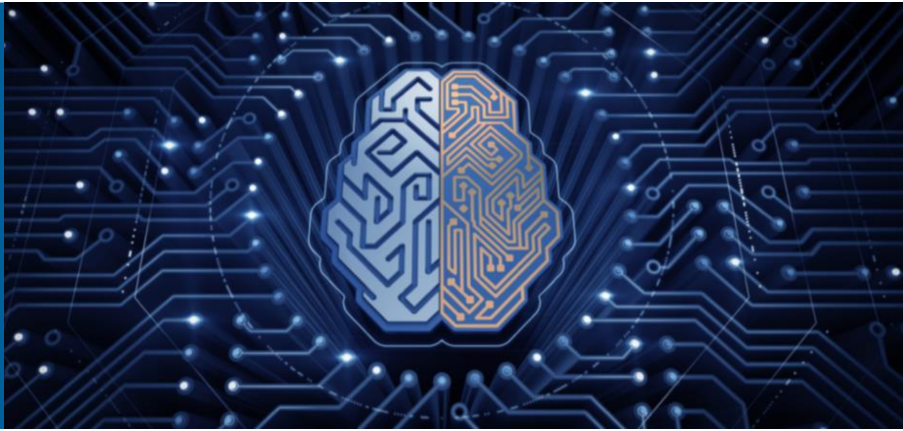
2022 DOE/NP AI-ML DATA SCIENCE PI EXCHANGE MEETING

Argonne NATIONAL LABORATORY | 75 1946–2021

USE OF AI-ML TO OPTIMIZE ACCELERATOR OPERATIONS & IMPROVE MACHINE PERFORMANCE



PRESENTER

Brahim Mustapha
Physics Division
Argonne National Laboratory

CONTRIBUTORS

Jose Martinez, Postdoc (ANL/PHY)
Ian Sugrue, Student (NIU, ANL/HEP)
Anthony Tran, Student (MSU/FRIB)

John Power (ANL/HEP)
Philippe Piot (NIU, ANL/APS)
Nathan Krislock (NIU)
Yue Hao (MSU/FRIB)

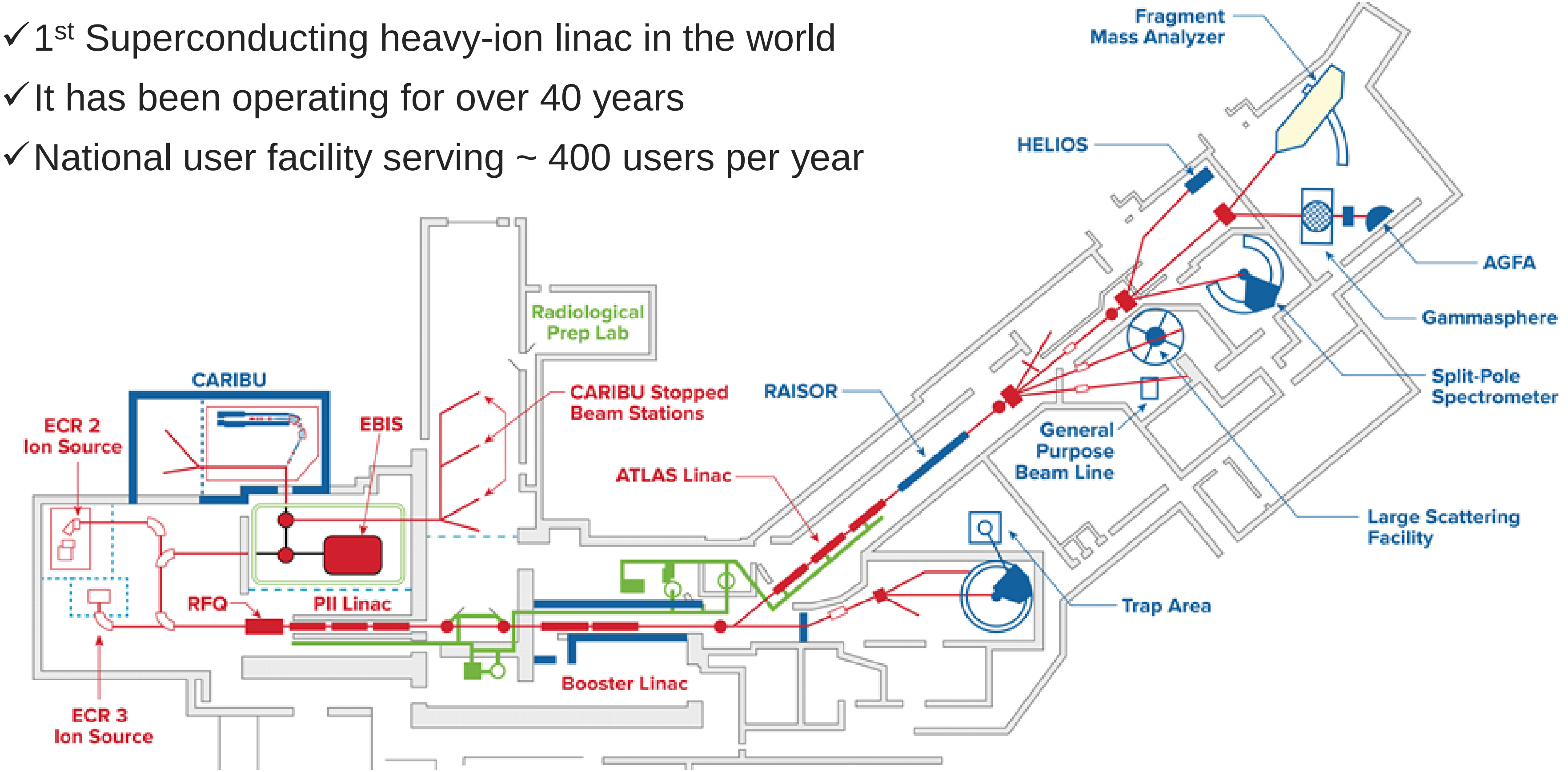
November 30th, 2022
DOE/NP (Via Zoom)

OUTLINE

- ❑Brief Overview of the Project and the Team
- ❑Project Status and Summary of Progress
- ❑Progress & Highlights at ATLAS
- ❑Progress Highlights at AWA and FRIB
- ❑Summary & Future Plans

ATLAS: ARGONNE TANDEM LINEAR ACCELERATOR SYSTEM

- ✓ 1st Superconducting heavy-ion linac in the world
- ✓ It has been operating for over 40 years
- ✓ National user facility serving ~ 400 users per year



BRIEF OVERVIEW OF THE PROJECT

Use of artificial intelligence to optimize accelerator operations and improve machine performance

- ❑ At ATLAS, we switch ion beam species every 3-4 days … → Using AI could streamline beam tuning & help improve machine performance
- ❑ The main project goals are:
  - **Data collection, organization and classification, towards a fully automated and electronic data collection for both machine and beam data**… established
  - **Online tuning model to optimize operations and shorten beam tuning time in order to make more beam time available for the experimental program** … made good progress
  - **Virtual model to enhance understanding of machine behavior in order to improve performance and optimize particular/new operating modes** ... started

THE TEAM / COLLABORATION

- ANL / PHY: B. Blomberg, D. Stanton, J. Martinez and C. Dickerson
  - J. Martinez is a postdoc focused on ATLAS

- MSU / FRIB: Y. Hao and A. Tran (PhD student started in May'21)
  - ATLAS and FRIB have a lot in common, any development for ATLAS will be useful for FRIB and vice versa

- ANL / AWA: J. Power, P. Piot and I. Sugrue (PhD student started in Jan'21)
  - AWA can serve as test bed for AI tools development and testing. Being a test facility, more beam time is available for testing tools useful for ATLAS

- ANL / DSL & ALCF: A. Ramanathan and V. Vishwanath
  - Consult & advise on ML/AI modeling, HP computing and data storage at ALCF

PROGRESS & HIGHLIGHTS - ATLAS

PROGRESS & HIGHLIGHTS AT ATLAS

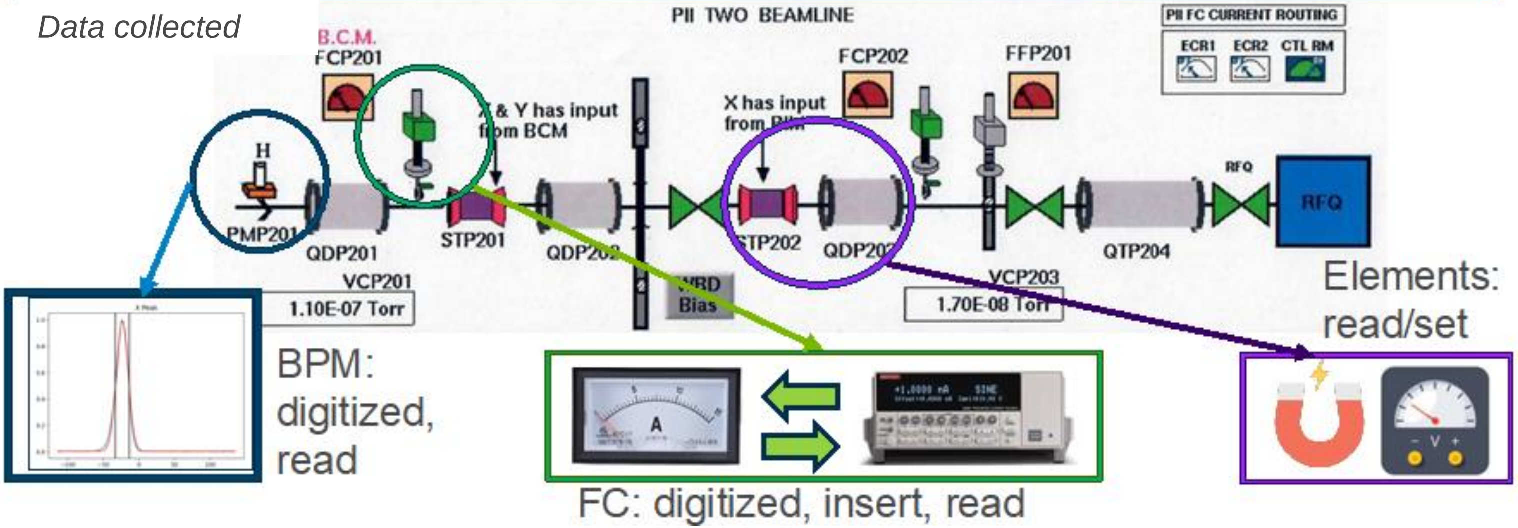
- ❑ Automated data collection established.
- ❑ Bayesian Optimization (BO) used for online beam tuning.
- ❑ AI-ML supporting the commissioning of a new beamline (AMIS)
- ❑ Transfer learning from one beam to another. (BO)
- ❑ Transfer learning from simulation to online model (BO with Deep Kernel Learning.
- ❑ Reinforcement Learning for online beam tuning – promising results

AUTOMATED DATA COLLECTION ESTABLISHED

- ✓ Beam currents and beam profiles digitized
- ✓ A python interface developed to collect the data automatically

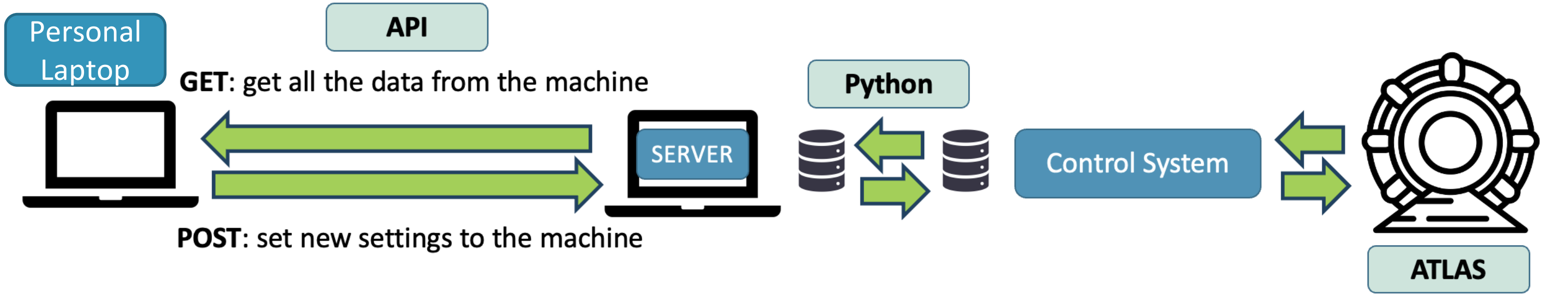
Schematic of data collection interface

Data collected



Now working on reducing acquisition time …

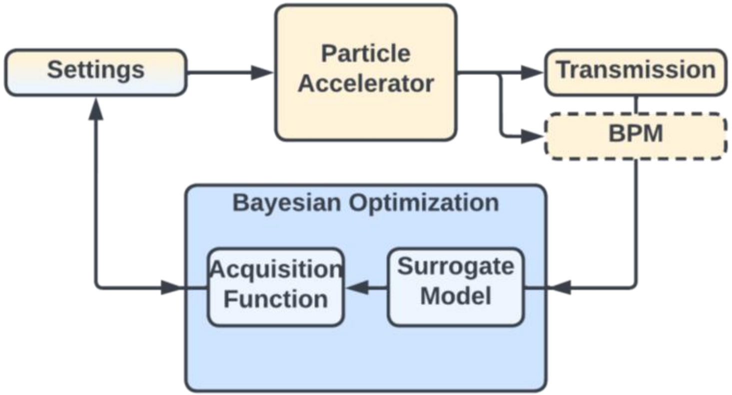
ONLINE – COMMUNICATION WITH CONTROL SYSTEM



OFFLINE – INTERFACE WITH TRACK SIMULATION

- ✓ Python wrapper for TRACK (Simulation Code)
- ✓ Generation of simulation data
- ✓ Different conditions and inputs
- ✓ Integration with AI/ML modeling

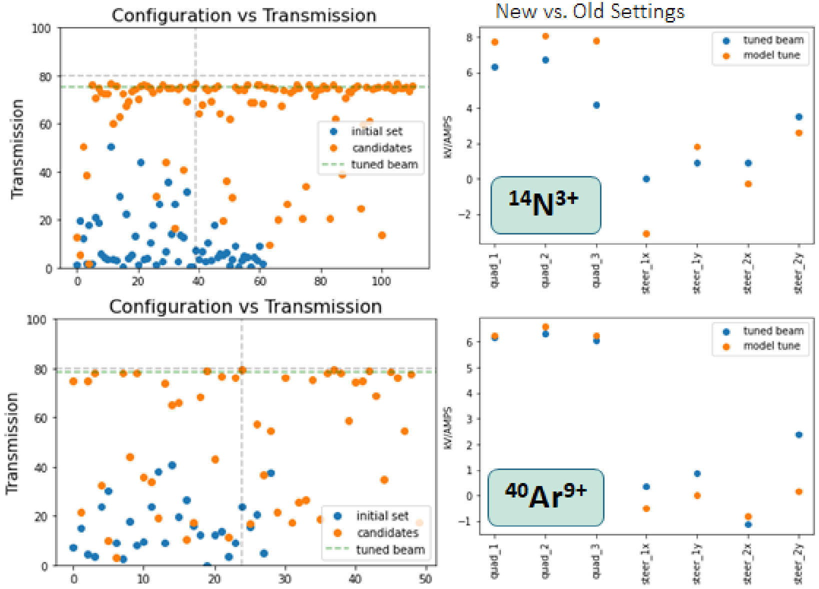
BAYESIAN OPTIMIZATION USED FOR BEAM TUNING



○ **<u>Surrogate Model</u>: A probabilistic model approximating the objective function** [Gaussian Process with RBF Kernel and Gaussian likelihood]

○ **<u>Acquisition Function</u> tells the model where to query the system next for more likely improvement** [EI]

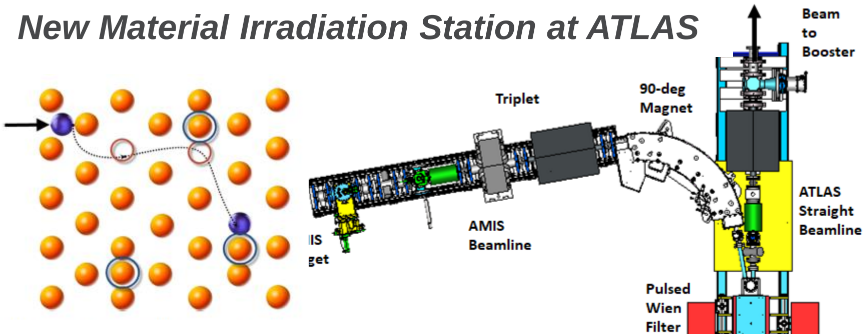
○ **Bayesian Optimization with Gaussian Processes gives a reliable estimate of uncertainty and guides the model**



○ **7 varied parameters (3 quads + 2 steerers)**
○ **Optimization of beam transmission**
○ **Case of $^{14}N^{3+}$: 29 historical + 33 random tunes**
○ **Case of $^{40}Ar^{9+}$: 29 historical tunes**

AI/ML SUPPORTING AMIS LINE COMMISSIONING

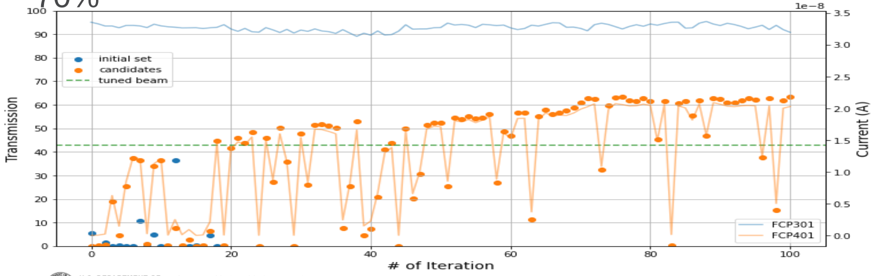
New Material Irradiation Station at ATLAS



Low-energy heavy-ion beams ~ 1 MeV/u can effectively emulate material damage in nuclear reactors, in both fuel and structural materials.

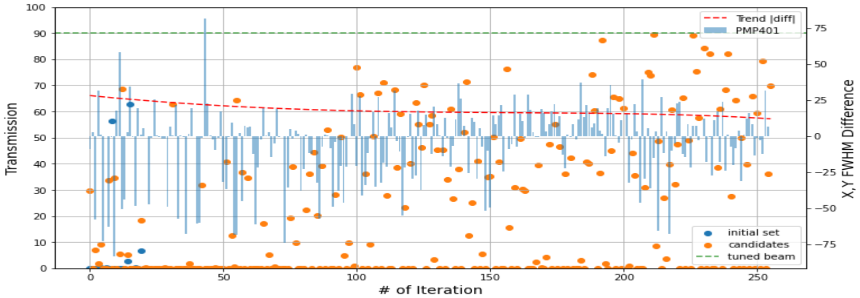
Improving Beam Transmission

Problem: Maximize beam transmission by varying a triplet, two dipoles and two steerers [BO]; **Results**: 40 → 70%

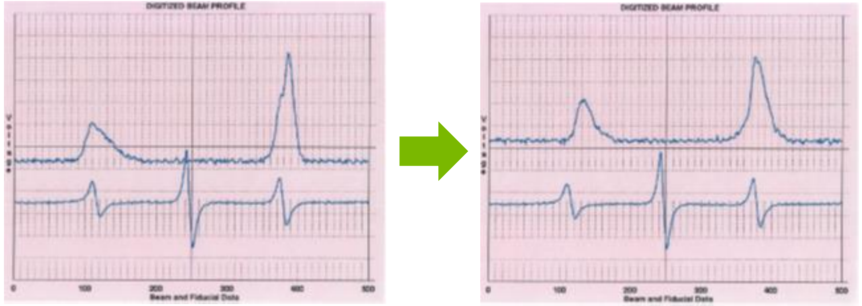


Improving Beam Profiles

Problem: Produce symmetric beam profiles by varying a triplet and a steerer [BO]



Training online, slow convergence but steady progress. Competition between nice profiles and beam transmission!

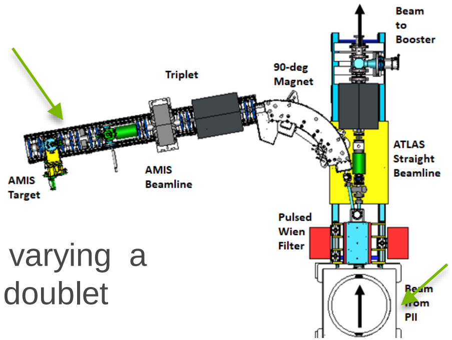


Very encouraging first results!

MULTI-OBJECTIVE BAYESIAN OPTIMIZATION

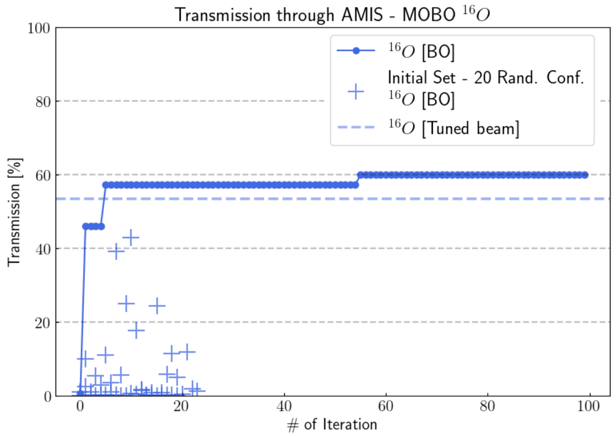
Multi-Objective Problem: Optimize transmission and beam profiles on target - Not easy for an operator!

Improving Beam Transmission & Improving Beam Profiles

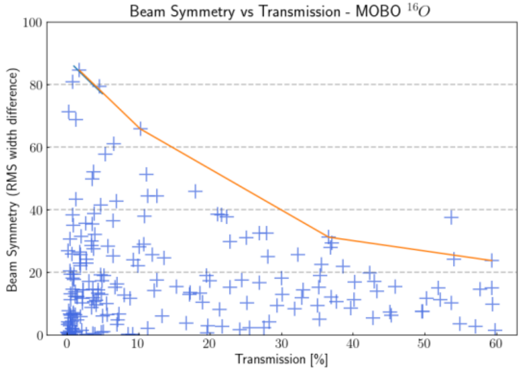


AMIS line: varying a triplet and a doublet

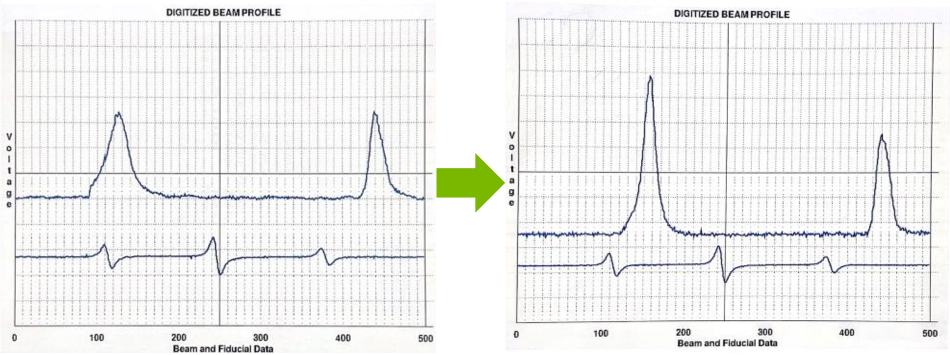
MOBO Results: 53→ 60% Beam transmis.



MOBO Results: Pareto Front



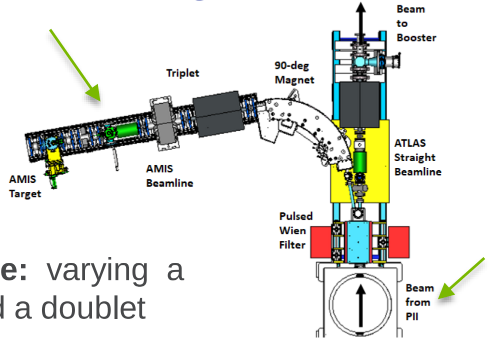
MOBO Results: More symmetric beam profiles



TRANSFER LEARNING FROM 16O TO 22NE - BO

Goal: Train a model using one beam then use it to tune another beam → Faster switching and tuning

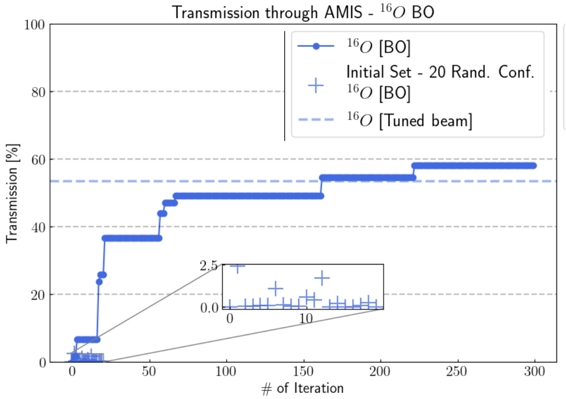
Training model on 16O



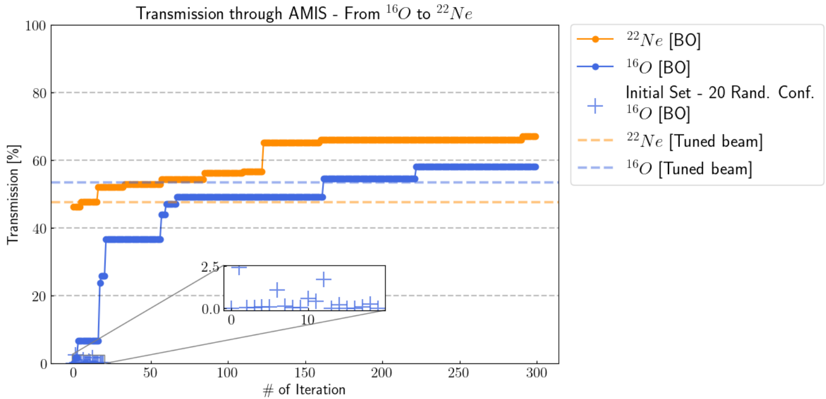
AMIS line: varying a triplet and a doublet

BO Training:
Over 300 iterations
53 → ~ 60%
Beam transmis.

Model saved & exported



Applying same model to 22Ne



16O Model loaded for 22Ne: Initial transmission improved in 7 iterations: 48 → 55 %

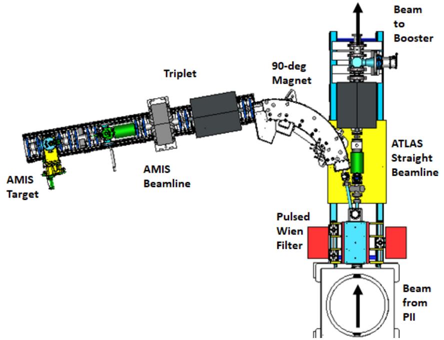
With more training for 22Ne: 48 → 67%

Scaling was applied from 16O to 22Ne, but re-tuning is always needed because of different initial beam distributions

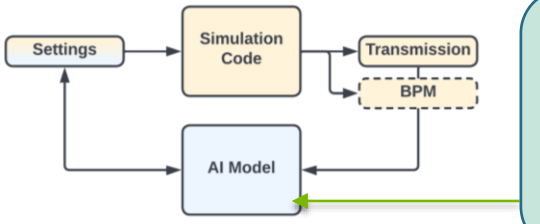
TRANSFER LEARNING FROM SIMULATION TO ONLINE

Goal: Train a model using simulations then use it for online tuning → Less training & fast convergence online

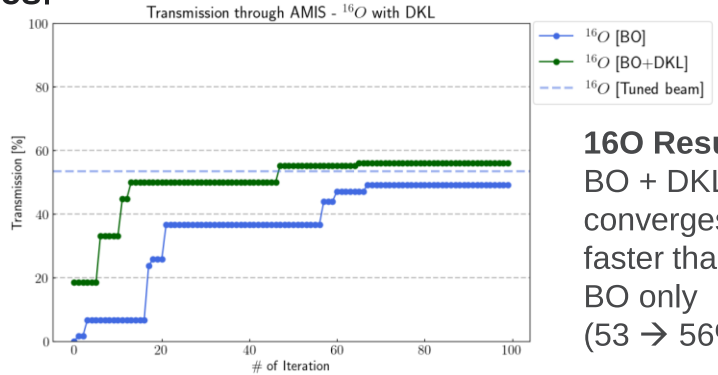
Method: Deep kernel learning (DKL) to **combine the representational power of neural networks with the reliable uncertainty estimates of Gaussian processes**.



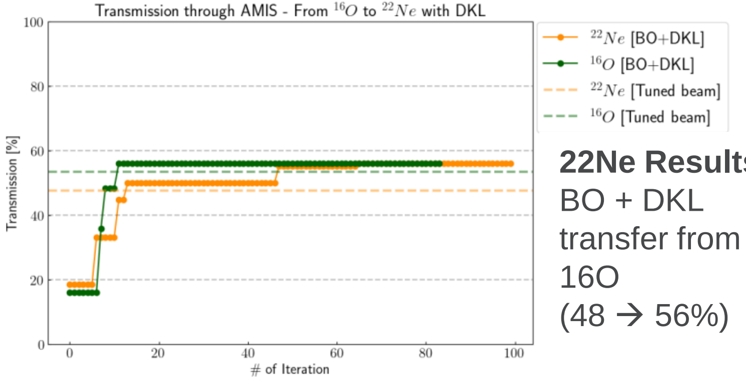
AMIS Line: Maximize beam transmission by varying a triplet [BO+DKL]



NN trained offline with TRACK simulations [4k training set /1k evaluation set]

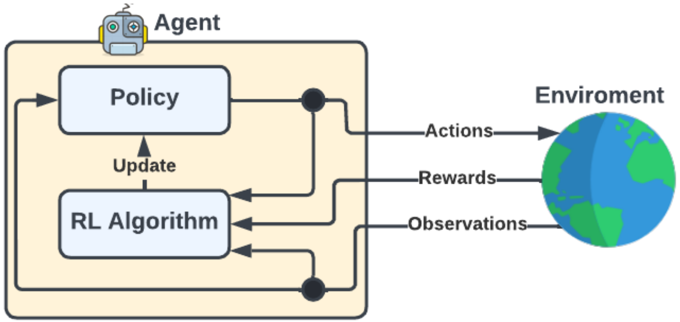


16O Results: BO + DKL converges faster than BO only (53 → 56%)

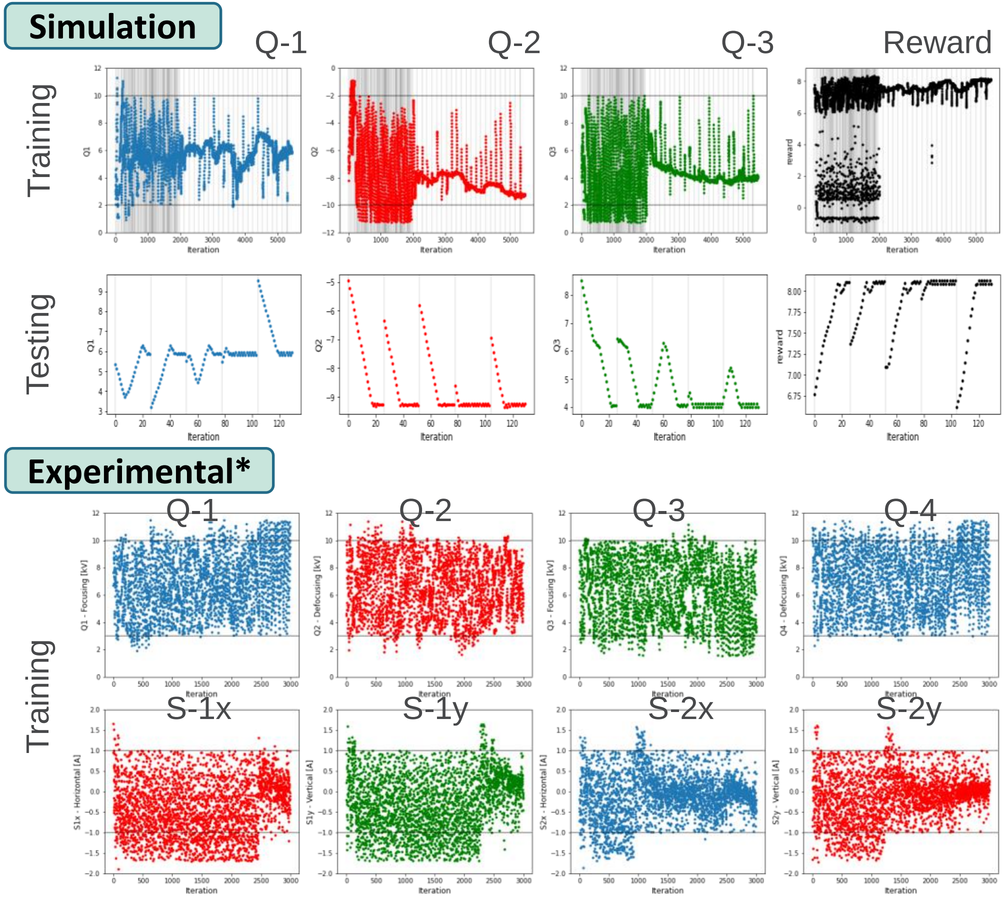


22Ne Results: BO + DKL transfer from 16O (48 → 56%)

REINFORCEMENT LEARNING FOR FINE TUNING



- ✓ **Method**: Deep Deterministic Policy Gradient (DDPG); Actor-Critic Approach
- ✓ **Simulation Case**: Focusing beam on target using a triplet (3 Quadrupoles)
- ✓ **Experimental Case**: Maximizing beam transmission using 4 quads and 2 steerers
- ✓ Electrostatic Quadrupoles :
  - 2 kV to 10 kV
  - Max action +/- 0.25 kV
- ✓ Steering Magnets:
  - -1 A to 1 A
  - Max action +/- 0.25 A



PROGRESS HIGHLIGHTS - AWA

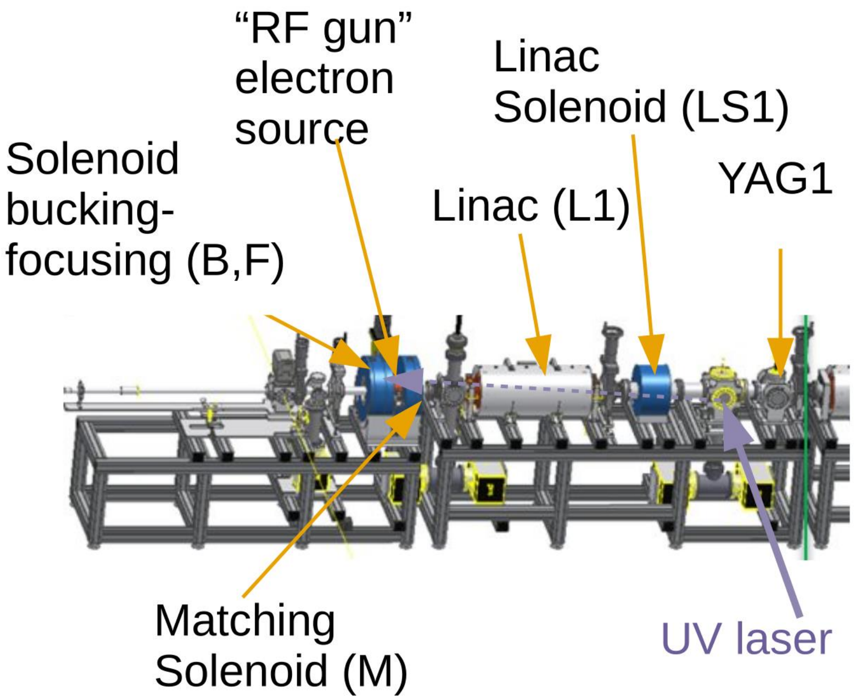
PROGRESS SUMMARY - AWA WORK

Idea: AWA can be used as a testbed for ML-based machine tuning and virtual diagnostics development

Progress made so far

- ❑ Improved surrogate model for beam image prediction: Improved simulation data and PCA decomposition
- ❑ Least squares minimalization applied to retrieve the actual beamline elements settings for a given beam image
- ❑ Method tested first on simulation data with known settings and added noise to image – controlled or supervised fitting
- ❑ When tested on experimental data, some parameters are predicted very well but not the rest – work in progress

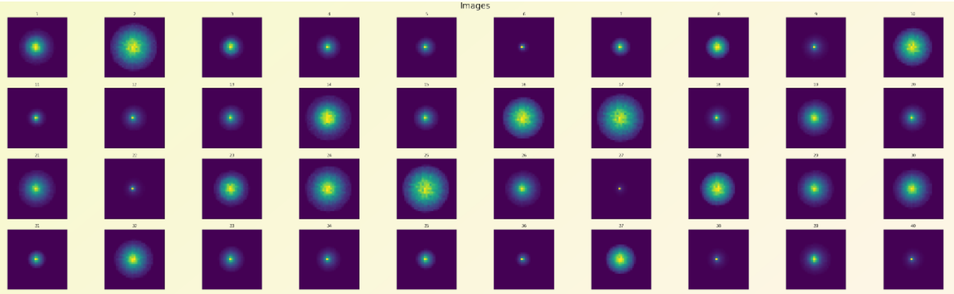
Lattice & beamline parameters

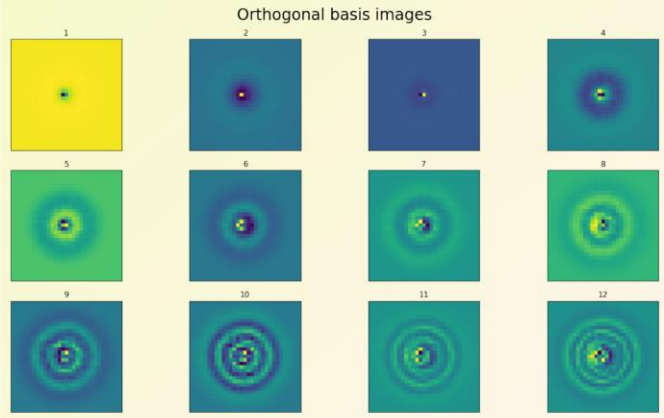


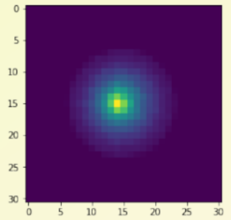
IMPROVED SURROGATE MODEL FOR BEAM IMAGE PREDICTION

Goal: Associate a given image to given input lattice parameters

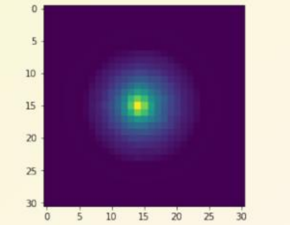
Improved simulation data & PCA decomposition





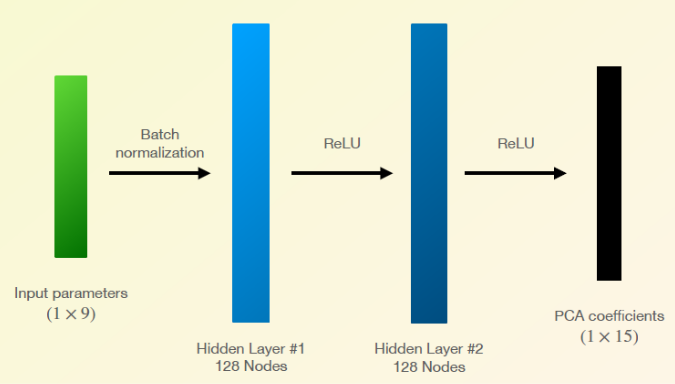


- Before PCA



- After PCA (15 Components)

Surrogate model: NN architecture



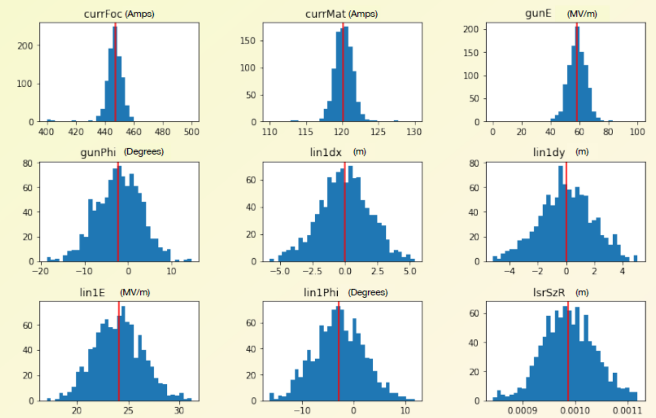
- ✓ 9 Input lattice parameters
- ✓ Images reduced to 15 PCA components
- ✓ Two hidden layers of 128 nodes each
- ✓ ~ 500 epochs, default batch size (32), MSE loss function

LEAST SQUARES MINIMIZATION: TEST ON SIMULATION DATA

Problem: What are the real lattice parameters for given beam image?

- **Method**: Minimize $\|f(x) - y\|_2^2$, where f(x) is the surrogate model output with input parameters x, and y is the PCA coefficients of the image.
- The initial input is the vector of experimental parameters x_0, and the result of the least squares optimization is an approximate solution of the true parameters.
- Test the optimization by pretending we have experimental data (noisy input) and that we know the true parameters (true input).
- Let v be a vector of random noise in R^9, and let x_t be the true input parameters. Minimize $\|f(x_t) - f(x)\|^2$ when $x0 = x_t(1+ v)$.
- We run this 1000 times, each with a different noise vector. The results are shown next …

Results of Least squares Minimization



Blue is the least squares estimate.

Red is the true parameter value

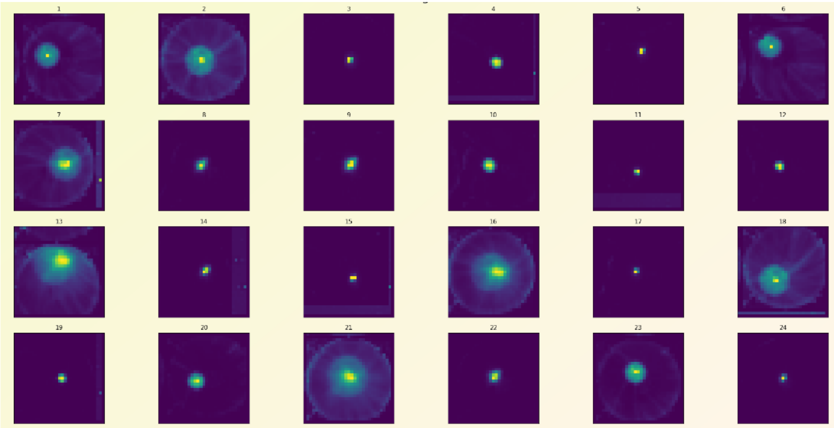
Results seems to be much closer to reality for the first 3 parameters than for the rest of them!

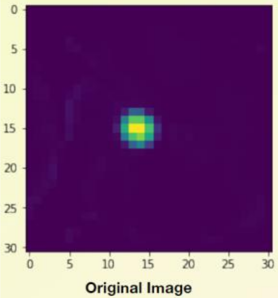
There seem to be large uncertainties on the misalignment parameters of the first linac cavity; lin1dx and 1lin1dy

LEAST SQUARE MINIMIZATION: TESTS ON EXPERIMENTAL DATA

Problem: What are the real lattice parameters for given beam image?

Experimental beam images & Related PCA





100 experimental images of size 31×31, flattened to a single array of size 100×961

Simulated Data = X

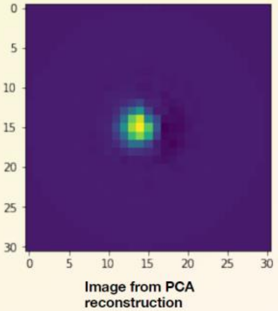
$\hat{X} = X - \mu$ (mean of each row of X)

$\hat{X} = U \Sigma V^{\top}$

Array of coefficients is $C = \hat{X} V_k$ (first k rows of V)

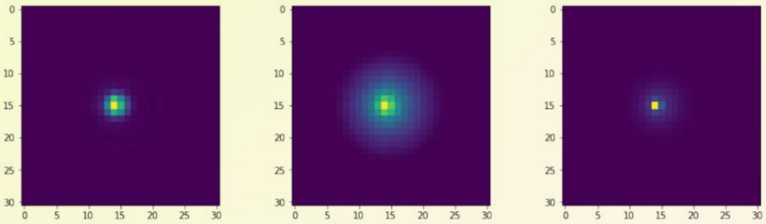
Experimental Data = Y

PCA coefficients using existing basis images is $C_{\text{exp}} = Y V_k$



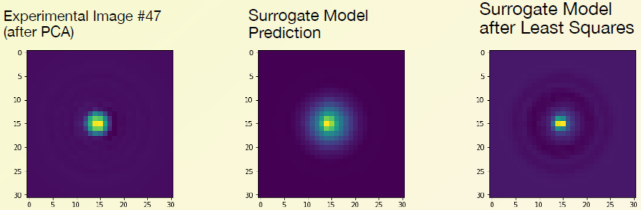
$V_k^{\top} C_{\text{exp}}^{\text{l}} + \mu$

Experimental Test #1



- Experimental Image #7 (after PCA)
- Surrogate Model Prediction
- Surrogate Model after Least Squares

Experimental Test #2



	Experimental Params	Least Squares Result	Difference
currFoc	4.981189243296668110e+02	4.981189243296668110e+02	0.000000000000000000e+00
currMat	1.416873704809135006e+02	1.416873704809135006e+02	0.000000000000000000e+00
gunPhi	6.000000000000000000e+01	6.000000000000000000e+01	0.000000000000000000e+00
gunE	0.000000000000000000e+00	0.000000000000000000e+00	0.000000000000000000e+00
Lin1dx	0.000000000000000000e+00	4.705160544383231809e+02	4.705160544383231809e+02
Lin1dy	0.000000000000000000e+00	-2.253368202526503410e+02	-2.253368202526503410e+02
Lin1Phi	2.200000000000000000e+01	2.200000000000000000e+01	0.000000000000000000e+00
Lin1E	0.000000000000000000e+00	-1.879956681709598598e+01	-1.879956681709598598e+01
lsrSzR	1.000000000000000021e-03	1.000000000000000021e-03	0.000000000000000000e+00

PROGRESS HIGHLIGHTS - FRIB

PROGRESS ON FRIB WORK

Idea: Transfer learning between ATLAS and FRIB based on similarities

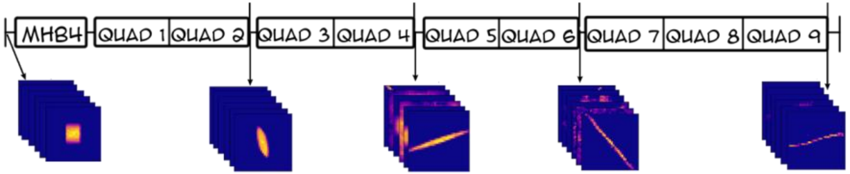
Progress made so far

- ❑ Improved surrogate model to predict beam loss given an initial beam distribution, original implementation using full beam distribution (6 projections)

- ❑ Model using less data – 3 beam projections instead of 6 – produced similar results as the first model → possibility of implementation with limited beam diagnostics

- ❑ Future: 4D beam tomography using measured beam profiles at different locations

Lattice & beamline parameters



- ✓ Current modeling is based on TRACK simulations of the ATLAS Low-energy beam transport (LEBT) line – 9 electrostatic quadrupoles

- ✓ Plan to transfer model to the FRIB front-end

- ✓ Apply experimentally at ATLAS or FRIB

IMPROVED SURROGATE MODEL FOR BEAM LOSS PREDICTION

Goal: Given an initial beam distribution, predict beam loss

Model Layout – Input & Output

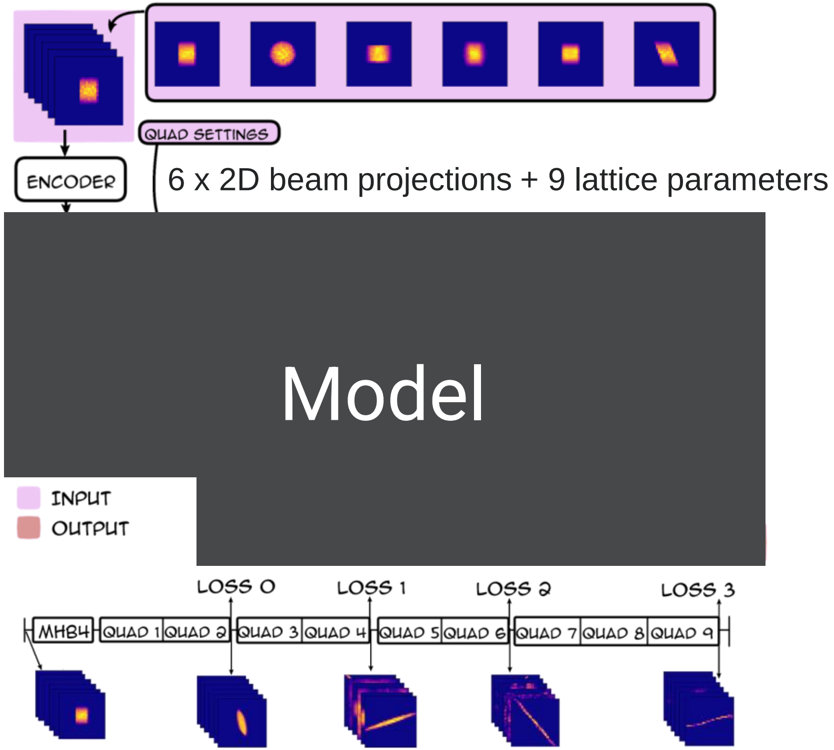
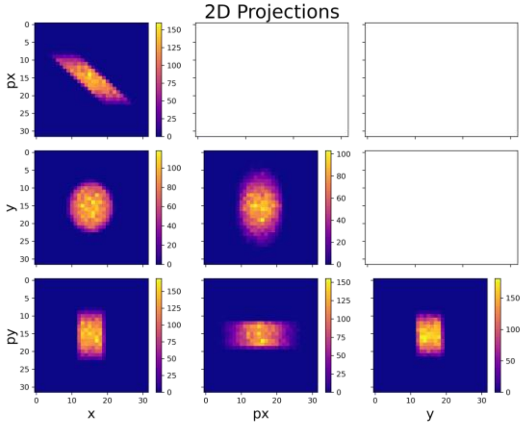
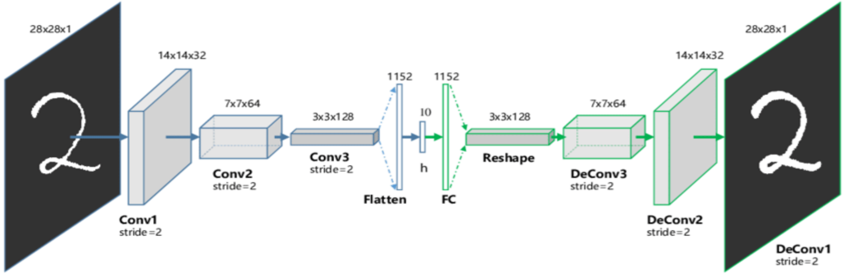


Image Reduction: Convolutional Autoencoder

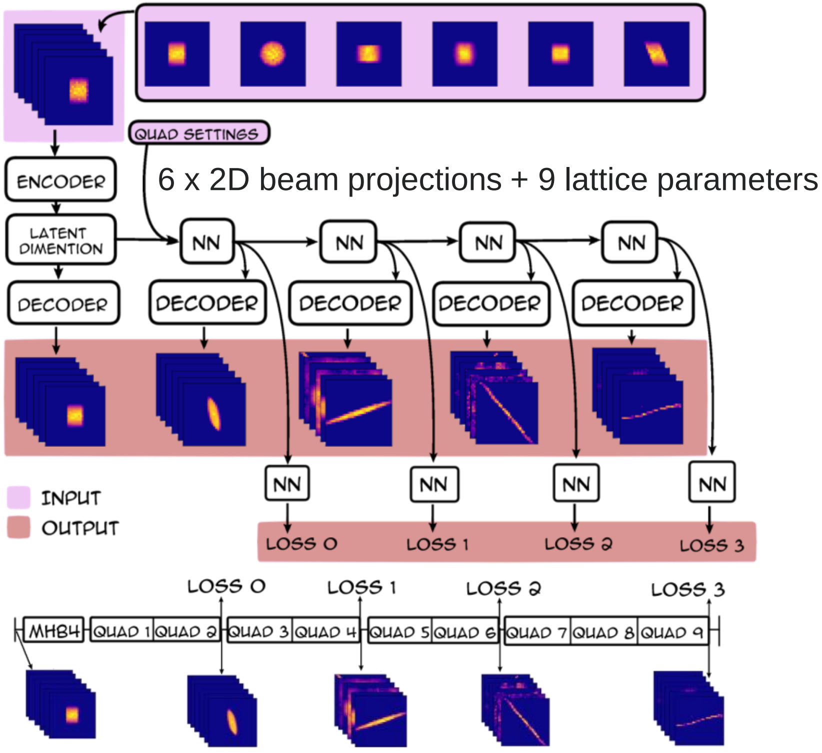




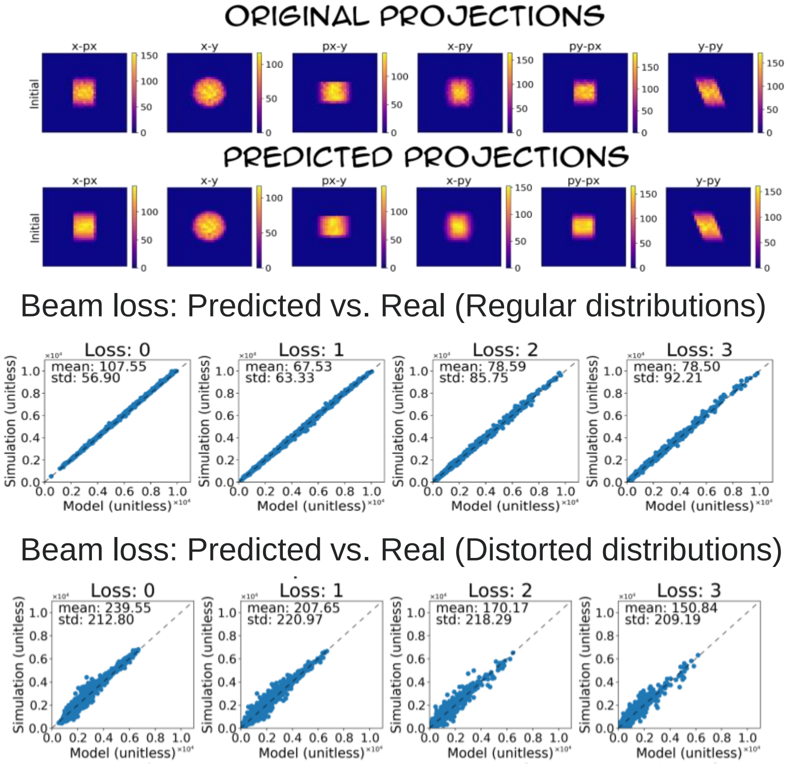
SURROGATE MODEL PREDICTION – 6 BEAM PROJECTIONS

Goal: Given an initial beam distribution, predict beam loss

Actual Model Workflow



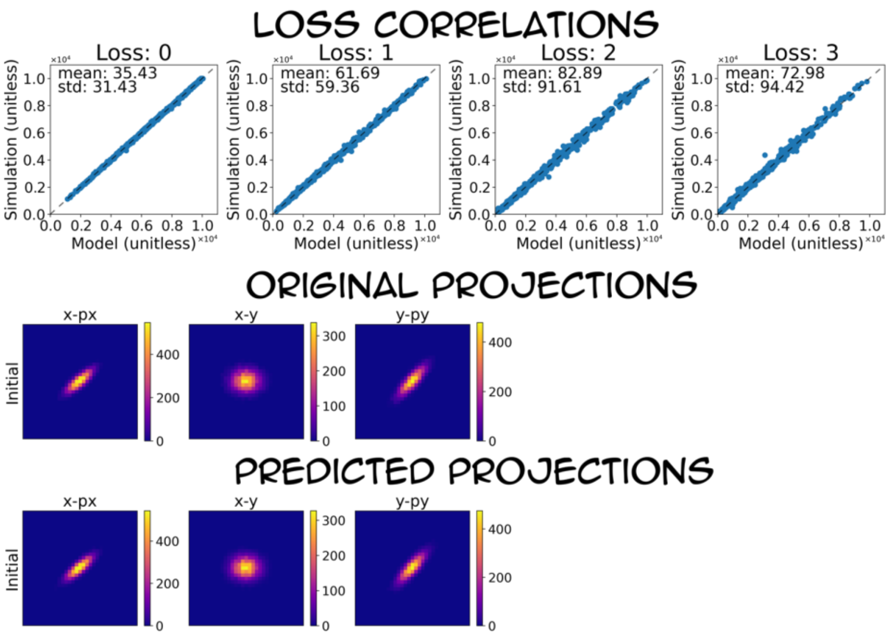
Results – Using 6 Projections



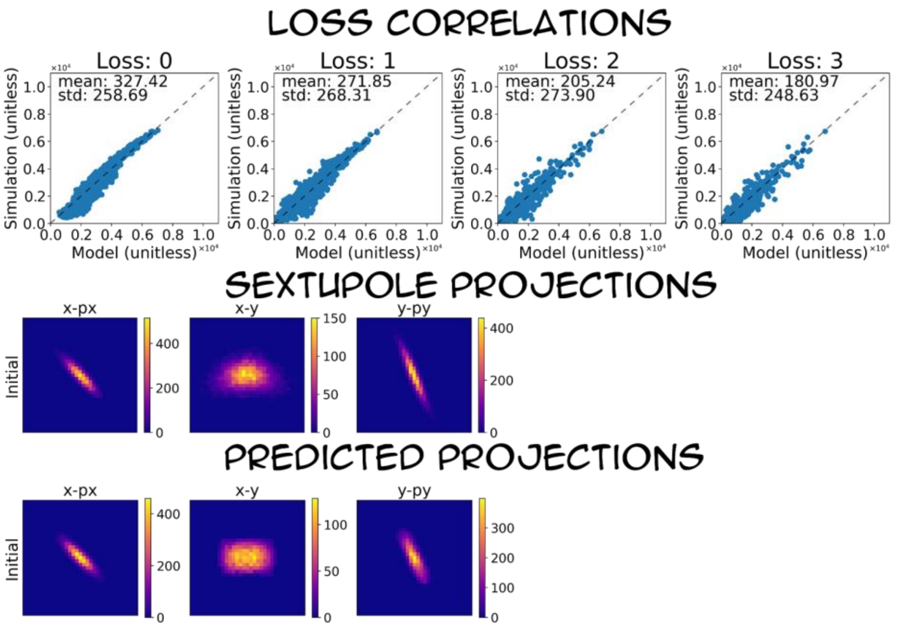
SURROGATE MODEL PREDICTION – 3 BEAM PROJECTIONS

Goal: Given an initial beam distribution, predict beam loss

Regular Beam Distributions



Distorted Beam Distributions



The model produced similar results despite the information reduction from 6 to 3 beam projections

SUMMARY & FUTURE PLANS

CHALLENGES & FUTURE OPPORTUNITIES

- Challenges
  - Need to limit the number of random or non-physical settings to avoid unintentional damage to beamline components, power supplies, …
  - Need faster data acquisition and collection to speed-up the process
  - Not enough diagnostics and data to characterize the initial beam distribution from the source

- Future Plans:
  - The recent progress shows proof of concept for short linac sections with limited number of parameters, need scaling to other sections and more parameters
  - Implement sequential tuning from one section to the next, possibly back and forth
  - Develop virtual diagnostics tools based on a validated virtual machine model
  - Start preparing for the operation of the ATLAS multi-user upgrade with two different beams accelerated in the linac and delivered to two different experimental areas

RECENT TALKS AND PUBLICATIONS

- “Reinforcement Learning and Bayesian Optimization for Ion Linac Operations”, J. Martinez, B. Mustapha et al, Invited talk at the Heavy Ion Accelerator Technology (HIAT) Conference, Darmstadt, Germany, June 27 - July 1 2022

- “Machine Learning to support the ATLAS Linac Operations at Argonne”, B. Mustapha et al, Poster & Paper at NAPAC’22, August 7-12th, 2022, Albuquerque, New Mexico & ICFA Workshop on Machine Learning for Accelerators, Nov. 1-4, Chicago, Illinois

- “Machine Learning Tools to support the ATLAS Ion Linac Operations at Argonne”, J. Martinez, B. Mustapha et al, Talk at the ICFA Workshop on Machine Learning for Accelerators, Nov. 1-4, Chicago, Illinois

- “Model-based Calibration of Control Parameters at the Argonne Wakefield Accelerator”, I. Sugrue et al, NAPAC’22, August 7-12th, 2022, Albuquerque, New Mexico

- “Predicting beam transmission using 2-dimensional phase space projections of hadron Accelerators”, A. Tran et al, Front. Phys. 10:955555. doi: 10.3389/fphy.2022.955555

MANY THANKS TO

- ATLAS Controls Team:

  D. Stanton, K. Bunnell and C. Dickerson

- ATLAS Operations Team:

  B. Blomberg, E. Letcher, G. Dunn and M. Hendriks

- ATLAS Liaison and beam time schedular:

  D. Santiago

- …

THANK YOU