

TMVA, student projects #3 & 4



Jefferson Lab

Friday July 2, 2021

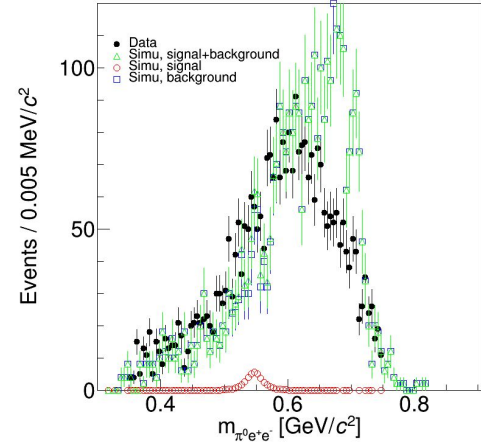
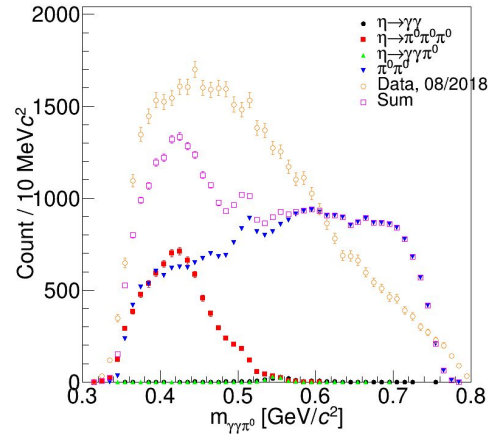
Background suppression via TMVA

Process: $\gamma p \rightarrow \eta p$

- $\eta \rightarrow 2\gamma\pi^0 \rightarrow 4\gamma$ (Chase)
- $\eta \rightarrow 2\gamma e^- e^+$ (Joanna)

Three sets of variables:

- Baseline selection
- Input variables
- Spectator variables



Baseline selection

Process: $\gamma p \rightarrow \eta p$

- 4 γp (Chase)
- 2 γ 2ep (Joanna)

- Particle ID
- χ^2
- Extra energy
- Unused tracks
- Missing mass
- Coplanarity
- Vertex: R & z

Naively, I suspect that MVA-algorithm just looks at one variable that is drastically different between signal and background (say χ^2), and this makes a decision. In this case, would it may be have sense to pre-select events on this variable, so that algorithm will try to discriminate better for the “signal-like” region.

This list is not definitive, we will test different combination of baseline selection & input variables

Input variables

Which ones?

- Number of insert clusters
- Mandelstam variable $t = (P4[\text{proton at rest}] - P4[\text{recoil proton}])$
- Helicity, look at polar angle distribution of decay products in the η frame
- Transverse momentum
- ...

This list is not definitive, we will test different combination of baseline selection & input variables

Spectator variable is the invariant mass: 4γ (Chase) and $2\gamma 2e$ (Joanna)

Method

- Cut-based analysis
- Fisher discriminant (fastest method, but the separation is not as good in general cases)
- (Boosted) Decision Tree (good separation in general, fast training speed compare to neural network)
- Deep neural network
 - MLP (multi-layer perceptron)
 - CNN(convolutional neural network)
 - etc.

I am building test trees based on MC simulation of the signal and background for an initial study