)\right)
\tag{B.2}
$$

$$
H^2(X, T\mathcal{A}|_X) \cong H^2(X, S) \oplus \mathrm{Ker}\left(\mathbb{C}^m \to H^3(X, S)\right) .
\tag{B.3}
$$

The part of $H^{1,1}(X)$ which descends from the second ambient space cohomology corresponds to the $\mathbb{C}^m$ term in the second equation. Hence, the precise conditions for the CICY X to be favourable are

$$\text{Coker}\left(H^1(X,S) \to H^1(X,N)\right) = 0 \,, \quad H^2(X,S) = 0 \,. \tag{B.4}$$

In particular, this means a CICY with $h^{1,1}(X) > m$ or $h^1(X,S) < h^1(X,N)$ or $h^2(X,S) > 0$ is not favourable. A sufficient, however slightly too strong, condition for X to be favourable is

$$h^1(X,N) = h^2(X,S) = 0 \,, \tag{B.5}$$

where the first of these conditions guarantees that the Coker in (B.4) vanishes. Eq. (B.5) can be checked relatively easily since it only involves cohomologies of line bundles on X and we, therefore, adopt it as our practical definition of favourability.

C The distribution of models according to $(X, |\Gamma|)$

X, $|\Gamma|$	GUT models	no $\overline{\mathbf{10}}$ multiplets	no $\overline{\mathbf{10}}$s and at least one $\mathbf{5} - \overline{\mathbf{5}}$ pair	no $\overline{\mathbf{10}}$s and equivariance check for individual line bundles
7484, 4	1	1	1	1
7669, 3	2	2	0 (2)	2
7669, 9	1	1	0 (1)	1
7735, 8	1	1	1	0
7745, 8	1	1	1	0

Table 7. Number of models on CICYs with $h^{1,1}(X) = 3$.

X, $|\Gamma|$	GUT models	no $\overline{\mathbf{10}}$ multiplets	no $\overline{\mathbf{10}}$s and at least one $\mathbf{5} - \overline{\mathbf{5}}$ pair	no $\overline{\mathbf{10}}$s and equivariance check for individual line bundles
6784, 2	12	10	10	10
6784, 4	70	59	59	55
6828, 2	6	6	6	6
6828, 4	35	33	33	31
6831, 2	2	2	2	2
7204, 2	22	14	14	14
7218, 2	11	11	11	11
7241, 2	11	9	9	9
7245, 2	4	4	4	4
7247, 3	59	42 (14)	22 (4)	38
7270, 2	22	18	18	18
7403, 2	6	4 (2)	0 (3)	2
7435, 2	2	2	2	2
7435, 4	10	9	9	7
7462, 2	6	6	6	6
7462, 4	30	16	16	14
7468, 2	7	6	5	6
7491, 2	2	2	2	2
7491, 4	10	4	4	4
7522, 2	6	6	6	6
7522, 4	30	21	21	17
7719, 2	26	24	24	24
7736, 2	13	12	12	12
7742, 2	13	12	12	12
7862, 2	10	10	8	10
7862, 4	58	53	46	44
7862, 8	64	52	36	10
7862, 16	5	5	4	0

Table 8. Number of models on CICYs with $h^{1,1}(X) = 4$.

X, $|\Gamma|$	GUT models	no $\overline{\mathbf{10}}$ multiplets	no $\overline{\mathbf{10}}$s and at least one $\mathbf{5} - \overline{\mathbf{5}}$ pair	no $\overline{\mathbf{10}}$s and equivariance check for individual line bundles
5256, 2	763	625 (12)	480 (65)	625
5256, 4	2128	1812 (23)	1485 (167)	1444
5301, 2	191	178 (3)	87 (40)	178
5301, 4	534	504 (6)	323 (82)	406
5452, 2	762	547 (11)	497 (25)	547
5452, 4	2122	1624 (17)	1518 (71)	1278
6024, 3	509	244 (69)	215 (29)	237
6204, 2	119	96 (14)	76 (17)	93
6225, 2	229	137 (21)	118 (17)	133
6715, 2	184	170 (0)	138 (4)	170
6715, 4	847	711 (4)	539 (76)	457
6724, 2	39	32 (7)	20 (10)	21
6732, 2	880	667 (6)	532 (60)	667
6770, 2	330	271 (0)	197 (39)	271
6777, 2	880	587 (6)	549 (32)	587
6788, 2	184	155 (0)	147 (4)	155
6788, 4	848	621 (4)	579 (28)	397
6802, 2	877	786 (6)	524 (128)	786
6804, 2	141	108 (4)	99 (5)	101
6834, 2	441	371 (3)	283 (47)	371
6836, 2	46	37 (0)	36 (1)	37
6836, 4	214	151 (1)	147 (4)	97
6836, 8	136	109 (0)	97 (9)	14
6836, 16	3	3 (0)	2 (1)	0
6890, 2	1750	1245 (12)	1091 (83)	1245
6896, 2	441	421 (3)	232 (88)	421
6927, 2	276	243 (0)	218 (6)	243
6927, 4	1264	983 (6)	856 (67)	628
6927, 8	798	659 (5)	510 (79)	81
6947, 2	46	45 (0)	30 (1)	45
6947, 4	214	196 (1)	105 (35)	127
6947, 8	136	125 (0)	44 (19)	21
6947, 16	3	3 (0)	2 (1)	0
7279, 2	218	109 (49)	96 (10)	108
7447, 2	93	89 (0)	45 (15)	89
7447, 4	430	396 (2)	182 (77)	306
7447, 5	0	0 (0)	0 (0)	0
7447, 10	81	76 (0)	12 (19)	0
7447, 20	0	0 (0)	0 (0)	0
7487, 2	459	319 (0)	261 (28)	319
7487, 4	2115	1505 (8)	1257 (94)	1136

Table 9. Number of models on CICYs with $h^{1,1}(X) = 5$.

X, $|\Gamma|$	GUT models	no $\overline{\mathbf{10}}$ multiplets	no $\overline{\mathbf{10}}$s and at least one $\mathbf{5} - \overline{\mathbf{5}}$ pair	no $\overline{\mathbf{10}}$s and equivariance check for individual line bundles
3413, 3	1737	709 (516)	599 (98)	698
4190, 2	1145	540 (195)	473 (57)	429
5273, 2	6753	4154 (934)	3292 (701)	3757
5302, 2	6294	4130 (246)	3291 (456)	4130
5302, 4	17329	13242 (82)	10174 (1678)	9235
5425, 2	3128	1946 (533)	1358 (409)	1802
5958, 2	246	215 (23)	103 (66)	179
6655, 5	161	143 (15)	67 (64)	1
6738, 2	4243	1846 (743)	1599 (169)	1763

Table 10. Number of models on CICYs with $h^{1,1}(X) = 6$.

References

[1] D.J. Gross, J.A. Harvey, E.J. Martinec and R. Rohm, *The Heterotic String*, *Phys. Rev. Lett.* **54** (1985) 502 [INSPIRE].

[2] D.J. Gross, J.A. Harvey, E.J. Martinec and R. Rohm, *Heterotic String Theory. 1. The Free Heterotic String*, *Nucl. Phys.* **B 256** (1985) 253 [INSPIRE].

[3] D.J. Gross, J.A. Harvey, E.J. Martinec and R. Rohm, *Heterotic String Theory. 2. The Interacting Heterotic String*, *Nucl. Phys.* **B 267** (1986) 75 [INSPIRE].

[4] P. Candelas, G.T. Horowitz, A. Strominger and E. Witten, *Vacuum Configurations for Superstrings*, *Nucl. Phys.* **B 258** (1985) 46 [INSPIRE].

[5] B.R. Greene, K.H. Kirklin, P.J. Miron and G.G. Ross, *A Three Generation Superstring Model. 1. Compactification and Discrete Symmetries*, *Nucl. Phys.* **B 278** (1986) 667 [INSPIRE].

[6] B.R. Greene, K.H. Kirklin, P.J. Miron and G.G. Ross, *A Three Generation Superstring Model. 2. Symmetry Breaking and the Low-Energy Theory*, *Nucl. Phys.* **B 292** (1987) 606 [INSPIRE].

[7] V. Braun, P. Candelas and R. Davies, *A Three-Generation Calabi-Yau Manifold with Small Hodge Numbers*, *Fortsch. Phys.* **58** (2010) 467 [arXiv:0910.5464] [INSPIRE].

[8] V. Braun, P. Candelas, R. Davies and R. Donagi, *The MSSM Spectrum from (0,2)-Deformations of the Heterotic Standard Embedding*, *JHEP* **05** (2012) 127 [arXiv:1112.1097] [INSPIRE].

[9] J. Distler and B.R. Greene, *Aspects of (2,0) String Compactifications*, *Nucl. Phys.* **B 304** (1988) 1 [INSPIRE].

[10] J. Distler and S. Kachru, *(0,2) Landau-Ginzburg theory*, *Nucl. Phys.* **B 413** (1994) 213 [hep-th/9309110] [INSPIRE].

[11] S. Kachru, *Some three generation (0,2) Calabi-Yau models*, *Phys. Lett.* **B 349** (1995) 76 [hep-th/9501131] [INSPIRE].

[12] V. Braun, Y.-H. He, B.A. Ovrut and T. Pantev, *A Heterotic standard model*, *Phys. Lett.* **B 618** (2005) 252 [hep-th/0501070] [INSPIRE].

[13] V. Braun, Y.-H. He, B.A. Ovrut and T. Pantev, *A Standard model from the $E_8 \times E_8$ heterotic superstring*, *JHEP* **06** (2005) 039 [hep-th/0502155] [INSPIRE].

[14] V. Braun, Y.-H. He, B.A. Ovrut and T. Pantev, *The Exact MSSM spectrum from string theory*, *JHEP* **05** (2006) 043 [hep-th/0512177] [INSPIRE].

[15] V. Bouchard and R. Donagi, *An* SU(5) *heterotic standard model*, *Phys. Lett.* **B 633** (2006) 783 [hep-th/0512149] [INSPIRE].

[16] R. Blumenhagen, S. Moster and T. Weigand, *Heterotic GUT and standard model vacua from simply connected Calabi-Yau manifolds*, *Nucl. Phys.* **B 751** (2006) 186 [hep-th/0603015] [INSPIRE].

[17] R. Blumenhagen, S. Moster, R. Reinbacher and T. Weigand, *Massless Spectra of Three Generation U(N) Heterotic String Vacua*, *JHEP* **05** (2007) 041 [hep-th/0612039] [INSPIRE].

[18] L.B. Anderson, Y.-H. He and A. Lukas, *Heterotic Compactification, An Algorithmic Approach*, *JHEP* **07** (2007) 049 [hep-th/0702210] [INSPIRE].

[19] L.B. Anderson, Y.-H. He and A. Lukas, *Monad Bundles in Heterotic String Compactifications*, *JHEP* **07** (2008) 104 [arXiv:0805.2875] [INSPIRE].

[20] L.B. Anderson, J. Gray, Y.-H. He and A. Lukas, *Exploring Positive Monad Bundles And A New Heterotic Standard Model*, *JHEP* **02** (2010) 054 [arXiv:0911.1569] [INSPIRE].

[21] W. Buchmüller, K. Hamaguchi, O. Lebedev and M. Ratz, *Supersymmetric standard model from the heterotic string*, *Phys. Rev. Lett.* **96** (2006) 121602 [hep-ph/0511035] [INSPIRE].

[22] W. Buchmüller, K. Hamaguchi, O. Lebedev and M. Ratz, *Supersymmetric Standard Model from the Heterotic String (II)*, *Nucl. Phys.* **B 785** (2007) 149 [hep-th/0606187] [INSPIRE].

[23] O. Lebedev, H.P. Nilles, S. Raby, S. Ramos-Sanchez, M. Ratz et al., *A Mini-landscape of exact MSSM spectra in heterotic orbifolds*, *Phys. Lett.* **B 645** (2007) 88 [hep-th/0611095] [INSPIRE].

[24] O. Lebedev, H.P. Nilles, S. Raby, S. Ramos-Sanchez, M. Ratz et al., *The Heterotic Road to the MSSM with R parity*, *Phys. Rev.* **D 77** (2008) 046013 [arXiv:0708.2691] [INSPIRE].

[25] O. Lebedev, H.P. Nilles, S. Ramos-Sanchez, M. Ratz and P.K. Vaudrevange, *Heterotic mini-landscape. (II). Completing the search for MSSM vacua in a Z(6) orbifold*, *Phys. Lett.* **B 668** (2008) 331 [arXiv:0807.4384] [INSPIRE].

[26] J.E. Kim, J.-H. Kim and B. Kyae, *Superstring standard model from Z(12-I) orbifold compactification with and without exotics and effective R-parity*, *JHEP* **06** (2007) 034 [hep-ph/0702278] [INSPIRE].

[27] S. Groot Nibbelink, J. Held, F. Ruehle, M. Trapletti and P.K. Vaudrevange, *Heterotic Z(6-II) MSSM Orbifolds in Blowup*, *JHEP* **03** (2009) 005 [arXiv:0901.3059] [INSPIRE].

[28] M. Blaszczyk, S. Groot Nibbelink, M. Ratz, F. Ruehle, M. Trapletti et al., *A Z2xZ2 standard model*, *Phys. Lett.* **B 683** (2010) 340 [arXiv:0911.4905] [INSPIRE].

[29] M. Blaszczyk, S. Groot Nibbelink, F. Ruehle, M. Trapletti and P.K. Vaudrevange, *Heterotic MSSM on a Resolved Orbifold*, *JHEP* **09** (2010) 065 [arXiv:1007.0203] [INSPIRE].

[30] R. Kappl, B. Petersen, S. Raby, M. Ratz, R. Schieren et al., *String-Derived MSSM Vacua with Residual R Symmetries*, *Nucl. Phys.* **B 847** (2011) 325 [arXiv:1012.4574] [INSPIRE].

[31] K.R. Dienes, *Statistics on the heterotic landscape: Gauge groups and cosmological constants of four-dimensional heterotic strings*, *Phys. Rev.* **D 73** (2006) 106010 [hep-th/0602286] [INSPIRE].

[32] K.R. Dienes, M. Lennek, D. Senechal and V. Wasnik, *Supersymmetry versus Gauge Symmetry on the Heterotic Landscape*, *Phys. Rev.* **D 75** (2007) 126005 [arXiv:0704.1320] [INSPIRE].

[33] K.R. Dienes, M. Lennek, D. Senechal and V. Wasnik, *Is SUSY Natural?*, *New J. Phys.* **10** (2008) 085003 [arXiv:0804.4718] [INSPIRE].

[34] B. Assel, K. Christodoulides, A.E. Faraggi, C. Kounnas and J. Rizos, *Exophobic Quasi-Realistic Heterotic String Vacua*, *Phys. Lett.* **B 683** (2010) 306 [arXiv:0910.3697] [INSPIRE].

[35] K. Christodoulides, A.E. Faraggi and J. Rizos, *Top Quark Mass in Exophobic Pati-Salam Heterotic String Model*, *Phys. Lett.* **B 702** (2011) 81 [arXiv:1104.2264] [INSPIRE].

[36] G. Cleaver, A. Faraggi, J. Greenwald, D. Moore, K. Pechan et al., *Investigation of Quasi-Realistic Heterotic String Models with Reduced Higgs Spectrum*, *Eur. Phys. J.* **C 71** (2011) 1842 [arXiv:1105.0447] [INSPIRE].

[37] B. Gato-Rivera and A. Schellekens, *Heterotic Weight Lifting*, *Nucl. Phys.* **B 828** (2010) 375 [arXiv:0910.1526] [INSPIRE].

[38] B. Gato-Rivera and A. Schellekens, *Asymmetric Gepner Models II. Heterotic Weight Lifting*, *Nucl. Phys.* **B 846** (2011) 429 [arXiv:1009.1320] [INSPIRE].

[39] M. Maio and A. Schellekens, *Permutation orbifolds of heterotic Gepner models*, *Nucl. Phys.* **B 848** (2011) 594 [arXiv:1102.5293] [INSPIRE].

[40] L.B. Anderson, J. Gray, A. Lukas and E. Palti, *Two Hundred Heterotic Standard Models on Smooth Calabi-Yau Threefolds*, *Phys. Rev.* **D 84** (2011) 106005 [arXiv:1106.4804] [INSPIRE].

[41] L.B. Anderson, J. Gray, A. Lukas and E. Palti, *Heterotic Line Bundle Standard Models*, *JHEP* **06** (2012) 113 [arXiv:1202.1757] [INSPIRE].

[42] S.-T. Yau, *On the Ricci curvature of a compact Kahler manifold and the complex Monge-Ampere equation. I,*, *Comm. Pure Appl. Math.* **31** (1978) 339.

[43] S. Donaldson, *Anti-self-dual Yang-Mills connections over complex algebraic surfaces and stable vector bundles*, *Proc. Lond. Math. Soc.* **50** (1985) 1 [INSPIRE].

[44] K. Uhlenbeck and S.-T. Yau, *On the existence of hermitian Yang-Mills connections in stable vector bundles*, *Pure Appl. Math.* **39** (1986) 257.

[45] M.R. Douglas, R.L. Karp, S. Lukic and R. Reinbacher, *Numerical solution to the hermitian Yang-Mills equation on the Fermat quintic*, *JHEP* **12** (2007) 083 [hep-th/0606261] [INSPIRE].

[46] L.B. Anderson, V. Braun, R.L. Karp and B.A. Ovrut, *Numerical Hermitian Yang-Mills Connections and Vector Bundle Stability in Heterotic Theories*, *JHEP* **06** (2010) 107 [arXiv:1004.4399] [INSPIRE].

[47] L.B. Anderson, V. Braun and B.A. Ovrut, *Numerical Hermitian Yang-Mills Connections and Kähler Cone Substructure*, *JHEP* **01** (2012) 014 [arXiv:1103.3041] [INSPIRE].

[48] L.B. Anderson, J. Gray, A. Lukas and B. Ovrut, *Stabilizing All Geometric Moduli in Heterotic Calabi-Yau Vacua*, *Phys. Rev.* **D 83** (2011) 106011 [arXiv:1102.0011] [INSPIRE].

[49] L.B. Anderson, J. Gray, A. Lukas and B. Ovrut, *The Atiyah Class and Complex Structure Stabilization in Heterotic Calabi-Yau Compactifications*, *JHEP* **10** (2011) 032 [arXiv:1107.5076] [INSPIRE].

[50] L.B. Anderson, J. Gray, A. Lukas and B. Ovrut, *Vacuum Varieties, Holomorphic Bundles and Complex Structure Stabilization in Heterotic Theories*, *JHEP* **07** (2013) 017 [arXiv:1304.2704] [INSPIRE].

[51] V. Bouchard, M. Cvetič and R. Donagi, *Tri-linear couplings in an heterotic minimal supersymmetric standard model*, *Nucl. Phys.* **B 745** (2006) 62 [hep-th/0602096] [INSPIRE].

[52] A. Lukas and K. Stelle, *Heterotic anomaly cancellation in five-dimensions*, *JHEP* **01** (2000) 010 [hep-th/9911156] [INSPIRE].

[53] R. Blumenhagen, G. Honecker and T. Weigand, *Loop-corrected compactifications of the heterotic string with line bundles*, *JHEP* **06** (2005) 020 [hep-th/0504232] [INSPIRE].

[54] The database of 63325 GUT line bundle models on favourable Cicys with freely-acting symmetries can be found at http://www-thphys.physics.ox.ac.uk/projects/CalabiYau/linebundlemodels/index.html.

[55] D. Huybrechts, *Complex Geometry: An Introduction*, Universitext (1979), Springer (2005).

[56] J.W. Morgan, *Holomorphic bundles over elliptic manifolds.*, in *Moduli spaces in algebraic geometry. Lecture notes of the school on algebraic geometry*, Trieste, Italy, Jul. 26 – Aug. 13 (1999), pp. 135, Trieste: The Abdus Salam International Centre for Theoretical Physics (ICTP) (2000).

[57] P. Candelas, A. Dale, C. Lütken and R. Schimmrigk, *Complete Intersection Calabi-Yau Manifolds*, *Nucl. Phys.* **B 298** (1988) 493 [INSPIRE].

[58] P. Candelas and R. Davies, *New Calabi-Yau Manifolds with Small Hodge Numbers*, *Fortsch. Phys.* **58** (2010) 383 [arXiv:0809.4681] [INSPIRE].

[59] V. Braun, *On Free Quotients of Complete Intersection Calabi-Yau Manifolds*, *JHEP* **04** (2011) 005 [arXiv:1003.3235] [INSPIRE].

[60] L.B. Anderson, J. Gray, Y.-H. He and A. Lukas, *Exploring Positive Monad Bundles And A New Heterotic Standard Model*, *JHEP* **02** (2010) 054 [arXiv:0911.1569] [INSPIRE].

[61] Y.-H. He, S.-J. Lee and A. Lukas, *Heterotic Models from Vector Bundles on Toric Calabi-Yau Manifolds*, *JHEP* **05** (2010) 071 [arXiv:0911.0865] [INSPIRE].

[62] L.B. Anderson, Y.-H. He and A. Lukas, *Monad Bundles in Heterotic String Compactifications*, *JHEP* **07** (2008) 104 [arXiv:0805.2875] [INSPIRE].

[63] J. Gray, Y.-H. He, A. Ilderton and A. Lukas, *A New Method for Finding Vacua in String Phenomenology*, *JHEP* **07** (2007) 023 [hep-th/0703249] [INSPIRE].

[64] L.B. Anderson, Y.-H. He and A. Lukas, *Heterotic Compactification, An Algorithmic Approach*, *JHEP* **07** (2007) 049 [hep-th/0702210] [INSPIRE].

[65] D. Huybrechts and M. Lehn, *The Geometry of Moduli Spaces of Sheaves*, Cambridge Mathematical Library, Cambridge University Press (2010).

[66] M. Maruyama, *On boundedness of torsion free sheaves*, *J. Math. Kyoto Univ.* **21** (1981) 673.

[67] A. Langer, *Semistable principal G-bundles in positive characteristic*, `math/0312260` .

[68] P.S. Green, T. Hubsch and C.A. Lütken, *All Hodge Numbers of All Complete Intersection Calabi-Yau Manifolds*, *Class. Quant. Grav.* **6** (1989) 105 [INSPIRE].

[69] The database of 7890 complete intersection Calabi-Yau three-folds in products of projective spaces (CICYs) can be accessed at
http://www-thphys.physics.ox.ac.uk/projects/CalabiYau/cicylist/index.html.

[70] The database of heterotic line bundle standard models obtained in the previous scan can be accessed at
http://www-thphys.physics.ox.ac.uk/projects/CalabiYau/linebundlemodels/index.html.

[71] R. Friedman, *Algebraic Surfaces and Holomorphic Vector Bundles*, Universitext - Springer-Verlag, Springer Verlag (1998).