Machine Learning for Energy Reconstruction at ATLAS

University of Victoria

ATLAS EXPERIMENT

Luke Polson

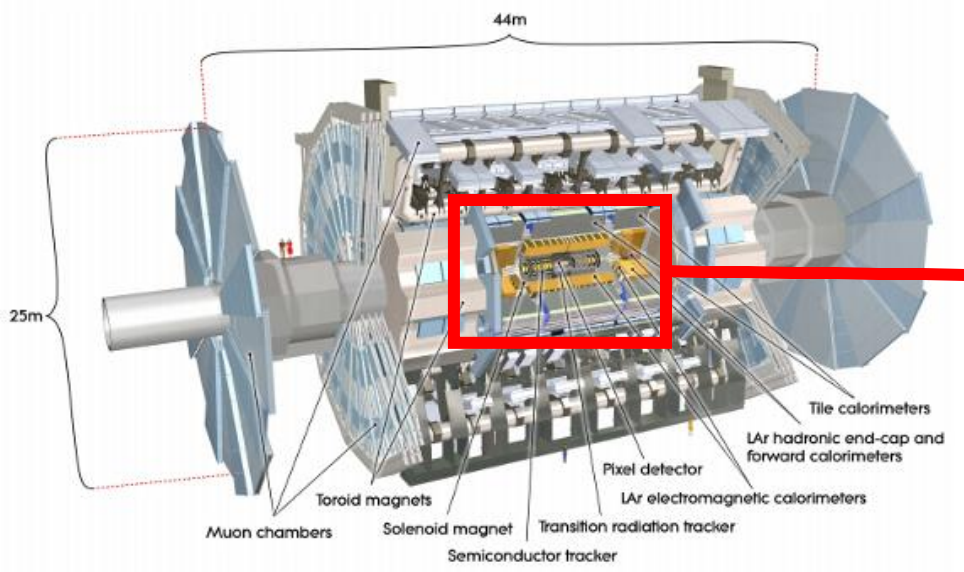
Supervisor: Dr. Michel Lefebvre

February 9-12, 2021

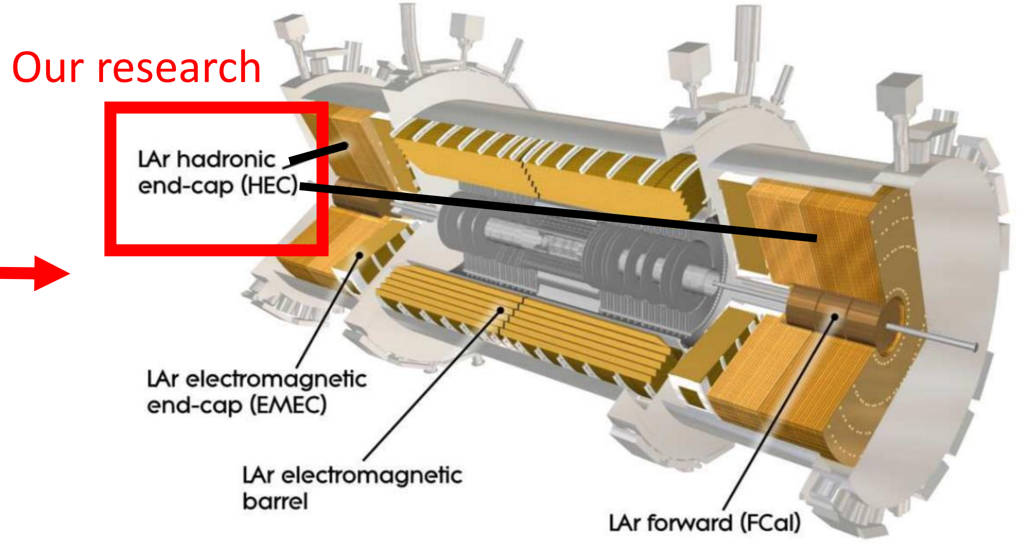
ATLAS and LAr Calorimeter

- The ATLAS detector is one of the primary detector experiments at CERN
  - Composed of many subdetectors: one such subdetector is the **Liquid Argon Calorimeter**
  - The **Liquid Argon Calorimeter** itself consists of many components; this research focuses on the **Hadronic End-Cap**
  - Primary purpose of the **Hadronic End-Cap** is to reconstruct the energy of hadronically interacting particles

ATLAS



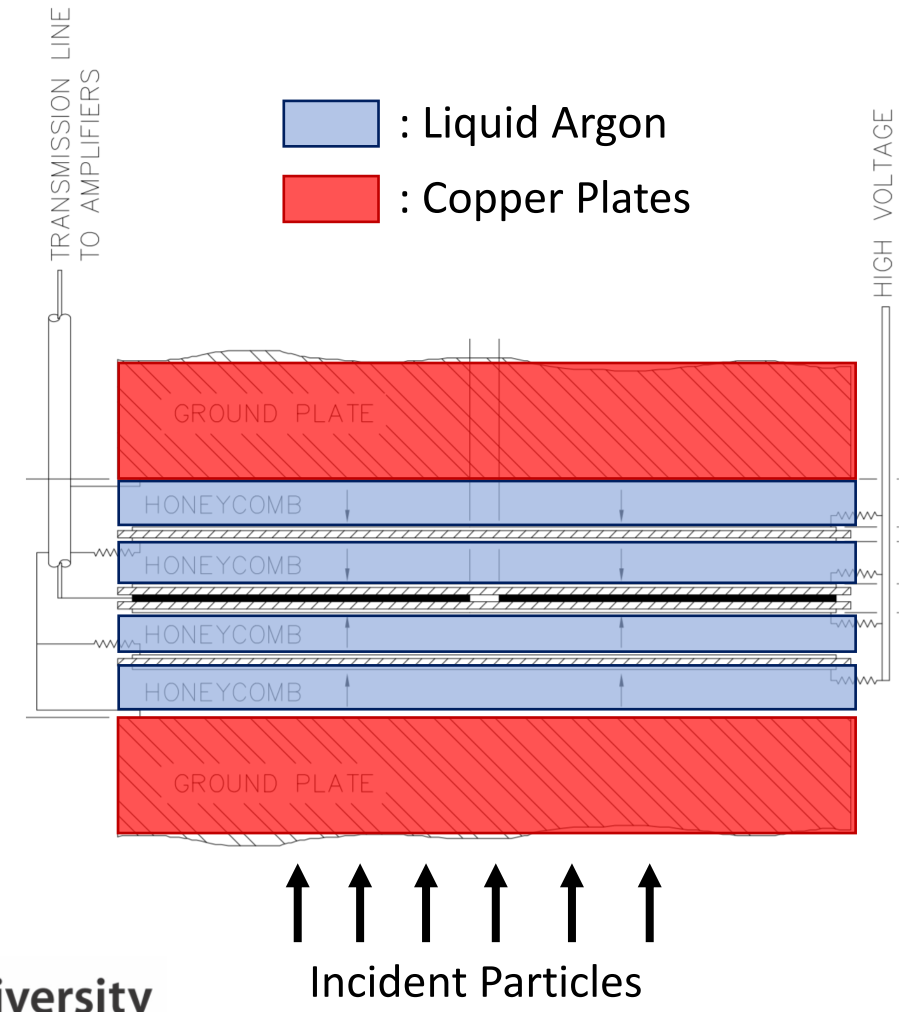
Liquid Argon Calorimeter



University of Victoria

Energy Reconstruction in HEC Cell

- The entire LAr Calorimeter consists of approximately 183000 distinct "**cells**" placed cylindrically around the beam line. The following is a cell in the HEC

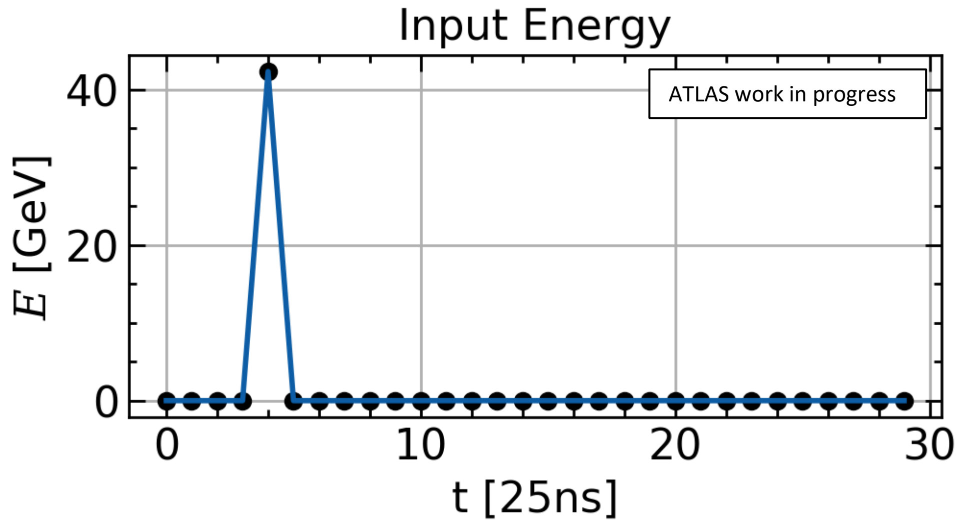


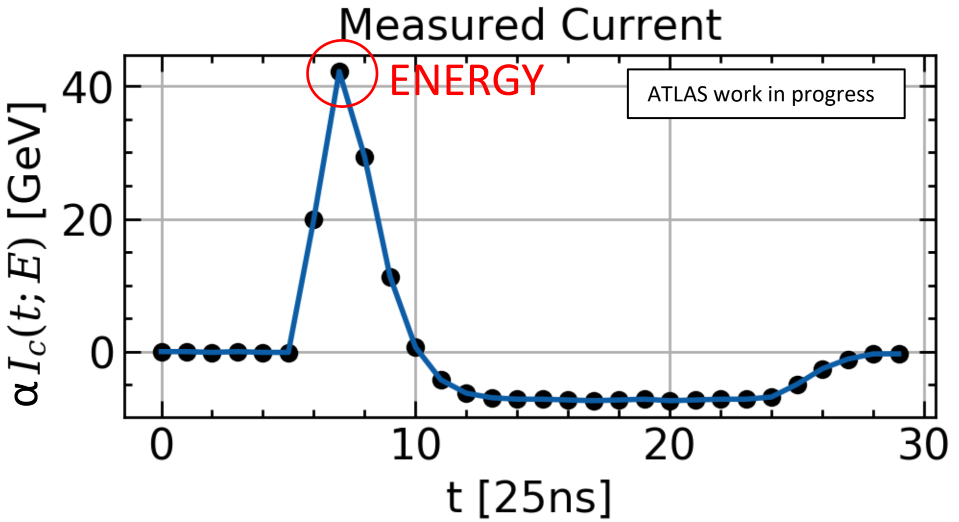
1. Incident particles strike copper plates and the particle shower further develops
2. The shower is sampled by the liquid argon, where charged particles ionize the argon.
3. An electric field in the liquid argon gaps causes the released electrons to drift into the transmission line
4. The drifting electrons induce a current, which can be measured, and which is proportional to the charged particle tracks in the gaps
5. For a shower contained in the calorimeter, the total current is proportional to the incoming particle energy

University of Victoria

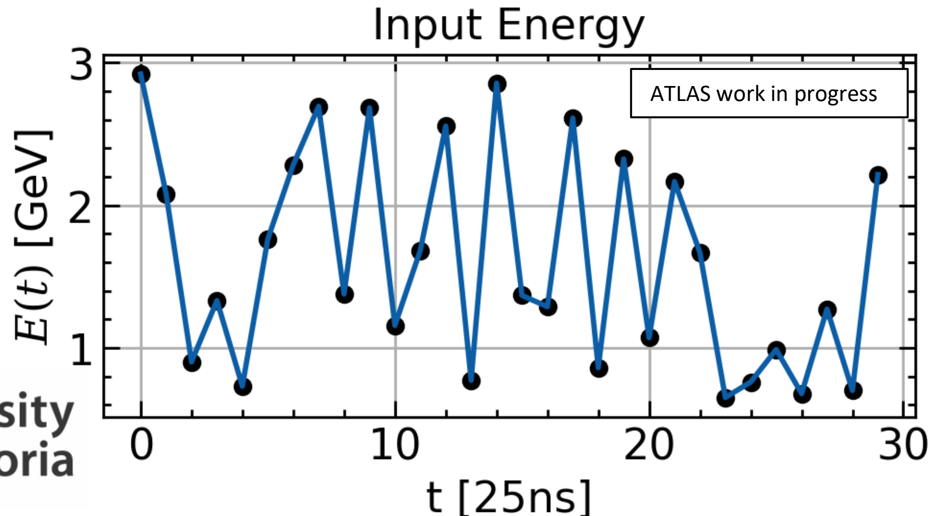
Energy Reconstruction in HEC Cell

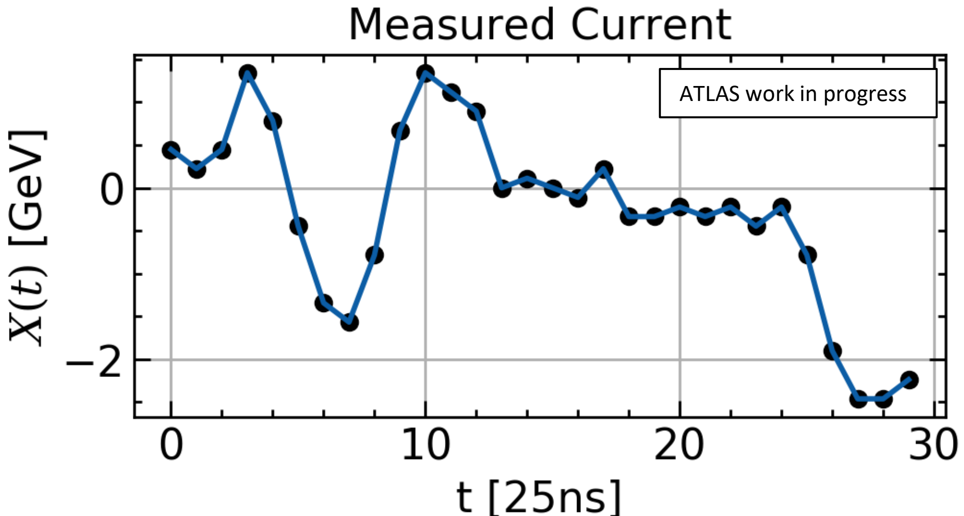
- For a **single incident particle**





- For a **continuous series of incident particles…**

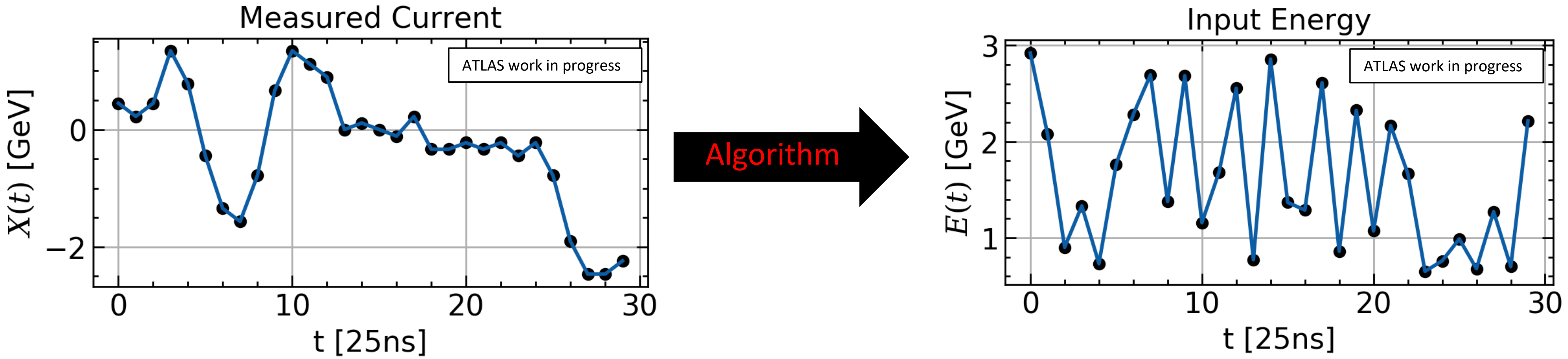




University of Victoria

Energy Reconstruction Theory

- ATLAS Researchers want to extract **incident energy time series** after measuring **measured current time series**



- Energy data in this thesis is Monte Carlo generated; measured current is obtained through the detailed ATLAS electronic readout simulation AREUS

Currently Used:

$$E_t = a_0 X_t + a_1 X_{t+1} + \dots + a_n X_{t+n} + b$$

- n = 25 coefficients obtained through optimal filter technique
- filter depth of 25

My Research:

Convolutional Neural Networks

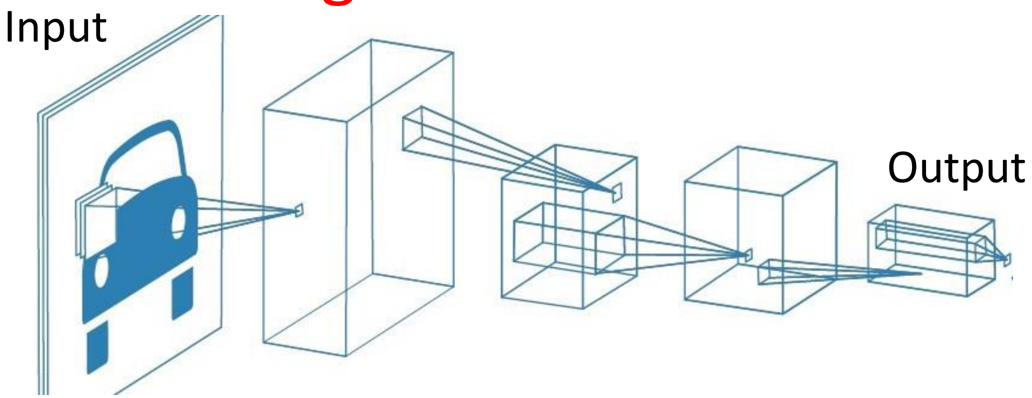
- less than 100 parameters
- filter depth of 25

University of Victoria

Machine Learning Specifications

- The Data set: For each detector cell…
  - 30 million discrete times; training set constitutes 80% and validation constitutes 20%. All results are shown for validation data set.
- The Model
  - A convolutional neural network is a sequence of mappings (3D convolutions) followed by a non-linear activation function at each step

In general



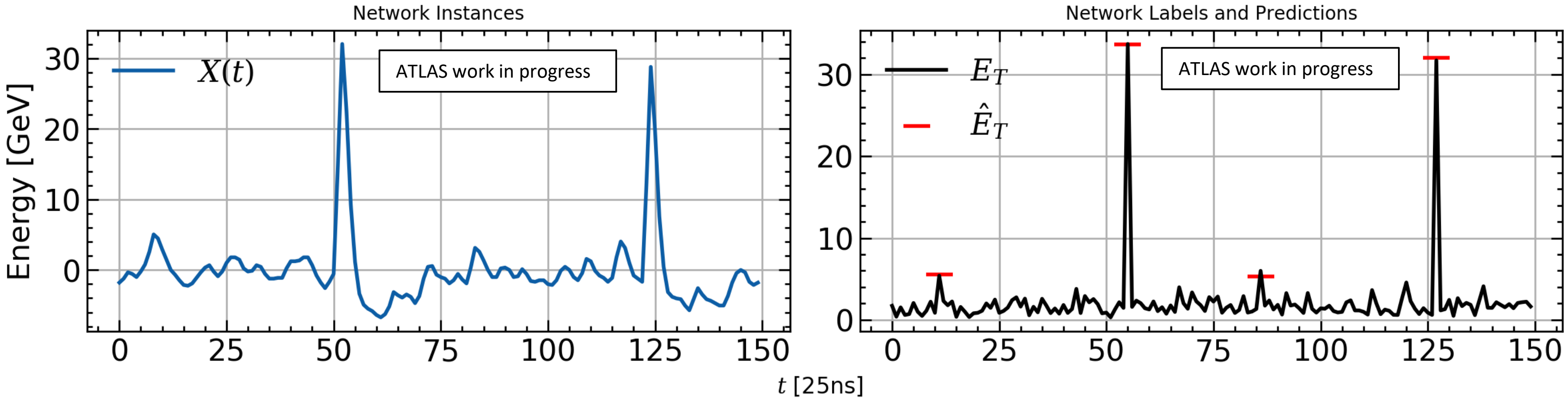
Our usage

1. The initial "image" is 1 dimensional (time series) and has no color-channel dimension (univariate)
2. The output is also a 1 dimensional time series the same length as the input

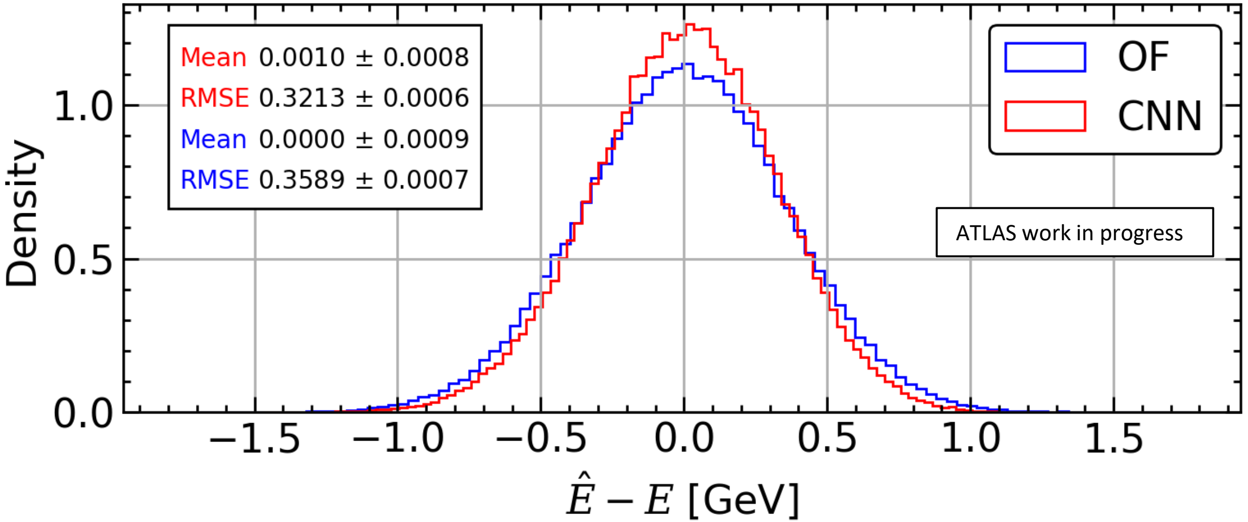
University of Victoria

Basic Results

- Optimal Filter and Convolutional Neural Network only make predictions at times during simulated physics events of interest



- Basic Convolutional Neural Networks have a lower RMSE across all predictions



$$\mathrm{RMSE} = \sqrt{\left\langle (\hat{E} - E)^2 \right\rangle}$$

University of Victoria

Basic Results

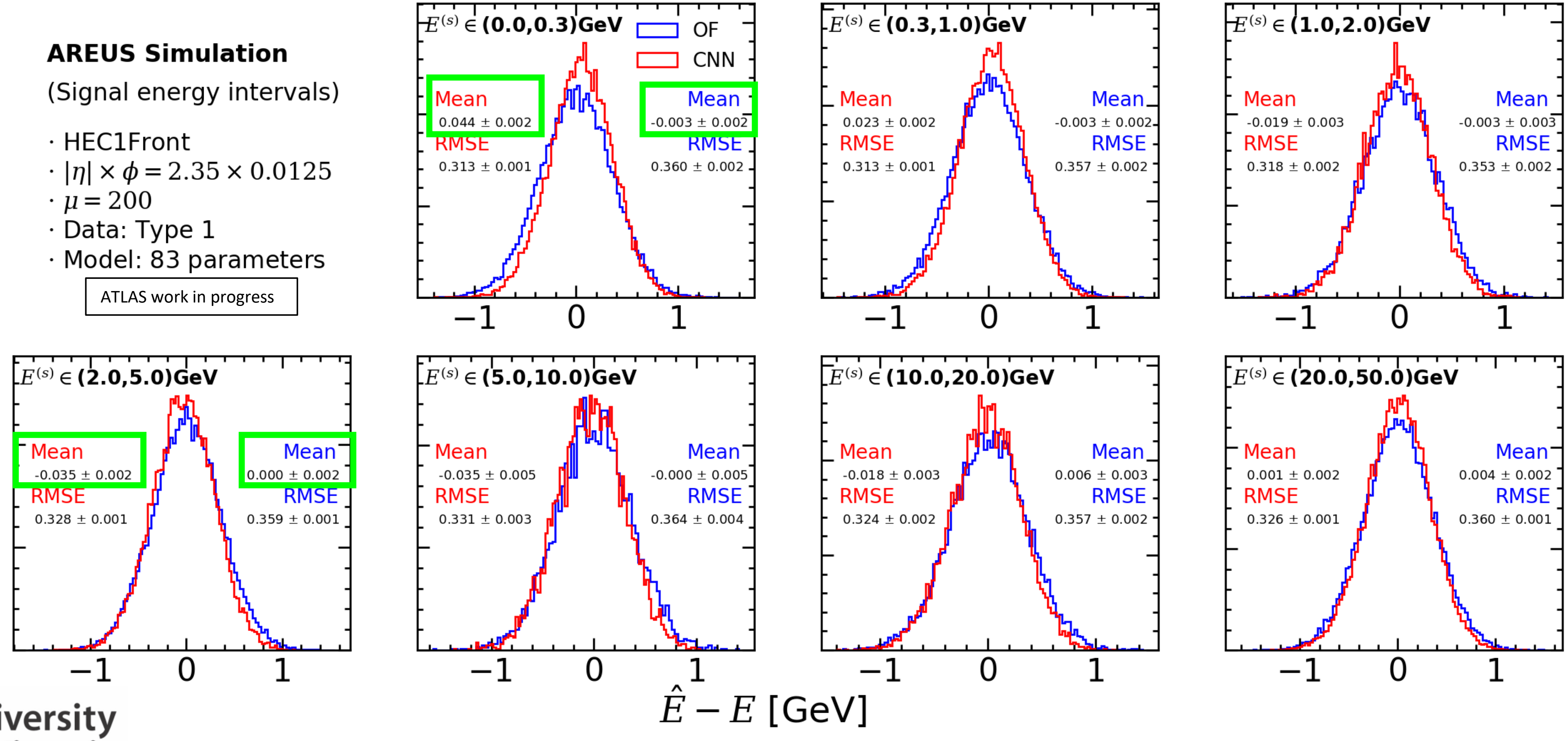
- While RMSE provides one metric of the strength of a model, **model linearity** is another important metric.
  - Model Linearity is the extent to which a model makes non-biased predictions in all energy intervals

AREUS Simulation

(Signal energy intervals)

· HEC1Front
· $|\eta| \times \phi = 2.35 \times 0.0125$
· $\mu = 200$
· Data: Type 1
· Model: 83 parameters

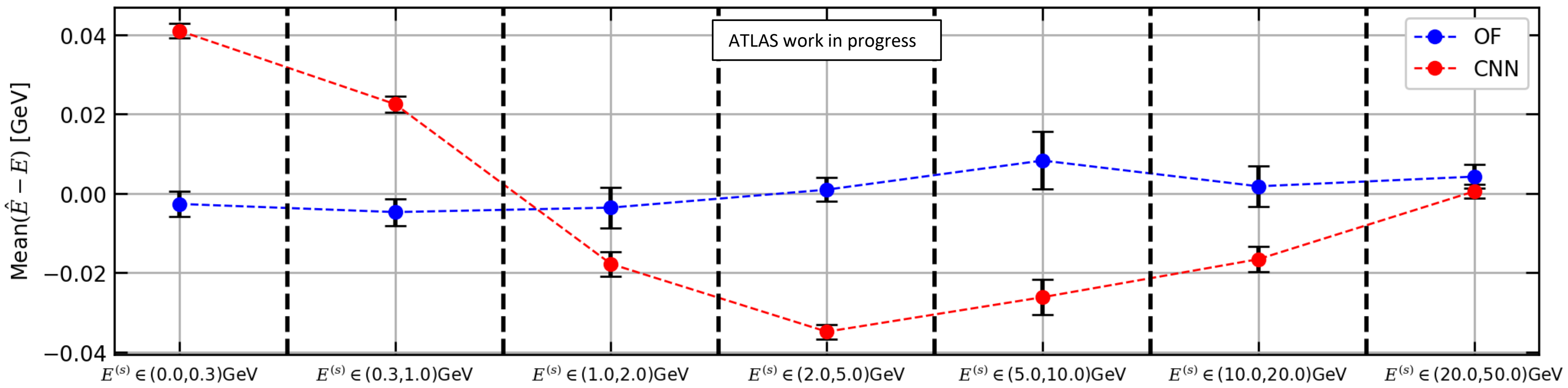
ATLAS work in progress



University of Victoria

Improved Loss Function

- Convolutional Neural Networks trained using a RMSE loss function tend to <u>overpredict small energies</u> and <u>underpredict intermediate energies</u>!



- When training the Convolutional neural network, "punish" not only according to RMSE, but also according to **<u>the squared sum of deviations</u>** on the plot above

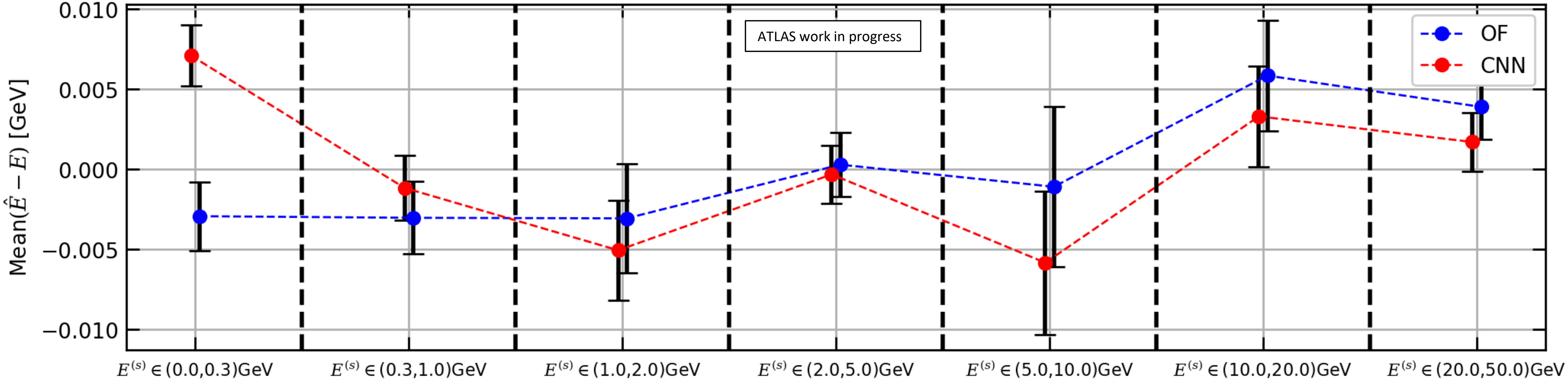
Sum of squared deviations from zero on plot above

RMSE

Custom Loss: $$L_1(P) = \boxed{L_0} + \alpha \boxed{\sum_{j=1}^{n} \left(\frac{1}{|I_{S_j}|} \sum_{t \in I_{S_j}} (\hat{E}_t - E_t) \right)^2}$$

Improved Loss Function Results

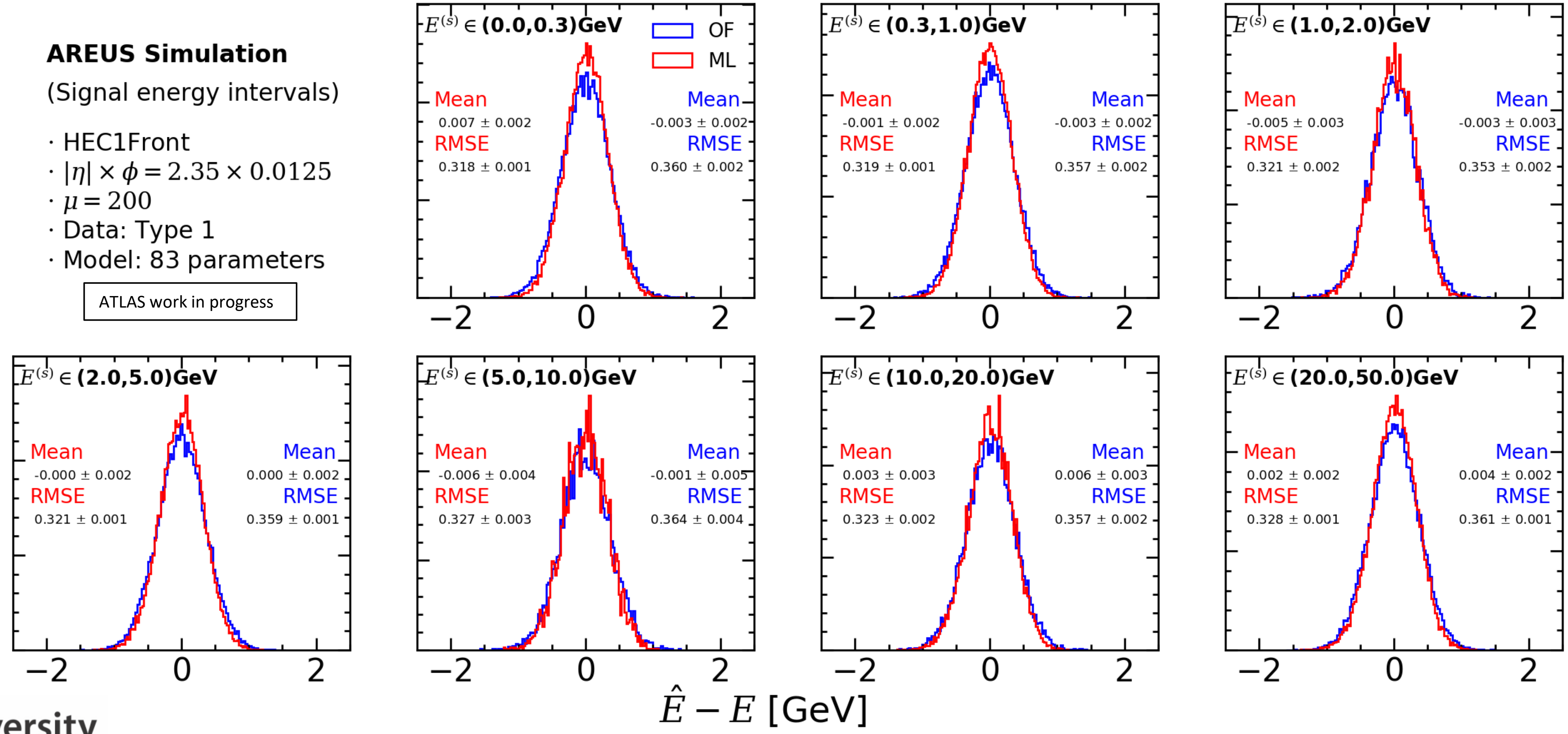
- Model Linearity greatly improves, and Convolutional Neural Network still outperforms Optimal Filter in terms of RMSE
  - Small price to pay: RMSE worsens a little bit



University of Victoria

Improved Loss Function Results

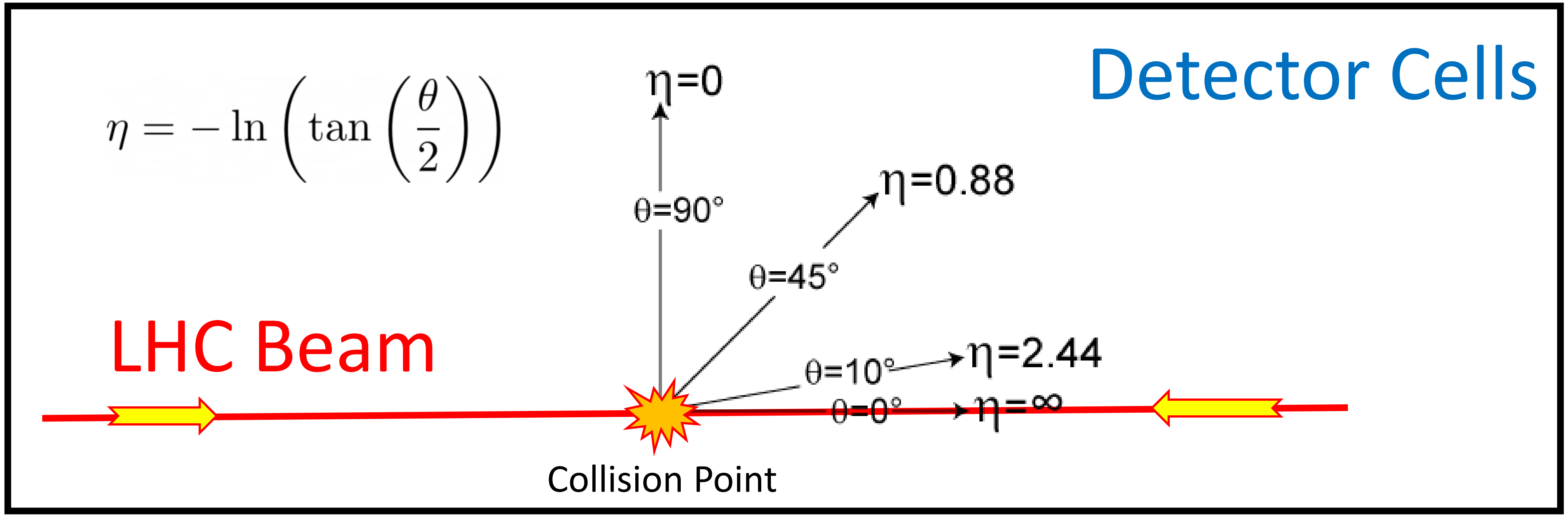
- Model Linearity greatly improves, and Convolutional Neural Network still outperforms Optimal Filter in terms of RMSE
  - Small price to pay: RMSE worsens a little bit



University of Victoria

Results for Entire Hadronic End-Cap

- Can examine results as a function of pseudorapidity η (related to polar angle) from the beam line in different regions of the HEC
  - Cells at larger η experience more background particle collisions, and hence predictions tend to have a higher RMSE

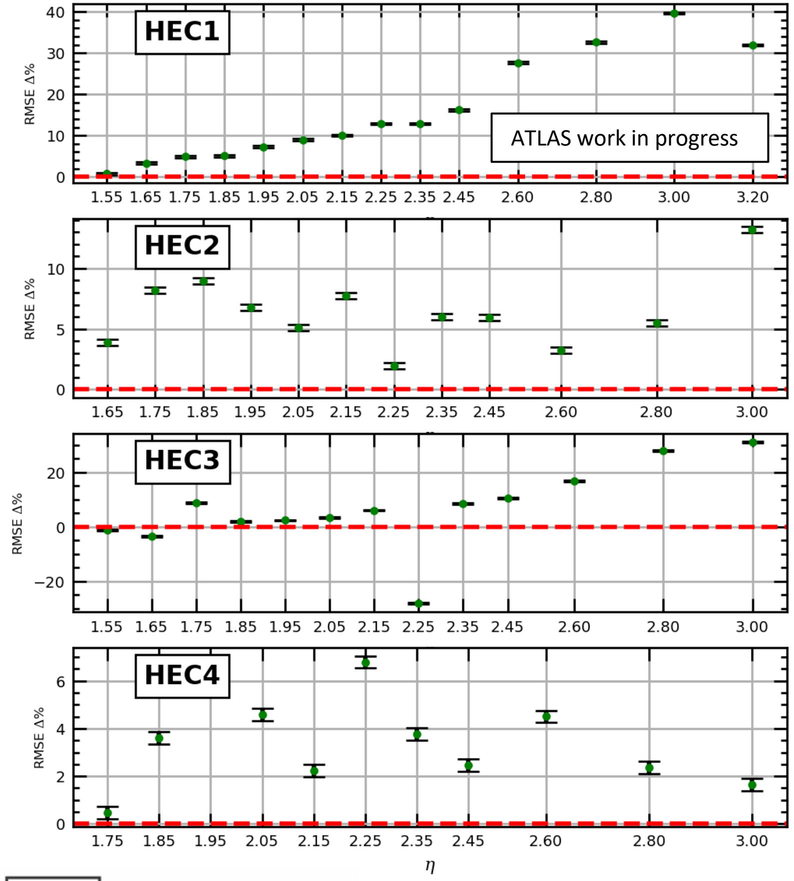


University of Victoria

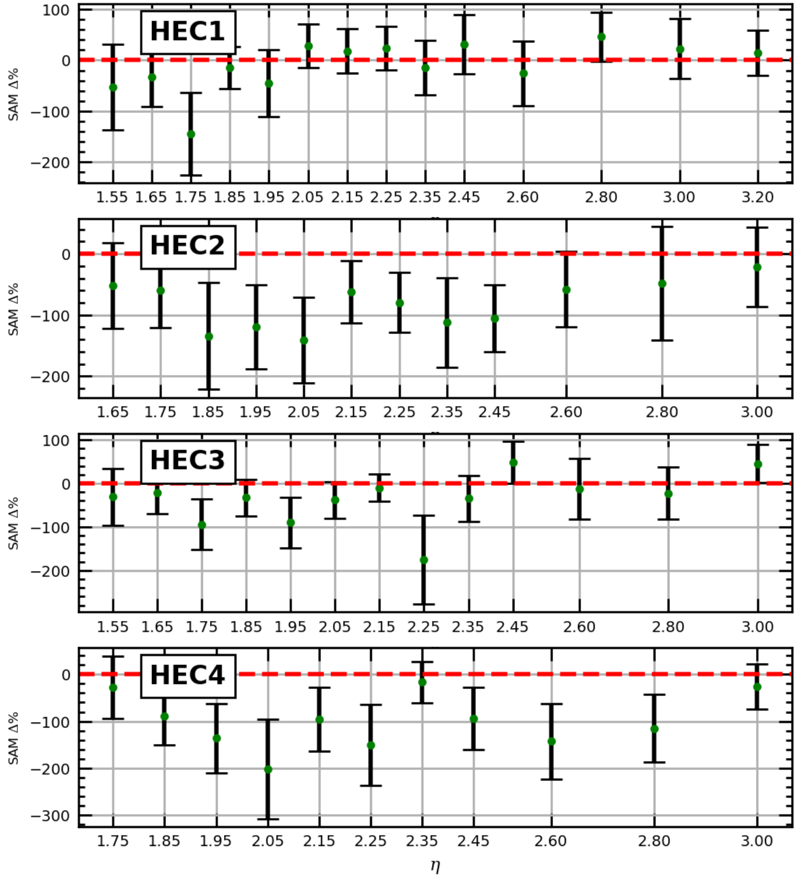
Results for Entire Hadronic End-Cap

- For each detector cell, ten neural networks were trained
  - Shown is the model with the best model linearity

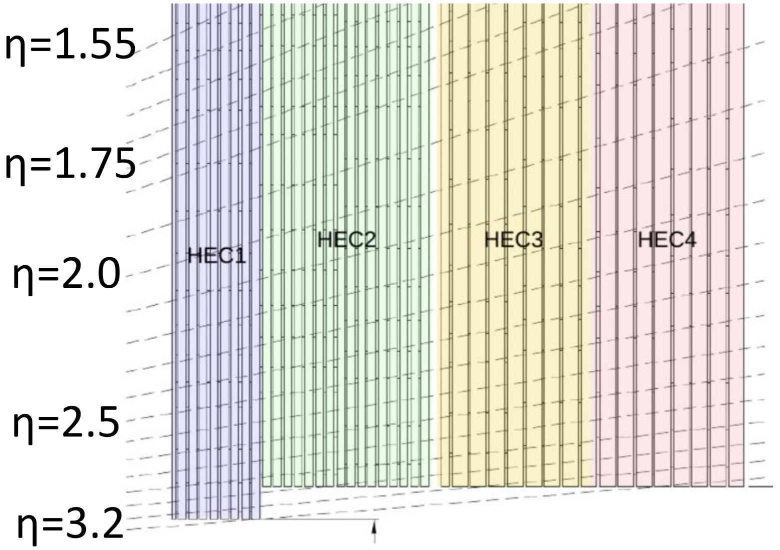
RMSE



Model Linearity



$$\Delta\% = 100\% \times \frac{OF - CNN}{OF}$$



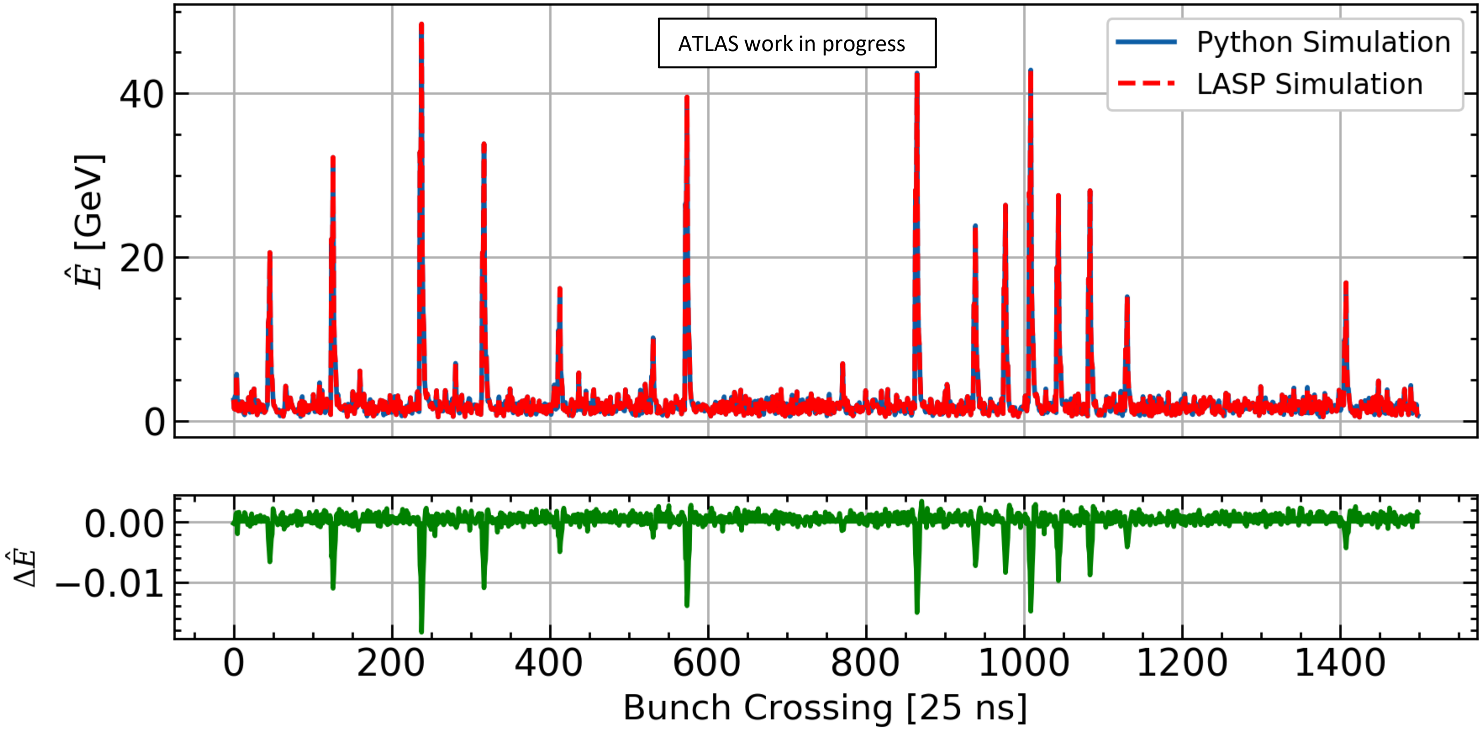
University of Victoria

Road to Implementation

- In the ATLAS phase-II upgrade, signal processing will be implemented on Field Programmable Gate Arrays (FPGAs)
  - The Liquid Argon Signal Processing (LASP) group of ATLAS has developed software to implement and simulate various signal processing techniques, including basic neural networks
  - CNNs converted appropriately and tested using professional LASP simulation software
    - Convert weights from floating point to binary (necessary loss of precision!)

Using LASP simulation:

STEP 1: Input ionization current time series
STEP 2: Simulate neural network and obtain energy time series output using LASP simulation software



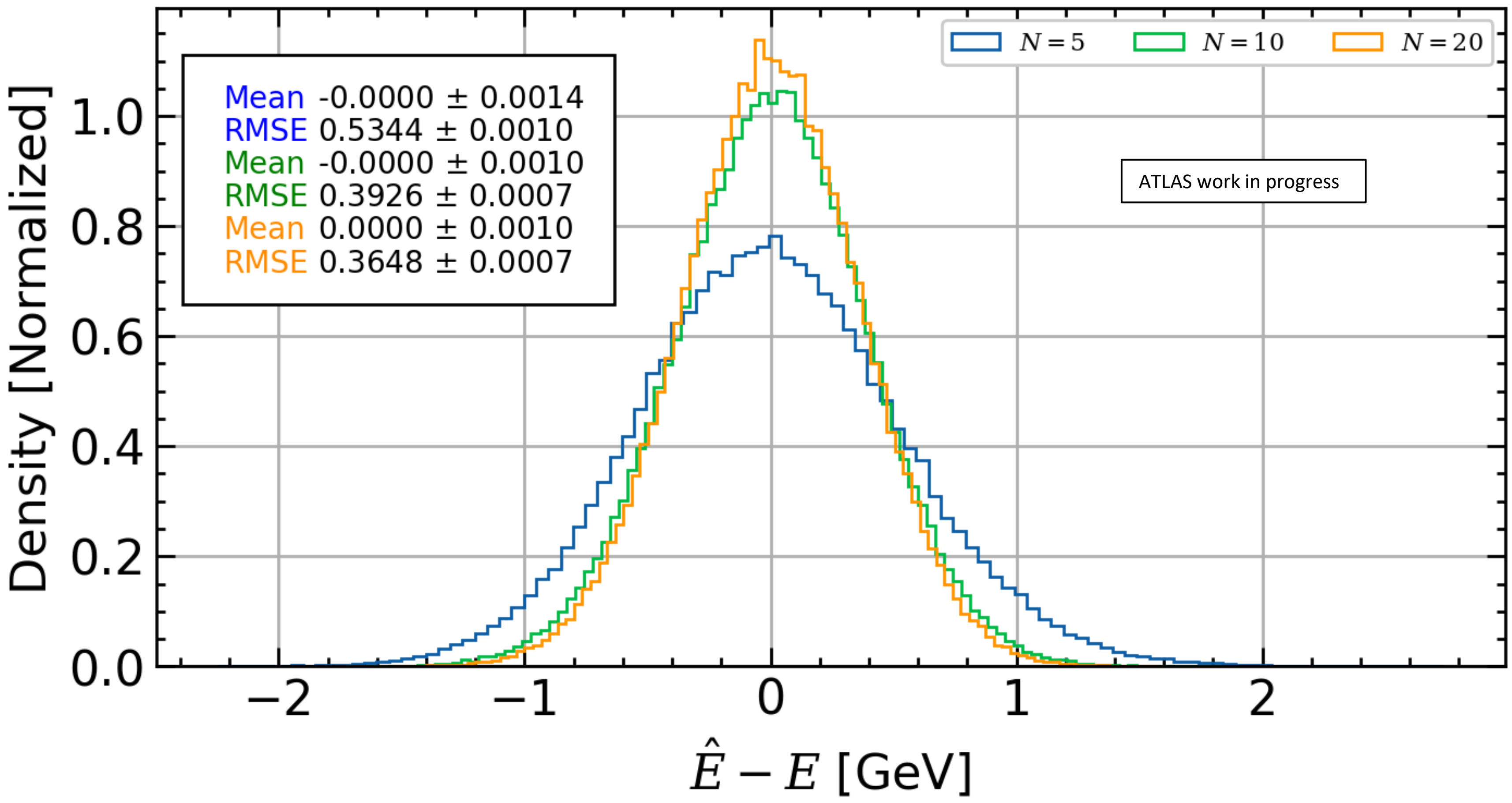
University of Victoria

Conclusion

- Convolutional neural networks outperform the traditional optimal filtering technique in terms of RMSE, while maintaining (and sometimes outperforming) in terms of model linearity
- The convolutional neural networks used in this analysis have been successfully simulated on professional ATLAS Liquid Argon Signal Processing libraries. Such libraries will be used to implement signal processing in the Phase-II upgrade

Backup

- N = 5 is the current optimal filter depth implemented in ATLAS
  - See residuals for η=2.35 below for different values of N



University of Victoria