Machine Learning Landscape - ICTP Trieste

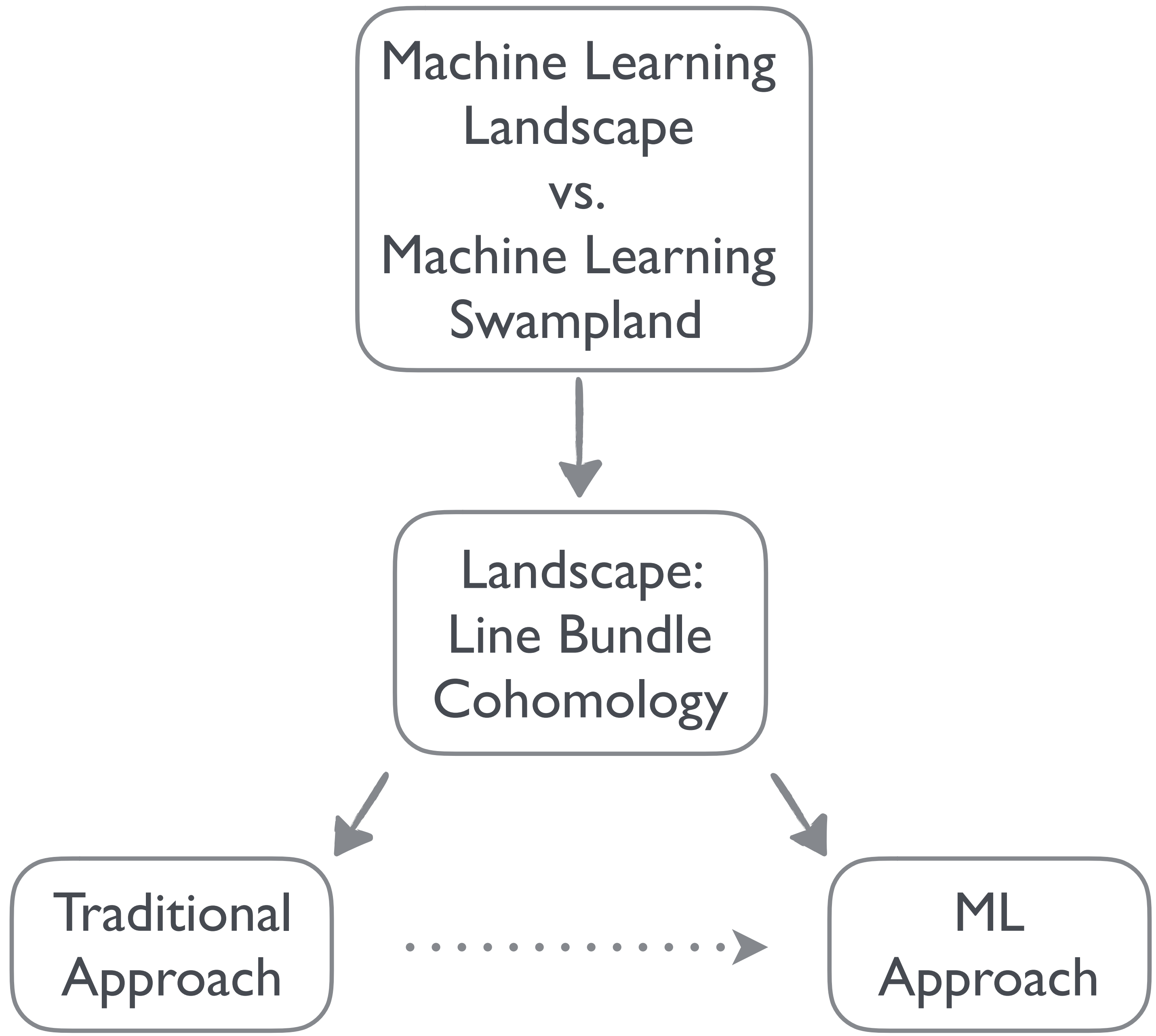
Machine Learning Line Bundle Cohomology

Daniel Kläwer

Max-Planck-Institut für Physik

DK, Lorenz Schlechter — arXiv:1809.02547

Outline


Machine Learning Landscape vs. Machine Learning Swampland
Landscape: Line Bundle Cohomology
Traditional Approach
ML Approach

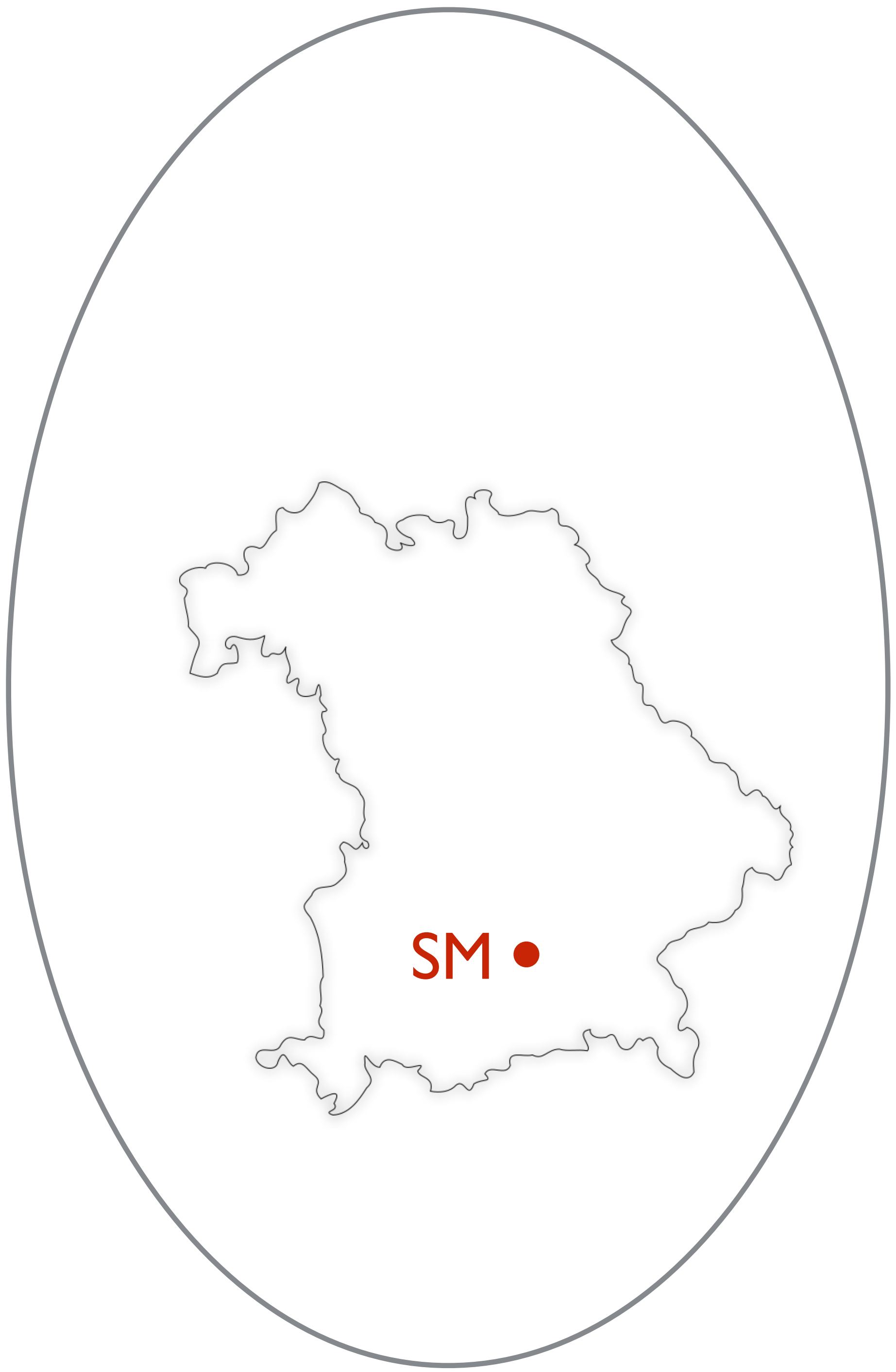
Size of the Landscape

- The String Landscape is **vast.**
- Classic estimate for # of type IIB flux vacua: $\# > 10^{500}$
- Exist even much larger numbers in the literature: $\# > 10^{272{,}000}$ Taylor, Wang 2015
- Proof of finiteness does not exist! But it is known that the # of elliptically fibered CY threefolds is finite — Grassi 1991; Gross 1993
- Recent work suggests that the # of Calabi-Yau threefolds (needed for compactification of type II/het.) could in fact be finite, because "almost all" are elliptically fibered — Anderson, Gao, Gray, Lee 2017; Huang, Taylor 2018

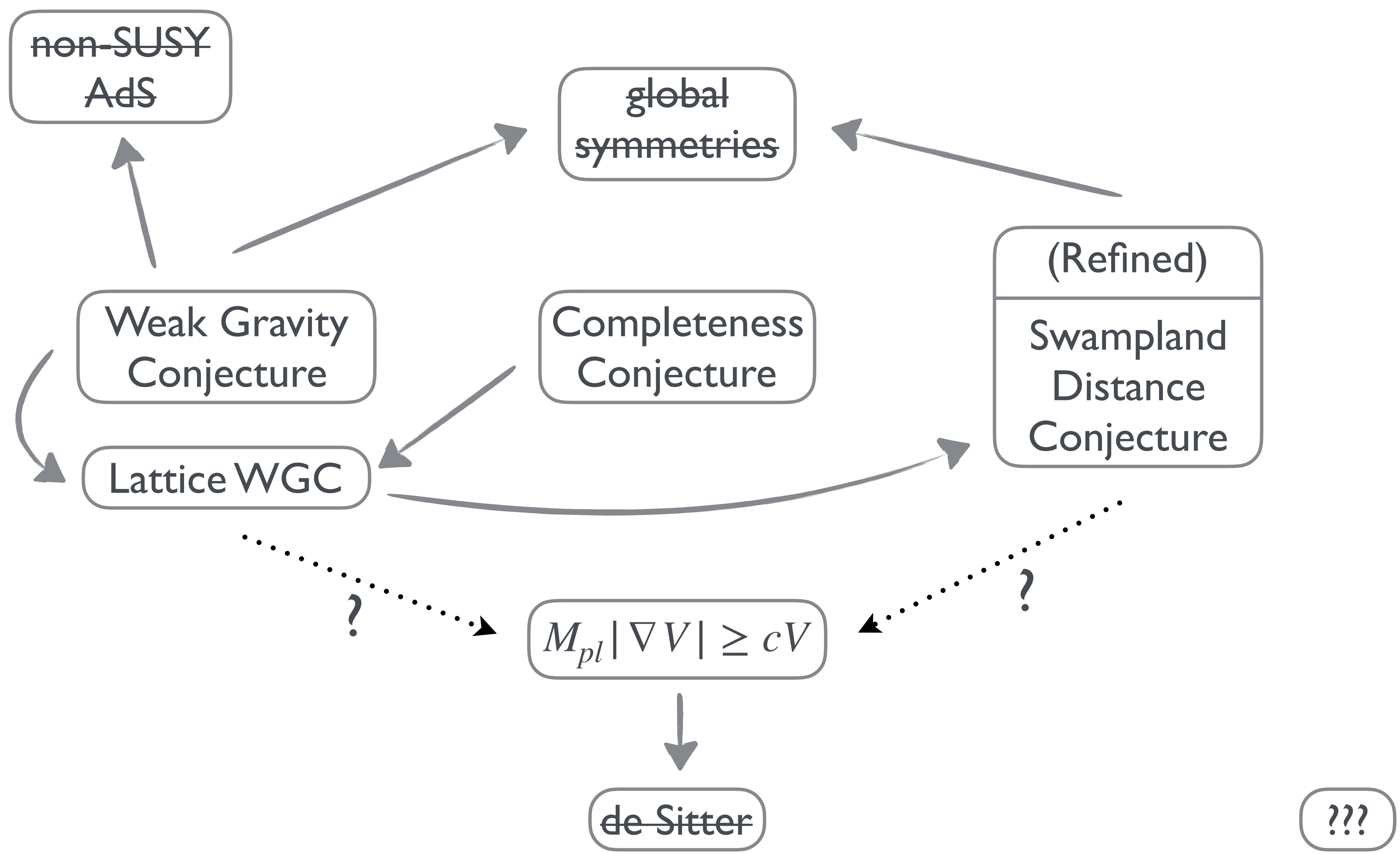
Landscape vs. Swampland

Vafa 2005

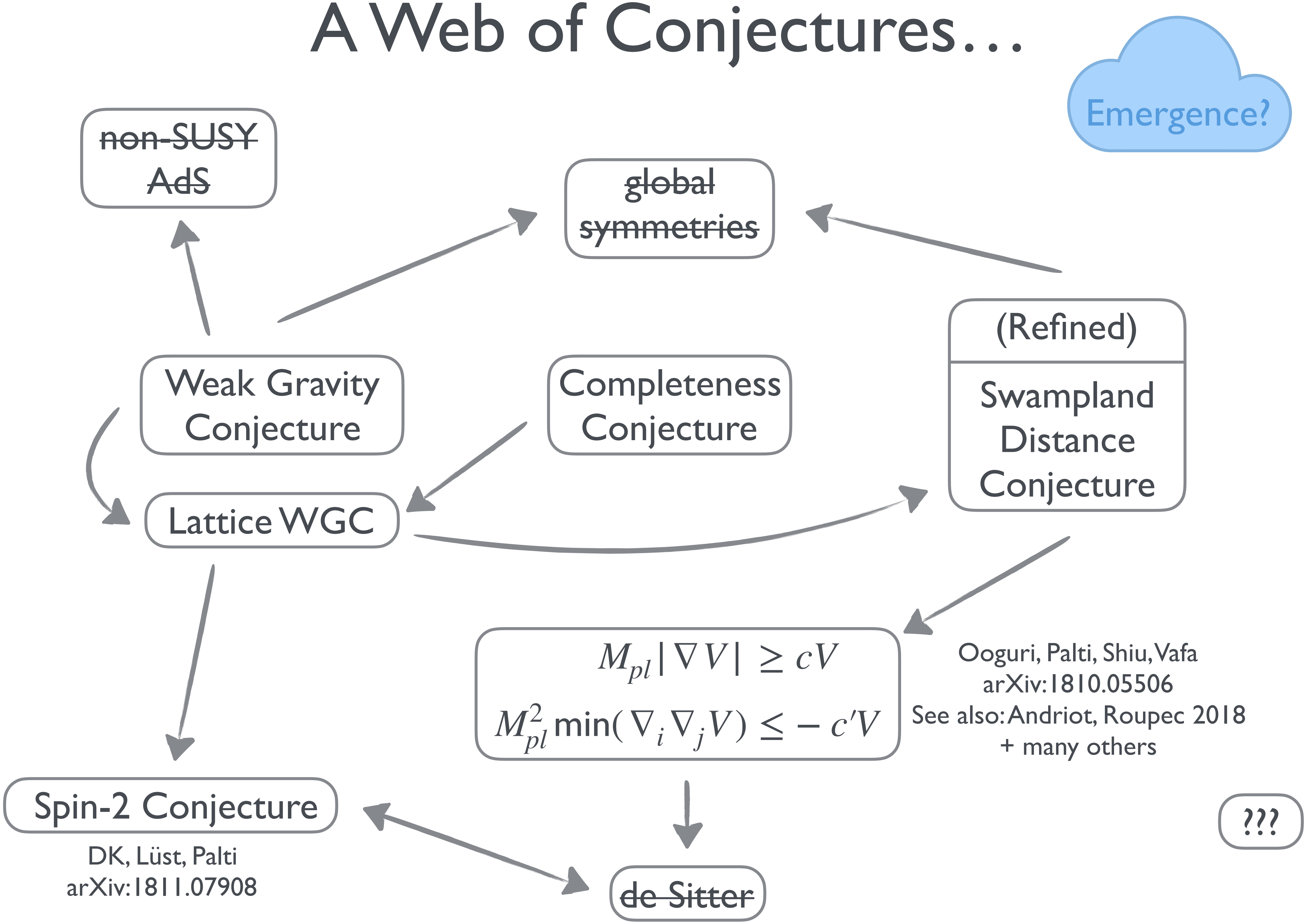
- **String Landscape:** Set of effective field theories that can be UV completed to a string theory vacuum
- **String Swampland:** The complement of seemingly consistent effective field theories which do not arise from string theory
- **Swampland Conjectures:** Conjectured properties of theories that are able to discriminate between the landscape and swampland



A Web of Conjectures…



A Web of Conjectures…



Problems of the Landscape Type

- Often study a highly nonlinear map of the type:

input data $\mathbb{Z}^I$ $\longrightarrow$ properties of vacua $\mathbb{Z}^O$

Input data	Properties of vacua
Toric data	SUSY?
Fluxes	Gauge group
Vector Bundles	# representations
# branes	Vacuum energy
$\vdots$	$\vdots$

- Possible questions:
  - Classify inequivalent input data
  - SM problem: find pre-images of given $U(x) \subset \mathbb{Z}^O$
  - Find approximate parameter distributions that emulate the distribution in string vacua
  - $\vdots$

Problems of the Swampland Type

- The known swampland conjectures are mostly of the following form:
  - A bound on a certain quantity in the Landscape $Q(\mathbb{Z}^I) < \mathcal{O}(1)$
  - Exclusion of a property $(\neg\text{SUSY}) \cup (V < 0) \cup \text{stable} = \text{False}$
- Possible Questions:
  - Can we violate it parametrically?
  - Can we bend it? What is the $\mathcal{O}(1)$ number?
  - How likely is it that the inequality is $\approx$ saturated?
- Very important to give generic predictions from string theory!

CYs and Toric Varieties

- Anti-canonical hypersurfaces in toric varieties → many Calabi-Yau manifolds
- Prototype: $\mathbb{C}P^{I-1}$ $[x_1 : \ldots : x_I]$ $\sum_{i=1}^{5} x_i^5 = 0$ in $\mathbb{C}P^4$
- More general: toric variety described in terms of I homogeneous coordinates and R scaling relations
- Example: K3 hypersurface in weighted projective space $\mathbb{P}^3_{1112}$

$$[x_1 : x_2 : x_3 : x_4 : x_5] \sim [\alpha x_1 : \alpha x_2 : \alpha x_3 : x_4 : \alpha^2 x_5]$$
$$[x_1 : x_2 : x_3 : x_4 : x_5] \sim [x_1 : x_2 : x_3 : \beta x_4 : \beta x_5]$$

- Line bundles = divisors: $\mathscr{L} = \mathscr{O}_X(D) = \mathscr{O}_X(m, n)$ $D = mD_1 + nD_2$

$D_1 = \{x_1 = 0\} \sim \{x_2 = 0\} \sim \{x_3 = 0\}$ $D_2 = \{x_4 = 0\}$

CYs and Toric Varieties

- Anti-canonical hypersurfaces in toric varieties $\longrightarrow$ many Calabi-Yau manifolds
- Prototype: $\mathbb{C}P^{I-1}$ $[x_1 : \ldots : x_I]$ $\sum_{i=1}^{5} x_i^5 = 0$ in $\mathbb{C}P^4$
- More general: toric variety described in terms of I homogeneous coordinates and R scaling relations
- Example: K3 hypersurface in weighted projective space $\mathbb{P}^3_{1112}$

$$[x_1 : x_2 : x_3 : x_4 : x_5] \sim [\alpha x_1 : \alpha x_2 : \alpha x_3 : x_4 : \alpha^2 x_5]$$
$$[x_1 : x_2 : x_3 : x_4 : x_5] \sim [x_1 : x_2 : x_3 : \beta x_4 : \beta x_5]$$

- Line bundles = divisors: $\mathscr{L} = \mathcal{O}_X(D) = \mathcal{O}_X(m,n)$ $D = mD_1 + nD_2$

$D_1 = \{x_1 = 0\} \sim \{x_2 = 0\} \sim \{x_3 = 0\}$ $D_2 = \{x_4 = 0\}$

Line Bundle Cohomology - Why?

- CY compactifications of the heterotic string or F-theory: interested in sheaf cohomology groups $H^i(X, \mathscr{L})$, where X is the CY itself or a sub-manifold of it.
- Dimensions $h^i(X, \mathscr{L})$ count massless modes in the 4d theory
- Heterotic String/F-theory: e.g. chiral fermions
- For us, they simply define a non-linear map of the type discussed before

$$h^i\left(X, \mathcal{O}_X(m_1, \ldots, m_I)\right) : \ \mathbb{Z}^I \to \mathbb{Z}^{dim(X)}, \quad i = 1, \ldots, dim(X)$$

Line bundle integers

Cohomology dimensions

Line Bundle Cohomology - Why?

- CY compactifications of the heterotic string or F-theory: interested in sheaf cohomology groups $H^i(X, \mathscr{L})$, where X is the CY itself or a sub-manifold of it.
- Dimensions $h^i(X, \mathscr{L})$ count massless modes in the 4d theory
- Heterotic String/F-theory: e.g. chiral fermions
- For us, they simply define a non-linear map of the type discussed before

$$h^i\left(X, \mathcal{O}_X(m_1, \ldots, m_I)\right) : \mathbb{Z}^I \to \mathbb{Z}^{dim(X)}, \quad i = 1, \ldots, dim(X)$$

Line bundle integers

Cohomology dimensions

The Traditional Approach

- Several algorithms are known for the computation of line bundle cohomology
- In various cases, exact formulae for toric varieties are known
- Not so much for hypersurfaces
- Analytic results for CICYs in products of projective spaces — Lukas, Constantin 2018
- Bott formula: $h^0\left(\mathbb{C}P^n, \mathcal{O}_{\mathbb{C}P^n}(k)\right) = \binom{k+n}{n}$ — Polynomial of degree n!
- We use the cohomCalg algorithm — Blumenhagen, Jurke, Rahn, Roschy 2010
- Computes for us cohomology of line bundles on the toric **ambient space** X

From Ambient Space to Hypersurface

- Line bundles on the ambient space $\mathcal{O}_X(D)$ pull back to the hypersurface $\mathcal{O}_H(D)$
- Relation is given by the Koszul sequence:

$$0 \to \mathcal{O}_X(D-H) \stackrel{\text{m}}{\to} \mathcal{O}_X(D) \stackrel{\text{res}}{\to} \mathcal{O}_H(D) \to 0$$

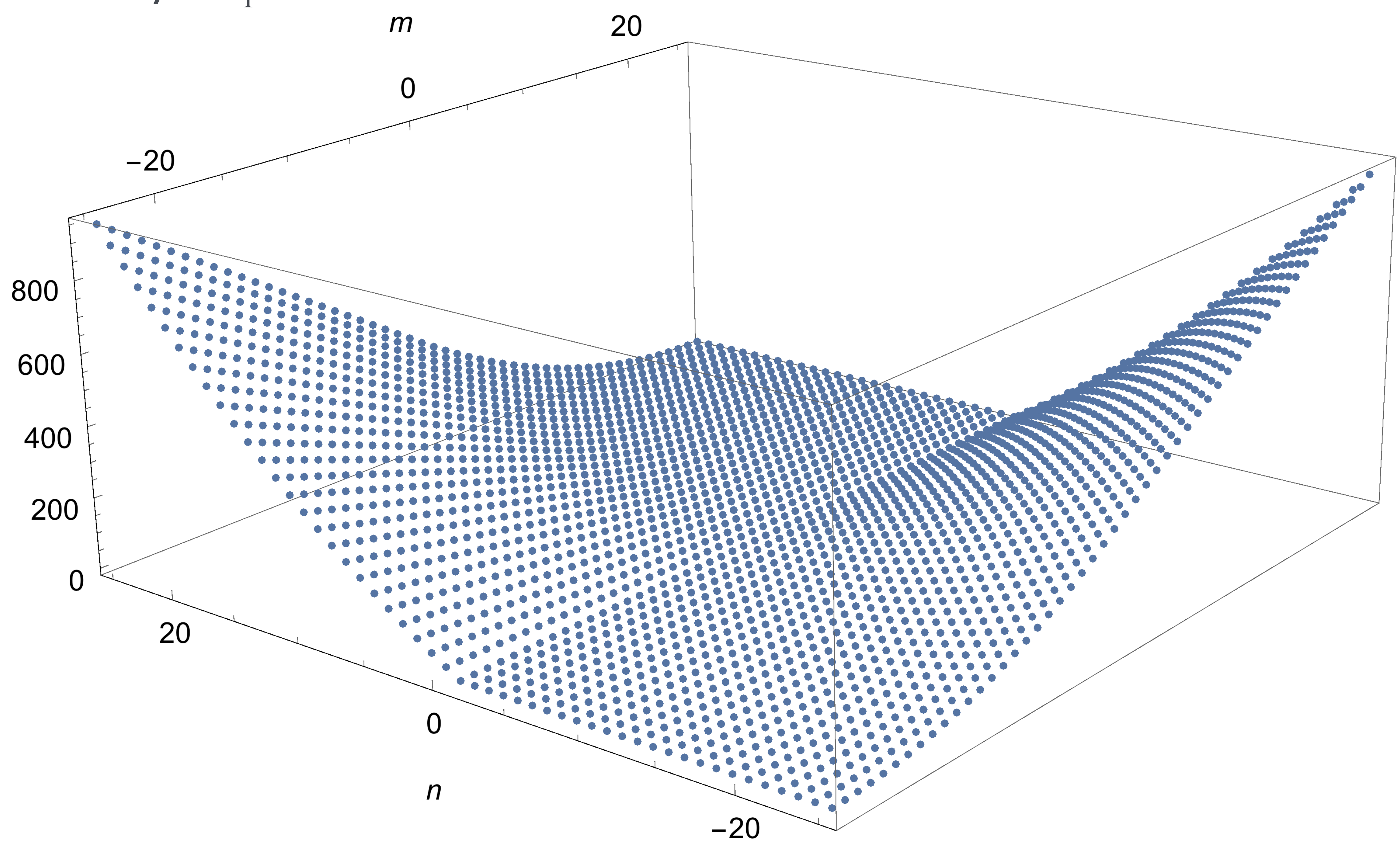
- Induces long exact sequence of cohomology groups

$$\cdots \to H^i(\mathcal{O}_X(D-H)) \stackrel{\text{m}}{\to} H^i(\mathcal{O}_X(D)) \stackrel{\text{res}}{\to} H^i(\mathcal{O}_H(D)) \stackrel{\delta}{\to} H^{i+1}(\mathcal{O}_X(D-H)) \to \cdots$$

- Strategy:
  1. Cut long sequence into short ones if we hit a zero
  2. Use the fact that the alternating sum of dimensions is zero
  3. Try to solve this linear system for the $h^i(\mathcal{O}_H(D))$ in terms of the $h^i(\mathcal{O}_X(D))$ and $h^i(\mathcal{O}_X(D-H))$
  4. If this procedure does not give a unique result, need to introduce further cuts and need info about maps!

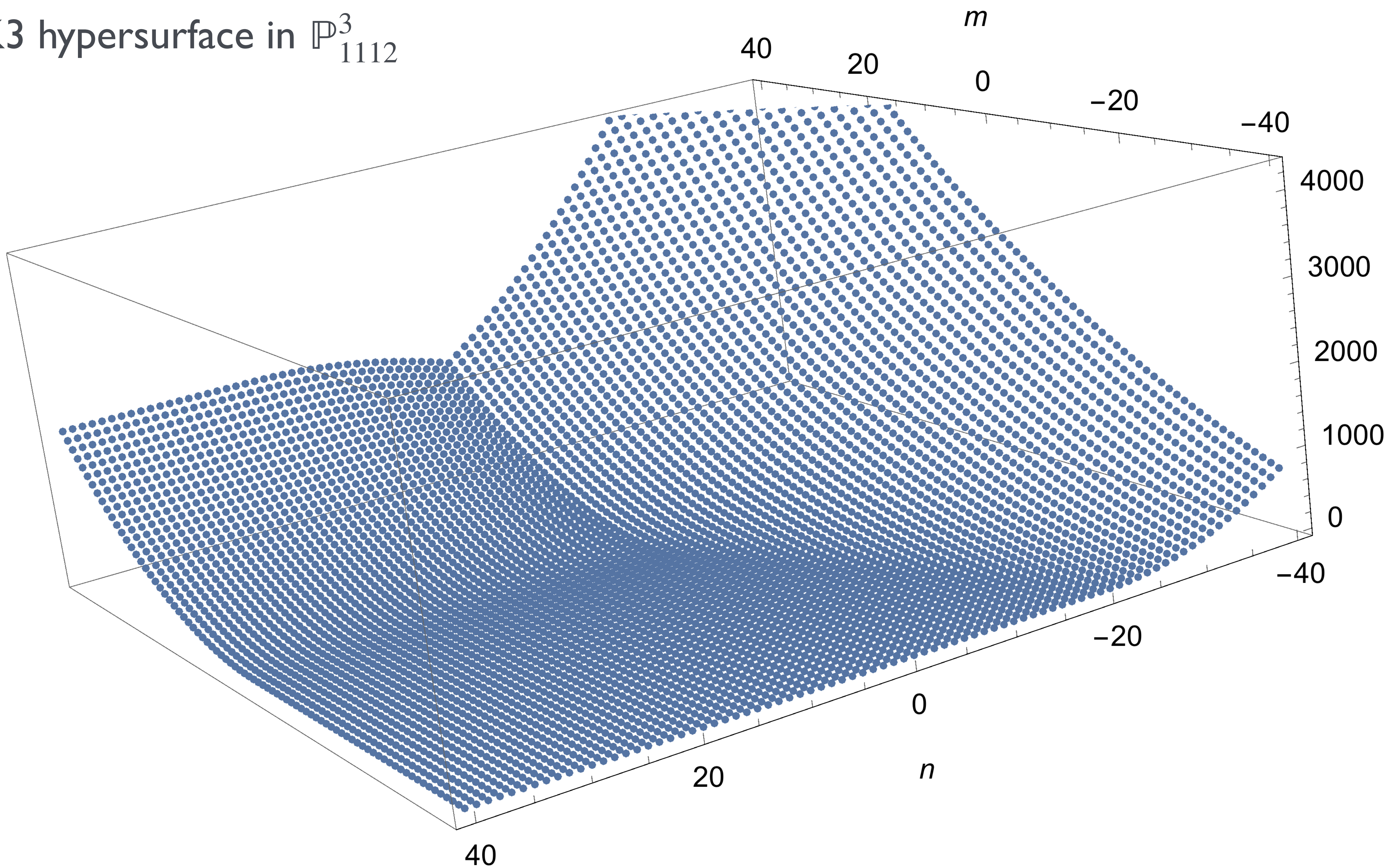
The Data

Toric variety: dP_1



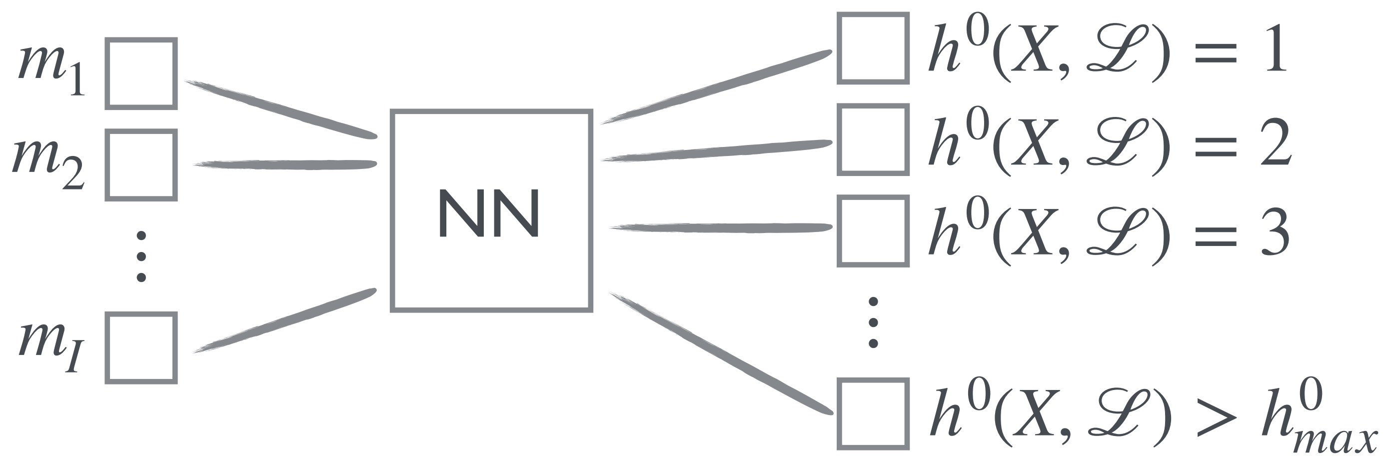
The Data

K3 hypersurface in $\mathbb{P}^3_{1112}$



NN Approaches for Learning Cohomolgy

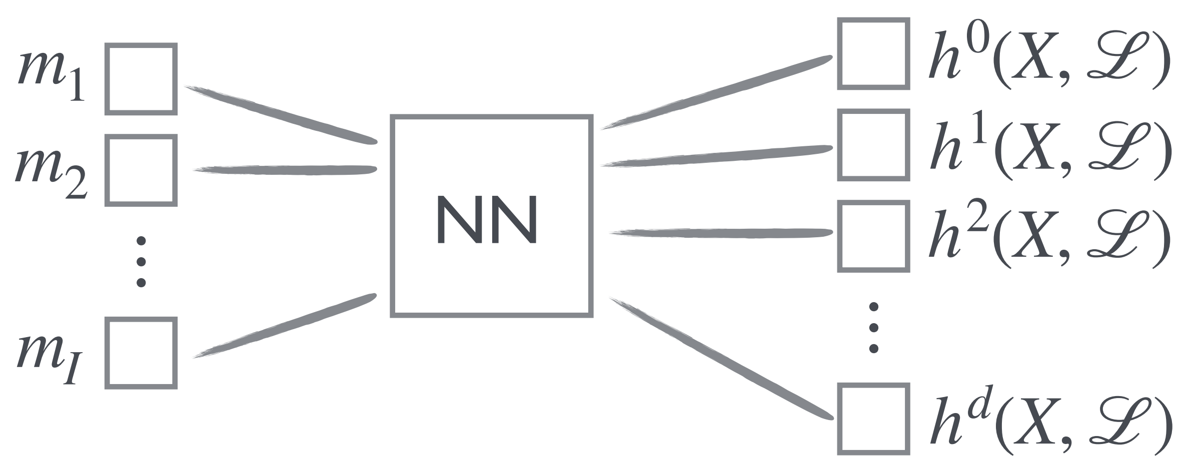
1. Classification Bull, He, Jejjala, Mishra 2018



- Output is probability for class membership
- Bad: Have to bin/cut off + output scales with m_i

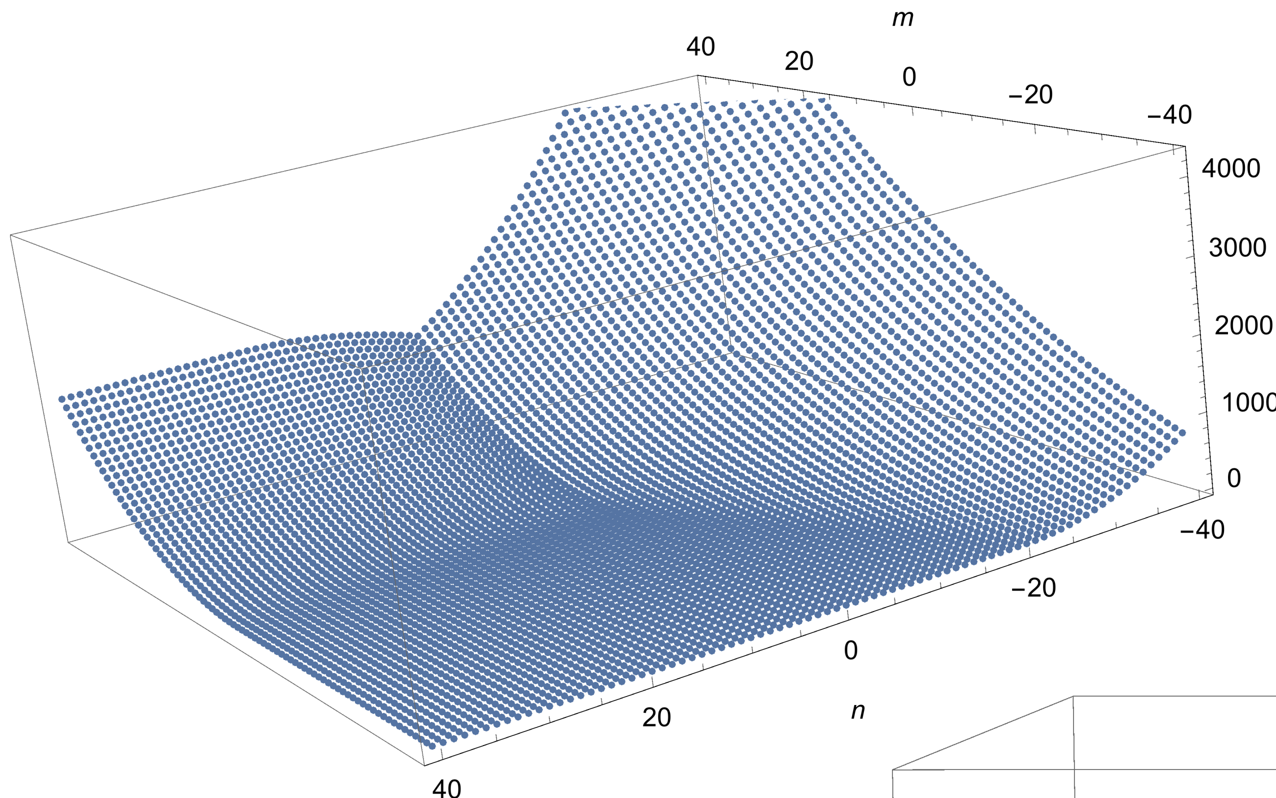
→ Cannot extrapolate to larger m_i than training set

2. Regression Ruehle 2017



- Output should reproduce directly the h^i
- Round to nearest integer
- Extrapolation limited by used float point precision
- Can make use of correlations between h^i

Regression by Neural Networks

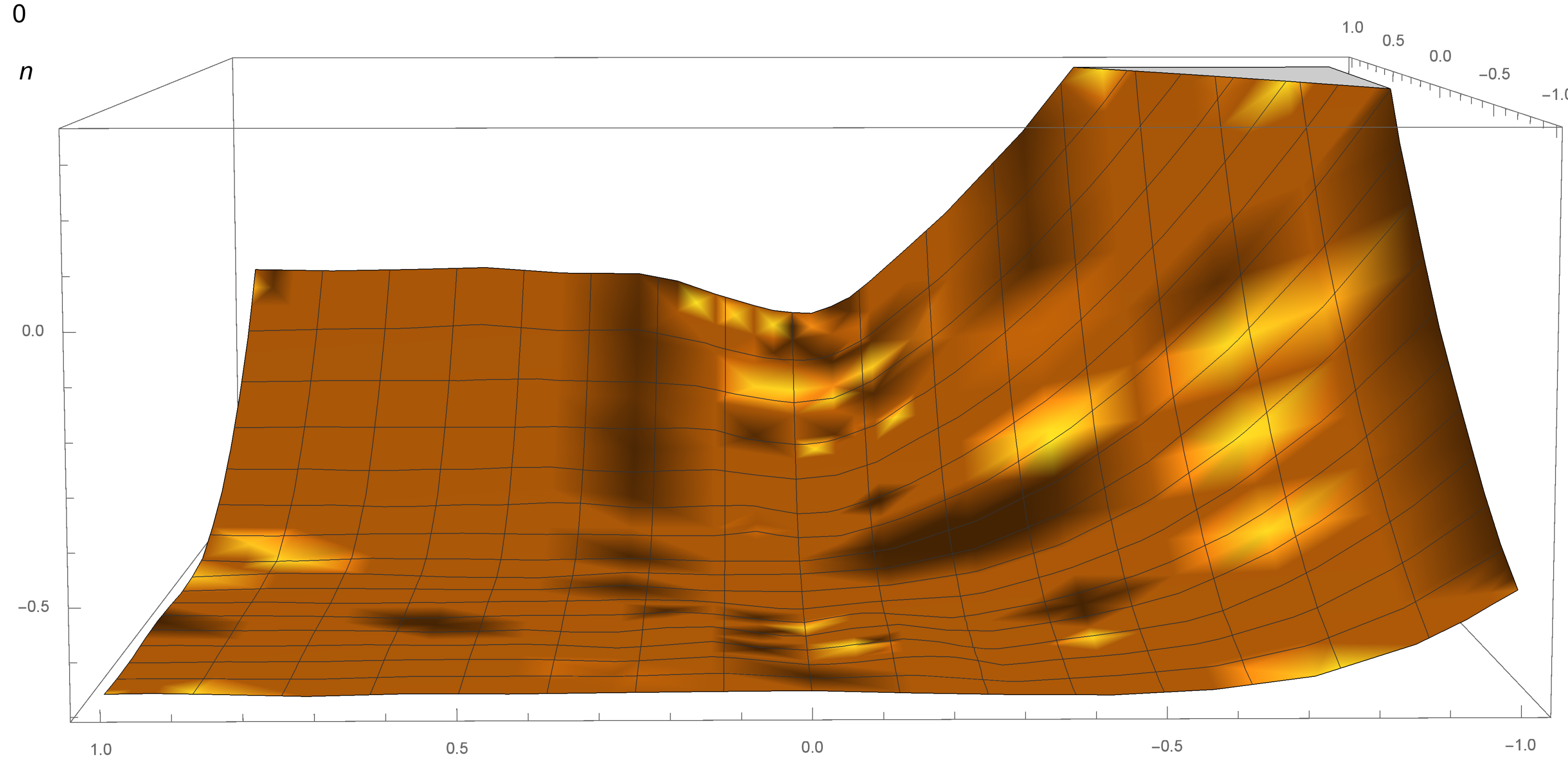


3 fully connected leaky ReLU layers
with 100 neurons
81^2=6561 data points
60% used for training
Normalized data
Batch size: 300
Training time: 2min on Desktop CPU

K3 hypersurface in $\mathbb{P}^3_{1112}$

$\approx$50% accuracy

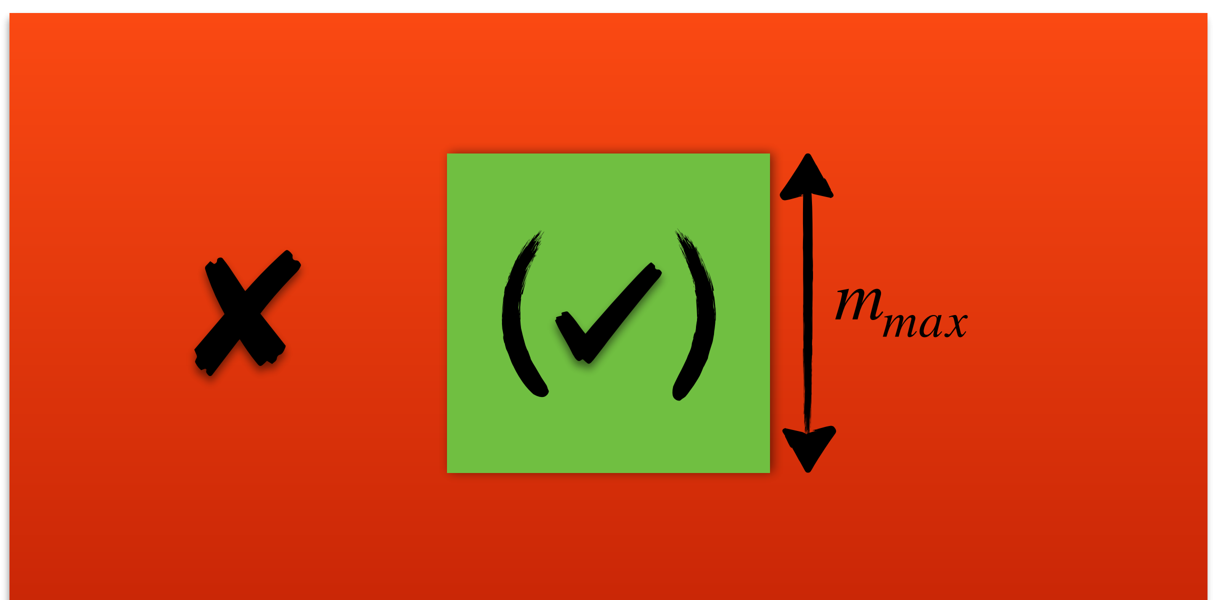
but: easy to learn
the zeros…



Regression by Neural Networks

Problems:

- Regression works for simple examples (almost 100% accuracy for a dP1 toric ambient space), but fails for more complicated ones (hypersurfaces).
- Later: partly due to high frequency oscillation in data
- Extrapolation fails also due to finite float precision



Can we do better? General lesson: **Yes**, if we understand our data…

Hirzebruch-Riemann-Roch

- The HRR theorem allows us to easily compute an analytic expression for the holomorphic Euler characteristic of a line bundle

$$\chi(X, \mathscr{L}) = \int_X ch(\mathscr{L}) td(X)$$

- The Euler characteristic is just the alternating sum of ranks

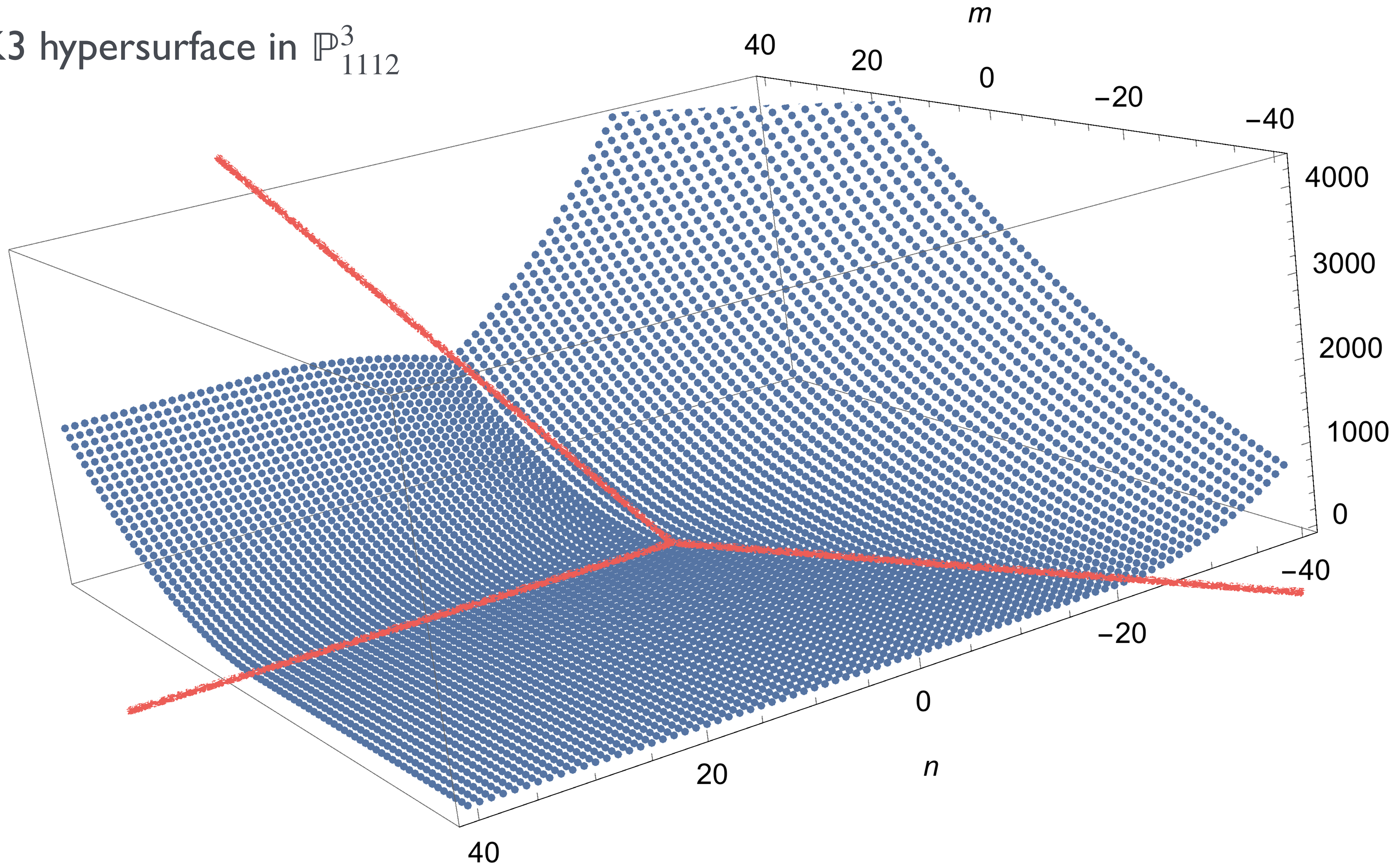
$$\chi(X, \mathscr{L}) = \sum_{i=0}^{dim_{\mathbb{C}}(X)} (-1)^i h^i(X, \mathscr{L})$$

- It is a **polynomial** of degree $dim_{\mathbb{C}}(X)$ in the line bundle integers

- If the alternating sum of cohomologies is that simple, it is reasonable to expect that locally they are also individually polynomial, with pairwise discontinuities or kinks at loci of codimension ≥ 1

A Closer Look at the Data…

K3 hypersurface in $\mathbb{P}^3_{1112}$



The Classification Problem

- Visually: Data is locally described by polynomial ✓
- Aim: identify the phase boundaries in the space of line bundle integers
- Look for locations where the map

$$h^i\left(X, \mathcal{O}_X(m_1, \ldots, m_I)\right) : \ \mathbb{Z}^I \to \mathbb{Z}^{dim(X)}, \quad i = 1,\ldots, dim(X)$$

  is not differentiable
- Get a cone structure in input space → not quite actually… (see later)
- Use unsupervised learning for this!
- Once done, just fit a polynomial within each phase!

The Algorithm

- The polynomials are of degree at most $d = dim_{\mathbb{C}}(X)$
- The (d+1)st derivatives of the map are only non-vanishing at phase boundary
- Run a classification on data set

$$\left\{ \vec{m}\,, \frac{\partial^{d+1} h^i}{\partial^{d+1} m_1}\,, \frac{\partial^{d+1} h^i}{\partial^{d} m_1 \partial m_2}\,, \ldots\ldots, \frac{\partial^{d+1} h^i}{\partial^{d+1} m_R} \right\}$$

- This separates the data into a large interior class and boundary classes
- After getting rid of the boundary classes, classify the data set

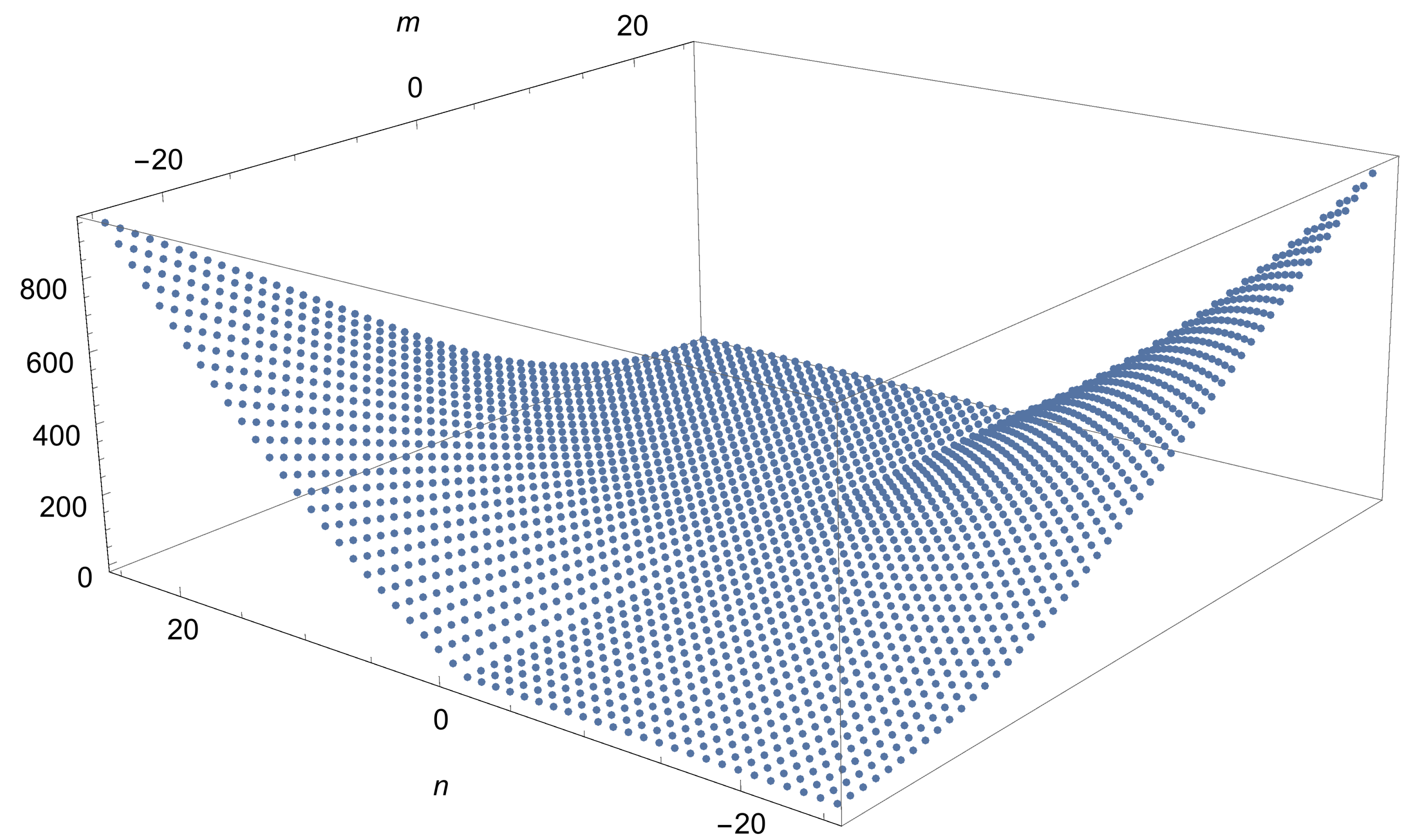
$$\left\{ \vec{m}\,, \frac{\partial^{d} h^i}{\partial^{d} m_1}\,, \frac{\partial^{d} h^i}{\partial^{d} m_1 \partial m_2}\,, \ldots\ldots, \frac{\partial^{d} h^i}{\partial^{d} m_R} \right\}$$

- This leads to a classification of interior phases with different polynomials

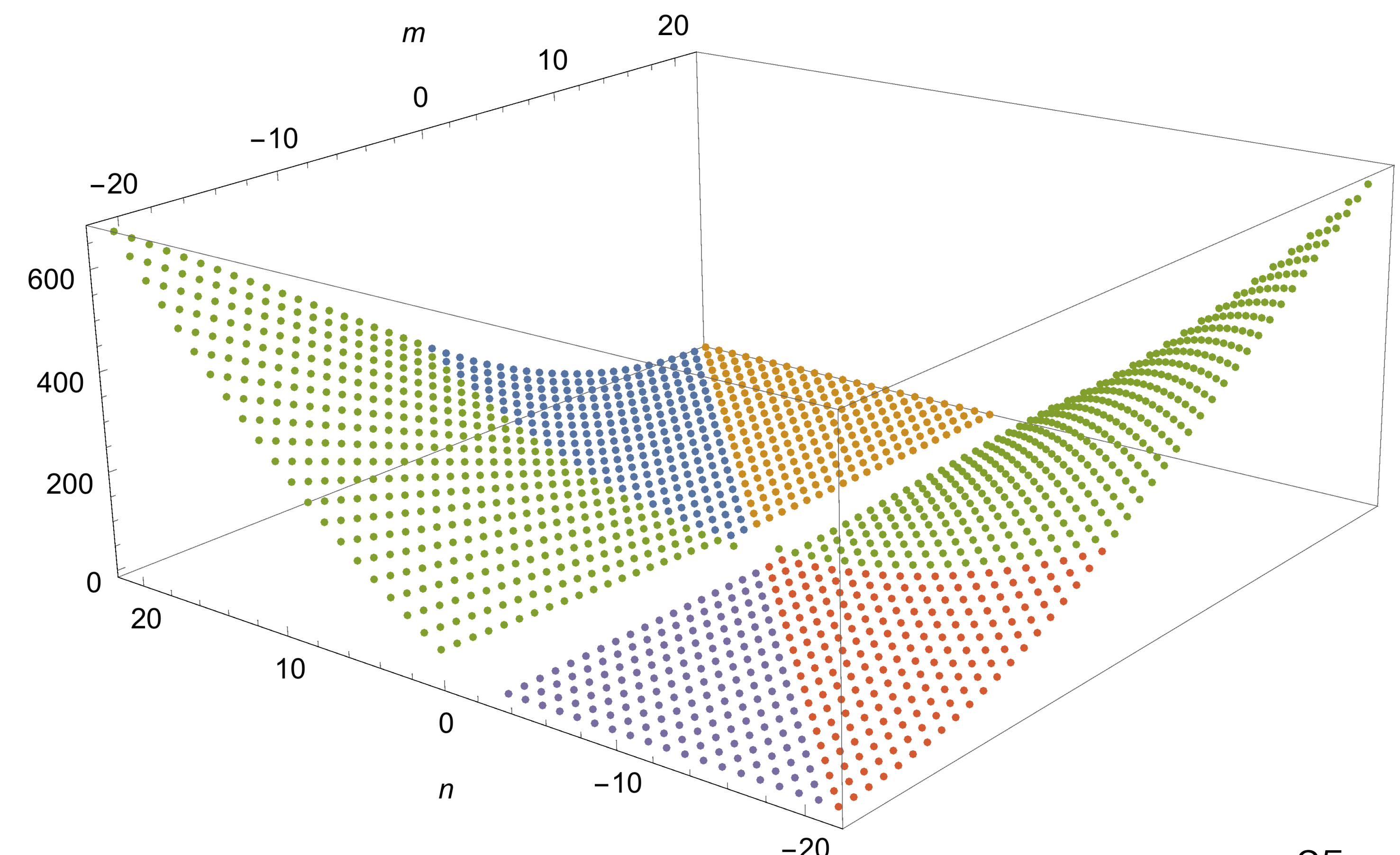
Implementation

- We used Mathematica 11.3
- Can use the Classify[] function with Method: NeuralNetwork
- It turned out that classical clustering algorithms like KMeans are more efficient
- The ClusterClassify[] function is used with options Method: KMeans
PerformanceGoal: Quality
- Use LinearModelFit[] to fit a polynomial of deg=dimension

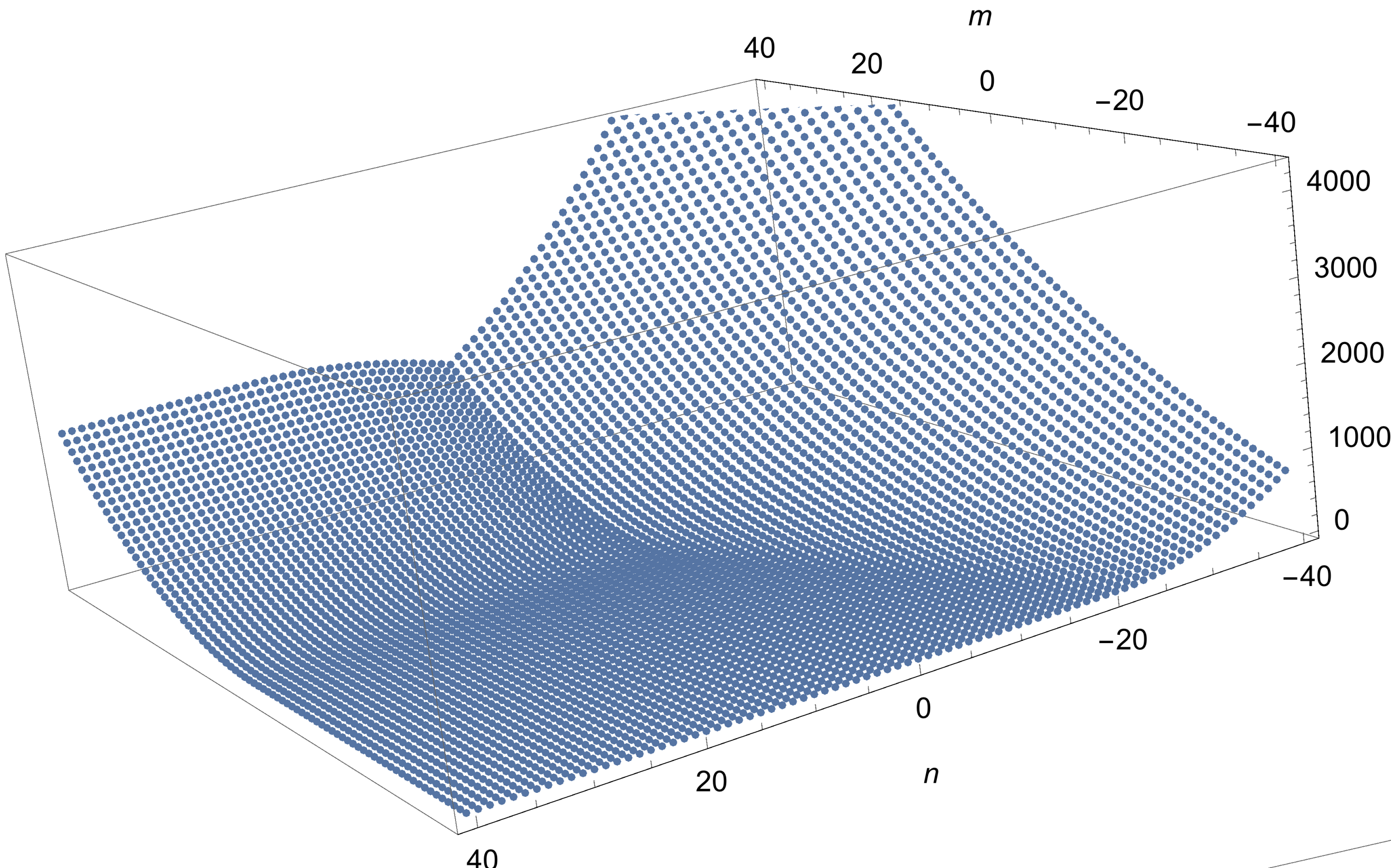
Example 1: dP_1



Works nicely, but so does the regression with a NN

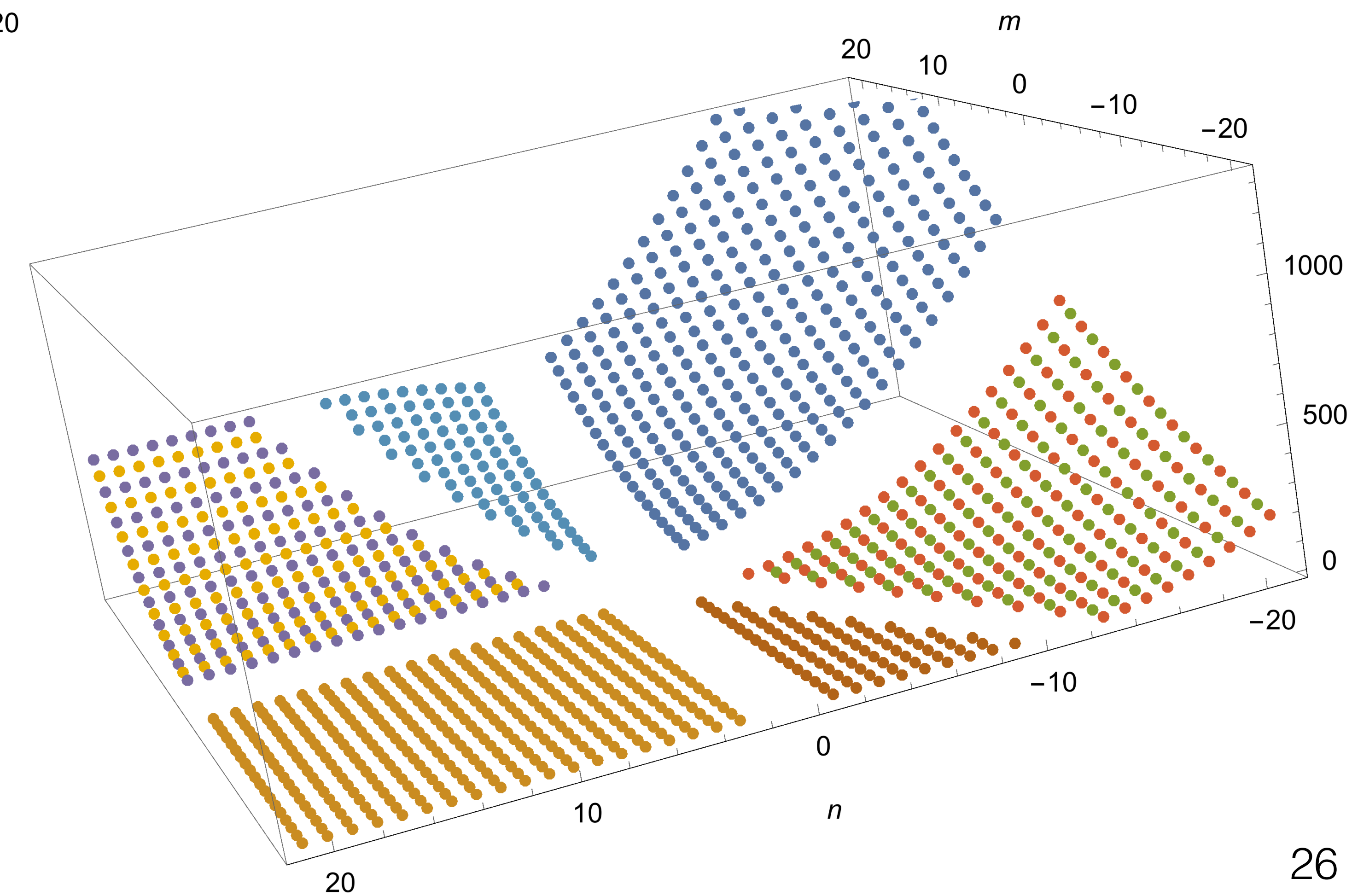


Example 2: K3 hypersurface in $\mathbb{P}^3_{1112}$

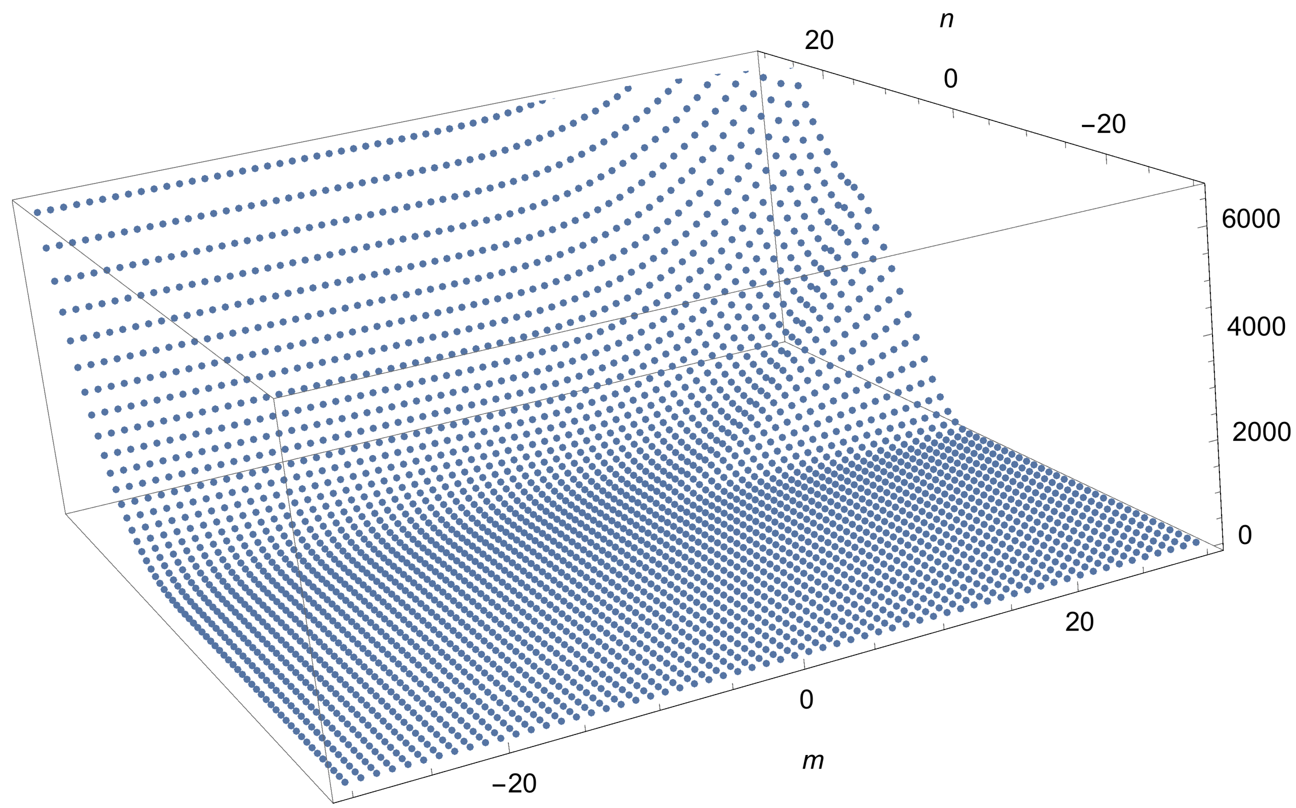


Remarkably, we find more than the visible three phases. In particular there is a $\mathbb{Z}_2$ modulation

Simple regression NN not well suited to reproduce the alternating phases! They are a generic feature.

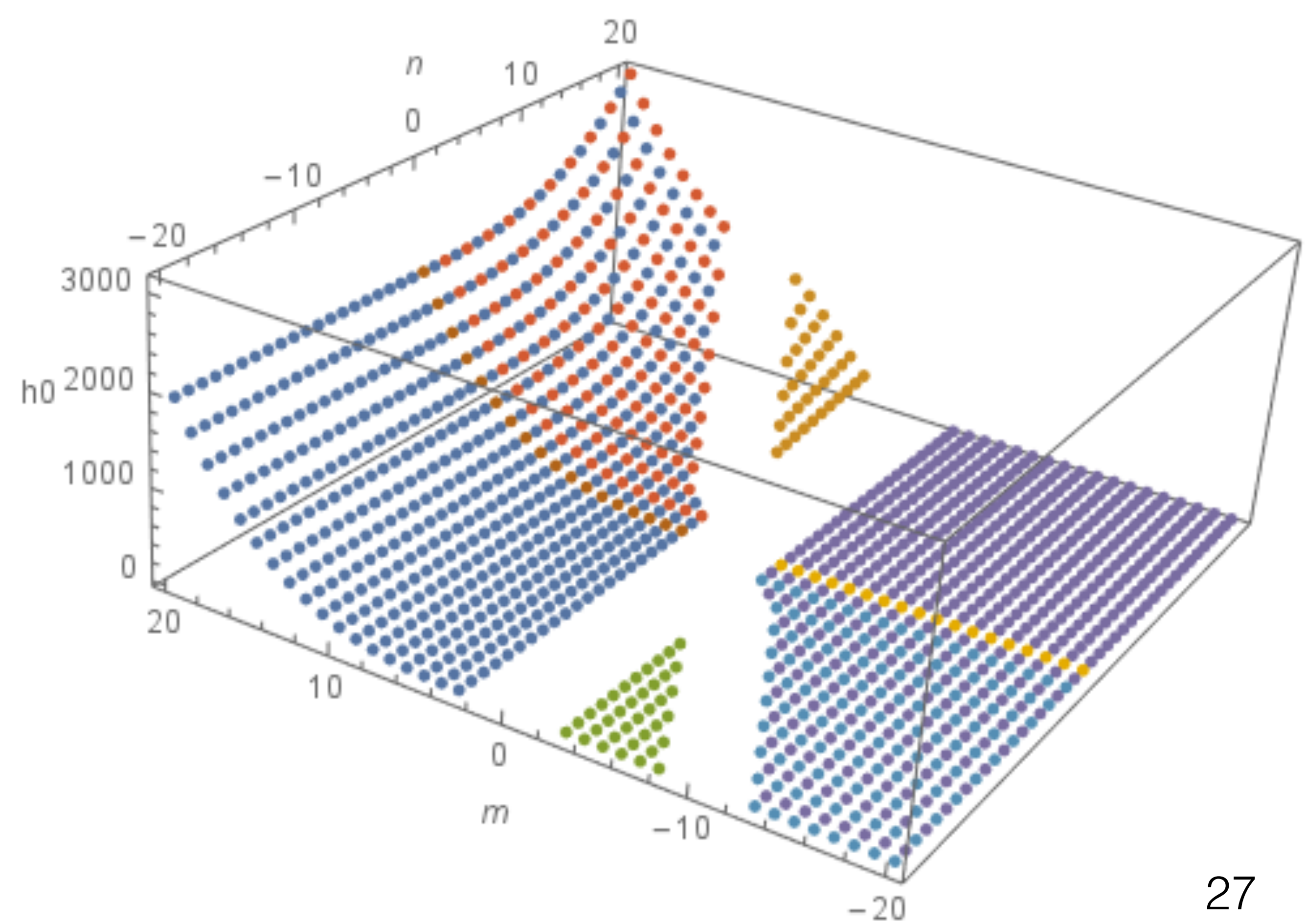


Example 3: CY3 hypersurface in $\mathbb{P}^4_{11222}$



Again, the $\mathbb{Z}_2$ modulation. For more complicated examples, we expect also $\mathbb{Z}_N$

$\mathbb{Z}_2$ modulation in constant phase needed for other h^i



Cross-check: HRR

- We fit polynomials with **rational coefficients.** This guarantees that we can extrapolate to arbitrarily large inputs if the fit is correct.
- If the alternating sum of these coefficients agrees with Hirzebruch-Riemann-Roch, it is a strong indication that the result is correct
- Check for K3 hypersurface in $\mathbb{P}^3_{1112}$

Phase	h^0	h^1	h^2
I	0	0	$-n^2+nm+m^2+2$
II	$-n^2+nm+m^2+2$	0	0
III	$\frac{5m^2}{4}+\frac{7}{4}$	$3n^2-3nm-3n+\frac{3m^2}{4}+\frac{3m}{2}+\frac{3}{4}$	$2n^2-2nm-3n+\frac{m^2}{2}+\frac{3m}{2}+1$
IV	$2n^2-2nm+3n+\frac{m^2}{2}-\frac{3m}{2}+1$	$3n^2-3nm+3n+\frac{3m^2}{4}-\frac{3m}{2}+\frac{3}{4}$	$\frac{5m^2}{4}+\frac{7}{4}$
V	$2n^2-2nm+3n+\frac{m^2}{2}-\frac{3m}{2}+1$	$3n^2-3nm+3n+\frac{3m^2}{4}-\frac{3m}{2}+1$	$\frac{5m^2}{4}+2$
VI	$\frac{5m^2}{4}+2$	$3n^2-3nm-3n+\frac{3m^2}{4}+\frac{3m}{2}+1$	$2n^2-2nm-3n+\frac{m^2}{2}+\frac{3m}{2}+1$
VII	0	$3n^2-3nm-3n+3m$	$2n^2-2nm-3n+m^2+3m+2$
$VIII$	$2n^2-2nm+3n+m^2-3m+2$	$3n^2-3nm+3n-3m$	0

- Agrees perfectly with HRR: $\chi\left(X, \mathcal{O}_X(m,n)\right) = m^2+mn-n^2+2$

Possible Generalizations

- Our algorithm applies to line bundles on toric hypersurfaces
- One is also interested more generally in vector bundles
- Can be constructed via the monad bundle construction from line bundles

$$0 \to \bigoplus_{i=1}^{r_A} \mathcal{O}_X(m_i) \hookrightarrow \bigoplus_{i=1}^{r_B} \mathcal{O}_X(n_i) \twoheadrightarrow U \to 0$$

- Again, the bundle is determined by the set of integers (m_i, n_i)
- Would be interesting to figure out if our algorithm works here
- Another possibility is to check higher codimension complete intersections
- Results from the literature indicate similar piecewise polynomial behaviour

Lukas, Constantin 2018

Conclusions

- Crucial to further our understanding of possible swampland constraints. ML?
- Landscape: computing line bundle cohomology of toric hypersurfaces.
- NNs fail in many respects and do not solve interesting question
- Understanding the data $\longrightarrow$ reduction to clustering + simple polynomial fit
- Analytic expressions! Can we understand them in terms of topological data?

String Theorists = Fruit Flies?

Science News *from research organizations*

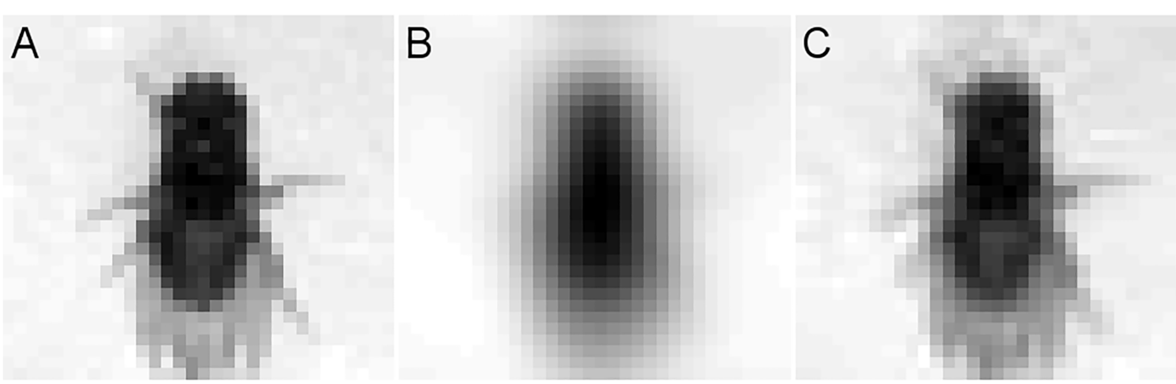
Artificial fly brain can tell who's who

Date: October 25, 2018

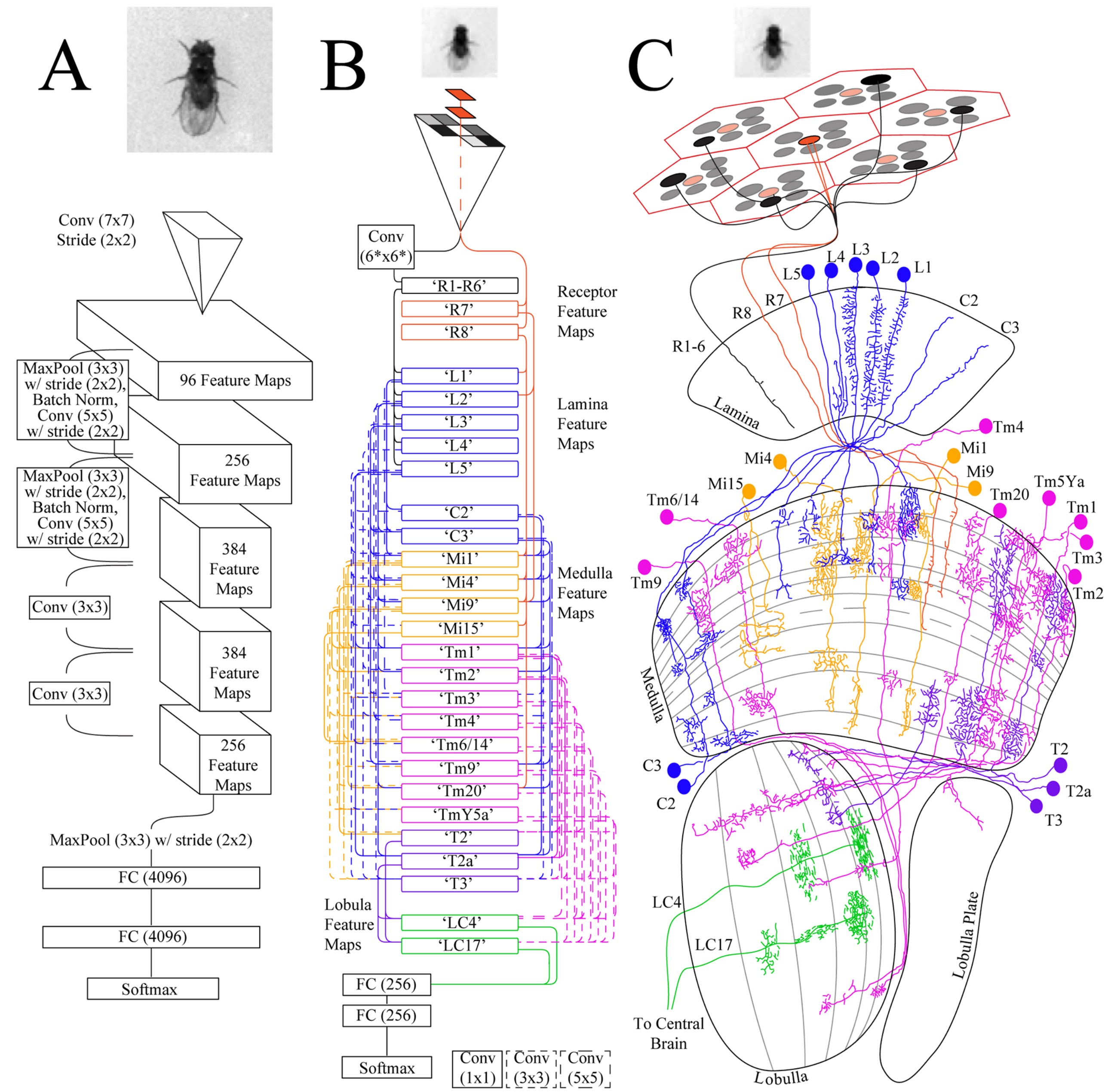
Source: Canadian Institute for Advanced Research

Summary: Researchers have built a neural network that mimics the fruit fly's visual system and can distinguish and re-identify flies. This provides evidence that the humble fruit fly's vision is clearer than previously thought.

Source: https://www.sciencedaily.com/releases/2018/10/181025142010.htm



25,000 artificial neurons



Schneider, Murali, Taylor, Levine 2018

April 1st-3rd 2018

Registration is open!

Thank You