MACHINE LEARNING THE DIMENSION OF A FANO VARIETY

TOM COATES [ID], ALEXANDER M. KASPRZYK [ID], AND SARA VENEZIALE [ID]

ABSTRACT. Fano varieties are basic building blocks in geometry – they are 'atomic pieces' of mathematical shapes. Recent progress in the classification of Fano varieties involves analysing an invariant called the quantum period. This is a sequence of integers which gives a numerical fingerprint for a Fano variety. It is conjectured that a Fano variety is uniquely determined by its quantum period. If this is true, one should be able to recover geometric properties of a Fano variety directly from its quantum period. We apply machine learning to the question: does the quantum period of X know the dimension of X? Note that there is as yet no theoretical understanding of this. We show that a simple feed-forward neural network can determine the dimension of X with 98% accuracy. Building on this, we establish rigorous asymptotics for the quantum periods of a class of Fano varieties. These asymptotics determine the dimension of X from its quantum period. Our results demonstrate that machine learning can pick out structure from complex mathematical data in situations where we lack theoretical understanding. They also give positive evidence for the conjecture that the quantum period of a Fano variety determines that variety.

1. INTRODUCTION

Algebraic geometry describes shapes as the solution sets of systems of polynomial equations, and manipulates or analyses a shape X by manipulating or analysing the equations that define X. This interplay between algebra and geometry has applications across mathematics and science; see e.g. [3, 22, 53, 57]. Shapes defined by polynomial equations are called *algebraic varieties*. Fano varieties are a key class of algebraic varieties. They are, in a precise sense, atomic pieces of mathematical shapes [45, 46]. Fano varieties also play an essential role in string theory. They provide, through their 'anticanonical sections', the main construction of the Calabi–Yau manifolds which give geometric models of spacetime [6, 30, 55].

The classification of Fano varieties is a long-standing open problem. The only one-dimensional example is a line; this is classical. The ten smooth two-dimensional Fano varieties were found by del Pezzo in the 1880s [19]. The classification of smooth Fano varieties in dimension three was a triumph of 20th century mathematics: it combines work by Fano in the 1930s, Iskovskikh in the 1970s, and Mori–Mukai in the 1980s [24, 38–40, 51, 52]. Beyond this, little is known, particularly for the important case of Fano varieties that are not smooth.

A new approach to Fano classification centres around a set of ideas from string theory called Mirror Symmetry [7, 15, 31, 35]. From this perspective, the key invariant of a Fano variety is its *regularized quantum period* [8]

$$\widehat{G}_X(t) = \sum_{d=0}^{\infty} c_d t^d \tag{1}$$

This is a power series with coefficients $c_0 = 1$, $c_1 = 0$, and $c_d = r_d d!$, where r_d is a certain Gromov–Witten invariant of X. Intuitively speaking, r_d is the number of rational curves in X of degree d that pass through a fixed generic point and have a certain constraint on their complex structure. In general r_d can be a rational number, because curves with a symmetry group of order k are counted with weight $1/k$, but in all known cases the coefficients c_d in (1) are integers.

It is expected that the regularized quantum period $\widehat{G}_X$ uniquely determines X. This is true (and proven) for smooth Fano varieties in low dimensions, but is unknown in dimensions four and higher, and for Fano varieties that are not smooth.

(a) $x^2 + y^2 = z^2 + 1$ (b) $x^2 + y^2 = z^2$
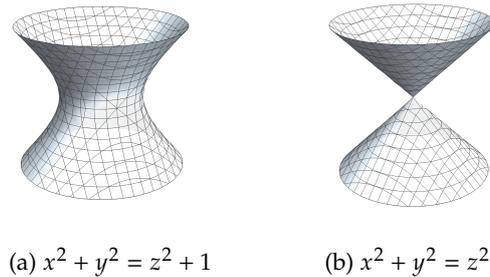
Figure 1. Algebraic varieties and their equations: (a) a smooth example; (b) an example with a singular point.

In this paper we will treat the regularized quantum period as a numerical signature for the Fano variety X, given by the sequence of integers $(c_0, c_1, \ldots)$. *A priori* this looks like an infinite amount of data, but in fact there is a differential operator L such that $L\widehat{G}_X \equiv 0$; see e.g. [8, Theorem 4.3]. This gives a recurrence relation that determines all of the coefficients c_d from the first few terms, so the regularized quantum period $\widehat{G}_X$ contains only a finite amount of information. Encoding a Fano variety X by a vector in $\mathbb{Z}^{m+1}$ given by finitely many coefficients $(c_0, c_1, \ldots, c_m)$ of the regularized quantum period allows us to investigate questions about Fano varieties using machine learning.

In this paper we ask whether the regularized quantum period of a Fano variety X knows the dimension of X. There is currently no viable theoretical approach to this question. Instead we use machine learning methods applied to a large dataset to argue that the answer is probably yes, and then prove that the answer is yes for toric Fano varieties of low Picard rank. The use of machine learning was essential to the formulation of our rigorous results (Theorems 5 and 6 below). This work is therefore proof-of-concept for a larger program, demonstrating that machine learning can uncover previously unknown structure in complex mathematical datasets. Thus the Data Revolution, which has had such impact across the rest of science, also brings important new insights to pure mathematics [18, 21, 34, 49, 58, 59]. This is particularly true for large-scale classification questions, e.g. [1, 10, 14, 17, 47], where these methods can potentially reveal both the classification itself and structural relationships within it.

2. Results

Algebraic varieties can be smooth or have singularities. Depending on their equations, algebraic varieties can be smooth (as in Figure 1(a)) or have singularities (as in Figure 1(b)). In this paper we consider algebraic varieties over the complex numbers. The equations in Figures 1(a) and 1(b) therefore define complex surfaces; however, for ease of visualisation, we have plotted only the points on these surfaces with co-ordinates that are real numbers.

Most of the algebraic varieties that we consider below will be singular, but they all have a class of singularities called *terminal quotient singularities*. This is the most natural class of singularities to allow from the point of view of Fano classification [46]. Terminal quotient singularities are very mild; indeed, in dimensions one and two, an algebraic variety has terminal quotient singularities if and only if it is smooth.

The Fano varieties that we consider. The fundamental example of a Fano variety is projective space $\mathbb{P}^{N-1}$. This is a quotient of $\mathbb{C}^N \setminus \{0\}$ by the group $\mathbb{C}^\times$, where the action of $\lambda \in \mathbb{C}^\times$ identifies the points $(x_1, x_2, \ldots, x_N)$ and $(\lambda x_1, \lambda x_2, \ldots, \lambda x_N)$. The resulting algebraic variety is smooth and has dimension $N - 1$. We will consider generalisations of projective spaces called *weighted projective spaces* and *toric varieties of Picard rank two*. A detailed introduction to these spaces is given in §A.

To define a weighted projective space, choose positive integers $a_1, a_2, \ldots, a_N$ such that any subset of size $N - 1$ has no common factor, and consider

$$\mathbb{P}(a_1, a_2, \ldots, a_N) = (\mathbb{C}^N \setminus \{0\})/\mathbb{C}^\times$$

where the action of $\lambda \in \mathbb{C}^\times$ identifies the points

$$(x_1, x_2, \ldots, x_N) \quad \text{and} \quad (\lambda^{a_1} x_1, \lambda^{a_2} x_2, \ldots, \lambda^{a_N} x_N)$$

in $\mathbb{C}^N \setminus \{0\}$. The quotient $\mathbb{P}(a_1, a_2, \ldots, a_N)$ is an algebraic variety of dimension $N - 1$. A general point of $\mathbb{P}(a_1, a_2, \ldots, a_N)$ is smooth, but there can be singular points. Indeed, a weighted projective space $\mathbb{P}(a_1, a_2, \ldots, a_N)$ is smooth if and only if $a_i = 1$ for all i, that is, if and only if it is a projective space.

To define a toric variety of Picard rank two, choose a matrix

$$\begin{pmatrix} a_1 & a_2 & \cdots & a_N \\ b_1 & b_2 & \cdots & b_N \end{pmatrix} \tag{2}$$

with non-negative integer entries and no zero columns. This defines an action of $\mathbb{C}^\times \times \mathbb{C}^\times$ on $\mathbb{C}^N$, where $(\lambda, \mu) \in \mathbb{C}^\times \times \mathbb{C}^\times$ identifies the points

$$(x_1, x_2, \ldots, x_N) \quad \text{and} \quad (\lambda^{a_1} \mu^{b_1} x_1, \lambda^{a_2} \mu^{b_2} x_2, \ldots, \lambda^{a_N} \mu^{b_N} x_N)$$

in $\mathbb{C}^N$. Set $a = a_1 + a_2 + \cdots + a_N$ and $b = b_1 + b_2 + \cdots + b_N$, and suppose that (a, b) is not a scalar multiple of (a_i, b_i) for any i. This determines linear subspaces

$$S_+ = \{(x_1, x_2, \ldots, x_N) \mid x_i = 0 \text{ if } b_i/a_i < b/a\}$$
$$S_- = \{(x_1, x_2, \ldots, x_N) \mid x_i = 0 \text{ if } b_i/a_i > b/a\}$$

of $\mathbb{C}^N$, and we consider the quotient

$$X = (\mathbb{C}^N \setminus S)/(\mathbb{C}^\times \times \mathbb{C}^\times) \tag{3}$$

where $S = S_+ \cup S_-$. The quotient X is an algebraic variety of dimension $N - 2$ and second Betti number $b_2(X) \leq 2$. If, as we assume henceforth, the subspaces S_+ and S_- both have dimension two or more then $b_2(X) = 2$, and thus X has Picard rank two. In general X will have singular points, the precise form of which is determined by the weights in (2).

There are closed formulas for the regularized quantum period of weighted projective spaces and toric varieties [9]. We have

$$\widehat{G}_{\mathbb{P}}(t) = \sum_{k=0}^{\infty} \frac{(ak)!}{(a_1 k)!(a_2 k)! \cdots (a_N k)!} t^{ak} \tag{4}$$

where $\mathbb{P} = \mathbb{P}(a_1, \ldots, a_N)$ and $a = a_1 + a_2 + \cdots + a_N$, and

$$\widehat{G}_X(t) = \sum_{(k,l) \in \mathbb{Z}^2 \cap C} \frac{(ak + bl)!}{(a_1 k + b_1 l)! \cdots (a_N k + b_N l)!} t^{ak+bl} \tag{5}$$

where the weights for X are as in (2), and C is the cone in $\mathbb{R}^2$ defined by the equations $a_i x + b_i y \geq 0$, $i \in \{1, 2, \ldots, N\}$. Formula (4) implies that, for weighted projective spaces, the coefficient c_d from (1) is zero unless d is divisible by a. Formula (5) implies that, for toric varieties of Picard rank two, $c_d = 0$ unless d is divisible by $\gcd\{a, b\}$.

Data generation: weighted projective spaces. The following result characterises weighted projective spaces with terminal quotient singularities; this is [43, Proposition 2.3].

Proposition 1. *Let $X = \mathbb{P}(a_1, a_2, \ldots, a_N)$ be a weighted projective space of dimension at least three. Then X has terminal quotient singularities if and only if*

$$\sum_{i=1}^{N} \{ka_i/a\} \in \{2, \ldots, N-2\}$$

for each $k \in \{2, \ldots, a - 2\}$. Here $a = a_1 + a_2 + \cdots + a_N$ and $\{q\}$ denotes the fractional part $q - \lfloor q \rfloor$ of $q \in \mathbb{Q}$.

A simpler necessary condition is given by [42, Theorem 3.5]:

Proposition 2. *Let $X = \mathbb{P}(a_1, a_2, \ldots, a_N)$ be a weighted projective space of dimension at least two, with weights ordered $a_1 \leq a_2 \leq \ldots \leq a_N$. If X has terminal quotient singularities then $a_i/a < 1/(N - i + 2)$ for each $i \in \{3, \ldots, N\}$.*

Weighted projective spaces			Rank-two toric varieties		
Dimension	Sample size	Percentage	Dimension	Sample size	Percentage
1	1	0.001			
2	1	0.001	2	2	0.001
3	7	0.005	3	17	0.009
4	8 936	5.957	4	758	0.379
5	23 584	15.723	5	6 050	3.025
6	23 640	15.760	6	19 690	9.845
7	23 700	15.800	7	35 395	17.698
8	23 469	15.646	8	42 866	21.433
9	23 225	15.483	9	47 206	23.603
10	23 437	15.625	10	48 016	24.008
Total	150 000		Total	200 000	

Table 1. The distribution by dimension in our datasets.

Weighted projective spaces with terminal quotient singularities have been classified in dimensions up to four [41, 43]. Classifications in higher dimensions are hindered by the lack of an effective upper bound on a.

We randomly generated 150 000 distinct weighted projective spaces with terminal quotient singularities, and with dimension up to 10, as follows. We generated random sequences of weights $a_1 \leq a_2 \leq \ldots \leq a_N$ with $a_N \leq 10N$ and discarded them if they failed to satisfy any one of the following:

   (i) for each $i \in \{1, \ldots, N\}$, $\gcd\{a_1, \ldots, \widehat{a_i}, \ldots, a_N\} = 1$, where $\widehat{a_i}$ indicates that a_i is omitted;
   (ii) $a_i/a < 1/(N - i + 2)$ for each $i \in \{3, \ldots, N\}$;
   (iii) $\sum_{i=1}^{N}\{ka_i/a\} \in \{2, \ldots, N - 2\}$ for each $k \in \{2, \ldots, a - 2\}$.

Condition (i) here was part of our definition of weighted projective spaces above; it ensures that the set of singular points in $\mathbb{P}(a_1, a_2, \ldots, a_N)$ has dimension at most $N - 2$, and also that weighted projective spaces are isomorphic as algebraic varieties if and only if they have the same weights. Condition (ii) is from Proposition 2; it efficiently rules out many non-terminal examples. Condition (iii) is the necessary and sufficient condition from Proposition 1. We then deduplicated the sequences. The resulting sample sizes are summarised in Table 1.

Data generation: toric varieties. Deduplicating randomly-generated toric varieties of Picard rank two is harder than deduplicating randomly generated weighted projective spaces, because different weight matrices in (2) can give rise to the same toric variety. Toric varieties are uniquely determined, up to isomorphism, by a combinatorial object called a *fan* [25]. A fan is a collection of cones, and one can determine the singularities of a toric variety X from the geometry of the cones in the corresponding fan.

We randomly generated 200 000 distinct toric varieties of Picard rank two with terminal quotient singularities, and with dimension up to 10, as follows. We randomly generated weight matrices, as in (2), such that $0 \leq a_i, b_j \leq 5$. We then discarded the weight matrix if any column was zero, and otherwise formed the corresponding fan F. We discarded the weight matrix unless:

   (i) F had N rays;
   (ii) each cone in F was simplicial (i.e. has number of rays equal to its dimension);
   (iii) the convex hull of the primitive generators of the rays of F contained no lattice points other than the rays and the origin.

Conditions (i) and (ii) together guarantee that X has Picard rank two, and are equivalent to the conditions on the weight matrix in (2) given in our definition. Conditions (ii) and (iii) guarantee that X has terminal quotient singularities. We then deduplicated the weight matrices according to the isomorphism type of F, by putting F in normal form [32, 48]. See Table 1 for a summary of the dataset.

(a) $\log c_d$ for $\mathbb{P}(5,5,11,23,28,29,33,44,66,76)$.

(b) $\log c_d$ for Example 3.
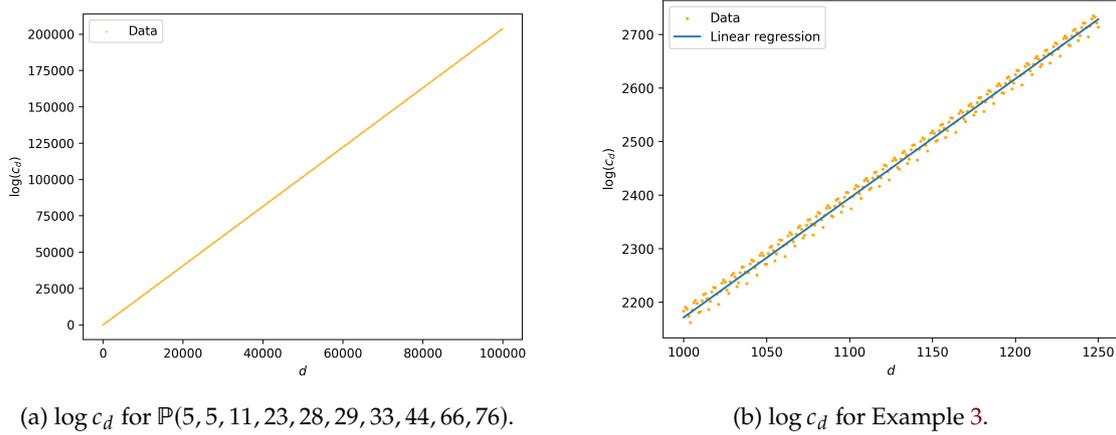
Figure 2. The logarithm of the non-zero period coefficients c_d: (a) for a typical weighted projective space; (b) for the toric variety of Picard rank two from Example 3.


(a)

(b)
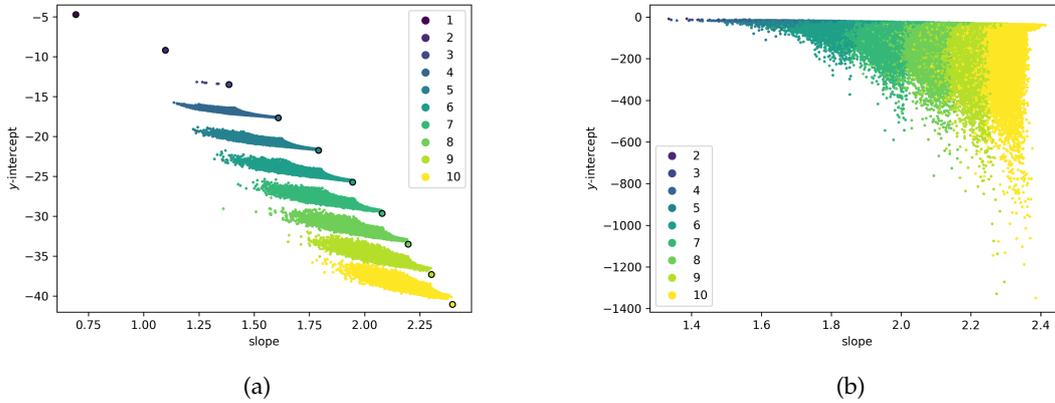
Figure 3. The slopes and y-intercepts from the linear models: (a) for weighted projective spaces with terminal quotient singularities. The colour records the dimension of the weighted projective space and the circled points indicate projective spaces. (b) for toric varieties of Picard rank two with terminal quotient singularities. The colour records the dimension of the toric variety.

Data analysis: weighted projective spaces. We computed an initial segment $(c_0, c_1, \ldots, c_m)$ of the regularized quantum period for all the examples in the sample of $150\,000$ terminal weighted projective spaces, with $m \approx 100\,000$. The non-zero coefficients c_d appeared to grow exponentially with d, and so we considered $\{\log c_d\}_{d \in S}$ where $S = \{d \in \mathbb{Z}_{\geq 0} \mid c_d \neq 0\}$. To reduce dimension we fitted a linear model to the set $\{(d, \log c_d) \mid d \in S\}$ and used the slope and intercept of this model as features; see Figure 2(a) for a typical example. Plotting the slope against the y-intercept and colouring datapoints according to the dimension we obtain Figure 3(a): note the clear separation by dimension. A Support Vector Machine (SVM) trained on 10% of the slope and y-intercept data predicted the dimension of the weighted projective space with an accuracy of 99.99%. Full details are given in §§B–C.

Data analysis: toric varieties. As before, the non-zero coefficients c_d appeared to grow exponentially with d, so we fitted a linear model to the set $\{(d, \log c_d) \mid d \in S\}$ where $S = \{d \in \mathbb{Z}_{\geq 0} \mid c_d \neq 0\}$. We used the slope and intercept of this linear model as features.

Example 3. In Figure 2(b) we plot a typical example: the logarithm of the regularized quantum period sequence for the nine-dimensional toric variety with weight matrix

$$\begin{pmatrix} 1 & 2 & 5 & 3 & 3 & 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 4 & 4 & 1 & 2 & 2 & 3 & 4 \end{pmatrix}$$
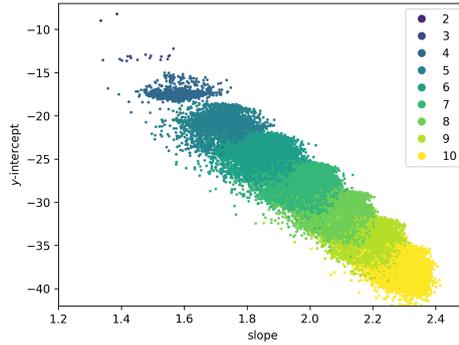
Figure 4. The slopes and y-intercepts from the linear model. This is as in Figure 3(b), but plotting only data points for which the standard error s_{int} for the y-intercept satisfies $s_{\text{int}} < 0.3$. The colour records the dimension of the toric variety.

along with the linear approximation. We see a periodic deviation from the linear approximation; the magnitude of this deviation decreases as d increases (not shown).

To reduce computational costs, we computed pairs $(d, \log c_d)$ for $1000 \leq d \leq 20\,000$ by sampling every 100th term. We discarded the beginning of the period sequence because of the noise it introduces to the linear regression. In cases where the sampled coefficient c_d is zero, we considered instead the next non-zero coefficient. The resulting plot of slope against y-intercept, with datapoints coloured according to dimension, is shown in Figure 3(b).

We analysed the standard errors for the slope and y-intercept of the linear model. The standard errors for the slope are small compared to the range of slopes, but in many cases the standard error s_{int} for the y-intercept is relatively large. As Figure 4 illustrates, discarding data points where the standard error s_{int} for the y-intercept exceeds some threshold reduces apparent noise. This suggests that the underlying structure is being obscured by inaccuracies in the linear regression caused by oscillatory behaviour in the initial terms of the quantum period sequence; these inaccuracies are concentrated in the y-intercept of the linear model. Note that restricting attention to those data points where s_{int} is small also greatly decreases the range of y-intercepts that occur. As Example 4 and Figure 5 suggest, this reflects both transient oscillatory behaviour and also the presence of a subleading term in the asymptotics of $\log c_d$ which is missing from our feature set. We discuss this further below.

Example 4. Consider the toric variety with Picard rank two and weight matrix

$$\begin{pmatrix} 1 & 10 & 5 & 13 & 8 & 12 & 0 \\ 0 & 0 & 3 & 8 & 5 & 14 & 1 \end{pmatrix}$$

This is one of the outliers in Figure 3(b). The toric variety is five-dimensional, and has slope 1.637 and y-intercept -62.64. The standard errors are 4.246×10^{-4} for the slope and 5.021 for the y-intercept. We computed the first 40 000 coefficients c_d in (1). As Figure 5 shows, as d increases the y-intercept of the linear model increases to -28.96 and s_{int} decreases to 0.7877. At the same time, the slope of the linear model remains more or less unchanged, decreasing to 1.635. This supports the idea that computing (many) more coefficients c_d would significantly reduce noise in Figure 3(b). In this example, even 40 000 coefficients may not be enough.

Computing many more coefficients c_d across the whole dataset would require impractical amounts of computation time. In the example above, which is typical in this regard, increasing the number of coefficients computed from 20 000 to 40 000 increased the computation time by a factor of more than 10. Instead we restrict to those toric varieties of Picard rank two such that the y-intercept standard error s_{int} is less than 0.3; this retains 67 443 of the 200 000 datapoints. We used 70% of the slope and y-intercept data in the restricted dataset for model training, and the rest for validation. An SVM model predicted the dimension of the toric variety with an accuracy of 87.7%, and a Random Forest Classifier (RFC) predicted the dimension with an accuracy of 88.6%.
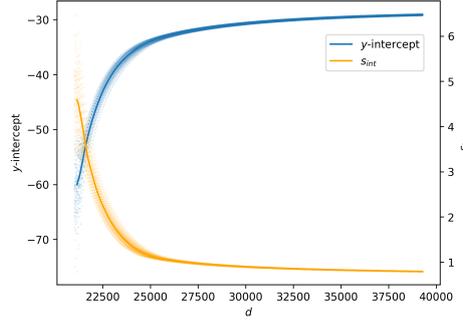
Figure 5. Variation as we move deeper into the period sequence. The y-intercept and its standard error s_{int} for the toric variety from Example 4, as computed from pairs $(k, \log c_k)$ such that $d - 20\,000 \leq k \leq d$ by sampling every 100th term. We also show LOWESS-smoothed trend lines.

Neural networks. Neural networks do not handle unbalanced datasets well. We therefore removed the toric varieties of dimensions 3, 4, and 5 from our data, leaving 61 164 toric varieties of Picard rank two with terminal quotient singularities and $s_{\text{int}} < 0.3$. This dataset is approximately balanced by dimension.

A Multilayer Perceptron (MLP) with three hidden layers of sizes $(10, 30, 10)$ using the slope and intercept as features predicted the dimension with 89.0% accuracy. Since the slope and intercept give good control over $\log c_d$ for $d \gg 0$, but not for small d, it is likely that the coefficients c_d with d small contain extra information that the slope and intercept do not see. Supplementing the feature set by including the first 100 coefficients c_d as well as the slope and intercept increased the accuracy of the prediction to 97.7%. Full details can be found in §§B–C.

From machine learning to rigorous analysis. Elementary "out of the box" models (SVM, RFC, and MLP) trained on the slope and intercept data alone already gave a highly accurate prediction for the dimension. Furthermore even for the many-feature MLP, which was the most accurate, sensitivity analysis using SHAP values [50] showed that the slope and intercept were substantially more important to the prediction than any of the coefficients c_d: see Figure 6. This suggested that the dimension of X might be visible from a rigorous estimate of the growth rate of $\log c_d$.

In §3 we establish asymptotic results for the regularized quantum period of toric varieties with low Picard rank, as follows. These results apply to any weighted projective space or toric variety of Picard rank two: they do not require a terminality hypothesis. Note, in each case, the presence of a subleading logarithmic term in the asymptotics for $\log c_d$.

Theorem 5. *Let X denote the weighted projective space $\mathbb{P}(a_1, \ldots, a_N)$, so that the dimension of X is $N - 1$. Let c_d denote the coefficient of t^d in the regularized quantum period $\widehat{G}_X(t)$ given in (4). Let $a = a_1 + \cdots + a_N$ and $p_i = a_i / a$. Then $c_d = 0$ unless d is divisible by a, and non-zero coefficients c_d satisfy*

$$\log c_d \sim Ad - \frac{\dim X}{2} \log d + B$$

as $d \to \infty$, where

$$A = -\sum_{i=1}^{N} p_i \log p_i$$

$$B = -\frac{\dim X}{2} \log(2\pi) - \frac{1}{2} \sum_{i=1}^{N} \log p_i$$

Note, although it plays no role in what follows, that A is the Shannon entropy of the discrete random variable Z with distribution $(p_1, p_2, \ldots, p_N)$, and that B is a constant plus half the total self-information of Z.
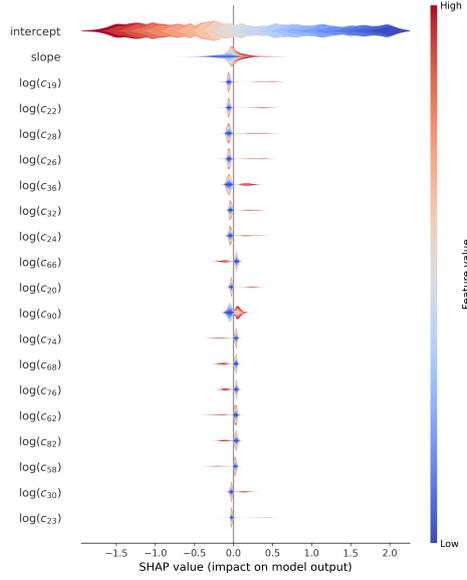
Figure 6. Model sensitivity analysis using SHAP values. The model is an MLP with three hidden layers of sizes (10,30,10) applied to toric varieties of Picard rank two with terminal quotient singularities. It is trained on the slope, y-intercept, and the first 100 coefficients c_d as features, and predicts the dimension with 97.7% accuracy.

Theorem 6. *Let X denote the toric variety of Picard rank two with weight matrix*

$$\begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_N \\ b_1 & b_2 & b_3 & \cdots & b_N \end{pmatrix}$$

so that the dimension of X is $N-2$. Let $a = a_1 + \cdots + a_N$, $b = b_1 + \cdots + b_N$, and $\ell = \gcd\{a, b\}$. Let $[\mu : \nu] \in \mathbb{P}^1$ be the unique root of the homogeneous polynomial

$$\prod_{i=1}^{N}(a_i\mu + b_i\nu)^{a_i b} - \prod_{i=1}^{N}(a_i\mu + b_i\nu)^{b_i a}$$

such that $a_i\mu + b_i\nu \geq 0$ for all $i \in \{1, 2, \ldots, N\}$, and set

$$p_i = \frac{\mu a_i + \nu b_i}{\mu a + \nu b}$$

Let c_d denote the coefficient of t^d in the regularized quantum period $\widehat{G}_X(t)$ given in (5). Then non-zero coefficients c_d satisfy

$$\log c_d \sim Ad - \frac{\dim X}{2}\log d + B$$

as $d \to \infty$, where

$$A = -\sum_{i=1}^{N} p_i \log p_i$$

$$B = -\frac{\dim X}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{N}\log p_i - \frac{1}{2}\log\left(\sum_{i=1}^{N}\frac{(a_i b - b_i a)^2}{\ell^2 p_i}\right)$$

Theorem 5 is a straightforward application of Stirling's formula. Theorem 6 is more involved, and relies on a Central Limit-type theorem that generalises the De Moivre–Laplace theorem.

Theoretical analysis. The asymptotics in Theorems 5 and 6 imply that, for X a weighted projective space or toric variety of Picard rank two, the quantum period determines the dimension of X. Let us revisit the clustering analysis from this perspective. Recall the asymptotic expression $\log c_d \sim Ad - \frac{\dim X}{2}\log d + B$ and the formulae for A and B from Theorem 5. Figure 7(a) shows the values of A

(a)   (b)
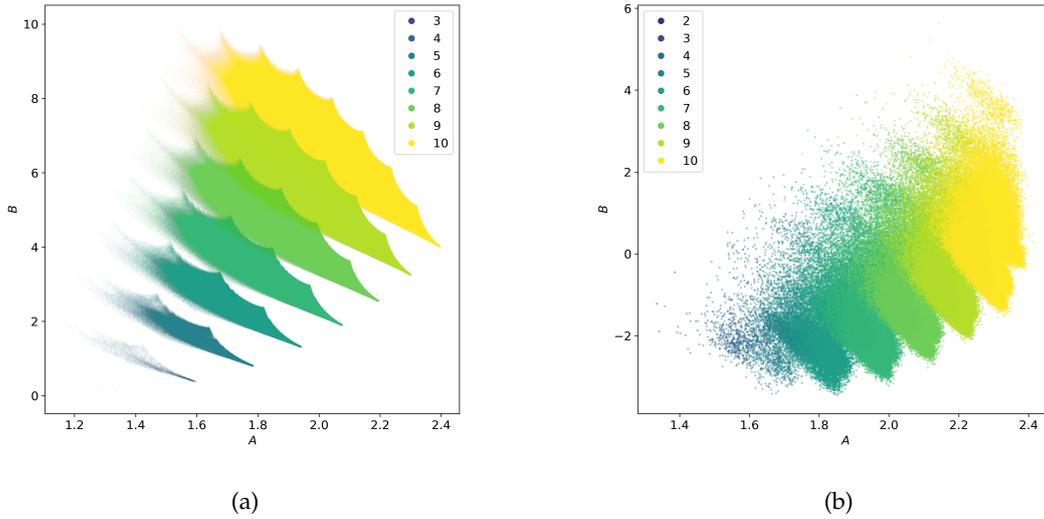
Figure 7. The values of the asymptotic coefficients A and B: (a) for all weighted projective spaces $\mathbb{P}(a_1, \ldots, a_N)$ with terminal quotient singularities and $a_i \leq 25$ for all i. The colour records the dimension of the weighted projective space. (b) for toric varieties of Picard rank two in our dataset. The colour records the dimension of the toric variety.

(a)   (b)
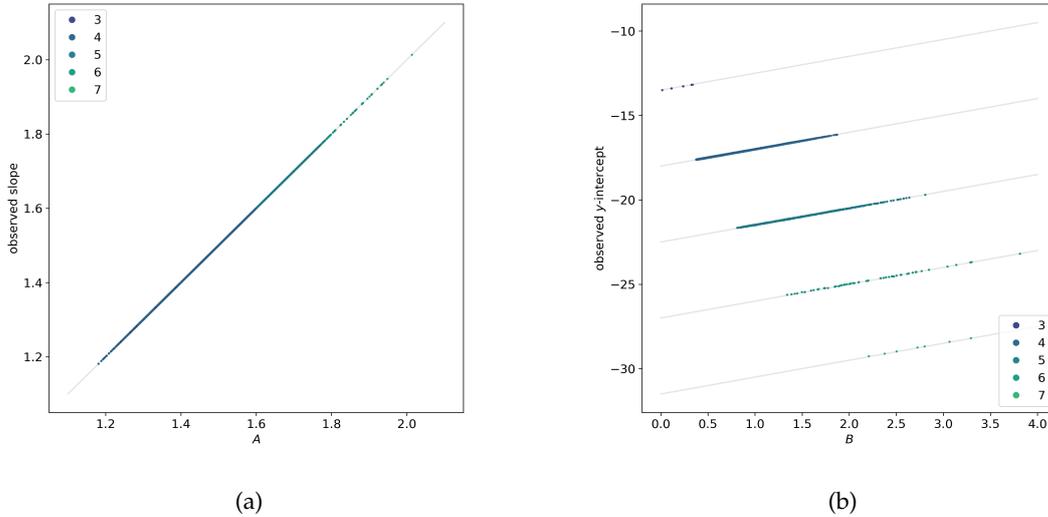
Figure 8. For weighted projective spaces, the asymptotic coefficients A and B are closely related to the slope and y-intercept. (a) Comparison between A and the slope from the linear model, for weighted projective spaces that occur in both Figure 3(a) and Figure 7(a), coloured by dimension. The line slope = A is indicated. (b) Comparison between B and the y-intercept from the linear model, for weighted projective spaces that occur in both Figure 3(a) and Figure 7(a), coloured by dimension. In each case the line y-intercept = $B - \frac{9}{2} \dim X$ is shown.

and B for a sample of weighted projective spaces, coloured by dimension. Note the clusters, which overlap. Broadly speaking, the values of B increase as the dimension of the weighted projective space increases, whereas in Figure 3(a) the y-intercepts decrease as the dimension increases. This reflects the fact that we fitted a linear model to $\log c_d$, omitting the subleading $\log d$ term in the asymptotics. As Figure 8 shows, the linear model assigns the omitted term to the y-intercept rather than the slope.
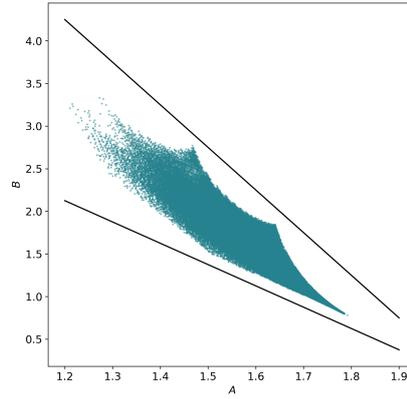
Figure 9. Linear bounds for the cluster of five-dimensional weighted projective spaces in Figure 7(a). The bounds are given by Proposition 7.

The slope of the linear model is approximately equal to A. The y-intercept, however, differs from B by a dimension-dependent factor. The omitted log term does not vary too much over the range of degrees ($d < 100\,000$) that we considered, and has the effect of reducing the observed y-intercept from B to approximately $B - \frac{9}{2}\dim X$, distorting the clusters slightly and translating them downwards by a dimension-dependent factor. This separates the clusters. We expect that the same mechanism applies in Picard rank two as well: see Figure 7(b).

We can show that each cluster in Figure 7(a) is linearly bounded using constrained optimisation techniques. Consider for example the cluster for weighted projective spaces of dimension five, as in Figure 9.

Proposition 7. *Let X be the five-dimensional weighted projective space $\mathbb{P}(a_1, \ldots, a_6)$, and let A, B be as in Theorem 5. Then $B + \frac{5}{2}A \geq \frac{41}{8}$. If in addition $a_i \leq 25$ for all i then $B + 5A \leq \frac{41}{40}$.*

Fix a suitable $\theta \geq 0$ and consider

$$B + \theta A = -\frac{\dim X}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{N}\log p_i - \theta \sum_{i=1}^{N} p_i \log p_i$$

with $\dim X = N - 1 = 5$. Solving

$$\min(B + \theta A) \quad \text{subject to} \quad p_1 + \cdots + p_6 = 1$$
$$p_1, \ldots, p_6 \geq 0$$

on the five-simplex gives a linear lower bound for the cluster. This bound does not use terminality: it applies to any weighted projective space of dimension five. The expression $B + \theta A$ is unbounded above on the five-simplex (because B is) so we cannot obtain an upper bound this way. Instead, consider

$$\max(B + \theta A) \quad \text{subject to} \quad p_1 + \cdots + p_6 = 1$$
$$\epsilon \leq p_1 \leq p_2 \leq \cdots \leq p_6$$

for an appropriate small positive ϵ, which we can take to be $1/a$ where a is the maximum sum of the weights. For Figure 9, for example, we can take $a = 124$, and in general such an a exists because there are only finitely many terminal weighted projective spaces. This gives a linear upper bound for the cluster.

The same methods yield linear bounds on each of the clusters in Figure 7(a). As the Figure shows however, the clusters are not linearly separable.

Discussion. We developed machine learning models that predict, with high accuracy, the dimension of a Fano variety from its regularized quantum period. These models apply to weighted projective spaces and toric varieties of Picard rank two with terminal quotient singularities. We then established rigorous asymptotics for the regularized quantum period of these Fano varieties. The form of the asymptotics implies that, in these cases, the regularized quantum period of a Fano variety X determines the dimension of X. The asymptotics also give a theoretical underpinning for the success of the machine learning models.

Perversely, because the series involved converge extremely slowly, reading the dimension of a Fano variety directly from the asymptotics of the regularized quantum period is not practical. For the same reason, enhancing the feature set of our machine learning models by including a $\log d$ term in the linear regression results in less accurate predictions. So although the asymptotics in Theorems 5 and 6 determine the dimension in theory, in practice the most effective way to determine the dimension of an unknown Fano variety from its quantum period is to apply a machine learning model.

The insights gained from machine learning were the key to our formulation of the rigorous results in Theorems 5 and 6. Indeed, it might be hard to discover these results without a machine learning approach. It is notable that the techniques in the proof of Theorem 6 – the identification of generating functions for Gromov–Witten invariants of toric varieties with certain hypergeometric functions – have been known since the late 1990s and have been studied by many experts in hypergeometric functions since then. For us, the essential step in the discovery of the results was the feature extraction that we performed as part of our ML pipeline.

This work demonstrates that machine learning can uncover previously unknown structure in complex mathematical data, and is a powerful tool for developing rigorous mathematical results; cf. [18]. It also provides evidence for a fundamental conjecture in the Fano classification program [8]: that the regularized quantum period of a Fano variety determines that variety.

3. METHODS

In this section we prove Theorem 5 and Theorem 6. The following result implies Theorem 5.

Theorem 8. *Let X denote the weighted projective space $\mathbb{P}(a_1, \ldots, a_N)$, so that the dimension of X is $N - 1$. Let c_d denote the coefficient of t^d in the regularized quantum period $\widehat{G}_X(t)$ given in (4). Let $a = a_1 + \ldots + a_N$. Then $c_d = 0$ unless d is divisible by a, and*

$$\log c_{ka} \sim ka \left[\log a - \frac{1}{a} \sum_{i=1}^{N} a_i \log a_i \right] - \frac{\dim X}{2} \log(ka) + \frac{1 + \dim X}{2} \log a - \frac{\dim X}{2} \log(2\pi) - \frac{1}{2} \sum_{i=1}^{N} \log a_i$$

That is, non-zero coefficients c_d satisfy

$$\log c_d \sim Ad - \frac{\dim X}{2} \log d + B$$

as $d \to \infty$, where

$$A = - \sum_{i=1}^{N} p_i \log p_i \quad B = -\frac{\dim X}{2} \log(2\pi) - \frac{1}{2} \sum_{i=1}^{N} \log p_i$$

and $p_i = a_i / a$.

Proof. Combine Stirling's formula

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e} \right)^n$$

with the closed formula (4) for c_{ka}. $\qquad \square$

Toric varieties of Picard rank 2. Consider a toric variety X of Picard rank two and dimension $N - 2$ with weight matrix

$$\begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_N \\ b_1 & b_2 & b_3 & \cdots & b_N \end{pmatrix}$$

as in (2). Let us move to more invariant notation, writing α_i for the linear form on $\mathbb{R}^2$ defined by the transpose of the ith column of the weight matrix, and $\alpha = \alpha_1 + \cdots + \alpha_N$. Equation 5 becomes

$$\widehat{G}_X(t) = \sum_{k \in \mathbb{Z}^2 \cap C} \frac{(\alpha \cdot k)!}{\prod_{i=1}^{N}(\alpha_i \cdot k)!} t^{\alpha \cdot k}$$

where C is the cone $C = \{x \in \mathbb{R}^2 \mid \alpha_i \cdot x \geq 0 \text{ for } i = 1, 2, \ldots, N\}$. As we will see, for $d \gg 0$ the coefficients

$$\frac{(\alpha \cdot k)!}{\prod_{i=1}^{N}(\alpha_i \cdot k)!} \quad \text{where } k \in \mathbb{Z}^2 \cap C \text{ and } \alpha \cdot k = d$$

are approximated by a rescaled Gaussian. We begin by finding the mean of that Gaussian, that is, by minimising

$$\prod_{i=1}^{N}(\alpha_i \cdot k)! \quad \text{where } k \in \mathbb{Z}^2 \cap C \text{ and } \alpha \cdot k = d.$$

For k in the strict interior of C with $\alpha \cdot k = d$, we have that

$$(\alpha_i \cdot k)! \sim \left(\frac{\alpha_i \cdot k}{e}\right)^{\alpha_i \cdot k}$$

as $d \to \infty$.

Proposition 9. *The constrained optimisation problem*

$$\min \prod_{i=1}^{N}(\alpha_i \cdot x)^{\alpha_i \cdot x} \quad \text{subject to} \quad \begin{cases} x \in C \\ \alpha \cdot x = d \end{cases}$$

has a unique solution $x = x^*$. Furthermore, setting $p_i = (\alpha_i \cdot x^*)/(\alpha \cdot x^*)$ we have that the monomial

$$\prod_{i=1}^{N} p_i^{\alpha_i \cdot k}$$

depends on $k \in \mathbb{Z}^2$ only via $\alpha \cdot k$.

Proof. Taking logarithms gives the equivalent problem

$$\min \sum_{i=1}^{N}(\alpha_i \cdot x) \log(\alpha_i \cdot x) \qquad \text{subject to} \quad \begin{cases} x \in C \\ \alpha \cdot x = d \end{cases} \tag{6}$$

The objective function $\sum_{i=1}^{N}(\alpha_i \cdot x) \log(\alpha_i \cdot x)$ here is the pullback to $\mathbb{R}^2$ of the function

$$f(x_1, \ldots, x_N) = \sum_{i=1}^{N} x_i \log x_i$$

along the linear embedding $\varphi : \mathbb{R}^2 \to \mathbb{R}^N$ given by $(\alpha_1, \ldots, \alpha_N)$. Note that C is the preimage under φ of the positive orthant $\mathbb{R}_+^N$, so we need to minimise f on the intersection of the simplex $x_1 + \cdots + x_N = d$, $(x_1, \ldots, x_N) \in \mathbb{R}_+^N$ with the image of φ. The function f is convex and decreases as we move away from the boundary of the simplex, so the minimisation problem in (6) has a unique solution x^* and this lies in the strict interior of C. We can therefore find the minimum x^* using the method of Lagrange multipliers, by solving

$$\sum_{i=1}^{N} \alpha_i \log(\alpha_i \cdot x) + \alpha = \lambda \alpha \tag{7}$$

for $\lambda \in \mathbb{R}$ and x in the interior of C with $\alpha \cdot x = d$. Thus

$$\sum_{i=1}^{N} \alpha_i \log(\alpha_i \cdot x^*) = (\lambda - 1)\alpha$$

and, evaluating on $k \in \mathbb{Z}^2$ and exponentiating, we see that

$$\prod_{i=1}^{N} (\alpha_i \cdot x^*)^{\alpha_i \cdot k}$$

depends only on $\alpha \cdot k$. The result follows. $\qquad\square$

Given a solution x^* to (7), any positive scalar multiple of x^* also satisfies (7), with a different value of λ and a different value of d. Thus the solutions x^*, as d varies, lie on a half-line through the origin. The direction vector $[\mu : \nu] \in \mathbb{P}^1$ of this half-line is the unique solution to the system

$$\prod_{i=1}^{N} (a_i \mu + b_i \nu)^{a_i b} = \prod_{i=1}^{N} (a_i \mu + b_i \nu)^{b_i a} \tag{8}$$

$$\begin{pmatrix} \mu \\ \nu \end{pmatrix} \in C$$

Note that the first equation here is homogeneous in μ and ν; it is equivalent to (7), by exponentiating and then eliminating λ. Any two solutions x^*, for different values of d, differ by rescaling, and the quantities p_i in Proposition 9 are invariant under this rescaling. They also satisfy $p_1 + \cdots + p_N = 1$.

We use the following result, known in the literature as the "Local Theorem" [29], to approximate multinomial coefficients.

Local Theorem. *For $p_1, \ldots, p_n \in [0,1]$ such that $p_1 + \cdots + p_n = 1$, the ratio*

$$d^{\frac{n-1}{2}} \binom{d}{k_1 \cdots k_n} \prod_{i=1}^{n} p_i^{k_i} : \frac{\exp(-\frac{1}{2} \sum_{i=1}^{n} q_i x_i^2)}{(2\pi)^{\frac{n-1}{2}} \sqrt{p_1 \cdots p_n}} \to 1$$

as $d \to \infty$, uniformly in all k_i's, where

$$q_i = 1 - p_i \qquad\qquad x_i = \frac{k_i - d p_i}{\sqrt{d p_i q_i}}$$

and the x_i lie in bounded intervals.

Let B_r denote the ball of radius r about $x^* \in \mathbb{R}^2$. Fix $R > 0$. We apply the Local Theorem with $k_i = \alpha_i \cdot k$ and $p_i = (\alpha_i \cdot x^*)/(\alpha \cdot x^*)$, where $k \in \mathbb{Z}^2 \cap C$ satisfies $\alpha \cdot k = d$ and $k \in B_{R\sqrt{d}}$. Since

$$x_i = \frac{\alpha_i \cdot (k - x^*)}{\sqrt{d p_i q_i}}$$

the assumption that $k \in B_{R\sqrt{d}}$ ensures that the x_i remain bounded as $d \to \infty$. Note that, by Proposition 9, the monomial $\prod_{i=1}^{N} p_i^{k_i}$ depends on k only via $\alpha \cdot k$, and hence here is independent of k:

$$\prod_{i=1}^{N} p_i^{k_i} = \prod_{i=1}^{N} p_i^{\alpha_i \cdot x^*} = \prod_{i=1}^{N} p_i^{d p_i}$$

Furthermore

$$\sum_{i=1}^{N} q_i x_i^2 = \frac{(k - x^*)^T A (k - x^*)}{d}$$

where A is the positive-definite 2×2 matrix given by

$$A = \sum_{i=1}^{N} \frac{1}{p_i} \alpha_i^T \alpha_i$$

Thus as $d \to \infty$, the ratio

$$\frac{(\alpha \cdot k)!}{\prod_{i=1}^{N} (\alpha_i \cdot k)!} : \frac{\exp\left(-\frac{1}{2d}(k - x^*)^T A (k - x^*)\right)}{(2\pi d)^{\frac{N-1}{2}} \prod_{i=1}^{N} p_i^{d p_i + \frac{1}{2}}} \to 1 \tag{9}$$

for all $k \in \mathbb{Z}^2 \cap C \cap B_{R\sqrt{d}}$ such that $\alpha \cdot k = d$.

Theorem 6. *Let X be a toric variety of Picard rank two and dimension $N - 2$ with weight matrix*

$$\begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_N \\ b_1 & b_2 & b_3 & \cdots & b_N \end{pmatrix}$$

Let $a = a_1 + \cdots + a_N$ and $b = b_1 + \cdots + b_N$, let $\ell = \gcd(a, b)$, and let $[\mu : \nu] \in \mathbb{P}^1$ be the unique solution to (8). Let c_d denote the coefficient of t^d in the regularized quantum period $\widehat{G}_X(t)$. Then non-zero coefficients c_d satisfy

$$\log c_d \sim Ad - \frac{\dim X}{2} \log d + B$$

as $d \to \infty$, where

$$A = - \sum_{i=1}^{N} p_i \log p_i$$

$$B = - \frac{\dim X}{2} \log(2\pi) - \frac{1}{2} \sum_{i=1}^{N} \log p_i - \frac{1}{2} \log \left(\sum_{i=1}^{N} \frac{(a_i b - b_i a)^2}{\ell^2 p_i} \right)$$

and $p_i = \dfrac{\mu a_i + \nu b_i}{\mu a + \nu b}$.

Proof. We need to estimate

$$c_d = \sum_{\substack{k \in \mathbb{Z}^2 \cap C \\ \text{with } \alpha \cdot k = d}} \frac{(\alpha \cdot k)!}{\prod_{i=1}^{N} (\alpha_i \cdot k)!}$$

Consider first the summands with $k \in \mathbb{Z}^2 \cap C$ such that $\alpha \cdot k = d$ and $k \notin B_{R\sqrt{d}}$. For d sufficiently large, each such summand is bounded by $cd^{-\frac{1+\dim X}{2}}$ for some constant c – see (9). Since the number of such summands grows linearly with d, in the limit $d \to \infty$ the contribution to c_d from $k \notin B_{R\sqrt{d}}$ vanishes.

As $d \to \infty$, therefore

$$c_d \sim \frac{1}{(2\pi d)^{\frac{N-1}{2}} \prod_{i=1}^{N} p_i^{dp_i + \frac{1}{2}}} \sum_{\substack{k \in \mathbb{Z}^2 \cap C \cap B_{R\sqrt{d}} \\ \text{with } \alpha \cdot k = d}} \exp \left(- \frac{(k - x^*)^T A (k - x^*)}{2d} \right)$$

Writing $y_k = (k - x^*)/\sqrt{d}$, considering the sum here as a Riemann sum, and letting $R \to \infty$, we see that

$$c_d \sim \frac{1}{(2\pi d)^{\frac{N-1}{2}} \prod_{i=1}^{N} p_i^{dp_i + \frac{1}{2}}} \sqrt{d} \int_{L_\alpha} \exp \left(-\tfrac{1}{2} y^T A y \right) dy$$

where L_α is the line through the origin given by $\ker \alpha$ and dy is the measure on L_α given by the integer lattice $\mathbb{Z}^2 \cap L_\alpha \subset L_\alpha$.

To evaluate the integral, let

$$\alpha^\perp = \frac{1}{\ell} \begin{pmatrix} b \\ -a \end{pmatrix} \quad \text{where } \ell = \gcd\{a, b\}$$

and observe that the pullback of dy along the map $\mathbb{R} \to L_\alpha$ given by $t \mapsto t\alpha^\perp$ is the standard measure on $\mathbb{R}$. Thus

$$\int_{L_\alpha} \exp \left(-\tfrac{1}{2} y^T A y \right) dy = \int_{-\infty}^{\infty} \exp \left(-\tfrac{1}{2} \theta x^2 \right) dx = \sqrt{\frac{2\pi}{\theta}}$$

where $\theta = \sum_{i=1}^{N} \frac{1}{\ell^2 p_i} (\alpha_i \cdot \alpha^\perp)^2$, and

$$c_d \sim \frac{1}{(2\pi d)^{\frac{\dim X}{2}} \prod_{i=1}^{N} p_i^{dp_i + \frac{1}{2}} \sqrt{\theta}}$$

Taking logarithms gives the result. □

(a) (b)
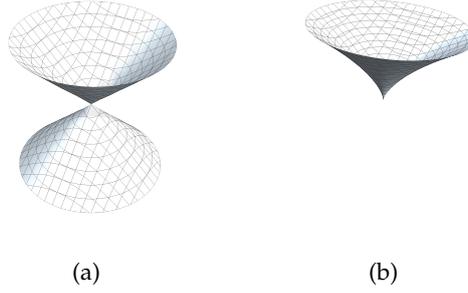
Figure 10. Singular charts on the weighted projective space $\mathbb{P}(1,2,3)$: (a) the real-valued points in the chart U_y. (b) the real-valued points in the chart U_z.

APPENDIX A. SUPPLEMENTARY NOTES

We begin with an introduction to weighted projective spaces and toric varieties, aimed at non-specialists.

Projective spaces and weighted projective spaces. The fundamental example of a Fano variety is two-dimensional projective space $\mathbb{P}^2$. This is a quotient of $\mathbb{C}^3 \setminus \{0\}$ by the group $\mathbb{C}^\times$, where the action of $\lambda \in \mathbb{C}^\times$ identifies the points (x, y, z) and $(\lambda x, \lambda y, \lambda z)$ in $\mathbb{C}^3 \setminus \{0\}$. The variety $\mathbb{P}^2$ is smooth: we can see this by covering it with three open sets U_x, U_y, U_z that are each isomorphic to the plane $\mathbb{C}^2$:

$$U_x = \{(1, Y, Z)\} \quad \text{given by rescaling } x \text{ to } 1$$
$$U_y = \{(X, 1, Z)\} \quad \text{given by rescaling } y \text{ to } 1$$
$$U_z = \{(X, Y, 1)\} \quad \text{given by rescaling } z \text{ to } 1$$

Here, for example, in the case U_x we take $x \neq 0$ and set $Y = y/x$, $Z = z/x$.

Although the projective space $\mathbb{P}^2$ is smooth, there are closely related Fano varieties called weighted projective spaces [20, 36] that have singularities. For example, consider the weighted projective plane $\mathbb{P}(1, 2, 3)$: this is the quotient of $\mathbb{C}^3 \setminus \{0\}$ by $\mathbb{C}^\times$, where the action of $\lambda \in \mathbb{C}^\times$ identifies the points (x, y, z) and $(\lambda x, \lambda^2 y, \lambda^3 z)$. Let us write

$$\mu_n = \{e^{2\pi k i/n} \mid k \in \mathbb{Z}\}$$

for the group of nth roots of unity. The variety $\mathbb{P}(1, 2, 3)$ is once again covered by open sets

$$U_x = \{(1, Y, Z)\} \quad \text{given by rescaling } x \text{ to } 1$$
$$U_y = \{(X, 1, Z)\} \quad \text{given by rescaling } y \text{ to } 1$$
$$U_z = \{(X, Y, 1)\} \quad \text{given by rescaling } z \text{ to } 1$$

but this time we have $U_x \cong \mathbb{C}^2$, $U_y \cong \mathbb{C}^2/\mu_2$, and $U_z = \mathbb{C}^2/\mu_3$. This is because, for example, when we choose $\lambda \in \mathbb{C}^\times$ to rescale (x, y, z) with $z \neq 0$ to $(X, Y, 1)$, there are three possible choices for λ and they differ by the action of μ_3. In particular this lets us see that $\mathbb{P}(1, 2, 3)$ is singular. For example, functions on the chart $U_y \cong \mathbb{C}^2/\mu_2$ are polynomials in X and Z that are invariant under $X \mapsto -X$, $Z \mapsto -Z$, or in other words

$$U_y = \operatorname{Spec} \mathbb{C}[X^2, XZ, Z^2]$$
$$= \operatorname{Spec} \mathbb{C}[a, b, c]/(ac - b^2)$$

Thus the chart U_y is the solution set for the equation $ac - b^2 = 0$, as pictured in Figure 10(a). Similarly, the chart $U_z \cong \mathbb{C}^2/\mu_3$ can be written as

$$U_z = \operatorname{Spec} \mathbb{C}[X^3, XY, Y^3]$$
$$= \operatorname{Spec} \mathbb{C}[a, b, c]/(ac - b^3)$$

and is the solution set to the equation $ac - b^3 = 0$, as pictured in Figure 10(b). The variety $\mathbb{P}(1, 2, 3)$ has singular points at $(0, 1, 0) \in U_y$ and $(0, 0, 1) \in U_z$, and away from these points it is smooth.

There are weighted projective spaces of any dimension. Let $a_1, a_2, \ldots, a_N$ be positive integers such that any subset of size $N - 1$ has no common factor, and consider

$$\mathbb{P}(a_1, a_2, \ldots, a_N) = (\mathbb{C}^N \setminus \{0\})/\mathbb{C}^\times$$

where the action of $\lambda \in \mathbb{C}^\times$ identifies the points

$$(x_1, x_2, \ldots, x_N) \quad \text{and} \quad (\lambda^{a_1} x_1, \lambda^{a_2} x_2, \ldots, \lambda^{a_N} x_N)$$

in $\mathbb{C}^N \setminus \{0\}$. The quotient $\mathbb{P}(a_1, a_2, \ldots, a_N)$ is an algebraic variety of dimension $N - 1$. A general point of $\mathbb{P}(a_1, a_2, \ldots, a_N)$ is smooth, but there can be singular points. Indeed, $\mathbb{P}(a_1, a_2, \ldots, a_N)$ is covered by N open sets

$$U_i = \{(X_1, \ldots, X_{i-1}, 1, X_{i+1}, \ldots, X_N)\} \qquad i \in \{1, 2, \ldots, N\}$$

given by rescaling x_i to 1; here we take $x_i \neq 0$ and set $X_j = x_j/x_i$. The chart U_i is isomorphic to $\mathbb{C}^{N-1}/\mu_{a_i}$, where μ_{a_i} acts on $\mathbb{C}^{N-1}$ with weights a_j, $j \neq i$. In Reid's notation, this is the cyclic quotient singularity $\frac{1}{a_i}(a_1, \ldots, \widehat{a_i}, \ldots, a_N)$; it is smooth if and only if $a_i = 1$.

The topology of weighted projective space is very simple, with

$$H^k\big(\mathbb{P}(a_1, a_2, \ldots, a_N); \mathbb{Q}\big) = \begin{cases} \mathbb{Q} & \text{if } 0 \leq k \leq 2N - 2 \text{ and } k \text{ is even}; \\ 0 & \text{otherwise.} \end{cases}$$

Hence every weighted projective space has second Betti number $b_2 = 1$. There is a closed formula [9, Proposition D.9] for the regularized quantum period of $X = \mathbb{P}(a_1, a_2, \ldots, a_N)$:

$$\widehat{G}_X(t) = \sum_{k=0}^{\infty} \frac{(ak)!}{(a_1 k)!(a_2 k)! \cdots (a_N k)!} t^{ak} \tag{10}$$

where $a = a_1 + a_2 + \cdots + a_N$.

Toric varieties of Picard rank 2. As well as weighted projective spaces, which are quotients of $\mathbb{C}^N \setminus \{0\}$ by an action of $\mathbb{C}^\times$, we will consider varieties that arise as quotients of $\mathbb{C}^N \setminus S$ by $\mathbb{C}^\times \times \mathbb{C}^\times$, where S is a union of linear subspaces. These are examples of *toric varieties* [16, 25]. Specifically, consider a matrix

$$\begin{pmatrix} a_1 & a_2 & \cdots & a_N \\ b_1 & b_2 & \cdots & b_N \end{pmatrix} \tag{11}$$

with non-negative integer entries and no zero columns. This defines an action of $\mathbb{C}^\times \times \mathbb{C}^\times$ on $\mathbb{C}^N$, where $(\lambda, \mu) \in \mathbb{C}^\times \times \mathbb{C}^\times$ identifies the points

$$(x_1, x_2, \ldots, x_N) \quad \text{and} \quad (\lambda^{a_1}\mu^{b_1} x_1, \lambda^{a_2}\mu^{b_2} x_2, \ldots, \lambda^{a_N}\mu^{b_N} x_N)$$

in $\mathbb{C}^N$. Set $a = a_1 + a_2 + \ldots + a_N$ and $b = b_1 + b_2 + \cdots + b_N$, and suppose that (a, b) is not a scalar multiple of (a_i, b_i) for any i. This determines linear subspaces

$$S_+ = \{(x_1, x_2, \ldots, x_N) \mid x_i = 0 \text{ if } b_i/a_i < b/a\}$$
$$S_- = \{(x_1, x_2, \ldots, x_N) \mid x_i = 0 \text{ if } b_i/a_i > b/a\}$$

of $\mathbb{C}^N$, and we consider the quotient

$$X = (\mathbb{C}^N \setminus S)/(\mathbb{C}^\times \times \mathbb{C}^\times) \tag{12}$$

where $S = S_+ \cup S_-$. See e.g. [5, §A.5].

These quotients behave in many ways like weighted projective spaces. Indeed, if we take the weight matrix (11) to be

$$\begin{pmatrix} a_1 & a_2 & \cdots & a_N & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

then X coincides with $\mathbb{P}(a_1, a_2, \ldots, a_N)$. We will consider only weight matrices such that the subspaces S_+ and S_- both have dimension two or more; this implies that the second Betti number $b_2(X) = 2$, and hence X is not a weighted projective space. We will refer to such quotients (12) as *toric varieties of Picard rank two*, because general theory implies that the Picard lattice of X has rank two. The dimension of X is $N - 2$. As for weighted projective spaces, toric varieties of Picard rank two can have singular points, the precise form of which is determined by the weights (11). There is also a closed

Dimension			
1	2	3	4
$\mathbb{P}^1$	$\mathbb{P}^2$	7 cases see [41]	28 686 cases see [43]

Table 2. The classification of low-dimensional weighted projective spaces with terminal quotient singularities.

formula [9, Proposition C.2] for the regularized quantum period. Let C denote the cone in $\mathbb{R}^2$ defined by the equations $a_i x + b_i y \geq 0$, $i \in \{1, 2, \ldots, N\}$. Then

$$\widehat{G}_X(t) = \sum_{(k,l) \in \mathbb{Z}^2 \cap C} \frac{(ak + bl)!}{(a_1 k + b_1 l)!(a_2 k + b_2 l)! \cdots (a_N k + b_N l)!} t^{ak+bl} \tag{13}$$

Classification results. Weighted projective spaces with terminal quotient singularities have been classified in dimensions up to four; see Table 2 for a summary. There are 35 three-dimensional Fano toric varieties with terminal quotient singularities and Picard rank two [41]. There is no known classification of Fano toric varieties with terminal quotient singularities in higher dimension, even when the Picard rank is two.

Appendix B. Supplementary Methods 1

Data analysis: weighted projective spaces. We computed an initial segment $(c_0, c_1, \ldots, c_m)$ of the regularized quantum period, with $m \approx 100\,000$, for all the examples in the sample of $150\,000$ weighted projective spaces with terminal quotient singularities. We then considered $\{\log c_d\}_{d \in S}$ where $S = \{d \in \mathbb{Z}_{\geq 0} \mid c_d \neq 0\}$. To reduce dimension we fitted a linear model to the set $\{(d, \log c_d) \mid d \in S\}$ and used the slope and intercept of this model as features. The linear fit produces a close approximation of the data. Figure 11 shows the distribution of the standard errors for the slope and the y-intercept: the errors for the slope are between 3.9×10^{-8} and 1.4×10^{-5}, and the errors for the y-intercept are between 0.0022 and 0.82. As we will see below, the standard error for the y-intercept is a good proxy for the accuracy of the linear model. This accuracy decreases as the dimension grows – see Figure 11(c) – but we will see below that this does not affect the accuracy of the machine learning classification.

Data analysis: toric varieties of Picard rank 2. We fitted a linear model to the set $\{(d, \log c_d) \mid d \in S\}$ where $S = \{d \in \mathbb{Z}_{\geq 0} \mid c_d \neq 0\}$, and used the slope and intercept of this linear model as features. The distribution of standard errors for the slope and y-intercept of the linear model are shown in Figure 12. The standard errors for the slope are small compared to the range of slopes, but in many cases the standard error for the y-intercept is relatively large. As Figure 13 illustrates, discarding data points where the standard error s_{int} for the y-intercept exceeds some threshold reduces apparent noise. As discussed above, we believe that this reflects inaccuracies in the linear regression caused by oscillatory behaviour in the initial terms of the quantum period sequence.

Example 10. Let us consider in more detail the toric variety from Example 3. In Figure 14 we plot $\log c_d$ along with its linear approximation. Figure 14(a) shows only the first 250 terms, whilst Figure 14(b) shows the interval between the 1000th and the 1250th term. We see considerable deviation from the linear approximation among the first 250 terms; the deviation reduces for larger d.

Appendix C. Supplementary Methods 2

We performed our experiments using scikit-learn [54], a standard machine learning library for Python. The computations that produced the data shown in Figure 7(a) were performed using Mathematica [37]. All code required to replicate the results in this paper is available from Bitbucket under an MIT license [13].

(a)   (b)

(c)
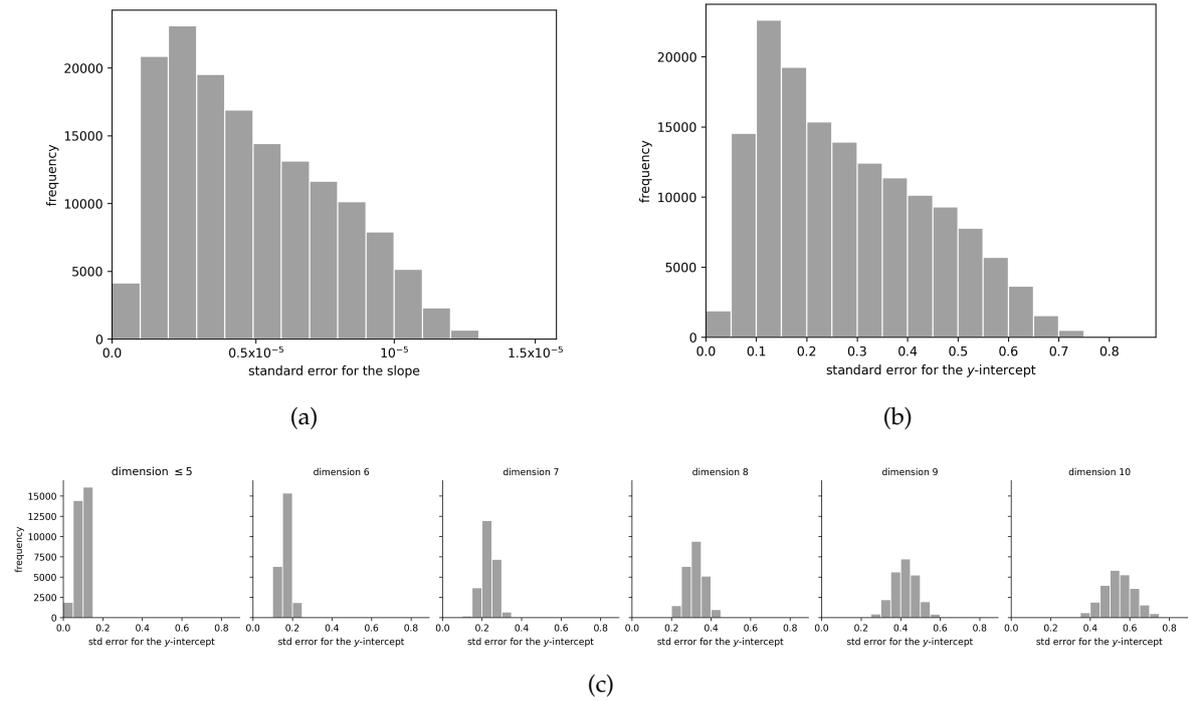
Figure 11. Standard errors for the slope and y-intercept. The distribution of standard errors for the slope and y-intercept from the linear model applied to weighted projective spaces X with terminal quotient singularities: (a) standard error for the slope. (b) standard error for the y-intercept. (c) standard error for the y-intercept by dimension.

(a)   (b)
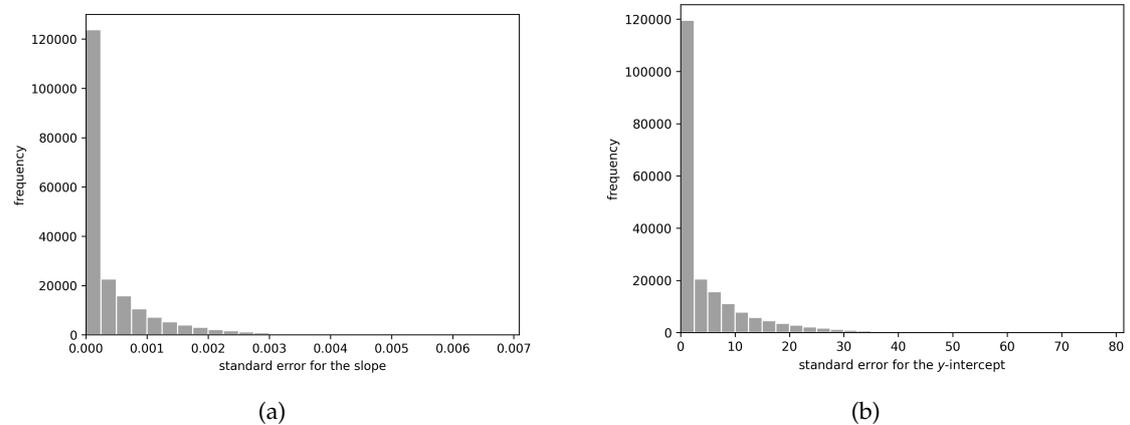
Figure 12. Standard errors for the slope and y-intercept. The distribution of standard errors for the slope and y-intercept from the linear model applied to toric varieties of Picard rank two with terminal quotient singularities: (a) standard error for the slope. (b) standard error for the y-intercept.

Weighted projective spaces. We excluded dimensions one and two from the analysis, since there is only one weighted projective space in each case (namely $\mathbb{P}^1$ and $\mathbb{P}^2$). Therefore we have a dataset of 149 998 slope-intercept pairs, labelled by the dimension which varies between three and ten. We standardised the features, by translating the means to zero and scaling to unit variance, and applied a Support Vector Machine (SVM) with linear kernel and regularisation parameter $C = 10$. By looking at different train–test splits we obtained the learning curves shown in Figure 15. The figure displays the mean accuracies for both training and validation data obtained by performing five random test-train splits each time: the shaded areas around the lines correspond to the 1σ region, where σ denotes the

(a)

(b)

(c)
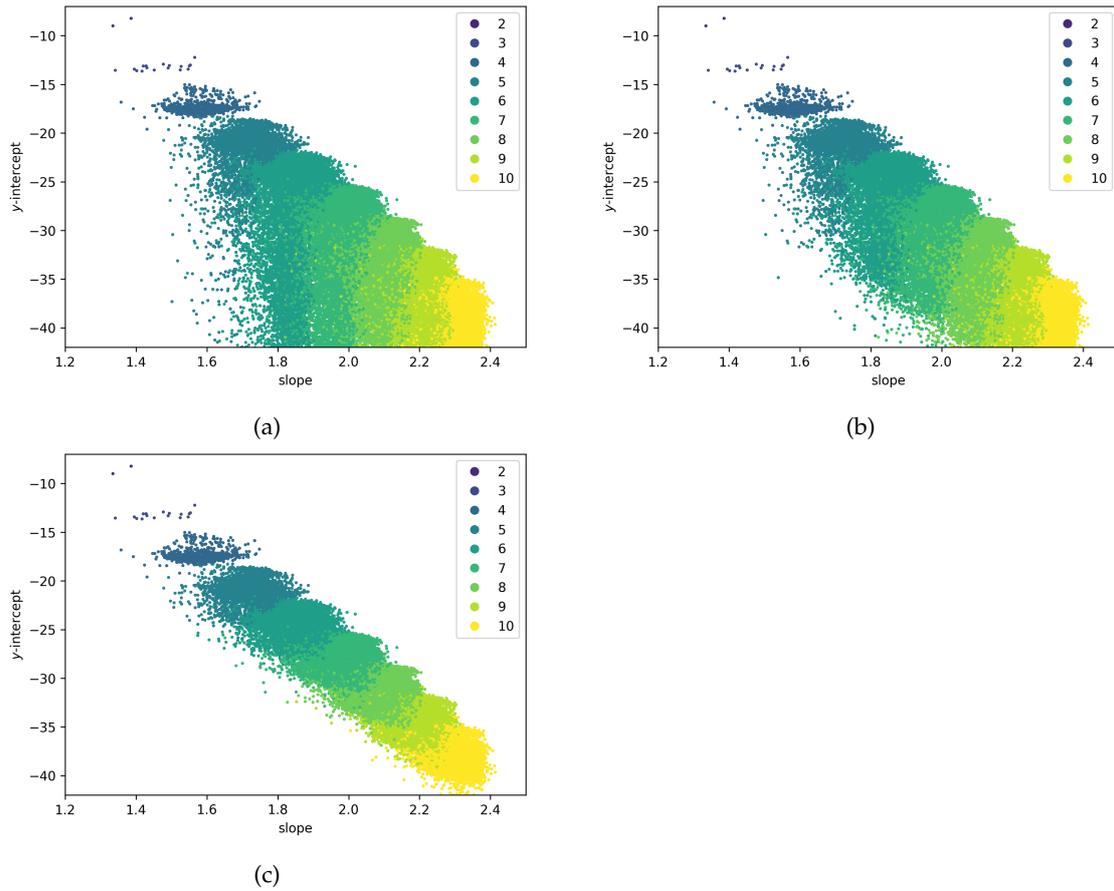
Figure 13. The slopes and y-intercepts from the linear model applied to toric varieties of Picard rank two with terminal quotient singularities. Data points are selected according to the standard error s_{int} for the y-intercept. The colour records the dimension of the toric variety. (a) All data points. (b) Points with $s_{\text{int}} < 1$: $101\,183/200000$ points. (c) Points with $s_{\text{int}} < 0.3$: $67\,445/200000$ points.

(a)

(b)
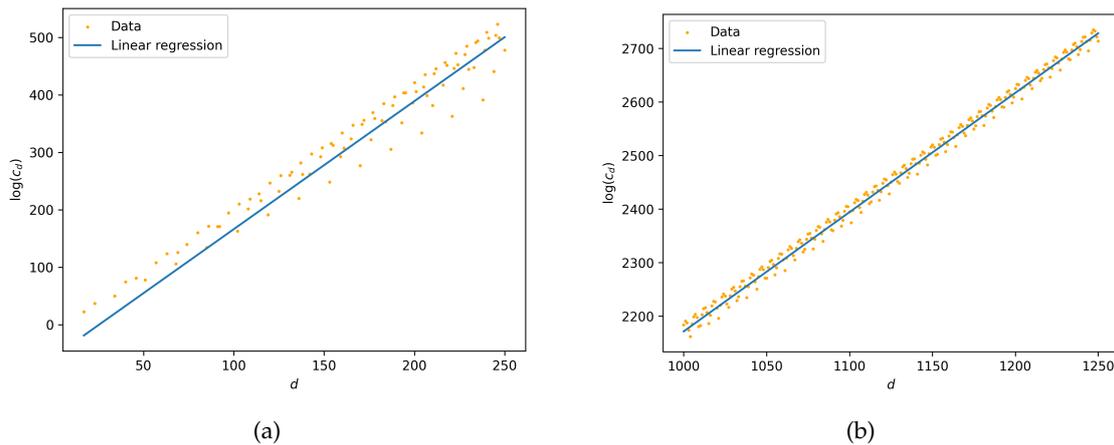
Figure 14. The logarithm of the non-zero coefficients c_d for Example 3: (a) the first 250 terms. (b) terms between $d = 1000$ and $d = 1250$. In each case, the linear approximation is also shown.
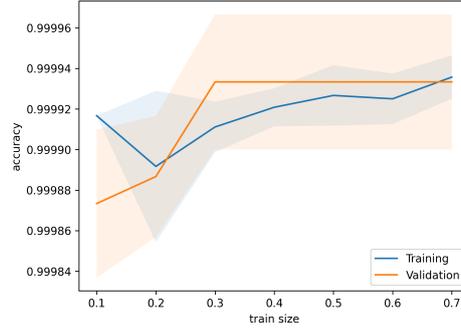
Figure 15. Learning curves for a Support Vector Machine with linear kernel applied to the dataset of weighted projective spaces. The plot shows the means of the training and validation accuracies for five different random train–test splits. The shaded regions show the 1σ interval, where σ is the standard deviation.
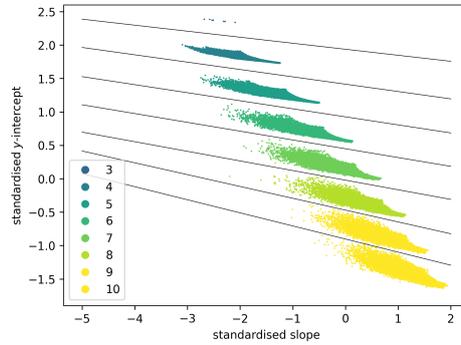
Figure 16. Decision boundaries computed from a Support Vector Machine with linear kernel trained on 70% of the dataset of weighted projective spaces. Note that the data has been standardised.

standard deviation. Using 10% (or more) of the data for training we obtained an accuracy of 99.99%. In Figure 16 we plot the decision boundaries computed by the SVM between neighbouring dimension classes.

Toric varieties of Picard rank 2. In light of the discussion above, we restricted attention to toric varieties with Picard rank two such that the y-intercept standard error s_{int} is less than 0.3. We also excluded dimension two from the analysis, since in this case there are only two varieties (namely, $\mathbb{P}^1 \times \mathbb{P}^1$ and the Hirzebruch surface $\mathbb{F}_1$). The resulting dataset contains 67 443 slope-intercept pairs, labelled by dimension; the dimension varies between three and ten, as shown in Table 3.

Support Vector Machine. We used a linear SVM with regularisation parameter $C = 50$. By considering different train–test splits we obtained the learning curves shown in Figure 17, where the means and the standard deviations were obtained by performing five random samples for each split. Note that the model did not overfit. We obtained a validation accuracy of 88.2% using 70% of the data for training. Figure 18 shows the decision boundaries computed by the SVM between neighbouring dimension classes. Figure 19 shows the confusion matrices for the same train–test split.

Random Forest Classifier. We used a Random Forest Classifier (RFC) with 1500 estimators and the same features (slope and y-intercept for the linear model). By considering different train–test splits we obtained the learning curves shown in Figure 20; note again that the model did not overfit. Using 70% of the data for training, the RFC gave a validation accuracy of 89.4%. Figure 21 on page 22 shows confusion matrices for the same train–test split.
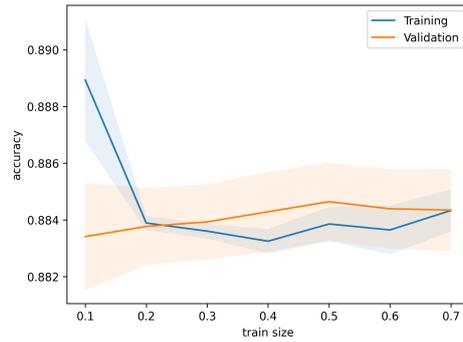
Figure 17. Learning curves for a Support Vector Machine with linear kernel applied to the dataset of toric varieties of Picard rank two. The plot shows the means of the training and validation accuracies for five different random train–test splits. The shaded regions show the 1σ interval, where σ is the standard deviation.
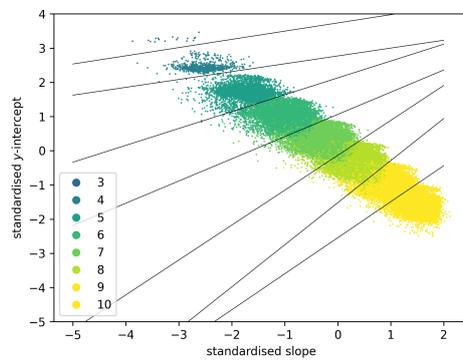
Figure 18. Decision boundaries computed from a Support Vector Machine with linear kernel trained on 70% of the dataset of toric varieties of Picard rank two. Note that the data has been standardised.
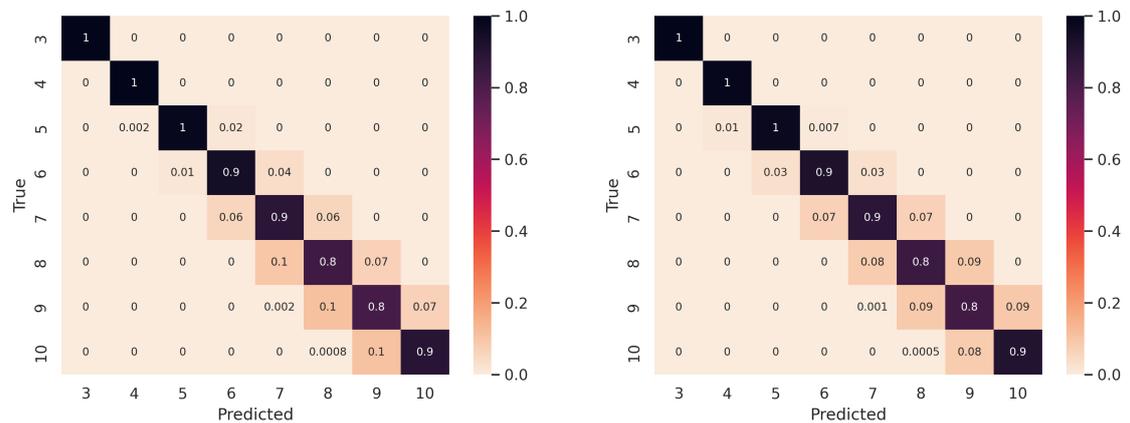
(a) Confusion matrix normalised with respect to the true values.

(b) Confusion matrix normalised with respect to the predicted values.

Figure 19. Confusion matrices for a Support Vector Machine with linear kernel trained on 70% of the dataset of toric varieties of Picard rank two.

Rank-two toric varieties with $s_{\text{int}} < 0.3$		
Dimension	Sample size	Percentage
3	17	0.025
4	758	1.124
5	5 504	8.161
6	12 497	18.530
7	16 084	23.848
8	13 701	20.315
9	10 638	15.773
10	8 244	12.224
Total	67 443	

Table 3. The distribution by dimension among toric varieties of Picard rank two in our dataset with $s_{\text{int}} < 0.3$.
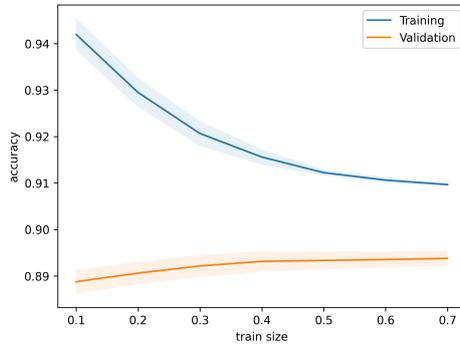
Figure 20. Learning curves for a Random Forest Classifier applied to the dataset of toric varieties of Picard rank two. The plot shows the means of the training and validation accuracies for five different random train–test splits. The shaded regions show the 1σ interval, where σ is the standard deviation.
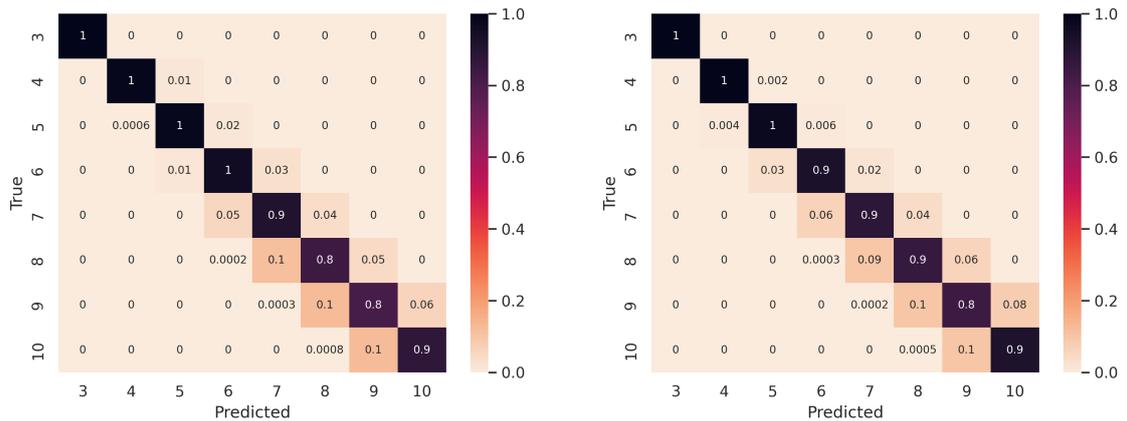
(a) Confusion matrix normalised with respect to the true values.

(b) Confusion matrix normalised with respect to the predicted values.

Figure 21. Confusion matrices for a Random Forest Classifier trained on 70% of the dataset of toric varieties of Picard rank two.
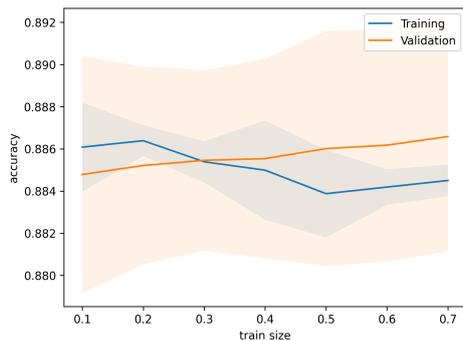
Figure 22. Learning curves for a Multilayer Perceptron classifier MLP_2 applied to the dataset of toric varieties of Picard rank two and dimension at least six, using just the regression data as features. The plot shows the means of the training and validation accuracies for five different random train–test splits. The shaded regions show the 1σ interval, where σ is the standard deviation.
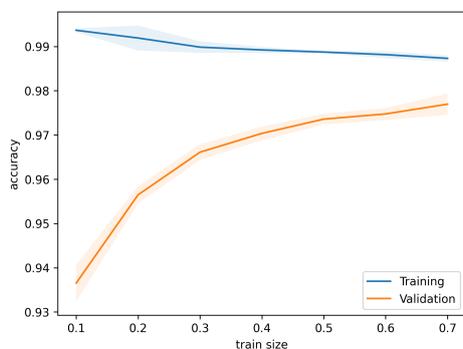
Figure 23. Learning curves for a Multilayer Perceptron classifier MLP_{102} applied to the dataset of toric varieties of Picard rank two and dimension at least six, using as features the regression data as well as $\log c_d$ for $1 \leq d \leq 100$. The plot shows the means of the training and validation accuracies for five different random train–test splits. The shaded regions show the 1σ interval, where σ is the standard deviation.

Feed-forward neural network. As discussed above, neural networks do not handle unbalanced datasets well, and therefore we removed the toric varieties with dimensions three, four, and five from our dataset: see Table 3. We trained a Multilayer Perceptron (MLP) classifier on the same features, using an MLP with three hidden layers $(10, 30, 10)$, Adam optimiser [44], and rectified linear activation function [2]. Different train–test splits produced the learning curve in Figure 22; again the model did not overfit. Using 70% of the data for training, the MLP gave a validation accuracy of 88.7%. One could further balance the dataset, by randomly undersampling so that there are the same number of representatives in each dimension (8244 representatives: see Table 3). This resulted in a slight decrease in accuracy: the better balance was outweighed by loss of data caused by undersampling.

Feed-forward neural network with many features. We trained an MLP with the same architecture, but supplemented the features by including $\log c_d$ for $1 \leq d \leq 100$ (unless c_d was zero in which case we set that feature to zero), as well as the slope and y-intercept as before. We refer to the previous neural network as MLP_2, because it uses 2 features, and refer to this neural network as MLP_{102}, because it uses 102 features. Figure 23 shows the learning curves obtained for different train–test splits. Using 70% of the data for training, the MLP_{102} model gave a validation accuracy of 97.7%.

We do not understand the reason for the performance improvement between MLP_{102} and MLP_2. But one possible explanation is the following. Recall that the first 1000 terms of the period sequence were excluded when calculating the slope and intercept, because they exhibit irregular oscillations

ML models			
SVM	RFC	MLP_2	MLP_{102}
87.7%	88.6%	88.7%	97.7%

Table 4. Comparison of model accuracies. Accuracies for various models applied to the dataset of toric varieties of Picard rank two and dimension at least six: a Support Vector Machine with linear kernel, a Random Forest Classifier, and the neural networks MLP_2 and MLP_{102}.

that decay as d grows. These oscillations reduce the accuracy of the linear regression. The oscillations may, however, carry information about the toric variety, and so including the first few values of $\log(c_d)$ potentially makes more information available to the model. For example, examining the pattern of zeroes at the beginning of the sequence (c_d) sometimes allows one to recover the values of a and b – see (13) for the notation. This information is relevant to estimating the dimension because, as a very crude approximation, larger a and b go along with larger dimension. Omitting the slope and intercept, however, and training on the coefficients $\log c_d$ for $1 \le d \le 100$ with the same architecture gave an accuracy of only 62%.

Comparison of models. The validation accuracies of the SVM, RFC, and the neural networks MLP_2 and MLP_{102}, on the same data set ($s_{int} < 0.3$, dimension between six and ten), are compared in Table 4. Their confusion matrices are shown in Table 5. All models trained on only the regression data performed well, with the RFC slightly more accurate than the SVM and the neural network MLP_2 slightly more accurate still. Misclassified examples are generally in higher dimension, which is consistent with the idea that misclassification is due to convergence-related noise. The neural network trained on the supplemented feature set, MLP_{102}, outperforms all other models. However, as discussed above, feature importance analysis using SHAP values showed that the slope and the intercept were the most influential features in the prediction.

Appendix D. Supplementary Discussion

Comparison with Principal Component Analysis. An alternative approach to dimensionality reduction, rather than fitting a linear model to $\log c_d$, would be to perform Principal Component Analysis (PCA) on this sequence and retain only the first few principal components. Since the vectors (c_d) have different patterns of zeroes – c_d is non-zero only if d is divisible by the Fano index r of X – we need to perform PCA for Fano varieties of each index r separately. We analysed this in the weighted projective space case, finding that for each r the first two components of PCA are related to the growth coefficients (A, B) from Theorem 5 by an invertible affine-linear transformation. That is, our analysis suggests that the coefficients (A, B) contain exactly the same information as the first two components of PCA. Note, however, that the affine-linear transformation that relates PCA to (A, B) varies with the Fano index r. Using A and B as features therefore allows for meaningful comparison between Fano varieties of different index. Furthermore, unlike PCA-derived values, the coefficients (A, B) can be computed for a single Fano variety, rather than requiring a sufficiently large collection of Fano varieties of the same index.

Towards more general Fano varieties. Weighted projective spaces and toric varieties of Picard rank two are very special among Fano varieties. It is hard to quantify this, because so little is known about Fano classification in the higher-dimensional and non-smooth cases, but for example this class includes only 18% of the $\mathbb{Q}$-factorial terminal Fano toric varieties in three dimensions. On the other hand, one can regard weighted projective spaces and toric varieties of Picard rank two as representative of a much broader class of algebraic varieties called toric complete intersections. Toric complete intersections share the key properties that we used to prove Theorems 5 and 6 – geometry that is tightly controlled by combinatorics, including explicit expressions for genus-zero Gromov–Witten invariants in terms of hypergeometric functions – and we believe that the rigorous results of this paper will generalise to the toric complete intersection case. All smooth two-dimensional Fano varieties and 92 of the 105 smooth

Model	True confusion matrix	Predicted confusion matrix
SVM		
RFC		
MLP$_2$		
MLP$_{102}$		
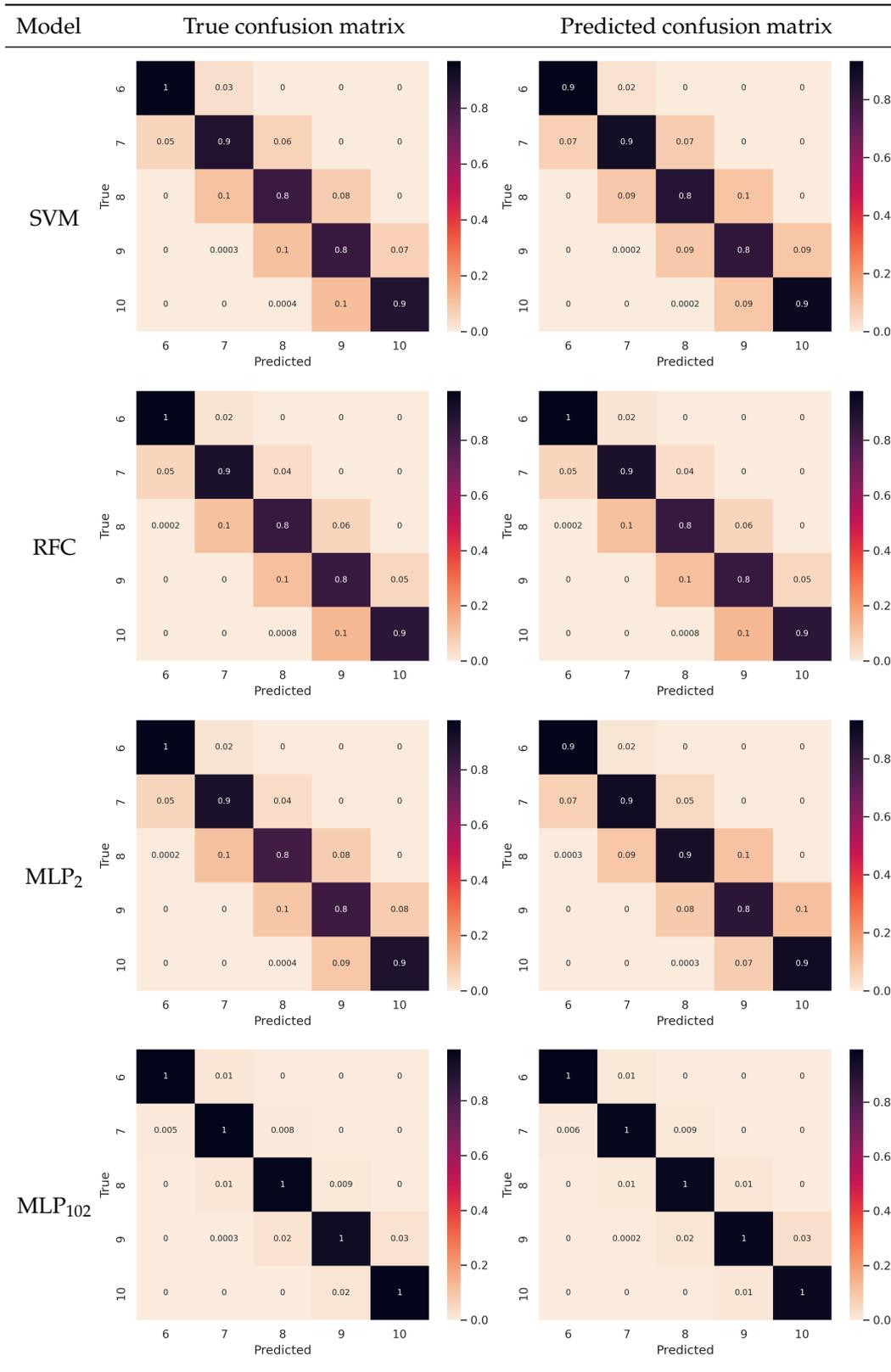
Table 5. Comparison of confusion matrices. Confusion matrices for various models applied to the dataset of toric varieties of Picard rank two and dimension at least six: a Support Vector Machine with linear kernel, a Random Forest Classifier, and the neural networks MLP$_2$ and MLP$_{102}$.

three-dimensional Fano varieties are toric complete intersections [9]. Many theorems in algebraic geometry were first proved for toric varieties and later extended to toric complete intersections and more general algebraic varieties; cf. [26, 27, 33] and [28, 56].

The machine learning paradigm presented here, however, applies much more broadly. Since our models take only the regularized quantum period sequence as input, we expect that whenever we can calculate $\widehat{G}_X$ – which is the case for almost all known Fano varieties – we should be able to apply a machine learning pipeline to extract geometric information about X.

REFERENCES

[1] Jeffrey Adams, Annegret Paul, Ran Cui, Susana Salamanca-Riba, Peter Trapa, Marc van Leeuwen, and David Vogan. Atlas of Lie groups and representations. Online, 2016. http://www.liegroups.org.

[2] Abien Fred Agarap. Deep learning using rectified linear units (ReLU). arXiv:1803.08375 [cs.NE], 2018.

[3] M. F. Atiyah, N. J. Hitchin, V. G. Drinfeld, and Yu. I. Manin. Construction of instantons. *Phys. Lett. A*, 65(3):185–187, 1978. doi:10.1016/0375-9601(78)90141-X.

[4] Wieb Bosma, John Cannon, and Catherine Playoust. The Magma algebra system. I. The user language. *J. Symbolic Comput.*, 24(3-4):235–265, 1997. doi:10.1006/jsco.1996.0125.

[5] Gavin Brown, Alessio Corti, and Francesco Zucconi. Birational geometry of 3-fold Mori fibre spaces. In *The Fano Conference*, pages 235–275. Univ. Torino, Turin, 2004.

[6] P. Candelas, Gary T. Horowitz, Andrew Strominger, and Edward Witten. Vacuum configurations for superstrings. *Nuclear Phys. B*, 258(1):46–74, 1985. doi:10.1016/0550-3213(85)90602-9.

[7] Philip Candelas, Xenia C. de la Ossa, Paul S. Green, and Linda Parkes. A pair of Calabi-Yau manifolds as an exactly soluble superconformal theory. *Nuclear Phys. B*, 359(1):21–74, 1991. doi:10.1016/0550-3213(91)90292-6.

[8] Tom Coates, Alessio Corti, Sergey Galkin, Vasily Golyshev, and Alexander M. Kasprzyk. Mirror symmetry and Fano manifolds. In *European Congress of Mathematics*, pages 285–300. Eur. Math. Soc., Zürich, 2013.

[9] Tom Coates, Alessio Corti, Sergey Galkin, and Alexander M. Kasprzyk. Quantum periods for 3-dimensional Fano manifolds. *Geom. Topol.*, 20(1):103–256, 2016. doi:10.2140/gt.2016.20.103.

[10] Tom Coates and Alexander M. Kasprzyk. Databases of quantum periods for Fano manifolds. *Scientific Data*, 9(1):163, 2022.

[11] Tom Coates, Alexander M. Kasprzyk, and Sara Veneziale. A dataset of 150000 terminal weighted projective spaces. *Zenodo*, 2022. doi:10.5281/zenodo.5790079.

[12] Tom Coates, Alexander M. Kasprzyk, and Sara Veneziale. A dataset of 200000 terminal toric varieties of Picard rank 2. *Zenodo*, 2022. doi:10.5281/zenodo.5790096.

[13] Tom Coates, Alexander M. Kasprzyk, and Sara Veneziale. Supporting code. https://bitbucket.org/fanosearch/mldim, 2022.

[14] J. H. Conway, R. T. Curtis, S. P. Norton, R. A. Parker, and R. A. Wilson. 𝔸𝕋𝕃𝔸𝕊 *of finite groups*. Oxford University Press, Eynsham, 1985. Maximal subgroups and ordinary characters for simple groups. With computational assistance from J. G. Thackray.

[15] David A. Cox and Sheldon Katz. *Mirror symmetry and algebraic geometry*, volume 68 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1999. doi:10.1090/surv/068.

[16] David A. Cox, John B. Little, and Henry K. Schenck. *Toric varieties*, volume 124 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2011. doi:10.1090/gsm/124.

[17] John Cremona. The L-functions and modular forms database project. *Found. Comput. Math.*, 16(6):1541–1553, 2016. doi:10.1007/s10208-016-9306-z.

[18] Alex Davies, Petar Veličković, Lars Buesing, Sam Blackwell, Daniel Zheng, Nenad Tomašev, Richard Tanburn, Peter Battaglia, Charles Blundell, András Juhász, Marc Lackenby, Geordie Williamson, Demis Hassabis, and Pushmeet Kohli. Advancing mathematics by guiding human intuition with AI. *Nature*, 600:70–74, 2021. doi:10.1038/s41586-021-04086-x.

[19] Pasquale Del Pezzo. Sulle superficie dell'n^{mo} ordine immerse nello spazio ad n dimensioni. *Rend. del Circolo Mat. di Palermo*, 1:241–255, 1887.

[20] Igor Dolgachev. Weighted projective varieties. In *Group actions and vector fields (Vancouver, B.C., 1981)*, volume 956 of *Lecture Notes in Math.*, pages 34–71. Springer, Berlin, 1982. doi:10.1007/BFb0101508.

[21] Harold Erbin and Riccardo Finotello. Inception neural network for complete intersection Calabi–Yau 3-folds. *Machine Learning: Science and Technology*, 2(2):02LT03, 2021.

[22] Nicholas Eriksson, Kristian Ranestad, Bernd Sturmfels, and Seth Sullivant. Phylogenetic algebraic geometry. In *Projective varieties with unexpected properties*, pages 237–255. Walter de Gruyter, Berlin, 2005.

[23] European Organization For Nuclear Research and OpenAIRE. Zenodo, 2013. doi:10.25495/7GXK-RD71.

[24] Gino Fano. Nuove ricerche sulle varietà algebriche a tre dimensioni a curve-sezioni canoniche. *Pont. Acad. Sci. Comment.*, 11:635–720, 1947.

[25] William Fulton. *Introduction to toric varieties*, volume 131 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 1993. doi:10.1515/9781400882526.

[26] Alexander Givental. A mirror theorem for toric complete intersections. In *Topological field theory, primitive forms and related topics (Kyoto, 1996)*, volume 160 of *Progr. Math.*, pages 141–175. Birkhäuser Boston, Boston, MA, 1998.

[27] Alexander B. Givental. Equivariant Gromov-Witten invariants. *Internat. Math. Res. Notices*, (13):613–663, 1996. doi:10.1155/S1073792896000414.

[28] Alexander B. Givental. Semisimple Frobenius structures at higher genus. *Internat. Math. Res. Notices*, (23):1265–1286, 2001. doi:10.1155/S1073792801000605.

[29] Boris V Gnedenko. *Theory of probability*. Routledge, 2018.

[30] B. R. Greene. String theory on Calabi-Yau manifolds. In *Fields, strings and duality (Boulder, CO, 1996)*, pages 543–726. World Sci. Publ., River Edge, NJ, 1997.

[31] B. R. Greene and M. R. Plesser. Duality in Calabi-Yau moduli space. *Nuclear Phys. B*, 338(1):15–37, 1990. doi:10.1016/0550-3213(90)90622-K.

[32] Roland Grinis and Alexander M. Kasprzyk. Normal forms of convex lattice polytopes. arXiv:1301.6641 [math.CO], 2013.

[33] Mark Gross and Bernd Siebert. Intrinsic mirror symmetry. arXiv:1909.07649 [math.AG], 2019.

[34] Yang-Hui He. Machine-learning mathematical structures. *International Journal of Data Science in the Mathematical Sciences*, 1:23–47, 2023.

[35] Kentaro Hori and Cumrun Vafa. Mirror symmetry. arXiv:hep-th/0002222, 2000.

[36] A. R. Iano-Fletcher. Working with weighted complete intersections. In *Explicit birational geometry of 3-folds*, volume 281 of *London Math. Soc. Lecture Note Ser.*, pages 101–173. Cambridge Univ. Press, Cambridge, 2000.

[37] Wolfram Research, Inc. Mathematica, Version 13.1. Champaign, IL, 2022.

[38] V. A. Iskovskih. Fano threefolds. I. *Izv. Akad. Nauk SSSR Ser. Mat.*, 41(3):516–562, 717, 1977.

[39] V. A. Iskovskih. Fano threefolds. II. *Izv. Akad. Nauk SSSR Ser. Mat.*, 42(3):506–549, 1978.

[40] V. A. Iskovskih. Anticanonical models of three-dimensional algebraic varieties. In *Current problems in mathematics, Vol. 12 (Russian)*, pages 59–157, 239 (loose errata). VINITI, Moscow, 1979.

[41] Alexander M. Kasprzyk. Toric Fano three-folds with terminal singularities. *Tohoku Math. J. (2)*, 58(1):101–121, 2006. doi:10.2748/tmj/1145390208.

[42] Alexander M. Kasprzyk. Bounds on fake weighted projective space. *Kodai Math. J.*, 32(2):197–208, 2009. doi:10.2996/kmj/1245982903.

[43] Alexander M. Kasprzyk. Classifying terminal weighted projective space. arXiv:1304.3029 [math.AG], 2013.

[44] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. arXiv:1412.6980 [cs.LG], 2014.

[45] János Kollár. The structure of algebraic threefolds: an introduction to Mori's program. *Bull. Amer. Math. Soc. (N.S.)*, 17(2):211–273, 1987. doi:10.1090/S0273-0979-1987-15548-0.

[46] János Kollár and Shigefumi Mori. *Birational geometry of algebraic varieties*, volume 134 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1998. doi:10.1017/CBO9780511662560.

[47] Maximilian Kreuzer and Harald Skarke. Complete classification of reflexive polyhedra in four dimensions. *Adv. Theor. Math. Phys.*, 4(6):1209–1230, 2000. doi:10.4310/ATMP.2000.v4.n6.a2.

[48] Maximilian Kreuzer and Harald Skarke. PALP: a package for analysing lattice polytopes with applications to toric geometry. *Comput. Phys. Comm.*, 157(1):87–106, 2004. doi:10.1016/S0010-4655(03)00491-0.

[49] Jesse SF Levitt, Mustafa Hajij, and Radmila Sazdanovic. Big data approaches to knot theory: Understanding the structure of the Jones polynomial. *Journal of Knot Theory and Its Ramifications*, 31(13):2250095, 2022.

[50] Scott M Lundberg and Su-In Lee. A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30:4765–4774, 2017.

[51] Shigefumi Mori and Shigeru Mukai. Classification of Fano 3-folds with $B_2 \geq 2$. *Manuscripta Math.*, 36(2):147–162, 1981/82. doi:10.1007/BF01170131.

[52] Shigefumi Mori and Shigeru Mukai. Erratum: "Classification of Fano 3-folds with $B_2 \geq 2$". *Manuscripta Math.*, 110(3):407, 2003. doi:10.1007/s00229-002-0336-2.

[53] Harald Niederreiter and Chaoping Xing. *Algebraic geometry in coding theory and cryptography*. Princeton University Press, Princeton, NJ, 2009.

[54] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[55] Joseph Polchinski. *String theory. Vol. II*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, Cambridge, 2005. Superstring theory and beyond, Reprint of 2003 edition.

[56] Constantin Teleman. The structure of 2D semi-simple field theories. *Invent. Math.*, 188(3):525–588, 2012. doi:10.1007/s00222-011-0352-5.

[57] Jacobus H. van Lint and Gerard van der Geer. *Introduction to coding theory and algebraic geometry*, volume 12 of *DMV Seminar*. Birkhäuser Verlag, Basel, 1988. doi:10.1007/978-3-0348-9286-5.

[58] Adam Zsolt Wagner. Constructions in combinatorics via neural networks. arXiv:2104.14516 [math.CO], 2021.

[59] Yue Wu and Jesús A De Loera. Turning mathematics problems into games: Reinforcement learning and Gröbner bases together solve integer feasibility problems. arXiv:2208.12191 [cs.LG], 2022.

DEPARTMENT OF MATHEMATICS, IMPERIAL COLLEGE LONDON, 180 QUEEN'S GATE, LONDON, SW7 2AZ, UK
Email address: t.coates@imperial.ac.uk

SCHOOL OF MATHEMATICAL SCIENCES, UNIVERSITY OF NOTTINGHAM, NOTTINGHAM, NG7 2RD, UK
Email address: a.m.kasprzyk@nottingham.ac.uk

DEPARTMENT OF MATHEMATICS, IMPERIAL COLLEGE LONDON, 180 QUEEN'S GATE, LONDON, SW7 2AZ, UK
Email address: s.veneziale21@imperial.ac.uk