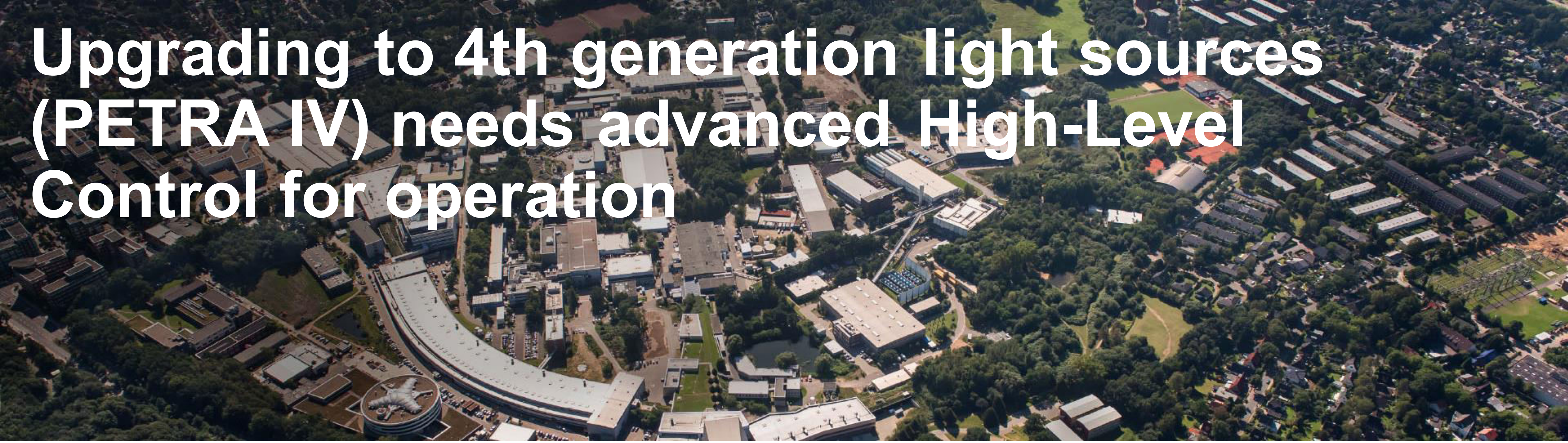
Physics-Based Deep Neural Networks for Beam Dynamics in Charged Particle Accelerators

AMALEA Helmholtz Innovation Pool project

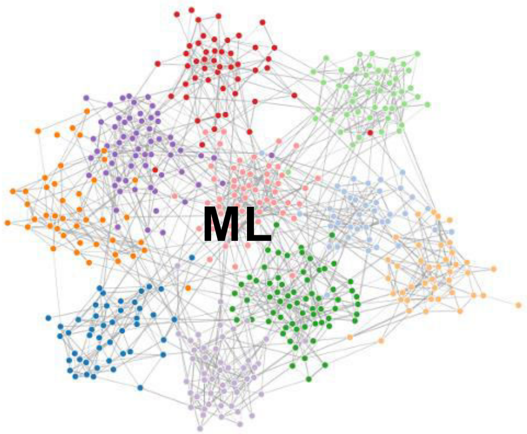
Andrei Ivanov
Hamburg, 18.06.2020

Upgrading to 4th generation light sources (PETRA IV) needs advanced High-Level Control for operation



While PETRAIII is stable in operation and don’t require ML, the 4th generation rings (PETRAIV) faces certain difficulties:

- The resolution gap between 1-10 nanometers
- Reliability demands grow (95% -> 99%)
- Machines are more sensitive with larger number of components
- High nonlinearities



Big Data is all about finding correlations

The model will be only as good or as bad as the data you have

- well-posed problem
- high performance computing
- ubiquitous data

https://xkcd.com/1838/

In the real-world, application of ML is more difficult than in research:

- learning on the real system from **limited samples**
- **high-dimensional** continuous state and action spaces.
- **safety constraints** that should never or at least rarely be violated
- tasks that may be **partially observable**, alternatively viewed as non-stationary or stochastic
- system operators who desire **explainable** policies and actions
- inference that must happen in **real-time** at the control frequency of the system

Novel approach for constructing deep neural networks for beam dynamics

with the following key features:

- accurate simulation of dynamics **without training**
- model **fine-tuning** with limited measurements

ECAI 2020

oral presentation at the European Conference on Artificial Intelligence

Novel approach for constructing deep neural networks for beam dynamics

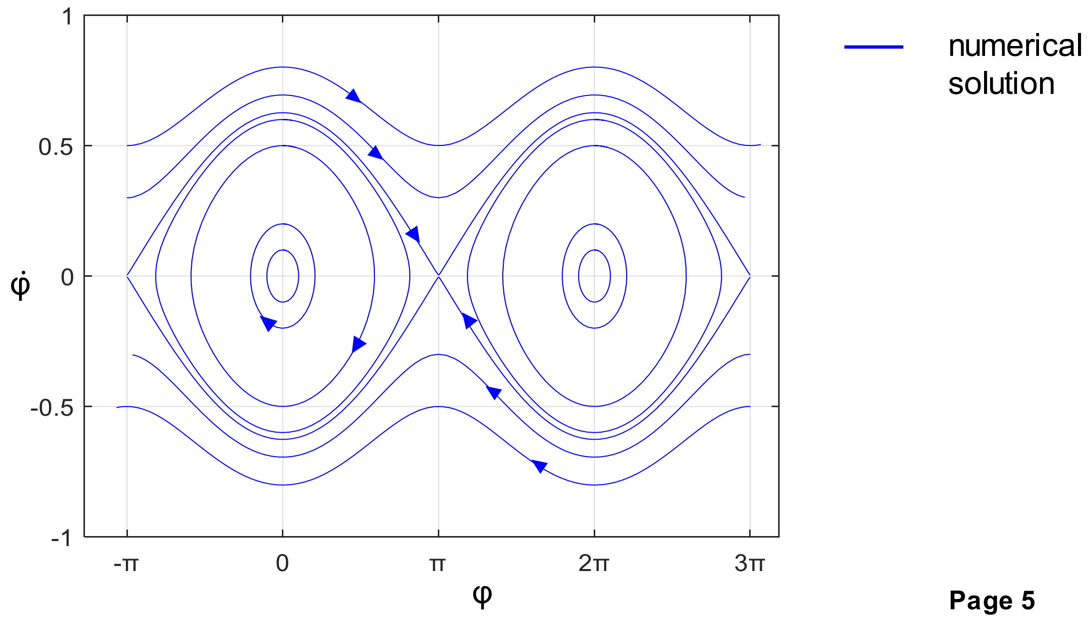
with the following key features:

- accurate simulation of dynamics **without training**
- model **fine-tuning** with limited measurements

oral presentation at the European Conference on Artificial Intelligence

The key idea: If the dynamics of a system approximately follows a given differential equation, the Taylor mapping technique can be used to initialize the weights of a polynomial neural network

Pendulum oscillation: $\ddot{\varphi} = -\omega^2 \sin\varphi$



Novel approach for constructing deep neural networks for beam dynamics

with the following key features:

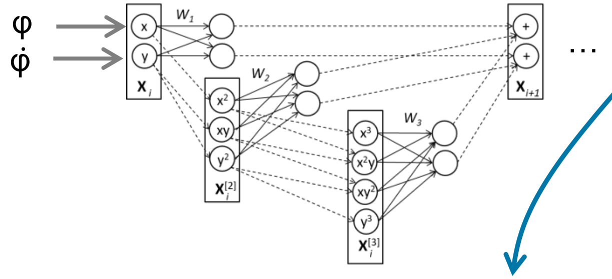
- accurate simulation of dynamics **without training**
- model **fine-tuning** with limited measurements

ECAI 2020

oral presentation at the European Conference on Artificial Intelligence

The key idea: If the dynamics of a system approximately follows a given differential equation, the Taylor mapping technique can be used to initialize the weights of a polynomial neural network

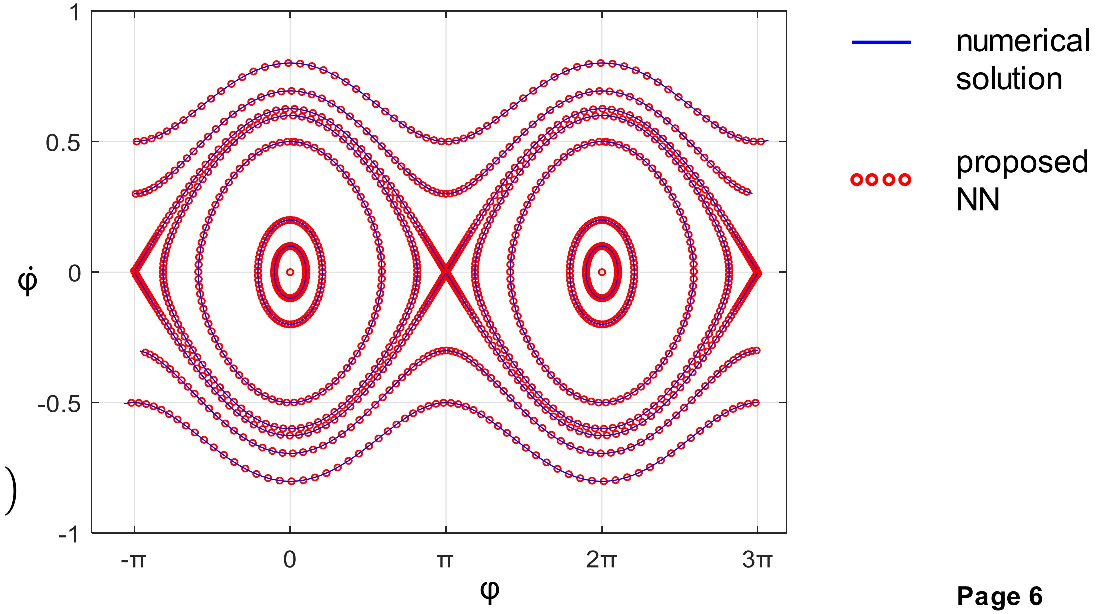
Pendulum oscillation: $\ddot{\varphi} = -\omega^2 \sin\varphi$



$$W_1 = \begin{pmatrix} 1 & 0.099 & -7.64E{-}06 & 0 \\ 0 & 1.000 & -1.54E{-}04 & 0 \end{pmatrix}$$

$$W_2 = \begin{pmatrix} 0\ 0\ 0\ 0\ 0\ 0\ -1.58E{-}05\ 0\ 6.11E{-}10\ 0 \\ 0\ 0\ 0\ 0\ 0\ 0\ -4.80E{-}04\ 0\ 2.47E{-}08\ 0 \end{pmatrix}$$

$$W_3 = \begin{pmatrix} 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 2.46E{-}05\ 0\ -2.28E{-}09\ 0\ 7.61E{-}10\ 5.87E{-}14\ 0\ -1.95E{-}14\ 0 \\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 9.95E{-}04\ 0\ -1.15E{-}07\ 0\ 3.84E{-}08\ 3.56E{-}12\ 0\ -1.18E{-}12\ 0 \end{pmatrix}$$



Novel approach for constructing deep neural networks for beam dynamics

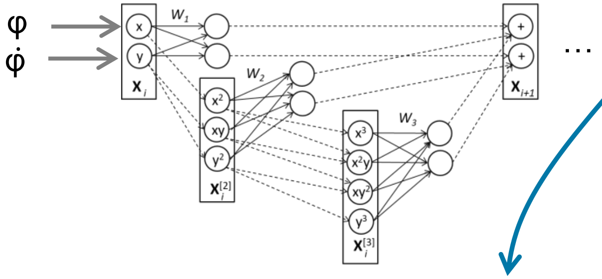
with the following key features:

- accurate simulation of dynamics **without training**
- model **fine-tuning** with limited measurements

oral presentation at the
European Conference
on Artificial Intelligence

The key idea: If the dynamics of a system approximately follows a given differential equation, the Taylor mapping technique can be used to initialize the weights of a polynomial neural network

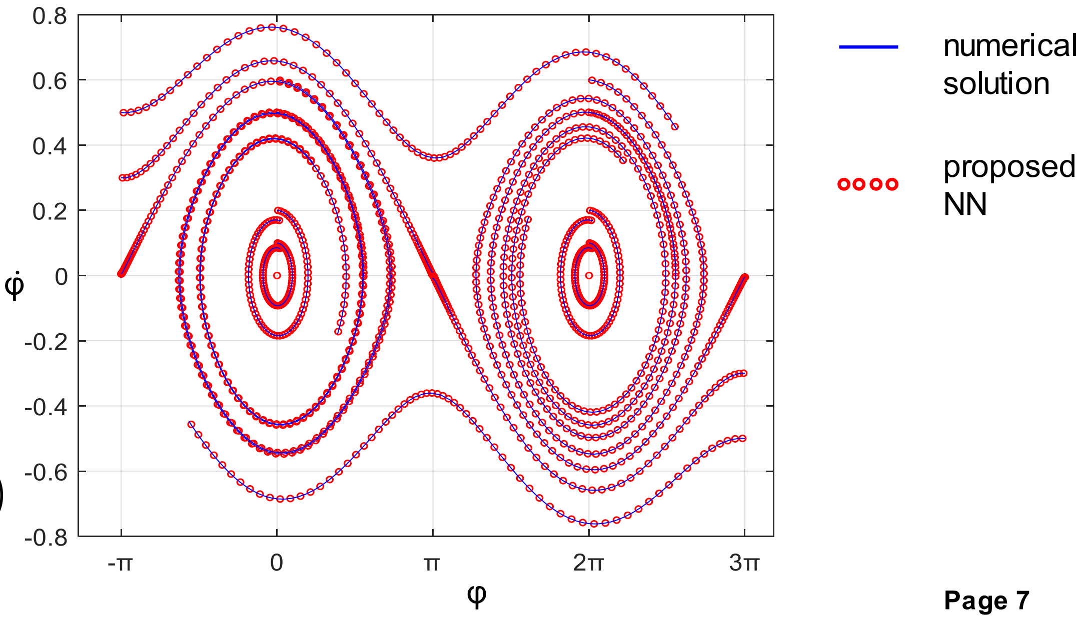
Pendulum oscillation: $\ddot{\varphi} = -\omega^2 \sin\varphi \boxed{- 2\gamma\dot{\varphi}}$



$$W_1 = \begin{pmatrix} 1 & 0.099 & -7.79E-06 & 0 \\ 0 & 0.990 & -1.55E-04 & 0 \end{pmatrix}$$

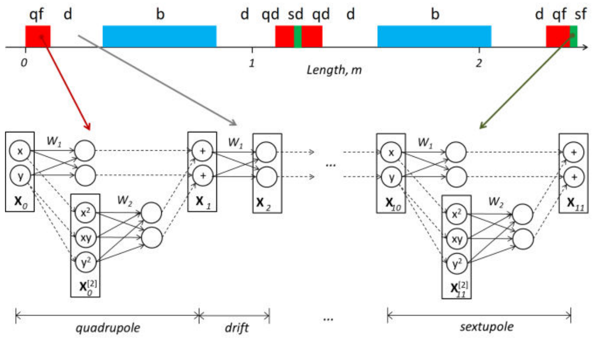
$$W_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & -1.63E-05 & 0 & 6.36E-10 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -4.89E-04 & 0 & 2.54E-08 & 0 \end{pmatrix}$$

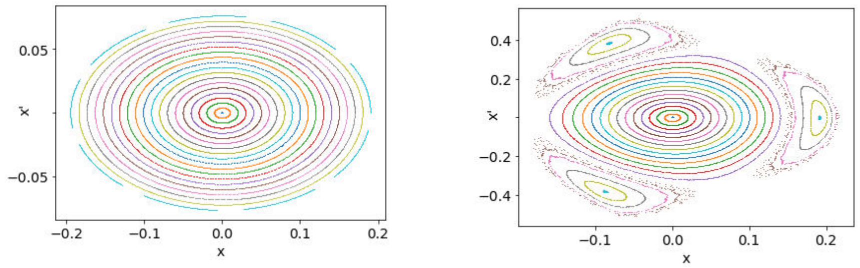
$$W_3 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2.56E-05 & 0 & -2.39E-09 & 0 & 7.99E-10 & 6.23E-14 & 0 & -2.07E-14 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.02E-03 & 0 & -1.19E-07 & 0 & 3.99E-08 & 3.74E-12 & 0 & -1.24E-12 & 0 \end{pmatrix}$$



Translating lattices of the storage rings into deep neural networks

FODO: neural network with 12 layers represents **resonance**

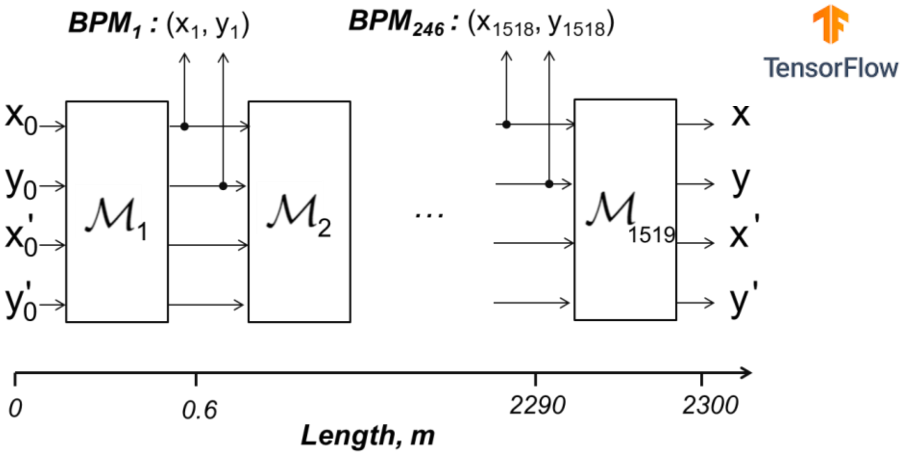




Initialized NN accurately represents the parametric dependency of dynamics on magnet strength, such as the appearance of a third-integer resonance

PETRAIII: deep neural network with <u>1519 layers</u> represents ideal lattice with fair accuracy

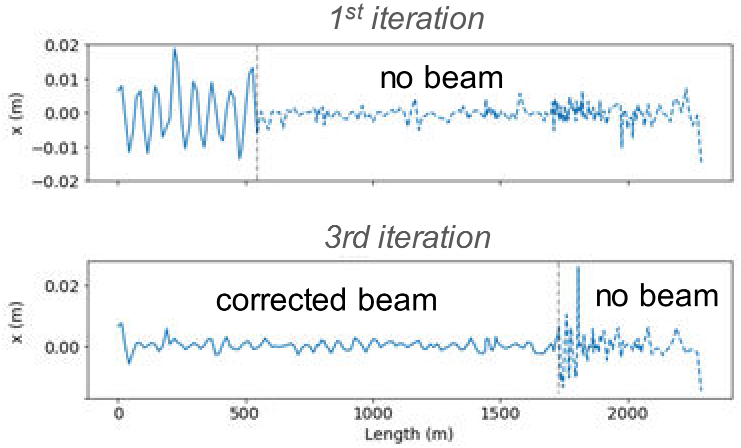
- **2,3** km length with **1519** magnets
- **210** horizontal and 194 vertical correctors
- **246** beam position monitors



One-short learning of PETRAIII in experiments

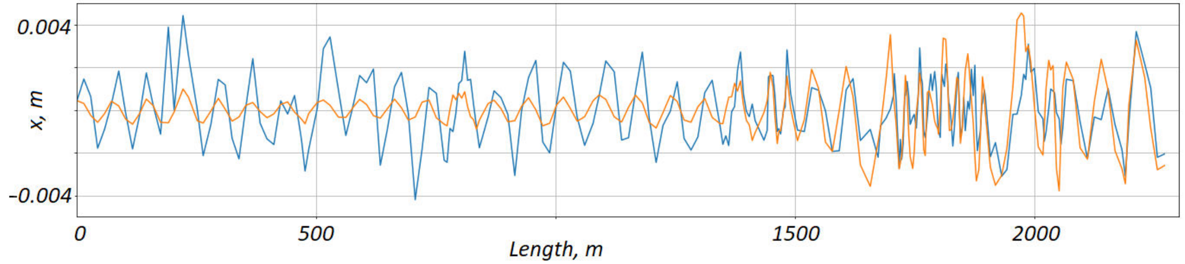
Beam threading

1. All corrector magnets are switched off
2. Beam is able to travel through only a part of the ring
3. **Neural Network predicts** an optimal control policy for beam propagation



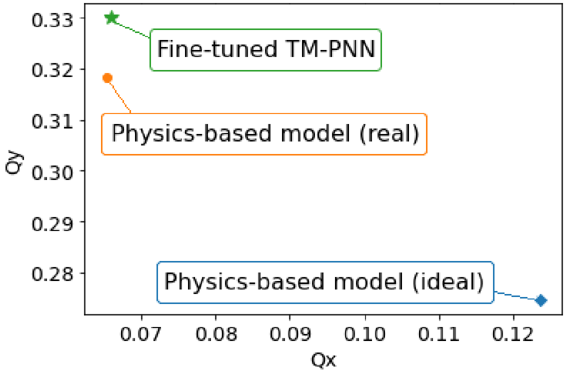
Tune recovering

1. Tune is the main multi-turn frequency of beam oscillation in the storage ring
2. The affected magnets cause the tune change from the designed values.
3. **Neural Network** is trained with only a **single-turn measurement** and estimates tunes with 95% accuracy.



— beam oscillation in ideal lattice
— beam oscillation in imperfect lattice

training NN with a single-turn measurements



95% accuracy of the multi-turn prediction

To handle problem of limited observations we implemented special regularization methods

Regularization aims to reduce the number of free parameters (weights) of the NN to ovoid overfitting. The traditional methods (L1-L2 norms) do not reflect physics and just try to reduce the absolute magnitude of the weights during training.

Symplectic regularization

For Hamiltonian systems representing single-particle beam dynamics, the symplectic property can be used. The Hamiltonian structure of each layer is preserved for all new inputs which has a large impact on generalization.

$$W_1 = \begin{pmatrix} {w_1}^{11} & {w_1}^{12} \\ {w_1}^{21} & {w_1}^{22} \end{pmatrix}, \quad W_2 = \begin{pmatrix} {w_2}^{11} & {w_2}^{12} & {w_2}^{13} \\ {w_2}^{21} & {w_2}^{22} & {w_2}^{23} \end{pmatrix} \xrightarrow{\text{symplectic property}}$$

$$\begin{aligned} &w_1^{11}w_1^{22} - w_1^{12}w_1^{21} - 1 = 0, \\ &w_2^{11}w_2^{23} - w_2^{13}w_2^{21} = 0, \\ &w_2^{12}w_2^{23} - w_2^{13}w_2^{22} = 0, \end{aligned} \qquad \begin{aligned} &w_1^{11}w_2^{22} - w_1^{21}w_2^{12} + 2w_1^{22}w_2^{11} - 2w_1^{12}w_2^{21} = 0, \\ &w_1^{22}w_2^{12} - w_1^{12}w_2^{22} + 2w_1^{11}w_2^{23} - 2w_1^{21}w_2^{13} = 0. \end{aligned}$$

QUBO-based regularization (quadratic unconstrained binary optimization)

Since physical systems that are described by ODEs often lead to sparse weights, this should be preserved during training:

$$W_1 = \begin{pmatrix} 1 & 0.099 & -7.64\text{E}{-06} & 0 \\ 0 & 1.000 & -1.54\text{E}{-04} & 0 \end{pmatrix}$$

$$W_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & -1.58\text{E}{-05} & 0 & 6.11\text{E}{-10} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -4.80\text{E}{-04} & 0 & 2.47\text{E}{-08} & 0 \end{pmatrix}$$

$\xrightarrow{\text{QUBO problem}}$

Problem: fit weights with data and maximize number of zero elements

Solution: combinatorial problem that can be solved with Quantum Annealers

Results

01 Novel architecture of deep NN incorporating physical knowledge from ODEs

Deep NN without training simulates the dynamics and a possibility to fine-tune the weights on limited experimental data. This is in sharp contrast with most of the existing works on learning of dynamical systems, where complex neural architectures are trained on large datasets

02 New regularization methods are suggested

symplectic regularization	→	restriction to Hamiltonian systems	→	easy to implement
QUBO-based regularization	→	general purpose	→	combinatorial optimization (Quantum Annealers)

03 The NN is validated on both simulation of PETRAIV and experiments in PETRAIII

The proposed NN allows fine-tuning with single trajectory (one-short learning) of noisy measurements.

04 Beyond accelerator physics

Since ODEs and\or PDEs generally arise in physical problems, the proposed NN may be helpful for solving and speeding up real physical problems in various domains.

ECAI 2020

oral presentation at the European Conference on Artificial Intelligence (Accepted)

APS physics

Paper in Physical Review AB (Under Review)

Thank you

Contact

DESY. Deutsches
Elektronen-Synchrotron

Andrei Ivanov
andrei.ivanov@desy.de

www.desy.de