Learning non-Higgsable gauge groups in 4D F-theory

Upcoming work with Zhibai Zhang

Yi-Nan Wang

CTP, MIT

Workshop on data science and string theory, LMU Munich
Mar. 27th, 2018

The problem and the solution

- Problem: in 4D F-theory, given local geometric data near a divisor D on the base B, decide what's the (geometric) non-Higgsable gauge group on D

- Input data (feature): local triple intersection numbers between D and its neighbors

- Output data (label): the non-Higgsable gauge group on D, only 10 possible choices

- Solution: supervised machine learning

F-theory

• Physical setup: 4D F-theory compactification on an elliptic Calabi-Yau fourfold X with complex threefold base B.

• F-theory is a geometric description of strongly coupled IIB superstring theory.

• The elliptic fibration X over B is described by a Weierstrass form:

$$y^2 = x^3 + fx + g \tag{1}$$

• 7-branes locates at the cod-1 locus of $\Delta = 4f^3 + 27g^2 = 0$, where the elliptic fiber is singular.

• Non-Abelian gauge group$\leftrightarrow$ order of vanishing of (f, g, Δ) and other information.

F-theory

$\text{ord}(f)$	$\text{ord}(g)$	$\text{ord}(\Delta)$		Gauge group
0	0	2	I_2	$SU(2)$
0	0	$n \geq 3$	I_n	$Sp\lfloor \frac{n}{2} \rfloor$ or $SU(n)$
1	≥ 2	3	III	$SU(2)$
≥ 2	2	4	IV	$SU(2)$ or $SU(3)$
≥ 2	≥ 3	6	I_0^*	G_2 or $SO(7)$ or $SO(8)$
2	3	$6+n$	I_n^*	$SO(8+2n)$ or $SO(7+2n)$
≥ 3	4	8	IV^*	F_4 or E_6
3	≥ 5	9	III^*	E_7
≥ 4	5	10	II^*	E_8

Non-Higgsable phase of F-theory

• We require that the elliptic fibration is "generic", hence f and g are general holomorphic sections of line bundles $-4K_B$ and $-6K_B$.

• The gauge groups in the 4D supergravity model are minimal (geometric non-Higgsable). The only possible NH gauge groups are
$\varnothing$, SU(2), SU(3), G_2, SO(7), SO(8), F_4, E_6, E_7, E_8

• The number of complex structure moduli $h^{3,1}(X)$ is maximal.

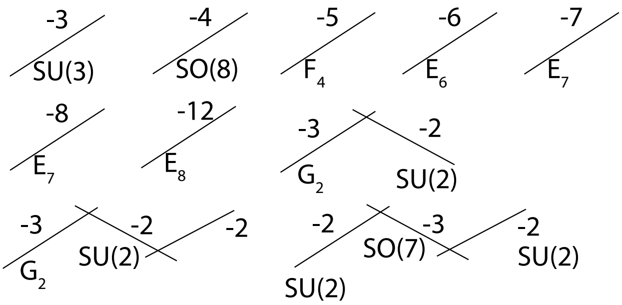
• From the non-Higgsable phase, we can tune f and g to get bigger gauge groups, such as GUT SU(5).

• Non-Higgsable gauge group structures are good characterization of the base geometry, e.g. non-Higgsable clusters in 6D F-theory(Morrison, Taylor 12')/atomic classification of 6D (1,0) SCFT(Heckman, Morrison, Rudelius, Vafa 15').

6D non-Higgsable clusters

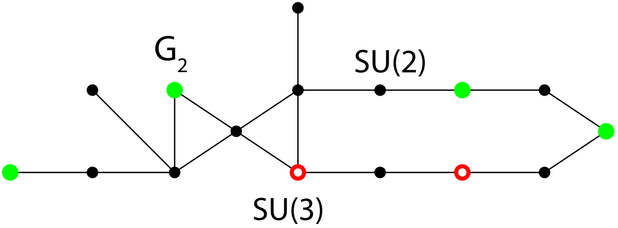
• In 6D F-theory, B is a complex surface, and the geometric data are the intersection numbers between curves.

• If we require that (f, g) does not vanish to order (4,6) or higher on a point, then the non-Higgsable clusters are:



- In 4D F-theory, B is a complex threefold, the intersection structure is much more complicated.

- An example found in (Taylor, YNW 15'):



- It seems that the notation of "cluster" may not be very useful.

Resolvable bases and good bases

- In 6D F-theory, if (f,g) vanishes to order (4,6) at cod-2 locus, it means that there's a strongly coupled (1,0) SCFT sector decoupled with gravity.

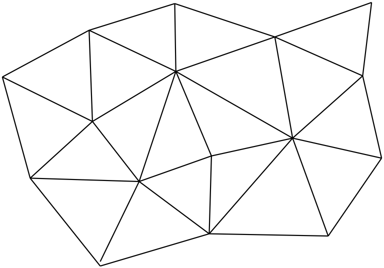
- In 4D F-theory, one may expect that cod-2 (4,6) locus will give rise to 4D $\mathcal{N}=1$ SCFT. There can be subtleties if we include Euclidean D3-brane effects (Apruzzi, Heckman, Morrison, Tizzano 18')or G_4 flux.

- In general, we accept all the bases with cod-2 (4,6) locus that can be blown up to get rid of them: resolvable bases.

- good bases: bases without cod-2 (4,6) locus; may have cod-3 (4,6) but not as high as (8,12).

Toric threefolds

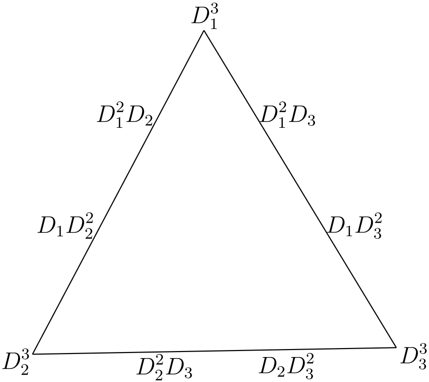
• For simplicity, we only study the smooth toric base threefolds. Toric fan: a collection of 3D, 2D, 1D simplicial cones in the lattice $\mathbb{Z}^3$.



point $\leftrightarrow$ divisor; line $\leftrightarrow$ curve; triangle $\leftrightarrow$ point.

Toric threefolds

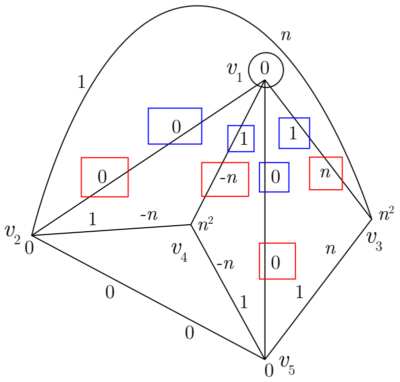
The triple intersection numbers are labeled as



$D_1 \cdot D_2 \cdot D_3 = 1$

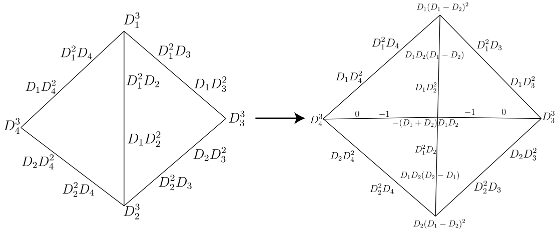
Toric threefolds

- The Hirzebruch threefold $\tilde{\mathbb{F}}_n$ ($\mathbb{P}^1$ over $\mathbb{P}^2$):

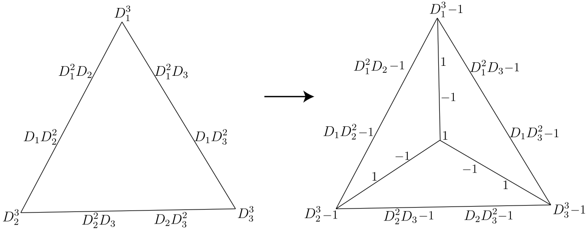


- blue: normal bundle; red: self-intersection of curves

Blow ups



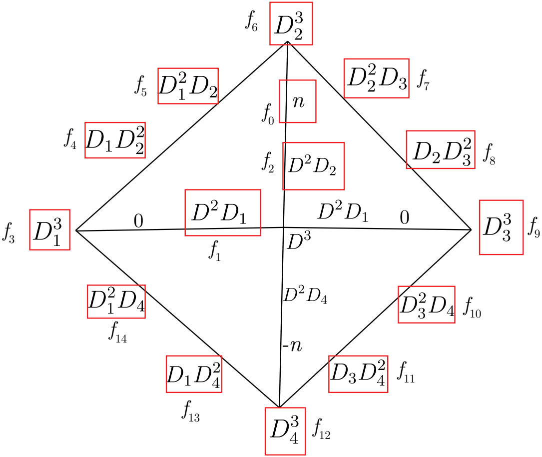
D_1^3
$D_1^2D_4$
$D_1^2D_3$
$D_1D_4^2$
$D_1^2D_2$
$D_1D_3^2$
D_4^3
D_3^3
$D_1D_2^2$
$D_2D_4^2$
$D_2D_3^2$
$D_2^2D_4$
$D_2^2D_3$
D_2^3
$D_1(D_1-D_2)^2$
$D_1^2D_4$
$D_1D_2(D_1-D_2)$
$D_1^2D_3$
$D_1D_4^2$
$D_1D_2^2$
$D_1D_3^2$
D_4^3
0
-1
-1
0
D_3^3
$-(D_1+D_2)D_1D_2$
$D_1^2D_2$
$D_2D_4^2$
$D_2D_3^2$
$D_1D_2(D_2-D_1)$
$D_2^2D_4$
$D_2^2D_3$
$D_2(D_1-D_2)^2$



D_1^3
$D_1^2D_2$
$D_1^2D_3$
$D_1D_2^2$
$D_1D_3^2$
D_2^3
$D_2^2D_3$
$D_2D_3^2$
D_3^3
D_1^3-1
$D_1^2D_2-1$
1
$D_1^2D_3-1$
-1
$D_1D_2^2-1$
1
$D_1D_3^2-1$
-1
-1
1
1
D_2^3-1
$D_2^2D_3-1$
$D_2D_3^2-1$
D_3^3-1

Input vector for machine learning

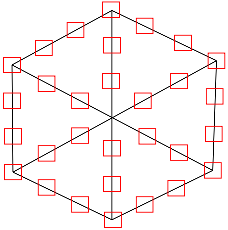
- For divisors $D = \mathbb{F}_n$ (Hirzebruch surface) with 4 neighbors:



Generation of data set

- Separate divisors according to the number of neighbors n: $h^{1,1}(D) = n - 2$.

- Compute the triple intersection numbers between divisors. The number of local triple intersection numbers is different for each n



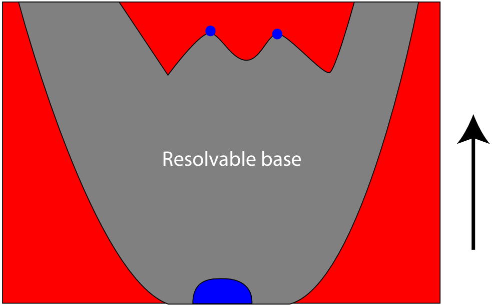
- Compute f, g and the gauge group on each divisor using toric geometry machinery

Generation of data set

- We generate our train set from two different set of good bases:

(1) The end point bases (Taylor, YNW 17'): start with $\mathbb{P}^3$, blow up until one can't; allowing resolvable bases in the process. The end point is always good.

(2) The good bases from random walk (Taylor, YNW 15'): start with $\mathbb{P}^3$, do a random blow up/down sequence; do not allow bases with toric (4,6) curves.



Choosing the machine learning method

- Training set: $\mathbb{F}_n$ on end point bases, ~ 3 million samples.

G	$\varnothing$	SU(2)	SU(3)	G_2	SO(8)	F_4	E_8
$N(G)$	2053638	520783	24	286592	8	66934	4374

- Highly unbalanced; we do a resampling and equalize $N(G) \approx N(G)_{max}/100$.

- On the resampled data set, we separate (train:test)=(0.75:0.25).

Choosing the machine learning method

Classification method	Accuracy	Training time
Logistic Regression	77.33%	28.6
Decision Tree	99.17%	1.2
Random Forest	99.32%	2.7
Support Vector Machine	97.05%	4.9
Feedforward Neural Network	96.27%	279

- Untrimmed decision tree is the best method in our case:

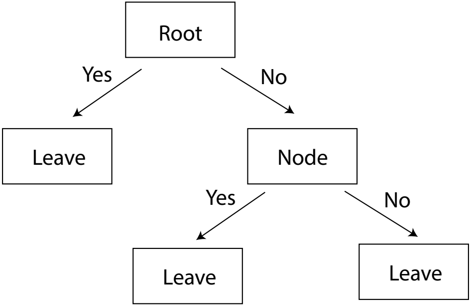
(1) Accuracy as high as random forest

(2) Great interpretability

(3) Fast

- Seems that overfitting does not occur

Decision tree



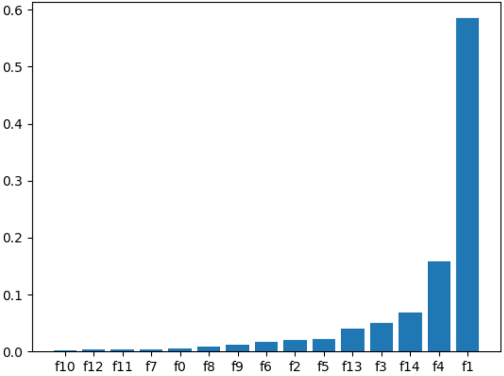
- At each step, we try to split the data set with a feature which maximize

$$\text{Information gain} = I - \frac{N_{left}}{N} I_{left} - \frac{N_{right}}{N} I_{right}, \tag{2}$$

- I is the Gini index:

$$I = \sum_i p_i(1 - p_i), \tag{3}$$

Feature Importance



- $f_1 = D^2 D_1$ is the most important one.

Results

- For the actual mixed data set from end point bases and other bases, $\sim$ 6,300,170 samples.

- We do a resampling $N(G) \approx N(G)_{max}/10$; (train:test)=(0.75:0.25).

- The decision tree has 66441 nodes and 33221 leaves, and the maximal depth is $d_{\max} = 49$.

- The in sample and out of sample accuracies on the resampled data set are 98.22% and 97.79%. On the original set, it is $A = 97.86\%$.

- Generated a lot of analytic rules in terms of inequalities, for example: if $f_1 = D^2 D_1 \leq -9$, then the gauge group is E_8.

Results

d	$\lvert S(l)\rvert$	f_0	f_1	f_2	f_3	f_4	f_5	f_6	f_{13}	other f_i	G
2	6831	-	≤ -9	-	-	-	-	-	-	-	E_8
7	4698	-	-6	-	-	-	≤ 0	-	-	$f_7 \geq 0$	F_4
8	52528	-	-5	-	-	-	≤ 3	≤ 5	≤ -1	$f_7 \geq 0,\ f_{12} \leq -1$	F_4
10	4478	-	-5	$-7 \sim -4$	≥ 2	-	≥ 4	≤ 3	≤ -1	-	F_4
9	4285	0	-5	-	-	-	≤ 2	≥ 6	≤ -1	$f_7 \geq 0$	F_4
9	5680	-	-4	≥ -12	-	$1 \sim 3$	≥ 5	-	-	-	G_2
10	4249	-	-4	≥ -12	≥ 5	$1 \sim 4$	≤ 4	-	-	$f_7 \geq 0$	G_2
12	4248	≥ 7	-4	≤ -13	-	3	-	-	-	$f_7 \geq 0,\ f_{10} \geq 0$	G_2
8	160214	-	-3	-	-	≥ 6	≥ -5	-	-	$f_{12} \leq -1$	G_2
9	42972	-	-3	-	-	≤ 2	≥ -5	-	-	$f_{11} = 0,\ f_{12} \leq -2$	G_2
9	42871	-	-3	-	-	$3 \sim 5$	≥ -5	≥ 7	-	$f_{12} \leq -1$	G_2
16	17985	-	-3	-2	≥ 6	$0 \sim 2$	≥ 12	$-4 \sim 2$	-	$f_{12} \geq 0,\ f_{14} \leq -1$	SU(2)
12	11200	-	-3	≤ -3	-	≥ 7	≥ -5	-	-	$f_{12} \geq 0,\ f_{14} \geq -2$	G_2
14	8483	-	-3	≥ -12	-	≤ 1	$-5 \sim 11$	-	≤ -2	$f_{12} \geq -1,\ f_{14} = 0$	G_2
13	3751	≥ 1	-3	≥ -5	-	$3 \sim 5$	≥ -5	≤ 6	-	$f_7 \geq 0,\ f_{10} \geq 0,\ f_{12} \leq -1$	G_2
16	3536	-	-3	≤ -13	≤ 6	$-7 \sim 2$	$-5 \sim 3$	≥ 10	-	$f_{11} \geq 0,\ f_{12} \geq -1$	SU(2)
16	3355	≥ 2	-3	$-12 \sim -7$	≥ -11	2	$-5 \sim 11$	-	≤ -1	$f_7 \geq 0,\ f_{12} \geq -1$	G_2
20	3230	-	-3	≥ -4	≥ 3	2	$3 \sim 11$	-	≤ -1	$f_7 \geq 0,\ f_{10} \geq 0,\ f_{11} \geq 0,\ f_{12} \geq -1,\ f_{14} \geq 0$	SU(2)
12	245756	-	-2	≤ -3	-	≥ 16	-	-	≤ -1	$f_7 \geq -1,\ f_8 \leq -2$	99.9996% SU(2), 0.0004%∅
14	11713	≥ 3	-2	≤ -3	-	$14 \sim 51$	-	-	≤ -1	$f_7 \geq -1,\ f_8 \geq -1$	99.94% SU(2), 0.06%∅
12	9063	-	-2	≥ -11	≤ 0	$5 \sim 12$	≤ 2	-	≤ -3	$f_{12} \leq -1,\ f_{14} \leq 0$	SU(2)
16	8989	≥ 1	-2	≥ -11	≥ 1	$10 \sim 12$	≥ 0	≥ 13	-	$f_9 \leq -2,\ f_{12} \leq -1$	SU(2)
13	8710	-	-2	≥ -11	≥ -18	≤ 4	-	-	≤ -3	$f_{10} \geq 0,\ f_{12} \leq -1,\ f_{14} = 0$	SU(2)
13	6218	≥ 2	-2	≤ -2	-	$14 \sim 15$	-	-	≤ -1	$f_7 \geq -1,\ f_8 \leq -2$	SU(2)
19	3017	≥ 1	-2	≥ -11	≥ 1	$10 \sim 12$	$0 \sim 2$	$8 \sim 12$	-	$f_9 \leq -2,\ f_{12} \leq -1,\ f_{14} \leq -1$	SU(2)

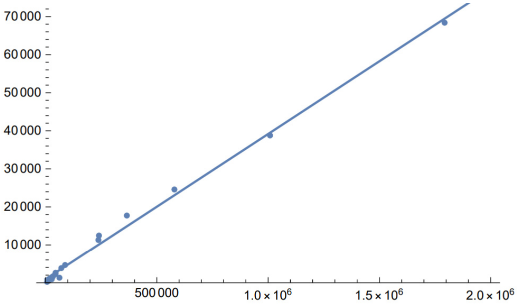
If the rule is simple enough, then it can be derived by formula (Morrison,Taylor 14'):

$$f_k \in \mathcal{O}(-4K_D + (4-k)N_D - \sum_{D \bigcap D_j \neq \varnothing} \phi_j C_{ij}), \tag{4}$$

$$g_k \in \mathcal{O}(-6K_D + (6-k)N_D - \sum_{D \bigcap D_j \neq \varnothing} \gamma_j C_{ij}). \tag{5}$$

A Universality

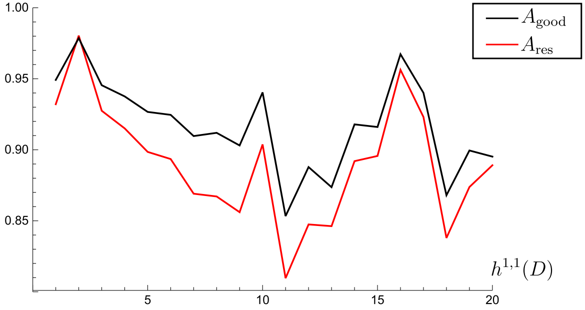
- For each different $h^{1,1}(D)$, train a different decision tree
- $N_{\text{nodes}} \propto N_{\text{samples}}$



Applications

(1) Cross-check on resolvable bases

• The resolvable bases with toric (4,6) curves are generated from random walks starting from $\mathbb{P}^3$

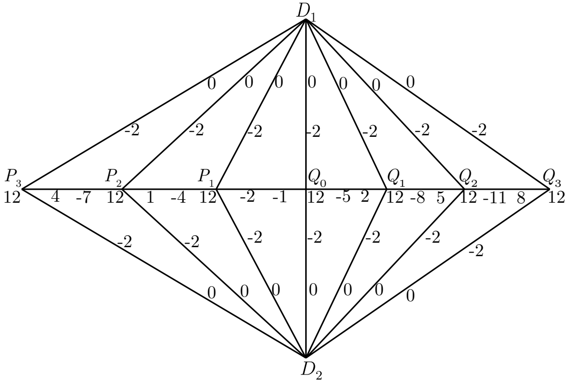


• The decision tree trained from good bases works on resolvable bases as well!

Applications

(2) Constructing local geometric configurations

- Assemble the rules compatible with each other



- An infinite? SU(3) chain with a non-Higgsable SU(3) at each P_n and Q_n