Predicting the Orbifold Origin of the MSSM

Erik Parr, Patrick K.S. Vaudrevange,* and Martin Wimmer

MSSM-like string models from the compactification of the heterotic string on toroidal orbifolds (of the kind $\mathbb{T}^6/P$) have distinct phenomenological properties, like the spectrum of vector-like exotics, the scale of supersymmetry breaking, and the existence of non-Abelian flavor symmetries. We show that these characteristics depend crucially on the choice of the underlying orbifold point group P. In detail, we use boosted decision trees to predict P from phenomenological properties of MSSM-like orbifold models. As this works astonishingly well, we can utilize machine learning to predict the orbifold origin of the MSSM.

1. Introduction

String theory compactified to four-dimensional space-time naturally provides a unified framework for quantum gravity and gauge interactions with chiral matter. This fact raises the obvious question whether string theory can incorporate the Standard Model (SM) of particle physics (or its Minimal Supersymmetric Extension, the MSSM). A definite answer to this question would be given by an explicit construction of a string compactification that is in agreement with all experimental facts from particle physics (and, if one is even more ambitious, with all cosmological observations). However, due to the enormous number of four-dimensional string models[1,2] and the computational complexity[3] a naive search in the string landscape for the MSSM is very likely to fail. New methods seem to be unavoidable to narrow down the string landscape towards realistic particle physics.

In recent years, big data and machine learning (ML) has entered the field of strings.[4–9] Various tasks have been addressed, for example, to identify the topological structure of the string landscape using persistent homology,[10] to predict the Hodge numbers of complete intersection Calabi-Yau manifolds (CICYs) with large $h^{1,1}$,[11] to find consistent type IIA D6-brane configurations that yield MSSM-like models using Deep

Reinforcement Learning,[12] to identify whether a given CICY is elliptically fibered or not,[13] to explore the landscape of type IIB flux vacua using genetic algorithms,[14] to find numerical metrics of Calabi-Yau manifolds by combining conventional curve fitting and techniques from supervised learning,[15] and to approximate Kähler metrics for type IIB Calabi-Yau compactifications using generative adversarial networks (GANs).[16] Hence, encouraged by these results, techniques from big data and ML are expected to yield new insights into the string landscape.

In the context of heterotic orbifolds,[17–19] the Mini-Landscape of $\mathbb{Z}_6$-II orbifold models has been a test ground for model searches: First, individual models have been identified and analyzed.[20–23] Then, larger scans have been performed.[24–27] Finally, new methods from ML have been applied to this test ground, like an autoencoder neural network to automatically identify fertile islands in the $\mathbb{Z}_6$-II Mini-Landscape[28] and techniques from contrast data mining to reduce the landscape by extracting new features of orbifold models that correlate with their phenomenological property of being MSSM-like.[29] However, there are in total 138 Abelian orbifolds with $\mathcal{N} = 1$ supersymmetry.[30] So, the heterotic orbifold landscape is much wider than $\mathbb{Z}_6$-II. Consequently, how do we know that we will find the most promising models in the $\mathbb{Z}_6$-II region of the landscape? Actually, can ML algorithms predict the orbifold geometry which most likely reproduces a certain MSSM-like model? As shown in this paper using a boosted decision tree, the answer to this question seems to be positive.

This paper is organized as follows: Section 2 begins with a brief review of the Orbifold-Landscape of all known MSSM-like orbifold models constructed so far. In addition, motivated by some generic properties of these models, we define phenomenological features that characterize MSSM-like models in general. Then, in Section 3 we discuss boosted decision trees and present the resulting predictions for the orbifold origin of the MSSM in Section 4. Finally, Section 5 gives conclusions and outlook.

E. Parr, P. K. Vaudrevange, M. Wimmer
Physik Department T75
Technische Universität München
James-Franck-Straße 1, Garching 85748, Germany
E-mail: patrick.vaudrevange@tum.de

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/prop.202000032

2. Phenomenology of the Orbifold-Landscape

The aim of this paper is to construct a machine learning (ML) model that predicts the orbifold origin of MSSM-like bottom-up models. In more detail, motivated by the generic features of MSSM-like string models in the Orbifold-Landscape, we train an ML model to predict the orbifold point group that has the highest probability to reproduce a given MSSM-like model, see **Figure 1**.
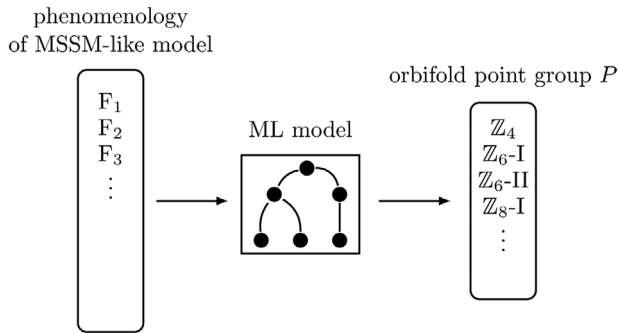
ADVANCED
SCIENCE NEWS

www.advancedsciencenews.com

Fortschritte
der Physik
Progress
of Physics

www.fp-journal.org

Figure 1. We use a machine learning (ML) model to predict the orbifold point group P that most likely is able to reproduce in string theory a given MSSM-like bottom-up model, which we specify by some phenomenological features F_i.

To train such an ML model, we need a large dataset of MSSM-like orbifold models based on various different orbifold point groups. Then, we have to define and compute some phenomenological features that yield a basic characterization of MSSM-like models. These features are taken by the ML algorithm as input, while the output of the ML algorithm is the prediction of the corresponding orbifold point group. Thus, in the following we first discuss our dataset of MSSM-like orbifold models and, afterwards, we define our phenomenological features.

2.1. The Dataset of MSSM-Like Orbifold Models

For the analysis we use a large dataset of 126,783 inequivalent MSSM-like string models originating from the $E_8 \times E_8$ heterotic string, compactified on various orbifolds $\mathbb{O}$ with Abelian point group P (e.g. in the case without roto-translations, $\mathbb{O} = \mathbb{T}^6/P$, where P is either $\mathbb{Z}_N$ or $\mathbb{Z}_N \times \mathbb{Z}_M$, see Ref. [30]). This dataset is obtained as follows: First, we take the MSSM-like orbifold models from the searches performed in Refs. [26, 27, 29]. Then, we construct new $\mathbb{Z}_N \times \mathbb{Z}_M$ MSSM-like models using the enhanced search algorithm based on contrast patterns.[29] Finally, these individual datasets are merged using the orbifolder[31] resulting in 126,783 inequivalent MSSM-like orbifold models.[1] In the first row of Table A.1 we list the orbifold point groups P that yield MSSM-like string models and the second row gives the number of inequivalent models based on the respective point group. Note that from the point of view of data analysis, our dataset is highly imbalanced: Some point groups (like $\mathbb{Z}_4$) give rise to only a few hundred MSSM-like orbifold models, while others (like $\mathbb{Z}_2 \times \mathbb{Z}_4$) yield several ten-thousand MSSM-like models. Moreover, since we only have one MSSM-like orbifold model based on the $\mathbb{Z}_7$ point group, see Refs. [27, 33], we decide to leave out the $\mathbb{Z}_7$ point group from our prediction task.

In the following, we analyze our dataset of MSSM-like orbifold models for their generic phenomenological properties and, by doing so, we identify some universal features of MSSM-like models.

2.2. SM Singlets

MSSM-like orbifold models typically yield $\mathcal{O}(100)$, i.e. on average 146, SM singlets s^0 with SM charges $(\mathbf{1}, \mathbf{1})_0$ (see the eighth row in Tables A.1 and A.2). Generically, they are charged under several hidden U(1) factors and sometimes even under a non-Abelian hidden sector gauge group G_{hidden}. The existence of these SM singlets has several implications: As we will discuss in detail in Sections 2.3 and 2.7, they can acquire non-vanishing vacuum expectation values (VEVs) $\langle s^0 \rangle \neq 0$ without breaking supersymmetry, i.e. $F = D = 0$. Moreover, interpreting these SM singlets as right-handed neutrinos, they can give rise to a seesaw mechanism with a model-dependent seesaw scale that is typically somewhat below the string scale.[34] Since, their presence is so generic in MSSM-like string models from heterotic orbifolds, we include their number into our feature space.

2.3. Vector-Like Exotics

Beside three (chiral) generations of quarks and leptons, a Higgs-pair and SM singlets, MSSM-like orbifold models are generically equipped with vector-like exotics (i.e. matter that is charged under the SM and has a mass-partner transforming in the complex conjugate representation with respect to the SM). All types of vector-like exotics that appear in the Orbifold-Landscape are listed in Tables A.1 and A.2 in appendix A. Since they appear so frequently, we choose the numbers of vector-like exotics of all kinds as features (except for the vector-like exotics that only appear in the unique MSSM-like $\mathbb{Z}_7$ orbifold model).

The MSSM itself is defined without any vector-like exotics beside the Higgs-pair. So, what are the phenomenological consequences if vector-like exotics are present? Let us denote a pair of vector-like exotics by X and $\bar{X}$ in the following. In many cases, these exotics can become very massive through terms in the superpotential of the form $\mathcal{W} \supset (M_{\text{Planck}})^{1-p} (s^0)^p X \bar{X}$. Here, $p \in \mathbb{N}_+$ and s^0 denotes an MSSM singlet $(\mathbf{1}, \mathbf{1})_0$ that can acquire a large non-vanishing VEV. This would render the vector-like exotics X and $\bar{X}$ massive, with a mass that can be close to the Planck scale M_{Planck} depending on the size of $\langle s^0 \rangle$ and $p \in \mathbb{N}_+$. Still, one can argue that the presence of vector-like exotics might be a virtue or a problem: On the one hand, new elementary particles with spin 0 or 1 (called leptoquarks) have been proposed, especially to address some flavor anomalies, see for example Refs. [35–38] and references therein. In this scenario, the leptoquark has to be light compared to the Planck scale. On the other hand, the presence of vector-like exotics can affect gauge coupling unification and it can yield severe cosmological problems. Thus, we decide to look for "(almost) perfect MSSM-like models" that have no (or a minimal amount of) vector-like exotics. In summary, our basic features correspond to the numbers of all types of vector-like exotics that appear in the Orbifold-Landscape.

As discussed next, we extend our feature space by some additional properties of MSSM-like orbifold models, in order to obtain a more complex feature space. Moreover, these additional features give a notion of MSSM-like models that are more promising from a phenomenological point of view.

---

[1] The MSSM-like orbifold models can be found as arXiv ancillary files, see Refs. [29] and [32].

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

Fortschritte
der Physik
Progress
of Physics

www.fp-journal.org

2.4. Heavy Top Quark from Bulk

There is a large hierarchy between the top quark mass and the quark masses of the first and second generation. In the MSSM, it is explained by a renormalizable coupling

$$\mathcal{W} \supset \gamma_{ij}\, q_i\, \bar{u}_j\, h_u \tag{1}$$

in the superpotential $\mathcal{W}$. To first approximation, the matrix of Yukawa couplings y_{ij} has to have rank one in order to define the large top quark mass.

In string theory, the top Yukawa coupling can be related to the ten-dimensional gauge coupling constant.[39–42] In more detail, from a ten-dimensional perspective Equation (1) originates from a supersymmetric $E_8 \times E_8$ gauge interaction in ten-dimensions. Then, compactifying to four dimensions, the coupling Equation (1) is present if the left-chiral top quark doublet q_3, its right-chiral top quark partner $\bar{u}_3$ and the up-type Higgs h_u distribute among the three so-called untwisted sectors U_a, for $a = 1, 2, 3$, respectively. Here, the untwisted sector U_a, where $a = 1, 2, 3$, is defined by the (complexified) internal component $A^a \sim A^{2+2a} + iA^{3+2a}$ of the ten-dimensional $E_8 \times E_8$ gauge bosons A^M, $M = 0, \ldots, 9$. This mechanism gives an appealing explanation for the large hierarchy in the up-type quark masses and, hence, we append our feature space by a feature "heavy top from bulk".

2.5. Non-Abelian Flavor Symmetries from Vanishing Wilson Lines

Non-Abelian flavor symmetries of the four-dimensional effective theory can emerge in heterotic orbifolds from the localization of certain strings in the extra-dimensions of the orbifold. In more detail, string interactions are constrained by so-called string selection rules that describe the ability of strings to split, stretch and join while they propagate on the surface of the orbifold.[43,44] These constraints can be formulated in terms of Abelian discrete symmetries.[45] If certain background fields (i.e. Wilson lines[19]) vanish, an additional permutation symmetry of some of the localized strings can emerge such that the full flavor symmetry becomes non-Abelian.[20,46] In the heterotic orbifold construction, there are two main types of non-Abelian flavor groups for MSSM-like orbifold models[27] (being $\Delta(54)$ and D_8, where D_8 denotes the dihedral group of order 8, sometimes also denoted by D_4). Since non-Abelian flavor symmetries are phenomenologically appealing[47] and related to a vanishing Wilson line in the context of heterotic orbifolds, we extend our feature space by the number of vanishing Wilson lines, see also Refs. [26, 27]. In other words, a non-zero value of the feature "# vanishing Wilson lines" signals the presence of a non-Abelian flavor symmetry.

2.6. Hidden Sector Gaugino Condensation and Supersymmetry Breaking

Supersymmetry breaking through hidden sector gaugino condensation is correlated to the hidden sector gauge group and its light matter content.[48–51] The $E_8 \times E_8$ heterotic string is especially suitable for this mechanism, as it contains, beside the observable E_8 that hosts the MSSM, a hidden E_8 factor, which generically yields supersymmetry breaking at low energies, see Refs. [26, 52] and also [53, 54].

For each MSSM-like orbifold model, we compute for each hidden sector non-Abelian gauge group factor G_{hidden} the chiral part of the spectrum with respect to $SU(3)_C \times SU(2)_L \times U(1)_Y \times G_{\text{hidden}}$ (assuming that the vector-like part decouples). Then, the resulting beta-function coefficient is given by $b = 3C_2 - \sum_r \ell(\boldsymbol{r})$ of G_{hidden}. Here, the summation is performed over chiral matter transforming in a representation $\boldsymbol{r}$ of G_{hidden}. Furthermore, we have $C_2 = N$ and $\ell(\boldsymbol{N}) = 1/2$ for $SU(N)$, while we get $C_2 = 2(N-1)$ and $\ell(2\boldsymbol{N}) = 1$ for $SO(2N)$. Then, the gauge coupling $g_{\text{hidden}}(\mu)$ of G_{hidden} depends on the energy scale μ, where b determines the one-loop RGE, being

$$\frac{\partial g}{\partial \ln \mu} = -b \frac{g^3}{16\pi^2}. \tag{2}$$

After solving this differential equation, one can compute the energy scale Λ at which the coupling $g_{\text{hidden}}(\mu)$ diverges, i.e. when $1/g_{\text{hidden}}(\Lambda) \to 0$. It is given by

$$\Lambda = M_{\text{GUT}} \exp\left(-\frac{16\pi^2}{2\, b\, g^2(M_{\text{GUT}})}\right), \tag{3}$$

with $M_{\text{GUT}} \approx 3 \cdot 10^{16}\,\text{GeV}$ and $g^2(M_{\text{GUT}}) \approx 1/2$. Furthermore, we assume in Equation (3) that the gauge coupling constants of the MSSM coincide at the GUT scale M_{GUT} approximately with the one of the hidden sector gauge group and we neglect string threshold corrections.[55–57] Then, at the energy scale Λ, the gauginos of G_{hidden} form condensates and, consequently, supersymmetry is broken spontaneously by a non-vanishing F-term of the dilaton. Assuming dilaton stabilization by non-perturbative effects and gravity mediation to the observable sector, the gravitino mass can be estimated as

$$m_{3/2} \approx \frac{\Lambda^3}{M_{\text{Planck}}^2}. \tag{4}$$

Thus, the feature "hidden sector beta-function", specified by the coefficient b, gives a rough estimate of the scale of supersymmetry breaking.

In our MSSM-like orbifold models, the distribution of Λ is given in **Figure 2** for all point groups and in **Figure 3** for $\mathbb{Z}_6$-II only. We are interested in the case of supersymmetry breaking around (at least) the TeV scale. A very rough estimate gives the constraint $\Lambda \gtrsim 10^{13\pm1}$ GeV,[52] which translates to $b \gtrsim 15$.

As a remark, the beta-function coefficient b is closely related to the number of unbroken roots N_{ur}, i.e. the number of roots from the hidden E_8 factor that survives the orbifold projection conditions. Interestingly, the contrast patterns developed in Ref. [29] showed that large values of N_{ur} correlate with a higher production rate of MSSM-like orbifold models. Moreover, there are many MSSM-like orbifold models with large values of N_{ur} that are (practically) inaccessible by traditional search algorithms and were uncovered recently in Ref. [29].
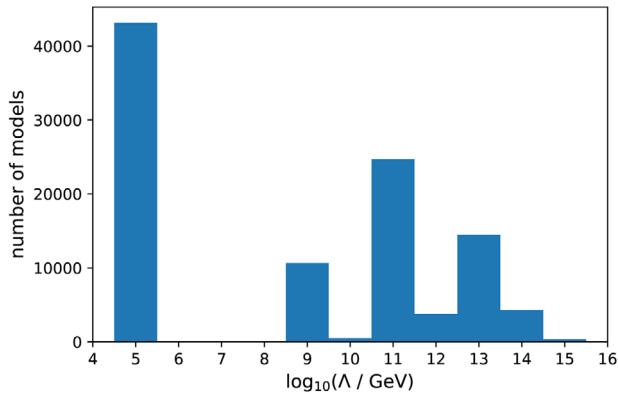
ADVANCED
SCIENCE NEWS
www.advancedsciencenews.com

Fortschritte
der Physik
Progress
of Physics

www.fp-journal.org

Figure 2. Hidden sector gaugino condensation scale Λ for MSSM-like orbifold models based on all point groups. Note that there are a few models with even smaller Λ (and models with $b = 0$ are excluded). Moreover, the peak at $\Lambda = 10^5$ GeV consists mainly of MSSM-like orbifold models with point groups $\mathbb{Z}_2 \times \mathbb{Z}_4$ and $\mathbb{Z}_4 \times \mathbb{Z}_4$ (each contributing $\approx 19,000$ models). For these models, the scale of supersymmetry breaking is far too low.
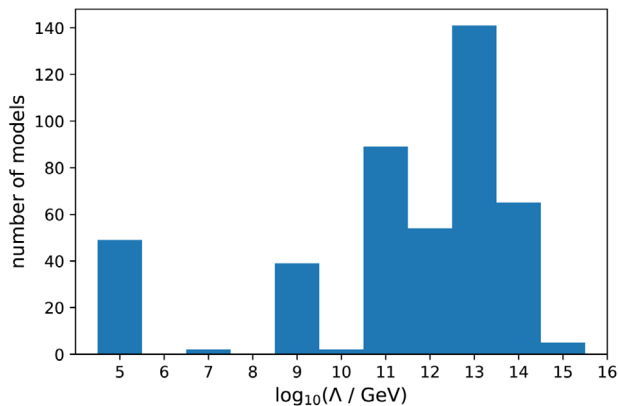

Figure 3. Hidden sector gaugino condensation scale Λ for MSSM-like orbifold models based on the $\mathbb{Z}_6$-II $(1,1)$ orbifold geometry (models with $b = 0$ are excluded), see [52].

2.7. Anomalous U(1)

99% of the MSSM-like orbifold models have an additional $U(1)_{\text{anom}}$ gauge factor beside hypercharge that appears to be anomalous, where the anomaly is canceled by a universal Green–Schwarz mechanism.[58] This induces a Fayet–Iliopoulos D-term (FI-term), which sets the scale for a Froggatt–Nielsen-like mechanism,[59] where some SM singlets s^0 develop non-vanishing VEVs $\langle s^0 \rangle$ in order to satisfy $D_{\text{anom}} = 0$. Consequently, these VEVs spontaneously break the additional U(1) factors, generate masses of the lighter quarks and leptons, and decouple (at least some of) the vector-like exotics, see Section 2.3. Hence, the existence of an "anomalous U(1)" is a good feature that characterizes promising orbifold models.

2.8. Comments on Our Feature Space

This concludes our feature space. Let us remark that our dataset of 126,783 inequivalent MSSM-like orbifold models corresponds

to a set of 106,027 inequivalent feature vectors, which we call $\mathbb{D}$. The main reason for the decrease from 126,783 to 106,027 is associated to the hidden sector: In order to distinguish between inequivalent MSSM-like orbifold models, the observable and the hidden sector are taken into account.[31] On the other hand, the feature vector is supposed to characterize only the properties of a given model with respect to the MSSM. Thus, models with different hidden sectors can yield the same feature vector. Taking this into account, our features give a good measure to distinguish between inequivalent phenomenological properties of MSSM-like orbifold models.

3. Boosted Decision Tree

A boosted decision tree is built up by an ensemble of single decision trees. Each decision tree is a so-called weak learner: on its own, it typically yields a poor performance. However, a weighted majority vote of many weak learners tends to perform much better.

In more detail, the idea of boosting is that one combines the predictions of many weak learners (e.g. decision trees) to get a more powerful estimator. This is achieved by a successive training of many weak learners, where the misclassified training data of the previous weak learner is weighted with a higher value for the next weak learner in order to enforce him to classify these data points correctly (and the weights of the data that has been classified correctly are decreased accordingly). In our case, this procedure is repeated 1,500 times. Finally, we combine the individual decision trees to a much more powerful estimator: the boosted decision tree. For further details on boosting, see for example Ref. [60].

3.1. How to Measure the Performance of ML Models for Imbalanced Datasets

The performance of a predictive ML model can be measured by the accuracy that is defined by the number of correct predictions divided by the total number of all predictions. However, for an imbalanced dataset, as the one we have in our case, this normal accuracy measurement can be misled in the following way: Assume a classification task with two classes A and B, where the class A builds the majority of the dataset with 99% of all instances. Any prediction method can now achieve an accuracy of 99% simply by predicting class A always, but never class B.

In order to avoid such a behavior, there exist several metrics for imbalanced classification tasks. The one we are using is based on three types of predictions involving the point group P_i, i.e. for each point group P_i we define

$$\text{\# of true positives: } \text{TP}_i := \left| \{ X \in \mathbb{D} \mid Y_{\text{correct}}(X) = P_i \text{ and} \right.$$

$$\left. Y_{\text{predicted}}(X) = P_i \} \right|,$$

$$\text{\# of false positives: } \text{FP}_i := \left| \{ X \in \mathbb{D} \mid Y_{\text{correct}}(X) = P_j \text{ but} \right.$$

$$\left. Y_{\text{predicted}}(X) = P_i \} \right|,$$

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

Fortschritte
der Physik
**Progress
of Physics**

www.fp-journal.org

# of false negatives: $\mathrm{FN}_i := \Big|\{X \in \mathbb{D} \mid Y_{\mathrm{correct}}(X) = P_i$ but

$$Y_{\mathrm{predicted}}(X) = P_j\}\Big|.$$

Here, the indices $i \neq j$ label the 14 different orbifold point groups P_i, X is a feature vector from our dataset $\mathbb{D}$, and $Y_{\mathrm{correct}}(X)$ ($Y_{\mathrm{predicted}}(X)$) label the correct (predicted) point group of X, respectively. Counting for each point group P_i the numbers TP_i, FP_i and FN_i allows us to define three different metrics,

$$\text{precision:} \quad \mathrm{p}_i := \frac{\mathrm{TP}_i}{\mathrm{TP}_i + \mathrm{FP}_i}, \tag{5a}$$

$$\text{recall:} \quad \mathrm{r}_i := \frac{\mathrm{TP}_i}{\mathrm{TP}_i + \mathrm{FN}_i}, \tag{5b}$$

$$\text{f1-score:} \quad \mathrm{f1}_i := 2\frac{\mathrm{p}_i \, \mathrm{r}_i}{\mathrm{p}_i + \mathrm{r}_i}. \tag{5c}$$

Finally, one can define the "f1-macro" as the average of the f1-scores for all point groups. Then, to deal with our highly imbalanced data, we use the f1-macro to measure the performance of our ML model.

Let us briefly illustrate the benefit of using the f1-macro on the example from the beginning of this section. In this case, we have $\mathrm{f1}_A \approx 1$ and $\mathrm{f1}_B = 0$ (for the extreme case of predicting always A and never B, the precision value of B is actually undefined. However, one can easily rewrite the f1-score directly in terms of $\mathrm{TP}_B = \mathrm{FP}_B = 0$ and $\mathrm{FN}_B \neq 0$ in order to see that $\mathrm{f1}_B = 0$). Consequently, the f1-macro is given by $1/2(\mathrm{f1}_A + \mathrm{f1}_B) \approx 0.5$, which rates the naive classification as insufficient.

3.2. Training of the ML Model

We start by splitting our dataset $\mathbb{D}$ and the corresponding target values Y_{correct}, i.e. the orbifold point groups P_i, of our 106,027 feature vectors into 80% training and 20% test data. The test set $\mathbb{T}$ is held back for the evaluation of the trained ML model. On the other hand, the training set is used to perform a grid search for the optimal hyperparameters, i.e. each set of hyperparameters in the grid is used to train an ML model using 5-fold cross-validation (CV). Then, the best ML model is chosen based on the f1-macro. In more detail, the module `GridSearchCV` from the scikit-learn library[61] is utilized for hyperparameter search and the LightGBM implementation[62] for boosted decision trees.

After an extensive hyperparameter search, it turns out that the following hyperparameters give the best f1-macro performance:

- `class_weight=‘balanced’`: The argument "balanced" weights the 14 classes of orbifold point groups inversely proportional to their frequency of occurrence in the input dataset at the beginning of the training. In detail, the dominant classes (like $\mathbb{Z}_2 \times \mathbb{Z}_4$ and $\mathbb{Z}_4 \times \mathbb{Z}_4$) get weights smaller than 1 and the small classes (like $\mathbb{Z}_4$) get weights larger than 1.
- `learning_rate=0.2`

- `min_child_samples=8`: The minimal number of samples per child node, where the default value is 20. This is beneficial since there are nodes with a small number of samples.
- `min_child_weight=0.01`: A regularization measure to stop splitting a node if its purity is high.
- `n_estimators=1500`: The number of individual decision trees.
- `num_leaves=50`: The maximum number of leaf nodes for each decision tree.

To evaluate the predictive power of the final ML model with optimized hyperparameters, we use the test set $\mathbb{T}$. The results of this evaluation will be discussed in Section 3.3.

In addition to a boosted decision tree (i.e. LightGBM), we also tried various alternative classification algorithms. To be specific: k-nearest-neighbors, linear and quadratic discriminant analysis, logistic regression, random forest, support vector machines, and fully connected neural networks with softmax classification.[2] Non of those alternative ML models individually performed on the level of LightGBM. Only XGBoost,[64] a different implementation of boosted decision trees, yields comparable results. In addition, we build an ensemble of these different estimators, where the prediction of the ensemble is a weighted linear combination of the predictions of each individual estimator. However only a combination of LightGBM with different neural network architectures shows a slight improvement, where the f1-macro increases from 0.856 for LightGBM as a single classifier to 0.867 for the ensemble. Since this improvement is small and our main results do not change, we decide to keep things simple and use LightGBM as a single estimator only. In addition, the usage of LightGBM as a single estimator allows us to read out and interpret the inner structure of the boosted decision tree by visualizing the feature importance, see **Figure 4**.

3.3. Performance of the ML Model

Next, we analyze the performance of our trained ML model on the test set $\mathbb{T}$. This is quantified in a classification report, see **Table 1**. We see that for all orbifold point groups (even for the minority classes like $\mathbb{Z}_4$ and $\mathbb{Z}_6$-I) the f1-score is very high. In addition, we also state the confusion matrix in **Table 2**. These results indicate that our boosted decision tree is well balanced. It is intriguing how well our ML model can predict the orbifold point group using only the spectrum of vector-like exotics and some additional phenomenologically appealing features. Note also that the MSSM-like orbifold models from the training set are only unique among a certain orbifold point group. In some cases, MSSM-like orbifold models from different point groups yield similar (or even identical) feature vectors. Then, the classifier has to decide to which orbifold point group this part of the feature space consists more likely. This introduces some intrinsic uncertainty to our ML model and some misclassifications are unavoidable.

---

2 The implementations from scikit-learn[61] are used for the non-neural network algorithms, whereas the Keras API[63] is used for the neural networks.
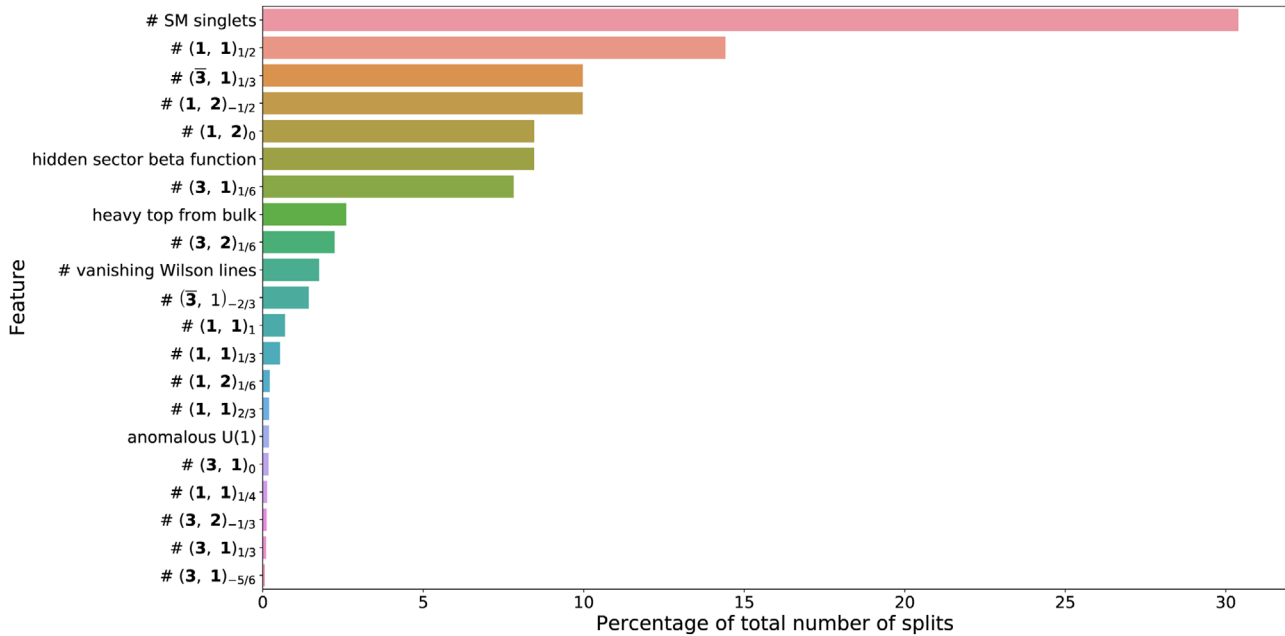
ADVANCED
SCIENCE NEWS
www.advancedsciencenews.com

Fortschritte
der Physik
Progress
of Physics
www.fp-journal.org

Figure 4. The importance of the features (defined in Section 2) for the classification task performed by our boosted decision tree: For each decision tree of the ensemble of trees, we count the number how often each feature was used in the nodes of the tree. Then, for each feature we take the average over all trees. Finally, we give the numbers as percentage and rank the features accordingly. As a result, the feature "# SM singlets" turns out to be used in 30% of the nodes and is the most important feature for our boosted decision tree.

Table 1. Classification report of our boosted decision tree, evaluated on the test set $\mathbb{T} \subset \mathbb{D}$ of 21,206 feature vectors of MSSM-like orbifold models, see Equation (5) for definitions. The f1-macro (i.e. the average of f1-scores) is 0.86. The support for each point group P_i is defined as $|\{X \in \mathbb{T} \mid Y_{\text{correct}}(X) = P_i\}|$. The best performance is achieved for MSSM-like orbifold models with point groups $\mathbb{Z}_3 \times \mathbb{Z}_3$, $\mathbb{Z}_4 \times \mathbb{Z}_4$ and $\mathbb{Z}_2 \times \mathbb{Z}_4$.

Point group	Precision	Recall	f1-score	
P_i	p_i	r_i	$f1_i$	Support
$\mathbb{Z}_4$	0.85	0.85	0.85	33
$\mathbb{Z}_6$-I	0.88	0.88	0.88	8
$\mathbb{Z}_6$-II	0.82	0.75	0.79	305
$\mathbb{Z}_8$-I	0.75	0.78	0.76	125
$\mathbb{Z}_8$-II	0.86	0.84	0.85	444
$\mathbb{Z}_{12}$-I	0.95	0.90	0.92	232
$\mathbb{Z}_{12}$-II	0.69	0.68	0.69	82
$\mathbb{Z}_2 \times \mathbb{Z}_2$	0.88	0.83	0.85	255
$\mathbb{Z}_2 \times \mathbb{Z}_4$	0.97	0.98	0.98	8,155
$\mathbb{Z}_2 \times \mathbb{Z}_6$-I	0.84	0.59	0.69	211
$\mathbb{Z}_3 \times \mathbb{Z}_3$	0.99	1.00	0.99	483
$\mathbb{Z}_3 \times \mathbb{Z}_6$	0.87	0.88	0.87	1,019
$\mathbb{Z}_4 \times \mathbb{Z}_4$	0.98	0.98	0.98	9,113
$\mathbb{Z}_6 \times \mathbb{Z}_6$	0.93	0.96	0.94	741
average	0.88	0.85	0.86	21,206

4. Predicting the Stringy Origin of the MSSM

After we have shown that the predictions of our classifier are reasonable for the given feature vectors of MSSM-like orbifold models, the next step is to extrapolate from these results towards the MSSM: we give a feature vector without any vector-like exotics to the ML algorithm in order to identify its most likely origin from orbifold compactifications. Extrapolations with ML models are in general rather difficult: we try to make a prediction for a feature vector that is different to the data from $\mathbb{D}$. However, since we use additional features beside the numbers of vector-like exotics (like the number of SM singlets and the hidden sector beta-function coefficient, see Section 2), the prediction for the MSSM also includes non-trivial features that are embedded in the orbifold dataset. In other words, the feature vector of the MSSM gets closer to the feature vectors of our MSSM-like orbifold models by using these additional features. Moreover, our experience with regularized decision trees (see Ref. [29]) indicates that they are suitable for extrapolations: In general, each decision tree divides up the feature space into smaller subspaces and assigns a prediction (i.e. an orbifold point group) to each subspace. However, a decision tree necessarily leaves subspaces of infinite volume at the boundary of the training set. Since the feature vector of the MSSM is outside the region of feature vectors of MSSM-like orbifold models, the MSSM will lie in one of these infinite subspaces, but still gets a meaningful prediction assigned. Then, an ensemble of decision trees gives an additional regularization that improves the generalization to a wider range in feature space. In this way, boosted decision trees can be used to get meaningful predictions, even for extrapolations.

In detail, in order to predict the orbifold origin of the MSSM, we generate a feature vector for each value of the hidden sector beta-function coefficient $b \in \{0, \ldots, 36\}$ (i.e. between the minimal and maximal value of b in our dataset $\mathbb{D}$) and for each number of Standard Model singlets: # SM singlets $\in \{20, \ldots, 350\}$. In this

ADVANCED
SCIENCE NEWS

www.advancedsciencenews.com

Fortschritte
der Physik
Progress
of Physics

www.fp-journal.org

Table 2. Confusion matrix of our boosted decision tree, evaluated on the test set $\mathbb{T} \subset \mathbb{D}$ of 21,206 feature vectors of MSSM-like orbifold models.

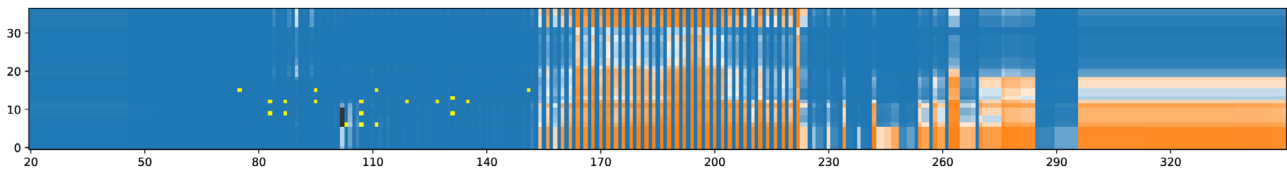|		Predicted orbifold point group													
		$\mathbb{Z}_4$	$\mathbb{Z}_6$-I	$\mathbb{Z}_6$-II	$\mathbb{Z}_8$-I	$\mathbb{Z}_8$-II	$\mathbb{Z}_{12}$-I	$\mathbb{Z}_{12}$-II	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$\mathbb{Z}_2 \times \mathbb{Z}_6$-I	$\mathbb{Z}_3 \times \mathbb{Z}_3$	$\mathbb{Z}_3 \times \mathbb{Z}_6$	$\mathbb{Z}_4 \times \mathbb{Z}_4$	$\mathbb{Z}_6 \times \mathbb{Z}_6$
Correct orbifold	$\mathbb{Z}_4$	28	0	0	1	0	0	0	0	4	0	0	0	0	0
point group	$\mathbb{Z}_6$-I	0	7	0	0	0	0	0	0	0	0	0	0	1	0
	$\mathbb{Z}_6$-II	0	0	229	0	14	0	4	0	55	2	1	0	0	0
	$\mathbb{Z}_8$-I	1	0	0	97	1	2	0	0	0	0	0	1	23	0
	$\mathbb{Z}_8$-II	0	0	15	0	375	0	10	2	41	0	0	0	1	0
	$\mathbb{Z}_{12}$-I	0	0	0	4	0	209	0	0	0	0	2	0	17	0
	$\mathbb{Z}_{12}$-II	0	0	4	0	10	0	56	0	12	0	0	0	0	0
	$\mathbb{Z}_2 \times \mathbb{Z}_2$	0	0	0	0	2	0	0	212	41	0	0	0	0	0
	$\mathbb{Z}_2 \times \mathbb{Z}_4$	3	0	24	4	31	1	7	27	8,004	21	0	5	26	2
	$\mathbb{Z}_2 \times \mathbb{Z}_6$-I	0	0	2	1	3	1	3	0	69	124	0	2	4	2
	$\mathbb{Z}_3 \times \mathbb{Z}_3$	0	0	0	0	0	2	0	0	0	0	481	0	0	0
	$\mathbb{Z}_3 \times \mathbb{Z}_6$	0	0	1	0	2	0	0	0	0	0	0	899	93	24
	$\mathbb{Z}_4 \times \mathbb{Z}_4$	1	1	3	22	0	5	1	0	21	1	0	116	8,914	28
	$\mathbb{Z}_6 \times \mathbb{Z}_6$	0	0	0	0	0	0	0	0	0	0	0	16	12	713

Figure 5. Stringy origin of the MSSM with "heavy top from bulk" but without a non-Abelian flavor symmetry. The colors are associated to point groups as follows: blue for $\mathbb{Z}_2 \times \mathbb{Z}_4$, orange for $\mathbb{Z}_4 \times \mathbb{Z}_4$ and black for $\mathbb{Z}_8$-I. In yellow, we highlight 20 MSSM-like orbifold models that have no vector-like exotics beside some Higgs-pairs, see Section 4.1. See also Figure 6 for additional cases and further details.

way, we obtain $37 \cdot 331$ feature vectors that we give to our trained ensemble of decision trees in order to obtain a prediction for each of them. We can illustrate the results in a two–dimensional plot, where different colors correspond to different orbifold point groups and the transparency of the color indicates the degree of accuracy of the corresponding prediction, see **Figure 5** and the caption of **Figure 6** for further details.

Furthermore, we turn on and off the properties of non-Abelian flavor symmetries (corresponding to the "# vanishing Wilson lines") as well as "heavy top from bulk". In this way, we obtain three additional cases and we display for each case the predictions of the $37 \cdot 331$ feature vectors in Figure 6. It turns out that in general the predictions are dominated by two classes: $\mathbb{Z}_2 \times \mathbb{Z}_4$ and $\mathbb{Z}_4 \times \mathbb{Z}_4$. Interestingly, these classes build up the majority of MSSM-like orbifold models with 55,429 and 48,812 MSSM-like models, respectively, and they achieve the highest f1-scores, see Table 1. In addition, for MSSMs without a heavy top from the bulk, the $\mathbb{Z}_2 \times \mathbb{Z}_2$ orbifold point group (with 1,711 MSSM-like models) occupies a large fraction of the prediction area, especially in cases with less than ≈ 80 SM singlets, see Figures 6(c) and 6(d). Finally, for MSSMs without a heavy top from the bulk and without non-Abelian flavor symmetries, several orbifold point groups become relevant in distinct areas of Figure 6(d): for MSSMs with 90 to 115 SM singlets the $\mathbb{Z}_8$-I point group appears, while for MSSMs with large b and ≈ 90 or 140 SM singlets the point groups

$\mathbb{Z}_{12}$-I or $\mathbb{Z}_6$-I get predicted, respectively. Moreover, MSSMs with many (≈ 300) SM singlets are predicted to originate from orbifolds with $\mathbb{Z}_3 \times \mathbb{Z}_6$ or $\mathbb{Z}_6 \times \mathbb{Z}_6$ point groups for $b \lesssim 5$ or $5 \lesssim b \lesssim 10$, respectively.

4.1. An Almost Perfect MSSM-Like Orbifold Model

In this section, we present an explicit "almost perfect" MSSM-like orbifold model, based on the point group $\mathbb{Z}_2 \times \mathbb{Z}_4$ (i.e. the orbifold geometry is labeled $\mathbb{Z}_2 \times \mathbb{Z}_4$ $(2, 4)$ in the nomenclature of Ref. [30]; also see Ref. [65] for string model building based on the $\mathbb{Z}_2 \times \mathbb{Z}_4$ orbifold). In this case, the six-torus is non-factorizable and both rotational generators of the orbifold are roto-translations, defined as

$$\left(\theta, \frac{1}{2}(e_1 + e_2 + e_6)\right) \quad \text{and} \quad \left(\omega, \frac{1}{2}(e_1 + e_2 + e_6)\right), \quad (6)$$

where $\theta^2 = \omega^4 = \mathbb{1}$. The $\mathbb{Z}_2 \times \mathbb{Z}_4$ shift vectors associated to θ and ω are chosen as

$$V_1 = \left(0, 0, 0, 0, \frac{1}{2}, \frac{1}{2}, \frac{3}{2}, \frac{3}{2}\right)\left(0, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, 2\right), \quad (7a)$$

$$V_2 = \left(-\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, 1, -1, 0, -1, \frac{1}{4}\right)\left(-\frac{3}{4}, 0, 0, 0, 0, 0, \frac{1}{2}, -\frac{3}{4}\right), \quad (7b)$$

ADVANCED
SCIENCE NEWS

www.advancedsciencenews.com

Fortschritte
der Physik
Progress
of Physics

www.fp-journal.org
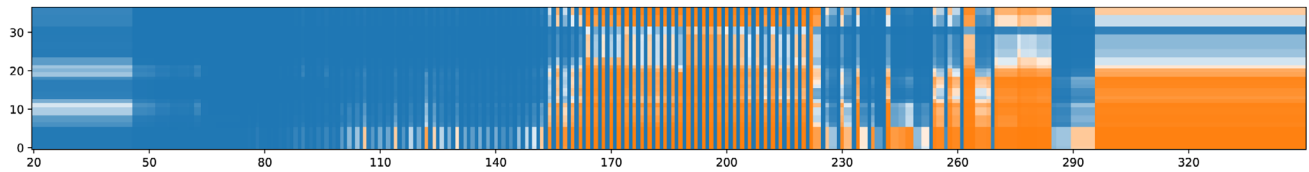
(a) *Stringy origin of the MSSM: with heavy top from bulk and a non-Abelian flavor symmetry.*
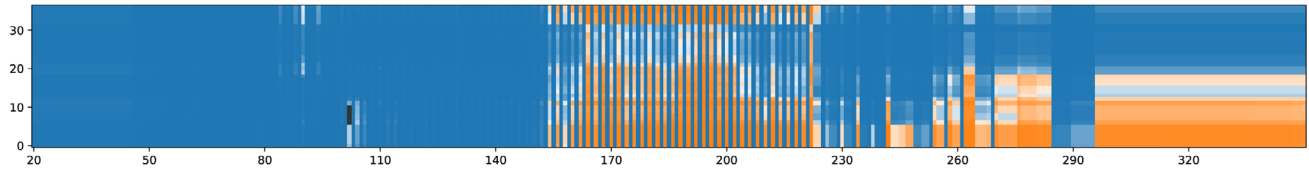
(b) *Stringy origin of the MSSM: with heavy top from bulk but without a non-Abelian flavor symmetry. The small black region for an MSSM with ≈ 100 SM singlets corresponds to $\mathbb{Z}_8$-I.*
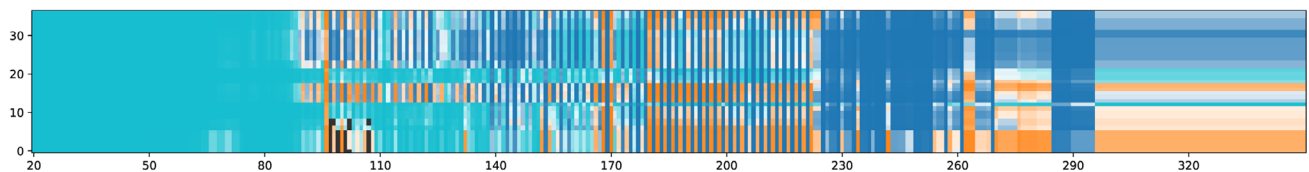
(c) *Stringy origin of the MSSM: without heavy top from bulk but with a non-Abelian flavor symmetry. The small black regions correspond to either $\mathbb{Z}_6$-II or $\mathbb{Z}_8$-I.*
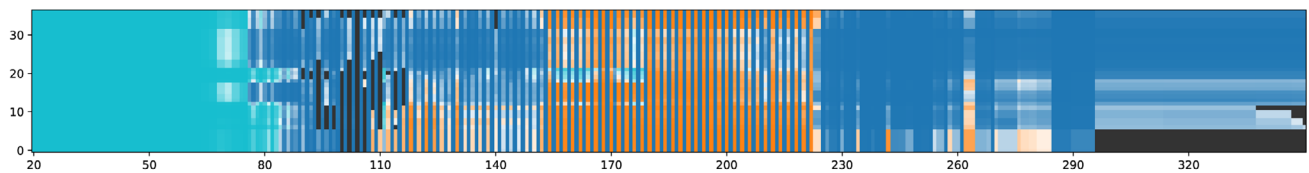
(d) *Stringy origin of the MSSM: without heavy top from bulk and without a non-Abelian flavor symmetry. The black regions correspond to either $\mathbb{Z}_{12}$-I, $\mathbb{Z}_3 \times \mathbb{Z}_6$, $\mathbb{Z}_6$-I, $\mathbb{Z}_6 \times \mathbb{Z}_6$ or $\mathbb{Z}_8$-I.*
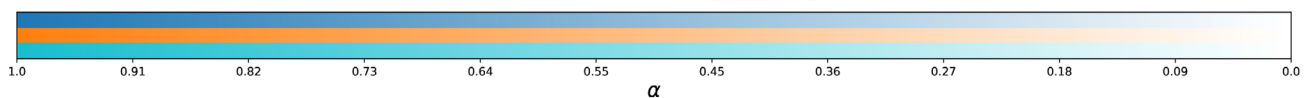
Figure 6. Predicted origin of the MSSM for four different cases (a) - (d) depending on the features "heavy top from bulk", see Section 2.4 and "non-Abelian flavor symmetry", see Section 2.5: in each case, the x-axis corresponds to the number of SM singlets (i.e. right-handed neutrinos), while the y-axis gives the beta-function coefficient b that sets the scale of supersymmetry breaking via hidden sector gaugino condensation.
The colors are associated to point groups as follows:
(i) blue: $\mathbb{Z}_2 \times \mathbb{Z}_4$,
(ii) orange: $\mathbb{Z}_4 \times \mathbb{Z}_4$,
(iii) turquoise: $\mathbb{Z}_2 \times \mathbb{Z}_2$.
Black indicates a different point group than the three dominant ones. The transparency $\alpha := p(1^{\text{st}}) - \frac{3}{4}p(2^{\text{nd}})$ of each pixel indicates the difference between the highest and the second highest probabilities of the predictions, $p(1^{\text{st}})$ and $p(2^{\text{nd}})$, respectively. Note that $\sum_i p(P_i) = 1$. Hence, the color gets very transparent if $p(1^{\text{st}}) \approx p(2^{\text{nd}})$.

Table 3. Massless matter spectrum of an "almost perfect" MSSM-like orbifold model originating from the $\mathbb{Z}_2 \times \mathbb{Z}_4$ (2, 4) orbifold geometry. The four-dimensional gauge group is $SU(3)_{flavor} \times SU(3)_C \times SU(2)_L \times SU(5)_{hidden} \times U(1)_Y \times U(1)^6$. The untwisted sectors U_a, $a = 1, 2, 3$, are defined in Section 2.4, while $T_{(1,3)}$ denotes the $\theta\,\omega^3$ twisted sector. There are three generations of quarks and leptons, three Higgs-pairs but no additional vector-like exotics charged under the SM. $SU(3)_{flavor}$ is a gauged flavor symmetry, where f_i^0 are so-called flavons: SM singlets charged as triplets under $SU(3)_{flavor}$.

Sector	#	Irrep	Labels
U_1	1	$(\mathbf{3}; \mathbf{3}, \mathbf{2}; \mathbf{1})_{\frac{1}{6}}$	q_i
	1	$(\mathbf{1}; \mathbf{1}, \mathbf{2}; \mathbf{1})_{-\frac{1}{2}}$	ℓ_i or h_d
	2	$(\mathbf{1}; \mathbf{1}, \mathbf{1}; \mathbf{1})_0$	s_i^0
U_2	1	$(\bar{\mathbf{3}}; \mathbf{1}, \mathbf{2}; \mathbf{1})_{-\frac{1}{2}}$	ℓ_i or h_d
	1	$(\mathbf{3}; \mathbf{1}, \mathbf{2}; \mathbf{1})_{\frac{1}{2}}$	h_u
U_3	1	$(\mathbf{3}; \bar{\mathbf{3}}, \mathbf{1}; \mathbf{1})_{-\frac{2}{3}}$	$\bar{u}_i$
	1	$(\mathbf{1}; \bar{\mathbf{3}}, \mathbf{1}; \mathbf{1})_{\frac{1}{3}}$	$\bar{d}_i$
	1	$(\mathbf{3}; \mathbf{1}, \mathbf{1}; \mathbf{1})_1$	$\bar{e}_i$
	2	$(\mathbf{1}; \mathbf{1}, \mathbf{1}; \bar{\mathbf{5}})_0$	s_i^0
	1	$(\mathbf{1}; \mathbf{1}, \mathbf{1}; \mathbf{1})_0$	s_i^0
$T_{(1,3)}$	22	$(\mathbf{1}; \mathbf{1}, \mathbf{1}; \mathbf{1})_0$	s_i^0
	2	$(\mathbf{1}; \bar{\mathbf{3}}, \mathbf{1}; \mathbf{1})_{\frac{1}{3}}$	$\bar{d}_i$
	2	$(\mathbf{1}; \mathbf{1}, \mathbf{2}; \mathbf{1})_{-\frac{1}{2}}$	ℓ_i or h_d
	10	$(\bar{\mathbf{3}}; \mathbf{1}, \mathbf{1}; \mathbf{1})_0$	f_i^0
	2	$(\mathbf{1}; \mathbf{1}, \mathbf{1}; \mathbf{5})_0$	s_i^0

respectively. Furthermore, the non-trivial Wilson lines W_i associated to the six independent directions e_i of the compact orbifold space read

$$W_3 = \left(-\frac{1}{4}, \frac{1}{4}, \frac{5}{4}, -\frac{5}{4}, \frac{5}{4}, \frac{5}{4}, -\frac{5}{4}, -\frac{5}{4} \right)$$
$$\times \left(-\frac{3}{4}, -\frac{5}{4}, -\frac{1}{4}, -\frac{1}{4}, \frac{1}{4}, \frac{5}{4}, \frac{9}{4}, \frac{3}{4} \right), \quad (8a)$$

$$W_5 = (0, 0, 0, 0, 0, 0, 0, 0)\left(\frac{3}{2}, -\frac{1}{2}, -1, 2, 0, -\frac{3}{2}, 1, \frac{1}{2} \right), \quad (8b)$$

while $W_3 = W_4 = W_6$ and $W_1 = W_2 = (0^{16})$ are fixed due to geometric constraints. These shifts and Wilson lines act on the $E_8 \times E_8$ gauge degrees of freedom. Using the orbifolder,[31] the resulting four-dimensional gauge group reads

$$SU(3)_{flavor} \times SU(3)_C \times SU(2)_L \times SU(5)_{hidden} \times U(1)_Y \times U(1)^6, \quad (9)$$

where $SU(3)_{flavor}$ is a gauged $SU(3)$ flavor symmetry and one of the U(1)'s appears anomalous. Gauge symmetry breaking is non-local in this model, i.e. it is associated to a non-trivial fundamental group of the orbifold with a Wilson line that breaks an intermediate SU(5) GUT to the SM. Consequently, there exist stable winding strings (with masses related to the compactification scale) which can contribute to the dark matter relic density.[66] Finally, the massless matter spectrum is given in **Table 3**. From the point of view of the MSSM, this model contains exactly three gen-

erations of quarks and leptons, three Higgs-pairs (h_u, h_d) and in total 75 SM singlets, but no additional (vector-like) exotics. Importantly, there are ten flavons f^0 that transform as triplets of $SU(3)_{flavor}$ (contributing $10 \times 3 = 30$ SM singlets from the point of view of the SM). Their VEVs will be important to explain quark and lepton masses and mixings, see Section 2.7. Concerning hidden sector gaugino condensation, we analyze the chiral spectrum with respect to the hidden sector gauge group factor $SU(5)_{hidden}$: The massless spectrum contains two **5**-plets and two $\bar{\mathbf{5}}$-plets. They can decouple without breaking the SM or $SU(5)_{hidden}$. Hence, there is no light matter charged under $SU(5)_{hidden}$ that enters the beta-function coefficient and we obtain $b = 15$.

In total, we have identified 20 almost perfect MSSM-like orbifold models, see the arXiv ancillary files.[32] All of them are based on the point group $\mathbb{Z}_2 \times \mathbb{Z}_4$, one from the (1, 6) orbifold geometry and 19 from the (2, 4) orbifold geometry. These 20 almost perfect MSSM-like orbifold models are very similar from the point of view of the massless matter spectrum:

- there is a gauged $SU(3)_{flavor}$ flavor symmetry,
- all quark doublets (q), up-type quarks ($\bar{u}$), electrons ($\bar{e}$) and up-type Higgses (h_u) are triplets of $SU(3)_{flavor}$ and live in the bulk of the orbifold,
- one down-type quark ($\bar{d}$) is a singlet of $SU(3)_{flavor}$ and lives in the bulk of the orbifold,
- two down-type quarks ($\bar{d}$) and some lepton doublets are localized either in the $T_{(1,3)}$ or in the $T_{(0,2)}$ twisted sector.

We marked these 20 almost perfect MSSM-like orbifold models as yellow points in Figure 5.

However, it is important to note that a detailed phenomenological study of these models is not possible at the moment. The reason is a missing understanding of the R-symmetries for these orbifold geometries. Even though R-symmetries are in general very well understood in heterotic orbifold compactifications,[67–69] there are two exceptions for orbifolds with $\mathbb{Z}_2 \times \mathbb{Z}_4$ point group, being the (1, 6) and (2, 4) orbifold geometries.[70] Hence, in order to analyze the phenomenological properties of our almost perfect MSSM-like orbifold models in more detail, the R-symmetries have to be re-analyzed for these orbifolds.

Finally, note that in our prediction task the $\mathbb{Z}_2 \times \mathbb{Z}_4$ point group is one of the dominant classes (with 48,812 MSSM-like orbifold models in total; however, the (1, 6) and (2, 4) orbifold geometries of this point group lead to only 82 and 320 MSSM-like orbifold models, respectively). Since the almost perfect models are based on the $\mathbb{Z}_2 \times \mathbb{Z}_4$ point group, it is reasonable that our ML model predicts $\mathbb{Z}_2 \times \mathbb{Z}_4$ as the orbifold origin of the MSSM for a wide range in feature space.

5. Conclusions and Outlook

In conclusion, in this paper we show that the huge number of possible string models in the heterotic orbifold landscape gets subdivided into several sub-landscapes, each with their own distinct phenomenological properties. These properties include most prominently the appearance of various types of vector-like exotics (see **Tables A.1** and **A.2**), the number of SM singlets, the existence of non-Abelian flavor symmetries, and the hidden sec-

tor beta function relevant for supersymmetry breaking via gaugino condensation, among others.

In detail, we demonstrate that the choice of the point group P that underlies an orbifold geometry leaves a particular imprint, for example, on the particle spectrum of vector-like exotics for MSSM-like orbifold models. This imprint can be exploited in order to predict with high certainty the most likely point group that is able to reproduce a specific MSSM-like particle spectrum, see Table 1. In more detail, this is achieved using a machine learning (ML) algorithm known as "boosted decision tree" that is particularly suitable for our classification task. We train and test our boosted decision tree on the largest known set of 126,783 distinct MSSM-like orbifold models. Then, we dissect our trained ensemble of decision trees in order to identify the most important phenomenological properties used by the decision trees to classify the point group of a given MSSM-like particle spectrum, see Figure 4.

After training and evaluating our ML model, we apply our boosted decision tree to the MSSM in order to predict the stringy origin of the MSSM. For this task, we assume that supersymmetry is broken in the MSSM by hidden sector gaugino condensation, and we extend the MSSM by a large number of SM singlets (i.e. right-handed neutrinos). Then, we vary over (i) the beta-function coefficient b of the hidden sector gauge group and (ii) the number of SM singlets, and predict in each case the most probable stringy origin of the MSSM. The result is shown in Figure 5. In a nutshell, we find that for up to ≈ 140 SM singlets and

all ranges of b, orbifolds with $\mathbb{Z}_2 \times \mathbb{Z}_4$ point group seem to be the most promising orbifold compactifications of the heterotic string to yield the MSSM. For an even larger number of SM singlets, the boosted decision tree predicts orbifolds with $\mathbb{Z}_4 \times \mathbb{Z}_4$ point group. As one can see in Figure 6 these predictions depend on a few additional phenomenological features that are inspired by the MSSM-like orbifold models of our dataset. Varying these features (in addition to b and the number of SM singlets) yields further promising point groups, like $\mathbb{Z}_2 \times \mathbb{Z}_2$ and in small corners of the parameter space also $\mathbb{Z}_6$-II, $\mathbb{Z}_8$-I and $\mathbb{Z}_3 \times \mathbb{Z}_6$. Hence, we suggest to focus heterotic orbifold model building on one of these most promising orbifold geometries.

Furthermore, we present the first "almost perfect" MSSM-like orbifold models, see Table 3 for one example. These orbifold models, which were unknown in the literature, have exactly three generations of quarks and leptons, either three or five Higgs-pairs, but no additional vector-like exotics. They originate from the $\mathbb{Z}_2 \times \mathbb{Z}_4$ (1,6) and (2,4) orbifold geometries. These orbifolds are equipped with non-local GUT breaking,[30] relevant for stringy dark matter.[66] However, since R-symmetries are not under control just for these two $\mathbb{Z}_2 \times \mathbb{Z}_4$ orbifold geometries,[70] our findings should encourage to continue the efforts of Refs. [67–69] to study the R-symmetries for these special orbifold geometries.

Appendix: Vector-Like Exotics in the Orbifold-Landscape

Table A.1. Percentage of MSSM-like orbifold models with certain types of vector-like exotics. Hypercharge is normalized such that $(\mathbf{3}, \mathbf{2})_{1/6}$ is a left-chiral quark-doublet. The row "# MSSM" lists the number of inequivalent MSSM-like orbifold models in our dataset. A complex representation has to be amended by its complex conjugate, e.g. $(\mathbf{3}, \mathbf{2})_{1/6}$ stands for $(\mathbf{3}, \mathbf{2})_{1/6} \oplus (\bar{\mathbf{3}}, \mathbf{2})_{-1/6}$.

	$\mathbb{Z}_4$	$\mathbb{Z}_6$-I	$\mathbb{Z}_6$-II	$\mathbb{Z}_7$	$\mathbb{Z}_8$-I	$\mathbb{Z}_8$-II	$\mathbb{Z}_{12}$-I	$\mathbb{Z}_{12}$-II	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$\mathbb{Z}_2 \times \mathbb{Z}_6$-I	$\mathbb{Z}_3 \times \mathbb{Z}_3$	$\mathbb{Z}_3 \times \mathbb{Z}_6$	$\mathbb{Z}_4 \times \mathbb{Z}_4$	$\mathbb{Z}_6 \times \mathbb{Z}_6$
# MSSM	212	62	1,870	1	865	2,844	1,250	435	1,711	55,429	1,095	3,337	5,153	48,812	3,707
$(\mathbf{3}, \mathbf{2})_{1/6}$	1.89%	0%	31.60%	0%	4.05%	25.00%	4.00%	24.14%	8.53%	14.04%	31.60%	14.59%	19.04%	15.31%	39.60%
$(\bar{\mathbf{3}}, \mathbf{1})_{-2/3}$	50.47%	0%	28.66%	0%	5.55%	44.13%	3.28%	35.63%	8.36%	14.35%	22.92%	18.91%	18.84%	14.93%	29.92%
$(\bar{\mathbf{3}}, \mathbf{1})_{1/3}$	100%	100%	99.95%	100%	99.54%	100%	100%	100%	100%	98.19%	99.73%	94.76%	99.83%	99.03%	99.00%
$(\mathbf{1}, \mathbf{2})_{-1/2}$	96.23%	35.48%	92.19%	100%	93.99%	94.94%	78.96%	91.72%	92.17%	92.20%	96.26%	90.62%	99.20%	95.07%	95.50%
$(\mathbf{1}, \mathbf{1})_1$	1.89%	0%	28.66%	0%	5.55%	44.09%	3.28%	35.63%	8.36%	14.36%	22.92%	11.30%	18.92%	14.98%	29.92%
$(\mathbf{1}, \mathbf{1})_0$	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
$(\mathbf{3}, \mathbf{2})_{-1/2}$	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0.06%	0%	0%	0%
$(\mathbf{3}, \mathbf{2})_{1/2}$	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0.18%	0%	0%	0%
$(\mathbf{3}, \mathbf{2})_{-1/3}$	0%	0%	0%	0%	0.23%	1.05%	0.24%	0%	0%	2.27%	0.09%	0%	0.43%	0.53%	0.32%
$(\mathbf{3}, \mathbf{2})_{-1/6}$	0%	0%	0%	0%	0%	0%	0%	0.64%	0%	0%	0%	7.19%	0.02%	0%	0%
$(\mathbf{3}, \mathbf{2})_{-1/12}$	2.83%	0%	0%	0%	0%	0%	0%	0%	0%	0.00%	0%	0%	0%	0.00%	0%
$(\mathbf{3}, \mathbf{1})_0$	0%	0%	6.84%	0%	0%	0%	25.12%	0%	0%	0%	0%	85.77%	9.99%	0%	0.35%
$(\mathbf{3}, \mathbf{1})_{-1/2}$	0%	0%	0%	0%	0%	0%	1.12%	0%	0%	0%	0%	0%	0.06%	0%	0%
$(\mathbf{3}, \mathbf{1})_{1/2}$	0%	0%	0%	0%	0%	0%	0.80%	0%	0%	0%	0%	0%	0.04%	0%	0%
$(\mathbf{3}, \mathbf{1})_{-2/3}$	0%	0%	0%	0%	0%	0%	0.32%	0%	0%	0%	0%	5.60%	0.06%	0%	0%
$(\mathbf{3}, \mathbf{1})_{1/3}$	0%	0%	4.71%	0%	0%	0%	18.72%	0%	0%	0%	0%	73.24%	6.85%	0%	0.27%
$(\mathbf{3}, \mathbf{1})_{-5/6}$	0%	0%	2.89%	0%	0%	0.74%	0%	0%	0%	0.12%	0%	0%	0.62%	0.44%	0.76%
$(\mathbf{3}, \mathbf{1})_{-1/6}$	0%	0%	0.21%	0%	0%	0%	4.32%	0%	0%	0%	0%	0%	0.45%	0%	0%
$(\mathbf{3}, \mathbf{1})_{1/6}$	67.45%	93.55%	66.47%	0%	87.63%	74.30%	69.20%	66.90%	85.21%	90.85%	86.58%	0%	84.84%	95.75%	98.17%

(Continued)

Table A.1. Continued.

# MSSM	$\mathbb{Z}_4$ 212	$\mathbb{Z}_6$-I 62	$\mathbb{Z}_6$-II 1,870	$\mathbb{Z}_7$ 1	$\mathbb{Z}_8$-I 865	$\mathbb{Z}_8$-II 2,844	$\mathbb{Z}_{12}$-I 1,250	$\mathbb{Z}_{12}$-II 435	$\mathbb{Z}_2 \times \mathbb{Z}_2$ 1,711	$\mathbb{Z}_2 \times \mathbb{Z}_4$ 55,429	$\mathbb{Z}_2 \times \mathbb{Z}_6$-I 1,095	$\mathbb{Z}_3 \times \mathbb{Z}_3$ 3,337	$\mathbb{Z}_3 \times \mathbb{Z}_6$ 5,153	$\mathbb{Z}_4 \times \mathbb{Z}_4$ 48,812	$\mathbb{Z}_6 \times \mathbb{Z}_6$ 3,707
$(\mathbf{3}, \mathbf{1})_{-7/12}$	2.83%	0%	0%	0%	0%	0%	0.32%	0%	0%	0%	0%	0%	0%	0.01%	0%
$(\mathbf{3}, \mathbf{1})_{-1/12}$	41.04%	0%	0%	0%	0.23%	0.56%	1.76%	0%	0%	0.08%	0%	0%	0%	0.13%	0%
$(\mathbf{3}, \mathbf{1})_{5/12}$	4.72%	0%	0%	0%	0%	0.14%	0.80%	0%	0%	0.03%	0%	0%	0%	0.06%	0%
$(\mathbf{3}, \mathbf{1})_{2/21}$	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
$(\mathbf{3}, \mathbf{1})_{5/21}$	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
$(\mathbf{1}, \mathbf{2})_0$	82.55%	100%	85.94%	0%	99.31%	93.95%	76.32%	90.34%	84.57%	96.49%	98.17%	0%	89.93%	99.26%	99.57%
$(\mathbf{1}, \mathbf{2})_{1/3}$	0%	0%	0.43%	0%	0%	0%	4.72%	0%	0%	0%	0%	0%	0.66%	0%	0%
$(\mathbf{1}, \mathbf{2})_{2/3}$	0%	0%	0%	0%	0%	0%	0.80%	0%	0%	0%	0%	0%	0.02%	0%	0%
$(\mathbf{1}, \mathbf{2})_{1/4}$	41.98%	0%	0%	0%	0.23%	0.70%	2.40%	0%	0%	0.08%	0%	0%	0%	0.14%	0%
$(\mathbf{1}, \mathbf{2})_{3/4}$	2.83%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
$(\mathbf{1}, \mathbf{2})_{1/6}$	0%	0%	8.98%	0%	0%	0%	26.72%	0%	0%	0%	0%	96.55%	10.58%	0%	0.35%
$(\mathbf{1}, \mathbf{2})_{5/6}$	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0.21%	0%	0%	0%
$(\mathbf{1}, \mathbf{2})_{1/14}$	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
$(\mathbf{1}, \mathbf{2})_{5/14}$	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
$(\mathbf{1}, \mathbf{1})_{1/2}$	100%	100%	92.09%	0%	100%	100%	76.96%	100%	95.79%	99.16%	99.36%	0%	90.28%	99.99%	99.65%
$(\mathbf{1}, \mathbf{1})_{1/3}$	0%	0%	8.98%	0%	0%	0%	29.20%	0%	0%	0%	0%	99.52%	10.58%	0%	0.35%
$(\mathbf{1}, \mathbf{1})_{2/3}$	0%	0%	8.98%	0%	0%	0%	28.72%	0%	0%	0%	0%	95.62%	10.58%	0%	0.35%
$(\mathbf{1}, \mathbf{1})_{1/4}$	52.36%	0%	0%	0%	0.23%	0.95%	3.20%	0%	0%	0.19%	0%	0%	0%	0.17%	0%
$(\mathbf{1}, \mathbf{1})_{3/4}$	42.92%	0%	0%	0%	0.23%	0.42%	1.92%	0%	0%	0.12%	0%	0%	0%	0.14%	0%
$(\mathbf{1}, \mathbf{1})_{1/6}$	0%	0%	1.07%	0%	0%	0%	6.32%	0%	0%	0%	0%	0%	0.85%	0%	0%
$(\mathbf{1}, \mathbf{1})_{5/6}$	0%	0%	0%	0%	0%	0%	1.12%	0%	0%	0%	0%	0%	0.04%	0%	0%
$(\mathbf{1}, \mathbf{1})_{1/7}$	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
$(\mathbf{1}, \mathbf{1})_{2/7}$	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
$(\mathbf{1}, \mathbf{1})_{3/7}$	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
$(\mathbf{1}, \mathbf{1})_{4/7}$	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

Table A.2. Average numbers of vector-like exotics for MSSM-like orbifold models. Hypercharge is normalized such that $(\mathbf{3}, \mathbf{2})_{1/6}$ is a left-chiral quark-doublet. The row "# MSSM" lists the number of inequivalent MSSM-like orbifold models in our dataset. A complex representation has to be amended by its complex conjugate, e.g. $(\mathbf{3}, \mathbf{2})_{1/6}$ stands for $(\mathbf{3}, \mathbf{2})_{1/6} \oplus (\bar{\mathbf{3}}, \mathbf{2})_{-1/6}$.

# MSSM	$\mathbb{Z}_4$ 212	$\mathbb{Z}_6$-I 62	$\mathbb{Z}_6$-II 1,870	$\mathbb{Z}_7$ 1	$\mathbb{Z}_8$-I 865	$\mathbb{Z}_8$-II 2,844	$\mathbb{Z}_{12}$-I 1,250	$\mathbb{Z}_{12}$-II 435	$\mathbb{Z}_2 \times \mathbb{Z}_2$ 1,711	$\mathbb{Z}_2 \times \mathbb{Z}_4$ 55,429	$\mathbb{Z}_2 \times \mathbb{Z}_6$-I 1,095	$\mathbb{Z}_3 \times \mathbb{Z}_3$ 3,337	$\mathbb{Z}_3 \times \mathbb{Z}_6$ 5,153	$\mathbb{Z}_4 \times \mathbb{Z}_4$ 48,812	$\mathbb{Z}_6 \times \mathbb{Z}_6$ 3,707
$(\mathbf{3}, \mathbf{2})_{1/6}$	0.04	0	0.48	0	0.05	0.31	0.09	0.46	0.09	0.19	0.51	0.17	0.23	0.18	0.46
$(\bar{\mathbf{3}}, \mathbf{1})_{-2/3}$	0.52	0	0.41	0	0.08	0.49	0.06	0.71	0.09	0.21	0.35	0.23	0.23	0.18	0.36
$(\bar{\mathbf{3}}, \mathbf{1})_{1/3}$	3.52	3.52	5.22	4.00	4.41	4.36	4.21	7.58	7.31	5.66	7.96	3.64	6.57	5.72	6.25
$(\mathbf{1}, \mathbf{2})_{-1/2}$	2.64	1.68	4.03	3.00	3.41	3.06	2.92	4.28	3.34	3.97	5.23	2.86	5.83	4.46	5.50
$(\mathbf{1}, \mathbf{1})_1$	0.04	0	0.41	0	0.08	0.49	0.06	0.71	0.09	0.21	0.35	0.15	0.23	0.18	0.36
$(\mathbf{1}, \mathbf{1})_0$	67.37	136.58	106.06	32.00	108.74	89.78	85.99	101.78	122.54	132.64	139.59	106.48	177.56	166.38	186.95
$(\mathbf{3}, \mathbf{2})_{-1/2}$	0	0	0	0	0	0	0	0	0	0	0	0.00	0	0	0
$(\mathbf{3}, \mathbf{2})_{1/2}$	0	0	0	0	0	0	0	0	0	0	0	0.00	0	0	0
$(\mathbf{3}, \mathbf{2})_{-1/3}$	0	0	0	0	0.00	0.01	0.00	0	0	0.03	0.00	0	0.01	0.01	0.00
$(\mathbf{3}, \mathbf{2})_{-1/6}$	0	0	0	0	0	0	0.01	0	0	0	0	0.08	0.00	0	0
$(\mathbf{3}, \mathbf{2})_{-1/12}$	0.03	0	0	0	0	0	0	0	0	0.00	0	0	0	0.00	0
$(\mathbf{3}, \mathbf{1})_0$	0	0	0.19	0	0	0	0.69	0	0	0	0	3.32	0.36	0	0.01
$(\mathbf{3}, \mathbf{1})_{-1/2}$	0	0	0	0	0	0	0.01	0	0	0	0	0	0.00	0	0
$(\mathbf{3}, \mathbf{1})_{1/2}$	0	0	0	0	0	0	0.01	0	0	0	0	0	0.00	0	0
$(\mathbf{3}, \mathbf{1})_{-2/3}$	0	0	0	0	0	0	0.00	0	0	0	0	0.06	0.00	0	0

(Continued)

Table A.2. Continued.

	$\mathbb{Z}_4$	$\mathbb{Z}_6$-I	$\mathbb{Z}_6$-II	$\mathbb{Z}_7$	$\mathbb{Z}_8$-I	$\mathbb{Z}_8$-II	$\mathbb{Z}_{12}$-I	$\mathbb{Z}_{12}$-II	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$\mathbb{Z}_2 \times \mathbb{Z}_6$-I	$\mathbb{Z}_3 \times \mathbb{Z}_3$	$\mathbb{Z}_3 \times \mathbb{Z}_6$	$\mathbb{Z}_4 \times \mathbb{Z}_4$	$\mathbb{Z}_6 \times \mathbb{Z}_6$
# MSSM	212	62	1,870	1	865	2,844	1,250	435	1,711	55,429	1,095	3,337	5,153	48,812	3,707
$(3,1)_{1/3}$	0	0	0.13	0	0	0	0.36	0	0	0	0	1.95	0.14	0	0.01
$(3,1)_{-5/6}$	0	0	0.04	0	0	0.01	0	0	0	0.00	0	0	0.01	0.01	0.01
$(3,1)_{-1/6}$	0	0	0.00	0	0	0	0.06	0	0	0	0	0	0.01	0	0
$(3,1)_{1/6}$	1.99	3.29	2.38	0	3.14	2.55	2.00	1.69	3.83	4.59	4.87	0	3.47	5.37	6.57
$(3,1)_{-7/12}$	0.03	0	0	0	0	0	0.00	0	0	0	0	0	0	0.00	0
$(3,1)_{-1/12}$	1.10	0	0	0	0.00	0.01	0.04	0	0	0.00	0	0	0	0.00	0
$(3,1)_{5/12}$	0.08	0	0	0	0	0.00	0.01	0	0	0.00	0	0	0	0.00	0
$(3,1)_{2/21}$	0	0	0	3.00	0	0	0	0	0	0	0	0	0	0	0
$(3,1)_{5/21}$	0	0	0	4.00	0	0	0	0	0	0	0	0	0	0	0
$(1,2)_0$	7.78	8.58	7.99	0	9.49	8.23	5.84	6.10	7.62	13.58	14.25	0	10.97	15.88	19.27
$(1,2)_{1/3}$	0	0	0.01	0	0	0	0.08	0	0	0	0	0	0.01	0	0
$(1,2)_{2/3}$	0	0	0	0	0	0	0.01	0	0	0	0	0	0.00	0	0
$(1,2)_{1/4}$	1.19	0	0	0	0.00	0.02	0.07	0	0	0.00	0	0	0	0.01	0
$(1,2)_{3/4}$	0.03	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$(1,2)_{1/6}$	0	0	0.44	0	0	0	1.18	0	0	0	0	7.75	0.81	0	0.02
$(1,2)_{5/6}$	0	0	0	0	0	0	0	0	0	0	0	0.00	0	0	0
$(1,2)_{1/14}$	0	0	0	5.00	0	0	0	0	0	0	0	0	0	0	0
$(1,2)_{5/14}$	0	0	0	1.00	0	0	0	0	0	0	0	0	0	0	0
$(1,1)_{1/2}$	18.79	23.81	14.60	0	23.19	20.92	12.45	13.59	18.80	29.37	29.20	0	25.33	37.34	41.21
$(1,1)_{1/3}$	0	0	3.14	0	0	0	7.72	0	0	0	0	42.99	5.22	0	0.23
$(1,1)_{2/3}$	0	0	0.43	0	0	0	1.25	0	0	0	0	8.30	0.84	0	0.03
$(1,1)_{1/4}$	14.90	0	0	0	0.07	0.17	0.68	0	0	0.03	0	0	0	0.10	0
$(1,1)_{3/4}$	1.18	0	0	0	0.00	0.01	0.04	0	0	0.00	0	0	0	0.01	0
$(1,1)_{1/6}$	0	0	0.13	0	0	0	0.63	0	0	0	0	0	0.17	0	0
$(1,1)_{5/6}$	0	0	0	0	0	0	0.01	0	0	0	0	0	0.00	0	0
$(1,1)_{1/7}$	0	0	0	22.00	0	0	0	0	0	0	0	0	0	0	0
$(1,1)_{2/7}$	0	0	0	16.00	0	0	0	0	0	0	0	0	0	0	0
$(1,1)_{3/7}$	0	0	0	11.00	0	0	0	0	0	0	0	0	0	0	0
$(1,1)_{4/7}$	0	0	0	9.00	0	0	0	0	0	0	0	0	0	0	0

Conflict of Interest

The authors have declared no conflict of interest.

[1] W. Lerche, D. Lüst, A. N. Schellekens, *Nucl. Phys.* **1987**, *B287*, 477, [,477(1986)].

[2] M. R. Douglas, *JHEP* **2003**, *05*, 046, arXiv:hep-th/0303194 [hep-th].

[3] J. Halverson, F. Ruehle, *Phys. Rev.* **2019**, *D99*, 046015, arXiv:1809.08279 [hep-th].

[4] Y.-H. He, *Deep-Learning the Landscape*, **2017**, arXiv:1706.02714 [hep-th].

[5] D. Krefl, R.-K. Seong, *Phys. Rev.* **2017**, *D96*, 066014, arXiv:1706.03346 [hep-th].

[6] F. Ruehle, *JHEP* **2017**, *08*, 038, arXiv:1706.07024 [hep-th].

[7] J. Carifio, J. Halverson, D. Krioukov, B. D. Nelson, *JHEP* **2017**, *09*, 157, arXiv:1707.00655 [hep-th].

[8] Y.-H. He, *The Calabi-Yau Landscape: From Geometry, to Physics, to Machine-Learning*, **2018**, arXiv:1812.02893 [hep-th].

[9] F. Ruehle, *Physics Reports* **2020**, *839*, 1, Data science applications to string theory, http://www.sciencedirect.com/science/article/pii/S0370157319303072.

[10] A. Cole, G. Shiu, *JHEP* **2019**, *03*, 054, arXiv:1812.06960 [hep-th].

[11] K. Bull, Y.-H. He, V. Jejjala, C. Mishra, *Phys. Lett.* **2019**, *B795*, 700, arXiv:1903.03113 [hep-th].

[12] J. Halverson, B. Nelson, F. Ruehle, *JHEP* **2019**, *06*, 003, arXiv:1903.11616 [hep-th].

[13] Y.-H. He, S.-J. Lee, *Phys. Lett.* **2019**, *B798*, 134889, arXiv:1904.08530 [hep-th].

[14] A. Cole, A. Schachner, G. Shiu, *JHEP* **2019**, *11*, 045, arXiv:1907.10072 [hep-th].

[15] A. Ashmore, Y.-H. He, B. A. Ovrut, *Machine learning Calabi-Yau metrics*, **2019**, arXiv:1910.08605 [hep-th].

[16] J. Halverson, C. Long, *Statistical Predictions in String Theory and Deep Generative Models*, **2020**, arXiv:2001.00555 [hep-th].

[17] L. J. Dixon, J. A. Harvey, C. Vafa, E. Witten, *Nucl. Phys.* **1985**, *B261*, 678, [,678(1985)].

[18] L. J. Dixon, J. A. Harvey, C. Vafa, E. Witten, *Nucl. Phys.* **1986**, *B274*, 285.

[19] L. E. Ibáñez, H. P. Nilles, F. Quevedo, *Phys. Lett.* **1987**, *B187*, 25.

[20] T. Kobayashi, S. Raby, R.-J. Zhang, *Nucl. Phys.* **2005**, *B704*, 3, hep-ph/0409098.

[21] W. Buchmüller, K. Hamaguchi, O. Lebedev, M. Ratz, *Phys. Rev. Lett.* **2006**, *96*, 121602, hep-ph/0511035.

[22] W. Buchmüller, K. Hamaguchi, O. Lebedev, M. Ratz, *Nucl. Phys.* **2007**, *B785*, 149, hep-th/0606187.

[23] O. Lebedev, H. P. Nilles, S. Raby, S. Ramos-Sánchez, M. Ratz, P. K. S. Vaudrevange, A. Wingerter, *Phys. Rev.* **2007**, *D77*, 046013, arXiv:0708.2691 [hep-th].

[24] O. Lebedev, H. P. Nilles, S. Raby, S. Ramos-Sánchez, M. Ratz, P. K. S. Vaudrevange, A. Wingerter, *Phys. Lett.* **2007**, *B645*, 88, hep-th/0611095.

[25] O. Lebedev, H. P. Nilles, S. Ramos-Sánchez, M. Ratz, P. K. S. Vaudrevange, *Phys. Lett.* **2008**, *B668*, 331, arXiv:0807.4384 [hep-th].

[26] H. P. Nilles, P. K. S. Vaudrevange, *Mod. Phys. Lett.* **2015**, *A30*, 1530008, arXiv:1403.1597 [hep-th].

[27] Y. Olguín-Trejo, R. Pérez-Martínez, S. Ramos-Sánchez, *Phys. Rev.* **2018**, *D98*, 106020, arXiv:1808.06622 [hep-th].

[28] A. Mütter, E. Parr, P. K. S. Vaudrevange, *Nucl. Phys.* **2019**, *B940*, 113, arXiv:1811.05993 [hep-th].

[29] E. Parr, P. K. S. Vaudrevange, *Nucl. Phys.* **2020**, *B952*, 114922, arXiv:1910.13473 [hep-th].

[30] M. Fischer, M. Ratz, J. Torrado, P. K. S. Vaudrevange, *JHEP* **2013**, *01*, 084, arXiv:1209.3906 [hep-th].

[31] H. P. Nilles, S. Ramos-Sánchez, P. K. S. Vaudrevange, A. Wingerter, *Comput. Phys. Commun.* **2012**, *183*, 1363, arXiv:1110.5229 [hep-th].

[32] E. Parr, M. Wimmer, P. K. S. Vaudrevange, *The model-files for the orbifolder, which contain the gauge embeddings of all MSSM-like* $\mathbb{Z}_N \times \mathbb{Z}_M$ *orbifold models, can be found as arXiv ancillary files of this paper*, **2020**.

[33] S. Ramos-Sánchez, *Fortsch. Phys.* **2009**, *10*, 907, arXiv:0812.3560 [hep-th], Ph.D. Thesis (Advisor: H.P. Nilles).

[34] W. Buchmüller, K. Hamaguchi, O. Lebedev, S. Ramos-Sanchez, M. Ratz, *Phys. Rev. Lett.* **2007**, *99*, 021601, arXiv:hep-ph/0703078 [HEP-PH].

[35] W. Buchmüller, R. Ruckl, D. Wyler, *Phys. Lett.* **1987**, *B191*, 442, [Erratum: Phys. Lett.B448,320(1999)].

[36] M. Bauer, M. Neubert, *Phys. Rev. Lett.* **2016**, *116*, 141802, arXiv:1511.01900 [hep-ph].

[37] B. Diaz, M. Schmaltz, Y.-M. Zhong, *JHEP* **2017**, *10*, 097, arXiv:1706.05033 [hep-ph].

[38] Particle Data Group, M. Tanabashi et al., *Phys. Rev.* **2018**, *D98*, 030001.

[39] A. E. Faraggi, *Phys. Lett. B* **1992**, *274*, 47.

[40] A. E. Faraggi, *Phys. Lett. B* **1996**, *377*, 43, hep-ph/9506388.

[41] G. Burdman, Y. Nomura, *Nucl. Phys.* **2003**, *B656*, 3, arXiv:hep-ph/0210257 [hep-ph].

[42] P. Hosteins, R. Kappl, M. Ratz, K. Schmidt-Hoberg, *JHEP* **2009**, *07*, 029, arXiv:0905.3323 [hep-ph].

[43] S. Hamidi, C. Vafa, *Nucl. Phys.* **1987**, *B279*, 465.

[44] L. J. Dixon, D. Friedan, E. J. Martinec, S. H. Shenker, *Nucl. Phys.* **1987**, *B282*, 13.

[45] S. Ramos-Sánchez, P. K. S. Vaudrevange, *JHEP* **2019**, *01*, 055, arXiv:1811.00580 [hep-th].

[46] T. Kobayashi, H. P. Nilles, F. Plöger, S. Raby, M. Ratz, *Nucl. Phys.* **2007**, *B768*, 135, arXiv:hep-ph/0611020 [hep-ph].

[47] F. Feruglio, A. Romanino, *Neutrino Flavour Symmetries*, **2019**, arXiv:1912.06028 [hep-ph].

[48] H. P. Nilles, *Phys. Lett.* **1982**, *115B*, 193.

[49] S. Ferrara, L. Girardello, H. P. Nilles, *Phys. Lett.* **1983**, *125B*, 457.

[50] J. P. Derendinger, L. E. Ibáñez, H. P. Nilles, *Phys. Lett.* **1985**, *155B*, 65.

[51] M. Dine, R. Rohm, N. Seiberg, E. Witten, *Phys. Lett.* **1985**, *156B*, 55.

[52] O. Lebedev, H.-P. Nilles, S. Raby, S. Ramos-Sánchez, M. Ratz, P. K. S. Vaudrevange, A. Wingerter, *Phys. Rev. Lett.* **2007**, *98*, 181602, arXiv:hep-th/0611203 [hep-th].

[53] T. P. T. Dijkstra, L. R. Huiszoon, A. N. Schellekens, *Nucl. Phys.* **2005**, *B710*, 3, arXiv:hep-th/0411129 [hep-th].

[54] K. R. Dienes, *Phys. Rev.* **2006**, *D73*, 106010, hep-th/0602286.

[55] V. S. Kaplunovsky, *Nucl. Phys.* **1988**, *B307*, 145, arXiv:hep-th/9205068 [hep-th], [Erratum: Nucl. Phys.B382,436(1992)].

[56] L. J. Dixon, V. Kaplunovsky, J. Louis, *Nucl. Phys.* **1991**, *B355*, 649.

[57] V. Kaplunovsky, J. Louis, *Nucl. Phys.* **1995**, *B444*, 191, arXiv:hep-th/9502077 [hep-th].

[58] M. B. Green, J. H. Schwarz, *Phys. Lett.* **1984**, *149B*, 117.

[59] C. D. Froggatt, H. B. Nielsen, *Nucl. Phys.* **1979**, *B147*, 277.

[60] T. Hastie, R. Tibshirani, J. Friedman, *The Elements of Statistical Learning*, Springer Series in Statistics, Springer New York Inc., New York, NY, USA **2001**.

[61] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, *Journal of Machine Learning Research* **2011**, *12*, 2825.

[62] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, T.-Y. Liu, *NIPS* **2017**.

[63] F. Chollet et al., *Keras*, https://keras.io, **2015**.

[64] T. Chen, C. Guestrin, *CoRR* **2016**, abs/1603.02754, 1603.02754, http://arxiv.org/abs/1603.02754.

[65] D. K. Mayorga Peña, H. P. Nilles, P.-K. Oehlmann, *JHEP* **2012**, *12*, 024, arXiv:1209.6041 [hep-th].

[66] A. Mütter, P. K. S. Vaudrevange, *String scale interacting dark matter from* π_1, **2019**, arXiv:1912.09909 [hep-ph].

[67] N. G. Cabo Bizet, T. Kobayashi, D. K. Mayorga Peña, S. L. Parameswaran, M. Schmitz, I. Zavala, *JHEP* **2013**, *05*, 076, arXiv:1301.2322 [hep-th].

[68] H. P. Nilles, S. Ramos-Sánchez, M. Ratz, P. K. S. Vaudrevange, *Phys. Lett.* **2013**, *B726*, 876, arXiv:1308.3435 [hep-th].

[69] N. G. Cabo Bizet, T. Kobayashi, D. K. Mayorga Peña, S. L. Parameswaran, M. Schmitz, I. Zavala, *JHEP* **2014**, *02*, 098, arXiv:1308.5669 [hep-th].

[70] M. Schmitz, *R-Symmetries from the Orbifolded Heterotic String*, Ph.D. thesis, Bonn U., **2014**, http://hss.ulb.uni-bonn.de/2014/3821/3821.htm.