Machine Learning Explorations in the String Landscape

A thesis submitted in partial fulfillment of the requirements for the award of the degree of

Doctor of Philosophy

Elli Heyes

May 2024

Department of Mathematics,

City, University of London

Contents

List of Figures

List of Tables

Acknowledgements

First and foremost, I would like thank my mother, Franki Heyes, who worked hard to give me a good education and every opportunity in life and for this I am eternally grateful. And to my father, John Heyes, and my brother, Adam Heyes, for their love and support in all my endeavours.

Secondly, I would like to extend my gratitude to my supervisor, Professor Yang-Hui He, for his mentorship during my PhD. He has not only helped me to develop my knowledge of the subject, but has made the process enjoyable and instilled in me a deep love for all areas of mathematics and physics. Thanks to him, I have received opportunities to join collaborations, coauthor book chapters, present my work at conferences and participate in workshops all around the world, all of which has allowed me to develop as an academic and for that I'm very grateful.

I would further like to thank all my collaborators, from whom I have and continue to learn so much: Daattavya Aggarwal, Guillermo Arias-Tamargo, Anthony Ashmore, Jiakang Bao, Per Berglund, Federico Carta, Siqi Chen, Man-Wai Cheung, Francesco Costantino, Pierre-Philippe Dechant, Edward Hirst, Vishnu Jejjala, Jianrong Li, Andre Lukas, David Marsh, Liam McAllister, Burt Ovrut, Dmitrii Riabchenko, Diego Rodriguez-Gomez, Henrique Sá Earp, Tomás Silva. I would like to acknowledge in particular my academic brother, Dr. Edward Hirst, for his support, mentorship and encouragement. I look forward to many more years of friendship and collaboration.

I also would like to give my thanks to the London Institute for Mathematical Sciences, who have welcomed me as a visiting member during my time as a PhD student. I have met many great people here and have learnt a lot from my discussions with members.

Finally, I would like thank all my incredible friends who support me and enrich my life, and to my yoga and callisthenics teachers who help me to get out of my head and into my body when I'm stuck in my research.

My graduate studies would not have been possible without the financial support of a studentship grant from the School of Science and Technology at City, University of London and a postgraduate grant from the States of Jersey.

Declaration

I hereby certify that the work, which is being presented in the project report, entitled **Machine Learning Explorations in the String Landscape**, in fulfillment of the requirement for the award of **Doctor of Philosophy** and submitted to the **Department of Mathematics at City, University of London** is an authentic record of my own work carried out during the period **10/2021 – 05/2024** under the supervision of **Professor Yang-Hui He**. The matter presented in this report has not been submitted elsewhere for the award of any other degree or diploma from any institutions.

Throughout the course of this PhD I have co-authored seven published papers and two preprints across a range of topics in AI for pure mathematics and string theory:

Published Papers

[Paper1] G. Arias-Tamargo, Y.-H. He, E. Heyes, E. Hirst, D. Rodriguez-Gomez, Brain Webs for Brane Webs, Physics Letters B 833 (2022). doi:10.1016/j.physletb.2022.137376.

[Paper2] P.-P. Dechant, Y.-H. He, E. Heyes, E. Hirst, Cluster Algebras: Network Science and Machine Learning, Journal of Computational Algebra 8 (2023). doi:10.1016/j.jaca.2023.100008.

[Paper3] M.-W. Cheung, P.-P. Dechant, Y.-H. He, E. Heyes, E. Hirst, J.-R. Li, Clustering Cluster Algebras with Clusters, Accepted in Advances in Theoretical and Mathematical Physics (2022). arXiv:2212.09771.

[Paper4] A. Ashmore, Y.-H. He, E. Heyes, B. Ovrut, Numerical Spectra of the Laplacian for Line Bundles on Calabi-Yau Hypersurfaces, Journal of High Energy Physics 164 (2023). doi:10.1007/JHEP07(2023)164.

[Paper5] P. Berglund, Y.-H. He, E. Heyes, E. Hirst, V. Jejjala, A. Lukas, New Calabi-Yau Manifolds from Genetic Algorithms, Phys. Lett. B 850 (2024) 138504. doi:10.1016/j.physletb.2024.138504.

[Paper6] S. Chen, P.-P. Dechant, Y.-H. He, E. Heyes, E. Hirst, D. Riabchenko, Machine Learning Clifford Invariants of ADE Coxeter Elements, Adv. Appl. Clifford Algebras 34 (3) (2024) 20. doi:10.1007/s00006-024-01325-y.

[Paper7] D. Aggarwal, Y.-H. He, E. Heyes, E. Hirst, H. N. S. Earp, T. S. R. Silva, Machine Learning Sasakian and G2 Topology on Contact Calabi-Yau 7-Manifolds, Phys. Lett. B 850 (2024) 138517. doi:10.1016/j.physletb.2024.138517.

Preprints

[Preprint1] F. Costantino, Y.-H. He, E. Heyes, E. Hirst, Learning 3-Manifold Triangulations (2024). arXiv:2405.09610.

[Preprint2] P. Berglund, G. Butbaia, Y.-H. He, E. Heyes, E. Hirst, V. Jejjala, Generating Triangulations and Fibrations with Reinforcement Learning (2024). arXiv:2405.21017.

Furthermore, I have co-authored two book chapters, reviewing recent work in the field of Machine Learning, Geometry and Physics:

Book Chapters

[Chapter1] J. Bao, Y.-H. He, E. Heyes, E. Hirst, Machine Learning Algebraic Geometry for Physics (2022). arXiv:2204.10334.

[Chapter2] Y.-H. He, E. Heyes, E. Hirst, Machine Learning in Physics and Geometry (2023). arXiv:2303.12626.

Due the broad range of topics I will not discuss all of my papers in this report. Instead I will focus on three papers, namely [Paper4, Paper5, Paper7], which are all based on the topic of compactification.

Date: 21/05/2024 **Signature of the Candidate**

Abstract

In this thesis we explore various aspects of Calabi-Yau (CY) and G_2 manifolds and string compactifications over them. We do so using techniques taken from the fields of data science and machine learning (ML).

At first we focus on generating CY manifolds, built as hypersurfaces in toric varieties, with genetic algorithms. We find new examples CY fourfolds with topology distinct from fourfolds in existing datasets. We also use these algorithms to generate CY manifolds that satisfy certain phenomenology constraints.

Knowledge of the Ricci-flat metric on a CY is required to compute particle properties of the resulting four-dimensional effective field theory. Using ML approximations of the metric, we compute the Yukawa couplings arising from the $E_8 \times E_8$ heterotic string compactified on a CY threefold X with holomorphic vector bundle V by computing the bundle-valued harmonic modes of the Laplacian operator. We consider the particular case where X is a hypersurface in a single ambient space and V is a line bundle sum.

Finally, we build a dataset of G_2 manifolds via certain contact CY manifolds, called CY links, and study their topology. We show how neural networks can be used to predict aspects of the topology of these manifolds, namely their Sasakian Hodge number $h^{2,1}$, from the list of weights $\mathbf{w}$ defining the CY link. Using symbolic regression we construct an approximate formula for $h^{2,1}$ in terms of $\mathbf{w}$. This serves as the first application of ML to G_2 geometry.

Keywords: String Theory, Calabi-Yau Manifolds, G_2 Manifolds, Machine-Learning, String Compactification

1 Introduction

In 2024 physics finds itself is in a difficult situation. The Standard Model of particle physics and the theory of general relativity are able to predict experimental data to extraordinary high degree of accuracy, however we lack a complete theory of physics that incorporates both these theories. Despite its tremendous success at predicting the behaviour of particles and forces, the Standard Model fails to describe gravity and is inconsistent with general relativity. There are also many free parameters in the Standard Model whose values are not given by the theory and have to be found experimentally. Ideally, we would like to have a more fundamental theory that explains why these parameters have the values that they do. Furthermore, the Standard Model does not explain the origin and nature of dark matter and dark energy, or why we have an excess of matter over antimatter, or why neutrinos have tiny yet nonzero masses, or why the cosmological constant is so small. These are important questions whose answers will no doubt fundamentally change our understanding of the physical world.

Out of several attempts to combine the Standard Model and general relativity into a "theory of everything" string theory remains the most promising candidate. In fact, string theory was originally introduced as a theory describing the strong interaction [1, 2], and it was only later discovered to be a theory of quantum gravity [3–5]. In the theory, zero-dimensional point particles are replaced by one-dimensional objects called strings, whose vibrational modes give rise to the different fundamental particles of the Standard Model in some perturbative sense. At low energies these strings appear point-like and the extended structure becomes apparent only at the string scale, which is assumed to be about the size of the Planck length.

Quantum field theory (QFT) is a theoretical framework, connecting special relativity and quantum mechanics, that treats particles as excited states of their underlying quantum fields [6, 7]. Just as particles sweep out worldlines, strings sweep out surfaces called worldsheets and we can formulate a quantised string theory in terms of a two-dimensional QFT on the worldsheet. Quantising an extended object places constraints on this two-dimensional field theory. In particular, the field theory must be a conformal field theory (CFT). There is a one-to-one correspondence between states of the CFT and operators that describe deformations in the background. These deformations include deformations of the

metric, i.e., they include gravitons. As such this theory includes perturbative quantum gravity and therefore overcomes the non-renormalisability problem in quantum field theory. Moreover, the extended nature of strings "smears" out the location of interactions which removes problematic ultraviolet divergences.

Research into string theory is well-motivated since it remains the strongest candidate to provide a unified fundamental theory of particle physics. There are however several reasonable arguments against the theory. Firstly, a requirement for the theory to be self-consistent is that there are extra dimensions of space, in addition to the three that we experience. Secondly, there are multiple formulations and solutions of string theory. This lack of uniqueness raises concerns about whether it can offer a definitive description of our universe. Finally, the theory makes predictions that are often difficult to test experimentally due to the extremely high energies required to observe its effects. As a result, critics argue that it lacks empirical evidence to support its validity. Despite these difficulties, we argue that string theory remains a worthwhile field of study since it provides a framework to explore connections between seemingly disparate areas of physics. In the worst case scenario where string theory fails completely as a realistic theory, the concepts and techniques can still contribute to our understanding of mathematics and fundamental physics. The classic example is the discovery of mirror symmetry, which was first discovered by physicists in the context of string theory [8] but is now its own field of study in mathematics.

A key ingredient of most string theory models is supersymmetry since string theory compactifications that preserve supersymmetry are devoid of problematic tachyons [9] . Superstring theories are ten-dimensional; one time dimension and nine spatial dimensions. There are five superstring theories [10, 11]: Type I, Type IIA, Type IIB, Heterotic $SO(32)$ and Heterotic $E_8 \times E_8$. All these theories are connected to one another and to eleven-dimensional M-theory via S and T-dualities. Superstring compactification consists of reducing the ten-dimensions of superstring theory by choosing a background of the form $M_{10} = \mathbb{R}^{1,3} \times M_6$, where M_{10} is the ten-dimensional background spacetime, $\mathbb{R}^{1,3}$ is the usual 4-dimensional Minkowski space and M_6 is a some six-dimensional compact manifold which serves as the so-called target space. The process of compactification requires a choice of compact manifold M_6 and the properties of M_6 fix almost all features of the lower dimensional theory, including the masses and charges of particles, the strength of forces and other physical constants. The low-energy effective field theories that arise from

compactifications of superstring theories are given by supergravity theories.

Since we are interested in a theory of gravity, which requires a metric $g_{\mu\nu}$, M_6 must be Riemannian. Furthermore, in order for the associated two-dimensional field theory to be conformally invariant $g_{\mu\nu}$ must be Ricci-flat. The requirement of the existence of chiral fermions enforces that M_6 be irreducible. By Berger's classification [12] this restricts the choice of holonomy of M_6 to be $U(3)$, $SU(3)$ or $SO(6)$. Finally, if we require that $\mathcal{N} = 1$ supersymmetry is preserved this fixes the holonomy of M_6 to be $SU(3)$. All together, these conditions imply that M_6 is a Calabi-Yau manifold of three complex dimensions, often called a Calabi-Yau threefold. A consistent string model, therefore, with four flat Minkowski spacetime directions can be built using any Calabi-Yau threefold CY_3 as the internal target space. If we take the radius of such a Calabi-Yau manifold to be small then the ten-dimensional spacetime $\mathbb{R}^{1,3} \times CY_3$ will effectively look just like $\mathbb{R}^{1,3}$ and hence is consistent with observations.

To understand the size of the string landscape (the space of all possible string vacua), one needs to determine the number of topologically distinct Calabi-Yau threefolds that give rise to inequivalent effective theories.[1] It is widely believed that the number of topologically distinct Calabi-Yau manifolds at any given dimension is finite although no general proof is known. If this is indeed the case it is clear however, that the set of all Calabi-Yau threefolds is vast. Wall's theorem [13] says that simply connected Calabi-Yau threefolds with torsion-free cohomology are completely classified by: the Hodge numbers $h^{1,1}$ and $h^{1,2}$, the triple intersection numbers κ_{ijk} and the second Chern class c_2. Two such threefolds with equivalent data are diffeomorphic and give rise to equivalent effective theories upon compactification. Using the result of Wall's theorem, recent works have provided an estimate of $\sim 10^{400}$ for the number of topologically-inequivalent smooth, simply connected Calabi-Yau threefolds constructed as hypersurfaces in toric varieties [14, 15]. This is a huge amount and this is just one construction, there are many other methods for constructing Calabi-Yau manifolds. Clearly there are too many possibilities for us to study each one by hand! This motivates an algorithmic approach to searching the landscape and identifying Calabi-Yau threefolds with desirable properties.

[1] Note that Calabi-Yau threefolds are just one type of compactification space. The full string landscape contains all types of compactifications.

We cannot simply state that the Standard Model arises from string theory, we must engineer a geometry that will yield exactly the required particle content, masses, couplings, etc. This is what is referred to as "string phenomenology" [8]. Below is a non-exhaustive list of information about a particular Calabi-Yau threefold that a string phenomenologist would like to have in order to study the effective theory:

- The Ricci-flat metric $g_{\mu\nu}$ on the manifold.
- Information on sub-structure of the manifold. This includes cycles that can be wrapped by branes to produce important non-perturbative physics. In particular:
  - Divisors (complex codimension one sub-varieties)
  - Curves (complex codimension two sub-varieties)
  - Special Lagrangian cycles (of real dimension three) or other non-complex subspaces
- Symmetries of the manifold. Calabi-Yau manifolds do not admit continuous isometries, but discrete symmetries can and do arise, and can play an important physical role. Applications of such symmetries include:
  - The existence of discrete Wilson lines (i.e. gauge fields for which $A_\mu \neq 0$ but $\mathcal{F}_{\mu\nu} = 0$).
  - Orbifolds and orientifolds (i.e., related to singular Calabi-Yau geometries and non-perturbative physics).
  - Flavor/family symmetries/R-symmetry.
- The existence of differential forms. For example, the type of gauge fields, 2-forms (or more generally, n-forms) that can arise on the threefold. At higher order, interactions in the potential (e.g. Yukawa couplings) can be computed as trilinear couplings of such forms.

While many string models have been explored over the years, a model with exactly the right properties to reproduce the Standard Model remains elusive. One of the first and most successful approaches to this challenge has been provided by heterotic string theory. A heterotic string model is specified by four pieces of data, a Calabi-Yau threefold X, the observable and hidden holomorphic vector bundles V and $\tilde{V}$ on X, each with a structure group contained in E_8, and a holomorphic curve $C \subset X$ with associated homology class

$W = [C] \in H_2(X, \mathbb{Z})$. Physically, the curve C is wrapped by five-branes stretching across the four-dimensional uncompactified space-time. On this data three physical constraints have to be imposed:

1. **Anomaly cancellation**: Anomaly cancellation in the heterotic string imposes a topological condition which relates the Calabi-Yau manifold X, the two vector bundles and the five-brane class W. For the case of bundles with $c_1(V) = c_1(\tilde{V}) = 0$ it can be written as $c_2(TX) - c_2(V) - c_2(\tilde{V}) = W$.

2. **Effectiveness**: To ensure four-dimensional $\mathcal{N} = 1$ supersymmetry the five-brane has to wrap a holomorphic curve. Hence, the five-brane class W must be chosen such that it indeed has a holomorphic curve representative C, with $W = [C]$. Classes $W \in H_2(X, \mathbb{Z})$ with this property are called effective.

3. **Stability**: The Donaldson-Uhlenbeck-Yau theorem guarantees the existence of a connection satisfying the Hermitian-Yang Mills equations (and, hence, preserving $\mathcal{N} = 1$ supersymmetry) on a holomorphic vector bundle, provided this bundle is (poly-)stable. Hence, both V and $\tilde{V}$ must be (poly-)stable holomorphic vector bundles on X.

Just as Calabi-Yau threefolds are the target space on which ten-dimensional string theory compactifies to give a four-dimensional effective field theory with minimal supersymmetry, seven-dimensional G_2 holonomy manifolds are the equivalent target space for eleven-dimensional M-theory.

Despite the lack of an explicit bound and a classification, the birational geometry of Calabi-Yau manifolds is relatively well-understood via the minimal model program. In particular, working within the category of algebraic varieties gives us a precise control of the Kähler cone and more refined linear and homological structures beyond just the cohomology of the underlying topological space. On the other hand, the state-of-the-art results for the geometry of compact G_2 manifolds pale in comparison, primarily due to the lack of analogous algebraic techniques. There is no analogous classification program, finiteness bounds on the cohomology, or a wealth of constructions. It is only recently that some families of G_2 manifolds have been constructed.

M-theory compactifications on G_2 holonomy manifolds do however have the advantage of

being largely geometric. The compactification geometry in superstring theory often needs to be supplemented with additional data. The presence of these additional structures can complicate the computation of various physical properties in the four-dimensional effective theory. For instance, we will see later how the computation of Yukawa couplings in $E_8 \times E_8$ heterotic string theory is non-trivial due to the presence of the holomorphic vector bundle. This explains one of the advantages of working with G_2 compactifications in M-theory.

1.1 A New Era

Before the advent of computers, all of mathematics and physics was done by hand. When computers arrived on the scene they were quickly adopted by mathematicians and physicists to speed up long and difficult computations. Since that time technology has developed tremendously and now most of us own personal computers that can carry out computations in a matter of seconds that would take years to compute by hand. Computers have undoubtedly changed the way we carry out research in science and mathematics, but in recent years a new tool has emerged that is arguably just as revolutionary - artificial intelligence. In contract to traditional computer programming where the user tells the computer exactly what to do, machine learning algorithms are capable of solving problems without explicit instructions. These tools have been used extensively for many years in experimental physics, engineering, and other areas of applied physics; by making predictions and extracting patterns from large datasets generated from simulations of experimental observations. Notably, they have aided the analysis of large datasets generated by particle colliders like the Large Hadron Collider, identifying particle signatures amidst noisy data.

Machine learning techniques are often called blackbox, meaning that they are difficult to interpret and understand due to them having a large number of parameters. One might think, therefore, that these techniques are ill-suited for application in fields such as theoretical physics and pure mathematics that prioritize rigor and understanding. On the contrary, however, machine learning has been used with great success in these fields in the past few years. The first recorded use of machine learning in string theory was in 2017 [16–19], where neural networks were used to predict properties of Calabi-Yau manifolds such as Hodge numbers of volumes, ranks of gauge groups and bundle cohomologies. Since then there has been an explosion of interest in applying neural networks and other machine learning algorithms to study Calabi-Yau manifolds and other objects in string theory and algebraic

geometry. Examples include using siamese neural networks to classify five-dimensional superconformal field theories [20], studying exchange graphs of cluster algebras with methods from network science [21, 22], learning properties of Clifford algebra invariants with principal component analysis [23], using natural language processing techniques to study Pachner graphs of three-manifold triangulations [24] and many more [25–31]. For a review of machine learning in physics and geometry see [32, 33].

One of the main motivations for bringing modern data science methods to string theory is to deal with the vast amount of data. We just discussed that there exist too many Calabi-Yau threefolds for us to study the compactification on each one by hand. The question is: Can machine learning help us identify models within the landscape with desirable properties? Furthermore, we know that in order to determine the low-dimensional effective field theory one needs to know the Ricci-flat metric on the Calabi-Yau. Unfortunately, Yau's proof for the existence of Calabi-Yau manifolds is not constructive and we do not have a method for constructing an analytical Calabi-Yau metric. The lack of explicit metrics forms a major obstacle. With little to no chance of ever discovering analytic expressions for the metrics, there has been considerable focus on using numerical methods to compute these objects. Numerous algorithms have been devised for numerically determining Ricci-flat metrics on Calabi-Yau manifolds, including position-space techniques [34] and spectral methods [35–39]. These numerical methods allow one to approximate the metric, however they are slow and computationally expensive. Another approach that has developed in recent years is to apply machine learning techniques, such as neural networks, to approximate the metric [40–48]. Results so far have been very promising, showing impressive accuracy and requiring far less computing power than existing numerical methods.

Combining these two efforts we are in a position to, for the first time, efficiently pick out suitable candidate theories from the vast landscape of string models, approximate the Ricci-flat metric on the Calabi-Yau and with this compute properties of the effective field theory that arises from this compactification and check if these properties truly match those of observed particle physics.

1.2 Outline

The research presented in this thesis is interdisciplinary across three subjects: mathematics, physics and computer science, which makes it difficult to give a thorough background in all areas. We assume that the reader is of a mathematical physics background rather than computer science, and thus our explanation of the machine learning tools used will be more rudimentary, whereas the string theory and algebraic geometry content will be pitched at the level of a graduate student.

We begin in Section 2 by giving a general introduction to string theory and string compactification, and explain why Calabi-Yau manifolds and G_2 manifolds are the appropriate compactification spaces to achieve minimal supersymmetry in four-dimensions from superstring and M-theory respectively. We also discuss in more detail the process of string model building for the particular case of the $E_8 \times E_8$ heterotic string.

In Section 3 we review the necessary parts of algebraic and differential geometry in order to understand the construction and properties of Calabi-Yau manifolds and G_2 manifolds. We begin with Berger's classification of holonomy groups for Riemannian manifolds that connects to the discussion in the previous section as to why Calabi-Yau manifolds and G_2 manifolds are the compactification spaces that preserve supersymmetry. We then proceed to discuss Hodge decomposition and de Rham cohomology. Following this we introduce Kähler manifolds, Dolbeault cohomology and holomorphic vector bundles on complex manifolds. From Kähler manifolds we move on to discuss the specific case of Calabi-Yau manifolds. We cover the definition of a Calabi-Yau manifold, their construction, their topological properties and moduli spaces of Calabi-Yau manifolds. Finally, we lay the foundations of G_2 geometry, specifically we define a G_2 structure and its torsion.

Section 4 gives a brief introduction to machine learning, focusing on the algorithms used to produce the results shown in the later sections, namely: neural networks, genetic algorithms and symbolic regression. We discuss the training and evaluation procedures as well as hyperparameter choices in each case. Specific details of the algorithms used are given in the later sections.

The subject of Section 5 is the generation of Calabi-Yau manifolds with machine learning

and presents the results published in [49]. We focus on the construction of Calabi-Yau manifolds as hypersurfaces in toric varieties built from reflexive polytopes. We show how reflexive polytopes can be generated using genetic algorithms and how constraints can be added to the fitness function to generate Calabi-Yau manifolds with desired properties. This introduces a new approach to string phenomenology, whereby the user specifies properties of the string model which they would like and uses machine learning algorithms to construct appropriate compactification spaces.

In Section 6, we switch from Calabi-Yau manifolds to G_2 manifolds and outline the results of [50]. We consider a particular type of G_2 manifold, called a Calabi-Yau link, and firstly build a dataset of such manifolds. On this dataset we study two types of topological invariants: the Crowley-Nordström homotopy invariant ν and the Hodge numbers $\{h^{3,0}, h^{2,1}\}$. Using a simple feed forward neural network we are able to predict $h^{2,1}$ from a set of input weights $\mathbf{w}$ defining the Calabi-Yau link with a high degree of accuracy. Motivated by this success we apply symbolic regression to this problem and approximate a formula $h^{2,1}$ in terms of $\mathbf{w}$. It is possible that a precise formula exists but none is known and so our results provide a good starting point for discovering such a formula.

Section 5 and 6 focused on Calabi-Yau and G_2 manifolds directly. In Section 7, however, we instead focus on computing properties of the resulting low-dimensional effective field theory given some compactification space. Specifically, we look at computing the Yukawa couplings λ_{IJK} from compactifying the $E_8 \times E_8$ heterotic string on a Calabi-Yau threefold X with holomorphic vector bundle V. Using machine-learning approximations of the Ricci-flat metric on X, we are able to calculate the harmonic modes of the bundle-valued Dolbeault Laplacian operator $\Delta_{\bar{\partial}_V}$ which are the necessary ingredients for computing λ_{IJK}. Our choices for X and V are too simple to give physically relevant string models, however this calculation serves as an example of how one can use the recent machine learning results of approximating the Calabi-Yau metric to compute Yukawa couplings in the effective field theory. The results of this section are based on the following published paper [51].

In Section 8, we conclude by emphasising the most important findings and discuss possible directions for future work.

2 Physics Background

In this section, we give a very brief overview of some of the basic facts about string theory, summarising the five superstring theories and M-theory. We introduce the notion of string compactification and describe the process of string model building for the specific case of heterotic $E_8 \times E_8$ string theory compactified on a Calabi-Yau threefold with holomorphic vector bundle. We assume the reader is familiar with the basics of string theory and if not refer them to the following great introductions to the subject [52, 53]. We will mostly follow [54].

2.1 The Standard Model

First we must mention the Standard Model (SM) of particle physics. This theory describes the three fundamental non-gravitational forces: the electromagnetic, the weak and the strong force, as well as all elementary particles. Any candidate for a "theory of everything", including string theory, must ultimately be consistent with the SM.

A key concept of the SM is gauge symmetry of the fields describing the elementary particles. The gauge group of the SM is

$$SU(3) \times SU(2) \times U(1). \tag{1}$$

This is a local symmetry, meaning its action on the fields is local. All fields under the same orbit of this action describe the same particle and therefore the particles are characterised by the representation and charges of the gauge group.[2]

The global symmetry in the SM is the isometry group of the Minkowski spacetime, which is the Poincaré group. The Poincaré group is a semidirect product of translations and the Lorentz group

$$\mathbb{R}^{1,3} \rtimes SO(1,3), \tag{2}$$

where the Lorentz group $SO(1,3)$ is the group of all changes of orthonormal frames in Minkowski spacetime. Therefore the particles are also in representations of this group.

Fermionic fields are spinor fields on spacetime so fermions split into the two irreducible spin representation of the Lorentz group, which are called as "left-handed" and

[2]Throughout this thesis, we will make several references to groups and their representations. For the sake of space, we will not introduce this topic here but we refer the reader to the following excellent references for an introduction to group and representation theory [55–57], suitable for physicists.

"right-handed" fermions.

Bosonic fields are tensorial fields on spacetime. At rank 0, we have two Higgs fields H^u and H^d as scalar fields which transform trivially under the Lorentz group. These fields are responsible for breaking electroweak symmetry ($SU(2) \times U(1)$) and giving mass to the fermions. At rank 1, we have a vector representation of the Lorentz group. This is associated with gauge bosons whose fields are vector fields on spacetime with values in the Lie algebra of the gauge group. They act as mediators for the fundamental forces. The list stops here for the SM but in string theory we have graviton bosons in the rank 2 tensor representation.

Another ingredient of the SM is the framework of Lagrangians. If ϕ (or ψ) is a bosonic field (or fermionic field), then a term in the Lagrangian of the quadratic form $m^2\phi^2$ (or $m\psi\bar{\psi}$) is called a mass term for the particle with mass m. This can be generalised to n fields with the form $m_{ij}^2\phi_i\phi_j$ (or $m_{ij}\psi_i\bar{\psi}_j$) where the masses are eigenvalues of m_{ij}, and the eigenvectors are called mass eigenstates. At the cubic level, we have terms of the form $Y_{ijk}\psi_i\bar{\psi}_j\phi_k$ where the scalar Y_{ijk} is called a Yukawa coupling and determines the strength of the interaction between two fermions ψ_i and ψ_j and a scalar ϕ_k.

2.2 String Theory

String theory is a theory of one-dimensional fundamental objects, called strings, that propagate in some target space. These strings can either be open or closed, having the topology of a line or a circle respectively. Just as a particle traces out a worldline, a string traces out a two-dimensional worldsheet Σ. We study the embedding of the two-dimensional surfaces Σ into a D-dimensional ambient spacetime M with coordinates X^μ, where $\mu = 0, 1, 2, ..., D-1$.

$$X : \Sigma \to M. \tag{3}$$

We choose a parameterisation (σ, τ) and describe the embedding by the functions $X^\mu(\sigma, \tau)$.

The distinguishing feature of open strings as opposed to closed strings is the existence of two end points. The dynamics of an open string are determined by the action S but we must also know the boundary conditions that tell us how the end points move. There are two kinds of boundary conditions:

- **Neumann**: $\partial_\sigma X^\mu = 0$ at $\sigma = 0, \pi$.

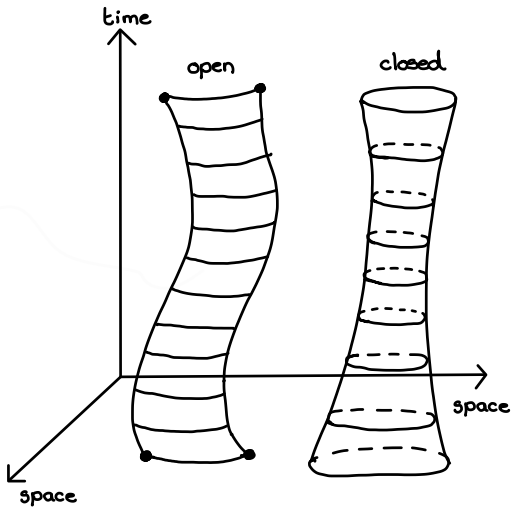


Figure 1: Worldsheet diagrams of open and closed strings.

- **Dirichlet**: $\delta X^\mu = 0$ at $\sigma = 0, \pi$.

In the Neumann boundary condition case the ends of the string move freely, whereas in the Dirichlet boundary condition case the end points of the string at fixed at some constant position in space.

If we have Dirichlet boundary conditions for some coordinates and Neumann boundary conditions for the others

$$\partial_\sigma X^a = 0 \text{ for } a = 0, ..., p \tag{4}$$

$$X^I = c^I \text{ for } I = p+1, ..., D-1 \tag{5}$$

then the end points of the string are fixed on a $(p+1)$-dimensional hypersurface. This hypersurface is called a Dp-brane. In perturbative string theory, a D-brane is a non-perturbative object and instead is a feature of the background. Gauge fields localised on D-branes are fields in a localised super Yang-Mills theory with unitary gauge group.

If we view the amplitude as a function of the worldsheet of the string, string theory is a two-dimensional field theory where the amplitude of a given worldsheet configuration is given by e^{iS} where S is the action of the corresponding two-dimensional field theory. The action S

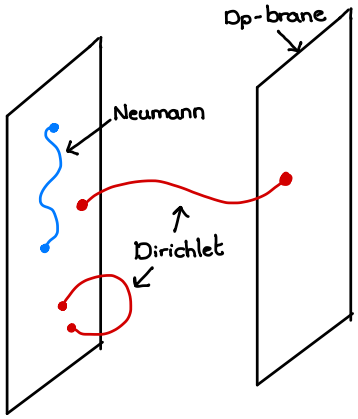


Figure 2: Diagram of two Dp-branes with open strings living along the brane, satisfying Neumann boundary conditions, as well as open strings with both end attached to one or two branes, satisfying Dirichlet boundary conditions.

is proportional to the area of the worldsheet

$$S = -T \int dA, \tag{6}$$

where dA is the area element and T is a constant called the string tension, with units of energy per unit length.

In this two-dimensional field theory, the fields ϕ_i are scalar fields whose values correspond to the values of the coordinates of the points of the worldsheet in its embedding into spacetime. One can enhance the worldsheet field theory from a classical theory to a quantum theory. Excitations of the scalar fields correspond to oscillations of the string. The most remarkable feature of string theory is that it is a theory of quantum gravity because one of the excitations of the quantum relativistic string gives rise to a massless spin-2 particle, which one can identify as the graviton.

The action S is invariant with respect to Weyl rescalings of the metric $g_{\alpha\beta}(\sigma) \to e^{\lambda(\sigma)} g_{\alpha\beta}(\sigma)$ and reparametrisations of the worldsheet $\sigma^\alpha = \tilde{\sigma}^\alpha(\sigma)$ where $\alpha, \beta = 0, 1$ and $(\sigma^0, \sigma^1) \equiv (\tau, \sigma)$. There is a special class of reparametrisations under which the metric is rescaled by a scalar function of σ. Such transformations are called conformal transformations.

These transformations of the metric can be undone by a Weyl transformation, which does not affect $X^\mu(\sigma)$. Therefore, there is a combined symmetry of $S(X, g)$ that fixes the $g_{\alpha\beta}$ and is a coordinate transformation acting on X^μ. If the action $S(\phi_i)$, where ϕ_i denotes the fields, is invariant under such coordinate transformations then the field theory is said to be conformally invariant. Thus string theory can be associated with a two-dimensional conformal field theory.

The original string theories are called bosonic string theories since they only contain a bosonic spectrum. In this case, conformal symmetry of the worldsheet requires that spacetime is 26-dimensional. In this case if we look at the ground state of a single free string we find a negative mass squared. Such particles are called tachyons. The masses of particles in perturbation theory give the second derivative of the background potential energy at the point about which the perturbation series is developed, therefore the existence of a tachyon indicates that the perturbation theory is set up at a point of unstable vacuum.

At the next level up in the spectrum we have massless states. According to Wigner's classification [58], massless particles in d-dimensional spacetime are associated with finite-dimensional irreducible representations (irreps) of the group $SO(d-2)$. We can therefore decompose the massless states, which are irreps of $SO(24)$, into a symmetric-traceless irrep, an anti-symmetric irrep and a singlet. The symmetric-traceless states are the propagating modes of a massless spin 2 particle, i.e. the graviton. Each massless particle of the string is therefore associated to a spacetime field in 26-dimensions

$$G_{\mu\nu}, \qquad B_{\mu\nu}, \qquad \phi, \tag{7}$$

where $G_{\mu\nu}$ is the spacetime metric, $B_{\mu\nu}$ is a 2-form gauge potential called the B-field and ϕ is a scalar field called the dilaton. Beyond this states have masses on the order of the string scale and so their dynamics are irrelevant at low energies and can be neglected in the low energy action.

The quantum bosonic string exhibits many interesting physical phenomena, but cannot be probed beyond tree level due to the presence of the tachyon. From now on we will consider a set of alternative models, collectively known as the superstring which are free of tachyons. For superstring theory, the local symmetries on the worldsheet are enhanced from the ordinary conformal group to the $\mathcal{N} = 1$ superconformal group with new (fermionic) generators. This new symmetry group reduces the appropriate spacetime dimension of the theory from 26 to

10.

There are two types of supersymmetry: worldsheet supersymmetry and spacetime supersymmetry. Worldsheet supersymmetry is the symmetry of the worldsheet action S under the worldsheet supersymmetry transformations. On the other hand, spacetime supersymmetry is a symmetry between tensor fields (bosons) and spinor fields (fermions). As a well-defined supersymmetric transformation requires an invariant spinor, the spacetime supersymmetry directly relates to the holonomy of the spacetime.

Closed strings are periodic in their spatial direction, and therefore the fields living on the worldsheet must also respect this periodicity. For the fermionic fields, this condition plus Lorentz invariance allows two possible boundary conditions; periodic $\psi^\mu(\sigma^0 + 2\pi, \sigma^1) = \psi^\mu(\sigma^0, \sigma^1)$ or antiperiodic $\psi^\mu(\sigma^0 + 2\pi, \sigma^1) = -\psi^\mu(\sigma^0, \sigma^1)$, and similarly for the antiholomorphic fields. This can be summarised clearly by introducing the cylindrical coordinate $w = \sigma^0 + i\sigma^1$:

$$\psi^\mu(w + 2\pi) = e^{2\pi i\nu}\psi^\mu(w) \tag{8}$$

$$\tilde{\psi}^\mu(\bar{w} + 2\pi) = e^{-2\pi i\tilde{\nu}}\tilde{\psi}^\mu(\bar{w}), \tag{9}$$

where ν and $\tilde{\nu}$ take the values 0 and 1/2. The periodic case where $\nu = 0$ is called the Ramond (R) boundary condition and the antiperiodic case where $\nu = 1/2$ is called the Neveu-Schwarz (NS) boundary condition. Thus, there are four possible ways of putting fermions on the closed string, giving rise to four sectors in the Hilbert space, the NS-NS, NS-R, R-NS, R-R sectors.

2.3 Compactification

For superstring theories, one can consider the background spacetime as having the form

$$\mathbb{R}^{1,k-1} \times X_{10-k}, \tag{10}$$

where X_{10-k} is a compact oriented Riemannian manifold, called the internal space. This means we have one time dimension, $k-1$ extended spatial dimensions and $10-k$ curled up spatial dimensions.

As we've mentioned, spacetime supersymmetry requires an invariant spinor and this is directly related to the holonomy of the spacetime. In the compactification scenario, the

compactified background preserves some of the supersymmetry if there exists a covariantly constant spinor on X_{10-k}; that is, there exists some spinor ξ such that

$$\nabla_I \xi = 0, \tag{11}$$

where ∇_I is the covariant derivative and I is an internal space vector index. The generic holonomy group for a Riemannian d-dimensional manifold is $SO(d)$. Special choices of the internal space will lead to "reduced holonomy", where the holonomy group is some subgroup of $SO(d)$. See the discussion on Berger's classification in Section 3.1 for more details. A covariantly constant spinor on X_d exists if the minimal spinor representation of d-dimensions contains a singlet in the decomposition under the holonomy group.

Below we list the relevant manifolds, enumerated by their dimensionality, which preserve supersymmetry. In Table 1, we present the amount of supersymmetry preserved in the five superstring theories for various compactifications. This shows how we can construct theories with $4, 8, 16, 32$ supercharges in four-dimensions corresponding to $\mathcal{N} = 1, 2, 4, 8$ respectively.

- $d = 1$ ($k = 9$): All 1d manifolds have trivial holonomy.
- $d = 2$ ($k = 8$):The holonomy group is $SO(2) \simeq U(1)$, and only the trivial subgroup gives rise to a covariantly constant spinor. The only X_2 with trivial holonomy is T^2.
- $d = 3$ ($k = 7$): The holonomy group is $SO(3) \simeq SU(2)$ and again the only relevant subgroup is the trivial subgroup. The only X_3 with trivial holonomy is T^3.
- $d = 4$ ($k = 6$): The holonomy group is $SO(4) \simeq SU(2) \times SU(2)$. Again T^4 has trivial holonomy. A manifold with $SU(2) \simeq Sp(1)$ holonomy is known as a K3 surface. Thus, the two supersymmetry preserving options are T^4 and K3.
- $d = 5$ ($k = 5$): The holonomy group is $SO(5)$. The analysis is identical to $d = 4$, where we simply tensor with a circle. Thus the two options are $K3 \times S^1$ (or some twisted product) and T^5.
- $d = 6$ ($k = 4$): The holonomy group is $SO(6) \simeq SU(4)$. The obvious examples are T^6 and $K^3 \times T^2$. $SU(3) \subset SO(6)$ is another subgroup. A manifold with $SU(N)$ holonomy is known as a Calabi-Yau N-fold. The options are therefore CY_3, T^6 and $K3 \times T^2$ (or some twisted product).

- $d = 7$ ($k = 3$): The holonomy group is $SO(7)$. The obvious examples are $T^7, K3 \times T^3$ and $CY_3 \times S^1$. There is an additional manifold, the so-called G_2-manifold, whose holonomy group is $G_2 \subset SO(7)$.

- $d = 8$ ($k = 2$): The holonomy group is $SO(8)$. In this case, there are two reduced holonomy subgroups: $SU(4)$, which correspond to CY_4 and Spin(7), which leads to Spin(7)-manifolds. Additionally, we have the obvious product manifolds such as $G_2 \times S^1$ and $CY_3 \times T^2$.

k	X_{10-k}	fraction of supercharges	Type IIA	Type IIB	Heterotic / Type I
6	$K3$	1/2	$\mathcal{N}_6 = (1,1)$	$\mathcal{N}_6 = (2,0)$	$\mathcal{N}_6 = (1,0)$
5	$K3 \times S^1$	1/2	$\mathcal{N}_5 = 2$	$\mathcal{N}_5 = 2$	$\mathcal{N}_5 = 1$
4	$K3 \times T^2$	1/2	$\mathcal{N}_4 = 4$	$\mathcal{N}_4 = 4$	$\mathcal{N}_4 = 2$
	CY_3	1/4	$\mathcal{N}_4 = 2$	$\mathcal{N}_4 = 2$	$\mathcal{N}_4 = 1$
3	$CY_3 \times S^1$	1/4	$\mathcal{N}_3 = 4$	$\mathcal{N}_3 = 4$	$\mathcal{N}_3 = 2$
	G_2	1/8	$\mathcal{N}_3 = 2$	$\mathcal{N}_3 = 2$	$\mathcal{N}_3 = 1$

Table 1: Reduced supersymmetry in various superstring compactifications.

For a more thorough introduction to string compactification, we refer the reader to the great sets of lecture notes [59, 60].

Wrapped Branes

Branes are higher dimensional objects and so can wrap internal dimensions or extend across external dimensions. The case where a brane extends across all $k = 4$ external dimensions, while also possibly wrapping internal dimensions, is called the "spacetime-filling" case. Alternatively, branes can extend over only some (< 4) of the external dimensions, as well as possibly some of the internal dimensions, or it can be entirely internal, existing at a point in external spacetime.

One is interested in brane configurations that do not break Poincaré symmetry and preserve supersymmetry of the effective field theory in the external spacetime. Note that gauge fields can take on non-trivial background expectation values in the internal space without violating Lorentz invariance (but not in the external spacetime). In order to preserve supersymmetry in the external spacetime, the expectation values of gauge fields localised on

D-branes must satisfy certain conditions. In the case where the brane wraps an even-dimensional submanifold of the internal space (holomorphic cycle), the background gauge fields must satisfy the Hermitian Yang-Mills equations. On the other hand, in the case where the brane wraps an odd-dimensional submanifold (special Lagrangian cycle) of the internal space, the background gauge fields must vanish.

Kaluza-Klein Reduction

Consider the scenario of $d+1$-dimensional spacetime with one compact dimension $X \times S^1$. Let x^μ for $\mu \in \{0, ..., d-1\}$ denote the non-compact X coordinates and x^d the compact S^1 coordinate with periodicity condition $x^d \sim x^d + 2\pi R$. The spacetime momentum operator $\hat{\mathbf{p}}_d$ is a generator of translation and therefore we have the identity

$$e^{i2\pi R\hat{\mathbf{p}}_d} \equiv \mathbf{1}. \tag{12}$$

From this we see that the eigenvalues of momentum along the circle are $p_d = n/R$, where $n \in \mathbb{Z}$. We can thus decompose a massless $d+1$-dimensional field ϕ into an infinite set of Fourier modes along the compact dimension

$$\phi(x^0, ..., x^d) = \sum_{n\in\mathbb{Z}} \phi_n(x^0, ..., x^{d-1}) e^{\frac{inx^d}{R}}, \tag{13}$$

where the ϕ_n are d-dimensional fields.

The mass formula for ϕ_n is

$$E^2 = \sum_{\mu=0}^{d} p_\mu p^\mu = \frac{n^2}{R^2} + \sum_{\mu=0}^{d-1} p_\mu p^\mu. \tag{14}$$

In other words, the ϕ_n field has developed a non-zero mass $m^2 = n^2/R^2$. These modes are called Kaluza-Klein (KK) modes and arise in more general compactifications $E \to X$[3]. In this case the masses of the fields are inversely proportional to the metric volume of F, where F denotes the fiber over X. In model building these volumes are chosen to be very small so that the masses of the KK modes are too small to be detectable by current experiments.

[3]A fiber bundle is a space E that locally resembles a product space $B \times F$ but globally can have a different topological structure. It is defined by a continuous surjective map $\pi : E \to B$, that in small regions of E behaves just like a projection from regions of $B \times F$ to B. The map π is called the projection, E the total space, B the base space and F the fiber.

Since in general relativity, the size and shape of F is dynamical, generically the effective field theories contain fields, called moduli, which parameterise the geometry of F. For example, the dilaton which encodes the total volume of F. Since these moduli fields are not observed in experiment, naive KK models are ruled out by experiment. There are however ways for the moduli to obtain masses and for the KK models to become viable. This is called moduli stabilisation.

2.4 Five Superstring Theories

There are five known superstring theories called type I, type IIA, type IIB, heterotic $SO(32)$ and heterotic $E_8 \times E_8$. These theories contain different combinations of open and closed strings and branes, have different amounts of supersymmetry, different gauge symmetry and different matter content. The content of each of these five theories is briefly summarised below.

Type I

Type I theory is the only superstring theory which contains both open and closed strings. This theory also contains D1-, D5- and D9-branes. The low energy effective theory has $\mathcal{N} = 1$ supersymmetry and $\mathrm{Spin}(32)/\mathbb{Z}_2$ gauge group. The bosonic fields from the closed sector are the NS-NS graviton G, the dilaton ϕ, the R-R 2-form B, and from the open sector we also have NS gauge bosons A.

Type II

Type II theories are theories of closed strings as well as D-branes (D(-1)-, D1-, D3-, D5-, D7- and D9-branes in IIB and D0-, D2-, D4-, D6-, D8-branes in IIA). The worldsheet theory is split into left and right moving sectors. In type IIA they have opposite chirality ($\mathcal{N} = (1,1)$) and in type IIB they have the same chirality ($\mathcal{N} = (2,0)$). The left-moving and right-moving sectors contribute equally and so both theories have $\mathcal{N} = 2$ supersymmetry in their low energy effective field theory.

The bosonic fields for type IIA string theory are the graviton G, the NS-NS 2-form B, the dilaton ϕ, and the R-R 1-form C_1 and the 3-form C_3. For type IIB the bosonic fields are the graviton G, the NS-NS 2-form B, the dilaton ϕ, the R-R 0-form a, 2-form C_2 and 4-form C_4 (with self-dual field strength).

We can consider type II string compactifications on the circle, i.e. $\mathbb{R}^{1,8} \times S^1$. Under T-duality, type IIA theory compactified on the circle with radius R is mapped to type IIB theory compactified on the circle with radius $1/R$. Mirror symmetry is a generalisation of T-duality where the content of type IIA string theory on one Calabi-Yau threefold is exactly the same content of type IIB string theory compactified on the mirror manifold. Roughly speaking, mirror symmetry says that for every Calabi-Yau threefold X there exists a mirror Calabi-Yau threefold Y such that

$$h^{1,1}(X) = h^{2,1}(Y) \qquad h^{2,1}(X) = h^{1,1}(Y). \tag{15}$$

Heterotic

In type II superstring theories, we have ten bosons and ten fermions in both the right and left moving sectors. In heterotic superstring theories we keep the right-handed bosons and fermions but in the left sector we take the bosonic string bosons. Therefore, heterotic superstring theories are a heterosis of the 26-dimensional bosonic and the 10-dimensional superstring theory. The worldsheet theory consists of a left-moving CFT and a right-moving SCFT.

The matter content consists of 10 right-moving bosons, 10 right-moving fermions, 10 left-moving bosons and 32 left-moving fermions, which cancel the Virasoro anomaly by bosonisation. For each of these fermions we need to choose the boundary conditions, i.e. antiperiodic (NS) or periodic (R). To achieve an anomaly-free theory with $\mathcal{N} = 1$ spacetime supersymmetry, there are just two options:

- $\mathbf{SO(32)}$ All of the fermions are in the same sector (NS or R). In this case we can mix these fermions by an $SO(32)$ transformation without changing the theory and thus $SO(32)$ is a symmetry of the theory.
- $\mathbf{E_8 \times E_8}$: Half of the fermions are in the same sector and the other half are in the same sector. Therefore there are four possibilities: NS-NS, NS-R, R-NS, R-R. Since 16 of the fermions live in the same sector, we can mix them by an $SO(16)$ transformation, and similarly for the other 16. The $SO(16)$ representations of the massless excitations are:
  - NS-NS: $(120, 1) + (1, 120)$
  - NS-R: $(1, 128)$
  - R-NS: $(128, 1)$

So the massless states live in the $120 + 128$ representation of $SO(16)$, but this is not the adjoint representation of $SO(16)$ and therefore $SO(16) \times SO(16)$ is not the right gauge group. Instead the gauge group is $E_8 \times E_8$ since $SO(16)$ is a subgroup of E_8 and E_8 is $120 + 128 = 248$-dimensional.

Heterotic superstring theories are theories of only closed strings and no D-branes. For both heterotic $E_8 \times E_8$ and $SO(32)$, the bosonic field content consists of the graviton G, the 2-form B, the dilaton ϕ and the gauge bosons A.

	SUSY	bosonic content	gauge group
Type IIA	$\mathcal{N} = (1,1)$ $N_{SUSY} = 32$	NS-NS:ϕ, B, G RR:C_1, C_3	-
Type IIB	$\mathcal{N} = (2,0)$ $N_{SUSY} = 32$	NS-NS:ϕ, B, G RR:a, C_2, C_4	-
Heterotic $E_8 \times E_8$	$\mathcal{N} = (1,0)$ $N_{SUSY} = 16$	NS-NS:ϕ, B, G RR:$A \in \mathfrak{e}_8 \oplus \mathfrak{e}_8$	$E_8 \times E_8$
Heterotic $SO(32)$	$\mathcal{N} = (1,0)$ $N_{SUSY} = 16$	NS-NS:ϕ, B, G RR:$A \in \mathfrak{so}(32)$	$\mathrm{Spin}(32)/\mathbb{Z}_2$
Type I	$\mathcal{N} = (1,0)$ $N_{SUSY} = 16$	NS-NS:ϕ, G RR:B NS+:$A \in \mathfrak{so}(32)$	$\mathrm{Spin}(32)/\mathbb{Z}_2$

Table 2: Summary of the supersymmetry, bosonic content and gauge group of five string theories.

Dualities

The five superstring theories that we have just described: type I, type IIA, type IIB, heterotic $SO(32)$ and heterotic $E_8 \times E_8$, are all related to each other via various dualities. Observables can be mapped between two theories that are dual to one another to yield equivalent predictions.

String dualities often link quantities that appear to be separate. For instance, T-duality, which we saw connects the two type II theories, links large and small distances. There also exists a T-duality connecting the two heterotic theories. Another duality is S-duality which

links strong g_s and weak $1/g_s$ coupling. A theory with weak coupling can be understood by means of perturbation theory but a theory with strong coupling cannot. S-duality therefore allows us to push past the perturbation theory and learn more about the underlying non-perturbative theory.

Type I theory is S-dual to heterotic $SO(32)$ theory. Furthermore, type IIA theory at the strong coupling limit $g_s \to \infty$ behaves like an eleven-dimensional theory, where the dilaton plays the role of the eleventh dimension. In particular, it behaves like a theory on $M_{10} \times S^1$ where the radius of the circle S^1 goes to infinity. The full eleven-dimensional theory is called M-theory. The radius limit is a T-duality, and the strong coupling limit is an S-duality.

M-theory has no dimensionless coupling constant. In the duality with type IIA, both the gauge coupling g_s and the string tension T of type IIA string theory become part of the geometry. Moreover, the type IIA graviton, dilaton and R-R 1-form are all lifted to be part of the metric in eleven-dimensions. By KK reduction on a circle, we can connect the M-theory fields with the fields in IIA, as shown in Table 3.

M-Theory		IIA	
G_{MN}	$\to$	G_{MN}	graviton
$G_{M,10}$	$\to$	C_1	R-R 1-form
$G_{10,10}$	$\to$	ϕ	dilaton
C_{MNP}	$\to$	C_3	R-R 3-form
C_{MN10}	$\to$	B_2	NS-NS 2-form

Table 3: M-theory to IIA by KK reduction on S^1.

We have already seen that M-theory is T-dual to type IIA and type IIA is T-dual to type IIB. Combining these dualities we see that we can get type IIB by first compactifying M-theory on a torus $S^1_A \times S^1_B$ to get type IIA on a circle and then T-dualising that circle. In the limit where the radius of one of the S^1s goes to infinity, the area of the torus goes to zero. In this way, one can think of type IIB as a 12d theory where two of the directions look like a zero-area torus. F-theory can be viewed as a non-perturbative compactification of type IIB. More generally, this torus can be taken to be an elliptic curve and this may vary over the base space as an elliptic fibration.

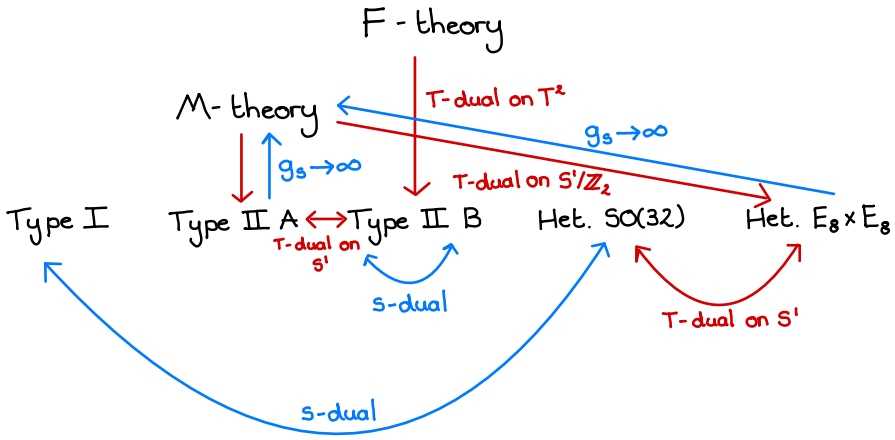


Figure 3: Web of string dualities.

2.5 $E_8 \times E_8$ Heterotic Model Building

One of the standard approaches to heterotic string model building is to compactify on a Calabi-Yau threefold X that admits a principal $G \subset E_8 \times E_8$ vector bundle V, where V is a direct sum of an observable sector bundle V^1, whose structure group G^1 is embedded in the first E_8 factor, and a hidden sector bundle V^2, with structure group G^2 embedded in the second E_8.

In order for the bundle V^1 to preserve $\mathcal{N} = 1$ supersymmetry, its field strength F must satisfy the hermitian Yang-Mills equations:

$$F_{ij} = F_{\bar{i}\bar{j}} = 0, \qquad g^{i\bar{j}} F_{i\bar{j}} \propto \mathrm{Id}, \tag{16}$$

where i, j label the holomorphic coordinates on X, Id is the identity element of G^1 and the constant of of proportionality $\mu(V^1)$ is a real number known as the slope and is determined by the choice of V^1. The second equation in (16) can be rephrased in terms of the hermitian fiber metric H:

$$g^{i\bar{j}} F_{i\bar{j}} = -g^{i\bar{j}} \partial_{\bar{j}} A_i = -g^{i\bar{j}} \partial_{\bar{j}} (H^{-1}(\partial_i H)) \propto \mathrm{Id}. \tag{17}$$

A metric that satisfies this equation is said to be Hermite-Einstein. Therefore a solution to the Yang-Mills equations is equivalent to the vector bundle V^1 being holomorphic and admitting a hermitian metric H on its fibers which is Hermite-Einstein. Whether there exists a Hermite-Einstein metric H on V^1 depends on the so-called stability of the bundle [61, 62]. This can often be checked by somewhat laborious algebraic calculations, though the guarantee of existence is not constructive – even if a given bundle is stable, it is often impossible to find

an explicit expression for the corresponding Hermite-Einstein metric.

Assuming that X admits a principal G-bundle with $G \subset E_8$, the gauge group in four-dimensions is given by the commutant of G in E_8. Using the Dynkin diagram of E_8, as shown in Figure 4, we can get the E_7 and E_6 subgroups by contracting nodes.

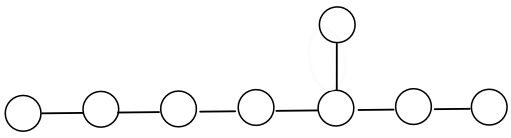


Figure 4: E_8 dynkin diagram.

Standard Embedding

The scenario when V^1 is chosen to be the holomorphic tangent bundle TX of the Calabi-Yau threefold X with structure group $G^1 = SU(3)$, is called the standard embedding. In this case, E_8 is broken to E_6 since $SU(3)$ is the commutant of E_6 within E_8. Note that:

$$SU(3) \times SU(2) \times U(1) \subset SU(5) \subset SO(10) \subset E_6. \tag{18}$$

The breaking of the E_8 gauge group to the Standard Model gauge group happens in two stages; firstly E_8 is broken to the GUT group E_6 and then E_6 is then broken down further to the Standard Model gauge group.

When E_8 is broken down as

$$E_8 \to E_6 \oplus SU(3), \tag{19}$$

the adjoint representation of the E_8, namely **248**, decomposes as

$$\mathbf{248} \to \bigoplus_{\mathbf{r},\mathbf{R}} (\mathbf{r}, \mathbf{R}) = (\mathbf{8}, \mathbf{1}) \oplus (\mathbf{1}, \mathbf{78}) \oplus (\mathbf{3}, \mathbf{27}) \oplus (\overline{\mathbf{3}}, \overline{\mathbf{27}}), \tag{20}$$

where we use $\mathbf{r}$ to denote a representation of $SU(3)$ and $\mathbf{R}$, that of E_6. From this, we observe that the low-energy theory can contain matter transforming as the $\mathbf{1}$, $\mathbf{27}$ or $\overline{\mathbf{27}}$ of E_6,

corresponding to bundle moduli, families and anti-families. By KK reduction one finds that these matter fields come from harmonic TX bundle-valued $(0,1)$-forms $\psi_{\mathbf{r}}$ [10, 11]. In other words, the number of **27** families present in four-dimensions is counted by the number of harmonic TX-valued $(0,1)$-forms on X. The number of these harmonic $(0,1)$-forms is equal to the dimension of $H^1(X,TX)$. Similarly, the number of $\overline{\mathbf{27}}$ anti-families is given by the dimension of $H^1(X,T^*M)$. Thus we have

$$\text{generations of particles} \sim H^1(X,TX) \tag{21}$$

$$\text{anti-generations of particles} \sim H^1(X,T^*X) \tag{22}$$

The effective theory is determined by a superpotential and a Kähler potential. These objects are fixed by the geometry of the compactification in the following way. Let $\psi^I_{\mathbf{r}}$ be a basis for $H^1(X,TX)$ which is not necessarily harmonic. If there is a singlet in the product $\mathbf{r}\times\mathbf{r}'\times\mathbf{r}''$, there can be a holomorphic Yukawa coupling of the form

$$\lambda_{IJK}(\mathbf{r},\mathbf{r}',\mathbf{r}'') = \int_X \Omega\wedge \mathrm{Tr}(\psi^I_{\mathbf{r}}\wedge\psi^J_{\mathbf{r}'}\wedge\psi^K_{\mathbf{r}''}), \tag{23}$$

where Ω is the holomorphic $(3,0)$-form on X, and the trace indicates a projection to the $SU(3)$ singlet. Using the above decomposition (20), we denote the four-dimensional chiral superfields associated via KK reduction to each $\psi^I_{\mathbf{r}}$ by $C^I_{\mathbf{R}}$. The superpotential for these chiral superfields is then given by

$$W = \lambda_{IJK}(\mathbf{R},\mathbf{R}',\mathbf{R}'')C^I_{\mathbf{R}}C^J_{\mathbf{R}'}C^K_{\mathbf{R}''}, \tag{24}$$

where we have relabelled the Yukawa couplings by the four-dimensional gauge group, i.e. their E_6 representations. Given that a singlet appears in $\mathbf{8}^3$, $\mathbf{8}\cdot\mathbf{3}\cdot\overline{\mathbf{3}}$, $\mathbf{3}^3$ and their conjugates, the possible types of Yukawa couplings are $\mathbf{1}^3$, $\mathbf{1}\cdot\mathbf{27}\cdot\overline{\mathbf{27}}$, $\mathbf{27}^3$ and $\overline{\mathbf{27}}^3$.

The above are often known as holomorphic Yukawa couplings as they are quasi-topological in the sense that λ_{IJK} can be computed using representatives of $H^1(X,TX)$ which are not harmonic (this follows straightforwardly from $d\Omega=0$). However, the physical Yukawa couplings depend on the normalisation of the kinetic terms for the chiral superfields. This normalisation is fixed by the matter-field Kähler metric, given by

$$\mathcal{G}_{IJ} = \int_X \bar{\star}_{TX}\psi^I_{\mathbf{r}}\wedge\psi^J_{\mathbf{r}}, \tag{25}$$

which is simply the inner product between harmonic representatives of $H^1(X,TX)$. Due to the need for harmonic forms and the presence of the Hodge star on bundle-valued forms $\star_{TX}$, this depends on knowledge of the Calabi–Yau metric on X, the Hermite-Einstein metric on the fibres of TX and the zero modes of the Dolbeault Laplacian on $(0,1)$-forms valued in TX.

Non-Standard Embedding

For the standard embedding, many quantities of interest can be computed using the techniques of special geometry without needing an explicit metric on X. Unfortunately, it is difficult to find acceptable minimal supersymmetric Standard Model (MSSM) like physics in these simple models, so one is forced to consider more general vector bundles. By taking V^1 not to be TX but, for example, a stable $SU(4)$ or $SU(5)$ bundle, one could obtain the more interesting commutant $SU(10)$ or $SU(5)$ GUTs. This is known as non-standard embedding. Just as in the standard embedding case, the low energy particle spectrum is determined through the cohomology of V^1 and $(V^1)^\vee$ according to the decomposition of the **248** adjoint representation of E_8.

Various non-standard embedding models have been proposed over the years, including the spectral cover construction over elliptically fibered Calabi-Yau threefolds [63–66], monad bundles [67, 68], extension bundles [69, 70] and sums of line bundles [71–73]. In Section 7 we will consider the line bundle case and attempt to compute the particle spectrum in this scenario.

In this section we have discussed the role of Calabi-Yau manifolds in compactification in string theory without defining them. In the next section we will define explicitly what it means for a complex manifold to be labelled Calabi-Yau, discuss properties of these manifolds and ways to construct them. We will also discuss G_2 manifolds which are the analogous compactificaiton space for eleven-dimensional M-theory.

3 Maths Background

In this section we review parts of algebraic and differential geometry with the goal of understanding the nature of Calabi-Yau and G_2 manifolds. For the sake of space we assume that the reader is familiar with the basics, such as complex manifolds, fibre bundles, etc. For a more thorough introduction to the topics discussed in this section we refer the reader to the following references: for an overview of complex manifolds and Kähler geometry see the lecture notes by Joyce [74], for a introduction to Calabi-Yau manifolds (suitable for physicists) see [75] and for an overview of G_2 manifolds we refer the reader to [76].

3.1 Riemannian Holonomy Groups

Let ∇^E be a connection on a vector bundle $E \to X$ and let $\gamma : [0,1] \to X$ be a smooth curve with $\gamma(0) = x$ and $\gamma(1) = y$. Then $\gamma^*(\nabla^E)$ is a connection on $\gamma^*(E) \to [0,1]$. For each $e \in E_x$ there is a unique section s of $\gamma^*(E)$ with $s(0) = e$ and $\gamma^*(\nabla^E) \equiv 0$. Define $P_\gamma(e) = s(1)$. Then $P_\gamma : E_x \to E_y$ is the parallel transport map.

Definition 3.1 *Let ∇^E be a connection on a vector bundle $E \to X$. Then the holonomy group $Hol_x(\nabla^E)$ of ∇^E is defined as*

$$Hol_x(\nabla^E) := \{P_\gamma : \gamma \text{ is a piecewise-smooth loop based at } x\} \subset GL(E_x), \tag{26}$$

where P_γ is the action on the vector space E_x induced by the parallel transport along γ.

The classification of Riemannian holonomy groups was given by Berger in 1955 [12].

Theorem 3.1 *Let M be a simply-connected manifold of dimension n and let g be an irreducible and non-symmetric Riemannian metric on M. Then exactly one of the following cases holds:*

- *$Hol(g) = SO(n)$.*
- *$n = 2m$ with $m \geq 2$, and $Hol(g) = U(M) \subset SO(2m)$. A manifold with $Hol(g) \subseteq U(m)$ is called a Kähler manifold.*
- *$n = 2m$ with $m \geq 2$, and $Hol(g) = SU(m) \subset SO(2m)$. A manifold with $Hol(g) \subseteq SU(m)$ is called a Calabi-Yau manifold.*
- *$n = 4m$ with $m \geq 2$ and $Hol(g) = Sp(m) \subset SO(4m)$. A manifold with $Hol(g) \subseteq Sp(m)$ is called a hyperkähler manifold.*

- $n = 7$ *and* $Hol(g) = G_2 \subset SO(7)$.
- $n = 8$ *and* $Hol(g) = Spin(7) \subset SO(8)$.

In string theory the interest in holonomy groups comes from supersymmetry and the existence of invariant spinors - the requirement for preserving n supersymmetries is the existence of n invariant spinors [77]. Let M be an orientable, connected, simply-connected spin n-manifold for $n \geq 3$ and let g be an irreducible Riemannian metric on M. Define N to be the dimension of parallel spinors on M and if n is even, define $N_\pm$ to be the dimensions of the spaces of parallel spinors in $C^\infty(S_\pm)$, so that $N = N_+ + N_-$. Suppose $N \geq 1$. Then, fixing the orientation, exactly one of the following holds:

- $n = 4m$ for $m \geq 1$ and $\mathrm{Hol}(g) = SU(2m)$, with $N_+ = 2$ and $N_- = 0$.
- $n = 4m$ for $m \geq 2$ and $\mathrm{Hol}(g) = Sp(m)$, with $N_+ = m + 1$ and $N_- = 0$.
- $n = 4m + 2$ for $m \geq 1$ and $\mathrm{Hol}(g) = SU(2m + 1)$, with $N_+ = 1$ and $N_- = 1$.
- $n = 7$ and $\mathrm{Hol}(g) = G_2$, with $N = 1$.
- $n = 8$ and $\mathrm{Hol}(g) = \mathrm{Spin}(7)$, with $N_+ = 1$ and $N_- = 0$.

With this we see that Calabi-Yau threefolds and G_2 manifolds are the spaces on which heterotic/type I superstring and M-theory compactify preserving $\mathcal{N} = 1$ supersymmetry in four-dimensions.

3.2 Hodge Theory

Let (X, g) be a compact, oriented Riemannian n-manifold. Then the Hodge star operator $*$ acts on k-forms:

$$* : C^\infty(\Lambda^k T^*X) \to C^\infty(\Lambda^{n-k} T^*X). \tag{27}$$

It satisfies $*^2 = (-1)^{k(n-k)}$, so $*^{-1} = \pm *$. We define the codifferential

$$d^* : C^\infty(\Lambda^k T^*X) \to C^\infty(\Lambda^{k-1} T^*X), \tag{28}$$

by $d^* = (-1)^k *^{-1} d*$, and the Laplacian on k-forms $\Delta_d = dd^* + d^*d$.

If $\alpha, \beta \in C^\infty(\Lambda^k T^*X)$ we define the L_2 inner product as

$$\langle \alpha, \beta \rangle_{L^2} = \int_X (\alpha, \beta) dV_g, \tag{29}$$

where (α, β) is the pointwise inner product of k-forms using g, and dV_g the volume form of g. The Hodge star is defined so that if α, β are k-forms then $\alpha \wedge *\beta = (\alpha, \beta)dV_g$. Thus

$$\langle \alpha, \beta \rangle_{L^2} = \int_X \alpha \wedge *\beta. \tag{30}$$

We write d_k, d_k^* for d, d^* acting on k-forms, and $\mathcal{H}^k$ for $\mathrm{Ker}\Delta_d$ on k-forms, i.e. $\mathcal{H}^k(X) = \{\alpha \in C^\infty(\Lambda^k T^*X) | \Delta_d \alpha = 0\}$.

Theorem 3.2 (Hodge Decomposition [78]) *Let (X, g) be a compact, oriented Riemannian manifold. Then*

$$C^\infty(\Lambda^k T^*M) = \mathcal{H}^k \oplus Im(d_{k-1}) \oplus Im(d_{k+1}^*). \tag{31}$$

In other words, any k-form ω on X can be uniquely written as

$$\omega = d_{k-1}\beta + d_{k+1}^*\gamma + \omega', \tag{32}$$

where β is a $(k-1)$-form, γ is a $(k+1)$-form and ω' is a harmonic k-form ($\omega' \in \mathrm{Ker}\Delta_d$). In particular, if ω is closed then it is not hard to show that γ vanishes and hence we have

$$\omega = d_{k-1}\beta + \omega'. \tag{33}$$

It follows that

$$H_{dR}^k(X, \mathbb{R}) = \mathrm{Ker}d_k / \mathrm{Im}d_{k-1} \tag{34}$$

$$= (\mathcal{H}^k \oplus \mathrm{Im}(d_{k-1}))/\mathrm{Im}d_{k-1} \cong \mathcal{H}^k \tag{35}$$

Theorem 3.3 *Every de Rham cohomology class on X contains a unique harmonic representative.*

The Hodge star gives an isomorphism $* : \mathcal{H}^k \to \mathcal{H}^{n-k}$, thus $H_{dR}^k(X, \mathbb{R}) \cong H_{dR}^{n-k}(X, \mathbb{R})$. This is a form of Poincaré duality.

3.3 Kähler Manifolds

Let (X, J) be a complex manifold and g be a Riemannian metric on X. We call g Hermitian if $g_{ab} = J_a^c J_b^d g_{cd}$. That is, for all vector fields v, w we have $g(v, w) = g(Jv, Jw)$. Let g be a Hermitian metric on (X, J) and define a 2-tensor $\omega = \omega_{ab}$ by $\omega_{ab} = J_a^c g_{cb}$. That is, $\omega(v, w) = g(Jv, w)$ for all vector fields v, w. We have $\omega_{ba} = -\omega_{ab}$ and therefore ω is a 2-form which we call the Hermitian form of g.

Definition 3.2 *Let g be a Hermitian metric on a complex manifold (X, J), with Hermitian form ω. We call g Kähler if ω is closed, $d\omega = 0$. In which case, we call (X, J, g) a Kähler manifold and ω the Kähler form.*

In local coordinates, the fact that $d\omega = 0$ for a Kähler manifold implies

$$d\omega = (\partial + \bar{\partial}) i g_{i\bar{j}} dz^i \wedge d\bar{z}^{\bar{j}} = 0. \tag{36}$$

This implies that

$$\frac{\partial g_{i\bar{j}}}{\partial z^l} = \frac{\partial g_{l\bar{j}}}{\partial z^i}, \tag{37}$$

and similarly with z and $\bar{z}$ interchanged. From this we see that locally we can express $g_{i\bar{j}}$ as

$$g_{i\bar{j}} = \frac{\partial^2 K}{\partial z^i \partial \bar{z}^{\bar{j}}}. \tag{38}$$

That is, $J = i\partial\bar{\partial}K$, where K is a locally defined function in the patch whose local coordinates we are using. K is known as the Kähler potential.

Let (X, J, g) be a Kähler manifold with Kähler form ω. Then ω is a closed real 2-form, so it has a cohomology class $[\omega]$ in the de Rham cohomology $H^2_{dR}(X, \mathbb{R})$. We call $[\omega]$ the Kähler class of g.

Consider (p, q)-forms $C^\infty(\Lambda^{p,q}X)$ on X and define the operators

$$\partial : C^\infty(\Lambda^{p,q}X) \to C^\infty(\Lambda^{p+1,q}X) \tag{39}$$

$$\bar{\partial} : C^\infty(\Lambda^{p,q}X) \to C^\infty(\Lambda^{p,q+1}X) \tag{40}$$

by $d = \partial + \bar{\partial}$, and their adjoints

$$\partial^* : C^\infty(\Lambda^{p,q}X) \to C^\infty(\Lambda^{p-1,q}X) \tag{41}$$

$$\bar{\partial}^* : C^\infty(\Lambda^{p,q}X) \to C^\infty(\Lambda^{p,q-1}X), \tag{42}$$

by $\partial^* = - * \partial *$ and $\bar{\partial}^* = - * \bar{\partial} *$. With these operators we can define the d, ∂ and $\bar{\partial}$-Laplacians by $\Delta_d = dd^* + d^*d$, $\Delta_\partial = \partial\partial^* + \partial^*\partial$ and $\Delta_{\bar{\partial}} = \bar{\partial}\bar{\partial}^* + \bar{\partial}^*\bar{\partial}$. For Kähler manifolds we have $\Delta_\partial = \Delta_{\bar{\partial}} = \frac{1}{2}\Delta_d$.

Let $\mathcal{H}^{p,q}$ denote the kernel of Δ_d acting on (p, q)-forms and $\bar{\partial}_{p,q}, \bar{\partial}^*_{p,q}$ denote $\bar{\partial}, \bar{\partial}^*$ acting on (p, q)-forms. Then we have the following version of the Hodge decomposition theorem:

Theorem 3.4 *Let (X, J, g) be a compact Kähler manifold. Then*

$$C^\infty(\Lambda^{p,q}X) = \mathcal{H}^{p,q} \oplus Im(\bar\partial_{p,q-1}) \oplus Im(\bar\partial^*_{p,q+1}). \tag{43}$$

In other words, any (p,q)-form $\omega^{p,q}$ on X can by written as

$$\omega^{p,q} = \bar\partial\alpha^{p,q-1} + \bar\partial^*\beta^{p,q+1} + \omega'^{p,q}, \tag{44}$$

where ω' is harmonic with respect to the Laplacian Δ_d, and Dolbeault cohomology satisfies

$$H^{p,q}_{\bar\partial}(X) = \mathrm{Ker}\bar\partial_{p,q}/\mathrm{Im}\bar\partial_{p,q-1} \tag{45}$$

$$= (\mathcal{H}^{p,q} \oplus \mathrm{Im}(\bar\partial))/\mathrm{Im}\bar\partial_{p,q-1} \cong \mathcal{H}^{p,q}. \tag{46}$$

If we write $H^{p,q}(X)$ for the subspace of $H^{p+q}(X;\mathbb{C})$ represented by forms in $\mathcal{H}^{p,q}$, then we have

$$H^k(X;\mathbb{C}) = \bigoplus_{p+q=k} H^{p,q}(X), \tag{47}$$

and $H^{p,q}(X) \cong H^{p,q}_{\bar\partial}(X)$. Hence

$$H^k(X;\mathbb{C}) = \bigoplus_{p+q=k} H^{p,q}_{\bar\partial}(X). \tag{48}$$

Observe that complex conjugation takes $\mathcal{H}^{p,q}$ to $\mathcal{H}^{q,p}$. Since $\mathcal{H}^{p,q} \cong H^{p,q}_{\bar\partial}(X)$, this implies that

$$H^{p,q}_{\bar\partial}(X) \cong \overline{H^{q,p}_{\bar\partial}(X)} \tag{49}$$

Furthermore $*$ gives

$$* : \mathcal{H}^{p,q} \to \overline{\mathcal{H}^{n-p,n-q}}, \tag{50}$$

which implies the Poincaré duality style isomorphisms

$$H^{p,q}(X) \cong H^{n-p,n-q}(X)^* \tag{51}$$

$$H^{p,q}_{\bar\partial}(X) \cong H^{n-p,n-q}_{\bar\partial}(X)^* \tag{52}$$

The Betti numbers of X are $b^k(X) = \dim_{\mathbb{C}} H^k_{dR}(X;\mathbb{C})$, and the Hodge numbers of X are $h^{p,q}(X) = \dim_{\mathbb{C}} H^{p,q}_{\bar\partial}(X)$. We have

$$b^k(X) = \sum_{p+q=k} h^{p,q}(X). \tag{53}$$

The Hodge star operator ensures

$$h^{p,q}(X) = h^{n-p,n-q}(X), \tag{54}$$

and using complex conjugation and Kählerity, we also have

$$h^{p,q}(X) = h^{q,p}(X). \tag{55}$$

3.4 Holomorphic Vector Bundles

Let X be a complex manifold of complex dimension n and let V be a rank-r holomorphic vector bundle over X. We denote by $\Omega^{p,q}(V)$ the space of V-valued (p,q)-forms. Explicitly, let $\{E_a\}$ be a holomorphic frame for V, i.e. on each patch of X, the E_a form a basis for $\mathbb{C}^r$ and have holomorphic transition functions valued in $GL(r,\mathbb{C})$. A bundle-valued (p,q)-form α can then be written locally as

$$\alpha = \sum_{a=1}^{r} \alpha^a \otimes_{\mathbb{C}} E_a, \tag{56}$$

where $\alpha^a \in \Omega^{p,q}(X)$ are standard (p,q)-forms on X.

$\bar{\partial}$-Operators and Connections

There is a natural $\bar{\partial}$ operator on V:

$$\bar{\partial}_V : \Omega^{0,0}(V) \to \Omega^{0,1}(V), \tag{57}$$

such that $\alpha \in \Omega^{0,0}(V)$ is holomorphic iff $\bar{\partial}_V \alpha = 0$. It satisfies the Leibnitz rule

$$\bar{\partial}_V(fe) = f \cdot \bar{\partial}_V e + \bar{\partial} f \otimes_{\mathbb{C}} e, \tag{58}$$

for all $e \in \Omega^{0,0}(V)$ and smooth $f : X \to \mathbb{C}$. Given $\bar{\partial}_V$ there are unique extensions

$$\bar{\partial}_V^{p,q} : \Omega^{p,q}(V) \to \Omega^{p,q+1}(V) \tag{59}$$

with $\bar{\partial}_V = \bar{\partial}_V^{0,0}$, such that

$$\bar{\partial}_V^{p,q}(\alpha \otimes_{\mathbb{C}} e) = \bar{\partial}_V e \wedge \alpha + e \otimes_{\mathbb{C}} \bar{\partial}\alpha \tag{60}$$

for $e \in \Omega^{0,0}(V)$ and $\alpha \in \Omega^{p,q}(X)$.

For a holomorphic vector bundle we have $\bar{\partial}_V^{p,q+1} \circ \bar{\partial}_V^{p,q} = 0$ for all p, q. For any $\bar{\partial}_V$, the operator

$$\bar{\partial}_V^{0,1} \circ \bar{\partial}_V : \Omega^{0,0}(V) \to \Omega^{0,2}(V) \tag{61}$$

is of the form $e \mapsto F_V^{0,2} \cdot e$ for unique $F_V^{0,2} \in \Omega^{0,2}(\mathrm{End}(V))$.

A connection on V is defined by a differential operator

$$D : \Omega^0(V) \to \Omega^1(V), \tag{62}$$

such that

$$D(fe) = f \cdot De + e \otimes df \tag{63}$$

for all locally defined holomorphic sections e of E and all locally defined holomorphic functions f on X.

Let E be a complex vector bundle over (X, J), and G a Hermitian metric on the fibres of E. A connection D on V is said to be Hermitian if

$$d(G(s_1, s_2)) = G(Ds_1, s_2) + G(s_1, Ds_2), \tag{64}$$

or equivalently $DG = 0$ for all $s_1, s_2 \in \Omega^0(V)$. There always exists Hermitian connections.

Now let E be a holomorphic vector bundle over X and D a Hermitian connection with respect to the Hermitian fiber metric G. $F_V^{0,2}$ is then $(0,2)$-component of the curvature $F_V = \Omega^2(\mathrm{End}(V))$ and since E is holomorphic $F_V^{0,2} = 0$. Furthermore, as D preserves G,

$$F_V \in \Omega^2(\mathrm{Herm}^-(V)), \tag{65}$$

where $\mathrm{Herm}^-(V) \subset \mathrm{End}(V)$ is the set of anti-Hermitian transformations with respect to G. This implies that the $(2,0)$-component of F_V is conjugate to the $(0,2)$-component and thus is also zero. Therefore, we conclude that F_V is of type $(1,1)$.

Note that compatibility with the holomorphic and Hermitian structures uniquely determines the connection as the Chern connection of G, which is characterised by a local one-form A. With respect to a local holomorphic frame, A is given as

$$A = G^{-1}\partial G, \tag{66}$$

such that $D = d + A$. Therefore,

$$D^{1,0} = \partial_V = \partial + A \tag{67}$$

$$D^{0,1} = \bar{\partial}_V = \bar{\partial}. \tag{68}$$

Dolbeaut-Type Cohomology

Let E be a holomorphic vector bundle over a complex manifold X of complex dimension n, with metric g and with the Hermitian fiber metric G. The generalised Hodge star $\bar{\star}_V$ maps $\Omega^{p,q}(V)$ to $\Omega^{n-p,n-q}(V^*)$ and defines an inner product $\langle \cdot, \cdot \rangle$ as

$$\langle \alpha, \beta \rangle = \int_X \bar{\star}_V \alpha \wedge \beta = \int_X (\alpha, \beta) dV_g, \tag{69}$$

where dV_g is the volume form of g and (α, β) is given by

$$(\alpha, \beta) = \frac{1}{p!q!}\overline{(\alpha)^a_{i_1\ldots i_p\bar{j}_1\ldots\bar{j}_q}}(\beta)^b_{k_1\ldots k_p\bar{l}_1\ldots\bar{l}_q} g^{\bar{i}_1 k_1}\ldots g^{j_1\bar{l}_1}\ldots G_{\bar{a}b}. \tag{70}$$

The adjoint $\bar{\partial}^*_V$ of $\bar{\partial}_V$ is then defined relative to this inner product as $\langle\bar{\partial}_V\alpha, \beta\rangle = \langle\alpha, \bar{\partial}^*_V\beta\rangle$ and is given explicitly by

$$\bar{\partial}^*_V = (-1)^{np+1} \star_V \partial \star_V , \tag{71}$$

where ∂ is the standard Dolbeault differential.

Using these ingredients, we define the Dolbeault Laplacian as

$$\Delta_{\bar{\partial}_V} = \bar{\partial}^*_V\bar{\partial}_V + \bar{\partial}_V\bar{\partial}^*_V, \tag{72}$$

which is self-adjoint with respect to $\langle\cdot,\cdot\rangle$. Following from this, one has the Bochner–Kodaira–Nakano identity [79–82] which relates the $\bar{\partial}_V$-Laplacian to the ∂_V-Laplacian as

$$\Delta_{\bar{\partial}_V} = \Delta_{\partial_V} + [F, \Lambda], \tag{73}$$

where Λ is contraction with the Kähler form ω on X. When V is trivial, so that the curvature F vanishes, this reduces to the usual relation between the ∂- and $\bar{\partial}$-Laplacians on a Kähler manifold, i.e., $\Delta_{\bar{\partial}} = \Delta_{\partial}$.

We define the Dolbeault cohomologies of V by

$$H^{p,q}(X, V) = \frac{\mathrm{Ker}(\bar{\partial}^{p,q}_V : \Omega^{p,q}(V) \to \Omega^{p,q+1}(V)}{\mathrm{Im}(\bar{\partial}^{p,q-1}_V : \Omega^{p,q-1}(V) \to \Omega^{p,q}(V)} \tag{74}$$

By Hodge theory we have

$$\mathcal{H}^{p,q} := \mathrm{Ker}\Delta_{\bar{\partial}_V} \cong H^{p,q}(X, V), \tag{75}$$

A bundle-valued (p, q)-form α is called harmonic (or a zero mode) if

$$\Delta_{\bar{\partial}_V}\alpha = 0. \tag{76}$$

On a compact manifold, harmonic is equivalent to being both $\bar{\partial}_V$- and $\bar{\partial}^*_V$-closed, with the harmonic forms giving the harmonic representatives of the Dolbeault cohomologies $H^{p,q}(X, V)$. We define $h^{p,q}(V) = \dim H^{p,q}(X, V)$ as the dimension of the V-valued (p, q)-form cohomology.

3.5 Calabi-Yau Manifolds

Recall, from Berger's classification 3.1, a complex n-dimensional Calabi-Yau manifold is one whose Riemannian metric has holonomy group $SU(n)$. Since $SU(n) \subseteq U(n)$, a Calabi-Yau manifold is also a Kähler manifold. If g is Kähler, then $\mathrm{Hol}(g) \subseteq SU(n)$ if and only if g is Ricci-flat.

Definition 3.3 *A Calabi-Yau n-fold of complex dimension n is a compact, Ricci-flat Kähler manifold (X, J, g) with holonomy group $SU(n)$.*

Calabi's conjecture from 1954 [83] states:

Conjecture 3.5 *Let X be a compact, complex manifold admitting Kähler metrics. Suppose ρ is a real, closed $(1,1)$-form on X with $[\rho] = 2\pi c_1(X)$ in $H^2_{dR}(X;\mathbb{R})$. Then in each Kähler class on X there exists a unique Kähler metric g with Ricci form ρ.*

Calabi's conjecture was proved in 1978 by Yau [84]. A consequence of Yau's proof is the following: Suppose (X, J) is a compact complex Kähler manifold with vanishing first Chern class. Then every Kähler class on X contains a unique Ricci-flat Kähler metric g, and thus (X, J, g) is Calabi-Yau manifold.

The utility of this theorem is that it is generally quite hard to directly determine whether or not X admits a Ricci-flat metric g. On the other hand, it is a simple matter to compute the first Chern class of X, and in particular, to find example with vanishing first Chern class. The downside is that Yau's proof is not constructive, meaning that it does not provide the explicit Ricci flat metric. In fact, no explicit Ricci-flat metrics are known on any non-trivial Calabi-Yau manifolds.

Topological Properties

Since the holonomy of Calabi-Yau n-folds is $SU(n)$, it can be shown that $h^{0,s} = h^{s,0} = 0$ for $1 < s < n$ and $h^{0,n} = h^{n,0} = 1$. Using the fact that the space X is connected we also have that $h^{0,0} = 1$. The Hodge numbers for the cases $n = 1, 2, 3$ are given in Table 4.

Another topological invariant of interest is the triple intersection form which takes three elements of $H^{1,1}_{\bar\partial}(X)$ and produces a real number:

$$I^{1,1} : H^{1,1}_{\bar\partial}(X) \times H^{1,1}_{\bar\partial}(X) \times H^{1,1}_{\bar\partial}(X) \to \mathbb{R}, \tag{77}$$

$$
\begin{array}{c|ccccccccccccccc}
 & & & & h^{0,0} & & & & & & & & 1 & & & \\
n=1 & & & h^{1,0} & & h^{0,1} & & & = & & & 1 & & 1 & & \\
 & & & & h^{1,1} & & & & & & & & 1 & & & \\
\hline
 & & & & h^{0,0} & & & & & & & & 1 & & & \\
 & & & h^{1,0} & & h^{0,1} & & & & & & 0 & & 0 & & \\
n=2 & & h^{2,0} & & h^{1,1} & & h^{0,2} & & = & & 1 & & 20 & & 1 & \\
 & & & h^{2,1} & & h^{1,2} & & & & & & 0 & & 0 & & \\
 & & & & h^{2,2} & & & & & & & & 1 & & & \\
\hline
 & & & & h^{0,0} & & & & & & & & 1 & & & \\
 & & & h^{1,0} & & h^{0,1} & & & & & & 0 & & 0 & & \\
 & & h^{2,0} & & h^{1,1} & & h^{0,2} & & & & 0 & & h^{1,1} & & 0 & \\
n=3 & h^{3,0} & & h^{2,1} & & h^{1,2} & & h^{0,3} & = & 0 & & h^{2,1} & & h^{2,1} & & 0 \\
 & & h^{3,1} & & h^{2,2} & & h^{1,3} & & & & 0 & & h^{1,1} & & 0 & \\
 & & & h^{3,2} & & h^{2,3} & & & & & & 0 & & 0 & & \\
 & & & & h^{3,3} & & & & & & & & 1 & & &
\end{array}
$$

Table 4: Hodge diamonds for Calabi-Yau n-folds, where $n = 1, 2, 3$.

via

$$I^{1,1}(A, B, C) = \int_X A \wedge B \wedge C. \tag{78}$$

It is called the triple intersection form because it can be equally well phrased in terms of homology in which case the integral just counts the common points of intersection of the three four-cycles, dual to these three two-forms. The integral is a topological invariant of X.

To determine if two Calabi-Yau manifolds are the equivalent we need the topology, and deformation invariant numbers that characterise a manifold independent of a choice of metric.

- **n=1** (elliptic curve): The genus of complex curves is the only topological invariant.

$$g = 1 \Leftrightarrow c_1 = 0. \tag{79}$$

- **n=2** ($K3$ surfaces):

$$c_1 = 0, \quad c_2 = 24, \quad h^{1,1} = 20. \tag{80}$$

- **n=3** (CY threefolds): Wall's theorem [13] states that if two Calabi-Yau threefolds have the same following invariants then they are diffeomorphic and therefore equivalent:

$$(h^{1,1}(X), h^{1,2}(X), c_2(X), I^{1,1}(A, B, C)). \tag{81}$$

Moduli Spaces

Calabi-Yau manifolds are typically part of continuous families, each differing from others in the same family by the precise form of the defining equation and also the precise form of the Kähler class.

If X is Calabi-Yau, then X admits a metric g such that $R_{i\bar{j}} = 0$. Now, given such a g, can we continuously perturb to a new metric $g + \delta g$ such that the Ricci tensor still vanishes? There are two basic types of perturbations δg that we can consider: those with pure and those with mixed type indices:

$$\delta g = \delta g_{ij} dz^i dz^j + \delta g_{i\bar{j}} dz^i d\bar{z}^{\bar{j}} + c.c. \tag{82}$$

As g is a Hermitian metric, the perturbations with mixed type indices preserve the original index structure of g while those of pure type do not. Plugging these perturbations of the metric into the curvature tensor and demanding preservation of Ricci-flatness imposes severe restrictions on δg. In particular, it turns out that $\delta g_{i\bar{j}} dz^i \wedge d\bar{z}^{\bar{j}}$ must be harmonic and hence is uniquely associated to an element of $H^{1,1}_{\bar{\partial}}(X)$. Those deformations of the metric of mixed type therefore correspond to deformations of the Kähler class J of X.

Deformations to the metric with pure type indices yield a metric $g + \delta g$ which is no longer Hermitian. However, by a suitable change of variables, this new metric can be put back into Hermitian form - with only mixed type indices. This change of variables, however, is necessarily not holomorphic as holomorphic coordinate changes cannot affect the index structure of a tensor. What this means is that the new metric is Hermitian with respect to a different complex structure on X - a new set of complex coordinates which are not holomorphic functions of the original coordinates. Those deformations of the metric of pure type which are associated to elements of $H^{2,1}_{\bar{\partial}}(X)$ therefore correspond to deformations of the complex structure of X.

Any two representatives in the same cohomology class yield metric perturbations that can be undone by coordinate redefinitions. Hence, the cohomology classes capture the non-trivial Ricci-flat metric deformations. The parameter space of those Calabi-Yau manifolds continuously connected to some initial one X thereby consists of the possible choices of complex and Kähler structures on the underlying differentiable manifold. We refer to this parameter space as the moduli space of a Calabi-Yau manifold.

Constructions

Calabi-Yau manifolds X can be constructed as algebraic varieties: We embed X in a simple (complex) geometry $\mathcal{A}$ as the zero locus of a set of polynomials

$$p_1(x_i), p_2(x_i), ..., p_n(x_i) = 0, \tag{83}$$

where $p_j(x_i)$'s are polynomials in complex coordinates, x_i of $\mathcal{A}$.

We first need to find a suitable complex ambient space $\mathcal{A}$, in which to define our set of polynomial equations. Since we are after compact Calabi-Yau manifolds it is easier if the ambient space we start with is compact. We therefore choose $\mathcal{A}$ to be complex projective space $\mathbb{CP}^n$ (often abbreviated to $\mathbb{P}^n$), where $\mathbb{CP}^n$ is defined as the quotient $(\mathbb{C}^{n+1} - \{0\})/\sim$, where $\sim$ is the equivalence relation

$$(z_0, z_1, ..., z_n) \sim \lambda(z_0, z_1, ..., z_n), \quad \lambda \neq 0. \tag{84}$$

The z_i are referred to as homogeneous coordinates.

We begin by considering a single hypersurface in $\mathbb{P}^n$, defined by one holomorphic equation:

$$p(z^i) = 0 \tag{85}$$

of degree m. It is a common convention to denote this manifold by the shorthand $\mathbb{P}^n[m]$.

A particularly popular example is the quintic $\mathbb{P}^4[5]$, defined by any homogeneous degree 5 polynomial in the coordinates of $\mathbb{P}^4$. For example the so-called Fermat quintic is defined by the polynomial equation:

$$p = z_0^5 + \cdots + z_4^5 = 0. \tag{86}$$

In total there are $\begin{pmatrix} 5+4 \\ 4 \end{pmatrix} = 126$ degree 5 monomials in the homogeneous coordinates of $\mathbb{P}^4$. The $PGL(5, \mathbb{C})$ action on $\mathbb{P}^4$ generates $(25 - 1)$ coordinate redefinitions. Moreover, the overall scale doesn't effect the polynomial equation. This leaves $126-(25-1)-1 = 101$ distinct coefficients of the defining equation. By varying these coefficients we obtain different manifolds. This is the complex structure moduli of the Calabi-Yau threefold. One family choice of degree and ambient space leads to a parametric family of manifolds, all of which can be smoothly deformed into one another by varying the complex structure.

There is precisely one further metric modulus, namely the overall volume of the quintic hypersurface in $\mathbb{P}^4$. This is the Kähler modulus.

The results for a single hypersurface in a projective space can be easily generalised to complete intersections of polynomial equations $p_j(z_i) = 0$. The zero locus of a set of polynomial equations p_j of degree m_i in $\mathbb{P}^n$ will be Calabi-Yau when

$$n + 1 = \sum_i m_i. \tag{87}$$

We can start with a bigger ambient space - say $\mathbb{P}^5$ - and use two defining equations to define a Calabi-Yau threefold. In this case we have $\mathbb{P}^5[3,3]$ or $\mathbb{P}^5[4,2]$. In fact there are only 5 manifolds (referred to as "cyclic" Calabi-Yau manifolds). In addition to the those above, we also have $\mathbb{P}^6[2,2,3]$ and $\mathbb{P}^7[2,2,2,2]$.

This construction can be generalised in several ways. The first we can consider is to take an ambient space which consists of products of simple projective spaces. To build a Calabi-Yau threefold in a product of p projective spaces, $\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_p}$, we need

$$\sum_{r=1}^{p} n_{n_r} - K = 3, \tag{88}$$

where K is the number of defining equations, $p_j(z) = 0$ in the complete intersection. Each of the defining homogeneous polynomials p_j can be characterised by its multi-degree $\mathbf{m}_j = (m_j^1, ..., m_j^p)$, where m_j^r specifies the degree of the p_j in the homogeneous coordinates $\boldsymbol{F}^{(r)}$ of the factor $\mathbb{P}^{n_r}$ in $\mathcal{A}$. A convenient way to encode this information is by a configuration matrix

$$\begin{bmatrix} \mathbb{P}^{n_1} | & m_1^1 & m_2^1 & \cdots & m_K^1 \\ \mathbb{P}^{n_2} | & m_1^2 & m_2^2 & \cdots & m_K^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbb{P}^{n_p} | & m_1^p & m_2^p & \cdots & m_K^p \end{bmatrix} \tag{89}$$

In order that the resulting manifold be Calabi-Yau, the condition

$$\sum_{j=1}^{K} m_j^r = n_r + 1 \tag{90}$$

has to be satisfied for all $r = 1, ..., p$. Those Calabi-Yau manifolds that are constructed in products of projective spaces are called complete intersection Calabi-Yau manifolds, or CICYs for short.

Another generalisation to is to consider weighted projective spaces $W\mathbb{P}^n$ for the ambient space, where instead the equivalence relation becomes

$$(z_0, z_1, ..., z_2) \sim (\lambda^{w_1} z_0, \lambda^{w_2} z_1, ..., \lambda^{w_n} z_n). \tag{91}$$

Again the Calabi-Yau condition links the weights of the projective coordinates with the degree of the polynomials: $\sum w_i = m$, where m is the maximum degree of the polynomials.

We have seen weighting and taking products as ambient spaces, but these ideas can be generalised still further with interesting $\mathbb{C}^*$-scaling relations on coordinates in an ambient space known as a toric variety. Whilst a (weighted) projective space of complex dimension n is $\mathbb{C}^{n-1}$ (minus the origin), modulo the equivalence relation of the form (91), a toric variety of complex dimension n is $\mathbb{C}^{n+k}$ (minus a point set furnished by the so-called Stanley-Reisner ideal), modulo a set of k equivalence relations (encoded by a charge matrix). All of this data can be conveniently repackaged into lattice cones and polytopes in $\mathbb{R}^n$.

The number of Calabi-Yau n-folds starts small but grows rapidly as you increase n:

- **n=1**: Only one manifold, the elliptic curve (complex curve of genus, $g = 1$, same topology as T^2).
- **n=2**: Only one manifold, the so-called $K3$ surface. This is the unique non-trivial CY twofold (excluding the trivial case of $T^2 \times T^2$).
- **n=3**: There are 7890 CICY threefolds [85] and 473,800,776 families of Calabi-Yau hypersurfaces in toric varieties [86]. It is not known whether the number of Calabi-Yau threefolds is finite.
- **n=4**: There are 921,497 CICY fourfolds [87] and no complete classification of Calabi-Yau fourfolds as toric hypersurfaces has been constructed (although some partial datasets have been generated [88]).

3.6 G_2 Manifolds

Moving on from Calabi-Yau manifolds we now turn our attention to G_2 manifolds, which are those manifolds with dimension 7 that appear in Berger's classification [12].

From [89] we take the definition the group G_2:

Definition 3.4 *Let $\{e^i\}_{i=1,\dots,7}$ denote the canonical basis of $(\mathbb{R}^7)^*$ and $e^{ij} := e^i \wedge e^j$. We define the group G_2 to be the subgroup of $GL(7)$ preserving the 3-form*

$$\varphi_0 = (e^{12} - e^{34}) \wedge e^5 + (e^{13} - e^{42}) \wedge e^6 + (e^{14} - e^{23})^7 + e^{567} \tag{92}$$

under the standard pull-back action on $\Lambda^3(\mathbb{R}^7)^*$, i.e.,

$$G_2 := \{g \in GL(7) | g^*\varphi_0 = \varphi_0\}. \tag{93}$$

Moreover, we define a G_2 structure as:

Definition 3.5 *A G_2-structure on a 7-manifold M is a 3-form $\varphi \in \Omega^3(M)$ such that, at every point $p \in M$, $\varphi_p = f_p^*(\varphi_0)$ for some frame $f_p : T_pM \to \mathbb{R}^7$.*

Given a G_2-structure φ, the set of all global frames f satisfying $\varphi = f^*\varphi_0$ is a principal subbundle of the frame bundle $F \to M$, with fibre G_2.

The G_2-structure φ induces a Riemannian metric g_φ and a volume form vol_φ via the following identity:

$$(X \lrcorner \varphi) \wedge (Y \lrcorner \varphi) \wedge \varphi = -6 g_\varphi(X, Y) \text{vol}_\varphi, \tag{94}$$

for any vector fields X and Y on M, where $\lrcorner$ denotes the interior product.

Let $\star_\varphi$ denote the Hodge star operator associated to the metric g_φ and volume form vol_φ. We denote by ψ the dual 4-form $\psi = \star_\varphi \varphi$.

The torsion of φ is defined as the covariant derivative $\nabla\varphi$, where ∇ is the Levi-Civita connection coming from the induced metric g. We say that φ is torsion-free if $\nabla\varphi = 0$.

Definition 3.6 *A G_2 manifold is defined by the pair (M, g), where M is a 7-manifold and φ is a torsion free G_2-structure on M.*

The following is a theorem by Fernández and Gray [90]:

Theorem 3.6 *Let M be a 7-manifold with G_2-structure φ and induced metric g; then the following are equivalent:*

1. *$\nabla\varphi = 0$, i.e., M is a G_2-manifold;*

2. $Hol_g(M) \subset G_2$*;*

3.

$$d\varphi = d \star_\varphi \varphi = 0, \tag{95}$$

From this theorem we see that a G_2-manifold has holonomy contained in G_2 and not necessarily full G_2. The holonomy of a G_2 manifold must be one of the following:

- $\mathrm{Hol}_g(M) = \{1\} \Leftrightarrow b_1(M) = 7$
- $\mathrm{Hol}_g(M) = SU(2) \Leftrightarrow b_1(M) = 3$
- $\mathrm{Hol}_g(M) = SU(3) \Leftrightarrow b_1(M) = 1$
- $\mathrm{Hol}_g(M) = G_2 \Leftrightarrow b_1(M) = 0$

In M-theory we are interested in constructing manifolds with full holonomy G_2, and not a strictly smaller subgroup, in which case the fundamental group $\pi_1(M)$ is finite.

In this section we provided an overview of Calabi-Yau and G_2 manifolds which are the compactificaiton spaces that achieve minimal supersymmetry in four-dimensions for ten dimensional (super)string theory and eleven-dimensional M-theory respectively. In the results presented later we apply machine learning techniques to generate these geometries, study their properties and compute aspects of the low energy effective field theories that arise. Therefore, in the next section we complete our background information by introducing the subject of machine learning and describe the algorithms that will be used later.

4 Machine Learning Background

In this section we provide a brief introduction to some fundamental terms and concepts used in machine learning. For more information we refer the reader to the following resources. The standard textbooks [91, 92] focus on the theory and foundations of deep learning. For a gentler introduction to the subject a great online resource is Andrew Ng's course "Machine Learning Specialization" [93]. Finally, [94] is an excellent introduction to machine learning for physicists, which includes notebooks with practical demonstrations. We give specific details on neural networks, genetic algorithms, reinforcement learning and symbolic regression as these are the algorithms that we will encounter in the later sections.

Machine learning is a subset of artificial intelligence, whereby a system learns and improves from experience without being explicitly programmed. There are three main categories:

- **Supervised learning** is used when one is provided with input data labelled with the desired output. The aim of supervised learning is to approximate the mapping function from input to output so that given a new set of input data one can accurately predict the output. Supervised learning can be broken down into regression problems, where the output is a continuous numerical value, and classification problems, where the output is a discrete class. Examples of supervised learning algorithms include linear regression, neural networks, random forests, and support vector machines.

- **Unsupervised learning**, on the other hand, is applied to datasets consisting of input data without labelled responses. In this case, unsupervised learning algorithms look for underlying structures and patterns within the data. These methods can roughly be divided into clustering algorithms, such as k-means, and dimensionality reduction algorithms, such as principal component analysis.

- **Reinforcement Learning** is about learning the optimal behavior in an environment in order to obtain a maximum reward. This optimal behavior is learned through an agent interacting with the environment and receiving rewards or penalties based on its actions.

4.1 Neural Networks

In supervised learning, we are presented with a set of input samples $X_\mu \in \mathbb{R}^p$ each with an associated output label $y_\mu \in \mathbb{R}^d$. The goal is to find a function f so that when we are given a new sample X_{new} without a label, we are able to approximate well the label using the function

$f(X_{\text{new}})$.

In order to test the performance of f one splits the dataset $\{X_\mu, y_\mu\}$ into two sets: a training set and a test set. The usual train:test split is 80:20. The training dataset is used to learn f and after training f is used to make predictions on the unseen test set. Cross-validation is a method commonly used to get an unbiased evaluation, whereby the data is shuffled and split into k groups, then each group acts as the test group once and the remaining groups are combined to create the training set. Each time the model is trained on the training set, evaluated on the test set and the evaluation scores are recorded. The mean and standard deviation of the evaluation scores are then calculated and used to measure the model performance.

To explain neural networks, we begin by introducing the building block of any neural network; a neuron. A neuron takes in a set of input data $\{x_i\} \in \mathbb{R}^n$ and produces a single numerical output $y \in \mathbb{R}$. We can divide the function of the neuron into three parts:

1. Firstly, each input x_i is multiplied by a weight $w_i \in \mathbb{R}$: $w_i x_i$.
2. Next, all the weighted inputs are summed and a bias $b \in \mathbb{R}$ is added: $\sum_i w_i x_i + b$.[4]
3. Finally, the sum is passed through an activation function $a : \mathbb{R} \to \mathbb{R}$ which produces an output: $y = a(\sum_i w_i x_i + b)$.

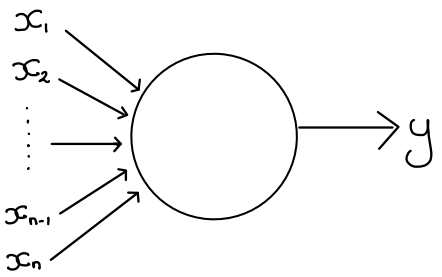


Figure 5: A neuron.

A neural network is simply a collection of neurons connected together in a series of layers:

$$f_w(X_\mu) = a^{(L)}(W^{(L)} \cdots a^{(2)}(W^{(2)} a^{(1)}(W^{(1)} X_\mu + b^{(1)}) + b^{(2)}) + b^{(L)}), \tag{96}$$

[4]The bias b is a constant added to shift the activation function to the left or right, which can help with successful learning. See [95] for a detailed explanation.

where L is the depth (i.e. number of layers) of the network, $W^{(i)} \in \mathbb{R}^{r_i \times r_{i-1}}$ are the weight matrices, where r_i for $1 \leq i \leq L-1$ is the width (i.e. number of neurons) of the i-th layer, $b^{(i)}$ are the biases, $a^{(i)}$ are the activation functions and $w = \{W^{(1)}, b^{(1)}, ..., W^{(L)}, b^{(L)}\}$ denotes the set of all weights and biases. Neural networks with $L > 2$ are called deep neural networks and machine learning on these networks is called deep learning.

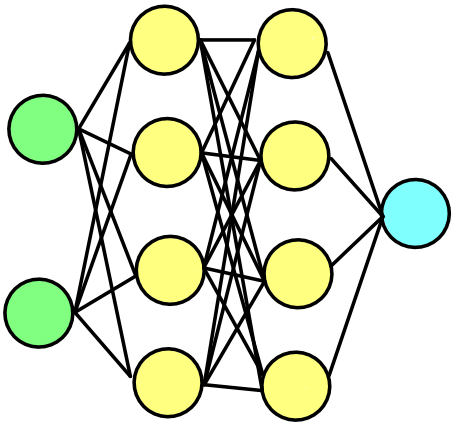
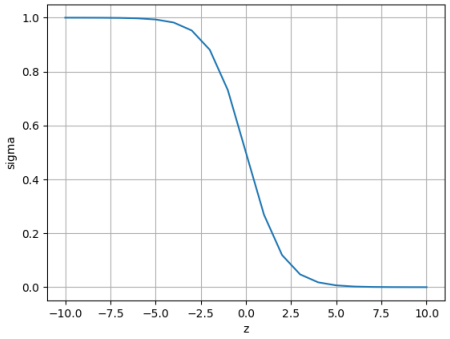


Figure 6: A neural network consisting of an input layer layer (green) with two neurons, two hidden layers (yellow) each with four neurons, and an output layer (blue) with one neuron. The edges between neurons represent the flow of information between layers, where the output of one neuron is fed as input to the next neuron.

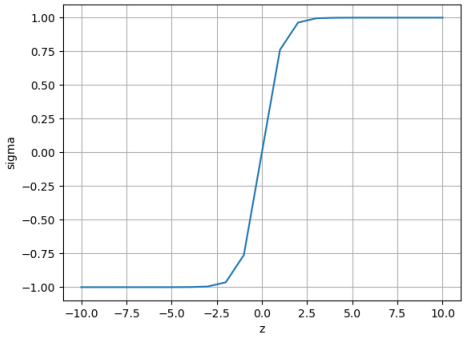
The activation functions $a^{(i)}$ play a crucial role in neural networks. There are many variants of activation functions, including binary, linear, and non-linear functions. Using linear activation functions means that the output of the model would simply be a linear function of the input. Using non-linear activation functions, on the other hand, allows the network to learn complex patterns and relationships in the data. A few commonly used non-linear activation functions are Sigmoid, Tanh, Rectified Linear Unit (ReLU), and Softmax shown in Figure 7. It isn't always obvious which activation function one should use and finding the right one is often a case of trial and error. The activation function in the final output layer, however, is usually chosen based on the type of problem, as the diagram in Figure 8 shows.

Training

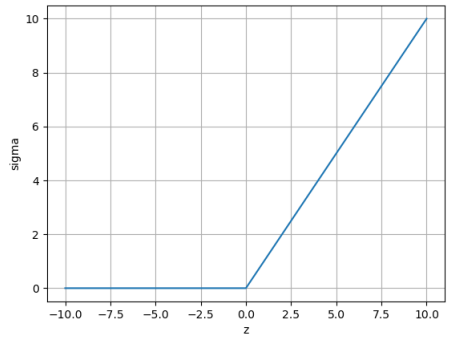
The process of training a neural network involves repeatedly calculating the "error" of the model outputs $f_w(X_\mu)$ so that the weights can be updated in order to reduce the error. Computing



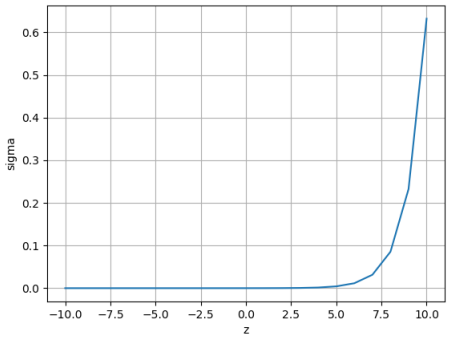
(a) Sigmoid: $\sigma(z) = \frac{1}{1+e^z}$



(b) Tanh: $\sigma(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$



(c) ReLU: $\sigma = \max 0, z$



(d) Softmax: $\sigma(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$

Figure 7: Common activation functions for neural networks.

the error requires a choice of loss function $\mathcal{L}[f_w(X_\mu), y_\mu]$. Typical loss functions for regression problems are mean absolute error (MAE):

$$MAE = \frac{1}{n}\sum_{\mu=1}^{n} |f_w(X_\mu) - y_\mu|, \tag{97}$$

and mean square error (MSE):

$$MSE = \frac{1}{n}\sum_{\mu=1}^{n} (f_w(X_\mu) - y_\mu)^2. \tag{98}$$

For classification problems the default loss function is cross entropy (CE):

$$CE = -\sum_{\mu=1}^{n} y_\mu \log\left(f_w(X_\mu)\right) \tag{99}$$

for multiclass classification, and binary cross entropy (BCE):

$$BCE = -\sum_{\mu=1}^{n} (y_\mu \log\left(f_w(X_\mu)\right) + (1 - y_\mu), \log\left(1 - f_w(X_\mu)\right)) \tag{100}$$

for binary classification.

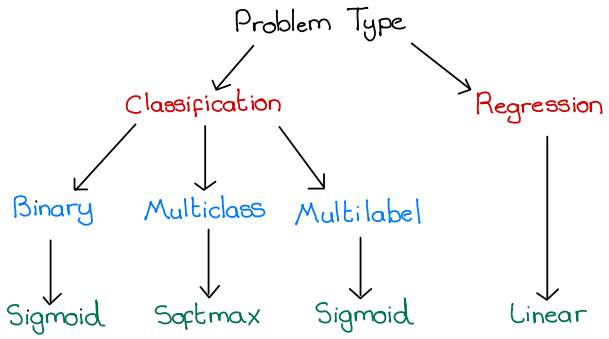


Figure 8: Output layer activation choices for supervised learning problems.

During training the weights w are adjusted in order to minimise the total loss. The method by which we change the weights is called the optimisation algorithm. The simplest optimisation algorithm is called gradient descent, which determines in which direction the model parameters should be altered so that the loss function can reach the minima. In the gradient descent optimisation algorithm, the model weights are adjusted as follows:

$$w_{t+1} = w_t - \alpha \frac{\partial \mathcal{L}}{\partial w} \tag{101}$$

where α is a parameter called the learning rate that controls the speed and accuracy of training. A larger α allows the model to learn faster, but at the cost of possibly converging to a sub-optimal set of parameters. A smaller α may allow the model to learn a more optimal set of parameters but may take significantly longer to train. The default starting value for the learning rate is 0.01, from which one can search for the optimal value. In mini-batch gradient descent, the training dataset is divided into smaller batches, and the model parameters are updated after every batch. If the batch size is equal to the size of the entire training dataset, then this is called batch gradient descent. On the other hand, if the size of the batches is 1 this is called stochastic gradient descent (SGD).

These gradient descent methods of optimisation present some challenges. Firstly, choosing a value for the learning rate can be difficult. A value that is too small can lead to painfully slow convergence. On the other hand, a value too large can cause the loss function to overshoot the minimum and in some cases diverge. Another problem that arises in gradient descent methods is getting trapped at local minima. There exist more advanced optimisation methods that tackle these challenges, some of which we describe here.

First of all, we introduce momentum [96]. Instead of updating the parameters depending

only on the current gradient, we use an exponential moving average over past gradients where recent gradients are weighted higher. In doing so, it reduces the chance of getting stuck at local minima. The model parameters are updated as

$$w_{t+1} = w_t - \alpha m_t, \tag{102}$$

where

$$m_t = \beta m_{t-1} + (1-\beta)\frac{\partial \mathcal{L}}{\partial m_t} \tag{103}$$

and m_t is initialised to 0. The parameter β takes values between 0 and 1 represents the degree of weighting increase. A smaller β discounts older observations faster, whereas a higher β accommodates more past gradients. Therefore, β is usually set to 0.9 in most cases.

So far we have considered the learning rate α kept constant. However, keeping the learning rate fixed might not be the best option for all parameters, and could cause problems with convergence. Some parameters might need more frequent updates to hasten convergence, while others might need smaller changes to prevent overshooting the ideal value. We now consider an optimisation algorithm called Adagrad (adaptive gradient) [97], which changes the learning rate of the parameters θ at each time step, t. In this case, the parameters are updated as

$$w_{t+1} = w_t - \frac{\alpha}{\sqrt{v_t + \epsilon}}\frac{\partial \mathcal{L}}{\partial w_t}, \tag{104}$$

where v is cumulative sum of the current and past gradients squared

$$v_t = v_{t+1} + \left(\frac{\partial \mathcal{L}}{\partial w_t}\right)^2, \tag{105}$$

and v_t is initialised to 0. The ϵ parameter is a small number, usually set around 10^{-7}, added to ensure that we do not encounter a zero in the denominator. Adagrad adjusts the learning rate for each parameter based on the parameter's significance to the optimisation process. This lessens the requirement for manual adjustment of learning rates.

A similar optimisation algorithm to Adagrad is RMSprop (root mean squared propagation) [98] where we introduce the exponential moving average to update v

$$v_t = \beta v_{t-1} + (1-\beta)\left(\frac{\partial \mathcal{L}}{\partial w_t}\right)^2 \tag{106}$$

Again v is initialised to 0.

The last optimisation method we shall consider is Adam [99] (adaptive moment estimation), which is a combination of RMSprop and momentum. The model parameters in this optimisation algorithm are updated as

$$w_{t+1} = w_t - \frac{\alpha}{\sqrt{\hat{v}_t} + \epsilon}\hat{m}_t \tag{107}$$

where

$$\hat{m}_t = \frac{m_t}{1 - \beta_1} \tag{108}$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2} \tag{109}$$

and

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1)\frac{\partial \mathcal{L}}{\partial w_t} \tag{110}$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2)\left(\frac{\partial \mathcal{L}}{\partial w_t}\right) \tag{111}$$

m and v are both initially set to 0.

The advanced optimisation algorithms we have discussed here are more computationally expensive than simple gradient descent methods, but for complex systems perform much better. Adam is the most commonly used optimisation algorithm and is the optimisation method we use in later in Section 6.

Evaluation

For regression tasks, typical performance metrics include MAE (97), MSE (98) and the R^2 score, which is defined as the proportion of the variance in the dependent variable that is predictable from the independent variable(s). Therefore, an R^2 score close to 1 means the regression model is a good fit, whereas a score close to 0 means the model is a poor fit. The equation for computing R^2 is

$$R^2 = 1 - \frac{\sum_{\mu=1}^{n}(y_\mu - f_w(X_\mu))^2}{\sum_{\mu=1}^{n}(y_\mu - \overline{f_w(X_\mu)})^2}, \tag{112}$$

where

$$\overline{f_w(X_\mu)} = \frac{1}{n}\sum_{\mu=1}^{n} f_w(X_\mu). \tag{113}$$

For classification tasks, the default performance metric is accuracy

$$acc = \frac{TP + TN}{TP + FP + TN + FN}, \tag{114}$$

where TP, FP, TN, FN denote the true positive, false positive, true negative, and false negative counts respectively. Accuracy, however, can sometimes be misleading. For example consider the binary classification situation where the first class occurs 99% of the time and the second class occurs only 1% of the time. An algorithm that just predicts the first class every time would give an accuracy score of 99% which looks great but this isn't actually a good model. A better metric in this case is Mathew's correlation coefficient (MCC)

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}}, \tag{115}$$

which is similar to the Pearson correlation coefficient[5], where a score of 1 indicates complete agreement between predictions and truth, 0 indicates that the predictions are no better than random guessing, and -1 indicates complete disagreement between the predicted and truth. MCC is a better choice of metric when dealing with classes of different sizes.

4.2 Genetic Algorithms

Genetic algorithms (GAs) are optimization algorithms, first put forward in the 1950s [100, 101] and later formalized by Holland [102] in the 1970s. The algorithms mimic the process of natural selection in Darwin's theory of evolution [103], whereby the fittest individuals reproduce to create the next generation. This process iterates and in the end there will be a generation of the fittest individuals.

The first step of a GA evolution is to randomly generate, from an environment E, an initial population P_0 that contains a certain number n_{pop} of individuals. Individuals are made up of a set of genes joined together. Usually, binary values are used to represent the genes in which case individuals are represented by bitstrings (a list of 0s and 1s).

The next step is to define some fitness function $f : E \to \mathbb{R}$ and assign a fitness score to each of the individuals in the population. The fitness score of an individual will determine how likely it is to be selected for breeding.

The genetic evolution then consists of a sequence:

$$P_0 \to P_1 \to \cdots\cdots \to P_{n_{\text{gen}}-1} \to P_{n_{\text{gen}}} \tag{116}$$

[5]The Pearson correlation coefficient is a coefficient that measures the linear correlation between two sets of data. It is defined as the ratio between the covariance of two variables and the product of their standard deviations.

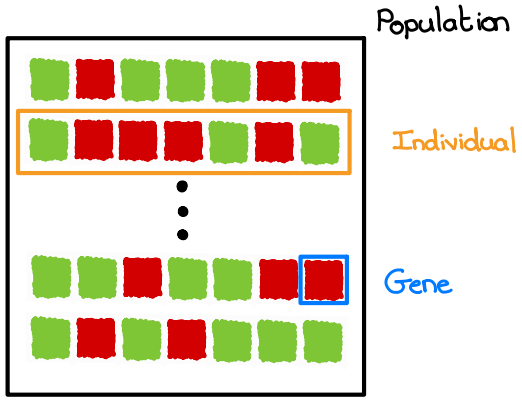


Figure 9: Genetic algorithm population made up of bitstrings representing individuals.

of further n_{gen} populations. The basic evolutionary process, $P_k \to P_{k+1}$, to obtain population $k+1$ from population k is carried out in three steps, namely, (i) selection, (ii) crossover, and (iii) mutation. We describe these three steps in turn.

(i) Selection: In the selection phase pairs of fit individuals are chosen for reproduction. Parents are selected using a probability distribution $p_k : P_k \to [0,1]$, based on the fitness scores. Individuals with a high fitness score are more likely to be selected. To keep the size of the population constant, $n_{\text{pop}}/2$ pairs are selected for breeding. [6]

(ii) Crossover: For each pair selected in step (i), the genes of the parents are combined in some way to create two offspring. Carrying this out for all pairs generates a set of new individuals, which make up the new population $\tilde{P}_{k+1}$.

(iii) Mutation: In the final step, a certain fraction, r_{mut}, of genes in the individuals of the population $\tilde{P}_{k+1}$ from step (ii) are subjected to mutation. Mutation occurs in order to maintain diversity within the population and prevent premature convergence.

A common addition to the above algorithm is elitism which means that the fittest individual from population P_k is copied to the population P_{k+1} unchanged. The algorithm terminates either after a set number of generations, or once a satisfactory fitness level of the population has been reached.

In summary, a GA evolution is subject to the following hyperparameter choices: the population size n_{pop}, the number of generations n_{gen} and the mutation rate r_{mut}. Typically

[6] Note that pairs of individuals are selected independently and so individuals can be selected more than once, in which case they contribute to more than one child.

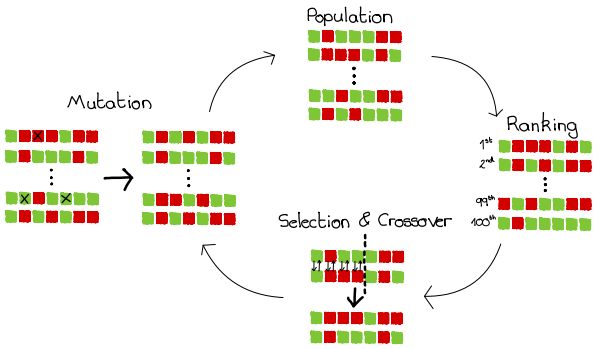


Figure 10: Diagram representing the evolution of a genetic algorithm.

many GA evolutions, each with a new randomly sampled initial population P_0, are carried out.

4.3 Symbolic Regression

Symbolic regression (SR) is the subfield of machine learning which infers symbolic mathematical expressions from data. These methods allow one to obtain accurate predictive models and more importantly the learned models are interpretable. Of course in physics we would like our models to be interpretable as this allows us to understand the underlying phenomenon. Although SR is fit to dataset it aims to find a symbolic model that can generalise to unseen data.

In this section, we will introduce the premise of SR and give a short description of a few of the methods, focusing more attention on the genetic algorithm tree-based method since this is the method that will be used later in Section 6.3. For more detail on the various other methods, we refer the reader to the review [104].

Problem Definition

The problem of symbolic regression can be defined in terms of the following:

- **Data**: A dataset $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^n$ of input $\mathbf{x}_i \in \mathbb{R}^d$ output $y_i \in \mathbb{R}$ pairs.
- **Function Class**: A function class $\mathcal{F} = \{f : \mathbb{R}^d \to \mathbb{R}\}$.
- **Loss Function**: A loss function

$$l(f) := \sum_{i=1}^{n} l(f(\mathbf{x}_i), y_i), \tag{117}$$

for $f \in \mathcal{F}$.

A common choice of loss function is the squared difference: $l(f) = \sum_i (y_i - f(\mathbf{x}_i))^2$. The optimisation task is to find $f^* \in \mathcal{F}$ that minimises the loss function:

$$f^* = \arg\min_{f \in \mathcal{F}} l(f). \tag{118}$$

To define the function class $\mathcal{F}$ we must specify a library of elementary arithmetic operations, mathematical functions and variables. An element $f \in \mathcal{F}$ is a function obtained by the composition of objects in the library. For example, let L be the library defined as

$$L = \{\mathrm{Id}(\cdot), \mathrm{add}(\cdot,\cdot), \mathrm{sub}(\cdot,\cdot), \mathrm{mul}(\cdot,\cdot), +1, -1\}. \tag{119}$$

With such a library the function class $\mathcal{F}$ consists of all polynomials, in one variable x, with integer coefficients. In practice, L may include many other mathematical functions, such as logarithm and exponential, sine and cosine, square root, etc. It is advantageous to have prior domain knowledge in defining a minimal library since it reduces the search space.

Methods

There are many variants of SR but methods can be categorised into the following classes: linear and non-linear regression-based methods, expression tree-based methods, physics-inspired and mathematics-inspired methods. These are summarised in Table 5.

Method	Tool	Expression Form	Unkown	Search Space
Linear	Uni-D Linear System	$y = \sum_i \theta_i f_i(\mathbf{x})$	$\{\theta\}_i$	$\mathbb{R}$
	Multi-D Linear System	$y_i = \sum_j f_j(\mathbf{x})$	$(\{\theta\}_i)_j)$	$\mathbb{R}$
Non-linear	Neural Network	$y = f(W \cdot \mathbf{x} + b)$	$\{W\}$	$\mathbb{R}$
Expression-Tree	Genetic Algorithm	Expression Tree		Trees
	Transformers	Set2Seq Mapping	$\{W_q, W_k, W_v\}$	$\mathbb{R}$
	Reinforcement Learning	Set2Seq Mapping	$\pi(\theta)$	$\mathbb{R}$
Physics-Inspired	AI-Feynman	$y = f(\mathbf{x}, \theta)$	-	-
Mathematics-Inspired	Symbolic Metamodels	$G(\mathbf{x}, \theta)$	θ	$\mathbb{R}$

Table 5: Summary of symbolic regression methods.

The linear SR method defines the functional form as a linear combination of (possibly) non-linear functions of x from the library L:

$$f(x, \theta) = \sum_{j=1}^{m} \theta_j h_j(x), \tag{120}$$

where $h_j \in L$. In this case, the optimisation problem amounts to finding the set of parameters θ_j that minimise the loss function (117).

Non-linear SR, on the other hand, defines the function form as:

$$f(x, W) = \sigma\left(\sum_i W_i \sigma\left(\sum_j W_j \sigma\left(\cdots \sum_\ell W_\ell x\right)\right)\right), \tag{121}$$

where σ is a non-linear activation function and the W's are the learnable parameters of a neural network. Here the problem amounts to finding a set of parameters W_I which minimises the loss function (117).

Expression tree-based methods treat mathematical expressions as unary-binary trees whose internal nodes are operators and terminals are operands (variables or constants), as illustrated in Figure 11. This category comprises GA-based, transformer-based and RL-based methods. In GA-based methods, a set of transition rules (e.g. mutation, crossover, etc.) is defined over the tree space and applied to an initial population of tress throughout many iterations until the loss function is minimised. The crossover operation involves exchanging the content of two individuals, for instance, by swapping one random subtree of one individual with a random subtree of the other individual, as shown in Figure 12a. The mutation operation involves random variations to an individual, for instance, by replacing one random subtree with another randomly generated subtree, as shown in Figure 12b.

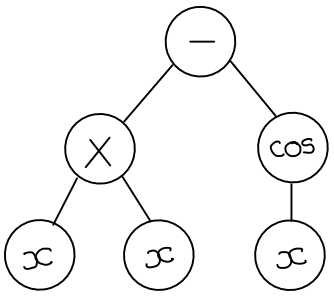
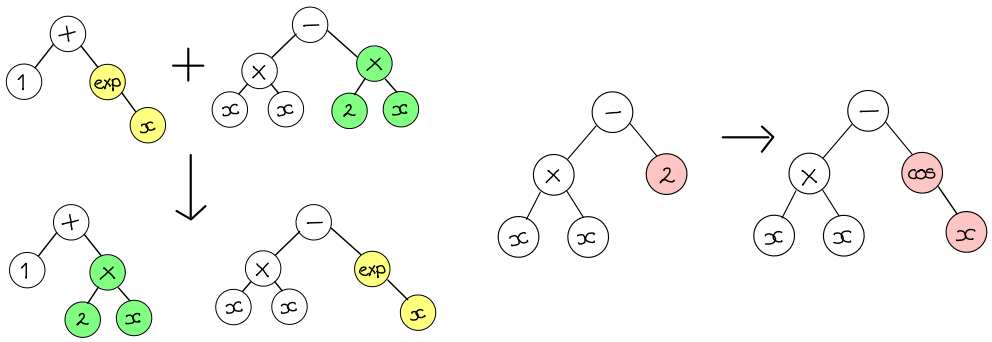


Figure 11: Expression tree representing the function $f(x) = x^2 - \cos x$.

In this section we have introduced the fundamentals of machine learning and provided brief summaries of the algorithms we will encounter in the next sections, namely: neural networks (NNs), symbolic regression (SR) and genetic algorithms (GAs). In the following section we will



(a) (b)

Figure 12: Crossover (a) and mutation (b) operations on expression trees in genetic algorithm based symbolic regression.

show how GAs can be used to generate Calabi-Yau manifolds with certain properties and in Section 6 we will see how NNs can be applied to learn topological properties of G2 manifolds from algebraic input data describing the manifold. Furthermore we see that SR can be used to approximate functions for the topological properties given the same input data.

5 Generating Calabi-Yau Manifolds

The term "string landscape" refers to the set of all effective field theories (EFTs) obtained from dimensional reduction of string theory. As we learnt in Section 2.3, Calabi-Yau (CY) threefolds provide the appropriate compactification space for ten-dimensional superstring theory to preserve $\mathcal{N} = 1$ supersymmetry in four-dimensions. Similarly, CY fourfolds are the appropriate compactification space for twelve-dimensional F-theory. Estimates on the size of the landscape include the 10^{500} type IIB flux compactifications [105, 106], as well as $10^{272,000}$ F-theory flux compactifications on a single elliptically fibered CY fourfold [107]. Random sampling is most likely to fail at identifying models in the landscape whose low-energy symmetry and particle content match those of the Standard Model (SM), and systematic searching is beyond computational capabilities. Without a method of identifying SM like models within the landscape, the usual procedure so far has been to choose a CY manifold mostly at random, generate vast amounts of EFTs only to throw away large numbers of these which don't satisfy the often stringent constraints. Clearly this procedure is inefficient and motivates a more targeted algorithmic approach to choosing a CY. In this work, we use machine learning methods to address this problem over the space of CYs generated from reflexive polytopes and their triangulations.

In [108, 109], Batyrev and Borisov showed how mirror pairs of $(n-1)$ complex dimensional CY manifolds can be realized from hypersurfaces in toric varieties constructed from n-dimensional reflexive polytopes. These toric varieties are generally singular, but one can resolve the singularities by taking fine regular star triangulations (FRSTs) of the polytope. Motivated by Batyrev and Borisov's discovery, Kreuzer and Skarke (KS) devised an algorithm to generate all reflexive polytopes in n dimensions [110]. The algorithm consists of two steps. First, a set S of "maximal" polytopes is constructed such that any reflexive polytope is a subpolytope of a polytope in S. These maximal polytopes are defined by a so-called "weight system" or a combination of weight systems. Second, all subpolytopes of polytopes in S are constructed and checked for reflexivity. The complete classification of 4319 three-dimensional reflexive polytopes with this algorithm was accomplished in [111]. From these we obtain K3 surfaces, which is to say, CY twofolds. Proceeding to dimension three, the 184, 026 weight systems giving rise to four-dimensional reflexive polytopes were presented in [112], and the resulting 473, 800, 776 four-dimensional reflexive polytopes, leading to CY threefolds, were listed in [86]. In five-dimensions the total number of reflexive polytopes is prohibitively large,

and Schöller and Skarke were only able to run the first stage of the algorithm to calculate all $322,383,760,930$ weight systems corresponding to maximal polytopes [88]. They found that $185,269,499,015$ of these weight systems give rise to reflexive polytopes directly.[7][8]

A polytope is said to be reflexive when it satisfies a set of conditions: it must be a lattice polytope (that is, its vertices must lie on integer lattice points), it must satisfy the IP property (that is, it must have only a single interior point) and its dual must also be a lattice polytope that satisfies the IP property. Alternatively and equivalently, a polytope is reflexive if and only if it satisfies the IP property and all bounding hyperplanes of the polytope lie at unit hyperplane distance from the origin. With multiple criteria, finding reflexive polytopes is an example of a multiobjective optimisation problem (MOOP). Regression is unsuitable for this types of problem since any loss function will have local minima, where some but not all of the criteria are satisfied. Popular methods for solving MOOPs are genetic algorithms (GAs) and reinforcement learning (RL), which explore an "environment" in order to maximize a fitness or reward function. In [49], GAs were applied to search for reflexive polytopes. In this section we present the results of this paper.

Significant work has already been done on applying machine learning techniques to study polytopes [115–117]. GAs have also been shown to be successful at scanning for phenomenologically attractive string models [118, 119], but [49] is the first recorded use of these algorithms to search for reflexive polytopes. There was also some recent work done on generating smooth Fano polytopes [120] using sequential modeling. However, we found applying this methodology to reflexive polytopes without the smoothness condition did not yield good results. Inspired by [49], others have since used GAs to generate FRSTs of reflexive polytopes [121] and applied this to search for smooth CYs which optimise axion decay constants and axion-photon couplings in type IIB. In parallel to [121], the authors of [49] subsequently also designed an algorithm to generate FRSTs of reflexive polytopes using RL instead of GAs [31], and applied this to search for smooth CYs together with holomorphic vector bundles that satisfy anomaly cancellation and poly-stability conditions in

[7] It can be that distinct weight systems correspond to the same polytope. This means that the total number of unique five-dimensional reflexive polytopes found by Schöller and Skarke is less than $185,269,499,015$ (it is not known by how much).

[8] All data produced by the KS algorithm for reflexive polytopes in three and four dimensions can be found at [113]. The complete Schöller-Skarke dataset of five-dimensional reflexive polytopes can be found at [114].

heterotic compactification.

The rest of the chapter is organised as follows: in Section 5.1 we introduce some of the basics of toric geometry so that we can understand the construction of Calabi-Yau manifolds from reflexive polytopes. We then, in Section 5.2, give the details our GA model; we describe the data encoding of polytopes as bitstrings and the define the fitness function that determines how close a polytope is to being reflexive. After this (Section 5.3) we present the results of our GA searches for two-, three-, four- and five-dimensional reflexive polytopes. Finally, in Section 5.4 we finish with a discussion and prospectus.

5.1 Toric Geometry

Definition 5.1 *Let $M \cong \mathbb{Z}^n$ and $N = Hom(M, \mathbb{Z})$ be a dual pair of lattices with the pairing $\langle \cdot, \cdot \rangle : N \times M \to \mathbb{Z}$, and let $M_{\mathbb{Q}}$, $N_{\mathbb{Q}}$ be their rational extensions. A (convex) polytope Δ in $M_{\mathbb{Q}}$ has two equivalent definitions:*

1. ***Vertex Representation**: Convex hull of a finite number of points (vertices) $v_1, ..., v_m \in M_{\mathbb{Q}}$.*
2. ***Hyperplane Representation**: Intersection of a set of linear inequalities $\langle \boldsymbol{u}_i, \boldsymbol{x} \rangle \leq d_i$ where $\boldsymbol{u}_i \in N_{\mathbb{Q}}$ are primitive lattice points and $d_i \in \mathbb{R}$.*

The vertices can be combined into an $n \times m$ vertex matrix $V = (v_1, \ldots, v_m)$.

Definition 5.2 *A face θ of Δ is defined as*

$$\theta = \{m \in \Delta | \langle n, m \rangle = d\} \ , \tag{122}$$

for some $n \in N_{\mathbb{Q}}$ and $d \in \mathbb{Q}$.

In other words, a face of the polytope is the intersection of the polytope with one of the bounding hyperplanes. The value of d is the lattice distance of the hyperplane from the origin. A face of codimension 1 is referred to as a facet. We denote the set of all facets by $F(\Delta)$.

Definition 5.3 *The dual (or polar) polytope of Δ is defined as*

$$\Delta^* = \{n \in N_{\mathbb{Q}} | \langle n, m \rangle \geq -1 \ \forall \, m \in \Delta\} \ . \tag{123}$$

(a) Δ (b) Σ_Δ

Figure 13: Polytope Δ and normal fan Σ_Δ associated to the polytope defined by the vertex matrix (125).

A polytope $\Delta \in M_{\mathbb{Q}}$ is called a lattice polytope if all its vertices are integral, i.e. $v_i \in M$. Given an n-dimensional lattice polytope $\Delta \subset M$, one can construct a compact toric variety X_Δ of complex dimension n. In short, one constructs the normal fan Σ_Δ as follows: for a face θ of Δ, let $\sigma_\theta \subset N$ be the dual of the cone:

$$\sigma_\theta^\vee := \{\lambda(u - u') | u \in \Delta, u' \in \theta, \lambda \geq 0\} \subset M. \tag{124}$$

The normal fan is then given as $\Sigma_\Delta := \{\sigma_\theta\}$ for all faces θ of Δ.

Example 1 *Consider the two-dimensional lattice polytope $\Delta \subset \mathbb{Z}^2$, as shown in Figure 13a, defined by the vertex matrix:*

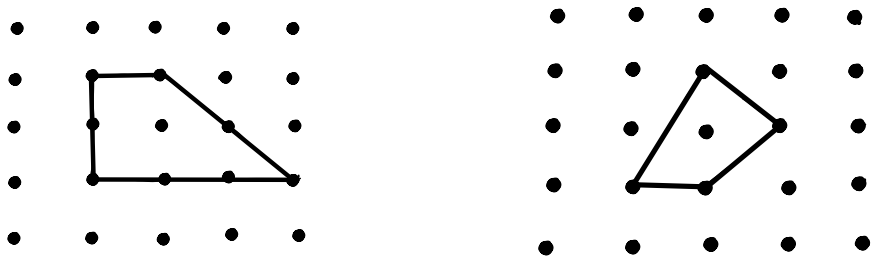
$$V = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 1 & 1 & 1 & -1 \end{pmatrix}. \tag{125}$$

The normal fan for this polytope is shown in Figure 13b.

From the normal fan, the construction of the compact projective toric variety X_Δ follows the usual procedure (see [122] for details), where each cone σ_θ gives rise to an affine toric variety $U_{\sigma_\theta} = \text{Spec}(\mathbb{C}[\sigma_\theta^\vee \cap M])$ and one glues these patches together.

A polytope is said to satisfy the IP property if the origin is its only interior lattice point.

Definition 5.4 *A lattice polytope Δ is called reflexive if it satisfies the IP property and if its dual Δ^* is also a lattice polytope that satisfies the IP property. Equivalently, Δ is reflexive if and only if it satisfies the IP property and if all its hyperplanes lie at a distance one from the origin, that is, if all $d_i = 1$.*



(a) $\Delta \subset M$ (b) $\Delta^* \subset N$

Figure 14: A dual pair (Δ, Δ^*) of two-dimensional reflexive polytopes.

The connection between Calabi-Yau (CY) manifolds and reflexive polytopes is given by the following theorem [108, 109]:

Theorem 5.1 *Let $\Delta \subset M$ be an n-dimensional lattice polytope and X_Δ the corresponding n complex dimensional toric variety. If Δ is reflexive then it follows that X_Δ is Gorenstein Fano with at most canonical singularities, and moreover the zero locus of a generic section of the anticanonical bundle $-K_X$ is a CY variety $\mathcal{M}$ of complex dimension $n-1$.*

The mirror CY $\mathcal{W}$ is similarly obtained from the dual polytope.

For CY $n-1$-folds, built from n-dimensional reflexive polytopes, there exist formulas to compute the Hodge numbers in terms of the polytope data such as [108, 109]:

$$h^{1,1}(X) = \ell(\Delta^*) - n - 1 - \sum_{\text{codim}\theta=1} \ell^*(\theta^*) + \sum_{\text{codim}\theta^*=2} \ell^*(\theta^*)\ell^*(\theta), \tag{126}$$

where θ and θ^* are the faces of Δ and Δ^* respectively, ℓ denotes the number of integer points and ℓ^* denotes the number of integer interior points.

5.2 Genetic Algorithm Model

To generate reflexive polytopes, we apply a genetic algorithm (GA) due to its simplicity and effectiveness. Here we just give the specific details of our GA; namely the environment, fitness function and selection criteria. For a general introduction to GAs, see Section 4.2.

In our case, the environment consists of lattice polytopes Δ in n dimensions, generated as the convex hull of m vectors $x_a \in \mathbb{Z}^n$, where $a = 1, ..., m$. The vectors are arranged into an

$n \times m$ matrix $X = (x_1, \ldots, x_m)$. In practice, we restrict the entries x_a^i of the matrix X to a finite range $[x_{\min}, x_{\max}]$. Our environment E therefore consists of all $n \times m$ integer matrices X with entries in this range. We convert the X matrices into bitstrings. Each integer x_a^i is converted into a bitstring of length ν and concatenating these leads to a bitstring of length $n_{\text{bits}} = n\,m\,\nu$. The environment E therefore contains a total of

$$2^{n_{\text{bits}}} = 2^{nm\nu} \tag{127}$$

states.

The basic fitness function is defined to give a score to a polytope based on how close it is to being reflexive:

$$f(\Delta) = w_1 \left(\text{IP}(\Delta) - 1\right) - \frac{w_2}{|F(\Delta)|} \sum_{\varphi \in F(\Delta)} |d_\Delta(\varphi) - 1| \,, \tag{128}$$

where $\text{IP}(\Delta)$ equals 1 if Δ has the IP property and is 0 otherwise and $d_\Delta(\varphi)$ denotes the distance of the φ facet hyperplane from the origin. The numbers $w_1, w_2 \in \mathbb{R}^{\geq 0}$ are weights which are typically chosen as $w_1 = w_2 = 1$. Note that $f(\Delta) \leq 0$ always and $f(\Delta) = 0$ if and only if Δ is reflexive. Accordingly, we set $f_{\text{term}} = 0$ so that the terminal states correspond to reflexive polytopes.

For some of our applications we are interested in generating reflexive polytopes Δ whose number $N_{\text{p}}(\Delta)$ of points equaled a certain target $N_{\text{p},0}$. To facilitate such targeted searches, we modify the fitness function (128) to

$$\tilde{f}(\Delta) = f(\Delta) - w_3 \left| N_{\text{p}}(\Delta) - N_{\text{p},0} \right| , \tag{129}$$

where $w_3 \in \mathbb{R}^{\geq 0}$ is a further weight which can be used to switch the additional requirement on and off. If $w_3 > 0$, then $\tilde{f}(\Delta) = 0$ and, hence, Δ is terminal, if and only if Δ is reflexive and has the target number $N_{\text{p},0}$ of points.

For the selection criteria we employ the so-called roulette wheel selection where the probability p_k of an individual $s \in P_k$ beign selected for breeding is defined as

$$p_k(s) = \frac{1}{n_{\text{pop}}} \frac{(\alpha - 1)\left(f(s) - \bar{f}\right) + f_{\max} - \bar{f}}{f_{\max} - \bar{f}} \,, \tag{130}$$

where $\bar{f}$ and $f_{\max}$ are the average and maximal fitness values on P_k, respectively. The parameter α, typically chosen in the range $\alpha \in [2, 5]$, indicates by which factor the fittest

individual in the population is more likely to be selected than the average one. We fix $\alpha = 3$ for all our investigations.

For each pair of individuals selected for breeding, a random location $l \in \{1, \ldots, n_{\text{bits}}\}$ along the bitstrings is chosen and the tails of the two strings are swapped, creating two offspring. In the final mutation step, a certain fraction, r_{mut}, of bits in the population $\tilde{P}_{k+1}$ are flipped. Finally, we employ elitism in our GA, carrying the fittest individual from the previous population to the next.

Once the GA has found a list of reflexive polytopes, possibly with additional properties, we are not yet finished, since we have to eliminate the redundancies which arise from transformation equivalences. This is done by computing the normal form of the vertex matrix, using the algorithm described in Appendix A.1.

For our applications, we use PALP [123] tools for polytope computation and additional c code combined into a lightweight and fast c code [124].

5.3 Generating Reflexive Polytopes

In this section we present the results of our GA at generating reflexive polytopes in various dimensions. We start with $n = 2, 3, 4$ dimensions where complete classification already exists. We then progress onto $n = 5$ dimensions where a complete classification is lacking. There are some common hyperparameter choices which we use for all following runs. In each case, we evolve populations for $n_{\text{gen}} = 500$ generations, we use a mutation rate of $r_{\text{mut}} = 0.005$, and the parameter α in (130) is set to $\alpha = 3$. Other hyperparameters, such as the population size n_{pop}, and environmental variables will be chosen to optimize results and their values for each case will be stated below.

Two- and Three-Dimensions

In two-dimensions, where there are 16 unique reflexive polytopes, we use an integer range $x_a^i \in [-4, 4]$ and $m = 6$ generators. Using a population size of $n_{\text{pop}} = 200$, the genetic algorithm finds all 16 reflexive polytopes after only one evolution, taking only a few seconds on a single CPU.

In three-dimensions, we use the coordinate range $x_a^i \in [-8, 8]$ and $m = 14$ generators. With population size set to $n_{\text{pop}} = 450$, the genetic algorithm finds all 4319 reflexive polytopes after 117251 evolutions. Considering the small fraction of reflexive polytopes in the environment this is a remarkable achievement.

Four-Dimensions

In four-dimensions, the total number of reflexive polytopes is large $(473, 800, 776)$ and, for our purpose it is not necessary to generate the complete set. Instead, we focus our attention on finding those polytopes Δ with the lowest numbers, $N_{\text{p},0} = 6, ..., 10$, of points. To facilitate such a search we use the modified fitness function (129) with certain targets $N_{\text{p},0}$ for the number of points. The integer range is taken to be $x_a^i \in [-4, 4]$ in all cases and the number of generators is set as $m = N_{\text{p}} - 1$, since the maximum number of vertices of a reflexive polytope with N_{p} points is $N_{\text{p}} - 1$ where all points except the origin are vertices. With these settings we performed multiple GA runs and the results are summarised in Table 6. It is remarkable that all states are found in all cases after a sufficient number of GA runs.

# points	# refl. poly.	n_{pop}	# GA runs
6	3	400	5
7	25	300	30
8	168	400	60
9	892	300	9378
10	3838	350	9593

Table 6: Results for four-dimensional reflexive polytopes with a small number of lattice points, as in the first column. The total number of unique reflexive polytopes for the given number of points (as taken from the Kreuzer–Skarke list) is given in the second column, the third column gives the population size and the last column gives the number of GA runs required to find all unique reflexive polytopes.

Five-Dimensions

In the previous sections, we saw that GAs can generate complete lists of reflexive polytopes in two, three, and four-dimensions. This is a valuable proof of principle demonstrating that GAs can successfully generate reflexive polytopes. However, the results are of limited practical use,

given the complete classifications in those dimensions. We now turn to reflexive polytopes in five-dimensions, where a complete classification is not available. The total number of (inequivalent) reflexive polytopes in dimensions $n = 1, 2, 3, 4$ is given by 1, 16, 4319, $473,800,776$, respectively. This sequence suggests the number of reflexive polytopes in five dimensions is extremely large and producing a complete catalog is intractable.[9] The partial list of Schöller and Skarke of $185,269,499,015$ weight systems that give rise to (not necessarily inequivalent) maximal five-dimensional reflexive polytopes [88] is currently the largest database.

Unlike the Schöller-Skarke list, the polytopes produced by the GA are not biased towards maximal polytopes, and can be used to search for polytopes with other properties. In fact, it is likely that the vertices of the largest polytopes are far from the origin, and the GA would struggle to find such cases. For this reason, we focus on generating small polytopes with a small number of points. We note that the "small" polytopes we generate are not necessarily dual to Skarke and Scholler's "maximal" polytopes. The maximal polytopes in Skarke and Scholler's list are those that contain every reflexive polytope as a subpolytope and so the dual of these will indeed have a small number of points. However, some of our small polytopes which have a small number of points might have a dual that is large (meaning it has large number of points) but it is still a subpolytope of an even larger maximal polytope, in which case the dual would not be in Skarke and Scholler's database.

In analogy with the four-dimensional case, we search for those five-dimensional reflexive polytopes with the lowest number of points, that is, $N_{\mathrm{p},0} \in \{7, 8, 9, 10, 11\}$. In each case, using the vertex coordinate range $[-4, 4]$, we perform as many GA runs as necessary until no new polytopes are found for 1000 evolutions. The results are summarised in Table 7. Of course we do not know with certainty which fraction of low-point polytopes we have found. It is possible that some five-dimensional reflexive polytopes with such numbers of points still exist inside or outside the defined GA environment. On the other hand, in view of the highly successful low-point searches in four dimensions, it seems likely we have found a large fraction of these polytopes.

The Schöller-Skarke list contains 8 weight systems corresponding to CY hypersurfaces with $h^{1,1} = 1$ and 33 weight systems corresponding to CY hypersurfaces with $h^{1,1} = 2$. In their

[9] Extrapolating this trend gives an estimate of 1.15×10^{18} five-dimensional reflexive polytopes [88].

# points	n_{pop}	# un. refl. poly.	# GA runs
7	350	9	36
8	350	115	1278
9	450	1385	7520
10	750	12661	31857
11	650	94556	376757

Table 7: Results for five-dimensional reflexive polytopes with a small number of lattice points, as in the first column. The population size is given in the second column, the third column lists the total number of unique reflexive polytopes found by the GA for the given number of points and the last column gives the number of GA runs at which the total list of reflexive polytopes saturates and after which no new reflexive polytopes are found for 1000 runs.

list there are also $8,409,140$ and $186,659,154$ weight systems with $h^{1,3} = 1, 2$ respectively, whose duals will have $h^{1,1} = 1, 2$. Taking a sample of $100,000$ weight systems in each case and computing the corresponding reflexive polytopes in their normal forms we find that there are only 7 and 47 unique polytopes. This is a huge reduction and highlights the large amount of redundancy in the Schöller-Skarke list of weight systems. Scanning the lists of five-dimensional reflexive polytopes obtained from the GA runs described above we find 15 polytopes with $h^{1,3} = 1$ whose dual polytopes have $h^{1,1} = 1$. Similarly, we find 195 polytopes with $h^{1,3} = 2$ whose dual polytopes have $h^{1,1} = 2$. By comparing normal forms, we see that these dual polytopes contain all $h^{1,1} = 1, 2$ polytopes from the Schöller-Skarke list. In addition, there are many new examples, with Hodge numbers which are not contained in the Schöller-Skarke list or in the list of four-dimensional CY manifolds realized as complete intersections in products of projective space (CICYs) [87, 125]. Two such examples with a new set of Hodge numbers are:

Example 2 *A new five-dimensional reflexive polytope giving rise to a four-dimensional CY hypersurface with* $h^{1,1} = 1$ *is given by the vertex matrix:*

$$\begin{pmatrix} 0 & 4 & -1 & 2 & -1 & -2 \\ 3 & -1 & -4 & 9 & 0 & -3 \\ 2 & 2 & -2 & 2 & -2 & 2 \\ 1 & -3 & 0 & 3 & 0 & -1 \\ 3 & -5 & 1 & -1 & 1 & -1 \end{pmatrix} . \tag{131}$$

This polytope has Hodge numbers $h^{1,2} = 0$, $h^{1,3} = 111$, $h^{2,2} = 492$, and Euler number $\chi = 720$.

Example 3 *A new five-dimensional reflexive polytope giving rise to a four-dimensional CY hypersurface with $h^{1,1} = 2$ is given by the vertex matrix:*

$$\begin{pmatrix} 0 & -2 & -4 & -1 & 10 & 8 & -2 \\ 4 & 0 & 2 & -1 & -6 & -4 & 0 \\ -1 & 1 & -3 & 0 & 1 & 3 & 1 \\ 2 & -2 & 0 & 0 & -2 & 0 & 1 \\ 2 & 0 & -2 & 1 & -6 & -2 & 3 \end{pmatrix} . \tag{132}$$

This polytope has Hodge numbers $h^{1,2} = 0$, $h^{1,3} = 111$, $h^{2,2} = 496$, and Euler number $\chi = 726$.

Targeted Search

To showcase the capability of our GA at generating CY manifolds with specific criteria, we present an example of a targeted search inspired by [126]. In that paper, the authors consider eleven-dimensional supergravity compactified on CY fourfolds with four-form flux and provide the conditions necessary to break supersymmetry from $\mathcal{N} = 2$ to $\mathcal{N} = 1$. In Appendix A they search for CY fourfolds with Euler number χ divisible by $\delta \in \{24, 224, 504\}$ which satisfy the $\mathcal{N} = 1$ condition. By searching the Schöller-Skarke list they find eight examples which they present in Table 1. To facilitate a GA search for such cases, we modify our fitness function to

$$\tilde{f}(\Delta) = f(\Delta) - w_4 \sum_{\delta} \chi(\Delta) \bmod \delta \,, \tag{133}$$

where w_4 is a weight and $\chi(\Delta)$ is the Euler number of Δ. In our search for such polytopes we set the number of generators to be $m = 10$ and use the integer coordinate range $[-4, 4]$. With population size $n_{\text{pop}} = 550$ and after 10000 evolutions the GA finds 1871 polytopes that satisfy the index condition and, comparing the Hodge numbers with those in Schöller and Skarke's list, we find that two of these are new. One such example is:

Example 4 *A new five-dimensional reflexive polytope giving rise to a four-dimensional CY hypersurface whose Euler number χ is divisible by 24, 224, and 504 is given by the vertex matrix:*

$$\begin{pmatrix} -1 & 0 & 2 & 0 & 2 & -1 & 0 & 1 \\ 0 & -1 & 2 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 3 & 0 & 2 & 2 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} . \tag{134}$$

This polytope has Hodge numbers $h^{1,1} = 331$, $h^{1,2} = 9$, $h^{1,3} = 6$, $h^{2,2} = 1374$, *and Euler number* $\chi = 2016$.

5.4 Outlook

In this section, we have shown how genetic algorithms (GAs) can be efficiently used to generate reflexive polytopes. In two, three and four-dimensions, by comparing with the existing classifications, we showed that the GA was able to generate complete datasets of reflexive polytopes. Due to the large number of reflexive polytopes in four-dimensions we focused only on those with a small number of points N_{p}. For the five-dimensional case, where only a partial classification of reflexive polytopes exists, we performed a GA search for $7 \leq N_{\mathrm{p}} \leq 11$. Our generated datasets included all known polytopes and indeed many more, previously unknown cases. This includes cases which lead to CY fourfolds with new sets of Hodge numbers. While the numbers of reflexive polytopes obtained in this way (see Table 7) might not be the true total they at least provide strong lower bounds.

It is perhaps not desirable, or even feasible, to generate the complete list of reflexive polytopes beyond four dimensions. Instead, in this work we propose an alternative approach, well-suited to the needs of string compactifications, of targeted searches for reflexive polytopes (and their associated CY manifolds) with certain prescribed properties. We demonstrated that GAs can be used for such targeted searches, by looking for cases with certain prescribed values of the Euler number. We expect the same approach will work for other targets, such as a certain desirable pattern of Hodge numbers.

The targeted search example illustrates how one can design a dedicated search for CY manifolds with prescribed properties, for example with certain topological properties satisfying certain constraints. This points to a different approach for dealing with large classes of geometries in string theory. Rather than producing complete lists of such geometries (which is often not even feasible), generative machine learning methods, such as GAs, can be used to search for geometries with prescribed properties, as required for the intended string compactification.

There are many possible directions for future research. In particular, by fine, star, regular triangulation of a (dual) reflexive polytope into simplices, we can resolve non-terminal

singularities and construct the CY hypersurface explicitly. This process is also amenable to attack with GAs and reinforcement learning (RL) as was shown in [121] and [31]. It would be worthwhile to create a combined GA or RL model that can generate reflexive polytopes and triangulations simultaneously. Targeted searches are another promising avenue. For example, it might be possible to design targeted searches which produce string theory compactifications with certain desirable phenomenological properties. It might also be interesting to apply RL to the problem of searching for reflexive polytopes and compare its performance to that of GAs. We leave this to future work.

In this chapter we have considered CY manifolds, including CY threefolds and CY fourfolds which are the appropriate compactification spaces of string theory and F-theory, respectively, that preserve $\mathcal{N} = 1$ supersymmetry in four-dimensions. In the next chapter, we will instead look at G_2 manifolds, which are the analogous compactification space for eleven-dimensional M-theory. We will show how simple neural networks can be used to predict, to a high degree of accuracy, certain topological invariants given input data that describes these manifolds (for a certain construction).

6 Learning G2 Geometry

As we discussed in Section 3.1, the number of spacetime supersymmetries is given by the number of covariantly constant spinors on the compactifying manifold X, and this is in turn related to the holonomy group of X. This was first discovered by Witten in 1981 [77]. The first non-trivial example came a year later in 1982 [127] by compactifying eleven-dimensional supergravity on the squashed S^7 whose G_2 holonomy yields $\mathcal{N} = 1$ supersymmetry in four-dimensions. Today G_2 manifolds continue to play an important role in eleven-dimensional M-theory for the same reason of preserving $\mathcal{N} = 1$ supersymmetry. Compactification on smooth G_2 manifolds does not give chiral fermions, however compactifications on singular G_2 manifolds can yield chiral models in four-dimensional Minkowski space and with realistic gauge groups [128, 129].

A metric with holonomy G_2 can be defined in terms of a torsion-free G_2 structure φ, which is a non-degenerate 3-form. A G_2-structure φ is said to be closed if $d\varphi = 0$ and coclosed if $d\psi = 0$, where $\psi := *_\varphi\varphi$, and the torsion-free condition $T = 0$ is equivalent to φ being both closed and coclosed. A torsion-free G_2 structure φ on a real seven-dimensional manifold M has $\mathrm{Hol}(g_\varphi) = G_2$ if and only if $\pi_1 M$ is finite. Finding torsion-free G_2 structures amounts to solving a difficult non-linear PDE. Instead we can consider just the G_2 structure as a topological residue of the holonomy G_2 metric.

In this section we work with G_2-structures on certain contact Calabi-Yau (cCY) 7-manifolds. These manifolds were introduced by Tomassini and Vezzoni in [130], and consist of Sasakian manifolds endowed with a closed basic complex volume form. It was shown in [131] that such a manifold carries naturally a coclosed (but not necessarily closed) G_2-structure. Despite their unsuitability to M-theory, torsionful G_2 structures retain relevance in the context of (3+7)-dimensional heterotic supergravity with flux, as demonstrated by [132–134]. Indeed, as shown by [135], one can explicitly solve the corresponding Strominger system on cCY 7-manifolds, by way of coclosed G_2-structures together yielding non-trivial scalar and G_2-instanton gauge fields, with constant dilaton, as well as an H-flux with prescribed Chern-Simons defect, in accordance to the anomaly-free condition referred to as the heterotic Bianchi identity.

A special class of cCY manifolds arises from Calabi-Yau (CY) links, which were first

discussed from the perspective of G_2 topology in [136]. In this case a 7-dimensional weighted link K_f is obtained as the intersection of $S^9 \subset \mathbb{C}^5$ with a weighted homogeneous affine variety (defined by the zero locus of the polynomial f) having an isolated singularity at the origin. Milnor showed that such links are 2-connected compact smooth manifolds [137], indeed K_f is the total space of a Hopf S^1-bundle over a weighted projective 3-orbifold in $\mathbb{P}^4(\mathbf{w})$. For appropriate choices of polynomial degree and weighted $\mathbb{C}^*$-action the projective 3-orbifold is a CY threefold and K_f is then called CY link. The dataset of possible weights that admit these CY threefolds consists of the 7555 cases classified in [138]. Therefore, we pursue the construction of a CY link for each of these weight systems, computing the following two types of topological invariants.

Firstly, considering the Sasakian topology of the CY links, we compute the Sasakian-Hodge numbers $\{h^{3,0}, h^{2,1}\}$. These numbers can be obtained as the dimensions of certain linear subspaces of the Milnor algebra $\mathbb{M}_f = \mathbb{C}[[z_1, \ldots, z_5]]/\mathcal{J}_f$, defined by the corresponding Jacobian ideal of f [139]. Secondly, considering their G_2-topology and building upon the calculations first carried out in [136], we compute the Crowley-Nordström (CN) homotopy invariant $\nu(\varphi) \in \mathbb{Z}/48\mathbb{Z}$ associated to the natural G_2 structure induced from the transverse CY structure for all CY links. This invariant was introduced in [140] in an effort to determine when two G_2 structures on a closed 7-manifold are deformation-equivalent, in other words they are related by homotopies and diffeomorphisms.

For CY threefolds defined as hypersurfaces in weighted projective space $\mathbb{P}^4(\mathbf{w})$, the weights $\mathbf{w}$ defining the ambient projective space are sufficient to uniquely determine the CY threefold's Hodge numbers [141]. Therefore, it is no surprise that machine learning (ML) methods perform well at the supervised learning task of predicting the Hodge numbers from the list of weights. However, in the 7-dimensional CY link case no such explicit formula for computing the Sasakian-Hodge numbers from the weights is known, and one would initially expect the specific CY polynomial coefficients chosen to change the topology. However, when we apply supervised ML techniques to learn the Sasakian-Hodge numbers from the weights we achieve surprisingly high accuracy which suggests the existence of a formula. Motivated by this success, we apply ML interpretability techniques to determine an approximate form for this formula. This formula provides new insights into Sasakian structures, as well as being vastly quicker to compute.

The results presented in this section are those published in [50]. The datasets of the weighted CY polynomials used in the link construction, with the computed Sasakian-Hodge numbers and CN invariants, as well as the ML scripts, are available on `GitHub` [142]. The rest of the chapter is organised as follows: in Section 6.1 we introduce contact and Sasakian geometry in order to understand the definition and construction of CY links. In Section 6.2 we explain how we constructed our dataset of CY links and discuss the distribution of their Sasakian hodge numbers and CN invariants. Then in Section 6.3 we present the ML results of predicting topological invariants using neural networks. Finally, in Section 6.4 we finish with a discussion and prospectus.

6.1 Calabi-Yau Links

Contact Geometry

In some sense, contact geometry is the odd-dimensional counterpart of symplectic geometry. Given an n-dimensional smooth manifold M, and a point $p \in M$, a contact element of M with contact point p is an $(n-1)$-dimensional linear subspace of the tangent space to M at p. A contact element can be given by the kernel of a linear function on T_pM. A contact structure on an odd $(2n+1)$-dimensional manifold K is a smooth distribution of contact elements, denoted by F, which is non-integrable. Let F_η be a contact structure, given by a differential 1-form η (called the contact form), i.e. a smooth section of the cotangent bundle. Then the non-integrability condition is given explicitly as:

$$\eta \wedge (d\eta)^n \neq 0, \tag{135}$$

where $(d\eta)^n = \underbrace{d\eta \wedge \cdots \wedge d\eta}_{n\text{-times}}$.

Definition 6.1 *A contact manifold (K, F_η) is a $(2n+1)$-dimensional manifold K equipped with a smooth non-integrable hyperplane field $F_\eta \subset TK$, i.e., locally $F_\eta = Ker(\eta)$, where η is a 1-form which satisfies $\eta \wedge (d\eta)^n \neq 0$.*

The restriction of the 2-form $\omega = d\eta$ to a hyperplane in F_η is a non-degenerate 2-form. This construction provides any contact manifold K with a natural symplectic bundle of rank one smaller than the dimension of K.

Definition 6.2 *Associated with a contact form η one has the Reeb vector field R_η* [10]*, defined by the equalities:*

[10] From this point onwards we will omit η when denoting the Reeb vector field and simply denote it as R.

1. $d\eta(R_\eta, \cdot) = 0$,

2. $\eta(R_\eta) = 1$.

In other words, if η is a contact form for a given contact structure, the Reeb vector field R can be defined as the unique element of the (one-dimensional) kernel of $d\eta$ such that $\eta(R) = 1$.

(Regular) Sasakian Geometry

Just as Kähler geometry is the natural intersection of symplectic and Riemannian geometry, Sasakian geometry is the the natural intersection of contact and Riemannian geometry. One may interpret the odd-dimensional structures of a contact manifold (K^{2n+1}, η) as even-dimensional structures transverse with respect to a S^1-action along the fibres of a submersion $S^1 \to K \to V$. In particular, Sasakian geometry may be seen as transverse Kähler geometry, corresponding to the reduction of the transverse holonomy group to $U(n)$.

A Sasakian manifold [143] is a contact manifold equipped with a special kind of Riemannian metric, called a Sasakian metric. A Sasakian metric is defined using the construction of the Riemannian cone. Given a Riemannian manifold (M, g), its Riemannian cone is the product

$$(M \times \mathbb{R}^{>0}), \tag{136}$$

of M with a half-line $\mathbb{R}^{>0}$, equipped with the cone metric

$$t^2 g + dt^2, \tag{137}$$

where t is the parameter in $\mathbb{R}^{>0}$.

A contact Reimannian manifold (K^{2n+1}, η) is Sasakian, if its Riemannian cone with the cone metric is a Kähler manifold with the Kähler form

$$t^2\, d\eta + 2t\, dt\, \eta. \tag{138}$$

As the cone is by definition Kähler, there exists a complex structure $J \in \mathrm{End}(TK)$. The complex structure J is such that

$$J \circ J = -I_{TK} + \eta \otimes R, \tag{139}$$

and furthermore the Reeb vector field on the Sasakian manifold is given as

$$R = -J\left(t\frac{\partial}{\partial t}\right). \tag{140}$$

Contact Calabi-Yau Manifolds

In the previous subsection we saw that Sasakian geometry may be seen as transverse Kähler geometry, corresponding to the reduction of the transverse holonomy group to $U(n)$. One can also consider Sasakian manifolds with special transverse holonomy $SU(n)$ in which case the manifold is an S^1 bundle over a Calabi-Yau (CY) manifold.

Definition 6.3 *A Sasakian manifold $(K^{2n+1}, \eta, R, J, \Upsilon)$ is said to be a contact Calabi–Yau manifold (cCY) if Υ is a nowhere-vanishing transverse form of horizontal type $(n, 0)$, such that*

$$\Upsilon \wedge \bar{\Upsilon} = (-1)^{\frac{n(n+2)}{2}} \omega^n \quad \textit{and} \quad d\Upsilon = 0, \quad \textit{with} \quad \omega = d\eta. \tag{141}$$

In the special case where $n = 3$, a cCY structure naturally induces a coclosed G_2-structure [131]:

Proposition 6.1 *Every cCY manifold $(K^7, \eta, R, J, \Upsilon)$ is an S^1-bundle $\pi : K \to V$ over a CY threefold (V, ω, Υ), with connection 1-form η and curvature*

$$d\eta = \omega, \tag{142}$$

and it carries a cocalibrated G_2-structure

$$\varphi := \eta \wedge \omega + Re\Upsilon, \tag{143}$$

with torsion $d\varphi = \omega \wedge \omega$ and Hodge dual 4-form

$$\psi = *\varphi = \frac{1}{2}\omega \wedge \omega + \eta \wedge Im\Upsilon.$$

Calabi-Yau Links

We turn our attention to a special kind of cCY manifold, called CY links which admit a G_2 structure. We begin first by defining the notion of a hypersurface link of an isolated singularity. We denote by $\bar{B}_\epsilon$ the closed ball of radius ϵ centered at the origin of $\mathbb{C}^{n+1}$, by $S^{2n+1}_\epsilon = \partial B_\epsilon(0)$ the boundary of this ball, and B_ϵ the corresponding open ball. Let $f : \mathbb{C}^{n+1} \to \mathbb{C}$ be a complex analytic map with $f(0) = 0$ and denote $\mathcal{V} := f^{-1}(0)$ and $K_f := \mathcal{V} \cap S^{2n+1}_\epsilon$ (Figure 15).

Now we focus on the particular case in which f is a weighted homogeneous polynomial with an isolated singularity at $0 \in \mathbb{C}^{n+1}$. In this case $\mathcal{V}$ intersects transversally every sphere S^{2n+1}_ϵ around the origin.

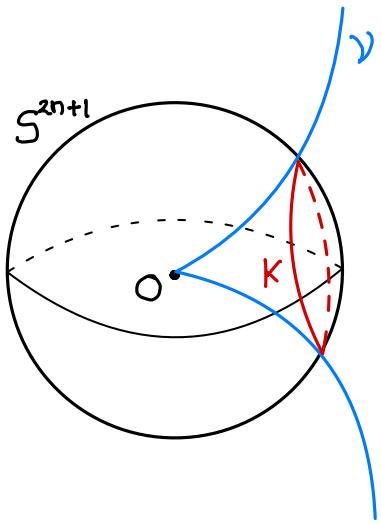


Figure 15: Link K of a hypersurface $\mathcal{V}$ in $\mathbb{C}^{n+1}$.

Definition 6.4 *A polynomial $f(z_0, ..., z_n)$ is called a weighted homogeneous polynomial of degree d and with weights $\boldsymbol{w} = (w_0, ..., w_n) \in \mathbb{Z}^{n+1}_{>0}$ if for any $\lambda \in \mathbb{C}^*$*

$$f(\lambda^{w_0} z_0, ..., \lambda^{w_n} z_n) = \lambda^d f(z_0, ..., z_n), \tag{144}$$

Such an f defines a non-singular affine hypersurface

$$\tilde{\mathcal{V}} := \{z \in \mathbb{C}^{n+1} | f(z) = 0\}. \tag{145}$$

which, in general, admits a singularity at the origin. Therefore, weighted homogeneous polynomials give rise to links, fibering by circles over weighted projective hypersurfaces:

Definition 6.5 *Let $f : \mathbb{C}^{n+1} \to \mathbb{C}$ be a $\boldsymbol{w}$-weighted homogeneous polynomial with an isolated critical point at 0. Then $K_f := \mathcal{V} \cap S^{2n+1}$ is called a weighted link of degree $\deg f$ and weight $\boldsymbol{w}$.*

We have the commutative diagram, where the horizontal arrows are Sasakian and Kählerian embeddings and the vertical arrows are principal S^1-orbibundles[11] and orbifold Riemannian submersions. As a complex orbifold, the hypersurface $V \subset \mathbb{P}^4(\mathbf{w})$ can be represented as the quotient $(\mathcal{V} - 0)/\mathbb{C}^*(\mathbf{w})$ where $\mathcal{V} = f^{-1}(0)$

$$\begin{array}{ccc} K & \to & S^9 \\ \downarrow & & \downarrow \\ V & \to & \mathbb{P}^4(\mathbf{w}) \end{array} \tag{146}$$

[11] Roughly speaking an orbibundle over an orbifold is generalisation of a fiber bundle over a manifold, where the structure group of the total space is required to be the same as the base orbifold.

Definition 6.6 *A weighted link K_f of degree d and weights $\boldsymbol{w} = (w_0, ..., w_n)$ is said to be a CY link if*

$$d = \sum_{i=0}^{n} w_i. \tag{147}$$

The condition $d - \sum_{i=0}^{n} w_i = 0$ means that the weighted projective V is a CY orbifold, thus CY links are nontrivial circle fibrations over CY 3-orbifolds.

Proposition 6.2 *Every CY link admits an S^1-invariant contact Calabi-Yau structure [131].*

From 6.2 and 6.1 we see that every CY link has a G_2-structure.

6.2 Data Generation and Analysis

The data for a CY link as defined in the previous section is given by a set of weights $\mathbf{w} = (w_0, ..., w_4)$ defining the ambient complex weighted projective space $\mathbb{P}^4(\mathbf{w})$ and a polynomial f defining the CY threefold hypersurface. The list of all 7555 weight vectors $\mathbf{w}$ which lead to unique weighted projective spaces $\mathbb{P}^4(\mathbf{w})$ whose anticanonical divisors are compact and Ricci-flat was classified in [138].

For each $\mathbb{P}^4(\mathbf{w})$ in this list, any Calabi-Yau hypersurface in the anticanonical divisor class can be represented as a weighted homogeneous polynomial f of degree $\sum_i w_i$. Within this class there is freedom in the choice of complex coefficients for each of the monomial terms in the defining equation. There are however sources of redundancy in the choice of coefficients, such coordinate transformations, coefficient normalisation, etc., allowing multiple sets of coefficients to define the same threefold. All CY hypersurfaces in $\mathbb{P}^4(\mathbf{w})$ will share the same Hodge numbers, but may otherwise be topologically distinct.

The dataset of CY links considered in this work were constructed using one CY from each of the respective 7555 $\mathbb{P}^4(\mathbf{w})$'s. In each case, the CY polynomial was first selected to have all monomial coefficients as 1. However 1484 out of the 7555 polynomial hypersurfaces intersected with singularities in the ambient space, leading to a higher-dimensional singularity structure on the links. To avoid this, for these 1484 cases other polynomials were sampled, with coefficients from $\{1, 2, 3, 4, 5\}$, until the singularity structure was exclusively the isolated singularity at the origin - as required for the link construction.

To understand this process, let's consider the example $\mathbb{P}^4(\mathbf{w})$ where $\mathbf{w}$ is the weight vector $\mathbf{w} = (22, 29, 49, 50, 75)$, whose degree is $d = 225$ $(= \sum_i w_i)$ and whose monomial basis has the following 7 terms:

$$z_1^8 z_3,\ z_1^4 z_2^3 z_3,\ z_1 z_2^7,\ z_1 z_2 z_3 z_4 z_5,\ z_2 z_3^4,\ z_4^3 z_5,\ z_5^3.$$

The CY polynomial equation is:

$$0 = a_1 z_1^8 z_3 + a_2 z_1^4 z_2^3 z_3 + a_3 z_1 z_2^7 + a_4 z_1 z_2 z_3 z_4 z_5 + a_5 z_2 z_3^4 + a_6 z_4^3 z_5 + a_7 z_5^3,$$

for the complex coefficient vector $(a_1, \ldots, a_7)$. We initialise the coefficients $a_1 = \cdots = a_7 = 1$, and check the singularity structure of the resulting hypersurface. In this case, the singular locus defined by this polynomial has dimension 0, which is the isolated singularity at the origin, there is hence no further singularity structure introduced. We therefore accept this CY threefold, adding it to our database (no further sampling of the a_i values is required).

For each of the 7555 CYs selected in this way, the topological properties of the corresponding links were calculated. Namely, the Sasakian Hodge numbers $\{h^{3,0}, h^{2,1}\}$, and the CN invariant. The polynomial generation and topological invariant computations were performed in `sagemath` [144], with the help of `macaulay2` [145] and `singular` [146]. Computation of each of the topological invariants required the respective Gröbner bases of the CY polynomials; these bases are notoriously expensive to compute, hence, as a side product of these computational efforts, the Gröbner basis for a selection of the Calabi-Yau polynomials considered (one for each possible weight vector) is provided, along with the corresponding topological quantities, on this work's respective `GitHub`.

Sasakian Hodge Numbers

The following theorem from [147, 148] allows us to compute certain mixed Hodge numbers $h^{p,q}(K_f)$ of a CY link K_f from the Milnor algebra.

Theorem 6.3 *Let f be a $\mathbf{w}$-weighted homogeneous polynomial in $\mathbb{C}^n$ of degree d. Given $p+q = n$, let $\ell = (p+1)d - \sum_i w_i$, and denote by $(\mathbb{M}_f)_\ell$ the linear subspace of the Milnor algebra consisting of degree ℓ elements.*

$$h^{p,q}(K_f) = \dim_{\mathbb{C}}(\mathbb{M}_f)_\ell. \tag{148}$$

When K_f is a Calabi-Yau link, the condition reduces to $\ell = pd$.

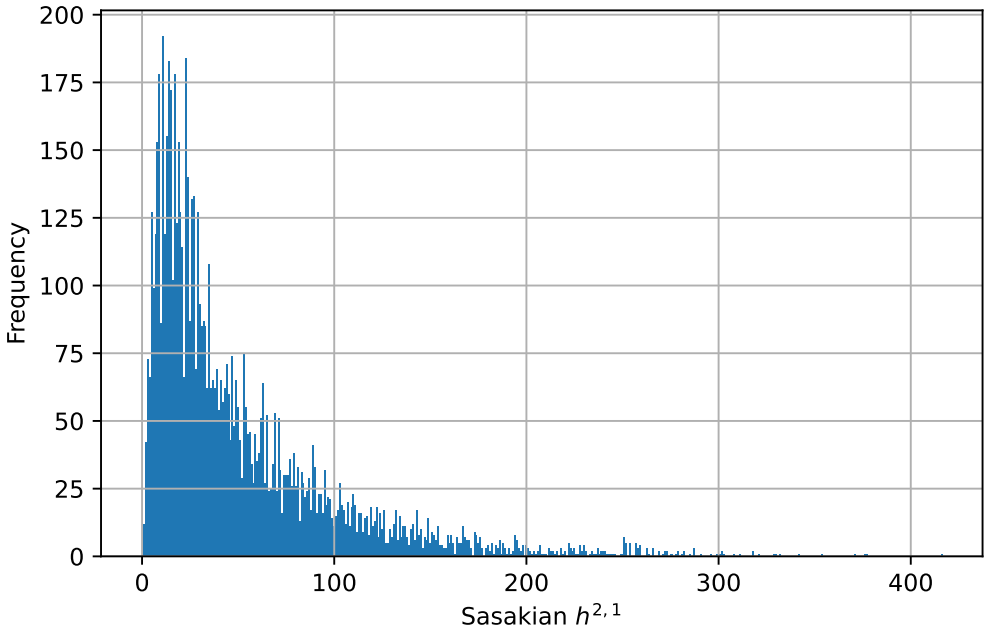


Figure 16: Histogram of Sasakian $h^{2,1}$ values for the 7555 CY link constructions computed.

The computation of the Hodge numbers $h^{3,0}$ and $h^{2,1}$ associated with each weighted homogeneous polynomial was carried out by following the algorithmic implementation of the explicit formula in Theorem 6.3, from [139]:

Algorithm 1 Computation of Sasakian Hodge Numbers.

$f(z_0, ..., z_4) \leftarrow$ a homogeneous polynomial in $\mathbb{C}^5$.

$\mathbf{w} = (w_0, \ldots, w_4) \leftarrow$ the weight vector associated with the polynomial f.

$A \leftarrow \mathbb{C}[z_0, \ldots, z_4]$

$d \leftarrow \deg(f)$

$J_f \leftarrow \left\langle \frac{\partial f}{\partial z_0}, \ldots, \frac{\partial f}{\partial z_4} \right\rangle$

$K \leftarrow \textsc{GröbnerBasis}\left(\frac{A}{J_f}\right)$

$h^{3,0} = \#\{x \in K : \deg(x) = 4d - \sum_i w_i\}$

$h^{2,1} = \#\{x \in K : \deg(x) = 3d - \sum_i w_i\}$

The Sasakian $h^{3,0}$ values for the 7555 CY links computed all take value 1, matching the value known for all Calabi-Yau threefolds, which corresponds to the unique holomorphic volume form. The Sasakian $h^{2,1}$ values range from 1 to 416, their frequency distribution is shown in Figure 16.

To compare directly the CY Hodge numbers with the Sasakian Hodge numbers of the links

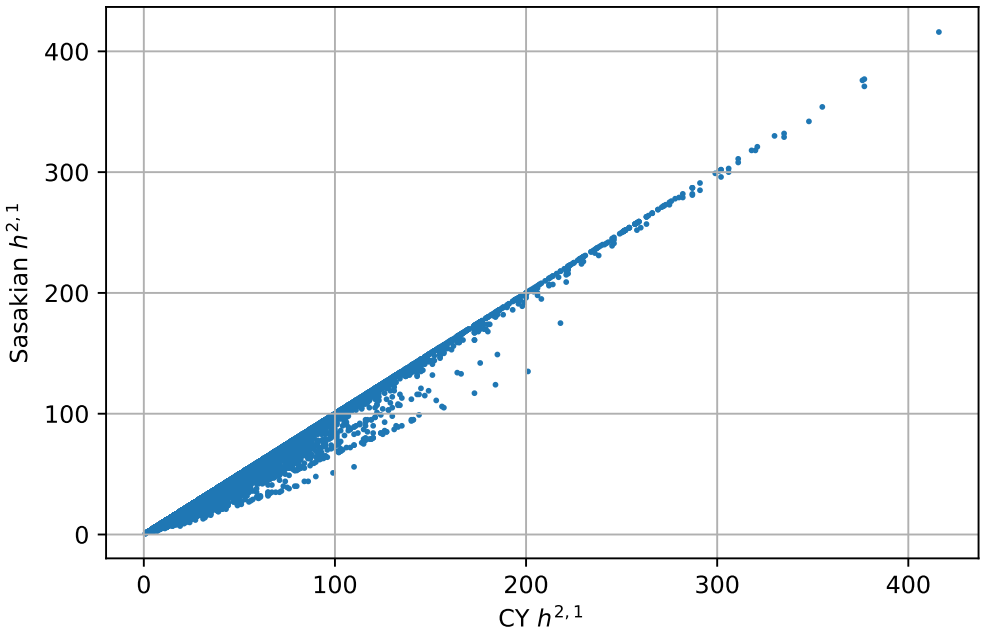


Figure 17: Scatter plot of the CY threefold $h^{2,1}$ values against the Sasakian transverse $h^{2,1}$ values for the 7555 CY links.

built from the same polynomials, a cross-plot of the respective $h^{2,1}$ values is given in Figure 17. This plot shows that these topological invariants are strongly correlated (PMCC ~ 0.99), and the Sasakian Hodge number is bounded above by the CY Hodge number – suggesting the following mathematical conjecture:

Conjecture 6.4 *The Sasakian Hodge number $h^{2,1}_S$ for a Calabi-Yau link is bounded above by the Hodge number $h^{2,1}_{CY}$ of the Calabi-Yau threefold built from the same* $\mathbf{w}$*-homogeneous polynomial:*

$$h^{2,1}_S \leq h^{2,1}_{CY} \ . \tag{149}$$

Crowley-Nordström Invariant

For an arbitrary closed 7-manifold with G_2-structure (Y^7, φ), Crowley and Nordström (CN) [140] have defined a $\mathbb{Z}/48\mathbb{Z}$-valued homotopy invariant $\nu(\varphi)$, which is a combination of topological data from a compact coboundary 8-manifold with Spin(7)-structure (W^8, Ψ) extending (Y^7, φ), in the sense that $Y = \partial W$ and $\Psi\,|_Y = \varphi$:

$$\nu(\varphi) := \chi(W) - 3\sigma(W) \bmod 48, \tag{150}$$

where χ the real Euler characteristic and σ is the signature.

A cCY structure naturally induces a G_2-structure. We therefore study the CN invariants associated to the CY links K_f. To compute ν for the CY links in our dataset, we modify a procedure developed and described in [136] which utilises Steenbrink's Signature theorem. Let

$$\{z^\alpha : \alpha = (\alpha_0, ..., \alpha_n) \in I \subset \mathbb{Z}_{>0}^{n+1}\} \tag{151}$$

be a set of monomials in $\mathbb{C}[z_0, ..., z_n]$ representing a basis over $\mathbb{C}$ for its Milnor algebra $\mathbb{M}_f = \frac{\mathbb{C}[[z_0,...,z_n]]}{\partial f/\partial z_0,...,\partial f/\partial z_n}$. For each $\alpha \in I$, define

$$l(\alpha) := \sum_{i=0}^{n} (\alpha_i + 1)\frac{w_i}{d}. \tag{152}$$

The CN invariant of a link can then be computed in terms of its degree and weights, along with the signature (μ_-, μ_0, μ_+) of the intersection form on $H^4(V_f, \mathbb{R})$:

$$\nu(\varphi) = \left(\frac{d}{w_1} - 1\right) ... \left(\frac{d}{w_5} - 1\right) - 3(\mu_+ - \mu_-) + 1, \tag{153}$$

Where the signature is computed as follows [147]:

$$\begin{aligned}
\mu_+ &= |\{\beta \in I : l(\beta) \notin \mathbb{Z}, \lfloor l(\beta) \rfloor \in 2\mathbb{Z}\}| \\
\mu_- &= |\{\beta \in I : l(\beta) \notin \mathbb{Z}, \lfloor l(\beta) \rfloor \notin 2\mathbb{Z}\}| \\
\mu_0 &= |\{\beta \in I : l(\beta) \in \mathbb{Z}\}|
\end{aligned}$$

We computed the CN invariant for all 7555 CY links. Their frequency distribution is shown in Figure 18, which exhibits an unexpected periodicity of 12 in the most populous invariant values (~ 500). The CN invariants computed fully span the range of possible values, which are odd integers from 1 to 47, cf. [136, Proposition 3.2]. In particular, we note the occurrence of CN invariants 27 and 35, where previous work had not identified examples in these topological classes. Below we provide an explicit example of a Calabi-Yau polynomial that leads to a link in each of these classes:

$$\begin{aligned}
&\nu : 27 \\
&\mathbf{w} : (22, 29, 49, 50, 75) \\
&f : z_1^8 z_3 + z_1^4 z_2^3 z_3 + z_1 z_2^7 + z_1 z_2 z_3 z_4 z_5 + z_2 z_3^4 + z_4^3 z_5 + z_5^3 = 0
\end{aligned} \tag{154}$$

$$\begin{aligned}
&\nu : 35 \\
&\mathbf{w} : (31, 35, 36, 42, 108) \\
&f : z_1^7 z_2 + z_1^2 z_2^2 z_3 z_4^2 + z_1 z_2 z_3^4 z_4 + z_1 z_2 z_3 z_4 z_5 + z_2^6 z_4 + z_3^7 \\
&\quad + z_3^4 z_5 + z_3 z_5^2 + z_4^6 = 0
\end{aligned} \tag{155}$$

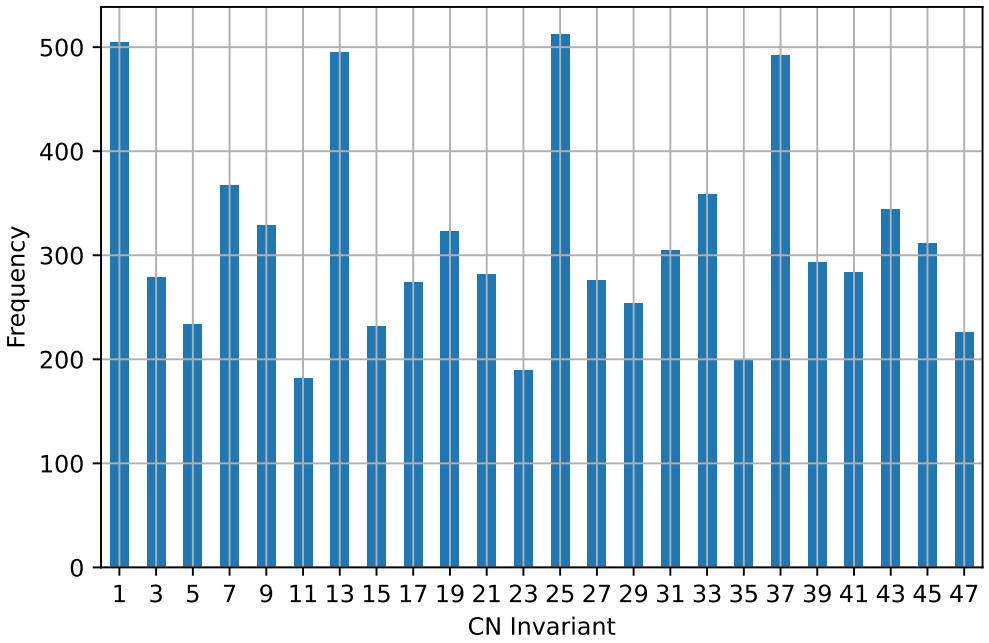


Figure 18: Histogram of CN invariants for the 7549 Calabi-Yau link constructions computed.

6.3 Predicting Topological Invariants of Calabi-Yau Links

Using the dataset of weights and topological invariants generated from the previous section we apply neural networks (NNs) to predict, from an input of weight vectors, the topological invariants (Sasakian Hodge number $h^{2,1}$ and the CN invariant) of CY links. Since the output invariants take a large range of values in each case, this is a regression style problem. The NNs used had the same architecture in each case. They had neuron layer sizes of $(16, 32, 16)$, ReLU activation, and were trained on a MSE loss using an Adam optimiser. These layer sizes and the other hyperparameters were set after some heuristic tuning. Each NN hence amounts to a map of the form:

$$\mathbb{R}^5 \xrightarrow{f_1} \mathbb{R}^{16} \xrightarrow{f_2} \mathbb{R}^{32} \xrightarrow{f_3} \mathbb{R}^{16} \xrightarrow{f_4} \mathbb{R}^1 \ , \tag{156}$$

such that each f_i acts via linear then non-linear action as $f(\mathbf{x}) = \text{ReLU}(\mathbf{W}\cdot\mathbf{x}+\mathbf{b})$. In each case, the NNs were trained using cross-validation on 5 different partitions of the dataset into 80:20 train:test splits, to provide statistical error on the metrics used to assess learning performance.

Crowley-Nordström Invariant

After training the network to predict the CN invariant from the weight vectors, the network performance on the test set was not good, giving an R^2 value of ~ 0.004. Even reducing the problem to a binary classification between the two most populous classes ($\nu = 1$ and $\nu = 25$) did not lead to accuracies much above 0.5, indicating no significant learning and highlighting

the highly non-trivial dependency of this invariant on the input polynomial and weight data.

Sasakian Hodge Numbers

Using the same NN architecture, the network was trained to predict the $h^{2,1}$ values for the CY links. The performance scores of the network on the test set after training were:

$$\begin{aligned} R^2 &= 0.969 \pm 0.003\,, \\ \text{MAE} &= \quad 5.53 \pm 0.22\,, \\ \text{Accuracy} &= 0.967 \pm 0.004\,. \end{aligned} \tag{157}$$

These results are very impressive and suggest that there exists a relation between the Sasakian Hodge numbers and the weights used to define the CY links, similar to the Poincaré polynomial for CY threefolds [149]. We now see whether we can approximate such a formula using symbolic regression (SR) which would allow one to bypass the Milnor algebra Gröbner basis computation.

The particular SR method we applied was the genetic algorithm tree-based SR method `PySR` developed by Cranmer [150]. We use the library of binary functions $\{+, -, *, /\}$ and the absolute distance loss function. We train the model over 100000 iterations with 18 populations of size 33, where each population is mutated 500 times per iteration.

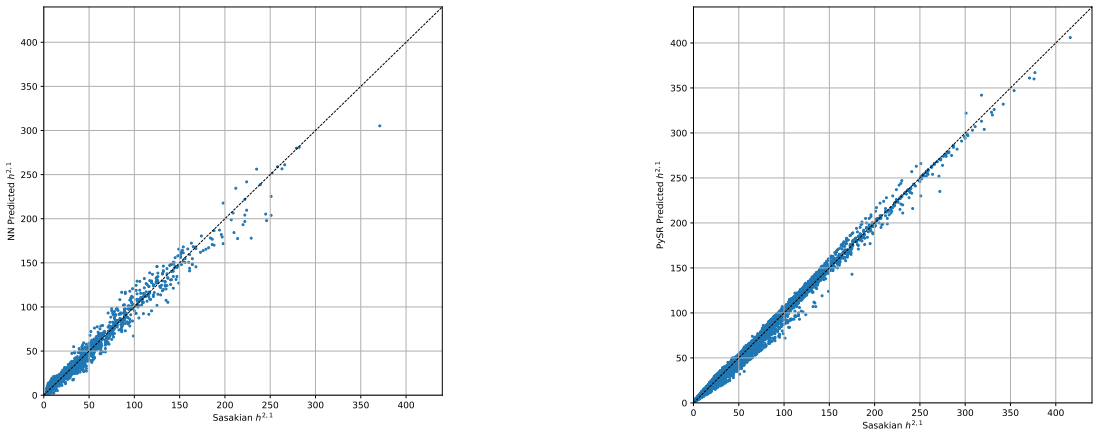
Again we use cross-validation on 5 different partitions of the dataset into 80:20 train:test splits. The (highest performing) equation on the independent test data to model Sasakian $h^{2,1}$ is presented in (158), achieving $R^2 \approx 0.99$ and MAE ≈ 2.6, exceeding the scores of the NN in (157).

$$h^{2,1}_{\texttt{PySR}}(w_0, \ldots, w_4) = \frac{14.91 w_1 \left(w_0 w_4 + w_3 \left(w_0 + w_3\right)\right)}{w_0 w_1 w_2 w_3} + \frac{10.02 w_2 w_3 \left(w_0 + w_4 + 0.77\right)}{w_0 w_1 w_2 w_3}\,, \tag{158}$$

The $h^{2,1}_S$ predicted values from (158) for each of the CY links are plotted against the true values in Figure 19. Additionally the plot shows the equivalent predictions from the trained NN.

6.4 Outlook

In this work, real 7-dimensional Calabi-Yau (CY) links were constructed from CY threefolds coming from the 7555 complex four-dimensional weighted projective spaces that admit them. The datasets of these invariants were statistically analysed, and neural networks were used successfully to predict the respective Sasakian $h^{2,1}$ values from the ambient $\mathbb{P}^4(\mathbf{w})$ weights



(a) Neural Network (b) Symbolic Regression

Figure 19: ML predictions of the $h^{2,1}_S$ values, against the true values, for the 7555 CY links from: (a) a trained NN; and (b) the symbolic regression best model of equation (158).

alone. By applying symbolic regression techniques we were able to achieve an approximate formula ((158)) for the $h^{2,1}$ in terms of the weights $\mathbf{w}$, which provides a good starting point for finding an exact formula (if one exists) in future work. As CY links posses a G_2 structure we also used neural networks to predict the Crowley-Nordström ν invariant [140] but unfortunately we did not achieve the same success. The exhaustive list of CY link data, as well as the `python` scripts used for their analysis and machine learning (ML) are made available at this work's corresponding `GitHub`.

Until [50] there had not been applications of ML to study G_2-geometry. This is no doubt due to the scarcity of databases for G_2-manifolds, which in turn reflects the difficulty in describing G_2-manifolds systematically in terms of algebraic discrete data. In this work we have only looked at a small number of topological properties of a very special kind of G_2 manifold. There are many more avenues for future research within the realm of "ML for G_2". An obvious next step is to construct more general CY links made through general toric varieties rather than weighted projective spaces and repeat the same experiments. With more data we hope, using symbolic regression, to obtain a more accurate formula for $h^{2,1}$ which may lead us to finding an exact formula.

Torsionful G_2 structures, such as the CY links studied here, have applications in heterotic systems, however in the context of M-theory we are interested in torsion-free G_2-structures

since these structures induce metrics with holonomy G_2. There has been a great deal of work in the past few years using ML methods to approximate Ricci-flat CY metrics [40–48] including modelling Ricci-flow [151]. As yet, no similar work has been done using ML to find Ricci-flat G_2 metrics. Like Ricci-flow for CY manifolds, whose critical point gives a Ricci-flat CY metric, there exist flows of G_2-structures, such as the Laplacian flow and coflow, whose critical points give torsion-free G_2 structures. Such flows on contact CY manifolds were recently studied in [152]. In future work we hope to apply ML methods to model flows of G_2-structures and find approximate Ricci-flat holonomy G_2 metrics.

In this chapter and the previous one we considered CY and G_2 manifolds from a purely mathematical point of view, ignoring the physics. As we learnt in Section 2, the geometry of these manifolds influence the physics of the low energy effective field theory that arise when they are chosen as the compactification space. In the next chapter, we will attempt to compute aspects of the effective field theory, namely the Yukawa couplings, that arise when compactifying the $E_8 \times E_8$ heterotic string on a CY threefold, together with a holomorphic vector bundle.

7 Computing Yukawa Couplings in Heterotic String Theory

Several attempts at connecting string theory and particle physics have been made since the first string revolution in the mid 80s. All of these approaches come with their own set of difficulties. The earliest and arguably one of the most promising proposals is the $E_8 \times E_8$ heterotic string compactified on a smooth Calabi-Yau threefold X, equipped with a Ricci-flat metric, together with a vector bundle V whose connection solves the hermitian Yang–Mills equation [8]. Upon compactifying on X, one finds a four-dimensional effective theory with $\mathcal{N} = 1$ supersymmetry governed by a Kähler potential and a superpotential. By judicious choices of the threefold and the vector bundle, one can find minimal supersymmetric Standard Model (MSSM) like models incorporating a variety of desirable features. In principle, the masses and couplings in these models can be computed directly from the geometry of X and V. However, even after decades of work, we are still unable to compute these numbers for all but the simplest examples. A substantial part of the difficulty can be attributed to the lack of explicit expressions for non-trivial Calabi-Yau metrics and hermitian Yang-Mills connections.

Some general features can be inferred from topological or algebraic data of the threefold X and bundle V. For example, we saw in Section 5 that the Euler number of X is related to the number of generations of elementary particles. However, the full details of the resulting four-dimensional physics are determined by a Kähler potential and a superpotential, both of which depend on the metric on X and connection on V. Without this data, it is generally not possible to compute masses or couplings, thereby preventing us from making precise particle physics predictions using string theory.

With new approximations of Ricci-flat metrics and hermitian Yang-Mills connection coming from both numerical [34–39, 153–156] and machine learning methods [40–48, 157], we are are a stage now where we can begin to compute the resulting particle physics. Among many possible applications, the computation carried out in [51], which we discuss in this section, is that of physical Yukawa couplings.

In order to derive the Yukawa couplings of the four-dimensional effective theory that descends from the $E_8 \times E_8$ heterotic string on a Calabi-Yau threefold X admitting a bundle V, one has to carry out the following steps:

1. Calculate the Calabi-Yau metric on X for a particular point in both complex and Kähler moduli space.

2. Calculate the hermitian Yang–Mills connection on V.

3. Calculate the zero modes of a certain twisted Dirac operator. Since X is Kähler, this is equivalent to finding bundle-valued differential forms which are harmonic with respect to the Dolbeault Laplacian $\Delta_{\bar{\partial}_V}$ associated to the twisted Dolbeault differential $\bar{\partial}_V$ [8, 158, 159].

4. Find an orthonormal basis for the harmonic modes.

5. Calculate the physical Yukawa couplings from integrals of wedge products of the normalised harmonic modes.

We will focus our attention on step three. Following on from [160, 161] we give the first numerical calculation of the spectrum and eigenmodes of the Laplacian acting on bundle-valued forms on a Calabi-Yau threefold. Specifically, we compute the approximate spectrum and eigenmodes of the Dolbeault Laplacian acting on bundle-valued scalars ($(0,0)$-forms) and $(0,1)$-forms. We restrict our attention to line bundles over Calabi-Yau n-folds constructed as hypersurfaces in a single ambient projective space. With the eigenmodes in hand, we are able to compute an orthonormal basis of harmonic modes and thus the correctly normalised physical Yukawa couplings. Unfortunately, the examples we consider are too simple to admit non-vanishing matter-field Yukawa couplings. However, this calculation serves as a proof of concept and represents a significant step towards calculating a Yukawa coupling in a physically relevant compactification.

The rest of the chapter is organised as follows: in Section 7.1 we begin by outlining the eigenvalue problem we wish to solve and describing how one can convert the eigenvalue problem into one of finite-dimensional linear algebra. Then in Section 7.2 we apply our numerical method to the toy example of line bundles over complex projective space. We present the known analytic results for the spectrum and then compare the exact and numerical results. In Section 7.3 we consider the simplest Calabi-Yau hypersurface, namely the flat torus described by a cubic equation in $\mathbb{P}^2$. Again, we present the known exact results for the spectrum and then compare these to results for the bundle-valued scalar and $(0,1)$-form spectra computed numerically. Finally in Section 7.4 we apply our numerical method to compute the spectrum for a line bundle over the Fermat quintic threefold. Though there are no analytic results to

compare with, we carry out a number of consistency checks that the spectrum should satisfy. In Section 7.5 we finish with a discussion and prospectus.

7.1 Eigenmodes of the Dolbeault Laplacian

Minimal $\mathcal{N} = 1$ supersymmetric compactifications of $E_8 \times E_8$ heterotic string theory without three-form flux are specified by a choice of Calabi-Yau threefold (X, g), where g is a Ricci-flat Kähler metric, and a principal G-bundle with $G \subset E_8 \times E_8$, whose curvature F satisfies the hermitian Yang-Mills equations (16). A solution to (16) is equivalent to the bundle V being holomorphic and admitting a Hermitian metric on its fibres which is Hermite-Einstein.

Certain $(0,1)$-form sheaf cohomologies count the number of four-dimensional matter fields

$$H^1(X, V) \cong H^{0,1}_{\bar{\partial}_V}(X, V). \tag{159}$$

These cohomologies are spanned by harmonic modes of the Dolbeault Laplacian $\Delta_{\bar{\partial}_V}$. We therefore wish to find the spectrum and eigenmodes of the Dolbeault Laplacian $\Delta_{\bar{\partial}_V}$ acting on (p, q)-forms valued in a vector bundle V. The particular examples we consider are those where the bundle V is a line bundle over a compact Kähler manifold X. Furthermore, we will focus on computing the $(0,0)$- and $(0,1)$-form spectra. The spectrum of bundle-valued scalars will be useful for comparing with known results when X is a projective space or a torus, while the $(0,1)$-form spectrum is what one needs to compute Yukawa couplings. For more details on the Dolbeault Laplacian acting on holomorphic bundle-valued (p, q)-forms, refer to Section 3.4.

The eigenmodes $\phi \in \Omega^{p,q}(V)$ and eigenvalues λ are defined by

$$\Delta_{\bar{\partial}_V} \phi = \lambda \phi, \tag{160}$$

where the eigenvalues λ are real and non-negative. The eigenmodes with zero eigenvalue, $\lambda = 0$, are the "harmonic" or "zero modes" which span $H^{p,q}_{\bar{\partial}_V}(X, V)$. Since X is assumed to be compact, the eigenvalues are discrete and have finite degeneracies. As we will see in examples, if the Kähler metric g on X admits either continuous or discrete symmetries, there may be multiple eigenmodes with the same eigenvalue. We will denote the n-th eigenvalue by λ_n with multiplicity ℓ_n starting from $n = 0$. Note that λ_0 always labels the smallest eigenvalue of $\Delta_{\bar{\partial}_V}$ even when λ_0 is not zero. As usual, the eigenvalues scale with the volume of X as $\lambda \sim \mathrm{Vol}(X)^{-2/d}$. We always normalise the volume of X to one in the examples that follow.

Let us make a few comments on the expected structure of the spectrum of $\Delta_{\bar{\partial}_V}$. First, Serre duality implies

$$h^{p,q}(V) = h^{n-p,n-q}(V^*), \tag{161}$$

so that the counting of zero modes of $\Delta_{\bar{\partial}_V}$ acting on $\Omega^{p,q}(V)$ and $\Omega^{n-p,n-q}(V^*)$ should agree. In fact, since the Hodge star with conjugation $\bar{\star}_V$ commutes with the Laplacian, $\bar{\star}_V \Delta_{\bar{\partial}_V} = \Delta_{\bar{\partial}_{V^*}} \bar{\star}_V$, there is a relation between the entire tower of eigenmodes and eigenvalues. Denoting the set of V-valued (p,q)-form eigenmodes by $\{\phi\}_V^{p,q}$ and the corresponding eigenvalues as $\{\lambda\}_V^{p,q}$, one has

$$\begin{aligned} \bar{\star}_V \{\phi\}_V^{p,q} &= \{\phi\}_{V^*}^{n-p,n-q}, \\ \{\lambda\}_V^{p,q} &= \{\lambda\}_{V^*}^{n-p,n-q}. \end{aligned} \tag{162}$$

Moreover, for $(0,q)$-forms, one can write this in terms of the canonical bundle K_X of X as $\{\lambda\}_V^{0,q} = \{\lambda\}_{K_X \otimes V^*}^{0,n-q}$. We will use these relations as a non-trivial check on the numerical spectra later.

For fixed (p,q), let $\{\alpha_A\}$ be a basis for the vector space of complex-valued (p,q)-forms valued in V on the manifold. This basis is infinite-dimensional, $A = 1, \ldots, \infty$, as we want to be able to express any element of $\Omega^{p,q}(V)$ as a linear combination of the basis with constant coefficients.[12] The basis is not assumed to be orthonormal; the inner product (69) defines a matrix O_{AB} as

$$O_{AB} \equiv \langle \alpha_A, \alpha_B \rangle = \int_X \bar{\star}_V \alpha_A \wedge \alpha_B, \tag{163}$$

which captures the non-orthonormality. Similarly, the matrix elements of $\Delta_{\bar{\partial}_V}$ with respect to this basis are

$$\Delta_{AB} \equiv \langle \alpha_A, \Delta_{\bar{\partial}_V} \alpha_B \rangle. \tag{164}$$

The eigenvalue equation (160) can then be written in terms of the matrix elements as

$$\Delta_{AB} \phi_B = \lambda \, O_{AB} \phi_B, \tag{165}$$

where $\phi = \phi_C \alpha_C$. This is then a generalised eigenvalue problem for (λ, ϕ_A), albeit an infinite-dimensional one. Upon truncating $\{\alpha_A\}$ to finite-dimensional basis, one is left with a standard linear algebra problem to determine the eigenvalues λ and the eigenvectors ϕ_A, which in turn give the spectrum of $\Delta_{\bar{\partial}_V}$ and the expansion of the eigenmodes in terms of the truncated basis. Of course truncating to a finite basis gives only an approximation of the spectrum and eigenmodes, with the dimension of the basis controlling the accuracy of the approximation.

[12] Recall that $\Omega^{p,q}(V)$ restricted to a point $x \in X$ is a finite-dimensional $\mathbb{C}$-vector space. If one does not restrict to a point but instead wants to describe the space of forms over the entire manifold, $\Omega^{p,q}(V)$ is an infinite-dimensional $\mathbb{C}$-vector space.

7.2 The spectrum of $\Delta_{\bar{\partial}_V}$ on $\mathbb{P}^3$

We begin with a study of three-dimensional complex projective space $\mathbb{P}^3$ equipped with the Fubini-Study (FS) metric. Since much is known explicitly about projective space, this will provide an arena where we can check our numerical methods against exact results.

Recall that the Fubini–Study metric is the unique (up to scale) Kähler metric on $\mathbb{P}^3$ with $SU(4)$ isometry, corresponding to the presentation of $\mathbb{P}^3$ as a symmetric space:

$$\mathbb{P}^3 = \frac{\mathrm{S}^7}{U(1)} = \frac{SU(4)}{\mathrm{S}(U(3) \times U(1))}. \tag{166}$$

The Fubini–Study metric is defined by $g_{i\bar{j}} = \partial_i \bar{\partial}_{\bar{j}} K$, where K is the Kähler potential

$$K = \frac{6^{1/3}}{2\pi} \log Z^I \bar{Z}_I. \tag{167}$$

Here $[Z^0 : \cdots : Z^3]$ are homogeneous coordinates on $\mathbb{P}^3$ where, for example, on the patch $U_0 = \{Z^0 = 1\}$, we have $Z^I = (1, z^i)$ with $i = 1, 2, 3$. The choice of prefactor in (167) ensures $\mathrm{Vol}(\mathbb{P}^3) = 1$.

The bundles we consider are line bundles $V = \mathcal{O}(m)$ on $\mathbb{P}^3$ for integer values of m. A hermitian metric on the fibres of $\mathcal{O}(m)$ is given by

$$G = (Z^I \bar{Z}_I)^{-m}. \tag{168}$$

Indeed, this is actually automatically Hermite–Einstein with respect to the Kähler metric defined by (167).

Analytic Results

The particular eigenvalue problem we want to solve is

$$\Delta_{\bar{\partial}_V} \phi = \lambda \phi, \tag{169}$$

where ϕ is an $\mathcal{O}(m)$-valued $(0,0)$- or $(0,1)$-form. At this point, we consider the general problem of $\mathbb{P}^N$ and specialise to $N = 3$ when presenting our numerical results. First, we note that global holomorphic sections of $\mathcal{O}(m)$ are counted by $H^{0,0}_{\bar{\partial}_V}(\mathbb{P}^N, \mathcal{O}(m)) \simeq H^0(\mathbb{P}^N, \mathcal{O}(m))$, which should match the number of harmonic modes of the above Laplacian. On $\mathbb{P}^N$, these cohomologies can

be computed using the Bott formula:

$$\begin{aligned}
\dim H^0(\mathbb{P}^N, \mathcal{O}(m)) &= \binom{N+m}{m}, \qquad m \geq 0, \\
\dim H^N\mathbb{P}^N, \mathcal{O}(m)) &= \binom{-m-1}{-N-m-1}, \qquad m \leq -N-1, \\
\dim H^p(\mathbb{P}^N, \mathcal{O}(m)) &= 0, \qquad \text{otherwise.}
\end{aligned} \tag{170}$$

On $\mathbb{P}^3$, we see that $\Omega^{0,0}(\mathcal{O}(m))$ will have zero modes, i.e. $\lambda = 0$, for $m \geq 0$, while there are no zero modes at all for $\Omega^{0,1}(\mathcal{O}(m))$.

Though finding the spectrum of $\Delta_{\bar{\partial}_V}$ on $\mathbb{P}^N$ where $V = \mathcal{O}(m)$ does not seem to have been considered in the literature before, there is a related problem from which we can extract the spectrum (at least for the scalar eigenmodes). Kuwabara [162] and Bykov and Smilga [163] analysed the spectrum of a Schrödinger operator on a line bundle $\mathcal{O}(m)$ over N-dimensional complex projective space equipped with a Fubini–Study metric with volume $(4\pi)^N/N!$. Given a connection D on $\mathcal{O}(m)$ with curvature F, they showed that the spectrum of the Bochner Laplacian $\Delta_D = DD^\dagger + D^\dagger D$ acting on $(0,0)$-forms (scalars) is spanned by the following eigenvalues λ_n with multiplicities ℓ_n for $n = 0, 1, \ldots$:

$$\lambda_n = \left(n + \frac{|m|}{2}\right)\left(n + \frac{|m|}{2} + N\right) - \frac{m^2}{4}, \tag{171}$$

$$\ell_n = \binom{n+N-1}{N-1}\binom{n+|m|+N-1}{N-1}\frac{2n+|m|+N}{N}. \tag{172}$$

From (73) we see that the Bochner Laplacian $\Delta_D \equiv \Delta_{\partial_V} + \Delta_{\bar{\partial}_V}$ acting on scalars is related to the $\bar{\partial}_V$-Laplacian as $\Delta_D = 2\Delta_{\bar{\partial}_V} + \Lambda F$, where Λ is contraction with the Kähler form on X. The fibre metric on $\mathcal{O}(m)$ is taken to be the unique Hermite–Einstein metric (168), so that $F = \frac{1}{2}m\omega$, which then implies $\Lambda F = Nm/2$. Thus, we expect the spectrum of the Dolbeault Laplacian to be given in terms of the Bochner Laplacian as

$$\Delta_{\bar{\partial}_V} = \frac{1}{2}\left(\Delta_D - \frac{Nm}{2}\right). \tag{173}$$

Finally, our convention that $\mathrm{Vol}(\mathbb{P}^N) = 1$ implies a rescaling of the eigenvalues by $4\pi/(N!)^{1/N}$. Putting this all together, we expect the $\Omega^{0,0}(\mathcal{O}(m))$ spectrum of the Dolbeault Laplacian to be given by

$$\lambda_n = \frac{2\pi}{(N!)^{1/N}}\left[n(n+N+|m|) + \frac{N(|m|-m)}{2}\right], \tag{174}$$

$$\ell_n = \binom{n+N-1}{N-1}\binom{n+|m|+N-1}{N-1}\frac{2n+|m|+N}{N}, \tag{175}$$

m	-3		-2		-1		0		1		2		3	
n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n
0	31.1	20	20.7	10	10.4	4	0	1	0	4	0	10	0	20
1	55.3	120	41.5	70	27.7	36	13.8	15	17.3	36	20.7	70	24.2	120
2	86.4	420	69.2	270	51.9	160	34.6	84	41.5	160	48.4	270	55.3	420
3	124	1120	104	770	83.0	500	62.2	300	72.6	500	83.0	770	93.4	1120
4	169	2520	145	1820	121	1260	96.8	825	111	1260	124	1820	138	2520

Table 8: Exact eigenvalues of $\Delta_{\bar{\partial}_V}$ and their multiplicities for $\mathcal{O}(m)$-valued scalars on $\mathbb{P}^3$.

for $n \geq 0$. As a check, we recall that the zero modes should appear with multiplicity predicted by (170). Indeed, for $n = 0$ the above reduces to

$$\lambda_0 = \frac{2\pi}{(N!)^{1/N}} \left[\frac{N(|m| - m)}{2} \right], \qquad \ell_0 = \binom{N + |m|}{|m|}, \tag{176}$$

so that one has zero modes, $\lambda = 0$, only for $m \geq 0$ with multiplicities agreeing with the Bott formula (170).

Using this exact expression, we give the first few eigenvalues in the spectrum for $N = 3$ and $m \in \{-3, \ldots, 3\}$ in Table 8.

We have not found exact expressions for the spectrum of $\mathcal{O}(m)$-valued $(0,1)$-forms on $\mathbb{P}^N$ in the literature. Instead, as a check of this spectrum, we will appeal to Serre duality (162) which implies that the $\mathcal{O}(m)$-valued $(0,1)$-form spectrum should agree with the $K_X \otimes \mathcal{O}(-m)$-valued $(0, N-1)$-form spectrum. In particular, for $N = 3$ we have $K_X \simeq \mathcal{O}(-4)$, so that the $\mathcal{O}(-m)$-valued $(0,1)$-form spectrum should agree with the honest $(0,2)$-form spectrum computed in previous work [161].

Numerical Results

We now want to find a basis of bundle-valued forms which can be used to approximate the space of eigenmodes of $\Delta_{\bar{\partial}_V}$ and calculate the spectrum via matrix (165). We first consider bundle-valued scalars for $m \geq 0$ and note that the set

$$\mathcal{F}^{0,0}_{k_\phi}(m) \equiv \frac{(\text{degree } k_\phi + m \text{ monomials in } Z^I) \otimes \overline{(\text{degree } k_\phi \text{ monomials in } Z^I)}}{(Z^I \bar{Z}_I)^{k_\phi}}, \tag{177}$$

gives a finite set of $\mathcal{O}(m)$-valued scalar functions α_A on $\mathbb{P}^N$, with the size of the set controlled by the non-negative integer parameter k_ϕ. Under the scaling $Z^I \to \nu Z^I$, the scalars transform as

$\alpha_A \to \nu^m \alpha_A$, and so they are naturally thought of as smooth sections of $\mathcal{O}(m)$. Upon increasing the degree k_ϕ, one has a series of inclusions

$$\mathcal{F}_0^{0,0}(m) \subset F_1^{0,0}(m) \subset \cdots \subset \Omega^{0,0}(\mathcal{O}(m)), \tag{178}$$

where $\mathcal{F}_0^{0,0}(m) \simeq H^0(X, \mathcal{O}(m))$, so that larger values of k_ϕ better approximate the (infinite-dimensional) space of $\mathcal{O}(m)$-valued scalar functions on $\mathbb{P}^N$, and so also the space of eigenfunctions of $\Delta_{\bar\partial_V}$. The eigenfunctions of $\Delta_{\bar\partial_V}$ on $\mathbb{P}^N$ are given by finite linear combinations of these functions at each degree.

There is a generalisation of (177) to give a finite set of (p,q)-forms at degree k_ϕ on $\mathbb{P}^N$. These are constructed by considering forms which are well defined on $\mathbb{P}^N$ under both the $\mathbb{R}^+$ and $U(1)$ action on the homogeneous coordinates. A simple extension of this produces a set $\mathcal{F}_{k_\phi}^{p,q}(m)$ of $\mathcal{O}(m)$-valued (p,q)-forms on $\mathbb{P}^N$ for $m \geq 0$:

$$\mathcal{F}_{k_\phi}^{p,q}(m) \equiv \frac{(\text{degree } k_\phi + m \ (p,0)\text{-forms in } Z^I) \otimes \overline{(\text{degree } k_\phi \ (0,q)\text{-forms in } Z^I)}}{(Z^I \bar Z_I)^{k_\phi}}, \tag{179}$$

where, for example, the degree-two $(1,0)$-forms are $\{Z^0 dZ^1 - Z^1 dZ^0, Z^0 dZ^2 - Z^2 dZ^0, \dots\}$ and so on. Unlike the scalars, there is some redundancy in this set, so one has to discard any $\alpha_A \in \mathcal{F}_{k_\phi}^{p,q}(m)$ which can be written as linear combinations of the remaining forms. Again, there is an inclusion of sets, $\mathcal{F}_0^{p,q}(m) \subset F_1^{p,q}(m) \subset \cdots \subset \Omega^{p,q}(\mathcal{O}(m))$, so that larger values of k_ϕ will better approximate the space of eigenmodes of $\Delta_{\bar\partial_V}$.

Note that the fibre metric (168) on $\mathcal{O}(1)$ pairs

$$G \colon \mathcal{O}(1) \times \overline{\mathcal{O}(1)} \to \mathbb{C}. \tag{180}$$

Here, our notation is that sections of $\mathcal{O}(1)$, $\mathcal{O}(-1)$ and $\overline{\mathcal{O}(1)}$ transform with factors of ν, ν^{-1} and $\bar\nu$ respectively under the scaling $Z^I \to \nu Z^I$. There is also a natural pairing between $\mathcal{O}(1)$ and its dual bundle $\mathcal{O}(1)^* \simeq \mathcal{O}(-1)$ such that $\mathcal{O}(1) \times \mathcal{O}(-1) \to \mathbb{C}$. Combining these two, we see there is a map between smooth sections of $\mathcal{O}(-1)$ and $\overline{\mathcal{O}(1)}$ of the form

$$\begin{aligned} &\overline{\mathcal{O}(1)} \to \mathcal{O}(-1) \\ &f(\bar Z) \mapsto \frac{f(\bar Z)}{Z^I \bar Z_I}, \end{aligned} \tag{181}$$

where $G = (Z^I \bar Z_I)^{-1}$ is the Hermite–Einstein metric on $\mathcal{O}(1)$. Given $f(\bar Z) \mapsto \bar\nu f(\bar Z)$ under the scaling of homogeneous coordinates on $\mathbb{P}^N$, we have $f(\bar Z)(Z^I \bar Z_I)^{-1} \mapsto \nu^{-1} f(\bar Z)(Z^I \bar Z_I)^{-1}$, and

so it transforms as a (smooth) section of $\mathcal{O}(-1)$. Extending this logic to all $m < 0$, this means the basis can be taken to be

$$\mathcal{F}^{p,q}_{k_\phi}(-|m|) \equiv \frac{(\text{degree } k_\phi \ (p,0)\text{-forms in } Z^I) \otimes \overline{(\text{degree } k_\phi + |m| \ (0,q)\text{-forms in } Z^I)}}{(Z^I \bar{Z}_I)^{k_\phi + |m|}}, \quad (182)$$

Denoting this basis by $\mathcal{F}^{p,q}_{k_\phi}(m) \equiv \mathcal{F}^{p,q}_{k_\phi}(-|m|)$ for $m < 0$, there is again an inclusion of sets, $\mathcal{F}^{p,q}_0(m) \subset F^{p,q}_1(m) \subset \cdots \subset \Omega^{p,q}(\mathcal{O}(m))$, so that the eigenmodes of $\Delta_{\bar{\partial}_V}$ on $\mathbb{P}^N$ for $m < 0$ are again given by finite linear combinations of these forms.

For $(p,q) = (0,0)$ and $k_\phi = 0$, the set $\mathcal{F}^{0,0}_0(-|m|)$ reduces to

$$\mathcal{F}^{0,0}_0(-|m|) = \frac{\overline{(\text{degree } |m| \text{ monomials in } Z^I)}}{(Z^I \bar{Z}_I)^{|m|}}, \quad (183)$$

which are never holomorphic, consistent with the absence of zero modes for $m < 0$ from (170). Similarly, for $(p,q) = (0,N)$ and $k_\phi = 0$, $\mathcal{F}^{0,N}_0(-|m|)$ is spanned by

$$\mathcal{F}^{0,N}_0(-|m|) = \frac{\overline{(\text{degree } |m| \ (0,N)\text{-forms in } Z^I)}}{(Z^I \bar{Z}_I)^{|m|}}, \quad (184)$$

where one includes only those that are linearly independent on $\mathbb{P}^N$. These are automatically $\bar{\partial}_V$-closed since there are no $(0, N+1)$-forms on a complex N-fold, and they actually give a basis for $H^N(\mathbb{P}^N, \mathcal{O}(-|m|))$ since they are also $\bar{\partial}^\dagger_V$-closed:

$$\bar{\partial}^\dagger_V \alpha_A \propto \star_V \partial \left((Z^I \bar{Z}_I)^{|m|} \alpha_A \right) = 0. \quad (185)$$

The number of these forms on $\mathbb{P}^N$ is $\dim \mathcal{F}^{0,N}_0(-|m|) = \binom{|m|-1}{|m|-N-1}$, again in agreement with the Bott formula.

Before presenting our numerical results, we recall the essential ingredients for computing the matrix elements Δ_{AB} and O_{AB} in (163) and (164) that determine the generalised eigenvalue problem for the spectrum.

One begins by choosing a truncated basis $\mathcal{F}^{0,0}_{k_\phi}(m) = \{\alpha_A\}$ of bundle-valued forms for some degree k_ϕ. Larger values of k_ϕ will give larger matrices which better approximate the action of the Laplacian on the space of bundle-valued forms. The matrix elements Δ_{AB} and O_{AB} are then computed relative to this basis by Monte Carlo integration on $\mathbb{P}^3$, where integrals over projective space are approximated by summing over N_ϕ random points $p_i \in \mathbb{P}^3$ according to

$$\int_{\mathbb{P}^3} f \ \text{vol} \simeq \frac{1}{N_\phi} \sum_{i=1}^{N_\phi} f(p_i). \quad (186)$$

m	-3		-2		-1		0		1		2		3	
n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n
0	31.1 ± 0.2	20	20.7 ± 0.1	10	10.37 ± 0.02	4	0	1	0	4	0	10	0	20
1	55.2 ± 0.7	120	41.4 ± 0.4	70	27.6 ± 0.2	36	13.8 ± 0.1	15	$17.3 \pm .1$	36	20.7 ± 0.2	70	24.2 ± 0.3	120
2	86 ± 2	420	69 ± 1	270	51.8 ± 0.8	160	34.6 ± 0.4	84	41.5 ± 0.6	160	48.3 ± 0.9	270	55 ± 1	420
3	125 ± 5	1120	104 ± 3	770	83 ± 2	500	62.3 ± 1.3	300	72.7 ± 1.9	500	83 ± 3	770	94 ± 4	1120

Table 9: Numerical eigenvalues λ_n of $\Delta_{\bar{\partial}_V}$ on $\mathbb{P}^3$ acting on $\mathcal{O}(m)$-valued scalars for $m \in \{-3, \ldots, 3\}$. We have also included their multiplicities ℓ_n. The quoted eigenvalues are the mean of the eigenvalues in a cluster, with the error given by the standard deviation of the cluster. We used $k_\phi = 3$ to allow us to compute the first four eigenspaces.

Here, vol is the volume form associated to the Fubini–Study metric on $\mathbb{P}^3$ and the distribution of the random points is chosen to reproduce this measure. With Δ_{AB} and O_{AB} in hand, one computes the eigenvalues and eigenvectors. The eigenvalues are the λ which appear in (169), with the eigenvectors determining the eigenmodes ϕ in (169) in the chosen basis $\{\alpha_A\}$.

The Bundle-Valued Scalar Spectrum

We begin with a numerical calculation of the scalar spectrum of bundle-valued eigenfunctions of $\Delta_{\bar{\partial}_V}$. The inputs are the exact Fubini–Study metric on $\mathbb{P}^3$ determined by the Kähler potential in (167), a bundle $\mathcal{O}(m)$ together with a hermitian structure determined by the Hermite–Einstein metric (168), a choice of degree k_ϕ which determines the size of the approximate basis (177) in which we expand the eigenfunctions, and the number of points N_ϕ that are used to discretise the integrals that appear in the matrix elements of the Laplacian. For the rest of this section, we fix $k_\phi = 3$ and $N_\phi = 10^6$ and compute the spectrum for $m \in \{-3, \ldots, 3\}$. The results are shown in Table 9 and Figure 20.

We see that the numerical results in Table 9 reproduce the exact results in Table 8 with excellent precision and the correct multiplicities. In particular, the mean of the numerical eigenvalues in each cluster match the exact results to better than 1% in all cases. One can also see this from Figure 20 which shows the numerical results and indicates the values of the exact eigenvalues; in all cases, the exact result is in the middle of the cluster of numerical eigenvalues.

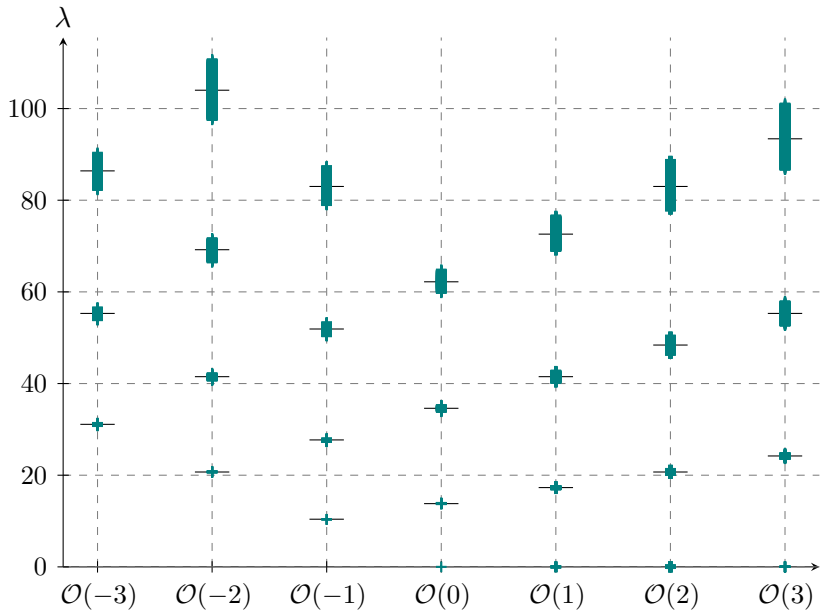


Figure 20: Numerical eigenvalues λ_n of $\Delta_{\bar{\partial}_V}$ on $\mathbb{P}^3$ acting on $\mathcal{O}(m)$-valued scalars for $m \in \{-3, \ldots, 3\}$. These were computed using the Fubini–Study metric on $\mathbb{P}^3$ and the associated Hermite–Einstein metric on $\mathcal{O}(m)$. Integrals were computed via Monte Carlo over $N_\phi = 10^6$ points. We used $k_\phi = 3$ for the basis functions, giving us access to the first four eigenspaces. The horizontal black lines indicate the exact analytic values from Table 8.

The Bundled-Valued $(0,1)$-Form Spectrum

Next, we make a numerical calculation of the $\Omega^{0,1}(\mathcal{O}(m))$ spectrum. This follows the scalar calculation almost exactly apart from using an appropriate basis of bundle-valued $(0,1)$-forms from (179). The results for $m \in \{-3, \ldots, 3\}$ are shown in Table 10 and Figure 21. Unlike the scalar spectrum, we do not have complete exact results to compare with.

As additional evidence that the spectra are correct, we recall that the $\bar{\partial}_V$ Hodge decomposition implies that a non-zero mode of $\Delta_{\bar{\partial}_V}$ must be either $\bar{\partial}_V$- or $\bar{\partial}_V^\dagger$-exact. Specifically, an $\mathcal{O}(m)$-valued $(0,1)$ eigenmode ϕ of the Laplacian with non-zero eigenvalue can be written as

$$\phi = \bar{\partial}_V \beta + \bar{\partial}_V^\dagger \gamma, \tag{187}$$

where β and γ are $\mathcal{O}(m)$-valued scalar and $(0,2)$-forms respectively. Crucially, since $\Delta_{\bar{\partial}_V}$ commutes with both $\bar{\partial}_V$ and $\bar{\partial}_V^\dagger$, β and γ will also be eigenmodes of the Laplacian with the same eigenvalue as ϕ. From this we see that the $\Omega^{0,1}(\mathcal{O}(m))$ spectrum must be some combination of the $\Omega^{0,0}(\mathcal{O}(m))$ and $\Omega^{0,2}(\mathcal{O}(m))$ spectra. In fact, using a further Hodge

m	-3		-2		-1		0		1		2		3	
n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n
0	20.74 ± 0.09	20	13.83 ± 0.03	6	10.37 ± 0.03	4	13.8 ± 0.1	15	17.3 ± 0.1	36	20.7 ± 0.2	70	24.2 ± 0.3	120
1	31.1 ± 0.2	20	20.74 ± 0.09	10	20.74 ± 0.09	20	27.7 ± 0.2	45	35.6 ± 0.3	84	41.4 ± 0.4	140	48.3 ± 0.6	216
2	41.5 ± 0.4	140	31.1 ± 0.2	64	27.7 ± 0.2	36	34.6 ± 0.4	84	41.5 ± 0.6	160	48.4 ± 0.9	270	55 ± 1	420
3	55.3 ± 0.7	120	41.5 ± 0.4	70	41.5 ± 0.4	140	51.9 ± 0.7	256	62 ± 1	420	73 ± 2	640	83 ± 2	924

Table 10: Numerical eigenvalues λ_n of $\Delta_{\bar{\partial}_V}$ on $\mathbb{P}^3$ acting on $\mathcal{O}(m)$-valued $(0,1)$-forms for $m \in \{-3, \ldots, 3\}$. We have also included their multiplicities ℓ_n. The quoted eigenvalues are the mean of the eigenvalues in a cluster, with the error given by the standard deviation of the cluster. We used $k_\phi = 2$ for $m < 0$ and $k_\phi = 3$ for $m \geq 0$.

decomposition for β and γ, it is simple to see that the $\Omega^{0,1}(\mathcal{O}(m))$ spectrum should consist of the entire $\Omega^{0,0}(\mathcal{O}(m))$ non-zero mode spectrum plus the $\Omega^{0,2}(\mathcal{O}(m))$ eigenvalues whose eigenmodes are $\bar{\partial}_V$-exact. We then have one final simplification: since $\bar{\star}_V$ commutes with the Laplacian, (162) implies that the $\Omega^{0,2}(\mathcal{O}(m))$ and $\Omega^{0,1}(\mathcal{O}(-4-m))$ spectra should match.

We can check these claims for the numerical spectrum we have calculated. For the $\bar{\partial}_V$-exact modes in (187), comparing Tables 9 and 10, we see that, for example, for $m = -2$ the eigenvalues $(20.7, 41.4)$ (to within numerical accuracy) appear in both the $(0,1)$ and $(0,0)$ spectra with the same multiplicities. A cursory glance at the rest of the results should assure the reader that this holds for the other values of m, with all the $(0,0)$ eigenvalues appearing in the $(0,1)$ spectra. For the $\bar{\partial}_V^\dagger$-exact modes in (187), for $m = -1$, for example, we expect that the remaining $\mathcal{O}(-1)$-valued $(0,1)$ eigenvalues should come from roughly half of $\mathcal{O}(-3)$-valued $(0,1)$ spectrum. Indeed, looking at Table 10, we see that both contain the eigenvalues $(20.7, 41.5)$ with the same multiplicities. Together, these constitute a non-trivial check that our numerical algorithm is correct for both the scalar and $(0,1)$ modes.

7.3 The spectrum of $\Delta_{\bar{\partial}_V}$ on T_2

We now apply our numerical method to calculate the spectrum of bundle-valued scalars and $(0,1)$-forms on Calabi–Yau manifolds. As a warm-up, and as another example where we can check things analytically, we consider a Calabi–Yau one-fold (a torus) defined by a single cubic equation in $\mathbb{P}^2$. Moving to Calabi–Yau three-folds is then just a matter of changing the dimension of the ambient projective space and the defining equation of the hypersurface (the algorithm does not change in any other way). With confidence that our algorithm is correct,

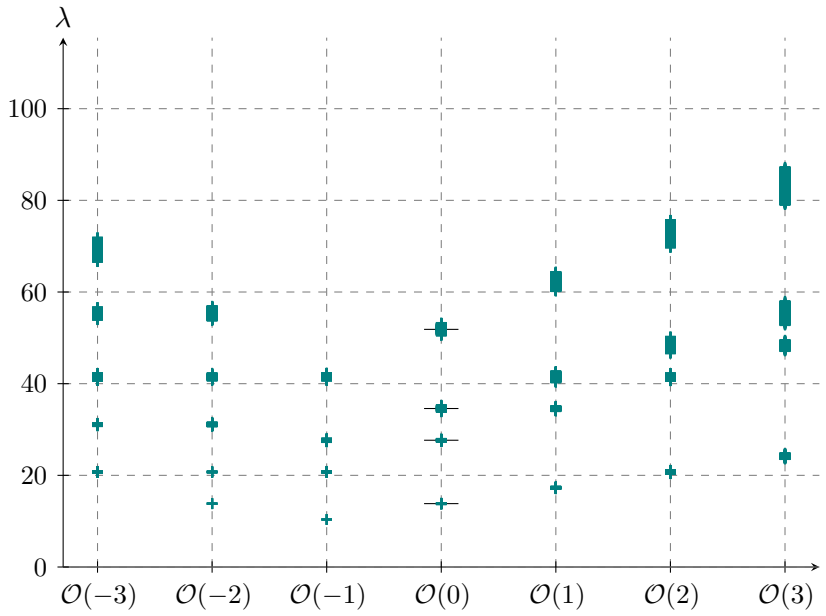


Figure 21: Numerical eigenvalues λ_n of $\Delta_{\bar{\partial}_V}$ on $\mathbb{P}^3$ acting on $\mathcal{O}(m)$-valued $(0,1)$-forms for $m \in \{-3, \ldots, 3\}$. These were computed using the Fubini–Study metric on $\mathbb{P}^3$ and the associated Hermite–Einstein metric on $\mathcal{O}(m)$. Integrals were computed via Monte Carlo over $N_\phi = 10^6$ points. We used $k_\phi = 2$ for $m < 0$ and $k_\phi = 3$ for $m \geq 0$ to allow us to compute the first four eigenspaces. The horizontal black lines indicate the exact analytic values for the $\Omega^{0,1}(X)$ spectrum.

in the next section we move to the more involved and physically relevant case of a Calabi–Yau threefold.

The particular one-fold that we will study is the Fermat cubic hypersurface X in $\mathbb{P}^2$ defined by the vanishing locus of the equation

$$Q \equiv Z_0^3 + Z_1^3 + Z_2^3 = 0. \tag{188}$$

Analytic Results

First, we note that since the canonical bundle of a Calabi–Yau is trivial, $K_X = \mathcal{O}$, the relations in (162) imply $\{\lambda\}_V^{0,0} = \{\lambda\}_{V^*}^{0,1}$. Thanks to this, once we compute the $\mathcal{O}(m)$-valued scalar spectrum for all m, we automatically have the bundle-valued $(0,1)$-form spectrum.

For $m = 0$, since $\Delta_{\bar{\partial}_V} = \frac{1}{2}\Delta$, the scalar spectrum is exactly one-half of that for the de Rham Laplacian on the torus. This is given by [164]

$$\lambda_{u,v} = 4\pi^2 b \left[\left(1 + \frac{a^2}{b^2}\right) u^2 - \frac{2a}{b^2} uv + \frac{v^2}{b^2} \right], \qquad u, v \in \mathbb{Z}, \tag{189}$$

where the complex structure $\tau \equiv a + ib$ is fixed to $e^{2\pi i/3}$ for the Fermat cubic. The eigenvalue multiplicities match the dimensions of irreducible representations of the symmetry group of X (automorphisms of Q together with complex conjugation), which is $(S_3 \times \mathbb{Z}_2) \rtimes (\mathbb{Z}_3 \times \mathbb{Z}_3)$ [165].

For $m \neq 0$, the exact scalar spectrum can be inferred from the results of Tejero Prieto [166]. There, they compute the eigenvalues and multiplicities for a Schrödinger-like operator

$$\hat{H} = \frac{\hbar^2}{2m} \Delta_D, \tag{190}$$

where D is a connection compatible with the hermitian metric on $V = \mathcal{O}(m)$, and $\Delta_D = D^\dagger D$ is the Bochner Laplacian for V. This can be related to the holomorphic structure on V as follows.

Given the Dolbeault operators ∂_V and $\bar{\partial}_V$, where $D = \partial_V + \bar{\partial}_V$, [166] gives the identity

$$\partial_V^\dagger \partial_V - \bar{\partial}_V^\dagger \bar{\partial}_V = \star F = \frac{e\hat{B}}{\hbar}, \tag{191}$$

where $F = eB/\hbar$ is the curvature of D and $B = \hat{B}\mathrm{vol}$. This implies

$$D^\dagger D = 2\bar{\partial}_V^\dagger \bar{\partial}_V + \frac{e\hat{B}}{\hbar}, \tag{192}$$

where $\hat{B}$ is related to the degree of the line bundle V by

$$\deg V = \frac{e\hat{B}}{2\pi\hbar}\mathrm{Vol}(X). \tag{193}$$

Remembering that our eigenvalue problem is for the Dolbeault Laplacian $\Delta_{\bar{\partial}_V} = \bar{\partial}_V^\dagger \bar{\partial}_V$, we can use (192) to relate the spectrum of $\hat{H}$ calculated in [166] with the spectrum of $\Delta_{\bar{\partial}_V}$. From Section 4.2 of that work, the spectrum (with multiplicity ℓ) of $\hat{H}$ is given by

$$\mathrm{spec}\hat{H} = \left\{ E_n = \frac{2\pi\hbar^2}{M\mathrm{Vol}(X)} |\deg V| \left(n + \tfrac{1}{2}\right),\ n \geq 0 \right\}, \tag{194}$$

$$\ell(E_n) = |\deg V|. \tag{195}$$

Equation (192) then implies

$$\mathrm{spec}\bar{\partial}_V^\dagger \bar{\partial}_V = \left\{ \lambda_n = \frac{2\pi|\deg V|}{\mathrm{Vol}(X)} \left(n + \tfrac{1}{2}(1 - \mathrm{sign}\deg V)\right),\ n \geq 0 \right\}. \tag{196}$$

We then recall that for a line bundle $V = \mathcal{O}(m)$ on a torus, Riemann–Roch implies that the degree is given by $\deg V = h^0(V) - h^0(V^*)$.[13] Thus, for $m > 0$, we have $\deg V = h^0(\mathcal{O}(m))$, while for $m < 0$ we have $\deg V = -h^0(\mathcal{O}(|m|))$, with $h^0(\mathcal{O}(|m|)) = 3|m|$. Finally, remembering that we always normalise the volume of the Calabi–Yau to one, the eigenvalues and multiplicities of $\Delta_{\bar{\partial}_V}$ for $m \neq 0$ should be

$$\lambda_n = \begin{cases} 6\pi mn & m > 0, \\ 6\pi|m|(n+1) & m < 0, \end{cases} \qquad n \geq 0, \tag{197}$$

$$\ell_n = 3|m|. \tag{198}$$

The spectra for $m \in \{-3, \ldots, 3\}$ are given in Table 11. In particular, we notice that there are no zero modes for $m < 0$, in agreement with $h^0(\mathcal{O}(m)) = 0$ for a negative-degree line bundle. The $\mathcal{O}(m)$-valued $(0,1)$-form spectra are then given by the $\mathcal{O}(-m)$-valued scalar spectra, corresponding to mirroring Table 11 about the $m = 0$ column.

Numerical results

Before presenting our numerical results, we quickly outline how the calculation on a Calabi-Yau hypersurface differs from that on projective space. Practically, the salient differences are:

[13] For a bundle V over a complex genus-g Riemann surface, the Riemann–Roch theorem implies

$$h^0(V) - h^1(V) = \deg V - (1-g)\mathrm{rank}V.$$

For a line bundle over a torus, $g = 1 = \mathrm{rank}V$ and the canonical bundle is trivial, so that $h^1(V) = h^0(V^*)$.

m	-3		-2		-1		0		1		2		3	
n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n
0	56.55	9	37.70	6	18.85	3	0.0	1	0.0	3	0.0	6	0.0	9
1	113.1	9	75.40	6	37.70	3	22.79	6	18.85	3	37.70	6	56.55	9
2	169.6	9	113.1	6	56.55	3	68.38	6	37.70	3	75.40	6	113.1	9
3	226.2	9	150.8	6	75.40	3	91.17	6	56.55	3	113.1	6	169.6	9
4	282.7	9	188.5	6	94.25	3	159.6	12	75.40	3	150.8	6	226.2	9
5	339.3	9	226.2	6	113.1	3	205.1	6	94.25	3	188.5	6	282.7	9
6	396.0	9	263.9	6	169.6	3	273.5	6	113.1	3	226.2	6	339.3	9

Table 11: Exact eigenvalues of $\Delta_{\bar{\partial}_V}$ and their multiplicities for $\mathcal{O}(m)$-valued scalars on the Fermat cubic. The spectrum of $\mathcal{O}(m)$-valued $(0,1)$-forms is given by reflecting the table about $m = 0$.

- The metric on the Calabi-Yau is not known analytically, but must be computed numerically. We compute the Calabi–Yau using the "energy functional" approach introduced by Headrick and Nassar [38]. In the case of the torus, the Calabi–Yau metric is simply the flat metric associated to the presentation of the torus as a quotient of $\mathbb{C}$. However, this metric looks non-trivial in the coordinates inherited from the ambient projective space. Thanks to this, and also to mimic the higher-dimensional case where there are no analytic results, we will compute the metric numerically.

- The set $F^{p,q}_{k_\phi}(m)$ defined in (179) is pulled back to the hypersurface to give an approximate basis on the Calabi–Yau. The set may be overcomplete in the sense that some elements are linearly dependent when restricted to the hypersurface. In practice, this means removing elements of $F^{p,q}_{k_\phi}(m)$ that are related by $Q = 0$. Choosing larger values of k_ϕ corresponds to using a larger basis of forms with which to approximate the eigenmodes of the Laplacian.

- The random points used to discretise integrals as in (186) should be distributed according to the Calabi–Yau measure rather than the Fubini–Study measure. This problem was solved for Calabi–Yau hypersurfaces by Douglas et al. [36] and Braun et al. [37].

The metric on X is given by a choice of complex structure, via the defining equation (188), and a choice of Kähler potential. This is approximated by an "algebraic metric" [35] with Kähler potential

$$K = \frac{1}{\pi k_h} \log s_\alpha h^{\alpha\bar{\beta}} \overline{s_\beta}, \tag{199}$$

where $h^{\alpha\bar{\beta}}$ is a hermitian matrix of parameters and $\{s_\alpha\}$ are a basis for the degree-k_h polynomials (sections of $\mathcal{O}(k_h)$) on $\mathbb{P}^2$ restricted to the hypersurface. Here, k_h is a positive

integer parameter which controls the complexity of the ansatz (199) – larger values of k_h should be thought of as including higher Fourier modes to better approximate the honest Calabi–Yau metric on X. The corresponding Kähler metric is $g_{i\bar{j}} = \partial_i \bar{\partial}_{\bar{j}} K$, where a pullback to the hypersurface on the $i, \bar{j}$ indices is implicit.

The bundles we consider are line bundles $V = \mathcal{O}(m)$ on the torus X for integer values of m. Since the approximate Calabi–Yau metric is defined by (199), similar to (168), a Hermite–Einstein metric on the fibres of $\mathcal{O}(m)$ is given by

$$G = \left(s_\alpha h^{\alpha\bar{\beta}} \overline{s_\beta}\right)^{-m/k_h}. \tag{200}$$

Again, one can check that this choice satisfies the hermitian Yang–Mills equation on X with the Kähler metric determined by (199). With these ingredients, we can now compute the numerical spectrum of the Dolbeault Laplacian on our first example of a Calabi–Yau hypersurface. In what follows, we computed the approximate Calabi–Yau metric at $k_h = 10$ corresponding to a "σ-measure" of $\sigma \approx 2 \times 10^{-15}$ [36]. Integrals were computed via Monte Carlo using $N_\phi = 10^6$ points.

The Bundle-Valued Scalar Spectrum

We begin with a numerical calculation of the spectrum of bundle-valued eigenfunctions of $\Delta_{\bar{\partial}_V}$. The inputs are the approximate Calabi–Yau metric on X determined by the Kähler potential in (199) with the parameters fixed by the "energy functional" approach [38], a bundle $\mathcal{O}(m)$ together with a Hermite–Einstein metric (200), a choice of degree k_ϕ which determines the size of the approximate basis (177) in which we expand the eigenfunctions, and the number of points $N_\phi = 10^6$ that are used to discretise the integrals that appear in matrix elements of the Laplacian. For the rest of this section, we fix $k_\phi = 3$ and compute the spectrum for $m \in \{-3, \ldots, 3\}$. Our numerical results are shown in Table 12 and Figure 22.

The numerical results in Table 12 reproduce the exact results in Table 11 with excellent precision and the correct multiplicities. This is also visible in Figure 22 which shows the numerical results and indicates the values of the exact eigenvalues; in all cases, the exact result lies in the middle of the cluster of numerical eigenvalues. For larger values of N_ϕ, the eigenvalues in Figure 22 become more tightly clustered. In the $N_\phi \to \infty$ limit, one recovers the $(S_3 \times \mathbb{Z}_2) \rtimes (\mathbb{Z}_3 \times \mathbb{Z}_3)$ symmetry of X, and the eigenvalues become exactly degenerate.

m	-3		-2		-1		0		1		2		3	
n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n
0	56.5 ± 0.2	9	37.7 ± 0.1	6	18.85 ± 0.05	3	0.0	1	0.0	3	0.0	6	0.0	9
1	113.1 ± 0.5	9	75.4 ± 0.3	6	37.7 ± 0.1	3	22.80 ± 0.08	6	18.85 ± 0.05	3	37.7 ± 0.1	6	56.5 ± 0.2	9
2	169.6 ± 0.7	9	113.1 ± 0.5	6	56.4 ± 0.2	3	68.4 ± 0.2	6	37.7 ± 0.1	3	75.4 ± 0.3	6	113.1 ± 0.5	9
3	226 ± 1	9	150.8 ± 0.7	6	75.6 ± 0.2	3	91.2 ± 0.4	6	56.6 ± 0.1	3	113.1 ± 0.5	6	169.6 ± 0.7	9
4	283 ± 2	9	188.5 ± 0.6	6	94.3 ± 0.3	3	159.7 ± 0.9	12	75.3 ± 0.1	3	150.8 ± 0.6	6	226 ± 1	9
5	340 ± 2	9	226.4 ± 0.8	6	113.0 ± 0.3	3	205.3 ± 0.9	6	94.3 ± 0.3	3	188.6 ± 0.7	6	283 ± 1	9
6	396 ± 2	9	264.0 ± 0.9	6	131.9 ± 0.4	3	274 ± 1	6	113.1 ± 0.2	3	226.1 ± 0.9	6	339 ± 1	9

Table 12: Numerical eigenvalues λ_n of $\Delta_{\bar{\partial}_V}$ on the Fermat cubic acting on $\mathcal{O}(m)$-valued scalars for $m \in \{-3, \ldots, 3\}$ with $k_\phi = 3$. We have also included their multiplicities ℓ_n. These were computed using a numerical Calabi–Yau metric computed at $k_h = 10$ and the associated Hermite–Einstein metric on $\mathcal{O}(m)$. Integrals were computed via Monte Carlo over $N_\phi = 10^6$ points. The quoted eigenvalues are the mean of the eigenvalues in a cluster, with the error given by the standard deviation of the cluster.

The Bundled-Valued $(0,1)$-Form Spectrum

Next, we have the numerical calculation of the $\Omega^{0,1}(\mathcal{O}(m))$ spectrum. Again, this follows the scalar calculation in the previous subsection almost exactly, apart from using an appropriate basis of bundle-valued $(0,1)$-forms from (179). The results for $m \in \{-3, \ldots, 3\}$ are shown in Table 13 and Figure 23.

As we mentioned at the start of this section, since X is Calabi–Yau, its canonical bundle is trivial, $K_X = \mathcal{O}$. The identity (162) then implies that the $\mathcal{O}(m)$-valued $(0,1)$-form spectrum should match the $\mathcal{O}(-m)$-valued scalar spectrum. Comparing Tables 12 and 13, we see this is indeed the case up to numerical accuracy. This is also apparent in Figure 23, where we have plotted the numerical $(0,1)$ eigenvalues and indicated the values that one infers from the exact $\mathcal{O}(m)$ scalar spectrum with black lines. In all cases, we see the two agree.

7.4 The spectrum of $\Delta_{\bar{\partial}_V}$ on the Quintic

In the previous section, we extended the numerical calculation of the bundle-valued scalar and $(0,1)$-form spectra to a torus defined as a hypersurface in projective space. From this

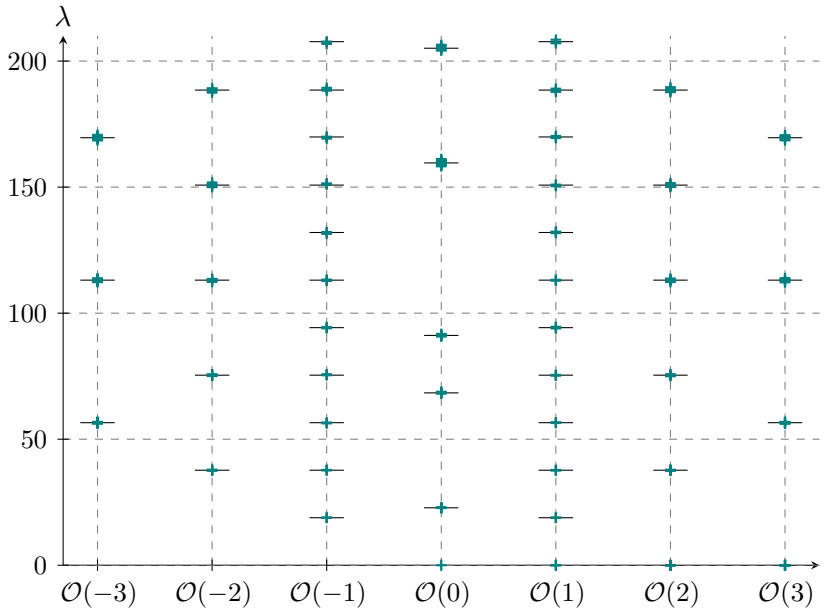


Figure 22: Numerical eigenvalues λ_n of $\Delta_{\bar{\partial}_V}$ on the Fermat cubic acting on $\mathcal{O}(m)$-valued scalars for $m \in \{-3, \ldots, 3\}$. These were computed using a numerical Calabi–Yau metric computed at $k_h = 10$ and the associated Hermite–Einstein metric on $\mathcal{O}(m)$. Integrals were computed via Monte Carlo over $N_\phi = 10^6$ points. We used $k_\phi = 3$ for the basis functions. The horizontal black lines indicate the exact analytic values from Table 11.

m	-3		-2		-1		0		1		2		3	
n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n
0	0.0	9	0.0	6	0.0	3	0.04	1	18.85 ± 0.05	3	37.7 ± 0.1	6	56.5 ± 0.2	9
1	56.5 ± 0.2	9	37.7 ± 0.1	6	18.85 ± 0.06	3	22.79 ± 0.08	6	37.7 ± 0.1	3	75.4 ± 0.3	6	113.1 ± 0.5	9
2	113.1 ± 0.5	9	75.4 ± 0.2	6	37.7 ± 0.1	3	68.4 ± 0.2	6	56.4 ± 0.2	3	113.1 ± 0.5	6	169.6 ± 0.7	9
3	169.6 ± 0.7	9	113.1 ± 0.4	6	56.5 ± 0.1	3	91.2 ± 0.2	6	75.6 ± 0.2	3	150.8 ± 0.7	6	226 ± 1	9
4	226 ± 1	9	150.8 ± 0.7	6	75.4 ± 0.1	3	159.5 ± 0.7	12	94.3 ± 0.3	3	188.5 ± 0.6	6	283 ± 2	9
5	283 ± 1	9	188.6 ± 0.6	6	94.3 ± 0.2	3	205.2 ± 0.7	6	113.0 ± 0.3	3	226.4 ± 0.8	6	340 ± 2	9
6	340 ± 1	9	226.3 ± 0.8	6	113.1 ± 0.3	3	274 ± 1	6	131.9 ± 0.4	3	264.0 ± 0.9	6	396 ± 2	9

Table 13: Numerical eigenvalues λ_n of $\Delta_{\bar{\partial}_V}$ on the Fermat cubic acting on $\mathcal{O}(m)$-valued $(0,1)$-forms for $m \in \{-3, \ldots, 3\}$ with $k_\phi = 3$ ($k_\phi = 4$ for $m = 0$). We have also included their multiplicities ℓ_n. These were computed using a numerical Calabi–Yau metric computed at $k_h = 10$ and the associated Hermite–Einstein metric on $\mathcal{O}(m)$. Integrals were computed via Monte Carlo over $N_\phi = 10^6$ points. The quoted eigenvalues are the mean of the eigenvalues in a cluster, with the error given by the standard deviation of the cluster. Thanks to (162), these eigenvalues should be related to those of Table 12 by $\{\lambda\}^{(0,0)}_{\mathcal{O}(m)} = \{\lambda\}^{(0,1)}_{\mathcal{O}(-m)}$, which simply reflects the table about $m = 0$.

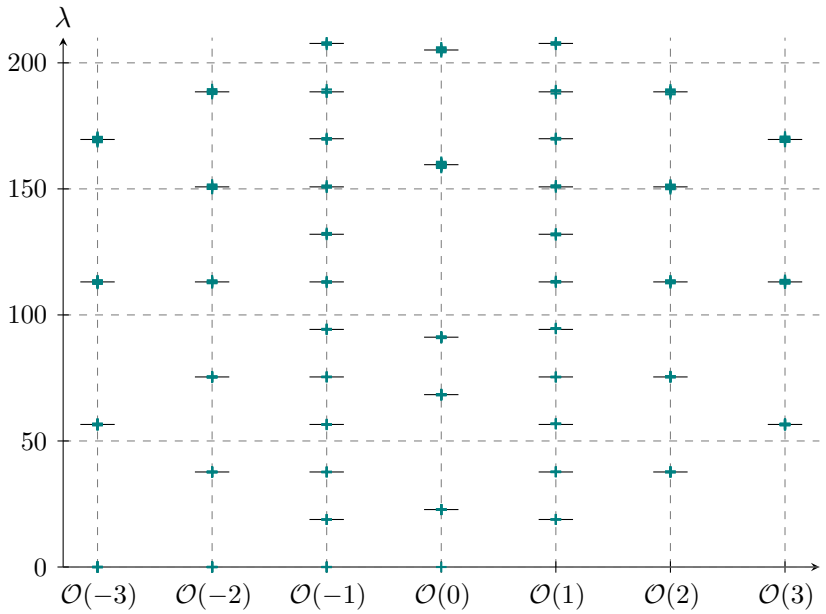


Figure 23: Numerical eigenvalues λ_n of $\Delta_{\bar{\partial}_V}$ on the Fermat cubic acting on $\mathcal{O}(m)$-valued $(0,1)$-forms for $m \in \{-3, \ldots, 3\}$. These were computed using a numerical Calabi–Yau metric computed at $k_h = 10$ and the associated Hermite–Einstein metric on $\mathcal{O}(m)$. Integrals were computed via Monte Carlo over $N_\phi = 10^6$ points. We used $k_\phi = 3$ for the basis functions. The horizontal black lines indicate the exact analytic values inferred from Table 11 and the identity (162).

toy example, it is simple to generalise to higher-dimensional Calabi–Yau manifolds defined as hypersurfaces. The particular example that we focus on is that of the Fermat quintic three-fold X defined as the vanishing locus in $\mathbb{P}^4$ of the equation

$$Q \equiv Z_0^5 + Z_1^5 + Z_2^5 + Z_3^5 + Z_4^5 = 0. \tag{201}$$

Unlike the previous examples, there are no analytic results to match to other than the dimensions of certain bundle-valued cohomologies which count zero modes. The results we present below are thus the first calculation of the spectrum of a bundle-valued Laplacian on a non-trivial Calabi–Yau manifold.

Before moving to the numerical results, we describe various constraints on bundle cohomologies on general Calabi–Yau manifolds. These will provide a consistency check for the count of zero modes. First, Serre duality relates the sheaf cohomologies as

$$H^p(X, V) = H^{n-p}(X, V^*). \tag{202}$$

Second, the Kodaira vanishing theorem states that on a Calabi–Yau X

$$H^p(X, V) = \{0\} \quad \text{for } p > 0 \text{ if } V \text{ is positive}, \tag{203}$$

where for manifolds with Picard rank one (such as the hypersurfaces in a single projective space that we consider), positive just means line bundles $V = \mathcal{O}(m)$ with $m > 0$. These constraints imply that on a threefold, such as the Fermat quintic, the only non-vanishing cohomologies are $h^0(\mathcal{O}(m))$ and $h^3(\mathcal{O}(-m))$ for $m > 0$. The scalar zero modes of $\Delta_{\bar{\partial}_V}$, counted by $h^0(\mathcal{O}(m))$, are the degree-m holomorphic monomials of the Z^I coordinates on $\mathbb{P}^4$ pulled back to the hypersurface. Since $h^1(\mathcal{O}(m)) = 0$ for all m, there are no bundle-valued $(0,1)$-form zero modes.

Numerical Results

Since there are no known explicit expressions for either the Calabi–Yau metric on the quintic nor Hermite–Einstein metrics on bundles over it, we must compute these numerically. The ansatz for the Kähler potential is again of the form (199), but now with degree-k_h polynomials on $\mathbb{P}^4$ restricted to the hypersurface. Similarly, the Hermite–Einstein metric on the fibres of $\mathcal{O}(m)$ over the quintic is given by (200). For what follows, we will use an approximate Calabi–Yau metric on the quintic computed at $k_h = 6$ using the "energy functional" approach of Headrick and Nassar [38], with a σ-measure of $\sigma \approx 2 \times 10^{-4}$ [36]. The

m	-3		-2		-1		0		1		2		3	
n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n
0	53.2 ± 0.2	35	35.4 ± 0.1	15	17.73 ± 0.05	5	0.0	1	0.0	5	0.0	15	0.0	35
1	71.9 ± 0.1	10	57.0 ± 0.1	20	39.5 ± 0.1	20	20.56 ± 0.07	20	21.79 ± 0.07	20	21.50 ± 0.06	20	17.94 ± 0.04	10
2	82.0 ± 0.4	60	62.6 ± 0.9	80	42.0 ± 0.2	30	39.4 ± 0.1	20	24.3 ± 0.1	30	26.6 ± 0.1	60	27.7 ± 0.1	60
3	85.8 ± 0.8	90	74.4 ± 0.1	15	51.9 ± 0.1	10	42.28 ± 0.06	4	33.75 ± 0.06	10	28.6 ± 0.1	20	30.7 ± 0.1	30
4	87.8 ± 0.1	5	80.2 ± 0.5	50	60.1 ± 0.1	15	47.3 ± 0.2	60	42.43 ± 0.08	15	38.8 ± 0.07	15	32.0 ± 0.2	60

Table 14: Numerical eigenvalues λ_n of $\Delta_{\bar{\partial}_V}$ on the Fermat quintic acting on $\mathcal{O}(m)$-valued scalars for $m \in \{-3, \ldots, 3\}$. These were computed using a numerical Calabi–Yau metric computed at $k_h = 6$ and the associated Hermite–Einstein metric on $\mathcal{O}(m)$. Integrals were computed via Monte Carlo over $N_\phi = 5 \times 10^6$ points. The approximate basis used $k_\phi = 3$, except for $m = \pm 3$ which were computed at $k_\phi = 2$. We have also included their multiplicities ℓ_n. The quoted eigenvalues are the mean of the eigenvalues in a cluster, with the error given by the standard deviation of the cluster.

numerical integrations were carried out using $N_\phi = 5 \times 10^6$ points. Unless otherwise stated, the spectra were computed using an approximate basis $\mathcal{F}^{p,q}_{k_\phi}(m)$ at $k_\phi = 3$.

The Bundle-Valued Scalar Spectrum

We have computed numerically the $\mathcal{O}(m)$-valued scalar spectrum of $\Delta_{\bar{\partial}_V}$ for $m \in \{-3, \ldots, 3\}$ on the Fermat quintic threefold. The results are shown in Table 14 and Figure 24. For $m = 0$, the eigenvalues are one-half of those computed in [161], as expected from the identity $\Delta_{\bar{\partial}_V} = \frac{1}{2}\Delta$ when $V = \mathcal{O}$. For $m > 0$, the zero modes of $\Delta_{\bar{\partial}_V}$ should be monomials of degree m in the homogeneous Z^I coordinates modulo the defining equation, $Q = 0$. The counting of these monomials, given by $h^0(\mathcal{O}(m)) = \binom{4+m}{m}$ for $0 < m < 5$, agrees with the number of zero modes in Table 14. For $m < 0$, the numerical results indicate there are no zero modes, in agreement with the vanishing of the relevant cohomologies that we mentioned above.

The Bundle-Valued $(0,1)$-Form Spectrum

Finally, we have the numerical calculation of the $\Omega^{0,1}(\mathcal{O}(m))$ spectrum on the Fermat quintic. Our results for $m \in \{-3, \ldots, 3\}$ are shown in Table 15 and Figure 25. We first note that there are no zero modes for any values of m, in agreement with the constraints from

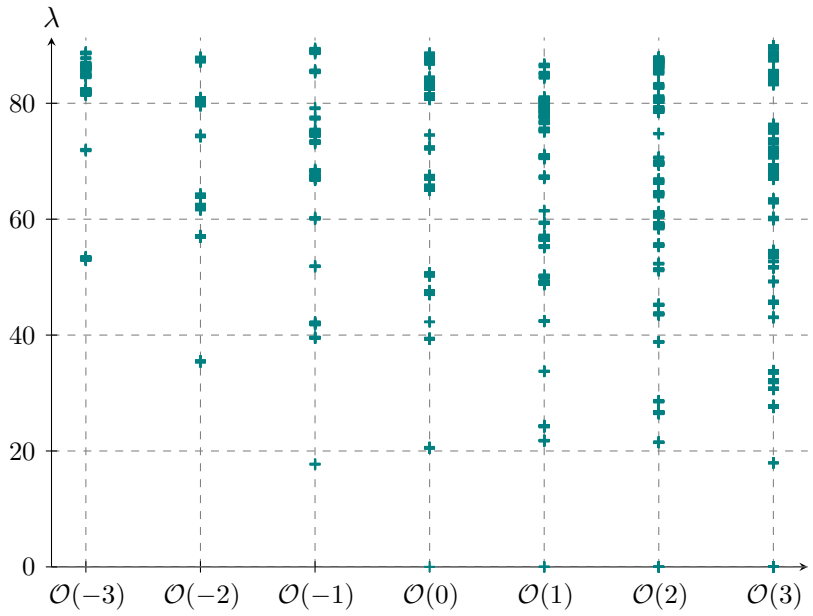


Figure 24: Numerical eigenvalues λ_n of $\Delta_{\bar{\partial}_V}$ on the Fermat quintic acting on $\mathcal{O}(m)$-valued scalars for $m \in \{-3, \ldots, 3\}$. These were computed using a numerical Calabi–Yau metric computed at $k_h = 6$ and the associated Hermite–Einstein metric on $\mathcal{O}(m)$. Integrals were computed via Monte Carlo over $N_\phi = 5 \times 10^6$ points. We used $k_\phi = 3$ for the basis functions, except for $m = \pm 3$ which were computed at $k_\phi = 2$.

Serre duality and the Kodaira vanishing theorem. As additional evidence that the spectra are consistent, we can again appeal to the $\bar{\partial}_V$ Hodge decomposition. This discussion mirrors that for projective space given around Equation (187). Since there are no zero modes, all $(0,1)$-form eigenmodes of $\Delta_{\bar{\partial}_V}$ must be either $\bar{\partial}_V$- or $\bar{\partial}_V^\dagger$-exact. The $\bar{\partial}_V$-exact eigenmodes must be of the form $\bar{\partial}_V \beta$, where β is an $\mathcal{O}(m)$-valued scalar eigenmode, while the $\bar{\partial}_V^\dagger$-exact modes are of the form $\bar{\partial}_V^\dagger \gamma$, where γ is $\bar{\partial}_V$-exact $\mathcal{O}(m)$-valued $(0,2)$ eigenmode. Since $\Delta_{\bar{\partial}_V}$ commutes with $\bar{\star}_V$ and the canonical bundle of a Calabi–Yau is trivial, the spectrum of $\mathcal{O}(m)$-valued $(0,2)$ eigenmodes agrees with the $\mathcal{O}(-m)$-valued $(0,1)$ spectrum. Putting this together, the spectrum of the Laplacian acting on $\Omega^{0,1}(\mathcal{O}(m))$ should be the union of the entire $\Omega^{0,0}(\mathcal{O}(m))$ spectrum and roughly half of the $\Omega^{0,1}(\mathcal{O}(-m))$ spectrum.

Comparing Tables 14 and 15, we see this appears to be the case, though with worse accuracy than we achieved for $\mathbb{P}^3$. For example, for $m = 1$, the $\Omega^{0,1}(\mathcal{O}(1))$ modes with eigenvalue 25.2 and multiplicity 50 originate from the $\Omega^{0,0}(\mathcal{O}(1))$ modes with eigenvalues $(21.8, 24.3)$ whose multiplicities sum to 50. Moving up the spectrum, the $\Omega^{0,1}(\mathcal{O}(1))$ modes with eigenvalue 31.7 and multiplicity 30 likely come from the $\Omega^{0,1}(\mathcal{O}(-1))$ modes with

eigenvalue 28.8 and the same multiplicity. Similarly, the $\Omega^{0,1}(\mathcal{O}(1))$ modes with eigenvalue 37.8 and multiplicity 10 likely come from the $\Omega^{0,0}(\mathcal{O}(1))$ modes with eigenvalue 33.8 and the same multiplicity.

A glance at the other results should convince the reader that this decomposition holds more generally, though the match is not perfect. This is likely due to inaccuracies introduced by the truncation at $k_\phi = 3$ to a finite-dimensional basis of forms. Recall that on projective space, the basis $\mathcal{F}^{p,q}_{k_\phi}(m)$ exactly spans the first k_ϕ eigenspaces of $\Delta_{\bar{\partial}_V}$. However, since the Calabi–Yau metric is not simply the pullback of Fubini–Study, the eigenspaces of the Laplacian are not exactly spanned by $\mathcal{F}^{p,q}_{k_\phi}(m)$ for finite k_ϕ, nor there is not a direct map between $(0,0)$ and $(0,1)$ modes at each degree k_ϕ. Instead, the approximate eigenmodes computed at some finite degree will receive corrections as k_ϕ is increased and the basis of forms is enlarged. We believe that upon moving to larger values of k_ϕ and increasing the number of integration points, the match between the $\Omega^{0,1}(\mathcal{O}(m))$ and the $\Omega^{0,0}(\mathcal{O}(m))$ and $\Omega^{0,1}(\mathcal{O}(-m))$ spectra will improve.

Regardless of this, one should remember that the lower-dimensional physics of a string compactification is determined by properties of harmonic/zero modes on the compactification manifold. These zero modes are, by definition, long wavelength and slowly varying, and likely to be very well approximated already at the modest values of k_ϕ that we have used. The same is certainly not true for massive modes higher up the spectrum; thankfully, these modes seem to be less relevant for low-energy physics questions.

Computing a Superpotential

The low-energy $\mathcal{N} = 1$ physics of a Calabi-Yau compactification is controlled by a superpotential and a Kähler potential. In particular, the matter sector is determined, to lowest order, by integrals of harmonic modes on the Calabi–Yau. In principle, the numerical method that we have presented gives us direct access to the data needed to compute all of this information. In practice, however, the line bundle and threefold we have considered are too simple to admit non-vanishing superpotential couplings. Let us see why this is the case.

The Fermat quintic threefold was constructed as a hypersurface in a single ambient projective space. This implies that the rank of the Picard lattice is one and so line bundles on

m	-3		-2		-1		0		1		2		3	
n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n	λ_n	ℓ_n
0	35.5 ± 0.1	40	23.7 ± 0.1	10	17.76 ± 0.04	5	21.60 ± 0.06	20	25.2 ± 0.1	50	29.7 ± 0.4	110	23.83 ± 0.05	20
1	53.3 ± 0.4	957	36.5 ± 0.1	15	28.77 ± 0.08	30	33.50 ± 0.08	30	31.7 ± 0.1	30	45.7 ± 0.1	15	33.9 ± 0.4	155
2	60.6 ± 0.3	30	43.1 ± 0.1	60	43.3 ± 0.5	110	36.7 ± 0.1	30	37.8 ± 0.1	10	47.1 ± 0.1	20	42.7 ± 0.1	40
3	62.6 ± 0.1	120	45.6 ± 0.1	40	49.7 ± 0.1	10	42.3 ± 0.1	34	42.8 ± 0.5	75	50.0 ± 0.2	60	44.73 ± 0.09	10
4	66.1 ± 0.1	15	51.9 ± 0.1	30	53.6 ± 0.2	115	48.0 ± 0.1	20	45.3 ± 0.1	10	51.8 ± 0.1	30	47.8 ± 0.1	15

Table 15: Numerical eigenvalues λ_n of $\Delta_{\bar{\partial}_V}$ on the Fermat quintic acting on $\mathcal{O}(m)$-valued $(0,1)$-forms for $m \in \{-3, \ldots, 3\}$. These were computed using a numerical Calabi–Yau metric computed at $k_h = 6$ and the associated Hermite–Einstein metric on $\mathcal{O}(m)$. Integrals were computed via Monte Carlo over $N_\phi = 5 \times 10^6$ points. We used $k_\phi = 3$ for the basis functions for $m = 0, \pm 1$ and $k_\phi = 2$ for $m = \pm 2, \pm 3$. We have also included their multiplicities ℓ_n. The quoted eigenvalues are the mean of the eigenvalues in a cluster, with the error given by the standard deviation of the cluster.

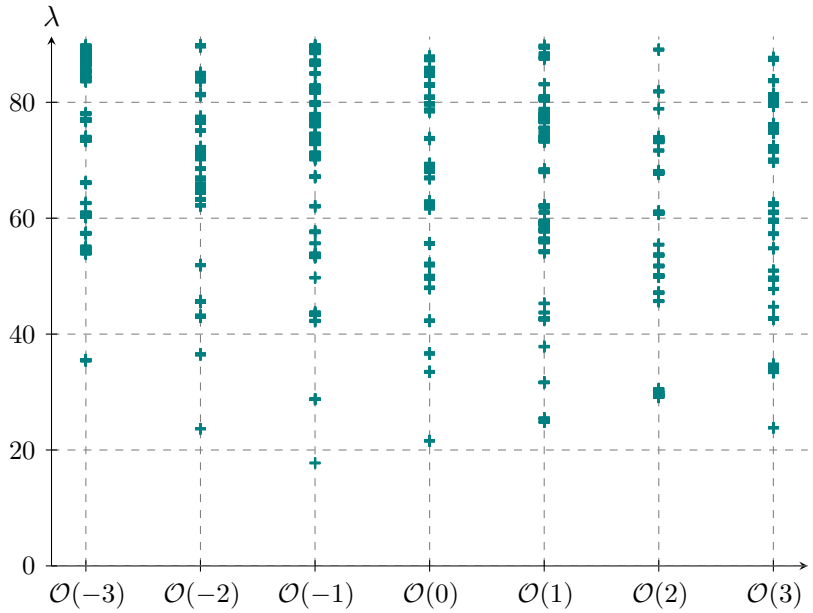


Figure 25: Numerical eigenvalues λ_n of $\Delta_{\bar{\partial}_V}$ on the Fermat quintic acting on $\mathcal{O}(m)$-valued $(0,1)$-forms for $m \in \{-3, \ldots, 3\}$. These were computed using a numerical Calabi–Yau metric computed at $k_h = 6$ and the associated Hermite–Einstein metric on $\mathcal{O}(m)$. Integrals were computed via Monte Carlo over $N_\phi = 5 \times 10^6$ points. We used $k_\phi = 3$ for the basis functions for $m = 0, \pm 1$ and $k_\phi = 2$ for $m = \pm 2, \pm 3$.

this quintic are of the form $\mathcal{O}(m)$ for some integer m. Now imagine trying to write down a non-vanishing superpotential coupling as

$$\lambda_{IJK}(m_1, m_2, m_3) = \int_X \Omega \wedge \psi^I_{m_1} \wedge \psi^J_{m_2} \wedge \psi^K_{m_3}, \tag{204}$$

where $\psi^I_{m_1} \in H^1(X, \mathcal{O}(m_1))$ is an $\mathcal{O}(m_1)$-valued harmonic $(0,1)$-form, and we have dropped a trace compared with (23) since the relevant group is abelian. Since Ω is an honest three-form, this integral vanishes whenever the degrees of the relevant line bundles do not sum to zero:

$$\lambda_{IJK}(m_1, m_2, m_3) = 0 \quad \text{if} \quad m_1 + m_2 + m_3 \neq 0. \tag{205}$$

In other words, the charges of the harmonic modes must sum to zero so that the integrand of (204) is an honest top-form. Thus, for a non-vanishing superpotential contribution, all the charges must be zero or at least one of them is negative. However, it is simple to argue that the requisite harmonic $(0,1)$ modes are not present in either case. When the charges are zero, we need harmonic $(0,1)$-forms. These are counted by the Hodge number $h^{0,1}$ which vanishes on the quintic, so the $(0,1)$ modes are not present. When there are both positive and negative charges, we can appeal to Serre duality and the Kodaira vanishing theorem. The first of these gives $h^{0,1}(\mathcal{O}(-m)) = h^{0,2}(\mathcal{O}(m))$, while the second implies $h^{0,1}(\mathcal{O}(m)) = h^{0,2}(\mathcal{O}(m)) = 0$ if $m > 0$. Together, these imply that there are no harmonic $\mathcal{O}(m)$-valued $(0,1)$-forms for either sign of m, in complete agreement with our numerical results in Table 15. We conclude that there are no non-vanishing superpotential couplings for matter coming from line bundles on the Fermat quintic.

7.5 Outlook

To summarise, the results of this section are:

- The construction of a finite basis of bundle-valued differential forms on complex projective space, which can be used to approximate the space of eigenmodes of the Dolbeault Laplacian on both $\mathbb{P}^N$ and hypersurfaces therein.

- Numerical calculations of the eigenmodes and eigenvalues of the Dolbeault Laplacian on $\mathcal{O}(m)$-valued scalars and $(0,1)$-forms on $\mathbb{P}^3$. We find perfect agreement with known exact results for the scalar spectrum and perform a number of consistency checks for the $(0,1)$ spectrum. We have not found exact results for the bundle-valued $(0,1)$-form spectrum in the literature, so, to the best of the authors' knowledge, our numerical calculation is the first of its kind.

- Numerical calculations of the eigenmodes and eigenvalues of the Dolbeault Laplacian on $\mathcal{O}(m)$-valued scalars and $(0,1)$-forms on a torus. The torus is a Calabi–Yau one-fold, allowing us to compare our numerical results with exact predictions. Again, we find perfect agreement with the known exact results.

- The first numerical calculations of the eigenmodes and eigenvalues of the Dolbeault Laplacian on $\mathcal{O}(m)$-valued scalars and $(0,1)$-forms on a Calabi–Yau three-fold. We focus on the Fermat quintic threefold. Here, there are no analytic results to compare with; instead, we perform a number of non-trivial checks on the spectrum which come from Serre duality and the Hodge decomposition.

Here we focused on the examples of line bundles over Calabi-Yau manifolds in a single ambient projective space. There are a number of ways to generalise our set-up to allow for interesting superpotential couplings. First is simply moving from line bundles to non-abelian bundles, such as the examples given in [153–155], where instead of the charges summing to zero, one requires a singlet in the antisymmetric product of the three representations appearing in the cubic coupling. Second, we can stay with line bundles but move to Calabi–Yau manifolds with higher-rank Picard lattices. Line bundles on these spaces are labelled by a vector of charges $\boldsymbol{m}$ and the corresponding vanishing theorems are less restrictive. In practice, this means moving to, for example, complete intersection Calabi–Yau (CICY) manifolds given as hypersurfaces in products of projective spaces. These advances should allow concrete computations of masses and couplings in top-down string models in the near future.

8 Conclusion

In this thesis we have addressed a number of issues related to string compactification, using techniques from data science and machine learning (ML), and in doing so proved that these tools are useful for performing research in string theory. This represents a paradigm shift in theoretical physics. From exploring the string landscape to probing fundamental interactions, ML techniques are reshaping our approach to understanding the nature of the universe and ushering in a new era of scientific discovery.

In Section 5 we considered with the problem of dealing with a large string landscape. We showed how, rather than randomly sampling Calabi-Yau (CY) manifolds as compactification spaces, one can use ML methods to generate CY manifolds that produce a desired set of properties in the effective field theory (EFT). Specifically, we looked at CY n-fold hypersurfaces in toric varieties built from $n+1$-dimensional reflexive polytopes. We demonstrated how genetic algorithms (GAs) are able to generate reflexive polytopes by exploring an environment with the goal of maximising some reward [49]. Within the results presented is an example of a targeted search, where CY fourfolds with desired properties and satisfying certain EFT constraints were generated by modifying the fitness function. This introduces a new more efficient approach in string phenomenology of sampling the string landscape. There are several obvious extensions of this work, for instance, one could investigate whether these methods are capable of generating fine regular star triangulations of reflexive polytopes or CY manifolds via other constructions, such as complete intersection Calabi-Yau (CICY) manifolds. Furthermore, by simple modifications to the fitness functions of our GA algorithm, one can use these algorithms to search for CY manifolds that satisfy a number of phenomenology constraints. For example, combining the reflexive polytope GA with the GA in [121] it would be interesting to see whether our a GA is capable of finding CY threefolds that optimise axion decay constants and axion-photon.

Whilst the study of CY manifolds with ML has been an active area of research since its inception in 2017 [16–19], until [50], there had not been applications of ML to study G_2-geometry. Unlike for CYs, where there are multiple large datasets, there are no such datasets for G_2 manifolds. This reflects the difficulty in describing these manifolds with algebraic discrete data. For the particular case of G_2 manifolds constructed as CY links, however, we can represent the manifold by a list of weights. In Section 6 we construct a

dataset of CY links and study their topological properties, namely their Sasakian Hodge numbers and Crowley-Nördtrom homotopy invariant. We successfully apply a neural network to learning $h^{2,1}$ of the CY links from the input weights. This serves as proof that G_2 manifolds are also amenable to ML and opens the door to a whole new field of G_2-ML research. Inspired by the many applications of machine learning to finding Ricci-flat metrics on CY manifolds [40–48], we hope in future work to apply similar methods to look for Ricci-flat metrics on G_2 manifolds. Such metrics will allow one to study the EFT arising from compactification of M-theory on such a manifold.

In the past string theory has been criticised due to the fact that it has not yet been connected to the Standard Model. In principle, the masses and couplings in string models compactified on CY manifolds can be computed directly from the geometry of the CY. However, even after decades of work, we are still unable to compute these numbers for all but the simplest examples. A substantial part of the difficulty can be attributed to the lack of explicit expressions for non-trivial CY metrics. However, with new approximations of Ricci-flat CY metrics coming from ML methods [40–48, 157], we are are a stage now where we can begin to compute the resulting particle physics. In Section 7, detailing work that was carried out in [51], we considered the $E_8 \times E_8$ heterotic string compactified on a CY threefold X with holomorphic vector bundle V and used ML approximations of the Ricci-flat metric on X to compute the resulting Yukawa couplings in the four-dimensional EFT. Specifically, we looked at line bundles over CY threefolds defined as hypersurfaces in a single ambient projective space. Unfortunately these examples are too simple to produce physically realistic Yukawa couplings. There are a number of ways to generalise our set-up to allow for physically relevant couplings. For example moving to CICY manifolds and non-abelian bundles.

Combining this with the results of Section 5, we are in a position to, for the first time, efficiently pick out suitable candidate theories from the vast landscape of string models and then check if their properties truly match those of observed particle physics. From this one may shed light on the question of whether string theory is indeed the correct description of nature.

A Appendix

A.1 Normal Form

There are two sources of redundancy when defining a reflexive polytope Δ in n dimensions by its $n \times N_{\rm v}$ vertex matrix V, whose columns are the $N_{\rm v}$ vertices. First of all, one can permute the vertices, leading to an $S_{N_{\rm v}}$ symmetry which permutes the columns of V. Secondly, one can perform a coordinate transformation on the n-dimensional lattice by acting on V with a $\mathrm{GL}(n,\mathbb{Z})$ matrix from the left. Altogether, this amounts to a transformation of the vertex matrix

$$\tilde{V} = GVP \ . \tag{206}$$

We consider two polytopes Δ and $\tilde{\Delta}$ related by (206) to be equivalence. The most efficient way to eliminate the redundancy due to this identification is to define a normal form for the vertex matrix, thereby selecting precisely one representative per equivalence class. It is known that the number of reflexive polytopes, after modding out by this identification, is finite in any given dimension n [167].

This is the approach that was used by Kreuzer and Skarke in constructing the complete classification of three- and four-dimensional reflexive polytopes [86, 111] and is included in the PALP software package [123]. If two polytopes Δ and $\tilde{\Delta}$ have the same normal form, then they are equivalent, in the sense that they are isomorphic with respect to a lattice automorphism. A detailed description of the how one computes the normal form is given in [168]. We shall give a short description here.

Let M be a n-dimensional lattice and $\Delta \subset M_{\mathbb{Q}}$ a n-dimensional lattice polytope with $N_{\rm v}$ vertices, $N_{\rm f} = |F(\Delta)|$ facets and vertex matrix V. We also define the supporting hyperplanes of Δ, associated to the facets $\varphi_i \in F(\Delta)$, as the set of all vectors v satisfying $\langle w_i, v\rangle = -c_i$, where $(w_i, c_i) \in M^* \times \mathbb{Z}$. The algorithm to compute the normal form is then as follows.

1. Compute the $N_{\rm f} \times N_{\rm v}$ vertex-facet pairing matrix PM:

$$PM_{ij} := \langle w_i, v_j\rangle + c_i \ . \tag{207}$$

2. Order the pairing matrix PM lexicographically to get the maximal matrix $PM^{\max}$.

3. Further rearrange the columns of $PM^{\max}$ to get M by the following:

$M \leftarrow PM^{\max}$

for $i = 1$ **to** $N_{\rm v}$ **do**

$k \leftarrow i$

for $j = i + 1$ **to** $N_{\rm v}$ **do**

if $c_M(j) < c_M(k) \vee (c_M(j) = c_M(k) \wedge s_M(j) < s_M(k))$ **then**

$k \leftarrow j$

end if

end for

$M \leftarrow \mathrm{SwapColumn}(M, i, k)$

end for

where $c_M(j) := \max(M_{ij} | 1 \leq i \leq N_{\rm f})$ and $s_M(j) := \sum_{i=1}^{N_{\rm f}} M_{ij}$, where $1 \leq j \leq N_{\rm v}$.

4. Let $\sigma_{\max}$ denote the associated element of $S_{N_{\rm f}} \times S_{N_{\rm v}}$ that transforms PM into M. Order the columns of V according to the restriction of $\sigma_{\max}$ to $S_{N_{\rm v}}$ to get the maximal vertex matrix $V^{\max}$. This removes the permutation degeneracy.

5. Compute the row style Hermite normal form of $V^{\max}$ to obtain the normal form NF. This step removes the $GL(n, \mathbb{Z})$ degeneracy.

Example 5 *To illustrate the above algorithm, we present an example in three dimensions. Let* Δ *be a lattice polytope defined by the vertex matrix:*

$$V = \begin{pmatrix} 0 & -2 & -1 & -1 & 1 & -3 & 2 & -2 \\ 1 & -3 & -1 & 0 & 0 & -3 & 1 & -1 \\ 1 & -3 & -2 & -1 & 0 & -4 & 3 & -2 \end{pmatrix} . \tag{208}$$

Computing the vertex-facet pairing matrix we get

$$PM = \begin{pmatrix} 2 & 0 & 0 & 0 & 2 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 2 & 2 & 2 & 1 \\ 0 & 6 & 6 & 0 & 0 & 1 & 3 & 2 \\ 2 & 1 & 2 & 2 & 0 & 0 & 0 & 1 \\ 3 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 2 & 5 & 6 & 0 & 0 & 0 & 2 & 1 \\ 3 & 2 & 3 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 & 0 & 1 & 0 & 2 \end{pmatrix} . \tag{209}$$

Ordering PM lexicographically we get the following maximal matrix:

$$PM^{max} = \begin{pmatrix} 6 & 6 & 3 & 2 & 1 & 0 & 0 & 0 \\ 6 & 5 & 2 & 1 & 0 & 2 & 0 & 0 \\ 3 & 2 & 1 & 0 & 0 & 3 & 1 & 0 \\ 2 & 1 & 0 & 1 & 0 & 2 & 0 & 2 \\ 1 & 0 & 0 & 0 & 0 & 3 & 1 & 1 \\ 0 & 1 & 2 & 1 & 2 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 & 1 & 2 & 2 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 & 0 & 3 \end{pmatrix}, \tag{210}$$

corresponding to the row and column permutations $(3, 6, 7, 4, 5, 2, 1, 8)$ and $(3, 2, 7, 8, 6, 1, 5, 4)$ respectively. Further ordering the columns by the procedure described in Step 3 above we get

$$M = \begin{pmatrix} 0 & 1 & 0 & 2 & 0 & 6 & 6 & 3 \\ 2 & 0 & 0 & 1 & 0 & 6 & 5 & 2 \\ 3 & 0 & 0 & 0 & 1 & 3 & 2 & 1 \\ 2 & 0 & 2 & 1 & 0 & 2 & 1 & 0 \\ 3 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 2 & 0 & 1 & 2 & 0 & 1 & 2 \\ 2 & 1 & 0 & 0 & 2 & 0 & 0 & 1 \\ 0 & 1 & 3 & 2 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{211}$$

corresponding to the column permutation $(6, 5, 8, 4, 7, 1, 2, 3)$. Ordering the columns in V correspondingly we get

$$V^{max} = \begin{pmatrix} -3 & 1 & -2 & -1 & 2 & 0 & -2 & -1 \\ -3 & 0 & -1 & 0 & 1 & 1 & -3 & -1 \\ -4 & 0 & -2 & -1 & 3 & 1 & -3 & -2 \end{pmatrix}. \tag{212}$$

Finally, computing the row style Hermite normal form of V^{max} we arrive at the following normal form:

$$NF = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & -2 & 1 & 0 \\ 0 & 1 & -1 & 0 & -1 & 1 & -3 & -3 \\ 0 & 0 & 0 & 1 & -1 & 0 & -2 & -2 \end{pmatrix}. \tag{213}$$

References

[1] G. Veneziano, Construction of a Crossing-Symmetric, Regge Behaved Amplitude for Linearly Rising Trajectories, Nuovo Cim. 57 (1968) 190–197. doi:10.1007/BF02824451.

[2] L. Susskind, Dual Symmetric Theory of Hadrons. 1., Nuovo Cim. 69 (1970) 457–496. doi:10.1007/BF02726485.

[3] T. Yoneya, Quantum Gravity and the Zero-Slope Limit of the Generalized Virasoro Model, Nuovo Cim. 8 (1973) 951–955. doi:10.1007/BF02727806.

[4] T. Yoneya, Connection of Dual Models to Electrodynamics and Gravidynamics, Progress of Theoretical Physics 51 (6) (1974) 1907–1920. doi:10.1143/PTP.51.1907.

[5] J. Scherk, J. H. Schwarz, Dual Models for Non-Hadrons, Nucl. Phys. B 81 (1) (1974) 118–144. doi:https://doi.org/10.1016/0550-3213(74)90010-8.

[6] S. Weinberg, The Quantum Theory of Fields, Cambridge University Press, 1995.

[7] M. E. Peskin, D. V. Schroeder, An Introduction to Quantum Field Theory, Westview Press, 1995.

[8] P. Candelas, G. T. Horowitz, A. Strominger, E. Witten, Vacuum Configurations for Superstrings, Nucl. Phys. B 258 (1985) 46–74. doi:10.1016/0550-3213(85)90602-9.

[9] P. Ramond, Dual Theory for Free Fermions, Phys. Rev. D 3 (1971) 2415–2418. doi:10.1103/PhysRevD.3.2415.

[10] M. B. Green, J. H. Schwarz, E. Witten, Superstring Theory 1: Introduction, Cambridge University Press, 1988.

[11] M. B. Green, J. H. Schwarz, E. Witten, Superstring Theory 2: Loop Amplitudes, Anomalies and Phenomenology, Cambridge University Press, 1988.

[12] M. Berger, Sur les groupes d'holonomie homogènes de variétés à connexion affine et des variétés riemanniennes, Bulletin de la Société Mathématique de France 83 (1955) 279–330. doi:10.24033/bsmf.1464.

[13] C. T. C. Wall, Classification Problems in Differential Topology. V, Inventiones mathematicae 1 (1966) 355–374. doi:10.1007/BF01389738.

[14] A. Chandra, A. Constantin, C. S. Fraser-Taliente, T. R. Harvey, A. Lukas, Enumerating Calabi-Yau Manifolds: Placing Bounds on the Number of Diffeomorphism Classes in the Kreuzer-Skarke List (2023). arXiv:2310.05909.

[15] N. Gendler, N. MacFadden, L. McAllister, J. Moritz, R. Nally, A. Schachner, M. Stillman, Counting Calabi-Yau Threefolds (2023). arXiv:2310.06820.

[16] Y.-H. He, Deep-Learning the Landscape (2017). arXiv:1706.02714.

[17] J. Carifio, J. Halverson, D. Krioukov, B. D. Nelson, Machine Learning in the String Landscape, JHEP 9 (2017) 157. doi:10.1007/JHEP09(2017)157.

[18] D. Krefl, R.-K. Seong, Machine Learning of Calabi-Yau Volumes, Phys. Rev. D 96 (6) (2017) 66014. doi:10.1103/PhysRevD.96.066014.

[19] F. Ruehle, Evolving Neural Networks With Genetic Algorithms to Study the String Landscape, JHEP 8 (2017) 38. doi:10.1007/JHEP08(2017)038.

[20] G. Arias-Tamargo, Y.-H. He, E. Heyes, E. Hirst, D. Rodriguez-Gomez, Brain Webs for Brane Webs, Phys. Lett. B 833 (2022) 137376. doi:10.1016/j.physletb.2022.137376.

[21] P.-P. Dechant, Y.-H. He, E. Heyes, E. Hirst, Cluster Algebras: Network Science and Machine Learning, J. Comput. Algebra 8 (2023). doi:10.1016/j.jaca.2023.100008.

[22] M.-W. Cheung, P.-P. Dechant, Y.-H. He, E. Heyes, E. Hirst, J.-R. Li, Clustering Cluster Algebras With Clusters, Adv. Theor. Math. Phys. 27 (3) (2023) 797–828. doi:10.4310/ATMP.2023.v27.n3.a5.

[23] S. Chen, P.-P. Dechant, Y.-H. He, E. Heyes, E. Hirst, D. Riabchenko, Machine Learning Clifford Invariants of ADE Coxeter Elements, Adv. Appl. Clifford Algebras 34 (3) (2024) 20. doi:10.1007/s00006-024-01325-y.

[24] F. Costantino, Y.-H. He, E. Heyes, E. Hirst, Learning 3-Manifold Triangulations (2024). arXiv:2405.09610.

[25] Y.-H. He, E. Hirst, T. Peterken, Machine-Learning Dessins D'Enfants: Explorations via Modular and Seiberg-Witten Curves, J. Phys. A 54 (7) (2021) 075401. doi:10.1088/1751-8121/abbc4f.

[26] J. Bao, S. Franco, Y.-H. He, E. Hirst, G. Musiker, Y. Xiao, Quiver Mutations, Seiberg Duality and Machine Learning, Phys. Rev. D 102 (8) (2020) 086013. doi:10.1103/PhysRevD.102.086013.

[27] J. Bao, O. Foda, Y.-H. He, E. Hirst, J. Read, Y. Xiao, F. Yagi, Dessins D'Enfants, Seiberg-Witten Curves and Conformal Blocks, JHEP 5 (2021) 65. doi:10.1007/JHEP05(2021)065.

[28] J. Bao, Y.-H. He, E. Hirst, J. Hofscheier, A. Kasprzyk, S. Majumder, Hilbert Series, Machine Learning, and Applications to Physics, Phys. Lett. B 827 (2022) 136966. doi:10.1016/j.physletb.2022.136966.

[29] J. Bao, Y.-H. He, E. Hirst, Neurons on Amoebae, J. Symb. Comput. 116 (2022) 1–38. doi:10.1016/j.jsc.2022.08.021.

[30] S. Chen, Y.-H. He, E. Hirst, A. Nestor, A. Zahabi, Mahler Measuring the Genetic Code of Amoebae (2022). arXiv:2212.06553.

[31] P. Berglund, G. Butbaia, Y.-H. He, E. Heyes, E. Hirst, V. Jejjala, Generating Triangulations and Fibrations With Reinforcement Learning (2024). arXiv:2405.21017.

[32] J. Bao, Y.-H. He, E. Heyes, E. Hirst, Machine Learning Algebraic Geometry for Physics (2022). arXiv:2204.10334.

[33] Y.-H. He, E. Heyes, E. Hirst, Machine Learning in Physics and Geometry, 2023. arXiv:2303.12626.

[34] M. Headrick, T. Wiseman, Numerical Ricci-flat Metrics on K3, Class. Quant. Grav. 22 (2005) 4931–4960. doi:10.1088/0264-9381/22/23/002.

[35] S. K. Donaldson, Some Numerical Results in Complex Differential Geometry (2005). arXiv:math/0512625.

[36] M. R. Douglas, R. L. Karp, S. Lukic, R. Reinbacher, Numerical Calabi-Yau Metrics, J. Math. Phys. 49 (2008) 032302. doi:10.1063/1.2888403.

[37] V. Braun, T. Brelidze, M. R. Douglas, B. A. Ovrut, Calabi-Yau Metrics for Quotients and Complete Intersections, JHEP 5 (2008) 80. doi:10.1088/1126-6708/2008/05/080.

[38] M. Headrick, A. Nassar, Energy Functionals for Calabi-Yau Metrics, Adv. Theor. Math. Phys. 17 (5) (2013) 867–902. doi:10.4310/ATMP.2013.v17.n5.a1.

[39] W. Cui, J. Gray, Numerical Metrics, Curvature Expansions and Calabi-Yau Manifolds, JHEP 5 (2020) 44. doi:10.1007/JHEP05(2020)044.

[40] L. B. Anderson, M. Gerdes, J. Gray, S. Krippendorf, N. Raghuram, F. Ruehle, Moduli-Dependent Calabi-Yau and SU(3)-Structure Metrics From Machine Learning, JHEP 5 (2021) 13. doi:10.1007/JHEP05(2021)013.

[41] A. Ashmore, Y.-H. He, B. A. Ovrut, Machine Learning Calabi-Yau Metrics, Fortsch. Phys. 68 (9) (2020) 2000068. doi:10.1002/prop.202000068.

[42] M. Douglas, S. Lakshminarasimhan, Y. Qi, Numerical Calabi-Yau Metrics From Holomorphic Networks, in: Proceedings of the 2nd Mathematical and Scientific Machine Learning Conference, Vol. 145 of Proceedings of Machine Learning Research, 2022, pp. 223–252.
URL https://proceedings.mlr.press/v145/douglas22a.html

[43] V. Jejjala, D. K. Mayorga Pena, C. Mishra, Neural Network Approximations for Calabi-Yau Metrics, JHEP 8 (2022) 105. doi:10.1007/JHEP08(2022)105.

[44] A. Ashmore, L. Calmon, Y.-H. He, B. A. Ovrut, Calabi-Yau Metrics, Energy Functionals and Machine-Learning, International Journal of Data Science in the Mathematical Sciences 1 (1) (2023) 49–61. doi:10.1142/S2810939222500034.

[45] M. Larfors, A. Lukas, F. Ruehle, R. Schneider, Learning Size and Shape of Calabi-Yau Spaces (2021). arXiv:2111.01436.

[46] M. Larfors, A. Lukas, F. Ruehle, R. Schneider, Numerical Metrics for Complete Intersection and Kreuzer-Skarke Calabi-Yau Manifolds, Mach. Learn. Sci. Tech. 3 (3) (2022) 035014. doi:10.1088/2632-2153/ac8e4e.

[47] M. Gerdes, S. Krippendorf, CYJAX: A Package for Calabi-Yau Metrics With JAX (2022). arXiv:2211.12520.

[48] P. Berglund, G. Butbaia, T. Hüubsch, V. Jejjala, D. Mayorga Peña, C. Mishra, J. Tan, Machine-learned Calabi-Yau Metrics and Curvature, Adv. Theor. Math. Phys. 27 (4) (2023) 1107–1158. doi:10.4310/ATMP.2023.v27.n4.a3.

[49] P. Berglund, Y.-H. He, E. Heyes, E. Hirst, V. Jejjala, A. Lukas, New Calabi-Yau Manifolds From Genetic Algorithms, Phys. Lett. B 850 (2024) 138504. doi:10.1016/j.physletb.2024.138504.

[50] Aggarwal, Daattavya and He, Yang-Hui and Heyes, Elli and Hirst, Edward and Earp, Henrique N. Sá and Silva, Tomás S. R., Machine Learning Sasakian and G2 Topology on Contact Calabi-Yau 7-Manifolds, Phys. Lett. B 850 (2024) 138517. doi:10.1016/j.physletb.2024.138517.

[51] A. Ashmore, Y.-H. He, E. Heyes, B. A. Ovrut, Numerical Spectra of the Laplacian for Line Bundles on Calabi-Yau Hypersurfaces, JHEP 7 (2023) 164. doi:10.1007/JHEP07(2023)164.

[52] D. Tong, Lectures on String Theory (2012). arXiv:0908.0333.

[53] B. Greene, String Theory on Calabi-Yau Manifolds (1997). arXiv:hep-th/9702155.

[54] N. B. Agmon, A. Bedroya, M. J. Kang, C. Vafa, Lectures on the String Landscape and the Swampland (2023). arXiv:2212.06187.

[55] H. F. Jones, Groups, Representations and Physics, 2nd ed., IOP Publishing, 1998.

[56] S. Sternberg, Group Theory and Physics, Cambridge University Press, 2008.

[57] A. Lukas, H. Tillim, Groups and Representations Lecture Notes. URL http://www-thphys.physics.ox.ac.uk/people/AndreLukas/GroupsandRepresentations/groupsrepslecturenotes.pdf

[58] E. P. Wigner, On Unitary Representations of the Inhomogeneous Lorentz Group, Annals Math. 40 (1939) 149–204. doi:10.2307/1968551.

[59] L. B. Anderson, M. Karkheiran, TASI Lectures on Geometric Tools for String Compactifications (2018). arXiv:1804.08792.

[60] J. Bao, Y.-H. He, E. Hirst, S. Pietromonaco, Lectures on the Calabi-Yau Landscape (2020). arXiv:2001.01212.

[61] S. K. Donaldson, Anti Self-Dual Yang-Mills Connections Over Complex Algebraic Surfaces and Stable Vector Bundles, Proceedings of the London Mathematical Society s3-50 (1) (1985) 1–26. doi:10.1112/plms/s3-50.1.1.

[62] K. Uhlenbeck, S. T. Yau, On the Existence of Hermitian-Yang-Mills Connections in Stable Vector Bundles, Communications on Pure and Applied Mathematics 39 (S1) (1986) S257–S293. doi:10.1002/cpa.3160390714.

[63] R. Friedman, J. W. Morgan, E. Witten, Vector Bundles Over Elliptic Fibrations (1997). arXiv:alg-geom/9709029.

[64] R. Donagi, A. Lukas, B. A. Ovrut, D. Waldram, Holomorphic Vector Bundles and Non-Perturbative Vacua in M-Theory, JHEP 1999 (06) (1999) 034–034. doi:10.1088/1126-6708/1999/06/034.

[65] R. Donagi, B. Ovrut, T. Pantev, D. Waldram, Standard-Model Bundles (2002). arXiv:math/0008010.

[66] L. B. Anderson, X. Gao, M. Karkheiran, Extending the Geometry of Heterotic Spectral Cover Constructions, Nucl. Phys. B 956 (2020) 115003. doi:10.1016/j.nuclphysb.2020.115003.

[67] L. Anderson, Y.-H. He, A. Lukas, Monad Bundles in Heterotic String Compactifications, JHEP 2008 (07) (2008) 104–104. doi:10.1088/1126-6708/2008/07/104.

[68] Y.-H. He, S.-J. Lee, A. Lukas, Heterotic Models From Vector Bundles on Toric Calabi-Yau Manifolds, JHEP 2010 (5) (2010). doi:10.1007/jhep05(2010)071.

[69] R. Blumenhagen, S. Moster, T. Weigand, Heterotic GUT and Standard Model Vacua from Simply Connected Calabi–Yau Manifolds, Nucl. Phys. B 751 (1–2) (2006) 186–221. doi:10.1016/j.nuclphysb.2006.06.005.

[70] V. Bouchard, R. Donagi, An SU(5) Heterotic Standard Model, Phys. Lett. B 633 (6) (2006) 783–791. doi:10.1016/j.physletb.2005.12.042.

[71] M. Larfors, D. Passaro, R. Schneider, Heterotic Line Bundle Models on Generalized Complete Intersection Calabi Yau Manifolds, JHEP 2021 (5) (2021). doi:10.1007/jhep05(2021)105.

[72] L. B. Anderson, J. Gray, A. Lukas, E. Palti, Heterotic line bundle standard models, JHEP 2012 (6) (2012). doi:10.1007/jhep06(2012)113.

[73] E. I. Buchbinder, A. Constantin, A. Lukas, The Moduli Space of Heterotic Line Bundle Models: A Case Study for the Tetra-Quadric, JHEP 2014 (3) (2014). doi:10.1007/jhep03(2014)025.

[74] D. Joyce, Complex Manifolds and Kähler Geometry (2022).
URL https://people.maths.ox.ac.uk/~joyce/KahlerGeom2022/index.html

[75] T. Hubsch, Calabi-Yau Manifolds: A Bestiary for Physicists, World Scientific, 1994.

[76] S. Karigiannis, Introduction to G2 Geometry, Springer US, 2020, pp. 3–50. doi:10.1007/978-1-0716-0577-6_1.

[77] E. Witten, Search for a Realistic Kaluza-Klein Theory, Nucl. Phys. B 186 (3) (1981) 412–428. doi:https://doi.org/10.1016/0550-3213(81)90021-3.

[78] W. V. D. Hodge, Harmonic Integrals: The Theory and Applications of Harmonic Integrals, Cambridge University Press, 1941.

[79] S. Bochner, Curvature and Betti Numbers, Annals Math. 49 (2) (1948) 379–390. doi:10.2307/1969287.

[80] K. Kodaira, On a Differential-Geometric Method in the Theory of Analytic Stacks, Proceedings of the National Academy of Sciences of the United States of America 39 (12) (1953) 1268–1273. doi:10.1073/pnas.39.12.1268.

[81] S. Nakano, On Complex Analytic Vector Bundles, Journal of the Mathematical Society of Japan 7 (1) (1955) 1–12. doi:10.2969/jmsj/00710001.

[82] J.-P. Demailly, Sur l'identite de Bochner-Kodaira-Nakano en geometrie hermitienne, in: P. Lelong, P. Dolbeault, H. Skoda (Eds.), Séminaire d'Analyse, Springer Berlin Heidelberg, 1986, pp. 88–97. doi:10.1007/BFb0077045.

[83] E. Calabi, The Space of Kähler Metrics, Proceedings of the International Congress of Mathematicians 2 (1954) 206–207.

[84] S. T. Yau, On the Ricci Curvature of a Compact Kähler Manifold and the Complex Monge–Ampère Equation. I, Communications on Pure and Applied Mathematics 31 (1978) 339–411. doi:10.1002/cpa.3160310304.

[85] P. Candelas, A. Dale, C. Lütken, R. Schimmrigk, Complete Intersection Calabi-Yau Manifolds, Nucl. Phys. B 298 (3) (1988) 493–525. doi:https://doi.org/10.1016/0550-3213(88)90352-5.

[86] M. Kreuzer, H. Skarke, Complete classification of Reflexive Polyhedra in Four Dimensions (2000). arXiv:hep-th/0002240.

[87] J. Gray, A. S. Haupt, A. Lukas, All Complete Intersection Calabi-Yau Four-Folds, JHEP 7 (2013) 70. doi:10.1007/JHEP07(2013)070.

[88] F. Schöller, H. Skarke, All Weight Systems for Calabi-Yau Fourfolds From Reflexive Polyhedra, Communications in Mathematical Physics 372 (2) (2019) 657–678. doi:10.1007/s00220-019-03331-9.

[89] S. Salamon, Riemannian Geometry and Holonomy Groups, Acta Applicandae Mathematica 20 (3) (1990) 309–311. doi:10.1007/BF00049573.

[90] M. Fernández, A. Gray, Riemannian Manifolds With Structure Group G2, Annali di Matematica Pura ed Applicata 132 (1982) 19–45. doi:10.1007/BF01760975.

[91] C. M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.
URL https://link.springer.com/book/9780387310732

[92] I. Goodfellow, Y. Bengio, A. Courville, Deep Learning, The MIT Press, 2016.

[93] A. Ng, G. Ladwig, A. Bagul, E. Shyu, Machine Learning Specialization, Machine-Learning Introduction Coursera Course.

[94] P. Mehta, M. Bukov, C.-H. Wang, A. G. Day, C. Richardson, C. K. Fisher, D. J. Schwab, A High-Bias, Low-Variance Introduction to Machine Learning for Physicists, Physics Reports 810 (2019) 1–124. doi:10.1016/j.physrep.2019.03.001.

[95] K. S. Ganesh, What's The Role Of Weights And Bias In a Neural Network?
URL https://towardsdatascience.com/whats-the-role-of-weights-and-bias-in-a-neural-ne

[96] N. Qian, On the Momentum Term in Gradient Descent Learning Algorithms, Neural Networks 12 (1) (1999) 145–151. doi:https://doi.org/10.1016/S0893-6080(98)00116-6.

[97] J. Duchi, E. Hazan, Y. Singer, Adaptive Subgradient Methods for Online Learning and Stochastic Optimization, J. Mach. Learn. Res. 12 (2011) 2121–2159.

[98] G. Hinton, Coursera Neural Networks for Machine Learning, lecture 6 (2018).
URL https://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf

[99] D. P. Kingma, J. Ba, Adam: A Method for Stochastic Optimization (2017). arXiv:1412.6980.

[100] A. M. Turing, Computing Machinery and Intelligence, Mind 59 (236) (1950) 433.

[101] N. A. Barricelli, Esempi Numerici di Processi di Evoluzione, Methodos 6 (21-22) (1954) 45–68.

[102] J. Holland, Adaptation in Natural and Artificial Systems, The MIT Press, 1992.

[103] C. Darwin, On the Origin of Species by Means of Natural Selection, or the Preservation of Favored Races in the Struggle for Life, Murray, 1859.

[104] N. Makke, S. Chawla, Interpretable Scientific Discovery With Symbolic Regression: A Review (2023). arXiv:2211.10873.

[105] M. R. Douglas, The Statistics of String/M Theory Vacua, JHEP 2003 (5) (2003) 46–46. doi:10.1088/1126-6708/2003/05/046.

[106] S. K. Ashok, M. R. Douglas, Counting Flux Vacua, JHEP 2004 (1) (2004) 60–60. doi:10.1088/1126-6708/2004/01/060.

[107] W. Taylor, Y.-N. Wang, The F-Theory Geometry With Most Flux Vacua, JHEP 2015 (12) (2015) 1–21. doi:10.1007/jhep12(2015)164.

[108] V. V. Batyrev, Dual Polyhedra and Mirror Symmetry for Calabi-Yau Hypersurfaces in Toric Varieties, J. Alg. Geom. 3 (1994) 493–545. arXiv:alg-geom/9310003.

[109] V. V. Batyrev, L. A. Borisov, On Calabi-Yau Complete Intersections in Toric Varieties (1994). arXiv:alg-geom/9412017.

[110] M. Kreuzer, H. Skarke, On the Classification of Reflexive Polyhedra, Communications in Mathematical Physics 185 (2) (1997) 495–508. doi:10.1007/s002200050100.

[111] M. Kreuzer, H. Skarke, Classification of Reflexive Polyhedra in Three Dimensions (1998). arXiv:hep-th/9805190.

[112] H. Skarke, Weight Systems for Toric Calabi-Yau Varieties and Reflexivity of Newton Polyhedra, Modern Physics Letters A 11 (20) (1996) 1637–1652. doi:10.1142/s0217732396001636.

[113] M. Kreuzer, H. Skarke, Calabi-Yau Data.
URL https://hep.itp.tuwien.ac.at/~kreuzer/CY/

[114] F. Schöller, H. Skarke, 5d Calabi-Yau Weight Systems.
URL https://huggingface.co/datasets/calabi-yau-data/ws-5d

[115] J. Bao, Y.-H. He, E. Hirst, J. Hofscheier, A. Kasprzyk, S. Majumder, Polytopes and Machine Learning, Math. Sci. 1 (2023) 181–211. doi:10.1142/S281093922350003X.

[116] D. S. Berman, Y.-H. He, E. Hirst, Machine Learning Calabi-Yau Hypersurfaces, Phys. Rev. D 105 (6) (2022) 066002. doi:10.1103/PhysRevD.105.066002.

[117] E. Hirst, T. S. Gherardini, Calabi-Yau Four-, Five-, Sixfolds as $\mathbb{P}^n_{\mathbf{w}}$ Hypersurfaces: Machine Learning, Approximation, and Generation, Phys. Rev. D 109 (10) (2024) 106006. doi:10.1103/PhysRevD.109.106006.

[118] S. A. Abel, J. Rizos, Genetic Algorithms and the Search for Viable String Vacua, JHEP 10 (2014) 1029–8479. doi:10.1007/JHEP08(2014)010.

[119] S. Abel, A. Constantin, T. R. Harvey, A. Lukas, String Model Building, Reinforcement Learning and Genetic Algorithms (2021). arXiv:2111.07333.

[120] B. I. U. Nødland, Generating Refelexive Polytopes via Sequence Modeling (2022). URL https://mathai2022.github.io/papers/3.pdf

[121] N. MacFadden, A. Schachner, E. Sheridan, The DNA of Calabi-Yau Hypersurfaces (2024). arXiv:2405.08871.

[122] W. Fulton, Introduction to Toric Varieties, Princeton University Press, 1993. URL https://www.jstor.org/stable/j.ctt1b7x7vc

[123] M. Kreuzer, H. Skarke, PALP: A Package for Analysing Lattice Polytopes With Applications to Toric Geometry, Computer Physics Communications 157 (1) (2004) 87–106. doi:10.1016/s0010-4655(03)00491-0.

[124] P. Berglund, Y.-H. He, E. Heyes, E. Hirst, V. Jejjala, A. Lukas, Reflexive Polytope Genetic Algorithm.
URL https://github.com/elliheyes/Polytope-Generation

[125] J. Gray, A. S. Haupt, A. Lukas, Complete Intersection Calabi-Yau Four-Folds.
URL https://www-thphys.physics.ox.ac.uk/projects/CalabiYau/Cicy4folds/index.html

[126] M. Berg, M. Haack, H. Samtleben, Calabi-Yau Fourfolds With Flux and Supersymmetry Breaking, JHEP 2003 (04) (2003) 046–046. doi:10.1088/1126-6708/2003/04/046.

[127] M. J. Duff, B. E. W. Nilsson, C. N. Pope, Spontaneous Supersymmetry Breaking by the Squashed Seven-Sphere, Phys. Rev. Lett. 50 (1983) 2043–2046. doi:10.1103/PhysRevLett.50.2043.

[128] B. Acharya, E. Witten, Chiral Fermions From Manifolds of G2 Holonomy (2001). arXiv:hep-th/0109152.

[129] B. Acharya, S. Gukov, M-Theory and Singularities of Exceptional Holonomy Manifolds, Physics Reports 392 (3) (2004) 121–189.

[130] A. Tomassini, L. Vezzoni, Contact Calabi-Yau Manifolds and Special Legendrian Submanifolds (2007). arXiv:math/0612232.

[131] G. Habib, L. Vezzoni, Some Remarks on Calabi-Yau and Hyper-Kähler Foliations, Differential Geometry and its Applications 41 (2015) 12–32. doi:10.1016/j.difgeo.2015.03.006.

[132] X. de la Ossa, M. Larfors, E. E. Svanes, Exploring $SU(3)$ Structure Moduli Spaces With Integrable G2 Structures, Adv. Theor. Math. Phys. 19 (2015) 837–903. doi:10.4310/ATMP.2015.v19.n4.a5.

[133] X. de la Ossa, M. Larfors, E. E. Svanes, The Infinitesimal Moduli Space of Heterotic G2 Systems, Commun. Math. Phys. 360 (2) (2018) 727–775. doi:10.1007/s00220-017-3013-8.

[134] X. de la Ossa, M. Larfors, E. E. Svanes, Restrictions of Heterotic G2 Structures and Instanton Connections, in: Nigel Hitchin's 70th Birthday Conference, 2017. arXiv:1709.06974.

[135] J. Lotay, H. S. Earp, The Heterotic G2 System on Contact Calabi-Yau 7-Manifolds, Transactions of the American Mathematical Society, Series B 10 (26) (2023) 907–943. doi:10.1090/btran/129.

[136] O. Calvo-Andrade, L. Rodríguez, H. N. Sá Earp, Gauge Theory and G2-Geometry on Calabi-Yau Links, Rev. Mat. Iberoam. 36 (6) (2020) 1753–1778.
URL https://www.ems-ph.org/journals/show_pdf.php?issn=0213-2230&vol=36&iss=6&rank=6

[137] J. Milnor, Singular Points of Complex Hypersurfaces, Princeton University Press, 1969.
URL http://www.jstor.org/stable/j.ctt1bd6kvv

[138] P. Candelas, M. Lynker, R. Schimmrigk, Calabi-Yau Manifolds in Weighted P4, Nucl. Phys. B 341 (2) (1990) 383–402. doi:https://doi.org/10.1016/0550-3213(90)90185-G.

[139] M. Itoh, Sasakian Manifolds, Hodge Decomposition and Milnor Algebras, Kyushu J. Math 58 (2004) 121–140.

URL https://www.jstage.jst.go.jp/article/kyushujm/58/1/58_1_121/_article/-char/ja/

[140] D. Crowley, J. Nordström, New invariants of G2-Structures, Geometry and Topology 19 (2015) 2949–2992.

[141] C. Vafa, String Vacua and Orbifoldized L-G Models, Mod. Phys. Lett. A 4 (1989) 1169. doi:10.1142/S0217732389001350.

[142] D. Aggarwal, Y.-H. He, E. Heyes, E. Hirst, H. N. S. Earp, T. S. R. Silva, MLcCY7 (2023). URL https://github.com/TomasSilva/MLcCY7

[143] S. Sasaki, Y. Hatakeyama, On Differentiable Manifolds With Certain Structures Which Are Closely Related to Almost Contact Structure, II, Tohoku Mathematical Journal 13 (2) (1961) 281–294. doi:10.2748/tmj/1178244304.

[144] T. S. Developers, W. Stein, D. Joyner, D. Kohel, J. Cremona, B. Eröcal, SageMath, version 9.0, http://www.sagemath.org (2020).

[145] D. R. Grayson, M. E. Stillman, Macaulay2, a software system for research in algebraic geometry, http://www.math.uiuc.edu/Macaulay2/ (1992).

[146] W. Decker, G.-M. Greuel, G. Pfister, H. Schönemann, SINGULAR 4-3-0 — A computer algebra system for polynomial computations, http://www.singular.uni-kl.de (2022).

[147] J. Steenbrink, Intersection Form for Quasi-Homogeneous Singularities, Compositio Mathematica 34 (2) (1977) 211–223.

[148] J. H. M. Steenbrink, Mixed Hodge Structures Associated With Isolated Singularities (1983). doi:10.1090/PSPUM/040.2/713277.

[149] V. V. Batyrev, On the Stringy Hodge Numbers of Mirrors of Quasi-Smooth Calabi-Yau Hypersurfaces (2020). arXiv:2006.15825.

[150] M. Cranmer, Interpretable Machine Learning for Science with PySR and SymbolicRegression.jl (2023). doi:10.48550/arXiv.2305.01582.

[151] J. Halverson, F. Ruehle, Metric Flows with Neural Networks (2023). arXiv:2310.19870.

[152] J. D. Lotay, H. N. Sá Earp, J. Saavedra, Flows of G2-Structures on Contact Calabi–Yau 7-Manifolds, Annals of Global Analysis and Geometry 62 (2) (2022) 367—389. doi:10.1007/s10455-022-09854-0.

[153] M. R. Douglas, R. L. Karp, S. Lukic, R. Reinbacher, Numerical Solution to the Hermitian Yang-Mills Equation on the Fermat Quintic, JHEP 12 (2007) 83. doi:10.1088/1126-6708/2007/12/083.

[154] L. B. Anderson, V. Braun, R. L. Karp, B. A. Ovrut, Numerical Hermitian Yang-Mills Connections and Vector Bundle Stability in Heterotic Theories, JHEP 6 (2010) 107. doi:10.1007/JHEP06(2010)107.

[155] L. B. Anderson, V. Braun, B. A. Ovrut, Numerical Hermitian Yang-Mills Connections and Kahler Cone Substructure, JHEP 1 (2012) 14. doi:10.1007/JHEP01(2012)014.

[156] W. Cui, Numerical Hermitian Yang-Mills Connection for Bundles on Quotient Manifold (2023). arXiv:2302.09622.

[157] A. Ashmore, R. Deen, Y.-H. He, B. A. Ovrut, Machine Learning Line Bundle Connections, Phys. Lett. B 827 (2022) 136972. doi:10.1016/j.physletb.2022.136972.

[158] A. Strominger, E. Witten, New Manifolds for Superstring Compactification, Commun. Math. Phys. 101 (1985) 341. doi:10.1007/BF01216094.

[159] A. Strominger, Yukawa Couplings in Superstring Compactification, Phys. Rev. Lett. 55 (1985) 2547. doi:10.1103/PhysRevLett.55.2547.

[160] V. Braun, T. Brelidze, M. R. Douglas, B. A. Ovrut, Eigenvalues and Eigenfunctions of the Scalar Laplace Operator on Calabi-Yau Manifolds, JHEP 7 (2008) 120. doi:10.1088/1126-6708/2008/07/120.

[161] A. Ashmore, Eigenvalues and Eigenforms on Calabi-Yau Threefolds (11 2020). arXiv:2011.13929.

[162] R. Kuwabara, Spectrum of the Schrödinger Operator on a Line Bundle Over Complex Projective Spaces, Tohoku Mathematical Journal 40 (2) (1988) 199–211. doi:10.2748/tmj/1178228026.

[163] D. Bykov, A. Smilga, Monopole Harmonics on $\mathbb{CP}^{n-1}$ (2023). arXiv:2302.11691.

[164] J. Milnor, Eigenvalues of the Laplace Operator on Certain Manifolds, PNAS 51 (4) (1964) 542–542. doi:10.1073/pnas.51.4.542.

[165] H. Ahmed, F. Ruehle, Level Crossings, Attractor Points and Complex Multiplication, JHEP 164 (6) (2023). doi:10.1007/JHEP06(2023)164.

[166] C. Tejero Prieto, Holomorphic Spectral Geometry of Magnetic Schrödinger Operators on Riemann Surfaces, Differential Geometry and its Applications 24 (3) (2006) 288–310. doi:https://doi.org/10.1016/j.difgeo.2005.09.001.

[167] J. C. Lagarias, G. M. Ziegler, Bounds for Lattice Polytopes Containing a Fixed Number of Interior Points in a Sublattice, Canadian Journal of Mathematics 43 (5) (1991) 1022–1035. doi:10.4153/CJM-1991-058-4.

[168] R. Grinis, A. Kasprzyk, Normal Forms of Convex Lattice Polytopes (2013). arXiv:1301.6641.