$SU(3)$-Holonomy and $SU(3)$-Structure Metrics from Machine Learning

Lara B. Anderson

Department of Physics, Virginia Tech

Work done in collaboration with:

J. Gray, M.Gerdis, S. Krippendorf, N. Raghuram, and F. Ruehle

arXiv:2012.04656

(See also the work of Douglas, et al 2012.04797 and Jejjala et al 2012.15821)

Numerical and Geometric Methods for Ricci-flat Metrics and Flows

May 25th, 2021

Determining the 4D theory in string compactification

- Physical quantities in low energy string theory depend on the metric and gauge connections in the extra dimensions.
- For example:
  - Yukawa couplings in Heterotic string theory descend from a term in the 10-dimensional action of the form $\sim \int d^{10}\sqrt{-g}\bar{\psi}A\psi$. Normalization of fields and coefficients of the superpotential depend on g.
  - Matter field Kahler potential unknown except for special cases.
  - Modes of V-twisted Dirac Operator: $\nabla\!\!\!/_X \Psi = 0$ depend on the metric and a connection on a vector bundle, $\mathcal{V}$ on X (gauge field vevs on X).
  - Problems in moduli stabilization
- **How to determine the metric and the connection?**
- Only current viable approach via $\Rightarrow$ numeric approximation.

- One definition of a Calabi-Yau three-fold: A complex 3-fold admitting a nowhere vanishing real two-form, J, and a complex three-form, Ω,such that:

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

$$dJ = 0 \qquad d\Omega = 0$$

- Yau's theorem guarantees existence of a Ricci-flat metric associated to such a structure that is unique if the manifold is fixed. It is related to the two form as $ig_{a\bar{b}} = J_{a\bar{b}}$
- Note that on the Calabi-Yau threefolds that will appear in this talk, an explicit expression is known for Ω.

Simple algebraic descriptions of CY manifolds

- Calabi-Yau threefolds are algebraic.
- Metric deformations parameterized by $h^{1,1}(X)$ (Kähler) and $h^{2,1}(X)$ (complex structure moduli).
- Simplest examples: complete intersection manifolds in complex projective spaces
- The "Quintic" hypersurface: $X = \mathbb{P}^4[5]$
- e.g. $p(\vec{z}) = z_0^5 + z_1^5 + z_2^5 + z_3^5 + z_4^5 + \psi z_0 z_1 z_2 z_3 z_4 = 0$
- Here the holomorphic $(3,0)$ form can constructed easily (Candelas, et al). On a patch where $z_a = 1$ and considering the coordinate z_b as an implicit function of the coordinates z_c with $c \neq a, b$:

$$\Omega = \frac{1}{\partial p_\psi(\vec{z})/\partial z_b} \bigwedge_{\substack{c=0,\dots 3,\\ c\neq a,b}} dz_c$$

The idea behind most metric algorithms used to date

- For Calabi-Yau (Kähler) manifolds we always have an embedding $X \in \mathbb{P}^n$ for some n.
- Kodaira embedding: Given an ample line bundle $\mathcal{L}$ on X then an embedding

$$i_k : \; X \to \mathbb{P}^{n_k - 1}, \quad (x_0, \ldots, x_2) \mapsto \left[s_0(x) : \ldots : s_{n_k - 1}(x)\right] \tag{1}$$

exists for all $\mathcal{L}^k = \mathcal{L}^{\otimes k}$ with $k \geq k_0$ for some k_0, where $s_\alpha \in H^0(X, \mathcal{L}^k)$.

- What do we know about metrics on $\mathbb{P}^n$? Fubini-Study:

$$(g_{FS})_{i\bar{j}} = \frac{i}{2}\partial_i \bar{\partial}_{\bar{j}} K_{FS} \quad \text{where} \quad K_{FS} = \frac{1}{\pi} \ln \sum_{i\bar{j}} h^{i\bar{j}} z_i \bar{z}_{\bar{j}} \tag{2}$$

and $h^{i\bar{j}}$ is any hermitian, non-singular matrix.

- FS metric restricted to X is not Ricci-flat. But...

The starting point

- In terms of the natural embedding line bundle, we can generalize the FS Kähler potential to

$$K_{h,k} = \frac{1}{k\pi} \ln \sum_{\alpha,\bar{\beta}=0}^{n_k-1} h^{\alpha\bar{\beta}} s_\alpha \bar{s}_{\bar{\beta}} = \ln ||s||^2_{h,k} \ . \tag{3}$$

  and restrict it to X.

- Geometrically, (3) defines an hermitian metric, h, on the line bundle $\mathcal{L}^{\otimes k}$ itself (i.e. a Kähler form $\omega = \partial_i \bar{\partial}_{\bar{j}}(h) dx^i \wedge d\bar{x}^{\bar{j}}$). It provides a natural inner product on the space of global sections

$$M_{\alpha\bar{\beta}} = \langle s_\beta | s_\alpha \rangle = \frac{n_k}{Vol_{CY}(X)} \int_X \frac{s_\alpha \bar{s}_{\bar{\beta}}}{||s||^2_h} dVol_{CY} \ , \tag{4}$$

  where $\mathrm{dVol}_{\mathrm{CY}} = \Omega \wedge \bar{\Omega}$ and Ω is the holomorphic (3,0) volume form on X.

Density

- Such metrics on ample line bundles can provide a "basis" of Kähler metrics on X.

Theorem (Tian)

Let $\{s_\alpha\}$ be a basis for $H^0(X, \mathcal{L}^k)$ for some ample line bundle $\mathcal{L}$. Then the space of all "algebraic" Kähler potentials,

$$K_{h,k} = \frac{1}{k\pi} \ln \sum_{\alpha,\bar{\beta}=0}^{n_k-1} h^{\alpha\bar{\beta}} s_\alpha \bar{s}_{\bar{\beta}} \tag{5}$$

where $k \in \mathbb{Z}$, is dense in the space of Kähler potentials.

- Idea: Generalized Fubini study metrics can approximate the metric of our choice. But how to find the right "path" to the Ricci flat metric?

Balanced Metrics

- The metric h on the line bundle $\mathcal{L}$ is called "balanced" if $(M_{\alpha\bar{\beta}})^{-1} = h^{\alpha\bar{\beta}}$
- "Orthonormal" basis for sections for which $M_{\alpha\beta} = \delta_{\alpha\beta} = h_{\alpha\beta}$
- Fixed point of Donaldson's "T-operator" $\leftrightarrow$ balanced point.

  $T(h)_{\alpha\bar{\beta}} = \frac{n_k}{Vol_{CY}(X)} \int_X \frac{s_\alpha \bar{s}_{\bar{\beta}}}{\sum_{\gamma\bar{\delta}} h^{\gamma\bar{\delta}} s_\gamma \bar{s}_{\bar{\delta}}} \mathrm{dVol}_{CY}$
- Many theorems about balanced metrics. This one gives a possible path:

Theorem (Donaldson)

For each $k \geq 1$, the balanced metric, h, on $\mathcal{L}^{\otimes k}$ exists and is unique. As $k \to \infty$, the sequence of metrics

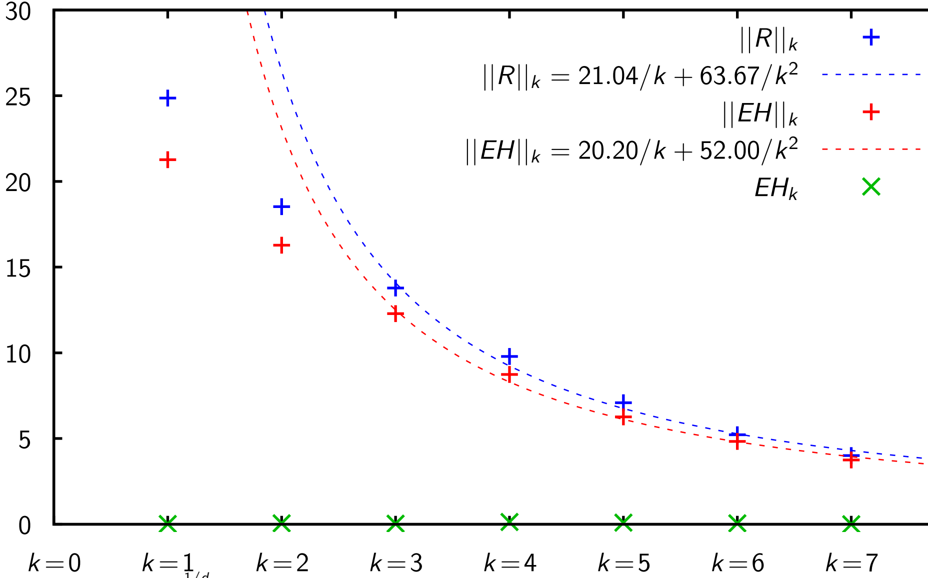
$$g_{i\bar{j}}^{(k)} = \frac{1}{k\pi} \partial_i \bar{\partial}_{\bar{j}} \ln \sum_{\alpha,\bar{\beta}=0}^{n_k-1} h^{\alpha\bar{\beta}} s_\alpha \bar{s}_{\bar{\beta}} \tag{6}$$

on X converges to the unique Ricci-flat metric for the given Kähler class and complex structure.

The metric algorithm implemented

(See work of Douglas, Karp, Reinbacher and Brelidze, Braun, Ovrut, et al)

1. Choose an ample line bundle $\mathcal{L}$ and a degree k (that is, a twisting of the line bundle $\mathcal{L}^k$) at which to compute the balanced metric which will approximate the Calabi-Yau metric.
2. Calculate a basis $\{s_\alpha\}_{\alpha=0}^{n_k-1}$ for $H^0(X, \mathcal{L}^k)$ at the chosen k.
3. Choose an initial non-singular, hermitian matrix, $h^{\gamma\bar{\delta}}$. Perform the numerical integration to compute the T-operator.
4. Set the new $h^{\alpha\bar{\beta}}$ to be $h^{\alpha\bar{\beta}} = (T_{\alpha\bar{\beta}})^{-1}$.
5. Return to item 3 and repeat until $h^{\alpha\bar{\beta}}$ approaches its fixed point. In practice, this convergence occurs in less than 10 iterations and does not depend on the initial choice of $h^{\alpha\bar{\beta}}$.
6. Increment k and repeat all steps until desired accuracy is reached.



With $||R||_k = \frac{Vol_k^{1/d}}{Vol_{CY}} \int |R_k| \, dVol_k$. Predicted fall-off: $||R||_k = \frac{a_0}{k} + \frac{a_1}{k^2} + \ldots$

- So, given the need for a certain accuracy of approximation to the Ricci-flat metric, one can choose a sufficiently high k and iterate the T-operator to convergence.
- What is a "good enough" level of accuracy? We don't really know yet – the methods are computationally intensive and no one has yet pushed the development further to the physical applications it is ultimately wanted for.
- Note: other approaches do exist in the literature (we'll touch on one briefly later)

Plan of the rest of the talk

- Supervised learning of moduli dependence of Calabi-Yau metrics using the Donaldson algorithm to generate training data.
- Direct learning of moduli dependent Calabi-Yau metrics both using the metric ansatz and without it.
- Direct learning of metrics associated to SU(3) structures with torsion.

Learning Moduli Dependence

- For most applications need to know the metric as a function of the moduli.
- The problem with using the Donaldson algorithm for this is that the moduli enter in a rather subtle way (through the choice of points on the Calabi-Yau in discretizing the integral in the T-operator). In addition the algorithm is computationally costly (so running it to high accuracy at many different values of the moduli is prohibitively slow).
- Idea: learn the moduli dependence using Supervised Learning with training data produced by the Donaldson algorithm.

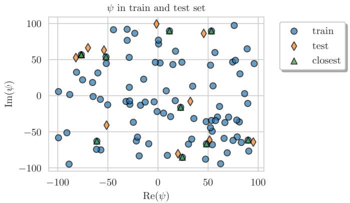
Procedure

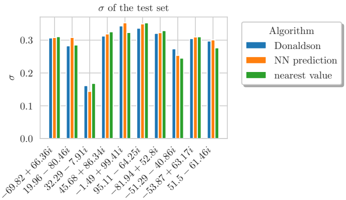
- Compute h for different choices of complex structure using Donaldson's algorithm as training data. Used a sample of $80,000$ points over the Calabi-Yau, at $k = 3$ and demanded iteration to converge to the 10^{-6} level.
- Input: $Re(\psi)$, $Im(\psi)$, $|\psi|$
- Output: $Re(h^{\alpha\bar{b}})$, $Im(h^{\alpha\bar{b}})$
- Not very sensitive to hyperparameter tuning, does not require complicated network architectures. Feed forward NN:

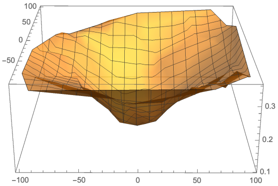
Layer	Number of Nodes	Activation	Number of Parameters
input	3	–	–
hidden 1	100	leaky ReLU	400
hidden 2	1000	leaky ReLU	101 000
hidden 3	1000	leaky ReLU	1 001 000
output	N_k^2	identity	$1000 \times N_k^2 + N_k^2$

Results

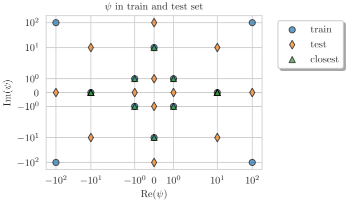
- Error Measure: $\sigma = \frac{1}{\int_X \Omega\wedge\overline{\Omega}} \int_X \left|1 + \frac{4i}{3} \frac{J^3}{\Omega\wedge\overline{\Omega}}\right|$

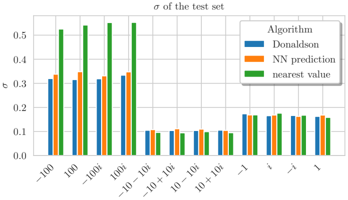


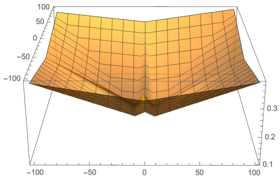




- Key point is that this interpolation continues to work even with a sparser training set







- This is important because running the Donaldson algorithm is so costly, especially if we wished to obtain results to higher accuracy (or if we wanted to work in cases with less symmetry...).

- There is no in principle problem with extending such methods to other Calabi-Yau threefolds and multiple moduli.
- The downside of what we just did is that we need to run Donaldson to obtain the training data. This ties us to the accuracy and computational cost of that algorithm, even if we can make do with a sparser set of training data than one may have thought...
- The obvious thing to try is to learn the moduli dependent metric directly, without using Donaldson's algorithm as a crutch.

Direct Learning of the Kahler Potential

- The balanced metric output by Donaldson's algorithm at given finite k is not necessarily the most accurate approximation to the Ricci-flat metric - maybe we can do better?
- So instead of using Donaldson algorithm data, we instead generate networks to find the parameters that are trained directly using a loss such as:

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

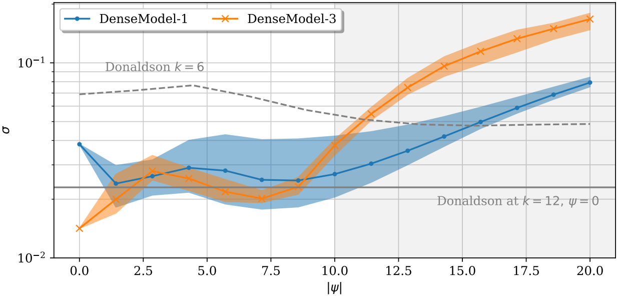
- C.f.: Headrick and Nassar (0908.2635), (although note that we are obtaining moduli dependent results and using ML techniques, rather than using a minimize function in Mathematica).

Procedure

- Input: $|\psi|$ and $arg(\psi)$
- Output: Cholesky decomposition of $h^{\alpha\bar{\beta}}$ at $k = 6$ ($42,025$ parameters)
- 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 similar types to that seen earlier (sigmoid activation function used this time)

Results

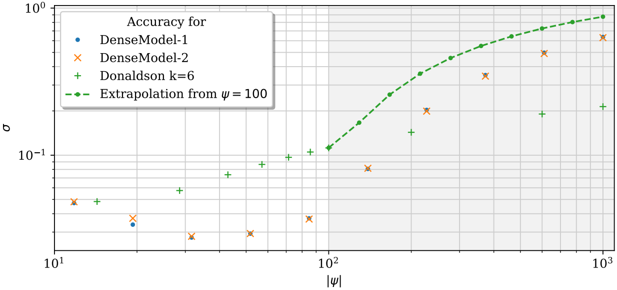
- These networks were optimized for $0 \leq |\psi| \leq 10$



(shaded region denotes extrapolation of the networks).

- Note Donaldson algorithm with $k = 12$ takes order days to run even for the single case of $\psi = 0$.
- This network at $k = 6$ takes only minutes and gives comparable accuracy

- Here are some networks trained over a larger range:



- We see again that we do better than Donaldson Alg. 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 work.

Direct learning of the metric

- Instead of learning parameters in an ansatz for the Kahler potential we can try to learn the CY metric directly.
- Why try?
  - Perhaps we can improve performance by not being tied to an ansatz at fixed k.
  - We will be able to generalize this approach to more complicated geometries.
- One disadvantage:
  - We now need losses to check that the metric is globally well defined and Kahler!
- We did a few different experiments with this – will only sketch one here.

Loss functions and Boosts

- We use

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

- Here $\mathcal{L}_{\rm MA}$ is the loss described before and we add to this

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

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

- Losses used to enforce that the two form is closed and globally consistent.
- We tried both additive and multiplicative boosts:

$$g_{CY} = g_{FS} + g_{NN} \qquad \text{vs} \qquad g_{CY} = g_{FS}(1 + g_{NN})$$

- Multiplicative boosting works best (linear boosting didn't do much better than just using the FS metric and depends sensitively on the λ's)

Procedure

- Input: $Re(z_i)$, $Im(z_i)$ (homogeneous coords describing pt in CY) $Re(\psi)$, $Im(\psi)$
- Output d^2 real and imaginary parts of a metric at that point.
- To give a concrete example, lets look at a case optimized at $\psi = 10$ on a data set of $10,000$ points.
- We split the points according to train:test=$90 : 10$ and we train for 20 epochs.
- Accuracy reaches same level as Donaldson Alg. at $k = 5$ (we expect more points and better architecture will easily improve this).

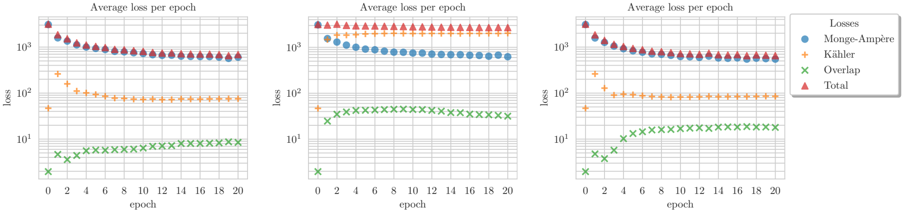


Figure: *Left:* Optimizing the NN with all three losses. *Middle:* Optimizing the NN without Kähler loss (i.e. $\lambda_2 = 0$). *Right:* Optimizing the NN without overlap loss (i.e. $\lambda_3 = 0$).

Learning $SU(3)$ structures from an anzatz

- One important class of geometries for $\mathcal{N} = 1$ compactifications: $SU(3)$ structure manifolds
- These are six-manifolds with a nowhere vanishing two form J and three form Ω obeying the same algebraic properties as the Calabi-Yau threefold case:

$$J \wedge \Omega = 0 \qquad\qquad J \wedge J \wedge J = \frac{3i}{4}\Omega \wedge \overline{\Omega} \tag{7}$$

But with different differential properties...

- An $SU(3)$ structure can be classified by its torsion classes:

$$
\begin{aligned}
dJ &= -\frac{3}{2}\mathrm{Im}(W_1\overline{\Omega}) + W_4 \wedge J + W_3 \\
d\Omega &= W_1 J \wedge J + W_2 \wedge J + W_5 \wedge \Omega \,,
\end{aligned}
$$

- Where torsion classes are given the defining forms:

$$
W_1 = -\frac{1}{6} i\Omega \lrcorner dJ = \frac{1}{12} J^2 \lrcorner d\Omega \ \ , \ \ W_4 = \frac{1}{2} J \lrcorner dJ \ \ , \ \ W_5 = -\frac{1}{2}\Omega_+ \lrcorner d\Omega_+
$$

- Given string theories place different constraints on the torsion classes for there to be an associated solution to the theory of the type we want.
- E.g. heterotic string theory: $W_1 = W_2 = 0$, $W_4 = \frac{1}{2} W_5 = d\phi$, W_3 free.
- Note that a CY structure is a special case: $W_i = 0 \ \forall i = 1, \ldots 5$.

Learning an anzatz

- Our goal was to start with some well-controlled/simple example.
- Observation: Some CY manifolds admit not only Ricci-flat metrics, but other $SU(3)$ structures as well.
- Here is an anzatz (generalization of that appearing in work by Larfors, Lukas and Ruehle, 1805.08499)

$$J = \sum_{i=1}^{h^{1,1}(x)} a_i J_i \qquad\qquad \Omega = A_1 \Omega_0 + A_2 \bar{\Omega}_0 \tag{8}$$

- The a_i are real functions and A_1 and A_2 are complex functions. CY taken to be a complete intersection in a product of projective spaces ($7,890$ such CICY e.g.s).

- Resulting torsion classes

$$
\begin{aligned}
W_1 &= 0 \\
W_2 &= -i\bar{\partial}A_1 \lrcorner \Omega_0 + i\partial A_2 \lrcorner \overline{\Omega}_0 + i\frac{\bar{\partial}(A_1+\overline{A}_2)}{A_1+\overline{A}_2} \lrcorner A_1\Omega_0 - i\frac{\partial(\overline{A}_1+A_2)}{\overline{A}_1+A_2} \lrcorner A_2\overline{\Omega}_0 \\
W_3 &= \sum_i (da_i - W_4)\wedge J_i \\
W_4 &= \frac{1}{2}\sum_i J_i \lrcorner (da_i \wedge J_i) \\
W_5 &= \frac{\bar{\partial}(A_1+\overline{A}_2)}{A_1+\overline{A}_2} + \frac{\partial(\overline{A}_1+A_2)}{\overline{A}_1+A_2}\ .
\end{aligned} \tag{9}
$$

- In principle we could use such an ansatz as a basis to learn appropriate SU(3) structures for string compactification.
- If we define Λ_{ijk} by $J_i \wedge J_j \wedge J_k = \frac{3}{4} i \Lambda_{ijk} \Omega_0 \wedge \overline{\Omega}_0$ then (7) implies $|A_1|^2 + |A_2|^2 = \sum_{i,j,k=1}^{m} \Lambda_{ijk} a_i a_j a_k$

- Then define the following $\mathcal{L}_{Strominger} = \gamma_1 \mathcal{L}_{SU(3)} + \gamma_2 \mathcal{L}_{W_2} + \gamma_3 \mathcal{L}_{W_4} + \gamma_4 \mathcal{L}_{W_5}$
- Where

$$\mathcal{L}_{SU(3)} = \left\| |A_1|^2 + |A_2|^2 - \sum_{i,j,k=1}^{m} \Lambda_{ijk} a_i a_j a_k \right\|_n \tag{10}$$

$$\mathcal{L}_{W_2} = \left\| d\Omega + (\frac{1}{2}\Omega_+ \lrcorner d\Omega_+) \wedge \Omega \right\|_n$$

$$\mathcal{L}_{W_4} = \| J \lrcorner dJ - d\phi \|_n$$

$$\mathcal{L}_{W_5} = \left\| -\frac{1}{2}\Omega_+ \lrcorner d\Omega_+ - d\phi \right\|_n$$

- This could be used to train a neural network with
  - Inputs: $Re(z_i)$, $Im(z_i)$ of point on manifold
  - Outputs: ϕ, h_i, a_i, A_1 and A_2
- We haven't actually done this! Perhaps the most important reason to show you this ansatz today is to introduce an example on the quintic...

- Quintic E.g.

$$a_1 = \frac{1}{\pi^3}\frac{|\nabla p|^2}{\sigma^4} \ , \ \ A_1 = a_1^2 \ , \ \ A_2 = 0 \ , \tag{11}$$

$$\text{where} \quad \sigma = \sum_{a=0}^{4} |X_a|^2 \ .$$

with p the defining equation of the hypersurface.

- 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 (this solution is from 1805.08499).
- I want to show you results where we reproduce this known analytic solution using machine learning techniques where we learn the metric directly.
- Such a check is particularly important in learning such metrics as we have no analogue of Yau's theorem to use to argue we are converging towards an exact/unique solution.

Direct learning of the $SU(3)$ structure metric

- Lets see if we can learn the two form J of the known analytic solution just presented.
- Because Ω is fixed in that case we know that

$$W_1 = W_2 = 0 \qquad\qquad W_5 = 2d(ln(a_1)) \tag{12}$$

- To try and force the network towards the known solution we need

$$W_3 = 0 \qquad\qquad W_4 = d(ln(a_1)) \tag{13}$$

- We are going to use the same losses as in the Calabi-Yau case, then, with the exception of replacing the Kahler loss by the following.

$$\mathcal{L}_{W_4} = \|dJ - dln(a_1 \wedge J)\|_n$$

- We ran this for the $\psi = 10$ quintic, using multiplicative boosting from g_{FS}.
- Input and outputs mimic our experiment for directly learning the Calabi-Yau metrics we saw earlier (except that now a non-Ricci flat metric is output).
- We can see by observing the evolution of the losses that we do indeed seem to be approaching an SU(3) structure solution...

Results

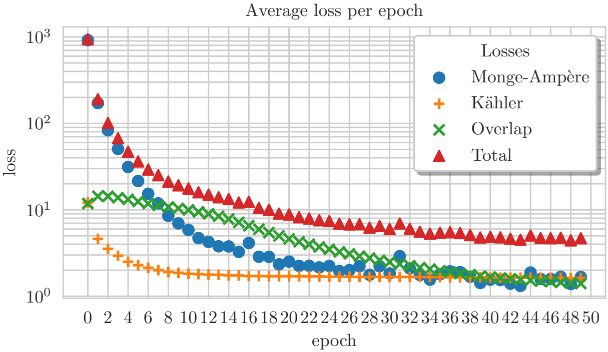


Figure: Change in loss during training for the $SU(3)$-structure example.

- Can we verify it's approaching the expected solution?...

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

$$\mathcal{E}_{\mathrm{known}} = ||g_{\mathrm{numeric}} - g_{\mathrm{known}}||_n \tag{14}$$

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

$$\mathcal{E}^{FS}_{known} = 0.511 \qquad \mathcal{E}^{NN}_{known} = 0.025 \tag{15}$$

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

Conclusions

We have presented a number of experiments.

- Supervised learning of moduli dependence of Calabi-Yau metrics.
- Direct learning of moduli dependent Calabi-Yau metrics (both using an ansatz and not)
- Direct learning of metrics associated to $SU(3)$ structures with torsion.
- The application of machine learning techniques in this arena seems very promising... Much more to consider. Stick around for the discussion session!