

A series of orange circles of varying sizes arranged in a descending pattern on the left side of the slide.

UPDATE OF PIDTOOLD IN THIS WORKSHOP

Masakazu Kurata

The University of Tokyo

HLReco Week, 07/06/2015-07/10/2015

PIDTOOLS

- First version was committed for ilcsoft v01-17-07
 - Using PIDHandler to store the PID results
 - PIDHandler looks work well
 - Only one(global PID)
- During this workshop:
 - Construct several kinds of PID
 - Basic Variables PID(E/P, Ecal/(Ecal+Hcal), Eyoke)
 - dEdx PID
 - Shower shapes PID
 - LikelihoodPID(global)
 - Output is based on likelihood, instead of posterior prob.
- Future plan:
 - Integrate low momentum μ/π separation(Hale is writing code)
 - Include new variables
 - New PID method?

DUMPEVENT

```
File Edit Options Buffers Tools Help
kurata@ccw08:~/myAna/myAna/slciofile/trackinfo/qqqh_eLpR

clusters ( [ id ] ): [00001851]
particle ids ( [id], PDG, (type)):
vertices: startVertex( id:[ 00000000]id_aRP: 00000000) endVertex( id:[00000000], id_aRP:[00000000]
[00002459] | 0 | 221-2,53e-01, -1,16e-01, -1,48e-01|3,15e-01|0,00e+00|0,00e+00|+0,00e+00, +0,00e+00, +0,00e+00|00000000| 0,00e+00)

covariance( px.py.pz.E ): (0,00e+00, 0,00e+00, 0,00e+00, 0,00e+00, 0,00e+00, 0,00e+00, 0,00e+00, 0,00e+00, 0,00e+00, 0,00e+00)
particles ( [ id ] ):
tracks ( [ id ] ):
clusters ( [ id ] ): [00001852]
particle ids ( [id], PDG, (type)):
vertices: startVertex( id:[ 00000000]id_aRP: 00000000) endVertex( id:[00000000], id_aRP:[00000000]
[00002460] | 0 | 221-1,66e-02, -1,40e-01, -2,54e-02|1,44e-01|0,00e+00|0,00e+00|+0,00e+00, +0,00e+00, +0,00e+00|00000000| 0,00e+00)

covariance( px.py.pz.E ): (0,00e+00, 0,00e+00, 0,00e+00, 0,00e+00, 0,00e+00, 0,00e+00, 0,00e+00, 0,00e+00, 0,00e+00, 0,00e+00)
particles ( [ id ] ):
tracks ( [ id ] ):
clusters ( [ id ] ): [00001853]
particle ids ( [id], PDG, (type)):
vertices: startVertex( id:[ 00000000]id_aRP: 00000000) endVertex( id:[00000000], id_aRP:[00000000]
[00002461] | 0 | 221-1,55e-01, -1,03e-01, +2,31e-01|2,37e-01|0,00e+00|0,00e+00|+0,00e+00, +0,00e+00, +0,00e+00|00000000| 0,00e+00)

covariance( px.py.pz.E ): (0,00e+00, 0,00e+00, 0,00e+00, 0,00e+00, 0,00e+00, 0,00e+00, 0,00e+00, 0,00e+00, 0,00e+00, 0,00e+00)
particles ( [ id ] ):
tracks ( [ id ] ):
clusters ( [ id ] ): [00001854]
particle ids ( [id], PDG, (type)):
vertices: startVertex( id:[ 00000000]id_aRP: 00000000) endVertex( id:[00000000], id_aRP:[00000000]

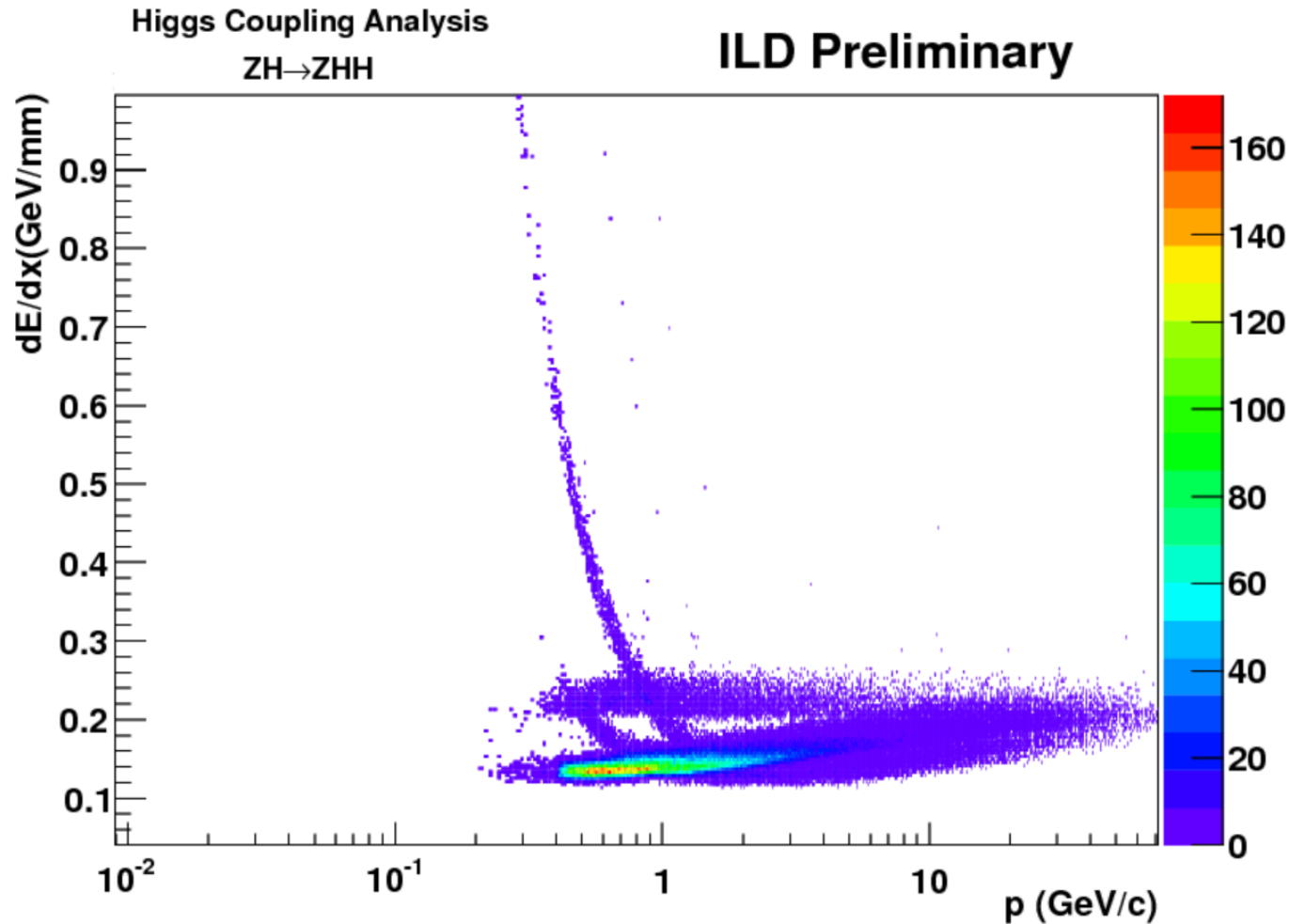
----- detailed PID info: -----
algorithms :
[id: 0] BasicVariablePID - params: electronLikelihood muonLikelihood pionLikelihood kaonLikelihood protonLikelihood MVAOutput_mupiSeparation electronProbability muonProbability pionProbability kaonProbability protonProbability
[id: 3] LikelihoodPID - params: electronLikelihood muonLikelihood pionLikelihood kaonLikelihood protonLikelihood MVAOutput_mupiSeparation electronProbability muonProbability pionProbability kaonProbability protonProbability
[id: 2] ShowerShapesPID - params: electronLikelihood muonLikelihood pionLikelihood kaonLikelihood protonLikelihood MVAOutput_mupiSeparation electronProbability muonProbability pionProbability kaonProbability protonProbability
[id: 1] dEdxPID - params: electronLikelihood muonLikelihood pionLikelihood kaonLikelihood protonLikelihood MVAOutput_mupiSeparation electronProbability muonProbability pionProbability kaonProbability protonProbability electron_dEdxdistance muon_dEdxdistance pion_dEdxdistance
e kaon_dEdxdistance proton_dEdxdistance

[particle] | PDG | likelihood | type | algoId | parameters :
[00000741]
[00000742] | 321 | -7,0865e+00 | 000000 | 0 | [ electronLikelihood : -8,95e+00, muonLikelihood : -1,35e+01, pionLikelihood : -7,34e+00, kaonLikelihood : -7,09e+00, protonLikelihood : -7,45e+00, MVAOutput_mupiSeparation : -1,00e+00, electronProbability : 2,08e-03, muon\
nProbability : 4,35e-05, pionProbability : 2,56e-01, kaonProbability : 5,55e-01, protonProbability : 1,87e-01,]
| 211 | -7,8910e+00 | 000000 | 1 | [ electronLikelihood : -1,71e+01, muonLikelihood : -1,44e+01, pionLikelihood : -7,89e+00, kaonLikelihood : -1,11e+01, protonLikelihood : -9,57e+00, MVAOutput_mupiSeparation : -1,00e+00, electronProbability : 7,82e-05, muon\
nProbability : 1,18e-03, pionProbability : 8,13e-01, kaonProbability : 3,40e-02, protonProbability : 1,52e-01, electron_dEdxdistance : -1,03e+01, muon_dEdxdistance : -1,44e+00, pion_dEdxdistance : -5,73e-01, kaon_dEdxdistance : 2,00e+00, proton_dEdxdistance : 7,99e-01,]
| 321 | -1,3454e+01 | 000000 | 2 | [ electronLikelihood : -1,56e+01, muonLikelihood : -1,69e+01, pionLikelihood : -1,37e+01, kaonLikelihood : -1,35e+01, protonLikelihood : -1,36e+01, MVAOutput_mupiSeparation : -1,00e+00, electronProbability : 9,62e-04, muon\
nProbability : 1,09e-04, pionProbability : 2,49e-01, kaonProbability : 4,43e-01, protonProbability : 3,07e-01,]
| 211 | -2,8935e+01 | 000000 | 3 | [ electronLikelihood : -4,17e+01, muonLikelihood : -4,48e+01, pionLikelihood : -2,89e+01, kaonLikelihood : -3,16e+01, protonLikelihood : -3,06e+01, MVAOutput_mupiSeparation : -1,00e+00, electronProbability : 3,17e-16, muon\
nProbability : 7,10e-12, pionProbability : 9,67e-01, kaonProbability : 3,88e-04, protonProbability : 3,16e-02,]

[00000747]
[00000748]
[00000749]
[0000074a]
[0000074b] | 13 | -5,1004e+00 | 000000 | 1 | [ electronLikelihood : -9,90e+00, muonLikelihood : -5,10e+00, pionLikelihood : -5,38e+00, kaonLikelihood : -1,23e+01, protonLikelihood : -8,76e+00, MVAOutput_mupiSeparation : -1,00e+00, electronProbability : 4,57e-03, muon\
nProbability : 5,35e-01, pionProbability : 4,23e-01, kaonProbability : 4,00e-04, protonProbability : 1,44e-02, electron_dEdxdistance : -1,31e+01, muon_dEdxdistance : -3,79e-01, pion_dEdxdistance : -4,41e-03, kaon_dEdxdistance : -5,20e+00, proton_dEdxdistance : -2,55e+01,]
| 13 | -5,5761e+00 | 000000 | 0 | [ electronLikelihood : -5,63e+00, muonLikelihood : -5,58e+00, pionLikelihood : -6,20e+00, kaonLikelihood : -7,15e+00, protonLikelihood : -5,82e+00, MVAOutput_mupiSeparation : -1,00e+00, electronProbability : 2,89e-01, muon\
nProbability : 3,66e-01, pionProbability : 1,04e-01, kaonProbability : 1,57e-02, protonProbability : 2,25e-01,]
| 13 | -6,1144e+00 | 000000 | 2 | [ electronLikelihood : -8,75e+00, muonLikelihood : -6,11e+00, pionLikelihood : -7,72e+00, kaonLikelihood : -7,86e+00, protonLikelihood : -7,98e+00, MVAOutput_mupiSeparation : -1,00e+00, electronProbability : 1,09e-02, muon\
nProbability : 5,83e-01, pionProbability : 1,95e-01, kaonProbability : 1,14e-01, protonProbability : 9,72e-02,]
| 13 | -2,0903e+01 | 000000 | 3 | [ electronLikelihood : -2,65e+01, muonLikelihood : -2,09e+01, pionLikelihood : -2,16e+01, kaonLikelihood : -2,97e+01, protonLikelihood : -2,48e+01, MVAOutput_mupiSeparation : -1,00e+00, electronProbability : 3,47e-07, muon\
nProbability : 8,29e-01, pionProbability : 1,70e-01, kaonProbability : 1,81e-09, protonProbability : 3,12e-05,]

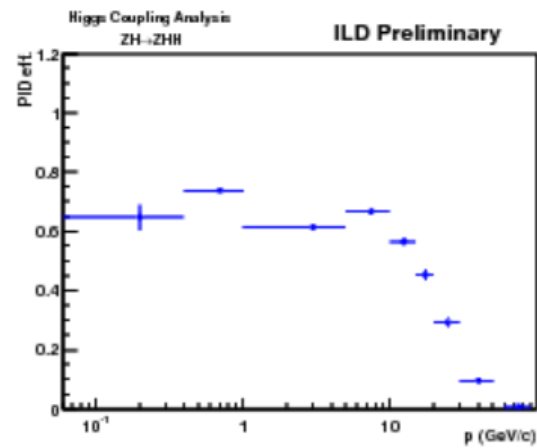
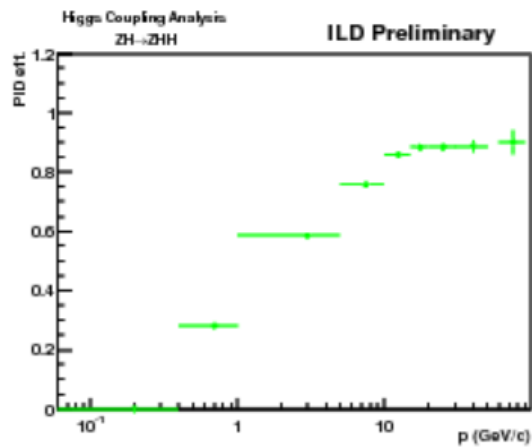
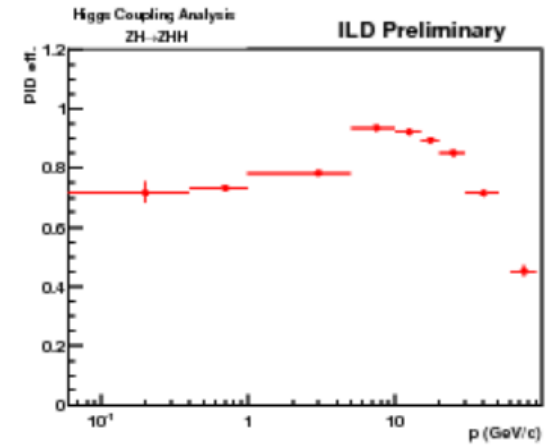
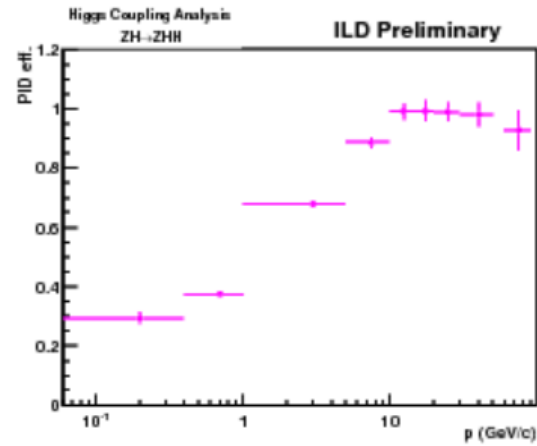
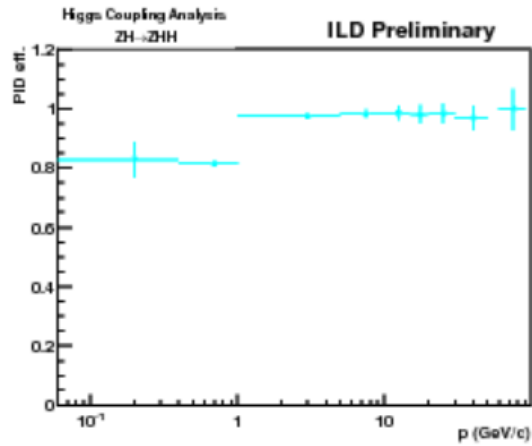
[00000750]
-UU:***-F1 check_log 70% [4471 (Fundamental)]
```

D \bar{E} DX DISTRIBUTION



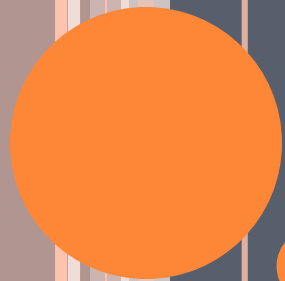
- More sophisticated plot later
 - But something like that?

PID EFFICIENCY



PROSPECTS

- Looks working – need many performance plots
 - dEdx slice plot
 - dEdx pull
 - With bethe-bloch lines
 - etc.etc.
- PID
 - Momentum dependence
 - Angle dependence
 - Fake rate
 - etc.etc...
- Neutral PID?



BACK UPS

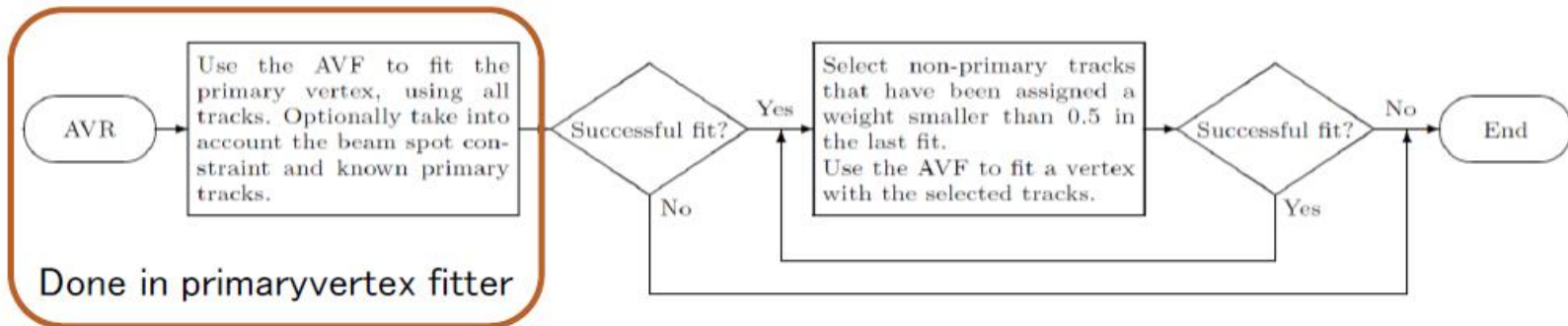


ADAPTIVE VERTEX FITTING

- To introduce the effect of multi-vertex fitting
 - Introduce weight function to estimate which vertex a track belongs to
 - Weight function definition: k-th track' s weight on n-th vertex

$$w_{nk} = \frac{e^{-\chi_{nk}^2/2T}}{e^{-\chi_{\text{cut}}^2/2T} + \sum_{i=1}^N e^{-\chi_{ik}^2/2T}}$$

- Parameter: temperature T
 - If T small, decision is like χ^2 minimization
 - If T large, multi-vertex effect becomes large(suppress the weight function)
- Apply it to associate IP tracks to secondary vertices:
- Algorithm: Adaptive Vertex Fitting



- Tracks will belong to the vertices when $w_{nk} > 0.5$ (k-th track belongs to n-th vertex)
 - Try to fit more tracks than nominal algorithm

VERTEX FINDING OF BJETS

- Common parameters are set at same values for comparison
- Same event sample(qqHH sample@500GeV) 19889 events
- 6 jet clustering, jet matching with MCtruth is performed
- Num. of jets

method	bjet with 2vtx	bjet with 1+1vtx	bjet with 1vtx	total
Nominal Algorithm	10577	9159	12804	32504
AVF&BNess	13461	6502	14256	34219

- Jets with vtx: ~5% increased
 - Jets with 2vtx: ~24% increased → move from 1+1
 - Jets with 1vtx: ~11% increased
- Fake track rate per vtx: how many are fake tracks contaminated to vertices?
 - Seems fake singletrk is increased → need opt. and more selection

method	bjet with 2vtx	bjet with 1+1vtx	bjet with 1vtx
Nominal Algorithm	0.011 ± 0.0007	0.007 ± 0.0006	0.025 ± 0.001
AVF&BNess	0.010 ± 0.0006	0.011 ± 0.0009	0.021 ± 0.001

VERTEX FINDING OF C JETS

- Common parameters are set at same values for comparison
- Same event sample(nnH sample@500GeV) 99432 events
 - H→cc: 6461 events
- 2 jet clustering, jet matching with MCtruth is performed
- Num. of vertices

method	cjet with 2vtx	cjet with 1+1vtx	cjet with 1vtx	total
Nominal Algorithm	43	165	6537	6745
AVF&BNess	84	215	6960	7259

- Total: $\sim 7\%$ increased
- Vertex mis-ID eff. is increased(but, 2vtx jet has pure vertices)
 - Though num. of vertices is small
 - need additional selection for singletrk? (e.g.)vertex mass?)
- Fake track rate per vtx:

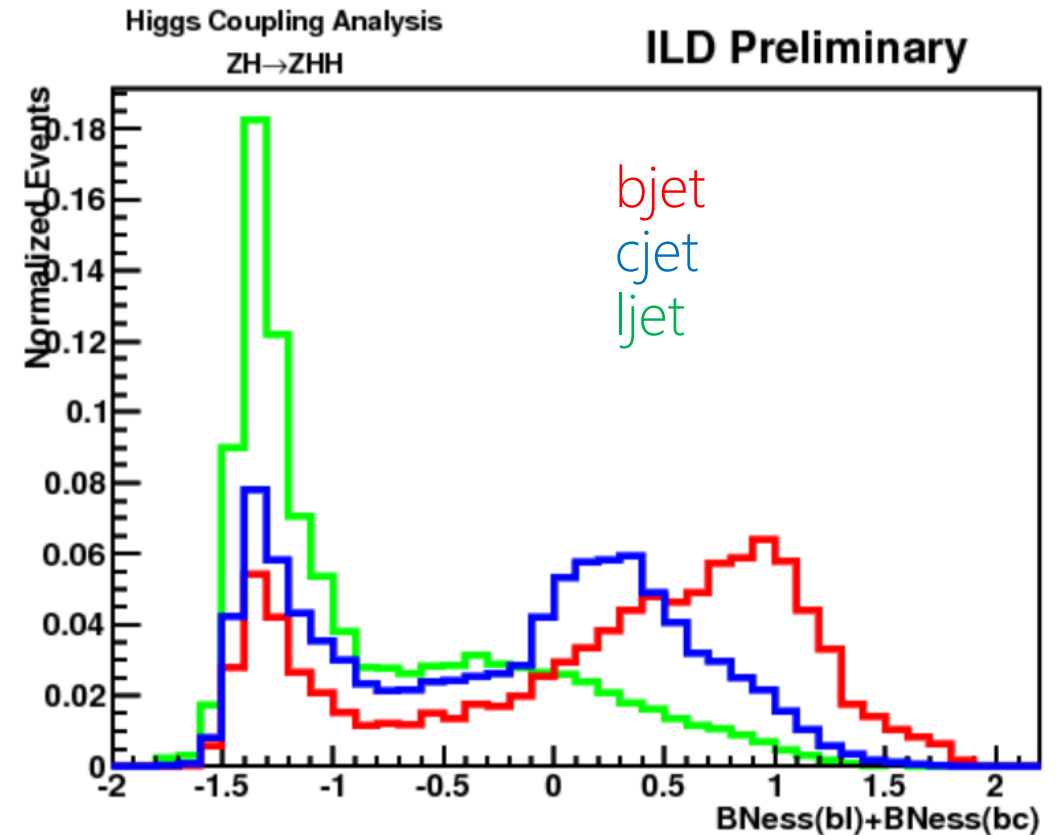
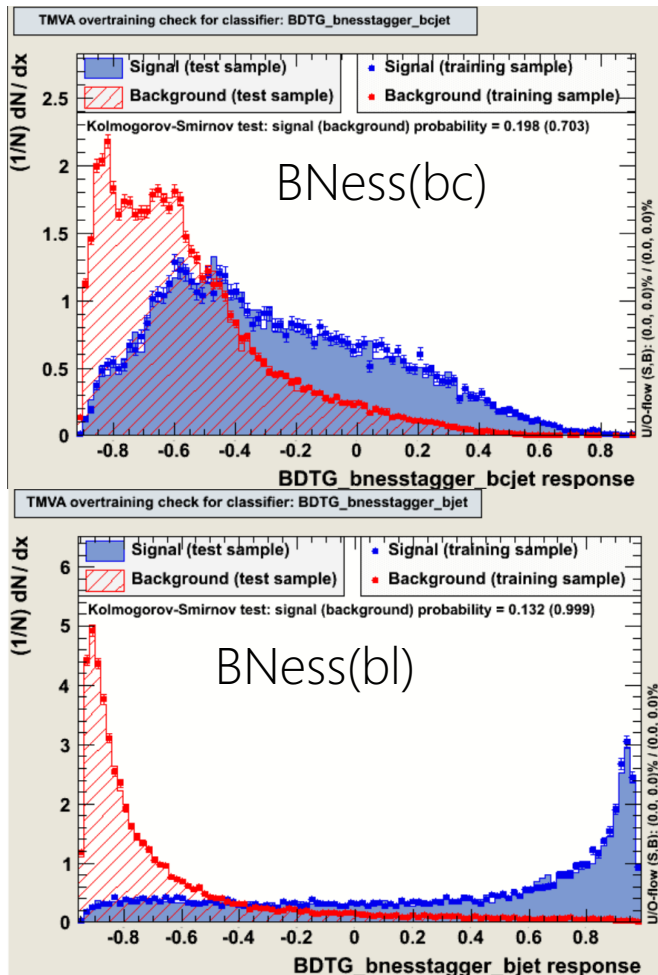
method	cjet with 2vtx	cjet with 1+1vtx	cjet with 1vtx
Nominal Algorithm	0.00 ± 0.00	0.012 ± 0.006	0.0014 ± 0.004
AVF&BNess	0.00 ± 0.00	0.018 ± 0.007	0.0013 ± 0.004

BNESS TAGGER

- Flavor separation of 0vtx jet is most difficult situation
 - Only impact parameter implies the existence of secondary vertices for flavor separation
- BNess tagger will be worth trying in this case!
 - Developed in CDF
 - Focus on individual tracks and evaluate jet flavor only using single track
 - Track's potential for coming from heavy flavor particle(D&B meson and baryons) should be evaluated(using MVA)
- Difficulty in ILC
 - In CDF, it is important to separate b and other flavor → c quark separation is not required
 - In ILC, separation among b, c and other is very important → bc separation is a key for flavor tagger
- How is bc separation using BNess tagger?

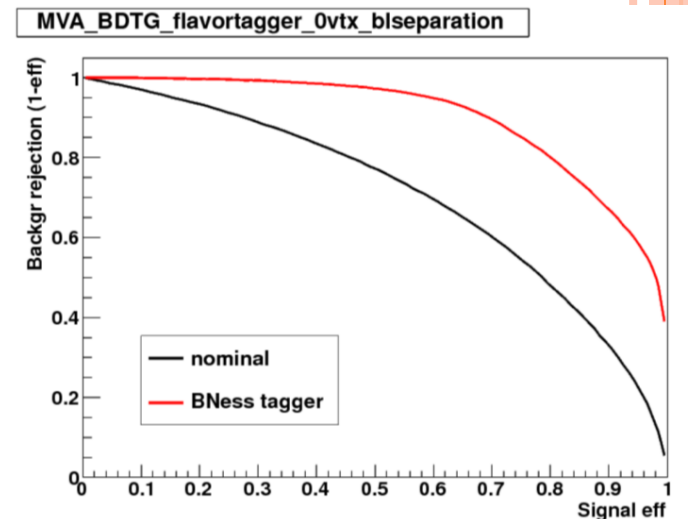
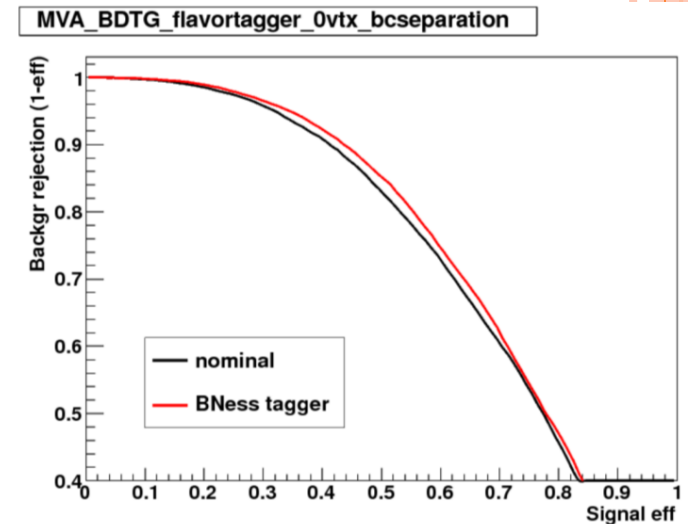
BNESS OUTPUT

- Collect **Highest score BNess** track in 0vtx jets
- Final BNess is defined as BNess(bl) + BNess(bc)
- Well separated between bjets and l jets
- Difference can be seen between bjets and cjets



RESULTS OF BNESSTAGGER ON FLAVOR TAGGING

- Construct a “toy” flavor tagger
 - Convert nominal input variables to BNesstagger variables
 - Compare with ROC curve
- For bc separation, some improvement can be obtained
- For bl separation, becomes too good? under investigation
- b-l separation will be very good!
- Need optimization
- Especially, precise study of b-c-l flavor separation is necessary

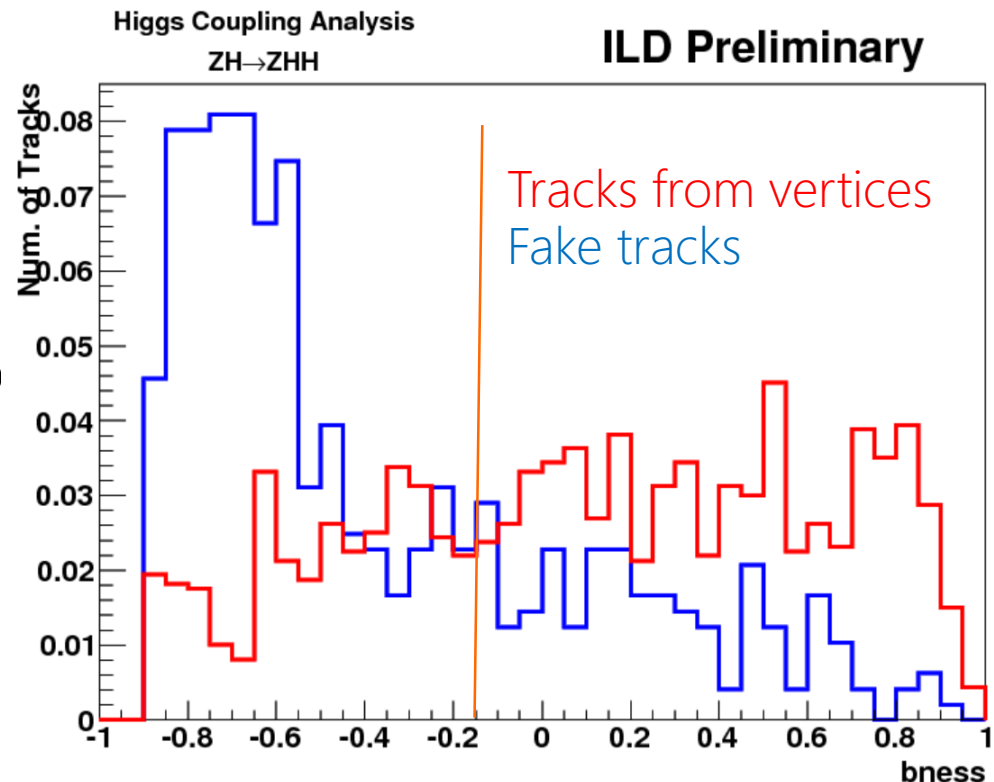
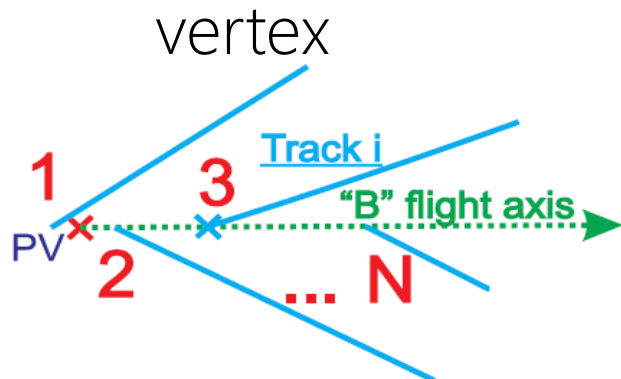


BNESS TAGGER FOR FAKE TRACK REJECTION

- Loosen the track selection to try to attach as many tracks as possible to vertices
 - Fake track rate will be increased
- To reject fakes, BNess tagger is used
 - So far, just use BNess(bl)
- So far, only BNess is checked

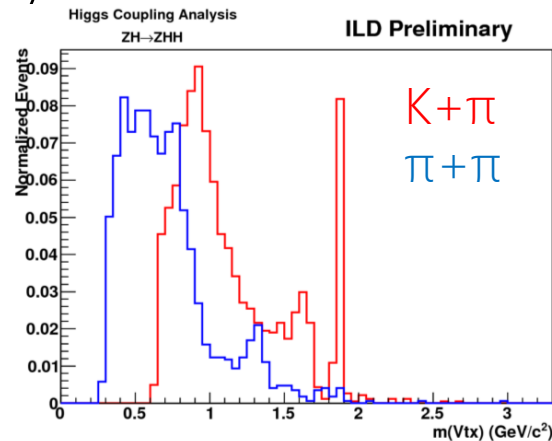
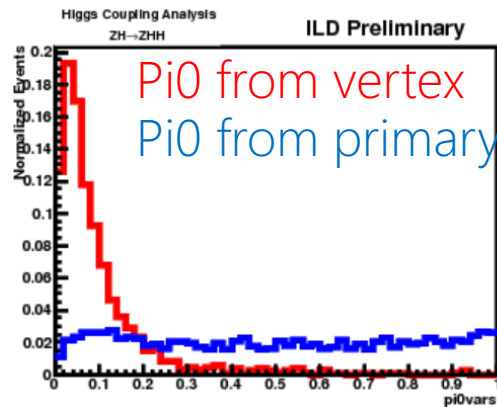
→ some bias for D meson tracks?

Example: looking for single track

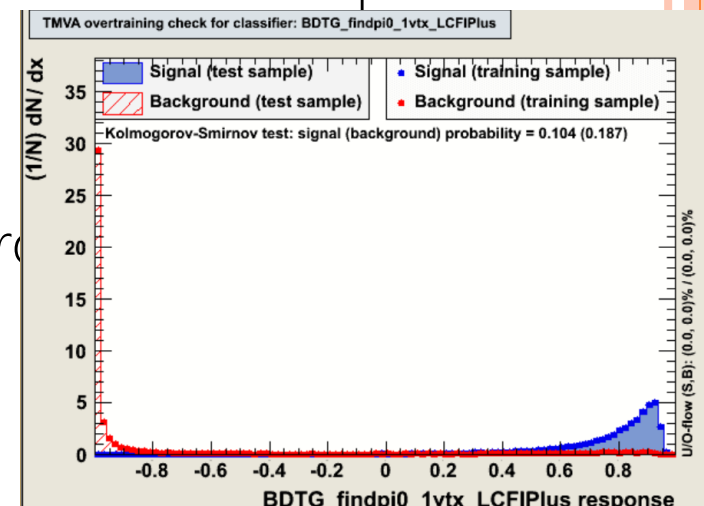


VERTEX MASS RECOVERY

- Using π^0 s which escape from vertices
 - Need to choose good π^0 candidates –construct π^0 vertex finder
 - Key issue – π^0 kinematics, very collinear to vertex direction

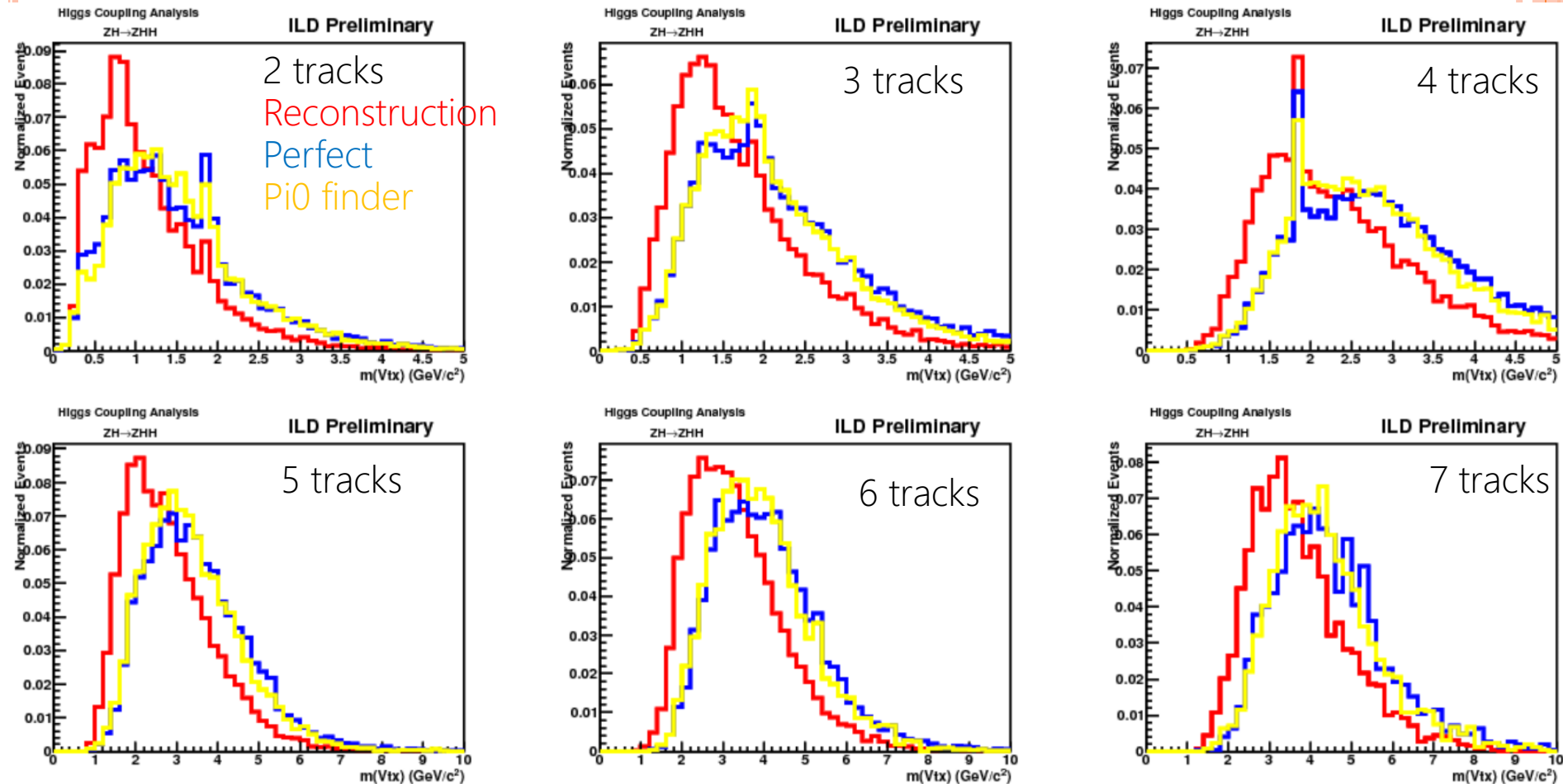


- Particle ID is the other key to classify vertices
 - Different particle patterns have different vertex mass patterns
- Construct π^0 Vertex finder using MVA
 - Identify which vertex π^0 s are coming from



VTX MASSES OF BJETS IN DOUBLE-HIGGS PROCESS

- Vtx mass distributions for each vertex pattern(ntrk)
 - These results are the outputs of LCFIPlus(unofficial ver.)!
 - Difference is limited by mis-pairing of gammas and mis-attachment of pi0s

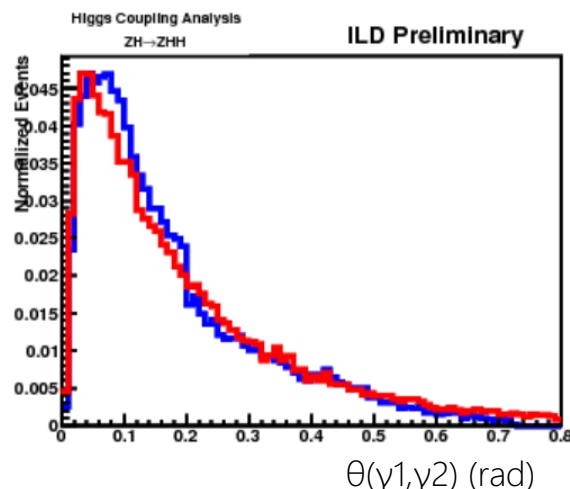
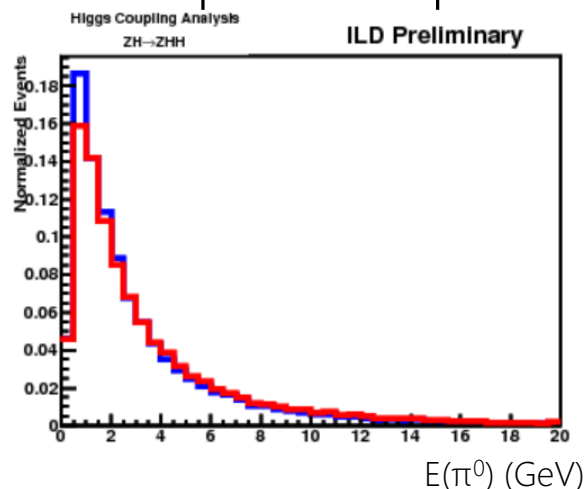


PI0 RECO USING NAÏVE BAYES FOR VERTEX MASS RECOVERY

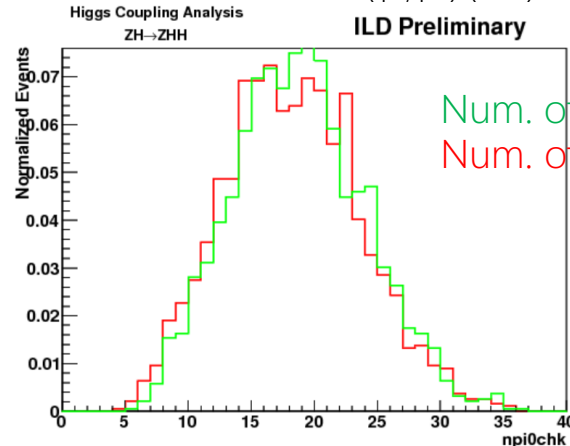
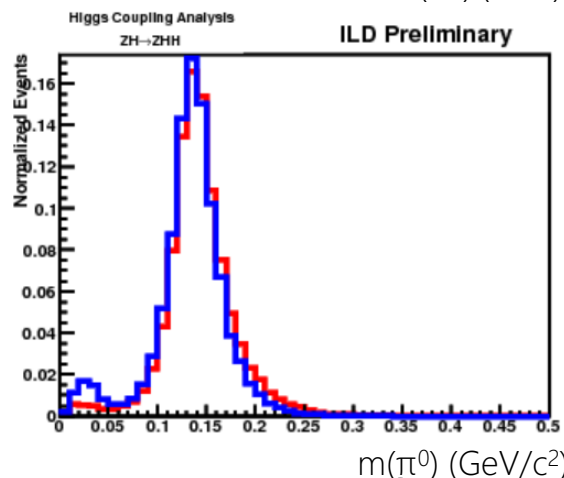
- Good pairing eff. & mis-pairing eff.

	Correct pair	Wrong pair
eff. (%)	46.0 ± 0.3	54.0 ± 0.4

- Kin. plots of pi0 reco. results



MC truth
Pi0 finder

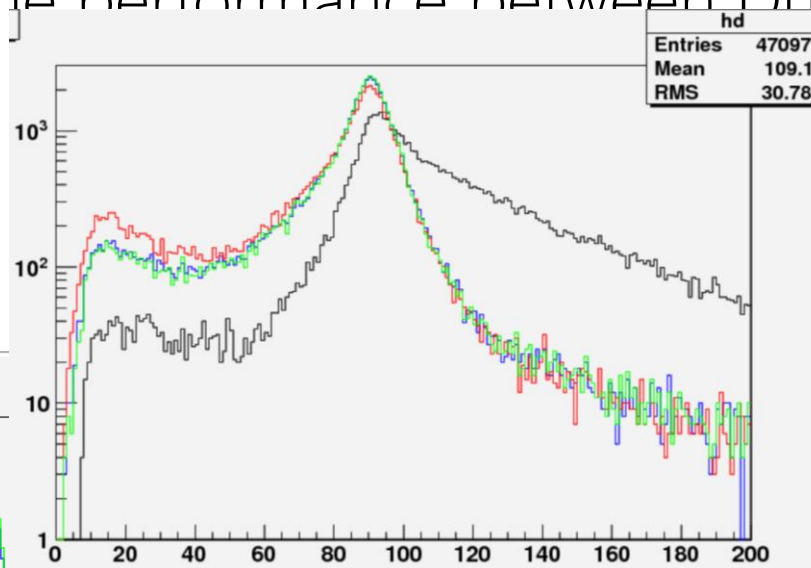


Num. of pi0s to be reconstructed
Num. of pi0s from pi0 finder

- Integrate pi0 reconstruction?

JET CLUSTERING WITH BEAM BACKGROUND REJECTION

- Now in LCFIPlus, Valencia jet clustering is available!
- We also include Durham jet clustering with beam b.g. rejection
 - Assumed very large energy jet exists in beam direction
- Compare the performance between Durham, Kt and Valencia



w/o beam b.g. rejection

Kt

Valencia

Durham

$\nu\nu Z$ @500GeV

2 jet clustering

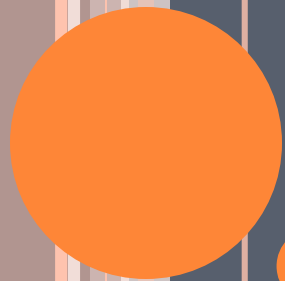
Parameters are tuned
for better result

OTHER TOPICS

- Other change from DBD LCFIPlus:
 - Automatic creation of Joint Probability plots
 - We can calibrate for that variable
 - Performance check for new joint probability is ongoing
- Paper available:
 - arxiv: 1506:08371
 - NIM paper has been submitted

SUMMARY AND PROSPECTS

- For flavor tagging improvement:
 - New vertexing algorithm(AVF) will provide better vertex finding efficiency
 - BNesstagger will give some improvement for 0vtx jet flavor separation
 - There seems hope for attaching π^0 s to vertices to recover vertex mass
- So far, AVF will provide 4-7% improvement of vertex finding in bjets
 - Need to check the bias of fake track rejection using BNesstagger
 - Vertex quality check is necessary
 - This study will lead to vertex charge assignment improvement
- 0vtx jet case will improve well not only b-c separation, but also b-l separation
 - More Precise study of b-c-l separation is necessary
- Vertex mass recovery is reasonable
 - Will provide better flavor tagger
 - Of course, many checks are necessary
- Finally, incorporate all the ideas and check the final flavor tagging effs.in LCFIPlus!



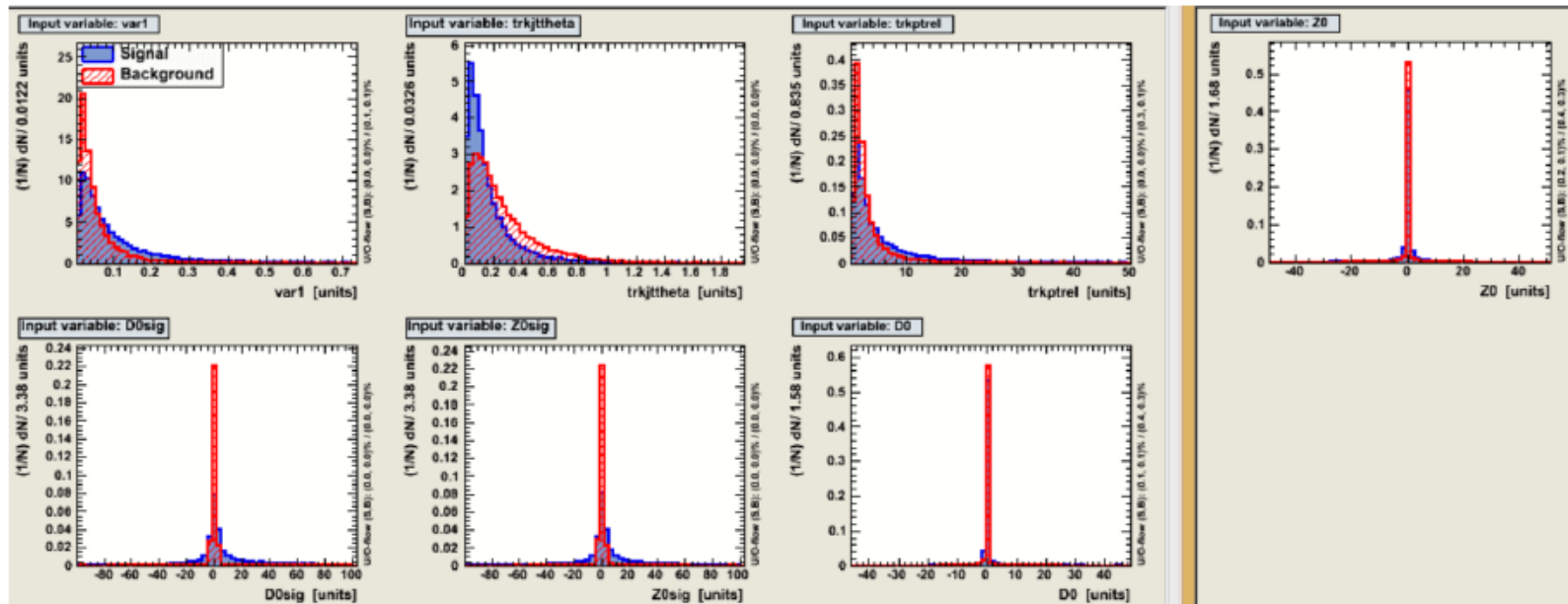
BACK UPS2

21

TRACK MVA(BNESS)

- To identify track which comes from heavy flavor particle
→ using MVA
 - Signal: tracks which come from B mesons or B baryons
 - Background: tracks produced in hadronization process
- Most significant tracks with both plus and minus signed impact parameters in a jet are collected

- Significance: $sig = \sqrt{\left(\frac{d_0}{\sigma}\right)^2 + \left(\frac{z_0}{\sigma}\right)^2}$



VTX MASSES

- Vtx mass distributions for each vertex pattern(ntrk)
 - not so bad
 - Difference is coming from **mis-pairing of gammas** and mis-attachment of pi0s

