Approximating Hermitian Yang–Mills connections on vector bundles

Anonymous Author(s)
Affiliation
Address
email

Abstract

Geometric objects of interest to physicists typically arise as the solution to challenging nonlinear systems of PDEs on manifolds. In this work, we propose a simple two–stage procedure to approximate one such system, the Hermitian Yang–Mills equations on a holomorphic vector bundle over a Kähler manifold. The main challenge here is developing an appropriate fully–differentiable parameterisation of the associated tensor fields in a manner which respects the symmetries and topology of the manifold.

1 Introduction

String theory is a leading candidate for a unified description of fundamental physics. However, it suffers great difficulty in making empirically verifiable physical predictions. One reason for this is that string theory does not provide unique predictions for observable physics, but instead offers a bewildering array of possibilities. This may be attributed to the intertwining between physics and geometry at the heart of string theory. In its original formulation, spacetime is ten—dimensional. To account for the missing six dimensions, spacetime is thought of as locally represented by a 'product' of two spaces. The first is the four-dimensional observable spacetime M_4 we inhabit, with its familiar set of low-energy physical laws. The second is a six-dimensional compact space X which is sufficiently 'small' to be hidden from direct observation. Schematically,

$$\{\text{Spacetime}\} \simeq M_4 \times X \ . \tag{1}$$

This construction asserts that the **physics of M_4 is completely determined by the geometry and topology of the extra dimensions of X** — the 'compactification manifold'. Connecting string models to observable physics is a matter of understanding the geometry of the hidden dimensions.

Informally, this geometric data consists of *tensor fields* $\mathscr{T}$ over the compactification manifold X, which may assume values in some vector bundle V over X. Tensors on manifolds are objects with an intrinsic geometric and physical meaning — however, for computational purposes, these may be (given an *arbitrary* choice of local coordinates on X) regarded as multidimensional arrays possessing certain transformation properties.[1]

The tensor fields of interest to us satisfy certain highly non–trivial systems of geometric partial differential equations which arise from the physics of string compactifications Greene et al. [1986]. Solving these systems exactly is a difficult problem that has defied four decades of effort, and we still do not understand how to compute the geometric data necessary to make physical predictions. This is a significant bottleneck to model—building as a good deal of important physics — for example, the precise values of particle masses and their interaction strengths depends on this geometric information. For this reason, it is natural to turn to general numerical techniques to compute these quantities. This

[1] It is important to emphasise that a tensor exists independently of any coordinate representation.

Submitted to 39th Conference on Neural Information Processing Systems (NeurIPS 2025). Do not distribute.

is not as much as a abdication of responsibility as may appear — numerical methods are necessary to address even seemingly elementary problems in physics such as n–body dynamics or solving the Schrödinger equation for the helium atom.

1.1 Hermitian Yang–Mills equations

In the sequel, for standard physical reasons, we will be interested in the scenario where X is a Kähler manifold of Calabi–Yau type Hübsch [1994] equipped with a Kähler form ω, and the vector bundle $V \to (X, \omega)$ is a stable holomorphic bundle Huybrechts [2005]. It will suffice for our purposes to note existence and uniqueness of the geometric data we are interested in is guaranteed in this setting. X should be regarded as the arena which the fields $\mathscr{T}$ inhabit, and V as an abstract space which interacts with $\mathscr{T}$. In particular, V admits a unique connection (Hermitian structure) $\nabla(H)$ and associated curvature form $F \in \Omega^{1,1}(\mathrm{End}(V))$ satisfying the **Hermitian Yang–Mills (HYM) equations** on $V \to (X, \omega)$.

$$\Lambda F := \omega^{\bar{j}i} F^a{}_{b\mu\bar{\nu}} = \lambda \mathbf{1}_V \ . \tag{2}$$

Here $\mathbf{1}_V$ is the identity matrix of the fibres of V and the Einstein constant λ is some real constant purely depending on the topology of V. This is a second order elliptic partial differential equation in the Hermitian bundle metric H. An explicit form for H is not known for any non–trivial example of $V \to X$ with non–Abelian structure group. From a computational perspective, the curvature tensor $F^a{}_{b\mu\bar{\nu}}$ is a four–index object where Latin indices $a = 1, \ldots, \mathrm{rank}\, V$ run over the fibre dimension and Greek indices $\mu = 1, \ldots, \dim X$ correspond to the manifold X.

Under certain technical conditions on the bundle $V \to (X, \omega)$, a proof in the mid 1980s emerged that a solution to this system exists and is unique. Donaldson [1985], Uhlenbeck and Yau [1986]. However, this theorem is non–constructive, and it is natural to turn to numerical approximations of the unique HYM connection on V.

2 Approach

Many interesting problems in differential geometry may be phrased as finding the 'optimal' representative in a given cohomology class, which will exist and be unique for the systems we are interested in. We will follow this approach to find the unique HYM connection on a stable bundle V. The curvature F^∇ is a closed $(1,1)$–form taking values in $\mathrm{End}(V)$. Starting with some reference connection ∇_0 on V, we will search for the HYM connection via a $\partial\bar{\partial}$–exact correction to the reference curvature,

$$F^\nabla = F^{\nabla_0} + \partial\bar{\partial}\beta \ , \quad \beta \in \Gamma(\mathrm{End}(V)) \ . \tag{3}$$

That the true solution to the HYM equations assumes this form is guaranteed by the $\bar{\partial}\partial$–lemma. Our hypothesis space for the HYM connection is now the space of all Hermitian endomorphisms of V w.r.t. the background Hermitian structure H_0, which is taken to be the generalised Fubini–Study metric on V Keller [2006]. We follow the general procedure outlined in Butbaia et al. [2025]:

1. Via some geometric ansatz, reduce the problem to finding a vector–valued **global function** $\{u : u_i \in C^\infty(X)\}$, s.t. the HYM condition (2) is locally satisfied. *i.e.*, u must be independent of the choice of local coordinates.
2. Develop a variational formulation such that finding a solution to the HYM equations is equivalent to minimisation of an objective functional over some function class $\mathcal{U}$ containing the true solution,
$$\Lambda F = \lambda \mathbf{1}_V \iff \min_{u \in \mathcal{U}} \mathscr{L}[u] \ . \tag{4}$$
3. Discretise the problem by parameterising u by some ansatz u_θ, typically a neural network. The variational objective is minimised in the parameter space of the restricted function class,
$$\beta^* := \operatorname*{argmin}_{\theta \in \Theta} \mathscr{L}(\cdots ; \theta) \ .$$

The true solution will not lie in this discretised function space, but 'correctness' of the solution should be an open condition — in the sense that hypotheses sharing similar values of the variational objective $\mathcal{L}$ should exhibit similar macroscopic properties. Assuming this is true, hypotheses close to the optimum (4) may be used as a substitute for the true solutions in subsequent computations. Previous

work concerned with finding optimal Riemannian metrics on X has found numerical evidence that this is indeed the case Butbaia et al. [2024], Berglund et al. [2024].

Ultimately, we would like to parameterise tensor fields $\mathscr{T}$ in a fully–differentiable manner, as they are subject to differential operators on X. These tensor fields are not merely local multidimensional arrays — their global form is significantly constrained; they must behave in a certain way under the symmetries of X and respect the topology of X. Any appropriate numerical approximation scheme should enforce these properties by construction, significantly constraining the search space that must be navigated by any parameterised ansatz.

2.1 Objective functional

To motivate our approach, first recall that a closed $(1,1)$ form ξ is harmonic w.r.t. (X,ω) if and only if the contraction $\Lambda\xi \in C^\infty(X)$ is constant. Define $\xi := \operatorname{Tr} F \in H^{1,1}(X;\mathbb{C})$ as the trace of the curvature form over the endomorphism indices. Then the curvature F satisfies the HYM condition if and only if ξ is harmonic and the trace–free part of F vanishes;

$$\Lambda F = \gamma \mathbf{1} \iff \left[\operatorname{Tr} F \text{ harmonic } \wedge\ F - \left(\frac{1}{\text{rank}(V)} \operatorname{Tr} F \right) \mathbf{1}_V = 0 \right] . \tag{5}$$

We begin with a background Hermitian structure H_0 and deform H_0 such that the final metric satisfies the Hermitian–Einstein equation (2). That is, we want to find the endomorphism $h \in \Gamma\left(\mathsf{End}(V)\right)$ taking $H_0 \mapsto hH_0$, the metric corresponding to the HYM connection. We phrase this as a two–stage optimisation process;

- First we constrain the trace of the colour matrix ΛF to the constant value determined by the HYM condition (2). This corresponds to simply solving the Poisson equation over X, a solution of which always exists on a compact manifold.
- Subsequently, we hold $\det h$ fixed and optimise our covariant ansatz for the endomorphism h, described in Section 2.2, to eliminate the non–Abelian trace–free part of the curvature.

2.1.1 Abelian part

The first stage seeks a harmonic representative of the cohomology $[\operatorname{Tr} F^{\nabla_0}]$. On a compact Kähler manifold, this can always be done through a conformal change to the metric, thanks to the $\partial\bar\partial$–lemma. Accordingly, we make the ansatz $H' = e^f H_0$, where $f \in C^\infty(X;\mathbb{R})$ is a learnable global function representing the conformal factor.

Recall that for the determinant line bundle, $\det V = \bigwedge^n V$, the curvature form is given by $F_{\det V} = \partial\bar\partial \log\det H = \operatorname{Tr} F$. Then our ansatz assumes the form

$$\begin{aligned} \eta &:= \partial\bar\partial \log\det H_0 + \partial\bar\partial \log\det h \\ &= \xi + (\operatorname{rank} V)\partial\bar\partial f \in H^{1,1}(X)\ ,\ f \in C^\infty(X,\mathbb{R}) \end{aligned} \tag{6}$$

To find the unique harmonic representative in $[\operatorname{Tr} F^{\nabla_0}]$, noting that η is $\bar\partial$–closed, our objective function is simply the norm of the codifferential of η. As the curvature of $\det V$ is also ∂–closed, this is equivalent to $\Lambda\eta =$ constant by the first Kähler identity

$$i\bar\partial^\dagger = [\partial, \Lambda] \implies (\bar\partial^\dagger \eta)_\lambda = \partial_\lambda (g^{\bar\nu\mu}\eta_{\mu\bar\nu})\ . \tag{7}$$

2.1.2 Non–Abelian part

In the second stage, we approximate the non–Abelian part of the HYM curvature form by modelling a section of the endomorphism bundle $h : H \mapsto h \cdot e^f \cdot H_0$, holding the conformal factor f fixed. Here $\det h$ is fixed to the constant function 1 to preserve harmonicity of η. First recall the relationship between any pair of Hermitian metrics (H, H_0) on V related by some smooth endomorphism h,

$$F^\nabla = F^{\nabla_0} + \bar\partial\left(h^{-1}(\partial_{H_0} h)\right)\ . \tag{8}$$

The task is to find the representative in $[F_{H_0}]$ with vanishing trace–free part via optimisation of some ansatz for h. Denote this trace–free part by

$$F_0 := F - \left(\frac{1}{\text{rank}(V)} \operatorname{Tr} F \right) \mathbf{1}_V\ .$$

There is a natural objective function for this stage; the L_2 norm of the trace–free component of the curvature, essentially a modified Yang–Mills energy functional;

$$\mathscr{E}_{\text{YM}}[\nabla] := \int_X \operatorname{Tr}\left(F_0^\nabla \wedge \star F_0^\nabla\right) \ . \tag{9}$$

2.2 Equivariant ansatz

Provided the ansatz for $h \in \Gamma(\text{End}(V))$ transforms correctly as a global section of the endomorphism bundle, the hypothesis (3) remains an honest curvature form at any point during optimisation. Expanding h in a basis of sections for V and $V^\vee$ renders the ansatz manifestly equivariant w.r.t. coordinate transformations by construction. Bundles of interest in the context of string compactifications ($c_1 = 0$, slope–stable), will not have global sections, and one must instead work on the twisted bundle $V(k) := V \otimes \mathcal{L}^k$, where $\mathcal{L}$ is some ample line bundle. Given a basis of global sections $\{S^a_m\}$ of $V(k)$, and a choice of frame $\{e_a\}$ for V, one may expand our ansatz as

$$\begin{aligned} h &= h^a_b e_a \otimes e^b = h^{mn}(S^a_m e_a) \otimes ((S^b)_{nb} e^b) \\ &= h^{mn} S^a_m (H_0)_{b\bar{c}} \bar{S}^{\bar{c}}_n e_a \otimes e^b \ . \end{aligned} \tag{10}$$

Here a, b denote colour indices and m, n enumerate the elements of the basis. We parameterise h^{mn} as a Hermitian matrix of manifestly global functions, using the spectral network construction Berglund et al. [2023].

2.3 Stable bundle over Fermat quintic

The first example we shall consider is a rank three stable vector bundle over the Fermat quintic $X \subset \mathbb{P}^4$ (*cf.* Douglas et al. [2006], Douglas et al. [2007]), defined by the short exact sequence

$$0 \longrightarrow O_X(-1) \xrightarrow{f} \bigoplus_{i=1}^{4} O_X \longrightarrow V \longrightarrow 0 \ . \tag{11}$$

The homomorphism f is given by four generic non–intersecting sections of $O_X(1)$, and we choose $f = (Z_0, \ldots, Z_3)$. Firstly, we approximate the Ricci–flat Calabi–Yau metric in the lone Kähler class $[\omega_0]$ on X defined by restriction of the ambient Fubini–Study form on $\mathbb{P}^4$. This procedure is described in Larfors et al. [2022].

Next, we parameterise the conformal factor f (6) using the spectral network construction Berglund et al. [2023] and fix the trace of the curvature form F to the appropriate constant value by minimising the codifferential of $\operatorname{Tr} F$ (7) with respect to the parameters of f. The trajectories of the variance and norm of the codifferential during optimisation are depicted in Figure 1. Fixing the parameters of the spectral network, we then carry out the non–Abelian part of the program by minimising the Yang–Mills energy functional (9) with respect to the parameters of the endomorphism ansatz (10). The trajectories of various quantities of interest for this stage are depicted in Figure 2.

We study how closely our approximation satisfies the HYM condition (2), by evaluating the average value of the function ΛF and the associated Monte Carlo error over an independent set of 500,000 points sampled from X, and find

$$\langle \Lambda F \rangle = \frac{1}{\text{Vol}_X} \int d\mu_\Omega \, \Lambda F = (2.000 \pm 0.002) \cdot \mathbf{I}_3 \pm \mathcal{O}(10^{-5}) \ . \tag{12}$$

We compute the variance elementwise as

$$\mathbf{V}\left[\Lambda F\right] = \frac{1}{\text{Vol}_X} \int_X d\mu_\Omega \, \left(\Lambda F - \langle \Lambda F \rangle\right)^2 \ ,$$

and find $\max \sigma\left(\Lambda F^\nabla\right) = 1.93 \times 10^{-3}$ along the diagonal, with the error for off–diagonal entries being $\mathcal{O}(10^{-4})$. We estimate the HYM condition is satisfied post–optimisation within an error of $\left(\max \sigma_{\Lambda F^\nabla}\right) / \mu_{\Lambda F^\nabla} \approx 0.097\%$. This is to be compared with the figure of merit for the background connection, at $\left(\max \sigma_{\Lambda F^{\nabla_0}}\right) / \mu_{\Lambda F^{\nabla_0}} \approx 20\%$. A compilation of these results may be found in Table 1. Recall the trace ΛF is a topological quantity; it is not surprising that this coincides for the background and post–optimisation connections under the ansatz (3) — what should be noted is that our optimised HYM connection ∇ reduces the variance of the trace by over four orders of magnitude relative to the background connection ∇_0.

References

P. Berglund, G. Butbaia, T. Hübsch, V. Jejjala, D. Mayorga Peña, C. Mishra, and J. Tan. Machine learned Calabi–Yau metrics and curvature. *Adv. in Th. Math. Phys.*, 27(4):1107–1158, 2023. doi: 10.4310/ATMP.2023.v27.n4.a3.

P. Berglund, G. Butbaia, T. Hübsch, V. Jejjala, D. Mayorga Peña, C. Mishra, and J. Tan. Precision String Phenomenology. 7 2024.

G. Butbaia, D. Mayorga Peña, J. Tan, P. Berglund, T. Hübsch, V. Jejjala, and C. Mishra. Physical Yukawa Couplings in Heterotic String Compactifications. 1 2024.

G. Butbaia, D. Mayorga Peña, J. Tan, P. Berglund, T. Hübsch, V. Jejjala, and C. Mishra. cymyc: Calabi-Yau Metrics, Yukawas, and Curvature. *JHEP*, 03:028, 2025. doi: 10.1007/JHEP03(2025)028.

S. K. Donaldson. Anti self-dual yang-mills connections over complex algebraic surfaces and stable vector bundles. *Proceedings of the London Mathematical Society*, s3-50(1):1–26, 1985. doi: https://doi.org/10.1112/plms/s3-50.1.1. URL `https://londmathsoc.onlinelibrary.wiley.com/doi/abs/10.1112/plms/s3-50.1.1`.

M. R. Douglas, R. Reinbacher, and S.-T. Yau. Branes, bundles and attractors: Bogomolov and beyond. 4 2006.

M. R. Douglas, R. L. Karp, S. Lukic, and R. Reinbacher. Numerical solution to the hermitian Yang-Mills equation on the Fermat quintic. *JHEP*, 12:083, 2007. doi: 10.1088/1126-6708/2007/12/083.

B. R. Greene, K. H. Kirklin, P. J. Miron, and G. G. Ross. A Superstring Inspired Standard Model. *Phys. Lett. B*, 180:69, 1986. doi: 10.1016/0370-2693(86)90137-1.

T. Hübsch. *Calabi–Yau Manifolds: a Bestiary for Physicists*. World Scientific Publishing Co. Inc., River Edge, NJ, 2nd edition, 1994.

D. Huybrechts. *Complex Geometry*. Springer, 2005. doi: 10.1007/b137952.

J. Keller. *Canonical metrics and Härder-Narasimhan filtration*, pages 132–148. Aug. 2006. doi: 10.1142/9789812701763_0013. URL `https://hal.science/hal-01282570`.

M. Larfors, A. Lukas, F. Ruehle, and R. Schneider. Numerical metrics for complete intersection and Kreuzer–Skarke Calabi–Yau manifolds. *Mach. Learn. Sci. Tech.*, 3(3):035014, 2022. doi: 10.1088/2632-2153/ac8e4e.

K. Uhlenbeck and S. T. Yau. On the existence of hermitian-yang-mills connections in stable vector bundles. *Communications on Pure and Applied Mathematics*, 39(S1):S257–S293, 1986. doi: https://doi.org/10.1002/cpa.3160390714. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/cpa.3160390714`.

A Optimisation trajectories

Table 1: HYM approximation measures over Fermat quintic bundle.

Quantity	Background ∇_0	Post–optimisation ∇
$\langle \Lambda F \rangle$	$\begin{pmatrix} 2.006 & 3.7\times10^{-4} & 9.4\times10^{-5} \\ 2.0\times10^{-4} & 1.996 & 3.4\times10^{-5} \\ 7.9\times10^{-5} & 9.6\times10^{-6} & 1.998 \end{pmatrix}$	$\begin{pmatrix} 2.000 & 9.3\times10^{-6} & 1.2\times10^{-5} \\ 9.3\times10^{-6} & 2.000 & -2.3\times10^{-5} \\ 1.2\times10^{-5} & -2.3\times10^{-5} & 2.000 \end{pmatrix}$
Max diagonal σ	0.41	1.9×10^{-3}
Max off-diagonal σ	0.096	9.0×10^{-5}
$\langle \Lambda \operatorname{Tr} F \rangle$	6.00 ± 1.0	6.000 ± 0.005

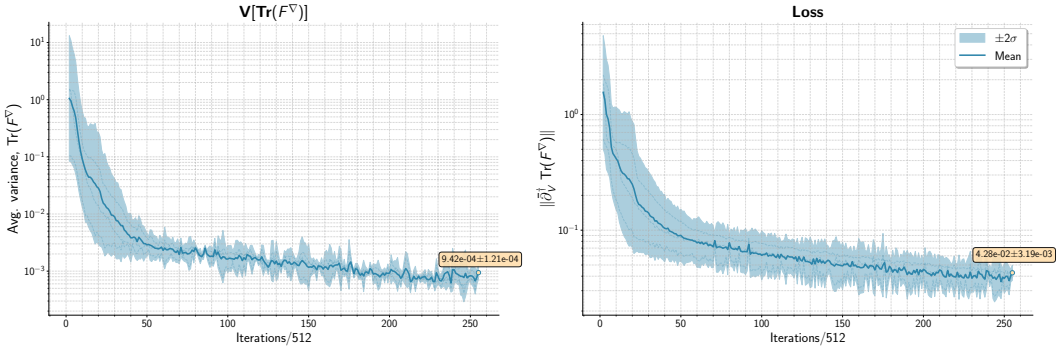

Figure 1: Evolution of the integrated variance of $\mathrm{Tr}\, F$ (left) and L_2 norm of the codifferential of $\mathrm{Tr}\, F$ (right) during first–stage optimisation for the stable bundle over the Fermat quintic with $c_1 = H$, evaluated on a separate validation set. Results are reported on three independent runs over the same dataset with respective learning rates `1e-4, 2e-4, 3e-4`.

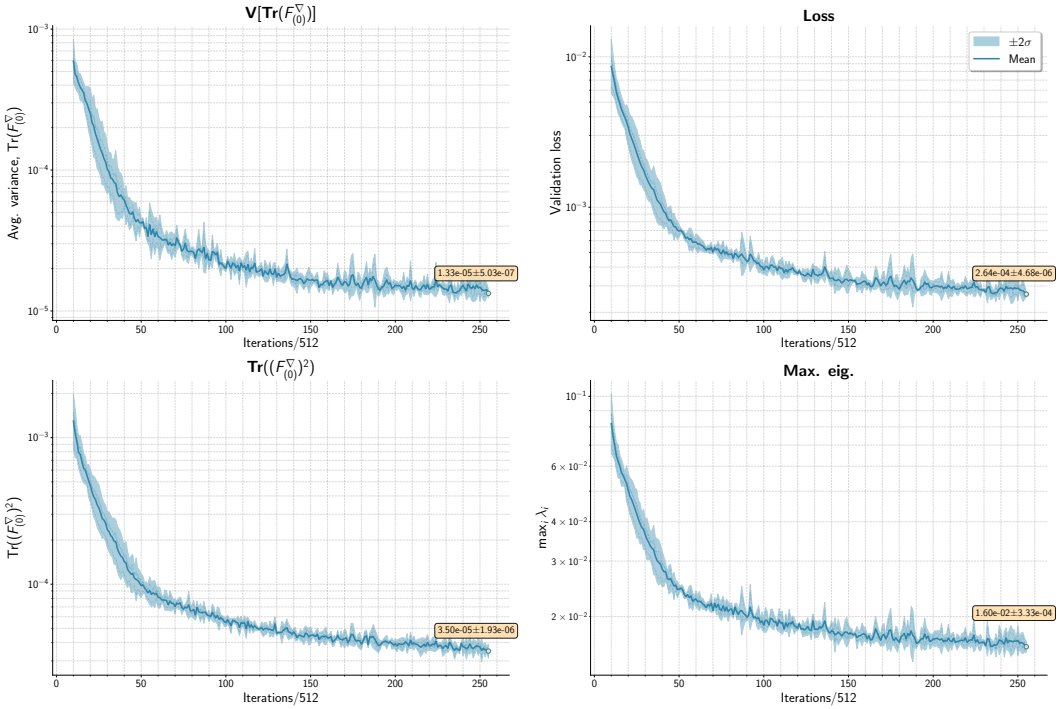

Figure 2: Evolution of various quantities during the second optimisation stage for the $c_1 = H$ Fermat quintic bundle, evaluated on a separate validation set. Clockwise from top left: *(a)*: variance of the diagonal elements of the contraction of the trace–free curvature ΛF_0, *(b)*: objective function — an upper bound on the variance of $\mathrm{Tr}\, \Lambda F_0$, *(c)*: maximum eigenvalue of F_0, *(d)*: trace of the (matrix) square of the contraction of the trace–free curvature, $\mathrm{Tr}(\Lambda F_0)^2$. Results are reported on three independent runs over the same dataset with respective learning rates `1e-4, 2e-4, 3e-4`.