

https://dspace.org/

Generating triangulations and fibrations with reinforcement learning

Per Berglund; Giorgi Butbaia; Yang-Hui He; Ellie Heyes; Edward Hirst ...

ELSEVIER SCIENCE BV
https://hdl.handle.net/10539/46982
WOS

Downloaded from a DSpace repository. https://dspace.org/

Phys. Lett. B 860 (2025) 139158

Contents lists available at ScienceDirect

Physics Letters B

journal homepage: www.elsevier.com/locate/physletb

Letter

Generating triangulations and fibrations with reinforcement learning

Per Berglund [a], Giorgi Butbaia [a], Yang-Hui He [b,c,d,e], Elli Heyes [b,c,*], Edward Hirst [f], Vishnu Jejjala [g]

[a] Department of Physics and Astronomy, University of New Hampshire, Durham, NH 03824, USA
[b] London Institute for Mathematical Sciences, Royal Institution, London, W1S 4BS, UK
[c] Department of Mathematics, City, University of London, EC1V 0HB, UK
[d] Merton College, University of Oxford, OX1 4JD, UK
[e] School of Physics, NanKai University, Tianjin, 300071, PR China
[f] Centre for Theoretical Physics, Queen Mary University of London, E1 4NS, UK
[g] Mandelstam Institute for Theoretical Physics, School of Physics, NITheCS, and CoE-MaSS, University of the Witwatersrand, Johannesburg, WITS 2050, South Africa

ARTICLE INFO

Editor: N. Lambert

ABSTRACT

We apply reinforcement learning (RL) to generate fine regular star triangulations of reflexive polytopes, that give rise to smooth Calabi-Yau (CY) hypersurfaces. We demonstrate that, by simple modifications to the data encoding and reward function, one can search for CYs that satisfy a set of desirable string compactification conditions. For instance, we show that our RL algorithm can generate triangulations together with holomorphic vector bundles that satisfy anomaly cancellation and poly-stability conditions in heterotic compactification. Furthermore, we show that our algorithm can be used to search for reflexive subpolytopes together with compatible triangulations that define fibration structures of the CYs.

Contents

1. Introduction . . . 2
2. The model . . . 2
3. Results . . . 3
3.1. Low $h^{1,1}$ . . . 4
3.2. Bundles . . . 4
3.3. Fibrations . . . 5
3.3.1. Subpolytopes . . . 5
3.3.2. Compatible triangulations . . . 6
4. Outlook . . . 7
Declaration of competing interest . . . 7
Acknowledgements . . . 7
Appendix A. Toric geometry . . . 7
A.1. Reflexive polytopes . . . 7
A.2. Fine regular star triangulations . . . 8
A.3. Fibration structures . . . 8
Appendix B. Reinforcement learning . . . 9
B.1. Q-learning . . . 9
Data availability . . . 10
References . . . 10

* Corresponding author.
E-mail addresses: per.berglund@unh.edu (P. Berglund), giorgi.butbaia@unh.edu (G. Butbaia), hey@maths.ox.ac.uk (Y.-H. He), elli.heyes@city.ac.uk (E. Heyes), e.hirst@qmul.ac.uk (E. Hirst), v.jejjala@wits.ac.za (V. Jejjala).
https://doi.org/10.1016/j.physletb.2024.139158
Received 4 September 2024; Received in revised form 14 November 2024; Accepted 25 November 2024
Available online 28 November 2024
0370-2693/© 2024 The Author(s). Published by Elsevier B.V. Funded by SCOAP³. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

There are many possible top down paths to low energy theories like the Standard Model. Typical ingredients for obtaining four-dimensional effective field theories starting from ten-dimensional superstring theory include: i) a Calabi–Yau (CY) threefold, ii) orientifold planes, iii) fluxes, and iv) D-branes with open strings. The landscape of possible effective field theories is vast, in part due to the large number of CY geometries, and we do not yet fully understand the principles that underlie a solution to the vacuum selection problem. One obstacle is that the number of explicit models is still rather small, and isolated to particular corners of the space of theories. In this work, we investigate the first building block (i) and use machine learning to propose new CY manifolds for string compactification.

The largest known class of compact CY threefolds is constructed as hypersurfaces in toric varieties: any fine regular star triangulation (FRST) of a four-dimensional reflexive polytope defines a toric Fano variety in which any generic anticanonical hypersurface is a smooth CY threefold [1,2]. Each such triangulation gives rise to a different ambient toric variety and potentially different CY hypersurfaces. However, triangulations of two different polytopes can also produce equivalent CY hypersurfaces.

In 2000, Kreuzer and Skarke (KS) constructed all 473,800,776 four-dimensional reflexive polytopes. The CY manifolds inherit their topological properties from the polytope, giving rise to threefolds with 30,108 distinct Hodge diamonds [3].[1] As we lack a complete classification of all FRSTs of the polytopes, we do not know how many different CY geometries there are. The total number of FRSTs of polytopes in the KS list has been estimated to be as large as 1.53×10^{928} [6], though the number of topologically inequivalent CY threefolds is known to be much smaller than this. Generalizing the discussion in [7], there has been revived interest in calculating diffeomorphism classes of CY threefolds [8–10]. A consequence of Wall's theorem [11] is that the CY hypersurfaces coming from two different FRSTs of the same reflexive polytope Δ that have the same restriction on the two-faces of Δ are topologically equivalent. It has been estimated that the number of non two-face equivalent FRSTs of four-dimensional reflexive polytopes is $< 1.65\times10^{428}$ [6]. Both the upper bounds are dominated by the single largest polytope with 496 lattice points and which produces CY threefolds with $h^{1,1}=491$.[2] For low values of $h^{1,1}$, it is possible to exhaustively compute FRSTs of four-dimensional reflexive polytopes [13], but it is clearly infeasible to sustain this brute force attack once the number of Kähler moduli increases. We therefore seek a method of efficiently and effectively sampling triangulations.

With multiple criteria, finding FRSTs is an example of a multiobjective optimisation problem (MOOP). Regression is unsuitable for this type of problem since any loss function will have local minima, where triangulations satisfy some but not all of the desired criteria. Popular methods for solving MOOPs are genetic algorithms (GAs) and reinforcement learning (RL), which explore an "environment" in order to maximize a fitness or reward function. GAs have been shown to generate reflexive polytopes [14] and FRSTs [15], and RL has achieved successes for string model building in other contexts [16–19]. Here, we extend this work to obtain FRSTs of reflexive polytopes with RL.

The organization of this paper is as follows. In Section 2, we present the RL model used in the analysis. In Section 3, we demonstrate that we can obtain FRSTs of reflexive polytopes for low $h^{1,1}$. We perform targeted searches for CY geometries whose intersection numbers are conducive to string model building and use our algorithm to find fibration structures of the CYs. In Section 4, we outline the next steps and conclude. Appendix A recalls relevant features of toric geometry. Appendix B reviews RL.

Our code, is available on GitHub [20].

2. The model

In our investigations, we apply deep Q-learning, a model-free reinforcement learning (RL) algorithm involving an optimal Q-function $Q : \mathcal{S}\times\mathcal{A}\to\mathbb{R}$ which gives the value of action $a\in\mathcal{A}$ in a particular state $s\in\mathcal{S}$. In short, given a state s_t, the Q-learning algorithm selects an action a_t, computes the reward and moves to the next state s_{t+1}. The Q-function is then updated based on the Bellman equation (B.4). In deep Q-learning as opposed to regular Q-learning the Q-function is represented by a deep neural network. Additional details about RL and Q-learning are in Appendix B.

The neural network used for the Q-function consists of two hidden layers of 100 and 200 neurons using ReLU activation function. The loss function used was mean squared error and the network was trained using Adam optimisation. In all our investigations the learning rate α and discount factor γ in (B.4) were set to 0.7 and 0.618, respectively.

Wall's theorem [11] establishes that two CY threefolds are topologically equivalent if they share the following invariants: i) Hodge numbers $h^{1,1}$ and $h^{1,2}$, ii) first Pontrjagin class p_1, or equivalently, the second Chern class c_2, and iii) triple intersection numbers d_{ijk}. The second Chern class and the triple intersection numbers are defined up to $SL(h^{1,1},\mathbb{Z})$ equivalence.

The Hodge numbers of CY threefolds constructed as hypersurfaces in toric varieties built from four-dimensional reflexive polytopes can be obtained from the polytope data alone. The second Chern class and the triple intersection numbers, however, depend on the triangulation of the two-faces. Thus, two FRSTs $\mathcal{T}_1$ and $\mathcal{T}_2$ of the same reflexive polytope Δ which give the same two-face restriction give rise to topologically equivalent CY threefolds [11,13,21]. With this in mind, we encode triangulations of the polytope Δ by their two-face restrictions. We first compute all the two-faces of Δ: $\{F_1,...,F_N\}$ and choose an ordering for these faces. Then for each two-face F_i, we compute all fine, regular triangulations $\{\mathcal{T}^i_1,...,\mathcal{T}^i_{M_i}\}$ and decide on an ordering for these triangulations. We then define a state $s\in\mathcal{S}$ as a collection of two-face triangulations given by the two-dimensional bitlist s^i_j where $s^i_j=1$ if the j-th triangulation $\mathcal{T}^i_j$ of F^i is chosen and 0 otherwise.

Example 1. Let Δ_1 be the three-dimensional reflexive polytope defined as

$$\Delta_1=\mathrm{Conv}\begin{pmatrix}1&1&1&-3\\0&2&0&-2\\0&0&2&-2\end{pmatrix}, \tag{2.1}$$

where Conv denotes the convex hull of the vertices (columns), as shown in Fig. 1. There are four two-faces of Δ_1, and each of these two-faces has four fine, regular triangulations. Fig. 2 shows the possible triangulations for one of the two-faces. Therefore, the two-face triangulation states $s\in\mathcal{S}$ will be represented as a 4×4 bitlist, *e.g.*,

$$s=[[1,0,0,0],[1,0,0,0],[0,0,1,0],[0,0,0,1]], \tag{2.2}$$

where, for example the sublist $[0,0,1,0]$ corresponds to the two-face triangulation represented in 2c.

With the state space just described, we define the action space $\mathcal{A}$ as all possible changes of two-face triangulations: $\mathcal{T}^i_j\to\mathcal{T}^i_k\ \forall\ i,j,k$.[3]

[1] The statistics of polytopes is discussed in [4], with supervised machine learning in [5].

[2] The numbers obtained from a naïve counting of effective field theories become still more astronomic in F-theory when we consider toric CY fourfolds [12].

[3] This includes the identity maps $\mathcal{T}^i_j\to\mathcal{T}^i_j$ since the action space must be fixed.

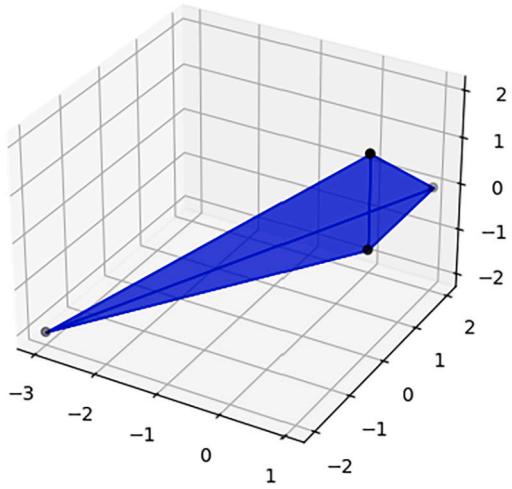


Fig. 1. Reflexive polytope from (2.1).

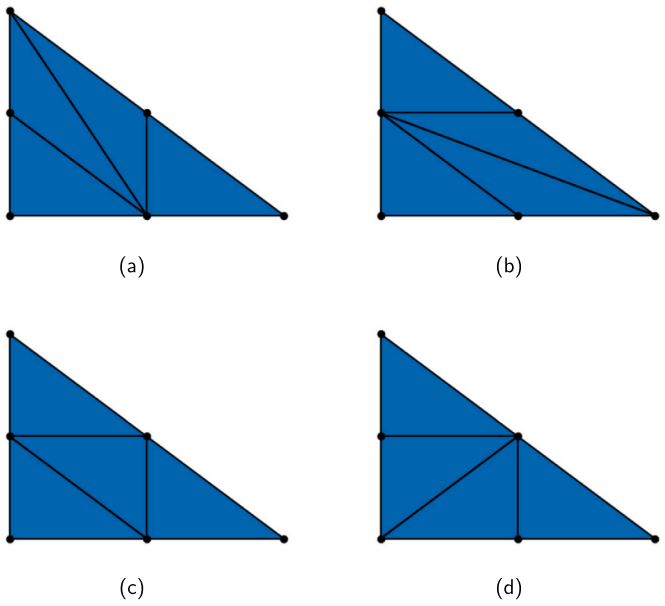


Fig. 2. All possible fine regular triangulations of a two-face of the polytope Δ defined by (2.1).

Finally, we wish to define the fitness function $f : \mathcal{S} \to [0, 1]$ that determines how close a state s is from defining an FRST of the polytope. To do so, we first combine the two-face triangulations of the state $s \in \mathcal{S}$ to get the full triangulation $\mathcal{T}$ of the polytope. We then check the fine, regular, and star conditions of $\mathcal{T}$, introduced in Appendix A.2. In doing so, we only need for $\mathcal{T}$ to be fine with respect to two-faces since the anticanonical divisor defining the CY does not intersect the divisors associated with points interior to facets. In our encoding, all the two-face triangulations $\mathcal{T}_j^i$ are defined to be fine and therefore $\mathcal{T}$ will always be fine with respect to the two-faces and we do not need to check this condition. Furthermore, any non-star, regular triangulation can be converted into a star triangulation by lowering the height of the origin until it appears as a vertex in all simplices. Therefore we also do not need to check the star condition, and it remains only to check regularity of the combined triangulation $\mathcal{T}$.

Recall that for a point configuration $\mathbf{A} = (\mathbf{a}_1, ..., \mathbf{a}_n)$, the secondary cone C of a regular triangulation $\mathcal{T}$ of $\mathbf{A}$ is the space of all height vectors $h = (h_1, ..., h_n)$ that produce $\mathcal{T}$. It follows that $\mathcal{T}$ is regular if and only if C is solid, *i.e.*, full-dimensional. Now, let $\mathbf{A}_1, ..., \mathbf{A}_n$ be multiple point configurations, with corresponding regular triangulations $\mathcal{T}_1, ..., \mathcal{T}_n$ and let $\mathbf{A} = \cup_{i=1}^n \mathbf{A}_i$. If we embed the secondary cones C_i of $\mathcal{T}_i$ into the height space of $\mathbf{A}$ and take the intersection $C = \cap_{i=1}^n C$, it follows that if C is solid then there exists a regular triangulation $\mathcal{T}$ of $\mathbf{A}$ that when restricted to $\mathbf{A}_1, ..., \mathbf{A}_n$ produces the regular triangulations $\mathcal{T}_1, ..., \mathcal{T}_n$. From this we can determine whether a two-face triangulation state s produces a regular triangulation of the ambient polytope. Firstly, we compute all the secondary cones C_i of the two-face triangulations and then compute the intersection $C = \cap_{i=1}^n C_i$. The fitness function of a state $s \in S$ is then defined as:

$$f_{\mathcal{T}}(s) = \frac{N_p(F_\Delta) - \dim C}{N_p(F_\Delta)}, \tag{2.3}$$

where $N_p(F_\Delta)$ is the number of points (not interior to two-faces) of the polytope Δ and therefore this is also the ambient dimension of the intersection cone C. Given a fitness function f the reward function of a state-action pair is defined as

$$R(s, a) = f(a(s)) - f(s). \tag{2.4}$$

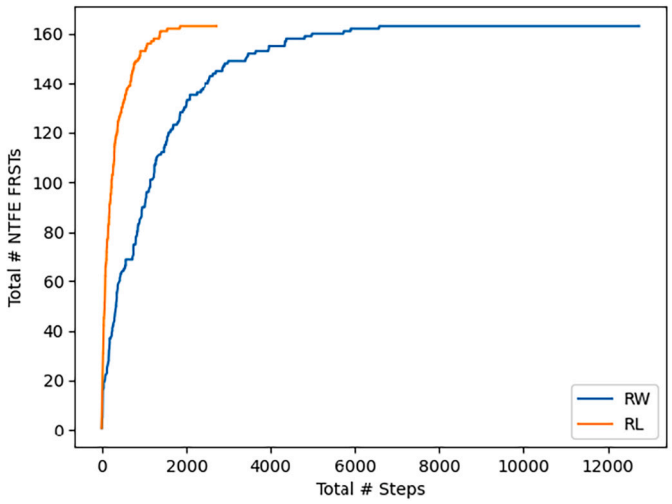


Fig. 3. Plot of the total number of non two-face equivalent fine regular star triangulations found against the number of steps taken for a random walk search (RW) and using the trained reinforcement learning model (RL).

3. Results

To showcase the capability of our RL model at generating datasets of non two-face equivalent (NTFE) FRSTs for a given reflexive polytope, we start with an example. Let Δ_2 be the four-dimensional reflexive polytope defined as:

$$\Delta_2 = \text{Conv}\begin{pmatrix} 0 & 4 & -2 & 0 & 0 & -2 & 1 \\ 0 & -2 & 2 & 0 & 1 & -2 & 0 \\ 1 & -1 & -1 & 0 & 0 & 2 & 0 \\ 0 & -2 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}. \tag{3.1}$$

Δ_2 has 17 points, 19 two-faces and the maximum number of fine regular triangulations of a two-face is 5. The Hodge numbers for the associated CYs are $h^{1,1} = h^{1,2} = 15$.

We trained our RL algorithm over 1000 episodes and then used the trained model to generate a dataset of NTFE FRSTs for Δ, terminating when no new triangulations are found for 1000 episodes. For comparison, we also run a random walk search. The total number of NTFE FRSTs found against total cumulative number of steps taken in both cases are shown in Fig. 3. We see that both searches reach the same total, highlighting that the RL does not introduce further bias towards generating certain FRSTs, and, moreover, RL is far more efficient and takes far fewer steps. Of course we do not know with certainty what fraction of FRSTs we have found. It is possible that there are further FRSTs that exist. On

Table 1
Counts of non two-face equivalent (NTFE) fine regular star triangulations (FRSTs) of reflexive polytopes with low $h^{1,1}$, with the number of steps taken by a random walk (RW) search and by the trained reinforcement learning (RL) model to generate them all. Also listed in the third and fifth columns are the total numbers of FRSTs and FRST classes, respectively, at each $h^{1,1}$ value.

$h^{1,1}$	# polys	# FRSTs	# NTFE FRSTs	# FRST classes	RW # Steps	RL # Steps
1	5	5	5	5	0	0
2	36	48	39	36	7	6
3	244	526	306	275	144	107
4	1197	5348	2022	1774	2658	2292
5	4990	57050	13814	11847	41902	33148

the other hand, since our RL search produces the same total as the random walk search, it seems likely that we have found a large fraction of all FRSTs.

The states in the RL model encode a choice of fine regular triangulation for each two-face which can be combined into a fine star triangulation of the polytope by adding the origin as a vertex to every simplex. However, not every combination of two-face triangulations will give a regular triangulation of the ambient polytope. The agent's job, therefore, is to search for the fine, regular, star triangulations of the polytope within a larger set of fine star triangulations. In this sense, the RL model does not triangulate the polytope itself but rather selects appropriate FRSTs from a set of pre-existing triangulations. Existing algorithms for finding triangulations, such as the `triangulations` function in SageMath [22] and the `all_triangulations` function in CYTools [23], which both rely on TOPCOM [24], either take too long to run or run out of memory and fail on large polytopes, such as this one. This is where our RL model, or the GA model in [15], provide a better alternative.

3.1. Low $h^{1,1}$

In a recent paper [10], the authors classified the number of equivalence classes of CY threefolds, constructed from four-dimensional reflexive polytopes with $h^{1,1} \in \{1,2,3,4,5\}$ and their triangulations. The authors determined the number of CY classes by, first enumerating all NTFE FRSTs for all polytopes at a given $h^{1,1}$ value, and then computing and comparing Hodge numbers, triple intersection numbers, and second Chern classes. The counts from [10] are presented in Table 1.

To test the capability of our RL algorithm at generating complete datasets of NTFE FRSTs for reflexive polytopes we trained the model separately for each polytope with $h^{1,1} \in \{1,2,3,4,5\}$ and used the trained model to search for FRSTs. We run the search until no new triangulations are found in 1000 episodes, with both methods generating the complete list of NTFE FRSTs are found for every polytope. The total cumulative number of steps taken by the RL model is given in Table 1 as well as the number of steps taken by a random walk search. These results show that RL can successfully generate complete datasets of NTFE FRSTs and is also more efficient that random walk searches, particularly for polytopes with larger $h^{1,1}$.

3.2. Bundles

Analogous to the targeted search presented in [14] for reflexive polytopes with GAs, we now present an example of a targeted search for FRSTs of reflexive polytopes by modifying our RL model. This search is inspired by [25], and complements other machine learning approaches to CY bundles [26–28]. In [25], the authors consider $E_8 \times E_8$ heterotic string compactified on smooth CY threefolds X with holomorphic vector bundles V. They showcase how GAs can be used, given a CY threefold X, to find suitable bundles V that satisfy a set of conditions, namely i) anomaly cancellation, ii) slope-stability, and iii) appropriate particle spectrum. They focus on the particular case where V is a sum of line bundles, in which case the conditions are straightforward to check. Specifically, V is taken to be a rank-5 line bundle sum $V = \bigoplus_{a=1}^{5} L_a$, where $L_a = \mathcal{O}_X(L_a)$, so that the resulting model has $SU(5) \times S(U(1)^5)$ symmetry. The line bundle L_a has first Chern class $c_1(L_a) = k_a^i J_i$, where $k_a \in \mathbb{Z}^h$ are integer vectors and $(J_1, ..., J_h)$ is a suitably chosen basis of $H^2(X, \mathbb{Z})$, where $h = h^{1,1}(X)$. Therefore, the line bundle L_a can be written as $\mathcal{O}_X(k_a^1, k_a^2, \dots, k_a^h)$. In general, k_a^I are a mixture of positive and non-positive integers. The five integer vectors $(k_1, ..., k_5)$ therefore uniquely specify the line bundle sum V.

In our targeted RL search we focus on the first three conditions (C1)–(C3) from [25]:

1. E_8 **Embedding**:
$$c_1(V) = \sum_{a=1}^{5} k_a = 0\,. \tag{3.2}$$
This condition ensures that the structure group of V is $S(U(1)^5)$ and not a smaller subgroup.
2. **Anomaly Cancellation**:
$$c_{2,i}(V) = -\frac{1}{2} d_{ijk} \sum_{a=1}^{5} k_a^j k_a^k \leq c_{2,i}(TX)\,, \tag{3.3}$$
for all $i = 1, ..., h$, where d_{ijk} denotes the triple intersection numbers and $c_2(TX)$ denotes the second Chern class of the tangent bundle of X relative to the basis $(J_1, ..., J_h)$.
3. **Poly-Stability**: There exists a non-trivial solution t^i to
$$\mu(L_a) = d_{ijk} k_a^i t^j t^k\,, \tag{3.4}$$
for all $a = 1, ..., 5$, such that $J = t^i J_i$ is in the interior of the Kähler cone, which in the case where $V = \bigoplus_{a=1}^{5} L_a$ corresponds to $t^i > 0$. This check is computationally expensive but can be replaced with the weaker condition that the matrices $M_a = (d_{ijk} k_a^i)$ for $a = 1, ..., 5$, and any linear combination $v^a M_a$, all have at least one negative and one positive entry. Practically, considering v_a, with entries in $\{-2, -1, 0, 1, 2\}$ provides a strong enough check.

In contrast to the GA search performed in [25], where the CY geometry was fixed, we search the space of CYs[4] and line bundles simultaneously. We extend the state encoding from before to include the matrix $(k_1, ..., k_5)$ that specifies the line bundle sum. Furthermore, we modify the fitness function (2.3) to include terms that penalise individuals that do not satisfy the constraints (C1)–(C3):

$$f_{\mathcal{T}-V}(s) = w_1 f_{\mathcal{T}}(s) + w_2 f_{\text{GUT}} + w_3 f_{\text{anom}} + w_4 f_{\text{slope}}\,, \tag{3.5}$$

where f_{GUT}, f_{anom}, f_{slope} are the contributions associated with the E_8 embedding, cancellation of anomalies, and slope stability, respectively, and the $w_i \in \mathbb{R}^{>0}$ are weights, which we set to $w_1 = w_2 = w_3 = w_4 = \frac{1}{4}$. The E_8 GUT embedding, anomaly cancellation, and slope stability condition contributions are given explicitly as

[4] The space of CYs is restricted to those originating from the same polytope but with different FRSTs.

$$f_{\text{GUT}} = \frac{1}{h}\sum_{i=1}^{h}\begin{cases}1 & \text{if} \quad \sum_{a=1}^{5} k_a^i = 0\,,\\ 0 & \text{if} \quad \sum_{a=1}^{5} k_a^i \neq 0\,,\end{cases} \tag{3.6}$$

$$f_{\text{anom}} = \frac{1}{h}\sum_{i=1}^{h}\begin{cases}1 & \text{if} \quad -\frac{1}{2}\sum_{a=1}^{5}\sum_{j=1}^{h}\sum_{k=1}^{h} d_{ijk}k_a^j k_a^k \leq c_2\,,\\ 0 & \text{if} \quad -\frac{1}{2}\sum_{a=1}^{5}\sum_{j=1}^{h}\sum_{k=1}^{h} d_{ijk}k_a^j k_a^k > c_2\,,\end{cases} \tag{3.7}$$

$$f_{\text{slope}} = \frac{1}{3125}\sum_{n=1}^{3125}\begin{cases}1 & \text{if } \max(v_n)^a M_a > 0 \text{ and } \min(v_n)^a M_a < 0\,,\\ 0 & \text{if not } \max(v_n)^a M_a > 0 \text{ and } \min(v_n)^a M_a < 0\,,\end{cases} \tag{3.8}$$

where $M_a = d_{ijk}k_a^i$ and $v_n \in \mathbb{Z}^5$ with entries in $\{-2,-1,0,1,2\}$. There are $5^5 = 3125$ v_n vectors which explains this value appearing in the f_{slope} formula. The $\frac{1}{h}$ and $\frac{1}{3125}$ factors ensure that the maximum value for each contribution is 1 and therefore each condition contributes equally to the fitness.

In [29], the authors determine all CY threefolds constructed from the KS database, which have non-trivial first fundamental group, and classify all $SU(5)$ and $SO(10)$ line bundle models on these manifolds which satisfy the conditions (C1)–(C3). We extract one of these example manifolds from [29], which we know admit appropriate $SU(5)$ bundles and apply our RL model to look for compatible $(\mathcal{T}, k)$ triangulation-bundle pairs. The example we consider is defined by[5]

$$\Delta_3 = \text{Conv}\begin{pmatrix} 0 & 0 & 0 & -2 & 0 & 0 & 2 & 0\\ -1 & -1 & 1 & 0 & 0 & 0 & 0 & -1\\ -1 & 1 & 0 & 0 & -1 & 1 & 0 & 0\\ -1 & 1 & 0 & -1 & 0 & 0 & 1 & 0\end{pmatrix}. \tag{3.9}$$

From [29], we know that all appropriate line bundle sums of these manifolds have k matrices with integers in the range $[-4,4]$. Therefore, we use this range to define our state space and train the RL model over 1000 episodes. An episode ends either when a $(\mathcal{T}, k)$ pair is found, such that $\mathcal{T}$ is an FRST of Δ and k defines a line bundle sum that satisfies all the constraints, or the number of steps taken reaches 1000 and no terminal state is found. Once the model has been trained, we use it to search for $(\mathcal{T}, k)$ pairs. An example of a triangulation and line bundle sum pair found by the RL after training is given below.

Example 2. An FRST of the reflexive polytope defined by (3.9) with compatible line bundle sum found by the trained RL model is given by:

$$\begin{aligned}\mathcal{T} = \{&\{0,1,3,4,5\},\{0,1,3,4,6\},\{0,1,3,5,7\},\{0,1,3,6,7\},\{0,1,4,5,8\},\\ &\{0,1,4,6,8\},\{0,1,5,7,8\},\{0,1,6,7,8\},\{0,2,3,4,5\},\{0,2,3,4,6\},\\ &\{0,2,3,5,7\},\{0,2,3,6,7\},\{0,2,4,5,8\},\{0,2,4,6,8\},\{0,2,5,7,8\},\\ &\{0,2,6,7,8\}\}\end{aligned} \tag{3.10}$$

$$k = \begin{pmatrix} 1 & -2 & -1 & 1 & 1\\ 0 & 1 & -1 & 0 & 0\\ -1 & 0 & 3 & -1 & -1\\ 0 & -1 & 1 & 0 & 0\end{pmatrix}, \tag{3.11}$$

where in $\mathcal{T}$ the 0 index corresponds to the origin and the indices $1-8$ correspond to the points whose coordinates are given in columns $1-8$ of (3.9).

3.3. Fibrations

Recently there has been progress in identifying the existence of certain fibration structures of a toric CY using the geometric data of the polytope [30–33]. Explicit construction of the fibration structure involves finding a reflexive subpolytope $\Delta' \subset \Delta$ and a triangulation compatible with Δ' and Δ. A brief overview of fibration structures is given in Appendix A.3. From the results of [31], existence of a two-dimensional reflexive subpolytope inside a four-dimensional reflexive polytope implies the existence of compatible triangulations yielding an elliptic-fibration structure in the corresponding family of CY threefolds. However, this in general is not the case. The only other exception being codim = 1 subpolytopes, where there always exist compatible triangulations. In this section, we apply the same RL model used in the previous section to search for reflexive subpolytopes and compatible triangulations. We are able to search for subpolytopes and compatible triangulations of any given dimension, however we are particularly interested in the case of two- and three-dimensional subpolytopes, which correspond to elliptic- and K3-fibrations respectively.

3.3.1. Subpolytopes

We begin by applying our RL model to search for reflexive subpolytopes. The encoding of the subpolytopes Δ' is outlined in Appendix A.3. An algorithm which uses this encoding to find fibration structures of toric CYs, by enumerating all possible subspaces, has been proposed in [32]. We instead develop a Q-learning based method to search for such structures. We define the fitness function by penalizing states that do not meet the required subpolytope conditions Δ'. More precisely, we have:

$$f_{\Delta'}(s) = f_{\text{reflex}}(s) + f_{\text{rank}}(s)\,, \tag{3.12}$$

where the two terms check whether the subpolytope is reflexive and has the appropriate dimension. Explicitly,

$$f_{\text{reflex}}(s) := \begin{cases}1 & \text{if } \Delta' \text{ is reflexive}\,,\\ -1 & \text{otherwise}\,,\end{cases} \tag{3.13}$$

$$f_{\text{rank}}(s) := -(\dim\Delta' - d)^2\,, \tag{3.14}$$

where Δ' is the subpolytope associated to the state s and d is the desired subpolytope dimension. This provides a "maze"-like environment for Q-learning. As an example, consider the four-dimensional reflexive polytope:

$$\Delta_4 = \text{Conv}\begin{pmatrix} 1 & 0 & 0 & 0 & -1 & -1\\ 0 & 1 & 0 & 0 & -1 & -1\\ 0 & 0 & 1 & 0 & 0 & -1\\ 0 & 0 & 0 & 1 & 0 & -1\end{pmatrix}. \tag{3.15}$$

The space of states S corresponding to virtual two-fibrations is two-dimensional. We plot the fitness $f_{\Delta'}$ and Q-learned policy maps in Fig. 4. The polytope (3.15) admits only a single virtual two-fibration structure given by the subpolytope:

$$\Delta_4' = \text{Conv}\begin{pmatrix} 1 & 0 & -1\\ 0 & 1 & -1\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}. \tag{3.16}$$

To demonstrate the generality of this approach, we consider the case when Δ is a five-dimensional reflexive polytope:

Example 3. Let $\Delta_5 \subset N \simeq \mathbb{Z}^5$ be a reflexive polytope:

$$\Delta_5 = \text{Conv}\begin{pmatrix} -2 & 0 & 0 & 1 & 1 & 2\\ 0 & -1 & 0 & 0 & 0 & 1\\ -1 & -1 & 1 & 0 & 0 & 1\\ 2 & 1 & 2 & -3 & -1 & -5\\ 5 & 0 & -1 & -3 & -3 & -2\end{pmatrix}. \tag{3.17}$$

The Q-learned two-dimensional reflexive subpolytopes are generated by vertex matrices:

[5] In [29] this CY is labeled as X_{13}.

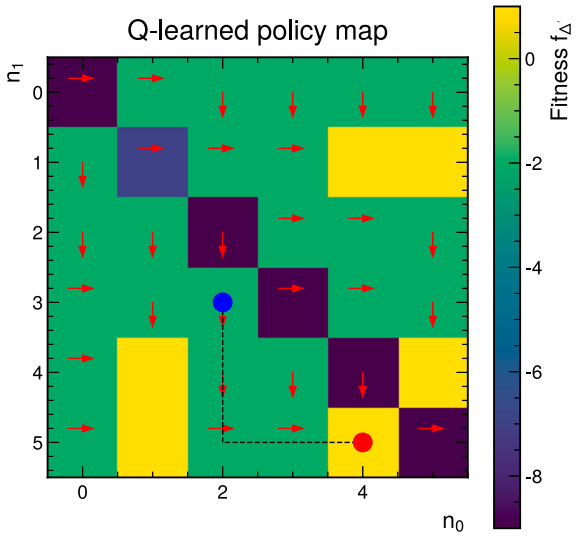


Fig. 4. The fitness and Q-learned policy maps on the state space of (3.15). The policies are denoted by the arrows $\{\rightarrow, \downarrow\}$. The terminal point of the path, starting at the blue dot and terminating at the red dot, corresponds to the elliptic fiber (3.16). The coordinates (n_0, n_1) correspond to vertices (v_{n_0}, v_{n_1}) spanning the plane of Δ'.

$$\begin{pmatrix} 0 & -2 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 2 & 2 & 2 & -2 \\ 0 & 5 & -1 & 2 & -2 \end{pmatrix}, \quad \begin{pmatrix} 0 & -1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & -2 & -2 & -2 \\ 0 & 2 & -3 & -1 & -2 \end{pmatrix}, \tag{3.18}$$

while the Q-learned three-dimensional reflexive subpolytopes are generated by:

$$\begin{pmatrix} 0 & -2 & 0 & -1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 2 & 2 & -2 & -2 & -2 \\ 0 & 5 & -1 & 2 & -3 & -1 & -2 \end{pmatrix},$$

$$\begin{pmatrix} 0 & 0 & 2 & -1 & 1 & 1 & 1 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -5 & 2 & -2 & -2 & -2 \\ 0 & 0 & -2 & 2 & -3 & -1 & -2 \end{pmatrix} \tag{3.19}$$

$$\begin{pmatrix} 0 & 1 & 1 & -1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -3 & -1 & 2 & -2 & -2 & -2 \\ 0 & -3 & -3 & 2 & -3 & -1 & -2 \end{pmatrix}. \tag{3.20}$$

Finally, the Q-learned four-dimensional reflexive subpolytopes are generated by:

$$\begin{pmatrix} 0 & -2 & 0 & 0 & 2 & -1 & 1 & 1 & 1 \\ 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 2 & -5 & 2 & -2 & -2 & -2 \\ 0 & 5 & 0 & -1 & -2 & 2 & -3 & -1 & -2 \end{pmatrix},$$

$$\begin{pmatrix} 0 & -2 & 0 & 1 & 1 & -1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 2 & -3 & -1 & 2 & -2 & -2 & -2 \\ 0 & 5 & -1 & -3 & -3 & 2 & -3 & -1 & -2 \end{pmatrix} \tag{3.21}$$

$$\begin{pmatrix} 0 & 0 & 1 & 1 & 2 & -1 & 1 & 1 & 1 \\ 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -3 & -1 & -5 & 2 & -2 & -2 & -2 \\ 0 & 0 & -3 & -3 & -2 & 2 & -3 & -1 & -2 \end{pmatrix}. \tag{3.22}$$

The four-dimensional virtual fibrations corresponding to the reflexive subpolytopes above are guaranteed to have a compatible set of fans by the results of [32].

3.3.2. Compatible triangulations

In order to incorporate triangulations into the search, we extend the state space. In particular, the state space for searching for d-fibration structures together with compatible triangulations, is given as a product: $\mathcal{S} = \mathcal{S}_{\text{triang.}} \times \mathcal{S}_{\text{fib.}}$, where $\mathcal{S}_{\text{triang.}}$ corresponds to the state space of triangulations defined in Section 2, and $\mathcal{S}_{\text{fib.}}$ corresponds to the state space of virtual fibrations defined in Appendix A.3. The action space $\mathcal{A}$ is constructed in the same manner. The fitness function is extended by adding a penalty ensuring compatibility of the fan corresponding to the triangulation with the virtual fibration structure:

$$f_{\Delta'-\mathcal{T}}(s_1, s_2) = f_{\Delta'}(s_1) + f_{\mathcal{T}}(s_2) + f_{\text{compat}}(s_1, s_2), \tag{3.23}$$

where s_1, s_2 describe the subpolytope and two-face triangulations respectively, and f_{compat} is the contribution associated to the compatibility condition. In particular,[6] [33].

$$f_{\text{compat}}(s_1, s_2) := \begin{cases} 1 & \text{if the restriction } \mathcal{T}|_{\Delta'} \text{ is an FRST of } \Delta', \\ 0 & \text{otherwise}. \end{cases} \tag{3.24}$$

As a proof-of-concept, an example using the reward function (3.23) is presented below.

Example 4. Let $\Delta_6 \subset N \simeq \mathbb{Z}^4$ denote the reflexive polytope:

$$\Delta_6 = \text{Conv} \begin{pmatrix} -1 & 0 & 0 & 1 & -1 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 & -1 & -1 & 1 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{3.25}$$

Two-fibration structures obtained using the trained Q-learning model, given by reflexive two-dimensional subpolytopes Δ' and Δ'' and their respective compatible triangulations are shown in Fig. 5. The vertices of Δ' and Δ'' are:

$$\Delta' = \text{Conv} \begin{pmatrix} 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \Delta'' = \text{Conv} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{3.26}$$

It is easy to see that both Δ' and Δ'' are indeed reflexive. We may also look for three-fibration structures of (3.25) (corresponding to K3 fibered CYs). In particular, Δ admits a three-dimensional reflexive subpolytope ΔK3 given by:

$$\Delta\text{K3} = \text{Conv} \begin{pmatrix} 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{3.27}$$

the visualization of a triangulation compatible with the virtual three-fibration $\Delta\text{K3} \hookrightarrow \Delta$ is shown in Fig. 6.

[6] Note that it might be the case that even if there exists a fibration structure, there is no compatible triangulation generated using two-face encoding. In those cases, for each state we compute neighboring triangulations using bistellar flips [34].

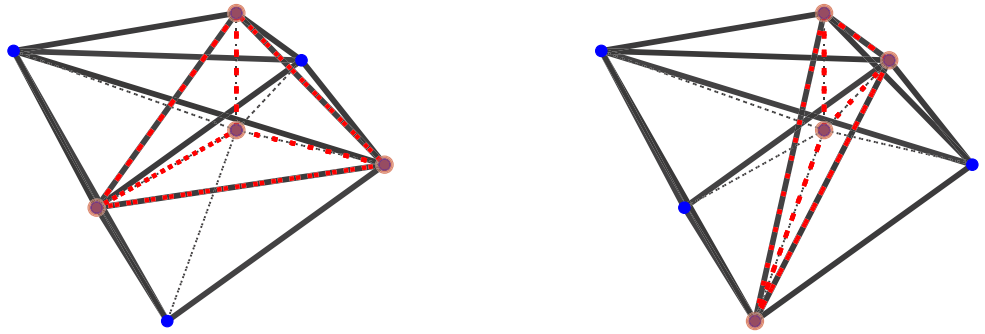


Fig. 5. Two-fibrations $\Delta', \Delta'' \hookrightarrow \Delta$ with compatible triangulations found using the trained Q-learning model with reward function (3.23). The vertices are projected onto $\mathbb{Z}^3$ by taking first three components in $\mathbb{Z}^4$. The red dots and edges correspond to the subpolytopes Δ' (a) and Δ'' (b) (3.26).

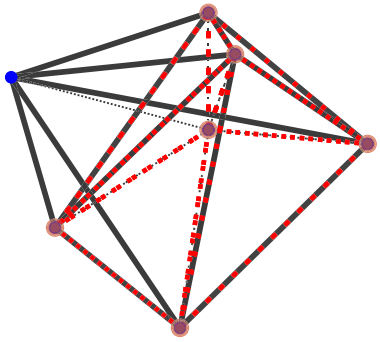


Fig. 6. A three-fibration $\Delta\mathrm{K3} \hookrightarrow \Delta$ with compatible triangulation computed using Q-learning with reward function (3.23).

4. Outlook

Following on from the work carried out in [14], in this paper we have shown how reinforcement learning (RL) can be efficiently used to generate fine regular star triangulations (FRSTs) of four-dimensional reflexive polytopes. Such triangulations provide resolutions of non-terminal singularities in the ambient toric variety, and consequently the Calabi-Yau (CY) threefold hypersurfaces, constructed from the corresponding polytope. As detailed in [6], the total number of FRSTs of four-dimensional reflexive polytopes in the complete classification [3], is as large as 1.53×10^{928}.[7] It is therefore not feasible to generate a complete list of triangulations and instead we would like a way of generating fair samples. In the context of string theory, we would like to generate FRSTs that produce CY threefolds satisfying certain physical constraints.

Our results show that RL can generate complete datasets of FRSTs for reflexive polytopes with low $h^{1,1}$, where the total number of FRSTs is manageable. Specifically, we regenerated all non two-face equivalent FRSTs with $h^{1,1} \in \{1,2,3,4,5\}$ matching the total found in [10]. Following this, we considered the particular case of $E_8 \times E_8$ heterotic string compactified on a smooth CY threefolds X with holomorphic vector bundles V. Fixing the reflexive polytope and searching the space of FRSTs and line bundle sums $V = \bigoplus_{a=1}^{5} L_a$, where $L_a = \mathcal{O}_X(L_a)$, we were able to find examples that satisfy anomaly cancellation and polystability constraints. We also demonstrated how our RL model can be used look for CY manifolds that admit certain fibration structures. Starting from a reflexive polytope, we were able to find three-dimensional reflexive subpolytopes and compatible FRSTs that together define smooth K3 fibered CYs. These targeted search examples, as well as those presented in [14,15], illustrate how one can design a dedicated search for CY manifolds with prescribed properties as required for the intended string compactification.

[7] Triangulations of the largest polytope, with $h^{1,1} = 491$, make up the majority of all triangulations. Some of the two-faces of this polytope are too large that it is not feasible to compute all two-face triangulations. It is for this reason that we have not presented results for generating triangulations of this polytope with RL. It is possible, as was done in [15], to take a sample of triangulations of the two-faces and use these to sample FRSTs of the ambient polytope, however this produces a biased sample.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Per Berglund reports financial support was provided by US Department of Energy grant DE-SC0020220. Yang-Hui He reports financial support was provided by UK Science and Technology Facilities Council grant ST/J00037X/2. Vishnu Jejjala reports financial support was provided by South African Research Chairs Initiative of the Department of Science and Innovation grant 78554. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Andre Lukas for collaboration in initial stages of this project. We are grateful to Nate MacFadden, Andreas Schachner, and Elijah Sheridan for discussions on two-face restrictions and for sharing a pre-arXiv draft of [15]. We thank Liam McAllister for comments on the preprint. We also thank the Pollica Physics Center, Italy, and the organizers of the 2023 Workshop on Machine Learning for a very stimulating environment which is where this project was initiated. PB and GB are supported in part by the Department of Energy grant DE-SC0020220. YHH is supported by STFC grant ST/J00037X/2. E. Heyes is supported by City, University of London and the States of Jersey. E. Hirst is supported by Pierre Andurand. VJ is supported by the South African Research Chairs Initiative of the Department of Science and Innovation and the National Research Foundation grant 78554.

Appendix A. Toric geometry

We briefly recall relevant aspects of toric geometry with the goal of understanding fine regular star triangulations of reflexive polytopes and fibration structures of the resulting toric varieties.

A.1. Reflexive polytopes

Given an n-dimensional lattice polytope $\Delta \subset M$, one can construct a compact toric variety X_Δ of complex dimension n. In short, one con-

structs the normal fan Σ_Δ as follows: for a face θ of Δ, let $\sigma_\theta \subset N$ be the dual of the cone:

$$\sigma_\theta^\vee := \{\lambda(u-u') | u \in \Delta, u' \in \theta, \lambda \geq 0\} \subset M\,. \tag{A.1}$$

The normal fan is then given as $\Sigma_\Delta := \{\sigma_\theta\}$ for all faces θ of Δ. From the normal fan, the construction of the compact projective toric variety X_Δ follows the usual procedure, where each cone σ_θ gives rise to an affine toric variety $U_{\sigma_\theta} = \mathrm{Spec}(\mathbb{C}[\sigma_\theta^\vee \cap M])$ and one glues these patches together.

A polytope is said to satisfy the *IP property* if it has only a single interior lattice point, taken to be the origin.

Definition A.1. A lattice polytope Δ is called **reflexive** if it satisfies the IP property and if its dual Δ^* is also a lattice polytope that satisfies the IP property. Equivalently, the lattice polytope is IP and all of its facets are a unit distance from the origin.

The reflexive polytope is only defined up to $GL(n,\mathbb{Z})$ transformations of the coordinates of the vertices; equivalent reflexive polytopes have the same normal form (which is a specific representative from the $GL(n,\mathbb{Z})$ equivalence class).

The connection between CY manifolds and reflexive polytopes is given by the following theorem due to Batyrev [1,2]:

Theorem A.1. *Let $\Delta \subset M$ be an n-dimensional lattice polytope and X_Δ the corresponding n complex dimensional toric variety. If Δ is reflexive then it follows that X_Δ is Gorenstein Fano with at most canonical singularities and moreover the zero locus of a generic section of the anticanonical bundle $-K_X$ is a CY variety $\mathcal{M}$ of complex dimension $n-1$.*

The mirror CY $\mathcal{W}$ is similarly obtained from the dual polytope.

A.2. Fine regular star triangulations

The variety X_Δ generated by the reflexive polytope Δ may be singular. If it is too singular, then the CY variety $\mathcal{M}$ may not be smooth. We wish to find a resolution of the singularities given by a birational morphism $\pi : \widetilde{X}_\Delta \to X_\Delta$ such that the desingularised space $\widetilde{X}_\Delta$ is smooth enough that any CY variety $\mathcal{M} \subset \widetilde{X}_\Delta$ can be chosen smooth.

Consider the particular case where $n=4$ and therefore $\mathcal{M}$ describes a CY threefold. Since $\mathcal{M}$ has dimension 3, it can be smoothly deformed around singular loci with codimension 3 (terminal singularities). Therefore, we only need to consider desingularisations that resolve everything up to terminal singularities. If $\widetilde{X}_\Delta$ contains terminal singularities (*i.e.*, $\widetilde{X}_\Delta$ is quasi-smooth) we say that $\widetilde{X}_\Delta$ is $\mathbb{Q}$-factorial. Moreover, to ensure the desingularised space $\widetilde{X}_\Delta$ remains Gorenstein Fano, and therefore projective, the desingularisation must be crepant.

We define a maximal projective crepant partial (MPCP) desingularisation $\pi : \widetilde{X}_\Delta \to X_\Delta$ to be one such that the pullback π^* is crepant, and $\widetilde{X}_\Delta$ is $\mathbb{Q}$-factorial. From [1,2], we have the following theorem:

Theorem A.2. *Let X_Δ be the toric Gorenstein Fano variety built from a reflexive polytope Δ with normal fan Σ. Then $\widetilde{X}_\Delta$ admits at least one MPCP desingularisation*

$$\pi : \widetilde{X}_\Delta \to X_\Delta\,. \tag{A.2}$$

Resolution of non-terminal singularities can be seen as refining the open cover $\mathcal{U}(\widetilde{X}_\Delta)$ on $\widetilde{X}_\Delta$. Since each cone $\sigma_\theta \subset N$ in the normal fan Σ_Δ corresponds to a coordinate patch $U \in \mathcal{U}(\widetilde{X}_\Delta)$, refining $\mathcal{U}(\widetilde{X}_\Delta)$ amounts to subdividing the cones Σ_Δ. This subdivision corresponds to a triangulation of the dual polytope $\Delta^* \subset N$.

To be precise, a *triangulation* of an n-dimensional reflexive polytope Δ^* consists of simplices of dimension n such that the intersection of any two simplices is a face of each and the union of all simplices recovers Δ^*.

A triangulation is said to be:

- *star* if the origin is a vertex of every full-dimensional simplex. We require the triangulations of Δ^* to be star so that each subdivision of Σ_Δ is still a convex rational polyhedral cone with a vertex at the origin.
- *fine* if it uses all points in the point configuration. We require the triangulations of Δ^* to be fine so that all non-terminal singularities are resolved and the resulting CY $\mathcal{M}$ is smooth.
- *regular* if it can be obtained from the following construction. Let $\mathbf{A} = (\mathbf{a}_1, ..., \mathbf{a}_m)$ be a point configuration in $\mathbb{R}^n$ and Δ the convex hull of $\mathbf{A}$. Define a height vector $h = (h_1, ..., h_m) \in \mathbb{R}^m$ and consider the lifted point configuration in $\mathbb{R}^{n+1}$

$$\mathbf{A}^h := \begin{pmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_m \\ h_1 & \cdots & h_m \end{pmatrix}. \tag{A.3}$$

  Compute the convex hull of $\mathbf{A}^h$ to obtain the polytope Δ^h and compute the "lower faces" of Δ^h, where the lower faces are those that have a non-vertical supporting hyperplane H and with Δ^h above H. Projecting down the lower faces of Δ^h to $\mathbb{R}^n$ produces a triangulation of Δ. We require the triangulations of Δ^* to be regular so that $\widetilde{X}_\Delta$ is projective, and hence Kähler; the CY hypersurfaces inherit this Kähler structure.

The heights generating some regular triangulation are not unique, and in fact, many heights can lead to the same triangulation. Let $h = (h_1, ..., h_m)$ be the set of heights defining the regular triangulation $\mathcal{T}$ of a polytope Δ. The set of all height vectors that generate the same triangulation form a polyhedral cone C called the secondary cone. A simple argument for this is that if h is a set of heights that generates $\mathcal{T}$, then ch also generates $\mathcal{T}$ for any $c > 0$ and if g is also a set of heights that generates $\mathcal{T}$ then $h+g$ generates $\mathcal{T}$.

A.3. Fibration structures

Given a CY hypersurface in a toric variety X_Σ associated to a polytope $\Delta \subset N$ with a fan Σ, the existence of a fibration of form:

$$X_{\Sigma_{\mathrm{fib}}} \hookrightarrow X_\Sigma \xrightarrow{\pi} B \tag{A.4}$$

over a toric base B with a fan Σ_b and $\Sigma_{\mathrm{fib}} \subset \Sigma$ is obstructed by the set of compatibility conditions on the subpolytope $\Delta_{\mathrm{fib}} \subset \Delta$ and the fans Σ, Σ_b and Σ_{fib} [32].

In the case when searching for $\mathrm{codim}(\Delta_{\mathrm{fib}} \subset \Delta) = 1$ fibration structures, it is sufficient to show that such Δ_{fib} is indeed a reflexive subpolytope [32]. Similarly, in the case when searching for elliptically fibered toric CY threefolds, it has been shown that the existence of a two-dimensional reflexive subpolytope implies existence of compatible fans [31]. However, in general, one must ensure that there indeed exists a set of compatible fans Σ, Σ_b, and Σ_{fib}.

We mainly follow the construction outlined in [32]. In particular, we have:

Definition A.2. Let $\Delta \subset N \simeq \mathbb{Z}^n$ be a reflexive polytope with a fan Σ. A **virtual toric d-fibration** of X_Σ is an embedding $\Delta' \hookrightarrow \Delta$ of a d-dimensional reflexive subpolytope Δ' into Δ.

Construction of a virtual toric d-fibration for a given polytope Δ, assuming one exists, can be formulated as a pure optimization problem by noting that the subpolytope Δ' can be defined using a set of d vertices of Δ. We shall first briefly describe the construction of Δ'. Let $\{v_0, \ldots, v_{k-1}\}$ be the vertices of Δ. Define a state s of the RL environment of virtual d-fibrations as a tuple:

$$s = (n_0, \ldots, n_{d-1}), \tag{A.5}$$

where $n_j \in \mathbb{Z}_k^d$ such that $n_j \neq n_k$ for $j \neq k$. Therefore, the state space $\mathcal{S}$ can be identified as a subset of $\mathbb{Z}_k^d$. A subpolytope Δ' associated to a state s is an intersection of the plane:

$$H(s) = \bigoplus_{i=0}^{d-1} \mathbb{R} v_{n_i}, \tag{A.6}$$

with the polytope Δ, that is: $\Delta' = \Delta \cap H(s)$.

Naturally, we may define an action on the state space $\mathcal{S}$ by the natural group action:

$$\mathbb{Z}_k^d \times \mathcal{S} \longrightarrow \mathcal{S} \tag{A.7}$$

given by the group product:

$$(m_0, \ldots, m_{d-1}) \cdot (n_0, \ldots, n_{d-1}) = (m_0 + n_0, \ldots, m_{d-1} + n_{d-1}). \tag{A.8}$$

Noting that $\mathbb{Z}_k^d$ is cyclic, we may restrict the set of actions on a given state s to the generators:

$$a_0 := (1, 0, \ldots, 0), \quad a_1 := (0, 1, \ldots, 0), \quad \cdots, \quad a_{d-1} := (0, 0, \ldots, 1), \tag{A.9}$$

which yields the set of actions $\mathcal{A} := \{a_0, \ldots, a_{d-1}\}$.

In order to construct the fitness function $f : \mathcal{S} \to \mathbb{R}$, note that there are two main conditions that need to be imposed on the subpolytope $\Delta' = \Delta \cap H(s)$:

1. Δ' must be a reflexive polytope;
2. $\dim \Delta' = d$.

Although the number of the vertices defining $H(s)$, and hence Δ', is d, it is not necessarily guaranteed that the rank of Δ' is equal to d. The fitness function defined in (3.12) penalizes the action on a given state if the resulting state fails to meet any of the conditions above.

Appendix B. Reinforcement learning

Reinforcement learning is an area of machine learning which works on the principal of trial and error. An agent performs actions on states in an environment with the goal of maximizing the cumulative reward. The action changes the state in some way and the agent receives a reward or penalty based on whether the new state is better or worse than the previous one. Over time the agent learns which actions to take given the current state in order to maximize the reward.

The main components of reinforcement learning are:

- a set of states $\mathcal{S}$;
- a set of actions $\mathcal{A}$; and
- a reward function $R(s, a)$ that gives a reward associated to the action a taken on the state s.

An agent interacts with the environment in discrete time steps; at each time step t the agent receives the current state s_t and a reward r_t. The agent then chooses an action from $\mathcal{A}$ and the environment moves to the next state s_{t+1} with reward r_{t+1}. The end goal of the agent is to learn a policy function $\pi : \mathcal{S} \times \mathcal{A} \to [0, 1]$, $\pi(s, a) = \Pr(a_t = a | s_t = s)$, which maximises the expected cumulative reward.

The state-value function $V_\pi(s)$ is defined as the expected discounted return starting at state s and successively following policy π:

$$V_\pi(s) = \mathbb{E}[G | S_0 = s] = \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t R_{t+1} | S_0 = s\right], \tag{B.1}$$

where R_{t+1} denoted the reward of moving from state S_t to S_{t+1} and $0 \leq \gamma < 1$ is the discount rate. Since $\gamma < 1$ it has the effect of valuing rewards received earlier higher than those received later. G denotes the discounted return, and is defined as the sum of future discounted rewards. The agent aims to learn a policy that maximises the expected discounted return. Such a policy is called the optimal policy.

The action-value of a state-action pair (s, a) under a policy π is defined as

$$Q^\pi(s, a) = \mathbb{E}[G | s, a, \pi], \tag{B.2}$$

where G now stands for the random discounted return associated with first taking action a in state s and then following π thereafter. Then if π^* is an optimal policy, the optimal action to take from a state s is the one with the highest action-value $Q^{\pi^*}(s, \cdot)$. The action-value function of such an optimal policy Q^{π^*} is called the optimal action-value function and is commonly denoted by Q^*. Knowledge of the optimal action-value function alone suffices to know how to act optimally.

B.1. Q-learning

One of the most common reinforcement learning algorithms is Q-learning. Q-learning is based on a Q-function that assigns quality for a given action when the environment is in a given state. This value function Q is then iteratively refined. C. Watkins first introduced the foundations of Q-learning in his thesis in 1989 [35] and gave further details in a 1992 publication titled Q-learning [36].

In Q-learning, the Q-function calculates the quality of a state-action combination:

$$Q : \mathcal{Q} \times \mathcal{A} \to \mathbb{R}. \tag{B.3}$$

Firstly Q is initialised to some constant function. Then at each time step t the agent selects an action a_t, observes a reward r_t and moves to a new state s_{t+1} and Q is updated.

Q is updated based on the Bellman equation:

$$Q^{new}(s_t, a_t) \leftarrow (1 - \alpha) Q(s_t, a_t) + \alpha \left(R_{t+1} + \gamma \max_a Q(s_{t+1}, a) \right), \tag{B.4}$$

where R_{t+1} is the reward received when moving from state s_t to s_{t+1}, $0 < \alpha < 1$ is the learning rate and $0 \leq \gamma \leq 1$ is the discount factor. Therefore, $Q^{new}(s_t, a_t)$ is the sum of three factors:

- $(1 - \alpha) Q(s_t, a_t)$: the current value (weighted by one minus the learning rate);
- αR_{t+1}: the reward if action a_t is taken on state s_t (weighted by the learning rate); and
- $\alpha \gamma \max_a Q(s_{t+1}, a))$: the maximum reward that can be obtained from state s_{t+1} (weighted by the learning rate and discount factor).

The learning rate α determines to what extent newly acquired information overrides old information. A factor of 0 makes the agent learn nothing (exclusively exploiting prior knowledge), while a factor of 1 makes the agent consider only the most recent information (ignoring prior knowledge to explore possibilities). In fully deterministic environments, a learning rate of $\alpha_t = 1$ is optimal. When the problem is stochastic, the algorithm converges under some technical conditions on the learning rate that require it to decrease to zero. In practice, often a constant learning rate is used, such as $\alpha_t = 0.1$ for all t.

In its simplest form the Q function data is stored in a table. The rows and columns of the table correspond to different states and actions respectively. The first step is to initialise the Q-table and then as the agent interacts with the environment and receives rewards, the values in the Q-table are updated. In cases where the set of states and action is very large the likelihood of the agent visiting a particular state and performing a particular action is small in which case Q-learning can be combined with function approximation. One solution is to use a deep feed-forward neural network to represent the function. This gives rise to deep Q-learning.

Data availability

We have shared the link to the GitHub in our paper which hosts the code used to generate the results.

References

[1] V.V. Batyrev, Dual polyhedra and mirror symmetry for Calabi-Yau hypersurfaces in toric varieties, arXiv:alg-geom/9310003, 1993.

[2] V.V. Batyrev, L.A. Borisov, On Calabi-Yau complete intersections in toric varieties, arXiv:alg-geom/9412017, 1994.

[3] M. Kreuzer, H. Skarke, Complete classification of reflexive polyhedra in four dimensions, arXiv:hep-th/0002240, 2000.

[4] Y.-H. He, V. Jejjala, L. Pontiggia, Patterns in Calabi–Yau distributions, Commun. Math. Phys. 354 (2) (2017) 477–524, https://doi.org/10.1007/s00220-017-2907-9, arXiv:1512.01579.

[5] J. Bao, Y.-H. He, E. Hirst, J. Hofscheier, A. Kasprzyk, S. Majumder, Polytopes and machine learning, Math. Sci. 01 (2023) 181–211, https://doi.org/10.1142/S281093922350003X, arXiv:2109.09602.

[6] M. Demirtas, L. McAllister, A. Rios-Tascon, Bounding the Kreuzer-Skarke landscape, Fortschr. Phys. (2020), https://doi.org/10.1002/prop.202000086.

[7] T. Hubsch, Calabi-Yau Manifolds: A Bestiary for Physicists, World Scientific, Singapore, 1994.

[8] V. Jejjala, W. Taylor, A. Turner, Identifying equivalent Calabi–Yau topologies: a discrete challenge from math and physics for machine learning, in: Nankai Symposium on Mathematical Dialogues: In Celebration of S.S.Chern's 110th Anniversary, 2022, arXiv:2202.07590.

[9] A. Chandra, A. Constantin, C.S. Fraser-Taliente, T.R. Harvey, A. Lukas, Enumerating Calabi-Yau manifolds: placing bounds on the number of diffeomorphism classes in the Kreuzer-Skarke list, Fortschr. Phys. 72 (5) (2024) 2300264, https://doi.org/10.1002/prop.202300264, arXiv:2310.05909.

[10] N. Gendler, N. MacFadden, L. McAllister, J. Moritz, R. Nally, A. Schachner, M. Stillman, Counting Calabi-Yau threefolds, arXiv:2310.06820, 10 2023.

[11] C.T.C. Wall, Classification problems in differential topology. v, Invent. Math. 1 (1966) 355–374, https://doi.org/10.1007/BF01389738.

[12] W. Taylor, Y.-N. Wang, The F-theory geometry with most flux vacua, J. High Energy Phys. 12 (2015) 164, https://doi.org/10.1007/JHEP12(2015)164, arXiv:1511.03209.

[13] R. Altman, J. Gray, Y.-H. He, V. Jejjala, B.D. Nelson, A Calabi-Yau database: threefolds constructed from the Kreuzer-Skarke list, J. High Energy Phys. 02 (2015) 158, https://doi.org/10.1007/JHEP02(2015)158, arXiv:1411.1418.

[14] P. Berglund, Y.-H. He, E. Heyes, E. Hirst, V. Jejjala, A. Lukas, New Calabi–Yau manifolds from genetic algorithms, Phys. Lett. B 850 (2024) 138504, https://doi.org/10.1016/j.physletb.2024.138504, arXiv:2306.06159.

[15] N. MacFadden, A. Schachner, E. Sheridan, The DNA of Calabi-Yau hypersurfaces, arXiv:2405.08871, 5 2024.

[16] J. Halverson, B. Nelson, F. Ruehle, Branes with brains: exploring string vacua with deep reinforcement learning, J. High Energy Phys. 06 (2019) 003, https://doi.org/10.1007/JHEP06(2019)003, arXiv:1903.11616.

[17] T.R. Harvey, A. Lukas, Quark mass models and reinforcement learning, J. High Energy Phys. 08 (2021) 161, https://doi.org/10.1007/JHEP08(2021)161, arXiv:2103.04759.

[18] A. Constantin, T.R. Harvey, A. Lukas, Heterotic string model building with monad bundles and reinforcement learning, Fortschr. Phys. 70 (2–3) (2022) 2100186, https://doi.org/10.1002/prop.202100186, arXiv:2108.07316.

[19] S. Abel, A. Constantin, T.R. Harvey, A. Lukas, Evolving heterotic gauge backgrounds: genetic algorithms versus reinforcement learning, Fortschr. Phys. 70 (5) (2022) 2200034, https://doi.org/10.1002/prop.202200034, arXiv:2110.14029.

[20] P. Berglund, G. Butbaia, Y.-H. He, E. Heyes, E. Hirst, V. Jejjala, Triangulation generation, https://github.com/elliheyes/Triangulation-Generation, 2024.

[21] N. MacFadden, Efficient algorithm for generating homotopy inequivalent Calabi-Yaus, arXiv:2309.10855, 9 2023.

[22] Sagemath, the sage mathematics software system (version 9.0) (2020).

[23] M. Demirtas, A. Rios-Tascon, L. McAllister, CYTools: a software package for analyzing Calabi-Yau manifolds, https://github.com/LiamMcAllisterGroup/cytools.

[24] J. Rambau, TOPCOM: triangulations of point configurations and oriented matroids, https://doi.org/10.1142/9789812777171_0035, pp. 330–340.

[25] S.A. Abel, A. Constantin, T.R. Harvey, A. Lukas, L.A. Nutricati, Decoding nature with nature's tools: heterotic line bundle models of particle physics with genetic algorithms and quantum annealing, Fortschr. Phys. 72 (2) (2024) 2300260, https://doi.org/10.1002/prop.202300260, arXiv:2306.03147.

[26] D. Klaewer, L. Schlechter, Machine learning line bundle cohomologies of hypersurfaces in toric varieties, Phys. Lett. B 789 (2019) 438–443, https://doi.org/10.1016/j.physletb.2019.01.002, arXiv:1809.02547.

[27] C.R. Brodie, A. Constantin, R. Deen, A. Lukas, Machine learning line bundle cohomology, Fortschr. Phys. 68 (1) (2020) 1900087, https://doi.org/10.1002/prop.201900087, arXiv:1906.08730.

[28] R. Deen, Y.-H. He, S.-J. Lee, A. Lukas, Machine learning string standard models, Phys. Rev. D 105 (4) (2022) 046001, https://doi.org/10.1103/PhysRevD.105.046001, arXiv:2003.13339.

[29] Y.-H. He, S.-J. Lee, A. Lukas, C. Sun, Heterotic model building: 16 special manifolds, J. High Energy Phys. 06 (2014) 077, https://doi.org/10.1007/JHEP06(2014)077, arXiv:1309.0223.

[30] Y.-C. Huang, W. Taylor, Mirror symmetry and elliptic Calabi-Yau manifolds, J. High Energy Phys. 04 (2019) 083, https://doi.org/10.1007/JHEP04(2019)083, arXiv:1811.04947.

[31] Y.-C. Huang, W. Taylor, Fibration structure in toric hypersurface Calabi-Yau threefolds, J. High Energy Phys. 03 (2020) 172, https://doi.org/10.1007/JHEP03(2020)172, arXiv:1907.09482.

[32] F. Rohsiepe, Fibration structures in toric Calabi-Yau fourfolds, arXiv:hep-th/0502138, 2 2005.

[33] J. Knapp, M. Kreuzer, Toric methods in f-theory model building, Adv. High Energy Phys. 2011 (2011) 513436, https://doi.org/10.1155/2011/513436, arXiv:1103.3358.

[34] M. Demirtas, A. Rios-Tascon, L. McAllister, CYTools, https://github.com/LiamMcAllisterGroup/cytools.

[35] C.J.C.H. Watkins, Learning from delayed rewards, Ph.D. thesis, King's College, Oxford, 1989.

[36] C.J.C.H. Watkins, P. Dayan, Q-learning, Mach. Learn. 8 (3) (1992) 279–292, https://doi.org/10.1007/BF00992698.