

# **LCFIPPlus performance test with new MC samples for ILD**

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

# What's LCFIPlus?

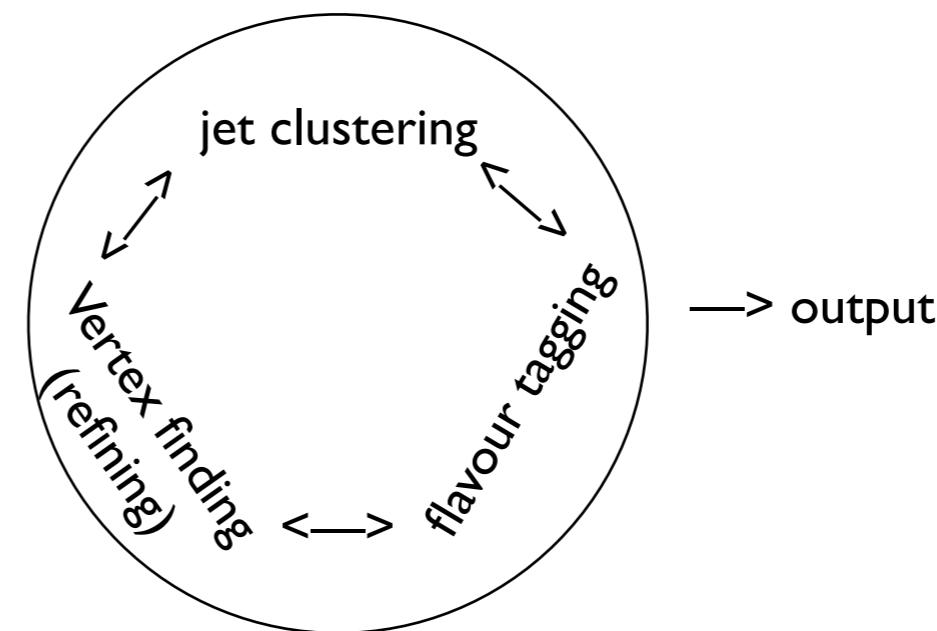
## ❖ A framework for jet flavour identification.

- ▶ does vertex finding, jet clustering, and flavour tagging
- ▶ each process is implemented as a modular algorithm.
  - ▶ gives flexibility to iterate or reverse the processes.

w/o LCFIPlus :

jet clustering —> vertex finding —> flavour tagging —> output

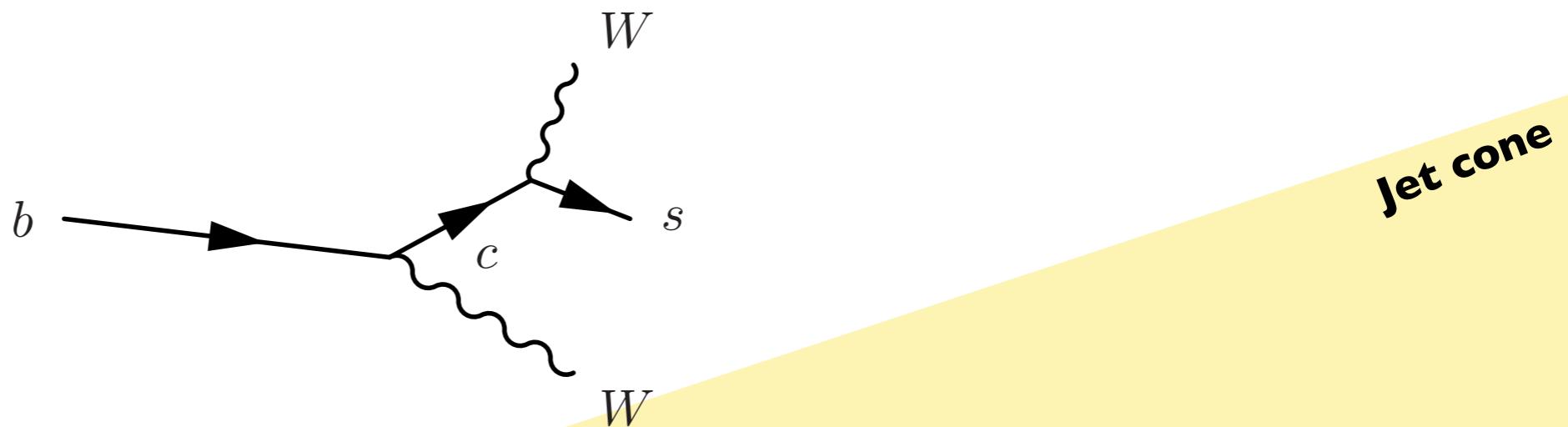
w/ LCFIPlus :



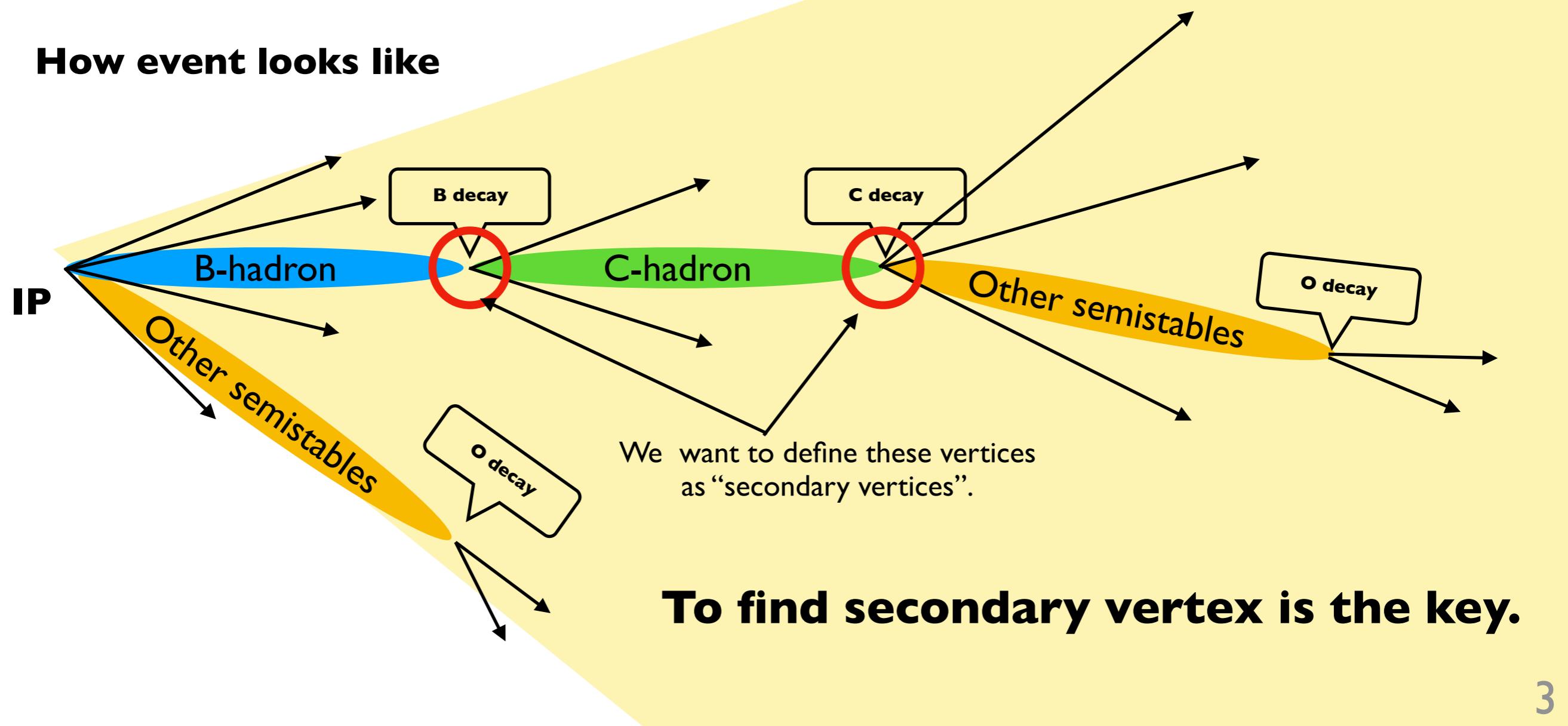
- ▶ It is flexible but typical usage is very simple;  
“vertex finding —> jet clustering —> vertex refining —>flavour tagging”
- ▶ originated from LCFIVertex (e.g. arXiv:0908.3019)

# Principle of b-tag and c-tag

Feynman diagram



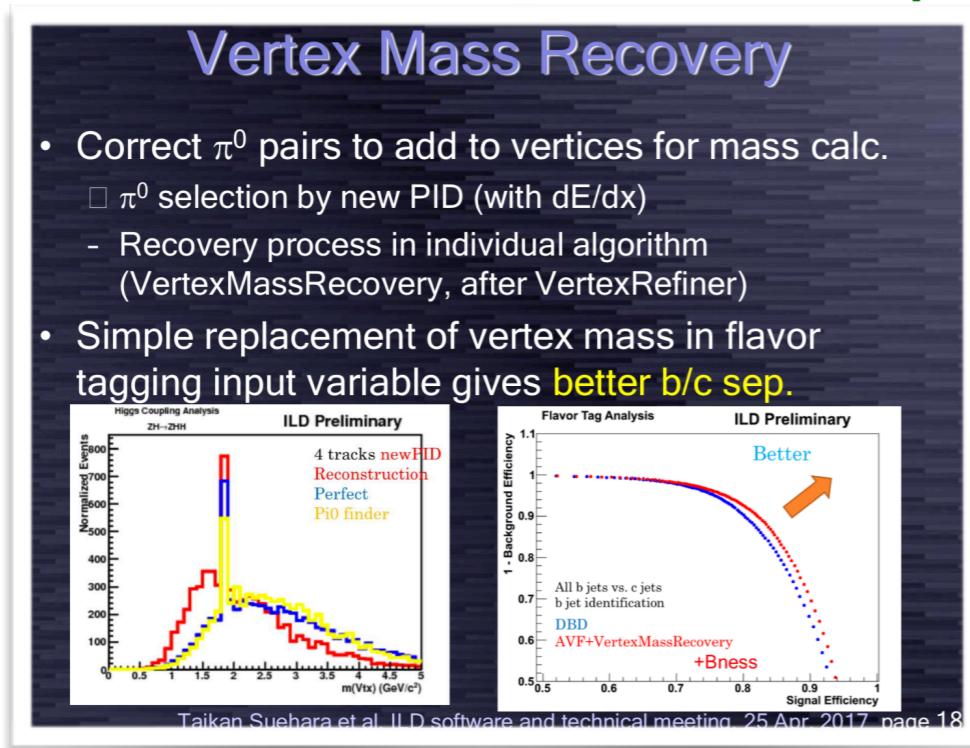
How event looks like



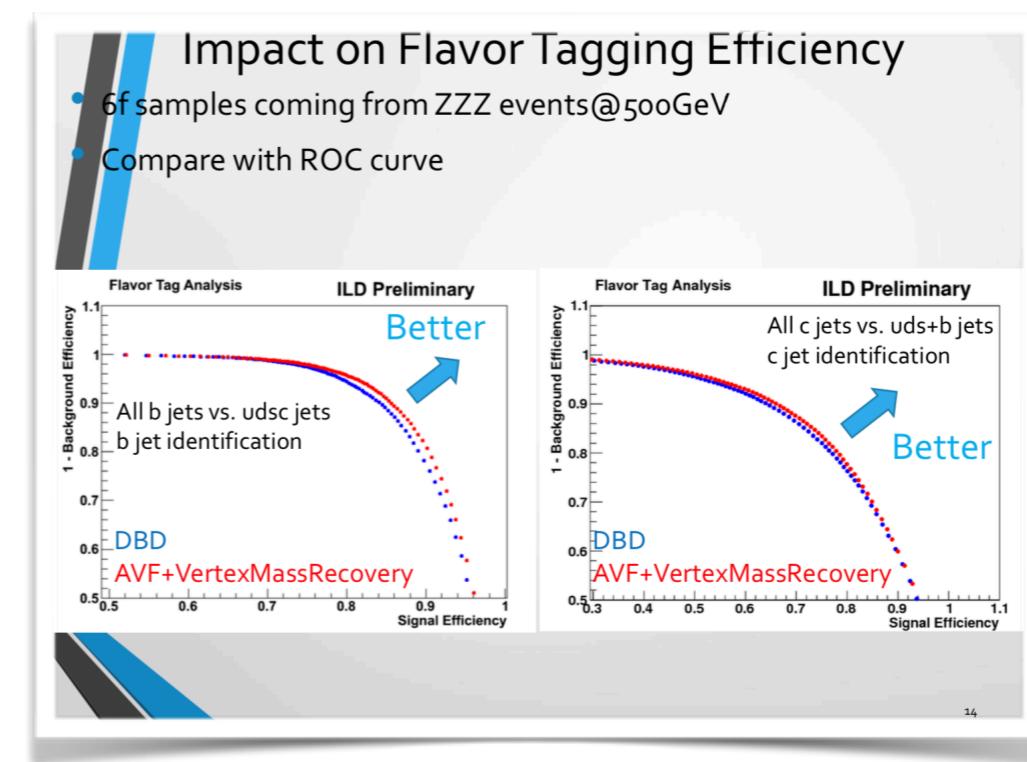
# Current concerns

## ❖ Flavour tagging performance

- ▶ Taikan and Masakazu have reported that new features (e.g. adaptive vertex fitting, vertex mass recovery, etc) improve the performance using previous ilcsoft versions. This should also be confirmed with the latest release. This is our next step.



ILD software and technical meeting, Apr. 2017



LCWS@strasbourg, Oct. 2017

- ▶ Test samples has been produced with newer versions.
  - ▶ To prevent problems mixed up, the new features above are off for now.

# Process overview

**Particle Flow Object (PFO)** **LCFIPlus input**

**Charged particles**      **Neutral particles**

**Vertex finding**

- primary vertex finding,
- secondary vertex finding,
- V0 vertex finding

**Jet clustering & Vertex refiner**

**Jet flavour tagging**

**Jets with b, c probabilities**

**LCFIPlus output**

**PFO** is an object created in Particle Flow Algorithm which aims for the best jet energy resolution with whole detector system.

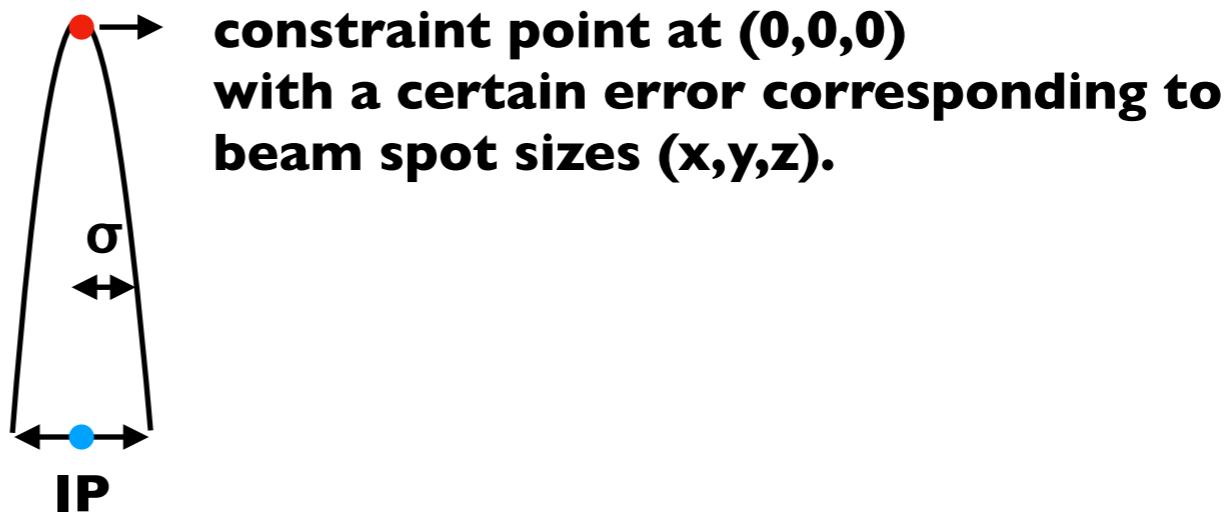
In this presentation, it will do just to take PFO as “Best estimated reconstructed particle”.

# How do we check ?

- ❖ **Compare with DBD sample at 91GeV and 500GeV**
  - ▶ DBD : Detailed Baseline Design report
  - ▶ Two energy samples :
    - ▶  $\sqrt{s}=91\text{GeV}$  , processes = 2b, 2c, 2q ( $q=u,d,s$ )
    - ▶  $\sqrt{s}=500\text{GeV}$  , processes = 6b, 6c, 6q ( $q=u,d,s$ )
- ❖ **Few words on new samples**
  - ▶ Two detector models; I5 (and s5).
  - ▶ “I5” (and “s5”) will be used to refer to these new samples.
  - ▶ IP = (0,0,0) for 91 GeV sample. (same as DBD)
  - ▶ IP is smeared in simulation for 500 GeV. (different from DBD)
- ❖ **For all samples (DBD, I5, s5)**
  - ▶ Use recent software to perform vertexing, jet clustering, vertex refining, and flavour tagging.

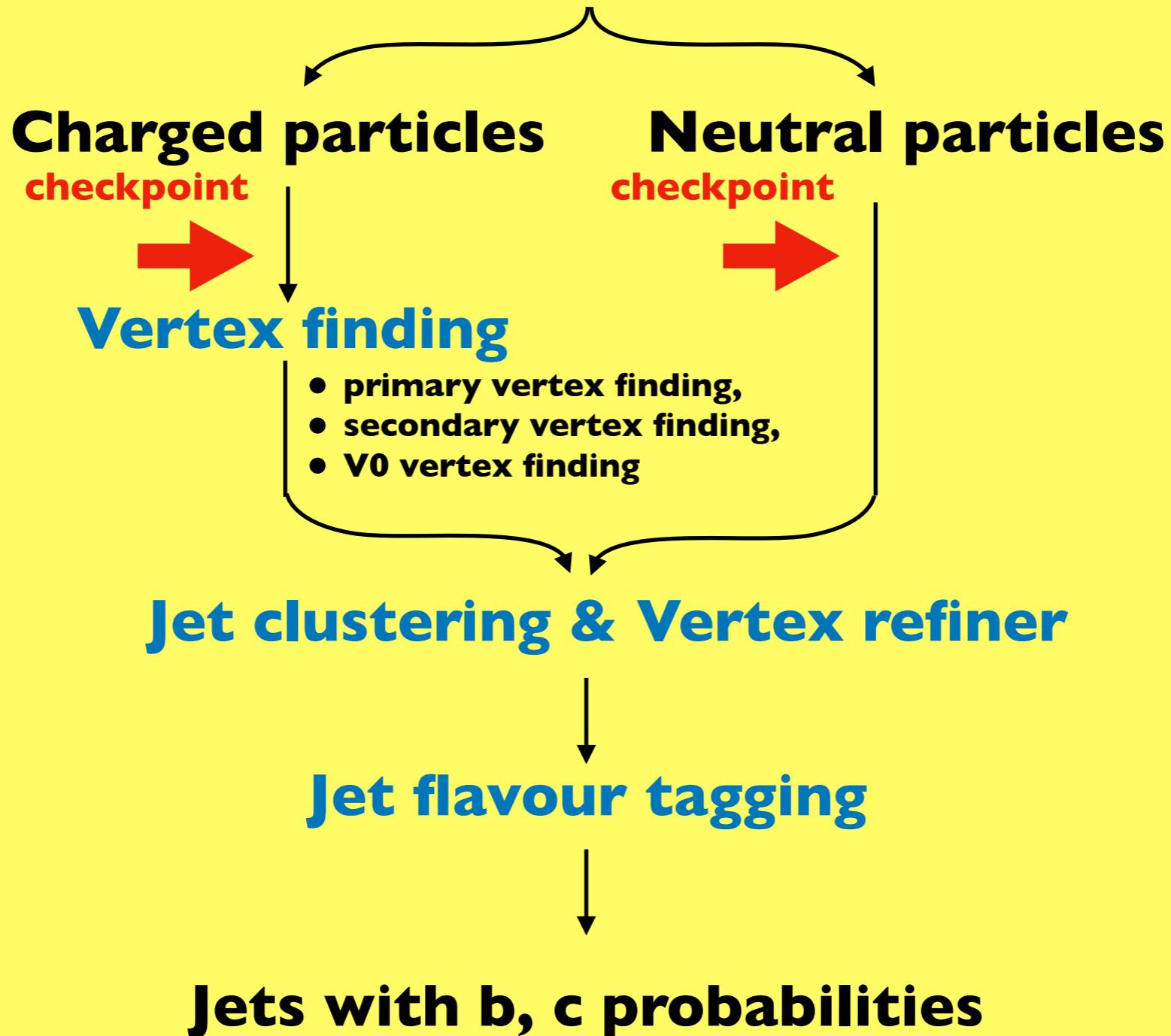
# IP smearing

- ❖ In primary vertex finding, beam constraint is useful to reject non-primary tracks. Use an additional point (=constraint point) in primary vertex fitting.

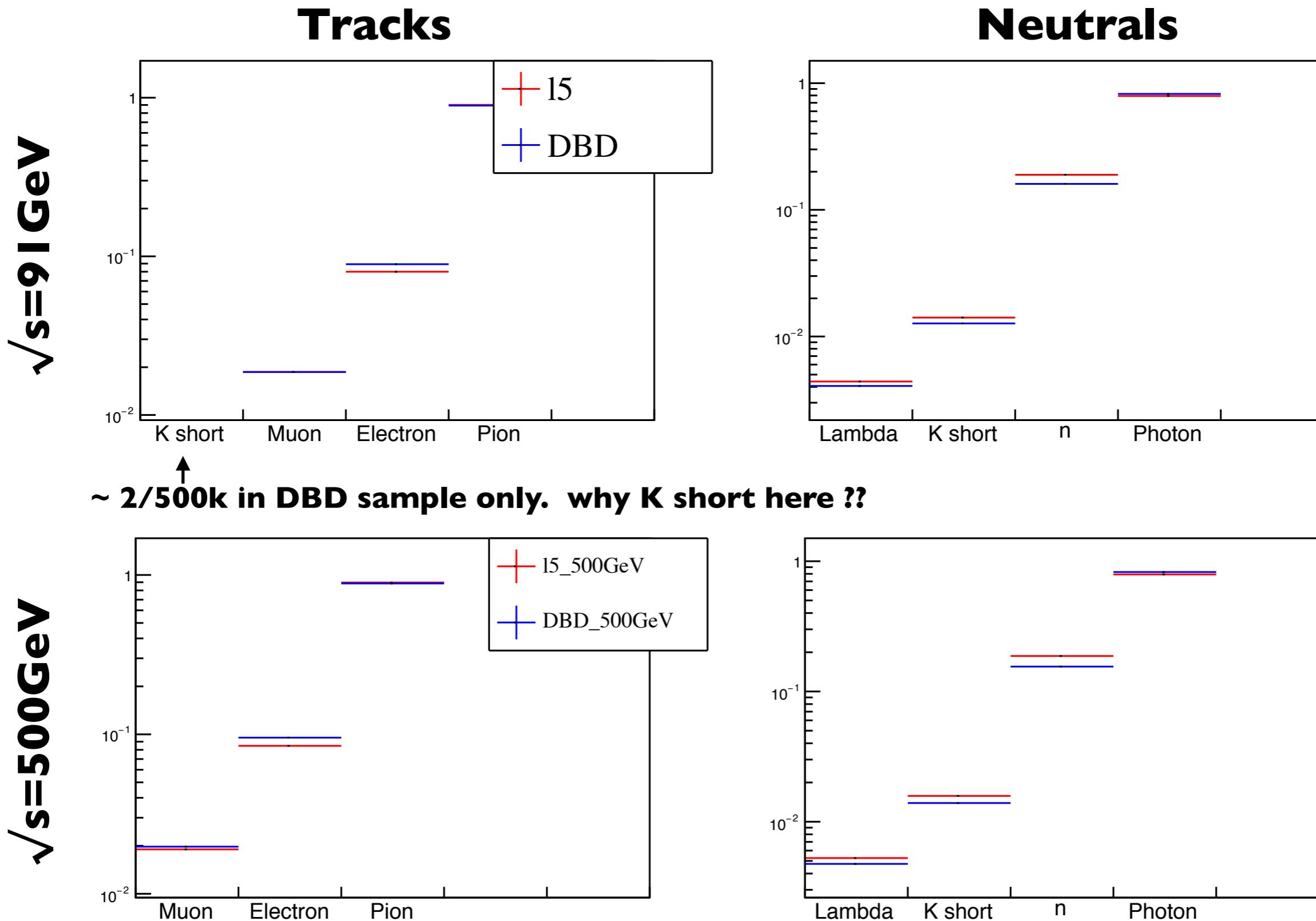


- ❖ In real experiment, IP is distributed in luminous region.
  - ▶ In DBD samples, IP is always fixed at (0,0,0)
  - ▶ Instead, we smear the centre point of the constraint with sigmas comparable to beam spot sizes in LCFIPlus.
  - ▶ In recent production, IP is smeared in simulation step. No need “constraint point” smearing in LCFIPlus.

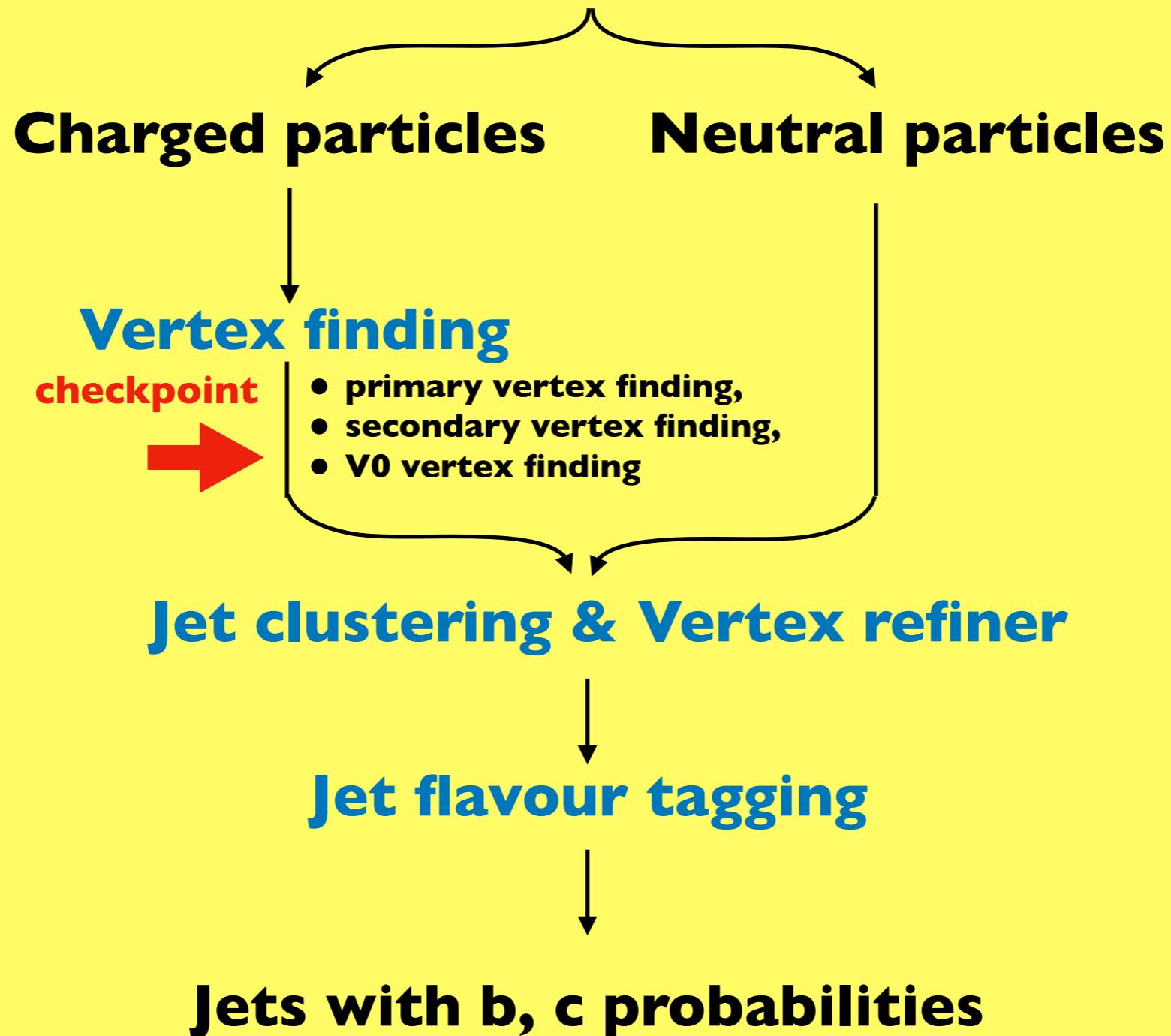
# Particle Flow Object (PFO)



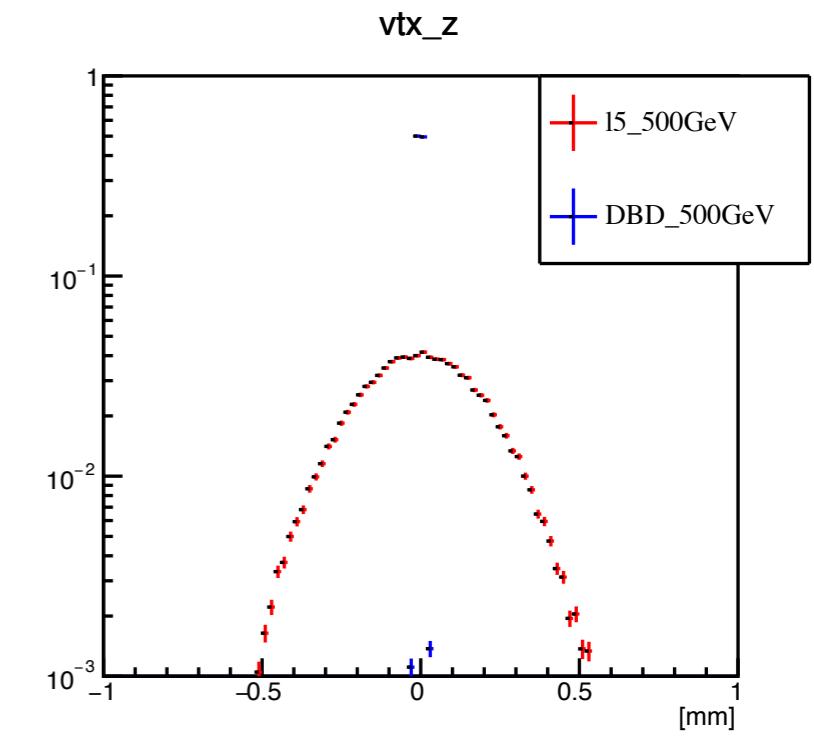
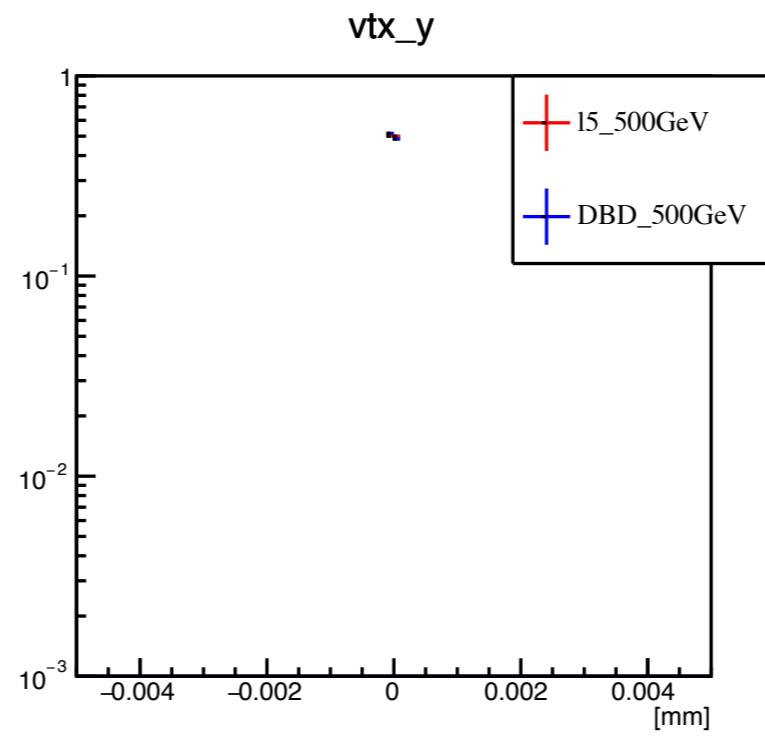
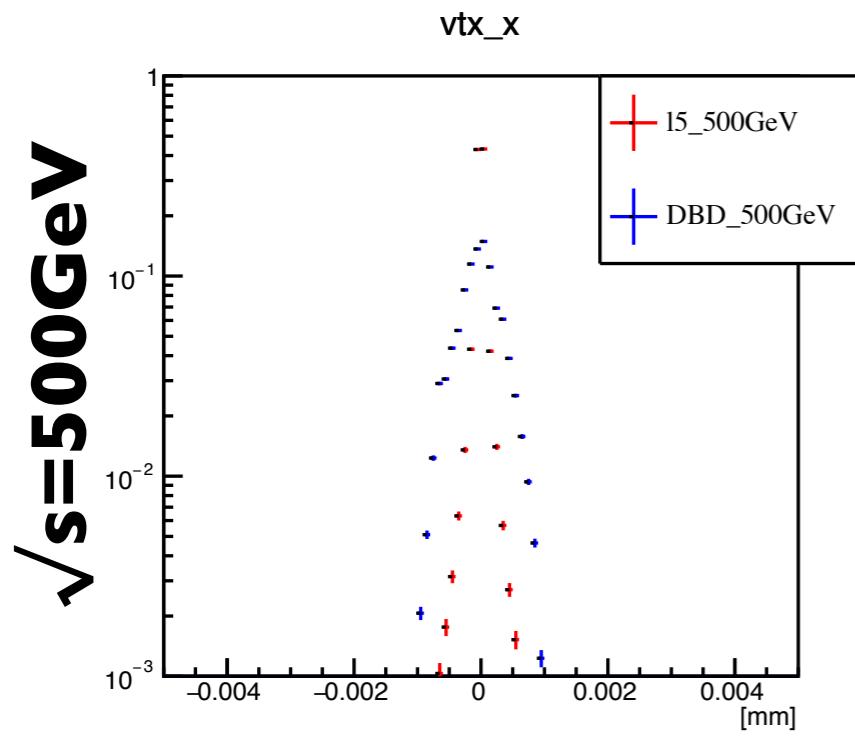
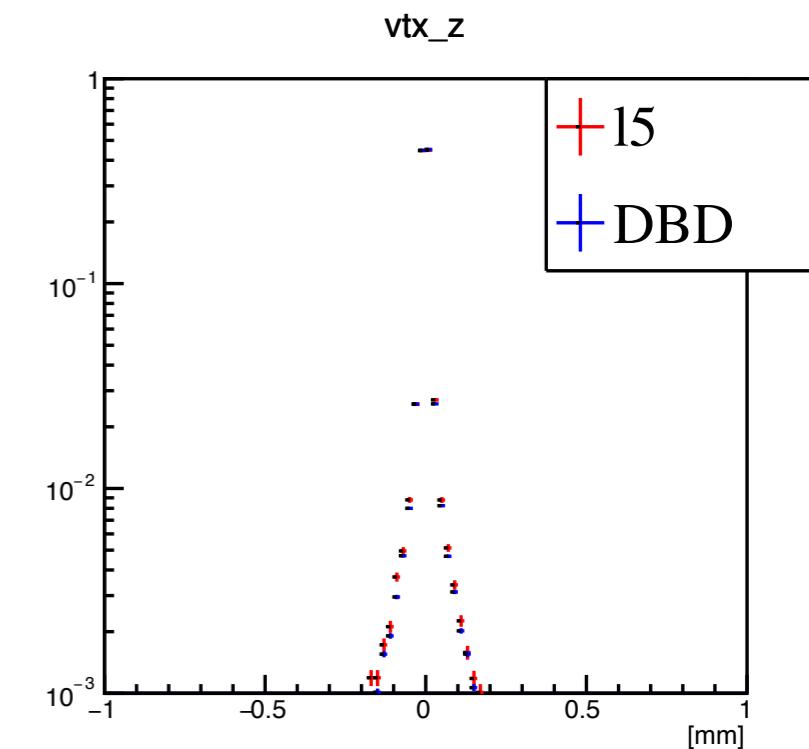
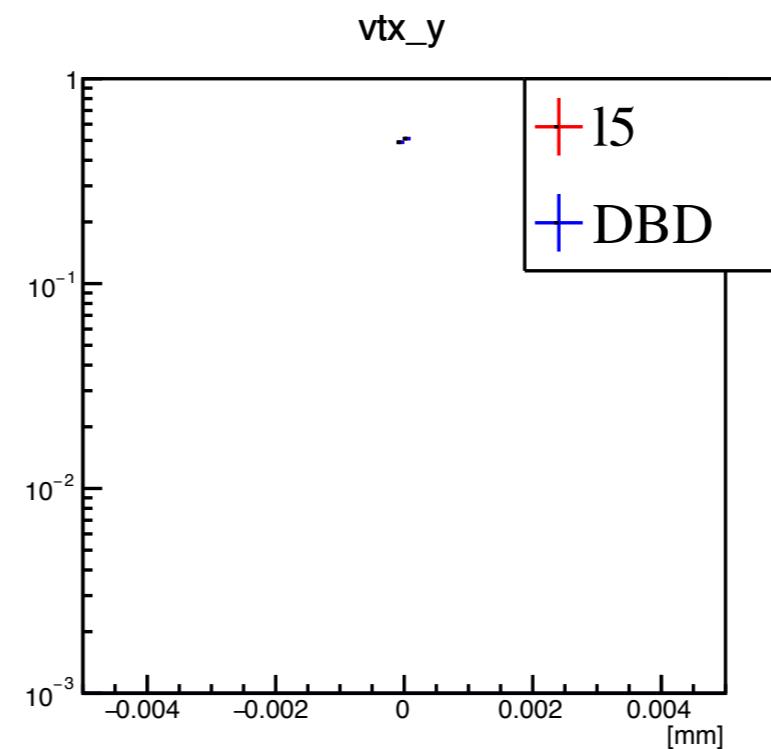
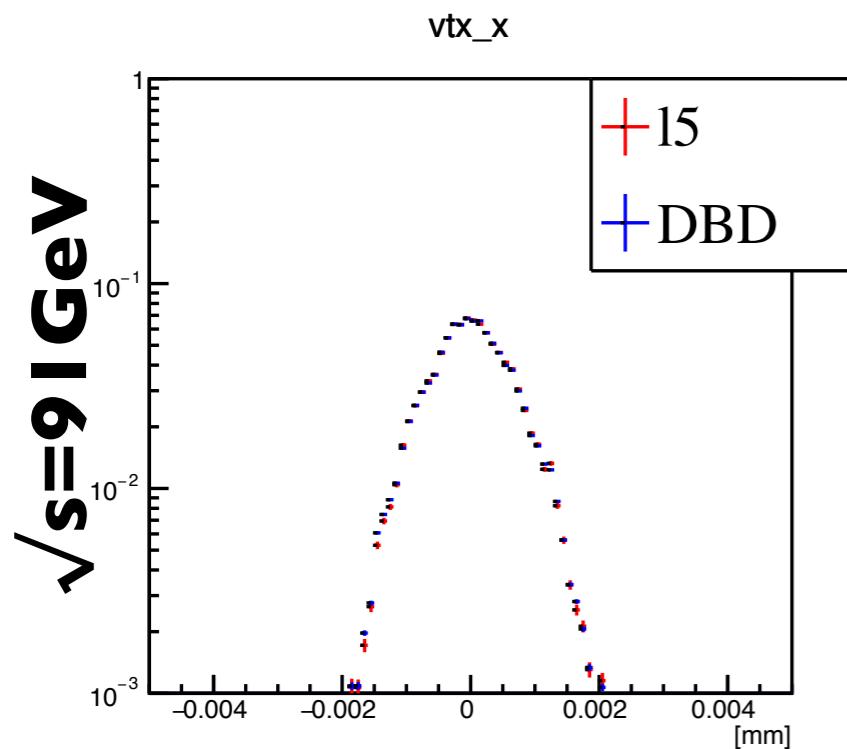
# PID profiles (PFO->getType())



# Particle Flow Object (PFO)

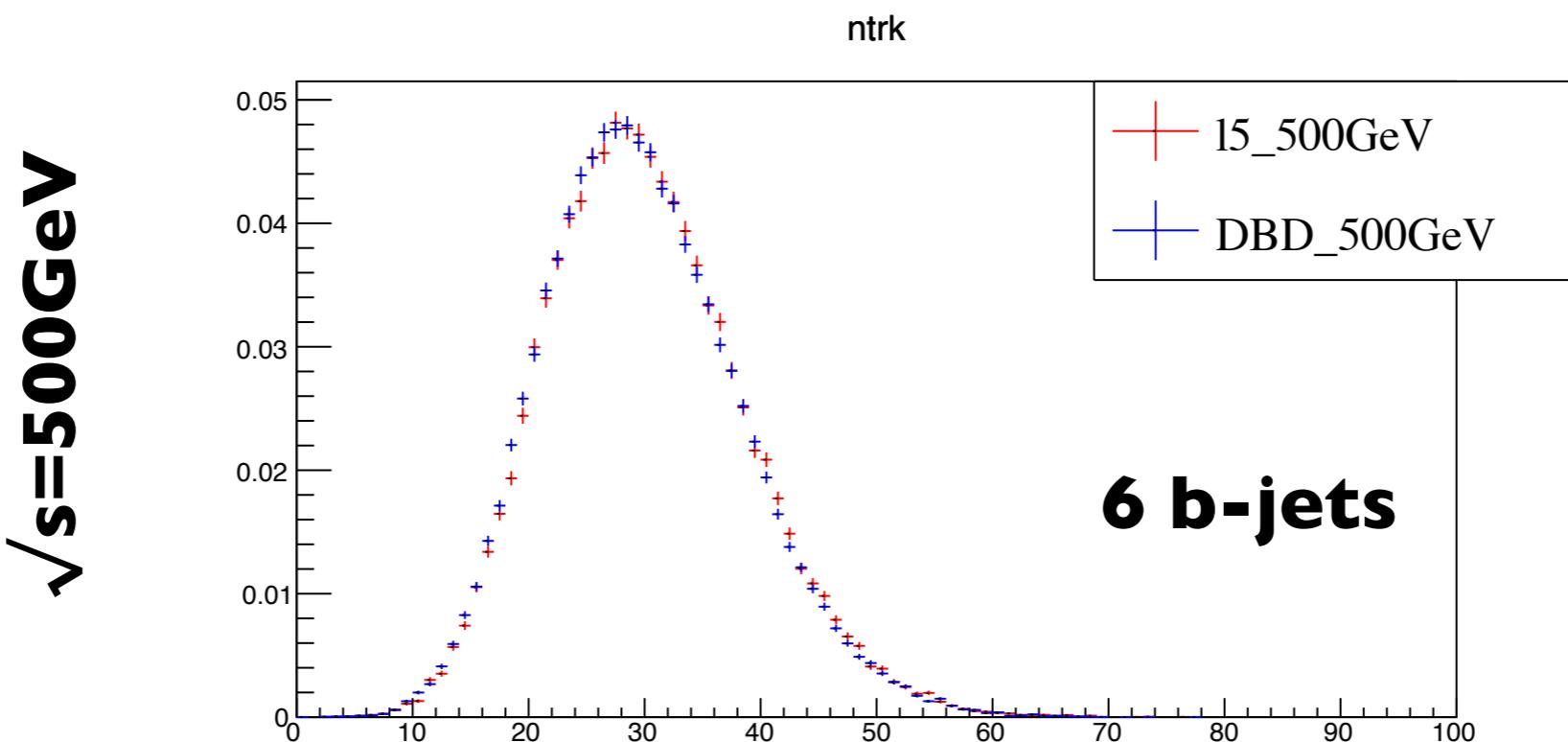
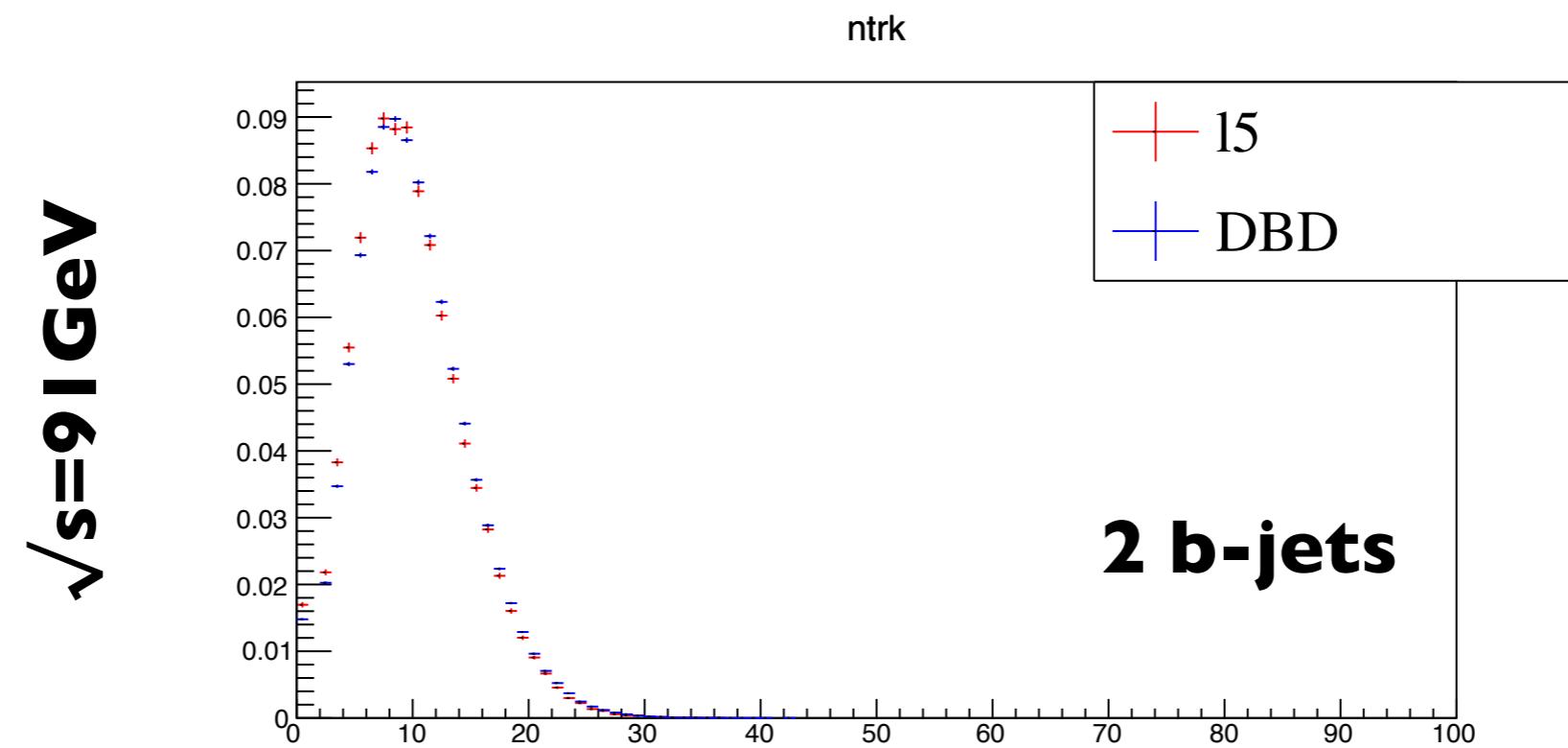


# Primary vertex finding (Position)

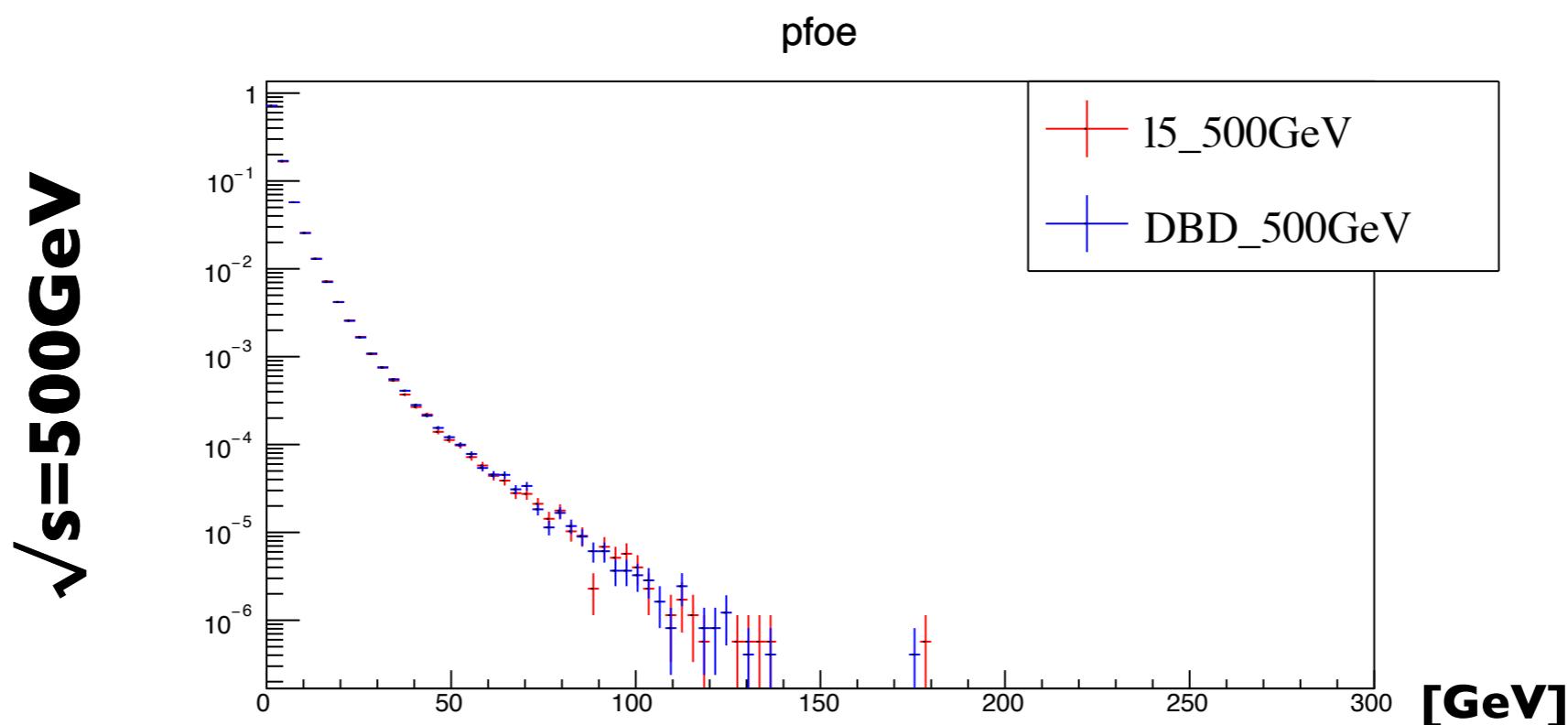
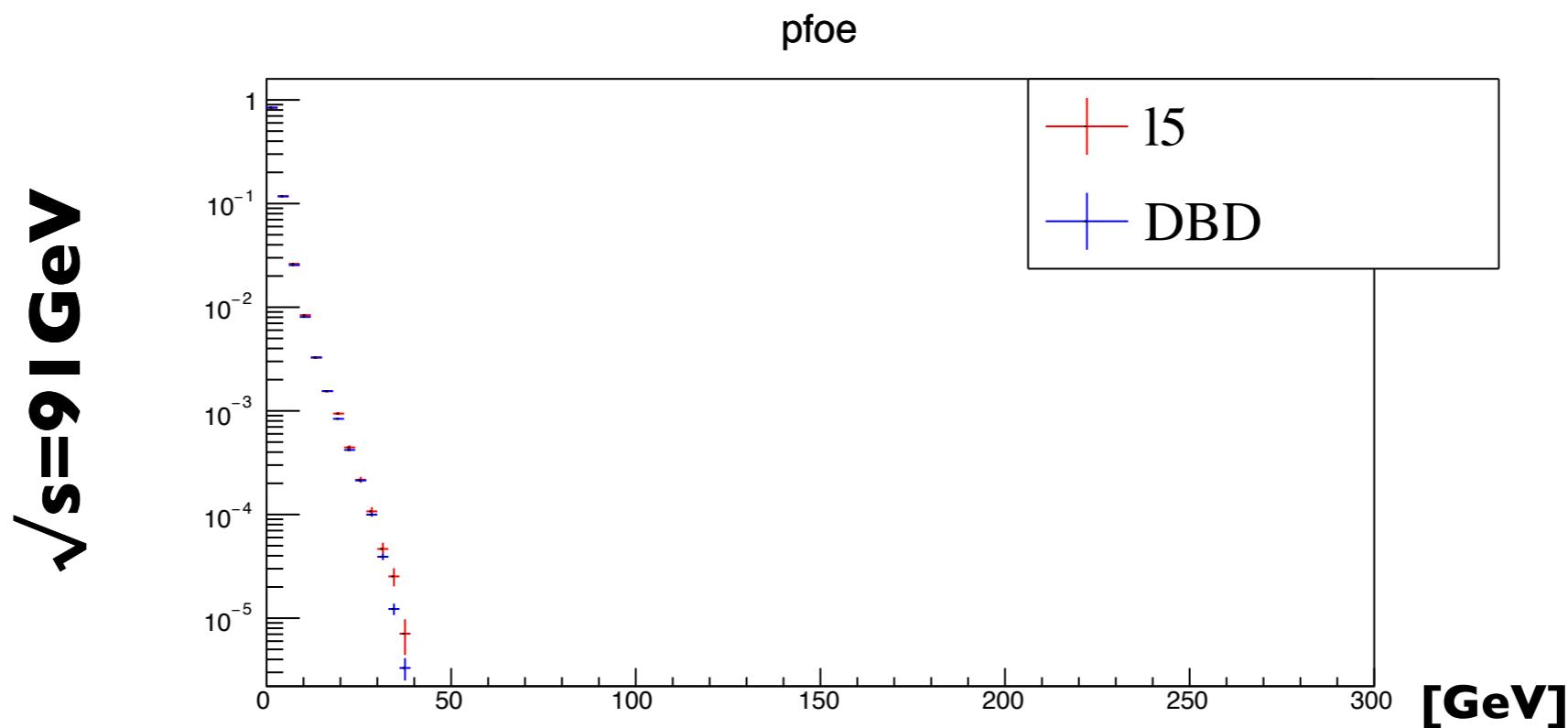


The difference comes from IP smearing.

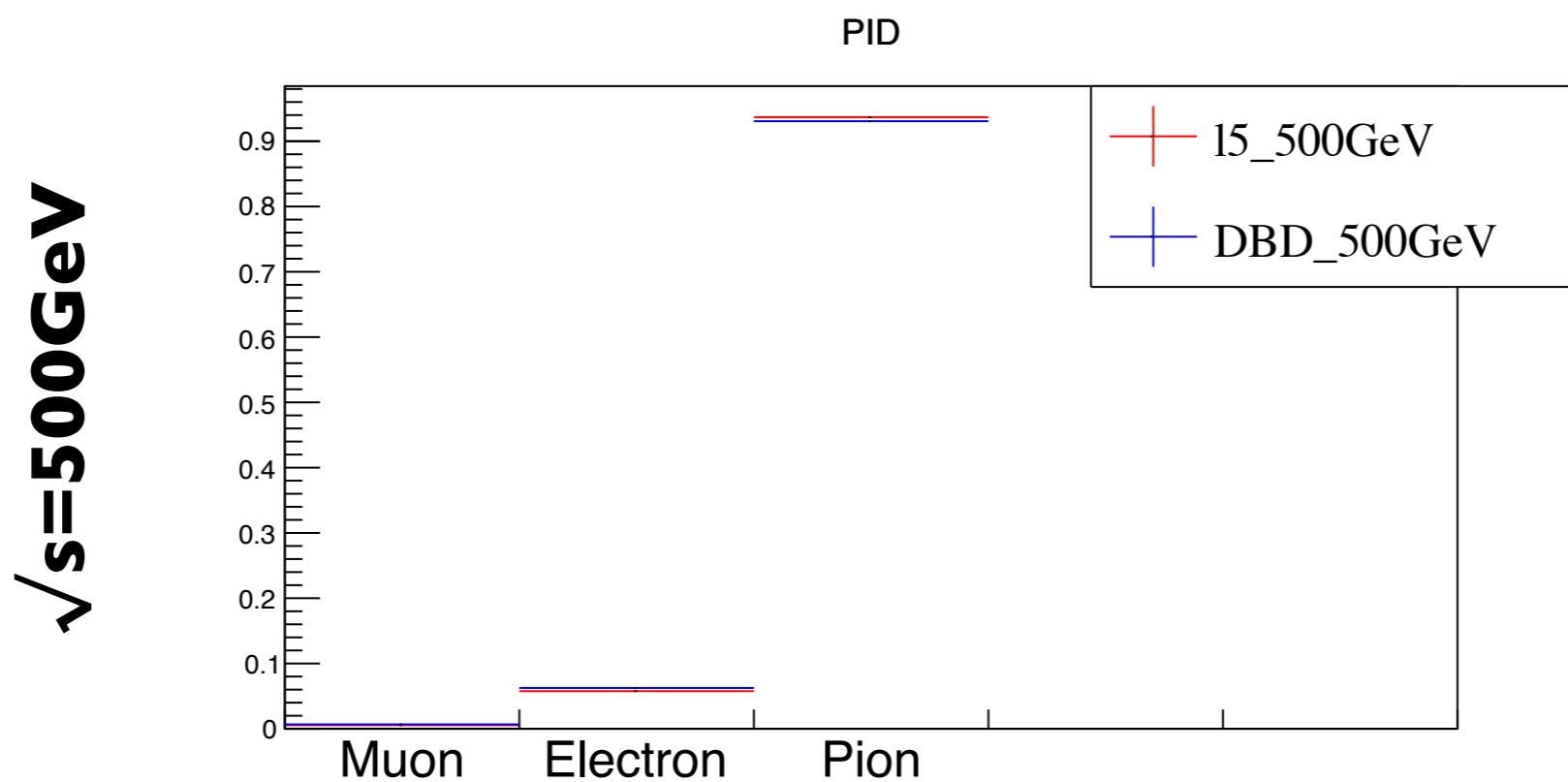
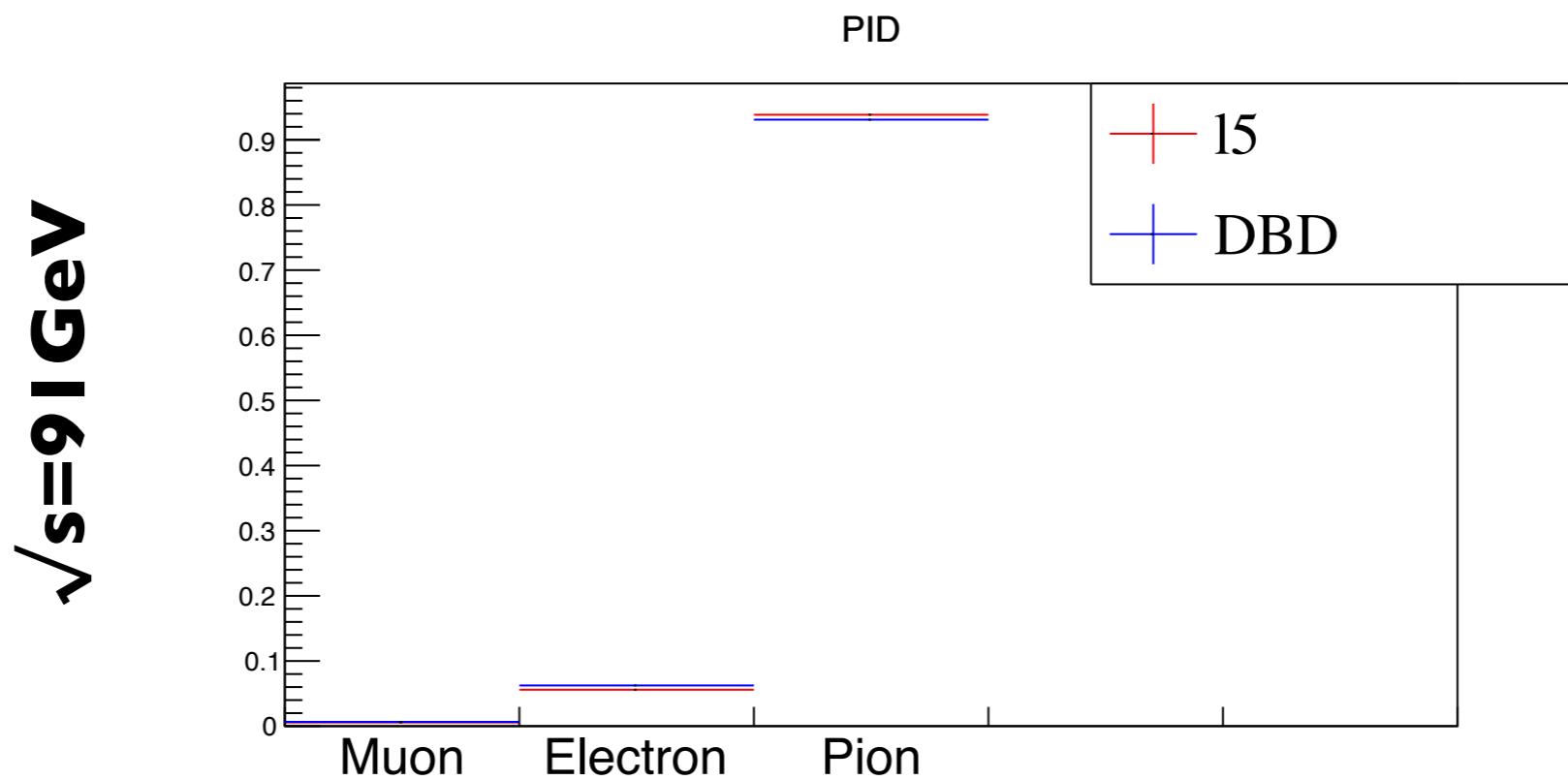
# Primary vertex finding (# of tracks)



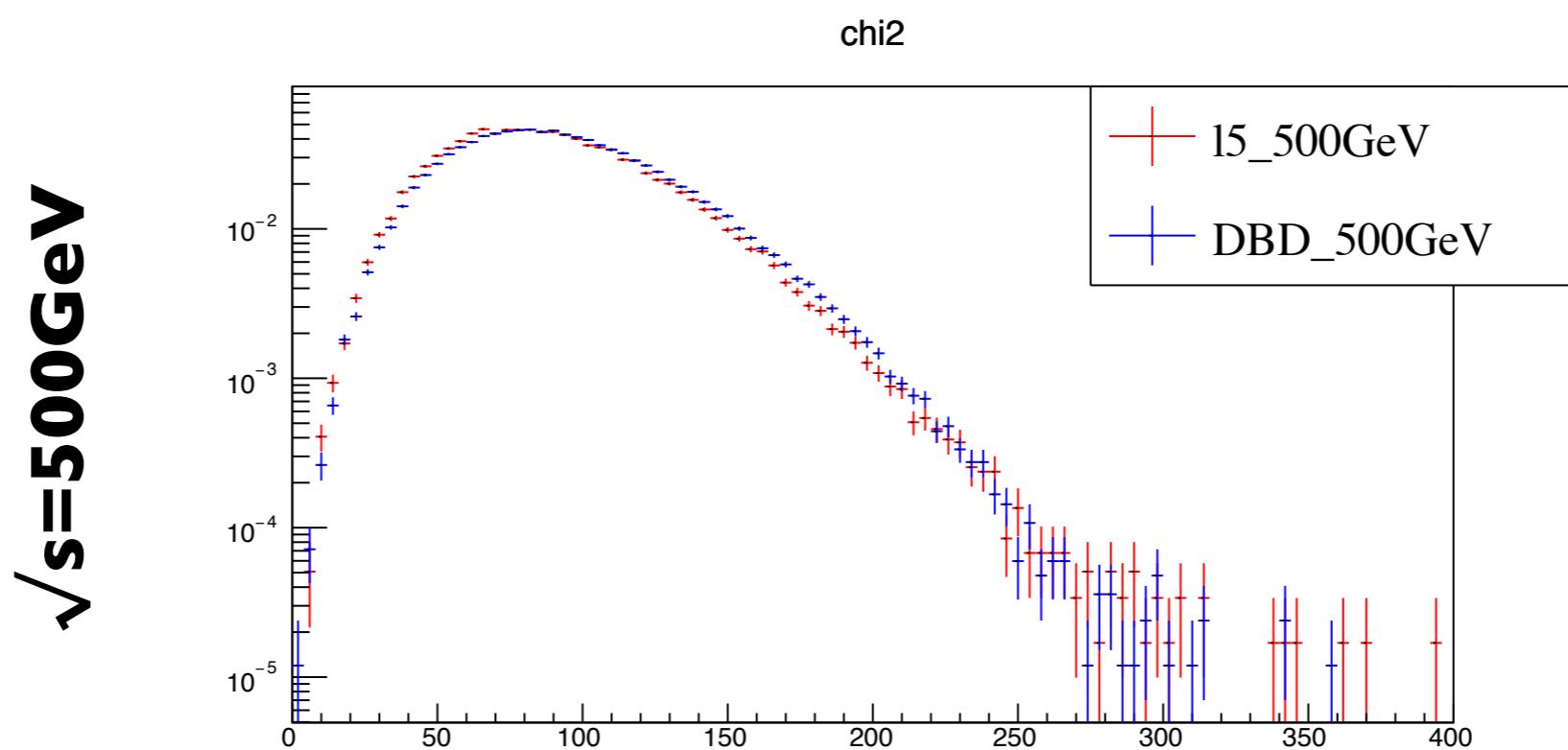
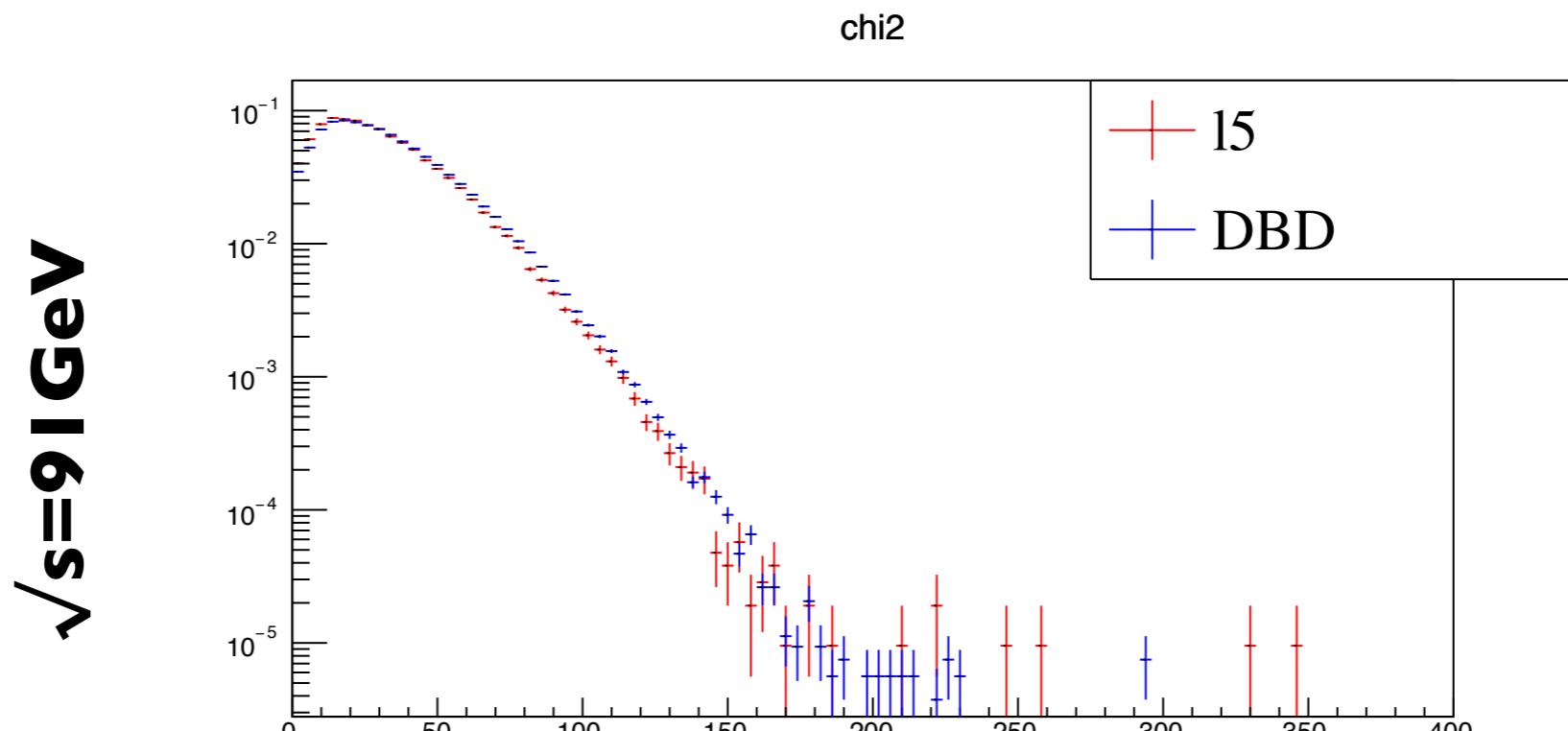
# Primary vertex finding (track energy sum)



# Primary vertex finding (track PID)

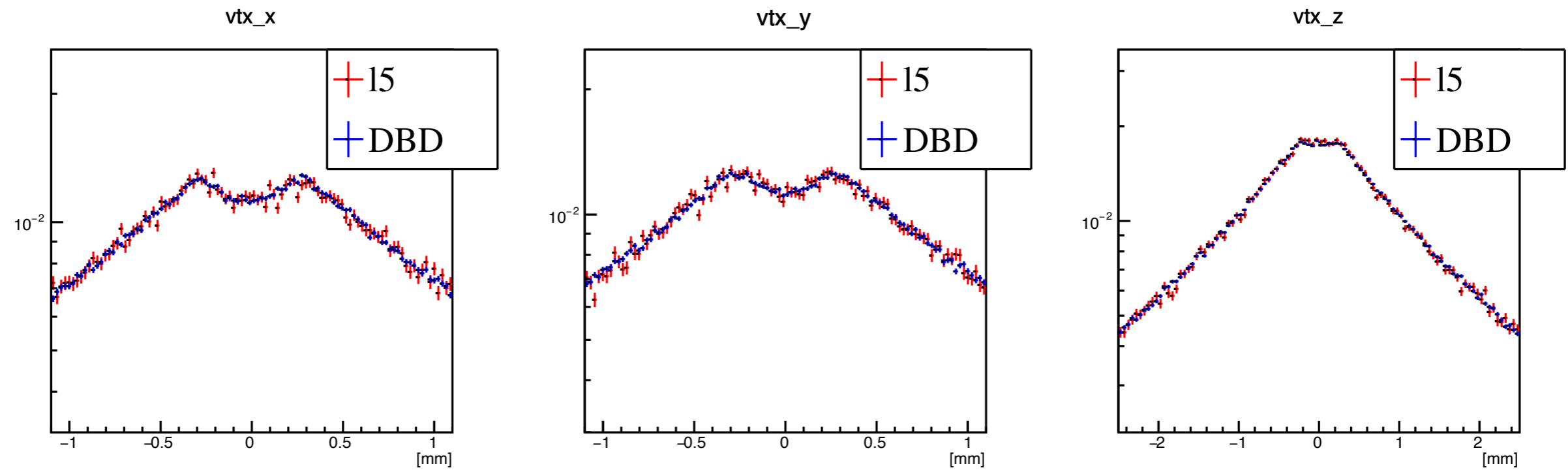


# Primary vertex finding (chi2)

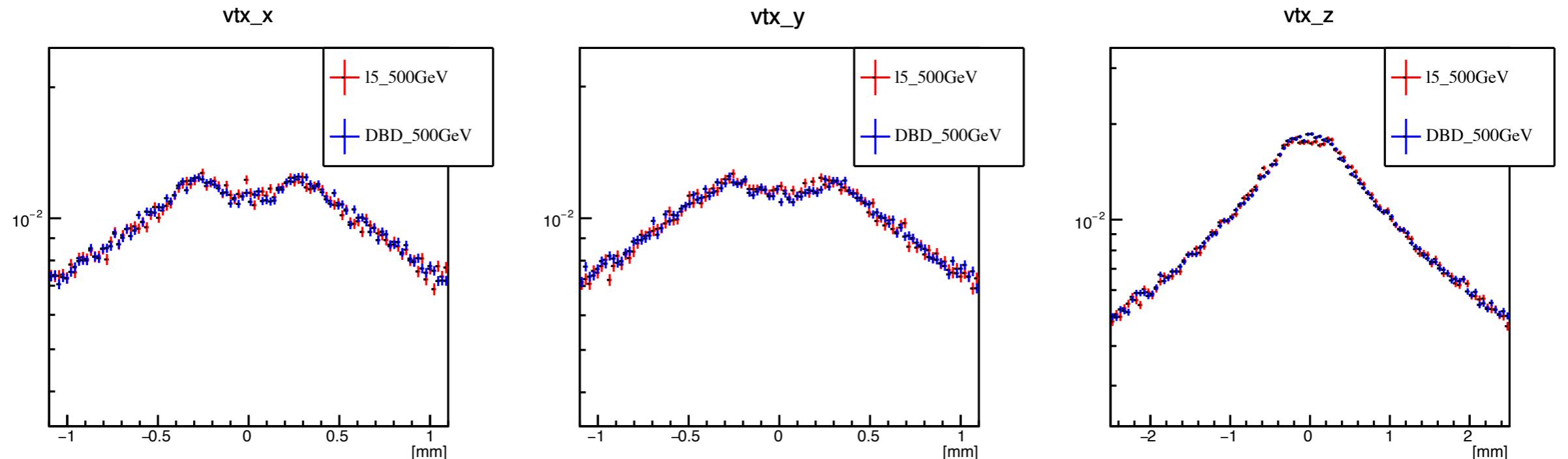


# Secondary vertex finding (Position)

$\sqrt{s} = 9 \text{ GeV}$

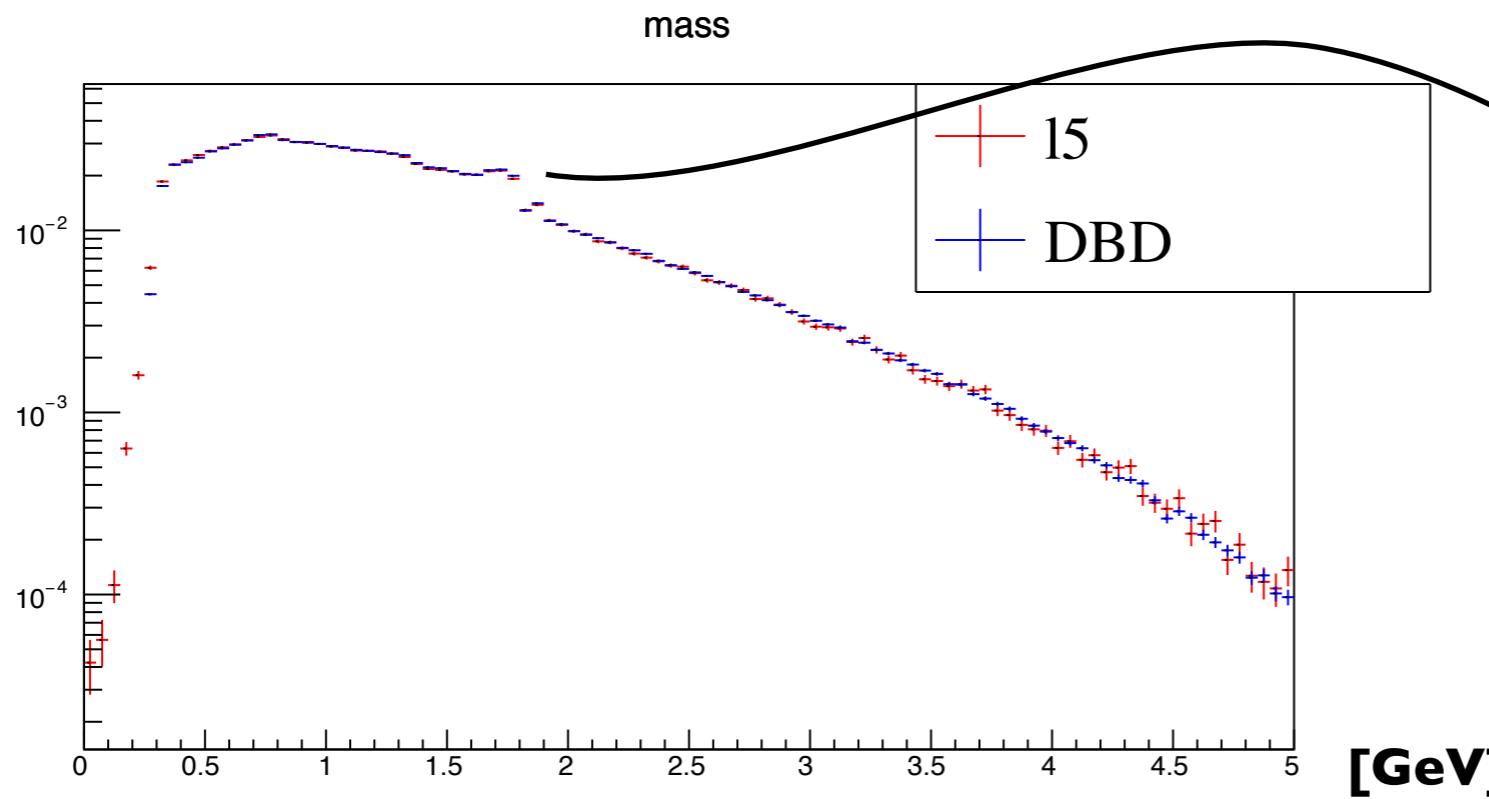


$\sqrt{s} = 500 \text{ GeV}$

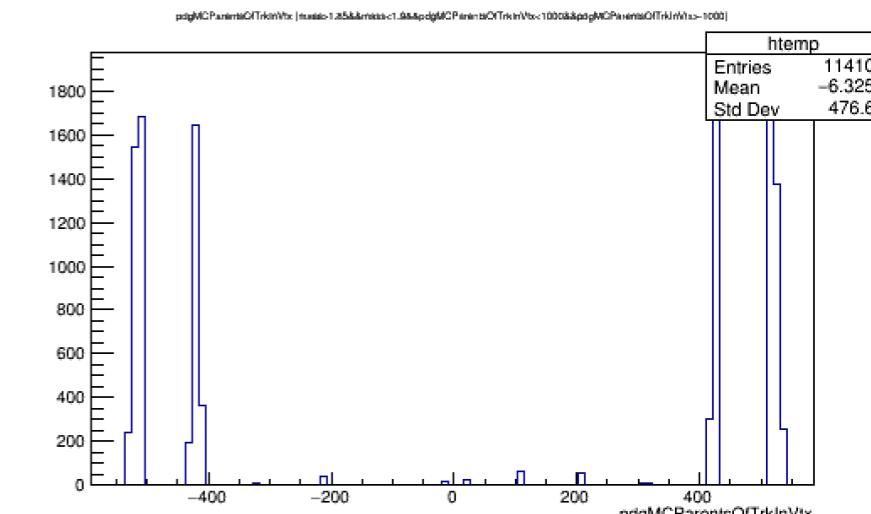


# Secondary vertex finding (vertex mass)

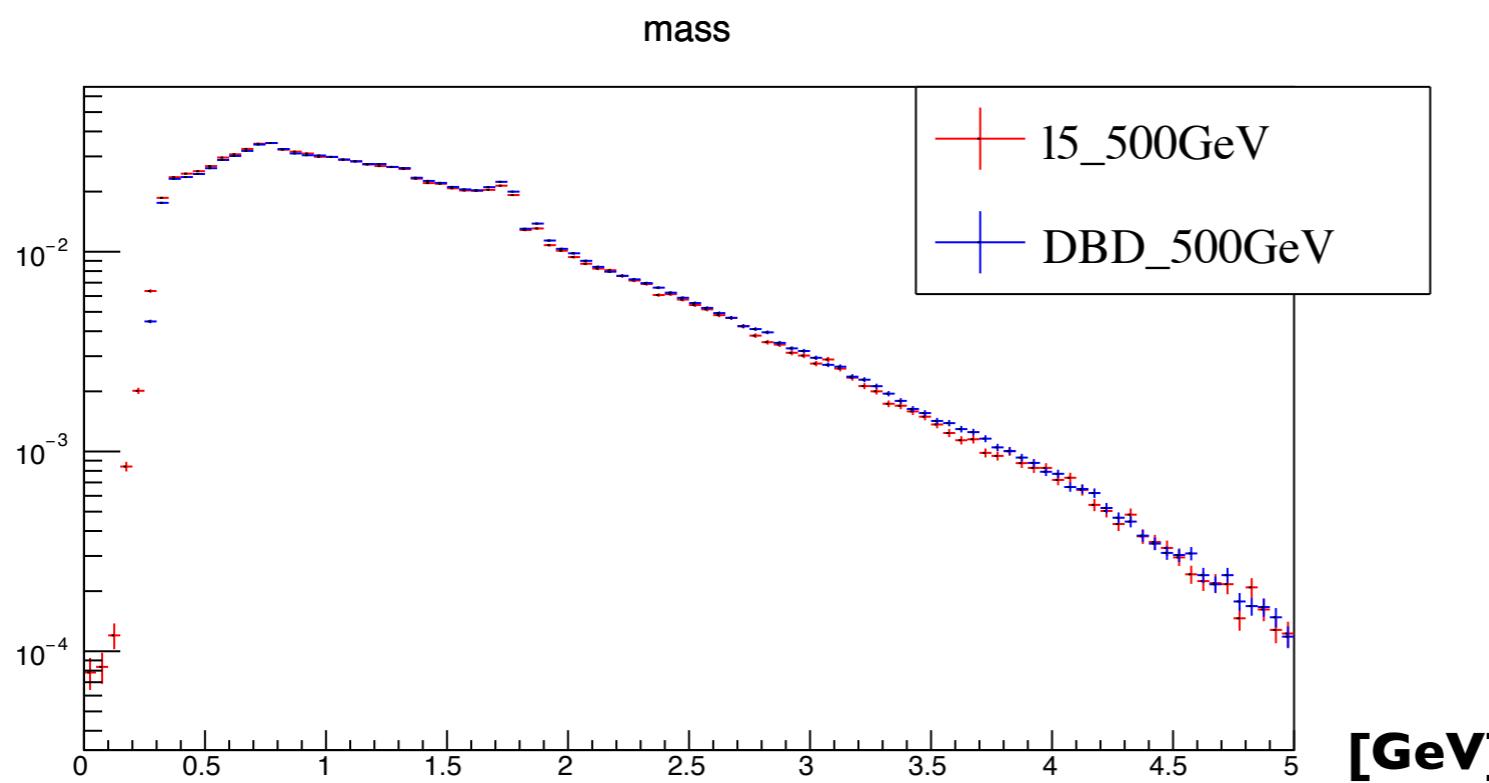
$\sqrt{s=9 \text{ GeV}}$



**track parent PID**

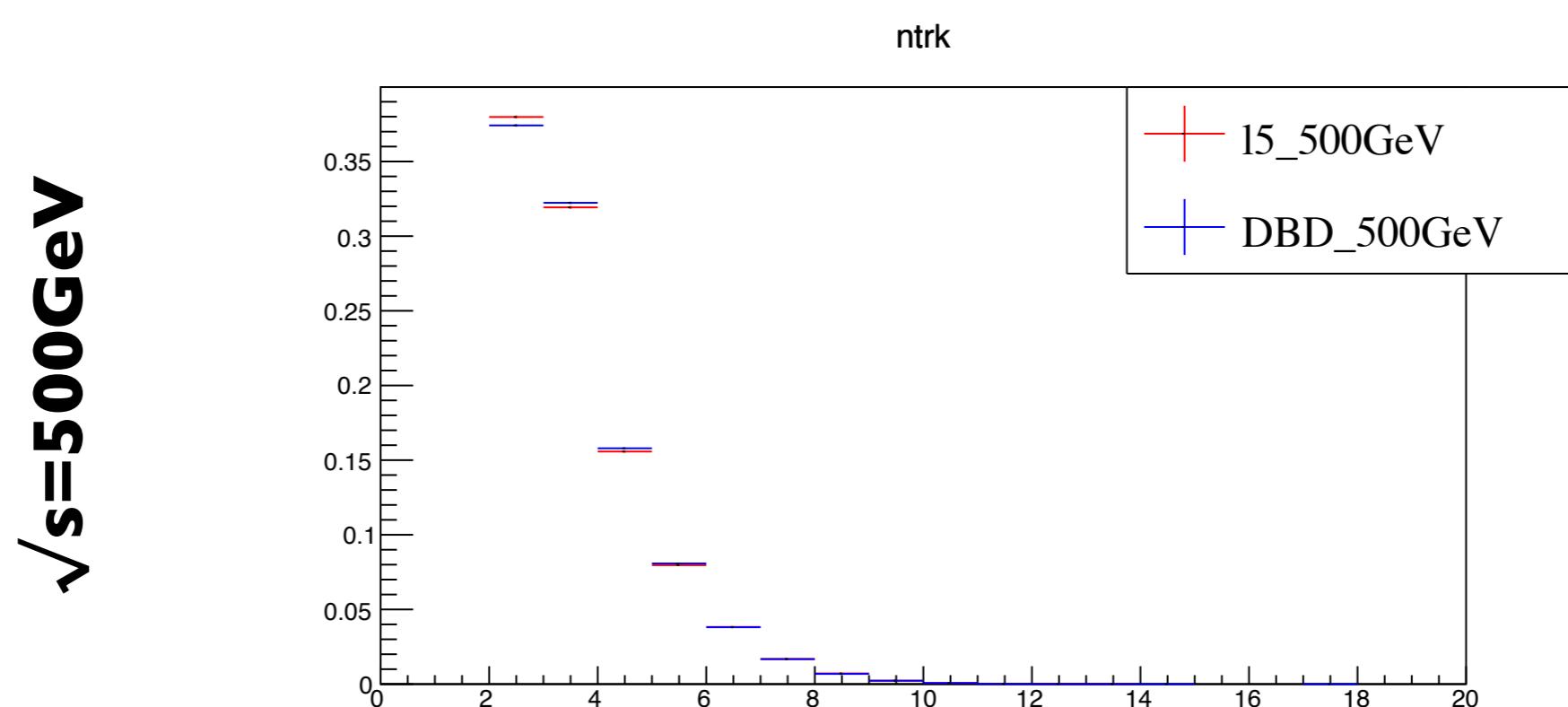
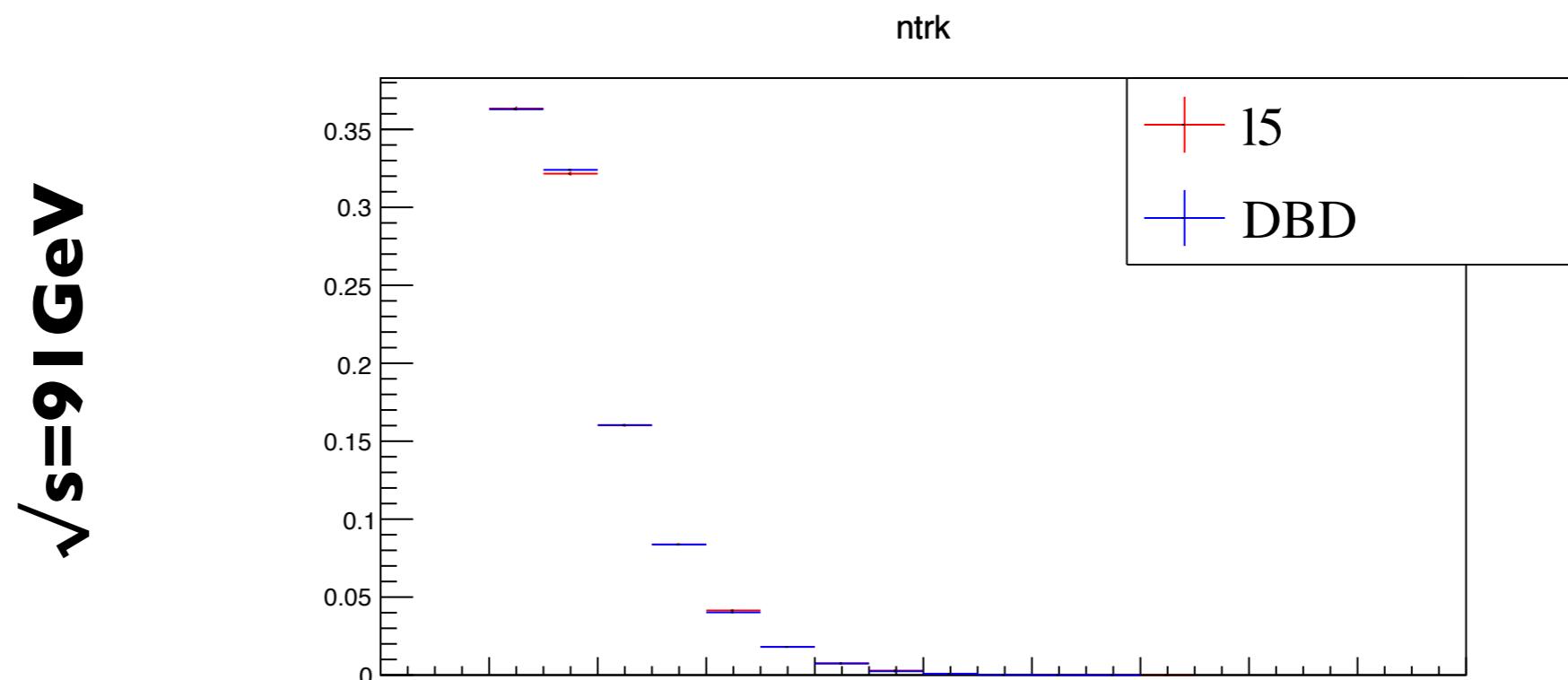


mass

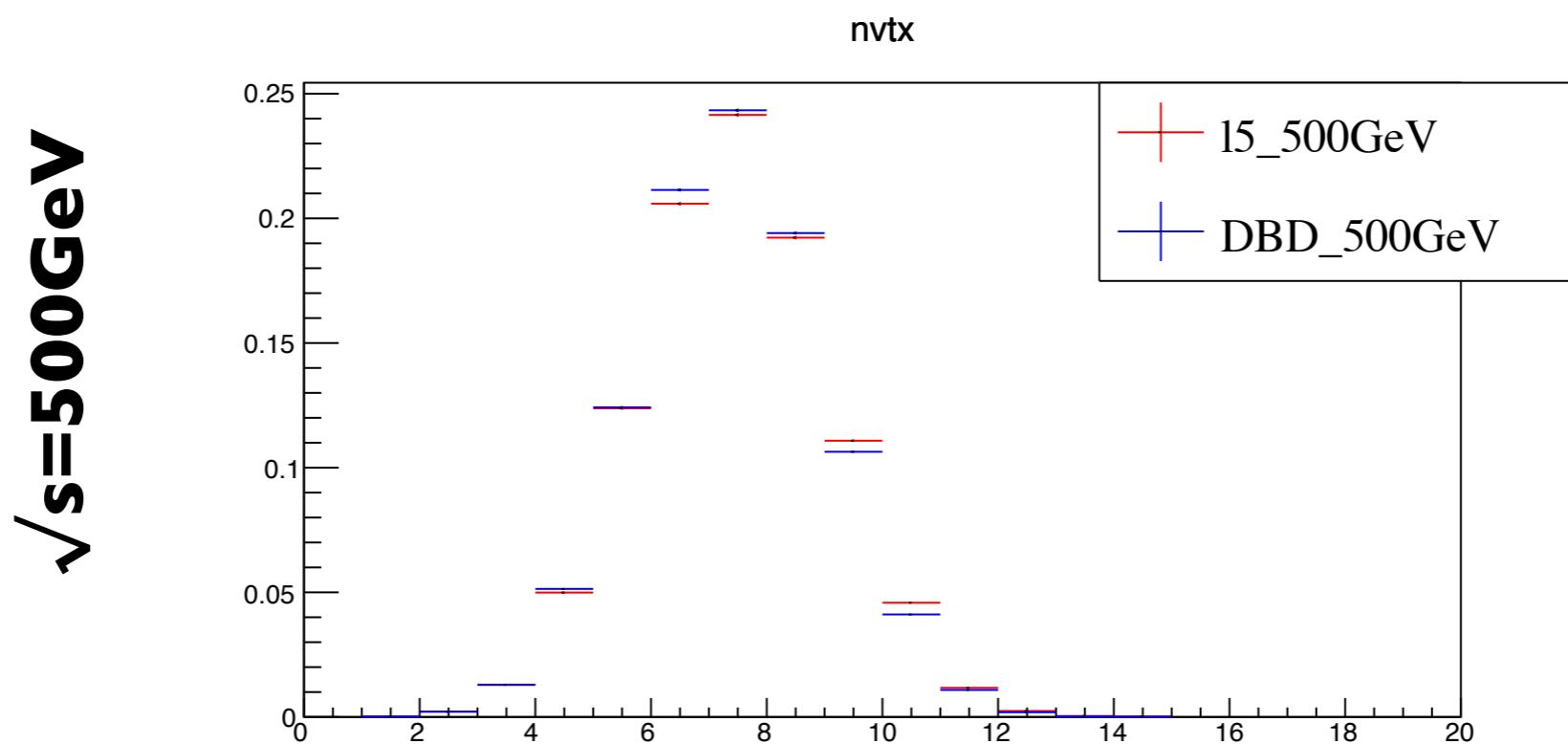
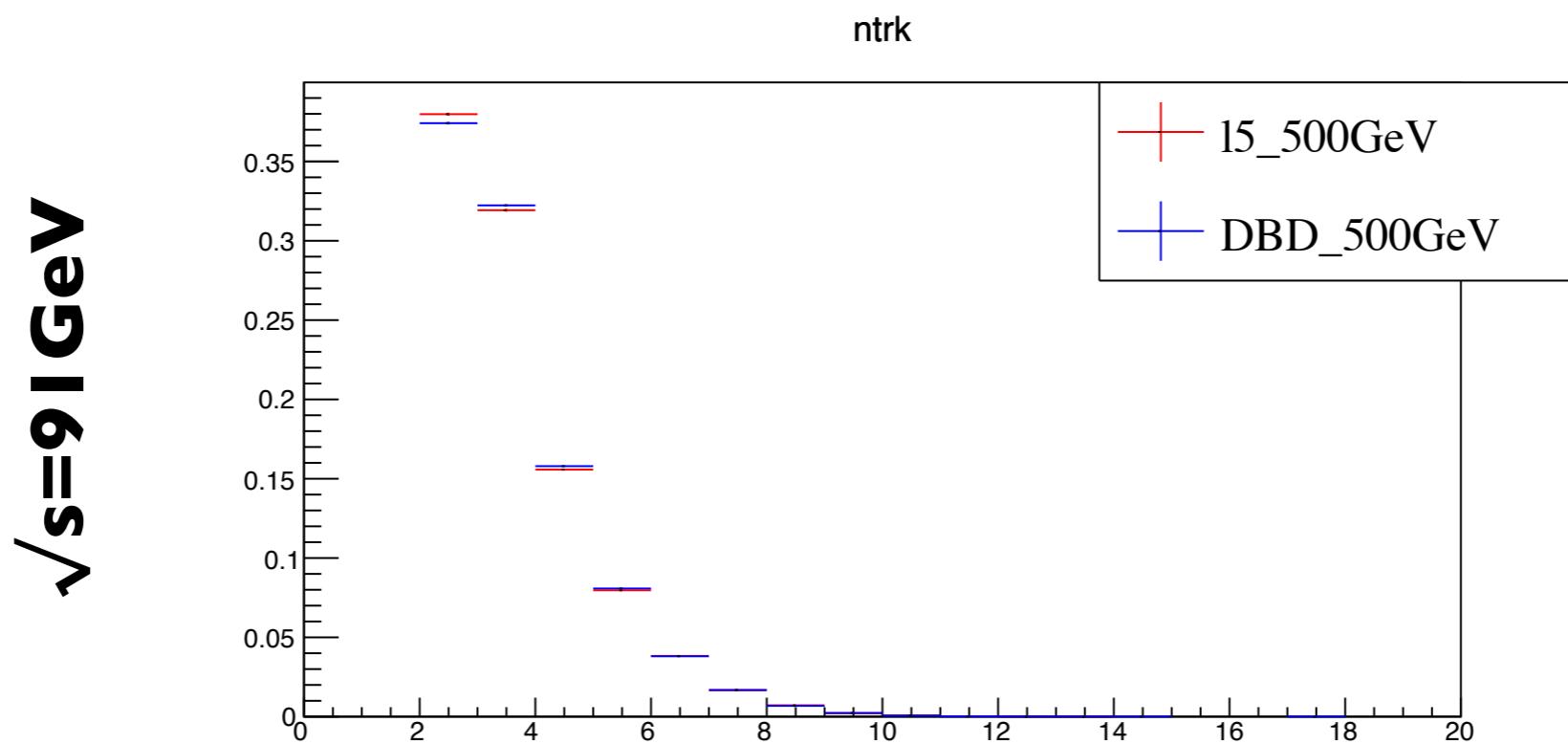


**Most of the entries correspond to B,C-meson**

# Secondary vertex finding (# of vertex tracks)

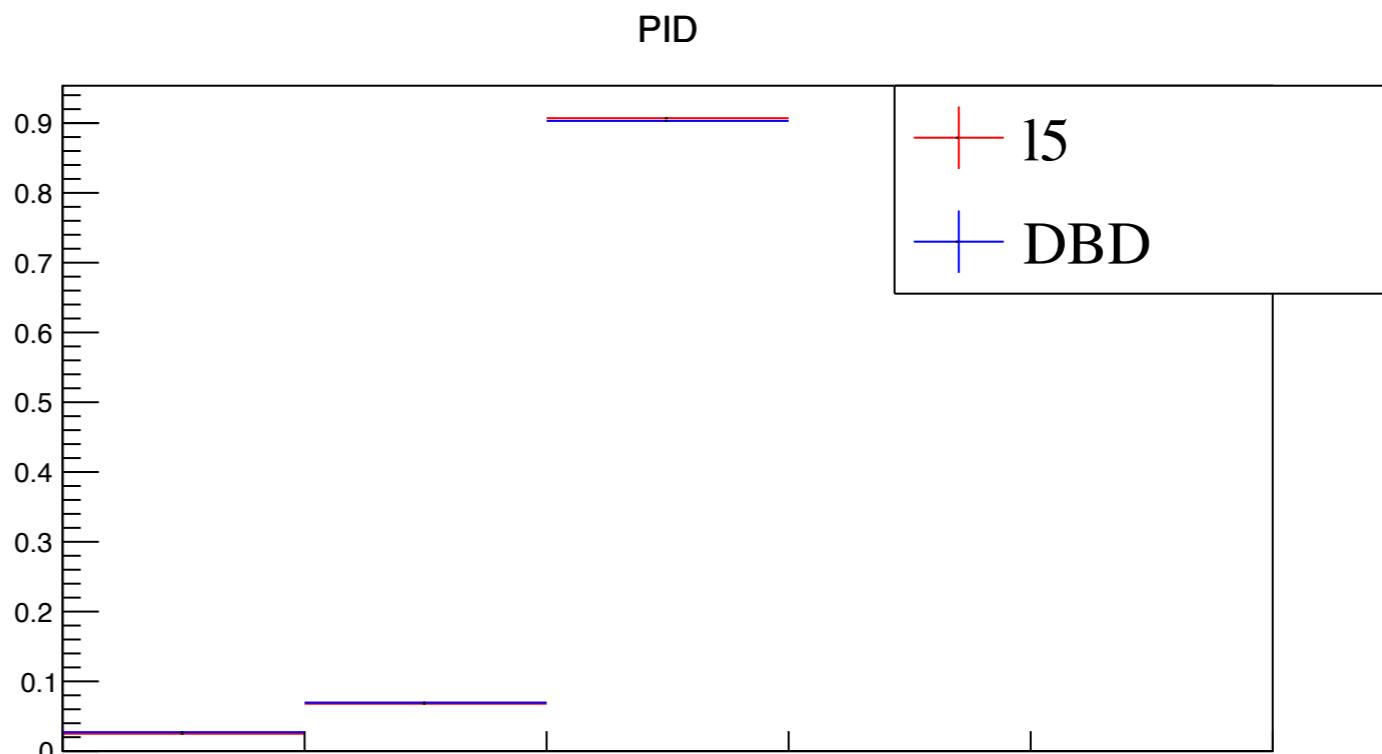


# Secondary vertex finding (# of vertex)

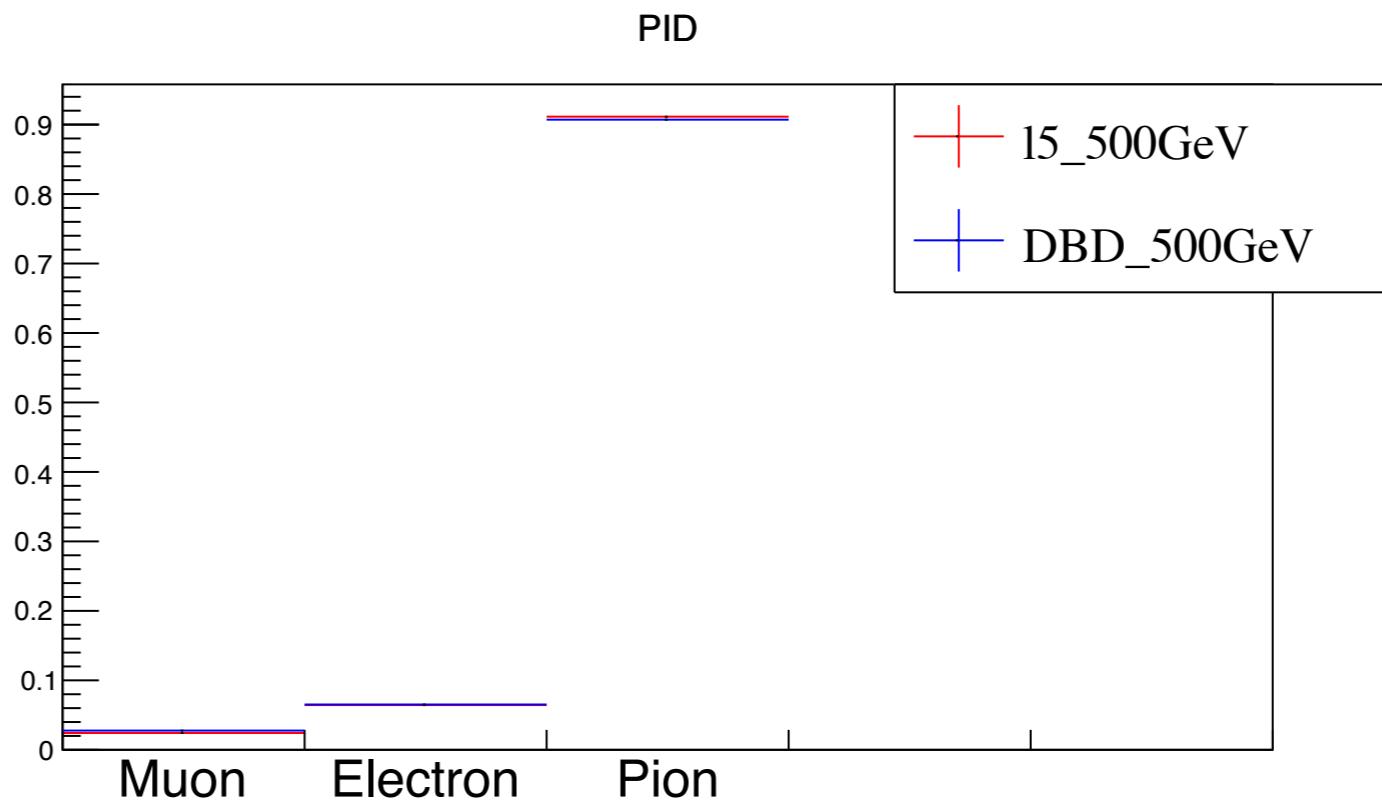


# Secondary vertex finding (track PID)

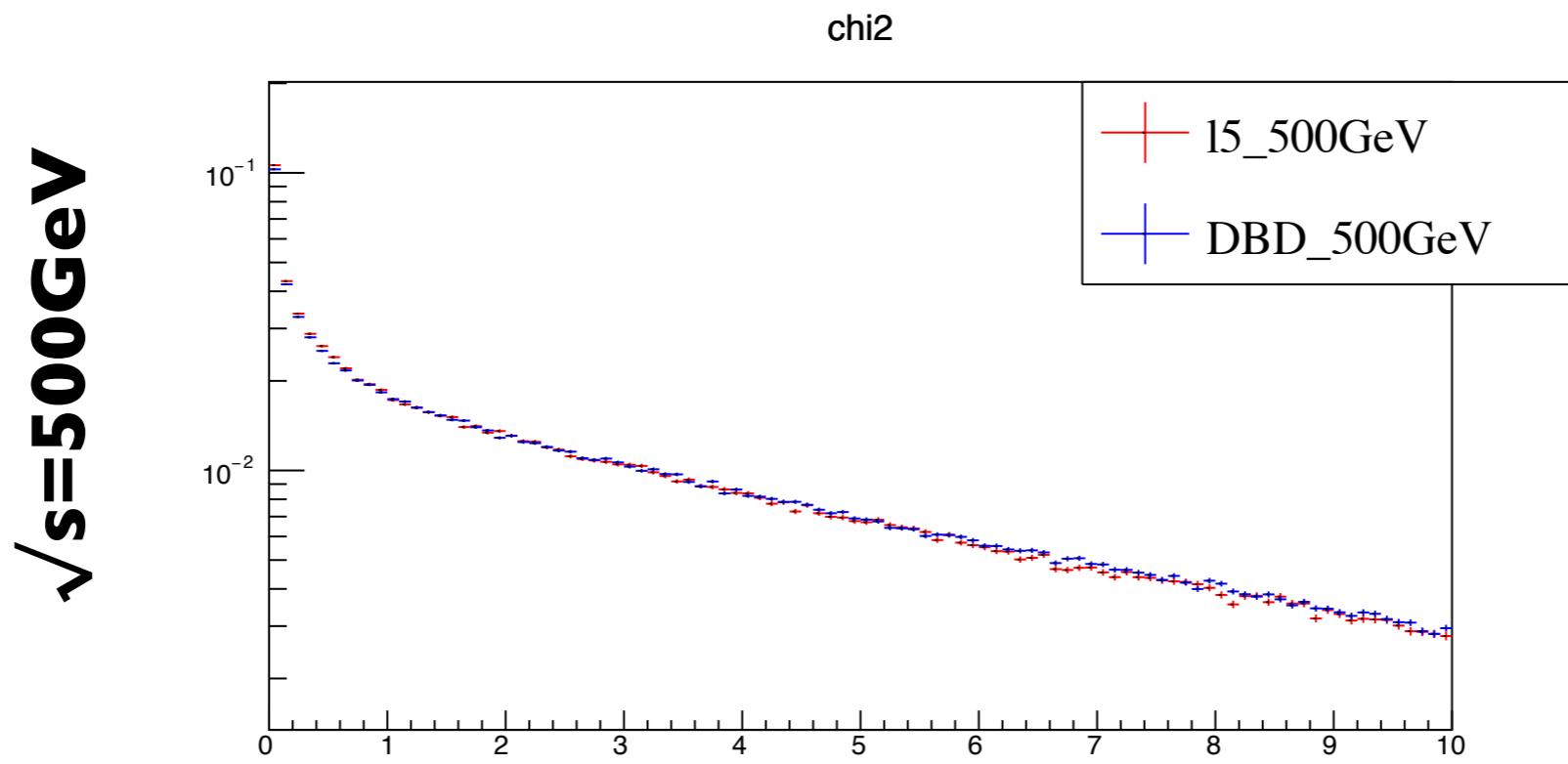
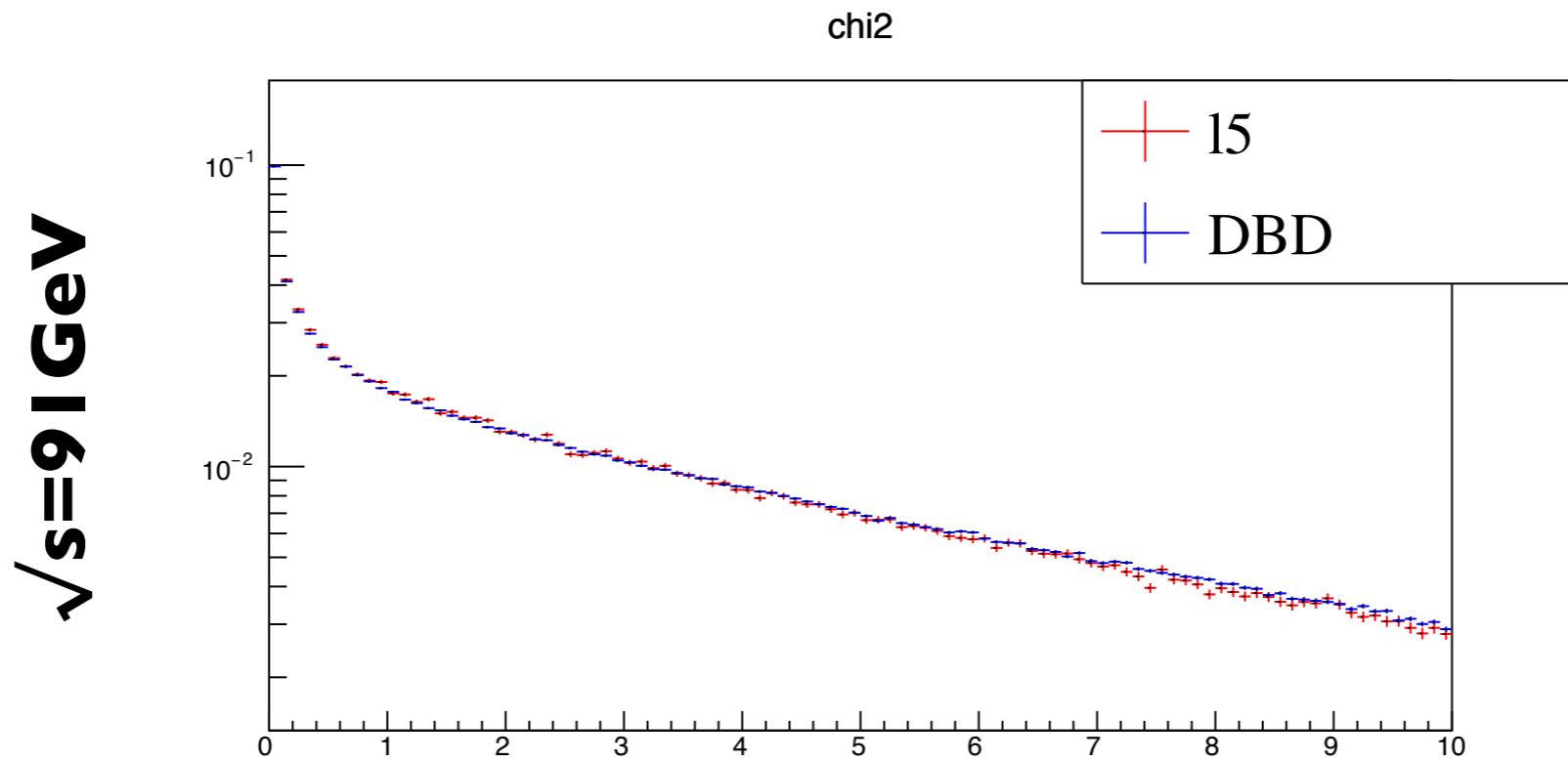
$\sqrt{s}=91\text{ GeV}$



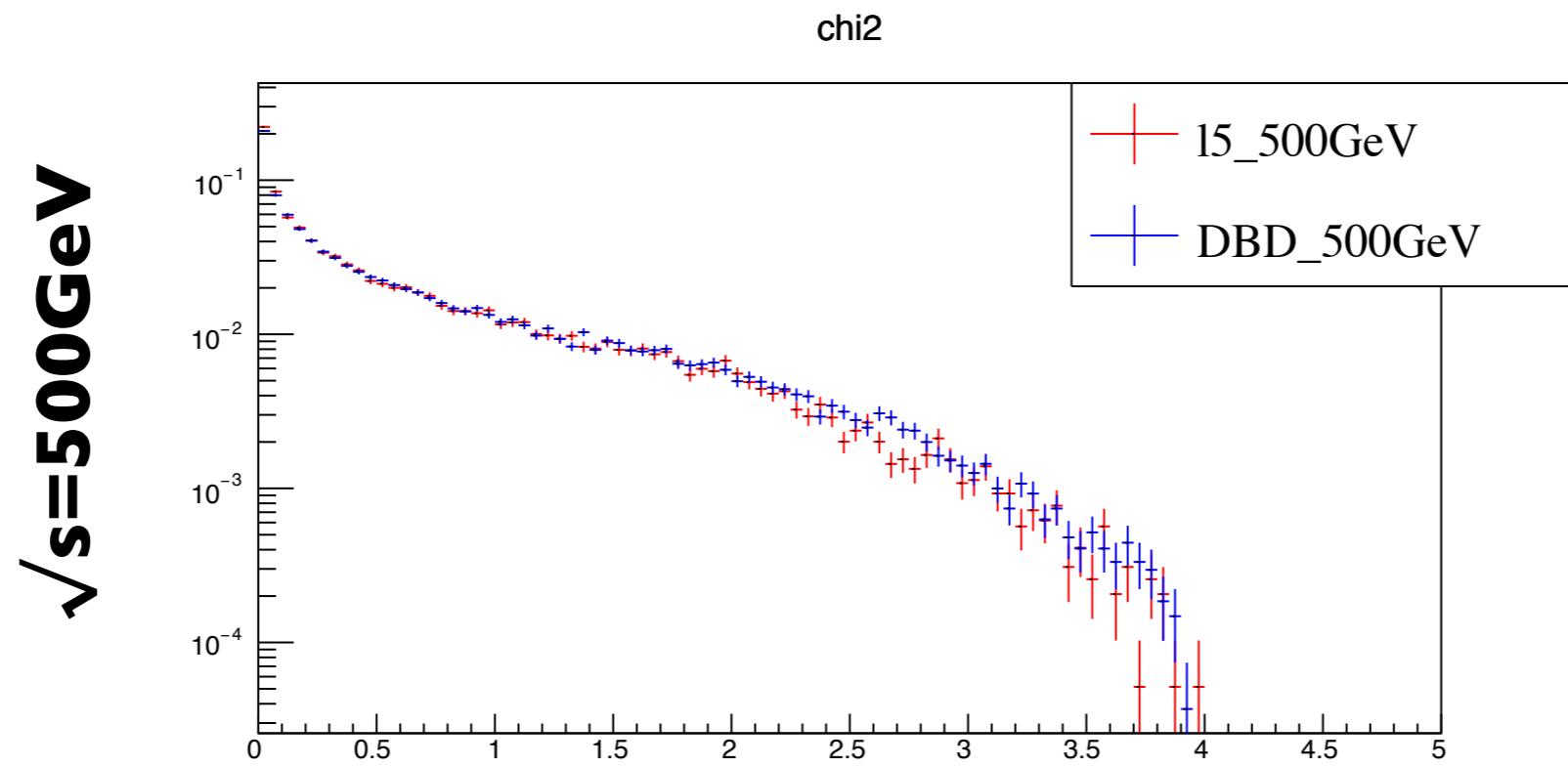
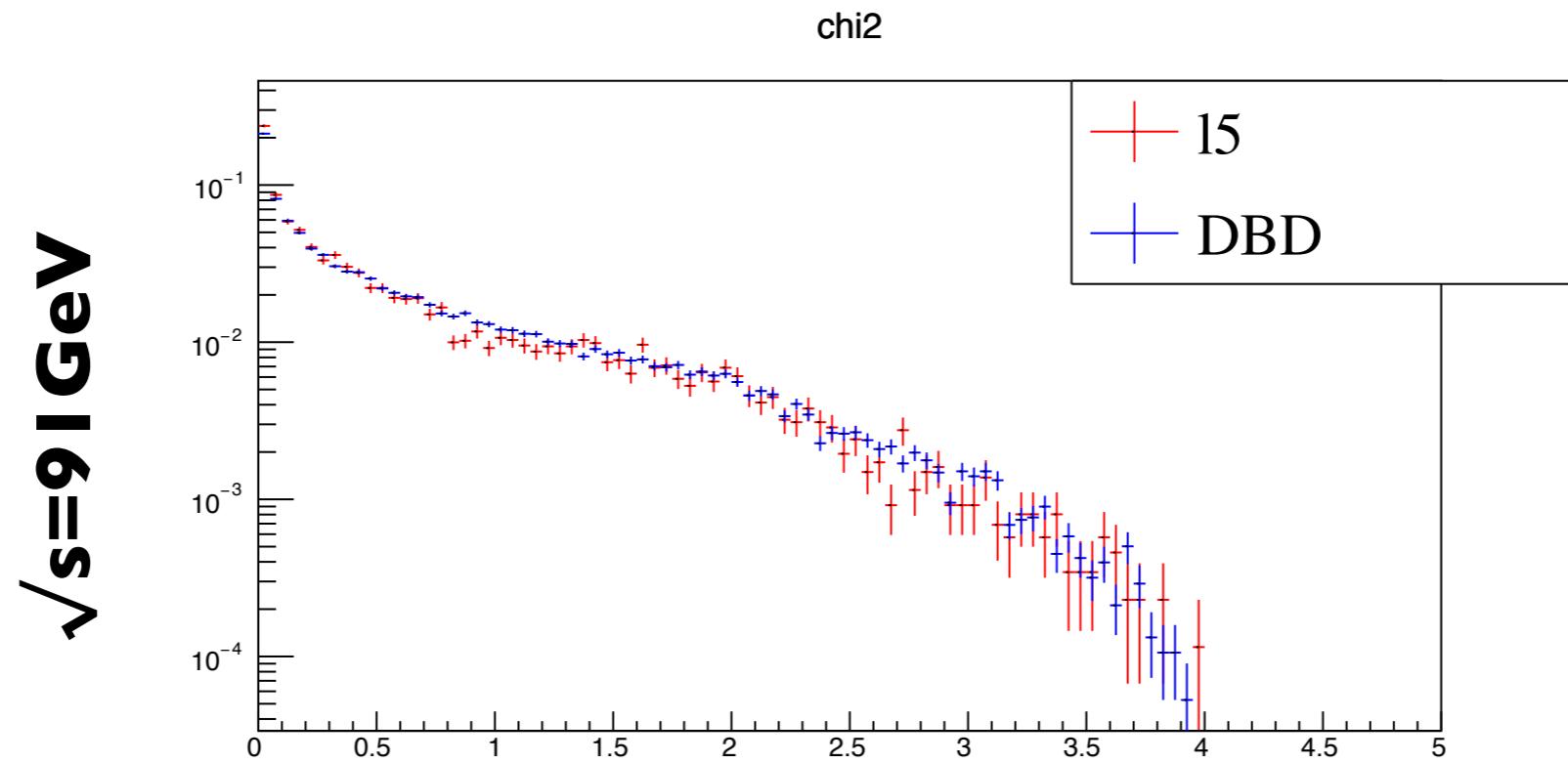
$\sqrt{s}=500\text{ GeV}$



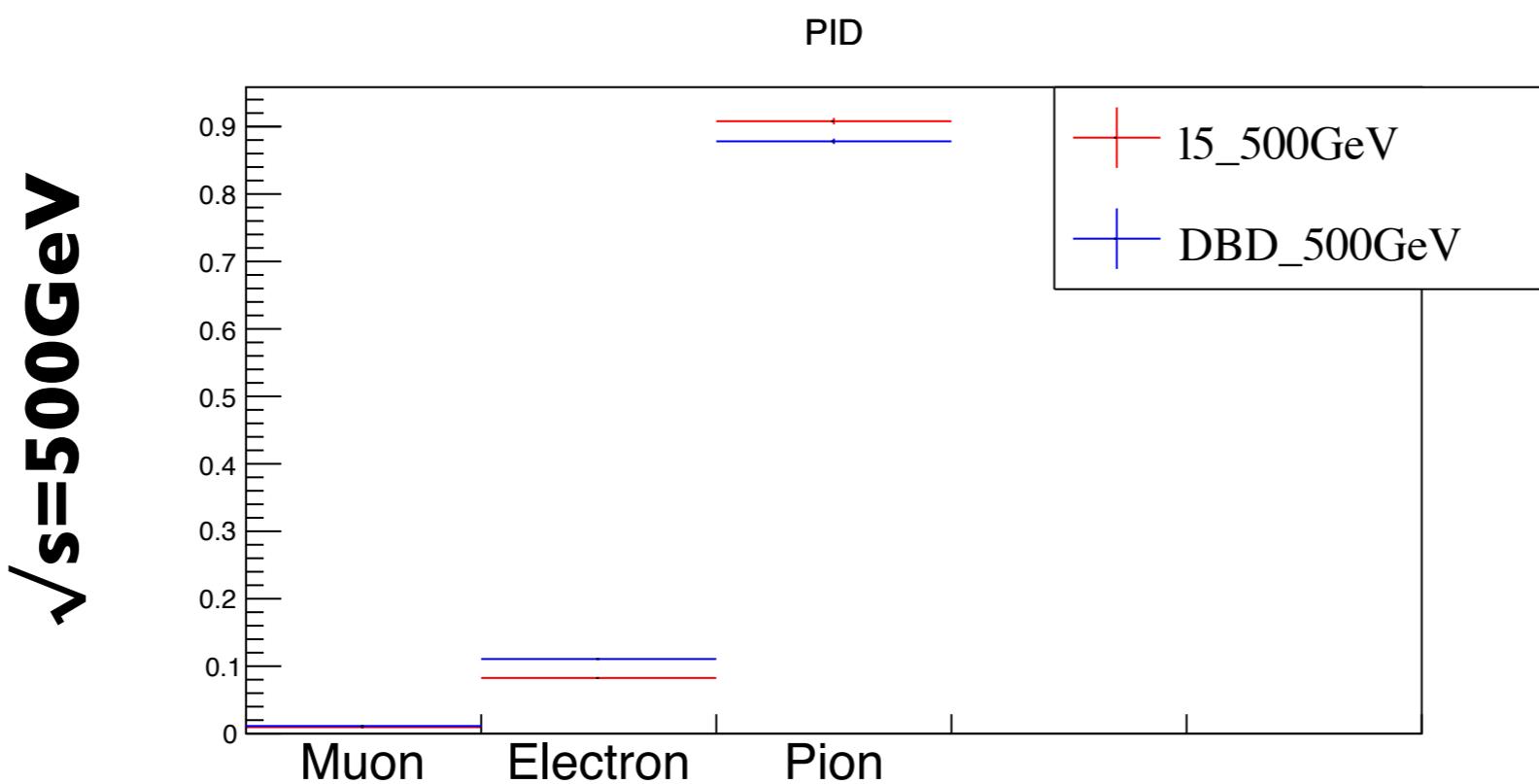
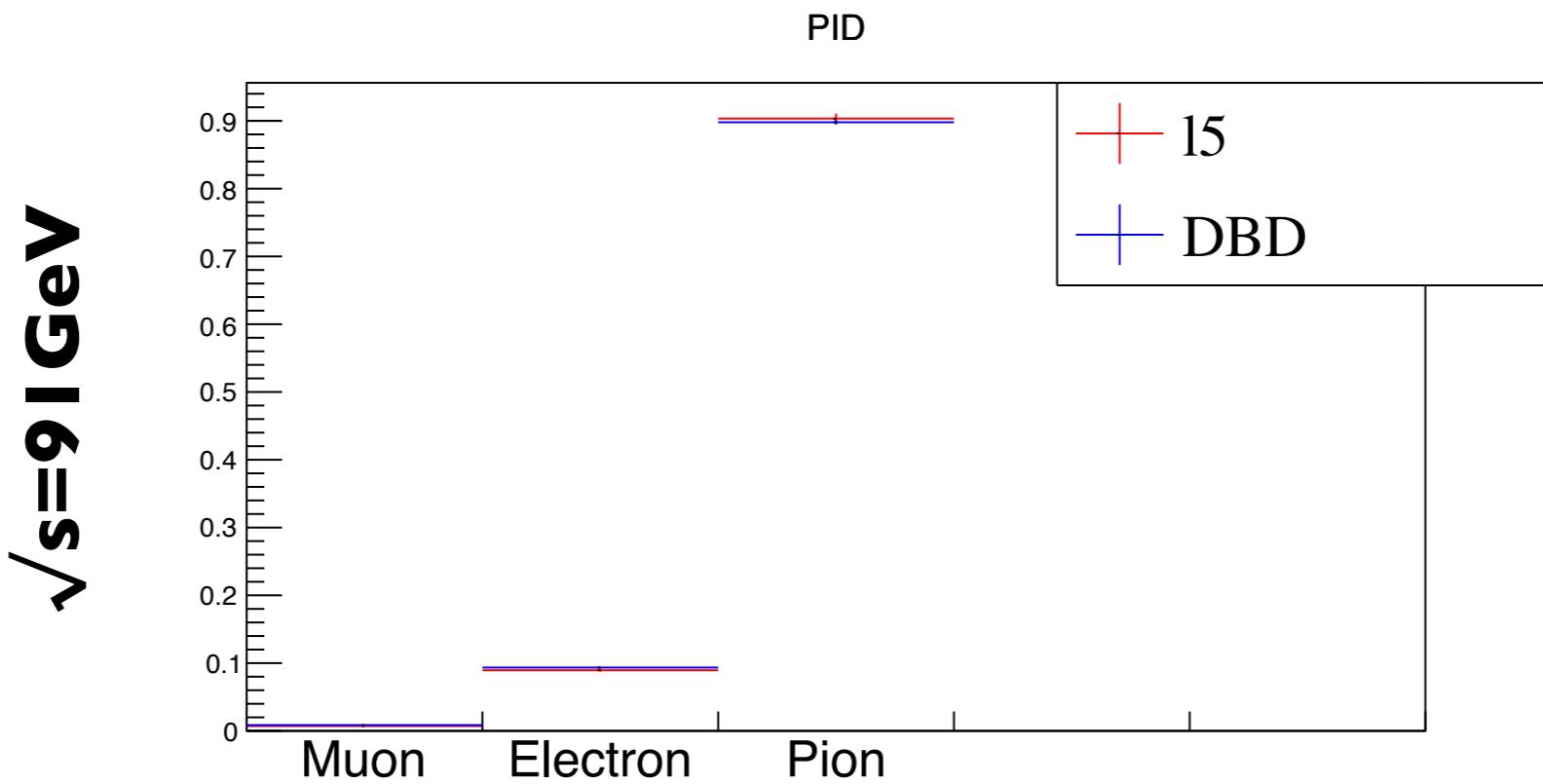
# Secondary vertex finding (chi2)



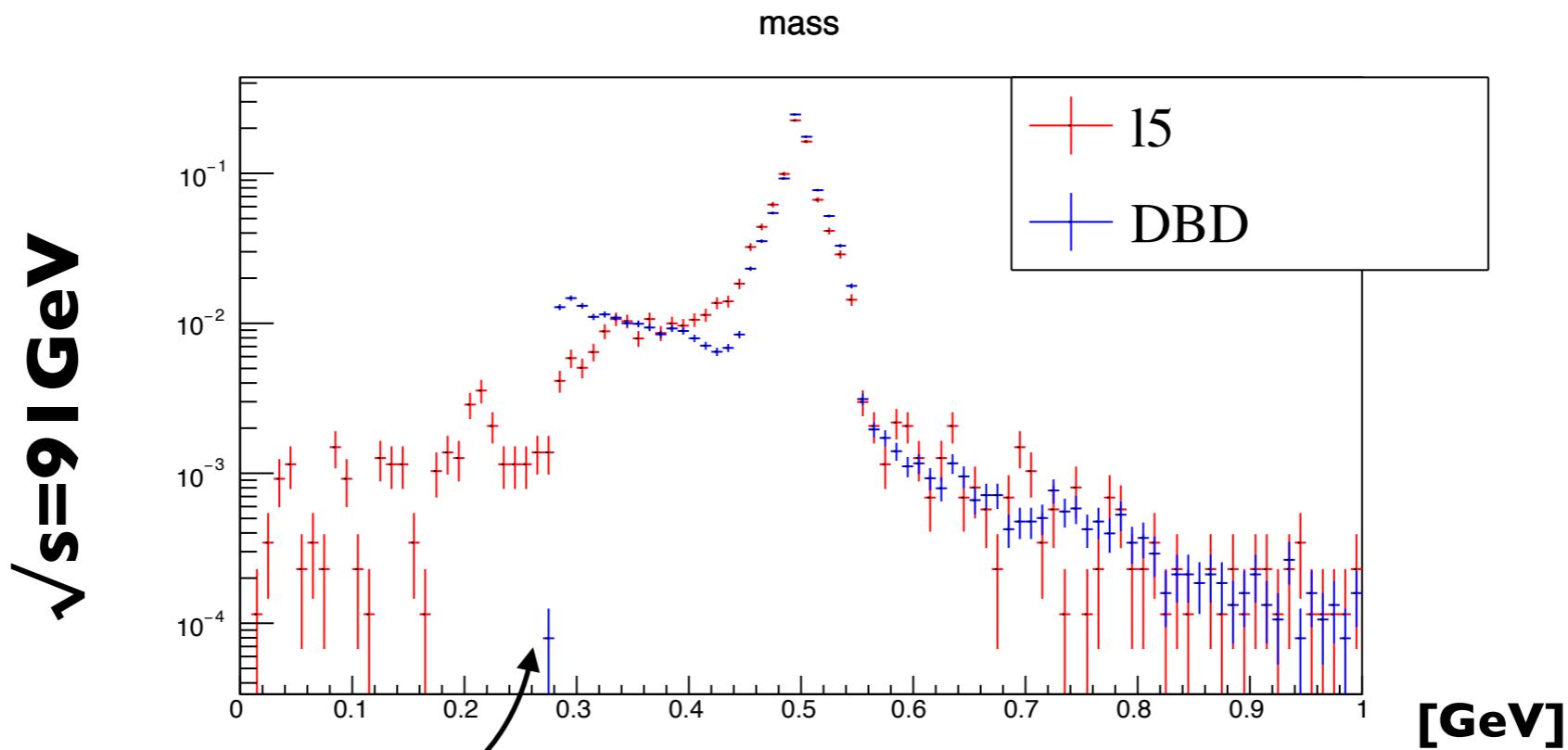
# V0 vertex finding (chi2)



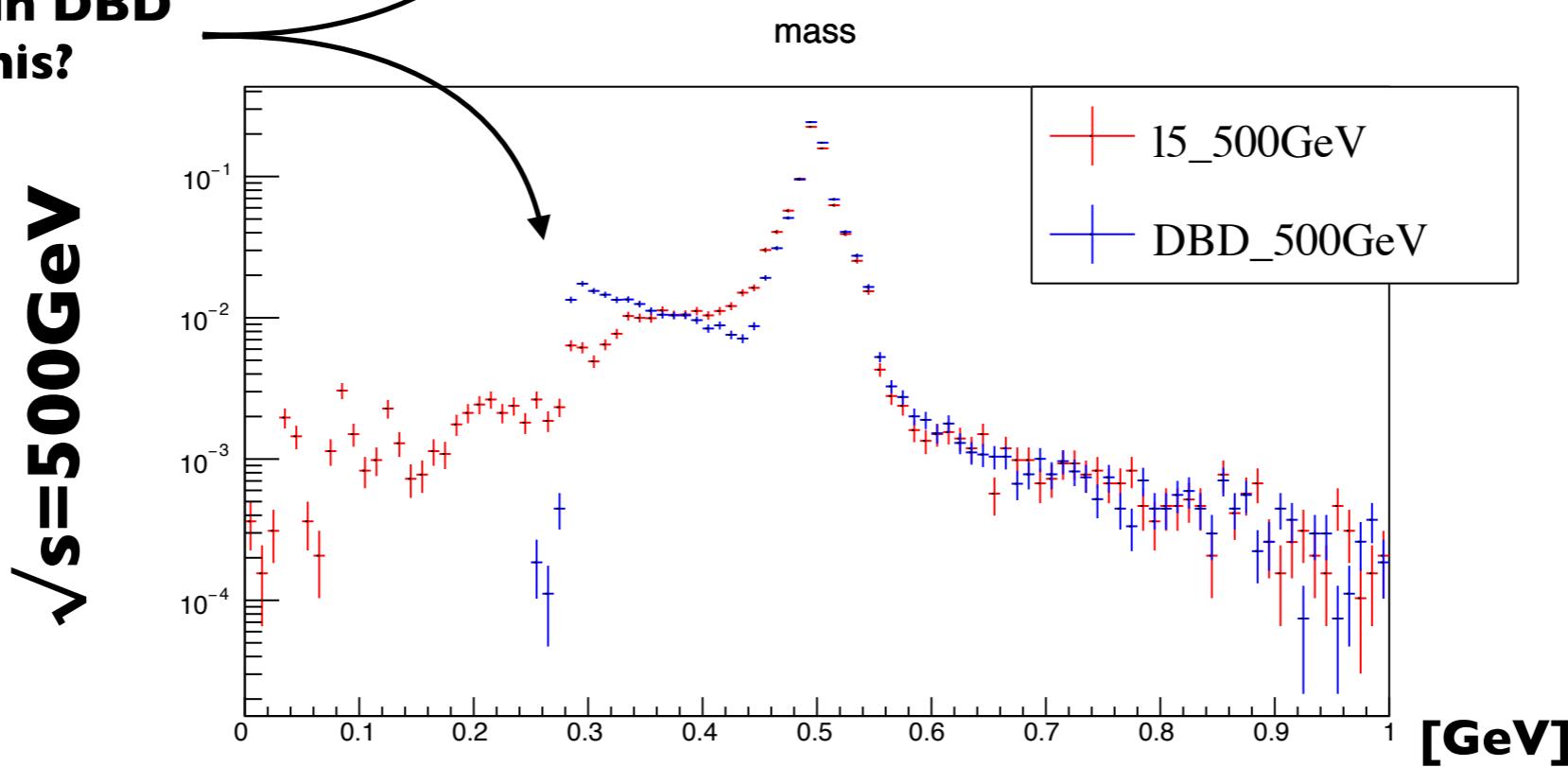
# V0 vertex finding (track PID)



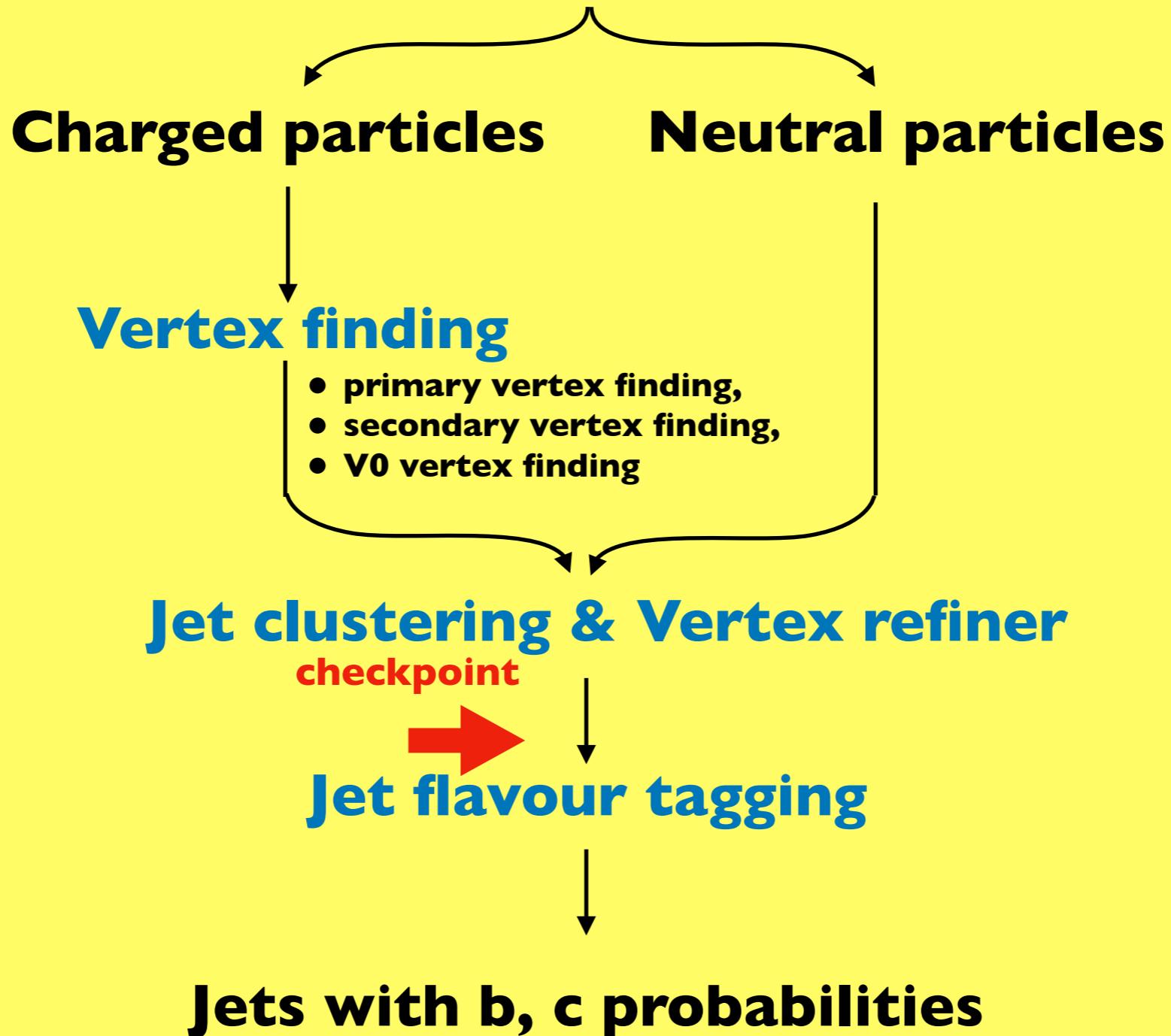
# V0 vertex finding (mass)



Rho mass cut in DBD  
explains this?

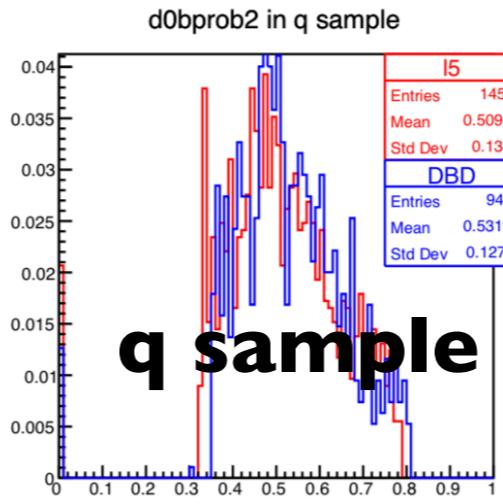
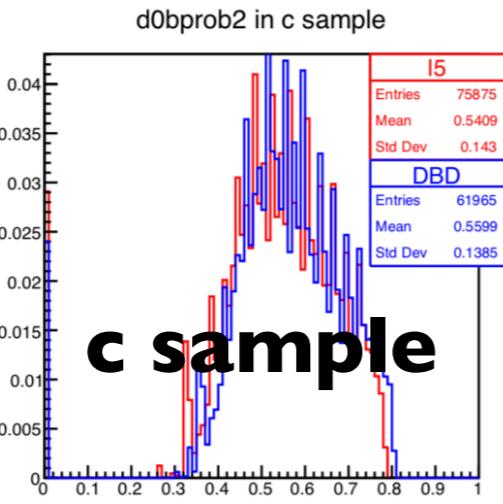
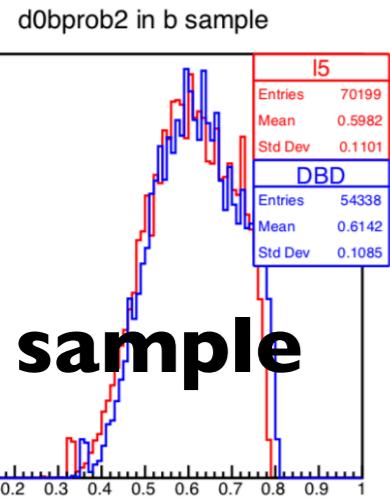


# Particle Flow Object (PFO)



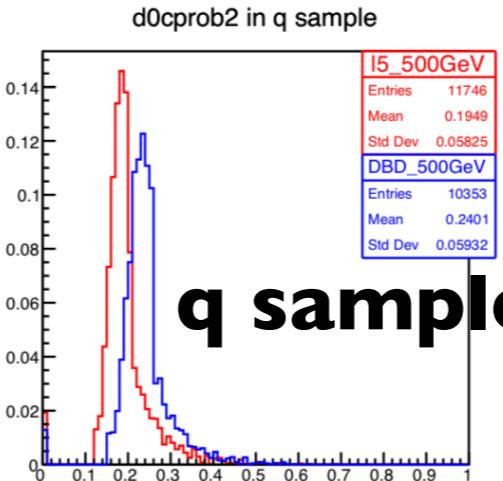
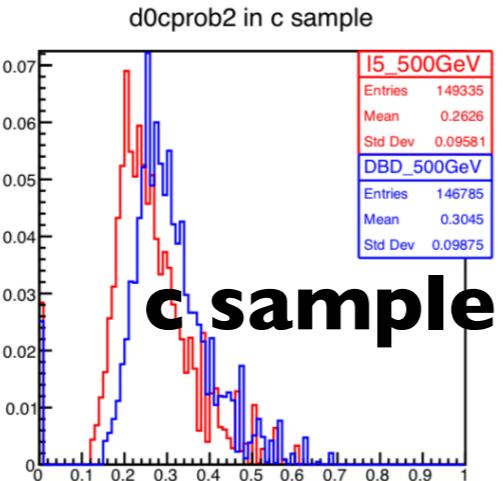
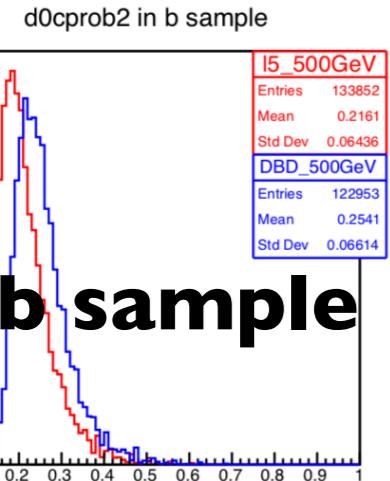
# Inputs for flavour tagging with TMVA

- ❖ **Categorize jets depends on their signature and do multi-variate analysis for each category.**
  - ▶ # of secondary vertices  $\geq 2$  (likely b-jet)
  - ▶ # of secondary vertices == 1 & # of pseudo-vertices == 1 (likely b-jet)
  - ▶ # of secondary vertices == 1 (likely c-jet, maybe b-jet)
  - ▶ # of secondary vertices == 0 (likely light flavour jet)
  - ▶ Input variable examples : d0, z0, track pt, vertex mass, # of vertex tracks
  - ▶ # of input variables  $> 30$  in total.



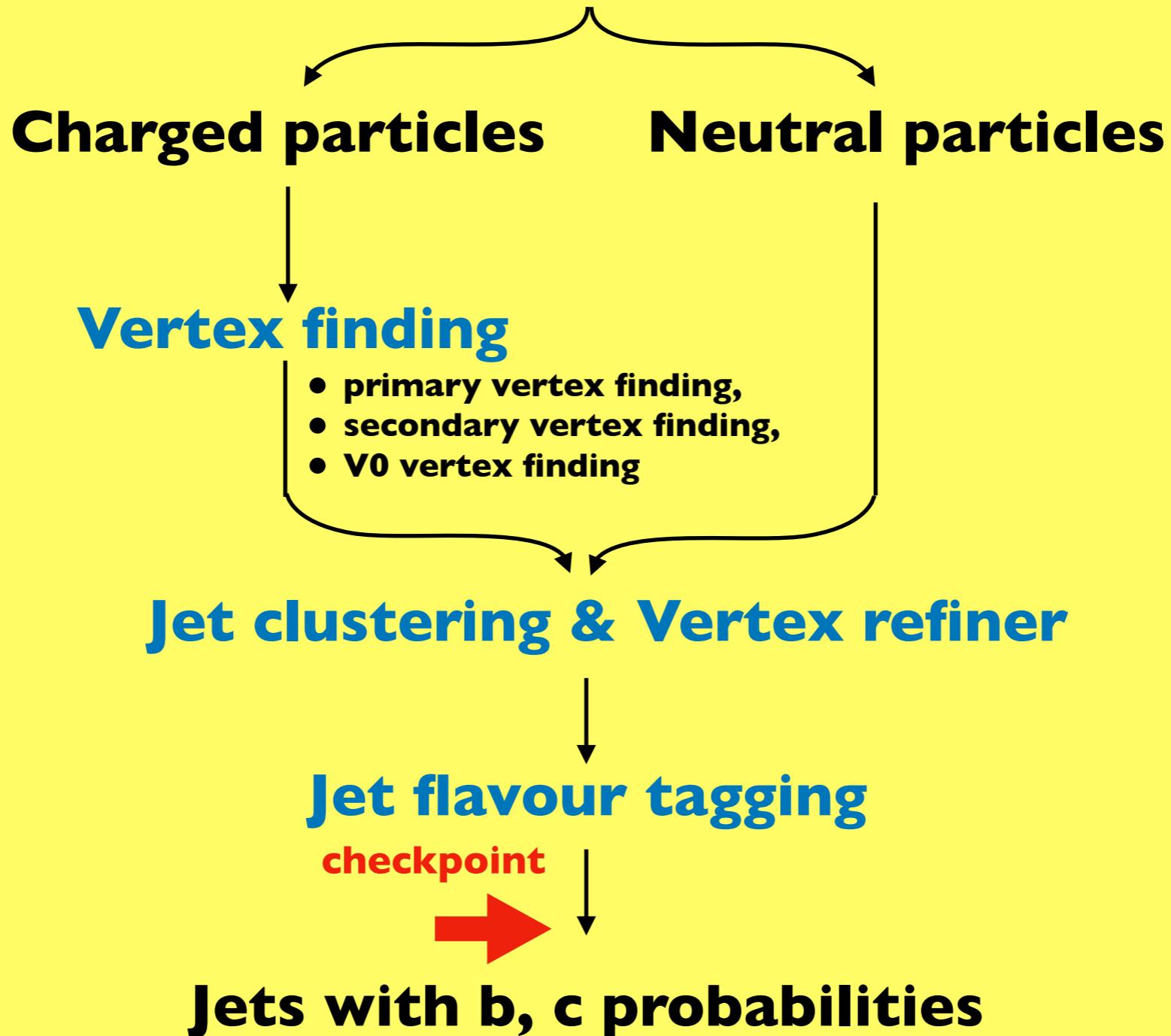
**91 GeV**

An example showing  
clear differences in 500 GeV sample,  
but not in 91 GeV sample.



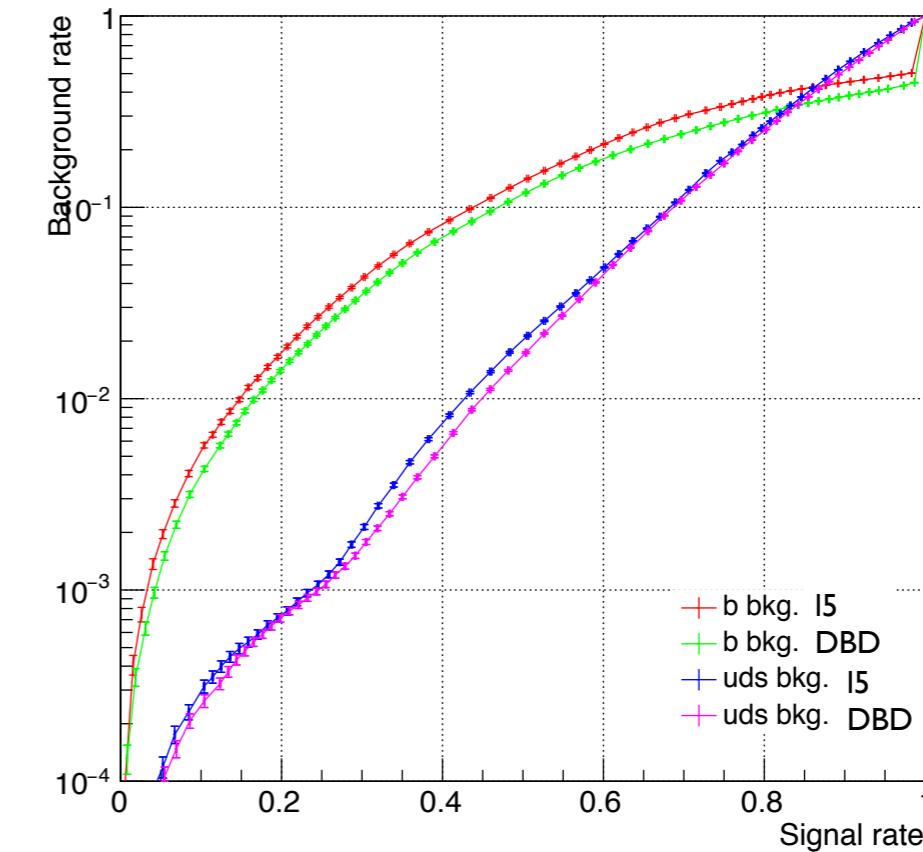
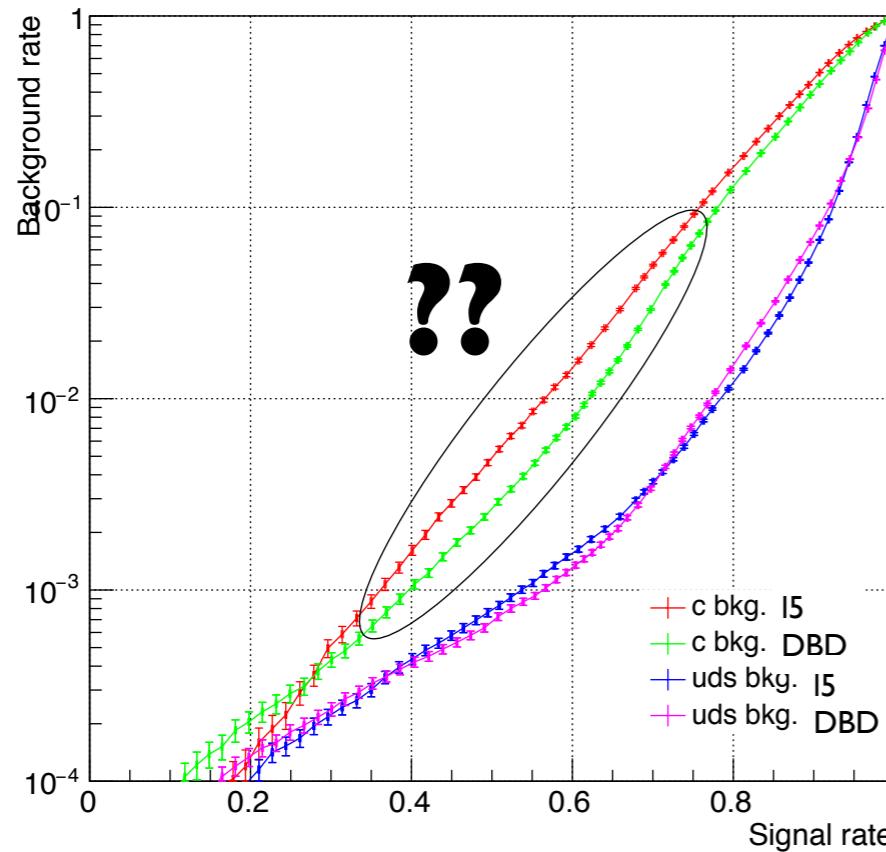
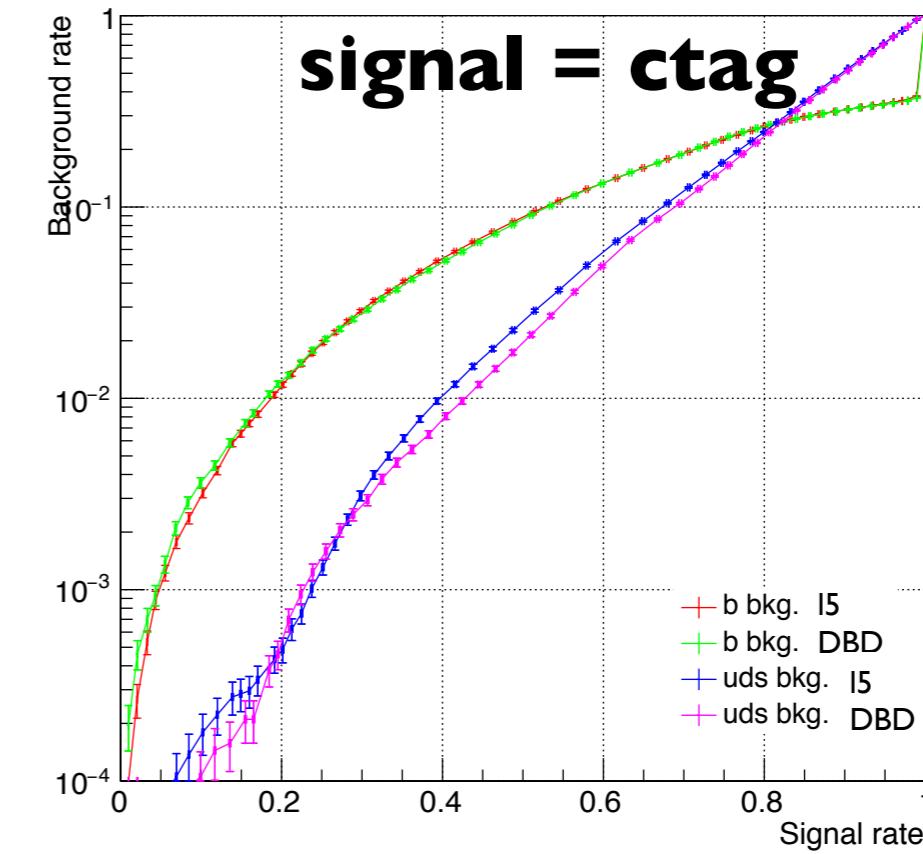
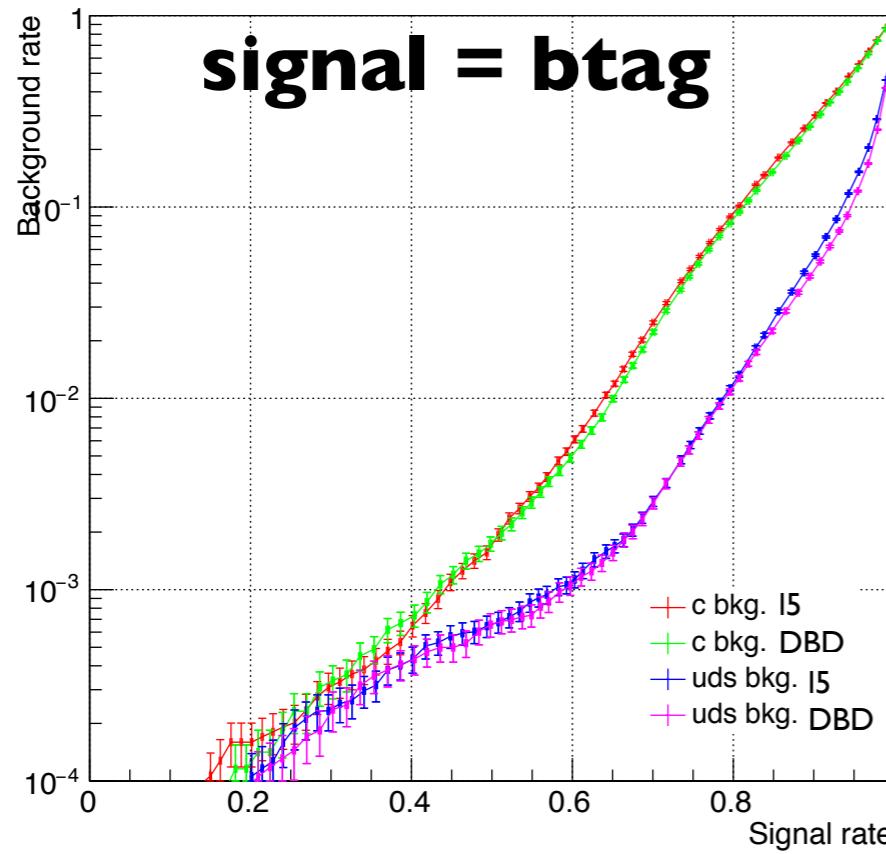
**500GeV**

# Particle Flow Object (PFO)



# Flavour tagging (Bkg. rate)

$\sqrt{s} = 9 \text{ GeV}$



# Summary

- ❖ **Flavour tag performance is being checked.**
  - ▶ comparison with DBD
  - ▶ Some difference in 500GeV sample
    - ▶ Some possible causes found in input distribution for flavour tagging.
    - ▶ The other possibilities :
      - ▶ flavour tagging parameters were optimized for DBD samples.
        - ▶ Parameter optimization for new samples is our next step.
      - ▶ multi-jet environment
      - ▶ IP smearing ( $\sigma z \sim 200\text{um}$ ). It was point-like at DBD era.
- ❖ **It would be helpful to disentangle the possible causes if we could have a small samples of 500GeV (6b, 6c, 6q) without IP smearing and beam bkg (= same condition as DBD).**
  - ▶ Could also be useful to optimize parameters of flavour tagging for new samples.

# **Backup**

# Technical details

## ❖ Sample used (at KEKCC)

► I5, 91GeV, no beam bkg.

/hsm/ilc/grid/storm/prod/ilc/mc-opt.dsk/ild/dst/91-nobeam/flavortag/ILD\_15\_o1\_v02\_nobg/v01-19-05-p01/00009199/000/

► DBD, 91GeV, no beam bkg.

/gpfs/group/ilc/suehara/reco/dbd/train-lcfiplus/bb91new,  
/gpfs/group/ilc/suehara/reco/dbd/train-lcfiplus/cc91new,  
/gpfs/group/ilc/suehara/reco/dbd/train-lcfiplus/qq91new

► I5, 500GeV, no beam bkg., IP smearing in simulation

/hsm/ilc/grid/storm/prod/ilc/mc-opt.dsk/ild/dst/500-TDR\_ws/flavortag/ILD\_15\_o1\_v02\_nobg/v01-19-06-p01

► DBD, 500GeV, no beam bkg.

/group/ilc/suehara/reco/dbd/train-lcfiplus/6q-jp/dst\_E500

## ❖ ilcsoft v01-19-06 used

## ❖ Beam size [mm]

```
<parameter name="BeamSizeX" type="float" value="303.e-6" />
<parameter name="BeamSizeY" type="float" value="2.38e-6" />
<parameter name="BeamSizeZ" type="float" value="196.e-3" />
```