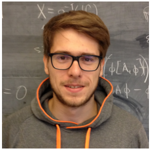
Reinforcement learning heterotic line bundle models

Magdalena Larfors

Uppsala University

String-Data 2020, CERN



Based on work with Robin Schneider (2003.04817)

see also ML–Schneider (1906.00392)

Heterotic Line Bundle Models

- Systematic semi-topological realisation of Standard like models (SLM)
  Anderson et al (1106.4804,1202.1757,1307.4787), Lukas et al (1309.0223, 1706.07688), ML–Passaro–Schneider (2010.09763).
- Large configuration space
  e.g. up to 10^{428} topological inequivalent CY 3-folds
  Demirtas–McAllister–Rios-Tascon (2008.01730)
- Topological computations resource-demanding (e.g. sequence-chasing) systematic constructions limited to CY with $h^{(1,1)} < 7$.

Well studied "lab" where we can (and should) try new search strategies

Heterotic Line Bundle Models

Well studied "lab" where we can try new search strategies

- Large configuration space
- Topological computations resource-demanding

Reinforcement Learning

- Solves large configuration spaces, e.g. AlphaZero win in Go
  Silver et al (Science, 2018)
- Learns with imperfect information, e.g. OpenAI win in DotA 2 (1912.06680)
- Learns type IIA intersecting brane models
  Halverson–Nelson–Ruehle (1903.11616)

Heterotic SLM via supervised and non-supervised machine learning:
Deen–He–Lee–Lukas (2003.13339), Mutter–Parr–Vaudrevange (1811.05993)

Heterotic Line Bundle Models

Well studied "lab" where we can try new search strategies

- Large configuration space
- Topological computations resource-demanding

Reinforcement Learning

- Solves large configuration spaces, e.g. AlphaZero win in Go
  Silver et al (Science, 2018)
- Learns with imperfect information, e.g. OpenAI win in DotA 2 (1912.06680)
- Learns type IIA intersecting brane models
  Halverson–Nelson–Ruehle (1903.11616)

Heterotic SLM via supervised and non-supervised machine learning:
Deen–He–Lee–Lukas (2003.13339), Mutter–Parr–Vaudrevange (1811.05993)

This talk in summary:

- build ML environments for heterotic line bundle models on CICY
- train A3C agents to find SLM

Results:

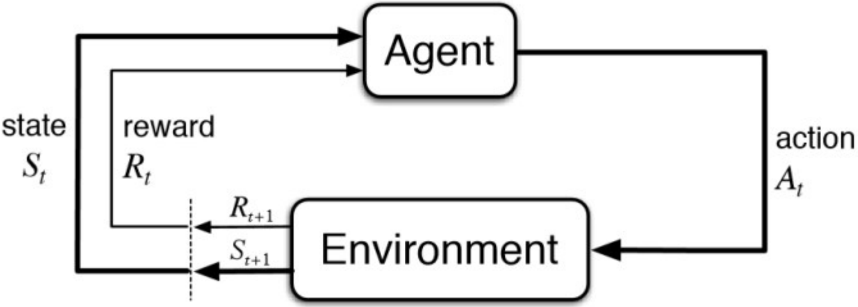
- consistently outperform random walkers
- after initial design phase, experiments run fast
- long-term strategy, seem to detect hidden structures, partly transferrable

Python packages:
`github.com/robin-schneider/gymCICY`; `pypi.org/project/pyCICY/`

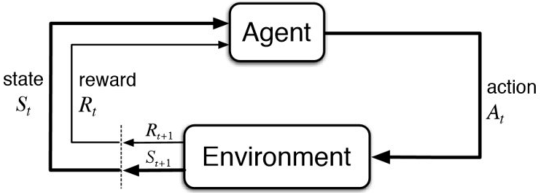
Outline

1. Reinforcement Learning
2. Agents
3. Environments
4. Experiments on fixed geometry
5. Transfer learning
6. Conclusions and Outlook

Reinforcement Learning



Reinforcement Learning



Agent takes action according to **policy** $\pi(a|s)$

$$\pi : \mathcal{S} \to \mathcal{A}$$

Goal: maximize **return** (accumulated future reward)

$$G_t = R_{t+1} + \gamma R_{t+2} + ... = \sum_k \gamma^k R_{t+k+1}, \qquad \text{with discount factor } \gamma \in [0, 1]$$

Vast state space $\rightsquigarrow$ use Neural Networks (deep reinforcement learning).

Agents

Asynchronous Advantage Actor Critic (A3C)

Mnih et al (1602.01783)

- use a NN as **Actor** to update policy $\pi(a_t|s_t;\theta_\pi)$,
- use a NN as **Critic** to update state value function $V(s_t;\theta_v) = \mathbb{E}[R_t|s = s_t]$,
- update parameters θ_π, θ_v using gradient descent from **Advantage** function,
- **Asynchronous** update of global parameters from local agents,
- can be trained on a single CPU,
- are stable and robust (on considered benchmarks).

Design and hyperparamters

cf. Halverson–Nelson–Ruehle (1903.11616).

Softmax $\pi(a_t|s_t;\theta_\pi)$, scalar $V(s_t;\theta_v)$ (see fig 4)
Hyperparameters tuned during design phase (see table 1)

Environments

Heterotic string compactification with three ingredients

1. Calabi Yau manifold M.
2. Line bundle sum $V = \oplus_{a=1}^{5} L_a$.
3. Freely acting discrete symmetry Γ for Wilson line.

Explored systematically $\longrightarrow$ 35 000 standard like models

Anderson et al (1106.4804,1202.1757,1307.4787).

For example:

$$\mathcal{M}_{5302} = \left[\begin{array}{c||ccc} 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \end{array}\right]^{6,30}_{-48} \quad \text{and } V = \begin{bmatrix} -1 & 0 & 0 & 0 & 1 \\ 4 & -3 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 1 & 1 & 0 & -2 & 0 \end{bmatrix}$$

and $|\Gamma| = 2$. There are a total of 6294 such models.

Environments

Heterotic string compactification with three ingredients

1. $\boxed{\text{Calabi Yau manifold } M}$ $\longrightarrow$ Fix to CICY with $h^{(1,1)} \geq 5$.
2. Line bundle sum $V = \oplus_{a=1}^{5} L_a$
3. Freely acting $\boxed{\text{discrete symmetry } \Gamma}$ for Wilson line. $\longrightarrow$ Fix to $|\Gamma| = 2$.

$$\mathcal{M}_{5302} = \left[\begin{array}{c||ccc} 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \end{array}\right]^{6,30}_{-48} \quad \text{and } \mathcal{M}_{5256} = \left[\begin{array}{c||cccc} 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 3 & 1 & 1 & 0 & 1 \end{array}\right]^{5,29}_{-48} .$$

Environments

Heterotic string compactification with three ingredients

1. Calabi Yau manifold M.
2. $\boxed{\text{Line bundle sum } V = \oplus_{a=1}^{5} L_a} \longrightarrow$ Environment.
3. Freely acting discrete symmetry Γ for Wilson line.

To qualify as a heterotic SLM, the bundle must satisfy physical constraints.

$\rightarrow$ Translate constraints to reward structure.

Environments: Reward Structure

(1. V is an $S(U(1)^5)$ bundle:	$c_1(V) = 0$	R_{max}=5)
2. Weak stability constraint:	$\mu(L_a) = 0, a = 1, ..., 5$	R_{max}=2
3. Three chiral families:	$-3\lvert\Gamma\rvert \leq \text{index}(L_a) \leq 0$	R_{max}= 10
4. Three chiral families.	$\text{ind}(V) \overset{!}{=} -3\lvert\Gamma\rvert$	R=10^2
5. Bianchi identity and Bogomolov bound:	$0 < c_2(V) \leq c_2(\mathcal{M})$	R=10^4
6. No Higgs triplets:	$\text{ind}(L_a \otimes L_b) < 0$	R=10^4
7. At least one Higgs doublet:	$h^2(\wedge^2 V) \geq 0$	R=10^6
8. No antigenerations:	$h^2(V) \overset{!}{=} 0$	R=10^7
9. V is slope stable:	$\mu(L_a) = 0$ in Kähler cone	—

Environments: Stacking

Precompile list L of n_{line} slope stable line bundles with $-3|\Gamma| \leq \text{index}(L_a) \leq 0$

Stack four of these, and adjust L_5 to satisfy $c_1(V) = 0$

$\longrightarrow$ *Constraints 1, 2, 3 are automatic*

States: The line bundle sum V. Hence $S_t \in \mathbb{Z}^{(5,\text{nProj})}$.

Actions: Pick $L_a \in L$ and replace one of L_{1-4}.

# of configurations: $n_{\text{conf}} = n_{\text{line}}^4$.

Example: $(\mathcal{M}_{5302},\ q_{\max} = 2,\ |\Gamma| = 2)$ gives $n_{\text{line}} = 2890$ and $n_{\text{conf}} \approx 7 \cdot 10^{13}$.

$$\begin{bmatrix} -1 & 0 & 0 & 0 & 1 \\ 2 & 0 & -1 & -1 & 0 \\ -2 & -1 & 0 & -2 & 5 \\ 0 & -1 & 0 & 2 & -1 \\ 0 & 2 & -2 & 2 & -2 \\ 2 & 2 & 1 & 0 & -5 \end{bmatrix} \xrightarrow{A_t} \begin{bmatrix} -1 & -2 & 0 & 0 & 3 \\ 2 & -2 & -1 & -1 & 2 \\ -2 & 0 & 0 & -2 & 4 \\ 0 & 2 & 0 & 2 & -4 \\ 0 & 0 & -2 & 2 & 0 \\ 2 & 2 & 1 & 0 & -5 \end{bmatrix}$$

Environments: Stacking

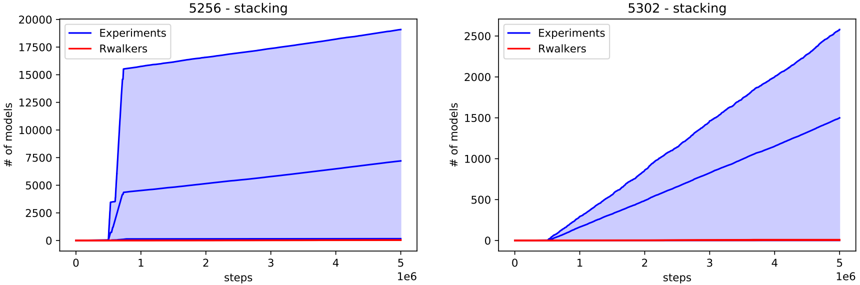


Figure: *Number of models found in two selected sets of stacking experiments (in blue) for the manifolds 5256 and 5302 plotted for comparison with each five random walkers (in red).*

Environments: Flipping

No precompiled list, instead "flip" individual charges in L_{1-4}

Still adjust L_5 to satisfy $c_1(V) = 0$

$\longrightarrow$ *Constraint 1 is automatic*

States: The line bundle sum V. Hence $S_t \in \mathbb{Z}^{(5,\text{nProj})}$.

Actions: Pick a charge q_i^j and add ± 1. Thus $A_t \in \{1, ..., 4 \cdot 2 \cdot \text{nProj}\}$.

# of configurations: $n_{\text{conf}} = (2 \cdot q_{\text{max}} + 1)^{4 \cdot h^{1,1}}$.

Example: ($\mathcal{M}_{5302}$, $q_{\text{max}} = 2$, $|\Gamma| = 2$) gives $n_{\text{conf}} \approx 5 \cdot 10^{16}$.

$$\begin{bmatrix} 1 & 1 & -1 & 0 & -1 \\ -1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & -1 & -1 \\ 1 & 1 & -1 & -1 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{A_t} \begin{bmatrix} 2 & 1 & -1 & 0 & -2 \\ -1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & -1 & -1 \\ 1 & 1 & -1 & -1 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

Flipping II

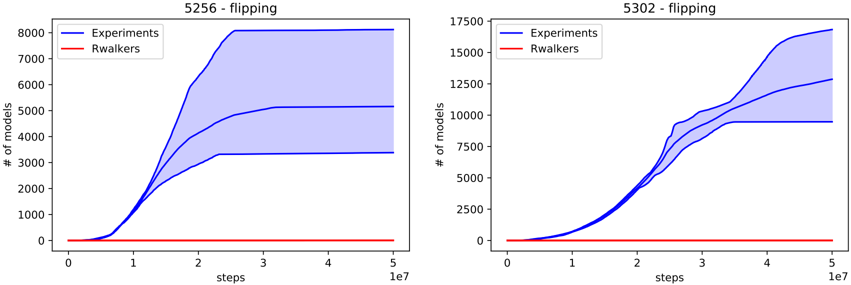


Figure: *Number of models found in two selected sets of flipping experiments (in blue) for the manifolds 5256 and 5302 plotted for comparison with each five random walkers (in red).*

Experiments on fixed geometry: Results

- *Stacking* close to human derived strategy of systematic scan
  - Runtime about 50 minutes on 32 cores
  - Moderately outperform random walker (factor 3-20)
  - Gets stuck in local minima $\longrightarrow$ low number of unique models.
- *Flipping* strategy different from systematic scan
  - Runtime about 3.5 hours on 32 cores
  - Rapid increase in performance followed by flattening at late times
  - Significantly outperform random walker (factor 300-1700)
  - Large number of unique models.

Experiments on fixed geometry

Experiments require significant design phase (see table 1)

But, eventually, hyperparameters are similar on the two explored geometries

$\boxed{\text{Has agent learned hidden structure?}}$

Successful flip experiments have slow start

$\boxed{\text{Speed up by smart initialization?}}$

Test agent on new manifold $\longrightarrow$ Transfer learning

Transfer learning:

Test 5256 flip agent on new manifold

- save the agent with the highest mean reward ($21 \cdot 10^6$ steps)
- run pretrained agent in evaluation mode (no learning)
- transfer run: allow last three layers to update (but low learning rate)

$$\mathcal{M}_{5452} = \left[\begin{array}{c||cccc} 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 2 & 0 & 0 & 0 \\ 1 & 0 & 0 & 2 & 0 \\ 3 & 1 & 1 & 1 & 1 \end{array}\right]^{5,29}_{-48} \quad \text{and } \mathcal{M}_{6890} = \left[\begin{array}{c||ccccc} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 2 \\ 1 & 0 & 0 & 2 & 0 & 0 \\ 4 & 1 & 1 & 1 & 1 & 1 \end{array}\right]^{5,37}_{-64}$$

Transfer learning: Results

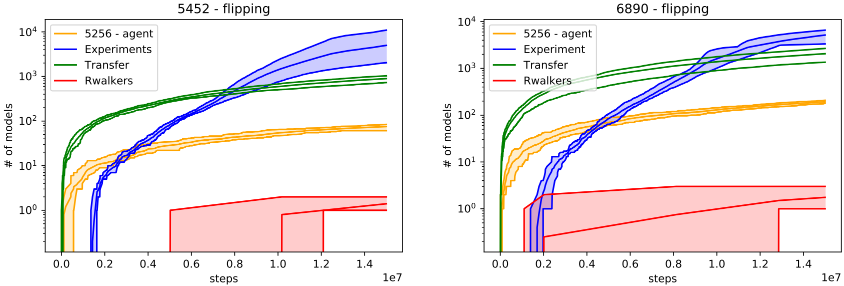


Figure: *Number of models found for selected sets of flipping experiments (in blue), random walkers (in red), pretrained agents (in yellow), and transfer agents (in green) on the manifolds 5452 and 6890. Note the logarithmic scale on the y-axis.*

Conclusions and Outlook

Results

- A3C agents successfully find heterotic SLM (outperform RW by 3-1700)
- Flipping environment $\rightsquigarrow$ new long term, "non-human" strategy
- Reproduce many models found in comprehensive scans (with less runtime)
- Computation time similar for $h^{(1,1)} = 5, 6$
- Transfer learning successful $\rightsquigarrow$ agent knowledge is partly general

Outlook

- More (exhaustive) experiments. Better choices for hyperparameters?
- More transfer experiments $\rightsquigarrow$ Explore models on CICY with $h^{(1,1)} \geq 7$
- Other RL developments: change networks or agents
- Explore other types of heterotic SLMs
- Combine with new physics/ML developments:
  CY Hodge numbers, formulas for line bundle cohomologies, CY metrics, ...

Conclusions and Outlook

Results

- A3C agents successfully find heterotic SLM (outperform RW by 3-1700)
- Flipping environment $\rightsquigarrow$ new long term, “non-human” strategy
- Reproduce many models found in comprehensive scans (with less runtime)
- Computation time similar for $h^{(1,1)} = 5, 6$
- Transfer learning successful $\rightsquigarrow$ agent knowledge is partly general

Outlook

- More (exhaustive) experiments. Better choices for hyperparameters?
- More transfer experiments $\rightsquigarrow$ Explore models on CICY with $h^{(1,1)} \geq 7$
- Other RL developments: change networks or agents
- Explore other types of heterotic SLMs
- Combine with new physics developments:
  analytical formulas for line bundle cohomologies, ML CY metrics, ...

Thank you for listening!

Extra slides: Implementation 1

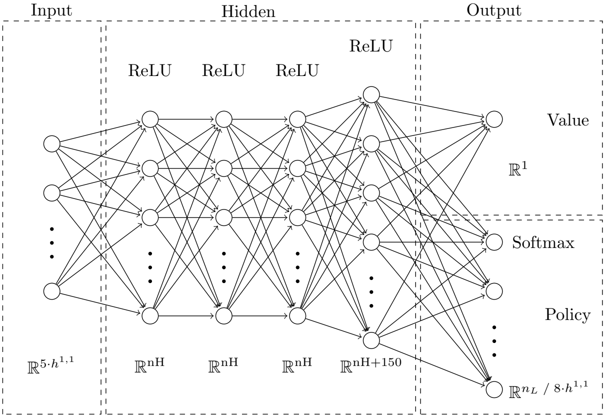


Figure: NN architecture

- implemented via ChainerRL
- Actor/Critic architecture differ only in output layer
- inspired by Halverson–Nelson–Ruehle (1903.11616)

Back to slide 8

Extra slides: Implementation 2

Hyperparameter tuning was not systematic.

Table: Hyperparameter values tuned in design phase:

	stack		flip	
parameter	5256	5302	5256	5302
t_{steps}	$5 \cdot 10^6$	$5 \cdot 10^6$	$50 \cdot 10^6$	$50 \cdot 10^6$
m_{steps}	30	30	200	300
lr	$5 \cdot 10^{-4}$	$5 \cdot 10^{-4}$	10^{-4}	10^{-4}
nH	100	100	75	100
β	1	0.1	1	1
γ	0.7	0.7	0.95	0.95

Common hyperparameters: local $t_{max} = 5$;
physics: $q_{max} = 2$, $|\Gamma| = 2$; RMSProp: $\alpha = 0.99$, $gc = 20$, $\epsilon = 0.0001$.

Back to slide 8