Decision trees, Brill–Noether theory, and vector-like spectra in F-theory

work in progress with M. Bies, M. Cvetič, R. Donagi, M. Liu, F. Ruehle

Ling Lin

CERN Theory Department

String Pheno
June 11, 2020

Motivation

- Classical problem of string pheno: find realization of (MS)SM in string landscape.
- In particular: need (massless) vector-like pair(s) to accommodate the Higgs.
- More generally: vector-like spectrum is charaterizing feature of 4d vacuum.

Motivation

- Classical problem of string pheno: find realization of (MS)SM in string landscape.
- In particular: need (massless) vector-like pair(s) to accommodate the Higgs.
- More generally: vector-like spectrum is charaterizing feature of 4d vacuum.
- In global F-theory compactifications: difficult to control due to non-topological nature.

Motivation

- Classical problem of string pheno: find realization of (MS)SM in string landscape.
- In particular: need (massless) vector-like pair(s) to accommodate the Higgs.
- More generally: vector-like spectrum is charaterizing feature of 4d vacuum.
- In global F-theory compactifications: difficult to control due to non-topological nature.
- How can machine learning techniques help?

Outline

1. Computing vector-like spectrum in F-theory
2. Learning cohomology jumps with Decision Trees
3. Application to toy example
4. “Moduli” space of jumps

Chiral matter in 4d F-theory

F-theory $\cong$ type IIB on B_3 with (p, q)-7-branes at finite g_s, by geometrizing backreaction on g_s in *elliptic CY-fibration* $\pi : Y_4 \to B_3$ (cf. 2017 TASI lectures by Weigand and Cvetič for recent reviews, see also C. Long's talk).

Chiral matter in 4d F-theory

F-theory $\cong$ type IIB on B_3 with (p, q)-7-branes at finite g_s, by geometrizing backreaction on g_s in *elliptic CY-fibration* $\pi : Y_4 \to B_3$ (cf. 2017 TASI lectures by Weigand and Cvetič for recent reviews, see also C. Long's talk).

$\mathcal{N} = 1$ gauge sector on 7-branes wrapped on (complex) surfaces $S \subset B_3$ with matter localized on curves $C \subset S$. For chiral spectrum: need to turn on gauge flux background $G_4 \in H^{2,2}(Y_4)$.

Chiral matter in 4d F-theory

F-theory $\cong$ type IIB on B_3 with (p, q)-7-branes at finite g_s, by geometrizing backreaction on g_s in *elliptic CY-fibration* $\pi : Y_4 \to B_3$ (cf. 2017 TASI lectures by Weigand and Cvetič for recent reviews, see also C. Long's talk).

$\mathcal{N} = 1$ gauge sector on 7-branes wrapped on (complex) surfaces $S \subset B_3$ with matter localized on curves $C \subset S$. For chiral spectrum: need to turn on gauge flux background $G_4 \in H^{2,2}(Y_4)$. Chiral excess is topological: can be computed via intersection theory.

[Donagi/Wijnholt, 09],[Braun/Collinucci/Valandro, 11], [Marsano/Schäfer-Nameki, 11], [Krause/Mayrhofer/Weigand, 11,12], [Grimm/Hayashi, 11], [Cvetič/Grimm/Klevers, 12], [Braun/Grimm/Keitel, 13], [Cvetič/Grassi/Klevers/Piragua, 13], [Borchmann/Mayrhofer/Palti/Weigand, 13], [LL/Mayrhofer/Till/Weigand, 15]

Chiral matter in 4d F-theory

F-theory $\cong$ type IIB on B_3 with (p, q)-7-branes at finite g_s, by geometrizing backreaction on g_s in *elliptic CY-fibration* $\pi : Y_4 \to B_3$ (cf. 2017 TASI lectures by Weigand and Cvetič for recent reviews, see also C. Long's talk).

$\mathcal{N} = 1$ gauge sector on 7-branes wrapped on (complex) surfaces $S \subset B_3$ with matter localized on curves $C \subset S$. For chiral spectrum: need to turn on gauge flux background $G_4 \in H^{2,2}(Y_4)$. Chiral excess is topological: can be computed via intersection theory.
[Donagi/Wijnholt, 09],[Braun/Collinucci/Valandro, 11], [Marsano/Schäfer-Nameki, 11], [Krause/Mayrhofer/Weigand, 11,12], [Grimm/Hayashi, 11], [Cvetič/Grimm/Klevers, 12], [Braun/Grimm/Keitel, 13], [Cvetič/Grassi/Klevers/Piragua, 13], [Borchmann/Mayrhofer/Palti/Weigand, 13], [LL/Mayrhofer/Till/Weigand, 15]

Allows explicit construction of "quadrillions" of MSSM(-like) three-family F-theory models.
[Cvetič/Klevers/Mayorga/Oehlmann/Reuter, 15], [LL/Weigand, 16], [Cvetič/LL/Liu/Oehlmann, 18], [Cvetič/Halverson/LL/Liu/Tian, 19]

Zero mode counting in global models

- Massless (anti-)chiral modes in representation $\mathbf{R}$ on *matter curves* $C_{\mathbf{R}}$ depend on C_3 rather than $G_4 = dC_3$. Encoded in *intermediate Jacobian* of Y_4 [Curio/Donagi, 98], [Donagi/Wijnholt, 12,13], [Anderson/Heckman/Katz, 13].

- Can be parametrized by Chow ring $\mathrm{CH}^2(Y_4)$ [Bies/Mayrhofer(/Pehle)/Weigand, 14,17]
$\Longrightarrow$ computationally more feasible: given $A \in \mathrm{CH}^2(Y_4)$, can extract for each $C_{\mathbf{R}}$ a line bundle $\mathcal{L}_{\mathbf{R}}$ such that

$$\text{massless chiral modes of} \longleftrightarrow H^0(C_{\mathbf{R}}, \mathcal{L}_{\mathbf{R}}) \,,$$
$$\text{massless anti-chiral modes} \longleftrightarrow H^1(C_{\mathbf{R}}, \mathcal{L}_{\mathbf{R}}) \,,$$

$\chi(\mathbf{R}) = h^0 - h^1$ topological invariant, depends only on $G_4 = [A] \in H^{2,2}(Y_4)$.

Zero mode counting in global models

- Massless (anti-)chiral modes in representation $\mathbf{R}$ on *matter curves* $C_{\mathbf{R}}$ depend on C_3 rather than $G_4 = dC_3$. Encoded in *intermediate Jacobian* of Y_4 [Curio/Donagi, 98], [Donagi/Wijnholt, 12,13], [Anderson/Heckman/Katz, 13].
- Can be parametrized by Chow ring $\mathrm{CH}^2(Y_4)$ [Bies/Mayrhofer(/Pehle)/Weigand, 14,17] $\Longrightarrow$ computationally more feasible: given $A \in \mathrm{CH}^2(Y_4)$, can extract for each $C_{\mathbf{R}}$ a line bundle $\mathcal{L}_{\mathbf{R}}$ such that

$$\text{massless chiral modes of } \longleftrightarrow H^0(C_{\mathbf{R}}, \mathcal{L}_{\mathbf{R}})\,,$$
$$\text{massless anti-chiral modes} \longleftrightarrow H^1(C_{\mathbf{R}}, \mathcal{L}_{\mathbf{R}})\,,$$

  $\chi(\mathbf{R}) = h^0 - h^1$ topological invariant, depends only on $G_4 = [A] \in H^{2,2}(Y_4)$.

- $\mathcal{L}_{\mathbf{R}}$ given as collection of points in $B_3 \Longrightarrow$ can be modeled as *coherent sheaf* on B_3.
- $H^i(C_{\mathbf{R}}, \mathcal{L}_{\mathbf{R}})$ computed via Ext groups; algorithm implemented in computer algebra system CAP [Bies, 17], [Bies/Posur, 19].

Computational challenges

- Sheaf description very general, e.g., $C_{\mathbf{R}}$ need not be smooth. However, implementation on computer extremely resource intensive, fails, e.g., if genus$(C_{\mathbf{R}})$ too large ($\gtrsim 10$).
  - Typical F-theory models have curves with $g > 20$ (oftentimes the would-be Higgs!).
- Cohomologies depend on complex structure of Y_4, of which there are in general $\mathcal{O}(100)$.
  - Determine (even just part of) complex structure dependence of vector-like spectrum tricky.
- In practice, sheaf description lacks type IIB-ish intuitions about localized matter.
  - Difficult to compute Yukawa couplings in global models [Cvetič/LL/Liu/Zhang/Zoccarato, 19].

First order questions: What is the "generic" value of h^0? When does it "jump"?

Machine Learning line bundle cohomology

- Surge of recent interest to study $H^i(X, \mathcal{L})$ with machine learning [Ruehle, 17], [Kläwer/Schlechter, 18], [Larfors/Schneider, 19,20], [Brodie/Constantin/Deen/Lukas, 19], mostly suited for heterotic compactifications: X is smooth with known $\text{Pic}(X) \ni \mathcal{L}$.

Machine Learning line bundle cohomology

- Surge of recent interest to study $H^i(X, \mathcal{L})$ with machine learning [Ruehle, 17], [Kläwer/Schlechter, 18], [Larfors/Schneider, 19,20], [Brodie/Constantin/Deen/Lukas, 19], mostly suited for heterotic compactifications: X is smooth with known $\text{Pic}(X) \ni \mathcal{L}$.
- Qualitatively different in F-theory: $C_{\mathbf{R}}$ and $\mathcal{L}_{\mathbf{R}}$ given by polynomials on B_3.
  - $C_{\mathbf{R}}$ can be singular.
  - $\text{Pic}(C_{\mathbf{R}})$ in general not known.
  - $\mathcal{L}_{\mathbf{R}}$ specified by sum of points $\sum_i \lambda_i p_i$, where $p_i \in B_3$ lie on $C_{\mathbf{R}}$.
    $\Longrightarrow$ which p_i are rationally equivalent? Even non-trivial if $\mathcal{L}_{\mathbf{R}}$ is pull-back bundle!
  - Both $C_{\mathbf{R}}$ and $\mathcal{L}_{\mathbf{R}}$ can simultaneously change with complex structure.

Machine Learning line bundle cohomology

- Surge of recent interest to study $H^i(X, \mathcal{L})$ with machine learning [Ruehle, 17], [Kläwer/Schlechter, 18], [Larfors/Schneider, 19,20], [Brodie/Constantin/Deen/Lukas, 19], mostly suited for heterotic compactifications: X is smooth with known $\text{Pic}(X) \ni \mathcal{L}$.
- Qualitatively different in F-theory: $C_{\mathbf{R}}$ and $\mathcal{L}_{\mathbf{R}}$ given by polynomials on B_3.
  - $C_{\mathbf{R}}$ can be singular.
  - $\text{Pic}(C_{\mathbf{R}})$ in general not known.
  - $\mathcal{L}_{\mathbf{R}}$ specified by sum of points $\sum_i \lambda_i p_i$, where $p_i \in B_3$ lie on $C_{\mathbf{R}}$.
    $\Longrightarrow$ which p_i are rationally equivalent? Even non-trivial if $\mathcal{L}_{\mathbf{R}}$ is pull-back bundle!
  - Both $C_{\mathbf{R}}$ and $\mathcal{L}_{\mathbf{R}}$ can simultaneously change with complex structure.
- Instead of training algorithm with optimal predictive power, we want to be able to interpret its strategy geometrically $\Longrightarrow$ use binary decision trees.

Decision trees

- A decision tree is a directed, connected graph with unique root vertex/node.
  Binary tree: each node has either 0 or 2 sub-nodes. Nodes with no sub-nodes are "leaves".
- Data organized by numeric features $\vec{x}$. Decision tree "classifies" input with *splitting criteria* at each node n:
  if $x_j \leq \kappa_j^{(n)}$, then input assigned to one sub-node, otherwise to the other sub-node.
- At the leaves, all assigned inputs ideally of same class (for us: h^0 "generic" or jumps).
  However, in general not possible; failure measured by *Gini impurity* ($\sim$how many different classes are assigned to node).
- For training: minimize Gini impurity for given training data.

The data set

- Data collected from curves with genus $1 \leq g \leq 6$ on $S = dP_3$.
  - Each *curve class* $[C] \longleftrightarrow \{P := \sum_k a_k m_k = 0\} \subset S$, with homogeneous monomials m_k.
  - For each class $[C]$, we consider up to 13 line bundles $L \in \text{Pic}(S)$.
  - For each pair $([C], L)$ (fixed χ), we compute $h^0(C_a, L|_{C_a})$ via CAP, where C_a is curve of class $[C]$, determined by a choice $a_k \in \{0, 1\}$ for the coefficients a_k of P.

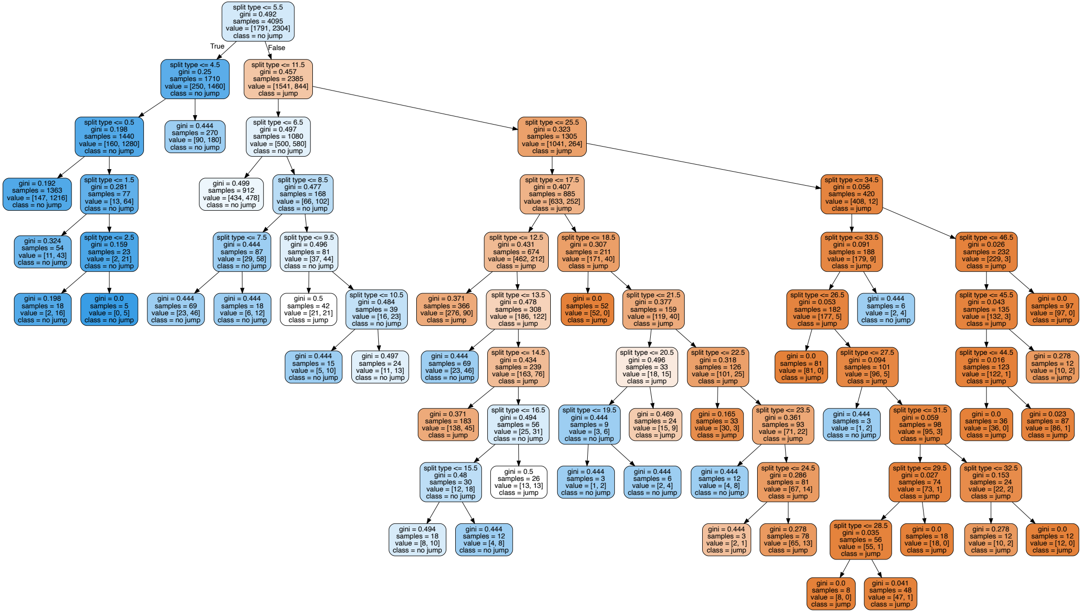
The data set

- Data collected from curves with genus $1 \leq g \leq 6$ on $S = dP_3$.
  - Each *curve class* $[C] \longleftrightarrow \{P := \sum_k a_k m_k = 0\} \subset S$, with homogeneous monomials m_k.
  - For each class $[C]$, we consider up to 13 line bundles $L \in \text{Pic}(S)$.
  - For each pair $([C], L)$ (fixed χ), we compute $h^0(C_a, L|_{C_a})$ via CAP, where C_a is curve of class $[C]$, determined by a choice $a_k \in \{0, 1\}$ for the coefficients a_k of P.
- For genus 1 only 127 data points per pair $([C], L)$, while for genus 6 we have $\sim 260\,000$. Lowest value h^0 considered generic, anything above classified as a "jump".

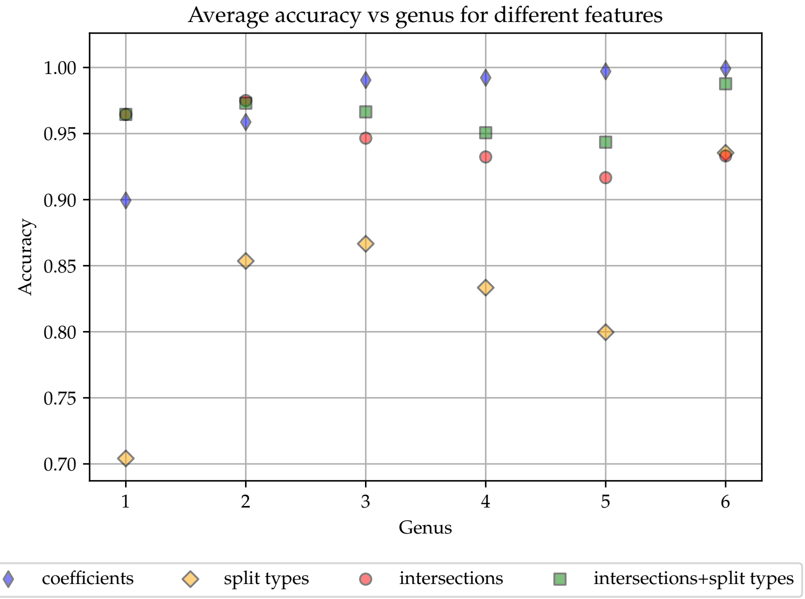
The data set

- Data collected from curves with genus $1 \leq g \leq 6$ on $S = dP_3$.
  - Each *curve class* $[C] \longleftrightarrow \{P := \sum_k a_k m_k = 0\} \subset S$, with homogeneous monomials m_k.
  - For each class $[C]$, we consider up to 13 line bundles $L \in \mathrm{Pic}(S)$.
  - For each pair $([C], L)$ (fixed χ), we compute $h^0(C_a, L|_{C_a})$ via CAP, where C_a is curve of class $[C]$, determined by a choice $a_k \in \{0, 1\}$ for the coefficients a_k of P.
- For genus 1 only 127 data points per pair $([C], L)$, while for genus 6 we have $\sim 260\,000$. Lowest value h^0 considered generic, anything above classified as a "jump".
- Consider different features of data: coefficients a_k, "split-type" (topology of C_a), "intersection" ($\Gamma_I \cdot L$, where Γ_I is component of C_a), "intersection + split-type".
- Based on feature, train tree to classify whether input is generic or has a jump. Training-testing data ratio: 90:10.

Example of tree trained on split-type ($g = 3, d = 3$)



Average accuracy



Interpretation of result

- Training on coefficients reach near perfect performance.
  Expected since coefficients entirely specifies setup, but no "intuitive" understanding.
- Topological criteria work surprisingly well (combining split-type and intersection numbers around and above 95% accuracy). Better suited for "extrapolation" to higher genus!
- Based on our data (without any further algebraic geometry considerations): $h^0(C_a, L|_{C_a})$ more likely to jump if $C_a = \tilde{C}_a \cup \mathbb{P}^1$.
- Small fraction of failure of topological criteria $\Longrightarrow$ other sources of jumps in cohomology. These are likely to be under-represented due to bias in our data set.

A toy F-theory model

- Compact fourfold $Y_4 \to B_3$ with B_3 a hypersurface in toric space. F-theory on Y_4 has $SU(5) \times U(1)$ gauge symmetry, with $SU(5)$ supported on $S \cong dP_3 \subset B_3$ [Bies, 17].
- In a "$U(1)$-flux" gauge background [Grimm/Weigand, 10], we have chiral spectrum:

$$\chi(\mathbf{10}_1) = 3\,, \quad \chi(\mathbf{5}_{-2}) = -18\,, \quad \chi(\mathbf{5}_3) = 15.$$

Focus on $C_{\mathbf{5}_3} \equiv C$, with $g = 24$, polynomial has 44 coefficients; $\deg(\mathcal{L}_{\mathbf{5}_3}) = 38$.

A toy F-theory model

- Compact fourfold $Y_4 \to B_3$ with B_3 a hypersurface in toric space. F-theory on Y_4 has $SU(5) \times U(1)$ gauge symmetry, with $SU(5)$ supported on $S \cong dP_3 \subset B_3$ [Bies, 17].
- In a "$U(1)$-flux" gauge background [Grimm/Weigand, 10], we have chiral spectrum:
$$\chi(\mathbf{10}_1) = 3\,, \quad \chi(\mathbf{5}_{-2}) = -18\,, \quad \chi(\mathbf{5}_3) = 15.$$
Focus on $C_{\mathbf{5}_3} \equiv C$, with $g = 24$, polynomial has 44 coefficients; $\deg(\mathcal{L}_{\mathbf{5}_3}) = 38$.
- Finding splitting $C \to \tilde{C} \cup \mathbb{P}^1$ easy in this case (dP_3 has 6 rigid divisors).
    - Only two of them ($E_{1,2}$) lead to jump when split off. They have $L \cdot E_{1,2} < -1$.
    - Splitting off E_1 and E_2 in combination, can get $h^0 = 17, 18, ..., 20$.

A toy F-theory model

- Compact fourfold $Y_4 \to B_3$ with B_3 a hypersurface in toric space. F-theory on Y_4 has $SU(5) \times U(1)$ gauge symmetry, with $SU(5)$ supported on $S \cong dP_3 \subset B_3$ [Bies, 17].
- In a "$U(1)$-flux" gauge background [Grimm/Weigand, 10], we have chiral spectrum:
$$\chi(\mathbf{10}_1) = 3\,, \quad \chi(\mathbf{5}_{-2}) = -18\,, \quad \chi(\mathbf{5}_3) = 15.$$
Focus on $C_{\mathbf{5}_3} \equiv C$, with $g = 24$, polynomial has 44 coefficients; $\deg(\mathcal{L}_{\mathbf{5}_3}) = 38$.
- Finding splitting $C \to \tilde{C} \cup \mathbb{P}^1$ easy in this case (dP_3 has 6 rigid divisors).
  - Only two of them ($E_{1,2}$) lead to jump when split off. They have $L \cdot E_{1,2} < -1$.
  - Splitting off E_1 and E_2 in combination, can get $h^0 = 17, 18, ..., 20$.
  - Cannot get $h^0 = 16$ this way!

Origin of jumps

- $C_a = \tilde{C} \cup C'$: count $s \in H^0(C_a, L|_{C_a})$ via "gluing" conditions along $\tilde{C} \cap C' \neq \emptyset$.
  In particular, sufficient condition for jump when $C' \cong \mathbb{P}^1$ and
  $L \cdot \mathbb{P}^1 < -1$, $D_L \cdot [C_a] > 2g - 2$.

Origin of jumps

- $C_a = \tilde{C} \cup C'$: count $s \in H^0(C_a, L|_{C_a})$ via "gluing" conditions along $\tilde{C} \cap C' \neq \emptyset$. In particular, sufficient condition for jump when $C' \cong \mathbb{P}^1$ and $L \cdot \mathbb{P}^1 < -1$, $D_L \cdot [C_a] > 2g - 2$.
  - In such cases: curve C_a is non-generic.
  - Occurs often when (many) coefficients of P set to 0.

Origin of jumps

- $C_a = \tilde{C} \cup C'$: count $s \in H^0(C_a, L|_{C_a})$ via "gluing" conditions along $\tilde{C} \cap C' \neq \emptyset$.
  In particular, sufficient condition for jump when $C' \cong \mathbb{P}^1$ and
  $L \cdot \mathbb{P}^1 < -1$, $D_L \cdot [C_a] > 2g - 2$.
  - In such cases: curve C_a is non-generic.
  - Occurs often when (many) coefficients of P set to 0.
- Cases not predicted by decision tree: $\mathcal{L} = L|_{C_a}$ non-generic.
  E.g.: $\mathcal{L} = p_1 - p_2$, becomes trivial when $p_1 = p_2$ in $\mathrm{Pic}(C_a)$.
  - No change of topological data explains failure of trained algorithm.
  - Requires precise "alignment" of L and C_a, unlikely to achieve with $a_k \in \{0, 1\}$.

Origin of jumps

- $C_a = \tilde{C} \cup C'$: count $s \in H^0(C_a, L|_{C_a})$ via "gluing" conditions along $\tilde{C} \cap C' \neq \emptyset$. In particular, sufficient condition for jump when $C' \cong \mathbb{P}^1$ and $L \cdot \mathbb{P}^1 < -1$, $D_L \cdot [C_a] > 2g - 2$.
  - In such cases: curve C_a is non-generic.
  - Occurs often when (many) coefficients of P set to 0.
- Cases not predicted by decision tree: $\mathcal{L} = L|_{C_a}$ non-generic. E.g.: $\mathcal{L} = p_1 - p_2$, becomes trivial when $p_1 = p_2$ in $\mathrm{Pic}(C_a)$.
  - No change of topological data explains failure of trained algorithm.
  - Requires precise "alignment" of L and C_a, unlikely to achieve with $a_k \in \{0, 1\}$.
  - Non-genericity quantified by *Brill–Noether* theory.

Brill–Noether theory

Assume C is a smooth Riemann surface of genus g. The *Abel–Jacobi map* φ gives a map

$$\mathrm{Pic}_d(C) \xrightarrow{\varphi_d} J(C) \cong \mathbb{C}^g/\Lambda\,.$$

For $n \in \mathbb{Z}_{\geq 0}$, define $G^n_d = \{\varphi_d(\mathcal{L}) \,|\, h^0(C,\mathcal{L}) = n\} \subset J(C)$. Then Brill–Noether theory says:

$$\dim(G^n_d) \geq \rho(n,d) := g - n(n-(d-g+1)) \equiv g - h^0\, h^1\,.$$

Brill–Noether theory

Assume C is a smooth Riemann surface of genus g. The *Abel–Jacobi map* φ gives a map

$$\mathrm{Pic}_d(C) \xrightarrow{\varphi_d} J(C) \cong \mathbb{C}^g/\Lambda\,.$$

For $n \in \mathbb{Z}_{\geq 0}$, define $G_d^n = \{\varphi_d(\mathcal{L}) \,|\, h^0(C,\mathcal{L}) = n\} \subset J(C)$. Then Brill–Noether theory says:

$$\dim(G_d^n) \geq \rho(n,d) := g - n(n - (d - g + 1)) \equiv g - h^0\, h^1\,.$$

If $\rho(n,d) > 0$, then there are degree d line bundles with $h^0 = n$ on C.

Brill–Noether theory

Assume C is a smooth Riemann surface of genus g. The *Abel–Jacobi map* φ gives a map

$$\mathrm{Pic}_d(C) \xrightarrow{\varphi_d} J(C) \cong \mathbb{C}^g/\Lambda\,.$$

For $n \in \mathbb{Z}_{\geq 0}$, define $G^n_d = \{\varphi_d(\mathcal{L}) \,|\, h^0(C, \mathcal{L}) = n\} \subset J(C)$. Then Brill–Noether theory says:

$$\dim(G^n_d) \geq \rho(n,d) := g - n(n - (d - g + 1)) \equiv g - h^0\, h^1\,.$$

If $\rho(n,d) > 0$, then there are degree d line bundles with $h^0 = n$ on C.

For the F-theory toy-model ($g = 24, d = 38$, generically $h^0 = 15$): $\rho(16,38) = 24 - 16 \cdot 1 = 8$. So jump by 1 expected (and found) on *smooth* curve.

"Moduli" space of jumps

- Number of vector-like pairs induces a *stratification* of complex structure moduli space.
  - Brill–Noether describes strata with *smooth* curves, provides upper bound on h^0 [Watari, 16].
  - Can be violated by curve splitting.

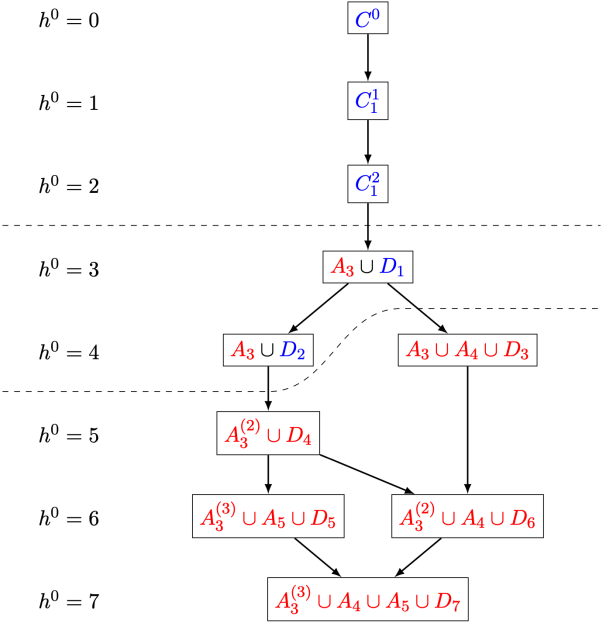
"Moduli" space of jumps

- Number of vector-like pairs induces a *stratification* of complex structure moduli space.
  - Brill–Noether describes strata with *smooth* curves, provides upper bound on h^0 [Watari, 16].
  - Can be violated by curve splitting.
- In global models: changing complex structure parameters affect genericity of line bundle (measured by Brill–Noether) and curve (determined by "split-type") democratically $\Longrightarrow$ strata of both types in moduli space.

"Moduli" space of jumps

- Number of vector-like pairs induces a *stratification* of complex structure moduli space.
  - Brill–Noether describes strata with *smooth* curves, provides upper bound on h^0 [Watari, 16].
  - Can be violated by curve splitting.
- In global models: changing complex structure parameters affect genericity of line bundle (measured by Brill–Noether) and curve (determined by "split-type") democratically $\Longrightarrow$ strata of both types in moduli space.
- For a given pair $([C], L)$, can be summarized in a Hasse-diagram.

Example ($g = 5, d = 4, \chi = 0$):



Summary

- Explicitly computing vector-like spectrum in global F-theory models is hard.
- Using machine learning techniques, can gain intuition about computationally challenging cases.
  - Qualitatively different than previous machine learning studies of line bundle cohomologies, because both line bundle and curve topology change simultaneously.
- Both changes source jumps in cohomologies, captured by Hasse-type diagrams.
  - Reflect fact that vector-like spectra induce stratification on complex structure moduli.

Summary

- Explicitly computing vector-like spectrum in global F-theory models is hard.
- Using machine learning techniques, can gain intuition about computationally challenging cases.
  - Qualitatively different than previous machine learning studies of line bundle cohomologies, because both line bundle and curve topology change simultaneously.
- Both changes source jumps in cohomologies, captured by Hasse-type diagrams.
  - Reflect fact that vector-like spectra induce stratification on complex structure moduli.

See also talk by Martin Bies in "Summer Series on String Pheno", June 16.

Open problems of ...

- ...technical nature: extend to non-pull-back & "fractional" pull-back bundles, combine stratification diagrams for several curves in one global model, ...
- ...conceptual nature: compute vector-like spectrum for pseudo-real representations, incorporate gauge backgrounds with non-vertical G_4 (flux moduli dependence!), (geometric) symmetries protecting vector-like pairs, ...
- ...practical nature: apply to model building, (S)CFTs, swampland program, ...

Open problems of ...

- ...technical nature: extend to non-pull-back & "fractional" pull-back bundles, combine stratification diagrams for several curves in one global model, ...
- ...conceptual nature: compute vector-like spectrum for pseudo-real representations, incorporate gauge backgrounds with non-vertical G_4 (flux moduli dependence!), (geometric) symmetries protecting vector-like pairs, ...
- ...practical nature: apply to model building, (S)CFTs, swampland program, ...

Thank you!