Physics Letters B 827 (2022) 136972

Contents lists available at ScienceDirect

Physics Letters B

www.elsevier.com/locate/physletb

Machine learning line bundle connections

Anthony Ashmore [a,b,*], Rehan Deen [c], Yang-Hui He [d,e,f,g], Burt A. Ovrut [h]

[a] Kadanoff Center for Theoretical Physics, University of Chicago, IL 60637, USA
[b] Sorbonne Université, CNRS, LPTHE, F-75005 Paris, France
[c] Rudolf Peierls Centre for Theoretical Physics, University of Oxford, OX1 3PU, UK
[d] London Institute for Mathematical Sciences, Royal Institution, W1S 4BS, UK
[e] Department of Mathematics, City, University of London, EC1V0HB, UK
[f] Merton College, University of Oxford, OX1 4JD, UK
[g] School of Physics, NanKai University, Tianjin, 300071, PR China
[h] Department of Physics, University of Pennsylvania, Philadelphia, PA 19104, USA

ARTICLE INFO

Article history:
Received 25 January 2022
Received in revised form 15 February 2022
Accepted 16 February 2022
Available online 22 February 2022
Editor: A. Volovich

Keywords:
Yang–Mills
Machine learning
Connections

ABSTRACT

We study the use of machine learning for finding numerical hermitian Yang–Mills connections on line bundles over Calabi–Yau manifolds. Defining an appropriate loss function and focusing on the examples of an elliptic curve, a K3 surface and a quintic threefold, we show that neural networks can be trained to give a close approximation to hermitian Yang–Mills connections.

1. Introduction and summary

Heterotic string theory on Calabi–Yau threefolds equipped with gauge bundles provides a large class of phenomenologically promising string models [1–12]. However, despite many decades of work, it is still not possible to compute the masses or couplings that appear in the resulting four-dimensional theories from first principles. A good deal of the problem can be traced to the lack of explicit expressions for non-trivial Calabi–Yau (CY) metrics or hermitian Yang–Mills connections. Let us recall why these are needed. Compactification of the heterotic string on a Calabi–Yau threefold X with its Ricci-flat metric gives a four-dimensional effective theory with $N = 1$ supersymmetry. To obtain MSSM-like theories, X should also carry a holomorphic vector bundle V whose connection solves the hermitian Yang–Mills (HYM) equations [13,14].

Generic details of the compactification, such as the number of generations or the vanishing of certain couplings, can be obtained from algebro-geometric results for the existence and topology of the threefold X and the bundle V [15–22]. These calculations do not need explicit expressions for either the metric or the connection. Unfortunately, the detailed four-dimensional physics is controlled by a Kähler potential and a superpotential, which depend on both the explicit Ricci-flat metric and the explicit HYM connection. Without this data, it is generally not possible to accurately compute masses or couplings, leaving us unable to make precise particle physics predictions from string theory.

With little hope of finding analytic expressions for the relevant metrics or connections, much progress has been made on finding *numerical* approximations. There is now a diverse range of algorithms for computing Ricci-flat metrics on Calabi–Yau manifolds numerically, including position space methods [23], spectral approaches [24–26] building on the work of Tian [27] and Donaldson [28], and, most recently machine learning [29] and neural networks tailored for the metric computation [30–33] (see [34] for a recent pedagogical review on Calabi–Yau manifolds and machine-learning).

Given these advances, it now seems appropriate to focus on computing hermitian Yang–Mills connections. This will be the subject of the present work. As with the Ricci-flatness condition for the metric, the HYM equations are a system of partial differential equations that are difficult to solve, and so one is again compelled to consider numerical approximations. Previous work [35–37] has used Wang's extension [38] of Donaldson's approach to compute

* Corresponding author.
E-mail addresses: ashmore@uchicago.edu (A. Ashmore), rehan.deen@gmail.com (R. Deen), hey@maths.ox.ac.uk (Y.-H. He), ovrut@elcapitan.hep.upenn.edu (B.A. Ovrut).

https://doi.org/10.1016/j.physletb.2022.136972

numerical HYM connections for a number of examples, including SU(n) bundles over threefolds.

Unfortunately, connections are a jump in computational complexity compared to the Ricci-flat metric, with a corresponding loss of speed and accuracy. With this in mind, we seek a faster and more accurate method by employing machine learning. The aim of this paper is to take the first step in applying machine learning to find HYM connections. Focusing on the simplest cases of connections on line bundles, we show that it is both feasible and promising to compute connections in this way. Though we do not tackle non-abelian bundles in the present work, we note that many Standard Model-like theories can be obtained from heterotic line bundle models [10,11,39,12,40–43]

Our approach builds on and extends the work of Douglas et al. [31] which presented a neural network for computing Calabi–Yau metrics (they have provided a TensorFlow implementation of their approach on GitHub [44]). We give three examples, namely line bundles over an elliptic curve, a K3 surface, and a quintic threefold. For each of these, one starts by computing a numerical approximation to the Calabi–Yau metric. One then constructs a neural network whose input is the coordinates on the Calabi–Yau and whose output is interpreted as the hermitian metric on the line bundle. By taking derivatives of the neural network, one can compute both the connection and the curvature defined by the hermitian metric. We give a loss function whose value is minimised for hermitian Yang–Mills connections, and then use this loss function to train the network. The resulting network encodes the hermitian metric that defines a HYM connection on the Calabi–Yau. In this way, we find that accurate HYM connections can be obtained in a straightforward manner. In our results, we examine how the accuracy of the numerical connections changes with varying network depth. We observe that deeper networks are generally more accurate (as expected since they contain more parameters), with this improvement more pronounced as the dimension of the Calabi–Yau increases.

There are a number of obvious extensions. First, one could consider line bundles over more complicated Calabi–Yau manifolds. This would involve generalising the code of Douglas et al. [31] to complete intersections in products of projective space. Second, one would want to move beyond abelian bundles to consider non-abelian bundles, defined by monads, extensions, and so on. Both of these are essential if one wants to make contact with the many constructions of the so-called heterotic Standard Model [45–47,4,10,48,11,39,12,49,40,50,51]. These advances, together with numerical metrics and Laplacians [52–54], and results for the matter-field Kähler potential [55–58], should enable real progress on computing masses and couplings in top-down string models. We will discuss these issues in future publications.

2. Hermitian Yang–Mills and line bundles on Calabi–Yau manifolds

Given a complex manifold X with a Kähler metric g (defined by a choice of complex structure and Kähler form J), a stable holomorphic vector bundle V admits a unique connection A whose curvature F solves the hermitian Yang–Mills equations:

$$F_{ij} = F_{\bar{i}\bar{j}} = 0, \qquad g^{i\bar{j}} F_{i\bar{j}} = \mu(V)\,\mathbf{1}. \tag{1}$$

Here $g^{i\bar{j}}$ is the inverse Kähler metric on X, $\mu(V)$ is a real constant known as the *slope* of V, and $\mathbf{1}$ is the $d \times d$ identity matrix on the fibres of the rank-d bundle V. The first two conditions are equivalent to the holomorphicity of V (and will be automatic in our construction). The third condition gives the HYM equations, a system of non-linear PDEs for the connection A.

The connection A can equivalently be described by a hermitian structure on V, which, more prosaically, is simply a hermitian inner product G on sections of V. Given a frame $\{e_a\}$ for V, the inner product is

$$(e_a, e_b) = G_{\bar{a}b}, \qquad G = G^\dagger. \tag{2}$$

In holomorphic gauge, the connection is determined by G as

$$A_i = G^{-1}\partial_i G, \qquad A_{\bar{i}} = 0, \tag{3}$$

with the curvature then given by

$$F_{i\bar{j}} = \partial_{\bar{j}}\partial_i \log G, \tag{4}$$

where we are using the shorthand notation $\partial_i \log G \equiv G^{-1}\partial_i G$. Given a Kähler metric, finding a solution to the HYM equations then reduces to choosing G such that (1) is satisfied. If this is the case, G is known as a *Hermite–Einstein* metric on V.

There exists a solution to the HYM equations on a Kähler manifold if and only if the holomorphic vector bundle V is (at least) polystable [13,14]. To check this, one begins by computing the slope of V via

$$\mu(V) \equiv \int_X c_1(V) \wedge J^{n-1}. \tag{5}$$

Note that we always normalise Vol_g, the volume of X as measured by the Kähler metric, to one. The bundle V is stable if $\mu(\mathcal{F}) < \mu(V)$ for all subsheaves $\mathcal{F} \subset V$ with $0 < \mathrm{rank}\,\mathcal{F} < \mathrm{rank}\,V$. Polystability is the statement that V is a direct sum of stable bundles, all with the same slope. Thanks to this, the *existence* of a HYM connection can be reduced to algebraic conditions on subsheaves of V. Notice however that this is in no way constructive; that is, knowing a HYM connection exists does not give any hint of how to find it explicitly. For this we must turn to numerical methods. The aim of the present work is to use a neural network to search for numerical HYM solutions for the simplest examples, namely line bundles.

Line bundles on CY manifolds are by now a well-understood ingredient in heterotic compactifications (see, for example, [10–12, 39] and references therein). Recall that a holomorphic line bundle L over a complex manifold is determined (up to torsion) by its first Chern class, $c_1(L)$. Thanks to this, we can associate a line bundle over X to a divisor $\mathcal{D}$ by taking

$$c_1(L) \equiv \frac{[F]}{2\pi} = \mathcal{D}, \tag{6}$$

where $[F]$ is the class of the curvature of the connection on L. The corresponding line bundle is then denoted by $\mathcal{O}_X(\mathcal{D})$, or often by $\mathcal{O}_X(k^I)$, where $\mathcal{D} = k^I\mathcal{D}_I$ and the basis of divisors is implicit. The slope of a line bundle is then

$$\mu(L) = \int_X c_1(L) \wedge J^{n-1}, \tag{7}$$

which depends on both the choice of line bundle via $c^1(L)$ and the choice of Kähler moduli via J. Since a line bundle has no subsheaves $\mathcal{F} \subset L$ with $0 < \mathrm{rank}\,\mathcal{F} < 1$, line bundles are always stable. This means that a line bundle will always admit a connection that solves the HYM equation, $g^{i\bar{j}}F_{i\bar{j}} = \mu(L)$. The problem is finding the explicit form of this connection.

Fig. 1. A $D=2$ network on an elliptic curve, whose output should be interpreted either as the Kähler potential, K, or log of the inverse bundle metric, $\log G^{-1}$, depending on whether one is computing the Calabi–Yau metric or the hermitian Yang–Mills connection. Here, "Bihom" refers to a bihomogenous layer which takes $z_i = (z_0, z_1, z_2)$ as input and outputs the real and imaginary parts of $z_i \bar{z}_{\bar{j}}$. "Square" is a dense layer with a quadratic activation function, $\vec{x} \mapsto (W_1 \vec{x})^2$, where W_1 is a general linear transformation of dimension $W^{(1)} \times 9$. "Log" is a dense layer with a log activation function, $\vec{x} \mapsto \log(W_2 \vec{x})$, where W_2 is a general linear transformation of dimension $1 \times W^{(1)}$.

3. Numerical metrics and connections from neural networks

In this section, we begin by reviewing the calculation of numerical Calabi–Yau metrics using a neural network following Douglas et al. [31].[1] We then discuss numerical HYM connections. Finally, we define a functional which is minimised on HYM connections and thus can act as a loss function for a suitable neural network whose output will be interpreted as $\log G^{-1}$.

3.1. Numerical metrics from neural networks

Consider a compact Calabi–Yau n-fold X defined as a hypersurface in $\mathbb{P}^{n+1}$ by the vanishing of a holomorphic equation $f(z) = 0$ of degree $n+2$ (for example, a threefold defined by a quintic equation in $\mathbb{P}^4$). The choice of defining equation f fixes the complex structure moduli of the Calabi–Yau. A choice of Kähler structure then determines the metric $g_{i\bar{j}}$ on X. In other words, the metric on X is fixed by a Kähler potential $K(z, \bar{z})$. Finding the Ricci-flat metric on X then amounts to choosing K such that the resulting Kähler metric is Ricci flat.

Apart from on the torus, there are no explicitly known Kähler potentials that give such Ricci-flat metrics. Instead, work has mostly focused on finding numerical approximations starting from a Fubini–Study-like ansatz for K:

$$K = \frac{1}{k\pi} \log s_\alpha h^{\alpha\bar{\beta}} \bar{s}_{\bar{\beta}}, \tag{8}$$

where the s_α are sections of $\mathcal{O}_X(k)$ (homogeneous functions of the coordinates z of degree k modulo $f = 0$), and $h^{\alpha\bar{\beta}}$ is a hermitian matrix of parameters. One then varies the parameters so that the resulting Kähler metric is as close as possible to Ricci flat. Increasing the degree k increases the size of the matrix $h^{\alpha\bar{\beta}}$, allowing a better approximation of the honest Calabi–Yau metric. There are now a variety of schemes for choosing $h^{\alpha\bar{\beta}}$, including via balanced metrics [28,24,25], direct optimisation [23,26] and neural networks [30].

The work of Douglas et al. [31] follows a similar path but uses a neural network to compute the Kähler potential directly (see also [30,32] for similar approaches). The network is a series of densely connected layers $L^{(i)}$ of depth D and width $W^{(i)}$ with quadratic activation functions $\theta^{(i)} : x \mapsto x^2$. The final layer, $L^{(D)}$, has width $W^{(D)} = 1$ and a log activation function, $\theta^{(D)} : x \mapsto \log x$. The output of the network can thus be thought of as the logarithm of a homogeneous scalar function of the inputs, with the coefficients that appear in this function fixed by the collective weights $\boldsymbol{\nu}$ of the network. A diagram of this network structure is shown in Fig. 1.

The inputs to the network are coordinates on the Calabi–Yau hypersurface, given as points $z_i = [z_0 : \ldots : z_{n+1}]$ in the ambient projective space $\mathbb{P}^{n+1}$, which can thus be thought of as sections of $\mathcal{O}_X(1)$, i.e. elements of $H^0(X, \mathcal{O}_X(1))$. In practice, the first layer is actually a "bihomogeneous layer" which converts the inputs z_i to the real and imaginary parts of $z_i \bar{z}_{\bar{j}}$, allowing one to work with real quantities. The successive layers have activation functions which square the output of each layer, so that the network essentially constructs the tensor product

$$\bigotimes_1^{D-1} \mathcal{O}_X(2) = \mathcal{O}_X(2^{D-1}). \tag{9}$$

Thus the output of the penultimate layer represents elements of $H^0(X, \mathcal{O}_X(2^{D-1}))$. Together with the final layer, the network output is K, the Kähler potential, with the precise way that elements of $H^0(X, \mathcal{O}_X(2^{D-1}))$ are combined fixed by the weights $\boldsymbol{\nu}$. The output of the network can then be used to compute a Kähler metric on the hypersurface. The aim is then to choose the weights $\boldsymbol{\nu}$ so that the resulting metric is as close as possible to Ricci flat.

The network is trained by minimising the pointwise difference between the measure defined by the (explicitly known) holomorphic $(n,0)$-form, $\mathrm{vol}_\Omega = \mathrm{i}\,\Omega \wedge \bar{\Omega}$, and the measure defined by the Kähler metric on X, $\mathrm{vol}_g = J^n$ (using K computed by the network). The two quantities agree pointwise if and only if the metric is the honest Ricci-flat metric.

In outline, training proceeds as follows. First, a training set and a test set, each containing 10,000 points lying on the Calabi–Yau hypersurface, are generated. The training points (and data about coordinate patches, the $(n,0)$-form and the point distribution) are passed to the network in batches of 1,000 in a training round. The Kähler metric defined by the network is given by the complex Hessian of the network output, $g_{i\bar{j}}(\boldsymbol{\nu}) \sim \partial_i \partial_{\bar{j}} K(\boldsymbol{\nu})$, where $\boldsymbol{\nu}$ denotes the weights of the network. The loss function is simply the mean absolute percentage error (MAPE), summed over the points in the training round:

$$\sigma(\boldsymbol{\nu}) = \int_X \left| 1 - \frac{\mathrm{vol}_g(\boldsymbol{\nu})}{\mathrm{vol}_\Omega} \right| \mathrm{vol}_\Omega, \tag{10}$$

where here, and in what follows, we normalise the integrated volumes, Vol_g and Vol_Ω, to one. Note that this is known as the "σ measure" in [25] and later work. One then searches for the minimum of this function in weight space, using stochastic gradient descent to update the weights after each training round. After 500 epochs, the network has usually converged to an approximately Ricci-flat Kähler potential. The accuracy of the resulting network can then be checked by evaluating $\sigma(\boldsymbol{\nu})$ on the test set.

It is simple to see how the numerical accuracy of the approximation can be increased. From (9), a deeper network provides a higher-degree expansion of the Kähler potential with more parameters (weights), both of which should allow a better approximation of the Ricci-flat metric. A wider network increases only the number of parameters (weights).

3.2. Numerical connections from neural networks

Following [35–37], one can calculate numerical HYM connections by starting with an ansatz similar in spirit to (8) but now for the hermitian structure G as

$$(G^{-1})^{a\bar{b}} = \sum_{\alpha,\beta}^{N_k} S^a_\alpha H^{\alpha\bar{\beta}} \bar{S}^{\bar{b}}_{\bar{\beta}}, \tag{11}$$

where S^a_α are sections of $V \otimes \mathcal{O}_X(k)$ and $H^{\alpha\bar{\beta}}$ is a hermitian matrix of parameters. In principle, one then varies these parameters to find an approximate solution to the HYM equation (1). The result of this is the hermitian metric, and hence connection, on the

[1] See also [33] for a discussion of the "holomorphic feedforward networks" that underlie this approach.

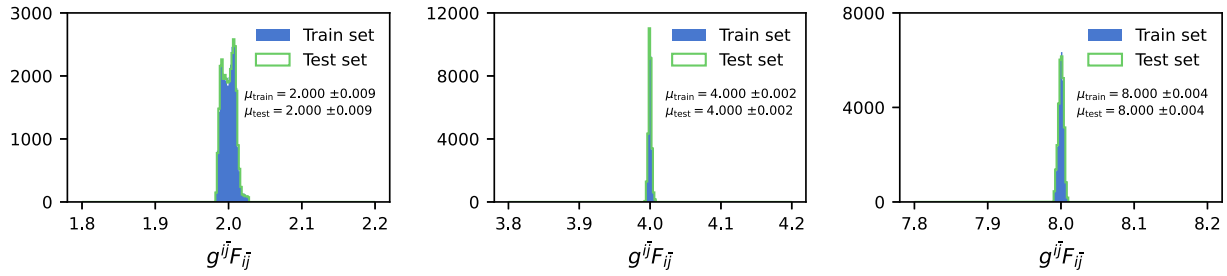

Fig. 2. Results for line bundle connections on an elliptic curve trained using Loss[$\boldsymbol{\nu}$]. The plots show the histogram of $g^{i\bar{j}}F_{i\bar{j}}$ evaluated for sample points on the elliptic curve in both the training and test sets. The left, middle and right plots are for networks of depth $D = 2, 3, 4$ respectively, which correspond to connections on the line bundles $\mathcal{O}_X(2)$, $\mathcal{O}_X(4)$ and $\mathcal{O}_X(8)$.

bundle $V(k) \equiv V \otimes \mathcal{O}_X(k)$. Since we are interested in the connection on V alone, one should subtract the contribution of $\mathcal{O}_X(k)$. As discussed in [36,37], the optimal way to do this is to take the metric on $\mathcal{O}_X(k)$ to be that induced by $\det G$. For the case where V is a line bundle, one does not encounter this complication as the connection on $V = \mathcal{O}_X(m)$ is simple to recover from the connection on $\mathcal{O}_X(m+k)$. As with the metric, increasing k increases the number of sections S^a_α and hence the number of parameters in $H^{\alpha\bar{\beta}}$, so that larger values of k allow for a better approximation to the honest HYM connection.

Our idea is to use the structure of a neural network to mimic the construction of G^{-1}, using the coordinates and activation functions to reproduce the sections, with the weights standing in for the parameters.

We will focus on the example of rank-one bundles, i.e. line bundles. In this case, the hermitian fibre metric on the line bundle is a scalar

$$G^{-1} = \sum_{\alpha,\beta}^{N_k} S_\alpha H^{\alpha\bar{\beta}} \bar{S}_{\bar{\beta}}, \tag{12}$$

with the curvature given by

$$F_{i\bar{j}} = \partial_{\bar{j}}\partial_i \log G = -\partial_{\bar{j}}\partial_i \log G^{-1}. \tag{13}$$

We will treat the output of the neural network as $\log G^{-1}$, from which it is simple to calculate $F_{i\bar{j}}$.

The structure of the connection network is the same as that of Fig. 1, with the depth of the network controlling the value of $m+k$. As with the metric network, the inputs are the points on the Calabi–Yau hypersurface, given as points on $\mathbb{P}^{n+1}$, and the output of the network should be identified with $\log G^{-1}$. For each training round, one computes $F_{i\bar{j}}$ as the complex Hessian of the output of the network. One then uses a previously trained metric network to compute the Ricci-flat metric, and combines this with $F_{i\bar{j}}$ into an appropriate loss function, which we give in the next subsection. Training then attempts to minimise this loss function to find a numerical approximation to the HYM connection (for a given choice of Ricci-flat metric). After sufficient training rounds, one has a neural network that is equivalent to $\log G^{-1}$ as a function of coordinates. A schematic of this structure is given in the appendices in Fig. A.5.

3.3. Loss function

We now introduce an accuracy measure for HYM connections that will serve as a loss function for the neural network. As a notational convenience, we define the contraction of g with F to be the scalar $F_g \equiv g^{i\bar{j}}F_{i\bar{j}}$ valued in endomorphisms of the gauge group; that is, for a rank-d bundle V, at a point on X, F_g is a $d \times d$ matrix. With this notation the HYM equation is simply

$$F_g = \mu(V)\,\mathbf{1}. \tag{14}$$

We also define the expectation $\langle O \rangle$ of a quantity O to be its average over the Calabi–Yau X using the exact CY measure vol_Ω – for example, the expectation of $\text{tr}\, F_g$ is defined to be

$$\langle \text{tr}\, F_g \rangle \equiv \int_X \text{vol}_\Omega\, \text{tr}\, F_g, \tag{15}$$

where we recall that we normalised $\text{Vol}_\Omega \equiv \int_X \text{vol}_\Omega = 1$.

Our connection network outputs $\log G^{-1}$, which in turn is used to compute F. Together with the data of an approximate Calabi–Yau metric g, this gives F_g as a function of the network weights $\boldsymbol{\nu}$. As we discuss in Appendix A, a suitable choice for the loss function of the connection network is

$$\text{Loss}[\boldsymbol{\nu}] \equiv \left\langle \text{tr}\, F_g^2(\boldsymbol{\nu}) \right\rangle - \frac{1}{d}\left\langle \text{tr}\, F_g(\boldsymbol{\nu}) \right\rangle^2. \tag{16}$$

Obviously, there are other loss functions that one could choose. For example, given that one can often compute the slope of $V(k)$ by algebraic means, one could instead minimise $|\langle \text{tr}\, F_g \rangle - \mu(V(k))|$, or any power of this.

4. Results

Having laid out our strategy, we now move to our results. The examples we consider are an elliptic curve, a K3 surface and a quintic threefold, all given as n-dimensional hypersurfaces in $\mathbb{P}^{n+1}$ defined by the zero locus of a degree-$(n+2)$ polynomial of the homogeneous coordinates $[z_0 : \ldots : z_{n+1}]$. In all examples, the Ricci-flat metric was first calculated using Douglas et al.'s metric network [31,44]. We then trained the connection network for a variety of network depths. In all examples, the loss function was taken to be Loss[$\boldsymbol{\nu}$] as in (16) and the networks were trained for 500 epochs using the Adam gradient-based optimisation algorithm [59]. The training sets and test sets each consisted of 10,000 random sample points on the relevant Calabi–Yau hypersurface. Note that with our choice of normalisations, the slope of the line bundle $V = \mathcal{O}_X(m)$ is given by $\mu(V) = m$, and so the HYM connection on V should satisfy $F_g = m$.

4.1. Elliptic curve

The defining equation of an elliptic curve can be expressed in the form[2]

$$f(z) = z_1^3 + a z_0^2 z_1 - z_0 z_2^2 + b z_0^3, \tag{17}$$

[2] This can be brought into the usual Weierstrass form $y^2 = x^3 + ax + b$ by defining $x = z_1/z_0$ and $y = z_2/z_0$.

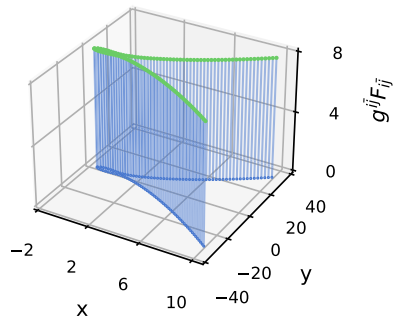

Fig. 3. Values of $g^{i\bar{j}}F_{i\bar{j}}$ for a line bundle connection calculated using a $D=4$ network on the elliptic curve trained using Loss[$\boldsymbol{\nu}$]. The plot shows the values of $g^{i\bar{j}}F_{i\bar{j}}$ on the z-axis sampled over points on the elliptic curve on the patch $z_0=1$ with $(x,y)=(z_1,z_2)$.

with the curve itself given by the zero locus of $f(z)$ in $\mathbb{P}^2$. The curve is non-singular if and only if the discriminant $\Delta = -16(4a^3+27b^2)$ is non-zero. The example we consider is the elliptic curve with $(a,b)=(-1,1)$. Since $\Delta\neq 0$, the curve is smooth and free from singularities. The approximate Calabi–Yau metric for this example was computed at $k=4$ using a network of depth $D=3$ with $W^{(i)}=(70,100,1)$. In the language of [25], the accuracy of this metric is $\sigma=0.001$.

To compute the connection, we considered neural networks of depth $D=2,3,4$ with intermediate layers of width $W^{(i)}=40$. Our results are shown in Fig. 2 with the training curves given in the appendices in Fig. B.9. In particular, we plot the histogram of $g^{i\bar{j}}F_{i\bar{j}}$ evaluated for points in both the training and test sets. One sees that the histogram of $g^{i\bar{j}}F_{i\bar{j}}$ is clustered around 2, 4 and 8 for the $D=2,3,4$ networks respectively, in agreement with $F_g=m$, with the distribution more peaked for $D=3,4$. In Fig. 3, we plot the values of $g^{i\bar{j}}F_{i\bar{j}}$ for the $D=4$ network over the elliptic curve on the patch $z_0=1$. As expected from the histograms, the values of $g^{i\bar{j}}F_{i\bar{j}}$ over the elliptic curve are very close to constant.

In order to compare the accuracy of the networks, we use the fact that the curvature of the HYM connection on $\mathcal{O}_X(m)$ and that on $\mathcal{O}_X(m+k)$ are related in a simple way since they are proportional. As an example, consider $V=\mathcal{O}_X(1)$ where we then twist by $\mathcal{O}_X(k)$ with $k=1,3,7$. Our neural network then computes the HYM connections on $\mathcal{O}_X(2)$, $\mathcal{O}_X(4)$ and $\mathcal{O}_X(8)$. We then untwist in order to obtain a connection on $V=\mathcal{O}_X(1)$ itself. We show the result of this in the left plot of Fig. 4. We see that all three networks are accurate, with the values of $g^{i\bar{j}}F_{i\bar{j}}$ within 1% of the expected result, i.e. one, though the $D=2$ network (corresponding to $\mathcal{O}_X(2)$) is the least accurate. The deeper $D=3$ and $D=4$ networks, however, have very similar accuracy to each other. This is not that surprising, since the Kähler metric on X was itself computed using a $D=3$ network, so the extra freedom allowed by the $D=4$ network is not necessary.

4.2. K3 surface

The K3 surface we consider is a smooth quartic hypersurface $f(z)=0$ in $\mathbb{P}^3$. The defining equation is

$$f(z)=z_0^4+z_1^4+z_2^4+z_3^4, \tag{18}$$

which gives the Fermat quartic. The approximate Calabi–Yau metric for this example was computed at $k=8$ using a network of depth $D=4$ with intermediate layers of width 100, i.e. $W^{(i)}=(100,100,100,1)$. In the language of [25], the resulting metric has a sigma measure of $\sigma=0.00035$.

We considered neural networks of depth $D=2,3,4$ with intermediate layers of width $W^{(i)}=100$. Our full results are given in Appendix B, with the histograms of $g^{i\bar{j}}F_{i\bar{j}}$ evaluated for both the training and test sets shown in Fig. A.7, and the training curves given in Fig. B.10. One sees that the histogram of $g^{i\bar{j}}F_{i\bar{j}}$ is tightly clustered around 2, 4 and 8 for the $D=2,3,4$ networks respectively, with the distribution more peaked for $D=3,4$.

In order to compare the accuracy of these three networks, we again treat the networks as computing connections on $\mathcal{O}_X(1+k)$, and then untwist in order to obtain a connection on $V=\mathcal{O}_X(1)$. We show the result of this in the middle plot of Fig. 4. We observe that all three networks are accurate, with the values of $g^{i\bar{j}}F_{i\bar{j}}$ within 2% of the expected result, i.e. one. Again, the $D=2$ network (corresponding to $\mathcal{O}_X(2)$) is the least accurate of the three, as it displays the largest spread in values, and the deepest, $D=4$ network gives the smallest spread of the three. The extra complexity allowed by the $D=4$ network leads to a small advantage over the $D=3$ network. Since the numerical Ricci-flat metric on X was itself computed using a $D=4$ network, it is somewhat surprising that the $D=4$ network does not show a larger increase in accuracy over $D=3$.

4.3. Quintic threefold

Finally, we consider a Calabi–Yau threefold given as smooth quintic hypersurface in $\mathbb{P}^4$. The defining equation is

$$f(z)=z_0^5+z_1^5+z_2^5+z_3^5+z_4^5+\tfrac{1}{2}z_0z_1z_2z_3z_4, \tag{19}$$

which gives a member of the Dwork family of quintics. The approximate Calabi–Yau metric for this example was computed at $k=8$ using a network of depth $D=4$ with $W^{(i)}=(100,100,100,1)$. The resulting metric has a sigma measure of $\sigma=0.001$.

We considered neural networks of depth $D=2,3,4$ with intermediate layers of width $W^{(i)}=100$. Our full results are given in Appendix B, with the histograms of $g^{i\bar{j}}F_{i\bar{j}}$ evaluated for points in both the training and test sets shown in Fig. B.8, and the training curves shown in Fig. B.11. In order to compare the accuracy of these three networks, we again untwist in order to obtain a connection on $V=\mathcal{O}_X(1)$. We show the result of this in the right plot of Fig. 4. We see that all three networks are accurate, with the values of $g^{i\bar{j}}F_{i\bar{j}}$ within 5% of the expected result, i.e. one. Note that we computed the approximate Ricci-flat metric using a $D=4$ network, and the curvature of the HYM connection should agree with the Kähler form of the Ricci-flat metric. Thanks to this, and the complexity of the Calabi–Yau metric on a threefold, it is not surprising that the $D=4$ network performs the best of the three.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

AA is supported by the EU's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 838776. YHH would like to thank STFC for grant ST/J00037X/1. BAO is supported in part by both the research grant DOE No. DESC0007901 and SAS Account 020-0188-2-010202-6603-0338.

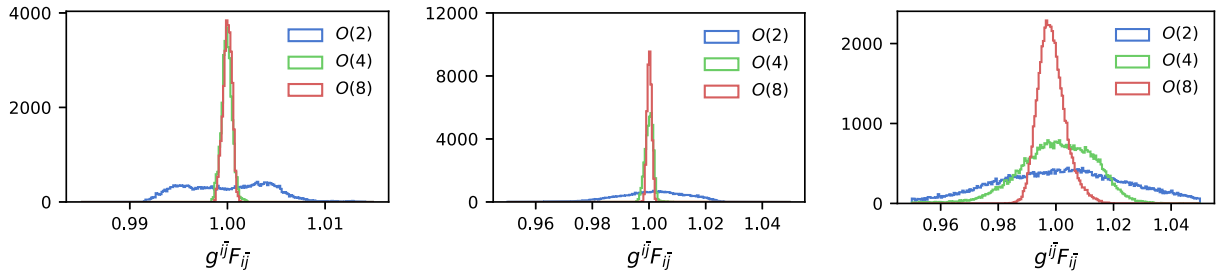

Fig. 4. Histograms of $g^{i\bar{j}}F_{i\bar{j}}$ on the line bundle $\mathcal{O}_X(1)$ for: (Left) an elliptic curve; (Middle) a K3 surface; (Right) a quintic threefold. In all cases, $g^{i\bar{j}}F_{i\bar{j}}$ is evaluated for points in the test set with the curvature given by untwisting the connection calculated by neural networks with $D = 2, 3, 4$, corresponding to $\mathcal{O}_X(2)$, $\mathcal{O}_X(4)$ and $\mathcal{O}_X(8)$.

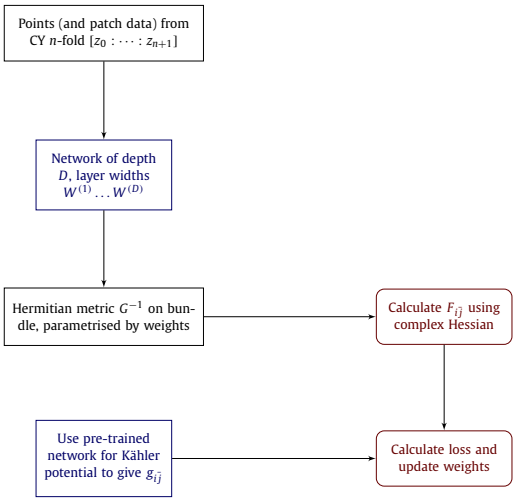

Fig. A.5. Network structure for determining the gauge connection of a line bundle on a CY n-fold.

Appendix A. Network structure and loss function

In this appendix, we further discuss the structure of the neural network for computing numerical connections on line bundles and also justify the choice of loss function in the main text.

A.1. The network structure

The network structure for determining a gauge connection that satisfies hermitian Yang–Mills is given in Fig. A.5. As we also describe in the main text, the inputs to the network are sets of points on the Calabi–Yau hypersurface, given as points on the ambient projective space. This information is fed into a linear, dense network of depth D with layer widths $W^{(i)}$, shown in Fig. 1. The first layer of the network is a bihomogeneous layer, followed by dense layers with quadratic activation functions and zero biases. The output of the network is the hermitian metric on the bundle (actually $\log G^{-1}$), parametrised by the weights $\boldsymbol{\nu}$ of the network. Next, we compute the curvature $F_{i\bar{j}}$ of the connection induced by this hermitian metric by taking the complex Hessian of the network. In practice, this computes the curvature as a tensor on the ambient space, so one must pull it back to the hypersurface using the Jacobian defined by the patches and the defining equation of the Calabi–Yau (see [24] for more details about this). The loss function of the network, whose discretised form is given in Equation (A.7), is defined by both the curvature computed by the network and a numerical Calabi–Yau metric. The latter comes from a pre-trained network whose output is a Kähler metric $g_{i\bar{j}}$ which is approximately Ricci flat. The loss function is then minimised by adjusting the weights $\boldsymbol{\nu}$ of the network using backpropagation. The result of this is a "trained" network whose output defines an approximate hermitian Yang–Mills connection.

A.2. The loss function

Consider the variance of the trace of F_g:

$$\mathrm{Var}[\mathrm{tr}\, F_g] = \langle (\mathrm{tr}\, F_g)^2 \rangle - \langle \mathrm{tr}\, F_g \rangle^2. \tag{A.1}$$

Clearly, $\mathrm{Var}[\mathrm{tr}\, F_g] \geq 0$ with equality only when $\mathrm{tr}\, F_g = \langle \mathrm{tr}\, F_g \rangle$. However, this is the case only if $\mathrm{tr}\, F_g$ is *constant* over X. Now consider the fact that for a $d \times d$ hermitian matrix M (such as $\langle \mathrm{tr}\, F_g \rangle$), one always has

$$d \cdot \mathrm{tr}\, M^2 \geq (\mathrm{tr}\, M)^2, \tag{A.2}$$

with equality if and only if M is proportional to the identity matrix, $M \propto \mathbf{1}$ (but with no constraint on the function relating the two). Putting together these two observations, we define the functional

$$E[F, g] = \langle \mathrm{tr}\, F_g^2 \rangle - \frac{1}{d} \langle \mathrm{tr}\, F_g \rangle^2, \tag{A.3}$$

which satisfies

$$0 \leq \mathrm{Var}[\mathrm{tr}\, F_g] \leq d \cdot E[F, g]. \tag{A.4}$$

Thus if one finds a connection such that $E[F, g] = 0$, it must be the case that (A.2) is saturated and $\mathrm{Var}[\mathrm{tr}\, F_g] = 0$, which imply $\mathrm{tr}\, F_g = \langle \mathrm{tr}\, F_g \rangle$ and $F_g \propto \mathbf{1}$ respectively. Taken together, these two conditions are equivalent to $F_g = c\,\mathbf{1}$ with c constant, and so one has found a HYM connection. Conversely, it is clear that a HYM connection satisfies both $\mathrm{Var}[\mathrm{tr}\, F_g] = 0$ and saturates the inequality (A.2). Hence, we have

$$F \text{ solves HYM} \quad \Leftrightarrow \quad E[F, g] = 0. \tag{A.5}$$

This motivates our choice of loss function, $\mathrm{Loss}[\boldsymbol{\nu}]$, in Equation (16). For the special case of a rank-one bundle, i.e. a line bundle with $d = 1$, one has

$$E[F, g] = \langle F_g^2 \rangle - \langle F_g \rangle^2, \tag{A.6}$$

which is simply the variance of F_g.

Written as a discrete sum over points of X, the loss function is

$$\mathrm{Loss}[\boldsymbol{\nu}] = \frac{\sum_p \mathrm{tr}(g^{i\bar{j}}(p) F_{i\bar{j}}(\boldsymbol{\nu}, p))^2 w_p}{\sum_p w_p} - \frac{1}{d} \frac{(\sum_p g^{i\bar{j}}(p)\, \mathrm{tr}\, F_{i\bar{j}}(\boldsymbol{\nu}, p) w_p)^2}{(\sum_p w_p)^2}, \tag{A.7}$$

where p denotes a point in the training set, and w_p is a mass which weights the sum over points to reproduce the integration measure defined by vol_Ω [24].

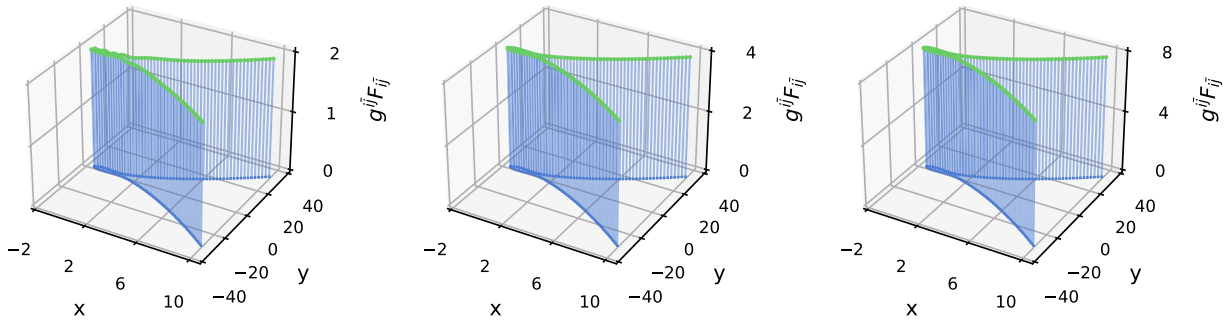

Fig. A.6. Results for line bundle connections on an elliptic curve trained using Loss[$\boldsymbol{\nu}$]. The plots show the values of $g^{i\bar{j}}F_{i\bar{j}}$ on the z-axis over the elliptic curve on the patch $z_0 = 1$ with $(x, y) = (z_1, z_2)$. The left, middle and right plots are for networks of depth $D = 2, 3, 4$ respectively, which correspond to connections on the line bundles $\mathcal{O}_X(2)$, $\mathcal{O}_X(4)$ and $\mathcal{O}_X(8)$.

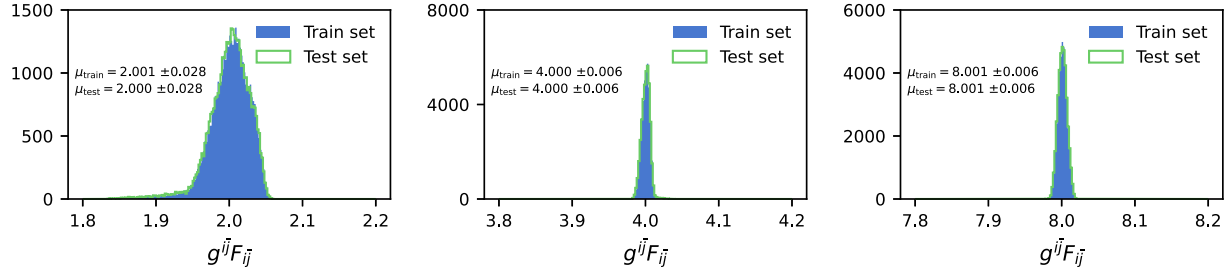

Fig. A.7. Results for line bundle connections on a K3 surface trained using Loss[$\boldsymbol{\nu}$]. The plots show the histogram of $g^{i\bar{j}}F_{i\bar{j}}$ evaluated for points in both the training and test sets, together with the mean and standard deviation. The plots are for networks of depth $D = 2, 3, 4$ which correspond to connections on the line bundles $\mathcal{O}_X(2)$, $\mathcal{O}_X(4)$ and $\mathcal{O}_X(8)$.

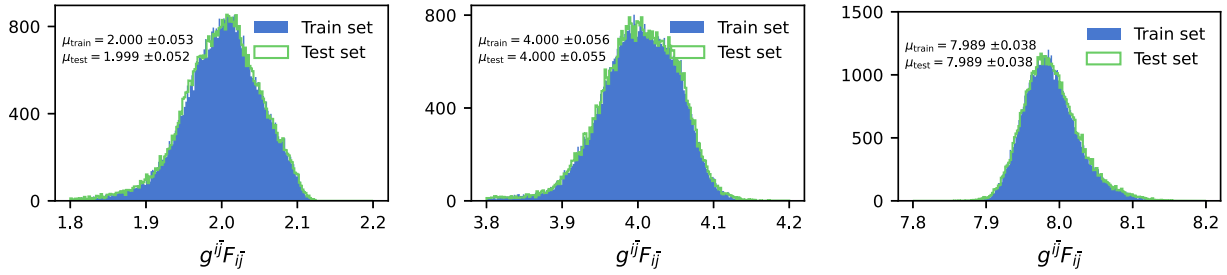

Fig. B.8. Results for line bundle connections on a quintic threefold trained using Loss[$\boldsymbol{\nu}$]. The plots show the histogram of $g^{i\bar{j}}F_{i\bar{j}}$ evaluated for points in both the training and test sets, together with the mean and standard deviation. The plots are for networks of depth $D = 2, 3, 4$ which correspond to connections on the line bundles $\mathcal{O}_X(2)$, $\mathcal{O}_X(4)$ and $\mathcal{O}_X(8)$.

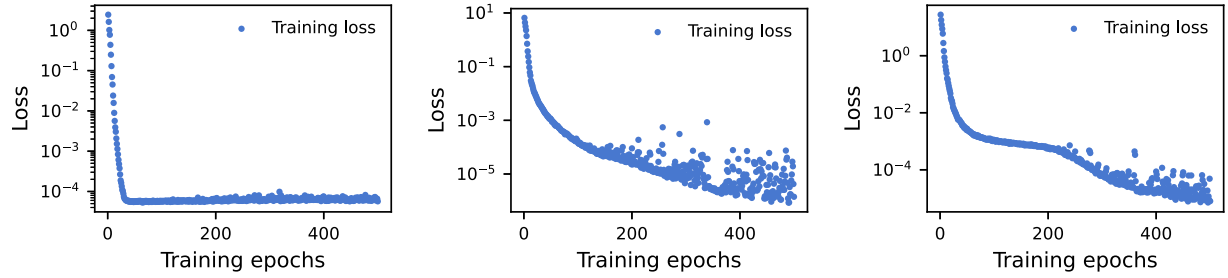

Fig. B.9. Training curves (the value of the loss function evaluated on the training set) as a function of training epoch for line bundle connections on an elliptic curve trained using Loss[$\boldsymbol{\nu}$]. The plots are for networks of depth $D = 2, 3, 4$ which correspond to connections on the line bundles $\mathcal{O}_X(2)$, $\mathcal{O}_X(4)$ and $\mathcal{O}_X(8)$.

Appendix B. Histograms and training curves

In this appendix, we display the histograms for $g^{i\bar{j}}F_{i\bar{j}}$ and the training curves for $D = 2, 3, 4$ networks computing line bundle connections on an elliptic curve, K3 surface, and quintic threefold, as in the main text.

Fig. A.6 shows the values of $g^{i\bar{j}}F_{i\bar{j}}$ for points on the elliptic curve (17) on the patch $z_0 = 1$ for $D = 2, 3, 4$ networks, completing Fig. 3 given in the main text. Figs. A.7 and B.8 show the histogram of $g^{i\bar{j}}F_{i\bar{j}}$ for line bundle connections for both the K3 and quintic threefold examined in the main text. Each figure displays the histograms for networks of depth $D = 2, 3, 4$, corresponding to

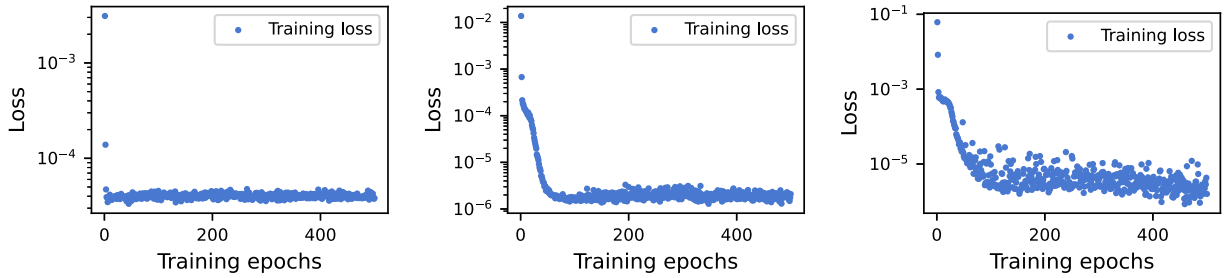

Fig. B.10. Training curves (the value of the loss function evaluated on the training set) as a function of training epoch for line bundle connections on a K3 surface trained using Loss[$\boldsymbol{v}$]. The plots are for networks of depth $D = 2, 3, 4$ which correspond to connections on the line bundles $\mathcal{O}_X(2)$, $\mathcal{O}_X(4)$ and $\mathcal{O}_X(8)$.

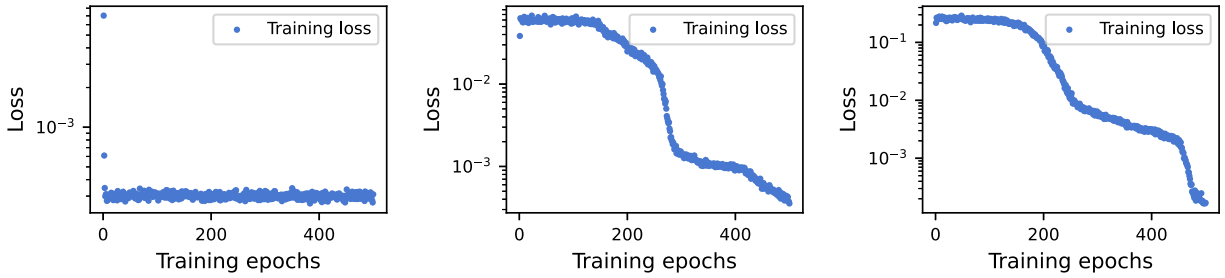

Fig. B.11. Training curves (the value of the loss function evaluated on the training set) as a function of training epoch for line bundle connections on a quintic threefold trained using Loss[$\boldsymbol{v}$]. The plots are for networks of depth $D = 2, 3, 4$ which correspond to connections on the line bundles $\mathcal{O}_X(2)$, $\mathcal{O}_X(4)$ and $\mathcal{O}_X(8)$.

computing the connection on $\mathcal{O}_X(2)$, $\mathcal{O}_X(4)$ and $\mathcal{O}_X(8)$. Both the training set and test set are included; there are no signs of overtraining.

Finally, Figs. B.9, B.10 and B.11 show the training curves (the value of the loss function evaluated on the training set) as a function of training epoch for line bundle connections on the elliptic curve (17), the K3 surface (18), and the quintic threefold (19), respectively. Each figure displays the training curves for networks of depth $D = 2, 3, 4$, corresponding to computing the connection on $\mathcal{O}_X(2)$, $\mathcal{O}_X(4)$ and $\mathcal{O}_X(8)$. The general pattern that one observes is that the shallow $D = 2$ networks reach a minimum very quickly, while the deepest $D = 4$ networks show decreasing losses all the way to the 500th epoch.

References

[1] V. Braun, Y.-H. He, B.A. Ovrut, T. Pantev, A Heterotic standard model, Phys. Lett. B 618 (2005) 252–258, https://doi.org/10.1016/j.physletb.2005.05.007, arXiv:hep-th/0501070.

[2] A. Lukas, B.A. Ovrut, K.S. Stelle, D. Waldram, The Universe as a domain wall, Phys. Rev. D 59 (1999) 086001, https://doi.org/10.1103/PhysRevD.59.086001, arXiv:hep-th/9803235.

[3] R. Donagi, B.A. Ovrut, T. Pantev, D. Waldram, Standard models from heterotic M theory, Adv. Theor. Math. Phys. 5 (2002) 93–137, https://doi.org/10.4310/ATMP.2001.v5.n1.a4, arXiv:hep-th/9912208.

[4] V. Bouchard, R. Donagi, An SU(5) heterotic standard model, Phys. Lett. B 633 (2006) 783–791, https://doi.org/10.1016/j.physletb.2005.12.042, arXiv:hep-th/0512149.

[5] R. Blumenhagen, S. Moster, T. Weigand, Heterotic GUT and standard model vacua from simply connected Calabi-Yau manifolds, Nucl. Phys. B 751 (2006) 186–221, https://doi.org/10.1016/j.nuclphysb.2006.06.005, arXiv:hep-th/0603015.

[6] O. Lebedev, H.P. Nilles, S. Raby, S. Ramos-Sanchez, M. Ratz, P.K.S. Vaudrevange, A. Wingerter, A mini-landscape of exact MSSM spectra in heterotic orbifolds, Phys. Lett. B 645 (2007) 88–94, https://doi.org/10.1016/j.physletb.2006.12.012, arXiv:hep-th/0611095.

[7] P. Candelas, X. de la Ossa, Y.-H. He, B. Szendroi, Triadophilia: a special corner in the landscape, Adv. Theor. Math. Phys. 12 (2) (2008) 429–473, https://doi.org/10.4310/ATMP.2008.v12.n2.a6, arXiv:0706.3134.

[8] O. Lebedev, H.P. Nilles, S. Ramos-Sanchez, M. Ratz, P.K.S. Vaudrevange, Heterotic mini-landscape. (II). Completing the search for MSSM vacua in a Z(6) orbifold, Phys. Lett. B 668 (2008) 331–335, https://doi.org/10.1016/j.physletb.2008.08.054, arXiv:0807.4384.

[9] D.K. Mayorga Pena, H.P. Nilles, P.-K. Oehlmann, A zip-code for quarks, leptons and Higgs bosons, J. High Energy Phys. 12 (2012) 024, https://doi.org/10.1007/JHEP12(2012)024, arXiv:1209.6041.

[10] L.B. Anderson, J. Gray, Y.-H. He, A. Lukas, Exploring positive monad bundles and a new heterotic standard model, J. High Energy Phys. 02 (2010) 054, https://doi.org/10.1007/JHEP02(2010)054, arXiv:0911.1569.

[11] L.B. Anderson, J. Gray, A. Lukas, E. Palti, Two hundred heterotic standard models on smooth Calabi-Yau threefolds, Phys. Rev. D 84 (2011) 106005, https://doi.org/10.1103/PhysRevD.84.106005, arXiv:1106.4804.

[12] L.B. Anderson, A. Constantin, J. Gray, A. Lukas, E. Palti, A comprehensive scan for heterotic SU(5) GUT models, J. High Energy Phys. 01 (2014) 047, https://doi.org/10.1007/JHEP01(2014)047, arXiv:1307.4787.

[13] S.K. Donaldson, Anti self-dual Yang-Mills connections over complex algebraic surfaces and stable vector bundles, Proc. Lond. Math. Soc. s3–50 (1) (1985) 1–26, https://doi.org/10.1112/plms/s3-50.1.1.

[14] K. Uhlenbeck, S.T. Yau, On the existence of Hermitian-Yang-Mills connections in stable vector bundles, Commun. Pure Appl. Math. 39 (S1) (1986) S257–S293, https://doi.org/10.1002/cpa.3160390714.

[15] B.R. Greene, K.H. Kirklin, P.J. Miron, G.G. Ross, A superstring inspired standard model, Phys. Lett. B 180 (1986) 69, https://doi.org/10.1016/0370-2693(86)90137-1.

[16] B.R. Greene, K.H. Kirklin, P.J. Miron, G.G. Ross, A three generation superstring model. 1. Compactification and discrete symmetries, Nucl. Phys. B 278 (1986) 667–693, https://doi.org/10.1016/0550-3213(86)90057-X.

[17] B.R. Greene, K.H. Kirklin, P.J. Miron, G.G. Ross, A three generation superstring model. 2. Symmetry breaking and the low-energy theory, Nucl. Phys. B 292 (1987) 606–652, https://doi.org/10.1016/0550-3213(87)90662-6.

[18] T. Matsuoka, D. Suematsu, Realistic models from the $E(8) \times E(8)$-prime superstring theory, Prog. Theor. Phys. 76 (1986) 886, https://doi.org/10.1143/PTP.76.886.

[19] B.R. Greene, K.H. Kirklin, P.J. Miron, G.G. Ross, 27**3 Yukawa couplings for a three generation superstring model, Phys. Lett. B 192 (1987) 111–118, https://doi.org/10.1016/0370-2693(87)91151-8.

[20] R. Donagi, B.A. Ovrut, T. Pantev, D. Waldram, Standard model bundles, Adv. Theor. Math. Phys. 5 (2002) 563–615, https://doi.org/10.4310/ATMP.2001.v5.n3.a5, arXiv:math/0008010.

[21] V. Braun, Y.-H. He, B.A. Ovrut, Yukawa couplings in heterotic standard models, J. High Energy Phys. 04 (2006) 019, https://doi.org/10.1088/1126-6708/2006/04/019, arXiv:hep-th/0601204.

[22] L.B. Anderson, J. Gray, B. Ovrut, Yukawa textures from heterotic stability walls, J. High Energy Phys. 05 (2010) 086, https://doi.org/10.1007/JHEP05(2010)086, arXiv:1001.2317.

[23] M. Headrick, T. Wiseman, Numerical Ricci-flat metrics on K3, Class. Quantum Gravity 22 (2005) 4931–4960, https://doi.org/10.1088/0264-9381/22/23/002, arXiv:hep-th/0506129.

[24] M.R. Douglas, R.L. Karp, S. Lukic, R. Reinbacher, Numerical Calabi-Yau metrics, J. Math. Phys. 49 (2008) 032302, https://doi.org/10.1063/1.2888403, arXiv:hep-th/0612075.

[25] V. Braun, T. Brelidze, M.R. Douglas, B.A. Ovrut, Calabi-Yau metrics for quotients and complete intersections, J. High Energy Phys. 05 (2008) 080, https://doi.org/10.1088/1126-6708/2008/05/080, arXiv:0712.3563.
[26] M. Headrick, A. Nassar, Energy functionals for Calabi-Yau metrics, Adv. Theor. Math. Phys. 17 (5) (2013) 867–902, https://doi.org/10.4310/ATMP.2013.v17.n5.a1, arXiv:0908.2635.
[27] G. Tian, On a set of polarized Kähler metrics on algebraic manifolds, J. Differ. Geom. 32 (1) (1990) 99–130, https://doi.org/10.4310/jdg/1214445039.
[28] S.K. Donaldson, Some numerical results in complex differential geometry, arXiv:math/0512625.
[29] A. Ashmore, Y.-H. He, B.A. Ovrut, Machine learning Calabi–Yau metrics, Fortschr. Phys. 68 (9) (2020) 2000068, https://doi.org/10.1002/prop.202000068, arXiv:1910.08605.
[30] 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, J. High Energy Phys. 05 (2021) 013, https://doi.org/10.1007/JHEP05(2021)013, arXiv:2012.04656.
[31] M.R. Douglas, S. Lakshminarasimhan, Y. Qi, Numerical Calabi-Yau metrics from holomorphic networks, arXiv:2012.04797.
[32] V. Jejjala, D.K. Mayorga Pena, C. Mishra, Neural Network Approximations for Calabi-Yau Metrics, arXiv:2012.15821.
[33] M.R. Douglas, Holomorphic feedforward networks, arXiv:2105.03991.
[34] Y.-H. He, The Calabi–Yau Landscape: From Geometry, to Physics, to Machine Learning, Lecture Notes in Mathematics, 2021, arXiv:1812.02893.
[35] M.R. Douglas, R.L. Karp, S. Lukic, R. Reinbacher, Numerical solution to the Hermitian Yang-Mills equation on the Fermat quintic, J. High Energy Phys. 12 (2007) 083, https://doi.org/10.1088/1126-6708/2007/12/083, arXiv:hep-th/0606261.
[36] L.B. Anderson, V. Braun, R.L. Karp, B.A. Ovrut, Numerical Hermitian Yang-Mills connections and vector bundle stability in heterotic theories, J. High Energy Phys. 06 (2010) 107, https://doi.org/10.1007/JHEP06(2010)107, arXiv:1004.4399.
[37] L.B. Anderson, V. Braun, B.A. Ovrut, Numerical Hermitian Yang-Mills connections and Kahler cone substructure, J. High Energy Phys. 01 (2012) 014, https://doi.org/10.1007/JHEP01(2012)014, arXiv:1103.3041.
[38] X. Wang, Canonical metrics on stable vector bundles, Commun. Anal. Geom. 13 (2) (2005) 253–285, https://doi.org/10.4310/CAG.2005.v13.n2.a1.
[39] L.B. Anderson, J. Gray, A. Lukas, E. Palti, Heterotic line bundle standard models, J. High Energy Phys. 06 (2012) 113, https://doi.org/10.1007/JHEP06(2012)113, arXiv:1202.1757.
[40] S. Groot Nibbelink, O. Loukas, F. Ruehle, (MS)SM-like models on smooth Calabi-Yau manifolds from all three heterotic string theories, Fortschr. Phys. 63 (2015) 609–632, https://doi.org/10.1002/prop.201500041, arXiv:1507.07559.
[41] S. Groot Nibbelink, O. Loukas, F. Ruehle, P.K.S. Vaudrevange, Infinite number of MSSMs from heterotic line bundles?, Phys. Rev. D 92 (4) (2015) 046002, https://doi.org/10.1103/PhysRevD.92.046002, arXiv:1506.00879.
[42] S. Groot Nibbelin, F. Ruehle, Line bundle embeddings for heterotic theories, J. High Energy Phys. 04 (2016) 186, https://doi.org/10.1007/JHEP04(2016)186, arXiv:1601.00676.
[43] A.P. Braun, C.R. Brodie, A. Lukas, Heterotic line bundle models on elliptically fibered Calabi-Yau three-folds, J. High Energy Phys. 04 (2018) 087, https://doi.org/10.1007/JHEP04(2018)087, arXiv:1706.07688.
[44] M.R. Douglas, S. Lakshminarasimhan, Y. Qi, MLGeometry, https://github.com/yidiq7/MLGeometry, 2021.
[45] V. Braun, Y.-H. He, B.A. Ovrut, T. Pantev, The exact MSSM spectrum from string theory, J. High Energy Phys. 05 (2006) 043, https://doi.org/10.1088/1126-6708/2006/05/043, arXiv:hep-th/0512177.
[46] V. Braun, Y.-H. He, B.A. Ovrut, T. Pantev, A Standard model from the $E(8) \times E(8)$ heterotic superstring, J. High Energy Phys. 06 (2005) 039, https://doi.org/10.1088/1126-6708/2005/06/039, arXiv:hep-th/0502155.
[47] V. Braun, Y.-H. He, B.A. Ovrut, T. Pantev, A heterotic standard model, Phys. Lett. B 618 (2005) 252–258, https://doi.org/10.1016/j.physletb.2005.05.007, arXiv:hep-th/0501070.
[48] V. Braun, P. Candelas, R. Davies, R. Donagi, The MSSM spectrum from (0, 2)-deformations of the heterotic standard embedding, J. High Energy Phys. 05 (2012) 127, https://doi.org/10.1007/JHEP05(2012)127, arXiv:1112.1097.
[49] S. Groot Nibbelink, O. Loukas, F. Ruehle, P.K.S. Vaudrevange, Infinite number of MSSMs from heterotic line bundles?, Phys. Rev. D 92 (4) (2015) 046002, https://doi.org/10.1103/PhysRevD.92.046002, arXiv:1506.00879.
[50] M. Blaszczyk, S. Groot Nibbelink, F. Ruehle, M. Trapletti, P.K.S. Vaudrevange, Heterotic MSSM on a resolved orbifold, J. High Energy Phys. 09 (2010) 065, https://doi.org/10.1007/JHEP09(2010)065, arXiv:1007.0203.
[51] B. Andreas, G. Curio, A. Klemm, Towards the Standard Model spectrum from elliptic Calabi-Yau, Int. J. Mod. Phys. A 19 (2004) 1987, https://doi.org/10.1142/S0217751X04018087, arXiv:hep-th/9903052.
[52] V. Braun, T. Brelidze, M.R. Douglas, B.A. Ovrut, Eigenvalues and eigenfunctions of the scalar Laplace operator on Calabi-Yau manifolds, J. High Energy Phys. 07 (2008) 120, https://doi.org/10.1088/1126-6708/2008/07/120, arXiv:0805.3689.
[53] A. Ashmore, Eigenvalues and eigenforms on Calabi-Yau threefolds, arXiv:2011.13929.
[54] N. Afkhami-Jeddi, A. Ashmore, C. Cordova, Calabi-Yau CFTs and Random Matrices, arXiv:2107.11461.
[55] A. Lukas, B.A. Ovrut, D. Waldram, Boundary inflation, Phys. Rev. D 61 (2000) 023506, https://doi.org/10.1103/PhysRevD.61.023506, arXiv:hep-th/9902071.
[56] J. McOrist, On the effective field theory of heterotic vacua, Lett. Math. Phys. 108 (4) (2018) 1031–1081, https://doi.org/10.1007/s11005-017-1025-0, arXiv:1606.05221.
[57] c. Blesneag, E.I. Buchbinder, A. Constantin, A. Lukas, E. Palti, Matter field Kähler metric in heterotic string theory from localisation, J. High Energy Phys. 04 (2018) 139, https://doi.org/10.1007/JHEP04(2018)139, arXiv:1801.09645.
[58] K. Ishiguro, T. Kobayashi, H. Otsuka, Hierarchical structure of physical Yukawa couplings from matter field Kähler metric, J. High Energy Phys. 07 (2021) 064, https://doi.org/10.1007/JHEP07(2021)064, arXiv:2103.10240.
[59] D.P. Kingma, J. Ba Adam, A Method for Stochastic Optimization, arXiv:1412.6980.