Topo-Clustering Classification with Deep Neural Networks

ATLAS Machine Learning Forum

Aviv Cukierman[1,2] **Michael Kagan**[2] **David Miller**[3]
Joakim Olsson[3] **Ariel Schwartzman**[2]

[1]**Stanford University** [2]**SLAC National Accelerator Laboratory** [3]**University of Chicago**

February 2, 2017

THE UNIVERSITY OF CHICAGO

Introduction

(Slides from ATLAS Hadronic Final State Forum, Dec 2016 @SLAC)

- *Is it possible to do local cluster calibration using machine learning?*
- *How well can we do compared to default LC calibration?*

- Start with a simpler problem: $\pi^{\pm}$ vs. π^{0} classifier
- Build images from cells associated with clusters for the 6 layers of the barrel calorimeter (LAr + Tile)
  - EMB1, EMB2, EMB3, TileBar0, TileBar1, TileBar2
- Selection
  - $|\eta_{\text{clus}}| < 0.7$ (Barrel only)
  - $0.5\ \text{GeV} < E_{\text{clus}} < 100\ \text{GeV}$

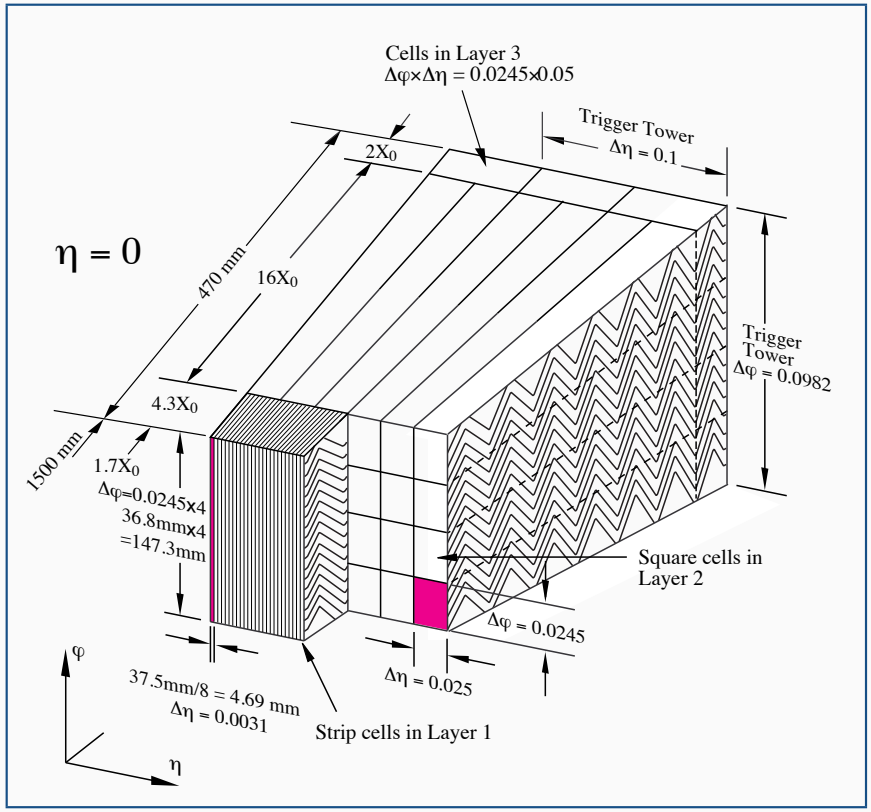
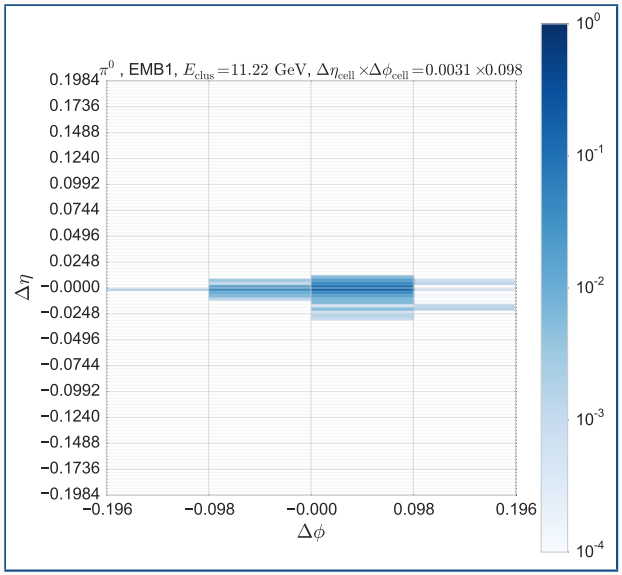


Figure 5.4: Sketch of a barrel module where the different layers are clearly visible with the ganging of electrodes in ϕ. The granularity in η and ϕ of the cells of each of the three layers and of the trigger towers is also shown.

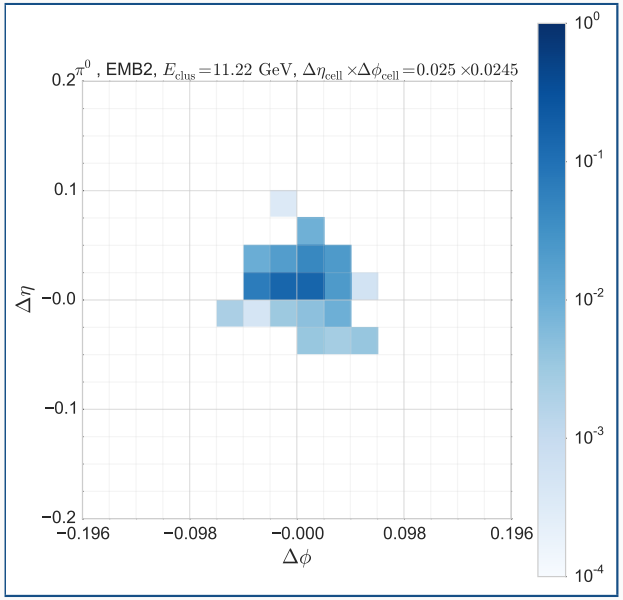
(LAr barrel module)

Preparing the Images

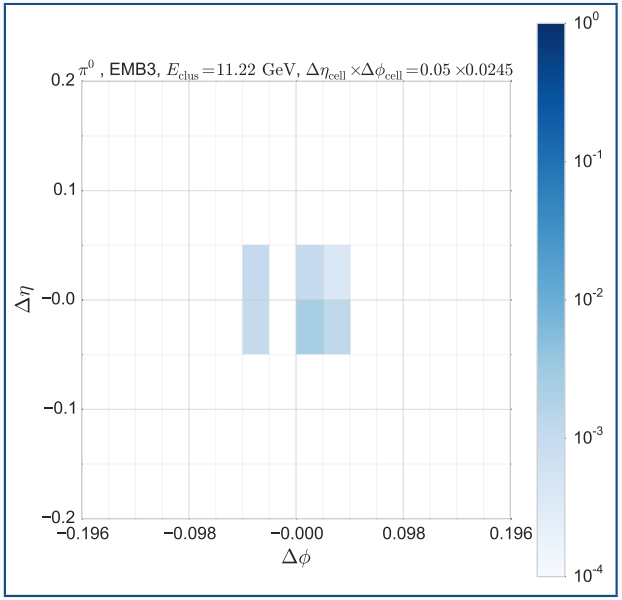
- Steps: ESD $\xrightarrow{\text{Athena}}$ TTree $\xrightarrow{\text{root\_numpy}}$ numpy arrays for Keras
- Source code: https://github.com/jmrolsson/MLTree



(a) EMB1



(b) EMB2

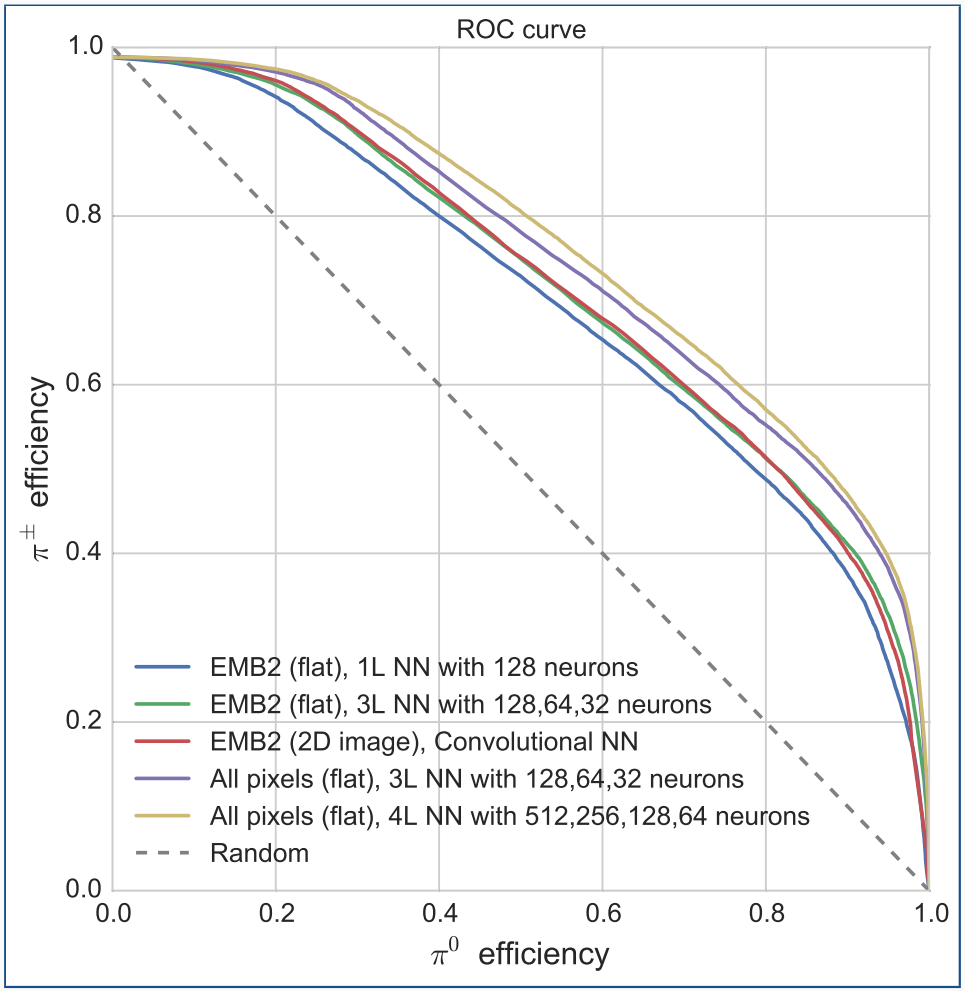


(c) EMB3

Training

- Training performed using Keras (with TensorFlow)
  - $600\,000$ (50/50 π^0 and $\pi^{\pm}$) training clusters, and $66\,000$ test clusters
- EMB2 pixels, flattened to a 1D-array + total cluster energy
  - 1 layer with 128 neurons (seems to converge after about 100 epochs)
  - 3 layers with 128, 64, and 32 neurons, respectively (takes longer to run, but converges after fewer epochs)
- Convolutional NN in EMB2, keeping the 2D information (but not including cluster energy)
  - Takes longer to train, but converges after only 3-5 epochs or so
- All pixels in all layers, flattened to a 1D-array + total cluster energy
  - 3 layers with 128, 64, and 32 neurons, respectively
  - 4 layers with 512, 256, 128, and 64 neurons, respectively

Result



- Very preliminary, but we're definitely learning some information

Conclusions and Future Plans

- We are definitely doing better than random!
  - The NN with 3 layers (128, 64, 32 neurons, respectively) does a little better than the one with just one 128 neuron layer
  - The ConvNet does as well as the simple NNs, despite not including the total cluster energy as a feature
  - Training on a feature vector of all the pixels (in all barrel layers) does improve things a little
- Next steps for cluster classification
  - Compare with what is currently used for distinguishing EM vs. HAD showers
  - Implement 2D DNN/ConvNet that takes all the (barrel) layers into account
  - Add tracking information (per layer)
  - Look at clusters in a more dense environment (inside jets, etc.)
- Main Goal: Cluster calibration using regression
  - Perform cluster calibration by training on cells associated with a cluster (using the truth cluster energy as labels)
  - How well we can do with a DNN compared to LC?

Thanks!