Investigation of photon identification, continuing from Will's work:
BCAL Reconstruction Update (W. Levine)

Separation of real photons and background events can be approached numerous ways. Placing hard cuts on some key reconstructed variables was previously investigated. We are looking into using a Boosted Decision Tree for comparison.

Looking for similar efficiency with higher background rejection and signal purity.

Based on current shower reconstruction code. May be possible to improve code using what we learn from the BDT or other studies.

Photon_discrim plugin: Uses truth and reconstructed information to identify photons – used to train the BDT with reconstructed variables.

Bcal_pi0 plugin: keeps particle combinations with two or more neutral particles in the final state. Runs each neutral particle through photon_discrim to get reconstructed variables for each. Each set of reconstructed variables is evaluated using the trained BDT. If both neutral particles are signal-like and the invariant mass of the two is similar to the pi0 mass, count this combination as a pi0 event.

Pre-training cut: z-cluster position < 400

Training done in eight groups, each corresponding to a range in theta for photons.

For a given theta range, train the BDT with bggen data, identifying photons with the photon_discrim plugin.

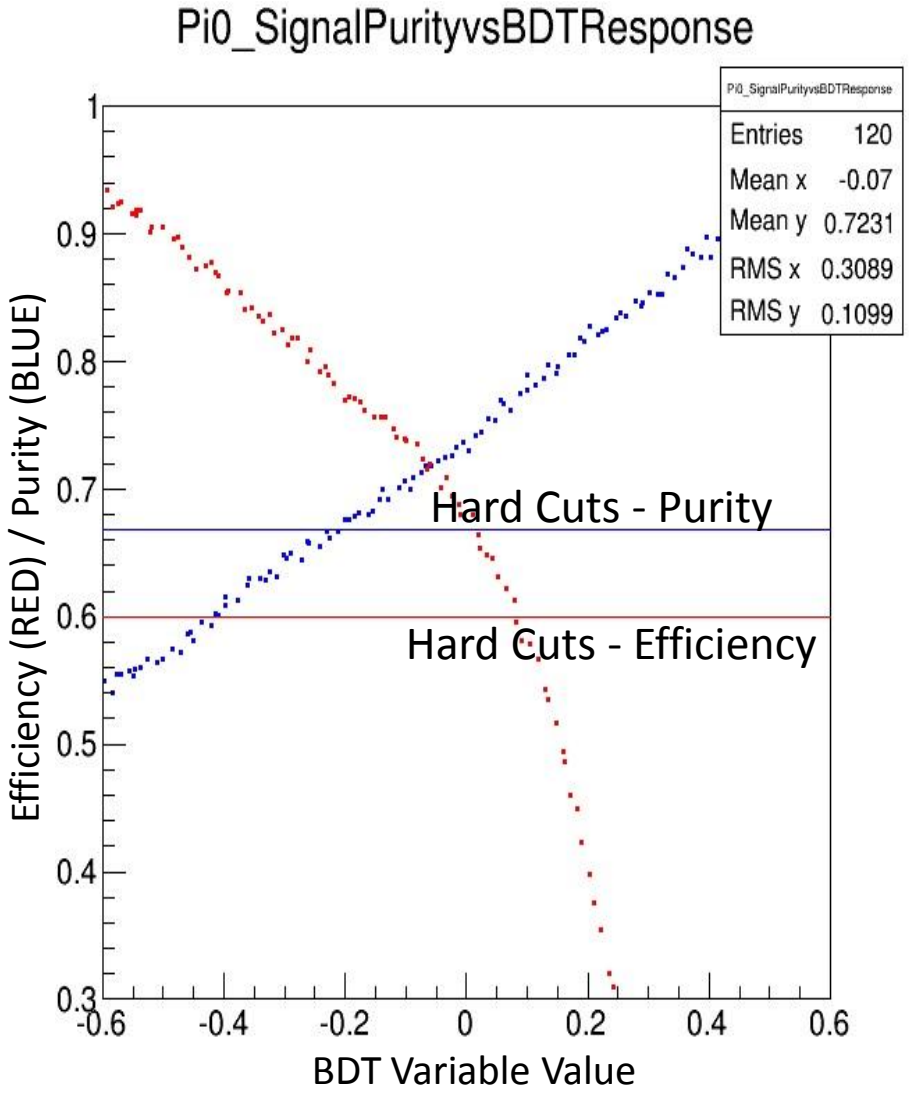
Run the bcal_pi0 plugin using the corresponding training for each theta range to extract BDT value cuts for 60%, 70%, 80%, and 90% efficiencies.

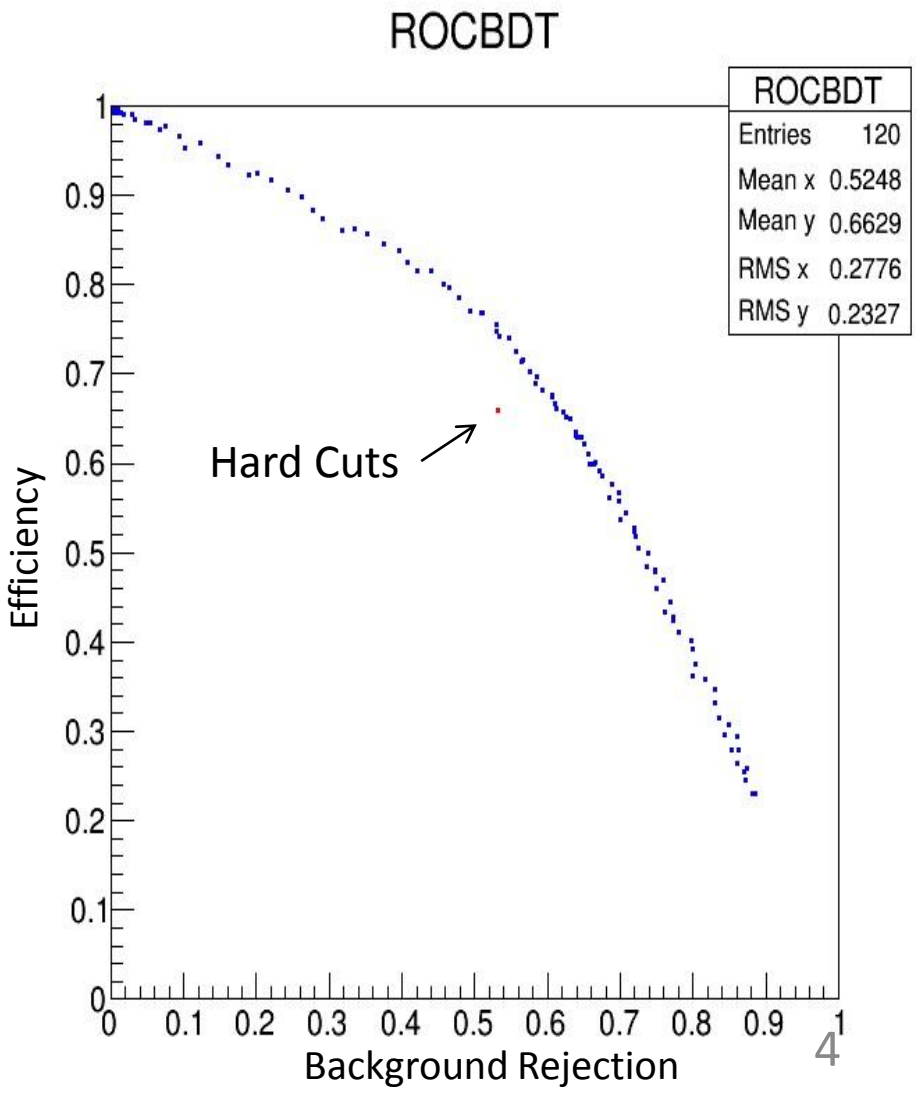
Input these value cuts for each theta range and run bcal_pi0 with all events, still using the appropriate training for each range.

This allows the BDT to work with theta dependencies implicitly.

Global Cut: z-cluster Position < 400

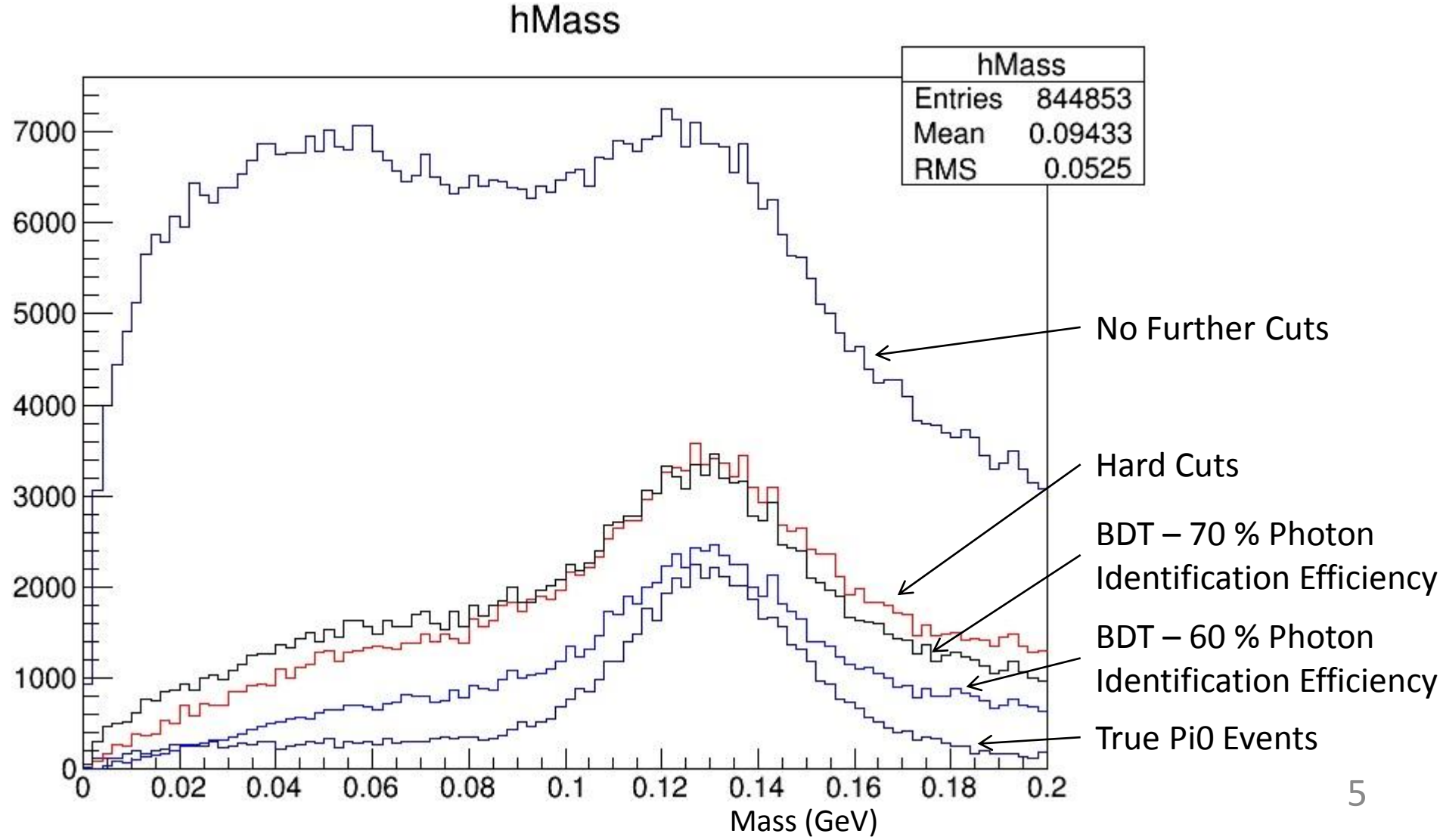
Hard Cuts: r <= 74, closest track dist >= 10, closest shower dist >= 13, tproj <= 1, Points > 1





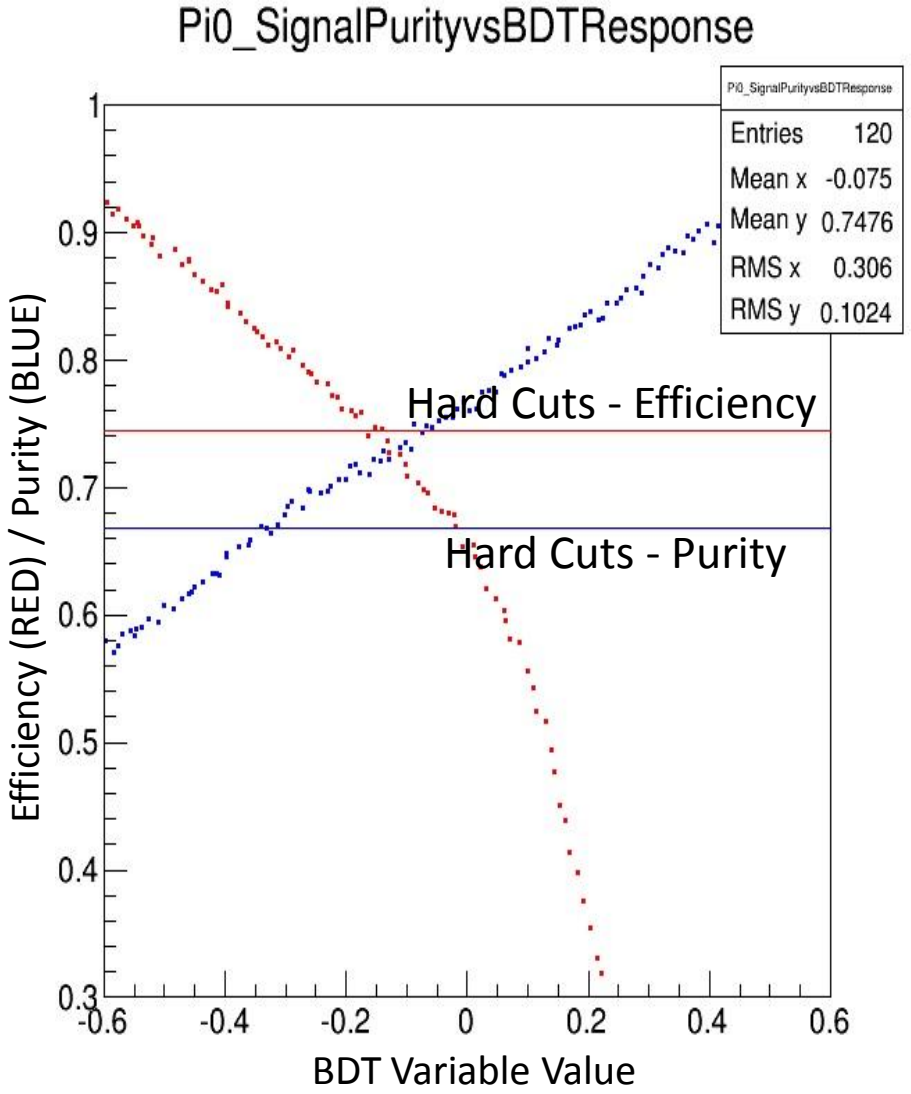
Global Cut: z-cluster Position < 400

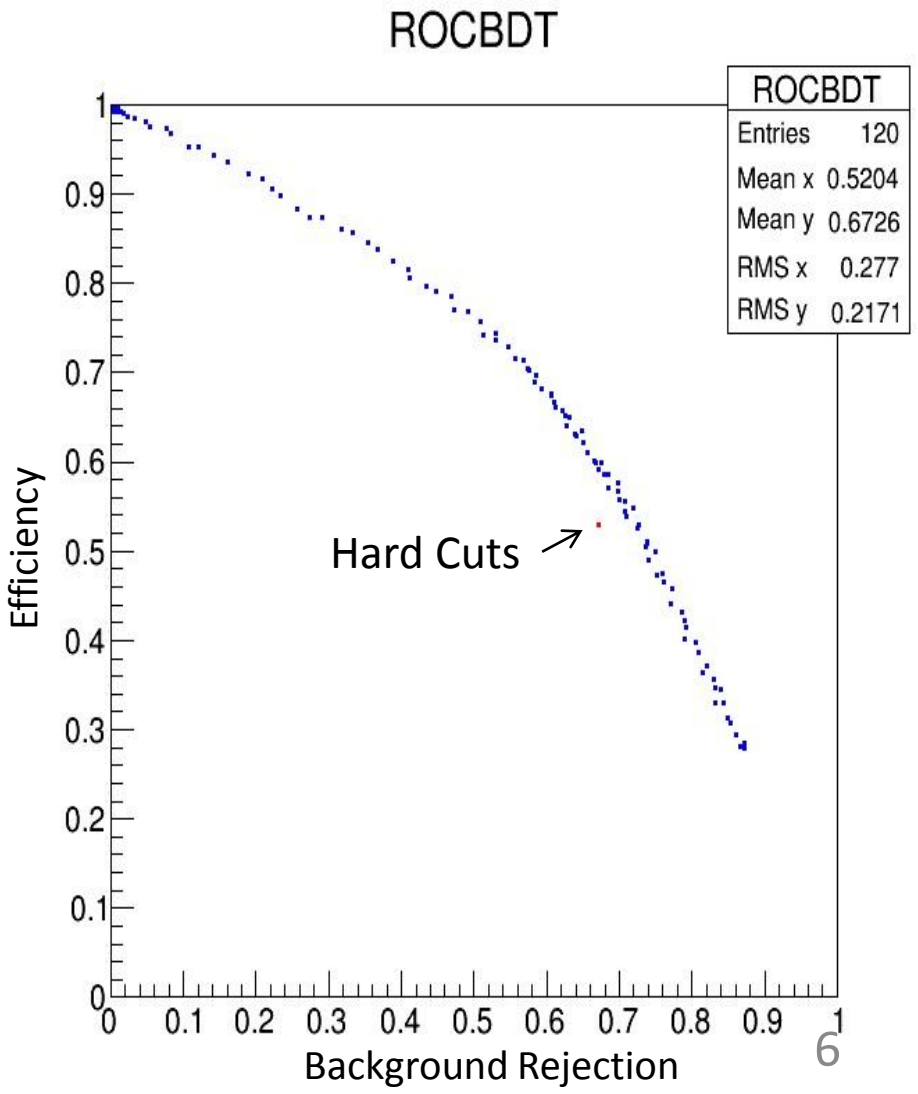
Hard Cuts: r <= 74, closest track dist >= 10, closest shower dist >= 13, tproj <= 1, Points > 1



Global Cut: Points > 1 and z-cluster Position < 400

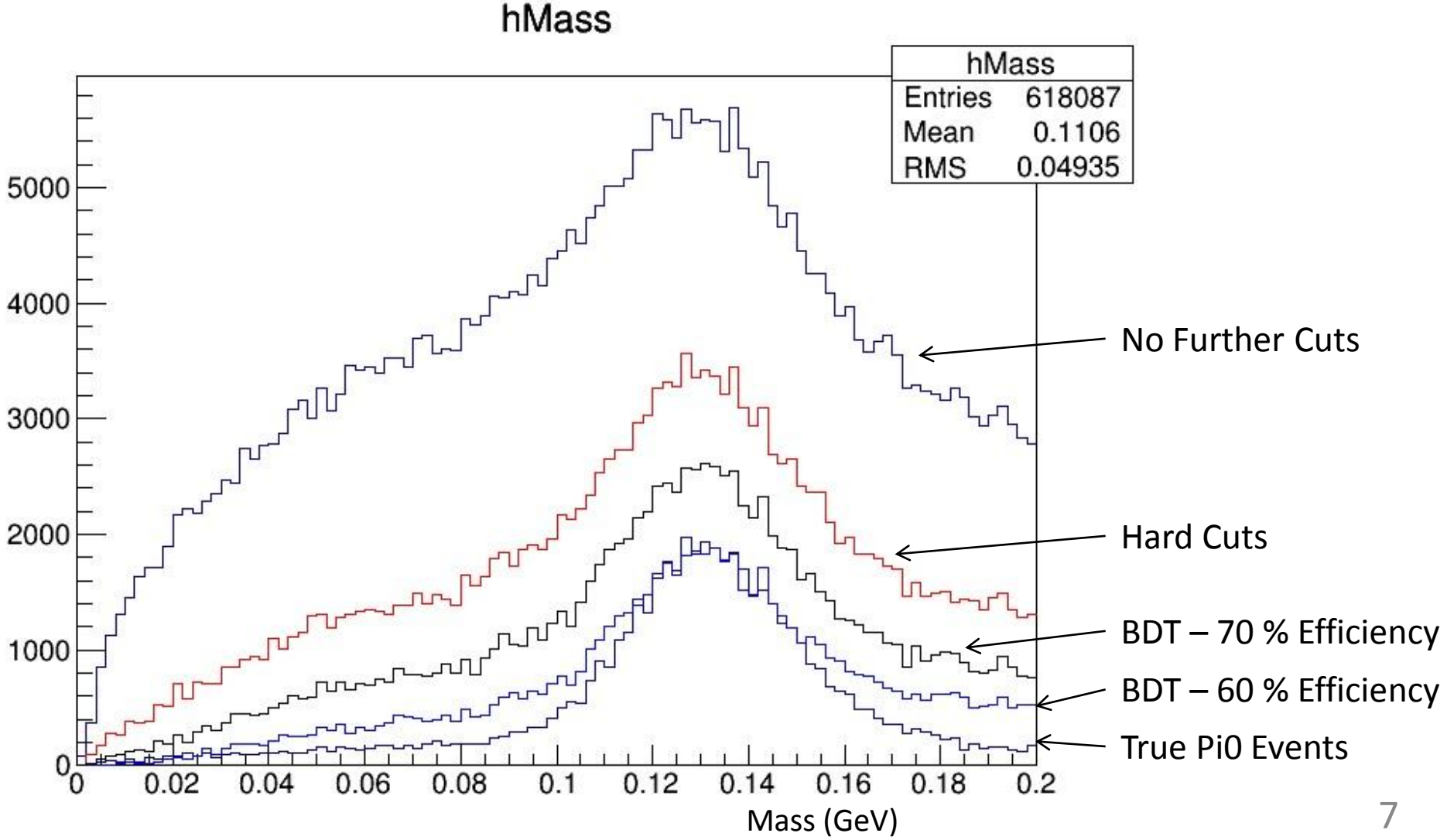
Hard Cuts: r <= 74, closest track dist >= 10, closest shower dist >= 13, tproj <= 1





Global Cut: Points > 1 and z-cluster Position < 400

Hard Cuts: r <= 74, closest track dist >= 10, closest shower dist >= 13, tproj <= 1



Looking in to identifying the types of events which 'fool' the BDT.

Possibly use DBCALTruthShowers or DBCALIncidentParticles to try to identify what shower or particle is closest to the reconstructed shower to determine the event type?

DBCALTruthShowers reports multiple showers for most single-photon events. Do we know why this is the case? Only look at 'primary' showers?

DBCALTruthShower:

ptype:	track:	primary:	phi:	r:	z:	t:	p:	E:
1	1	1	-2.986	65.042	47.5	2.246	1.000	1.000
3	3	0	-2.986	66.626	47.1	2.301	0.651	0.651
1	3	0	-2.986	66.995	47.0	2.314	0.001	0.001
1	3	0	-2.986	66.958	47.0	2.313	0.003	0.003
1	3	0	-2.986	66.825	47.1	2.308	0.001	0.001
1	3	0	-2.986	66.798	47.1	2.307	0.004	0.004
1	3	0	-2.986	66.711	47.1	2.304	0.244	0.244
2	2	0	-2.986	66.626	47.1	2.301	0.349	0.349
1	2	0	-2.986	66.781	47.1	2.306	0.105	0.105
1	2	0	-2.986	66.735	47.1	2.305	0.014	0.014

Variables used in BDT Training

recon_BCAL_r	R-position of shower
recon_BCAL_tproj	Shower time minus shower position / c
recon_BCAL_E	Shower energy
recon_BCAL_closest_shower_dist	Distance from shower to closest other shower
recon_BCAL_closest_shower_E	That shower's energy
recon_BCAL_closest_track_dist	Closest track, found with 'DistToRTwithTime()'
recon_BCAL_closest_track_inter	Same, found with 'GetIntersectionWithRadius()'
recon_BCAL_closest_track_inter_delz	Difference of shower z with closest track z
recon_BCAL_closest_track_inter_delphi	Difference of shower phi with closest track phi
recon_BCAL_Layer1_E	Sum of energies of hits in layer 1 of the BCAL
recon_BCAL_Layer2_E	Same for layer 2
recon_BCAL_Layer3_E	Same for layer 3
recon_BCAL_Layer4_E	Same for layer 4
recon_BCAL_Layer4bySumLayers_E	Layer 4 energy divided by sum of layer energies
recon_BCAL_Layer3bySumLayers_E	Same for layer 3
recon_BCAL_Layer2bySumLayers_E	Same for layer 2
recon_BCAL_Layer1bySumLayers_E	Same for layer 1
recon_BCAL_ZWidth	Energy-weighted second moments for Z, R, Phi, and Theta, found with 'Σ (E_i * Δx_i^2) / Σ E_i,' for sums over points. Gives information about shape of shower
recon_BCAL_RWidth	
recon_BCAL_PhiWidth	
recon_BCAL_ThetaWidth	
recon_BCAL_Points	Number of points in shower

Variable Ranking – Theta < 20 degrees – Global Points Cut

```
--- BDT       : ------------------------------------------------------------------------------------------------------------
--- BDT       : Rank : Variable                                    : Variable Importance
--- BDT       : ------------------------------------------------------------------------------------------------------------
--- BDT       :    1 : recon_BCAL_tproj                            : 1.599e-001
--- BDT       :    2 : recon_BCAL_r                                : 1.216e-001
--- BDT       :    3 : recon_BCAL_closest_track_dist               : 1.137e-001
--- BDT       :    4 : recon_BCAL_closest_track_inter              : 9.051e-002
--- BDT       :    5 : recon_BCAL_Layer4_E                         : 8.705e-002
--- BDT       :    6 : recon_BCAL_Layer1bySumLayers_E              : 6.333e-002
--- BDT       :    7 : recon_BCAL_Layer4bySumLayers_E              : 5.392e-002
--- BDT       :    8 : recon_BCAL_closest_shower_dist              : 4.134e-002
--- BDT       :    9 : recon_BCAL_E                                : 2.950e-002
--- BDT       :   10 : recon_BCAL_PhiWidth                         : 2.853e-002
--- BDT       :   11 : recon_BCAL_RWidth                           : 2.803e-002
--- BDT       :   12 : recon_BCAL_Layer1_E                         : 2.750e-002
--- BDT       :   13 : recon_BCAL_ZWidth                           : 2.709e-002
--- BDT       :   14 : recon_BCAL_Layer3bySumLayers_E              : 2.418e-002
--- BDT       :   15 : recon_BCAL_Layer2bySumLayers_E              : 2.282e-002
--- BDT       :   16 : recon_BCAL_Points                           : 1.912e-002
--- BDT       :   17 : recon_BCAL_closest_shower_E                 : 1.333e-002
--- BDT       :   18 : recon_BCAL_Layer2_E                         : 1.332e-002
--- BDT       :   19 : recon_BCAL_Layer3_E                         : 1.250e-002
--- BDT       :   20 : recon_BCAL_closest_track_inter_delphi       : 1.073e-002
--- BDT       :   21 : recon_BCAL_closest_track_inter_delz         : 6.690e-003
--- BDT       :   22 : recon_BCAL_ThetaWidth                       : 5.365e-003
--- BDT       : ------------------------------------------------------------------------------------------------------------
```

Variable Ranking – 35 < Theta < 50 degrees – Global Points Cut

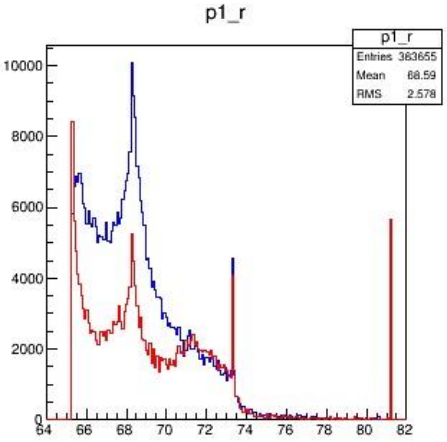
```
--- BDT       : ------------------------------------------------------------------------------------------------------------
--- BDT       : Rank : Variable                                    : Variable Importance
--- BDT       : ------------------------------------------------------------------------------------------------------------
--- BDT       :    1 : recon_BCAL_closest_track_dist               : 2.372e-001
--- BDT       :    2 : recon_BCAL_closest_track_inter              : 1.408e-001
--- BDT       :    3 : recon_BCAL_tproj                            : 1.313e-001
--- BDT       :    4 : recon_BCAL_r                                : 5.621e-002
--- BDT       :    5 : recon_BCAL_PhiWidth                         : 4.288e-002
--- BDT       :    6 : recon_BCAL_E                                : 3.599e-002
--- BDT       :    7 : recon_BCAL_RWidth                           : 3.430e-002
--- BDT       :    8 : recon_BCAL_Layer4bySumLayers_E              : 3.123e-002
--- BDT       :    9 : recon_BCAL_Layer2bySumLayers_E              : 3.073e-002
--- BDT       :   10 : recon_BCAL_closest_track_inter_delphi       : 2.853e-002
--- BDT       :   11 : recon_BCAL_ThetaWidth                       : 2.546e-002
--- BDT       :   12 : recon_BCAL_Layer1bySumLayers_E              : 2.450e-002
--- BDT       :   13 : recon_BCAL_Layer2_E                         : 2.373e-002
--- BDT       :   14 : recon_BCAL_closest_shower_dist              : 2.321e-002
--- BDT       :   15 : recon_BCAL_Layer1_E                         : 2.249e-002
--- BDT       :   16 : recon_BCAL_Layer3_E                         : 1.945e-002
--- BDT       :   17 : recon_BCAL_closest_shower_E                 : 1.872e-002
--- BDT       :   18 : recon_BCAL_ZWidth                           : 1.779e-002
--- BDT       :   19 : recon_BCAL_Layer3bySumLayers_E              : 1.658e-002
--- BDT       :   20 : recon_BCAL_Points                           : 1.580e-002
--- BDT       :   21 : recon_BCAL_closest_track_inter_delz         : 1.247e-002
--- BDT       :   22 : recon_BCAL_Layer4_E                         : 1.071e-002
--- BDT       : ------------------------------------------------------------------------------------------------------------
```

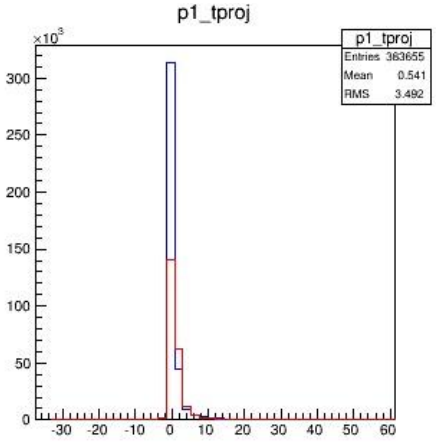
Variable Ranking – 80 < Theta < 95 degrees – Global Points Cut

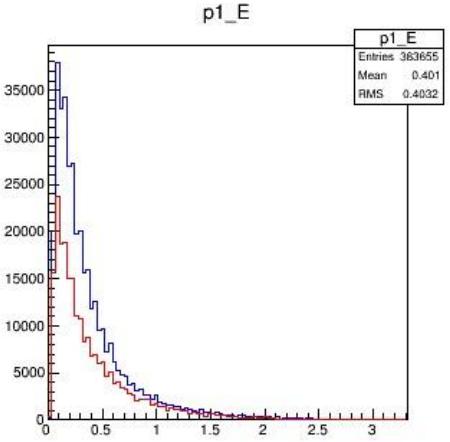
--- BDT	: Rank : Variable	: Variable Importance
--- BDT	: 1 : recon_BCAL_closest_track_inter	: 1.183e-001
--- BDT	: 2 : recon_BCAL_closest_track_dist	: 8.720e-002
--- BDT	: 3 : recon_BCAL_tproj	: 7.518e-002
--- BDT	: 4 : recon_BCAL_Layer2_E	: 5.667e-002
--- BDT	: 5 : recon_BCAL_E	: 5.621e-002
--- BDT	: 6 : recon_BCAL_closest_track_inter_delphi	: 4.680e-002
--- BDT	: 7 : recon_BCAL_r	: 4.653e-002
--- BDT	: 8 : recon_BCAL_closest_track_inter_delz	: 4.620e-002
--- BDT	: 9 : recon_BCAL_Layer1bySumLayers_E	: 4.441e-002
--- BDT	: 10 : recon_BCAL_ThetaWidth	: 4.280e-002
--- BDT	: 11 : recon_BCAL_PhiWidth	: 4.047e-002
--- BDT	: 12 : recon_BCAL_Layer3_E	: 3.912e-002
--- BDT	: 13 : recon_BCAL_Layer3bySumLayers_E	: 3.910e-002
--- BDT	: 14 : recon_BCAL_closest_shower_E	: 3.893e-002
--- BDT	: 15 : recon_BCAL_closest_shower_dist	: 3.795e-002
--- BDT	: 16 : recon_BCAL_Layer1_E	: 3.417e-002
--- BDT	: 17 : recon_BCAL_RWidth	: 3.355e-002
--- BDT	: 18 : recon_BCAL_Layer2bySumLayers_E	: 3.099e-002
--- BDT	: 19 : recon_BCAL_Layer4bySumLayers_E	: 2.900e-002
--- BDT	: 20 : recon_BCAL_Layer4_E	: 2.680e-002
--- BDT	: 21 : recon_BCAL_Points	: 2.218e-002
--- BDT	: 22 : recon_BCAL_ZWidth	: 7.384e-003

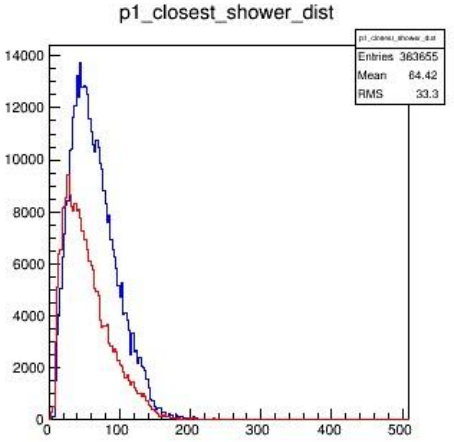
BDT Input Variables with Global Points Cut

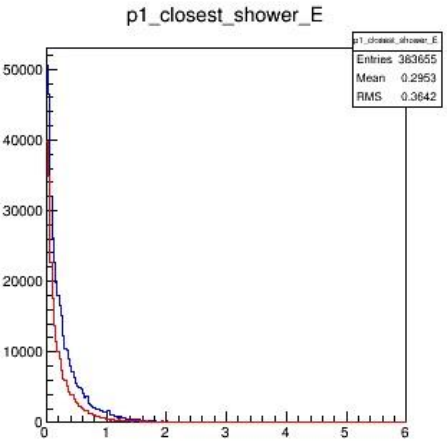
Photon 1 Variables

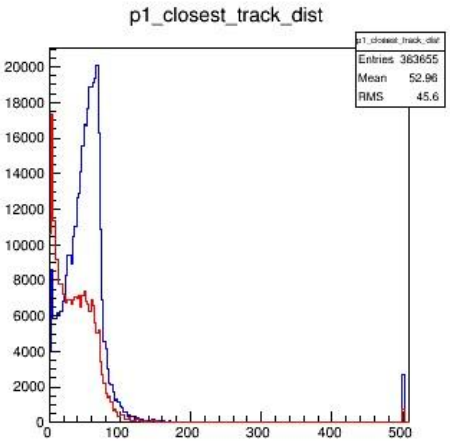


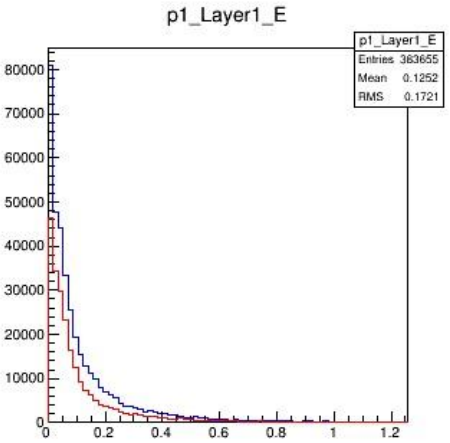


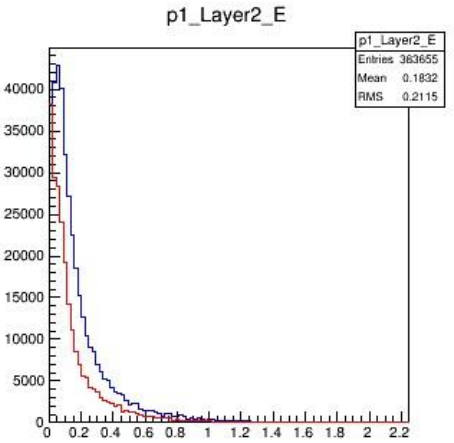






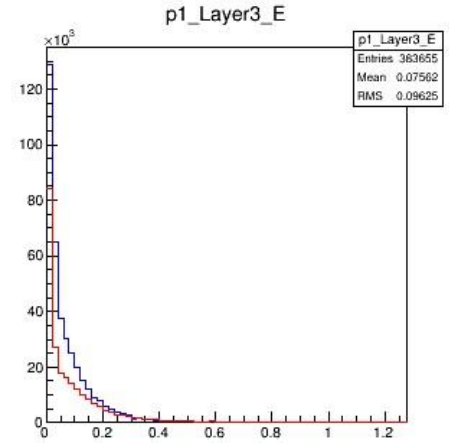


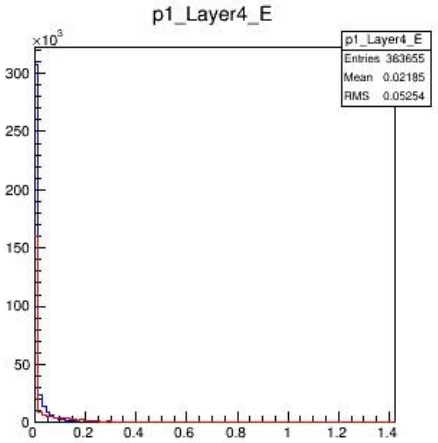


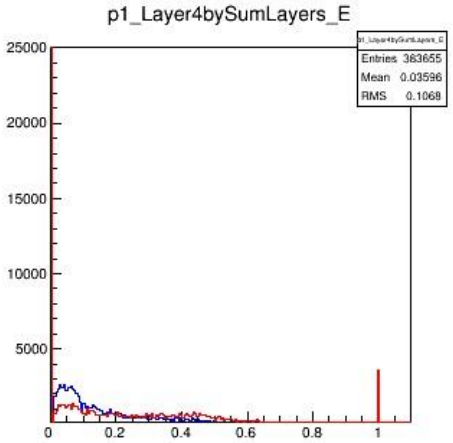


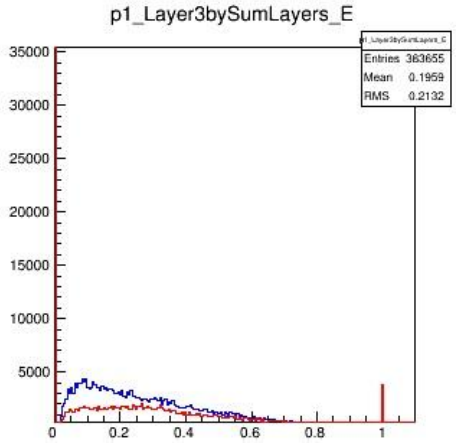
BDT Input Variables with Global Points Cut

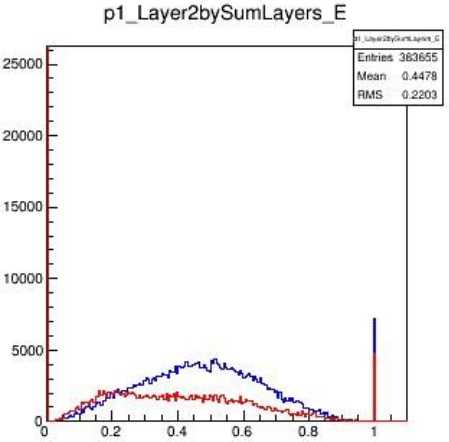
Photon 1 Variables

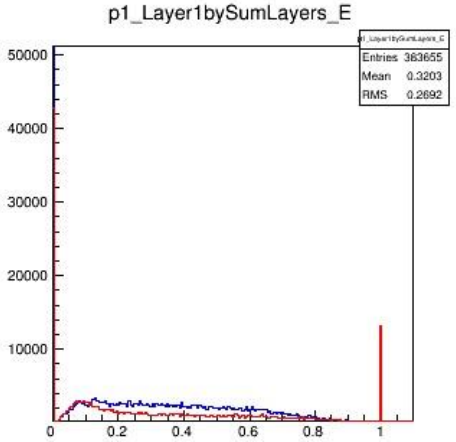


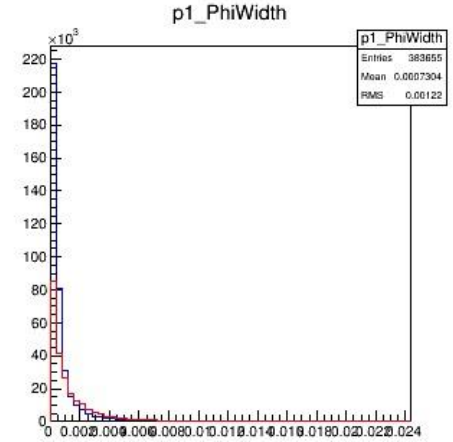


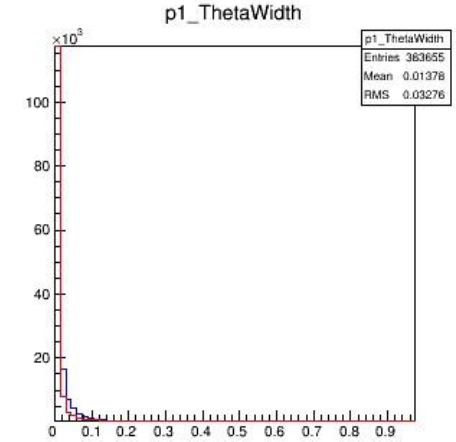












BDT Input Variables with Global Points Cut

Photon 1 Variables