Universes as Bigdata:

Superstrings, Calabi-Yau Manifolds & Machine-Learning

YANG-HUI HE

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

Merton College, **University of Oxford**

School of Physics, **NanKai University**

M-Theory & Mathematics, NYU Abu Dhabi, 2020

1984: $10 = 4 + 3 \times 2$

- Heterotic string [Gross-Harvey-Martinec-Rohm]: $E_8 \times E_8$ or $SO(32)$, 1984 - 6
- String Phenomenology [Candelas-Horowitz-Strominger-Witten]: 1985
  - E_8 accommodates SM

$$SU(3) \times SU(2) \times U(1) \subset SU(5) \subset SO(10) \subset E_6 \subset E_8$$

  - Standard Solution: $\mathbb{R}^{3,1} \times X$, X Calabi-Yau (Ricci-flat Kähler)
- *mathematicians were independently thinking of the same problem*:
  - Riemann Uniformization Theorem in $\dim_{\mathbb{C}} = 1$: Trichotomy $R < 0, = 0, > 0$
  - Euler, Gauss, Riemann Σ: $\dim_{\mathbb{R}} = 2, i.e., \dim_{\mathbb{C}} = 1$ (in fact Kähler)
  - $\chi(\Sigma) = 2 - 2g(\Sigma) = [c_1(\Sigma)] \cdot [\Sigma] = \frac{1}{2\pi} \int_\Sigma R = \sum_{i=0}^{2} (-1)^i h^i(\Sigma)$

1984: $10 = 4 + 3 \times 2$

- Heterotic string [Gross-Harvey-Martinec-Rohm]: $E_8 \times E_8$ or $SO(32)$, 1984 - 6
- String Phenomenology [Candelas-Horowitz-Strominger-Witten]: 1985
  - E_8 accommodates SM

$$SU(3) \times SU(2) \times U(1) \subset SU(5) \subset SO(10) \subset E_6 \subset E_8$$

  - Standard Solution: $\mathbb{R}^{3,1} \times X$, X Calabi-Yau (Ricci-flat Kähler)
- *mathematicians were independently thinking of the same problem*:
  - Riemann Uniformization Theorem in $\dim_{\mathbb{C}} = 1$: Trichotomy $R < 0, = 0, > 0$
  - Euler, Gauss, Riemann Σ: $\dim_{\mathbb{R}} = 2, i.e., \dim_{\mathbb{C}} = 1$ (in fact Kähler)
  - $\chi(\Sigma) = 2 - 2g(\Sigma) = [c_1(\Sigma)] \cdot [\Sigma] = \frac{1}{2\pi} \int_\Sigma R = \sum_{i=0}^{2} (-1)^i h^i(\Sigma)$

Generalizing a Classic Problem in Mathematics

- 

$g(\Sigma) = 0$	$g(\Sigma) = 1$	$g(\Sigma) > 1$
$\chi(\Sigma) = 2$	$\chi(\Sigma) = 0$	$\chi(\Sigma) < 0$
Spherical	Ricci-Flat	Hyperbolic
$+$ curvature	0 curvature	$-$ curvature
Fano	Calabi-Yau	General Type

- CONJECTURE [E. Calabi, 1954, 1957] / Thm [ST. Yau, 1977-8] M compact Kähler manifold (g, ω) and $([R] = [c_1(M)])_{H^{1,1}(M)}$. Then $\exists!(\tilde{g}, \tilde{\omega})$ such that $([\omega] = [\tilde{\omega}])_{H^2(M;\mathbb{R})}$ and $Ricci(\tilde{\omega}) = R$.
- Strominger & Yau were neighbours at IAS in 1985

Calabi-Yau Spaces as Algebraic Varieties

- $\dim_{\mathbb{C}} = 2$ (T^2 as cubic elliptic curve in $\mathbb{P}^2$); $\dim_{\mathbb{C}} = 2$? K3, quartic in $\mathbb{P}^3$;
- TMH: Homogeneous Eq in $\mathbb{P}^n$, degree $= n+1$ is Calabi-Yau of $\dim_{\mathbb{C}} = n-1$
- We get 5 CY3 immediately (Cyclic Manifolds)
  - Degree 5 in $\mathbb{P}^4$ (The Quintic Q)
  - Two degree 3 in $\mathbb{P}^5$
  - One degree 2 and one degree 4 in $\mathbb{P}^5$
  - Two degree 2 and one degree 3 in $\mathbb{P}^6$
  - Four degree 2 in $\mathbb{P}^7$
- Examples of Complete Intersection CY3

  dim(Ambient space) - #(defining Eq.) = 3 (complete intersection)

An Early Physical Challenge to Algebraic Geometry

- CY3 X, tangent bundle $SU(3) \Rightarrow$
  1. E_6 GUT: commutant $E_8 \to SU(3) \times E_6$, then
  2. Wilson-line/discrete symmetry to break E_6-GUT to some SUSY version of Standard Model (generalize later)
  3. Particle Spectrum: Generation $n_{27} = h^1(X, TX) = h^{2,1}_{\bar\partial}(X)$; Anti-Generation $n_{\overline{27}} = h^1(X, TX^*) = h^{1,1}_{\bar\partial}(X)$
- Net-generation: $\chi = 2(h^{1,1} - h^{2,1}) =$ Euler Number
- 1986 Question: Are there Calabi-Yau threefolds with Euler number ± 6?
- None of our 5 obvious ones ☹

  e.g., Quintic Q in $\mathbb{P}^4$ is CY3 $Q^{h^{1,1},h^{2,1}}_{\chi} = Q^{1,101}_{-200}$ so too may generations (even with quotient $-200 \notin 3\mathbb{Z}$)

An Early Physical Challenge to Algebraic Geometry

- CY3 X, tangent bundle $SU(3) \Rightarrow$
  1. E_6 GUT: commutant $E_8 \to SU(3) \times E_6$, then
  2. Wilson-line/discrete symmetry to break E_6-GUT to some SUSY version of Standard Model (generalize later)
  3. Particle Spectrum: Generation $n_{27} = h^1(X, TX) = h^{2,1}_{\bar\partial}(X)$; Anti-Generation $n_{\overline{27}} = h^1(X, TX^*) = h^{1,1}_{\bar\partial}(X)$
- Net-generation: $\chi = 2(h^{1,1} - h^{2,1}) =$ Euler Number
- 1986 Question: Are there Calabi-Yau threefolds with Euler number ± 6?
- None of our 5 obvious ones ☹

  e.g., Quintic Q in $\mathbb{P}^4$ is CY3 $Q^{h^{1,1},h^{2,1}}_{\chi} = Q^{1,101}_{-200}$ so too may generations (even with quotient $-200 \notin 3\mathbb{Z}$)

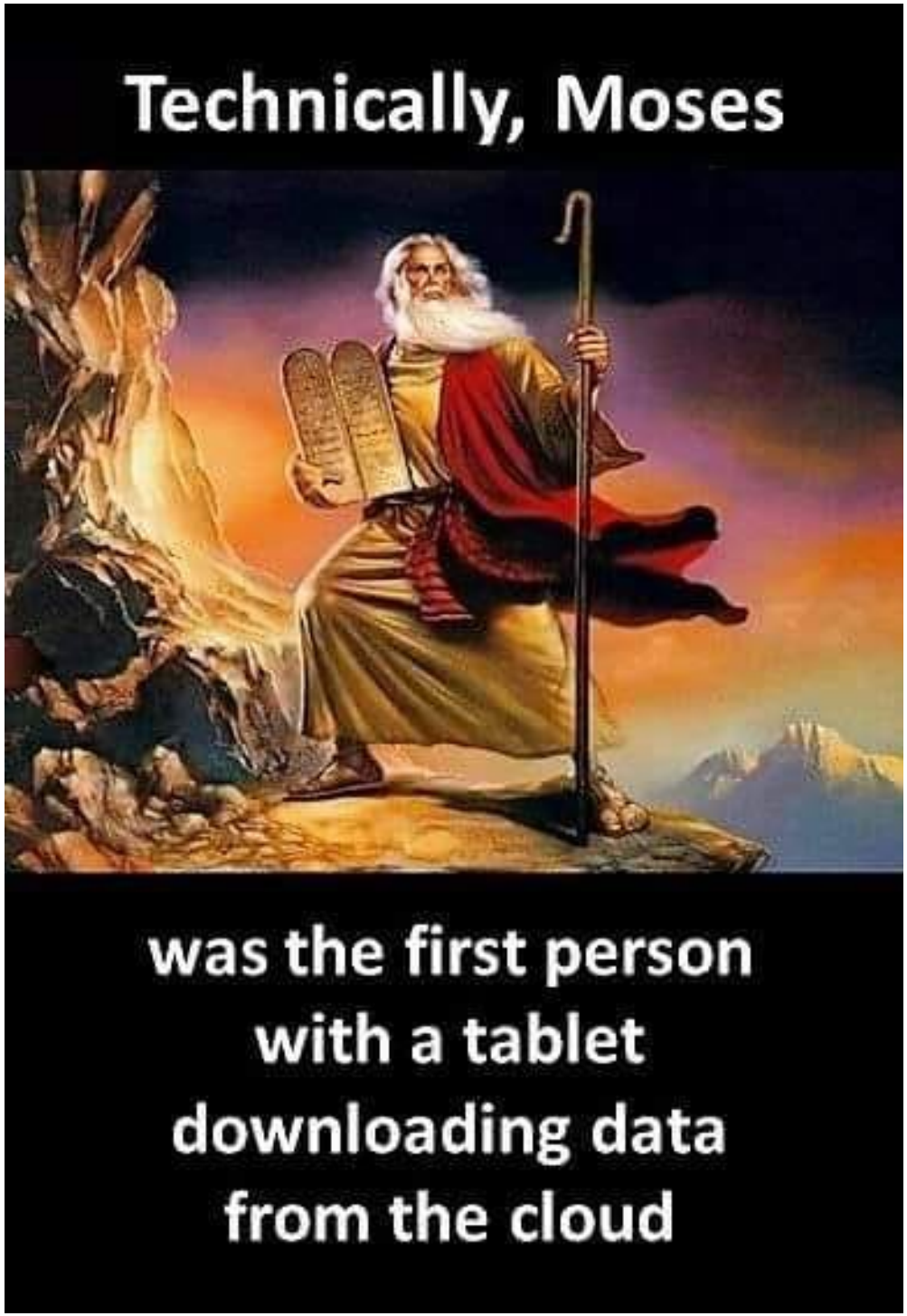
An Early Physical Challenge to Algebraic Geometry

- CY3 X, tangent bundle $SU(3) \Rightarrow$
  1. E_6 GUT: commutant $E_8 \to SU(3) \times E_6$, then
  2. Wilson-line/discrete symmetry to break E_6-GUT to some SUSY version of Standard Model (generalize later)
  3. Particle Spectrum: Generation $n_{27} = h^1(X, TX) = h^{2,1}_{\bar\partial}(X)$; Anti-Generation $n_{\overline{27}} = h^1(X, TX^*) = h^{1,1}_{\bar\partial}(X)$
- Net-generation: $\chi = 2(h^{1,1} - h^{2,1}) =$ Euler Number
- 1986 Question: Are there Calabi-Yau threefolds with Euler number ± 6?
- None of our 5 obvious ones ☹

  e.g., Quintic Q in $\mathbb{P}^4$ is CY3 $Q^{h^{1,1},h^{2,1}}_{\chi} = Q^{1,101}_{-200}$ so too may generations (even with quotient $-200 \notin 3\mathbb{Z}$)

The First Data-sets in Mathematical Physics/Geometry

- [Candelas-A. He-Hübsch-Lutken-Schimmrigk-Berglund] (1986-1990)
  - CICYs (complete intersection CYs) multi-deg polys in products of $\mathbb{CP}^{n_i}$ CICYs
  - Problem: *classify all configuration matrices*; employed the best computers at the time (**CERN supercomputer**); q.v. magnetic tape and dot-matrix printout in Philip's office
  - 7890 matrices, 266 Hodge pairs $(h^{1,1}, h^{2,1})$, 70 Euler $\chi \in [-200, 0]$
- [Candelas-Lynker-Schimmrigk, 1990]
  - Hypersurfaces in Weighted P4
  - 7555 inequivalent 5-vectors w_i, 2780 Hodge pairs, $\chi \in [-960, 960]$
- [Kreuzer-Skarke, mid-1990s - 2000] Reflexive Polytopes
  - Hypersurfaces in (Reflexive, Gorenstein Fano) Toric 4-folds
  - 6-month running time on dual Pentium SGI machine
  - at least 473,800,776, with 30,108 distinct Hodge pairs, $\chi \in [-960, 960]$



The age of data science in mathematical physics/string theory not as recent as you might think

of course, experimentals physics had been decades ahead in data-science/machine-learning

After 40 years of research by mathematicians and physicists

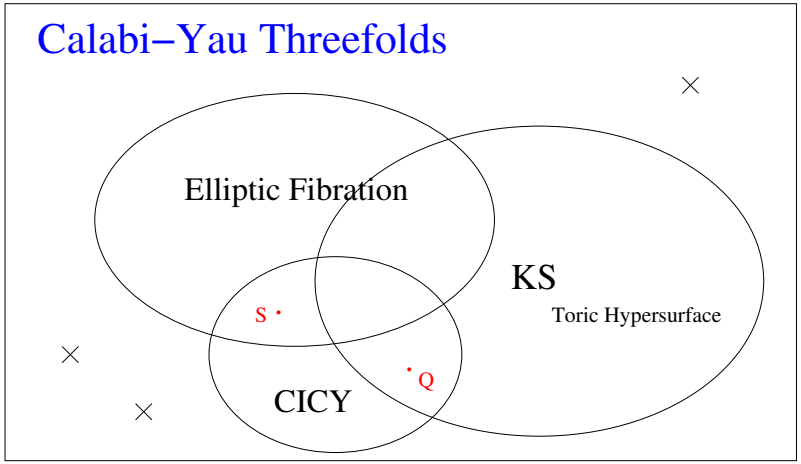
. . . . . .

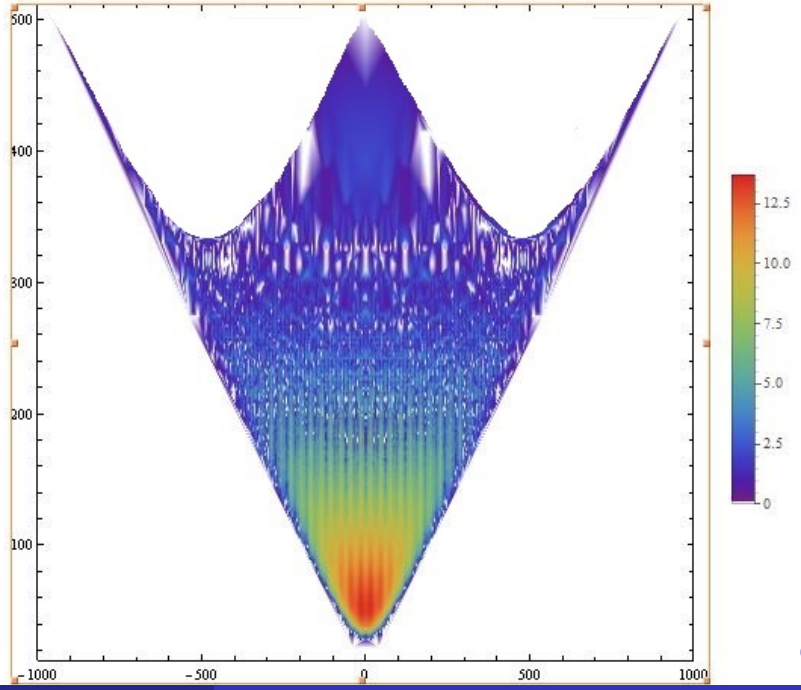
The Compact CY3 Landscape

cf. YHH, *The Calabi-Yau Landscape: from Geometry, to Physics, to Machine-Learning,* 1812.02893, Springer, to appear, 2019/20

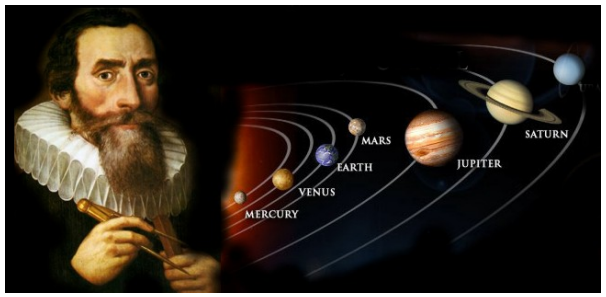
- $\sim 10^{10}$ data-points (and growing, still mined by many international collabs: London/Oxford, Vienna, Northeastern, Jo'burg, Munich, ,. . .)

  a Georgia O'Keefe Plot for Kreuzer-Skarke





The Geometric Origin of our Universe

- Each CY3 (+ bundles, discrete symmetries) X gives a 4-D universe
  - The geometry (algebraic geometry, topology, differential geometry etc.) of X determines the physical properties 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: $\begin{cases} \text{(1) probabilistic/anthropic?} \\ \text{(2) Sui generis/selection rule?} \\ \text{(3) one of multi-verse ?} \end{cases}$

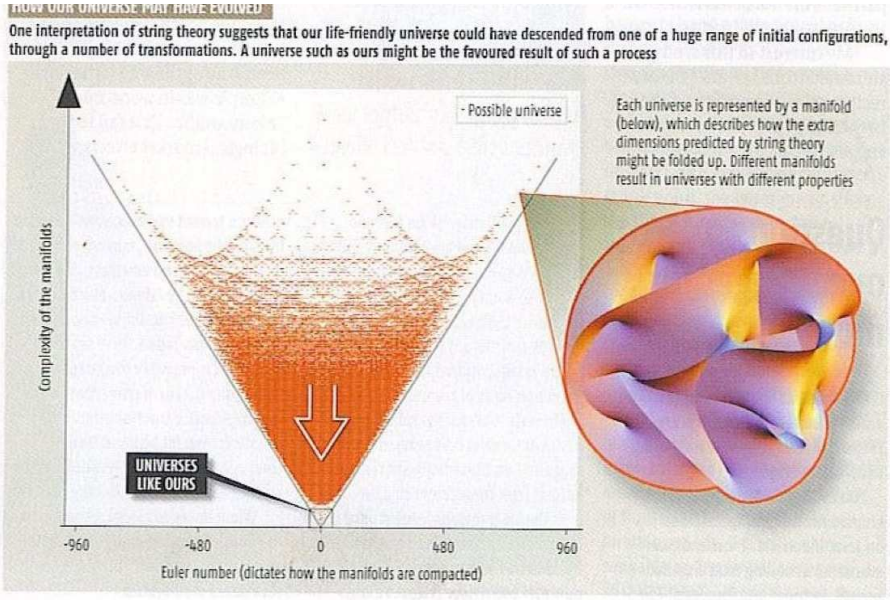
  cf. *Exo-planet/Habitable Zone search*

Triadophilia

Exact (MS)SM Particle Content from String Compactification

- [Braun-YHH-Ovrut-Pantev, Bouchard-Cvetic-Donagi 2005] first exact MSSM
- [Anderson-Gray-YHH-Lukas, 2007-] use alg./comp. algebraic geo & sift
- Anderson-Gray-Lukas-Ovrut-Palti ~ 200 in 10^{10} MSSM Stable Sum of Line Bundles over CICYs (Oxford-Penn-Virginia 2012-)

  Constantin-YHH-Lukas '19: 10^{23} exact MSSMs (by extrapolation on above set)?



A Special Corner

[New Scientist, Jan, 5, 2008 feature]

P. Candelas, X. de la Ossa, YHH, and B. Szendroi

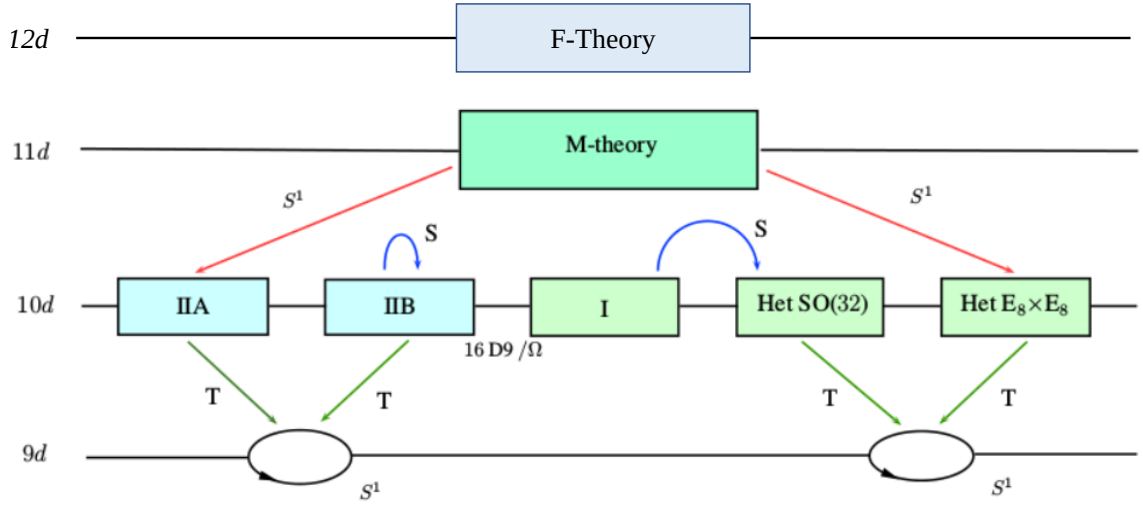
"Triadophilia: A Special Corner of the Landscape" ATMP, 2008

The Landscape Explosion

meanwhile . . . LANDSCAPE grew rapidly with

- D-branes Polchinski 1995
- M-Theory/G_2 Witten, 1995
- F-Theory/4-folds Katz-Morrison-Vafa, 1996
- AdS/CFT Maldacena 1998 Alg Geo of AdS/CFT
- Flux-compactification Kachru-Kallosh-Linde-Trivedi, 2003, . . .

The Vacuum Degeneracy Problem

- 

More solutions related by dualities

Fig. modified from `https://www.physics.uu.se/`

- String theory trades one hard-problem [quantization of gravity] by another [looking for the right compactification] (in many ways a richer and more interesting problem)
- KKLT 2003, Douglas, Denef 2005 - 6 at least 10^{500} possibilities

SUMMARY: Algorithms and Datasets in String Theory

- Growing databases and algorithms (many motivated by string theory): e.g., Singular, Macaulay2, GAP (10^7 finite groups), SAGE, Bertini, grdb, etc; "Periodic table of shapes Project" classify Fanos, KS $\sim 10^9$ CYs, Cremona's $\sim 10^7$ elliptic curves, ...
- Archetypical Problems
  - Classify configurations (typically integer matrices: polyotope, adjacency, ...)
  - Compute geometrical quantity algorithmically
    - toric $\rightsquigarrow$ combinatorics;
    - quotient singularities $\rightsquigarrow$ rep. finite groups;
    - generically $\rightsquigarrow$ ideals in polynomial rings;
    - Numerical geometry (homotopy continuation);
    - Cohomolgy (spectral sequences, Adjunction, Euler sequences)

Where we stand . . .

The Good Last 10-15 years: several international groups have bitten the bullet

Oxford, London, Vienna, Blacksburg, Boston, Johannesburg, Munich, . . .

computed many geometrical/physical quantities and compiled them into various databases **Landscape Data** ($10^{9\sim10}$ entries typically)

The Bad Generic computation HARD: dual cone algorithm (exponential), triangulation (exponential), Gröbner basis (double-exponential) . . . e.g., how to construct stable bundles over the $\gg 473$ million KS CY3? Sifting through for SM computationally impossible . . .

The ??? **Borrow new techniques from "Big Data" revolution**

A Wild Question

- Typical Problem in String Theory/Algebraic Geometry:

$$\overset{INPUT}{\boxed{\text{integer tensor}}} \longrightarrow \overset{OUTPUT}{\boxed{\text{integer}}}$$

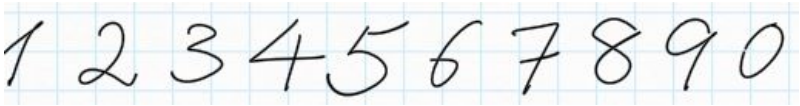
- Q: Can (classes of problems in computational) Algebraic Geometry be "learned" by AI ? , i.e., can we "machine-learn the landscape?"
- [YHH 1706.02714] Deep-Learning the Landscape, PLB 774, 2017:

  Experimentally, it seems to be the case for many situations
- 2017

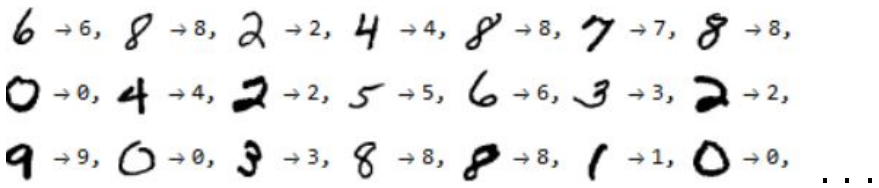
  YHH (1706.02714), Seong-Krefl (1706.03346), Ruehle (1706.07024), Carifio-Halverson-Krioukov-Nelson (1707.00655) Progress in String Theory

A Prototypical Question

- Hand-writing Recognition, e.g., my 0 to 9 is different from yours:



- How to set up a bijection that takes these to $\{1, 2, \ldots, 9, 0\}$? Find a clever Morse function? Compute persistent homology? Find topological invariants? <u>ALL are inefficient and too sensitive to variation.</u>
- What does your iPhone/tablet do? What does Google do? Machine-Learn
  - Take large sample, take a few hundred thousand (e.g. NIST database)

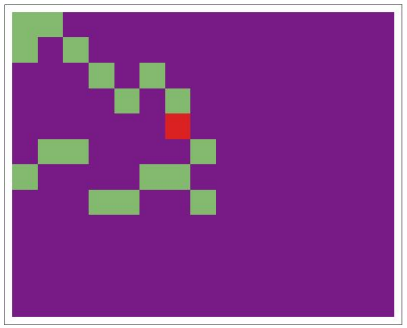


Supervised ML in 1 min

NN Doesn't Care/Know about Algebraic Geometry

- Hodge Number of a Complete Intersection CY is the association rule, e.g.

$$X = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}, \qquad h^{1,1}(X) = 8 \quad \rightsquigarrow$$



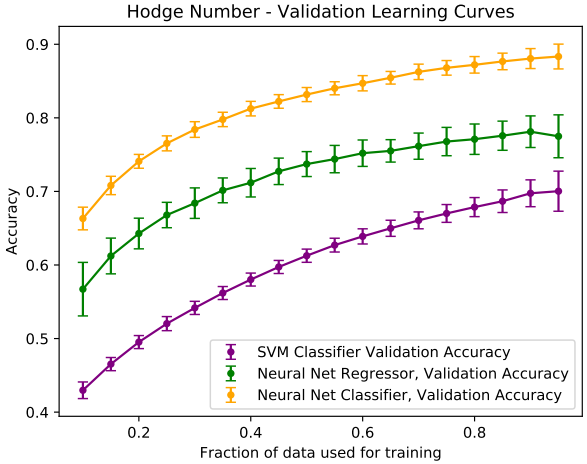
$$\longrightarrow 8$$

CICY is 12×15 integer matrix with entries $\in [0, 5]$ is simply represented as a 12×15 pixel image of 6 colours Proper Way

- Cross-Validation:
  - Take samples of $X \to h^{1,1}$
  - train a NN, or SVM
  - Validation on *unseen* $X \to h^{1,1}$

Deep-Learning Algebraic Geometry

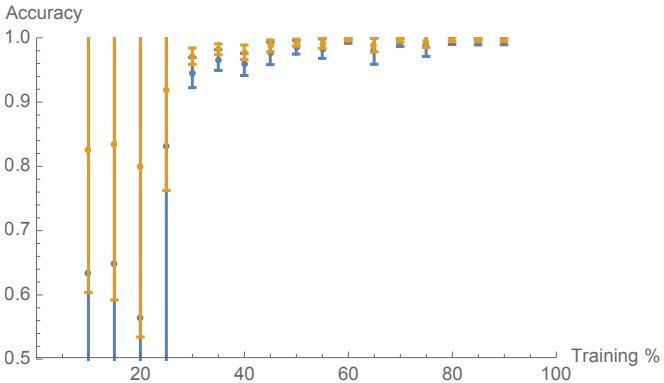
- YHH '17 Bull-YHH-Jejjala-Mishra '18:



Learning Hodge Number $h^{1,1} \in [0, 19]$ so can set up 20-channel NN classifer, regressor, as well as SVM,

bypass exact sequences

- YHH-SJ Lee'19: Distinguishing Elliptic Fibrations in CY3



bypass Oguiso-Kollar-Wilson Theorem/Conjecture

learning curves for precision and Matthews ϕ

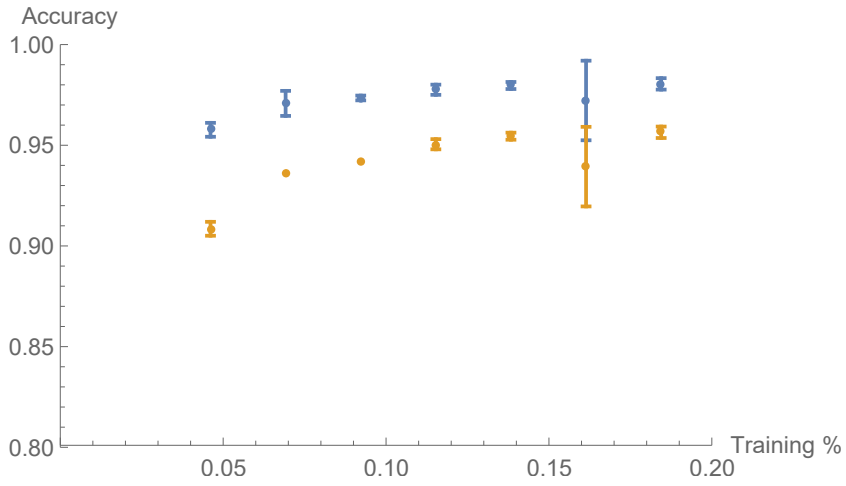
More Success Stories in Algebraic Geometry

- Ruehle '17: genetic algorithm for bundle cohomology
- Brodie-Constantin-Lukas '19: EXACT formulae for line-bundle coho / complex surfaces Interpolation vs Extrapolation $\rightsquigarrow$ Conjecture Formulation
- Ashmore-YHH-Ovrut '19: ML Calabi-Yau metric:
  - No known explicit Ricci-Flat Kähler metric (except T^n) (Yau's '86 proof non-constructive); Donaldson ['01-05] relatively fast method of *numerical* (balanced) such metrics
  - ML improves it to 10-100 times faster with equal/better accuracy (NB. *checking* Ricci-flat is easy)
- RMK: Alg Geo / $\mathbb{C}$ amenable to ML: core computations (Grobner bases, syzygies, long exact sequences, etc) $\sim$ integer (co-)kernel of matrices.

Why stop at string/geometry?

[YHH-MH. Kim 1905.02263] Learning Algebraic Structures

- When is a Latin Square (Sudoku) the Cayley (multiplication) table of a finite group? (rmk: there is a known quadrangle-thm to test this) NN/SVM find to 94.9% ($\phi = 0.90$) at 25-75 cross-validation.
- Can one look at the Cayley table and recognize a finite simple group?



- bypass Sylow and Noether Thm
- rmk: can do it via character-table T, but getting T not trivial
- SVM: space of finite-groups (point-cloud of Cayley tables), ?$\exists$ hypersurface separating simple/non-simple?

Further Explorations

- Alessandretti,Baronchelli,YHH 1911.02008 ML/TDA@Birch-Swinnerton-Dyer

  BSD: L-function $L(s,\mathcal{E})$ of elliptic curve $\mathcal{E}$ has $L(s\to 1,\mathcal{E})$ given in terms of precise quantities: rank r, torsion T, period Ω, Tate-Shaferevich group $Ш$, conductor N, regulator R, Tamagawa number c

  $Ш$ and Ω good with regression and boosted decision trees: RMS < 0.1

- a Reprobate: predicting primes

  YHH 1706.02714: tried supervised ML on $2\to 3,\ 2,3\to 5,\ 2,3,5\to 7$

  tried fixed window of $(\text{yes/no})_{1,2,\ldots,k}$ to $(\text{yes/no})_{k+1}$, no breaking banks yet.

Why stop at the mathematics/physics?

[YHH-Jejjala-Nelson] "hep-th" 1807.00735

- Word2Vec: [Mikolov et al., '13] NN which maps words in sentences to a vector space **by context** (much better than word-frequency, quickly adopted by Google); maximize (partition function) over all words with sliding window ($W_{1,2}$ weights of 2 layers, C_α window size, D # windows)

$$Z(W_1, W_2) := \frac{1}{|D|} \sum_{\alpha=1}^{|D|} \log \prod_{c=1}^{C_\alpha} \frac{\exp([\vec{x}_c]^T \cdot W_1 \cdot W_2)}{\sum_{j=1}^{V} \exp([\vec{x}_c]^T \cdot W_1 \cdot W_2)}$$

- We downloaded all $\sim 10^6$ titles of hep-th, hep-ph, gr-qc, math-ph, hep-lat from ArXiv since the beginning (1989) till end of 2017 Word Cloud

  (rmk: Ginzparg has been doing a version of linguistic ML on ArXiv)

  (rmk: abs and full texts in future)

Subfields on ArXiv has own linguistic particulars

- Linear Syntactical Identities

  bosonic + string-theory = open-string

  holography + quantum + string + ads = extremal-black-hole

  string-theory + calabi-yau = m-theory + g2

  space + black-hole = geometry + gravity . . .

- binary classification (Word2Vec + SVM) of formal (hep-th, math-ph, gr-qc) vs phenomenological (hep-ph, hep-lat) : 87.1% accuracy (5-fold classification 65.1% accuracy). ArXiv classifications

- Cf. **Tshitoyan et al.**, "Unsupervised word embeddings capture latent knowledge from materials science literature", **Nature** July, 2019: 3.3. million materials-science abstracts; uncovers structure of periodic table, predicts discoveries of new thermoelectric materials years in advance, and suggests as-yet unknown materials

Summary and Outlook

PHYSICS

- Use AI (Neural Networks, SVMs, Regressor . . .) as
  1. Classifier deep-learn and categorize landscape data
  2. Predictor estimate results beyond computational power

MATHS

- Not solving NP-hard problems, but stochastically bypassing the expensive steps of long sequence-chasing, Gröbner bases, dual cones/combinatorics **how is AI doing maths more efficiently without knowing any maths?**
- Hierarchy of Difficulty ML struggles with:

  numerical $<$ **algebraic geometry over** $\mathbb{C}$ $<$ **combinatorics/algebra** $<$ **number theory**

شكراً

- Boris Zilber [Merton Professor of Logic, Oxford]: "you've managed syntax without semantics. . ."
-

Alpha Go	$\to$	Alpha Zero
ML	$\to$	Voevodsky's Dream; Automated Thm Pf

- cf. Renner et al., PRL/Nature News, Nov, 2019: ML (*SciNet*, *autoencoder*) finds heliocentrism from Mars positions alone.

Sophia (Hanson Robotics, HK)

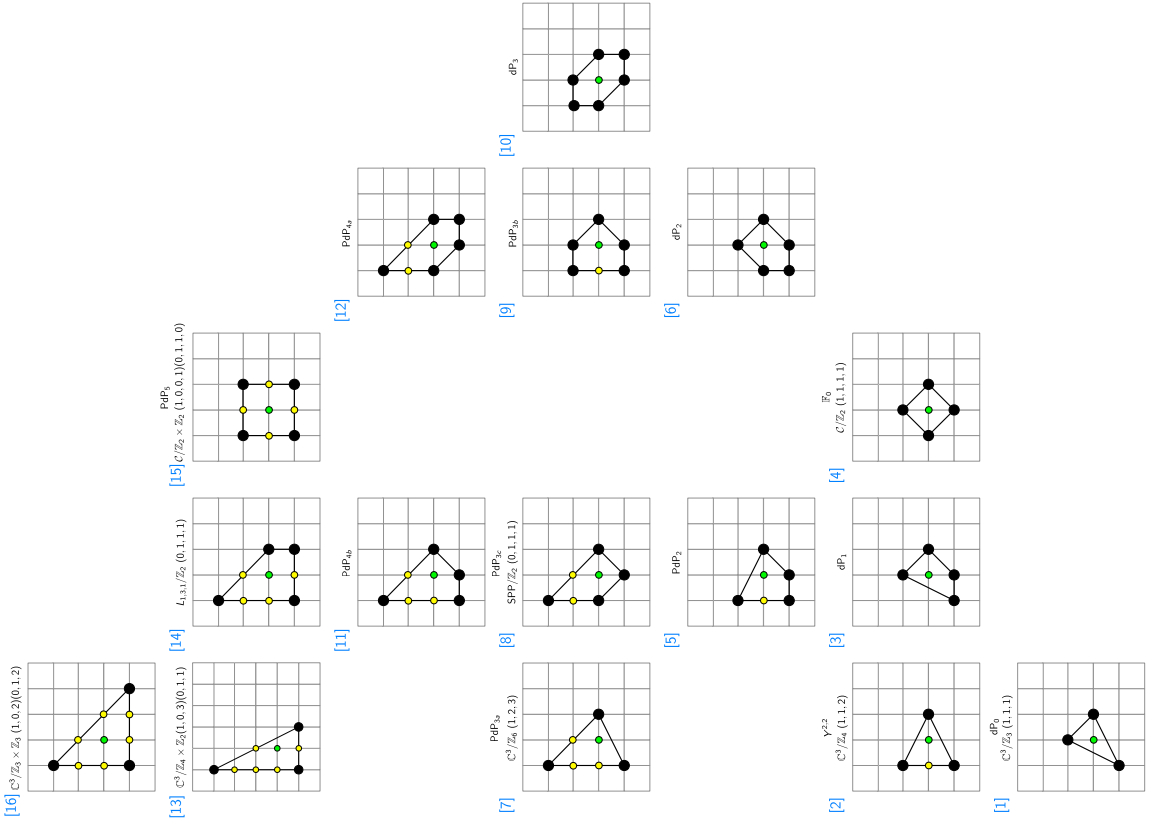
1st non-human citizen (2017, Saudi Arabia)

1st non-human with UN title (2017)

1st String Data Conference (2017)

16 Reflexive Polygons

Back to Reflexives



classify convex lattice polytopes with single interior point and all faces are distance 1 therefrom (up to $SL(n;\mathbb{Z})$)

Kreuzer-Skarke: 4319 reflexive polyhedra, 473,800,776 reflexive 4-polytopes,

Skarke: next number is at least 185,269,499,015.

Progress in String Theory Back to ML

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

CICYs

$$M = \left[\begin{array}{c|cccc} n_1 & q_1^1 & q_1^2 & \cdots & q_1^K \\ n_2 & q_2^1 & q_2^2 & \cdots & q_2^K \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ n_m & q_m^1 & q_m^2 & \cdots & q_m^K \end{array}\right]_{m\times K}$$

- Complete Intersection Calabi-Yau (CICY) 3-folds
- K eqns of multi-degree $q_j^i \in \mathbb{Z}_{\geq 0}$ embedded in $\mathbb{P}^{n_1} \times \ldots \times \mathbb{P}^{n_m}$
- $c_1(X) = 0 \rightsquigarrow \sum_{j=1}^{K} q_r^j = n_r + 1$
- M^T also CICY

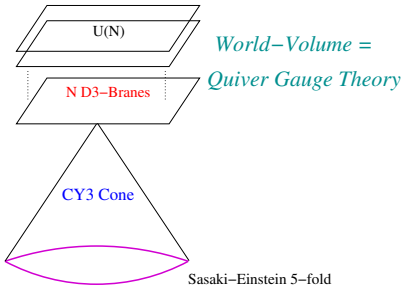
- The Quintic $Q = [4|5]^{1,101}_{-200}$ (or simply $[5]$);
- CICYs Central to string pheno in the 1st decade [Distler, Greene, Ross, et al.]

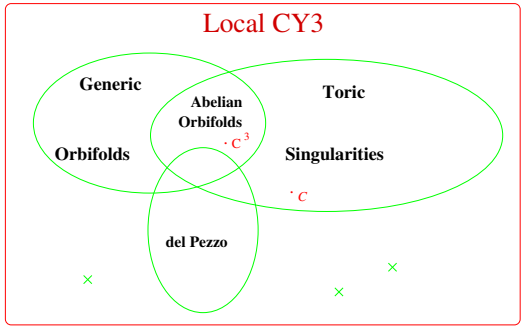
  E_6 GUTS unfavoured; Many exotics: e.g. 6 entire anti-generations

Back to CICYs

AdS/CFT as a Quiver Rep/Moduli Variety Corr.

a 20-year prog. joint with **A. Hanany**, S. Franco, B. Feng, et al.





D-Brane Gauge Theory

(SCFT encoded as quiver)

$\longleftrightarrow$

Vacuum Space as affine Variety

- ($\mathcal{N}=4$ SYM) $\left(\quad , W = \mathrm{Tr}([x,y],z) \right) \longleftrightarrow \mathbb{C}^3 = \mathrm{Cone}(S^5)$ [Maldacena]
- THM [(P) Feng, Franco, Hanany, YHH, Kennaway, Martelli, Mekareeya, Seong, Sparks, Vafa, Vegh, Yamazaki, Zaffaroni ...

  (M) R. Böckland, N. Broomhead, A. Craw, A. King, G. Musiker, K. Ueda ...] (coherent component of) representation variety of a quiver is toric CY3 iff quiver + superpotential graph dual to a bipartite graph on T^2 Back to Landscape

  combinatorial data/lattice polytopes $\longleftrightarrow$ gauge thy data as quivers/graphs

A Single Neuron: The Perceptron

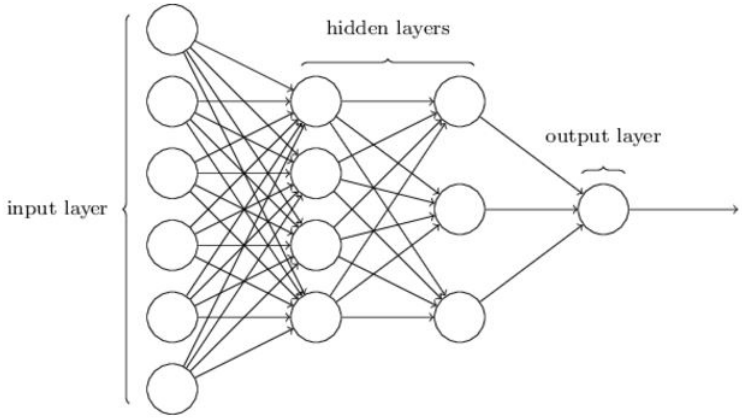
- began in 1957 (!!) in early AI experiments (using CdS photo-cells)
- DEF: Imitates a neuron: activates upon certain inputs, so define
  - Activation Function $f(z_i)$ for input tensor z_i for some multi-index i;
  - consider: $f(w_i z_i + b)$ with w_i weights and b bias/off-set;
  - typically, $f(z)$ is sigmoid, Tanh, etc.
- Given training data: $D = \{(x_i^{(j)}, d^{(j)}\}$ with input x_i and **known output** $d^{(j)}$, minimize

$$SD = \sum_j \left(f(\sum_i w_i x_i^{(j)} + b) - d^{(j)} \right)^2$$

  to find optimal w_i and $b \rightsquigarrow$ "learning", then check against Validation Data
- Essentially (non-linear) regression

The Neural Network: network of neurons $\rightsquigarrow$ the "brain"

- DEF: a connected graph, each node is a perceptron (*Implemented on Mathematica 11.1 + / TensorFlow-Keras on Python*)
  1. adjustable weights/bias;
  2. distinguished nodes: 1 set for input and 1 for output;
  3. iterated training rounds.



Simple case: forward directed only, called multilayer perceptron

- others: e.g., decision trees, support-vector machines (SVM), etc
- Essentially how brain learns complex tasks; **apply to our Landscape Data**

Back to Landscape

Computing Hodge Numbers: Sketch

- 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:

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

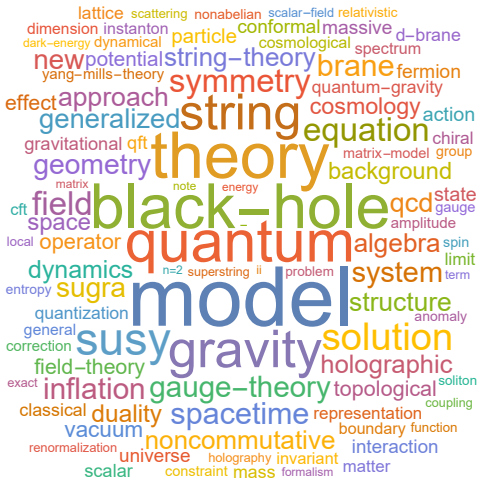
- Induces long exact sequence in cohomology:

$$\begin{array}{ccccccccc}
0 & \to & \cancelto{0}{H^0(X, T_X)} & \to & H^0(X, T_A|_X) & \to & H^0(X, N_X) & \to \\
 & \to & \boxed{H^1(X, T_X)} & \xrightarrow{d} & H^1(X, T_A|_X) & \to & H^1(X, N_X) & \to \\
 & \to & H^2(X, T_X) & \to & \dots & & &
\end{array}$$

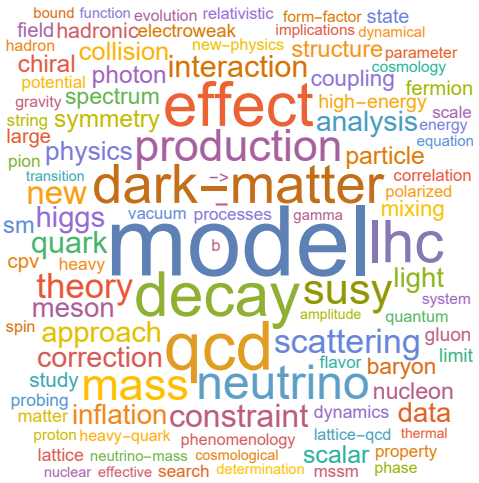
- Need to compute Rk(d), cohomology and $H^i(X, T_A|_X)$ (Cf. Hübsch)

Back to ML

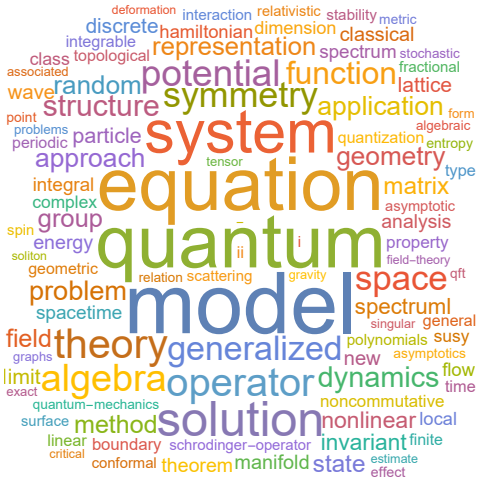
ArXiv Word-Clouds



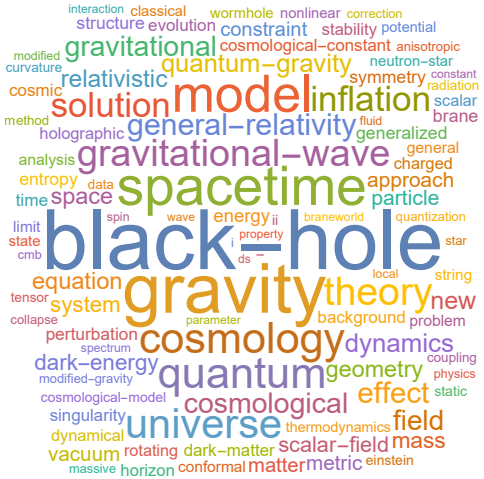
hep-th



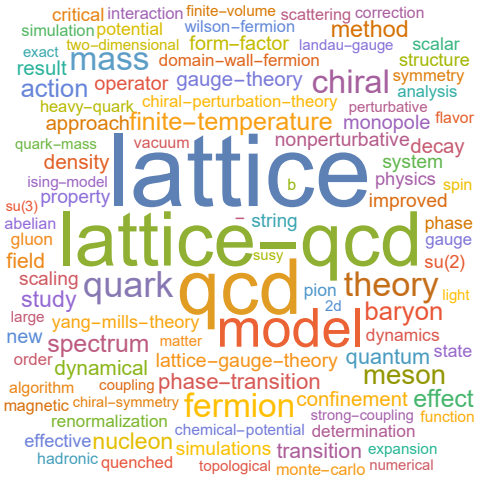
hep-ph



math-ph



hep-th



hep-ph

Back to Word2Vec

Classifying Titles

Compare, + non-physics sections, non-science (Times), pseudo-science (viXra)

Actual \ Word2Vec + SVM	1	2	3	4	5
1	40.2	6.5	8.7	24.0	20.6
2	7.8	65.8	12.9	9.1	4.4
3	7.5	11.3	72.4	1.5	7.4
4	12.4	4.4	1.0	72.1	10.2
5	10.9	2.2	4.0	7.8	75.1

1 : hep-th
2 : hep-ph
3 : hep-lat
4 : gr-qc
5 : math-ph

Actual \ NN	1	2	3	4	5	6	7	8	9	10
viXra-hep	11.5	47.4	6.8	13.	11.	4.5	0.2	0.3	2.2	3.1
viXra-qgst	13.3	14.5	1.5	54.	8.4	1.8	0.1	1.1	2.8	3.

6: cond-mat, 7: q-fin, 8: stat, 9: q-bio, 10: Times of India Back to Main