

Hilbert Series, Machine Learning, and Applications to Physics

Jiakang Bao,^{1,*} Yang-Hui He,^{1,2,3,†} Edward Hirst,^{1,‡} Johannes Hofscheier,^{4,§} Alexander Kasprzyk,^{4,¶} and Suvajit Majumder^{1,**}

¹*Department of Mathematics, City, University of London, EC1V 0HB, UK*

²*Merton College, University of Oxford, OX1 4JD, UK*

³*School of Physics, NanKai University, Tianjin, 300071, P.R. China*

⁴*School of Mathematical Sciences, University of Nottingham, Nottingham, NG7 2RD, UK*

We describe how simple machine learning methods successfully predict geometric properties from Hilbert series (HS). Regressors predict embedding weights in projective space to ~ 1 mean absolute error, whilst classifiers predict dimension and Gorenstein index to $> 90\%$ accuracy with $\sim 0.5\%$ standard error. Binary random forest classifiers managed to distinguish whether the underlying HS describes a complete intersection with high accuracies exceeding 95%. Neural networks (NNs) exhibited success identifying HS from a Gorenstein ring to the same order of accuracy, whilst generation of “fake” HS proved trivial for NNs to distinguish from those associated to the three-dimensional Fano varieties considered.

I. INTRODUCTION AND SUMMARY

The Hilbert series (HS) is an important invariant in the study of modern geometry. In physics, HS have recently become a powerful tool in high energy theory, appearing, for example, in the study of: Bogomol’nyi–Prasad–Sommerfield (BPS) operators of supersymmetric gauge theories [1, 2]; supersymmetric quantum chromodynamics (SQCDs) [3–6] and instanton moduli spaces [7, 8]; invariants of the standard model [9, 10]; polytopes which arise in string compactifications [11]; and explicit constructions of effective Lagrangians [12–17].

In parallel, a programme to use machine learning (ML) techniques to study mathematical structures has recently been proposed [18, 19]. The initial studies were inspired by timely and independent works [18, 20–23]. In these, the effectiveness of ML regressor and classifier techniques in various branches of mathematics and mathematical physics has been investigated. Applications of ML include: finding bundle cohomology on varieties [22, 24, 25]; distinguishing elliptic fibrations [26] and invariants of Calabi–Yau threefolds [27]; the Donaldson algorithm for numerical Calabi–Yau metrics [28]; the algebraic structures of groups and rings [29]; arithmetic geometry and number theory [30–32]; quiver gauge theories and cluster algebras [33]; patterns in particle masses [34]; statistical predictions and model-building in string theory [35–37]; and classifying combinatorial properties of finite graphs [38]. Here we apply ML techniques to the plethystic programme of using Hilbert series to understand structures of quantum field theory.

We examine databases of HS arising in geometry – see the Graded Ring Database (GRDB) [39] – and “fake” HS

randomly generated to imitate the “real” HS. Simple ML methods were able to successfully predict several geometry quantities associated to the HS, and were able to accurately distinguish real from fake HS.

Depending on the form of the HS, simple regression neural networks (NNs) managed to learn embedding weights in projective space to mean absolute error (MAE) ~ 1 ; whilst classification NNs predicted the dimension and the Gorenstein index to accuracy and Matthews correlation coefficient (MCC) > 0.9 .

Motivated by the question if ML can detect if real HS come from a Gorenstein ring, binary classifiers managed to identify whether a fake HS had a palindromic numerator to accuracy and MCC > 0.9 . Additionally, binary classifiers easily were able to distinguish the fake generated data from the dataset of HS associated to three-dimensional Fano varieties.

Moreover, a random forest classifier correctly predicted whether the HS describes a complete intersection (CI): to accuracy 0.9, and MCC 0.8 if the numerator (padded with 0’s) of the HS is used as input; whilst to accuracy 0.95, and MCC over 0.9 if the Taylor series (to order 100) of the HS is used as input.

Further work will investigate ML other more advanced properties of the plethystics. Code scripts for the subsequent investigations, as well as the datasets generated and used, are available from:

<https://github.com/edhirst/HilbertSeriesML.git>.

II. HILBERT SERIES AND PHYSICS

The HS is an important quantity that encodes numerical properties of a projective algebraic variety. It is not a topological invariant in that it depends on the embedding under consideration [40, Example 13.4]. Here we work with varieties defined over \mathbb{C} .

Given a complex projective variety X and ample divisor D there exists a natural embedding in a weighted projective space (w.p.s.) $\mathbb{P}_{\mathbb{C}}(p_0, \dots, p_k)$. We denote its homogeneous coordinate ring by R , i.e., $R = \mathbb{C}[X_0, \dots, X_k]/I$

* jiakang.bao@city.ac.uk

† hey@maths.ox.ac.uk

‡ edward.hirst@city.ac.uk

§ johannes.hofscheier@nottingham.ac.uk

¶ a.m.kasprzyk@nottingham.ac.uk

** suvajit.majumder@city.ac.uk

where I is the homogeneous ideal generated by the polynomials defining X . Notice the variables X_i have weights p_i which might coincide, so that the w.p.s. above can be rewritten as $\mathbb{P}_{\mathbb{C}}(p_0^{q_0}, \dots, p_s^{q_s})$ meaning that weight p_i appears q_i times. The embedding into the w.p.s. induces a grading on $R = \bigoplus_{i \geq 0} R_i$. We refer to [41] for details.

The HS is the generating function for the dimensions of the graded pieces of R :

$$H(t; X) = \sum_{i=0}^{\infty} (\dim_{\mathbb{C}} R_i) t^i$$

where $\dim_{\mathbb{C}} R_i$, the dimension of the i -th graded piece of the ring R , can be thought of as the number of independent degree i polynomials on the variety X . The map $i \mapsto \dim_{\mathbb{C}} R_i$ is called the *Hilbert function*.

The distinction between compact and non-compact can be relaxed by “decompactifying” a projective variety by considering its affine cone for which then the HS is computed by retaining the grading (resp. the weights).

By the Hilbert–Serre Theorem (see for example [42, Theorem 11.1]) there exists $P(t) \in \mathbb{Z}[t]$ such that

$$H(t; X) = \frac{P(t)}{\prod_{\ell=0}^s (1 - t^{p_{\ell}})^{q_{\ell}}} \quad (1)$$

where $p_{\ell} \in \mathbb{Z}_{>0}$ are the weights with corresponding repetitions $q_{\ell} \in \mathbb{Z}_{>0}$. Let j be the smallest positive multiple such that jD is very ample. We call j the *Gorenstein index*, and can rewrite (1) in the form:

$$H(t; X) = \frac{\tilde{P}(t)}{(1 - t^j)^{\dim + 1}} \quad (2)$$

Here \dim is the dimension of X , and $\tilde{P} \in \mathbb{Z}[t]$. If R is a Gorenstein graded ring then the numerator is a palindromic polynomial (by Serre duality). Recall that a polynomial $\sum_{i=1}^d a_i t^i$ is called palindromic if $a_i = a_{d-i}$ [43].

For example, consider the complex line $M = \mathbb{C}$ (regarded as the affine cone over a point) parameterised by a single complex variable x . Then the i -th graded piece R_i is generated by the single monomial x^i . Thus, $\dim_{\mathbb{C}} R_i = 1$ for all $i \in \mathbb{Z}_{\geq 0}$ so that the HS becomes $H(t; \mathbb{C}) = (1 - t)^{-1}$. In general, we have that $H(t; \mathbb{C}^n) = (1 - t)^{-n}$.

The Plethystic Programme. In supersymmetric gauge theories, when the vevs of scalars in different supermultiplets are turned on, the (vacuum) moduli spaces are non-trivial algebraic varieties [44–46] such as hyperkähler cones and (closures of) symplectic leaves. In this case HS are a powerful tool to enumerate gauge invariant operators (GIOs) at different orders.

A particularly useful application of HS to theoretical physics is the plethystic programme which reveals more information of the moduli spaces. We leave a detailed summary of the key formulae to Appendix A.

The multi-graded HS, i.e. the multi-variate series

$$H(t_1, \dots, t_k; X) = \sum_{\vec{i}=0}^{\infty} \dim_{\mathbb{C}}(X_{\vec{i}}) t_1^{i_1} \dots t_k^{i_k}$$

obtained by considering multi-graded rings with pieces $X_{\vec{i}}$ for $\vec{i} = (i_1, \dots, i_k)$, could fully determine how the GIOs transform under symmetry groups of gauge theories.

Duality and Moduli Spaces. HS have been well-studied in the context of quiver gauge theories. For Higgs branches in low dimensions, HS obtained from the Molien–Weyl integral enables us to systematically study the geometry of SQCDs [3]. Such methods can also be used to study the instanton moduli spaces [7, 8, 47]. As the spaces of dressed monopole operators, i.e., the Coulomb branches, receive quantum corrections, monopole formula [48] and Hall–Littlewood formula [49] are used to obtain the HS. This not only unveils the geometry of moduli spaces, but also provides tools and evidences to study three-dimensional mirror symmetry and duality including theories in higher dimensions.

Standard Model. Phenomenologically, HS have been applied to lepton and quark flavour invariants for the Standard Model in [9] as well as to the minimal supersymmetric Standard Model in [50, 51].

III. MACHINE LEARNING

In this section we describe our approaches to ML properties of the rational representations (1) and (2) by feeding in coefficients of the corresponding HS. **Keras** with the **TensorFlow** backend [52] was used for the investigations.

In §III A, “real” HS associated to certain three-dimensional Fano varieties are introduced and analysed. In §III B, “fake” HS, i.e. rational functions of the form (1) and (2), were randomly generated and properties of them were machine-learned. In §III C and §III D, binary classifiers were used to determine whether fake HS of the form (2) had palindromic numerator, and to determine fake HS from real HS, all with great success. Finally, in §III E we use ML to determine if a HS is associated to a complete intersection.

A. Acquiring HS

Example HS associated to algebraic varieties were retrieved from the GRDB [39]. We use a database of candidate HS conjecturally associated to three-dimensional \mathbb{Q} -Fano varieties with Fano index one, as constructed by Altınok–Brown–Reid [53]. Such varieties come with a natural choice of ample divisor $D = -K$, the anti-canonical divisor. We call these HS “real”. See Appendix B for the distributions of the parameters $d, \{a_i\}, s, \{p_{\ell}\}, \{q_{\ell}\}$ for this set of data. Here we are using notation as in (1), and write $P(t) = 1 + \sum_{i=1}^d a_i t^i$ for the numerator polynomial.

Example 1 Consider the three-dimensional \mathbb{Q} -Fano variety $X \subset \mathbb{P}(1^3, 2^2, 3^2)$ (number 11122 in the GRDB). This is of codimension 3, with $\mathcal{B} = \{\frac{1}{2}(1, 1, 1), 2 \times \frac{1}{3}(1, 1, 2)\}$ isolated orbifold points, and hence has Gorenstein index $j = 6$. Writing the HS in the form (1) gives:

$$H(t; X) = \frac{P(t)}{(1-t)^3(1-t^2)^2(1-t^3)^2}$$

where $P(t) = 1 - 2t^4 - 2t^5 + 2t^7 + 2t^8 - t^{12}$.

Rewriting this in the form (2) gives:

$$H(t; X) = \frac{\tilde{P}(t)}{(1-t^6)^4}$$

where $\tilde{P}(t) = 1 + 3t + 8t^2 + \dots + 8t^{21} + 3t^{22} + t^{23}$.

For the HS of this dataset, there are two competing phenomena that contribute to its coefficients: the initial part P_{ini} that coincides with the HS in small degrees and the ‘‘correction terms’’ $P_{\text{orb}}(Q)$ for each isolated orbifold point $Q = \frac{1}{r}(b_1, \dots, b_{\text{dim}})$ of X . More precisely, we have [54]

$$H(t; X) = P_{\text{ini}} + \sum_{Q \in \mathcal{B}} P_{\text{orb}}(Q)$$

where the sum is taken over the set \mathcal{B} of isolated orbifold points of X . P_{ini} and $P_{\text{orb}}(Q)$ ($Q = \frac{1}{r}(b_1, \dots, b_{\text{dim}})$) satisfy

$$P_{\text{ini}} = \frac{A(t)}{(1-t)^{\text{dim}+1}}, \quad P_{\text{orb}}(Q) = \frac{B_Q(t)}{(1-t)^{\text{dim}}(1-t^r)}$$

where $A(t), B_Q(t)$ are integral palindromic polynomials with degrees related via $\deg B_Q(t) - \deg A(t) = r - 1$. The initial plurigenera of H coincide with P_{ini} in degrees $\leq \lfloor \deg A(t)/2 \rfloor$, whilst in higher degrees the orbifold points start to contribute to the coefficients of H . Recall the coefficients of the HS are called plurigenera. Because of this phenomenon, extra care must be taken when computing parameters for the representations (1) and (2) from a finite set of coefficients of the HS. Our investigations indicate that ML can deal with this behaviour.

B. Generating and ML Fake HS

The ‘‘fake’’ HS generated take the forms (1) and (2), with numerators of the form $1 + \sum_{i=1}^d a_i t^i$. The numerator $\tilde{P}(t)$ of (2) is required to be palindromic (and, as a consequence, $a_d = 1$). Coefficient sets consisting of the parameters $d, \{a_i\}, s, \{p_\ell\}, \{q_\ell\}$, where $1 \leq i \leq d$ and $1 \leq \ell \leq s$, were randomly generated and the Taylor expansions of the resulting fake HS were computed to order ~ 1000 . If the parameters did not satisfy $\sum_\ell p_\ell q_\ell > d$, or if there were negative coefficients in the resulting Taylor expansion, then the parameters were discarded.

The resulting data were fed into a NN to learn the desired properties of the HS. The input was a vector of

Taylor expansion coefficients: either a vector of coefficients for low-order terms 0 to 100; or for high-orders terms 1000 to 1009. Although coefficients of low-order terms are easier to calculate, predictions based on those inputs are more error-prone as contributions from orbifold points take effect only for high-order terms (see §III A).

Fewer coefficients were required when learning from coefficients deeper in the Taylor expansion; geometric data are more readily extracted from larger plurigenera. We found the following analogy from toric geometry insightful. When counting the number of lattice points $c_m = |m\Delta \cap \mathbb{Z}^{\text{dim}}|$ in the m -th dilation of a polytope Δ then, for $m \gg 0$, $c_m \sim \text{Vol}(m\Delta) = m^{\text{dim}} \text{Vol}(\Delta)$. (This is a toric rephrasing of the HS, with Δ the polytope associated with an ample divisor D and $c_m = h^0(mD)$.)

The first investigation used supervised regressor NNs to learn $\{p_\ell\}$ for fake HS in the form (1). Supervised classifier NNs were trained to predict the Gorenstein index j and the dimension dim of fake HS in the form (2). Classifiers were used since the NN outputs were single numbers and hence associated well to classifier data structures.

HS Regressor Investigations. For this investigation $\sim 10\,000$ fake HS of the form (1) were uniformly drawn from a sample space given by $d = 3, s = 3, |a_i| \leq 10, p_\ell \leq 10$. This space was chosen to provide a sufficiently large range of HS whilst ensuring that its size was still feasible for ML training. The goal was to predict the values $\{p_\ell\}$ and $\{q_\ell\}$ of the form (1) from a given (finite) range of HS coefficients. This information was encoded into single vector where each p_ℓ was repeated q_ℓ times, and the entries were given in increasing order.

A 5-fold cross-validation (in the sense of [55]) was performed for a feed-forward regressor NN with 4 hidden dense layers of 1024 neurons each, using LeakyReLU activation (with $\alpha = 0.01$), in batches of 32 for 20 epochs over the full dataset. The NN had a final dense layer with as many neurons as p_ℓ 's (counting multiplicities). Dropout layers between the dense layers reduced the risk of overfitting (dropout factor 0.05). The NN was trained with the Adam optimiser [56] using a $\log(\cosh)$ loss function and the training performance was measured via MAE.

Table I summarises the averaged MAE, with standard error, over the 5-fold cross validation for two ranges of HS coefficients: the first 101 coefficients; and the coefficients of order 1000 to 1009. In both cases the MAE is below 2, i.e. the true denominator of the form (1) of the underlying HS could be extracted with reasonably good accuracy from the HS coefficients alone.

HS Classification Investigations. In this investigation a 5-fold cross-validation for a feed-forward classifier NN with the same layer structure as before was trained. We again used an Adam optimiser, but now with *sparse categorical cross entropy* loss to reflect the classification question. Training performance was measured with accuracy and MCC. The final dense layer now had as many neurons as classes in the investigation (5 in both cases),

Orders of Input	MAE
0 to 100	1.94 ± 0.11
1000 to 1009	1.04 ± 0.12

TABLE I. Averaged MAE, with standard error, of the 5-fold cross-validation of the NN learning the weights p_ℓ (with multiplicity) of the form (1) of the HS from input vectors of HS coefficients to the specified orders.

Parameter Learnt	Orders of Input	Performance Measures	
		Accuracy	MCC
j	0 to 100	0.934 ± 0.008	0.916 ± 0.010
	1000 to 1009	0.780 ± 0.018	0.727 ± 0.022
dim	0 to 100	0.995 ± 0.005	0.993 ± 0.006
	1000 to 1009	0.865 ± 0.024	0.822 ± 0.031

TABLE II. Averaged accuracy and MCC, with standard error, of the 5-fold cross-validation of the NN learning the Gorenstein index j , the dimension dim, and the form (2) with HS coefficients in the specified ranges of degrees as input.

with softmax activation, and neurons representing the values the learnt parameters could take.

This time $\sim 10\,000$ HS of the form (2) were uniformly drawn from a sample space given by $d = 5$, $|a_i| \leq 50$, $\dim \leq 5$. The goal this time was to train an NN to predict the Gorenstein index j , the dimension dim, and the form (2) from the HS coefficients in the same orders of degrees.

Note if coefficients in larger degrees were used as input, the larger values caused problems with the loss function. This issue was mitigated by log-normalising the HS coefficients, i.e., by taking the natural logarithm input values were scaled down to ranges the loss function and optimiser could handle. However some fake HS contained 0 coefficients and were therefore omitted, hence resulting in a full dataset of 8711 HS for the training with HS coefficients of larger degree. Note also that log-normalisation was only used in this case and in no other investigations.

Table II summarises the averaged accuracies and MCCs, with standard error, over the 5-fold cross-validation of the NN. These results show almost perfect classification of both the Gorenstein index, j , and the dimension, dim, from HS coefficients in low degrees. Interestingly the performance is worse when using terms deeper in the HS, presumably due to the required log-normalisation of the coefficients removing the finer structure of the coefficients required to determine the exact parameter value being learnt.

C. Identifying the Gorenstein Property

In this section we investigate the effectiveness of binary classifiers to detect if the numerator of form (2) of an HS is palindromic. Recall from Section II that the numerator is palindromic if the ring R is Gorenstein (by Serre dual-

Orders of Input	Performance Measures	
	Accuracy	MCC
0 to 100	0.844 ± 0.087	0.717 ± 0.155
1000 to 1009	0.954 ± 0.043	0.919 ± 0.073

TABLE III. Averaged accuracy and MCC, with standard error, of the 5-fold cross-validation of a NN learning whether the HS has palindromic numerator in form (2) from HS coefficients to the specified orders as input.

ity). Then the numerator of form (1) is palindromic too (possibly up to a sign); see Example 1 for an illustration. The goal was to use a NN to distinguish whether a HS is coming from a Gorenstein ring, i.e., the numerator polynomial of form (2) is palindromic. As before the NN's input were HS coefficients from the same ranges of degrees.

For the investigation two equally sized sets of fake HS, one with and the other without palindromic numerators, were uniformly drawn from a sample space given by $d = 9$, $|a_i| \leq 50$, $j = 5$, $\dim + 1 = 6$. The same reasons as before apply for this choice of space. The HS in each of the two sets were then labelled and together comprised the full dataset for a 5-fold cross-validation to be performed using a feed-forward classifier NN with the same layer structure as in the previous investigation. Also the same Adam optimiser was used for training, but now with *binary cross-entropy* loss to reflect the classification question. Training performance was measured with accuracy and MCC. The final dense layer of the NN now had 2 neurons corresponding to whether the HS comes from a Gorenstein ring or not.

Table III summarises the averaged accuracies and MCCs, with standard error, over the 5-fold cross-validation of the NN. The results show good success in detecting if a HS comes from a Gorenstein ring using HS coefficients alone. The classifier performed better on coefficients in larger degrees indicating that the palindromicity property is more readily evident from plurigenera deeper in the HS (possibly because of the bigger variation).

D. Differentiating Real and Fake HS

The investigation examined the success of a binary classifier in distinguishing whether a HS, represented by a finite set of HS coefficients, corresponds to a real HS from the GRDB, or a randomly generated fake HS. The dataset consisted of HS candidates conjecturally associated to 3-dimensional Fano polytopes from the GRDB, amounting to $\sim 29\,000$ HS, along with as many fake HS with the same structure which were randomly generated.

A 5-fold cross-validation for a feed-forward classifier NN with the same layer structure as in the previous investigations was performed. For training an Adam optimiser with a binary cross-entropy loss with the same parameters as before was used. Training performance was measured with accuracy and MCC. The final dense layer had 2

neurons corresponding to whether the inputted HS coefficients were associated to a real or fake HS.

The $\sim 29\,000$ fake HS were generated randomly using form (1) parameters drawn from probability distributions reflecting the real HS data as given in Appendix B. An equal number of real HS were taken from the GRDB to produce the full dataset, and as before HS coefficients to the same order of degrees were used as NN inputs.

In this investigation the averaged accuracies and MCCs exceeded 0.99 for both ranges of degrees of HS coefficients. Further analysis of the data showed that coefficients of fake HS were orders of magnitudes different to the real case which possibly made this classification far easier. This also highlights the uniqueness of real HS which come with a wealth of further impactful structure, e.g., on the parameters of the corresponding forms (1) and (2).

E. Detecting Complete Intersection

An important application of the plethystic logarithm (see Appendix A for details and references) is that it detects whether the underlying variety is a *complete intersection* (CI), i.e., the defining ideal (the ideal of polynomials vanishing on the variety) is generated by exactly codimension many polynomials. Such optimal intersection has been widely used in the physics literature, e.g., in string model-building [57, 58]. As can be seen from the definition, the PE^{-1} involves the number-theoretic μ -function, making the computation non-trivial. A natural question arises as to whether a trained classifier can identify whether X is CI, i.e., when PE^{-1} terminates as a Taylor series, by only “looking” at the the shape of the HS.

Suppose $X = \{f_1 = 0, \dots, f_c = 0\}$ defines a complete intersection in $\mathbb{P}_{\mathbb{C}}^k$ where each f_i is a homogeneous polynomial of degree m_i in a standard graded polynomial ring $R_{k+1} = \mathbb{C}[X_0, \dots, X_k]$, such that each variable X_i has degree 1. Then the HS of X takes the form

$$\frac{(1-t^{m_1}) \dots (1-t^{m_c})}{(1-t)^n} = \frac{1 + a_1 t + \dots + a_d t^d}{(1-t)^n}. \quad (3)$$

This follows by induction on the f_i using the additivity of HS and the the exact sequences

$$0 \rightarrow R_k^{[m_i]} \xrightarrow{f_i} R_k \rightarrow R_{k+1} \rightarrow 0$$

where $R_k^{[m_i]}$ denotes a standard graded polynomial ring with degrees shifted by m_i so that the first map becomes a morphism of graded rings. Notice X is a projective variety of codimension c in $\mathbb{P}_{\mathbb{C}}^k$, i.e. has dimension $\dim = k - c$.

This time ~ 5500 HS of the form (3) representing CIs were uniformly drawn from a sample space given by $c, m_i = 1, \dots, 10$ and $1 \leq n - \sum_i m_i \leq 5$. The HS representing non-CI were generated by drawing CI HS f from the sample space above and then adding a binomial preventing the result to factor as in (3). More precisely,

ML algorithm	Orders of Input	Performance Measures	
		Accuracy	MCC
NN	0 to 20	0.502 ± 0.001	0
	0 to 100	0.535 ± 0.007	0.069 ± 0.013
RF	0 to 20	0.720 ± 0.041	0.496 ± 0.013
	0 to 100	0.919 ± 0.067	0.910 ± 0.010

TABLE IV. Averaged accuracy and MCC, with standard error, of the 5-fold cross-validation of the random forest learning complete intersections the form (2) with HS coefficients in the specified ranges of degrees as input.

the non-CI HS was computed by

$$f + \frac{t^{k_0} + (-1)^c \cdot t^{d-k_0}}{(1-t)^n}$$

where k_0 denotes the smallest positive index of the non-vanishing coefficients a_i of the numerator polynomial of f . This procedure ensured that learning is non-trivial, because the resulting non-CI HS have a similar shape to the form (3), but do not correspond to HS of CI. The full dataset was comprised by ~ 5500 CI HS and ~ 5500 non-CI HS, resulting in a total of ~ 11000 samples.

We use the Taylor expansion of the HS (not the plethystic log of it) as input to see if the machine could identify complete intersections. We create a balanced dataset of size $\sim 11\,000$ for binary classification. We use a random forest classifier. If we truncate the Taylor series at some small order (e.g. order 20). The accuracy would be ~ 0.75 with MCC ~ 0.5 . However, if we include sufficiently high orders, the results would be greatly improved. If we use the Taylor expansions to order 100, at 80% training, the random forest model could give over 0.95 accuracy and over 0.9 MCC. A 5-fold cross validation would give $0.919(\pm 0.067)$ accuracy (with 95% confidence interval). Incidentally, one may as well try a neural network, but it could only give less than 0.6 accuracy which is almost random guessing. This is because the coefficients in the Taylor expansions are too big for the neural network, and the optimizers are not finding the minima of loss functions very well for our synthetic question. To resolve these problems, we apply principal component analysis (PCA), and we find that NNs could then also reproduce ~ 0.95 accuracy. Indeed, PCA shows a clear separation of CIs and non-CIs as in Figure 1. Therefore, it seems to indicate that the machine learning models could capture some features of Hilbert functions for complete intersections which may not be known before. Further study on such features is desired.

Acknowledgements. JB is supported by a CSC scholarship. YHH would like to thank STFC for grant ST/J00037X/1. EH would like to thank STFC for a PhD studentship. JH is supported by a Nottingham Research Fellowship. AK is supported by EPSRC Fellowship EP/N022513/1. SM is funded by a SMCSE Doctoral Studentship. This collaboration was made

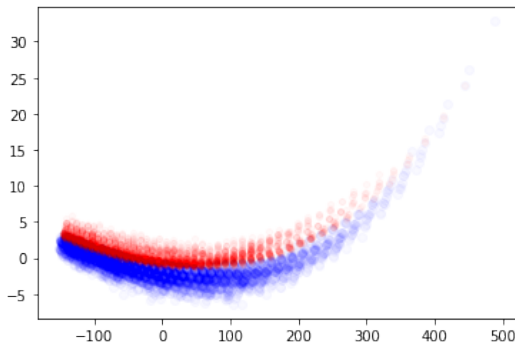


FIG. 1. The PCA for complete and non-complete intersections (in blue and red respectively).

possible by a Focused Research Workshop grant from the Heilbronn Institute for Mathematical Research.

-
- [1] S. Benvenuti, B. Feng, A. Hanany, and Y.-H. He, “Counting BPS operators in gauge theories: quivers, syzygies and plethystics,” *J. High Energy Phys.* no. 11, (2007) 050, 48.
- [2] B. Feng, A. Hanany, and Y.-H. He, “Counting gauge invariants: the plethystic program,” *J. High Energy Phys.* no. 3, (2007) 090, 42.
- [3] J. Gray, Y.-H. He, A. Hanany, N. Mekareeya, and V. Jejjala, “SQCD: a geometric aperçu,” *Journal of High Energy Physics* **2008** no. 05, (May, 2008) 099–099.
- [4] A. Hanany, N. Mekareeya, and G. Torri, “The Hilbert series of adjoint SQCD,” *Nuclear Phys. B* **825** no. 1-2, (2010) 52–97.
- [5] Y. Chen and N. Mekareeya, “The Hilbert series of U/SU SQCD and Toeplitz determinants,” *Nuclear Phys. B* **850** no. 3, (2011) 553–593.
- [6] N. Jokela, M. Järvinen, and E. Keski-Vakkuri, “New results for the SQCD Hilbert series,” *J. High Energy Phys.* no. 3, (2012) 048, front matter+30.
- [7] S. Benvenuti, A. Hanany, and N. Mekareeya, “The Hilbert series of the one instanton moduli space,” *J. High Energy Phys.* no. 6, (2010) 100, 40.
- [8] A. Hanany, N. Mekareeya, and S. S. Razamat, “Hilbert series for moduli spaces of two instantons,” *J. High Energy Phys.* no. 1, (2013) 070, front matter + 48.
- [9] A. Hanany, E. E. Jenkins, A. V. Manohar, and G. Torri, “Hilbert series for flavor invariants of the Standard Model,” *J. High Energy Phys.* no. 3, (2011) 096, 7.
- [10] L. Lehman and A. Martin, “Low-derivative operators of the Standard Model effective field theory via Hilbert series methods,” *Journal of High Energy Physics* **02** no. 81, (2016) .
- [11] V. Braun, “Counting points and Hilbert series in string theory,” in *Strings, gauge fields, and the geometry behind*, pp. 225–235. World Sci. Publ., Hackensack, NJ, 2013.
- [12] L. Lehman and A. Martin, “Hilbert series for constructing lagrangians: Expanding the phenomenologist’s toolbox,” *Phys. Rev. D* **91** (May, 2015) 105014.
- [13] B. Henning, X. Lu, T. Melia, and H. Murayama, “Hilbert series and operator bases with derivatives in effective field theories,” *Comm. Math. Phys.* **347** no. 2, (2016) 363–388.
- [14] A. Kobach and S. Pal, “Hilbert series and operator basis for NRQED and NRQCD/HQET,” *Phys. Lett. B* **772** (2017) 225–231.
- [15] Anisha, S. Das Bakshi, J. Chakraborty, and S. Prakash, “Hilbert series and plethystics: paving the path towards 2HDM- and MLRSM-EFT,” *J. High Energy Phys.* no. 9, (2019) 035, 113.
- [16] C. B. Marinissen, R. Rahn, and W. J. Waalewijn, “. . . , 83106786, 114382724, 1509048322, 2343463290, 27410087742, . . . efficient Hilbert series for effective theories,” *Phys. Lett. B* **808** (2020) 135632, 7.
- [17] L. Graf, B. Henning, X. Lu, T. Melia, and H. Murayama, “2, 12, 117, 1959, 45171, 1170086, . . . : A Hilbert series for the QCD chiral Lagrangian.” [arXiv:2009.01239](https://arxiv.org/abs/2009.01239) [[hep-ph](https://arxiv.org/abs/2009.01239)], 2020.
- [18] Y.-H. He, “Deep-learning the landscape.” [arXiv:1706.02714](https://arxiv.org/abs/1706.02714) [[hep-th](https://arxiv.org/abs/1706.02714)], 2017.
- [19] Y.-H. He, “The Calabi–Yau landscape: from geometry, to physics, to machine-learning.” [arXiv:1812.02893](https://arxiv.org/abs/1812.02893) [[hep-th](https://arxiv.org/abs/1812.02893)], 2018.
- [20] Y.-H. He, “Machine-learning the string landscape,” *Phys. Lett. B* **774** (2017) 564–568.
- [21] D. Krefl and R.-K. Seong, “Machine learning of Calabi–Yau volumes,” *Phys. Rev. D* **96** no. 6, (2017) 066014, 8.
- [22] F. Ruehle, “Evolving neural networks with genetic algorithms to study the string landscape,” *J. High Energy Phys.* no. 8, (2017) 038, front matter+19.
- [23] J. Carifio, J. Halverson, D. Krioukov, and B. D. Nelson, “Machine learning in the string landscape,” *J. High Energy Phys.* no. 9, (2017) 157, front matter+35.
- [24] C. R. Brodie, A. Constantin, R. Deen, and A. Lukas, “Machine learning line bundle cohomology,” *Fortschr. Phys.* **68** no. 1, (2020) 1900087, 14.
- [25] M. Larfors and R. Schneider, “Explore and exploit with heterotic line bundle models,” *Fortschr. Phys.* **68** no. 5, (2020) 2000034, 10.
- [26] Y.-H. He and S.-J. Lee, “Distinguishing elliptic fibrations with AI,” *Phys. Lett. B* **798** (2019) 134889, 5.
- [27] K. Bull, Y.-H. He, V. Jejjala, and C. Mishra, “Machine learning CICY threefolds,” *Phys. Lett. B* **785** (2018) 65–72.
- [28] A. Ashmore, Y.-H. He, and B. A. Ovrut, “Machine

- learning Calabi-Yau metrics,” *Fortschr. Phys.* **68** no. 9, (2020) 2000068, 23.
- [29] Y.-H. He and M. Kim, “Learning algebraic structures: Preliminary investigations.” [arXiv:1905.02263](https://arxiv.org/abs/1905.02263) [[cs.LG](https://arxiv.org/abs/1905.02263)], 2019.
- [30] L. Alessandretti, A. Baronchelli, and Y.-H. He, “Machine learning meets Number Theory: The Data Science of Birch–Swinnerton-Dyer.” [arXiv:1911.02008](https://arxiv.org/abs/1911.02008) [[math.NT](https://arxiv.org/abs/1911.02008)], 2019.
- [31] Y.-H. He, E. Hirst, and T. Peterken, “Machine-learning Dessins d’Enfants: Explorations via modular and Seiberg–Witten curves.” [arXiv:2004.05218](https://arxiv.org/abs/2004.05218) [[hep-th](https://arxiv.org/abs/2004.05218)], 2020.
- [32] Y.-H. He, K.-H. Lee, and T. Oliver, “Machine-learning the Sato–Tate conjecture.” [arXiv:2010.01213](https://arxiv.org/abs/2010.01213) [[math.NT](https://arxiv.org/abs/2010.01213)], 2020.
- [33] J. Bao, S. Franco, Y.-H. He, E. Hirst, G. Musiker, and Y. Xiao, “Quiver mutations, Seiberg duality and machine learning.” [arXiv:2006.10783](https://arxiv.org/abs/2006.10783) [[hep-th](https://arxiv.org/abs/2006.10783)], 2020.
- [34] Y. Gal, V. Jejjala, D. K. M. Pena, and C. Mishra, “Baryons from Mesons: A machine learning perspective.” [arXiv:2003.10445](https://arxiv.org/abs/2003.10445) [[hep-ph](https://arxiv.org/abs/2003.10445)], 2020.
- [35] R. Deen, Y.-H. He, S.-J. Lee, and A. Lukas, “Machine learning string standard models.” [arXiv:2003.13339](https://arxiv.org/abs/2003.13339) [[hep-th](https://arxiv.org/abs/2003.13339)], 2020.
- [36] J. Halverson, B. Nelson, and F. Ruehle, “Branes with brains: exploring string vacua with deep reinforcement learning,” *J. High Energy Phys.* no. 6, (2019) 003, 59.
- [37] J. Halverson and C. Long, “Statistical predictions in string theory and deep generative models,” *Fortschr. Phys.* **68** no. 5, (2020) 2000005, 13.
- [38] Y.-H. He and S.-T. Yau, “Graph Laplacians, Riemannian manifolds and their machine-learning.” [arXiv:2006.16619](https://arxiv.org/abs/2006.16619) [[math.CO](https://arxiv.org/abs/2006.16619)], 2020.
- [39] G. Brown and A. M. Kasprzyk, “The Graded Ring Database.” Online. <http://www.grdb.co.uk/>.
- [40] J. Harris, *Algebraic geometry*, vol. 133 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995. A first course, Corrected reprint of the 1992 original.
- [41] I. Dolgachev, “Weighted projective varieties,” in *Group actions and vector fields (Vancouver, B.C., 1981)*, vol. 956 of *Lecture Notes in Math.*, pp. 34–71. Springer, Berlin, 1982.
- [42] M. F. Atiyah and I. G. Macdonald, *Introduction to commutative algebra*. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1969.
- [43] R. P. Stanley, “Hilbert functions of graded algebras,” *Advances in Mathematics* **28** (1978) 57–83.
- [44] F. Buccella, J. P. Derendinger, S. Ferrara, and C. A. Savoy, “Patterns of symmetry breaking in supersymmetric gauge theories,” *Phys. Lett. B* **115** no. 5, (1982) 375–379.
- [45] M. A. Luty and W. Taylor, IV, “Varieties of vacua in classical supersymmetric gauge theories,” *Phys. Rev. D* (3) **53** no. 6, (1996) 3399–3405.
- [46] D. Mehta, Y.-H. He, and J. D. Hauenstein, “Numerical algebraic geometry: a new perspective on gauge and string theories,” *J. High Energy Phys.* no. 7, (2012) 018, front matter+31.
- [47] A. Hanany, N. Mekareeya, and S. S. Razamat, “Hilbert series for moduli spaces of two instantons,” *J. High Energy Phys.* no. 1, (2013) 070, front matter + 48.
- [48] S. Cremonesi, A. Hanany, and A. Zaffaroni, “Monopole operators and Hilbert series of Coulomb branches of $3d\mathcal{N}=4$ gauge theories,” *Journal of High Energy Physics* **5** (2014) .
- [49] S. Cremonesi, A. Hanany, N. Mekareeya, and A. Zaffaroni, “Coulomb branch Hilbert series and Hall–Littlewood polynomials,” *Journal of High Energy Physics* **178** (2014) .
- [50] Y.-H. He, V. Jejjala, C. Matti, and B. D. Nelson, “Veronese geometry and the electroweak vacuum moduli space,” *Phys. Lett. B* **736** (2014) 20–25.
- [51] Y. Xiao, Y.-H. He, and C. Matti, “Standard model plethystics,” *Phys. Rev. D* **100** (Oct, 2019) 076001.
- [52] e. a. Abadi, M., “TensorFlow: Large-scale machine learning on heterogeneous systems.” Online, 2015. <http://tensorflow.org/>.
- [53] S. Altınok, G. Brown, and M. Reid, “Fano 3-folds, $K3$ surfaces and graded rings,” in *Topology and geometry: commemorating SISTAG*, vol. 314 of *Contemp. Math.*, pp. 25–53. Amer. Math. Soc., Providence, RI, 2002.
- [54] A. Buckley, M. Reid, and S. Zhou, “Ice cream and orbifold Riemann-Roch,” *Izv. Ross. Akad. Nauk Ser. Mat.* **77** no. 3, (2013) 29–54.
- [55] T. Hastie, R. Tibshirani, and J. Friedman, *The elements of statistical learning*. Springer Series in Statistics. Springer, New York, second ed., 2009. Data mining, inference, and prediction.
- [56] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization.” [arXiv:1412.6980](https://arxiv.org/abs/1412.6980) [[cs.LG](https://arxiv.org/abs/1412.6980)], 2014.
- [57] P. Candelas, A. M. Dale, C. A. Lütken, and R. Schimmrigk, “Complete intersection Calabi-Yau manifolds,” *Nuclear Phys. B* **298** no. 3, (1988) 493–525.
- [58] L. B. Anderson, Y.-H. He, and A. Lukas, “Heterotic compactification, an algorithmic approach,” *J. High Energy Phys.* no. 7, (2007) 049, 34.

Appendix A: The Plethystic Programme

For a function $f(t) = \sum_{n=0}^{\infty} a_n t^n$, we can define the *plethystic exponential* (sometimes known as the Euler transform) as

$$\text{PE}[f(t)] := \exp\left(\sum_{n=1}^{\infty} \frac{f(t^n) - f(0)}{n}\right) = \prod_{n=1}^{\infty} (1 - t^n)^{-a_n}. \quad (\text{A1})$$

For instance, the mesonic BPS operators fall into two categories: single- and multi-trace. Then the HS is the generating function for counting the basic single-trace invariants. Moreover, the HS of the N -th symmetric product is given by $g_N(t; M) = f(t; \text{sym}^N(X))$, $\text{sym}^N(X) := M^N/S_N$, where the “grand-canonical” partition function is given by the fugacity-inserted plethystic exponential of the Hilbert series: $\text{PE}\nu[f(t)] := \prod_{n=0}^{\infty} (1 - \nu t^n)^{-a_n} = \sum_{N=0}^{\infty} g_N(t)\nu^N$. In gauge theory, this is considered to be at finite N and the expansion $g_N(t) = \sum_{n=0}^{\infty} b_n t^n$ gives the number b_n of operators of charge n .

There is also an analytic inverse function to PE, which

is the *plethystic logarithm*, given by

$$\mathrm{PE}^{-1}[g(t)] = \sum_{k=1}^{\infty} \frac{\mu(k)}{k} \log(g(t^k)) , \quad (\mathrm{A}2)$$

where $\mu(k)$ is the Möbius function. The first positive terms in the Taylor expansion of PE^{-1} encodes generators at different degrees, and the first negative terms give the relations among them. Higher order terms are known as the syzygies. In particular, if X is a complete intersection, then $\mathrm{PE}^{-1}[H(t)]$ is a polynomial of t (i.e. terminates at a finite order).

Appendix B: Real HS parameter distributions

The dataset of real HS associated to 3-dimensional Fano varieties considered in this paper [39] that was analysed to

produce distributions of the HS function form parameters $d, \{a_i\}, s, \{p_\ell\}, \{q_\ell\}$ as shown in Figures 2-6. These distributions, and their respective fittings were used to make fake HS generation more representative of the real HS data.

Fittings used sums of Gaussian distributions, reflecting a Central Limit Theorem motivation in analysis of this large dataset of $\sim 54\,000$ HS. In all cases the sum of 2 independent Gaussian distributions sufficed in making a visually accurate fit. Thus, using these distribution in fake HS generation would ideally produce HS of the same form. Interestingly, the fake HS still had quite different coefficient growth rates to the real HS, stabilising deeper in the series. This phenomena is further discussed in III D.

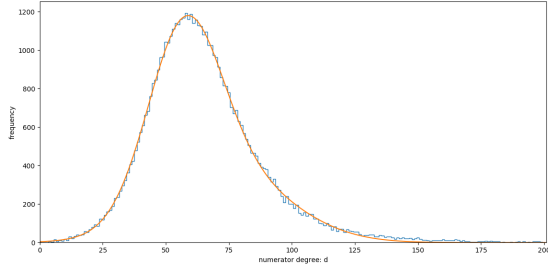


FIG. 2. Histogram of distribution of real HS numerator degrees d , with Gaussian fitting.

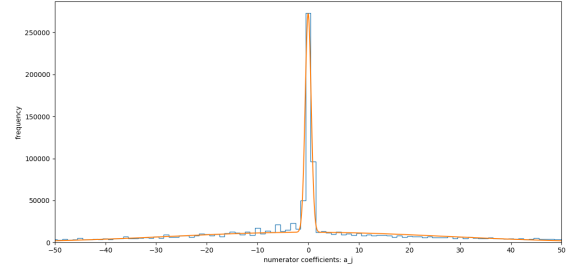


FIG. 3. Histogram of distribution of real HS numerator coefficient values a_i , with Gaussian fitting.

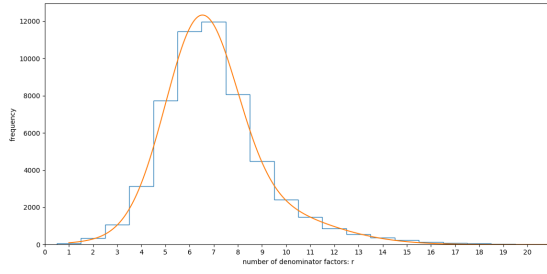


FIG. 4. Histogram of distribution of real HS number of denominator factors s , with Gaussian fitting.

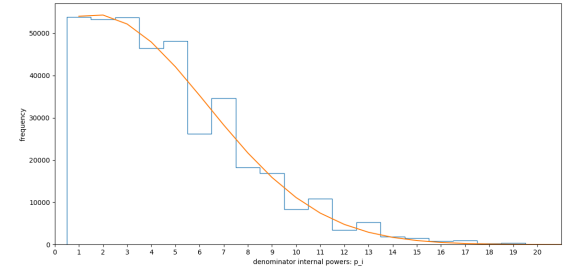


FIG. 5. Histogram of distribution of real HS denominator internal powers (i.e. denominator weights) p_l , with Gaussian fitting.

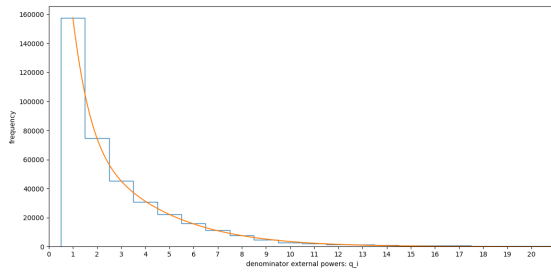


FIG. 6. Histogram of distribution of real HS denominator external powers (i.e. number of repetitions of each denominator weight) q_l , with Gaussian fitting.