Optimizing Beam Dynamics in LHC with Active Deep Learning

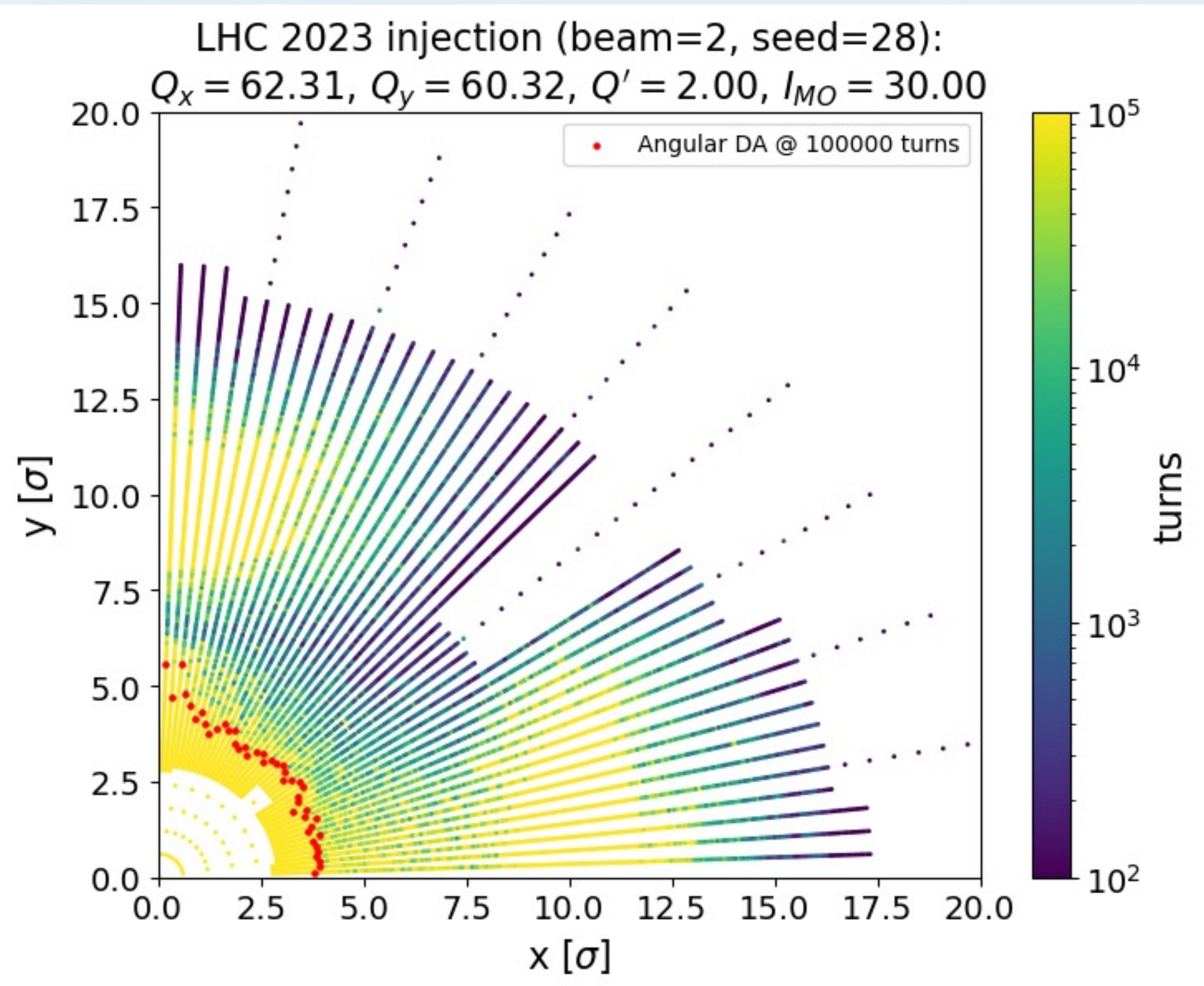
D. Di Croce[2], M. Giovannozzi[1], E. Krymova[3], T. Pieloni[2], M. Seidel[2,4], F. F. Van der Veken[1]
[1]CERN, [2]EPFL, [3]SDSC, [4]PSI

DYNAMIC APERTURE

Dynamic Aperture (DA) is essential for studying nonlinear beam dynamics in circular accelerators, providing insights into beam stability and lifetime. Traditional DA calculations are **computationally intensive**, especially for large accelerators such as the LHC.
Our previous study shown that **Deep Neural Networks** (DNNs) can efficiently predict DA for new machine setups, greatly speeding up computations (10.18429/JACoW-IPAC2023-WEPA097).
In this study, we incorporated the DNN model into an **Active Learning** (AL) framework. To achieve this, we introduced an error estimator alongside the DA regressor, enabling **uncertainty estimation**.

SIMULATED SAMPLES

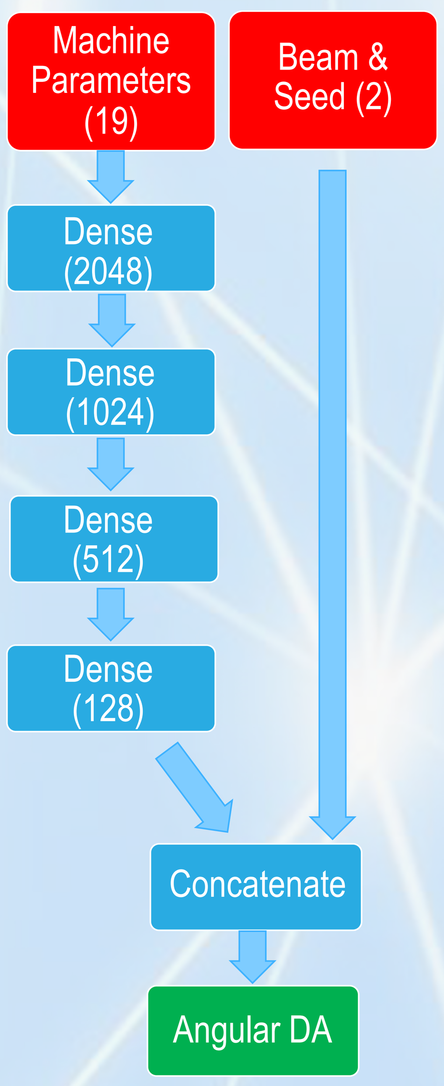
Sample generated using **MAD-X and 2023 LHC lattice at 450 GeV** by simulating various accelerator configurations.
6 machine parameters varied: **betatron tunes, chromaticities, Landau octupole strength, and magnetic field errors (seeds).**
Total of 10k sets of accelerator parameters tracked with **Xsuite.**
Phase space probed with 44 polar angles and 330 radial amplitudes.

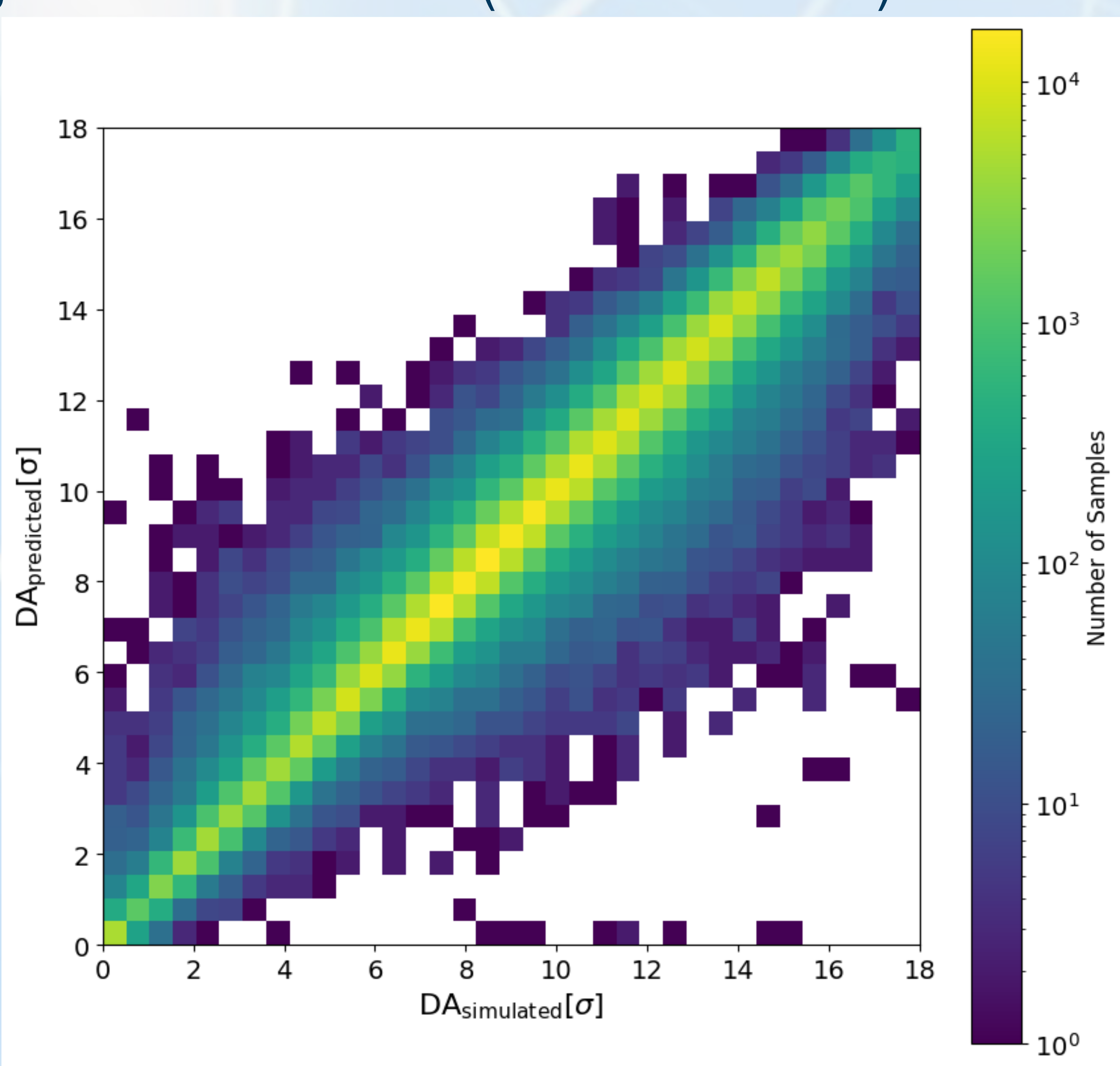


Goal is to regress the evolution of the stable region (**angular DA**) in 12 different number of turns (up to 10^5 turns).
Additional variables: anharmonicities up to second order (PTC), maximum values of α and β and phase-advance μ (x,y) at IP5

DNN ARCHITECTURE AND PERFORMANCE

Considering fully connected DNN for machine parameters with concatenate layer to gather discrete data (beam and seed).

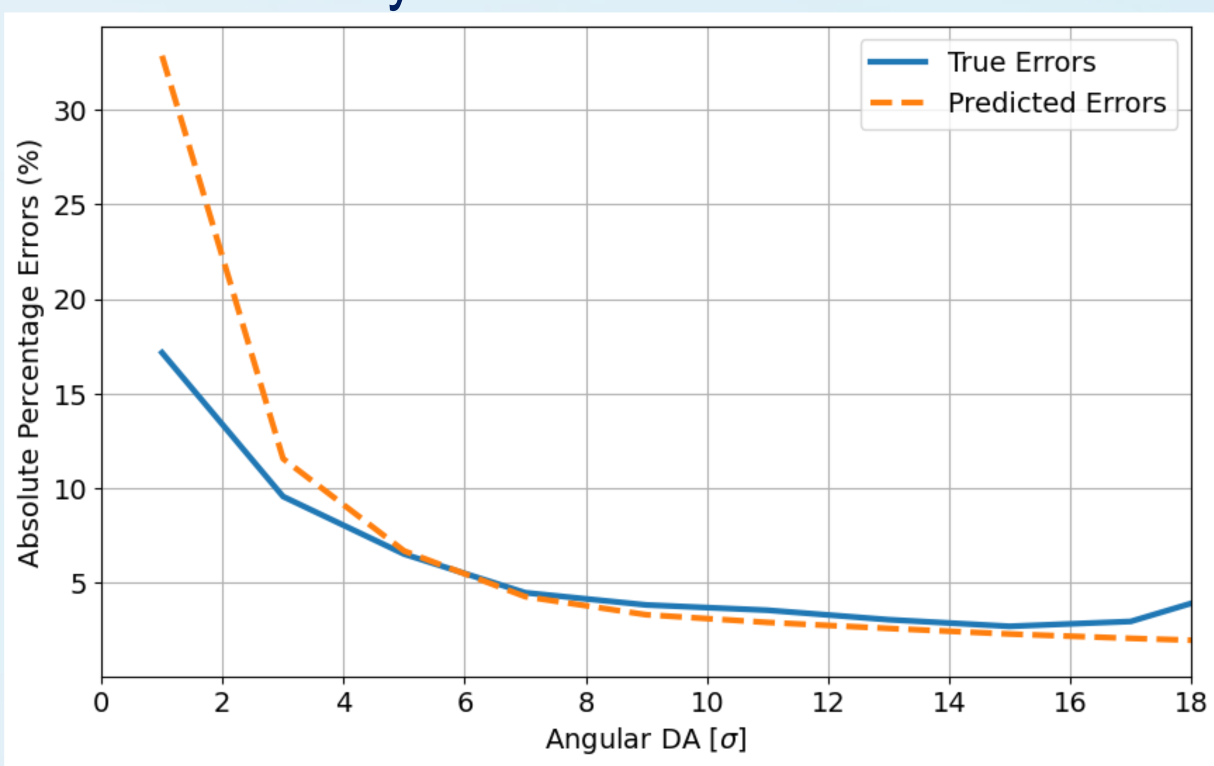




Mean Absolute Error (MAE) used as loss function.
Inference in 1 μs/angular DA (0.5 ms/ machine configuration).
MAE = 0.201 beam σ in test dataset.
Mean Absolute Percentage Error (MAPE) = 11.91%.

ERROR ESTIMATION: MC DROPOUT

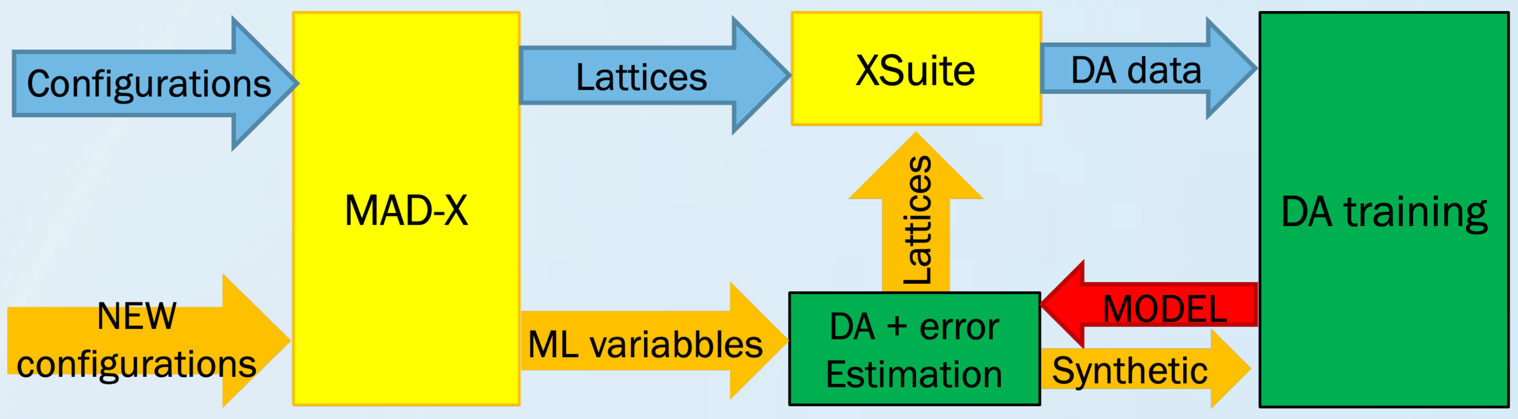
By leveraging dropout (randomly dropping information in some nodes) at inference time, we introduce diversity among the predictions. This technique is known as Monte Carlo (MC) dropout.
These variations are utilized to estimate uncertainties: dropout at 1% between the first hidden layers and 1 s.t.d. of 128 variations as error.



DA and error prediction in 0.75 s/machine configuration.

ACTIVE LEARNING FRAMEWORK

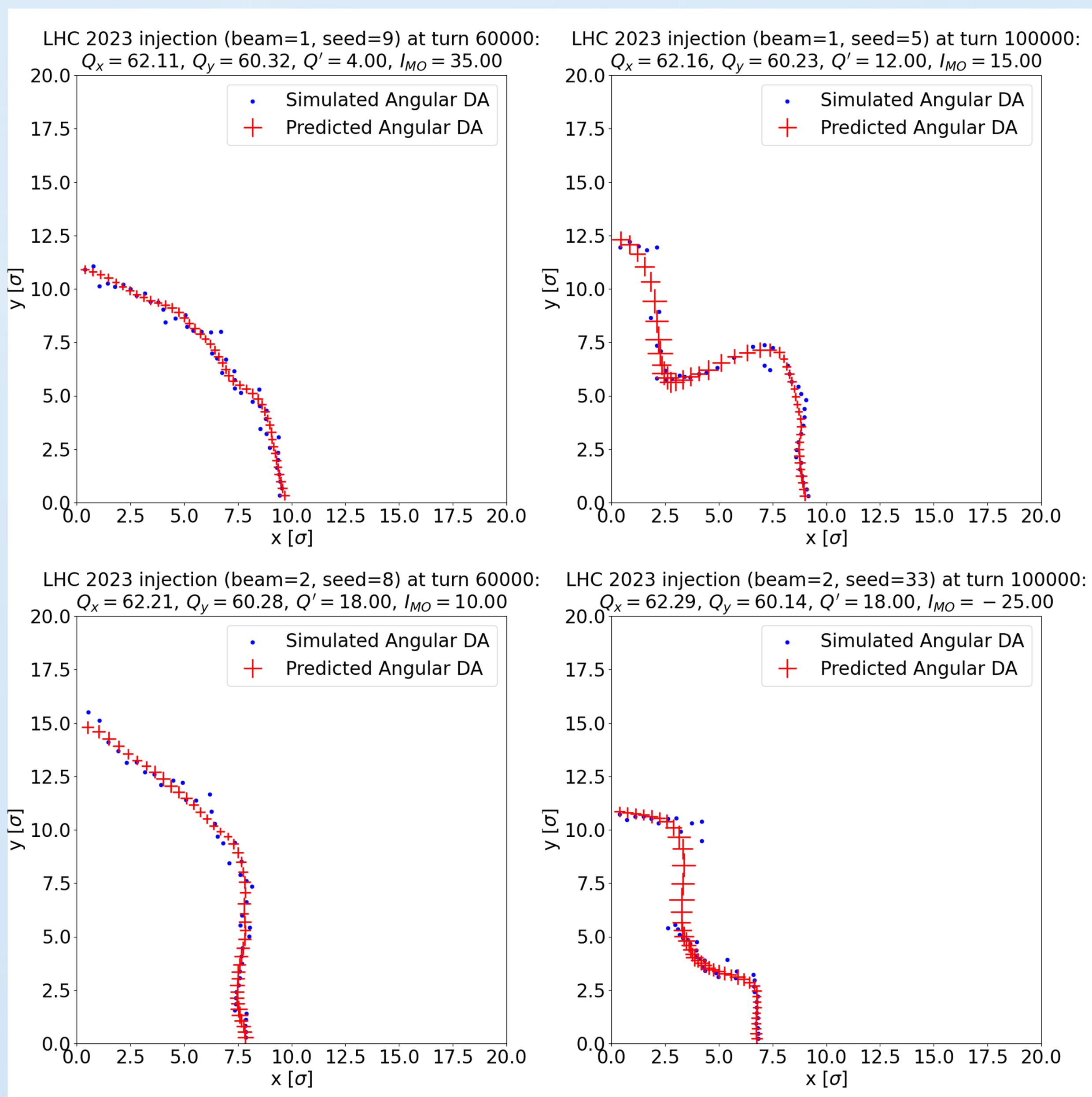
When the estimated error exceeds a percentage error of 10%, we employ the full simulation approach with tracking to accumulate a sufficient number of samples for subsequent training. It is important to note that we prioritize the **ranking of the predicted error** to determine which samples should undergo full simulation first, enabling us to improve our model more efficiently through **smart sampling**.



If the predicted error is lower than 10%, we leverage the angular DA value provided by the regressor, which can also be utilized as **synthetic data** for subsequent training.
By generating 1000 synthetic machine configurations, we retrained and improved the model to a MAPE of 9.57%.

DISCUSSION

Tracking on Xsuite takes 107 s/machine configurations (HT-Condor), while the AL framework, once trained, is **140 times faster!**



This AL framework not only allows for the retraining and updating of the model, but also facilitates **efficient data generation**.
Since the physics of unstable chaotic motion can only be observed through tracking, traditional simulations are integrated into the framework to assess the chaotic nature of initial conditions.

davide.dicroce@epfl.ch
This work is partially funded by the SDSC project C20-10