Computing the Uncomputable:

String Compactifications with Neural Networks

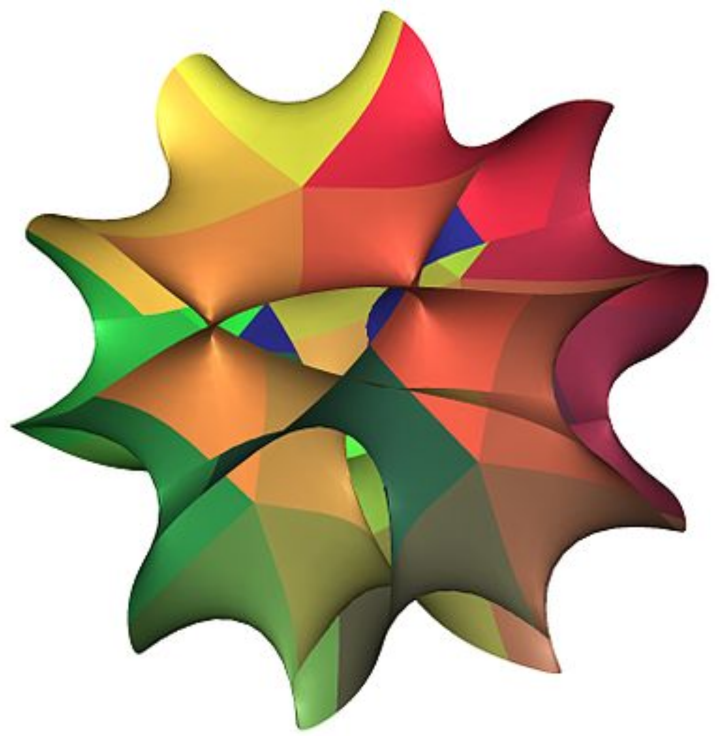
Thomas R. Harvey

Michigan - 2026

11th LCTP Spring Symposium: Theoretical Physics and AI

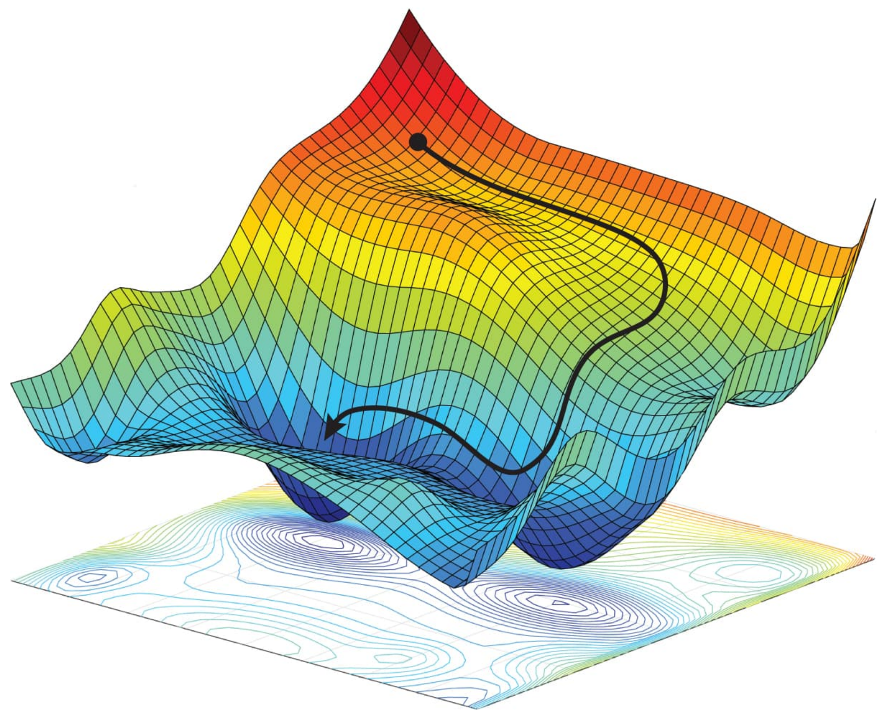
Massachusetts Institute of Technology

AI Applied to Physics



Physics

AI



Physics For AI

Other Work

FunSearch

[Submitted on 14 Mar 2025 (v1), last revised 17 Mar 2025 (this version, v2)]

Generative Modeling for Mathematical Discovery

Jordan S. Ellenberg, Cristofero S. Fraser-Taliente, Thomas R. Harvey, Karan Srivastava, Andrew V. Sutherland

[Submitted on 12 May 2025]

Symbolic Regression with Multimodal Large Language Models and Kolmogorov Arnold Networks

Thomas R. Harvey, Fabian Ruehle, Kit Fraser-Taliente, James Halverson

Optimisation

[Submitted on 3 Sep 2025]

The Optimiser Hidden in Plain Sight: Training with the Loss Landscape's Induced Metric

Thomas R. Harvey

[Submitted on 1 Apr 2026]

Sven: Singular Value Descent as a Computationally Efficient Natural Gradient Method

Samuel Bright-Thonney, Thomas R. Harvey, Andre Lukas, Jesse Thaler

Information Theory and Naturalness

[Submitted on 2 Mar 2026 (v1), last revised 1 May 2026 (this version, v2)]

Naturalness and Fisher Information

James Halverson, Thomas R. Harvey, Michael Nee

Computing the Uncomputable:

String Compactifications with Neural Networks

Thomas R Harvey
Michigan - 2026
11th LITP Spring Symposium: Theoretical Physics and AI

Massachusetts Institute of Technology

NSF

Aim of string compactifications: Understand the field theories that can come from string theory <u>and calculate their properties</u>

[Submitted on 1 Nov 2024]

Not So Flat Metrics

Kit Fraser-Taliente, Thomas R. Harvey, Manki Kim

[Submitted on 2 Feb 2024 (v1), last revised 2 Jul 2024 (this version, v2)]

Computation of Quark Masses from String Theory

Andrei Constantin, Kit Fraser-Taliente, Thomas R. Harvey, Andre Lukas, Burt Ovrut

(+ upcoming long paper)

String Theory

- The obvious problem: String theory (typically) exists in 10 (11) dimensions
- We need initial conditions (i.e. We need to specify a 10D geometry)
- Compactifications - at low energies and large distances:

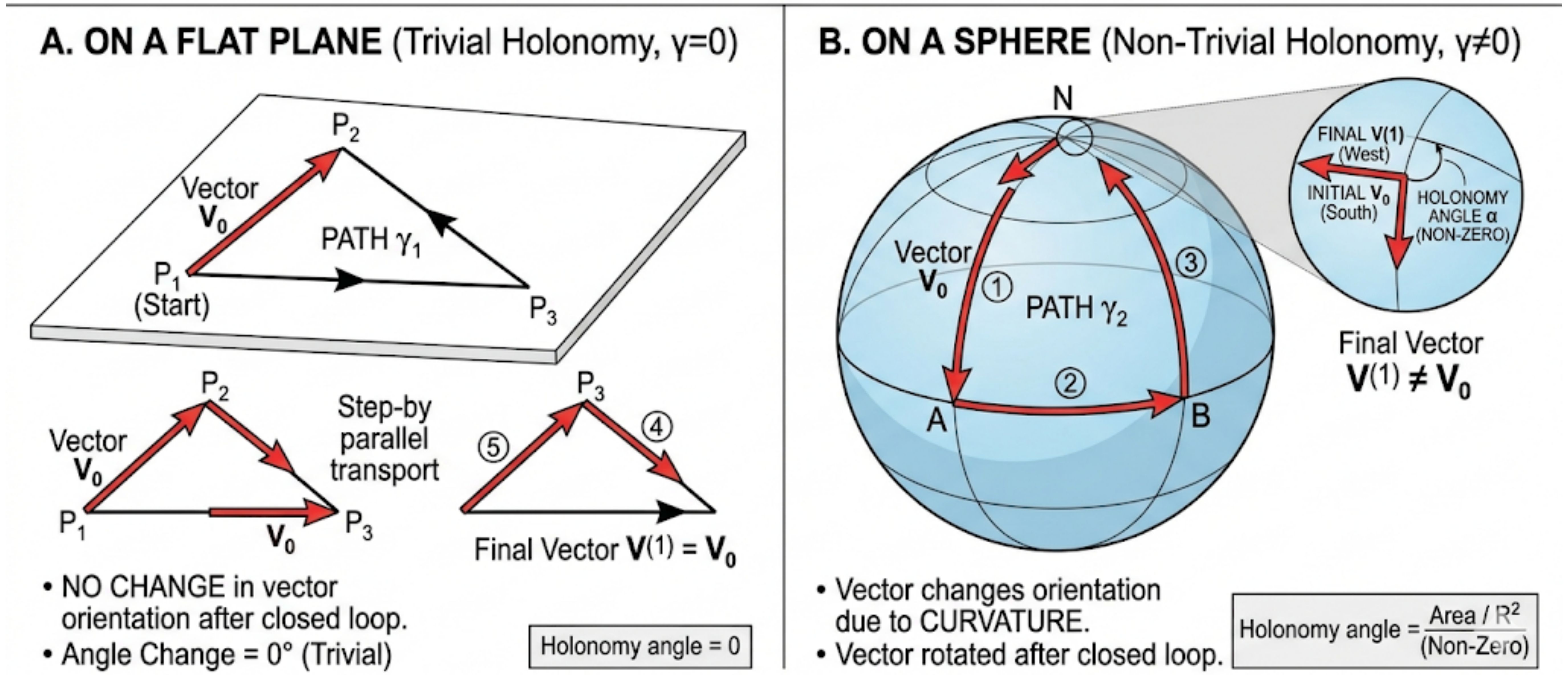
$$M_{10} = \mathbb{R}^{1,3} \times M_6 \xrightarrow{V_6 \to 0} \mathbb{R}^{1,3} \qquad S = \int_{M_{10}} d^{10}x \mathscr{L} \approx V_6 \int_{\mathbb{R}^{1,3}} d^4x \left[\mathscr{L}^{\Lambda}_{eff} + \mathcal{O}\left(\Lambda V_6^{1/6}\right) \right]$$

- Different choices of M_6, and field profiles of it, lead to different 4D Physics $\mathscr{L}^{\Lambda}_{eff}$
- We will make use of supersymmetry - this is not to say we have low energy SUSY
- For our purposes, this will mean that M_6 is a Calabi-Yau manifold (CY 3-fold)

Calabi-Yau Metrics

- Calabi-Yau Manifold: Kähler Manifold with vanishing first Chern Class
- What makes Calabi-Yau manifolds special?
  - There is a metric that is Ricci Flat and has SU(3) Holonomy

$$R_{\mu\nu} = 0$$



The recipe for a Heterotic string compactification

<u>Topology (Discrete Data)</u>

CY Manifold

Gauge Bundle

Wilson Line

Low Energy Massless Spectrum
(Discrete Data of theory)
(Cohomology)

The recipe for a Heterotic string compactification

Topology (Discrete Data)

CY Manifold

Gauge Bundle

Wilson Line

Equations of Motion
(PDEs in 6D)

Unique Solution

1. Yau's Theorem
2. Donaldson-Uhlenbeck-Yau Theorem

Geometry (Solutions to PDEs)

Metric $g_{\mu\nu}$

Gauge Connection A^{μ}

Low Energy Massless Spectrum
(Discrete Data of theory)
(Cohomology)

The recipe for a Heterotic string compactification

Topology (Discrete Data)

CY Manifold

Gauge Bundle

Wilson Line

Equations of Motion
(PDEs in 6D)

Unique Solution

1. Yau's Theorem
2. Donaldson-Uhlenbeck-Yau Theorem

Geometry (Solutions to PDEs)

Metric $g_{\mu\nu}$

Gauge Connection A^{μ}

Low Energy Massless Spectrum
(Discrete Data of theory)
(Cohomology)

Low Energy Couplings
(Continuous Data of Theory)

Calabi-Yau Metrics

- General Method: Physics-Informed-Neural-Networks (PINNs)
- Solving differential equations with neural networks
- The neural network $f_\theta(x)$ is the solution to the differential equations

$$\mathscr{L}[f_\theta] = \sum_i |PDE(f_\theta(x_i))|^2 + |BC(f_\theta(x_i))|^2$$

- We're going to solve the 6D Einstein equations like this!

Calabi-Yau Metrics

Moduli-dependent Calabi-Yau and SU(3)-structure metrics from Machine Learning

Lara B. Anderson, Mathis Gerdes, James Gray, Sven Krippendorf, Nikhil Raghuram, Fabian Ruehle

Numerical Calabi-Yau metrics from holomorphic networks

Michael R. Douglas, Subramanian Lakshminarasimhan, Yidi Qi

Learning Size and Shape of Calabi-Yau Spaces

Magdalena Larfors, Andre Lukas, Fabian Ruehle, Robin Schneider

Numerical Metrics for Complete Intersection and Kreuzer-Skarke Calabi-Yau Manifolds

Magdalena Larfors, Andre Lukas, Fabian Ruehle, Robin Schneider

CYJAX: A package for Calabi-Yau metrics with JAX

Mathis Gerdes, Sven Krippendorf

cymyc -- Calabi-Yau Metrics, Yukawas, and Curvature

Per Berglund, Giorgi Butbaia, Tristan Hübsch, Vishnu Jejjala, Challenger Mishra, Damián Mayorga Peña, Justin Tan

Calabi-Yau Metrics

- The simplest Calabi-Yau 3-fold manifold is the quintic $\mathscr{M}$ in $\mathbb{P}^4$

$$\mathbb{P}^4 = \frac{\mathbb{C}^5 - \vec{0}}{\mathbb{C}^*}$$

$$\overrightarrow{X} = (X_0, X_1, X_2, X_3, X_4) \sim (\lambda X_0, \lambda X_1, \lambda X_2, \lambda X_3, \lambda X_4) \quad \forall X \in \mathbb{P}^4 \quad \forall \lambda \in \mathbb{C}^*$$

$$\mathscr{M} \subset \mathbb{P}^4, \quad p = \sum_{n=0}^{4} X_n^5 = 0$$

- There is a “reference metric” we have on $\mathbb{P}^4$: $g_{a\bar{b}}^{FS} = \partial_a \partial_{\bar{b}} \log \left(\sum_n |X_n|^2 \right)$

- Related to the Ricci-flat metric on $\mathscr{M}$: $g_{i\bar{j}} = (g^{FS}|_{\mathscr{M}})_{i\bar{j}} + \partial_i \partial_{\bar{j}} \phi(X, \bar{X})$

Calabi-Yau Metrics

- This relationship will help us a lot! It makes the problem tractable
- Analogous expressions also exist for other CY manifolds

$$g_{i\bar{j}} = (g^{FS}|_{\mathscr{M}})_{i\bar{j}} + \partial_i \partial_{\bar{j}} \phi(X, \bar{X})$$

- We represent ϕ with a neural network, and it has some useful properties
  - Unique solution - No BCs needed!
  - Global function - Unlike a Kähler potential or full metric

(Equivariant) Projective Neural Networks

$$\mathbb{P}_1 = \frac{\mathbb{C}^2 - \vec{0}}{\mathbb{C}^*} \quad \Rightarrow \quad [x, y] = [\lambda x, \lambda y] \ \forall \lambda \in C^*$$

Functions: $f(\lambda x, \lambda y) = f(x, y)$

$$\pi_\theta : [x, y] \to \left[\frac{x\bar{y}}{|x|^2 + |y|^2}, \frac{x\bar{x}}{|x|^2 + |y|^2}, \frac{y\bar{y}}{|x|^2 + |y|^2} \right] \to \cdots\text{feed-forward network}\cdots \to \mathbb{R}$$

Sections: $\sigma \in \mathcal{O}_{\mathbb{P}_1}(n) \Rightarrow \sigma(\lambda x, \lambda y) = \lambda^n \sigma(x, y)$

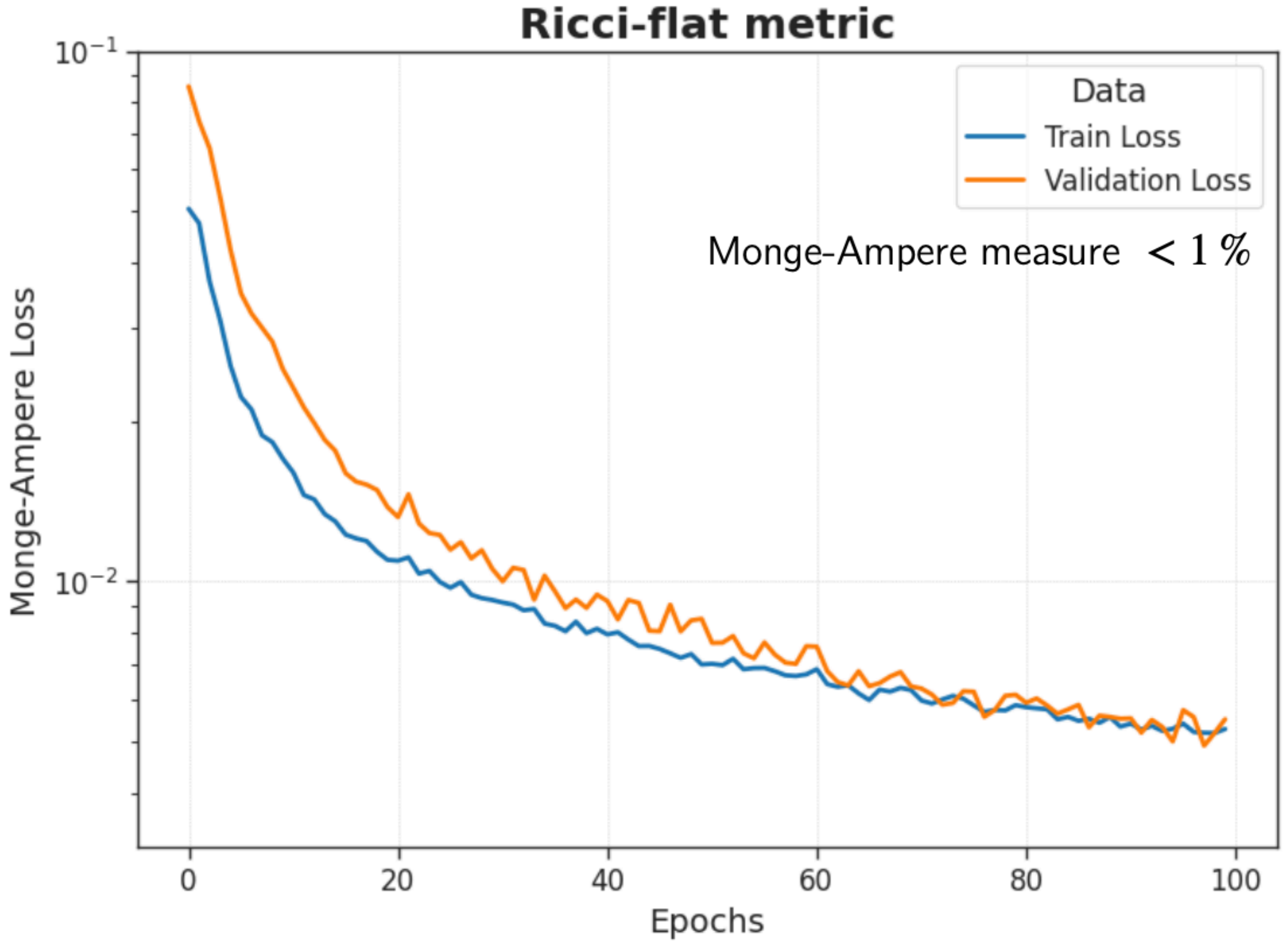
$$n = 2 \quad \pi_\theta : [x, y] \to \left[\frac{x\bar{y}}{|x|^2 + |y|^2}, \frac{x\bar{x}}{|x|^2 + |y|^2}, \frac{y\bar{y}}{|x|^2 + |y|^2} \right] \to \cdots\text{feed-forward network}\cdots \to \mathbb{R}^6 \to (a, b, c) \in C^3 \to ax^2 + bxy + cy^2$$

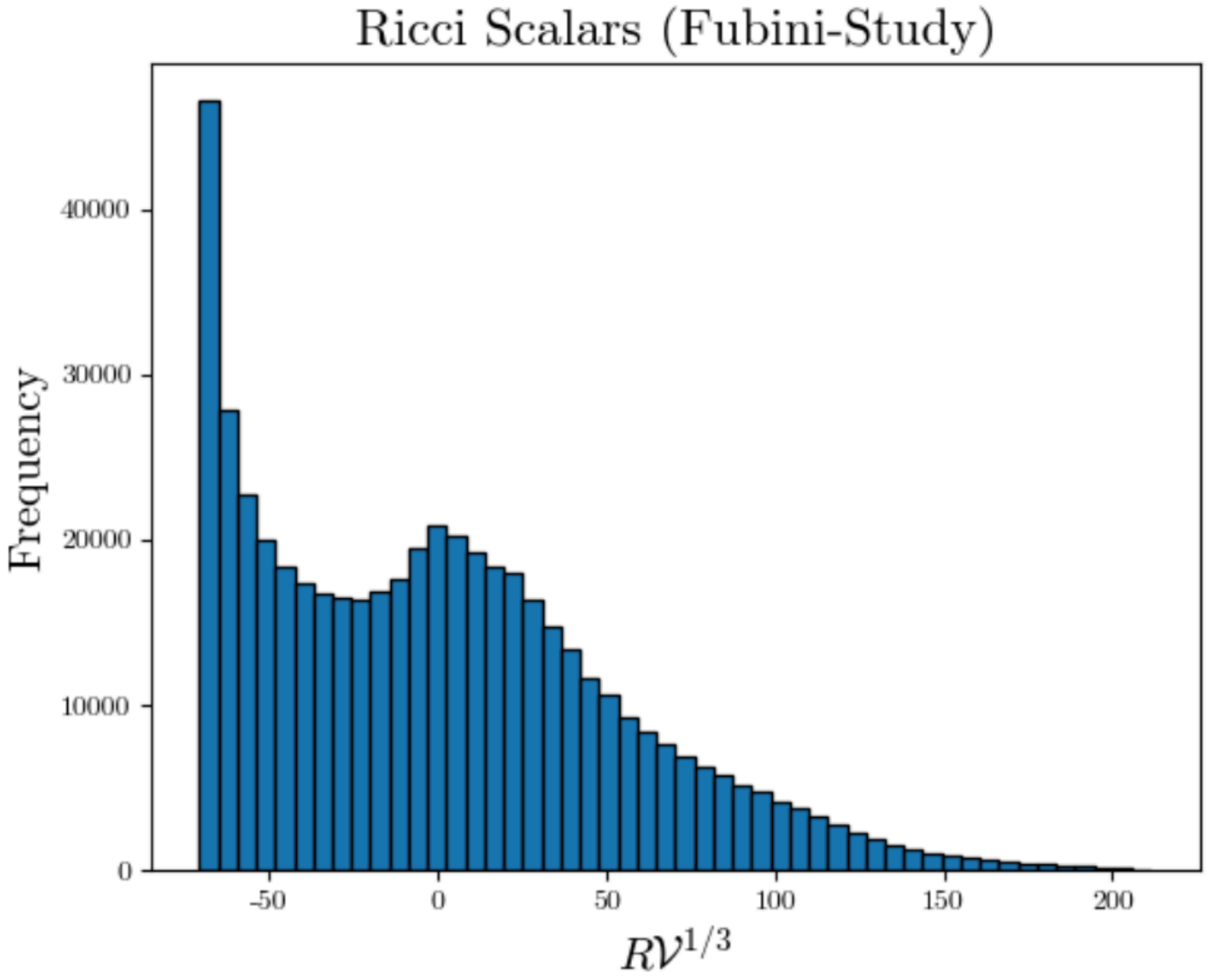
Generalises to other projective spaces and CYs

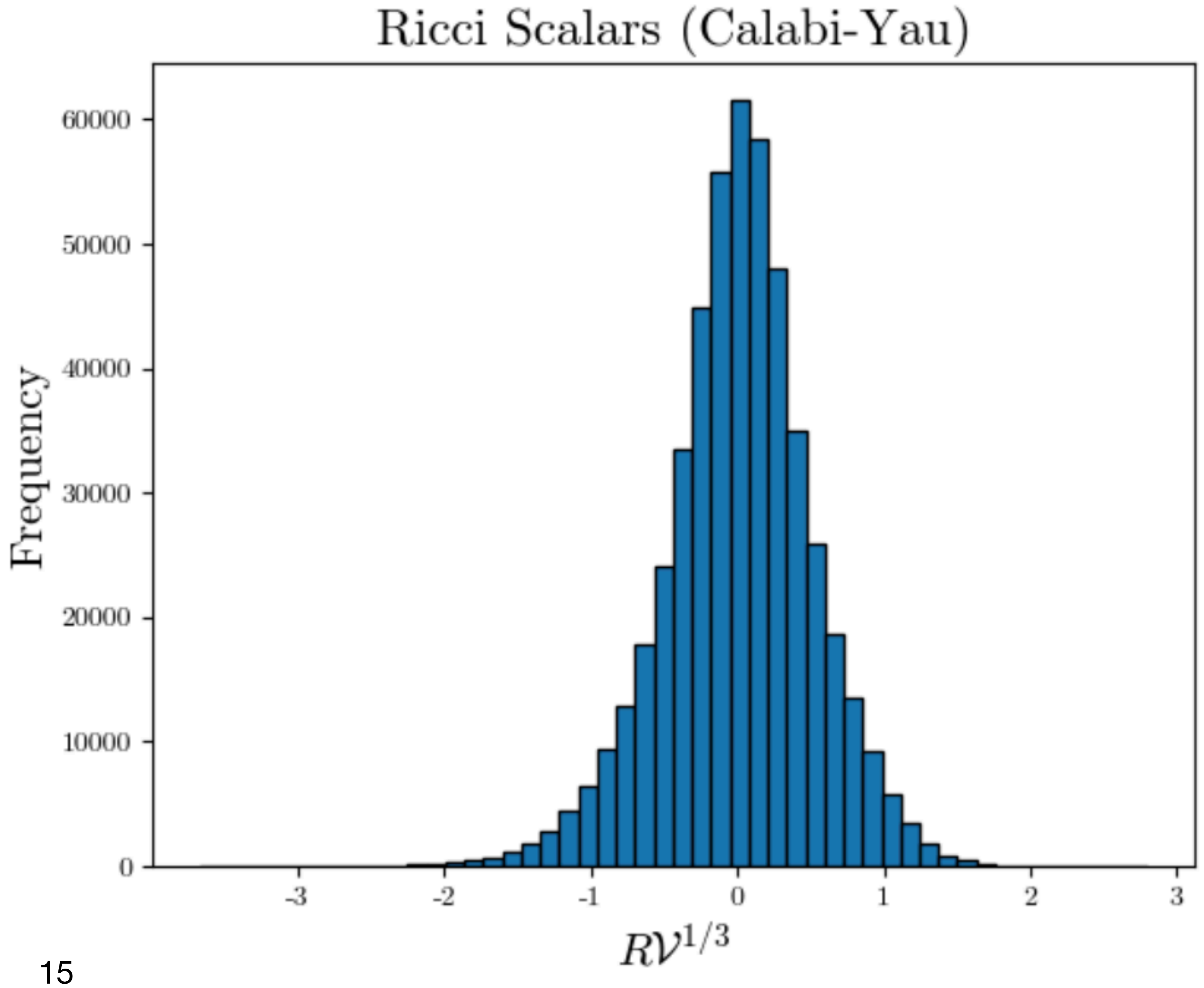
Calabi-Yau Metrics

- So what is the recipe?
    - 1. Sample the manifold with a known distribution (Shiffman & Zelditch)
    - 2. Find ϕ via gradient descent, insuring it is a global function, such that the manifold is Ricci Flat
- We actually do this with the Monge-Ampere Loss

$$R_{a\bar{b}}(g^{FS} + \partial\bar{\partial}\phi) = 0 \Longleftrightarrow 1 - \frac{1}{\kappa}\frac{(g^{FS} + \partial\bar{\partial}\phi)^{\wedge^3}}{\Omega \wedge \bar{\Omega}} = 0$$


Ricci-flat metric
Data
Train Loss
Validation Loss
Monge-Ampere measure $< 1\,\%$
Monge-Ampere Loss
10^{-1}
10^{-2}
0
20
40
60
80
100
Epochs


Ricci Scalars (Fubini-Study)
Frequency
40000
30000
20000
10000
0
-50
0
50
100
150
200
$R\mathcal{V}^{1/3}$


Ricci Scalars (Calabi-Yau)
Frequency
60000
50000
40000
30000
20000
10000
0
-3
-2
-1
0
1
2
3
$R\mathcal{V}^{1/3}$

Application 1:
EFT Validity

EFT Validity

- 10D SUGRA theory understood in two expansions
  - String Coupling: $\sum_n \mathscr{A}_n g_s^n$
  - Higher derivatives: $\sum_n \mathscr{B}_n \alpha'^n$
- For validity of the resulting 4D EFT, we must check both!

EFT Validity

10D EFT:

$$\mathcal{L} \propto R + \sum_{n=L} \alpha'^n f_n(R_{MNPQ}),$$

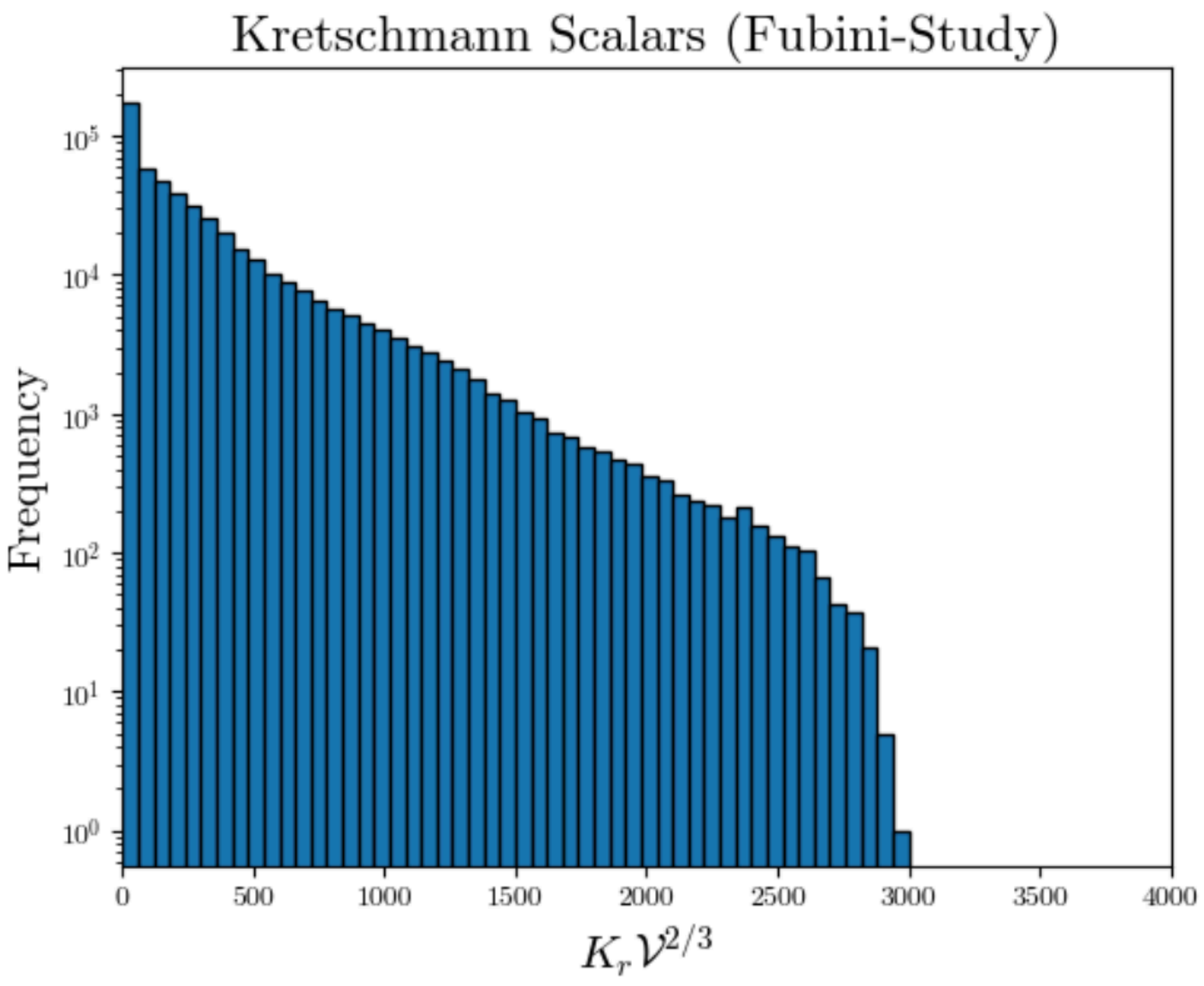
Solve zeroth order equations, then check $(2\pi)^2 \alpha'^{L+1} \mathcal{V}^{2/6} |f_L(R^0_{MNPQ})| \ll 1$

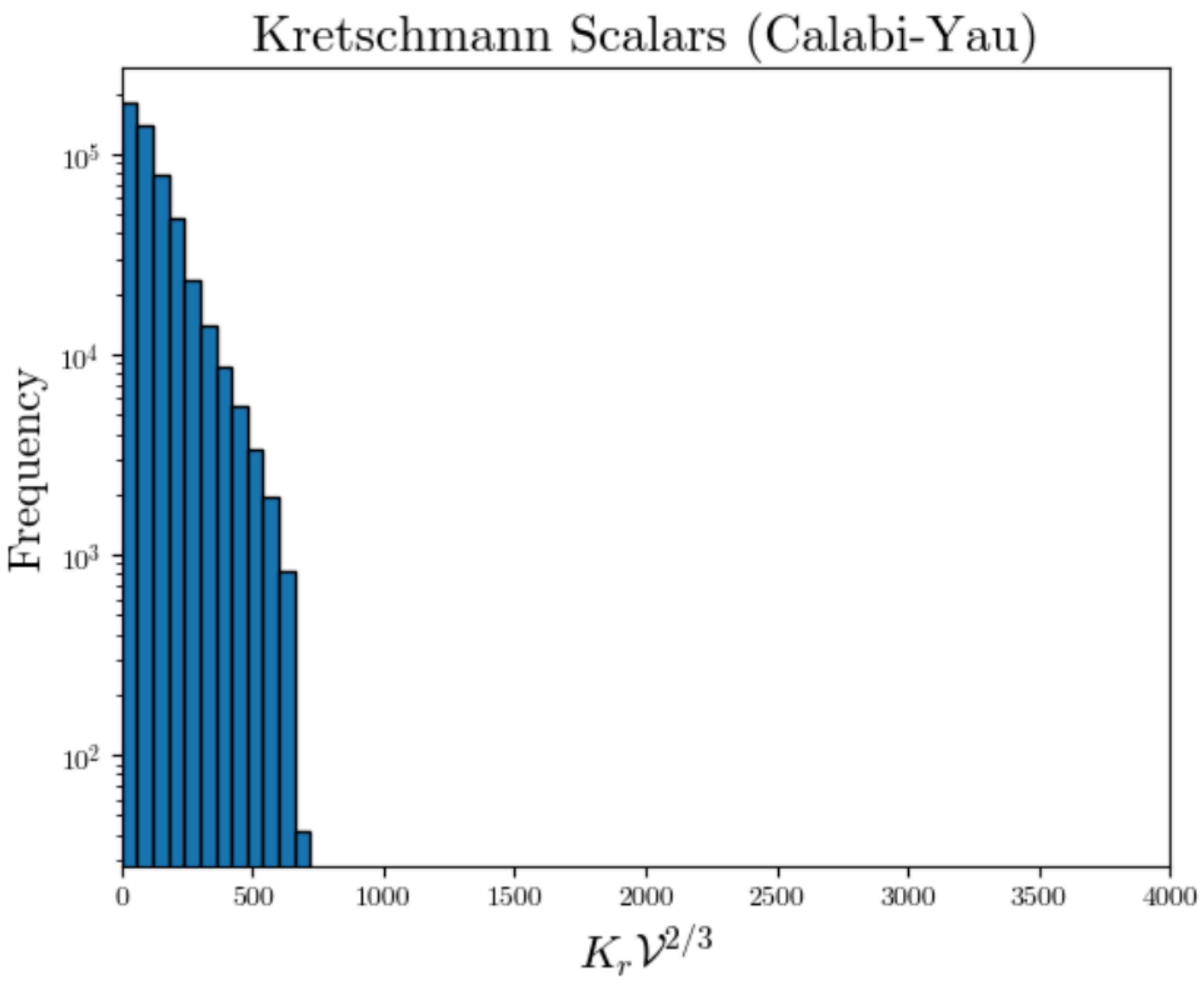
Without access to the metric, people usually just check: $\mathcal{V} \gg 1.$

EFT Validity

$$K_r = \frac{1}{2\pi} R^{\bar{m}n\bar{p}q} R_{n\bar{m}q\bar{p}},$$

This is for the quintic CY3





$$K_r \ll 1 \Rightarrow \mathscr{V} \gg 10^4 - 10^5$$

For Comparison: $\mathscr{V} \sim \alpha_{GUT}^{-1} \sim 25$

EFT Validity

Important caveat - this is not always so bad!

However, this is evidence that this is an important thing to check, especially when considering moduli stabilisation!

Candidate de Sitter Vacua

Liam McAllister,[a] Jakob Moritz,[b] Richard Nally,[a] and Andreas Schachner[a,c]

A long way to go before such a calculation can be completed there

(Ongoing work for LVS vacua)

Application 2:
Yukawa Couplings and Quark Masses

Yukawa Couplings in String Theory

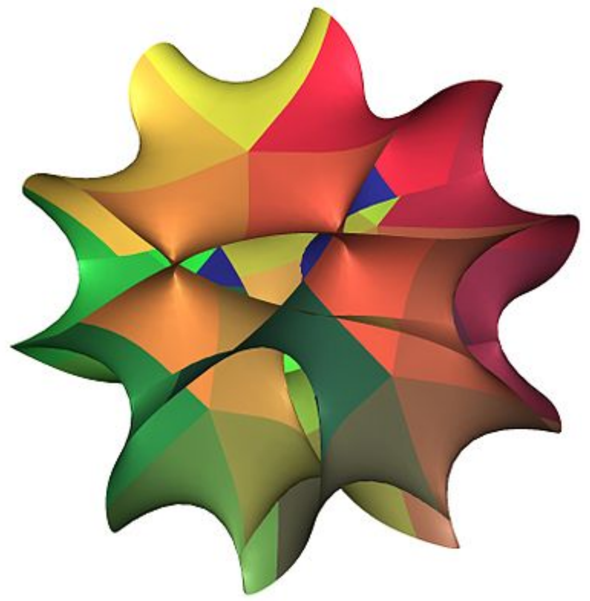
- Couplings require information about geometry (not just topology)
- This involves finding the metric on the Calabi-Yau and other field profiles
- Have recently managed to do this with these line bundle sums!
- First calculation of this kind - made possible by ML!

[Submitted on 2 Feb 2024 (v1), last revised 2 Jul 2024 (this version, v2)]

Computation of Quark Masses from String Theory

Andrei Constantin, Kit Fraser–Taliente, Thomas R. Harvey, Andre Lukas, Burt Ovrut

(+ upcoming long paper)



Why do you need geometry now?

- Fields are not canonically normalised in a string compactifications
- $\mathscr{L} = -K_{i\bar{j}}\bar{\psi}^j\gamma_\mu D^\mu\psi - K_{i\bar{j}}D_\mu\bar{\phi}^j D^\mu\phi - (\lambda_{ijk}\phi^i\psi^j\psi^k + h.c.) + \ldots$

Field space metric

Holomorphic Yukawa Couplings

$$K_{IJ} \sim \int_{CY} \nu_I \wedge \star_V \nu_J$$

$$\lambda_{IJ} \sim \int_{CY} \nu_I \wedge \nu_J \wedge \nu_K$$

Hodge Star - Depends on the Metric!

Most discussions on flavour physics in string theory have just assumed K is euclidean

What is v?

$$K_{IJ} \sim \int_{CY} \nu_I \wedge \star_V \nu_J$$

$$\lambda_{IJ} \sim \int_{CY} \nu_I \wedge \nu_J \wedge \nu_K$$

$$\triangle_{10}\, \phi(x,y) = (\, \triangle_4 + \triangle_6 \,)\phi(x,y)$$

$$\phi(x,y) = \sum_I \varphi_I(x)\nu_I(y), \quad \triangle_6\, \nu_I(y) = m_I^2 \nu_I(y)$$

$$\boxed{\triangle_4\, \phi_n(x) + m_n^2 \phi_n(x) = 0}$$

We only want the zero modes

The derivative is the gauge and gravity covariant derivative

Geometry: Metric → Gauge Field → Harmonic forms

Eleven neural networks in total!

The String Model

This model has the **MSSM Particle Content + Uncharged Moduli**

No extra vector-like pairs or chiral exotica

Hypersurface in A = $\mathbb{P}_1 \times \mathbb{P}_1 \times \mathbb{P}_1 \times \mathbb{P}_1 \sim S^2 \times S^2 \times S^2 \times S^2$

SU(5) Like structure but no GUT phase

$$p = \sum_{\text{even}} x_\alpha^2 y_\beta^2 u_\gamma^2 v_\delta^2 + \psi_0 \sum_{\text{odd}} x_\alpha^2 y_\beta^2 u_\gamma^2 v_\delta^2 + \psi\, x_0 x_1 y_0 y_1 u_0 u_1 v_0 v_1$$

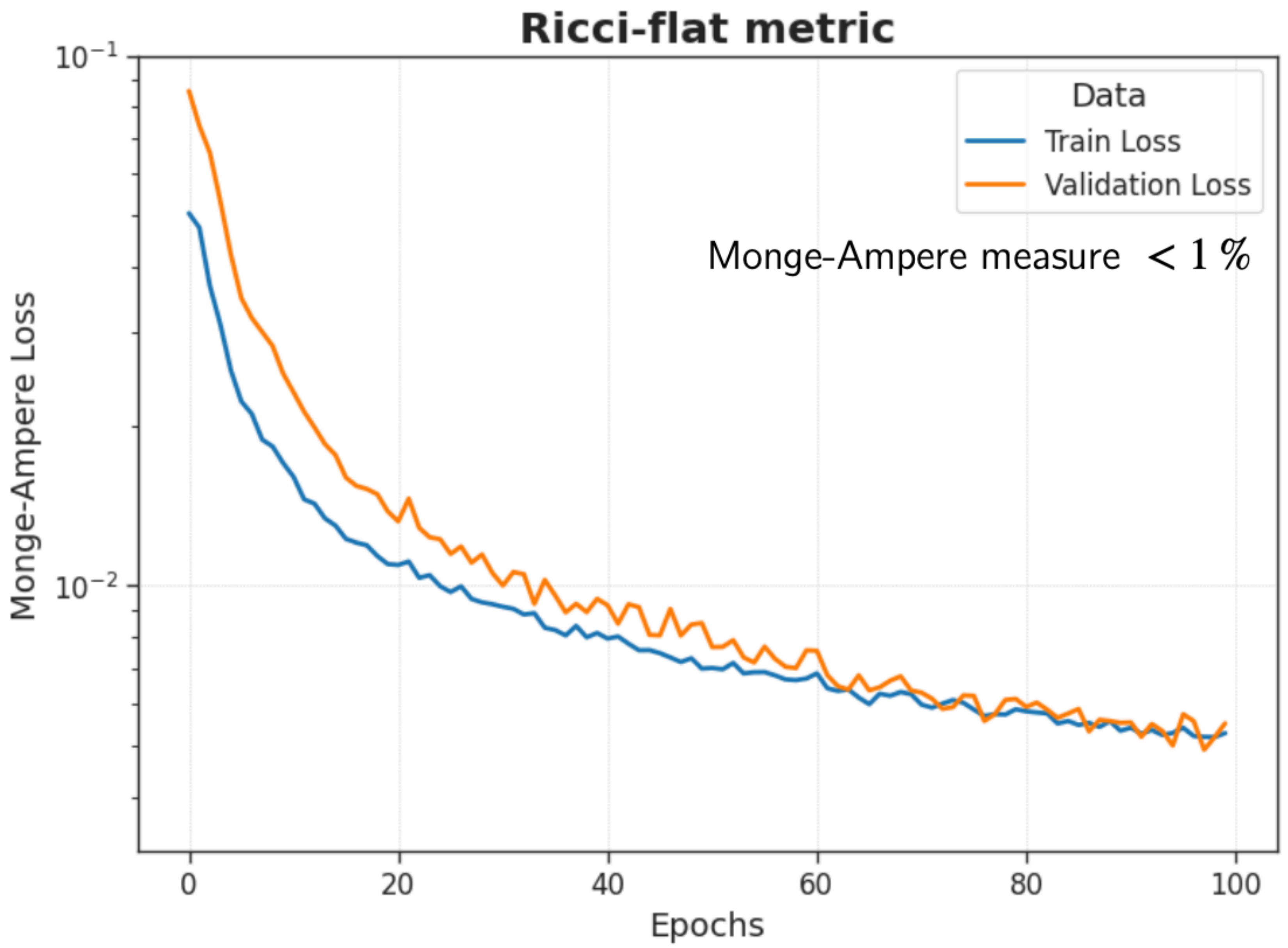
$$2\begin{pmatrix} Q_2 \\ U_2 \\ E_2 \end{pmatrix},\ \begin{pmatrix} Q_5 \\ U_5 \\ E_5 \end{pmatrix},\ \begin{pmatrix} D_{2,4} \\ L_{2,4} \end{pmatrix},\ 2\begin{pmatrix} D_{4,5} \\ L_{4,5} \end{pmatrix},\ H^d_{2,5},\ H^u_{2,5}\ .$$

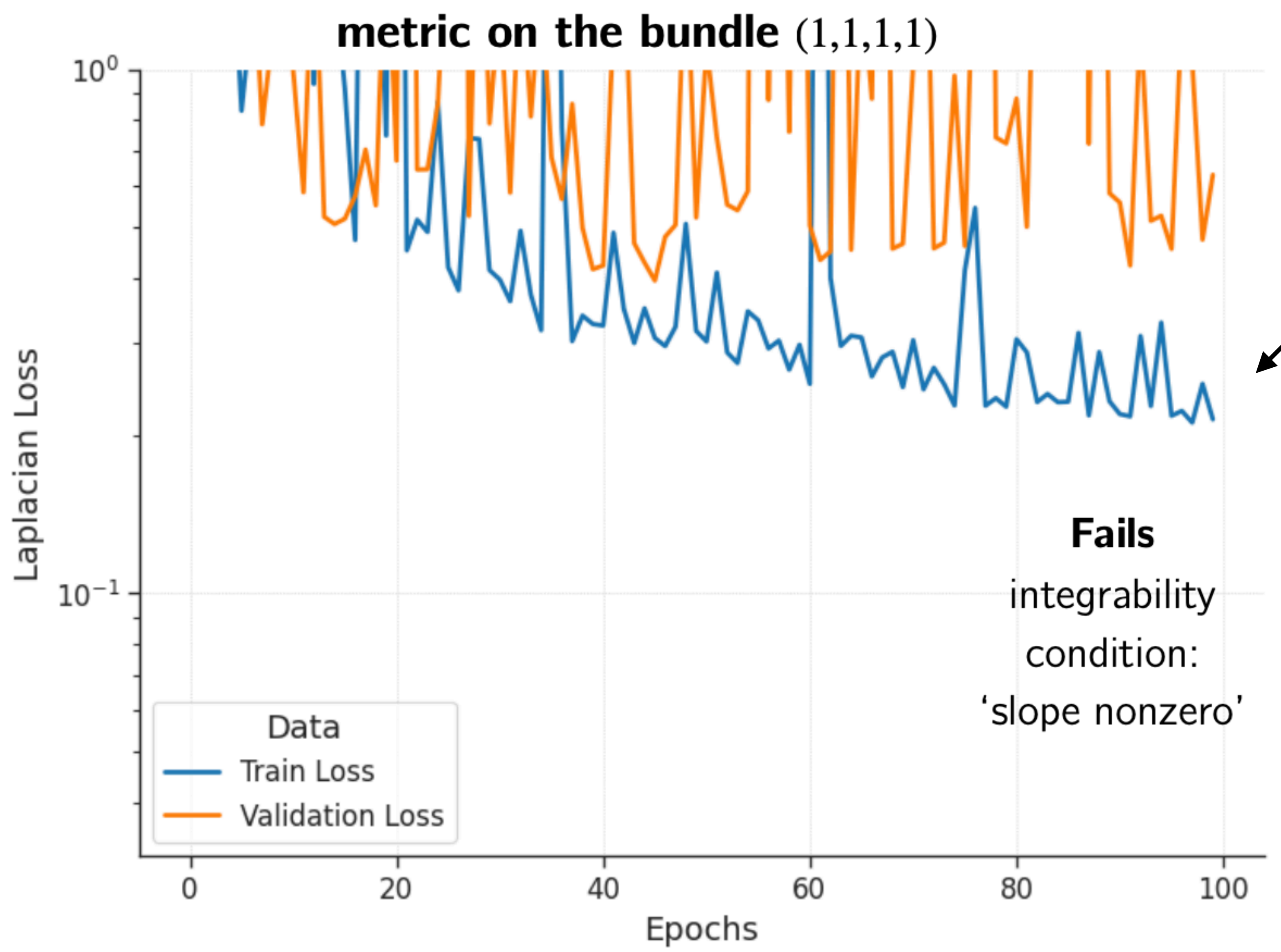
$$V = \mathcal{O}_X \begin{array}{c} \begin{array}{ccccc} L_1 & L_2 & L_3 & L_4 & L_5 \end{array} \\ \begin{pmatrix} -1 & -1 & 0 & 1 & 1 \\ 0 & -3 & 1 & 1 & 1 \\ 0 & 2 & -1 & -1 & 0 \\ 1 & 2 & 0 & -1 & -2 \end{pmatrix} \end{array}$$

$$Y_u \sim \begin{bmatrix} 0 & 0 & * \\ 0 & 0 & * \\ * & * & 0 \end{bmatrix} \qquad Y_d \sim \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

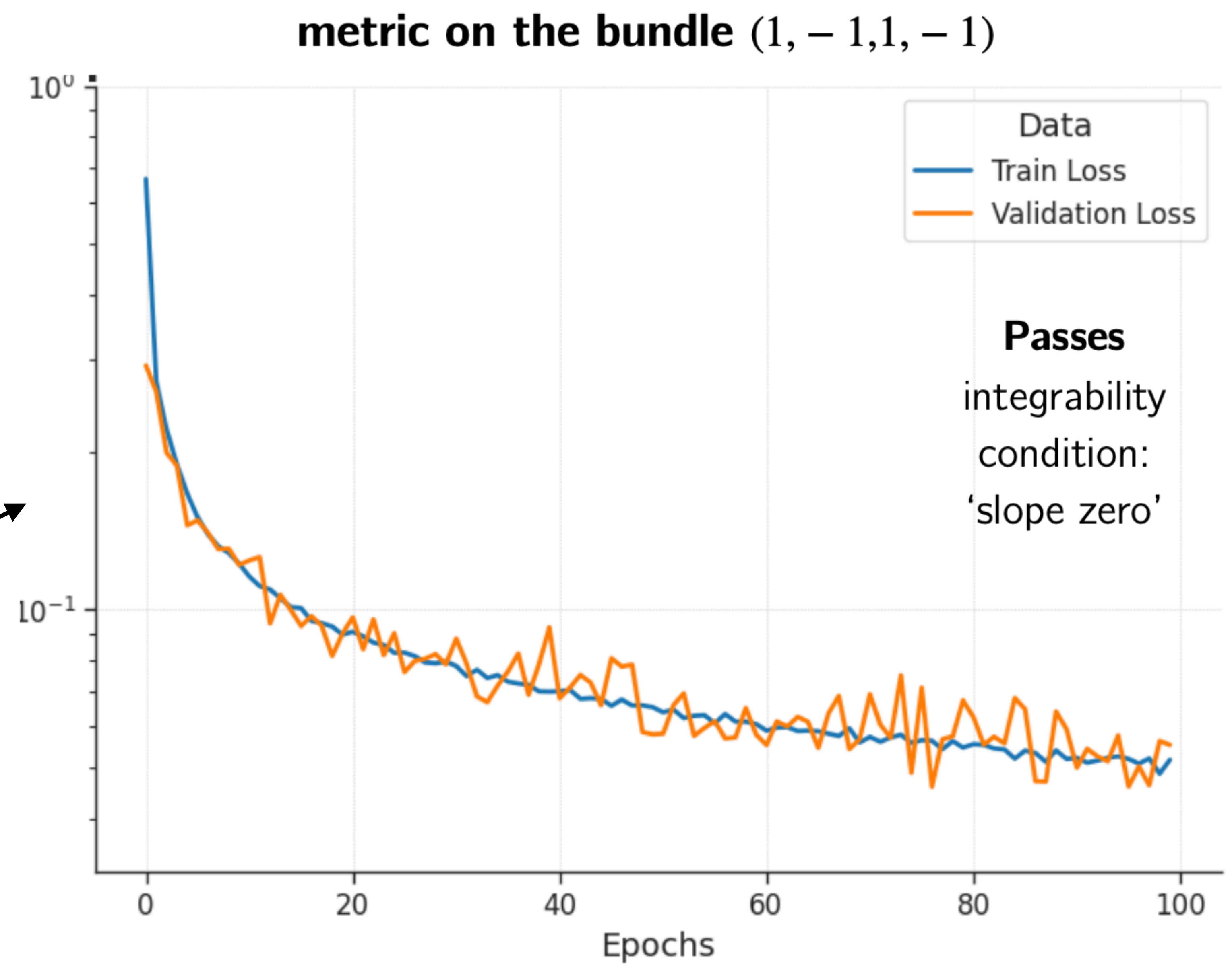
First Calculation of its kind! Proof of concept - do not expect realistic results

Training

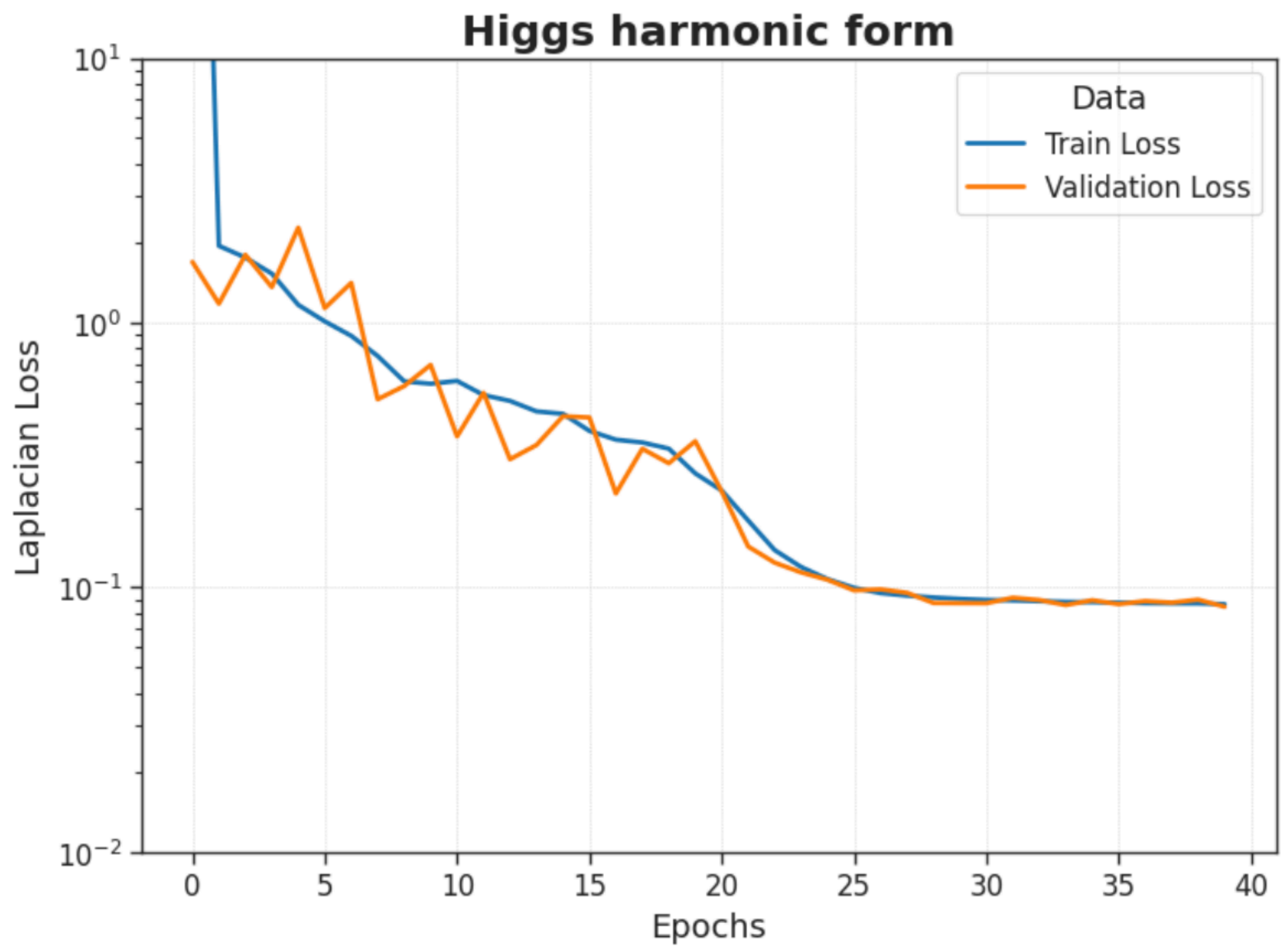




Topological obstruction to solving EOM



Topologically allowed to solve EOM



All satisfy the EOM to within a few percent

11 networks trained in total

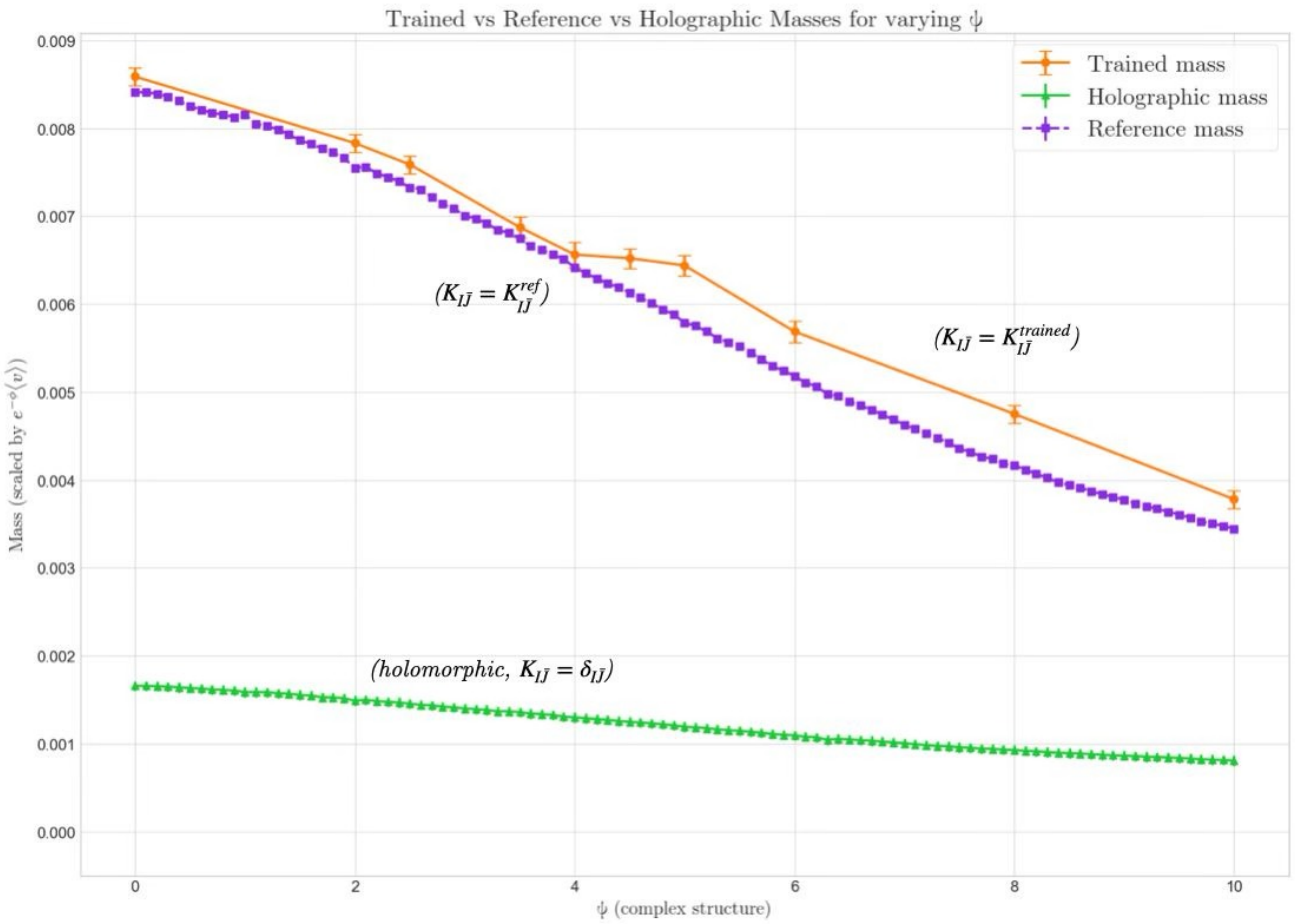
Results

From model structure, will always have one massless quark

From choice of one-parameter family of moduli two remaining masses are equal

We checked that this degeneracy is lifted for other choices of moduli

Statistically ~1% error



Conclusion

- Numerical methods, enabled by ML/AI, has enabled the calculation for previously inaccessible quantities in string compactifications
- Better analysis of EFT control - more refined than large volume approximation
  - These are important considerations for moduli stabilisation
- First, from first principles, calculation of quark masses from string theory!