LCFIPlus

BASIC IDEA AND TOWARDS IMPROVEMENT

Masakazu Kurata

The University of Tokyo

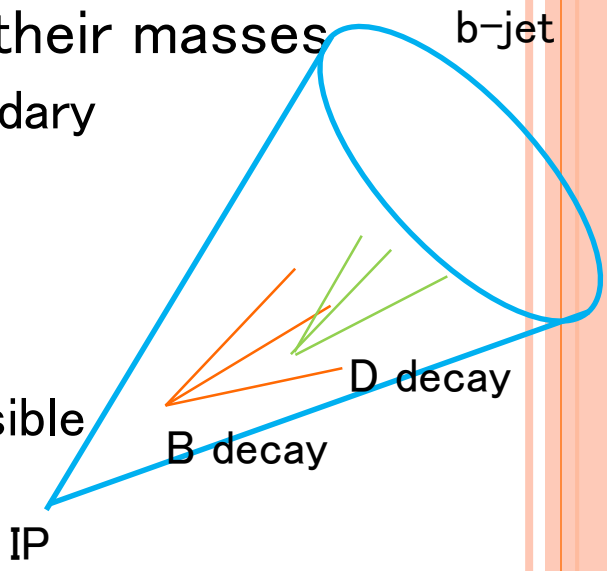
LCWS15, 11/02/2015-11/06/2015

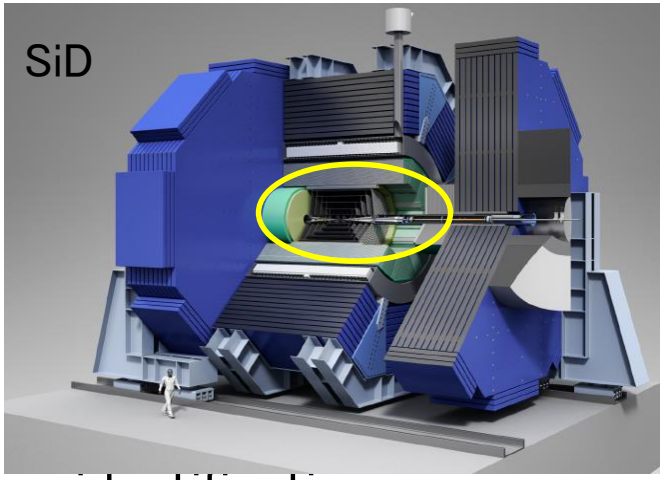
Introduction

- Flavor tagging is one of important analysis components for all the ILC physics programs:
  - For Higgs physics: H→bb, H→cc
  - For top physics: t→bW
  - For BSM: HA→bbbb
- Important topics for good flavor tag:
  - Well reconstruct secondary and tertiary vertices from B/D decays
  - Good operation in the environment with many jets
  - Good separation among b/c/uds jets
- LCFIPlus package has been developed:
  - Consists of:
    - Vertex finder – primary and secondary/tertiary vertex finding
    - Jet finder(Jet Clustering)
    - Flavor tagger
  - Optimized for ILC(e.g. Higgs self-coupling@500GeV)

Requirements for Flavor tagging

- B/D meson can fly before their decay due to their masses
  - We can identify heavy flavor jets by finding secondary /tertiary vertices in jets:
- So, we need:
  - Good impact parameter resolution
  - Good quality track measurement as much as possible
  - High purity lepton identification
- ILD/SiD have:
  - Vertex detector: good impact parameter measurement(<5 μ m)
  - Main tracker (ILD: TPC SiD: Silicon tracker): good momentum resolution
  - Calorimeters can be used for lepton identification





Flow chart for Flavor tagging in LCFIPlus

Primary Vertex Finder
- Good track selection for vertexing
- Find Primary vertex

Secondary Vertex Finder
- Find secondary vertex candidates

JetFinder(Jet Clustering)
- Cluster PFOs into Jets

Jet Vertex Refiner
- Integrate vertices in jets into upto 2

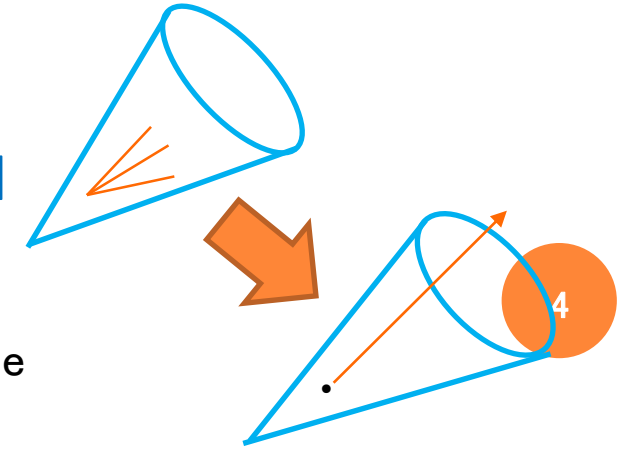
→corresponds to B/D vertices

Flavor Tagging
- Jet flavor is evaluated

→MVA is used for flavor separation

- Vertex finding strategy of LCFIPlus:
  - **Vertex finding first, Jet finding second**
  - All the vertex candidates in a event are checked
  - Jet finding with replaced particles

→ Reconstructed vertices are regarded as one particle



Impact on Vertex finding Strategy

- LCFIPlus: Vertex finding first, Jet finding second
  - This strategy provides:
    - Fewer fake track rate(coming from primary vertex)
    - Better vertex finding efficiency(both B/D vertices)
    - V0 rejection efficiency is improved

(a) $ZHH \rightarrow qqbbbb$	Track origin			
	Primary	b hadron	c hadron	Other
Number of all reconstructed tracks	67575	12912	15246	4087
Number of tracks used by ZVTOP	1162	8534	10404	999
...in *good* vertices	-	8248	10103	-
Number of tracks used by our original vertex finder	617	8717	10529	358
...in *good* vertices	-	8551	10333	-

(b) $t\bar{t} \rightarrow bbqqqq$	Track origin			
	Primary	b hadron	c hadron	Other
Number of all reconstructed tracks	74504	8945	12602	4219
Number of tracks used by ZVTOP	920	5999	8353	1024
...in *good* vertices	-	5830	8137	-
Number of tracks used by our original vertex finder	420	6161	8447	341
...in *good* vertices	-	6060	8279	-

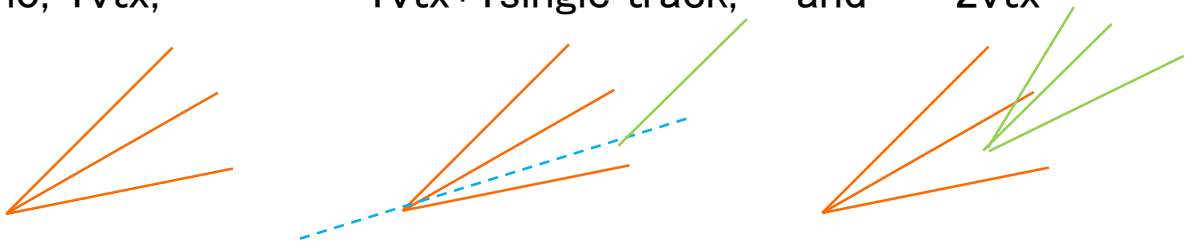
LoI

LCFIPlus

FLAVOR TAGGER

- Use Multivariate Analysis(MVA)
  - Classifier: Boosted Decision Tree(BDT)
  - For Flavor separation of b/c/uds jets
  - Categorize using vertex condition in a jet and train independently

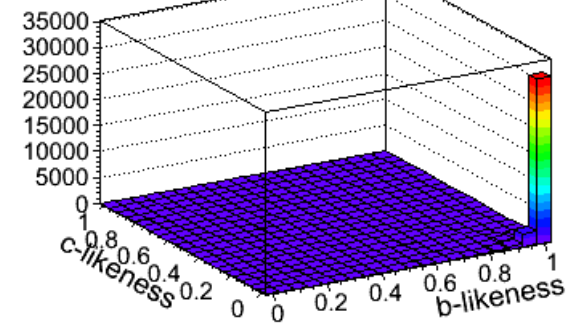
  →This categorization is coming from JetVertexRefiner
    - Vtx: no, 1vtx, 1vtx+1single track, and 2vtx



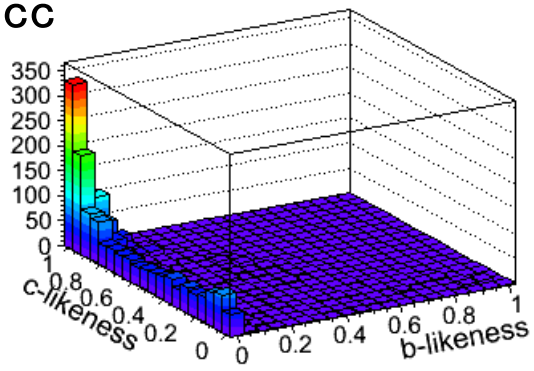
  - 2 output type: b-likeliness and c-likeliness

- Example: distribution of each Higgs decay mode:

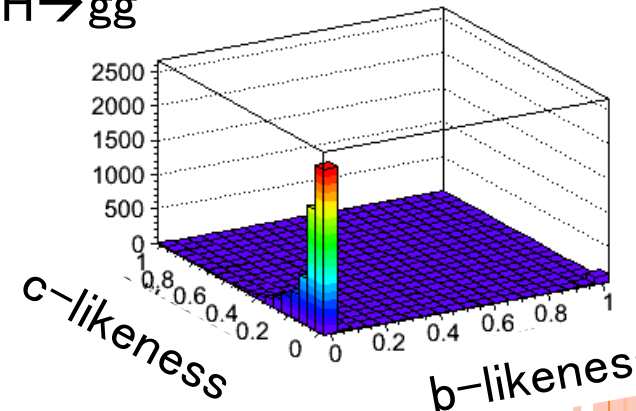
H→bb



H→cc

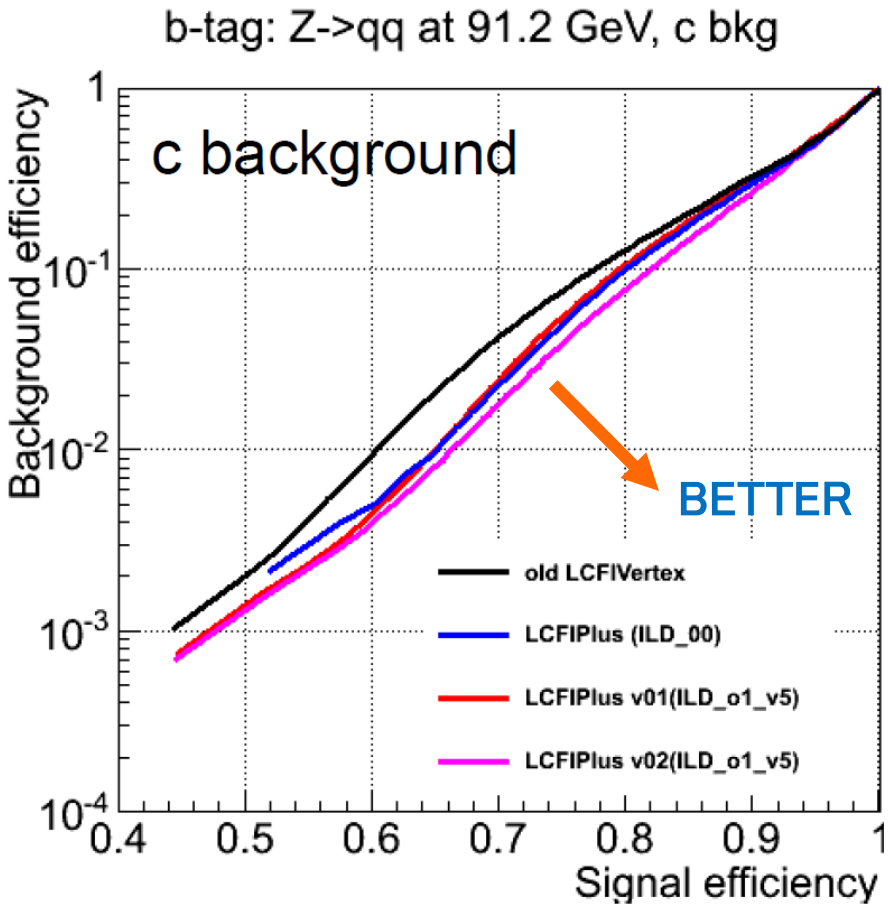


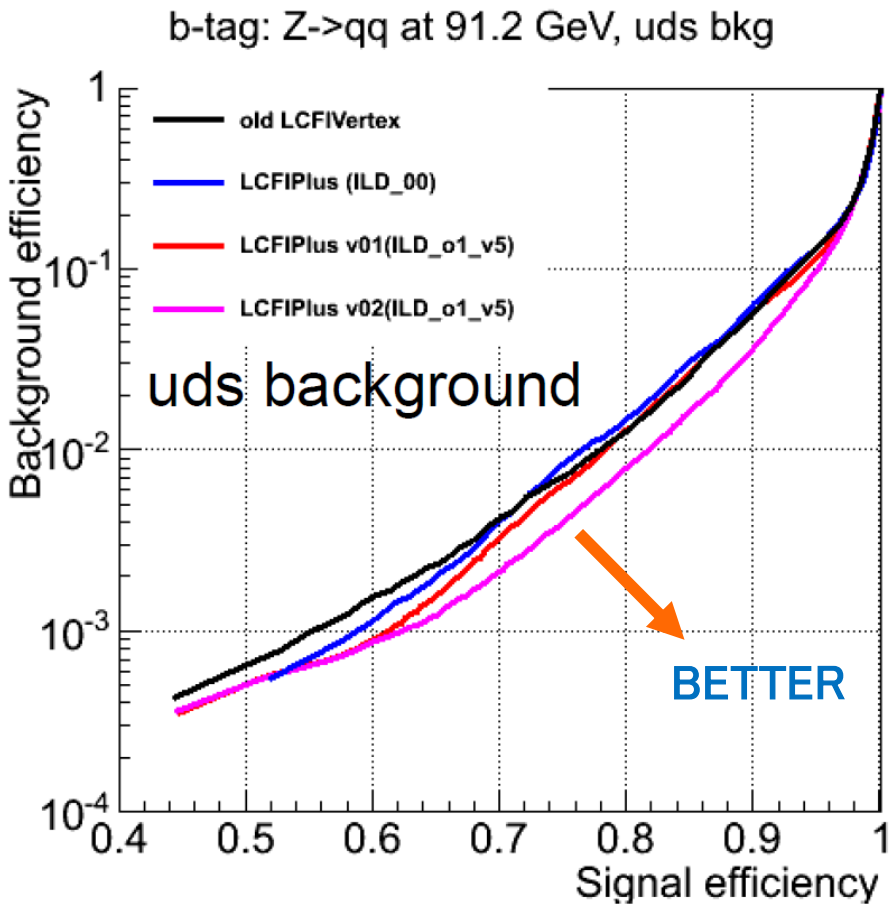
H→gg



PERFORMANCE@Z MASS

- Signal efficiency vs. background efficiency

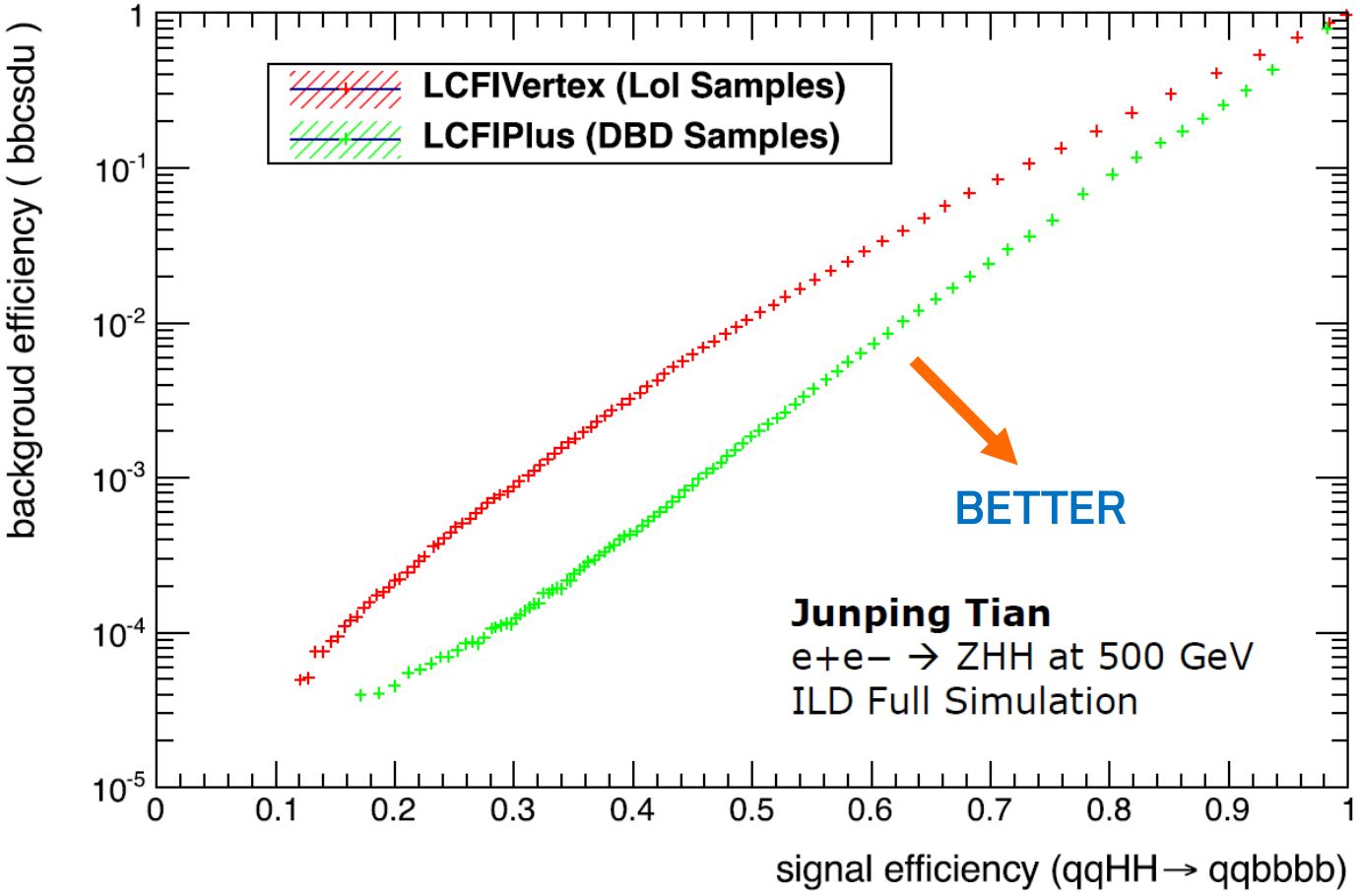




- Can form better classifier for any operation points

Impact on Higgs self-coupling

- B-tagging is crucial to suppress ttbar events in Higgs Self-coupling analysis
  - ttbar has huge cross section(S/N～1/3000 @500GeV)
  - Suppression using kinematics is not enough
  - HH→(bb)(bb): 4 b-tagging available　ttbar: tt→bb(xxxx) 2 b-tagging



Further Plans for flavor tagging Improvement

LCFIPLUS IMPROVEMENT

- For better flavor tagger, we need
  - Find secondary/tertiary vertices as many as possible
  - Perfect reconstruction of B/D meson mass
    - Perfect attachment of charged particles
    - Recover lost component(neutrals)
  - Better variables which can separate jet flavors

- We found that we can acquire flavor tagging improvement:
  - Vertex finding efficiency improvement → introduce a new algorithm
  - Vertex Mass Recovery → using escaping π^0s
  - Better flavor separation(don't mention in this talk)

- Much help is necessary!:
  - Particle ID is one of the key to flavor tagging improvement
    - To classify vertices
  - Pi0 reconstruction(γ pairing) is other key for vertex mass recovery
    - First of all, pi0 is necessary!

Adaptive Vertex Fitting

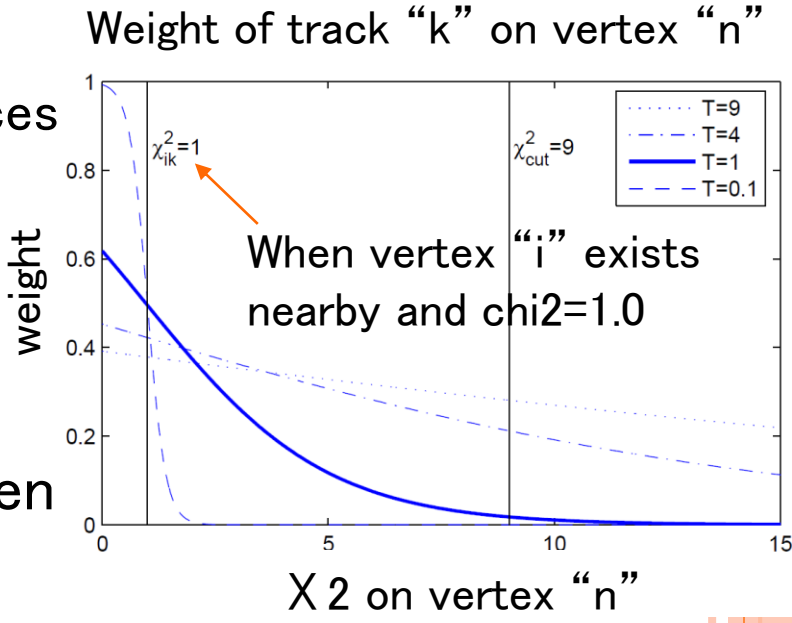
- To introduce the effect of multi-vertex fitting
  - Introduce weight function to estimate vertex which a track belongs to
  - Weight function definition: k-th track's weight on n-th vertex

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

  - Parameter: temperature T
    - If T very small, decision is like χ 2 minimization(almost same as DBD LCFIPlus)
    - If T large, multi-vertex effect becomes large
  - In multi-vertex environment, weight on certain vertex will degrade

→becomes harder to attach tracks to vertices in multi-vertex environment

→can reject fake tracks well!



- Thanks to weight function, we can loosen the track quality selection

→vertex finding eff. will be improved!

Impact of Adaptive Vertex Fitting

- 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 with vertex:

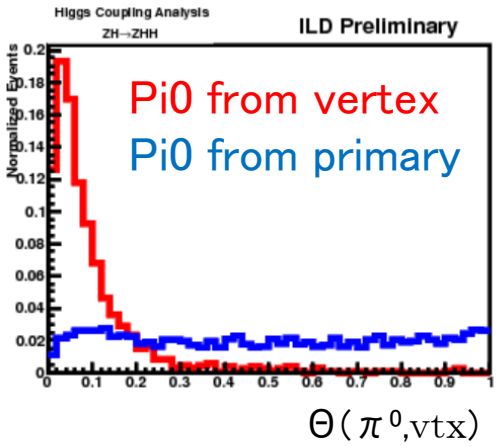
method	bjet with 2vtx	bjet with 1+1vtx	bjet with 1vtx	total
DBD LCFIPlus	10586	9111	12844	32541
AVF	13179	6360	13375	32914

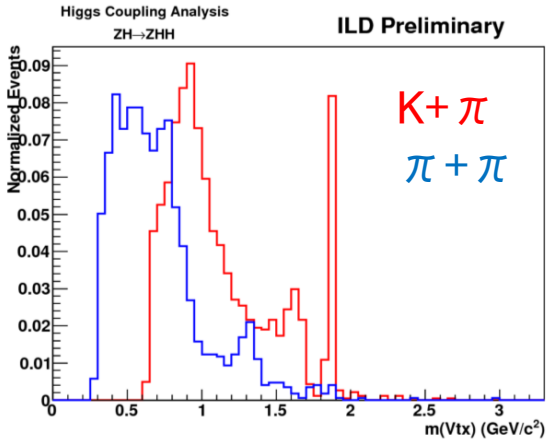
- Total jets with vtx: ～1.1% increased
  - Jets with 2vtx: ~21% increased → good for bjet ID!
  - Jets with 1vtx: ~3% increased → good for uds jet separartion!
- Fake track rate per vtx: how many fake tracks contaminate vertices?
  - Almost same – slightly better!

method	bjet with 2vtx	bjet with 1+1vtx	bjet with 1vtx
DBD LCFIPlus	0.029±0.001	0.013±0.0012	0.058±0.002
AVF	0.028±0.001	0.008±0.0008	0.058±0.002

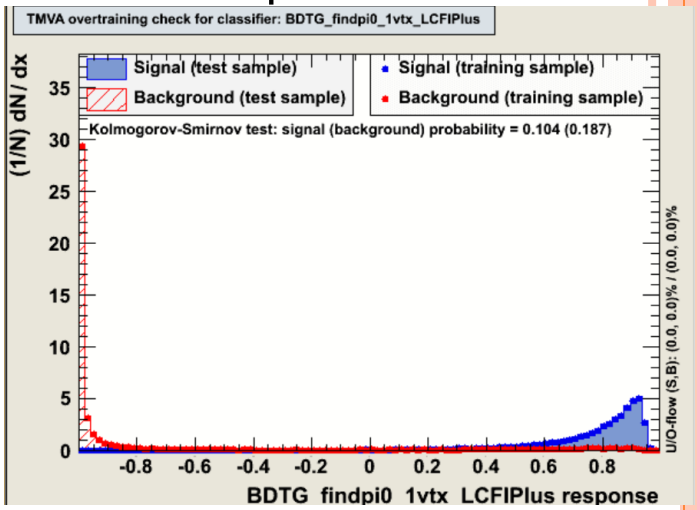
Vertex mass recovery using pi0s

- 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



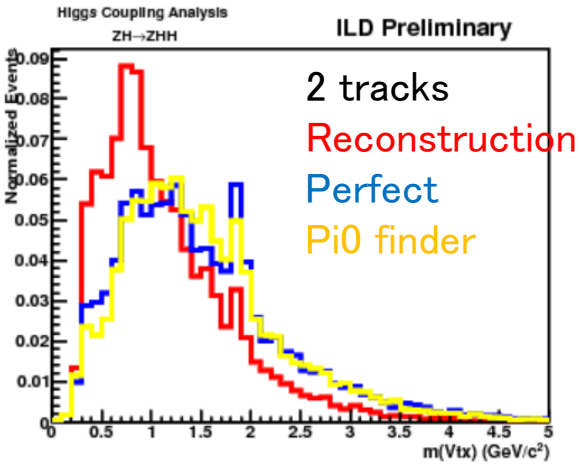


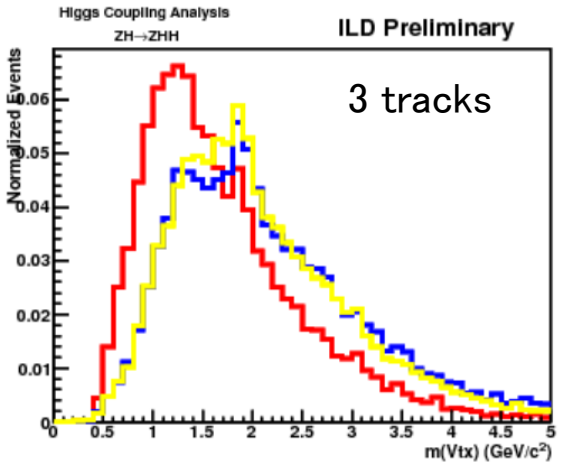
- Particle ID is the other key to classify vertices
  - Different particle patterns have different vertex mass patterns
- Construct Pi0 Vertex finder using MVA
  - Identify which vertex pi0s are coming from

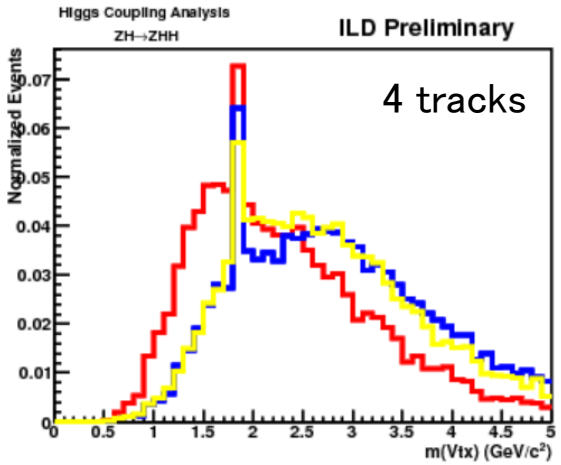


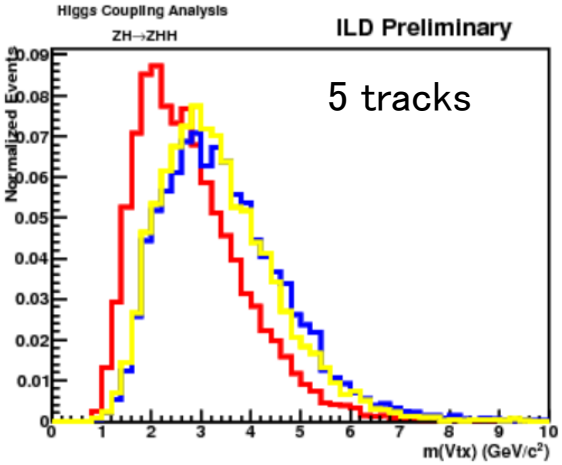
Vtx masses of bjets in Higgs self-coupling

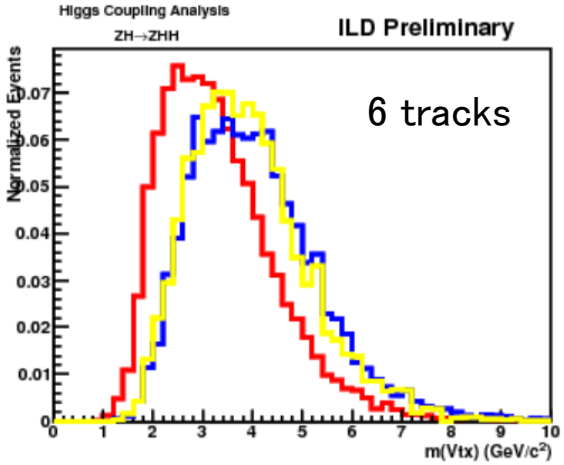
- Vtx mass distributions for each vertex pattern(ntrk)
  - These results are the outputs of LCFIPlus(unofficial ver.)!
  - Difference is limited by mis-pairing of gammas and mis-attachment of pi0s

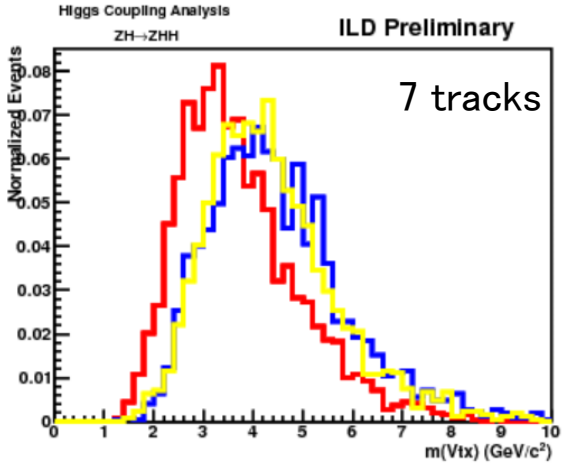






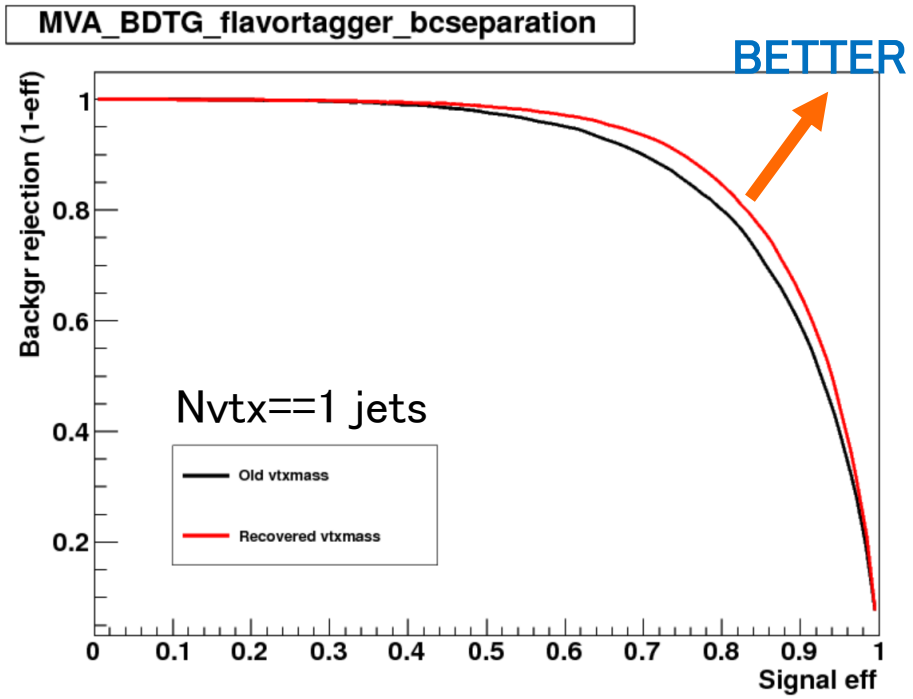


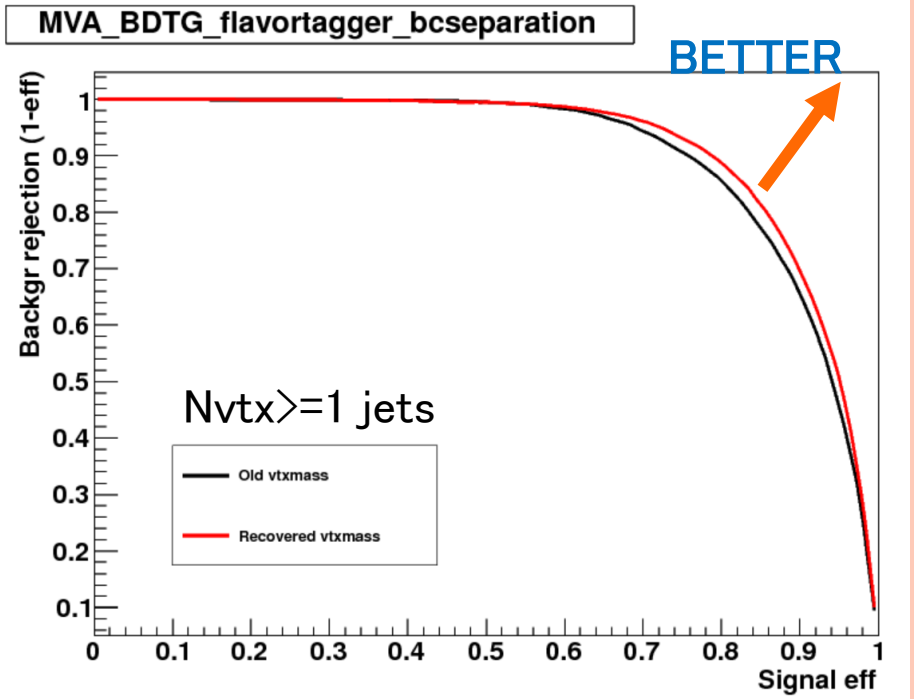




Vertex mass recovery effect on flavor tagging

- 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





- Vertex is created using DBD LCFIPlus vertex finding
  - need to check AVF case

Summary and prospects

- LCFIPlus has been successful for DBD physics analysis
  - Vertex finding first, jet finding second approach provides
    - Better vertex finding
    - Better fake Rejection
    - Finally, better jet flavor separation!

- Prospects:
  - There is much room to improve!
  - So far, AVF will provide ～1% improvement of vertex finding efficiency in bjets
    - Jets with 2vtx well increased – better for b jet ID!
    - Fake rejection will be same – slightly better!
  - Vertex mass recovery using pi0s is reasonable
    - Will provide better flavor tagger using recovered vertex mass
    - Pi0 reco. Improvement will give better vertex mass recovery!
  - Finally, incorporate all the ideas and check the final flavor tagging effs.in LCFIPlus!

Back ups

History of flavor tagger development

- ～2009 LCFIVertex [NIM A 610 573]
  - Used in ILC LoI and CLIC CDR physics analyses

- 2010～ LCFIPlus has been developed
  - Used in DBD physics analyses for both ILD and SiD
  - Reference: 1506:08371
  - NIM paper will be soon!

- Working group was re-organized for further improvements
  - Working for legacies from DBD ver. LCFIPlus
  - Tackling for flavor tagging improvement

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

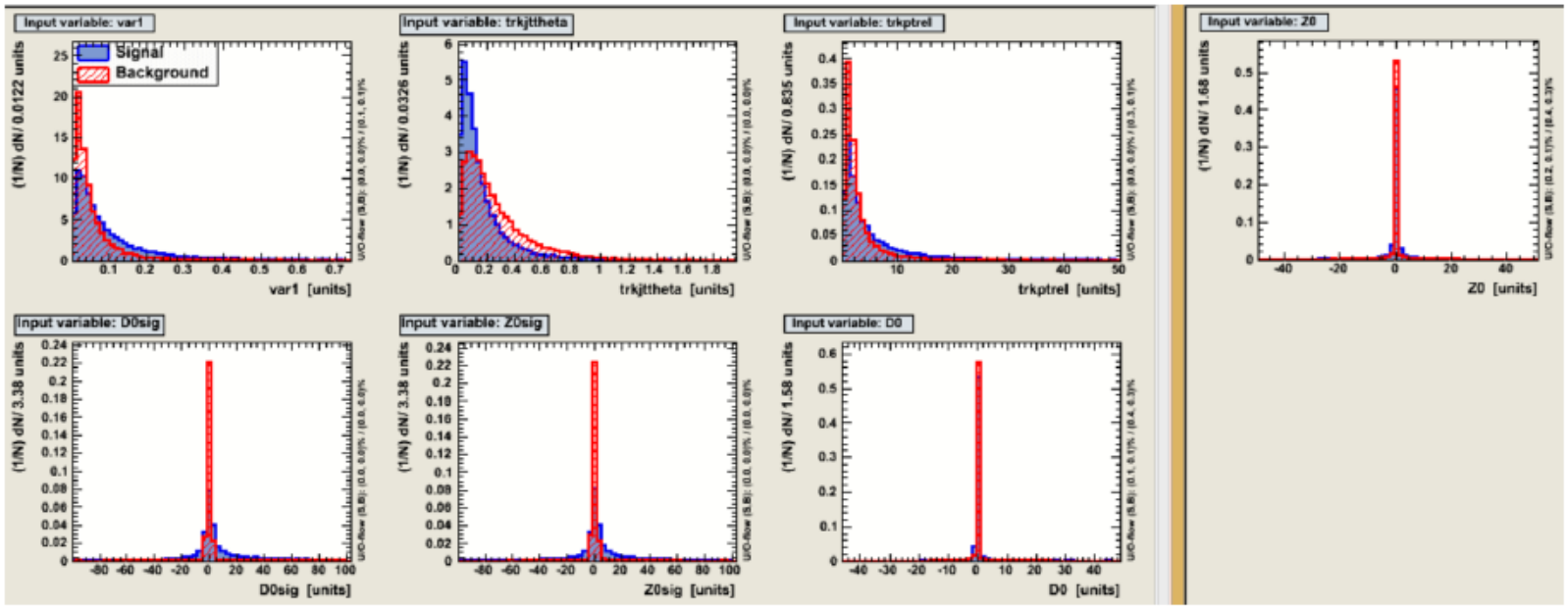
method	cjet with 2vtx	cjet with 1+1vtx	cjet with 1vtx	total
DBD LCFIPlus	48	149	6261	6458
AVF	59	141	6327	6527

  - Total: ～1% 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
BDB LCFIPlus	0.00±0.00	0.012±0.006	0.0014±0.004
AVF	0.00±0.00	0.018±0.007	0.0013±0.004

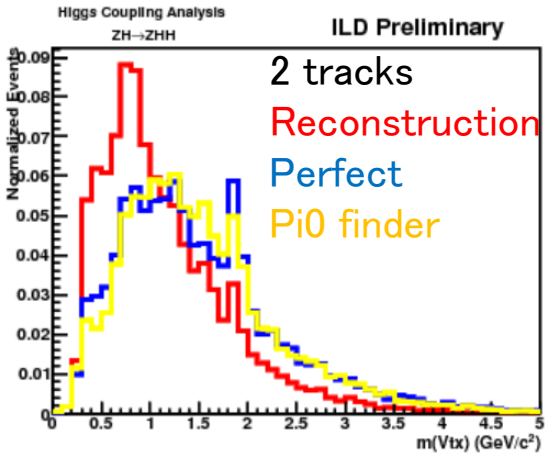
TRACK MVA(BNESS)

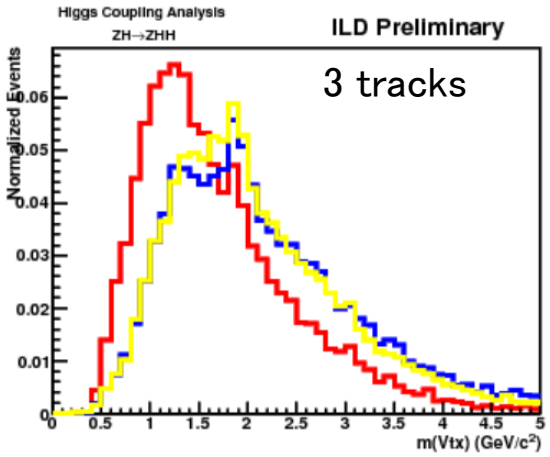
- To identify track which comes from heavy flavor particle
  →using MVA
  - Signal: tracks which come from B mesons or B baryons
  - Background: tracks produced in hadronization process
- Most significant tracks with both plus and minus signed impact parameters in a jet are collected
  - Significance: $sig = \sqrt{(\frac{d_0}{\sigma})^2+(\frac{z_0}{\sigma})^2}$

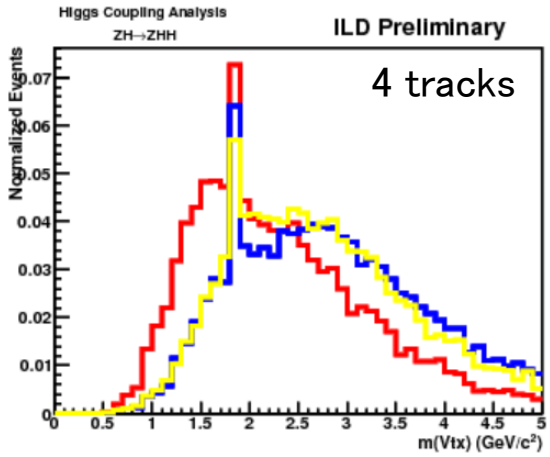


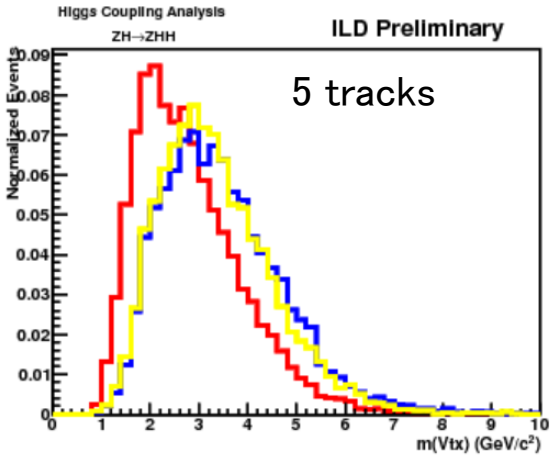
VTX MASSES

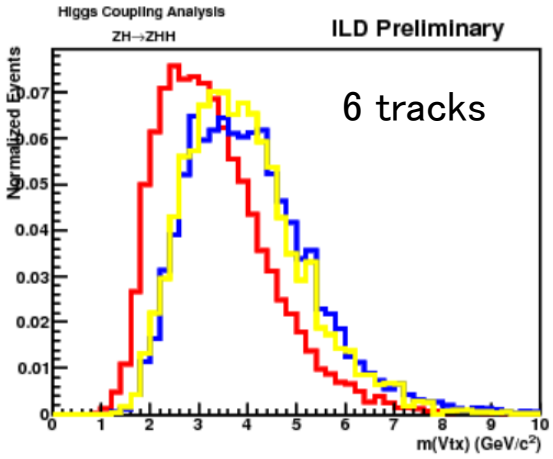
- Vtx mass distributions for each vertex pattern(ntrk)
  - not so bad
  - Difference is coming from mis-pairing of gammas and mis-attachment of pi0s

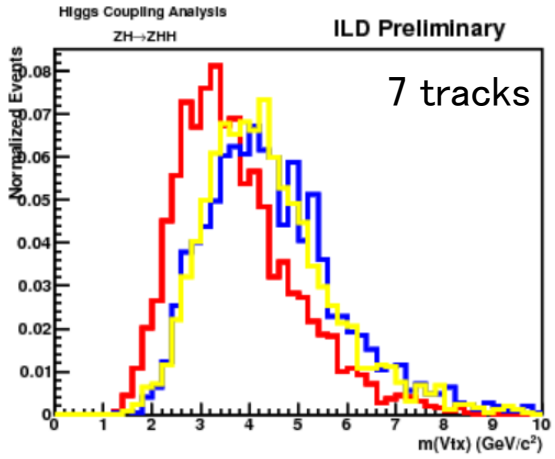






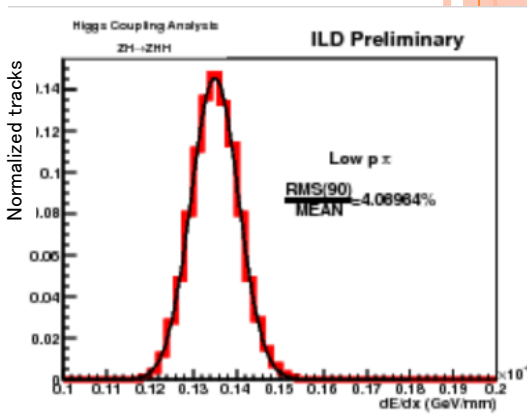




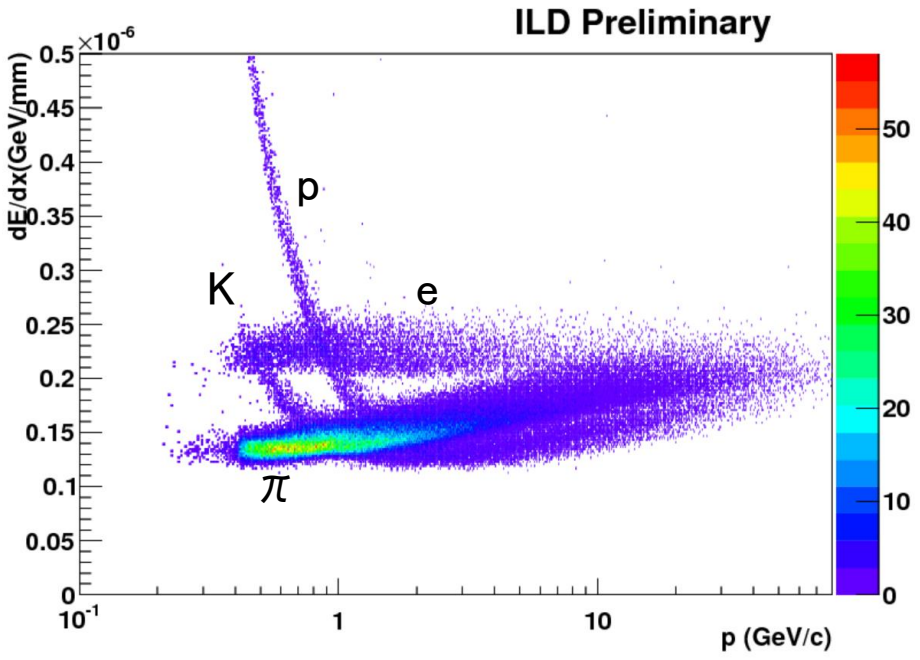


dE/dx from TPC

- Fluctuation of dE/dx using various type of tracks
  - Truncation method is used to avoid landau tail
  - Fluctuations of each particle/each momentum range in simulation: **3 – (<5)%!!** **TDR goal: 5%**
  - Including detector effect is necessary



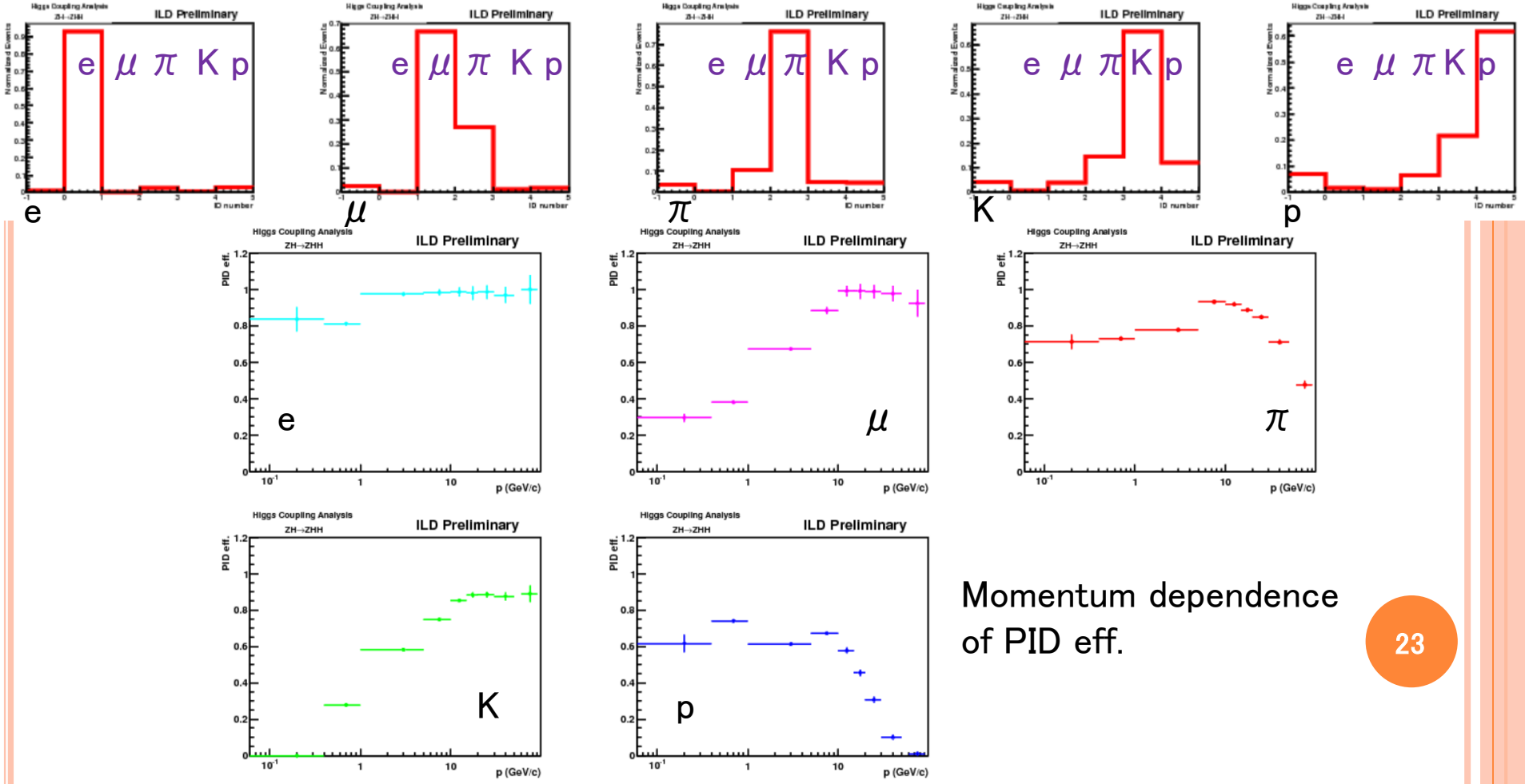
- Momentum dependence of dE/dx for each particle
  - Polar angle dependence corrected
  - Num. of Hits dependence corrected
  - Scale to $\left\langle \frac{dE}{dx} \right\rangle = 1.0$ for MIP pion



- Can check detector R effect on dE/dx easily(But, detector effect necessary?)

Particle ID

- New variables make Particle ID available –construct Particle ID
- Overall ID efficiency – using 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%



Momentum dependence of PID eff.