Poster/Demo Flash Talks

4th ICFA Beam Dynamics Mini-Workshop on Machine Learning Applications for Particle Accelerators

Distance Preserving Machine Learning for Uncertainty Aware Accelerator Capacitance Predictions

Steven Goldenberg, Malachi Schram, Kishansingh Rajput, Thomas Britton,
Chris Pappas, Majdi Radaideh, Jared Walden, Dan Lu, Sarah Cousineau

RESEARCH DESCRIPTION

- Capacitors in High-Voltage Converter Modulators (HVCMs) degrade over time which causes significant downtime
- Extensive simulation data based on available non-invasive sensor data is available
- Modeling HVCM capacitance values with Uncertainty Quantification is necessary to provide a reliable early indicator of failure

SNS DOWNTIME BY SYSTEM

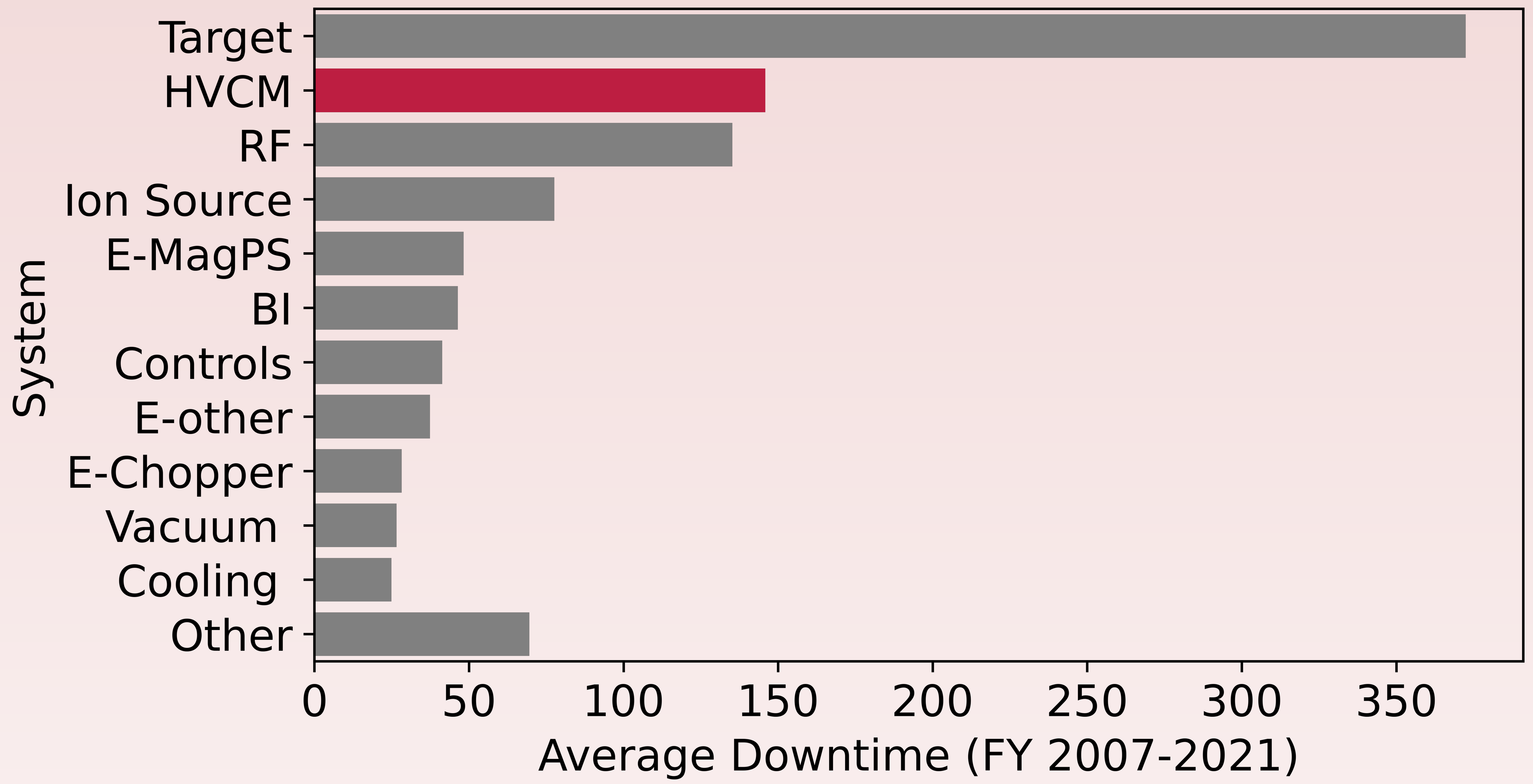


Figure 1: Average downtime by system in the Spallation Neutron Source (SNS) at the Oak Ridge National Lab.

Oak Ridge National Laboratory

Jefferson Lab

MODEL PERFORMANCE

- <1% in-distribution error
- <3.5% out-of-distribution error
- Increased model uncertainty for out-of-distribution samples

MODEL PREDICTION ERRORS

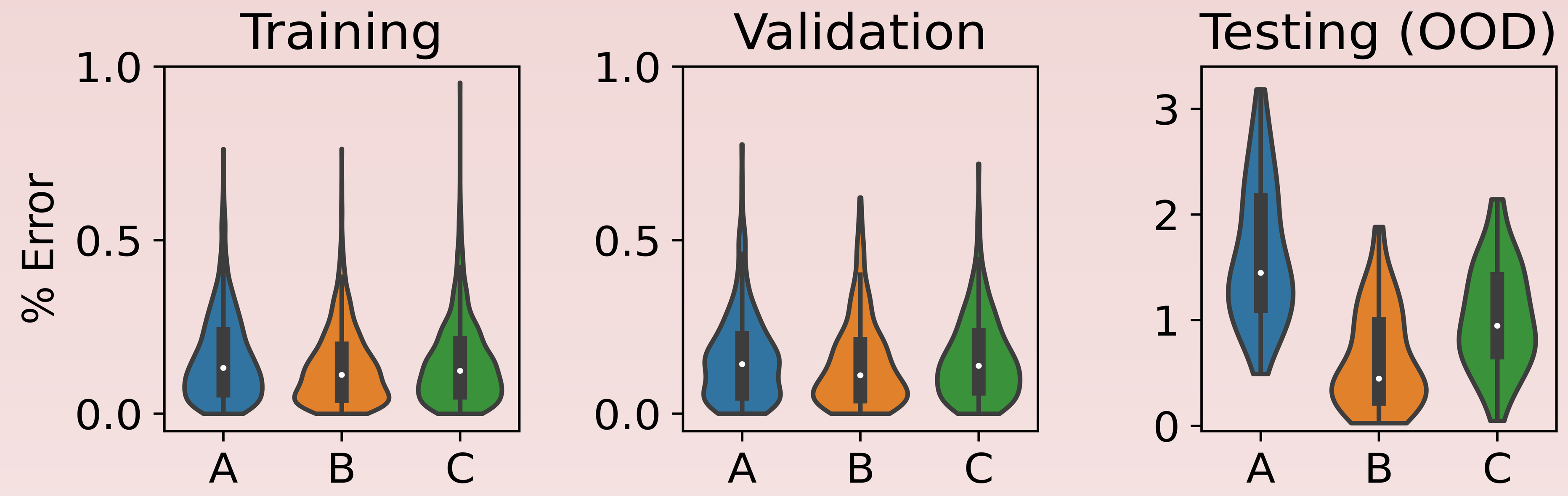


Figure 4: Violin plots of percent errors for each capacitor (A, B, C) over each of our data splits. The width of each plot denotes the density of predictions at that error level.

UNCERTAINTY FROM INPUT NOISE

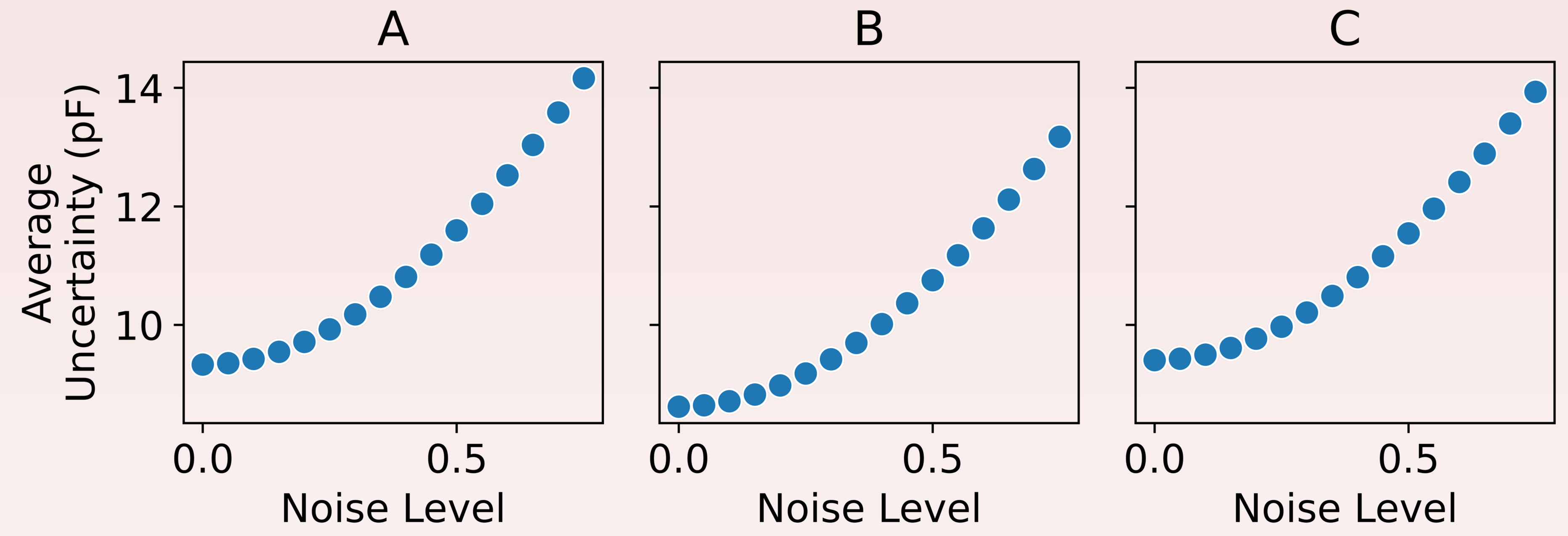
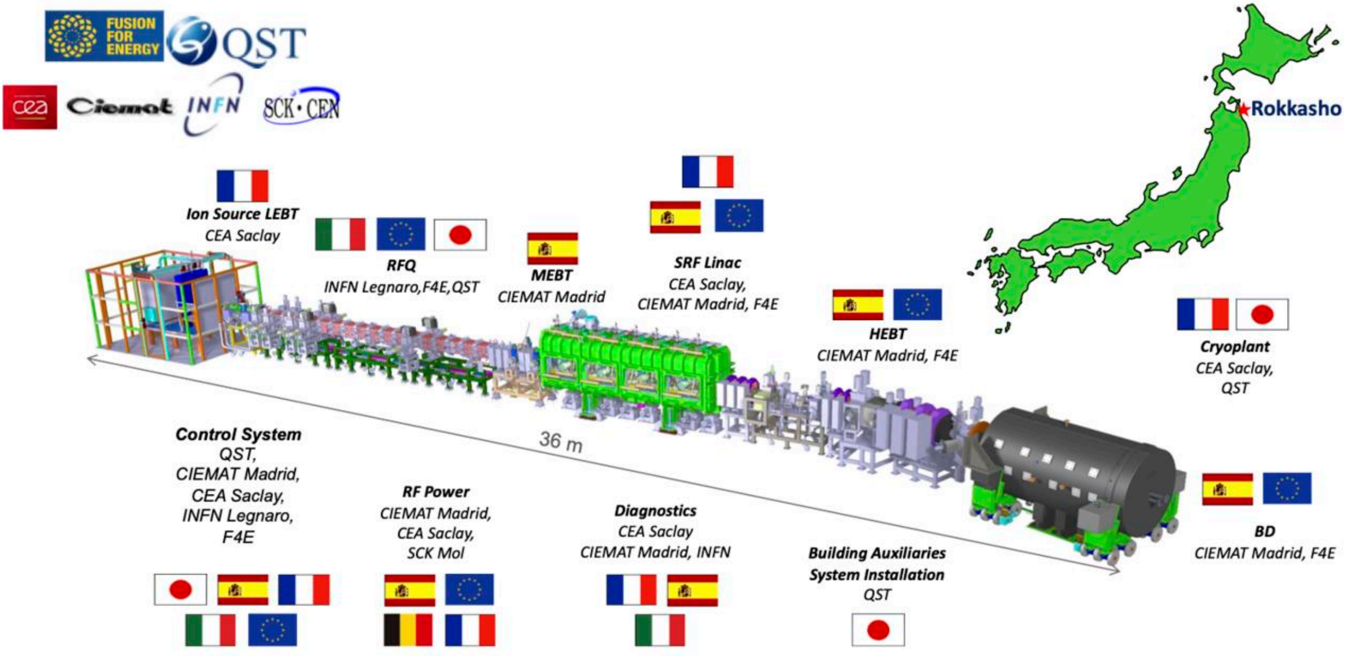


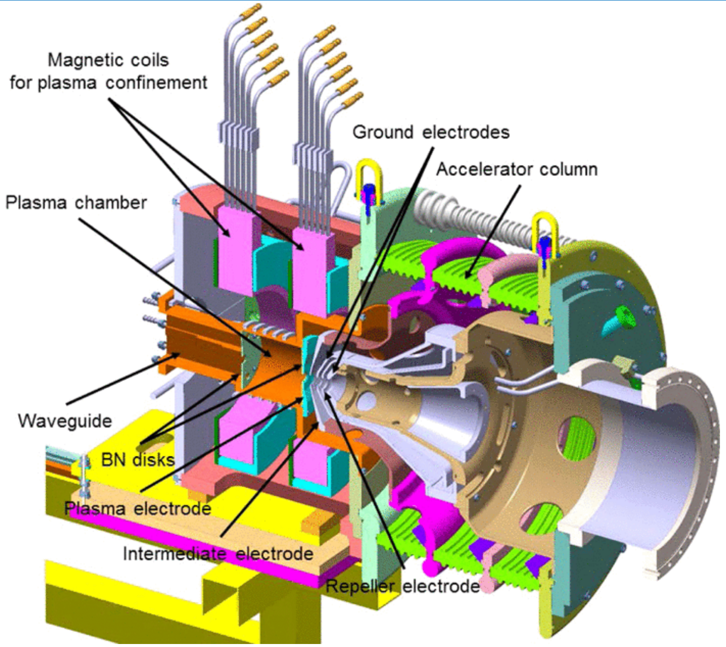
Figure 6: Average uncertainty returned by our model with increasing Gaussian noise applied to the training data.

ACKNOWLEDGEMENT

This research has been authored by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the US Department of Energy (DOE). The Jefferson Science Associates (JSA) operates the Thomas Jefferson National Accelerator Facility for the DOE under Contract No. DE-AC05-06OR23177. This research used resources at the Spallation Neutron Source, a DOE Office of Science User Facility operated by the Oak Ridge National Laboratory.

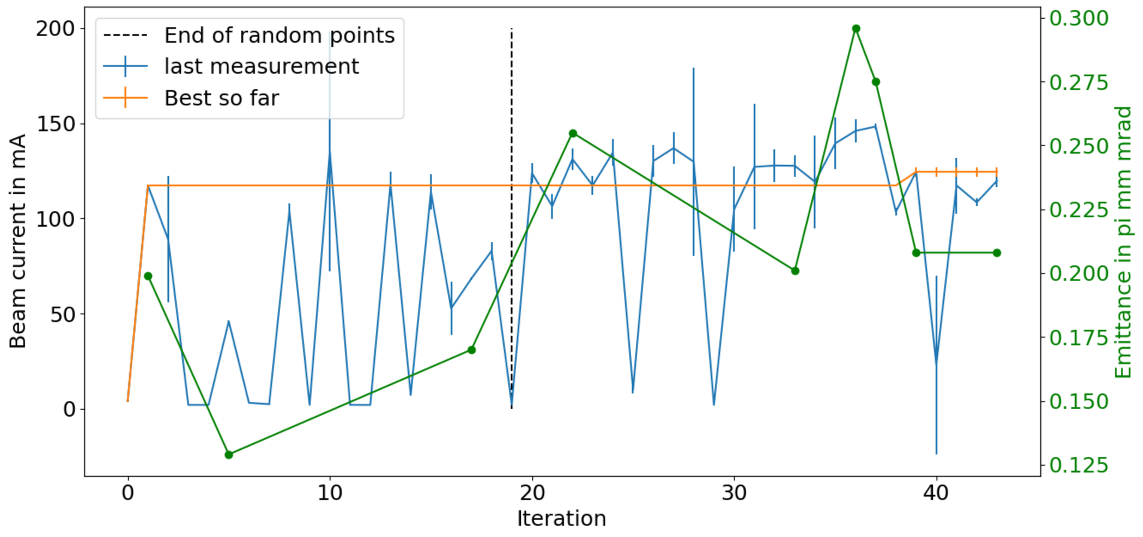
Multi Objective Bayesian Opt. of ECR Ion Source





Tuned 6 variables of ECR Ion Source to find:

- Max beam current
- Smallest beam instabilities
- Smallest emittance



Reshaping SRF Cavity Resonance Management with Smart Techniques

Faya Wang

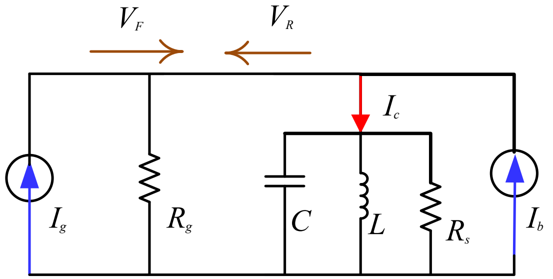
Mar 7, 2024

4th ICFA Beam Dynamics Workshop On Machine Learning Applications for Particle Accelerators, South Korea

SLAC NATIONAL ACCELERATOR LABORATORY

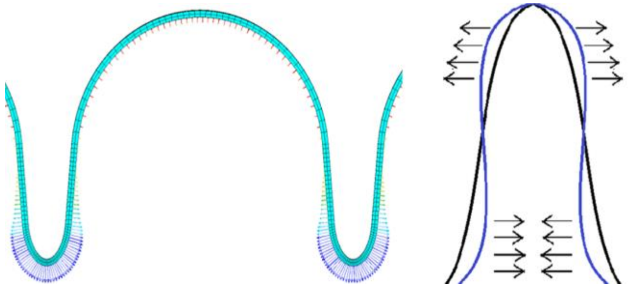
Stanford University | U.S. DEPARTMENT OF ENERGY

- **Cavity Circuit model**



Linear System

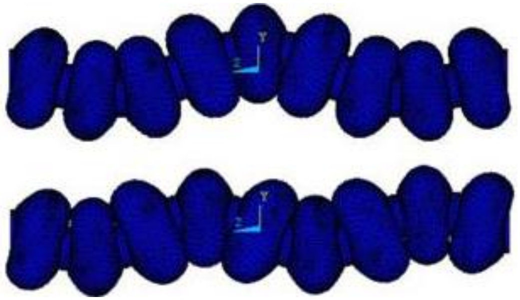
- **Lorentz pressure distribution on cavity wall**



Nonlinear Force

$$P = \frac{1}{4}\left(\varepsilon_0 E^2 - \mu_0 H^2\right)$$

- **Mechanical Modes: ω_m, Q_m, K_m**

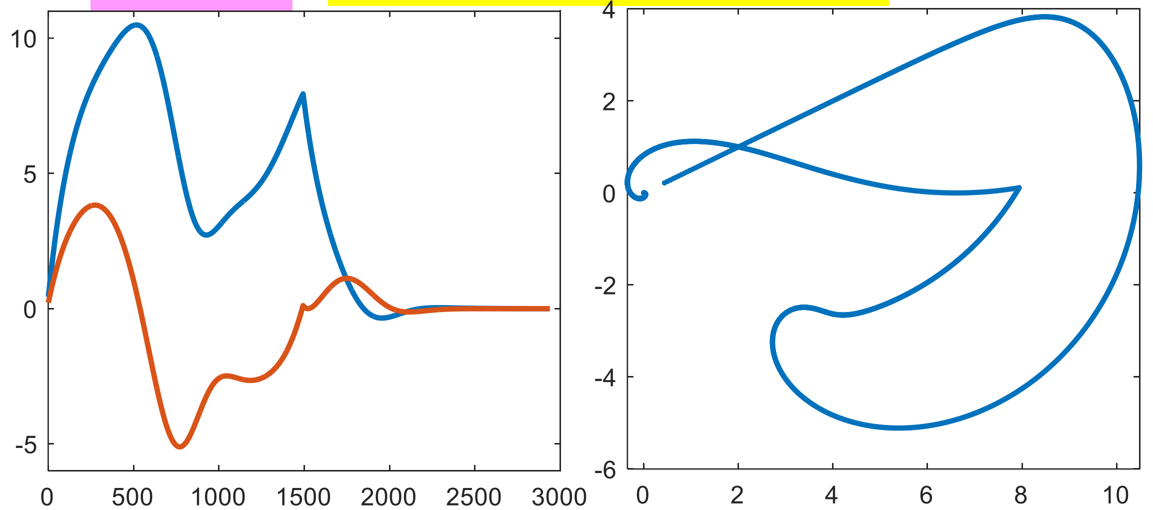


Electromagnetic field

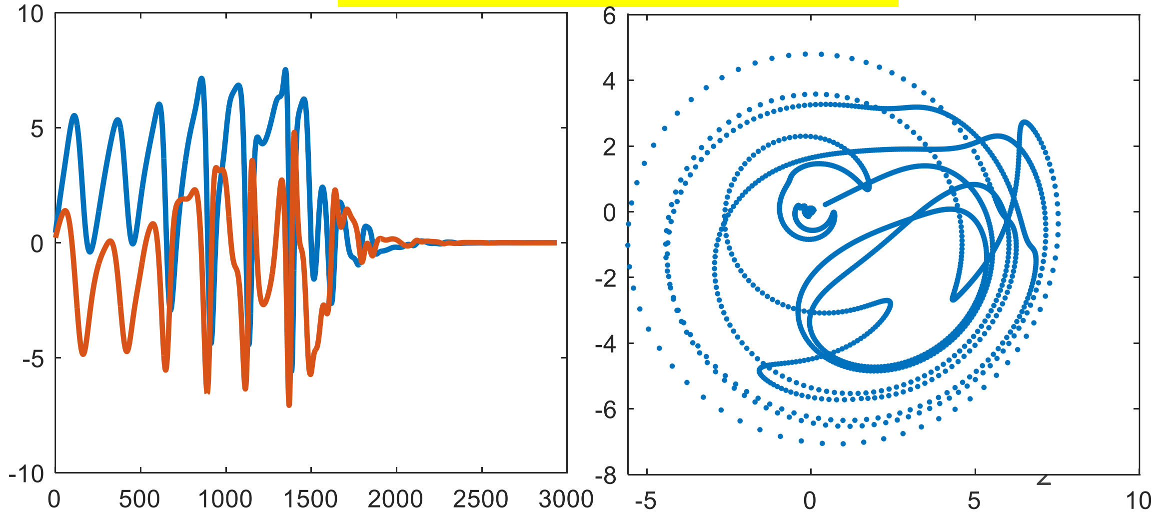
Vibration

$\times 5K_L$

Single Mechanical Mode



5 Mechanical Modes



- **DMD: Dynamic Mode Decomposition**

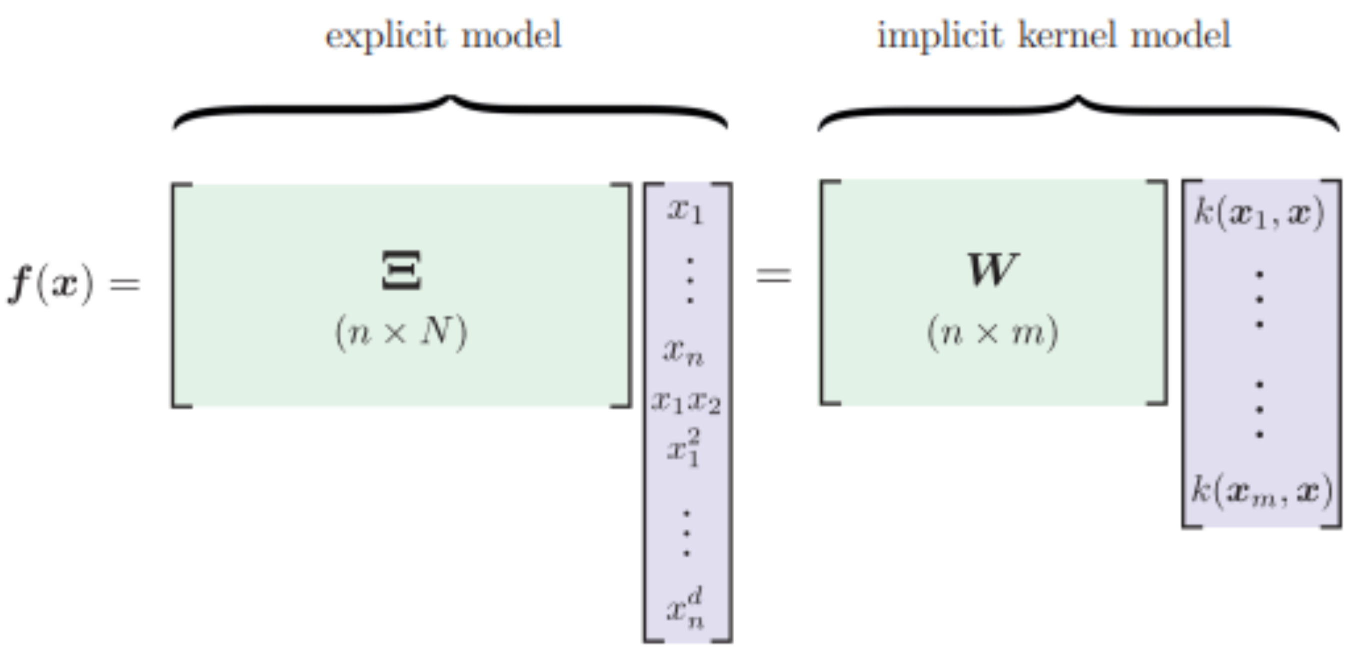
$$\frac{d}{dt}\boldsymbol{x}(t) = F(\boldsymbol{x}(t)) \qquad \boldsymbol{x_{k+1} = F(x_k)} \qquad F \approx f \text{ based on data}$$

$\boldsymbol{x_k} = [\boldsymbol{a_k}\ \ \boldsymbol{b_k}]^T$: system status and actuator inputs

Linear system: $\boldsymbol{x_{k+1} = W x_k}$

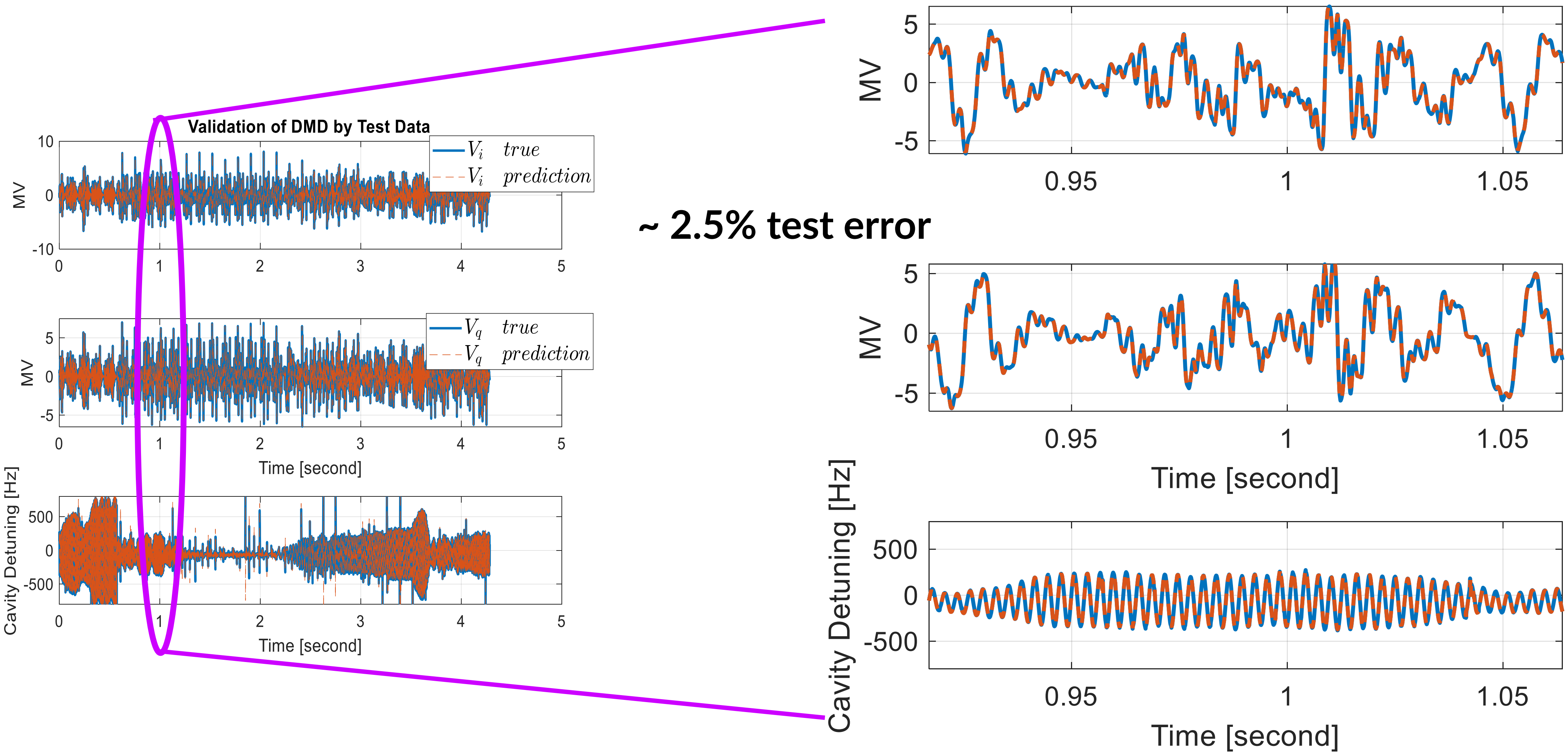
- **Mapping nonlinear problem in large state dimension with kernel function**

$$\boldsymbol{f \approx \sum_{j=1}^{N} \xi_j \phi_j(x) = \Xi\phi(x) = Wk(X, x)}$$



- $\boldsymbol{\phi}$: **the feature library of N candidate term that may describe the dynamics**
- $\boldsymbol{\Xi}$: **the coefficients that determine which feature terms are active and what proportions.**
- $\boldsymbol{k}$: **kernel function**
- **Data matrices:** $\boldsymbol{X = [x_1\ x_2\ \ldots x_m]}$

SRF Cavity DMD Test Results

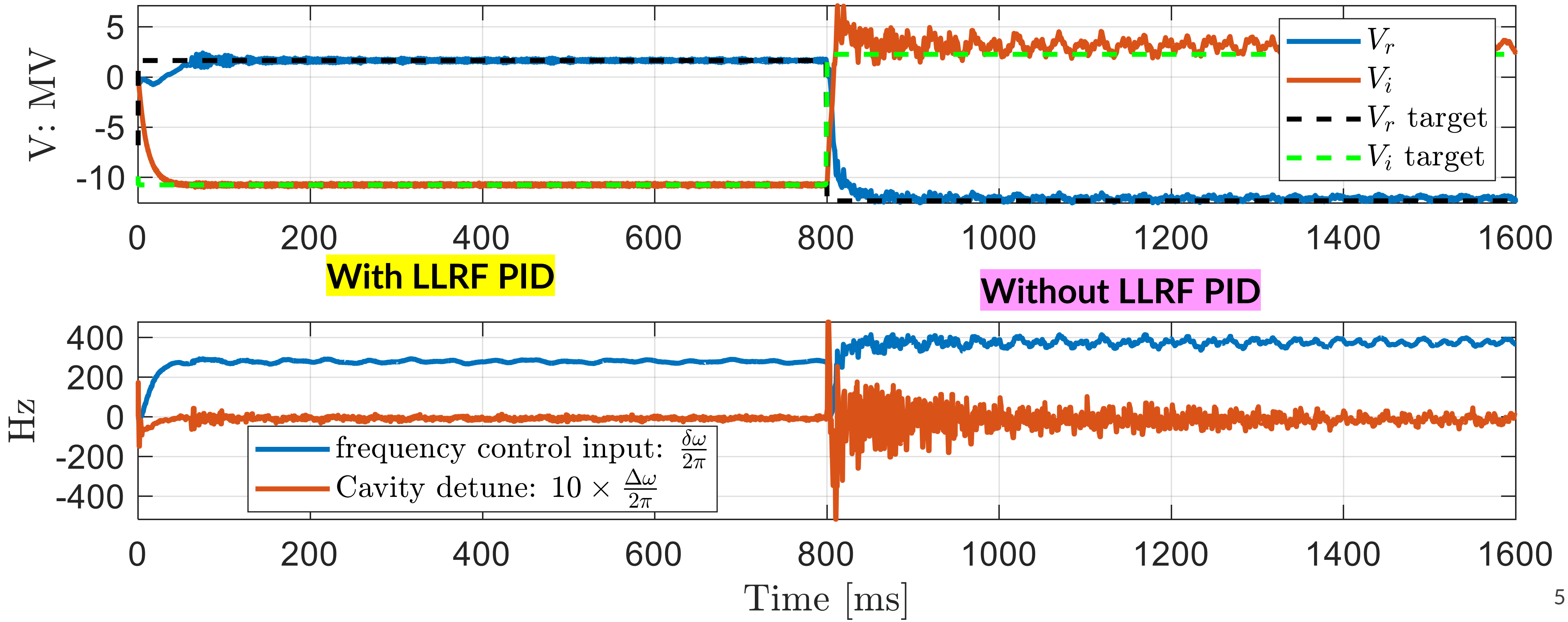


~ 2.5% test error

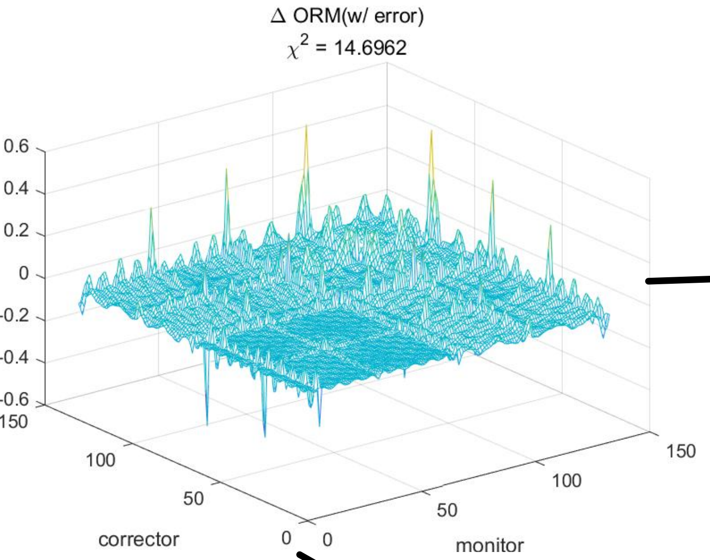
Faya Wang, "Enhancing SRF cavity stability and minimizing detuning with data-driven resonance control based on dynamic mode decomposition", AIP Advances 13, 075104 (2023)

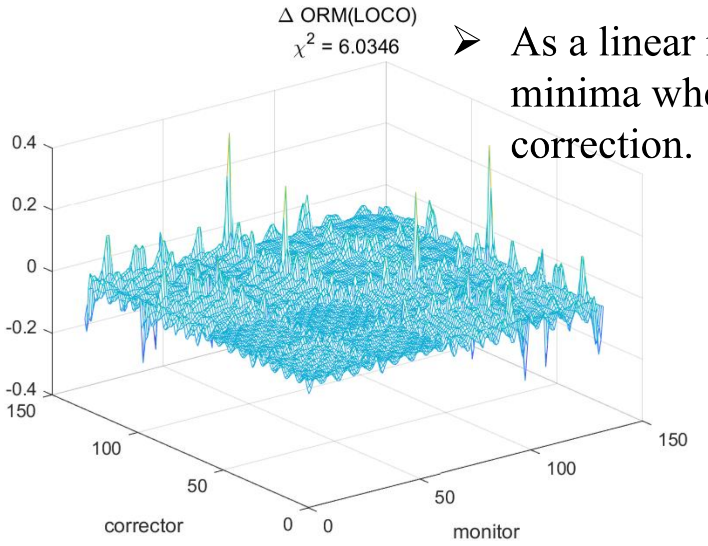
Active Resonance Controller

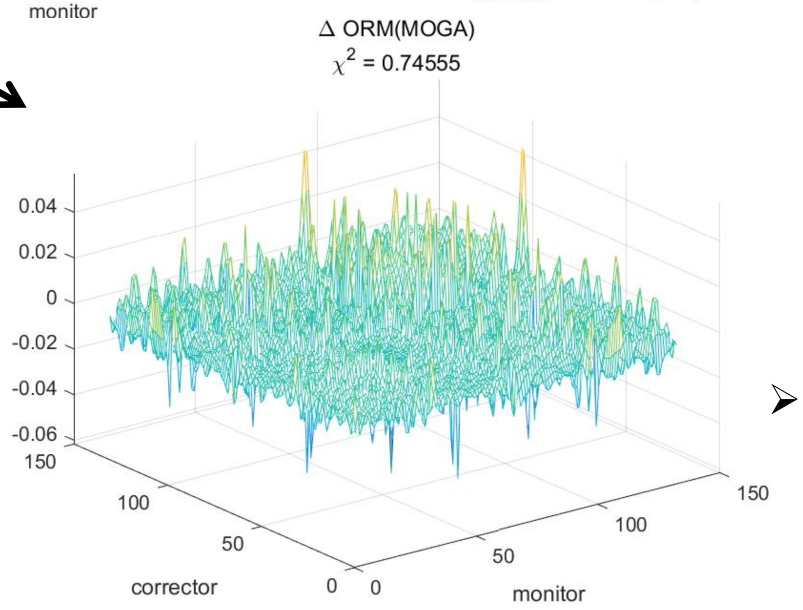
- **Simulation with 32 mechanical modes**
- **Cavity half bandwidth: 16.25 Hz**
- **Detune std: ~ 1 Hz**



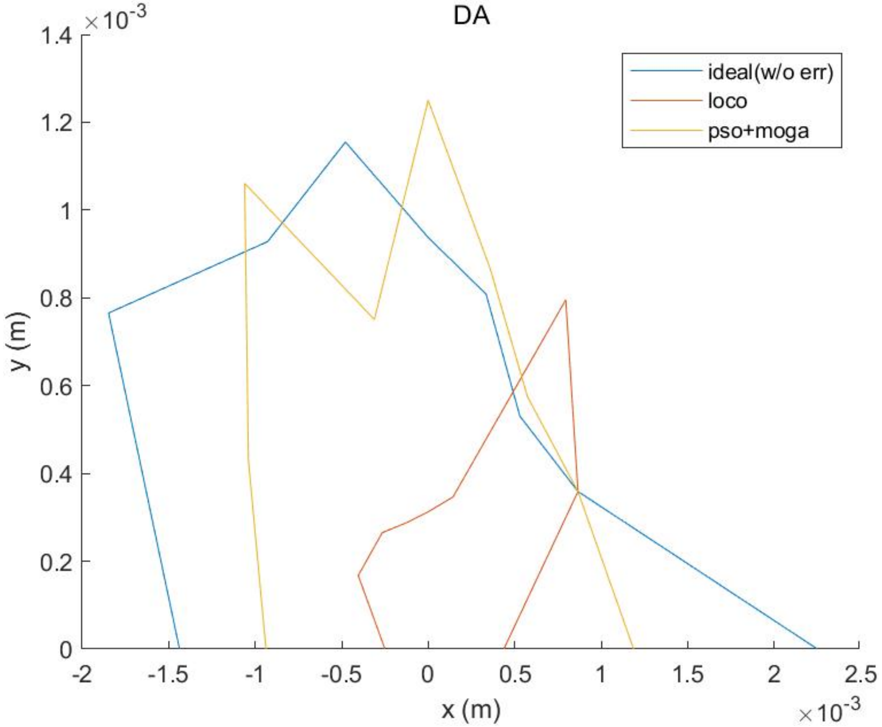
Orbit Response matrix correction based on exploration enhanced evolutionary algorithm





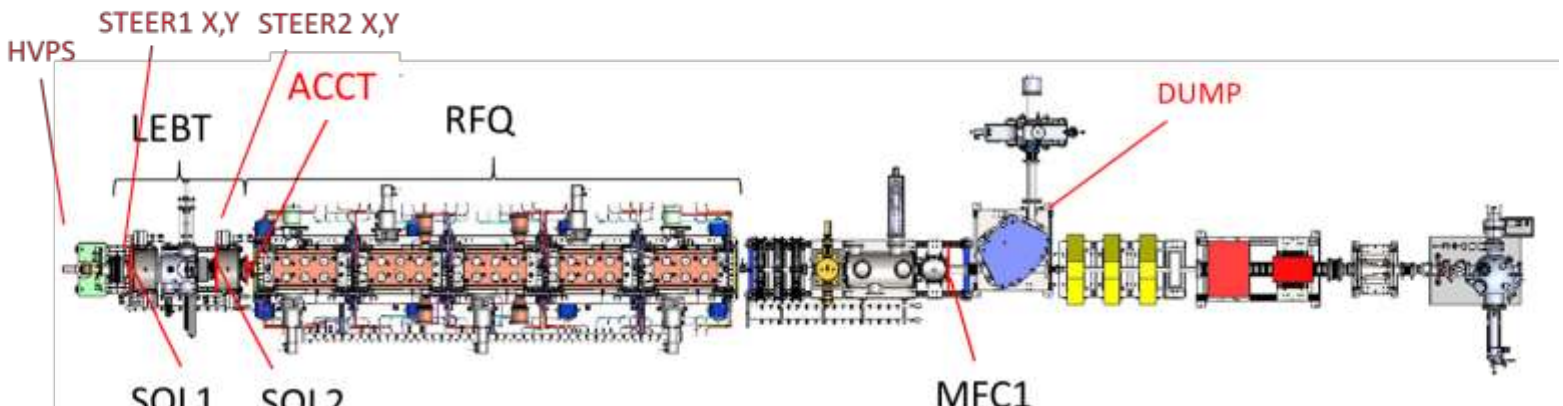


- As a linear method, LOCO is easily trapped into local minima when facing severe nonlinear response matrix correction.

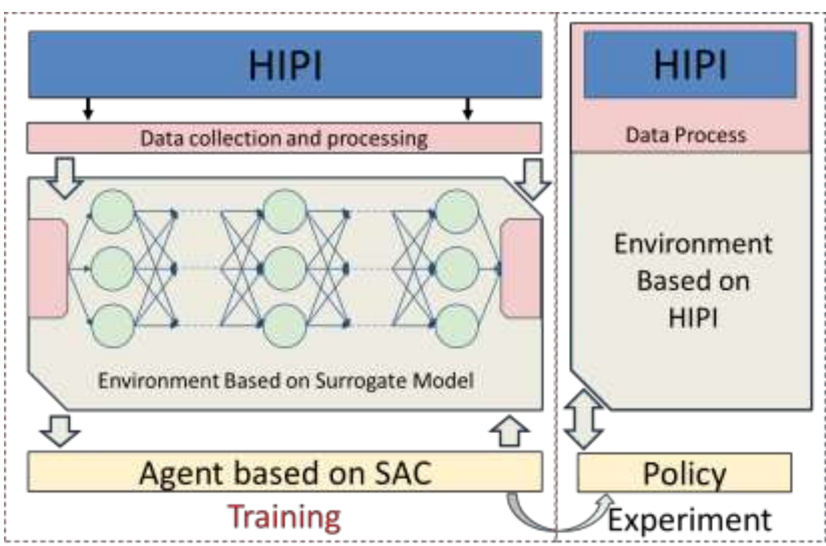


- Combining PSO and MOGA and introducing opposition based learning and Levy flight, it greatly improves the global exploration capability of the algorithm, which can significantly improve the effect of response matrix correction and obtain better dynamic performance.

• RL-Based Control Strategies for HIPI Accelerator



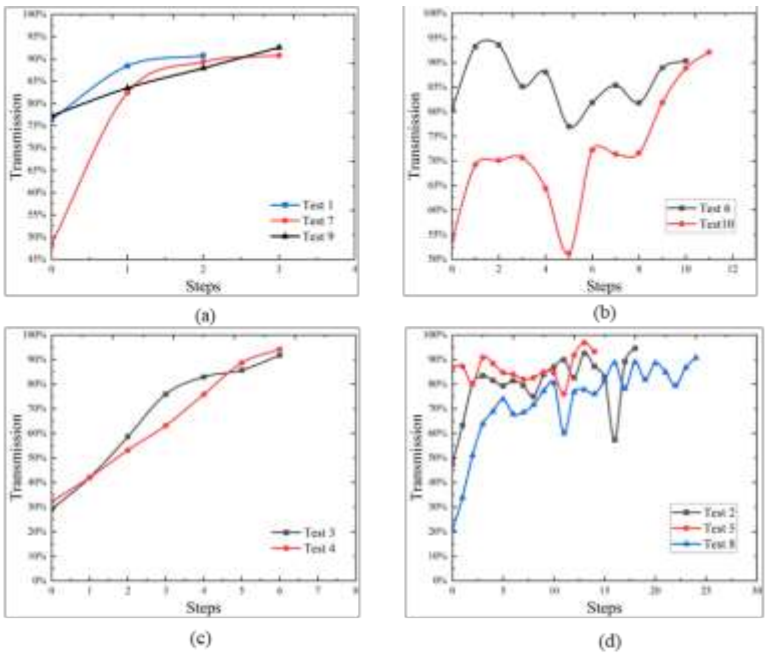
Layout of HIPI



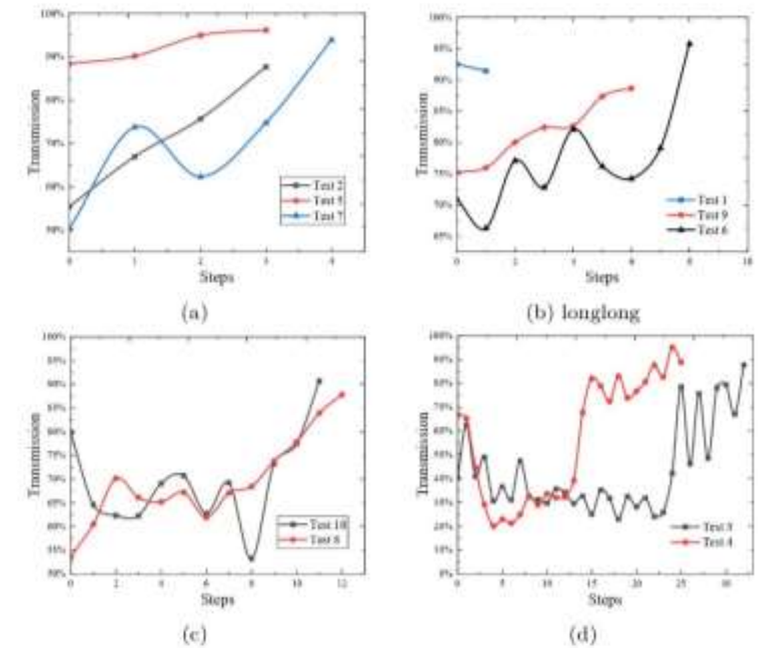
- Data Process
- Build a Surrogate Model
- Agent Trained Based on Surrogate Model

Target	High RFQ transmission with low beam loss in LEBT
Input Parameters	The current strength of the following 6 electromagnets, Solenoid 1, Solenoid 2,Steer 1X,Y, Steer 2 X,Y
Output Parameters	Current of ACCT and DUMP

• RL-Based Control Strategies for HIPI Accelerator



The Random Initial Values Policy Test with the Environment Based on Surrogate Model



The Random Initial Values Policy Test with the Environment Based on HIPI

The control strategy based on RL is faster than manual debugging in beam commission, completing hours of manual work in minutes.

Optimization design of photocathode injector assisted by deep gaussian process

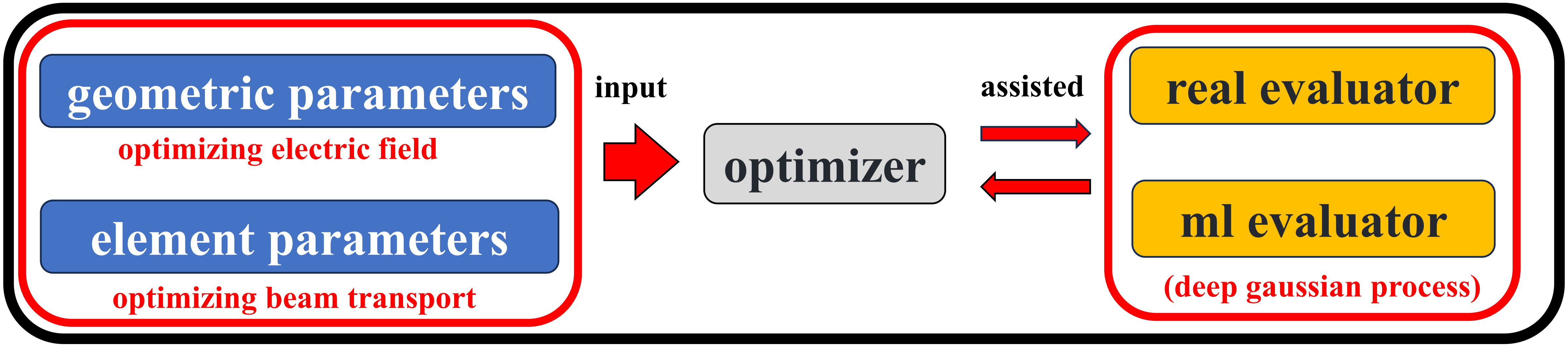
Sun Zheng， Xin Tianmu

Combined **geometric parameters** of radio frequency gun and **beam element parameters**

Challenge

too many parameters

evaluation process time-consuming



Objectives： reduce the emittance and bunch length at exit

Study of Orbit Correction by Neural Networks In Taiwan Photon Source

Mau-Sen Chiu

2024/03/07

Beam Dynamics Group, NSRRC

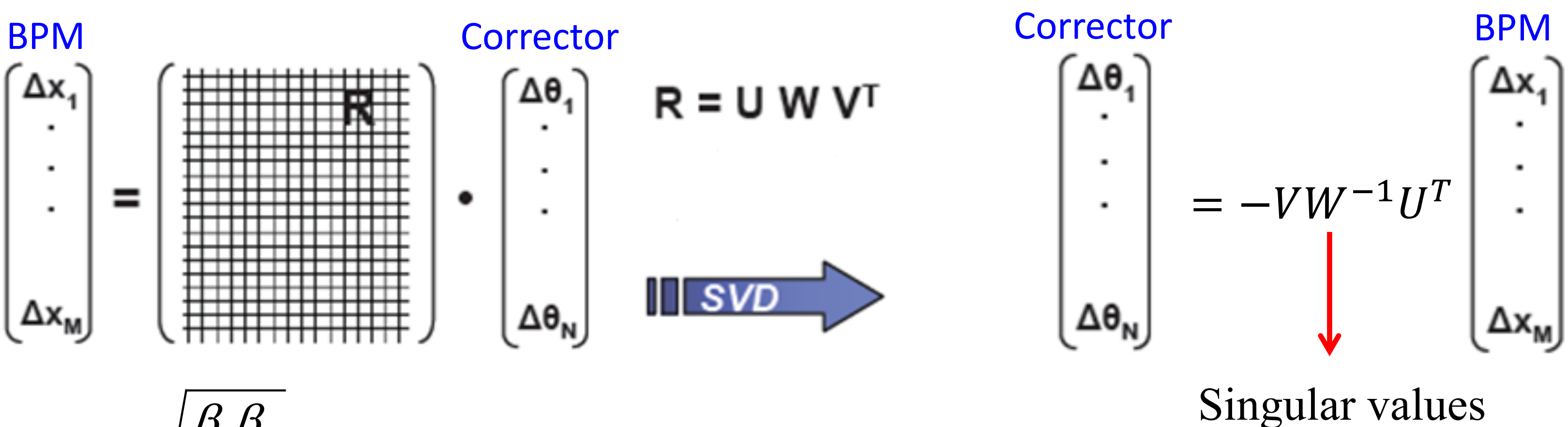
Abstract

The Taiwan Photon Source is designed as a 3 GeV synchrotron light source, encompassing a 518.4 m circumference. The lattice structure of the storage ring consists of 24 Double-Bend Achromat cells. The storage ring is equipped with 172 BPMs and 72/96 correctors to do orbit correction and control in horizontal and vertical planes, respectively. The correction algorithm uses a measured orbit response matrix and singular value decomposition (SVD) algorithm at present. This traditional method is rooted in physics and well-established principles of beam dynamics in particle accelerators. In this study, we use neural network model to do orbit correction. The training data for the neural networks is generated by accelerator toolbox (AT).

Orbit Correction by SVD (Traditional)

1. Establish reference orbit (Target Orbit)
2. Measure Orbit Response Matrix R between BPMs and correctors.
3. Apply SVD to decompose R, and select the number of singular values
4. Measure actual orbit - check for bad readings
5. Compute difference orbit
6. Compute corrector strength from → $\Delta\theta = -V \cdot diag(1/w_j)\cdot(U^T \cdot \Delta X)$

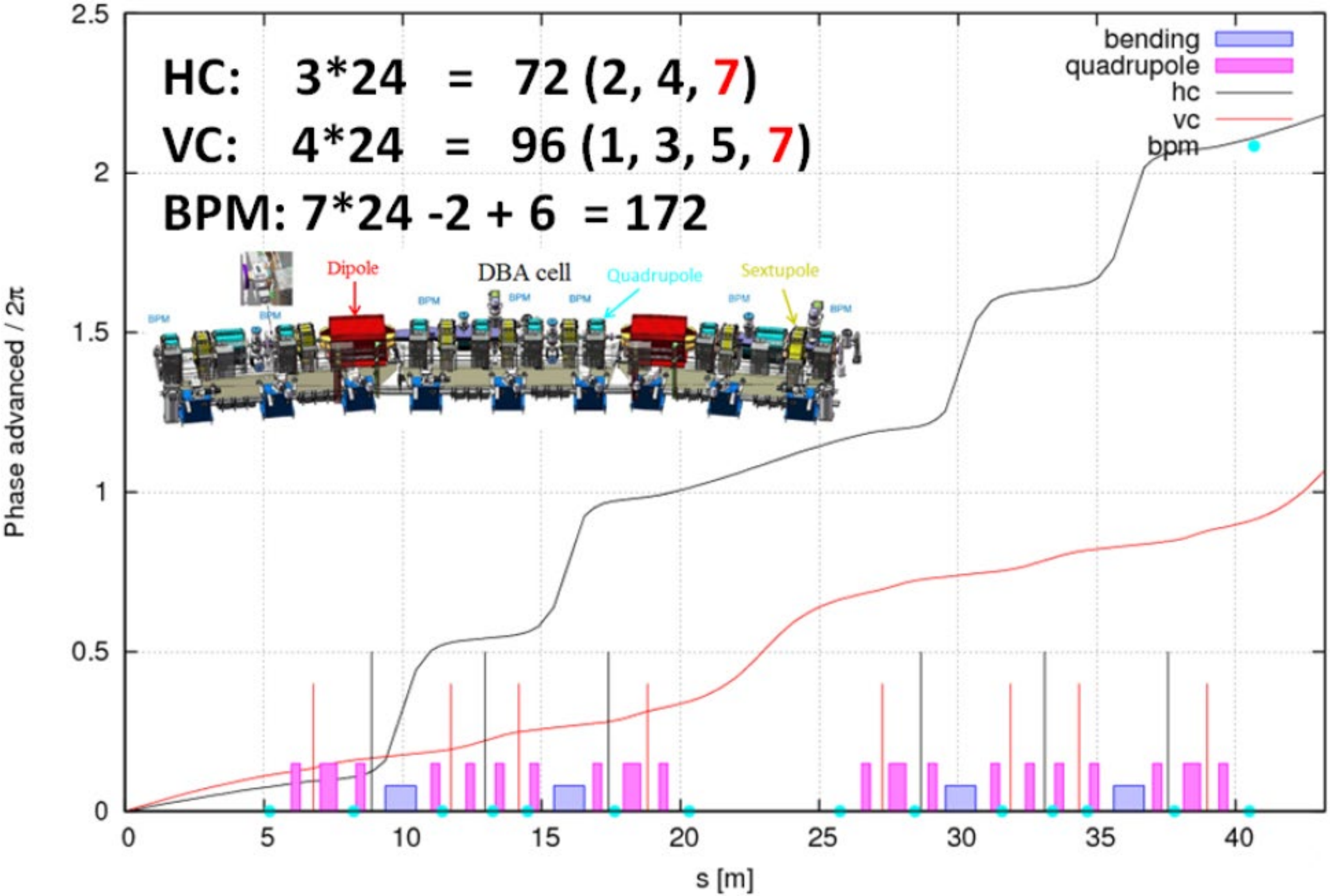
   ΔX: Difference Orbit
7. Check for corrector currents in reasonable range
8. Apply corrector currents



$$R_{ij} = \frac{\sqrt{\beta_i \beta_j}}{2\sin \pi\nu}\cos(|\phi_i - \phi_j| - \pi\nu)$$

It work with difference orbit and corrector changes rather than the absolute orbit and corrector values.

Orbit Correction Scheme in TPS Storage Ring



Simulation of Orbit Correction by Neural Networks

- **Training:**

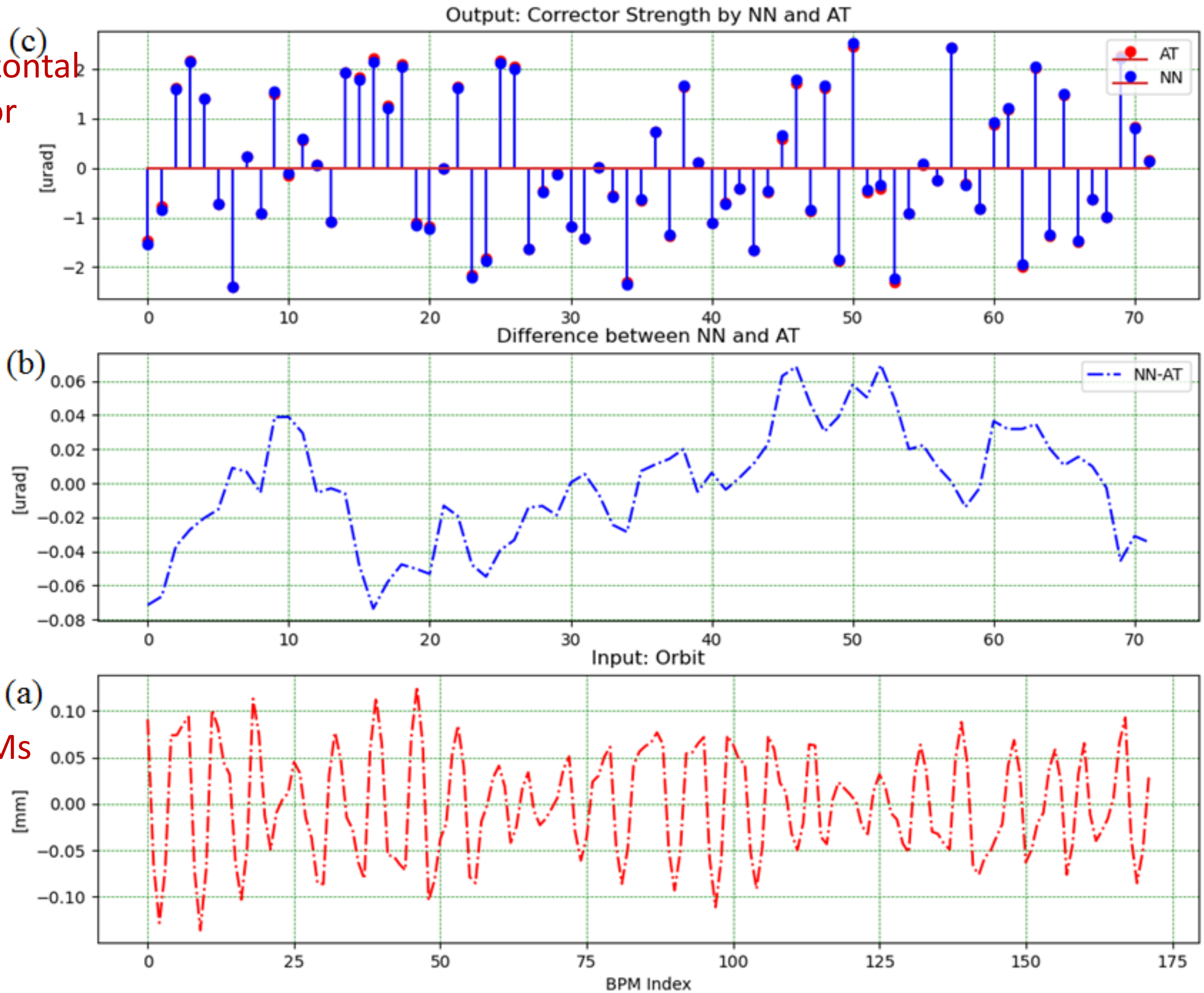
1. 72 horizontal correctors (HC) strengths within +/- 2.5 μrad are randomly assigned and then get orbits (172 BPMs) by AT: repeat 3000 times.
2. Build Model by keras: input layer is 172 nodes, hidden layer is 172 nodes, output layer is 72 nodes.
3. Train the model with AT simulation data.
4. Save the well-trained model of the neural networks.

- **Test:**

5. Generate many orbit distortions by randomly shifting 249 quadrupoles within +/- 3 μm in horizontal plane.
6. Load the well-trained model of the neural networks
7. Input the orbit distortions to the neural networks to get the predicted corrector strength
8. Use the predicted corrector strength to correct the orbit distortion generated by quadrupole misalignment
9. Iterate step 7 ~ 8: 3 times

Training Neural Networks (NN)

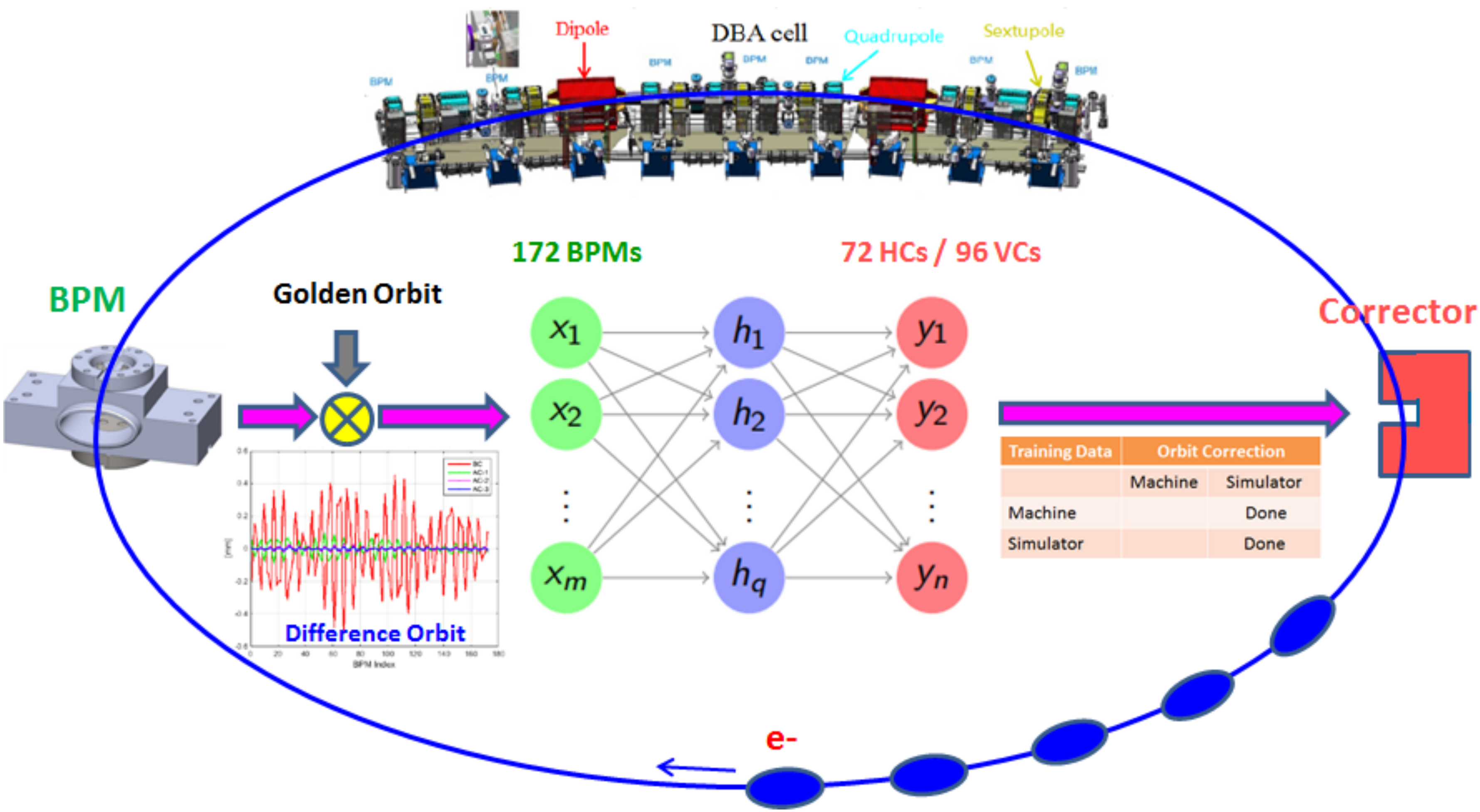
Output:
72 Horizontal
corrector



Input:
172 BPMs

Simulation of Orbit Correction by Neural Networks

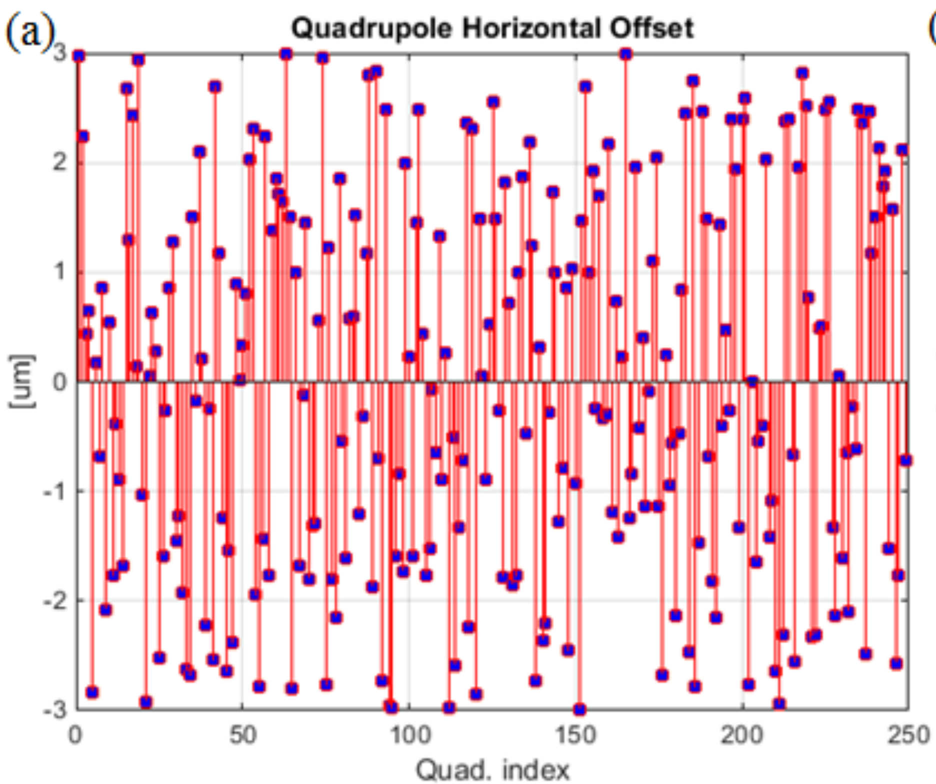
In TPS Storage Ring

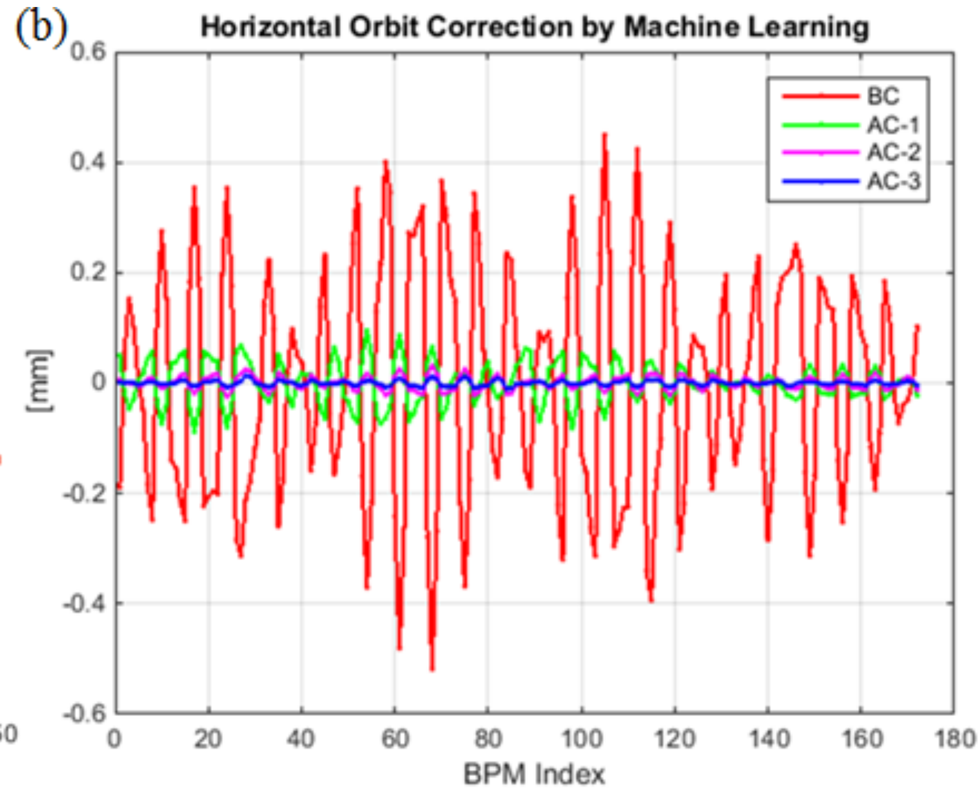


Misalignment quantities of 249 quadrupole magnets within +/- 3 μm to generate orbit distortion in TPS storage ring simulated by AT.

Simulation of Orbit Correction by Neural Networks

In TPS Storage Ring





Misalignment quantities of 249 quadrupole magnets within +/- 3 μm to generate orbit distortion in TPS storage ring simulated by AT.

Orbit correction by neural network: Red is the orbit before correction (BC), green, magenta, and blue are the orbit after correction (AC), iterate 3 times (AC-1, AC-2, AC-3).

Demonstration

APPENDIX

17th Int. Conf. on Acc. and Large Exp. Physics Control Systems ICALEPCS2019, New York, NY, USA JACoW Publishing
ISBN: 978-3-95450-209-7 ISSN: 2226-0358 doi:10.18429/JACoW-ICALEPCS2019-WEPHA138

ORBIT CORRECTION WITH MACHINE LEARNING TECHNIQUES AT THE SYNCHROTRON LIGHT SOURCE DELTA

115 m, 1.5 GeV

D. Schirmer*
Center for Synchrotron Radiation (DELTA), TU Dortmund University, Germany

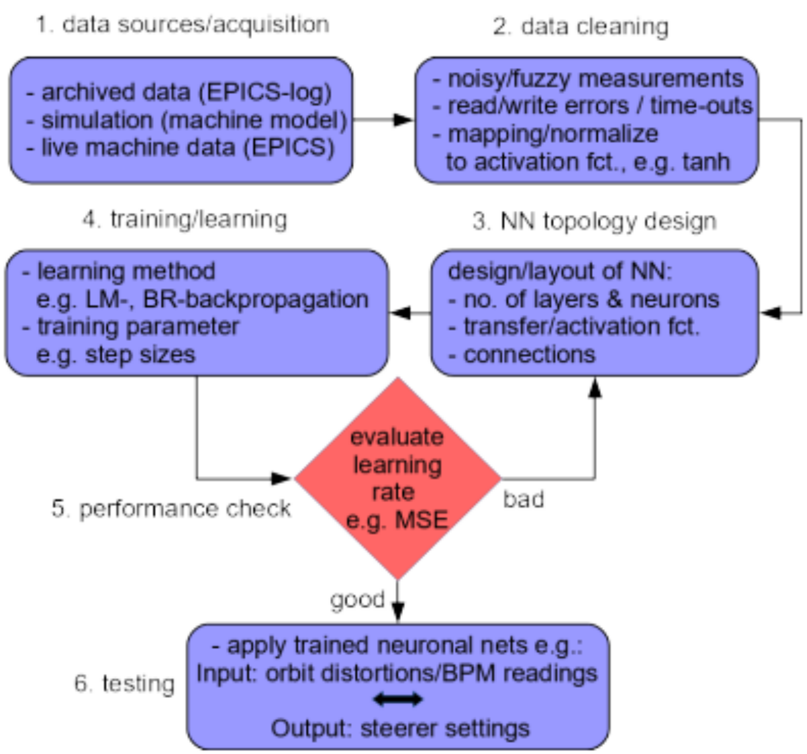


Figure 1: Development stages for an ML-based OC.

30 HC (±200 to 300 mA), 54 BPM, 1500 data sets

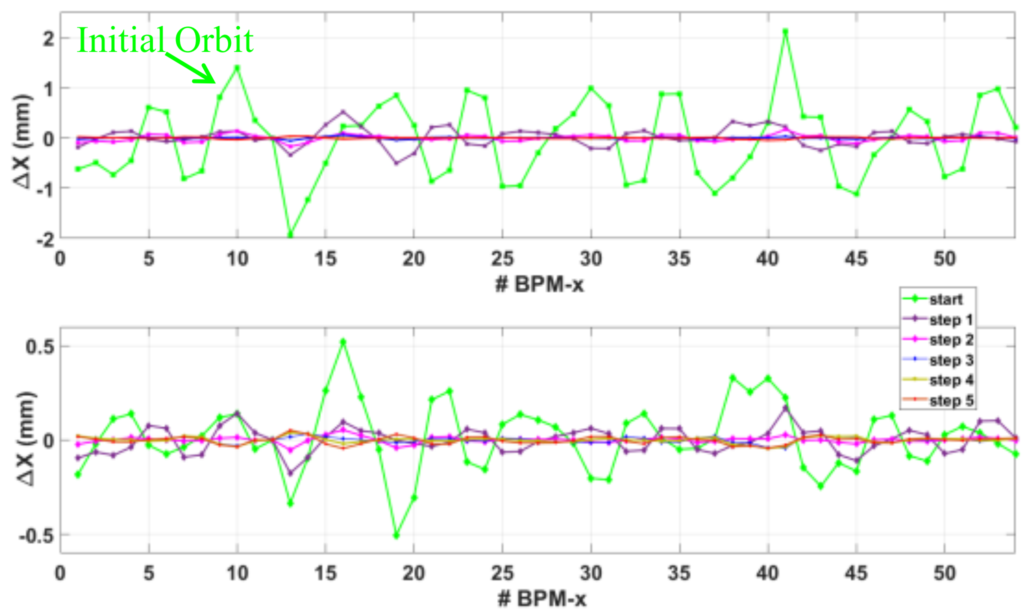
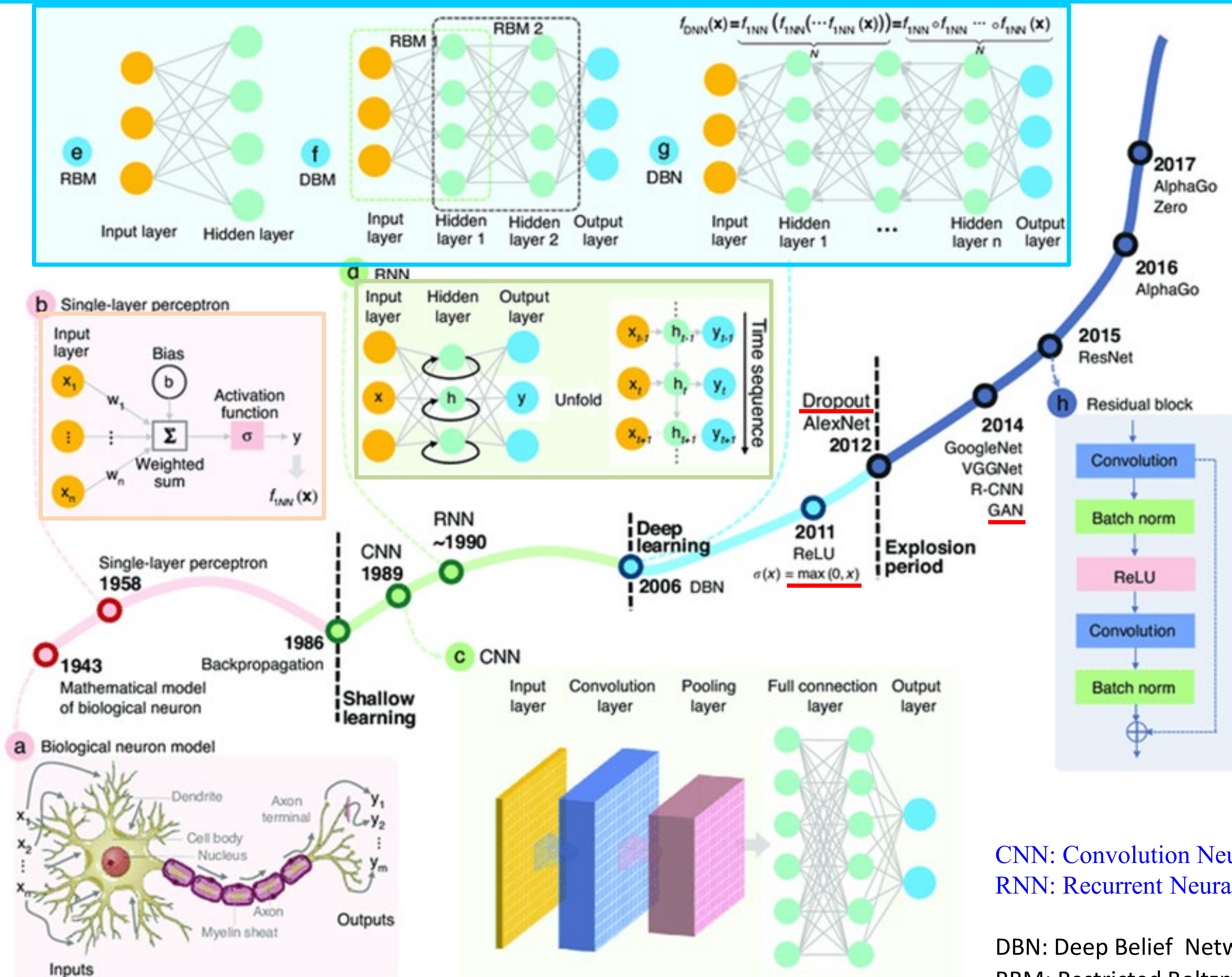


Figure 6: Iterative application of the pretrained FFNN referred to the previously corrected orbit, starting from a randomly disturbed orbit (start). After 3 successive correction steps, an error of < 200 µm was achieved.

History of Neural Networks



CNN: Convolution Neural Network
RNN: Recurrent Neural Network

DBN: Deep Belief Network
RBM: Restricted Boltzmann Machine

Ref: Deep learning in optical metrology: a review, Chao Zuoet al. Light: Science & Applications (2022) 11:39

Popular Deep Learning & Software

TABLE 2 | List of popular deep learning models, available learning algorithms (unsupervised, supervised) and software implementations in R or python.

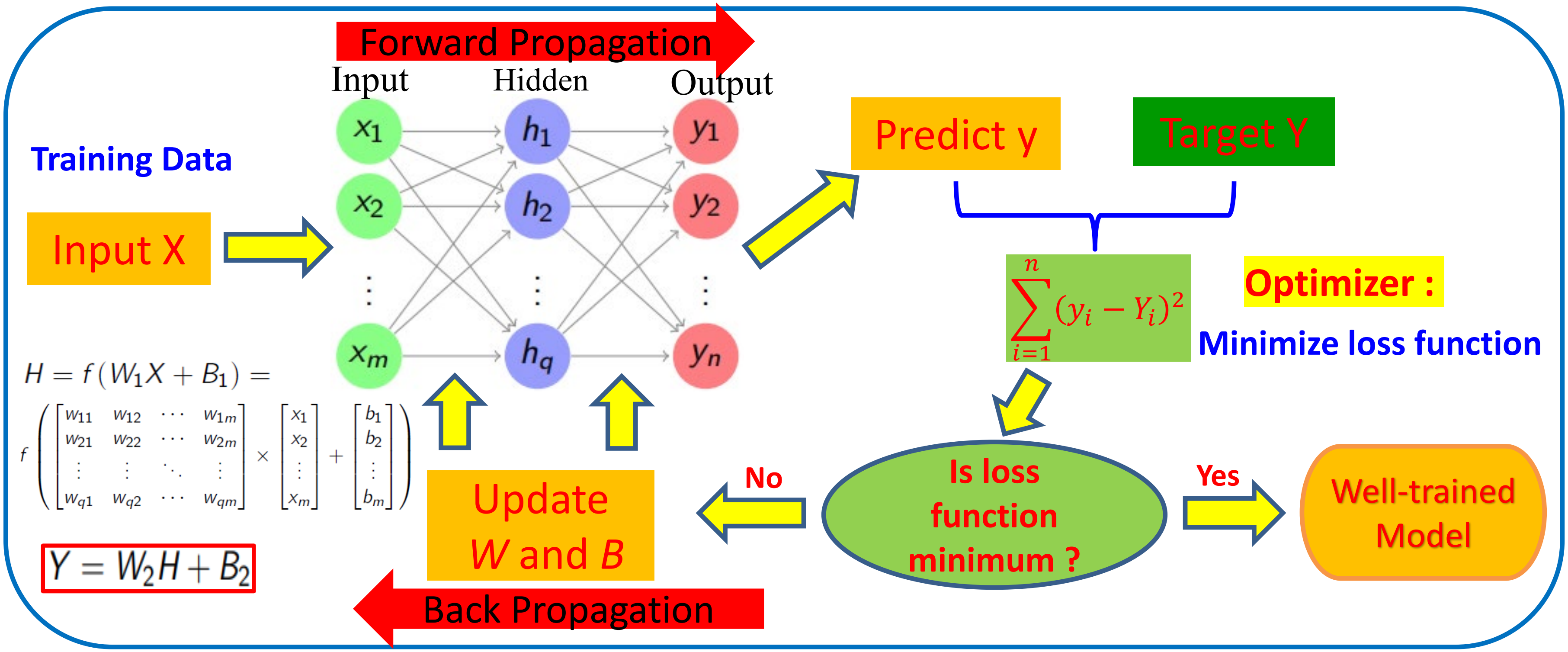
Model	Unsupervised	Supervised	Software
Autoencoder	✓		Keras (Chollet, 2015), R: dimRed (Kraemer et al., 2018), h2o (Candel et al., 2015), RcppDL (Kou and Sugomori, 2014)
Convolutional Deep Belief Network (CDBN)	✓	✓	R & python: TensorFlow (Abadi et al., 2016), Keras (Chollet, 2015), h2o (Candel et al., 2015)
Convolutional Neural Network (CNN)	✓	✓	R & python: Keras (Chollet, 2015) MXNet (Chen et al., 2015), Tensorflow (Abadi et al., 2016), h2O (Candel et al., 2015), fastai (python) (Howard and Gugger, 2018)
Deep Belief Network (DBN)	✓	✓	RcppDL (R) (Kou and Sugomori, 2014), python: Caffee (Jia et al., 2014), Theano (Theano Development Team, 2016), Pytorch (Paszke et al., 2017), R & python: TensorFlow (Abadi et al., 2016), h2O (Candel et al., 2015)
Deep Boltzmann Machine (DBM)		✓	python: boltzmann-machines (Bondarenko, 2017), pydbm (Chimera, 2019)
Denoising Autoencoder (dA)	✓		Tensorflow (R, python) (Abadi et al., 2016), Keras (R, python) (Chollet, 2015), RcppDL (R) (Kou and Sugomori, 2014)
Long short-term memory (LSTM)		✓	rnn (R) (Quast, 2016), OSTSC (R) (Dixon et al., 2017), Keras (R and python) (Chollet, 2015), Lasagne (python) (Dieleman et al., 2015), BigDL (python) (Dai et al., 2018), Caffe (python) (Jia et al., 2014)
Multilayer Perceptron (MLP)		✓	SparkR (R) (Venkataraman et al., 2016), RSNNS (R) (Bergmeir and Benítez, 2012), keras (R and python) (Chollet, 2015), sklearn (python) (Pedregosa et al., 2011), tensorflow (R and python) (Abadi et al., 2016)
Recurrent Neural Network (RNN)		✓	RSNNS (R) (Bergmeir and Benítez, 2012), rnn (R) (Quast, 2016), keras (R and python) (Chollet, 2015)
Restricted Boltzmann Machine (RBM)	✓	✓	RcppDL (R) (Kou and Sugomori, 2014), deepnet (R) (Rong, 2014), pydbm (python) (Chimera, 2019), sklearn (python) (Chimera, 2019), Pylearn2 (Goodfellow et al., 2013), TheanoLM (Enarvi and Kurimo, 2016)

Ref: An Introductory Review of Deep Learning for Prediction Models With Big Data, Frontiers in Artificial Intelligence, 28 Feb. 2020

Training by Backpropagation

How to determine weights and bias ?

- Initialize weights "randomly"
- For all training epochs
  - for all input-output in training set
    - using input and compute output : forward propagation
    - compare computed output with training output -> calculate loss function
    - update weights (backpropagation) to improve output -> minimize loss function
  - if accuracy is good enough, stop



https://www.analyticsvidhya.com/blog/2016/08/evolution-core-concepts-deep-learning-neural-networks/

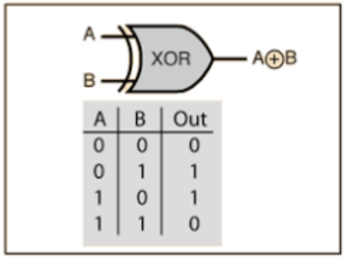
Workflow of Neural Networks

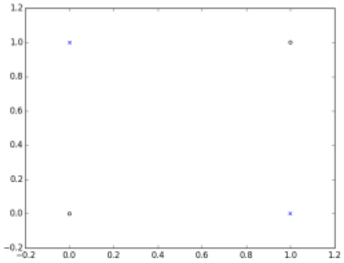
- **Software Packages**: Keras, Tensorflow, Python.
- **Data collection**: Scaling and normalizing data, then splitting data into training, validation and test sets.
- **Build a neural network**: Select an appropriate neural network architecture (e.g. feedforward, recurrent, convolution, *et al*) based on problem type (e.g. regression, classification, *et al*.), and assign the number of layers, neuron number in each layer, activation function (e.g. sigmoid, tanh, ReLu, *et al*.).
- **Compile the Model**: Specify the loss function (e.g. mean square error, *et al*.), optimizer (e.g. adam, sgd, *et al*.) that adjusts the model's weights and bias.
- **Fit (Training) Model (minimize loss function)**: Specify the batch size, the number of epochs (training iteration times), and using training set of data.
- **Evaluate Model**: Evaluate the model's performance by using validation data set.
- **Fine-Tuning Hyperparameter**: Training model with different learning rate (step size during training), batch size (number of data sets used in each iteration of training, , number of layers, neurons per layer, Epoch (training times of passing data sets through network model), to avoid underfitting and overfitting.
- **Make Predictions**: Use the trained model to make prediction on test data.

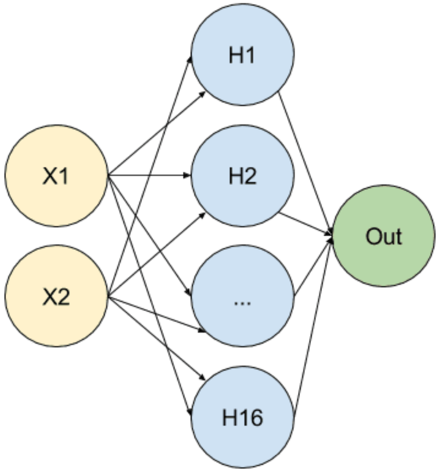
Python Code by Keras for XOR

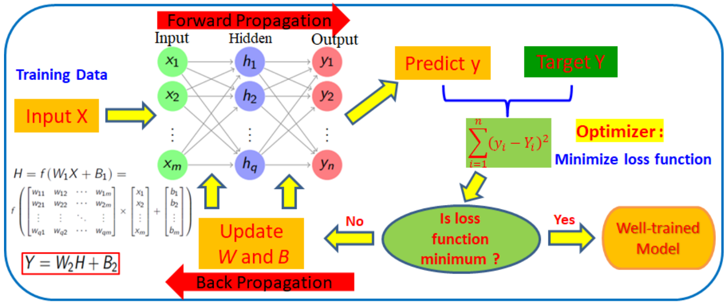
```python
import numpy as np
from keras.models import Sequential
from keras.layers.core import Dense

# the four different states of the XOR gate
training_data = np.array([[0,0],[0,1],[1,0],[1,1]], "float32")

# the four expected results in the same order
target_data = np.array([[0],[1],[1],[0]], "float32")
# Build a model
model = Sequential()
model.add(Dense(16, input_dim=2, activation='relu'))
model.add(Dense(1, activation='sigmoid'))

model.compile(loss='mean_squared_error',
          optimizer='adam',
          metrics=['binary_accuracy'])
# start to train
model.fit(x=training_data, y=target_data, nb_epoch=500, verbose=2)
# Prediction
print model.predict(training_data).round()
```

Dense: Fully connected









https://keras.io/api/metrics/

https://keras.io/api/optimizers/

https://keras.io/api/models/model_training_apis/

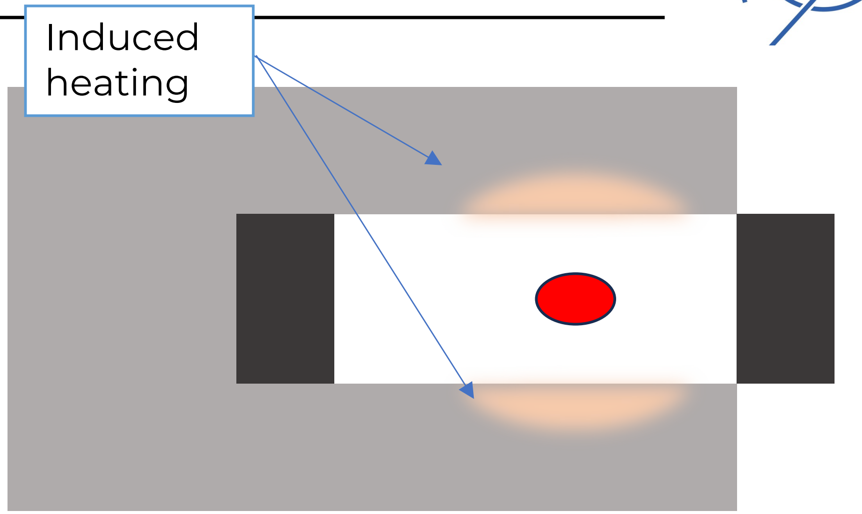
https://blog.thoughtram.io/machine-learning/2016/11/02/understanding-XOR-with-keras-and-tensorlow.html

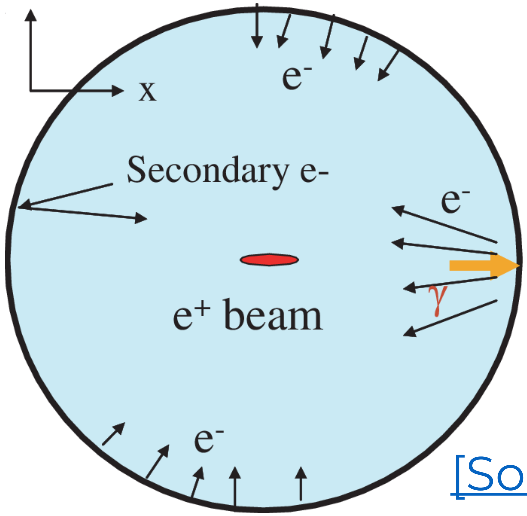
Dynamic vacuum and temperature predictions for informed anomaly detection at the CERN-SPS

F.M. Velotti, M. Barnes, G. Favia, V. Kain

Motivation

- Moving towards High Luminosity LHC → intensity effects induce high stress on sensitive equipment, e.g. kickers, septa…
- Electron cloud, beam induced heating main responsible
- Important source of unavailability of the CERN-SPS
  - If kickers get too hot, risk to not be able to inject or damage them ➔ cool down needed
  - If vacuum rises too much, high risk of breakdown in high voltage elements
- ~50% of call to stand-by service "just" for resets due to high vacuum/high temperature interlock trip

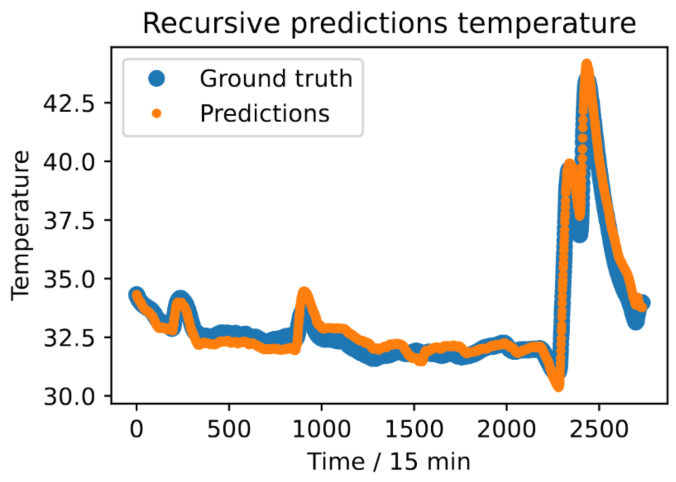




[Source]

Possible solutions

- Developing "intelligent" system to:
  - Predict machine behavior for given beam properties
  - Determine normal and abnormal operation
  - Identify breakdowns from vacuum readings
  - Define optimal machine usage

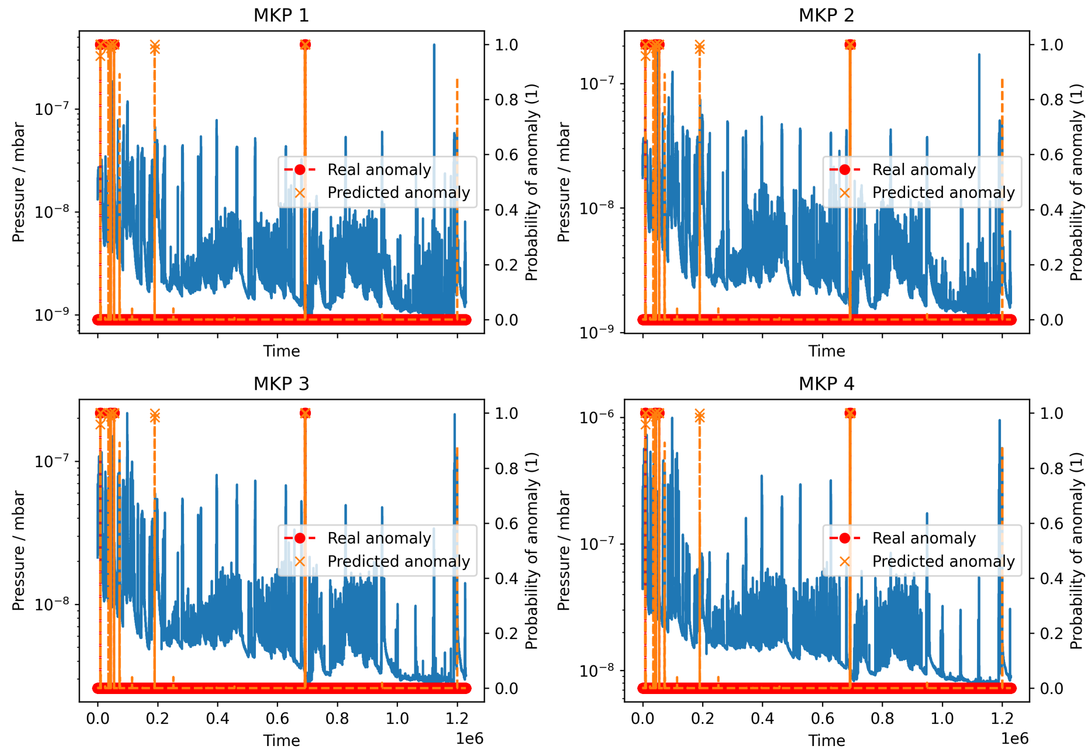


- Gradient boosting algorithms very successful at this:
  - Fast training and fast predictions
  - Capable of handle complex responses:

$$\Delta W = (f_0 e I_b N_b)^2 \sum (|\Gamma(k\omega_0)^2 R[Z_{||}(k\omega_0)])$$

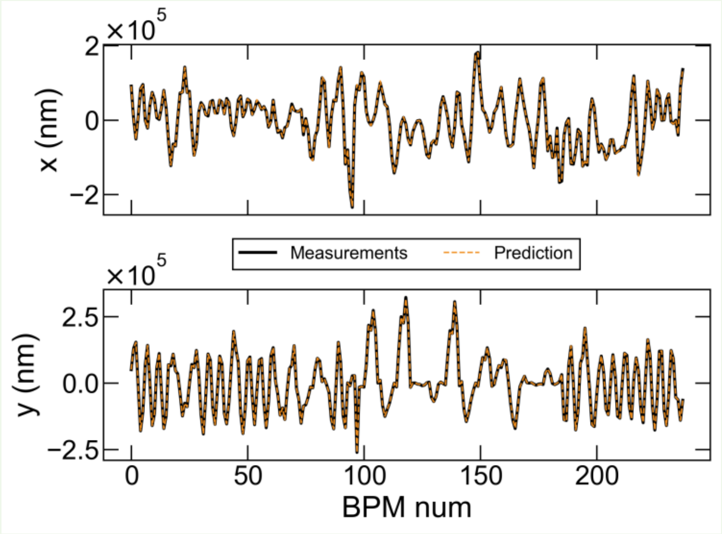
$$\frac{dT}{dt} = \frac{k}{\rho C_p}\frac{\partial^2 T}{\partial x^2} + \frac{\Delta W}{\rho C_p}$$

  - Possibility to implement in continuous learning → fight concept drift/conditioning of different elements



Experience with ML-driven applications at PETRA III

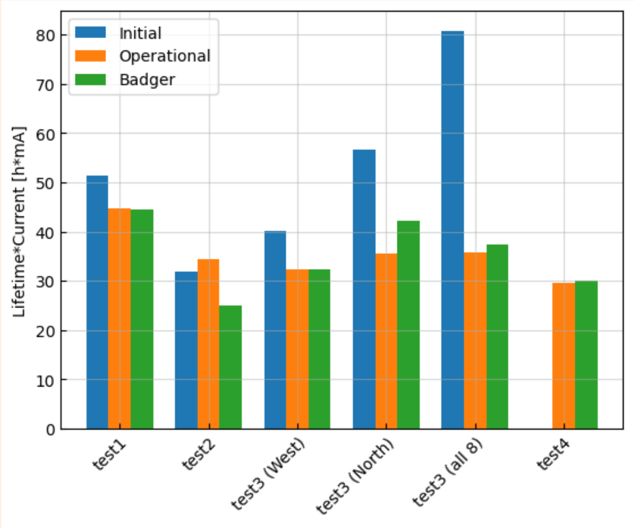
NNs for ID orbit distortion



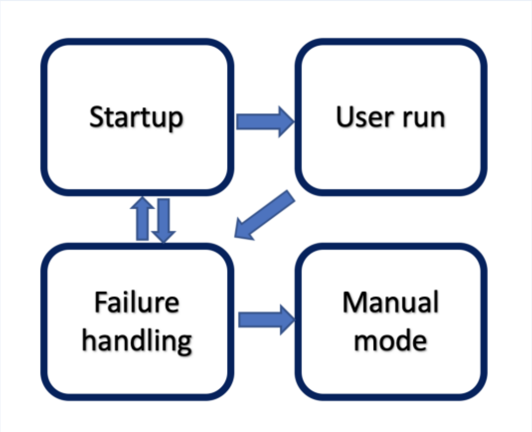
The combined impact of multiple IDs moving at the same time is non-linear. NNs accurately predict the transverse displacement of the beam along the ring for any given ID configuration.

Badger/Xopt optimization

Optimizations of lifetime and injection efficiency were performed and compared with the results obtained with the current manual procedures.



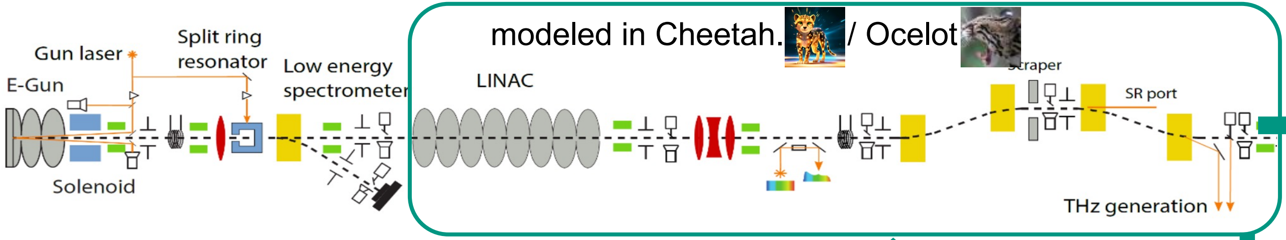
Pipeline for ML and control applications

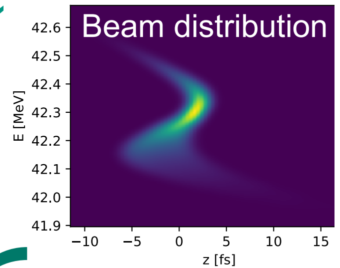


The ACSS (Accelerator Control and Simulation Services) allows for scheduling and orchestrating of multiple intelligent agents, training and tuning of ML models, handling of data streams and for software testing.

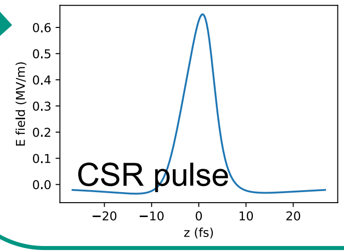
Using Reinforcement Learning for CSR radiation optimization at a Linac

Karlsruhe Institute of Technology

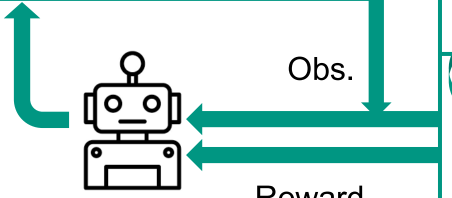




Bunch length σ_z



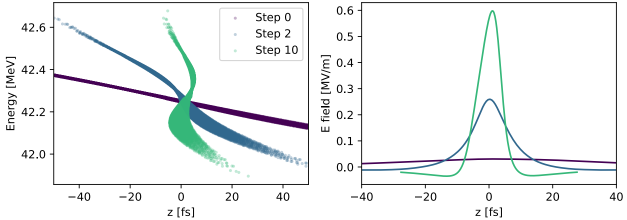
Peak field $E_{\max}$



Formulating the RL task

Observation	accelerator setting $[\phi_{RF}, \theta_{Chicane}]$ + beam parameters σ_z , $E_{\max}$
Action	$\Delta[\phi_{RF}, \theta_{Chicane}]$
Reward	• Bunch length: $\mathrm{elu}(\sigma_{threshold} - \sigma_z)$ • CSR field: $E_{\max}$

PPO agent trained in SB3

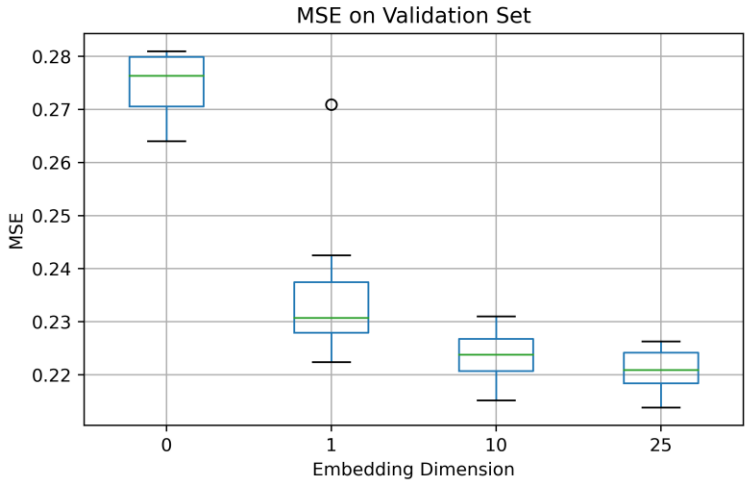


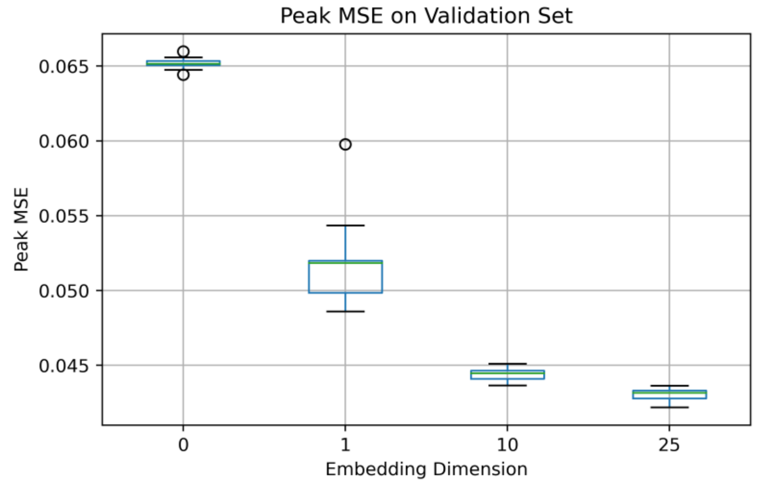
Improving Surrogate Model Performance for Sparse Outputs in the Spatial Domain

Accurate estimates of **where beam loss occurs** is important both for overall optimisation and personnel/machine protection

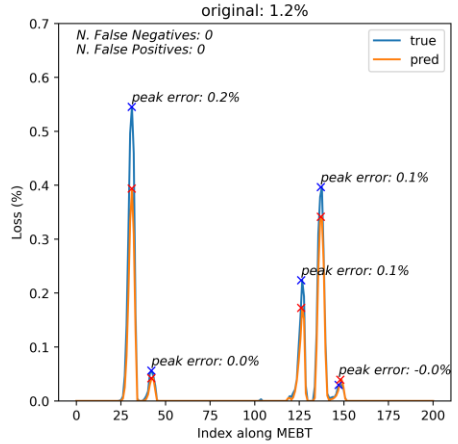
$$\gamma(x) = [cos(2\pi Bx), sin(2\pi Bx)]^T$$

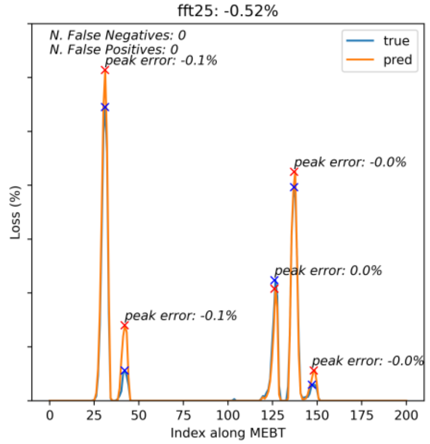
Where $B \in R^{m \times d}$, with m being the embedding dimension





Concatenating **1D-fourier feature mappings** (proven to be effective in 2D applications) of the spatial dimension with the inputs to the model to improve resolution

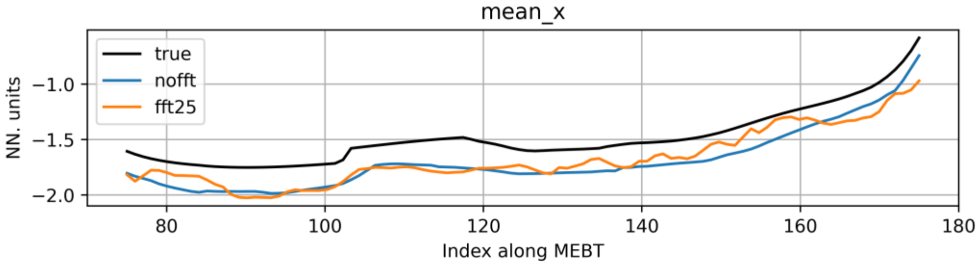




Increasing the Embedding Dimension (m) **decreases MSE and Peak Error**, as well as reducing errors between predicted peak maxima and the overall cumulative loss.

Also **improves stability** of peak predictions in sparse outputs

BUT **adds a high frequency component** to the output prediction which is non-physical for smoother functions.



Science and Technology Facilities Council

ISIS Neutron and Muon Source

www.isis.stfc.ac.uk

@isisneutronmuon

uk.linkedin.com/showcase/isis-neutron-and-muon-source

The Reinforcement Learning for Autonomous Accelerators collaboration

https://rl4aa.github.io/

Mission

- Connect RL enthusiasts in the particle accelerators community and share our experience.
- Teach fundamental RL concepts and show practical accelerator applications.
- Discuss the current challenges of developing RL algorithms for particle accelerators.
- Be a community!

Join our Discord server

Workshops every year

Reinforcement Learning for Intensity Tuning at Large FEL Facilities

4th ICFA Machine Learning Workshop

Jan Kaiser, Annika Eichler, Auralee Edelen, Daniel Ratner, Malachi Schram and Kishansingh Rajput

Gyeongju, 7 March 2024

HELMHOLTZ RESEARCH FOR GRAND CHALLENGES

Jefferson Lab

SLAC NATIONAL ACCELERATOR LABORATORY

DESY.

FEL Intensity Tuning at LCLS

Reinforcement learning-trained optimisation (RLO)

- **Maximise FEL intensity** using 14 quadrupole magnets
- Challenges:
  - Slow simulation -> **Cheetah** and **neural network surrogate modelling**
  - High dimensionality and easy failure -> **Curriculum learning**

Ultimately also transfer to European XFEL

RL Agent
(Stable Baselines3)

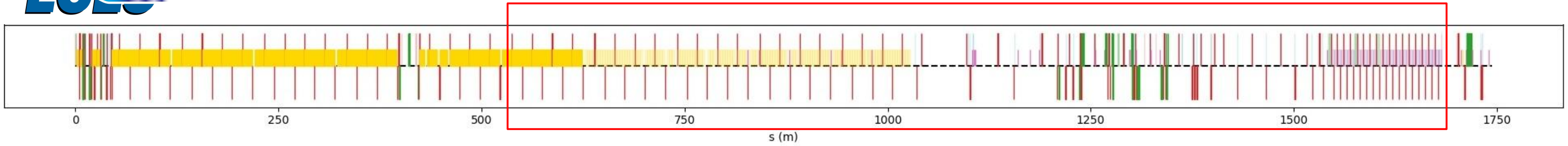
Action
(quadrupole settings)

Reward
(based on intensity)

Observation
(Quadrupole settings + intensity)

Backends: Bmad EPICS

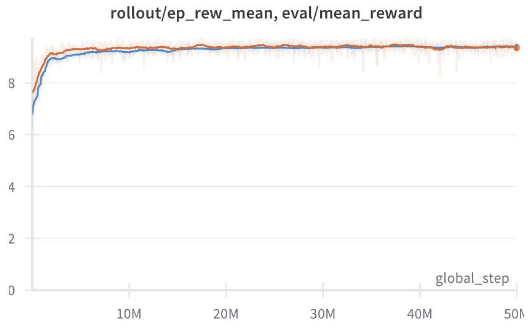
LCLS

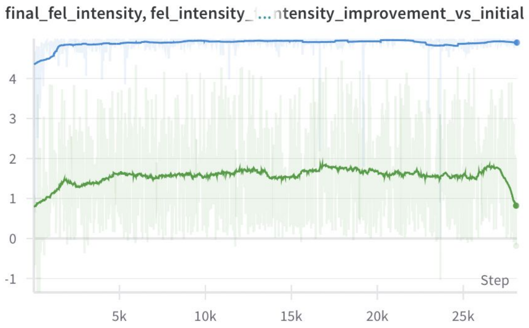


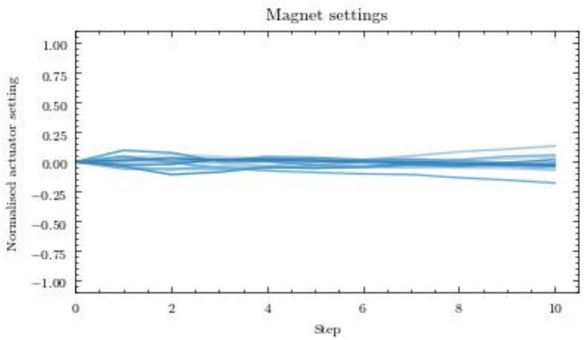
Results with PPO

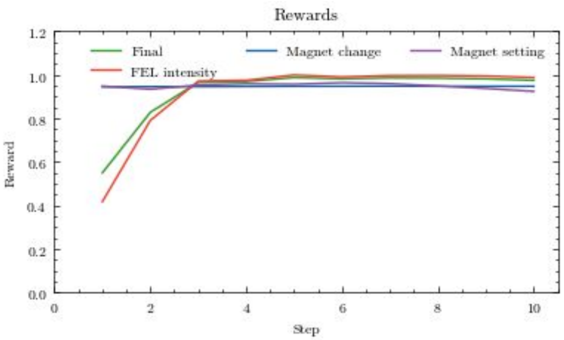
It works!

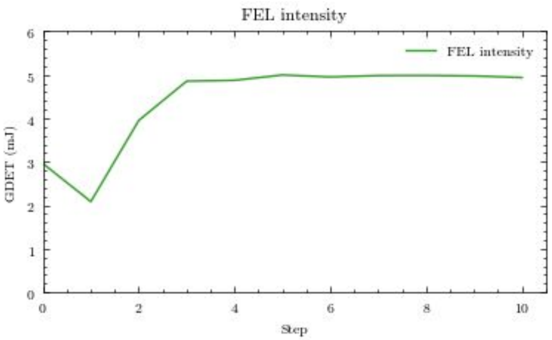
- **Proximal Policy Optimisation (PPO)** algorithm from Stable Baselines3
- Training for **50 Million** environment interactions
- **1 day 16 hours** on a HPC cluster node





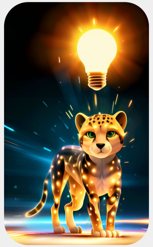






Gradient-based Reinforcement Learning

Significantly improving sample-efficiency and reducing training times

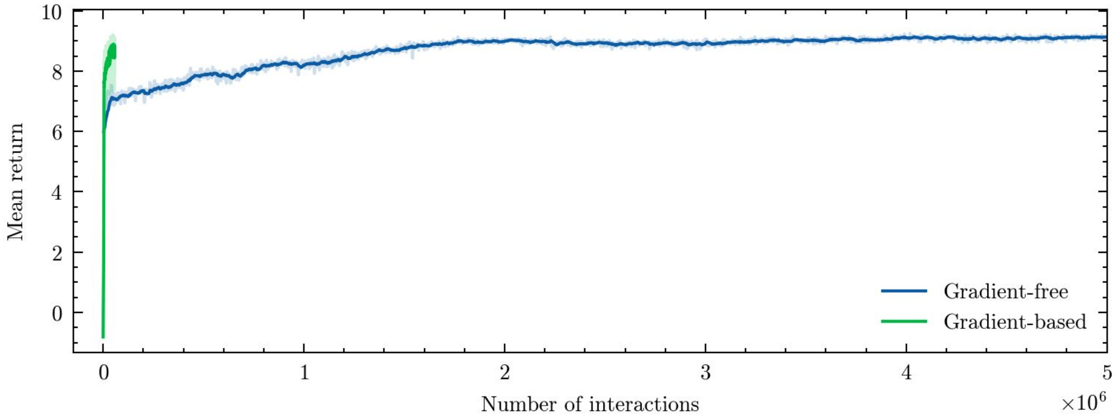


We are using **Cheetah!**

Cheetah supports **automatic** differentiation.

→ **Gradient-based RL** with true policy gradient.

→ Achieve same performance in **45x fewer samples**.



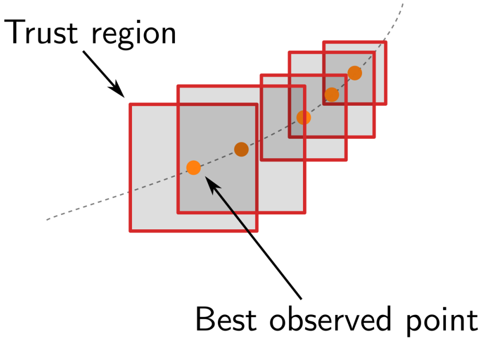
Contact

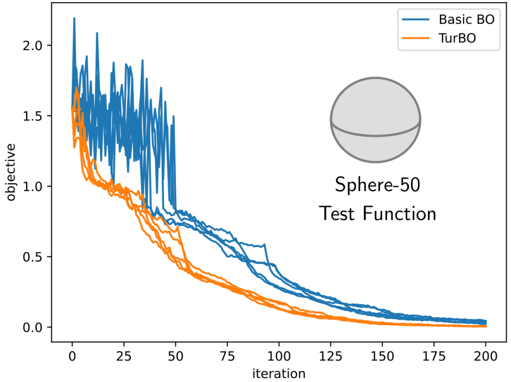
DESY. Deutsches
Elektronen-Synchrotron

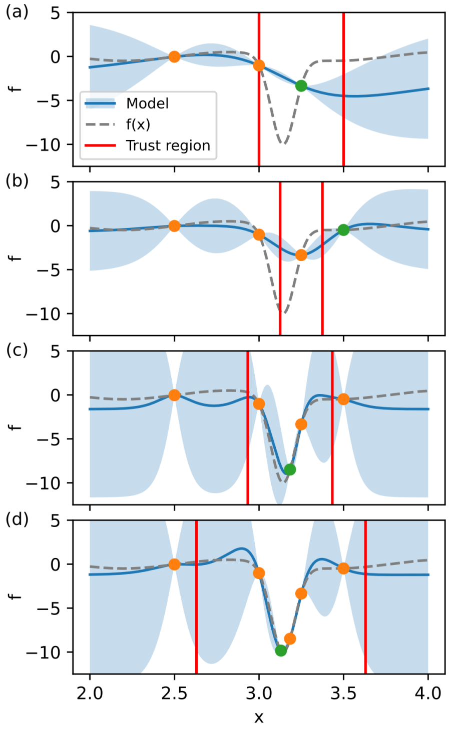
www.desy.de

Jan Kaiser
Machine Beam Controls (MSK)
jan.kaiser@desy.de

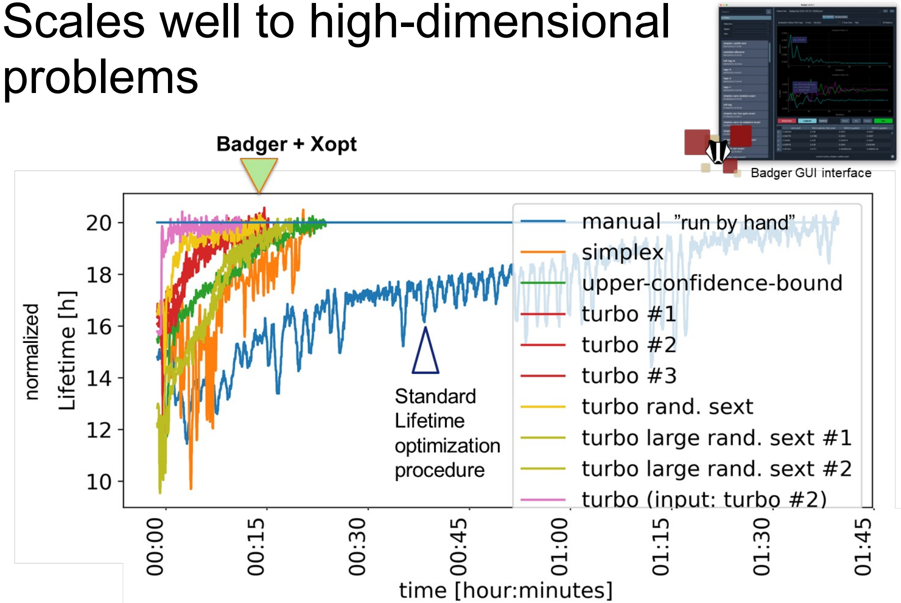
Trust Region Bayesian Optimization for Online Accelerator Control







Trust region BO (TurBO):
Global → local BO optimizer
Scales well to high-dimensional problems



Development of beam transport system optimization method using VAE and Bayesian optimization

Y. Morita, T. Nishi, T. Nagatomo, Y. Nakashima

Ultimate goal

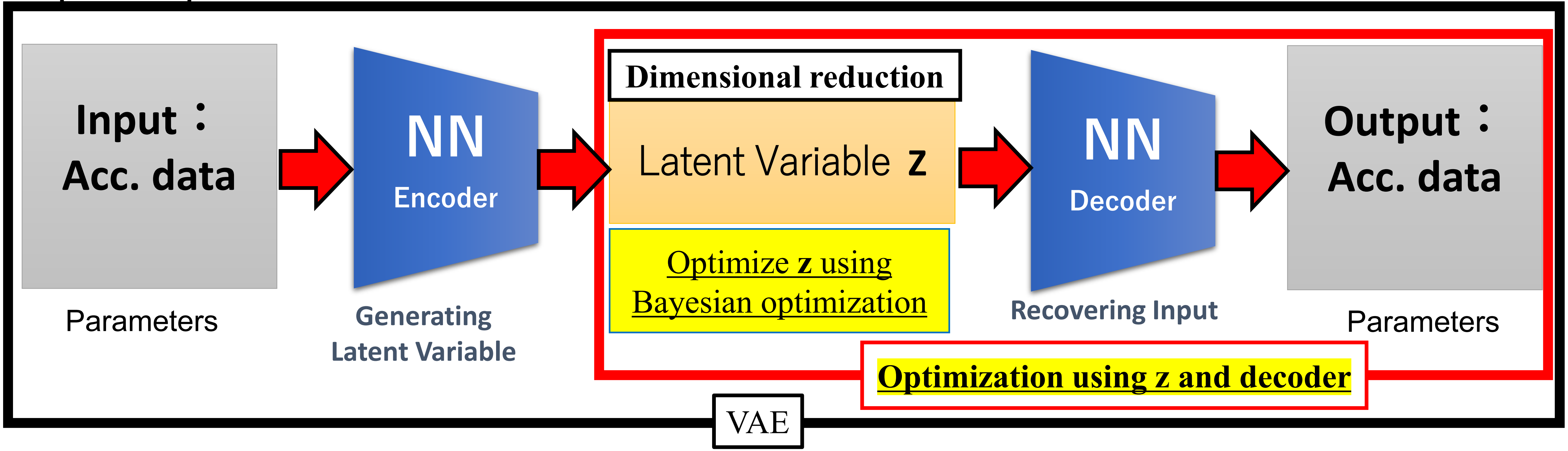
Optimize the entire accelerator facility at once

But …

There are too many parameters

Reduce the number of parameters using VAE

Method



We have already started experiments with Low Energy Beam Transport!

Simulation methods of 3D coupled storage ring based on SLIM formalism

ZHAO Jingyuan, TANG Chuanxiang, DENG Xiujie, PAN Zhilong, LI Zizheng, CHEN Liwei, Tsinghua University, Beijing, China
Alexander Chao, Tsinghua University, Beijing, China, also at Stanford University, Stanford, USA

SLIM Formalism

SLIM is a linear storage ring beam dynamic formalism based on transport matrix and eigen-analysis. It can self-consistently analyze linear coupled/uncoupled storage ring and give the following results without using any courant-snyder auxiliary functions.

- All the linear dynamics
- Closed orbit disortion
- Equilibrium beam size and shape

SLIM Formalism

$M(s_2|s_1)_{6\times6}$ → eigenvectors → linear dynamics

Courant-Snyder Formalism

$M(s_2|s_1)_{2\times2}$ → auxiliary functions → linear dynamics

We have extended SLIM and processed all elements with thick lens analysis, which speeds up SLIM code calculation. SLIM can thus be used as a linear self-consistent physical computing core for MOGA and machine learning.

Generalized longitudinal strong focusing

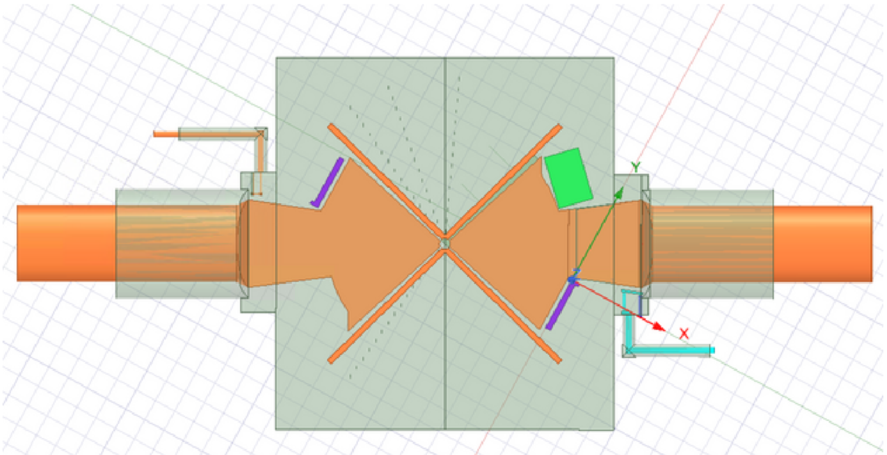
Generalized longitudinal strong focusing (GLSF) scheme aims to produce coherent EUV radiation turn by turn in laser-driven storage rings. It invokes transverse-longitudinal coupling and then attains a short bunch length with significantly reduced modulation laser power.

We present a method for analyzing local lattice design with SLIM, and combine SLIM and MOGA to carry out lattice design of GLSF. Finally, we achieve a bunch length of less than 5nm at the storage ring radiator.

Multiobjective Optimization of Cyclotron Cavity Model using Neural Network

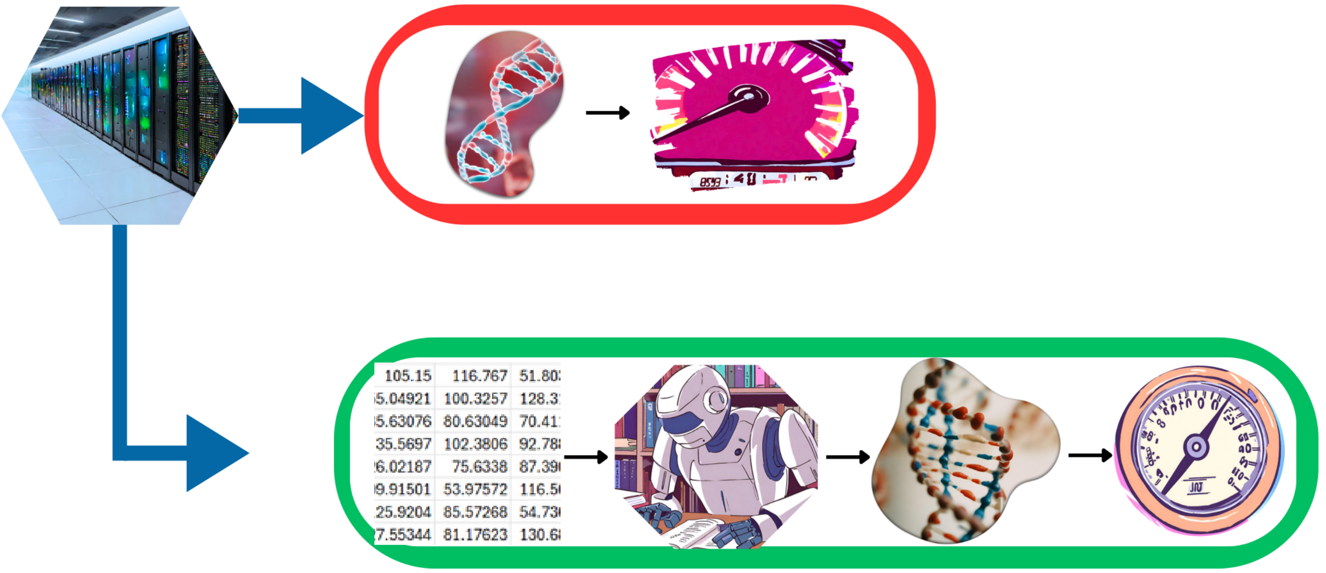
Ahsani Hafizhu Shali, Takafumi Hara, Tetsuhiko Yorita, Hiroki Kanda, Mitsuhiro Fukuda

Research Center for Nuclear Physics, Osaka University

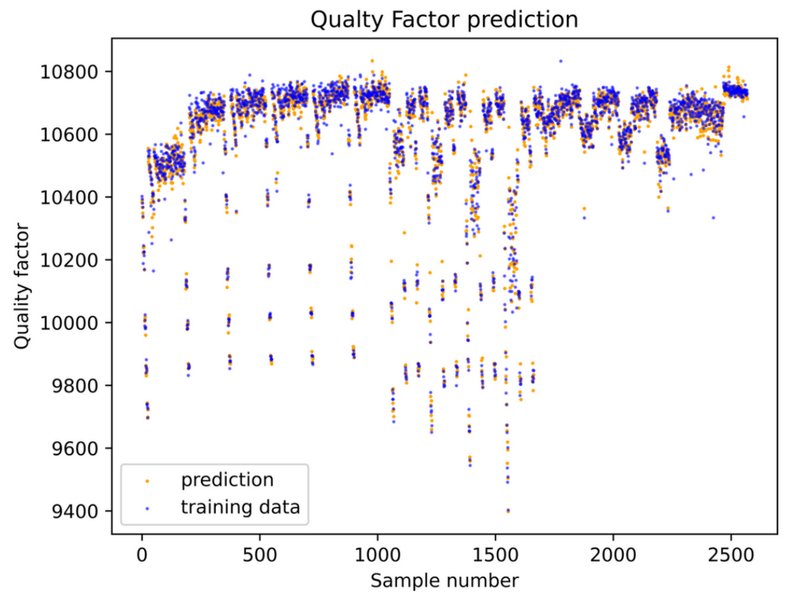


Smaller components modification to maximize Q factor and minimize return loss

Reducing the computational cost for cavity optimization



Variations limited to smaller components, should not affect beam dynamics



NN model could recognize the pattern for Q factor, frequency and return loss of a cavity, with relatively small error.

Coupler & compensator dimension and position

NN is trained to replace FEM simulation, for the specific case of RCNP AVF cyclotron cavity

Input – 6 Neurons → 120 Neurons - ReLU → 110 Neurons - ReLU → 100 Neurons - ReLU → 90 Neurons - ReLU → 80 Neurons - ReLU → 70 Neurons - ReLU → 60 Neurons - ReLU → 50 Neurons - ReLU → 40 Neurons - ReLU → 30 Neurons - ReLU → 20 Neurons - ReLU → 10 Neurons - ReLU → Output – 3 Neurons

10000 epochs

2569 samples are used for training

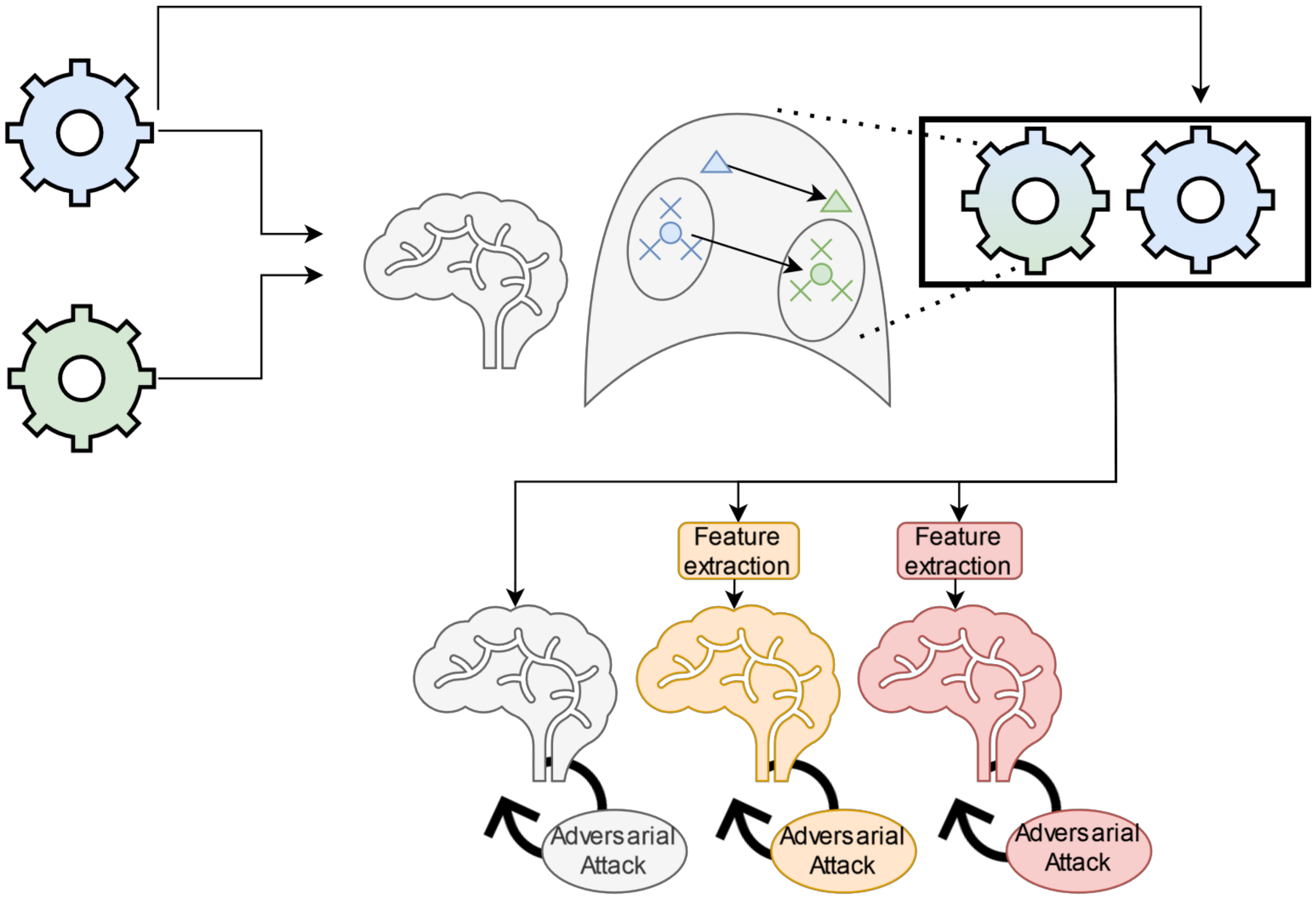
Research on Recognition of Quench and Flux Jump Based on Machine Learning

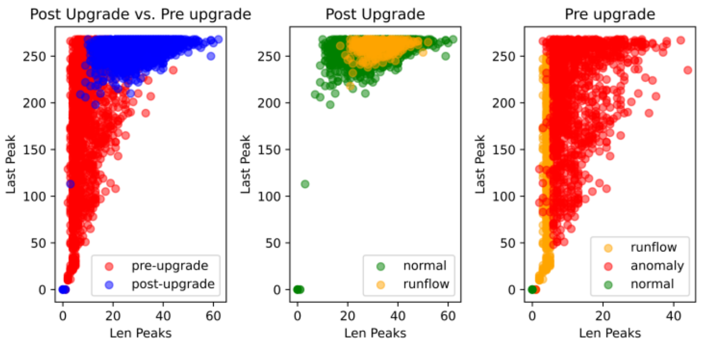
The Institute of Modern Physics is developing the Fourth generation of Electron Cyclotron Resonance (FECR), which requires Nb3Sn superconducting hexapole magnets with higher magnetic fields and composite structures. For Nb3Sn superconducting magnets, they exhibit significant thermal magnetic instability, known as "flux jump". This characteristic can generate random voltage spikes during the excitation process of the magnet, leading to misjudgment of the Quench Detection System (QDS) and seriously affecting the normal operation of FECR.

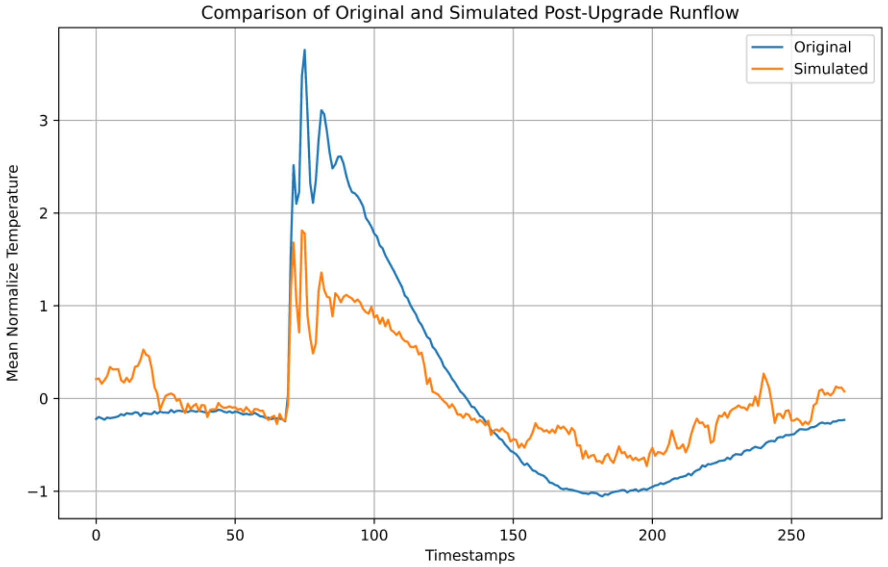
To solve this problem, this study uses machine learning algorithms and aims to build a simplified and efficient recognition model to effectively distinguish the phenomenon of overshoot and flux jump during the excitation process of Nb3Sn magnets. Based on the voltage data obtained from multiple excitation processes of Nb3Sn superconducting hexapole magnets,

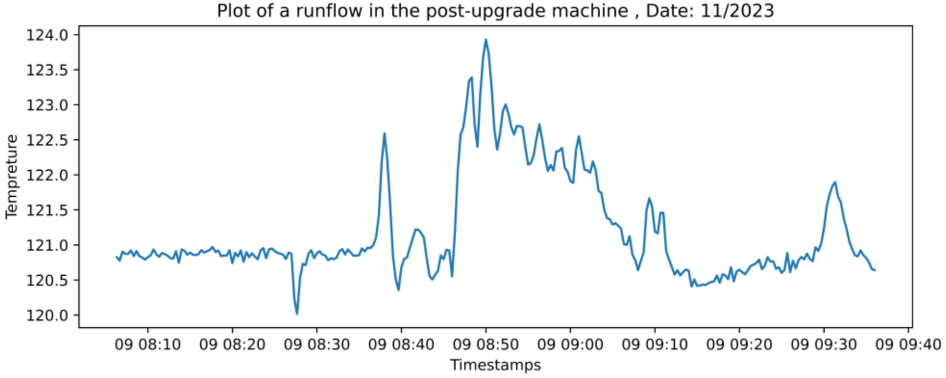
this paper extracted 27 quench samples and 25 flux jump samples, and extracted 33 features from each sample. Multiple machine learning algorithms were used to train and construct these data, and the accuracy of different algorithms was compared to ultimately explore the best recognition model. The experimental results show that the model only uses 5 features and achieves 100% classification accuracy on linear kernel SVM. By using this machine learning model, high accuracy and computational speed have been achieved in the recognition of magnetic flux jump and quench, which can provide reference for the optimization of subsequent FECR quench detection algorithms.

Analysis and Improvement of Generalisability of Anomaly Detection Methods









Science and Technology Facilities Council

ISIS Neutron and Muon Source

www.isis.stfc.ac.uk

@isisneutronmuon

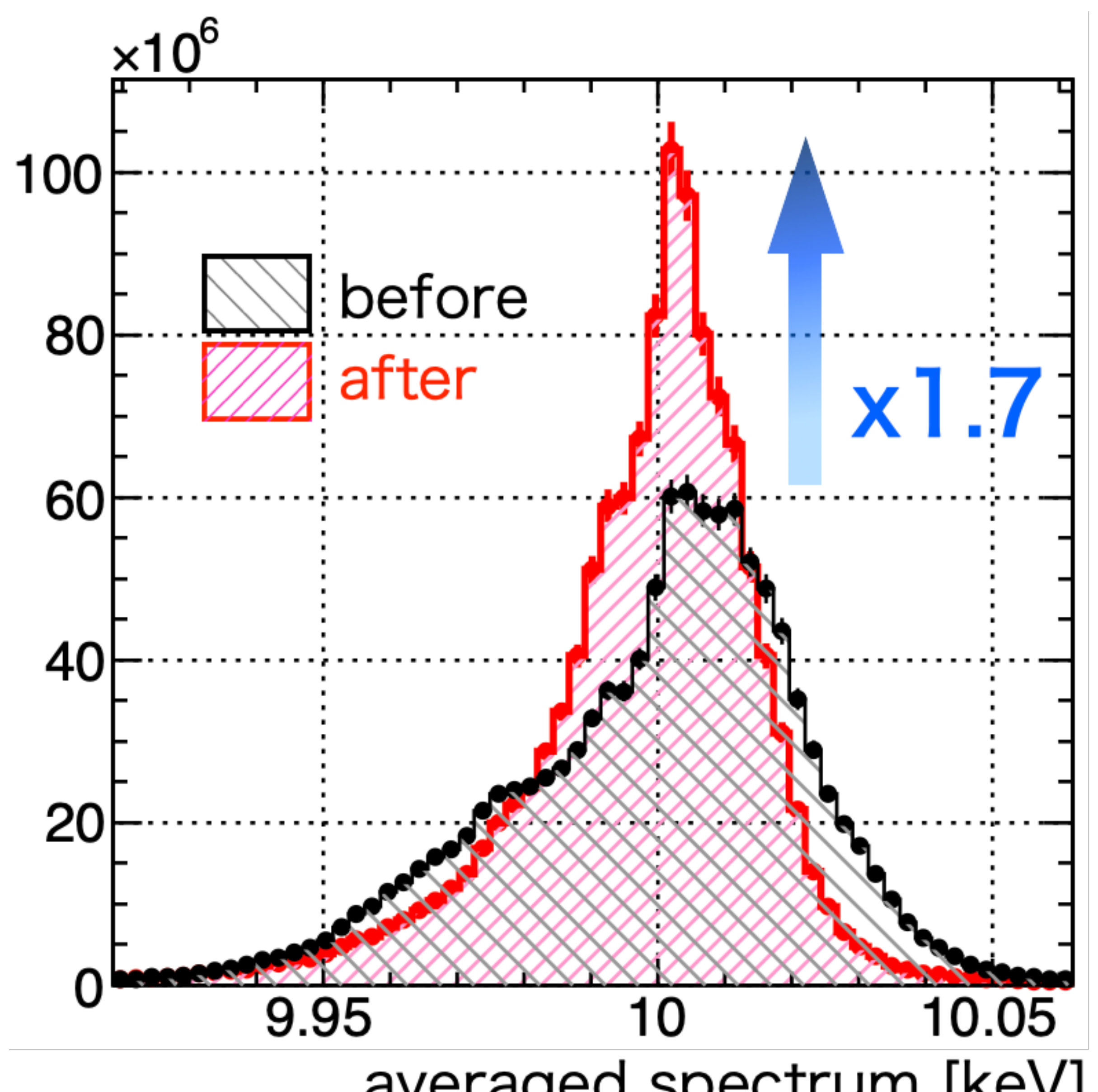
uk.linkedin.com/showcase/isis-neutron-and-muon-source

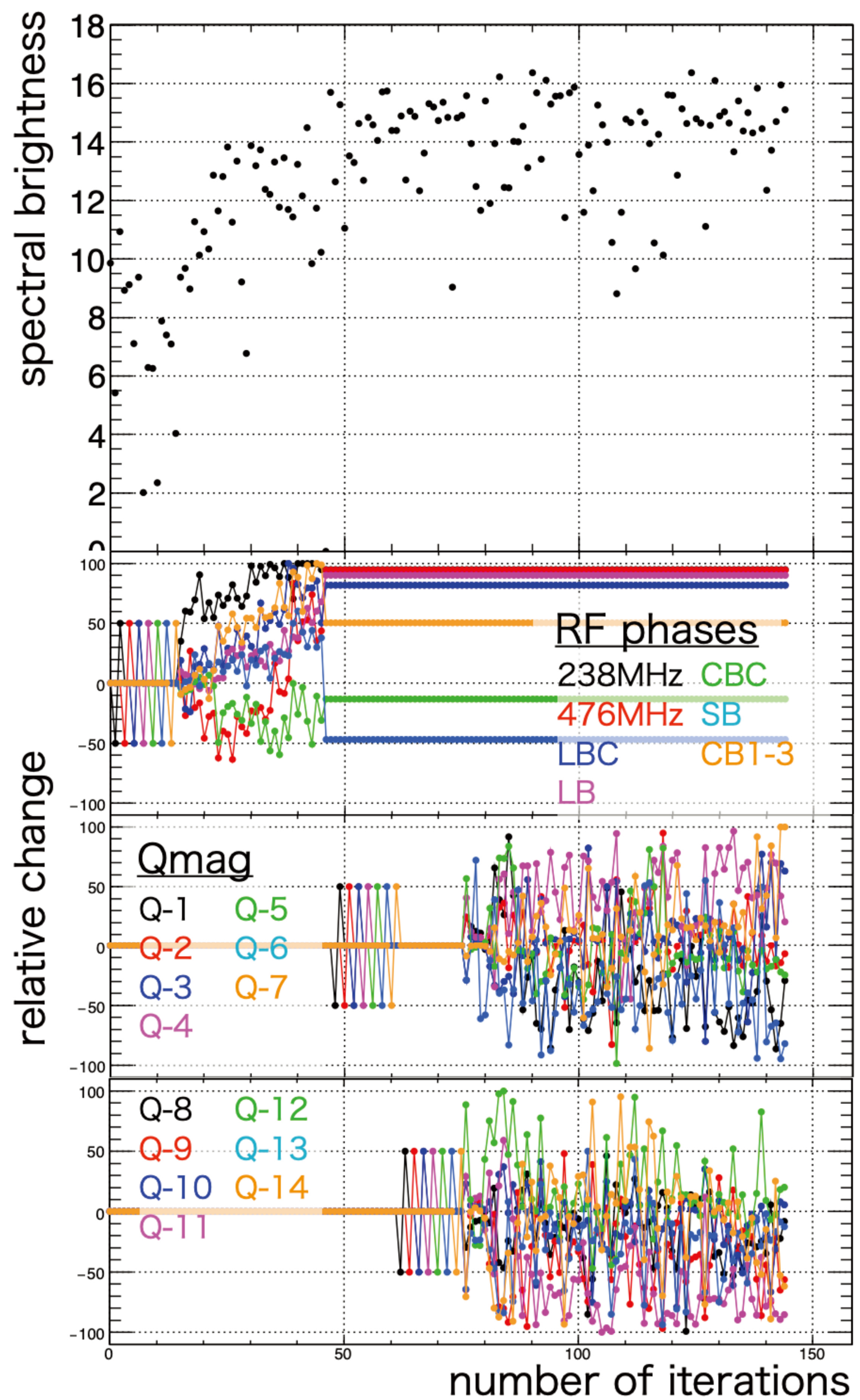
Application of Machine Learning to Accelerator Operations at SACLA/SPring-8

Hirokazu Maesaka[1,2], Eito Iwai[2,1], Ichiro Inoue[1]   1: RIKEN SPring-8 Center, 2: Japan Synchrotron Radiation Research Institute (JASRI)

XFEL Optimization

- We developed a Gaussian Process (GP) optimizer for XFEL.
- The optimizer succeeded in maximizing the XFEL pulse energy.
- We recently developed and installed a new high-resolution inline spectrometer.
- The new spectrometer enabled us to optimize the spectral brightness.
- The spectral brightness was improved by a factor of 1.7 over the pulse energy optimization.

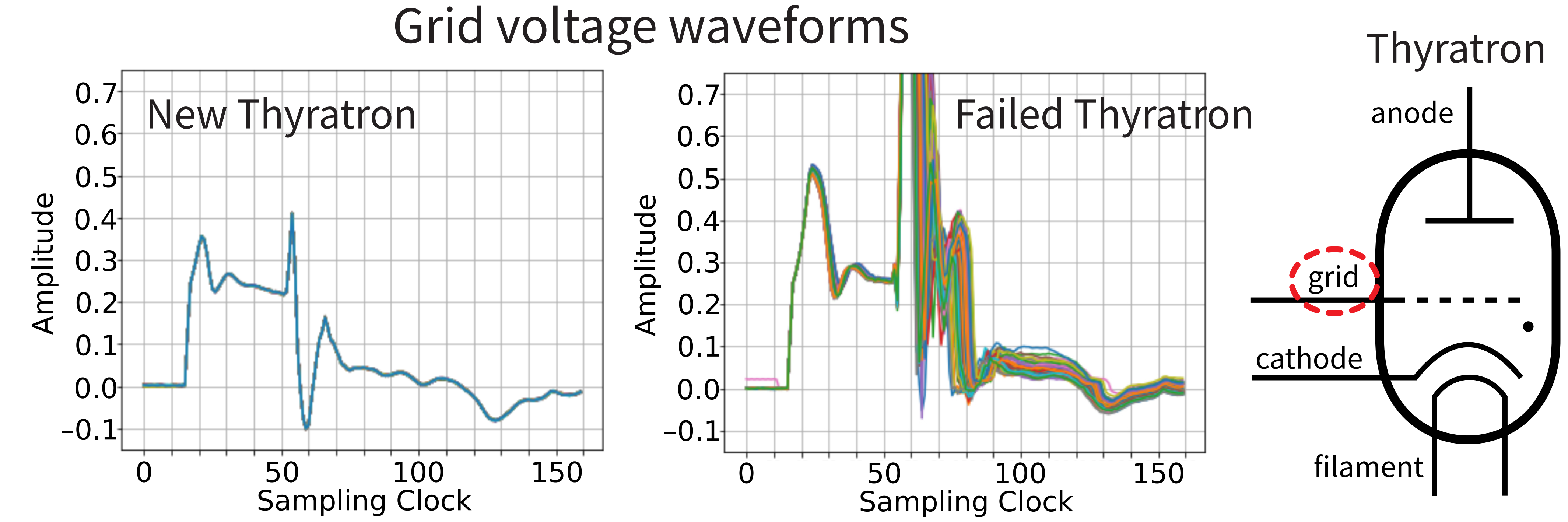




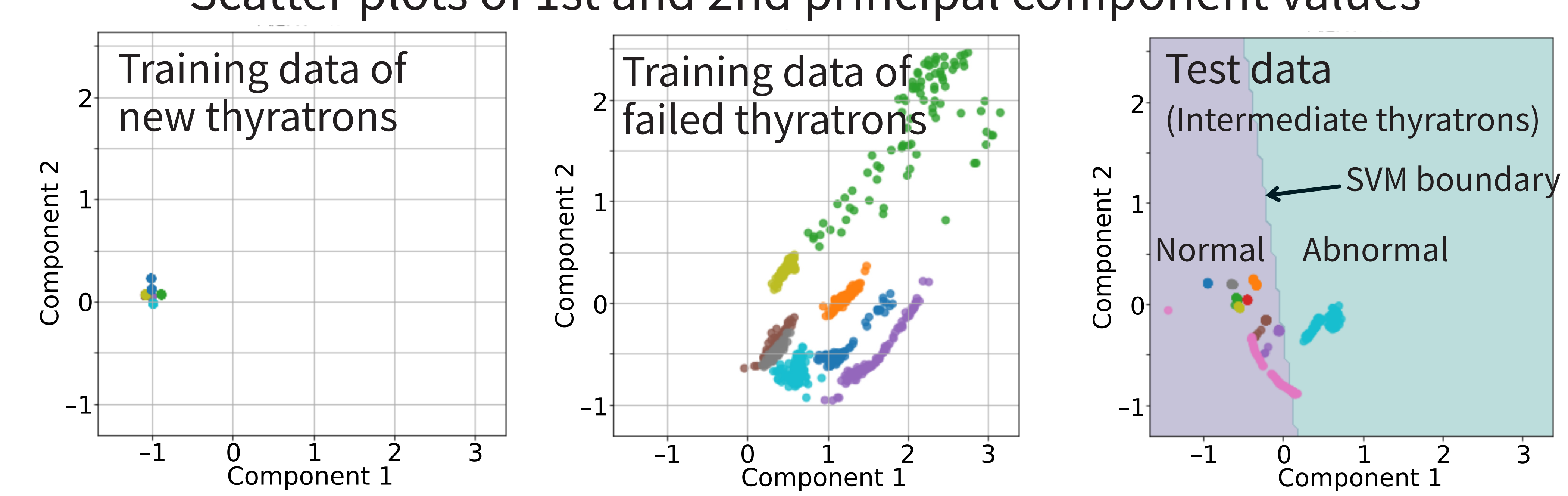
Failure Prognosis of Thyratrons

- A thyratron is a high-power switch to drive a pulsed klystron.
- The grid voltage waveform varies with its operation time.
- We are developing a health check algorithm by using the grid waveform.
- We applied principal component analysis (PCA) to reduce the dimensionality of the waveform data and support vector machine (SVM) to classify thyratrons into normal or abnormal.
- This system can emit warning if the grid waveform enters the abnormal area of the SVM result.

Grid voltage waveforms



Scatter plots of 1st and 2nd principal component values



ONLINE OPTIMISATIONS OF LIFETIME AND INJECTION EFFICIENCY IN THE ESRF EBS STORAGE RING

N. Carmignani, L. Carver, L. Hoummi, S. Liuzzo, T. Perron, S. White, ESRF, Grenoble, France

4th ICFA Beam Dynamics Mini-Workshop on Machine Learning for Particle Accelerators
Gyeongju, South Korea

Design of EBS:
Multi-objective optimisations of sextupoles and octupoles

Operation of EBS:
Online optimisations to make use the independent sextupoles and octupoles power supplies

Which KNOBS should we use for online optimisations?

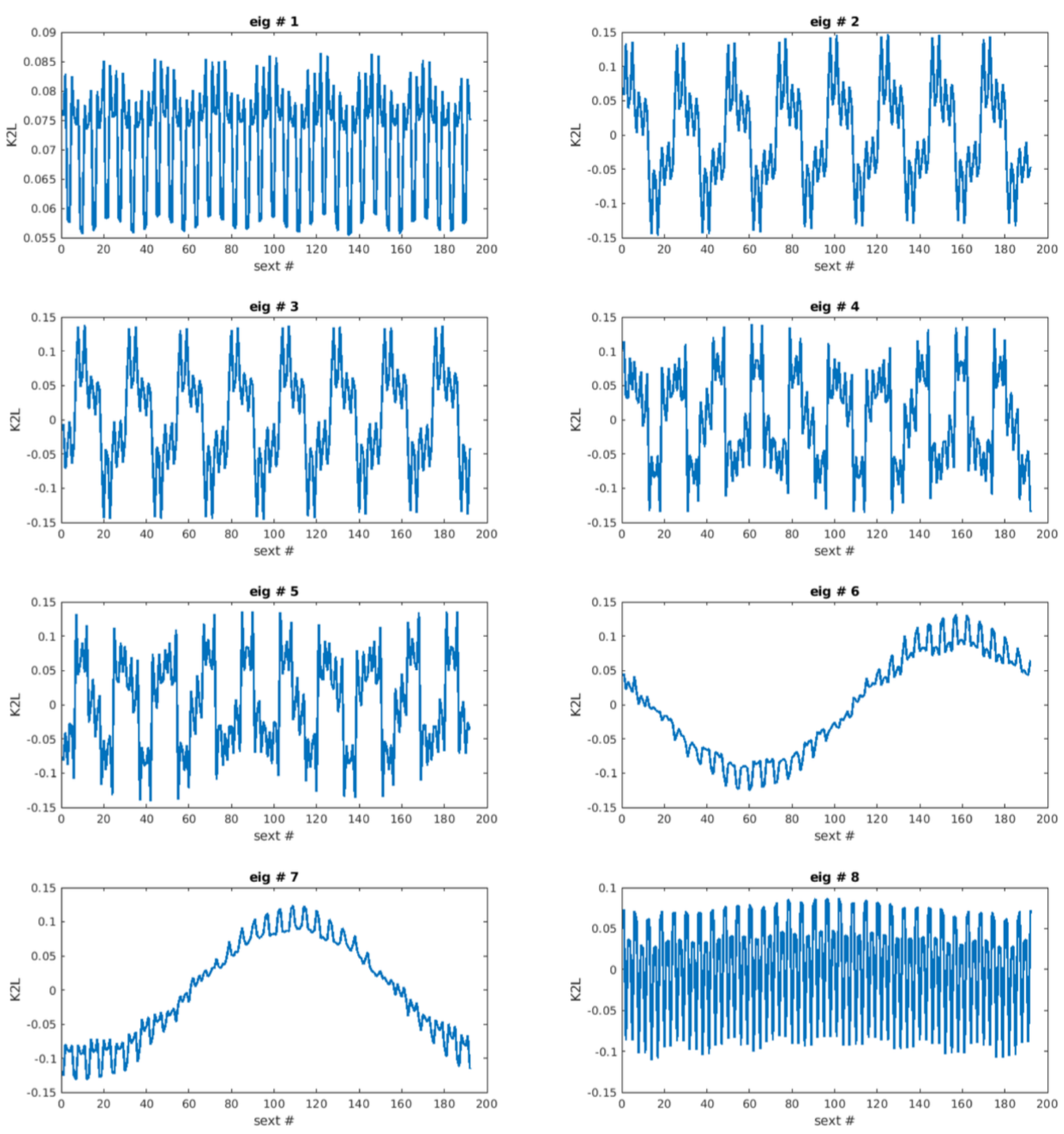
The idea is that with sextupoles we can correct off-energy linear optics.

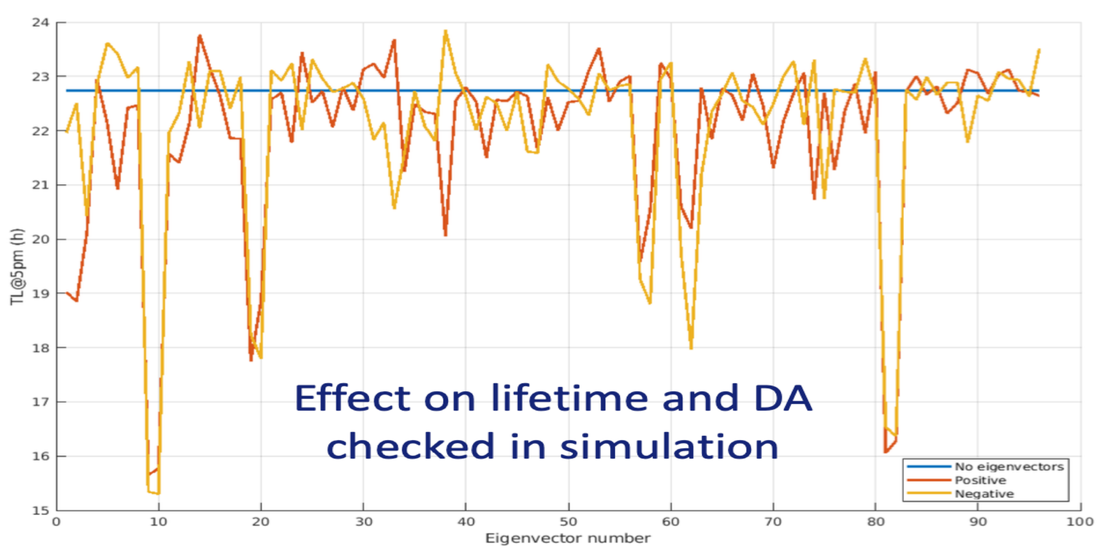
We can define some pseudo-sextupolar singular vectors:

$$J_{quad} = \frac{\delta ORM}{\delta K_{quad}}$$

$$K_{quad} \propto 2K_{sext}\eta_h$$

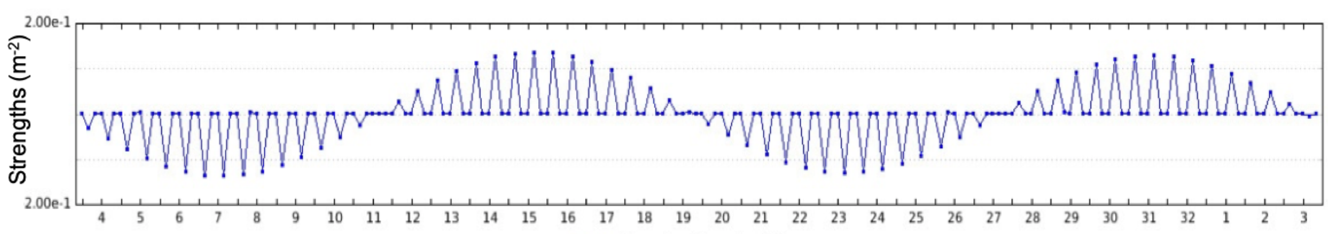
$$J_{sext} = J_{quad} \cdot 2\eta_h$$





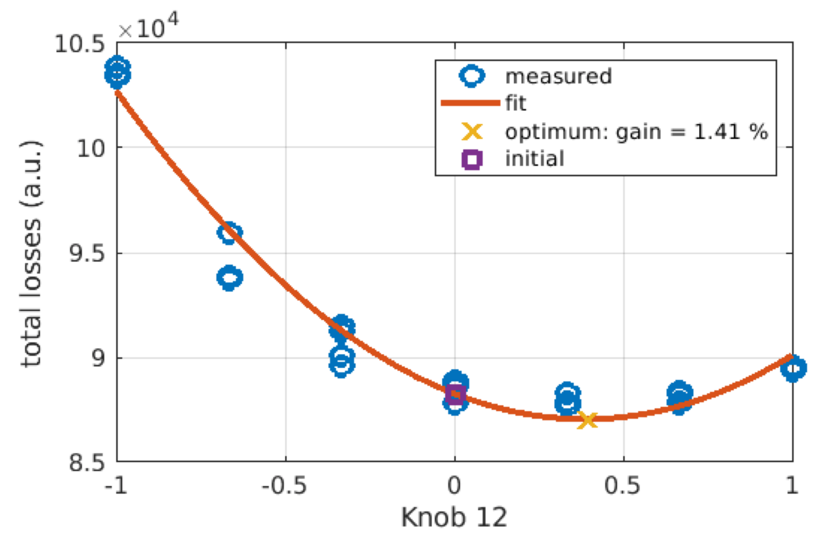
Other knobs tested:

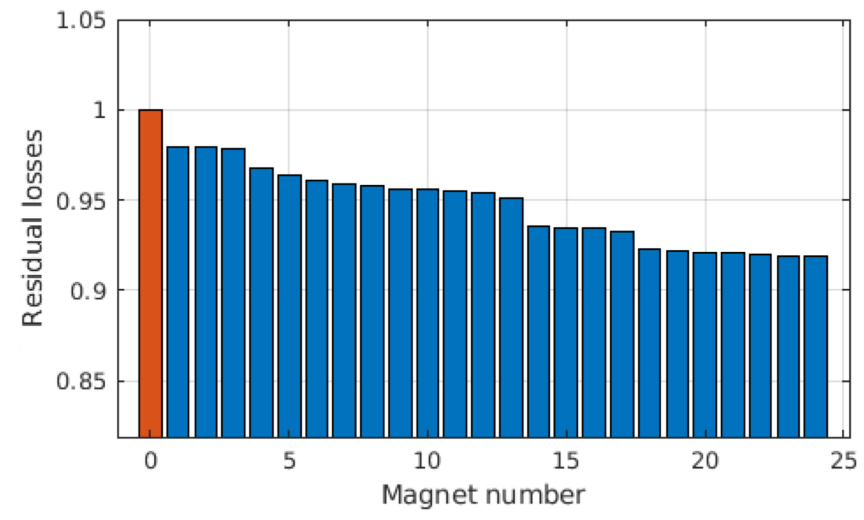
- Sine waves of sextupoles
- Sine waves of octupoles
- Skew quads eigenvalues for coupling

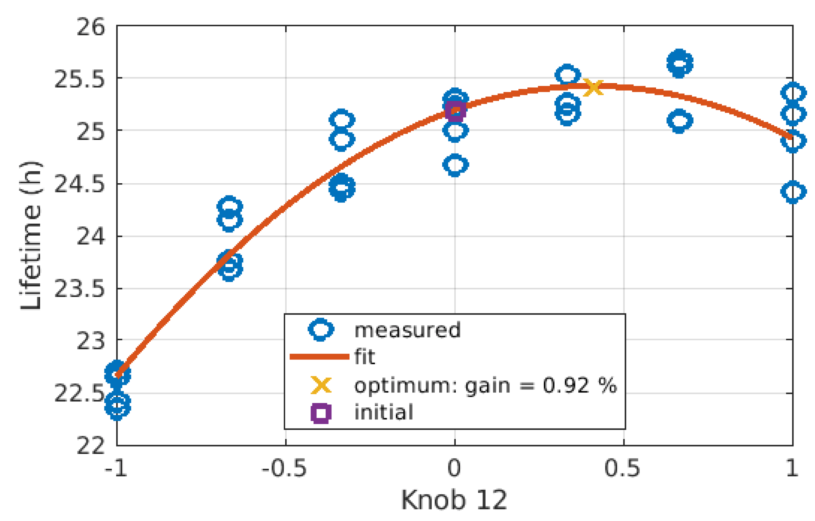


First optimizer

Matlab based code. It can scan knobs of all kind of magnets, minimizing losses measured with BLM.



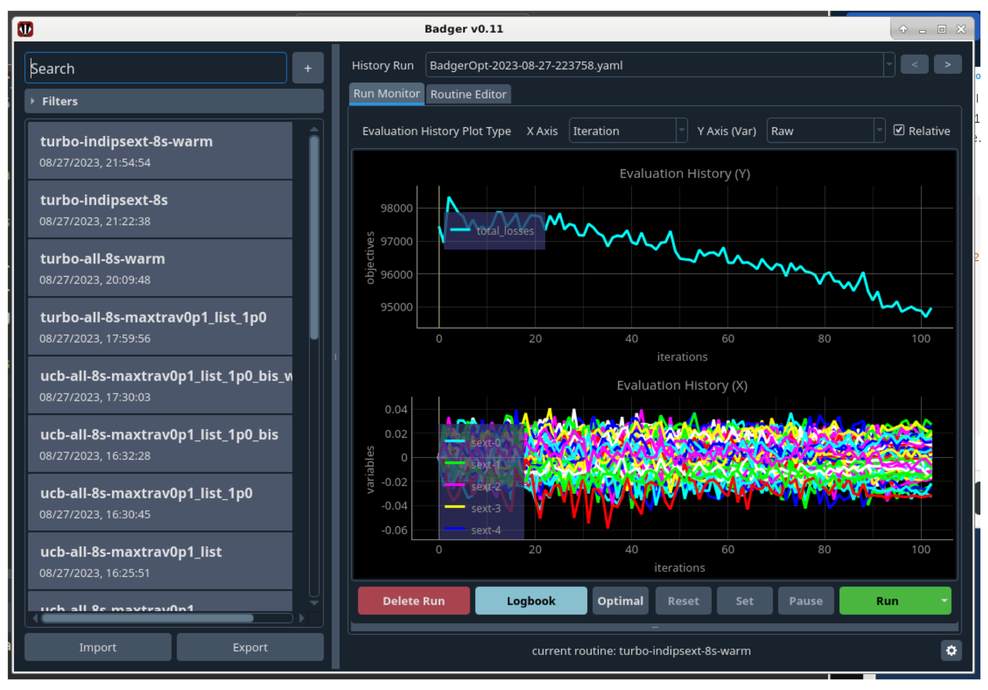




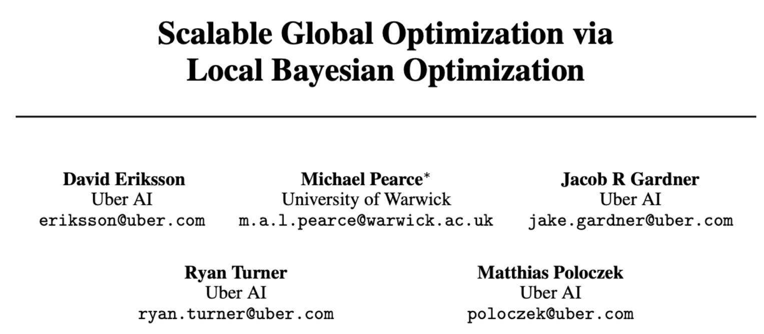
24 sextupolar knobs and 4 octupolar knobs selected

Badger + Xopt

Developed at SLAC

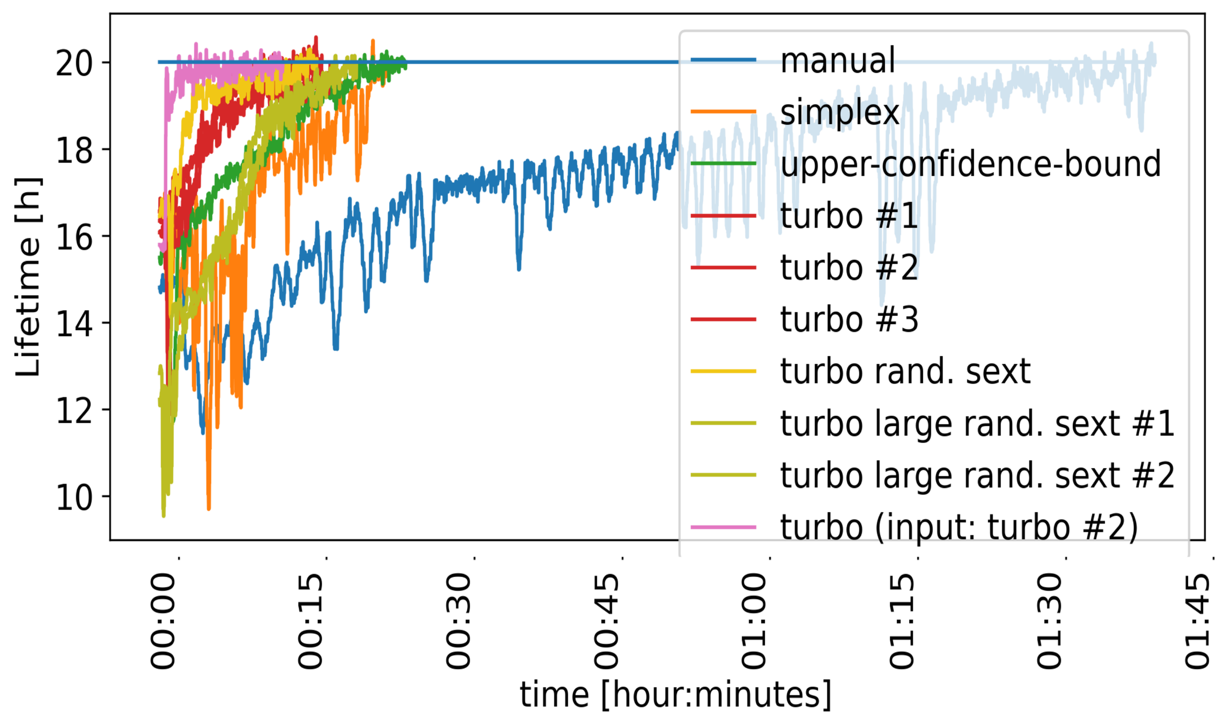


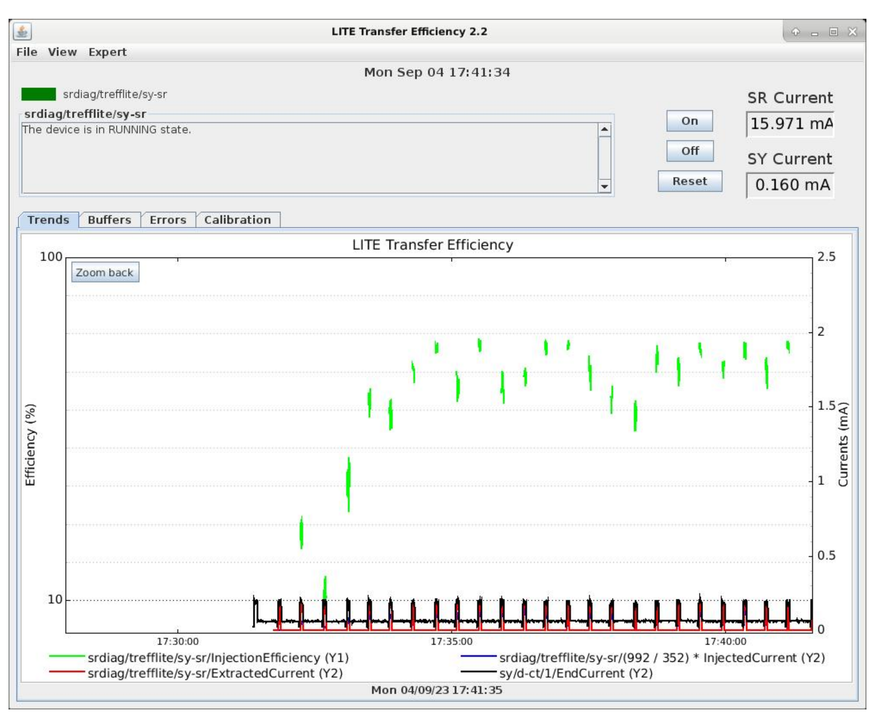
EASY INTERFACE
EASY SETUP
EASY INSTALLATION
MANY OPTIMIZERS


Scalable Global Optimization via Local Bayesian Optimization

David Eriksson, Uber AI; Michael Pearce, University of Warwick; Jacob R Gardner, Uber AI; Ryan Turner, Uber AI; Matthias Poloczek, Uber AI

Different optimisers tested within badger. Trust Region Bayesian Optimization (TuRBO) selected to be the most effective and fast.

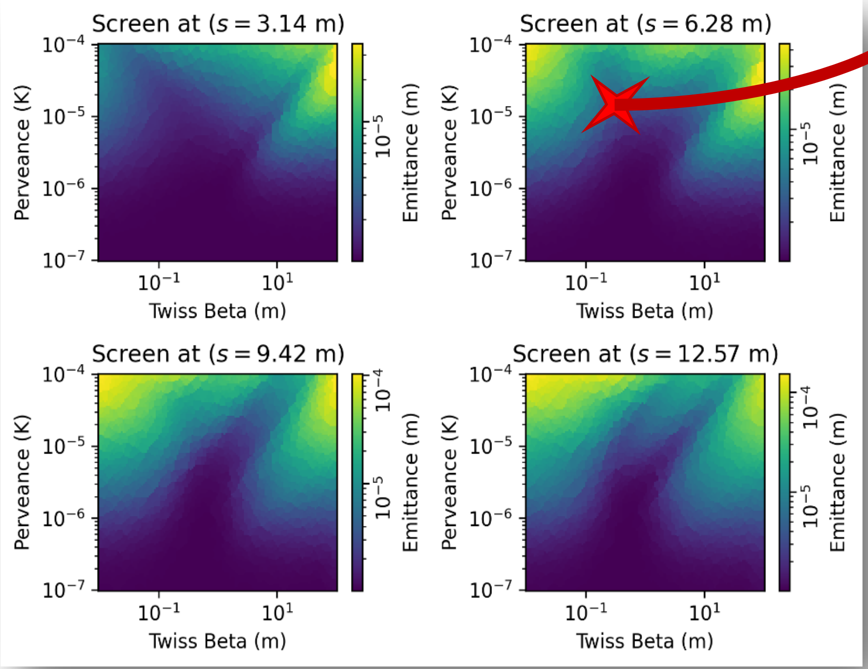




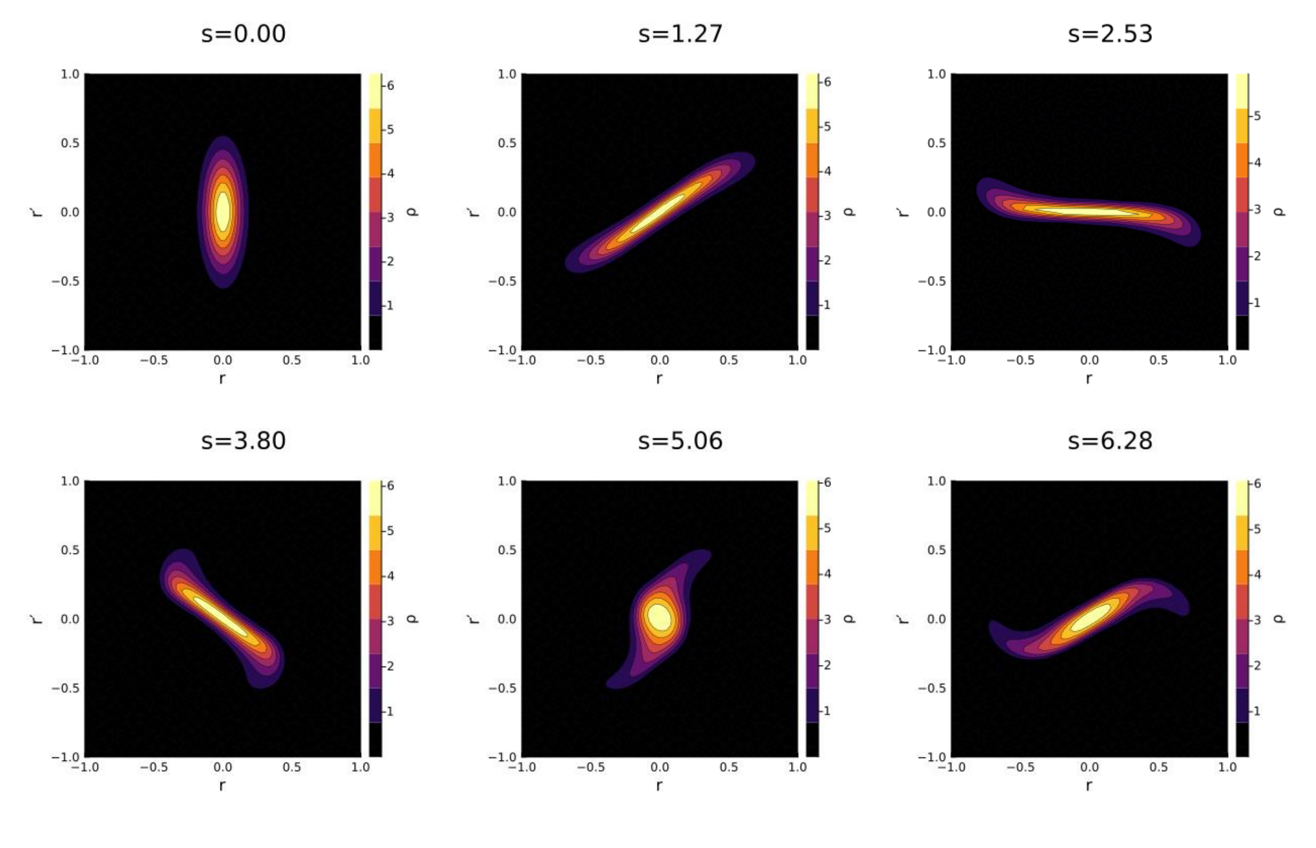
Badger is also used to optimise injection efficiency.
Injection elements, transfer line magnets, timing parameters can be optimized to maximize the injection efficiency measured with 8 injection shots.

Physics-Informed Surrogates Enhance Data-Driven Models with Knowledge of Physical Systems

Uniform Focusing Channel Dataset

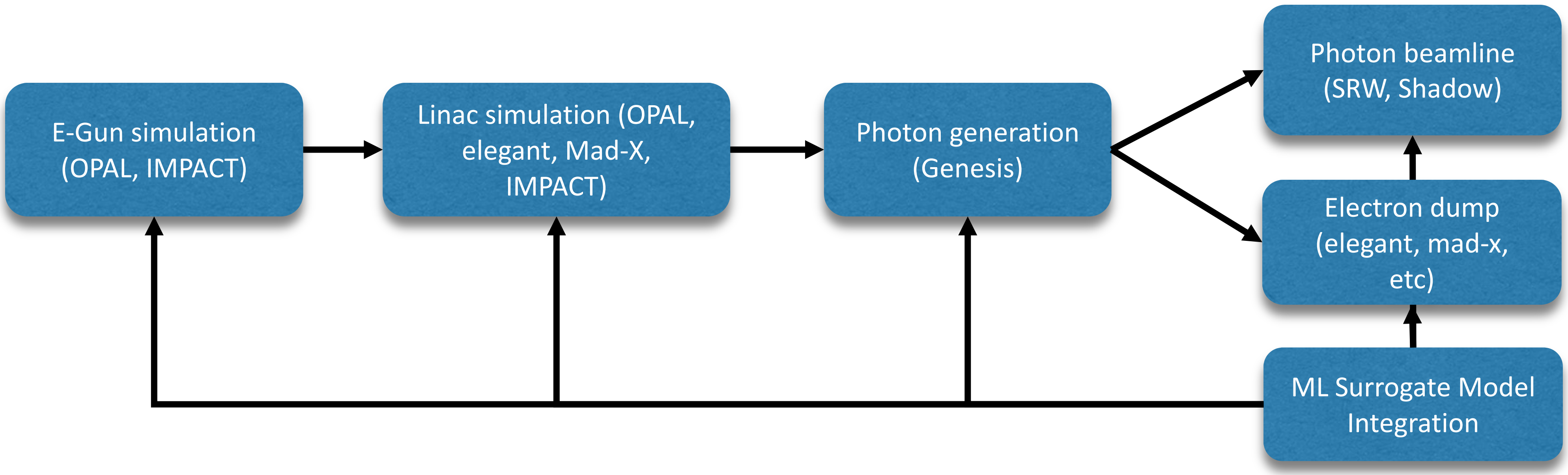


Trained PINN (K = 1.6e-5, Beta = 0.34)

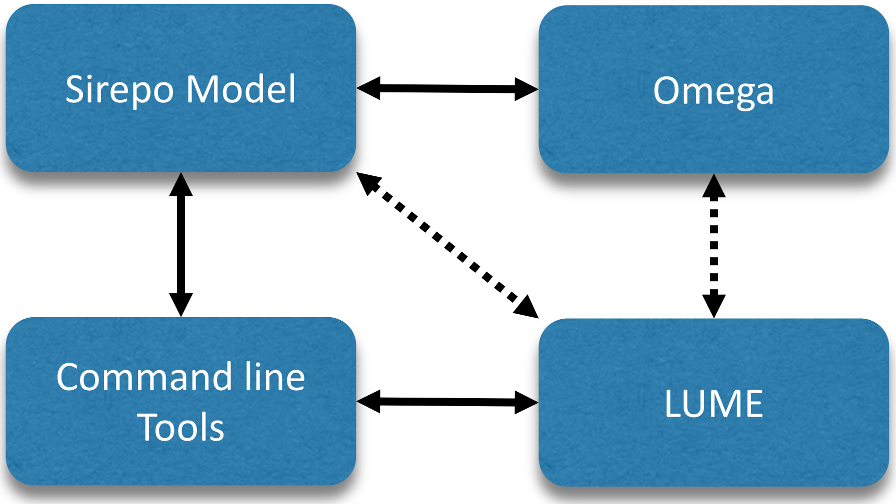


PINNs can be trained in challenging high space charge regime. Physics-informed priors significantly improve the accuracy of the surrogate model.

End-to-End Simulations and ML infrastructure for Light Sources



- LUME and Sirepo provide complementary integration tools
- Creating interoperability between these tools will improve the end-to-end simulation infrastructure
- Integrate machine-learning model infrastructure into Sirepo
- Demonstrate model deployment utilizing photon beamline test bench
- Provide integration with optimization workflows and controls



Supervised learning for nonlinear corrections in the LHC

Current state of nonlinear commissioning in the LHC **time consuming and iterative**

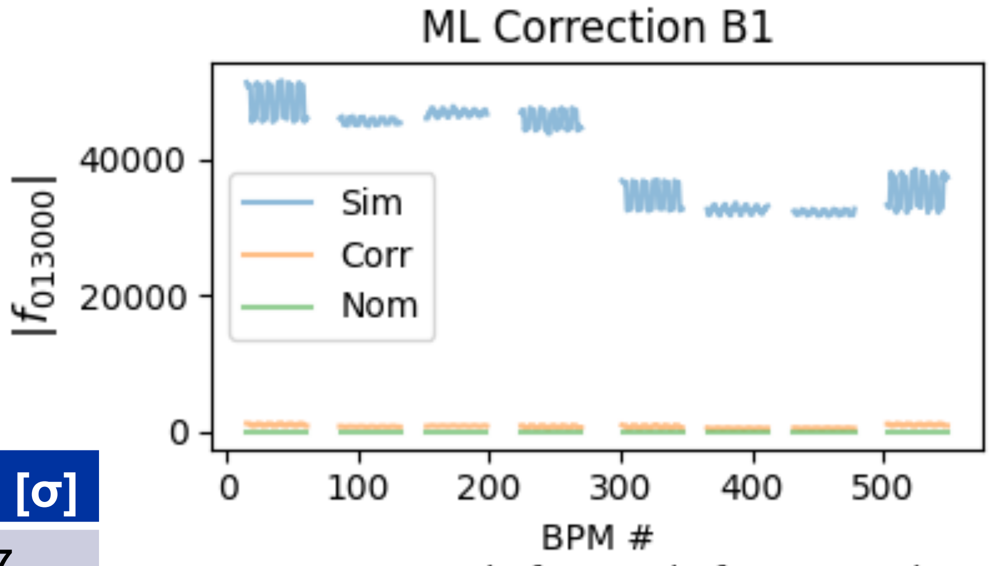
Can ML be used to correct multiple order errors at once using resonance driving terms (RDTs)?

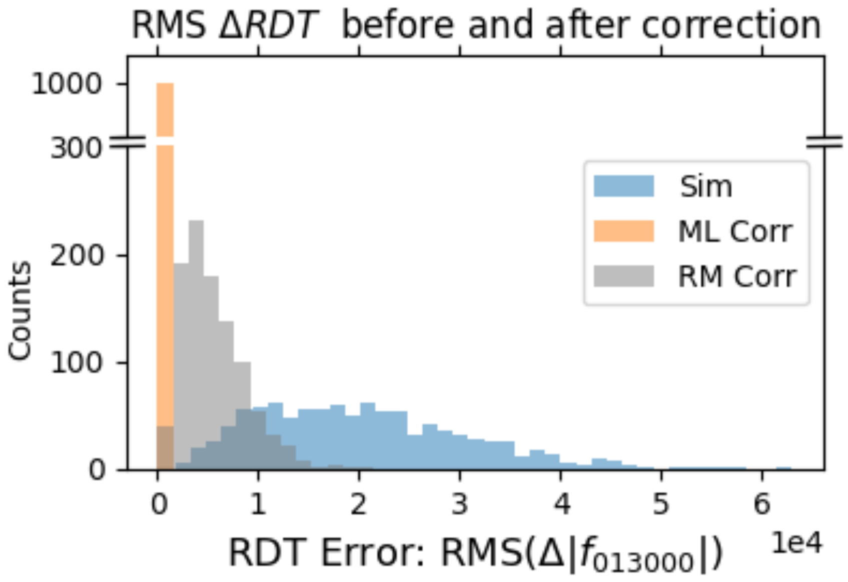
Using MADNG to generate up to 30k RDT samples with random errors 82 times faster than PTC!

Set	R2	MAE [σ]
Train	0.904	0.157
Test	0.883	0.173

Quadratic polynomial regression allows to model nonlinearities and collinearity in the variables

Performance yields better results than a traditional response matrix method for simultaneous correction of RDTs





Towards Natural Language-driven Autonomous Particle Accelerator Tuning

4th ICFA Machine Learning Workshop

Jan Kaiser, Annika Eichler and Anne Lauscher

Gyeongju, 7 March 2024

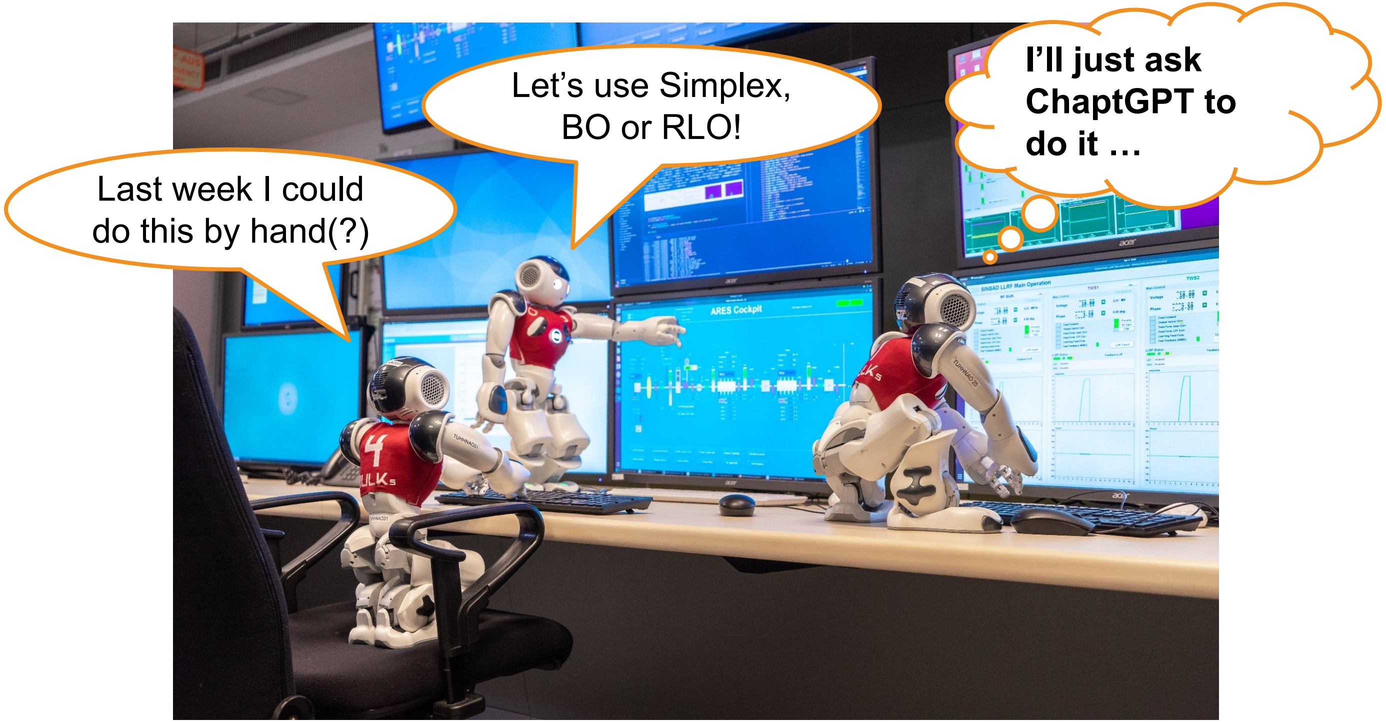
HELMHOLTZ RESEARCH FOR GRAND CHALLENGES

Universität Hamburg DER FORSCHUNG | DER LEHRE | DER BILDUNG

DESY.

An Oversimplified History of Autonomous Accelerator Tuning

From human intelligence over optimisation to artificial intelligence



Let's Ask ChatGPT to Do It …

Questions

large language models (LLMs)

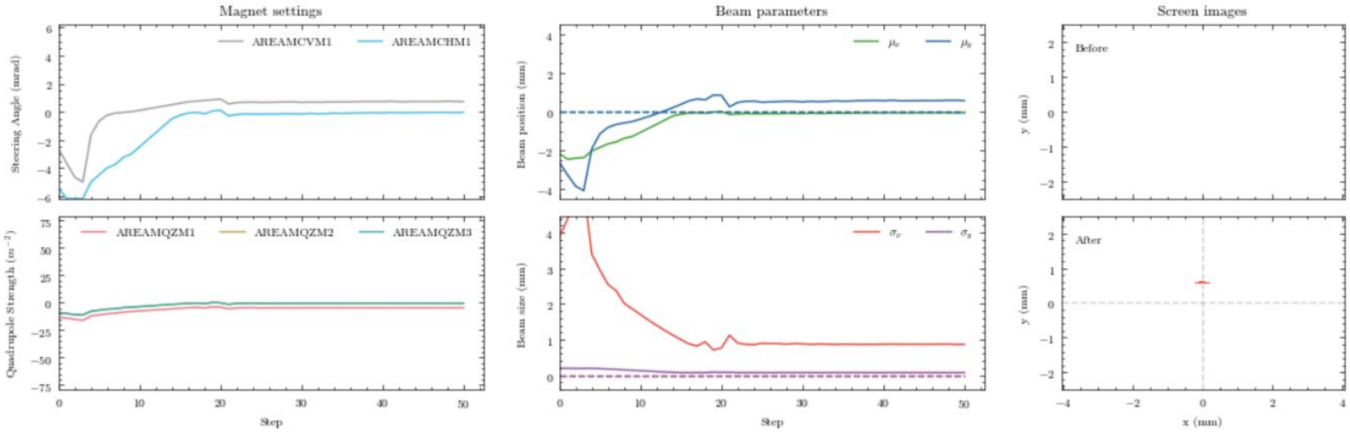
Can ~~ChatGPT~~ tune a particle accelerator?

How would that be implemented?

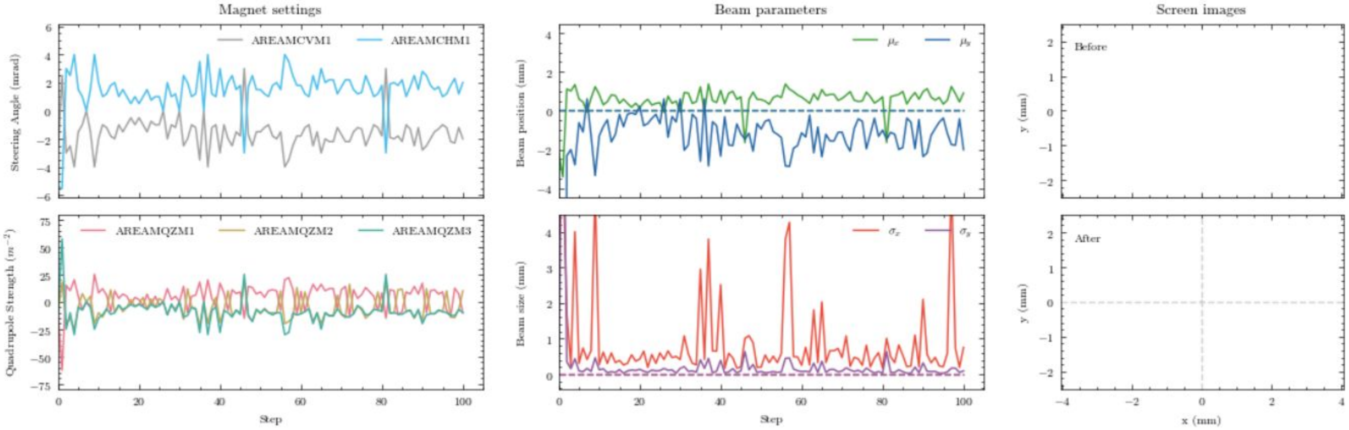
Teaser

Yes and no …

Yes … to some extent
GPT 4 (optimisation prompt)



But also no
GPT 3.5 Turbo (tuning prompt)



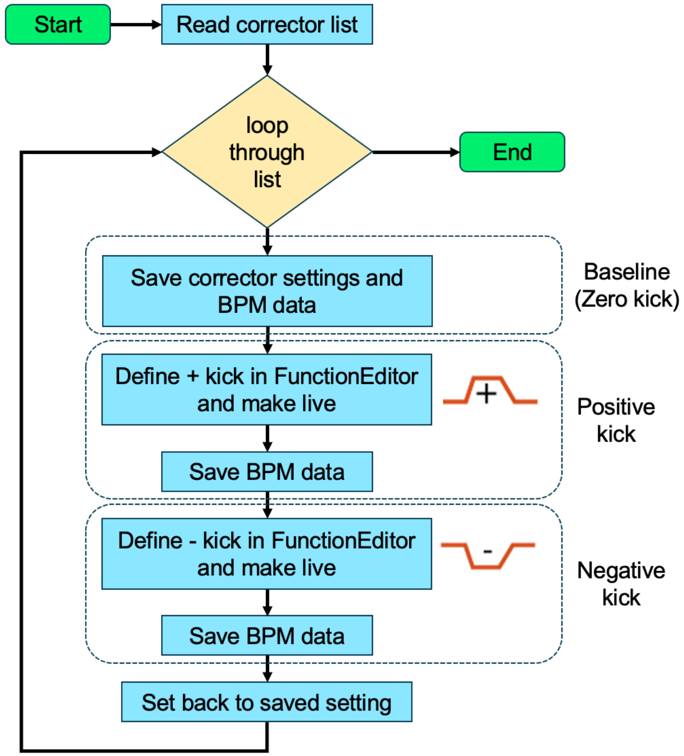
Contact

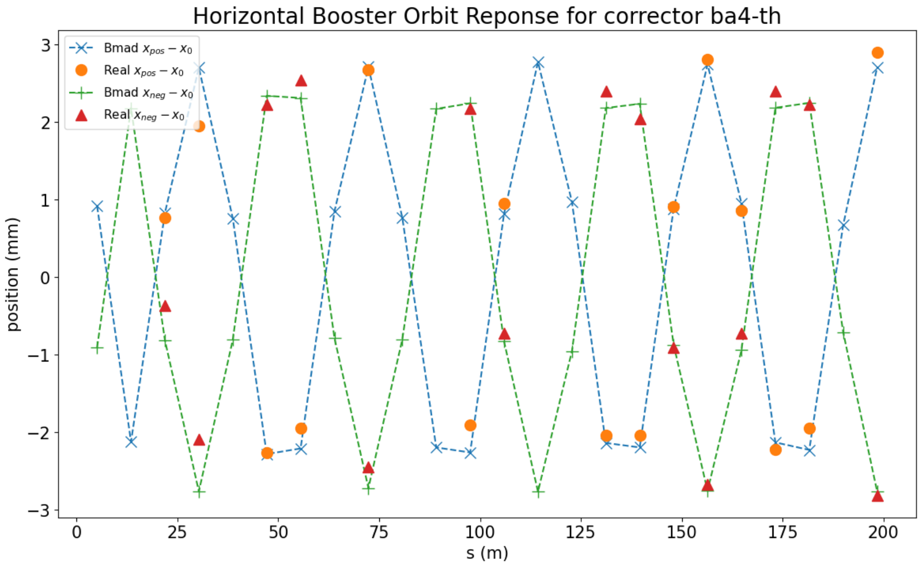
DESY. Deutsches Elektronen-Synchrotron

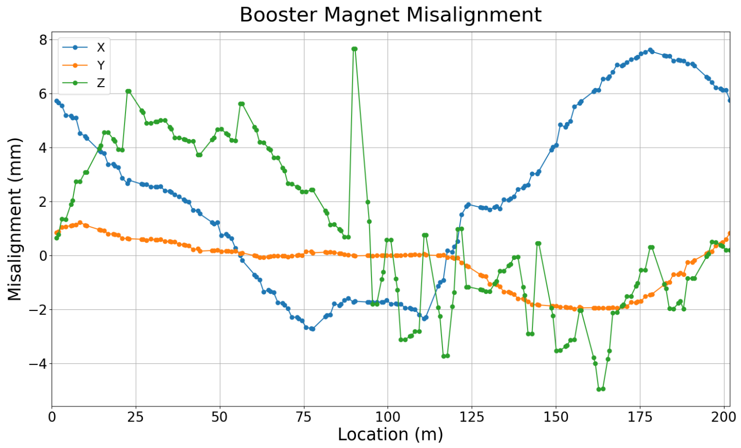
www.desy.de

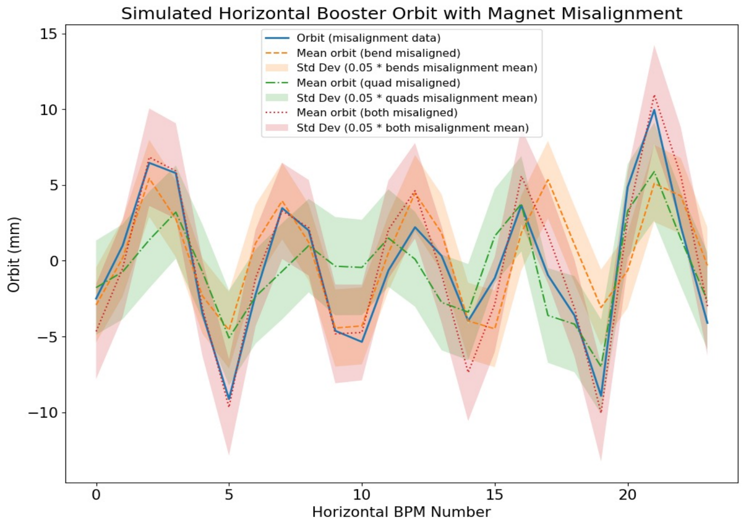
Jan Kaiser
Machine Beam Controls (MSK)
jan.kaiser@desy.de

Bayesian Optimal Experimental Design for AGS Booster Magnet Misalignment Estimation

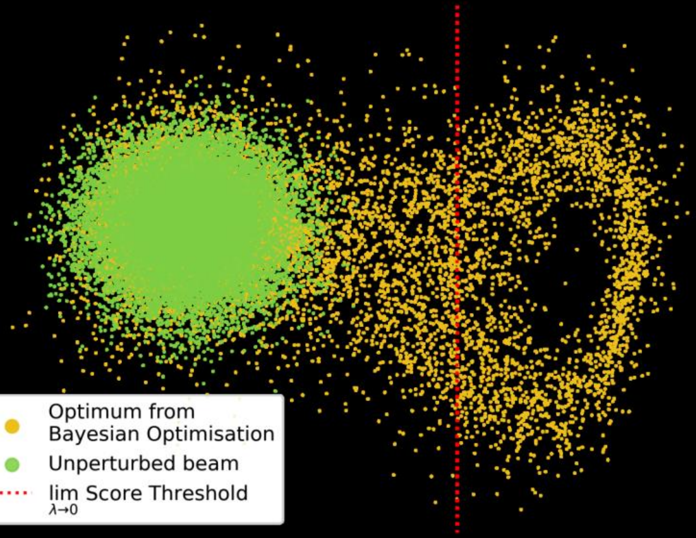




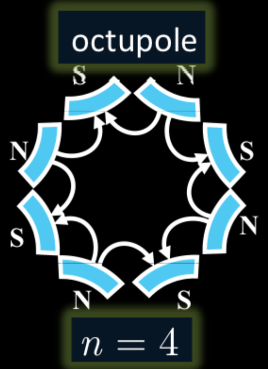


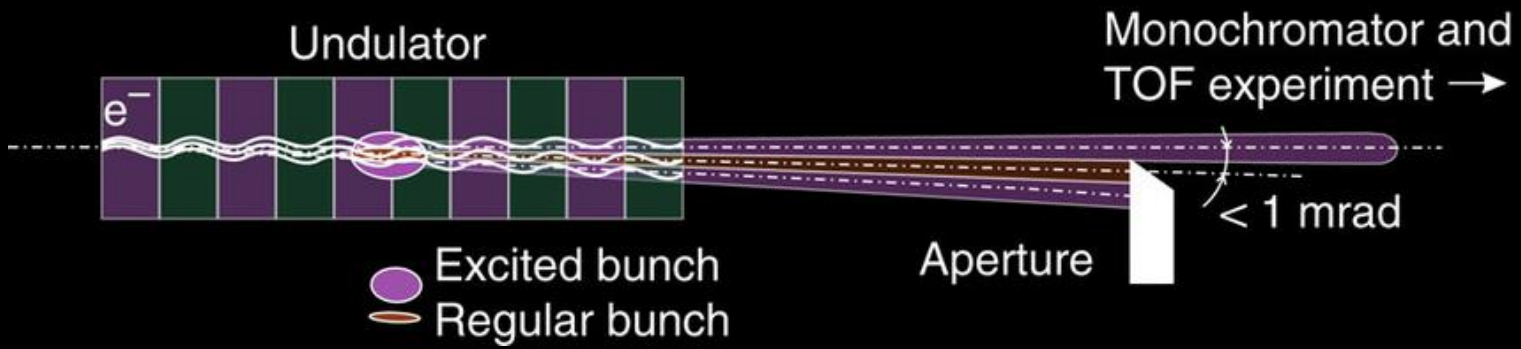


Brookhaven National Laboratory · CLASSE Cornell Laboratory for Accelerator-based ScienceS & Education · Stony Brook University · Rensselaer · SLAC National Accelerator Laboratory · Jefferson Lab Thomas Jefferson National Accelerator Facility



Comparing *Gradient Descent* vs Standard Methods (including *Bayesian Optimisation*) on a “real” Problem





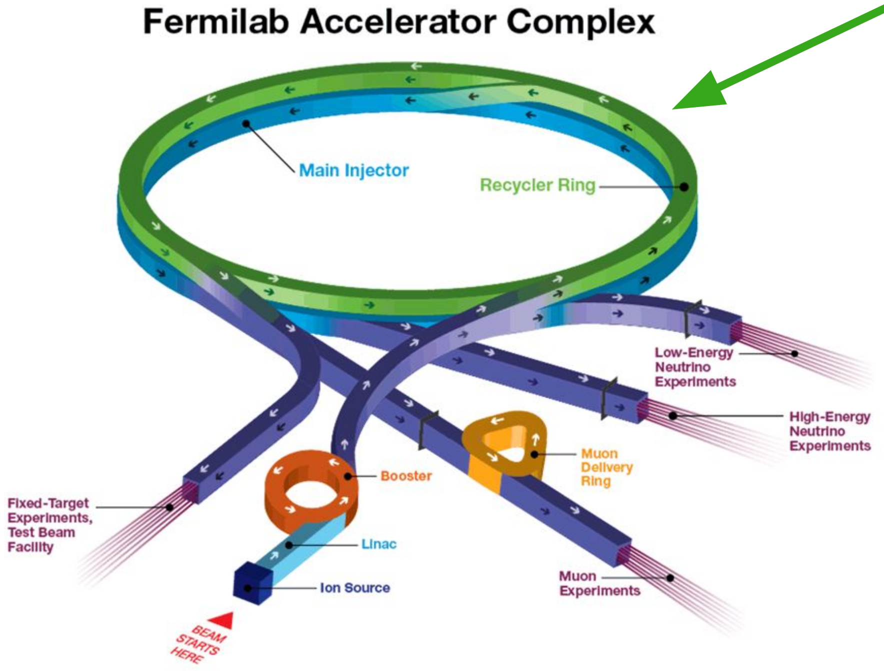
Rapid Tuning of Synchrotron Surrogate Models at the Recycler Ring

Jason St. John
Accelerator Division

4th ICFA Beam Dynamics Mini-Workshop Machine Learning Applications for Particle Accelerators
2024.03.05-08

FERMILAB-SLIDES-24-0026-AD

The Recycler Ring and High-Power Neutrino Beams



Recycler Ring is essential to Fermilab:

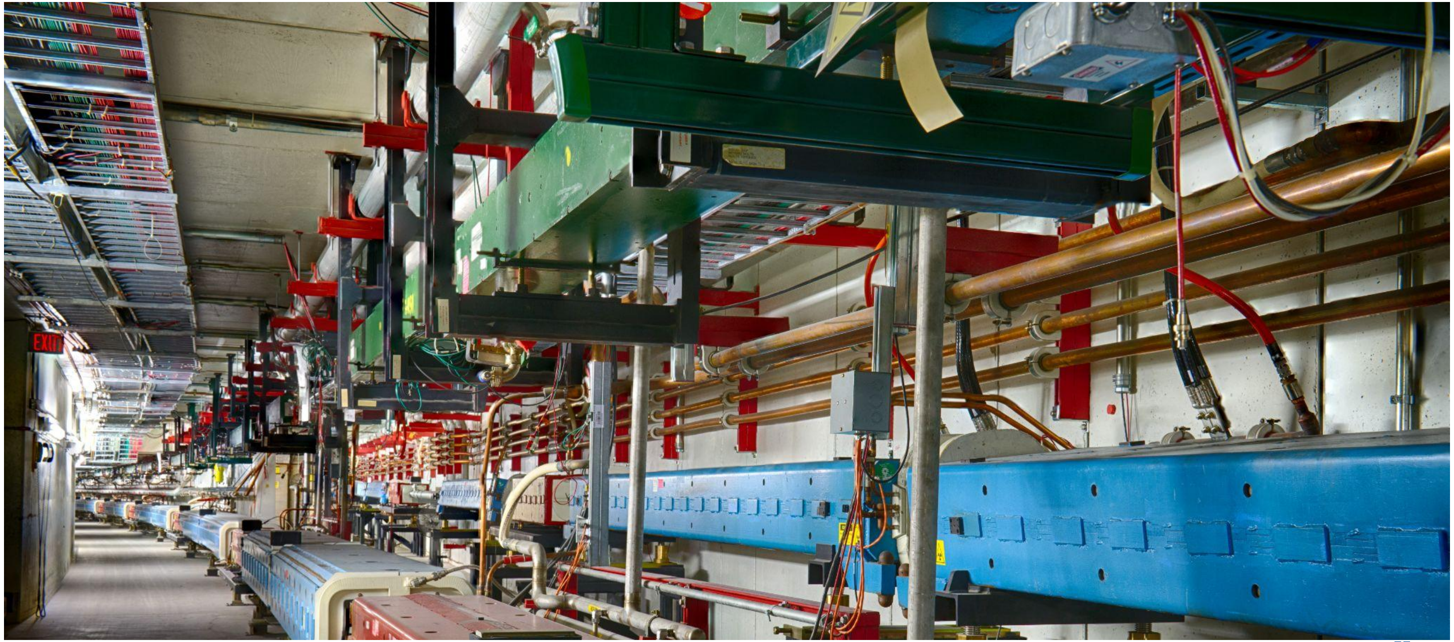
megawatt proton beams → high-intensity neutrino beams

This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics.

The Recycler Ring

Recycler Ring is a permanent-magnet storage ring

Matched to Main Injector 8 GeV proton KE



The Recycler Ring Multipole Shims

"Computer Generated End Shims for Recycler Ring Magnets" C.N. Brown, G.W. Foster, G. P. Jackson, J. T. Volk, Proceedings of the 1999 Particle Accelerator Conference, New York, 1999

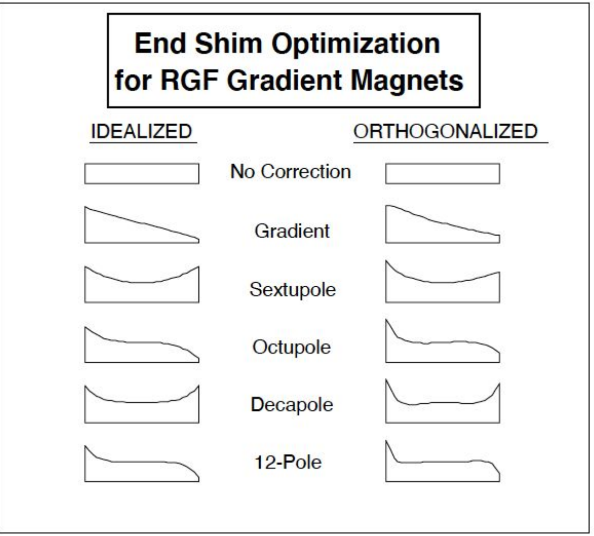
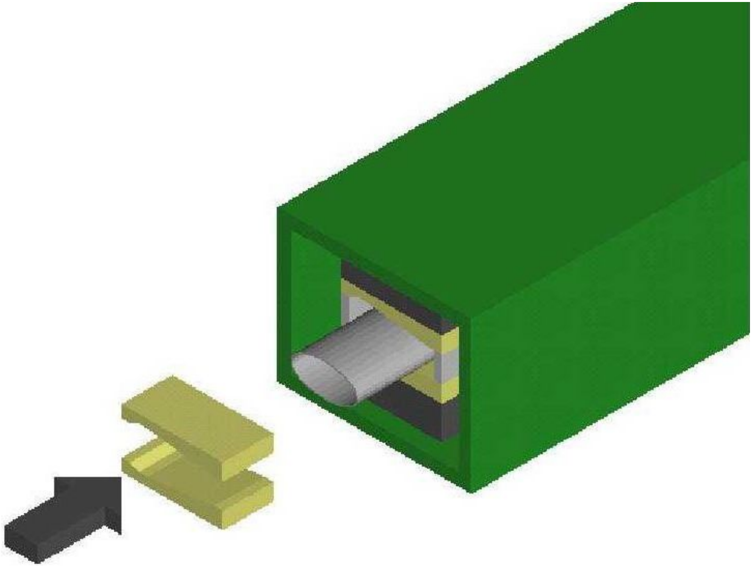
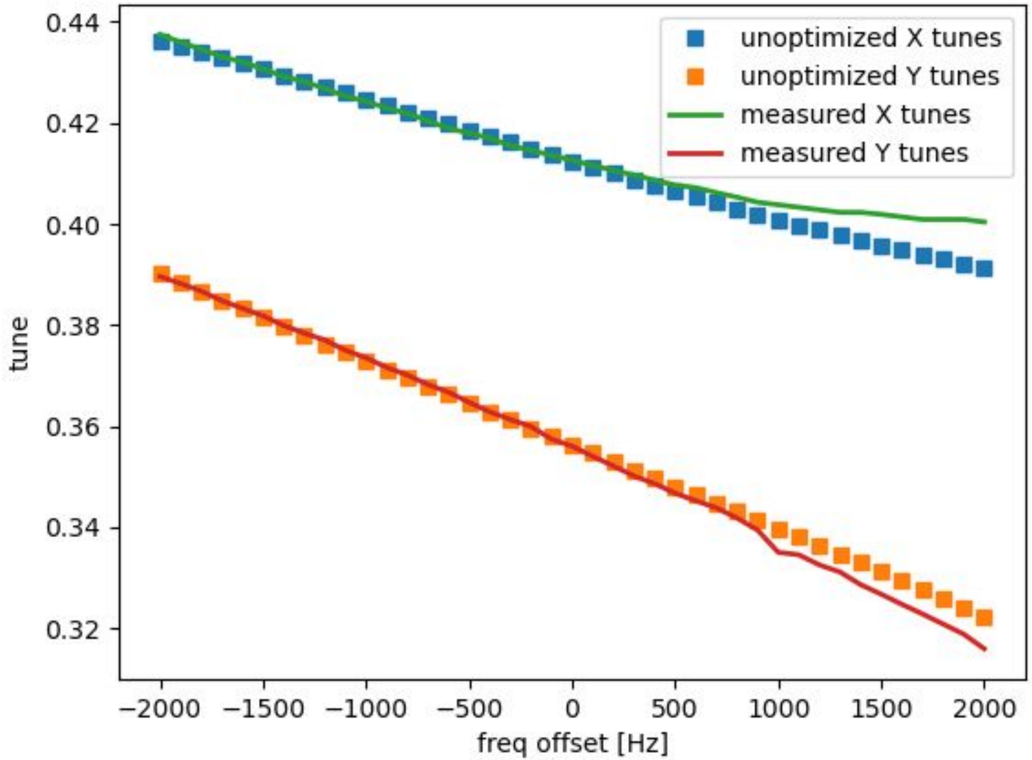


Fig. 2 – Elementary monomial Z-shim designs used as a starting point, and shim designs which were found to produce pure multipole shifts after re-orthogonalizing the multipole contributions from the elementary shims.



Magnetic shim plates were installed to correct for undesirable multiple moments observed in the Recycler.

The Recycler Ring Chromaticity

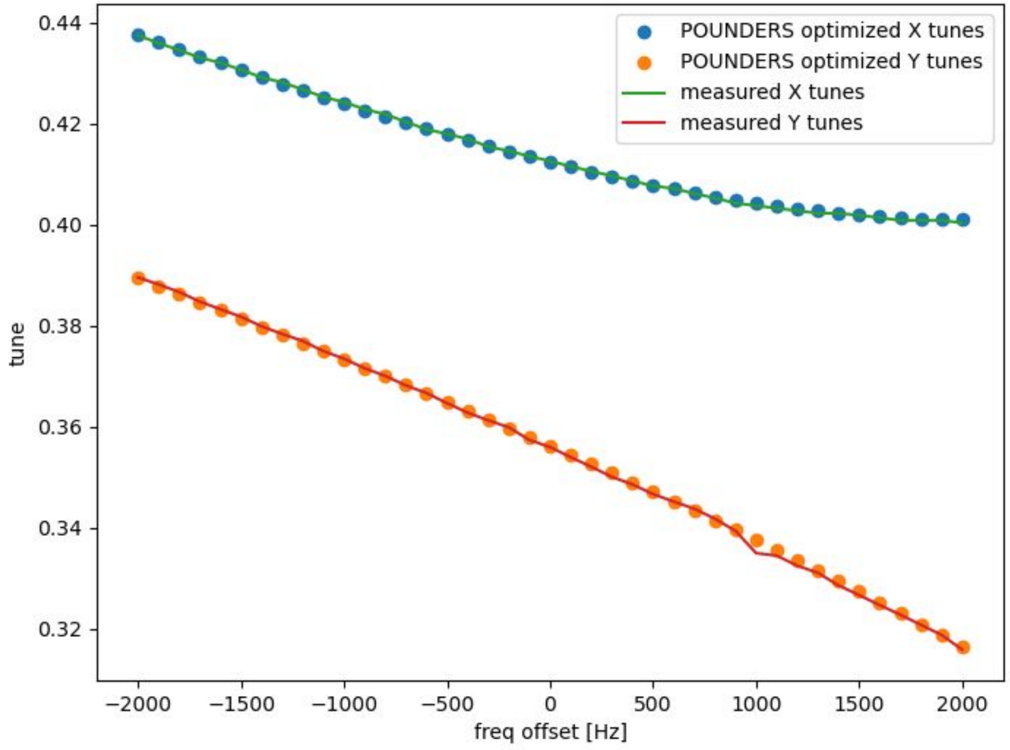


Challenge:

Tune simulated chromaticity to better match observed, using only additive corrections of multipole moments at shims.

Can small errors in shim plate shape account for the observed difference?

The Recycler Ring Chromaticity



Challenge:

Tune simulated chromaticity to better match observed, using only additive corrections of multipole moments at shims.

Can small errors in shim plate shape account for the observed difference?

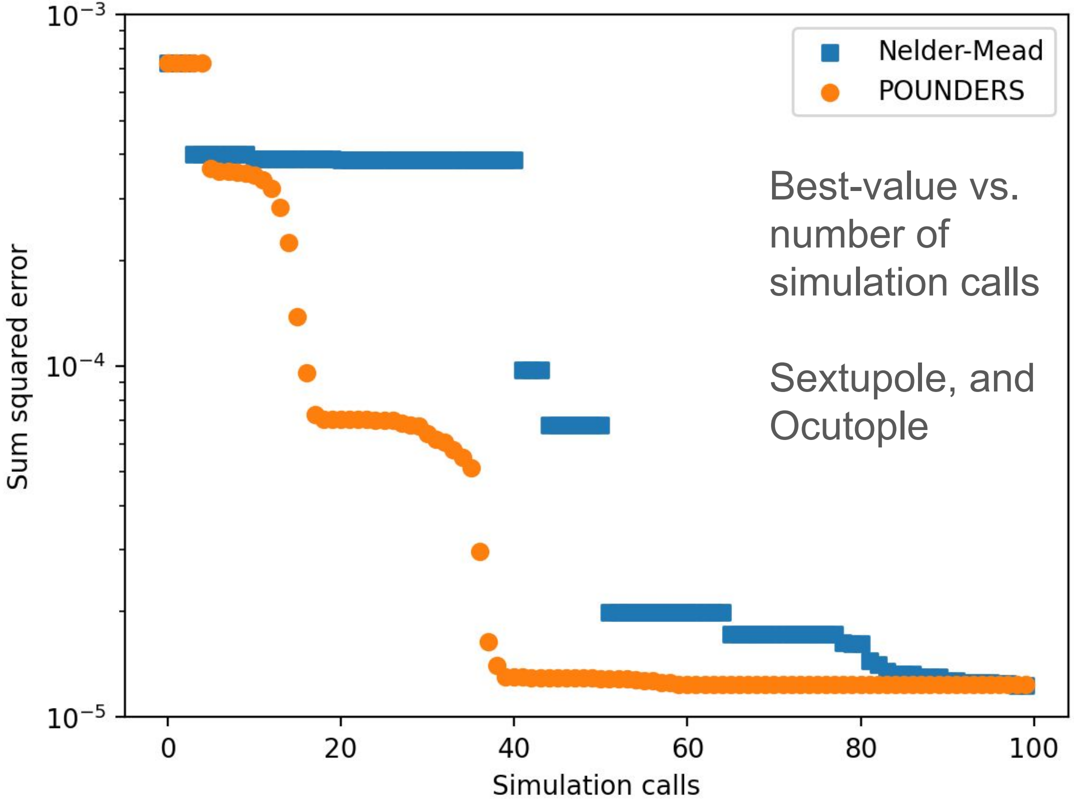
Yep! very fast with POUNDERS

Parameter Optimization with POUNDERS

Comparison to Nelder-Mead downhill simplex method (NM). Both are set to minimize the sum of squared errors.

Different inputs:
N-M: sum sq. err's
POUNDERS: vector of sq. err's

POUNDERS converges faster (~60 vs ~100 steps)



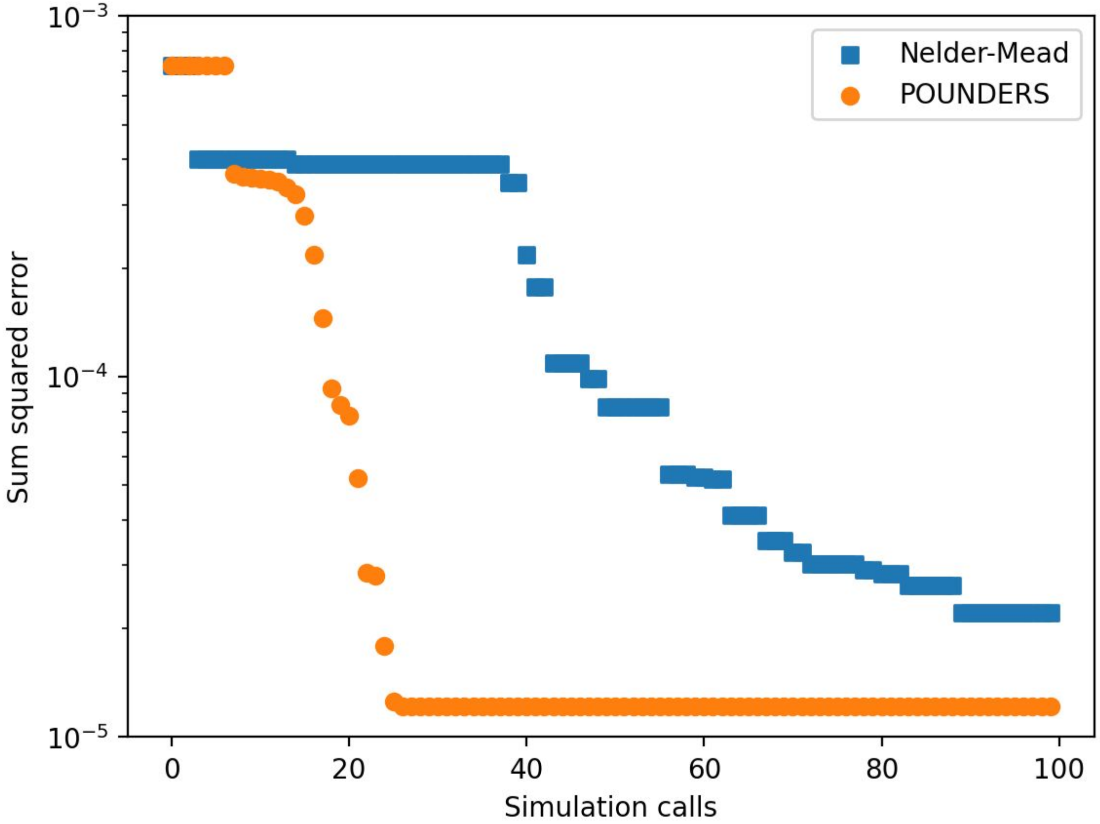
Best-value vs. number of simulation calls

Sextupole, and Ocutople

Final parameter values very close to NM result

	H sext.	V sext.	H octu.	V octu.
POUNDERS	-0.00107	-0.00099	0.37075	0.40921
NM	-0.00108	-0.00097	0.34959	0.45167

A tool for self-updating accelerator models?



Best-value vs. number of simulation calls

Sextupole, and Ocutople, and **Decapole**