Title	Generative AI Exploration of MSSM-Like Models in E8×E8 Heterotic String Z6-II Orbifold Compactifications
Author(s)	見村, 優太
Degree Grantor	北海道大学
Degree Name	博士(理学)
Dissertation Number	甲第16190号
Issue Date	2024-12-25
DOI	https://doi.org/10.14943/doctoral.k16190
Doc URL	https://hdl.handle.net/2115/94062
Type	doctoral thesis
File Information	Yuta_Mimura.pdf

博士学位論文

Generative AI Exploration of MSSM-Like Models in $E_8 \times E_8$ Heterotic String Z_6-II Orbifold Compactifications
(Z_6-II オービフォールドコンパクト化における $E_8 \times E_8$ ヘテロティック弦理論での MSSM 様モデルの生成 AI による探索に関する研究)

Yuta Mimura
見村　優太

Department of Physics, Hokkaido University, Sapporo 060-0810, Japan
北海道大学大学院理学院宇宙理学専攻

December, 2024
2024 年 12 月

Acknowledgement

This study was supported by JST SPRING, Grant Number JPMJSP2119. We would like to thank my supervisor, Associate Professor Kazuhiko Suehiro at Hokkaido University for his advice on this paper and review of it throughout this study. We would like to thank everyone at the laboratory for their guidance and consultation. For enhancing the clarity and accuracy of our paper, as well as for reviewing the code utilized in our experiments, we employed the assistance of ChatGPT [16].

Abstract

String theory is a leading candidate for unifying gravity with quantum mechanics, avoiding quantum anomalies. Due to its ten-dimensional nature, six extra dimensions must be compactified to unobservable scales to be consistent with our four-dimensional reality. This compactification process leads to a vast landscape of possible four-dimensional string models, estimated to be on the order of 10^{1500}. Among these, MSSM-like models, which resemble the Minimal Supersymmetric Standard Model, are of particular interest but represent only a small fraction of the landscape.

Traditional random scanning methods are often employed to search for these MSSM-like models within the string landscape. However, the relationship between compactification parameters and resulting spectra is complex, making this process computationally challenging. Recent research has proposed using autoencoders (AE) and Variational Autoencoders (VAE) to map the high-dimensional parameter space onto a lower-dimensional space, improving the search efficiency for MSSM-like models.

This dissertation explores the use of VAE-based generative models to generate new MSSM-like data, demonstrating their potential to efficiently explore the string landscape. We also introduce the "phased data augmentation" technique to enhance the training of these models on sparse datasets. The effectiveness of this approach is evaluated through experiments, with results showing promising advancements in the search for viable MSSM-like models.

Contents

Chapter 1

Introduction

String theory has been actively researched as one of the theories that can integrate gravitational theory without quantum anomalies. However, due to the consistency requirements of string theory, it is inherently a ten-dimensional theory, necessitating that the excess six dimensions not be observable at low energy scales to align with our four-dimensional reality. A widely adopted theoretical method to achieve this involves compactifying the extra dimensions to scales that are unobservable, a process known as compactification. This compactification is carried out based on symmetries, and by specifying parameters that dictate how these symmetries are applied, one can derive a four-dimensional effective theory with defined symmetries and spectra, provided the conditions for compactification are met. However, it was estimated soon after the advent of string theory that the number of inequivalent four-dimensional string models is at least on the order of 10^{1500}, a significantly large number [14]. Among these, models that resemble the Minimal Supersymmetric extensions of the Standard Model (MSSM-like models) represent only a fraction, and even within this group, models are not unique. This vast array of four-dimensional effective string theories is referred to as the string landscape [14].

The exploration of MSSM-like four-dimensional effective string theories has been a significant area of research within the field of string theory, as referenced in studies [1, 4, 5]. The driving force for these investigations arises from the endeavor to identify models that align with all existing experimental and observational data when framed within the Standard Model (SM) or the Minimal Supersymmetric Standard Model (MSSM), facilitated through the method of compactification. The discovery of a realistic MSSM-like model through compactification in string theory is considered significant, as it may potentially have applications in high-energy theoretical frameworks. For instance, identifying MSSM-like models that exhibit new phenomena or properties distinct from those found in conventional quantum field theory models could lead to the discovery of common characteristics that address theoretical issues within the SM or MSSM. This, in turn, could unveil new mechanisms that provide solutions to these theoretical challenges. Ultimately, such discoveries have the potential to pave the way for a deeper understanding of new physics, expanding our knowledge and possibly leading to groundbreaking insights into the nature of the universe [14].

Research in the string landscape primarily focuses on the gauge symmetries of the MSSM, $SU(3)_C \times SU(2)_L \times U(1)_Y$, the representations suitable for three generations of quarks and leptons, and at least one set of Higgs particles. The four-dimensional string landscape contains an immense number of non-equivalent models, making a comprehensive search impractical. As a result, researchers are unable to systematically categorize the entire string landscape. Due to the impracticality of a comprehensive classification of 4D string theory models, recent studies have shifted towards exploring different models within a small sector of the string landscape by randomly scanning for MSSM-like models, narrowing down the parameter space or specifying the manner of compactification [14].

The construction of four-dimensional string theory models through compactification typically involves a process where around 100 compactification parameters are identified. These parameters define aspects such as the geometry of the six-dimensional compactification space, fluxes, or world-sheet parameters. It is essential to check whether these parameters satisfy certain consistency conditions, such as quantization conditions, Bianchi identities, or the world-sheet modular invariance of the one-loop partition function. Therefore, random scans are frequently utilized in the following manner. These involve randomly selecting compactification parameters based on specific symmetries and within certain parameter ranges, followed by verifying whether they satisfy the consistency conditions. If found consistent, the gauge group and matter spectrum of the resulting four-dimensional string theory model are calculated. This model is then compared to see if it resembles the MSSM. While this conventional scanning method can identify MSSM-like models, the relationship between the compactification parameters and the resulting spectra is generally complex, leading to significant computational challenges. It is often difficult to pinpoint which areas of the compactification parameter space are more likely to yield a high number of MSSM-like models. Furthermore, a notable drawback of this approach is the need to search through an extensive parameter space to find a relatively small number of MSSM-like models [14].

In recent research, an approach employing autoencoder (AE) [7] has been proposed to map the high-dimensional compactification parameter space onto a two-dimensional continuous space, where islands of MSSM-like models emerge experimentally. This method suggests that by identifying these islands on the two-dimensional plane and subsequently exploring the corresponding original subspace, search efficiency for MSSM-like models can be enhanced. Unlike traditional exhaustive searches across specific regions of the string landscape, this technique involves conducting random scans over a subset of non-equivalent 4D string models to obtain a coarse sample of a specific area within the string landscape. This sample is then used to train an AE, which identifies islands on the corresponding two-dimensional continuous space where MSSM-like models are clustered. The method aims to pinpoint the original compactification parameter subspaces that are more likely to contain MSSM-like models with a higher probability. Experimental results showcase a two-dimensional plot, revealing that, despite the lack of explicit labeling, MSSM-like models are predominantly distributed within certain islands among multiple clusters of 4D models. This finding illustrates the potential of using AE to streamline the search process within the vast string landscape, making the identification of realistic MSSM-like models more efficient [14].

However, identifying MSSM-like models remains challenging, and extracting physically meaningful insights from multiple models is difficult. In response, we have pursued research aimed at leveraging generative AI to efficiently explore insights based on the similarities and differences among specific MSSM models. As an initial step in this direction within our research scope, we employed a Variational Auto Encoder (VAE) [12]-based generative AI, related to AE, demonstrating that it can generate new data not present in the training data and produce similar data, even with limited MSSM data. It is well-known in the field of deep learning that training AI with limited data is challenging, a difficulty that extends to the domain of image generation where research in generative AI is particularly active. Although there has been research proposing methods to address these challenges, focusing mainly on certain models, efforts specifically targeting likelihood-based models like VAEs related to AEs have been lacking. Consequently, I have previously developed a method to efficiently train AI on limited data, specifically for use with VAE-based models. In our experiments, we applied the phased approach to limited MSSM data, demonstrating that generative AI could capture the characteristics of the data without label information and generate new MSSM data not present in the training dataset. We were able to produce data with identical gauge symmetries but somewhat differing compactification parameters in a 4D string model.

The advantages of our study over the random scan method in orbifolder [15] are as follows: 1. Compared to random scans in tools like orbifolder, our approach can more readily identify data with at least similar gauge symmetries to a given data. If we use orbifolder, it is necessary to prepare an extensive list of data, including those similar to a given data, which proves challenging due to the difficulty of exhaustively finding MSSM-like data among a vast number of combinations of compactification parameters through random scanning. 2. By examining similarities and differences between MSSM models through interpolation among two or more points, we can efficiently explore insights into the relationships between MSSM models. 3. While orbifolder is limited to specific datasets, specific compactification likeZ_6-II orbifold compactification, our paper presents a methodology potentially applicable to any physical data, suggesting broader applicability. 4. Unlike orbifolder, which performs calculations based on a predetermined procedure, generative AI learns the distribution of data, allowing it to extract useful information from the relationships within the data. 5. Similar to previous studies on AE, our approach has the potential to find MSSM-like models with a higher probability, thus enhancing efficiency.

Regarding the utility of our research compared to previous studies on AE: While prior AE research suggested the presence of islands resembling MSSM models within the encoded dimensional compression space (latent space), indicating that unsupervised learning could capture data characteristics, it was limited to identifying these islands without generating similar data. In contrast, VAE-based models, through stochastic reconstruction and generation with regularization in the latent space, facilitate learning that clusters data with similar properties closer together in the latent space. We were able to generate data with identical gauge symmetries but slightly different compactification parameters in a 4D string model, supporting its advantages, particularly in identifying similarities in gauge symmetries. This development supports the benefits outlined in point 2 above.

Deep learning techniques such as supervised learning, unsupervised learning, and reinforcement learning have recently been employed in the field of particle physics. Supervised learning leverages label information for the efficient execution of classification tasks. Unsupervised learning, on the other hand, enables the learning of underlying relationships within data without the use of label information; however, methods like AE fall short of achieving complete continuity in their latent space representations, making sampling challenging. Reinforcement learning offers the flexibility to set reward systems and goals freely but requires the appropriate configuration of these systems and objectives. In this study, we utilize generative AI, learning the distribution of data without label information and constructing a more dense latent space that captures similarities and differences. This allows for efficient exploration through sampling of similar data within this space. In this study, while this application was specifically applied to certain particle physics data, it holds potential for broader application across different particle physics datasets. Our work represents a first step in the utilization of generative AI within the realm of particle physics. Furthermore, research applying generative AI to discrete numerical data, as opposed to image or text data, has not been as prevalent, making this study a notable example in this direction as well.

The remainder of this chapter is organized as follows: Chapter 2 begins by providing an overview of the orbifold compactification of $E_8 \times E_8$ heterotic string theory, which is the focus of this study. Chapter 3 reviews the methodological developments undertaken in preparation for experiments on generating AI data for MSSM-like compactification parameters. Preparing MSSM-like data through traditional random scanning has a difficulty and results in only a limited number of datasets. Moreover, training generative AI with limited data is generally challenging. In Chapter 4, we explain our previous work for addressing these challenges, "phased data augmentation" technique developed for effectively training generative AI, specifically VAE-based models, on limited datasets. In Chapter 5, we describe the main part of this study, including the application of phased data augmentation to generate similar MSSM-like data and the experimental results from generating new data that differs in the spectrum after compactification. Finally, Chapter 6 summarizes and discusses the position, results, and potential future directions of this research.

Chapter 2

Review of Orbifold Compactification of $E_8 \times E_8$ Heterotic String Theory

In this chapter, we provide a brief overview of the compactification of $E_8 \times E_8$ heterotic strings and the spectrum of the resulting four-dimensional effective theory. This chapter is based on [8].

2.1 The Standard Model of Particle Physics

The Standard Model (SM) of particle physics describes electromagnetic, weak, and strong interactions with remarkable precision across a wide range of energy scales. The SM is a quantum field theory (QFT) based on the following gauge symmetries:

$$SU(3)_C \times SU(2)_L \times U(1)_Y \tag{2.1}$$

Here, $SU(3)_C$ represents the strong interaction, while $SU(2)_L \times U(1)_Y$ describes the electroweak interaction.

The fermions in the SM consist of three generations of quarks q and leptons L, which are represented as complex two-component Weyl spinors. Along with the Higgs scalar field, the quantum numbers of the fermions under these gauge groups are summarized in Table 2.1. Here, the subscripts L and R indicate whether the fields transform as doublets or singlets under the $SU(2)_L$ gauge group. The subscript i indicates which of the three generations the particle belongs to. The bar notation represents the conjugate representation of a field. The left-handed quarks and leptons are denoted as follows:

$$Q_L = (U_L^i, D_L^i), \quad L = (\nu_L^i, E_L^i), \quad i = 1, 2, 3 \tag{2.2}$$

Here, U, D, ν, and E represent the up-type quarks, down-type quarks, neutrinos, and charged leptons, respectively.

The SM exhibits a chiral symmetry, where the left-handed and right-handed components of the fermions transform under different electroweak quantum numbers and satisfy different equations. Due to this chiral symmetry, Dirac mass terms for fermions are forbidden at high energies. As a result, the fermions remain massless at high energies. However, as the energy decreases and the Higgs field acquires the following vacuum expectation value (VEV), the electroweak symmetry is spontaneously broken:

$$\langle H^0 \rangle \approx 170\,\text{GeV} \tag{2.3}$$

This spontaneous symmetry breaking generates masses for the $W^\pm$ and Z bosons, as well as for the quarks and leptons through Yukawa interactions.

Despite the remarkable success of the SM in explaining a wide range of experimental data, it leaves several important questions unanswered, indicating the need for new physics at more fundamental scales. Notably, the SM does not incorporate gravitational interactions, which suggests that it is an effective theory with a potential cutoff at the Planck scale:

$$M_p = \frac{1}{\sqrt{8\pi G_N}} = 1.2 \times 10^{19}\,\text{GeV} \tag{2.4}$$

where G_N is Newton's gravitational constant. The challenge of reconciling gravity with quantum mechanics remains one of the most pressing issues in theoretical physics today.

	$SU(3)_C$	$SU(2)_L$	$U(1)_Y$
Q_L^i	$\mathbf{3}$	$\mathbf{2}$	$+\frac{1}{6}$
U_R^i	$\bar{\mathbf{3}}$	$\mathbf{1}$	$+\frac{2}{3}$
D_R^i	$\bar{\mathbf{3}}$	$\mathbf{1}$	$-\frac{1}{3}$
L_L^i	$\mathbf{1}$	$\mathbf{2}$	$-\frac{1}{2}$
E_R^i	$\mathbf{1}$	$\mathbf{1}$	-1
H	$\mathbf{1}$	$\mathbf{2}$	$-\frac{1}{2}$

Table 2.1: Gauge quantum numbers of the SM fermions and the Higgs field.

2.1.1 Grand Unified Theories (GUTs)

Grand Unified Theories (GUTs) extend the Standard Model by attempting to unify the electromagnetic, weak, and strong interactions under a single gauge group at high energy scales. One of the key advantages of GUTs is their ability to unify quarks and leptons into larger, more symmetric multiplets. Additionally, GUTs provide a natural framework for implementing the seesaw mechanism, which can generate the small neutrino masses that are observed experimentally. The unification of gauge couplings in GUTs is further enhanced in supersymmetric versions, where the coupling constants of the different forces converge at a single energy scale.

Despite these appealing features, GUTs do not address all fundamental questions. For instance, they do not incorporate gravitational interactions, nor do they explain the existence of exactly three generations of fermions, which limits their ability to fully account for the structure of fermion masses and mixings. These limitations motivate the exploration of even more comprehensive theories that might integrate gravity and address the full scope of particle physics.

2.1.2 The Fine-Tuning Puzzles in the Standard Model

One of the major issues the Standard Model (SM) faces when considered without extending it to a theory beyond the SM is the problem of the naturalness and fine-tuning of its parameters. Specifically, this issue arises because, when attempting to explain the measured values of fermion masses in the context of the SM, one must account for large Planck-scale corrections to the masses. The result is that the measured electroweak-scale masses seem to come from an unnatural cancellation of these large corrections, leading to the fine-tuning problem.

At high energies, the Standard Model respects both $SU(2)$ symmetry and chiral symmetry, resulting in massless fermions. However, as the energy scale drops below the electroweak scale, the spontaneous breaking of electroweak symmetry results in the loss of $SU(2)$ symmetry, and mass terms are generated for fermions. At this lower scale, chiral symmetry is also broken.

The Higgs vacuum expectation value (VEV) is inserted into the Higgs coupling terms as a constant, generating mass terms for the gauge bosons and fermions. Through the process of spontaneous symmetry breaking, the $SU(2)$ symmetry of the Higgs doublet is broken, and this symmetry breaking simultaneously gives mass to the gauge bosons, breaking the electroweak symmetry.

The Higgs scalar potential is given by the following expression:

$$V(H) = -\mu^2 H^\dagger H + \lambda (H^\dagger H)^2$$

The Higgs field obtains a vacuum expectation value (VEV) when the potential reaches a minimum, given by:

$$|H|^2 = \frac{\mu^2}{2\lambda}$$

As the energy scale decreases, the potential takes a nonzero value at the Higgs vacuum expectation value, leading to spontaneous symmetry breaking. By inserting the Higgs VEV into the Higgs Yukawa coupling terms, mass terms for the gauge bosons and fermions are generated. The Yukawa coupling terms are the following:

$$\mathcal{L}_{\text{Yuk}} = Y_U^{ij} \overline{Q}_L^i U_R^j H^* + Y_D^{ij} \overline{Q}_L^i D_R^j H + Y_L^{ij} \overline{L}_L^i E_R^j H + h.c.$$

This VEV is proportional to μ^2, and by substituting it into the interaction terms involving other fields, we obtain the corresponding mass terms.

However, the parameter μ^2, which represents the mass term of the Higgs field, receives quadratically divergent one-loop corrections as shown in Fig. 2.1.

These corrections are ultraviolet divergent and receive multiple contributions from the cutoff scale M_p. As a result, the observed small mass of the Higgs boson appears to be fine-tuned,

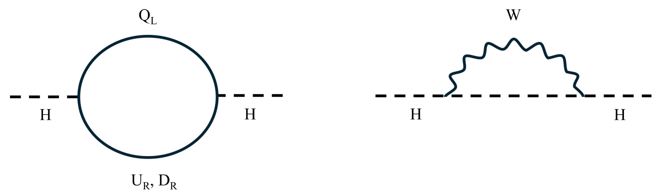

Figure 2.1: Quadratic one-loop corrections to the Higgs mass.

meaning that the difference between these large values must cancel out to an extremely precise degree, which seems unnatural given the large orders of magnitude involved.

One effective way to resolve this hierarchy and fine-tuning problem is to introduce $N = 1$ supersymmetry (SUSY). In a supersymmetric framework, the large loop corrections to the Higgs mass are canceled by the contributions from the loops of the supersymmetric partner particles.

Furthermore, the introduction of SUSY plays a crucial role in string theory, as it allows for the inclusion of fermions, which will be discussed in later sections.

2.1.3 A Fifth Dimension and Kaluza–Klein Compactification

In our exploration of physics beyond the Standard Model, one compelling idea is the existence of extra spatial dimensions. The simplest scenario involves a fifth dimension, which can be compactified into a small circle, proposed in the Kaluza–Klein (KK) theory. This section introduces the concept of KK compactification in five-dimensional (5d) theories and explains why such extra dimensions remain unobservable at low energies. While the ideas here are generalizable to more dimensions, we focus on the simplest cases involving scalar fields and gravitons.

The 5d Scalar Field

The action for a five-dimensional (5D) free massless scalar field $\phi(x^M)$, where $M = 0, \ldots, 4$, is given by

$$S_{5\mathrm{d},\phi} = \int d^5x \left(-\frac{1}{2} \partial_M \phi \, \partial^M \phi \right), \tag{2.5}$$

Here, we have omitted an appropriate dimensionful coefficient. The fifth dimension, denoted as $y \equiv x^4$, is compactified on a circle of radius R, with the coordinate y satisfying the following boundary condition:

$$\phi(x^\mu, y) = \phi(x^\mu, y + 2\pi R),$$

where $\mu = 0, \ldots, 3$. Given this periodicity, the 5D scalar field can be expanded in a Fourier series as:

$$\phi(x^\mu, y) = \sum_{k \in \mathbb{Z}} \phi_k(x^\mu) e^{iky/R}. \tag{2.6}$$

Substituting this expansion into the original 5D action and integrating over the fifth dimension, we obtain the following four-dimensional (4D) action:

$$S_{4\mathrm{d},\phi} = (2\pi R)\int d^4x \left(-\frac{1}{2}\partial_\mu\phi_0\,\partial^\mu\phi_0\right) - (2\pi R)\sum_{k=1}^{\infty}\int d^4x\left(\partial_\mu\phi_k\,\partial^\mu\phi_k^* + \frac{k^2}{R^2}\phi_k\phi_k^*\right), \tag{2.7}$$

where $\phi_k^* = \phi_{-k}$.

This theory includes a massless scalar field ϕ_0 and an infinite tower of massive scalar fields, known as Kaluza-Klein (KK) modes, with masses given by

$$m_k^2 = \frac{k^2}{R^2}, \quad k = 1, 2, 3, \ldots$$

At energies much lower than $1/R$, only the zero mode ϕ_0 is observable, effectively concealing the extra dimension.

The 5d Gravity

Next, we consider the Kaluza-Klein (KK) compactification of 5-dimensional gravity and its implications for 4-dimensional physics. The 4-dimensional Einstein-Hilbert action is given by

$$S_{4d} = \frac{1}{2\kappa_4^2}\int d^4x\sqrt{-g}R_{4d}, \tag{2.8}$$

where $g = \det(g_{\mu\nu})$, $\frac{1}{\kappa_4^2} = \frac{M_p^2}{8\pi}$, and R_{4d} is the 4-dimensional scalar curvature. Here, the 5-dimensional gravity action is expressed as

$$S_{5d} = \frac{M_5^3}{2}\int d^5x\sqrt{-G}R_{5d}, \tag{2.9}$$

where $G = \det(G_{MN})$, with $M, N = 0, \ldots, 4$, and R_{5d} is the 5-dimensional scalar curvature. We compactify $y \equiv x^4$ as a circle, then the Fourier-expanded metric is given by

$$G_{MN}(x^\mu, y) = \sum_{k\in\mathbb{Z}} G_{MN}^k(x^\mu)e^{iky/R}. \tag{2.10}$$

Substituting this into the action yields terms corresponding to the massless particles and a sum of massive graviton KK modes. The massless states correspond to the 4-dimensional graviton $g_{\mu\nu}$, a vector boson A_μ, and a scalar σ, known as the radion. The relation between these fields and the zero-mode metric is expressed as

$$\left(G_{MN}^0\right) = e^{\sigma/3}\begin{pmatrix} g_{\mu\nu}(x) + e^{-\sigma}A_\mu A_\nu & e^{-\sigma}A_\mu \\ e^{-\sigma}A_\nu & e^{-\sigma}\end{pmatrix}. \tag{2.11}$$

Here, the vector boson A_μ arises from the $G_{\mu 4}$ component of the 5-dimensional metric, and it exhibits a $U(1)$ gauge symmetry due to local reparametrizations on S^1 of the form $x^4 \to x^4 + \lambda(x^\mu)$, leading to the gauge transformation

$$A_\mu \to A_\mu - \partial_\mu\lambda(x^\mu). \tag{2.12}$$

The scalar field radion σ has a vanishing potential, and its vacuum expectation value (VEV) is arbitrary, parametrizing a microscopic parameter of the configuration, specifically the radius of S^1.

Despite the elegance of this approach, attempts to incorporate both gravity and the Standard Model (SM) gauge group via higher-dimensional theories often fail due to the absence of charged chiral fermions. However, in string theory, compactifications are successful in incorporating both the gravitational and SM gauge interactions, resolving this issue through mechanisms unique to string theory.

2.2 Introduction to Supersymmetry and the MSSM

Supersymmetry (SUSY) is a profound symmetry that unifies fermions and bosons. In string theory, SUSY plays a crucial role in the introduction of fermions and is essential for preventing the occurrence of tachyonic scalars, which can lead to instabilities in the theory. Additionally, in four-dimensional theories, SUSY is important for avoiding the electroweak hierarchy problem, making it a key feature in theories that extend the Standard Model (SM) to higher energy scales.

In this chapter, we first explain the basic structure of SUSY algebra. We then proceed to discuss the Minimal Supersymmetric Standard Model (MSSM), which is the minimal extension of the SM that incorporates SUSY and is the focus of this study.

2.2.1 Supersymmetry Algebra

Here, we introduce the concept of supersymmetry (SUSY), which is a symmetry between bosons and fermions. A SUSY generator Q acts as follows:

$$Q(\text{fermion}) = \text{boson}, \quad Q(\text{boson}) = \text{fermion}. \tag{2.13}$$

This relationship requires an equal number of fermionic and bosonic degrees of freedom. A simple supersymmetric action can be written using a Weyl fermion ψ_α and a complex scalar field Φ as follows:

$$S = \int d^4x \left(-\partial^\mu \Phi^* \partial_\mu \Phi - i\bar{\psi}\bar{\sigma}^\mu \partial_\mu \psi \right). \tag{2.14}$$

In this physical system, there exists a conserved current, known as the supercurrent, which is conserved on-shell, meaning it is conserved when the equations of motion are satisfied. The supercurrent is given by:

$$J^\mu_\alpha = \left(\partial_\nu \Phi^* \sigma^\nu \psi\right)_\alpha, \quad \partial_\mu J^\mu_\alpha = 0, \tag{2.15}$$

where α denotes the spinor index. The conservation of the supercurrent implies the conservation of the corresponding (super)charges:

$$Q_\alpha = \int d^3x\, J^0_\alpha, \quad \bar{Q}_{\dot{\alpha}} = \int d^3x\, \bar{J}^0_{\dot{\alpha}}. \tag{2.16}$$

These charge generators are fermionic because they contain spinor components and transform as Weyl spinors under the Lorentz group. Their algebra is generated by anticommutation relations, rather than commutators. The anticommutator of Q_α and $\bar{Q}_{\dot{\alpha}}$ must be a bosonic conserved quantity. The only candidate is the spacetime momentum P_μ, which, considering the spinorial structure of the anticomutator, must be contracted with $\sigma^\mu_{\alpha\dot{\alpha}}$. This leads to the following (super)algebra:

$$\{Q_\alpha, \bar{Q}_{\dot{\alpha}}\} = 2\sigma^\mu_{\alpha\dot{\alpha}} P_\mu. \tag{2.17}$$

The other (anti)commutators vanish:

$$\{Q_\alpha, Q_\beta\} = \{\bar{Q}_{\dot{\alpha}}, \bar{Q}_{\dot{\beta}}\} = [Q_\alpha, P_\mu] = [\bar{Q}_{\dot{\alpha}}, P_\mu] = 0. \tag{2.18}$$

Here, Q_α and $\bar{Q}_{\dot{\alpha}}$ are not simply generators of an internal symmetry independent of spacetime symmetries; rather, they intertwine with the Poincaré symmetry, connecting internal and spacetime symmetries.

In the following, we will primarily focus on the $4d\, N = 1$ SUSY algebra relevant to this work, while the cases with $N > 1$ will not be discussed in detail.

2.2.2 Low-Energy Supersymmetry and the MSSM

The superpotential is a component of the SUSY version of the Lagrangian that determines much of the interactions in a SUSY theory. Importantly, in 4d $N = 1$ SUSY, the superpotential does not receive renormalization corrections, meaning that the scalar masses remain stable even at high energy scales. This stability of the scalar masses provides a compelling reason to consider SUSY extensions of the Standard Model (SM) as a solution to the electroweak hierarchy problem.

These considerations require $N = 1$ SUSY, because extended SUSY models contain vector-like combinations of fermions, which would necessarily lead to non-chiral theories. The Minimal Supersymmetric Standard Model (MSSM) is an extension of the SM using $N = 1$ SUSY, where SM particles (gauge bosons, quarks, leptons, and Higgs doublets) are supplemented with their SUSY partners (gauginos, squarks, sleptons, and higgsinos).

In the MSSM, each particle is paired with its SUSY partner to form a supermultiplet. A supermultiplet is a multiplet of particles that transform into each other under SUSY transformations. The auxiliary fields introduced in the formalism of supermultiplet theory do not carry physical degrees of freedom and are neglected in the final physical stage where the supermultiplets are identified. The spectrum of the MSSM is summarized in the Table 2.2 where the tilde ($\sim$) denotes the SUSY partner of each particle.

Furthermore, due to the commutation of SUSY generators with gauge generators, each particle and its SUSY partner transform identically under the gauge group $SU(3)_C \times SU(2)_L \times U(1)_Y$. The Higgs multiplets H_u and H_d have opposite hypercharges, $Y = +\frac{1}{2}$ and $Y = -\frac{1}{2}$, respectively. This configuration is the minimal setup necessary to generate masses for the up-type quarks, down-type quarks, and leptons.

Table 2.2: Spectrum of the MSSM classified by spin and supermultiplet type

Supermultiplet Type	**Spin 1**	**Spin 1/2**	**Spin 0**
Vector Multiplets	g, $W^{\pm}$, W^0, B	$\tilde{g}$, $\tilde{W}^{\pm}$, $\tilde{W}^0$, $\tilde{B}$	-
Chiral Multiplets	-	q_L, q_R, l_L, l_R, $\tilde{H}_u$, $\tilde{H}_d$	$\tilde{q}_L$, $\tilde{q}_R$, $\tilde{l}_L$, $\tilde{l}_R$, H_u, H_d

2.3 Introduction to Orbifold Compactification in $E_8 \times E_8$ Heterotic String Theory

2.3.1 String Theory

In the Standard Model (SM) and its Grand Unified Theories (GUT) extensions, there are issues related to the possible divergence of coupling constants at high energies. Supersymmetry (SUSY) can partially control these ultraviolet divergences, but when considering supergravity theories, the issue of non-renormalizability still persists. Moreover, Einstein's theory of gravity, when treated as a quantum field theory, is also non-renormalizable.

String theory, on the other hand, offers a framework that can avoid these quantum divergences. It naturally incorporates the SM and has the potential to unify gravity with the SM without any inconsistencies.

Unlike point particles, strings are one-dimensional objects with a characteristic length $L_s = \frac{1}{M_s}$, where M_s is the string scale. In many models, this string scale is set to be much larger than the experimentally verified energy scales, typically on the order of the 4-dimensional Planck scale. At low energy limits, the resolution of the string becomes coarse, making it appear to behave like a point particle. As illustrated in Fig. 2.2, strings can be either closed or open, but this study primarily focuses on closed strings.

Different vibrational states of the string correspond to different physical states with varying masses, spin quantum numbers, and gauge quantum numbers. Each excitation of the string adds mass on the order of M_s. For energy scales much smaller than M_s, only the massless particles' behavior is observable.

As will be discussed later, considering closed bosonic strings leads to the presence of a massless state corresponding to a traceless symmetric tensor G_{MN} with two Lorentz indices. This naturally incorporates gravity into the theory.

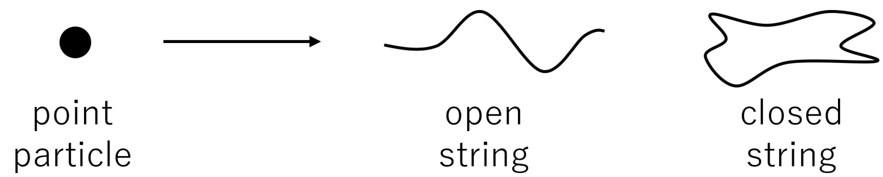

Figure 2.2: At low energy limits, the resolution of the string becomes coarse, making it appear to behave like a point particle. This figure illustrates the difference between open and closed strings, with a focus on the closed strings, which are the primary subject of this study.

The trajectory of a string's time evolution forms a two-dimensional surface called the worldsheet Σ, which is analogous to the worldline of a point particle. As depicted in Fig. 2.3, a point on this worldsheet is parametrized by coordinates (t, σ), where t represents the time coordinate, and σ represents the spatial direction along the string.

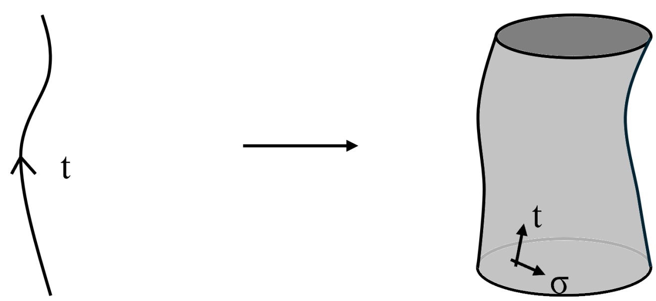

Figure 2.3: The worldsheet Σ formed by the trajectory of a string's time evolution, parametrized by coordinates (t, σ).

2.3.2 Closed Bosonic String

A closed bosonic string propagating in a flat D-dimensional Minkowski space is represented as an embedding $X^M(t, \sigma)$, with $M = 0, \ldots, D-1$.

The dynamics of the string are described by the action $S[X(t, \sigma)]$, which can be expressed as the total area of the worldsheet:

$$S_{NG} = -\frac{1}{2\pi\alpha'} \int_\Sigma dA$$

where $\frac{1}{2\pi\alpha'} \approx M_s^2$ is the string tension, representing the energy per unit length. This can be rewritten in terms of the embedding $X^M(t, \sigma)$ as:

$$S_{NG} = -\frac{1}{2\pi\alpha'} \int_\Sigma \sqrt{-\det(h)}\, d\sigma\, dt$$

Here, h is the induced metric on the 2-dimensional worldsheet, derived from the spacetime metric as:

$$h_{tt} = \partial_t X^M \partial_t X_M, \quad h_{\sigma\sigma} = \partial_\sigma X^M \partial_\sigma X_M, \quad h_{t\sigma} = \partial_t X^M \partial_\sigma X_M$$

This action is known as the Nambu-Goto action.

However, due to the square root in the action, quantization is challenging. Instead, we use the Polyakov action, which is classically equivalent but more amenable to quantization:

$$S_P[X, g] = -\frac{1}{4\pi\alpha'} \int_\Sigma d^2\xi \sqrt{-g}\, g^{ab} \partial_a X^M \partial_b X^N \eta_{MN},$$

where $g = \det(g_{ab})$. The worldsheet Σ of a closed string is parametrized by (t, σ), collectively denoted as ξ^a, where $a = 0, 1$. The metric $g_{ab}(t, \sigma)$ is an abstract 2-dimensional metric on the worldsheet, independent of the embedding $X(t, \sigma)$.

The physical vibrational modes of the string are constrained to propagate in directions perpendicular to the worldsheet, resulting in $D - 2$ physical degrees of freedom corresponding to the two transverse directions on the worldsheet.

From the equation of motion derived by varying the action with respect to the metric g_{ab}:

$$\frac{\delta S}{\delta g^{ab}} = 0 \Rightarrow -\frac{1}{2} g_{ab} g^{cd} \partial_c X^M \partial_d X_M + \partial_a X^M \partial_b X_M = 0$$

The Polyakov action possesses the following global and local symmetries:

- **Poincaré invariance**: The action is invariant under global Lorentz transformations and translations in the target space. The transformations are given by:

$$X'^M(\xi) = \Lambda^M_N X^N(\xi) + a^M,$$

$$g'_{ab}(\xi) = g_{ab}(\xi).$$

  where Λ^M_N is a Lorentz transformation matrix, and a^M represents a constant translation.

- **Local worldsheet coordinate reparameterization invariance**: The action remains invariant under the following reparameterizations of the worldsheet coordinates:

$$\xi'^a = \xi'^a(\xi),$$

$$X'^M(\xi') = X^M(\xi),$$

$$g'_{ab}(\xi') = \frac{\partial \xi^c}{\partial \xi'^a} \frac{\partial \xi^d}{\partial \xi'^b} g_{cd}(\xi).$$

- **Weyl invariance**: The action is also invariant under local rescalings of the worldsheet metric:

$$X'^M(\xi) = X^M(\xi),$$

$$g'_{ab}(\xi) = \Omega(\xi) g_{ab}(\xi).$$

By applying the gauge fixing known as the conformal gauge, the degrees of freedom in the metric g_{ab} can be removed, leading to:

$$g_{ab} = \eta_{ab},$$

where η_{ab} is the Minkowski metric. This gauge choice results in flat, orthogonal coordinates t and σ on the worldsheet. Let $\sigma = 0$ be a reference line along the t-axis, and denote l as the total length of the string in the σ direction.

With this gauge fixing, the Polyakov action simplifies to:

$$S_F = -\frac{1}{4\pi\alpha'} \int_\Sigma d^2\xi\, \eta^{ab} \partial_a X^M \partial_b X^N \eta_{MN},$$

The equations of motion derived from this action are the 2D wave equations:

$$(\partial_t^2 - \partial_\sigma^2) X^M = 0,$$

which admit general solutions in the form of left- and right-moving waves:

$$X^M(t, \sigma) = X_L^M(t + \sigma) + X_R^M(t - \sigma).$$

From the condition $\delta S/\delta g_{ab} = 0$, we obtain the following constraints:

$$(\partial_t X^M \partial_t X_M) + (\partial_\sigma X^M \partial_\sigma X_M) = 0,$$

$$\partial_t X^M \partial_\sigma X_M = 0.$$

Introducing the left- and right-moving worldsheet coordinates $\xi_L = t + \sigma$ and $\xi_R = t - \sigma$, these constraints can be rewritten as the Virasoro constraints:

$$(\partial_{\xi_L} X^M)(\partial_{\xi_L} X_M) = 0, \quad (\partial_{\xi_R} X^M)(\partial_{\xi_R} X_M) = 0.$$

After the gauge fixing discussed above, the metric $\mathrm{ds}^2 = -d\xi_L d\xi_R$ remains invariant under the coordinate transformations

$$\xi_L \to f_L(\xi_L), \quad \xi_R \to f_R(\xi_R),$$

along with an appropriate Weyl rescaling. This invariance is known as **conformal symmetry**.

Next, we proceed to quantization using light-cone coordinates.

The spacetime light-cone coordinates are defined as:

$$X^\pm = \frac{1}{\sqrt{2}}(X^0 \pm X^1),$$

with the remaining coordinates indexed by $i = 2, \ldots, D-1$. The metric of the D-dimensional Minkowski spacetime takes the form $\eta_{+-} = \eta_{-+} = -1$, $\eta_{ij} = \delta_{ij}$, and for a vector V^M, we have $V_- = -V^+$, $V_+ = -V^-$, and $V_i = V^i$.

From the solution to the equation of motion, X^+ is expressed as:

$$X^+(\xi) = X_L^+(\xi_L) + X_R^+(\xi_R).$$

Given that $t = (\xi_L + \xi_R)/2$ and using the conformal transformation, we obtain:

$$X^+(t, \sigma) = t.$$

In these light-cone coordinates, conformal symmetries are gauge-fixed, and Lorentz invariance is no longer manifestly preserved.

The Virasoro constraints simplify and become linear in X^-, expressed as:

$$\partial_{\xi_L} X_L^- = \frac{1}{2}(\partial_{\xi_L} X_L^i)^2, \quad \partial_{\xi_R} X_R^- = \frac{1}{2}(\partial_{\xi_R} X_R^i)^2.$$

The degrees of freedom in X^- are reduced to the center of mass term given by:

$$X^-(t) = -\frac{1}{l}\int_0^l d\sigma \, X^-(t, \sigma).$$

The Lagrangian then takes the form:

$$L = -\frac{l}{2\pi\alpha'}\partial_t x^-(t) - \frac{1}{4\pi\alpha'}\int_0^l d\sigma \left(-\partial_t X^i \partial_t X^i + \partial_\sigma X^i \partial_\sigma X^i\right).$$

The momentum conjugate to x^- is constant and is given by:

$$p_- = -p^+ = \frac{\partial L}{\partial(\partial_t x^-)} = \frac{l}{2\pi\alpha'}$$

where the string length l is related to p^+ as follows:

$$l = 2\pi\alpha' p^+.$$

The conjugate momentum associated with $X^i(t, \sigma)$ is:

$$\Pi_i(t, \sigma) = \frac{1}{2\pi\alpha'}\partial_t X^i(t, \sigma).$$

The Hamiltonian, representing the $D-2$ free bosons on a two-dimensional worldsheet, is given by:

$$H = p_- \partial_t x^-(t) + \int_0^l d\sigma \, \Pi_i(t, \sigma)\partial_t X^i(t, \sigma) - L = \frac{1}{2}\int_0^l d\sigma \left(2\pi\alpha' \Pi_i \Pi_i + \frac{1}{2\pi\alpha'}\partial_\sigma X^i \partial_\sigma X^i\right).$$

The left- and right-moving components X_L^i and X_R^i can be expanded in terms of oscillators. The periodicity condition for the closed string,

$$X^i(t, \sigma + l) = X^i(t, \sigma).$$

imposes the following form on the expansions of X_L^i and X_R^i:

$$X_L^i(t+\sigma) = \frac{x^i}{2} + \frac{p_i}{2p^+}(t+\sigma) + i\sqrt{\frac{\alpha'}{2}} \sum_{n\in\mathbb{Z}-\{0\}} \frac{\alpha_n^i}{n} e^{-2\pi i n(t+\sigma)/l}.$$

$$X_R^i(t-\sigma) = \frac{x^i}{2} + \frac{p_i}{2p^+}(t-\sigma) + i\sqrt{\frac{\alpha'}{2}} \sum_{n\in\mathbb{Z}-\{0\}} \frac{\tilde{\alpha}_n^i}{n} e^{-2\pi i n(t-\sigma)/l}.$$

$$X^i(t,\sigma) = x^i + \frac{p_i}{p^+}t + i\sqrt{\frac{\alpha'}{2}} \sum_{n\in\mathbb{Z}-\{0\}} \left[\frac{\alpha_n^i}{n} e^{-2\pi i n(t+\sigma)/l} + \frac{\tilde{\alpha}_n^i}{n} e^{-2\pi i n(t-\sigma)/l}\right].$$

where x^i and p^i are the center-of-mass position and momentum, respectively, and α_n^i, $\tilde{\alpha}_n^i$ represent the amplitudes of the left- and right-moving momentum modes of the string.

Treating the worldsheet degrees of freedom x^-, p^+, X^i, and Π^i as operators, we obtain the following canonical commutation relations:

$$[x^i, p_j] = i\delta_j^i, \quad [x^-, p^+] = -i,$$

$$[\alpha_m^i, \alpha_n^j] = [\tilde{\alpha}_m^i, \tilde{\alpha}_n^j] = m\delta^{ij}\delta_{m,-n}, \quad [\alpha_m^i, \tilde{\alpha}_n^j] = 0.$$

The Hamiltonian in its oscillator representation is given by:

$$H = \sum_{i=2}^{D-1} \left[\frac{p_i^2}{2p^+} + \sum_{n>0} \left(\alpha_{-n}^i \alpha_n^i + \tilde{\alpha}_{-n}^i \tilde{\alpha}_n^i\right) + E_0 + \tilde{E}_0\right],$$

where E_0 and $\tilde{E}_0$ are the zero-point energies of the harmonic oscillators.

The Hilbert space of string states is constructed by defining the ground state $|0\rangle$ such that:

$$\alpha_n^i|0\rangle = \tilde{\alpha}_n^i|0\rangle = 0 \quad \forall n > 0.$$

The excited states are then obtained by acting on the ground state with the creation operators α_{-n}^i and $\tilde{\alpha}_{-n}^i$ for $n > 0$.

Due to the symmetry along the $\sigma = 0$ reference line, the momentum operator in the σ direction, P_σ, must satisfy the condition $P_\sigma = 0$. This leads to the physical states having to satisfy the level-matching constraint:

$$N = \tilde{N},$$

where the oscillator number operators N and $\tilde{N}$ are defined as:

$$N = \sum_i \sum_{n>0} \alpha_{-n}^i \alpha_n^i, \quad \tilde{N} = \sum_i \sum_{n>0} \tilde{\alpha}_{-n}^i \tilde{\alpha}_n^i.$$

From the mass-shell condition, the squared mass M^2 of the string states is given by:

$$M^2 = -p^2 = 2p^+p^- - p^i p_i.$$

Substituting $p^- = i\partial_{x^+} = i\partial_t = H$, we obtain:

$$\frac{\alpha' M^2}{2} = N + \tilde{N} + E_0 + \tilde{E}_0.$$

The zero-point energies E_0 and $\tilde{E}_0$ are given by:

$$E_0 = \tilde{E}_0 = \sum_i \frac{1}{2} \sum_{n=1}^{\infty} n.$$

These are calculated using zeta function regularization and are given by:

$$E_0 = \tilde{E}_0 = -\frac{D-2}{24}.$$

The light spectrum of the string states is as follows:

$$N = \tilde{N} = 0: \quad |k\rangle, \quad \frac{\alpha' M^2}{2} = -\frac{D-2}{12},$$

$$N = \tilde{N} = 1: \quad a^i_{-1}\tilde{a}^j_{-1}|k\rangle, \quad \frac{\alpha' M^2}{2} = 2 - \frac{D-2}{12}.$$

In the case of 26 fields, the spectrum includes a tachyon with negative mass, the graviton G_{MN}, an antisymmetric background field B_{MN}, and the scalar dilaton field Φ. However, the tachyon does not appear in superstring theory.

To restore Lorentz symmetry in D dimensions, the critical dimension $D = 26$ is required. This critical dimension ensures the consistency of the theory and the preservation of Lorentz invariance across all dimensions.

2.3.3 One Loop Vaccum Amplitude and Modular Symmetry

Next, we explain how the ultraviolet divergence in the one-loop vacuum amplitude, also called partition function, is avoided due to modular invariance, which originates from the symmetry of the worldsheet.

The one-loop amplitude for a closed string corresponds to the geometry of a two-torus, which can be mapped onto a two-dimensional lattice. If we take the length of the string in the σ direction as l and the time evolution as $\tau_2 l$, where τ_2 represents the ratio of the time period to the string length, the geometry can be represented as shown in the left diagram of Fig. 2.4. Here, the coordinate z is expressed as $z = \sigma + it$, and the identifications $z \equiv z + \ell$ and $z \equiv z + i\tau_2\ell$ are made.

Due to the scaling symmetry, physically distinct cases arise only when the ratio τ_2 differs. Additionally, if there is a shift in the σ direction when completing a loop, this also represents a distinct geometry. Therefore, a general two-torus can be represented by $\tau = \tau_1 + i\tau_2$, with the identifications $z \equiv z + \ell$ and $z \equiv z + \tau\ell$, as shown in the right diagram of Fig. 2.4.

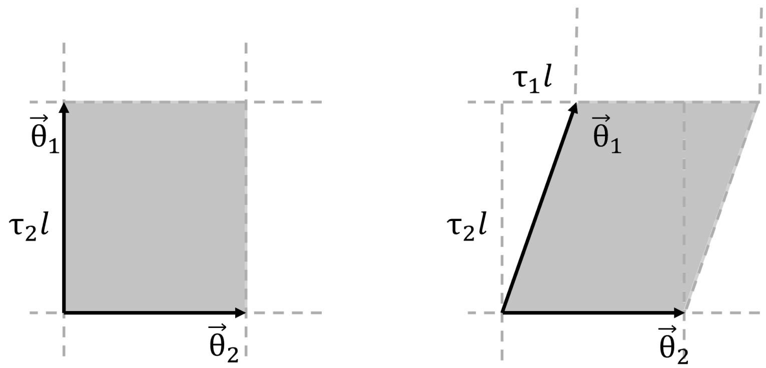

Figure 2.4: The geometry of the two-torus corresponding to the one-loop amplitude for a closed string, mapped onto a two-dimensional lattice. The vectors $\vec{\theta}_1$ and $\vec{\theta}_2$ represent the lattice vectors of the torus.

When the reference line in the σ direction overlaps after one complete loop, it represents the same geometry, giving rise to the symmetry of the geometries:

$$\tau \to \tau + 1.$$

Moreover, due to the symmetry of the worldsheet coordinates, there is a symmetry corresponding to interchanging the σ and t directions of the lattice, leading to the transformation:

$$\tau \to -\frac{1}{\tau}.$$

These transformations are part of the modular group of the two-torus and have the general $\mathrm{SL}(2,\mathbb{Z})$ form:

$$\tau \to \frac{a\tau + b}{c\tau + d}, \quad \text{with} \quad a, b, c, d \in \mathbb{Z}, \quad \text{and} \quad ad - bc = 1.$$

Under this modular symmetry, every point in the upper half of the complex plane is mapped to the shaded region in the Fig. 2.5.

Thus, when summing over all distinct geometries of the two-torus in the one-loop calculation, the geometry corresponding to $\tau_2 \to 0$, which would lead to ultraviolet divergence, is replaced by the geometry $\tau_2 \to \infty$. This prevents the occurrence of ultraviolet divergence in the one-loop correction to the vacuum expectation value.

2.3.4 Kaluza-Klein compactification

We now proceed to explain the Kaluza-Klein (KK) compactification, which forms the foundation for the compactifications considered in this study.

Consider a circle compactification of the 26th dimension. The coordinate X^{25} has the following periodic boundary condition:

$$X^{25}(t, \sigma + l) = X^{25}(t, \sigma) + 2\pi R w, \quad w \in \mathbb{Z},$$

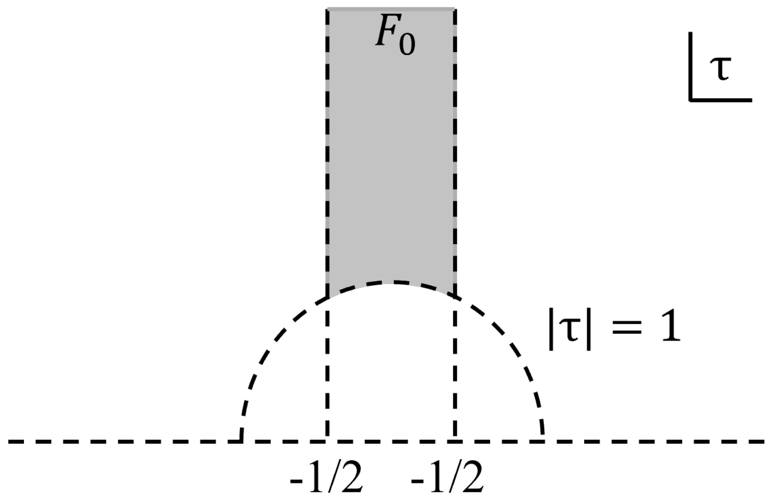

Figure 2.5: The fundamental domain under modular symmetry, where each point in the upper half-plane is mapped to the shaded region.

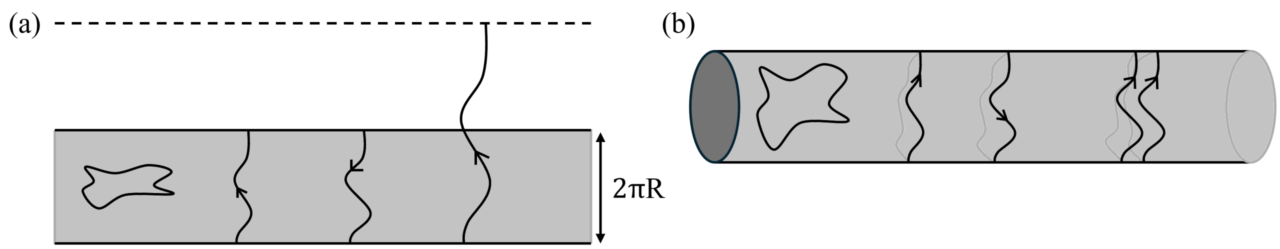

Figure 2.6: The winding number w corresponds to the number of times the string wraps around the compactified 26th dimension.

where w is known as the winding number, as shown in Fig. 2.6.

The oscillator expansion for X^{25} is given by:

$$X^{25}(t,\sigma) = x^{25} + \frac{k/R}{p^+}t + \frac{2\pi R w}{l}\sigma + i\sqrt{\frac{\alpha'}{2}}\sum_{n\neq 0}\left(\frac{\alpha_n^{25}}{n}e^{-2\pi i n(t+\sigma)/l} + \frac{\tilde{\alpha}_n^{25}}{n}e^{-2\pi i n(t-\sigma)/l}\right).$$

where the center-of-mass momentum is quantized as $p^{25} = \frac{k}{R}$ with $k \in \mathbb{Z}$.

The Hamiltonian is then given by:

$$H = H_{w=0} + \frac{1}{2}\int_0^l d\sigma \frac{1}{2\pi\alpha'}\left(\frac{2\pi R w}{l}\right)^2$$

$$= \sum_{i=2}^{24}\frac{p_i^2}{2p^+} + \frac{(k/R)^2}{2p^+} + \frac{R^2 w^2}{2\alpha'^2 p^+} + \frac{1}{\alpha' p^+}\left(N + \tilde{N} - 2\right).$$

where the level matching constraint is modified to:

$$P_\sigma = \int_0^l d\sigma\, \Pi_i \partial_\sigma X^i = \frac{2\pi}{l}(N - \tilde{N} + kw) = 0.$$

The mass formula for the compactified 25th dimension is:

$$M_{5d}^2 = \frac{k^2}{R^2} + \frac{w^2 R^2}{\alpha'^2} + \frac{2}{\alpha'}\Big(N + \tilde{N} - 2\Big).$$

Ignoring the tachyon, the massless modes are given by:

$$\begin{aligned}
\alpha_{-1}^{\mu}\tilde{\alpha}_{-1}^{\nu}|0\rangle &\to g_{\mu\nu}, b_{\mu\nu}, \phi, \\
\alpha_{-1}^{\mu}\tilde{\alpha}_{-1}^{25} + \alpha_{-1}^{25}\tilde{\alpha}_{-1}^{\nu}|0\rangle &\to A_\mu, \\
\alpha_{-1}^{\mu}\tilde{\alpha}_{-1}^{25} - \alpha_{-1}^{25}\tilde{\alpha}_{-1}^{\nu}|0\rangle &\to \tilde{A}_\mu, \\
\alpha_{-1}^{25}\tilde{\alpha}_{-1}^{25}|0\rangle &\to \sigma.
\end{aligned}$$

Here, μ, ν are indices in the 25-dimensional spacetime.

The oscillator expansions for X_L^{25} and X_R^{25} are given by:

$$X_L^{25}(t+\sigma) = \frac{x^{25}}{2} + \sqrt{\frac{2}{\alpha'}}\frac{p_L}{2p^+}(t+\sigma) + i\sqrt{\frac{\alpha'}{2}}\sum_{n\neq 0}\frac{\alpha_n^{25}}{n}e^{-2\pi i n(t+\sigma)/l},$$

$$X_R^{25}(t-\sigma) = \frac{x^{25}}{2} + \sqrt{\frac{2}{\alpha'}}\frac{p_R}{2p^+}(t-\sigma) + i\sqrt{\frac{\alpha'}{2}}\sum_{n\neq 0}\frac{\tilde{\alpha}_n^{25}}{n}e^{-2\pi i n(t-\sigma)/l}.$$

The left- and right-moving momenta are:

$$p_L = \sqrt{\frac{\alpha'}{2}}\left(\frac{k}{R} + \frac{wR}{\alpha'}\right), \quad p_R = \sqrt{\frac{\alpha'}{2}}\left(\frac{k}{R} - \frac{wR}{\alpha'}\right).$$

The expressions for H_L and H_R are:

$$H_L = \frac{1}{4p^+}\sum_{i=1}^{24} p_i^2 + \frac{1}{\alpha' p^+}\left(\frac{p_L^2}{2} + N + E_0\right),$$

$$H_R = \frac{1}{4p^+}\sum_{i=1}^{24} p_i^2 + \frac{1}{\alpha' p^+}\left(\frac{p_R^2}{2} + \tilde{N} + \tilde{E}_0\right).$$

The total mass squared is given by:

$$M^2 = M_L^2 + M_R^2,$$

where

$$\begin{aligned}
\frac{\alpha' M_L^2}{2} &= \frac{p_L^2}{2} + N - 1, \\
\frac{\alpha' M_R^2}{2} &= \frac{p_R^2}{2} + \tilde{N} - 1.
\end{aligned}$$

The level-matching constraint requires $M_L^2 = M_R^2$.

2.3.5 Heterotic String Theories

By introducing supersymmetry, we incorporate fermionic strings. Fermionic strings have a critical dimension of 10, which is required by the demand for Lorentz symmetry during the light-cone quantization process.

Next, we discuss heterotic strings. Heterotic strings are a theory of closed strings, where the left and right sectors are constructed from different string theories. However, the level-matching condition must still be satisfied.

Using the light-cone gauge, we define the right sector as X^i_R and Ψ^i_R for $i = 2, \ldots, 9$, and the left sector as X^I_L and X^I_L for $I = 1, \ldots, 16$.

Here, the index i refers to the 8 dimensions of the 10-dimensional spacetime with 10d Poincaré invariance, while the remaining 16 dimensions indexed by I correspond to an internal space compactified on a 16-dimensional torus of stringy size $R = \sqrt{\alpha'}$.

The harmonic oscillator expansion for X^i_R is given by:

$$X^i_R(t-\sigma) = x^i + \frac{p^i}{2p^+}(t-\sigma) + i\sqrt{\frac{\alpha'}{2}} \sum_{n\neq 0} \frac{\tilde{\alpha}^i_n}{n} e^{-2\pi i n(t-\sigma)/l}.$$

For fermionic fields, which appear quadratically in the action and other physical observables, there are two possible physical periodicities:

- Neveu-Schwarz (NS): $\Psi^i_R(t-\sigma+l) = -\Psi^i_R(t-\sigma)$
- Ramond (R): $\Psi^i_R(t-\sigma+l) = \Psi^i_R(t-\sigma)$

Thus, the harmonic oscillator expansion for Ψ^i_R is:

$$\Psi^i_R(t-\sigma) = i\sqrt{\frac{\alpha'}{2}} \sum_{r\in\mathbb{Z}} \tilde{\psi}^i_{r+\nu} e^{-2\pi i(r+\nu)(t-\sigma)/l}.$$

where $\nu = \frac{1}{2}$ for NS and $\nu = 0$ for R sectors.

These oscillators satisfy the following anticommutation relations:

$$\{\psi^i_{m+\nu}, \psi^j_{n+\nu}\} = \delta^{ij}\delta_{m+\nu,-(n+\nu)}.$$

The Hamiltonian for the right sector, including the number operators, is given by:

$$H_R = \frac{\sum_i p_i^2}{4p^+} + \frac{1}{\alpha' p^+}\left(\tilde{N}_B + \tilde{N}_F + \tilde{E}_0\right),$$

where the number operators are defined as:

$$\tilde{N}_B = \sum_i \sum_{n>0} \tilde{\alpha}^i_{-n}\tilde{\alpha}^i_n, \quad \tilde{N}_F = \sum_{r\geq 0}(r+\nu)\tilde{\psi}^i_{-r-\nu}\tilde{\psi}^i_{r+\nu}, \quad \tilde{E}_0 = -2\nu(1-\nu).$$

For the left sector, X^i_L follows the same structure as in the bosonic closed string theory.

The absence of a right-moving pair for X_L^I is due to the specific choice of coordinates in the 16-dimensional space, where $X_R^I \equiv 0$. Here, we have set

$$x^I \equiv 0, \quad \tilde{\alpha}_n^I \equiv 0, \quad k^I \equiv w^I, \quad R \equiv \sqrt{\alpha'}.$$

As a result, the harmonic oscillator expansion for X_L^I is given by:

$$X_L^I(t+\sigma) = \frac{P^I}{\sqrt{2\alpha'}\, p^+}(t+\sigma) + i\sqrt{\frac{\alpha'}{2}} \sum_{n\neq 0} \frac{\alpha_n^I}{n} e^{-2\pi i n(t+\sigma)/l}.$$

The Hamiltonian for the left sector is then given by:

$$H_L = \frac{\sum_i p_i^2}{4p^+} + \frac{\sum_I (P^I)^2}{2\alpha' p^+} + \frac{1}{\alpha' p^+}\left(N_B - 1\right),$$

where:

$$N_B = \sum_i \sum_{n>0} \alpha_{-n}^i \alpha_n^i + \sum_I \sum_{n>0} \alpha_{-n}^I \alpha_n^I.$$

The level-matching condition requires that $M_L^2 = M_R^2$, and the mass formulas are given by:

$$\frac{\alpha' M_R^2}{2} = \tilde{N}_B + \tilde{N}_F - 2\nu(1-\nu),$$

$$\frac{\alpha' M_L^2}{2} = N_B + \frac{P^2}{2} - 1,$$

where $P^2 = \sum_I P_I^2$.

The vacuum amplitude, represented by the partition function, must satisfy the symmetries related to the toroidal structure of the worldsheet, particularly modular invariance. This symmetry is crucial for maintaining quantum consistency in string theory and ensures that the partition function yields physically correct results.

The partition function is given by:

$$Z(\tau) = \left(4\pi\alpha'\tau_2\right)^{-4} |\eta(\tau)|^{-16} \bar{Z}_\psi(\bar{\tau}) Z_{X^I}(\tau).$$

In this expression, the factors corresponding to the center-of-mass momentum and the trace over the bosonic closed string's harmonic oscillators are analogous to those in the bosonic closed string theory. The new factor $\bar{Z}_\psi$ corresponds to the trace over the Ψ_R^i oscillators, which requires the presence of both NS and R sectors and imposes the GSO projection.

The GSO projection is an operation in string theory that eliminates unphysical states, yielding a physically meaningful spectrum. Specifically, it removes the tachyon in the NS sector, leaving only the state $\Psi^i_{-1/2}|0\rangle_{\text{NS}} = |8_V\rangle$ from the NS sector and the state $|8_C\rangle$ from the R sector. The state $|8_C\rangle$ is a spinor representation of SO(8), indicating an anti-chiral spinor.

The final factor is the trace over the X_L^I oscillators and left-moving momentum, with the 16-dimensional lattice Λ_{16}, given by:

$$Z_{X^I}(\tau) = \eta(\tau)^{-16} \sum_{P \in \Lambda_{16}} q^{P^2/2},$$

where

$$q = e^{2\pi i \tau},$$

and $\eta(\tau)$ is the Dedekind eta function, defined as:

$$\eta(\tau) = q^{1/24} \prod_{n=1}^{\infty} (1 - q^n).$$

The condition $\tau \to \tau + 1$ requires an even lattice, $P^2 \in 2\mathbb{Z}$, while $\tau \to -1/\tau$ requires a self-dual lattice, $\Lambda_{16}^* = \Lambda_{16}$.

These conditions are satisfied only by the $Spin(32)/\mathbb{Z}_2$ lattice and the $E_8 \times E_8$ lattice. In this dissertation, we focus on the latter, which is defined as a direct sum of two 8-dimensional lattices:

$$(n_1, ..., n_8) or (n_1 + \frac{1}{2}, ..., n_8 + \frac{1}{2}) \qquad with \quad n_I \in \mathbb{Z}, and \sum_I n_I = even$$

The E_8 symmetry includes the gauge groups used in Grand Unified Theories, $SU(5)$ or $SO(10)$, and the gauge symmetry of the MSSM as subgroups.

2.3.6 Heterotic Orbifold Compactification

In this study, we utilize toroidal orbifold compactifications. To explain this concept, let us first consider the simple example of a 2-dimensional torus.

The 2-torus T^2 is defined as $T^2 = \mathbb{R}^2/\Lambda_2$, where $\mathbf{e}_1$ and $\mathbf{e}_2$ are the lattice vectors of Λ_2.

In other words, a point $\mathbf{x} \in \mathbb{R}^2$ is identified on the 2-torus as follows:

$$\mathbf{x} \equiv \mathbf{x} + n_1 \mathbf{e}_1 + n_2 \mathbf{e}_2, \quad n_1, n_2 \in \mathbb{Z}.$$

Now, consider a $\mathbb{Z}_2$ orbifold of this torus. The generator of the $\mathbb{Z}_2$ discrete rotation group is denoted by θ, and the corresponding point group is denoted as P. This operation identifies points on T^2 under the action of θ.

There are four fixed points on the torus that remain invariant under this orbifold action, given by:

$$\mathbf{x}_f = \theta \mathbf{x}_f + \sum_a m_a^f \mathbf{e}_a^f \quad \text{for some } m_a^f \in \mathbb{Z}, \quad a = 1, 2.$$

These fixed points correspond to the coordinates $(0,0)$, $(1/2, 0)$, $(0, 1/2)$, and $(1/2, 1/2)$. The resulting $\mathbb{Z}_2$ orbifold is illustrated in Fig. 2.7.

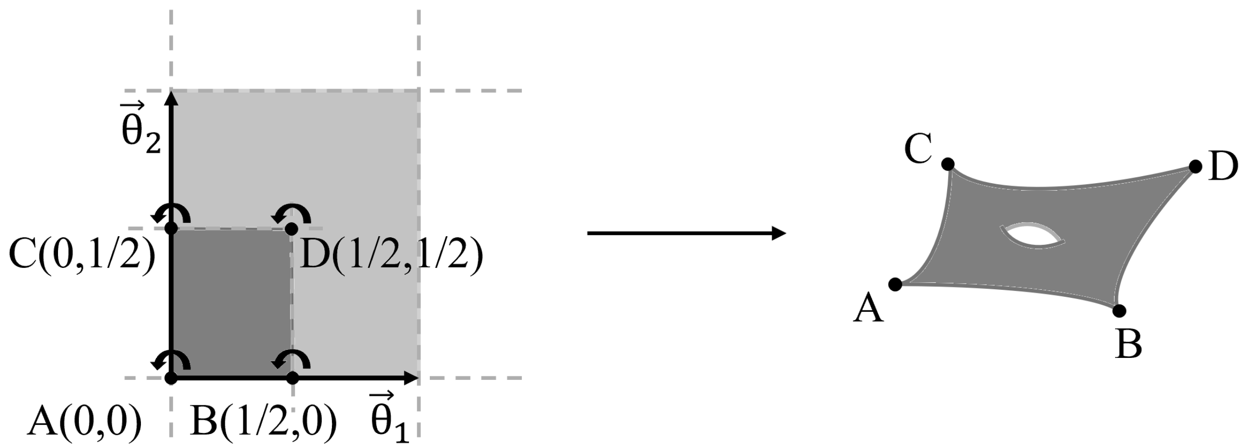

Figure 2.7: The fixed points of the $\mathbb{Z}_2$ orbifold are shown at the coordinates $(0,0)$, $(1/2,0)$, $(0,1/2)$, and $(1/2,1/2)$ on the 2-dimensional torus.

Thus, a toroidal orbifold can be viewed as $\mathbb{R}^2/S$, where S is the space group generated by the lattice translations $\mathbf{e}_a$ and the action of θ. It is important that the introduction of P must not distort the lattice structure, generally requiring that

$$\theta \mathbf{e}_a \in \Lambda_2.$$

The same approach applies to the 6-dimensional case.

Consider $T^6 = T^2 \times T^2 \times T^2$, where each 2-torus is mapped to complex coordinates $z^i = \frac{1}{\sqrt{2}}\left(x^{2i-1} + ix^{2i}\right)$, $i = 1,2,3$. The generator θ of $\mathbb{Z}_N$ is expressed in terms of the Cartan generators $J_{2i-1,2i}$ of $SO(6)$ as:

$$\theta = \exp\left[2\pi i \left(v_1 J_{12} + v_2 J_{34} + v_3 J_{56}\right)\right],$$

Here, the $J_{2i-1,2i}$ induces a rotation in the ith complex plane. Specifically, it acts as:

$$z_i \to e^{2\pi i v_i} z_i,$$

with the condition that after N applications of θ, the space returns to itself, implying

$$N v_i \in \mathbb{Z}.$$

Additionally, for the action to affect spinors, we require:

$$\sum_i N v_i = \text{even}.$$

By convention, we take $-1 < v_i < 1$.

To preserve 4d $N = 1$ supersymmetry after compactification, the following condition must hold:

$$v_1 + v_2 + v_3 = 0.$$

The possible combinations of $\mathbb{Z}_N$ and v satisfying these conditions are limited. For example, for $\mathbb{Z}_3$, we have $\frac{1}{3}(1,1,-2)$. For $\mathbb{Z}_6$, there are two possible cases: $\frac{1}{6}(1,1,-2)$ and $\frac{1}{6}(1,2,-3)$. The latter is the $\mathbb{Z}_6 - \mathrm{II}$ orbifold considered in this work. The $\mathbb{Z}_2$ case does not satisfy the conditions for $N=1$ supersymmetry, but for $N=2$, the configuration $\frac{1}{2}(1,-1,0)$ is possible.

We now consider heterotic compactification.

In this context, we denote indices as follows: the real compact space coordinates are labeled by $m, n, \cdots = 4, \ldots, 9$, the corresponding complex coordinates are labeled by $i, j, \cdots = 1, \ldots, 3$, and the dimensions corresponding to the oscillatory directions orthogonal to the light-cone coordinates in 4-dimensional spacetime are labeled by $\mu = 2, 3$.

Following the notation for representing the $\mathbb{Z}_N$ rotation in the complex mapping, we express the fields of the compact space as complex fields:

$$X^m \to Z^i, \quad \psi_R^m \to \psi_R^i.$$

These fields transform under the action of the $\mathbb{Z}_N$ rotation in the compact complex space.

Furthermore, the internal space field X^I is also subject to the $\mathbb{Z}_N$ generator θ, which acts as a shift:

$$X^I \to X^I + V^I.$$

The group generated by such transformations is denoted as G, and we consider the 16-dimensional torus T^{16,Het_L}, such that the compactification space can be expressed as:

$$\frac{T^6_{R+L}}{P} \otimes \frac{T_L^{16,\mathrm{Het}}}{G}.$$

To maintain the integrity of the lattice structure, we require:

$$NV \in \Lambda_{16},$$

where in this study, we take Λ_{16} to be the $E_8 \times E_8$ lattice.

In orbifold compactifications, as illustrated in Fig. 2.8, there are untwisted sectors, where closed strings remain invariant under the orbifold action, and twisted sectors, where closed strings acquire a twist under the orbifold action.

In both sectors, to simplify the calculations, the fermionic fields ψ^m are bosonized.

The boundary conditions in the untwisted sector and the corresponding oscillator expansions are as follows:

$$X^m(t, \sigma + l) = X^m(t, \sigma) + 2\pi R_m w^m, \quad \text{(no sum over } m\text{)}.$$

$$X^m(t, \sigma) = x_0^m + \frac{k_m}{R_m p^+} t + \frac{2\pi R_m}{l} w^m \sigma + \text{(oscillator terms)}.$$

where no summation is taken over the index m. Here, p^m represents the quantized momentum vector, and w^m represents the quantized winding vector.

The field X^I satisfies the same boundary conditions.

Next, we consider the twisted sector.

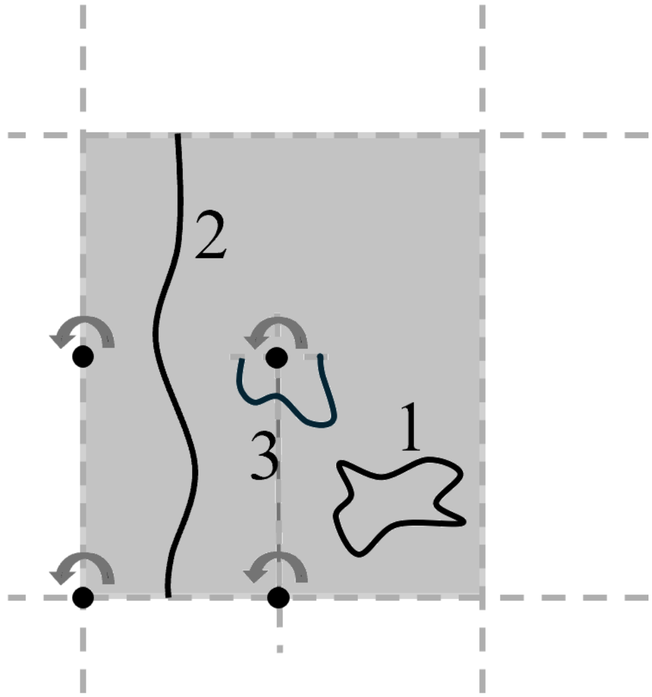

Figure 2.8: Illustration of untwisted and twisted sectors in orbifold compactifications. In the untwisted sectors, closed strings remain invariant under the orbifold action, while in the twisted sectors, closed strings acquire a twist.

The boundary conditions for X^m in the twisted sector are given by:

$$X^m(t, \sigma + l) = \theta X^m(t, \sigma) + 2\pi R_m w^m.(nosum)$$

Considering the complexified fields, the boundary conditions in the twisted sector for each field are as follows:

$$Z^i(t, \sigma + l) = e^{2\pi i v_i} Z^i(t, \sigma) + 2\pi R w^i,$$

$$\psi^i_R(t - \sigma + l) = \pm e^{2\pi i v_i} \psi^i_R(t - \sigma),$$

$$X^I(t + \sigma + l) = X^I(t + \sigma) + V^I,$$

where w^i is the complexified winding vector, and the sign in the fermion boundary condition is negative for the Neveu-Schwarz (NS) sector and positive for the Ramond (R) sector.

Since the physics is localized at the fixed points, there is no free motion, and the oscillator expansion does not include center-of-mass coordinates, momentum, or winding terms. Instead, the fixed-point coordinates Z^i_f appear, and the expansion is given by:

$$Z^i(t, \sigma) = Z^i_f + i\sqrt{\frac{\alpha'}{2}} \left[\sum_{n\neq 0} \frac{\alpha^i_{n-v_i}}{n - v_i} e^{-2\pi i(n-v_i)(t+\sigma)/l} + \sum_{n\neq 0} \frac{\tilde{\alpha}^i_{n+v_i}}{n + v_i} e^{-2\pi i(n+v_i)(t-\sigma)/l} \right],$$

where Z_f^i corresponds uniquely to the elements of S and is distinguished by them. Additionally, the oscillator modes are shifted due to the rotation. Similarly, the fermionic oscillators are modified, leading to terms like $\psi^i_{n-v_i}$ and $\tilde{\psi}^i_{n+v_i}$.

The oscillator expansion for the internal space X^I is different, as θ acts as a shift rather than a rotation:

$$X^I(t+\sigma) = \frac{P^I + V^I}{\sqrt{2\alpha' p^+}}(t+\sigma) + i\sqrt{\frac{\alpha'}{2}} \sum_{n\neq 0} \frac{\alpha_n^I}{n} e^{-2\pi i n(t+\sigma)/l}.$$

The mass in the untwisted sector is given by:

$$\frac{\alpha' M_R^2}{2} = \frac{p_R^2}{2} + \tilde{N}_B + \tilde{N}_{F_B} + \frac{r^2}{2} - \frac{1}{2},$$
$$\frac{\alpha' M_L^2}{2} = \frac{p_L^2}{2} + N_B + \frac{P^2}{2} - 1.$$

where $\tilde{N}_{F_B}$ is the number operator for the bosonized right-moving fermions. The vector $r = (r_1, r_2, r_3, r_0)$ represents the weights of the $SO(8)$ group, with the convention that r_0 corresponds to the two-dimensional part of $SO(8)$ not associated with the light-cone coordinates in 4-dimensional spacetime. For the Neveu-Schwarz (NS) sector, $r = (\pm 1, 0, 0, 0)$ represents the vector 8_V, and for the Ramond (R) sector, $r = \pm(1/2, 1/2, 1/2, -1/2)$ represents the spinor 8_C.

The level-matching condition requires that $M_R^2 = M_L^2$, and the massless states occur when $p_L = 0$, $N_B = 1$, $P = 0$, or $p_L = 0$, $N_B = 0$, $P^2 = 2$.

The following 4-dimensional fields are obtained:

- The 4-dimensional graviton $g_{\mu\nu}$, the two-index antisymmetric tensor $B_{\mu\nu}$, the dilaton ϕ, and their fermionic partners correspond to the states:

$$|(0,0,0,\pm 1)\rangle_R \otimes \alpha^\mu_{-1}|0\rangle_L, \quad |\pm\left(\frac{1}{2},\frac{1}{2},\frac{1}{2},-\frac{1}{2}\right)\rangle_R \otimes \alpha^\mu_{-1}|0\rangle_L.$$

- The 4d $N=1$ vector multiplet consisting of gauge bosons and gauginos corresponds to the states:

$$|(0,0,0,\pm 1)\rangle_R \otimes \alpha^I_{-1}|0\rangle_L, \quad |\pm\left(\frac{1}{2},\frac{1}{2},\frac{1}{2},-\frac{1}{2}\right)\rangle_R \otimes \alpha^I_{-1}|0\rangle_L.$$
$$|(0,0,0,\pm 1)\rangle_R \otimes |P^I\rangle_{P^2=2}, \quad |\pm\left(\frac{1}{2},\frac{1}{2},\frac{1}{2},-\frac{1}{2}\right)\rangle_R \otimes |P^I\rangle_{P^2=2}.$$

- The 4-dimensional complex scalar fields, which can be interpreted as zero modes of g_{mn} and B_{mn}, correspond to the states:

$$|(\underline{\pm 1,0,0},0)\rangle_R \otimes \alpha^m_{-1}|0\rangle_L, \quad |\pm\left(\underline{-\frac{1}{2},\frac{1}{2},\frac{1}{2}},\frac{1}{2}\right)\rangle_R \otimes \alpha^m_{-1}|0\rangle_L.$$

- The charged matter chiral multiplets C^J, $J=1,2,3$ correspond to the states:

$$|(\underline{\pm 1,0,0},0)\rangle_R \otimes |P^I\rangle_{P^2=2}, \quad |\pm\left(\underline{-\frac{1}{2},\frac{1}{2},\frac{1}{2}},\frac{1}{2}\right)\rangle_R \otimes |P^I\rangle_{P^2=2}.$$

where the J corresponds to the permutation of the elements in the underlined part of the states.

We now consider the spectrum in the twisted sector. Under bosonization, the mass formulae are given by:

$$\frac{\alpha' M_R^2}{2} = \tilde{N}_B + \tilde{N}_{F_B} + \frac{(r+nv)^2}{2} + E_0 - \frac{1}{2},$$
$$\frac{\alpha' M_L^2}{2} = N_B + \frac{(P+nV)^2}{2} + E_0 - 1.$$

Here, E_0 is shifted due to the fractional modes of the twisted oscillators, and is given by:

$$E_0 = \sum_i \frac{1}{2}|nv_i|(1-|nv_i|),$$

where $-1 < nv_i < 1$.

From the level-matching condition, which arises from modular symmetry, the following must be satisfied:

$$N(V^2 - v^2) = 0 \mod 2.$$

We introduce the internal components of the $E_8 \times E_8$ gauge fields A_m^I as background fields. These correspond to Wilson lines. Wilson lines can be considered along each non-trivial 1-cycle of the toroidal compactification dimensions. Let $\mathbf{e}_r$, $r = 1, \dots, 6$, be the lattice vectors of T^6. The Wilson lines are then defined as:

$$a_r^I = \int_{\mathbf{e}_r} dx^m A_m^I,$$

where a Wilson line along $\mathbf{e}$ (defined as $\mathbf{e} = \sum_r n_r \mathbf{e}_r$) acts on a state with momentum P^I as $\exp\left(2\pi i \sum_r n_r a_r^I P^I\right)$. This is embedded as a transformation of the gauge degrees of freedom.

Since the lattice vectors are commutative, the Wilson lines are also commutative.

For simplicity, we assume that the embedding of the Wilson lines and the orbifold twist is also commutative.

The embedding of S is then given by:

$$\text{Twist: } (\theta, 0; V^I), \quad \text{Wilson line: } (1, \mathbf{e}_r; a_r^I).$$

For the Wilson line corresponding to $\mathbf{e}_r$, since $\sum_n \theta^n \mathbf{e}_r = 0$ implies that $(\theta, \mathbf{e}_r)^N = (1, 0)$, the gauge embedding must satisfy:

$$(\theta, \mathbf{e}_r; V + a_r)^N = (1, 0; 0) \to N(V + a_r) \in \Lambda_{16d}, \quad \to N a_r \in \Lambda_{16d}.$$

If $\theta \mathbf{e}_r = \sum_s n_{rs} \mathbf{e}_s$, then $a_r = \sum_s n_{rs} a_s \mod \Lambda_{16d}$. For example, if $\mathbf{e}_2 = \theta \mathbf{e}_1$, then $a_2 = a_1$. More generally, not all Wilson lines are independent.

In the untwisted sector, a state with a gauge charge vector P experiences a shift in its 6-dimensional momentum, which contributes to the 4-dimensional mass as follows:

$$p_r \to p_r + P \cdot a_r.$$

The effect of the Wilson line on massless states is restricted by the condition:

$$P^I a_r^I \in \mathbb{Z}.$$

This restriction reduces the massless untwisted matter fields and the gauge group.

In the twisted sector, the boundary conditions at a fixed point f corresponding to an element $S = (\theta^n, \sum_r k_r^f \mathbf{e}_r)$ are given by:

$$X^m(t, \sigma + l) = (\theta^n X(t, \sigma))^m + \sum_r k_r^f e_r^m,$$

where e_r^m are the components of the lattice vectors $\mathbf{e}_r$.

The corresponding boundary condition for the gauge space coordinates is:

$$X^I(t + \sigma + l) = X^I(t + \sigma) + nV^I + \sum_r k_r^f a_r^I.$$

In this twisted sector at fixed point f, the gauge charge vector P is shifted as:

$$P \to P + nV + \sum_r k_r^f a_r.$$

Thus, in the presence of Wilson lines, the general replacement in the mass formulae is:

$$nV \to nV + \sum_r k_r^f a_r.$$

The left-moving mass formula is modified as follows:

$$\frac{\alpha' M_L^2}{2} = N_B + \frac{(P + nV + \sum_r k_r^f a_r)^2}{2} + E_0 - 1.$$

The twisted sectors at different fixed points receive different shifts, leading to different massless spectra. This diversity increases the number of viable models. The Wilson lines can reduce the gauge group to the Standard Model gauge group and also influence the number of generations.

Moreover, the Wilson lines are subject to constraints imposed by modular symmetry:

$$N\left[(nV + \sum_r k_r^f a_r)^2 - n^2 v^2\right] = 0 \mod 2.$$

The shifts in the twisted sector vary depending on the combination of fixed points on each torus, resulting in different gauge groups and different states. It is important to note that some fixed points are identified by shift vectors that are also identified under the orbifold action.

2.4 Orbifold Compactification of $E_8 \times E_8$ Heterotic String Theory in This Study

In this chapter, we provide a concise summary of the Z_6-II orbifold compactification data associated with the $E_8 \times E_8$ heterotic string theory, which forms the foundation of the generative AI-driven exploration of MSSM-like models in this research. This section is based on [8].

2.4.1 Orbifold Compactification in $E_8 \times E_8$ Heterotic String Theory

In this subsection, we introduce the compactification process of the $E_8 \times E_8$ heterotic string theory this study focuses on. A common approach in compactification, for the sake of theoretical consistency, involves endowing the extra dimensions with translational symmetry in those dimensions, creating a lattice-like periodicity that corresponds to a torus, and then folding these dimensions down to a stringy scale that is imperceptible to observation. On this torus, there exists a symmetry known as modular symmetry, with modular transformations associated with scale transformations, rendering the physics equivalent even when scaled down. The degrees of freedom of these reduced dimensions are treated as internal space degrees of freedom. Heterotic string theory combines elements from both bosonic string theory, which describes bosons, and superstring theory, which describes both bosons and fermions. Bosonic string theory, dictated by physical constraints, is a theory in 26-dimensional spacetime, whereas superstring theory, with the imposition of supersymmetry, operates in a 10-dimensional framework. In heterotic string theory, only closed strings are considered. These closed strings possess two distinct modes of vibration: right-moving and left-moving. The left-moving degrees of freedom are associated with the strings of bosonic string theory, while the right-moving degrees of freedom correspond to the strings of superstring theory that describe both bosons and fermions. Consequently, within the theory, there exist left-moving degrees of freedom in 26 dimensions and right-moving degrees of freedom in 10 dimensions. To maintain consistency with Lorentz symmetry, the extra 16 dimensions of left-moving degrees of freedom must be compactified. This compactification, to achieve a theory that integrates gravity without anomalies and maintains mathcalN = 1 supersymmetry, is only feasible through compactification on a 16-dimensional lattice corresponding to either $E_8 \times E_8$ or $SO(32)$. The choice of $E_8 \times E_8$ for this purpose provides the specified 16-dimensional degrees of freedom, and the lattice representing these 16 dimensions is described by the following expression:

$$(n_1, ..., n_8) or (n_1 + \frac{1}{2}, ..., n_8 + \frac{1}{2}) \qquad with \quad n_I \in \mathbb{Z}, and \sum_I n_I = even$$

These 16-dimensional degrees of freedom are considered internal ones and are treated as gauge degrees of freedom because they exhibit E_8 symmetry, which includes the gauge groups used in Grand Unified Theories, $SU(5)$ or $SO(10)$, and the gauge symmetry of the MSSM as subgroups.

The degrees of freedom in heterotic string theory, when compactified to 10 dimensions, must be further compactified down to four-dimensional spacetime degrees of freedom to align with the currently observed Standard Model (SM). This extra six-dimensional toroidal space is often treated as a product of three two-dimensional toroidal spaces, $T^2 \times T^2 \times T^2$. By introducing rotational symmetry, singularities and singular planes emerge in the space, and the compactification of the extra six-dimensional space in ten dimensions into a four-dimensional space with gauge symmetries in the six-dimensional internal space results from specifying and introducing rotational axes to these three tori. This process combines the original space's symmetries with translational symmetries, thereby generating an internal space that possesses both dimensional reduction and enriched gauge symmetries. Such spaces in which points are identified through introduced symmetries are referred to as orbifolds.

Specifically, in heterotic string compactifications, orbifolds are defined as the quotient space of the six-dimensional real space $\mathbb{R}^6$ divided by a so-called space group S, where the quotient is taken using the equivalence relation $X \sim gX$ for all $g \in S$ and $X \in \mathbb{R}^6$. The space group S is composed of two types of groups:

- The so-called point group P, which exhibits discrete rotations. For simplicity, P is taken to be Abelian. To achieve $\mathcal{N} = 1$ supersymmetry in four dimensions, P can be either Z_M or $Z_M \times Z_N$, generated by rotation matrices θ and ω, with $\omega = 1$ used for Z_M. These matrices are represented by so-called

- twisted vectors v_1 and v_2, which give the three rotational phases in the three complex planes, ensuring the sum over all entries is an integer to maintain $\mathcal{N} = 1$.

- The vectors $e_\alpha \in \mathbb{R}^6$ ($\alpha = 1, \ldots, 6$) generating translations, forming a 6D lattice Γ that corresponds to a six-torus. Elements of P must map the lattice Γ onto itself.

An element of the space group S has the form $g = \left(\theta^k \omega^l, n_\alpha e_\alpha\right) \in S$, where $k, l \in \mathbb{Z}$, $n_\alpha \in \mathbb{Z}$ (or in some cases $n_\alpha \in \mathbb{Q}$), with summation over $\alpha = 1, \ldots, 6$. This acts on $X \in \mathbb{R}^6$ as $gX = \theta^k \omega^l X + n_\alpha e_\alpha$. Using these definitions, all 6D Abelian and toroidal orbifolds can be addressed.

Due to modular symmetry, the geometry of this compactification must be embedded in a 16-dimensional gauge degrees of freedom spaces X^I, $I = 1, ..., 16$ to maintain consistency with the prior stage of 16-dimensional compactification. Here, we consider only the method of embedding known as shift embedding, where $\theta \mapsto V_1$, $\omega \mapsto V_2$ and $e_\alpha \mapsto W_\alpha$ for $\alpha = 1, \ldots, 6$:

$$gX = \theta^k \omega^l X + n_\alpha e_\alpha \to gX^I = X^I + kV_1^I + lV_2^I + n_\alpha W_\alpha^I$$

where, the vectors V_1, V_2 and W_α are called shifts and Wilson lines, respectively. They must satisfy the following conditions for the $E_8 \times E_8$ lattice:

$$NV_1 \in \Lambda, MV_2 \in \Lambda, N_\alpha W_\alpha \in \Lambda$$

where Λ is the $E_8 \times E_8$ weight lattice. Furthermore, to comply with modular symmetry, they must also fulfill the following conditions:

$$N(V_1^2 - v_1^2) = 0 \qquad mod \quad 2,$$

```
begin model
Label:Model_SM77
SpaceGroup:Geometry/Geometry_Z6-II_G2xSU3xSU2^2.txt
Lattice:E8xE8
Shifts and Wilsonlines:
    0,     0,     0,     0,     0,     0,     0,     0, -5/12, -5/12, -1/12, -1/12,  1/12,  1/12,  1/12,  1/12
    0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
    0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
    0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
 -2/3,     0,     0,     0,     0,     0,   1/3,   1/3,    -1,   2/3,     0,   1/3,    -1,  -1/3,     0,     0
 -2/3,     0,     0,     0,     0,     0,   1/3,   1/3,    -1,   2/3,     0,   1/3,    -1,  -1/3,     0,     0
   -2,  -3/2,   1/2,   1/2,   1/2,     1,     0,     2,   1/2,     1,   1/2,     1,     1,  -3/2,    -2,  -1/2
 -3/2,   3/2,     0,   3/2,     2,  -1/2,    -1,    -2,     0,     0,     0,     0,     0,     0,     0,     0
end model
```

Figure 2.9: An example of MSSM compactification data.

$$M(V_2^2 - v_2^2) = 0 \qquad mod \quad 2,$$

$$M(V_1 \cdot V_2 - v_1 \cdot v_2) = 0 \qquad mod \quad 2,$$

$$N_\alpha(W_\alpha \cdot V_i) = 0 \qquad mod \quad 2,$$

$$N_\alpha(W_\alpha^2) = 0 \qquad mod \quad 2,$$

$$Q_{\alpha\beta}(W_\alpha \cdot W_\beta) = 0 \qquad mod \quad 2 \qquad (\alpha \neq \beta)$$

where $Q_{\alpha\beta}$ denotes the greatest common divisor of N_α and N_β, $\beta = 1, ..., 6$.

While there are multiple choices for the symmetries used during compactification, we consider a specific orbifold resulting from dividing the six-dimensional spacetime R^6 by a space group S composed of a six-dimensional lattice T_6 and the discrete rotation group $Z_6{-}II$. Among the various types of compactifications in string theory, we focus on the $Z_6 - II$ compactification of the $E_8 \times E_8$ heterotic string, because of its relative success in identifying MSSM-like parameter sets, as MSSM-like configurations have been found comparatively frequently within this setup. In the case of $Z_6 - II$, the twist vector is given by $v = 1/6(1, 2, -3)$, and the specification uses the shift vector V_1 along with three Wilson lines W_3 $(= W_4)$, W_5, and W_6. These must similarly satisfy the lattice conditions of the $E_8 \times E_8$ symmetry and constraints from the modular conditions. As a result, the parameters specifying the compactification amount to 64, exemplified in Fig. 2.9 Such data is used as training data for our generative model.

2.4.2 Massless Spectra of Heterotic Orbifolds

In this subsection, we provide a very brief overview of the massless spectra of heterotic orbifolds. In heterotic string theory, closed strings on orbifolds are classified into two categories: strings that directly descend from the 10-dimensional theory before the orbifold action, such as vector multiplets and gravity multiplets of $E_8 \times E_8$ or $SO(32)$, known as the untwisted sector, and new strings that are localized around orbifold singularities and close only due to the orbifold action, referred to as the twisted sectors. Under orbifolds, non-trivial and distinct space group elements are identified as the constructing elements for twisted strings. For $g \in S$, we define a twisted boundary condition $X(\tau, \sigma + \pi) = gX(\tau, \sigma)$ on the string world-sheet. Utilizing standard conformal field theory (CFT), the equations for massless right- and left-movers, given the boundary condition g, are as follows:

$$\left(\frac{q+v_g}{2}\right)^2 - \frac{1}{2} + \delta c = 0 \quad \text{and} \quad \left(\frac{p+V_g}{2}\right)^2 + \tilde{N} - 1 + \delta c = 0$$

Here, p originates from the weight lattice of $E_8 \times E_8$ or $Spin(32)/\mathbb{Z}_2$ (in this study, we focus on $E_8 \times E_8$), while q comes from the weight lattice of either the vector or spinor representations of $SO(8)$. The term δc represents the shift in zero-point energy, and $\tilde{N}$ denotes the number of left-moving oscillators. Furthermore, we denote the local twist as $v_g = kv_1 + lv_2$ and the local shift as $V_g = kV_1 + lV_2 + n\alpha W_\alpha$. The massless states are represented as tensor products of massless right- and left-movers ' states, such that they are invariant under the entire orbifold action. This is expressed by the following equations:

$$|q+v_g\rangle_R \otimes |p+V_g\rangle_L \quad \text{or} \quad |q+v_g\rangle_R \otimes \tilde{\alpha}|p+V_g\rangle_L,$$

where $\tilde{\alpha}$ represents possible oscillator excitations. We define the shifted momenta $q_{\text{sh}} = q + v_g$ and $p_{\text{sh}} = p + V_g$, where p_{sh} describes the transformation properties under gauge transformations. The states in the above equations correspond to the massless fields of the 4D effective field theory. These fields carry gauge charges derived from p_{sh}, discrete R charges, modular weights obtained from q_{sh} and the possible oscillator excitations, as well as discrete non-R charges that come from the constructing element g of the space group S.

Chapter 3

Review of Method Development

In this chapter, we describe the method development undertaken for experiments of generating MSSM data in this study. We provide an overview of the methodology, explain the motivation for its use, and describe the adjustments made in preparation for the experiments.

3.1 Preprocessing of Training Data

Each element of the training data prepared using orbifolder [15] is a fraction within a finite range, and these were converted to integers by multiplying the elements corresponding to shift vector and Wilson lines on the compactification parameters by an integer factor. This process resulted in integers that fall within the range from -8 to 9. The input data, ranging from -8 to 9, represents integers with discrete properties where the order of values does not imply magnitude. Such data often undergo a process known as one-hot encoding, where they are converted into vectors composed solely of ones and zeros. The length of this vector is 18, corresponding to the number of possible integer values, with each dimension of the vector corresponding to each integer value. A value of one in the vector indicates the presence of the corresponding integer. By principle, a one-hot vector has only one element set to one, with all others being zeros.

3.2 Hierarchical Quantized Autoencoder Used in This Study

In this section, we describe the Hierarchical Quantized Autoencoder (HQA) [23] used as the generative model in our experiments. The HQA is based on the structure of Autoencoders (AEs) that incorporate a discrete latent space, specifically, the Vector Quantized Variational Autoencoders (VQ-VAEs) [22], and introduces stochastic quantization and a hierarchical latent structure into this framework. The application of VQ-VAEs, which handle discrete spaces, has been widespread and successful in the fields of information compression and generation. The HQA further incorporates stochastic quantization and a hierarchical structure, which, together, have been successful in achieving interpolation between two inputs. The choice of HQA for this study was motivated by the ease of training due to its being a likelihood-based model with a loss function that facilitates learning through the minimization process. This process iteratively updates the model parameters based on the gradient. The affinity of our training parameter space with discrete spaces and the capability of HQA to achieve interpolation were also considered. Additionally, the motivation for selecting the HQA, which is related to AE, was also influenced by prior research using AE for the $Z_6 - II$ compactification [14], which had achieved a certain level of success.

An overview of the HQA structure used in this study is depicted in Fig. 3.1. The rectangular cells in the figure represent vectors, with blue indicating continuous space and red indicating discrete space. In the input and output spaces of the entire HQA model, the vector dimension is either 18 or 19, depending on the one-hot encoding applied. In each discrete latent space, the vector dimension, referred to as the codebook dimension, and the number of possible vectors for each discrete vector (i.e., codebook slots) were set to 64 and 256, respectively, as per the original paper. The corresponding continuous space vectors also share the same dimension as the codebook dimension. Although the original HQA model was designed with a Convolutional Neural Network (CNN) architecture optimized for image data, we adapted the model to a more expressive neural network (NN) architecture to handle 64 discrete numerical parameters instead of image data. Given the limited number of elements in the original model's training data and the need for efficient experimentation, the hierarchical structure was designed to perform compression and restoration twice, with each compression rate set to 1/4. This is analogous to the original model, where image dimensions were halved in both the height and width during compression. Furthermore, the loss function used for the input space in this study was modified to Cross Entropy to accommodate one-hot encoded vectors, while the loss function for the latent space in the second layer training remained as the Mean Squared Error (MSE). In this study, the HQA model is trained in a manner similar to the original paper. First, the encoder, decoder, and discrete hidden vectors for the first layer are trained. Afterward, these parameters are fixed, and the second layer is added. The training of the second layer is then carried out with the first layer's parameters remaining unchanged.

The loss function in Hierarchical Quantized Autoencoders (HQA) is given by:

$$L = \underbrace{-\log p(x|z=k)}_{\text{reconstruction loss}} - \underbrace{H[q(z|x)]}_{\text{entropy}} + \underbrace{\mathbb{E}_{q(z|x)}||z_e(x) - e_z||_2^2}_{\text{probabilistic commitment loss}} .$$

In our research, since the input data consists of discrete values represented as one-hot vectors, the reconstruction loss term was implemented using Cross Entropy. In training the second layer, the encoded continuous parameters from the first layer were treated as input data, and, as in the original HQA methodology, Mean Squared Error (MSE) was employed as the reconstruction loss.

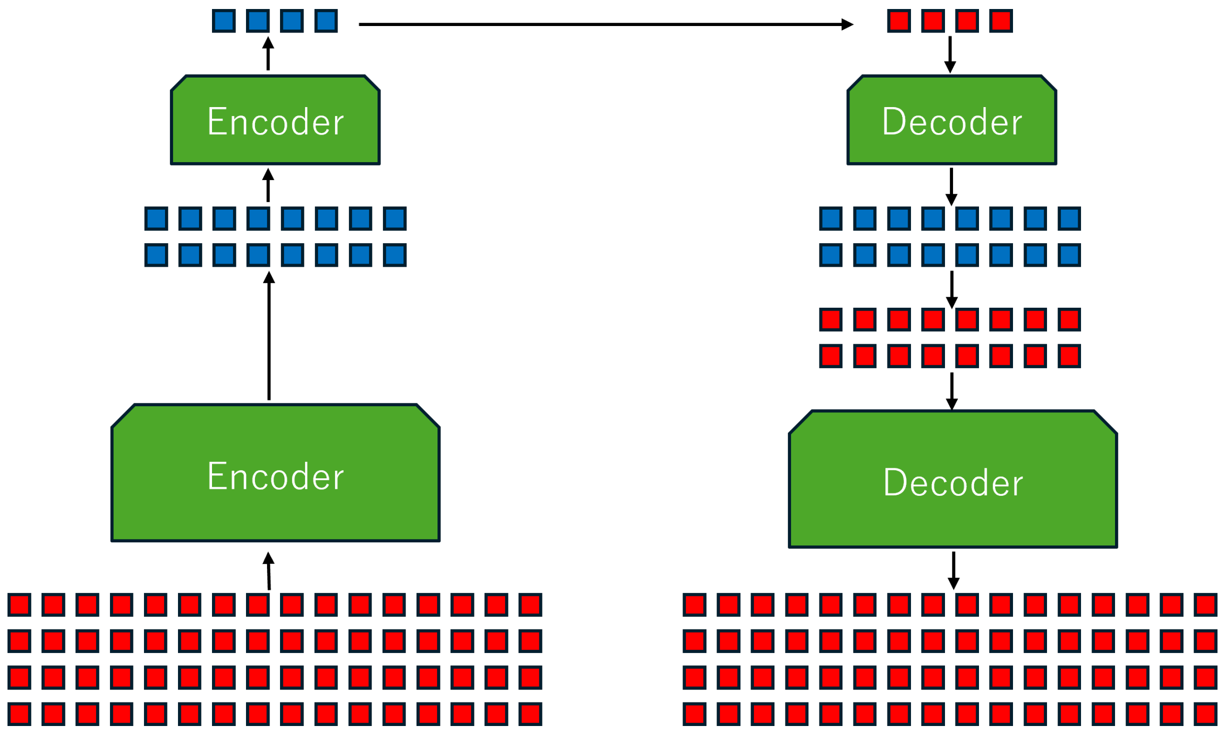

Figure 3.1: An overview of the HQA structure used in this study.

3.3 Emphasized Denoising Used in This Study

In the generation of discrete numerical data for this study, unlike continuous data such as image data, slight deviations in values are not permissible, and strict adherence to the compactification conditions is required, making the learning process more challenging. To enhance the model's fidelity in capturing the distribution of training data, we introduced emphasized denoising (ED) [19]. ED is implemented in Autoencoders (AEs) to prevent overfitting to an identity function-like trivial compression and reconstruction. During training, a certain proportion of the input data elements is randomly dropped out at each input, and the loss function is calculated in such a way that the model learns to reconstruct the dropped elements solely from the information of the remaining elements.

For simplicity, the original paper considers a linear autoencoder (LAE) represented as $B = UV^T \in \mathbb{R}^{m\times m}$, where $U, V \in \mathbb{R}^{m\times k}$ correspond to the encoder and decoder matrices, respectively. Let $Z^{(n)}_{u,i} = \delta_{u,i}X^{(n)}_{u,i}$ (no sum) denote the dropped–ut training data at epoch n, where $\delta_{u,i} \in \{0, 1\}$ follows a Bernoulli distribution applied for dropout.

The LAE is trained by minimizing the following objective:

$$\hat{B} = \arg\min_{B} \lim_{n\to\infty} \frac{1}{n} \left\| A^{(n)1/2} \odot \left(X^{(n)} - Z^{(n)} \cdot B\right)\right\|_F^2 ,$$

where $\|\cdot\|_F$ denotes the Frobenius norm and $A^{(n)}$ is a weighting matrix defined as:

$$A^{(n)}_{u,i} = \begin{cases} a, & \text{if } \delta_{u,i} = 0, \\ b, & \text{otherwise,} \end{cases}$$

with $a \geq b$.

During training, the reconstruction error for the elements that have been dropped out is up-weighted by a factor of a, thus promoting the learning of the dropped–ut elements. In contrast, the remaining elements are weighted by b, ensuring that they are learned less intensively, thereby preventing the model from simply memorizing the input data.

By setting $b = 0$, the reconstruction of certain information solely from other information is encouraged; therefore, this approach was adopted in our study. Furthermore, the loss function of HQA was modified to accommodate this scheme. With ED, elements of the training data are dropped out, and since the data is treated as one-hot vectors in this case, the size of the one-hot vector was increased by one to 19. This additional 19th element indicates a dropout when it is set to one. ED can be considered a form of data amplification with changing the original data distribution, and in this research, it is treated as normal data augmentation as opposed to phased data augmentation.

Chapter 4

Phased Data Augmentation

Within the vast compactification parameter space, MSSM-like models are relatively scarce, and traditional random scans using orbifolder [15] have only been able to prepare a limited number of datasets for training generative AI. Generally, training with limited data presents significant challenges. In this chapter, against such a backdrop, we will explain the "phased data augmentation" technique [13] we previously proposed as a method to effectively train generative AI, particularly those based on VAE, with limited data.

4.1 Necessity and Concept of Phased Data Augumentation

The advent of generative models has marked a significant milestone in the realm of artificial intelligence, showcasing an extraordinary ability to craft images of remarkable fidelity and diversity. These models' success is predominantly tethered to their training on extensive and varied datasets. Nonetheless, the acquisition of large, domain-specific datasets presents a formidable challenge, notably in fields where data collection is inherently costly, such as in the generation of medical images. The quest to surmount this obstacle underpins the essence of our research, wherein we unveil a novel, straightforward methodology to train generative models effectively, even when data resources are scant.

In the expansive domain of deep learning, the concept of transfer learning has been recognized as a powerful instrument, particularly valuable in scenarios where extensive datasets are scarce. This approach aims to boost the performance of a model within a target domain by leveraging the knowledge acquired from a related but distinct source task. However, it's crucial to acknowledge that transferring knowledge from a source that bears little resemblance to the target domain can inadvertently impair the model's performance in the target domain, an adverse effect known as negative transfer. Simultaneously, while the technique of data augmentation is widely adopted to artificially inflate the size of datasets, it can sometimes lead to a distortion of the original data's distribution. Moreover, the landscape of existing methodologies for efficient training with limited data [9,20,24] is predominantly geared towards Generative Adversarial Networks (GANs) [6], with other types of generative models receiving comparatively less attention.

Our prior work [13] introduces "phased data augmentation" as a pioneering step towards addressing these challenges. This technique strategically moderates the intensity of data augmentation to align with the model's progression through its learning phases. In the initial stages, it broadens the dataset's effective size, facilitating the model's learning of general patterns. As training progresses, we gradually refine the augmentation parameters, honing the model's focus on the salient features intrinsic to the original dataset. While rooted in conventional data augmentation strategies, our method distinguishes itself by its applicability to a wide array of generative models beyond the realm of GANs.

We demonstrate the efficacy of our approach through its application to a model that synergizes PixelCNNs [2, 21] with VQ-VAE-2 [18], designated as PC-VQ2. This model stands out for its ability to generate images that are not only sharper but also exhibit greater variety than many existing models, without relying on the adversarial architecture characteristic of GANs. The integration of PixelCNNs with the hierarchical structure of VQ-VAE-2 serves to further refine the precision of the generated images.

Our empirical analyses unequivocally affirm the superiority of the phased data augmentation technique over conventional data augmentation methods, both in quantitative benchmarks and qualitative evaluations. This underlines its potential as a highly effective strategy for training likelihood-based models in the face of limited data availability. The consistency of these results across diverse data domains and sampled datasets further underscores the robustness of our method, highlighting its capacity to significantly enhance model performance even when data resources are constrained.

4.2 Phased Data Augmentation

In this section, we delve into the intricacies of our novel training strategy, "phased data augmentation," elaborating on its application within the PC-VQ2 model and addressing the challenge of overfitting in generative models with constrained datasets.

4.2.1 Overfitting in Generative Models

Generative models thrive on large and diverse datasets, often necessitating tens of thousands of images for optimal performance. Limited datasets pose a significant challenge, leading to overfitting—a scenario where models, including GANs, fail to generalize beyond their training data, resulting in poor-quality image generation. This challenge extends to the PC-VQ2 model, which, in the absence of adequate data augmentation, tends to produce images resembling noise or, at best, unnaturally distorted outputs even with conventional augmentation techniques.

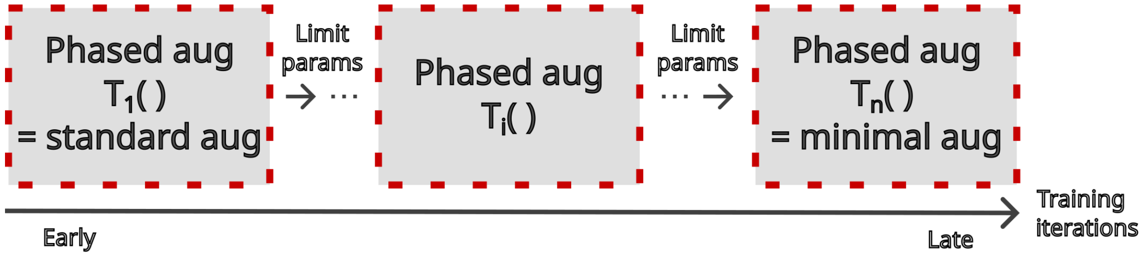

Figure 4.1: Graphical representation of phased data augmentation, as shown in [13].

This necessitates a more nuanced approach to data augmentation, as standard methods often fall short in preserving the natural diversity and quality of generated images.

4.2.2 Methodology of Phased Data Augmentation

To mitigate the challenges associated with training on limited datasets, we propose a novel yet intuitive training methodology known as "phased data augmentation." This strategy, initially introduced in our previous work [13], is underpinned by a simple yet effective principle, executed in a series of progressive steps designed to dynamically adapt the intensity of data augmentation in accordance with the model's evolving learning stages. The essence of phased data augmentation is encapsulated in the following sequential phases:

1. Initiation with Comprehensive Augmentation: The process commences with the application of a broad spectrum of label-preserving data augmentations. This initial phase is critical for introducing a diverse range of variations within the limited dataset, thereby expanding the model's exposure to a wider array of patterns and scenarios.

2. Progressive Restriction of Augmentation Intensity: As the model advances through its training journey, the scope of data augmentations is systematically narrowed down. This phased reduction is meticulously planned to occur in incremental stages, each designed to gradually align the training data closer to its original distribution, while still leveraging the benefits of augmentation to prevent overfitting.

3. Transition to Minimal Augmentation: The final phase of this strategy marks a shift towards minimal data augmentation, carefully calibrated to preserve the core distribution of the original training set. This ensures that the model, now finely tuned and more adept at recognizing the nuances of the data, focuses on mastering the detailed characteristics inherent to the dataset, thereby enhancing its predictive accuracy and generalization capabilities.

A graphical illustration of this phased approach is depicted in Fig. 4.1, visually summarizing the methodical reduction of data augmentation intensity across different stages of the training process.

What distinguishes phased data augmentation from conventional augmentation practices is its foundational principle of adaptive augmentation. Unlike the static application of standard data augmentation, our method introduces a dynamic element to the augmentation process, tailoring it to the model's learning curve. This flexibility allows phased data augmentation to be universally applicable across various model architectures, extending its utility beyond the confines of GANs. By replacing the one-size-fits-all approach of standard data augmentation with our phased strategy, we offer a versatile and effective tool for enhancing model performance, particularly in scenarios constrained by limited data availability.

In the previous our paper's experiments for generating images, this phased approach begins with extensive use of flipping, rotation, zooming, and color-space adjustments, with each phase strategically reducing the extent of these augmentations based on their impact on the data's distribution. However, rotations, zooming, and color adjustments are systematically scaled back, maintaining a balance that fosters both learning from diverse patterns and fine-tuning to the nuanced details of the original dataset.

Initial phases included broad transformations, with subsequent phases gradually dialing back on these parameters. This gradual restriction was informed by empirical observations and aimed at preserving the integrity of the original dataset's distribution. The phased approach not only combats overfitting by enriching the training dataset's diversity but also fine-tunes the model's sensitivity to the intrinsic details of the data, as illustrated by the progressive refinement of generated images through each phase of training.

The phased data augmentation strategy showcases a clear advantage over standard augmentation methods, particularly in the context of training generative models like PC-VQ2 on limited datasets. This approach ensures that the models are not just learning to replicate their training data but are also retaining and applying the nuanced patterns and characteristics essential for generating high-fidelity images. Through meticulous implementation and a strategic reduction of augmentation intensity, phased data augmentation emerges as a pivotal technique in enhancing the performance and output quality of generative models faced with the challenge of limited data.

4.3 Results and Discussions for Phased Data Augmentation

In this section, we present a comprehensive evaluation of the "phased data augmentation" strategy in our previous works [13], specifically its impact on the performance of the PC-VQ2 model trained with limited datasets. This analysis juxtaposes phased data augmentation against standard augmentation techniques, with a focus on quantitative metrics and qualitative assessments.

The evaluation employs the Fréchet Inception Distance (FID) score, a widely recognized metric for assessing the quality and diversity of generated images, to objectively measure the performance of the trained PC-VQ2 models. A lower FID score indicates closer similarity between generated and real images, thus reflecting better model performance.

Table 4.1: FID results over trained PC-VQ2 models, illustrating findings from (ref phased).

Method	FFHQ 0-99	10,000-10,099	20,000-20,099	AFHQ v2 Cat
Standard data augmentation	263.21	240.66	252.87	259.24
Phased data augmentation	**169.62**	**140.46**	**149.63**	**177.33**

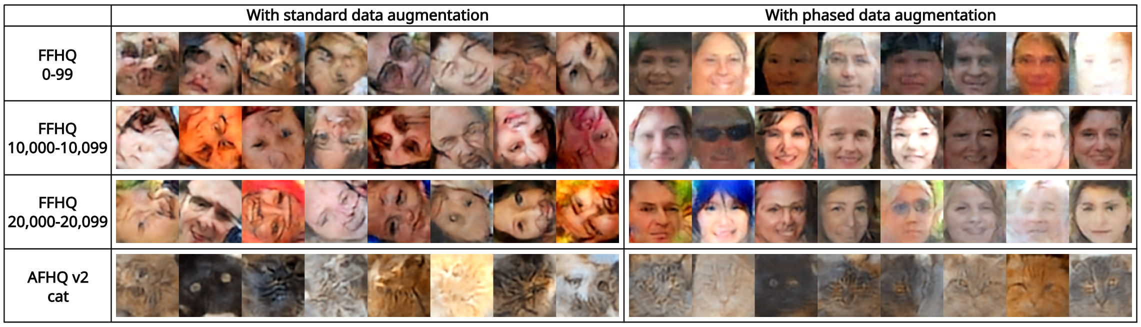

Figure 4.2: Generated human-face and cat-face images by the trained PC-VQ2 models, illustrating findings from [13].

For the empirical analysis, we utilized two high-quality datasets: the FFHQ dataset [11], comprising human-face images, and the AFHQ v2 cat dataset [3, 10, 17], consisting of cat-face images. We drew three distinct subsets from the FFHQ dataset (indices 0-99, 10,000-10,099, and 20,000-20,099) and a random selection of 100 images from the AFHQ v2 cat dataset to train and evaluate our models.

As indicated in Table 4.1, the phased data augmentation method consistently outperformed the standard augmentation approach across all subsets, achieving significantly lower FID scores. This pattern of results not only highlights the superiority of phased data augmentation in generating higher-quality images but also underscores its robustness across diverse data domains and limited datasets.

Figures 4.2 and 4.3 visually complement the quantitative FID scores by showcasing the qualitative differences in the images generated by the PC-VQ2 models trained with phased data augmentation versus standard or no data augmentation. The qualitative results support the quantitative results. Those generated with no augmentation are predominantly noisy and lack coherence.

4.4 Conclusions for Phased Data Augmentation

Our previous research [13] marks a first step in the realm of data-efficient training methodologies for generative models, without being constrained by the structure of GANs. By introducing "phased data augmentation", this study pioneers an approach tailored for enhancing the performance of likelihood-based generative models, such as PC-VQ2, within the constraints of limited

Figure 4.3: Generated images by the trained PC-VQ2 model with no data augmentation, illustrating findings from [13].

datasets. Our findings robustly affirm the superiority of phased data augmentation over conventional augmentation methods, demonstrating consistent performance enhancements across a spectrum of data domains and datasets. This underscores the method's adaptability and effectiveness in leveraging minimal data resources to achieve significant improvements in model output quality.

While the specific hyperparameters employed in this investigation were shown to be highly effective, surpassing traditional augmentation approaches across diverse datasets, it is acknowledged that these parameters may not represent the universally optimal set. Nonetheless, the observed consistency and efficacy underline the potential of phased data augmentation as a versatile and powerful tool, not only for the PC-VQ2 model but also potentially for a broad range of other generative models and in various transfer learning scenarios.

Building upon the foundation of conventional data augmentation, phased data augmentation offers a promising avenue for extending the utility of data-efficient training techniques to a wider array of generative models. While our results are compelling, the exploration of this strategy's applicability to other likelihood-based models remains an open avenue for future research. It is our aspiration that this work not only elevates the performance and understanding of the PC-VQ2 model but also catalyzes further exploration into likelihood-based models, enabling their more effective training with limited data resources.

In conclusion, phased data augmentation emerges as a potent strategy, capable of significantly enhancing the quality of generated images by likelihood-based models trained on limited datasets. This approach advocates for a phased adjustment of data augmentation intensity, aligning with the model's learning progression—a methodology that promises to enrich the toolkit available for training generative models under data scarcity.

Chapter 5

Generation of MSSM-like Data with Phased Data Augmentation

In this chapter, we present the experimental results and engage in discussion regarding the generation of MSSM data, which constitutes the main part of our study.

5.1 Phased Data Augmentation for Generating MSSM data

In this section, we discuss our approach to phased data augmentation [13] for dealing with the challenges of generating MSSM data from a limited dataset. It is widely acknowledged that learning from small datasets often does not yield successful outcomes. This challenge persists in the realm of generative tasks, where learning from limited datasets frequently fails to produce desirable results, prompting research into limited dataset learning methodologies. Data augmentation techniques have been widely utilized for learning from limited datasets, achieving successful outcomes. However, in the case of generative models for images, which have been more extensively studied, these techniques encounter a problem: they result in the generation of transformed data due to the data augmentation applied during training.

To mitigate the drawbacks of data augmentation and to harness the benefits of augmentation during learning from limited data, we introduced phased data augmentation in this study. While emphasized denoising (ED) is employed as a normal augmentation technique in this study, they involve dropping out information from the training data, subsequently attempting to approximate and restore this information, by averaging n encoded data points after randomly applying dropout for data, incurring computational costs, detailed in Appendix A. By applying phased data augmentation to ED, our goal was to eliminate the need for such approximation while still benefiting from the advantages offered by normal augmentation. Preliminary experiments confirmed that generating similar outcomes with no augmentation, even with 10,000 generated samples, was unsuccessful.

In the preliminary experiments, the original 422 MSSM data points were divided into 294 training and 128 validation data points for the learning process. For the normal augmentation condition, the dropout rate p was set to 0.5, while for the no augmentation condition, dropout was not applied ($p = 0$). In the phased augmentation condition, we employed a simplified approach where the dropout rate was reduced from 0.5 to 0.25, and finally to 0 in phases. We set mostly identical hyperparameters across three different conditions for comparison: phased data augmentation, normal data augmentation, and no data augmentation. The learning rate was uniformly set to 4e-4 for all conditions. The number of epochs was kept consistent across all conditions for both the training of the first and second layers. For the normal and no augmentation conditions, we used 3000 epochs. In contrast, for the phased augmentation condition, we fine-tuned the following phases using 10 additional epochs, starting from the pre-trained first layer of the normal augmentation condition. This approach was taken to simplify the comparison and experimental process.

During generation, a single original data point selected randomly was used across all conditions, and points for generation were produced by sampling 10,000 points from a normal distribution with a standard deviation (std) of four, centered around the point in the continuous latent space encoded from the original data.

The original data and the experimental results are shown in Figs. 5.1, 5.2, 5.3, and 5.4. We observed that both the phased data augmentation and the normal data augmentation support the generation of similar outputs with the same gauge symmetries but different compactification parameters. However, in the case of no data augmentation, we did not obtain results supporting the generation of similar outputs. In the case of the normal data augmentation, approximation is still being performed, resulting in additional computational costs.

```
begin model
Label:Model_SM1420
SpaceGroup:Geometry/Geometry_Z6-II_G2xSU3xSU2^2.txt
Lattice:E8xE8
Shifts and Wilsonlines:
    0,     0,   1/6,   1/6,   1/6,   1/6,   1/3,   1/3, -1/12, -1/12, -1/12, -1/12,  1/12,  1/12,  1/12,  1/12
    0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
    0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
    0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
    0,     0,   1/3,   1/3,     1,   4/3,  -2/3,     1,  -2/3,     0,   1/3,   1/3,  -2/3,     0,     0,   4/3
    0,     0,   1/3,   1/3,     1,   4/3,  -2/3,     1,  -2/3,     0,   1/3,   1/3,  -2/3,     0,     0,   4/3
 -1/4,   5/4,  -1/4,   1/4,   3/4,  -1/4,   7/4,   7/4,    -2,   1/2,     1,     1,   1/2,    -1,     0,    -2
    0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
end model

 Gauge group in vev-configuration "TestConfig1": SU(3) x SU(2) x SU(2) and SU(5) x SU(2) and U(1)^7
  First U(1) is anomalous with tr Q_anom = 16.00.
```

Figure 5.1: Parameters of the original data.

```
begin model
Label:Model_775
SpaceGroup:Geometry/Geometry_Z6-II_G2xSU3xSU2^2.txt
Lattice:E8xE8
Shifts and Wilsonlines:
    0,     0,   1/6,   1/6,   1/6,   1/6,   1/3,   1/3, -1/12, -1/12, -1/12, -1/12,  1/12,  1/12,  1/12,  1/12
    0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
    0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
    0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
    0,     0,   1/3,   1/3,     1,   4/3,    -1,   4/3,  -2/3,     0,   1/3,   1/3,  -2/3,     0,     0,   4/3
    0,     0,   1/3,   1/3,     1,   4/3,    -1,   4/3,  -2/3,     0,   1/3,   1/3,  -2/3,     0,     0,   4/3
 -1/4,   5/4,  -1/4,   1/4,   3/4,  -1/4,   7/4,   7/4,    -2,   1/2,     1,     1,   1/2,    -1,     0,    -2
    0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
end model

 Gauge group in vev-configuration "TestConfig1": SU(3) x SU(2) x SU(2) and SU(5) x SU(2) and U(1)^7
   First U(1) is anomalous with tr Q_anom = 16.00.
```

Figure 5.2: Parameters of the generated data with the phased data augmentation.

5.2 Generating New MSSM Data

This section details the experimental results of generating new MSSM data on post-compactification spectra not present in the training data.

The trained model with phased data augmentation in the previous section was used for this experiment. During generation, all the generated data were processed to retain only those differing from all the 422 original data, and then checked for the emergence of new data. Two data points selected randomly were input into the trained model to locate the midpoint in the latent space obtained, around which 100,000 samples were taken by normal distribution with a std of 10. As a result, two datasets differing from the original compactification spectra were obtained, with one confirming the presence of a $U(1)$ spectrum belonging to a different linear space than a similar original one.

The newly generated data and their corresponding similar original data's post-compactification massless matter field spectra and compactification parameters are as shown in Figs. 5.5, 5.6, 5.7, and 5.8. These were calculated using the orbifolder's print summary command [15]. The subscript $_l$ denotes left-chiral fields. The numbers appearing before the representations indicate multiplicities, followed by the dimensions of the representations in brackets for $SO(10)$, $SU(3)$, $SU(2)$, and $SU(2)$, respectively. The negative sign represents a "bar," indicating complex con-

```
begin model
Label:Model_3376
SpaceGroup:Geometry/Geometry_Z6-II_G2xSU3xSU2^2.txt
Lattice:E8xE8
Shifts and Wilsonlines:
    0,     0,   1/6,   1/6,   1/6,   1/6,   1/3,   1/3, -1/12, -1/12, -1/12, -1/12,  1/12,  1/12,  1/12,  1/12
    0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
    0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
    0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
    0,     0,   1/3,   1/3,     1,   4/3,    -1,   4/3,  -2/3,     0,   1/3,   1/3,  -2/3,     0,     0,   4/3
    0,     0,   1/3,   1/3,     1,   4/3,    -1,   4/3,  -2/3,     0,   1/3,   1/3,  -2/3,     0,     0,   4/3
 -1/4,   5/4,  -1/4,   1/4,   3/4,  -1/4,   7/4,   7/4,    -2,   1/2,     1,     1,   1/2,    -1,     0,    -2
    0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
end model

 Gauge group in vev-configuration "TestConfig1": SU(3) x SU(2) x SU(2) and SU(5) x SU(2) and U(1)^7
   First U(1) is anomalous with tr Q_anom = 16.00.
```

Figure 5.3: Parameters of the generated data with the normal data augmentation.

jugate representations. For instance, 3 bar, $\bar{3}$, signifies the complex conjugate representation of $SU(3)$. The numbers inside the parentheses on the right represent the charges for each $U(1)$. Since this generated spectrum belongs to a different linear space than the original, it does not revert to the same model with identical field types and charges under a $U(1)$ reinterpretation, thus constituting a distinct model.

Regarding the compactification parameters, the generated parameters differed from the original only in the fourth element from the left in the second-to-last row. However, it is not always guaranteed that altering one of the 64 elements will yield similar MSSM data within the explored parameter space. According to the compactification conditions, changing one value in one of the vectors requires satisfying both the congruence of the square of that vector and the congruences with other vectors. Restricting each value's range during typical sampling does not ensure new similar compactification parameters are obtained by changing just one element.

The generated and original MSSM spectra are identical. It is illustrated in Fig. 5.9.

This result was analyzed using the orbifolder's command, confirming that the VEV-configuration of the above orbifold model allows for vacua with the SM gauge group, $SU(3)_C \times SU(2)_L \times U(1)_Y$, including three generations of quarks and leptons and vector-like exotics [15].

The subscript n and the like on the right side denote indices for right-handed neutrinos and other SM particles. The subscripts v, w, and x denote indices for the other particles, respectively, with b representing the bar, indicating charge conjugated particles. Considering the numbers appearing before the representations indicate multiplicities, for the five types of fields in the SM and up-type and down-type Higgs doublets of MSSM, excluding right-handed neutrinos, the number of generations for fields, excluding charge conjugate pairs, is confirmed to be three in the figure. Although the MSSM spectrum remains the same, as mentioned, it does not revert to the same model with identical field types and charges under a $U(1)$ reinterpretation, resulting in a distinct model.

For the second set of generated data, while it has not been confirmed whether it belongs to a spectrum in a different linear space than a similar original, the resulting spectrum, as obtained directly from orbifolder, is different, possessing the same compactification group but significantly differing compactification parameters.

```
begin model
Label:Model_162
SpaceGroup:Geometry/Geometry_Z6-II_G2xSU3xSU2^2.txt
Lattice:E8xE8
Shifts and Wilsonlines:
 -2/3,     0,     0,     0,     0,   1/6,   1/6,   1/3,  -1/3,     0,     0,     0,     0,     0,     0,     0
    0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
    0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
    0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
  5/6,  -7/6,   1/6,   1/6,   1/6,  -7/6,   5/6,   3/2,  -1/3,  -2/3,   1/3,   1/3,   1/3,   1/3,   1/3,   2/3
  5/6,  -7/6,   1/6,   1/6,   1/6,  -7/6,   5/6,   3/2,  -1/3,  -2/3,   1/3,   1/3,   1/3,   1/3,   1/3,   2/3
  9/4,   3/4,  -5/4,  -1/4,   7/4,   3/4,   7/4,   1/4,     0,     0,     0,     0,     0,     0,     0,     0
    0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     2,    -2,     0,     0
end model

 Gauge group in vev-configuration "TestConfig1": SU(3) x SU(2) and SU(7) and U(1)^7
   First U(1) is anomalous with tr Q_anom = 578.67.

begin model
Label:Model_1579
SpaceGroup:Geometry/Geometry_Z6-II_G2xSU3xSU2^2.txt
Lattice:E8xE8
Shifts and Wilsonlines:
 -1/3,     0,     0,     0,   1/6,   1/6,   1/6,   1/6,     0,     0,     0,     0,     0,   1/6,   1/6,   2/3
    0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
    0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
    0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
 -5/6,  -1/2,  -1/2,   1/6,  -7/6,   1/6,   1/6,   7/6,    -1,     0,     0,     0,   1/3,     0,   2/3,  -2/3
 -5/6,  -1/2,  -1/2,   1/6,  -7/6,   1/6,   1/6,   7/6,    -1,     0,     0,     0,   1/3,     0,   2/3,  -2/3
  9/4,  -5/4,   7/4,  -1/4,  -5/4,  -3/4,   3/4,   3/4,  -3/2,   1/2,   1/2,   1/2,     1,    -2,    -1,    -1
    0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
end model

 Gauge group in vev-configuration "TestConfig1": SU(3) x SU(2) x SU(2) and SO(8) and U(1)^8
   First U(1) is anomalous with tr Q_anom = 74.67.

begin model
Label:Model_4087
SpaceGroup:Geometry/Geometry_Z6-II_G2xSU3xSU2^2.txt
Lattice:E8xE8
Shifts and Wilsonlines:
    0,     0,   1/6,   1/6,   1/6,   1/6,   1/2,   1/2,  -1/4, -1/12, -1/12, -1/12, -1/12,  1/12,  1/12,  1/12
    0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
    0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
    0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
 -2/3,     0,  -2/3,     0,   1/3,   1/3,   1/3,   1/3,  -7/6,  -1/2,   1/6,   1/6,   1/6,  -7/6,   1/6,   3/2
 -2/3,     0,  -2/3,     0,   1/3,   1/3,   1/3,   1/3,  -7/6,  -1/2,   1/6,   1/6,   1/6,  -7/6,   1/6,   3/2
    0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
    0,     1,   3/2,   1/2,   1/2,     2,    -1,   3/2,     0,     0,     0,     0,     0,     0,     0,     0
end model

 Gauge group in vev-configuration "TestConfig1": SU(3) x SU(2) and SO(8) x SU(3) and U(1)^7
   First U(1) is anomalous with tr Q_anom = 186.67.
```

Figure 5.4: Parameters of the generated data with no data augmentation.

```
Gauge group in vev-configuration"TestConfig1": SO(10) and SU(3) x SU(2) x SU(2) and U(1)^7
  First U(1) is anomalous with tr Q_anom = 1.5e+02.

   1 ( 10,  1,  1,  1)_l  U(1) : (  8/3,    0,  -10,   -8,    0,   24,    0)
   1 (  1, -3,  2,  2)_l  U(1) : (  2/3,    0,    0,    0,    0,  -23,   -2)
   1 (  1,  1,  1,  1)_l  U(1) : (  2/3,    0,    0,   21,    3,    6,    0)
   1 (  1,  3,  1,  1)_l  U(1) : (  2/3,    0,    0,   21,   -1,    6,    4)
   1 (  1,  3,  1,  1)_l  U(1) : (   -2,    0,    0,  -21,    1,   40,    0)
   1 (  1, -3,  1,  1)_l  U(1) : (    2,    0,    0,   21,   -1,  -40,    0)
   1 (  1,  3,  1,  1)_l  U(1) : (  4/3,    0,    0,    0,    0,  -46,   -4)
   1 (  1,  1,  2,  2)_l  U(1) : (  4/3,    0,    0,   21,   -1,  -17,    2)
   1 (  1,  1,  1,  2)_l  U(1) : (-16/9,    0,  -10,   -8,    0,  -45,    0)
   2 (  1,  1,  1,  1)_l  U(1) : (-34/9,    0,  -10,   -8,    0,   24,    0)
   2 (  1,  1,  2,  1)_l  U(1) : ( 20/9,    2,   -4,    1,    1,   20,   -2)
   2 (  1,  1,  1,  2)_l  U(1) : (  2/9,   -2,   14,  -14,    0,  -27,    0)
   1 (  1, -3,  1,  1)_l  U(1) : (-16/9,    0,  -10,   13,   -1,  -16,    0)
   2 (  1,  1,  1,  1)_l  U(1) : ( 20/9,    2,   -4,    1,    1,   20,    4)
   1 (  1,  1,  1,  1)_l  U(1) : (-16/9,   -2,   14,  -14,    0,   42,    0)
   2 (  1, -3,  1,  1)_l  U(1) : (  2/9,   -2,   14,    7,   -1,    2,    0)
   2 (  1,  1,  1,  1)_l  U(1) : (  8/9, -4/3,-52/3, 34/3,  2/3, 82/3,  8/3)
   2 (  1,  3,  2,  1)_l  U(1) : (-10/9,  2/3, 26/3,-17/3, -1/3, 28/3,  2/3)
   1 (  1,  1,  1,  1)_l  U(1) : (-16/9,  2/3, 26/3,-80/3,  2/3, 10/3,  8/3)
   2 (  1, -3,  1,  1)_l  U(1) : (  2/9,  2/3, 26/3,-17/3, -1/3,-110/3,  8/3)
   2 (  1,  1,  1,  1)_l  U(1) : (-22/9,  2/3, 26/3,-17/3, -1/3,166/3, -4/3)
   1 (  1, -3,  1,  1)_l  U(1) : ( -4/9,  2/3, 26/3, 46/3, -4/3, 46/3, -4/3)
   1 (  1,  1,  2,  1)_l  U(1) : (  8/9,  2/3, 26/3, 46/3, -4/3,-92/3,  2/3)
   3 (  1,  1,  1,  2)_l  U(1) : ( -4/9,  2/3, 26/3,-17/3, -1/3,-41/3, -4/3)
   2 (  1,  1,  1,  1)_l  U(1) : (-22/9, 10/3, 10/3,-55/3,  1/3, 50/3,  4/3)
   2 (  1,  1,  1,  2)_l  U(1) : (-10/9, 10/3, 10/3,  8/3, -2/3, -1/3, -8/3)
   2 (  1, -3,  1,  1)_l  U(1) : (-10/9, -8/3,-14/3, 14/3, -2/3,-88/3,  4/3)
   1 (  1, -3,  1,  1)_l  U(1) : ( -4/9, 10/3, 10/3,  8/3, -2/3,-70/3,  4/3)
   1 (  1,  1,  1,  1)_l  U(1) : (-28/9, -8/3,-14/3,-49/3,  1/3, 32/3,  4/3)
   1 (  1,  1,  2,  1)_l  U(1) : ( 26/9, -2/3,  4/3,-22/3,  4/3, 20/3, -2/3)
   1 (  1,  1,  1,  2)_l  U(1) : (-16/9, -8/3,-14/3, 14/3, -2/3,-19/3, -8/3)
   2 (  1,  1,  1,  1)_l  U(1) : ( 20/9, -2/3,  4/3, 41/3,  1/3,176/3,  4/3)
   2 (  1,  1,  2,  1)_l  U(1) : ( -7/3,    2,    1,    5,    1,    8,   -2)
   2 (  1,  1,  1,  1)_l  U(1) : (   -1,    2,    1,    5,   -1,  -38,   -4)
   2 (  1,  1,  1,  1)_l  U(1) : (    1,   -2,   -1,   -5,    1,   38,    4)
   2 (  1, -3,  1,  1)_l  U(1) : ( -7/3,    2,    1,    5,   -1,    8,    0)
   2 (  1,  1,  1,  2)_l  U(1) : (   -1,    2,    1,   26,    0,   -9,    0)
   2 (  1,  1,  1,  2)_l  U(1) : (    1,   -2,   -1,  -26,    0,    9,    0)
   4 (  1,  3,  1,  1)_l  U(1) : (  7/3,   -2,   -1,   -5,    1,   -8,    0)
   4 (  1,  1,  2,  1)_l  U(1) : (  7/3,   -2,   -1,   -5,   -1,   -8,    2)
   1 (  1,  1,  1,  1)_l  U(1) : (-20/9,  2/3, -4/3,-41/3, -1/3,-176/3, -4/3)
   2 (  1,  1,  2,  1)_l  U(1) : (-26/9,  2/3, -4/3, 22/3, -4/3,-20/3,  2/3)
   2 (  1,  1,  1,  1)_l  U(1) : ( 28/9,  8/3, 14/3, 49/3, -1/3,-32/3, -4/3)
   1 (  1,  3,  1,  1)_l  U(1) : ( 10/9,  8/3, 14/3,-14/3,  2/3, 88/3, -4/3)
   1 (  1,  1,  1,  2)_l  U(1) : ( 10/9,-10/3,-10/3, -8/3,  2/3,  1/3,  8/3)
   2 (  1,  1,  1,  2)_l  U(1) : ( 16/9,  8/3, 14/3,-14/3,  2/3, 19/3,  8/3)
   2 (  1,  3,  1,  1)_l  U(1) : (  4/9,-10/3,-10/3, -8/3,  2/3, 70/3, -4/3)
   1 (  1,  1,  1,  1)_l  U(1) : ( 22/9,-10/3,-10/3, 55/3, -1/3,-50/3, -4/3)
   2 (  1,  3,  1,  1)_l  U(1) : (  4/9, -2/3,-26/3,-46/3,  4/3,-46/3,  4/3)
   2 (  1,  1,  2,  1)_l  U(1) : ( -8/9, -2/3,-26/3,-46/3,  4/3, 92/3, -2/3)
   1 (  1,  3,  1,  1)_l  U(1) : ( -2/9, -2/3,-26/3, 17/3,  1/3,110/3, -8/3)
   1 (  1, -3,  2,  1)_l  U(1) : ( 10/9, -2/3,-26/3, 17/3,  1/3,-28/3, -2/3)
   1 (  1,  1,  1,  1)_l  U(1) : ( -8/9,  4/3, 52/3,-34/3, -2/3,-82/3, -8/3)
   1 (  1,  1,  1,  1)_l  U(1) : ( 22/9, -2/3,-26/3, 17/3,  1/3,-166/3,  4/3)
   2 (  1,  1,  1,  1)_l  U(1) : ( 16/9, -2/3,-26/3, 80/3, -2/3,-10/3, -8/3)
   3 (  1,  1,  1,  2)_l  U(1) : (  4/9, -2/3,-26/3, 17/3,  1/3, 41/3,  4/3)
   1 (  1,  3,  1,  1)_l  U(1) : ( -2/9,    2,  -14,   -7,    1,   -2,    0)
   1 (  1,  1,  1,  2)_l  U(1) : ( -2/9,    2,  -14,   14,    0,   27,    0)
   1 (  1,  1,  2,  1)_l  U(1) : (-20/9,   -2,    4,   -1,   -1,  -20,    2)
   1 (  1,  1,  1,  1)_l  U(1) : ( 34/9,    0,   10,    8,    0,  -24,    0)
   2 (  1,  1,  1,  1)_l  U(1) : ( 16/9,    2,  -14,   14,    0,  -42,    0)
   1 (  1,  1,  1,  1)_l  U(1) : (-20/9,   -2,    4,   -1,   -1,  -20,   -4)
   2 (  1,  3,  1,  1)_l  U(1) : ( 16/9,    0,   10,  -13,    1,   16,    0)
   2 (  1,  1,  1,  2)_l  U(1) : ( 16/9,    0,   10,    8,    0,   45,    0)
   2 (  1,  1,  1,  1)_l  U(1) : ( -1/9,    1,   -7,-49/2,  3/2,  -30,    0)
   2 (  1, -3,  1,  2)_l  U(1) : ( -1/9,    1,   -7, -7/2,  1/2,   -1,    0)
   2 (  1,  3,  1,  1)_l  U(1) : ( -1/9,    1,   -7, 35/2, -1/2,   28,    0)
   2 (  1,  1,  1,  1)_l  U(1) : ( -7/9,    1,   -7, -7/2,  1/2,   22,   -4)
   2 (  1,  1,  2,  1)_l  U(1) : ( -7/9,    1,   -7, -7/2,  1/2,   22,    2)
   2 (  1,  1,  1,  1)_l  U(1) : ( 11/9,    1,   -7, 35/2,  3/2,  -18,    0)
   4 (  1,  1,  1,  1)_l  U(1) : (  5/9,    1,   -7, -7/2, -3/2,  -24,    0)
   2 ( 10,  1,  1,  1)_l  U(1) : ( 17/9,    0,    5,    4,    0,  -12,    0)
   2 (  1,  1,  1,  1)_l  U(1) : ( 11/9,    4,   -3,    6,    0,  -18,    0)
   2 (  1,  1,  1,  1)_l  U(1) : ( -7/9,    0,   15,   12,    0,  -36,    0)
   4 (  1,  1,  1,  1)_l  U(1) : (  5/9,   -2,  -11,    8,    0,  -24,    0)
   2 (  1,  1,  1,  1)_l  U(1) : (-25/9, -1/3, 17/3,-61/6,  7/6,-104/3, -4/3)
   2 (  1,  1,  1,  1)_l  U(1) : (-19/9, -1/3, 17/3, 65/6,-11/6,-86/3, -4/3)
   2 (  1,  1,  1,  1)_l  U(1) : ( 11/9,  5/3, 35/3,-133/6,  7/6,  4/3, -4/3)
   2 (  1,  1,  1,  1)_l  U(1) : ( 17/9,  5/3, 35/3, -7/6,-11/6, 22/3, -4/3)
   2 (  1,  1,  1,  1)_l  U(1) : (-13/9,  8/3, 29/3, 61/3, -1/3,-68/3, -4/3)
   2 (  1,  1,  1,  1)_l  U(1) : (  5/9, -4/3, 23/3,-95/3, -1/3,-14/3, -4/3)
   2 (  1,  1,  1,  2)_l  U(1) : (  5/9, -4/3, 23/3,-32/3,  2/3, 73/3,  8/3)
   2 (  1,  1,  1,  1)_l  U(1) : ( -1/9,  2/3,-49/3, 49/3, -1/3,-32/3, -4/3)
   4 (  1,  1,  1,  1)_l  U(1) : ( 17/9, -4/3, 23/3, 31/3, -1/3, 22/3, -4/3)
   2 (  1,  1,  2,  1)_l  U(1) : ( -1/9, -5/3, -5/3,-71/6,  5/6, 26/3, 10/3)
   2 (  1,  1,  1,  2)_l  U(1) : ( -7/9, -5/3, -5/3,-71/6, -7/6, 95/3,  4/3)
   2 (  1,  1,  2,  2)_l  U(1) : ( 11/9, -5/3, -5/3, 55/6, -1/6,-25/3, -2/3)
   2 (  1,  1,  1,  1)_l  U(1) : ( 17/9, -5/3, -5/3, 55/6, 11/6,-94/3,  4/3)
   2 (  1,  1,  1,  1)_l  U(1) : (  5/9, -5/3, -5/3, 55/6,-13/6, 44/3, -8/3)
   2 (  1,  1,  1,  1)_l  U(1) : ( 11/9, -5/3, -5/3,-71/6, -7/6,-112/3,  4/3)
   2 (  1,  3,  1,  1)_l  U(1) : (  5/9, -5/3, -5/3, 55/6, -1/6, 44/3,  4/3)
   4 (  1,  1,  1,  1)_l  U(1) : ( -1/9, -5/3, -5/3,-71/6,  5/6, 26/3, -8/3)
   2 (  1, -3,  2,  1)_l  U(1) : (  5/9,  4/3,  7/3, -7/3,  1/3, 44/3, -2/3)
   2 (  1,  1,  1,  1)_l  U(1) : (  5/9,  4/3,  7/3, -7/3,  7/3, 44/3, -8/3)
   2 (  1,  3,  1,  1)_l  U(1) : (  5/9,  4/3,  7/3, -7/3, -5/3, 44/3,  4/3)
   2 (  1,  1,  1,  1)_l  U(1) : ( -1/9,  4/3,  7/3,-70/3, -2/3, 26/3, -8/3)
   2 (  1,  1,  1,  2)_l  U(1) : ( -1/9,  4/3,  7/3, -7/3,  1/3,113/3,  4/3)
   2 (  1,  1,  1,  1)_l  U(1) : ( 11/9,  4/3,  7/3, 56/3, -2/3, 62/3, -8/3)
   4 (  1,  1,  1,  1)_l  U(1) : ( 17/9,  4/3,  7/3, -7/3,  1/3,-94/3,  4/3)
```

Figure 5.5: Generated data's spectrum

```
Gauge group in vev-configuration "TestConfig1": SO(10) and SU(3) x SU(2) x SU(2) and U(1)^7
  First U(1) is anomalous with tr Q_anom = 1.5e+02.

    1 ( 10,  1,  1,  1)_l  U(1) : (  8/3,    0,  -10,   -8,    0,   24,    0)
    1 (  1, -3,  2,  2)_l  U(1) : (  2/3,    0,    0,    0,    0,  -23,   -2)
    1 (  1,  1,  1,  1)_l  U(1) : (  2/3,    0,    0,   21,    3,    6,    0)
    1 (  1,  3,  1,  1)_l  U(1) : (  2/3,    0,    0,   21,   -1,    6,    4)
    1 (  1,  3,  1,  1)_l  U(1) : (   -2,    0,    0,  -21,    1,   40,    0)
    1 (  1, -3,  1,  1)_l  U(1) : (    2,    0,    0,   21,   -1,  -40,    0)
    1 (  1,  3,  1,  1)_l  U(1) : (  4/3,    0,    0,    0,    0,  -46,   -4)
    1 (  1,  1,  2,  2)_l  U(1) : (  4/3,    0,    0,   21,   -1,  -17,    2)
    1 (  1,  1,  1,  2)_l  U(1) : (-16/9,    0,  -10,   -8,    0,  -45,    0)
    2 (  1,  1,  1,  1)_l  U(1) : (-34/9,    0,  -10,   -8,    0,   24,    0)
    2 (  1,  1,  2,  1)_l  U(1) : ( 20/9,    2,   -4,    1,    1,   20,   -2)
    2 (  1,  1,  1,  2)_l  U(1) : (  2/9,   -2,   14,  -14,    0,  -27,    0)
    1 (  1, -3,  1,  1)_l  U(1) : (-16/9,    0,  -10,   13,   -1,  -16,    0)
    2 (  1,  1,  1,  1)_l  U(1) : ( 20/9,    2,   -4,    1,    1,   20,    4)
    1 (  1,  1,  1,  1)_l  U(1) : (-16/9,   -2,   14,  -14,    0,   42,    0)
    2 (  1, -3,  1,  1)_l  U(1) : (  2/9,   -2,   14,    7,   -1,    2,    0)
    2 (  1,  1,  1,  1)_l  U(1) : (  8/9, -4/3,-52/3, 34/3,  2/3, 82/3,  8/3)
    2 (  1,  3,  2,  1)_l  U(1) : (-10/9,  2/3, 26/3,-17/3, -1/3, 28/3,  2/3)
    1 (  1,  1,  1,  1)_l  U(1) : (-16/9,  2/3, 26/3,-80/3,  2/3, 10/3,  8/3)
    2 (  1, -3,  1,  1)_l  U(1) : (  2/9,  2/3, 26/3,-17/3, -1/3,-110/3,  8/3)
    2 (  1,  1,  1,  1)_l  U(1) : (-22/9,  2/3, 26/3,-17/3, -1/3,166/3, -4/3)
    1 (  1, -3,  1,  1)_l  U(1) : ( -4/9,  2/3, 26/3, 46/3, -4/3, 46/3, -4/3)
    1 (  1,  1,  2,  1)_l  U(1) : (  8/9,  2/3, 26/3, 46/3, -4/3,-92/3,  2/3)
    3 (  1,  1,  1,  2)_l  U(1) : ( -4/9,  2/3, 26/3,-17/3, -1/3,-41/3, -4/3)
    2 (  1,  1,  1,  1)_l  U(1) : (-22/9, 10/3, 10/3,-55/3,  1/3, 50/3,  4/3)
    2 (  1,  1,  1,  2)_l  U(1) : (-10/9, 10/3, 10/3,  8/3, -2/3, -1/3, -8/3)
    2 (  1, -3,  1,  1)_l  U(1) : (-10/9, -8/3,-14/3, 14/3, -2/3,-88/3,  4/3)
    1 (  1, -3,  1,  1)_l  U(1) : ( -4/9, 10/3, 10/3,  8/3, -2/3,-70/3,  4/3)
    1 (  1,  1,  1,  1)_l  U(1) : (-28/9, -8/3,-14/3,-49/3,  1/3, 32/3,  4/3)
    1 (  1,  1,  2,  1)_l  U(1) : ( 26/9, -2/3,  4/3,-22/3,  4/3, 20/3, -2/3)
    1 (  1,  1,  1,  2)_l  U(1) : (-16/9, -8/3,-14/3, 14/3, -2/3,-19/3, -8/3)
    2 (  1,  1,  1,  1)_l  U(1) : ( 20/9, -2/3,  4/3, 41/3,  1/3,176/3,  4/3)
    2 (  1,  1,  2,  1)_l  U(1) : ( -7/3,    2,    1,    5,    1,    8,   -2)
    2 (  1,  1,  1,  1)_l  U(1) : (   -1,    2,    1,    5,   -1,  -38,   -4)
    2 (  1,  1,  1,  1)_l  U(1) : (    1,   -2,   -1,   -5,    1,   38,    4)
    2 (  1, -3,  1,  1)_l  U(1) : ( -7/3,    2,    1,    5,   -1,    8,    0)
    2 (  1,  1,  1,  2)_l  U(1) : (   -1,    2,    1,   26,    0,   -9,    0)
    2 (  1,  1,  1,  2)_l  U(1) : (    1,   -2,   -1,  -26,    0,    9,    0)
    4 (  1,  3,  1,  1)_l  U(1) : (  7/3,   -2,   -1,   -5,    1,   -8,    0)
    4 (  1,  1,  2,  1)_l  U(1) : (  7/3,   -2,   -1,   -5,   -1,   -8,    2)
    1 (  1,  1,  1,  1)_l  U(1) : (-20/9,  2/3, -4/3,-41/3, -1/3,-176/3, -4/3)
    2 (  1,  1,  2,  1)_l  U(1) : (-26/9,  2/3, -4/3, 22/3, -4/3,-20/3,  2/3)
    2 (  1,  1,  1,  1)_l  U(1) : ( 28/9,  8/3, 14/3, 49/3, -1/3,-32/3, -4/3)
    1 (  1,  3,  1,  1)_l  U(1) : ( 10/9,  8/3, 14/3,-14/3,  2/3, 88/3, -4/3)
    1 (  1,  1,  1,  2)_l  U(1) : ( 10/9,-10/3,-10/3, -8/3,  2/3,  1/3,  8/3)
    2 (  1,  1,  1,  2)_l  U(1) : ( 16/9,  8/3, 14/3,-14/3,  2/3, 19/3,  8/3)
    2 (  1,  3,  1,  1)_l  U(1) : (  4/9,-10/3,-10/3, -8/3,  2/3, 70/3, -4/3)
    1 (  1,  1,  1,  1)_l  U(1) : ( 22/9,-10/3,-10/3, 55/3, -1/3,-50/3, -4/3)
    2 (  1,  3,  1,  1)_l  U(1) : (  4/9, -2/3,-26/3,-46/3,  4/3,-46/3,  4/3)
    2 (  1,  1,  2,  1)_l  U(1) : ( -8/9, -2/3,-26/3,-46/3,  4/3, 92/3, -2/3)
    1 (  1,  3,  1,  1)_l  U(1) : ( -2/9, -2/3,-26/3, 17/3,  1/3,110/3, -8/3)
    1 (  1, -3,  2,  1)_l  U(1) : ( 10/9, -2/3,-26/3, 17/3,  1/3,-28/3, -2/3)
    1 (  1,  1,  1,  1)_l  U(1) : ( -8/9,  4/3, 52/3,-34/3, -2/3,-82/3, -8/3)
    1 (  1,  1,  1,  1)_l  U(1) : ( 22/9, -2/3,-26/3, 17/3,  1/3,-166/3,  4/3)
    2 (  1,  1,  1,  1)_l  U(1) : ( 16/9, -2/3,-26/3, 80/3, -2/3,-10/3, -8/3)
    3 (  1,  1,  1,  2)_l  U(1) : (  4/9, -2/3,-26/3, 17/3,  1/3, 41/3,  4/3)
    1 (  1,  3,  1,  1)_l  U(1) : ( -2/9,    2,  -14,   -7,    1,   -2,    0)
    1 (  1,  1,  1,  2)_l  U(1) : ( -2/9,    2,  -14,   14,    0,   27,    0)
    1 (  1,  1,  2,  1)_l  U(1) : (-20/9,   -2,    4,   -1,   -1,  -20,    2)
    1 (  1,  1,  1,  1)_l  U(1) : ( 34/9,    0,   10,    8,    0,  -24,    0)
    2 (  1,  1,  1,  1)_l  U(1) : ( 16/9,    2,  -14,   14,    0,  -42,    0)
    1 (  1,  1,  1,  1)_l  U(1) : (-20/9,   -2,    4,   -1,   -1,  -20,   -4)
    2 (  1,  3,  1,  1)_l  U(1) : ( 16/9,    0,   10,  -13,    1,   16,    0)
    2 (  1,  1,  1,  2)_l  U(1) : ( 16/9,    0,   10,    8,    0,   45,    0)
    2 (  1,  1,  1,  1)_l  U(1) : (-19/9,    1,    3,  9/2, -3/2,  -48,    0)
    2 (  1,  1,  1,  1)_l  U(1) : ( 17/9,    3,    9,-15/2, -3/2,  -12,    0)
    2 (  1,  1,  1,  1)_l  U(1) : ( -1/9,   -3,    1,-11/2,  1/2,   28,   -4)
    2 (  1,  1,  2,  1)_l  U(1) : ( -1/9,   -3,    1,-11/2,  1/2,   28,    2)
    2 (  1,  1,  1,  1)_l  U(1) : ( 17/9,   -3,    1, 31/2,  3/2,  -12,    0)
    2 (  1,  1,  1,  1)_l  U(1) : ( 11/9,   -3,    1,-11/2, -3/2,  -18,    0)
    2 ( 10,  1,  1,  1)_l  U(1) : ( 17/9,    0,    5,    4,    0,  -12,    0)
    2 (  1,  1,  1,  1)_l  U(1) : ( 11/9,    4,   -3,    6,    0,  -18,    0)
    2 (  1,  1,  1,  1)_l  U(1) : ( -7/9,    0,   15,   12,    0,  -36,    0)
    4 (  1,  1,  1,  1)_l  U(1) : (  5/9,   -2,  -11,    8,    0,  -24,    0)
    2 (  1,  1,  1,  2)_l  U(1) : ( -7/9, -1/3,-13/3,-109/6, -5/6, 37/3,  8/3)
    2 (  1, -3,  1,  2)_l  U(1) : ( -1/9, -1/3,-13/3, 17/6,  1/6, 55/3, -4/3)
    2 (  1,  1,  2,  2)_l  U(1) : ( 11/9, -1/3,-13/3, 17/6,  1/6,-83/3,  2/3)
    2 (  1,  3,  1,  1)_l  U(1) : (  5/9, -1/3,-13/3, 17/6,  1/6,-14/3,  8/3)
    2 (  1,  1,  2,  1)_l  U(1) : ( -7/9, -1/3,-13/3, 17/6,  1/6,124/3,  2/3)
    2 (  1,  1,  1,  1)_l  U(1) : ( 11/9, -1/3,-13/3,143/6,  7/6,  4/3, -4/3)
    4 (  1,  1,  1,  1)_l  U(1) : ( -1/9, -1/3,-13/3,-109/6,  7/6,-32/3, -4/3)
    4 (  1,  1,  1,  1)_l  U(1) : (  5/9, -1/3,-13/3, 17/6,-11/6,-14/3, -4/3)
    2 (  1,  1,  1,  1)_l  U(1) : (-13/9,  8/3, 29/3, 61/3, -1/3,-68/3, -4/3)
    2 (  1,  1,  1,  1)_l  U(1) : (  5/9, -4/3, 23/3,-95/3, -1/3,-14/3, -4/3)
    2 (  1,  1,  1,  2)_l  U(1) : (  5/9, -4/3, 23/3,-32/3,  2/3, 73/3,  8/3)
    2 (  1,  1,  1,  1)_l  U(1) : ( -1/9,  2/3,-49/3, 49/3, -1/3,-32/3, -4/3)
    4 (  1,  1,  1,  1)_l  U(1) : ( 17/9, -4/3, 23/3, 31/3, -1/3, 22/3, -4/3)
    2 (  1,  1,  1,  1)_l  U(1) : (  5/9,  7/3,-29/3,-59/6, -7/6,-130/3,  4/3)
    2 (  1,  3,  1,  1)_l  U(1) : ( -1/9,  7/3,-29/3, 67/6, -1/6, 26/3,  4/3)
    2 (  1,  1,  1,  1)_l  U(1) : (-25/9, -5/3, 25/3,-23/6,  5/6,-46/3, -8/3)
    2 (  1,  1,  1,  1)_l  U(1) : ( 11/9,  1/3, 43/3,-95/6,  5/6, 62/3, -8/3)
    2 (  1,  1,  1,  1)_l  U(1) : ( -7/9,  7/3,-29/3,-59/6,  5/6,  8/3, -8/3)
    2 (  1, -3,  2,  1)_l  U(1) : (  5/9,  4/3,  7/3, -7/3,  1/3, 44/3, -2/3)
    2 (  1,  1,  1,  1)_l  U(1) : (  5/9,  4/3,  7/3, -7/3,  7/3, 44/3, -8/3)
    2 (  1,  3,  1,  1)_l  U(1) : (  5/9,  4/3,  7/3, -7/3, -5/3, 44/3,  4/3)
    2 (  1,  1,  1,  1)_l  U(1) : ( -1/9,  4/3,  7/3,-70/3, -2/3, 26/3, -8/3)
    2 (  1,  1,  1,  2)_l  U(1) : ( -1/9,  4/3,  7/3, -7/3,  1/3,113/3,  4/3)
    2 (  1,  1,  1,  1)_l  U(1) : ( 11/9,  4/3,  7/3, 56/3, -2/3, 62/3, -8/3)
    4 (  1,  1,  1,  1)_l  U(1) : ( 17/9,  4/3,  7/3, -7/3,  1/3,-94/3,  4/3)
```

Figure 5.6: Original data's spectrum

```
begin model
Label:Model_4517
SpaceGroup:Geometry/Geometry_Z6-II_G2xSU3xSU2^2.txt
Lattice:E8xE8
Shifts and Wilsonlines:
  -1/2,  -1/3,     0,     0,     0,     0,     0,   1/6,  -1/3,  -1/3,     0,     0,   1/6,   1/6,   1/6,   1/6
     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
   4/3,    -1,     0,     0,     0,     0,     0,     1,   1/6,   7/6,   1/6,   5/6,  -7/6,  -1/6,  -1/6,   7/6
   4/3,    -1,     0,     0,     0,     0,     0,     1,   1/6,   7/6,   1/6,   5/6,  -7/6,  -1/6,  -1/6,   7/6
   1/2,   1/2,    -1,     1,     0,     0,     2,    -1,     0,   3/2,    -1,     0,     2,  -3/2,     2,    -2
     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
end model
```

Figure 5.7: Compactification parameters of the generated data

```
begin model
Label:Model_SM3321
SpaceGroup:Geometry/Geometry_Z6-II_G2xSU3xSU2^2.txt
Lattice:E8xE8
Shifts and Wilsonlines:
  -1/2,  -1/3,     0,     0,     0,     0,     0,   1/6,  -1/3,  -1/3,     0,     0,   1/6,   1/6,   1/6,   1/6
     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
   4/3,    -1,     0,     0,     0,     0,     0,     1,   1/6,   7/6,   1/6,   5/6,  -7/6,  -1/6,  -1/6,   7/6
   4/3,    -1,     0,     0,     0,     0,     0,     1,   1/6,   7/6,   1/6,   5/6,  -7/6,  -1/6,  -1/6,   7/6
   1/2,   1/2,    -1,     0,     0,     0,     2,    -1,     0,   3/2,    -1,     0,     2,  -3/2,     2,    -2
     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0,     0
end model
```

Figure 5.8: Compactification parameters of the original data

```
Gauge group in vev-configuration "SMConfig1": SU(3)_C and SU(2)_L and U(1)_Y

139 (  1,  1)_l  U(1) : (    0)  n_1 - n_86
  5 ( -3,  2)_l  U(1) : ( -1/6)  q_1 - q_4
  3 (  1,  1)_l  U(1) : (   -1)  be_1 - be_3
  3 (  3,  1)_l  U(1) : (  2/3)  bu_1 - bu_3
 15 (  3,  1)_l  U(1) : ( -1/3)  bd_1 - bd_15
 12 ( -3,  1)_l  U(1) : (  1/3)  d_1 - d_12
 10 (  1,  2)_l  U(1) : (  1/2)  l_1 - l_9
  7 (  1,  2)_l  U(1) : ( -1/2)  bl_1 - bl_7
  2 (  3,  2)_l  U(1) : (  1/6)  bq_1 bq_2
 16 (  1,  1)_l  U(1) : ( -1/2)  v_1 - v_16
  4 ( -3,  1)_l  U(1) : ( -1/6)  w_1 w_2
  4 (  3,  1)_l  U(1) : (  1/6)  bw_1 - bw_4
  8 (  1,  2)_l  U(1) : (    0)  x_1 - x_6
 16 (  1,  1)_l  U(1) : (  1/2)  bv_1 - bv_14
```

Figure 5.9: MSSM spectrum of the generated data and the original data

Chapter 6

Summary and Discussions

In the search for MSSM-like compactification parameters of $E_8 \times E_8$ heterotic strings' orbifold compactifications, traditional approaches have predominantly relied on random search. Even with a specific orbifold, finding a relatively small number of MSSM-like models from the vast compactification parameter space has been difficult. Focusing on $Z_6 - II$ orbifolds compactification, considered to yield comparatively more MSSM-like models, a research [14] has proposed using unsupervised learning techniques, specifically AE, to map to a two-dimensional continuous space. This approach aimed to identify "islands" where MSSM-like models cluster, thereby streamlining the search within corresponding subspaces of the compactification parameter space. However, identifying specific MSSM-like models remained difficult, making it challenging to extract physically meaningful insights across multiple models.

Our study focuses on $Z_6 - II$ orbifolds and employs generative AI to effectively obtain insights based on commonalities and differences among specific MSSM models. As a first step in this direction, we utilized a generative AI, HQA, related to AEs, specifically VAE-based AIs, to demonstrate the generation of new data that shares gauge symmetry with the compactification models prepared as training data but differs in the spectrum after compactification. Given the limited number of MSSM-like models prepared via random scans, training AI on such limited data is generally challenging. Therefore, we combined "phased data augmentation" techniques proven effective even with limited data in our previous research for VAE-based generative AI. The experiments also showed the generation of data similar to a particular MSSM model, sharing gauge symmetry but differing in compactification parameters.

This work also represents an initial step in utilizing generative AI within the field of particle physics. While application of supervised, unsupervised, and reinforcement learning have been explored in recent particle physics research, generative AI has the potential to learn data distributions without label information. It can capture similarities and differences in a denser latent space, having the potential to conduct an effective search based on similarities learned from the distribution of data. Although applied to specific physics data in this study, there is potential for broader application across other particle physics data. Moreover, applying generative AI to discrete numerical data, as done in this research, has significance beyond physics as an example of its application.

While the findings of this study represent only a first step, further research in this direction could lead to the generation of models that transition between two points of interest through interpolation. This process could provide an efficient method to glean insights from the similarities and differences between MSSM models, contributing to the search for MSSM models.

Appendix A

Comparison of Approximation Methods for Normal Augmentation in HQA

In this appendix, we supplement the explanation of the approximation methods used for normal augmentation in the Hierarchical Quantized Autoencoder (HQA) model. In this study, the difficulty of generating discrete numerical data necessitated the use of emphasized denoising (ED) during the training of HQA. However, using ED implies that the model trained with dropout-applied data must be used during generation, requiring an approximation method.

Unlike the linear ED used in the original paper, this study employs a nonlinear neural network for ED. Moreover, dropout is applied to the training data itself rather than the model parameters in ED, leading to the consideration of the following three generation methods:

1. **Generation with Dropout Applied (No Approximation)**: Only the information remaining after dropout is used from the original data, which may be inappropriate, especially for generating similar data.

2. **Generation without Dropout and without Data Correction (No Approximation)**: Since only dropout-applied data is used during training, this approach is also likely inappropriate for generating similar data.

3. **Averaging after n Encodes with Dropout Applied n Times (Approximation)**: By applying dropout multiple times and averaging the encoded outputs, this method is expected to approximate the full information from the original data. The learning process in HQA, which encourages similar data to be mapped close together in latent space, should positively impact this method. However, increasing n to improve approximation accuracy proportionally increases computational costs.

We conducted a comparative experiment with these three methods. To focus on the approximation method during generation with ED, the training data used was not limited to MSSM but included data from the Z_6 orbifold. A total of 56,384 training samples were used, with 512 samples set aside for validation during training.

For the experiment, 1,000 data samples were input into the trained model, and generation was conducted based on the latent space values. The comparison was made based on the number of generated data samples that met the compactification conditions, without explicitly comparing the similarity of the generated data.

In the original paper, continuous variable data was used, and an approximation was made by scaling the input data by the dropout probability p. However, in this study, the input data is discrete and trained using one-hot encoding, meaning the values in the input during training are only 0 and 1, making the original approximation method unsuitable.

The results of the simplified comparison experiment are as follows:

- **1. Generation with Dropout Applied:**
  - 3 orbifolds loaded.
  - Three types of orbifolds were generated.
  - However, only 50% of the original data's information was used, making this approach particularly inappropriate for generating similar data.
- **2. Generation without Dropout and without Data Correction:**
  - 0 orbifolds loaded.
  - This method failed to generate any valid orbifold data, confirming that it is inappropriate for generating similar data.
- **3. Averaging after n Encodes with Dropout Applied n Times:**
  - 6 orbifolds loaded.
  - Three types of orbifolds were generated.
  - Increasing n allows for a better coverage of all the original data's information, though at the cost of increased computational resources.

In method 2, no valid orbifold data could be generated, while method 3 successfully generated a greater variety of data, including those with 3 shift vectors and Wilson lines. The more vectors involved in a 16-dimensional embedding, the stricter the conditions to be met, making generation more difficult. However, generating data with more vectors is desirable as it increases the variety of MSSM that can be generated.

Based on these results, we adopted method 3 for data generation in this study.

Bibliography

[1] Volker Braun, Yang-Hui He, Burt A Ovrut, and Tony Pantev. A heterotic standard model. *Physics Letters B*, 618(1-4):252–258, 2005.

[2] Xi Chen, Nikhil Mishra, Mostafa Rohaninejad, and Pieter Abbeel. Pixelsnail: An improved autoregressive generative model. In *International Conference on Machine Learning*, pages 864–872. PMLR, 2018.

[3] Yunjey Choi, Youngjung Uh, Jaejun Yoo, and Jung-Woo Ha. Stargan v2: Diverse image synthesis for multiple domains. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2020.

[4] Tim Pieter Tjipko Dijkstra, LR Huiszoon, and Adrian Norbert Schellekens. Supersymmetric standard model spectra from rcft orientifolds. *Nuclear Physics B*, 710(1-2):3–57, 2005.

[5] Alon E Faraggi. Construction of realistic standard-like models in the free fermionic superstring formulation. *Nuclear Physics B*, 387(2):239–262, 1992.

[6] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in neural information processing systems*, 27, 2014.

[7] Geoffrey E Hinton and Ruslan R Salakhutdinov. Reducing the dimensionality of data with neural networks. *science*, 313(5786):504–507, 2006.

[8] Luis E Ibanez and Angel M Uranga. *String theory and particle physics: An introduction to string phenomenology.* Cambridge University Press, 2012.

[9] Tero Karras, Miika Aittala, Janne Hellsten, Samuli Laine, Jaakko Lehtinen, and Timo Aila. Training generative adversarial networks with limited data. *Advances in neural information processing systems*, 33:12104–12114, 2020.

[10] Tero Karras, Miika Aittala, Samuli Laine, Erik Härkönen, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. Alias-free generative adversarial networks. In *Proc. NeurIPS*, 2021.

[11] Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 4401–4410, 2019.

[12] Diederik P Kingma. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.

[13] Yuta Mimura. Phased data augmentation for training a likelihood-based generative model with limited data. *to be published in ITE Transactions on Media Technology and Applications*, 2025. arXiv:2305.12681.

[14] Andreas Mütter, Erik Parr, and Patrick KS Vaudrevange. Deep learning in the heterotic orbifold landscape. *Nuclear Physics B*, 940:113–129, 2019.

[15] Hans-Peter Nilles, Saúl Ramos-Sánchez, Patrick KS Vaudrevange, and Akin Wingerter. The orbifolder: A tool to study the low-energy effective theory of heterotic orbifolds. *Computer Physics Communications*, 183(6):1363–1380, 2012.

[16] OpenAI. ChatGPT (September 25 Version). `https://chat.openai.com/chat`, 2023. [Large language model].

[17] Gaurav Parmar, Richard Zhang, and Jun-Yan Zhu. On aliased resizing and surprising subtleties in gan evaluation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11410–11420, 2022.

[18] Ali Razavi, Aaron Van den Oord, and Oriol Vinyals. Generating diverse high-fidelity images with vq-vae-2. *Advances in neural information processing systems*, 32, 2019.

[19] Harald Steck. Autoencoders that don't overfit towards the identity. *Advances in Neural Information Processing Systems*, 33:19598–19608, 2020.

[20] Ngoc-Trung Tran, Viet-Hung Tran, Ngoc-Bao Nguyen, Trung-Kien Nguyen, and Ngai-Man Cheung. On data augmentation for gan training. *IEEE Transactions on Image Processing*, 30:1882–1897, 2021.

[21] Aaron Van den Oord, Nal Kalchbrenner, Lasse Espeholt, Oriol Vinyals, Alex Graves, et al. Conditional image generation with pixelcnn decoders. *Advances in neural information processing systems*, 29, 2016.

[22] Aaron Van Den Oord, Oriol Vinyals, et al. Neural discrete representation learning. *Advances in neural information processing systems*, 30, 2017.

[23] Will Williams, Sam Ringer, Tom Ash, David MacLeod, Jamie Dougherty, and John Hughes. Hierarchical quantized autoencoders. *Advances in Neural Information Processing Systems*, 33:4524–4535, 2020.

[24] Shengyu Zhao, Zhijian Liu, Ji Lin, Jun-Yan Zhu, and Song Han. Differentiable augmentation for data-efficient gan training. *Advances in neural information processing systems*, 33:7559–7570, 2020.