OPTIMIZING BEAM DYNAMICS IN LHC WITH ACTIVE DEEP LEARNING

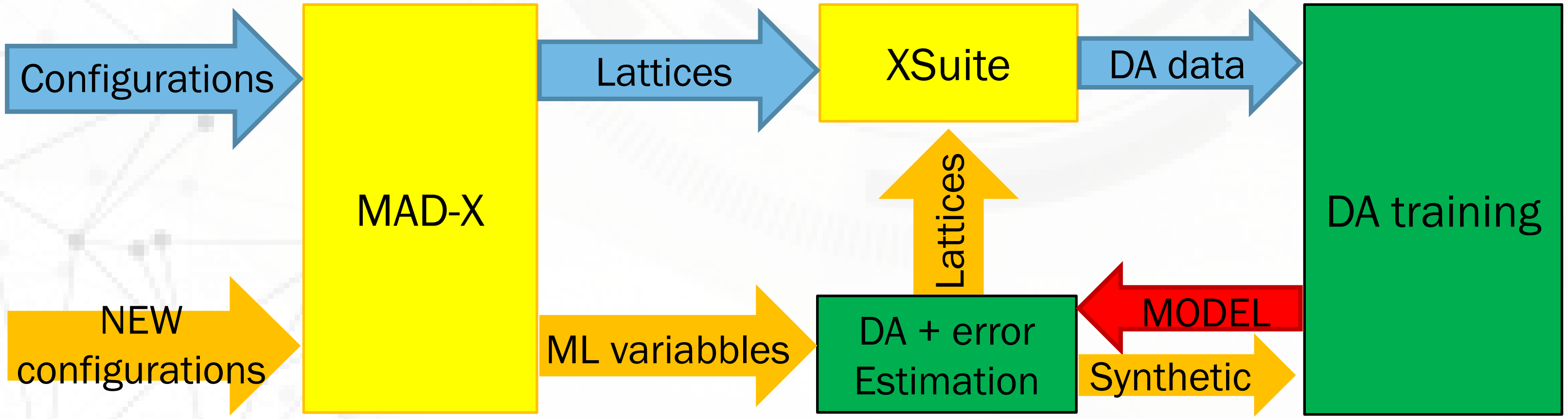
Davide Di Croce, Massimo Giovannozzi, Ekaterina Krymova, Tatiana Pieloni, Mike Seidel & Frederik F. Van der Veken

HB 2023 – FLASH PRESENTATIONS - OCT 12, 2023

This work is partially funded by the SDSC project C20-10

INTRODUCTION

- Dynamic Aperture (DA) is crucial for understanding non-linear beam dynamics in circular accelerators like the LHC, offering insights into beam stability and lifetime.
- Traditional DA calculation methods are computationally demanding, especially for large accelerators like the LHC.
- Our previous work has demonstrated that Deep Neural Networks (DNNs) can accurately predict the DA for new machine configurations (interpolation) while significantly accelerating computational processes.
- In this study we integrated the DNN model into an innovative Active Learning (AL) framework. For this purpose, we introduced an error estimator alongside the DA regressor, allowing uncertainty estimation.
- AL framework also enables smart sampling of simulations: by prioritising predictions with higher errors, it efficiently determines the sequence in which to simulate new machine configurations.



14th International Particle Accelerator Conference,Venice, Italy JACoW Publishing
ISBN: 978-3-95450-231-8 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2023-WEPA097

ACCELERATING DYNAMIC APERTURE EVALUATION USING DEEP NEURAL NETWORKS

D. Di Croce[2*], M. Giovannozzi[1], T. Pieloni[2], M. Seidel[2,3], F. F. Van der Veken[1]
[1]Beams Department - CERN, Geneva, Switzerland
[2]École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland
[3]Paul Scherrer Institut, Villigen, Switzerland

Abstract

The Dynamic Aperture is an important concept for the study of non-linear beam dynamics in a circular accelerator. The DA is defined as the extent of the phase-space region in which the particle's motion remains bounded over a given finite number of turns. Such a region is determined by the imperfections in the magnetic fields, beam-beam effects, electron lens, electron clouds, and other non-linear effects. The study of the DA provides insight into the mechanisms driving the beam lifetime, which is essential for the operation of existing circular accelerators, such as the CERN Large Hadron Collider, as well as for the design of future ones. The standard approach to numerical evaluation of the DA relies on the ability to accurately track initial conditions, distributed in phase space, on the required time scale, and this is computationally demanding. To accelerate the angular DA calculation, we propose the use of a Machine Learning technique for the angular DA regression based on simulated HL-LHC data. We demonstrate the implementation of a Deep Neural Network model by measuring the time and assessing the performance of the angular DA regressor, as well as carrying out studies with various hardware architectures including CPU, GPU, and TPU.

INTRODUCTION

The study of dynamic aperture (DA), defined as the extent of the connected phase-space region in which the single-particle dynamic is bounded, provides insight into the single-particle, non-linear beam dynamics and mechanisms driving the time evolution of beam losses [1], which is essential for the design and operation of existing [2, 3] and future circular accelerators [4].

The numerical calculation of the DA involves tracking a large number of initial conditions in phase space for many turns [5, 6]. This method is computationally demanding, especially for large accelerators such as the CERN Large Hadron Collider (LHC) [2], and for this analytical scaling laws have been studied for several years [6, 7]. In general, in the accelerator community, there is growing interest in developing methods to accelerate the DA calculation while maintaining its accuracy.

In recent years, Machine Learning (ML) techniques have emerged as a promising approach to accelerate DA evaluation (see, e.g. [8–11]). By training a model on a large data set of simulated initial conditions, an ML algorithm can learn the complex mapping between the initial conditions and the angular DA (defined below) and provide a fast and accurate prediction of the angular DA for new sets of initial conditions and machine configurations. This approach has the potential to reduce the computational cost of DA evaluation and enable faster accelerator parameter optimisation.

Here, we propose to use machine learning techniques to speed up angular DA evaluation based on simulated data obtained using the High Luminosity LHC (HL-LHC) lattice [3]. We investigated the use of a Deep Neural Network (DNN) model to regress the angular DA as a function of the initial conditions. We study the performance of this ML model on various hardware architectures and compare it with the standard simulation method.

SIMULATED SAMPLES

To train the ML model, we simulated several accelerator configurations using MAD-X [12] and the V1.0 HL-LHC lattice in the injection configuration at 450 GeV [13]. We varied six accelerator parameters, namely the betatron tunes Q_x, Q_y, chromaticities Q'_x, Q'_y, strength of the Landau octupoles (using the current, I_{MO}, powering the octupoles) and the realisations (sometimes also called seeds) of the magnetic field errors assigned to the various magnet families. Furthermore, both Beam 1 and Beam 2 have been considered in these studies. For this first study, we limited the parameters sampling to two Q_x, Q_y scans (8 Q_x values in $[0.255, 0.295]$ and 9 Q_y values in $[0.280, 0.325]$) and a Q', I_{MO} scan (15 Q' values in $[0, 15]$ and 17 I_{MO} values in $[-40, 40]$A) for Beam 1 and Beam 2 and 60 possible realisations of the magnetic errors. This resulted in a total of 29880 sets of accelerator parameters.

The phase space was probed by tracking with SixTrack [14] for 10^5 turns a set of initial conditions selected along 11 polar angles, evenly distributed in $]0, \pi/2[$ and 290 radial amplitudes, evenly distributed in $[0.0\sigma, 20\sigma]$. An example of the results of these computations in the $x-y$ space is shown in Fig. 1 for a specific accelerator configuration, in which the stability time, i.e. the time taken by the orbit to reach an amplitude corresponding to a numerical overflow, is provided for each initial condition.

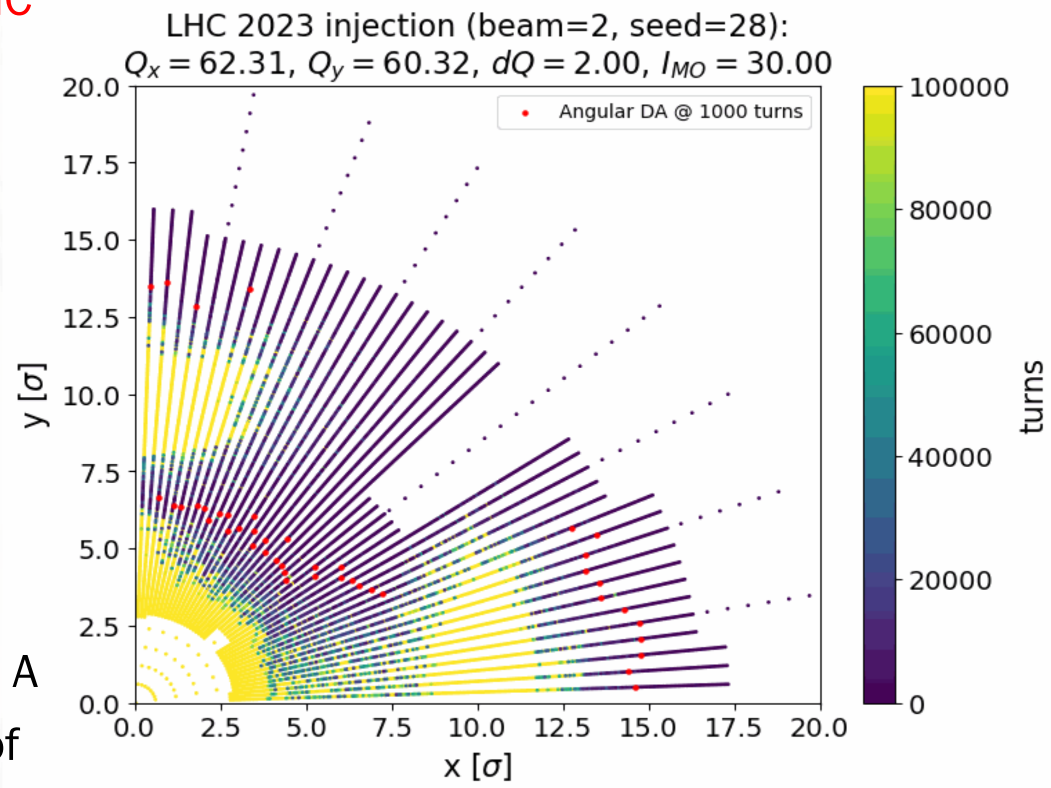
The input for the surrogate model is given by the parameters describing the accelerator configuration and the polar angle, the regressor will learn for each accelerator configuration the value of the last stable amplitude for that angle, which we call angular DA. When considering the angle as an additional parameter, the number of samples is increased to

* davide.dicroce@epfl.ch

10.18429/JACoW-IPAC2023-WEPA097

DATASET

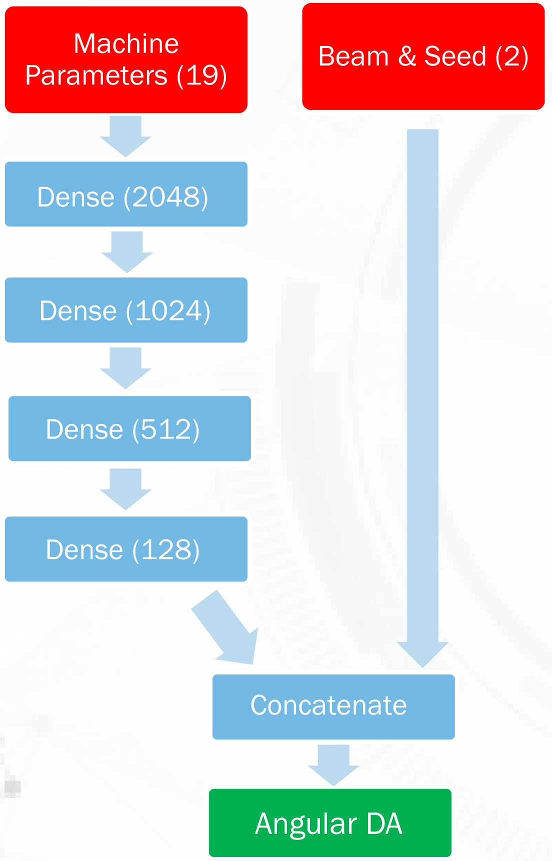
- The dataset is based on simulation (MADX) and tracking (xsuite) on LHC 2023 injection optics.
- Tracked the particles distributed in polar coordinates (44 angles and 0.06 σ radial steps) for every machine configuration.
- Goal is to regress the evolution of the stable region (angular DA) in 12 different number of turns (up to 10^5 turns) [Red points in the image].
- 10k sets of accelerator parameters generated using:
  - Normal random sampling 60 different seeds (magnet error realizations) for the 2 beams
  - Chromaticity (dQ) in interval of [0,30] in steps of 2 DQ
  - Octupole magnet current (I_MO) in interval of [-40,40] in steps of 5 A
  - Tune scan: Q_x [62.100,62.500] and Q_y [60.100,60.500] in steps of 0.05

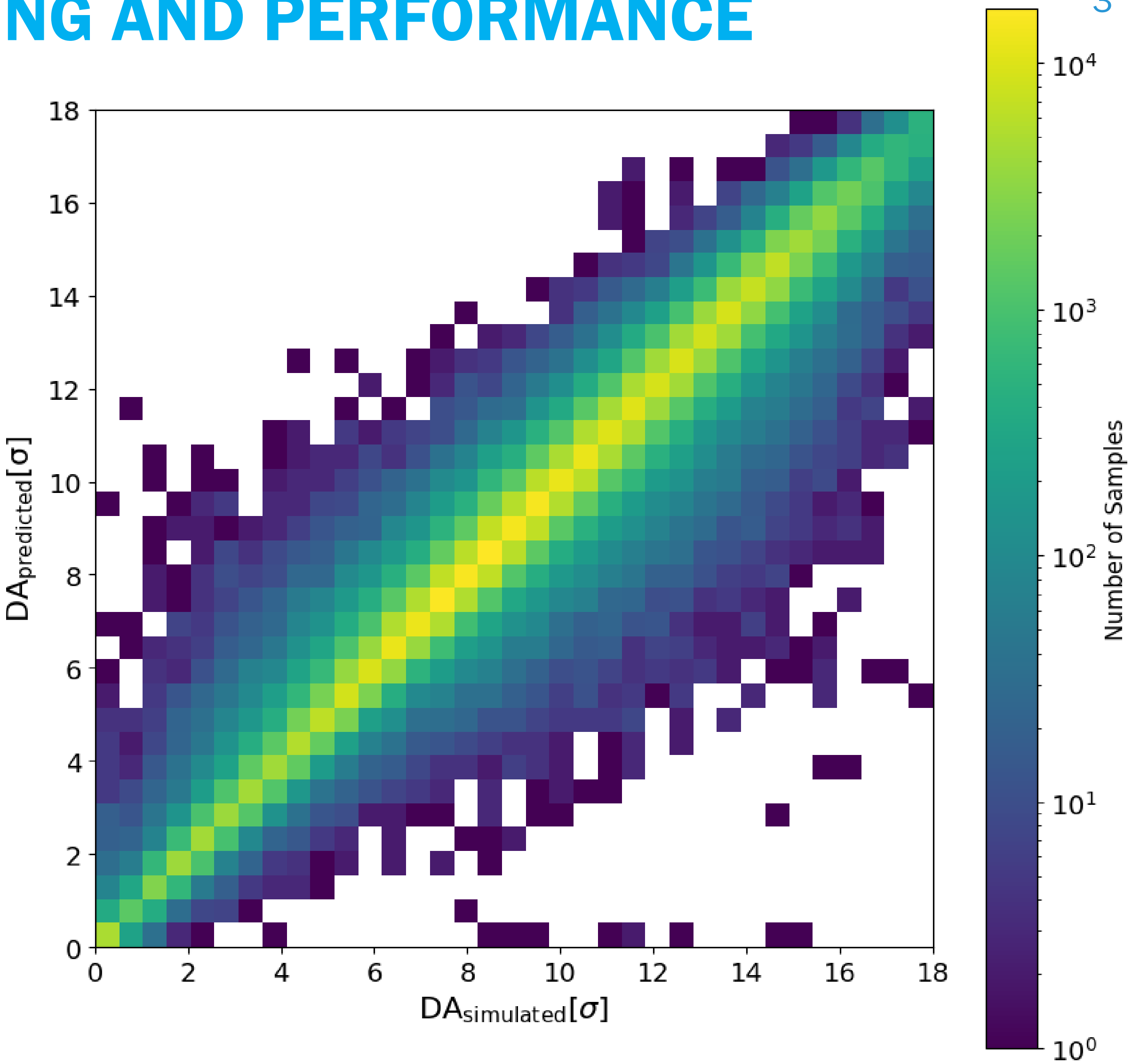


- Additional machine variables added into the dataset (total of 19 machine variables): 7 anharmonicities up to second order (PTC), maximum values of α and β and phase-advance μ (x,y) at IP5.

DNN ARCHITECTURE, TRAINING AND PERFORMANCE

- Considering fully connected DNN for machine parameters with concatenate layer (bias) to gather Beam and Seed labels.



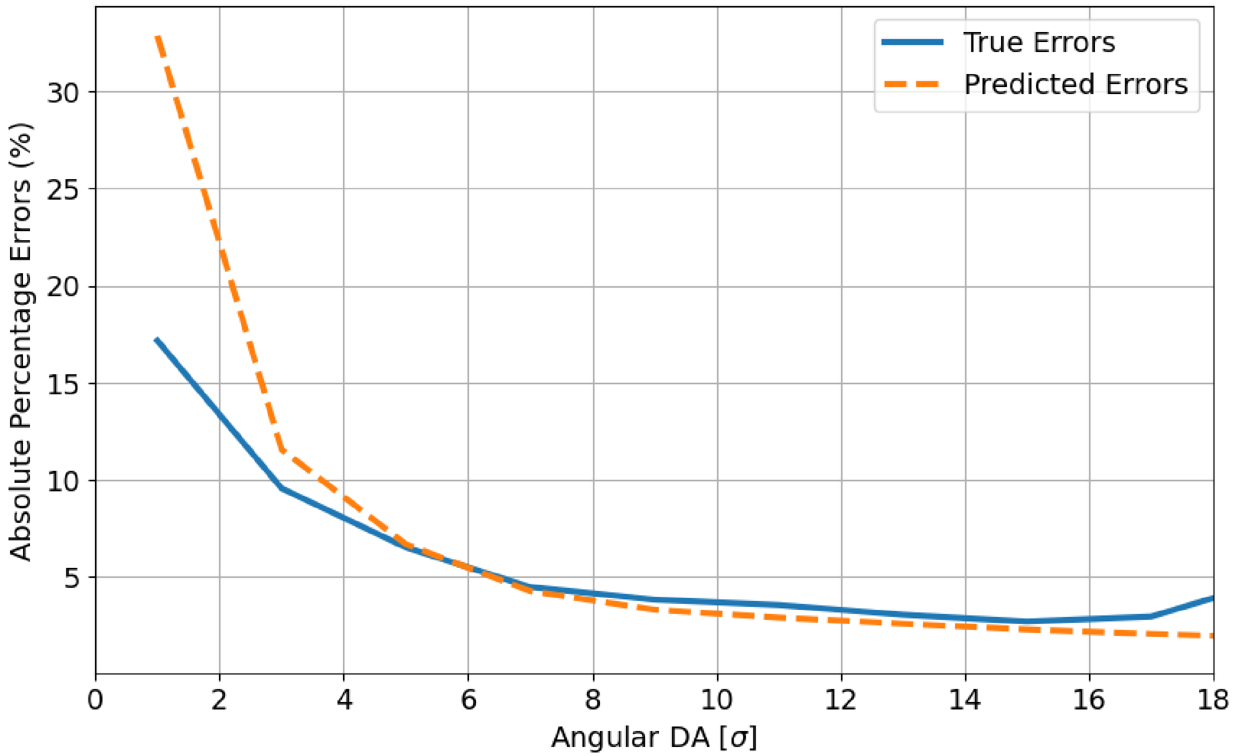


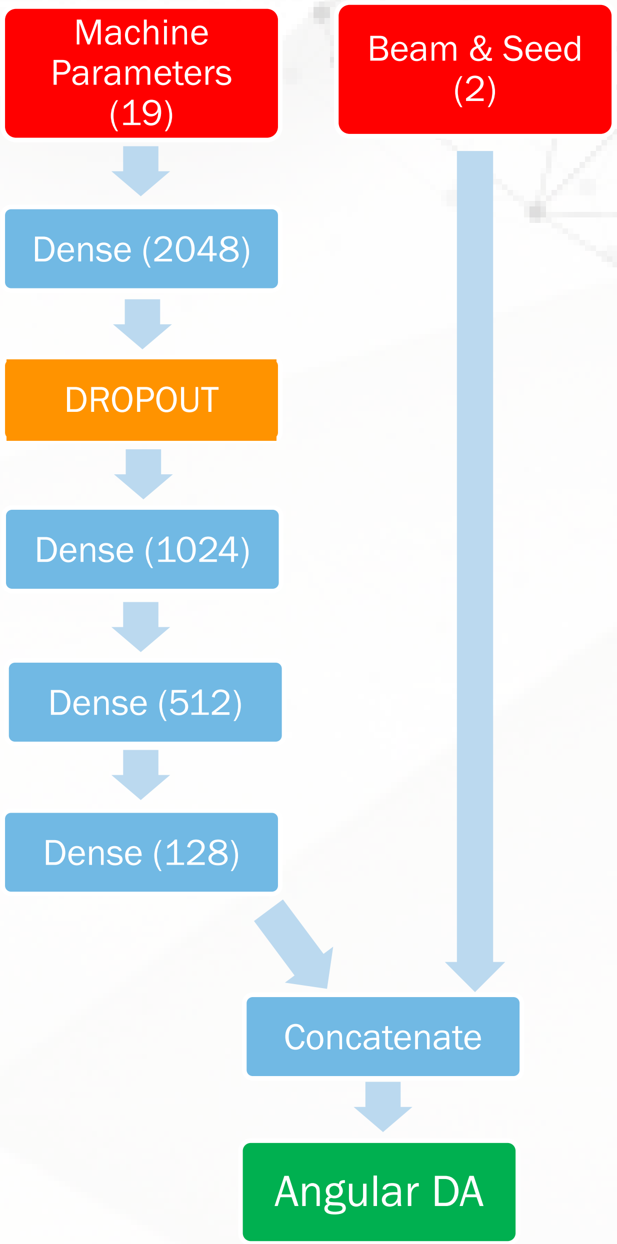
- Mean Absolute Error (MAE) used as Loss function.
- Inference of a single machine (12 different turns x 44 angles) in 0.5 ms (~1 μs/angular DA prediction)

- Test MAE = 0.201 beam σ and MAPE = 11.91 %.
- Improved performance due to the increase of variables (previous model MAE= 0.64 beam σ)

ERROR ESTIMATION: MONTE CARLO DROPOUT

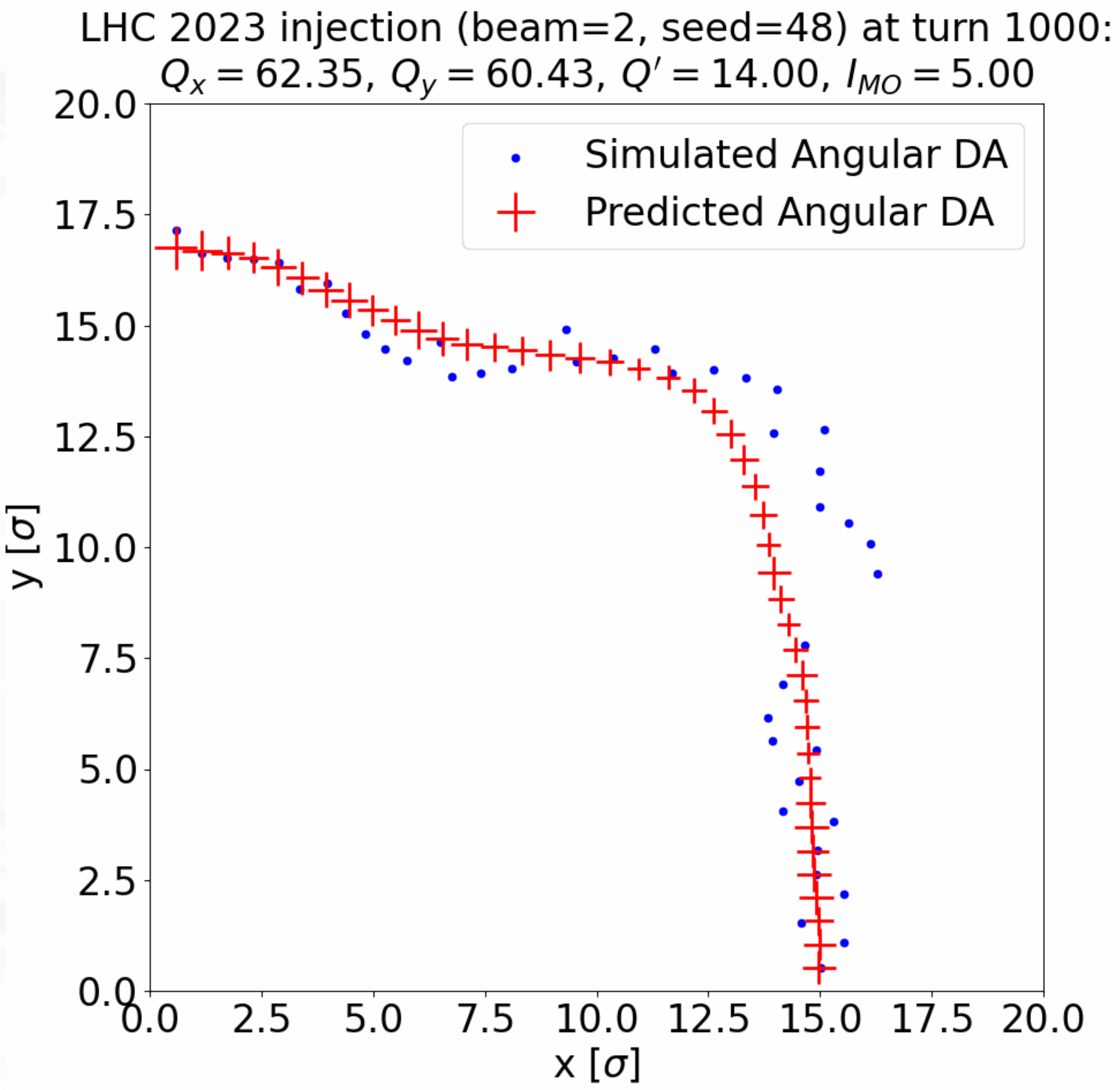
- Usually, dropout is a regularization technique to avoid overfitting during training (which randomly sets a fraction of nodes to zero).
- By leveraging dropout at inference time, we introduce diversity among the predictions (different angular DAs every time). This technique is known as Monte Carlo (MC) dropout.
- The variation in these predictions are utilized to estimate uncertainty: dropout at 1% between the first hidden layers and 1 std of 128 variations as error.
- DA and error prediction (129 inferences) in 0.75 s/machine configuration.





ACTIVE LEARNING FRAMEWORK

- Tracking on Xsuite takes 107s/machine configurations (using HT-Condor), while the AL framework, once trained, is approximately 140 times faster!
- AL demonstrated as a powerful tool for accelerating beam dynamics studies while maintaining precision.



For more details,
let's meet at the poster session!

THANK YOU!!