

## Notes: Machine learning flat directions

Bastien Duboeuf<sup>1</sup>, Camille Eloy<sup>1</sup> and Gabriel Larios<sup>2</sup>

<sup>1</sup> *ENS de Lyon, CNRS, LPENSL, UMR5672,  
69342, Lyon cedex 07, France*

<sup>2</sup> *Mitchell Institute for Fundamental Physics and Astronomy,  
Texas A&M University, College Station, TX, 77843, USA*

### Abstract

...

# Contents

1	Introduction	1
2	Supergravity setup	1
3	Numerical analysis	2
3.1	Gradient descent and local analysis . . . . .	2
3.2	Annealed Importance Sampling for polynomial symbolic regression . . . . .	2
4	Supergravity solutions	4
5	Conclusion	5

## 1 Introduction

## 2 Supergravity setup

- Context: 6d  $\text{AdS}_3 \times S^3$ , truncation to 3d, potential, conformal manifold.
- Choice of truncation from 32 to 13 to 5 variables.

**Potential** The 22 scalars of the theory can be parametrised by ( $22 = 32 - 10$ , with 10 scalars gauge fixed using translations in the gauge group)

- $m = \nu\nu^T \in \text{GL}(3, \mathbb{R})$  parametrizing the coset  $\text{GL}(3, \mathbb{R})/\text{SO}(3)$ ,
- $\phi$  a  $3 \times 3$  antisymmetric matrix,
- $\xi$  a  $3 \times 4$  matrix, and  $\xi^2 = \xi\xi^T$ ,
- a dilaton  $\tilde{\varphi}$ .

We can further restrict ourselves to a set of 13 scalars (see ref. [1]), with

$$\xi = \begin{pmatrix} 0 & 0 & 0 & x_1 \\ 0 & 0 & 0 & x_2 \\ 0 & 0 & 0 & x_3 \end{pmatrix}, \quad \phi = \begin{pmatrix} 0 & x_4 & x_5 \\ -x_4 & 0 & x_6 \\ -x_5 & -x_6 & 0 \end{pmatrix}, \quad \tilde{\varphi} = x_{13}$$

$$\nu = e^{(6x_7 + 3x_8 + \sqrt{3}x_9)/6} \begin{pmatrix} 1 & \frac{x_{10}}{\sqrt{2}} & \frac{x_{11}}{\sqrt{2}} + \frac{x_{10}x_{12}}{4} \\ 0 & e^{-x_8} & \frac{e^{-x_8}x_{12}}{\sqrt{2}} \\ 0 & 0 & e^{-(x_8 + \sqrt{3}x_9)/2} \end{pmatrix}. \quad (2.1)$$

Scalar potential from ref. [1]:

$$\begin{aligned}
V = & 4e^{-4\tilde{\phi}} + 2e^{-2\tilde{\phi}} \left[ -\text{tr}(m + m^{-1}) + \text{tr}(\phi m^{-1} \phi) - 2\text{tr}(\phi m^{-1} \xi^2) - 2\text{tr}(\xi^2) \right. \\
& - \text{tr}(\xi^2 m^{-1} \xi^2) + \frac{1}{2} \det(m^{-1}) \left( 1 - \text{tr}(\phi^2) - \text{tr}(\xi^4) + \text{tr}(\xi^2)^2 \right) \\
& + \frac{1}{2} \text{T}(m^{-1}(\xi^2 - \phi), (\xi^2 + \phi)m^{-1}, m + (\xi^2 + \phi)m^{-1}(\xi^2 - \phi) + 2\xi^2) \\
& \left. + \frac{1}{4} \text{T}(m^{-1}, m + (\xi^2 + \phi)m^{-1}(\xi^2 - \phi) + 2\xi^2, m + (\xi^2 + \phi)m^{-1}(\xi^2 - \phi) + 2\xi^2) \right],
\end{aligned} \tag{2.2}$$

where  $\text{T}(A, B, C) = \varepsilon_{mnp} \varepsilon_{qrs} A^{mq} B^{nr} C^{ps}$ .

As a first simplified example we have considered only the parameters  $x_1, x_2, x_4, x_8$  and  $x_{10}$ . The potential (2.2) becomes

$$\begin{aligned}
V = & \frac{1}{8} e^{-2x_8} \left[ 4 + 4x_1^4 + e^{4x_8} (2 + x_{10}^2)^2 (1 + x_2^4) \right. \\
& - 4e^{3x_8} (2 + x_{10}^2) \left( 2 - x_1^2 + \sqrt{2} x_1 x_{10} x_2^3 + x_2^4 + 2x_1 x_2 x_4 + x_4^2 - x_2^2 (1 - x_1^2 + x_4^2) \right) \\
& - 8e^{x_8} \left( 2 + x_1^4 + \sqrt{2} x_1^3 x_{10} x_2 - x_2^2 - 2x_1 x_2 x_4 + x_4^2 - x_1^2 (1 - x_2^2 + x_4^2) \right) \\
& + 4e^{2x_8} \left( 2 - 4x_1^2 + x_1^4 - 4x_2^2 + 4x_1^2 x_2^2 + x_2^4 + x_{10}^2 (1 + 3x_1^2 x_2^2) + 4x_4^2 + x_4^4 \right) \\
& \left. + 8\sqrt{2} x_{10} e^{2x_8} \left( x_1^3 x_2 + x_1^2 x_4 - x_2^2 x_4 + x_1 (x_2^3 - x_2 x_4^2) \right) \right].
\end{aligned} \tag{2.3}$$

### 3 Numerical analysis

#### 3.1 Gradient descent and local analysis

- Gradient descent: implementation, sampling, time needed, loss.
- Local PCA: graphs of number of patches with given local dimension, and with varying size of local patch, discuss optimal patch size (we want at least few points in each direction).
- Clustering: HDBScan (only parameter: minimal number of points in a cluster), and graphs of 3d slices of the 5d space.
- After PCA and clustering we know that the points after the gradient descent sample a three-dimensional manifold.

#### 3.2 Annealed Importance Sampling for polynomial symbolic regression

- We can convert the potential to a polynomial by converting the variables that appear in exponentials to logs.
- As the potential is a polynomial, the solutions satisfy polynomial equations. Discuss that fact that we could get polynomials directly by taking the gradient of the potential, but those will be too complicated to express the vacuum in a usable way. We search those using Annealed Importance Sampling.

- Annealed Importance Sampling: first explain general idea (construct density probability, role of temperature, links with Monte-Carlo), discuss  $\beta$  to control exploration and exploitation phases.
- Then more details: discuss hypotheses to compute the weights, choice of transformations, choice of  $\beta$ , choice of loss, prior, initial sampling, choice of representation for the polynomials.
- Discuss the fact that we allow float coefficients even though we know the coefficients will be only integers of square roots of integers?
- Analysis after AIS: select the best polynomials and do exploitation on the coefficients (without MC: we keep only the proposal of it betters the polynomial).
- Results for naive choices of parameters (numbers of iteration and particles, probabilities,  $\beta$ ) and motivate them (we want some exploration and then exploitation, not too long computations): for multiple runs we find multiple polynomials (good polynomial: after exploitation we convert the coefficients to integers and square roots, and recompute loss without regularisation, select with threshold). Quantify it nicely: success rates for each polynomials, and absolute number of success, failing rate. Total time needed, without cluster or fancy computers.

**Results** We have used the numerical analysis outlined in the previous sections to the search for extrema of the scalar potential (2.3). For 100 runs with parameters

$$n_{\text{iter}} = 1000, \quad n_{\text{particles}} = 1000, \quad \text{reg} = 10^3, \quad \beta = 10^{-10} + \left(\frac{i}{n_{\text{iter}}}\right)^5, \quad \sigma = 0.5 - \frac{0.45}{1 + \exp(10 - 20i/n_{\text{iter}})},$$

$$\text{sample size} = 10000, \quad p_{\text{add}} = 0, \quad p_{\text{remove}} = 0, \quad p_{\text{modify}} = 0.5, \quad p_{\text{multiply}} = 0.25, \quad p_{\text{divide}} = 0.25,$$
(3.1)

the code finds the 7 different following polynomials:

$$p_1 = -\sqrt{2}x_1 + \sqrt{2}x_1\tilde{x}_8 + x_2\tilde{x}_8x_{10} - \sqrt{2}x_2x_4\tilde{x}_8, \quad (3.2a)$$

$$p_2 = 2x_2 + 2x_2\tilde{x}_8 + \sqrt{2}x_1x_{10} + 2x_1x_4 - x_2\tilde{x}_8x_{10}^2, \quad (3.2b)$$

$$p_3 = 2x_2 - 2x_2\tilde{x}_8 + \sqrt{2}x_1\tilde{x}_8x_{10} + 2x_1x_4\tilde{x}_8 - 2x_2x_4^2\tilde{x}_8, \quad (3.2c)$$

$$p_4 = \sqrt{2}x_2 - \sqrt{2}x_2\tilde{x}_8 + \sqrt{2}x_1x_4 + x_1\tilde{x}_8x_{10} - x_2x_4\tilde{x}_8x_{10}, \quad (3.2d)$$

$$p_5 = -2x_1 - 2x_2x_4 - 2x_1x_4^2 + 2x_1\tilde{x}_8 + \sqrt{2}x_2x_{10} + x_1\tilde{x}_8x_{10}^2, \quad (3.2e)$$

$$p_6 = -2x_2x_4 - 2x_1x_4^2 + 2x_2x_4\tilde{x}_8 + \sqrt{2}x_2x_{10} - \sqrt{2}x_2\tilde{x}_8x_{10} + x_1\tilde{x}_8x_{10}^2, \quad (3.2f)$$

$$p_7 = -\sqrt{2}x_1^2x_4 + \sqrt{2}x_2^2x_4 + \sqrt{2}x_1x_2x_4^2 - x_1^2x_{10} - x_2^2x_{10}, \quad (3.2g)$$

where  $\tilde{x}_8 = e^{x_8}$ , with the distribution:

	$p_1$	$p_2$	$p_3$	$p_4$	$p_5$	$p_6$	$p_7$	$\emptyset$
Distribution	59%	25%	5%	9%	2%	4%	2%	13%

(3.3)

The sum of the percentages is higher than 100% because some runs find more than one polynomials. The code fails to find any polynomial in 13% of the time. The averaged time needed for a single run on a PC in  $\sim 1400$  s. **(ce: This timing is with analysis, it may be less without.)**

←

## 4 Supergravity solutions

- For each couple of candidate polynomials, we get the same expression for the solution, and it indeed defines a unique vacuum of the 3d SUGRA.
- Discussion of the vacuum: gauge group, Zham. metric, change of variables, 3d spectrum (and stability), spin 2 spectrum on  $S^3$ , one parameter seems to be a gauge parameter, uplift?

For every couples  $(p_i, p_j)$  of polynomials in eq. (3.2), the system

$$\begin{cases} p_i = 0, \\ p_j = 0, \end{cases} \quad (4.1)$$

gives the same solution

$$\begin{cases} x_1 = \frac{x_2}{\sqrt{2}} \frac{e^{x_8/2}}{e^{x_8} - 1} \left( -x_5 e^{x_8/2} + \sqrt{2 - 4e^{x_8} + e^{2x_8}(2 + x_{10}^2)} \right), \\ x_4 = \frac{e^{-x_8/2}}{\sqrt{2}} \sqrt{2 - 4e^{x_8} + e^{2x_8}(2 + x_{10}^2)}. \end{cases} \quad (4.2)$$

We checked that this defines a 3-parameter solution of three-dimensional half-maximal supergravity. Equivalently:

$$\begin{cases} \tilde{x}_8 = \frac{x_1^2 + x_2^2}{x_2^2 + (x_1 - x_2 x_4)^2}, \\ x_{10} = \sqrt{2} x_4 \frac{x_2^2 - x_1^2 + x_1 x_2 x_4}{x_1^2 + x_2^2}. \end{cases} \quad (4.3)$$

The solution preserves a  $U(1) \times U(1)$  gauged symmetry.

**Moduli space** The  $(x_1, x_2, x_4)$  moduli space is most nicely parametrised using the change of coordinates

$$x_1 = r \cos(\theta) \cos(\Phi), \quad x_2 = r \cos(\theta) \sin(\Phi) \quad \text{and} \quad x_4 = r \sin(\theta), \quad (4.4)$$

for which the Zamolodchikov metric reads

$$d^2 s_{\text{Zam.}} = -dr^2 - r^2 \left( d\theta^2 - r \cos(\theta) d\theta d\Phi + \sin(\theta) dr d\Phi + \frac{1}{2} (3 + r^2 - \cos(2\theta)) d\Phi^2 \right). \quad (4.5)$$

(ce: Tests other change of variables? Test choices of variables other than  $(x_1, x_2, x_4)$ ?)  $\Leftarrow$

**Bosonic spectrum** Vector fields:

$$\begin{aligned} m_{(1)} \ell_{\text{AdS}} : \quad & 0 \ [2], \quad -2 \ [5], \quad 2 \ [1], \\ & -1 \pm \sqrt{(1 + r^2)^2 - 2r^2 \cos(2\theta)} \ [2 + 2], \\ & 1 \pm \sqrt{1 + 4r^2 + r^4} \ [2 + 2]. \end{aligned} \quad (4.6)$$

The integers between square brackets indicate the multiplicity of each eigenvalue. The spectrum includes two massless vectors corresponding to the unbroken  $U(1) \times U(1)$  gauge symmetry, although in three dimensions they are non-propagating.

Scalars:

$$(m_{(0)}\ell_{\text{AdS}})^2 : \quad 0 \ [5], \quad 8 \ [1], \quad r^2 \ (4 + r^2) \ [8],$$

$$2r \left( 3r + r^3 \pm (2 + r^2) \sqrt{2 + r^2 - 2 \cos(2\theta)} - r \cos(2\theta) \right) \ [2 + 2]. \quad (4.7)$$

Gravitini:

$$m_{(3/2)}\ell_{\text{AdS}} : \quad \frac{1}{2} \left[ 1 \pm \sqrt{4 + 2r^2 + r^4 - 2r^2 \cos(2\theta)} \right] \ [4 + 4], \quad (4.8)$$

no SUSY enhancements other than  $r = 0$ .

No dependence on  $\Phi$ .

## 5 Conclusion

- Conclusion: good prospects of improvement to be able to access higher dimensional cases. Classify flat directions.
- Appendix with some details on the code?

## References

- [1] C. Eloy, G. Larios, and H. Samtleben, *Triality and the consistent reductions on  $AdS_3 \times S^3$* , *JHEP* **01** (2022) 055, [[arXiv:2111.01167](#)].