MACHINE LEARNING FOR BACKGROUND HIT REJECTION IN THE MU2E STRAW TRACKER

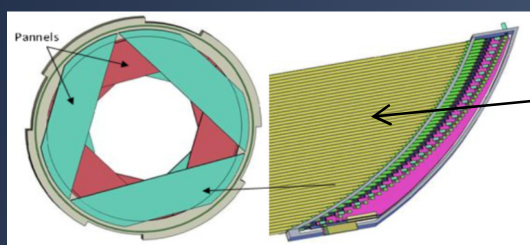
Digvijay Roy Varier *University of California, Berkeley*
Richard Bonventre *Lawrence Berkeley National Lab*

2021 Meeting of the Division of Particles and Fields of the American Physical Society

14 July 2021

Mu2e setup and objectives

Mu2e will search for Charged Lepton Flavor Violation, using a cylindrical straw tracker for detecting particles.



- Signal is a 105 MeV/c electron from CLFV conversion of muons. (we accept a window centered at 105 MeV/c, i.e. 80 MeV/c < MC truth momentum < 110 MeV/c)

- This talk is about developing a cut based on information from single straw hits to select only the signal hits, i.e. enhance the signal reconstruction by maximizing cut efficiency and signal acceptance.

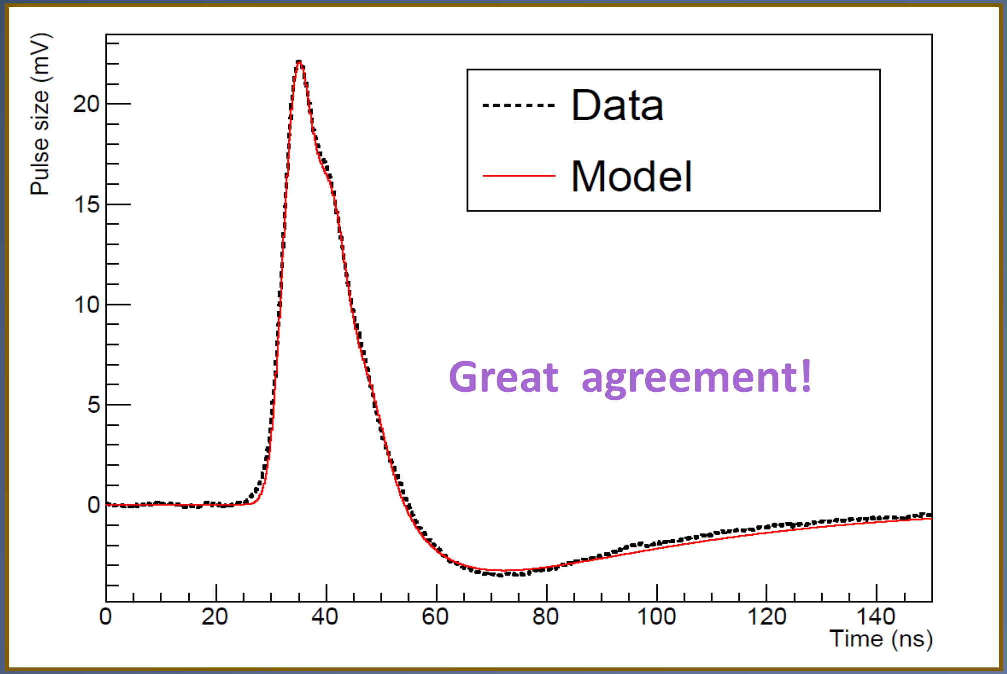
- We want to reject highly-ionizing tracker hits from other sources, like protons from nuclear capture.

Simulation of the physical processes

- We analyze simulation data that includes expected tracker measurements from conversion electron signals as well as from backgrounds (like protons).

- Geant4 simulation uses a detailed model of the Mu2e geometry, with customized simulation of ionization cluster creation, drifting & electronics response (including nonlinear contributions and channel to channel gain differences).

- Comparison of a measured pulse from the straw and the simulation output, for an Fe55 source event → →



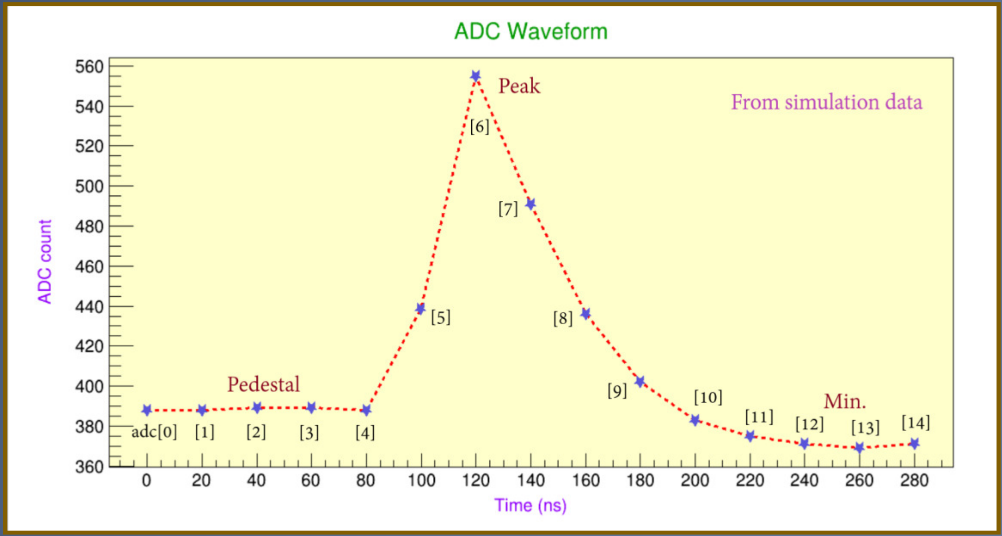
After each straw hit …

- The hit produces ionization clusters → drift to anode wire → current pulse travels as a time signal to the two ends (cal and HV) of the straw → two-sided readout determines position of the hit along the straw by comparing the two arrival times

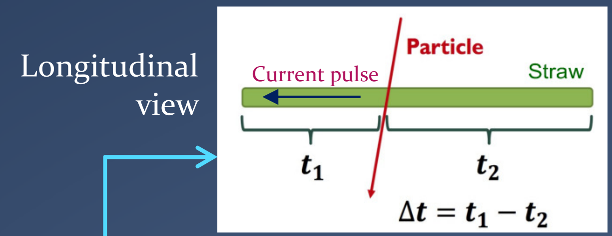
- Record duration for which time signal is over a pre-set threshold on each side

- Shaped waveform is digitized by ADCs every 20ns, over a period of ~300ns, thus giving an array : adc[15]

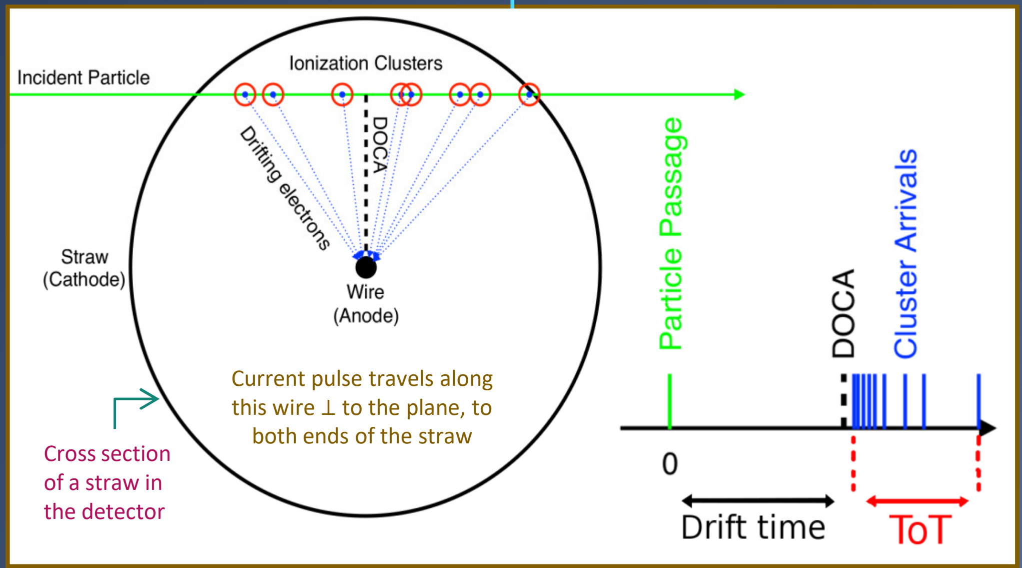
- Energy deposited in straw (edep) is estimated as peak minus pedestal of this wave - form, converted into keV



ToT : Time over Threshold



- Time for which the signal in straw stays above a threshold value
- Ranges from from 0 to 80ns, with a binning of 5ns.
- Correlated to path length of the particle through the straw (longer path → more distance b/w the clusters which are nearest and farthest from anode → larger ToT)



- So from simulations, numerically compute the path length (dx) from ToT and then get a measure of specific ionization (dE/dx) of each incident particle as edep ÷ dx.

Current Performance

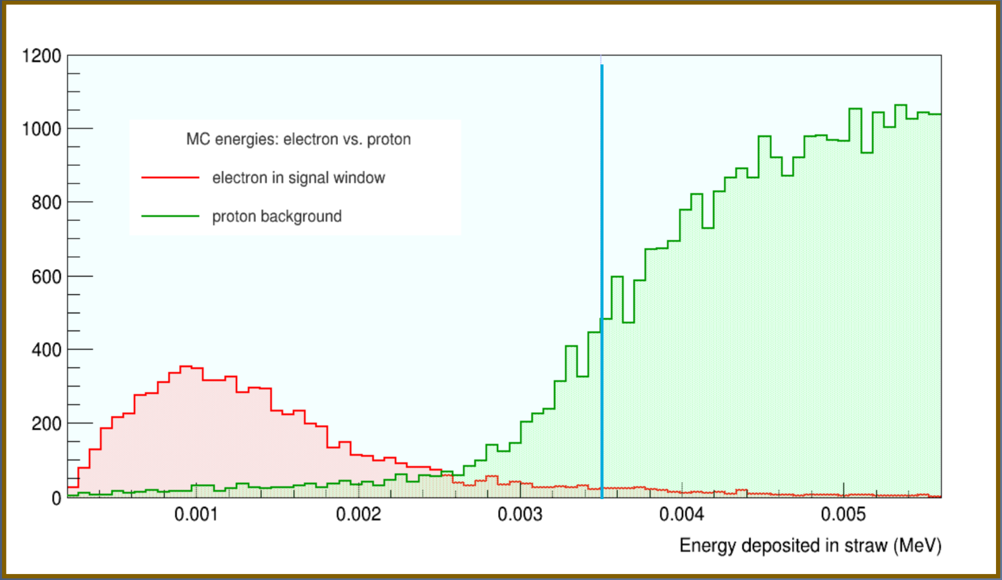
- A cut on the reconstructed energy deposited in each straw (edep) has been used so far to separate electrons and protons:

  edep < 3.5 KeV ➡ conversion electron

- Works because electrons at this energy are "minimum ionizing particles" while protons deposit a lot of energy. A part of the energy deposition distribution → →

- Testing on a simulation …
  Signal acceptance : 94.49 %
  & Proton rejection : 84.57 %

  Can be improved



TMVA

- Train a type of artificial neural network called Multi-Layer Perceptron (MLP) that outputs a value between 0 and 1. In our case an output of 0 = proton hit, and 1 = conversion electron hit with 80 < MC truth momentum < 110

- Want to separate high-quality measurement from low-quality measurements, so want output to strongly peak at 0 and 1.

- Specify desired cuts on input to tell TMVA what “signal” & “background” are. Define the input variables (in our case, edep, dE/dx, adc[15], totcal, tothv and/or calibration) for which weights are determined via the MLP method.

- TMVA automatically splits the input tree into a training and a testing sample.

Using alternative parameters to define the cut value

Some promising candidates to start with:

- **MODEL 1** : Combination of edep and dE/dx (reconstructed energy loss per distance, which is correlated to edep and ToT)
- **MODEL 2** : Combination of full ADC waveform and ToT
- **MODEL 3** : Simplify the adc information to lessen the no. of input variables in Model 2 from 17 to 5 → make the classification sequence run faster when used in the trigger, and also make the model less sensitive to small differences between simulation and real data . (Time drops by 70.5% when Model 3 is used instead of Model 2, while classifying 646215 simulated events)

  The simplified variables are Max(adc), Min(adc) and pedestal = 0.25*(adc[0]+adc[1]+adc[2]+adc[3]). Tothv and totcal are kept intact.

Continued on next slide ...

Including ADC channel calibrations

- Calibration used to convert the peak minus pedestal into units of KeV.
- Gain of each ADC channel has been varied by 20% to simulate a smearing effect that is observed in the real readout system.
- Model 4 assumes perfect calibration (i.e. the calibration factor exactly accounts for differences in gain across channels).

➢ **MODEL 4** : Try adding calibrations for each ADC channel, combined with simplified adc-tot variables in 2 different ways :-

- Model 4a : Divide peak and min(adc) by the normalized (i.e. value per mean) calibration
- Model 4b : Include calibration as extra input variable in additon to max(adc), min(adc), pedestal and ToT

Time taken is 27.7% (4a) and 28.4% (4b) of that for Model 2, when tested on the same set of 646215 events mentioned earlier.

TMVA Training Results

for the case when input variables are adc[15], totcal and tothv only

- At the end, the training and testing efficiencies are almost identical

| cut 0.01: 0.936, 0.938 |
| cut 0.10: 0.987, 0.987 |
| cut 0.30: 0.999, 0.999 |

- Looks like there has been no over-fitting

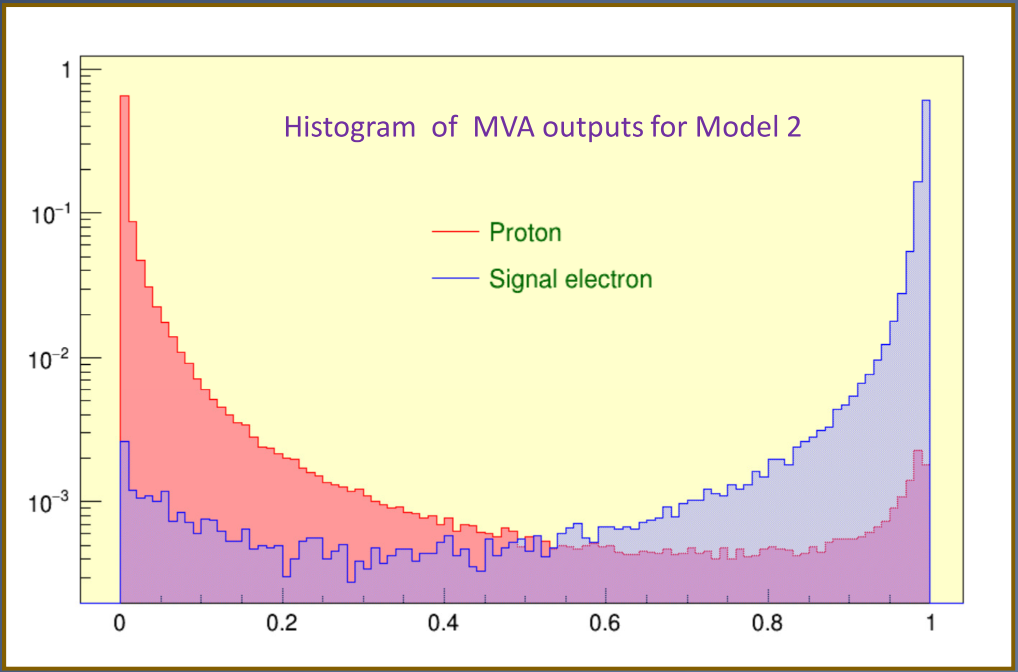
Rank	Variable	Importance
1	adc_12	10.95
2	adc_14	10.27
3	adc_11	9.433
4	adc_10	9.308
5	adc_6	8.256
6	adc_13	8.215
7	adc_7	6.541
8	adc_9	4.888
9	adc_8	3.939
10	tot_cal	3.704
11	tot_hv	2.858
12	adc_5	1.725
13	adc_2	1.664
14	adc_3	1.519
15	adc_0	1.484
16	adc_4	1.100
17	adc_1	0.985

Apply TMVA Classification

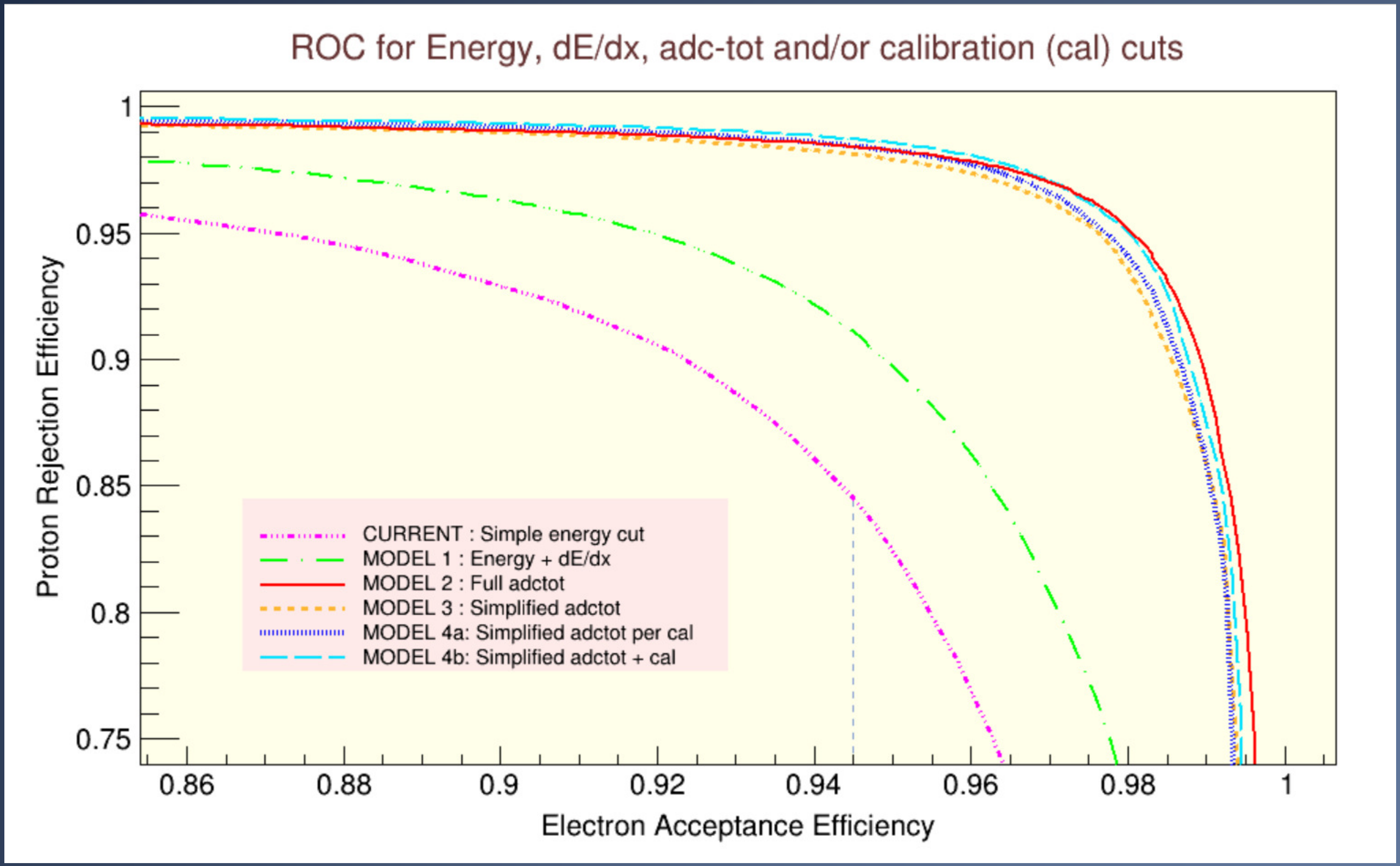
- Set up script to compute MVA outputs (0-1) of all events in a test file, using the weights; the distribution of these outputs form a histogram

- For each position of cut, area to left gives p+ rejection and area to right gives signal e- acceptance; we printed these to identify the most optimum cuts

- ROC plotted & overlaid in slide 13 for all the tested models.



Efficiency/acceptance curves (ROCs)



Conclusions

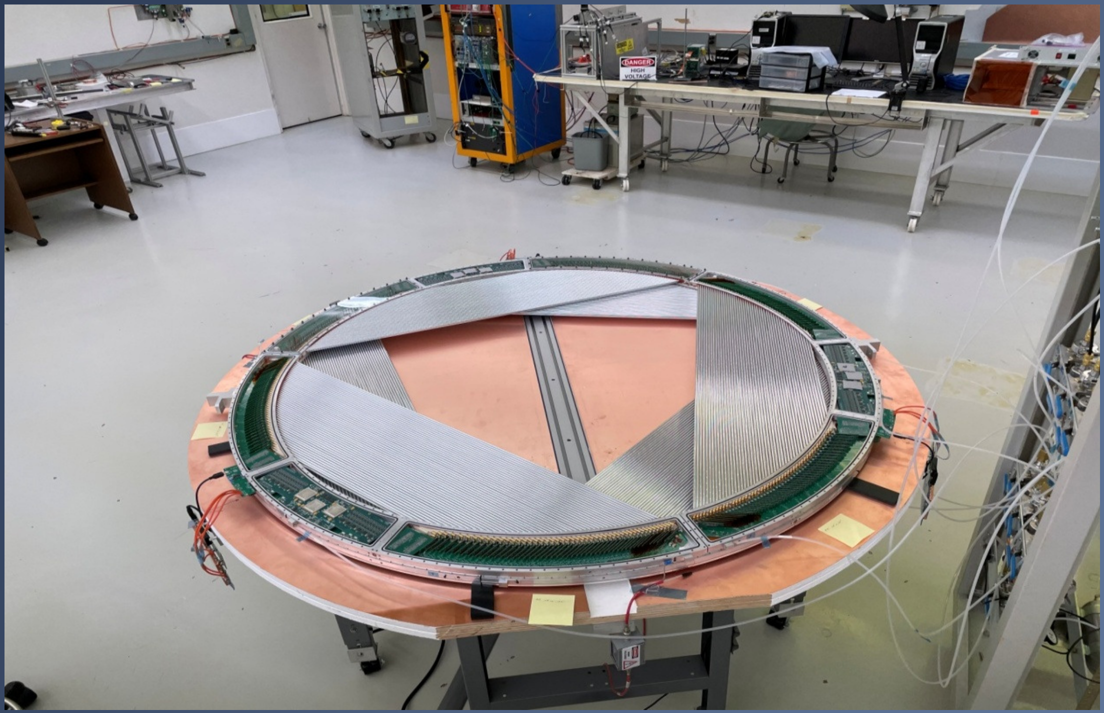
- Best performance when the full ADC waveform is combined with ToT readings (Model 2)

** Tested on a random ** simulation dataset	Current Model edep cut = 3.5 keV	Model 2 Full adc-tot cut = 0.55
Signal Acceptance	94.49 %	96.99 %
Proton Rejection	84.57 %	97.95 %

- The improvement is significant , so for now we have implemented this MVA in the official Mu2e Straw Hit Reconstruction module (had been using energy cuts so far)

- Good news : Simplifying adc inputs in TMVA does not depreciate the performance by much, and inclusion of calibrations improves those models. In fact, adding channel calibrations as a separate TMVA input variable along with the simplified adc waveform and tot (Model 4b) does almost as good as Model 2

Future work: “Real data” from Vertical Slice Test --- 6 production quality panels (~600 channels) operated together. We will use this extensive setup to perform a more detailed comparison between cosmic ray signals and our simulations.



Thank You