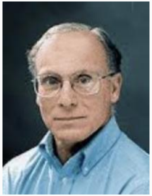
Machine learning for geometry and string compactifications

Anthony Ashmore

LPTHE, Sorbonne Université

Rencontres théoriciennes, 21 September 2023

Collaborators



Burt Ovrut

Yang-Hui He

Elli Heyes

Fabian Ruehle

Nima Afkhami-Jeddi

Clay Córdova

Overview

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

Numerical methods are the only way to access this data

Machine learning and neural networks provide a powerful set of tools to tackle geometric problems

Outline

Physics from geometry

Calabi–Yau metrics

Hermitian Yang–Mills connections

Machine learning and neural networks

Applications

Physics from geometry

Motivation from physics

Does string theory describe our universe? Many semi-realistic MSSM-like string models from M-theory / F-theory / heterotic [...; Cole et al. '21; Abel et al. '21; Loges, Shiu '21, '22;...]

- Focus on models from heterotic string on Calabi–Yau

Coarse details: correct gauge group, matter spectrum, etc.

- Topological – do not need details of geometry

How many of these string vacua are physically reasonable?

- Predicted masses and couplings depend intricately on underlying geometry, i.e. metric and gauge connection
- No analytically known (non-trivial) Calabi–Yau metrics or 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 equipped with Calabi–Yau metric g
- V admits hermitian Yang–Mills connection A
- Bianchi identity: $p_1(X) = p_1(V)$

Particle spectrum of low-energy theory determined by X and V

- e.g. standard embedding: SU(3) bundle gives E_6 GUT gauge group in 4d with $\frac{1}{2}\chi(X)$ particle generations
- Most interesting MSSM examples from non-standard embedding, but not so simple... [...;Donagi et al. '98; Braun et al. '05; Anderson et al. '11;...]

Low-energy physics

Compactification on X leads to 4d $N = 1$ effective theory with gauge + chiral multiplets.

- Chiral multiplets split into moduli fields and matter fields

Particle content comes from topology of X and V, e.g.

- SU(3) bundle V gives E_6 GUT group in 4d

$$E_8 \to E_6 \times SU(3)$$
$$\underline{248} \to \bigoplus_{\underline{r},\underline{R}} (\underline{r}, \underline{R}) = (\underline{78}, \underline{1}) \oplus (\underline{1}, \underline{8}) \oplus (\underline{27}, \underline{3}) \oplus (\overline{\underline{27}}, \overline{\underline{3}})$$

- 4d multiplets transforming in $\underline{r}$ come from $H^{0,1}(X, \underline{R})$, e.g. matter fields from $C^I \in H^{0,1}(X, \underline{3})$

Yukawa couplings

Yukawa terms in Standard Model include $\mathcal{L}_{\mathrm{SM}} \supset \mathcal{L}_{\mathrm{Yuk}} = Y^d_{ij} H Q^i d^j + \ldots$

4d $N = 1$ theory $\to$ superpotential and Kähler potential with moduli ϕ

$$W = \lambda_{IJK}(\phi) C^I C^J C^K + \ldots \qquad K = G_{I\bar{J}}(\phi) C^I \bar{C}^J + \ldots$$

- Perturbative superpotential from triple overlap of modes on X

$$\lambda_{IJK} = \int_X \Omega \wedge \mathrm{tr}(C^I \wedge C^J \wedge C^K)$$

- Matter field Kähler potential gives normalisation where C^I are harmonic

$$G_{IJ} = \int_X C^I \wedge \bar{\star}_V C^J$$

A string model wish list

MSSM spectrum, three families, etc. ✓

- Reduces to topology / algebraic methods

Superpotential couplings λ_{IJK} ✓

- Holomorphic – can use algebraic / differential methods

Harmonic modes and Kähler metric G_{IJ} on field space ✗

- Numerical methods

Supersymmetry breaking, moduli stabilisation, etc. ✗

- Soft masses and couplings c.f. $N = 1$ Kähler potential and normalised zero modes [Kaplunovsky, Louis '93; Blumenhagen et al. '09; ...]

The missing ingredients

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

Calabi–Yau metrics

Calabi–Yau geometry

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 metric g and closed two-form $J = \partial\bar{\partial}K$

$$\mathrm{vol}_g \equiv J \wedge J \wedge J$$

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

$$\mathrm{vol}_\Omega \equiv \mathrm{i}\,\Omega \wedge \bar{\Omega}$$

Example: Fermat quintic

Calabi–Yau threefold is quintic hypersurface X 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

$$\boxed{g_{i\bar{j}}(Z, \bar{Z}) = \partial_i \bar{\partial}_{\bar{j}} K(Z, \bar{Z})}$$

How do we measure accuracy?

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

$$\left.\frac{\mathrm{vol}_g}{\mathrm{vol}_\Omega}\right|_p = 1 \quad \Rightarrow \quad R_{i\bar{j}} = 0$$

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

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

The exact CY metric has $\sigma(K) = 0$

How to fix K?

Finding the "best" approximation to the Ricci-flat metric amounts to finding a $K(z, \bar{z})$ that minimises σ

Three approaches:

- "Balanced metric" – iterative procedure [Donaldson '05; Douglas '06; Braun '07]
- Minimise σ given "algebraic metric" ansatz [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, '22]

In all cases, numerical integrals carried out by Monte Carlo [Shiffman, Zelditch '98]

Hermitian Yang–Mills connections

Hermitian Yang–Mills

A hermitian metric G on fibers of vector bundle V defines a connection and curvature

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

We say A is hermitian Yang–Mills if

$$g^{i\bar{j}} F_{i\bar{j}} = \mu(V)\,\mathrm{Id}$$

G is then known as a Hermite–Einstein metric on V

- Nonlinear PDE for G with no closed-form solutions when X is Calabi–Yau
- HYM implies Yang–Mills: $\mathrm{d} \star F = 0$
- Supersymmetry in 10d requires HYM with $\mu(V) = 0$

Existence and stability

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

A holomorphic vector bundle V over a compact Kähler manifold (X, g) 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!

How do we measure accuracy?

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

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

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

Suitable choice of accuracy measure is

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

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

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

The goal

There is an iterative method to compute HYM connections, but slow, computationally intensive and relatively inaccurate [Wang '05; Douglas et al. '06; Anderson et al. '10]

Train a neural network to find solutions to the hermitian Yang–Mills equation

Machine learning and neural networks

Overview

New era of big data in string theory

- Vacuum selection problem, huge number of CYs, even larger number of flux vacua [Denef, Douglas '04; Taylor, Wang '15;...]

Many different types of machine learning

- Supervised – known inputs and outputs, e.g. recognise images, predict Hodge numbers [He '17; Bull et al. '18; Erbin, Finotello '20;...]
- Unsupervised – known inputs, e.g. looking for patterns or generate images
- Self-supervised – known inputs, output minimises a loss function, e.g. QM ground states, Ricci-flat metrics, HYM connections

Neural networks

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

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

Why does this work? Universal approximation theorem for NNs
[Cybenko '89]

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

Line bundles on CY manifolds

Line bundles crucial 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, so always admit a solution to HYM, $g^{i\bar{j}} F_{i\bar{j}} = \mu(L)$

We need the functional form of G to calculate harmonic representatives and the matter field Kähler metric

Bihomogenous networks on $X \subset \mathbb{P}^2$ [Douglas et al. '20]

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

$$Z_i \mapsto (\operatorname{re} Z_j \bar{Z}_k, \operatorname{im} Z_j \bar{Z}_k) \qquad \vec{x} \mapsto (W_1 \vec{x})^2 \qquad \vec{y} \mapsto \log(W_2 \vec{y})$$

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

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

A loss function

Network output is treated as $\log G^{-1}$, which defines F [AA, Deen, He, Ovrut '20]

- Together with approximate CY metric g, this gives $F_g[\vec{w}]$ as a function of the network weights $\vec{w}$

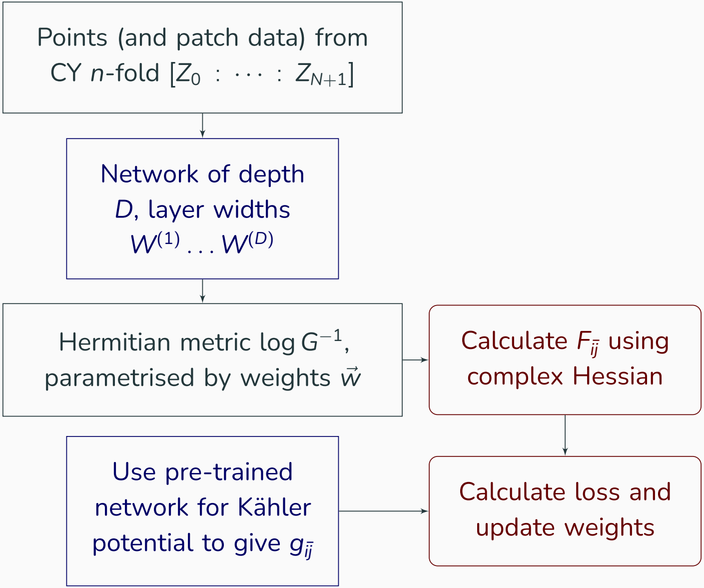
Loss function is

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

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

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

General strategy



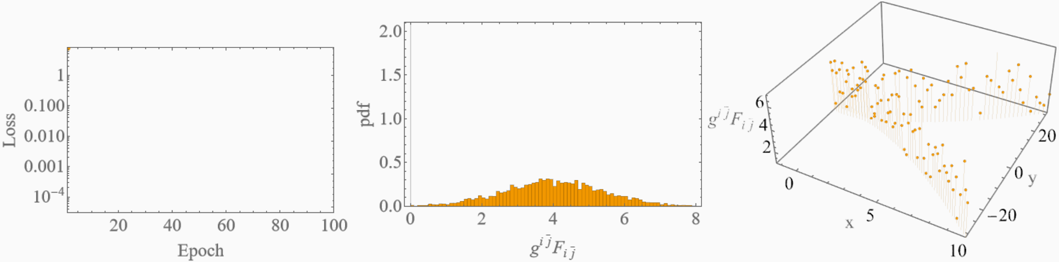
$\mathcal{O}_X(4)$ on elliptic curve

Line bundle $\mathcal{O}(4)$ over elliptic curve defined by

$$Q(Z) \equiv Z_1^3 - Z_0^2 Z_1 - Z_0 Z_2^2 + Z_0^3 = 0 \quad \subset \mathbb{P}^2$$

- Solution to HYM should give $g^{i\bar{j}} F_{i\bar{j}} = 4$ pointwise

Evolution of loss, pdf of $g^{i\bar{j}} F_{i\bar{j}}$ and values of $g^{i\bar{j}} F_{i\bar{j}}$ on elliptic curve



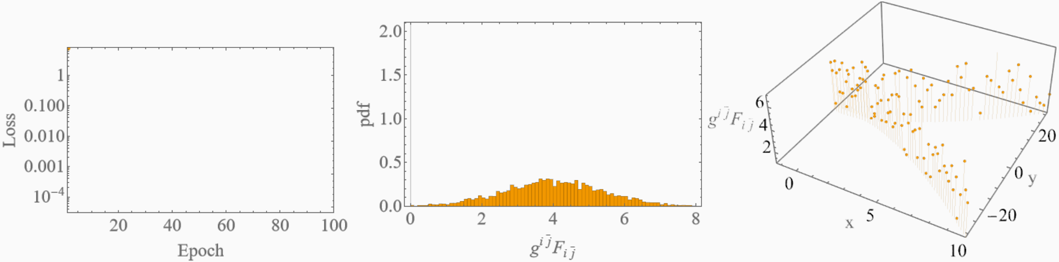
$\mathcal{O}_X(4)$ on elliptic curve

Line bundle $\mathcal{O}(4)$ over elliptic curve defined by

$$Q(Z) \equiv Z_1^3 - Z_0^2 Z_1 - Z_0 Z_2^2 + Z_0^3 = 0 \quad \subset \mathbb{P}^2$$

- Solution to HYM should give $g^{i\bar{j}} F_{i\bar{j}} = 4$ pointwise

Evolution of loss, pdf of $g^{i\bar{j}} F_{i\bar{j}}$ and values of $g^{i\bar{j}} F_{i\bar{j}}$ on elliptic curve



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

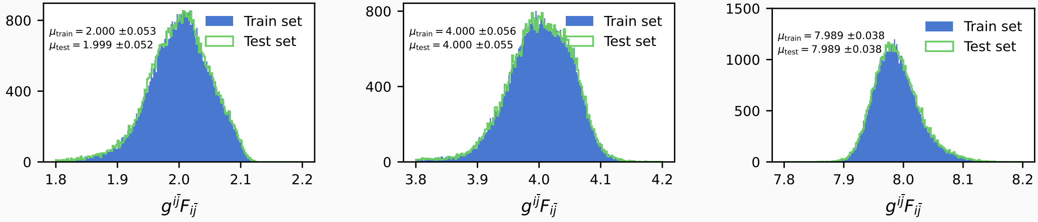
Dwork quintic defined by

$$Q(Z) \equiv Z_0^5 + \cdots + Z_4^5 + \frac{1}{2} Z_0 Z_1 Z_2 Z_3 Z_4 = 0 \quad \subset \mathbb{P}^4$$

Approximate CY metric computed with $\sigma = 0.001$

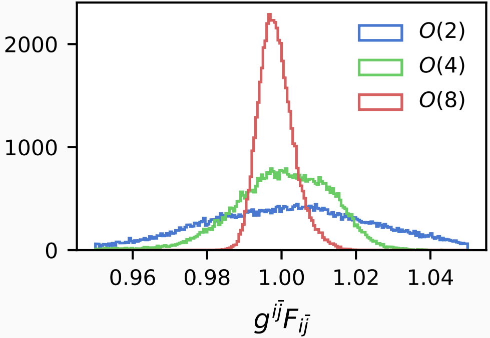
Neural networks of depth $D = 2, 3, 4$ with intermediate $W = 100$ layers

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



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

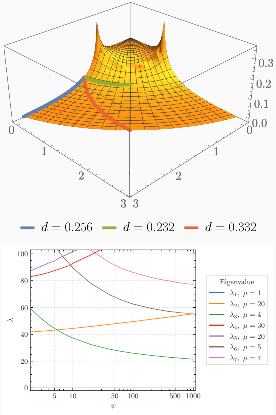
$D = 2, 3, 4$ networks give connections on $\mathcal{O}_X(2)$, $\mathcal{O}_X(4)$ and $\mathcal{O}_X(8)$ – untwist to give connections on $V = \mathcal{O}_X(1)$



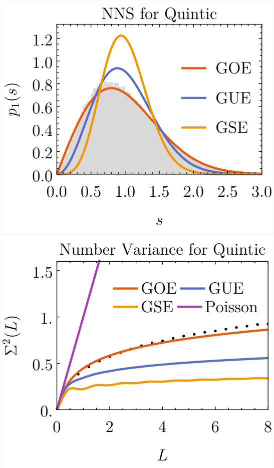
Loss curves show that $D = 2$ network is underparametrised, but all still within 5% of expected result $g^{i\bar{j}}F_{i\bar{j}} = 1$

Applications

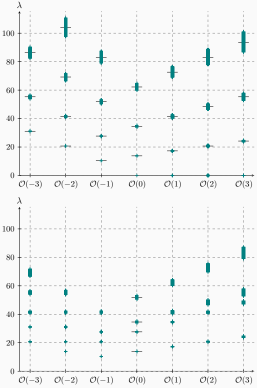
Applications



Swampland distance conjecture



CFT data and random matrices



Laplacian spectra

Matter fields and harmonic modes

Matter fields C^I are bundle-valued $(0,1)$-forms, harmonic wrt the Dolbeault Laplacian

$$\Delta_{\bar{\partial}_V} = \bar{\partial}_V^\dagger \bar{\partial}_V + \bar{\partial}_V \bar{\partial}_V^\dagger, \qquad \Delta_{\bar{\partial}_V} C_I = 0$$

- $\bar{\partial}_V \colon \Omega^{p,q}(V) \to \Omega^{p,q+1}(V)$ is Dolbeault operator
- λ_n are real and non-negative and can appear with multiplicity (continuous or finite symmetries)
- $\Delta_{\bar{\partial}_V}$ requires knowledge of both CY metric on manifold and HYM connection on bundle

Focus on case of hypersurface $X \subset \mathbb{P}^N$ with abelian bundle $V = \mathcal{O}(m)$ for $m \in \mathbb{Z}$

Dolbeault Laplacian [Braun et al. '08; AA '20; AA, He, Heyes, Ovrut '23]

Want both the spectrum $\{\lambda_n\}$ and the eigenmodes $\{\phi_n\}$

$$\boxed{\Delta_{\bar{\partial}_V}\phi_n = \lambda_n\phi_n}$$

QM of charged particle in monopole background [...; Tejero Prieto '06; ...; Bykov, Smilga '23]

Given a basis of modes $\{\alpha_A\}$, expand eigenmode as

$$\phi = \sum_A \langle\alpha_A, \phi\rangle\, \alpha_A = \sum_A \phi_A\, \alpha_A, \qquad A = 1, \dots, \infty$$

to give an eigenvalue problem for λ and ϕ_A

$$\Delta_{AB}\phi_B = \lambda\, O_{AB}\phi_B \qquad \text{where } O_{AB} \equiv \langle\alpha_A, \alpha_B\rangle = \int_X \bar{\star}_V\alpha_A \wedge \alpha_B$$

Approximate basis [Braun et al. '08; AA '20; AA, He, Heyes, Ovrut '23]

Basis $\{\alpha_A\}$ is infinite dimensional – truncate to a finite approximate basis at degree k_ϕ in Z^I. For example,

$$\{\alpha_A\} = \mathcal{F}^{0,0}_{k_\phi}(m) = \frac{(\text{degree } k_\phi + m \text{ in } Z)(\text{degree } k_\phi \text{ in } \bar{Z})}{(Z^I \bar{Z}_I)^{k_\phi}}$$

gives finite set of $\mathcal{O}_{\mathbb{P}^N}(m)$-valued scalars

- $\mathcal{F}^{0,0}_0(m) \subset \mathcal{F}^{0,0}_1(m) \subset \cdots \subset \Omega^{0,0}(\mathcal{O}_{\mathbb{P}^N}(m))$
- Larger values of k_ϕ better approximate the space – c.f. first k_ϕ-th eigenspaces on $\mathbb{P}^N$
- Can construct similar sets of modes for $m < 0$ and $(0,1)$-forms, etc.

Strategy

1. Specify the CY hypersurface by $Q = 0$ and compute metric numerically
2. Specify the bundle $V = \mathcal{O}(m)$ and compute the HYM connection numerically
3. Compute matrices Δ_{AB} and O_{AB} numerically at degree k_ϕ for $\mathcal{O}(m)$-valued $(0, 1)$-forms
4. Compute eigenvalues and eigenvectors to find harmonic modes

Warm-up: a torus as a Calabi–Yau one-fold

Two-dimensional flat tori are Calabi–Yau and their spectrum can be computed *explicitly* [Milnor '63, Tejero Prieto '06]

- Parametrised by $\tau \equiv a + ib$ where lattice generated by $(1, 0)$ and (a, b)

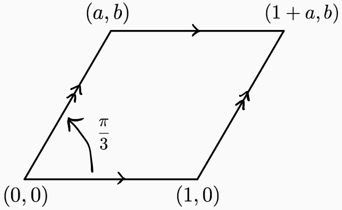
$\mathcal{O}(m)$-valued scalar eigenvalues

$$\{\lambda\}_m^{0,0} = \begin{cases} \frac{6\pi m n}{b} & m > 0, n \geq 0 \\ \frac{4\pi^2}{b}\left[(n_1^2 + n_2^2)m^2 - 2a\, n_1 n_2 + n_2^2\right] & m = 0, n_i \in \mathbb{Z} \\ \frac{6\pi |m|(n+1)}{b} & m < 0, n \geq 0 \end{cases}$$

- No zero-modes for $m < 0$
- Serre duality implies $\{\lambda\}_{-m}^{0,1} = \{\lambda\}_m^{0,0}$

Warm-up: a torus as a Calabi–Yau one-fold

The equilateral torus defined by $\tau = \mathrm{e}^{\mathrm{i}\pi/3}$ – $(1,0)$ and (a,b) generate a hexagonal lattice ($\mathbb{Z}_3$ symmetries)



Equivalent to the Fermat cubic – curve in $\mathbb{P}^2$ defined by

$$Q \equiv Z_0^3 + Z_1^3 + Z_2^3 = 0$$

- Can check numerics against *known* results

Warm-up: a torus as a Calabi–Yau one-fold

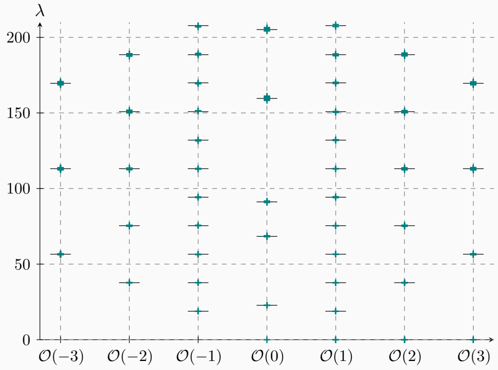
Assume we don't have the CY metric or HYM connecion

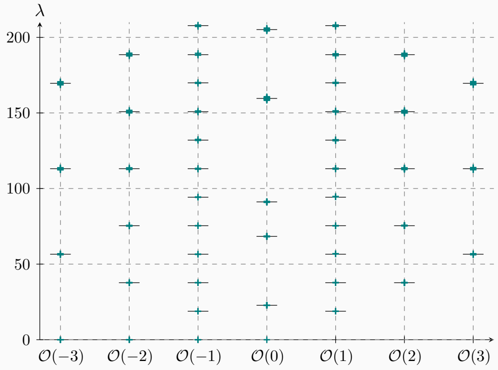
1. Specify the CY by $Q = 0$ and compute metric numerically
2. Specify the bundle $\mathcal{O}(m)$ and compute connection numerically
3. Pick a finite basis for $\mathcal{O}(m)$-valued $(0, 0)$- and $(0, 1)$-forms at some degree k_ϕ
4. Solve numerically for eigenvalues and eigenmodes of $\Delta_{\bar{\partial}_V}$ using Monte Carlo to evaluate integrals

Compute these using

- 10^6 points for metric, connection and Laplacian
- $k_\phi = 3$ and $m \in \{-3, \ldots, 3\}$

Scalars and (0, 1)-forms on Fermat cubic





$\{\lambda\}_m^{0,0} = \{\lambda\}_{-m}^{0,1}$ as expected ✓

Multiplicities match dimensions of irreps of $(S_3 \times \mathbb{Z}_2) \ltimes (\mathbb{Z}_3 \times \mathbb{Z}_3)$ [Ahmed, Ruehle '23] ✓

Example: Fermat quintic

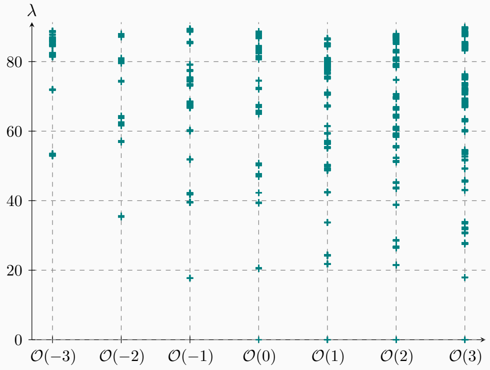
Recall the quintic hypersurface $Q \subset \mathbb{P}^4$

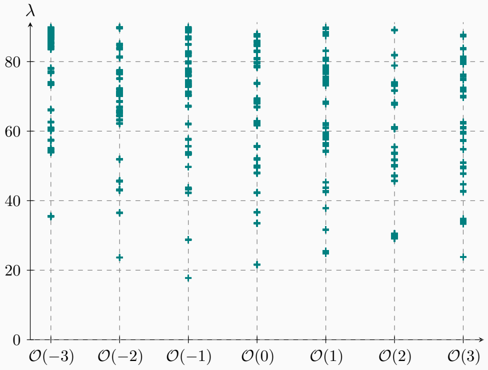
$$Q(z) \equiv Z_0^5 + Z_1^5 + Z_2^5 + Z_3^5 + Z_4^5 = 0$$

Metric not known, no analytic results for spectrum other than counts of zero-modes

- CY metric computed using energy functional method with $\sigma \approx 10^{-4}$
- Monte Carlo integration over 5×10^6 points
- Spectra computed at $k_\phi = 3$

Spectrum of scalars and (0, 1)-forms on Fermat quintic





Zero-modes counted by $h^0(\mathcal{O}(m)) = \binom{4+m}{m}$ for $0 < m < 5$ ✓

$\{\lambda\}^m_{0,1}$ is union of $\{\lambda\}^m_{0,0}$ and (half) of $\{\lambda\}^{-m}_{0,1}$ ✓

- e.g. $\lambda^1_{0,1} = 25.2$ come from $\lambda^1_{0,0} = 21.8, 24, 3$; $\lambda^1_{0,1} = 31.7$ come from $\lambda^{-1}_{0,1} = 28.8$

The superpotential

Consider

$$E_8 \to E_7 \times U(1)$$

where U(1) bundle $V = \mathcal{O}(m)$ gives E_7 GUT group in 4d

$$\underline{248} \to \underline{133}_0 \oplus \underline{56}_1 \oplus \underline{56}_{-1} \oplus \underline{1}_2 \oplus \underline{1}_1 \oplus \underline{1}_{-1}$$

4d matter comes from $C^I \in H^{0,1}(X, \mathcal{O}(m))$

- Numerics (or Kodaira vanishing + Serre duality) imply $H^{0,1}(X, \mathcal{O}(m)) = \{0\}$
- No superpotential matter couplings for this example – need non-abelian bundle or extend to CICY

Summary and outlook

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

Ongoing work: bundle-valued harmonic modes for CICYs, non-abelian bundles

- Compute Yukawa couplings, etc., at chosen point in moduli space

Future work

- SYZ conjecture? Non-Kähler metrics? G_2 metrics? Flux backgrounds? Neural networks as general PDE solvers?
- 2d CFTs? [Afkhami-Jeddi, AA, Córdova '21] Input for conformal bootstrap? [Lin et al. '15; Lin et al. '16;...]