LCFIPlus status report
(Flavour tagging and Jet clustering)

Masakazu Kurata (KEK)
Tomohiko Tanabe (The University of Tokyo)
Taikan Suehara (Kyushu University)
Jan Strube (PNNL)
Ryo Yonamine (Tohoku University)

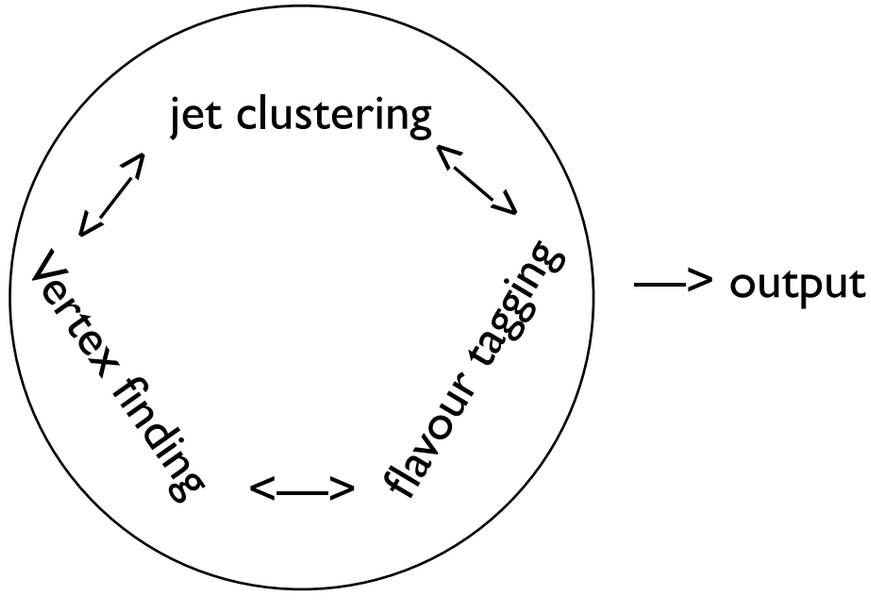
SW/Ana pre-meeting @ KEK, 19 Feb. 2018

LCFIPlus

- **A framework for jet flavour identification.**
  - originated from LCFIVertex (e.g. arXiv:0908.3019)
  - does vertex finding, jet clustering, flavour tagging,…
  - each process is implemented as a modular algorithm.
  - gives flexibility to iterate or reverse processes.

w/o LCFIPlus : jet clustering —> vertex finding —> flavour tagging —> output

w/ LCFIPlus :



  - typical flow is not so complicated;
    "vertex finding —> jet clustering —> vertex refining —>flavour tagging"
    to use vertex information for jet clustering.

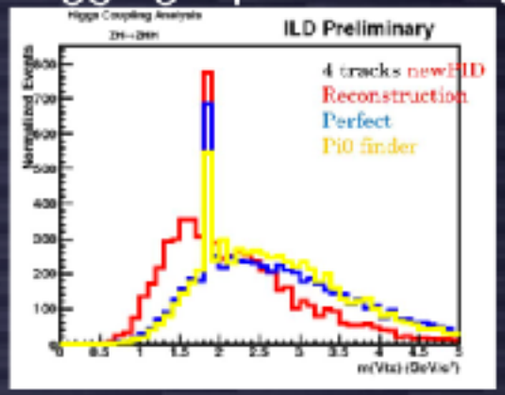
Concerns

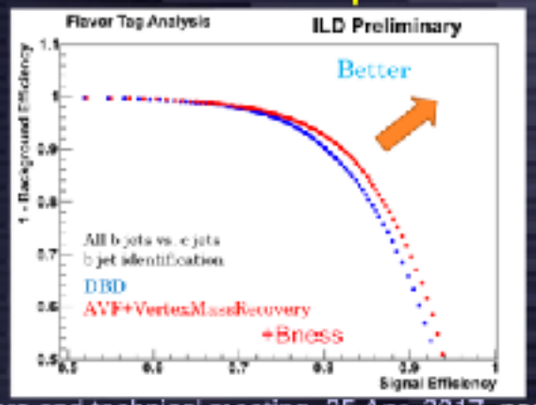
- **Crash problem related to ROOT6**
  - already fixed, but to be checked also with the latest release.
- **Vertexing performance**
  - This must be checked before coming mass-production.
- **Flavour tagging performance**
  - Taikan and Masakazu have reported that new features (e.g. adaptive vertex figging, vertex mass recovery, etc) improve the performance using previous ilcsoft versions. This should also be confirmed with the latest release.
- **Other idea/todo/hope (picked up from Taikan's slide)**
  - improve PID, π0 reconstruction
  - b/c separation (efficient finder = worse b/c separtion), need two std conf.
  - gamma gamma hadron background
  - tau finder
  - Jet substructure
  - vertex quality
  - systematic errors in flavour tagging
  - improve track selection

Developments since DBD improves the performance

Vertex Mass Recovery

- Correct π^0 pairs to add to vertices for mass calc.
  - π^0 selection by new PID (with dE/dx)
  - Recovery process in individual algorithm (VertexMassRecovery, after VertexRefiner)
- Simple replacement of vertex mass in flavor tagging input variable gives better b/c sep.

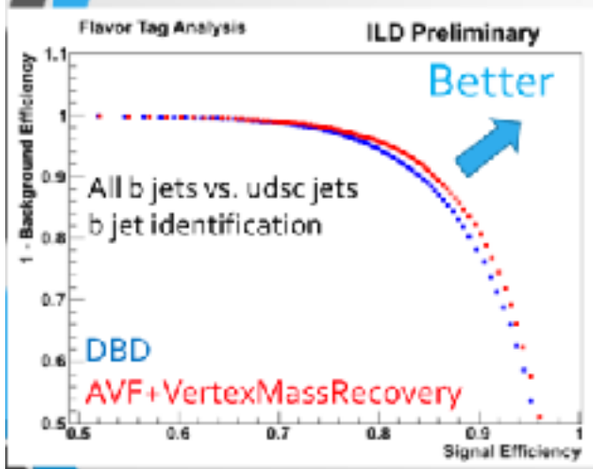


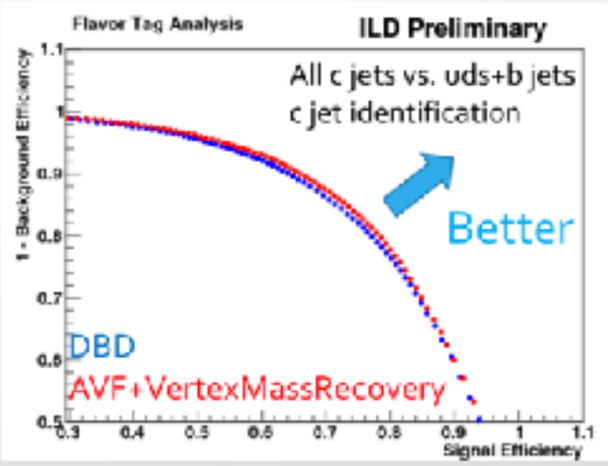


ILD software and technical meeting, Apr. 2017

Impact on Flavor Tagging Efficiency

- 6f samples coming from ZZZ events@500GeV
- Compare with ROC curve





LCWS@strasbourg, Oct. 2017

with v01-19-04

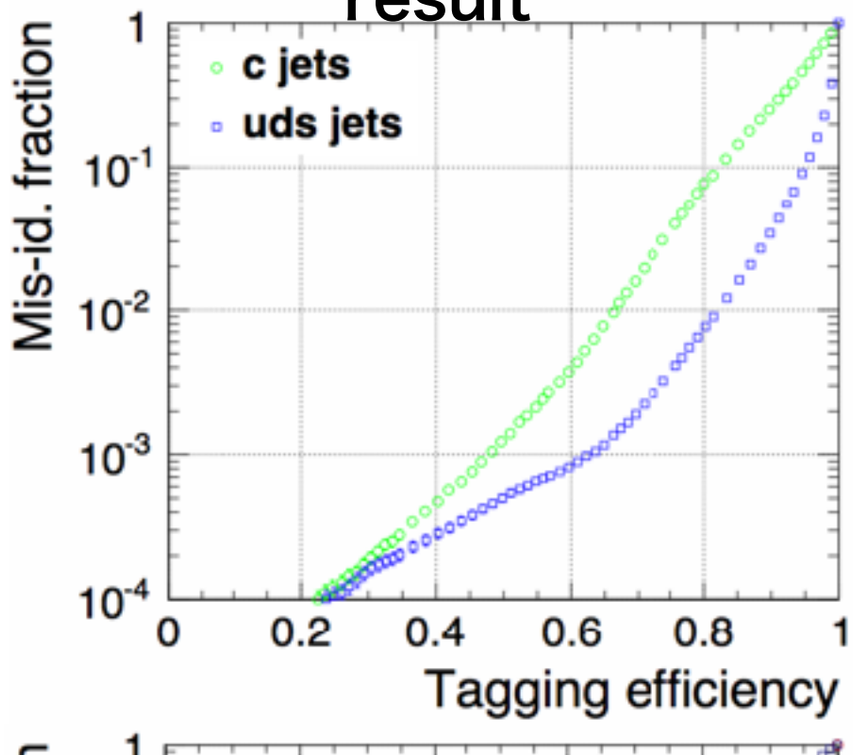
No validation plots with latest release yet

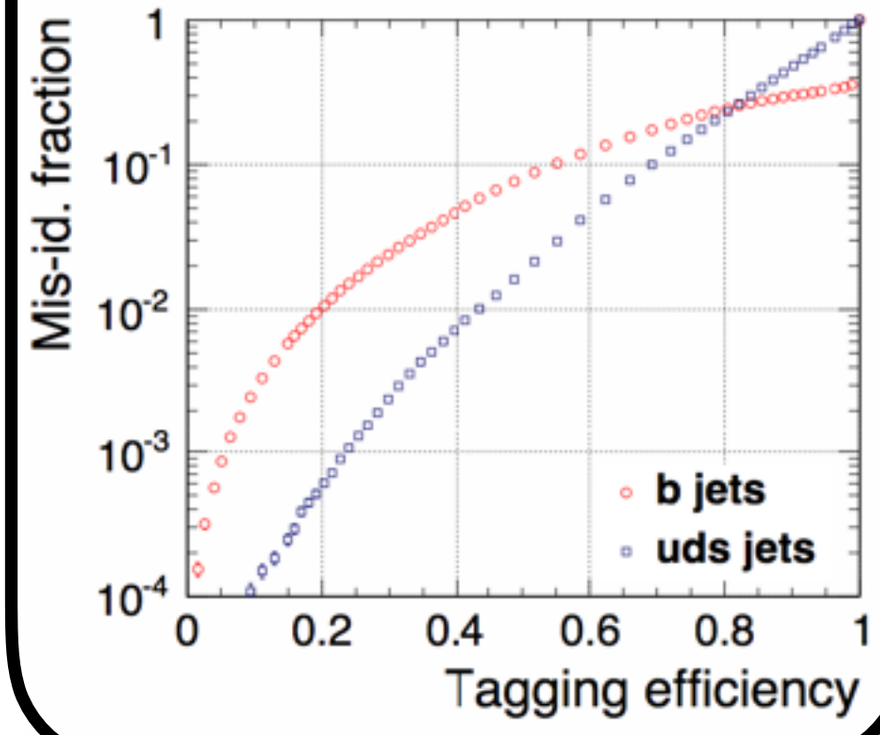
- **Problem is just I have not yet found the correct configuration/setup.**
  - I'm on the way to reproduce DBD plots. (I tried with v01-19-05, and the result is not exactly same. I will rerun with v01-16 just in case.)
  - When I compared results with my setup (same as above), flavour tagging performance looks worse than that of DBD studies.
  - Ichinoseki meeting is good chance to work together with Taikan and Masakazu.

- **Comparing v01-19-04 and v01-19-05 would help to see what we can expect for the moment.**
  - If there is no difference, we can expect same performance as what Taikan and Masakazu showed.
  - Relative comparison with these two can be easily done in a common condition.
  - Please do not worry, I will validate the "absolute" performance once I find the configuration/setup of Taikan/Masakazu.

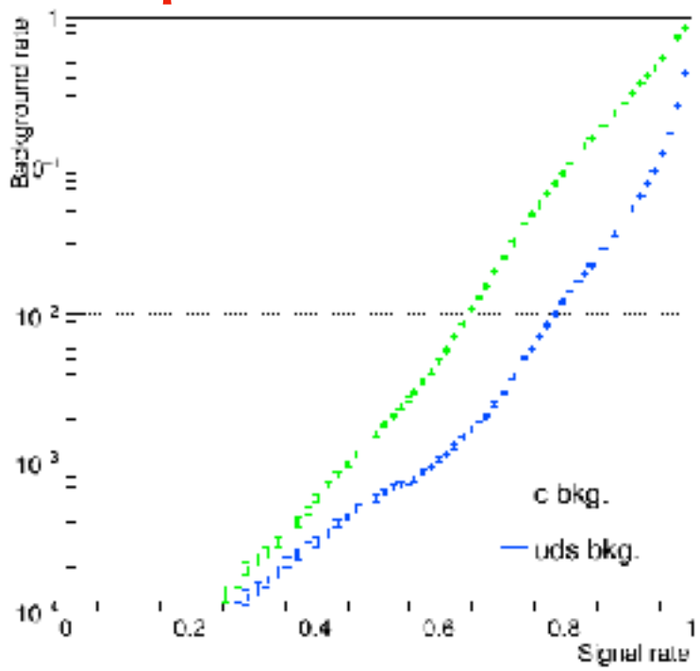
Comparison for DBD sample

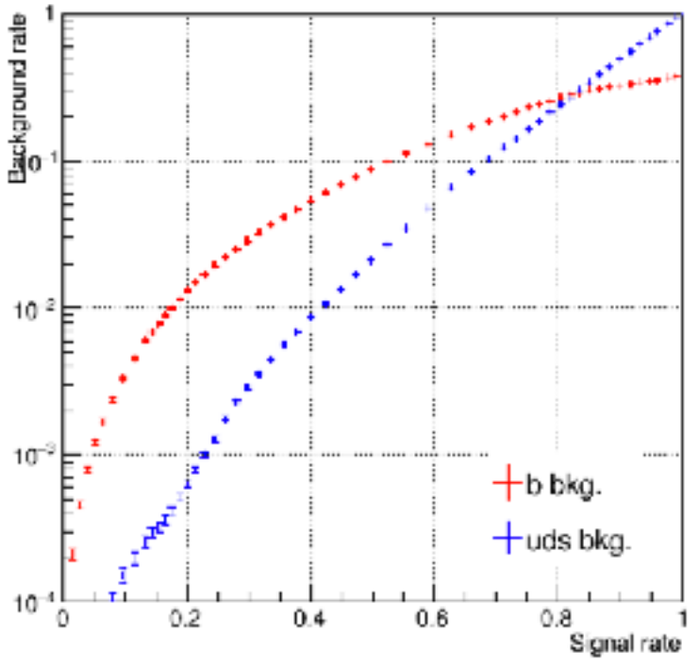
Taikan/Tomohiko's result





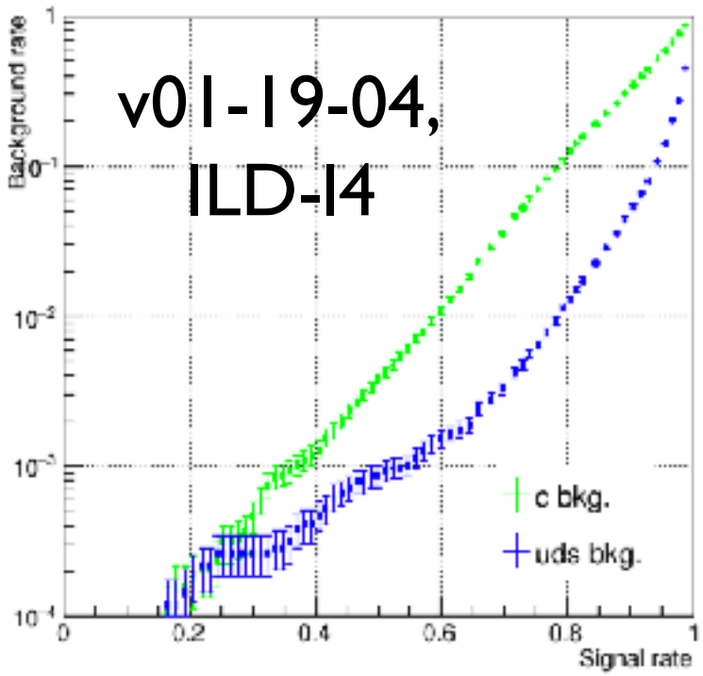
My result for DBD sample with v01-19-05

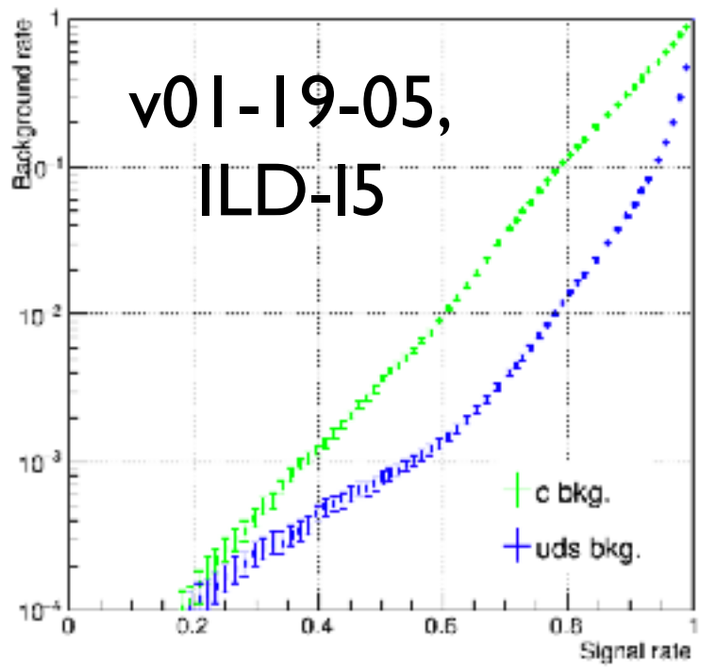




Comparison between v01-19-04 and v01-19-05

Btag performance

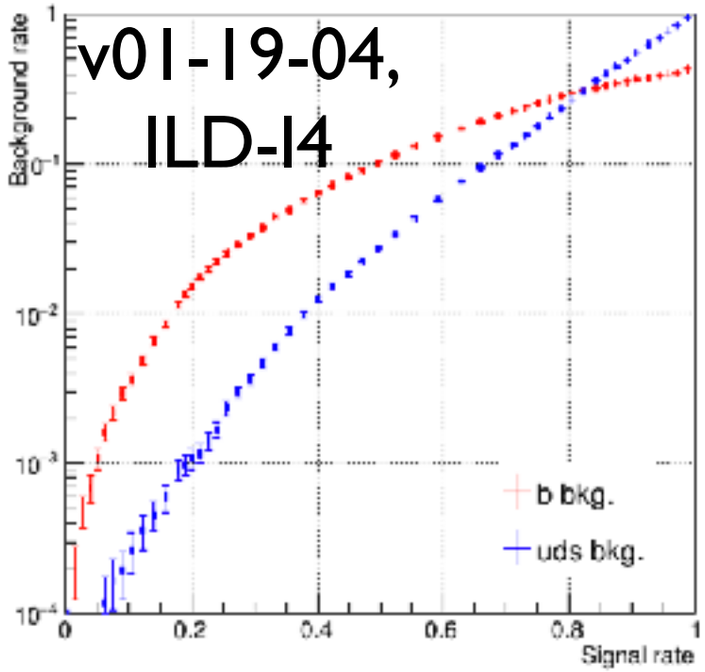


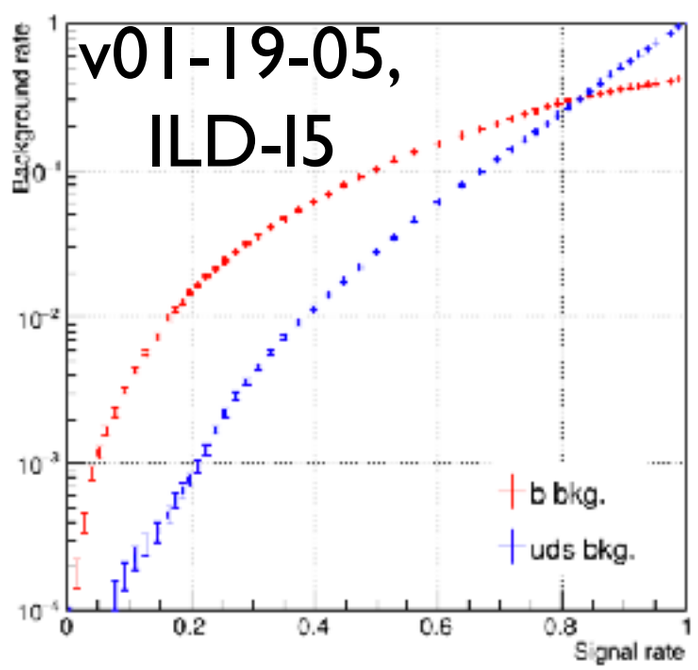


I confirmed no difference in
ILD-s4/-s5 comparison.

Worse than DBD for now.
I expect
my configuration/setup is not proper.

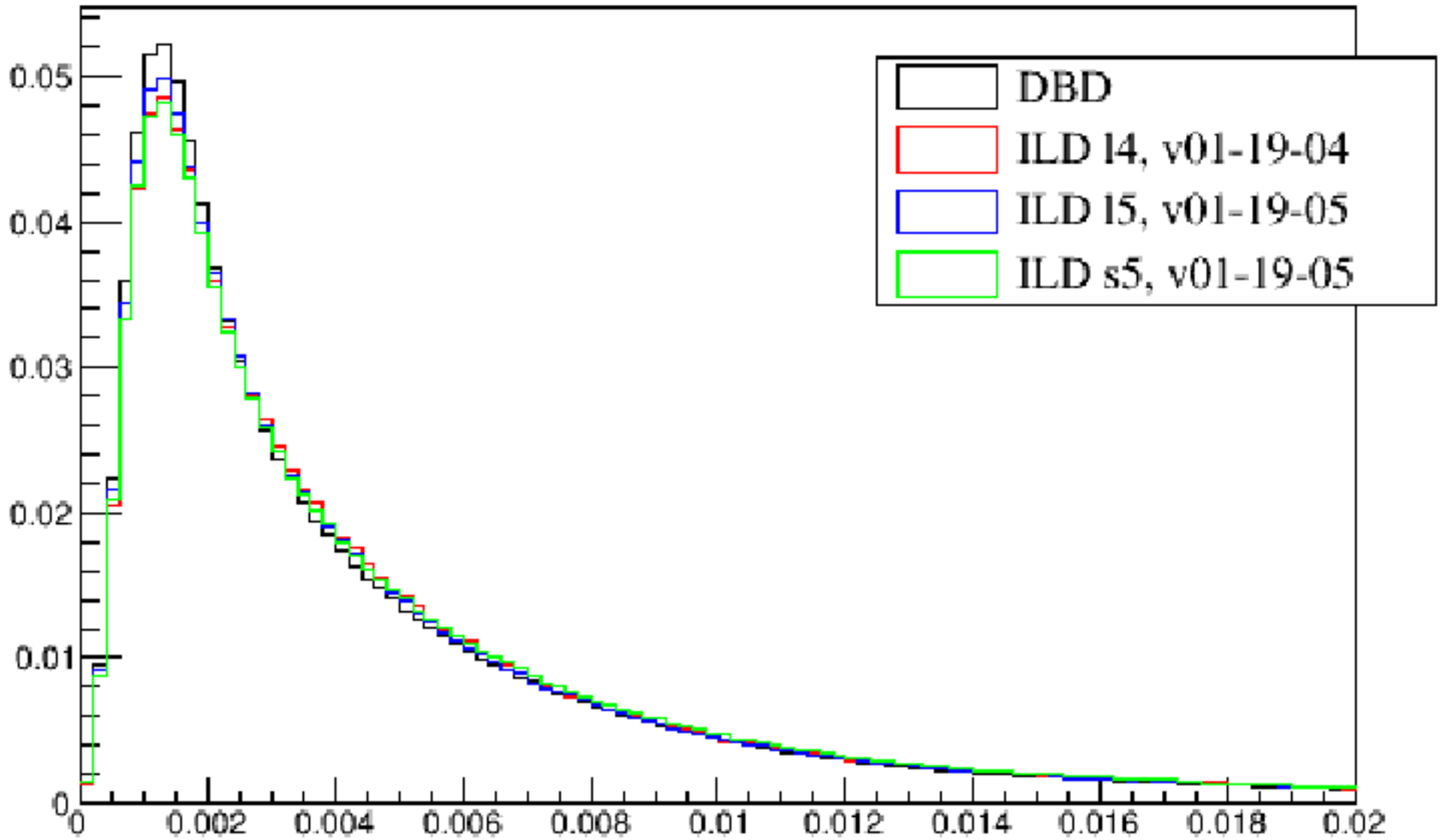
Ctag performance





Residual dist. of track momentum

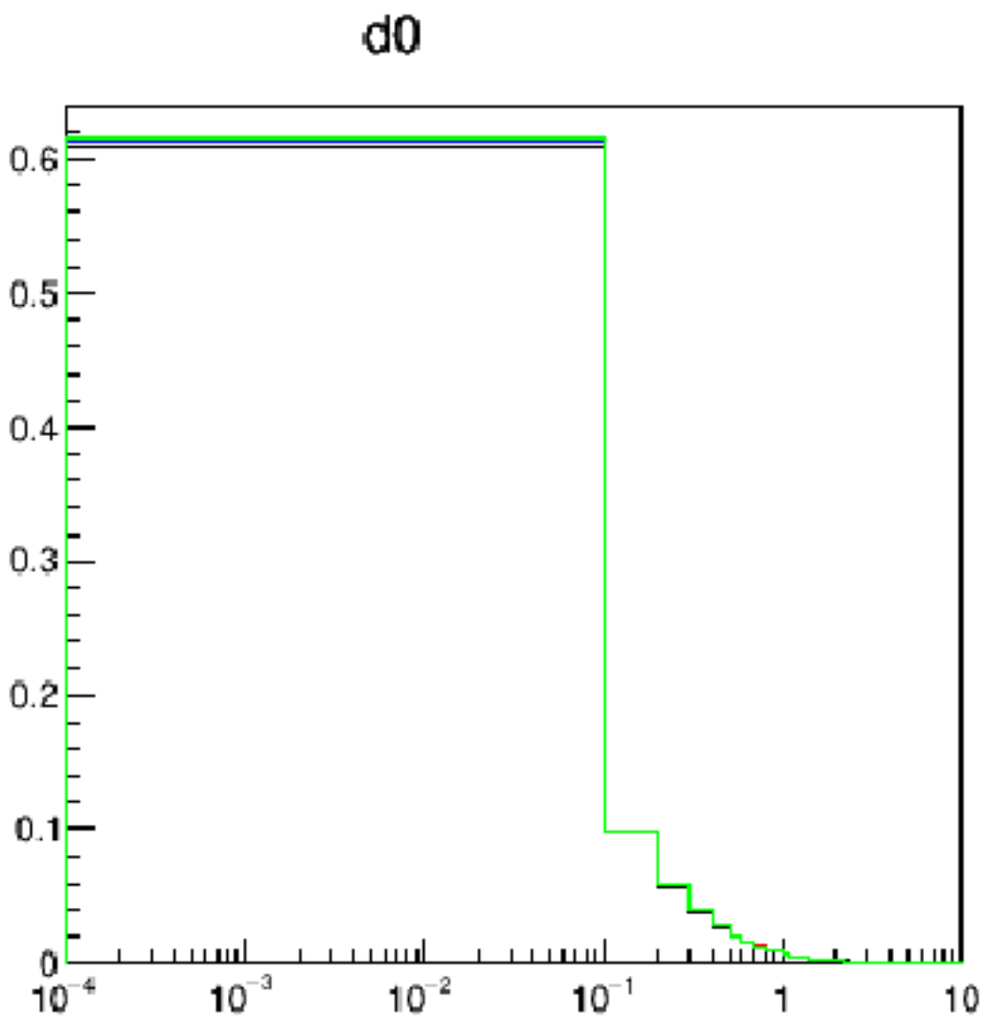
- bbar sample, no ISR, no beam bkg, $\sqrt{s}$=91.2GeV

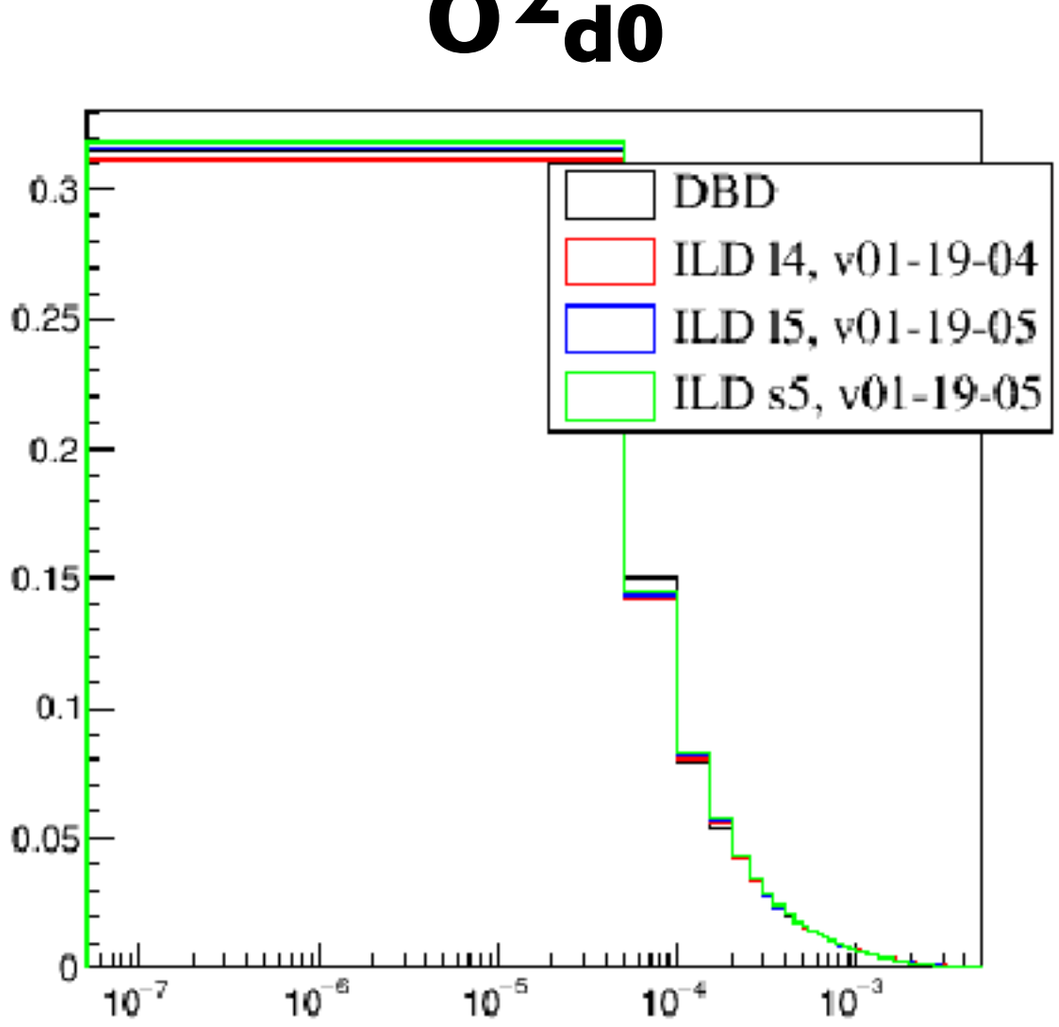


$|P_{rec} - P_{mc}|$

Track (helix) parameter (D0)

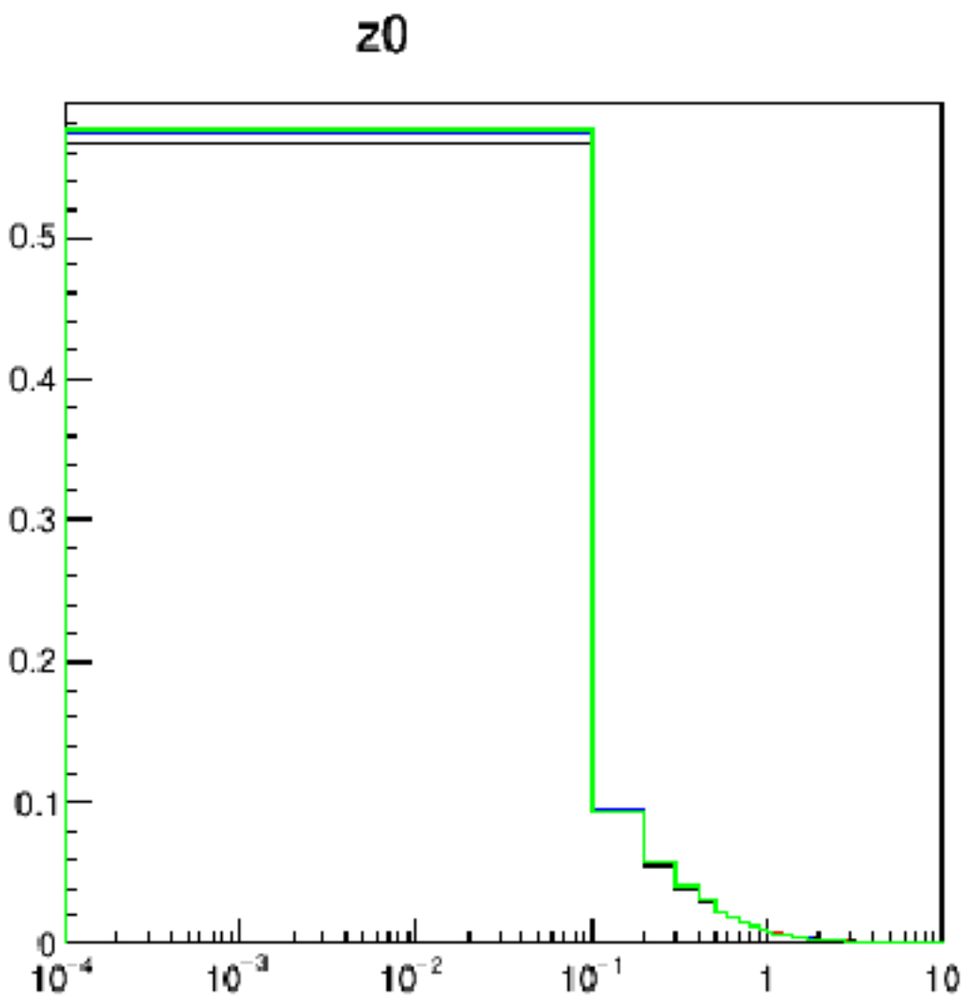
- bbar sample, no ISR, no beam bkg, $\sqrt{s}$=91.2GeV

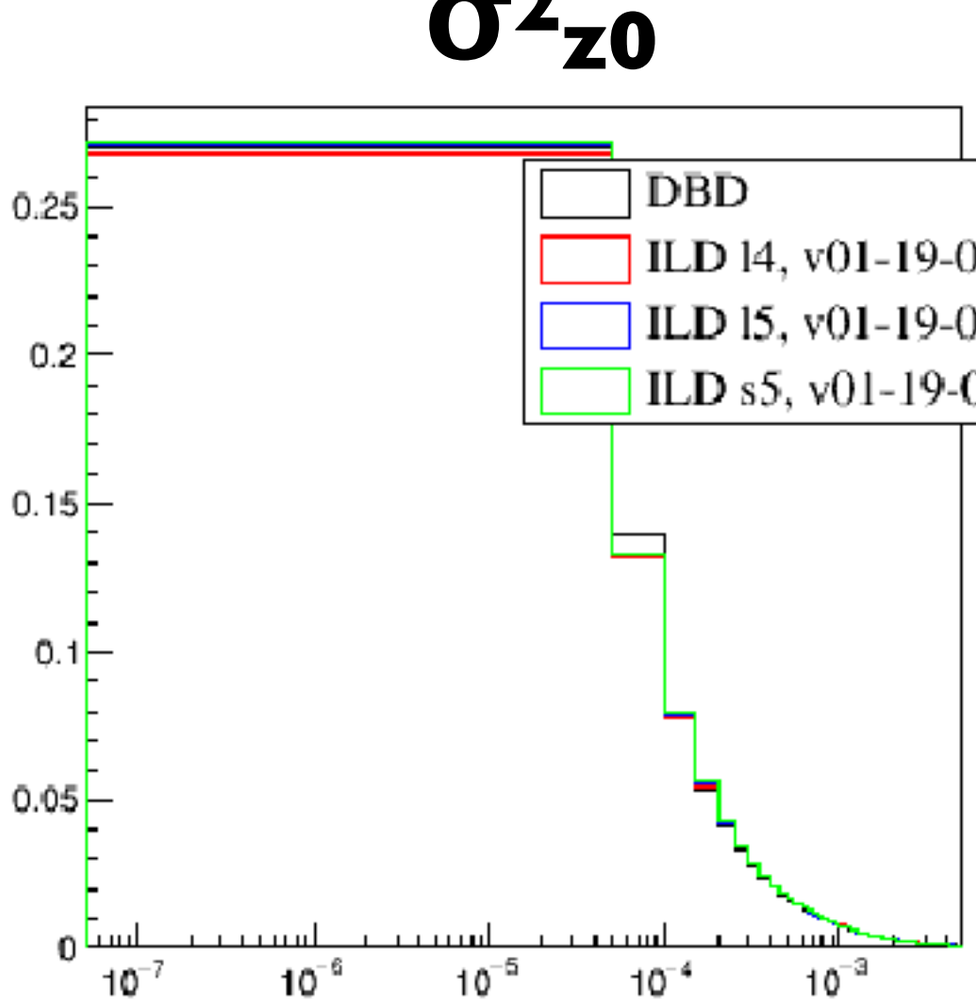




Track (helix) parameter (z0)

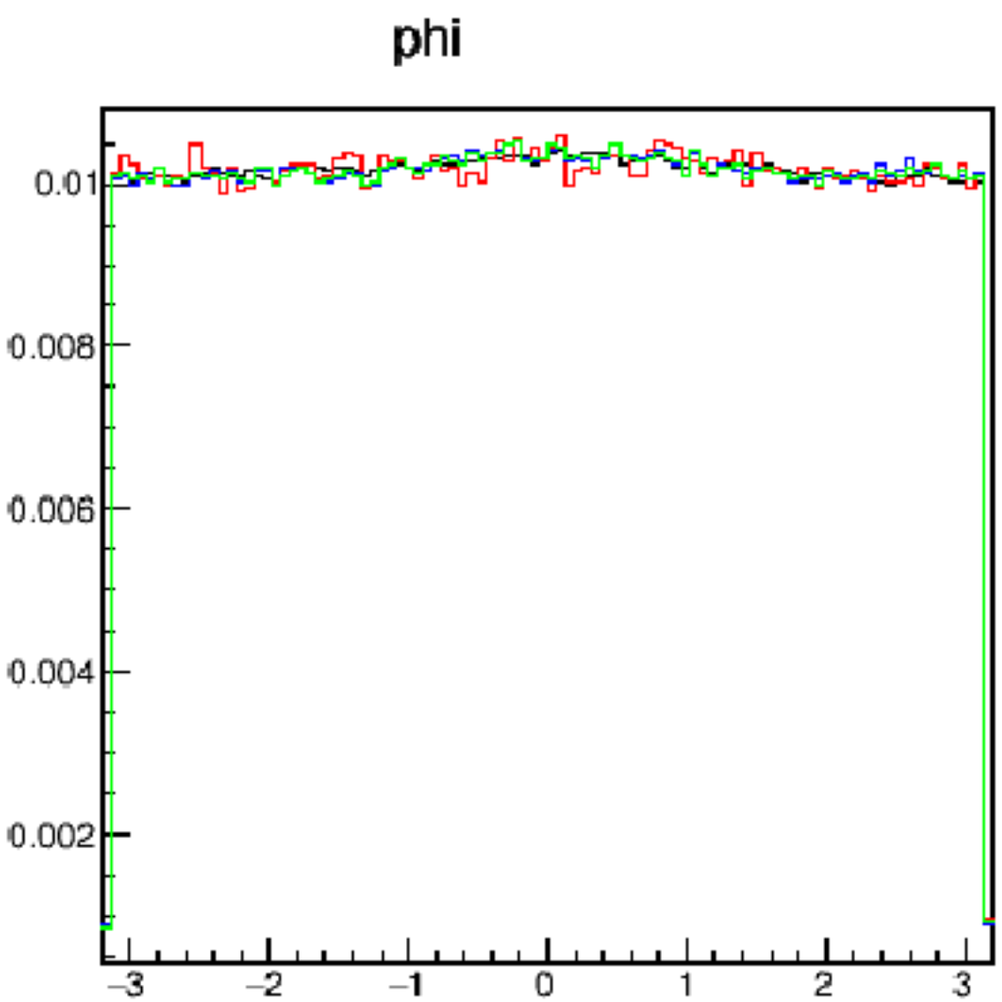
- bbar sample, no ISR, no beam bkg, $\sqrt{s}$=91.2GeV



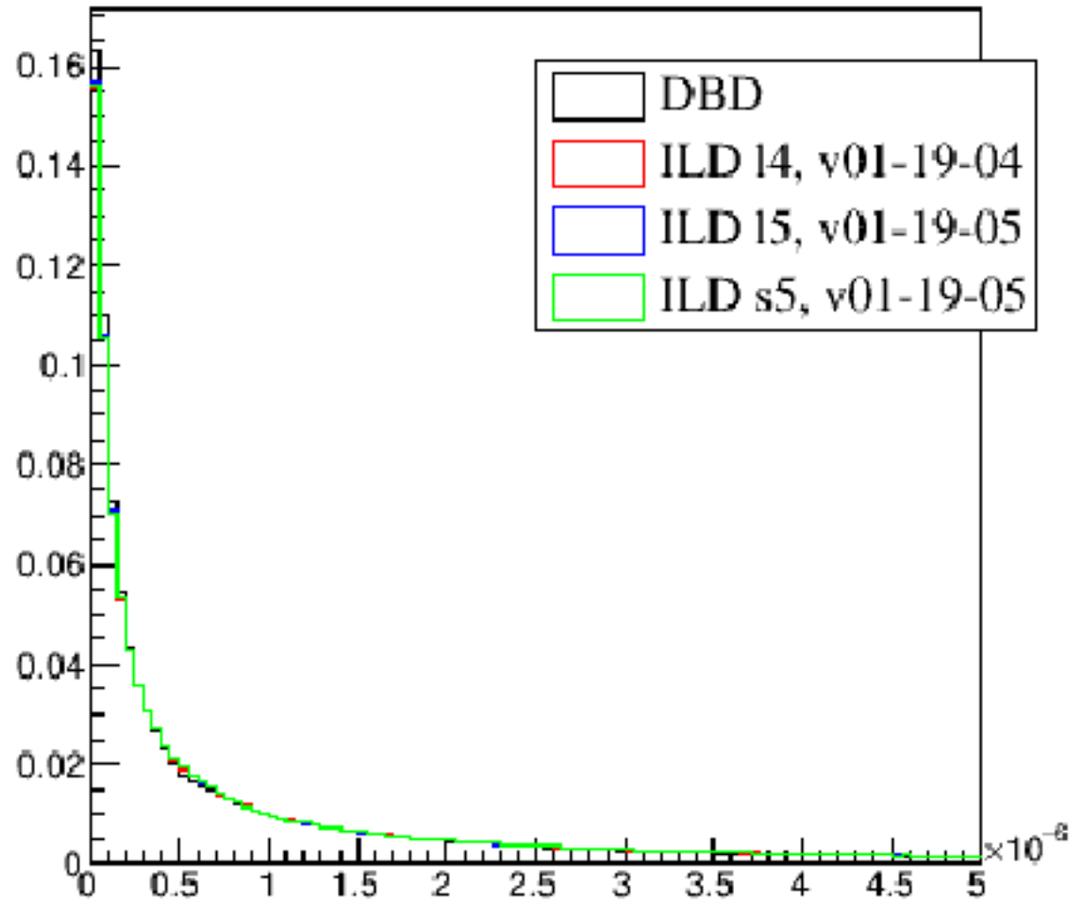


Track (helix) parameter (phi)

- bbar sample, no ISR, no beam bkg, $\sqrt{s}$=91.2GeV



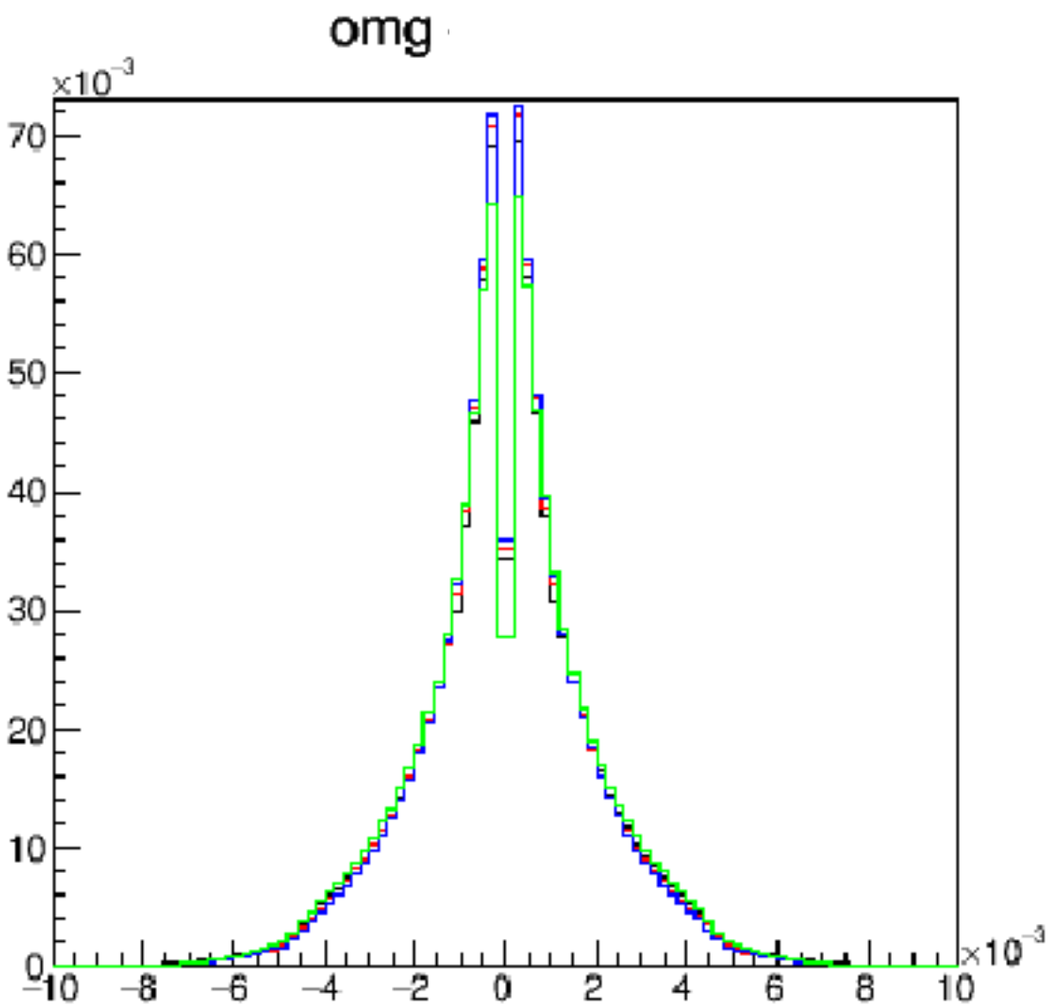
σ^2_{phi}

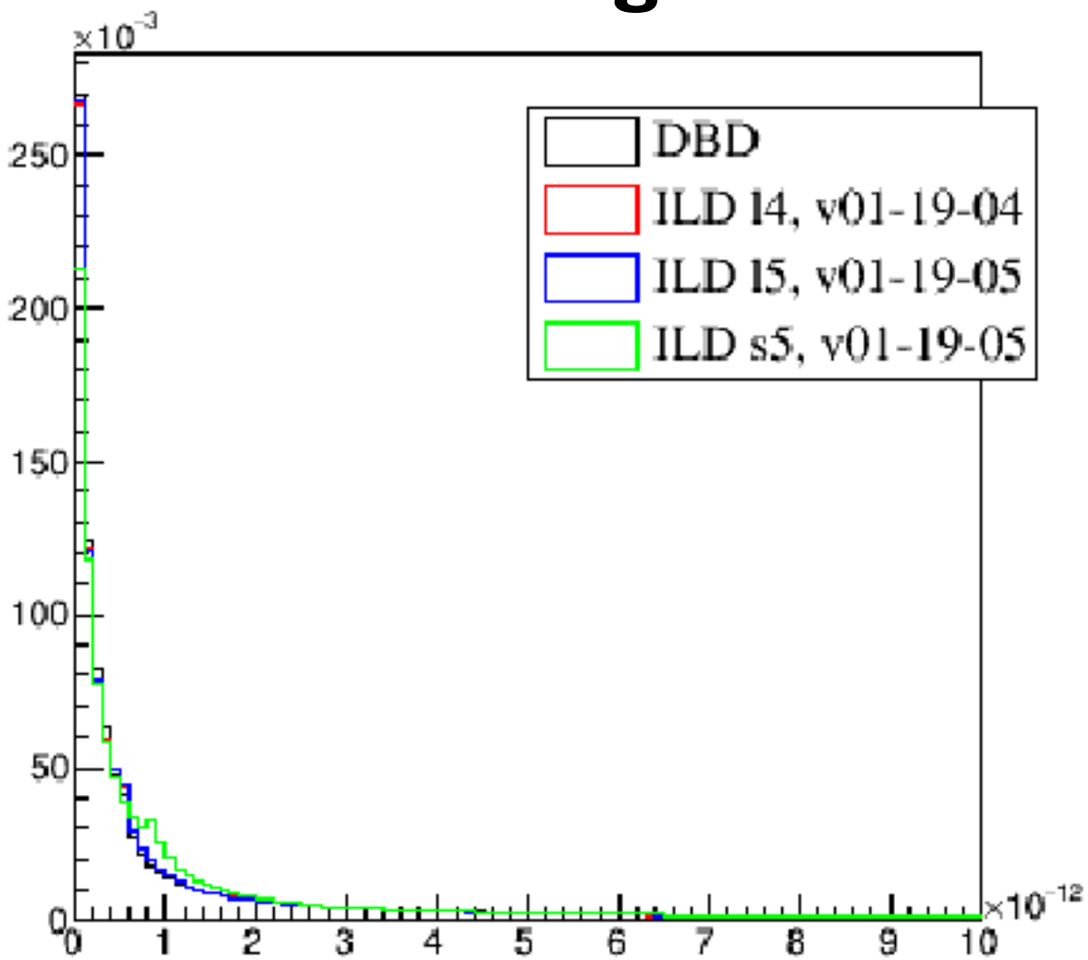


Track (helix) parameter (omega)

- bbar sample, no ISR, no beam bkg, $\sqrt{s}$=91.2GeV

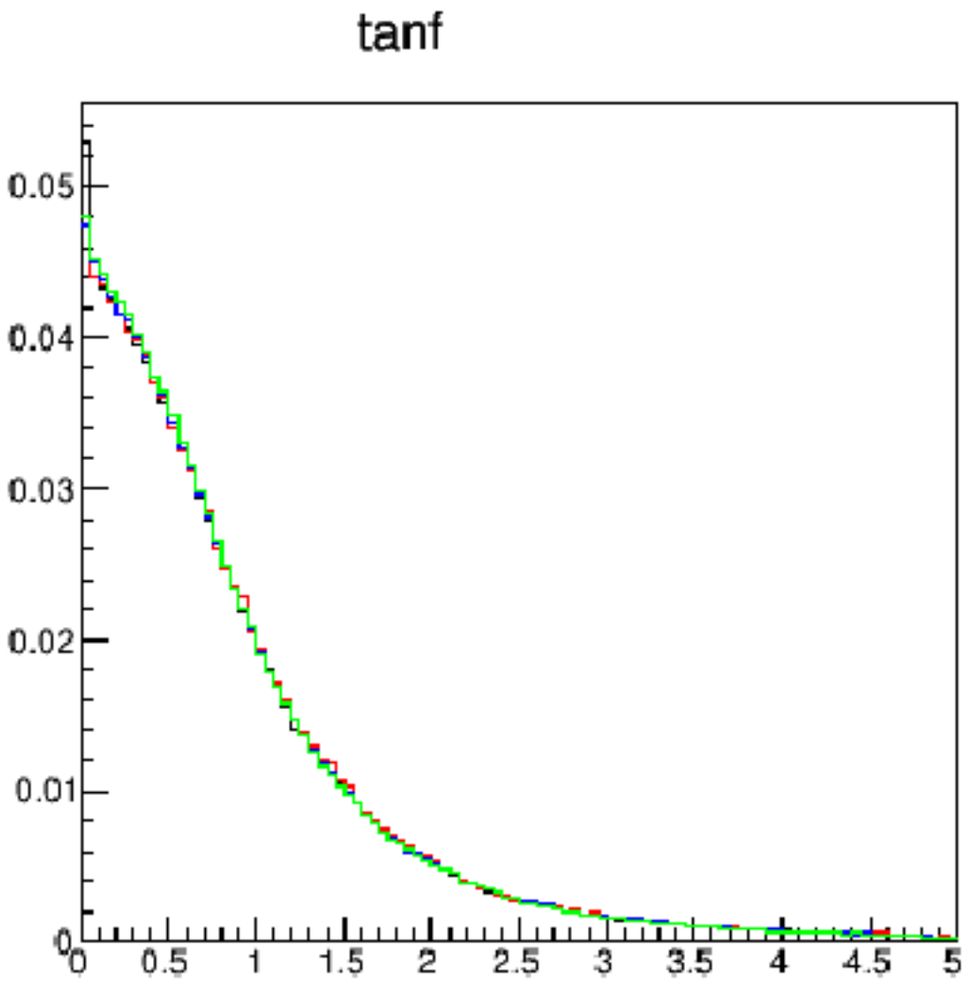
σ^2_{omg}



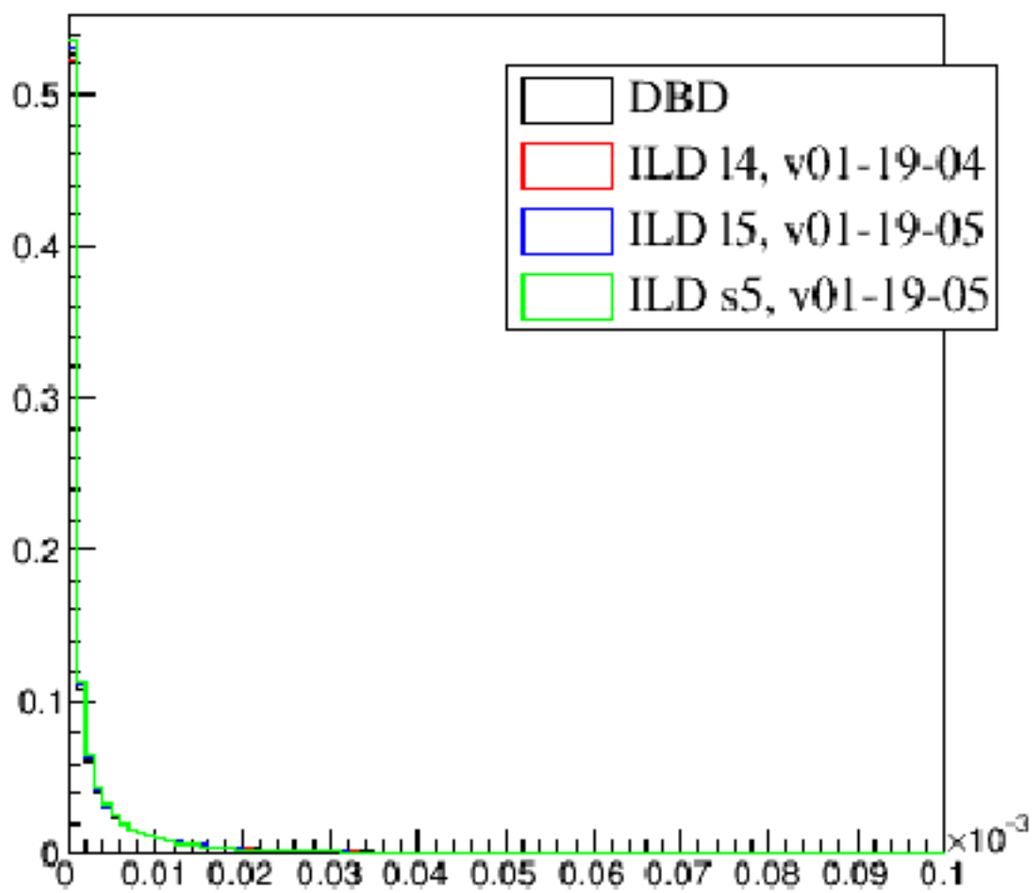


Track (helix) parameter (tanλ)

- bbar sample, no ISR, no beam bkg, $\sqrt{s}$=91.2GeV

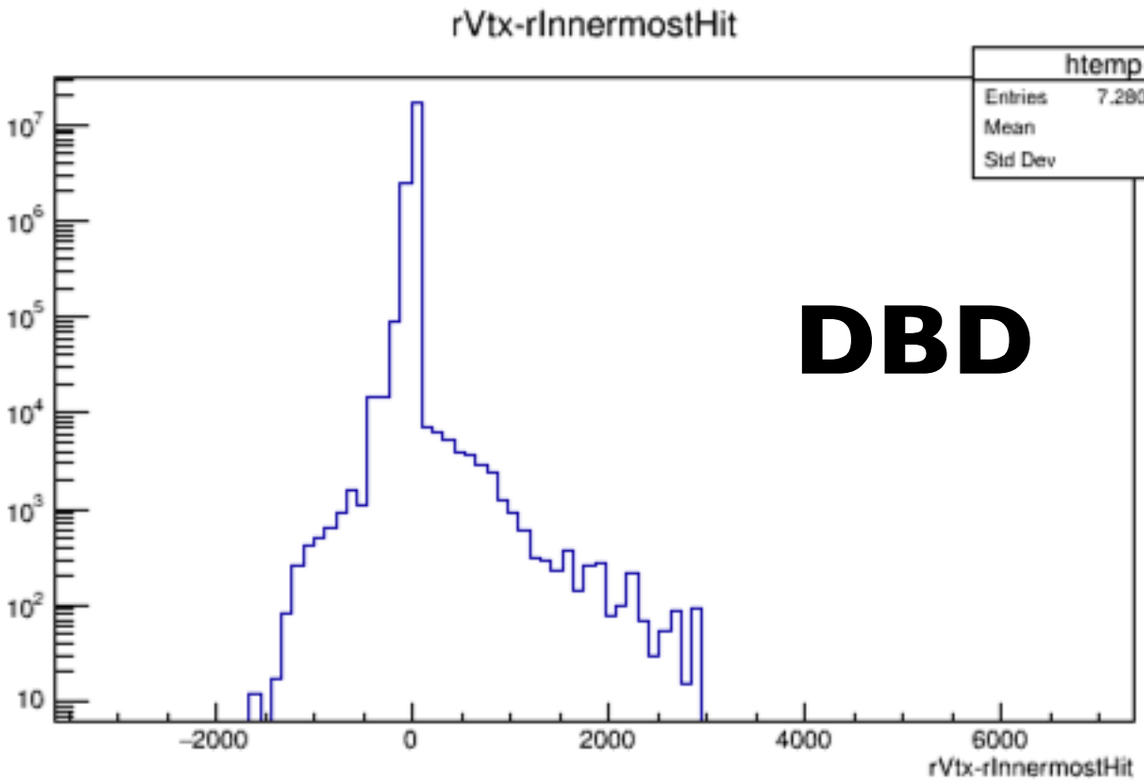


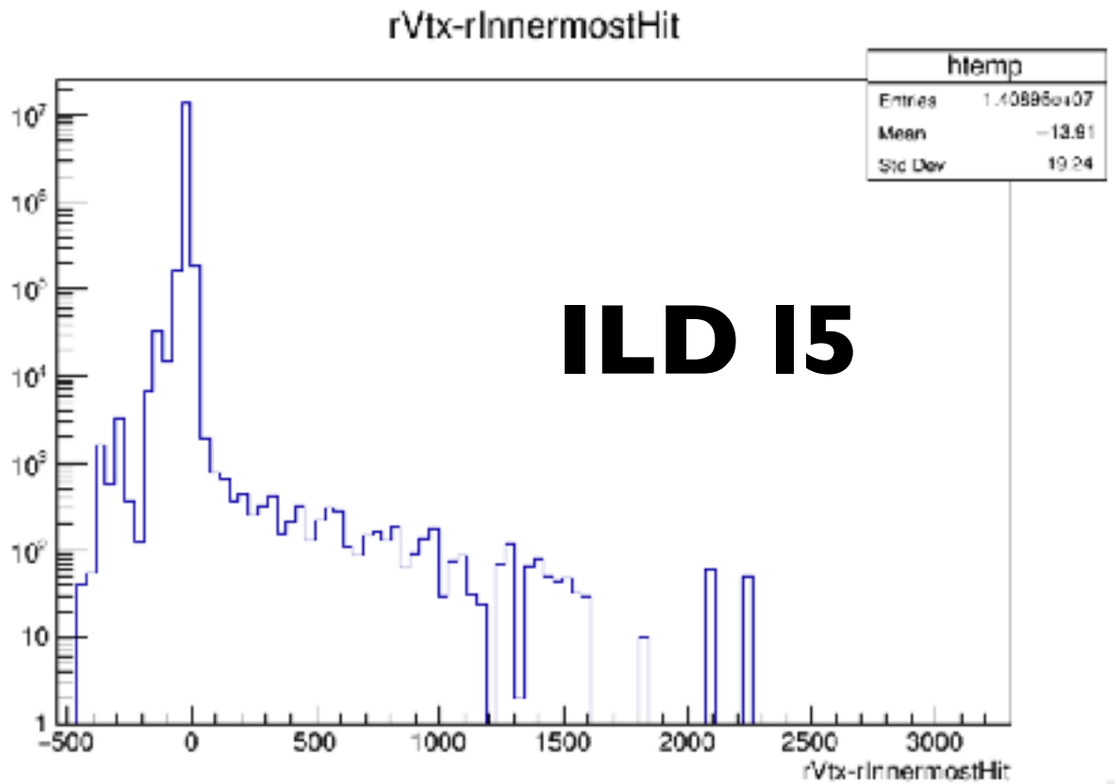
σ^2_{tanl}



Extra: Findings

R_{vertex position} - R_{innermost hit of the tracks associated to the vertex}





What's going on at the tails?

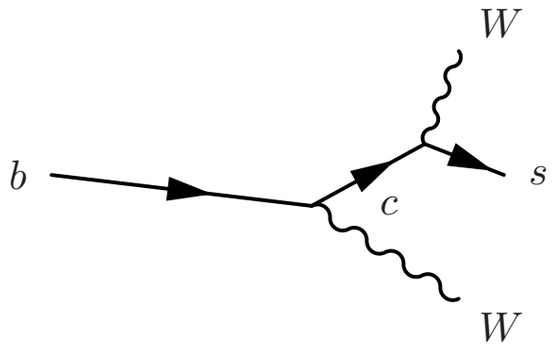
Conclusions

- **No technical problem was found with v01-19-05 (with ROOT6).**
  - works with DD4hep.
- **No significant differences between v01-19-04 and v01-19-05**
  - With previous results, I expect that v01-19-05 can achieve better performance than that of DBD studies.
  - We should confirm it in a direct way soon. Stay tuned.

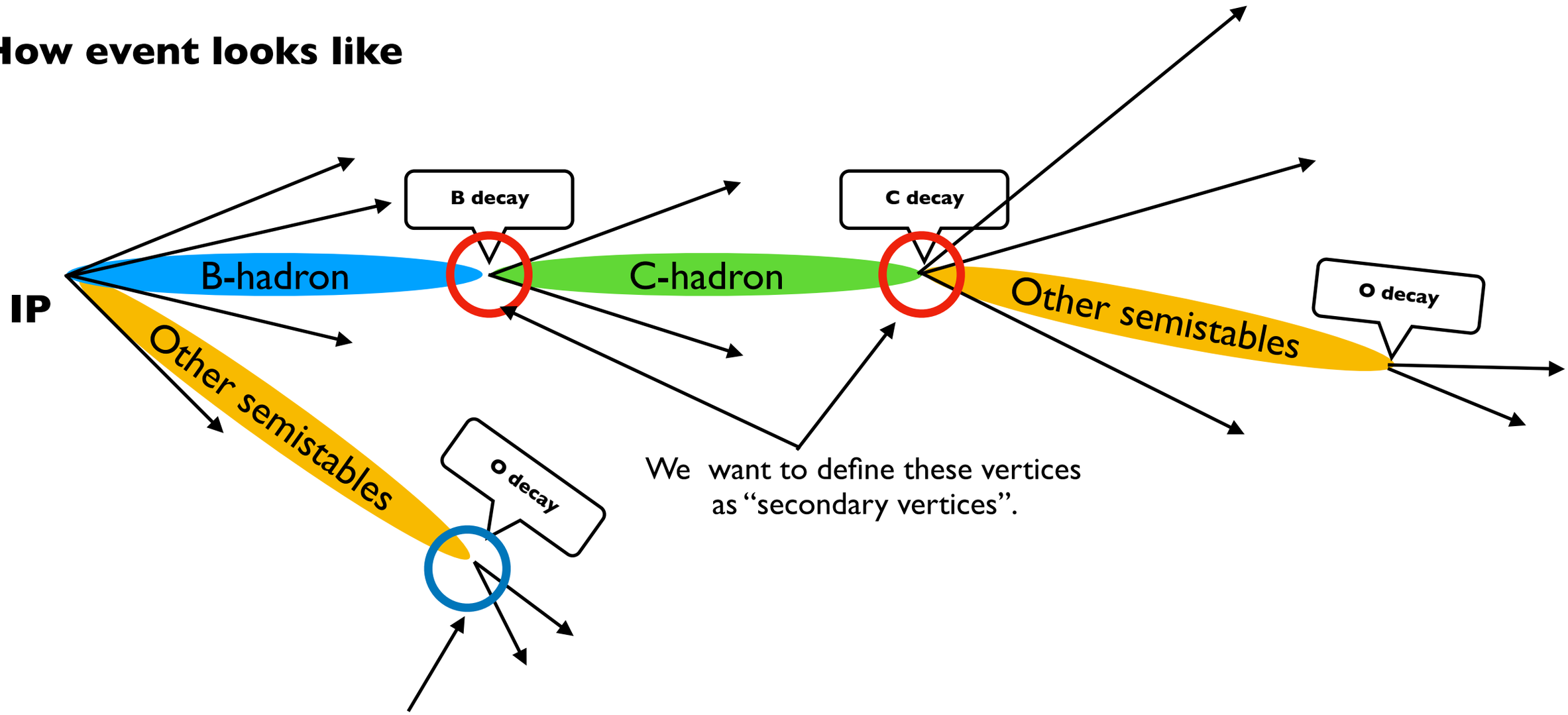
Backup

Event signature and term definition

Feynman diagram



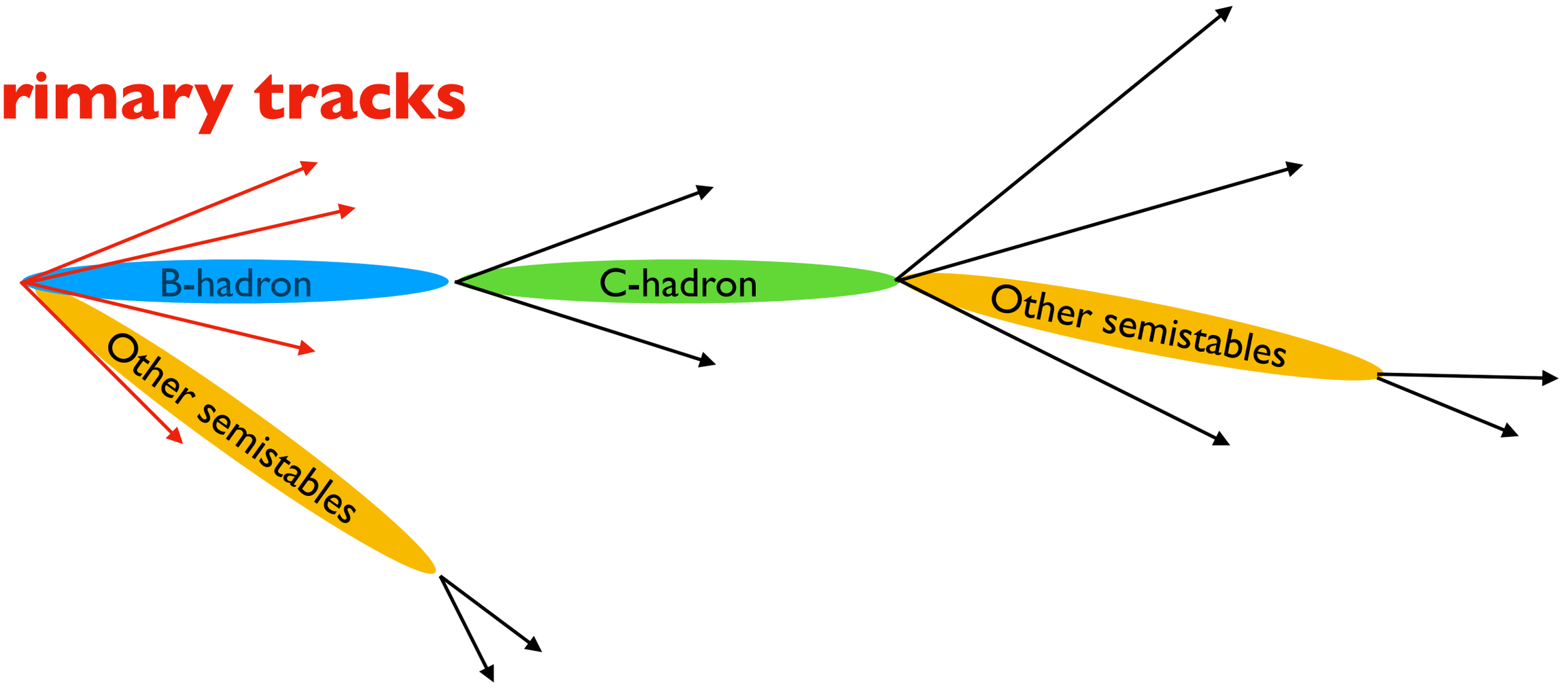
How event looks like



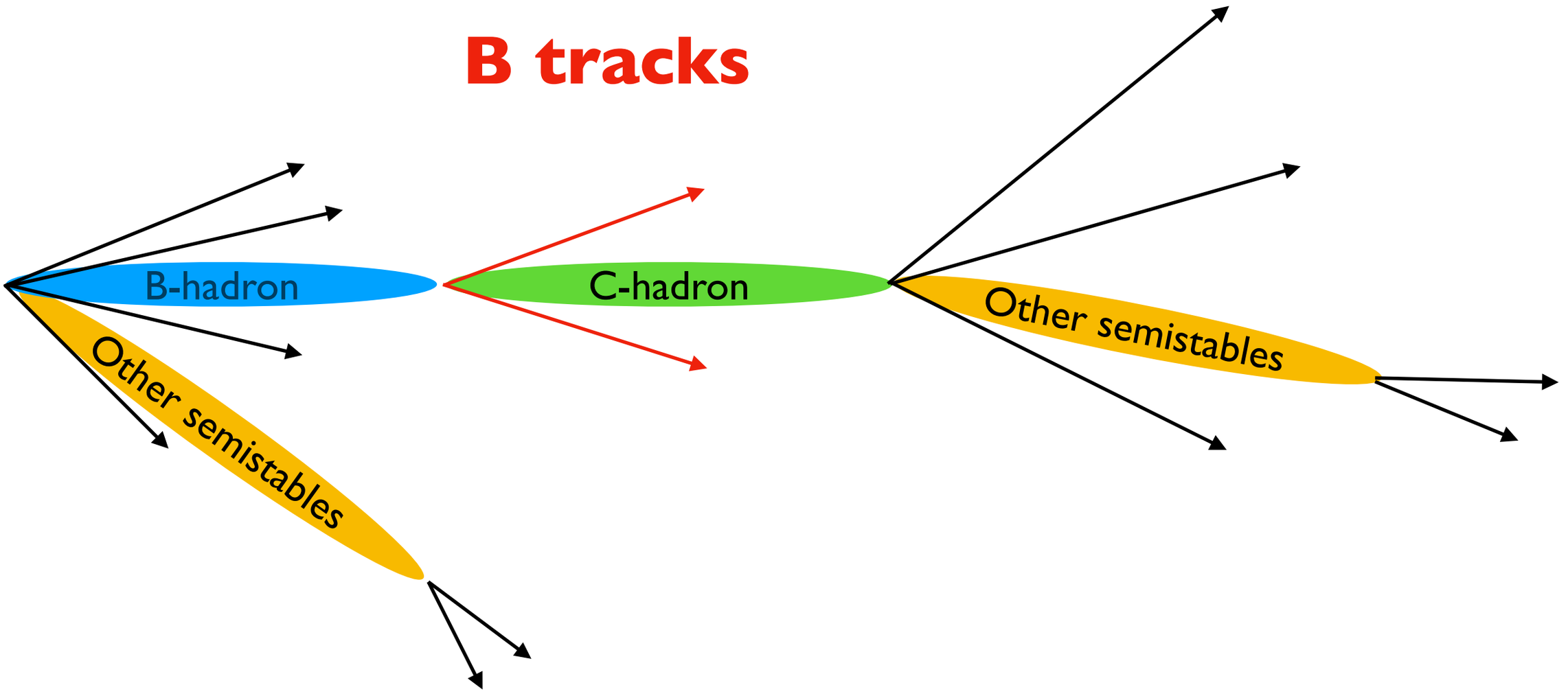
We want to define these vertices as "secondary vertices".

This vertex can actually be defined as secondary vertex,
but we are not interested in this vertex. (—> remove this vertex by V0 rejection algorithm)

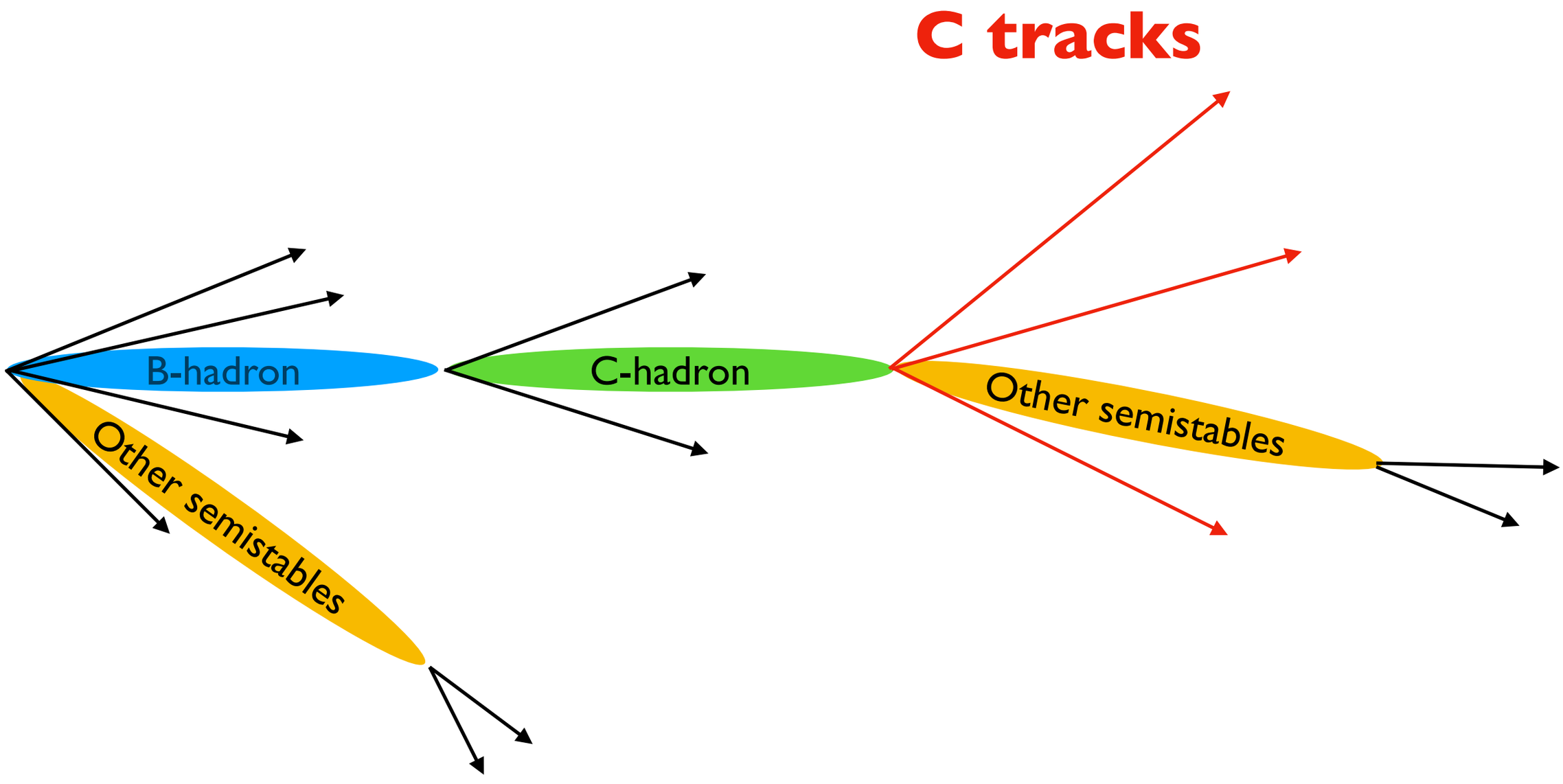
Event signature and term definition



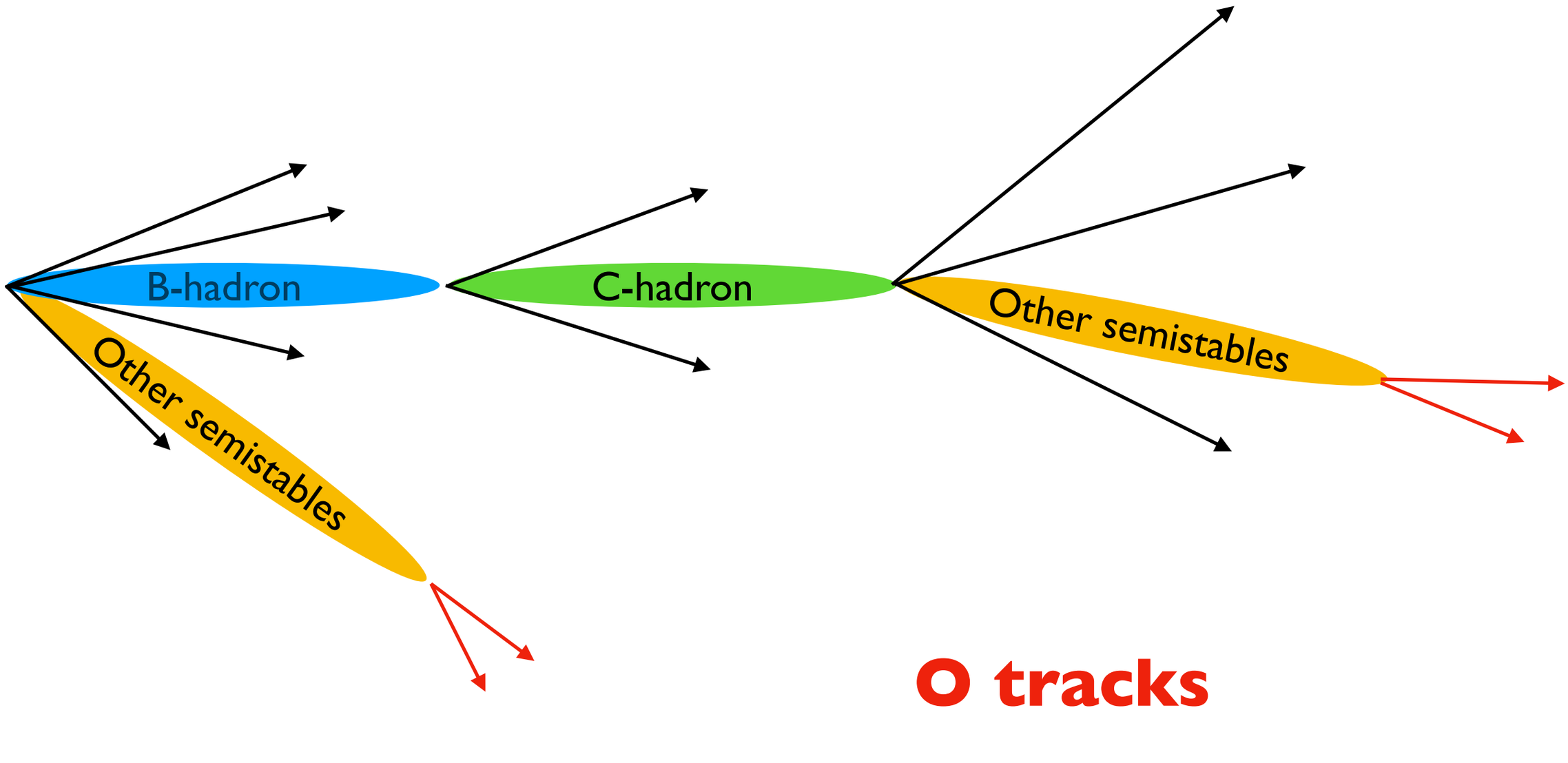
Event signature and term definition



Event signature and term definition



Event signature and term definition



0 tracks

Event signature and term definition

In summary,
we categorized tracks by their ancestors being semi-stables
(e.g. B, D, K, π,...) (if exist).

Definitions of semi-stables which define B, C, O tracks

```
const int semistableBs[] = { 511, 521, 531, 541, 5122, 5132, 5232, 5332};
const int semistableCs[] = { 411, 421, 431, 4122, 4132, 4232, 4332};
const int semistableOs[] = { 11, 13, 15, 22, 130, 211, 310, 321, 2112, 2212, 3112, 3122,
3212, 3222, 3312, 3322, 3334 };
```

of tracks in categories : New (ILD-l5 model) vs DBD

	New	DBD
nPrimaryTracks	= 913039	= 4845286
nPrimaryTracksInSecVtx	= 2023 (0.2 %)	= 11234 (0.2 %)
nBTracks	= 561479	= 2901676
nBTracksInSecVtx	= 334663 (59.6 %)	= 1683847 (58.0 %)
nBTracksInSecVtxCorrectDecayChain	= 331960 (59.1 %)	= 1667652 (57.5 %)
nBTracksInSecVtxCorrectParent	= 204521 (36.4 %)	= 999507 (34.4 %)
nCTracks	= 542794	= 2825299
nCTracksInSecVtx	= 340029 (62.6 %)	= 1697957 (60.1 %)
nCTracksInSecVtxCorrectDecayChain	= 318598 (58.7 %)	= 1588121 (56.2 %)
nCTracksInSecVtxCorrectParent	= 215529 (39.7 %)	= 1052934 (37.3 %)
nOTracks	= 113128	= 715640
nOTracksInSecVtx	= 1412 (1.2 %)	= 21363 (3.0 %)
nOTracksInSecVtxCorrectDecayChain	= 1403 (1.2 %)	= 21240 (3.0 %)
nOTracksInSecVtxCorrectParent	= 730 (0.6 %)	= 12536 (1.8 %)

How to look at this table?
The table divided into four categories (PrimaryTracks, BTracks, CTracks, OTracks) highlighted in green boxes.
- 1st line in each category shows the total number of tracks that were assigned to its category,
- 2nd line in each category shows the number of tracks that were assigned to any secondary vertices (percentage is the value for each 1st line, the same hereafter.),
- 3rd line in each category shows the number of tracks that were assigned to secondary vertices that are in correct decay chains.
- 4th line in each category shows the number of tracks that were assigned to correct vertices

of tracks in categories : New (ILD-s5 model) vs DBD

	New	DBD
nPrimaryTracks	= 889603	= 4845286
nPrimaryTracksInSecVtx	= 1967 (0.2 %)	= 11234 (0.2 %)
nBTracks	= 556241	= 2901676
nBTracksInSecVtx	= 332900 (59.8 %)	= 1683847 (58.0 %)
nBTracksInSecVtxCorrectDecayChain	= 330234 (59.4 %)	= 1667652 (57.5 %)
nBTracksInSecVtxCorrectParent	= 203824 (36.6 %)	= 999507 (34.4 %)
nCTracks	= 538712	= 2825299
nCTracksInSecVtx	= 338856 (62.9 %)	= 1697957 (60.1 %)
nCTracksInSecVtxCorrectDecayChain	= 317558 (58.9 %)	= 1588121 (56.2 %)
nCTracksInSecVtxCorrectParent	= 215088 (39.9 %)	= 1052934 (37.3 %)
nOTracks	= 108402	= 715640
nOTracksInSecVtx	= 1424 (1.3 %)	= 21363 (3.0 %)
nOTracksInSecVtxCorrectDecayChain	= 1413 (1.3 %)	= 21240 (3.0 %)
nOTracksInSecVtxCorrectParent	= 767 (0.7 %)	= 12536 (1.8 %)

Almost same results as ILD-I5 model.
ILD-I5 model seems slightly better (in terms of absolute values, not fraction).

500GeV(bbbbbb), ILD_l5 case
w/ beam background

		DBD 91GeV(bb)
nPrimaryTracks	= 66596	= 4845286
nPrimaryTracksInSecVtx	= 304 (0.5 %)	= 11234 (0.2 %)
nBTracks	= 1595853	= 2901676
nBTracksInSecVtx	= 1079292 (67.6 %)	= 1683847 (58.0 %)
nBTracksInSecVtxCorrectDecayChain	= 1061042 (66.5 %)	= 1667652 (57.5 %)
nBTracksInSecVtxCorrectParent	= 688894 (43.2 %)	= 999507 (34.4 %)
nCTracks	= 1516893	= 2825299
nCTracksInSecVtx	= 1105700 (72.9 %)	= 1697957 (60.1 %)
nCTracksInSecVtxCorrectDecayChain	= 1086130 (71.6 %)	= 1588121 (56.2 %)
nCTracksInSecVtxCorrectParent	= 731048 (48.2 %)	= 1052934 (37.3 %)
nOTracks	= 4241905	= 715640
nOTracksInSecVtx	= 87326 (2.1 %)	= 21363 (3.0 %)
nOTracksInSecVtxCorrectDecayChain	= 55674 (1.3 %)	= 21240 (3.0 %)
nOTracksInSecVtxCorrectParent	= 22263 (0.5 %)	= 12536 (1.8 %)

vertexing step doesn't looks that bad

500GeV(bbbbbb), lLD_s5 case
w/ beam background

		DBD 91GeV(bb)
nPrimaryTracks	= 62512	= 4845286
nPrimaryTracksInSecVtx	= 265 (0.4 %)	= 11234 (0.2 %)
nBTracks	= 1587520	= 2901676
nBTracksInSecVtx	= 1075895 (67.8 %)	= 1683847 (58.0 %)
nBTracksInSecVtxCorrectDecayChain	= 1058797 (66.7 %)	= 1667652 (57.5 %)
nBTracksInSecVtxCorrectParent	= 688357 (43.4 %)	= 999507 (34.4 %)
nCTracks	= 1510978	= 2825299
nCTracksInSecVtx	= 1102388 (73.0 %)	= 1697957 (60.1 %)
nCTracksInSecVtxCorrectDecayChain	= 1083469 (71.7 %)	= 1588121 (56.2 %)
nCTracksInSecVtxCorrectParent	= 730151 (48.3 %)	= 1052934 (37.3 %)
nOTracks	= 4125745	= 715640
nOTracksInSecVtx	= 84513 (2.0 %)	= 21363 (3.0 %)
nOTracksInSecVtxCorrectDecayChain	= 54151 (1.3 %)	= 21240 (3.0 %)
nOTracksInSecVtxCorrectParent	= 22884 (0.6 %)	= 12536 (1.8 %)

vertexing step doesn’t looks that bad

Flavour tagging

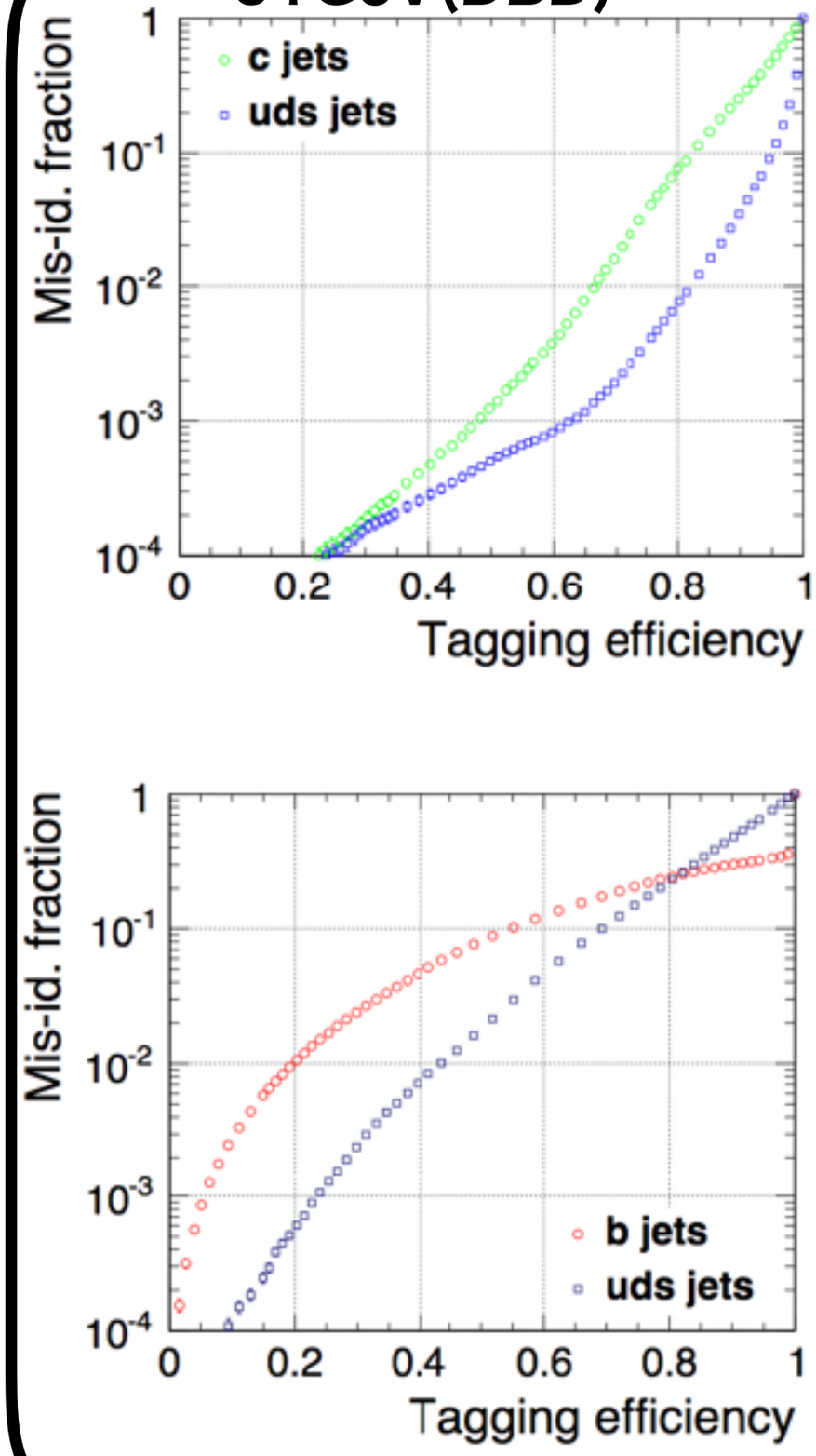
500GeV(6b,6c,6q), ILD_l5 case

(trained at 91GeV)

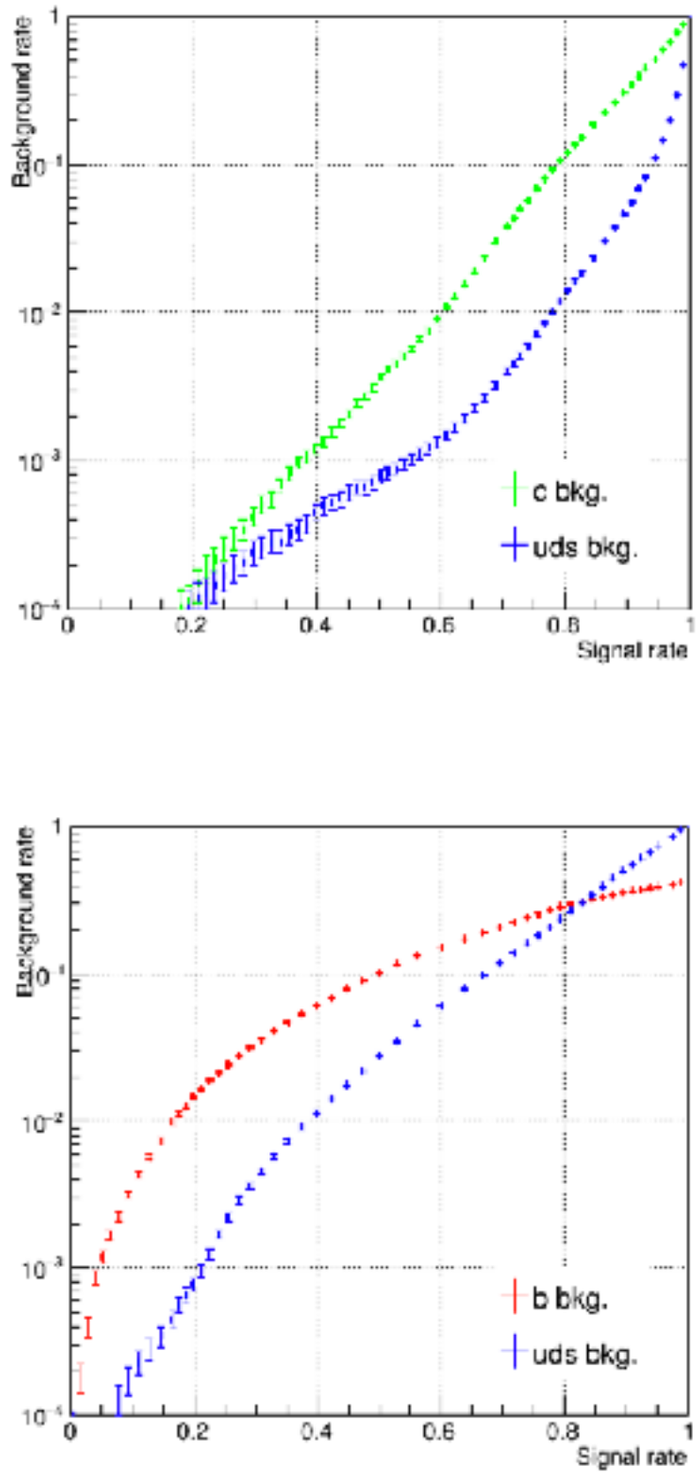
Beam bkg rejection off

(trained at 500GeV)

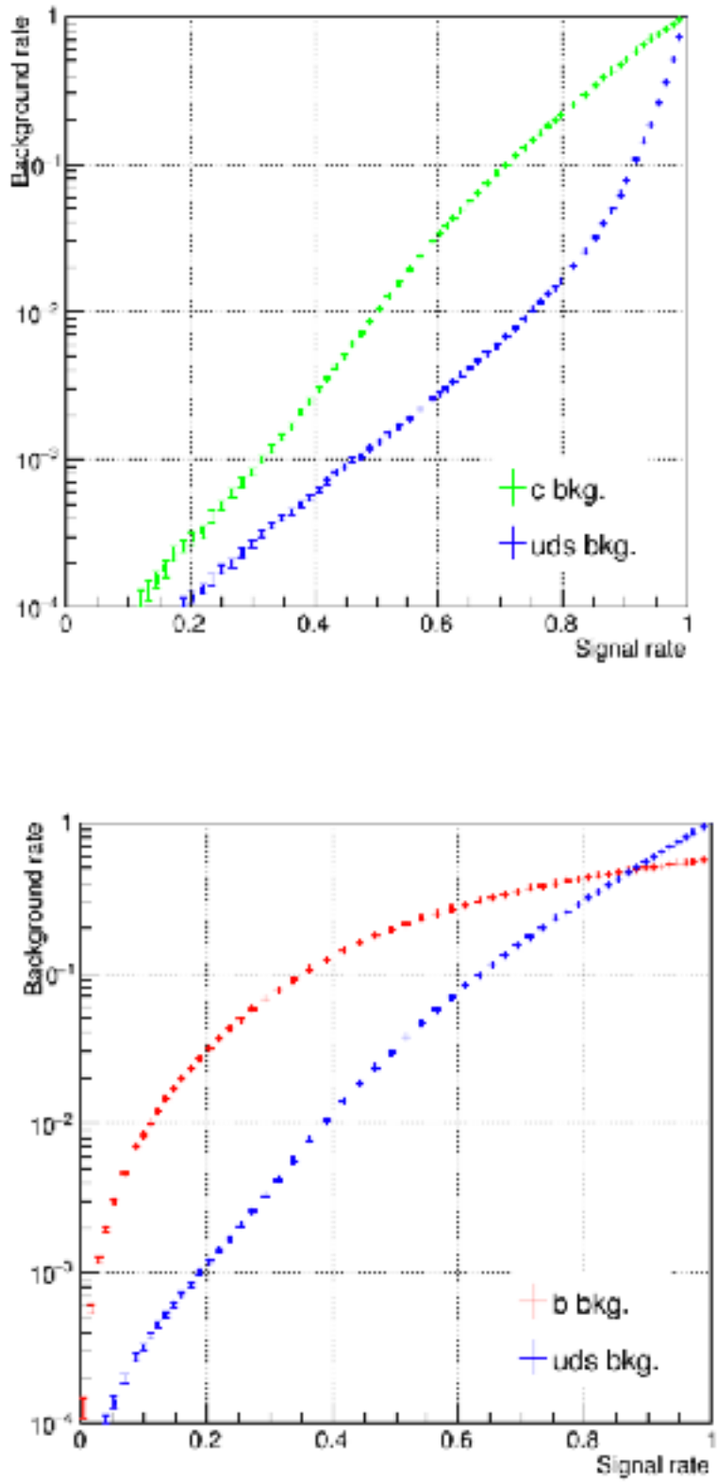
91GeV(DBD)



91GeV(l5)



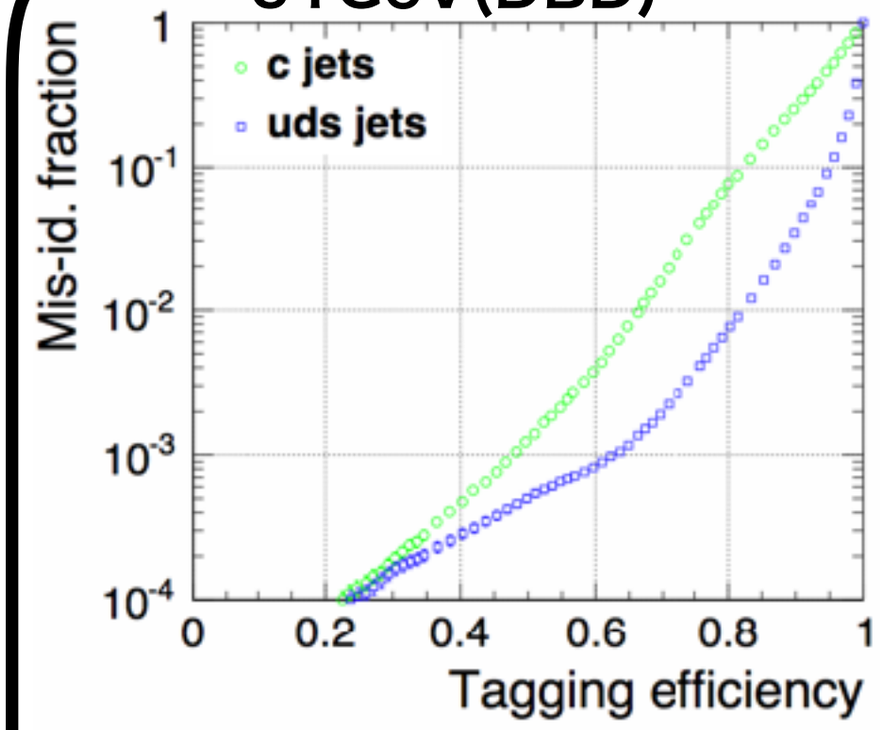
500 GeV(l5)

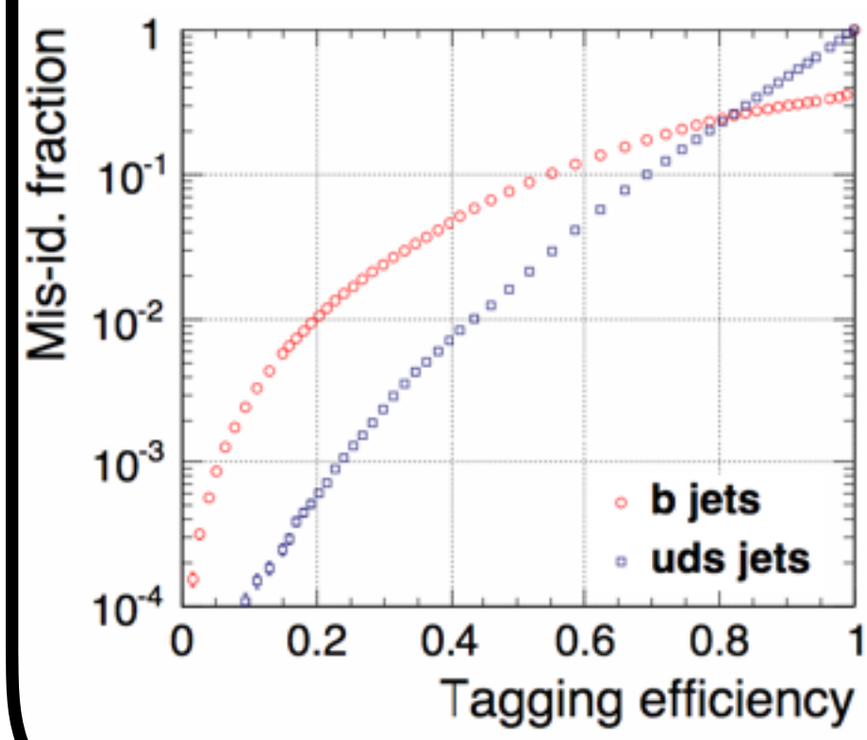


Flavour tagging

500GeV(6b,6c,6q), ILD_s5 case

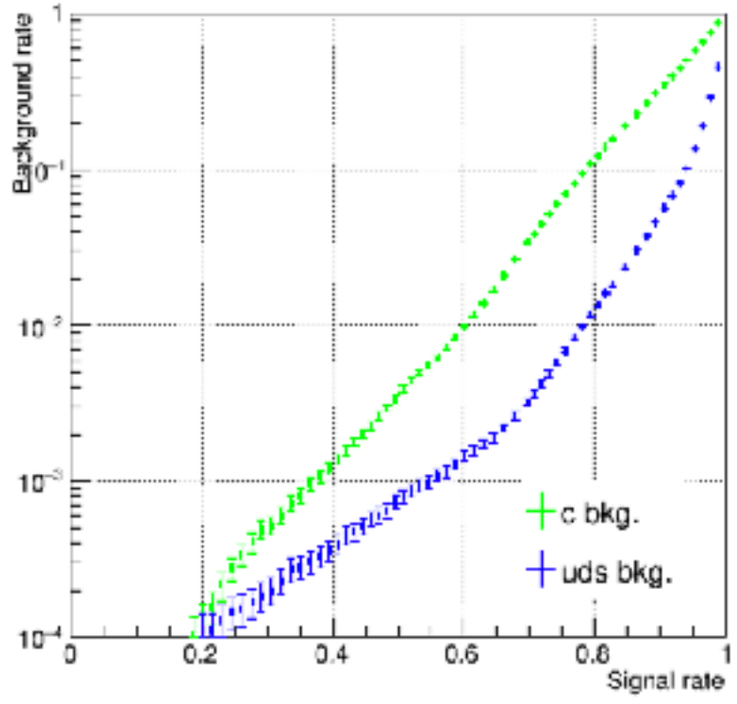
91GeV(DBD)

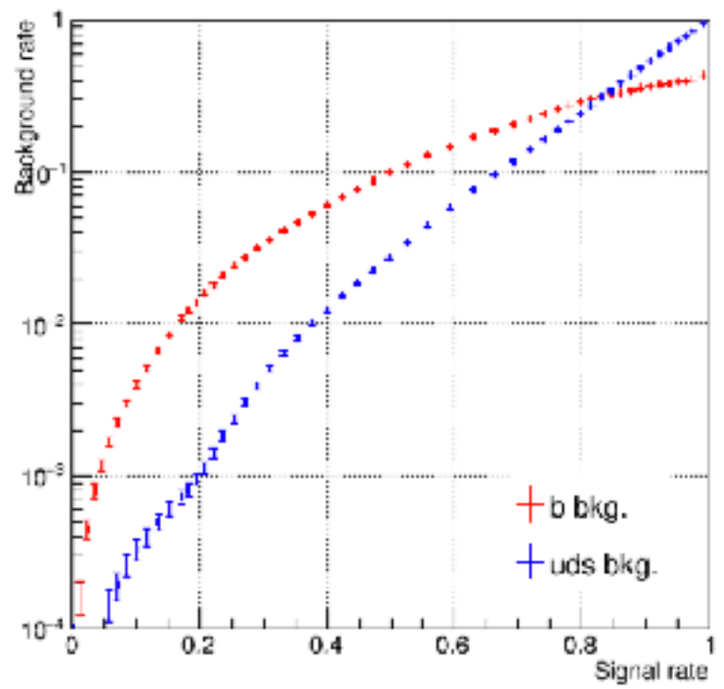




(trained at 91GeV)

91GeV(s5)





Beam bkg rejection off (trained at 500GeV)

500 GeV (s5)