Numerical CY metrics from holomorphic networks

Michael R. Douglas, Subramanian Lakshminarasimhan and Yidi Qi

Stony Brook / Harvard CMSA

MSML 2021

Abstract

Our MSML paper proposes machine learning inspired methods for computing numerical Ricci-flat Kähler metrics, and compares them with previous work. In this talk we review Kähler geometry, explain the Ricci flatness problem and embedding methods, and briefly describe our new results.

Introduction

According to superstring theory, there are six extra dimensions of space. We cannot observe them directly, because they form a very small compact manifold. But the topology and geometry of this manifold determine the spectrum of particles which we do observe.

Only very special types of manifold can be used in superstring compactification. For one thing, the manifold has to solve Einstein's equations, which are part of string theory. To a good approximation that means the manifold has to admit a Ricci flat metric (more on this below).

Very few manifolds are known to admit Ricci flat metrics. The best studied class which do is the Calabi-Yau manifolds. They can be studied in great detail using techniques of algebraic geometry. They are also favored in string theory because they allow for supersymmetry to be unbroken at low energies.

Introduction

According to superstring theory, there are six extra dimensions of space. We cannot observe them directly, because they form a very small compact manifold. But the topology and geometry of this manifold determine the spectrum of particles which we do observe.

Only very special types of manifold can be used in superstring compactification. For one thing, the manifold has to solve Einstein's equations, which are part of string theory. To a good approximation that means the manifold has to admit a Ricci flat metric (more on this below).

Very few manifolds are known to admit Ricci flat metrics. The best studied class which do is the Calabi-Yau manifolds. They can be studied in great detail using techniques of algebraic geometry. They are also favored in string theory because they allow for supersymmetry to be unbroken at low energies.

Introduction

According to superstring theory, there are six extra dimensions of space. We cannot observe them directly, because they form a very small compact manifold. But the topology and geometry of this manifold determine the spectrum of particles which we do observe.

Only very special types of manifold can be used in superstring compactification. For one thing, the manifold has to solve Einstein's equations, which are part of string theory. To a good approximation that means the manifold has to admit a Ricci flat metric (more on this below).

Very few manifolds are known to admit Ricci flat metrics. The best studied class which do is the Calabi-Yau manifolds. They can be studied in great detail using techniques of algebraic geometry. They are also favored in string theory because they allow for supersymmetry to be unbroken at low energies.

We expect that not many people in the audience will be familiar with string theory or higher dimensional algebraic geometry. We refer those who are to our paper, to arXiv:2105.03991 which adds mathematical details, and to

`https://sites.duke.edu/scshgap/michael-douglas-lectures/`

What we will do here is review Kähler geometry and describe the embedding method, before summarizing our new results.
Kähler geometry is a special case of Riemannian geometry, and many Riemannian manifolds are also Kähler manifolds. For example, every Riemann surface (a two dimensional orientable Riemannian manifold) is a Kähler manifold. Many higher dimensional manifolds are Kähler.

Kähler geometry is simpler than Riemannian geometry, but it already exhibits the most important properties of curved space, such as non-constant curvature. And projective Kähler manifolds (the case we discuss) have natural embeddings into Euclidean spaces, analogous to – but far simpler than – Laplacian eigenmaps and the like. These embeddings are a general and powerful tool.

Basic geometric definitions

Let us review some basic definitions of Riemannian geometry, and then compare with the corresponding definitions in complex and Kähler geometry.

Manifold – a space M in which every small region "looks like" a ball in D-dimensional real space $\mathbb{R}^D$. The simplest example is $\mathbb{R}^D$, and we can use as coordinates the D-component vector

$$\vec{X} = X^i = (X^1, X^2, \ldots, X^D).$$

The general definition of manifold uses coordinate charts, which are diffeomorphisms (maps) from subsets $U_i \subset M$ to subsets of $\mathbb{R}^D$. We can then define functions between manifolds $f : M \to N$, vector fields, differential forms and tensors, including the metric tensor.

$$V = \sum_{i=1}^{D} V^i(X) \frac{\partial}{\partial X^i};\ \omega = \sum_{i=1}^{D} \omega_i(X) dX^i;\ ds^2 = \sum_{1 \le i,j \le D} g_{ij}(X) dX^i dX^j.$$

Basic geometric definitions

Let us review some basic definitions of Riemannian geometry, and then compare with the corresponding definitions in complex and Kähler geometry.

Manifold – a space M in which every small region "looks like" a ball in D-dimensional real space $\mathbb{R}^D$. The simplest example is $\mathbb{R}^D$, and we can use as coordinates the D-component vector

$$\vec{X} = X^i = (X^1, X^2, \ldots, X^D).$$

The general definition of manifold uses coordinate charts, which are diffeomorphisms (maps) from subsets $U_i \subset M$ to subsets of $\mathbb{R}^D$. We can then define functions between manifolds $f : M \to N$, vector fields, differential forms and tensors, including the metric tensor.

$$V \equiv \sum_{i=1}^{D} V^i(X) \frac{\partial}{\partial X^i};\ \omega \equiv \sum_{i=1}^{D} \omega_i(X) dX^i;\ ds^2 \equiv \sum_{1 \le i,j \le D} g_{ij}(X) dX^i dX^j.$$

A complex manifold uses almost the same definitions but now the coordinates are complex variables

$$\vec{Z} = Z^i = (Z^1, Z^2, \ldots, Z^N).$$

The simplest example is N-dimensional complex space $\mathbb{C}^N$. From the point of view of differential geometry, this is the same manifold as $\mathbb{R}^{2N}$, as we could define

$$Z^1 = X^1 + iX^{N+1};\ Z^2 = X^2 + iX^{N+2};\ \ldots;\ Z^N = X^N + iX^{2N}$$

But when we use the Z's, we are using additional geometric structure, called complex structure. In the case $N = 1$, it is the same as conformal structure (local rescalings which preserve angles).

Complex coordinate transformations must be holomorphic, meaning a new coordinate W is a function of the Z's and not the complex conjugate $\bar{Z}$'s.

$$\frac{\partial W^j}{\partial \bar{Z}^{\bar{i}}} = 0; \qquad \frac{\partial}{\partial \bar{Z}^{\bar{i}}} \equiv \frac{1}{2}\left(\frac{\partial}{\partial X^i} + i\frac{\partial}{\partial X^{N+i}}\right).$$

A complex manifold uses almost the same definitions but now the coordinates are complex variables

$$\vec{Z} = Z^i = (Z^1, Z^2, \ldots, Z^N).$$

The simplest example is N-dimensional complex space $\mathbb{C}^N$. From the point of view of differential geometry, this is the same manifold as $\mathbb{R}^{2N}$, as we could define

$$Z^1 = X^1 + iX^{N+1};\ Z^2 = X^2 + iX^{N+2};\ \ldots\ ; Z^N = X^N + iX^{2N}$$

But when we use the Z's, we are using additional geometric structure, called complex structure. In the case $N = 1$, it is the same as conformal structure (local rescalings which preserve angles).

Complex coordinate transformations must be holomorphic, meaning a new coordinate W is a function of the Z's and not the complex conjugate $\bar{Z}$'s.

$$\frac{\partial W^j}{\partial \bar{Z}^{\bar{i}}} = 0; \qquad \frac{\partial}{\partial \bar{Z}^{\bar{i}}} \equiv \frac{1}{2}\left(\frac{\partial}{\partial \bar{X}^i} + i\frac{\partial}{\partial \bar{X}^{N+i}}\right).$$

In Riemannian geometry, the length of a tangent vector V at a point p is

$$||V|| = \sqrt{\sum_{1 \le i,j \le D} g_{ij}(p) V^i(p) V^j(p)}.$$

We can use the same definition in complex geometry, but usually we want to ensure that the length of a vector is a non-negative real number. This is true iff g_{ij} is hermitian, meaning

$$ds^2 = 2 g_{i\bar{j}} dZ^i d\bar{Z}^{\bar{j}}; \qquad g_{i\bar{j}} = g^*_{\bar{i}j};\ g_{ij} = g_{\bar{i}\bar{j}} = 0.$$

Example: the Riemann sphere is $\mathbb{C} \cup \{\infty\}$, with the round metric

$$ds^2 = \frac{4 dZ d\bar{Z}}{(R^2 + |Z|^2)^2} = 4 \frac{dX^2 + dY^2}{(R^2 + X^2 + Y^2)^2}$$

This is also the stereographic projection of $(x, y, z) \in \mathbb{R}^3$ as $Z = \frac{x+iy}{R-z}$.

In Riemannian geometry, the length of a tangent vector V at a point p is

$$||V|| = \sqrt{\sum_{1\leq i,j\leq D} g_{ij}(p) V^i(p) V^j(p)}.$$

We can use the same definition in complex geometry, but usually we want to ensure that the length of a vector is a non-negative real number. This is true iff g_{ij} is hermitian, meaning

$$ds^2 = 2g_{i\bar{j}} dZ^i d\bar{Z}^{\bar{j}}; \qquad g_{i\bar{j}} = g^*_{\bar{i}j};\ g_{ij} = g_{\bar{i}\bar{j}} = 0.$$

Example: the Riemann sphere is $\mathbb{C} \cup \{\infty\}$, with the round metric

$$ds^2 = \frac{4dZd\bar{Z}}{(R^2 + |Z|^2)^2} = 4\frac{dX^2 + dY^2}{(R^2 + X^2 + Y^2)^2}$$

This is also the stereographic projection of $(x, y, z) \in \mathbb{R}^3$ as $Z = \frac{x+iy}{R-z}$.

In Riemannian geometry, each component of the metric tensor g_{ij} is an independent function, so specifying a metric requires $D(D+1)/2$ functions. On the other hand one can redefine the coordinates to fix D of these functions. So the data of a metric is comparable to $D(D-1)/2$ functions, but making this precise is very complicated.

In complex geometry, a hermitian metric is specified by N^2 independent real functions, but the coordinate redefinitions must be holomorphic functions. So this problem becomes easier.

A Kähler metric is one for which

$$g_{i\bar{j}} = \frac{\partial^2 K}{\partial Z^i \partial \bar{Z}^{\bar{j}}}$$

for a (locally defined) real function $K(Z, \bar{Z})$, the Kähler potential. For the Riemann sphere the Kähler potential is

$$K = 2\log\left(R^2 + |Z|^2\right).$$

In Riemannian geometry, each component of the metric tensor g_{ij} is an independent function, so specifying a metric requires $D(D+1)/2$ functions. On the other hand one can redefine the coordinates to fix D of these functions. So the data of a metric is comparable to $D(D-1)/2$ functions, but making this precise is very complicated.

In complex geometry, a hermitian metric is specified by N^2 independent real functions, but the coordinate redefinitions must be holomorphic functions. So this problem becomes easier.

A Kähler metric is one for which

$$g_{i\bar{j}} = \frac{\partial^2 K}{\partial Z^i \partial \bar{Z}^{\bar{j}}}$$

for a (locally defined) real function $K(Z, \bar{Z})$, the Kähler potential. For the Riemann sphere the Kähler potential is

$$K = 2\log\left(R^2 + |Z|^2\right).$$

In Riemannian geometry, each component of the metric tensor g_{ij} is an independent function, so specifying a metric requires $D(D+1)/2$ functions. On the other hand one can redefine the coordinates to fix D of these functions. So the data of a metric is comparable to $D(D-1)/2$ functions, but making this precise is very complicated.

In complex geometry, a hermitian metric is specified by N^2 independent real functions, but the coordinate redefinitions must be holomorphic functions. So this problem becomes easier.

A Kähler metric is one for which

$$g_{i\bar{j}} = \frac{\partial^2 K}{\partial Z^i \partial \bar{Z}^{\bar{j}}}$$

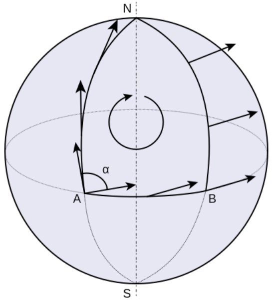
for a (locally defined) real function $K(Z, \bar{Z})$, the Kähler potential. For the Riemann sphere the Kähler potential is

$$K = 2 \log \left(R^2 + |Z|^2 \right).$$

What motivates this definition?

$$ds^2 = 2dZ^i d\bar{Z}^{\bar{j}} \frac{\partial^2 K}{\partial Z^i \partial \bar{Z}^{\bar{j}}}$$

The answer is rather deep and involves the metric compatible connection and its holonomy. In Riemannian geometry, a connection allows carrying tensors along a path from one point to another. If one carries a vector around a closed loop, in general it will undergo a rotation. The holonomy group consists of all such rotations, and for a general curved manifold one can get a general rotation in $SO(D)$.



For a complex manifold, one might expect the holonomy group to be restricted to the unitary rotations, $U(N) \subset SO(2N)$. But it turns out that this restriction only holds for a Kähler manifold.

The simplest higher dimensional Kähler manifolds are the complex projective spaces $\mathbb{CP}^N$. They are most easily defined as the quotient of $\mathbb{C}^{N+1} - \{\vec{0}\}$ by an equivalence relation,

$$(Z^1, \ldots, Z^{N+1}) \cong \lambda(Z^1, \ldots, Z^{N+1}) \qquad \forall \lambda \neq 0 \in \mathbb{C}.$$

One can use patches U_i in which $Z^i = 1$, *etc.*, related by $\lambda_{i \to j} = 1/Z_j$. In fact $\mathbb{CP}^1 \cong S^2$, but the higher dimensional cases are not equivalent to spheres, even as topological spaces.

The most symmetric metric on $\mathbb{CP}^N$ is the Fubini-Study metric, with Kähler potential

$$K = 2 \log \left(\sum_{i=1}^{N+1} |Z^i|^2 \right).$$

This does not define a function because it depends on λ. However, if we compare the Kähler potentials before and after taking $\vec{Z} \to \lambda \vec{Z}$, we find

$$K \to K + 2 \log |\lambda|^2; \qquad g_{i\bar{j}} = \partial_i \partial_{\bar{j}} K \text{ is invariant.}$$

The simplest higher dimensional Kähler manifolds are the complex projective spaces $\mathbb{CP}^N$. They are most easily defined as the quotient of $\mathbb{C}^{N+1} - \{\vec{0}\}$ by an equivalence relation,

$$(Z^1, \ldots, Z^{N+1}) \cong \lambda(Z^1, \ldots, Z^{N+1}) \qquad \forall \lambda \neq 0 \in \mathbb{C}.$$

One can use patches U_i in which $Z^i = 1$, *etc.*, related by $\lambda_{i\to j} = 1/Z_j$. In fact $\mathbb{CP}^1 \cong S^2$, but the higher dimensional cases are not equivalent to spheres, even as topological spaces.

The most symmetric metric on $\mathbb{CP}^N$ is the Fubini-Study metric, with Kähler potential

$$K = 2\log\left(\sum_{i=1}^{N+1} |Z^i|^2\right).$$

This does not define a function because it depends on λ. However, if we compare the Kähler potentials before and after taking $\vec{Z} \to \lambda\vec{Z}$, we find

$$K \to K + 2\log|\lambda|^2; \qquad g_{i\bar{j}} = \partial_i \partial_{\bar{j}} K \text{ is invariant.}$$

Many complex manifolds can be defined as submanifolds of $\mathbb{CP}^N$ – these are called projective manifolds. The simplest are the hypersurfaces in $\mathbb{CP}^N$.

A hypersurface is the subset of points satisfying a single equation $f(x) = 0$. For example the sphere is a real hypersurface in Euclidean space,

$$0 = f(X) = (X^1)^2 + (X^2)^2 + \ldots + (X^n)^2 - R^2.$$

A hypersurface will be a manifold if there is no p on the surface (so, satisfying $f(p) = 0$) with all $\partial f/\partial X|_p = 0$.
Examples of complex hypersurfaces: the Fermat hypersurfaces in $\mathbb{CP}^2$,

$$(Z^1)^n + (Z^2)^n + (Z^3)^n = 0.$$

These are Riemann surfaces of genus $g = (n-1)(n-2)/2$. On the next slide we show $n = 3$, the cubic elliptic curve, from Bozlee and Amethyst 2019 (ICERM Illustrating Mathematics),

https://im.icerm.brown.edu/portfolio/visualizing-complex-points-of-elliptic-curves/

Many complex manifolds can be defined as submanifolds of $\mathbb{CP}^N$ – these are called projective manifolds. The simplest are the hypersurfaces in $\mathbb{CP}^N$.

A hypersurface is the subset of points satisfying a single equation $f(x) = 0$. For example the sphere is a real hypersurface in Euclidean space,

$$0 = f(X) = (X^1)^2 + (X^2)^2 + \ldots + (X^n)^2 - R^2.$$

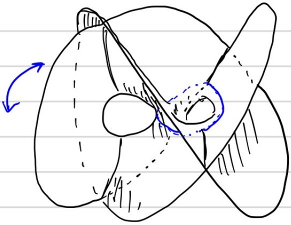
A hypersurface will be a manifold if there is no p on the surface (so, satisfying $f(p) = 0$) with all $\partial f/\partial X|_p = 0$.
Examples of complex hypersurfaces: the Fermat hypersurfaces in $\mathbb{CP}^2$,

$$(Z^1)^n + (Z^2)^n + (Z^3)^n = 0.$$

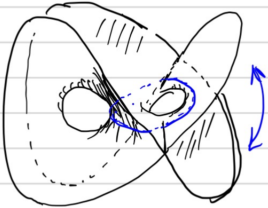
These are Riemann surfaces of genus $g = (n-1)(n-2)/2$. On the next slide we show $n = 3$, the cubic elliptic curve, from Bozlee and Amethyst 2019 (ICERM Illustrating Mathematics),

https://im.icerm.brown.edu/portfolio/visualizing-complex-points-of-elliptic-curves/

Seeing the torus:

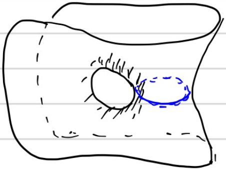


unfold left flip

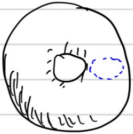


0:06 in film

unfold right flip



close up circle outside



The cubic curve ($n = 3$, genus one, the torus) is special because it admits a flat metric.
By comparison, the case $n = 2$ is topologically a sphere (and equivalent to $\mathbb{CP}^1$ as a complex manifold), while the cases $n \geq 4$ are higher genus surfaces. In both cases one can see just on topological grounds (using the Gauss-Bonnet theorem) that there could be no nonsingular flat metric.

The generalization to higher dimensions is that the degree $n = N + 1$ hypersurface in $\mathbb{CP}^N$ admits a Ricci flat metric (more about this later). For $n = 4$ and $N = 3$, so $(Z^1)^4 + (Z^2)^4 + (Z^3)^4 + (Z^4)^4 = 0$, we have a K3 surface. It is the only simply connected topological four real dimensional manifold which admits a Ricci flat metric.

The next case is the Fermat quintic with $n = 5$ and $N = 4$. The Kähler manifolds which satisfy the topological condition necessary to have a Ricci flat metric are called Calabi-Yau manifolds. Even in six real dimensions, there are many topologically distinct CY manifolds. All are projective and a large number are "hypersurfaces in toric varieties."

The cubic curve ($n = 3$, genus one, the torus) is special because it admits a flat metric.
By comparison, the case $n = 2$ is topologically a sphere (and equivalent to $\mathbb{CP}^1$ as a complex manifold), while the cases $n \geq 4$ are higher genus surfaces. In both cases one can see just on topological grounds (using the Gauss-Bonnet theorem) that there could be no nonsingular flat metric.

The generalization to higher dimensions is that the degree $n = N + 1$ hypersurface in $\mathbb{CP}^N$ admits a Ricci flat metric (more about this later). For $n = 4$ and $N = 3$, so $(Z^1)^4 + (Z^2)^4 + (Z^3)^4 + (Z^4)^4 = 0$, we have a K3 surface. It is the only simply connected topological four real dimensional manifold which admits a Ricci flat metric.

The next case is the Fermat quintic with $n = 5$ and $N = 4$. The Kähler manifolds which satisfy the topological condition necessary to have a Ricci flat metric are called Calabi-Yau manifolds. Even in six real dimensions, there are many topologically distinct CY manifolds. All are projective and a large number are "hypersurfaces in toric varieties."

The cubic curve ($n = 3$, genus one, the torus) is special because it admits a flat metric.
By comparison, the case $n = 2$ is topologically a sphere (and equivalent to $\mathbb{CP}^1$ as a complex manifold), while the cases $n \geq 4$ are higher genus surfaces. In both cases one can see just on topological grounds (using the Gauss-Bonnet theorem) that there could be no nonsingular flat metric.

The generalization to higher dimensions is that the degree $n = N + 1$ hypersurface in $\mathbb{CP}^N$ admits a Ricci flat metric (more about this later). For $n = 4$ and $N = 3$, so $(Z^1)^4 + (Z^2)^4 + (Z^3)^4 + (Z^4)^4 = 0$, we have a K3 surface. It is the only simply connected topological four real dimensional manifold which admits a Ricci flat metric.

The next case is the Fermat quintic with $n = 5$ and $N = 4$. The Kähler manifolds which satisfy the topological condition necessary to have a Ricci flat metric are called Calabi-Yau manifolds. Even in six real dimensions, there are many topologically distinct CY manifolds. All are projective and a large number are "hypersurfaces in toric varieties."

All of the formulas of Riemannian geometry for the connection and for the curvature become simpler for a Kähler metric. In particular, the Ricci tensor, which for a Riemannian metric would be

$$\begin{aligned} R_{ij} = & -\frac{1}{2}\sum_{a,b=1}^{n}\left(\frac{\partial^2 g_{ij}}{\partial x^a \partial x^b} + \frac{\partial^2 g_{ab}}{\partial x^i \partial x^j} - \frac{\partial^2 g_{ib}}{\partial x^j \partial x^a} - \frac{\partial^2 g_{jb}}{\partial x^i \partial x^a}\right) g^{ab} \\ & + \frac{1}{2}\sum_{a,b,c,d=1}^{n}\left(\frac{1}{2}\frac{\partial g_{ac}}{\partial x^i}\frac{\partial g_{bd}}{\partial x^j} + \frac{\partial g_{ic}}{\partial x^a}\frac{\partial g_{jd}}{\partial x^b} - \frac{\partial g_{ic}}{\partial x^a}\frac{\partial g_{jb}}{\partial x^d}\right) g^{ab} g^{cd} \\ & - \frac{1}{4}\sum_{a,b,c,d=1}^{n}\left(\frac{\partial g_{jc}}{\partial x^i} + \frac{\partial g_{ic}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^c}\right)\left(2\frac{\partial g_{bd}}{\partial x^a} - \frac{\partial g_{ab}}{\partial x^d}\right) g^{ab} g^{cd}. \end{aligned}$$

becomes

$$R_{i\bar{j}} = \frac{\partial^2}{\partial Z^i \partial \bar{Z}^{\bar{j}}} \log \det_{k,\bar{l}} g_{k\bar{l}}.$$

Thus it is entirely determined by the local volume element $\det g_{k\bar{l}}$.

Essentially, the Ricci tensor measures the variation of the volume element on moving in a direction v (it decreases for positive curvature).

All of the formulas of Riemannian geometry for the connection and for the curvature become simpler for a Kähler metric. In particular, the Ricci tensor, which for a Riemannian metric would be

$$
\begin{aligned}
R_{ij} = & -\frac{1}{2}\sum_{a,b=1}^{n}\left(\frac{\partial^2 g_{ij}}{\partial x^a \partial x^b} + \frac{\partial^2 g_{ab}}{\partial x^i \partial x^j} - \frac{\partial^2 g_{ib}}{\partial x^j \partial x^a} - \frac{\partial^2 g_{jb}}{\partial x^i \partial x^a}\right) g^{ab} \\
& + \frac{1}{2}\sum_{a,b,c,d=1}^{n}\left(\frac{1}{2}\frac{\partial g_{ac}}{\partial x^i}\frac{\partial g_{bd}}{\partial x^j} + \frac{\partial g_{ic}}{\partial x^a}\frac{\partial g_{jd}}{\partial x^b} - \frac{\partial g_{ic}}{\partial x^a}\frac{\partial g_{jb}}{\partial x^d}\right) g^{ab} g^{cd} \\
& - \frac{1}{4}\sum_{a,b,c,d=1}^{n}\left(\frac{\partial g_{jc}}{\partial x^i} + \frac{\partial g_{ic}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^c}\right)\left(2\frac{\partial g_{bd}}{\partial x^a} - \frac{\partial g_{ab}}{\partial x^d}\right) g^{ab} g^{cd}.
\end{aligned}
$$

becomes

$$
R_{i\bar{j}} = \frac{\partial^2}{\partial Z^i \partial \bar{Z}^{\bar{j}}} \log \det_{k,\bar{l}} g_{k\bar{l}}.
$$

Thus it is entirely determined by the local volume element $\det g_{k\bar{l}}$.

Essentially, the Ricci tensor measures the variation of the volume element on moving in a direction v (it decreases for positive curvature).

The Ricci tensor is the geometric basis for Einstein's equation of general relativity,

$$R_{ij} - \frac{1}{2} g_{ij} g^{kl} R_{kl} = 8\pi T_{ij}.$$

The right hand side is the stress-energy of matter and in the vacuum it is zero. Thus one can rewrite the vacuum Einstein's equation as the Ricci flatness condition,

$$R_{ij} = 0$$

Specializing this to Kähler geometry and using the formulas from the previous slide, it becomes

$$\frac{\partial^2}{\partial Z^i \partial \bar{Z}^{\bar{j}}} \log \det_{k,\bar{l}} \left(\frac{\partial^2 K}{\partial Z^k \partial \bar{Z}^{\bar{l}}} \right) = 0.$$

(as discussed in the paper it can be simplified a bit more). This is one elliptic PDE for one unknown function K, so on a contractable region and with given boundary conditions it should have a unique solution.

The Ricci tensor is the geometric basis for Einstein's equation of general relativity,

$$R_{ij} - \frac{1}{2} g_{ij} g^{kl} R_{kl} = 8\pi T_{ij}.$$

The right hand side is the stress-energy of matter and in the vacuum it is zero. Thus one can rewrite the vacuum Einstein's equation as the Ricci flatness condition,

$$R_{ij} = 0$$

Specializing this to Kähler geometry and using the formulas from the previous slide, it becomes

$$\frac{\partial^2}{\partial Z^i \partial \bar{Z}^{\bar{j}}} \log \det_{k,\bar{l}} \left(\frac{\partial^2 K}{\partial Z^k \partial \bar{Z}^{\bar{l}}} \right) = 0.$$

(as discussed in the paper it can be simplified a bit more). This is one elliptic PDE for one unknown function K, so on a contractable region and with given boundary conditions it should have a unique solution.

For a compact Kähler manifold which satisfies the topological condition, and given the Kähler class (the volumes of a basis of homology two-cycles), Yau proved in 1978 that there is a unique solution of this equation, and thus a Ricci flat metric. This is by far the most intricate solution to Einstein's equations known to exist – in fact the proof is nonconstructive and it is generally believed that no closed form expression for it exists.
String theory predicts that there are six "hidden" dimensions of space which must satisfy Einstein's equations, and the Calabi-Yau metrics are the leading candidates to consider.

From the point of view of numerical methods, this is an interesting and challenging PDE. It is of Monge-Ampere type: the terms with most derivatives are nonlinear. It describes highly nontrivial curved spaces about which a lot is known mathematically, so there are many checks on the results. But very little structure is given *a priori* – basic questions such as the best way to discretize space or otherwise reduce to a finite dimensional problem are wide open.

For a compact Kähler manifold which satisfies the topological condition, and given the Kähler class (the volumes of a basis of homology two-cycles), Yau proved in 1978 that there is a unique solution of this equation, and thus a Ricci flat metric. This is by far the most intricate solution to Einstein's equations known to exist – in fact the proof is nonconstructive and it is generally believed that no closed form expression for it exists.
String theory predicts that there are six “hidden” dimensions of space which must satisfy Einstein's equations, and the Calabi-Yau metrics are the leading candidates to consider.

From the point of view of numerical methods, this is an interesting and challenging PDE. It is of Monge-Ampere type: the terms with most derivatives are nonlinear. It describes highly nontrivial curved spaces about which a lot is known mathematically, so there are many checks on the results. But very little structure is given *a priori* – basic questions such as the best way to discretize space or otherwise reduce to a finite dimensional problem are wide open.

Rather than discretize the manifold M, many works (following Donaldson arXiv: math/0512625) use an embedding method. The general idea is to postulate an embedding into a higher dimensional ambient space $\mathbb{R}^K$ or $\mathbb{C}^K$ and a family of metrics $g[W]$ on this space. The restriction of the ambient metric then gives a family of metrics on M, over which one can minimize an energy function.

Of course, if M is a hypersurface in $\mathbb{CP}^N$, then it is defined as an embedding and we can restrict the Fubini-Study potential on $\mathbb{CP}^N$ to M,

$$K = 2 \log \sum_{i,\bar{j}} h_{i\bar{j}} Z^i \bar{Z}^{\bar{j}},$$

Although on $\mathbb{CP}^N$ we can set h to the identity using a coordinate transformation, this is not the case on M (it would change the equation $f = 0$).

This is a natural family of metrics which depends on $(N+1)^2$ real parameters – a good start but we want more parameters to improve the approximation.

Rather than discretize the manifold M, many works (following Donaldson arXiv: math/0512625) use an embedding method. The general idea is to postulate an embedding into a higher dimensional ambient space $\mathbb{R}^K$ or $\mathbb{C}^K$ and a family of metrics $g[W]$ on this space. The restriction of the ambient metric then gives a family of metrics on M, over which one can minimize an energy function.

Of course, if M is a hypersurface in $\mathbb{CP}^N$, then it is defined as an embedding and we can restrict the Fubini-Study potential on $\mathbb{CP}^N$ to M,

$$K = 2 \log \sum_{i,\bar{j}} h_{i\bar{j}} Z^i \bar{Z}^{\bar{j}},$$

Although on $\mathbb{CP}^N$ we can set h to the identity using a coordinate transformation, this is not the case on M (it would change the equation $f = 0$).
This is a natural family of metrics which depends on $(N+1)^2$ real parameters – a good start but we want more parameters to improve the approximation.

Rather than discretize the manifold M, many works (following Donaldson arXiv: math/0512625) use an embedding method. The general idea is to postulate an embedding into a higher dimensional ambient space $\mathbb{R}^K$ or $\mathbb{C}^K$ and a family of metrics $g[W]$ on this space. The restriction of the ambient metric then gives a family of metrics on M, over which one can minimize an energy function.

Of course, if M is a hypersurface in $\mathbb{CP}^N$, then it is defined as an embedding and we can restrict the Fubini-Study potential on $\mathbb{CP}^N$ to M,

$$K = 2\log \sum_{i,\bar{j}} h_{i\bar{j}} Z^i \bar{Z}^{\bar{j}},$$

Although on $\mathbb{CP}^N$ we can set h to the identity using a coordinate transformation, this is not the case on M (it would change the equation $f = 0$).

This is a natural family of metrics which depends on $(N+1)^2$ real parameters – a good start but we want more parameters to improve the approximation.

To get a family of metrics with more parameters, one can replace Z^i with a complete basis of holomorphic sections of a line bundle on M, defining a Kodaira embedding of M. By taking higher degree line bundles one can get as many parameters as needed.
For the special case of a submanifold $M \hookrightarrow \mathbb{CP}^N$, one can think of this as a composition of the defining embedding with a Veronese embedding. This is defined by taking all the degree k monomials in the original coordinates, so

$$\left(Z^1, \ldots, Z^n\right) \to \left((Z^1)^k, (Z^1)^{k-1}(Z^2), (Z^1)^{k-1}(Z^3), \ldots, (Z^2)^k, \ldots\right).$$

The resulting family of Kähler potentials is

$$K = 2\log \sum_{I,\bar{J}} H_{I\bar{J}}(Z^{I_1} Z^{I_2} \ldots Z^{I_k})(\bar{Z}^{\bar{J}_1} \bar{Z}^{\bar{J}_2} \ldots \bar{Z}^{\bar{J}_k})$$

The precise number of parameters depends on M, but asymptotically is $\mathcal{O}(k^{\dim_{\mathbb{R}} M})$. As $k \to \infty$ these metrics are dense in L_2.

The idea of embedding a manifold into higher dimensions to get a simpler, more linear description is familiar from machine learning:

- On the one hand, one can embed a graph into higher dimensional Euclidean space using the eigenfunctions of its Laplacian (the "Laplacian eigenmaps" technique, Belkin and Niyogi 2002). This can be done for a manifold in the same way (Bérard, Besson and Gallot 1985).
- On the other hand, one can take a set of points in high dimensions and look for a low dimensional submanifold which contains them.

Laplacian eigenmaps is a geometric embedding – it is determined by giving a metric. But in our problem we do not know the metric *a priori*.

The Kodaira embedding is also a geometric embedding, but it requires less structure to determine: a complex structure and a choice of holomorphic line bundle. In algebraic geometry one relates this second choice to the Kähler class of the manifold, so all of the additional structure of the embedding is already given to us *a priori*.

The idea of embedding a manifold into higher dimensions to get a simpler, more linear description is familiar from machine learning:

- On the one hand, one can embed a graph into higher dimensional Euclidean space using the eigenfunctions of its Laplacian (the "Laplacian eigenmaps" technique, Belkin and Niyogi 2002). This can be done for a manifold in the same way (Bérard, Besson and Gallot 1985).
- On the other hand, one can take a set of points in high dimensions and look for a low dimensional submanifold which contains them.

Laplacian eigenmaps is a geometric embedding – it is determined by giving a metric. But in our problem we do not know the metric *a priori*.

The Kodaira embedding is also a geometric embedding, but it requires less structure to determine: a complex structure and a choice of holomorphic line bundle. In algebraic geometry one relates this second choice to the Kähler class of the manifold, so all of the additional structure of the embedding is already given to us *a priori*.

Our results

- We have implemented a Tensorflow/Keras package for working with Kähler metrics on hypersurfaces in projective space, and for finding Ricci flat metrics.
- It uses the embedding method and represents a manifold M using (1) the defining equation $f = 0$ and (2) a set of points sampled from M. The Ricci flatness condition can be interpreted as finding the K such that $\det \partial\bar{\partial}K$ interpolates the constant function.
- It represents the Kähler potential using a bihomogeneous network, a FFN with inputs $Z^i \bar{Z}^{\bar{j}}$ and activation function $z \to z^2$. This defines a parameterized subset of embedded Fubini-Study metrics of very high degree $k = 2^{\text{depth}-1}$ with a controllable number of parameters $\mathcal{O}(\text{depth} \times \text{width}^2)$.

$$\mathrm{K} = \log \mathrm{W}^{(\ell)} \circ \theta_{\mathrm{D}_\ell} \circ \mathrm{W}^{(\ell-1)} \circ \ldots \circ \theta_{\mathrm{D}_2} \circ \mathrm{W}^{(2)} \circ \theta_{\mathrm{D}_1} \circ \mathrm{W}^{(1)} \circ (\mathrm{Re}\, \mathrm{Z}^{\mathrm{I}} \bar{\mathrm{Z}}^{\bar{\mathrm{J}}}, \mathrm{Im}\, \mathrm{Z}^{\mathrm{I}} \bar{\mathrm{Z}}^{\bar{\mathrm{J}}})$$

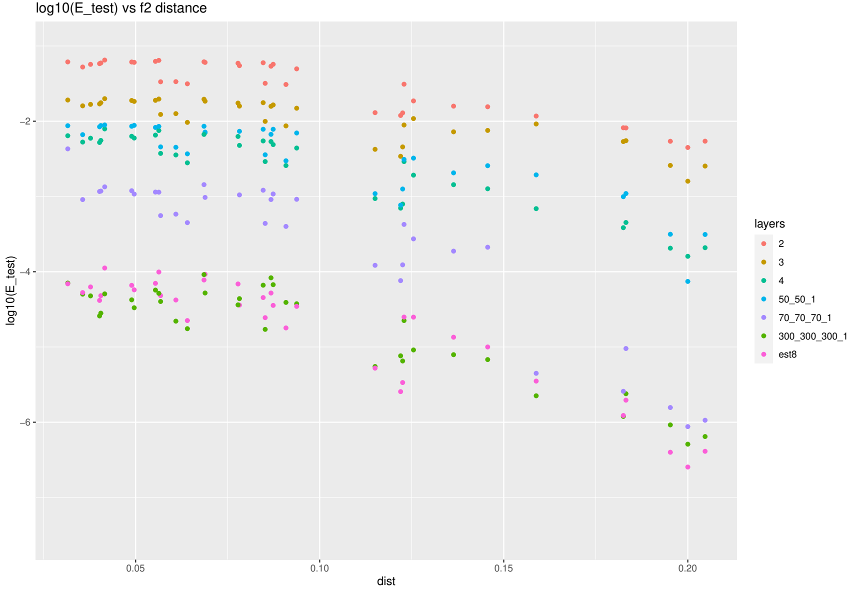
with $\theta(\mathrm{z}) = \mathrm{z}^2$ and the $\mathrm{W}^{(\mathrm{n})}$'s are real matrices of weights.

- We studied Calabi-Yau (Ricci flat Kähler) metrics on quintic hypersurfaces in $\mathbb{CP}^4$, looking at dependence on network depth and width, the symmetry group of the manifold, and the shortest length scale on the manifold (distance to discriminant locus).
- The bihomogeneous networks can often represent Ricci flat metrics with many fewer parameters than the general Fubini-Study metric. However, we have evidence that this is not always the case – metrics with no symmetry and short scales require as many parameters as the general Fubini-Study metric.

On the graph on the following slide, the y-axis is the log mean square error for the approximate Ricci flat metric, and the x-axis is (the $3/2$ power of) the shortest length scale (vanishing cycle radius) on the manifold. The colors indicate different models for the metric: $2, 3, 4$ are Fubini-Study metrics of that degree, $\mathrm{est}8$ estimates the error for degree 8 as explained in the paper, and the others are depth 2,3,4 bihomogeneous networks with the given widths. While 70_70_70_1 has fewer parameters, the 300_300_300_1 has almost as many parameters as the general degree 8 metric.

- We studied Calabi-Yau (Ricci flat Kähler) metrics on quintic hypersurfaces in $\mathbb{CP}^4$, looking at dependence on network depth and width, the symmetry group of the manifold, and the shortest length scale on the manifold (distance to discriminant locus).
- The bihomogeneous networks can often represent Ricci flat metrics with many fewer parameters than the general Fubini-Study metric. However, we have evidence that this is not always the case – metrics with no symmetry and short scales require as many parameters as the general Fubini-Study metric.

On the graph on the following slide, the y-axis is the log mean square error for the approximate Ricci flat metric, and the x-axis is (the $3/2$ power of) the shortest length scale (vanishing cycle radius) on the manifold. The colors indicate different models for the metric: $2, 3, 4$ are Fubini-Study metrics of that degree, $\mathrm{est}8$ estimates the error for degree 8 as explained in the paper, and the others are depth 2,3,4 bihomogeneous networks with the given widths. While 70_70_70_1 has fewer parameters, the 300_300_300_1 has almost as many parameters as the general degree 8 metric.


log10(E_test) vs f2 distance
log10(E_test)
dist
−2
−4
−6
0.05
0.10
0.15
0.20
layers
2
3
4
50_50_1
70_70_70_1
300_300_300_1
est8