Machine Learning Processing of BNL AC-LGAD Sensors Readout with Signal Sharing

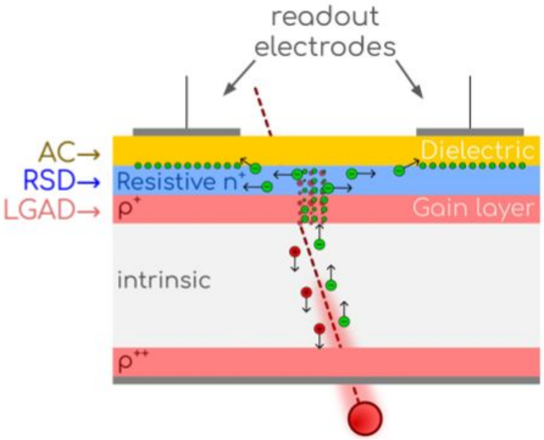
Gaetano Barone[1], Spencer Ellis[1], Gabriele Giacomini[3], Ben Kilminster[2], **Daniel Li[1]**, Anna Macchiolo[2], Matias Sengar[2], Alessandro Tricoli[3], Nikhilesh Venkatasubramanian[1]

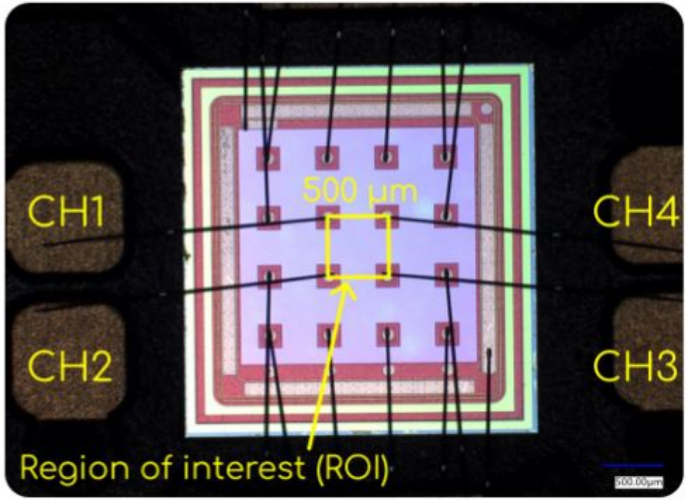
Brown University[1], Brookhaven National Laboratory[2], University of Zurich[3]

Introduction and Overview

Goal: Accurate and precise prediction of charged particle hit coordinates using machine-learning methods taking AC-LGAD waveforms as inputs

- Unique exploit is charge sharing between electrodes → improved spatial resolution
- Limited ROI for study utilizing four electrodes → 570 μm × 570 μm





Specifications:

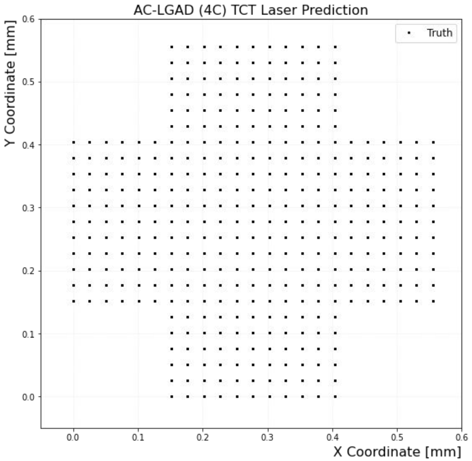
- Manufactured at BNL
- Active thickness: 30 μm
- Pad size: 200 μm
- Pitch: 500 μm

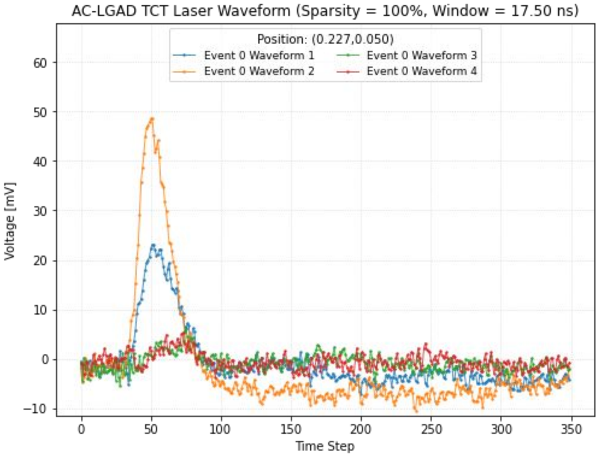
Transient Current Technique Laser Waveforms

Using TCT laser to extract waveforms (without Landau fluctuations)

- TCT waveform dataset collected by Dr. Matias Senger from UZH
- Two available intensities (through attenuation) → “high” and “low”
- 385 grid positions with $25 \times 25\ \mu m^2$ spacing and 111 events per position

Charge sharing exemplified in waveform shape → correlated to hit-to-pad proximity





Specifications:

- 1064 nm laser
- Laser spot Gaussian with $\sigma \sim 9\ \mu m$
- $\sim 1\ \mu m$ spatial resolution
- Laser intensity match 1 MIP
- Two pulses separated by 100 ns

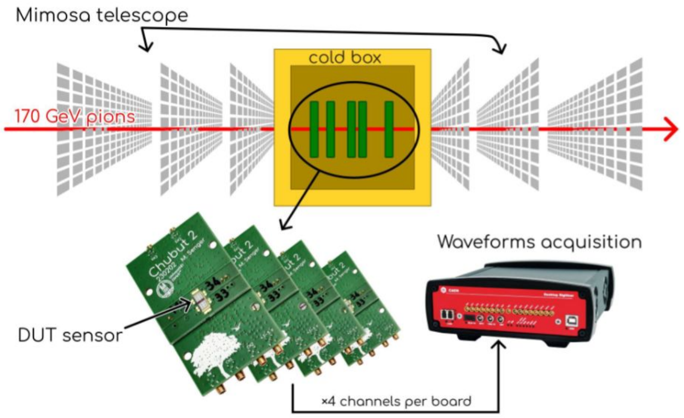
Summary of Previous Studies with Test Beam

Dr. Senger previously studied a DNN trained on "higher-level" waveform information

- Using relative amplitudes
- Best TCT median MSE achieved ~20 μm
- Comparison to analytic charge sharing methods

CERN H6 test beam with O(100) MeV pions → "AIDAinnova"

- Mimosa telescope
- Chubut 2, 4 channels readout board
- CAEN DT5742 digitizer, 500 MHz @ 5 GS/s
- Cold box (-12 °C) for irradiated DUTs



Method	TCT	Test Beam
Analytic Peak Relative Amplitude	~25 μm	~50 μm
DNN Peak Relative Amplitude	~20 μm	~44 μm

Proof of Concept with Recurrent Neural Networks

Pre-processing of waveforms to remove noise and keep "interesting" portions around peak

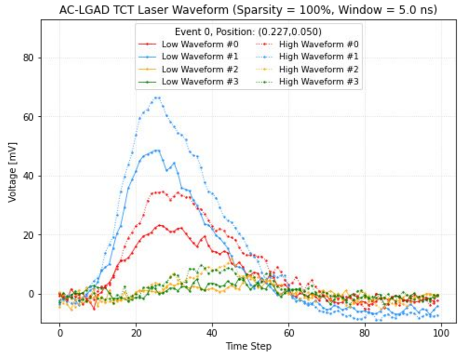
- 100 time steps for each (4) waveform
- Train on both high and low intensity waveforms

Long Short-Term Memory (LSTM) layer # weights scales with series length

- Fewer weights desirable for future application (latency, memory usage)

Model architecture "lightly" tuned to have fewer weights → training regulators included

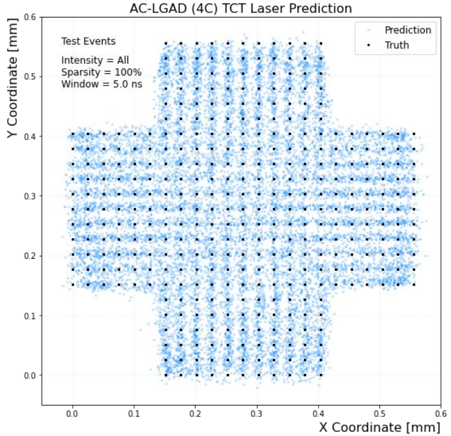
- Converges ~150 epochs with Adam(0.001)
- Batch size ≧ # grid points → 2^9 (512)



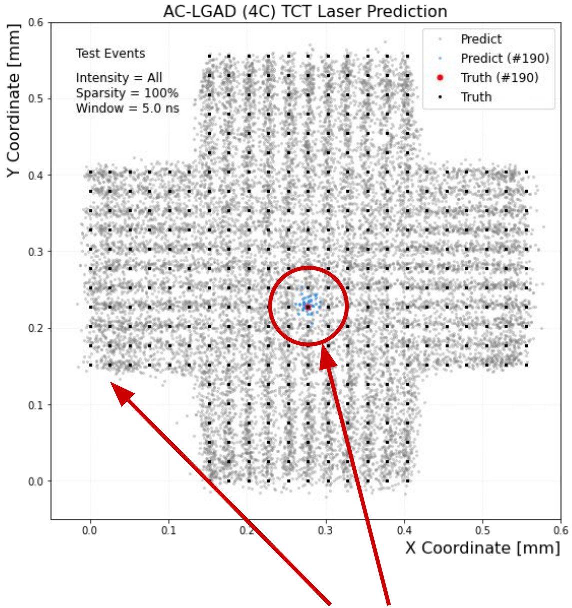
Fixed input series length

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	[(None, 4, 100)]	0
lstm (LSTM)	(None, 32)	17024
dense (Dense)	(None, 32)	1056
dense_1 (Dense)	(None, 32)	1056
dense_2 (Dense)	(None, 32)	1056
dense_3 (Dense)	(None, 2)	66

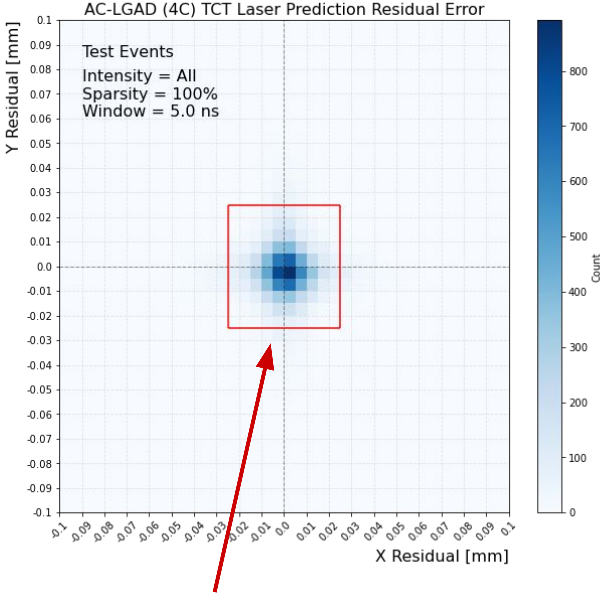
Total params: 20258 (79.13 KB)
Trainable params: 20258 (79.13 KB)
Non-trainable params: 0 (0.00 Byte)



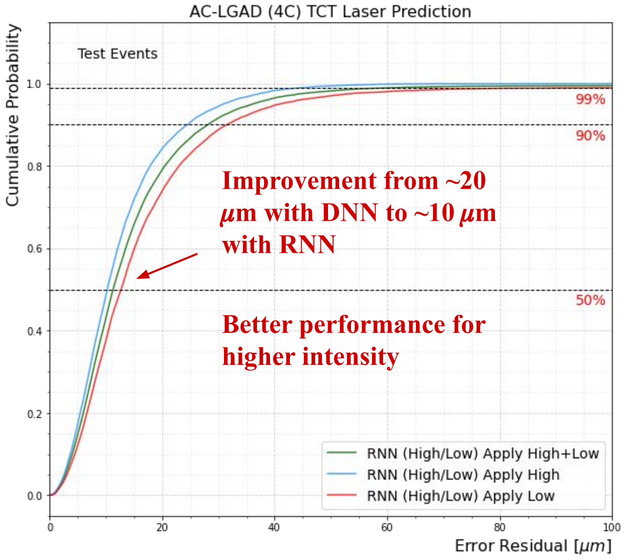
Metrics for Performance



Central regions perform better than the edges



Accurate and precise enough to distinguish two grid points



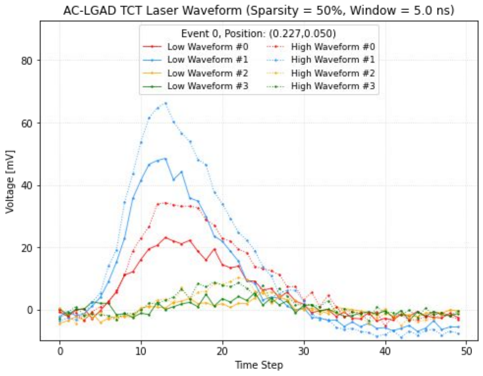
Quantitative metric to compare various methods

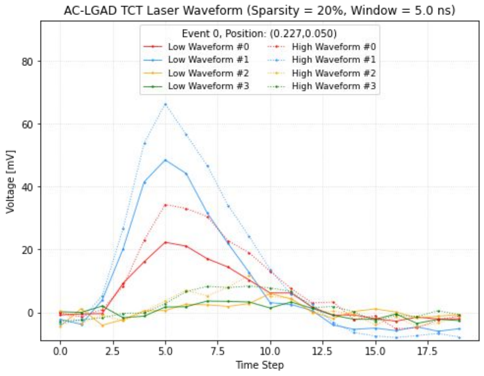
Waveform Rasterization Studies

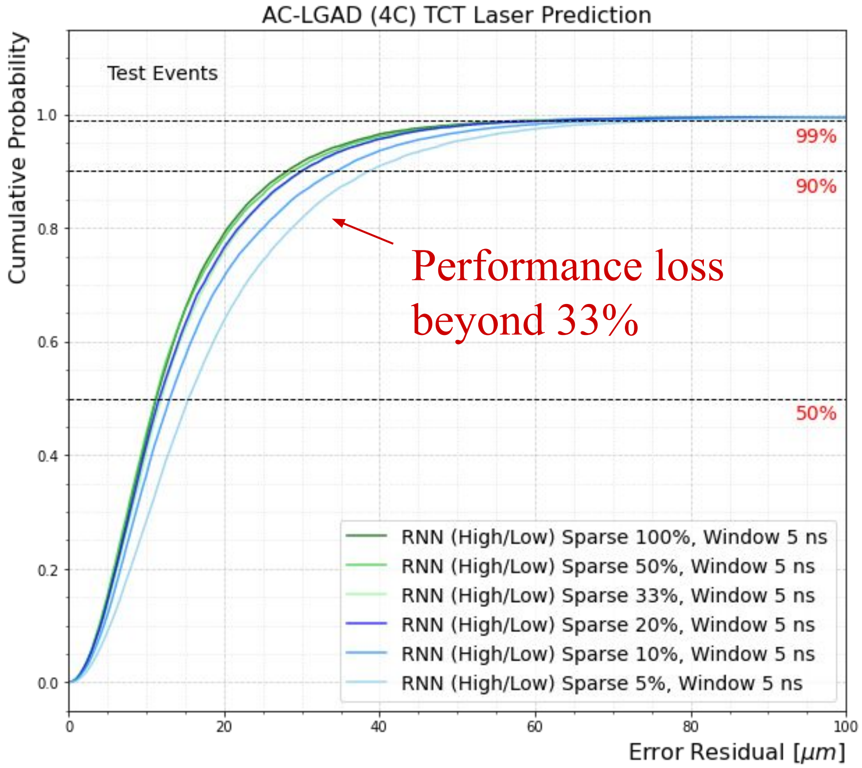
Main parameter of optimization is amount of information necessary to keep from waveforms

- Benefits having fewer LSTM weights

Rasterize the waveform input at a few thresholds of 100/50/33/20/10/5%







Comparison to Analytic Methods

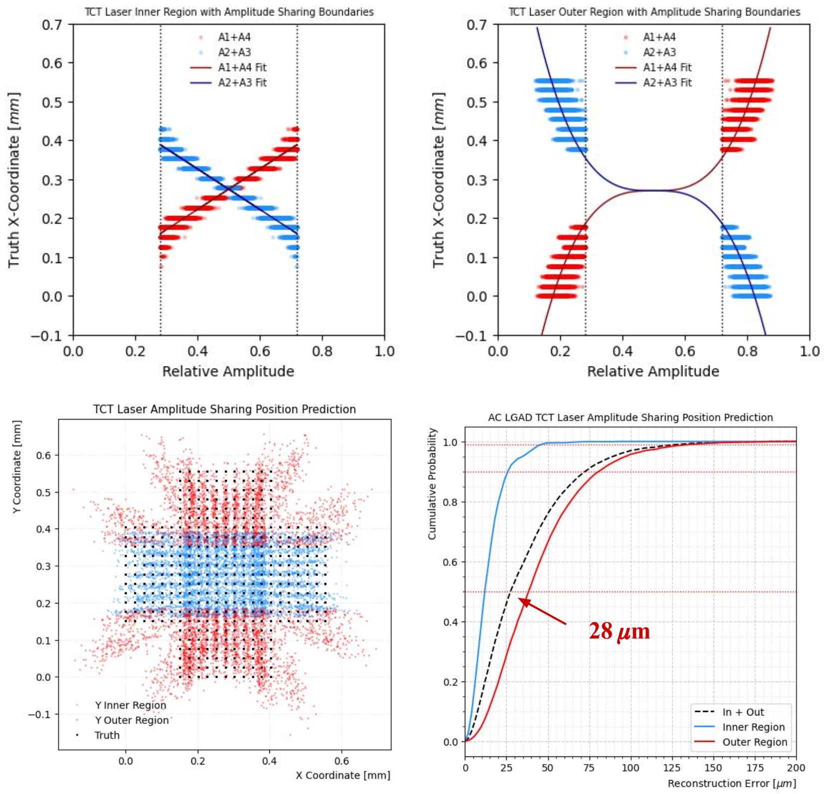
Traditional approach of estimating hit position with strip sensors and charge sharing uses *only* the waveform peak value

- 1-dimensional problem with linear regression using relative amplitudes

For 2D grid, higher-order (3rd) regression required in “outer” region

- Define regions by relative amplitudes

Inner region predictions expectedly accurate but outer regions (majority) very inaccurate



Summary and Outlook

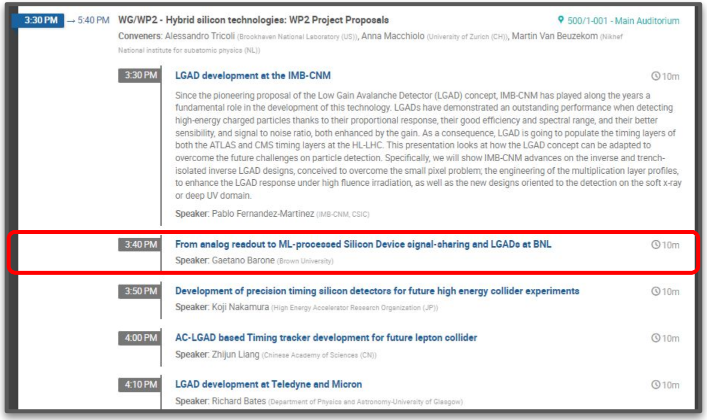
Improved hit reconstruction with RNN using "full" waveform at median MSE ~10 μm

- Peak amplitude with DNN ~20 μm
- Peak amplitude analytic method ~25 μm
- Projected improved performance with test beam

Performance maintained at 33% rasterization and spatial interpolation achievable (not shown)

Future and ongoing work:

- Expanding beyond 4 channels
- Test beam and incorporating Landau noise
- HLS conversion and application on FPGA


3:30 PM → 5:40 PM WG/WP2 - Hybrid silicon technologies: WP2 Project Proposals — 500/1-001 - Main Auditorium

Conveners: Alessandro Tricoli (Brookhaven National Laboratory (US)), Anna Macchiolo (University of Zurich (CH)), Martin Van Beuzekom (Nikhef National institute for subatomic physics (NL))

3:30 PM LGAD development at the IMB-CNM — 10m

Since the pioneering proposal of the Low Gain Avalanche Detector (LGAD) concept, IMB-CNM has played along the years a fundamental role in the development of this technology. LGADs have demonstrated an outstanding performance when detecting high-energy charged particles thanks to their proportional response, their good efficiency and spectral range, and their better sensibility, and signal to noise ratio, both enhanced by the gain. As a consequence, LGAD is going to populate the timing layers of both the ATLAS and CMS timing layers at the HL-LHC. This presentation looks at how the LGAD concept can be adapted to overcome the future challenges on particle detection. Specifically, we will show IMB-CNM advances on the inverse and trench-isolated inverse LGAD designs, conceived to overcome the small pixel problem; the engineering of the multiplication layer profiles, to enhance the LGAD response under high fluence irradiation, as well as the new designs oriented to the detection on the soft x-ray or deep UV domain.

Speaker: Pablo Fernandez-Martinez (IMB-CNM, CSIC)

3:40 PM From analog readout to ML-processed Silicon Device signal-sharing and LGADs at BNL — 10m

Speaker: Gaetano Barone (Brown University)

3:50 PM Development of precision timing silicon detectors for future high energy collider experiments — 10m

Speaker: Koji Nakamura (High Energy Accelerator Research Organization (JP))

4:00 PM AC-LGAD based Timing tracker development for future lepton collider — 10m

Speaker: Zhijun Liang (Chinese Academy of Sciences (CN))

4:10 PM LGAD development at Teledyne and Micron — 10m

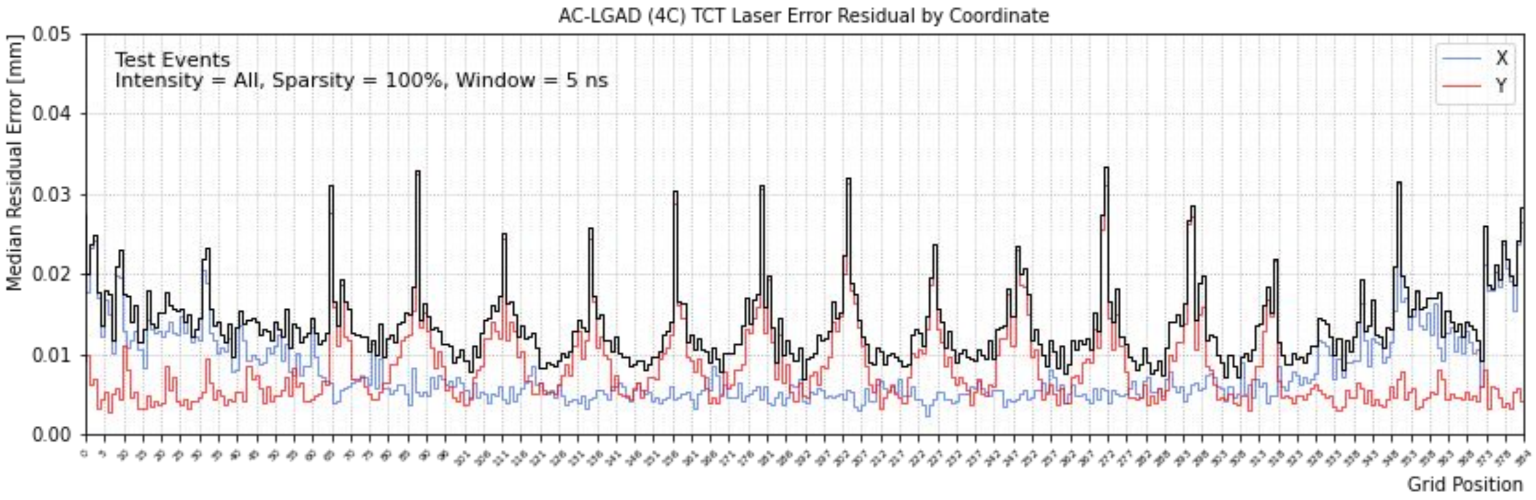
Speaker: Richard Bates (Department of Physics and Astronomy-University of Glasgow)

Join Gaetano Barone's talk this afternoon for a deeper discussion regarding the "bigger" picture and more technical hardware details

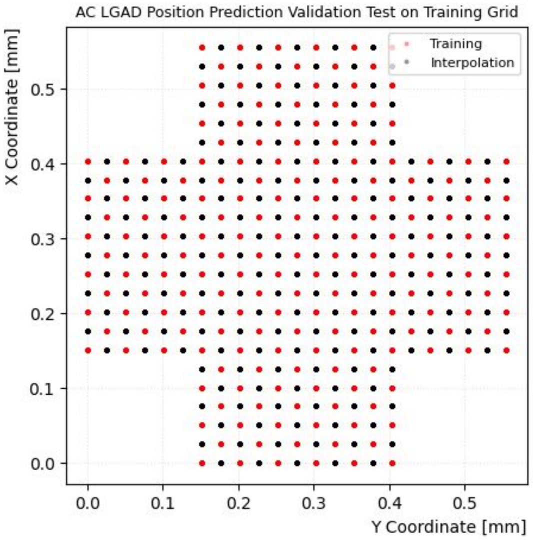
Thank you for your attention–Questions?

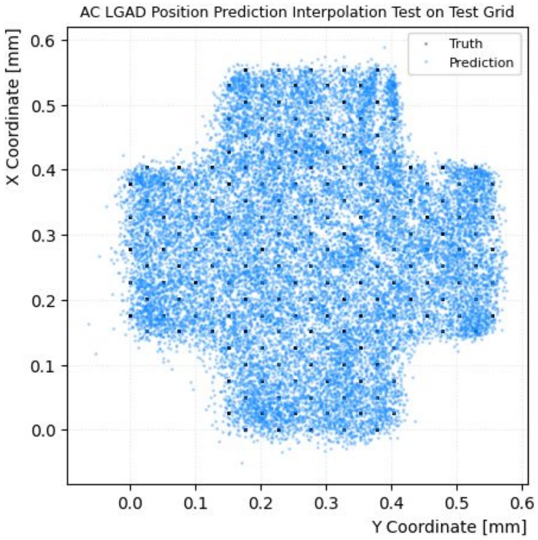
Backup Slides

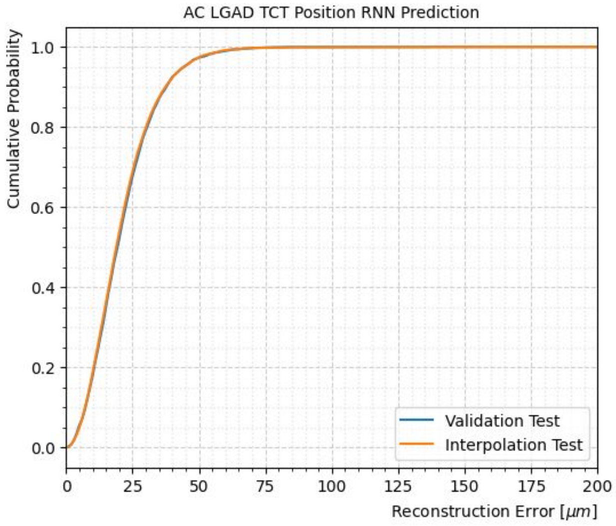
Characterization of RNN Prediction



Interpolation Studies

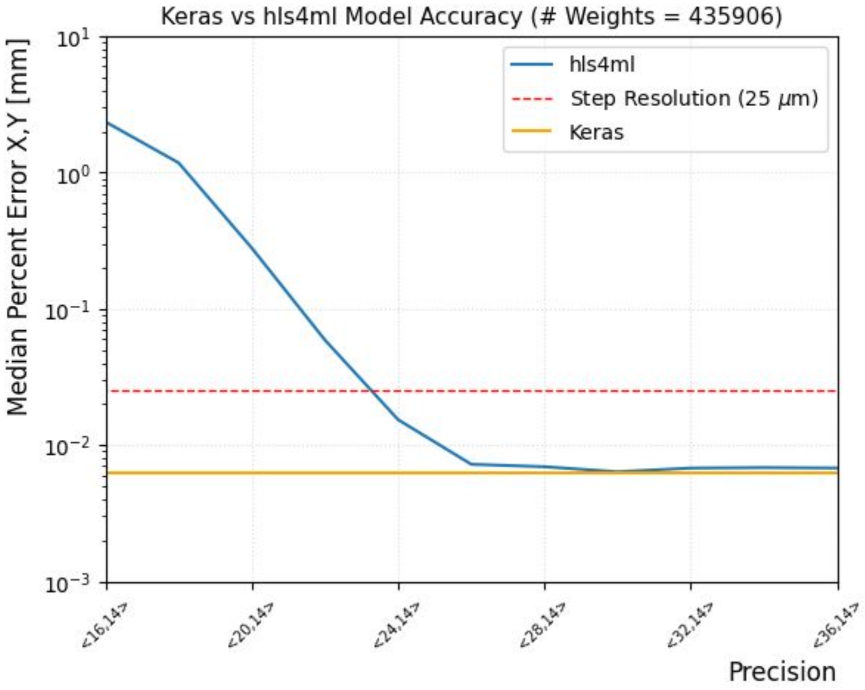


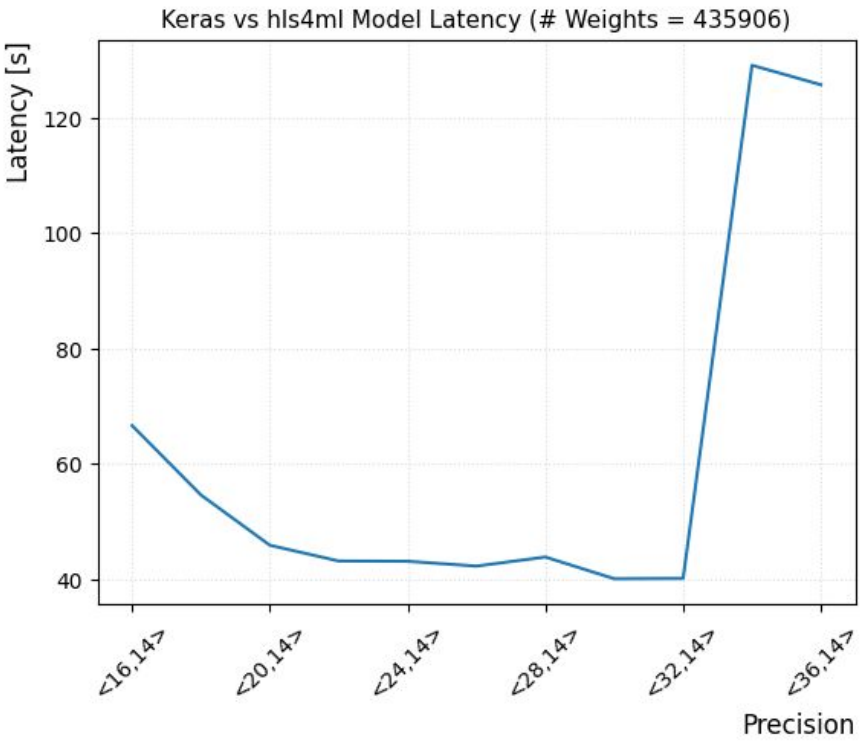




The network is trained to interpolate within a 25 μm step size $\rightarrow$ can target finer levels of precision with closer grid spacing

hls4ml Studies





hls4ml Keras Conversion