Progress Update

Trevor Towstego
UofT Neutrino/DM Meeting
September 12, 2018

TMVA memory issue resolved

- I was creating new branches for every event in the loop
  - fixed by moving the branch creating to the initialization phase before entering into the event loop
- Memory requirements are no longer an issue
- Previously, had split the TMVA input trees into 10 different files
  - this was because of previous memory issues when filling the trees
  - now only using 1 file, which results in a different splitting of events into training/testing
    - FOM results are a bit different due to statistical fluctuations, but I performed consistency checks to ensure that when training/testing events are split consistently, FOM remains the same
      - Brings in question (again) of error in FOM value

Statistical Error in FOM

- Is this approach correct?
  - uncertainty in signal = (sum of squared weights)$^{1/2}$
  - uncertainty in background = (sum of squared weights)$^{1/2}$
  - uncertainty in FOM is given by:

$$\delta \mathrm{FOM} = \sqrt{\left(\frac{\partial \mathrm{FOM}}{\partial S}\delta S\right)^2 + \left(\frac{\partial \mathrm{FOM}}{\partial B}\delta B\right)^2}$$

- I've estimated the error to be ~0.02 to 0.03, but have yet to do calculations with actual event weights (took the average weight, which underestimates the uncertainty)

BDT v0 trial 9 results

- Preliminary cuts:
  - FCFV
  - not 1Re/1Re1de
  - 1/2 sub-events
    - separate samples
  - $E_{rec}(1e,1\pi) < 1.5$ GeV

v0 trial 9 (NTrees=1500, MaxDepth=5, MinNodeSize=0.5, nCuts=50)					
	Signal	**Bkgd**	**Purity**	**Eff**	**FOM**
0 decay e	0.58	0.60	49.4%	30.6%	0.536
1 decay e	2.00	0.83	70.5%	57.9%	1.187
v1 trial 8 (NTrees = 1500, MaxDepth=5, MinNodeSize=0.5, nCuts=50)					
	Signal	**Bkgd**	**Purity**	**Eff**	**FOM**
0 decay e	0.53	0.25	68.1%	27.7%	0.598
1 decay e	2.16	1.01	68.3%	62.7%	1.215

	BDT variables								
	1R v 1R nll	1R v 2R nll	2R v 2R nll	2R v 3R nll	3R v 3R nll	3R v 4R nll	1R+2R kinematics	E_{rec}, towall e, towall π, p_{low}, $m_{\pi 0}$, (*d2se*)	1R+2R+ 3R fit indices
Trial 8	■	■	■	■	■		■	■	■
Trial 9	■	■	■	■	■	■	■	■	■

Starting to look into systematics

- Going through Sophie's slides from her last plenary talk
  - neutrino pion production error
    - compare kinematics from different generators (NEUT, NUWRO, GENIE, alt. GENIE)
  - multi-pion production error
    - compare different multi-π production/DIS models (NEUT, varying pion multiplicity in NEUT, deuterium fits, AGKY model in GENIE)
  - pion hadronic interaction error (FSI and secondary)
    - use fit to external pion scattering data in TN325, and vary pion hadronic interaction cross sections
- Is this a good starting point?
  - Still need to address detector systematics, effect of near detector constraint?