Particle ID study and its application (@TPC session)

Masakazu Kurata
The University of Tokyo

ALCW15, 04/20/2015-04/24/2015

For analysis improvement

- All the analyses are saturated within the present framework
  - Needs new idea
  - Especially, need to improve the results of Higgs self-coupling@500GeV
- Fundamental new variables might provide improvements of analysis tools @ILD, but not yet used well
  - dE/dx in TPC
  - Shower profiles in the calorimeters

- Particle ID will be available using those variables

- Will those variables give improvements to other analysis components?
  - Isolated lepton ID → of course!
  - Energy correction using PID → it is OK!
  - Flavor tagging using PID? → looks hopeful!
  - Hope for jet clustering? → need to try

→it is necessary to study them

dE/dx from TPC

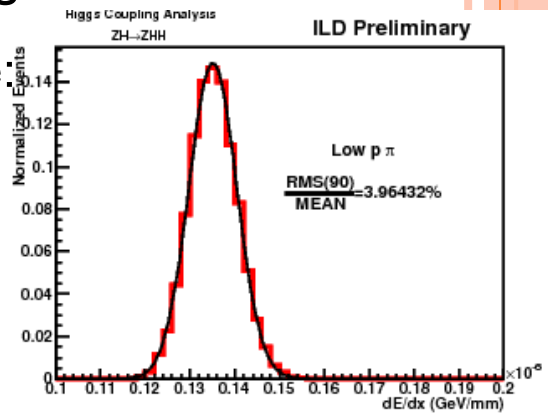
- For improvement, using dE/dx is one of the powerful tools
  - Particle ID for each track will give a large impact to the analysis
  - Application to general analysis component is very wide
    - Lepton ID
    - Track energy correction
    - Flavor tagging
    - Jet clustering?
- Important factor to use dE/dx is: fluctuation
  - TDR: measurement resolution is **5%**
  - So, natural fluctuation from simulation is within 5% without detector effect
- dE/dx definition:
  - $\frac{dE}{dx} = \frac{energy\ deposit}{flight\ path\ in\ the\ hit(TPC)}$
  - dE/dx can be calculated at any hit point
  - Truncated mean is calculated as track dE/dx

    $\left\langle \frac{dE}{dx} \right\rangle = \frac{1}{n}\sum_i^n \frac{dE_i}{dx_i}$ upper 30%, lower 8%(important!) hits are discarded to avoid Landau tail(next slide)

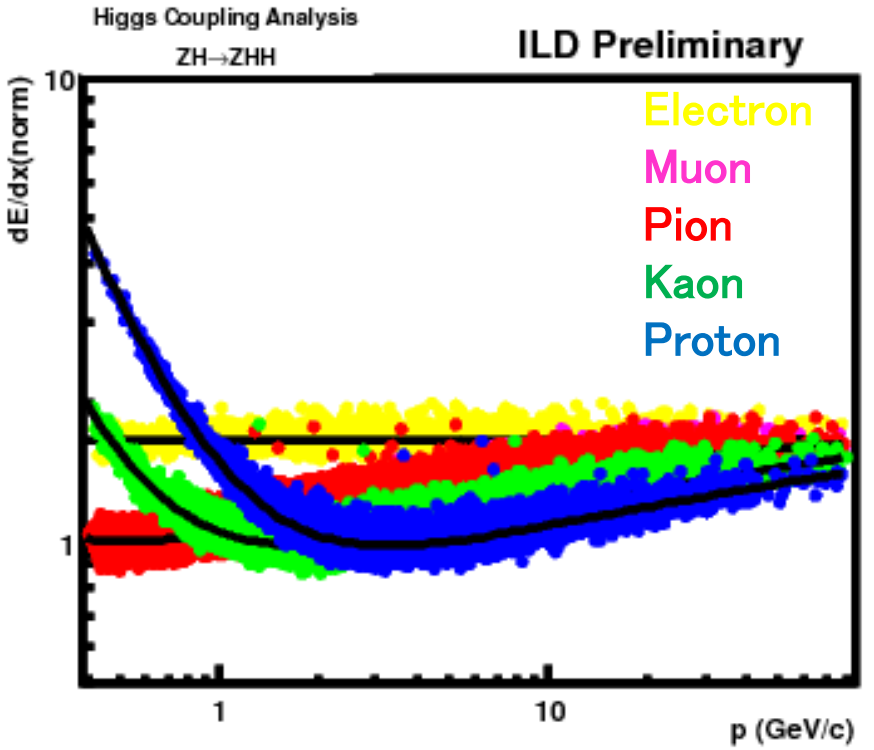
    →optimization is necessary

dE/dx from TPC @ILD

- Fluctuation of dE/dx using various type of tracks
  - Fluctuations of each particle/each momentum range:
    3 – (<5)%!!
  - Including detector effect is necessary
    - Do you have any idea?

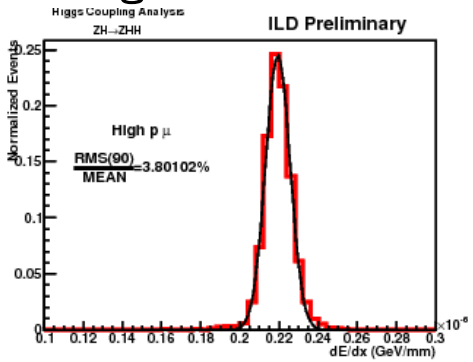


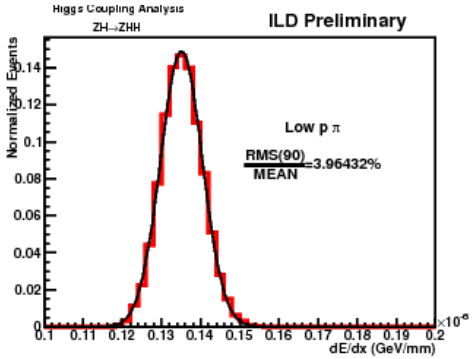
- Momentum dependence of dE/dx for each particle
  - Polar angle dependence corrected
  - Num. of Hits dependence corrected
  - Scale to $\left\langle \frac{dE}{dx} \right\rangle = 1.0$ for MIP pion

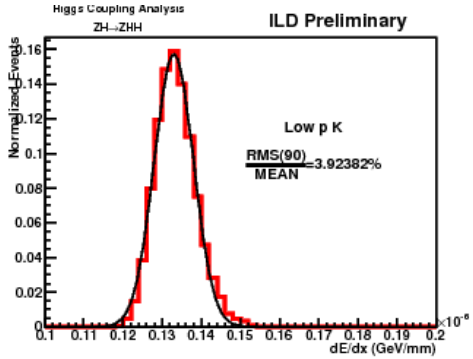


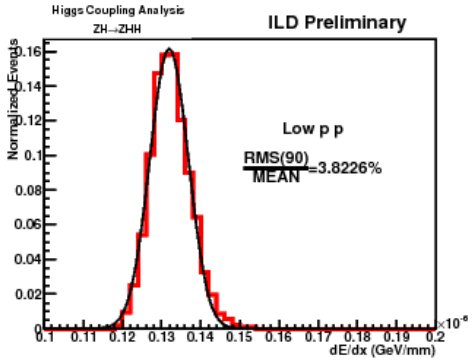
dE/dx fluctuation

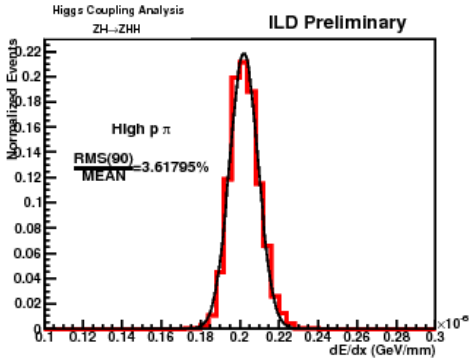
- Fluctuation of dE/dx using various type of tracks
  - Using truncated mean

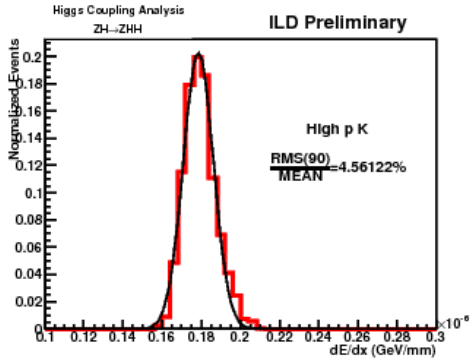


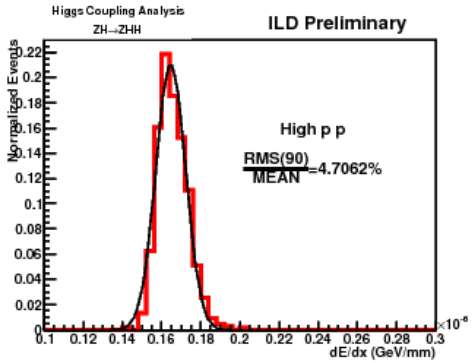












High: p >20GeV/c

Low: π 0.3GeV/c<p<0.6GeV/c

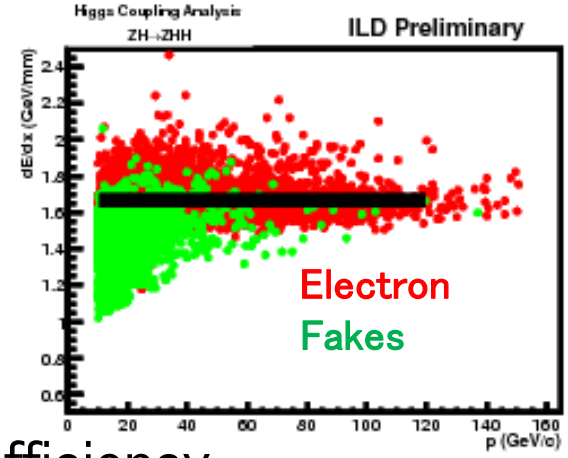
K 1.0GeV/c <p<3.0GeV/c

p 2.0GeV/c<p<4.0GeV/c

RMS(90) is taken for fluctuation

First Application – Isolated lepton ID

- Lepton ID for single lepton – using likelihood method
  - Lepton likeliness: $L = \frac{\prod s}{\prod s + \prod b}$,
  - Variables: traditional variables(Ecal/(Ecal+Hcal), E/P, D0, Z0, cone energy)
  - And using dE/dx(convert to χ^2) & shower profiles
- Signal is HH→(bb)(WW*)→(bb)(l ν jj)



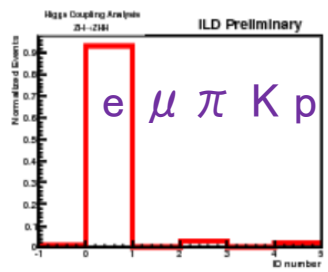
- Signal detection efficiency – set almost same efficiency
- Background rejection efficiency:

Single lepton ID	Cut based	Old likelihood	New likelihood
Signal(%)	98.1	98.1	97.8
ttbar – all hadronic(%)	7.9	3.1	2.3

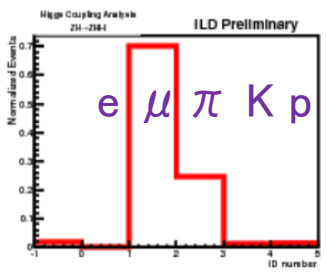
- Improvement of all hadronic event rejection: ～30%
- Note: lepton energy threshold is loosened on likelihood_new
  - From E(lep)>15GeV → E(lep)>10GeV

Particle ID @ILD

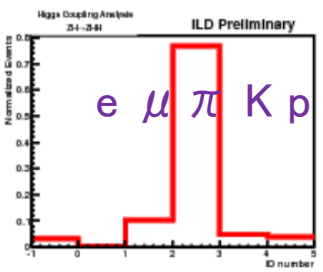
- New variables make Particle ID available
  - How are particles identified as each particle type?
- Construct Particle ID algorithm:
  - Based on Bayesian approach: define posterior probability
  - Make "rejected" category:
    - Track is rejected if its posterior probability is below threshold
    - Those tracks are moved to pions
- Overall ID efficiency of tracks in jets:
  - Electron can be identified almost perfectly(>90%)
  - Muon ID eff. is ～70% →due to low energy muons(μ / π separation)
  - Hadron ID effs. are 62%～75%



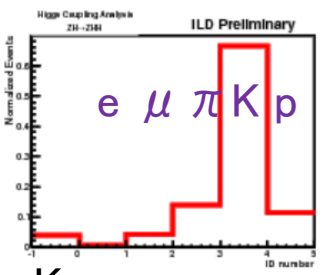
e



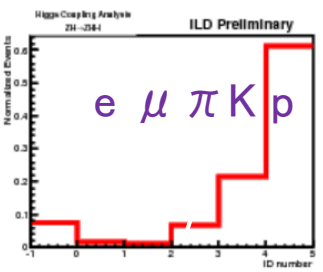
μ



π



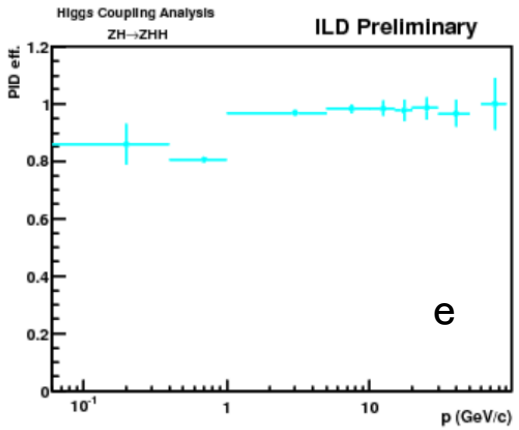
K

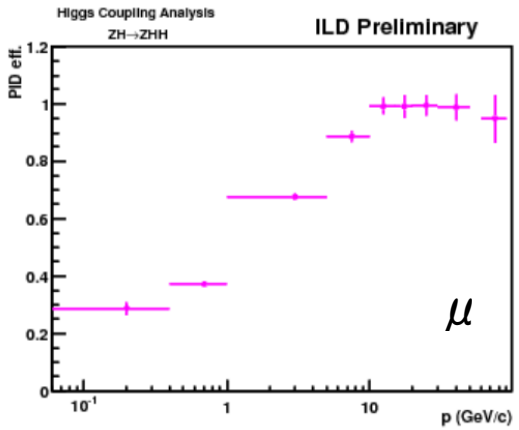


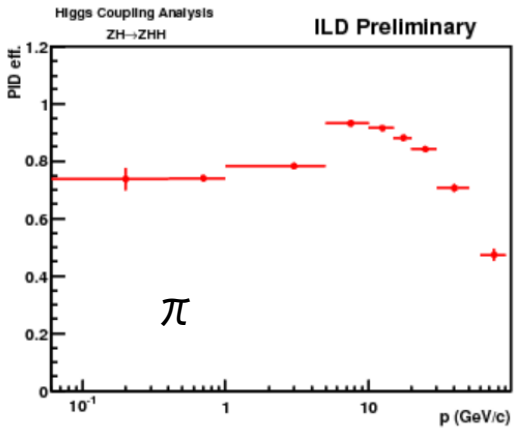
p

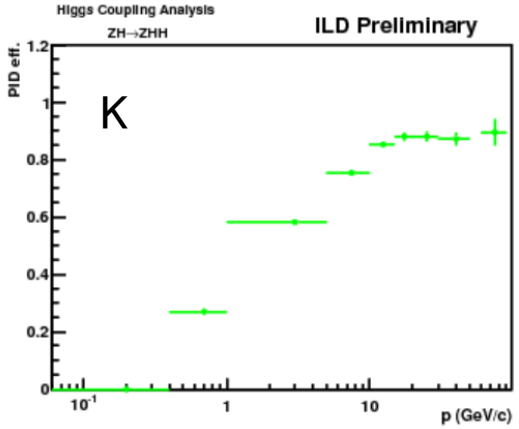
Look more

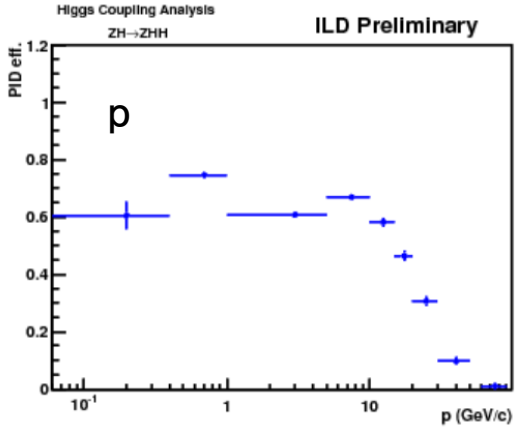
- Momentum Dependence of Particle ID efficiency
  - Momentum ranges where PID is good/bad







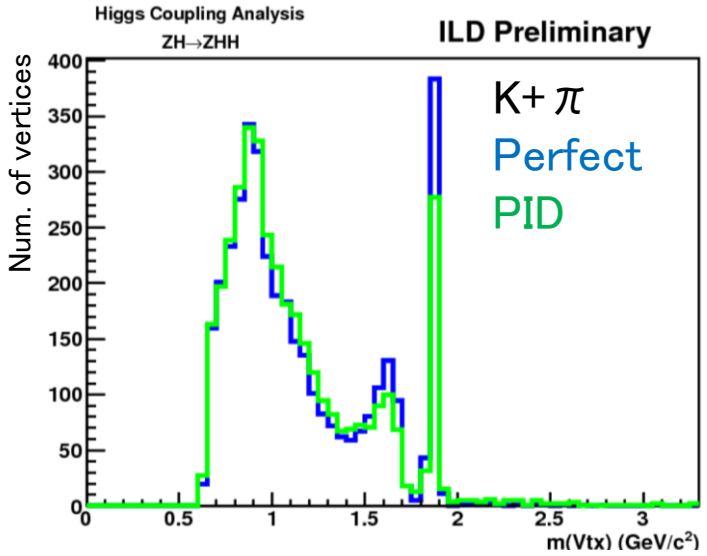




  - Electron ID is good
  - PID efficiency is >60% @1GeV/c～20GeV/c
  - Low momentum μ / π separation is difficult
  - Too low momentum PID is not effective(tracking is good?)

Vertex mass using PID

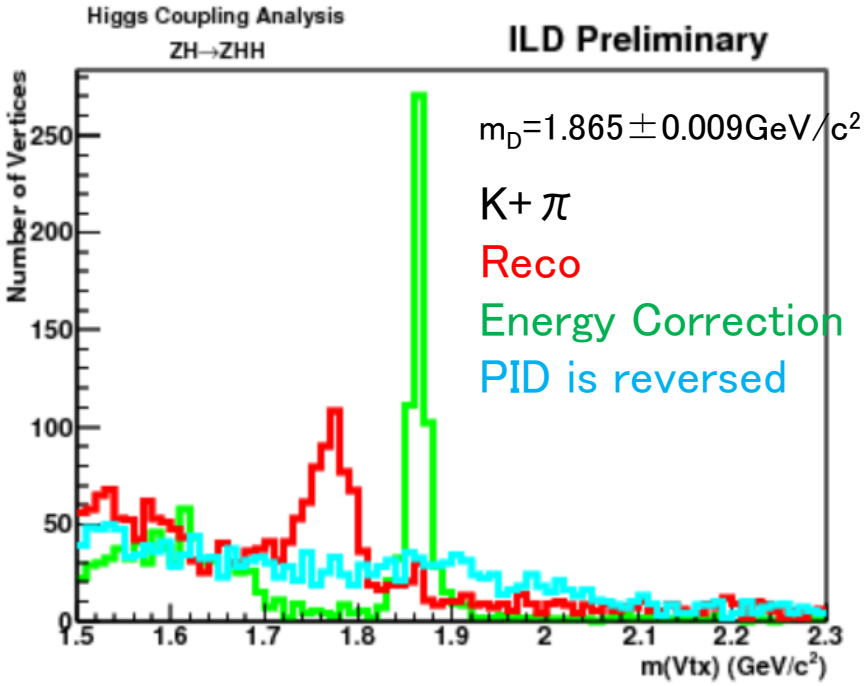
- Can Particle ID be used for flavor tagging improvement?
  - Checking vertex mass distribution
  - Vertex is from LCFIPlus
  - How much effect on vertex mass?



- Check D meson reconstruction
  - Track energy correction using PID
  - How much D meson mass is close to PDG value($1.869 \pm 0.0001\,GeV/c^2$)?
  - How much does wrong PID destroy D meson mass?

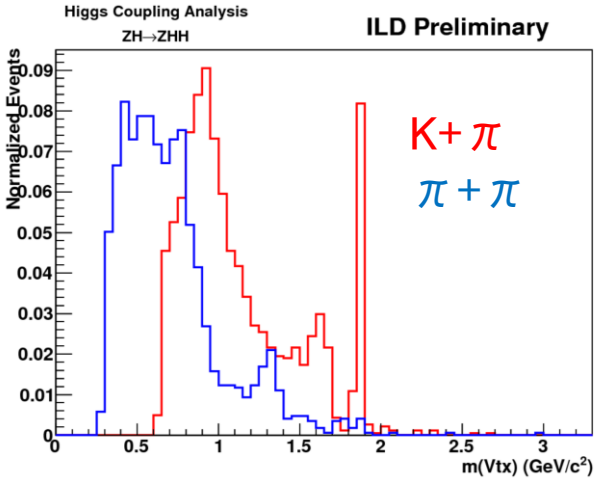
$m_D = 1.865 \pm 3\sigma$ is defined as D meson mass range

status	Inside	outside
PID Correct(num. of vtx)	550	(6940)
PID reversed(num. of vtx)	83	77
Reversed PID is near nominal D mass	22	77

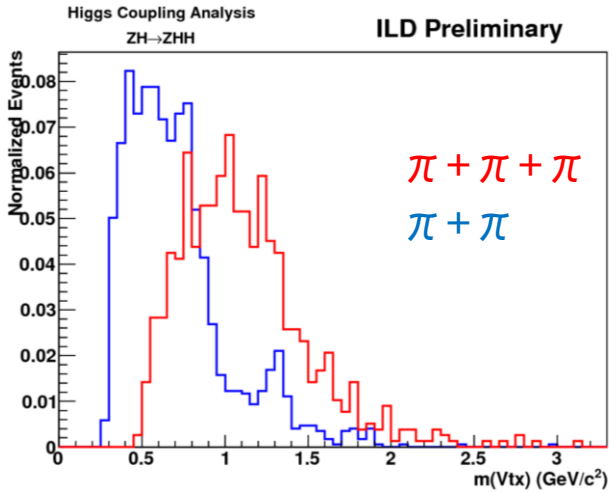


Vertex classification – A clue for next step

- Different vertex patterns have different vertex mass patterns
- e.g. 1) same num. of tracks with different particle patterns
  - $K+\pi$ vs. $\pi+\pi$
  - From third vertex in bjet



- e.g. 2) different num. of tracks with same particle
  - $\pi+\pi$ vs. $\pi+\pi+\pi$
  - From third vertex in bjet



Hope for flavor tagging improvement

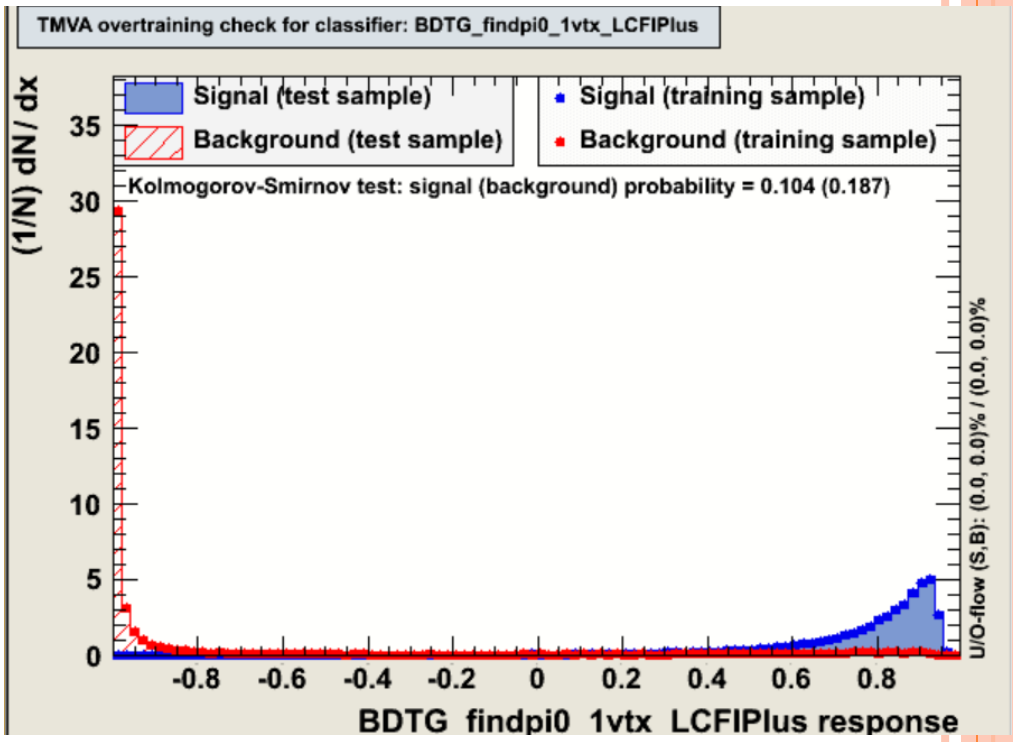
- For flavor tagging improvement
  - Vertex mass is the key to separate heavy/light flavor vertex
  - Many pi0s will escape from B/D vertex → checked that using MC truth
  - Mass resolution will be degrade due to escaping neutrals
  - Is there possibility to recover pi0s which escape from vertices?

- We are studying the possibility of vertex mass recovery using pi0s
  - Pi0 vertex finder – which vertex is the π^0 coming?
- Finding vertex of pi0s
  - Very difficult to identify vertex – depends on detector configuration
  - Making the best of decay kinematics
  - Using TMVA to find pi0 candidates from the vertex
  - Comparing vertex mass distribution
  - Sample: using qqHH@500GeV samples(so many tracks & pi0s in events)

- Goal: flavor tagging efficiency improvement!

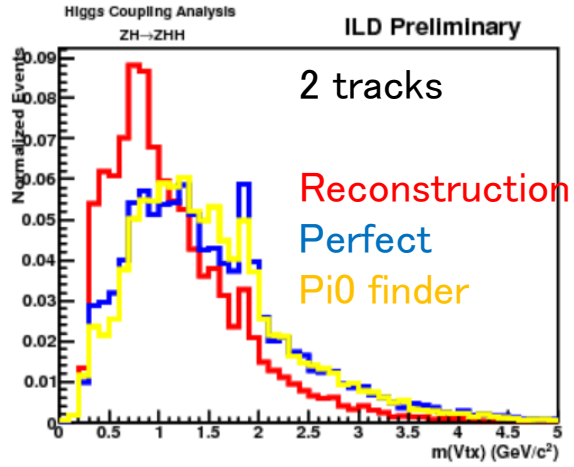
Input variables to construct a general classifier

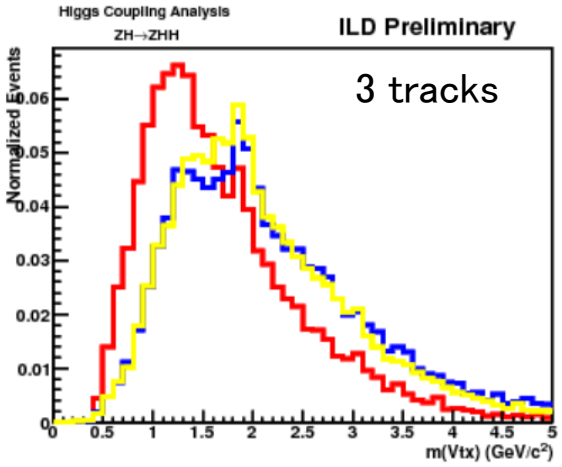
- Getting general – num. of particles are used as input variables
  - Num. of e/ μ / π /K/p in the vertices – using particle ID
  - Those variables will work as variables for vertex classification in the MVA classifier
- I have constructed the 3 types of MVA classifiers:
  - Jets with 2 vtx – secondary and third vertex
  - Jets with 1 vtx – secondary vertex only
  - Using b jets
- MVAoutput example
  - Set operation point using num. of pi0s to be attached to vertices

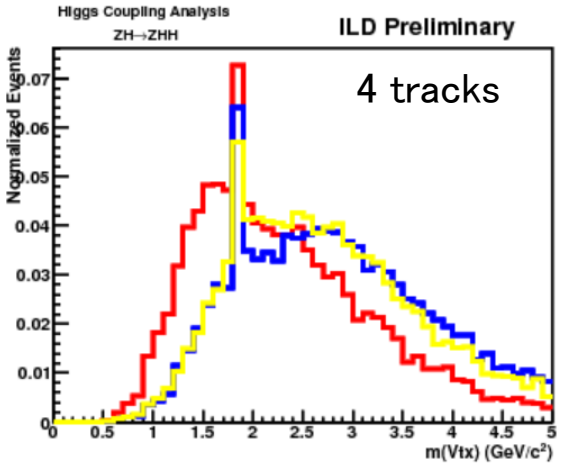


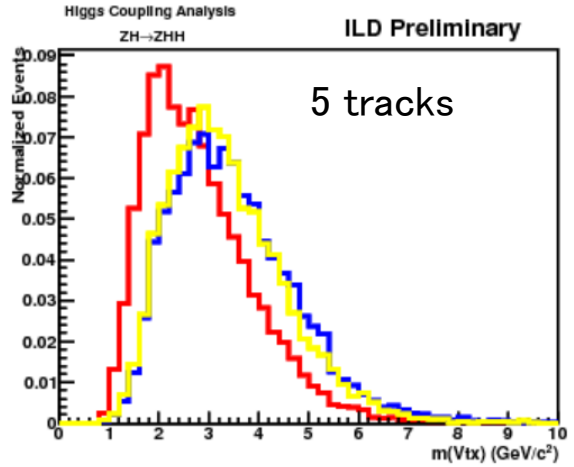
Vtx masses

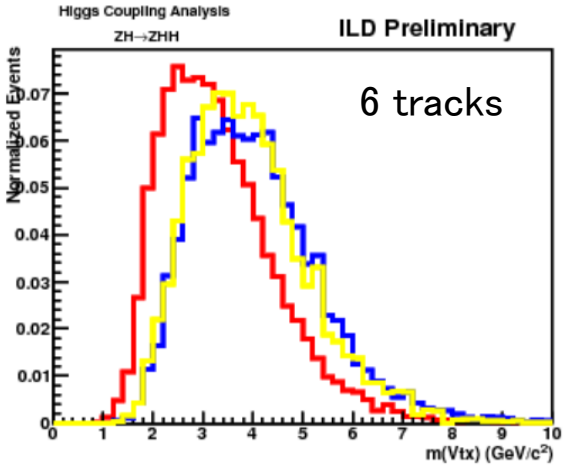
- Vtx mass distributions for each vertex pattern(ntrk)
  - not so bad
  - Difference is mainly coming from combinatorial problem of gamma pairing

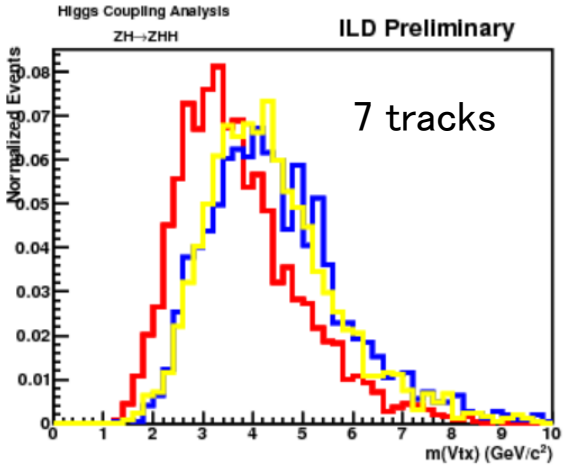






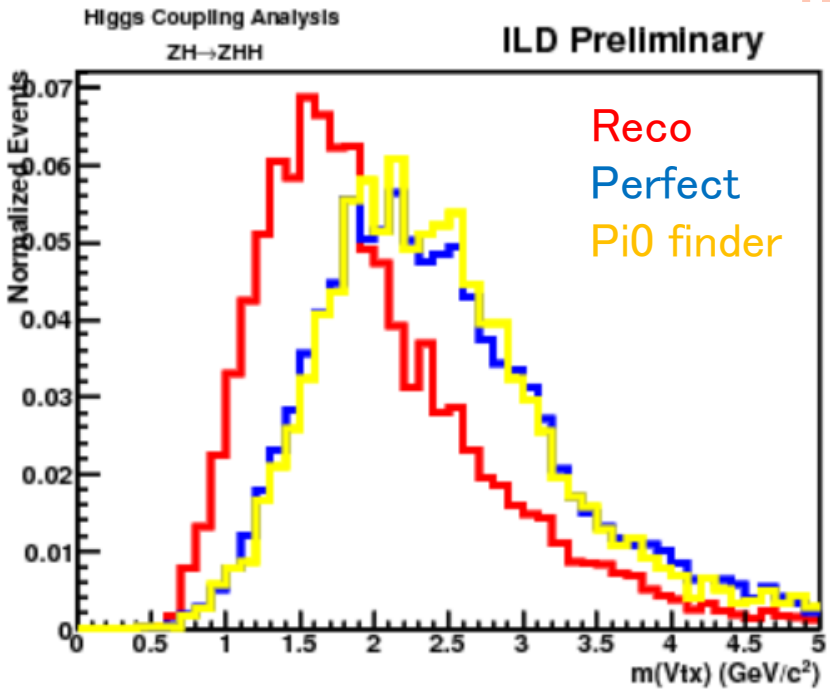




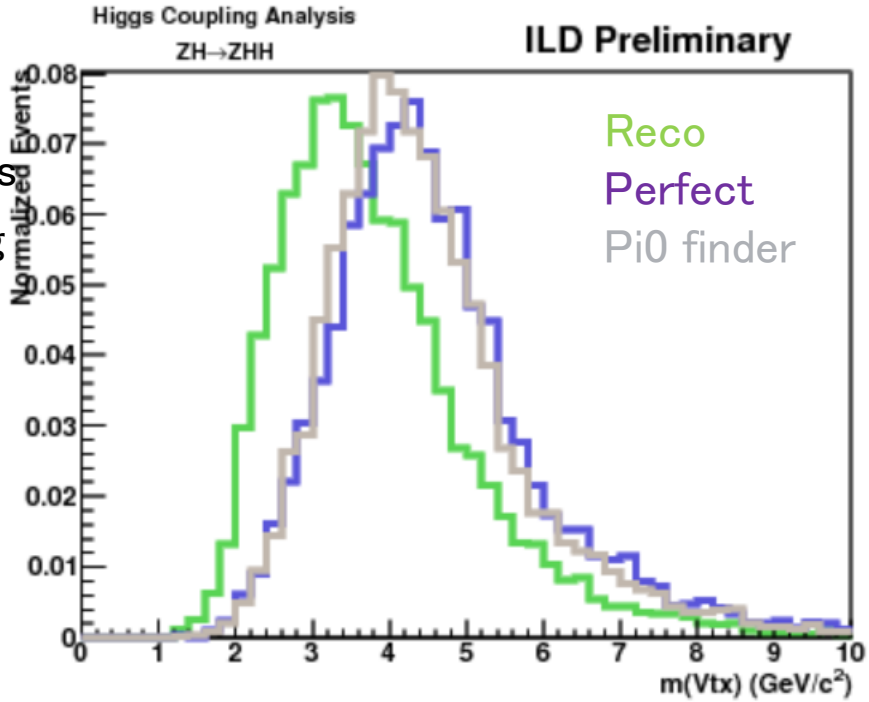


GLANCE AT OTHER CASE

- 2 vertices in bjet
  - Secondary vertex – 4tracks case

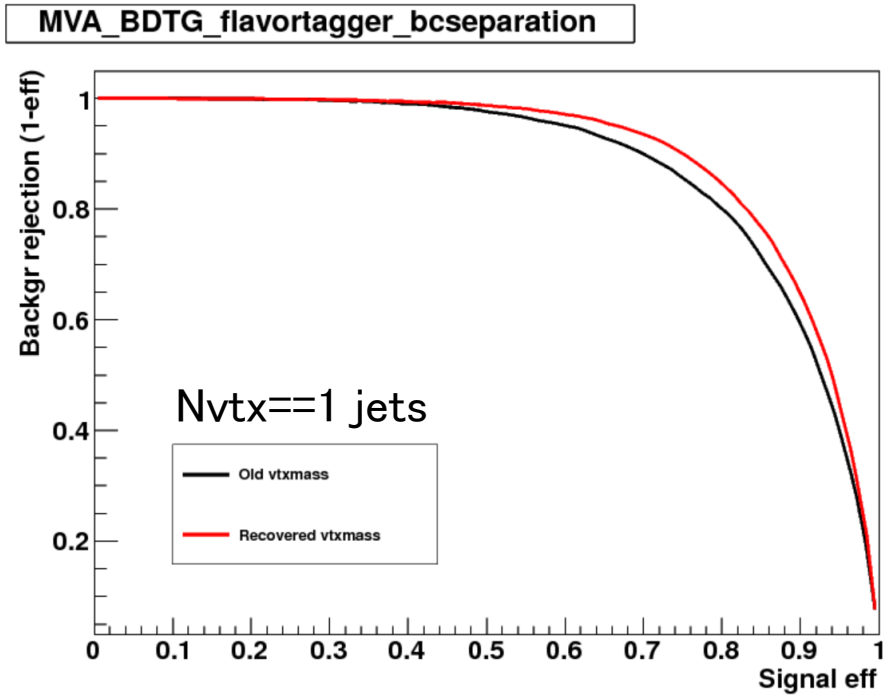


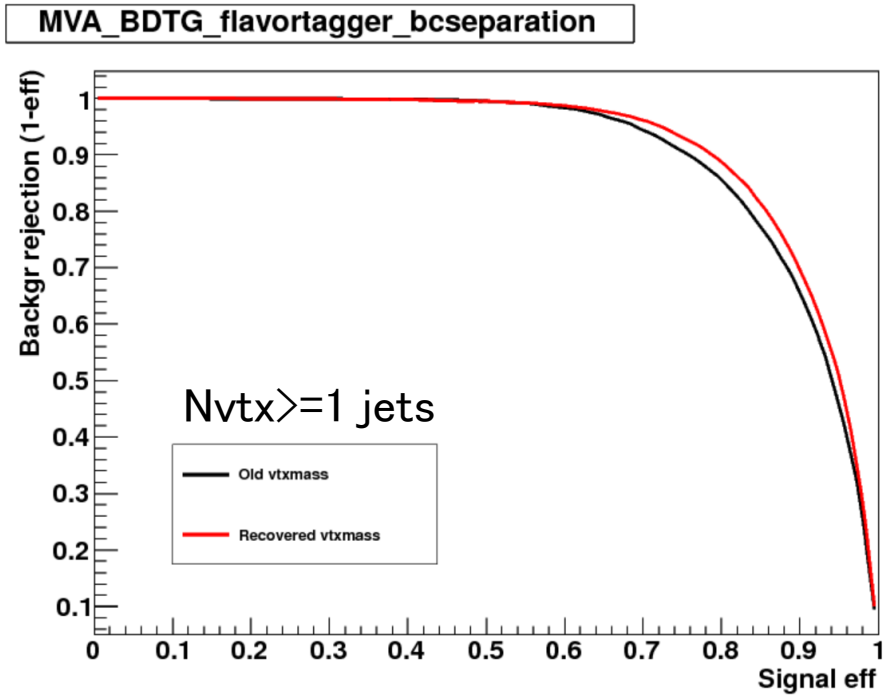
  - Merging with third vertex
    - Third vertices allow all the track patterns
    - Attach pi0s to both of the vertices using pi0 vertex finder



Vertex mass recovery effect on flavor tagging

- Construct a “toy” flavor tagger
  - Input variables are obtained from LCFIPlus
  - Input variable selection is too primitive!
  - Only vertex mass is replaced to recovered vertex mass
  - Compare with ROC curve





- For more precise study, need to step into LCFIPlus

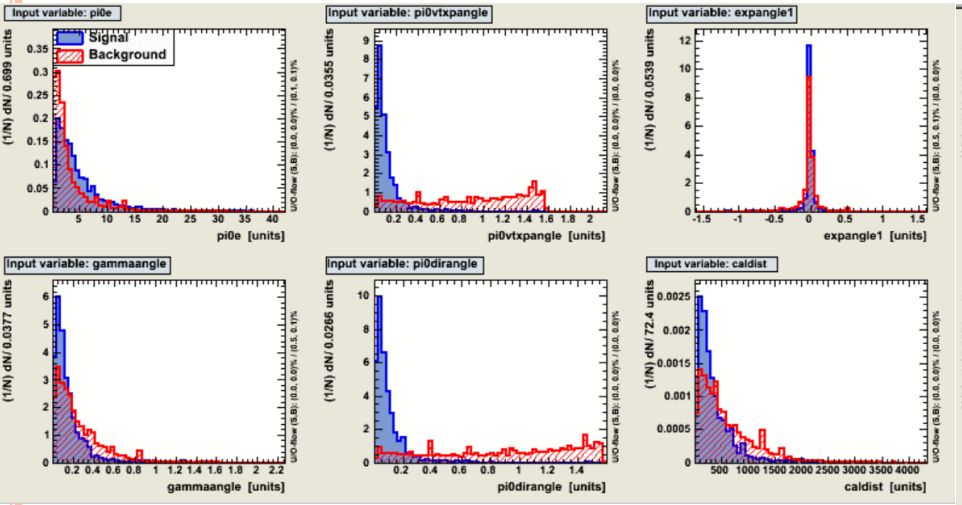
Summary, Problems and Prospects

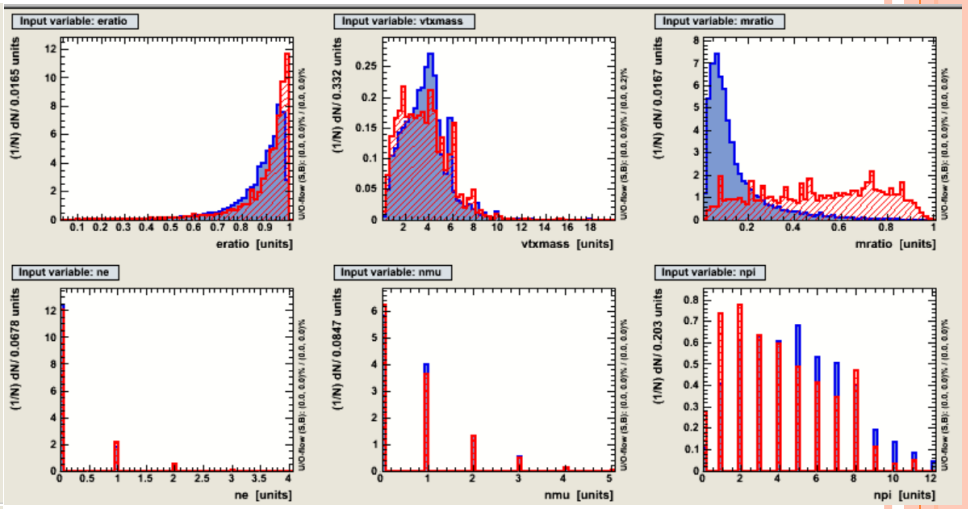
- Explore some fundamental variables for analysis improvement
  - dE/dx in TPC and shower profile
- dE/dx and shower profile information provide ～30% improvement for Isolated lepton ID@10-20GeV/c
- Studying particle ID:
  - Hadron ID eff. is 62%～75%
  - Energy correction effect is very good for D meson reconstruction
  - Vertex mass recovery is hopeful using particle ID
- Flavor tagger improvement:
  - There seems hope for attaching pi0s to vertices
    - Vertex mass recovery is reasonable
    - Of course, many checks are necessary
- Vertex mass recovery will provide better separation on b/c jets!
  - Recovered vertex mass seems to bring better flavor tagger!
  - Finally, check the flavor tagging effs. in LCFIPlus!
- What is a next step using PID? – advantage of TPC!

BACKUPS

MVA – USING TMVA

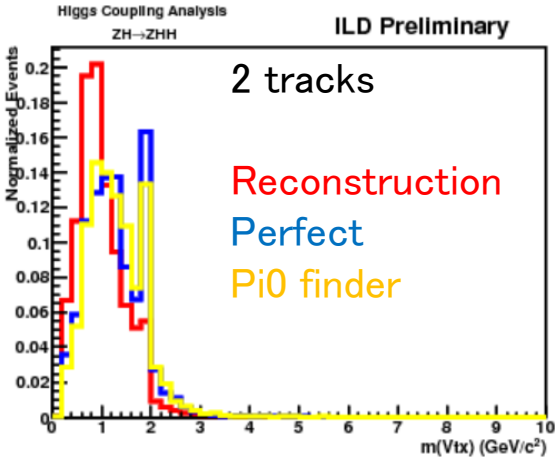
- Input variables to be used
  - Secondary vertices which don't have third vertex

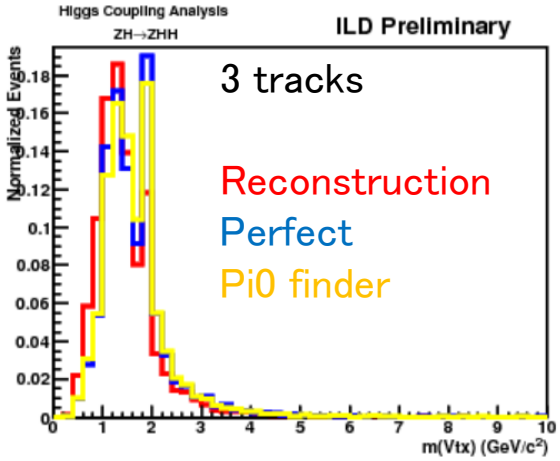


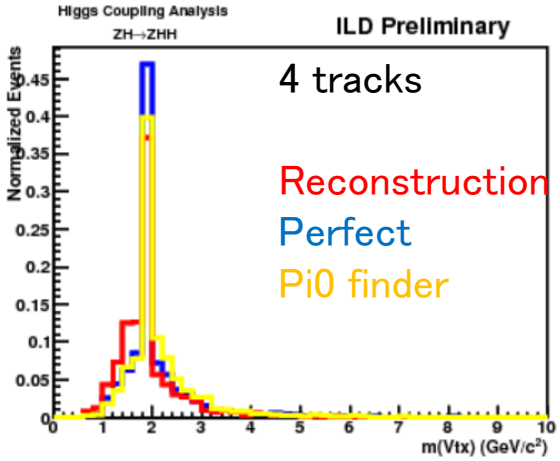


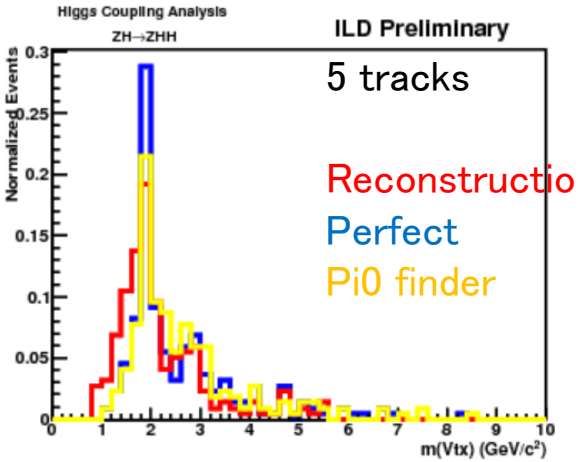
Testing of c vertex case

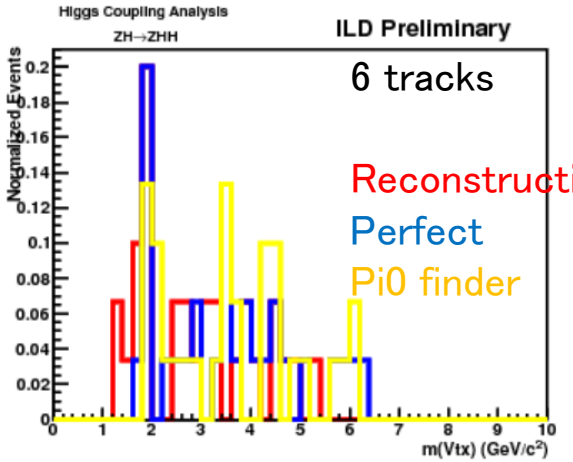
- Attaching pi0s to c vertex using same classifier
  - So far, no strange behavior

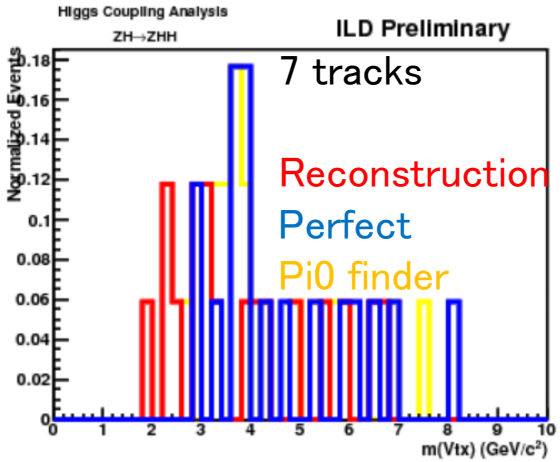






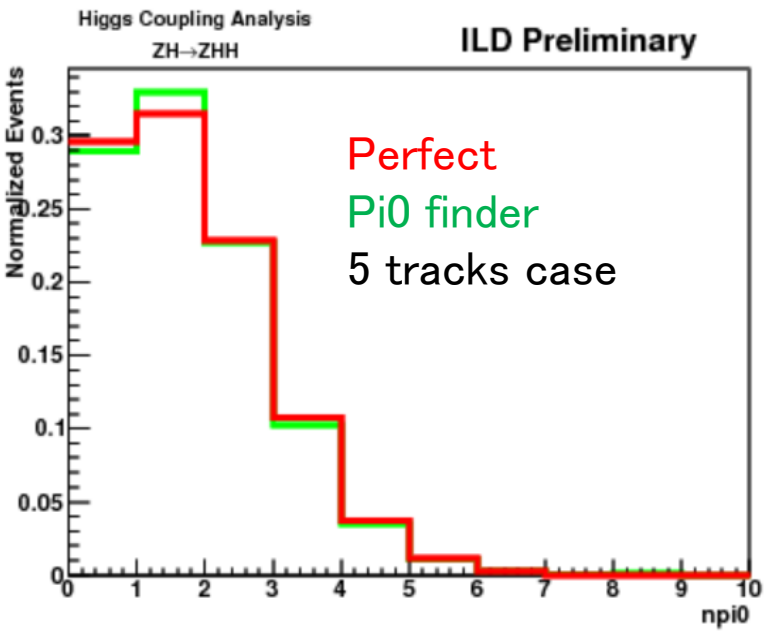


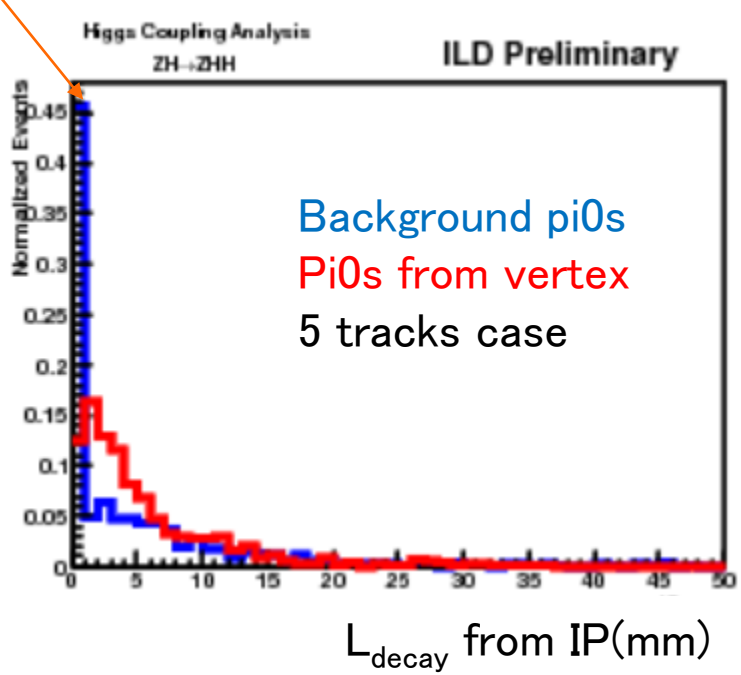




Some plots

- Num. of pi0s to be attached →determine MVAcut by it
- Where do pi0s really come from?
  - Many pi0s from primary are mis-attached to the vertices
  - Now, that is limited by detector configuration(can't determine exact gamma direction)
  - To some extent, an idea to catch gamma direction is necessary





The most realistic situation

- After an event occurs, we only measure:
  - Charged particle information – 4-momentum, and particle type(PID)
  - Neutral particle information – 4-momentum of gamma or stable hadrons
  - We have no direct information of pi0s
- We need to get pi0 information from gammas!
  - Gamma finder – choosing gamma candidates from neutral particles
  - Pi0 reconstruction – gamma pairing from gamma candidates
- In such situation, how is the vertex mass recovery?
  - How is neutral hadron contamination effect?
  - How is gamma mis-pairing effect?
- About pi0 reconstruction, I have already talked at previous talk
- By using that pi0 reconstruction, attaching pi0 candidates and compare the vertex mass