Machine Learning, Calabi-Yau Metrics and Beyond 2

James Gray

(Virginia Tech)

Anything related to my work is based on arXiv:2012.04656 with Lara Anderson, Mathis Gerdes, Sven Krippendorf, Nikhil Raghuram and Fabian Ruehle

See also arXiv:2012.04797 Michael Douglas, Subramanian Lakshminarasimhan and Yidi Qi

VirginiaTech
Invent the Future

NSF

Last time:

- We looked at a simple set of Calabi-Yau d-folds. For today let's specialize to threefolds. For example:

The quintic:

The zero set of a degree five polynomial…

$$X = \left[\mathbb{P}^4 \middle| 5\right]$$

… in the homogeneous coordinates of $\mathbb{P}^4$

Often in what follows we will consider a defining relation of the following form:

A metric ansatz

- One can make the following ansatz for a Kahler potential on the open set intersected with the Calabi-Yau.

- Here, h is a Hermitian matrix of tunable parameters.
- The are a basis of sections of
  . $\mathbb{P}^4$
- In other words, the are polynomials in the coordinates of of degree where redundancies involving polynomials that restrict to

Functional minimization

- Define a functional

  where:

  and

  This functional is minimized when

  i.e. when J is the Ricci flat Kahler form.

Using ML 1: Replace the minimization procedure

- Keep the metric ansatz.
- Generate neural networks to find the parameters h that are trained directly using a loss such as:

$$\mathcal{L}_{\mathrm{MA}} = \left| 1 + \frac{4}{3} i \frac{J^3}{\Omega \wedge \overline{\Omega}} \right|^2$$

- Essentially, we are simply following the approach of Headrick and Nassar, but we are using modern implementations of the ADAM algorithm to do the minimization!

Network giving h :

- Input: $|\psi|$ and $\arg(\psi)$
- Output: Cholesky decomposition of at (42025 parameters)

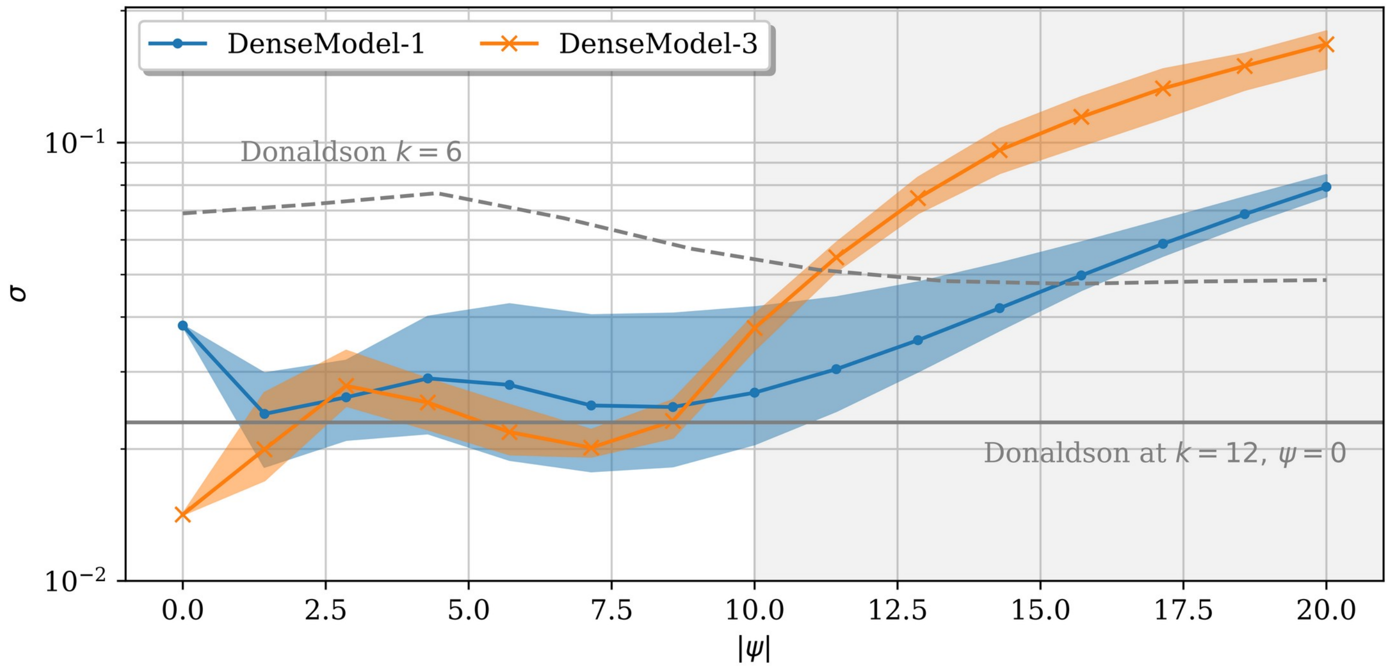
  Note that we do not need to impose symmetries!

  The Cholesky decomposition is used in the output as it makes it easy to ensure h is positive definite. It also seems to lead to slightly better results than utilizing the real and imaginary parts of h directly.

- Architecture is of a very simple form: a feedforward network with two hidden layers of dimension 205 each with sigmoid activation function.

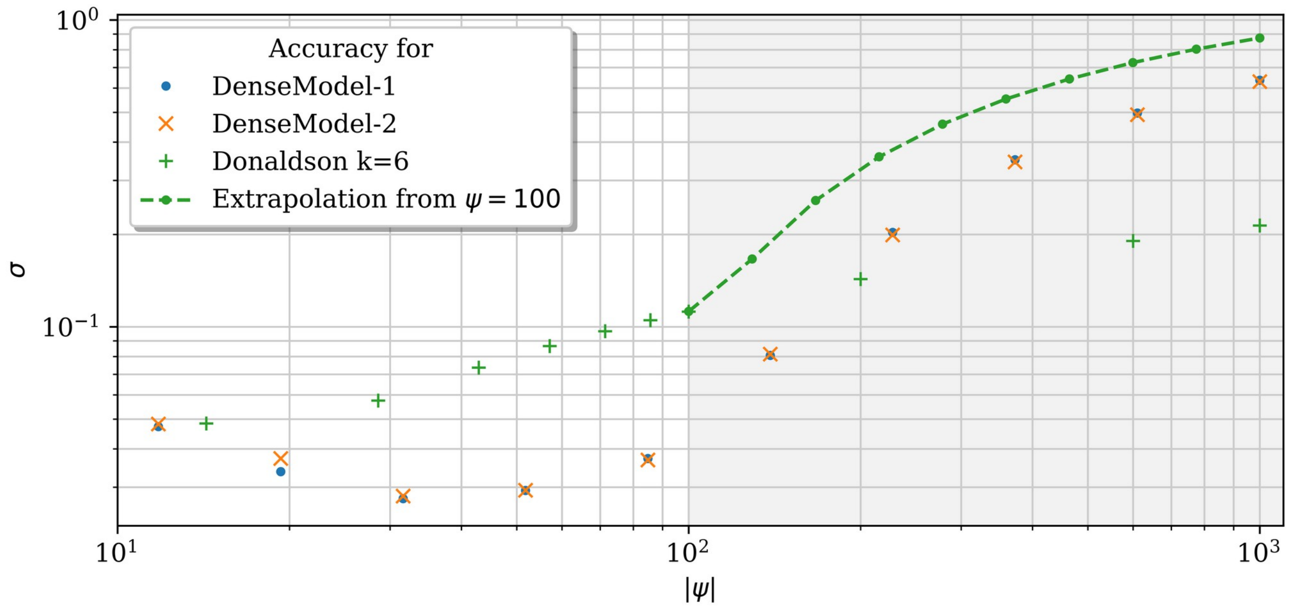
Results:

- These networks were optimized for $0 \leq |\psi| \leq 10$



- Note Donaldson with $\psi = 0$ takes order days to run even for the single case .
- This network at takes only minutes ψ and gives comparable accuracy for a whole range of

- Here are some networks trained over a larger range:



- We see again that we do better than Donaldson at $k = 6$ and that this improvement extends up to $|\psi| \simeq 175$, nearly a factor of 2 beyond the regime used during training.
- Results are strongly dependent on architecture (including extrapolation) – optimization left for future

Using ML 2: Replace metric ansatz

- Instead of learning parameters in an ansatz for the Kahler potential we can try to learn the CY metric directly.

- Neural networks are universal function approximators: they can approximate any function on a compact domain in

  (see for example:
  Cybenko, Math. Cont. Sig. Syst. 1 (1989) 4, 303-314)

- So, if we have a neural network output the metric components that could in principle work!

- Let's keep the replacement of the minimization by ADAM too to deal with the number of parameters…

Why try?:

- Perhaps we can improve performance by not being tied to an ansatz at fixed n.
- We will be able to generalize this approach to non-Kahler geometries.

One Disadvantage:

- We now need losses to check that the metric is globally well defined and Kahler (in the Calabi-Yau case)!

- We did a few different experiments with this – I will just describe one.

Loss functions:

- We use

$$\mathcal{L} = \lambda_1 \mathcal{L}_{\mathrm{MA}} + \lambda_2 \mathcal{L}_{dJ} + \lambda_3 \mathcal{L}_{\mathrm{overlap}}$$

- Here $\mathcal{L}_{\mathrm{MA}}$ is the loss we had before and

$$\mathcal{L}_{dJ} = \frac{1}{2} ||dJ||_1$$

$$\mathcal{L}_{\mathrm{overlap}} = \frac{1}{d} \sum_{k,j} \left|\left| g^{(k)}(\vec{z}) - T_{jk}(\vec{z}) \cdot g^{(j)}(\vec{z}) \cdot T^{\dagger}_{jk}(\bar{\vec{z}}) \right|\right|_1$$

- Are the losses that will be used to enforce that the two form is closed and global consistency.

Boosts

- We tried additive boosts:

$$g_{\mathrm{CY}} = g_{\mathrm{FS}} + g_{\mathrm{NN}}$$

- And multiplicative boosts:

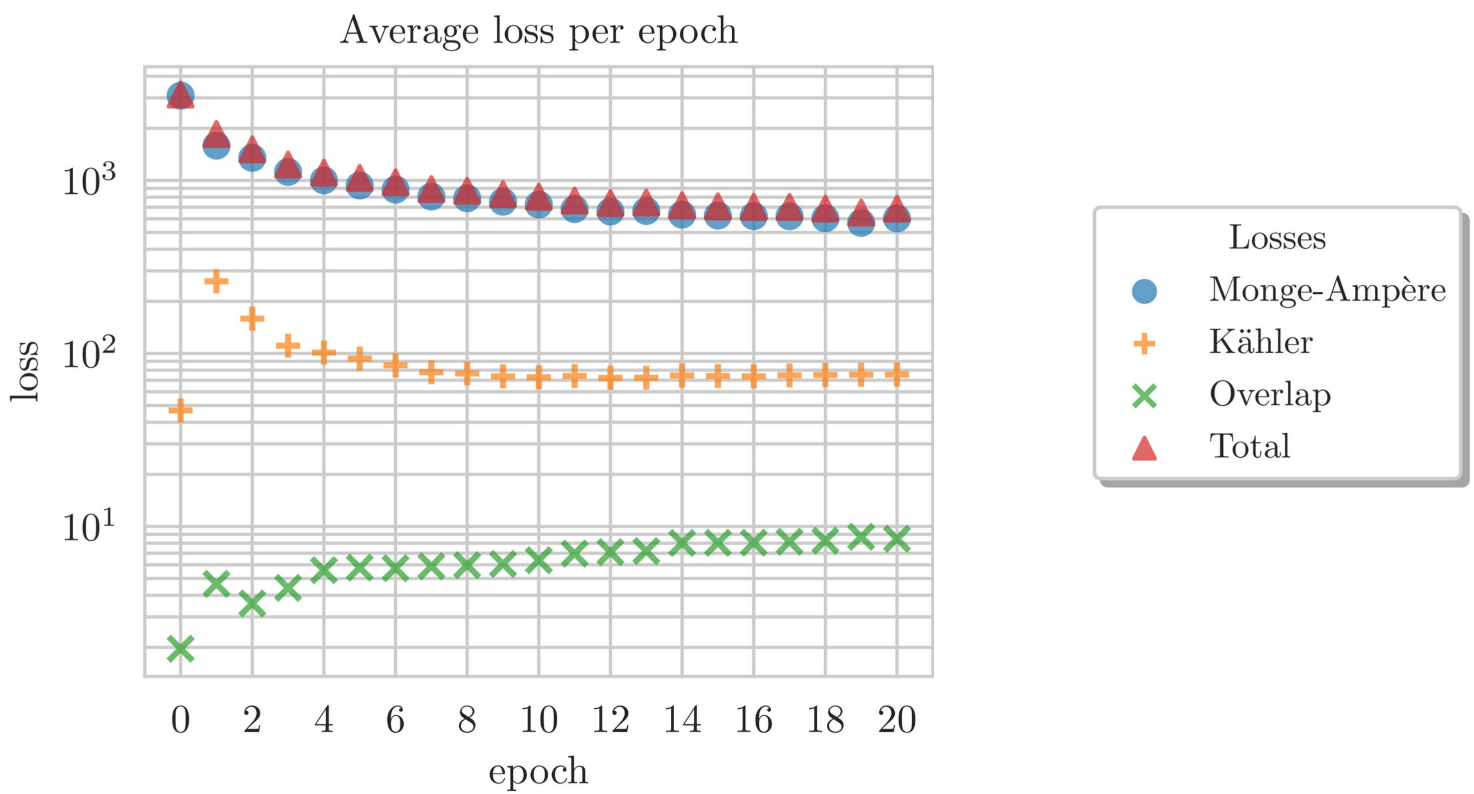
$$g_{CY} = g_{\mathrm{FS}}(1 + g_{\mathrm{NN}})$$

We have found that multiplicative boosting works best (linear boosting didn't do much better than just using the FS metric and depends sensitively λ on the ‘s)

Network giving :

- Input (17 nodes):
  - Real and imaginary parts of homogeneous coordinates in describing one of our points on X (*10 degrees of freedom*)
  - Real and imaginary parts of (*2 degrees of freedom*)
  - True/False identifiers for each of 5 patches (*5 degrees of freedom*)
- Simple feed forward network, 3 hidden layers with 100 nodes each, leaky ReLU activation function.
- Output: d^2 real and imaginary parts of metric at the input point.

- To give a concrete example, lets look at a case optimized at $\psi = 10$ on a data set of 50,000 points.
- We split the points according to train:test=90:10 and we train for 20 epochs.
- For this simple experiment accuracy reaches same level as Donaldson's algorithm at n=5.



These days several packages are available!

- cymetric:

  https://github.com/pythoncymetric/cymetric

- MLgeometry:

  https://github.com/yidiq7/MLGeometry

- cyjax:

  https://github.com/ml4physics/cyjax

- cymyc:

  https://github.com/Justin-Tan/cymyc

Kahler modulus dependence

- It is possible to augment the ansatz to include Kahler moduli dependence:

  Constantin, Lukas and Nutricati 2603.12384

- On hypersurfaces in products of projective spaces they take an ansatz of the form:

  where

  and is a globally defined sum over Laplacian eigenfunctions just like the one we saw last lecture

- They directly learn the metric for a whole set of discrete values of the Kahler moduli using the method I just described.

- They then match to the ansatz just given and extrapolate the parameters that appear in by using symbolic regression to get them as a function of the Kahler moduli.

→ Result: simple analytic form approximating the Ricci flat metric as a function of the Kahler moduli.

You can go further…

- A similar analysis had already been done for a complex structure modulus:

Lee and Lukas 2506.15766

(see also Mirjanic and Mishra 2412.19778)

- Combine the two…

  Constantin, Lee, Lukas and Nutricati 2606.28487

- And you get an example of an approximate, modulus dependent, Ricci flat metric:

Beyond Kahler geometry

- Once we are using neural networks as our ansatz for the metric we are no longer tied to Kahler geometry.
  - So far looked at case where the geometry admits nowhere vanishing two and three forms,
    and such that:

Algebraic Constraints: $$J \wedge \Omega = 0 \qquad J \wedge J \wedge J = \frac{3}{4} i \Omega \wedge \overline{\Omega}$$

Differential Constraints: $$dJ = 0 \qquad d\Omega = 0$$

Metric: $$i g_{a\bar{b}} = J_{a\bar{b}}$$

  - In explicit constructions is often known analytically, while has to be found numerically.

SU(3) Structures

Chiossi and Salamon arXiv:math/0202282

- String theories often have other, equally interesting, solutions based on more general SU(3) structures.
- The geometry admits nowhere vanishing two and three forms,

  and such that:

Algebraic Constraints: $J \wedge \Omega = 0 \qquad J \wedge J \wedge J = \frac{3}{4} i \Omega \wedge \overline{\Omega}$

Differential Constraints:

$$dJ = -\frac{3}{2}\mathrm{Im}(W_1\overline{\Omega}) + W_4 \wedge J + W_3$$

$$d\Omega = W_1 J \wedge J + W_2 \wedge J + W_5 \wedge \Omega$$

The forms are no longer closed.

- The W's are called torsion classes and obey certain constraints to make the decompositions of and unique. For example:

- You can extract the torsion classes if you know and using formulae like this:

$$W_1 = -\frac{1}{6} i\Omega \lrcorner dJ = \frac{1}{12} J^2 \lrcorner d\Omega$$

$$W_4 = \frac{1}{2} J \lrcorner dJ \qquad W_5 = -\frac{1}{2} \Omega_+ \lrcorner d\Omega_+$$

This slide and the previous one are the definition of a *general* SU(3) structure…

SU(3) Structures for string theory

- Asking for a solution to a given string theory places constraints on the torsion classes of the SU(3) structure.
- For example, heterotic string theory:

$$W_1 = W_2 = 0 \qquad W_4 = \frac{1}{2} W_5 = d\phi$$

$$W_3 = \text{anything}$$

Strominger Nucl. Phys. B274 (1986) 253

Cardoso, Curio, Dall'Agata, Castellani, Lust hep-th/0211118

- Note that the Calabi-Yau structure is a special case of this:

$$W_1 = W_2 = W_3 = W_4 = W_5 = 0$$

but far from the most general.

So why do people talk about Calabi-Yau so much?

1. Methods of algebraic geometry are often more applicable in the Calabi-Yau case.

   Given how hard it is to carry out explicit differential geometric computations on a six dimensional manifold with no continuous symmetries, this is perhaps one of the most serious issues.

2. Existence has been proven (Yau) given simple topological constraints on the manifold.

3. That the equations I have described are what is wanted for a good solution to string theory rests on the assumption that all cycles in the manifold are “large” on some scale.

   This is easy to achieve in Calabi-Yau manifolds because the overall scale of the metric is a free parameter. This is not necessarily the case for more general SU(3) structures.

- More on point 3): Heterotic string theory also requires that the following Bianchi Identity be satisfied:

- In that expression, is the field strength on an
gauge bundle that is present in the construction.

- The two sides of the Bianchi Identity transform differently with the overall scale of the of the 2-form/metric. Often this leads to the overall scale being fixed.

An ansatz for SU(3) structure

- Let's look at a slightly more complex set of Calabi-Yau manifolds:

$$X = \left[\begin{array}{c|ccc} \mathbb{P}^{n_1} & q_1^1 & \dots & q_K^1 \\ \vdots & \vdots & \vdots & \vdots \\ \mathbb{P}^{n_m} & q_1^m & \dots & q_K^m \end{array} \right]$$

Ambient space is product of projective spaces

K defining equations instead of just one

- Dimension of manifold: $\displaystyle\sum_{i=1}^{m} n_i - K$

- Calabi-Yau condition: $\displaystyle\sum_{r} q_r^i = n_i + 1 \;\; \forall \;\; i$

- These manifolds all admit a Ricci-flat metric, but they also admit other SU(3) structures!
- Here is an ansatz (generalization of that appearing in Larfors, Lukas and Ruehle 1805.08499)

$$J = \sum_{i=1}^{m} a_i J_i \qquad \qquad \Omega = A_1 \Omega_0 + A_2 \overline{\Omega}_0$$

Restriction of algebraic Kahler forms on each ambient projective factor to Calabi-Yau.

Standard form for the holomorphic three form of the SU(3) holonomy structure (and its conjugate)

- The a_i are real functions and A_1 and A_2 are complex functions.

- If we define Λ_{ijk} by: $J_i \wedge J_j \wedge J_k = \frac{3}{4} i \Lambda_{ijk} \Omega_0 \wedge \overline{\Omega}_0$

Algebraic conditions

- Resulting Torsion classes:

$$
\begin{aligned}
W_1 &= 0 \\
W_2 &= -i\overline{\partial} A_1 \lrcorner \Omega_0 + i \partial A_2 \lrcorner \overline{\Omega}_0 + i \frac{\overline{\partial}(A_1 + \overline{A}_2)}{A_1 + \overline{A}_2} \lrcorner A_1 \Omega_0 - i \frac{\partial(\overline{A}_1 + A_2)}{\overline{A}_1 + A_2} \lrcorner A_2 \overline{\Omega}_0 \\
W_3 &= \sum_i (da_i - W_4) \wedge J_i \\
W_4 &= \frac{1}{2} \sum_i J_i \lrcorner (da_i \wedge J_i) \\
W_5 &= \frac{\overline{\partial}(A_1 + \overline{A}_2)}{A_1 + \overline{A}_2} + \frac{\partial(\overline{A}_1 + A_2)}{\overline{A}_1 + A_2} \ .
\end{aligned}
$$

- You can use these expressions to try and find solutions to the heterotic string theory conditions

Analytic example to try and reproduce:

Larfors, Lukas and Ruehle arXiv:1805.08499

Work on the quintic again but do not take the Calabi-Yau structure. Instead:

Restriction of algebraic Kahler forms from ambient projective space.

holomorphic three form of the Calabi-Yau structure.

where

and $\sigma = \sum_{i=0}^{4} |z_i|^2$

- This has torsion classes

$$W_1 = W_2 = W_3 = 0 \ \ , \ \ W_5 = 2W_4 = 2d(\ln a_1)$$

and thus provides a solution to heterotic string theory .

Direct learning of the SU(3) structure metric

- In principle we can employ very similar methods indeed! We would just need to:

    - Modify the differential loss.
    - Perhaps learn as well if one is not known in a given example.
    - Potentially work on different manifolds.

- For these early experiments we wanted to be careful: since there is no existence proof for the desired metric, we wanted some way of checking we were converging to a true solution:

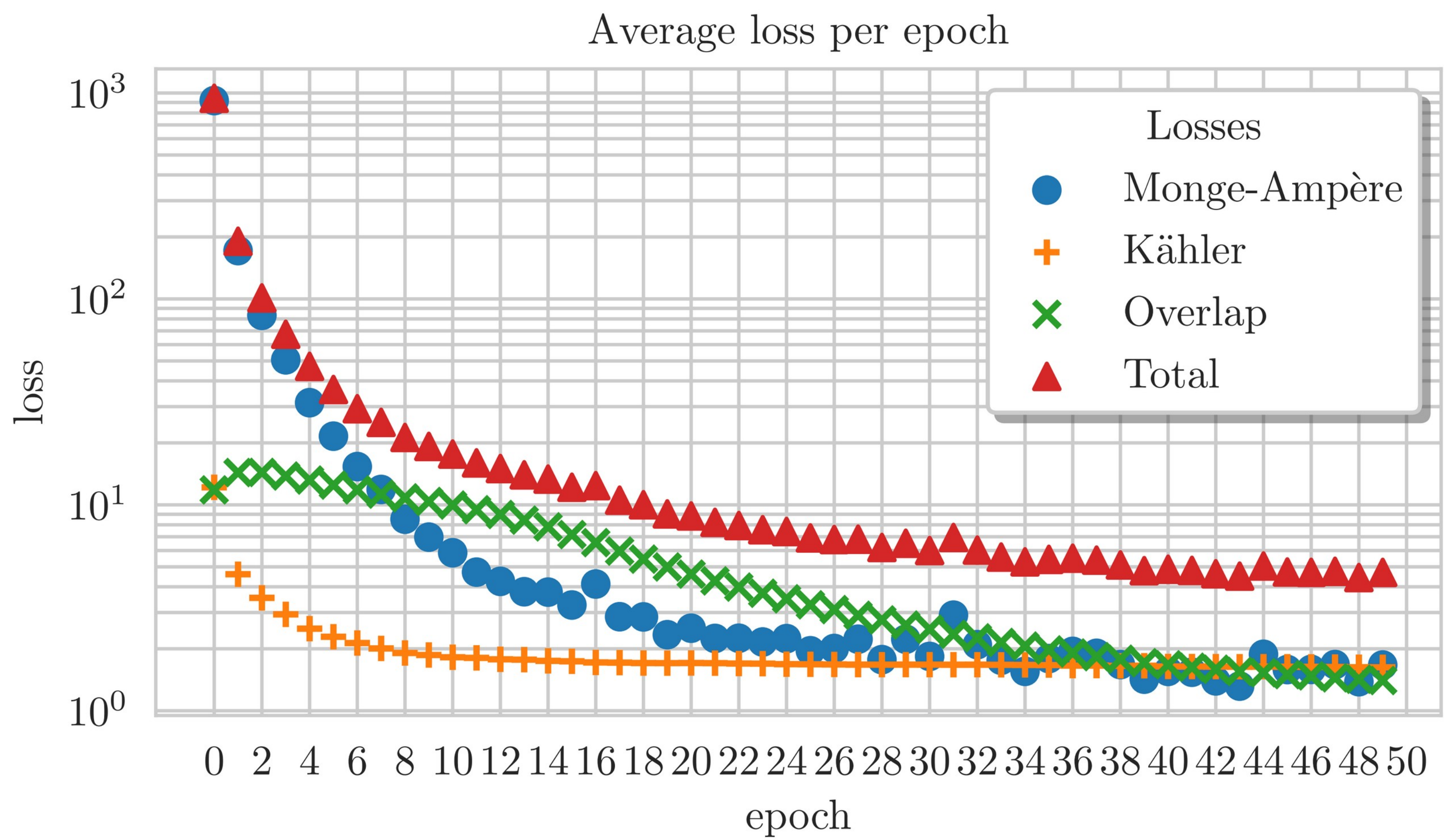
$\longrightarrow$ Try to learn known analytic solution taken from the literature.

Procedure:

- We take Ω to be known, as it was in the Calabi-Yau case.
- Take neural network that mimics our experiment for directly learning the Calabi-Yau metrics we saw earlier.
- Take same losses as in the Calabi-Yau case, with the exception of replacing the Kahler loss by:

- We can see by observing the evolution of the losses that we do indeed seem to be approaching an SU(3) structure solution…

Results:



Can we verify it is approaching the *known* solution?..

- Define an error measure that measures the difference to the known solution:

$$\mathcal{E}_{\text{known}} = ||g_{\text{numeric}} - g_{\text{known}}||_n$$

- For n=1 we can compare the output of the trained neural network to the Fubini-Study metric:

$$\mathcal{E}^{\text{FS}}_{\text{known}} = 0.511 \qquad \mathcal{E}^{\text{NN}}_{\text{known}} = 0.025$$

- The improvement is even more pronounced at higher n, showing that the neural network has fewer outlying regions far from the correct value.

What could be done going forward:

- Don’t try and hit a known case - be more flexible with target torsion classes. For example, include the following in the loss:

- Learn as well as …..

- Try and find examples that don’t have small cycles:
  - Could include losses that penalize small cycles.
  - Could include a loss that keeps the scale as a free parameter! For example:

In heterotic case:

So **naively** cases other than the Calabi-Yau could have a free overall scale of the metric.

Other topics

- There are many other areas people have been investigating:
  - Solutions to Hermitian Yang-Mills over the Ricci-flat metric.
  - Finding harmonic bundle valued forms.
  - Computing normalized Yukawa couplings
  - Computing massive spectra.
  - …

Summary

- CY state of the art:

  Numeric approximations to Ricci-flat metrics depending on a handful of complex structure and Kahler moduli.

  Approximate metrics have been used in a variety of ways.

- More general SU(3) structures:

  Subject is still basically untouched.

  Differential geometry approach using ML seems feasible.

  Currently no examples known with no small cycles.