Machine Learning for Calabi–Yau Compactifications

Anthony Ashmore

University of Chicago & Sorbonne Université

2110.12483 AA, R. Deen, Y-H. He, B. Ovrut

Overview

Calabi–Yau metrics and hermitian Yang–Mills connections are crucial for string phenomenology

Numerical methods can give us access to this data

Machine learning and neural networks provides a powerful set of tools to tackle geometric problems, such as finding line bundle connections

Motivation from physics

Does string theory describe our universe?

- Heterotic string on Calabi–Yau comes closest to realistic MSSM models

Usually focus on getting correct gauge group, matter spectrum, etc.

- Do not need details of metric or connection for these

How many of these string vacua are physically reasonable?

- 4d physics depends on metric and connection
- No explicitly known non-trivial Calabi–Yau metrics or hermitian Yang–Mills connections!

Calabi–Yau compactifications

Minimal supersymmetry on $\mathbb{R}^{1,3} \times X$ with $E_8 \times E_8$ bundle V [Candelas et al. '85]

- No H flux $\Rightarrow X$ is Calabi–Yau
- V admits hermitian Yang–Mills connection

Particle content comes from choice of X and V

- $SU(3)$ bundle gives E_6 GUT gauge group in 4d, matter fields in $H^1(X, V_R)$

Physics in 4d determined by geometry of X and V – zero modes determine low-energy physics

The problem

How do we calculate Calabi–Yau metrics or hermitian Yang–Mills connections?

Outline

Calabi–Yau metrics

Hermitian Yang–Mills connections

Machine learning and neural networks

Example: HYM connections on line bundles

Calabi–Yau metrics

Calabi–Yau basics

Calabi–Yau manifolds are Kähler and admit Ricci-flat metrics

- Existence but no explicit constructions
- Kähler + $c_1(X) = 0 \Rightarrow$ there exists a Ricci-flat metric [Yau '77]

Kähler $\Rightarrow$ Kähler potential K gives (real) closed two-form $J = \partial\bar{\partial}K$

$c_1(X) = 0 \Rightarrow$ (complex) nowhere-vanishing (3,0)-form Ω

$$\mathrm{vol}_J \equiv J \wedge J \wedge J, \qquad \mathrm{vol}_\Omega \equiv \mathrm{i}\,\Omega \wedge \bar{\Omega}.$$

Example: Fermat quintic

Quintic hypersurface Q in $\mathbb{P}^4$

$$Q(z) \equiv z_0^5 + z_1^5 + z_2^5 + z_3^5 + z_4^5 = 0$$

(3,0)-form Ω determined by Q, e.g. in $z_0 = 1$ patch

$$\Omega = \frac{\mathrm{d}z_2 \wedge \mathrm{d}z_3 \wedge \mathrm{d}z_4}{\partial Q / \partial z_1}$$

Metric g and Kähler form J determined by Kähler potential

$$g_{i\bar{j}}(z,\bar{z}) = \partial_i \bar{\partial}_{\bar{j}} K(z,\bar{z}), \qquad \mathrm{vol}_J \sim \det g_{i\bar{j}} \mathrm{d}^6 z$$

How do we measure accuracy?

The Ricci-flat metric is given by a K that satisfies (c.f. Monge–Ampère)

$$\left.\frac{\mathrm{vol}_J}{\mathrm{vol}_\Omega}\right|_p = \text{constant} \quad \Rightarrow \quad R_{i\bar{j}} = 0$$

Define a functional of K [Douglas et al. '06]

$$\sigma(K) = \int_X \left| 1 - \frac{\mathrm{vol}_J}{\mathrm{vol}_\Omega} \right| \mathrm{vol}_\Omega$$

Finding the Ricci-flat metric reduces to finding $K(z, \bar{z})$ that minimises σ

- The exact CY metric has $\sigma = 0$

Algebraic metrics [Tian '90; Donaldson '05]

Natural Kähler metric on $\mathbb{P}^4$ given by

$$K_{\mathrm{FS}} = \ln \sum_{i=0}^{4} z_i \bar{z}_{\bar{i}}$$

Can generalise this with a hermitian matrix $h^{i\bar{j}}$

$$K(h) = \ln \sum_{i,\bar{j}=0}^{4} z_i h^{i\bar{j}} \bar{z}_{\bar{j}}$$

Restricting to $Q \subset \mathbb{P}^4$ gives Kähler metric but not Ricci-flat

- 25 real parameters in $h^{i\bar{j}}$ that we can vary

Algebraic metrics [Tian '90; Donaldson '05]

Replace coordinates z_i with homogeneous polynomials s_α of degree k

$$\text{e.g.} \quad k = 2 : \qquad s_\alpha = (z_0^2, z_0 z_1, z_0 z_2, \ldots)$$

Kähler potential is then

$$K(h) = \ln \sum_{\alpha, \bar\beta = 0}^{14} s_\alpha h^{\alpha\bar\beta} \bar{s}_{\bar\beta}, \qquad h^{\alpha\bar\beta} \sim 225 \text{ parameters}$$

At degree k have $\mathcal{O}(k^8)$ parameters, so can approximate the Ricci-flat metric to arbitrary precision

- Spectral method as $\{s_\alpha \bar{s}_{\bar\beta}\}$ gives basis for eigenspaces on $\mathbb{P}^4$

How to fix K?

Finding the "best" approximation to the Ricci-flat metric amounts to finding K so that σ is minimised

Three approaches:

- Iterative procedure for $h^{\alpha\bar{\beta}}$ [Donaldson '05; Douglas '06; Braun '07]
- Minimise σ directly as function of $h^{\alpha\bar{\beta}}$ [Headrick, Nassar '09; Anderson et al. '20]
- Find K or $g_{i\bar{j}}$ directly by treating σ as a loss function for a neural network [Headrick, Wiseman '05; Douglas et al. 20; Anderson et al. '20; Jejjala et al. '20; Larfors et al. '21]

In all cases, numerical integrals carried out by Monte Carlo

Hermitian Yang–Mills connections

Hermitian Yang–Mills

Given Kähler manifold (X, g) , a connection A with curvature $F = \mathrm{d}A + A \wedge A$ on a holomorphic vector bundle V is hermitian Yang–Mills if

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

HYM implies Yang–Mills: $\mathrm{d} \star F = 0$

- Equations of motion in 10d require Yang–Mills
- Supersymmetry in 10d requires HYM with $\mu(V) = 0$

Hermitian Yang–Mills

Connection A defined by hermitian structure on V, i.e. hermitian inner product G on sections of V

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

Explicitly

$$A_i = G^{-1}\partial_i G, \quad A_{\bar{i}} = 0 \quad \Rightarrow \quad F_{i\bar{j}} = \partial_{\bar{j}}(G^{-1}\partial_i G).$$

If F solves HYM, G is known as a Hermite–Einstein metric on V

Hermitian Yang–Mills

Existence of HYM solutions [Donaldson ‘85; Uhlenbeck, Yau ‘86]

A holomorphic vector bundle V over a compact Kähler manifold (X, J) admits a Hermite–Einstein metric iff V is slope polystable

Slope of V

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

V is stable if $\mu(\mathcal{F}) < \mu(V)$ for all $\mathcal{F} \subset V$ (or polystable if sum of stable bundles with same slope)

- Algebraic condition (like $c_1(X) = 0$), but not constructive!

Numerical connections

Ansatz for hermitian structure [Wang '05; Douglas et al. '06; Anderson et al. '10]

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

where $S^a_\alpha \in H^0(X, V \otimes L^k)$ and $H^{\alpha\bar{\beta}}$ is a hermitian matrix of parameters

Varying $H^{\alpha\bar{\beta}}$, one can find an approximate HYM connection for $V \otimes L^k$ and then V itself

Increasing k increases the number of sections $\{S^a_\alpha\}$ and hence the number of parameters in $H^{\alpha\bar{\beta}}$

How do we measure accuracy?

Defining $F_g \equiv g^{i\bar{j}} F_{i\bar{j}}$, the HYM equation is $F_g = \mu(V)\,\mathbf{1}$

The average over the the Calabi–Yau is defined using the exact CY measure vol_Ω, e.g.

$$\langle \operatorname{tr} F_g \rangle \equiv \int_X \mathrm{vol}_\Omega \operatorname{tr} F_g$$

Suitable choice of accuracy measure is

$$E[F, g] = \langle \operatorname{tr} F_g^2 \rangle - \frac{1}{d} \langle \operatorname{tr} F_g \rangle^2$$

$E[F, g]$ is positive semi-definite and vanishes on HYM solutions

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

Line bundles on CY manifolds

Line bundles feature in many string models [Anderson, Gray, Lukas, Palti '11;...]

Holomorphic line bundle L determined by $c_1(L)$. Given a basis of divisors $\mathcal{D}_I$ on X, denote by $\mathcal{O}_X(m^I)$ the line bundle with $c_1(L) = m^I \mathcal{D}_I$

Line bundles are automatically stable (no rank > 0 subsheaves)

- L always admits a solution to

$$g^{i\bar{j}} F_{i\bar{j}} = \mu(L)$$

- How do we find the explicit form of A?

The goal

Train a neural network to find solutions to hermitian Yang–Mills equation on line bundles

Machine learning and neural networks

Overview

New era of big data in string theory

- Vacuum selection problem, huge number of CYs, larger number of flux vacua (at least $10^{272,000}$? [Denef, Douglas '04; Taylor, Wang '15])

Different types of machine learning

- Supervised learning – known inputs and outputs, e.g. recognise images, predict $h^{1,1}$ [He '17; Bull, He Jejala, Mishra '18; Erbin, Finotello '20]
- Unsupervised – known inputs, e.g. looking for patterns or generate images
- Self-generative – known inputs, output minimises a loss function, e.g. ground states, Ricci-flat metrics, HYM connections

Neural networks

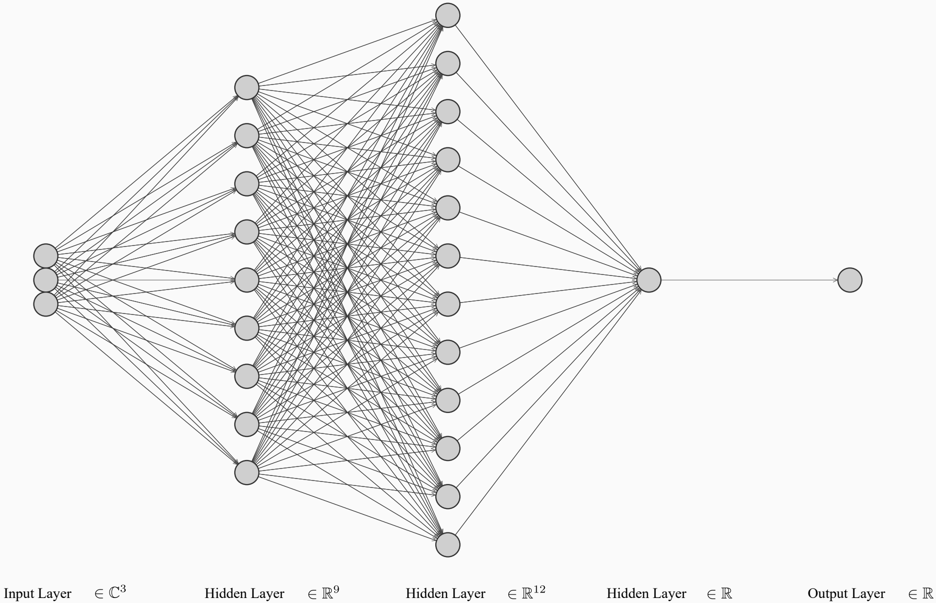
Neural networks (NN) convert inputs to outputs: $\vec{x} \mapsto f(\vec{x}, \vec{v})$

- Network built from connected nodes called neurons
- Weights $\vec{v}$ are parameters in network (strength of connections)
- Non-linear activation functions
- Training attempts to minimise a loss function computed from NN

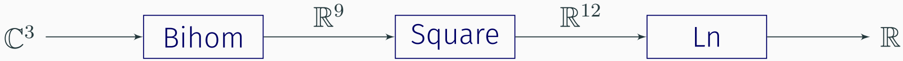
More interested in the network itself than the actual values!

- Universal approximation theorem for NNs
- NN gives a variational ansatz for some function you want to find, e.g. Hermite–Einstein metric G that solves HYM equation

Example: $D = 2$, $W^{(i)} = (12, 1)$ network



Bihomogenous networks [Douglas et al. '20]



$$\mathbb{C}^3 \to \mathbb{R}^9 \qquad \mathbb{R}^9 \to \mathbb{R}^{12} \qquad \mathbb{R}^{12} \to \mathbb{R}$$

$$z_i \mapsto (\mathrm{re}\, z_j \bar{z}_k, \mathrm{im}\, z_j \bar{z}_k) \qquad \vec{x} \mapsto (W_1 \vec{x})^2 \qquad \vec{y} \mapsto \ln(W_2 \vec{y})$$

Parameters in W_1 and W_2 are weights, collectively denoted by $\vec{v}$

First implemented for CY metrics in TensorFlow by [Douglas et al. '20] at

`https://github.com/yidiq7/MLGeometry`

A loss function

Network output is treated as $\ln G^{-1}$, which defines F

- Together with approximate CY metric g, this gives $F_g[\vec{v}]$ as a function of the network weights $\vec{v}$
- Mimics previous ansatz for hermitian structure

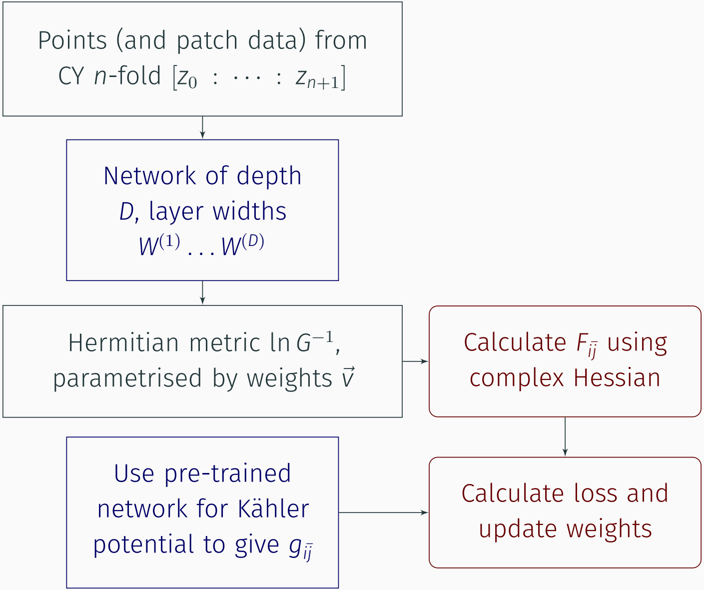
Loss function is

$$\text{Loss}[F, g] = E[F, g] \equiv \langle \text{tr}\, F_g^2 \rangle - \frac{1}{d} \langle \text{tr}\, F_g \rangle^2$$

After training, the network gives a NN-based representation of the HYM connection

- Effectively the functional form of G or A or F (can take derivatives, etc.)

General strategy



Example: HYM connections on line bundles

Examples

Consider Calabi–Yau n-fold X given as hypersurface in $\mathbb{P}^{n+1}$ defined by zero locus of single degree-$(n+2)$ polynomial

- e.g. elliptic curve in $\mathbb{P}^2$, K3 surface in $\mathbb{P}^3$, quintic threefold in $\mathbb{P}^4$

One Kähler modulus, so line bundles labelled by degree $\mathcal{O}_X(m)$

- Normalise so that $V = \mathcal{O}_X(m)$ has slope $\mu(V) = m$, so honest HYM connection should be constant over X

$$F_g \equiv g^{i\bar{j}} F_{i\bar{j}} = m$$

Outline

Train networks with loss function $\text{Loss}[F, g]$ – minimised on HYM solutions

- Networks of depth D with intermediate layers of width $W^{(i)} = (W^{(1)}, \ldots)$
- Points $[z_0 : \ldots]$ on X are inputs
- Approximate CY metric g needed for loss (assume already calculated)
- Training and test sets of 10,000 points each

Depth D network gives connection on $\mathcal{O}_X(2^{D-1})$

- Wider and deeper network has more parameters

$\mathcal{O}_X(m)$ on elliptic curve

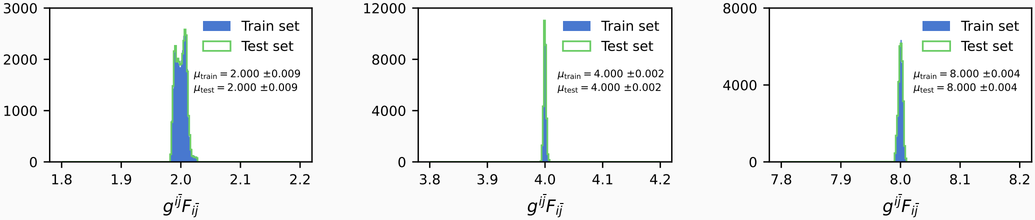
Elliptic curve defined by

$$f(z) \equiv z_1^3 - z_0^2 z_1 - z_0 z_2^2 + z_0^3 = 0 \quad \subset \mathbb{P}^2$$

Approximate CY metric computed with $\sigma = 0.0001$

Neural networks of depth $D = 2, 3, 4$ with intermediate $W = 40$ layers

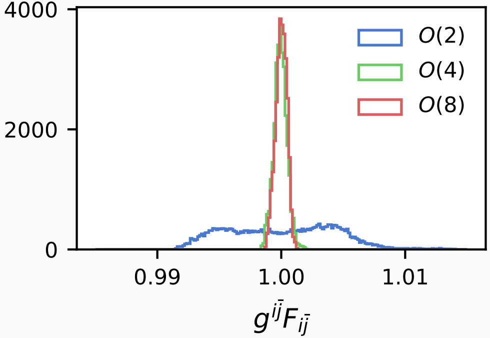
- Histogram of values of $g^{i\bar{j}} F_{i\bar{j}}$ – should be constant over X



$\mathcal{O}_X(1)$ on elliptic curve

$D = 2, 3, 4$ networks give connections on $\mathcal{O}_X(2)$, $\mathcal{O}_X(4)$ and $\mathcal{O}_X(8)$ – how do we compare their accuracy?

Untwist to give connections on $V = \mathcal{O}_X(1)$



All within 1% of expected result $g^{i\bar{j}}F_{i\bar{j}} = 1$

$\mathcal{O}_X(m)$ on quintic threefold

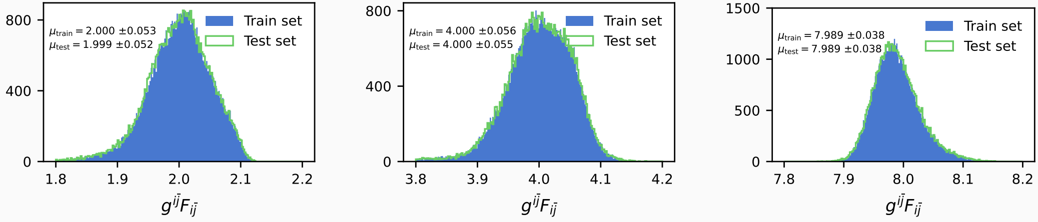
Dwork quintic defined by

$$f(z) \equiv z_0^5 + \dots z_4^5 + \tfrac{1}{2} z_0 z_1 z_2 z_3 z_4 = 0 \qquad \subset \mathbb{P}^4$$

Approximate CY metric computed with $\sigma = 0.001$

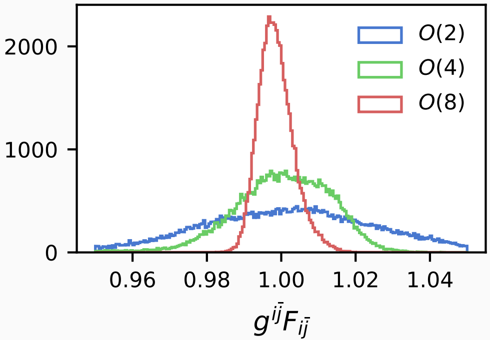
Neural networks of depth $D = 2, 3, 4$ with intermediate $W = 100$ layers

- Histogram of values of $g^{i\bar{j}} F_{i\bar{j}}$ – should be constant over X



$\mathcal{O}_X(1)$ on quintic threefold

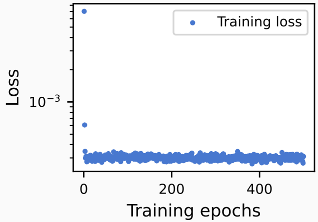
$D = 2, 3, 4$ networks give connections on $\mathcal{O}_X(2)$, $\mathcal{O}_X(4)$ and $\mathcal{O}_X(8)$ – untwist to give connections on $V = \mathcal{O}_X(1)$

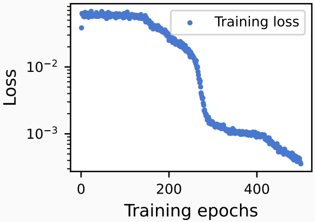


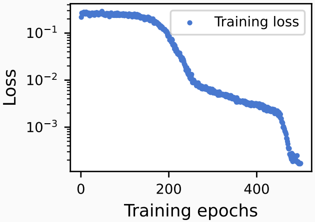
All within 5% of expected result $g^{i\bar{j}}F_{i\bar{j}} = 1$

$\mathcal{O}_X(m)$ on quintic threefold

Loss curves show that $D = 2$ network is not sufficiently complex to capture HYM connection – underparametrised







(reader beware – scale not normalised)

Summary and outlook

Calabi–Yau metrics and HYM connections are accessible with numerical methods and machine learning

- Can extend to other line bundles (and non-Abelian?)
- Can also compute Laplacian spectrum – source of new, non-BPS information (KK spectrum [AA '20]), and essential for string models
- Mirror symmetry? 2d CFTs? [Afkhami-Jeddi, AA, Córdova '21] Many other geometric questions!
- Neural networks as general PDE solvers? G_2 metrics? Non-Kähler metrics?