

Heterotic line bundle standard models

Lara B. Anderson,^a James Gray,^b Andre Lukas^c and Eran Palti^d

^aCenter for the Fundamental Laws of Nature, Jefferson Laboratory, Harvard University,
17 Oxford Street, Cambridge, MA 02138, U.S.A.

^bArnold-Sommerfeld-Center for Theoretical Physics, Department für Physik,
Ludwig-Maximilians-Universität München,
Theresienstraße 37, 80333 München, Germany

^cRudolf Peierls Centre for Theoretical Physics, Oxford University,
1 Keble Road, Oxford, OX1 3NP, U.K.

^dCentre de Physique Theorique, Ecole Polytechnique, CNRS,
91128 Palaiseau, France

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

ABSTRACT: In a previous publication, [arXiv:1106.4804](https://arxiv.org/abs/1106.4804), we have found 200 models from heterotic Calabi-Yau compactifications with line bundles, which lead to standard models after taking appropriate quotients by a discrete symmetry and introducing Wilson lines. In this paper, we construct the resulting standard models explicitly, compute their spectrum including Higgs multiplets, and analyze some of their basic properties. After removing redundancies we find about 400 downstairs models, each with the precise matter spectrum of the supersymmetric standard model, with one, two or three pairs of Higgs doublets and no exotics of any kind. In addition to the standard model gauge group, up to four Green-Schwarz anomalous U(1) symmetries are present in these models, which constrain the allowed operators in the four-dimensional effective supergravity. The vector bosons associated to these anomalous U(1) symmetries are massive. We explicitly compute the spectrum of allowed operators for each model and present the results, together with the defining data of the models, in a database of standard models accessible [here](#). Based on these results we analyze elementary phenomenological properties. For example, for about 200 models all dimension four and five proton decay violating operators are forbidden by the additional U(1) symmetries.

KEYWORDS: Superstrings and Heterotic Strings, Superstring Vacua

ARXIV EPRINT: [1202.1757](https://arxiv.org/abs/1202.1757)

Contents

1	Introduction	1
2	Line bundle standard models	4
2.1	The gauge group	4
2.2	The matter field sector	5
2.3	The moduli sector	6
2.4	The effective action	7
2.5	D-terms	9
2.6	Green-Schwarz anomaly cancellation	10
2.7	Masses of U(1) gauge bosons	11
3	The model database	12
3.1	An example model	12
3.1.1	The gravitational sector	12
3.1.2	The matter field sector	14
3.1.3	The D-terms	15
3.1.4	The singlet superpotential: F-terms and neutrino Majorana masses	17
3.1.5	The Higgs sector	18
3.1.6	Yukawa couplings	19
3.1.7	Proton decay	19
3.1.8	Bilinear R-parity violating operators and superpotential neutrino Dirac masses	20
3.1.9	Neutrino Kähler potential Dirac masses	20
3.1.10	The Giudice-Masiero term	20
3.1.11	Kinetic terms and soft masses	21
3.2	General phenomenology overview of database	21
4	The geometry of split heterotic models	23
4.1	General formalism	23
4.2	Group theory	25
4.3	Split bundles and stability	27
4.4	Spectrum of GUT theory	29
4.5	Discrete symmetries, equivariance and downstairs spectrum	30
4.6	Anomalies	32
4.7	Connecting different splitting types	33
4.8	The conditions imposed on the geometry	35
4.8.1	Consistency conditions on the bundles	35
4.8.2	Conditions on spectrum of GUT theory	36

5	Heterotic line bundle models on CICYs	38
5.1	Complete intersection Calabi-Yau three-folds	39
5.2	Line bundles on favourable CICYs	40
5.3	The scan	41
6	Further consequences of split heterotic geometry	41
6.1	Line bundle models with non-trivial equivariant blocks	41
6.2	Up-type Yukawa couplings	43
6.3	Realizing a Peccei-Quinn symmetry	44
7	Summary	45
A	Equivariant structures and Wilson lines	46
B	Representation content of cohomology	50

1 Introduction

Compactifications of the heterotic string on Calabi-Yau manifolds, despite being the oldest approach to string phenomenology [1, 2], remains one of the most promising and well-understood paths to obtaining realistic string vacua. These models can combine the attractive ideas of grand unification with a large top Yukawa coupling, features which have proved to be difficult to realize in other types of models, particularly those based on type II string theory. In essence, this leaves the heterotic string, F-theory and the lesser studied G_2 compactifications of M-theory as primary starting points for string phenomenology.

Traditionally, heterotic Calabi-Yau model building has been based on the standard embedding [3–5] whereby the Bianchi identity is solved by setting the internal gauge bundle V equal to the tangent bundle, TX , of the Calabi-Yau manifolds X . However, over the past decade it has been realized that this approach is too restrictive and the focus has shifted to the wider class of non-standard embedding models [6]–[18], where V is a more general bundle over X . Only a relatively small number of models exhibiting a realistic massless spectrum have been constructed in this way, some with [6] and some without [5, 11, 18] antibranes, reflecting the considerable technical problems associated with vector bundles on smooth Calabi-Yau manifolds. They are complemented by models found in related heterotic constructions such as those based on orbifolds [19–28], on the free fermionic strings [29–31], and on Gepner models [32–34].

This situation has changed with the results published in ref. [35] where 200 heterotic Calabi-Yau GUT models were presented. By verifying a number of general criteria it was shown that each of these models leads to heterotic standard models upon suitable quotienting by discrete symmetries and including Wilson lines. This progress has been possible for two main reasons. Firstly, rather than following a “model building approach” by trying to fine-tune individual models for the right phenomenological properties, systematic scans,

using methods of computational algebraic geometry, have been performed over large classes of models and unsuitable candidates have been successively filtered out. The considerable mathematical and computational tools necessary for such systematic scans have been built up over a number of years [16–18, 36–45]. The second reason is related to the nature of the vector bundles V used in the construction. Previous model building attempts [6–18] have mostly focused on vector bundles V with a non-Abelian structure group. However, once we move away from the standard embedding, the complexity of the constructions rather motivates studying the simplest bundle choices, that is, bundles with Abelian structure groups. Using such Abelian bundles is one of the key ideas underlying the work in ref. [35] and the present paper. The technical simplifications which arise in the case of Abelian bundles greatly facilitate the systematic scanning and the construction of a sizeable number of promising models.

The 200 models given in ref. [35] were essentially constructed at the “upstairs” GUT level. The structure group of the bundle V on the Calabi-Yau manifolds X was chosen to be $S(U(1)^5) \subset E_8$ so that the low-energy gauge group is $SU(5) \times S(U(1)^5)$, with the additional $U(1)$ symmetries being Green-Schwarz anomalous (and hence, spontaneously broken) in most cases. The GUT matter spectrum for all models consists of $3|\Gamma|$ $\mathbf{10}$ and $\mathbf{\bar{5}}$ multiplets, some number of $\mathbf{5}-\mathbf{\bar{5}}$ vector-like pairs and a number of $SU(5)$ singlets. Here $|\Gamma|$ is the order of a freely-acting discrete symmetry Γ on X . It is clear that quotienting these models by Γ and including Wilson lines in order to break $SU(5)$ can lead to a low-energy theory with the standard model group (times anomalous $U(1)$ symmetries with massive associated gauge bosons) and three families of quarks and leptons. Further, provided certain constraints on the number of $\mathbf{5}-\mathbf{\bar{5}}$ pairs hold one can ensure that all Higgs triplets can be projected out and at least one pair of Higgs doublets can be kept. In ref. [35] it was shown that these constraints are indeed satisfied, so that all 200 models lead to heterotic standard models without any exotic fields charged under the standard model group.

In the present paper we go one step further and construct the downstairs standard models which result from the 200 GUT models of ref. [35] explicitly. We compute the complete spectrum, including Higgs multiplets and gauge singlet fields, for each model, thereby determining the $S(U(1)^5)$ charges for all multiplets. Taking into account all different choices of quotienting the bundle and including the Wilson lines, this leads to tens of thousands of downstairs models. In this paper we focus on the four-dimensional spectrum of particles and operators and, hence, we identify models which descend from the same upstairs theory if they lead to the same four-dimensional fields. After removing these and some other redundancies we find about 400 models, each with the standard model gauge group times $S(U(1)^5)$, precisely three families of quarks and leptons, between one and three pairs of Higgs doublets and no exotic fields charged under the standard model group of any kind. In addition, we have a number of standard model singlet fields, S^α , which are charged under $S(U(1)^5)$. To the best of our knowledge, this is the largest set of string models with precisely the standard model spectrum found to date. Details of all models can be found in the standard model database [46].

From a 10-dimensional point of view the singlet fields S^α can be interpreted as bundle moduli, where vanishing vacuum expectation values for S^α correspond to the original

Abelian gauge bundle and non-zero vacuum expectation values indicate a deformation to a bundle with non-Abelian structure group. We would like to stress that, despite the presence of the additional $S(U(1)^5)$ symmetry, there is no problem with additional massless vector bosons. For most models, all additional $U(1)$ symmetries are Green-Schwarz anomalous [2, 47–52] and, hence, the associated gauge bosons are super-heavy. If one of the $U(1)$ symmetries remains non-anomalous (and the associated gauge boson is massless), as happens in some cases, it can easily be spontaneously broken by turning on vacuum expectation values for the singlet fields S^α . As discussed above, this corresponds to deforming the gauge bundle to a one with a non-Abelian structure group.

Despite the similarity of their low-energy field content our models are distinct in a number of ways. Most importantly, the $S(U(1)^5)$ charges of matter and Higgs multiplets can vary between models. In addition, the numbers and $S(U(1)^5)$ charges of the singlets S^α are model-dependent, as is the number of Higgs doublet pairs. Taking this into account, we find 74 different spectra among the 400 models. However, even models with an identical four-dimensional spectrum have a different higher-dimensional origin and can, therefore, be expected to differ at a more sophisticated level, for example in the values of their coupling constants. For this reason, we have kept all 400 models in our database [46].

Our models fall within a general class of four-dimensional $N = 1$ supergravity theories obtained from heterotic line bundle compactifications which we would like to refer to as *line bundle standard models*. From a four-dimensional point of view, these models are characterized by an NMSSM-type spectrum (however, with generally many rather than just one singlet field), the presence of an additional Green-Schwarz anomalous $S(U(1)^5)$ symmetry and a specific pattern of charges under this symmetry. The presence of these additional (spontaneously broken) symmetries constrains the allowed operators in the four-dimensional theory and thereby facilitates the study of phenomenological properties beyond the computation of the matter spectrum. They can be phenomenologically helpful, for example by forbidding proton decay violating operators, or phenomenologically dangerous, for example if they force all Yukawa couplings to vanish. A wide range of phenomenological issues, including flavour physics, proton decay, the μ term, R-parity violation and neutrino masses can be addressed in this way. In ref. [35] this was carried out for a particular example. In the present paper, we compute the allowed set of operators in the four-dimensional theory for all 400 models and the results are listed in the database [46]. These results allow for a more detailed study of the models' phenomenology and we discuss a number of generic features based on these results. For example, we find that 45 of our models allow for an up Yukawa matrix with non-vanishing rank, before switching on singlet vacuum expectation values. For about 200 of our models, all dimension four and five proton-decay violating operators are forbidden by the $S(U(1)^5)$ symmetry. We have 259, 72 and 63 models, respectively, with one, two and three pairs of Higgs doublets. Requiring precisely one pair of Higgs doublets, the absence of all dimension four and five proton decay violating operators, an up-Yukawa matrix with non-zero rank and no massless $U(1)$ vector boson (in the absence of singlet VEVs), 13 models remain.

Because of the somewhat technical nature of the underlying 10-dimensional construction we have split the paper into two parts which can largely be read separately. The first

part, which consists of sections 2 and 3, describes heterotic line bundle models purely from the perspective of the four-dimensional $N = 1$ supergravity theory. In section 2, we set up the general structure of these four-dimensional models. Section 3 presents an example model from the database [46]. This chosen example is not selected for its phenomenological viability, but simply used as an explicit case to discuss various phenomenological issues and explain the structure of the data files. We end the section with an overview of basic phenomenological properties among our 400 standard models. The remainder of the paper describes the construction of the models starting with the 10-dimensional theory. In section 4, we set up the general formalism for heterotic Calabi-Yau compactifications in the presence of vector bundles with split structure groups. We also explain our scanning criteria and procedure in general. Section 5 describes our specific arena for the construction of models, that is, complete intersection Calabi-Yau manifolds (CICYs) and line bundles thereon, as well as details of the scanning procedure. A number of specific issues which arise in heterotic Calabi-Yau models with split bundles is discussed in section 6. Our summary and outlook is presented in section 7. Appendices A and B contain additional technical information on the construction of equivariant structures and the computation of equivariant cohomology.

2 Line bundle standard models

In this section, we introduce the general class of four-dimensional $N = 1$ supergravity theories with a standard model spectrum, derived from heterotic line bundle compactifications on Calabi-Yau manifolds. We will refer to this class of supergravities as line bundle standard models. This sets the scene for the discussion in section 3, where we present an explicit example from our standard model data base and a general phenomenological overview of our models. In addition, this class of supergravities provides a general framework for string phenomenology within a purely four-dimensional setting. Indeed, we expect many more line bundle models to exist than are currently available in our database [46], constructed by considering more general line bundles and other Calabi-Yau manifolds. All of these models will be described by a supergravity of the type introduced below.

2.1 The gauge group

The gauge group of line bundle standard models is given by the standard model group $G_{\text{SM}} = \text{SU}_c(3) \times \text{SU}_W(2) \times U_Y$ times the additional gauge symmetry $\mathcal{J} = \text{S}(\text{U}(1)^5)$. We can think of the elements of \mathcal{J} as given by five phases $(e^{i\eta^1}, \dots, e^{i\eta^5})$ subject to the “determinant one” condition $\sum_{a=1}^5 \eta^a = 0$. Although $\mathcal{J} \cong \text{U}(1)^4$ it will be more convenient for our purposes to work with \mathcal{J} rather than $\text{U}(1)^4$. Irreducible \mathcal{J} representations can be labelled by an integer vector $\mathbf{q} = (q_a)_{a=1, \dots, 5}$. However, due to the determinant one condition two such vectors, \mathbf{q} and $\tilde{\mathbf{q}}$ refer to the same representation and, hence, have to be identified iff

$$\mathbf{q} - \tilde{\mathbf{q}} \in \mathbb{Z}(1, 1, 1, 1, 1). \quad (2.1)$$

In particular, this means that a four-dimensional operator is \mathcal{J} invariant precisely if the five entries in its charge vector are identical. All standard model multiplets carry \mathcal{J} charges

which follow a specific pattern originating from the underlying string construction. This structure of charges will be introduced below.

We stress that the four gauge bosons associated to \mathcal{J} do not cause a phenomenological problem. In most cases, all U(1) symmetries are Green-Schwarz anomalous and, hence, the gauge bosons receive a super-heavy Stueckelberg mass. In cases where some of the U(1) symmetries are non-anomalous masses for the associated gauge bosons can be generated by spontaneously symmetry breaking through VEVs of standard model singlet fields. This will be discussed in more detail in the section on U(1) vector boson masses below.

2.2 The matter field sector

Matter fields transform linearly under \mathcal{J} , that is,

$$\Phi \rightarrow \exp(i\mathbf{q}\cdot\boldsymbol{\eta})\Phi \tag{2.2}$$

for a matter field Φ with \mathcal{J} charge $\mathbf{Q}(\Phi) = \mathbf{q}$. Although there is no four-dimensional SU(5) GUT symmetry it turns out that the \mathcal{J} charge is always identical for all fields in a given SU(5) multiplet. For this reason, it is useful to combine the three standard model families into SU(5) multiplets and introduce the notation $(\mathbf{10}^p) = (Q^p, u^p, e^p)$ and $(\bar{\mathbf{5}}^p) = (d^p, L^p)$, where $p, q, \dots = 1, 2, 3$ are family indices. Their pattern of \mathcal{J} charges is given by

$$\mathbf{Q}(\mathbf{10}^p) = \mathbf{e}_{a_p}, \quad \mathbf{Q}(\bar{\mathbf{5}}^p) = \mathbf{e}_{b_p} + \mathbf{e}_{c_p}, \tag{2.3}$$

where $a_p, b_p, c_p = 1, \dots, 5$ and $b_p < c_p$. Here \mathbf{e}_a denotes the a^{th} standard unit vector in five dimensions. Hence, $\mathbf{10}$ families have charge one under precisely one of the five U(1) symmetries in \mathcal{J} , while $\bar{\mathbf{5}}$ multiplets have charge one with respect to two of the U(1) symmetries. Apart from these rules, the precise pattern of charges across the three families is model dependent. For example, for the three $\mathbf{10}$ families, there are models with all three \mathcal{J} charges the same, two charges the same and the third one different or all three charges different. To specify explicit models it will be convenient to introduce a simple notation for the \mathcal{J} charge. We do this by adding a charge label as a subscript to the multiplet's name so that, for example $\mathbf{10}_2$ denotes a $\mathbf{10}$ multiplet with charge $\mathbf{e}_2 = (0, 1, 0, 0, 0)$ and $\bar{\mathbf{5}}_{1,4}$ denotes a $\bar{\mathbf{5}}$ multiplet with charge $\mathbf{e}_1 + \mathbf{e}_4 = (1, 0, 0, 1, 0)$.

In addition, we have one (or, in some cases, more than one) pair of Higgs doublets H, \bar{H} with \mathcal{J} charges of the type

$$\mathbf{Q}(H) = \mathbf{e}_h + \mathbf{e}_g, \quad \mathbf{Q}(\bar{H}) = -\mathbf{e}_{\bar{h}} - \mathbf{e}_{\bar{g}}, \tag{2.4}$$

where $h, g, \bar{h}, \bar{g} = 1, \dots, 5$ and $h < g, \bar{h} < \bar{g}$. As before, we attach the \mathcal{J} charge as a subscript so that, for example, a down Higgs $H_{2,3}$ has charge $\mathbf{e}_2 + \mathbf{e}_3 = (0, 1, 1, 0, 0)$ and an up-Higgs $\bar{H}_{3,5}$ has charge $-\mathbf{e}_3 - \mathbf{e}_5 = (0, 0, -1, 0, -1)$.

Finally, line bundle standard models come with standard model singlet fields, which we denote by S^α . Their number is model-dependent and, for typical examples, varies between a few and a few $\times 10$. Their \mathcal{J} charges have the form

$$\mathbf{Q}(S^\alpha) = \mathbf{e}_{d_\alpha} - \mathbf{e}_{f_\alpha}, \tag{2.5}$$

where $d_\alpha, f_\alpha = 1, \dots, 5$. Following the convention for the other fields we append this charge as a subscript so that, for example, $S_{2,5}$ has charge $\mathbf{e}_2 - \mathbf{e}_5 = (0, 1, 0, 0, -1)$ and $S_{4,1}$ has charge $\mathbf{e}_4 - \mathbf{e}_1 = (-1, 0, 0, 1, 0)$. As mentioned in the introduction, from a 10-dimensional point of view, these singlet fields can be interpreted as gauge bundle moduli. Vanishing VEVs for all singlets correspond to Abelian gauge bundles while non-vanishing VEVs indicate a deformation to non-Abelian structure groups. The singlets also play an important role from the viewpoint of the four-dimensional theory since they always carry a non-trivial \mathcal{J} charge. This means that non-vanishing singlet VEVs can spontaneously break $U(1)$ symmetries in \mathcal{J} , thereby giving mass to the vector bosons associated to non-anomalous $U(1)$ factors which have not received a mass from the Stueckelberg mechanism.

In summary, the matter spectrum of line bundle standard models is that of a generalized NMSSM, typically with a number of singlet fields rather than just a single one, and with a specific pattern of $\mathcal{J} = S(U(1)^5)$ charges, as explained above.

2.3 The moduli sector

The gravitational moduli of the models consist of the dilaton, $\mathcal{S} = s + i\sigma$, a certain number of Kahler moduli, denoted by $T^i = t^i + 2i\chi^i$, and complex structure moduli generically denoted by Z . All the moduli are singlets under the standard model group. The complex structure moduli are also singlets under the $U(1)$ symmetries in \mathcal{J} but the Kahler moduli and the dilaton have non-linear transformations, acting on their respective axionic components χ^i and σ as

$$\delta\chi^i = -k_a^i \eta^a, \quad \delta\sigma = -2k_a^i \beta_i \eta^a. \tag{2.6}$$

Here, $\mathbf{k}_a = (k_a^i)$ and $\boldsymbol{\beta} = (\beta_i)$ are numbers which are fixed for a given string construction and can be determined from the underlying topology, as will be discussed in section 4. The special unitarity of the gauge group $\mathcal{J} = S(U(1)^5)$ means that the vectors \mathbf{k}_a are subject to the constraint

$$\sum_{a=1}^5 \mathbf{k}_a = 0. \tag{2.7}$$

The 10- or 11-dimensional origin of our theories implies certain constraints on the moduli fields which are necessary for the validity of the four-dimensional effective theory. In particular, it is necessary that

$$t^i \gg 1, \quad \frac{\beta_i t^i}{s} \ll 1. \tag{2.8}$$

The first of these constraints ensures that the internal Calabi-Yau volume and the volume of cycles therein is sufficiently large for the supergravity approximation to be valid. The second constraint is necessary for the strong coupling expansion [53, 54] of the 11-dimensional theory to be valid.

In addition, the model can have moduli associated to the hidden E_8 sector and to five-branes (if present in the construction), all of which are standard model singlets. They will not play an essential role for the subsequent discussion.

2.4 The effective action

We begin by writing down the generic form for the superpotential which we split up as

$$W = W_Y + W_R + W_5 + W_{\text{sing}} + W_{\text{np}} . \tag{2.9}$$

The first four terms are perturbative while W_{np} contains the non-perturbative contributions. The standard Yukawa couplings and the μ -term are contained in W_Y , W_R consists of the R-parity violating terms and W_5 consists of the order five terms in standard model fields. The pure singlet field terms are collected in W_{sing} . Schematically, these perturbative parts can be written as

$$W_Y = \mu H \bar{H} + Y_{pq}^{(d)} H \bar{5}^p \mathbf{10}^q + Y_{pq}^{(u)} \bar{H} \mathbf{10}^p \mathbf{10}^q \tag{2.10}$$

$$W_R = \rho_p \bar{H} L^p + \lambda_{pqr} \bar{5}^p \bar{5}^q \mathbf{10}^r \tag{2.11}$$

$$W_5 = \lambda'_{pqrs} \bar{5}^p \mathbf{10}^q \mathbf{10}^r \mathbf{10}^s \tag{2.12}$$

$$W_{\text{sing}} = \tau_{\alpha\beta\gamma} S^\alpha S^\beta S^\gamma . \tag{2.13}$$

For simplicity, we have expressed the operators in terms of GUT multiplets, wherever possible. Since the U(1) charges in \mathcal{J} commute with SU(5) this will be sufficient to discuss the pattern implied by \mathcal{J} -invariance, which is our main purpose. It should, however, be kept in mind that the precise values of the allowed couplings will, in general, break SU(5). This means, for example, that the standard SU(5) GUT relation between tau and bottom Yukawa couplings may not be satisfied. All couplings above should be thought of as functions of moduli. As usual, they cannot depend on the dilaton, \mathcal{S} , and the Kahler moduli T^i thanks to their axionic shift symmetries (some of which are even gauged according to eq. (2.6)). However, they are, in general, functions of the complex structure moduli Z and the singlet fields (bundle moduli) S^α .

In this paper, for the most part, we will be interested in studying the theory for the locus in moduli space where all singlet fields are small, so $|S^\alpha| \ll 1$. From a 10-dimensional point of view this means we are considering gauge bundles with Abelian structure group or small non-Abelian deformations thereof. On this locus, all couplings above can be expanded in powers of S^α around the ‘‘Abelian locus’’ $S^\alpha = 0$. For example, for the μ -term we can write¹

$$\mu = \mu_0 + \mu_{1,\alpha} S^\alpha + \mu_{2,\alpha\beta} S^\alpha S^\beta + \dots , \tag{2.14}$$

and similarly for all other couplings. In general, the expansion coefficients μ_0 , $\mu_{1,\alpha}$, etc. should still be considered functions of the complex structure moduli. Their pattern is restricted by the \mathcal{J} charges of the standard model fields and the singlet fields S^α and it is this structure which we will mainly analyze in the following. Also note that the zeroth order μ -term, μ_0 , in eq. (2.14) vanishes even if the Higgs pair is vector-like under \mathcal{J} since all our models have an exactly massless Higgs pair at the Abelian locus $S^\alpha = 0$.

In the rest of the paper, we will not consider the non-perturbative superpotential W_{np} but a few remarks concerning its structure may be in order. Generally, one expects two

¹For models with multiple pairs of Higgs doublets the μ -term of course generalizes to a matrix of μ -terms.

types of non-perturbative effects to contribute: string instanton effects leading to terms of the form $P(Z, S^\alpha) \exp(-n_i T^i)$, and gaugino condensation leading to terms of the form $Q(Z, S^\alpha) \exp(-c(\mathcal{S} \pm \beta_i T^i))$. Here, c and n_i are positive constants (related to the beta function of the condensing gauge group and the instanton number, respectively) and P, Q are functions, typically rational, of the moduli Z and S^α . The main point is that the presence of the $U(1)$ gauge symmetries in \mathcal{J} significantly constrains the allowed non-perturbative terms, in view of the transformations (2.6) and (2.2). Specifically, the phase change of the non-perturbative exponentials due to the axion transformations (2.6) has to be cancelled by the phase change of the pre-factors P, Q due to the linear transformations of the singlet fields S^α . In ref. [55] this has been analyzed for the special case when singlet fields S^α are absent. The more general case with singlets remains to be considered in detail and this will clearly be central for the discussion of moduli stabilization and supersymmetry breaking in heterotic line bundle models.

Let us now move on to the general structure of the Kahler potential. As usual, it can be written as a sum

$$K = K_{\text{mod}} + K_{\text{mat}} \tag{2.15}$$

of the moduli superpotential K_{mod} and the matter superpotential K_{mat} . For the former, we have

$$K_{\text{mod}} = -\ln(\mathcal{S} + \bar{\mathcal{S}}) - \ln(\kappa) + K_{\text{cs}} + \dots, \tag{2.16}$$

where K_{cs} is the standard special geometry Kahler potential for complex structure moduli [56], and the dots stand for contributions from other moduli. The quantity κ is defined as

$$\kappa = d_{ijk} t^i t^j t^k, \tag{2.17}$$

with numbers d_{ijk} . From a 10-dimension viewpoint κ is proportional to the Calabi-Yau volume and d_{ijk} are the triple intersection numbers of the Calabi-Yau manifold. It is also useful to introduce the Kahler metric for the T^i moduli which follows from the above Kahler potential. It is given by

$$G_{ij} \equiv -\frac{1}{4} \frac{\partial^2}{\partial t^i \partial t^j} \ln \kappa = -\frac{3}{2} \left(\frac{\kappa_{ij}}{\kappa} - \frac{3}{2} \frac{\kappa_i \kappa_j}{\kappa^2} \right). \tag{2.18}$$

with $\kappa_i = d_{ijk} t^j t^k$ and $\kappa_{ij} = d_{ijk} t^k$.

The matter field Kahler potential has the structure

$$K_{\text{mat}} = K_{p\bar{q}}^{(10)} \mathbf{10}^p \mathbf{10}^{\bar{q}\dagger} + K_{p\bar{q}}^{(5)} \bar{\mathbf{5}}^p \bar{\mathbf{5}}^{\bar{q}\dagger} + K^{(u)} \bar{H} \bar{H}^\dagger + K^{(d)} H H^\dagger + \left(\tilde{\mu} H \bar{H} + \tilde{\rho}_p L^p \bar{H} + \hat{\rho}_p L^p H^\dagger + \text{c.c.} \right) + K_{\text{sing}}, \tag{2.19}$$

where K_{sing} is the singlet superpotential which depends on the singlets S^α and their conjugates but not on the other matter fields. The couplings in K_{mat} should be considered as functions of the moduli, more specifically of $\mathcal{S} + \bar{\mathcal{S}}, T^i + \bar{T}^i, Z, Z^\dagger, S^\alpha$ and $S^{\alpha\dagger}$. As before, for small S^α we can expand all couplings around the locus $S^\alpha = 0$, for example

$$K^{(u)} = K_0^{(u)} + \left(K_{1,\alpha}^{(u)} S^\alpha + \text{c.c.} \right) + \dots, \tag{2.20}$$

and similarly for the other couplings. The expansion coefficients are still functions of the other moduli and, as for the superpotential, they are restricted by \mathcal{J} invariance.

Some general remarks about the constraints implied by \mathcal{J} invariance are in order. Of course we know that \mathcal{J} non-invariant terms must be absent from the action. A \mathcal{J} invariant term will typically be present with a coupling which is of order one for generic values of the complex structure moduli. However, it is still possible that this coupling vanishes for specific values of the complex structure moduli. The term might even be forbidden altogether for reasons unrelated to the \mathcal{J} symmetry, for example, because of the presence of an additional discrete symmetry in the model. We can, therefore, safely draw conclusions from the absence of certain terms due to \mathcal{J} non-invariance, but we have to keep this limitation in mind when we rely on the presence of \mathcal{J} -invariant operators. In principle, we can improve on this point since many of the couplings can be explicitly computed from the underlying string theory [42]. This task is beyond the scope of the present paper and will be addressed in future publications.

The gauge kinetic function for the standard model group is universal, as is usually the case in heterotic theories, and given by

$$f = \mathcal{S} + \beta_i T^i, \tag{2.21}$$

with the topological numbers β_i identical to the ones which appear in the transformations (2.6) of the axions. In view of these, the gauge kinetic function transforms non-trivially under the U(1) symmetries in \mathcal{J} , namely

$$\delta f = -4ik_a^i \beta_i \eta^a. \tag{2.22}$$

As we will see, this non-trivial classical variation cancels the mixed $\mathcal{J}G_{\text{SM}}G_{\text{SM}}$ triangle anomaly in a four-dimensional realization of the Green-Schwarz mechanics. The gauge kinetic function for the U(1) vector fields in \mathcal{J} is given by [48]

$$f_{ab} = f\delta_{ab} + \frac{2}{3}d_{ijk}k_a^j k_b^k T^i. \tag{2.23}$$

Note that the second term represents a kinetic mixing between the U(1) symmetries. In the presence of anomalous U(1) symmetries in the hidden sector this kinetic mixing becomes more complicated and involves cross terms between hidden and observable U(1) symmetries. Since we are here focusing on the observable matter field sector we will not consider this explicitly. The general form of the gauge kinetic function, including hidden-observable mixing, can be found in ref. [48]. The variation of (2.23) leads to a cancellation of the \mathcal{J}^3 triangle anomaly.

This concludes our general set-up of heterotic line bundle models. It remains to discuss a number of generic features of these theories which are all related to the presence of the additional U(1) symmetries in \mathcal{J} .

2.5 D-terms

In this subsection we would like to discuss the D-terms associated to the U(1) gauge symmetries in \mathcal{J} . They can be computed from the linear matter fields transformations (2.2)

and the non-linear transformations (2.6) of the dilaton and the T-moduli using standard supergravity methods [57]. Explicitly they are given by

$$D_a = \frac{3k_a^i \kappa_i}{\kappa} + \frac{\beta_i k_a^i}{s} - \sum_{P, \bar{Q}} q_{Pa} K_{P\bar{Q}} C^P \bar{C}^{\bar{Q}}. \quad (2.24)$$

Here, C^P collectively denote all matter fields with \mathcal{J} charges q_{Pa} and $K_{P\bar{Q}}$ is their Kahler metric as computed from eq. (2.19). In particular, these matter fields include the singlets S^α . Since the gauge group \mathcal{J} is special unitary there are in fact only four-independent D-terms. Indeed, as a consequence of eq. (2.7) and the structure of the matter field charges the above D-terms satisfy the relation

$$\sum_{a=1}^5 D_a = 0. \quad (2.25)$$

For a supersymmetric vacuum at or near the locus $S^\alpha = 0$ we need to solve the D-term equations $D_a = 0$ along with the F-term equations which follow from the singlet superpotential W_{sing} in (2.9). In general, this requires specific knowledge of the singlet superpotential and the matter field part in the D-term (2.24). For a given model in our database both will normally be highly constrained by \mathcal{J} invariance so that this analysis can be carried out explicitly on a case-by-case basis. However, since we think of our models as being defined near $S^\alpha = 0$ we should first ensure that a supersymmetric vacuum exists at this Abelian locus. In this case, the F-term equations for S^α are automatically satisfied and the matter field contributions to the D-terms vanish. In other words, we have to ensure that the FI terms, corresponding to the first two terms in eq. (2.24), vanish. Evidently, this imposes restrictions on the dilaton and the T-moduli. To this end, let us introduce the corrected T-moduli $\tilde{t}_i = 3\kappa_i/\kappa + \beta_i/s$. Note that in view of the constraints (2.8) on the moduli, the second term in this definition is indeed a small correction. Then the D-term equations can be written as

$$D_a = k_a^i \tilde{t}_i = 0. \quad (2.26)$$

A non-trivial solution to these equations exists only if

$$(\text{number of lin. independent } \mathbf{k}_a) < (\text{number of T-moduli}). \quad (2.27)$$

Hence, for models with less than five Kahler moduli further linear dependencies, in addition to (2.7), must exist between the charge vectors \mathbf{k}_a . This implies a significant model-building constraint for models with a small number of Kahler moduli.

2.6 Green-Schwarz anomaly cancellation

The U(1) symmetries in \mathcal{J} are generically anomalous in our models. In particular, this means that the mixed $\mathcal{J}G_{\text{SM}}G_{\text{SM}}$ triangle anomalies between a \mathcal{J} gauge boson and two standard model gauge bosons as well as the cubic \mathcal{J}^3 anomaly between three \mathcal{J} gauge bosons are typically non-vanishing. The Green-Schwarz mechanism, in its four-dimensional

version, implies that these triangle anomalies are cancelled due to the non-trivial \mathcal{J} -transformations (2.21), (2.23) of the gauge-kinetic functions.

We begin, by discussing this explicitly for the mixed $\mathcal{J}G_{\text{SM}}G_{\text{SM}}$ anomalies. Using the charges (2.3) for the **10** and $\bar{\mathbf{5}}$ families the coefficients of these triangle anomalies are given by

$$\mathbf{A} = \sum_{p=1}^3 (3\mathbf{e}_{a_p} + \mathbf{e}_{b_p} + \mathbf{e}_{c_p}) . \quad (2.28)$$

For these to be cancelled by the transformation (2.22) of the gauge kinetic function we have to require that

$$(A_a - k_a^i \beta_i)_{a=1,\dots,5} \in \mathbb{Z}(1, 1, 1, 1, 1) . \quad (2.29)$$

For the models in our database these relations are automatically satisfied due to the Green-Schwarz mechanism in the underlying 10-dimensional theory. However, from a bottom-up point of view this constitutes a significant constraint, relating the charge choices for the matter fields and the moduli fields with the parameters β_i which determine the size of the threshold correction to the gauge kinetic function.

Similarly, the \mathcal{J}^3 triangle anomaly must be cancelled by the variation of the U(1) gauge kinetic functions (2.23). This leads to constraints analogous to eq. (2.29) which, however, also depend on the spectrum of singlet fields S^α . For this reason they are of less practical importance and we will not present them explicitly.

2.7 Masses of U(1) gauge bosons

The mass terms for the U(1) vector bosons arise from the kinetic terms for the axions σ and χ^i as a consequence of the non-linear transformations (2.6) and, for non-vanishing VEVs for the singlets S^α , also from the kinetic terms of those fields. At the Abelian locus, $S^\alpha = 0$, only the former contribution is present and results in a mass matrix

$$M_{ab} = \mathbf{k}_a^T \tilde{G} \mathbf{k}_b \quad \text{where} \quad \tilde{G}_{ij} = G_{ij} + \frac{\beta_i \beta_j}{4s^2} \quad (2.30)$$

is the corrected Kahler metric for the T-moduli. Since \tilde{G}_{ij} is non-degenerate this means that the number of massless U(1) vector bosons at the locus $S^\alpha = 0$ is given by

$$(\text{number of massless U(1) vector bosons}) = 4 - \text{rank}(k_a^i) . \quad (2.31)$$

Such a massless linear combination of vector bosons, characterized by a vector v^a satisfying $\mathbf{k}_a v^a = 0$, corresponds to a non-anomalous U(1) symmetry, as can be seen, in the case of the mixed anomaly, from eq. (2.29). Combining the above result with eq. (2.27) we learn that

$$(\text{number of massless U(1) vector bosons}) > 4 - (\text{number of T-moduli}) . \quad (2.32)$$

In particular, for models with less than five Kahler moduli, there necessarily exists at least one massless U(1) vector boson at the Abelian locus. On the other hand, for five or more Kahler moduli all U(1) vector bosons will be generically massive.

Non-anomalous U(1) symmetries can of course be easily broken spontaneously, thereby giving masses to the associated vector bosons, by switching on S^α VEVs. For this reason,

there is no serious phenomenological problem with the presence of additional massless U(1) symmetries at the Abelian locus and we have included such models in our database. In a detailed analysis it has of course to be checked that this spontaneous breaking is consistent with supersymmetry, that is, that it can be achieved for vanishing F- and D-terms.

3 The model database

After this general set-up we will now present the line bundle standard models from heterotic compactifications which are accessible from the database [46]. This will be done mainly from the viewpoint of the four-dimensional effective theories, following the set-up of the previous section. The underlying 10-dimensional construction will be explained in the following section. We begin by presenting one specific example model from the database. There is no implication that this particular model is phenomenologically favoured or even viable. It has merely been chosen as a useful example to explain the contents of the database and to illustrate the possible phenomenological applications of heterotic line bundle models.

In the second part of this section, we will discuss the distribution of basic phenomenological properties in our database. For example, we will count the number of models with one, two and three pairs of Higgs doublets, the number of models with vanishing dimension four and five proton-decay inducing terms and similar properties.

3.1 An example model

We will now present an example model from the database [46], namely model number 7 on the Calabi-Yau manifold with number 6732. First we discuss the gravitational sector and then move on to the matter fields and the detailed spectrum of allowed operators in the four-dimensional effective theory.

3.1.1 The gravitational sector

The database entry for the Calabi-Yau manifold underlying our example model is shown in figure 1. The data given in the figure defines a Calabi-Yau three fold X with a freely-acting symmetry group Γ . The actual Calabi-Yau manifold underlying the model is the quotient space $\hat{X} = X/\Gamma$. The details of the construction will be explained in the next section. Here we merely mention the properties which are required to extract the relevant information about the four-dimensional theory. We first note that the freely-acting symmetry for our example is $\Gamma = \mathbb{Z}_2$, so that the symmetry order is $|\Gamma| = 2$. For the number of Kahler moduli, T^i , we have²

$$\#(\text{Kahler moduli}) = h^{1,1}(\hat{X}) = h^{1,1}(X). \tag{3.1}$$

The number of complex structure moduli, Z , is then given by

$$\#(\text{complex structure moduli}) = h^{2,1}(\hat{X}) = h^{1,1}(X) - \frac{\eta(X)}{2|\Gamma|}. \tag{3.2}$$

²The number of Kahler moduli is given by the Hodge number $h^{1,1}(\hat{X})$ of the quotient manifold. It turns out that for all models in the database this number equals $h^{1,1}(X)$, although this is not true in general.

■ Cicy 6732, Symmetry 1

$$X = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$\eta(X) = -64 \quad h^{1,1}(X) = 5 \quad h^{2,1}(X) = 37 \quad c_2(TX) = (24, 24, 24, 24, 56)$$

$$\kappa = 12t_1t_2t_3 + 12t_1t_2t_4 + 12t_1t_3t_4 + 12t_2t_3t_4 + 12t_1t_2t_5 + 24t_1t_3t_5 + 24t_2t_3t_5 + 24t_1t_4t_5 + 24t_2t_4t_5 + 24t_3t_4t_5 + 12t_1t_5^2 + 12t_2t_5^2 + 24t_3t_5^2 + 24t_4t_5^2 + 8t_5^3$$

symmetry: 1 order: 2

Abelian: True block diagonal: True factors: [2]

$$\text{Action on coordinates: } \left\{ \begin{pmatrix} -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \right\}$$

$$\text{Action on polynomials: } \left\{ \begin{pmatrix} -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \right\}$$

Figure 1. Example database entry for complete intersection Calabi-Yau three-fold 6732.

From figure 1 we have $\eta(X) = -64$ and together with $h^{1,1}(X) = 5$ and $|\Gamma| = 2$ this implies that the model has 21 complex structure moduli. Another relevant quantity which can be read off from figure 1 is κ , defined in eq. (2.17), which determines the Kahler potential (2.16) for the Kahler moduli $T^i = t^i + 2i\chi^i$. For our example it is given by

$$\begin{aligned} \kappa = & 12t_1t_2t_3 + 12t_1t_2t_4 + 12t_1t_3t_4 + 12t_2t_3t_4 + 12t_1t_2t_5 + 24t_1t_3t_5 + 24t_2t_3t_5 \\ & + 24t_1t_4t_5 + 24t_2t_4t_5 + 24t_3t_4t_5 + 12t_1t_5^2 + 12t_2t_5^2 + 24t_3t_5^2 + 24t_4t_5^2 + 8t_5^3. \end{aligned} \quad (3.3)$$

The database entry for our model which defines the vector bundle is shown in figure 2. As before, we defer the details of the construction to later and focus on how to extract the relevant low-energy quantities. The charge vector \mathbf{k}_a which determine the \mathcal{J} transformations (2.6) of the axions are given by the column vectors of the matrix V in figure 2. For our example this means

$$(\mathbf{k}_1, \dots, \mathbf{k}_5) = \begin{pmatrix} 1 & 1 & 0 & 0 & -2 \\ 1 & 1 & 0 & -2 & 0 \\ 1 & -1 & -2 & 1 & 1 \\ 0 & 1 & 1 & -1 & -1 \\ -1 & -1 & 0 & 1 & 1 \end{pmatrix}. \quad (3.4)$$

■ **Model number 7, Identifier {6732, 2, 1}**

■ **Basic properties**

standard model? **True** massless U(1): **1** number of $5 \bar{5}$ pairs: **2** $c_2(V) = (12, 12, 12, 12, 32)$

$$V: (k'_a) = \begin{pmatrix} 1 & 1 & 0 & 0 & -2 \\ 1 & 1 & 0 & -2 & 0 \\ 1 & -1 & -2 & 1 & 1 \\ 0 & 1 & 1 & -1 & -1 \\ -1 & -1 & 0 & 1 & 1 \end{pmatrix}$$

Cohomology of V:

L_3	= {0, 0, -2, 1, 0}	$h[L_3]$	= {0, 2, 0, 0}	$h[L_3, R]$	= {{0, 0}, {1, 1}, {0, 0}, {0, 0}}
L_4	= {0, -2, 1, -1, 1}	$h[L_4]$	= {0, 2, 0, 0}	$h[L_4, R]$	= {{0, 0}, {1, 1}, {0, 0}, {0, 0}}
L_5	= {-2, 0, 1, -1, 1}	$h[L_5]$	= {0, 2, 0, 0}	$h[L_5, R]$	= {{0, 0}, {1, 1}, {0, 0}, {0, 0}}
$L_1 \times L_2$	= {2, 2, 0, 1, -2}	$h[L_1 \times L_2]$	= {0, 2, 0, 0}	$h[L_1 \times L_2, R]$	= {{0, 0}, {1, 1}, {0, 0}, {0, 0}}
$L_3 \times L_4$	= {0, -2, -1, 0, 1}	$h[L_3 \times L_4]$	= {0, 1, 1, 0}	$h[L_3 \times L_4, R]$	= {{0, 0}, {1, 0}, {1, 0}, {0, 0}}
$L_3 \times L_5$	= {-2, 0, -1, 0, 1}	$h[L_3 \times L_5]$	= {0, 1, 1, 0}	$h[L_3 \times L_5, R]$	= {{0, 0}, {1, 0}, {1, 0}, {0, 0}}
$L_4 \times L_5$	= {-2, -2, 2, -2, 2}	$h[L_4 \times L_5]$	= {0, 4, 0, 0}	$h[L_4 \times L_5, R]$	= {{0, 0}, {2, 2}, {0, 0}, {0, 0}}
$L_1 \times L_2^*$	= {0, 0, 2, -1, 0}	$h[L_1 \times L_2^*]$	= {0, 0, 2, 0}	$h[L_1 \times L_2^*, R]$	= {{0, 0}, {0, 0}, {1, 1}, {0, 0}}
$L_2 \times L_4^*$	= {1, 3, -2, 2, -2}	$h[L_2 \times L_4^*]$	= {0, 0, 8, 0}	$h[L_2 \times L_4^*, R]$	= {{0, 0}, {0, 0}, {4, 4}, {0, 0}}
$L_2 \times L_5^*$	= {3, 1, -2, 2, -2}	$h[L_2 \times L_5^*]$	= {0, 0, 8, 0}	$h[L_2 \times L_5^*, R]$	= {{0, 0}, {0, 0}, {4, 4}, {0, 0}}
$L_3 \times L_4^*$	= {0, 2, -3, 2, -1}	$h[L_3 \times L_4^*]$	= {0, 0, 4, 0}	$h[L_3 \times L_4^*, R]$	= {{0, 0}, {0, 0}, {2, 2}, {0, 0}}
$L_3 \times L_5^*$	= {2, 0, -3, 2, -1}	$h[L_3 \times L_5^*]$	= {0, 0, 4, 0}	$h[L_3 \times L_5^*, R]$	= {{0, 0}, {0, 0}, {2, 2}, {0, 0}}
$L_4 \times L_5^*$	= {2, -2, 0, 0, 0}	$h[L_4 \times L_5^*]$	= {0, 3, 3, 0}	$h[L_4 \times L_5^*, R]$	= {{0, 0}, {1, 2}, {1, 2}, {0, 0}}

Wilson line: **{{0}, {1}}** Equivariant structure: **{{0}, {0}, {0}, {1}, {1}}** Higgs pairs: **2**

Downstairs spectrum: **{10₃, 10₄, 10₅, $\bar{5}_{1,2}$, $2\bar{5}_{4,5}$, H_{3,4}, $\bar{H}_{3,4}$, H_{3,5}, $\bar{H}_{3,5}$, S_{2,1}, 4 S_{4,2}, 4 S_{5,2}, 2 S_{4,3}, 2 S_{5,3}, S_{4,5}, S_{5,4}}** Phys. Higgs: **{H_{3,5}, $\bar{H}_{3,4}$ }**

Transfer format: **{{8, 2, 3, 4, 7, 0, 9, 9, 18, 1, 11, 13, 15, 17, 19}, {7, 7, 8, 8, -1, -1}}**

$\text{rk}(Y^{(u)}) = (2, 2)$ $\text{rk}(Y^{(d)}) = (1, 1)$ dim. 4 operators absent: **{False, False}** dim. 5 operators absent: **{False, False}**

Figure 2. Example database entry specifying basic properties of a line bundle standard model on the three-fold in figure 1.

The transformation of the dilatonic axion, σ , in eq. (2.6) also depends on the numbers β_i which enter the gauge kinetic function (2.21). They can be computed from

$$\beta_i = \frac{1}{|\Gamma|} \left(c_{2i}(V) - \frac{1}{2} c_{2i}(TX) \right). \quad (3.5)$$

From figure 1 we read off $c_2(TX) = (24, 24, 24, 24, 56)$ and from figure (2) we have $c_2(V) = (12, 12, 12, 12, 32)$. With $|\Gamma| = 2$ this means that for our example

$$\beta = (0, 0, 0, 0, 2). \quad (3.6)$$

3.1.2 The matter field sector

All of the models in the database have a massless spectrum which includes the gauge and matter spectrum of the MSSM. However, some models have additional massless fields at the Abelian locus, $S^\alpha = 0$, which can include additional vector-like pairs of Higgs doublets and one additional massless U(1) gauge boson, related to the non-anomalous part of the \mathcal{J} gauge symmetry. Masses for these fields may be generated by non-vanishing S^α VEVs and for this reason such models have been included in the database.

Our example model has one additional massless U(1) vector field as stated at the top of figure 2. Alternatively, this follows from the general result (2.31) and the fact that only

three of the vectors \mathbf{k}_a in eq. (3.4) are linearly independent. The matter field spectrum at the Abelian locus, $S^\alpha = 0$, can be read off from the database entry entitled “Downstairs spectrum” and, from figure 2, for our example model it is given by

$$\mathbf{10}_3, \mathbf{10}_4, \mathbf{10}_5, \bar{\mathbf{5}}_{1,2}, 2\bar{\mathbf{5}}_{4,5}, H_{3,4}, \bar{H}_{3,4}, H_{3,5}, \bar{H}_{3,5}, S_{2,1}, 4S_{4,2}, 4S_{5,2}, 2S_{4,3}, 2S_{5,3}, S_{4,5}, S_{5,4}. \quad (3.7)$$

Here, we follow the notation introduced in the previous section. In particular, we have grouped the standard model particles into their standard SU(5) representations for ease of notation. We recall that the subscripts indicate the \mathcal{J} charge of a multiplet. For example, $\mathbf{10}_3$ denotes a $\mathbf{10}$ multiplet of SU(5) with \mathcal{J} charge $\mathbf{Q}(\mathbf{10}_3) = \mathbf{e}_3$, while $\bar{\mathbf{5}}_{1,2}$ denotes a $\bar{\mathbf{5}}$ multiplet of SU(5) with \mathcal{J} charge $\mathbf{Q}(\bar{\mathbf{5}}_{1,2}) = \mathbf{e}_1 + \mathbf{e}_2$. The \mathcal{J} charge for a down Higgs is, for example, $\mathbf{Q}(H_{3,4}) = \mathbf{e}_3 + \mathbf{e}_4$ while we have $\mathbf{Q}(\bar{H}_{3,4}) = -\mathbf{e}_3 - \mathbf{e}_4$ for an up Higgs. The standard model singlet fields are denoted by S and their \mathcal{J} charge pattern is exemplified by $\mathbf{Q}(S_{2,1}) = \mathbf{e}_2 - \mathbf{e}_1$.

The mixed $\mathcal{J}G_{\text{SM}}G_{\text{SM}}$ triangle anomaly can be computed from eq. (2.28). For the above spectrum we easily find

$$\mathbf{A} = (1, 1, 3, 5, 5). \quad (3.8)$$

Further, using the values of the charge vectors (3.4) and of β in eq. (3.6) it follows that

$$(k_a^i \beta_i)_a = (-2, -2, 0, 2, 2). \quad (3.9)$$

Therefore, the anomaly constraint (2.29) is indeed satisfied for our example model, as it must be due to the Green-Schwarz mechanism. This simple calculation provides a useful consistency check for our models.

The spectrum (3.7) shows that the example model contains two massless pairs of Higgs doublets at the locus $S^\alpha = 0$. In cases such as these a physical pair of Higgs doublets is chosen and separate models are generated for each possible choice. For the case at hand, the choice is $H_{3,5}, \bar{H}_{3,4}$, as the “Phys. Higgs” entry in figure 2 indicates. For consistency, the other Higgs doublet should then obtain a mass from non-zero singlet VEVs if we are to recover exactly the standard model charged spectrum with the chosen Higgs doublet. The relevant mass operators will be discussed in section 3.1.5. Another possibility is, of course, to consider phenomenological models with two or three Higgs doublets. In the following, we will discuss the various types of \mathcal{J} invariant operators allowed in the effective action and their possible phenomenological relevance. For our example model, the relevant database entry listing these operators is shown in figure 3.

3.1.3 The D-terms

The general form of the D-term has been given in eq. (2.24). The first term in this expression is the leading part of the FI term. Since we have already specified κ , for our example model given by eq. (3.3), it suffices to provide the expressions $k_a^i \kappa_i$ in order to fix this first term. In our database, these expressions are listed under the heading “FI-terms”

In order to specify the dilaton-dependent correction to the FI-term, which corresponds to the second term in eq. (2.24), we need to provide the vector $(k_a^i \beta_i)_a$. For the example model, this vector has already been determined from other database entries and is given in eq. (3.9). Finally, we need the last term in eq. (2.24) which represents the matter field contribution. Here, we write down the parts of these matter field D-terms which depend only on the singlet fields S^α . In the database it is listed under the heading “singlet D-terms”. Figure 3 shows that for our example model it is given by

$$\left(\sum_{\alpha, \beta} q_{a\alpha} S^\alpha \bar{S}^\beta \right) = \begin{pmatrix} -S_{2,1} S_{2,1}^\dagger \\ S_{2,1} S_{2,1}^\dagger - S_{4,2} S_{4,2}^\dagger - S_{5,2} S_{5,2}^\dagger \\ -S_{4,3} S_{4,3}^\dagger - S_{5,3} S_{5,3}^\dagger \\ S_{4,2} S_{4,2}^\dagger + S_{4,3} S_{4,3}^\dagger + S_{4,5} S_{4,5}^\dagger - S_{5,4} S_{5,4}^\dagger \\ -S_{4,5} S_{4,5}^\dagger + S_{5,2} S_{5,2}^\dagger + S_{5,3} S_{5,3}^\dagger + S_{5,4} S_{5,4}^\dagger \end{pmatrix}. \quad (3.11)$$

In writing these expressions, we have omitted the Kahler moduli space metric of the singlet fields which should appear but is not explicitly known. However, for the analysis of D-flat directions with $S^\alpha \neq 0$ is it usually sufficient to know that this metric is positive definite.

At the locus $S^\alpha = 0$ where all singlet VEVs vanish, the systems we consider in this paper admit a solution to the D-term equations. A necessary condition for this to be possible is that the basic constraint (2.27) holds, that is, that we have more Kahler moduli than linearly independent charge vectors \mathbf{k}_a . For our example model, which has five Kahler moduli and three linearly independent charge vectors, this is certainly satisfied. Note also that, since the overall scaling of the Kähler moduli does not enter the D-term equations, such a solution can always be scaled to the large volume regime (that is, the regime where all values of the Kahler moduli t^i are large).

For non-vanishing singlet VEVs, $S^\alpha \neq 0$, the existence of D-flat directions depends on the details of the above singlet matter field terms in the D-term and has to be analyzed case by case. Of course, for supersymmetric vacua with $S^\alpha \neq 0$ we also need to check the F-term equations which follow from the singlet superpotential W_{sing} in eq. (2.9). We now turn to a discussion of this singlet superpotential.

3.1.4 The singlet superpotential: F-terms and neutrino Majorana masses

In the database, the singlet superpotential is denoted by W_{sing} . A quick glance at figure 3 shows that for the example model it is given as

$$W_{\text{sing}} \sim S_{4,5}^2 S_{5,4}^2, \quad (3.12)$$

with possible higher dimension operators omitted. It is important to note that the singlet fields with a given \mathcal{J} charge can appear with multiplicity greater than one, as is evident from the spectrum (3.7). For simplicity, the sum over these multiplicities has been suppressed in the above expression. We also re-iterate from our general discussion in the previous section, that \mathcal{J} -invariance of singlet operators does not necessarily imply their presence in W_{sing} , as they might be forbidden for other reasons. An example of this is provided by the gauge-invariant quadratic terms in the singlets, $S_{4,5} S_{5,4}$ for our example model. We

know these terms must vanish since the underlying string construction shows that all of the singlets are indeed massless at the locus $S^\alpha = 0$ (see section 4.7 for the geometric discussion). Given the uncertainty in the exact coefficients of the terms it is not possible to give an explicit solution to the F-terms where the contribution of one operator cancels against another. However, one can argue for the existence of such a solution assuming generic coefficients. It is also possible to show that, for a given combination of singlet VEVs, the contribution of each operator to the F-terms vanishes separately. For the above example (3.12) it is clear that the (global) singlet F-terms vanish as long as the VEVs of either $S_{4,5}$ or $S_{5,4}$ are zero.

The standard model singlet fields are also attractive candidates for right-handed neutrinos (RHNs). In this context, the role of the singlet superpotential is to generate Majorana masses for the RHNs due to non-vanishing singlet VEVs. For example, for the superpotential (3.12) a non-zero VEV for $S_{5,4}$ (with the VEV of $S_{4,5}$ still vanishing to satisfy the F-term equations) generates a Majorana mass term for $S_{4,5}$ which might then play the role of a RHN. Of course a realisation of the see-saw mechanism also requires the presence of an associated Dirac mass. This will be discussed in section 3.1.8.

3.1.5 The Higgs sector

The only part of the spectrum charged under the standard model that varies within the database is the Higgs sector - some examples contain more than one set of Higgs doublets. For such models we identify a particular pair of weak doublets to play the role of the Higgs fields. This pair will then be used to calculate all of the relevant phenomenological operators such as the Yukawa couplings. The remaining doublets will be considered as exotic fields which must obtain a large mass. Another option would be to consider theories with multiple light Higgs doublet pairs. To cover all possibilities, we have generated a separate database entry for each possible choice of a Higgs doublet pair among the available doublets. The most straightforward way for the additional doublets to obtain a mass is through couplings of the form $\mu_{ij}(S)H^i\bar{H}^j$. To study these couplings, for models with multiple doublet pairs, the database contains the 3×3 mass matrix μ_{ij} for up to 3 pairs of Higgs fields. For our example model in figure 3 this matrix is given by

$$\mu_{ij} = \begin{pmatrix} \{1, S_{4,5}S_{5,4}, S_{4,5}^2S_{5,4}^2\} & \{S_{4,5}, S_{4,5}^2S_{5,4}, S_{4,5}^3S_{5,4}^2\} & \{0\} \\ \{S_{5,4}, S_{4,5}S_{5,4}^2, S_{4,5}^2S_{5,4}^3\} & \{1, S_{4,5}S_{5,4}, S_{4,5}^2S_{5,4}^2\} & \{0\} \\ \{0\} & \{0\} & \{0\} \end{pmatrix}. \quad (3.13)$$

For the case at hand one row and one column vanish because the model only contains two doublet pairs. Our notation is such that each matrix entry (i, j) lists singlet operator which can couple to $H^i\bar{H}^j$ in a \mathcal{J} -invariant way. The indices i, j run over the massless doublet pairs in the order in which they are given in the spectrum (3.7). For a given singlet VEVs it is possible to study if the additional doublets can indeed obtain a large mass while keeping the chosen Higgs pair light. The diagonal entries in the above matrix contain entries 1, consistent with the \mathcal{J} -invariance of the operator $H^i\bar{H}^i$. However, these entries should be ignored since, by construction, all doublets are indeed exactly massless at the locus $S^\alpha = 0$. This is another example of a set of operators absent for reasons unrelated to \mathcal{J} -invariance.

Keeping models with multiple Higgs pairs in the database (while we did not keep models with additional massless matter which could similarly be made massive by GUT singlets) is primarily motivated by the possibility of realising an approximate $U_{PQ}(1)$ symmetry. As will be discussed in detail in section 4.1, for the simplest cases which we have scanned over, a single Higgs pair is always vector-like under \mathcal{J} so that the \mathcal{J} -symmetry does not contain a $U(1)_{PQ}$ symmetry. However, such a symmetry may be useful for forbidding proton decay operators, for generating attractive flavour structures, and for additional control over the Higgs mass. Given two pairs of doublets, where each pair has different \mathcal{J} charges, it is possible to identify an off-diagonal combination as the Higgs fields thereby inducing a $U(1)_{PQ}$. Indeed, this has been done for the example model in figure 2 where $H_{3,5}, \bar{H}_{3,4}$ has been chosen as the physical Higgs pair. Of course the remaining doublets, $\bar{H}_{3,5}$ and $H_{3,4}$ in the example, must obtain a mass due which breaks the $U(1)_{PQ}$ symmetry. If this breaking is sufficiently controlled, for example due to a breaking scale well below the string scale, the remaining approximate $U(1)_{PQ}$ symmetry might still be useful.

For the example model, a VEV for $S_{4,5}$ gives mass to the additional doublets while keeping the physical Higgs fields massless.

3.1.6 Yukawa couplings

The database contains the Yukawa couplings under the headings $Y^{(u)}$ and $Y^{(d)}$. For the example model in figure 3 they are given by

$$\begin{aligned}
 Y^{(u)} &= \begin{pmatrix} \{S_{5,3}, S_{4,3}S_{5,4}, S_{4,5}S_{5,3}S_{5,4}\} & \{S_{5,4}, S_{4,5}S_{5,4}^2\} & \{1, S_{4,5}S_{5,4}\} \\ \{S_{5,4}, S_{4,5}S_{5,4}^2\} & \{0\} & \{0\} \\ \{1, S_{4,5}S_{5,4}\} & \{0\} & \{0\} \end{pmatrix} \\
 Y^{(d)} &= \begin{pmatrix} \{S_{5,3}, S_{4,3}S_{5,4}, S_{4,5}S_{5,3}S_{5,4}\} & \{0\} & \{0\} \\ \{S_{5,4}, S_{4,5}S_{5,4}^2\} & \{0\} & \{0\} \\ \{1, S_{4,5}S_{5,4}\} & \{0\} & \{0\} \end{pmatrix}, \quad (3.14)
 \end{aligned}$$

where we have dropped terms higher than cubic in S^α . For each model we also list the generic rank of these Yukawa matrices as in figure 2. For the above matrices we have $\text{rk}(Y^{(u)}) = \{2, 2\}$ and $\text{rk}(Y^{(d)}) = \{1, 1\}$. Here, the first (second) entry denotes the generic rank if all $S^\alpha = 0$ (if all S^α are non-zero). We note that a non-vanishing top Yukawa coupling of order one is possible in these models, even at the Abelian locus $S^\alpha = 0$.

3.1.7 Proton decay

Proton decay forms one of the classic constraints on extensions of the standard model. Dimension four proton decay operators are tightly constrained by experiments to have coefficients less than 10^{-6} for any combination of generation indices [58]. Often these are forbidden by imposing the R-parity of the MSSM. However, within the context of top-down model building from string theory, such an R-parity does not necessarily have to be realized. It is, therefore, important to consider whether such operators can be forbidden using the $U(1)$ symmetries of our models. The dimension four and five proton decay operators in (2.11) and (2.12) are denoted by λ_{pqr} and λ'_{pqrs} , respectively, and are listed under those headings in the database. For our example model they can be found in figure 3.

3.1.8 Bilinear R-parity violating operators and superpotential neutrino Dirac masses

An important set of operators which relates directly to neutrino physics are operators of the type $\rho_p \bar{H} L^p$ which appear in the superpotential (2.11). Such R-parity violating operators, among other things, lead directly to large neutrino masses through mixing with the Higgs fields and so are constrained to be very small (around 10^{-22} in Planck units). For the example model in figure 3 these coefficients vanish

$$\rho = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{3.15}$$

Therefore, in this case non-vanishing singlet VEVs cannot generate any R-parity violating operators. As ever, a bare quadratic term in the superpotential is forbidden by construction in these models.

If we take some of the standard model singlets to be RHNs then these same terms also play the role of the superpotential neutrino Dirac masses. These may then be combined with the pure singlet terms discussed in section 3.1.4 to realize the see-saw mechanism. For the example model no such superpotential Dirac masses are allowed.

3.1.9 Neutrino Kähler potential Dirac masses

In the absence of superpotential Dirac or Majorana neutrino masses there is a natural way to induce neutrino masses of the correct magnitude through a Kähler potential operator [59]. The relevant operator in the matter Kahler potential (2.19) is

$$\hat{\rho}_p L^p H^\dagger. \tag{3.16}$$

When the up-type Higgs develops a VEV, v , it induces an F-term for the down-type Higgs which, from the above operator, leads to Dirac neutrino masses. For our example model in figure 3 these couplings are given by

$$\hat{\rho} = \begin{pmatrix} \{0\} \\ \left\{ S_{4,5} S_{4,3}^\dagger, S_{4,2} S_{4,3}^\dagger S_{5,2}^\dagger, S_{4,5} S_{5,4} S_{5,3}^\dagger \right\} \\ \left\{ S_{4,5} S_{4,3}^\dagger, S_{4,2} S_{4,3}^\dagger S_{5,2}^\dagger, S_{4,5} S_{5,4} S_{5,3}^\dagger \right\} \end{pmatrix}. \tag{3.17}$$

Following the discussion in section 3.1.4 we may consider $S_{4,5}$ as a RHN. Then giving a VEV to $S_{4,3}$ induces Dirac neutrino mass. Note that we allow for conjugates of the singlets to appear since we are dealing with a Kähler potential operator.

3.1.10 The Giudice-Masiero term

A well-known way to induce a μ -term within gravity mediated supersymmetry breaking is through the Giudice-Masiero mechanism [60]. The relevant operator in the matter field Kahler potential (2.19) is

$$\tilde{\mu} H \bar{H}. \tag{3.18}$$

If $\tilde{\mu}$ depends on the conjugate, $S^{\alpha\dagger}$, of a singlet field which breaks supersymmetry, a μ -term of the right order of magnitude is generated. Hence, we should list all gauge invariant operators of the above form which involve at least one singlet appearing as a conjugate. For the example model in figure 3 this leads to the operators

$$\tilde{\mu} \supset \left\{ S_{4,5}^\dagger, S_{5,2} S_{4,2}^\dagger, S_{5,3} S_{4,3}^\dagger, \left(S_{4,5}^\dagger \right)^2 S_{5,4}^\dagger \right\}. \quad (3.19)$$

3.1.11 Kinetic terms and soft masses

One of the useful properties of the U(1) symmetries is that they allow us to gain a handle on the form of the kinetic terms of the matter fields. The kinetic terms enter the determination of the physical Yukawa couplings from the Yukawa couplings in the superpotential and are therefore of great importance. However, due to their non-holomorphic nature, they are rather difficult to calculate from first principles. For our example model in figure 3 we have

$$K^{(\bar{5})} = \begin{pmatrix} 1 & S_{2,1}^\dagger S_{4,2}^\dagger S_{5,2}^\dagger & S_{2,1}^\dagger S_{4,2}^\dagger S_{5,2}^\dagger \\ S_{2,1} S_{4,2} S_{5,2} & 1 & 1 \\ S_{2,1} S_{4,2} S_{5,2} & 1 & 1 \end{pmatrix},$$

$$K^{(10)} = \begin{pmatrix} 1 & S_{4,2} & S_{5,3} \\ S_{4,3}^\dagger & 1 & S_{4,5}^\dagger \\ S_{5,3}^\dagger & S_{4,5} & 1 \end{pmatrix}. \quad (3.20)$$

For simplicity we have only displayed the leading term for each operator while figure 3 shows the full list including terms involving up to 3 standard model singlets.

The same gauge invariant combinations are also relevant for constraining the possible soft supersymmetry breaking masses that can appear in the potential. Understanding their flavour structure is important, especially within gravity mediation, due to the possibility of inducing flavour changing neutral currents (FCNCs). It is well known that it is possible to use U(1) global symmetries to constrain the flavour off-diagonal terms of such soft masses [82] and the structure of the operators in eq. (3.20) will be conducive to realizing such a scenario.

3.2 General phenomenology overview of database

Having described the output data for an individual model it is interesting to consider how various basic phenomenological properties are distributed within the database as a whole. To this end, we present some statistics of models in the database. The given numbers are not meant as a comprehensive statistical analysis of heterotic line bundle models but merely as a rough indication of how difficult it might be to achieve certain phenomenological properties within this class.

The models presented in the database [46] descend from the 202 SU(5) GUT models constructed in ref. [35] by quotienting the Calabi-Yau three-fold and Wilson line breaking. This process breaks the GUT group to the standard model group and projects out certain unwanted states, in particular the Higgs triplets still present in the GUT theory. Depending on the symmetry by which we divide (either \mathbb{Z}_2 or $\mathbb{Z}_2 \times \mathbb{Z}_2$ in all cases) there are between

standard models	no massless U(1)	1 Higgs pair	2 Higgs pairs	3 Higgs pairs	$\text{rk}(Y^{(u)}) > 0$	proton decay operators $\lambda = \lambda' = 0$	1 Higgs, $\text{rk}(Y^{(u)}) > 0$, $\lambda = \lambda' = 0$, U(1)s massive
399	233	259	72	63	45	195	13

Table 1. Statistics of basic properties in the standard model database [46].

order 100 and 1000 choices per GUT model on how to realize this breaking. Not all of these choices lead to a phenomenologically viable spectrum (for example, in some cases Higgs triplets are still present) and, for a given model, many choices result in the same spectrum. In our scan, we have only kept the cases which lead to an acceptable spectrum and we have chosen one representative model per spectrum generated. This leads to a total of 2113 line bundle standard models which originate from the 202 GUT models. A list of these models is available as a data file at [46]. Upon inspection it turns out that many of these models are closely related in that they have the same spectrum and are based on the same (or equivalent) Calabi-Yau manifolds and the same bundle. Two models related in this way look identical for the purposes of this paper, although, since they are generally based on different symmetries of the Calabi-Yau manifold, they may differ at a more detailed level. We have eliminated these redundancies in the explicit printout of the models, in order to keep the size manageable. This results in 399 models available in the printed lists at [46]. The statistics of phenomenological properties below is based on these 399 models.

The results are summarized in table 1 below. A few comments on what precisely is being counted are in order. The number of massless U(1) vector fields and the number of Higgs pairs is determined at the Abelian locus $S^\alpha = 0$ where all singlet VEVs vanish. As discussed earlier, massless U(1) vector bosons can acquire a mass when singlet VEVs are switched on. This means that the 166 models with such a massless vector boson are not necessarily ruled out but have to be analyzed in more detail. A similar remark applies to models with more than one Higgs pair. The rank of the up Yukawa matrix $Y^{(u)}$ in column six of the table has also been determined for vanishing singlet VEVs. It can be shown that the U(1) symmetries in \mathcal{J} never allow an up Yukawa matrix with rank one and, it turns out there are no examples with $\text{rk}(Y^{(u)}) = 3$ in our list. This means all 45 models mentioned in column six have $Y^{(u)} = 2$ while all remaining models have an entirely vanishing up Yukawa matrix for vanishing singlet VEVs. A positive rank for $Y^{(u)}$ is, of course, desirable since we would like a top Yukawa coupling of order one, however, it would be preferable to have $\text{rk}(Y^{(u)}) = 1$. This can, in fact, be achieved for related constructions, to be discussed in the second part of the paper, which lead to fewer U(1) symmetries in the low-energy theory.

The second last column in the table gives the number of models for which all proton decay operators in (2.11) and (2.12) vanish, that is, $\lambda_{pqr} = 0$ and $\lambda'_{pqr} = 0$ for all values of the family indices and in the presence of generic singlet VEVs. Evidently, this is a fairly strong condition which is sufficient but not necessary to guarantee that such operators do not destabilize the proton. For example, some terms for the second and third family might be allowed, particularly if they are suppressed by small singlet VEVs. This has to be studied in detail on a case-by-case basis. At any rate, it is encouraging that we remain with 13 models even when all conditions are imposed simultaneously, as in the last column of table 1.

4 The geometry of split heterotic models

In this section, we introduce the necessary formalism to study compactifications of the $E_8 \times E_8$ heterotic theory on smooth Calabi-Yau manifolds [2] and the associated low-energy particle physics. In particular, we will discuss split bundles, that is, bundles with a direct product structure group. Bundles of this type, with the simplest splitting into a structure group $S(U(1)^5)$, underly the standard models presented in the first part of this paper. For reason which will become clear we will keep our discussion more general to cover all splittings into unitary factors. The construction of specific models based on this formalism will be presented in the next section.

4.1 General formalism

We begin by briefly reviewing the structure and constraints of generic heterotic Calabi-Yau compactifications. The geometric data required to specify a heterotic Calabi-Yau compactification which preserves four-dimensional $N = 1$ supersymmetry consist of a Calabi-Yau three-fold, X , two holomorphic, poly-stable vector bundles, V and \tilde{V} , with zero slope over X and a holomorphic curve C with second homology class $[C]$. The two vector bundles are associated to the observable and hidden E_8 sectors of the theory and their structure groups, H and \tilde{H} , must be sub-groups of E_8 . In the present paper we will take these structure groups to be $SU(n)$ (hence $c_1(V) = 0$), typically with $n = 5$ for the observable sector, or sub-groups thereof. The holomorphic curve C is wrapped by five-branes (NS five-branes in the weakly coupled limit, M five-branes in the 11-dimensional strong-coupling picture) whose other directions stretch across the four-dimensional uncompactified space-time.

This data has to satisfy a series of consistency conditions in order to obtain a well-defined vacuum. We will outline the conditions briefly here and study them in more depth in the following subsections. The first condition on the geometry is the well-known heterotic anomaly cancellation condition [2],

$$c_2(TX) - c_2(V) - c_2(\tilde{V}) = [C]. \tag{4.1}$$

In the subsequent discussion, we will focus on the observable bundle V . The hidden bundle \tilde{V} and the five-brane curve C will not be constructed explicitly but we will ensure, by an appropriate choice of V , that a consistent completion of the model exists. Usually, we will do this by requiring $c_2(TX) - c_2(V)$ to be an effective class, $[C] \in H_2(X)$. Hence, in this case we can obtain a consistent completion of the model by adding an appropriate amount of five-branes while choosing the hidden bundle \tilde{V} to be trivial.

The presence of the vector bundle V (that is, the presence of non-trivial gauge field VEVs over the Calabi-Yau three-fold, X) breaks the visible sector E_8 symmetry to a sub-group, $G \subset E_8$. The gauge group, G , is given by the commutant of H in E_8 . For example, choosing the structure group to be $H = SU(5)$ produces the commutant of $G = SU(5)$ so that we obtain a minimal GUT theory in four dimensions. If H is a proper rank four sub-group of $SU(5)$, as we will consider in this paper, the low-energy gauge group enhances to $G = SU(5) \times \mathcal{J}$, where $\mathcal{J} \cong U(1)^{f-1}$ consists of a product of $U(1)$ factors. As will be reviewed in the following sections, it is well-known, that some or all of these $U(1)$ factors

are anomalous in the Green-Schwarz sense and are, consequently, spontaneously broken at a high scale with associated massive vector bosons. The various types of low-energy multiplets in the GUT theory are obtained by decomposing the **248** adjoint representation of E_8 into representations of $H \times G$ and the number of each multiplet can be computed from the bundle-valued cohomology of V and its tensor powers.

In order to produce realistic four-dimensional models, it is necessary to further break the GUT group to the Standard Model. To this end, we will also introduce Wilson lines. However, Wilson lines can only be defined over a Calabi-Yau manifold, X , which is not simply connected (i.e. $\pi_1(X) \neq 0$). Since there are few known Calabi-Yau geometries which have a non-trivial fundamental group by construction, we shall explicitly construct such manifolds from simply connected ones, by forming quotient manifolds X/Γ where Γ is a discrete group. To this end, we require the existence of a symmetry Γ acting freely on the Calabi-Yau three-fold X , so that the quotient $\hat{X} = X/\Gamma$ is smooth and has a non-trivial first fundamental group (for example, for X/\mathbb{Z}_n , $\pi_1(X) = \mathbb{Z}_n$). In order for the bundle, V , to descend to a bundle, \hat{V} , on the quotient Calabi-Yau \hat{X} the group Γ must act consistently on the bundle. This means, there must be a group action of Γ on V which commutes with the projection $\pi : V \rightarrow X$ and satisfies a certain co-cycle condition. Such a group action is referred to as an “equivariant structure” and a bundle that admits such a structure is called “equivariant” with respect to Γ . In summary, the “downstairs” Calabi-Yau manifold \hat{X} is defined by a multi-sheeted cover $q : X \rightarrow X/\Gamma$ and all vector bundles on \hat{X} can be pulled back to equivariant bundles $V = q^*\hat{V}$ on X . That is, if V is equivariant, \hat{V} is well-defined on \hat{X} .

With the addition of non-trivial Wilson lines, the full “downstairs” bundle on \hat{X} is $\hat{V} \oplus \mathcal{W}$, where \mathcal{W} is a flat rank one bundle representing an Abelian Wilson line. Its structure group can be embedded into hypercharge $U_Y(1) \subset SU(5)$ in order to break $SU(5)$ into the standard model group. The downstairs zero-mode spectrum can be obtained from the bundle cohomology of $\hat{V} \oplus \mathcal{W}$ which, in practice, can be computed from the cohomology of the upstairs bundle V and its equivariant structure.

With this framework in hand, we turn now to the central point of this paper. We will consider vector bundles V with a split structure group of the form

$$H = S(U(n_1) \times \cdots \times U(n_f)) , \tag{4.2}$$

where $n_a \geq 1$ are integers. In order to ensure that the non-Abelian part of the low-energy gauge group is given by the GUT group $SU(5)$ we will demand that $\sum_{a=1}^f n_a = 5$. We will be especially interested in the case of “maximal splitting” when $n_a = 1$ for all a . In this case, V is simply a direct sum of five line bundles, and $SU(5)$ splits as $H = S(U(1)^5)$. However, other patterns will be of interest as well so that we keep the formalism general for now and characterize a particular pattern by the integer vector $\mathbf{n} = (n_1, \dots, n_f)$. We will now explain in detail how the general formalism for heterotic Calabi-Yau compactifications outlined above applies to such split bundles. We begin with some simple group theoretical considerations.

H	H_s	\mathcal{J}	f	\mathbf{n}
$S(U(4) \times U(1))$	$SU(4)$	$U(1)$	2	(4, 1)
$S(U(3) \times U(2))$	$SU(3) \times SU(2)$	$U(1)$	2	(3, 2)
$S(U(3) \times U(1)^2)$	$SU(3)$	$U(1)^2$	3	(3, 1, 1)
$S(U(2)^2 \times U(1))$	$SU(2)^2$	$U(1)^2$	3	(2, 2, 1)
$S(U(2) \times U(1)^3)$	$SU(2)$	$U(1)^3$	4	(2, 1, 1, 1)
$S(U(1)^5)$	1	$U(1)^4$	5	(1, 1, 1, 1, 1)

Table 2. The six splittings H of $SU(5)$ considered in this paper, together with the maximal semi-simple sub-groups $H_s \subset H$, the maximal Abelian sub-groups $\mathcal{J} \subset H$, the number f of factors and the split vector \mathbf{n} .

4.2 Group theory

In this section we lay down the necessary notation to discuss both the gauge symmetries of the four-dimensional theory, as well as the structure group of the visible sector bundle, V . We will find it convenient to introduce two sub-groups of the bundle structure group H , namely the maximal semi-simple sub-group H_s and the maximal Abelian sub-group \mathcal{J} . Explicitly, they are given by

$$H_s = SU(n_1) \times \cdots \times SU(n_f), \quad \mathcal{J} \cong \{(e^{i\eta^1}, \dots, e^{i\eta^f}) \mid \sum_a n_a \eta^a = 0\} \cong U(1)^{f-1}, \quad (4.3)$$

where η^a are the $U(1)$ group parameters and the sum condition in the definition of \mathcal{J} accounts for the fact that H consists of special unitary matrices. The different possible splittings of $SU(5)$, together with the sub-groups H_s and \mathcal{J} are listed in table 2. For the subsequent discussion it will be useful to label representations of the group H by the H_s and \mathcal{J} representations they induce. We denote by \mathcal{F}_a (Adj_a) the representation of H_s which transforms as a fundamental (adjoint) of the $SU(n_a)$ factor in H_s and as a singlet under all other factors. Representations of \mathcal{J} are specified by a charge vector $\mathbf{q} = (q_1, \dots, q_f)$. As a consequence of the constraint in the definition (4.3) of \mathcal{J} , in order to get a one-to-one correspondence between charge vectors and \mathcal{J} representations, we have to identify two such vectors \mathbf{q} and $\tilde{\mathbf{q}}$ if

$$\mathbf{q} - \tilde{\mathbf{q}} \in \mathbb{Z}\mathbf{n}. \quad (4.4)$$

Finally, an H representation which transforms under the representation R of H_s and carries \mathcal{J} charge \mathbf{q} is denoted by $R_{\mathbf{q}}$. Using this notation we can write down rules for the branching of $SU(5)$ representations into H representations. For the $SU(5)$ representations relevant to our discussion these branching rules read explicitly

$$\begin{aligned}
 \mathbf{10} &\rightarrow (\bigoplus_{a=1}^f (\wedge^2 \mathcal{F}_a)_{2\mathbf{e}_a}) \oplus (\bigoplus_{a<b} (\mathcal{F}_a \otimes \mathcal{F}_b)_{\mathbf{e}_a + \mathbf{e}_b}) & \bar{\mathbf{5}} &\rightarrow \bigoplus_{a=1}^f (\bar{\mathcal{F}}_a)_{-\mathbf{e}_a} \\
 \bar{\mathbf{10}} &\rightarrow (\bigoplus_{a=1}^f (\wedge^2 \bar{\mathcal{F}}_a)_{-2\mathbf{e}_a}) \oplus (\bigoplus_{a<b} (\bar{\mathcal{F}}_a \otimes \bar{\mathcal{F}}_b)_{-\mathbf{e}_a - \mathbf{e}_b}) & \mathbf{5} &\rightarrow \bigoplus_{a=1}^f (\mathcal{F}_a)_{\mathbf{e}_a} \\
 \mathbf{24} &\rightarrow \bigoplus_a (\text{Adj}_a)_0 \oplus \bigoplus_{a \neq b} (\mathcal{F}_a \otimes \bar{\mathcal{F}}_b)_{\mathbf{e}_a - \mathbf{e}_b}
 \end{aligned} \quad (4.5)$$

where \mathbf{e}_a denotes the a^{th} standard unit vector.

$H \cong H_s \times \mathcal{J}$ repr. $R_{\mathbf{q}}$	$SU(5) \times \mathcal{J}$ repr. $r_{\mathbf{q}}$	ass. bundle $U_{R_{\mathbf{q}}}$	$G_{\text{SM}} \times \mathcal{J}$ repr.	symbol	name
Adj_0	$\mathbf{1}_0$	$U_a \otimes U_a^*$	$(1, 1)_{0,0}$	S_a	bundle modulus
$(\mathcal{F}_a \otimes \bar{\mathcal{F}}_b)_{\mathbf{e}_a - \mathbf{e}_b}, a \neq b$	$\mathbf{1}_{\mathbf{e}_a - \mathbf{e}_b}$	$U_a \otimes U_b^*$	$(1, 1)_{0, \mathbf{e}_a - \mathbf{e}_b}$	S_{ab}	bundle modulus
$(\wedge^2 \mathcal{F}_a)_{2\mathbf{e}_a}$	$\bar{\mathbf{5}}_{2\mathbf{e}_a}$	$\wedge^2 U_a$	$(\bar{\mathbf{3}}, 1)_{2, 2\mathbf{e}_a}$ $(1, 2)_{-3, 2\mathbf{e}_a}$	d_a, T_a L_a, H_a	RH d quark/Higgs triplet LH lepton/d Higgs
$(\mathcal{F}_a \otimes \mathcal{F}_b)_{\mathbf{e}_a + \mathbf{e}_b}$ $a < b$	$\bar{\mathbf{5}}_{\mathbf{e}_a + \mathbf{e}_b}$	$U_a \otimes U_b$	$(\bar{\mathbf{3}}, 1)_{2, \mathbf{e}_a + \mathbf{e}_b}$ $(1, 2)_{-3, \mathbf{e}_a + \mathbf{e}_b}$	d_{ab}, T_{ab} L_{ab}, H_{ab}	RH d quark/Higgs triplet LH lepton/d Higgs
$(\mathcal{F}_a)_{\mathbf{e}_a}$	$\mathbf{10}_{\mathbf{e}_a}$	U_a	$(1, 1)_{6, \mathbf{e}_a}$ $(\bar{\mathbf{3}}, 1)_{-4, \mathbf{e}_a}$ $(3, 2)_{1, \mathbf{e}_a}$	e_a u_a Q_a	RH electron RH u quark LH quarks
$(\wedge^2 \bar{\mathcal{F}}_a)_{-2\mathbf{e}_a}$	$\mathbf{5}_{-2\mathbf{e}_a}$	$\wedge^2 U_a^*$	$(3, 1)_{-2, -2\mathbf{e}_a}$ $(1, 2)_{3, -2\mathbf{e}_a}$	\tilde{d}_a, \bar{T}_a \tilde{L}_a, \bar{H}_a	RH mirror d/Higgs triplet LH mirror lepton/u Higgs
$(\bar{\mathcal{F}}_a \otimes \bar{\mathcal{F}}_b)_{-\mathbf{e}_a - \mathbf{e}_b}$ $a < b$	$\mathbf{5}_{-\mathbf{e}_a - \mathbf{e}_b}$	$U_a^* \otimes U_b^*$	$(3, 1)_{-2, -\mathbf{e}_a - \mathbf{e}_b}$ $(1, 2)_{3, -\mathbf{e}_a - \mathbf{e}_b}$	$\tilde{d}_{ab}, \bar{T}_{ab}$ $\tilde{L}_{ab}, \bar{H}_{ab}$	RH mirror d/Higgs triplet LH mirror lepton/u Higgs
$(\bar{\mathcal{F}}_a)_{-\mathbf{e}_a}$	$\bar{\mathbf{10}}_{-\mathbf{e}_a}$	U_a^*	$(1, 1)_{-6, -\mathbf{e}_a}$ $(3, 1)_{4, -\mathbf{e}_a}$ $(\bar{\mathbf{3}}, 2)_{-1, -\mathbf{e}_a}$	\tilde{e}_a \tilde{u}_a \tilde{Q}_a	RH mirror electron RH mirror u quark LH mirror quark

Table 3. Representation content of the **248** adjoint of E_8 under the subgroup $H \times SU(5)$. The first column provides the representation of the bundle structure group H in terms of $H_s \times \mathcal{J}$, the second column the representation under the low-energy GUT gauge group $SU(5) \times \mathcal{J}$, with the subscripts indicating the \mathcal{J} charge. The bundle associated to each representation is given in column three. Column four provides the break-up into representations of standard model group $G_{\text{SM}} = SU_c(3) \times SU_W(2) \times U_Y(1) \times \mathcal{J}$, with the first subscripts denoting the charge $3Y$ and the second subscripts the \mathcal{J} charge.

Let us now embed H into E_8 via the embedding chain $H \subset SU(5) \subset E_8$. The commutant of the so-embedded H within E_8 , that is the low-energy gauge group, is given by $G = SU(5) \times \mathcal{J}$. As discussed before, the $U(1)$ factors in \mathcal{J} may be Green-Schwarz anomalous in which case their associated gauge bosons are massive. In order to find the multiplet types in the resulting $SU(5)$ GUT theory we need to decompose the **248** adjoint representation of E_8 . We begin with its well-known branching under $SU(5) \times SU(5)$ given by

$$\mathbf{248} \rightarrow (\mathbf{1}, \mathbf{24}) \oplus (\mathbf{24}, \mathbf{1}) \oplus (\mathbf{10}, \bar{\mathbf{5}}) \oplus (\mathbf{5}, \mathbf{10}) \oplus (\bar{\mathbf{10}}, \mathbf{5}) \oplus (\bar{\mathbf{5}}, \bar{\mathbf{10}}). \quad (4.6)$$

Here, we think of the first $SU(5)$ as the internal and the second $SU(5)$ as the external gauge group. If we replace the internal $SU(5)$ representations with the branching rules in (4.5) we immediately obtain the desired branching of **248** into representations of $H \times SU(5)$. The resulting multiplets together with other relevant information are listed in table 3.

4.3 Split bundles and stability

We would like to construct vector bundles V with the required structure group $H = S(U(n_1) \times \dots \times U(n_f))$. Starting with vector bundles U_a on the Calabi-Yau three-fold X , each with structure group $U(n_a)$, we set

$$V = \bigoplus_{a=1}^f U_a . \tag{4.7}$$

In order to ensure that the structure group is *special* unitary we also impose the vanishing of the first Chern class³

$$c_1(V) = \sum_{a=1}^f c_1(U_a) \stackrel{!}{=} 0 . \tag{4.8}$$

Relative to a basis $\{\omega_i\}$ of harmonic two-forms on X , where $i = 1, \dots, h^{1,1}(X)$, we expand the first Chern classes as $c_1(U_a) = c_1^i(U_a)\omega_i$ and, in order to make contact with the four-dimensional discussion in the previous section, introduce the vectors \mathbf{k}_a by setting

$$k_a^i = c_1^i(U_a) . \tag{4.9}$$

Now we need to discuss the conditions on such bundles V which follow from the requirement of preserving four-dimensional $N = 1$ supersymmetry. As discussed above, in order for this bundle to be supersymmetric, it needs to be poly-stable with zero slope.

To understand these conditions, we must define the notions of slope, stability and poly-stability which we introduce in turn. The slope $\mu(\mathcal{G})$ of a coherent sheaf, \mathcal{G} , on the Calabi-Yau three-fold X is defined by

$$\mu(\mathcal{G}) = \frac{1}{\text{rk}(\mathcal{G})} \int_X c_1(\mathcal{G}) \wedge J \wedge J = \frac{1}{\text{rk}(\mathcal{G})} d_{ijk} c_1^i(\mathcal{G}) t^j t^k , \tag{4.10}$$

where J is the Kahler form of X . For the second equality we have expanded $J = t^i \omega_i$ and $c_1(\mathcal{G}) = c_1^i(\mathcal{G}) \omega_i$ and introduced the triple intersection numbers $d_{ijk} = \int_X \omega_i \wedge \omega_j \wedge \omega_k$ of X . A holomorphic vector bundle V is now called (slope-) *stable* if

$$\mu(\mathcal{G}) < \mu(V) \quad \text{for all coherent sub-sheaves } \mathcal{G} \subset V \quad \text{with } 0 < \text{rk}(\mathcal{G}) < \text{rk}(V) . \tag{4.11}$$

Note that due to the restriction on the rank in this definition line bundles are always stable. Further, V is called poly-stable if

$$V = \bigoplus_a V_a \quad \text{such that } V_a \text{ stable and } \mu(V_a) = \mu(V) \quad \forall a . \tag{4.12}$$

Hence, a poly-stable bundle consists of a direct sum of stable bundles, each with the same slope. Since supersymmetry also requires that $\mu(V) = 0$ (which is automatic in our case

³If there are additional conditions between the first Chern classes of the U_a the structure group might reduce further and become a proper sub-group of one of the structure groups given in table 2. In this case, the non-Abelian part of the low-energy gauge group might be larger than $SU(5)$. We will not consider this case explicitly in our general set-up and avoid models of this type in our discussion of examples later on.

since we consider bundles V with $c_1(V) = 0$) this means that the slope of all constituent bundles V_a must vanish. A poly-stable bundle V with zero slope has no global sections since the trivial line bundle \mathcal{O}_X has slope zero and is, hence, a potentially de-stabilising sub-bundle. Therefore, \mathcal{O}_X cannot inject into V and we must have $H^0(X, V) \cong \text{Hom}(\mathcal{O}_X, V) = 0$. If V is poly-stable, its dual V^* , is also poly-stable with zero slope, so that on a Calabi-Yau manifold $H^3(X, V) \cong H^0(X, V^*) = 0$. In conclusion, poly-stable bundles with slope zero on a Calabi-Yau manifold have vanishing zeroth and third cohomology.⁴

Let us now apply these general statements to the bundle V in (4.7). For V to be supersymmetric it needs to be poly-stable with zero slope which is equivalent to saying that all U_a must be stable with zero slope, $\mu(U_a) = 0$. Stability is automatic if U_a is a line bundle but has to be checked explicitly for higher-rank bundles (this can be carried out explicitly following the procedure outlined in ref. [50, 73]). For the full bundle V to be poly-stable, the stable loci for the various U_a must have a non-trivial intersection in the Kahler cone. On this intersection V is poly-stable. The second condition for supersymmetry, the vanishing of the slope for each U_a , can from eq. (4.10) be expressed as

$$\mu(U_a) \sim k_a^i \kappa_i = 0 \quad \forall a, \tag{4.13}$$

where $\kappa_i = d_{ijk} t^j t^k$ are “dual” Kahler moduli space coordinates. Hence, the vanishing slope conditions lead to additional constraints on the Kahler moduli space of X which have to be combined with the ones following from stability. As can be seen from eq. (2.24), in the four-dimensional effective theory all these constraints are enforced via D-terms associated to the anomalous U(1) symmetries in \mathcal{J} and additional anomalous U(1) symmetries which may appear at particular loci in Kahler moduli space when one or more of the bundles U_a split up further [63]. The bundle is supersymmetric only in the part of the Kahler moduli space where all of these conditions are satisfied simultaneously. As is clear from eq. (4.13), in order for a common solution to the zero slope conditions to exist it is necessary that

$$(\text{number of lin. independent } \mathbf{k}_a) < h^{1,1}(X). \tag{4.14}$$

In the context of the four-dimensional discussion we have seen the same condition (2.27), for the case of purely Abelian splittings, appear from the D-term equations.

For Calabi-Yau three-folds with a small Hodge number $h^{1,1}(X)$, the slope zero conditions (4.13) are an important model building constraint on the bundle, V . For example, for $h^{1,1}(X) = 1$ there are no solutions at all, while for $h^{1,1}(X) = 2$ all first Chern classes \mathbf{k}_a must be multiples of each other.

⁴If a line bundle appears in the sum (4.12) the above argument breaks down. However, in this case, one can still conclude that a poly-stable, zero slope bundle V satisfies $H^0(X, V) = H^3(X, V) = 0$ by invoking the vanishing Theorem (1.24) in ref. [61]. It states that a line bundle L has no global sections if $\mu(L) < 0$ (and, on a Calabi-Yau manifold, it has vanishing third cohomology if $\mu(L) > 0$) somewhere in the Kahler cone of X . Since we require that $\mu(L) = 0$, all line bundles except the trivial one have points in the Kahler cone where $\mu(L) > 0$ and $\mu(L) < 0$, so that the theorem applies and $H^0(X, L) = H^3(X, L) = 0$. The one exception is the trivial bundle \mathcal{O}_X which has vanishing slope everywhere in the Kahler cone and satisfied $H^0(X, \mathcal{O}_X) = H^3(X, \mathcal{O}_X) = 1$. However, we are not interested in cases for which \mathcal{O}_X appears in the direct sum in (4.7) since this leads to the case of enhanced symmetry described in footnote 3. Hence, for our considerations, all line bundles L will indeed have vanishing zeroth and third cohomology.

As we have seen, stability and vanishing slope of each U_a implies that $H^0(X, U_a) = H^3(X, U_a) = 0$. This means the chiral asymmetries associated to the bundles U_a (that is the chiral asymmetry of the $\mathbf{10}$ and $\overline{\mathbf{10}}$ multiplets with charges $\pm \mathbf{e}_a$) can be computed from the index, so that $\text{ind}(U_a) = -h^1(X, U_a) + h^2(X, U_a)$.

An analogous argument holds for $\wedge^2 V$ and its index. To see this, note that if V is a poly-stable bundle with slope zero then it follows that $\wedge^2 V$ (and $V \otimes V^*$) are also poly-stable with slope zero [74]. As a result, each indecomposable term, $\wedge^2 U_a, U_a \otimes U_b$ in $\wedge^2 V$ is a properly stable bundle with slope zero. Following the same line of argument once again, such a term either has vanishing zeroth and third cohomology, or consists of a trivial bundle in which case its zeroth and third cohomologies are equal. Either way we have that $\text{ind}(\wedge^2 U_a) = -h^1(X, \wedge^2 U_a) + h^2(X, \wedge^2 U_a)$ (and similarly for $U_a \otimes U_b$), so that the index counts the chiral asymmetry of $\overline{\mathbf{5}}$ and $\mathbf{5}$ multiplets with charges $\pm(\mathbf{e}_a + \mathbf{e}_b)$.

4.4 Spectrum of GUT theory

As discussed above, the four-dimensional gauge group is $\text{SU}(5) \times \mathcal{J}$, where $\mathcal{J} \simeq \text{U}(1)^{f-1}$ consists of $\text{U}(1)$ factors which are normally Green-Schwarz anomalous. From table 3, the multiplets $r_{\mathbf{q}}$ under this gauge group are given by

$$\mathbf{1}_{\mathbf{e}_a - \mathbf{e}_b}, \overline{\mathbf{5}}_{\mathbf{e}_a + \mathbf{e}_b}, \mathbf{10}_{\mathbf{e}_a}, \mathbf{5}_{-\mathbf{e}_a - \mathbf{e}_b}, \overline{\mathbf{10}}_{-\mathbf{e}_a} \quad (4.15)$$

where we recall that the sub-script indicates the \mathcal{J} charge, an integer vector subject to the identification (4.4). The corresponding internal H representation, $R_{\mathbf{q}}$, for each of these multiplets is listed in the first column of table 3. Given that the bundles U_a are associated to the H -representations $(\mathcal{F}_a)_{\mathbf{e}_a}$, the associated bundles for each of the multiplets $r_{\mathbf{q}}$ in (4.15) are easily worked out by first identifying the corresponding H representation $R_{\mathbf{q}}$ and then taking appropriate tensor products of $(\mathcal{F}_a)_{\mathbf{e}_a}$. The result for the associated bundles $U_{R_{\mathbf{q}}}$ is listed in the third column of table 3. The number, $n(r_{\mathbf{q}})$ of multiplets with $\text{SU}(5) \times \mathcal{J}$ representation $r_{\mathbf{q}}$ is then given by the first cohomology, $H^1(X, U_{R_{\mathbf{q}}})$, of these associated bundle. Serre duality on a Calabi-Yau manifold implies that $H^1(X, U) \cong H^2(X, U^*)$ for any bundle U . Hence, the chiral asymmetries $N(r_{\mathbf{q}}) = n(r_{\mathbf{q}}) - n(\overline{r}_{-\mathbf{q}})$ can be expressed in terms in terms of the topological index of each respective bundle as

$$\begin{aligned} N(\mathbf{1}_{\mathbf{e}_a - \mathbf{e}_b}) &= n(\mathbf{1}_{\mathbf{e}_a - \mathbf{e}_b}) - n(\mathbf{1}_{\mathbf{e}_b - \mathbf{e}_a}) = -\text{ind}(U_a \otimes U_b^*) \\ &= -\text{rk}(U_b)\text{ind}(U_a) + \text{rk}(U_a)\text{ind}(U_b) - \text{ch}_1(U_a)\text{ch}_2(U_b) + \text{ch}_2(U_a)\text{ch}_1(U_b) \end{aligned} \quad (4.16)$$

$$N(\overline{\mathbf{5}}_{2\mathbf{e}_a}) = n(\overline{\mathbf{5}}_{2\mathbf{e}_a}) - n(\mathbf{5}_{-2\mathbf{e}_a}) = -\text{ind}(\wedge^2 U_a) \quad (4.17)$$

$$= (4 - \text{rk}(U_a))\text{ind}(U_a) - \text{ch}_1(U_a) \left(\text{ch}_2(U_a) + \frac{1}{4}c_2(TX) \right) \quad (4.18)$$

$$\begin{aligned} N(\overline{\mathbf{5}}_{\mathbf{e}_a + \mathbf{e}_b}) &= n(\overline{\mathbf{5}}_{\mathbf{e}_a + \mathbf{e}_b}) + n(\mathbf{5}_{-\mathbf{e}_a - \mathbf{e}_b}) = -\text{ind}(U_a \otimes U_b) \\ &= -(\text{rk}(U_b)\text{ind}(U_a) + \text{rk}(U_a)\text{ind}(U_b) + \text{ch}_1(U_a)\text{ch}_2(U_b) + \text{ch}_2(U_a)\text{ch}_1(U_b)) \end{aligned} \quad (4.19)$$

$$N(\mathbf{10}_{\mathbf{e}_a}) = n(\mathbf{10}_{\mathbf{e}_a}) - n(\overline{\mathbf{10}}_{-\mathbf{e}_a}) = -\text{ind}(U_a) . \quad (4.20)$$

The total chiral asymmetry of $\overline{\mathbf{5}}$ and $\mathbf{10}$ multiplets, summed over all \mathcal{J} charges, is given by the usual formula

$$N(\mathbf{10}) = -\text{ind}(V) = -\text{ind}(\wedge^2 V) = N(\overline{\mathbf{5}}) \quad (4.21)$$

for poly-stable rank five bundles with $c_1(V) = 0$.

4.5 Discrete symmetries, equivariance and downstairs spectrum

As in generic heterotic compactifications, in order to break the visible GUT symmetry to $G_{\text{SM}} = \text{SU}(3) \times \text{SU}(2) \times U_Y(1)$, we must introduce Wilson lines, on a non-simply connected space. To this end, we will quotient the Calabi-Yau three-fold and bundle by a freely-acting discrete symmetry Γ . A bundle U over X descends to a bundle \hat{U} on the quotient $\hat{X} = X/\Gamma$ if and only if it is equivariant, that is, if the symmetry Γ can be “lifted” to the bundle. The mathematical definitions and the details on how to construct equivariant structures are described in appendix A. Here we will merely need a few facts about such bundles. First, the indices of an equivariant bundle U and its downstairs counterpart \hat{U} are related by

$$\text{ind}(\hat{U}) = \text{ind}(U)/|\Gamma|. \tag{4.22}$$

Also the cohomologies, $H^q(X, U)$, of an equivariant bundle U form representations under the group Γ . It is therefore useful to define the “graded” cohomologies $H^q(X, U, R)$ which are the subspaces of $H^q(X, U)$ which transform under the Γ representations R . Further, $h^q(X, U, R) = \dim(H^q(X, U, R))$ is the corresponding graded dimension and the graded index is defined by

$$\text{ind}(U, R) = \sum_q (-1)^q h^q(X, U, R). \tag{4.23}$$

As explained earlier, we focus on models where each constituent bundle U_a has an equivariant structure on its own. The obvious ambiguity in choosing such equivariant structures consists of an overall phase for each U_a which can be encoded in Γ -representations or characters, denoted by χ_a^* . We should now discuss how to obtain the downstairs spectrum. Eq. (4.22) applied to each U_a means that the downstairs spectrum consists of a certain number of complete $\bar{\mathbf{5}}$ and $\mathbf{10}$ representations plus vector-like multiplets. In particular, each downstairs $\bar{\mathbf{5}}$ ($\mathbf{10}$) multiplet descends as a whole from a specific sector of \mathcal{J} charge. This means that the constituents of each such multiplet carry the same charge under the anomalous $U(1)$ symmetries in \mathcal{J} .

In order to find the vector-like spectrum downstairs one has to study the equivariant structure in more detail. To do so, we must choose a Wilson line \mathcal{W} (that is, a flat rank one bundle on \hat{X} , embedded into $U_Y(1)$ in order to break the GUT symmetry to the standard model) which induces a representation of the discrete group Γ contained in $U_Y(1)$. Each standard model multiplet, ψ , in table 3 is, therefore, associated to an Wilson line bundle \mathcal{W}_ψ and carries a Γ representation R_ψ which is determined by its weak hypercharge. If we focus on a particular such multiplet, ψ , contained within the GUT multiplet associated to U then we have the following relation between cohomologies.

$$H^1(\hat{X}, \hat{U} \oplus \mathcal{W}_\psi) \cong (H^1(X, U) \otimes R_\psi)_{\text{inv}} \tag{4.24}$$

That is, the downstairs spectrum in the presence of the Wilson line can be computed from the upstairs cohomology by tensoring with the various Γ representations of the standard model multiplets and extracting the Γ invariant part from this tensor product.

Let us now explain this procedure in more detail, focusing on our models with split structure group. Here we will only discuss Abelian discrete symmetries

$$\Gamma = \bigotimes_r \mathbb{Z}_{m_r}, \tag{4.25}$$

typically with either a single or with two \mathbb{Z}_m factors. We define $\alpha_r = \exp(2\pi i/m_r)$ and the \mathbb{Z}_{m_r} representations $R_{p_r}^{(m_r)}(g) = \alpha_r^{p_r g}$, where $g \in \{0, \dots, m_r - 1\}$. Then, an embedding of \mathbb{Z}_{m_i} into $U_Y(1) \subset \text{SU}(5)$ can be written as

$$g \rightarrow \text{diag}(\alpha_r^{p_r g}, \alpha_r^{p_r g}, \alpha_r^{p_r g}, \alpha_r^{\tilde{p}_r g}, \alpha_r^{\tilde{p}_r g}), \tag{4.26}$$

where p_r, \tilde{p}_r are integers satisfying $3p_r + 2\tilde{p}_r = 0 \pmod{m_r}$. Further, to actually break $\text{SU}(5)$ to G_{SM} we need that $p_r \neq \tilde{p}_r$ for at least one r . So in summary, the viable $\text{SU}(5)$ breaking Wilson lines can be obtained by solving

$$3p_r + 2\tilde{p}_r = 0 \pmod{m_r} \text{ for all } r, \quad p_i \neq \tilde{p}_r \text{ for at least one } r, \tag{4.27}$$

where $p_r, \tilde{p}_r \in \{0, \dots, m_r - 1\}$. In particular, this means that a single \mathbb{Z}_m can break $\text{SU}(5)$ to the standard model provided that $n \neq 5$. If Γ is a direct product of \mathbb{Z}_m factors is it sufficient that one of the factors is different from \mathbb{Z}_5 . For a solution of the above equations, we define the two Γ representations $W = \bigotimes_r R_{p_r}^{(m_r)}$ and $\tilde{W} = \bigotimes_r R_{\tilde{p}_r}^{(m_r)}$. The relevant $\text{SU}(5)$ representations decompose under $\Gamma \times G_{\text{SM}}$ as

$$\mathbf{5} \rightarrow (\tilde{W}, \mathbf{1}, \mathbf{2})_3 \oplus (W, \mathbf{3}, \mathbf{1})_{-2} \tag{4.28}$$

$$\bar{\mathbf{5}} \rightarrow (\tilde{W}^*, \mathbf{1}, \mathbf{2})_{-3} \oplus (W^*, \bar{\mathbf{3}}, \mathbf{1})_2 \tag{4.29}$$

$$\mathbf{10} \rightarrow (\tilde{W} \otimes \tilde{W}, \mathbf{1}, \mathbf{1})_6 \oplus (W \otimes W, \bar{\mathbf{3}}, \mathbf{1})_{-4} \oplus (W \otimes \tilde{W}, \mathbf{3}, \mathbf{2})_1 \tag{4.30}$$

$$\bar{\mathbf{10}} \rightarrow (\tilde{W}^* \otimes \tilde{W}^*, \mathbf{1}, \mathbf{1})_{-6} \oplus (W^* \otimes W^*, \mathbf{3}, \mathbf{1})_4 \oplus (W^* \otimes \tilde{W}^*, \bar{\mathbf{3}}, \mathbf{2})_{-1}. \tag{4.31}$$

Then, the downstairs cohomology can be expressed in terms of graded cohomologies with respect to the Γ representations W, \tilde{W} , which describe the Wilson line and the characters χ_a , which encode the freedom in choosing the equivariant structure. The result is summarized in table 4.

For the ‘‘physics’’ models we consider in this paper not all of the graded cohomologies in table 4 need to be computed explicitly. The upstairs spectrum of such models consists of $3|\Gamma|$ $\mathbf{10}$ multiplets and has no $\bar{\mathbf{10}}$ multiplets. By virtue of eq. (4.22) this will guarantee precisely three $\mathbf{10}$ multiplets downstairs so there is no need to check graded cohomologies in this sector. Further, from eq. (4.21) we know that the chiral asymmetry of $\bar{\mathbf{5}}$ and $\mathbf{5}$ multiplets is also $3|\Gamma|$, so that we are guaranteed three chiral $\bar{\mathbf{5}}$ multiplets downstairs. In addition, we have to check that all Higgs triplets can be projected out and at least one pair of Higgs doublets remains in the spectrum. This can be done by computing the number of $\bar{T}_{a,b}$ and $\bar{H}_{a,b}$ multiplets from the associated cohomologies in table (4). The details of how to compute graded cohomologies for line bundles on the particular Calabi-Yau manifolds used in our constructions are explained in appendix B.

SU(5) repr.	G_{SM} repr.	name	cohomology
$\mathbf{10}_{\mathbf{e}_a}$	$(\mathbf{3}, \mathbf{2})_1$	Q_a	$h^1(X, U_a, \chi_a \otimes W^* \otimes \tilde{W}^*)$
	$(\bar{\mathbf{3}}, \mathbf{1})_{-4}$	u_a	$h^1(X, U_a, \chi_a \otimes W^* \otimes W^*)$
	$(\mathbf{1}, \mathbf{1})_6$	e_a	$h^1(X, U_a, \chi_a \otimes \tilde{W}^* \otimes \tilde{W}^*)$
$\bar{\mathbf{5}}_{\mathbf{e}_a+\mathbf{e}_b}$	$(\bar{\mathbf{3}}, \mathbf{1})_2$	$d_{a,b}, T_{a,b}$	$h^1(U_a \otimes U_b, \chi_a \otimes \chi_b \otimes W)$
	$(\mathbf{1}, \mathbf{2})_{-3}$	$L_{a,b}, H_{a,b}$	$h^1(U_a \otimes U_b, \chi_a \otimes \chi_b \otimes \tilde{W})$
$\bar{\mathbf{5}}_{2\mathbf{e}_a}$	$(\bar{\mathbf{3}}, \mathbf{1})_2$	d_a, T_a	$h^1(\wedge^2 U_a, \chi_a \otimes \chi_a \otimes W)$
	$(\mathbf{1}, \mathbf{2})_{-3}$	L_a, H_a	$h^1(\wedge^2 U_a, \chi_a \otimes \chi_a \otimes \tilde{W})$
$\mathbf{5}_{-\mathbf{e}_a-\mathbf{e}_b}$	$(\mathbf{3}, \mathbf{1})_{-2}$	$\bar{T}_{a,b}$	$h^2(U_a \otimes U_b, \chi_a \otimes \chi_b \otimes W)$
	$(\mathbf{1}, \mathbf{2})_3$	$\bar{H}_{a,b}$	$h^2(U_a \otimes U_b, \chi_a \otimes \chi_b \otimes \tilde{W})$
$\mathbf{5}_{-2\mathbf{e}_a}$	$(\mathbf{3}, \mathbf{1})_{-2}$	\bar{T}_a	$h^2(\wedge^2 U_a, \chi_a \otimes \chi_a \otimes W)$
	$(\mathbf{1}, \mathbf{2})_3$	\bar{H}_a	$h^2(\wedge^2 U_a, \chi_a \otimes \chi_a \otimes \tilde{W})$
$\mathbf{1}_{\mathbf{e}_a-\mathbf{e}_b}$	$(\mathbf{1}, \mathbf{1})_0$	$S_{a,b}$	$h^1(U_a \otimes U_b^*, \chi_a \otimes \chi_b^*)$

Table 4. Cohomologies which compute the downstairs spectrum. The Wilson line is characterized by the two representations W and \tilde{W} of the freely-acting Abelian symmetry Γ and χ_a^* are the characters of U_a . The number of mirror particles is obtained by the second cohomology of the same bundle and the same representation.

4.6 Anomalies

A characteristic feature of our split models is the presence of the Green-Schwarz anomalous U(1) symmetries in \mathcal{J} . For the case of line bundle sums and from a four-dimensional perspective this has already been discussed in section 2.6. Here, we will provide a general discussion, valid for all splitting types and from a geometric viewpoint.

We begin by computing the anomaly coefficients in the GUT model, focusing on the mixed $\mathcal{J} \text{SU}(5)^2$ contribution. With the group theoretical indices $c(\bar{\mathbf{5}}) = 1$ and $c(\mathbf{10}) = 3$ for the relevant SU(5) representations the $\mathcal{J} \text{SU}(5)^2$ triangle anomaly is proportional to

$$\mathbf{A} = \sum_{a=1}^f (2N(\bar{\mathbf{5}}_{2\mathbf{e}_a}) + 3N(\mathbf{10}_{\mathbf{e}_a})) \mathbf{e}_a + \sum_{a < b} N(\bar{\mathbf{5}}_{\mathbf{e}_a+\mathbf{e}_b}) (\mathbf{e}_a + \mathbf{e}_b). \quad (4.32)$$

Here, we recall that $N(r_{\mathbf{q}}) = n(r_{\mathbf{q}}) - n(\bar{r}_{-\mathbf{q}})$ is the chiral asymmetry of a certain representation $r_{\mathbf{q}}$. The above anomaly coefficient is an f -dimensional vector which is defined only up to the identification (4.4) of \mathcal{J} charges, so we should explain how to extract unambiguous anomaly coefficients from this result. Any particular U(1) symmetry within \mathcal{J} can be represented by a vector $\mathbf{Q} = (Q_1, \dots, Q_f)$ satisfying $\mathbf{n} \cdot \mathbf{Q} = 0$. The anomaly coefficient for this U(1) symmetry is then given by $\mathbf{Q} \cdot \mathbf{A}$ and is, hence, independent of the identification (4.4).

We can use the results (4.16)–(4.20) for the chiral asymmetries to rewrite the above expression for \mathbf{A} in terms of topological data. This leads to

$$(A_a - \beta_i k_a^i)_{a=1, \dots, f} \in \mathbb{Z}\mathbf{n}, \quad \beta_i = c_{2i}(V) - \frac{1}{2} c_{2i}(TX), \quad (4.33)$$

where β is the coefficient which appears in the one-loop correction to the four-dimensional gauge-kinetic function (2.21). We have phrased the above results for the upstairs Calabi-Yau X and the upstairs bundle V . However, analogous equations, with the chiral asymmetries in (4.32) interpreted as the downstairs chiral asymmetries and the replacements $U_a \rightarrow \hat{U}_a$, $V \rightarrow \hat{V}$ and $X \rightarrow \hat{X}$ in (4.33) hold on the quotient Calabi-Yau \hat{X} , provided all bundles U_a are equivariant individually, as we are assuming here. The subsequent inclusion of a Wilson line does, of course, not affect the chiral asymmetries and, hence, leaves the anomaly coefficients unchanged.

It is important to also consider the masses of the $f - 1$ vector fields associated to the potentially anomalous $U(1)$ symmetries in \mathcal{J} . They are given by eq. (2.30), but with the more general definition (4.9) of the first Chern classes. The mass matrix (2.30) has size f while we only have $f - 1$ vector fields. As usual, this mismatch is corrected by imposing the constraint $\mathbf{n} \cdot \mathbf{v} = 0$ on f dimensional vectors $\mathbf{v} = (v^a)$ representing gauge bosons in this basis. Hence, following the discussion in section 2.7, massless vector $U(1)$ bosons are characterized by the equation

$$\sum_{a=1}^f \mathbf{k}_a v^a = 0 \text{ for } \mathbf{n} \cdot \mathbf{v} = 0. \tag{4.34}$$

Comparison with eq. (4.33) shows that a $U(1)$ symmetry with a massless vector boson is necessarily anomaly-free. Conversely, while a $U(1)$ symmetry with a massive vector boson can generically be expected to be anomalous, it may be anomaly-free in special cases. This happens if $k_a^i v^a \beta_i = 0$ while $k_a^i v^a \neq 0$. This means the number of massless $U(1)$ vector bosons is given by

$$(\text{number of massless } U(1) \text{ vector bosons}) = f - 1 - \text{rank}(k_a^i), \tag{4.35}$$

and combining this result with (4.14) we learn that

$$(\text{number of massless } U(1) \text{ vector bosons}) \geq f - h^{1,1}(X). \tag{4.36}$$

This means that for Calabi-Yau manifolds with small $h^{1,1}(X) < f$ we necessarily have at least one massless $U(1)$ vector boson. However it is frequently still possible to spontaneously break these symmetries by giving a vacuum expectation value to appropriate singlet fields. The phenomenological implications of such choices were explored in section 3. In the next section, we will explore the geometric meaning of such singlet VEVs.

4.7 Connecting different splitting types

Split bundles of the type considered in this paper typically arise at a special locus in Kahler and bundle moduli space and, provided certain fields exist, one can move away from this special locus to obtain a less-split bundle or even recover a full $SU(5)$ structure group. This process corresponds to moving upwards in the list of structure groups given in table 2. Conversely, setting certain moduli to zero one can go back to the special locus and obtain a finer splitting of the structure group. Such movement in moduli space has

been studied in detail in recent literature [50–52, 62–64] and here we will simply review some of the relevant facts as they apply to the models under investigation.

The moduli of the split bundle, V , in eq. (4.7) are described by

$$H^1(X, V \otimes V^*) = \bigoplus_{a,b} H^1(X, U_a \otimes U_b^*) . \tag{4.37}$$

The $a = b$ terms on the right-hand side clearly correspond to moduli of the bundle U_a . It is suggestive (and in fact correct) to relate the terms for $a \neq b$ to deformations away from the split bundles $U_a \oplus U_b$. To see this in more detail, let us focus on two constituent bundles U_a, U_b with structure groups $U(n_a), U(n_b)$. Possible bundles U which correspond to moving away from the split bundle $U_a \oplus U_b$ can be described by the following two extension sequences

$$0 \rightarrow U_a \rightarrow U \rightarrow U_b \rightarrow 0 . \tag{4.38}$$

$$0 \rightarrow U_b \rightarrow \tilde{U} \rightarrow U_a \rightarrow 0 \tag{4.39}$$

Any infinitesimal smooth, slope-stable deformation of $U_a \oplus U_b$ can be described by a combination of the bundles U and \tilde{U} above [65]. For example, the space of inequivalent bundles U (respectively \tilde{U}) described by this sequence is given by $\text{Ext}^1(U_b, U_a) \cong H^1(X, U_a \otimes U_b^*)$ (respectively $\text{Ext}^1(U_a, U_b) \cong H^1(X, U_b \otimes U_a^*)$), so precisely the terms for $a \neq b$ in the sum (4.37). Comparing with table 3 we see that these cohomology groups are associated to the $SU(5)$ singlet fields S_{ab} which carry charge $\mathbf{e}_a - \mathbf{e}_b$ under the anomalous $U(1)$ symmetries in \mathcal{J} . Hence, once we have computed the complete spectrum of a given split bundle V we can immediately decide which deformations are possible.

If $H^1(X, U_a \otimes U_b^*) = H^1(X, U_b \otimes U_a^*) = 0$ then no relevant singlet fields exist and the split bundle $U = U_a \oplus U_b$ is the only bundle possible in the extension sequences (4.38), (4.39). On the other hand, if at least one of $H^1(X, U_a \otimes U_b^*)$ and $H^1(X, U_b \otimes U_a^*)$ is non-zero, non-trivial extensions U whose structure group is “larger” than $U(n_a) \times U(n_b)$ and is contained in $U(n_a + n_b)$ exist.

In the four-dimensional effective theory these infinitesimal deformations are described by the D-terms associated to the anomalous $U(1)$ symmetries. For the case of line bundle sums this has been discussed in section 2.5. For general splittings and after setting all non-singlet fields to zero, the schematic structure of these D-terms is

$$D_a = \frac{3\mu(U_a)}{\kappa} + \frac{1}{n_a} \sum_{\substack{b \\ b \neq a}} (|S_{ab}|^2 - |S_{ba}|^2) , \tag{4.40}$$

where S_{ab} are bundle moduli singlets, $\kappa = d_{ijk} t^i t^j t^k$ is the Calabi-Yau volume and the slope, μ , has been defined in (4.10).⁵ Given these expressions for the D-terms, we can now interpret the geometric discussion above from a four-dimensional point of view.

⁵For simplicity of notation we have dropped possible Kahler metrics in front of the singlet matter field terms. Also, note that, in order to keep our notation more “covariant”, we have written down f D-terms for only $f - 1$ anomalous $U(1)$ symmetries. This does not lead to an additional constraint since $\sum_a n_a D_a = 0$, as a consequence of $c_1(V) = \sum_a c_1(U_a) = 0$. The actual four-dimensional D-terms are $f - 1$ linearly independent combinations of the D_a .

For a split bundle $U = U_a \oplus U_b$ all the relevant singlet fields (if they even exist) should be zero in vacuum, that is, $\langle S_{ab} \rangle = \langle S_{ba} \rangle = 0$. In this case, in order to satisfy the D-term equations, we should require that $\mu(U_a) = \mu(U_b) = 0$. This corresponds precisely to the zero-slope poly-stability conditions (4.13) on the split bundle which constrain the Kahler moduli. If no fields S_{ab} or S_{ba} exist this is the only way to satisfy the D-term equations. In this case, some of the Kahler moduli are stabilised and the bundle is necessarily split. On the other hand, if fields S_{ab} or S_{ba} do exist they may be given a vacuum expectation value, provided this is consistent with D- and F-flatness of the four-dimensional theory. This corresponds to a non-trivial extension U (respectively, \tilde{U}) of $U_a \oplus U_b$, with structure group $U(n_a + n_b)$. Solving the D-term equations then typically requires that $\mu(U_a) \neq 0$ and $\mu(U_b) \neq 0$, that is, at the same time as moving in the bundle moduli space one is forced to move in the Kahler moduli space away from the special zero-slope locus. Due to the non-trivial singlet VEVs some or all of the anomalous $U(1)$ symmetries are broken spontaneously. While this merely generates an additional breaking for the $U(1)$ symmetries which are anomalous this can give masses to the $U(1)$ symmetries which have remained massless at the split locus. Also, note that the direction in which one can move away from the slope-zero locus is dictated by which singlet fields are present. If there are fields S_{ab} with $a < b$ only, then one can only leave the locus in directions characterised by $\mu(U_a) < 0$ and $\mu(U_b) > 0$. For fields S_{ba} only we have the opposite inequalities and, finally, if both types of fields are present there need not be a restriction in Kahler moduli space.

One must also be careful in this context to consider the constraints arising from bundle holomorphy. While these conditions are satisfied automatically by ensuring that the component bundles U_a are holomorphic, it is not automatic that all choices of singlet VEVs lead to holomorphic bundles. These obstructions to the deformations appear as non-trivial F-terms in the four-dimensional theory, as discussed in sections 2.4 and 3.1.4. Such F-term obstructions must be explored on a case-by-case basis (see [52, 63, 66–68] for further discussions).

In summary, depending on the choice of bundle V , various or possibly all of the different splitting types in table 2 may be connected by deformations and may, hence, be part of the same moduli space. Whether bundles can be connected in such a way, that is, whether one can move away from a given zero-slope locus in Kahler moduli so that the $U(1)$ symmetries are spontaneously broken, is of obvious relevance for the phenomenology of our models. It is important that we can answer these questions, not only by exploring the effect of singlet VEVs in the four-dimensional effective theory, but also explicitly by a direct calculation of the cohomology groups $H^1(X, U_a \otimes U_b^*)$.

With this brief overview of deformation theory in place, we have completed our exploration of the consistency conditions on split gauge bundles in heterotic compactifications. We briefly summarize these conditions below, before turning to phenomenologically interesting examples.

4.8 The conditions imposed on the geometry

4.8.1 Consistency conditions on the bundles

Here we will concisely summarize the conditions for a consistent heterotic compactification discussed in the previous section. We assume that we have given a Calabi-Yau manifold

X with a freely-acting discrete symmetry Γ . In order to be able to break $SU(5)$ to the standard model group, for $\Gamma = \mathbb{Z}_n$ we should have $n \neq 5$ and for Γ given by a product of \mathbb{Z}_n factors at least one factor should be different from \mathbb{Z}_5 . As discussed earlier, such a Wilson line can be characterized by two Γ representations W and \tilde{W} , subject to the constraints (4.27). Further, on the Calabi-Yau manifold X , we have a vector bundle $V = \oplus_{a=1}^f U_a$ with $c_1(V) = 0$, where U_a are bundles with structure group $U(n_a)$ and first Chern class $k_a^i = c_1^i(U_a)$. Then, the full structure group is given by $H = S(U(n_1) \times \dots \times U(n_f))$ and will frequently be referred to by the split vector $\mathbf{n} = (n_1, \dots, n_f)$. On this data we impose the following conditions.

- (C1) There should be an anomaly-free completion of the model, that is we should be able to satisfy eq. (4.1). In practice, we will guarantee this by requiring that

$$c_2(TX) - c_2(V) \stackrel{!}{\in} \text{Mori cone of } X, \tag{4.41}$$

so that a complete model can be obtained by adding a suitable number of five-branes and choosing the hidden bundle to be trivial.

- (C2) The bundle V must be holomorphic and poly-stable with zero slope, somewhere in the Kahler cone of X . As discussed, the second of these constraints means that every constituent bundle U_a is stable with vanishing slope and that the resulting constraints on the Kahler moduli have a common solution. In particular, all of the zero slope conditions

$$\mu(U_a) \sim d_{ijk} k_a^i t^j t^k = k_a^i \kappa_i \stackrel{!}{=} 0 \tag{4.42}$$

should have a common solutions. This necessarily means that the number of linearly independent Chern classes \mathbf{k}_a is less than $h^{1,1}(X)$, the number of Kahler moduli.

- (C3) The bundle V needs to be equivariant under the discrete symmetry Γ for it to descend to a bundle \hat{V} on the quotient manifolds $\hat{X} = X/\Gamma$. In order to preserve the splitting type, we realize this by asking all U_a to be Γ -equivariant individually. Hence, they descend to bundles \hat{U}_a , so that $\hat{V} = \oplus_{a=1}^f \hat{U}_a$. In such cases, we can characterize the U_a equivariant structures by Γ -representations χ_a^* .

4.8.2 Conditions on spectrum of GUT theory

In preparation for considering specific examples, we list here a set of conditions that we must impose on split heterotic models that could lead to standard model-like physics in the associated four-dimensional theory. Some of the physical conditions on the spectrum can already be formulated in terms of the underlying GUT theory and are, therefore, somewhat easier to check.

- (S1) In order to have a chiral asymmetry of three families we require that

$$-\text{ind}(V) = \sum_a N(\mathbf{10}_{e_a}) \stackrel{!}{=} 3|\Gamma|. \tag{4.43}$$

Since the index divides by the group order, $|\Gamma|$, as we descend to the downstairs bundle \hat{V} this guarantees a chiral asymmetry of three $\mathbf{10}$ multiplets. The total chiral asymmetry of $\bar{\mathbf{5}}$ multiplets equals $-\text{ind}(\wedge^2 V)$. However, since $\text{ind}(\wedge^2 V) = \text{ind}(V)$ for an $\text{SU}(5)$ bundle no additional condition is needed for the $\bar{\mathbf{5}}$ multiplets.

(S2) We would like to exclude any mirror families originating from $\bar{\mathbf{10}}$ multiplets. It is possible (although perhaps not likely) that an entire $\mathbf{10}$ multiplet is projected out by the Wilson line. Also, $\mathbf{10}\text{-}\bar{\mathbf{10}}$ pairs might be lifted away from the split locus, when singlet VEVs are switched on. However, we would like to use a cleaner approach and remove $\bar{\mathbf{10}}$ multiplets completely at the GUT level and on the split locus. Hence, we require that

$$\sum_a n(\bar{\mathbf{10}}_{-\mathbf{e}_a}) = h^2(X, V) \stackrel{!}{=} 0. \quad (4.44)$$

(S3) In the downstairs spectrum we would like to have at least one pair of Higgs doublets. A necessary condition for this is the existence of at least one $\mathbf{5}\text{-}\bar{\mathbf{5}}$ pair in the GUT theory. This means that

$$\sum_a n(\mathbf{5}_{-2\mathbf{e}_a}) + \sum_{a < b} n(\mathbf{5}_{-\mathbf{e}_a - \mathbf{e}_b}) = h^2(X, \wedge^2 V) \stackrel{!}{>} 0. \quad (4.45)$$

If all conditions above are satisfied we are guaranteed a model with the standard model gauge group (times $\text{U}(1)$ symmetries, possibly anomalous) and precisely three families of quarks and leptons (and no mirror families). The only additional multiplets are whatever remains from the $\mathbf{5}\text{-}\bar{\mathbf{5}}$ pairs after including the Wilson line. We need to ensure that the Higgs triplets are projected out and at least one pair of Higgs doublets survives.

(H1) In order to remove all Higgs triplets we demand that

$$n(\bar{T}_a) = h^2(X, \wedge^2 U_a, \chi_a \otimes \chi_a \otimes W) \stackrel{!}{=} 0, \quad n(\bar{T}_{a,b}) = h^2(X, U_a \otimes U_b, \chi_a \otimes \chi_b \otimes W) \stackrel{!}{=} 0. \quad (4.46)$$

As mentioned earlier, for an Abelian group Γ , a sufficient (although not strictly necessary) condition for the Higgs triplets to be projected out in a particular sector $\wedge^2 U_a$ or $U_a \otimes U_b$ is

$$h^2(X, \wedge^2 U_a) \stackrel{!}{<} |\Gamma|, \quad h^2(X, U_b \otimes U_c) \stackrel{!}{<} |\Gamma|. \quad (4.47)$$

(H2) In order to keep at least one pair of Higgs doublets we demand that

$$n(\bar{H}_a) = h^2(X, \wedge^2 U_a, \chi_a \otimes \chi_a \otimes \tilde{W}) \stackrel{!}{>} 0 \quad n(\bar{H}_{a,b}) = h^2(X, U_a \otimes U_b, \chi_a \otimes \chi_b \otimes \tilde{W}) \stackrel{!}{>} 0 \quad (4.48)$$

in at least one sector $\wedge^2 U_a$ or $U_a \otimes U_b$.

If these two further conditions are satisfied, we have a standard model spectrum, that is a standard model gauge group (times $\text{U}(1)$ factors), precisely three families of quarks and leptons, one or more pairs of Higgs doublets and no exotic fields charged under the standard model group.

5 Heterotic line bundle models on CICYs

In this section, we will begin a systematic and algorithmic search for split heterotic model with a standard model particle spectrum, that is, models which satisfy the constraints (C1)–(C3), (S1)–(S3) and (H1), (H2) described in the previous section. As our starting point, we will focus on “maximally split” models defined by a sum of five line bundles

$$V = \bigoplus_{a=1}^5 L_a . \tag{5.1}$$

That is, we consider purely Abelian internal gauge fields such that all higher-rank factors H_s in the split structure group given in (4.3) vanish. This choice is motivated by starting with the most basic splitting type and by the relative technical simplicity of line bundles. In particular, the slope-stability conditions of section 4.3 reduce to the straightforward zero-slope condition of (4.13) in this case. Starting from line-bundle models we still have a window into a larger moduli space of geometries via the bundle deformations discussed in section 4.7. Moreover, as described in section 3, it is frequently of phenomenological interest to turn on singlet VEVs, thereby continuing to other splitting types.

In order to systematically construct a large-scale data set of line bundle models, we must specify an explicit class of Calabi-Yau geometries and the available line bundles over them. In this paper, we will take as our arena perhaps the simplest set of Calabi-Yau three-folds, namely complete intersections in products of projective spaces (CICYs), and line bundles defined over them. The set of CICYs have been completely classified [69, 70] and there exists a list [71] of 7890 such spaces together with their basic properties.

In addition, the two main technical issues relevant for the construction of heterotic line bundle models have been resolved for CICYs. First, freely acting discrete symmetries of CICY manifolds have been classified recently in ref. [72].⁶ Secondly, efficient methods to compute line bundle cohomology on CICYs have been described and implemented by the authors in ref. [16, 17, 73]. Our constructions will heavily rely on both of these technical results.

We begin by reviewing the relevant properties of CICYs and line bundles over them before we move on to the construction of models. Since the Picard number, that is, the dimension of the space of available line bundles, is given by the Hodge number $h^{1,1}(X)$ of the underlying three-fold, X , one expects the number of models and their complexity to increase with $h^{1,1}(X)$. We will, therefore, start with CICYs with the smallest possible Picard number $h^{1,1}(X) = 2$ (recall that the vanishing slope constraint (4.13) cannot be satisfied for $h^{1,1}(X) = 1$) and work our way up to and including $h^{1,1}(X) = 5$ (the smallest Picard number for which models free of massless U(1) vector fields can be constructed). Of course, on each of those CICYs we must study a large number of line bundle sums in order to find those with a phenomenologically interesting particle spectrum.

⁶This classification covers discrete groups which can be constructed via a linear action on the coordinates of the ambient projective spaces.

5.1 Complete intersection Calabi-Yau three-folds

Complete intersection Calabi-Yau manifolds X are defined as the common zero locus of K homogeneous polynomials on the ambient space $\mathcal{A} = \mathbb{P}^{N_1} \times \dots \times \mathbb{P}^{N_m}$, given by a product of m projective spaces with dimensions N_i . To obtain three-folds from such complete intersections we obviously need $\sum_{i=1}^m N_i - K = 3$. The defining polynomials, p_α , are characterised by their multi-degrees $\mathbf{q}_j = (q_\alpha^1, \dots, q_\alpha^m)$, where q_α^i specifies the degree of p_α in the coordinates of the factor \mathbb{P}^{N_i} in \mathcal{A} . This information can be summarised by the configuration matrix

$$\begin{bmatrix} \mathbb{P}^{N_1} & q_1^1 & q_2^1 & \dots & q_K^1 \\ \mathbb{P}^{N_2} & q_1^2 & q_2^2 & \dots & q_K^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbb{P}^{N_m} & q_1^m & q_2^m & \dots & q_K^m \end{bmatrix}_{m \times K} \quad (5.2)$$

whose columns are the multi-degrees of the defining polynomials. In order to obtain a Calabi-Yau manifold, that is, a manifold with vanishing first Chern class, the conditions $\sum_{\alpha=1}^K q_\alpha^i = N_i + 1$ need to be satisfied for all i , so that each row of the configuration matrix sums up to the dimension of the associated projective space plus one.

As stated above, the classification of CICYs has led to a list of 7890 such configuration matrices (although not all of them correspond to different Calabi-Yau three-folds [17]). For the purpose of this paper, we will only be interested in the low Picard number cases within this set, satisfying $2 \leq h^{1,1}(X) \leq 5$, which, in addition have at least one symmetry of the form classified in [72]. There are 65 such configurations although only 39 of them turn out to be inequivalent.⁷ These 65 manifolds, together with their main properties are available at the website listed in ref. [71].

We need to collect a few general properties of these manifolds which will be used in the following. It turns out that all of the 65 manifolds relevant to our initial scan are “favourable” in the sense that their complete second cohomology descends from the ambient space. Hence, if we introduce the standard Kahler forms J_i , normalised as $\int_{\mathbb{P}^{N_i}} J_i^{N_i} = 1$, on the projective factors of the ambient space their restriction to X (which we denote by the same symbol) spans the complete second cohomology of X . If we introduce the forms $\rho_X = \wedge_{\alpha=1}^K (\sum_i q_\alpha^i J_i)$ the triple intersection numbers in the basis $\{J_i\}$ can be easily computed from the formula

$$d_{ijk} = \int_X J_i \wedge J_j \wedge J_k = \int_{\mathcal{A}} J_i \wedge J_j \wedge J_k \wedge \rho_X \quad (5.3)$$

Elements of the fourth cohomology (and, by Poincaré duality, the second homology) of X will often be given relative to a basis $\{\nu^i\}$ of four forms dual to $\{J_i\}$, satisfying $\int_X J_i \wedge \nu^j = \delta_i^j$. For example, we expand the second Chern class of X as $c_2(TX) = c_{2i}(TX)\nu^i$ and the values of the coefficients $c_{2i}(TX)$ are explicitly given in the data available from the database [46]. Kahler forms J on X can be expanded as $J = t^i J_i$, where t^i are the Kahler

⁷It is still useful valuable to have the same Calabi-Yau three-fold represented by different CICY configurations since the classification of discrete symmetries, which are restricted to act linearly on the ambient space, can depend on the representation used.

moduli. For favourable CICYs, the Kahler cone is given by $t^i > 0$ for all i and effective classes in the second homology of X corresponds to positive linear combinations of the four-forms ν^i .

5.2 Line bundles on favourable CICYs

Line bundles on Calabi-Yau three-folds are classified by their first Chern class (modulo discrete torsion which determines Wilson Lines) and, on favourable CICYs, they can therefore be labelled by an integer vector $\mathbf{k} = (k^1, \dots, k^m)$. We denote by $L = \mathcal{O}_X(\mathbf{k})$ the line bundle with first Chern class $c_1(\mathcal{O}_X(\mathbf{k})) = k^i J_i$. This line bundle can also be thought of as the restriction of the ambient space line bundle $\mathcal{O}_{\mathcal{A}}(\mathbf{k}) = \mathcal{O}_{\mathbb{P}^{N_1}}(k^1) \otimes \dots \otimes \mathcal{O}_{\mathbb{P}^{N_m}}(k^m)$ to X . Its dual, L^* , is simply given by $L^* = \mathcal{O}_X(-\mathbf{k})$. With the Todd class of a Calabi-Yau three fold given by $\text{Td}(TX) = 1 + c_2(TX)/12$ it is straightforward to find the index

$$\text{ind}(L) \equiv \sum_{q=0}^3 (-1)^q h^q(X, L) = \int_X \text{ch}(L) \wedge \text{Td}(X) = \frac{1}{6} \left(d_{ijkl} k^i k^j k^k k^l + \frac{1}{2} k^i c_{2i}(TX) \right) \quad (5.4)$$

of L . However, computing the line bundle cohomology groups $H^q(X, L)$ individually is more involved. One useful relation is provided by Serre duality [76, 77] which, for Calabi-Yau three-folds, states that $H^q(X, L) \cong H^{3-q}(X, L^*)$. Positive (ample) line bundles in the present context are those for which all $k^i > 0$. To such positive line bundles we can apply Kodaira's vanishing theorem [76, 77] which states that $H^q(X, L) = 0$ for $q > 0$. Hence, for positive line bundles the zeroth cohomology is the only non-vanishing one which can be computed from the above index formula alone. Similarly, negative line bundles are those for which all $k^i < 0$. Their only non-vanishing cohomology is the third which again can be computed from the index. Unfortunately, positive and negative line bundles are not useful for our model building purposes. In fact, for such line bundles the zero slope condition (4.13) cannot be satisfied, given that the Kahler cone for our three-folds is given by $t^i > 0$. Hence, we need to understand the cohomology of "mixed" line bundles. In this context, there exist weaker vanishing theorems which can sometimes be helpful. For example, the vanishing condition mentioned in footnote 4, which states that if the slope $\mu(L) = d_{ijk} c_1^i(L) t^k$ is negative somewhere in the Kahler cone, then $H^0(X, L) = 0$ [61], can be useful in determining the cohomology of many line bundles of interest to us.

To extract further information about line bundle cohomology one can consider Koszul resolutions combined with Bott-Borel-Weil representations [78] of ambient space cohomology, as described in ref. [73]. These methods, combined with the above vanishing theorems, allow us to calculate the vast majority of line bundle cohomology groups and underlie most of the cohomology results in this paper. For a small number of line bundles these methods do not lead to a complete answer and one has to resort to more explicit methods, such as computing Cech cohomology [76, 77]. The method for computing graded line bundle cohomology on CICYs is outlined in appendix B.

In our systematic scans for heterotic line bundle standard models on CICYs, we will construct rank five bundles V with structure group $S(U(1)^5)$ by considering sums (5.1) of five line bundle $L_a = \mathcal{O}_X(k_a^i)$, satisfying $c_1(V) = \sum_a c_1(L_a) = 0$. This corresponds

to the splitting type $\mathbf{n} = (1, 1, 1, 1, 1)$ with $f = 5$ from the general set-up in section 4.1 and all basic results from this section, in particular the content of tables 3 and 4, apply with the identification $U_a = L_a$. In addition, one special feature of line bundle models has to be taken into account, namely the vanishing of the bundles $\wedge^2 U_a, \wedge^2 U_a^*$ and, hence, the absence of the corresponding $\mathbf{5}_{2e_a}, \mathbf{5}_{-2e_a}$ multiplets given in table 3. For the Chern characters and the index of such a sum of line bundles one has

$$\text{ch}_1(V) = \sum_a k_a^i \stackrel{!}{=} 0 \tag{5.5}$$

$$\text{ch}_2(V) = \frac{1}{2} d_{ijkl} \sum_a k_a^j k_a^l \tag{5.6}$$

$$\text{ind}(V) = \frac{1}{6} d_{ijkl} \sum_a k_a^i k_a^j k_a^l. \tag{5.7}$$

5.3 The scan

With the methods described above, we were able to perform the scan described in previous sections, over approximately 10^{12} different line bundle sums k_a^i , to build line bundle standard models over CICYs. The GUT models from this scan have already been presented and discussed in ref. [35]. They were obtained by considering all favourable CICYs with $h^{1,1}(X) \leq 5$ and with freely-acting symmetries. No phenomenologically viable model were found for $h^{1,1}(X) = 2, 3$. The scan, which ranged over $-3 \leq k_a^i \leq 3$ for $h^{1,1}(X) = 4$ and $-2 \leq k_a^i \leq 2$ for $h^{1,1}(X) = 5$, led to 202 GUT models on 13 CICYs. Here, we have calculated the explicit downstairs models with standard model gauge group from all these GUT models and the results are presented in the database [46]. Phenomenological properties of these models have already been discussed in section 3.

6 Further consequences of split heterotic geometry

In this section we explore the possibility of more complicated equivariant structures on the bundle V , such that not every constituent bundle U_a is equivariant by itself. Such models are relatively rare at least among line bundle models. For example, the scan in ref. [35] which led to 202 phenomenologically promising GUT models, only produced four models with non-trivial equivariant building blocks. Nevertheless, such constructions might be of interest for a number of phenomenological reasons, in particular in relation to the up-type Yukawa matrix and the possibility of a Peccei-Quinn symmetry. We begin by outlining the theoretical issues which arise in these constructions and discuss possible phenomenological applications towards the end of the section.

6.1 Line bundle models with non-trivial equivariant blocks

In this section, we explore sums of line bundles $V = \bigoplus_a L_a$, for which not all of the L_a admit an equivariant structure. Such sums can still descend to the quotient manifold, \hat{X} , if V splits into equivariant blocks, that is, into non-trivial line bundle sums which are equivariant as a whole even though their constituent line bundles are not. A simple example of this

type of non-trivial structure can be found by considering the line bundle $\mathcal{O}(1)$ on the quintic hypersurface in \mathbb{P}^4 . This manifold admits a well-known $\Gamma = \mathbb{Z}_5 \times \mathbb{Z}_5$ freely acting discrete symmetry. On the homogenous coordinates of \mathbb{P}^4 , Γ can be represented as

$$\mathbb{Z}_5^1 : x_k \rightarrow x_{k+1} \quad , \quad \mathbb{Z}_5^2 : x_k \rightarrow \alpha^k x_k \tag{6.1}$$

where α is a fifth root of unity, so $\alpha^5 = 1$. It is straightforward to verify that $\mathcal{O}(1)$ is not equivariant with respect to the symmetry (6.1). Indeed, not only is $\text{ind}(\mathcal{O}(1)) = 5$ which is clearly not divisible by $|\Gamma| = 25$, but it is straightforward to show that, for the induced action of (6.1) on $H^0(X, \mathcal{O}(1))$, the two \mathbb{Z}_5 symmetries do not commute and instead form a representation of the order 125 Heisenberg group [75] (see appendix A for mathematical details on equivariance). However, a sum of five such line bundles

$$V = \mathcal{O}(1)^{\oplus 5} \tag{6.2}$$

does admit an equivariant structure. The required morphisms $\phi \in \text{Hom}(V, V)$ are no longer a simple group character (as was the case for individually equivariant line bundles), but instead are generated by the following two matrices acting fiber-wise on the sum in (6.2)

$$\phi_1 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \phi_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \alpha^2 & 0 & 0 & 0 \\ 0 & 0 & \alpha^4 & 0 & 0 \\ 0 & 0 & 0 & \alpha^1 & 0 \\ 0 & 0 & 0 & 0 & \alpha^3 \end{pmatrix}. \tag{6.3}$$

It can be verified that under the combined action (A.3) of (6.1) and (6.3) the sections $H^0(X, \mathcal{O}(1)^{\oplus 5})$ carry a representation of $\Gamma = \mathbb{Z}_5 \times \mathbb{Z}_5$.

In this example, it is clear that V in (6.2) is indeed an equivariant bundle and equal to the pull-back of some rank five bundle, \hat{V} on $\hat{X} = X/\Gamma$. However, in this case the “downstairs” bundle, \hat{V} , is *not a sum of line bundles*, rather is some generically indecomposable rank five vector bundle on \hat{X} . Since $H^0(X, \hat{V}^* \otimes \hat{V}) = 1$, \hat{V} is in fact a simple bundle (in the algebro-geometric sense) [76, 77]. In other words, given the covering map $q : X \rightarrow X/\Gamma$, the operation of pulling back, $q^*(\hat{V})$, need not preserve the structure group. In particular, a non-split bundle may pull-back to a split one.

Such bundles are of interest to us in the present context. Although they represent more complicated, higher-rank objects over \hat{X} , their properties are completely determined by a sum of line bundles over X and, hence, are easily analyzed using the techniques we have described in this work. Although we expect relative few models with non-trivial equivariant building blocks they may have several features of phenomenological interest, as we will see. Before we begin to explore these features in detail, however, it should be noted that for such models the number of Green-Schwarz anomalous U(1) symmetries will be different in the “upstairs” and “downstairs” theories. For example, consider a sum of two line bundles $L_1 \oplus L_2 \subset V = \bigoplus_{a=1}^5 L_a$ which are equivariant as a pair but not individually. If $c_1(V) = 0$ there will be four U(1) symmetries upstairs, but only three associated to the downstairs bundle, which has the split form

$$\hat{V} = U_1 \oplus L_3 \oplus L_4 \oplus L_5 \tag{6.4}$$

with U_1 an indecomposable, $U(2)$ bundle. In the extreme case where all five line bundles form an equivariant block (as in the above toy example on the quintic), there would in fact be no enhanced $U(1)$ symmetries left in the downstairs theory. It seems in such cases one has lost not only all additional $U(1)$ symmetries but also the constraints they normally imply for the structure of the four-dimensional theory. However, since the downstairs theory associated to (\hat{X}, \hat{V}) is derived as the Γ -invariant part of the theory on (X, V) , the $U(1)$ symmetries which appear in the upstairs theory can still constrain the theory. For example, the downstairs Yukawa couplings $\lambda_{pq}^{(u)}$ for $\mathbf{10}^p \mathbf{10}^q \mathbf{5}^r$ (see eq. (2.10)) are determined by the Yoneda pairing

$$Y_{\text{downstairs}} = \int_{\hat{X}} H^1(\hat{X}, \hat{V}) \wedge H^1(\hat{X}, \hat{V}) \wedge H^1(\hat{X}, \wedge^2 \hat{V}^*). \quad (6.5)$$

This can be written in terms of the invariant part of the upstairs Yukawa couplings as

$$\text{Invariant}_{\Gamma, \phi} \left(\int_X H^1(X, q^*(\hat{V})) \wedge H^1(X, q^*(\hat{V})) \wedge H^1(X, q^*(\wedge^2 \hat{V}^*)) \right). \quad (6.6)$$

While $H^1(X, q^*(\hat{V}))$ need not be the cohomology of a simple line bundle L on X , the integral in (6.6) will in general be constrained to be $U(1)$ gauge invariant for all the anomalous $U(1)$ symmetries which arise in the upstairs theory. With this observation in hand we turn now to look at some of the phenomenological features which arise in models with non-trivial equivariant blocks.

6.2 Up-type Yukawa couplings

As discussed in section 2 (see eq. (2.10)), for line bundle models with $V = \bigoplus_a L_a$, the up-type Yukawa couplings

$$\lambda_{pq}^{(u)} \bar{H} \mathbf{10}^p \mathbf{10}^q \quad (6.7)$$

are allowed only if the Higgs charge $-\mathbf{Q}(\bar{H}) = \mathbf{e}_{\bar{h}} + \mathbf{e}_{\bar{g}}$ equals $\mathbf{Q}(\mathbf{10}^p) + \mathbf{Q}(\mathbf{10}^q) = \mathbf{e}_{a_p} + \mathbf{e}_{b_q}$. For pure line bundle models, we always have $\bar{h} < \bar{g}$ since the sector $H^1(X, \wedge^2 U_a)$ which may lead to up-type Higgs doublets with charge $-2\mathbf{e}_a$ is absent for line bundle models. Since the Yukawa terms in (6.7) are symmetric in family space it follows that up Yukawa couplings for such models without equivariant blocks cannot have rank one. This conclusion can obviously change away from the Abelian locus when singlet VEVs are switched on or when multiple Higgs pairs are considered.

Another possibility to generate phenomenologically desirable rank one up Yukawa matrices at the perturbative level is to consider line bundle models with non-trivial equivariant blocks. For example, consider the case where two line bundles, $L_1 \oplus L_2$ are associated to the pull-back of a non-Abelian rank two bundle, \hat{U} on the quotient space, \hat{X} . Then the downstairs Yukawa couplings may contain a terms of the form

$$Y_{\text{downstairs}} \sim \int_{\hat{X}} H^1(\hat{X}, \hat{U})_{\mathbf{e}_a} \wedge H^1(\hat{X}, \hat{U})_{\mathbf{e}_a} \wedge H^1(\hat{X}, \wedge^2 \hat{U}^*)_{-2\mathbf{e}_a} \quad (6.8)$$

which can lead to a rank one mass matrix. Such a term is possible since each of the downstairs cohomology groups above, pulls back not to a single line bundle cohomology on

X but a sum of two line bundles. That is, $Y^{(u)}$ is given by the invariant part of

$$\int_X \begin{pmatrix} H^1(X, L_1) & \\ & H^1(X, L_2) \end{pmatrix} \wedge \begin{pmatrix} H^1(X, L_1) & \\ & H^1(X, L_2) \end{pmatrix} \wedge H^1(X, L_1^* \otimes L_2^*) \quad (6.9)$$

where the 2×2 matrices schematically represent the group invariant elements of cohomology that are obtained from non-trivial combinations of $H^1(X, L_1)$ and $H^1(X, L_2)$.

6.3 Realizing a Peccei-Quinn symmetry

Peccei-Quinn $U(1)$ symmetries, denoted by $U_{PQ}(1)$, may be used to explain some of the more puzzling aspects of supersymmetric GUTs. Of particular interest is their role in addressing the μ problem, that is, the problem of why the superpotential Higgs mass term is so small given that it is a super-renormalizable operator. The class of models explored in this work, perhaps point at least in the right direction since the Higgs doublets are precisely massless at the split locus of the bundle. However this feature is not guaranteed to persist away from the split locus when non-trivial singlet VEVs are switched on and moduli stabilization might well drive some models to such a region in moduli space. However, if the μ term is forbidden by a $U_{PQ}(1)$ symmetry it is more robustly protected. Peccei-Quinn symmetries can also be motivated from proton stability. Assuming neutrality of the Yukawa-couplings it is the easy to show that any $U(1)$ symmetry which forbids the dimension five proton decay operators $\bar{\mathbf{5}}^p \mathbf{10}^q \mathbf{10}^r \mathbf{10}^s$ (see eq. (2.12)) must in fact be a $U_{PQ}(1)$ symmetry. For these reasons, it is important to study whether such a symmetry can be realized in split heterotic models.⁸

We will now argue that in split models with bundle $V = \bigoplus_a U_a$ and each U_a equivariant individually, no $U_{PQ}(1)$ can arise. For such models, each constituent bundle U_a descends to a bundle \hat{U}_a on the quotient space $\hat{X} = X/\Gamma$ and it follows that $\text{ind}(\hat{U}_a \otimes \hat{U}_b) = \text{ind}(U_a \otimes U_b)/|\Gamma|$ (and similarly for $\wedge^2 U_a$), that is the index is divisible by the group order $|\Gamma|$ in each sector of $\bar{\mathbf{5}}_{\mathbf{e}_a + \mathbf{e}_b}$ and $\mathbf{5}_{-\mathbf{e}_a - \mathbf{e}_b}$ multiplets. Of course, the chiral asymmetry in each such sector must be zero or negative, $\text{ind}(\hat{U}_a \otimes \hat{U}_b) \leq 0$, or else this sector leads to complete $\mathbf{5}$ multiplets downstairs which contain phenomenologically unacceptable Higgs triplets. The $\bar{\mathbf{5}}$ excess in each sector will contribute towards the families, while divisibility of the index implies that any remainders from the $\bar{\mathbf{5}}-\mathbf{5}$ pairs have to come in vector-like pairs. In conclusion, if all constituent bundles U_a are equivariant individually and we require Higgs triplets to be projected out by Wilson-line breaking, then the Higgs doublets can always be grouped into vector-like pairs under the $U(1)$ symmetries.

We would now like to argue, for the case of line bundle sums $V = \bigoplus_a L_a$, that this statement does not necessarily hold any more in the presence of non-trivial equivariant blocks. Suppose that the line bundle sum $L_1 \oplus L_2 \subset V$ constitutes a non-trivial equivariant block and is the pull-back of a non-decomposable rank two bundle U on \hat{X} , that is $L_1 \oplus L_2 = q^*(\hat{U})$. Now consider a third line bundle, L_a in V such that $\text{ind}((L_1 \oplus L_2) \otimes L_a) < 0$ and $h^2((L_1 \oplus L_2) \otimes L_a) < |\Gamma|$. The Higgs multiplets which arise from $(L_1 \oplus L_2) \otimes L_a$ can be obtained from the up-stairs cohomology by extracting the part which corresponds to the

⁸For analyses of similar issues in an F-theory context see for example [79–81].

appropriate Γ -representation, along the lines of section 4.5. If this process leads to Higgs multiplets

$$H \in H^1(X, L_2 \otimes L_a), \quad \bar{H} \in H^1(X, L_1^* \otimes L_a^*), \quad (6.10)$$

a pattern which presumably can be arranged for appropriate model building choices, then it is clear that a linear combination of the upstairs $U(1)$ symmetries will behave as a $U_{PQ}(1)$ symmetry and forbid the μ -term as well as dangerous dimension five proton decay operators. It is worth noting that in this case, although the μ -terms of the final four-dimensional theory are effectively forbidden by the $U(1)$ invariance described above, there is no corresponding $U(1)$ gauge symmetry in the four-dimensional theory.

It should be clear from the above discussion that many features of the four-dimensional theory can be affected by non-trivial equivariant blocks, often in subtle and surprising ways. We hope to explore the consequences of such behaviour in more detail in future work. For now, it should be noted that, unfortunately, the interesting features described above do not arise in the four models with non-trivial equivariant blocks in our current data set. All four models suffer from the problem that the color triplets cannot be projected out. If desired, one could modify the scanning criteria we have employed to generate models with a $U_{PQ}(1)$ using the methods described in this paper.

7 Summary

In this paper, we have developed the formalism for heterotic models with split vector bundles, with particular emphasis on the maximal splitting into line bundle sums. We have shown that such heterotic line bundle models are a promising arena for building particle physics models within string theory and we have presented a database [46] of about 400 heterotic line bundle standard models. All of these models have a standard model gauge group times four additional and frequently Green-Schwarz anomalous $U(1)$ symmetries, the exact matter spectrum of the MSSM, one or more pairs of Higgs doublets, a spectrum of bundle moduli which are standard model singlets and no exotics charged under the standard model group of any kind.

The additional $U(1)$ symmetries constrain the allowed operators in the four-dimensional theory and provide an interesting tool for low-energy phenomenology. For our line bundle standard models we have worked out the spectrum of $U(1)$ invariant four-dimensional operators, including operators with insertions of the bundle moduli singlets. The results which are presented in the database [46] can be used to discuss a large variety of phenomenological issues, including the structure of Yukawa couplings, proton stability, the μ -term and neutrino masses. In section 3 these issues have been illustrated for a particular example model from the database. We have emphasized that this specific model has not been put forward as the most attractive one, but merely as an example to demonstrate the phenomenological possibilities of heterotic line bundle models. We hope that the database [46] provides a starting point for exploring these phenomenological questions in depth and, eventually, to isolate a true string standard model which reproduces all known features of low-energy physics.

The present work is merely the first step in this direction. First and foremost we need to extend the current scans for line bundle standard models. The present models, collected in the database [46] are defined on a relatively small set of Calabi-Yau three-folds and rely on line bundles with a fairly restricted range of integer entries. Both of these restrictions can be lifted, at least to a certain extent, and we expect to find many more line bundle standard models in this case. Ideally, this should eventually lead to a classification of all line bundle standard models on the known sets of Calabi-Yau three-folds. Such a data set, which will likely contain tens of thousands or possibly even more models, would be a legitimate starting point for the search for a true string standard model.

Acknowledgments

L.A. is supported by the Fundamental Laws Initiative of the Center for the Fundamental Laws of Nature, Harvard University. J.G. would like to acknowledge support by the NSF-Microsoft grant NSF/CCF-1048082. A.L. is supported in part by the EC 6th Framework Programme MRTN-CT-2004-503369 and EPSRC network grant EP/I02784X/1. He would like to thank the string theory group at LMU Munich for hospitality. The work of EP is supported by a Marie Curie Intra European Fellowship within the 7th European Community Framework Programme.

A Equivariant structures and Wilson lines

One of the requirements we impose on line bundles sums U is that it respects the freely-acting symmetry, Γ , by which we quotient our Calabi-Yau three-folds. More precisely, we need that our bundle $U \xrightarrow{\pi} X$ descends to a bundle \hat{U} on X/Γ , in the sense that $U \cong q^*\hat{U}$. For a bundle to descend to the quotient space it is necessary that the automorphisms Γ of X “lift” to automorphisms of the bundle U over X . In other words, for each $g \in G$ there must exist a bundle morphism, $\phi_g : U \rightarrow U$ which commutes with the projection $\pi : U \rightarrow X$ and covers the action $g : X \rightarrow X$ on the base. Such a lifting of the group action is known as an invariant structure on U . We can express this concisely by the commutativity of the following diagram for all $g \in \Gamma$.

$$\begin{array}{ccc}
 U & \xrightarrow{\phi_g} & U \\
 \pi \downarrow & & \downarrow \pi \\
 X & \xrightarrow{g} & X
 \end{array} \tag{A.1}$$

In addition to invariance we must require that the ϕ_g satisfy what is called the co-cycle condition, namely that $\forall g, h \in \Gamma$,

$$\phi_g \circ \phi_h = \phi_{gh} . \tag{A.2}$$

An invariant structure which satisfies the cocycle condition is called an equivariant structure on U . If U admits such a set of morphisms it is said to admit an equivariant structure and, in this case, it descends to a bundle \hat{U} on X/Γ . Indeed, the set of vector bundles on X/Γ is in one-to one correspondence with the set of equivariant vector bundles on X .

How does one decide whether a sum of line bundles admits such an equivariant structure? The bundle morphisms ϕ_g induce linear maps Φ_g on the global sections, $s \in H^0(X, U)$, of U which are defined by

$$s \rightarrow \Phi_g(s) = \phi_g \circ s \circ g^{-1} . \quad (\text{A.3})$$

From eq. (A.2), these induced actions on the sections satisfy

$$\Phi_g \circ \Phi_h = \Phi_{gh} , \quad (\text{A.4})$$

that is, they form a Γ representation on the sections of U . If the bundle U is globally generated by its sections this provides a practical way of checking equivariance (see ref. [18] for a more detailed discussion of this point). In such a case one can reconstruct the bundle morphisms ϕ_g from the action on section Φ_g and so U is equivariant if and only if such linear maps on the sections exist. Since the sections of U are usually given by polynomials in the ambient space homogeneous coordinates, checking for the existence of such maps on the sections is a practical possibility. If the bundle U is not globally generated, as will frequently be the case for our examples, this method does not immediately apply. However, in this case, one can consider the twisted bundle $U \otimes L^p$, where L is an ample line bundle known to be equivariant and $p > 0$ is a sufficiently large integer such that the twisted bundle is globally generated. Equivariance of $U \otimes L^p$ can then be checked on its sections as described above and it is equivariant if and only if the original bundle U is.

To consider the actions of symmetries on sections, we shall start by considering symmetry actions on homogeneous polynomials. The two cases of interest to us here are where Γ is a single \mathbb{Z}_m factor, or a direct product of two such groups. In the discussion to follow much of the complication which will occur is only relevant for the case where we have more than one Abelian factor, and even then only for line bundles which must appear with a non-trivial multiplicity if they are to admit an equivariant structure. Here, we will discuss the general case as, once this is understood, specializing to the more straightforward examples is trivial. In particular it is very easy to see, by simplifying the following analysis, that for $\Gamma = \mathbb{Z}_m$ all line bundles admit an equivariant structure.

Consider a group $\Gamma = \mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2}$ acting on the ambient space $\mathcal{A} = \mathbb{P}^{N_1} \times \dots \times \mathbb{P}^{N_m}$. We will use indices i, j, \dots to label projective space factors and r, s, \dots to label the Abelian factors in Γ . For such an action to be well defined on the product of projective spaces we have,

$$M_r^{m_r} \begin{pmatrix} a_0^i \\ \vdots \\ a_{N_i}^i \end{pmatrix} = \lambda_r^i \begin{pmatrix} a_0^i \\ \vdots \\ a_{N_i}^i \end{pmatrix} \quad \forall a, i , \quad (\text{A.5})$$

where the a^i are the homogeneous coordinates on the i 'th projective space and M_r is the action of the \mathbb{Z}_{m_r} factor in Γ . In addition, we have the following to ensure that the Abelian factors commute when considered as an action on the projective spaces.

$$M_r M_s \begin{pmatrix} a_0^i \\ \vdots \\ a_{N_i}^i \end{pmatrix} = \Gamma_{rs}^i M_s M_r \begin{pmatrix} a_0^i \\ \vdots \\ a_{N_i}^i \end{pmatrix} \quad (\text{A.6})$$

Note that by definition we have $\Gamma_{sr}^i = (\Gamma_{rs}^i)^{-1}$. Given such an action on the homogeneous coordinates one can compute the action on a vector of polynomials of multi-degree $d^\alpha = [d_1^\alpha, \dots, d_m^\alpha]$, which we denote P_{d^α} (note here that α labels the polynomial and $i = 1, \dots, m$ the projective spaces).⁹

$$M_r M_s P_{d^\alpha} = \gamma_{rs}^\alpha M_s M_r P_{d^\alpha} \tag{A.7}$$

$$M_r^{m_r} P_{d^\alpha} = \Lambda_r^\alpha P_{d^\alpha} \tag{A.8}$$

$$\gamma_{rs}^\alpha = \prod_{i=1}^m (\Gamma_{rs}^i)^{d_i^\alpha} \tag{A.9}$$

$$\Lambda_r^\alpha = \prod_{i=1}^m (\lambda_r^i)^{d_i^\alpha} \tag{A.10}$$

Now that we have understood the action of our symmetries on vectors of polynomials we can proceed to consider such polynomials to be the sections of globally generated sums line bundles on the products of projective spaces. Using the global generation property to induce the group action on the fibres from those of the sections we see that, in general, the naive induced action given in (A.7) and (A.8) need not give rise to an equivariant structure on the bundle. Equivalently, because of the possibility of projective rescalings, equations (A.7) and (A.8) do not obey a cocycle condition such as (A.4) when considered as an action on sections. In order to “fix up” these transformations we include an additional bundle morphism \hat{M}_r for each group action, which acts as a $GL(N)$ transformation on the sum of N line bundles.¹⁰ Note that in the case of a single \mathbb{Z}_m we would only have equation (A.8), and as such we could always lay down an equivariant structure on any single line bundle, simply by taking the associated bundle morphism to be Λ_r^{1/m_r} multiplied by the identity map.

Returning to the case of two Abelian factors in Γ , we then find the following requirements on the bundle morphisms from (A.4).

$$(\hat{M}_r M_r)^{m_r} P_d = P_d \tag{A.11}$$

$$\Rightarrow \hat{M}_r^{m_r} = \begin{pmatrix} \Lambda_r^1 & & \\ & \Lambda_r^2 & \\ & & \ddots \end{pmatrix} \text{ and} \tag{A.12}$$

$$(\hat{M}_r M_r)(\hat{M}_s M_s) P_d = (\hat{M}_s M_s)(\hat{M}_r M_r) P_d \tag{A.13}$$

$$\Rightarrow \hat{M}_r \hat{M}_s \begin{pmatrix} \gamma_{rs}^1 & & \\ & \gamma_{rs}^2 & \\ & & \ddots \end{pmatrix} M_s M_r P_d = \hat{M}_s \hat{M}_r M_s M_r P_d \tag{A.14}$$

Thus demanding a good group action on the sum of line bundles on the ambient space, we obtain the conditions on our bundle morphisms \hat{M}_r , determined purely in terms of the

⁹Strictly, we should use a different symbol for the action on vectors of polynomials and the homogenous coordinates. In this appendix, however, we will use M_a for both in order to induce a less cluttered notation. Which action is meant is unambiguous from context.

¹⁰To make a connection to the notation in our general discussion, the M_r actions correspond to the g^{-1} factor in (A.3) and the \hat{M}_r correspond to the bundle morphisms ϕ_g which we must choose such that a cocycle condition of the form (A.4) is satisfied.

original action on the homogenous coordinates, which follow.

$$\hat{M}_r^{m_r} \begin{pmatrix} \Lambda_r^1 & & \\ & \Lambda_r^2 & \\ & & \ddots \end{pmatrix} = 1 \quad \forall a, \quad \hat{M}_r \hat{M}_s \begin{pmatrix} \gamma_{rs}^1 & & \\ & \gamma_{rs}^2 & \\ & & \ddots \end{pmatrix} = \hat{M}_s \hat{M}_r \quad \forall r \neq s. \quad (\text{A.15})$$

In this paper we restrict ourselves to the case of individually equivariant line bundles or non-trivial equivariant blocks composed from a sum of the same line bundle.¹¹ Specializing to one such sum of identical line bundles for simplicity, we can set $d^\alpha = d^\beta \quad \forall \alpha, \beta$. This simplification allows us to write $\Lambda_r^\alpha \cong \Lambda_r$ and $\gamma_{rs}^\alpha \cong \gamma_{rs} \quad \forall \alpha$. Equations (A.15) then simplify to,

$$\hat{M}_r^{m_r} \Lambda_r = 1 \quad \text{and} \quad \hat{M}_r \hat{M}_s \gamma_{rs} = \hat{M}_s \hat{M}_r \quad (\text{A.16})$$

It is easy to show that if the equations (A.16) are to have a solution then $\gamma_{rs}^{\hat{l}} = 1$ where $\hat{l} = \text{LCM}(m_r, m_s)$. We can, without loss of generality, choose \hat{M}_1 to be diagonal. Combining these two observations, it is then possible to show that a solution to equations (A.16), giving rise to a good group action, only exists if there we have a sum of \hat{m} identical line bundles, where \hat{m} is the minimal integer such that $\gamma^{\hat{m}} = 1$ (where $\gamma = \gamma_{12}$). In such a case the solution to the system takes the following form

$$\hat{M}_1 = \begin{pmatrix} \tilde{M}_1 & & \\ & \tilde{M}_1 & \\ & & \ddots \end{pmatrix} \quad \text{where} \quad \tilde{M}_1 = \Lambda_1^{-\frac{1}{N_1}} \begin{pmatrix} 1 & & \\ & \gamma & \\ & & \ddots \\ & & & \gamma^{\hat{m}-1} \end{pmatrix}, \quad (\text{A.17})$$

$$\hat{M}_2 = \begin{pmatrix} \tilde{M}_2 & & \\ & \tilde{M}_2 & \\ & & \ddots \end{pmatrix} \quad \text{where} \quad \tilde{M}_2 = \begin{pmatrix} 0 & X_1 & 0 & 0 & \dots \\ 0 & 0 & X_2 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ X_{\hat{m}} & 0 & 0 & 0 & \dots \end{pmatrix}$$

with $(\prod_{i=1}^{m_2} X_i)^{\frac{m_2}{\hat{m}}} = \Lambda_2^{-1}$. The above discussion allows us to lay down an equivariant structure on any sum of globally generated line bundles on products of projective spaces which admits an equivariant structure. Non-globally generated line bundles can be dealt with using this formalism by employing the technique of twisting by an ample equivariant line bundle as discussed earlier in this section. Finally we should consider equivariant structures on the restriction of these line bundles to the Calabi-Yau three-fold, defined as a complete intersection in the ambient space \mathcal{A} . Fortunately, for our favourable manifolds, due to the structure of the Koszul sequence and thanks to the fact that the normal bundle to the Calabi-Yau itself admits an equivariant structure, equivariant line bundle sums on the three-fold are in one-to-one correspondence with their ambient space counterparts.

¹¹In fact for the situation at hand, due to the structure of zeroth cohomologies of line bundles on projective spaces, this simplification is true in generality. It is easy to show that, if one has a morphism ϕ_g which mixes different line bundles, its inverse $\phi_{g^{-1}}$ can not exist as one of the required homomorphism groups between line bundles will vanish. This of course means that we can not construct an equivariant structure on such a mixed sum of line bundles.

B Representation content of cohomology

Based on the method to find an explicit equivariant structure on U discussed in the previous appendix, we must now explain how to extract the representation content for cohomologies of U . If a bundle admits a Γ -equivariant structure, then the cohomology groups $H^q(X, U)$ form Γ -representations. We can then define the cohomology groups $H^q(X, U, R)$, the subspaces of $H^q(X, U)$ which transform under the Γ -representation R , and their dimensions $h^q(X, U, R) = \dim(H^q(X, U, R))$. As in table 4, the physical properties of our models can then be expressed in terms of these “graded” cohomologies.

How do we compute these cohomologies? First of all, we relate the bundle U to its ambient space counterpart \mathcal{U} by the Koszul resolution. For illustration purposes let us consider a co-dimension one Calabi-Yau manifolds X so that the Koszul resolution is given by a short exact sequence

$$0 \rightarrow \mathcal{N}^* \otimes \mathcal{U} \rightarrow \mathcal{U} \rightarrow U \rightarrow 0, \tag{B.1}$$

where \mathcal{N} is a bundle whose restriction to X is the normal bundle of the Calabi-Yau manifold. The cohomologies of the ambient space bundles \mathcal{U} and $\mathcal{N}^* \otimes \mathcal{U}$ (and of bundles $\wedge^p \mathcal{N}^* \otimes \mathcal{U}$ for higher co-dimensions) can be expressed in terms of polynomials, using the Bott-Borel-Weil representation of cohomology on projective spaces [78]. On those polynomial representatives we can explicitly act with the symmetry transformations in Γ and, together with the equivariant structure on U , determined as explained in appendix A, the Γ -characters $\chi_{p,q}$ of $H^q(\mathcal{A}, \wedge^p \mathcal{N}^* \otimes \mathcal{U})$ can be computed. These characters are related to the characters χ_U^q of $H^q(X, U)$ by the long exact sequence in cohomology, associated to the Koszul resolution (B.1). As an example, if $H^1(\mathcal{A}, \mathcal{N}^* \otimes \mathcal{U}) = H^2(\mathcal{A}, \mathcal{N}^* \otimes \mathcal{U}) = 0$, so that $H^1(X, U) \cong H^1(\mathcal{A}, \mathcal{U})$, then we have $\chi_U^1 = \chi^{0,1}$. In general, the characters χ_U^q can be expressed in terms of the ambient space characters $\chi^{p,q}$ by an obvious generalization of spectral sequence methods to characters.¹²

The multiplicity of representations can, in general, be computed from the scalar product between two characters χ and ψ defined as

$$(\chi, \psi) = \frac{1}{|\Gamma|} \sum_{g \in \Gamma} \chi(g) \bar{\psi}(g). \tag{B.2}$$

Under this scalar product, the characters χ_α of the irreducible Γ -representations form an orthonormal system, that is, $(\chi_\alpha, \chi_\beta) = \delta_{\alpha\beta}$. This means that the multiplicity n_α^q of the α^{th} irreducible representation in $H^q(X, U)$ can be extracted by

$$n_\alpha^q = (\chi_\alpha, \chi_U^q). \tag{B.3}$$

For Abelian groups Γ , the case considered in this paper, all irreducible representations are one-dimensional so that the graded cohomologies are simply given by these multiplicities, that is $h^q(X, U, R_\alpha) = n_\alpha^q$.

¹²In general, to get complete results, one needs to calculate the characters of images or kernels of maps in the spectral sequence. However, given the knowledge of the total cohomology dimensions $h^q(X, U)$, and the relative simplicity of our examples after imposing the relevant physical constraints, this is frequently unnecessary.

Open Access. This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References

- [1] P. Candelas, G.T. Horowitz, A. Strominger and E. Witten, *Vacuum configurations for superstrings*, *Nucl. Phys. B* **258** (1985) 46 [[INSPIRE](#)].
- [2] M.B. Green, J.H. Schwarz and E. Witten, *Superstring theory. Vol. 2: Loop amplitudes, anomalies and phenomenology*, *Cambridge Monographs On Mathematical Physics*, Cambridge University Press, Cambridge U.K. (1987), pg. 596.
- [3] 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](#)].
- [4] 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](#)].
- [5] 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](#)].
- [6] 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](#)].
- [7] 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](#)].
- [8] V. Braun, Y.-H. He, B.A. Ovrut and T. Pantev, *Vector bundle extensions, sheaf cohomology and the heterotic standard model*, *Adv. Theor. Math. Phys.* **10** (2006) 4 [[hep-th/0505041](#)] [[INSPIRE](#)].
- [9] J. Distler and B.R. Greene, *Aspects of (2,0) String Compactifications*, *Nucl. Phys. B* **304** (1988) 1 [[INSPIRE](#)].
- [10] S. Kachru, *Some three generation (0,2) Calabi-Yau models*, *Phys. Lett. B* **349** (1995) 76 [[hep-th/9501131](#)] [[INSPIRE](#)].
- [11] V. Bouchard and R. Donagi, *An SU(5) heterotic standard model*, *Phys. Lett. B* **633** (2006) 783 [[hep-th/0512149](#)] [[INSPIRE](#)].
- [12] 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](#)].
- [13] 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](#)].
- [14] 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](#)].
- [15] 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](#)].

- [16] L.B. Anderson, Y.-H. He and A. Lukas, *Heterotic compactification, an algorithmic approach*, *JHEP* **07** (2007) 049 [[hep-th/0702210](#)] [[INSPIRE](#)].
- [17] L.B. Anderson, Y.-H. He and A. Lukas, *Monad bundles in heterotic string compactifications*, *JHEP* **07** (2008) 104 [[arXiv:0805.2875](#)] [[INSPIRE](#)].
- [18] 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](#)].
- [19] 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](#)].
- [20] 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](#)].
- [21] O. Lebedev et al., *A mini-landscape of exact MSSM spectra in heterotic orbifolds*, *Phys. Lett. B* **645** (2007) 88 [[hep-th/0611095](#)] [[INSPIRE](#)].
- [22] 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](#)].
- [23] O. Lebedev et al., *The Heterotic Road to the MSSM with R parity*, *Phys. Rev. D* **77** (2008) 046013 [[arXiv:0708.2691](#)] [[INSPIRE](#)].
- [24] 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](#)].
- [25] S. Nibbelink Groot, 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](#)].
- [26] M. Blaszczyk et al., *A $Z2 \times Z2$ standard model*, *Phys. Lett. B* **683** (2010) 340 [[arXiv:0911.4905](#)] [[INSPIRE](#)].
- [27] M. Blaszczyk, S. Nibbelink Groot, F. Ruehle, M. Trapletti and P.K. Vaudrevange, *Heterotic MSSM on a Resolved Orbifold*, *JHEP* **09** (2010) 065 [[arXiv:1007.0203](#)] [[INSPIRE](#)].
- [28] R. Kappl et al., *String-Derived MSSM Vacua with Residual R Symmetries*, *Nucl. Phys. B* **847** (2011) 325 [[arXiv:1012.4574](#)] [[INSPIRE](#)].
- [29] 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](#)].
- [30] 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](#)].
- [31] G. Cleaver et al., *Investigation of Quasi-Realistic Heterotic String Models with Reduced Higgs Spectrum*, *Eur. Phys. J. C* **71** (2011) 1842 [[arXiv:1105.0447](#)] [[INSPIRE](#)].
- [32] M. Maio and A. Schellekens, *Permutation orbifolds of heterotic Gepner models*, *Nucl. Phys. B* **848** (2011) 594 [[arXiv:1102.5293](#)] [[INSPIRE](#)].
- [33] B. Gato-Rivera and A. Schellekens, *Heterotic Weight Lifting*, *Nucl. Phys. B* **828** (2010) 375 [[arXiv:0910.1526](#)] [[INSPIRE](#)].
- [34] B. Gato-Rivera and A. Schellekens, *Asymmetric Gepner Models II. Heterotic Weight Lifting*, *Nucl. Phys. B* **846** (2011) 429 [[arXiv:1009.1320](#)] [[INSPIRE](#)].

- [35] 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](#)].
- [36] J. Gray, Y.-H. He, A. Ilderton and A. Lukas, *STRINGVACUA: A Mathematica Package for Studying Vacuum Configurations in String Phenomenology*, *Comput. Phys. Commun.* **180** (2009) 107 [[arXiv:0801.1508](#)] [[INSPIRE](#)].
- [37] 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](#)].
- [38] 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](#)].
- [39] L.B. Anderson, Y.-H. He and A. Lukas, *Monad bundles in heterotic string compactifications*, *JHEP* **07** (2008) 104 [[arXiv:0805.2875](#)] [[INSPIRE](#)].
- [40] 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](#)].
- [41] L.B. Anderson, Y.-H. He and A. Lukas, *Heterotic compactification, an algorithmic approach*, *JHEP* **07** (2007) 049 [[hep-th/0702210](#)] [[INSPIRE](#)].
- [42] L.B. Anderson, J. Gray, D. Grayson, Y.-H. He and A. Lukas, *Yukawa Couplings in Heterotic Compactification*, *Commun. Math. Phys.* **297** (2010) 95 [[arXiv:0904.2186](#)] [[INSPIRE](#)].
- [43] J. Gray, *A Simple Introduction to Grobner Basis Methods in String Phenomenology*, *Adv. High Energy Phys.* **2011** (2011) 217035 [[arXiv:0901.1662](#)] [[INSPIRE](#)].
- [44] R. Blumenhagen, B. Jurke, T. Rahn and H. Roschy, *Cohomology of Line Bundles: A Computational Algorithm*, *J. Math. Phys.* **51** (2010) 103525 [[arXiv:1003.5217](#)] [[INSPIRE](#)].
- [45] T. Rahn, *Target Space Dualities of Heterotic Grand Unified Theories*, [arXiv:1111.0491](#) [[INSPIRE](#)].
- [46] The database of 400 heterotic line bundle standard models can be accessed at <http://www-thphys.physics.ox.ac.uk/projects/CalabiYau/linebundlemodels/index.html>.
- [47] M. Dine, N. Seiberg and E. Witten, *Fayet-Iliopoulos Terms in String Theory*, *Nucl. Phys. B* **289** (1987) 589 [[INSPIRE](#)].
- [48] A. Lukas and K. Stelle, *Heterotic anomaly cancellation in five-dimensions*, *JHEP* **01** (2000) 010 [[hep-th/9911156](#)] [[INSPIRE](#)].
- [49] 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](#)].
- [50] L.B. Anderson, J. Gray, A. Lukas and B. Ovrut, *Stability walls in heterotic theories*, *JHEP* **09** (2009) 026 [[arXiv:0905.1748](#)] [[INSPIRE](#)].
- [51] L.B. Anderson, J. Gray, A. Lukas and B. Ovrut, *The Edge Of Supersymmetry: Stability Walls in Heterotic Theory*, *Phys. Lett. B* **677** (2009) 190 [[arXiv:0903.5088](#)] [[INSPIRE](#)].
- [52] L.B. Anderson, J. Gray and B.A. Ovrut, *Transitions in the Web of Heterotic Vacua*, *Fortsch. Phys.* **59** (2011) 327 [[arXiv:1012.3179](#)] [[INSPIRE](#)].
- [53] E. Witten, *Strong coupling expansion of Calabi-Yau compactification*, *Nucl. Phys. B* **471** (1996) 135 [[hep-th/9602070](#)] [[INSPIRE](#)].

- [54] A. Lukas, B.A. Ovrut and D. Waldram, *Nonstandard embedding and five-branes in heterotic M-theory*, *Phys. Rev. D* **59** (1999) 106005 [[hep-th/9808101](#)] [[INSPIRE](#)].
- [55] 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](#)].
- [56] P. Candelas and X. de la Ossa, *Moduli space of Calabi-Yau manifolds*, *Nucl. Phys. B* **355** (1991) 455 [[INSPIRE](#)].
- [57] J. Wess and J. Bagger, *Supersymmetry and supergravity*, Princeton University Press, Princeton U.S.A. (1992), pg. 259.
- [58] R. Barbier et al., *R-parity violating supersymmetry*, *Phys. Rept.* **420** (2005) 1 [[hep-ph/0406039](#)] [[INSPIRE](#)].
- [59] N. Arkani-Hamed, L.J. Hall, H. Murayama, D. Tucker-Smith and N. Weiner, *Small neutrino masses from supersymmetry breaking*, *Phys. Rev. D* **64** (2001) 115011 [[hep-ph/0006312](#)] [[INSPIRE](#)].
- [60] G. Giudice and A. Masiero, *A Natural Solution to the mu Problem in Supergravity Theories*, *Phys. Lett. B* **206** (1988) 480 [[INSPIRE](#)].
- [61] S. Kobayashi, *Mathematical Society of Japan Publications. Vol. 15: Differential geometry of complex vector bundles*, Princeton University Press, Princeton U.S.A. (1987).
- [62] E.R. Sharpe, *Kähler cone substructure*, *Adv. Theor. Math. Phys.* **2** (1999) 1441 [[hep-th/9810064](#)] [[INSPIRE](#)].
- [63] L.B. Anderson, J. Gray and B. Ovrut, *Yukawa textures from heterotic stability walls*, *JHEP* **05** (2010) 086 [[arXiv:1001.2317](#)] [[INSPIRE](#)].
- [64] M. Kuriyama, H. Nakaajima and T. Watari, *Theoretical Framework for R-parity Violation*, *Phys. Rev. D* **79** (2009) 075002 [[arXiv:0802.2584](#)] [[INSPIRE](#)].
- [65] J. Li and S.-T. Yau, *The existence of supersymmetric string theory with torsion*, [hep-th/0411136](#) [[INSPIRE](#)].
- [66] P. Berglund et al., *On the instanton contributions to the masses and couplings of E_6 singlets*, *Nucl. Phys. B* **454** (1995) 127 [[hep-th/9505164](#)] [[INSPIRE](#)].
- [67] L.B. Anderson, J. Gray, A. Lukas and B. Ovrut, *Stabilizing the complex structure in heterotic Calabi-Yau vacua*, *JHEP* **02** (2011) 088 [[arXiv:1010.0255](#)] [[INSPIRE](#)].
- [68] 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](#)].
- [69] P. Candelas, A. Dale, C. Lütken and R. Schimmrigk, *Complete Intersection Calabi-Yau Manifolds*, *Nucl. Phys. B* **298** (1988) 493 [[INSPIRE](#)].
- [70] M. Gagnon and Q. Ho-Kim, *An exhaustive list of complete intersection Calabi-Yau manifolds*, *Mod. Phys. Lett. A* **9** (1994) 2235 [[INSPIRE](#)].
- [71] The CALABI-YAU Home Page, <http://www.th.physik.uni-bonn.de/th/Supplements/cy.html>.
- [72] V. Braun, *On free quotients of complete intersection Calabi-Yau manifolds*, *JHEP* **04** (2011) 005 [[arXiv:1003.3235](#)] [[INSPIRE](#)].
- [73] L.B. Anderson, *Heterotic and M-theory compactifications for string phenomenology*, [arXiv:0808.3621](#) [[INSPIRE](#)].

- [74] R. Friedman, *Algebraic Surfaces and Holomorphic Vector Bundles*, Springer-Verlag, Heidelberg Germany (1998).
- [75] R. Donagi, B.A. Ovrut, T. Pantev and R. Reinbacher, *SU(4) instantons on Calabi-Yau threefolds with $Z(2) \times Z(2)$ fundamental group*, *JHEP* **01** (2004) 022 [[hep-th/0307273](#)] [[INSPIRE](#)].
- [76] R. Hartshorne, *GTM 52: Algebraic Geometry*, Springer Verlag, Heidelberg Germany (1977).
- [77] P. Griffith and J. Harris, *Principles of algebraic geometry*, (1978).
- [78] T. Hubsch, *Calabi-Yau manifolds: A Bestiary for physicists*, first edition World Scientific, Singapore (1992), pg. 362 [second edition, World Scientific, Singapore (1994), pg. 374].
- [79] E. Dudas and E. Palti, *On hypercharge flux and exotics in F-theory GUTs*, *JHEP* **09** (2010) 013 [[arXiv:1007.1297](#)] [[INSPIRE](#)].
- [80] J. Marsano, *Hypercharge Flux, Exotics and Anomaly Cancellation in F-theory GUTs*, *Phys. Rev. Lett.* **106** (2011) 081601 [[arXiv:1011.2212](#)] [[INSPIRE](#)].
- [81] M.J. Dolan, J. Marsano and S. Schäfer-Nameki, *Unification and phenomenology of F-theory GUTs with $U(1)_{PQ}$* , *JHEP* **12** (2011) 032 [[arXiv:1109.4958](#)] [[INSPIRE](#)].
- [82] Y. Nir and N. Seiberg, *Should squarks be degenerate?*, *Phys. Lett. B* **309** (1993) 337 [[hep-ph/9304307](#)] [[INSPIRE](#)].