ML for Beam Polarization Increase

Accepted ML / AI Proposal to DOE-NP FOA

Georg Hoffstaetter de Torquat
Collider-Accelerator Department, BNL and Cornell University
Georg.Hoffstaetter@cornell.edu

CAD MAC December 20, 2023

@BrookhavenLab

DE-FOA-0002875 : ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR AUTONOMOUS OPTIMIZATION AND CONTROL OF ACCELERATORS AND DETECTORS

Title: Higher RHIC polarization by Physics-informed Bayesian Learning

Budget: $1.5M, duration 2 years, start 09/01/2023 to BNL, Cornell, JLAB, SLAC, RPI

Funding through DOE-NP DE SC-0024287, contr.# 2023-BNL-AD060-FUND

Funding officer Manouchehr Farkhondeh

Requested topics:

- Efficiently extract critical and strategic information from large complex data sets
- Address the challenges of autonomous control and experimentation
- **Efficiency of operation of accelerators and scientific instruments**
- AI for data reduction of large experimental data

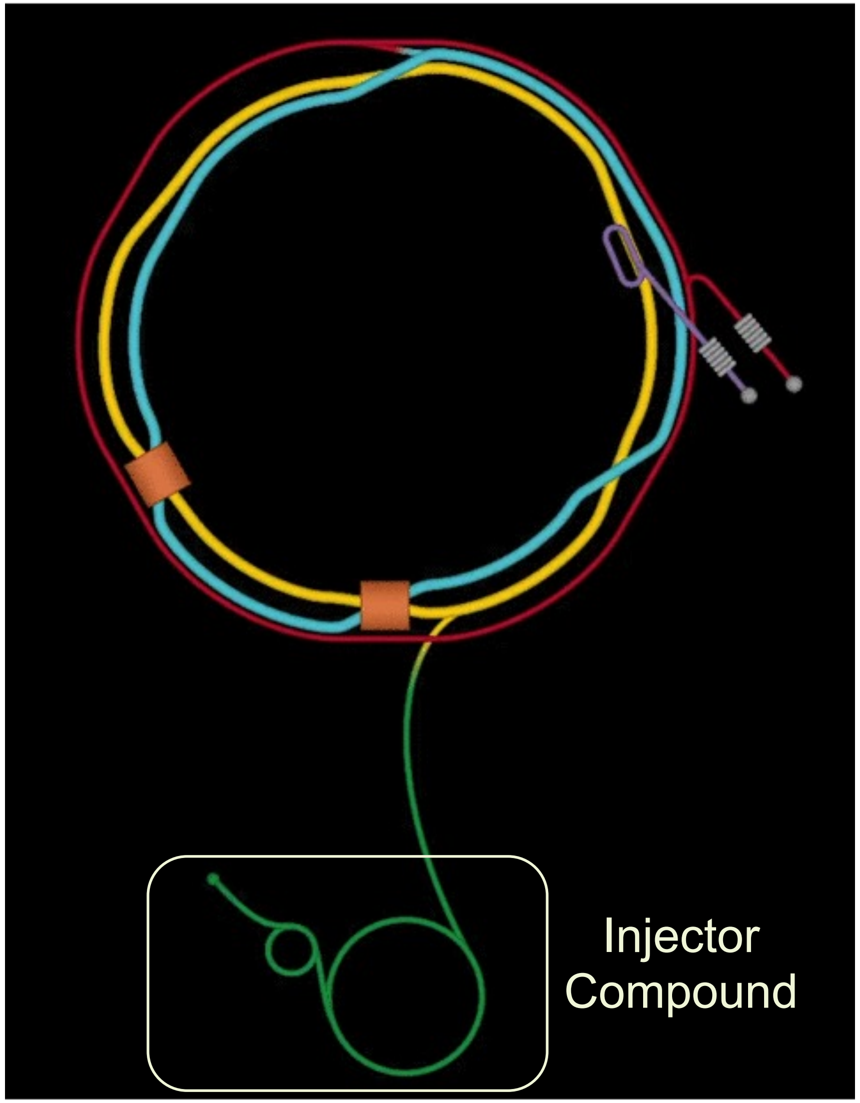
Desired result: higher proton polarization

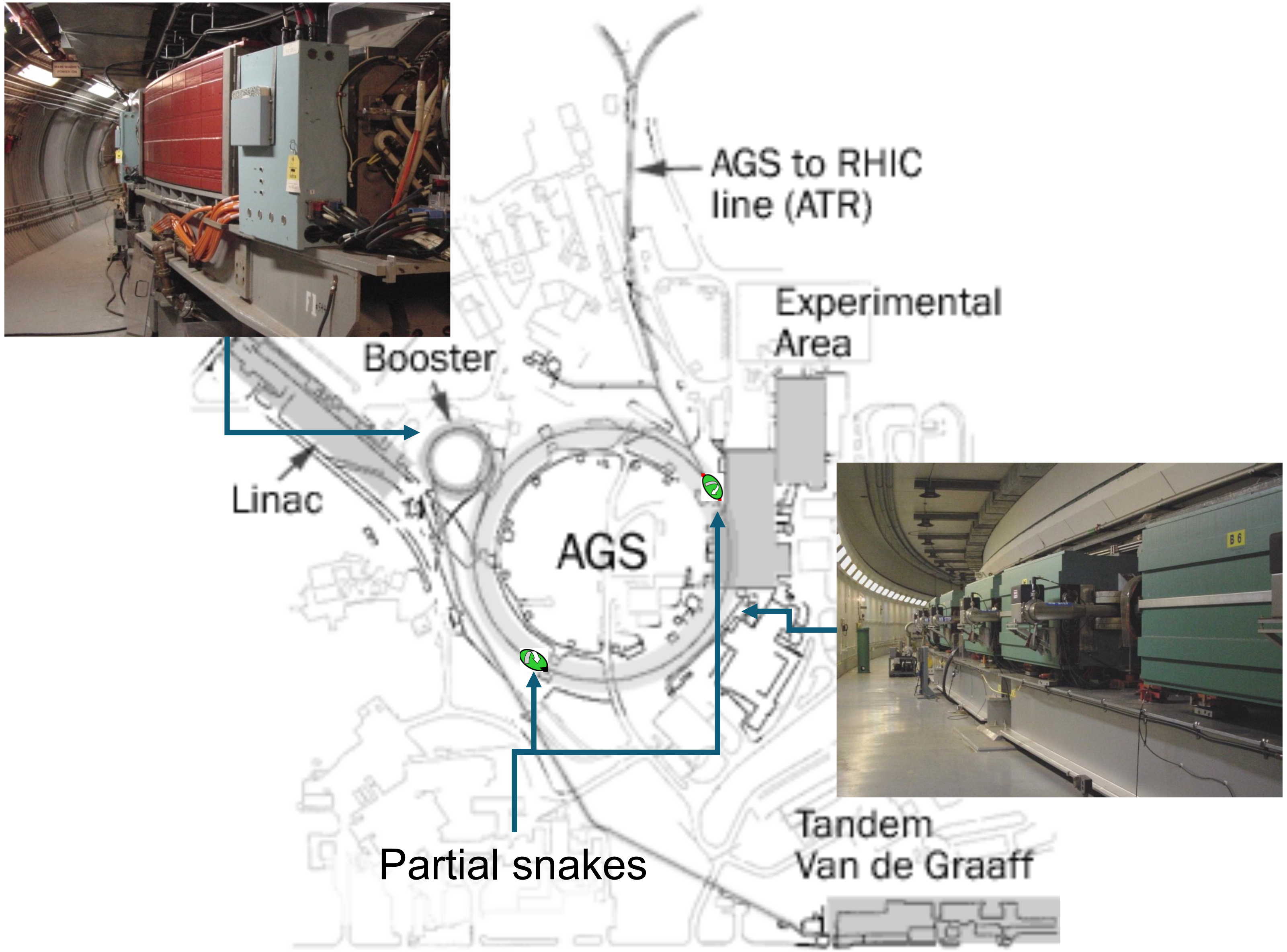
- What high-impact operational challenge can be addressed by MI/AI?
  ➔ Polarized protons.
- From the source to high energy RHIC experiments, 20% polarization is lost.
- Polarized luminosity for longitudinal collisions scales with P^4, i.e., a factor of 2 reduction!
- The proton polarization chain depends on a hose of delicate accelerator settings form Linac to the Booster, the AGS, and the RHIC ramp.
- Even 5% more polarization would be a significant achievement.

Outline

- Objective of proposed work: higher proton polarization in RHIC and the EIC.
- Polarized-proton acceleration chain.
- Potential avenues toward higher proton polarization.
  (1) Emittance reduction
  (2) More accurate timing of timed elements
  (3) Reduction of resonance driving terms
- Started Activities
- Gaussian Process (GP) Bayesian Optimization (BO) and physics informed learning.
- When is ML/AI better for accelerator operations than other feedbacks and optimizers?

The polarized proton accelerator chain



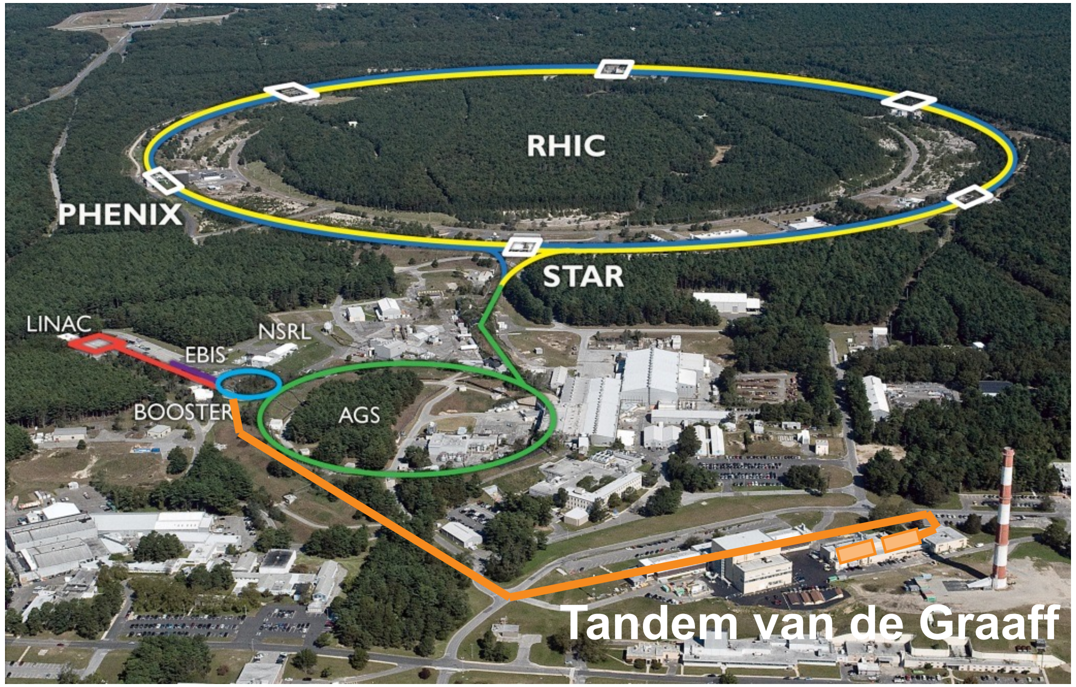


Polarization at RHIC

	Max Energy [GeV]	Pol. At Max Energy [%]	Polarimeter
Source+Linac	1.1	82-84	
Booster	2.5	~80-84	
AGS	23.8	67-70	p-Carbon
RHIC	255	55-60	Jet, full store avg*

Loss in polarization along the chain

	Relative Ramp Polarization Loss (Run 17, full run avg)
AGS	17 %
RHIC	8 %



Polarimetry available at:

- Source
- End of Linac (200 MeV)
- AGS extraction
- RHIC injection energy
- RHIC flattop

No Booster polarimeter

Topics that can improve polarization

(1) Emittance reduction

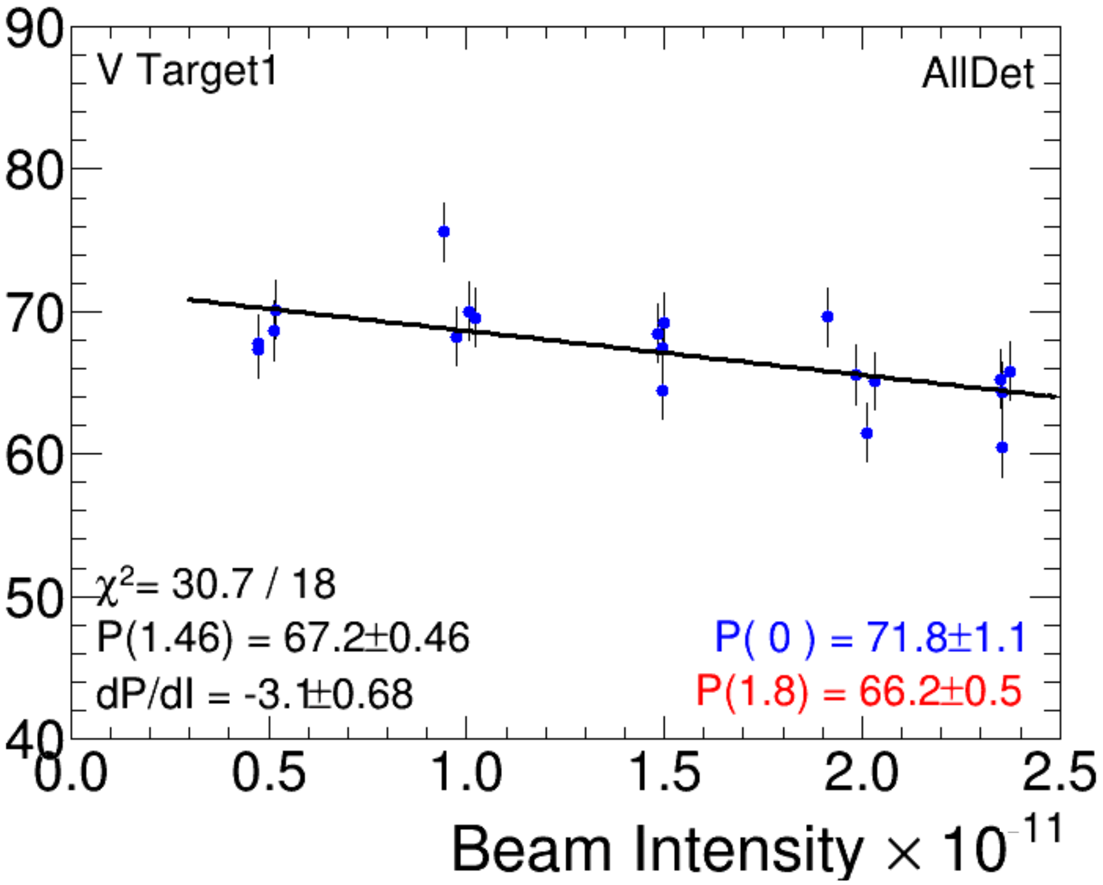
(2) More accurate timing of tune jumps

(3) Reduction of resonance driving terms

Improve Polarization at RHIC

- Figure-of-merits (FOM) for the project ("experimental outputs"): emittance, beam intensity, polarization
- Trade-offs in optimizing **FOMs**:
  - Emittance ↓ Beam intensity ↑ Polarization ↑
- Trade-offs between **controls**:
  - Beam intensity ↑ → Emittance ↑
  - Emittance ↑ → Polarization ↓
- Main areas to optimize:
  - Booster injection / capture
  - AGS bunch splitting / merging scheme
  - AGS spin resonance compensation

AGS Polarization vs. Beam Intensity



Emittance reduction → less depolarization

- Optimized Linac to Booster transfer
- Optimized Booster to AGS transfer
- Optics and orbit correction in Booster and AGS
- Beam-based alignment & calibration from orbit response in Booster and AGS.
- Bunch splitting in the Booster for space charge reduction and bunch re-coalescing at AGS top energy.

Linac to Booster transfer

Parameters to vary:

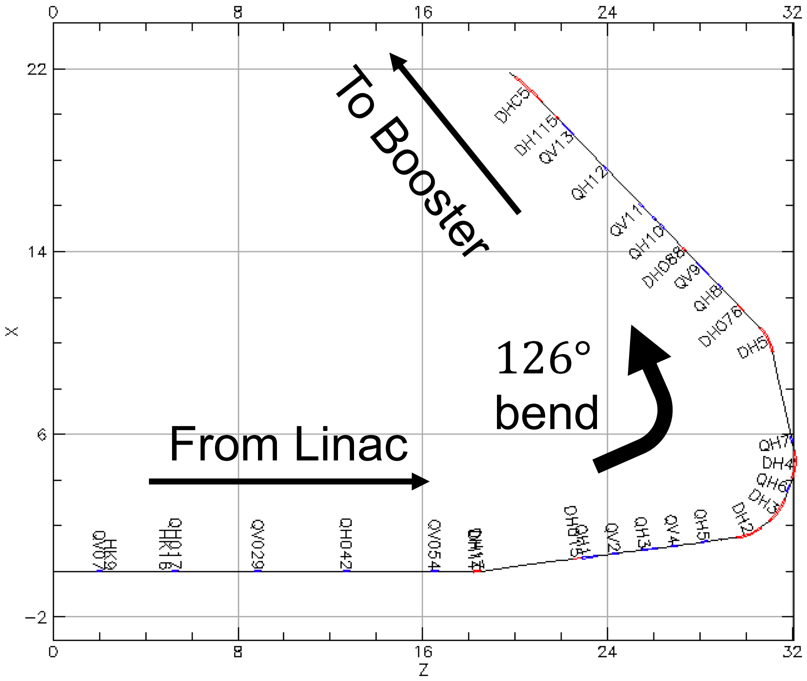
- Transfer line steers
- Main Booster dipol90e field
- Booster beta wave (stop-band quadrupoles) for tune toward ½ and minimum on the foil
- Last two linac phases
- Injection bump elements and their time profile
- Scraper amplitudes

Observables to optimize:

- Transfer efficiency linac ➔ Booster early ramp (2% absolute)
- Emittance from multi wires of the AGS transfer line (5% relative)

Ongoing project (a): Booter injection/capture

- Booster injection/early acceleration process sets maximum beam brightness for rest of acceleration though RHIC
- Linac pulse of 300 us, H^- beam ~6-9x10^{11} protons, strip through a carbon foil. Intentional horizontal and vertical scraping reduce emittance (and intensity) to RHIC requirements ~2.5x10^{11} protons
- Goal: minimize beam loss at scraper
- Controls: Linac to Booster (LtB) transfer line optics, beam size on ionization foil
- Method: Bayesian Optimization



Progress: Setup injection model incl. foil, create OpenAI Gym function for optimization training.

Ongoing project (b): AGS injection

Parameters to vary:

- Transfer line steerers
- Main AGS dipole field, RF phase, injection bumps, tunes.
- Horizontal orbit in the snakes and their optics and orbit correction.

Observables to optimize:

- Transfer efficiency Booster ➔ AGS early ramp (2% absolute)
- Emittance from two IPMs (10% relative)

Progress

- Setup detailed injection and AGS model, including symplectic snake tracking

Ongoing project (c): Booster model calibration

$$\left(I_{quad}, I_{corr}, \theta\right) \overset{model}{\Longleftarrow\!=\!=\!=\!=\!\Longrightarrow} \left(X_{BPM}^{(I_{quad}^{(1)}, I_{corr}^{(1)})}, X_{BPM}^{(I_{quad}^{(2)}, I_{corr}^{(2)})}, \dots\right)$$

- Control: power supply currents of quadrupoles and correctors
- Parameter θ: parameters that affect the orbit but not in our control → (magnet misalignments, magnet transfer functions, etc.)
- Output: orbit at the BPMs with certain current configuration
- Invert from measured BPM data to simulation model parameters, update model for digital twin.

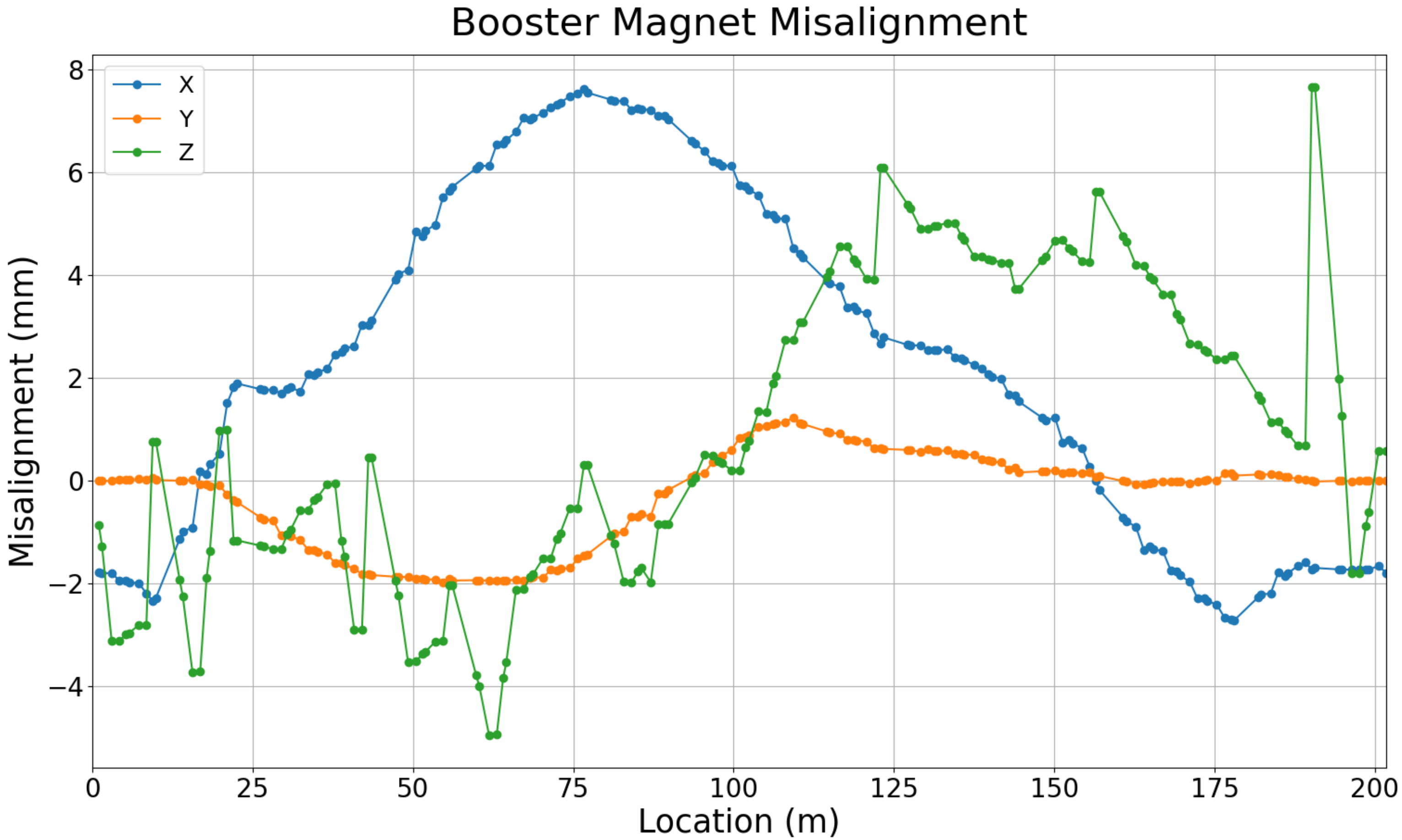
Progress: Detailed Booster model, compared to survey data and BPM readings. Write system parameter search as OpenAI gym function for ML optimization.

$$X_{BPM} = m\left(I_{quad}, I_{corr}; \theta\right) + \epsilon, \qquad \epsilon \sim N(0, \sigma)$$

$$I_{quad}^{*}, I_{corr}^{*} = \operatorname{argmax} F\left(m\left(I_{quad}, I_{corr}; \theta\right)\right)$$

Booster magnet misalignment

- Magnet location in real machine from 2015 and 2022 survey data for quadrupoles and dipoles
- Trouble with making physics simulation with misalignment agree with measured orbit data



Challenge: How well can we determine the alignment by orbit-response evaluation?

Where do we put AI/ML?

- ORM will give us
  - BPM and Corrector Anomalies (Trust Analysis)
  - Gradient errors for given conditions
  - Beta-deviations from model
- Dispersion measurements give us
  - BPM Consistency check for given dp/p (BPM Anomalies)
  - Coupling through longitudinal motion (very slow, typically)
- Tune measurements
  - Betatron tune and coupling = destructive measurement in Booster/AGS
  - Tune, Chrom, coupling, emittance, dp/p from RHIC Schottky
- Chromaticity measurements – need to change energy and measure tune
- Orbit Measurements – parasitic = most are time averaged, some turn by turn
- Linear model + small nonlinearities with NN model

Ongoing project (d): AGS model calibration

Parameters to vary:

- Corrector coils (24 per Booster plane)
- Corrector coils (48 per AGS plane)

Observables to optimize:

- BPM readings (24 x&y in the Booster) (100um accuracy)
- BPM readings (72 x&y in the AGS) (100um for 2mm size at 25GeV)

Progress:

- Detailed model of AGS incl. differentiable snakes, symplectic tracking, orbit and optics compensation of snakes for all energies.

Space-charge emittance increase

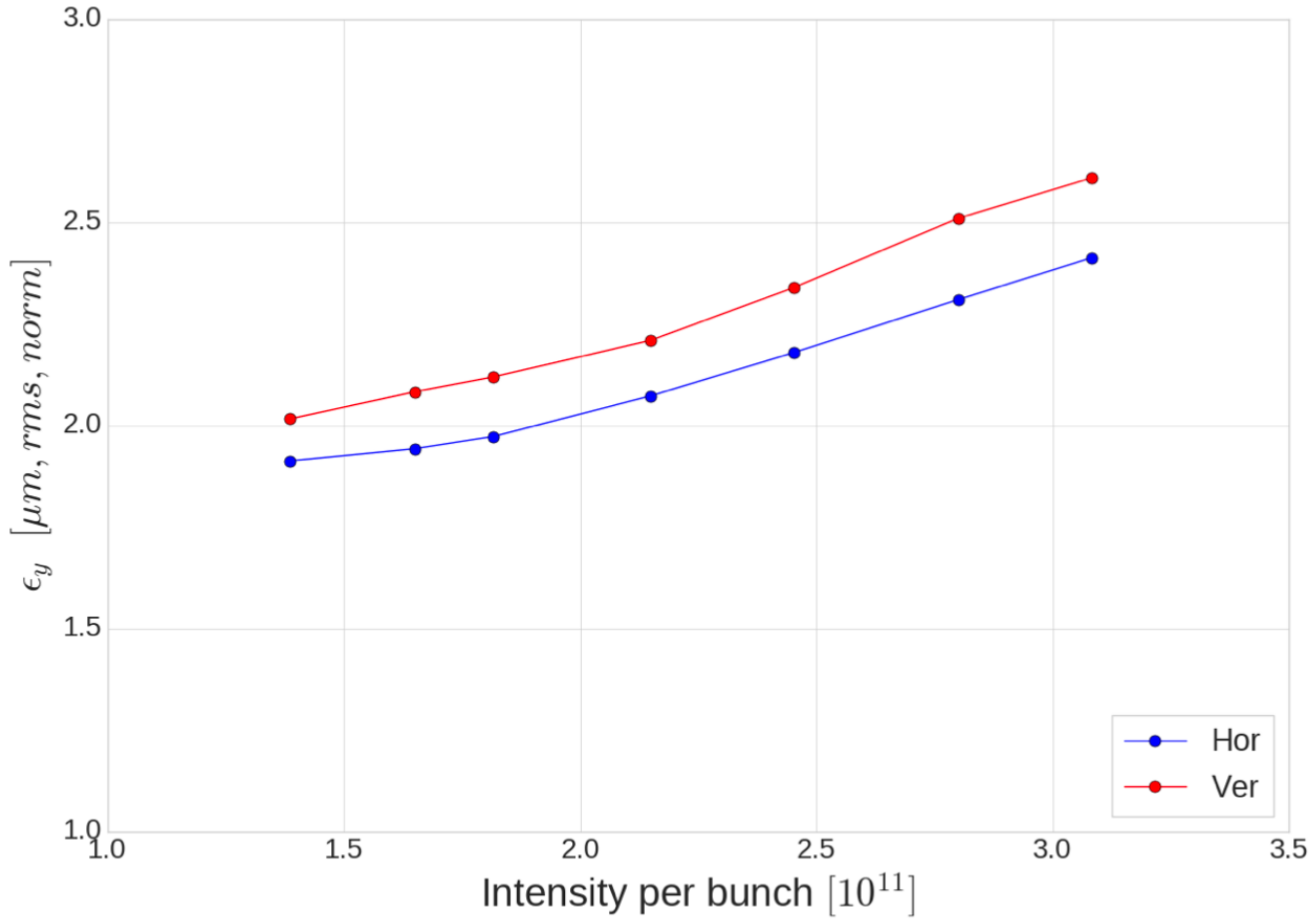


Figure 3.168: Normalized transverse emittances of polarized proton beam at AGS extraction energy ($\gamma = 25.5$) as a function of intensity.

➔ Splitting bunches before AGS acceleration can reduce the emittance.

Ongoing project (e): Bunch splitting / coalescing

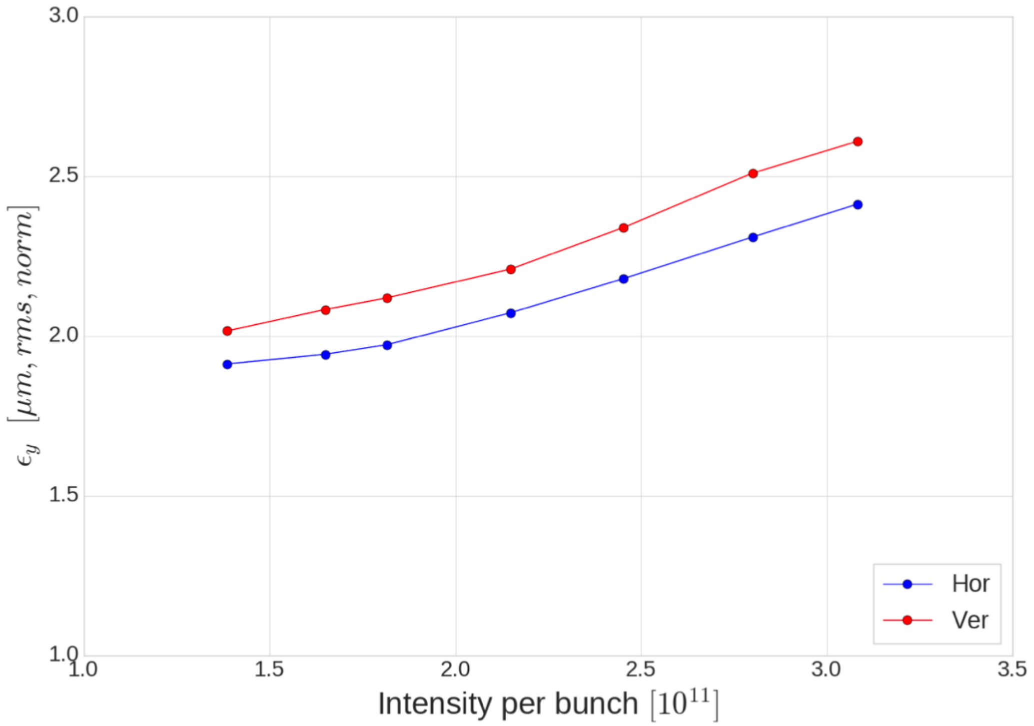


Figure 3.168: Normalized transverse emittances of polarized proton beam at AGS extraction energy ($\gamma = 25.5$) as a function of intensity.

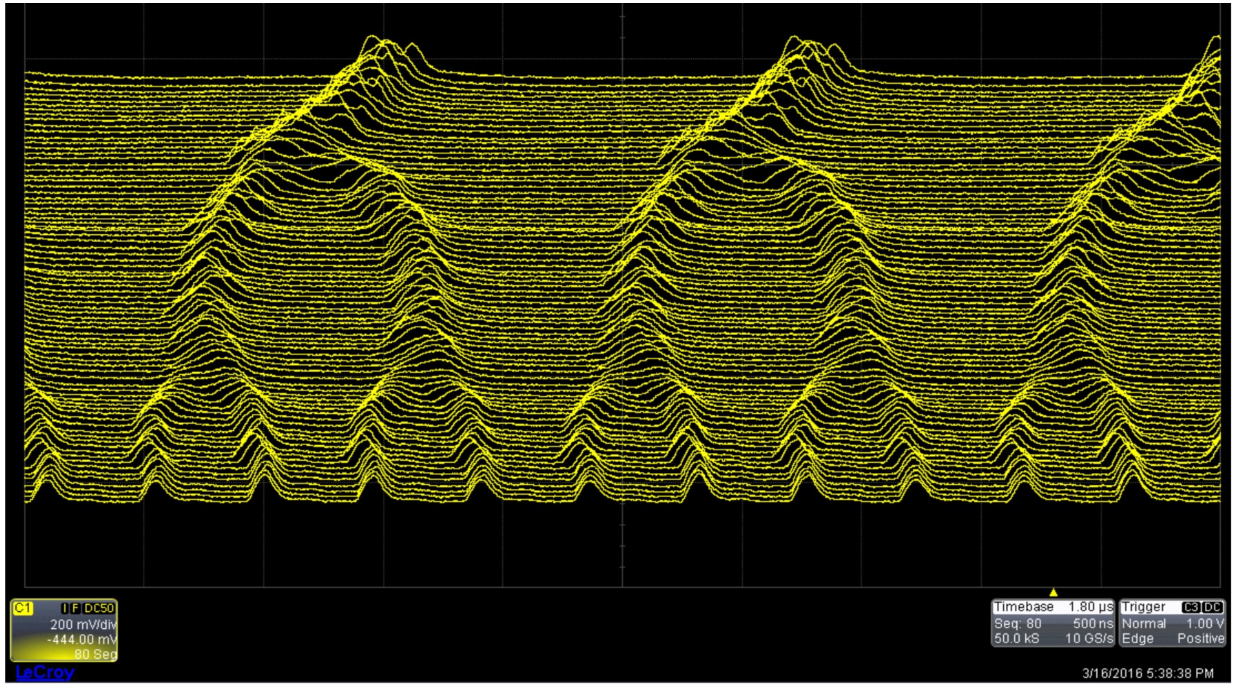


Figure 3: Mountain range display of the wall current monitor signal for the 4:2:1 Booster merge used for EBIS Au.

- ➔ Splitting bunches before AGS acceleration can reduce the emittance.

Progress: Setup Bmad model of bunch merging to find adiabatic parameters without emittance increase.

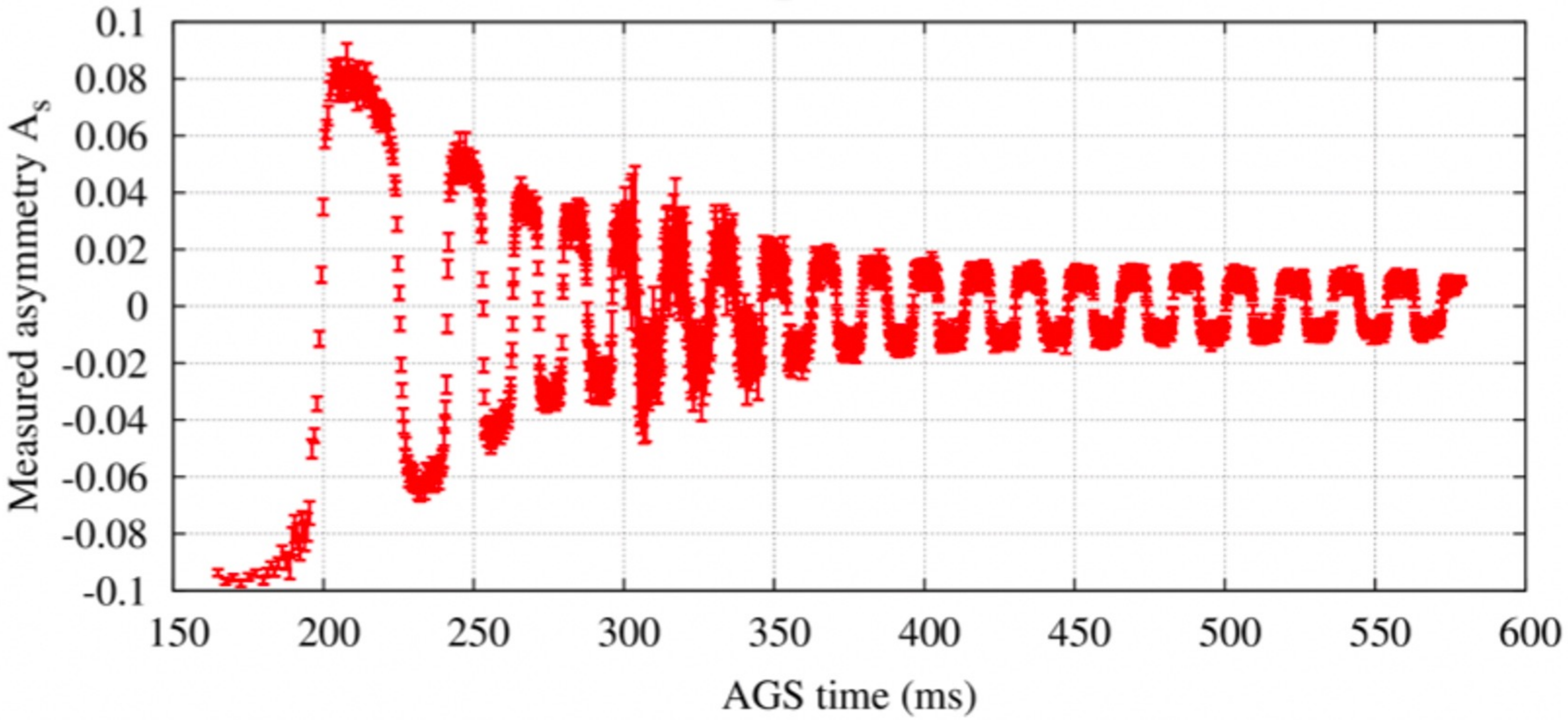
Future project (f): Timing of tune jumps

The G-gamma meter and accurate energy vs. time

(1) Measure the energy by orbit + revolution frequency measurement

(2) Measure of energy by field + revolution frequency measurement

(3) Measure energy by spin flip at every integer spin tune



Combined optimization

➔ better timing

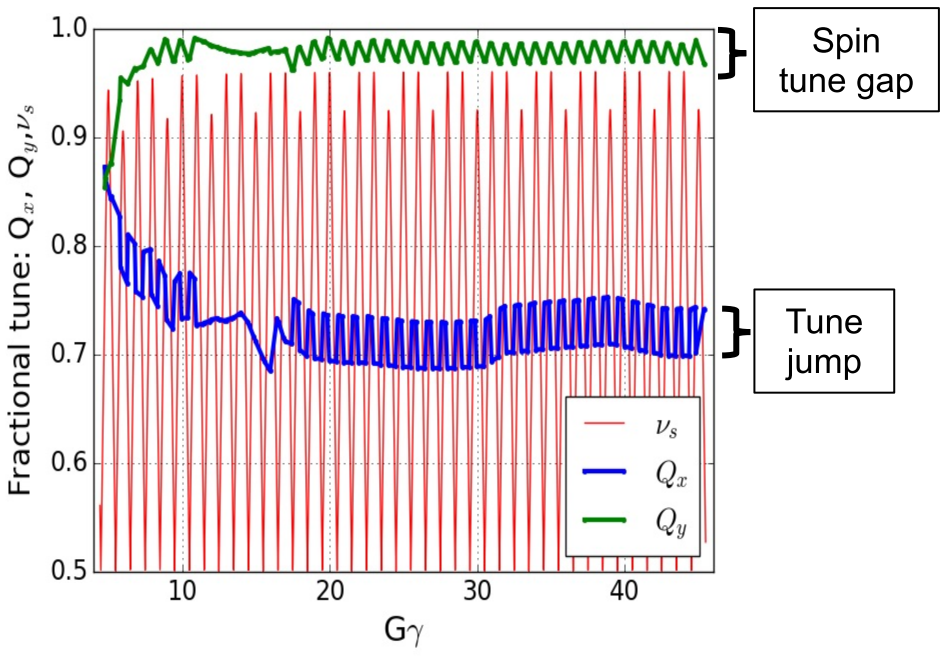
➔ higher polarization

Ongoing project (g): AGS resonance compensation

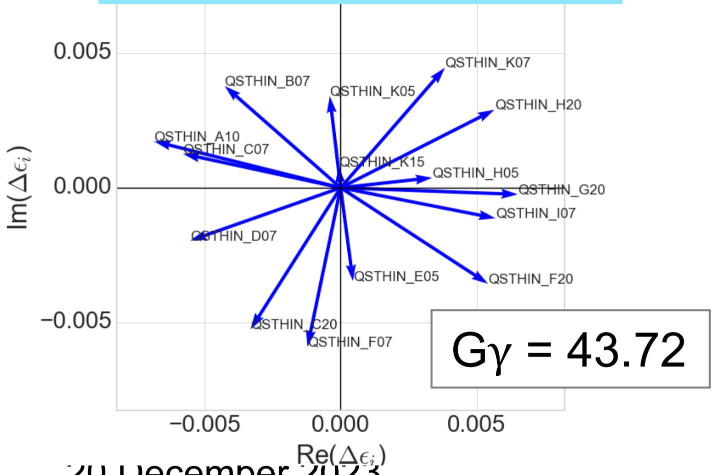
- Partial snakes in the AGS keep the spin tune away from the integer (>0.96), avoiding vertical resonances
- Horizontal resonances remain, currently ‘jumped’ by moving the horizontal tune through the resonance
    - Each resonance is weak (~0.1% polarization loss)
    - But there are many of them (82)
- Proposal to use 15 pulsed skew quadrupoles to eliminate residual resonances (as described by Vincent Schoefer)
- Goal: minimize resonance strengths by timed skew quads
- Method: Reinforcement Learning / Bayesian Optimization

Progress: detailed Bmad model incl. differentiable snake model, symplectic tracking, orbit and optics correction, and various methods of resonance strength evaluation.

Betatron and spin tunes during AGS ramp

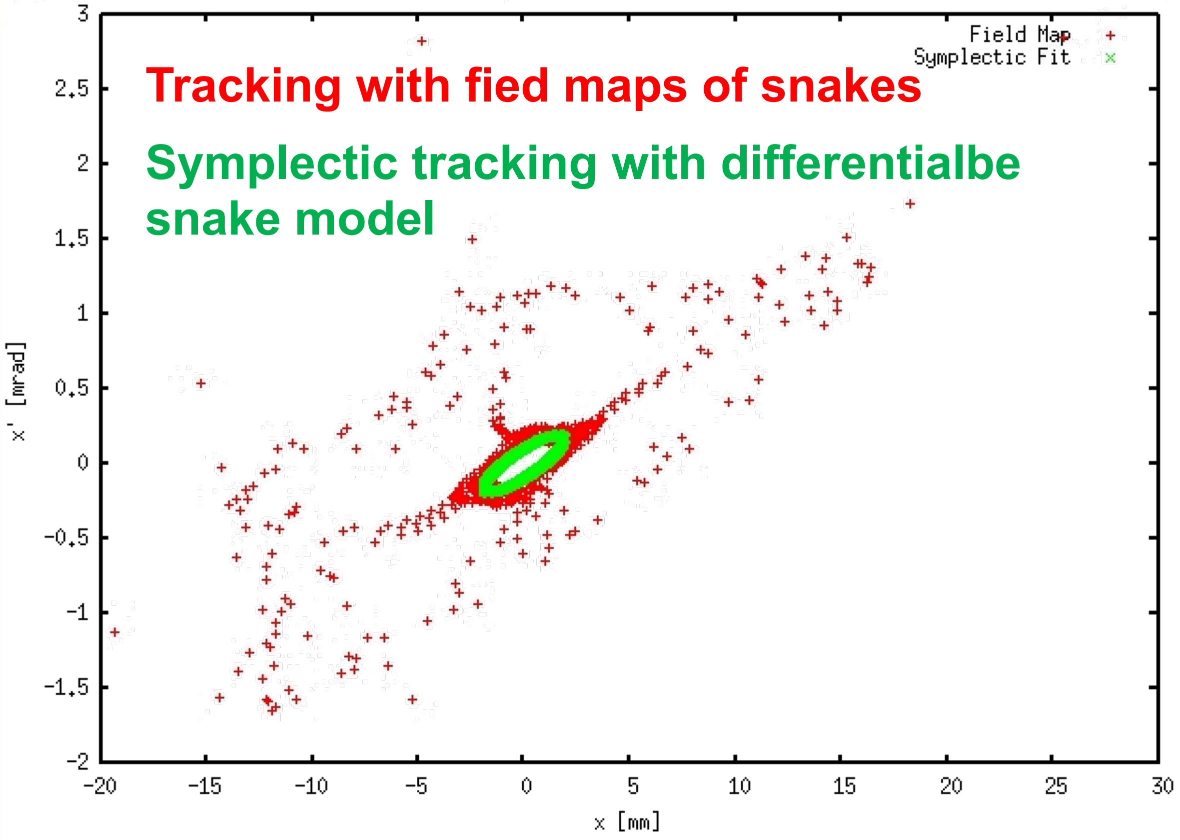


Spin resonance terms from skew quads in AGS

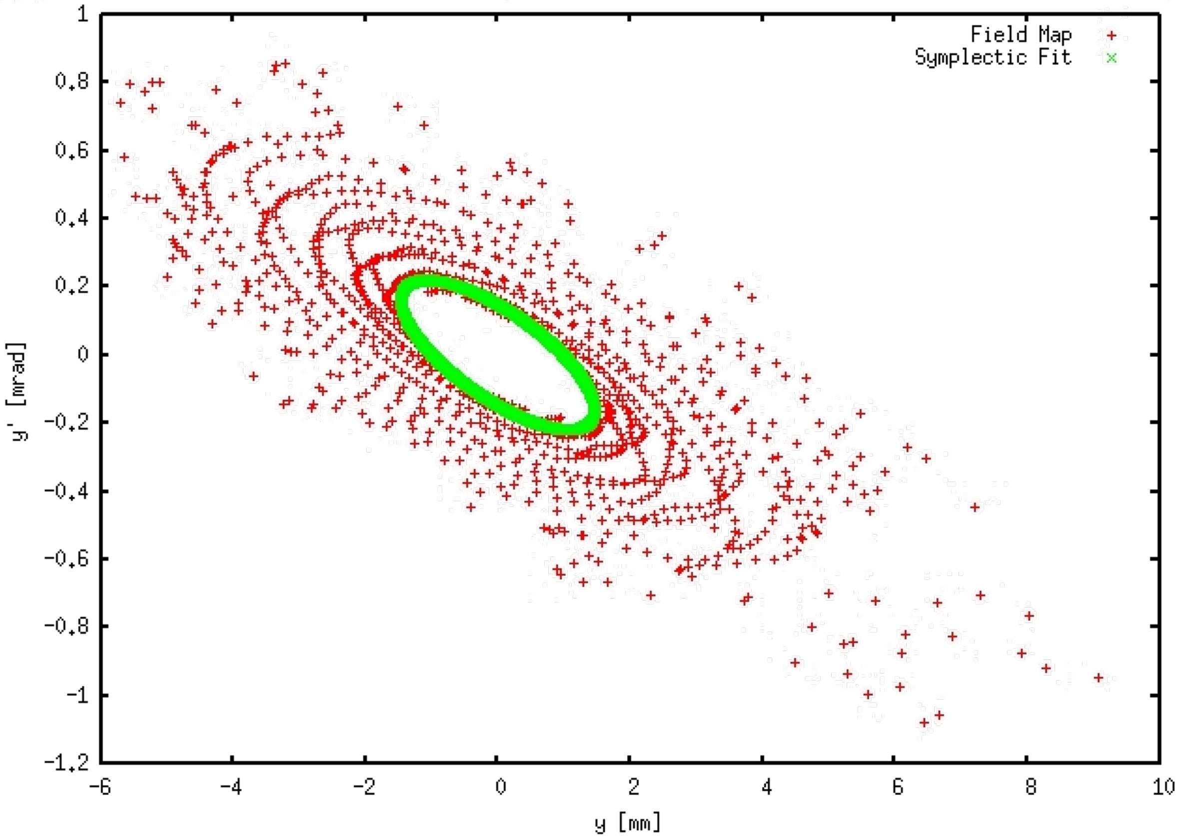


For ongoing project (g): Importance of snake modeling

Horizontal Phase Space



Vertical Phase Space



Note: All 10,000 turns of symplectic tracking were used in the green

For ongoing project (g): AGS snake modeling

I used a series expansion of solutions to Maxwell's equations such that each term is an exact solution.

This allows me to truncate the expansion at any order and still precisely satisfy Maxwell's equations.

The following is an example of 1 term in such a series solution:

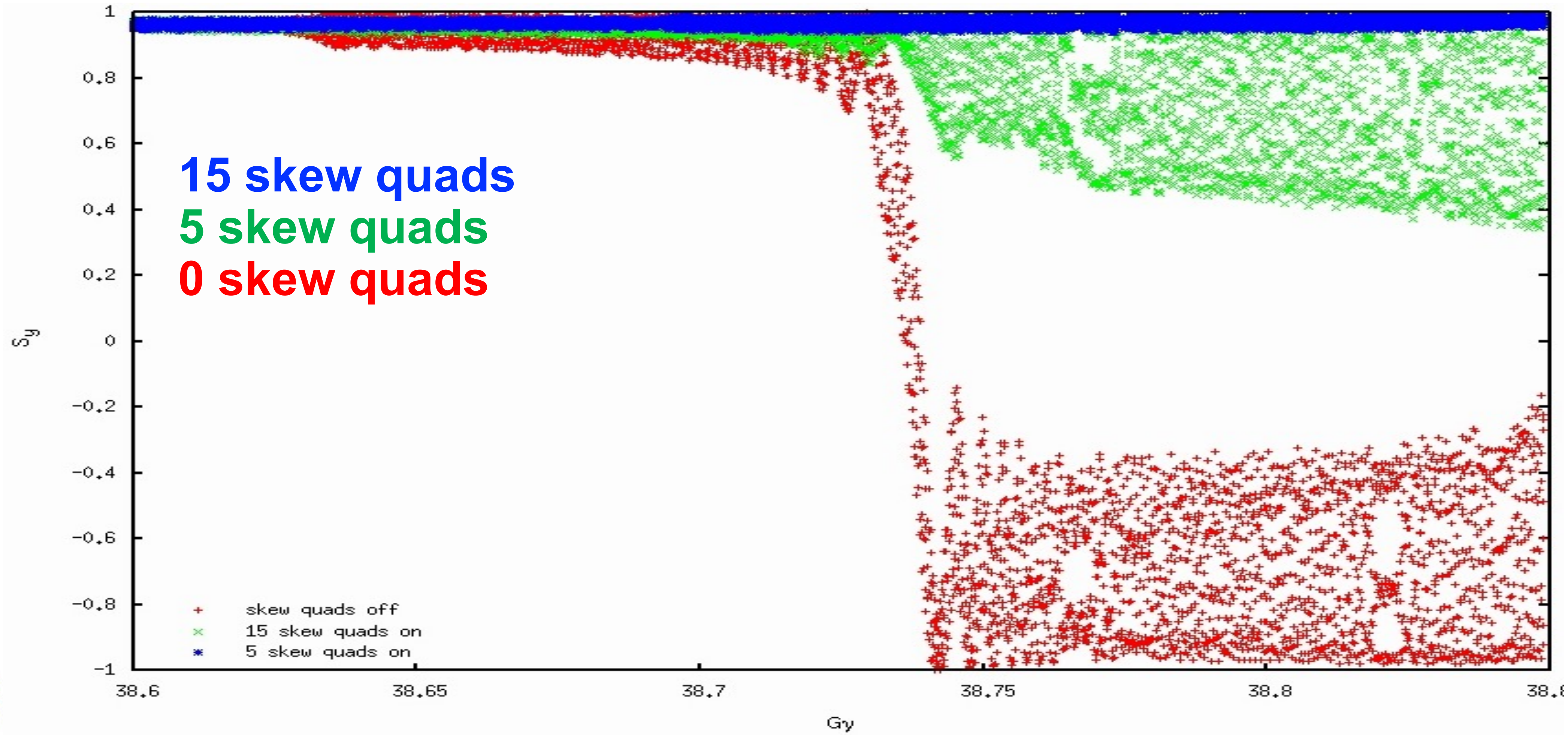
$$
\begin{aligned}
B_x &= \quad A \quad \cosh(k_x(x+x_0))\,\cos(k_y(y+y_0))\,\cos(k_z z+\phi_z)\\
B_y &= -A\,\frac{k_y}{k_x}\,\sinh(k_x(x+x_0))\,\sin(k_y(y+y_0))\,\cos(k_z z+\phi_z)\\
B_s &= -A\,\frac{k_z}{k_x}\,\sinh(k_x(x+x_0))\,\cos(k_y(y+y_0))\,\sin(k_z z+\phi_z)
\end{aligned}
$$

By fitting the field data to 300 such terms, I recover an expression that exactly solves Maxwell's equations (i.e. satisfies symplecticity for any particle trajectory) and agrees with the simulated data with an RMS deviation of 0.02 T in the relevant region. Max field is 2T.

With a differentiable model, TPSA (DA) techniques become available, e.g. normal forms

For ongoing project (g): AGS polarization tracking

Tracking through a single horizontal spin resonance with very large emittance to visualize depolarization. *This is not the real case since each resonance is much weaker and causes less than 0.1% depolarization*:

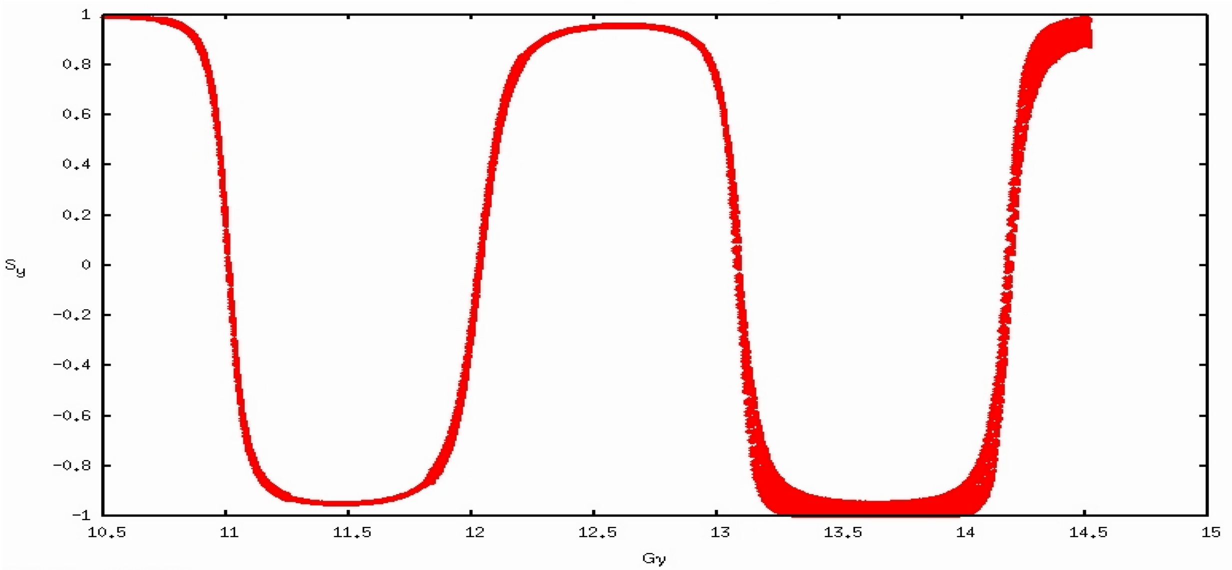


For ongoing project (g): Need of optics correction

In the following, we track through a few resonances with realistic transverse emittance before transition energy (at Gγ~15).

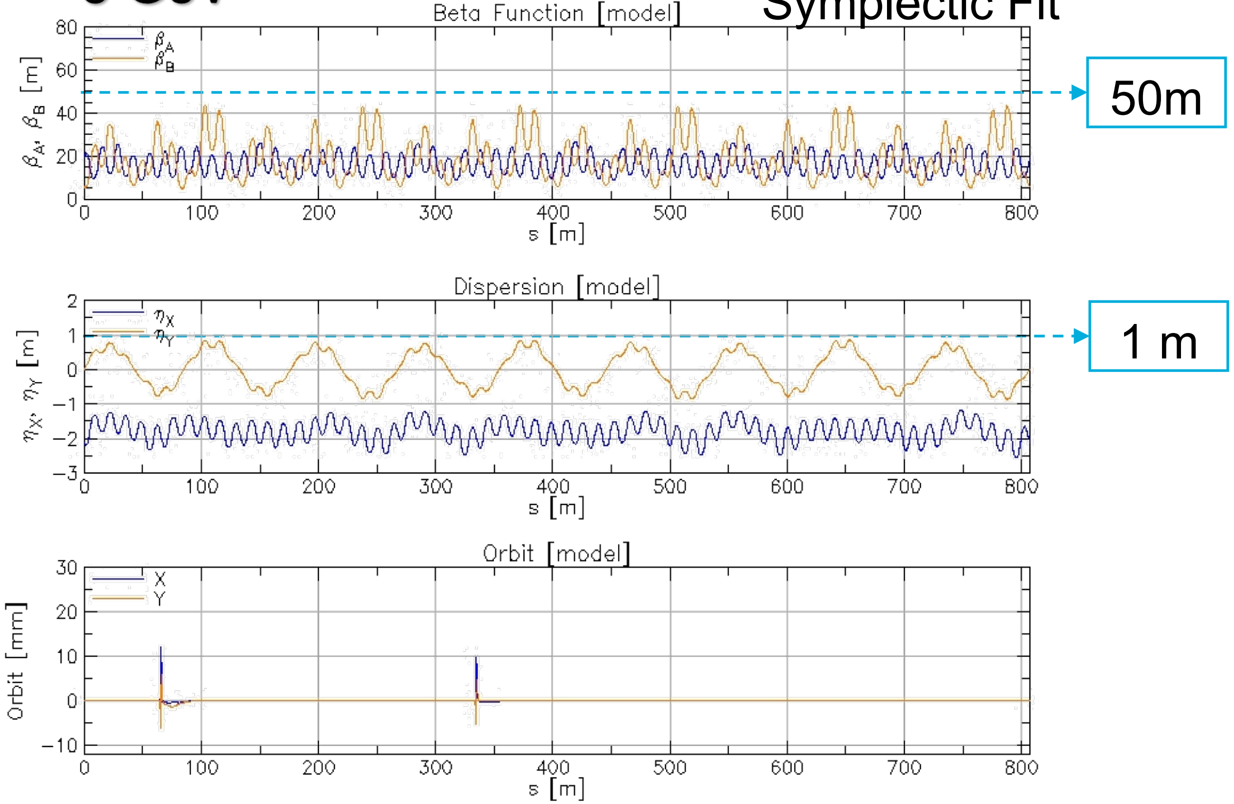
The transverse emittance blows up due to vertical dispersion and optics errors after Gγ~13, decreasing polarization.

In response to yesterday's question, no problem has yet been observed at 1/3 spin tune.



8 GeV

Symplectic Fit



Activities

Kickoff Collaboration meeting @ Cornell, August 25, 2023 – successful in person

Weekly meetings Mondays 3:30pm of all collaborators

Interface with weekly meeting Wednesdays 11am on digital twins from Linac to AGS

Semi-weekly ML/AI software meeting Friday's on beam & ML computer standards

Teams have formed for projects

(a) Booster injection, (b) AGS injection

(c) Booster model calibration, (d) AGS model calibration

(e) Bunch splitting/coalescing, *(f) Timing*

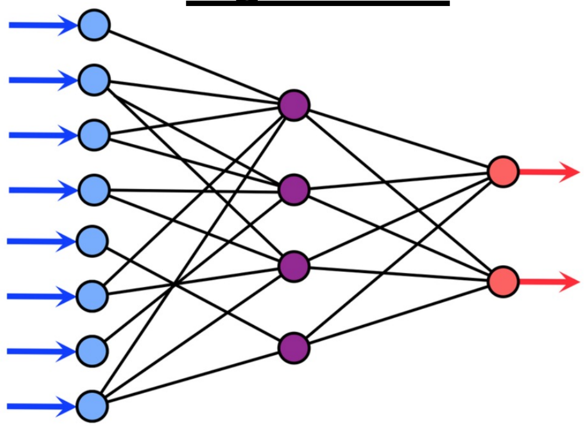
(g) Resonance minimization.

(h) Combined and verified evaluation of existing emittance measurements (Radiasoft)

Future: Digital twin and Optimal control

Control parameters + Environmental factors

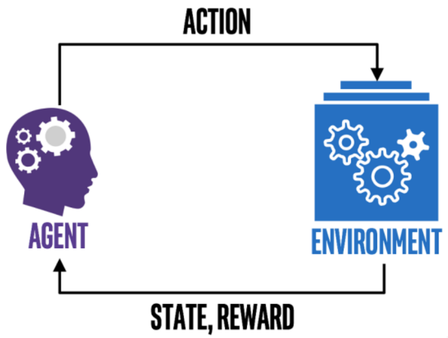
Digital Twin

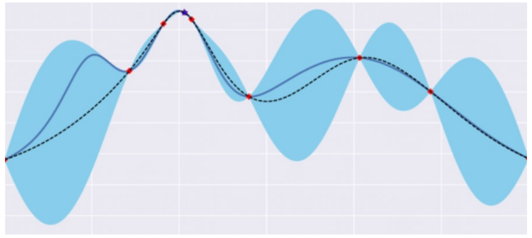


Evaluation / Analytic Metric (e.g., agreement with real measurements)

Find best settings for desired machine state

AI/ML Methods (RL, BO, etc.)





Dominant Participants

- Kevin Brown, Bhawin Dhital, Yuan Gao, Levente Hajdu, Kiel Hock, Natalie Isenberg, Chuyu Liu, Linh Nguyen, Vincent Schoefer, Nathan Urban, Keith Zeno
- Eiad Hamwi, Georg Hoffstaetter de Torquat, David Sagan
- Weining Dai, Bohong Huang, Thomas Robertazzi
- Yinan Wang
- Auralee Edelen
- Malachi Schram
- Radiasoft: Nathan Cook, Jon Edelen, Chris Hall

Summary

- A proposal to DOE-NP has been accepted for the enhancement of proton polarization using ML/AI. Goal: 5%.

- Several accelerator optimizations can impact polarization.

- These topics are of the type suitable for Bayesian Optimization

- Excellent team has formed, items being addressed:
- Emittance reduction (orbit, optics, bunch splitting)
- More accurate timing of quadrupole jumps (G-gamma meter)
- Reduction of resonance driving terms (Horizontal spin matching with skew quads)
- Accelerator studies in the next run will be aimed to improve emittances and polarization toward the run-end already.

Publications

ML/AI efforts at BNL/CAD

- 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

Polarized proton beams at BNL

- K. Zeno, An overview of Booster and AGS Polarized Proton Operations during Run 17, BNL-114742-2017-TECH (10/2017)
- K. Zeno, Run 16 Tandem gold performance in the injectors and possible improvement with AGS type 6:3:1 bunch merge in the Booster, C-A/AP/576, (10/2016)
- V. H. Ranjbar, Approximations for crossing two nearby spin resonances, Phys. Rev. ST-AB 18, 014001 (2015)
- Y. Dutheil, L. Ahrens, H. Huang, F. Méot, A. Poblaguev, V. Schoefer, K. Yip, Energy Calibration and Tune Jumps Efficiency in the PP APS, in Proc. PAC2014, Dresden/Germany (2014)
- V. H. Ranjbar, M. Bai, H. Huang, A. Marusic, M. Minty, V. Ptitsyn, Experimental Effects of Orbit on Polarization Loss in RHIC, Proc. IPAC2012, New Orleans/LA (2012)
- V. Schoefer, L. Ahrens, K.A. Brown, J.W. Glenn, H. Huang, Optics Error Measurements in the AGS for Polarized Proton Operation, Proc. PAC2011, New York/NY (2011)
- V. Schoefer, Using betatron coupling to suppress horizontal intrinsic spin resonances driven by partial snakes, Phys. Rev. AB 24, 031001 (2021)
- V. Schoefer, AGS Horizontal Resonance Compensation Overview, Presentation at BNL (2022)

Thank you and Questions?

Backup slides

Started Activities

(a) Emittance reduction by linac 2 booster optimization:

i) Adjust the model to hardware and alignment data for the L2B line

ii) Simulate scraping of beam in the Booster, including ionization foil with Bmad

iii) Write Bmad function that produces loss rate vs. last two L2B correctors and bring to OpenAI Gym format.

iv) Apply established ML codes to this function.

(b) Emittance reduction by B2A transfer:

i) Adjust the Bmad model to hardware and alignment data for the B2A line

ii) Simulate Booster phase space transfer through this line and compare to measured harp profiles

Started Activities

(c) Emittance reduction by booster optimization

i) Compare Bmad to established Zgoubi results and compare to BPM/alignment measurements

ii) Establish booster-orbit response to corrector changes and produce function that gets minimized by optimizing system parameters, e.g. element alignments, bring into OpenAI gym format.

iii) Apply established ML codes to this function.

(d) Emittance reduction by re-bucketing

i) Adapt Bmad bunch-merging code for the EIC’s RCS to the Booster (bunch splitting) and AGS (bunch merging)

ii) Compare Bmad to established Python re-bucketing code for Booster and AGS

iii) Write a function that characterizes bunch splitting/merging efficiency from RF parameters for ML/AI optimization.

Started Activities

(f) *Improved timing by combining 3 gamma meter measurements*

(g) Depolarization reduction by Skew-quad resonance minimization in the AGS:

i) Detailed Bmad model of the AGS at all energies overcoming:

1) Non-symplecticity of tracking through field maps of snakes
2) Introduce differentiable models of snake fields and match these to their field maps
3) Compensate the closed orbit, optics, coupling, and dispersion (esp. vertical) for these maps at all energies

ii) Compute resonance strength of the Bmad model and compare to established results

1) Compare resonance strength to previous measurements

2) Define optimized skew-quad compensation schemes at all 82 resonances, minimizing vertical dispersion, coupling, optics errors, and speed of skew-quad changes

(h) Combined and verified evaluation of existing emittance measurement techniques (through Radiasoft)

a) Technique discussed, Team is formed

Optimizers for different applications

less ← assumed knowledge of machine → more

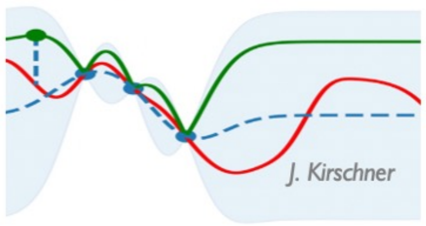
Model-Free Optimization

Observe performance change after a setting adjustment

→ estimate direction or apply heuristics toward improvement

gradient descent
simplex
ES

Model-guided Optimization

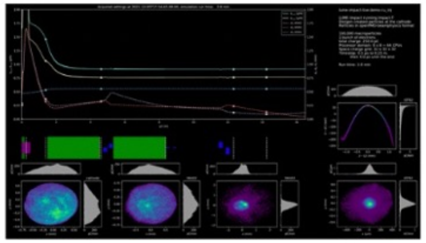


Update a model at each step

→ use model to help select the next point

Bayesian optimization
reinforcement learning

Global Modeling + Feed-forward Corrections



Make fast system model

→ provide initial guess (i.e. warm start) for settings or fast compensation

ML system models +
inverse models

Courtesy Auralee Edelen

Characteristics of involved optimizations

1. Optimal parameter settings are hard to find, and the optimum is difficult to maintain.
2. The data to optimize on has significant uncertainties.
3. Models of the accelerator exist.
4. A history of much data is available and can be stored.

Is this type of problem suitable for Machine Learning?

Why would ML be better suited than other optimizers and feedbacks?

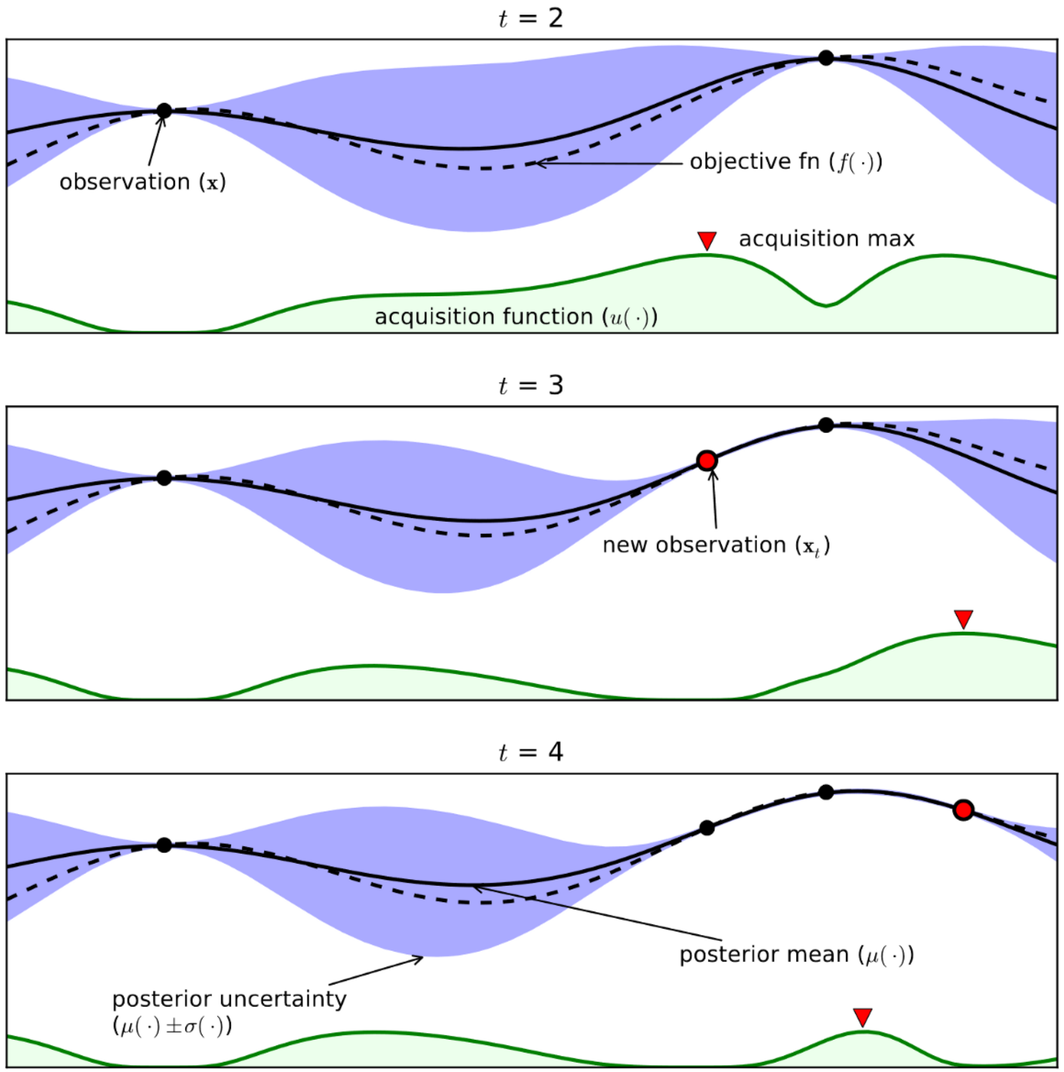
Acquisition Function

[2]

- Guide how input space should be explored during optimization

- Combine predicted mean and variance from Gaussian Process model
  - Probability Improvement (PI)
  - Expected Improvement (EI)
  - **Upper Confidence Bound (UCB)**

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



Advantages of Bayesian Optimization

Summary of optimization methods

	Nelder-Mead	Gradient descent	Powell / RCDS	L-BFGS	Genetic algorithm	Bayesian optimization
Sample efficiency	Medium	Medium	Medium/high	Medium/high	Low	High
Computational cost of picking the next point	Low/Medium	Low	Low	Low	Medium (e.g. sorting)	High (esp. in high dimensions)
Multi-objective	No (but can use scalarization)	No (but can use scalarization)	No (but can use scalarization)	No (but can use scalarization)	Yes	Yes
Sensitivity to local minima	High (but can use multi-start)	High (but can use multi-start)	High (but can use multi-start)	High (but can use multi-start)	Low	Low (builds a **global** model of f)
Sensitivity to noise	High	High	High (Powell) Low (RCDS)	High	Medium	Low (can model noise itself)

Summary of optimization methods

	Nelder-Mead	Gradient descent	Powell / RCDS	L-BFGS	Genetic algorithm	Bayesian optimization
Requires to compute or estimate derivatives of f	No	Yes	No	Yes	No	No
Evaluations of f *inherently* done in parallel	No	No	No	No	Yes	No
Hyper-parameters	Initial simplex	Step size: α (+momentum: β)	# fit points Noise level	Accuracy of hessian estimate	• Population size • Mutation rate • Cross-over rate • Number of generations	• Kernel function • Kernel length scales, amplitude • Noise level • Acquisition function

Why is Bayesian Optimization suitable?

1. The data to optimize on has significant uncertainties

➔ No derivatives have to be computed.

2. Models of the accelerator exist

➔ the expected functional form can be included in the function search (Physics-informed learning)

3. A history of much data is available and can be stored

➔ All past data are included to model the function to be optimized.