Universes as Big-Data: Strings, Geometry & AI

YANG-HUI HE

London Institute for Mathematical Sciences, Royal Institution

Dept of Mathematics, City, **University of London**

Merton College, **University of Oxford**

School of Physics, **NanKai University**

Planck 2022, Paris

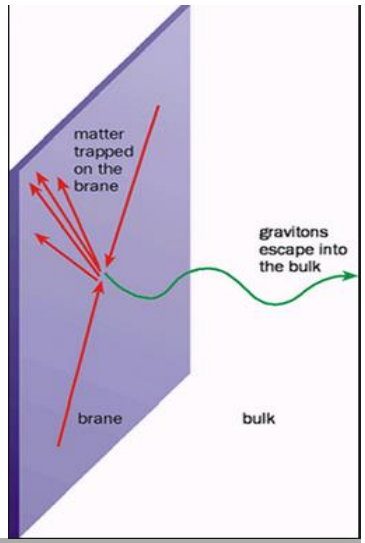
String Phenomenology

Superstring Theory 9+1 d

Unified theory of quantum gravity

I. 6 Large Dim

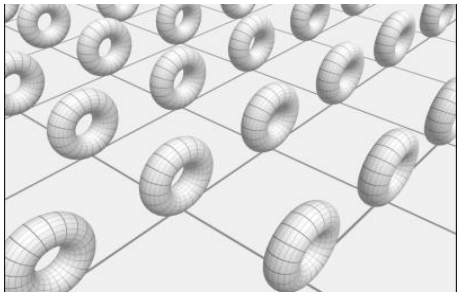
AdS/CFT
Brane World

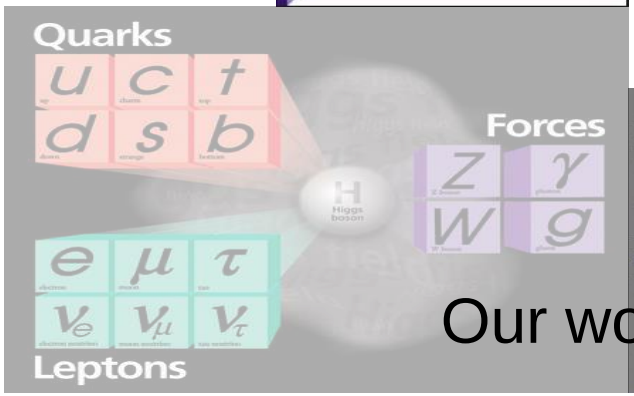


1. *Reduce Dim: 10 = 6+4*
2. *Break SUSY*

II. 6 small dim

Compactification





Our world 3+1d SU(3)xSU(2)xU(1) SM + GR

Enriching the Maths/Physics Dialogue

- Alg./diff. Geometry/topology - Rep. Theo : the right language for physics
  - Gravity $\sim$ Ricci 2-form of Tangent bundles;
  - Elementary Particles $\sim$ irred reps of the Lorentz group and sections of bundles with Lie structure group; Interactions $\sim$ Tensor products of sections ...
  - **String theory: brain-child of gauge-gravity geometrization tradition**
- A new exciting era for synergy with (pure & computational) geometry, group theory, combinatorics, number theory: *Sage, M2, GAP, LMFDB, GrDB* are becoming indispensible tools for physicists
- Interdisciplinary enterprise: cross-fertilisation of particle/string theory, phenomenology, pure mathematics, computer algorithms, data-bases, ...

Vacuum Degeneracy

Perhaps the biggest theoretical challenge to string theory:

selection criterion??? metric on the landscape???

- Douglas (2003): Statistics of String vacua
- Kachru-Kallosh-Linde-Trivedi (2003): type II/CY estimates of 10^{500}
- Taylor-YN Wang (2015-7): F-theory estimates 10^{3000} to 10^{10^5}
- Basic Reason:

Algebraic Geometry $\rightsquigarrow$ Combinatorial Geometry $\rightsquigarrow$ Exponential Growth in dim

The most famous Calabi-Yau dataset

Searching the Standard Model

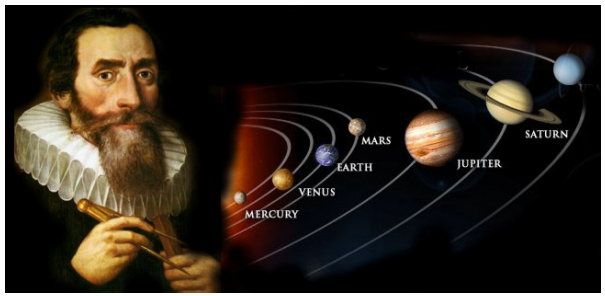
SM places some constraints but still not enough:

- Braun-YHH-Ovrut; Bouchard-Cvetic-Donagi (2005): exact MSSM particles
- Gmeiner-Blumenhagen-Honecker-Lüst-Weigand (2005):1 in 10^9 in D-brane MSSM modles
- Candelas-de la Ossa-YHH-Szendroi (2007): Triadophilia $\Rightarrow$ "des res"?
- Anderson-Gray-Lukas-Palti (2012-3): Het line bundle MSSM: 200 in 10^{10}

Recent estimates

- Constatin-YHH-Lukas; Deen-YHH-SJ Lee-Lukas (2018-9) MSSM from heterotic line bundles: 10^{23} from CICYs; 10^{723} from KS
- Cvetic-Halverson-Lin-Liu-Tian (2019): 10^{15} F-theory MSSMs

The Geometric Origin of our Universe

- Each geometry X gives a 4-D universe
  - The geometry of X determines the physics of the 4-D world
  - particles and interactions $\sim$ cohomology theory; masses $\sim$ metric; Yukawa $\sim$ Triple intersections/integral of forms over X
- 

  Ubi materia, ibi geometria

  – Johannes Kepler (1571-1630)

- Our Universe: $\begin{cases} \text{(1) probabilistic/anthropic?} \\ \text{(2) Sui generis/selection rule?} \\ \text{(3) one of multi-verse ?} \end{cases}$

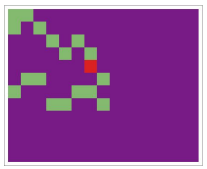
cf. *Exo-planet/Habitable Zone search*

Recasting Geometry

- A typical calculation:

$$X = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} \longrightarrow \boxed{\text{What Bourbaki teaches us}} \longrightarrow h^{2,1}(X) = 22$$

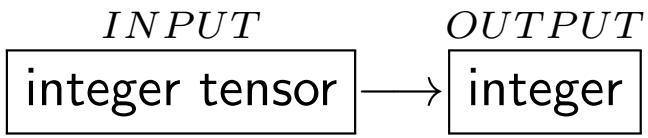
- Key to computational Algebraic Geometry: Gröbner basis, double-exponential complexity (unlike Gaussian elimination which is generalizes)
- [YHH 1706.02714] Deep-Learning the Landscape, *PLB 774, 2017*; (cf. Feature in *Science*, Aug, vol 365 issue 6452, 2019): think of it as an image processing problem



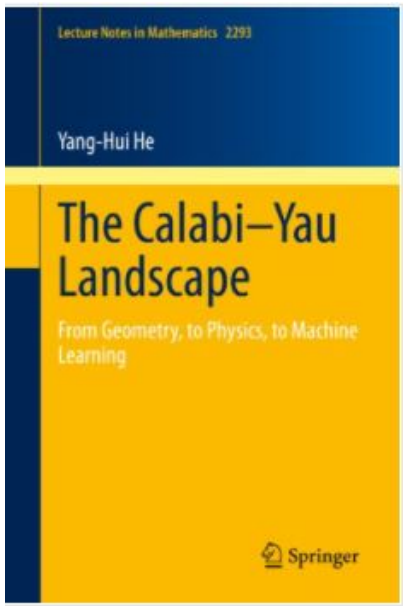
$\longrightarrow$ **What Machine-Learning teaches us** $\longrightarrow 22$

A Wild Question: "deep-learn the landscape?"

- Typical Problem in String Theory/Algebraic Geometry:



- Experimentally, it seems so for many situations in geometry and beyond.



- Initially tried a shallow MLP on CICY dataset
- improved over the years: Bull-YHH-Mishra-Jejjala, Krippendorf-Syvaeri , Erban-Finotello, Lukas, Constantin, Schneider, Larfors, ... > 99.96% accuracy

2017: String Theory enters the Machine-Learning Era

YHH (1706.02714);

q.v. Krefl-Seong (1706.03346); Ruehle (1706.07024); Carifio-Halverson-Krioukov-Nelson (1707.00655)

Of course, we are 30 years behind experimentalists!

CERN: 1990, first ML + HEP seminar series $\sim$ same time as first CY data!

Sophia: Hanson Robotics HongKong

- Beginning of `String_Data`
- How can ML and modern data-science help with the vacuum degeneracy problem??
- Meanwhile . . . Sophia becomes a "human" citizen (in Saudi Arabia)

2017: String Theory enters the Machine-Learning Era

YHH (1706.02714);

q.v. Krefl-Seong (1706.03346); Ruehle (1706.07024); Carifio-Halverson-Krioukov-Nelson (1707.00655)

Of course, we are 30 years behind experimentalists!

CERN: 1990, first ML + HEP seminar series $\sim$ same time as first CY data!

Sophia: Hanson Robotics HongKong

- Beginning of String_Data
- How can ML and modern data-science help with the vacuum degeneracy problem??
- Meanwhile . . . Sophia becomes a "human" citizen (in Saudi Arabia)

Progress in String Theory

Major International Annual Conference Series

1986- First "Strings" Conference

2002- First "StringPheno" Conference

2006 - 2010 String Vacuum Project (NSF)

2011- First "String-Math" Conference

2014- First String/Theoretical Physics Session in SIAM Conference

2017- First "String-Data" Conference

Mathematical Data

- NOISELESS Data: different from real-world data to which ML is usually applied; If I gave you 100,000 cases of

$$\text{e.g. } \begin{pmatrix} 5&3&4&3&5&1&4&4&1&2\\ 5&0&4&5&2&4&4&2&2&4\\ 1&1&2&2&0&4&1&4&5&0\\ 5&0&1&1&0&2&0&5&0&1\\ 2&5&0&1&1&3&2&3&0&3\\ 3&2&2&3&0&2&2&2&1&0\\ 3&2&5&1&4&4&0&0&1&2\\ 5&0&0&4&5&0&4&1&1&1\\ 4&3&4&3&3&1&0&0&2&5\\ 2&0&5&0&3&0&4&4&1&5 \end{pmatrix}, \quad \text{or, labeled data e.g. } \begin{pmatrix} 5&3&4&3&5&1&4&4&1&2\\ 5&0&4&5&2&4&4&2&2&4\\ 1&1&2&2&0&4&1&4&5&0\\ 5&0&1&1&0&2&0&5&0&1\\ 2&5&0&1&1&3&2&3&0&3\\ 3&2&2&3&0&2&2&2&1&0\\ 3&2&5&1&4&4&0&0&1&2\\ 5&0&0&4&5&0&4&1&1&1\\ 4&3&4&3&3&1&0&0&2&5\\ 2&0&5&0&3&0&4&4&1&5 \end{pmatrix} \longrightarrow 3$$

- Q: Is there a pattern? Can one conjecture & then prove a formula?
- Q: What branch of mathematics does it come from?
- Since 1997 - with Jiakang Bao, Elli Hayes, Ed Hirst; K. Bull, J. Ipina, T. Peterken, S. Majumdar, L. Calmon, Y. Xiao, ... and friends S-T. Yau, M. Kim, S. Lal, A. Kaspryzyk, A. Ashmore, B. Ovrut, A. Lukas, C. Mishra, V. Jejjala, ... ML different data from alg geo, diff geo, number thy, graph thy, combinatorics, rep thy, group thy, ...

Mathematical Data

- NOISELESS Data: different from real-world data to which ML is usually applied; If I gave you 100,000 cases of

$$\text{e.g. } \begin{pmatrix} 5&3&4&3&5&1&4&4&1&2\\ 5&0&4&5&2&4&4&2&2&4\\ 1&1&2&2&0&4&1&4&5&0\\ 5&0&1&1&0&2&0&5&0&1\\ 2&5&0&1&1&3&2&3&0&3\\ 3&2&2&3&0&0&2&2&1&0\\ 3&2&5&1&4&4&0&0&1&2\\ 5&0&0&0&4&5&0&4&1&1\\ 4&3&4&3&3&1&0&0&2&5\\ 2&0&5&0&3&0&4&4&1&5 \end{pmatrix}, \quad \text{or, labeled data e.g. } \begin{pmatrix} 5&3&4&3&5&1&4&4&1&2\\ 5&0&4&5&2&4&4&2&2&4\\ 1&1&2&2&0&4&1&4&5&0\\ 5&0&1&1&0&2&0&5&0&1\\ 2&5&0&1&1&3&2&3&0&3\\ 3&2&2&3&0&0&2&2&1&0\\ 3&2&5&1&4&4&0&0&1&2\\ 5&0&0&0&4&5&0&4&1&1\\ 4&3&4&3&3&1&0&0&2&5\\ 2&0&5&0&3&0&4&4&1&5 \end{pmatrix} \longrightarrow 3$$

- Q: Is there a pattern? Can one conjecture & then prove a formula?
- Q: What branch of mathematics does it come from?
- Since 1997 - with Jiakang Bao, Elli Hayes, Ed Hirst; K. Bull, J. Ipina, T. Peterken, S. Majumdar, L. Calmon, Y. Xiao, ... and friends S-T. Yau, M. Kim, S. Lal, A. Kaspryzyk, A. Ashmore, B. Ovrut, A. Lukas, C. Mishra, V. Jejjala, ... ML different data from alg geo, diff geo, number thy, graph thy, combinatorics, rep thy, group thy, ...

Mathematical Data

- NOISELESS Data: different from real-world data to which ML is usually applied; If I gave you 100,000 cases of

$$\text{e.g. } \begin{pmatrix} 5&3&4&3&5&1&4&4&1&2\\ 5&0&4&5&2&4&4&2&2&4\\ 1&1&2&2&0&4&1&4&5&0\\ 5&0&1&1&0&2&0&5&0&1\\ 2&5&0&1&1&3&2&3&0&3\\ 3&2&2&3&0&0&2&2&1&0\\ 3&2&5&1&4&4&0&0&1&2\\ 5&0&0&0&4&5&0&4&1&1\\ 4&3&4&3&3&1&0&0&2&5\\ 2&0&5&0&3&0&4&4&1&5 \end{pmatrix}, \quad \text{or, labeled data e.g. } \begin{pmatrix} 5&3&4&3&5&1&4&4&1&2\\ 5&0&4&5&2&4&4&2&2&4\\ 1&1&2&2&0&4&1&4&5&0\\ 5&0&1&1&0&2&0&5&0&1\\ 2&5&0&1&1&3&2&3&0&3\\ 3&2&2&3&0&0&2&2&1&0\\ 3&2&5&1&4&4&0&0&1&2\\ 5&0&0&0&4&5&0&4&1&1\\ 4&3&4&3&3&1&0&0&2&5\\ 2&0&5&0&3&0&4&4&1&5 \end{pmatrix} \longrightarrow 3$$

- Q: Is there a pattern? Can one conjecture & then prove a formula?
- Q: What branch of mathematics does it come from?
- Since 1997 - with Jiakang Bao, Elli Hayes, Ed Hirst; K. Bull, J. Ipina, T. Peterken, S. Majumdar, L. Calmon, Y. Xiao, ... and friends S-T. Yau, M. Kim, S. Lal, A. Kaspryzyk, A. Ashmore, B. Ovrut, A. Lukas, C. Mishra, V. Jejjala, ... ML different data from alg geo, diff geo, number thy, graph thy, combinatorics, rep thy, group thy, ...

Mathematical Data

- NOISELESS Data: different from real-world data to which ML is usually applied; If I gave you 100,000 cases of

$$\text{e.g. } \begin{pmatrix} 5&3&4&3&5&1&4&4&1&2\\ 5&0&4&5&2&4&4&2&2&4\\ 1&1&2&2&0&4&1&4&5&0\\ 5&0&1&1&0&2&0&5&0&1\\ 2&5&0&1&1&3&2&3&0&3\\ 3&2&2&3&0&0&2&2&1&0\\ 3&2&5&1&4&4&0&0&1&2\\ 5&0&0&0&4&5&0&4&1&1\\ 4&3&4&3&3&1&0&0&2&5\\ 2&0&5&0&3&0&4&4&1&5 \end{pmatrix}, \quad \text{or, labeled data e.g. } \begin{pmatrix} 5&3&4&3&5&1&4&4&1&2\\ 5&0&4&5&2&4&4&2&2&4\\ 1&1&2&2&0&4&1&4&5&0\\ 5&0&1&1&0&2&0&5&0&1\\ 2&5&0&1&1&3&2&3&0&3\\ 3&2&2&3&0&0&2&2&1&0\\ 3&2&5&1&4&4&0&0&1&2\\ 5&0&0&0&4&5&0&4&1&1\\ 4&3&4&3&3&1&0&0&2&5\\ 2&0&5&0&3&0&4&4&1&5 \end{pmatrix} \longrightarrow 3$$

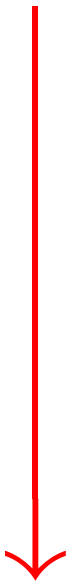
- Q: Is there a pattern? Can one conjecture & then prove a formula?
- Q: What branch of mathematics does it come from?
- Since 1997 - with Jiakang Bao, Elli Hayes, Ed Hirst; K. Bull, J. Ipina, T. Peterken, S. Majumdar, L. Calmon, Y. Xiao, ... and friends S-T. Yau, M. Kim, S. Lal, A. Kaspryzyk, A. Ashmore, B. Ovrut, A. Lukas, C. Mishra, V. Jejjala, ... ML different data from alg geo, diff geo, number thy, graph thy, combinatorics, rep thy, group thy, ...

Clearly useful for maths and physics

- Conjecture Formulation e.g.,
  - '19 YHH-Kim: separating hyperplane - simple/non-simple groups; open
  - '19 Brodie-Constantin-Lukas: exact formulae for cohomo surf.; proved.
  - '20 YHH-Lee-Oliver: L-coefs and integer pt./torsion on ell; proved.
  - '20 Craven-Jejjala-Par: Jones poly best-fit function; open
  - '22 DeepMind Collab bounds on volume conjecture for knots
  - ...
- Speed up & Improve Accuracies e.g.,
  - computing/estimating (top.inv., charges, etc) MUCH FASTER
  - '19 Ashmore-YHH-Ovrut: speed up Donaldson alg@CY metric 10-100
  - '20 Douglas et al., Anderson et al. accuracy improvement on Donaldson 10-100 times
  - ...

An Inherent Hierarchy?

- ML the structure of mathematics: YHH 2101.06317
- In decreasing precision/increasing difficulty:



numerical

string theory $\rightarrow$ algebraic geometry over $\mathbb{C} \sim$ arithmetic geometry

algebra

string theory $\rightarrow$ combinatorics

analytic number theory

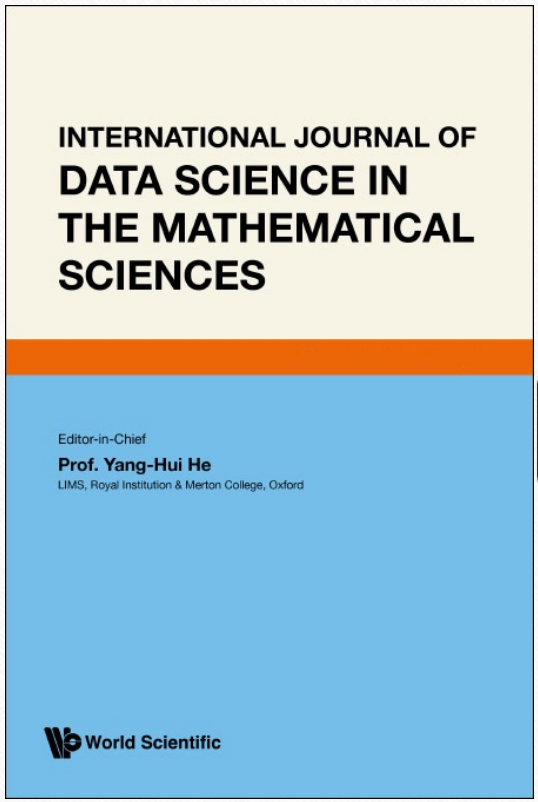
Launching in 2023

IJDSMS

Calling for Papers

Editor-in-Chief
Yang-Hui He
London Institute for Mathematical Sciences
& Merton College, University of Oxford
email: hey@maths.ox.ac.uk

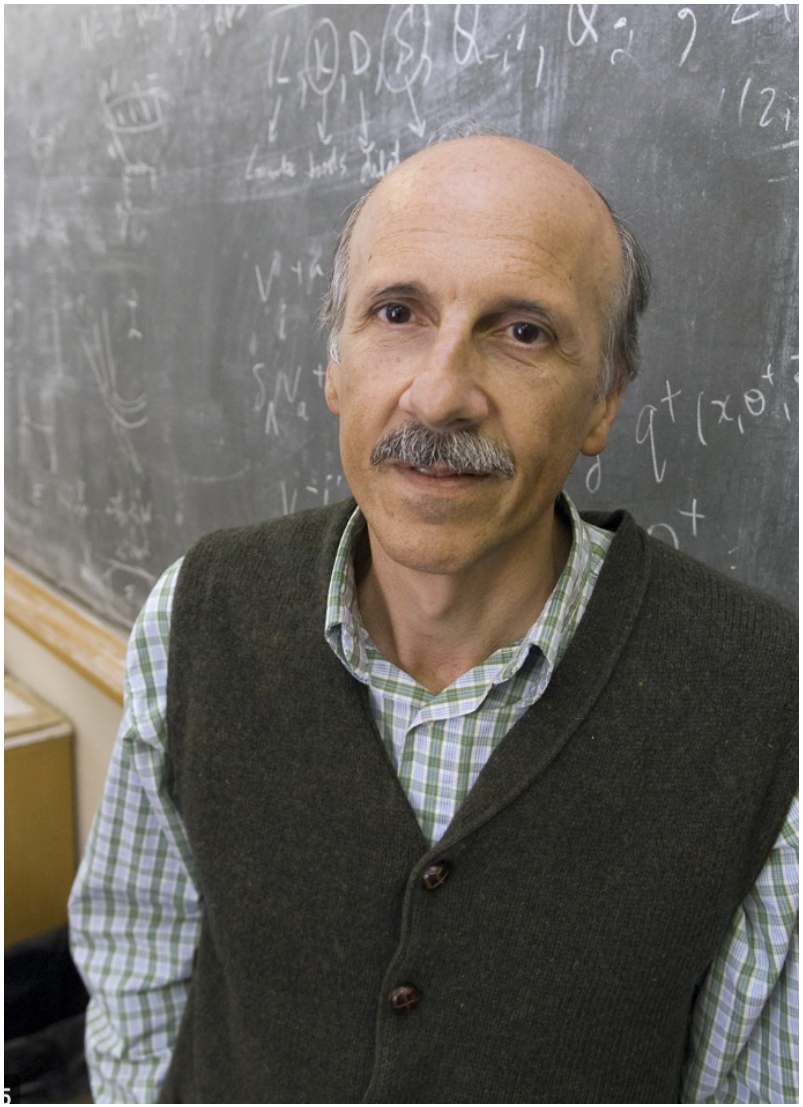
More Information:
https://www.worldscientific.com/worldscinet/ijdsms



World Scientific
Connecting Great Minds

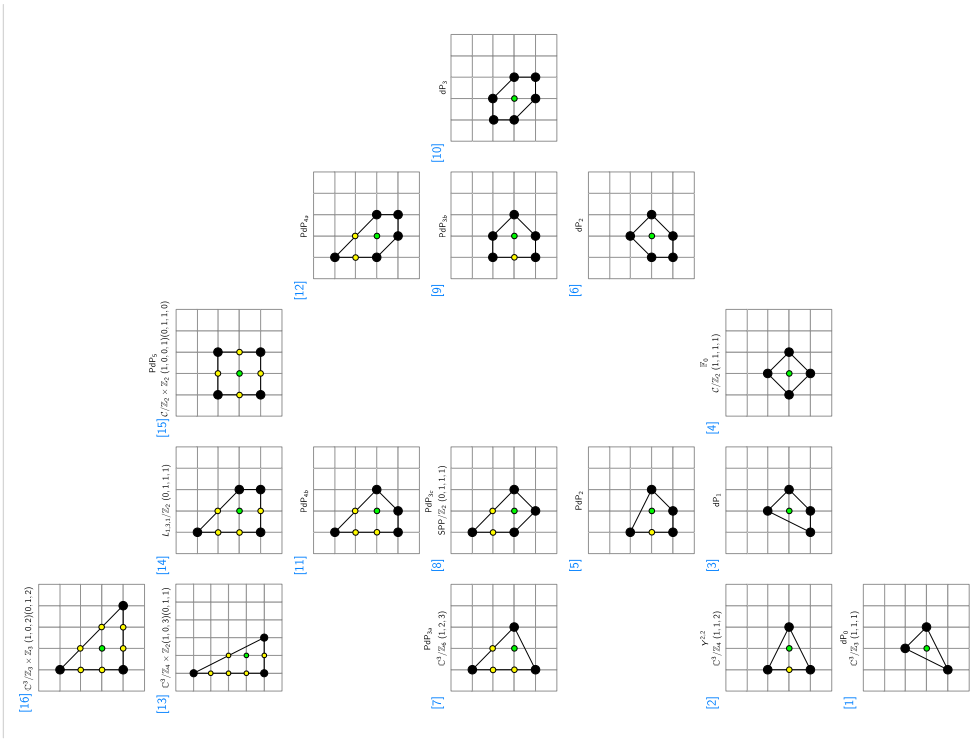
Merci!!

A toast to Prof. Ignatios Antoniadis



Borisov-Batyrev & Kreuzer-Skarke

Back to Landscape



- Reflexive Polyhedra $\rightsquigarrow$ CY: anticanonical hypersurface in toric variety from Δ
- Dim 2: 16 up to $SL(2;\mathbb{Z})$ (Italian School 1890s)
- Dim 3: 4139 up to $SL(3;\mathbb{Z})$ (KS, 1999)
- Dim 4: 473800776 up to $SL(4;\mathbb{Z})$ (KS, 2000)
- Dim > 4: Open ??

Altman-Carifio-Halverson-Nelson (2018): estimated 10^{10^4} triangulations

Altman-Gray-YHH-Jejjala-Nelson (2014): brute-force: $\sim 10^6$ up to $h^{1,1} = 6$

The Proper Way $\mathcal{O}(e^{e^d})$

- Recall Hodge decomposition $H^{p,q}(X) \simeq H^q(X, \wedge^p T^\star X) \rightsquigarrow$

$$H^{1,1}(X) = H^1(X, T_X^\star), \qquad H^{2,1}(X) \simeq H^{1,2} = H^2(X, T_X^\star) \simeq H^1(X, T_X)$$

- Euler Sequence for subvariety $X \subset A$ is short exact:

$$0 \to T_X \to T_M|_X \to N_X \to 0$$

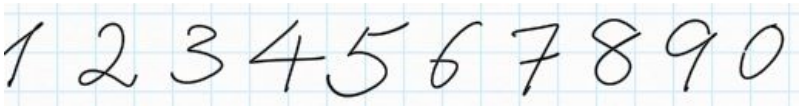
- Induces long exact sequence in cohomology:

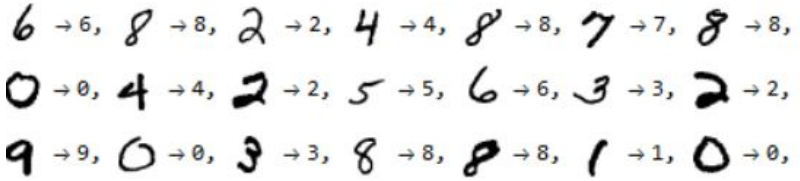
$$\begin{array}{ccccccc} 0 & \to & \cancel{H^0(X, T_X)}^{\,0} & \to & H^0(X, T_A|_X) & \to & H^0(X, N_X) \quad \to \\ & \to & \boxed{H^1(X, T_X)} & \xrightarrow{d} & H^1(X, T_A|_X) & \to & H^1(X, N_X) \quad \to \\ & \to & H^2(X, T_X) & \to & \cdots & & \end{array}$$

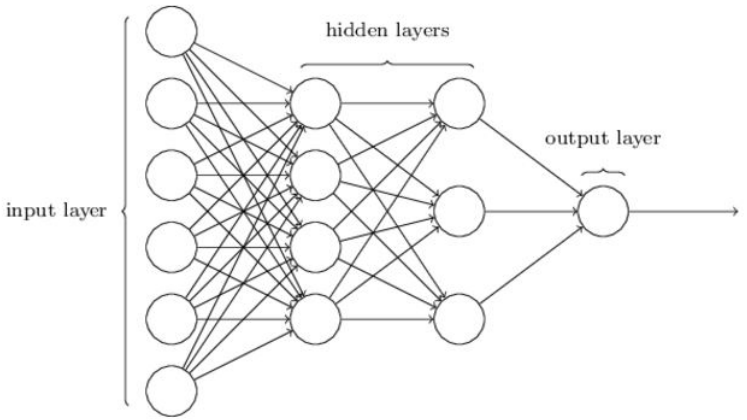
- Need to compute Rk(d), cohomology and $H^i(X, T_A|_X)$ (Cf. Hübsch)

Back to Alg Geo

The Neural Network Approach

- Bijection from  to $\{1, 2, \ldots, 9, 0\}$?
- Take large sample, take a few hundred thousand (e.g. NIST database)





- Data = Training Data ⊔ Validation Data

  Test trained NN on validations data to see accuracy performance

Universal Approximation Theorems

Large Depth Thm: (Cybenko-Hornik) For every continuous function $f : \mathbb{R}^d \to \mathbb{R}^D$, every compact subset $K \subset \mathbb{R}^d$, and every $\epsilon > 0$, there exists a continuous function $f_\epsilon : \mathbb{R}^d \to \mathbb{R}^D$ such that $f_\epsilon = W_2(\sigma(W_1))$, where σ is a fixed continuous function, $W_{1,2}$ affine transformations and composition appropriately defined, so that $\sup_{x \in K} |f(x) - f_\epsilon(x)| < \epsilon$.

Large Width Thm: (Kidger-Lyons) Consider a feed-forward NN with n input neurons, m output neuron and an arbitrary number of hidden layers each with $n + m + 2$ neurons, such that every hidden neuron has activation function φ and every output neuron has activation function the identity. Then, given any vector-valued function f from a compact subset $K \subset \mathbb{R}^m$, and any $\epsilon > 0$, one can find an F, a NN of the above type, so that $|F(x) - f(x)| < \epsilon$ for all $x \in K$.

ReLU Thm: (Hanin) For any Lebesgue-integral function $f : \mathbb{R}^n \to \mathbb{R}$ and any $\epsilon > 0$, there exists a fully connected ReLU NN F with width of all layers less than $n + 4$ such that $\int_{\mathbb{R}^n} |f(x) - F(x)| dx < \epsilon$.

Back to NN@Alg Geo