

IDEAS FOR FLAVOR TAGGING IMPROVEMENT

Masakazu Kurata

The University of Tokyo

06/10/2015

CURRENT STATUS FOR LCFIPLUS IMPROVEMENT

- There are much room to improve flavor tagging:
 - Vertex finding itself
 - Vertex mass recovery using π^0 s
 - Flavor separation in 0vtx jet case
 - etc. etc...

- We are trying to introduce some ideas:
 - Using Adaptive Vertex Fitting
 - Using BNess tagger
 - Using Particle ID & π^0 reco. for vertex mass recovery

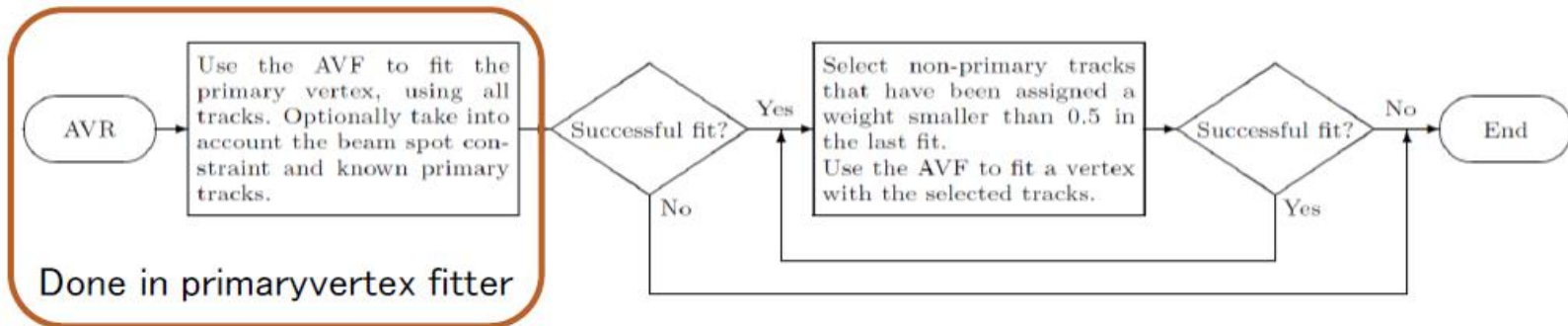
- Summarize current status of those studies and show some performances

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

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

- 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



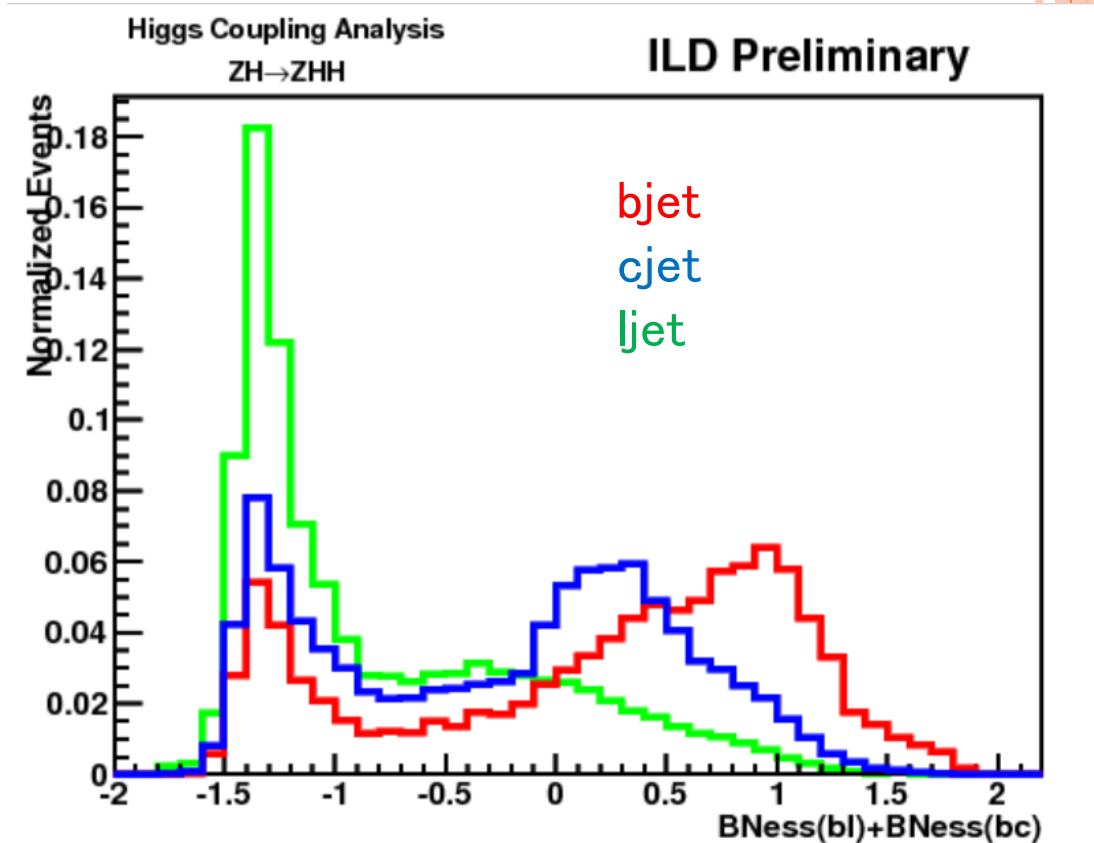
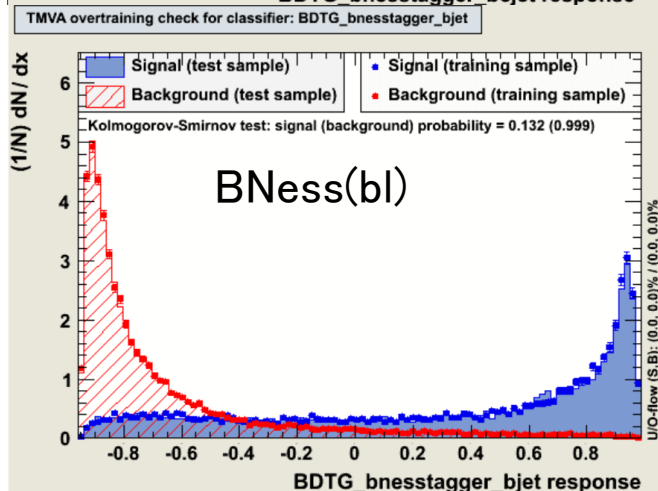
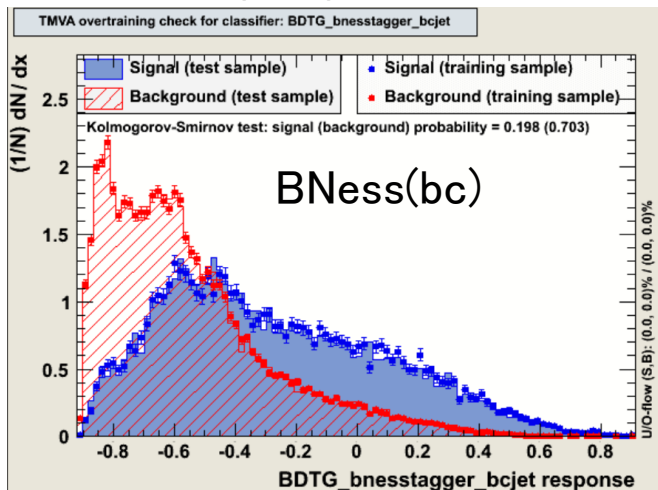
- Tracks will belong to the vertices when $w_{nk} > 0.5$ (k-th track belongs to n-th vertex)
 - Try to fit more tracks than nominal algorithm

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

- 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

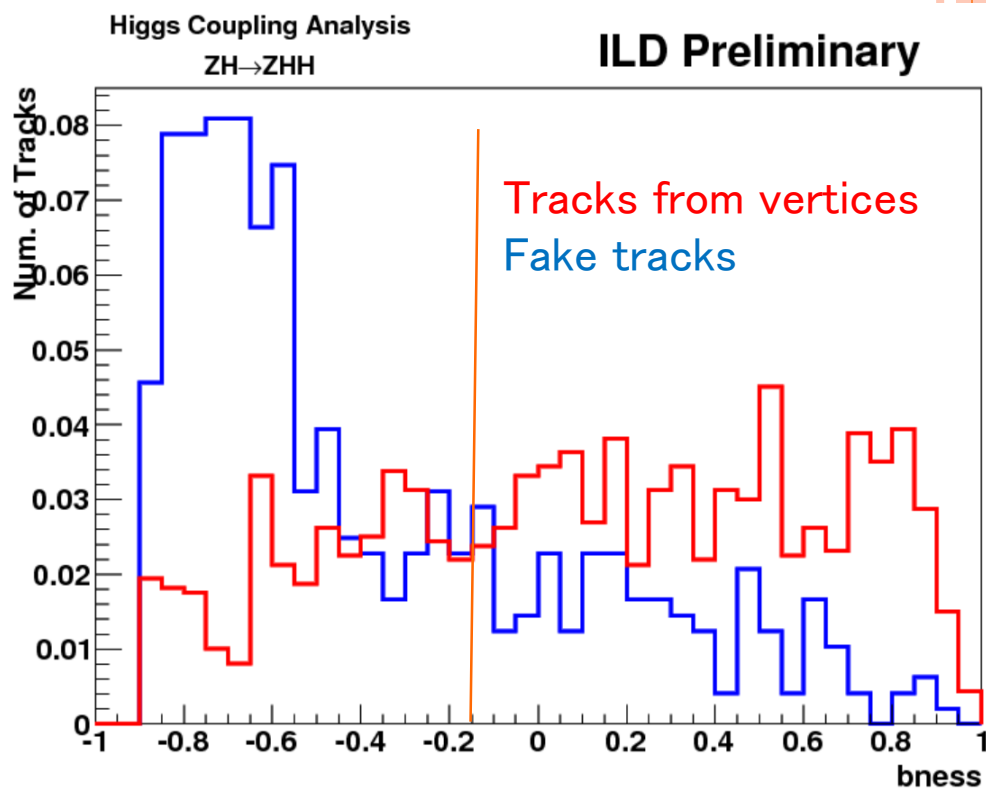
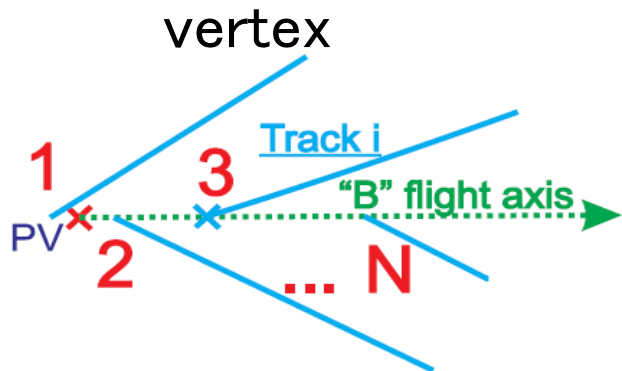


BNESS TAGGER FOR FAKE TRACK REJECTION

- 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

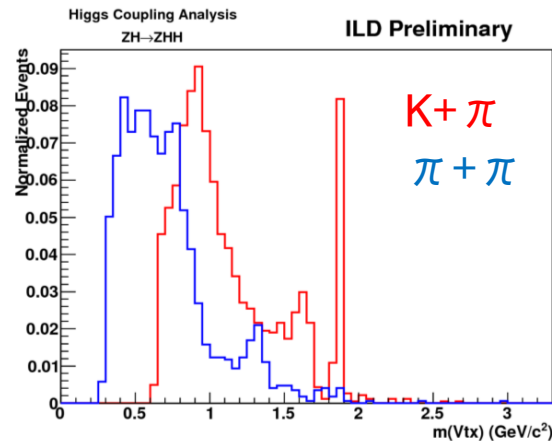
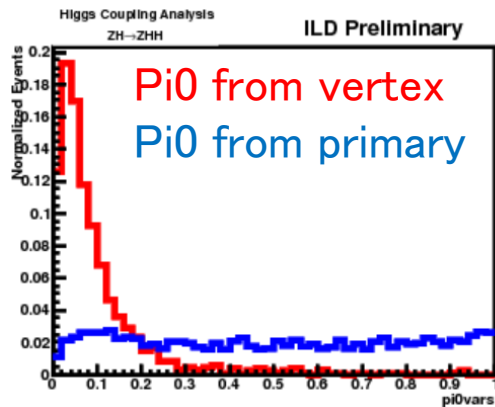
→ some bias for D meson tracks?

Example: looking for single track

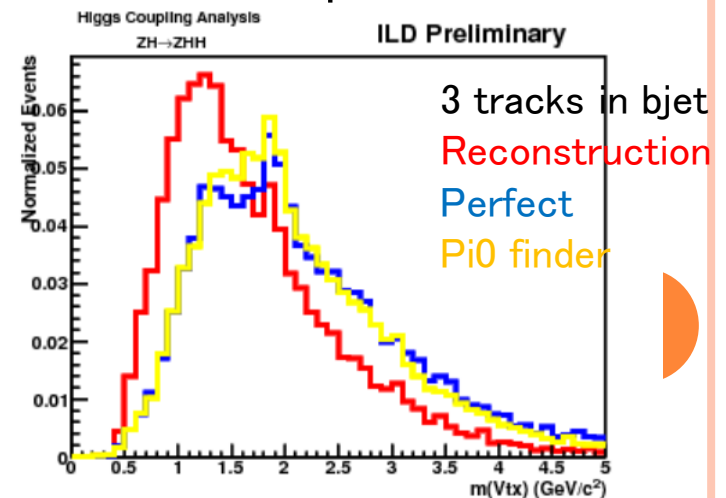


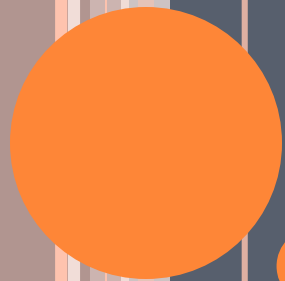
VERTEX MASS RECOVERY

- 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





SOME PERFORMANCE RESULTS

8

VERTEX FINDING OF BJETS IN SIMPLE SITUATION

- Common parameters are set at same values for comparison
- Same event sample (nnH sample@500GeV) 99432 events
 - $H \rightarrow bb$: 55474 events
- 2 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	28215	24049	35616	87880
AVF&BNess	36841	17755	36588	91184

- This case, Total: $\sim 4\%$ increased
 - Jets with 2vtx are drastically increased ($\sim 26\%$): move from 1+1
 - Jets with 1vtx slightly increased ($\sim 3\%$)
- Fake rate per vtx: in this case, a little worse (need opt.)
 - Due to single trk vertices \rightarrow under investigation

method	bjet with 2vtx	bjet with 1+1vtx	bjet with 1vtx
Nominal Algorithm	0.009 ± 0.0004	0.005 ± 0.0005	0.011 ± 0.005
AVF&BNess	0.011 ± 0.0004	0.008 ± 0.0005	0.011 ± 0.005

VERTEX FINDING OF C JETS

- Common parameters are set at same values for comparison
- Same event sample (nnH sample@500GeV) 99432 events
 - $H \rightarrow cc$: 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
 - \rightarrow need additional selection for singletrk? (e.g.) vertex mass?
- Fake 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

VERTEX FINDING OF BJETS IN SEVERE SITUATION

- 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

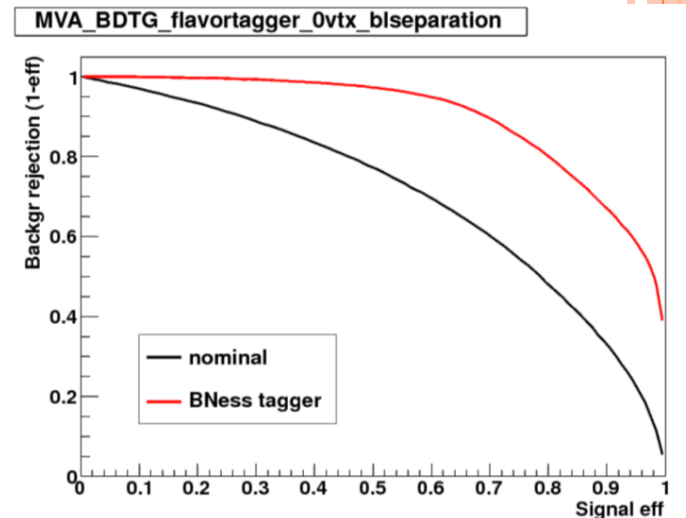
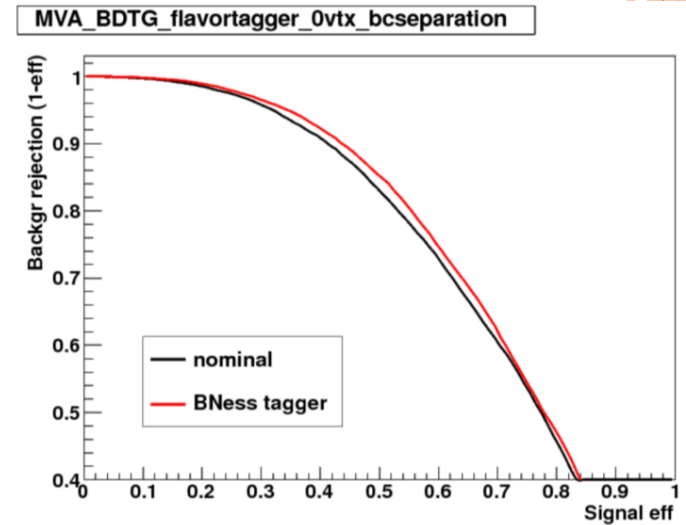
method	bjet with 2vtx	bjet with 1+1vtx	bjet with 1vtx	total
Nominal Algorithm	10577	9159	12804	32504
AVF&BNess	13461	6502	14256	34219

- Jets with vtx: **~5%** increased
 - Jets with 2vtx: **~24%** increased →move from 1+1
 - Jets with 1vtx: **~11%** increased
- Fake rate per vtx:
 - 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

RESULTS OF BNESS TAGGER ON FLAVOR TAGGING

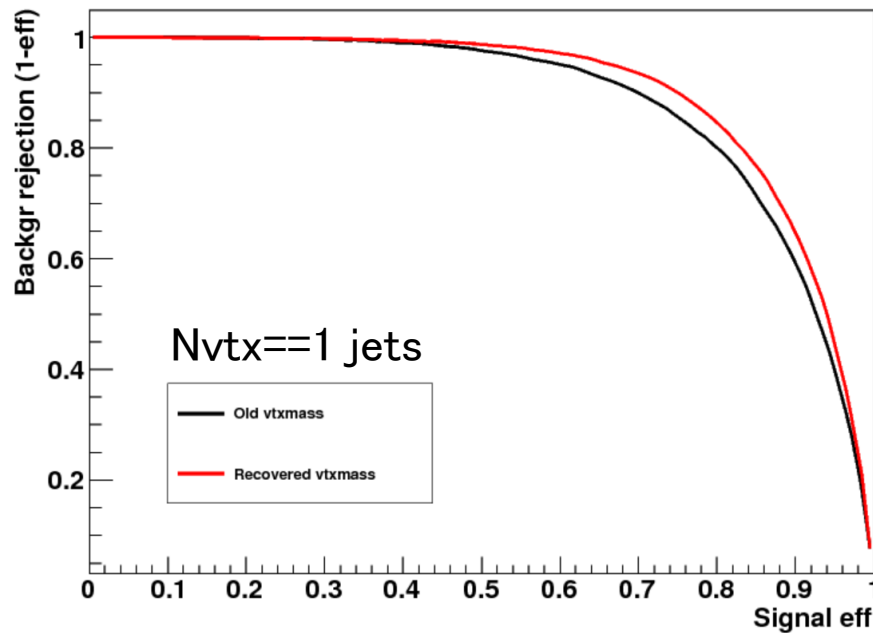
- Construct a “toy” flavor tagger
 - Convert nominal input variables to BNeStagger variables
 - Compare with ROC curve
- For bc separation, some improvement can be obtained
- For bl separation, becomes too good? under investigation
- b-l separation will be very good!
- Need optimization
- Especially,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



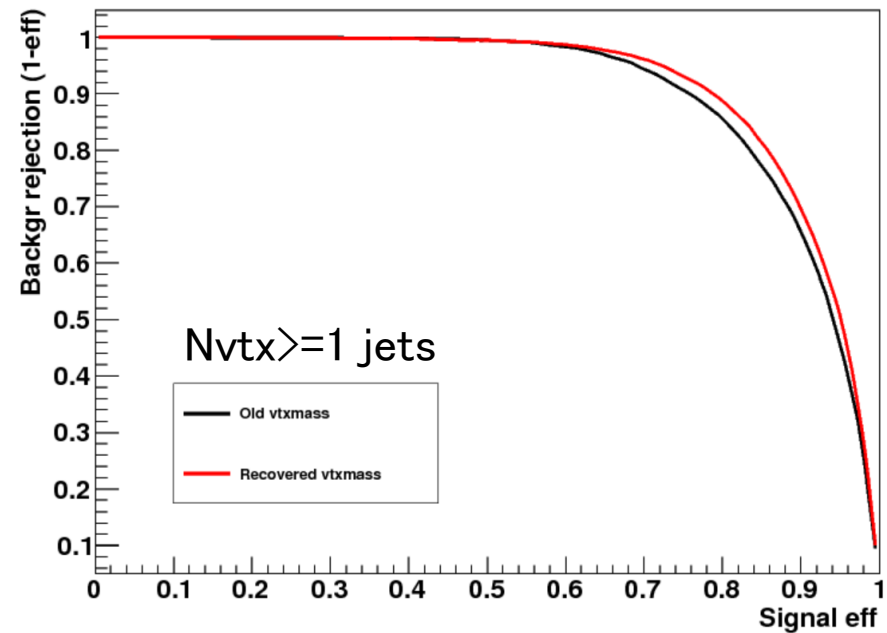
VERTEX MASS RECOVERY EFFECT ON FLAVOR TAGGING

- 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

MVA_BDTG_flavortagger_bcseparation



MVA_BDTG_flavortagger_bcseparation



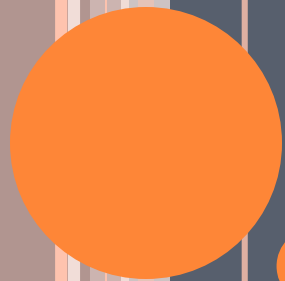
- Vertex is created using nominal algorithm
 - need to check AVF case

SOFTWARE RELATED ISSUE

- As far as I think, we can construct same structure as present LCFIPlus modules.
 - Because LCFIPlus has flexible structure for extension
 - We can provide same output of flavor tagging using updated method
- It is necessary to make flag or switch to choose specific method, for example, nominal algorithm or AVF or etc.
 - It takes a little time to construct such structure, but it can be
 - So, Including AVF and BNess tagger will be possible
 - I don't know about vertex mass recovery – need time for public use
- I don't think whole updated LCFIPlus can be constructed within a month because final flavor tagging is not yet studied
 - It takes some time...
 - In this time, update will not directly lead to flavor tagging improvement
 - But, very near future, it can be!

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
- Vertex mass recovery is reasonable
 - Will provide better flavor tagger
 - Of course, many checks are necessary
- 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
- So far, AVF will provide **4-7%** improvement of vertex finding in bjets or c jets
 - 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
- Finally, incorporate all the ideas and check the final flavor tagging effs.in LCFIPlus!
 - Not within a month, but very near future!



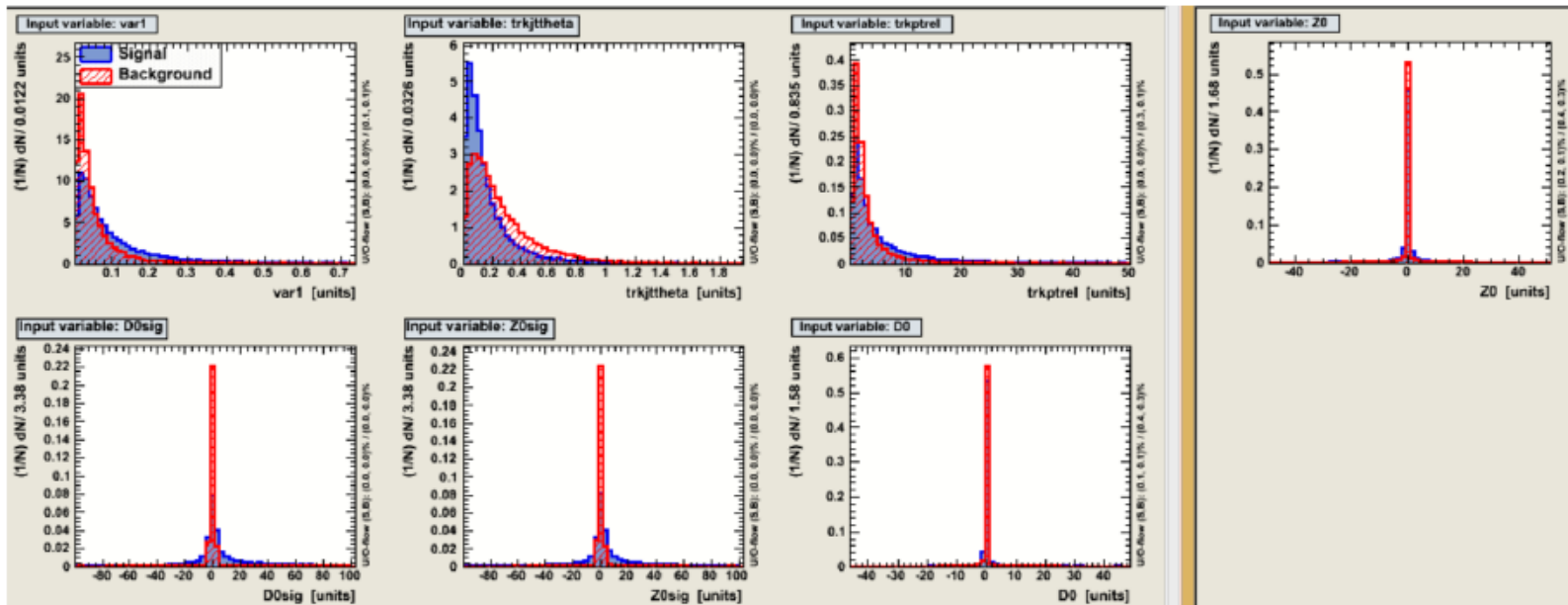
BACK UPS



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



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

