Machine Learning Projects at C-AD

Kevin Brown
Collider-Accelerator Department, BNL
brownk@bnl.gov

MAC 2022 December 14, 2022

Quick Listing of efforts

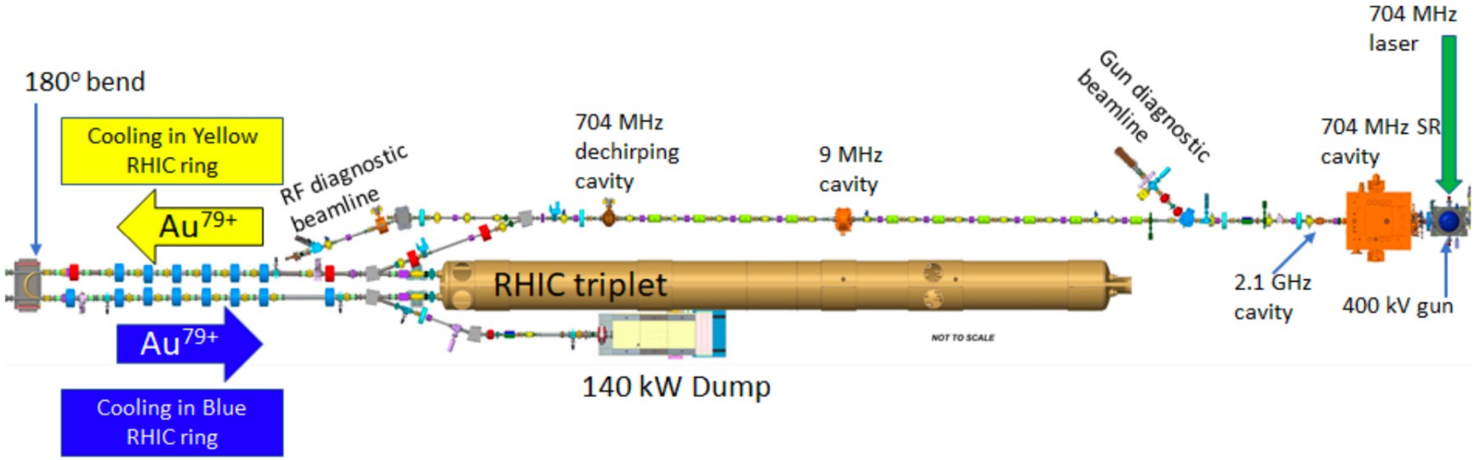
- Gaussian Process (GP) Bayesian Optimization (BO) for LEReC
- Emittance Measurement Speedup with Machine Learning at CeC
- G-gamma meter in AGS - how to learn the correct jump quad timing
- Accelerator self-diagnosis and automating ORMs
- Reconstructing transfer functions using beam-based analysis (Booster)
- Natural Language Processing (NLP) for elogs and other apps
- Optimization of a Longitudinal Bunch Merge Gymnastic with ML
- Ionization Profile Monitor Channel Gain Calibration with ML
- Adopting and using XOpt and Badger (and eval of Geoff and COI)
- Collaboration with FNAL/JPARC for slow spill control
- RadiaSoft - A Browser Based Toolkit for Improved Accelerator Controls

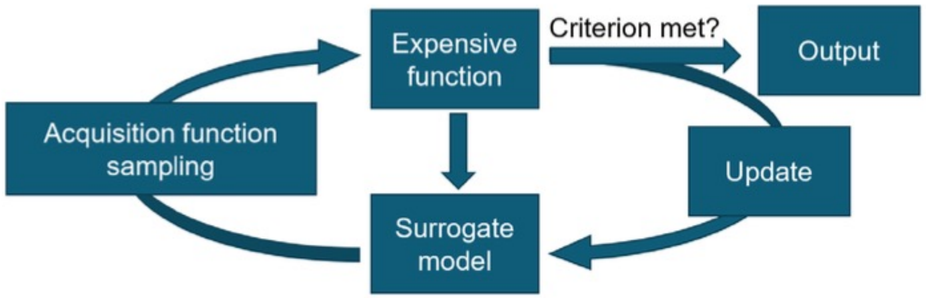
Collaborations & Sharing

- FNAL & KEK/J-PARC for Slow Extraction (US/Japan funded)
- RadiaSoft SBIR (now applying for Phase IIb)
- Collaborating with Cornell, supporting PhD student
- Collaborating with Univ. of New Mexico, supported Post-doc and PhD students
- Supporting PhD students at Stony Brook Univ.
- Strongly supportive community with active communication with SLAC, FNAL, J-Lab, CERN, GSI, and DESY
- Strongly supportive BNL community with support from Computational Science Initiative group, NSLS II, Physics, ATRO/ATF

Highlights

Bayesian optimization experiment for trajectory alignment: LEReC





Bayesian optimization process

- LEReC is used to increase the luminosity, it was successfully improved the luminosity multifold in 2020 and 2021 runs;
- 704 MHz e-bunches (grouped into 9 MHz macro-bunches) are produced from the photocathode and accelerated in the SRF cavity to the designed energy (1.6 MeV, 2 MeV);
- Those e-bunches are delivered to the cooling sections (20 meter), where they co-travel with ion bunches.

Bayesian Optimization (BO): a powerful tool for finding the extrema of objective functions that are expensive to evaluate

$$P(f|\mathcal{D}_{1:t}) \propto P(\mathcal{D}_{1:t}|f)P(f)$$

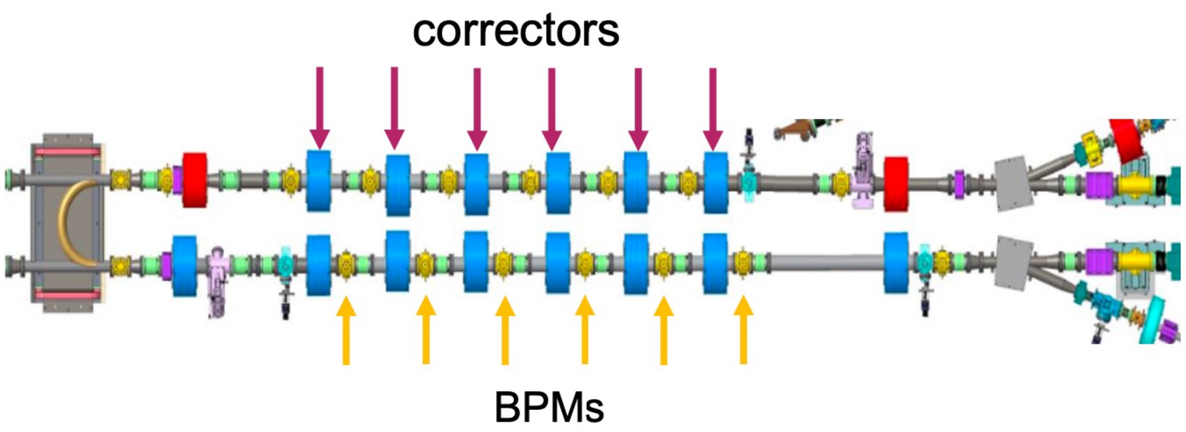
Upper Confidence Bound

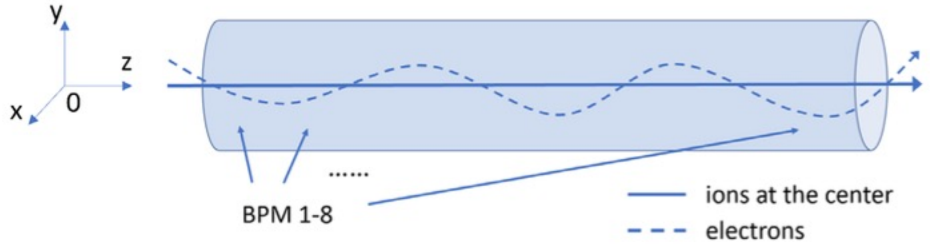
$$\mathbf{UCB}(\boldsymbol{x}) = \boldsymbol{\mu}(\boldsymbol{x}) + \boldsymbol{\kappa\sigma}(\boldsymbol{x})$$

Brookhaven National Laboratory CLASSE

Y. Gao, W. Lin, K. A. Brown, X. Gu, G. H. Hoffstaetter, J. Morris, and S. Seletskiy, *Bayesian optimization experiment for trajectory alignment at the low energy RHIC electron cooling system*, Phys. Rev. Accel. Beams 25, 014601 – Published 7 January 2022

Experiment Settings





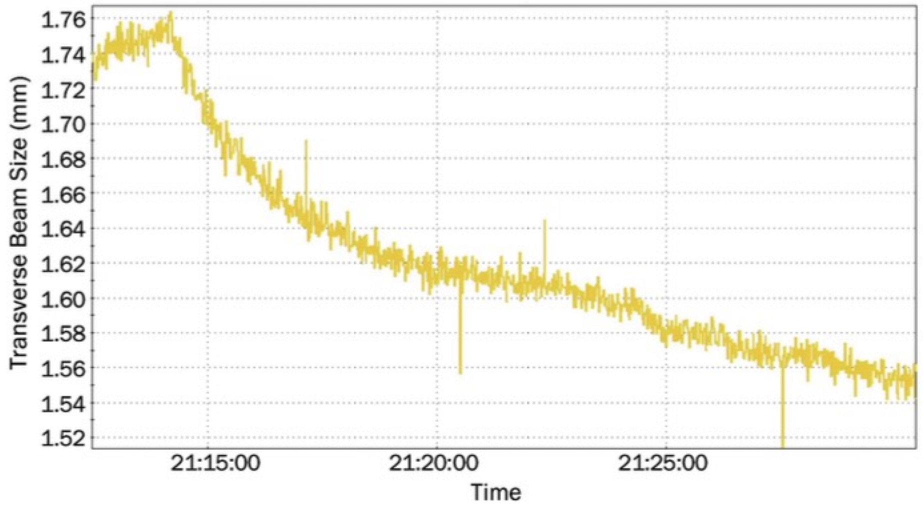
electrons travel through the cooling sections and are monitored by 8 BPMs

The <u>Goal</u> is to use BO to tune electron trajectories to maximize the ion cooling rate.

- Ions are assumed in the center position, only the first 4 BPMs are considered;
- Decreasing speed of transverse ion beam size:

$$\lambda = (1/\delta)(d\delta/dt)$$

- Cooling performance is measured by (- λ), a more negative λ means a faster cooling rate;

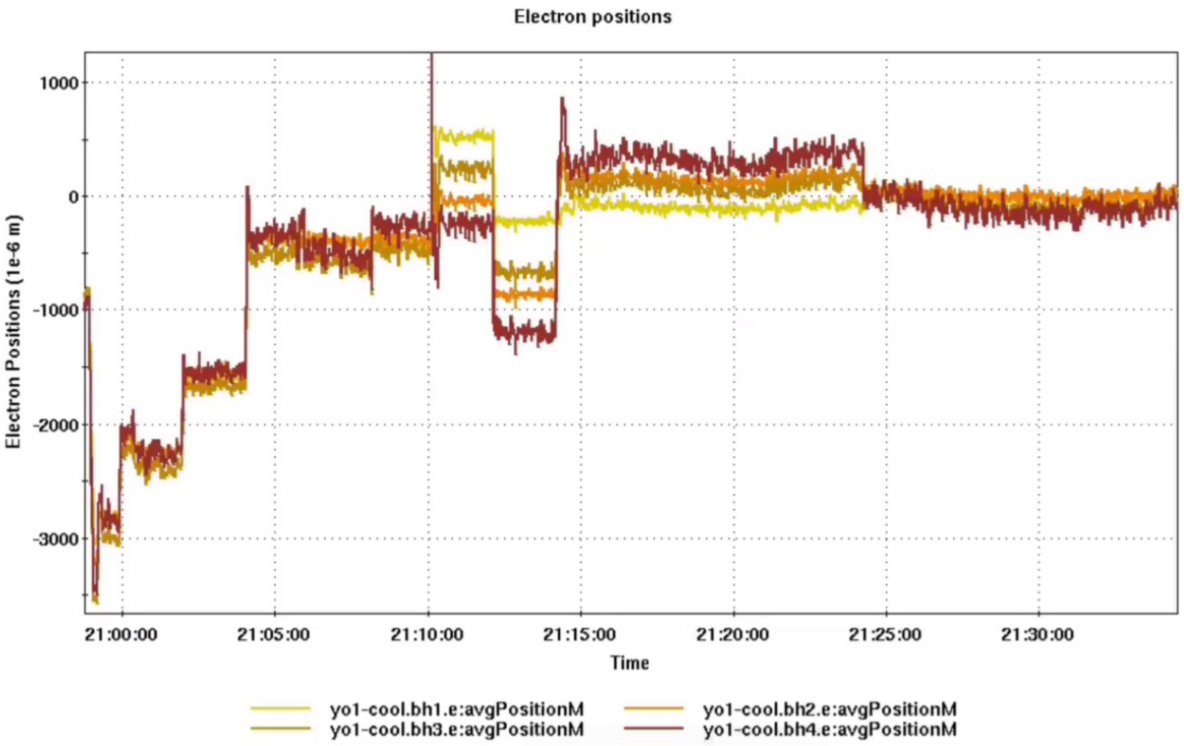


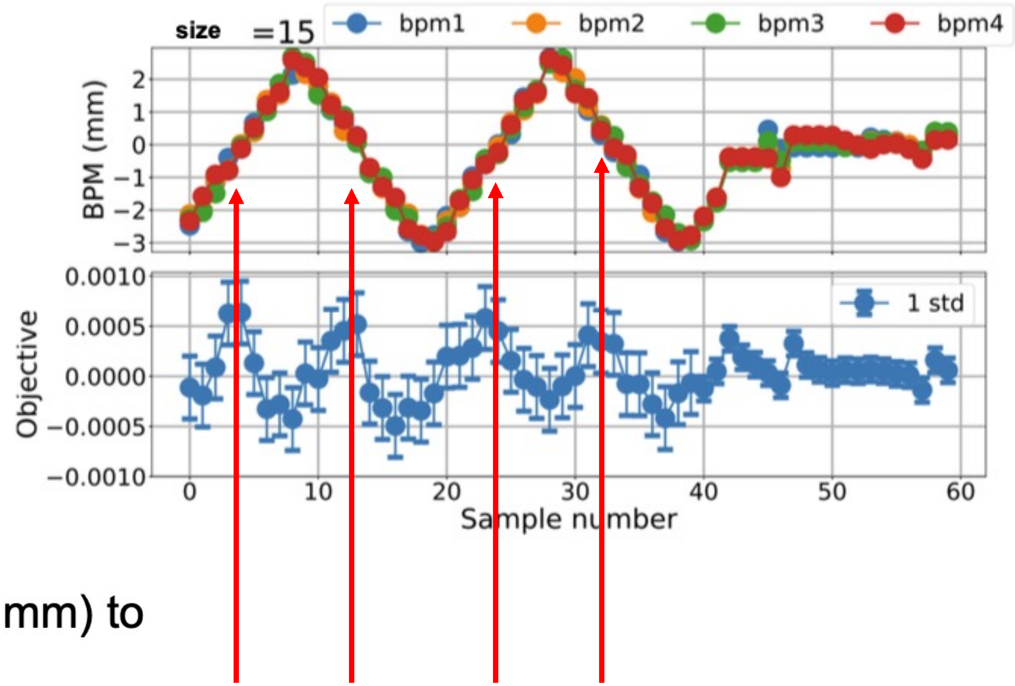
An example period of transverse ion beam size data are fetched from the real system during the experiment. It shows the data are noisy.

Brookhaven National Laboratory

Electron Positions Controlled by the BO

Transverse cooling rate λ is defined as the decreasing speed of the transverse ion beam size δ, which is calculated as $\lambda = (\frac{1}{\delta})(\frac{d\delta}{dt})$. The objective function is basically λ, using average beam sizes in an interval to reduce noise.





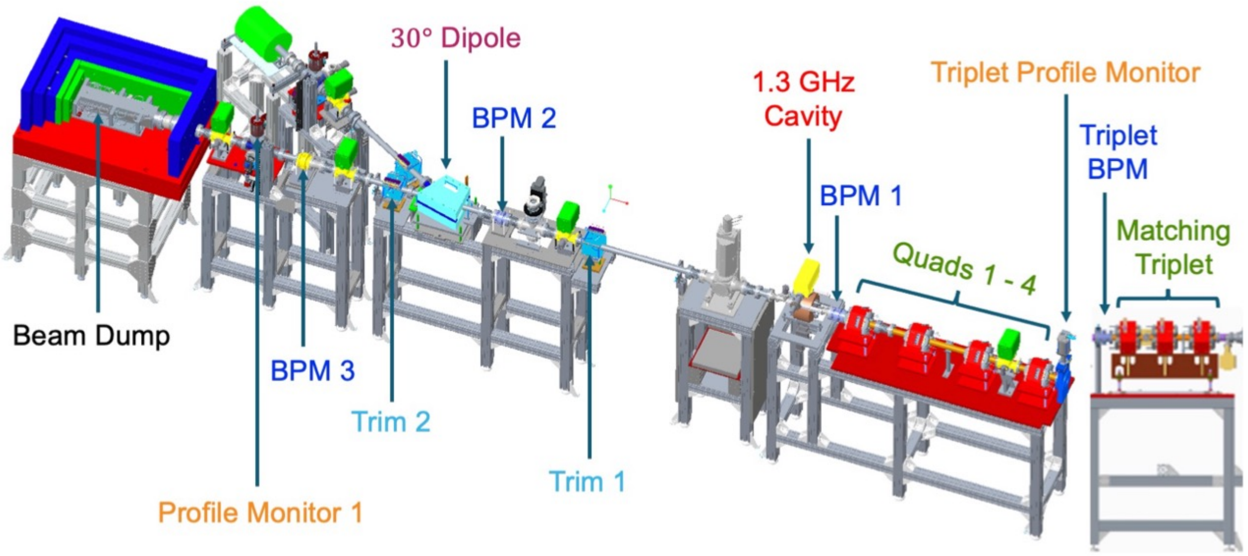
Maximum cooling rate when <x> near 0.

- Electron trajectories reported by 4 BPMs;
- The algorithm can tune the trajectories from the farthest points (-3 mm) to the center position and maintain them.

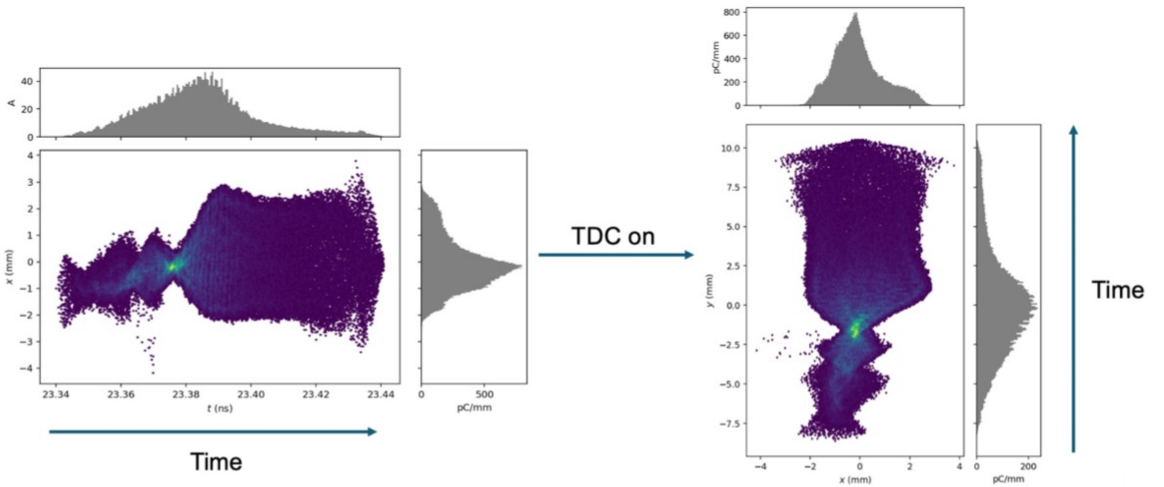
Emittance Measurement Speedup with Machine Learning at CeC

Time-resolved Diagnostic Beamline (TRDBL)

- Capable of evaluate electron beam quality with time resolution of 1 ps
- Fully characterize transverse and longitudinal beam profiles



- A **transverse deflecting cavity** (TDC) provides a time dependent transverse kick to the beam
- After TDC, the beam's longitudinal profile converts to Y direction, which is measurable on YAG screen



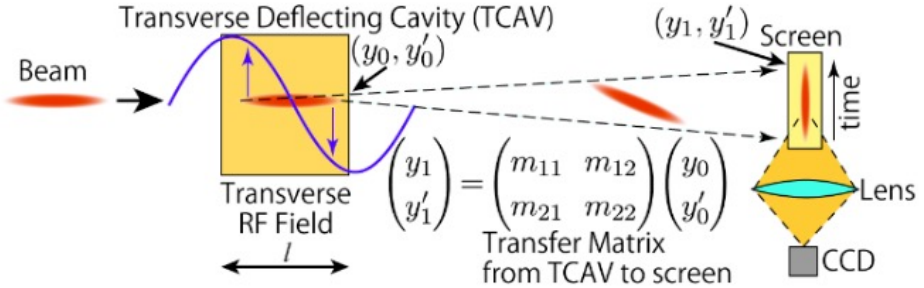
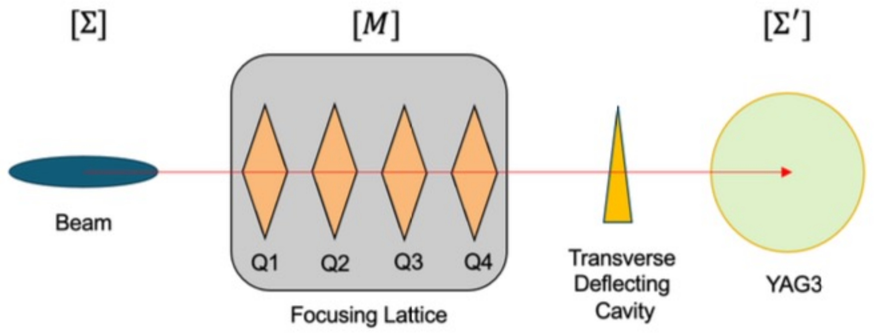
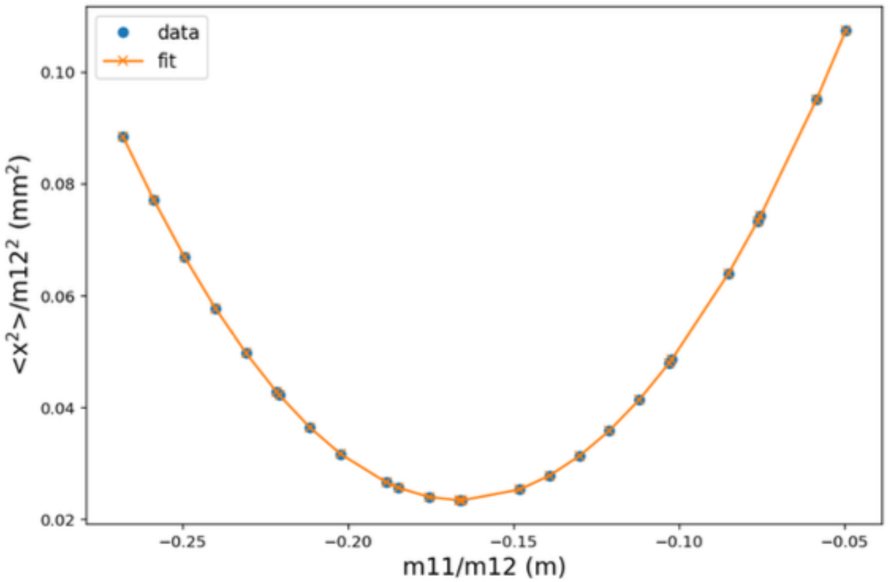


Diagram from:
H. Maesaka, T. Asaka, T. Ohshima, H. Tanaka, Y. Otake, S. Matsubara
MOCLA02, Proceedings of IBIC2015, Melbourne, Australia

W. Lin, A. Sampson, Y. Jing, K. Shih, G. H. Hoffstaetter
3rd ICFA Beam Dynamics Mini-Workshop on Machine Learning Applications for Particle Accelerators
Chicago, IL, November 1st – 4th, 2022

Emittance Measurement in Diagnostic Line





Use **quadrupole scan** to measure emittance:

$$\because\ [\Sigma'] = [M][\Sigma][M]^T$$

$$\therefore \sigma'_{11} = m_{11}^2\sigma_{11} + m_{11}m_{12}2\sigma_{12} + m_{12}^2\sigma_{22}$$

$$\text{Define } \nu = \frac{m_{11}}{m_{12}}\ ,\ \sigma'_\nu = \frac{\sigma'_{11}}{m_{12}^2}$$

$$\sigma'_\nu(\nu) = \sigma_{11}\nu^2 + 2\sigma_{12}\nu + \sigma_{22}$$

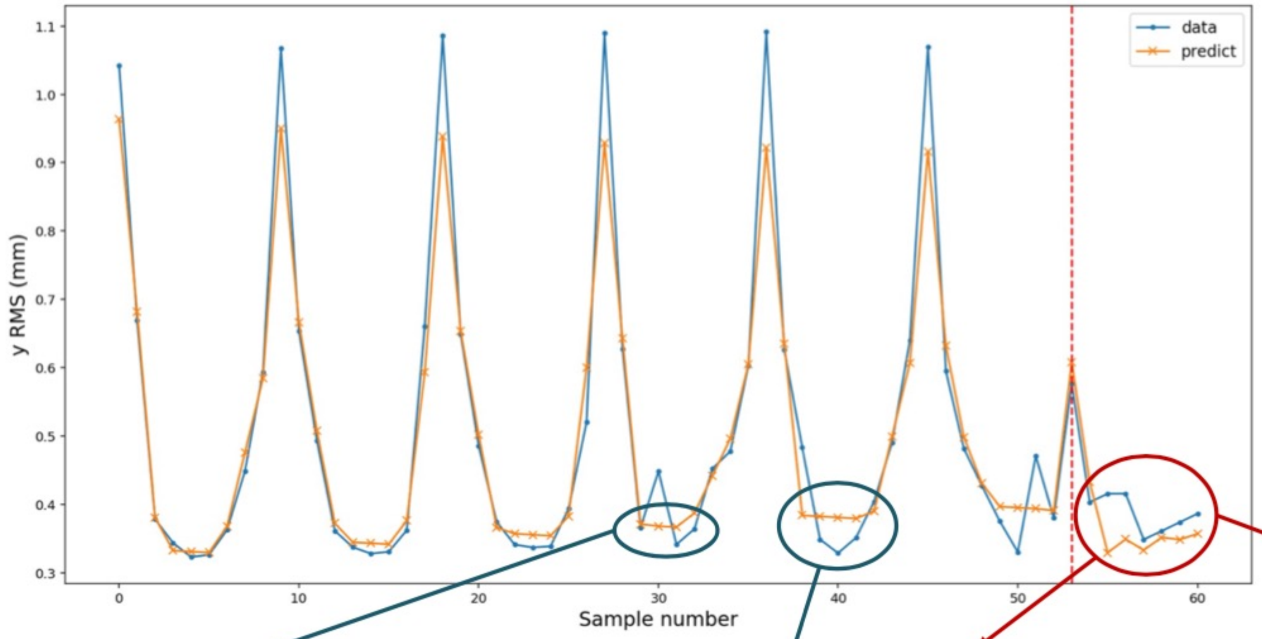
Use parabola fit parameters to get emittance:

$$\varepsilon = \sqrt{\sigma_{11}\sigma_{22} - \sigma_{12}^2}$$

- Use **Q3** and **Q4** for quadrupole scan
- For each Q3 value, find Q4 value for best vertical focusing at **YAG3**
- Turn on **deflecting cavity**, convert longitudinal beam info to Y direction
- Slice beam vertically for slice emittance
- Old routine:
  - Scan 13 Q3 settings
  - For each Q3, scan 9 Q4 settings
  - **> 1 hour** for entire scan

Brookhaven National Laboratory

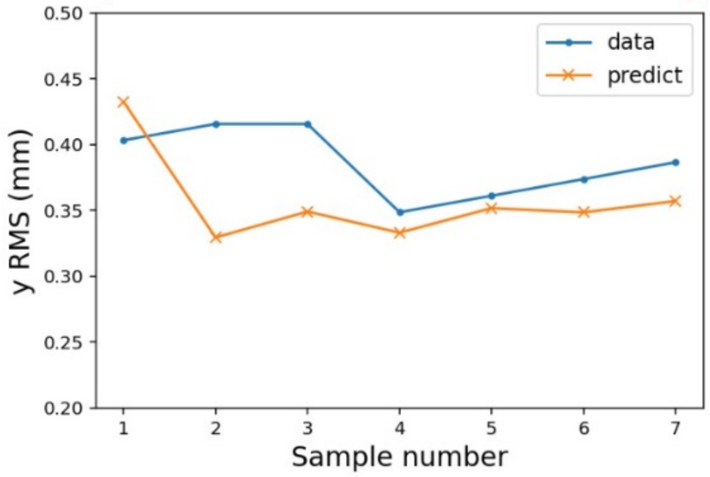
First 6 rounds: 54 saved data points with old script



New Routine with **Neural Network (NN)**:
- Scan 6 Q3 settings, save data
- Train NN model with saved data
- NN predicts best Q4 settings for remaining 7 Q3 settings
- Load predicted settings to beamline

Remaining 7 rounds: 7 data points using Q3-Q4 settings predicted by NN model

Need to fix: trouble getting the small Y RMS region features



Cut Scan Time by **50%**

- **Satisfactory preliminary results:**
  - Obtained 7 Y RMS values all around 0.3 – 0.4 mm range
- Future work: incorporate new routine into control system

Natural Language Processing for ELogs

Motivation

- The elog search feature only provides exactly what a user enters, what if there are other entries that do not include those exact words/characters BUT also are related to these things
- Can we determine what the user is interested in viewing?
- Eventually provide custom sets of entries based on users' interactions with the system

Brookhaven National Laboratory

ML techniques help customize logbook settings and views, including specification of favorite logbooks

æ Feb 25 11:26 cp ph

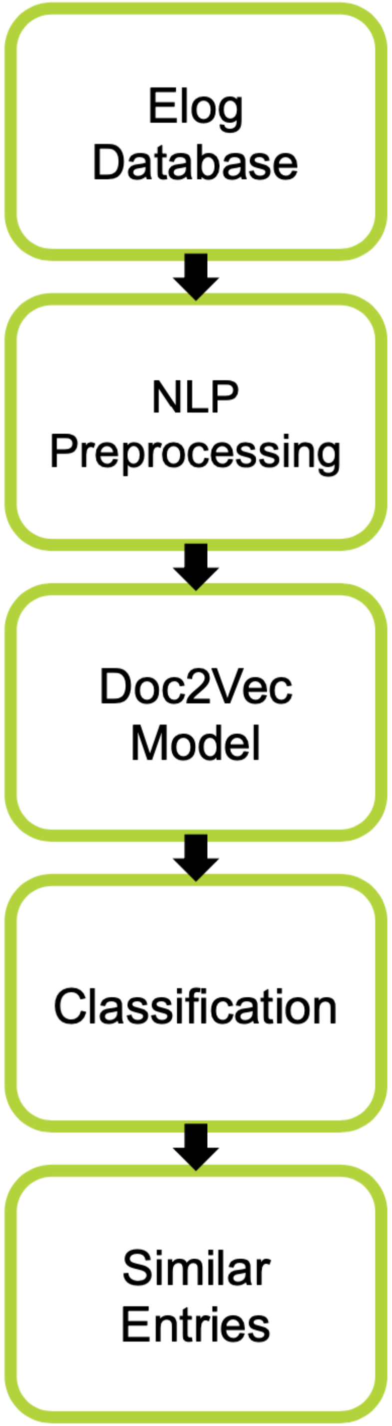
I'm not convinced that this is the wrong behavior but I dumping some pictures just for documentation purposes.

Feb 25 11:26 ph

Steve called worried that some of the polarimeter planes rdbks are railed at 100000. In the these pictures horiz is updating as expected, vertical is not. This is true for all 4 polarimeter targets (if horiz is updating, vert is not and vice versa).
The HW limits the operations to only a single plane, ie. horiz would need to be at home before vert coud be moved. I'm going back through archives and the rdbks are consistent.
Steve is going to do a little more investigating on the HW side but I think things look right on the SW side.

Feb 25 16:17 ph

After making a tunnel access to verify that things looked OK Steve enabled the motion and that triggered the updates.



Status

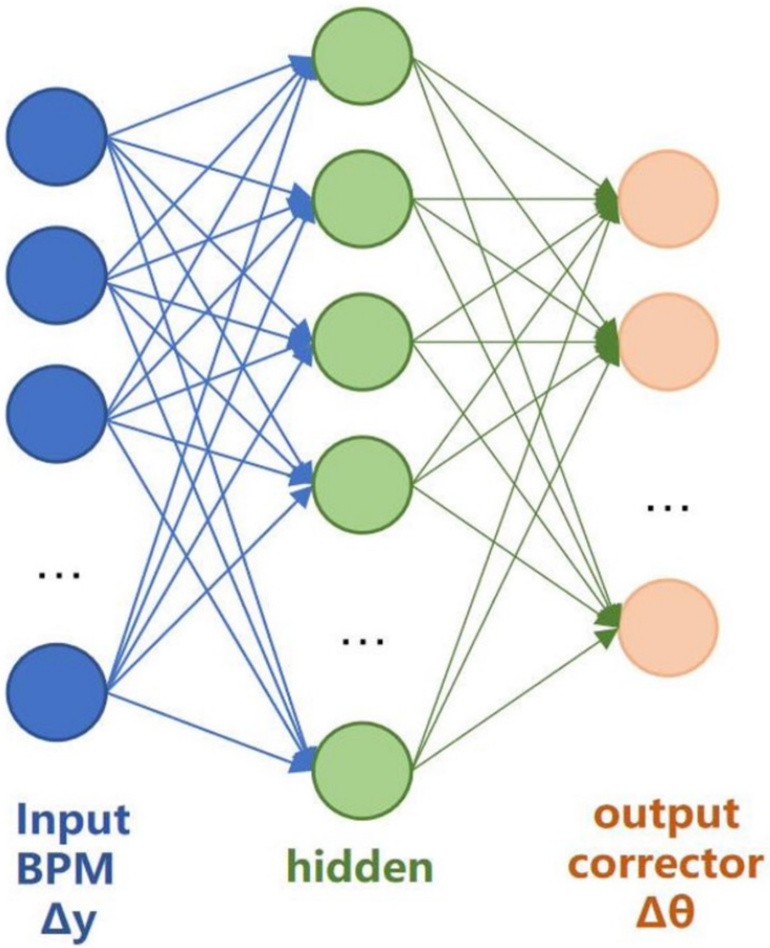
- System in place to determine what entries each user interacts with
- During run time, elog entries are often automated. studying the impact on the suggested entries
- Investigating optical character recognition to help include images in entries for similarity classification
- Evaluating how typos are dealt with
- Web interface implemented into the elog system

- Note: this work was reported at the last ICFA BD Mini-workshop on ML in Accelerators and generated a lot of excitement = new collaborations are forming

Accelerator self-diagnosis and automating ORMs

(an ongoing project in collaboration with Cornell Univ.)

Neural Network for real-time ORM

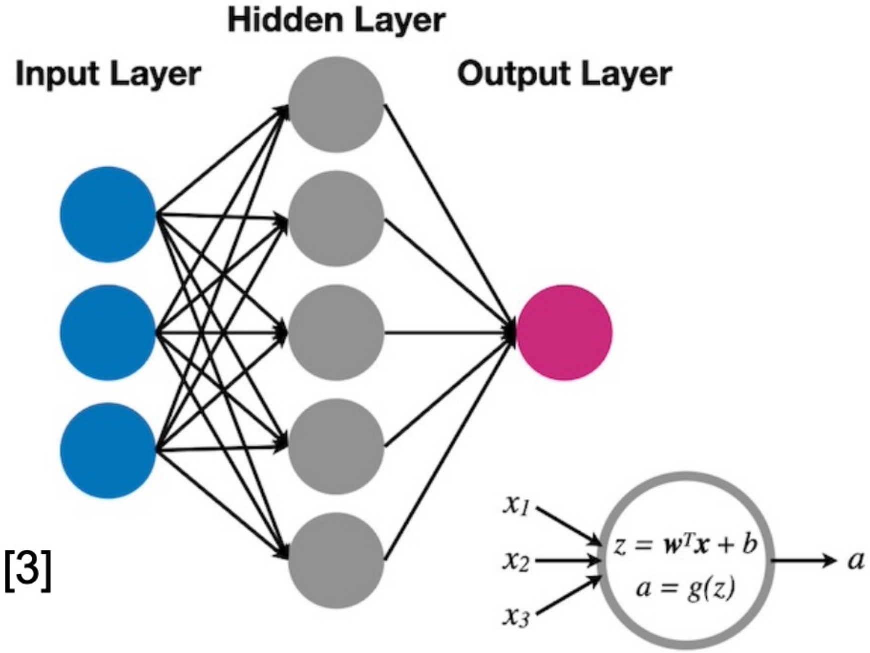


[2]

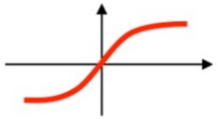
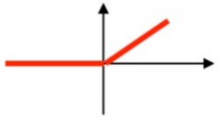
- Need dedicated machine time to measure ORM $\underline{R}_{measured}$: at least 30 min
- Pre-measured $\underline{R}_{measured}$ gets less accurate with time → orbit drift / brightness drop
- Update ORM with real-time data: build neural network model for $\underline{R}_{measured}$ or $\underline{R}^{-1}_{measured}$
- Can be used to calculate $\Delta\vec{R}$ for machine error reconstruction

Method: Feed Forward Neural Network

- Neural Network (NN) built with PyTorch library
- Fully connected layers: output = activation(dot(input, weight) + bias)
- Activation function: Hyperbolic Tangent (Tanh) and Rectified Linear Unit (ReLU)
- Feed forward neural network (FFNN): most common, no feedback route



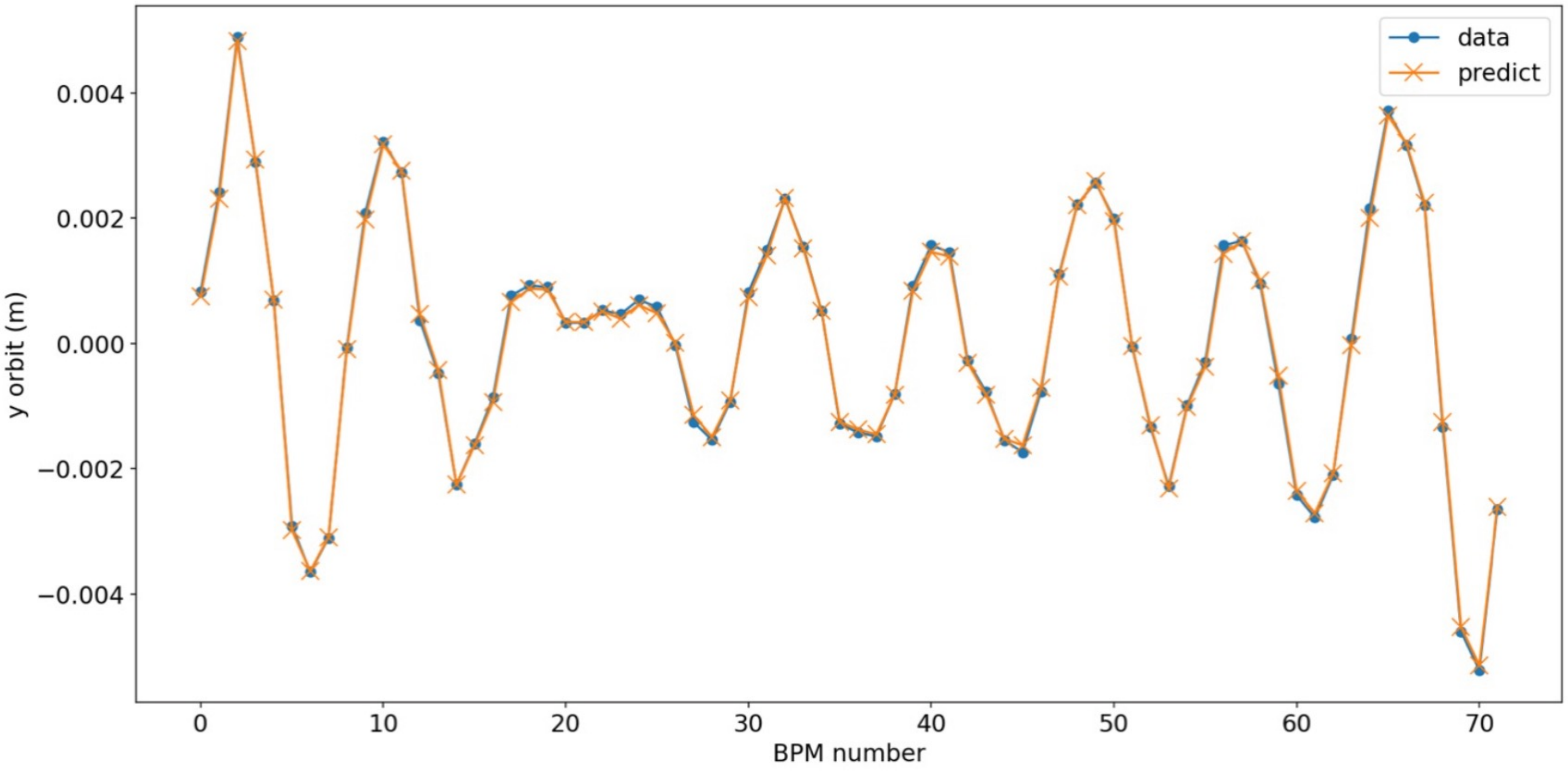
[3]

Hyperbolic tangent	$\phi(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$	Multi-layer Neural Networks	
Rectifier, ReLU (Rectified Linear Unit)	$\phi(z) = max(0, z)$	Multi-layer Neural Networks	

ORM NN model: training results

- Input 48 vertical corrector kick → Output 72 y orbit measured at BPM
- FFNN with one hidden layer and Tanh activation
- Trained on 800 data pairs, tested on 200 data pairs: R^2 score = 0.998

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \overline{y}_i)^2}$$

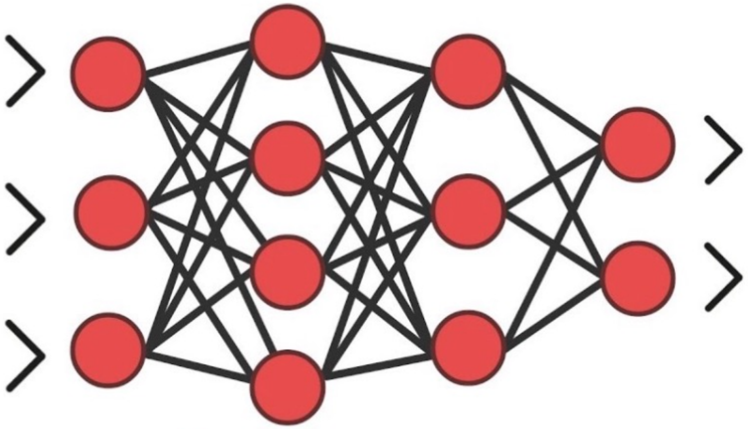


Brookhaven National Laboratory

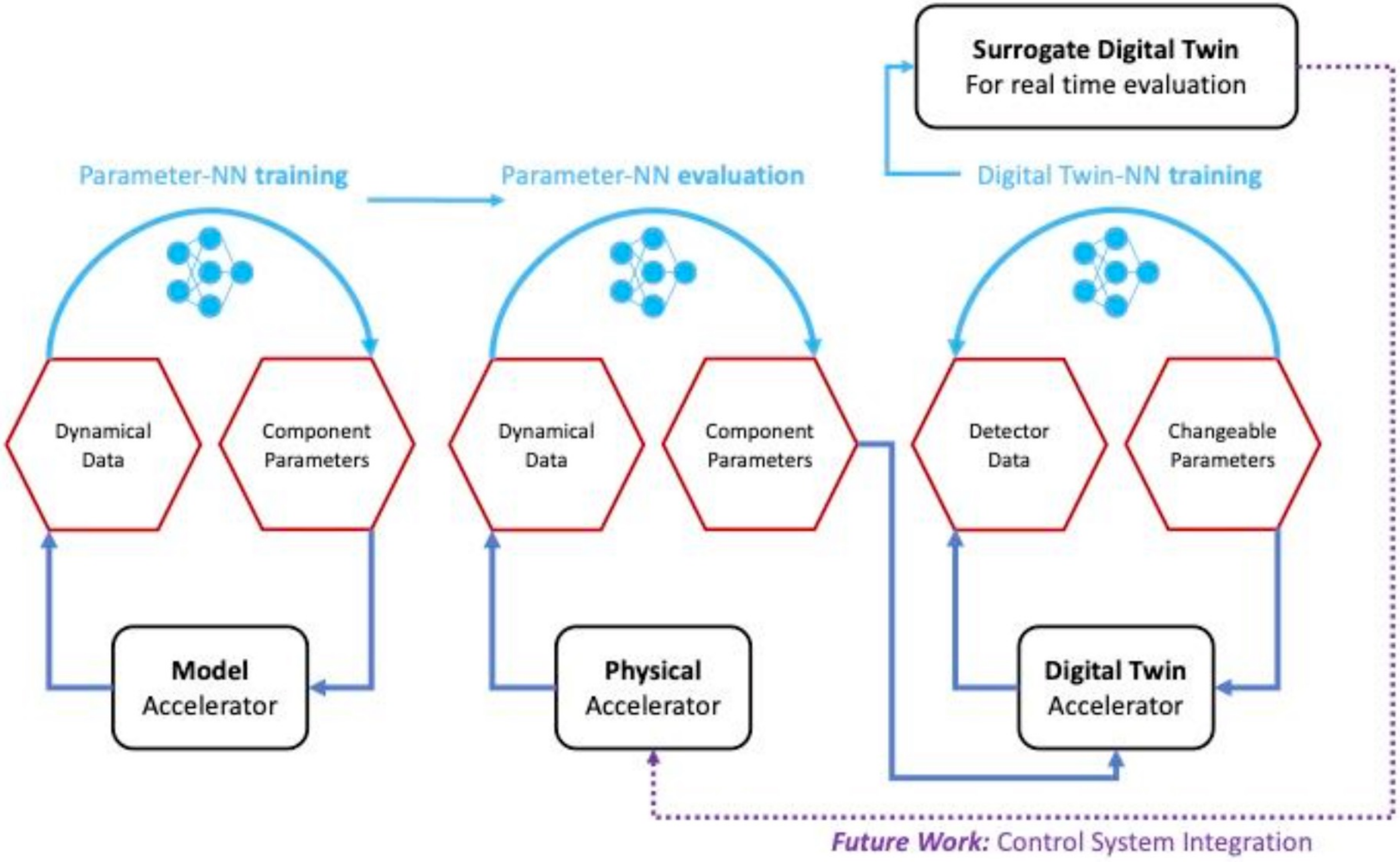
Sensitivity studies for ORM

- Scan through some common sources of error to see how much ORM changes
- Find relevant parameters to include for building error-detecting model
- **<u>Goal</u>**: establish a neural network that identify error source given a measured ORM

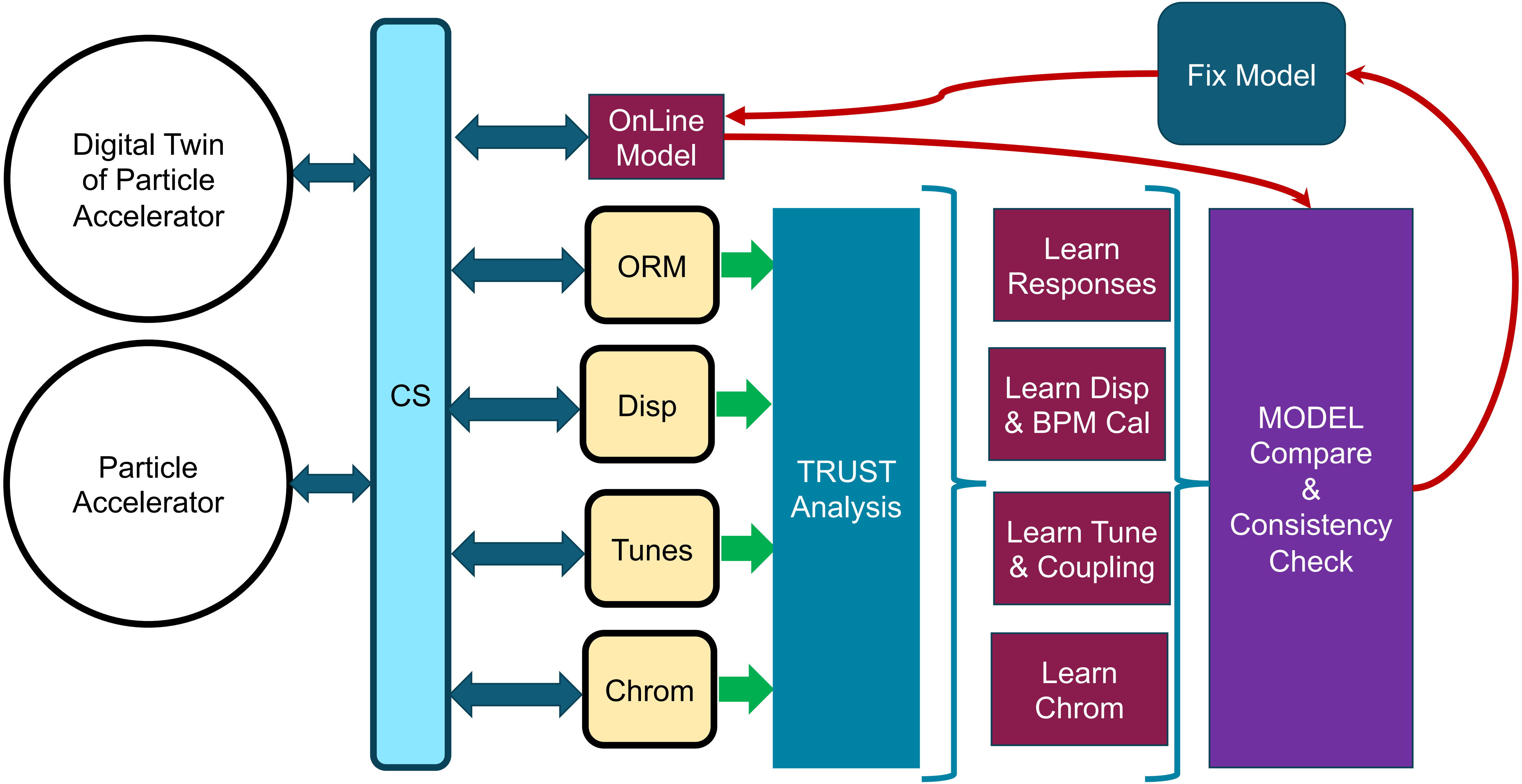
$$\begin{pmatrix} \Delta\nu_1 \\ \Delta\nu_2 \\ \cdots \\ \Delta\nu_{N-1} \\ \Delta\nu_N \end{pmatrix} = J^+_{model} \begin{pmatrix} \Delta R_{11} \\ \Delta R_{12} \\ \cdots \\ \Delta R_{n(m-1)} \\ \Delta R_{nm} \end{pmatrix}$$



Toward Self-diagnosis and Virtual diagnostics



Schematic of how the accelerator model, physical accelerator, and DT of the accelerator are related. The Parameter-NN is trained on the model accelerator dynamical data. This NN is then used to map the dynamic data of the physical accelerator to component parameters of the DT. A separate NN is trained on the output data of the DT, acting as a quick-to-evaluate surrogate of the DT. This Digital-Twin NN maps simulated component parameters to physical accelerator parameters.



Brookhaven National Laboratory

Computing?

Physics models have different computing requirements

- Offline Lattice analysis = single cpu (fast with deep memory)
- Online Lattice analysis = complete computation in ~10msec for real-time feedback, multiple fast independent cpu's
- Dynamic aperture = gpu's, HPC level
- Spin tracking = HPC, can still take days

AI/ML models can also vary

- Physics model informed Bayesian Optimization is very fast – single fast cpu works most of the time
- Deep NN requires HPC level resources

Accelerator control systems do not use HPC resources. Our paradigm needs to shift to combine Online (fast but simple) models with Offline (slow, includes more physics) models.

Summary

ML Methods we have been using and investigating

Bayesian Optimization for Gaussian Processes

- Applied to LEReC to optimized relative ion-electron trajectories and maximize cooling rate
- Planning to use in other areas: longitudinal bunch merge, IPM channel gain calibration, AGS Injection, polarization optimizations
- The method is useful for,
    - Speeding up and maintaining optimization for linear systems
    - Optimizing objectives in noisy linear systems
    - Optimizing objectives in non-linear systems

Use of Neural Networks as surrogate models → physics models, virtual diagnostics

Natural Language Processing

- Applied to elogs to provide more intelligent search
- Planning to expand use and functionality, based on user feedback
- Has vast applicability, as
    - Classifying data/information
    - Finding and classifying text in images
    - Enhances Help functionality

Brookhaven National Laboratory

- Problems we are working on next
  - Anomaly detection for operations using autoencoders
    - building temperatures, beam loss patterns, etc.
  - Learning magnet transfer functions (in AGS Booster) using ORMs with NN surrogate models
  - Improving physics models using beam-based measurements with surrogate models
  - Fast emittance measurement in CeC
  - Optimizing polarization (see Vincent's and Georg's talks)
  - Optimization of a Longitudinal Bunch Merge Gymnastic with ML
  - Ionization Profile Monitor Channel Gain Calibration with ML
  - Improving the NASA Space Radiation Laboratory beam spill (collaboration with FNAL and JPARC)
  - ML in Sirepo (RadiaSoft) for optimizing and tuning
- Community collaboration and engagement is critical to our success. Community tools, such as Xopt, Badger, and Sirepo, among others, will allow us to both benefit from the community experience but also to contribute in our own unique ways.

Publications

- B. Huang, C. González-Zacarías, S. Sosa Güitrón, A. Aslam, S. G. Biedron, K. Brown, T. Bolin, *Artificial Intelligence-Assisted Design and Virtual Diagnostic for the Initial Condition of a Storage-Ring-Based Quantum Information System*, IEEE Access, Volume 10, 2022, pp.14350-14358
- Y. Gao, W. Lin, K. A. Brown, X. Gu, G. H. Hoffstaetter, J. Morris, and S. Seletskiy, *Bayesian optimization experiment for trajectory alignment at the low energy RHIC electron cooling system*, Phys. Rev. Accel. Beams 25, 014601 – Published 7 January 2022
- Y. Gao, J. Chen, T. Robertazzi, and K. A. Brown, *Reinforcement learning based schemes to manage client activities in large distributed control systems*, Phys. Rev. Accel. Beams 22, 014601 – Published 2 January 2019
- W. Lin, M. A. Sampson, Y.C. Jing, K. Shih, G. H. Hoffstaetter, J. A. Crittenden, *Simulation Studies and Machine Learning Applications at the Coherent Electron Cooling Experiment at RHIC*, IPAC2022, Bangkok, Thailand
- Y. Gao, K. A. Brown, X. Gu, J. Morris, S. Seletskiy, W. Lin, G. H. Hoffstaetter, J. A. Crittenden, *Experiment Of Bayesian Optimization For Trajectory Alignment At Low Energy RHIC Electron Cooler,* IPAC2022, Bangkok, Thailand
- Y. Gao, K. A. Brown, P. Dyer, S. Seletskiy, H. Zhao, *Applying Machine Learning to Optimization of Cooling Rate at Low Energy RHIC Electron Cooler,* IPAC2021, Campinas, SP, Brazil

Thank you.

BNL

Kevin Brown, Ian Blackler, Sam Clark, Wenge Fu, Yuan Gao, Xiaofeng Gu, Natalie Isenberg, Y. Jing, Yongjun Li, Jennefer Maldonado, François Meot, John Morris, Sergei Seletskiy, Vincent Schoefer, Reid Smith, Nathan Urban, Dale Yu

Cornell University

Georg Heinz Hoffstaetter, Lucy Lin, David Sagan, A. Sampson

Stony Brook University

Weining Dai, Bohong Huang, Thomas Robertazzi, K. Shih

Brookhaven National Laboratory