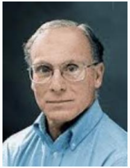
Machine Learning for String Compactifications

Anthony Ashmore

University of Chicago

Collaborators



Burt Ovrut

Yang-Hui He

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

Application: moduli dependence and the swampland

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 explicitly known (non-trivial) Calabi–Yau metrics or appropriate connections!

Calabi–Yau compactifications

Minimal supersymmetry on $\mathbb{R}^{1,3} \times X$ with $\mathrm{E}_8 \times \mathrm{E}_8$ bundle V [Candelas et al. '85]

- No H flux $\Rightarrow X$ equipped with Calabi–Yau metric g
- V admits hermitian Yang–Mills connection A
- Anomaly condition: $p_1(X) = p_1(V)$

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

$$\mathrm{E}_8 \to \mathrm{E}_6 \times \mathrm{SU}(3)$$
$$\underline{248} \to \bigoplus_{\underline{r},\underline{R}} (\underline{r}, \underline{R}) = (\underline{78}, \underline{1}) \oplus (\underline{1}, \underline{8}) \oplus (\underline{27}, \underline{3}) \oplus (\overline{\underline{27}}, \overline{\underline{3}})$$

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

Yukawa couplings

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

4d $N = 1$ theory governed by superpotential and Kähler potential

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

- Holomorphic superpotential from triple overlap of zero modes on X

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

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

$$G_{I\bar{J}} = \int_X C^I \wedge \star_V \bar{C}^J$$

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 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

$$\Omega = \frac{dz_2 \wedge dz_3 \wedge dz_4}{\partial Q / \partial z_1}$$

Metric g and Kähler form J determined by Kähler potential

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

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 $\sigma(K) = 0$

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

Algebraic metrics [Tian '90; Donaldson '05]

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

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

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

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

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 monomials 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

$$K(h) = \log \sum_{\alpha,\beta=1}^{15} s_\alpha h^{\alpha\bar{\beta}} \bar{s}_{\bar{\beta}}, \qquad h^{\alpha\bar{\beta}} \sim 225 \text{ parameters}$$

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

- Spectral method: $\frac{s_\alpha \bar{s}_{\bar{\beta}}}{(z_i \bar{z}_i)^k}$ gives basis for eigenspaces on $\mathbb{P}^4 = S^7 / U(1)$

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, '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 explicit 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.

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

Suitable choice of accuracy measure is

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

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

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

Numerical approximations

Ansatz for hermitian metric [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 on V itself

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

The goal

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, 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
- Universal approximation theorem for NNs [Cybenko '89]

More interested in the network itself than the actual values!

- 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 [Douglas et al. '20]

$$\mathbb{C}^3 \longrightarrow \boxed{\text{Bihom}} \xrightarrow{\mathbb{R}^9} \boxed{\text{Square}} \xrightarrow{\mathbb{R}^{12}} \boxed{\text{Log}} \longrightarrow \mathbb{R}$$

$$\begin{array}{ccc} \mathbb{C}^3 \to \mathbb{R}^9 & \mathbb{R}^9 \to \mathbb{R}^{12} & \mathbb{R}^{12} \to \mathbb{R} \\ z_i \mapsto (\operatorname{re} z_j \bar{z}_k, \operatorname{im} z_j \bar{z}_k) & \vec{x} \mapsto (W_1 \vec{x})^2 & \vec{y} \mapsto \log(W_2 \vec{y}) \end{array}$$

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

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 $\log G^{-1}$, which defines F

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

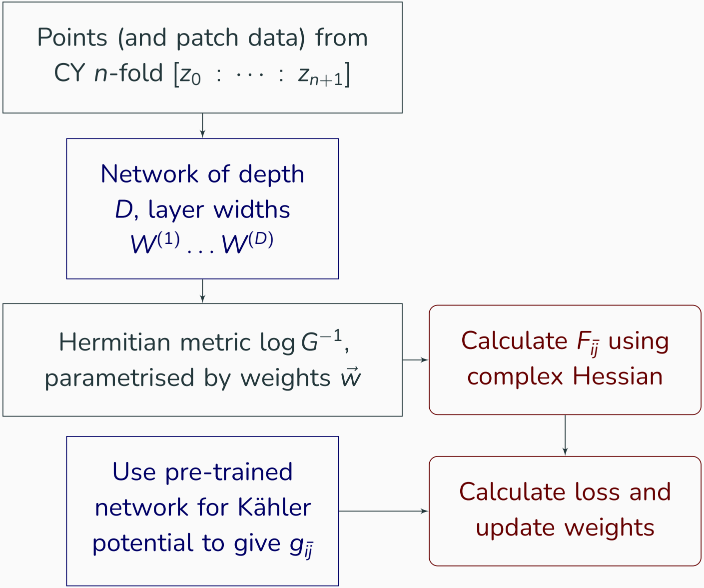
Loss function is

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

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



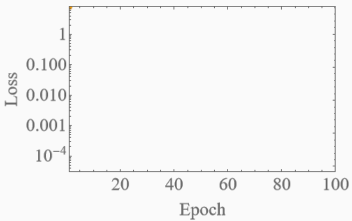
$\mathcal{O}_X(4)$ on elliptic curve

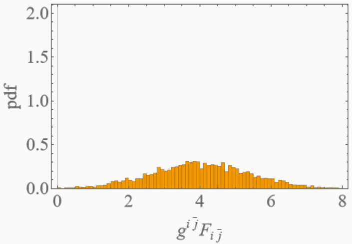
Line bundle $\mathcal{O}(4)$ over elliptic curve defined by

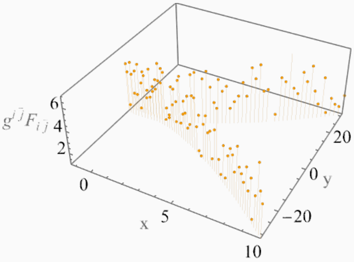
$$Q(z) \equiv z_1^3 - z_0^2 z_1 - z_0 z_2^2 + z_0^3 = 0 \qquad \subset \mathbb{P}^2$$

- 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







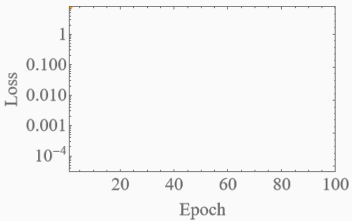
$\mathcal{O}_X(4)$ on elliptic curve

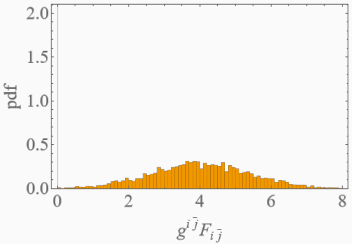
Line bundle $\mathcal{O}(4)$ over elliptic curve defined by

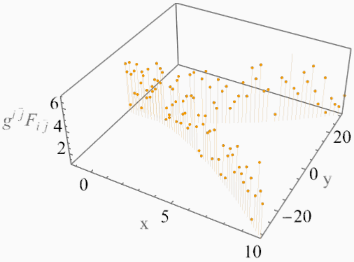
$$Q(z) \equiv z_1^3 - z_0^2 z_1 - z_0 z_2^2 + z_0^3 = 0 \quad \subset \mathbb{P}^2$$

- 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

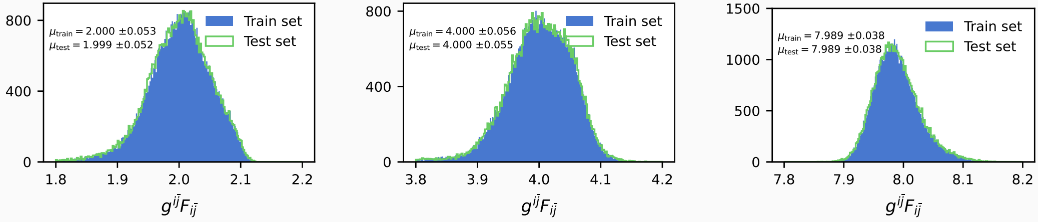
Dwork quintic defined by

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

Approximate CY metric computed with $\sigma = 0.001$

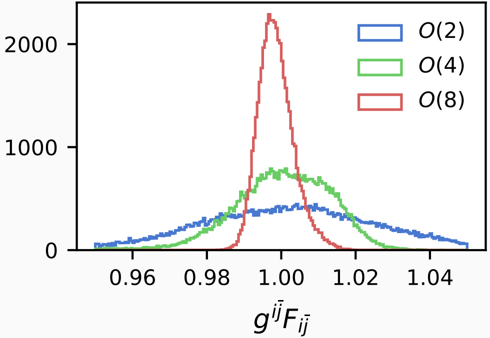
Neural networks of depth $D = 2, 3, 4$ with intermediate $W = 100$ layers

- 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 $V = \mathcal{O}_X(1)$



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

Application: moduli dependence and the swampland

The swampland distance conjecture

As an application of these numerical techniques, we can compute the spectrum of the Laplacian on a Calabi–Yau. This will teach us something about its KK spectrum and effective field theory.

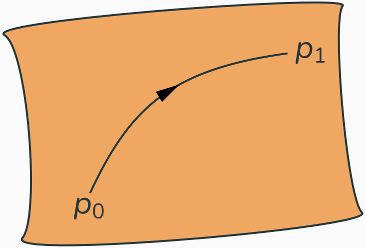
The Swampland Distance Conjecture [Ooguri, Vafa '06]

For an effective theory coupled to gravity with a moduli space parametrised by vevs of some fields, moving an infinite distance in moduli space brings down a tower of massless states that spoil the initial effective theory.

Finite version: there is a relation between the mass of a tower of states and the distance you move in moduli space

The swampland distance conjecture

Compare the effective theory at two points p_0 and p_1 in moduli space which are a geodesic distance $d(p_0, p_1)$ apart



Geodesic γ with $\gamma(\tau_i) = p_i$, distance along curve given by

$$d(p_0, p_1) = \int_{\tau_1}^{\tau_2} \sqrt{G_{ab}\dot{\gamma}^a\dot{\gamma}^b}$$

where G_{ab} is metric on moduli space

Conjecture implies that a tower of states becomes light on moving from p_0 to p_1 with masses

$$m(p_1) \sim m(p_0) e^{-\alpha\, d(p_0,p_1)}$$

where α is order one coefficient

The spectrum with varying complex structure [AA, Ruehle '21]

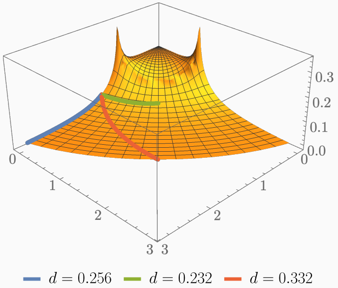
Focus on complex structure moduli space for quintic

$$Q_\psi \equiv z_0^5 + \cdots + z_4^5 - 5\psi z_0 \ldots z_4$$

1. Compute the moduli space metric [Candelas et al. '98; Keller, Lukic '09]

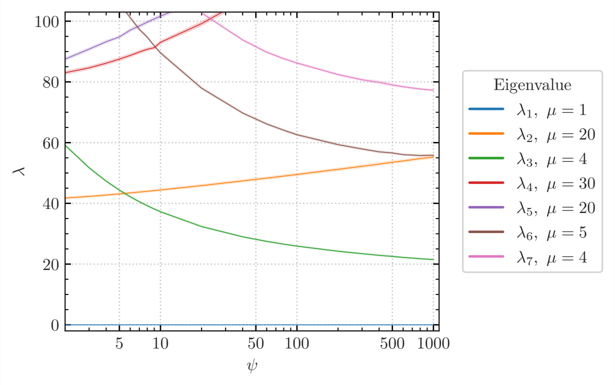
$$e^{-\mathcal{K}_{cs}} = i \int_{Q_\psi} \Omega \wedge \bar{\Omega}, \quad G_{a\bar{b}} = \partial_a \bar{\partial}_{\bar{b}} \mathcal{K}_{cs}$$

2. Compute geodesics and distances in moduli space for varying ψ
3. Compute the numerical Calabi–Yau metric and spectrum of the Laplacian Δ – scalars give spin-two modes



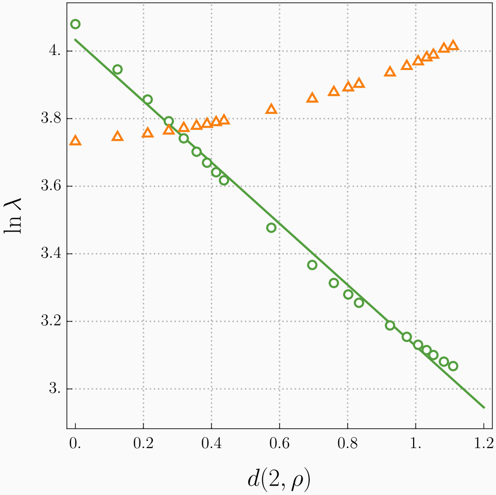
Scalar spectrum with varying ψ

Spectrum of $\Delta\phi = \lambda\phi$



1. Zero mode always present; massive modes appear with multiplicities given by irreps of $(S_5 \times \mathbb{Z}_2) \ltimes (\mathbb{Z}_5)^3$
2. Eigenvalues with small degeneracy become lighter
3. Local maximum in spectral gap for $\psi \approx 4$

A check of the distance conjecture



1. λ_3 falls exponentially as

$$\lambda_3 = 56.4\, e^{-(0.906 \pm 0.034)\, d(2,\rho)}$$
$$\sim e^{-2\alpha\, d(p_0, p_1)}$$

2. We see $\alpha \approx 0.45$ which is indeed order one
3. Almost saturates conjectured lower bound of $1/\sqrt{6} \approx 0.41$

   [Andriot et al. '20]

Summary and outlook

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

- Can extend to other bundles, compute (p, q)-form Laplacian spectrum – source of new non-BPS data [AA '20]

Ongoing work: bundle-valued harmonic forms, compute Yukawa couplings at a given point in moduli space

- SYZ conjecture? Non-Kähler metrics? G_2 metrics? 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]