Hidden structures in the landscape of heterotic line bundle models

Hajime Otsuka (KEK)

References :

H. O. and K. Takemoto, JHEP **05** (2020) 047, arXiv : 2003.11880
H. O. work in progress

Outline

1. **Why machine learning (ML) in string theory ?**
   - Deal with topological data and computational complexity

2. **Brief review of ML and autoencoder**

3. **Application of ML to string model building**

   H. Otsuka and K. Takemoto, JHEP **05** (2020) 047, arXiv : 2003.11880.

4. **Conclusion**

Superstring theory

Candidate of

- Quantum Gravity
- Unified theory of gauge and gravitational interactions

- (Perturbative) superstring theory predicts the extra 6D space

$$10=4+6$$

- 6D compactification → Degrees of freedom
  --- fluxes (VEVs of gauge fields)
  --- branes (wrapping sub-manifolds)

<u>Huge number of 4D stable vacua (landscape)</u>

Candidates of 6D spaces

- If SUSY is preserved in 4D,

$$\delta_{\mathrm{SUSY}}(\mathrm{Gravitino}) = \nabla_M \epsilon = 0$$

$M = 0,1,\cdots,9$

$\Downarrow$

$$R_{\mu\nu} = 0 \text{ and } R_{ij} = 0$$

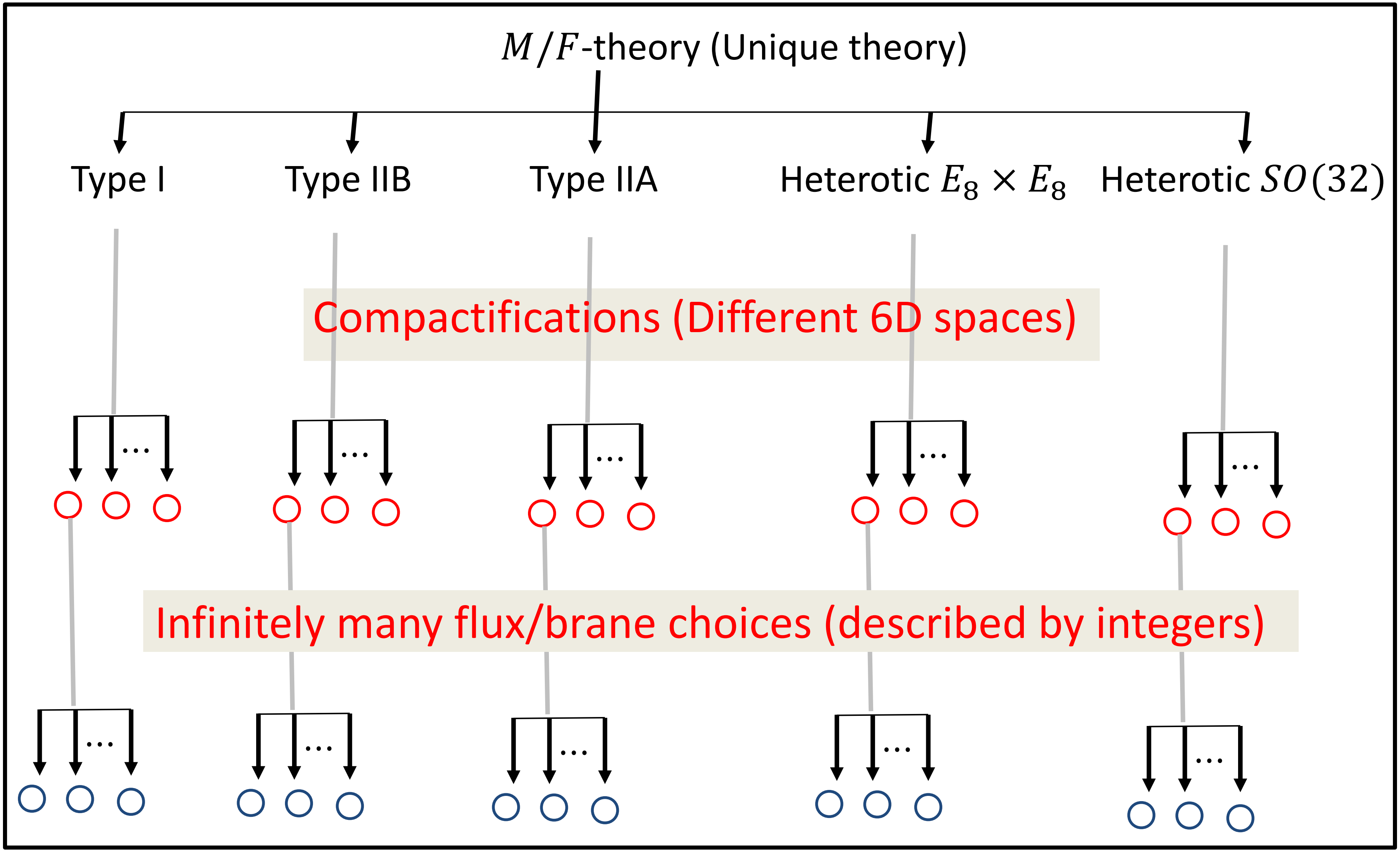
$\mu,\nu = 0,1,2,3$
$i,j = 4,5,\cdots,9$

- 10D(=4+6) spacetime :

4D spacetime (if maximally symmetric) = Minkowski

6D space = Ricci-flat Kahler manifold (Calabi-Yau manifold)

So far, we know $O(10^8)$ 6D CYs
(Unknown whether there are infinitely many CYs)



- $O(10^{500})$ Type IIB flux vacua *Ashok-Douglas ('04)*
- $O(10^{272,000})$ F-theory flux vacua *Taylor-Wang ('15)*
- $O(10^{662})$ MSSM-like models in Heterotic on CYs *Constantin-He-Lukas ('18)*

Computational complexity

- Big data (Topological data) is extremely constrained by

  - SUSY conditions for fluxes/branes
  - Charge cancellation conditions for brane charges
  - Phenomenological constraints
    - (i) SM gauge group : $SU(3) \times SU(2) \times U(1)$
    - (ii) 3 generations of quarks/leptons, Realistic Yukawa couplings, ...

Motivation for Machine Learning :

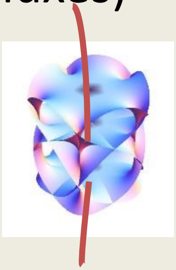
ML can deal with

- Gigantic number of topological data (CYs, flux/brane,...)
- Computationally complexity

How to apply ML ?

- In June 2017, 4 groups proposed the ML applications to string th.
  He, Krefl-Seung, Ruehle, Carifio-Halverson-Knoukov-Nelson

Topological data of CY
(geometrical quantities, fluxes)



⇨

- SM gauge group
- <u>3 generations of quarks/leptons</u>
- Yukawa and gauge couplings

Question :

Which topological data is important for 3 generations of quarks/leptons

<u>ML would reveal the hidden structure in the string landscape</u>

Outline

1. **Why machine learning (ML) in string theory ?**
   - Deal with topological data and computational complexity

2. **Brief review of ML and autoencoder**

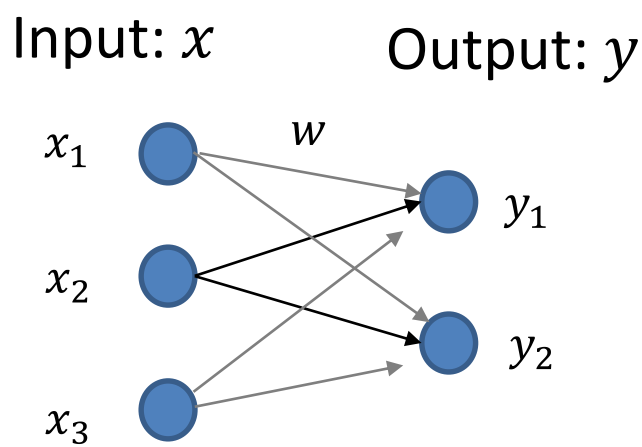
3. **Application of ML to string model building**

   H. Otsuka and K. Takemoto, JHEP **05** (2020) 047, arXiv : 2003.11880.

4. **Conclusion**

▪ (Fully connected) neural networks

Layout :



Output data : $y_i = h(\underbrace{w_{ij} \cdot x_j + b_i}_{\text{Linear}})$

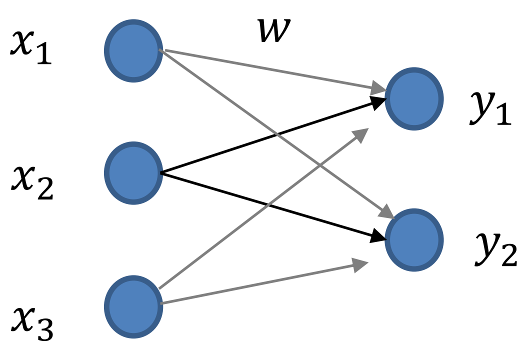
w_{ij}: Weight (linear map)
b_i : Bias

h : Activation function (Non-linear func.)
(Analogous to activate the neuron)

▪ (Fully connected) neural networks

Layout :

Input: x Output: y

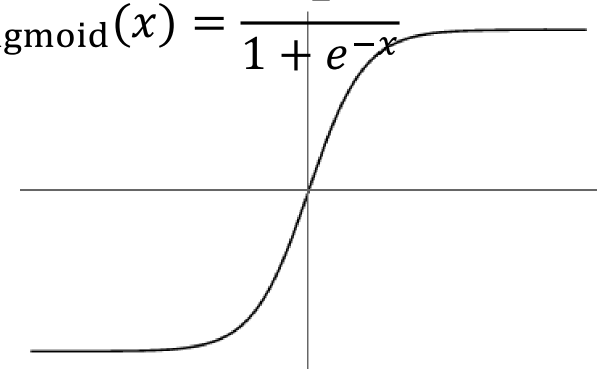


Output data : $y_i = \color{red}{h}(\underbrace{w_{ij} \cdot x_j + b_i}_{\text{Linear}})$

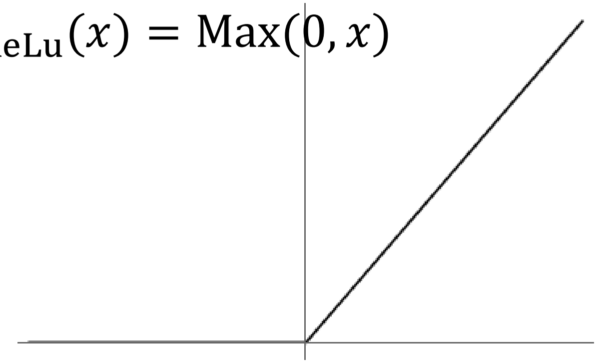
w_{ij}: Weight (linear map)
b_i : Bias

h : Activation function (Non-linear func.)
(Analogous to activate the neuron)

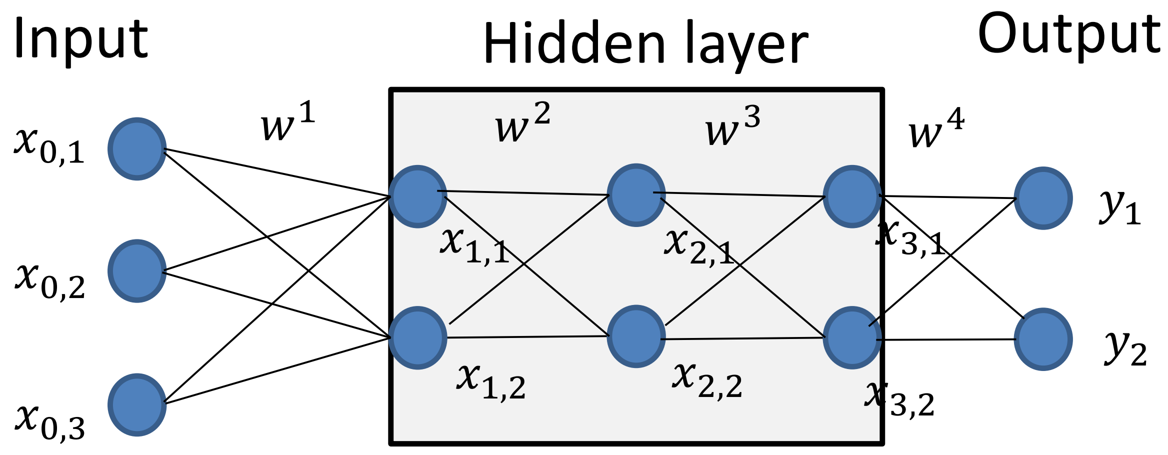
$$h_{\text{sigmoid}}(x) = \frac{1}{1+e^{-x}}$$



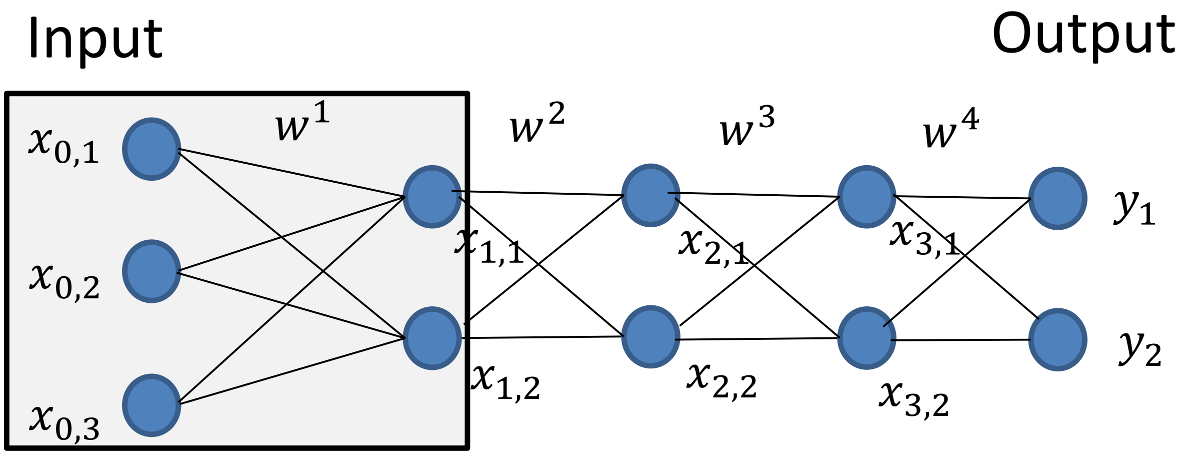
$$h_{\text{ReLu}}(x) = \text{Max}(0, x)$$



- (Fully connected) neural networks



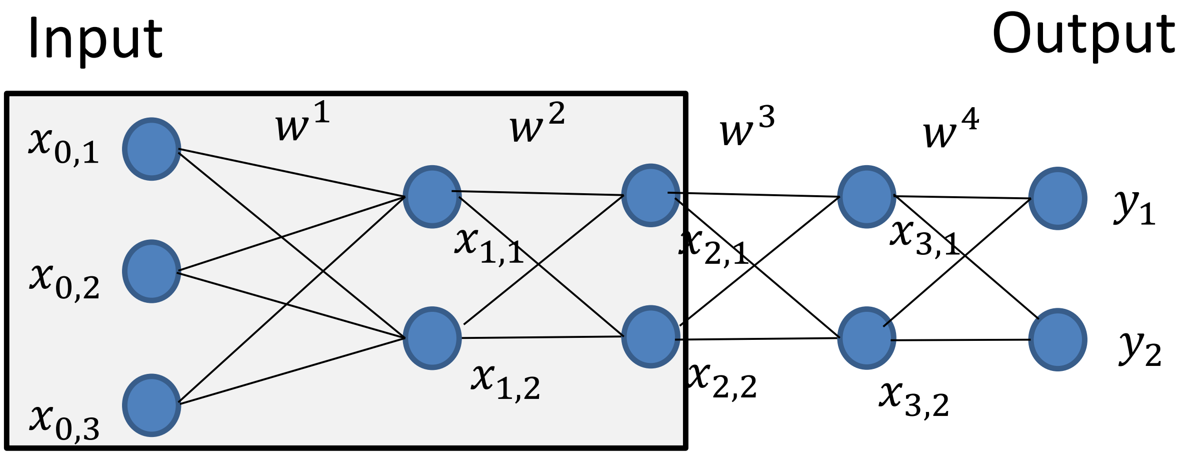
- (Fully connected) neural networks



- Output : $\vec{y} =$ $h_1\left(w^1 \cdot \vec{x}_0 + \vec{b}^1\right)$

w^1: Weights
$\vec{b}^1$: Bias
h_1 : Activation function

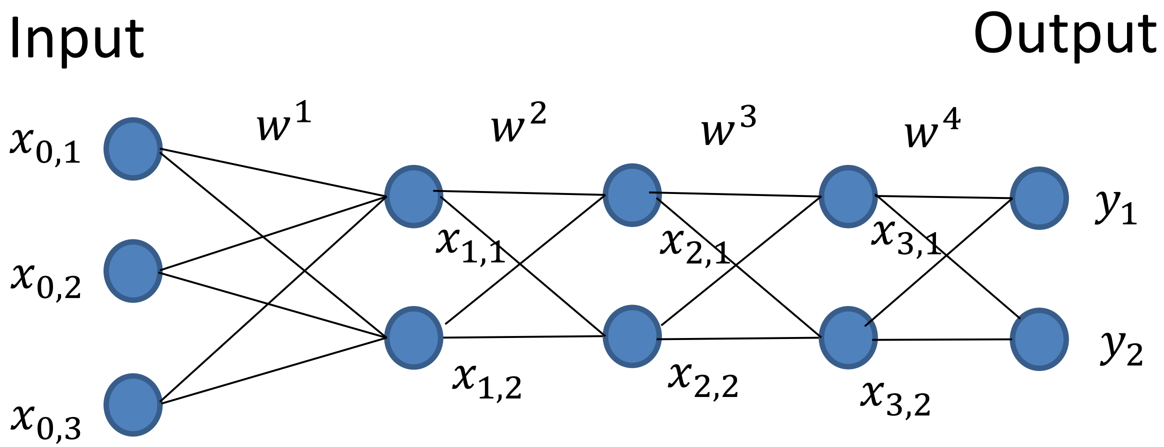
- (Fully connected) neural networks



- Output : $\vec{y} = h_2\left(w^2 h_1\left(w^1 \cdot \vec{x}_0 + \vec{b}^1\right) + \vec{b}^2\right)$

w^n: Weights
$\vec{b}^n$: Bias
h_n : Activation function

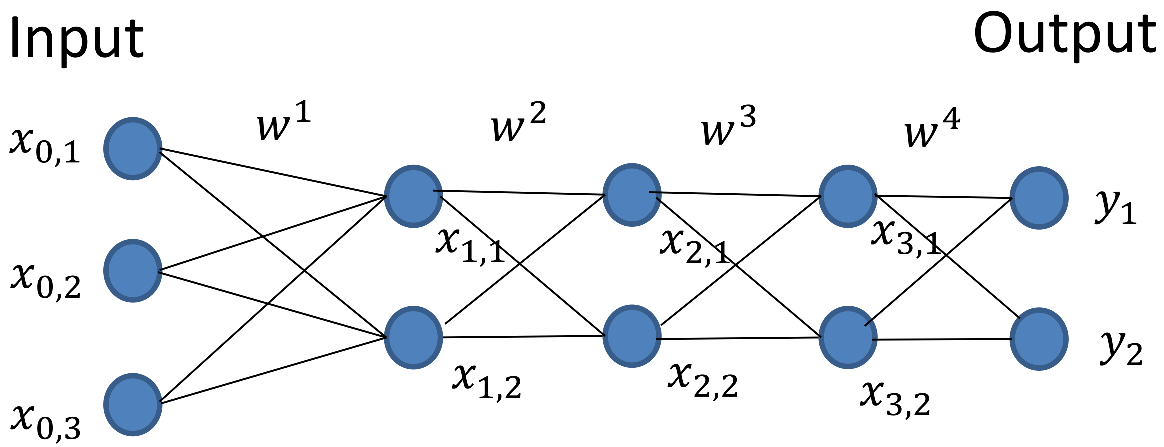
- (Fully connected) neural networks



- Output : $\vec{y} = h_4(w^4 h_3(\ldots h_2\left(w^2 h_1\left(w^1 \cdot \vec{x}_0 + \vec{b}^1\right) + \vec{b}^2\right) \cdots + \vec{b}^4)$

w^n: Weights
$\vec{b}^n$: Bias
h_n : Activation function

▪ (Fully connected) neural networks



- Output : $\vec{y} = h_4(w^4 h_3(\ldots h_2\left(w^2 h_1\left(w^1 \cdot \vec{x}_0 + \vec{b}^1\right) + \vec{b}^2\right)\cdots + \vec{b}^4)$

w^n: Weights
$\vec{b}^n$: Bias
h_n : Activation function

- "Learning" (for supervised ML)
  —Training data $(\vec{x}_d, \vec{y}_d)$ $(d = 1,2, \ldots N_d)$
  —Find parameters $\{\theta_a\}$=$\{w_{ij}^n, b_i^n\}$ by minimizing the error function

E.g., $$\mathrm{err}(\theta) = \frac{1}{N_d}\sum_d |\vec{y}_d - \vec{y}_{NN}(\vec{x}_d, \theta_a)\,|^2$$

Machine learning in string theory

- Basically, three types of ML have been used so far.

See for a review, [F. Ruehle '20; Y-Hui He '20; Tanaka-Tomiya-Hashimoto '20]

1. Bypass computations (Supervised ML)

— Deep neural networks, Support vector machines

[Wang-Zhang '18; Bull-He-Jejjala-Mishra '18; Klaewer-Schlechter '18; He '18; Jejjala-Kar-Parrikar '19; He-Lee '19]

2. Search the landscape (Semi-supervised ML)

— Reinforcement Learning

[Carifio-Halverson-Krioukov-Nelson '17; Altman-Carifio-Halverson-Nelson '18,….]

3. Vacuum structure (Unsupervised ML)

— Clustering, Feature extraction, Topological data analysis

[Cole-Shiu, '17, '18; Mutter-Parr-Vaudrevange, '18; Otsuka-Takemoto '20; Deen-He-Lee-Lukas '20],…

Machine learning in string theory

- Basically, three types of ML have been used so far.

See for a review, [F. Ruehle '20; Y-Hui He '20; Tanaka-Tomiya-Hashimoto '20]

1. Bypass computations (Supervised ML)

— Deep neural networks, Support vector machines

[Wang-Zhang '18; Bull-He-Jejjala-Mishra '18; Klaewer-Schlechter '18; He '18; Jejjala-Kar-Parrikar '19; He-Lee '19]

2. Search the landscape (Semi-supervised ML)

— Reinforcement Learning

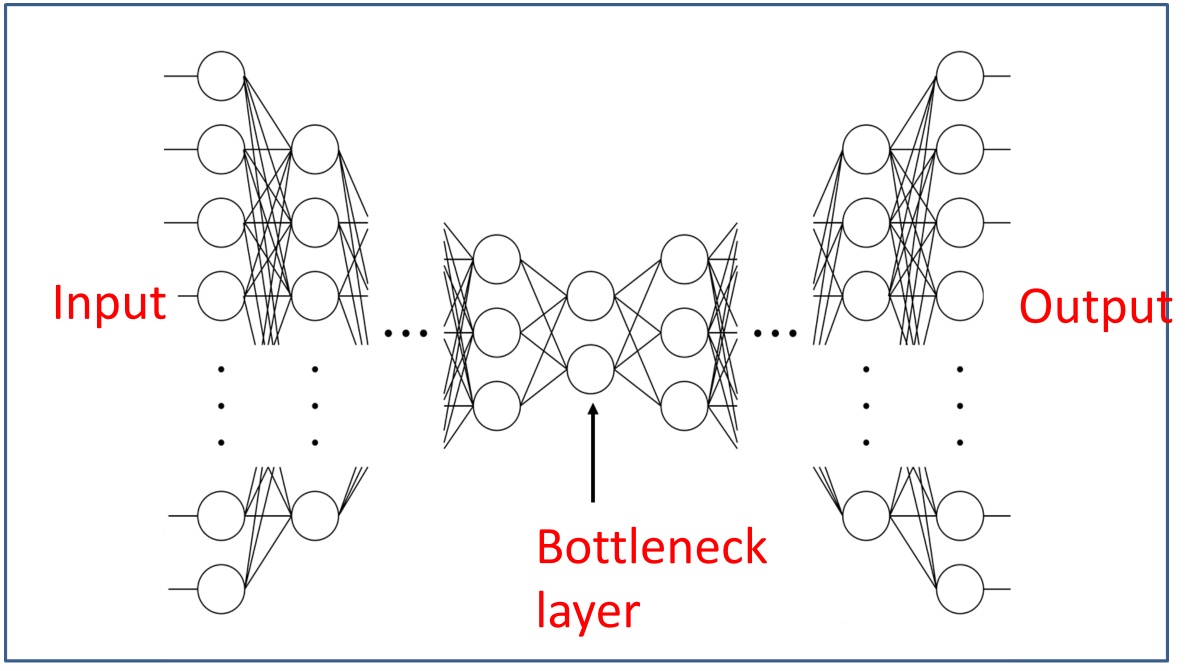
[Carifio-Halverson-Krioukov-Nelson '17; Altman-Carifio-Halverson-Nelson '18,....]

3. Vacuum structure (Unsupervised ML)

— Clustering, Feature extraction, Topological data analysis

[Cole-Shiu, '17, '18; Mutter-Parr-Vaudrevange, '18; Otsuka-Takemoto '20; Deen-He-Lee-Lukas '20],...

Autoencoder



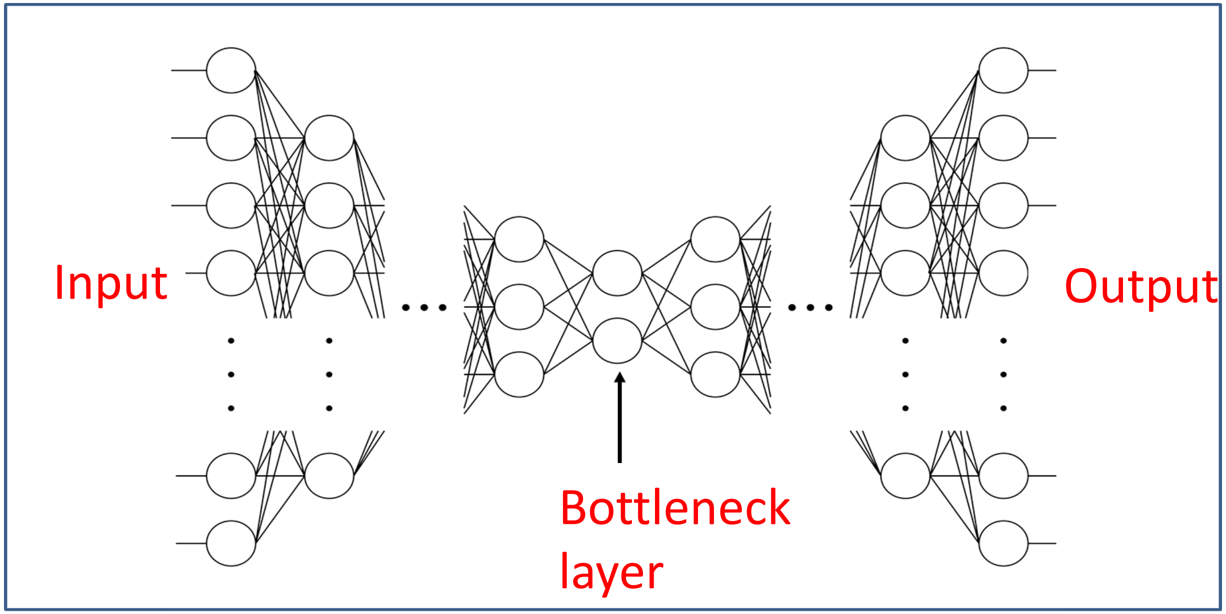
- NN is designed to Output ～Input by minimizing the error function
- Advantage :
  reducing input data to the compressed data in the 2D bottleneck layer
  (possible to extract characteristic features of the string data)
- We discuss the generation of quarks/leptons by taking
  Input data as parameters determining 4D EFT

Vacuum structure (Unsupervised ML)

The autoencoder was applied to the heterotic Z_{6-II} orbifold landscape

Mutter-Parr-Vaudrevange, 1811.05993

Input : (26 compactification parameters) $\times$(37 breaking patterns of E_8)=962-dim.
O(7×10^5) Z_{6-II} models (randomly constructed by “orbifolder” package)



Vacuum structure (Unsupervised ML)

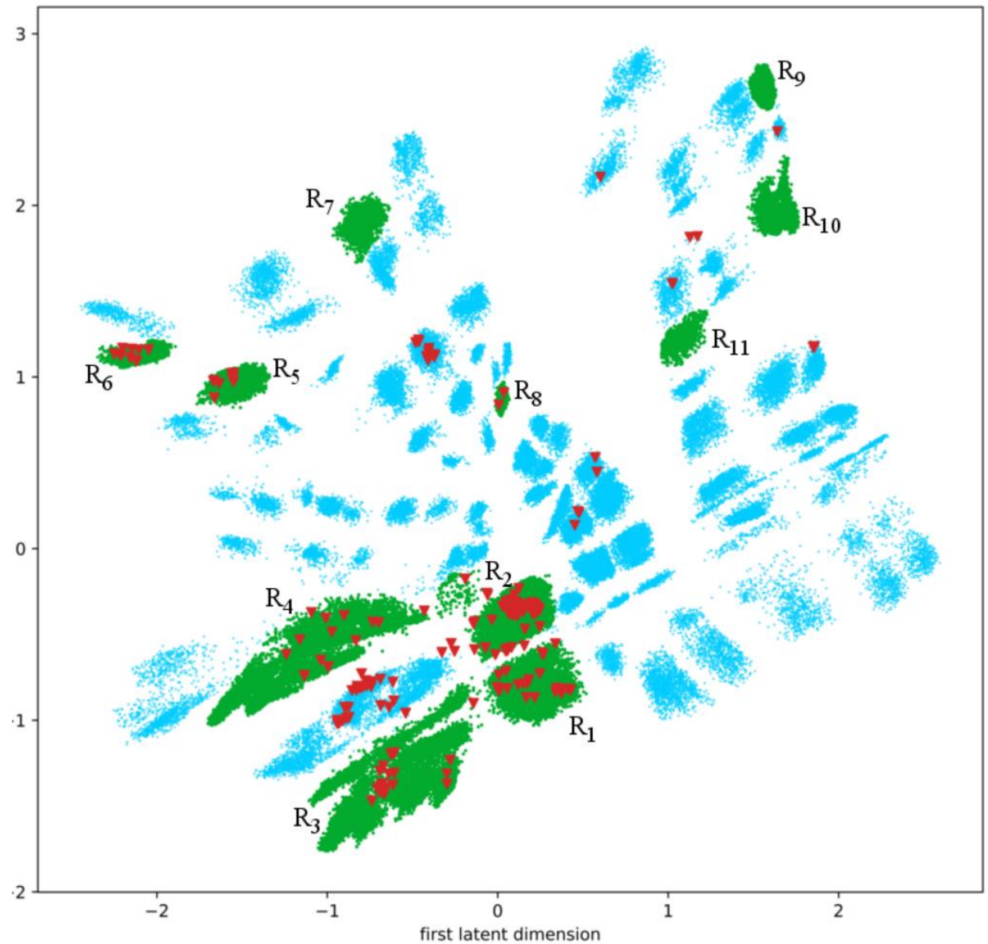
Mutter-Parr-Vaudrevange, 1811.05993

- MSSM-like models are clustered in 11 islands at the bottleneck layer

- If the input data is outside this fertile Islands, it is difficult to find the MSSM-like models

Extract only promising models

NN was trained without the knowledge of whether a model is MSSM-like or not.



Location of the MSSM-like models from the Mini-Landscape (red triangles) within the eleven fertile islands R_i *(green) and the whole* $\mathbb{Z}_6$*-II landscape (blue). As in figure 4, the MSSM-like models from the Mini-Landscape clearly prefer the fertile islands, especially islands* R_1, R_2 *and* R_3*, that were identified using our coarse sample only.*

Outline

1. **Why machine learning (ML) in string theory ?**
   — Deal with topological data and computational complexity

2. **Brief review of ML and autoencoder**

3. **Application of ML to string model building**

   H. Otsuka and K. Takemoto, JHEP **05** (2020) 047, arXiv : 2003.11880.

4. **Conclusion**

ML applications to heterotic string vacua with line bundles

Topological data
(geometrical quantities on CY, fluxes) ⇒

- SM gauge group,
- 3 generations of quarks/leptons
- Yukawa and gauge couplings

Wall's theorem : [C. T. C. Wall, '66]

If the following quantities are different for two CYs, they are *not diffeo.*

$$(h^{1,1}, h^{2,1}, c_2, \kappa_{\alpha\beta\gamma})$$

$h^{1,1}$: # of two-cycles
$h^{2,1}$: # of three-cycles
$c_2 = -\frac{1}{2(2\pi)^2}\mathrm{tr}R^2$ (Second Chern number)
$\kappa_{\alpha\beta\gamma}$: Intersection numbers among two-cycles

ML applications to heterotic string vacua with line bundles

Topological data :

$$(h^{1,1}, h^{2,1}, c_2, \kappa_{\alpha\beta\gamma}) + Fluxes$$

$h^{1,1}$: # of two-cycles
$h^{2,1}$: # of three-cycles
$c_2 = -\frac{1}{2(2\pi)^2}\mathrm{tr}R^2$ (Second Chern number)
$\kappa_{\alpha\beta\gamma}$: Intersection numbers

- These data determine the generation of quarks/leptons

Question :

Which topological data is important for 3 generations of quarks/leptons

Heterotic string on smooth CY with line bundles

- U(1) Internal gauge fluxes F in two-cycles Σ_i of CY ($i = 1,2,\cdots,h^{1,1}$)

$$\frac{1}{2\pi}\int_{\Sigma_i} F = m^{(i)} \in \mathbb{Z}$$

E.g., Hypercharge flux

$$SU(5) \to SU(3) \times SU(2) \times U(1)_Y$$

$$< F_{U(1)_Y} > \propto \begin{pmatrix} 2 & & & & \\ & 2 & & & \\ & & 2 & & \\ & & & -3 & \\ & & & & -3 \end{pmatrix}$$

- Popular in the F-theory context *Beasley-Heckman-Vafa, Donagi-Wijnholt ('08)*
- Direct flux breaking scenario is applicable in the Heterotic context

Blumenhagen-Honecker-Weigand ('05)

Heterotic string on smooth CY with line bundles

- U(1) Internal gauge fluxes F in two-cycles Σ_i of CY ($i = 1,2,\cdots,h^{1,1}$)

$$\frac{1}{2\pi}\int_{\Sigma_i} F = m^{(i)} \in \mathbb{Z}$$

- Gauge symmetry breaking $E_8 \times E_8$ or $SO(32) \to G_{\rm SM} \times G_{\rm hid}$

- Chiral and net-number of zero-modes, given by

$$\chi = \frac{1}{(2\pi)^3}\int_{\rm CY}\left[\frac{1}{6}{\rm tr}(F^3) + \frac{1}{12}{\rm tr}(R^2)\wedge {\rm tr}(F)\right]$$

Background curvatures F and R can lead to

Q, L, u^c, d^c, e^c : $\chi = -3$

No chiral exotics : $\chi = 0$

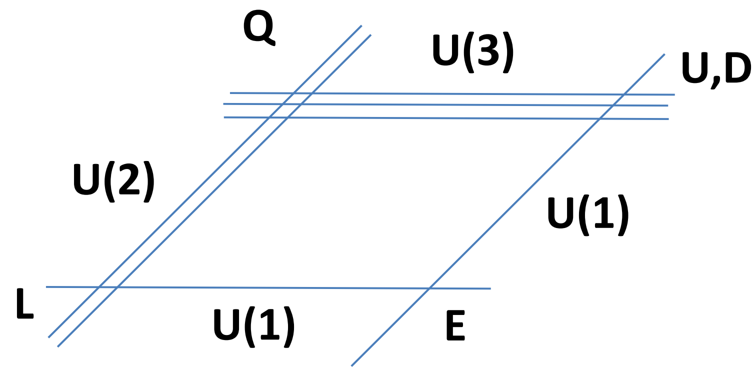
Heterotic string on smooth CY with line bundles

- $E_8 \times E_8$ heterotic Standard Models are well studied by

  Donagi-Ovrut-Pantev-Waldram ('00), Blumenhagen-Honecker-Weigand ('05)
  Anderson-Gray-Lukas-Palti ('12),….

  but Wilson lines are required to obtain the SM gauge group
  (applicable to the restricted CYs)

- SO(32) heterotic Standard Models
  S- and T-dual to Intersecting D6-brane models in type IIA string
  (Several stacks of D-branes → MSSM or Pati-Salam model)



Our research:
<u>SO(32) heterotic SM(MSSM) vacua</u>
<u>directly with the SM gauge group from smooth CYs</u>

Setup : SO(32) heterotic string with line bundles

JHEP **05** (2020) 047 (ArXiv:2003.11880) with Kenta Takemoto

Input : at most 161-dimensional parameters $SO(32) \to G_{\rm SM} \times \Pi_{a=1}^{5} U(1)_a \times SO(16)_{\rm hid}$

U(1) fluxes and topological data of 1477 classes of (complete intersection) CYs

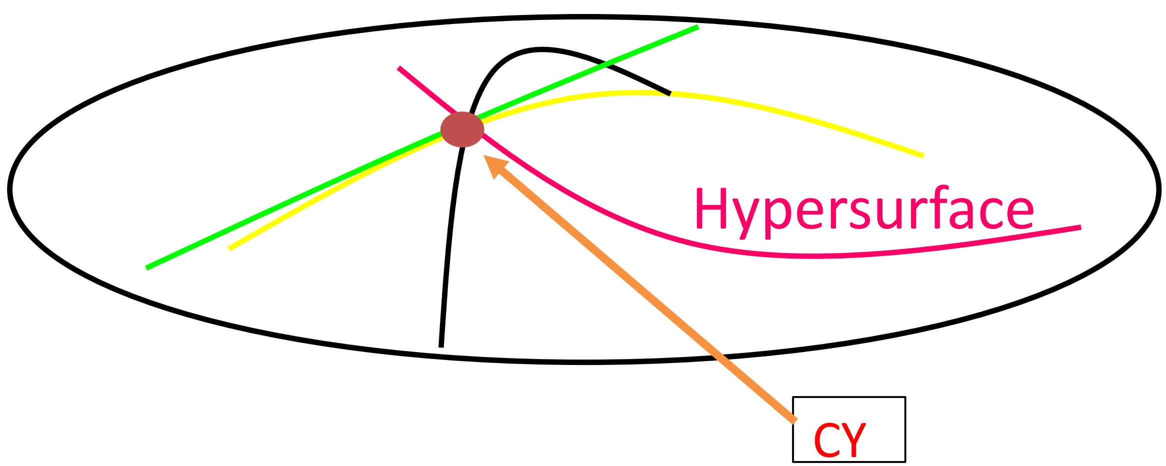
$1 \le h^{1,1} \le 5$

E.g., $h^{1,1} = 4$

$$\begin{matrix} \mathbb{P}^2 \\ \mathbb{P}^2 \\ \mathbb{P}^2 \\ \mathbb{P}^2 \end{matrix} \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

Ambient Spaces

Four $\mathbb{P}^2$



Setup : SO(32) heterotic string with line bundles

JHEP **05** (2020) 047 (ArXiv:2003.11880) with Kenta Takemoto

Input : at most 161-dimensional parameters $SO(32) \to G_{\rm SM} \times \Pi_{a=1}^{5} U(1)_a \times SO(16)_{\rm hid}$

U(1) fluxes and topological data of 1477 classes of (complete intersection) CYs

$1 \leq h^{1,1} \leq 5$

General CICY :

$$\begin{matrix} \mathbb{P}^{n_1} \\ \mathbb{P}^{n_2} \\ \vdots \\ \mathbb{P}^{n_m} \end{matrix} \begin{bmatrix} q_1^1 & q_2^1 & \cdots & q_K^1 \\ q_1^2 & q_2^2 & \cdots & q_K^2 \\ \vdots & \vdots & \ddots & \vdots \\ q_1^m & q_2^m & \cdots & q_K^m \end{bmatrix}_{m\times K}$$

$q_j^r \in \mathbb{Z}_{\neq 0}$

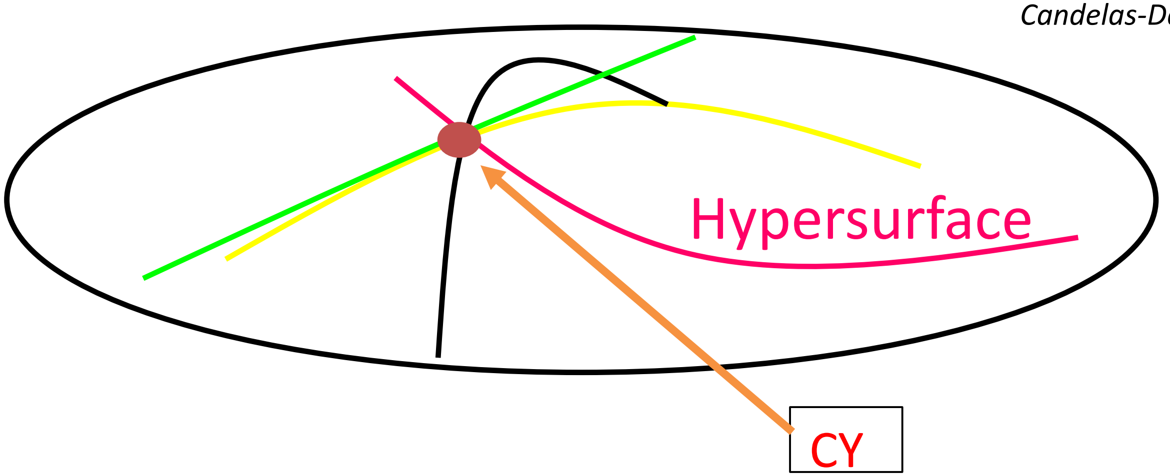
$\sum_{r=1}^{m} n_r - K = 3$

CY conditions

$\sum_{j=1}^{K} q_j^r = n_r + 1$

$r = 1, \cdots, m$

Candelas-Dale-Lutken-Schimmrigk ('88)



Setup : SO(32) heterotic string with line bundles

JHEP **05** (2020) 047 (ArXiv:2003.11880) with Kenta Takemoto

Input : at most 161-dimensional parameters $SO(32) \to G_{\mathrm{SM}} \times \Pi_{a=1}^{5} U(1)_a \times SO(16)_{\mathrm{hid}}$

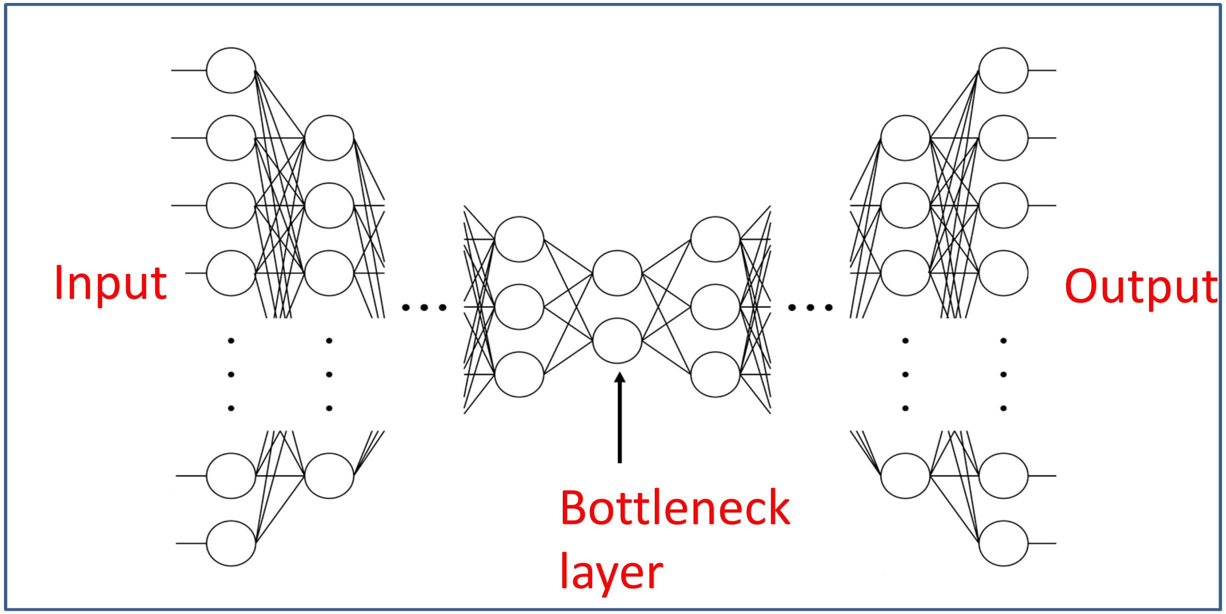
U(1) fluxes and topological data of 1477 classes of (complete-intersection) CYs

$1 \leq h^{1,1} \leq 5$

lead to $O(10^6)$ *n*-generation models (randomly constructed by our algorithms) satisfying SUSY and Tadpole cancellation conditions, masslessness $U(1)_Y$,...

Totally, 14 layers

Learning = Adam-Optimizer In TensorFlow

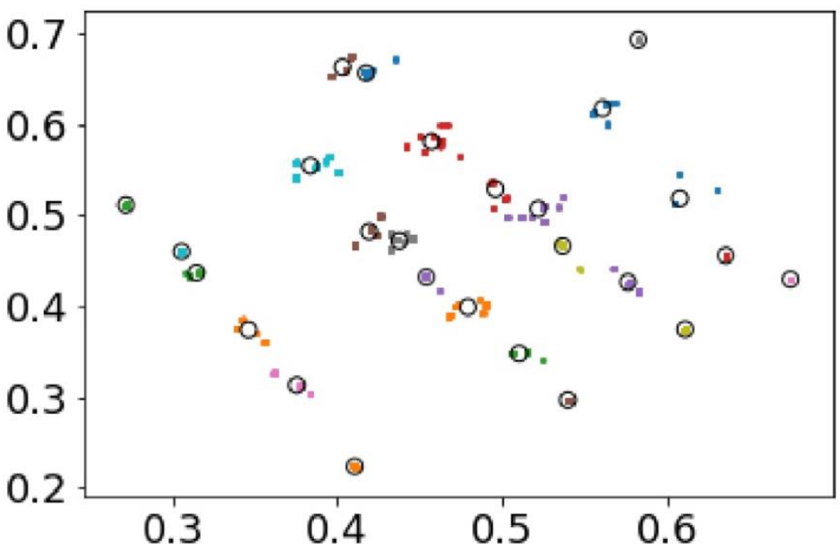


We apply the K-means clustering to the data at the bottleneck layer

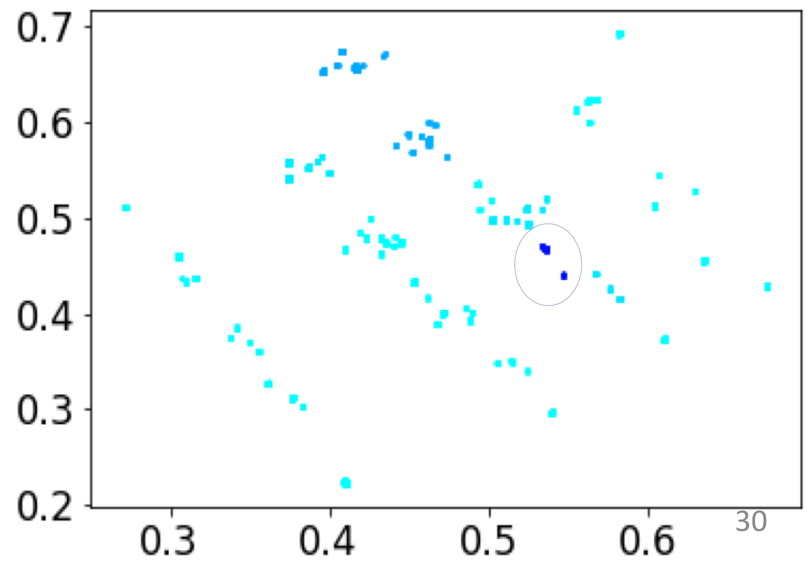
Result (I)

Hodge number of CY: $h^{1,1} = 3$, #of Clusters $= 26$

$-2 \leq \text{Flux quanta} \leq 2$



Result of AE and K-means clustering



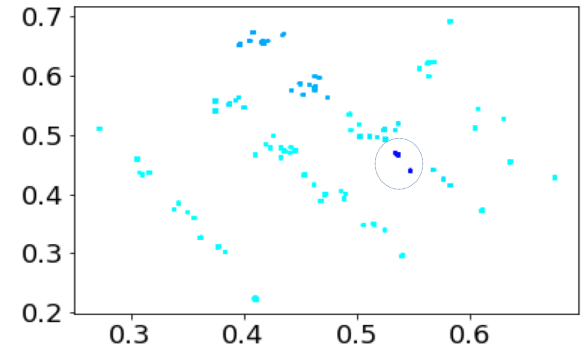
Ratio of 3-generation models to $n \neq 0$-gen. models in each cluster

The higher the ratio, the darker the color is.

- 3-generation models are clustered in the specific island ("3-generation island")
- Clustering will be universal phenomena

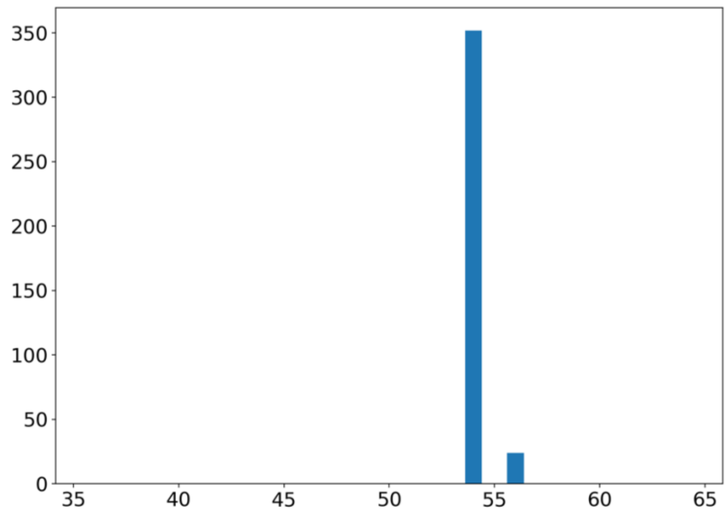
Result (II)

$$c_2 = -\frac{1}{2(2\pi)^2}\mathrm{tr}R^2 = c_{2,i}\widehat{w}_i \qquad \widehat{w}_i : \text{four-forms } (i = 1,2,3)$$

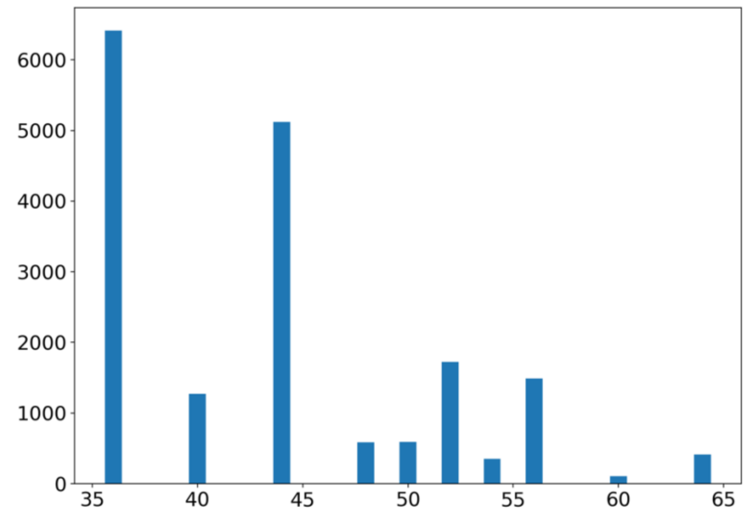


Histogram of $c_{2,3}$ (Curvature of CY)

"3-generation island"



All the region



$h^{1,1}$	N_{cl}	Favored $c_{2,i}$
3	26	(36,36,54)

- 3 generation

$\simeq$ Curvature of CY (c_2) = $18\mathbb{Z}$

Result (II)

Search	$h^{1,1}$	$N_{\rm cl}$	Favored $c_{2,i}$
(I)	3	26	(36,36,54)
	4	30	(24,24,36,36)
	5	42	(24,36,36,36,36)
(II)	3	40	(36,36,36)
	4	30	(24,36,36,36)

$-2 \leq$ Flux quanta ≤ 2

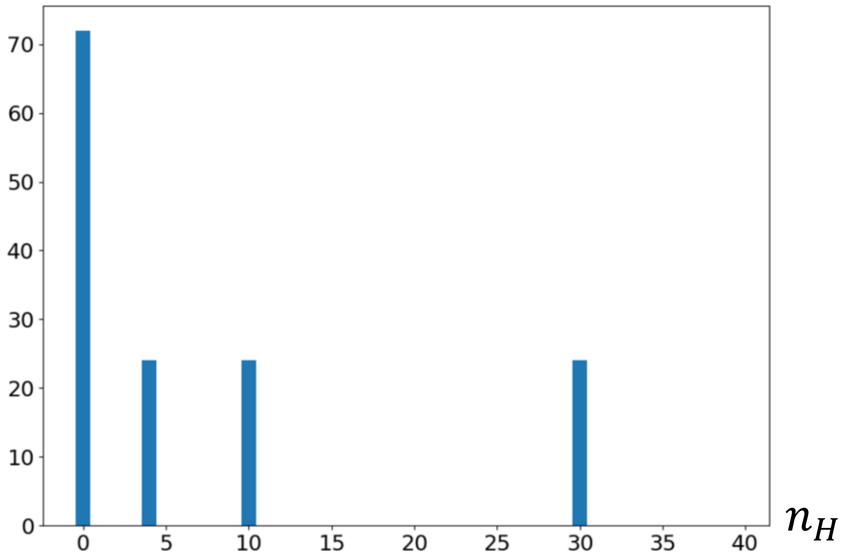
$-3 \leq$ Flux quanta ≤ 3

- 3-generation island is strongly correlated with the second Chern number of CY, compared with other topological data

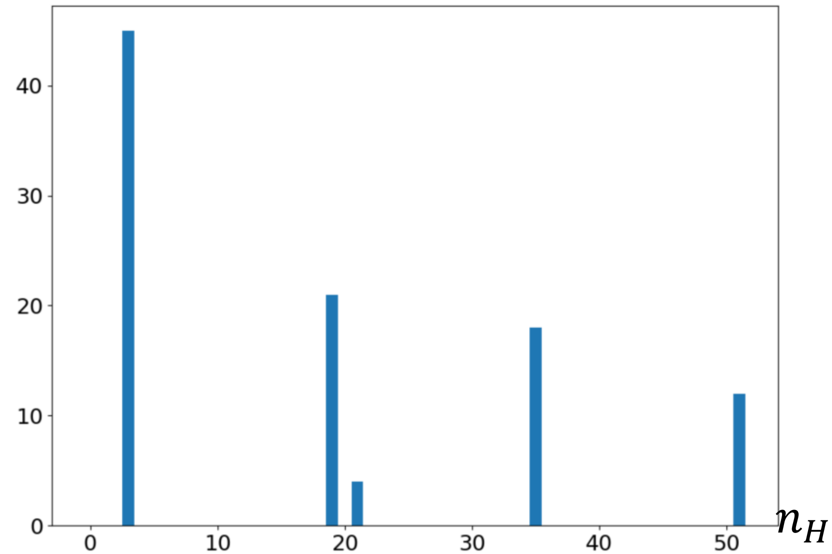
Result (III) : Generations of Higgs

- We count the number of Higgs pairs (n_H)
  — vector-like under SM gauge group, but chiral w.r.t. extra U(1)s

# of models in "3-generation island"



$h^{1,1} = 3, -2 \leq \text{Flux quanta} \leq 2$



$h^{1,1} = 3, -3 \leq \text{Flux quanta} \leq 3$

- In our limited search, 1-pair Higgs model is disfavored
- Generic property : a large number of Higgs pairs

Application to other heterotic string theories

- So far, we analyze *SO(32)* heterotic string line bundle models

- We are trying to analyze $E_8 \times E_8$ and $SO(16) \times SO(16)$ heterotic string theories more rigorously

- We have observed that similar clustering phenomena indeed exist in $E_8 \times E_8$ heterotic line bundle models (Work in progress)

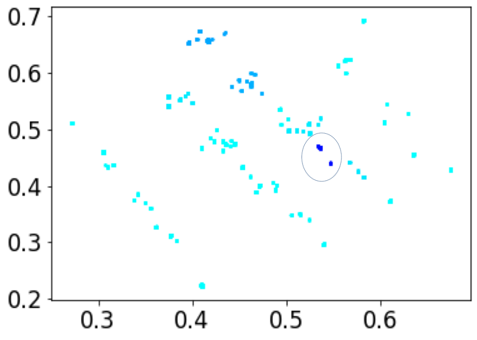
Conclusion

ML can deal with

- Gigantic number of topological data (CYs, flux/brane,...)
- Computationally complexity

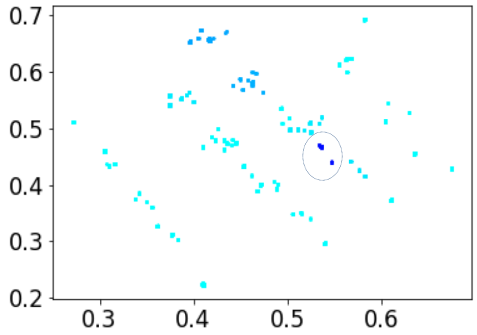
For SO(32) heterotic string on CY with line bundles,

- 3-generation models are clustered in the specific island "3-generation island"
  similar to the toroidal orbifold landscape



- 3 generation $\simeq$ Curvature of CY $=18\mathbb{Z}$

Discussion



- Why cluster ?

- Applications of our method to other string theory

  — $E_8 \times E_8$ heterotic string (work in progress)
  — D-brane models (Type IIB/IIA)

Other ML techniques are also useful to reveal hidden structures in the string landscape from gigantic number of topological data