Learning Integrable Models.
From numerics to exact result

Evgeny Sobko

London Institute for Mathematical Sciences (LIMS)

based on Shailesh Lal, Suvajit Majumder, ES 2304.07247 (MLST) & 241x.xxxxx

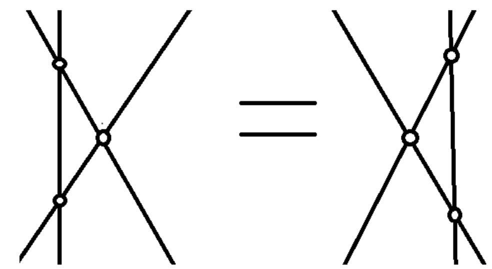
EuroStrings 2024 meets FPUK in Southampton

Plan

- Introduction. Integrable S-matrix Bootstrap and NNs
- Integrable Spin Chains. Overview
- Machine Learning the R-Matrix
- Extraction of exact results
- Work in progress

- Bootstrap approach : make focus on symmetries & self consistency
- S-matrix Bootstrap = analyticity + unitarity +crossing+…
- 2D Integrable Bootstrap = S-matrix Bootstrap + Factorisation $\to$ Exact solution Zamolodchikov,Zamolodchikov,1979

$$S_{12}(\theta_{12})S_{13}(\theta_1)S_{23}(\theta_2) = S_{23}(\theta_2)S_{13}(\theta_1)S_{12}(\theta_{12})$$



- Practically one have to solve a system of functional equations in the certain functional space. In general it is a non-convex problem.
- Neural Networks(NN) are good at minimazing non-convex functions. Strategy : parametrise unknown functions through the NNs, encode eqns into the loss function and train.

Spin-chains. Ultra quick overview

- Hilbert space : $\mathbb{V} = V_1 \otimes ... \otimes V_L, \quad V_i \sim V = \mathbb{C}^d$
- Hamiltonian : $H = \sum_{i=1}^{L} H_{i,i+1}, \qquad H_{L,L+1} \equiv H_{L,1}$
- Example: spin-1/2 XYZ magnet : $H = \sum_{i=1}^{L} \sum_{\alpha} J^\alpha S_i^\alpha S_{i+1}^\alpha$
- R-matrix operator : $R_{ij}(u) \;: V_i \otimes V_j \to V_i \otimes V_j$
- Yang-Baxter equation : $\underline{R_{ij}(u-v)R_{ik}(u)R_{jk}(v) = R_{jk}(v)R_{ik}(u)R_{ij}(u-v)}$
- Regularity : $\underline{R_{ij}(0) = P_{ij}}$
- Monodromy matrix : $\mathcal{T}_a(u) = R_{a,L}(u)R_{a,L-1}(u)\dots R_{a,1}(u)$
- Transfer matrix : $T(u) = \mathrm{tr}_a(\mathcal{T}_a(u))$

- $R\mathcal{T}\mathcal{T}$ relation : $R_{12}(u-v)\mathcal{T}_1(u)\mathcal{T}_2(v) = \mathcal{T}_2(v)\mathcal{T}_1(u)R_{12}(u-v)$

- Commutativity of transfer matrices : $[T(u), T(v)] = 0$

- Family of committing charges : $\log T(u) = \sum_{n=0}^{\infty} \mathbb{Q}_{n+1} \frac{u^n}{n!}$

$$\mathbb{Q}_{n+1} = \frac{d^n}{du^n} \log T(u)|_{u=0} = \frac{d^{n-1}}{du^{n-1}} \left(T^{-1}(u) \frac{d}{du} T(u) \right) \Big|_{u=0}$$

- Hamiltonian density : $\underline{H_{i,i+1} = R^{-1}_{i,i+1}(0)} \frac{d}{du} \underline{R_{i,i+1}(u)|_{u=0} = P_{i,i+1}} \frac{d}{du} \underline{R_{i,i+1}(u)|_{u=0}}$

- One can impose further restrictions on the level of R-matrix or Hamiltonian, such as unitarity $R_{12}(u)R_{21}(-u) = g(u)$, crossing, hermicity, symmetries etc.

- Given a solution for the Yang-Baxter equation, one can generate a whole family of solutions by acting with the following transformations :

  - $(\Omega \otimes \Omega)R(u)(\Omega^{-1} \otimes \Omega^{-1})$, $\Omega \in Aut(V)$: $\mathbb{Q}_n \to (\otimes^L \Omega)\mathbb{Q}_n(\otimes^L \Omega^{-1})$
  - rescaling of spectral parameter $u \to cu, \ \forall c \in \mathbb{C}$: $\mathbb{Q}_n \to c^{n-1}\mathbb{Q}_n$
  - $R(u) \to f(u)R(u), \quad f(0) = 1$
  - $PR(u)P, \ R(u)^T, \ PR^T(u)P \quad \to \quad P\mathcal{H}P, P\mathcal{H}^T P, \mathcal{H}^T$

There is no general way to solve YB

- There is no general way to find integrable Hamiltonian/solve YB. Roughly speaking there are three approaches :

  i. assume an symmetry (Lie algebra, Yangian, etc) and use it
  ii. directly solve functional or related dif equation R.S.Vieira ‘17‘19

  Drinfeld, Faddeev, Reshetikhin, Kulish,..

- iii. use boost operator $\mathcal{B}=\sum_{a=-\infty}^{\infty} aH_{a,a+1}$ to generate higher charges $\mathbb{Q}_{r+1}=[\mathcal{B},\mathbb{Q}_r]$

  impose commutativity $[\mathbb{Q}_i,\mathbb{Q}_j]=0$ and solve the resulting algebraic eqns.

  Marius de Leeuw et al ’19’20’20’21'23

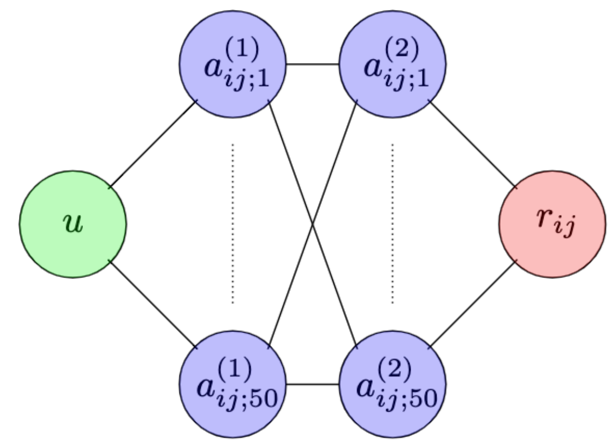
  De Leeuw, Pribytok, Ryan '19 classified all integrable 2d spin chains of difference form. They found 14 different classes in total, 8 of XYZ type and 6 of non-XYZ type :

$$H_{\mathrm{XYZ\,type}}=\begin{pmatrix} a_1 & 0 & 0 & d_1 \\ 0 & b_1 & c_1 & 0 \\ 0 & c_2 & b_2 & 0 \\ d_2 & 0 & 0 & a_2 \end{pmatrix}, \qquad H_{\mathrm{non-XYZ\,type}}=\begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ 0 & b_1 & b_3 & b_3 \\ 0 & c_1 & c_2 & c_3 \\ 0 & 0 & 0 & d_1 \end{pmatrix}$$

Machine Learning the R-Matrix

- We restrict the spectral parameter u to the interval (-1,1) and approximate each of real and imaginary parts of $R_{ij}(u)$ by NN with two hidden layers consisting of 50 neurons :

$\mathcal{R}_{ij}\left(u\right) = a_{ij}\left(u\right) + i\ b_{ij}\left(u\right)\ :$



$;\quad r_{ij} = \{a_{ij}, b_{ij}\}\ .$

- YB loss function :

$$\mathcal{L}_{YBE} = \sum_{u,v\in(-1,1)} ||\mathcal{R}_{12}(u-v)\mathcal{R}_{13}(u)\mathcal{R}_{23}(v) - \mathcal{R}_{23}(v)\mathcal{R}_{13}(u)\mathcal{R}_{12}(u-v)||$$

where $||A|| = \sum_{\alpha,\beta=1}^{n} |A_{\alpha\beta}|$ and u and v run over 20000 random points in (-1,1)

- Regularity : $\mathcal{L}_{reg} = ||\mathcal{R}\left(0\right) - P||$

Hamiltonian : $\mathcal{L}_H = ||P\frac{d}{du}\mathcal{R}(u)|_{u=0} - H||$

Hermicity : $\mathcal{L}_\dagger = ||H - H^\dagger||$

Total loss :

$$\mathcal{L} = \mathcal{L}_{YBE} + \mathcal{L}_{reg} + \lambda_H\mathcal{L}_H + \lambda_\dagger\mathcal{L}_\dagger + ...$$

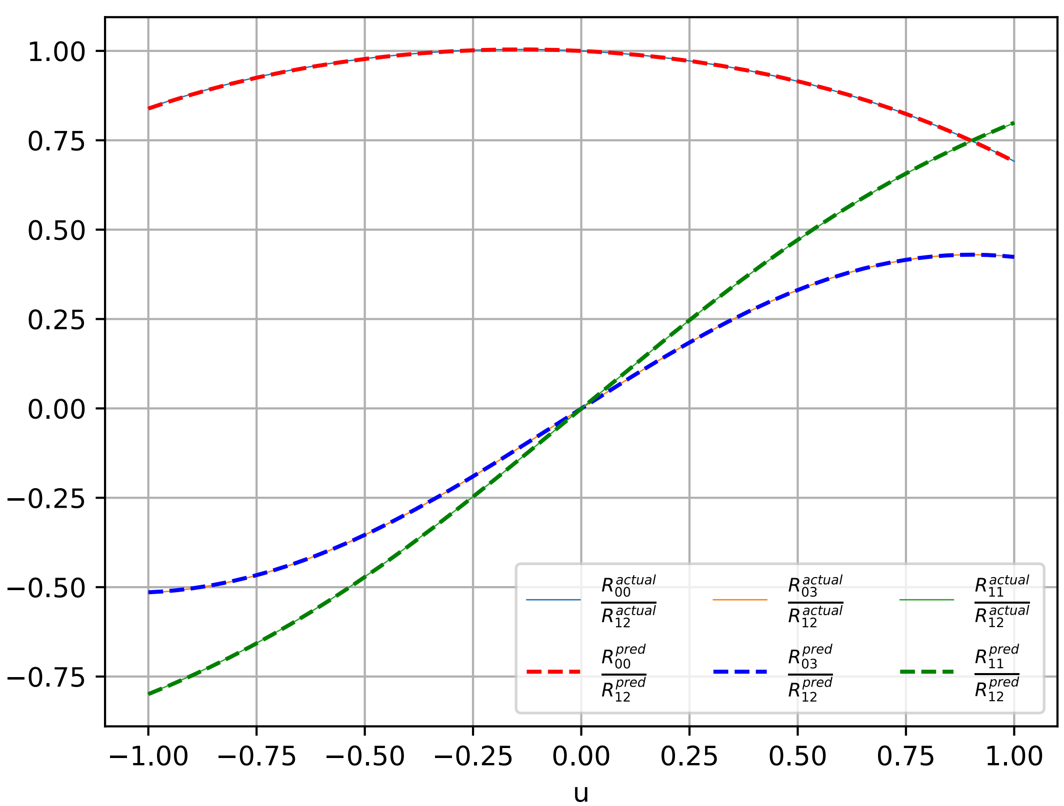
Hermitian XYZ model

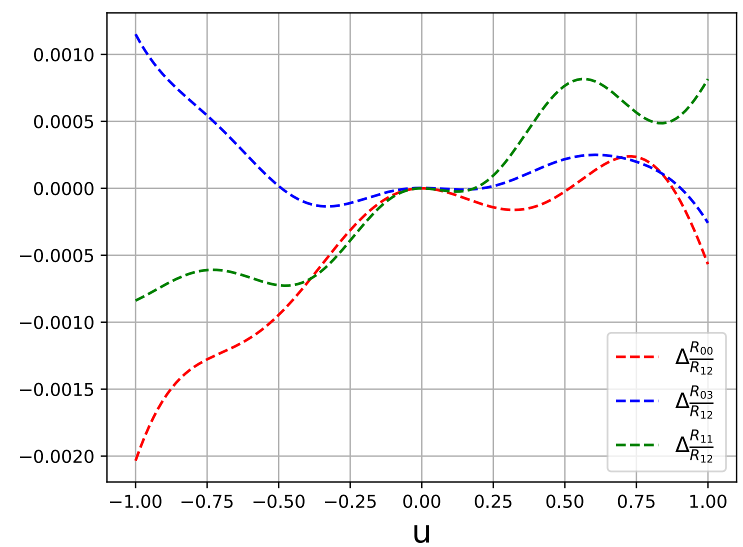
$$H_{XYZ}(J_x, J_y, J_z) = J_x S_1^x S_2^x + J_y S_1^y S_2^y + J_z S_1^z S_2^z = \begin{pmatrix} J_z & 0 & 0 & J_x - J_y \\ 0 & -J_z & 2 & 0 \\ 0 & 2 & -J_z & 0 \\ J_x - J_y & 0 & 0 & J_z \end{pmatrix}$$

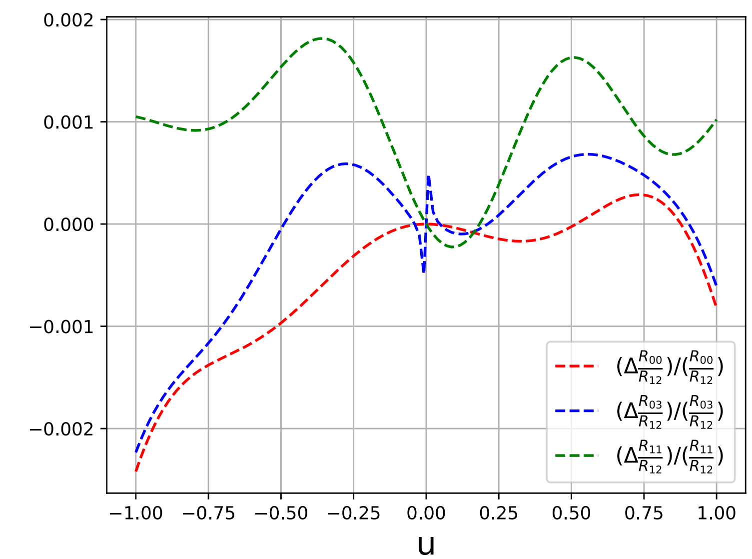
$$J_x = 1 + \sqrt{m}\ \mathrm{sn}(2\eta|m)/2$$
$$J_y = 1 - \sqrt{m}\ \mathrm{sn}(2\eta|m)/2$$
$$J_z = \mathrm{cn}(2\eta|m)\mathrm{dn}(2\eta|m)$$

$$R(u) = \begin{pmatrix} a & 0 & 0 & d \\ 0 & b & c & 0 \\ 0 & c & b & 0 \\ d & 0 & 0 & a \end{pmatrix}$$

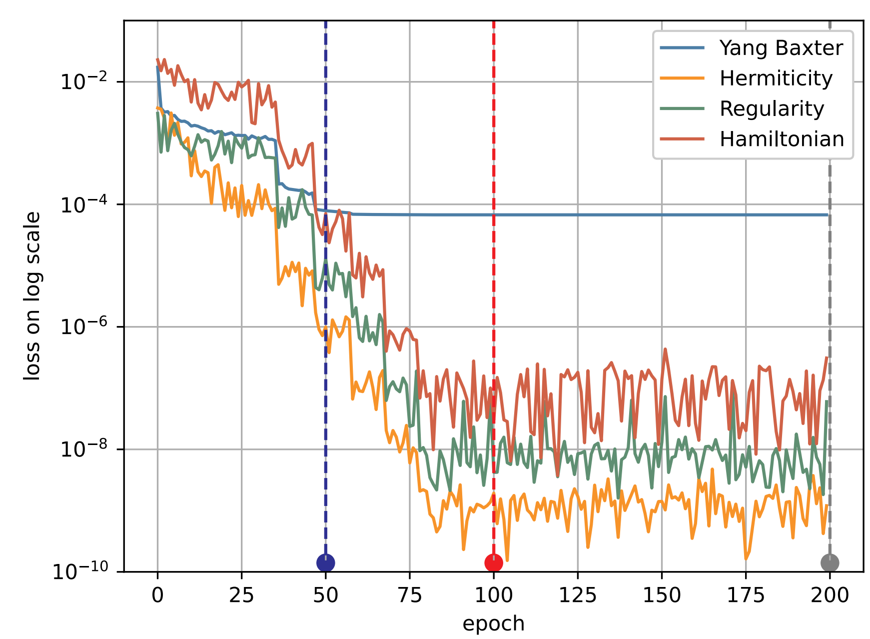
, where $a(u),\ b(u),\ c(u),\ d(u)$ are given in terms of Jacobi elliptic functions







$\eta = \pi/3, \quad m = 0.6$

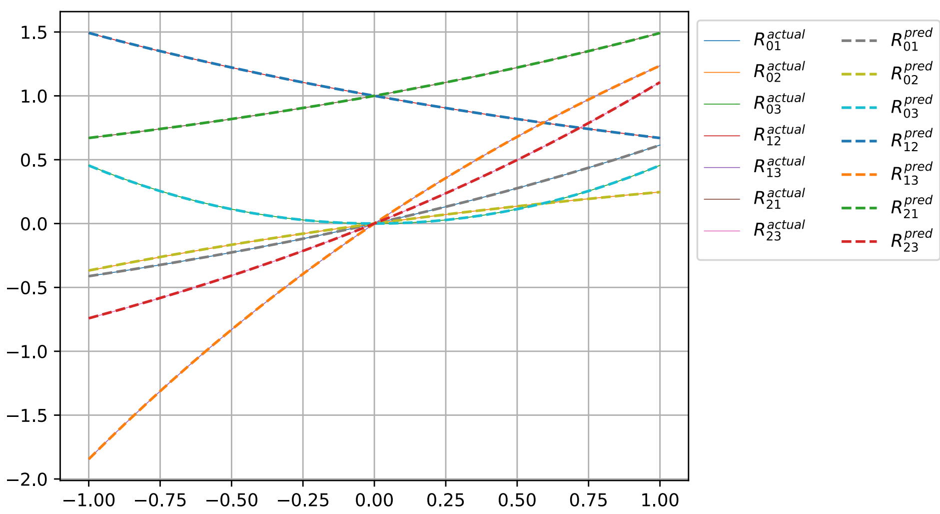


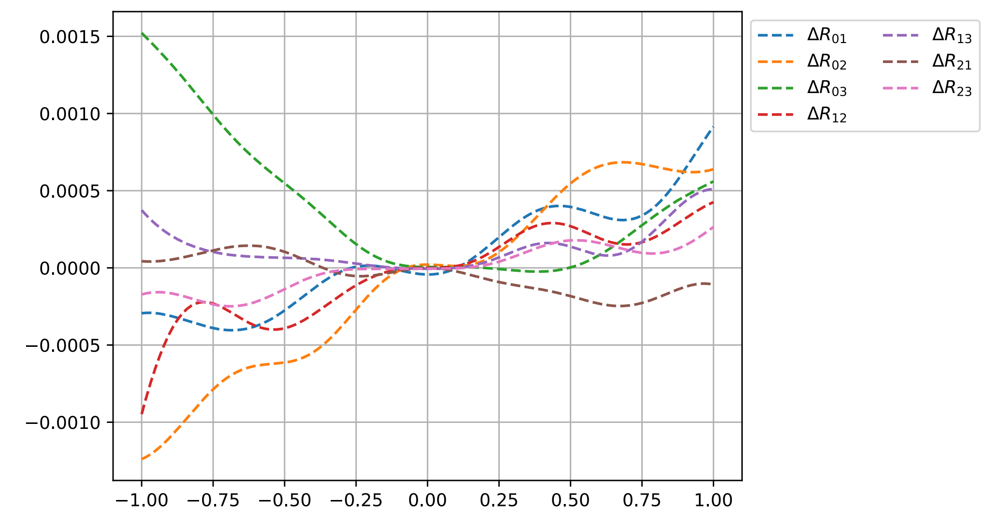
$$H_{class-1} = \begin{pmatrix} 0 & a_1 & a_2 & 0 \\ 0 & a_5 & 0 & a_3 \\ 0 & 0 & -a_5 & a_4 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$a_1 a_3 = a_2 a_4$$

$$R_{\text{class}-1}(u) = \begin{pmatrix} 1 & \frac{a_1(e^{a_5 u}-1)}{a_5} & \frac{a_2(e^{a_5 u}-1)}{a_5} & \frac{(a_1 a_3 + a_2 a_4)}{a_5^2}(\cosh(a_5 u) - 1) \\ 0 & 0 & e^{-a_5 u} & \frac{a_4(1-e^{-a_5 u})}{a_5} \\ 0 & e^{a_5 u} & 0 & \frac{a_3(1-e^{-a_5 u})}{a_5} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

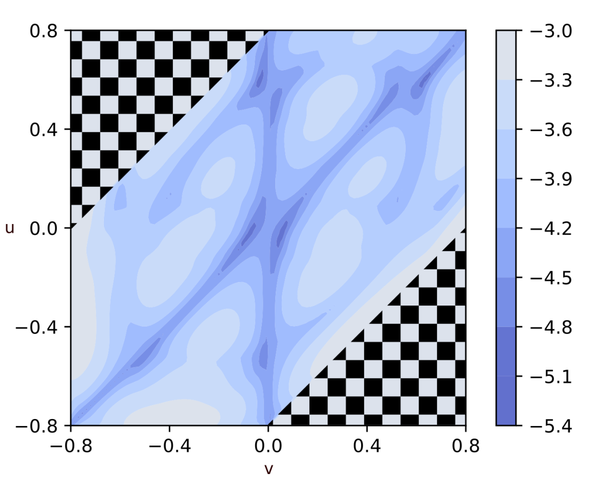
$a_1 = 0.5$
$a_2 = 0.3$
$a_3 = 0.9$
$a_4 = 1.5$
$a_5 = 0.4$



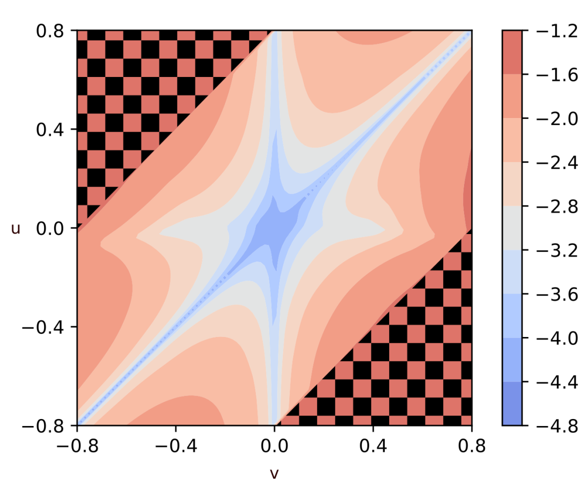


Similar $10^{-3} - 10^{-4}$ precision for all 14 classes

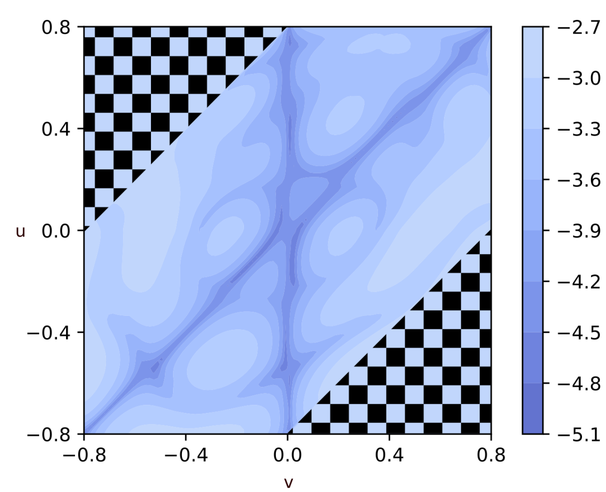
Without exact solution it is reasonable to define normalised YB : $\tilde{\mathcal{L}} = \frac{||\mathcal{R}_{12}(u-v)\mathcal{R}_{13}(u)\mathcal{R}_{23}(v) - \mathcal{R}_{23}(v)\mathcal{R}_{13}(u)\mathcal{R}_{12}(u-v)||}{||\mathcal{R}_{12}(u-v)\mathcal{R}_{13}(u)\mathcal{R}_{23}(v)||}$



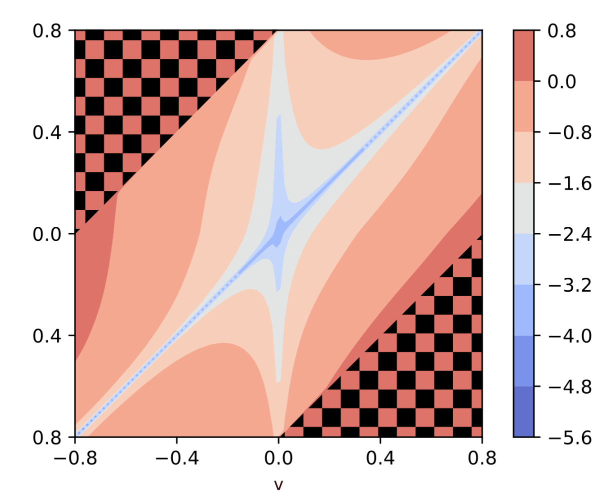
$H_{6v,1}$



$H_{6v,1} \quad non-int\ deform$



$H_{class-4}$



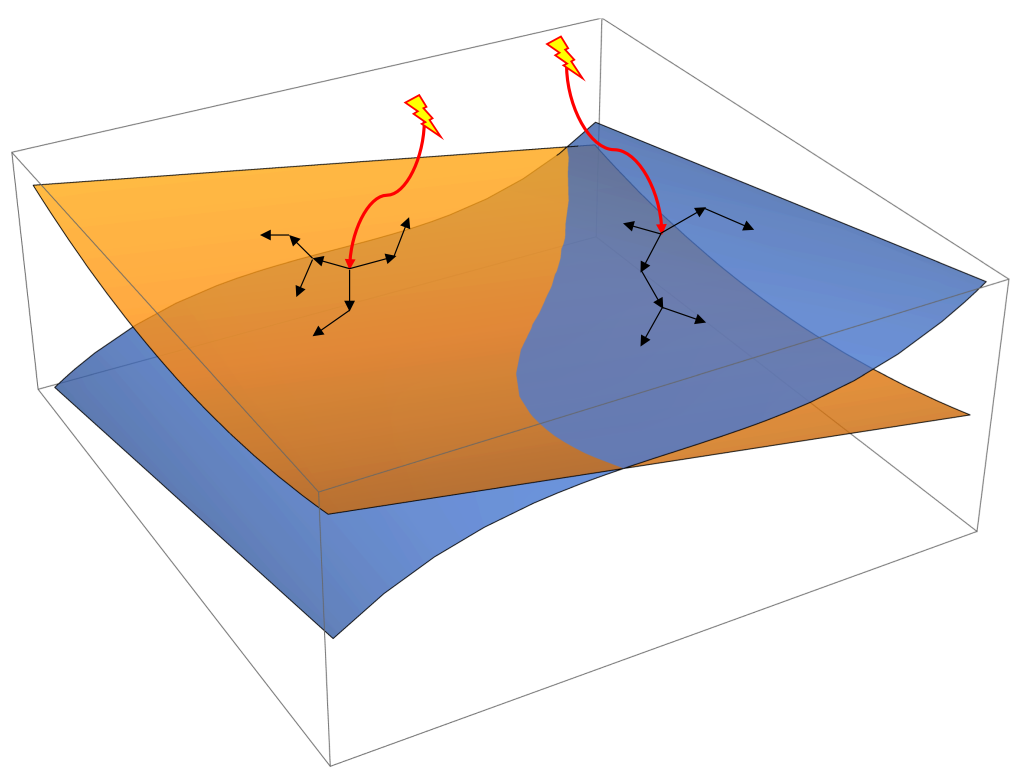
$H_{class-4} \quad non-int\ deform$

Explorer

- In explorer mode we i) don't provide a certain integrable Hamiltonian, instead we ask the NN to find an example of integrable model. It can be either general search without any restrictions at all or we can narrow the search to a certain class of hamiltonians imposing symmetry on the level of R-matrix or Hamiltonian, specifying ansatz etc. ii) Secondly we use *warm-start initialisation* and *repulsion* to find new integrable models in the vicinity of the already learnt one :

$$\mathcal{L}_{repulsion} = \exp\left(-||H - H_o||/\sigma\right)$$

- Explorer



Families of Hamiltonians as projective varieties

- Charges from Boost operator :

$$\mathbb{Q}_2 = \sum H_{a,a+1} \qquad \mathcal{B} = \sum_{a=-\infty}^{\infty} a H_{a,a+1} \qquad \mathbb{Q}_{r+1} = [\mathcal{B}, \mathbb{Q}_r] \qquad \mathbb{Q}_3 = \sum_{l=1}^{L} [H_{l,l+1}, H_{l+1,l+2}]$$

- Conjecture since 90th (no proof, no counterexamples) : $[\mathbb{Q}_2, \mathbb{Q}_3] = 0$ guarantees integrability

$$[\mathbb{Q}_2, \mathbb{Q}_3] = \sum_{n,l=1}^{L} [H_{n,n+1}, [H_{l,l+1}, H_{l+1,l+2}]] = 0 \qquad L = 2 + 3 - 1 = 4$$

- d^8 algebraic homogenous cubic equations depending on d^4 variables. All coefs are **integers**!

Families of integrable Hamiltonians $\longleftrightarrow$ $SL(d)$ factors of the primary decomposition of the ideal generated by the eqns $[\mathbb{Q}_2, \mathbb{Q}_3] = 0$

- d=2 : 256 eqns, 16 variables but the solution is extremely simple. There are just 14 families , 12 of them are linear and 2 have one/two quadratic relations!

$$H_{\text{XYZ type}} = \begin{pmatrix} a_1 & 0 & 0 & d_1 \\ 0 & b_1 & c_1 & 0 \\ 0 & c_2 & b_2 & 0 \\ d_2 & 0 & 0 & a_2 \end{pmatrix}, \qquad H_{\text{non}-\text{XYZ type}} = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ 0 & b_1 & b_3 & b_3 \\ 0 & c_1 & c_2 & c_3 \\ 0 & 0 & 0 & d_1 \end{pmatrix}$$

d=3 : 6561 eqns, 81 variables - already infeasible for existing methods

Our NN based approach to discover new integrable spin-chains analytically

Write an initial ansatz (with all desirable properties) fixing at least $d^2 - 1$ zeros

$$R_{initial} = \begin{pmatrix} * & * & * & * & * & * & * & * & * \\ 0 & * & * & * & * & * & * & * & * \\ 0 & 0 & * & * & * & * & * & * & * \\ 0 & * & 0 & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & * & * & * & * \\ 0 & 0 & * & 0 & 0 & 0 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & * & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * \end{pmatrix}$$

→ Solve YB with NN →

Get smth numerical from NN (usually with less number of vertices)

$$R_{from\ NN} \sim \begin{pmatrix} * & * & 0 & * & 0 & 0 & 0 & * & * \\ 0 & 0 & * & * & 0 & * & * & 0 & * \\ 0 & 0 & 0 & * & * & 0 & * & * & 0 \\ 0 & * & 0 & * & * & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & * & * & * & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & * & * & * & * \\ 0 & 0 & * & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & * & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * \end{pmatrix}$$

→ H=P R'(0) →

Typical precision for found Hamiltonians is 10^{-3}

$$H_{from\ NN} = PR'(0) \sim \begin{pmatrix} * & * & 0 & * & 0 & 0 & 0 & * & * \\ 0 & * & 0 & * & * & 0 & 0 & * & 0 \\ 0 & 0 & * & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & * & * & 0 & * & * & 0 & * \\ 0 & 0 & 0 & 0 & * & * & * & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & * & 0 & 0 & * \\ 0 & 0 & 0 & * & * & 0 & * & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * \end{pmatrix}$$

↓ Use it as a starting point to solve alg eqns [Q2,Q3]=0 with precision $\sim 10^{-8}$

Refined Hamiltonian with precision $\sim 10^{-8}$ or higher if necessary

← Generate new solutions of [Q2,Q3]=0 around the learnt one making small perturbations

Hundreds/thousands of Hamiltonians from the same family with precision $\sim 10^{-8}$

← Use these Hamiltonians to extract exact polynomial relations between $\{h_{ij}\}$

Exact polynomial relations between $\{h_{ij}\}$

↓ Plug them into the [Q2,Q3]=0 and solve

Exact Hamiltonian!

Typical example :

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2+\frac{h[3,5]^2}{h[3,7]}-2h[3,7]-h[4,4] & 0 & -h[3,7] & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3+\frac{2h[3,5]^2}{h[3,7]}-3h[3,7]-2h[4,4] & 0 & h[3,5] & 0 & h[3,7] & 0 & 0 \\ 0 & -h[3,7] & 0 & h[4,4] & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & h[3,5] & 0 & 1+\frac{h[3,5]^2}{h[3,7]}-2h[3,7] & 0 & h[3,5] & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2+\frac{h[3,5]^2}{h[3,7]}-2h[3,7]-h[4,4] & 0 & -h[3,7] & 0 \\ 0 & 0 & h[3,7] & 0 & h[3,5] & 0 & -1+h[3,7]+2h[4,4] & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -h[3,7] & 0 & h[4,4] & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Work in progress and future directions

- We proposed a novel ML based approach to discover new integrable models. First i) we should impose desirable restrictions at the level of R-matrix and Hamiltonian, ii) then NN will find an numerical solution and iii) using auxiliary system of algebraic equations coming from the method of the Boost operator we extract exact analytical formulas.

- Interesting observation : all algebraic varieties found so far are rational!

- Speed up the training making “batching of equations”. System of eqns is overdetermined so we need to train just a small fraction choosing them randomly (and wisely!). Check how far we can go

- S-matrices for 2d QFT. The only new ingredient is to fix analytical properties in rapidity plane (poles, cuts)

- R-matrices of nondifference form. Conceptually similar, technically more complicated.

- S-matrices of integrable strings on AdS background : nondiference form + analyticity

- gYB appears in many other problems : long-range interaction, knot invariants, quantum gates etc

Thank you for your attention!