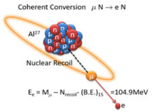
Machine Learning for Background Hit Rejection in the Mu2e Straw Tracker

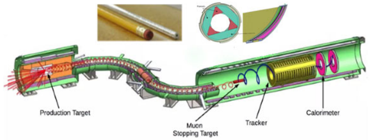
Digvijay Roy Varier (University of California, Berkeley)
Richard Bonventre (Lawrence Berkeley National Lab)

The Mu2e Experiment at Fermilab

- Search for evidence of Charged Lepton Flavor Violation via the neutrino-less conversion of a muon to a mono-energetic electron, while in orbit around an Aluminium nucleus



- Signal is a 105 MeV/c electron, detected using a cylindrical straw tracker and electromagnetic calorimeter



- We enhance signal reconstruction by developing a cut, based on information from single straw hits, to select only the signal hits and reject highly-ionizing hits from sources like protons from muon nuclear capture.

Simulation of the physical processes

We analyze Geant4 simulations of the tracker measurements. Simulation takes into account :

- Detailed Mu2e geometry
- Ionization clustering effects
- Dispersion, reflection
- Saturation, attenuation
- Gain differences between the fifteen ADC channels

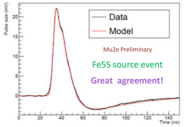


Figure 1 : Comparison between a measured pulse from the straw, and the simulation output

After each straw hit …

- (Fig. 2) Hit produces ionization clusters → drift to anode wire → current pulse travels as a time signal to the 2 ends (cal and HV) of the straw → readout determines position of the hit along the straw by comparing the two arrival times.
- Record Time over Threshold (ToT) for the signal on each side (may range from 0 to 80ns, with a binning of 5ns).

- A shaped waveform (Fig. 2) is digitized by ADCs every 20ns, over a period of ~300ns; giving an array (named *adc[15]*)
- Energy deposited in the straw (edep) by each hit is reconstructed as the peak minus pedestal of the digitized waveform
- Longer path of particle inside straw → larger ToT
- So from simulations, we can numerically compute the path length (dx) from ToT, and hence estimate the specific ionization (dE/dx) as edep ÷ dx.

Figure 2 : Example ADC Waveform

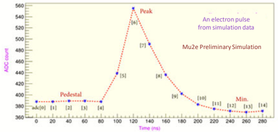
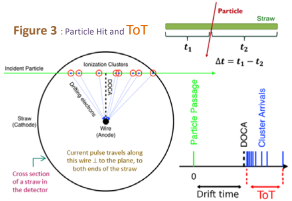


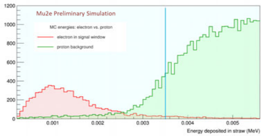
Figure 3 : Particle Hit and ToT



Current Performance

- A cut on edep has been in use to separate electrons and protons.

Figure 4 : A part of the energy deposition distribution



- e- at this energy are minimum ionizing particles, while p+ deposit a lot of energy (Fig.4); edep < 3.5 KeV implies a hit by a conversion electron.
- For this simulated dataset :
  Signal acceptance : 94.49 %
  Proton rejection : 84.57 %

Can be improved

TMVA (Toolkit for Multivariate Data Analysis with ROOT)

- Train an artificial neural network called Multi-Layer Perceptron (MLP) that outputs a value between 0 and 1 based on the specified criteria for "signal" and "background"
- 0 = proton hit, and 1 = electron hit in the signal window 80 < MC truth momentum (MeV/c) < 110. For clear separation, we want the distribution of outputs to strongly peak at 0 and 1.
- Define the input variables (in our case, edep, dE/dx, adc[15], totcal, tothv and/or calibration) for which the MLP determines suitable weights.

Using alternative parameters to define the cut value

Some promising candidates to start with:

- **MODEL 1** : Combination of edep and dE/dx
- **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 .

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

Figure 5 : Effect of variables on TMVA classification using Model 2

- ❖ Classification sequence will run faster when used in the trigger (Model 3 vs Model 2 on 646215 simulated events → Time drops by 70.5%)
- ❖ Model will be less sensitive to small differences between simulation and real data.
- ❖ Referring to Fig. 2 & 5, the agglomerating inputs suggest the representative variables :- Max(adc), Min(adc), tothv, totcal, and pedestal = 0.25*(adc[0]+adc[1]+adc[2]+adc[3]).

Including ADC channel calibrations

- Simulation varies the gain of each ADC channel by 20% to account for a smearing observed in the real readout system.
- Calibration is used to convert the peak minus pedestal into units of keV. Model 4 assumes that the calibration factor exactly accounts for differences in gain across all channels
- **MODEL 4** : The set of calibrations for each ADC channel was combined with the simplified Model 3 in two 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

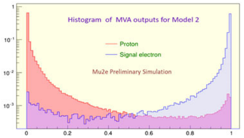
Time taken is 27.7% (4a) and 28.4% (4b) of that for Model 2, (same 646215 events as before).

Apply TMVA Classification

- TMVA weights used to compute MVA outputs (0-1) of all events in a test file → identify the cut-value that maximizes p+ rejection efficiency and signal e- acceptance ratio (e.g. it is 0.55 for Model 2).
- We plotted these pairs of numbers to get cut efficiency vs. acceptance curves (ROC); overlaid six curves to compare the performance of the current model and the five proposed models (Fig. 7).

Model 2 achieves best separation of protons and conversion electrons, thus we have implemented this MVA in the official Mu2e Straw Hit Reconstruction module, replacing the energy-only cuts that had been in use so far.

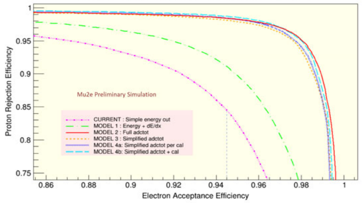
Figure 6 : Separation based on weighted input variables



Good result : Simplifying adc inputs in TMVA does not significantly decrease the performance. In fact, after the inclusion of ADC channel calibrations, Model 4b does almost as well as Model 2, with the added promise of faster run-time and better adaption when used on real data (due to greater degrees of freedom).

** 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 %

Figure 7 : Looks like we should implement Model 4b eventually
ROC for Energy, dE/dx, adc-tot and/or calibration (cal) cuts



Acknowledgements

We are grateful for the vital contributions of the Fermilab staff and the technical staff of the participating institutions. This work was supported by the US Department of Energy; the Italian Istituto Nazionale di Fisica Nucleare; the Science and Technology Facilities Council, UK; the Ministry of Education and Science of the Russian Federation; the US National Science Foundation; the Thousand Talents Plan of China; the Helmholtz Association of Germany; and the EU Horizon 2020 Research and Innovation Program under the Marie Sklodowska-Curie Grant Agreement No.690835. This document was prepared by members of the Mu2e Collaboration using the resources of the Fermi National Accelerator Laboratory (Fermilab), a U.S. Department of Energy, Office of Science, HEP User Facility. Fermilab is managed by Fermi Research Alliance, LLC (FRA), acting under Contract No. DE-AC02-07CH11359.