Current status of LCFIPlus study for flavor tagging improvement

Masakazu Kurata

The University of Tokyo

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

LCFIPlus improvement

- Flavor tagging is one of the most important analysis tools
- DBD LCFIPlus, which is a flavor tagging module, has been successful
- We need to go to next step, flavor tagging improvement
  - There is much room to improve!
- Now, focusing on
  - Vertex finding efficiency improvement itself
  - Flavor separation in the case of 0vtx jet
  - Vertex Mass Recovery using pi0s

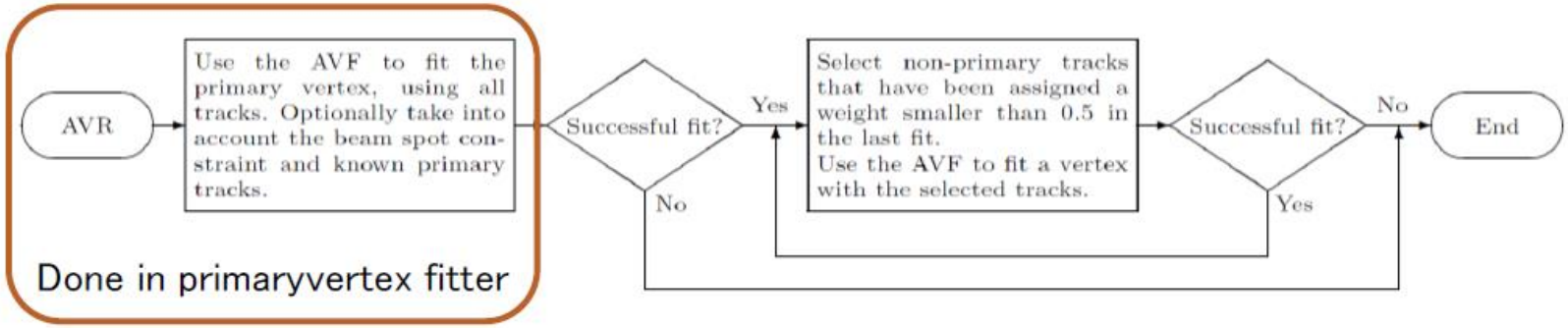
- Particle ID is one of the key to flavor tagging improvement
- Pi0 reco. is other key　for vertex mass recovery

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

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

  - 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: ～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:

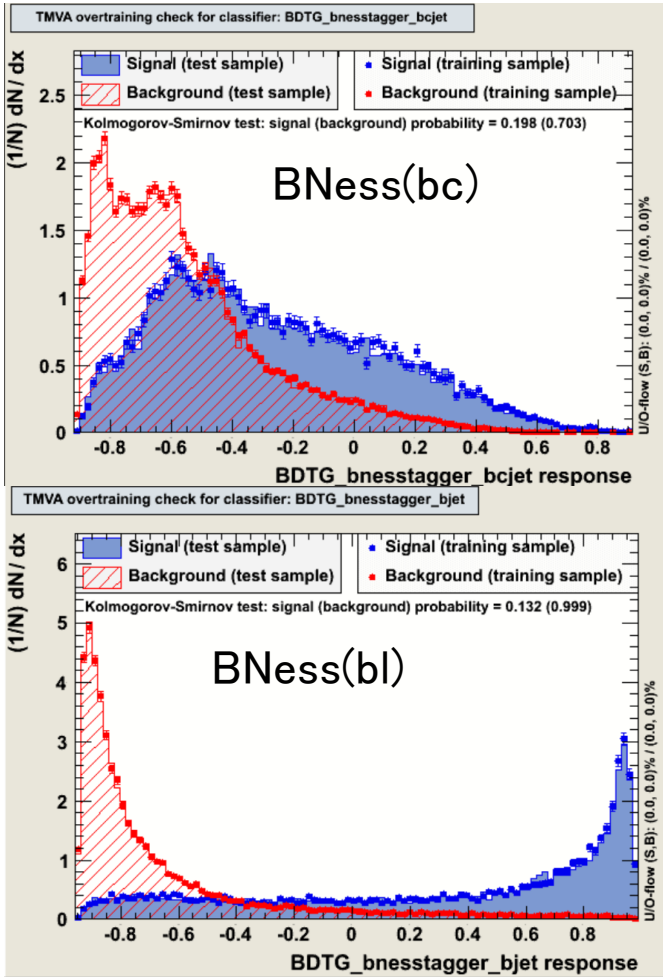
method	cjet with 2vtx	cjet with 1+1vtx	cjet with 1vtx
Nominal Algorithm	0.00±0.00	0.012±0.006	0.0014±0.004
AVF&BNess	0.00±0.00	0.018±0.007	0.0013±0.004

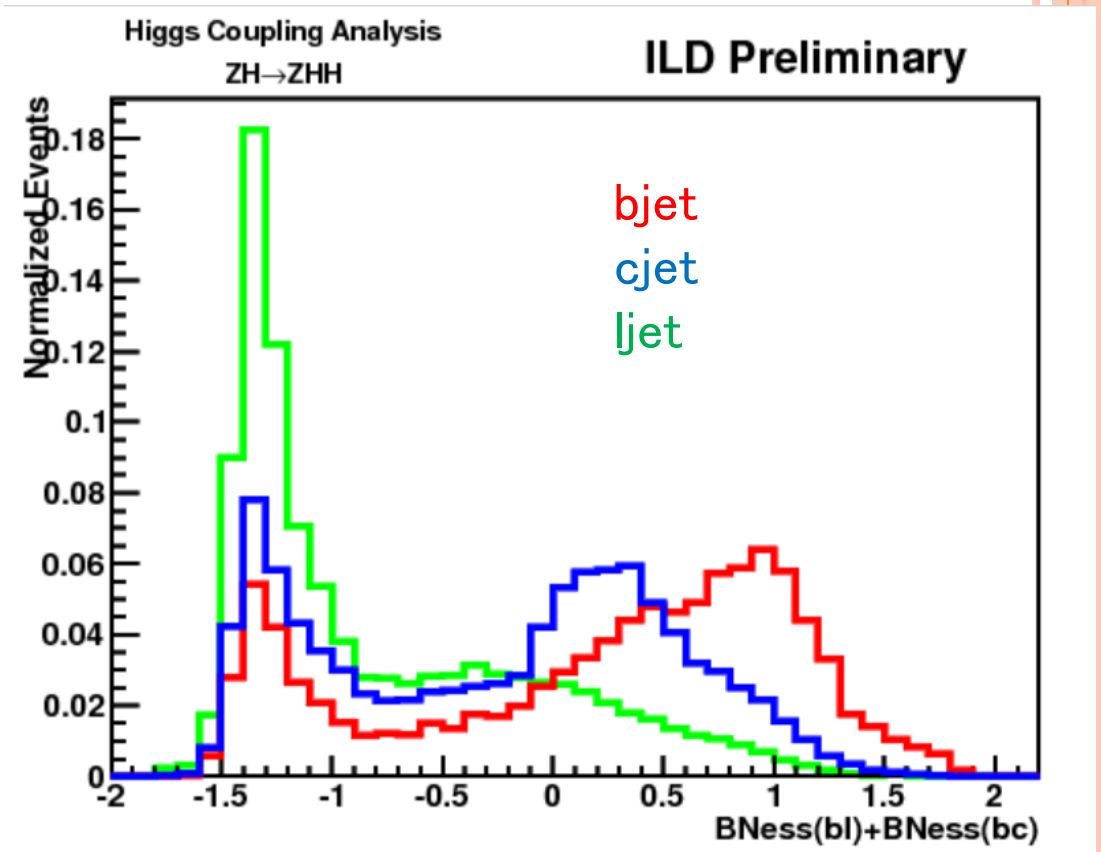
BNESS TAGGER

- 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

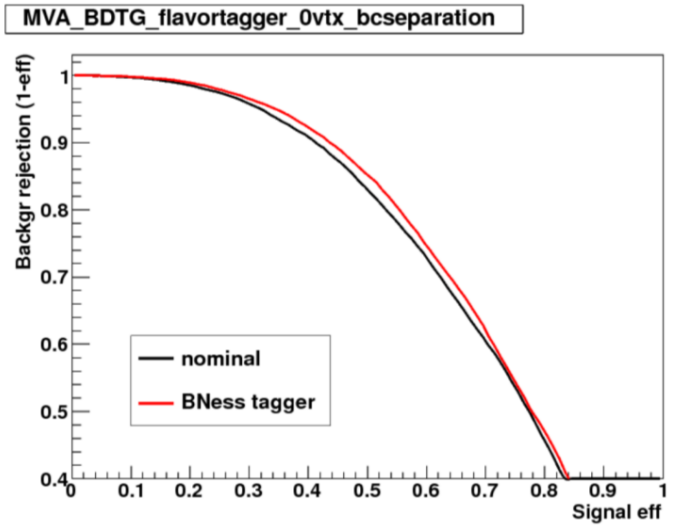
- Collect Highest score BNess track in 0vtx jets
- Final BNess is defined as BNess(bl)+ BNess(bc)
- Well separated between bjets and l jets
- Difference can be seen between bjets and cjets



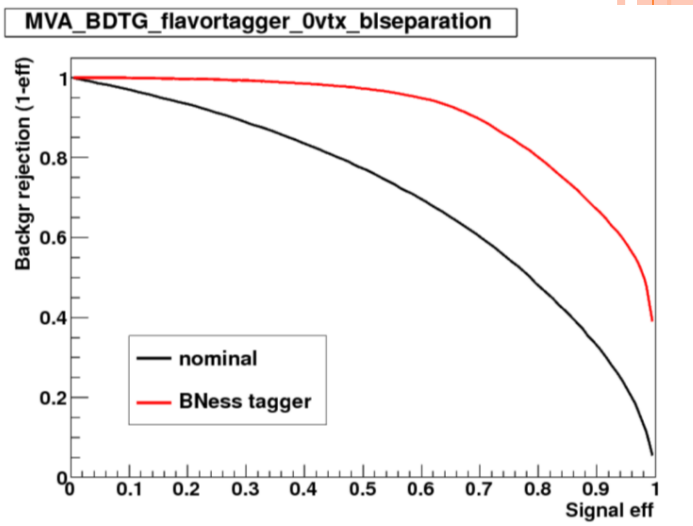


Results of BNess tagger on flavor tagging

- 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 sepraration, becomes too good? under investigation
- b–l separation will be very good!



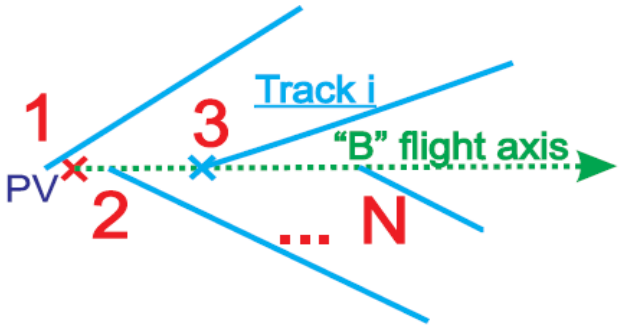
- Need optimization
- Especially, precise study of b–c–l flavor separation is necessary

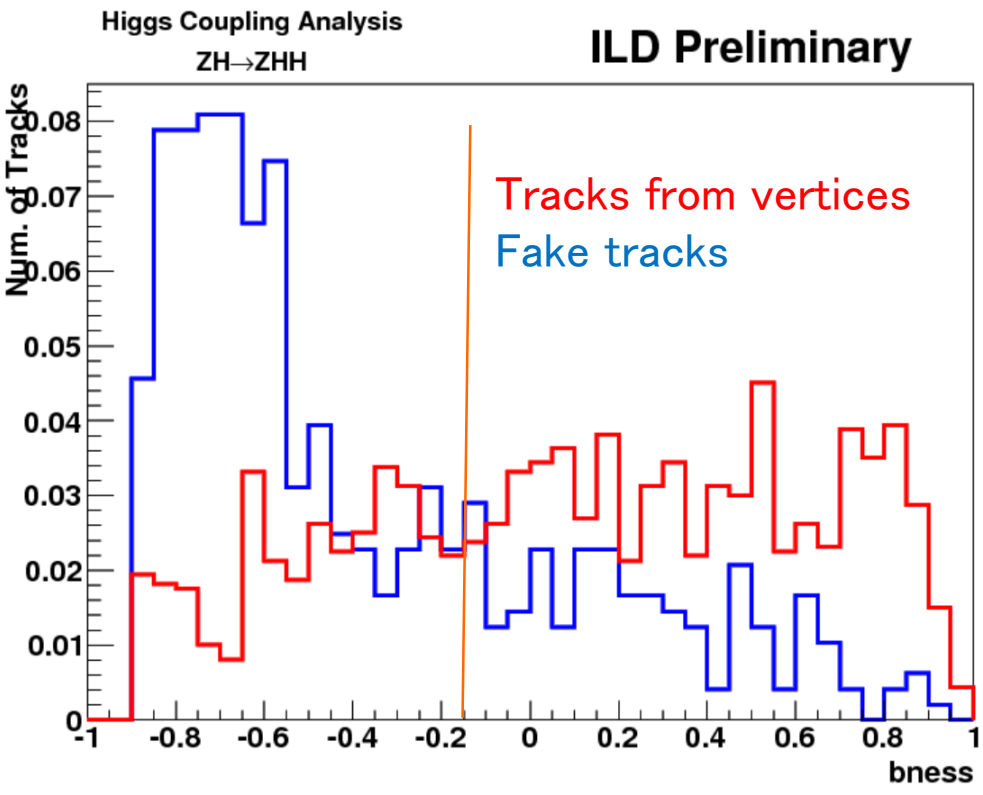
BNESS TAGGER FOR FAKE TRACK REJECTION

- 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?

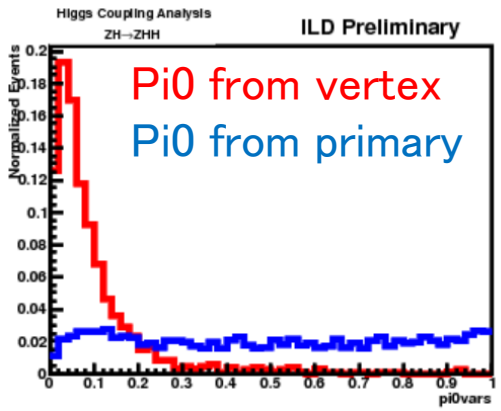
Example: looking for single track

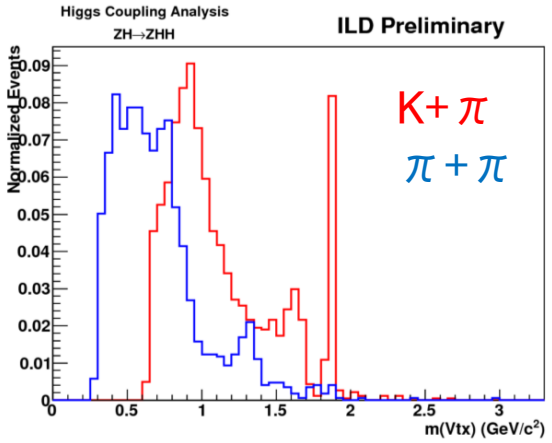




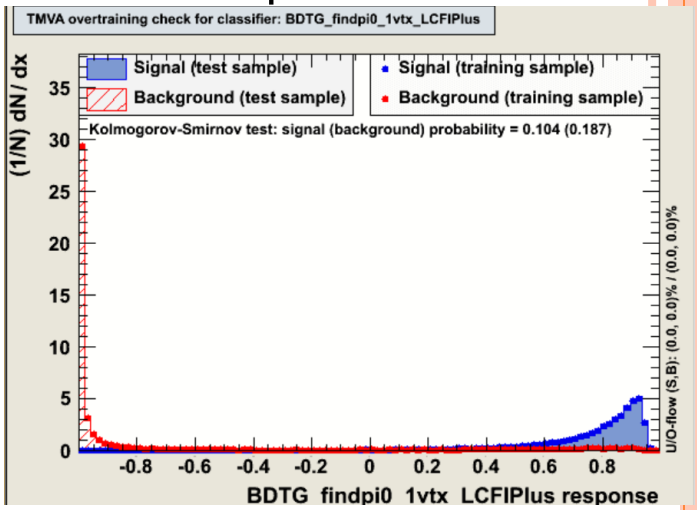
VERTEX MASS RECOVERY

- Using pi0s which escape from vertices
  - Need to choose good pi0 candidates –construct pi0 vertex finder
  - Key issue –pi0 kinematics, very collinear to vertex direction



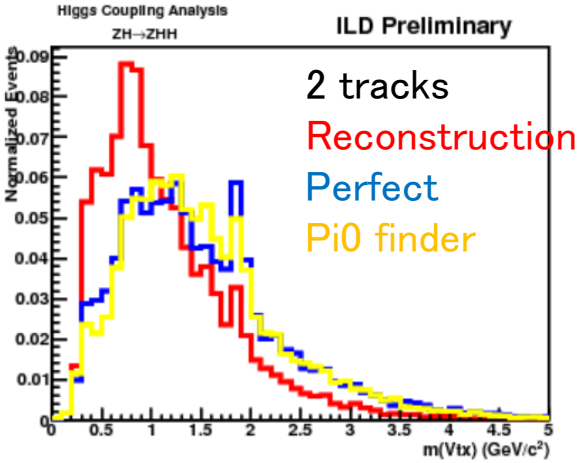


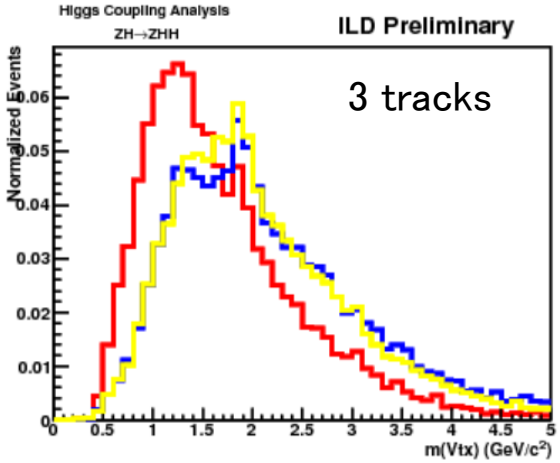
- Particle ID is the other key to classify vertices
  - Different particle patterns have different vertex mass patterns
- Construct Pi0 Vertex finder using MVA
  - Identify which vertex pi0s are coming from

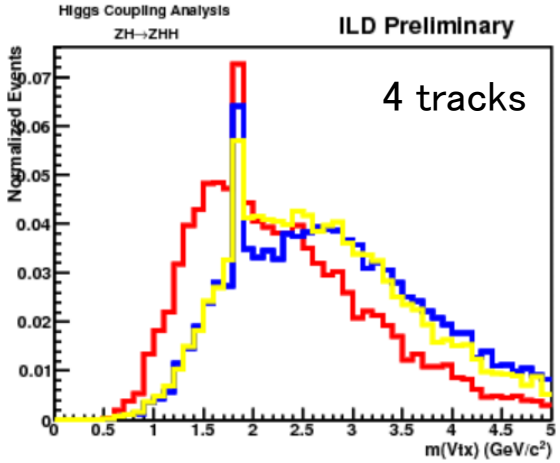


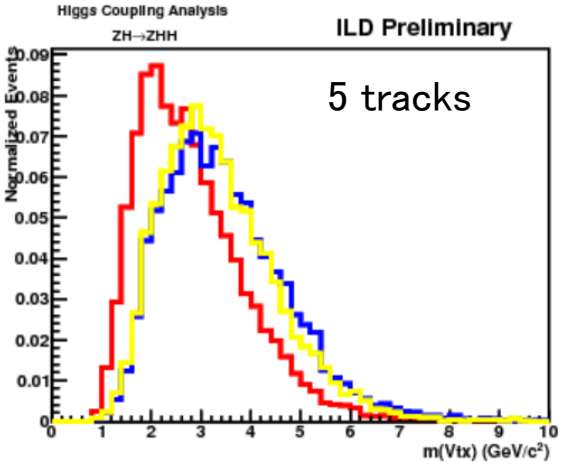
Vtx masses of bjets in double-Higgs process

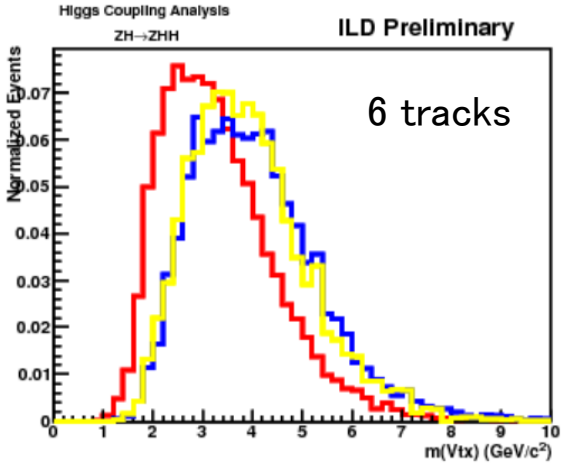
- Vtx mass distributions for each vertex pattern(ntrk)
  - These results are the outputs of LCFIPlus(unofficial ver.)!
  - Difference is limited by mis-pairing of gammas and mis-attachment of pi0s

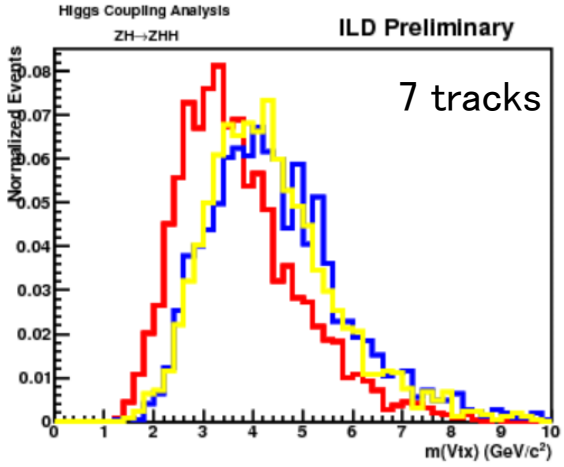










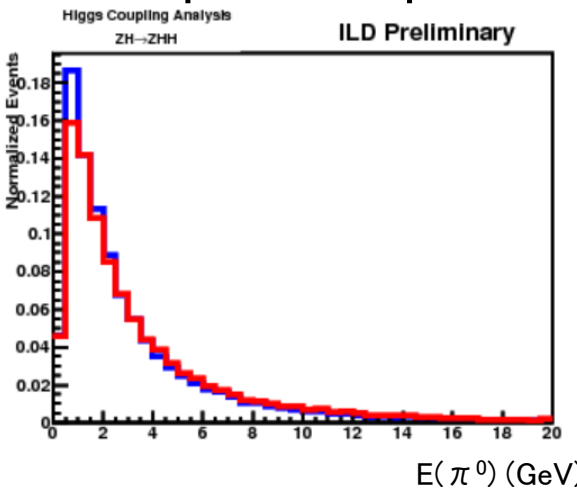


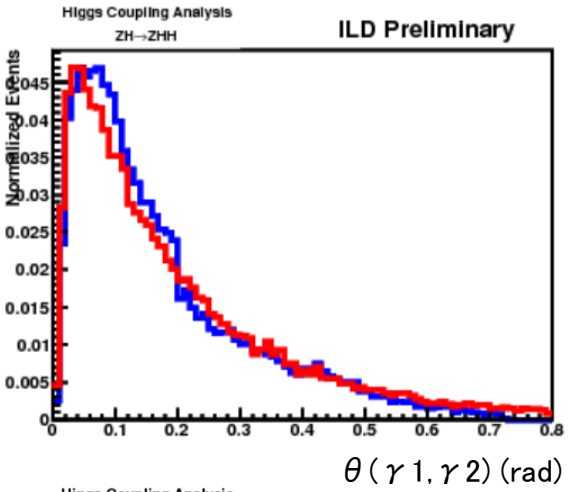
Pi0 reco using Naïve Bayes for vertex mass recovery

- Good pairing eff. & mis-pairing eff.

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

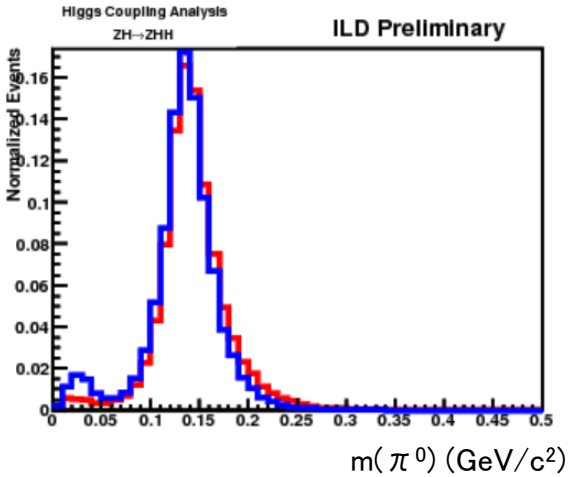
- Kin. plots of pi0 reco. results

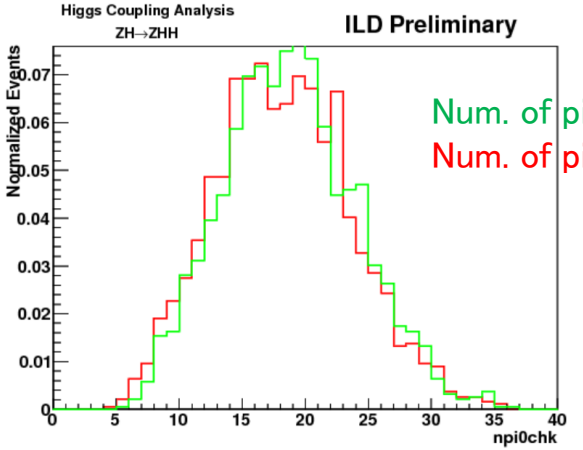




MC truth

Pi0 finder





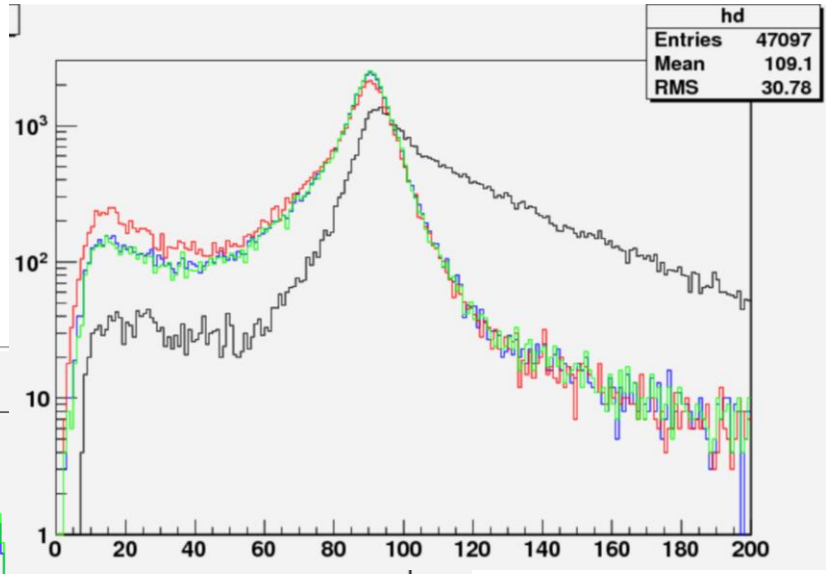
Num. of pi0s to be reconstructed

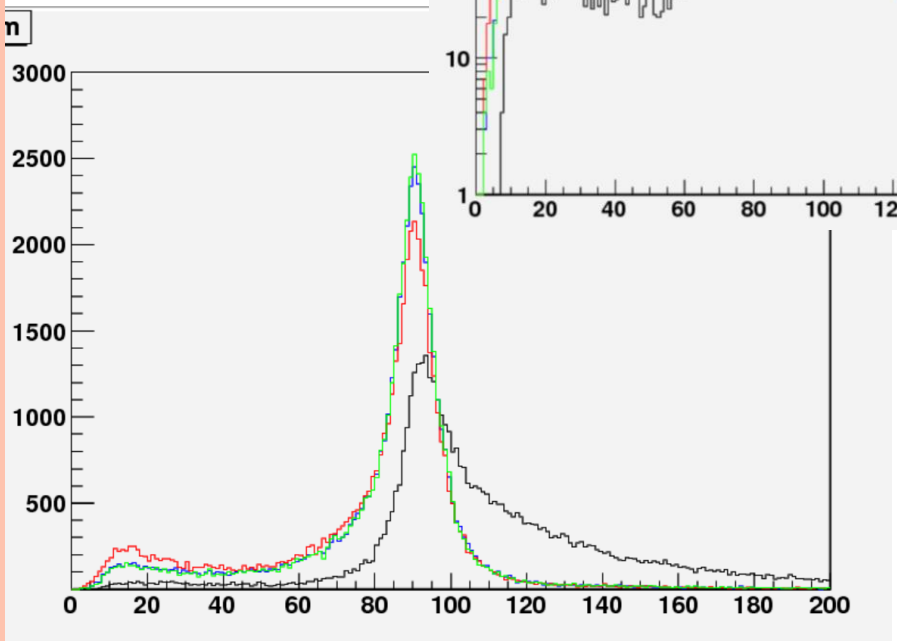
Num. of pi0s from pi0 finder

- 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

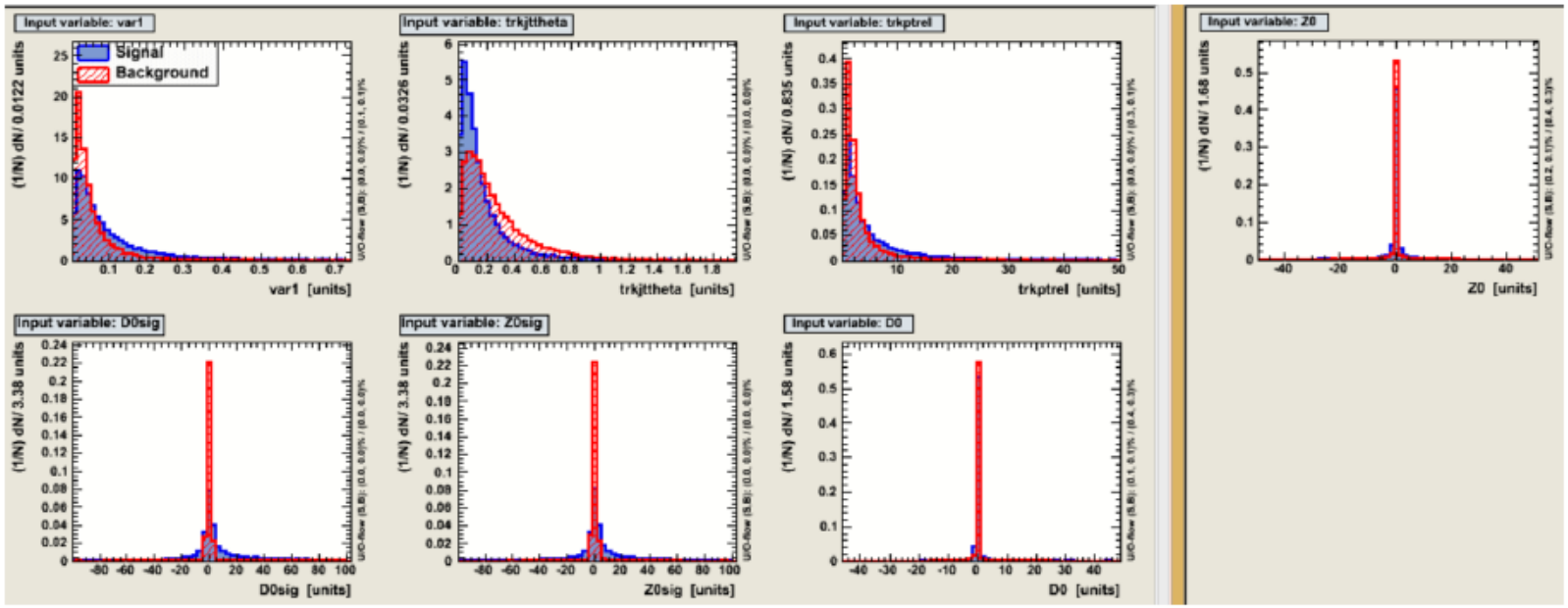
- For flavor tagging improvement:
  - New vertexing algorithm(AVF) will provide better vertex finding efficiency
  - BNesstagger will give some improvement for 0vtx jet flavor separation
  - There seems hope for attaching pi0s to vertices to recover vertex mass
- So far, AVF will provide 4–7% improvement 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 ups

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



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