Machine Learning and String Theory

Andre Lukas
University of Oxford

MLNosh, Oxford, 2 December 2024

collaborators:
Steve Abel, Lara Anderson, Andreas Braun, Evgeny Buchbinder, Per Berglund, Callum Brodie, Andrei Constantin, James Gray, Thomas Harvey, Yang-Hui He, Elli Heyes, Edward Hirst, Vishnu Jejjala, Magdalena Larfors, Seung-Joo Lee, Challenger Mishra, Luca Nutricati, Burt Ovrut, Eran Palti, Fabian Ruehle, Robin Schneider, Sebastian von Hausegger

String theory and ML started about 7 years ago in Oxford TP . . .

Fabian Ruehle

Sven Krippendorf

Yang-Hui He

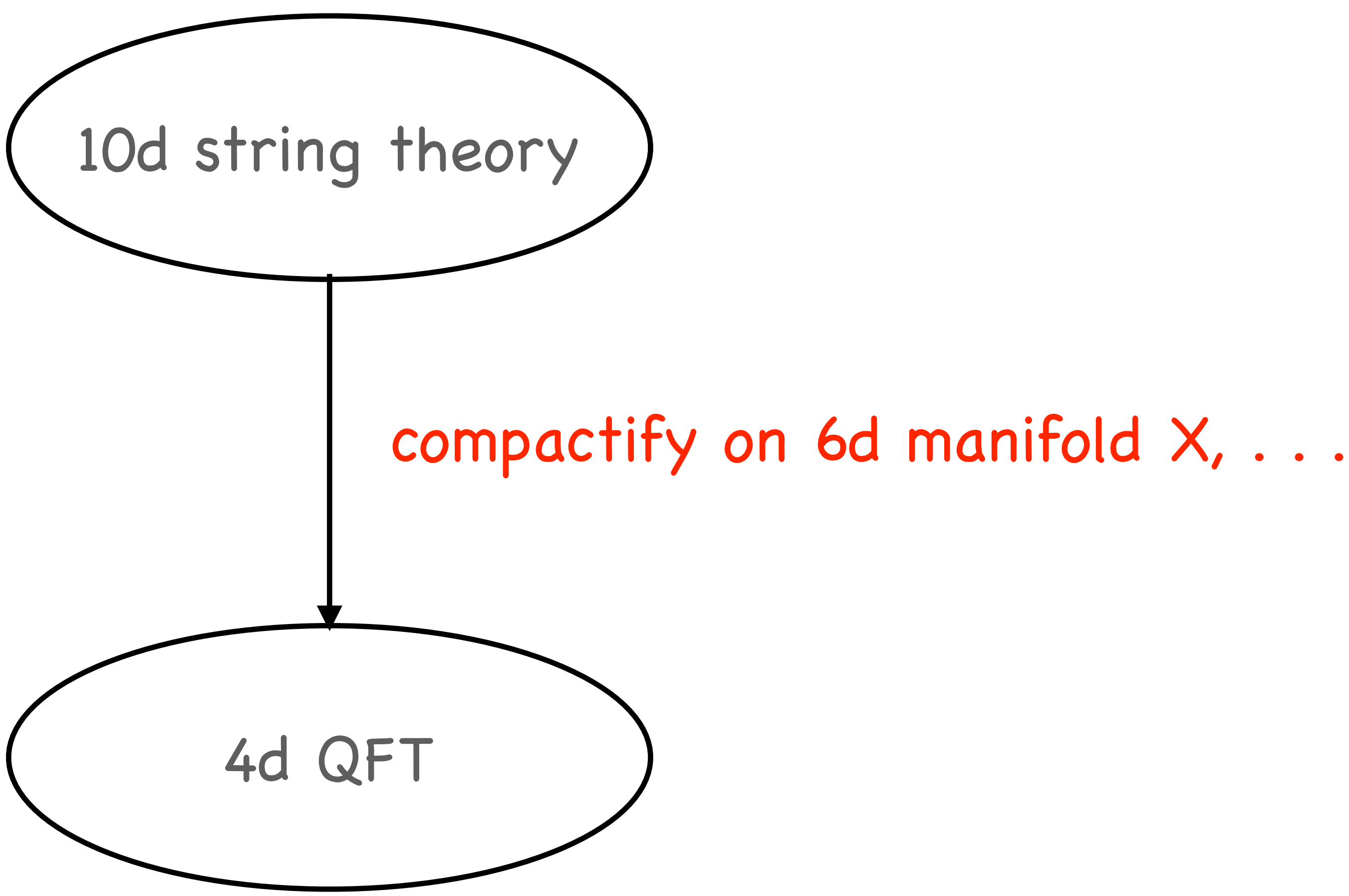
A burst of activity since . . .

Fabian Ruehle: ``Data Science Applications to String Theory'', Phys. Rept. 839 (2020) 1-117.

Why might ML techniques be useful in string theory?

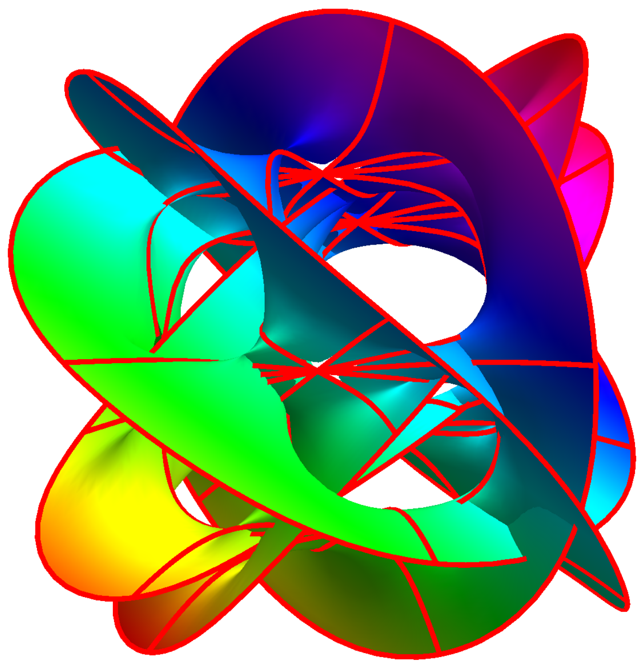
Some basic features of string theory

- String theory is a consistent theory which contains gauge theories and (quantum) gravity.
- But it is defined in 10 space-time dimensions.
- To make contact with physics we need to compactify (``curl up”) 6 dimensions.



But we need to satisfy the (10d) Einstein equations, so X needs to carry a metric with vanishing Ricci tensor.

Yau's theorem: ``Ricci-flat metrics exist (and are unique under certain extra conditions) on Calabi-Yau (CY) manifolds."

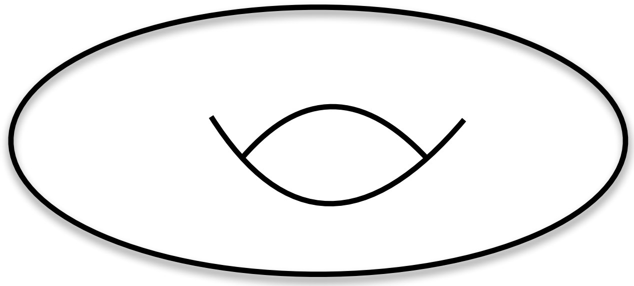
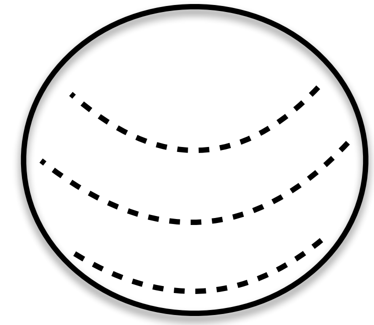


= bi-cubic

There is a huge number ($10^{100\,n}$, $10^{1000\,n}, \ldots$) of possibilities for X, all leading to 4d theories. Only a small fraction leads to theories close to the correct one....

The 10d theory is (basically) unique, but the 4d theory depends on X.

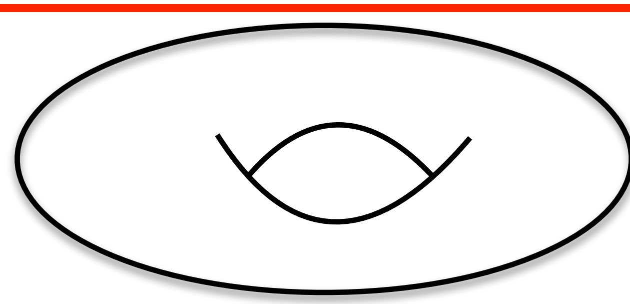
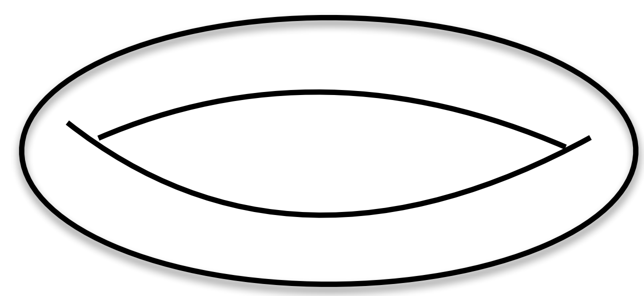
How does the 4d theory depend on X?

topology :  or  ?

-> determines structure of 4d theory: forces, matter content, . . .
(Maths: Algebraic Geometry)

Many compactification of string theory known which lead to the forces and particle content of the standard model (SM) of particle physics!

main focus so far -> example later

shape :  or  ?

-> determines couplings/particle masses in 4d theory
(Maths: Differential Geometry)

Can string theory also explain the couplings and masses in the SM?
``Can string theory explain the electron mass?"

to tackle this we need the Ricci-flat metric g on X <-> shape

Coming back to: Why might ML methods be useful in string theory?

- String theory contains large (mathematical) data sets, with entries typically of the type "geometrical object -> topological property" -> supervised learning

main approach initially

- The huge ``landscape" of string theory leads to large search problems, e.g. for realistic models -> heuristic search methods such as reinforcement learning and genetic algorithms

- The calculation of properties for any given compactification can be hugely challenging.
  - difficult algebraic computations -> supervised learning ???
  - difficult differential computations -> solving diff. eqs. with ML

Example 1: Supervised learning of line bundle cohomology

Line bundles $L \to X$ over CY manifolds can be labelled by integer vector $L = \mathcal{O}_X(k)$ and their cohomology dimensions $h^q(X, \mathcal{O}_X(k))$ are hard to compute and of interest in string theory.

training set: $\{(k, h^q(X, \mathcal{O}_X(k))\}$

supervised learning works well, but: error unacceptable, hard to verify (Fabian Ruehle 1706.07024)

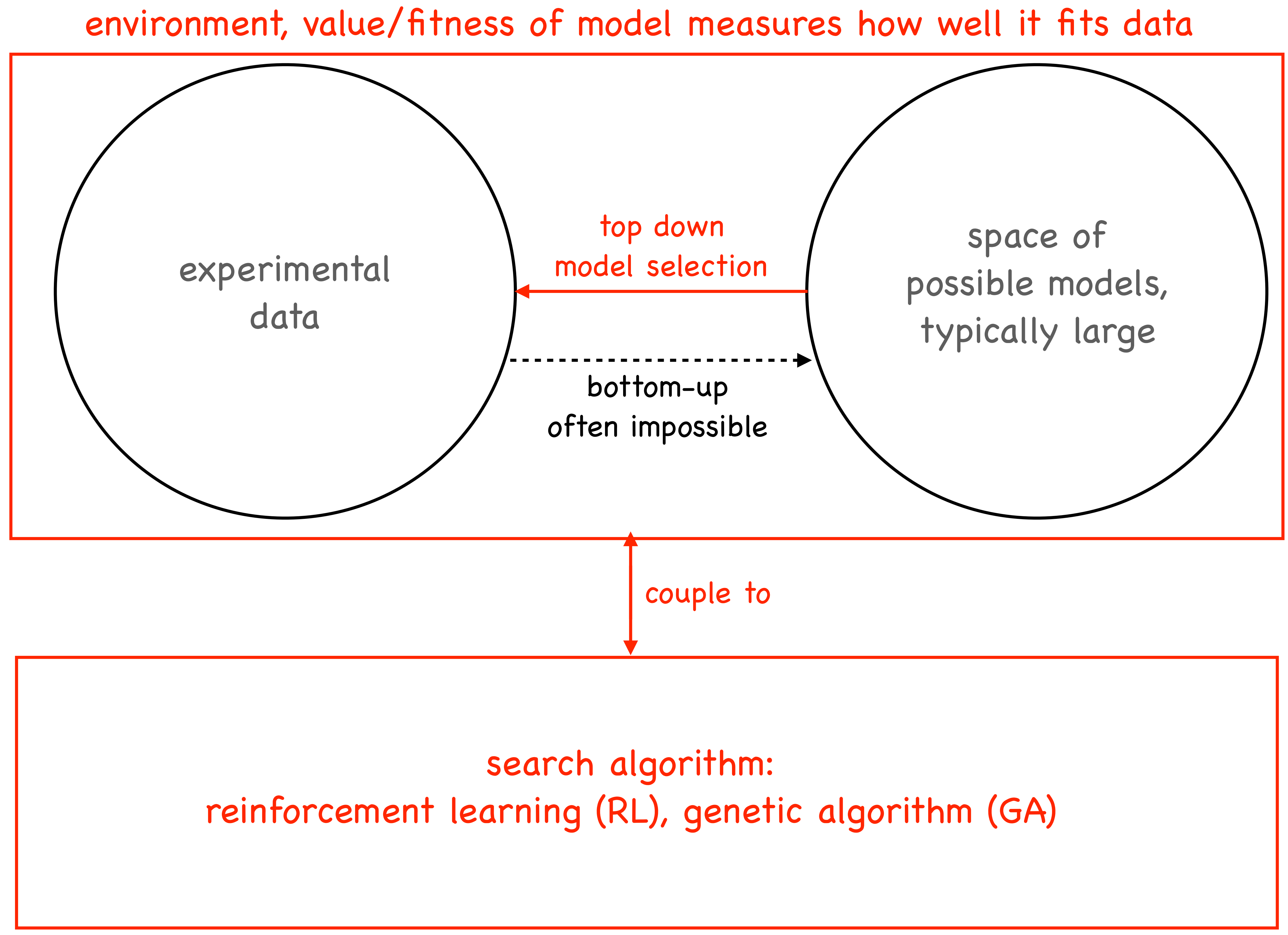
More dedicated network, ``opened up" allows for read-out of formula. This supports a conjecture:

(Constantin, Lukas, 1808.09992, Brodie, Constantin, Lukas, 2010.06597)

Conjecture: (Line) bundle cohomology dimensions on n-dim. Kahler manifolds are given by piecewise polynomial expressions with polynomial degrees less equal n.

Example 2: Model search with RL or GAs

Basic idea:



Applied to: heterotic CY models with flux (=bundles)

(A. Constantin, AL, T. Harvey, 2108.07316)

Consider bi-cubic CY X with flux = vector bundle V defined by

$$0 \to V \to B \xrightarrow{f} C \to 0 \qquad V \cong \mathrm{Ker}(f)$$

"monad"

$$B = \bigoplus_{a=1}^{r_B} \mathcal{O}_X(\mathbf{b}_a) \qquad C = \bigoplus_{\alpha=1}^{r_C} \mathcal{O}_X(\mathbf{c}_\alpha) \qquad r_B - r_C = 4$$

2d integer vectors

A model is described by a $2 \times (r_B + r_C)$ integer matrix

$$\{\ (\mathbf{b}_1, \ldots, \mathbf{b}_{r_B}, \mathbf{c}_1, \ldots, \mathbf{c}_{r_C})\ \} = \text{environment}$$

The particle content of a 4d model can be computed from this matrix.
(But it's complicated!)

Goal: Find models which lead to a SM spectrum -> Diophantine eqs. in $\mathbf{b}_a, \mathbf{c}_\alpha$

size of environment: $\sim 10^{2(r_B+r_C)} = 10^{12}$ (for $r_B = 5,\ r_C = 1$)

Example RL run for bi-cubic (actor-critic): $r_B = 6,\ r_C = 2 \quad \rightarrow \quad \#\text{states} \sim 10^{16}$

Training: about 1h on a single CPU

(a) Loss vs batch number.

(b) Policy loss vs batch number.

(c) Value loss vs batch number.

(d) TD return vs batch number.

(e) Number of terminal states vs episode number.

(f) Terminal fraction vs episode number.

Figure 6: Training metrics for the bicubic monad environment with $(r_B, r_C) = (6, 2)$.

Results: O(500) candidate models -> 18 new models with SM spectrum

Example 3: Ricci-flat CY metrics from ML

(M. Larfors, AL, F. Ruehle, R. Schneider, 2211.010436, 2205.13408)

Not a single Ricci-flat metric on (compact, three-fold) CY known analytically -> numerical methods

First consider lattice methods: $(20 \text{ points/dim})^6 = 6.4 \times 10^7$ points

ML approach:

- Generate point sample (x_i), $i = 1, \ldots, N$, on CY X (self-supervised learning)
- Use fully-connected NN F_θ
- Loss function $L(\theta) = \frac{1}{N} \sum_i |\mathrm{Ricci}(g(x_i))|^2 + \cdots$
- Perform gradient descent

Training

!!!!

(3 hidden layer, width 64, GELU activation, 100000 points each, Adam optimiser)

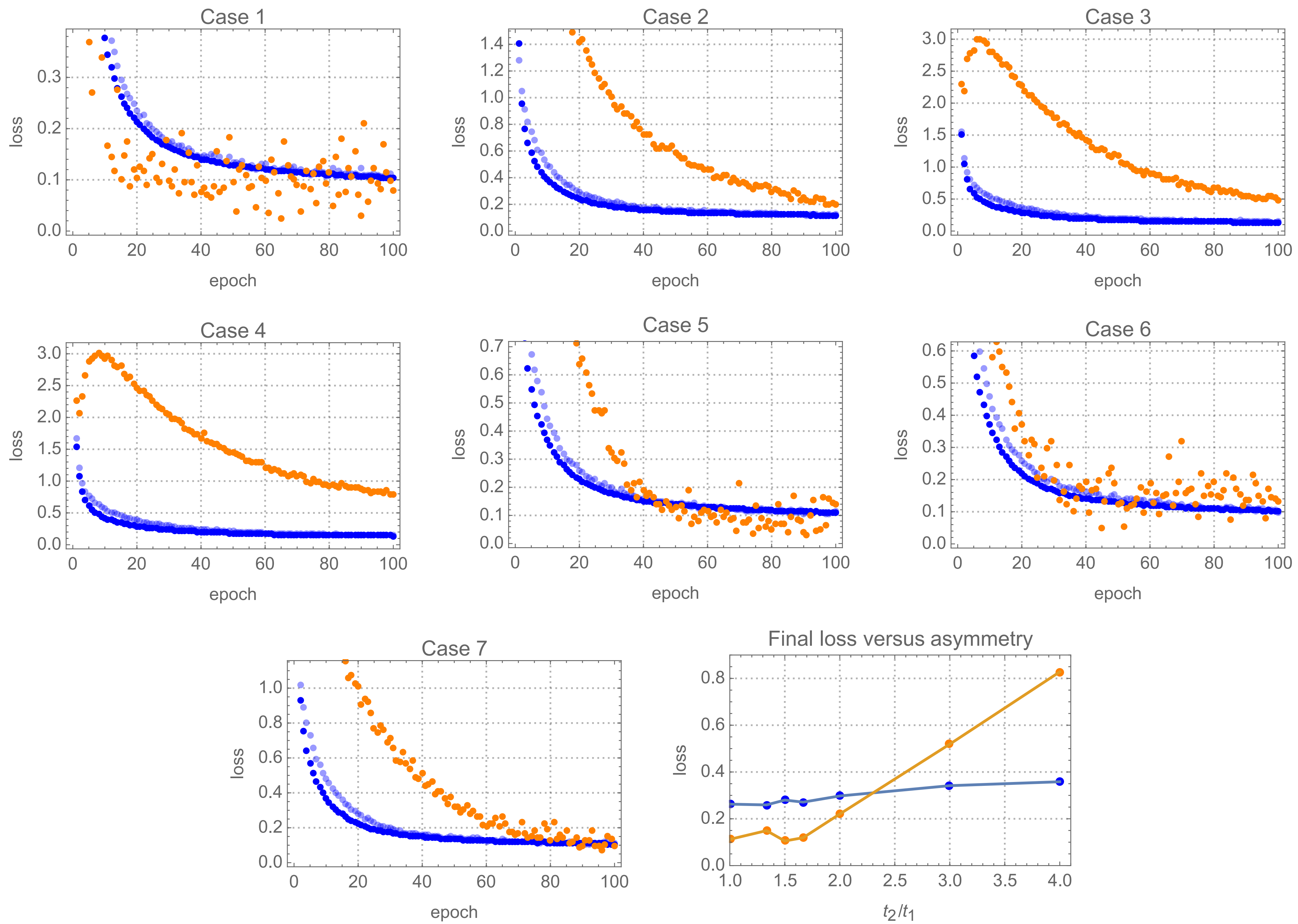


Figure 2: Bi-cubic training curves for the seven choices of Kähler parameters in Table 2. The last plot represents the final loss, obtained by averaging over the last 10 epochs, as a function of t^2/t^1 (orange: $\mathcal{L}_{\text{Kclass}}$, blue: $4 \times \mathcal{L}_{\text{MA}}$, both on training data, light-blue: $4 \times \sigma$ measure on validation data).

What can we do with the Ricci-flat metric? -> Yukawa couplings

(A. Constantin, K. Fraser-Taliente, T. Harvey, AL, B. Ovrut 2402.01615)

For example, up-quark Yukawa couplings:

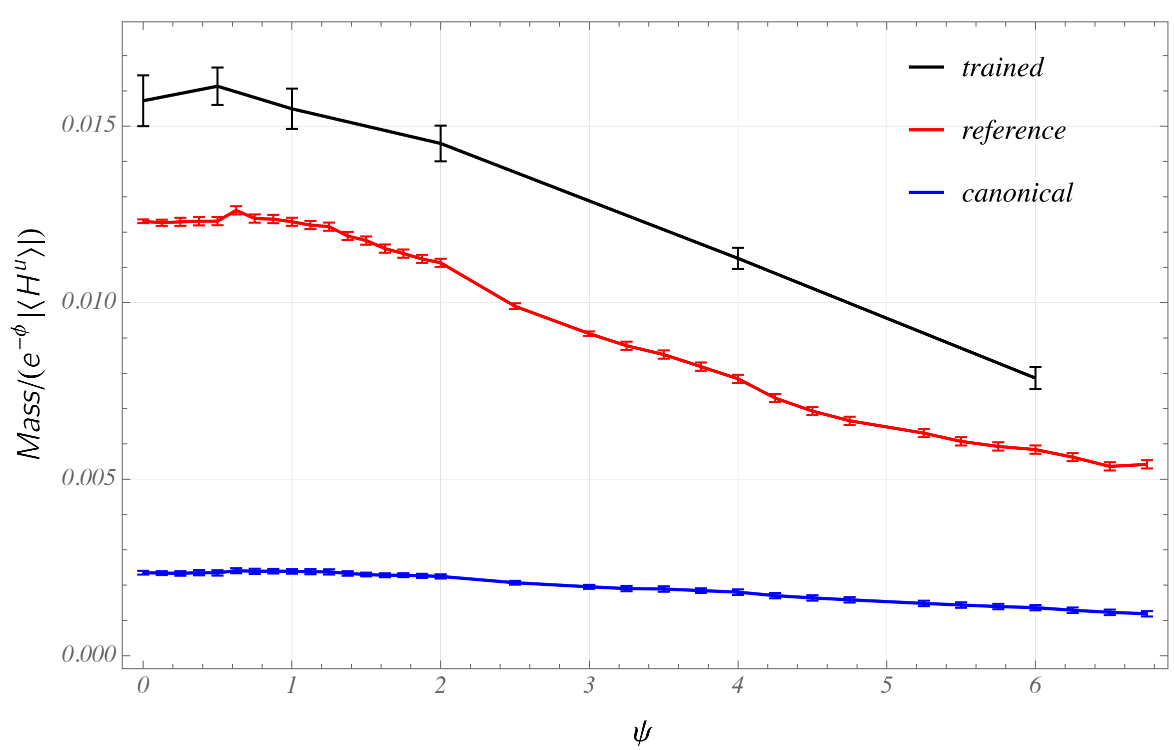
$$W_{\rm u} = Y^u_{ij} H^u Q^i U^j \qquad K = K^Q_{ij} Q^i \bar{Q}^j + K^u_{ij} U^i \bar{U}^j + k H^u \bar{H}^u$$

holomorphic Yukawa couplings, quasi-topological

field space metric (``wave fct. normalisation"), calculation requires Ricci-flat metric etc.

physical Yukawa couplings

First full computation for a quasi-realistic string model, using ML:



Conclusion

- ML techniques can be useful in string theory:
  - supervised learning of math. data sets -> conjecture generating
  - heuristic model searches in string landscape -> RL, GAs, . . .
  - solving non-linear diff. eqs. om manifolds -> self-supervised

- Distant dream: Data science techniques will allow us to explore the entire string landscape.

- Any implications for machine learning?

Thanks!