Doctoral Thesis

Structure of String Landscape on Toroidal Orientifolds and Its Phenomenological Implications

Keiya Ishiguro*

School of High Energy Accelerator Science,
Department of Particle and Nuclear Physics
The Graduate University for Advanced Studies, SOKENDAI

March, 2024

*ishigu@post.kek.jp

Contents

Chapter 1

Introduction

1.1 Exploring String Landscape

String phenomenology [1] is one of the most challenging corners of modern physics. String theory, which is the theory including one-dimensional "string" as perturbatively fundamental dynamical objects, is an anticipated candidate for the theory of everything. The theory is presumed to reside at an ultraviolet energy scale as high as the reduced Planck scale $M_{\mathrm{pl}} \sim 10^{18}$ GeV and to complete low energy effective theories into the ultraviolet scale. On the other hand, there is the Standard Model (SM) for particle physics as the best knowledge established by exceptionally precise experiments [2]. Regarding the electro-weak scale is around 10^2 GeV, one has to develop theories that fill the vast "dessert" spanning as much as $\sim 10^{16}$ GeV. This attempt to find the theories beyond SM has been performed from both bottom-up and top-down approaches. The latter one, which pursues SM or its extension from string theory viewpoints, is called string phenomenology. Since superstring theory predicts 10-dimensional (10d) spacetime, one often starts phenomenology by compactifying the extra 6d space to focus on the 4d spacetime. If the geometrical treatment is possible, one may obtain the 4d field theory via the usual Kaluza-Klein method [3, 4]. The resultant 4d theory is expected to include SM or its extension in a higher energy scale.

Then, is it easy to obtain SM from string theory? Has the rigorously phenomenological theory already been extracted from string theory? The answer is simply no. It requires strenuous efforts to be made, and indeed relevant studies date back to the 1980s (see for the milestones, e.g., [5–14]) with a focus on heterotic theory. Although significant progress including the discovery of D-brane [15] and the success of type II flux compactifications (see for very early works, e.g., [16, 17], for type IIB/F-theory flux compactifications, [18–21], and for reviews, [1, 22–24]) but there is still no known theory that is completely reliable and consistent with the experimental results while being originated from string theory. The set of low-energy effective theories which can be embedded in string theory is called the **string landscape**. The field of study is still aiming to state that the landscape contains the theory of our world.

What makes matters worse is the fact that the landscape is estimated to be tremendous. As there exist several types of string theory, namely heterotic $E_8 \times E_8$, heterotic $SO(32)$, type I, type IIA, and type IIB, one studies compactifications and resultant low energy effective field theories for each of them while the different corners of string theory are connected via the duality web [25]. Even if it is limited to discussions of effective theories based on a particular string theory, one still finds enormous degrees of freedom in choosing the topology of extra dimensions, brane configurations, and other external backgrounds including fluxes. Indeed, in the context of type II **intersecting D-brane models** (see early work of [26], and for reviews, [1, 23, 27, 28]), massless spectra of chiral fermions are determined by the brane configurations. The fluxes on backgrounds are also introduced to stabilize **moduli fields**, which are originally massless scalar fields, and parameterize deformations of the backgrounds. Since such massless scalar fields can mediate a fifth force that is severely constrained by experiments [29, 30], one has to stabilize them and give appropriate masses, i.e., in the context of **flux compactifications**, the string landscape is associated with a set of vacua of stabilized moduli fields and fluxes. The well-known result for counting the number of flux vacua predicts $\mathcal{O}(10^{272{,}000})$ vacua emerge from F-theory on a single elliptic four-fold [31]. Furthermore, it is only a fraction of the whole moduli space which can be addressed in general string compactification approaches. Although the recent studies on the swampland conjectures (see for a review, [32]) have proposed and partially proved that the string landscape is finite [33–35], the question of how we can choose a theory corresponding to our world in a natural way[1] [37] (see for relevant reviews, [38, 39]) from the string landscape remains since the fundamental degrees of freedom which label elements of the landscape e.g., fluxes cannot be fixed in the current framework.

Under these circumstances, one of the appropriate attitudes to deal with the landscape is to rely on **statistics**. Investigating what theories are contained in the vast landscape and clarifying the structures of the landscape through statistics will shed light on the study of string phenomenology. Indeed, many studies have been accomplished in this way (see for early works, e.g., [40–43] and an early review paper, [44]). Our contributions which will be summarized in this thesis are also aiming to be a part of those statistical studies, especially focusing on the type IIB flux landscape and type IIA intersecting D-brane models. In the latter models, we will challenge ourselves to partially go beyond traditional statistical research methods by using **machine learning** (see for a review of its applications on string phenomenology, [45]) and indeed find that a physical quantity characterizes the landscape.

We would like to close this section by noting a new notion for beyond SM physics suggested in recent bottom-up literature, namely **modular flavor symmetry** [46] (see [47–55] for other earlier studies and reviews, [56, 57]). The flavor symmetry has been discussed since the 1970s to explain the origin of mixing angles and mass hierarchies among fermions. Traditionally, non-Abelian symmetries are pointed out to be confident candidates for those symmetries, which include S_3, S_4, A_4, and A_5 for example. Notice that these finite non-Abelian symmetries are obtained as finite subgroups of the $SL(2, \mathbb{Z})$ modular symmetry. Indeed, the author of [46] investigated the possibility of describing the neutrino masses by assuming the modular symmetry as the flavor sector, and found the assumption brought a systematic framework for the flavor physics and strong predictive power. It is indeed surprising. The low-energy effective field theories from string theory often exhibit modular symmetry, which is associated with the moduli fields. Let us recall that the moduli fields can be stabilized by introducing the background

[1] Although there are criticisms, it is claimed that the numerous landscape does not mean the loss of predictivity [36].

fluxes. The stabilization indicates a breaking of the modular symmetry. Then, an important question arises: *Is it possible to embed the modular flavor symmetric models into the string landscape?* If this is the case, one shall find a bridge between the bottom-up and top-down approaches. Note that the modular flavor symmetry indeed happens in Super Yang-Mills theory with magnetized extra tori[2], and it has been intensively studied (see for recent study, e.g., [65]). For string compactifications, there are systems [66,67] where the Yukawa matrices are described using the theta functions on type II orientifold and it supports the embeddings. Indeed, many modular flavor symmetric models are also derived in the context of magnetized D-brane models. Although we do not address the concrete embeddings and incorporation with other problems e.g., moduli stabilizations in this thesis, we simply assume that the embedding can be done and discuss the implications on the modular flavor symmetry from the string landscape as one of the main topics.

This thesis is mainly based on our four studies [68–71]. The first three studies investigate distributions of moduli vacuum expectation values on toroidal orientifolds, within the framework of type IIB flux compactifications. In particular, by constructing the whole flux vacua yields a concept of the probability on the landscape. As long as the flux quanta are effectively treated as free parameters, the higher probability predicts that the corresponding vacuum is likely to be chosen among the landscape. Type IIA perspectives are considered in the last one. There, a certain machine learning approach is adopted to analyze hidden structures in the landscape of Intersecting D-brane models. While we briefly touched on those topics in Section 1.1, we will address them again in each corresponding section. Conclusions are added at the end of each chapter in order to avoid possible duplication. Besides the main papers, as our other papers [58, 59, 69, 72–75] are related to our topic of interest, we sometimes refer to them.

[2]On the higher dimensions tori or Calabi-Yau three-folds, the modular symmetry is promoted to the symplectic group. Correspondingly, the modular form is promoted to the Siegel modular form. See for this kind of discussion, e.g., [58–64].

Chapter 2

Type IIB Finite Flux Landscape, Symmetry, and Statistics

In this chapter, we start Section 2.1 with a short review of the status of type IIB flux compactifications by focusing on string phenomenology. Formulations that are relevant to our work are summarized in Section 2.1.1 and Section 2.1.2 while we concentrate on orientifold compactifications with Calabi-Yau three-folds in the latter section. We also mention known difficulties and recent progress relevant to our purpose in Section 2.1.3. Then we briefly review the earlier work on the statistics of flux landscape in Section 2.1.4. We see a quantity that is called **tadpole charge** as well as the **modular symmetry** plays a key role in these sections.

After that, we move to describe our works [68–70] which are relevant to the statistical approach. We focus on the toroidal orientifolds and obtain the finite flux landscape explicitly for each orientifold we consider. In Section 2.2, we consider the type IIB flux compactification on $T^6/(\mathbb{Z}_2 \times \mathbb{Z}_2')$ orientifold, which is the simplest example of the toroidal orientifolds. The distributions of moduli vacuum expectation values (VEVs) are shown there, and we discuss their implications on the modular flavor symmetric models and the CP violation in low energy effective field theory. Taking into account that the deviations from the fixed points are phenomenologically important, in Section 2.3 we further consider the KKLT-type non-perturbative effect and the uplifting term to stabilize the Kähler moduli. One of the drawbacks of the procedure is that the complex-structure moduli and the axio-dilaton slightly deviate their values from the initial VEVs, but we treat the deviations as the source of the residual flavor symmetry. We observe the distributions of moduli VEVs are slightly modified by these effects, and obtain an explicit distributions of the deviations. Finally, we consider other toroidal orientifolds in Section 2.4, namely $T^6/\mathbb{Z}_{6-\mathrm{II}}$ and $T^6/\mathbb{Z}_2 \times \mathbb{Z}_4$ orientifolds on certain lattices. We again show the distributions of VEVs and compare them to the previous cases. This discussion implies that criteria for choosing toroidal orientifolds must exist if the modular flavor symmetric models are the case and they are embedded into string theory.

2.1 Type IIB String Landscape from Flux Compactifications

NOTE: This section (Section 2.1) is fully dedicated for a review.

2.1.1 Type IIB Flux Compactifications

Preliminaries: type IIB string theory
Before describing the type IIB flux compactifications, let us firstly review type IIB superstring theory itself [1, 76–79]. Superstring theory is formulated in terms of supersymmetric conformal field theory (SCFT) on a two-dimensional (2d) worldsheet. The worldsheet represents the trajectory of a string $X(\tau, \sigma)$, which is the fundamental 1d dynamical object with its undetermined scale denoted as $l_s = 2\pi\sqrt{\alpha'}$. For string theory which contains open strings as well as closed strings, the worldsheets of free strings are classified into two types depending on its topology, while it can take more complex forms when interactions of the strings are considered. A coupling constant of strings, namely g_s is the other free parameter of string theory. As strings propagate in a background spacetime, the worldsheet of them and superstring theory as a 2d SCFT are embedded into the spacetime, i.e., the worldsheet is considered as its submanifolds and string theory describes the physics of submanifolds in the target space. It is well-known but very surprising that superstring theory itself requires its target space to be a ten-dimensional (10d) spacetime by consistency.

In 10d spacetime, superstring theory has several consistent formulations including type IIA/IIB superstring theory which we concentrate on in this thesis. Let us turn back to the worldsheet description where the action of string theory is defined to clarify the differences between them. Since we assume superstring theory, both bosonic and fermionic degrees of freedom (d.o.f.) for closed strings exist there (type II string theory is peturbatively formulated by closed strings on the worldsheet). The equation of motion for them turns out to be a set of 2d wave equations (Dirac equation for fermionic ones). Hence, as usual, the solution for it decomposes into independent left-moving and right-moving sectors. By analogy with quantum mechanics, one can expand each mode and impose quantization conditions. Then, we can define Hamiltonian and quantum states in terms of the oscillation modes and oscillation mass as well as energy. The states transform into representations of $SO(8)$ when the state is massless and of $SO(9)$ when it is massive, and we can relate the states to 10-dimensional fields in the target space. The effect of the formulation on a worldsheet appears a severe constraint on the partition function, namely modular invariance from the worldsheet topology. Indeed, the modular invariance plays a crucial role in the formulation of string theory, e.g., anomaly cancellation.

In this procedure, there remains some unconstrained d.o.f. in boundary conditions of fermionic partners of closed strings, which are classified into periodic (Ramond: R) condition and anti-periodic (Neveu-Schwartz: NS) one. As one can treat the left-moving and the right-moving sector independently, the states are decomposed into four sectors, namely NS-NS, NS-R, R-NS, R-R ones. The difference between type IIA and type IIB superstring theory is determined by the difference in the way of building the modular invariant partition function.

The modular invariance indeed requires the four sectors to co-exist but leaves two ways to define the invariant partition function for each chirality. Modding out by the reflection, there remain exactly two consistent formulations of the theory, and those correspond to type IIA/IIB superstring theory. This assembly of the left- and right- states are interpreted in terms of the Gliozzi-Scherk-Olieve (GSO) projection [80, 81] on the Hilbert space, and it leads spacetime fermions and eliminates possible tachyons from the theory.

The resultant bosonic part of the massless spectrum for type IIB superstring theory is summarized as follows in terms of the 10d fields:

$$\{\phi, B_{MN}, G_{MN}, C_0, C_{MN}, C_{MNPQ}\}, \tag{2.1}$$

where ϕ is a dilaton scalar field, B_{MN} is a rank two antisymmetric tensor, G_{MN} is also a rank two tensor (which is called a graviton) but traceless and symmetric, and C_0, C_{MN}, C_{MNPQ} are a pseudo-scalar, rank two, and rank four antisymmetric tensors, respectively. The tensors we introduced are naturally associated with differential forms, such as G, B_2, C_0, C_2, and C_4. A field strength of the four-form C_4 is defined to be self-dual in the Hodge star sense. Note that, ϕ, B_{MN}, G_{MN} are in NS-NS sector and C_0, C_2, C_4 are in R-R sector. Hence, in this thesis, we call fluxes induced by B_2 and C_2 NS-NS-(NS) and R-R-(RR) fluxes, respectively. Note that, we can safely concentrate on these massless states for our purpose since the string scale is considered to be very massive and only the massless states contribute to low-energy effective field theory.

For later convenience, let us mention also the massless spectrum for type IIA superstring theory. Instead of that for type IIB, we have odd-rank antisymmetric tensors in its R-R sector, namely C_M and C_{MNP}. Thus corresponding field strengths have always even orders as differential forms. Note that 10d $N = 2$ supersymmetry corresponds to the both spectra.

Type IIB string theory on Calabi-Yau orientifolds

So far, we shortly reviewed the world-sheet formulation of type IIB superstring theory and its 10-dimensional massless spectrum. While string theory is a consistent theory in the sense of mathematics, we expect it to be a physical theory which is a candidate for quantum gravity. Hence, the predicted 10d spacetime must include our 4d spacetime, nevertheless, the current experimental results do not observe the existence of such extra dimensions. One should deal with the extra dimensions in accordance with the experimental observation. Compactifying the extra dimensions into space as tiny as experimentally undetectable is one of the candidate treatments, and that is called string compactification. The number of supersymmetries (SUSYs) is also an important constraint on the compactification. The breaking of SUSY is known to be related to the holonomy of compactified space, as follows. In the 10d theory, a supercharge of the 4d supersymmetry is defined globally on the extra dimensions. As the supercharge has an index of 6d spinor, it means the supercharge must be a Killing spinor of the 6d space. Note that the Killing spinors are defined as exactly constant spinors with regard to parallel transportations in the literature. In particular, a parallel translation along a closed curve is characterized by the holonomy (group) of the extra dimensions. When type II string theory is compactified on 6d tori, the so-called maximal $N = 8$ supersymmetries remain in the resultant 4d effective action. In the following, we concentrate on the type IIB string theory that is compactified on a (complex) three-dimensional Calabi-Yau manifold [8], which equips $SU(3)$ holonomy and its orientifold projected theory, which is known to give a 4d chiral $N=1$ theory. Our starting point is the bosonic part of 10d type IIB low-energy effective action from the massless states, which is represented in the 10d Einstein frame [1] (see also [21]):

$$\begin{aligned} S_{\mathrm{IIB,10d}} &= S_{\mathrm{NS,R}} + S_{\mathrm{CS}} + S_{\mathrm{local}} \\ &= \frac{1}{2\kappa_{10}^2}\int d^{10}x\sqrt{-g}\left(R - \frac{\partial_M S\partial^M \bar{S}}{2\mathrm{Im}S} - \frac{1}{2}|F_1|^2 - \frac{|G_3|^2}{2\mathrm{Im}S} - \frac{1}{4}|\tilde{F}_5|^2\right) \\ &+ \frac{1}{2\kappa_{10}^2}\frac{1}{4i\mathrm{Im}S}\int C_4\wedge G_3\wedge \bar{G}_3 + S_{\mathrm{local}}, \end{aligned} \tag{2.2}$$

with $2\kappa_{10}^2 = (2\pi)^7\alpha'^4$ and non-vanishing NS-, RR- fluxes H_3 and F_3. Initially, they are defined as $F_3 = dC_2$ and $H_3 = dB_2$. Other notations are as follows. The complexified axio-dilaton S is determined as $S := a + ie^{-\phi}$ where $a := C_0$ is axionic scalar and ϕ is the dilaton field. G_3 is a linear combination of the background three-form fluxes, $G_3 = F_3 - SH_3$. We do not specify S_{local} at this time, but it is a set of contributions from possible localized sources. The p-form fields C_p have their gauge transformations $C_p \to C_p + d\lambda_{p-1}$ and thus these fields are actually the generalized versions of gauge potential. Then, we may consider the electrical coupling between C_p and a submanifold of 10d spacetime. $|\cdot|^2$ is defined with the Hodge operator, while it includes the factorial $1/n!$. $\tilde{F}_5$ is defined as a self-dual five-form as mentioned i.e., $\star_{10d}\tilde{F}_5 = \tilde{F}_5$ with $\star$ being a 10d Hodge star operator, and $\tilde{F}_5$ is not a $F_5 = dC_4$ but

$$\tilde{F}_5 := F_5 - \frac{1}{2}C_2\wedge H_3 + \frac{1}{2}B_2\wedge F_3. \tag{2.3}$$

S-duality

One may find that there is a universal $SL(2,\mathbb{R})$ symmetry where S and F_3, H_3 in 2.2 transform as:

$$S \to \frac{aS+b}{cS+d}, \quad \begin{pmatrix} F_3 \\ H_3 \end{pmatrix} \to \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} F_3 \\ H_3 \end{pmatrix} \tag{2.4}$$

with $a,b,c,d \in \mathbb{R}$ which satisfies $ad - bc = 1$. In fact, a certain subgroup of $SL(2,\mathbb{R})$, namely $SL(2,\mathbb{Z})$ is considered to be a self-duality of type IIB string theory itself and is called S-duality (see also [82]). Under the duality, the string coupling g_s is transformed into g_s^{-1}, i.e., the duality connects the weak and the strong coupling limits of the theory.

Since this symmetry exactly matches that of a complex-structure modulus on a torus, which parameterizes its "shape", thus it leads to a certain geometrical unification in the theory. It is called F-theory, and it turns out to provide a certain non-perturbative description of type IIB string theory.

Orientifold projection

Let us move to review the orientifold projections on Calabi-Yau (CY) in type II context [1,23]. As we mentioned in Chapter 1, we have to compactify the theory, and the favorable $N=1$ SUSY requires internal manifolds to equip $SU(3)$ holonomy. The CY threefolds and toroidal orbifolds, which are the tori modded by some appropriate discrete symmetries [9, 10] are the known candidates as the extra dimensions. However, it is known that 4d $N = 2, 8$ SUSY remains for CY and toroidal compactifications, respectively. One should find a remedy to truncate the extra SUSYs. Indeed, the orientifold compactifications, which are the compactifications modded out by an additional so-called orientifold projection, enable us to truncate them desirably. Although we skipped the description of fermionic massless states, the identification under the worldsheet parity halves the number of gravitino (as states in the NS-R and R-NS sectors linearly combined to be invariant). Furthermore, it also leads to a new local and non-dynamical object which is known as the orientifold plane (O-plane). The O-planes have opposite charges in their signs compared with that of D-branes, and they crucially contribute to a consistency condition, which is known as the tadpole cancellation condition. Although we consider only toroidal orientifolds, which are the orientifold-projected versions of the toroidal orbifold compactifications, let us focus on the CY compactifications at first since the former ones are considered as a certain (orbifold) limit of the latter ones.

The orientifold projection is defined as a "dressed" group action, $\hat{\Omega}\mathcal{R}(-1)^{F_L}$. The symbols are defined as follows. First, $\hat{\Omega}$ is a worldsheet parity operator, which interchanges the left-moving and right-moving sectors of a closed string and does two endpoints of an open string. $\mathcal{R}$ is defined as a second-order cyclic group $\mathbb{Z}_2$ action on the CY geometry. The definition of the group action $\mathcal{R}$ is defined up to restrictions imposed so that one can obtain 4d $N = 1$ SUSY. For our purpose, $\mathcal{R}$ is understood as an $\mathbb{Z}_2$ isometry of the internal space. Note that the $\mathcal{R}$ is not even needed to be a holomorphic operation on $\mathscr{M}$, where $\mathscr{M}$ denotes a CY threefold. F_L is the number of left moving fermions in spacetime, and the additional phase factor $(-1)^{F_L}$ is included if one determines the projection using a relation between type IIA/IIB orientifolds, namely the T-duality but it ensures the square of projection is exactly an identity.

Corresponding to the concrete action of $\mathcal{R}$, there emerges a set of fixed points (or, loci) under $\mathcal{R}$ on $\mathscr{M}$. The fixed locus is called the O_p-plane, where p denotes its spatial dimensions. Since $\mathcal{R}$ is defined as an action on $\mathscr{M}$, the remaining 4d spacetime, which is expected to be our 4d Minkowski spacetime, is automatically invariant under the action. Thus, the O_p-planes with $p \geq 3$ span the whole of our spacetime and are interpreted as $p-3$-dimensional object on $\mathscr{M}$.

In this chapter, we focus on the type IIB orientifold with O3-planes and D3-branes. The models can be obtained by setting $\mathcal{R}$ as the reversal of all complex coordinates of $\mathscr{M}$[1]. Then the concrete actions of $\mathcal{R}$ on geometric quantities on $\mathscr{M}$:

$$\mathcal{R}J = J, \quad \mathcal{R}\Omega = -\Omega \quad \text{(O3 orientifold projection)}, \tag{2.5}$$

where J denotes a Kähler form of $\mathscr{M}$ and Ω without a hat does a unique holomorphic three-form on $\mathscr{M}$. It is also known that the 10d dynamical objects in the action (2.2) transform

[1]More precisely, under the orientifold action the Dolbeault cohomology group $H^{p,q}$ remains invariant. As a result, one gravity multiplet, $h^{2,1}_+$ vector multiplets, $h^{2,1}_- + h^{1,1} + 1$ chiral multiplets remain in the spectrum for the case. Note that the systems with both O3/O7 (and D3/D7) also exist. For the IIA case, it transforms $H^{p,q} \to H^{q,p}$.

under the $\Omega'(-1)^{F_L}$ as (nicely summarized in [83]):

$$\begin{aligned}\Omega'(-1)^{F_L}B&=-B, \quad \Omega'(-1)^{F_L}C_{2p}=(-1)^pC_{2p},\\ \Omega'(-1)^{F_L}G&=+G, \quad \Omega'(-1)^{F_L}\phi=+\phi.\end{aligned} \tag{2.6}$$

We would like to emphasize again that the theory is indeed projected out by this symmetry, but not the internal space itself. Hence, we can introduce the gauge field strengths of these odd potentials.

Introducing three-form fluxes

The background three-forms, F_3 and H_3, are quantized on cycles of an internal manifold as

$$\frac{1}{(2\pi)^2\alpha'}\int_C F_3=f_C\in\mathbb{Z}, \quad \frac{1}{(2\pi)^2\alpha'}\int_C H_3=h_C\in\mathbb{Z}. \tag{2.7}$$

If there are no sources, the equations

$$dF_3=0, \quad dH_3=0 \tag{2.8}$$

become the equations of motion. We consider the compactifications with these equations (which are consistent with $d^2=0$, hence F_3,H_3 can be expanded by a cohomology basis) and $F_1=0$ throughout this thesis. However, the presence of Chern-Simons action $S_{\rm CS}$ leads us to the 10d condition which is known as the tadpole cancellation condition, as follows. With no $S_{\rm local}$, the Bianchi identity for $\tilde{F}_5$ turns out to be

$$d\tilde{F}_5=H_3\wedge F_3, \tag{2.9}$$

and we define

$$N_{\rm flux}=\frac{1}{(2\pi)^4\alpha'^2}\int_{\mathscr{M}} H_3\wedge F_3. \tag{2.10}$$

$\mathscr{M}$ denotes an internal 6d manifold. Here and in what follows, we set $l_s=\frac{1}{(2\pi)\sqrt{\alpha'}}=1$ (or we can change the mass dimensions of F_3,H_3 alternatively) and $M_{\rm pl}=1,\kappa_4=1$. The integral of the Bianchi identity makes $N_{\rm flux}$ to be zero. In the moduli stabilization with $N_{\rm flux}=0$ at least SUSY Minkowski vacua cannot be achieved [72]. Then, we would like to consider the theory including sources to cancel $N_{\rm flux}$ nontrivially. Indeed, the orientifold compactifications provide a natural source of (RR) charge which enters in the Bianchi identity $d\tilde{F}_5$, and that is the O-plane (and D-branes).

D-branes and O-planes

Although type II string theory is constructed by the worldsheet closed strings, the theory turns out to possess a certain object where open strings whose boundaries are tied, which are called D-branes [15]. We skip the detail, but one can define a state of D-brane as a coherent state in terms of Boundary CFT (BCFT) [84], and it consists of oscillators and modes associated with closed strings. Then one can define an open string $X(\tau, \sigma)$ whose length of l and both endpoints are on the D_p-brane (spanning $(p+1)$-spacetime), with imposing the Neumann-Neumann (NN) boundary condition (this is satisfied by a free open-string with no presence of D-branes) $\partial_\sigma X^I(\tau, \sigma)|_{\sigma=0,l} = 0$ and the Dirichlet-Dirichlet (DD) boundary condition $\partial_\tau X^T(\tau, \sigma)|_{\sigma=0,l} = 0$ on each of the 2d worldsheet string with its indices directing spanned space (I) and transverse space (T). From the equation of motion, one finds that the open string can be exactly defined and propagates the $(p+1)$-spacetime (worldvolume) spanned by the D-brane. Thus, the usual procedure leads us to the massless spectrum of the open strings, but the GSO projection allows only odd p for type IIB string theory whereas only even p is allowed for type IIA. The massless spectrum contains a gauge boson, (real) scalar fields, and fermions, and importantly, the massless spectrum forms a $U(1)$ vector multiplet in the worldvolume. The presence of D-branes is a key point in considering phenomenology from type II string theory.

The effective action of D-branes is split into two pieces, where the first one called the Dirac-Born-Infeld (DBI) action describes couplings to the NS fields while the other, the Chern-Simons (CS) action describes those to the RR fields, which becomes the source of the RR charge. Indeed, this CS action includes the term (we use the notation in [1])

$$S_{\mathrm{CS},D_p} \supset \mu_p \int_{W_{p+1}} C_{p+1}, \tag{2.11}$$

and $p=3$ case this contributes to the Bianchi identity. W_{p+1} represents worldvolume spanned by the D_p-brane, and μ_p denotes the RR charge of the D_p-brane.

The O_p-plane is known to have its CS action S_{CS,O_p}, and the CS action is the same as that of D_p brane, up to constant multiplication in the charge μ_p. It turns out the constant factor is negative compared to the charge if there remains a spacetime supersymmetry, and it is a constant depending on p.

RR tadpole cancellation condition

The N_{flux} is called the D3 charge of fluxes since localized sources such as D-branes and O-planes contribute to this Bianchi identity as follows. Suppose there are N_{D3} D3-branes and N_{O3}-planes in 4d spacetime. Since the four form C_4 can couple to such objects, we obtain the following contribution

$$S_{\text{local}} \supset \int_M \frac{1}{2}\left(N_{\mathrm{D3}} - \frac{1}{2}N_{\mathrm{O3}}\right) C_4, \tag{2.12}$$

with M being $\mathbb{R}^{1,3}$. Here, the numerical factors reflect the relative factor mentioned above. This is just a generalization of the electric coupling of a gauge field. Then the Bianchi identity becomes

$$d\tilde{F}_5 = H_3 \wedge F_3 - 2\left(N_{\mathrm{D3}} - \frac{1}{2}N_{\mathrm{O3}}\right)\delta(M) \tag{2.13}$$

with $\delta(M)$ being a Poincaré dual of M. Integrating both sides of the equation, we obtain the cancellation condition of RR charges, which is called the RR tadpole cancellation condition[2]. Indeed, this important consistency condition plays a key role in the whole of our thesis.

Warped compactification ansatz

Before moving to 4d discussion, we review the compactification ansatz for supergravity solution with the presence of the flux background in the manner of Giddings-Kachru-Polchinski (GKP) [20]. We concentrate on the solution which maintains $N = 1$ SUSY. The favorable 4d Poincaré invariance leads us to parameterize the Einstein frame metric as

$$ds^2 = e^{2A(y)}\eta_{\mu\nu}dx^\mu dx^\nu + e^{-2A(y)}\tilde{g}_{mn}dy^m dy^n, \tag{2.14}$$

where $A(y)$ is called a warp factor, and we consider a compactification with a warped product based on $M \times \mathscr{M}$. x and y are the coordinates for M and $\mathscr{M}$, respectively. $\tilde{g}_{mn}$ is a metric on the internal manifold $\mathscr{M}$. Although this discussion holds without specifying $\mathscr{M}$, in this thesis, we mainly focus on the case $\mathscr{M}$ is a CY orientifold (or, toroidal orientifold). The 4d Poincaré invariance also requires $\tilde{F}_5$ to take the form:

$$\tilde{F}_5 = (1 + \star_{10d})(d\alpha \wedge dx^0 \wedge dx^1 \wedge dx^2 \wedge dx^3), \tag{2.15}$$

where $\alpha = \alpha(y)$ is a function on the internal space and is defined up to a constant. Then the Einstein equation and the Bianchi identity for $\tilde{F}_5$ are reduced to a pair of Laplace equations for A and α. These Laplace equations contain the contributions from the fluxes G_3, and thus the warped factor generally depends on the fluxes and the fluxes control the warp factor. From an integrability condition, one finds

$$\alpha = e^{4A}. \tag{2.16}$$

The physical meaning of this equation is that the tension and charge of fluxes are coincide if (2.16) is satisfied. Moreover, under this circumstance one find that the three-form flux is imaginary self-dual (ISD), namely

$$\star_{6d}G_3 = iG_3. \tag{2.17}$$

The ISD condition (2.17) restricts N_{flux} to be nonnegative and thus negatively-charged objects are needed. This motivates us to introduce O3-planes, and the cancellation condition is obtained as we reviewed above. Note that, derivation of the above results indeed needs the detail of localized sources. The authors of [20] pointed out that O5 planes can violate an assumption needed. We further restrict ourselves to a large-volume regime of the CY where the effective field theory is trusted and the compactification is nearly identified with the compactification on the exact CY.

[2] Of course, the considered condition here is only for the charge carried by the D3-branes. In general, one finds the tadpole cancellation conditions for each possible D_p-branes.

Four-dimensional effective action and moduli

Next, we move to the four-dimensional properties of the effective theory (see for details, [21,85]). Since there remains the N = 1 SUSY, the resulting theory can be described in terms of $N = 1$ SUGRA. The ingredients that constitute the SUGRA are (i) the Kähler potential K, (ii) the superpotential W and (iii) gauge kinetic functions f. These are the functions of chiral superfields Φ, but dependencies differ from each other. The superpotential and gauge kinetic functions are holomorphic functions of Φ and thus, complex functions. On the other hand, the Kähler potential is not a holomorphic function but a real function of Φ. As usual, the Kähler potential and the superpotential control kinetic terms and mass terms of Φ in the 4d Lagrangian, respectively. By using K and W, the (F-term) scalar potential of 4d $N = 1$ SUGRA can be summarized in the form

$$V = e^K (K^{I\bar{J}} D_I W D_{\bar{J}} \overline{W} - 3|W|^2), \tag{2.18}$$

where indices I, J correspond to the chiral superfields, and $K^{I\bar{J}}, D_I$ are introduced later. The term $-3e^K|W|^2$ is the SUGRA effect, which does not exist in a global supersymmetric theory without gravity. This scalar potential can stabilize moduli fields, which we introduce in the following.

Again, we turn back to the 10d description. Initially, the moduli fields are massless scalar fields. After compactifications, these fields appear anywhere in the resulting Lagrangian reflecting the internal manifold's geometry. Precisely speaking, the moduli parameterize continuous deformations of the internal manifolds. If we adopt a CY space as an internal space, this parameterization continuously changes its form. Since CYs are defined as Kähler and Ricci-flat with vanishing first Chern class, we can formulate the condition as

$$R_{mn}(\tilde{g}) = 0 = R_{mn}(\tilde{g} + \delta\tilde{g}), \tag{2.19}$$

under the small transformation of the metric $\delta\tilde{g} \to \tilde{g} + \delta\tilde{g}$ (see for details, [86]). By using the knowledge of Kähler geometry, this turns out to be an equation which is known as the Lichnerowicz equation and indices of $\delta\tilde{g}$ are restricted to pure-type or mixed-type deformations. For the mixed-type deformation, namely $\delta\tilde{g}_{m\bar{n}}$, we can associate an element of $H^{1,1}(\mathscr{M})$ and for the pure-type deformation, namely $\delta\tilde{g}_{\bar{m}\bar{n}}$, we can do so an element of $H^{2,1}(\mathscr{M})$, where $H^{1,1}$ and $H^{2,1}$ denote the Dolbeault cohomology groups of CY. The moduli parameterize the former space and latter space are called the Kähler moduli T_i and the complex structure moduli U_i. The geometrical meanings of these moduli are that the Kähler moduli parameterize the size of CY and the complex-structure moduli does the deformations of CY. The Kähler form J of CY transforms under the pure-type metric transformation. On the other hand, we can associate the holomorphic coordinate transformations and the complex structure moduli, hence it parameterizes the deformations. For example, some of the complex structure moduli appear as degrees of freedom of defining equations of CY. In this thesis, we mainly focus on the complex structure moduli as well as the axio-dilaton.

Let us introduce the holomorphic three-form Ω which is no-where vanishing and uniquely (up to Kähler transformations which we introduce later) defined three-form. It is known that certain period integrals of Ω on CY three-cycles give well-defined coordinates of CY moduli space. Since we defined the complex structure moduli as parameters of the deformations, or equivalently, parameters of the complex structure moduli space, the complex structure moduli appear in the period integrals. Thus, Ω is an intrinsic quantity that characterizes the complex structure moduli space. Note that, $\int \Omega \wedge \bar{\Omega}$ is proportional to CY volume, which can be expressed by the Kähler form as $V \sim \int J \wedge J \wedge J$. Then, the superpotential of moduli is given by the following Gukov-Vafa-Witten type superpotential [18]

$$W_{\rm GVW} = \int_{\mathscr{M}} G_3 \wedge \Omega. \tag{2.20}$$

This form of the superpotential can be obtained in the dimensional reduction of the 10d action. We can alternatively say, if we consider the (2.20), the resulting scalar potential of moduli is organized as a form of $N = 1$ SUGRA potential (2.18). The Kähler potential is also given by

$$K = K_{\rm ad} + K_{\rm cs} + K_{\rm vol} = -\log\bigl(-i(S - \bar{S})\bigr) - \log\left(-i\int_{\mathscr{M}} \Omega \wedge \bar{\Omega}\right) - 2\log \mathcal{V}, \tag{2.21}$$

with $\mathcal{V}$ being the volume of CY and it can be expressed as a function of Kähler moduli. At large volume region, $\mathcal{V}$ has an expression $\mathcal{V} = \frac{1}{6}\kappa_{ijk}T^iT^jT^k$ where κ_{ijk} denotes a Yukawa coupling between moduli. This leads us to the no-scale structure, where F-term the scalar potential reduces to

$$V_{F,\ \text{no-scale}} = e^K (K^{I\bar{J}} D_I W D_{\bar{J}} \overline{W}), \tag{2.22}$$

with I, J being axio-dilaton and complex-structure moduli only. This is a positive-definite potential, and its minima are Minkowski, as we can see. However, since the Kähler moduli is unstabilized with the no-scale potential, we have to seek other mechanisms which stabilize that. In type IIB flux compactifications, many studies include stabilizations with non-perturbative effects [87] and α'-corrections [88, 89], and non-geometric fluxes [90, 91] have been made. For the D-term potential, we assume that is canceled via the tadpole cancellation condition.

The Kähler metric $K_{I\bar{J}}$ is defined as

$$K_{I\bar{J}} = \partial_I \partial_{\bar{J}} K, \tag{2.23}$$

with I, J being all moduli we consider. $K^{I\bar{J}}$ denotes the inverse of Kähler metric. Note that, the Kähler potential (2.21) does not contain real-parts of moduli, since those moduli enjoy the axionic symmetry. Then $K_I = \partial_I K$ becomes pure imaginary.

2.1.2 Details of The Formulation: Moduli Space with Special Geometry

We start this section by describing the moduli space of CY [86, 92] and the special geometry in detail for general purposes. The fact that we will explain below also holds on toroidal orientifolds as they are considered singular limits of smooth CYs. Then, we also review the type IIB flux compactifications with the special geometry on which the chapter is based.

Special geometry on Calabi-Yau threefolds

CY threefolds and CFT whose central charge $c = 9$ admit the geometrical description of their moduli space, which is called special geometry [86, 93]. Since those backgrounds play crucial roles in string compactifications, we need to understand the structure of special geometry.

With special geometry, there exists the symplectic basis $(A^a, B_b) \in H_3(\mathscr{M}, \mathbb{Z})$ which satisfies

$$A^a \cap B_b = \delta^a_b, \quad B_b \cap A^a = -\delta^a_b, \tag{2.24}$$

$$A^a \cap A^b = 0, \quad B_a \cap B_b = 0, \tag{2.25}$$

where $a, b = 0, 1, \ldots, h^{2,1}$. $h^{2,1}$ denotes the dimension of $H^{2,1}(\mathscr{M})$ i.e. one element of the Hodge numbers. Then we can define a cohomology basis (α^a, β^b) of $H^3(\mathscr{M}, \mathbb{Z})$ so that the basis becomes the dual of (A^a, B_b) as follows.

$$\int_{A^a} \alpha_b = \int_{\mathrm{CY}} \alpha_b \wedge \beta^a = \delta^a{}_b, \quad \int_{B_a} \beta^b = \int_{\mathrm{CY}} \beta^b \wedge \alpha_a = -\delta^b{}_a, \tag{2.26}$$

i.e. $\{\alpha_a\}$ and $\{\beta^b\}$ are Poincaré dual. As we mentioned, the nowhere-vanishing three form Ω is uniquely defined on each of the CY manifolds. We can think Ω as a function of the complex structure moduli and space coordinates of CY. In other words, Ω should transform under the holomorphic coordinate transformations, which are parameterized by the complex structure moduli. It is even known that the (half of) period integrals of Ω span the moduli space:

$$F_a := \int_{B_a} \Omega, \quad U^a := \int_{A^a} \Omega, \tag{2.27}$$

where we take $\{U^a\}$ as coordinates of the moduli space. The number of $\{U^a\}$ is one more than the dimension of the moduli space, $h^{2,1}$. Indeed, $\{U^a\}$ are projective coordinates of the moduli space i.e. only the ratios between $\{U^a\}$ matter. Note that, derivatives by these moduli and integrals on CY coordinates (which we write $\mathscr{M}$ explicitly) are interchangeable. We gather these period integrals into one vector:

$$\Pi \equiv \begin{pmatrix} F_a \\ U^a \end{pmatrix} \equiv \begin{pmatrix} \int_{B_a} \Omega \\ \int_{A^a} \Omega \end{pmatrix}. \tag{2.28}$$

Π is called the period vector and its length is $(2h^{2,1} + 2)$. Then, Ω is expanded as

$$\Omega = U^a \alpha_a - F_a \beta^a. \tag{2.29}$$

Since we can consider $\{U^a\}$ as coordinates of the moduli space, F_a must be functions of $\{U^a\}$. We can introduce the prepotential, which is defined as

$$F = \frac{1}{2} U^a F_a, \tag{2.30}$$

thus F_a are expressed as $F_a = \partial_a F$. Note that $u^0 F_0 = 2F - u^i F_i$ holds by the definition. The prepotential has an expansion around the large complex structure (LCS) regime ($\mathrm{Im} U^a \gg 1$) as [94]

$$\frac{F}{(U^0)^2} = \frac{1}{6} \kappa^*_{abc} \frac{U^a U^b U^c}{(U^0)^3} + \frac{1}{2} \kappa^*_{ab} \frac{U^a U^b}{(U^0)^2} + \kappa^*_a \frac{U^a}{u^0} + \frac{1}{2} \kappa_0 + F_{\mathrm{inst.}}, \tag{2.31}$$

where the coefficients are given by

$$\kappa^*_{abc} = \int_{\mathscr{M}^*} J^*_a \wedge J^*_b \wedge J^*_c, \quad \kappa^*_{ab} = \frac{1}{2} \int_{\mathscr{M}^*} J^*_i \wedge J^{*2}_j, \tag{2.32}$$

$$\kappa^*_a = -\frac{1}{24} \int_{\mathscr{M}^*} c_2(M^*) \wedge J^*_i, \quad \kappa_0 = -\frac{\zeta(3)\chi(\mathscr{M}^*)}{(2\pi i)^3}. \tag{2.33}$$

Here, the symbol * denotes a mirror manifold of $\mathscr{M}$, and J^*_i are (1, 1)-forms on $\mathscr{M}^*$. c_2 and χ are the second Chern class and the Euler number, respectively. $F_{\text{inst.}}$ contains geometrical corrections and it is expressed as a power series of $q_a = e^{2\pi i U_a}$. Note that on a toroidal orientifold that is at its orbifold limit and has no twisted moduli contributions, the first term of the above expansion only contributes to the exact prepotential.

In this thesis, it is sufficient to introduce only the relationship between relevant differential forms of CY. Indeed, the Kodaira theory [86, 95] leads

$$\frac{\partial \Omega}{\partial \tilde{u}^a} \in H^{3,0}(\mathscr{M}, \mathbb{C}) \oplus H^{2,1}(\mathscr{M}, \mathbb{C}). \tag{2.34}$$

Here, $\{\tilde{U}^a\}$ are defined as

$$\tilde{U}^a = \frac{U^a}{U^0} \ (a \neq 0) \tag{2.35}$$

on the local patch $U^0 \neq 0$ and are called flat coordinates and usually called the complex-structure moduli fields[3]. This implies that Ω can be expressed as

$$\frac{\partial \Omega}{\partial \tilde{U}^a} = k_a \Omega + \chi_a, \tag{2.36}$$

where $\chi_a \in H^{2,1}(\mathscr{M}, \mathbb{C})$ denotes a harmonic (2, 1)- form of CY. The term $k_a\Omega$ is a result of the uniqueness of Ω. Note that this yields an alternative expression of the Kähler metric,

$$K_{a\bar{b}} = -\frac{\int \chi_a \wedge \bar{\chi}_{\bar{b}}}{\int \Omega \wedge \Omega}. \tag{2.37}$$

Precisely speaking, the Ω is not exactly unique, but a holomorphic section of $\mathcal{H} \otimes L$. Here, $\mathcal{H}$ is a $Sp(2n+2, \mathbb{R})$ vector bundle which is associated with the degree of freedom of choosing basis. L denotes a line bundle which is associated with the Kähler transformation:

$$\Omega \to f(U)\Omega. \tag{2.38}$$

[3]In toroidal compactifications, the flat coordinates are exactly τ-parameters, which characterize the lattice constitutes the torus. However, on general CY, the flat coordinates do not take simple forms but are expressed as functions of several variables. Then the variables are often called moduli fields, thus the flat coordinates do not always correspond directly to physical moduli fields.

Then, the $\{\Omega, \chi_a, \bar{\chi}_{\bar{b}}, \Omega\}$ with $a, b = 1, \ldots, h^{2,1}$ constitute the harmonic basis of $H^3(\mathscr{M}, \mathbb{C})$. There are relationships between the elements of this basis by differentiation. One of the relationships is (2.36). Since Ω is defined only up to the bundle, it is more natural to introduce Kähler covariant derivatives:

$$D_a\Omega = (\partial_a + K_a)\Omega, \tag{2.39}$$

with $\partial_a \coloneqq \frac{\partial}{\partial u^a}$ and $K_a \coloneqq \partial_a K$. The covariant derivative of Ω transforms as

$$D_a\Omega \to f(U)D_a\Omega \tag{2.40}$$

under the Kähler transformation, hence this is exactly the covariant derivative. This covariant derivative is generalized as usual,

$$D_a\psi^{(m,n)} = (\partial_a + mK_a)\psi^{(m,n)}, \tag{2.41}$$

$$D_{\bar{b}}\psi^{(m,n)} = (\partial_{\bar{b}} + nK_{\bar{b}})\psi^{(m,n)}, \tag{2.42}$$

where (m, n) means that the quantity with that indices transforms as

$$\psi^{(m,n)} \to f^m \bar{f}^n \psi^{(m,n)}, \quad \psi^{(m,n)} \to f^m \bar{f}^n \psi^{(m,n)} \tag{2.43}$$

under the Kähler transformation. For tensors, we replace ∂_a by ∇_a, where connections are given by Christoffel symbols on Kähler manifolds. Note that only Γ^a_{bc} and $\Gamma^{\bar{a}}_{\bar{b}\bar{c}}$ are nonzero on Kähler manifolds. We also impose the usual metric condition on the Kähler metric $K_{a\bar{b}}$ as

$$D_c K_{a\bar{b}} = 0. \tag{2.44}$$

We summarize the relationships between the basis $\{\Omega, \chi_a, \bar{\chi}_{\bar{b}}, \Omega\}$ in terms of the Kähler covariant derivatives, as follows. As we can see, (2.36) is equivalent to

$$D_a\Omega = \chi_a. \tag{2.45}$$

Since Ω is a holomorphic function of the complex structure moduli,

$$D_a\bar{\Omega} = \partial_a\bar{\Omega} = 0 \tag{2.46}$$

also follows. To obtain the remaining expressions, we have to introduce the Yukawa coupling [86, 96] of moduli

$$\kappa_{abc} = \int_{\mathscr{M}} \Omega \wedge D_a D_b D_c \Omega. \tag{2.47}$$

The $(0, 3)$-part of $D_a D_b D_c \Omega$ can contribute to the integration and this reduces to

$$\kappa_{abc} = \int_{\mathscr{M}} \Omega \wedge \partial_a\partial_b\partial_c\Omega, \tag{2.48}$$

and hence κ_{abc} is totally symmetric quantity with respect to the change of its indices.

It is convenient to introduce the inner product $\langle \alpha, \beta \rangle$ which is defined as

$$\langle \alpha, \beta \rangle = -i \int_{\mathscr{M}} \alpha \wedge \beta. \tag{2.49}$$

Then the expression of Yukawa coupling reduces to

$$\kappa_{abc} = i \langle \Omega, \partial_a \partial_b \partial_c \Omega \rangle. \tag{2.50}$$

Eq. (2.45) means that, $D_a : H^{3,0} \to H^{2,1}$ holds. This is also the case for $D_a \chi_b, D_a \bar{\chi}_{\bar{b}}$ i.e. $D_a : H^{2,1} \to H^{1,2}$ and $D_a : H^{1,2} \to H^{0,3}$, respectively. Following the lines of Ref. [86], we obtain

$$D_a \chi_b = -i e^{K_{\mathrm{cs}}} \kappa_{ab}{}^{\bar{c}} \bar{\chi}_{\bar{c}}, \tag{2.51}$$

$$D_a \bar{\chi}_{\bar{b}} = K_{a\bar{b}} \bar{\Omega}. \tag{2.52}$$

2.1.2.1 Type IIB flux compactification with special geometry

Similarly, the background fluxes F_3, H_3 are also expanded as

$$F_3 = f^a \alpha_a - f_a \beta^a, \tag{2.53}$$

$$H_3 = h^a \alpha_a - h_a \beta^a, \tag{2.54}$$

thanks to the special geometry[4]. Here, we defined integral quanta by

$$f^a := \int_{A^a} F_3, \;\; f_a := \int_{B_a} F_3, \tag{2.55}$$

$$h^a := \int_{A^a} H_3, \;\; h_a := \int_{B_a} H_3. \tag{2.56}$$

From these quantities, we can determine the explicit form of K and W in terms of the cycles. Here, let us consider the Kähler potential as an example. Recall that the complex structure sector of Kähler potential K is given by

$$K_{\mathrm{cs}} = -\log\left(-i \int \Omega \wedge \bar{\Omega}\right), \tag{2.57}$$

and this reduces to

$$K_{\mathrm{cs}} = -\log\left(-i \Pi \Sigma \Pi\right), \tag{2.58}$$

with

$$\Sigma = \begin{pmatrix} 0 & \mathbf{1}_{h^{2,1}+1} \\ -\mathbf{1}_{h^{2,1}+1} & 0 \end{pmatrix}. \tag{2.59}$$

[4]In this thesis, we use the notation different from the literature where $F_3 = f^a \alpha_a + f_a \beta^a$ as well as H_3. Our notation matches the expansion of Ω: $\Omega = u^a \alpha_a - F_a \beta^a$.

Here, we used

$$\int \alpha \wedge \beta = -\sum_I \left(\int_{A^I} \alpha \int_{B_I} \beta - \int_{A^I} \beta \int_{B_I} \alpha \right) \tag{2.60}$$

$$= \left((\textstyle\int_{B_I} \alpha) \quad (\int_{A^I} \alpha) \right) \Sigma \begin{pmatrix} (\int_{B_I} \beta) \\ (\int_{A^I} \beta) \end{pmatrix} \tag{2.61}$$

holds for $\alpha, \beta \in H^3(\mathscr{M}, \mathbb{Z})$.

Note that the Kähler potential is invariant under the symplectic transformation of the cohomology basis. We can obtain the explicit form of K_{cs} by using the period (2.28). In general, the explicit form of the period vector highly depends on geometry (see for example [97,98] where the integral is explicitly calculated).

The D3-charge of fluxes N_{flux} is now given by

$$N_{\text{flux}} = \int_{\mathscr{M}} H_3 \wedge F_3 = -h^a f_a + f^a h_a. \tag{2.62}$$

Now we consider the scalar potential V (2.18) as a function of the axio-dilaton, the complex structure moduli, and the Kähler moduli. Then the moduli stabilization can be achieved by solving extremum conditions of V. For stable minima, we impose them to have positive masses (which are eigenvalues of Hessian matrices). Let us introduce a special class of minima, the SUSY minima with $D_I W = 0$ (the F-term flatness). If this condition and its complex conjugate are satisfied by all the moduli, the scalar potential is minimized. If we denote $W = W_{\text{RR}} - W_{\text{NS}}$ with

$$W_{\text{RR}} = \int F_3 \wedge \Omega, \tag{2.63}$$

$$W_{\text{NS}} = \int H_3 \wedge \Omega, \tag{2.64}$$

the $N = 1$ SUSY conditions i.e. the F-term flatness $D_I W = 0$ for the axio-dilaton and the complex structure moduli reduce to

$$D_S W = K_S W - W_{\text{NS}}, \tag{2.65}$$

$$D_a W = \int G_3 \wedge \chi_a, \tag{2.66}$$

where $D_S = \partial_S + K_S$ and a runs all of the complex structure moduli.

2.1.3 General Limitations Constrain Our Studies

We have reviewed the formulation of type IIB flux compactifications. Since this formulation can address the moduli stabilization as well as configurations of D-branes and O-planes, it has attracted the interest of many studies. Nevertheless, as mentioned in Section 1.1, we have not obtained any theory from the compactifications that are completely consistent with the experiments. This problem is often related to the fundamental difficulties that are almost inevitable. In this section, we briefly review the current difficulties and the swampland program which has attempted to make the boundary of string landscape clear. Since the issues are diverse, we will discuss them only as they pertain to our thesis.

Let us first refer to the problem of the non-existence of de Sitter vacua (see for reviews, [32, 99, 100]). Currently, it is a well-known fact that the cosmological constant has a tiny but positive value [101]. Since it is also known that type IIB flux vacua favors Minkowski or anti-de Sitter (AdS) vacua, many studies have been devoted to constructing a de Sitter (dS) vacuum in the string landscape. Methods considered there include the KKLT scenario, which uses non-perturbative effects on the superpotential [87], the Large Volume Scenario, which uses both non-perturbative effects and perturbative α' corrections [89], so-called F-term uplifting where spontaneous SUSY breaking occurs and matter fields possibly contribute to the F-term as a correction [102–104], and D-term uplifting based on magnetized D-brane models [105]. However, due to the numerous critical points of view which are summarized in [99], we may understand that there is no established consistent dS construction. This is one of the central topics of the swampland program [32, 106] that aims to theoretically characterize the boundary of string landscape. In this thesis, we will adopt the KKLT scenario to investigate the resultant deviations of already-stabilized moduli VEVs, but we do not consider further the dS property of those vacua.

Even if we tentatively escape from the dS constructions as well as Kähler moduli stabilizations, still we must be careful whether the resultant vacua are consistent in the formulation. Indeed, moduli VEVs associated with some vacua should reside in the sufficiently weak coupling regime, the large volume regime which may ensure separation between the KK scale and moduli masses, and the regime where the expansion of prepotential is reliable. The tree-level vacua always suffers from non-perturbative and perturbative corrections. The tadpole cancellation conditions need to be satisfied while it turns out to be severe. Indeed, anyone who has tried to fix moduli may find that it is difficult to stabilize many complex-structure moduli due to the tadpole conditions. This general observation is precisely formulated as the tadpole conjecture [107]. The conjecture states the (D3) tadpole charge $N_{\rm flux}$ which is contributed by flux is bounded as $N_{\rm flux} > \alpha \cdot h^{2,1}$, where $\alpha = 1/3$ is conjectured and $h^{2,1}_{(-)}$ complex-structure moduli are stabilized. As one can see from this criterion, it is difficult to analyze flux landscape with large $h^{2,1}$. In this thesis, we only consider the toroidal orientifolds whose $h^{2,1}$ is sufficiently small.

Other contributions may originated from unstabilized moduli including not only Kähler moduli but also twisted moduli and open-string moduli (see also [100]). These various contributions make it almost impossible to have a crystal clear discussion on the four-dimensional vacua. In this thesis, we indeed ignore some needed checks of these consistencies, and sometimes we intentionally violate them. This paper purely aims to investigate the structure of solutions to the F-term equations, but in a rigorous phenomenological discussion, we have to keep in mind that we need to be sufficiently careful when we would like to discuss phenomenology rigorously.

In this chapter, we will concentrate on the moduli sector of theory. As mentioned in Chapter 1, the modular flavor symmetric models attract great interest from both bottom-up and top-down perspectives. However, while we can independently design the flavor symmetry in the framework of magnetized D-brane models and stabilize the moduli, it is pointed out that performing those simultaneously turns out to be very difficult, due to the tadpole cancellation condition [108]. It is interesting to pursue that point. Indeed, a relation between generations of chiral fermions and the moduli VEVs via the tadpole cancellation condition is also found in [75].

2.1.4 Statistics of Flux Landscape

As stated in Section 1.1, we consider the statistics to be one of the safest ways to handle the vast string landscape. Indeed, many studies have been performed, and they have revealed some important structures of the flux landscape. In this section, we briefly review the established progress and check whether our direction of study in the following is correct or not.

Due to the Dirac quantization, the flux landscape is discretized and can be labeled. Then one naturally hopes to reveal distributions of some physical quantities and count the number of flux vacua. Roughly speaking, there are three ways to study statistics of flux landscape: (i) continuous flux approximation, (ii) random sampling, and (iii) exactly obtaining whole flux vacua. For the method (i), which was suggested in [109] and extensively studied in [40, 42, 43], one obtains analytic expressions formulae for the number of flux vacua, with a certain bound on the flux quanta and the continuous approximation. The bound is understood as the flux tadpole charge $N_{\rm flux}$. Note that the assumption of the continuous approximation only holds on a regime $N_{\rm flux} \gg h^{2,1}$, while it matches the claim of tadpole conjecture. The validity of the results based on the method (i) have been explicitly checked in e.g., [110–113]. The method (ii) has been evolved along with the evolution of computational machinery. In general, it is very difficult to numerically solve the F-term equations or $\partial_I V = 0$ itself when the number of moduli is large besides the tadpole problem. Even if one assumes Minkowski vacua, then one has to solve $\partial_I \tilde{V} = 0$ with $V = e^K \tilde{V}$ for all moduli, but one cannot expect an analytic solution at all. For one of the simplest toroidal orientifolds with $h^{2,1} = 3$, the equations become non-linear simultaneous sextic equations. If the superpotential is not a polynomial, the problem would be even worse. Hence there are many studies based on random sampling, and a more sophisticated framework is recently proposed [114]. At last, the method (iii) is the most fundamental approach, but it is also the hardest one. One might wonder whether the number of vacua is finite in the first place. It has been believed that it is the case [115, 116] at least for vacua consistent with the experiments. Let us recall that $N_{\rm flux}$ has a quadratic form: $N_{\rm flux} \sim h \cdot \Sigma \cdot f$. If either the RR or the NS flux is turned off for a direction, the pair of the flux seems to be allowed to take an infinite number of choices. Nevertheless, the recent progress [33–35] proved that the number is indeed finite[5,6]. Note that the above observation with zero directions is inconsistent with the tadpole conjecture. Then the next obstruction is: could we construct the finite vacua explicitly? This is actually a difficult problem, but it can be obtained at least on the toroidal orientifolds (a constructive method is exhibited in [83, 111]). More recently, the study that completely scanned the flux vacua on a certain CY appeared [113]. The method (iii) has of course the

[5] To the best of our knowledge, the finiteness is proven for SUSY vacua.

[6] For finiteness of CY three-folds, see e.g., [117, 118].

advantage that it does not need the assumption made in the continuous approximation nor does the random sampling reflect the exact structure of landscape. We adopt method (iii) to analyze the distributions of moduli VEVs and their degeneracy in the following.

2.2 Distributions and Dualities in Flux Landscape: $T^6/(\mathbb{Z}_2 \times \mathbb{Z}_2')$ orientifold

Note: The whole discussion and results in this chapter are based on our work [68]. The figures and tables would be shown are already shown in Ref. [68] whence we will cite them constantly.

The introduction for this chapter is given in Chapter 1. We study the distribution of vacua in $T^6/(\mathbb{Z}_2 \times \mathbb{Z}_2')$ toroidal orientifold compactification, where T^6 is factorizable i.e. $T^6 = T_1^2 \times T_2^2 \times T_3^2$. If we denote the complex coordinates of $T^6/(\mathbb{Z}_2 \times \mathbb{Z}_2')$ as z_1, z_2, z_3, the orbifold action is defined as $\theta : (z_1, z_2, z_3) \to (-z_1, -z_2, z_3), \theta' : (z_1, z_2, z_3) \to (z_1, -z_2, -z_3)$ and $\mathcal{R}$ of the orientifold action is given by $(z_1, z_2, z_3) \to (-z_1, -z_2, -z_3)$. Note that, the orientifold action also makes the flux quanta be multiples of 8. In this orientifold, there survive three complex structures τ_i (i.e. the dimension of $H_-^{2,1}$ is three) under the orbifold and orientifold action where i denotes each torus. This section is organized as follows. In Section 2.2.1, we briefly review geometry of the $T^6/(\mathbb{Z}_2 \times \mathbb{Z}_2')$ toroidal orientifold. There is a simple modular group that relates different points in modular space. In Section 2.2.2, we solve SUSY conditions on the orientifold and obtain analytic solutions that are already known. Then, in Section 2.2.3, we utilize the modular transformation of moduli to count independent vacua and show the exact distributions of the number of vacua in the moduli space. Finally, we compare our results with the modular flavor symmetric models for illustrative purposes in Section 2.2.4.

2.2.1 Modular symmetry in the $T^6/(\mathbb{Z}_2 \times \mathbb{Z}_2')$ orientifold

The toroidal orientifolds have a similar structure of moduli space as the CY. The modular group $SL(2,\mathbb{Z})_i$ $(i = 1, 2, 3)$ appears for each i-th torus T^2 corresponding to the symplectic group of CY complex structure moduli space. Thus, it is easier to describe the geometry of tori than to describe it of CY. Let us summarize the modular group on tori since it plays an important role in the following sections.

Suppose there is a torus T^2. Since we consider the factorizable torus, it is sufficient to consider only one T^2. The torus is defined as a quotient space $\mathbb{C}/\Lambda$ with Λ being a lattice. The lattice is defined to cover the whole $\mathbb{C}$ by its repeating, hence they are defined up to the modular transformation. Thus, the effective action must be invariant under the modular group $SL(2,\mathbb{Z})$[7]. We define the basis of the lattice as (e_x, e_y) where the complex coordinates of torus z is expressed as

$$z = e_x + U e_y. \tag{2.67}$$

Here, U denotes the complex structure modulus of T^2 and it is defined as the ratio of e_y to e_x:

$$U = \frac{e_y}{e_x}. \tag{2.68}$$

[7]Since general background fluxes can break the invariance, we must impose it by hand.

If we denote z by u which is a complex coordinates of the original $\mathbb{C}$ at the same point, z is also expressed as

$$z = \frac{u}{e_x}. \tag{2.69}$$

We always define U and z by the same expressions. On the other hand, the modular transformation of lattice is represented as

$$\begin{pmatrix} e'_y \\ e'_x \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e_y \\ e_x \end{pmatrix}, \tag{2.70}$$

where $ad - bc = 1$ i.e.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{Z}). \tag{2.71}$$

Hence the modular transformations of U and z are given by

$$U = \frac{e_y}{e_x} \to U' = \frac{e'_y}{e'_x} = \frac{aU + b}{cU + d}, \tag{2.72}$$

$$z = \frac{u}{e_x} \to z' = \frac{u'}{e'_x} = \frac{z}{cU + d}, \tag{2.73}$$

where we use the convention of Ref. [119–121]. Thus, the complex structure modulus U transforms under the modular group $SL(2, \mathbb{Z})$ as $U \to R(U)$, with the action $R(U)$ being

$$R(U) = \frac{aU + b}{cU + d}. \tag{2.74}$$

Precisely speaking, the exact group that we consider is $PSL(2, \mathbb{Z}) = SL(2, \mathbb{Z}) \backslash \{\mathbf{1}, -\mathbf{1}\}$ since the change $\{a, b, c, d\} \to \{-a, -b, -c, -d\}$ has no effect on the transformation (2.74). Note that under the modular transformations, a holomorphic function $f_k(w)$ which transforms as

$$f_k\left(\frac{aw + b}{cw + d}\right) = (cw + d)^k f(w) \tag{2.75}$$

is called a modular form of modular weight k, and the complex coordinates z itself is a modular form of modular weight k.

The modular group is generated by $S, T \in SL(2, \mathbb{Z})$ which satisfy the algebra

$$S^2 = \mathbf{1}, \quad (ST)^3 = \mathbf{1}. \tag{2.76}$$

We can also represent S, T as[8]

$$S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \tag{2.77}$$

[8] Although this matrix representation of S leads to $S^2 = -1$, this makes no problem since we identify -1 with 1 i.e., actually we do not consider $SL(2, \mathbb{Z})$ but $PSL(2, \mathbb{Z}) = SL(2, \mathbb{Z})/\{\mathbf{1}, -\mathbf{1}\}$.

and these generators correspond to

$$S : U \to -\frac{1}{U}, \quad T : U \to U + 1, \tag{2.78}$$

respectively.

In group theory, a set of group elements $\{g \cdot x | g \in G\}$ is called the orbit of $x \in G$, where G denotes a group. Then, the transformation of U (2.72) through the change of lattice basis relates different points in the moduli space and defines the orbit of U. Since a change of basis does not affect the effective action, we focus on the quotient space $\mathcal{M}\backslash R = \mathcal{M}^R$ ($\mathcal{R}$ denotes the moduli space). In other words, it is sufficient to consider only a set of representatives. A domain where each orbit intersects only once is called the fundamental domain and it is what we are interested in here. The fundamental domain F_D of the complex structure modulus space is given by

$$F_D = \left\{ U \middle| -\frac{1}{2} \leq \mathrm{Re}U \leq 0, |U|^2 \geq 1 \right\} \bigcup \left\{ U \middle| 0 < \mathrm{Re}U < \frac{1}{2}, |U|^2 > 1 \right\}. \tag{2.79}$$

In general, it is possible to define the fundamental domain in this way for any field that transforms under the $SL(2,\mathbb{Z})$. Indeed, we consider distributions of the axio-dilaton as well as complex structure moduli in their fundamental domains in the later sections.

Next, we describe the invariance of the effective action on $T^6 = T_1^2 \times T_2^2 \times T_3^3$. The Kähler potential K and the superpotential W which we reviewed in Section 2.1.1 are determined by the symplectic cohomology basis. In the tori T^6, the basis is given by

$$\begin{array}{ll} \alpha_0 = dx^1 \wedge dx^2 \wedge dx^3, & \beta^0 = dy^1 \wedge dy^2 \wedge dy^3, \\ \alpha_1 = dy^1 \wedge dx^2 \wedge dx^3, & \beta^1 = -dx^1 \wedge dy^2 \wedge dy^3, \\ \alpha_2 = dy^2 \wedge dx^3 \wedge dx^1, & \beta^2 = -dx^2 \wedge dy^3 \wedge dy^1, \\ \alpha_3 = dy^3 \wedge dx^1 \wedge dx^2, & \beta^3 = -dx^3 \wedge dy^1 \wedge dy^2, \end{array} \tag{2.80}$$

where we normalized the volume of T^6 to one. Moreover, the holomorphic three-form Ω takes the simple form

$$\Omega = dz^1 \wedge dz^2 \wedge dz^3. \tag{2.81}$$

Then the Kähler potential K becomes

$$K_{\mathrm{cs}} = -\log\big(i(U_1 - \bar{U}_1)(U_2 - \bar{U}_2)(U_3 - \bar{U}_3)\big), \tag{2.82}$$

and the superpotential W is also determined by the background fluxes F_3, H_3 and Ω in (2.81).

To assure the invariance of the scalar potential, it is enough to impose the invariance on $e^K|W|^2$. Note that, this implies the change of overall sign $W \to -W$ has no effect on the effective action. First, the Kähler potential transforms as

$$K \to K + \sum_i \log|cU_i + d|^2. \tag{2.83}$$

Thus the superpotential must transform as

$$W \to \prod_i (cU_i + d)^{-1} W, \tag{2.84}$$

i.e. W is a modular function of weight -1 for each $SL(2,\mathbb{Z})$.

In the following sections, we use the notation in Section 2.1.1. The superpotential is given by

$$\begin{aligned} W = {} & (f^0 - Sh^0)U_1U_2U_3 - ((f^1 - Sh^1)U_2U_3 + (f^2 - Sh^2)U_3U_1 + (f^3 - Sh^3)U_1U_2) \\ & + (f_0 - Sh_0) + (f_1 - Sh_1)U_1 + (f_2 - Sh_2)U_2 + (f_3 - Sh_3)U_3, \end{aligned} \tag{2.85}$$

and the scalar potential takes the no-scale structure

$$V_{\text{no-scale}} = e^K (K^{I\bar{J}} D_I W D_{\bar{J}} \overline{W}), \tag{2.22}$$

where I, J denotes the axio-dilaton S and $U_{1,2,3}$.

Furthermore, we assume the isotropic torus where the complex structure moduli and fluxes are identified as

$$U \equiv U_1 = U_2 = U_3 \tag{2.86}$$

and

$$f^1 \equiv f^1 = f^2 = f^3, \quad f_1 \equiv f_1 = f_2 = f_3, \quad h^1 \equiv h^1 = h^2 = h^3, \quad h_1 \equiv h_1 = h_2 = h_3 \tag{2.87}$$

to simplify our discussion. Thus, the superpotential reduces to

$$W = (f^0 - Sh^0)U^3 - 3(f^1 - Sh^1)U^2 + 3(f_1 - Sh_1)U + (f_0 - Sh_0). \tag{2.88}$$

In the isotropic torus, the modular invariance of the effective action requires that the background fluxes transform as

- $S : S \to -\frac{1}{S}$

$$\begin{aligned} & f^0 \to -h^0, \quad f^1 \to -h^1, \quad f_1 \to -h^1, \quad f_0 \to -h_0 \\ & h^0 \to f^0, \quad h^1 \to f^1, \quad h_1 \to f_1, \quad h_0 \to f_0. \end{aligned} \tag{2.89}$$

- $T : S \to S + 1$

$$\begin{aligned} & f^0 \to f^0 + h^0, \quad f^1 \to f^1 + h^1, \quad f_1 \to f_1 + h_1, \quad f_0 \to f_0 + h_0 \\ & h^0 \to h^0, \quad h^1 \to h^1, \quad h_1 \to h_1, \quad h_0 \to h_0. \end{aligned} \tag{2.90}$$

for the $SL(2,\mathbb{Z})_S$ transformation of the axio-dilaton S, which we mentioned in (2.4). Next, for the $SL(2,\mathbb{Z})_U$ transformation of the complex structure modulus U the flux quanta must transform as

- $S : U \to -\frac{1}{U}$

$$\begin{aligned} &f^0 \to -f^0, \quad f^1 \to -f_1, \quad f_1 \to f^1, \quad f_0 \to f^0, \\ &h^0 \to -h_0, \quad h^1 \to -h_1, \quad h_1 \to h^1, \quad h_0 \to h^0. \end{aligned} \tag{2.91}$$

- $T : U \to U \to U + 1$

$$\begin{aligned} &f^0 \to f^0, \quad f^1 \to f^1 + a^0, \quad f_0 \to f_0 + 2f^1 + f^0, \quad f_0 \to f_0 - 3f_1 - 3f^1 - f^0, \\ &h^0 \to h^0, \quad h^1 \to h^1 + h^0, \quad h_1 \to h_1 - 2f_1 + h^0, \quad h_0 \to h_0 - 3h_1 - 3h^1 - h^0. \end{aligned} \tag{2.92}$$

These transformations can be found in Ref. [91, 111]. Note that, in the previous research [122], the distributions of vacua is already determined. The authors showed the distributions for various maximum values of the D3-flux charge $N_{\rm flux}$, which we denote $N_{\rm flux}^{\rm max}$. However, in this thesis, we focus on distributions of the degrees of degeneracy whereas the authors of Ref. [122] studied the distributions of moduli VEVs. The reason to study it is that we would like to study points in the moduli space which are chosen naturally as we mentioned in Chapter 1. Henceforth, we call distributions of the degrees of degeneracy just distributions of vacua.

2.2.2 Analytic solutions of SUSY conditions

Next, we describe analytic solutions of the SUSY conditions. More concretely, we consider the following equations:

$$\partial_S W = 0, \quad \partial_U W = 0, \quad W = 0. \tag{2.93}$$

The last equation is a consequence of $D_T W = 0$, where T denotes an unstabilized Kähler modulus[9]. The known method to solve $W = 0$ is as follows. First of all, the conditions require that the following cubic polynomials vanish at minima:

$$W_{\rm RR} = \int F_3 \wedge \Omega = f^0 U^3 - 3f^1 U^2 + 3f^1 U + f_0 \equiv 0, \tag{2.94}$$

$$W_{\rm NS} = \int H_3 \wedge \Omega = h^0 U^3 - 3h^1 U^2 + 3h^1 U + h_0 \equiv 0. \tag{2.95}$$

Since the positive VEV of ImU is needed, each equation must have a pair of roots which are complex conjugate with each other. The remaining root must be real, thus we can factorize these polynomials

$$W_{\rm RR} = (rU + s)P(U), \tag{2.96}$$

$$W_{\rm NS} = (uU + v)P(U), \tag{2.97}$$

with $P(U)$ is a quadratic polynomial and a common factor of $W_{\rm RR,NS}$. Hence the relevant VEV of U is a root of $P(U)$ with a positive imaginary part. Here, we followed the lines of [83, 111] where $\{m, l, n, u, v, s, r\}$ are real integers which are determined by the flux quanta.

[9] Although we solved the SUSY condition for T, T is still unstabilized and we have to consider a mechanism which stabilizes the T modulus.

Then, the SUSY condition for the complex structure modulus can be solved immediately. Indeed, $\partial_U W = 0$ reduces to

$$P(U)\partial_U\{(rU+s) - S(uU+v)\} + \{(rU+s) - S(uU+v)\}\partial_U P(U) = 0 \tag{2.98}$$

and $P(U)$ vanishes for the VEV of U[10]. It leads to the VEV of axio-dilaton

$$S = \frac{rU+s}{uU+v}. \tag{2.99}$$

The VEV of complex structure modulus U is determined as

$$U = \frac{-m+\sqrt{m^2-4ln}}{2l} \quad (l,n>0), \tag{2.100}$$

$$U = \frac{-m-\sqrt{m^2-4ln}}{2l} \quad (l,n<0), \tag{2.101}$$

for a general quadratic polynomial $P(U) = lU^2 + mU + n$. Thus, VEVs of moduli are determined analytically. Furthermore, these vacua are automatically stable except for the unstabilized Kähler modulus since SUSY remains at vacua. However, there remain equivalence relations in the moduli space, namely $SL(2,\mathbb{Z})_{S,U}$. We have to get rid of the degeneracy in counting the vacua. Note that, since the flux quanta are multiple of 8, the D3 flux charge N_{flux} becomes a multiple of 64. Moreover, as discussed in Ref. [83], there is an additional factor 3 so that N_{flux} is a multiple of 192.

2.2.3 Distributions of flux vacua

In the previous section, we obtained the analytic expressions of moduli VEV. In this section, we show the distributions of the degrees of degeneracy in the moduli space. To compare with known results, we follow the lines of Ref. [122] and change values of upper bound for N_{flux}. Furthermore, we set the lower bound for ImS, which gives the inverse of the string coupling g_s.

We set the maximum value $N_{\text{flux}}^{\text{max}}$[11] as

$$N_{\text{flux}}^{\text{max}} = 192 \times \{10, 100, 250, 500, 1000\} \tag{2.102}$$

and take into account all the possible configurations which satisfy each upper bound. The vacua are all physically-distinct from each other i.e., no possible double count remains. The number of such vacua is summarized in Table 2.1 and the number behaves against $N_{\text{flux}}^{\text{max}}$ as

$$\#\text{ of stable vacua} \simeq 2.89527 \left(\frac{N_{\text{flux}}^{\text{max}}}{192}\right)^2, \tag{2.103}$$

which is different from the result in Ref. [122]. This is because our procedures for counting the vacua are different from the previous research.

[10] We omit the symbol of VEV for simplicity here and what in follows.
[11] We omit the factor 192 in the following discussion.

	$N_{\text{flux}}^{\text{max}} = 10$	$N_{\text{flux}}^{\text{max}} = 100$	$N_{\text{flux}}^{\text{max}} = 250$	$N_{\text{flux}}^{\text{max}} = 500$	$N_{\text{flux}}^{\text{max}} = 1000$
# of stable vacua	312	29218	178191	720710	2896221

Table 2.1: Number of stable vacua for each upper bound $N_{\text{flux}}^{\text{max}}$ determining the range of flux quanta [68].

2.2.3.1 Distribution on the complex structure modulus plane

First, we show the distributions of the complex structure modulus without fixing the VEV of axio-dilaton i.e. we analyze the vacua in the complex structure plane. The distributions are drawn in Figure 2.1 for the cases $N_{\text{flux}}^{\text{max}} = 10$ (left panel) and $N_{\text{flux}}^{\text{max}} = 10$ (right panel). The dark color point indicates that there exist many degenerated vacua at that point. We find that most of the vacua are clustered in the region with small values of ImU. Since we can find this tendency more clearly in the left panel, the cluster is a result of the tadpole cancellation condition. If we restrict our discussion to the region with small N_{flux}, we cannot obtain U that resides at $U \sim i\infty$.

There are specific points in the moduli space that are favored in the Landscape. It is known that the $SL(2,\mathbb{Z})$ has its stabilizer, namely the fixed points under its action. There are three stabilizers in the fundamental domain F_D, as

$$U = i, \omega, i\infty \tag{2.104}$$

with $\omega = e^{2\pi i/3}$. We call the i and $i\infty$ $\mathbb{Z}_2$ fixed points and ω $\mathbb{Z}_3$ fixed points, respectively.

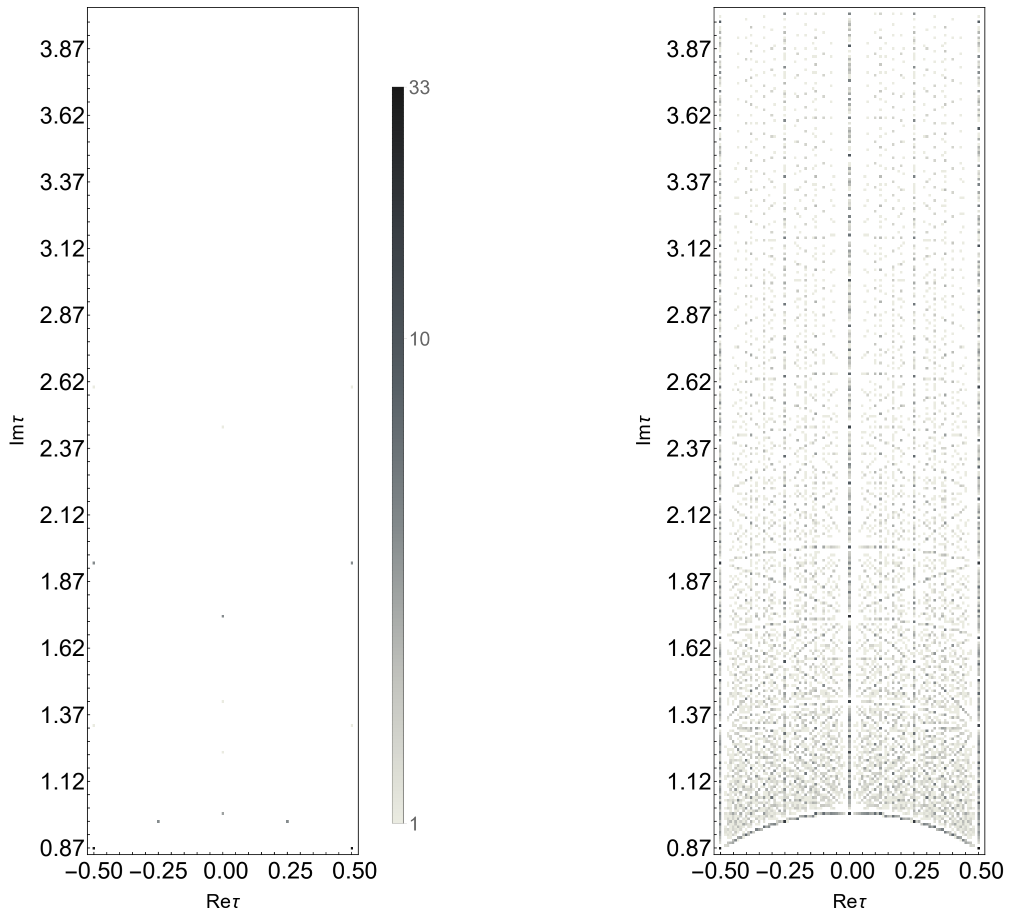

Figure 2.1: The numbers of SUSY vacua on $(\mathrm{Re}\, U, \mathrm{Im}\, U)$ plane for $N_{\rm flux}^{\rm max} = 10$(left) and $N_{\rm flux}^{\rm max} = 1000$(right), respectively [68].

$(\mathrm{Re}\, U, \mathrm{Im}\, U)$	$(-\frac{1}{2}, \frac{\sqrt{3}}{2})$	$(0, \sqrt{3})$	$(-\frac{1}{2}, \frac{\sqrt{15}}{2})$	$(-\frac{1}{4}, \frac{\sqrt{15}}{4})$	$(0, 1)$	$(-\frac{1}{2}, \frac{3\sqrt{3}}{2})$	$(0, \sqrt{\frac{3}{2}})$	$(0, \sqrt{6})$	$(0, \sqrt{2})$	$(-\frac{1}{2}, \frac{\sqrt{7}}{2})$
Probability (%)	62.3	7.55	7.55	7.55	5.66	1.89	1.89	1.89	1.89	1.89

Table 2.2: Probabilities of SUSY vacua as functions of $(\mathrm{Re}\, U, \mathrm{Im}\, U)$ in the descending order, the $N_{\rm flux}^{\rm max} = 10$ case [68].

$(\mathrm{Re}\, U, \mathrm{Im}\, U)$	$(-\frac{1}{2}, \frac{\sqrt{3}}{2})$	$(0, \sqrt{3})$	$(-\frac{1}{2}, \frac{\sqrt{15}}{2})$	$(-\frac{1}{4}, \frac{\sqrt{15}}{4})$	$(0, 1)$	$(-\frac{1}{2}, \frac{\sqrt{7}}{2})$	$(0, \sqrt{6})$	$(0, \sqrt{\frac{3}{2}})$	$(0, \sqrt{2})$	$(-\frac{1}{2}, \frac{3\sqrt{3}}{2})$
Probability (%)	40.3	7.55	4.85	4.85	3.79	2.43	1.88	1.88	1.88	1.49

Table 2.3: Probabilities of SUSY vacua as functions of $(\mathrm{Re}\, U, \mathrm{Im}\, U)$ in the descending order, for the $N_{\rm flux}^{\rm max} = 1000$ case [68].

The results of the probabilities of SUSY vacua are shown in Table 2.2 and 2.3 [12]. From

[12] Although the lines $ReX = \frac{1}{2}(X = S, U)$ are identified with the lines $ReX = -\frac{1}{2}(X = S, U)$, we add these lines to our figures. In the calculation of probabilities, these lines are excluded. In the histograms, the size of bin is set to (0.01×0.01)

these results, we can understand the cluster at $\mathbb{Z}_3$ fixed point quantitatively. The tendency becomes clear for small $N_{\rm flux}$, which is more favorable since the large $N_{\rm flux}$ leads to considerable backreactions through gravitated fluxes. Let us comment the other clustered point in the Landscape. First, $(\mathrm{Re}U, \mathrm{Im}U) = (0, \sqrt{k})$ $(k = 1, 2, 3, 6)$ points are statistically favored. Note that, since the fundamental region of $GL(2, \mathbb{Z})$ is half of F_D, these points are on edge of the fundamental region of $GL(2, \mathbb{Z})$. Next, we focus on the absolute values of U. These points satisfy

$$|U|^2 = \frac{l}{2}, \tag{2.105}$$

with l being a positive integer. There are other points

$$(\mathrm{Re}\,U, \mathrm{Im}\,U) = \left(\pm\frac{1}{4}, \frac{\sqrt{8k-1}}{4}\right) = \left(\pm\frac{1}{4}, \frac{\sqrt{23}}{4}\right), \left(\pm\frac{1}{4}, \frac{\sqrt{31}}{4}\right), \left(\pm\frac{1}{4}, \frac{\sqrt{39}}{4}\right), \cdots, \tag{2.106}$$

which are not listed on Table 2.2 and 2.3 with $\mathcal{O}(0.1 - 1)\%$ are found. The absolute values of these points are again expressed as (2.105).

Why the VEVs with their absolute values satisfying (2.105) are favored? It is also unclear why these discrete points are selected from those semicircles. Through the analytic expressions, we find that U satisfies

$$|U|^2 = \frac{n}{l}, \tag{2.107}$$

where n, l are integers in (2.100) and (2.101). Thus, all the VEVs of U are on the (semi)circles whose radii are given in (2.107). Although this does not imply directly that each semicircle itself becomes a set of vacua since the vacua are discrete at the first point, we can discuss some tendency through the fact. Suppose the vacua distributions are given by (2.107). Then, the parts of those semicircles that extend beyond the boundary of F_D are folded back through T-transformations. The patterns that are generated by this phenomenon can be seen in Figure 2.1. Since such folded back semicircles intersect other semicircles, the intersection points are statistically favored.

We also compared our exact results with the results which are obtained in the statistical methods in Ref. [40–42] and the methods predict [111]

$$N_{\rm vacua} \sim \frac{\pi^2}{108}\frac{(N_{\rm flux}^{\rm max})^2}{(t(m^2 - 4ln))^2}, \tag{2.108}$$

for the SUSY Minkowski vacua that we consider in the thesis. Here, l, m, n are integers that parameterize the VEV of U. $t(x)$ is defined as

$$t(x) = \begin{cases} x & (x \equiv 0), \qquad (2.109) \\ 3x & (\text{otherwise}) \qquad (2.110) \end{cases}$$

modulo 3. The predicted values through this equation (2.108) is listed on Table 2.4. One can see the $\mathbb{Z}_3$ point is also selected out in this method. Note that, the point $(0, \sqrt{3})$ is again predicted with $O(1)\%$ probability. On the other hand, the statistical method predicts a too large probability for the $\mathbb{Z}_2$ point. In our exact results, we can see that the probability of $\mathbb{Z}_2$ is lower than $(\frac{-1}{4}, \frac{\sqrt{15}}{4})$ which does not possess any additional symmetry. Furthermore, Eq. (2.108)implies the smaller number of vacua in the LCS regime. Thus, the statistical method roughly matches our results, but our results describe the structure of the landscape in more detail.

$(\mathrm{Re}\,U, \mathrm{Im}\,U)$	$(-\frac{1}{2}, \frac{\sqrt{3}}{2})$	$(0, \sqrt{3})$	$(0, 1)$	$(-\frac{1}{4}, \frac{\sqrt{15}}{4})$	$(-\frac{1}{2}, \frac{\sqrt{15}}{2})$	$(-\frac{1}{2}, \frac{\sqrt{7}}{2})$	$(0, \sqrt{6})$	$(0, \sqrt{\frac{3}{2}})$	$(0, \sqrt{2})$	$(-\frac{1}{2}, \frac{3\sqrt{3}}{2})$
Probability (%)	63.9	3.99	3.99	2.56	2.56	1.30	1.00	1.00	1.00	0.79

Table 2.4: Probabilities of SUSY vacua as functions of $(\mathrm{Re}\,U, \mathrm{Im}\,U)$ shown in Table 2.3, predicted by the statistical approach [123] [68].

The resulting vacua included all of the VEVs for the axio-dilaton S. Since our discussion is restricted to the perturbative framework, it is important to fix S in the weak coupling regime $\mathrm{Im}S \gg 1$ for consistency. Moreover, the known non-perturbative corrections often take the form which is suppressed in the weak coupling regime. Then, we impose the lower bound for $\mathrm{Im}S$ and summarize the distributions in Table 2.5 and 2.6. We find that in the $N_{\text{flux}}^{\text{max}} = 10$ case which gives a strict upper bound for N_{flux}, the Z_3 point is selected exclusively as $\mathrm{Im}S$ becomes larger. It implies that the $\mathbb{Z}_3$ point is only an allowed point in the well-controlled landscape. On the other hand, the probabilities are distributed uniformly for the $N_{\text{flux}}^{\text{max}} = 1000$ case. Note that, even with $N_{\text{flux}}^{\text{max}} = 1000$ the probability of $\mathbb{Z}_2$ point is less than one-tenth of the probability of $\mathbb{Z}_3$ point. The result suggests that even if a large number of vacua exist (see Table 2.1), only those with a specific VEV exist in the well-controlled landscape. Then the string landscape would have predictive power.

From a phenomenological point of view, the result that we find would affect bottom-up research. Indeed, there are many modular flavor symmetric models around $\mathbb{Z}_3, \mathbb{Z}_2$ fixed points in F_D (see for, e.g. [46,51,55,124–126]). These fixed points are preferred in such models because there remain the residual symmetries on those fixed points and one can utilize the residual symmetries to obtain Yukawa couplings whose forms are phenomenologically preferred. Then, our results which indicate such fixed points are statistically favored support those models. Furthermore, the statistical advantage of $\mathbb{Z}_3$ fixed point over $\mathbb{Z}_2$ fixed point may constrain the construction of such models.

$(\mathrm{Re}\,U,\,\mathrm{Im}\,U)$ \ $\mathrm{Im}\,S$	All	$\frac{\sqrt{3}}{2}$	1	2	3	4	5
$(-\frac{1}{2}, \frac{\sqrt{3}}{2})$	62.3	62.3	63.4	77.8	77.8	100	100
$(0, \sqrt{3})$	7.55	7.55	7.32	5.56	11.1	0	0
$(-\frac{1}{2}, \frac{\sqrt{15}}{2})$	7.55	7.55	7.32	0	0	0	0
$(-\frac{1}{4}, \frac{\sqrt{15}}{4})$	7.55	7.55	4.88	5.56	11.1	0	0
$(0, 1)$	5.66	5.66	7.32	5.56	0	0	0
$(-\frac{1}{2}, \frac{\sqrt{7}}{2})$	1.89	1.89	2.44	0	0	0	0
$(0, \sqrt{6})$	1.89	1.89	2.44	0	0	0	0
$(0, \sqrt{\frac{3}{2}})$	1.89	1.89	2.44	5.56	0	0	0
$(0, \sqrt{2})$	1.89	1.89	2.44	0	0	0	0
$(-\frac{1}{2}, \frac{3\sqrt{3}}{2})$	1.89	1.89	0	0	0	0	0

Table 2.5: Probabilities (%) of the SUSY vacua as functions of $(\mathrm{Re}\,U, \mathrm{Im}\,U)$ and $\mathrm{Im}\,S$ in the descending order with respect to $(\mathrm{Re}\,U, \mathrm{Im}\,U)$ in the $N_{\rm flux}^{\rm max} = 10$ case. Note that the column in "All" corresponds to the probability of Table 2.2 (i.e. no bound) [68].

$(\mathrm{Re}\,U,\,\mathrm{Im}\,U)$ \ $\mathrm{Im}\,S$	All	$\frac{\sqrt{3}}{2}$	1	2	3	4	5
$(-\frac{1}{2}, \frac{\sqrt{3}}{2})$	40.3	40.3	40.3	40.7	41.0	41.4	41.7
$(0, \sqrt{3})$	7.56	7.56	7.57	7.62	7.67	7.71	7.77
$(-\frac{1}{2}, \frac{\sqrt{15}}{2})$	4.85	4.85	4.85	4.88	4.91	4.93	4.99
$(-\frac{1}{4}, \frac{\sqrt{15}}{4})$	4.85	4.85	4.85	4.88	4.92	4.92	4.98
$(0, 1)$	3.79	3.79	3.82	3.85	3.86	3.89	3.90
$(-\frac{1}{2}, \frac{\sqrt{7}}{2})$	2.44	2.44	2.44	2.45	2.47	2.46	2.50
$(0, \sqrt{6})$	1.88	1.88	1.89	1.90	1.90	1.90	1.92
$(0, \sqrt{\frac{3}{2}})$	1.88	1.88	1.89	1.90	1.89	1.92	1.90
$(0, \sqrt{2})$	1.88	1.88	1.89	1.90	1.91	1.92	1.91
$(-\frac{1}{2}, \frac{3\sqrt{3}}{2})$	1.49	1.49	1.49	1.50	1.51	1.52	1.49

Table 2.6: Probabilities (%) of the SUSY vacua as functions of $(\mathrm{Re}\,U, \mathrm{Im}\,U)$ and $\mathrm{Im}\,S$ in the descending order with respect to $(\mathrm{Re}\,U, \mathrm{Im}\,U)$ in the $N_{\rm flux}^{\rm max} = 1000$ case. Note that the column in "All" corresponds to the probability of Table 2.3 (i.e. no bound) [68] .

2.2.3.2 Distribution of the complex structure modulus with the fixed axio-dilaton

Next, we study whether there exist correlations between the VEVs of two moduli or not in the Landscape. Namely, we take into account the VEV of axio-dilaton in this section. As a result, there exist strong correlations between them.

First, we showed the distribution of the axio-dilaton VEV on the axio-dilaton plane for each $N_{\rm flux}^{\rm max}$ in Figure 2.2. The black dots indicate that a large number of vacua degenerate at that point. As in Figure 2.1, we find that black dots cluster in the region with small imaginary part.

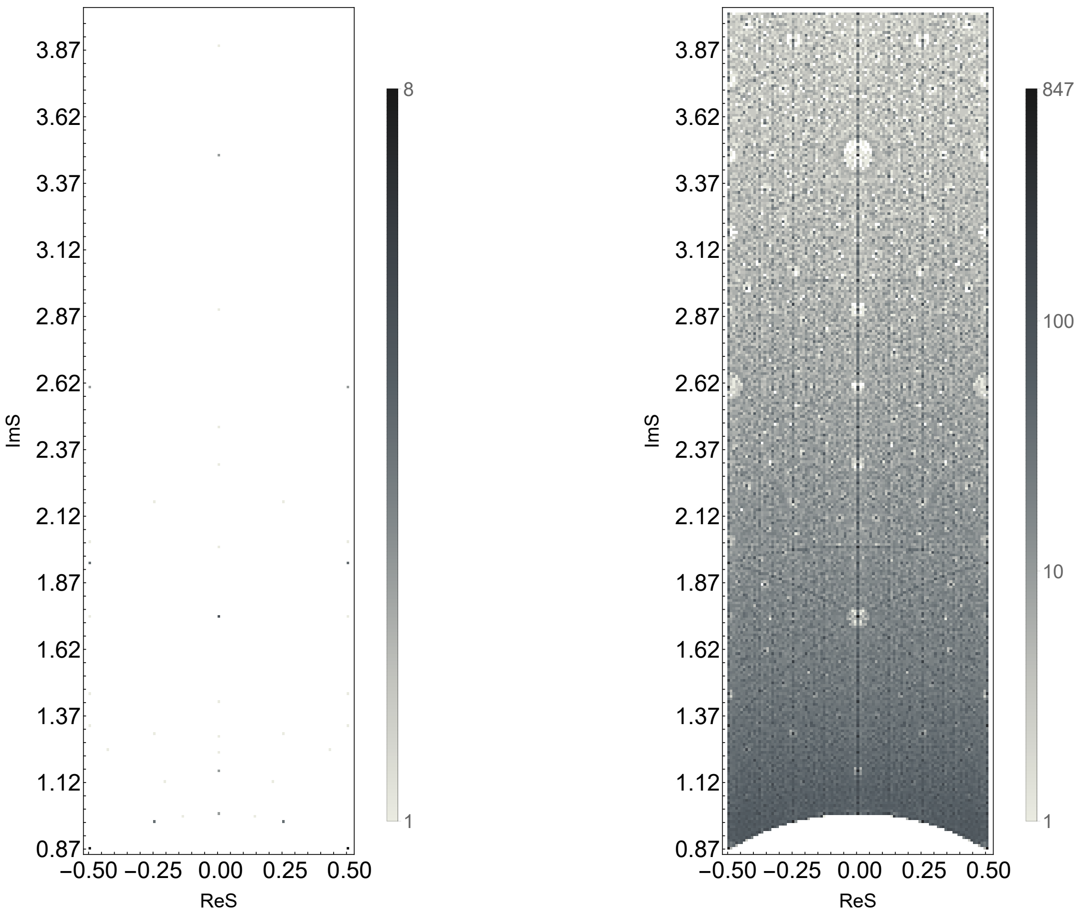

Figure 2.2: The numbers of the SUSY vacua on $(\mathrm{Re}\,S, \mathrm{Im}\,S)$ plane for $N_{\rm flux}^{\rm max} = 10$ (left) and $N_{\rm flux}^{\rm max} = 1000$ (right), respectively [68].

However, compared to Figure 2.1, it can be seen that the distributions are not limited to the fixed points but rather, it is uniformly dispersed. In fact, we can see that the probabilities of $\mathbb{Z}_3$ point shown in Table 2.7 and 2.8 are only about $\frac{1}{4}$ for the $N_{\rm flux}^{\rm max} = 10$ case and $\frac{1}{400}$ for the $N_{\rm flux}^{\rm max} = 1000$ case compared to the corresponding probabilities in the complex structure modulus space.

To find out whether there are correlations between the VEVs of two moduli or not, we selected 10 points in the descending order of the probability and extracted the distribution of the complex structure modulus by selecting only those that have each of the selected S as a pair. The results are listed on Table 2.9 and 2.10. We can find that unique distributions can appear when the distributions are cut by a specific S-plane. For example, let us consider the $\mathbb{Z}_2$ fixed point in the $N_{\rm flux}^{\rm max} = 10$ case. Although we concluded the previous section with a statement that claims $\mathbb{Z}_2$ is not a natural point in the landscape, the result shows that if we fix the axio-dilaton to specific values the $\mathbb{Z}_2$ fixed point is selected out exclusively. It is still nontrivial from a viewpoint of the string landscape since we need to fix the axio-dilaton to such specific values, but there remains a possibility to adopt phenomenological models around the $\mathbb{Z}_2$ fixed point. Thus, we have to consider a mechanism that stabilizes the axio-dilaton at such points.

$(\mathrm{Re}\,S, \mathrm{Im}\,S)$	$(-\frac{1}{2}, \frac{\sqrt{3}}{2})$	$(0, \sqrt{3})$	$(-\frac{1}{2}, \frac{\sqrt{15}}{2})$	$(-\frac{1}{4}, \frac{\sqrt{15}}{4})$	$(-\frac{1}{2}, \frac{3\sqrt{3}}{2})$	$(0,1)$	$(0, 2\sqrt{3})$	$(0, \frac{2}{\sqrt{3}})$	$(0, 5\sqrt{3})$	$(0, \frac{5}{\sqrt{3}})$
Probability (%)	15.1	7.55	5.66	5.66	3.77	3.77	3.77	3.77	1.89	1.89

Table 2.7: Probabilities of the SUSY vacua as functions of $(\mathrm{Re}\,S, \mathrm{Im}\,S)$ in the descending order for $N_{\rm flux}^{\rm max} = 10$ [68].

$(\mathrm{Re}\,S, \mathrm{Im}\,S)$	$(-\frac{1}{2}, \frac{\sqrt{3}}{2})$	$(0, \sqrt{3})$	$(-\frac{1}{2}, \frac{\sqrt{15}}{2})$	$(-\frac{1}{4}, \frac{\sqrt{15}}{4})$	$(-\frac{1}{2}, \frac{\sqrt{7}}{2})$	$(0, 2\sqrt{3})$	$(0, \frac{2}{\sqrt{3}})$	$(-\frac{1}{2}, \frac{3\sqrt{3}}{\sqrt{2}})$	$(0,1)$	$(0,2)$
Probability (%)	0.12	0.12	0.073	0.073	0.069	0.066	0.066	0.063	0.061	0.057

Table 2.8: Probabilities of the SUSY vacua as functions of $(\mathrm{Re}\,S, \mathrm{Im}\,S)$ in the descending order for $N_{\rm flux}^{\rm max} = 1000$ [68].

$(\mathrm{Re}\,U, \mathrm{Im}\,U)$ \ $(\mathrm{Re}\,S, \mathrm{Im}\,S)$	All	$(-\frac{1}{2}, \frac{\sqrt{3}}{2})$	$(0, \sqrt{3})$	$(-\frac{1}{2}, \frac{\sqrt{15}}{2})$	$(-\frac{1}{4}, \frac{\sqrt{15}}{4})$	$(-\frac{1}{2}, \frac{\sqrt{7}}{2})$	$(0, 2\sqrt{3})$	$(0, \frac{2}{\sqrt{3}})$	$(-\frac{1}{2}, \frac{3\sqrt{3}}{\sqrt{2}})$	$(0,1)$	$(0,2)$
$(-\frac{1}{2}, \frac{\sqrt{3}}{2})$	62.3	75	75	0	0	0	50	50	100	0	0
$(0, \sqrt{3})$	7.55	12.5	25	0	0	0	50	50	0	0	0
$(-\frac{1}{2}, \frac{\sqrt{15}}{2})$	7.55	0	0	66.7	33.3	0	0	0	0	0	0
$(-\frac{1}{4}, \frac{\sqrt{15}}{4})$	7.55	0	0	33.3	66.7	0	0	0	0	0	0
$(0,1)$	5.66	0	0	0	0	0	0	0	0	100	100
$(-\frac{1}{2}, \frac{\sqrt{7}}{2})$	1.89	0	0	0	0	100	0	0	0	0	0
$(0,6)$	1.89	0	0	0	0	0	0	0	0	0	0
$(0, \sqrt{\frac{3}{2}})$	1.89	0	0	0	0	0	0	0	0	0	0
$(0, \sqrt{2})$	1.89	0	0	0	0	0	0	0	0	0	0
$(-\frac{1}{2}, \frac{3\sqrt{3}}{2})$	1.89	12.5	0	0	0	0	0	0	0	0	0

Table 2.9: Probabilities of the SUSY vacua as functions of $(\mathrm{Re}\,U, \mathrm{Im}\,U)$ and $(\mathrm{Re}\,S, \mathrm{Im}\,S)$ in the descending order with respect to $(\mathrm{Re}\,U, \mathrm{Im}\,U)$ in the $N_{\rm flux}^{\rm max} = 10$ case. Note that the column in "All" corresponds to the probability of Table 2.2 (i.e. no cut) [68].

$(\mathrm{Re}\,U, \mathrm{Im}\,U)$ \ $(\mathrm{Re}\,S, \mathrm{Im}\,S)$	All	$(-\frac{1}{2}, \frac{\sqrt{3}}{2})$	$(0, \sqrt{3})$	$(-\frac{1}{2}, \frac{\sqrt{15}}{2})$	$(-\frac{1}{4}, \frac{\sqrt{15}}{4})$	$(-\frac{1}{2}, \frac{\sqrt{7}}{2})$	$(0, 2\sqrt{3})$	$(0, \frac{2}{\sqrt{3}})$	$(-\frac{1}{2}, \frac{3\sqrt{3}}{\sqrt{2}})$	$(0,1)$	$(0,2)$
$(-\frac{1}{2}, \frac{\sqrt{3}}{2})$	40.3	73.0	36.8	0	0	0	33.6	33.6	100	0	0
$(0, \sqrt{3})$	7.56	9.01	27.5	0	0	0	25.4	25.4	17.6	0	0
$(-\frac{1}{2}, \frac{\sqrt{15}}{2})$	4.85	0	0	33.3	32.7	0	0	0	0	0	0
$(-\frac{1}{4}, \frac{\sqrt{15}}{4})$	4.85	0	0	32.7	33.3	0	0	0	0	0	0
$(0,1)$	3.79	0	0	0	0	0	0	0	0	47.6	25.8
$(-\frac{1}{2}, \frac{\sqrt{7}}{2})$	2.44	0	0	0	0	35.7	0	0	0	0	0
$(0,6)$	1.88	0	0	0	0	0	0	0	0	0	0
$(0, \sqrt{\frac{3}{2}})$	1.88	0	0	0	0	0	0	0	0	0	0
$(0, \sqrt{2})$	1.88	0	0	0	0	0	0	0	0	0	0
$(-\frac{1}{2}, \frac{3\sqrt{3}}{2})$	1.88	8.17	12.1	0	0	0	10.6	10.6	16.2	0	0

Table 2.10: Probabilities of the SUSY vacua as functions of $(\mathrm{Re}\,U, \mathrm{Im}\,U)$ and $(\mathrm{Re}\,S, \mathrm{Im}\,S)$ in the descending order with respect to $(\mathrm{Re}\,U, \mathrm{Im}\,U)$ in the $N_{\rm flux}^{\rm max} = 1000$ case. Note that the column in "All" corresponds to the probability of Table 2.3 (i.e. no cut) [68].

2.2.3.3 Distributions of CP-breaking vacua in Landscape

As discussed in [58], the CP-breaking vacua are characterized by the non-vanishing VEVs of the axionic fields: ReS and ReU. In this case, there are additional VEVs which do not break CP. Indeed, the boundaries of fundamental domains are CP-invariant since there is a T-transformation:

$$\mathrm{Re}U = \pm\frac{1}{2} \xrightarrow{\mathcal{CP}} \mp\frac{1}{2} \xrightarrow{T} \pm\frac{1}{2}. \tag{2.111}$$

Hence, the Re$S = 0 =$ ReU and the boundaries of $SL(2,\mathbb{Z})_{S,U}$ are CP-conserving regions. Note that these regions are known as the fundamental domain of $GL(2,\mathbb{Z})$. We have the analytic expressions of VEVs and we can utilize them to search the distributions of CP-breaking vacua with the general scalar potential which is not invariant under $\mathcal{CP}$ i.e. we do not restrict our discussion to the spontaneous CP violation. We summarized our results on this search in Figure 2.3, Table 2.11 , 2.12 and 2.13. We find that the regions Re$U = 0, \pm\frac{1}{2}$ are statistically favored in F_D. For example, the sum of the probabilities of CP-conserving vacua excesses 92% of the landscape with $N_{\rm flux}^{\rm max} = 10$. Thus, CP-violation is not likely to happen in the complex structure modulus, which is consistent with [127] where it was pointed out that there is no spontaneous CP violation in the toroidal backgrounds.

We also find that Re$U = \pm\frac{1}{4}$ are statistically favored among the CP-breaking vacua. These correspond to the $\mathcal{O}(1)$% points in the distribution of the complex structure modulus (2.106). Moreover, the specific values of Re$S = \pm\frac{1}{4}$ leads to statistically favored CP-breaking vacua at Re$U = \pm\frac{1}{4}$. We can also see this effect in Table 2.9 and 2.10. Thus, a mechanism of stabilizing the axio-dilaton to specific values is needed to achieve a natural CP-breaking distribution in the landscape.

The above correlation is not a phenomenon that is realized only on the particular points in the moduli space. Indeed, we can find the lines

$$\mathrm{Re}U = \pm\mathrm{Re}S \tag{2.112}$$

are favored in the right panel of Figure 2.3. As we mentioned in Chapter 1, the VEVs of ReS, U contribute to different CP violations in SM. However, from a viewpoint of the string landscape, the VEVs are correlated; If ReU violates CP, then ReS tends to violate CP. Thus, it suggests that the different CP-violating phases in SM may have a common origin at a high energy scale. We can find a scenario that leads to this conclusion in Ref [128]. It is interesting to construct a concrete CP-violating scenario that is consistent with the correlations we observed here.

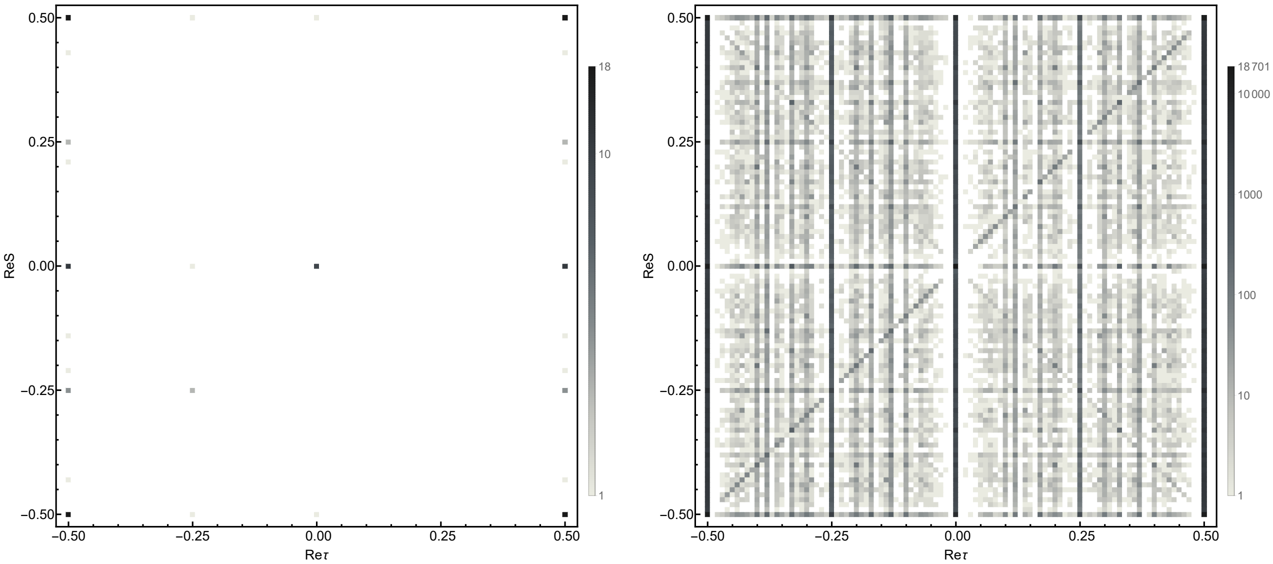

Figure 2.3: The numbers of SUSY vacua on the $(\mathrm{Re}\,U, \mathrm{Re}\,S)$ plane for $N_{\rm flux}^{\rm max} = 10$ (left) and $N_{\rm flux}^{\rm max} = 1000$ (right), respectively [68].

$(\mathrm{Re}\,U, \mathrm{Re}\,S)$	$(-\frac{1}{2}, -\frac{1}{2})$	$(-\frac{1}{2}, 0)$	$(0,0)$	$(-\frac{1}{2}, -\frac{1}{4})$	$(-\frac{1}{2}, \frac{1}{4})$	$(-\frac{1}{4}, -\frac{1}{4})$	$(0, -\frac{1}{2})$	$(-\frac{1}{2}, \frac{3}{7})$	$(-\frac{1}{2}, -\frac{3}{7})$	$(-\frac{1}{2}, -\frac{3}{14})$
Probability (%)	34.0	20.8	17.0	5.66	3.77	3.77	1.89	1.89	1.89	1.89

Table 2.11: Probabilities of the SUSY vacua as functions of $(\mathrm{Re}\,U, \mathrm{Re}\,S)$ in the descending order, for the $N_{\rm flux}^{\rm max} = 10$ case [68] .

$(\mathrm{Re}\,U, \mathrm{Re}\,S)$	$(0,0)$	$(-\frac{1}{2}, -\frac{1}{2})$	$(-\frac{1}{2}, 0)$	$(0, -\frac{1}{2})$	$(-\frac{1}{2}, -\frac{1}{4})$	$(-\frac{1}{2}, \frac{1}{4})$	$(0, -\frac{1}{3})$	$(0, \frac{1}{3})$	$(-\frac{1}{4}, -\frac{1}{2})$	$(0, -\frac{1}{4})$
Probability (%)	2.74	2.18	1.98	1.02	0.830	0.801	0.468	0.447	0.380	0.355

Table 2.12: Probabilities of the SUSY vacua as functions of $(\mathrm{Re}\,U, \mathrm{Re}\,S)$ in the descending order, for the $N_{\rm flux}^{\rm max} = 1000$ case [68] .

$N_{\rm flux}^{\rm max}$	10	100	250	500	1000
$\mathrm{Re}\,S$	22.6	66.3	78.7	85.3	90.0
$\mathrm{Re}\,U$	7.55	12.4	15.0	16.5	17.6

Table 2.13: Probabilities (%) of the CP-breaking VEV with respect to $\mathrm{Re}S, \mathrm{Re}U$ for various $N_{\rm flux}^{\rm max}$ [68].

2.2.4 Modular symmetric flavor models in Landscape

So far, we have explored the distributions of moduli VEVs with degeneracy explicitly shown in the Landscape on $T^6/\mathbb{Z}_2 \times \mathbb{Z}_2'$ orientifold. As already mentioned, such distributions of degeneracy are important in the sense of naturalness. However, as mentioned in Section 1.1, the other approach which is called modular flavor symmetry, has been developed recently in the context of the bottom-up approach. There, originally of the lepton sector, matter fields and Yukawa couplings are taken to be functions of a "modulus" field, which transforms under a $(P)SL(2,\mathbb{Z})$ group action. In particular, the functions are assumed to be modular forms. Since transformation properties of modular forms can easily be classified by their weights and there exists a finite basis for such functions, it provides ones a powerful tool to construct theory. Then, one can construct models with Lagrangian of the lepton sector being invariant under the modular group $SL(2,\mathbb{Z})$. Symmetries in those models are controlled by the modulus field.

A notable phenomenological benefit does emerge in this approach originates, from the fact that the modular group has certain non-Abelian discrete groups as its subgroups. Indeed, they have been considered as candidates for groups of flavor symmetry for a decade. Then, as shown later explicitly, such traditional flavor symmetries are unified into models with modular symmetry. In that sense, the modular symmetry in the models is called the modular flavor symmetry.

Of course, those flavor symmetries must be broken in the end. The existence of modular symmetry of the models also implies that if the modulus field is stabilized to acquire a suitable VEV, then the flavor symmetries are broken at least partially. The framework by which flavor symmetries become controllable and systematically constructive is realized by the modular flavor symmetry.

2.2.4.1 Brief Review of Modular Flavor Symmetry

2.2.4.2 Modular A_4 models

In this section, we introduce concrete modular A_4 models in Ref. [129] as examples.

One of the reasons why the modular group is preferred in phenomenological models is that the modular group has finite non-Abelian subgroups in it, as mentioned in Section 1.1. It is known that the following principal congruence subgroup $\Gamma(N)$ is a subgroup of $\Gamma \equiv SL(2,\mathbb{Z})$:

$$\begin{aligned} \Gamma(N) &= \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma \mid a = f \equiv 1, \quad b = c \equiv 0 (\mathrm{mod} N) \right\}, \\ \Gamma'_N &\equiv \Gamma/\Gamma(N) = \left\langle S, T \mid S^4 = (ST)^3 = T^N = \mathbb{I}, S^2 T = T S^2 \right\rangle , \end{aligned} \tag{2.113}$$

where Γ'_N is defined to be a quotient group of the principal congruence subgroup. As we commented in Section 2.2.1, the exact group that transforms the complex structure modulus is $\bar{\Gamma} = \Gamma \backslash \{\pm \mathbf{1}\}$. Hence we define the following quantities

$$\begin{aligned} \bar{\Gamma} &\equiv \Gamma \backslash \pm \mathbf{1}, \\ \bar{\Gamma}(N) &\equiv \Gamma(N) \backslash \pm \mathbf{1}, \\ \Gamma_N &\equiv \bar{\Gamma}/\bar{\Gamma}(N). \end{aligned} \tag{2.114}$$

Note that, Γ_N is generated as

$$\Gamma_N \equiv \left\langle S^2 = (ST)^3 = T^N = \mathbf{1} \right\rangle , \tag{2.115}$$

and $T^N = 1$ appears to close this subgroup. Moreover, it holds that [130]

$$\Gamma_2 \simeq S_3 \quad, \Gamma_3 \simeq A_4, \quad \Gamma_4 \simeq S_4, \quad \Gamma_5 \simeq A_5, \tag{2.116}$$

and these non-Abelian discrete groups S_3, A_4, S_4, and A_5 have geometrical interpretations. For example, A_4 is a group of tetrahedrons. These groups are associated with the SM particles (those particles transform as an irreducible representation of corresponding Γ_N).

As we mentioned, in the modular symmetric flavor models the Yukawa couplings are modular forms of the complex structure modulus U. In the following, we focus on the modular A_4 models which we compare with our results in Section 2.2.4.3. We follow the lines of Ref. [129].

First, we have to generalize the modular form f to have a level N as

$$f\left(\frac{aU+b}{cU+d}\right) = (cU+d)^k f(U) \tag{2.117}$$

with γ being an element of $\Gamma(N)$:

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{2.118}$$

Thus, what we have done is just replacing $SL(2,\mathbb{Z})$ by $\Gamma(N)$. One of the main features of the modular form is that it spans a linear space. Then, by choosing the appropriate basis, we can define the following multiplet of the modular forms which transforms as an irreducible representation of Γ_N:

$$\mathbf{F}(\gamma U) = (cU+d)^k \rho(\gamma)\mathbf{F}(U), \tag{2.119}$$

where $\gamma \in \bar{\Gamma}(3)$, $\mathbf{F} = (f_1, f_2, \ldots,)^T$ and $\rho(U)$ is a unitary representation matrix.

It is known that the $A_4 \simeq \Gamma_3$ group has three singlet representations $\mathbf{1}, \mathbf{1}', \mathbf{1}''$ as well as a triplet representation $\mathbf{3}$. Here, these three singlet representations correspond to S, T transformations as follows:

$$\mathbf{1} \leftrightarrow S=1,\ T=1, \quad \mathbf{1}' \leftrightarrow S=1,\ T=\omega^2, \quad \mathbf{1}'' \leftrightarrow S=1,\ T=\omega. \tag{2.120}$$

On the other hand, the triplet representation gives the explicit form of S, T as

$$S = \frac{1}{3}\begin{pmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega^2 & 0 \\ 0 & 0 & \omega \end{pmatrix}, \tag{2.121}$$

where the basis is chosen as [129].

Then, the authors of Ref. [129] introduced the (global) SUSY modular A_4 models, where chiral superfields Φ_I which constitute the action transform as

$$\Phi_I \to (cU+d)^{-k_I}\rho_I(\gamma)\Phi_I. \tag{2.122}$$

The chiral field transforms as in (2.122) is called a chiral superfield of weight k_I where the sign of k_I is opposite of our original definition. The superpotential of the modulus and the chiral fields is determined as

$$W = \sum_n Y_{I_1 \ldots I_n} \Phi_{I_i} \ldots \Phi_{I_n} \tag{2.123}$$

where n denotes the number of chiral superfields. Since the action must be invariant under the transformations, we can assign the allowed weights for chiral superfields. If the Yukawa couplings transform as

$$Y(\gamma U) = (cU + d)^{k_Y} \rho_Y(\gamma) Y(U), \tag{2.124}$$

$k_Y = \sum_i k_{I_i}$ is required. The product of unitary representation is also constrained. Then, in Ref. [129] the charge assignment which respects these requirements is performed as following Table 2.14.

	L	N^c	E_1^c, E_2^c, E_3^c	H_u	H_d
$SU(2)$	2	1	1	2	2
A_4	**3**	**3**	all combinations of **1**,**1**′, **1**″	**1**	**1**
$-k_I$	$-k_L$	$-k_{N^c}$	$-k_{E_1}$,$-k_{E_2}$,$-k_{E_3}$	0	0

Table 2.14: Charge assignment of the lepton sector and Higgs doublets which is performed in Ref. [129] [68].

The models which we utilize to compare with our results are all based on these configurations. The concrete allowed region in the moduli space is given in the next section.

2.2.4.3 Comparison between modular A_4 models and landscape

Finally, we compare our results and the models $\mathcal{B}_9, \mathcal{B}_{10}, \mathcal{D}_{5,6,7,8,9,10}$ in Ref. [129] which reproduces tiny neutrino masses through the type-I seesaw mechanism. At first, we summarize the allowed regions for the modulus U in reproducing favorable values of phenomenological quantities for each of those models in Table 2.15. Note that, the unit semicircle in the moduli space was outside the scope of the study in Ref. [129]. Thus, the most statistically favorable points in the landscape are excluded from the beginning. We summarized our results in Figure 2.4, Table 2.16 and 2.17. There, we do not consider the parts of the allowed regions that are out of the fundamental domain reflecting the method which is utilized in Ref. [129]. In Figure 2.4, we find the broad allowed region does not always mean that the corresponding model is statistically favored due to the clustered structure. It is phenomenologically interesting to focus on the small N_{flux} case i.e. $N_{\text{flux}}^{\text{max}} = 10$. As one can see in Table 2.16, models $\mathcal{D}_{5,8,9,10}$ are completely excluded in the landscape. Thus, we conclude that $\mathcal{B}_{9,10}$ and $\mathcal{D}_{6,7}$ models are natural models in the landscape. Note that, the other models become the candidates again with the large N_{flux} case i.e. $N_{\text{flux}}^{\text{max}} = 1000$. If, conversely, a model that is discarded with the small N_{flux} is confirmed experimentally, the result means the framework of our discussion is not sufficient for phenomenology since we neglected the flux backreactions entirely. Note that the models $\mathcal{D}_{5,6,8}$ and $\mathcal{B}_{10}$ contain the excluded semicircle in their allowed region. Therefore, it indicates that we may not be able to discard these models, but usually, the deviation from those fixed points is important in phenomenological models and we consider the above selection valid.

Models	Allowed regions	
	for ReU	for ImU
Model $\mathcal{B}_9$	$[0,\ 0.368]$	$[1.351,\ 1.856]$
Model $\mathcal{B}_{10}$	$[0,\ 0.431]$	$[0.905, 1.168] \bigcup [1.305, 1.861]$
Model $\mathcal{D}_5$	$[0.248,\ 0.300]$	$[0.957,\ 1.056]$
Model $\mathcal{D}_6$	$[0,\ 0.300]$	$[0.957,\ 1.394]$
Model $\mathcal{D}_7$	$[0.026,\ 0.5]$	$[1.468,\ 3.006]$
Model $\mathcal{D}_8$	$[0.424,\ 0.5]$	$[0.872,\ 0.964]$
Model $\mathcal{D}_9$	$[0.033, 0.056] \bigcup [0.440, 0.469]$	$[0.887, 0.908] \bigcup [2.0, 2.282]$
Model $\mathcal{D}_{10}$	$[0.0307,\ 0.1175]$	$[1.996,\ 2.50]$

Table 2.15: Phenomenologically allowed regions of $(\mathrm{Re}\, U, \mathrm{Im}\, U)$ [129]. [68]

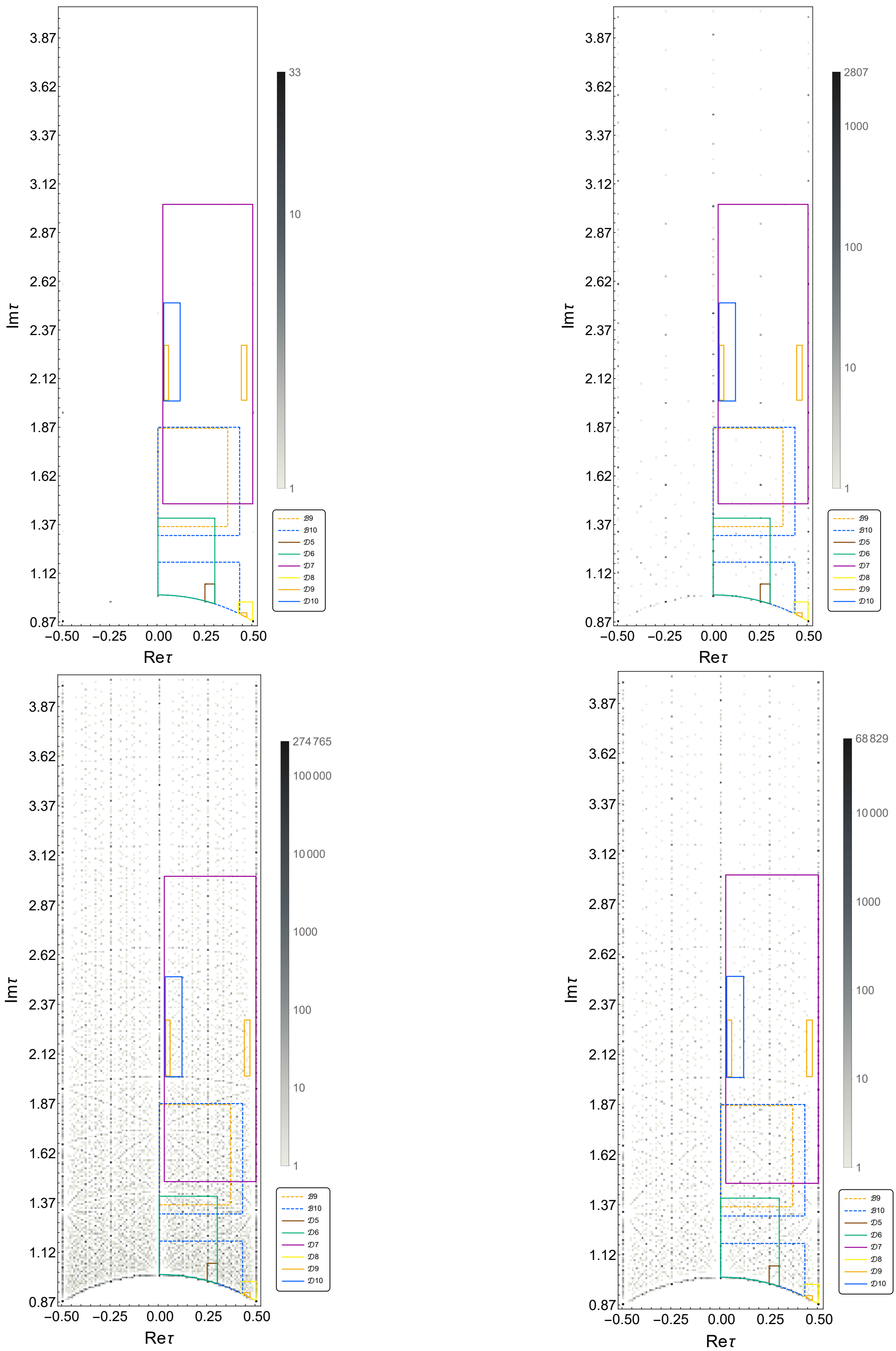

Figure 2.4: Phenomenologically predicted moduli values vs. the complex structure modulus from the string landscape for $N_{\text{flux}}^{\text{max}} = 10, 100, 500, 1000$, clockwise from top left. [68]

Models	Favorable (ReU, ImU)	Probabilities (%)		
		(i)	(ii)	(iii)
Model $\mathcal{B}_9$	$(0, \sqrt{3})$	80.000	7.547	9.434
Model $\mathcal{B}_{10}$	$(0, \sqrt{3})$	80.000	7.547	9.434
Model $\mathcal{D}_5$	None	-	-	0.000
Model $\mathcal{D}_6$	$(0, \sqrt{3/2})$	100.000	1.887	1.887
Model $\mathcal{D}_7$	$(1/2, \sqrt{15}/2)$	80.000	7.547	9.434
Model $\mathcal{D}_8$	None	-	-	0.000
Model $\mathcal{D}_9$	None	-	-	0.000
Model $\mathcal{D}_{10}$	None	-	-	0.000

Table 2.16: The theoretically favored VEVs of $(\mathrm{Re}\,U, \mathrm{Im}\,U)$ within each region for the $N_{\mathrm{flux}}^{\mathrm{max}} = 10$ case. We also show the probabilities of: (i) the favorable value in each model, (ii) the favorable value in the whole fundamental domain of $SL(2,\mathbb{Z})$, and (iii) the allowed moduli region in the string landscape. [68]

Models	Favorable (ReU, ImU)	Probabilities (%)		
		(i)	(ii)	(iii)
Model $\mathcal{B}_9$	$(0, \sqrt{3})$	64.560	7.556	11.704
Model $\mathcal{B}_{10}$	$(0, \sqrt{3})$	54.506	7.556	13.863
Model $\mathcal{D}_5$	$(3/10, \sqrt{111}/10)$	58.133	0.0886	0.152
Model $\mathcal{D}_6$	$(0, \sqrt{3/2})$	4.927	1.885	4.495
Model $\mathcal{D}_7$	$(1/2, \sqrt{15}/2)$	38.170	4.850	12.706
Model $\mathcal{D}_8$	$(1/2, \sqrt{11/12})$	56.216	0.0597	0.106
Model $\mathcal{D}_9$	$(1/22, \sqrt{2331}/22)$	4.762	0.000	0.003
Model $\mathcal{D}_{10}$	$(1/10, \sqrt{399}/10)$	16.466	0.006	0.037

Table 2.17: The theoretically favored VEVs of $(\mathrm{Re}\,U, \mathrm{Im}\,U)$ within each region for the $N_{\mathrm{flux}}^{\mathrm{max}} = 1000$ case. We also show the probabilities of: (i) the favorable value in each model, (ii) the favorable value in the whole fundamental domain of $SL(2,\mathbb{Z})$, and (iii) the allowed moduli region in the string landscape. [68]

2.3 Breaking Enhanced Symmetries in Landscape with Non-Perturbative Effects

Note: The whole discussion and results in this chapter are based on our work [69]. The figures and tables would be shown are already shown in Ref. [69] whence we will cite them constantly.

In the previous section, we investigated the toroidal landscape only with a tree scalar potential. There are $SL(2,\mathbb{Z})$ dualities in the complex-structure modulus space and the axio-dilaton space. With fixed $N_{\rm flux}$, the Landscape becomes finite, and thus probability is defined on vacua. Then, we also discussed compatibility between the predicted VEVs of the complex-structure modulus and the concrete modular flavor models.

However, there exist possible corrections for the scalar potential as mentioned in Section 2.1.1, and they may change our result drastically. For this reason, we must clarify the tolerance of moduli VEVs under such corrections. One possible yield is stabilized Kähler moduli, which has been neglected so far due to the no-scale structure of the scalar potential 2.22. In this section, we introduce KKLT-type non-perturbative effects [87] (see for a review, e.g., Section 5.6 in [23]) into the scalar potential, and show changes in Landscape statistics.

In the original set-up of KKLT, the superpotential is assumed to take the form

$$W = W_{\rm flux}(S,U) + Ce^{iaT}, \tag{2.125}$$

where the latter term is introduced by D-brane instanton effects or gaugino condensation on D7-branes. Here, C and a are numerical constants which are originated from the above effects and we do not determine their concrete origins. In this setup, we need a sufficiently small vacuum expectation value $W_0 := \langle W_{\rm flux}(S,U)\rangle$ in terms of the flux compactifications. Naively, it is difficult to realize such a small flux superpotential in that context, but several strategies have been proposed. (See for relevant studies, e.g., [131, 132].) We adopt a more generalized superpotential, which is written as

$$W = W_{\rm flux}(S,U) + W_{\rm np}(S,T), \tag{2.126}$$

where $W_{\rm np}$ takes the form

$$W_{\rm np} = \sum_m C_m e^{ia_m T + ib_m S}, \tag{2.127}$$

which is pointed out in [133] and further studied in [134]. There are D-brane instanton effects which lead $a_m, b_m = 2\pi$ and strong dynamics on D7-branes which lead $a_m = 2\pi/N$ for the rank of the relevant gauge group N (see for details, [133]). In the following, we do not consider their concrete origins, but we assume there is only one effect i.e., no summation, for simplicity.

KKLT setup originally assumes that the flux superpotential already has a vacuum expectation value (VEV), and it effectively integrates out the complex-structure moduli and the axio-dilaton. In that process, the VEVs of them are postulated to be nearly invariant, but we are interested in the small deviations of their VEVs in this study.

In the above and in what follows, it is assumed that internal space is approximated by the single Kähler modulus T whose Kähler potential is given by $K = -3\log\bigl(-i(T-\bar{T})\bigr)$ in the large volume regime.

2.3.1 Stabilization of volume moduli by non-perturbative effects

In this section, we discuss the deviations due to the Kähler moduli stabilization via the non-perturbative effect. The starting point is our previous study: the SUSY Minkowski vacua. Since we have the whole dataset of the flux landscape, we can determine a distribution of the deviations around the original SUSY Minkowski VEVs.

Hence, $\langle W_{\rm flux}\rangle = 0$, and the KKLT procedure gives an effective superpotential, which is written by

$$W_{\rm eff} = W_{\rm np}(\langle S\rangle, T) \tag{2.128}$$

where $\langle S\rangle$ is given by Eq. (2.99). Although it is absent here, $\langle U\rangle$ is similarly given by Eqs. (2.101) and (2.100). Henceforth, $W_{\rm np}$ is assumed to be the exponentially suppressed quantity

$$W_{\rm np} \simeq \langle e^{ibS}\rangle + Ce^{iaT}, \tag{2.129}$$

in the large volume regime. This yields the F-term of T as

$$D_T W_{\rm eff} := \partial_T W_{\rm eff} + K_T W_{\rm eff} = 0, \tag{2.130}$$

and the vacuum expectation value of T at this level is given by

$$a\langle T\rangle \simeq \log\frac{C}{W_0}, \tag{2.131}$$

where W_0 is defined as $W_0 := \langle e^{ibS}\rangle$. In the following, we further assume that $C = 1$ is a typical value.

As mentioned, the true vacuum is given as a solution for

$$D_I W = 0, \tag{2.132}$$

where W is a sum of the flux superpotential and the non-perturbative effect. Here, I runs S, U, and T. Thus we have deviations:

$$\begin{aligned} U &= \langle U\rangle + \delta U, \\ S &= \langle S\rangle + \delta S, \\ T &= \langle T\rangle + \delta T, \end{aligned} \tag{2.133}$$

where the VEVs are given by Eqs. (2.99), (2.101),(2.100), and (2.131). Furthermore, these deviations are exponentially suppressed as

$$\left.\frac{\delta\Phi^I}{\Phi^I}\right|_{\langle\rangle} \simeq \mathcal{O}(W_{\rm eff}). \tag{2.134}$$

So far, we considered the SUSY vacua which satisfies the F-term equations $D_I W = 0$. Then the broken no-scale structure leads to anti-de Sitter (AdS) vacua. Following the lines of the KKLT scenario, we introduce the SUSY-violating uplifting term via an anti-D3-brane. Note that, the other uplifting scenarios, such as F-term uplifting, might be incorporated into the discussion here. The uplifting term is given by

$$V_{\mathrm{up}} = \frac{D}{(-i(T-\overline{T}))^3}, \tag{2.135}$$

where the constant D depends on the number of the anti-D3-branes and the warp factor of the flux compactification. As the number of D3-branes should be changed by N_{flux} via the tadpole cancellation condition and the warp factor is controllable by varying some fluxes a là GKP solution, the constant D can be tuned to realize an ideal uplifting.

In the following, we numerically calculate these deviations from their initial VEVs, where the parameters are chosen as $a = b = 2\pi$ for concrete calculations, and D is chosen to realize Minkowski vacua.

2.3.2 Moduli values at nearby fixed points

In this section, we show our results of the numerical calculation. To analyze the deviations, we concentrate on the VEVs of U at $\langle U\rangle = \omega, i, 2i$ where the first two VEVs are statistically favored as observed in the previous section. Here, to see the vast distribution, the maximum value of N_{flux} is set as $N_{\mathrm{flux}}^{\mathrm{max}} = 1000$. Then, the solutions for $\partial_I V = 0$ with

$$V = e^K \left(K^{I\bar{J}} D_I W \overline{D_J W} - 3|W|^2 \right) + V_{\mathrm{up}} \tag{2.136}$$

are shown in Figures 2.5 to 2.7 [69] (because we cite the figures from our paper, the notation is slightly different from this thesis: τ denotes U i.e., the complex-structure modulus.). There, we show the absolute deviations and their angular dependencies. As shown, all the reference points favor $\sim O(10^{-5})$ absolute deviations, with regard to their peaks. On the other hand, there seems no radial dependency of the deviations. This means the CP symmetry $U \to -\overline{U}$ is violated at vacua (if $\arg \delta U$ is favored at 0 or π, we have CP-conserved vacua since the scalar potential is invariant under $SL(2,\mathbb{Z})_U$.) We additionally calculated the gravitino mass $m_{3/2} := e^{K/2} W$, and showed them in those figures with the box plots. We observe the typical SUSY breaking scale depends on the deviations linearly.

In our original paper [69], a concrete (but, bottom-up) A_4 modular flavor symmetric model is constructed. However, given the purpose of this paper, we will not go into the details here. In short, as far as we focus on the concrete model which is fitted to the observational and typically favored data (see for details of the fitting, [69] (our paper) and [135–137]), the observed deviations around $\tau = \omega$ (Figure 2.5) is not suitable for matching the deviations in the model. Furthermore, the deviation of U around $\langle U\rangle = i$ is also important to introduce a quasi-stable candidate of dark matter (DM) [138]. There, the natural upper bound of the allowed deviations $|\delta U| \lesssim 5.57 \times 10^{-19}$ is obtained with regard to the age of the universe 10^{17} sec, and this resides in the region that we obtained.

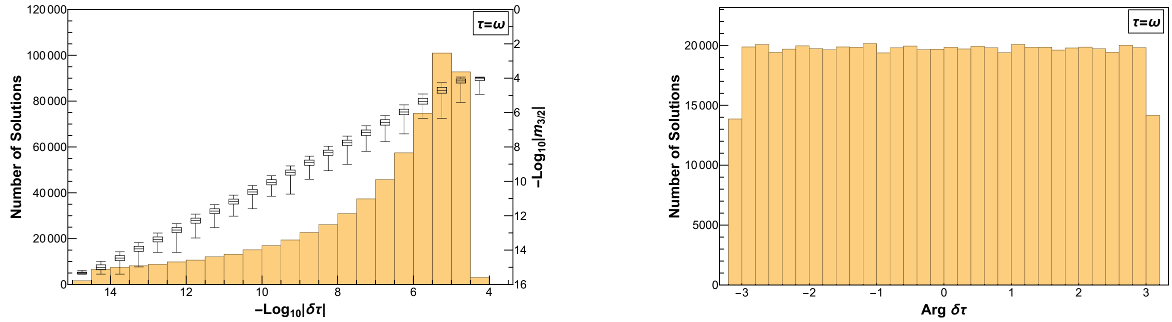

Figure 2.5: The number of vacua around $\langle U \rangle = \omega$. In the left panel, we show the absolute values $|\delta U|$, while the radial dependence $\mathrm{Arg}(\delta U)$ is shown in the right panel. We also show the gravitino mass in the left panel as a function of $|\delta U|$. Note that we cite these figures from [69], the notation is slightly different, where τ denotes U.

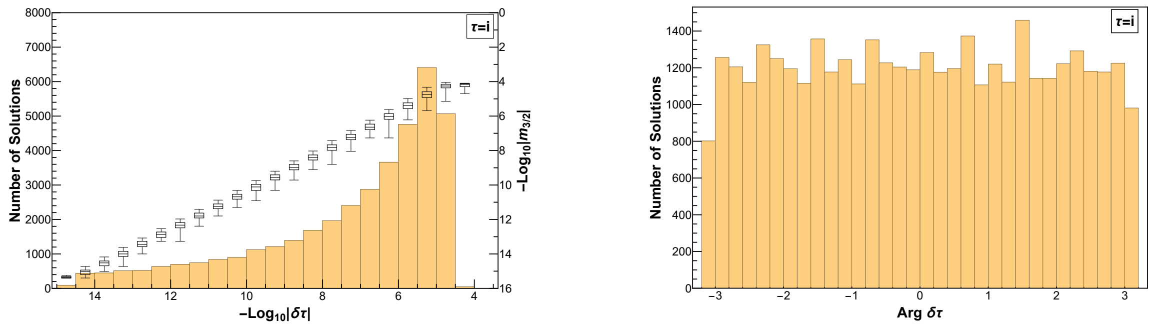

Figure 2.6: The number of vacua around $\langle U \rangle = i$. In the left panel, we show the absolute values $|\delta U|$, while the radial dependence $\mathrm{Arg}(\delta U)$ is shown in the right panel. We also show the gravitino mass in the left panel as a function of $|\delta U|$. Note that we cite these figures from [69], the notation is slightly different, where τ denotes U.

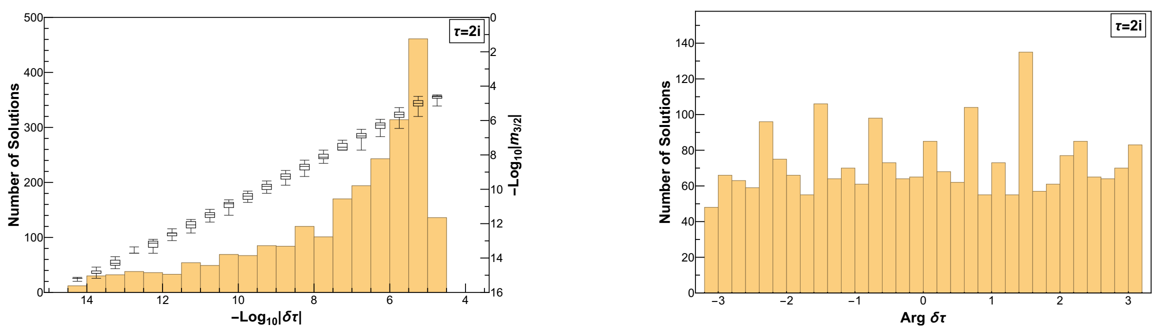

Figure 2.7: The number of vacua around $\langle U \rangle = 2i$. In the left panel, we show the absolute values $|\delta U|$, while the radial dependence $\mathrm{Arg}(\delta U)$ is shown in the right panel. We also show the gravitino mass in the left panel as a function of $|\delta U|$. Note that we cite these figures from [69], the notation is slightly different, where τ denotes U.

2.4 Exploring Other Toroidal Orientifolds

Note: The whole discussion and results in this chapter are based on our work [70]. The figures and tables would be shown are already shown in Ref. [70] whence we will cite them constantly.

So far, we have focused on the toroidal orientifold which probably has the simplest structure with the presence of complex-structure moduli as its factorizability indicates. We have explored the explicit distributions of moduli VEVs in the finite flux landscape and obtained some implications on the bottom-up modular flavor symmetric studies. For instance, the $T^6/(\mathbb{Z}_2 \times \mathbb{Z}_2')$ orientifold admits $\langle U \rangle = \omega$ and $\langle U \rangle = i$ in its flux landscape (henceforth, we omit the VEV symbol $\langle\rangle$). These values have been playgrounds of the bottom-up models, while the string landscape favors the former point statistically. To obtain $U = i$ in a natural way, we would have to find some specific mechanisms that lead to stabilizations of S on certain points, due to the correlation which was observed in Table 2.9.

However, we can naturally ask a question: do these results also hold for other background orientifolds? This is a fairly difficult question. Although we may have the finiteness theorem with regard to flux compactifications on general CY three-folds, it does not yield an explicit method for obtaining the whole flux landscape, as mentioned. The explicit distributions of flux landscape can only obtained when we know its geometrical details, especially its period integral. Thus, it is difficult to guess the flux landscape on general CY threefolds, while it is known that modular symmetry is usually promoted to symplectic symmetries, and the resultant fundamental functions will be given in terms of the Siegel modular forms. It is surely interesting to study the explicit distribution of moduli VEVs on CY three-folds. (See for an example on the mirror-dual of $\mathbb{P}_{11114}$, [113].)

On the other hand, the toroidal orientifolds, which have been studied for decades, have suitably fruitful geometrical structures while having highly controllable properties. Since the $T^6/(\mathbb{Z}_2 \times \mathbb{Z}_2')$ has a highly symmetric structure as its orbifold group suggests, we depart for a space with a more general orbifolding group. With this philosophy in mind, we decide to analyze $T^6/\mathbb{Z}_{6-\mathrm{II}}$ and $T^6/(\mathbb{Z}_2 \times \mathbb{Z}_4)$ orientifolds with certain lattices in this study (see for relevant papers, [14, 139–144]). Indeed, they are the simplest toroidal orientifolds as it has $h^{2,1} = 1$ i.e., only one complex-structure modulus. Analyzing flux vacua on these orientifolds has another advantage that we can stabilize all of the moduli (except for Kähler moduli) analytically. Furthermore, the former orientifold is non-factorizable whereas the latter one is factorizable. Thus, we can compare three orientifolds also in that point of view. As a result, the $T^6/\mathbb{Z}_{6-\mathrm{II}}$ orientifold turns out not to have the $SL(2,\mathbb{Z})$ modular symmetry with regard to U. Instead, the symmetry is partially broken into its certain subgroup, as its S-transformation disappears. Nevertheless, the finiteness of the flux landscape turns out to be ensured due to the presence of symmetry under a S-like transformation, S', which satisfies $S'^3 = 1$. Thus, we consider this orientifold a good starting point to discuss more general toroidal orientifolds and possible CY threefolds. In this study, we again ignore the strict satisfaction of the tadpole cancellation condition as the flux vacua have been studied in the same way [68,122]. We assume the existence of F-theory uplifting or some other justifications. However, it does not mean that the tadpole cancellation condition is irrelevant to our study at all. Indeed, the quantity which measures the flux contribution N_{flux} appears in the moduli stabilization in a manner of [83,111] (more generally, [72]), and VEVs of the axio-dilaton and the complex-structure modulus are related to N_{flux}. N_{flux} generically appears in the proof of the finiteness, as we will see later. Studying flux vacua on more general toroidal orientifolds is left for future work, but the quantity must appear in the constructive method of obtaining the vacua.

Moreover, this generalization to other toroidal orientifolds in this study seems to make new implications for the modular flavor symmetric models. As we will see later, the toroidal orientifolds with $h^{2,1}$ yields a linear superpotential with regard to the modulus (while there are couplings with the axio-dilaton). Then, the nontrivial orbifold group gives rise to complexified integer coefficients in the superpotential. Thus the SUSY Minkowski vacua do not exist on $U = \omega$ and $U = i$ simultaneously, because the modulus VEV takes its value in a quadratic field $\mathbb{Q}(\sqrt{n})$. In terms of modular symmetry, this structure forbids the simultaneous realization of two elliptic points in the landscape. Since there have been proposed many bottom-up studies around these elliptic points, this study implies that we can choose one of them by fixing the background orientifold. Furthermore, the modular symmetry can be broken, as we will see later on $T^6/\mathbb{Z}_{6-\text{II}}$ orientifold. As a result, the most favored point in the landscape, which has the highest degeneracy of vacua, is not an elliptic point of $SL(2,\mathbb{Z})$ modular group. Nonetheless, the favored point is still an elliptic point, but that of the broken certain congruence subgroup. Furthermore, on that favored point, a $\mathbb{Z}_2$ enhanced symmetry is observed and it might be related to some phenomenological studies. We will also see that an outer semidirect product of the congruence subgroup is the entire symmetry of the moduli space, and it will lead to a generalization of the usual modular flavor symmetric models which use the modular functions and their symplectic generalizations as fundamental building blocks. Constructing modular forms and modular flavor symmetric models on the orientifolds is beyond the scope of our current study, but we expect that our results will provide new insights into the context.

This section is organized as follows. In section 2.4.1, we review the geometry of the $T^6/\mathbb{Z}_{6-\text{II}}$ orientifold with a certain lattice and the flux compactification on it. Throughout this study, we concentrate on the SUSY Minkowski solutions, and we reveal all the dualities that the F-term equations possess. There, we show that the symplectic basis of the orientifold is only consistent with the certain congruence subgroup and its outer semidirect product, as mentioned. Note that, we call symmetries of equations of motions dualities (in the sense it does not mean a symmetry of the action or the theory). The modular symmetry is a symmetry in this sense since we fix the backgrounds as orientifolds whose geometries respect the modular symmetry. The finiteness of the landscape is also shown, via the S-duality and the "broken-and-enlarged" symmetry of the complex-structure modulus. Then, in section 2.4.2, we show the whole flux landscape on the orientifold, which is physical in the sense that they are modded by the symmetries. As mentioned, the most favored point is at an elliptic point which is not of the modular group $SL(2,\mathbb{Z})$. Rather, that is an elliptic point of the outer semidirect product. In section 2.4.4, we analyze the $T^6/(\mathbb{Z}_2 \times \mathbb{Z}_4)$ orientifold whose geometry is factorizable. On the orientifold, the usual modular group $SL(2,\mathbb{Z})$ for the complex-structure modulus turns out to be realized, and it is compared with the landscape of the former orientifold. Note that, both orientifolds yield nontrivial liner superpotentials, where two elliptic points cannot be obtained simultaneously.

2.4.1 Flux compactifications on $T^6/\mathbb{Z}_{6-\mathrm{II}}$

For the moment, we concentrate on the orientifold theory of $T^6/\mathbb{Z}_{6-\mathrm{II}}$ with $SU(6)\times SU(2)$ lattice. The flux compactification on this orientifold gives rise to nontrivial flux quanta.

2.4.1.1 Geometry

We first briefly review the geometry (see for details, [142,144]). The Hodge numbers of the orientifold are known to be $(h^{1,1},h^{2,1})=(25,1)$, and there are only untwisted contribution to $h^{2,1}$.

Although we have already studied $T^6/(\mathbb{Z}_2\times\mathbb{Z}_2')$ orientifold in the previous sections, we have some knowledge of the moduli space of toroidal orientifolds. Nevertheless, we need its details in order to construct a well-defined cohomology group on the orbifold. Therefore, let us summarize the basic geometry once again.

On the underlying tori T^6, we have 20 real three-forms which are elements of a third-cohomology basis,

$$\begin{aligned}
\alpha_0 &= dx^1\wedge dx^3\wedge dx^5, & \beta^0 &= dx^2\wedge dx^4\wedge dx^6,\\
\alpha_1 &= dx^2\wedge dx^3\wedge dx^5, & \beta^1 &= -dx^1\wedge dx^4\wedge dx^6,\\
\alpha_2 &= dx^1\wedge dx^4\wedge dx^5, & \beta^2 &= -dx^2\wedge dx^3\wedge dx^6,\\
\alpha_3 &= dx^1\wedge dx^3\wedge dx^6, & \beta^3 &= -dx^2\wedge dx^4\wedge dx^5,\\
\\
\gamma_1 &= dx^1\wedge dx^2\wedge dx^3, & \delta^1 &= -dx^4\wedge dx^5\wedge dx^6,\\
\gamma_2 &= dx^1\wedge dx^2\wedge dx^5, & \delta^2 &= -dx^3\wedge dx^4\wedge dx^6,\\
\gamma_3 &= dx^1\wedge dx^3\wedge dx^4, & \delta^3 &= -dx^2\wedge dx^5\wedge dx^6,\\
\gamma_4 &= dx^3\wedge dx^4\wedge dx^5, & \delta^4 &= -dx^1\wedge dx^2\wedge dx^6,\\
\gamma_5 &= dx^1\wedge dx^5\wedge dx^6, & \delta^5 &= -dx^2\wedge dx^3\wedge dx^4,\\
\gamma_6 &= dx^3\wedge dx^5\wedge dx^6, & \delta^6 &= -dx^1\wedge dx^2\wedge dx^4.
\end{aligned} \tag{2.137}$$

Then, we fix the orientation of the toroidal space by

$$\int_{T^6} dx^1\wedge dx^2\wedge dx^3\wedge dx^4\wedge dx^5\wedge dx^6 = -1. \tag{2.138}$$

The above basis satisfies

$$\int_{T^6}\alpha_i\wedge\beta^j=\delta^i_j,\quad \int_{T^6}\gamma_i\wedge\delta^j=\delta^i_j, \tag{2.139}$$

under the orientation. To define the unique holomorphic three-form Ω, we need coordinates. Indeed, the real coordinates are defined via the lattice vectors, which are transformed by the orbifold group $\Gamma=\mathbb{Z}_{6-\mathrm{II}}$ (we call it a twist). In other words, the lattice vector space equips nontrivial representations under the orbifold group, and we can find the survived singlets among them. The orbifold group is characterized via the Coxeter element Q of the corresponding root lattice. For the $SU(6) \times SU(2)$ case, it is given by the following actions:

$$\begin{aligned} Q(e_i) &= e_{i+1} \quad (i = 1, 2, 3, 4), \\ Q(e_5) &= -e_1 - e_2 - e_3 - e_4 - e_5, \\ Q(e_6) &= -e_6, \end{aligned} \tag{2.140}$$

where e_i $(i = 1, 2, \dots, 6)$ are the lattice vectors of the orientifold. To illustrate the orbifold group, we define a matrix representation by $Q(e_i) \coloneqq Q_{ji} e_j$ (thus, the contravariant vectors (1-forms) dx^i are transformed as $d\mathbf{x}' = Q d\mathbf{x}$). The representation is written by

$$Q = \begin{pmatrix} 0 & 0 & 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}. \tag{2.141}$$

while the $\mathbb{Z}_6$ property is seen as $Q^6 = 1$.

Having fixed the orbifold group actions, we can move to the complex coordinates as[13]

$$\begin{aligned} dz^1 &= dx^1 + e^{-\frac{2\pi i}{6}} dx^2 + e^{-\frac{2\pi i}{3}} dx^3 - dx^4 + e^{\frac{2\pi i}{3}} dx^5, \\ dz^2 &= dx^1 + e^{-\frac{2\pi i}{3}} dx^2 + e^{\frac{2\pi i}{3}} dx^3 + dx^4 + e^{-\frac{2\pi i}{3}} dx^5, \\ dz^3 &= \frac{1}{2\sqrt{3}} \left[\frac{1}{3} (dx^1 - dx^2 + dx^3 - dx^4 + dx^5) + U dx^6 \right]. \end{aligned} \tag{2.142}$$

by requiring that the complex coordinates transform as diagonally, i.e., $z^i \to e^{2\pi i v^i} z^i$ with $(v^1, v^2, v^3) = \left(-\frac{1}{6}, -\frac{2}{6}, \frac{3}{6}\right)$, where $v^1 + v^2 + v^3 = 0$ is a result of the $N = 1$ SUSY condition on toroidal orientifolds[14]. We see in Eq. 2.142 that a single complex-structure modulus U is not fixed by the orbifold action and remains as a dynamical d.o.f. The non-vanishing and unique holomorphic three-form Ω is defined as

$$\Omega \coloneqq dz^1 \wedge dz^2 \wedge dz^3, \tag{2.143}$$

for the complex coordinates, as usual. Note that our notation yields

$$-i \int \Omega \wedge \bar{\Omega} = 2 \mathrm{Im} U, \tag{2.144}$$

while the overall constant does not make sense due to the Kähler transformation.

The orientifold action $\hat{\Omega} \mathcal{R} (-1)^{F_L}$ is defined in the same way as in the previous studies, i.e., $\mathcal{R} : (z^1, z^2, z^3) \to -(z^1, z^2, z^3)$ as $\mathbb{Z}_2$ isometry of the orbifold. Then, the cohomology group H^2 and H^3 are decomposed into their even/odd parts. In the orientifold which we consider, we have only $h^{1,1}_{\text{untw.},+}$ and $h^{2,1}_{\text{untw.},-}$. In the following, we constantly suppress the labels $\pm$ and untw.

[13] We do not normalize in the form which respects a so-called volume factor and adopt the normalization of [142]. See for a review of the former normalization, [145].

[14] The vanishing summation of the twists v^i is a not result, but rather a result of our requirement.

2.4.1.2 Three-form flux G_3 and the scalar potential

Next, we summarize the details of the flux compactification on the orientifold. The $N = 1$ SUGRA ingredients are defined as we explained in Sec. 2.1.1, and Ω is given by Eq. (2.143).

The flux quanta and G_3 need some calculations. Because of the Dirac quantization, F_3, H_3 are expanded by the basis of H^3_- [15]. Following the lines of [142], we define the coefficients of the superpotential by

$$W = A + BS + U\left[C + DS\right] := \int G_3 \wedge \Omega. \tag{2.145}$$

This can be done since there is only one complex-structure modulus. Note that this is equivalent to the definitions

$$A + CU := \int F_3 \wedge \Omega, \quad -(B + DU) := \int H_3 \wedge \Omega, \tag{2.146}$$

where C, D are defined as the derivatives of W_{RR} and W_{NS}, with respect to U. As explicitly checked via the basis, we have the Hodge decomposition of the Dolbeault cohomology. Thus G_3 can be expanded as

$$\begin{aligned} G_3 = \frac{i}{2\mathrm{Im}U}\big[&- (\overline{A} + \overline{B}S + \overline{U}(\overline{C} + \overline{D}S))\omega_{A_0} \\ &+ (A + BS + U(C + DS))\omega_{B_0} \\ &+ (\overline{A} + \overline{B}S + U(\overline{C} + \overline{D}S))\omega_{A_3} \\ &- (A + BS + \overline{U}(C + DS))\omega_{B_3}\big], \end{aligned} \tag{2.147}$$

for the complex basis

$$\begin{aligned} \omega_{A_0} &= dz^1 \wedge dz^2 \wedge dz^3, \\ \omega_{A_3} &= dz^1 \wedge dz^2 \wedge d\overline{z}^3, \\ \omega_{B_3} &= \overline{\omega_{A_3}}, \\ \omega_{B_0} &= \overline{\omega_{A_0}}. \end{aligned} \tag{2.148}$$

Note that the above integrals are defined on the ambient (underlying) tori, but this only affects the results by an overall constant. However, the flux quantization conditions become more subtle. We define the flux quanta as integrals on the ambient space:

$$\begin{aligned} &\int_{T^6} (F_3, H_3) \wedge \alpha_i \equiv (-b_i, -d_i), \quad \int_{T^6} (F_3, H_3) \wedge \beta^i \equiv (a^i, c^i) \\ &\int_{T^6} (F_3, H_3) \wedge \gamma_i \equiv (-f_i, -h_i), \quad \int_{T^6} (F_3, H_3) \wedge \delta^i \equiv (e^i, g^i). \end{aligned} \tag{2.149}$$

[15] We would like to emphasize the total parity of fluxes becomes eventually even, since there is another contribution from the intrinsic parity $\Omega'(-1)^{F_L}$ (see for a nice summary, [83]).

The domain of the integration is important. Since there are no twisted contributions, we do not have to care about a problem which occurs in the presence of them (see for details, [146]). On the other hand, since we consider the orbifolding, cycles with shorter lengths naturally arise on the orbifold (such cycles are called fractional). Then let us define the cycles on the orbifold[16] in the following.

As mentioned, there are certain combinations of the cohomology basis (or, equivalently, the homology basis) that transform as singlets under the twist. To obtain singlets, we need to simply sum up each orbit of the orbifold group, as follows[17]:

$$\begin{aligned} \mathbf{1}_1 &\coloneqq \sum \Gamma(\alpha_0), \\ \mathbf{1}_2 &\coloneqq \sum \Gamma(\alpha_1), \\ \mathbf{1}_3 &\coloneqq -\left(\sum \Gamma(\beta^1) + \sum \Gamma(\beta^0)\right), \\ \mathbf{1}_4 &\coloneqq \sum \Gamma(\beta^0), \end{aligned} \tag{2.150}$$

where $G(x) \coloneqq \{gx \mid g \in G\}$ denotes an orbit of $x \in \mathbb{X}$ under a group G and the action of Γ on the three-forms is induced by Q (2.141). Note that, sometimes the term "orbit" means a length $|G|$ collection of all gx (not a *set*), or it's summation in physics literature. Then, the resultant cycles are inherited from the ambient space, and the fractional cycles are defined as the "orbits" with possible normalization factors. These cycles can exist when the intersection matrix between the "orbits" is not unimodular since we can define an integral (co-)homology group on the orbifold.

Indeed, the singlets satisfy

$$\int_{T^6} \mathbf{1}_1 \wedge \mathbf{1}_4 = 6, \quad \int_{T^6} \mathbf{1}_2 \wedge \mathbf{1}_3 = -6, \tag{2.151}$$

where other intersections are zero. Thus, those singlets form an unimodular matrix when they are measured on the orbifold (the integrals are divided by six). In other words, the singlets form a symplectic basis of the cohomology group on the orbifold. Since there is the identification under the twist, all the three-forms that belong to the same singlet must give rise to the same flux number. Thus, the expansion of G_3 in this singlet basis is given by

$$G_3 = \frac{1}{3}(a^0 - c^0 S)\mathbf{1}_1 + [(a^1 - c^1 S) + \frac{1}{3}(a^0 - c^0 S)]\mathbf{1}_2 - \frac{1}{2}(b_1 - d_1 S)\mathbf{1}_3 + (b_0 - d_0 S)\mathbf{1}_4, \tag{2.152}$$

where the representatives are chosen as $(a^0, a^1, b_0, b_1, c^0, c^1, d_0, d_1)$. Here, the coefficients are determined via the lengths of the orbits:

$$|\Gamma(\alpha_0)| = 6, |\Gamma(\alpha_1)| = 3, |\Gamma(\beta^0)| = 2, |\Gamma(\beta^1)| = 6, \tag{2.153}$$

[16] We first find well-defined cycles on the orbifold. Since there is no $h^{2,1}_+$ d.o.f. on the orientifold, the orientifold projection is almost trivial.

[17] The relative signs among the definitions are chosen to replicate the result of [142], and the representatives can be chosen freely.

and their intersection matrix.

The relation of the singlets to the real basis and the complex basis are as follows:

$$\begin{aligned}
\mathbf{1}_1 &= 3\alpha_0 - \alpha_1 - \alpha_2 + \beta^3 - \delta^5 + \delta^6 + \gamma_1 - \gamma_2 - \gamma_3 + \gamma_4, \\
\mathbf{1}_2 &= \alpha_1 + \alpha_2 + \beta^3 - \delta^6 - \gamma_2 - \gamma_3, \\
\mathbf{1}_3 &= -2\beta^1 - \beta^2 - \delta^1 + \delta^2 + 2\delta^3 + \delta^4, \\
\mathbf{1}_4 &= -\alpha_3 + \beta^0 + \gamma_5 - \gamma_6,
\end{aligned} \tag{2.154}$$

$$\begin{pmatrix} \omega_{A_0} \\ \omega_{A_3} \\ \omega_{B_3} \\ \omega_{B_0} \end{pmatrix} = \frac{1}{6} \begin{pmatrix} i & \sqrt{3} & \sqrt{3}U & -3iU \\ i & \sqrt{3} & \sqrt{3}\overline{U} & -3i\overline{U} \\ -i & \sqrt{3} & \sqrt{3}U & 3iU \\ -i & \sqrt{3} & \sqrt{3}\overline{U} & 3i\overline{U} \end{pmatrix} \begin{pmatrix} \mathbf{1}_1 \\ \mathbf{1}_2 \\ \mathbf{1}_3 \\ \mathbf{1}_4 \end{pmatrix}. \tag{2.155}$$

Thus, the coefficients A, B, C, D of the superpotential are expressed in terms of the flux quanta as

$$\begin{aligned}
A &= -ib_0 - \frac{\sqrt{3}}{2}b_1, \quad C = -ia^0 - \left(\sqrt{3}a^1 + \frac{1}{\sqrt{3}}a^0\right), \\
B &= id_0 + \frac{\sqrt{3}}{2}d_1, \qquad D = ic^0 + \left(\sqrt{3}c^1 + \frac{1}{\sqrt{3}}c^0\right).
\end{aligned} \tag{2.156}$$

We need to quantize the fluxes on the orbifold due to the identification. Then, the fundamental cycles are the singlets, and the fluxes are quantized on those cycles. Further, we follow [83, 144, 147, 148] and consider only even flux quanta since it is pointed out that the odd quanta may lead to exotic O3-planes. Besides the factor of 2 from this requirement, there are additional factors $N/|\Gamma(\xi)|$ that come from the definition of the flux quanta (2.149). Here, N is the order $\mathbb{Z}_N$ and ξ is a three-form whose Poincaré dual is the cycle where the considering flux is quantized[18]. For the $Z_{6-\mathrm{II}}$ orbifold with $SU(6) \times SU(2)$ lattice, we obtain

$$(b_0, d_0) \in 2\mathbb{Z}, (b_1, d_1) \in 4\mathbb{Z}, (a_0, c_0) \in 6\mathbb{Z}, (a_1, c_1) \in 2\mathbb{Z}. \tag{2.157}$$

Note that these non-aligned quantization conditions come from the nontrivial twist and do not occur on the $\mathbb{Z}_2 \times \mathbb{Z}_2'$ orbifold.

This quantization condition implies

$$\begin{aligned}
N_{\text{flux}} &= \left(\int \Omega \wedge \overline{\Omega}\right)^{-1} \left[\left(\overline{\int F_3 \wedge \Omega}\right)\left(\int H_3 \wedge \Omega\right) - (K_{U\overline{U}})^{-1}\left(D_{\overline{U}}\overline{\int F_3 \wedge \Omega}\right)\left(D_U \int H_3 \wedge \Omega\right)\right] \\
&= -2\mathrm{Re}(A\overline{D}) + 2\mathrm{Re}(B\overline{C}) \qquad (2.158) \\
&= 2b_0c^0 + b_1c^0 + 3b_1c^1 - 2a^0d_0 - a^0d_1 - 3a^1d_1 \in 24\mathbb{Z}, \qquad (2.159)
\end{aligned}$$

where $N_{\text{flux}} > 0$ follows by the ISD condition[19], and it implies a large contribution to the tadpole cancellation condition.

[18] The factor is ensured to be an integer via the Lagrange's theorem.

[19] $N_{\text{flux}} = 0$ is not allowed at SUSY Minkowski vacua in general [72].

2.4.1.3 Dualities in the F-term equations

Having obtained the fluxes, we can discuss the F-term equations. Besides the manifest dualities of the theory, we have several dualities by which we do not mod out the landscape, as discussed below. We need to clarify the dualities to obtain the finite landscape, as pointed out in [40][20].

Manifest Dualities
The overall signs of the fluxes do not make sense:

$$F_3 \to -F_3, \quad H_3 \to -H_3. \tag{2.160}$$

We mod out the landscape by this symmetry. Indeed, all relevant solutions and distributions of quantities are simply doubled by this symmetry, except for the trivial case where all the fluxes vanish.

As mentioned in Section 2.1.1, there is manifest S-duality in the effective action[21]. Considering integer fluxes, this appears as the modular group $SL(2,\mathbb{Z})_S$ under which the axio-dilaton and the flux quanta transform as

$$S \to S' = \frac{aS+b}{cS+d}, \quad \begin{pmatrix} F_3 \\ H_3 \end{pmatrix} \to \begin{pmatrix} F_3' \\ H_3' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} F_3 \\ H_3 \end{pmatrix}, \tag{2.161}$$

where the integers a, b, c, d satisfies $ad - bc = 1$. Under the transformation, the coefficients of W (2.145) transform as

$$\begin{pmatrix} A' \\ B' \end{pmatrix} = \begin{pmatrix} a & -b \\ -c & d \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix}, \quad \begin{pmatrix} C' \\ D' \end{pmatrix} = \begin{pmatrix} a & -b \\ -c & d \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix}. \tag{2.162}$$

Broken $SL(2,\mathbb{Z})$ for U
On the $\mathbb{Z}_2 \times \mathbb{Z}_2'$ orbifold, there is the apparent modular $SL(2,\mathbb{Z})_U$ symmetry in the complex-structure moduli space. However, this symmetry turns out broken on the orbifold as shown in the following. To see this, consider the F-term equation with respect to U:

$$D_U W = -\frac{1}{U - \overline{U}} \left(A + BS + \overline{U}(C + DS) \right) = 0, \tag{2.163}$$

[20] In the context, one often uses the term "duality" referring to a symmetry by which one mods out the landscape. In principle, one cannot mod out by dualities since they connect different theories. For example, the theory has a manifest S-duality but we cannot discuss the strong coupling regime in this context. Thus, we do not mod out by the S-duality but simply ignore inside the unit circle in the first place (and we will discuss sufficiently weak coupling regimes later). Although the use of the term might be misleading, we call symmetries of the action or equations of motion simply dualities. However, we will mod out the landscape by only the symmetries of the theory.

[21] Note that, we do not consider the non-geometric fluxes in this thesis. See [149] for details of self S-dual theory in the compactifications with the non-geometric fluxes.

and the transformation of the equation under $SL(2,\mathbb{Z})_U$ transformation

$$U \to U' = \frac{aU+b}{cU+d}. \tag{2.164}$$

Then, one finds that the equation is invariant when the flux quanta are transformed as[22]

$$\begin{pmatrix} A' \\ C' \end{pmatrix} = \begin{pmatrix} a & -b \\ -c & d \end{pmatrix} \begin{pmatrix} A \\ C \end{pmatrix}, \quad \begin{pmatrix} B' \\ D' \end{pmatrix} = \begin{pmatrix} a & -b \\ -c & d \end{pmatrix} \begin{pmatrix} B \\ D \end{pmatrix}, \tag{2.165}$$

where A, B, C, D again refer to the coefficients of W.

In terms of the real flux quanta(2.149), we have the following transformations under above dualities:
For $SL(2,\mathbb{Z})_S$,

- $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in SL(2,\mathbb{Z})_S$

$$\begin{aligned} &a^0 \to -c^0, \quad && b_0 \to -d_0, \\ &a^1 \to -c^1, && b_1 \to -d_1, \\ &c^0 \to a^0, && d_0 \to b_0, \\ &c^1 \to a^1, && d_1 \to b_1. \end{aligned} \tag{2.166}$$

- $T^q = \begin{pmatrix} 1 & q \\ 0 & 1 \end{pmatrix} \in SL(2,\mathbb{Z})_S$

$$\begin{aligned} &a^0 \to a^0 + qc^0, \quad && b_0 \to b_0 + qd_0, \\ &a^1 \to a^1 + qc^1, && b_1 \to b_1 + qd_1, \\ &c^0 \to c^0, && d_0 \to d_0, \\ &c^1 \to c^1, && d_1 \to d_1. \end{aligned} \tag{2.167}$$

For $SL(2,\mathbb{Z})_U$,

- $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in SL(2,\mathbb{Z})_U$

$$\begin{aligned} &a^0 \to -b_0, \quad && b_0 \to a^0, \\ &a^1 \to \frac{b_0}{3} - \frac{b_1}{2}, && b_1 \to 2a^1 + \frac{2}{3}a_0, \\ &c^0 \to -d_0, && d_0 \to c_0, \\ &c^1 \to \frac{d_0}{3} - \frac{d_1}{2}, && d_1 \to \frac{2}{3}c^0 + 2c^1. \end{aligned} \tag{2.168}$$

[22]Since this is a duality of the equation, there is a d.o.f. of an overall factor. As a result, the transformation rules are defined up to the undetermined factor. To determine it, one has to consider the period vector since the Kähler transformation of Ω would explain the factor.

- $T^q = \begin{pmatrix} 1 & q \\ 0 & 1 \end{pmatrix} \in SL(2,\mathbb{Z})_U$

$$\begin{aligned} a^0 &\to a^0, \quad b_0 \to b_0 - qa^0, \\ a^1 &\to a^1, \quad b_1 \to b_1 - q\left(\frac{2}{3}a^0 + 2a^1\right), \\ c^0 &\to c^0, \quad d_0 \to d_0 - qc^0, \\ c^1 &\to c^1, \quad d_1 \to d_1 - q\left(\frac{2}{3}c^0 + 2c^1\right). \end{aligned} \tag{2.169}$$

One can see that non-trivial rational numbers appear as coefficients in the transformation rules for $SL(2,\mathbb{Z})_U$, as shown in in Eqs. (2.205) and (2.169). As the flux quanta are integers, this means the F-term equation with respect to U is not invariant under the whole modular $SL(2,\mathbb{Z})_U$ group. Indeed, one can check that the flux quantization condition (2.157) is violated by the $S \in SL(2,\mathbb{Z})_U$.

One possible workaround is assuming only the flux quanta that are sufficiently large to cancel out the denominators. However, it is peculiar from the viewpoint of fundamental cycles. Indeed, both the flux quanta and the complex-structure modulus are defined in terms of the integrals on the cycles. Thus, it is more natural to review this symmetry at the period vector level.

The period vector is given by

$$\Pi = \begin{pmatrix} \int_{A^0}\Omega \\ \int_{A^1}\Omega \\ \int_{B^0}\Omega \\ \int_{B^1}\Omega \end{pmatrix} = \begin{pmatrix} \int_{T^6/\mathbb{Z}_{6-\mathrm{II}}} \Omega \wedge \mathbf{1}_4 \\ \int_{T^6/\mathbb{Z}_{6-\mathrm{II}}} \Omega \wedge (-\mathbf{1}_3) \\ \int_{T^6/\mathbb{Z}_{6-\mathrm{II}}} \Omega \wedge \mathbf{1}_1 \\ \int_{T^6/\mathbb{Z}_{6-\mathrm{II}}} \Omega \wedge \mathbf{1}_2 \end{pmatrix} = -\frac{1}{6}\begin{pmatrix} i \\ \sqrt{3} \\ 3iU \\ \sqrt{3}U \end{pmatrix}, \tag{2.170}$$

where the Poincaré dual cycles form the corresponding symplectic basis. For convenience, we switch to the basis where the definition of U and the derivative of the prepotential become clear. After a Kähler transformation and the symplectic transformation $Sp(b_3,\mathbb{Z})$, we obtain $\Pi' = (1, -\sqrt{3}iU, 3U, \sqrt{3}i)^T$. Then, its transformation under the modular $SL(2,\mathbb{Z})_U$ is given by

$$\Pi' \to (cU+d)^{-1}\begin{pmatrix} d & 0 & \frac{c}{3} & 0 \\ 0 & a & 0 & -b \\ 3b & 0 & a & 0 \\ 0 & -c & 0 & d \end{pmatrix}\Pi' =: (cU+d)^{-1}M\Pi'. \tag{2.171}$$

One can see the fractional coefficient also enters in the transformation matrix of the period vector. For general c we have $M \in Sp(b_3,\mathbb{Q})$ instead of $Sp(b_3,\mathbb{Z})$ and the fractional coefficient in the flux transformation rules (2.168, 2.169) are come from the matrix M as

$$\int G_3 \wedge \Omega = \mathbb{G}_3'^T \Sigma \Pi : \text{ invariant} \Rightarrow \mathbb{G}_3' \to M^T \mathbb{G}_3', \tag{2.172}$$

where Σ is given by Eq. (2.59) and $\mathbb{G}_3'$ denotes the period integral of G_3 on this basis ('). Since the basis is constructed to be symplectic, for a symmetry we would have found $M \in Sp(b_3,\mathbb{Z})$.

Thus, the transformation implies that $H_3(T^6/\mathbb{Z}_{6-\mathrm{II}}, \mathbb{Z})$ is not compatible with the modular symmetry, and it is more natural to postulate that $c \equiv 3 \pmod 3$ is required for the symmetry. This subgroup of the modular symmetry has its own name, the Hecke congruence subgroup of level n with $n = 3$[23], where the family is defined as

$$\Gamma_0(n) \equiv \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2,\mathbb{Z}) \middle| \, c \equiv 0 \mod n \right\}. \tag{2.173}$$

Although S-transformation is not contained in the congruence subgroup, the $\Gamma_0(3)$ is generated by its two generators[24], T and S':

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad S' = \begin{pmatrix} -1 & 1 \\ -3 & 2 \end{pmatrix}. \tag{2.174}$$

Here, T is the unbroken usual T-transformation and $S' = S^{-1}T^{-3}S^{-1}T^{-1}$ satisfies $(S')^3 = -\mathbf{1}_{2\times 2}$. Note that,

$$U = \omega' := -\frac{1}{2} + i\frac{1}{2\sqrt{3}} \tag{2.175}$$

is the fixed point under S'. Since the modular transformation of U is actually realized as PSL transformation, the residual symmetry at a fixed point is $\mathbb{Z}_3$. The transformation rules of the flux quanta for $\Gamma_0(3)$ is summarized as follows: For the S'-transformation of $\Gamma_0(3)_U$,

- $S' = \begin{pmatrix} -1 & 1 \\ -3 & 2 \end{pmatrix} \in \Gamma_0(3)_U$

$$\begin{aligned} a^0 &\to 2a^0 + 3b_0, & b_0 &\to -b_0 - a^0, \\ a^1 &\to 2a^1 - b_0 + \frac{3}{2}b_1, & b_1 &\to -b_1 - \frac{2}{3}a_0 - 2a_1, \\ c^0 &\to 2c^0 + 3d_0, & d_0 &\to -d_0 - c^0, \\ c^1 &\to 2c^1 - d_0 + \frac{3}{2}d_1, & d_1 &\to -d_1 - \frac{2}{3}c^0 - 2c^1, \end{aligned} \tag{2.176}$$

and the transformation rules for the T-transformation is given by Eq. (2.169). Indeed, for arbitrary fluxes, the flux quantization condition (2.157) is preserved under these transformations. Moreover, the combination of $(S')^3 = -1$ and the overall sign reversion of fluxes becomes consistent with the $\overline{\Gamma}_0(3)$ symmetry that effectively acts on the complex-structure modulus.

Note that, under the flux transformations which are compatible with the symplectic transformations, the tadpole charge N_{flux} remains invariant.

[23]Note that it was pointed out in [150] that the Hecke congruence subgroups and similar ones appear in the partition functions for $T^6/\mathbb{Z}_N$ with $N = 4, 6$, while the authors of [150] chose a different lattice for $N = 6$.

[24]The generators and corresponding presentations of the Fuchsian subgroup $\overline{\Gamma}_0(N)$ of $PSL(2,\mathbb{Z})$ are shown in [151], The `GAP` system [152] also gives the explicit generators.

"Scaling" duality $S_{(3)} \coloneqq U \to -\frac{1}{3U}$
Besides $\Gamma_0(3)$, we find that another duality is realized on the complex-structure modulus space, namely "scaling" duality. We denote it by $S_{(3)}$ in the following and it acts on U as

$$S_{(3)} : U \to -\frac{1}{3U}, \tag{2.177}$$

i.e., it induces a $\mathbb{Z}_2$ action. Indeed, N_{flux} remains invariant, and $\mathbb{G}_3$ transforms appropriately to preserve the flux quantization condition. At the period vector level, this symmetry comes from a d.o.f. in constructing the transformation matrix (2.171) in terms of integers a, b, c, d. Although the original transformation matrix is not an element of $SL(2,\mathbb{Z})$, by arranging the integers in different positions, up to a Kähler transformation, we obtain a symplectic transformation $\in Sp(b_3, \mathbb{Z})$ as

$$\Pi \to (cU+d)^{-1}(i\sqrt{3}) \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \Pi'. \tag{2.178}$$

Thus, $S_{(3)}$ is a symmetry by which we mod out the landscape. Nevertheless, the $S_(3)$ is exceptional from the view point of $SL(2,\mathbb{Z})$, as $S_{(3)} \notin SL(2,\mathbb{Z})$. Rather, it can be interpreted as an element of $SL(2,\mathbb{R})$: $a, d = 0$ and $b = -\frac{1}{\sqrt{3}}, c = \sqrt{3}$ [25]. For general $SL(2,\mathbb{R})$ elements, $Sp(4,\mathbb{Z})$ would not be compatible. However, since we have a d.o.f. to include an overall factor in the linear transformation of U, a specific element might be compatible when there exists a number which yields integers by multiplying with the real-valued parameters. The $S_{(3)}$ is thus an example of such cases.

We can also show that these two dualities, $\Gamma_0(3)$ and $S_{(3)}$, are the most general dualities that the landscape equips. As they are non-commutable, the resultant symmetry is the semi-direct product:

$$\overline{\Gamma}_0(3) \rtimes_{\phi(S_{(3)})} \mathbb{Z}_2, \tag{2.179}$$

with ϕ being a group conjugation of $\overline{\Gamma}_0(3)$ by $S_{(3)}$. We emphasize again that the nontrivial symmetry originated from the $\mathbb{Z}_6$ twist.

The flux transformations under the $S_{(3)}$ is summarized as

- $S_{(3)} = \begin{pmatrix} 0 & -\frac{1}{\sqrt{3}} \\ \sqrt{3} & 0 \end{pmatrix} \in \Gamma_0(3)_U$

[25]Thus, the center of the isometric circle is at origin, and we call $S_{(3)}$ the scaling group. Note that, $S_{(3)}$ is classified as an elliptic transformation, as $0 = |a+d| < 2$ for $a, d \in \mathbb{R}$ ($a+d=0$ leads $\mathbb{Z}_2$). The finite fixed points under an elliptic transformation are called elliptic points, and $U = \pm\frac{1}{\sqrt{3}}i$ for the case. See for details, [153]

$$
\begin{aligned}
a^0 &\to -\frac{3}{2}b_1, & b_0 &\to a^1 + \frac{1}{3}a^0, \\
a^1 &\to b_0 + \frac{1}{2}b_1, & b_1 &\to -\frac{2}{3}a_0, \\
c^0 &\to -\frac{3}{2}d_1, & d_0 &\to c^1 + \frac{1}{3}c^0, \\
c^1 &\to d_0 + \frac{1}{2}d_1, & d_1 &\to -\frac{2}{3}c^0.
\end{aligned}
\tag{2.180}
$$

Other dualities and symmetries

So far, we have shown the symmetries by which we mod out the landscape. On the other hand, there are other symmetries and dualities on the landscape, by which we do not mod out the landscape further. Those dualities and symmetries are summarized in the following list:

- Orientifold action
  Since the overall sign flipping of the fluxes can be regarded as the orientation reversion of the integrated cycles, the orientifold action $\hat{\Omega}\mathcal{R}(-1)^{F_L}$ on the orbifold geometry will be absorbed by the flipping. In fact, the total action sends $\Omega \to -\Omega$, $F_3 \to -F_3$, and $H_3 \to H_3$. Thus, the superpotential W transforms as $W \to -W$ while $\mathbb{G}_3 \to -\mathbb{G}_3$, and Ω is invariant. Hence the landscape which we consider is already modded out by the flipping.

- "pseudo-S transformation" on U
  As explained, the sufficiently large fluxes cancel the denominators of the fractional coefficients in Eq. (2.168). Then, the flux quantization condition is also satisfied after the transformation, i.e., we can define a "pseudo"-S transformation on U: $U \to -\frac{1}{U}$. However, as explained, this transformation corresponds to a $Sp(4, \mathbb{Q})$ transformation of the period vector as Eq. (2.171). Since the underlying cohomology basis must be symplectic with integer coefficients, we do not consider this transformation as a symmetry.

- Rescaling of the fluxes
  Since the F-term conditions have a trivial d.o.f. of an overall factor, we can freely change the flux quanta as

  $$(F_3, H_3) \to k(F_3, H_3). \tag{2.181}$$

  When the rescaling is originated from the fluxes, k must be an integer and N_{flux} transforms as

  $$N_{\text{flux}} \to k^2 N_{\text{flux}}, \tag{2.182}$$

  reflecting its quadratic form. This is indeed a useful duality because the solutions with a small N_{flux} are embedded in the solutions with a larger N'_{flux}, if the two charges differ by a factor of k^2. Of course, it is not a symmetry of the theory which is also characterized by the incompatibility with $Sp(4, \mathbb{Z}_2)$. The duality is used in a possible option of the way to construct all flux vacua on the $Z_{6-\mathrm{II}}$ orbifold.

- Different transformations for each of F_3 and H_3
  In the transformation of $\mathbb{G}_3$: $\mathbb{G}_3 \to M^T \mathbb{G}_3$, the assumption that the (period of) two fluxes F_3, H_3 are transformed with the same M. However, in principle, we can consider the case that they transform in an independent way, but their linear combination transforms as $\mathbb{G}_3 \to M^T \mathbb{G}_3$ in some specific cases. On the other hand, the symplectic transformation is a change of basis and F_3, H_3 are invariant. Thus, we have a uniform transformation of the integrals $\int X_3 \wedge (\alpha_I, \beta^J)$ for arbitrary X. Hence, this is not a symmetry.

- Any other transformations which cannot be interpreted as symplectic matrices
  If there are some additional symmetries in the period vector, we might obtain an incidental symmetry in the theory. Since the symplectic transformation is the most fundamental symmetry, we do not further mod out the landscape.

- $GL(2,\mathbb{Z})_U$ and its congruence subgroups
  As a result of the stabilization, the distribution of vacua may become symmetrical under $U \to -\bar{U}$ on the U-plane. This requires additional transformations of the fluxes and the axio-dilaton, but let us focus on the transformation of U. If there is modular symmetry, the corresponding group is given by $GL(2,\mathbb{Z}) \simeq SL(2,\mathbb{Z}) \rtimes \mathbb{Z}_2$, and the fundamental region is halved due to the additional transformation (if this is a symmetry). For the current congruence subgroup, we can construct its generalized group in the same way. The additional $\mathbb{Z}_2$ is generated by the matrix:

$$R = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{2.183}$$

  Then, one can check that under R the period vector is transformed by M which satisfies

$$M^T \Sigma M = -\Sigma = \Sigma^T , \tag{2.184}$$

  while the $\mathbb{Z}_2$ is not a normal subgroup of the symplectic group: $\mathbb{Z}_2 \ntrianglelefteq Sp(b_3 = 2h^{2,1}+2, \mathbb{Z})$. Thus, M is the element of the group

$$GSp(2h^{2,1}+2,\mathbb{Z}) = Sp(2h^{2,1}+2,\mathbb{Z}) \rtimes \mathbb{Z}_2 , \tag{2.185}$$

  as obtained in [58,59]. Indeed, the GSp group is linked to the generalized CP-transformation (see for details, [96, 111, 154, 155]).

  Since this is not the symmetry of the moduli space, we do not mod out the landscape by the duality, and we do not discuss the CP in the following of this thesis.

2.4.1.4 Moduli stabilization

Since the superpotential is linear in U, the F-term equations are easily solved. As in the previous sections, we additionally impose the SUSY condition on the Kähler moduli T_A as $D_{T_A} W = 0$. Due to the no-scale structure, we have the following equations:

$$\begin{cases} D_U W = \partial_U W + K_U W = 0 \\ D_S W = \partial_S W + K_S W = 0 \\ \quad W = 0 \end{cases} \Rightarrow \begin{cases} C + DS = 0 \\ B + DU = 0 \\ A + CU = 0 \end{cases} , \tag{2.186}$$

and thus we obtain an overdetermined system:

$$\langle S\rangle = -\frac{C}{D}, \quad \langle U\rangle = -\frac{B}{D}, \quad AD - BC = 0. \tag{2.187}$$

Here, $B, C \to -B, -C$ does not change N_{flux} and the quantization condition. Thus, in the landscape, we will have the solution with $-\overline{S}, -\overline{U}$, but we do not consider this is a symmetry by which we should mod out the theory.

Furthermore, at SUSY minima we have a relation of N_{flux} to ImS, ImU in general IIB compactifications (see [72]), and indeed

$$N_{\text{flux}} = 4\text{Im}S\text{Im}U|D|^2. \tag{2.188}$$

Note that the correlation between ImS and ImU that is observed in Section 2.2 and in [122] might be originated from this relation.

2.4.1.5 Identifications under the dualities

As in Section 2.2, we identify the sets of (flux, VEVs) via the symmetries we discussed in Section 2.4.1.3.

Schematically, the definition of the physically-equivalent solutions is,

$$\begin{aligned} {}^\exists\mathcal{A} \in PSL(2,\mathbb{Z})_S \times (\overline{\Gamma}_0(3)_U \rtimes_{\varphi(S_{(3)}} \mathbb{Z}_2) \text{ s.t. } (\text{fluxes}', \text{VEVs}') = (\mathcal{A}(\text{fluxes}), \mathcal{A}(\text{VEVs})) \\ \Leftrightarrow (\text{fluxes}, \text{VEVs}) \sim (\text{fluxes}', \text{VEVs}'), \end{aligned} \tag{2.189}$$

where $\mathcal{A}$(fluxes) denotes the group action on the flux space and the overall sign-flipping is also modded out since we consider the PSL group.

N_{flux} is the invariant constant under the symmetries, as it represents the charge contribution of the fluxes. Then, we obtain the whole physically-distinct landscape for given N_{flux}. We also summarize the technical details in Section 2.A.1. Indeed, the number of the whole vacua turns out finite also in this case, and thus we can define the probabilities on each VEV as we performed in Section 2.2.

2.4.2 Distributions of the moduli VEVs and degeneracies in the finite landscape on the $T^6/Z_{6-\mathrm{II}}$ orientifold

As in the previous analysis in Section 2.2, we simply ignore the tadpole cancellation condition and focus on the analysis of the solutions of the F-term equations. All of our analysis is analogous to that in Section 2.2, but due to the nontrivial symmetry in the complex-structure modulus space, the fundamental region has changed. Sine the modular symmetry is partially broken, the fundamental region indeed expanded to include Im$U \geq \frac{1}{2\sqrt{3}}$. Therefore the nontrivial distributions are observed, and the most favored point is not one of the elliptic points of $SL(2,\mathbb{Z})$ on this landscape.

2.4.2.1 Distributions of solutions projected on the U-plane

In the actual procedure of obtaining the dataset, we fix the flux bound as $N_{\rm flux}^{\rm max} \leq 24 \times 100$, and we start the next section by focusing on the $N_{\rm flux}^{\rm max} \leq 24 \times 20$ case. We find 229 solutions, and the VEVs are also found in the enlarged region as shown in Figure 2.9. As for the points which are favored by the modular flavor symmetric models, we have no solution on $U = i$ for all $N_{\rm flux}$, whereas $U = \omega$ is found. We also there are some solutions with Im$U \gg 1$ where $q = e^{2\pi i U}$ becomes sufficiently small.

Next, we also clarified the degeneracy of the solutions on each point. As in Section 2.2, the solutions are projected onto the U-plane, and thus different S and flux quanta are associated with the same point. We illustrate the degeneracy in Figure 2.9 and in Table 2.18. As one can see, the most favored point is $U = \frac{1}{\sqrt{3}}i$, which is not the elliptic point of $SL(2,\mathbb{Z})$. Instead, it is the elliptic point of the outer automorphism group, $S_{(3)}$. This result is quite different from the previous $T^6/(\mathbb{Z}_2 \times \mathbb{Z}_2')$ case in that the most preferred points appeared in the extended region and were not the elliptic points of the modular group. For the elliptic point of $\overline{\Gamma}_0(3)$, we find only $U = \omega'$ with the second highest probability. On the other hand, the ω, which is the elliptic point of $SL(2,\mathbb{Z})$ has the third highest probability, whereas there are three other points that have the same probability. Note that, these four points with the probability of 3.49% are placed on the same circle around its center at $U = \frac{2}{\sqrt{3}}i$. We can see from Table 2.18 that Re$U = 0$ is statistically favored. This suggests the well-known "void" structure (see e.g., [42]) does exist again in this case.

We also varied the maximum flux contribution $N_{\rm flux}^{\rm max}$ where as the small $N_{\rm flux}^{\rm max}$ is reliable due to the tadpole cancellation condition. We summarize the number of solutions for various $N_{\rm flux}^{\rm max}$ in Table 2.19 and in Figure 2.10. We again clearly observe $\# \sim (N_{\rm flux}^{\rm max}/24)^2$ where this quadratic behavior is predicted by the continuous approximation method. For the small $N_{\rm flux}^{\rm max}$, only a few solutions are obtained as in Section 2.2. Thus, the landscape becomes more predictive in such cases. Note that the solutions only for $N_{\rm flux} \geq 48$ whereas the quantization condition makes $N_{\rm flux} \geq 24\mathbb{Z}$. In [144], there are also listed the solutions with $N_{\rm flux} \in 24\mathbb{Z}$, but they are AdS solutions that we have not considered here.

We also summarize the most preferred VEVs and their probabilities for each bound $N_{\rm flux}^{\rm max}$ on Table 2.20. In particular, for the smallest bound, there are only two solutions and one is the elliptic point of the scaling duality $S_{(3)}$, while the other is the elliptic point of $\overline{\Gamma_0}(3)$. The usual playgrounds $U = \omega$ and $U = i$ did not appear in that landscape. Instead, the only one point which is in the fundamental region of $SL(2,\mathbb{Z})$ is $U = \frac{2}{\sqrt{3}}i$. Thus, we can conclude that the elliptic points are favored in the landscape of SUSY Minkowski solutions as usual, but the favored points turn out different from those of the modular $SL(2,\mathbb{Z})$ group.

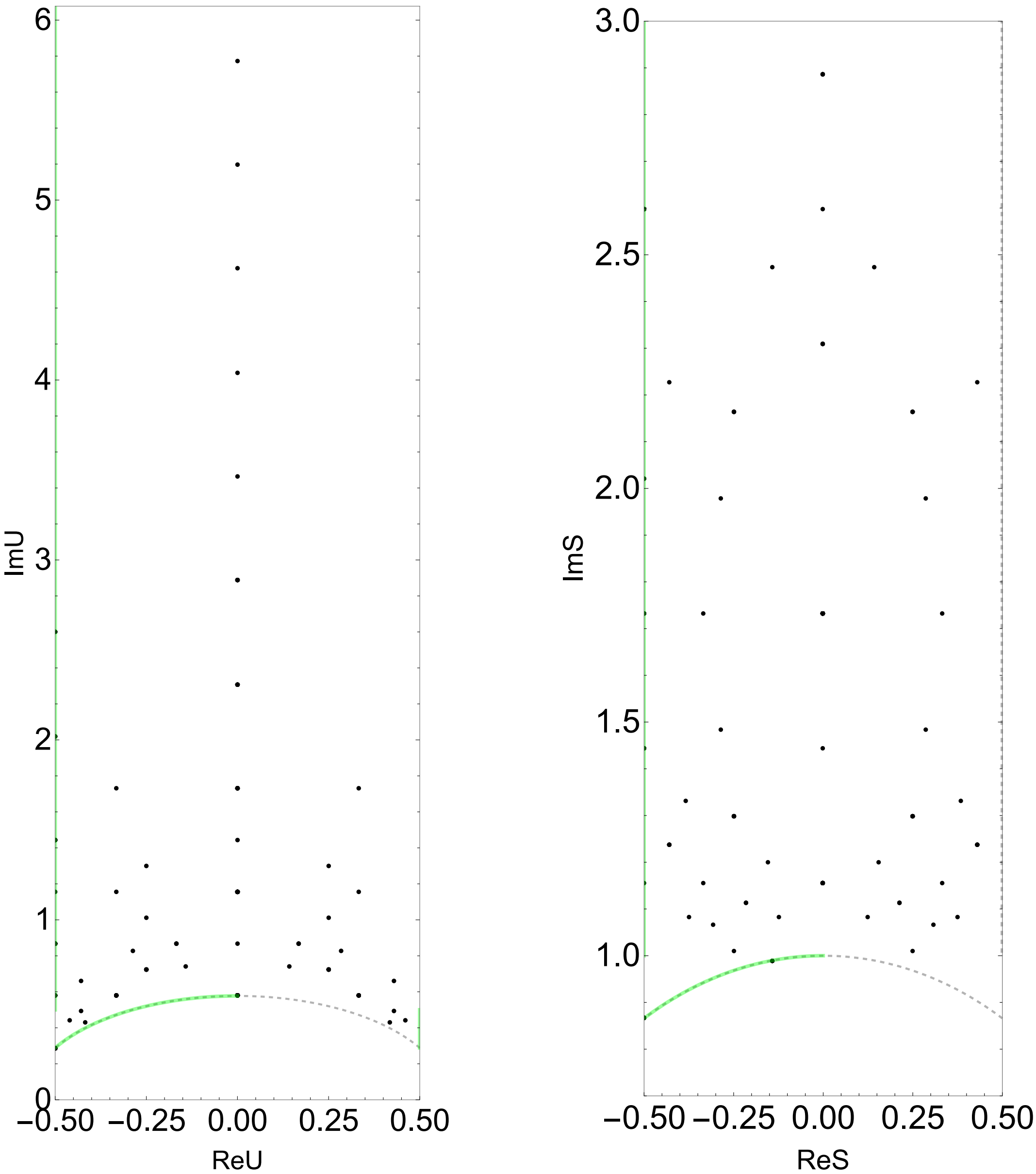

Figure 2.8: The distribution of $\langle U \rangle$ (left panel) and that of $\langle S \rangle$ (right panel) for the $N_{\rm flux}^{\rm max} = 24 \times 20$ case. Here, the boundary of the fundamental region of $\overline{\Gamma}_0(3) \rtimes_{\varphi(S_{(3)})} \mathbb{Z}_2$ is highlighted in green. The dashed curves denote the boundaries which we do not include. [70]

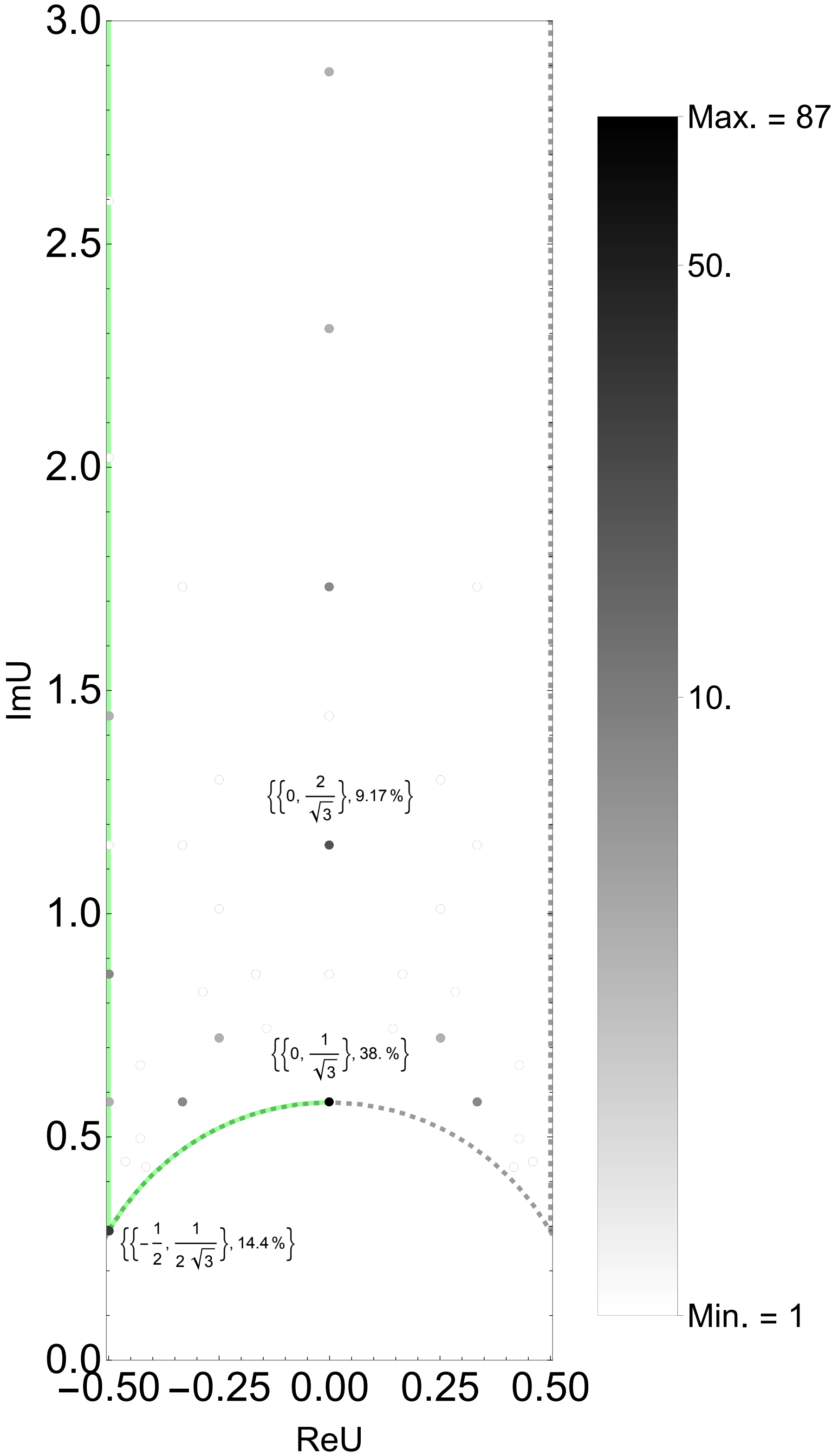

Figure 2.9: The distribution of physically-distinct solutions on $\langle U \rangle$-plane, with the degeneracy as labeled for the $N_{\text{flux}}^{\text{max}} = 24 \times 20$ case. Note that this is an enlarged view that concentrates on the region $\text{Im}U \leq 3.0$. We added the labels for the points whose probabilities are greater than 5.0 %. As the boundary, we colored the half of the isometric circle in green, which is given by $|U| = \frac{1}{3}$ and $U = -\frac{1}{2}$ (a line). [70]

Ratio	38.0%	14.4%	9.17%	3.49%	1.75%
U	$\frac{1}{\sqrt{3}}i$	$-\frac{1}{2}+\frac{1}{2\sqrt{3}}i$	$\frac{2}{\sqrt{3}}i$	$\frac{1}{3}+\frac{1}{\sqrt{3}}i$	$-\frac{1}{2}+\frac{1}{\sqrt{3}}i$
				$\sqrt{3}i$	$-\frac{1}{4}+\frac{5}{4\sqrt{3}}i$
				$-\frac{1}{3}+\frac{1}{\sqrt{3}}i$	$\frac{5}{\sqrt{3}}i$
				$-\frac{1}{2}+\frac{\sqrt{3}}{2}i$	$\frac{4}{\sqrt{3}}i$
					$\frac{1}{4}+\frac{5}{4\sqrt{3}}i$
					$-\frac{1}{2}+\frac{5}{2\sqrt{3}}i$

Table 2.18: The probability of each VEV on U-plane, for the case of $N_{\text{flux}}^{\text{max}} = 24 \times 20$. and we have included the probabilities up to the fifth largest. [70]

$N_{\text{flux}}^{\text{max}}$	24×2	24×4	24×10	24×20	24×40	24×80	24×100
# of solutions	2	7	44	229	1053	4823	7818

Table 2.19: The number of solutions for each bound $N_{\text{flux}}^{\text{max}}$. [70]

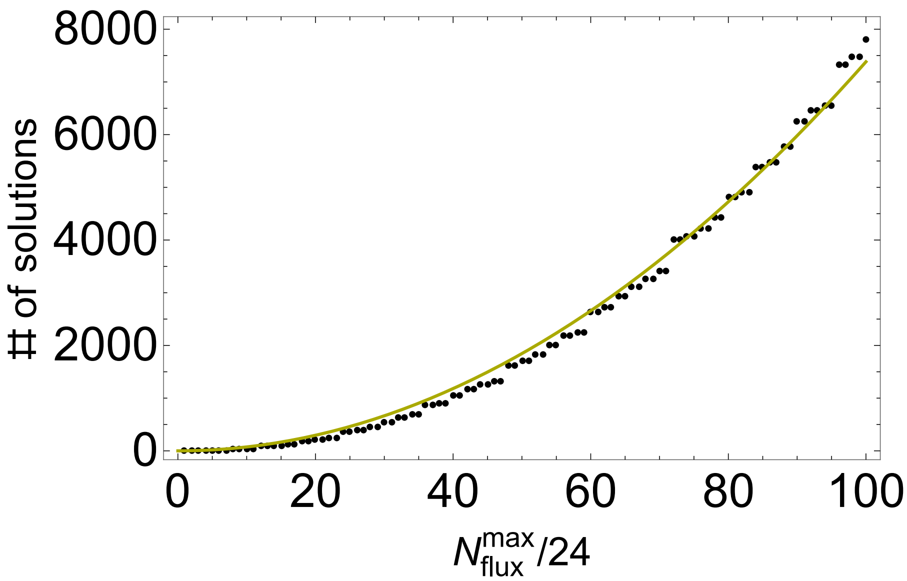

Figure 2.10: The number (#) of the whole solutions with $N_{\text{flux}} \leq N_{\text{flux}}^{\text{max}}$. The fitting curve with $\# = 0.738(N_{\text{flux}}^{\text{max}}/24)^2$. is colored in yellow. [70]

U \ $N_{\text{flux}}^{\max}$	24×2	24×4	24×10	24×20	24×40	24×80	24×100
$\frac{1}{\sqrt{3}}i$	50%	57.1%	47.7%	38.0%	32.2%	27.8%	26.6%
$-\frac{1}{2} + \frac{1}{2\sqrt{3}}i$	50%	28.6%	20.5%	14.4%	11.5%	9.62%	9.13 %
$\frac{2}{\sqrt{3}}i$		14.3%	9.09%	9.17%	8.26%	7.03%	6.68 %

Table 2.20: The probabilities of several dominant (> 5 %) VEVs. Here, $N_{\text{flux}}^{\max} = 20$. For the small bounds, we see that the elliptic point of $S_{(3)}$ is the most favored point. On the other hand, for the smallest bound we have two choices with the same probability. [70]

2.4.3 Distributions of solutions with specific S

Then, let us focus on the projection of the solutions onto the S-plane. As usual, larger $\mathrm{Im}S = g_s^{-1}$ is required for consistency. Nevertheless, we first illustrate all the solutions of the F-term equations, then we restrict ourselves to more reliable regions.

The whole solution is illustrated in Figure 2.11. There, one can see that $S = \sqrt{3}i$ is the most preferred point in the landscape. On the S-plane, $S = \omega$ has the second highest probability while it is not favored from the viewpoint of its strong coupling. In general, one finds that $\mathrm{Re}S = 0$ is favored. One can also observe the solutions with relatively large $\mathrm{Im}S$ and with probabilities as high as the solution on $S = \omega$.

On the other hand, $S = i$ does not realize because of the correlation (2.188) with $\mathrm{Im}U$. Indeed, one can check that the F-term equations have no solutions at $S = i$ or $U = i$, and this result is completely different from that in Section 2.2, where we observed that all of the elliptic points appeared on both the S and U-planes. We would like to emphasize that this phenomenon occurs even if the symmetry group of the axio-dilaton is usual $SL(2, \mathbb{Z})$.

Next, in order to analyze the weak coupling regime, we impose some bounds for $\mathrm{Im}S$ on the solutions. The result is summarized in Table 2.21 and fig. 2.12. Interestingly, we see that the dominant VEVs become more dominant as the bound increases. In particular, $U = \frac{i}{\sqrt{3}}$ is the only solution which is allowed in the sufficiently weak region $\mathrm{Im}S \geq 9$, as shown in Figure 2.12.

Although we see that there is the exclusive VEV of U with $g_s \ll 1$, there might be other acute correlations. Indeed, in Section 2.2 (in particular, see 2.9), we saw that the specific VEV of U was realized exclusively when the VEV of axio-dilatation was chosen to a specific VEV. To study this phenomenon, we list the probabilities of some dominant pairs of VEVs in Table 2.22 for the $N_{\text{flux}}^{\max} = 20$ case. As a result, we cannot observe any strong correlations among the list, in contrast to the $T^6/(\mathbb{Z}_2 \times \mathbb{Z}_2')$ case.

On the other hand, it is worthwhile to note that one might obtain $U = \frac{i}{\sqrt{3}}$ exclusively in another way. Indeed, there are 41 specific VEVs of the axio-dilaton, which gives only $U = \frac{i}{\sqrt{3}}$. They are not listed on the Table 2.22 because each of them has only a ratio of 0.44% among the landscape. However, the total is 17.9%. Therefore, $U = \frac{i}{\sqrt{3}}$ is automatically obtained if one finds a natural mechanism for obtaining such specific VEVs of S. For other Us, no value has a ratio greater than 25%, even if certain VEVs S are chosen. This result is completely different from the $T^6/(\mathbb{Z}_2 \times \mathbb{Z}_2')$ case.

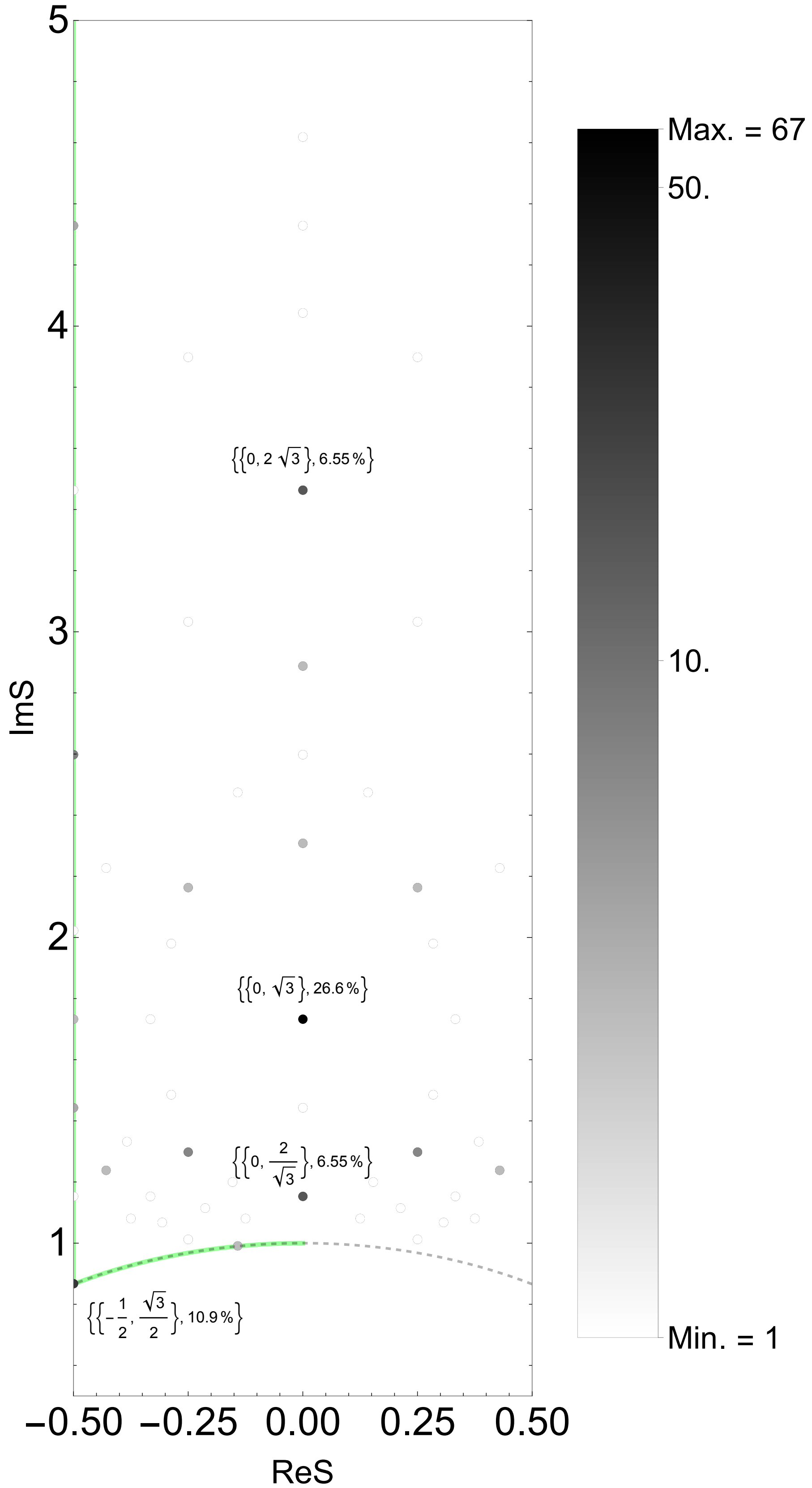

Figure 2.11: The distribution of physically-distinct solutions on $\langle S \rangle$-plane, with the degeneracy is colorized for the $N_{\text{flux}}^{\text{max}} = 24 \times 20$ case. Note that this is an enlarged view which concentrates on the region $\text{Im}S \leq 5.0$. We added the labels for the points whose probabilities are greater than 5.0 %. The boundary of the region is colorized in green. [70]

U \ $\mathrm{Im}S_{\text{bound}}$	$\frac{\sqrt{3}}{2}$ (All)	1	$\frac{2}{\sqrt{3}}$	$\sqrt{3}$	$2\sqrt{3}$	$3\sqrt{3}$
$\frac{1}{\sqrt{3}}i$	38.0%	41.3 %	39.3%	36.0%	52.2 %	47.6 %
$\frac{-1}{2}+\frac{1}{2\sqrt{3}}i$	14.4%	12.9 %	12.6%	12.7%	15.2%	23.8%
$\frac{2}{\sqrt{3}}i$	9.17%	9.45%	9.95%	9.86%	8.70%	14.3%

Table 2.21: The probabilities of several dominant (>5 %) VEVs, for the $N_{\text{flux}}^{\text{max}}=20$ case. For the small bounds, we see that the elliptic point of $S_{(3)}$ is the most favored point. [70]

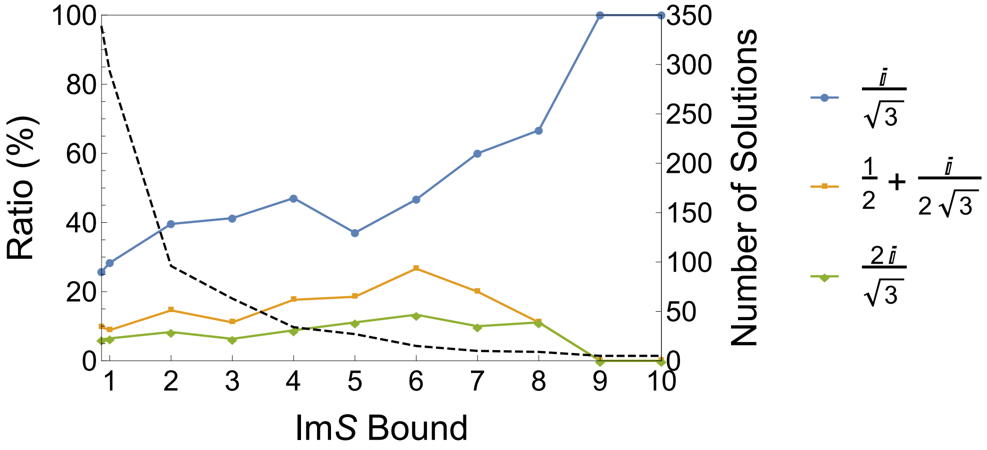

Figure 2.12: The dependency of the ratios which is associated with the three dominant (>5 %) VEVs: $U=\frac{i}{\sqrt{3}}, \frac{-1}{2}+\frac{1}{2\sqrt{3}}i$ and $\frac{2i}{\sqrt{3}}$, on the certain bounds with respect to $\mathrm{Im}S$ (solid line). The number of solutions above certain $\mathrm{Im}S_{\text{bound}}$ is represented by the dashed line. Note that the initial point of the lines is located on $\mathrm{Im}S_{\text{bound}}=\frac{\sqrt{3}}{2}$ (corresponds to all regions). [70]

U \ S_{section}	$\sqrt{3}i$	$-\frac{1}{2}+\frac{\sqrt{3}}{2}i$	$2\sqrt{3}i$	$\frac{2}{\sqrt{3}}i$
$\frac{1}{\sqrt{3}}i$	13.1%	12%	33.3%	33.3%
$-\frac{1}{2}+\frac{1}{2\sqrt{3}}i$	4.92 % (*)	24 %	6.67 % (*)	6.67 % (*)
$\frac{2}{\sqrt{3}}i$	8.20%	4.0 % (*)	6.67 % (*)	6.67 % (*)
$\omega=-\frac{1}{2}+\frac{\sqrt{3}}{2}i$	4.92 %	8.0%	0 %	0%

Table 2.22: The ratios of several dominant ($>5\%$) VEVs where the values are measured on the certain section of S, S_{section}. $N_{\text{flux}}^{\text{max}}$ is chosen as 24×20. (*) denotes that there are other VEV(s) which have the same or higher probabilities. For example, all 6.67% in the list are strictly equal. [70]

At last, we mention the void structure in the landscape. As the fundamental region enlarged, the void structure which has been observed on the other simple toroidal orientifolds [40, 122]. We illustrate the vacua pattern with $N_{\rm flux}^{\rm max} = 24 \times 100$ in Figure 2.13. One can see that the void structure is also seen, and is widened to include the enlarged regions of the fundamental region. There, the radii of the voids are related to $N_{\rm flux}$. Hence, this is phenomenologically interesting since it controls the deviations around points and small deviations are needed to explain the bottom-up modular flavor symmetric models. We will not study this point any further in this thesis.

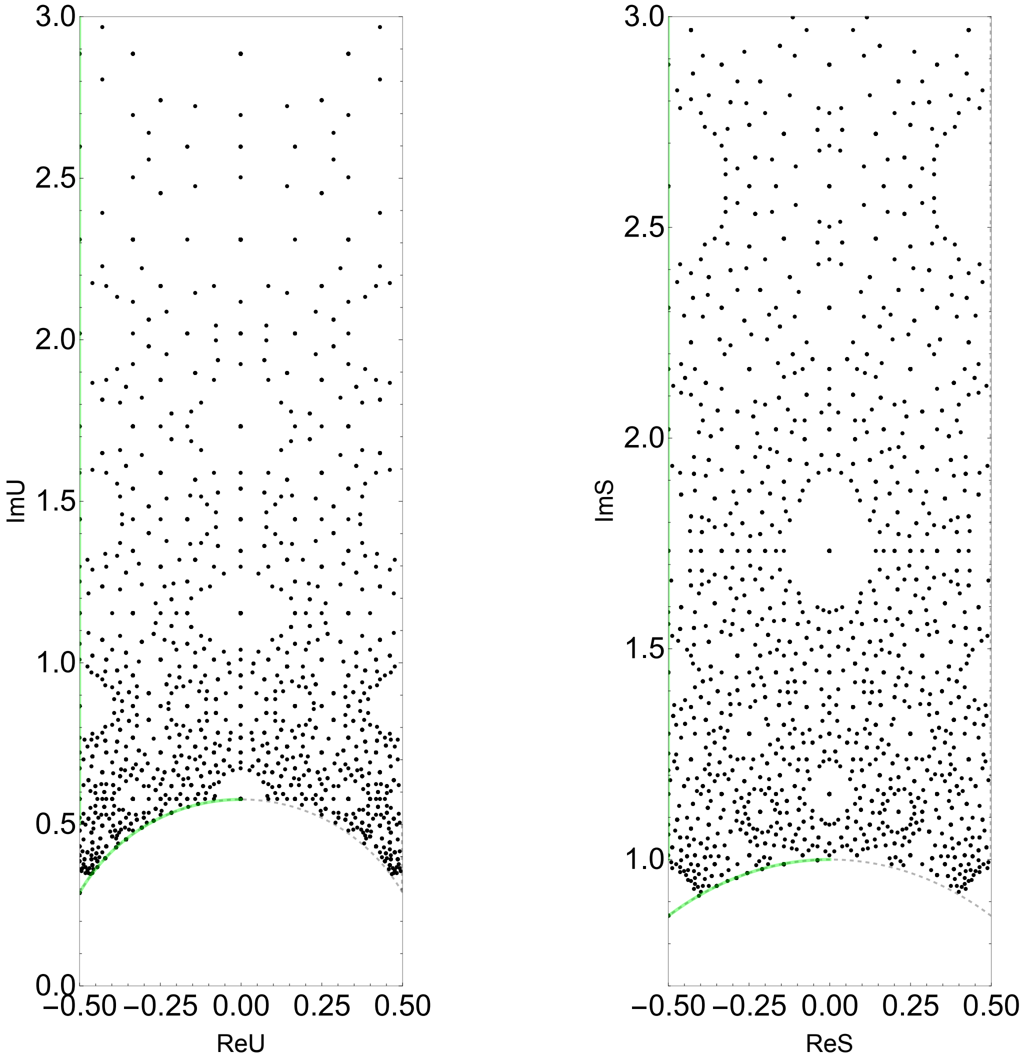

Figure 2.13: The distributions of the two moduli within Im$U \leq 3$ and Im$S \leq 3$. $N_{\rm flux}^{\rm max}$ is chosen as 24×100. One can see that the clustering occurs even in the enlarged region. [70]

2.4.4 Other orientifolds: $\mathbb{Z}_2 \times \mathbb{Z}_4$ orientifold

So far, we analyzed the $T^6/\mathbb{Z}_{6-II}$ orientifold with $\mathrm{SU}(6) \times \mathrm{SU}(2)$ lattice, and found many new features of the landscape in terms of the distributions of the moduli VEVs. Since it is straightforward to extend the discussion to another toroidal orientifold as long as it has only one complex-structure modulus ($h^{2,1}_{\text{untw.}} = 1$). In the following, we study the $T^6/(\mathbb{Z}_2 \times \mathbb{Z}_4)$ orientifold with $\mathrm{SU}(2)^2 \times \mathrm{SO}(5)^2$ as such example (it has $(h^{1,1}, h^{2,1}) = (61, 1)$). The notation is the same with Sec. 2.4.1, and we start the discussion again by constructing a proper basis of the cohomology.

2.4.4.1 Flux compactification on the $T^6/(\mathbb{Z}_2 \times \mathbb{Z}_4)$ orientifold

In this case, the Coxeter element Q which corresponds to the Lie algebra is given by

$$\begin{aligned}
Q_1(e_i) &= e_i \quad (i = 1, 2), \\
Q_1(e_i) &= e_i \quad (i = 3, 4), \\
Q_2(e_i) &= e_i + 2e_{i+1} \quad (i = 3), \\
Q_2(e_i) &= -e_{i-1} - e_i \quad (i = 4), \\
Q_2(e_i) &= -e_i - 2e_{i+1} \quad (i = 5), \\
Q_2(e_i) &= e_{i-1} + e_i \quad (i = 6),
\end{aligned} \tag{2.190}$$

where $Q = Q_1 Q_2$ as indicated by the two orbifold groups.

Then, its representation $Q : e_i \to Q_{ji} e_i$ is written by

$$Q_{(\mathrm{SU}(2))^2 \times (\mathrm{SO}(5))^2} = \begin{pmatrix} -1 & & & & & \\ & -1 & & & & \\ & & 1 & -1 & & \\ & & 2 & -1 & & \\ & & & & -1 & 1 \\ & & & & -2 & 1 \end{pmatrix}, \tag{2.191}$$

and the complex diagonal basis $z_i \to e^{2\pi i v_a^i} z_i$ $(a = 1, 2, \text{no sum})$ satisfies

$$v_1 = \frac{1}{2}(1, 0, -1), \quad v_2 = \frac{1}{4}(0, 1, -1). \tag{2.192}$$

Then the complex basis takes the explicit form of

$$\begin{aligned}
dz^1 &= dx^1 - U dx^2, \\
dz^2 &= dx^3 - \frac{1}{2}(1 - i) dx^4, \\
dz^3 &= dx^5 - \frac{1}{2}(1 + i) dx^6,
\end{aligned} \tag{2.193}$$

with $\int \Omega \wedge \overline{\Omega} = U - \overline{U}$ for $\Omega := dz^1 \wedge dz^2 \wedge dz^3$.

The procedure to find well-defined cycles is the same as in the previous case. It turns out that the four independent orbits of length two exist and the intersection matrix becomes an unimodular matrix up to a multiplied factor $\frac{1}{4}$ Then, we obtain the well-defined symplectic basis on the orbifold as

$$\begin{aligned} \mathbf{1}_1 &= 2\sum \Gamma(\beta^2) \\ \mathbf{1}_2 &= 2\sum \Gamma(\alpha_2) \\ \mathbf{1}_3 &= 2\sum \Gamma(\alpha_1) + 2\sum \Gamma(\beta^2) \\ \mathbf{1}_4 &= 2\sum \Gamma(\alpha_0) = -\Gamma(\beta^1). \end{aligned} \tag{2.194}$$

Here, $\Gamma = \mathbb{Z}_4 \times \mathbb{Z}_2$ and $\Gamma(X)$ is an orbit of X under the Γ as defined in the previous case. The basis satisfies the correct normalization:

$$\int_{T^6} \mathbf{1}_1 \wedge \mathbf{1}_4 = 8, \quad \int_{T^6} \mathbf{1}_2 \wedge \mathbf{1}_3 = 8. \tag{2.195}$$

Then, the fluxes F_3, H_3 become the elements of $H^3(T^6/(\mathbb{Z}_2 \times \mathbb{Z}_4), \mathbb{Z})$, and they are quantized on the Poincaré dual cycles.

The flux quanta measured on the orbifold are given by

$$\begin{aligned} &\int_{T^6/(\mathbb{Z}_2\times\mathbb{Z}_4)} (F_3, H_3) \wedge \mathbf{1}_1 = \frac{1}{8}\int_{T^6} (F_3, H_3) \wedge \mathbf{1}_1 = \frac{1}{4}\sum \int_{T^6} (F_3, H_3) \wedge \Gamma(\beta^2) = \frac{1}{2}(a^2, c^2), \\ &\int_{T^6/(\mathbb{Z}_2\times\mathbb{Z}_4)} (F_3, H_3) \wedge \mathbf{1}_2 = \frac{1}{2}(-b_2, -d_2), \\ &\int_{T^6/(\mathbb{Z}_2\times\mathbb{Z}_4)} (F_3, H_3) \wedge \mathbf{1}_3 = \frac{1}{2}\left\{(-b_1, -d_1) + (a^2, c^2)\right\}, \\ &\int_{T^6/(\mathbb{Z}_2\times\mathbb{Z}_4)} (F_3, H_3) \wedge \mathbf{1}_4 = \frac{1}{4}(-a^1, -c^1), \end{aligned} \tag{2.196}$$

where we chose the representative of $\mathbf{1}_4$ as $-\beta^1$. The expansion of the three-form flux G_3 is

$$G_3 = -\frac{1}{4}(a^1 - Sc^1)\mathbf{1}_1 + \frac{1}{2}[(-b_1 + a^2) - S(-d_1 + c^2)]\mathbf{1}_2 + \frac{1}{2}(b_2 - Sd_2)\mathbf{1}_3 - \frac{1}{2}(a^2 - Sc^2)\mathbf{1}_4, \tag{2.197}$$

and the flux quantization condition is written by

$$(b_1, d_1) \in 4\mathbb{Z}, (b_2, d_2) \in 4\mathbb{Z}, (a_1, c_1) \in 8\mathbb{Z}, (a_2, c_2) \in 4\mathbb{Z}, \tag{2.198}$$

where we require even fluxes on the orientifold.

The relevant quantities are as follows:

- N_{flux}

$$N_{\text{flux}} = a^2 c^1 - a^1 c^2 + 2b_2 c^2 - 2b_2 d_1 - 2a^2 d_2 + 2b_1 d_2 \in 32\mathbb{Z}. \tag{2.199}$$

- The period vector

$$\Pi \equiv \begin{pmatrix} \int \Omega \wedge \mathbf{1}_4 \\ \int \Omega \wedge \mathbf{1}_3 \\ \int \Omega \wedge \mathbf{1}_1 \\ \int \Omega \wedge \mathbf{1}_2 \end{pmatrix} = \frac{1}{8} \begin{pmatrix} 2U \\ 2i \\ -2+2i \\ (2+2i)U \end{pmatrix}, \tag{2.200}$$

where integrals are performed on the orbifold.

- Complexified flux quanta in Eq. (2.145)

$$\begin{aligned} A &= \frac{-1+i}{2}\left[a^1 + (-1+i)b_2\right], & C &= \frac{-1+i}{2}\left[(-1+i)a^2 + (-2i)b_1\right], \\ B &= \frac{-1+i}{2}(-1)\left[c^1 + (-1+i)d_2\right], & D &= \frac{-1+i}{2}(-1)\left[(-1+i)c_2 + (-2i)d_1\right]. \end{aligned} \tag{2.201}$$

The F-term equations are again linear due to $h^{2,1} = 1$. Thus, in terms of the quadratic field, it is impossible to obtain $U = \omega$ and $U = i$ in the landscape. Indeed, this leads that the solutions S, U should be rational numbers on the $\mathbb{Z}_2 \times \mathbb{Z}_4$ orientifold.

Moreover, on the $\mathbb{Z}_2 \times \mathbb{Z}_4$ orientifold the modular transformation is indeed realized:

$$\Pi \to (cU+d)^{-1} \begin{pmatrix} a & b & -b & 0 \\ -c & d & 0 & c \\ -2c & 0 & d & c \\ 0 & 2b & -b & a \end{pmatrix} \Pi \equiv (cU+d)^{-1} M \Pi, \tag{2.202}$$

where $M \in Sp(b_3, \mathbb{Z})$ is the most general symmetry of the cohomology.

The flux transformation rules for the real flux quanta are summarized as follows:
Under the $SL(2,\mathbb{Z})_S$,

- $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in SL(2,\mathbb{Z})_S$

$$\begin{aligned} a^1 &\to -c^1, & b_0 &\to -d_1, \\ a^2 &\to -c^2, & b_1 &\to -d_2, \\ c^1 &\to a^1, & d_0 &\to b_1, \\ c^2 &\to a^2, & d_1 &\to b_2. \end{aligned} \tag{2.203}$$

- $T^q = \begin{pmatrix} 1 & q \\ 0 & 1 \end{pmatrix} \in SL(2,\mathbb{Z})_S$

$$\begin{aligned} &a^1 \to a^1 + qc^1, && b_1 \to b_1 + qd_1, \\ &a^2 \to a^2 + qc^2, && b_2 \to b_2 + qd_2, \\ &c^1 \to c^1, && d_1 \to d_1, \\ &c^2 \to c^2, && d_2 \to d_2. \end{aligned} \tag{2.204}$$

Under the $SL(2,\mathbb{Z})_U$,

- $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in SL(2,\mathbb{Z})_U$

$$\begin{aligned} &a^1 \to -2b_1, && b_1 \to \frac{1}{2}a^1, \\ &a^2 \to a^1 - b_2, && b_2 \to a^2 - 2b_1, \\ &c^1 \to -2d_1, && d_1 \to \frac{1}{2}c^1, \\ &c^2 \to c^1 - d_2, && d_2 \to c^2 - 2d_1. \end{aligned} \tag{2.205}$$

- $T^q = \begin{pmatrix} 1 & q \\ 0 & 1 \end{pmatrix} \in SL(2,\mathbb{Z})_U$

$$\begin{aligned} &a^1 \to a^1 + 2qb_1, && b_1 \to b_1, \\ &a^2 \to a^2, && b_2 \to b_2 - q(a^2 - 2b_1), \\ &c^1 \to c^1 + 2qd_1, && d_1 \to d_1, \\ &c^2 \to c^2, && d_2 \to d_2 - q(c^2 - 2d_1). \end{aligned} \tag{2.206}$$

2.4.4.2 Distributions and Degeneracies in the Finite landscape

Following the lines of the $\mathbb{Z}_{6-\mathrm{II}}$ case, we can obtain the whole SUSY Minkowski flux vacua on the $\mathbb{Z}_2 \times \mathbb{Z}_4$ orientifold in an explicit way (which is analogous to the method in Appendix 2.A.1). The upper bound of N_{flux} is defined as $N_{\text{flux}} \leq 32 \times 100$. There, we again define the probability for each point in the moduli space. Having generated the dataset, we find that the smallest N_{flux} that yields a solution turns out 64, which typically exceeds the tadpole cancellation condition. Nevertheless, our first aim is to analyze the solutions of the F-term equations and we continue the discussion. Note that even if a certain F-theory uplifting on the orientifold turns out possible, the possible back-reaction makes the smaller N_{flux} a more reliable region.

We start our discussion with $N_{\text{flux}}^{\text{max}} = 32 \times 20$. Then, the number of solutions turns out 123. As in the previous case, we illustrate the projections of the solutions onto U- and S-plane in Figure 2.14, and one immediately finds that they are the same distributions.

This phenomenon originated from a duality in the superpotential, which interchanges U and S with a suitable flux transformation. Indeed, the flux transformation

$$\{c^1, c^2, d_1, d_2, a^1, a^2, b_1, b_2\} \to \{2b_1, c^2, d_1, -a^2 + 2b_1, a^1, c^1 - d_2, \frac{1}{2}c^1, b_2\} \tag{2.207}$$

is incorporated, while it does not contradict the flux quantization condition (2.198). The transformation also keeps N_{flux} invariant, thus one might think this is a symmetry of the theory. However, it turns out that there is no compatible way to induce an appropriate symplectic transformation of the cycles. Thus, as claimed in 2.4.1.3, we do not mod out the landscape via the duality. Let us focus on the projection on U-plane using this duality.

The results of the projection are summarized in Table 2.23 and fig. 2.15. The most favored point turns out $U = i$, and this is one of the elliptic points of $SL(2, \mathbb{Z})$. Taking into account $U = \omega, i\infty$ cannot be realized on the framework, we can conclude that $U = i$ is the only fixed point which can be obtained on the orientifold.

Next, we vary $N_{\text{flux}}^{\text{max}}$ as in the previous case. We illustrate the resultant number of solutions in Table 2.24 and fig. 2.16. As predicted by the continuous approximation, we again observe the quadratic behavior with regard to $N_{\text{flux}}^{\text{max}}$.

Note that $U = i$ naturally induces $S = i$, although this is inconsistent with the field-theoretic discussion. Hence, let us discuss the correlation between S and U in the landscape. In Figure 2.17, we illustrate the dependence of U on certain bounds on ImS.

As shown, although there are only $\mathcal{O}(10)$ solutions but if there are some natural mechanisms to stabilize the axio-dilaton in the weak-coupling limit $g_s \ll 1$, the VEV $U = i$ becomes exclusive among the landscape. Although we do not investigate such mechanisms, if there is such a compactification (more generally, on a different orientifold), the flux contribution $|D|^2$ in Eq. (2.188) should be suppressed in a natural way.

Ratio	39.0%	17.1%	6.50%	3.25%
U	i	$2i$	$3i$	$4i$
			$-\frac{1}{2} + \frac{3}{2}i$	$-\frac{1}{2} + i$
				$5i$
				$-\frac{1}{2} + \frac{5}{2}i$

Table 2.23: The probability of each VEV on U-plane, for the case of $N_{\text{flux}}^{\text{max}} = 32 \times 20$. and we have included the probabilities up to the fourth largest. [70]

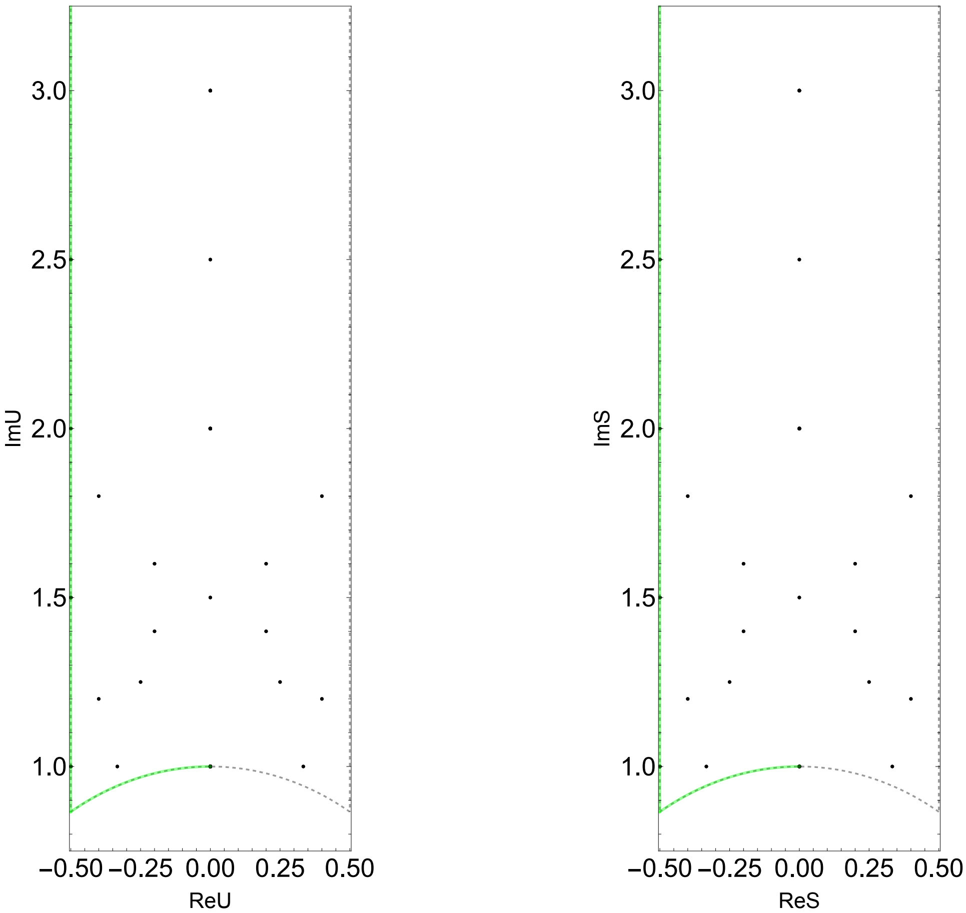

Figure 2.14: The distribution of $\langle U\rangle$ (left panel) and that of $\langle S\rangle$ (right panel) for the $N_{\text{flux}}^{\max} = 32\times 20$ case. The boundaries of the fundamental region are colored green, while the dashed curves denote the boundaries which we do not include. Apparently, these two distributions are exactly the same. Note that the illustrated ranges are $\mathrm{Im}(U,S)\geq\frac{\sqrt{3}}{2}$. [70]

$N_{\text{flux}}^{\max}$	32×2	32×4	32×10	32×20	32×40	232×80	32×100
# of solutions	1	4	23	123	589	2784	4562

Table 2.24: The number of the solutions for each bound $N_{\text{flux}}^{\max}$. [70]

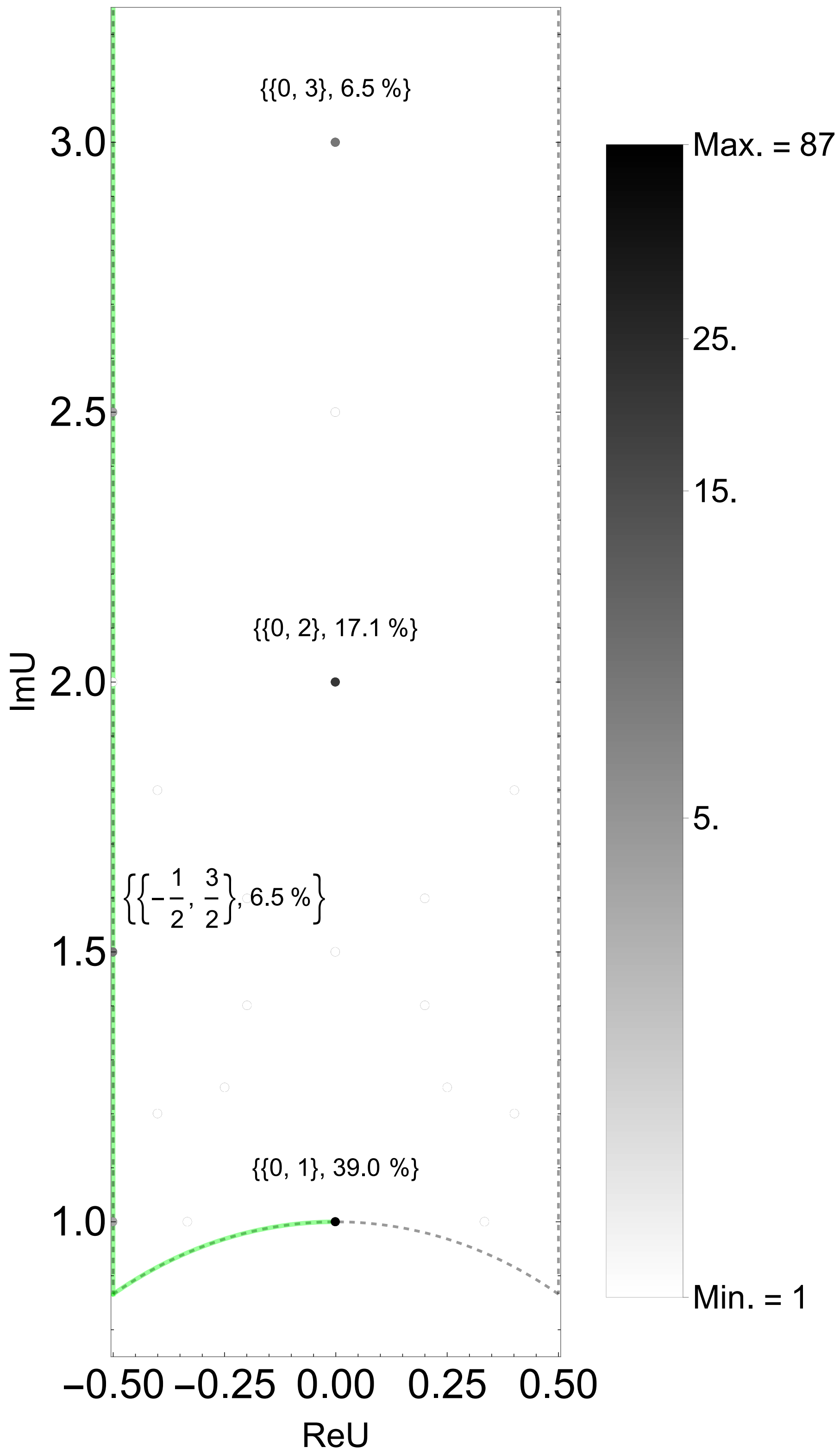

Figure 2.15: The distribution of physically-distinct solutions on $\langle U \rangle$-plane with the figure cropped to Im$U \leq 3.25$ is illustrated. Here, $N_{\text{flux}}^{\text{max}}$ is 32×20. The colors are determined by the associated number of degeneracies among the physically-distinct solutions. Here, we additionally listed the positions and ratios of the points whose ratio $> 5.0\%$. [70]

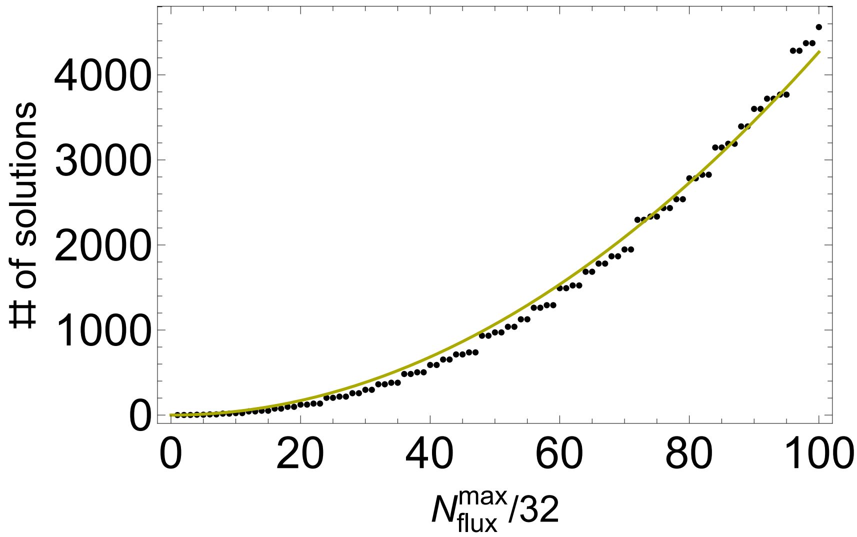

Figure 2.16: The number (#) of the whole solutions with $N_{\text{flux}} \leq N_{\text{flux}}^{\text{max}}$. The fitting curve with $\# = 0.427(N_{\text{flux}}^{\text{max}}/32)^2$ is colored in yellow. [70]

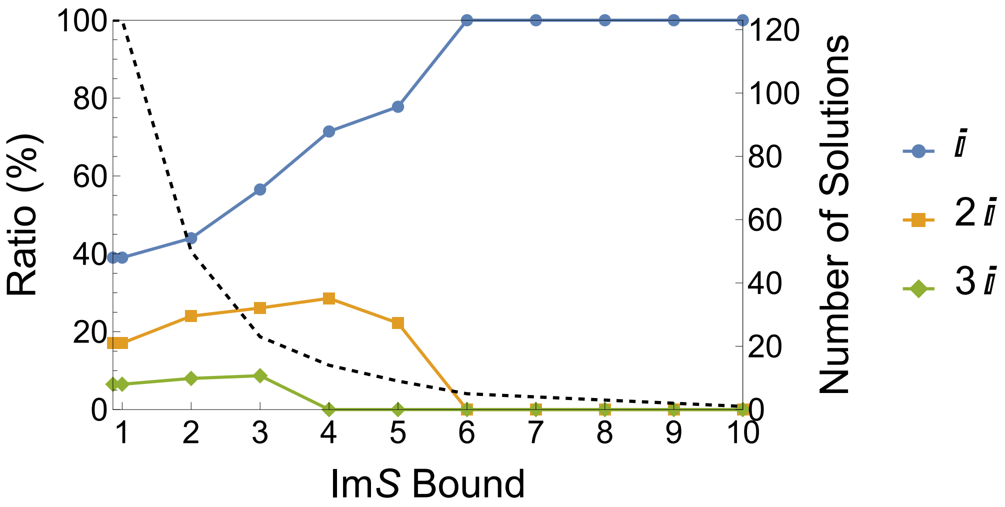

Figure 2.17: The dependency of the ratios which is associated with the three dominant (> 5 %) VEVs: $U = i, 2i$ and $3i$, on the certain bounds with respect to ImS (solid line). The number of solutions above certain ImS_{bound} is represented by the dashed line. Note that the initial points of the lines are located on Im$S_{\text{bound}} = \frac{\sqrt{3}}{2}$ (corresponds to all regions). [70]

The void structure exists also on the toroidal orientifold, as illustrated in Figure 2.18 where $N_{\rm flux}^{\rm max} = 32 \times 100$ is chosen. One can see the elliptic point $U = i$ appears as one of the centers of many circles.

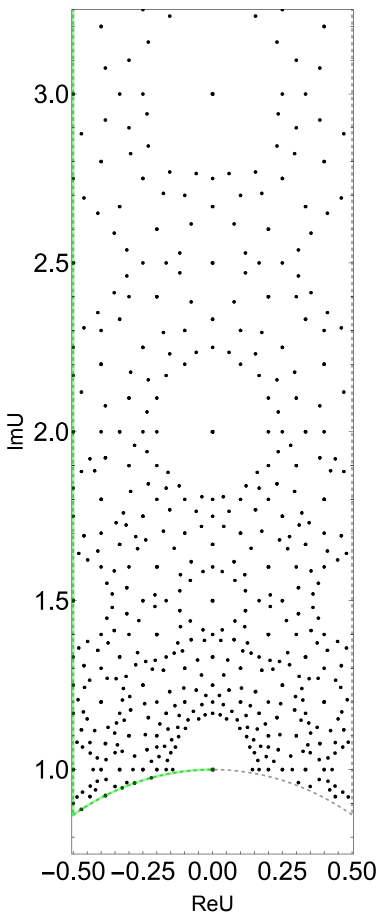

Figure 2.18: The distribution of $\langle U \rangle$ with $N_{\rm flux}^{\rm max} = 32 \times 100$. The region illustrated is $\mathrm{Im}U \leq 3.25$. One can see many circles emerge around the points which have high degeneracies. [70]

2.5 Conclusions of the Chapter 1

In this chapter, we have studied the explicit distributions of the moduli VEVs on the toroidal orientifolds. In particular, we have also clarified the degeneracy of the flux vacua, which leads a concept of the probability on the moduli space. Each discretized vacuum is possibly identified with another vacuum due to the symmetries, which mainly refer to modular symmetries. Since we study the compactifications on the modular invariant spaces, we can safely mod out the resultant landscape by those symmetries.

To be specific, we have chosen three toroidal orientifolds to analyze the landscape. The first one, namely the $T^6/\mathbb{Z}_2 \times \mathbb{Z}_2'$ orientifold, was studied in Sections 2.2 and 2.3.

In Section 2.2, we analyzed the SUSY Minkowski vacua without any corrections. There, we explicitly obtained the whole flux vacua, and we defined the probabilities (the ratios in the landscape) on the moduli space. As a result, we found that $U = \omega$ generically had a high probability, while $U = i$ is not favored when it is compared to the ratio of $U = \omega$. Since these two points, which are the elliptic points of $SL(2, \mathbb{Z})$, are favored in the bottom-up modular flavor symmetric models, it is interesting that a difference between the probabilities of those two points is predicted from the viewpoint of string theory. The other interesting result was about CP-violation. Due to the fact [96, 156, 157] that 4d CP-transformation and the orientation-reversing action on the 6d space are incorporated into the 10d proper Lorentz symmetry, we interpreted the 4d CP symmetry in terms of the VEVs of the real parts of the complex-structure moduli and the axio-dilaton. Then, in the vacua, the characteristic correlation between those two fields was observed. Indeed, it yields an interesting implication that different CP phases may have a common origin, while such a scenario is studied in [128]. Finally, we compare the landscape and a concrete A_4 model for illustrative purposes, without any concrete embedding.

So far, we have focused on the no-scale scalar potential. However, there are in general several possible corrections, where we have considered the KKLT-type non-perturbative effects in Section 2.3. Although the original KKLT scenario simply assumes that the complex-structure moduli and the axio-dilaton are effectively integrated out when the Kähler modulus is stabilized, it generally leads to small deviations from the original VEVs. We further considered the uplifting term which is generated by an anti-D3 brane, and numerically evaluated the typical order of the deviations around their VEVs. As a result, the absolute values of the deviations had their peak around $\mathcal{O}(10^{-5})$, while they have no radial dependence. It is interesting that the deviations do not depend on the point where the moduli are originally stabilized. Although we omitted the whole section, comparing this result with a concrete A_4 model, the initial VEV $U = \omega$ turns out not compatible with the A_4 model that is fitted to observational values.

Having intensively studied the landscape of the simplest toroidal orientifold, it is a natural direction to consider more nontrivial toroidal orientifolds. For example, the conclusions in the above studies claim that $U = \omega$ is statistically favored. However, the details of the F-term equation originated from the period vector, i.e., geometry. Therefore, we had to extend our study and discuss the robustness of our previous studies.

Then, we investigated in Section 2.4 the flux landscape of the $T^6/\mathbb{Z}_{6-II}$ orientifold and the $T^6/\mathbb{Z}_2 \times \mathbb{Z}_4$ orientifold with certain lattices. We explicitly constructed the invariant cycles and found a period vector, although the results have been already known, e.g., [142, 143].

Nevertheless, we carefully classified the symmetries and the dualities, and found the whole symmetry in the complex-structure modulus space is a certain semi-direct product group of a congruence subgroup. Then, under the whole group, the fundamental region was defined and the finite flux vacua was obtained. On the orientifold with $h^{2,1} = 1$, we cannot obtain $U = \omega$ and $U = i$ simultaneously due to the linearity of the F-term equations. Thus, it is naturally expected that the resultant landscape is different from that of $T^6/\mathbb{Z}_2 \times \mathbb{Z}_2'$ orientifold. Indeed this is the case, we found the most favored point is at an elliptic point of the outer automorphism group, which can be interpreted as an enhanced symmetry. Furthermore, we have the large degeneracy at an elliptic point, but not of the modular group $SL(2, \mathbb{Z})$, due to the fact that the outer automorphism group is not an element of $SL(2, \mathbb{Z})$.

To compare the landscape further, clarifying the landscape of $T^6/\mathbb{Z}_2 \times \mathbb{Z}_4$ orientifold is also important since it has a nontrivial combination of the orbifold groups while it has a factorizable structure. Then, we found that the exact modular group $SL(2, \mathbb{Z})$ is realized there, but $U = i$ is the only allowed VEV among the elliptic points. Therefore, we conclude that the situation highly depends on the toroidal orientifold.

There are two directions which we are interested in:

1. Systematic classification of symmetries/dualities in the landscape
   As we discussed, to find the symmetries by which we mod out the landscape it is sufficient to study the period vector. Since the period vector is determined through the Coxeter element (or, equivalently, the root lattice), we may systematically classify all possible symmetries or dualities in the landscape. All of the examples exhibit the high degeneracy on the fixed point, or more specifically, on the elliptic point. Thus we expect that we can roughly grasp the landscape by knowing its symmetry. It is of our interest and we will report the result in the upcoming paper.

2. Void structure and its implications on the modular flavor symmetric models We mentioned the appearance of the void structure in several examples. From the theoretical point of view, the mathematics under the structure is surely interesting. On the other hand, the fact that the voids form around a point with high degeneracy forbids direct stabilization at a point close to the fixed point (since the fixed point has typically high degeneracy). Thus, in the context of moduli stabilization, we have to include some corrections to explain the deviations. The quantity that controls the radii of the voids is likely to be the tadpole charge (see for the numerical relation between the voids and the tadpole charge, e.g., [122]). The tadpole charge is also related to D-branes, i.e., it governs both moduli stabilization and D-brane configurations. To discuss this point further, considering some D-brane models such as the magnetized D-branes is quite an interesting direction.

We consider our work as a first step to study further the string landscape from the phenomenological point of view.

Appendix

2.A Appendices For This Chapter

2.A.1 Finite physically-distinct vacua on $T^6/\mathbb{Z}_{6-\mathrm{II}}$

Algorithm 1
The scaling duality enables us to adopt the method that is essentially equivalent to the method of [40, 42, 111]. Henceforth, we discuss a concrete method that generates all physically-distinct solutions of the F-term equations (2.186) on the $\mathbb{Z}_{6-\mathrm{II}}$ orientifold.

As the continuous approximation method implies, let us define the maximum flux contribution to the tadpole cancellation condition as

$$N_{\text{flux}} \leq N_{\text{flux}}^{\text{max}}. \tag{2.208}$$

Here and in what follows, the term physically-distinct is defined via the identification (2.189). Then, as a result, it turns out that the landscape is finite.

To fix the identification, we choose some representatives of the flux quanta appropriately. The T-transformation of the fluxes is summarized in 2.A.1 for convenience. Then, $T \in \Gamma_0(3)$ is fixed by choosing

$$\frac{1}{2}d_0 = 0, 1, \ldots, \frac{1}{2}|c^0| - 1, \tag{2.209}$$
$$\frac{1}{4}d_1 = 0, 1, \ldots, \frac{1}{4}|2c^1| - 1 \quad (c^0 = 0), \tag{2.210}$$

where the second line is for the $c^0 = 0$ case. Since there is a symmetry under the overall sign-flipping of the fluxes, we can choose $c^0 = 0$ or $c^0 = 0, c^1 > 0$. Note that $c^0 = c^1 = 0$ is forbidden since it leads $D = 0$.

Analogously, we can fix T-transformation of $SL(2,\mathbb{Z})_S$ by choosing

$$\frac{1}{6}a^0 = 0, 1, \ldots, \frac{1}{6}|c^0| - 1, \tag{2.211}$$
$$\frac{1}{2}a^1 = 0, 1, \ldots, \frac{1}{2}|c^1| - 1 \quad (c^0 = 0). \tag{2.212}$$

As mentioned, strictly speaking, we do not mod out the landscape by the S-transformation of $SL(2,\mathbb{Z})_S$. Instead, we simply ignore the inside of (all translations of) unit-circle and consider only T-connected regions and mod out by T.

fluxes	$T^q \in SL(2,\mathbb{Z})_S$	$T^q \in \Gamma_0(3)_U$	$m\mathbb{Z}$
a^0	$a^0 + qc^0$	a^0	6
a^1	$a^1 + qc^1$	a^1	2
b_0	$b_0 + qd_0$	$b_0 - qa^0$	2
b_1	$b_1 + qd_1$	$b_1 - q\left(\frac{2}{3}a^0 + 2a^1\right)$	4
c^0	c^0	c^0	6
c^1	c^1	c^1	2
d_0	d_0	$d_0 - qc^0$	2
d_1	d_1	$d_1 - q\left(\frac{2}{3}c^0 + 2c^1\right)$	4

Table 2.A.1: The flux transformation rules under the T-transformation $m\mathbb{Z}$ denotes the flux quantization. [70]

The fundamental region $\mathcal{F}_{SL(2,\mathbb{Z})_S}$ is chosen as

$$\mathcal{F}_{SL(2,\mathbb{Z})_S} := \left\{ S \middle| -\frac{1}{2} \leq \mathrm{Re}S \leq 0, |S|^2 \geq 1 \right\} \bigcup \left\{ S \middle| 0 < \mathrm{Re}S < \frac{1}{2}, |S|^2 > 1 \right\}. \tag{2.213}$$

On the other hand, the fundamental region of $\Gamma_0(3)$ also exists and is constructed as follows. Since S-transformation is broken and the generator S' is given by $S' := S^{-1}T^{-3}S^{-1}T^{-1}$, we have four independent actions which map the $SL(2,\mathbb{Z})$ fundamental region into another independent region, and the fundamental region of $\Gamma_0(3)$ is thus given by

$$\mathcal{F}_{\Gamma_0(3)_U} = \mathcal{F}_{SL(2,\mathbb{Z})_U} \bigcup S\mathcal{F}_{SL(2,\mathbb{Z})_U} \backslash \{-\bar{\omega}\} \bigcup (ST)\mathcal{F}_{SL(2,\mathbb{Z})_U} \bigcup (ST^{-1})\mathcal{F}_{SL(2,\mathbb{Z})_U}. \tag{2.214}$$

Here, $U = -\bar{\omega}$ is excluded, to clarify that $T \in \Gamma_0(3)$ and $S\omega = T\omega = -\overline{\omega}$. The number of independent regions is in general described in terms of the index of the group in $PSL(2,\mathbb{Z})$, which reads

$$[PSL(2,\mathbb{Z}) : \Gamma_0(n)] = n \prod_{p|n} \left(1 + \frac{1}{p}\right). \tag{2.215}$$

(See for details, [158].)

The fundamental region of $\Gamma_0(3)$ (strictly speaking, the projected one $\overline{\Gamma}_0(3)$) is illustrated in Figure 2.A.1. There are two cusps $0, i\infty$, and the region inevitably includes $U \simeq 0$ regime.

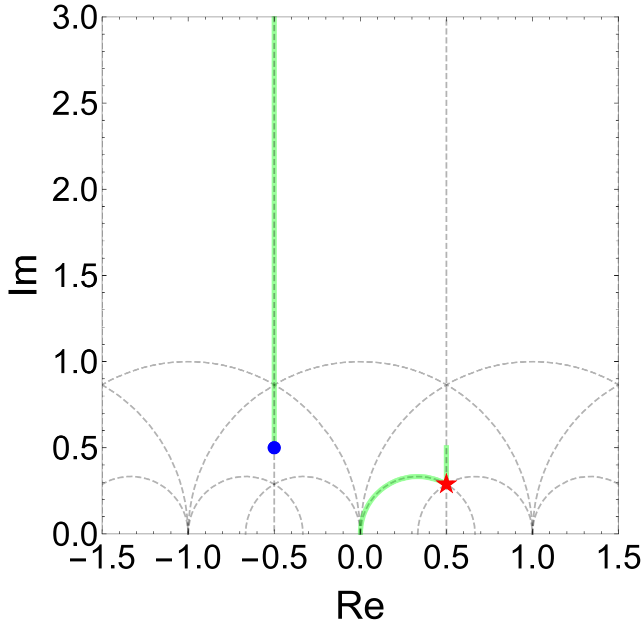

Figure 2.A.1: The fundamental region of $\overline{\Gamma}_0(3)$. We colored the boundaries in green. The blue dot denotes $-\frac{1}{2}+i\frac{1}{2}$, where (naturally) the boundary is split. The point is an image of i under S. On the other hand, the red star denotes $\frac{1}{2}+\frac{i}{2\sqrt{3}}$. This is the fixed point under S', where $\mathbb{Z}_3$ enhanced symmetry occurs. The bottom right green arc is $|U-\frac{1}{3}|=\frac{1}{3}$ within the region $0<\mathrm{Re}U\leq\frac{1}{2}$. [70]

The whole group is not the $\Gamma_0(3)$ but its outer semidirect product group with the outer automorphism, which we call the scaling duality: $U\to-\frac{1}{3U}$.

Then, suppose the outer semi-direct product group and mode out the landscape by the scaling duality. Indeed, the fixed cycles under the duality are written by two[26] one-parameter families:

$$(\mathrm{Re}U-\alpha)^2+\mathrm{Im}U^2=\frac{1}{3}+\alpha^2, \tag{2.216}$$

$$\mathrm{Re}U^2+(\mathrm{Im}U-\alpha)^2=-\frac{1}{3}+\alpha^2. \tag{2.217}$$

Here, $\alpha\in\mathbb{R}$ takes the all possible values. Among those cycles, the cycle $|\sqrt{3}U|=1$ which is obtained by defining $\alpha=0$, is called the isometric cycle. Here, the term isometric refers to the fact that the length of a line element is transformed by multiplying $|cU+d|^2$ under a general linear transformation: $U\to\frac{aU+b}{cU+d}$. Thus, this cycle behaves like a mirror and the resultant fundamental region is illustrated in Figure 2.A.2. Note that $\mathrm{Im}U\simeq 0$ region is excluded by the additional identification.

[26]For general elliptic transformations, there is only one such family. The first one emerges in this case since $a+d=0$.

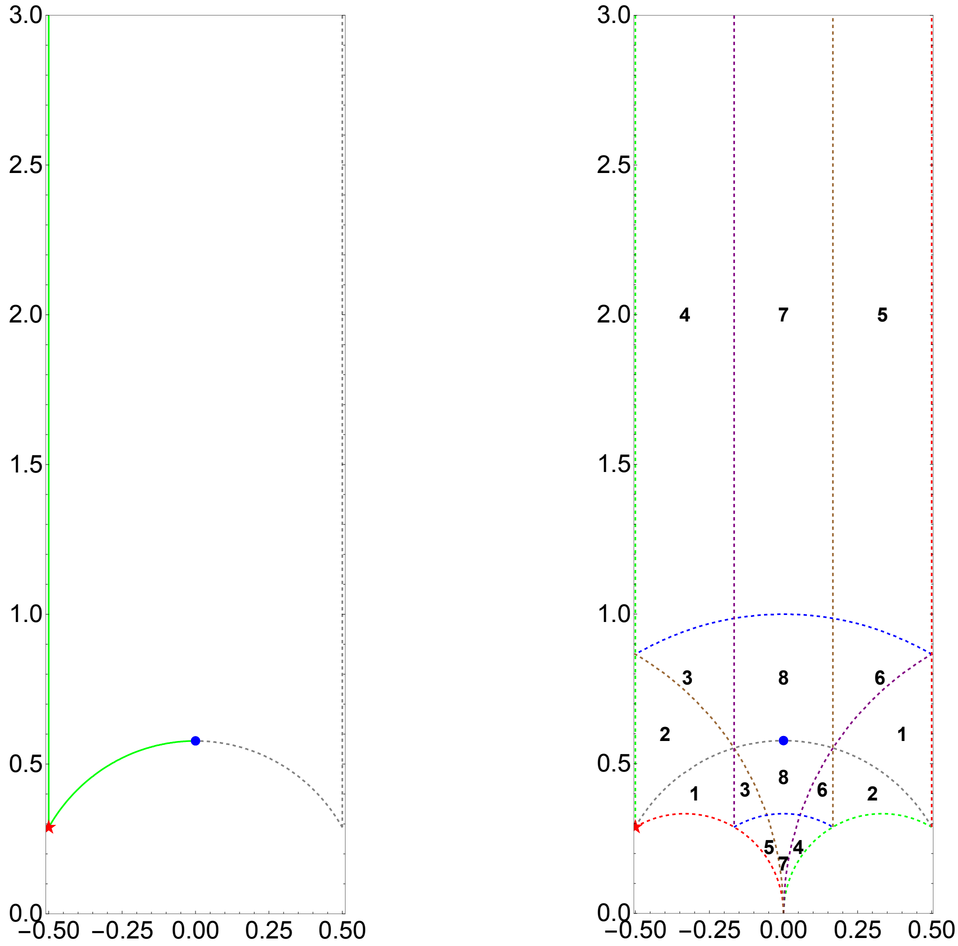

Figure 2.A.2: The boundary of the fundamental region of the total group, $\overline{\Gamma}_0(3) \rtimes \mathbb{Z}_2$ (left panel). Here, the blue dot denotes $\frac{i}{\sqrt{3}}$ that is the fixed point of the scaling duality: $U \to -\frac{1}{3U}$. On the other hand, the red star denotes $-\frac{1}{2} + i\frac{\sqrt{3}}{6}$, which is the fixed point of S'. Respectively, $\mathbb{Z}_2$ and $\mathbb{Z}_3$ enhanced symmetries are realized on those points. In the right panel, we show the correspondences between the detailed small regions, under the scaling duality. The integers label the small regions. Two regions with the same label are identified. Furthermore, the identified edges are shown in the same color. [70]

By solving the F-term equations, we obtain

$$\mathrm{Im}S = -\frac{\mathrm{Im}(C\overline{D})}{|D|^2}, \quad \mathrm{Im}U = -\frac{\mathrm{Im}(B\overline{D})}{|D|^2} \tag{2.218}$$

in terms of A, B, C and D in Eq. (2.156). In the beginning, we impose

$$\mathrm{Im}S \geq \frac{\sqrt{3}}{2} \tag{2.219}$$

since it is easier than ignoring the inside of the unit circle precisely.

Indeed, this assumption works well; First, we obtained all T-independent solutions on S-plane since we fixed the T-transformation. Then, we can choose only the solutions with $\mathrm{Im}S \geq \frac{\sqrt{3}}{2}$ without loss of generality. For the solutions with $\mathrm{Im}S < \frac{\sqrt{3}}{2}$, we can map them into the solutions with $\mathrm{Im}S \geq \frac{\sqrt{3}}{2}$, by an appropriate sequence of S- and T-transformations. Since we have all T-independent solutions with $\mathrm{Im}S \geq \frac{\sqrt{3}}{2}$, the newly obtained ones are already included in the former set. Then, we map the remaining solutions with $\mathrm{Im}S \geq \frac{\sqrt{3}}{2}$ into the solutions within $-\frac{1}{2} \leq \mathrm{Re}S < \frac{1}{2}$. Since we have not fixed S-transformation yet, we can simply ignore the solutions which are outside of the $SL(2,\mathbb{Z})$ fundamental region.

However, without the scaling duality, we cannot make an analogous assumption on U, since the fundamental region of $\Gamma_0(3)$ contains all $\mathrm{Im}U > 0$. Then, in this procedure, the finiteness is not ensured This difficulty comes from the fact that there are infinite regions which are equivalent to the $\mathcal{F}_{SL(2,\mathbb{Z})}$ within $|\mathrm{Re}U| \leq \frac{1}{2}$. Thus, the scaling duality is necessary to ensure finiteness if we rely on this method, as it gives the bound

$$\mathrm{Im}U \geq \frac{1}{2\sqrt{3}}. \tag{2.220}$$

Let us call the region $\mathcal{F} := |\mathrm{Re}U| \leq \frac{1}{2} \wedge \mathrm{Im}U \geq \frac{1}{2\sqrt{3}}$. The bound on $\mathrm{Im}S$ implies $|D|^2 \leq N_{\mathrm{flux}}$ via Eq. (2.188). Then, together with the quantization condition, it implies that c^0 and c^1 are multiples of 6 and 2 respectively, and satisfy

$$6\left\lceil -\frac{1}{6}\sqrt{N_{\mathrm{flux}}}\right\rceil \leq c^0 \leq 6\left\lfloor \frac{1}{6}\sqrt{N_{\mathrm{flux}}}\right\rfloor, \tag{2.221}$$

$$2\left\lceil -\frac{1}{2}\frac{\sqrt{3}\sqrt{N_{\mathrm{flux}}-(c^0)^2}+c^0}{3}\right\rceil \leq c^1 \leq 2\left\lfloor \frac{1}{2}\frac{\sqrt{3}\sqrt{N_{\mathrm{flux}}-(c^0)^2}-c^0}{3}\right\rfloor. \tag{2.222}$$

Here, we generate the possible values of c^0 from the first line. Then, we substitute each possible value into the second line, and we obtain the possible values of c^1.

On the other hand, due to the quantization condition and the definitions of the complexified fluxes, we see that $-\frac{1}{\sqrt{3}}\mathrm{Im}(C\overline{D})$ and $-2\sqrt{3}\mathrm{Im}(B\overline{D})$ are positive integers in this region. Then, from the relationship between the tadpole charge and moduli VEVs, we have

$$\frac{3|D|^2 N_{\mathrm{flux}}}{6} = \left(-\frac{1}{\sqrt{3}}\mathrm{Im}(C\overline{D})\right)\left(-2\sqrt{3}\mathrm{Im}(B\overline{D})\right). \tag{2.223}$$

Thus, we can determine B by the inequality

$$3|D|^2 \leq \mathrm{Im}D(2\sqrt{3}\mathrm{Re}B) - \mathrm{Im}B(2\sqrt{3}\mathrm{Re}D) \leq 3|D|^2\frac{N_{\mathrm{flux}}}{6}. \tag{2.224}$$

Note that D is already determined in terms of $c^{0,1}$.

Next, we fix C via Eq. (2.223). Of course, $-2\sqrt{3}\mathrm{Im}(B\overline{D})$ must divide out $\frac{3|D|^2 N_{\mathrm{flux}}}{6}$. Then, requiring also $\mathrm{Im}S \geq \frac{\sqrt{3}}{2}$ yields

- $c^0 > 0$ case
  d_0, a_0: fixed. We obtain

$$4\left[\frac{1}{4}\frac{|D|^2+\mathrm{Im}B2\sqrt{3}\mathrm{Re}D}{3\mathrm{Im}D}\right] \leq d_1 \leq 4\left\lfloor\frac{1}{4}\frac{3|D|^2\frac{N_{\text{flux}}}{6}+\mathrm{Im}B2\sqrt{3}\mathrm{Re}D}{3\mathrm{Im}D}\right\rfloor, \tag{2.225}$$

  and

$$a^1 = -\frac{1}{c^0}\left(\frac{3|D|^2N_{\text{flux}}}{6}\frac{1}{-2\sqrt{3}\mathrm{Im}(B\overline{D})} - a^0c^1\right). \tag{2.226}$$

  Here, $a^1 \in 2\mathbb{Z}$ must be checked due to the quantization condition.

- $c^0 = 0$, $c^1 > 0$ case
  d_1, a^1: fixed. We obtain

$$2\left[-\frac{1}{2}\frac{3|D|^2\frac{N_{\text{flux}}}{6}-2\sqrt{3}\mathrm{Im}D\mathrm{Re}B}{2\sqrt{3}\mathrm{Re}D}\right] \leq d_0 \leq 2\left\lfloor\frac{1}{2}\frac{-|D|^2+2\sqrt{3}\mathrm{Im}D\mathrm{Re}B}{2\sqrt{3}\mathrm{Re}D}\right\rfloor, \tag{2.227}$$

  and

$$a^0 = \frac{1}{c^1}\left(\frac{|D|^2N_{\text{flux}}}{6}\frac{1}{-2\sqrt{3}\mathrm{Im}(B\overline{D})} + a^1c^0\right), \tag{2.228}$$

  Here, $a^0 \in 6\mathbb{Z}$ must be checked due to the quantization condition.

Finally, A can be fixed via the F-term equation (2.187) $AD - BC = 0$ as

$$b_0 = -\mathrm{Im}\frac{CB}{D} = \frac{-1}{|D|^2}\left(\mathrm{Im}B\mathrm{Im}C\mathrm{Im}D - \mathrm{Re}B\mathrm{Re}C\mathrm{Im}D + \mathrm{Re}B\mathrm{Im}C\mathrm{Re}D + \mathrm{Im}B\mathrm{Re}C\mathrm{Re}D\right), \tag{2.229}$$

$$\begin{aligned} b_1 &= -\frac{2}{\sqrt{3}}\mathrm{Re}\frac{CB}{D} \\ &= \frac{-2}{\sqrt{3}|D|^2}\left(\mathrm{Re}B\mathrm{Im}C\mathrm{Im}D + \mathrm{Im}B\mathrm{Re}C\mathrm{Im}D - \mathrm{Im}B\mathrm{Im}C\mathrm{Re}D + \mathrm{Re}B\mathrm{Re}C\mathrm{Re}D\right). \end{aligned} \tag{2.230}$$

Here, $b^0 \in 2\mathbb{Z}$ and $b^1 \in 4\mathbb{Z}$ must be checked due to the quantization condition.

Thus, we have generated the whole flux vacua within $\mathcal{F}$ for given N_{flux}. They are representatives of T-transformation, but we have to deal with d.o.f. which is relevant to S'. It is surely easy, due to the correspondences in Figure 2.A.2. Thus we can simply drop the solutions which are inside the isometric circle.

Algorithm 2

Although we have not used that, but there is another algorithm to generate all possible flux vacua that does not need the scaling duality S' and is explained in the following for reference. For the fundamental region of $SL(2,\mathbb{Z})$, we can obtain the whole flux vacua in the region via Algorithm 1. As explained, the symmetry under the S-transformation is broken, and if one virtually transforms flux quanta by S, one would obtain fractional quanta. However, one can transform them to integral quanta by multiplying an appropriate integer factor k. Of course, the transformed fluxes have $k^2 N_{\text{flux}}$ as their charge when the original fluxes have N_{flux}. On the other hand, the symmetry under T-transformation is not broken. Hence, one can associate a set of solutions (N_{flux}) in the other parts of the fundamental region for $\overline{\Gamma}_0(3)$: $S\mathcal{F}_{SL(2,\mathbb{Z})}\backslash -\overline{\omega}$, $ST\mathcal{F}_{SL(2,\mathbb{Z})}$, and $ST^{-1}\mathcal{F}_{SL(2,\mathbb{Z})}$ with the set of solutions in $\mathcal{F}_{SL(2,\mathbb{Z})}$ whose charge is $k^2 N_{\text{flux}}$. Since it has been proven that the number of physically-distinct solutions in $\mathcal{F}_{SL(2,\mathbb{Z})}$ is finite, we can conclude that the number of those solutions in the whole fundamental region $\mathcal{F}_{\overline{\Gamma}_0(3)}$ is also finite.

More concretely, we may obtain the whole vacua by the following procedure:

1. Obtain all solutions in $\mathcal{F}_{SL(2,\mathbb{Z})}$ for some $3^2 N_{\text{flux}}$.

2. From the above solutions, one may find the solutions with

$$(b_0, d_0) \in 18\mathbb{Z}, \quad (b_1, d_1) \in 12\mathbb{Z}, \quad (a_0, c_0) \in 6\mathbb{Z}, \quad (a_1, c_1) \in 2\mathbb{Z}. \tag{2.231}$$

3. Then, the S image of the selected solutions can satisfy the correct quantization condition (see Eq. (2.205)) and their complex-structure moduli VEVs are in the $S\mathcal{F}$. Since this is not a sufficient condition, one must check if the obtained solutions are indeed quantized correctly.

4. For other regions, similar embeddings can be found.

This algorithm is apparently inefficient. Indeed, for the increase of N_{flux}, the maximum allowed values of c^0, c^1 behave as $\sim \sqrt{N_{\text{flux}}}$, while that of d_1 behaves $\sim N_{\text{flux}}^{3/2}$. Nevertheless, we can show the finiteness of the landscape without the scaling duality $S_{(3)}$, though the lack of the additional duality seems to make the discussion which is based on the bound for $\mathrm{Im}U$ invalid.

Counting the solutions

In counting the physically-distinct solutions with their degeneracy, one should be careful not to count the same solution more than once. This overcounting can happen when there is an additional symmetry on the VEV concerned. The additional symmetries are called the enhanced symmetries, and an enhanced symmetry emerges as a stabilizer subgroup of the background (modular) group. We call a point of the upper half-plane $\mathbb{H}$ a fixed point if there is a non-trivial stabilizer subgroup on the point. As for the modular $PSL(2,\mathbb{Z})$ group, there are three fixed points: ω, i, and $i\infty$ in the fundamental region. In particular, the former two points are named the elliptic points since they are the fixed points under elliptic transformations. Since we can stabilize the complex-structure moduli at these elliptic points exactly ($i\infty$ is only approximately realized in terms of $q = e^{2\pi i U}$), we focus on them.

The $\overline{\Gamma}_0(3) \rtimes \mathbb{Z}_2$ also has two elliptic points: $\omega' := \frac{1}{2} + i\frac{\sqrt{3}}{6}$ and $\frac{i}{\sqrt{3}}$. We denote an enhanced symmetry by adding the corresponding point to its subscript. Hence we have $\mathbb{Z}_3^{(\omega)}$ and $\mathbb{Z}_2^{(i)}$ for the $PSL(2,\mathbb{Z})$, $\mathbb{Z}_3^{(\omega')}$ and $\mathbb{Z}_2^{\left(\frac{i}{\sqrt{3}}\right)}$ for the $\overline{\Gamma}_0(3) \rtimes \mathbb{Z}_2$. Note that, either of $S, U = i$ cannot be realized on the landscape as pointed out. In fact, the solutions to the F-term equations (2.187) are

$$
\begin{aligned}
\langle S \rangle &= -\frac{C}{D} = \frac{\left(\sqrt{3}a^1 + \frac{1}{\sqrt{3}}a^0\right) + ia^0}{\left(\sqrt{3}c^1 + \frac{1}{\sqrt{3}}c^0\right) + ic^0} \\
\langle U \rangle &= -\frac{B}{D} = \frac{\frac{\sqrt{3}}{2}d_1 + id_0}{\left(\sqrt{3}c^1 + \frac{1}{\sqrt{3}}c^0\right) + ic^0}.
\end{aligned}
\tag{2.232}
$$

Thus, taking into account the quantization condition (2.157), we see the solutions are in the quadratic field $\mathbb{Q}(\sqrt{-3})$. This phenomenon occurs only in the orientifolds with $h^{2,1} = 1$.

On those elliptic points, we have to divide the number of solutions further by some constant factors, even after we obtained all solutions using Algorithm 1 or Algorithm 2 since we fixed the d.o.f of the modular groups only partially. Basically, the factor is given by the order of the enhanced group as it indicates the length of the orbit of the stabilizer group. Thus, we have to divide the numbers of solutions on $(S,U) = (\omega, \cdot)$ and $(\cdot, \omega)$ by three. However, $(S,U) = (\omega, \omega')$ must be divided by three, not nine. This is because ω and ω' are $SL(2,\mathbb{Z})$ equivalent points and $h^{2,1} = 1$. More concretely, $\langle W \rangle$ on (ω, ω') is written by

$$
\langle W \rangle = A + B\omega + \omega'[C + D\omega] = 0, \tag{2.233}
$$

but a $SL(2,\mathbb{Z})_S$ transformation maps it into

$$
\langle W \rangle = A' + B'\omega' + \omega'[C + D\omega'] = 0. \tag{2.234}
$$

Apparently, there emerges the symmetry $B' \leftrightarrow C'$. Then, let us recall that each of (A, C) (RR) and (B, D) (NS) forms a doublet under the modular transformation of U. Schematically, under the $g \in \Gamma_0(3)$ transformation, we have

$$
\begin{pmatrix} A' \\ C' \end{pmatrix} \to \rho(g) \begin{pmatrix} A' \\ C' \end{pmatrix}, \quad \begin{pmatrix} B' \\ D' \end{pmatrix} \to \rho(g) \begin{pmatrix} B' \\ D' \end{pmatrix}, \tag{2.235}
$$

for the representation matrix $\rho(g)$ (2.165). Then, let us suppose that $g \in \Gamma_0(3)^{(\omega')}$. The $B' \leftrightarrow C'$ symmetry implies that the above transformation is also written by

$$
\begin{pmatrix} A' \\ B' \end{pmatrix} \to \rho(g) \begin{pmatrix} A' \\ B' \end{pmatrix}, \quad \begin{pmatrix} C' \\ D' \end{pmatrix} \to \rho(g) \begin{pmatrix} C' \\ D' \end{pmatrix}. \tag{2.236}
$$

(A, B) and (C, D) form doublets under the S-duality. Indeed, $\rho(g)$ is also the representation of $SL(2,\mathbb{Z})$. The two stabilizer groups $\Gamma_0(3)^{(\omega')}$ and $SL(2,\mathbb{Z})^{(\omega')}$ are equivalent, i.e., we have $\Gamma_0(3)^{(\omega')} \simeq SL(2,\mathbb{Z})^{(\omega')}(\simeq \mathbb{Z}_3)$ on ω'. This can be checked via the relation $S'_{\Gamma_0(3)_U} = (T^{-1}S)_{SL(2,\mathbb{Z})_S}$ (up to a sign) that is satisfied at (ω,ω'). Therefore, the dividing factor is three, not nine.

This is in contrast to the other elliptic point $\frac{i}{\sqrt{3}}$, which is not equivalent to the other elliptic points of $PSL(2,\mathbb{Z})$. On the other hand, the overall sign-flip $(F_3,H_3) \to -(F_3,H_3)$ is equivalent to $\Gamma_0(3)^{(\frac{i}{\sqrt{3}})}$ on the elliptic point $U = \frac{i}{\sqrt{3}}$. Since we have already modded out the landscape by the sign-flip in Algorithm 1, we do not have to divide the solutions on $U = \frac{i}{\sqrt{3}}$ further.

2.A.2 The most general duality in reparametrizations of U

In this section, we give an explicit proof of the fact that the most general symmetry of the $\mathbb{Z}_{6-\mathrm{II}}$ period vector (2.170) is the $\overline{\Gamma}_0(3) \rtimes S_{(3)}$. Here, a symmetry is considered as a candidate when the symplectic transformation $Sp(b_3 = 4,\mathbb{Z})$ of the third-cohomology is compatible with the symmetry. Since we assume that such symmetries act on the complex-structure modulus, we begin the discussion by determining the most general action in terms of the symplectic transformation.

A pair of a G_3 flux and a period matrix, (G_3,Π), forms an orbit under the $Sp(b_3,\mathbb{Z})$ up to a Kähler transformation:

$$(G_3,\Pi) \to (MG_3, kM\Pi), \tag{2.237}$$

with $k := k(U)$ being a factor which is independent of the background coordinates and $M \in Sp(4,\mathbb{Z})$. The pairs (G_3,Π) which enter the same orbit are identified under the symmetry. If we write a representative of the pairs as (G_3,Π), the two pairs (G_3,Π) and (MG_3,Π) are not in the same class (M is again an arbitrary $Sp(4,\mathbb{Z})$ matrix). G_3 is independent of the complex-structure modulus U. Then, the compatibility means that these different classes can be further identified by a reparameterization of U: $U \to U' := f(U)$. The action of $\overline{\Gamma}_0(3)$ and the scaling duality $S_{(3)}$ are examples of the case. In other words, the reparameterization $U' = f(U)$ and the basis transformation of the third cohomology $Sp(b_3,\mathbb{Z})$ are compatible if and only if

$$^\exists f : U' = f(U), \left(^\exists M \in Sp(4,\mathbb{Z}), ^\exists k = k(U) \in \mathbb{C} \text{ s.t. } \Pi(U') = \Pi\left(f(U)\right) = k(U)M\Pi(U)\right). \tag{2.238}$$

If this is the case, the two orbit classes are identified as $[MG_3,\Pi(U')] = [MG_3, k(U)M\Pi(U)]$. This reparameterization connects the two solutions with U and U' on the VEVs plane, and we have to choose one representative from each total orbit. Note that the reparametrizations and $Sp(b_3,\mathbb{Z})$ do not need to be in one-to-one correspondence.

Then, what is the most general reparameterization of U where there is a possibility to be compatible with the basis transformation $Sp(4,\mathbb{Z})$? In fact, the reparameterization has to be a simple linear fractional transformation with integer coefficients[27],

$$U \to \frac{aU+b}{cU+d}. \tag{2.239}$$

[27] Although we can choose the coefficients to satisfy $ad - bc = 1$ and to be real (i.e., $SL(2,\mathbb{R})$), we do not redefine a,b,c,d so because we can discuss the following argument when the parameters are integers.

This comes from the definition of U:

$$U \propto \frac{\int_B \Omega}{\int_A \Omega}. \tag{2.240}$$

Then, one can see the linear fractional transformation is induced via the basis transformation

$$\begin{pmatrix} \int_A' \Omega \\ \int_B' \Omega \end{pmatrix} = M \begin{pmatrix} \int_A \Omega \\ \int_B \Omega \end{pmatrix}, \tag{2.241}$$

with $M \in Sp(4, \mathbb{Z})$. Hence, we can find the most general transformation U that is compatible with the basis transformation by finding conditions[28] on a, b, c, d to make M an appropriate symplectic matrix.

To derive the identities, it is important to note that the factor $k = k(U)$ can be factorized as $k(U) = (cU + d)^{-1}k'$, where k' is a numerical factor that is independent of U. This can be shown as follows. The compatibility of the linear fractional transformation and the basis transformation and a possible Kähler transformation leads

$$\begin{pmatrix} cU + d \\ -i\sqrt{3}(aU + b) \\ 3(aU + b) \\ i\sqrt{3}(cU + d) \end{pmatrix} = k' \begin{pmatrix} x_{11} + i\sqrt{3}x_{14} + U(-x_{12}i\sqrt{3} + 3x_{13}) \\ x_{21} + i\sqrt{3}x_{24} + U(-x_{22}i\sqrt{3} + 3x_{23}) \\ x_{31} + i\sqrt{3}x_{34} + U(-x_{32}i\sqrt{3} + 3x_{33}) \\ x_{41} + i\sqrt{3}x_{44} + U(-x_{42}i\sqrt{3} + 3x_{43}) \end{pmatrix}, \tag{2.242}$$

where $k' = (cU + d)k$ and (x_{ij})s are matrix elements of $X \in Sp(4, \mathbb{Z})$ if the transformation of U is a symmetry. The basis of the period matrix is chosen to be $\Pi = (1, -\sqrt{3}iU, 3U, \sqrt{3}i)^T$. As Eqs. (2.242) must hold everywhere of the complex-structure modulus space, it is sufficient to assume that $k' \sim U$ or $k' \sim \frac{1}{U}$ if it depends on U. For each dependence, the resultant identities imply that $-x_{i2}i\sqrt{3} + 3x_{i3} = 0$ and $x_{i1} + i\sqrt{3}x_{i4} = 0$ $(\forall i)$ respectively. Both of these relations make the determinant of X vanish, hence it contradicts the assumption. Henceforth, we denote the numerical part k' as k.

Therefore, we are left with the following identities:

$$\begin{aligned} x_{14} &= \frac{1}{\sqrt{3}k}\left(-i(cU + d) + ik(x_{11} + 3Ux_{13}) + \sqrt{3}kUx_{12}\right), \\ x_{24} &= \frac{1}{3k}\left(-3(aU + b) + ik\sqrt{3}(x_{21} + Ux_{23}) + 3kUx_{22}\right), \\ x_{34} &= \frac{1}{3}\left(3(x_{21} + 3Ux_{23} + Ux_{32}) + i\sqrt{3}(x_{31} - 3Ux_{22} + 3x_{24} + 3Ux_{33})\right), \\ x_{44} &= \frac{1}{\sqrt{3}}\left(\sqrt{3}(x_{11} + 3Ux_{13} + Ux_{42}) + i(x_{41} + 3x_{14} - 3Ux_{12} + 3Ux_{43})\right), \end{aligned} \tag{2.243}$$

[28]Since the reparameterization must be defined on the whole modulus space, we obtain the conditions as identities with regard to U.

where the last two identities make the form of X

$$X = \begin{pmatrix} x_{44} & x_{43} & -\frac{x_{42}}{3} & -\frac{x_{41}}{3} \\ x_{34} & x_{33} & -\frac{x_{32}}{3} & -\frac{x_{31}}{3} \\ x_{31} & x_{32} & x_{33} & x_{34} \\ x_{41} & x_{42} & x_{43} & x_{44} \end{pmatrix}. \tag{2.244}$$

On the other hand, the first two identities imply that

$$\begin{aligned} -\frac{\sqrt{3}x_{12}}{c} = \frac{\sqrt{3}x_{14}}{d} = \frac{x_{21}}{\sqrt{3}b} = \frac{\sqrt{3}x_{23}}{a} = \mathrm{Im}\frac{1}{k}, \\ \frac{3x_{13}}{c} = \frac{x_{11}}{d} = -\frac{x_{24}}{b} = \frac{x_{22}}{a} = \mathrm{Re}\frac{1}{k}. \end{aligned} \tag{2.245}$$

when a, b, c, d do not vanish. If some of them are zero, the identities force the (x_{ij})s that enter the corresponding numerators to be zero. The argument is divided into cases according to which of a, b, c, or d are 0, as follows.

Case I. $a, b, c, d \neq 0$
In this case, X becomes

$$X = \begin{pmatrix} \frac{d}{a}x_{33} & \frac{c}{3a}x_{32} & \frac{c}{3a}x_{33} & -\frac{d}{3a}x_{32} \\ -\frac{b}{a}x_{32} & x_{33} & -\frac{x_{32}}{3} & -\frac{b}{a}x_{33} \\ \frac{3b}{a}x_{33} & x_{32} & x_{33} & -\frac{b}{a}x_{32} \\ \frac{d}{a}x_{32} & -\frac{c}{a}x_{33} & \frac{c}{3a}x_{32} & \frac{d}{a}x_{33} \end{pmatrix}. \tag{2.246}$$

where all elements must be integers. The condition that X is symplectic: $X^T \Sigma X = \Sigma$ leads

$$(ad - bc)\frac{x_{32}^2 + 3x_{33}^2}{3a^2} = 1. \tag{2.247}$$

Since a, b, c, d are all integers and (x_{ij})s are real, we have $ad - bc \geq 1$.

The case $ad - bc = 1$ is rather simple. (x_{ij})s are all integers and we have $x_{32} = \frac{3a}{c}m, x_{33} = \frac{3a}{c}n$ $(m, n \in \mathbb{Z})$ and

$$\frac{3(m^2 + 3n^2)}{c^2} = 1. \tag{2.248}$$

Therefore c must be $c \in 3\mathbb{Z}$ and the linear fractional transformation forms $\overline{\Gamma}_0(3)$. Since we have proven a linear fractional transformation with non-vanishing a, b, c, d and $ad - bc = 1$ is an element of $\overline{\Gamma}_0(3)$ and $\overline{\Gamma}_0(3)$ is compatible with the basis transformation in Section 2.4.1.3, the whole $\overline{\Gamma}_0(3)$ is the most general symmetry in this case.

Let us next consider $ad - bc > 1$ case. Following the previous lines, we have

$$3(m^2 + 3n^2)(ad - bc) = c^2, \tag{2.249}$$

and thus $c \in 3\mathbb{Z}$. For the case $x_{33} \neq 0$ i.e., $n \neq 0$, Eq. (2.247) leads

$$\begin{aligned} x_{32}^2 + 3x_{33}^2 &= \frac{3a^2}{ad - bc} \\ &= \frac{3}{\frac{d}{a} - (-\frac{b}{a})3(-\frac{c}{3a})} \\ &= \frac{3x_{33}^2}{\left(\frac{d}{a}x_{33}\right)x_{33} - \left(-\frac{b}{a}x_{33}\right)3\left(-\frac{c}{3a}x_{33}\right)} \quad (x_{33} \neq 0). \end{aligned} \tag{2.250}$$

Here, all $e := \frac{d}{a}x_{33}, f := -\frac{b}{a}x_{33}, g := -\frac{c}{3a}x_{33}$ must be integer due to the explicit form of X (2.246). It implies that

$$ex_{33} - 3fg = \frac{3x_{33}^2}{x_{32}^2 + 3x_{33}^2} \tag{2.251}$$

must be an integer. Both side of the above equation are integers i.e. $x_{32}^2 + 3x_{33}^2 \leq 3x_{33}^2$, and it leads $x_{32} = 0$. Similarly, we have $x_{33} = 0$ when $x_{32} \neq 0$ is assumed. Hence the following two cases are left:

Case I-i. $ad - bc > 1$ **and** $x_{33} \neq 0, x_{32} = 0$
Here, X is given by

$$X = x_{33} \begin{pmatrix} \frac{d}{a} & 0 & \frac{c'}{a} & 0 \\ 0 & 1 & 0 & -\frac{b}{a} \\ \frac{3b}{a} & 0 & 1 & 0 \\ 0 & -\frac{3c'}{a} & 0 & \frac{d}{a} \end{pmatrix} \tag{2.252}$$

with $c' := \frac{c}{3} \in \mathbb{Z}$ and $X^T \Sigma X = \Sigma$ gives

$$|x_{33}| = \frac{|a|}{\sqrt{ad - bc}}. \tag{2.253}$$

The requirement that X is an integral matrix implies that the following four quantities

$$\tilde{a} = \frac{a}{\sqrt{ad - bc}}, \tilde{b} = \frac{b}{\sqrt{ad - bc}}, \tilde{c}' = \frac{c'}{\sqrt{ad - bc}}, \tilde{d}' = \frac{d}{\sqrt{ad - bc}} \tag{2.254}$$

are all integers and these satisfy $\tilde{a}\tilde{d} - 3\tilde{b}\tilde{c}' = 1$. However, the existence of these integers implies $\gcd(a, b, c, d) \geq \sqrt{ad - bc} > 1$ and a redefinition of a, b, c, d can reduce the problem to the $ad - bc = 1$ case. Hence, we have again $\Gamma_0(3)$ for the case $x_{32} = 0$.

Case I-ii. $ad - bc > 1$ **and** $x_{32} \neq 0, x_{33} = 0$
In this case, X becomes

$$X = \frac{x_{32}}{3} \begin{pmatrix} 0 & \frac{c}{a} & 0 & -\frac{d}{a} \\ -\frac{3b}{a} & 0 & -1 & 0 \\ 0 & 3 & 0 & -\frac{3b}{a} \\ \frac{3d}{a} & 0 & \frac{c}{a} & 0 \end{pmatrix}. \tag{2.255}$$

The requirement that X is an integral matrix implies that the four quantities $m_1 := \frac{c}{3a}x_{32}, m_2 := \frac{d}{3a}x_{32}, m_3 := \frac{b}{a}x_{32}, m_4 := \frac{x_{32}}{3}$ are all integers. On the other hand, the condition $X^T\Sigma X = \Sigma$:

$$|x_{32}| = \frac{\sqrt{3}|a|}{\sqrt{ad-bc}} \tag{2.256}$$

gives

$$m_4 = \frac{1+m_1m_3}{3m_2}, b = \frac{a}{3m_4}, c = -3(bm_1^2 - am_1m_2), d = -3(bm_1m_2 - am_2^2) \tag{2.257}$$

and $a^2 = 3m_4^2(ad-bc)$. Therefore a, c, d are multiples of three and we have

$$X = \frac{1}{\sqrt{3a'd'-bc'}}\begin{pmatrix} 0 & c' & 0 & -d' \\ -b & 0 & -a' & 0 \\ 0 & 3a' & 0 & -b \\ 3d' & 0 & c' & 0 \end{pmatrix} \tag{2.258}$$

for $a := 3a', c := 3c', d := 3d'$. Therefore

$$\tilde{a}' = \frac{a'}{\sqrt{3a'd'-bc'}}, \tilde{b} = \frac{b}{\sqrt{3a'd'-bc'}}, \tilde{c}' = \frac{c'}{\sqrt{3a'd'-bc'}}, \tilde{d}' = \frac{d'}{\sqrt{3a'd'-bc'}}. \tag{2.259}$$

must be integers and these satisfy $3\tilde{a}'\tilde{d}' - \tilde{b}\tilde{c}' = 1$. Using these variables X is rewritten by

$$X = \begin{pmatrix} 0 & \tilde{c}' & 0 & -\tilde{d}' \\ -\tilde{b} & 0 & -\tilde{a}' & 0 \\ 0 & 3\tilde{a}' & 0 & -\tilde{b} \\ 3\tilde{d} & 0 & \tilde{c}' & 0 \end{pmatrix}, \tag{2.260}$$

which is an integral symplectic matrix. This X can be expressed as a product of a $\Gamma_0(3)$ transformation and the scaling duality $U \to -\frac{1}{3U}$. Let $\begin{pmatrix} v & w \\ 3x & y \end{pmatrix}$ a matrix in $\Gamma_0(3)$. Then, the consecutive transformation on the period matrix leads

$$\begin{pmatrix} y & 0 & x & 0 \\ 0 & v & 0 & -w \\ 3w & 0 & v & 0 \\ 0 & -3x & 0 & y \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & y & 0 & x \\ v & 0 & -w & 0 \\ 0 & 3w & 0 & v \\ -3x & 0 & d & 0 \end{pmatrix}, \tag{2.261}$$

or

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} y & 0 & x & 0 \\ 0 & v & 0 & -w \\ 3w & 0 & v & 0 \\ 0 & -3x & 0 & y \end{pmatrix} = \begin{pmatrix} 0 & v & 0 & -w \\ y & 0 & x & 0 \\ 0 & -3x & 0 & y \\ 3w & 0 & v & 0 \end{pmatrix}. \tag{2.262}$$

In the former equation, we can put

$$v = -\tilde{b}, w = \tilde{a}', x = -\tilde{d}', y = \tilde{c}' \tag{2.263}$$

and obtain the X in Eq. (2.260). Similarly, the latter one reduces to the X.

Since $\Gamma_0(3)$ and the scaling duality are independent, the most general symmetries of U in this case are $\overline{\Gamma}_0(3)$ and the scaling duality, and the total symmetry is given by their outer semidirect product.

Case II. $a, d = 0, b, c \neq 0$
In the usual $SL(2,\mathbb{Z})$, we will find the S-transformation in this case. Now, X is written by

$$X = \begin{pmatrix} 0 & -\frac{c}{3b}x_{34} & \frac{c}{9b}x_{31} & 0 \\ x_{34} & 0 & 0 & -\frac{x_{31}}{3} \\ x_{31} & 0 & 0 & x_{34} \\ 0 & -\frac{c}{3b}x_{31} & -\frac{c}{3b}x_{34} & 0 \end{pmatrix}, \tag{2.264}$$

and there must be integers m, n such that $x_{31} = \frac{9b}{c}m, x_{34} = \frac{3b}{c}n,\ (m, n \in \mathbb{Z})$. The condition $X^T \Sigma X = \Sigma$:

$$-3b(3m^2 + n^2) = c \tag{2.265}$$

leads $c \in 3\mathbb{Z}$ and $bc < 0$.

Let us denote $\frac{3b}{c}$ by q^{-1}, where $q = -(3m^2 + n^2) \in -\mathbb{N}$. Then, $x_{31} = \frac{3m}{q}, x_{34} = \frac{n}{q}$ and we find that $x_{31} \in 3\mathbb{Z}$. This is consistent with the definition of q if and only if $m = 0, n = 1$. We have $b = -\frac{c}{3}$, and X is reduced to the scaling duality. Hence, S-transformation is broken and the scaling duality remains.

Case III. $b, c = 0$ **,** $a, d \neq 0$ X is written by

$$X = \begin{pmatrix} \frac{d}{a}x_{33} & 0 & 0 & -\frac{d}{3a}x_{32} \\ 0 & x_{33} & -\frac{x_{32}}{3} & 0 \\ 0 & x_{32} & x_{33} & 0 \\ \frac{d}{a}x_{32} & 0 & 0 & \frac{d}{a}x_{33} \end{pmatrix}. \tag{2.266}$$

Putting $x_{32} = -\frac{3a}{d}m$ and $x_{33} = \frac{a}{d}n$,

$$\frac{a}{d}(3m^2 + n^2) = 1 \tag{2.267}$$

is imposed by $X^T \Sigma X = \Sigma$. Hence, $q := \frac{d}{a} = 3m^2 + n^2 \in \mathbb{N}$ and $x_{32} \in 3\mathbb{Z}, x_{33} \in \mathbb{Z}$ hold. The case $m = 0, n = 1$ is only consistent with the definition of q, and it gives the trivial transformation $X = \mathbf{1}, U \to U$.

Case IV. $a = 0$**,** $b, c, d \neq 0$

$$X = \begin{pmatrix} \frac{d}{3b}x_{31} & \frac{c}{3d}x_{41} & -\frac{c}{9b}x_{31} & -\frac{x_{41}}{3} \\ -\frac{b}{d}x_{41} & 0 & 0 & -\frac{x_{31}}{3} \\ x_{31} & 0 & 0 & -\frac{b}{d}x_{41} \\ x_{41} & \frac{c}{3b}x_{31} & \frac{c}{3d}x_{41} & \frac{d}{3b}x_{31} \end{pmatrix} \tag{2.268}$$

holds and $X^T\Sigma X = \Sigma$ implies

$$\begin{aligned} x_{31}x_{41} &= 0, \\ \frac{c}{9b}x_{31}^2 - \frac{bc}{3d^2}x_{41}^2 &= 1. \end{aligned} \tag{2.269}$$

Hence we have $x_{31} = 0$ or $x_{41} = 0$. When x_{41} vanishes, we have a product of two integers:

$$\frac{c}{9b}x_{31}^2 = \left(-\frac{c}{3b}x_{31}\right)\left(-\frac{x_{31}}{3}\right) = 1. \tag{2.270}$$

However, this contradicts the requirement $\frac{c}{9b}x_{31}$ is an integer.

Let us assume $x_{31} = 0$. Similarly, we have

$$-\frac{bc}{3d^2}x_{41}^2 = \left(\frac{c}{3d}x_{41}\right)\left(-\frac{b}{d}x_{41}\right) = 1, \tag{2.271}$$

which implies $b = -\frac{c}{3}$. $d^2 = b^2x_{41}^2$ and $x_{41} \in 3\mathbb{Z}$ imply that $d \in 3\mathbb{Z}$. Then, the reparameterization becomes

$$U \to -\frac{m}{3(mU+n)}, \tag{2.272}$$

where $c = 3m$ and $d = 3n$, $m, n \in \mathbb{Z}$. From $(-\frac{x_{41}}{3})^2 = \frac{d^2}{c^2} \in \mathbb{Z}$, we also find $\frac{n}{m} \in \mathbb{Z}$ and the reparameterization is reduced to

$$U \to U = -\frac{1}{3(U+q)} \quad q \in \mathbb{Z}, \tag{2.273}$$

which is a product of the T-transformation and the scaling duality.

Case V. $b = 0$**,** $a, c, d \neq 0$

X is written by

$$X = \begin{pmatrix} \frac{d}{a}x_{33} & \frac{c}{3a}x_{32} & \frac{c}{3a}x_{33} & -\frac{d}{3a}x_{32} \\ 0 & x_{33} & -\frac{x_{32}}{3} & 0 \\ 0 & x_{32} & x_{33} & 0 \\ \frac{d}{a}x_{32} & -\frac{c}{a}x_{33} & \frac{c}{3a}x_{32} & \frac{d}{a}x_{33} \end{pmatrix}, \tag{2.274}$$

and one can show that $X^T\Sigma X = \Sigma$ implies

$$U \to \frac{U}{3qU+1} \quad q \in \mathbb{Z}, \tag{2.275}$$

in the same manner as case III. Hence we have again $\overline{\Gamma}_0(3)$ on U.

Case VI. $c = 0$, $a, b, d \neq 0$
Following the lines of the previous cases, one obtains $U' = U + q,\ q \in \mathbb{Z}$. Thus, we find the usual T-transformation.

Case VII. $d = 0$, $a, b, c \neq 0$
Similarly, one finds $U' = \frac{3qU-1}{3q} = -\frac{1}{3U} + q,\ q \in \mathbb{Z}$. Thus, U' is a composition of the T-transformation and the scaling duality.

Chapter 3

Machine Learning on Landscape

In the previous chapter, we obtained the whole flux landscape on the toroidal orientifolds and discussed its statistical properties. However, there is another approach that can analyze structures of the landscape, namely machine learning. In this paper, machine learning refers to the use of neural networks that are highly expressive with a non-linear structure, including contributions from activation functions. Among them, neural networks with many layers called deep neural networks, have become easily accessible due to recent improvements in the computational power of computers. Although machine learning has become ubiquitous in the modern era, it is not limited to real-world applications such as industrial applications. In fact, machine learning can be used for more general tasks. In other words, it can also be used in theoretical physics. The actual use of machine learning in particle physics has been more broadened (see for reviews, e.g., [159, 160] in particle physics and [161] in more general physics.). We analyze the string landscape via the machine learning approach in this chapter and find some observations.

We start with a brief review of type IIA intersecting D-brane models in Section 3.1. Then, in Section 3.2, we shortly mention the previous studies on this topic and explain an architecture that we will consider later. After the reviews, we have the motivation to study a machine-learning-based analysis on the landscape of intersecting D-brane models and discuss our observations there via the so-called autoencoder models in Section 3.3.

3.1 Some Basic Ingredients in Intersecting D-brane Models and Applications of Machine Learning

NOTE: This section (Section 3.1) is fully dedicated for a review.

In this section, we first outline some basic ingredients of intersecting D-brane models (see for reviews, [1, 23, 27]).

3.1.1 Parallel and Coincident D-branes

For the later sections, we start with an overview of a certain class of D-brane configurations, where the D-branes are aligned in parallel and possibly coincident with each other. At first glance, it seems peculiar since D-branes are charged objects. Nevertheless in the supersymmetric case the gravitational (attractive) interaction where the DBI action plays a role and the Coulomb (repulsive) interaction where the CS action plays a role exactly cancel each other.

As a single D-brane leads a $U(1)$ vector multiplet, a set of N parallel (but separated) D-branes lead to $U(1)^n$ gauge bosons in the massless spectrum. If we label the parallel D-branes by $a, b, \ldots$, the gauge bosons come from open strings connecting between a-th and a-th D-branes (aa-sector). For the ab-sector, the open strings become massive as we expected by their tension, and they are charged by $U(1)_a \times U(1)_b$ as $(+1, -1)$.

On the other hand, the coincident N parallel D-branes cannot be distinguished from each other. Then the $U(1)\times U(1)$ gauge group enhances to a single $U(n)$ group. Note that superstring theory has a virtue in this point, since in bosonic string theory one cannot obtain a conserved charge associated with the endpoints of open strings.

Furthermore, the flat potential between parallel but separated D-branes and coincident ones naturally explains the Higgs mechanism, as the above gauge groups imply.

3.1.2 Intersecting D6-branes

Next, we discuss a system of D-branes intersecting at angles, where broad phenomenological models have been constructed. In the following, we restrict ourselves to consider configurations only of D6-branes which wrap three-cycles of type IIA CY orientifolds (again, the results are the same also for the $N = 1$ toroidal orientifold), as these will be the focus of our interest.

For illustrative purposes, let us assume two intersecting D6-branes at a nonzero angle in a flat 10d space whose 6d space is decomposed into triple $\mathbb{R}^2$ planes. When the angle between two D6-branes on a plane is θ, we have an open string of length l satisfying the following boundary conditions (which is simply obtained by setting direction i to be parallel to the first D6-brane and θ to be oriented from the first one to the second one):

$$\partial_\sigma X^i\big|_{\sigma=0} = 0, \qquad \big(\cos\theta\partial_\sigma X^i + \sin\theta\partial_\sigma X^{i+1}\big)\big|_{\sigma=l} = 0, \tag{3.1}$$

$$\partial_\tau X^{i+1}\big|_{\sigma=0} = 0, \quad \big(-\sin\theta\partial_\tau X^i + \cos\theta\partial_\tau X^{i+1}\big)\big|_{\sigma=l} = 0, \tag{3.2}$$

where (τ, σ) are coordinates of the worldsheet and $(i, i+1)$ are directions which we now consider. Each line comes from the aforementioned NN and DD boundary conditions, respectively. The

Polyakov action with the above action gives the usual 2d wave equation $(-\partial_\tau^2+\partial_\sigma^2)X^{i,i+1}(\tau,\sigma)=0$, and the solutions are independently obtained for their left- and right-moving d.o.f. Following the lines of [26], we complexify them into $Z^i = X^i + iX^{i+1}$ and obtain its mode expansion. Recasting the boundary conditions into those of Z makes it easy to guess the mode expansion:

$$\partial_\sigma \mathrm{Re} Z^i\big|_{\sigma=0} = 0, \quad \partial_\sigma \left[\mathrm{Re}(e^{i\theta}Z^i)\right]\big|_{\sigma=l} = 0, \tag{3.3}$$

$$\partial_\tau \mathrm{Im} Z^i\big|_{\sigma=0} = 0, \quad \partial_\tau \left[\mathrm{Im}(e^{i\theta}Z^i)\right]\big|_{\sigma=l} = 0. \tag{3.4}$$

Due to the definition, for $\theta = 0$, the real part and the imaginary part satisfy NN ($\sim$ cos) and DD ($\sim$ sin) conditions respectively. In the presence of θ, to explain the phase shift, the mode n is shifted as $n + \theta/\pi$ (in general it is not a (half-)integer). The same is true for the fermion part. The linear terms which appear in the usual NN open string are absent, and the motion of the center of mass is frozen to the intersecting point of the D-branes. Thus they are localized at 4d intersecting region, and corresponding states include fermions in the bifundamental representation $(\mathbf{N}_a, \overline{\mathbf{N}_b})$ of $U(N_a) \times U(N_b)$ ($U(1)$ charges are implicitly understood as $+1, -1$ respectively) when the intersecting two D-branes promote to two stacks of N_a, N_b D-branes. Here, the relative sign comes from the fact the endpoints have opposite orientations to each other while it can be already seen in bosonic string theory. Since these states are associated with each intersection point, we have a multiplicity of I_{ab}, where I_{ab} denotes the intersection number of two stacks a, b. Considering the initial $(\mathbb{R}^2)^3$, the $N = 1$ SUSY condition is given by $\theta_1 + \theta_2 + \theta_3 = 0 \pmod{2\pi}$, where $\theta_{1,2,3}$ are angles in each $\mathbb{R}^2$. In the orientifold compactifications on general CY three-folds, the SUSY condition imposed on a stack of D6-brane to wrap a certain class of three-cycles, which is called special Lagrangian (sLag) cycles. The compactifications also require the tadpole cancellation condition as in the type IIB O3/D3 case.

Although our aim is not to construct a concrete model, we note that there are additional representations in the presence of O-planes, due to the appearance of a mirror stack under $\mathcal{R}$ for each stack of D-branes. On the $T^6/\mathbb{Z}_2 \times \mathbb{Z}_2'$ toroidal orientifold with no discrete torsion (there is an ambiguity in the definition of the torsion, see for the concrete discussion [162]), a stack of N D-branes coincident with an O6-plane yields $USp(2N)$ (which also has an ambiguity in its definition, and we define $USp(2) \simeq SU(2)$). The other intersections with a mirror stack can be used to introduce fermions in (anti-)symmetric representation.

Further details including the consistency conditions will be summarized in the actual study on the $T^6/(\mathbb{Z}_2 \times \mathbb{Z}_2')$ orientifold.

3.2 Some Machine Learning Techniques and Its Applications in String Theory

NOTE: This section (Section 3.2) is fully dedicated to a review.

The birth of the concept of machine learning (ML) in the modern sense dates back more than 60 years ago (for a historical review, see [163], and for a review of fundamental concepts, see a textbook [164]). These early studies gave rise to the concept of ML based on neural networks, the most promising candidate for a paradigm shift in the 21st century while ML is a very broad term. Here, a neural network is usually defined as a sequence of pairs of a linear map and a nonlinear function which is called an activation function. The method based on the neural networks is supported by the so-called universal approximation theorem (see, e.g., [165] which is a monumental paper in this context)[1]. Indeed, even a shallow one, a neural network is claimed to be able to express almost any functions (at least continuous functions in the L^p $(1 \leq p < \infty)$ space are included). In our analysis, we use the neural-network-based method in certain type IIA intersecting D-brane models.

As ML becomes more widely used, it is indeed becoming increasingly applied to string theory. See for very early works based on genetic algorithms (a different notion from ML based on the neural network), e.g., [169, 170], and those based on the neural networks, [171–173]. Furthermore, the combination of two methods is performed in [174]. In this section, we give a brief review of some ML techniques. Due to the large number of studies in this context, only a few comprehensive reviews are available. See for more reviews of ML-based studies in string theory, e.g., [45, 175–178] and Section 2 in [179].

3.2.1 Methods based on Machine Learning

As mentioned, the term machine learning has very broad meanings. In general, it is classified into three major paradigms i.e., supervised learning, unsupervised learning, and reinforcement learning [180]. While there are other paradigms including semi-supervised learning, let us focus on these main paradigms and explain them shortly. The actual use cases in string theory are comprehensively surveyed in [45].

Supervised learning

Supervised learning is perhaps the simplest concept in machine learning. Indeed it includes classification problems and regression problems, as its central uses. Basically, supervised learning requires a “labeled” dataset, which means one already knows a correspondence between two variables, or observed values (x, y), and one of them is interpreted as a label of another.

For example, let us assume there is a collection of handwritten numbers. We can tell which number it is by looking at it, but if we combine it with the image data as a label, we have

[1]Note that these discussions depend on the choice of activation functions. The paper which we cited assumes the sigmoid function for it, but our work will be based on mainly the SeLU (stands for Scaled Exponential Linear Unit) function [166] which is a generalization of the ReLU function (Rectified Linear Unit, which was firstly used in [167]). For the ReLU function, the universal approximation property is also proven [168].

the dataset. Then, the supervised learning architecture attempts to learn the correspondence. In terms of the neural networks, the matrices of linear maps, which are called "weights" are adjusted to predict the label data, if the given input data is the set of handwritten images in the example.

A simple regression problem is another example. Since data which is obtained by experiments contains noise intrinsically, one often guesses a relation between variables via regression. The simplest case is that there are only two variables (x, y), then finding a relation $y = f(x)$ is solved by the supervised learning, and it is generalized to higher dimensions.

In this concept, the performance of machine learning will be measured by its generalization ability. The ability is closely related to the notion of overfitting. Suppose that the learning reproduces the correspondences by 100% accuracy in the first example. Then, when fed with a new handwritten image, does the learned model correctly tell us the correct label? In cases where training data can be trained with high accuracy, extra details of the training data have been learned, often resulting in poor prediction performance for new data. Methods such as artificially adding noise to the data, terminating the learning process itself at an early stage, or cross-validation are possible workarounds. In regression problems, there are two approaches: maximum likelihood estimation and Bayesian estimation. The former method usually incorporates a regularization term, which adds a negative bias when the weights take somewhat large values (they often indicate overfitting), while the latter method assumes a prior probability distribution on the weights and restricts the range of weights.

Unsupervised learning

Learning with no valid labels is called unsupervised learning. The important areas of applications are dimensional reduction and clustering. We will perform unsupervised learning as a method for dimensional reduction and clustering in the following sections via an *autoencoder* model, so let us mention these frameworks briefly.

Dimensional reduction in this context is a procedure that reduces d.o.f. in given data by neglecting its unnecessary parts. Principal component analysis (PCA) which finds a basis where variations of data are effectively explained along its few axes is also included in this category as a typical classical method. The autoencoder[2], which is a model based on a neural network, may approach this kind of problem in a more sophisticated way. Their input data and output data are the same. Then, by configuring the dimensions of its inner layers to be small, the autoencoder attempts to extract the necessary features of given data. In principle, such features are accumulated in its inner layers. The dimensions of the innermost layer, which is called the latent layer, are usually made very small, such as two or three dimensions. An ideal autoencoder model which learns the needed features will exhibit the extracted features in its latent layer. If the dimensions of the latent layer are sufficiently small, one can even visualize the features as a distribution of output data from that layer. The distribution often forms a set of discrete islands, which are called clusters, and it explains the differences in the data (e.g., cats and dogs are classified in two different islands when the given data is a collection of their

[2] As using the input data as its label, it may classified into supervised learning, but the way the autoencoder attempts to learn in its latent layer is often classified into unsupervised learning. See for a review, [181]

images). Due to these properties of autoencoder models, the first part of an auto-encoder is called the encoder, the middle part is called the latent layer, and the latter part is called the decoder. The clustering method is also used to classify the islands not by humans, but by an automatic and not arbitrary way, which is based on some statistical and geometrical data e.g., densities of islands, distances between points, and averages over some regions. However, there is a large set of clustering methods, and they often have hyperparameters. Therefore, in the following sections, we will modify the results of the automatic clustering by hand, but we assume that the main points of discussion will not be lost.

Reinforcement learning

Reinforcement learning (abbreviated as RL, and see for a review, [182]) is completely different from the previous two methods in point that the data is not need to be prepared and fixed already, i.e., RL has the concept of "environment" and the process of interacting with the environment.

Suppose there is a set of whole "states" of a system, $\mathcal{S}$. On $\mathcal{S}$, a set $\mathcal{A}$ is defined as a collection of all possible "actions" A, where A acts on S and maps to $S' \in \mathcal{S}$. Then a transition matrix $T_{S'S} := \langle S' \,|\, A \,|\, S \rangle$ is a set of probabilities, where $\langle S' \,|\, A \,|\, S \rangle$ gives a probability of transition from a state S to a state S' via an action A. On the other hand, a random quantity called "reward" R_A which has an expectation value $\langle S' \,|\, R_A \,|\, S \rangle \in \mathbb{R}$ is also given after the action. Then let us consider a process consisting of sequential actions, states, and rewards labeled by time t. Basically, RL is formulated on the Markov decision process (MDP), and the aim of learning is finding a process which maximizes an expected value of total reward over the process.

RL is also employed in string theory, while the genetic algorithm which is first used in this context [169] is also similar to RL in that both are formulated on a process. Since RL is not performed below, further details are omitted.

3.3 Autoencoder-Driven Clustering of IIA Intersecting D-brane Models via Tadpole Charge

NOTE: This section (Section 3.3) is fully based on our results in [71]. In particular, the figures and tables exhibited in this section are the same as those in [71], and we simply cite them from the paper.

In this section, we present an application of a certain type of machine learning method, which is called autoencoder, in type IIA intersecting D-brane model. The section is organized as follows. In Section 3.3.1, we first give an introduction for this research and motivate the following analysis. Then in Section 3.3.2, we explain the explicit procedure to obtain the whole set of D-brane configurations, within a bound on the complex-structure moduli U. Further constraints which are not explained in Section 3.1 will be also given explicitly there. Specifically, we use concrete models in [183] which is cited in Section 3.3.2.3 to generate the dataset. Then in Section 3.4.1, we define the autoencoder models with some modifications and train them for the dataset in the following sections. Finally, we attempt to extract features hidden in the landscape from the latent layer of the autoencoder models and find a certain quantity that characterizes the observed clusters. We conclude this topic in Section 3.5.

3.3.1 Motivation

Extracting features of intersecting D-branes landscape
As explained, type IIA string theory provides us with a phenomenologically useful framework with intersecting D6-branes. One can directly construct a minimal supersymmetric Standard Model (MSSM)-like model or its generalization and many studies have been built phenomenological models (see for early works, e.g., [1,23,27,28]). Although these models are phenomenologically interesting, but we emphasize that there are usually large d.o.f. to configure D-branes. As the D6-branes wrap three-cycles of internal 6d space, one may wonder if the number of configurations is infinite. In fact, the number of configurations preserving $N = 1$ SUSY on the $T^6/\mathbb{Z}_2 \times \mathbb{Z}_2$ orientifold turns out to reach 134 billion [184]. It is fairly said a huge number. Nevertheless, the authors of [183] reported that for given concrete models, the probability of obtaining a three-family MSSM model among general configurations is only 10^{-9}, or one in a billion. Hence, it means we may obtain only $\mathcal{O}(10)$ three-family models on the vast landscape of intersecting D-brane models on the orientifold. Moreover, due to the Diophantine-type nature of constraints, it is hard to obtain those phenomenological configurations concretely. It naturally motivates a study based on the statistical approach (see, e.g., [185]) while brute-force scanning is also performed (see, e.g., [186]).

Machine-learning-based approaches have also been studied. So far, the reinforcement learning (RL) and the genetic algorithm mentioned in Section 3.2.1 are used to find a new configuration which is phenomenologically interesting in [187,188]. While the deterministic approach is computationally difficult, these machine-learning-based nonlinear approaches enable one to know more broader string landscape.

Although we are also interested in finding new configurations satisfying some models, we have another viewpoint for this string landscape at present. If the three-family configurations are rare in the landscape, we wonder which physical properties or physical quantities associated with the configurations, are distinct compared to other configurations. In fact, the number three is different from other integers, but the number of generations is a nonlinear function of the wrapping numbers of D-branes with many nonlinear constraints, which are explicitly shown later. Thus, controlling the number of generations by positioning D6-branes is very difficult and we wonder if there are other characterizations of the phenomenological models in terms of some physical quantities. Indeed, when it is compared to other family numbers, the three-family model has considerably smaller probabilities [183], which may suggest that there are hidden properties unique to the three-family model. In the first place, the same hidden structure may be suggested for configurations where the number of generations is well-defined. As shown later, it is highly nontrivial that the numbers of generations of quarks and leptons are aligned. Moreover, those of left-handed and right-handed fermions are not aligned in general. Characterizing the phenomenological configurations in terms of some physical quantities is interesting, and this is the goal of this research.

Indeed, there already exist statistical studies on this topic [183, 185]. The authors of [183] focused on the phenomenological models and studied distributions of some physical quantities. However, this study further considers non-phenomenological configurations in order to find the characteristics of phenomenological ones, and uses machine learning-based techniques rather than statistical methods. One problem with the statistical approach is that it is vague about which quantity should we focus on. In this study, we avoid this problem by using autoencoder models, which extract important features of input data into their latent layers. We define several concrete autoencoder architectures and feed them with the D-brane configurations, whether they are phenomenological or not. Then ideal autoencoder models should extract features which are necessary to reproduce from their small latent layers, and we will observe them as a set of discrete islands. Indeed, the autoencoder-based method is already used to visualize the clustering of type IIA intersecting D-brane models (but different models) [189, 190]. Our goal is to go beyond the visualization and find differences between those clusters in terms of more fundamental physical quantities[3]. Note that an autoencoder model whose latent layer has sufficiently small dimensions should focus on the most important features of given data. Thus, we can conclude that a physical quantity characterizes the data if we can explain differences between clusters by that quantity, and this is one of the big differences from statistical studies.

Limitations and possible workarounds

It is fair to note the possible disadvantages of this method at this level. We have to keep in mind these limitations and find workarounds for each limitation.

1. The three-family model configuration is not obtainable.
   Unfortunately, we could not find a three-family configuration for models in [183] in this study, as suggested by its extraordinarily small probability. Thus, our initial goal of characterizing the three-family configurations has already failed. On the other hand, a configuration where the numbers of generations of quarks and leptons are the same (we

[3]This methodology is utilized on different phenomenological corners of string theory, in particular heterotic string theory. See [191–194] for this kind of studies.

will call it "aligned") is already a nontrivial one, thus we decide to weaken the goal of this study. Our goal in this study is to characterize the aligned configurations among the whole configurations.

2. "Curse of dimensionality"
   This term refers to several nontrivial phenomena which occur due to the nature of higher-dimensional space. The mass distribution of a higher-dimensional sphere is an example, since its center is not a center of the mass distribution, i.e., almost all mass is concentrated in its outer shell. This nature of higher-dimensional space makes it difficult to quantify clustering. Thus, we restrict the dimensions of latent layers to be sufficiently small. As a result, only a few important features will be extracted.

3. Robustness
   The machine-learning-based studies usually possess a problem with the robustness of their results. In the first place, we can almost freely define an ML model. For neural networks, not only the architecture of nodes and layers but also their activation functions and hyperparameters relevant to training are chosen freely. Thus, it is natural to ask a question: are the results robust under changing the used ML model? Although it is true that we do not have to pursue this kind of robustness strictly, because other ML models may not be able to solve the given task. However, if clustering (in our case) is observed in outputs from several ML models, it is not natural to choose only one of them. The other robustness we have to check is, as explained later, the robustness under certain permutations. In general, physical data is obtained as a set, at least partially. Assume a particle moving under gravity. If we measure its 3d velocity over a certain time period and values of the gravitational field, the labels of the spatial directions must be interchanged freely. However, general neural networks will not learn these permutation symmetries due to their structure. As explained in more detail in the later section, we attempt to ensure the robustness via the so-called *positional encoding* technique [195].

4. Interpretability
   The interpretability is also a major difficulty of ML models. The ML models make decisions that turn out to be correct or useful, but why the ML models answered the solutions is highly nontrivial. This black-box nature of ML models also limits our study, because we have to analyze its latent layer. The clustering will occur in a mysterious way, thus it is difficult to find a feature between islands systematically. In this study, we simply take a brute-force approach.

Brief overview of our results in this study

As a result of our analysis, we find that clustering occurs, but it does not distinguish perfectly the aligned configurations from the others. Rather, the observed clusters gather both configurations in a nontrivial way. The rough clustering caused due to the small dimensions of the latent layer. However, it does not mean a failure of this study. Indeed, we will see that the tadpole charge of the hidden sector is a characteristic quantity in the landscape as the islands that contain aligned configurations have distinct distributions for the quantity. Furthermore, we observe the islands form a checkerboard-like pattern which suggests that the other quantity that characterizes the clustering exists.

3.3.2 Generating D-brane configurations

3.3.2.1 Model building constraints

This section is dedicated to an overview of constraints that should be satisfied by D-brane configurations. (See for a summary of these constraints, [183, 196].)

RR Tadpole Cancellation Condition
Due to the CS action, the orientifold compactification generates the tadpole cancellation condition:

$$\sum_a N_a \left([\Pi_a] + [\Pi'_a]\right) = 4\,[\Pi_{O6}]\,, \tag{3.5}$$

with N_a being the number of D6-branes. $a, b, \ldots$ label stacks of D-branes. The prime (') on $[\Pi'_a]$ represents the mirror cycle of $[\Pi_a]$ with respect to the orientifold action. Note that we do not consider any fluxes in the following.

As explained in Section 3.1, open strings which localize at an intersection point of two stacks of N_a, N_b D-branes leads states which transform in the bifundamental representation $(\mathbb{N}_a, \mathbb{N}_b)$. Together with other representations because of the existence of O6-planes, we summarize the possible spectrum in Table 3.3.1.

Representation	Multiplicity
$(\mathbf{N}_a, \overline{\mathbf{N}}_b)$	I_{ab}
$(\mathbf{N}_a, \mathbf{N}_b)$	$I_{ab'}$
Sym^2_a (symmetric)	$\frac{1}{2}(I_{aa'} - I_{aO6})$
Λ^2_a (anti-symmetric)	$\frac{1}{2}(I_{aa'} + I_{aO6})$

Table 3.3.1: Massless states and their multiplicities.

To calculate multiplicity, we need a geometric knowledge of an orientifold. As we consider an orientifold theory invariant under $\Omega\bar{\sigma}$ on a manifold $\mathcal{M}$ with the special geometry, we have the orientifold-even and -negative parts of $H_3(\mathcal{M})$. Let us denote the basis of the three cycles as $\alpha_I \in H_3^+(\mathcal{M})$ and $\beta_J \in H_3^-(\mathcal{M})$, as they satisfy $\alpha_I \circ \beta_J = \delta_{IJ}$ (otherwise zero). Here, $\circ$ denotes the intersections of the given two cycles. Then, we can expand three-cycles $[\Pi_a] \in H_3(\mathcal{M}, \mathbb{Z})$ and those of O6-planes as

$$[\Pi_a] = \vec{X}_a\vec{\alpha} + \vec{Y}_a\vec{\beta},\;\; [\Pi'_a] = \vec{X}_a\vec{\alpha} - \vec{Y}_a\vec{\beta},\;\; [\Pi_{O6}] = \frac{1}{2}\vec{L}\vec{\alpha}, \tag{3.6}$$

where we omit the indices as $\vec{X}_a\vec{\alpha} = \sum_{I=0}^{h^{2,1}} X_a^I \alpha_I$. Using this expansion, the intersection number I_{ab} is calculated as

$$I_{ab} = [\Pi_a] \circ [\Pi_b] = \vec{X}_a\vec{Y}_b - \vec{Y}_a\vec{X}_b. \tag{3.7}$$

SUSY conditions

The sLag cycles on a general manifold are given by [197]:

$$J|_{\Pi_a} = 0, \ \mathrm{Im}(e^{i\phi_a}\Omega_3)|_{\Pi_a} = 0, \tag{3.8}$$

where D6-branes need to wrap sLag cycles to preserve 4d $N = 1$ SUSY. The notation is the same as in previous sections. $X|_{\Pi_a}$ denotes that restriction on X to $[\Pi_a]$. The phase ϕ_a determines a preserved SUSY for each D-brane, and for the SUSY case it parameterizes the cycle's volume as $\int_{\Pi_a} \mathrm{Re}\left(e^{i\phi_a}\Omega_3\right)$. In the presence of O6-planes, $\phi_a = 0$ for all a corresponds to $N = 1$ SUSY. The sLag condition corresponds to the previous SUSY condition $\theta_1 + \theta_2 + \theta_3 = 0$ in a flat space, and thus it ensures that D6-branes wrap straight lines on each T^2.

Using the definition of the prepotential, $phi_a = 0$ leads

$$\vec{Y}_a \vec{F}(U) = 0, \qquad \vec{X}_a \vec{U} > 0, \tag{3.9}$$

where $\vec{F}$ denotes the vector of the derivatives F_I.

K-theory charge cancellation

There is an additional charge cancellation condition, which is called the K-theory condition which was found in [198]. Beyond the usual tadpole charge cancellation, a charge which is $\mathbb{Z}_2$-valued possessed by D6-branes, and it is fortunately known to be measured by inserting a probe D6-brane [199]. The probe D-brane is of $USp(2)$ D-brane which is on top of O6-planes in the toroidal orientifold which we will consider, and the K-theory condition ensures the cancellation of global $USp(2N)$ anomalies [200]. The explicit form of the condition is:

$$\sum_a N_a \left[\Pi_a\right] \circ \left[\Pi_{\text{probe } Sp(2)}\right] \equiv 0 \ (\mathrm{mod}\, 2), \tag{3.10}$$

where a runs over all stacks which exist in a system.

Gauge Anomalies and Massless U(1)

Apart from the gauge anomalies which are canceled via the tadpole cancellation condition, the Green-Schwartz mechanism ensures the absence of mixed (and cubic-) abelian gauge anomalies. In the cancellation, one finds that some $U(1)$ become massive by the Stueckelberg mechanism. The remaining massless $U(1)$ is given by [201]

$$U(1)_{\text{massless}} = \sum_a f_a U(1)_a \tag{3.11}$$

is given by

$$\sum_a f_a N_a \vec{Y}_a = 0. \tag{3.12}$$

3.3.2.2 The $T^6/(\mathbb{Z}_2 \times \mathbb{Z}_2')$ orientifold

Further calculations need the explicit structure of $H_3\mathcal{M}, \mathbb{Z}$. In this section, we outline the details of the cycle structure of the $T^6/(\mathbb{Z}_2 \times \mathbb{Z}_2')$ orientifold. However, due to its possible tilt and torsion d.o.f., the complete discussion is beyond the aim of this section. See for detailed discussions, [162]. Our situation corresponds to the without case (where "without" means "with" depending on literature) $\eta = 1$ case in the reference.

Orbifold Geometry

Basically, the geometry of the orientifold is inherited from that of the underlying orbifold. On the $T^6/(\mathbb{Z}_2 \times \mathbb{Z}_2')$ orbifold, we have the following orbifold actions:

$$\theta : (z_1, z_2, z_3) \to (-z_1, -z_2, z_3), \qquad \omega : (z_1, z_2, z_3) \to (z_1, -z_2, -z_3), \tag{3.13}$$

where the action of the orientifold action $\Omega\mathcal{R}$ is given by

$$\Omega\mathcal{R} : (z_1, z_2, z_3) \to (\bar{z}_1, \bar{z}_2, \bar{z}_3). \tag{3.14}$$

Note that the cycle structure is equivalently described in terms of the cohomology structure due to the Poincarè duality. Let us define a basis on each complex plane of T^2. On the i-th torus, the complex coordinates are expanded as $z_i = x_i e_i^x + y_i e_i^y$ using the basis e_i^x, e_i^y and have a periodicity $z_i \sim z_i + k e_i^x + l e_i^y$ $(k, l \in \mathbb{Z})$. Note that, the basis e_i^x, e_i^y is not needed to be the orthogonal basis of $\mathbb{C}$, and indeed the relation is given by

$$e_i^x = 2\pi(R_i^x + i b_i R_i^y), \qquad e_i^y = i 2\pi R_i^y, \tag{3.15}$$

where R^x, R^y denotes the two directions on $\mathbb{C}$. Here, b_i denotes the d.o.f for tilting the i-th torus and takes values $b_i \in \{0, \frac{1}{2}\}$. When $b_i = 0$, the i-th torus is a usual rectangular one and in the actual training of autoencoder models, we only consider the rectangular tori. The assumption does not lose much generality because it was pointed out in [183] that the number of aligned configurations realized on the tilted tori is dominated by that on the non-tilted tori. In this section, we describe the formulation which is also valid when there are tilted tori.

The next task is to obtain the basis of $H_3^{\pm}$. On T^6, the basis takes the simple form hence we expect that the orbifold (without torsion) does so. However, due to the orbifold actions, we need to avoid overcounting of intersections and define cycles which exist on the orbifold precisely. More details on this kind of problem can be found in [162, 196, 202]. Let us first define $\{[\pi_{2i-1}], [\pi_{2i}]\}$ as the fundamental one-cycles on the i-th torus along with e_i^x, e_i^y. The basis satisfies the intersection relations of (α_I, β_J). Then, the integral coefficients which appear in the expansion are called the wrapping numbers (n, m). We assume that they are coprime by interpreting a common factor as the number of D6-branes on the cycle. The intersection number I_{ab} is given by these wrapping numbers. The resulting three-cycle is expanded as

$$\left[\Pi_a^T\right] = \bigotimes_{i=1}^{3} (n_i^a [\pi_{2i-1}] + m_i^a [\pi_{2i}]), \tag{3.16}$$

where the upper "T" denotes that the cycle is defined on the underlying tori. The correspondences with (α_I, β_J) and (X^I, Y^I) is given by

$$\begin{aligned} \alpha^0 &= [\pi_1][\pi_3][\pi_5], \qquad & \beta^i &= [\pi_{2i}][\pi_{2j-1}][\pi_{2k-1}], \\ \alpha^i &= [\pi_{2i-1}][\pi_{2j}][\pi_{2k}], & \beta^0 &= [\pi_2][\pi_4][\pi_6], \end{aligned} \tag{3.17}$$

and

$$\begin{aligned} X^0 &= n_1 n_2 n_3, \quad & Y^i &= m_i n_j n_k, \\ X^i &= n_i m_j m_k, & Y^0 &= m_1 m_2 m_3. \end{aligned} \tag{3.18}$$

For the tilted tori, we have another useful basis

$$[\tilde{\pi}_{2i-1}] = [\pi_{2i-1}] - b_i[\pi_{2i}], \tag{3.19}$$

as the definition is motivated by the action

$$\Omega\mathcal{R} : [\pi_{2i-1}] \to [\pi_{2i-1}] - 2b_i[\pi_{2i}], \qquad \Omega\mathcal{R} : [\pi_{2i}] \to -[\pi_{2i}]. \tag{3.20}$$

The new basis is $\{[\tilde{\pi}_{2i-1}], [\pi_{2i}]\}$ and reduces to the original one when $b_i = 0$. Then we define the new wrapping numbers $(n_i, \tilde{m}_i)$ with $\tilde{m}_i = m_i + b_i n_i$ as

$$\left[\Pi_a^T\right] = \bigotimes_{i=1}^{3} \left(n_i^a[\tilde{\pi}_{2i-1}] + \tilde{m}_i^a[\pi_{2i}]\right). \tag{3.21}$$

To obtain an expansion of the three-cycles with integer coefficients, we modify the definition of (α_I, β_J) as:

$$\begin{aligned} \tilde{\alpha}^0 &= [\tilde{\pi}'_1][\tilde{\pi}'_3][\tilde{\pi}'_5], & \tilde{\beta}^i &= -[\pi'_{2i}][\tilde{\pi}'_{2j-1}][\tilde{\pi}'_{2k-1}], \\ \tilde{\alpha}^i &= -[\tilde{\pi}'_{2i-1}][\pi'_{2j}][\pi'_{2k}], & \tilde{\beta}^0 &= [\pi'_2][\pi'_4][\pi'_6]. \end{aligned} \tag{3.22}$$

Here, we re-defined the one-cycles as $[\tilde{\pi}'_{2i-1}] = \hat{b}^{-\frac{1}{3}}[\tilde{\pi}_{2i-1}]$ and $[\pi'_{2i}] = \hat{b}^{-\frac{1}{3}}[\pi_{2i}]$. This basis satisfies

$$\begin{aligned} \left[\tilde{\pi}'_{2i-1}\right] \circ [\pi'_{2j}] = -[\pi'_{2j}] \circ \left[\tilde{\pi}'_{2i-1}\right] &= -\hat{b}^{-\frac{2}{3}}\delta_{ij}, \\ \tilde{\alpha}_I \circ \tilde{\beta}_J = -\tilde{\beta}_I \circ \tilde{\alpha}_J &= \hat{b}^{-2}\delta_{IJ}, \end{aligned} \tag{3.23}$$

instead of the usual symplectic intersection. Then we have the expansion of the three-cycle $[\Pi_a^T]$ with integer coefficients

$$\left[\Pi_a^T\right] = \hat{\vec{X}}_a \hat{\vec{\alpha}} + \hat{\vec{Y}}_a \hat{\vec{\beta}}, \tag{3.24}$$

where we defined $\hat{\vec{\alpha}} = (\tilde{\alpha}^0, \tilde{\alpha}^1, \tilde{\alpha}^2, \tilde{\alpha}^3)$, $\hat{b} := \prod_i (1 - b_i)^{-1}$, and

$$\begin{aligned} \hat{X}^0 &= \hat{b} n_1 n_2 n_3, & \hat{Y}^i &= -\hat{b}\tilde{m}_i n_j n_k, \\ \hat{X}^i &= -\hat{b} n_i \tilde{m}_j \tilde{m}_k, & \hat{Y}^0 &= \hat{b}\tilde{m}_1 \tilde{m}_2 \tilde{m}_3. \end{aligned}$$

Here, $\hat{b}$ counts the number (#) of tilted tori as $2^{\#}$.

Note that $(\hat{X}, \hat{Y})$ satisfies the same algebra with that of (X, Y):

$$\begin{aligned} &\hat{X}^I \hat{Y}^I = \hat{X}^J \hat{Y}^J \ (\forall I, J), \\ &\hat{X}^I \hat{X}^J = -\hat{Y}^K \hat{Y}^L, \qquad \hat{X}^L (\hat{Y}^L)^2 = -\hat{X}^I \hat{X}^J \hat{X}^K, \quad \hat{Y}^L (\hat{X}^L)^2 = -\hat{Y}^I \hat{Y}^J \hat{Y}^K, \end{aligned} \tag{3.25}$$

where (I, J, K, L) denotes possible permutations of $(0, 1, 2, 3)$. In the following, we denote it by $(I, J, K, L) = \overline{(0, 1, 2, 3)}$.

The intersection number I_{ab}^T is given by

$$\begin{aligned} I_{ab}^T &= \left[\Pi_a^T\right] \circ \left[\Pi_b^T\right] = \vec{X}_a \vec{Y}_b - \vec{Y}_a \vec{X}_b = \hat{b}^{-2} (\hat{\vec{X}}_a \hat{\vec{Y}}_b - \hat{\vec{Y}}_a \hat{\vec{X}}_b), \\ I_{aa'}^T &= -\vec{X}_a \vec{Y}_a - \vec{Y}_a \vec{X}_a = -\hat{b}^{-2} (\hat{\vec{X}}_a \hat{\vec{Y}}_a + \hat{\vec{Y}}_a \hat{\vec{X}}_a), \end{aligned} \tag{3.26}$$

where "T" again indicates that the numbers are measured on the underlying tori.

O6-planes

The cycles which are invariant under the orientifold actions correspond to O6-planes. We have the following four invariant cycles:

$$\begin{aligned} \Omega\mathcal{R} &: 2^3 [\tilde{\pi}_1][\tilde{\pi}_3][\tilde{\pi}_5], & \Omega\mathcal{R}\omega &= -2^{3-2(b_2+b_3)} [\tilde{\pi}_1][\pi_4][\pi_6], \\ \Omega\mathcal{R}\theta &: -2^{3-2(b_1+b_2)} [\pi_2][\pi_4][\tilde{\pi}_5], & \Omega\mathcal{R}\theta\omega &= -2^{3-2(b_1+b_2)} [\pi_2][\tilde{\pi}_3][\pi_6], \end{aligned} \tag{3.27}$$

where the subscripts indicate the corresponding orientifold action and the relative signs are introduced to explain the negative charge. Then the total three-cycle wrapped by the O6-plane $[\Pi_{O6}^T]$ is given by

$$\left[\Pi_{O6}^T\right] = 2^3 ([\tilde{\pi}_1][\tilde{\pi}_3][\tilde{\pi}_5] - 2^{-2(b_2+b_3)} [\tilde{\pi}_1][\pi_4][\pi_6] - 2^{-2(b_1+b_2)} [\pi_2][\pi_4][\tilde{\pi}_5] - 2^{-2(b_3+b_1)} [\pi_2][\tilde{\pi}_3][\pi_6]), \tag{3.28}$$

where 2^3 corresponds to the number of orientifold planes.

Thus we have

$$\hat{\vec{L}} = 16 \left(\hat{b}, \frac{1}{1 - b_i} \right). \tag{3.29}$$

in Equation (3.6). The intersection number I_{aO6}^T with $[\Pi_a^T]$ becomes

$$I_{aO_6}^T = -2^3 \hat{b}^{-2} \left(\hat{Y}_a^0 \hat{b} + \sum_i \hat{Y}_a^i \frac{1}{1 - b_i} \right). \tag{3.30}$$

D-branes and O-planes on the orbifold

Let us define three-cycles on the orbifold (not the orientifold). Under general orbifold actions, a three-cycle is mapped into another one and it is a part of the three-cycle on the orbifold. In other words, we define a three-cycle on an orbifold as a sum of orbits of the orbifold action.

For the $T^6/(\mathbb{Z}_2 \times \mathbb{Z}_2')$ case, the wrapping numbers do not change and the "bulk" three-cycle on the orbifold is given by

$$\left[\Pi_a^{\mathrm{B}}\right] = 4\left[\Pi_a^{\mathrm{T}}\right] = 4\bigotimes_{i=1}^{3}\left(n_i^a[\tilde{\pi}_{2i-1}] + \tilde{m}_i^a[\pi_{2i}]\right). \tag{3.31}$$

The subscript "B" denotes "bulk", and the cycle is defined on the orbifold. On the other hand, we should slightly modify the topological intersection operator $\circ$ to give the intersection number

$$I_{ab}^{\mathrm{B}} = I_a \circ_{\mathrm{B}} I_b := \frac{1}{4}\left(4\left[\Pi_a^{\mathrm{T}}\right]\right) \circ \left(4\left[\Pi_b^{\mathrm{T}}\right]\right) = 4I_{ab}^{\mathrm{T}}. \tag{3.32}$$

The $\circ_{\mathrm{B}}$ is defined on the orbifold and the extra factor $1/4$ is needed since $\circ$ counts four times for each single intersection.

However, Equation (3.32) implies the homology group is not the integral one, and there are additional shorter cycles on the orbifold. Those cycles are called the fractional cycles and are defined as

$$\left[\Pi_a^{\mathrm{F}}\right] = \frac{1}{2}\left[\Pi_a^{\mathrm{B}}\right] = 2\left[\Pi_a^{\mathrm{T}}\right]. \tag{3.33}$$

With these fractional cycles we obtain the correct homology group $H_3(T^6/(\mathbb{Z}_2 \times \mathbb{Z}_2'), \mathbb{Z})$. The intersection of the fractional cycles is given by $\left[\Pi_a^{\mathrm{F}}\right] \circ_B \left[\Pi_b^{\mathrm{F}}\right] = \left[\Pi_a^{\mathrm{T}}\right] \circ \left[\Pi_b^{\mathrm{T}}\right]$. These fractional cycles account for the most general cycles on the orbifold without discrete torsion (when there is torsion, additional "exceptional" cycles also appear).

The O6-planes are already defined as a summation of the fixed cycles, and we do not have to define their shorter cycles. Thus,

$$\left[\Pi_{O6}\right] = \frac{1}{4}\left[\Pi_{O6}^{\mathrm{B}}\right] = \left[\Pi_{O6}^{\mathrm{T}}\right]. \tag{3.34}$$

Intersection numbers on the orientifold

The intersection numbers are defined on the orbifold (which the subscript O below means), and it is convenient to summarize the result of intersection numbers:

$$\begin{aligned}
I_{ab}^{\mathrm{O}} &= \left[\Pi_a^{\mathrm{F}}\right] \circ_B \left[\Pi_b^{\mathrm{F}}\right] = \frac{1}{4}I_{ab}^{\mathrm{B}} = I_{ab}^{\mathrm{T}} = \vec{X}_a\vec{Y}_b - \vec{Y}_a\vec{X}_b = \hat{b}^{-2}(\hat{\vec{X}}_a\hat{\vec{Y}}_b - \hat{\vec{Y}}_a\hat{\vec{X}}_b), \\
I_{aa'}^{\mathrm{O}} &= -\hat{b}^{-2}(\hat{\vec{X}}_a\hat{\vec{Y}}_a + \hat{\vec{Y}}_a\hat{\vec{X}}_a) = -2\hat{b}^{-2}\hat{\vec{X}}_a\hat{\vec{Y}}_a, \\
I_{aO6}^{\mathrm{O}} &= \frac{1}{2}I_{aO6}^{\mathrm{T}} = -2^2\hat{b}^{-2}\left(\hat{Y}_a^0\hat{b} + \sum_i \hat{Y}_a^i\frac{1}{1-b_i}\right).
\end{aligned} \tag{3.35}$$

In the following, we simply note the explicit forms of the constraints to generate the D-brane configurations. All results are consistent with the previous literature, in particular, [183].

Tadpole cancellation condition
The fundamental d.o.f is the fractional cycles. We obtain

$$\sum_a N_a \hat{X}_a^I = \frac{1}{2}\hat{L}^I, \tag{3.36}$$

or equivalently,

$$\begin{aligned} \sum_a N_a \hat{X}_a^0 &= 8\hat{b}, \\ \sum_a N_a \hat{X}_a^i &= \frac{8}{1-b_i}. \end{aligned} \tag{3.37}$$

SUSY conditions
The sLag cycle conditions are interpreted as

$$\sum_I \hat{Y}^I \frac{1}{U^I} = 0, \qquad \sum_I \hat{X}^I U_I > 0, \tag{3.38}$$

via the period integrals. Here, the complex structure moduli are defined as

$$U^0 = R_1^x R_2^x R_3^x, \tag{3.39}$$
$$U^1 = (1-b_2)(1-b_3) R_1^x R_2^y R_3^y, \tag{3.40}$$
$$U^2 = (1-b_1)(1-b_3) R_1^y R_2^x R_3^y, \tag{3.41}$$
$$U^3 = (1-b_1)(1-b_2) R_1^y R_2^y R_3^x. \tag{3.42}$$

K-theory conditions
The probe $USp(2)$ D6-branes are coincident with the O6-planes [162]. Thus invariant cycles under each of $\Omega\mathcal{R}\theta^k\omega^l$ $k,l \in \{0,1\}$[4] are given by Equation (3.27):

$$\begin{aligned} \left[\Pi^{\mathrm{T}}_{\Omega\mathcal{R}}\right] &= \hat{b}\,[\tilde{\pi}_1]\,[\tilde{\pi}_3]\,[\tilde{\pi}_5]\,, \\ \left[\Pi^{\mathrm{T}}_{\Omega\mathcal{R}\omega}\right] &= -\frac{1}{1-b_1}\,[\tilde{\pi}_1]\,[\pi_4]\,[\pi_6]\,, \\ \left[\Pi^{\mathrm{T}}_{\Omega\mathcal{R}\theta\omega}\right] &= -\frac{1}{1-b_2}\,[\pi_2]\,[\tilde{\pi}_3]\,[\pi_6]\,, \\ \left[\Pi^{\mathrm{T}}_{\Omega\mathcal{R}\theta}\right] &= -\frac{1}{1-b_3}\,[\pi_2]\,[\pi_4]\,[\tilde{\pi}_5]\,, \end{aligned} \tag{3.43}$$

[4] As mentioned, the normalization is defined to obtain $USp(2N)$ group by N D6-branes wrapping these cycles

where expressed in terms of wrapping numbers,

$$\begin{aligned}
\left[\Pi^{\mathrm{T}}_{\Omega\mathcal{R}}\right] &\leftrightarrow \hat{X}^0 = \hat{b}^2, \hat{X}^{1,2,3} = 0, \\
\left[\Pi^{\mathrm{T}}_{\Omega\mathcal{R}\omega}\right] &\leftrightarrow \hat{X}^1 = \frac{\hat{b}}{1-b_1}, \hat{X}^{0,2,3} = 0, \\
\left[\Pi^{\mathrm{T}}_{\Omega\mathcal{R}\theta\omega}\right] &\leftrightarrow \hat{X}^2 = \frac{\hat{b}}{1-b_2}, \hat{X}^{0,1,3} = 0, \\
\left[\Pi^{\mathrm{T}}_{\Omega\mathcal{R}\theta}\right] &\leftrightarrow \hat{X}^3 = \frac{\hat{b}}{1-b_3}, \hat{X}^{0,1,2} = 0,
\end{aligned} \tag{3.44}$$

and $Y^I = 0$ $({}^\forall I)$.

The prefactors are introduced to define closed cycles for general b_i. Note that, the fractional cycles and the torus cycles of the $USp(2)$ branes are the same since they are invariant under the orbifold actions.

Thus, the K-theory constraints is given by

$$\sum_a N_a \hat{Y}^0_a \equiv 0, \tag{3.45}$$

$$(1-b_j)(1-b_k)\sum_a N_a \hat{Y}^i_a \equiv 0, \tag{3.46}$$

where (i,j,k) is a cyclic permutation of $(1,2,3)$ and a runs over the stacks (USp stacks do not contribute to the condition).

Coprime condition

The coprime condition on (n,m) is implies the following condition on (X,Y):

$$(Y^0)^2 = \prod_{i=1}^{3} \gcd(X^i, Y^0) \tag{3.47}$$

for non-vanishing $\hat{X}^I$.

3.3.2.3 Concrete MSSM-like models on $T^6/(\mathbb{Z}_2 \times \mathbb{Z}'_2)$

In this section, we simply cite the MSSM-like models in Ref. [183] to generate the configurations. The authors of [183] constructed four models, but we cite only three of them because we could generate the aligned ($\mathrm{gen}(Q) = \mathrm{gen}(L)$) configurations for only those models.

- Model 1: $U(3)_a \times USp(2)_b \times U(1)_c \times U(1)_d$ in the visible sector with a hypercharge Q_Y^S

Matter	Representation	Multiplicity
Q_L	$(\mathbf{3}, \overline{\mathbf{2}})_{0,0}$	I_{ab}
u_R	$(\overline{\mathbf{3}}, 1)_{-1,0} + (\overline{\mathbf{3}}, 1)_{0,-1}$	$I_{a'c} + I_{a'd}$
d_R	$(\overline{\mathbf{3}}, 1)_{1,0} + (\overline{\mathbf{3}}, 1)_{0,1}$	$I_{a'c'} + I_{a'd'}$
d_R	$\left(\overline{\mathbf{3}}_A, 1\right)_{0,0}$	$\frac{1}{2}\left(I_{aa'} + I_{aO6}\right)$
L	$(1, \mathbf{2})_{-1,0} + (1, \mathbf{2})_{0,-1}$	$I_{bc} + I_{bd}$
e_R	$(\mathbf{1}, \mathbf{1})_{2,0}$	$\frac{1}{2}\left(I_{cc'} - I_{cO6}\right)$
e_R	$(\mathbf{1}, \mathbf{1})_{0,2}$	$\frac{1}{2}\left(I_{dd'} - I_{dO6}\right)$
e_R	$(\mathbf{1}, \mathbf{1})_{1,1}$	$I_{cd'}$

Here, the formulae which give generations are

$$
\begin{aligned}
\mathrm{gen}(Q_L) &= I_{ab}, \\
\mathrm{gen}(Q_R) &= I_{a'c} + I_{a'd} = I_{a'c'} + I_{a'd'} + \frac{1}{2}(I_{aa'} + I_{aO6}), \\
\mathrm{gen}(L_L) &= I_{bc} + I_{bd}, \\
\mathrm{gen}(L_R) &= \frac{1}{2}(I_{cc'} - I_{cO6}) + \frac{1}{2}\left(I_{dd'} - I_{dO6}\right) + I_{cd'}.
\end{aligned}
\tag{3.48}
$$

- Model 2: $U(3)_a \times U(2)_b \times U(1)_c \times U(1)_d$ in the visible sector with a hypercharge Q_Y^S

Q_L	$(\mathbf{3}, \overline{\mathbf{2}})_{0,0}$	I_{ab}
Q_L	$(\mathbf{3}, \mathbf{2})_{0,0}$	$I_{ab'}$
u_R	$(\overline{\mathbf{3}}, 1)_{-1,0} + (\overline{\mathbf{3}}, 1)_{0,-1}$	$I_{a'c} + I_{a'd}$
d_R	$(\overline{\mathbf{3}}, 1)_{1,0} + (\overline{\mathbf{3}}, 1)_{0,1}$	$I_{a'c'} + I_{a'd'}$
d_R	$\left(\overline{\mathbf{3}}_A, 1\right)_{0,0}$	$\frac{1}{2}\left(I_{aa'} + I_{aO6}\right)$
L	$(1, \mathbf{2})_{-1,0} + (1, \mathbf{2})_{0,-1}$	$I_{bc} + I_{bd}$
L	$(1, \overline{\mathbf{2}})_{-1,0} + (1, \overline{\mathbf{2}})_{0,-1}$	$I_{b'c} + I_{b'd}$
e_R	$(\mathbf{1}, \mathbf{1})_{2,0}$	$\frac{1}{2}\left(I_{cc'} - I_{cO6}\right)$
e_R	$(\mathbf{1}, \mathbf{1})_{0,2}$	$\frac{1}{2}\left(I_{dd'} - I_{dO6}\right)$
e_R	$(\mathbf{1}, \mathbf{1})_{1,1}$	$I_{cd'}$

Here, the formulae that give generations are

$$
\begin{aligned}
\mathrm{gen}(Q_L) &= I_{ab} + I_{ab'}, \\
\mathrm{gen}(Q_R) &= I_{a'c} + I_{a'd} = I_{a'c'} + I_{a'd'} = \frac{1}{2}(I_{aa'} + I_{aO6}), \\
\mathrm{gen}(L_L) &= (I_{bc} + I_{bd}) + (I_{b'c} + I_{b'd}), \\
\mathrm{gen}(L_R) &= \frac{1}{2}(I_{cc'} - I_{cO6}) + \frac{1}{2}\left(I_{dd'} - I_{dO6}\right) + I_{cd'}.
\end{aligned}
\tag{3.49}
$$

- Model 3: $U(3)_a \times U(2)_b \times U(1)_c \times U(1)_d$ in the visible sector with a different hypercharge $Q_Y^{(1)}$

Q_L	$(\mathbf{3}, \overline{\mathbf{2}})_{0,0}$	I_{ab}
u_R	$(\overline{\mathbf{3}}_A, 1)_{0,0}$	$\frac{1}{2}(I_{aa'} + I_{aO6})$
d_R	$(\overline{\mathbf{3}}, 1)_{-1,0} + (\overline{\mathbf{3}}, 1)_{0,-1}$	$I_{a'c} + I_{a'd}$
d_R	$(\overline{\mathbf{3}}, 1)_{1,0} + (\overline{\mathbf{3}}, 1)_{0,1}$	$I_{a'c'} + I_{a'd'}$
L	$(1, \mathbf{2})_{-1,0} + (1, \mathbf{2})_{0,-1}$	$I_{bc} + I_{bd}$
L	$(1, \overline{\mathbf{2}})_{-1,0} + (1, \overline{\mathbf{2}})_{0,-1}$	$I_{bc'} + I_{bd'}$
e_R	$(\mathbf{1}, \overline{\mathbf{1}}_A)_{0,0}$	$-\frac{1}{2}(I_{bb'} + I_{bO6})$

Here, the formulae that give generations are

$$\begin{aligned}
\mathrm{gen}(Q_L) &= I_{ab}, \\
\mathrm{gen}(Q_R) &= \frac{1}{2}(I_{aa'} + I_{aO6}) = (I_{a'c} + I_{a'd}) + (I_{a'c'} + I_{a'd'}), \\
\mathrm{gen}(L_L) &= (I_{bc} + I_{bd}) + (I_{bc'} + I_{bd'}), \\
\mathrm{gen}(L_R) &= -\frac{1}{2}(I_{bb'} - I_{bO6}).
\end{aligned} \tag{3.50}$$

For each model, the $U(1)$ hypercharge are defined as

$$Q_Y^S = \frac{1}{6}Q_a + \frac{1}{2}Q_c + \frac{1}{2}Q_d, \tag{3.51}$$

$$Q_Y^{(1)} = -\frac{1}{3}Q_a - \frac{1}{2}Q_b. \tag{3.52}$$

Correspondingly, the massless $U(1)_Y$ conditions are given by

$$\hat{Y}_a + \hat{Y}_b + \hat{Y}_d = 0, \tag{3.53}$$

$$\hat{Y}_a + \hat{Y}_b = 0. \tag{3.54}$$

3.3.2.4 Method

Following the lines of [183,203], we generate the D-brane configurations. As shown in [203], the number of configurations is finite for given complex structure moduli.

Tadpole Cancellation and SUSY

The tadpole cancellation conditions and SUSY conditions are summarized as

$$\sum_a S_a = C, \tag{3.55}$$

$$\sum_I \frac{1}{2}\hat{L}^I U_I \geq \sum_I \hat{X}_a^I U_I > 0 \ (\forall a), \tag{3.56}$$

$$\sum_I \hat{Y_a}^I \frac{1}{U_I} = 0 \ (\forall a), \tag{3.57}$$

where C and S_a are defined as

$$C := \sum_I \frac{1}{2}\hat{L}^I U_I, \tag{3.58}$$

$$S_a := \sum_I N_a \hat{X}_a^I U_I. \tag{3.59}$$

Generally speaking, to obtain the possible configurations we need to find integer partitions for given C. Then we need bounds on $\hat{X}^I$ for given U_I, and generate all allowed X^I in the range. For the set of possible $\{\hat{X}_a^I\}$, we find allowed $\{N_a\}$ which satisfies the integer partition Equation (3.55). Furthermore, without loss of generality, we can assume that all U_I are integers from Equation (3.57).

Exact expressions of $\hat{X}$ and $\hat{Y}$

Due to the algebra in Equation (3.25), possible $(\hat{X}, \hat{Y})$ are further constrained. For the above bounds on $\hat{X}^I$, it is related to the finiteness proof in [203], we review the proof in the next paragraph.

One can see that the algebra and the SUSY condition allow $\hat{X}_I$ only in the following categories (for each a):

- Nonzero $\hat{X}^I$
  In this case, we have $\hat{Y}$ via the algebra as

$$(\hat{Y}^L)^2 = -\frac{\hat{X}^I\hat{X}^J\hat{X}^K}{\hat{X}^L}, \tag{3.60}$$

  for $\hat{X}^I, \hat{X}^J = 0$ for $(I,J,K,L) = \overline{(0,1,2,3)}$. Hence there must be odd numbers of negative $\hat{X}$ and $\hat{Y}$s do not vanish. Note Equation (3.60) does not constrain the relative signs among $\hat{Y}$. They are fixed by the algebraic relation

$$\hat{Y}^i = \frac{\hat{X}^0}{\hat{X}^i}\hat{Y}^0, \tag{3.61}$$

  and we put

$$\hat{Y}^0 := \sqrt{-\frac{\hat{X}^I\hat{X}^J\hat{X}^K}{\hat{X}^L}}. \tag{3.62}$$

- Two $\hat{X}^I$s vanishing
  Suppose that $\hat{X}^I, \hat{X}^J = 0$ for $(I,J,K,L) = \overline{(0,1,2,3)}$. Then $\hat{Y}^I, \hat{Y}^J \neq 0$ is implied by the algebraic relation

$$\hat{X}^K\hat{X}^L = -\hat{Y}^I\hat{Y}^J. \tag{3.63}$$

Contrary to the pairs of vanishing $\hat{X}^I$, those of nonzero $\hat{X}^I$ vanish due to the relation

$$\hat{X}^K\hat{Y}^K = \hat{X}^I\hat{Y}^I = 0. \tag{3.64}$$

In this case, the SUSY condition becomes

$$\hat{Y}^I U_J + \hat{Y}^J U_I = 0 \text{ (no sum.)} \tag{3.65}$$

Then we obtain

$$\begin{aligned} \hat{Y}^I &= \pm\sqrt{\frac{U_I}{U_J}}\sqrt{\hat{X}^K\hat{X}^L}, \\ \hat{Y}^J &= \mp\sqrt{\frac{U_J}{U_I}}\sqrt{\hat{X}^K\hat{X}^L}. \end{aligned} \tag{3.66}$$

Although we obtained the solutions of the algebra, the intrinsic wrapping numbers impose the coprime condition on $(\hat{X}, \hat{Y})$. In Equation (3.47), we assumed nonzero $\hat{X}^I$ but that is modified in this case. For general b_i it becomes tedious because we need to express everything in terms of X^Is (with no hat). In the following computer search, we assume non-tilted tori, hence we fix all b_i to be zero. Then the coprime condition (Equation (3.47)) is divided into the following two cases:

- $X^0 = 0$ and $X^i = 0$ case
  In this case, n_i vanishes, and m_i is fixed to be one since we impose the coprime condition to separate wrapping numbers and the number of D-branes on the stack appropriately. Then we obtain

$$\gcd(X^j, Y^0)\gcd(X^k, Y^0) = |Y^0|. \tag{3.67}$$

- $X^i = 0$ and $X^j = 0$ case
  We have $(m_k, n_k) = (0, 1)$ for $k \neq i, j$. Then we obtain

$$\gcd(X^0, Y^i)\gcd(X^0, Y^j) = |X^0|. \tag{3.68}$$

Here, we defined gcd to be positive. These conditions turn out to be highly restrictive in generating the configurations.

- Three $\hat{X}^I$s vanishing
  The D6-branes with three zero $\hat{X}^I$s are $USp(2)$ branes. Thus we already know the wrapping numbers in Equation (3.43).

Finiteness of brane configurations
Following the lines of [203], we review the proof of the finiteness[5]. In the previous paragraph, we observed there are three types of possible $(\hat{X}^I, \hat{Y}^I)$. For each of the cases except for the $USp(2N)$ case, the finiteness is shown as follows:

[5] In fact, it was shown in [185] that there are finite configurations which are physically-distinct for all U^Is, but it is sufficient to follow the lines of [203] for our purpose.

- Nonzero $\hat{X}^I$
  The odd numbers of negative $\hat{X}^I$s imply that we can assume $(\hat{X}^A)(\hat{X}^{I\neq A}) < 0$ for some A. The SUSY condition 3.57 implies that we have
$$0 = \frac{\hat{Y}^A}{U_A}\left(1 + \sum_{I\neq A}\frac{U_A\hat{Y}^I}{U_I\hat{Y}^A}\right) = \frac{\hat{Y}^A}{U_A}\left(1 + \sum_{I\neq A}\frac{U_A\hat{Y}^I}{U_I\hat{Y}^A}\right) = \frac{\hat{Y}^A}{U_A}\left(1 + \sum_{I\neq A}\frac{U_A\hat{X}^A}{U_I\hat{X}^I}\right). \quad (3.69)$$
  From the most right-hand side, we obtain
$$|U_A\hat{X}^A|/|U_I\hat{X}^I| < 1 \quad (\text{no sum}, \forall I). \quad (3.70)$$
  Then the SUSY condition on $\hat{X}^I$ (3.56) allows only the $\hat{X}^A < 0$ case. Due to the tadpole cancellation condition,
$$0 < \hat{X}^{I\neq A}U_I \leq \frac{1}{2}\sum_J \hat{L}^J U_J \quad (3.71)$$
  follows for arbitrary I (no sum). Thus we can conclude that
$$1 \leq \hat{X}^{I\neq A} \leq \frac{1}{2}\sum_J \hat{L}^J\frac{U_J}{U_I}, \quad (3.72)$$
  i.e., there are finite configurations. We use this range to generate $\hat{X}^I$.

- Two $\hat{X}^I$s vanishing
  Let $\hat{X}^{K,L}$ to be zero. Then, the SUSY condition and the algebra $\hat{X}^I\hat{X}^J = -\hat{Y}^K\hat{Y}^L$ lead $\hat{Y}^K\hat{Y}^L < 0$. Hence, it follows that
$$\hat{X}^I > 0, \hat{X}^J > 0, \quad (3.73)$$
  and the discussion becomes the same as that for nonzero $\hat{X}^I$.

3.3.2.5 The Dataset

As mentioned, the SUSY condition (3.57) implies that it is sufficient to consider integral complex-structure moduli. Due to technical limitations, we further restrict us to the region $U^0 = 1, |U| := \sqrt{(U^1)^2 + (U^2)^2 + (U^3)^2} \leq 12$ with $U^1 \leq U^2 \leq U^3$, and $b_{1,2,3} = 0$. We obtain all possible configurations in this range by defining the visible sector for each concrete model. The total numbers of configurations are summarized in Table 3.3.2. Note that we ignore possible permutation symmetries in generating configurations.

Model	Model 1	Model 2	Model 3
Aligned/Total	1350/39551	1030/24336	816/171908

Table 3.3.2: The number of SUSY configurations which we obtained for each model. The algebra and massless $U(1)_Y$ conditions are taken into account. The left-side numbers represent the number of the aligned configurations, which satisfy gen(Q) = gen(L) whereas the right-side numbers do that of total configurations.

For later convenience, we note some technical details here.

- Zero-padding
  To fix autoencoder architectures, we need to fix the length of the data. Due to the actual architectures, the length is related to the maximum number of D-brane stacks in the dataset. It turns out that those numbers are 13, 11, and 12, for each model. We fill with zero the configurations whose length is less than the maximum.

- Normalization
  Although $(\hat{X}, \hat{Y})$ can be naturally normalized (since they correspond to the direction of single D-brane), we cannot normalize the number of D-branes N_a. Thus, we do not normalize the dataset in the following.

3.4 Feature Extraction with Autoencoder Models

In the previous sections, we obtained the datasets of D-brane configurations for each model. In this section, we define autoencoder models for each model in Section 3.4.1 and train them with the corresponding dataset. We architect the multiple models to ensure the robustness. As mentioned, we feed the autoencoder models with both the aligned configurations and the not-aligned ones and attempt to extract features in the configurations from their latent layers. The first result of training and the comparison of autoencoder models are given in Section 3.4.2. Section 3.4.3 is dedicated to the actual feature extractions for each model.

3.4.1 The Autoencoder Models and Additional Data Pre-processing

Model Architectures

In the following, we define three broad categories of models, which are named AE-0, 1, and 2. We also introduce positional encoding for the latter two models, as described later. The architectures for these models are summarized in Tables 3.4.1 to 3.4.3. All of the layers are dense layers and the notation in those tables is as follows. The symbols "E" and "D" represent the encoder part and the decoder part, respectively. The latent layer appears twice in each table, but they represent the same single layer. The dimensions of the input and output layers which are shown in the tables are for Model 1 ($USp(2)$ in the visible sector). For other models, they are changed appropriately depending on the maximum number of D-brane stacks (13, 11, and 12) for each model while the other layers remain unchanged. In particular, the latent layers are always defined as 2d dense layers. Since the typical length of each configuration is $\mathcal{O}(10^2)$, it is impossible to extract all possible features. Rather, as mentioned, the autoencoder models will extract a few important features, and clustering of them will occur in a nontrivial way to reproduce the input data efficiently. Note that, in each autoencoder model, there are many parameters that can be tuned, such as the dimensions of the layers and the activation functions. For the activation functions, we try several combinations of the activation functions for AE-2, as shown in Table 3.4.4. We mainly adopt the SeLU function [166] and tanh for the activation functions. One reason for employing these activation functions is to prevent saturation. The SeLU function is a generalization of the well-known ReLU function, as taking nonzero values also for negative input. The program is coded using `TensorFlow` and the layers are defined by `tensorflow.keras`.

Layer	Input	E1	E2	E3	E4	E5	Latent
Dimension	$\mathbb{Z}^{104}+\mathbb{Z}^{13}+\mathbb{Z}^{4}$	$\mathbb{R}^{88}$	$\mathbb{R}^{55}$	$\mathbb{R}^{22}$	$\mathbb{R}^{11}$	$\mathbb{R}^{6}$	$\mathbb{R}^{2}$
Layer	Latent	D1	D2	D3	D4	D5	Output
Dimension	$\mathbb{R}^{2}$	$\mathbb{R}^{6}$	$\mathbb{R}^{11}$	$\mathbb{R}^{22}$	$\mathbb{R}^{55}$	$\mathbb{R}^{88}$	$\mathbb{Z}^{121}$

Table 3.4.1: The architecture of AE-0. The dimensions of the input layer and the output layer are set at 121 since we have 13 sets of (X_a, Y_a), 13 N_as, and 4 moduli ($121 = 13 \times 8 + 13 + 4$). We adopt tanh for the latent layer and SeLU for the other layers as activation functions.

Layer	Input	Pre	E1	E2	E3	E4	E5	E6	Latent
Dimension	$\mathbb{Z}^{104}+\mathbb{Z}^{4}$	$\mathbb{R}^{108}$	$\mathbb{R}^{80}$	$\mathbb{R}^{48}$	$\mathbb{R}^{20}$	$\mathbb{R}^{10}$	$\mathbb{R}^{6}$	$\mathbb{R}^{4}$	$\mathbb{R}^{2}$
Layer	Latent	D1	D2	D3	D4	D5	D6	Output	
Dimension	$\mathbb{R}^{2}$	$\mathbb{R}^{2}+\mathbb{Z}^{4}$	$\mathbb{R}^{8}$	$\mathbb{R}^{16}$	$\mathbb{R}^{24}$	$\mathbb{R}^{48}$	$\mathbb{R}^{80}$	$\mathbb{Z}^{104}+\mathbb{Z}^{4}$	

Table 3.4.2: The architecture of AE-1 (without positional encoding). The dimensions of the input layer and the output layer are set at 108 since we have the $104 = 8 \times 13$ sets of (N_aX_a, N_aY_a) and four moduli. We adopt tanh for the latent layer and SeLU for the other layers as activation functions, except for the "Pre" layer. No activation function is adopted for Layer Pre. As hints of conditional autoencoder, the four complex-structure moduli are additionally input at D1 ($\mathbb{R}^2 + \mathbb{Z}^4$). If there is the positional encoding, all of the input and output elements become $\mathbb{R}$.

Layer	Input	Pre	E1	E2	E3	E4	E5	Latent	
Dimension	$13\mathbb{Z}^{12}$	$13\mathbb{R}^{4}$	$\mathbb{R}^{52}+\mathbb{Z}^{4}$	$\mathbb{R}^{40}$	$\mathbb{R}^{20}$	$\mathbb{R}^{10}$	$\mathbb{R}^{4}$	$\mathbb{R}^{2}$	
Layer	Latent	D1	D2	D3	D4	D5	D6	Post	Output
Dimension	$\mathbb{R}^{2}$	$\mathbb{R}^{2}+\mathbb{Z}^{4}$	$\mathbb{R}^{8}$	$\mathbb{R}^{16}$	$\mathbb{R}^{24}$	$\mathbb{R}^{36}$	$\mathbb{R}^{52}$	$13\mathbb{R}^{4}$	$13\mathbb{Z}^{8}$

Table 3.4.3: The architecture of AE-2 (without positional encoding). The input layer contains 13 sets of (N_aX_a, N_aY_a, U), while the outputs contains 13 sets of (N_aX_a, N_aY_a). In Layer Pre, each (N_aX_a, N_aY_a, U) is reduced to a four-dimensional vector $\mathbb{R}^4$. As hints, the four complex-structure moduli are given at E1 and D1. If there is the positional encoding, all of the input and output elements become $\mathbb{R}$.

Layer	E4	E5	Latent	D1	D2	Others
"tanh-1"	SeLU	SeLU	tanh	SeLU	SeLU	SeLU
"SeLU-1"	SeLU	SeLU	SeLU	SeLU	SeLU	SeLU
"tanh-2"	SeLU	tanh	tanh	tanh	SeLU	SeLU

Table 3.4.4: The three patterns of the activation functions for AE-2 which we compared.

Among the three models, one will see that the first one (AE-0) is the most simplest and traditional autoencoder. Unfortunately, we did not observe any clustering in the latent layer after a sufficiently long training session (set to 100 epochs). Since the purpose of our study depends on some clustering occurring, we decided to define other autoencoders to see if clustering would occur. This is the motivation for defining AE-1 and AE-2.

AE-1 and AE-2 are defined as a type of *conditional autoencoder* (but not variational) that is classified as a semi-supervised auto-encoder. Indeed, a conditional encoder is an autoencoder that is given a "hint of the correct answer" during the process of a neural network. For classification questions, the correct answer label should be given as a hint. The question is, in this case, what to choose as a hint. In this study, we choose the complex-structure moduli as the label, as a set of quantities that is universal to some extent and restricts the configurations non-trivially by a small number of variables. In AE-1, the hint is given at Layer D1, and in AE-2, it is given at Layers E1 and D1. Indeed, these models exhibit the desired clustering as shown in the following sections.

Embedding symmetries via positional encoding

As explained in Section 3.2.1, usually physical data has some permutation symmetries as its structure. This is true also for the case since the D-brane stacks (a) and the directions (I) can be freely interchanged while the autoencoder models do not know that fact and interpret the configurations as *sequences*, not *sets*. Hence, we have to make the autoencoder models learn the symmetries and there are three options for us to do so:

1. Design a neural network with a set as the input.
   In a naive way, a normal neural network cannot learn the properties of a set as a set. This point has been studied for many years, and a candidate for such neural network is proposed in [204]. However, our case is slightly more complex, since the permutation symmetry is not embedded globally in a configuration. Instead, we can interchange a D-brane stack as a single unit. When interchanging Is, we have to relabel for each direction. Thus, a configuration is not a set, but it possesses the property of a set partially. To the best of our knowledge, there is no universal framework where we can freely design neural networks for partial symmetries. We do not take this option in this study.

2. Feed the autoencoder model with permutated configurations.
   This is a more traditional way to make ML models learn such symmetries. However, as we mentioned there are typically $\mathcal{O}(10)$ number of D-brane stacks for each configuration. This would require about 10! configurations each, which is practically impossible.

3. Giving up the rigorous consistency under the symmetries.
   Then the final option that we can take is simply giving up embedding the symmetries. Unfortunately, in this study, we generally choose this final option.

Thus, it seems that we should give up embedding the symmetry of D-brane stacks. On the other hand, there is the other symmetry of Is as mentioned. We attempt to embed this symmetry via the *positional encoding*, which was originally invented in [195] in the context of Natural Language Processing (NLP). *Transformer*, the model proposed in the paper, aims to transform an input vector into an output vector to solve linguistic tasks such as translation. In such NLP models, the order of the input vector must be learned because it corresponds to a sentence. To embed the order efficiently, the authors of [195] proposed to add periodic values to the original elements of given data, whose phase corresponds to the position of the elements (see for a review, [205]). The periodic numerical values are defined to be obtained by a certain combination of trigonometric functions, respecting their linear property under a translation. Although the original use of the positional encoding method is to distinguish the different positions in the data, we consider if the same value is added to positions that are transformed together by symmetry, the model considers them to belong to the same class and distinguishes them from positions to which different common values are added. Unfortunately, we do not know any rigorous mathematical reasons to prove the efficiency of the positional encoding method. We simply try it, given the fact that Transformer successfully learns the positional information and is utilized in several services such as the famous *ChatGPT*.

To examine its effects, we define two ways of the positional encoding, PE-i/ii. For the PE-i, we also include the information of stacks but for PE-ii, we ignore the symmetry of D6-brane completely. The concrete values that are added to the original data are as follows:

- PE-i: respecting both of the symmetries of stacks a and I in each stack (AE-1/AE-2).

  For pos-th stack, add $A\sin\left(\frac{pos}{T^{2i/4}}\right)$ at position $I = 2i$ and $A\cos\left(\frac{pos}{T^{2i/4}}\right)$ at $I = 2i+1$. A is the amplitude and is set to one by default. PE-i is modified for each autoencoder model. For AE-1, those values with $pos = 0$ are added to U^Is. For AE-2, since it has $\hat{X}^I$, $\hat{Y}^I$ and U^I together in each stack, we add the same value as that for $\hat{X}^I, \hat{Y}^I$ to each U^I.

- PE-ii: respecting only the symmetry of I (AE-2)
  Since AE-2 has the pre-training layer for each stack, giving the information about stacks may not make sense. Hence we define PE-ii by ignoring the order of D-brane stacks by setting $pos := 1$ for all stacks. Later, we also vary the amplitude and compare the results.

In general, it is expected that positional encoding has both advantages and disadvantages. While it will make it easier for the autoencoder models to learn the information of the positions, the added values are just noise from the point of view of the algebra and the conditions. Thus, we have to decide whether to include the positional encoding or not by observing the result outputs of the latent layers.

3.4.2 Comparison of Methods (Model 1)

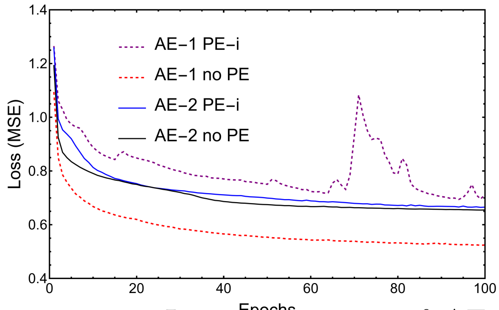

Figure 3.4.1: Loss in training of AE-1/2 with and without PE-i. [71]

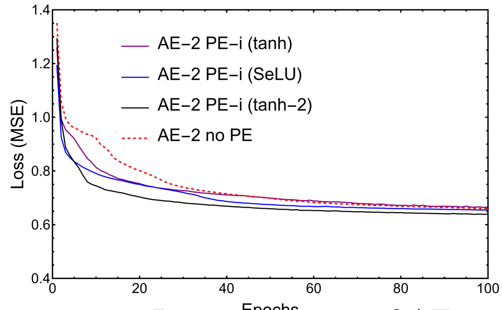

Figure 3.4.2: Loss in training of AE-2 with different activation functions (PE-i). [71]

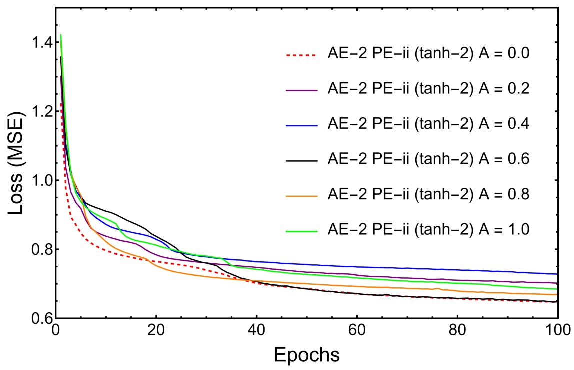

Figure 3.4.3: Loss in training of AE-2 with different amplitudes for PE-ii. [71]

In order to compare the various autoencoder models that have been defined, we trained them on the Model 1 data. From the results, we decide which autoencoder model to use from the perspective of loss and clustering for the following discussion.

Losses

We summarized the losses of training in Figures 3.4.1 to 3.4.3 [71]. Here, the mean squared error is chosen to be the loss function. From the absolute value of the loss in Figure 3.4.1, one might consider that AE-1 without positional encoding (PE) is the best model. On the other hand, the loss of AE-1 with PE-i is not stable during the training process. This is in contrast to AE-2 cases and implies that AE-2 is more robust than AE-1. In Figure 3.4.2, we vary the activation functions of AE-2. As a result, the tanh-2 model scored the least loss while the differences are relatively smaller than those between AE-1 without PE and AE-2. Thus, both AE-1 and AE-2 with tanh-2 models are candidates. We also illustrate the results of varying the amplitudes of PE-ii in Figure 3.4.3. We see that there are only small differences between the losses, but it is interesting that the larger amplitude does not generally lead to the larger loss. Since the differences are relatively small, we do not choose one amplitude to analyze.

It is also fair to say that these losses are large in general. Since the smallest nonzero element in the dataset is naturally one, the $\mathcal{O}(1)$ noises may spoil the input data. However, the actual data will be more complex and we are interested only in the most important features, we continue the discussion. Therefore, at this level, we were not able to pick up a few desirable autoencoder models.

Clusters in latent layers

Since our goal is feature extraction, whether we observe clusters or not is another important criterion. We illustrate the outputs from the latent layers in Figures 3.4.4 to 3.4.18 [71][6]. There, the red dots represent the 2d vectors corresponding to the aligned configurations and the blue ones do those of the non-aligned configurations. As one can see, the clusters appear in several models while it does not occur in other models. Since it is sufficient to choose the models where clusters are observed, we simply ignore the other models such as the saturated one, which is illustrated in Figure 3.4.15. Taking into account both losses and the clusters, we decided to choose the model AE-2 with no positional encoding in Figure 3.4.13 at first. Ideally, if clusters are observed in the several models, the results should be the same regardless of which one is selected. To ensure robustness, we will also analyze the outputs from AE-2 with PE-ii ($A = 1.0$) in Figure 3.4.18 later. As a result, the conclusion turns out to be invariant under the two choices.

[6]We omit the citation symbol in those captions, but they are all listed on [71]

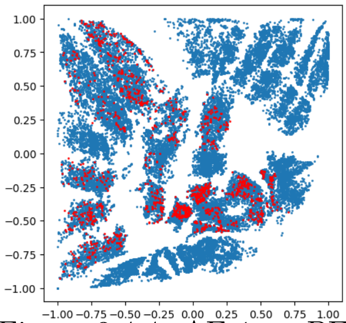

Figure 3.4.4: AE-1 no PE.

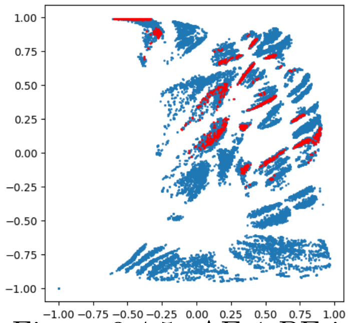

Figure 3.4.5: AE-1 PE-i.

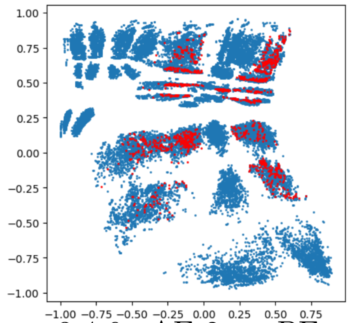

Figure 3.4.6: AE-2 no-PE tanh-1.

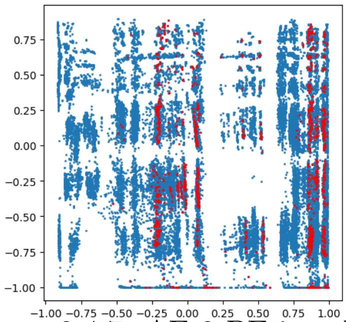

Figure 3.4.7: AE-2 PE-i tanh-1.

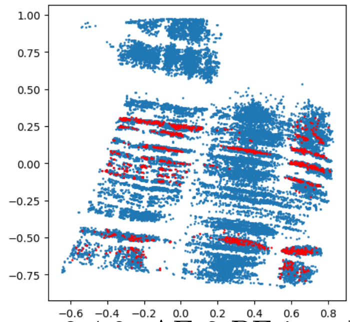

Figure 3.4.8: AE-2 PE-i tanh-2.

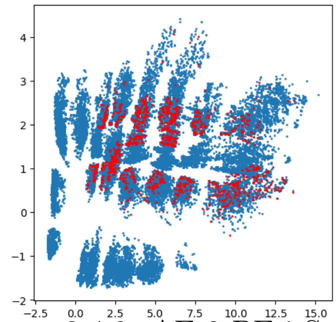

Figure 3.4.9: AE-2 PE-i SeLU.

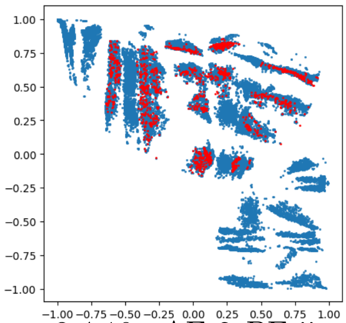

Figure 3.4.10: AE-2 PE-ii tanh-1.

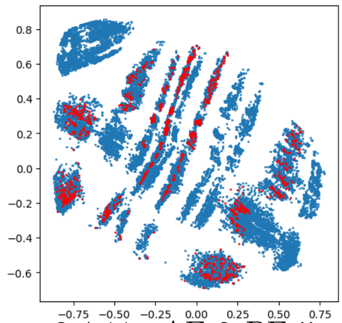

Figure 3.4.11: AE-2 PE-ii tanh-2.

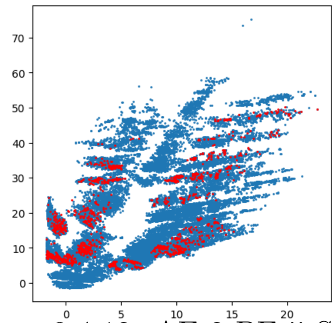

Figure 3.4.12: AE-2 PE-ii SeLU.

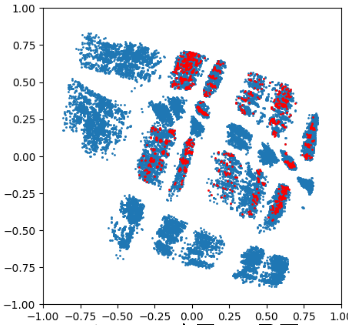

Figure 3.4.13: AE-2 PE-ii $A = 0.0$.

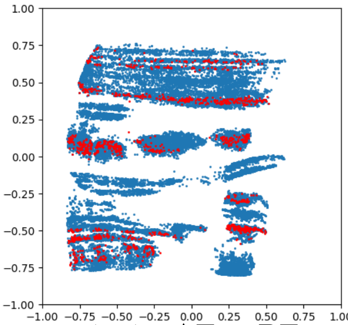

Figure 3.4.14: AE-2 PE-ii $A = 0.2$.

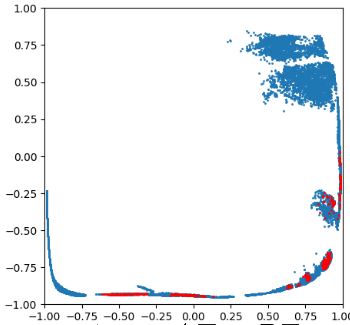

Figure 3.4.15: AE-2 PE-ii $A = 0.4$.

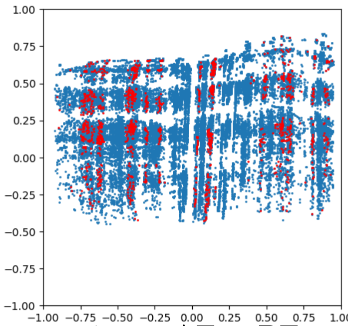

Figure 3.4.16: AE-2 PE-ii $A = 0.6$.

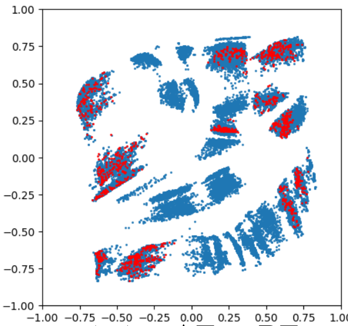

Figure 3.4.17: AE-2 PE-ii $A = 0.8$.

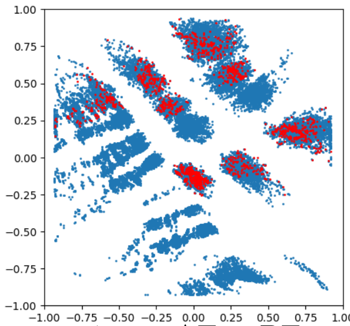

Figure 3.4.18: AE-2 PE-ii $A = 1.0$.

3.4.3 Feature Extraction

3.4.3.1 Model 1

Clustering and its modification

The first model which we consider is AE-2 (tanh-2) with no positional encoding where we observe the clusters in Figure 3.4.13. However, it is difficult to group them in a reasonable way because the size of the blank space between islands varies, as does the density of points on each island. One possible way to classify islands is to use existing clustering methods. We tried several methods which are built-in methods of `FindClusters` in `Mathematica v11.3` and summarize the results in Figures 3.4.19 to 3.4.21 [71]. One can see that the results highly depend on the adopted method. Indeed, those clustering methods usually have several hyperparameters, and we tuned them in Figures 3.4.19 to 3.4.21 to reproduce the islands which we observed with our eyes. It is fair to say that the clustering method does not provide a reasonable way to group islands in a strict sense. Thus, we will modify the result of the grouping later. In the following, we choose the `JarvisPatrick` method whose the radius parameter is tuned to 0.137 (Figure 3.4.20). The resultant number of clusters is 20. The numbers of configurations in each island are listed on Table 3.4.5.

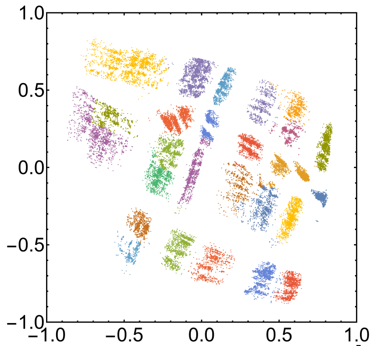

Figure 3.4.19: `MeanShift`. [71]

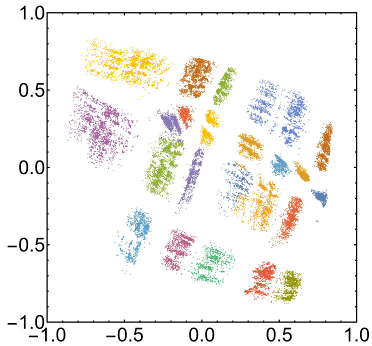

Figure 3.4.20: `JarvisPatrick` with radius 0.137. [71]

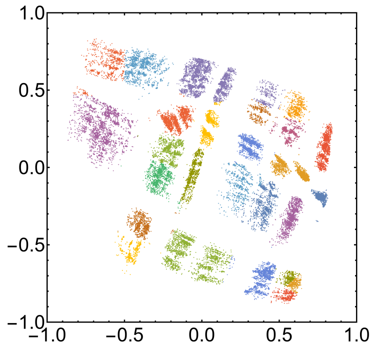

Figure 3.4.21: `KMeans` with 27 clusters. [71]

Then, we classify the clusters into two categories: containing or not containing the aligned configurations. The number of the configurations in each cluster is summarized in Table 3.4.5.

Cluster No.	1	2	3*	4*	5	6	7	8	9	10
Aligned/Total	0/1426	137/1449	137/3058	10/2544	297/3211	0/1649	0/2593	0/2745	0/1318	0/1529
Cluster No.	11	12*	13	14*	15	16	17	18	19	20
Aligned/Total	209/2384	86/2186	0/2202	11/1868	56/1344	134/1881	95/1351	75/1877	0/921	93/2015

Table 3.4.5: The numbers of the configurations in each cluster. The notation is the same with Table 3.3.2. Here, the symbol * is added to the clusters which we modify in the following.

However, if one looks closely, one notices that there are several instances in the inter-island regions where the grouping differs from the intuition. We separate the islands into two categories which contain the aligned configurations and do not contain the aligned ones. The categories are colorized and shown in Figure 3.4.22. Then, one can see that the borders between Clusters 3, and 4 and Clusters 12, and 14 are different from the intuition (Figure 3.4.23).

The differences caused by this grouping is at most 5.7% of a cluster, but we redefine the clustering artificially. Indeed, we see a checkerboard pattern in the latent layer and it provides us a natural way to redefine the clusters, as shown in Figure 3.4.24. This artificial clustering is justified because we have already tuned the hyperparameters of the clustering methods. The numbers of the configurations after the modification are summarized in Table 3.4.6.

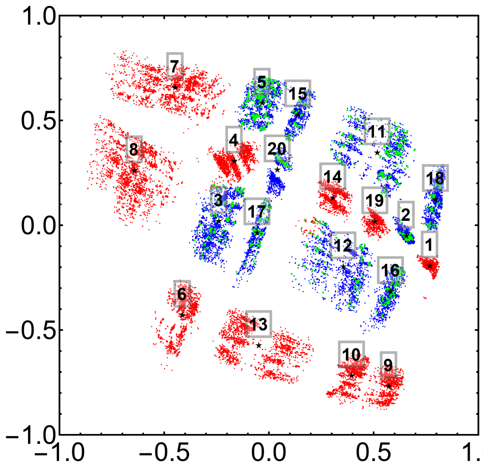

Figure 3.4.22: The two categories. The category containing the aligned configurations is colored blue, while the other category (which does not contain the aligned ones) is colored red. In particular, the aligned configurations are colored green. The integer labels represent the cluster numbers, which are listed on Table 3.4.5. Clusters 4 and 14 are colored red, as they do not contain the aligned ones after the modification. [71]

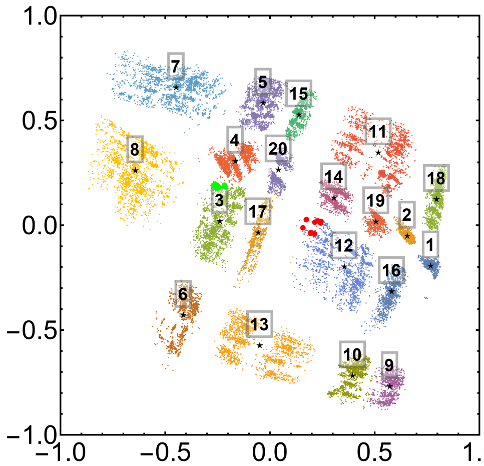

Figure 3.4.23: The outliers due to the clustering. Initially, the green dots were in Cluster 4, and the red dots were in Cluster 14. We consider them as outliers, and Clusters 4 and 14 are classified into the category not containing the aligned configurations. [71]

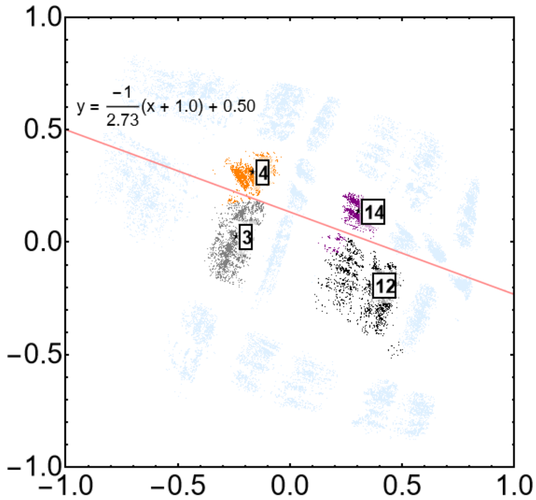

Figure 3.4.24: The modification by drawing a line. The equation of the line is shown in the figure, where x, y denote the coordinates for horizontal- and vertical-axes. Clusters 4, 14, 3, and 12 are colored orange, purple, gray, and black, respectively. We define the line as a new boarder and modify the clustering. [71]

Cluster No.	1	2	3*	4*	5	6	7	8	9	10
Aligned/Total	0/1426	137/1449	147/3140	0/2462	297/3211	297/1649	0/2593	0/2745	0/1318	0/1529
Cluster No.	11	12*	13	14*	15	16	17	18	19	20
Aligned/Total	209/2384	107/2292	0/2202	0/1762	56/1344	134/1881	95/1351	75/1877	0/921	93/2015

Table 3.4.6: The numbers of the configurations in each cluster after the modification.

Distributions of the hidden tadpole charges

In the following section, we will actually perform feature extraction, which is the purpose of this study. As explained, the autoencoder models that we defined have no interpretability, and thus we have to extract features in a brute-force way. Taking into account the quantities whose distributions were studied in [183], we analyzed the distributions of many quantities over the clusters. Those quantities include the number of hidden gauge groups, their total rank, net chirality [183], and the generations of quarks or leptons if they are aligned. Nevertheless, we do not observe characteristic distributions, except for one quantity, namely the tadpole charges of hidden sectors:

$$Q^I_{\text{hid}} \coloneqq \sum_{a:\text{hidden}} N_a \hat{X}^I_a = 8 - \sum_{a:\text{visible}} N_a \hat{X}^I_a, \quad \forall I. \tag{3.74}$$

In the following, we call them the hidden tadpole charges. Indeed, 455 different (sets of) hidden tadpole charges are found among the dataset, but 337 of them are found in the category containing the aligned configurations, and 206 of them are found in the category that does not contain the aligned ones. The number of duplicates is only 88, and the distributions of the hidden tadpole charges for each category are illustrated in Figure 3.4.25 [71].

There, the horizontal axis denotes the (set of) hidden tadpole charges. Then, the blue bar chart represents the number of configurations that are included in the category containing the aligned configurations, with given hidden tadpole charges. On the other hand, the red bar chart represents the number of configurations that are included in the other category, with given hidden tadpole charges. The green bar chart represents that of the aligned configurations (hence it is a part of the blue bar chart).

One can see that those blue and red charts have their peaks at the different points. It means that the autoencoder model weight on the hidden tadpole charges to reproduce the input data. Indeed, 69.4% and 73.4% of the configurations respectively in the two categories (blue and red ones) have unique hidden tadpole charges for their categories. Further, all of the aligned configurations have the hidden tadpole charges which are unique for them.

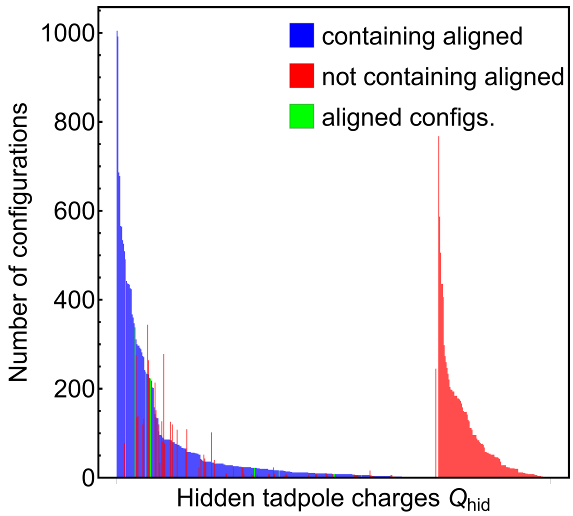

Figure 3.4.25: The distributions of the numbers of configurations for given hidden tadpole charge. The blue bar chart and the red bar chart correspond to the two categories (containing/not containing the aligned configurations), as shown in the figure. The green bar chart corresponds to the aligned configurations (hence it is a part of the blue bar chart). The horizontal axis represents the hidden tadpole charges. They are sorted in such a way that configurations in the first category have a higher number. [71]

As for the 88 hidden tadpole charges common to the two categories, we could not explain the overlap. However, the clustering pattern suggests that there is at least one other feature that is weighted by the autoencoder models as much as the hidden tadpole charges, as a checkerboard pattern appeared in the latent layer. This brute-force approach does not have enough power to find the other feature. Assuming the existence of another feature equally important to the hidden tadpole charge, we will conclude the discussion. We choose several hidden tadpole charges and illustrate their distributions in the latent layer in Figures 3.4.26 to 3.4.29 [71].

Note that, as observed in Figure 3.4.26 and Figure 3.4.27, it is likely that multiple straight belts will contain configurations with the same hidden tadpole charge. The feature is also seen in the following analysis of the different D-brane models, and it supports the existence of the checkerboard pattern.

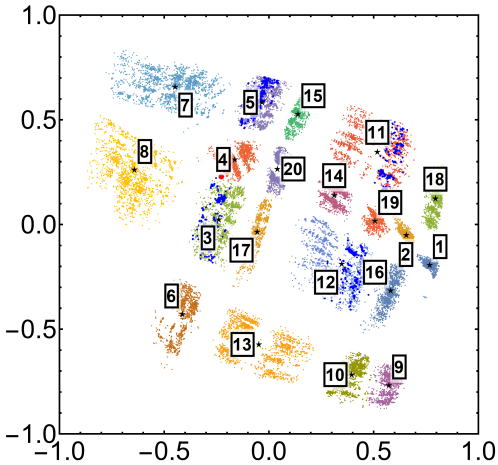

Figure 3.4.26: Hidden tadpole charge: $Q_{\mathrm{hid}} = (1, 7, 6, 5)$. All 1,005 configurations with the charge, except for one, are in the category containing aligned configurations (blue dots). [71]

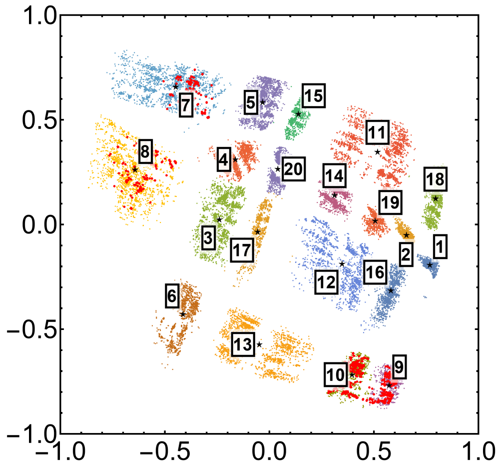

Figure 3.4.27: Hidden tadpole charge: $Q_{\mathrm{hid}} = (2, 6, 2, 7)$. All of the 768 configurations are in the same category which does not contain the aligned configurations (red dots). [71]

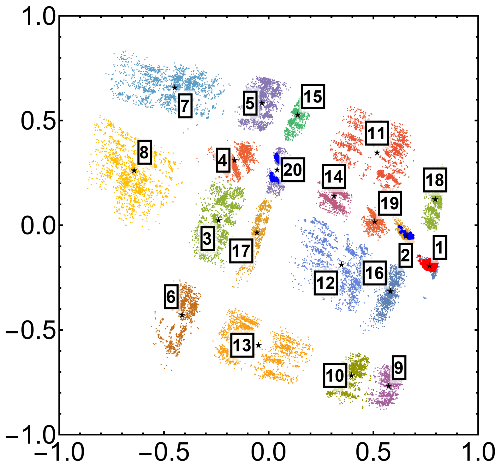

Figure 3.4.28: Hidden tadpole charge: $Q_{\mathrm{hid}} = (6, 7, 5, 1)$. The blue and red distributions are closely placed, but both categories have the same hidden tadpole charge. [71]

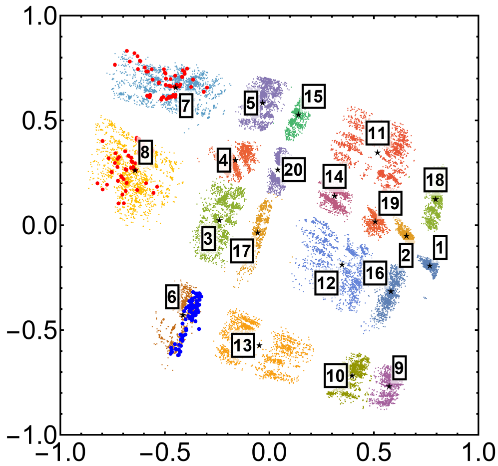

Figure 3.4.29: Hidden tadpole charge: $Q_{\mathrm{hid}} = (3, 2, 6, 6)$. The blue and red distributions appear independent. [71]

Robustness: other methods

In this paragraph, we consider the same Model 1 via the other method, AE-2 (tanh-2) with PE-ii positional encoding (Figure 3.4.18) so that we ensure the robustness of our result. The amplitude was chosen to be 1, i.e., the largest value among those employed.

Again, in this case, we observe the clear clusters and the checkerboard pattern. We again classify them into two groups and the results are illustrated in Figures 3.4.30 and 3.4.31 [71]. Compared to the previous result, the boundaries are more vague, and we have to perform fine-tuning of the hyperparameters to obtain the clustering which is close to our intuition. Thus, the clustering is again obtained in an artificial way. The checkerboard pattern and the fine-tuning justify an additional modification of the clustering by drawing lines, as in the previous paragraph. Our discussion in the following is based on the `JarvisPatrick` with the radius 0.31 (Figure 3.4.31), and its modification which we perform by hand is shown in Figure 3.4.32. The resultant numbers of the configurations in each cluster are summarized in Table 3.4.7.

Here, the number of clusters is 19. Ideally, 20 clusters should be found, the same as in the previous paragraph, but this should be reasonably considered within an error margin.

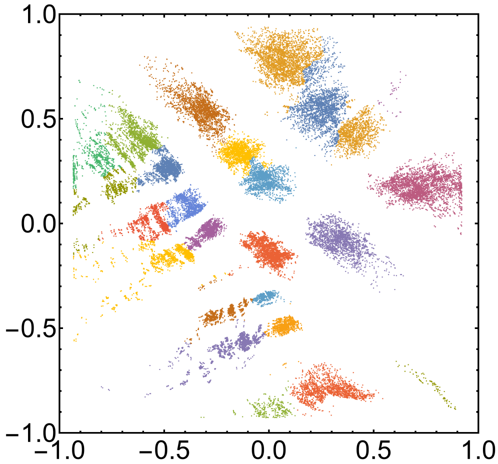

Figure 3.4.30: `MeanShift`.

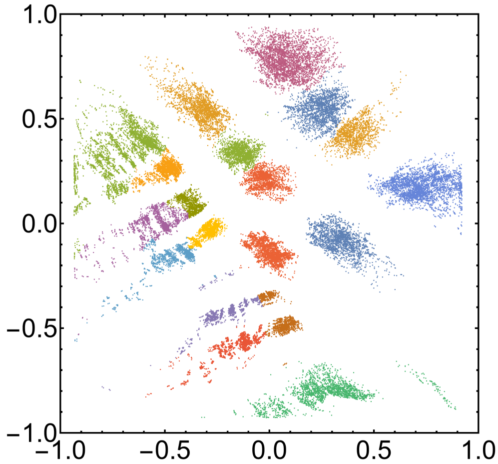

Figure 3.4.31: `JarvisPatrick` with radius 0.31.

The data is classified into two categories: those that contain the aligned configurations and those that do not, as in the previous case. Figure 3.4.33 shows the distributions of the hidden tadpole charges corresponding to Figure 3.4.25. As before, we can see that the two distributions peaked at different locations. Therefore, we can conclude that our analysis is robust to the positional encoding and that the hidden tadpole charge is the crucial factor in this case as well.

We again choose the hidden tadpole charges in Figures 3.4.26 and 3.4.27 and illustrate their distributions in the latent layer in Figures 3.4.34 and 3.4.35, respectively. Note that the configurations with the same hidden tadpole charges are again likely to be placed along the lines.

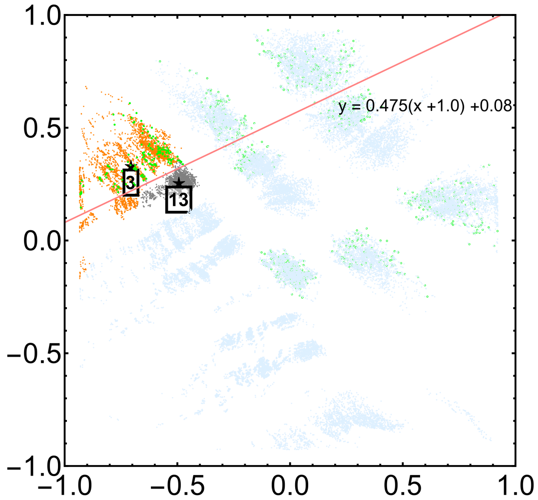

Figure 3.4.32: The modification by drawing a line. The equation of the line is shown in the figure, where x, y denote the coordinates for horizontal- and vertical-axes. Clusters 3 and 13 are colored orange and gray, respectively. We define the line as a new border and modify the clustering. The line which we drew also passes through the blank areas between other islands.

Cluster No.	1	2	3*	4	5	6	7	8	9	10
Aligned/Total	136/2329	0/1494	147/3173	297/3171	0/1132	0/1681	0/1326	0/1332	0/1606	0/1048
Cluster No.	11	12	13*	14	15	16	17	18	19	
Aligned/Total	0/1720	233/3431	0/2386	189/3412	0/2205	108/2173	147/2214	93/1949	0/1769	

Table 3.4.7: The numbers of the configurations in each cluster after the modification. The symbol * is added to the clusters which we modified.

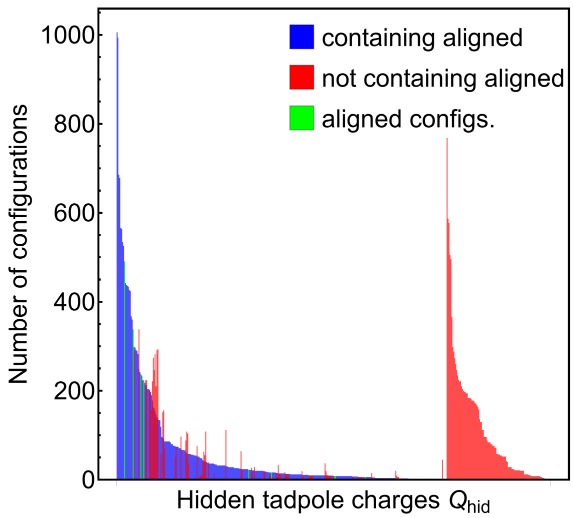

Figure 3.4.33: The distributions of the numbers of configurations for given hidden tadpole charge.

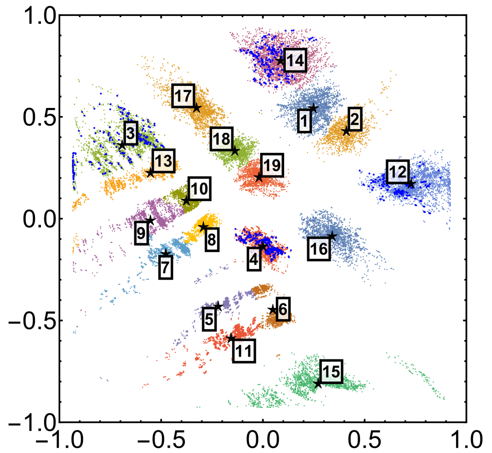

Figure 3.4.34: Hidden tadpole charge: $Q_{\text{hid}} = (1, 7, 6, 5)$. All 1,005 configurations with the charge are in the category containing aligned configurations (blue dots). [71]

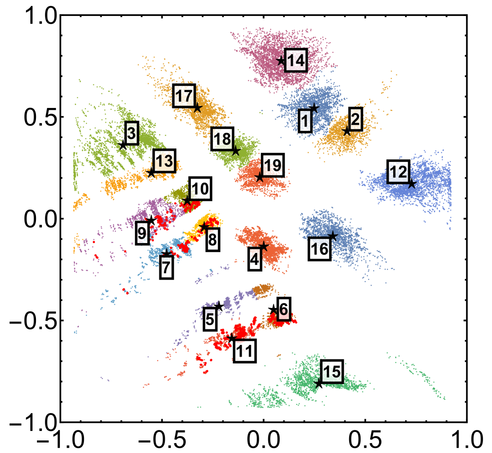

Figure 3.4.35: Hidden tadpole charge: $Q_{\text{hid}} = (2, 6, 2, 7)$. All of the 768 configurations with the charge are in the category that does not contain the aligned configurations (red dots). [71]

3.4.3.2 Models 2 and 3

Clustering

Although the hidden tadpole charge is implied to be an important feature of the landscape, the analysis is restricted to Model 1. In the following, we will see that the quantity again turns out to be the important feature in both Model 2 and 3 cases, while the clustering becomes vague. To compare with the previous results, we fix the architecture as AE-2 (tanh-2) and vary the amplitudes A. We illustrate the losses in Figures 3.4.36 and 3.4.37 [71] for both Model 2 and 3. As in the case of Model 1, the losses did not provide guidance on which model to choose. The outputs from the latent layers are summarized in Figures 3.4.38 to 3.4.43 for Model 2 and Figures 3.4.44 to 3.4.49 for Model 3 [71]. One can see that the clusters become vague compared to the previous one, but the red points, which denote the aligned configurations, are not scattered universally over the 2d spaces. Thus, we observe that the aligned configurations are distinguished from the others by the autoencoder models as in the previous case.

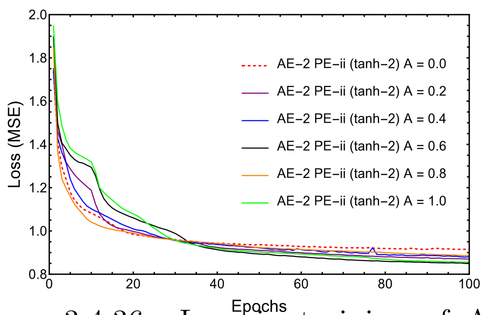

Figure 3.4.36: Loss in training of AE-2 with the amplitudes for PE-ii. Model 2 is used.

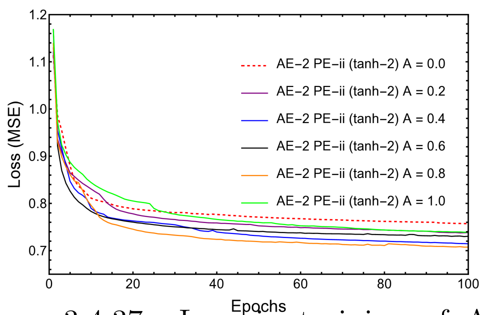

Figure 3.4.37: Loss in training of AE-2 with the amplitudes for PE-ii. Model 3 is used.

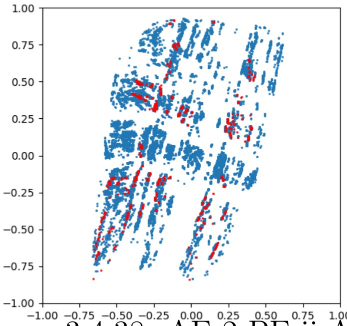

Figure 3.4.38: AE-2 PE-ii A=0.0 Model 2.

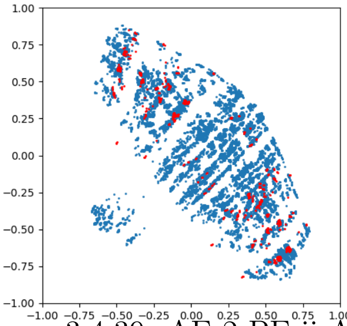

Figure 3.4.39: AE-2 PE-ii A=0.2 Model 2.

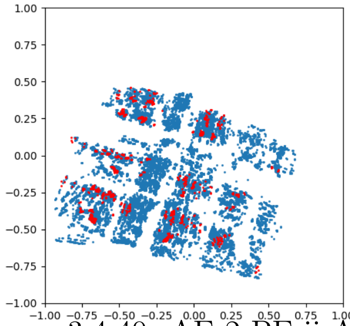

Figure 3.4.40: AE-2 PE-ii A=0.4 Model 2.

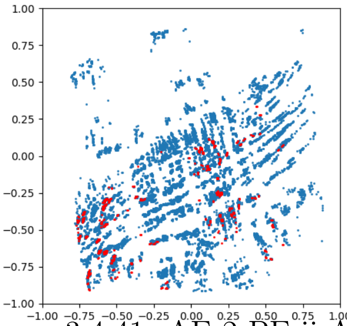

Figure 3.4.41: AE-2 PE-ii A=0.6 Model 2.

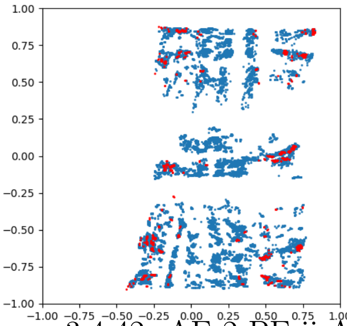

Figure 3.4.42: AE-2 PE-ii A=0.8 Model 2.

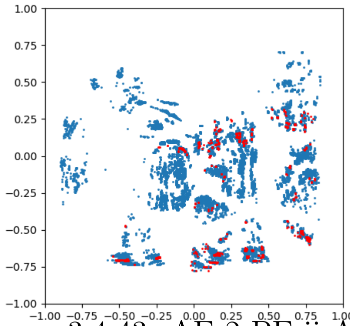

Figure 3.4.43: AE-2 PE-ii A=1.0 Model 2.

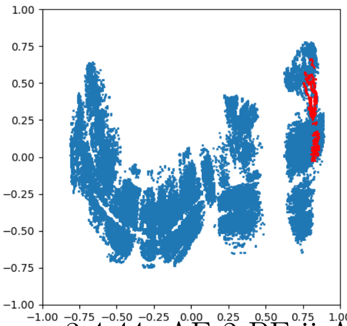

Figure 3.4.44: AE-2 PE-ii A=0.0 Model 3.

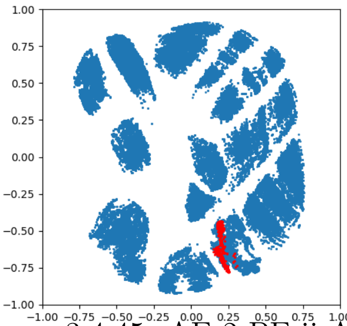

Figure 3.4.45: AE-2 PE-ii A=0.2 Model 3.

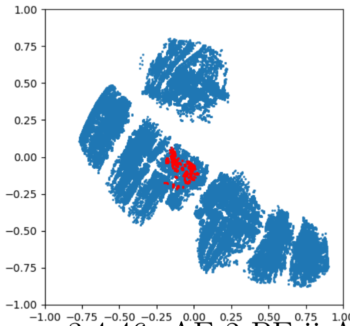

Figure 3.4.46: AE-2 PE-ii A=0.4 Model 3.

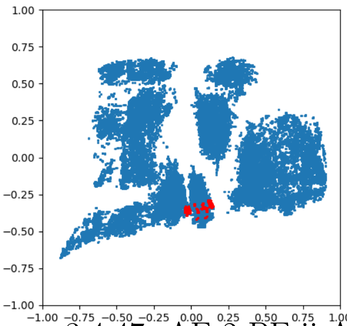

Figure 3.4.47: AE-2 PE-ii A=0.6 Model 3.

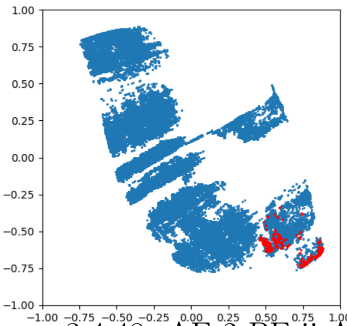

Figure 3.4.48: AE-2 PE-ii A=0.8 Model 3.

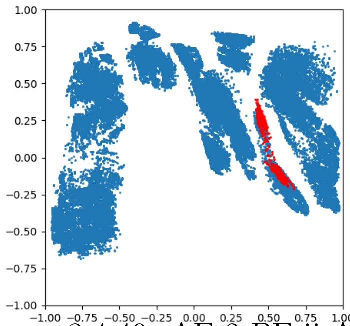

Figure 3.4.49: AE-2 PE-ii A=1.0 Model 3.

Model 2

For Model 2, we focus on the $A = 0.4$ case, which exhibits the clusters most clearly. Although it is vague compared to the previous case, we observe that the checkerboard pattern is weakly realized. Then we classify the clusters by drawing lines in the first place, instead of tuning the hyperparameters carefully.

The clustering that we obtained is illustrated in Figures 3.4.50 and 3.4.51. The resultant numbers of configurations in each cluster in Table 3.4.8.

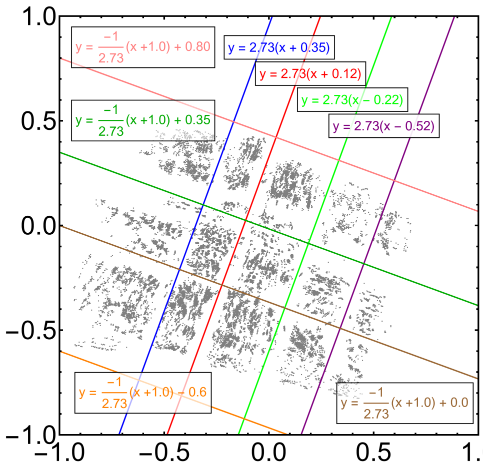

Figure 3.4.50: The clusters are artificially classified based on the drawn lines. [71]

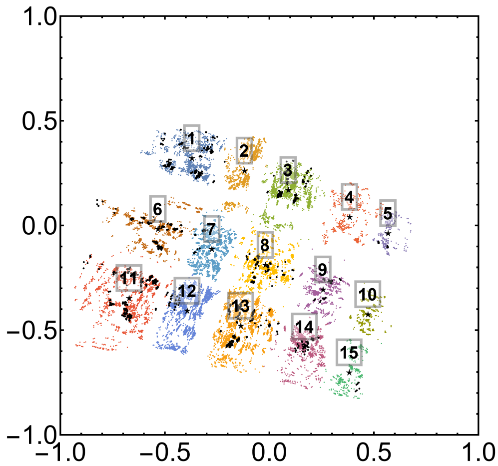

Figure 3.4.51: The resultant classification. The aligned configurations are colored black.

Cluster No.	1	2	3	4	5	6	7	8	9	10
Aligned/Total	146/1457	4/1617	150/2569	0/475	6/316	122/1512	14/2436	132/2667	24/963	0/393
Cluster No.	11	12	13	14	15					
Aligned/Total	179/2072	18/2883	190/3347	39/1270	6/359					

Table 3.4.8: The numbers of the configurations in each cluster.

Then, we again classify the clusters into two categories and analyze their differences. As the boundaries between the clusters become vague, it is difficult to perform the binary classification in a natural way. In fact, some clusters have only a few aligned configurations. One possible option is to ignore such outliers in the analysis. The results with and without treatment of outliers are shown in Figures 3.4.25 and 3.4.33. When we impose the threshold at 2% in the binary classification, the displacement is also observed in Model 2, but if we rigorously classify the clusters in the two categories, the two bar charts completely overlap and the displacement cannot be observed.

Therefore, we can conclude that the autoencoder model also weights the hidden tadpole charge in this case, but the other feature we did not discover becomes more important than in the previous case.

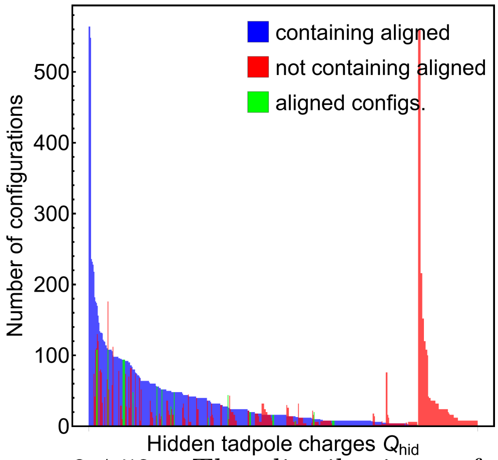

Figure 3.4.52: The distributions of the number of configurations for given hidden tadpole charge. The "not containing aligned" category is a set of Clusters 2, 4, 5, 7, 10, 12, and 15, with the threshold set at 2%. [71]

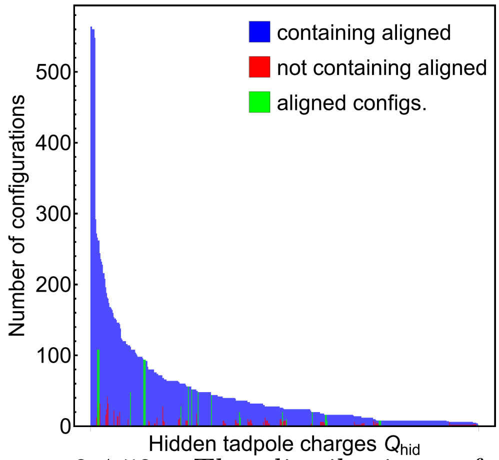

Figure 3.4.53: The distributions of the number of configurations for given hidden tadpole charge. The "not containing aligned" category is a set of Clusters 4 and 10, with the threshold is strictly set as zero. We do not observe the displacement for this classification. [71]

The conclusion that the autoencoder weights the hidden tadpole charge is supported by the observation that the same hidden tadpole charges are likely to be placed along the lines. To see this, we define the angular similarity matrix as

$$S_{ij} := \arccos \frac{\mathbf{n}_i \cdot \mathbf{n}_j}{|\mathbf{n}_i||\mathbf{n}_j|}, \tag{3.75}$$

where $\mathbf{n}_i$ is a vector of the numbers of different configurations for each hidden tadpole charge and i labels the clusters. The a-th element of $\mathbf{n}_i$ is the number of configurations in i-th cluster that have the a-th hidden tadpole charge (a is a label for the possible hidden tadpole charge as the horizontal axes of Figures 3.4.25, 3.4.33 and 3.4.52). By the similarity matrix, we quantify the similarity between the clusters in terms of the hidden tadpole charges.

The angular similarity matrix is shown in Table 3.4.9. The values that are the smallest to the third in each row have been highlighted in blue to emphasize the near clusters. With the exception of the position at (14, 4), it can be observed that the blue positions are aligned diagonally. For instance, in the first row, 1, 6, and 11 are colored blue. To see the meaning of these colored elements, let us remind Figure 3.4.51. Indeed, Clusters 1, 6, and 11 are placed diagonally in the latent layer. It is true for other triads, $(2, 7, 12)$, $(3, 8, 13)$, $(4, 9, 14)$ and $(5, 10, 15)$ as one can see the diagonal pattern in the figure.

Thus, the angular similarity matrix quantifies the diagonal structure, which is also observed in Model 1. This vertical correlation is likely related to the emergence of the checkerboard pattern and suggests the presence of other important factors.

It is worth noting the exception at (14, 4), where the element at (14, 12) has a slightly smaller value. The value at (9, 4) is colored, but it is very close to (9, 12). This irregularity may arise because Cluster 4 lacks the aligned configurations that other clusters possess. Therefore, we can conclude that the absence of aligned configurations reduces the similarity.

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	**0**	0.476	0.361	0.418	0.466	**0.340**	0.475	0.446	0.428	0.477	**0.272**	0.473	0.387	0.443	0.479
2	0.476	**0**	0.48	0.449	0.453	0.47	**0.234**	0.49	0.474	0.469	0.476	**0.263**	0.482	0.478	0.469
3	0.361	0.48	**0**	0.426	0.468	0.42	0.484	**0.255**	0.457	0.477	0.399	0.477	**0.194**	0.459	0.463
4	0.418	0.449	0.426	**0**	0.49	0.44	0.466	0.444	**0.399**	0.489	0.466	0.455	0.452	**0.392**	0.494
5	0.466	0.453	0.468	0.49	**0**	0.41	0.47	0.465	0.472	**0.295**	0.412	0.473	0.447	0.473	**0.279**
6	**0.340**	0.47	0.42	0.44	0.41	**0**	0.479	0.387	0.467	0.408	**0.285**	0.462	0.373	0.45	0.421
7	0.475	**0.234**	0.484	0.466	0.47	0.479	**0**	0.487	0.415	0.452	0.475	**0.165**	0.481	0.422	0.463
8	0.446	0.49	**0.255**	0.444	0.465	0.387	0.487	**0**	0.464	0.468	0.427	0.478	**0.213**	0.458	0.471
9	0.428	0.474	0.457	**0.399**	0.472	0.467	0.415	0.464	**0**	0.486	0.445	0.4	0.459	**0.179**	0.481
10	0.477	0.469	0.477	0.489	**0.295**	0.408	0.452	0.468	0.486	**0**	0.429	0.464	0.459	0.479	**0.256**
11	**0.272**	0.476	0.399	0.466	0.412	**0.285**	0.475	0.427	0.445	0.429	**0**	0.48	0.323	0.464	0.415
12	0.473	**0.263**	0.477	0.455	0.473	0.462	**0.165**	0.478	0.4	0.464	0.48	**0**	0.473	0.384	0.459
13	0.387	0.482	**0.194**	0.452	0.447	0.373	0.481	**0.213**	0.459	0.459	0.323	0.473	**0**	0.46	0.449
14	0.443	0.478	0.459	0.392	0.473	0.45	0.422	0.458	**0.179**	0.479	0.464	**0.384**	0.46	**0**	0.486
15	0.479	0.469	0.463	0.494	**0.279**	0.421	0.463	0.471	0.481	**0.256**	0.415	0.459	0.449	0.486	**0**

Table 3.4.9: The angular similarity matrix in units of radian. The rows as well as the columns correspond to the clusters, as indicated by the numbers of the clusters. The smallest values to the third nearest cluster in each row are shown in blue.

Model 3

For Model 3, the checkerboard pattern does not appear in Figures 3.4.44 to 3.4.49 while the aligned configurations cluster at some small regions. Thus, it is impossible to find a clustering in a natural way, and we simply adopt the automatic methods which are used for Model 1. Choosing $A = 0.2$ case, the results with two different methods are shown in Figures 3.4.54 and 3.4.55 [71]. For our purpose, it is important to distinguish the islands that contain the aligned configurations from the other clusters in a reasonable way. Since Cluster 14 in Figure 3.4.55 contains the aligned ones and is distinguished clearly, we choose `JarvisPatrick` with radius 0.127 as a clustering method.

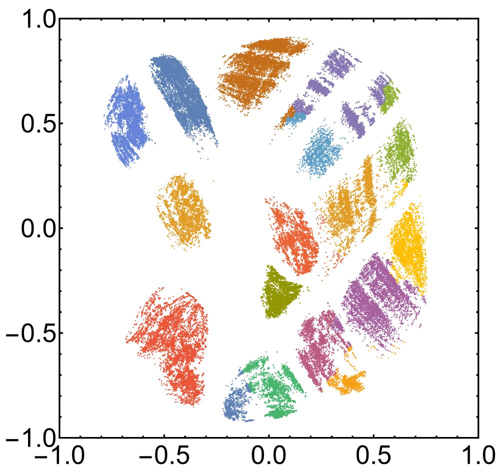

Figure 3.4.54: `MeanShift`.

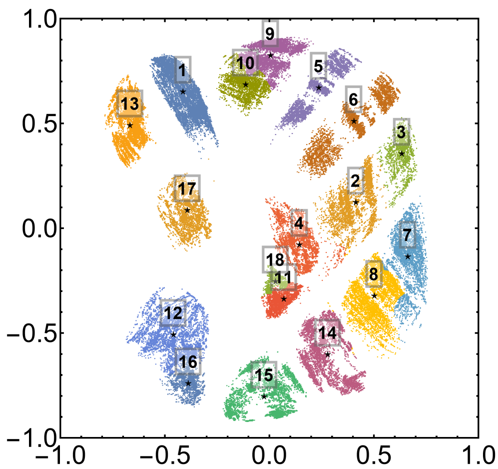

Figure 3.4.55: `JarvisPatrick` with radius 0.127. The numbers label the clusters.

Then, we study the distributions of the hidden tadpole charges for both categories as in the previous cases. Since Cluster 14 is the only one that contains the aligned configurations, we classify in this way and illustrate the result in Figure 3.4.56. As shown, the two distributions completely overlap again and the displacement is not observed. However, if we assume Cluster 15 and Cluster 14 are in the same category, the result becomes as Figure 3.4.57. In this case, we observe the displacement again, and the hidden tadpole charge is a characteristic quantity of the landscape if we adopt this classification.

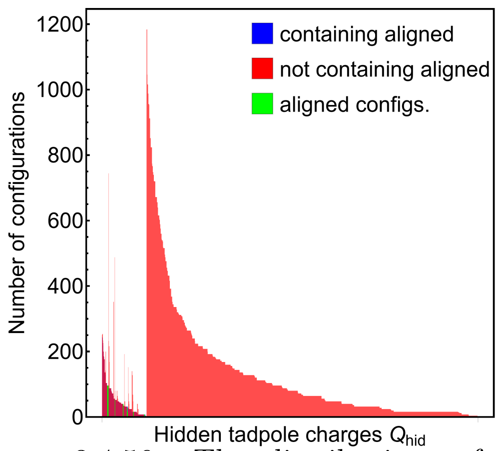

Figure 3.4.56: The distributions of the number of configurations for given hidden tadpole charge. The "containing aligned" category is Cluster 14, with the threshold is strictly set as zero. In this case, the displacement of two distributions does not appear. [71]

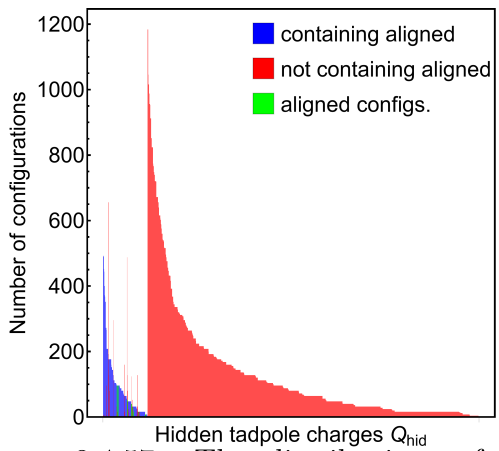

Figure 3.4.57: The distributions of the number of configurations for given hidden tadpole charge. However, the "containing aligned" category is a set of Cluster 14 and 15. The reason for choosing them is explained in the following. [71]

To quantify the situation, we again calculate the angular similar matrix, where the result is shown in Table 3.4.10. One can see that the elements in each row are clustered into two or three groups, where we used a typical clustering method `DBSCAN` with the radius $= 0.15$. Although there is no clear checkerboard pattern (thus the diagonal pattern in the matrix vanishes), the claim remains unchanged: the clusters form new clusters, respecting their associated hidden tadpole charges. The matrix shows that Cluster 14 and 15 are very similar and this is the reason why the classification in Figure 3.4.57 exhibits the displacement. Note that Cluster 14 is only similar to Cluster 15. Additionally, the aligned configurations have unique hidden tadpole charges. Based on these results, we conclude that the autoencoder model places significant weight on the hidden tadpole charge.

As in the previous cases, we were unable to identify the other hidden feature implied by the existence of the clusters of clusters. This feature will explain why Cluster 14 and 15 are eventually separated into the two different islands.

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	**0**	**0.264**	**0.282**	0.482	0.473	0.473	0.468	0.479	0.475	0.475	**0.206**	0.49	0.488	0.499	0.499	0.487	0.485	**0.245**
2	**0.264**	**0**	**0.217**	0.493	0.464	0.464	0.46	0.478	0.469	0.477	0.459	0.498	0.497	0.499	0.499	0.496	0.497	0.473
3	**0.282**	**0.217**	**0**	0.489	0.473	0.473	0.467	0.478	0.477	0.468	0.458	0.496	0.495	0.498	0.498	0.494	0.494	0.475
4	0.482	0.493	0.489	**0**	0.499	0.499	0.494	0.496	0.496	0.493	0.482	0.483	0.487	0.500	0.500	0.496	**0.00793**	0.482
5	0.473	0.464	0.473	0.499	**0**	**0.000**	0.48	0.467	0.473	0.472	0.492	0.498	0.498	0.491	0.492	0.498	0.499	0.494
6	0.473	0.464	0.473	0.499	**0.000**	**0**	0.48	0.467	0.473	0.472	0.492	0.498	0.498	0.491	0.492	0.498	0.499	0.494
7	0.468	0.460	0.467	0.494	0.480	0.480	**0**	**0.166**	**0.131**	**0.117**	0.488	0.500	0.499	0.491	0.494	0.499	0.495	0.492
8	0.479	0.478	0.478	0.496	0.467	0.467	**0.166**	**0**	**0.0639**	**0.165**	0.49	0.500	0.500	0.484	0.486	0.499	0.496	0.493
9	0.475	0.469	0.477	0.496	0.473	0.473	**0.131**	**0.0639**	**0**	**0.176**	0.491	0.500	0.500	0.487	0.490	0.499	0.497	0.493
10	0.475	0.477	0.468	0.493	0.472	0.472	**0.117**	**0.165**	**0.176**	**0**	0.487	0.500	0.499	0.490	0.493	0.498	0.493	0.491
11	**0.206**	0.459	0.458	0.482	0.492	0.492	0.488	0.490	0.491	0.487	**0**	0.490	0.488	0.500	0.500	0.486	0.483	**0.174**
12	0.490	0.498	0.496	0.483	0.498	0.498	0.500	0.500	0.500	0.5	0.490	**0**	**0.07**	0.500	0.500	**0.194**	0.483	0.488
13	0.488	0.497	0.495	0.487	0.498	0.498	0.499	0.500	0.500	0.499	0.488	**0.070**	**0**	0.500	0.500	**0.124**	0.487	0.486
14	0.499	0.499	0.498	0.500	0.491	0.491	0.491	0.484	0.487	0.490	0.500	0.500	0.500	**0**	**0.0799**	0.500	0.500	0.500
15	0.499	0.499	0.498	0.500	0.492	0.492	0.494	0.486	0.490	0.493	0.500	0.500	0.500	**0.0799**	**0**	0.500	0.500	0.500
16	0.487	0.496	0.494	0.496	0.498	0.498	0.499	0.499	0.499	0.498	0.486	**0.194**	**0.124**	0.500	0.500	**0**	0.496	0.485
17	0.485	0.497	0.494	**0.00793**	0.499	0.499	0.495	0.496	0.497	0.493	0.483	0.483	0.487	0.500	0.500	0.496	**0**	0.483
18	**0.245**	0.473	0.475	0.482	0.494	0.494	0.492	0.493	0.493	0.491	**0.174**	0.488	0.486	0.500	0.500	0.485	0.483	**0**

Table 3.4.10: The angular similarity matrix in units of radian. The rows as well as the columns correspond to the clusters, as indicated by the numbers of the clusters. In each row, the clusters form new clusters, respecting their similarities. Those clusters are indicated by their colors. Note that the clustering occurs in each row, and clusters with the same color in different rows do not belong to the same cluster.

3.5 Conclusions of Chapter 2

This study investigates how neural networks, specifically autoencoder models, learn intersecting D-brane models and extract important features of the landscape. Based on the models in Ref. [183], we generated D-brane configurations that align and do not align with the generations of quarks and leptons in each specific model. The goal is to classify input data and extract important features by training autoencoder models on the datasets. The features are deemed important as they are crucial for replicating the input data from the latent layer, and those will serve as constraints on the phenomenologically-favored aligned configurations.

Our initial expectation was that the autoencoder models would be able to classify each data set into aligned and non-aligned configurations. However, due to the difficulty of clustering, we restrict the dimensions of the latent layers to be small and it makes the expressive abilities not sufficient to replicate all features, which is constrained by many nonlinear relations, such as the algebra and the SUSY condition. Then, it is expected that the only important features are extracted by the autoencoder models.

Indeed, the dataset is separated into many islands, and the aligned and the non-aligned configurations are clustered together in a nontrivial way, which may respect their macroscopic properties.

Due to the lack of interpretability, we had to take a brute-force analysis that relied on statistics. However, we would like to emphasize that this autoencoder-driven method partially goes beyond the usual statistical studies, because if the characteristic value is found then it will be one of the most important features of the D-brane configurations.

In fact, we found that the hidden tadpole charge, which is the tadpole charge of the D6-branes in the hidden sector, has characteristic distributions over the clusters. In particular, for Models 1 and 3, the hidden tadpole charges of the aligned configurations are unique for them. The checkerboard pattern and the clusters of the clusters also show the characterization via the hidden tadpole charge. It is interesting that the non-aligned configurations are also characterized by the quantity. Thus we concluded that the hidden tadpole charge characterizes the landscape of type IIA intersecting D-brane configurations.

At the same time, the interesting patterns observed suggest that there is at least one more important quantity, as well as a hidden tadpole charge. Unfortunately, due to the principal difficulty, we could not find that feature. It is interesting to find that in future work, but it seems not efficient to study using these ML models which have no interpretability.

One possible candidate for finding the other feature is explainable artificial intelligence (XAI) (see for a review, e.g., [206]). However, it is not easy to realize an expedient model for specific tasks.

Furthermore, it is also interesting to study the roles of tadpole charges in a more general string landscape. As explained in the previous chapter, and pointed out in our previous work [72], the tadpole charge usually plays an important role in a general context. Considering both the flux compactification, which is another source of RR charges, and phenomenological D-brane models simultaneously may shed light on this kind of problem.

We close this section by discussing our further motivation for studying physics with ML. One important application which we are currently interested in is a Transformer-based alternative approach to solving dynamics. Indeed, there are many studies about analyzing the solutions of equations of motion without solving them directly. Notice that the dynamical process is essentially equivalent to a time series data, Transformer can be used to deal with this kind of task (see, e.g., [207]). We are currently interested in analyzing the moduli stabilization problem via a Transformer-based approach, and we will report the result in the upcoming paper.

Chapter 4

Conclusion of the Thesis

In this thesis, we have discussed the type II string landscape by focusing on the statistics, symmetries, and dualities, while all the topics are motivated by phenomenological interest. We do not mention the details, since we have already concluded in each chapter.

Unfortunately, but rather trivially (since there is no effort to do so), we could not address the main problem that we mentioned in Chapter 1: Did we obtain a model in this thesis that is completely consistent with the experiment? To answer this question with yes, we have tremendous difficulties that we must overcome in the future.

In fact, we have encountered important quantities that we have mostly ignored. One of those quantities is the tadpole charge with various sources. Indeed, it plays crucial roles in canceling anomalies, ensuring the finiteness of the flux vacua and stability of moduli fields, mixing the axio-dilaton and the complex-structure moduli, controlling the void radius, connecting the moduli sector with the matter physics, and characterizing intersecting D-brane models. Furthermore, it will be related to the vacuum energy, as it measures the flux contributions.

Of course, we should have considered it for consistency. Moreover, since we will need all of the machinery to discuss the rigorous phenomenology, the fact that the quantity appears in all the corners of string phenomenology motivates us to study the relationships between separated frameworks in terms of the tadpole charge.

We would like to conclude this thesis with a remark that emphasizes the importance of the tadpole charge and motivates our future work after this thesis.

Acknowledgements I (K.I.) would like to express my deepest gratitude to my gracious mentor, Hajime Otsuka, for his continuous and patient support during the six years since I was an undergraduate student. Without his thoughtful guidance, I would not have been able to continue my doctoral studies. I would like to thank all my collaborators, Tatsuo Kobayashi, Hiroshi Okada, Kai Takafumi, and Satsuki Nishimura, who supported the works on which this thesis is based. I would like to thank Tatsuo Kobayashi for his constructive suggestions and kind advice, which I have taken even though I am not his student. I am especially thankful to my supervisor, Shun'ya Mizoguchi, who kindly taught me what I did not understand due to my lack of effort. I would like to express my appreciation to Hiroyuki Abe, who gave me educational advice in my early days. I also express my appreciation to those who discussed with me, Yoshihiko Abe, Shuntaro Aoki, Yuma Furuta, Yuta Hamada, Tetsutaro Higaki, Atsushi Hinata, Masaki Honda, Satoshi Iso, Ryuichiro Kitano, Yuichi Koga, Rinto Kuramochi, Sota Nakajima, Takumi Oikawa, Yutaka Sakamura, Shotaro Yata, and Haruki Takahashi. Finally, I would like to thank all members of the KEK Theory Center.

Bibliography

[1] L.E. Ibanez and A.M. Uranga, *String theory and particle physics: An introduction to string phenomenology*, Cambridge University Press (2, 2012).

[2] Particle Data Group collaboration, *Review of Particle Physics*, *PTEP* **2022** (2022) 083C01.

[3] T. Kaluza, *Zum Unitätsproblem der Physik*, *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)* **1921** (1921) 966 [1803.08616].

[4] O. Klein, *Quantum Theory and Five-Dimensional Theory of Relativity. (In German and English)*, *Z. Phys.* **37** (1926) 895.

[5] L. Alvarez-Gaume and E. Witten, *Gravitational Anomalies*, *Nucl. Phys. B* **234** (1984) 269.

[6] M.B. Green and J.H. Schwarz, *Anomaly Cancellation in Supersymmetric D=10 Gauge Theory and Superstring Theory*, *Phys. Lett. B* **149** (1984) 117.

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

[8] P. Candelas, G.T. Horowitz, A. Strominger and E. Witten, *Vacuum configurations for superstrings*, *Nucl. Phys. B* **258** (1985) 46.

[9] L.J. Dixon, J.A. Harvey, C. Vafa and E. Witten, *Strings on Orbifolds*, *Nucl. Phys. B* **261** (1985) 678.

[10] L.J. Dixon, J.A. Harvey, C. Vafa and E. Witten, *Strings on Orbifolds. 2.*, *Nucl. Phys. B* **274** (1986) 285.

[11] K.S. Narain, *New Heterotic String Theories in Uncompactified Dimensions < 10*, *Phys. Lett. B* **169** (1986) 41.

[12] L.E. Ibanez, H.P. Nilles and F. Quevedo, *Orbifolds and Wilson Lines*, *Phys. Lett. B* **187** (1987) 25.

[13] A. Strominger, *Superstrings with Torsion*, *Nucl. Phys. B* **274** (1986) 253.

[14] L.E. Ibanez, J. Mas, H.-P. Nilles and F. Quevedo, *Heterotic Strings in Symmetric and Asymmetric Orbifold Backgrounds*, *Nucl. Phys. B* **301** (1988) 157.

[15] J. Polchinski, *Dirichlet Branes and Ramond-Ramond charges*, *Phys. Rev. Lett.* **75** (1995) 4724 [hep-th/9510017].

[16] J. Polchinski and A. Strominger, *New vacua for type II string theory*, *Phys. Lett. B* **388** (1996) 736 [hep-th/9510227].

[17] J. Michelson, *Compactifications of type IIB strings to four-dimensions with nontrivial classical potential*, *Nucl. Phys. B* **495** (1997) 127 [hep-th/9610151].

[18] S. Gukov, C. Vafa and E. Witten, *CFT's from Calabi-Yau four folds*, *Nucl. Phys. B* **584** (2000) 69 [hep-th/9906070].

[19] K. Dasgupta, G. Rajesh and S. Sethi, *M theory, orientifolds and G - flux*, *JHEP* **08** (1999) 023 [hep-th/9908088].

[20] S.B. Giddings, S. Kachru and J. Polchinski, *Hierarchies from fluxes in string compactifications*, *Phys. Rev. D* **66** (2002) 106006 [hep-th/0105097].

[21] T.W. Grimm and J. Louis, *The Effective action of N = 1 Calabi-Yau orientifolds*, *Nucl. Phys. B* **699** (2004) 387 [hep-th/0403067].

[22] M. Grana, *Flux compactifications in string theory: A Comprehensive review*, *Phys. Rept.* **423** (2006) 91 [hep-th/0509003].

[23] R. Blumenhagen, B. Kors, D. Lust and S. Stieberger, *Four-dimensional String Compactifications with D-Branes, Orientifolds and Fluxes*, *Phys. Rept.* **445** (2007) 1 [hep-th/0610327].

[24] M.R. Douglas and S. Kachru, *Flux compactification*, *Rev. Mod. Phys.* **79** (2007) 733 [hep-th/0610102].

[25] E. Witten, *String theory dynamics in various dimensions*, *Nucl. Phys. B* **443** (1995) 85 [hep-th/9503124].

[26] M. Berkooz, M.R. Douglas and R.G. Leigh, *Branes intersecting at angles*, *Nucl. Phys. B* **480** (1996) 265 [hep-th/9606139].

[27] R. Blumenhagen, M. Cvetic, P. Langacker and G. Shiu, *Toward realistic intersecting D-brane models*, *Ann. Rev. Nucl. Part. Sci.* **55** (2005) 71 [hep-th/0502005].

[28] D. Lust, *String Landscape and the Standard Model of Particle Physics*, in *11th Marcel Grossmann Meeting on General Relativity*, pp. 148–176, 7, 2007, DOI [0707.2305].

[29] P. Touboul et al., *MICROSCOPE Mission: First Results of a Space Test of the Equivalence Principle*, *Phys. Rev. Lett.* **119** (2017) 231101 [1712.01176].

[30] E.G. Adelberger, B.R. Heckel and A.E. Nelson, *Tests of the gravitational inverse square law*, *Ann. Rev. Nucl. Part. Sci.* **53** (2003) 77 [hep-ph/0307284].

[31] W. Taylor and Y.-N. Wang, *The F-theory geometry with most flux vacua*, *JHEP* **12** (2015) 164 [1511.03209].

[32] E. Palti, *The Swampland: Introduction and Review*, *Fortsch. Phys.* **67** (2019) 1900037 [1903.06239].

[33] T.W. Grimm, *Moduli space holography and the finiteness of flux vacua*, *JHEP* **10** (2021) 153 [2010.15838].

[34] B. Bakker, T.W. Grimm, C. Schnell and J. Tsimerman, *Finiteness for self-dual classes in integral variations of Hodge structure*, 2112.06995.

[35] T.W. Grimm and J. Monnee, *Finiteness Theorems and Counting Conjectures for the Flux Landscape*, 2311.09295.

[36] J. Polchinski, *String theory to the rescue*, 12, 2015 [1512.02477].

[37] L. Susskind, *The Anthropic landscape of string theory*, hep-th/0302219.

[38] A.N. Schellekens, *The Emperor's Last Clothes? Overlooking the String Theory Landscape*, *Rept. Prog. Phys.* **71** (2008) 072201 [0807.3249].

[39] A.N. Schellekens, *Life at the Interface of Particle Physics and String Theory*, *Rev. Mod. Phys.* **85** (2013) 1491 [1306.5083].

[40] S. Ashok and M.R. Douglas, *Counting flux vacua*, *JHEP* **01** (2004) 060 [hep-th/0307049].

[41] M.R. Douglas, *The Statistics of string / M theory vacua*, *JHEP* **05** (2003) 046 [hep-th/0303194].

[42] F. Denef and M.R. Douglas, *Distributions of flux vacua*, *JHEP* **05** (2004) 072 [hep-th/0404116].

[43] F. Denef and M.R. Douglas, *Distributions of nonsupersymmetric flux vacua*, *JHEP* **03** (2005) 061 [hep-th/0411183].

[44] F. Denef, M.R. Douglas and S. Kachru, *Physics of String Flux Compactifications*, *Ann. Rev. Nucl. Part. Sci.* **57** (2007) 119 [hep-th/0701050].

[45] F. Ruehle, *Data science applications to string theory*, *Phys. Rept.* **839** (2020) 1.

[46] F. Feruglio, *Are neutrino masses modular forms?*, *arXiv e-prints* (2017) arXiv:1706.08749 [1706.08749].

[47] T. Kobayashi, K. Tanaka and T.H. Tatsuishi, *Neutrino mixing from finite modular groups*, *Phys. Rev. D* **98** (2018) 016004 [1803.10391].

[48] J.T. Penedo and S.T. Petcov, *Lepton Masses and Mixing from Modular S_4 Symmetry*, *Nucl. Phys. B* **939** (2019) 292 [1806.11040].

[49] J.C. Criado and F. Feruglio, *Modular Invariance Faces Precision Neutrino Data*, *SciPost Phys.* **5** (2018) 042 [1807.01125].

[50] T. Kobayashi, N. Omoto, Y. Shimizu, K. Takagi, M. Tanimoto and T.H. Tatsuishi, *Modular A_4 invariance and neutrino mixing*, *JHEP* **11** (2018) 196 [1808.03012].

[51] P.P. Novichkov, J.T. Penedo, S.T. Petcov and A.V. Titov, *Modular S_4 models of lepton masses and mixing*, *JHEP* **04** (2019) 005 [1811.04933].

[52] P.P. Novichkov, J.T. Penedo, S.T. Petcov and A.V. Titov, *Modular A_5 symmetry for flavour model building*, *JHEP* **04** (2019) 174 [1812.02158].

[53] F.J. de Anda, S.F. King and E. Perdomo, *$SU(5)$ grand unified theory with A_4 modular symmetry*, *Phys. Rev. D* **101** (2020) 015028 [1812.05620].

[54] H. Okada and M. Tanimoto, *CP violation of quarks in A_4 modular invariance*, *Phys. Lett. B* **791** (2019) 54 [1812.09677].

[55] P.P. Novichkov, S.T. Petcov and M. Tanimoto, *Trimaximal Neutrino Mixing from Modular A4 Invariance with Residual Symmetries*, *Phys. Lett. B* **793** (2019) 247 [1812.11289].

[56] T. Kobayashi and M. Tanimoto, *Modular flavor symmetric models*, 7, 2023 [2307.03384].

[57] G.-J. Ding and S.F. King, *Neutrino Mass and Mixing with Modular Symmetry*, 2311.09282.

[58] K. Ishiguro, T. Kobayashi and H. Otsuka, *Spontaneous CP violation and symplectic modular symmetry in Calabi-Yau compactifications*, 2010.10782.

[59] K. Ishiguro, T. Kobayashi and H. Otsuka, *Symplectic modular symmetry in heterotic string vacua: flavor, CP, and R-symmetries*, *JHEP* **01** (2022) 020 [2107.00487].

[60] G.-J. Ding, F. Feruglio and X.-G. Liu, *Automorphic Forms and Fermion Masses*, *JHEP* **01** (2021) 037 [2010.07952].

[61] A. Baur, M. Kade, H.P. Nilles, S. Ramos-Sanchez and P.K.S. Vaudrevange, *Siegel modular flavor group and CP from string theory*, *Phys. Lett. B* **816** (2021) 136176 [2012.09586].

[62] H.P. Nilles, S. Ramos-Sanchez, A. Trautner and P.K.S. Vaudrevange, *Orbifolds from Sp(4,Z) and their modular symmetries*, *Nucl. Phys. B* **971** (2021) 115534 [2105.08078].

[63] S. Kikuchi, T. Kobayashi, K. Nasu, S. Takada and H. Uchida, *Modular symmetry in magnetized T^{2g} torus and orbifold models*, 2309.16447.

[64] S. Kikuchi, T. Kobayashi, K. Nasu, S. Takada and H. Uchida, *Zero-modes in magnetized T6/ZN orbifold models through Sp(6,Z) modular symmetry*, *Phys. Rev. D* **108** (2023) 036005 [2305.16709].

[65] S. Kikuchi, T. Kobayashi, K. Nasu, S. Takada and H. Uchida, *$Sp(6, Z)$ modular symmetry in flavor structures: quark flavor models and Siegel modular forms for $\widetilde{\Delta}(96)$*, 2310.17978.

[66] D. Cremades, L. Ibanez and F. Marchesano, *Yukawa couplings in intersecting D-brane models*, *JHEP* **07** (2003) 038 [hep-th/0302105].

[67] D. Cremades, L. Ibanez and F. Marchesano, *Computing Yukawa couplings from magnetized extra dimensions*, *JHEP* **05** (2004) 079 [hep-th/0404229].

[68] K. Ishiguro, T. Kobayashi and H. Otsuka, *Landscape of Modular Symmetric Flavor Models*, 2011.09154.

[69] K. Ishiguro, H. Okada and H. Otsuka, *Residual flavor symmetry breaking in the landscape of modular flavor models*, *JHEP* **09** (2022) 072 [2206.04313].

[70] K. Ishiguro, T. Kai, T. Kobayashi and H. Otsuka, *Flux Landscape with enhanced symmetry not on SL(2, $\mathbb{Z}$) elliptic points*, *JHEP* **02** (2024) 099 [2311.12425].

[71] K. Ishiguro, S. Nishimura and H. Otsuka, *Autoencoder-Driven Clustering of Intersecting D-brane Models via Tadpole Charge*, 2312.07181.

[72] K. Ishiguro and H. Otsuka, *Sharpening the boundaries between flux landscape and swampland by tadpole charge*, *JHEP* **12** (2021) 017 [2104.15030].

[73] K. Ishiguro, T. Kobayashi and H. Otsuka, *Hierarchical structure of physical Yukawa couplings from matter field Kähler metric*, *JHEP* **07** (2021) 064 [2103.10240].

[74] K. Ishiguro, T. Kai, H. Okada and H. Otsuka, *Flavor, CP and metaplectic modular symmetries in Type IIB chiral flux vacua*, *JHEP* **12** (2023) 136 [2305.19155].

[75] K. Ishiguro, T. Kai, S. Nishimura, H. Otsuka and M. Takeuchi, *Upper bound on the Atiyah-Singer index from tadpole cancellation*, 2308.12421.

[76] J. Polchinski, *String theory. Vol. 1: An introduction to the bosonic string*, Cambridge Monographs on Mathematical Physics, Cambridge University Press (12, 2007), 10.1017/CBO9780511816079.

[77] J. Polchinski, *String theory. Vol. 2: Superstring theory and beyond*, Cambridge Monographs on Mathematical Physics, Cambridge University Press (12, 2007), 10.1017/CBO9780511618123.

[78] K. Becker, M. Becker and J. Schwarz, *String theory and M-theory: A modern introduction*, Cambridge University Press (12, 2006).

[79] R. Blumenhagen, D. Lüst and S. Theisen, *Basic concepts of string theory*, Theoretical and Mathematical Physics, Springer, Heidelberg, Germany (2013), 10.1007/978-3-642-29497-6.

[80] F. Gliozzi, J. Scherk and D.I. Olive, *Supergravity and the Spinor Dual Model*, *Phys. Lett. B* **65** (1976) 282.

[81] F. Gliozzi, J. Scherk and D.I. Olive, *Supersymmetry, Supergravity Theories and the Dual Spinor Model*, *Nucl. Phys. B* **122** (1977) 253.

[82] A. Sen, *Strong - weak coupling duality in four-dimensional string theory*, *Int. J. Mod. Phys. A* **9** (1994) 3707 [hep-th/9402002].

[83] S. Kachru, M.B. Schulz and S. Trivedi, *Moduli stabilization from fluxes in a simple IIB orientifold*, *JHEP* **10** (2003) 007 [hep-th/0201028].

[84] R. Blumenhagen and E. Plauschinn, *Introduction to conformal field theory: with applications to String theory*, vol. 779 (2009), 10.1007/978-3-642-00450-6.

[85] T.W. Grimm, *The Effective action of type II Calabi-Yau orientifolds*, *Fortsch. Phys.* **53** (2005) 1179 [hep-th/0507153].

[86] P. Candelas and X. de la Ossa, *Moduli Space of Calabi-Yau Manifolds*, *Nucl. Phys. B* **355** (1991) 455.

[87] S. Kachru, R. Kallosh, A.D. Linde and S.P. Trivedi, *De Sitter vacua in string theory*, *Phys. Rev. D* **68** (2003) 046005 [hep-th/0301240].

[88] S. Kachru, J. Pearson and H.L. Verlinde, *Brane / flux annihilation and the string dual of a nonsupersymmetric field theory*, *JHEP* **06** (2002) 021 [hep-th/0112197].

[89] V. Balasubramanian, P. Berglund, J.P. Conlon and F. Quevedo, *Systematics of moduli stabilisation in Calabi-Yau flux compactifications*, *JHEP* **03** (2005) 007 [hep-th/0502058].

[90] J. Shelton, W. Taylor and B. Wecht, *Nongeometric flux compactifications*, *JHEP* **10** (2005) 085 [hep-th/0508133].

[91] J. Shelton, W. Taylor and B. Wecht, *Generalized Flux Vacua*, *JHEP* **02** (2007) 095 [hep-th/0607015].

[92] P. Candelas, *Yukawa Couplings Between (2,1) Forms*, *Nucl. Phys. B* **298** (1988) 458.

[93] A. Strominger, *SPECIAL GEOMETRY*, *Commun. Math. Phys.* **133** (1990) 163.

[94] S. Hosono, A. Klemm, S. Theisen and S.-T. Yau, *Mirror symmetry, mirror map and applications to complete intersection Calabi-Yau spaces*, *AMS/IP Stud. Adv. Math.* **1** (1996) 545 [hep-th/9406055].

[95] K. Kodaira, *Complex manifolds and deformation of complex structures*, Springer (2006).

[96] A. Strominger and E. Witten, *New Manifolds for Superstring Compactification*, *Commun. Math. Phys.* **101** (1985) 341.

[97] P. Candelas, X.C. De La Ossa, P.S. Green and L. Parkes, *A Pair of Calabi-Yau manifolds as an exactly soluble superconformal theory*, *AMS/IP Stud. Adv. Math.* **9** (1998) 31.

[98] P. Candelas, E. Derrick and L. Parkes, *Generalized Calabi-Yau manifolds and the mirror of a rigid manifold*, *Nucl. Phys. B* **407** (1993) 115 [hep-th/9304045].

[99] U.H. Danielsson and T. Van Riet, *What if string theory has no de Sitter vacua?*, *Int. J. Mod. Phys. D* **27** (2018) 1830007 [1804.01120].

[100] L. McAllister and F. Quevedo, *Moduli Stabilization in String Theory*, 2310.20559.

[101] Supernova Search Team collaboration, *Observational evidence from supernovae for an accelerating universe and a cosmological constant*, *Astron. J.* **116** (1998) 1009 [astro-ph/9805201].

[102] M. Gomez-Reino and C.A. Scrucca, *Locally stable non-supersymmetric Minkowski vacua in supergravity*, *JHEP* **05** (2006) 015 [hep-th/0602246].

[103] O. Lebedev, H.P. Nilles and M. Ratz, *De Sitter vacua from matter superpotentials*, *Phys. Lett. B* **636** (2006) 126 [hep-th/0603047].

[104] H. Abe, T. Higaki, T. Kobayashi and Y. Omura, *Moduli stabilization, F-term uplifting and soft supersymmetry breaking terms*, *Phys. Rev. D* **75** (2007) 025019 [hep-th/0611024].

[105] C.P. Burgess, R. Kallosh and F. Quevedo, *De Sitter string vacua from supersymmetric D terms*, *JHEP* **10** (2003) 056 [hep-th/0309187].

[106] G. Obied, H. Ooguri, L. Spodyneiko and C. Vafa, *De Sitter Space and the Swampland*, 1806.08362.

[107] I. Bena, J. Blåbäck, M. Graña and S. Lüst, *The tadpole problem*, *JHEP* **11** (2021) 223 [2010.10519].

[108] R. Blumenhagen, S. Moster and E. Plauschinn, *Moduli Stabilisation versus Chirality for MSSM like Type IIB Orientifolds*, *JHEP* **01** (2008) 058 [0711.3389].

[109] R. Bousso and J. Polchinski, *Quantization of four form fluxes and dynamical neutralization of the cosmological constant*, *JHEP* **06** (2000) 006 [hep-th/0004134].

[110] A. Giryavets, S. Kachru and P.K. Tripathy, *On the taxonomy of flux vacua*, *JHEP* **08** (2004) 002 [hep-th/0404243].

[111] O. DeWolfe, A. Giryavets, S. Kachru and W. Taylor, *Enumerating flux vacua with enhanced symmetries*, *JHEP* **02** (2005) 037 [hep-th/0411061].

[112] J.P. Conlon and F. Quevedo, *On the explicit construction and statistics of Calabi-Yau flux vacua*, *JHEP* **10** (2004) 039 [hep-th/0409215].

[113] E. Plauschinn and L. Schlechter, *Flux vacua of the mirror octic*, 2310.06040.

[114] A. Dubey, S. Krippendorf and A. Schachner, *JAXVacua – A Framework for Sampling String Vacua*, 2306.06160.

[115] C. Vafa, *The String landscape and the swampland*, hep-th/0509212.

[116] B.S. Acharya and M.R. Douglas, *A Finite landscape?*, hep-th/0606212.

[117] S.-T. Yau, *Open problems in geometry*, in *Proc. Symp. Pure Math*, vol. 54, pp. 1–28, 1993.

[118] V. Jejjala, W. Taylor and A. Turner, *Identifying equivalent Calabi–Yau topologies: A discrete challenge from math and physics for machine learning*, in *Nankai Symposium on Mathematical Dialogues: In celebration of S.S.Chern's 110th anniversary*, 2, 2022 [2202.07590].

[119] A. Hebecker, P. Henkenjohann and L.T. Witkowski, *Flat Monodromies and a Moduli Space Size Conjecture*, *JHEP* **12** (2017) 033 [1708.06761].

[120] T. Kobayashi and H. Otsuka, *Classification of discrete modular symmetries in Type IIB flux vacua*, *Phys. Rev. D* **101** (2020) 106017 [2001.07972].

[121] S. Kikuchi, T. Kobayashi, H. Otsuka, S. Takada and H. Uchida, *Modular symmetry by orbifolding magnetized $T^2 \times T^2$: realization of double cover of Γ_N*, *JHEP* **11** (2020) 101 [2007.06188].

[122] P. Betzler and E. Plauschinn, *Type IIB flux vacua and tadpole cancellation*, *Fortsch. Phys.* **67** (2019) 1900065 [1905.08823].

[123] O. DeWolfe, A. Giryavets, S. Kachru and W. Taylor, *Type IIA moduli stabilization*, *JHEP* **07** (2005) 066 [hep-th/0505160].

[124] H. Okada and M. Tanimoto, *Towards unification of quark and lepton flavors in A_4 modular invariance*, 1905.13421.

[125] G.-J. Ding, S.F. King, X.-G. Liu and J.-N. Lu, *Modular S_4 and A_4 symmetries and their fixed points: new predictive examples of lepton mixing*, *JHEP* **12** (2019) 030 [1910.03460].

[126] H. Okada and M. Tanimoto, *Modular invariant flavor model of A_4 and hierarchical structures at nearby fixed points*, *Phys. Rev. D* **103** (2021) 015005 [2009.14242].

[127] T. Kobayashi and H. Otsuka, *Challenge for spontaneous CP violation in Type IIB orientifolds with fluxes*, *Phys. Rev. D* **102** (2020) 026004 [2004.04518].

[128] T. Kobayashi and H. Otsuka, *Common origin of the strong CP and CKM phases in string compactifications*, *Phys. Lett. B* **807** (2020) 135554 [2002.06931].

[129] G.-J. Ding, S.F. King and X.-G. Liu, *Modular A_4 symmetry models of neutrinos and charged leptons*, *JHEP* **09** (2019) 074 [1907.11714].

[130] R. de Adelhart Toorop, F. Feruglio and C. Hagedorn, *Finite Modular Groups and Lepton Mixing*, *Nucl. Phys. B* **858** (2012) 437 [1112.1340].

[131] M. Demirtas, M. Kim, L. Mcallister and J. Moritz, *Vacua with Small Flux Superpotential*, *Phys. Rev. Lett.* **124** (2020) 211603 [1912.10047].

[132] Y. Honma and H. Otsuka, *Small flux superpotential in F-theory compactifications*, *Phys. Rev. D* **103** (2021) 126022 [2103.03003].

[133] H. Abe, T. Higaki and T. Kobayashi, *KKLT type models with moduli-mixing superpotential*, *Phys. Rev. D* **73** (2006) 046005 [hep-th/0511160].

[134] H. Abe, T. Higaki and T. Kobayashi, *Remark on integrating out heavy moduli in flux compactification*, *Phys. Rev. D* **74** (2006) 045012 [hep-th/0606095].

[135] F. Björkeroth, F.J. de Anda, I. de Medeiros Varzielas and S.F. King, *Towards a complete $A_4\times$ SU(5) SUSY GUT*, *JHEP* **06** (2015) 141 [1503.03306].

[136] I. Esteban, M.C. Gonzalez-Garcia, A. Hernandez-Cabezudo, M. Maltoni and T. Schwetz, *Global analysis of three-flavour neutrino oscillations: synergies and tensions in the determination of θ_{23}, δ_{CP}, and the mass ordering*, *JHEP* **01** (2019) 106 [1811.05487].

[137] I. Esteban, M.C. Gonzalez-Garcia, M. Maltoni, T. Schwetz and A. Zhou, *The fate of hints: updated global analysis of three-flavor neutrino oscillations*, *JHEP* **09** (2020) 178 [2007.14792].

[138] T. Kobayashi, H. Okada and Y. Orikasa, *Dark matter stability at fixed points in a modular A_4 symmetry*, 2111.05674.

[139] D.G. Markushevich, M.A. Olshanetsky and A.M. Perelomov, *Description of a Class of Superstring Compactifications Related to Semisimple Lie Algebras*, *Commun. Math. Phys.* **111** (1987) 247.

[140] Y. Katsuki, Y. Kawamura, T. Kobayashi, N. Ohtsubo, Y. Ono and K. Tanioka, *Z(N) ORBIFOLD MODELS*, *Nucl. Phys. B* **341** (1990) 611.

[141] T. Kobayashi and N. Ohtsubo, *Geometrical aspects of Z(N) orbifold phenomenology*, *Int. J. Mod. Phys. A* **9** (1994) 87.

[142] D. Lust, S. Reffert, W. Schulgin and S. Stieberger, *Moduli stabilization in type IIB orientifolds (I): Orbifold limits*, *Nucl. Phys. B* **766** (2007) 68 [hep-th/0506090].

[143] D. Lust, S. Reffert, E. Scheidegger, W. Schulgin and S. Stieberger, *Moduli Stabilization in Type IIB Orientifolds (II)*, *Nucl. Phys. B* **766** (2007) 178 [hep-th/0609013].

[144] A. Font, *Z(N) orientifolds with flux*, *JHEP* **11** (2004) 077 [hep-th/0410206].

[145] S. Reffert, *Toroidal Orbifolds: Resolutions, Orientifolds and Applications in String Phenomenology*, Ph.D. thesis, Munich U., 2006. hep-th/0609040.

[146] R. Blumenhagen, D. Lust and T.R. Taylor, *Moduli stabilization in chiral type IIB orientifold models with fluxes*, *Nucl. Phys. B* **663** (2003) 319 [hep-th/0303016].

[147] A.R. Frey and J. Polchinski, *N=3 warped compactifications*, *Phys. Rev. D* **65** (2002) 126009 [hep-th/0201029].

[148] J.F.G. Cascales and A.M. Uranga, *Chiral 4d string vacua with D branes and NSNS and RR fluxes*, *JHEP* **05** (2003) 011 [hep-th/0303024].

[149] G. Aldazabal, P.G. Camara, A. Font and L.E. Ibanez, *More dual fluxes and moduli fixing*, *JHEP* **05** (2006) 070 [hep-th/0602089].

[150] P. Mayr and S. Stieberger, *Threshold corrections to gauge couplings in orbifold compactifications*, *Nucl. Phys. B* **407** (1993) 725 [hep-th/9303017].

[151] A. Lascurain Orive, *Some presentations for\overline* $\{\Gamma\}(n)$, *Conformal Geometry and Dynamics of the American Mathematical Society* **6** (2002) 33.

[152] The GAP Group, *GAP – Groups, Algorithms, and Programming, Version 4.12.2*, 2022.

[153] L. Ford, *Automorphic Functions*, Chelsea Publishing Company (1972).

[154] H.P. Nilles, M. Ratz, A. Trautner and P.K.S. Vaudrevange, *$\mathcal{CP}$ violation from string theory*, *Phys. Lett. B* **786** (2018) 283 [1808.07060].

[155] P.P. Novichkov, J.T. Penedo, S.T. Petcov and A.V. Titov, *Generalised CP Symmetry in Modular-Invariant Models of Flavour*, *JHEP* **07** (2019) 165 [1905.11970].

[156] M. Dine, R.G. Leigh and D.A. MacIntire, *Of CP and other gauge symmetries in string theory*, *Phys. Rev. Lett.* **69** (1992) 2030 [hep-th/9205011].

[157] K.-w. Choi, D.B. Kaplan and A.E. Nelson, *Is CP a gauge symmetry?*, *Nucl. Phys. B* **391** (1993) 515 [hep-ph/9205202].

[158] F. Diamond and J.M. Shurman, *A first course in modular forms*, vol. 228, Springer (2005).

[159] D. Guest, K. Cranmer and D. Whiteson, *Deep Learning and its Application to LHC Physics*, *Ann. Rev. Nucl. Part. Sci.* **68** (2018) 161 [1806.11484].

[160] K. Albertsson et al., *Machine Learning in High Energy Physics Community White Paper*, *J. Phys. Conf. Ser.* **1085** (2018) 022008 [1807.02876].

[161] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby et al., *Machine learning and the physical sciences*, *Rev. Mod. Phys.* **91** (2019) 045002 [1903.10563].

[162] S. Forste and G. Honecker, *Rigid D6-branes on $T^6/(Z_2xZ_{2M}x\Omega R)$ with discrete torsion*, *JHEP* **01** (2011) 091 [1010.6070].

[163] A.L. Fradkov, *Early history of machine learning*, *IFAC-PapersOnLine* **53** (2020) 1385.

[164] C.M. Bishop and N.M. Nasrabadi, *Pattern recognition and machine learning*, vol. 4, Springer (2006).

[165] G. Cybenko, *Approximation by superpositions of a sigmoidal function*, *Mathematics of Control, Signals and Systems* **2** (1989) 303.

[166] G. Klambauer, T. Unterthiner, A. Mayr and S. Hochreiter, *Self-Normalizing Neural Networks*, in *Advances in Neural Information Processing Systems*, vol. 30, Curran Associates, Inc., 2017.

[167] K. Fukushima, *Cognitron: A self-organizing multilayered neural network*, *Biological Cybernetics* **20** (1975) 121.

[168] H. Mhaskar and C.A. Micchelli, *Approximation by superposition of sigmoidal and radial basis functions*, *Advances in Applied Mathematics* **13** (1992) 350.

[169] B.C. Allanach, D. Grellscheid and F. Quevedo, *Genetic algorithms and experimental discrimination of SUSY models*, *JHEP* **07** (2004) 069 [hep-ph/0406277].

[170] S. Abel and J. Rizos, *Genetic Algorithms and the Search for Viable String Vacua*, *JHEP* **08** (2014) 010 [1404.7359].

[171] Y.-H. He, *Deep-Learning the Landscape*, 1706.02714.

[172] Y.-H. He, *Machine-learning the string landscape*, *Phys. Lett. B* **774** (2017) 564.

[173] D. Krefl and R.-K. Seong, *Machine Learning of Calabi-Yau Volumes*, *Phys. Rev. D* **96** (2017) 066014 [1706.03346].

[174] F. Ruehle, *Evolving neural networks with genetic algorithms to study the String Landscape*, *JHEP* **08** (2017) 038 [1706.07024].

[175] E. Parr, *Machine Learning in String Theory*, Ph.D. thesis, Munich, Tech. U., 2020.

[176] E. Hirst, *Machine-Learning and Data Science Techniques in String and Gauge Theories*, Ph.D. thesis, London, City U., 3, 2023.

[177] Y.-H. He, *The Calabi–Yau Landscape: From Geometry, to Physics, to Machine Learning*, Lecture Notes in Mathematics (5, 2021), 10.1007/978-3-030-77562-9, [1812.02893].

[178] Y.-H. He, *From the String Landscape to the Mathematical Landscape: A Machine-Learning Outlook*, *Springer Proc. Math. Stat.* **396** (2022) 21 [2202.06086].

[179] J. Carifio, J. Halverson, D. Krioukov and B.D. Nelson, *Machine Learning in the String Landscape*, *JHEP* **09** (2017) 157 [1707.00655].

[180] E. Bernard, *Introduction to Machine Learning*, Wolfram Media (2021).

[181] D. Bank, N. Koenigstein and R. Giryes, *Autoencoders*, *Machine Learning for Data Science Handbook: Data Mining and Knowledge Discovery Handbook* (2023) 353.

[182] R. Nian, J. Liu and B. Huang, *A review On reinforcement learning: Introduction and applications in industrial process control*, *Computers & Chemical Engineering* **139** (2020) 106886.

[183] F. Gmeiner, R. Blumenhagen, G. Honecker, D. Lust and T. Weigand, *One in a billion: MSSM-like D-brane statistics*, *JHEP* **01** (2006) 004 [hep-th/0510170].

[184] G.J. Loges and G. Shiu, *134 billion intersecting brane models*, *JHEP* **12** (2022) 097 [2206.03506].

[185] M.R. Douglas and W. Taylor, *The Landscape of intersecting brane models*, *JHEP* **01** (2007) 031 [hep-th/0606109].

[186] W. He, T. Li and R. Sun, *The complete search for the supersymmetric Pati-Salam models from intersecting D6-branes*, *JHEP* **08** (2022) 044 [2112.09632].

[187] J. Halverson, B. Nelson and F. Ruehle, *Branes with Brains: Exploring String Vacua with Deep Reinforcement Learning*, *JHEP* **06** (2019) 003 [1903.11616].

[188] G.J. Loges and G. Shiu, *Breeding Realistic D-Brane Models*, *Fortsch. Phys.* **70** (2022) 2200038 [2112.08391].

[189] T. Li, A. Mansha and R. Sun, *Revisiting the supersymmetric Pati–Salam models from intersecting D6-branes*, *Eur. Phys. J. C* **81** (2021) 82 [1910.04530].

[190] A. Mansha, T. Li and L. Wu, *The hidden sector variations in the $\mathcal{N} = 1$ supersymmetric three-family Pati–Salam models from intersecting D6-branes*, *Eur. Phys. J. C* **83** (2023) 1067 [2303.02864].

[191] A. Mütter, E. Parr and P.K.S. Vaudrevange, *Deep learning in the heterotic orbifold landscape*, *Nucl. Phys. B* **940** (2019) 113 [1811.05993].

[192] H. Otsuka and K. Takemoto, *Deep learning and k-means clustering in heterotic string vacua with line bundles*, *JHEP* **05** (2020) 047 [2003.11880].

[193] R. Deen, Y.-H. He, S.-J. Lee and A. Lukas, *Machine learning string standard models*, *Phys. Rev. D* **105** (2022) 046001 [2003.13339].

[194] E. Escalante-Notario, I. Portillo-Castillo and S. Ramos-Sanchez, *Autoencoding heterotic orbifolds with arbitrary geometry*, 2212.00821.

[195] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez et al., *Attention is All you Need*, in *Advances in Neural Information Processing Systems*, vol. 30, Curran Associates, Inc. [1706.03762v7].

[196] R. Blumenhagen and E. Plauschinn, *Intersecting D-branes on shift Z(2) x Z(2) Orientifolds*, *JHEP* **08** (2006) 031 [hep-th/0604033].

[197] K. Becker, M. Becker and A. Strominger, *Five-branes, membranes and nonperturbative string theory*, *Nucl. Phys. B* **456** (1995) 130 [hep-th/9507158].

[198] E. Witten, *D-branes and K-theory*, *JHEP* **12** (1998) 019 [hep-th/9810188].

[199] A.M. Uranga, *D-brane probes, RR tadpole cancellation and K-theory charge*, *Nucl. Phys. B* **598** (2001) 225 [hep-th/0011048].

[200] E. Witten, *An SU(2) Anomaly*, *Phys. Lett. B* **117** (1982) 324.

[201] L.E. Ibanez, F. Marchesano and R. Rabadan, *Getting just the standard model at intersecting branes*, *JHEP* **11** (2001) 002 [hep-th/0105155].

[202] R. Blumenhagen, M. Cvetic, F. Marchesano and G. Shiu, *Chiral D-brane models with frozen open string moduli*, *JHEP* **03** (2005) 050 [hep-th/0502095].

[203] R. Blumenhagen, F. Gmeiner, G. Honecker, D. Lust and T. Weigand, *The Statistics of supersymmetric D-brane models*, *Nucl. Phys. B* **713** (2005) 83 [hep-th/0411173].

[204] M. Zaheer, S. Kottur, S. Ravanbakhsh, B. Poczos, R.R. Salakhutdinov and A.J. Smola, *Deep sets*, *Advances in neural information processing systems* **30** (2017) [1703.06114].

[205] P. Dufter, M. Schmitt and H. Schütze, *Position information in transformers: An overview*, *Computational Linguistics* **48** 733 [2102.11090v2].

[206] L.H. Gilpin, D. Bau, B.Z. Yuan, A. Bajwa, M. Specter and L. Kagal, *Explaining explanations: An overview of interpretability of machine learning*, in *2018 IEEE 5th International Conference on data science and advanced analytics (DSAA)*, pp. 80–89, IEEE, 2018 [1806.00069].

[207] N. Geneva and N. Zabaras, *Transformers for modeling physical systems*, *Neural Networks* **146** (2022) 272 [2010.03957].