Heavy Flavour group working meeting.
ee→ bb

A. Irles, 14th January 2018

Analysis group meeting

Reminder: last progresses

- Btag optimization.
- New categories for the charge calculation methods (BcBc, KcKc, BcKc, KcBc, BcKc same jets) + differential p-q correction
  - selection efficiency x2
- Proper error propagation during the p-q method application is applied
- Proper definition of the angle → The b-quark polar angle is defined as a polar angle of the vector

$$\vec{p}_{b\bar{b}} = \vec{p}_b - \vec{p}_{\bar{b}},$$

- Full BKG study.
- Use of generalized kt algorithm instead of the Durham. Bkg reduced by a factor 1.5 – 2.
  - Allows for the study of event shapes (a.k.a. sphericity) and jet shapes (y23) variables. Potential improvement of S/B of factor 2 without much loss of S thought it is not applied.
- Processor for Kaon identification using dEdx:
  - For bb@500GeV we get similar values of efficiency and purity than the quoted by S.B. for 250@GeV although a bit smaller.
  - Not used for the DBD samples.

Today

- Revisiting the Vertex/Track recovery.

- Study the issue with the shape of the distribution that gives bad chi2 in the fit.
  - Corrected with a fudge factor from MC.
  - Origin? We thought that it was caused by a lack of efficiency in the barrel region (due to loss tracks, PFOs?) but it seems to be (also) a resolution effect.

- Results.

- **We had strong suspicions that we were not applying it well...**
  - Partially because of problems on the steering file (QQbarProcessor was using the jets without the recovered tracks)
  - Partially because of missus of the processor.
  - Steering file attached in the indico. Technical slides from S.B. attached too,

- A key point:

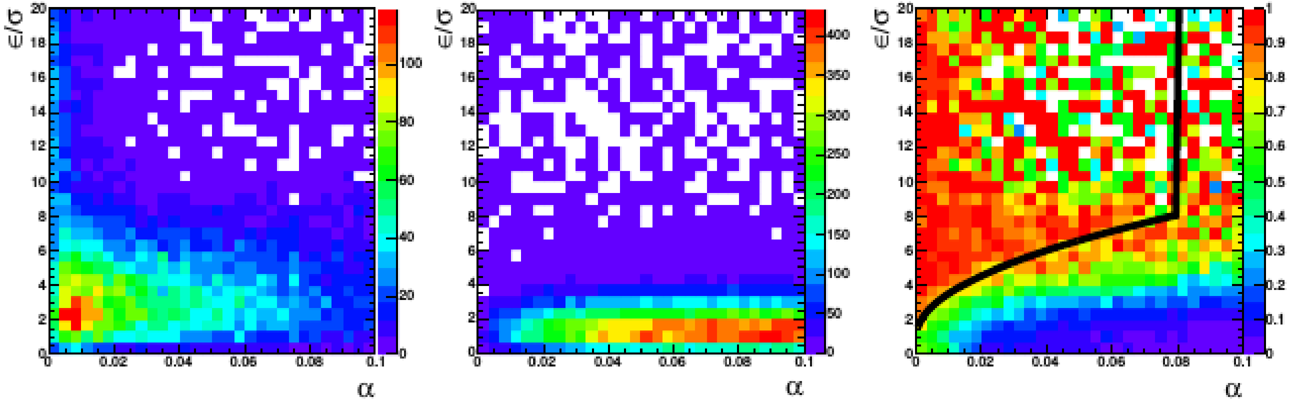
In the VertexRestorer part you have to modify the following lines:

```
<parameter name="NotUsedTracksCollectionName" type="string"> TracksFailBothCanFormPfoFlags </parameter>
by:  <parameter name="NotUsedTracksCollectionName" type="string"> MarlinTrkTracks </parameter>
```

and

```
<parameter name="UseTracks" type="int" value="0"/>
by: <parameter name="UseTracks" type="int" value="1"/>
```

- **S.B. ttbar (500GeV, DBD). Variables used to tune the recovery.** I did the same plots for a single file of ttbar (fully leptonic, without lepton tagging) and they look the same… unfortunately, I removed the file by mistake

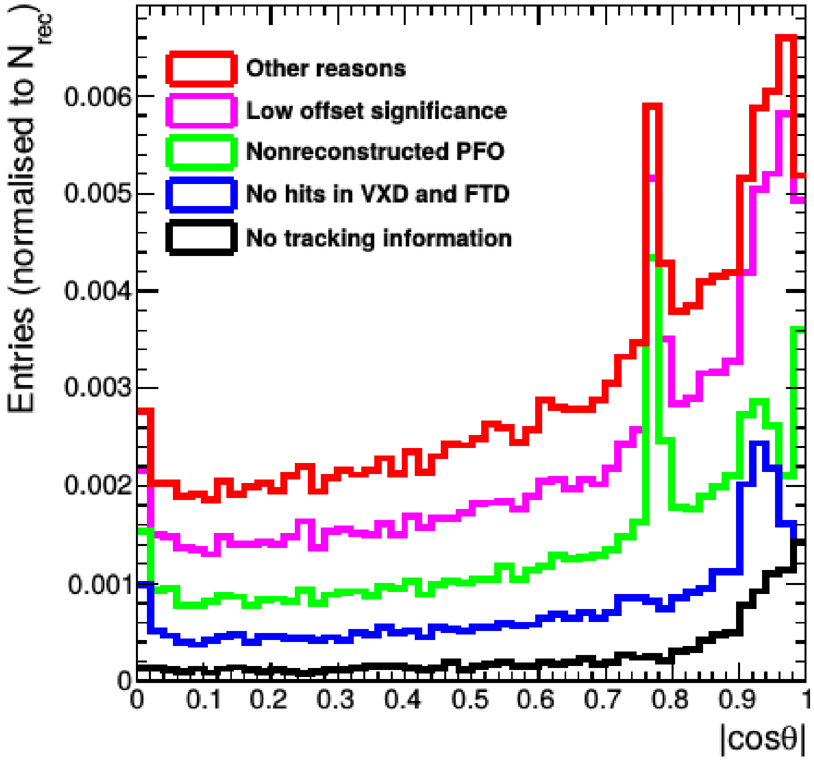


(a) Missing particles (b) Background (c) Purity map

Figure 9.14: *Distribution of the separation variables, the angle α and the offset significance ϵ/σ for the missing generated prongs and the background charged particles. Purity map shows the highest concentration of the missing generated prongs as compare to all charged particles. The black line demonstrates the chosen cut function.*

Track/Vertex Recovery (Reminder S.B)

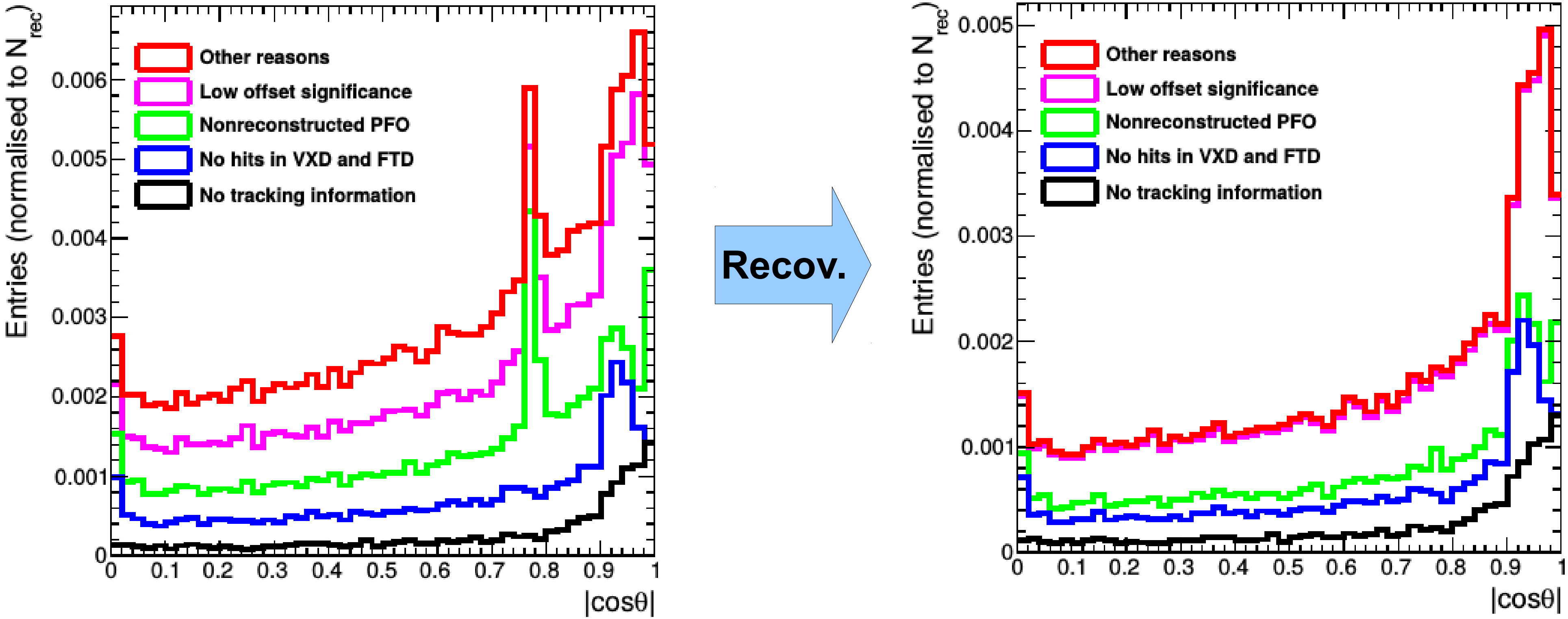
- **S.B. ttbar (500GeV, DBD). Lost tracks and recovered tracks.** I did the same plots for a single file of ttbar (fully leptonic, without lepton tagging) and they look the same… unfortunately, I removed the file by mistake.



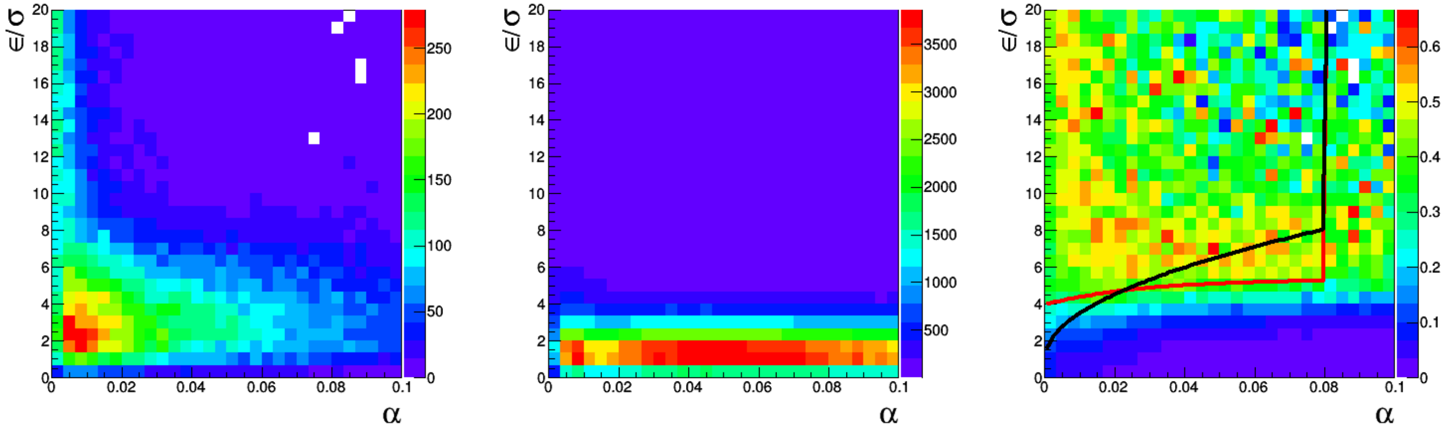
- No tracking information - the MarlinTrk algorithms fails to reconstruct the track. This category is tiny - only 0.93% of the generated prongs;
- No associated hits in the VXD or FTD - the track segment from the Vertex Detector or Forward Tracking Disks was not connected to the long TPC track segment. These reconstructed particles have large uncertainties on the impact parameters, which makes them not suitable for vertexing algorithms. They constitute 2.% of the generated prongs;
- No reconstructed PFO - the PandoraPFA fails to create the PFO from a reconstructed track. These tracks are discarded by the LCFI+ algorithms - 3.2% of the generated prongs;
- Low generated momentum or offset - the reconstructed particle was produced with impact parameters below the detector resolution - 3.1% of the generated prongs;
- Other reasons connected to vertex fitting problems - 1.7% of the generated prongs.

Track/Vertex Recovery (Reminder S.B)

- **S.B. ttbar (500GeV, DBD). Lost tracks and recovered tracks.** I did the same plots for a single file of ttbar (fully leptonic, without lepton tagging) and they look the same… unfortunately, I removed the file by mistake.

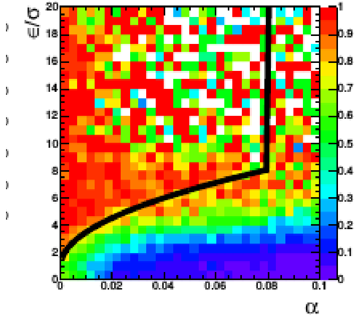


- A. Irles (bb, DBD)



- **Old parametization:**
  - E/sigma > 25sqrt(alpha)+1
  - + and arbitrary cut in a variable called "observable" which seems a function of alpha.
- **New parametrization**
  - E/sigma > 4 atan(alpha)+4

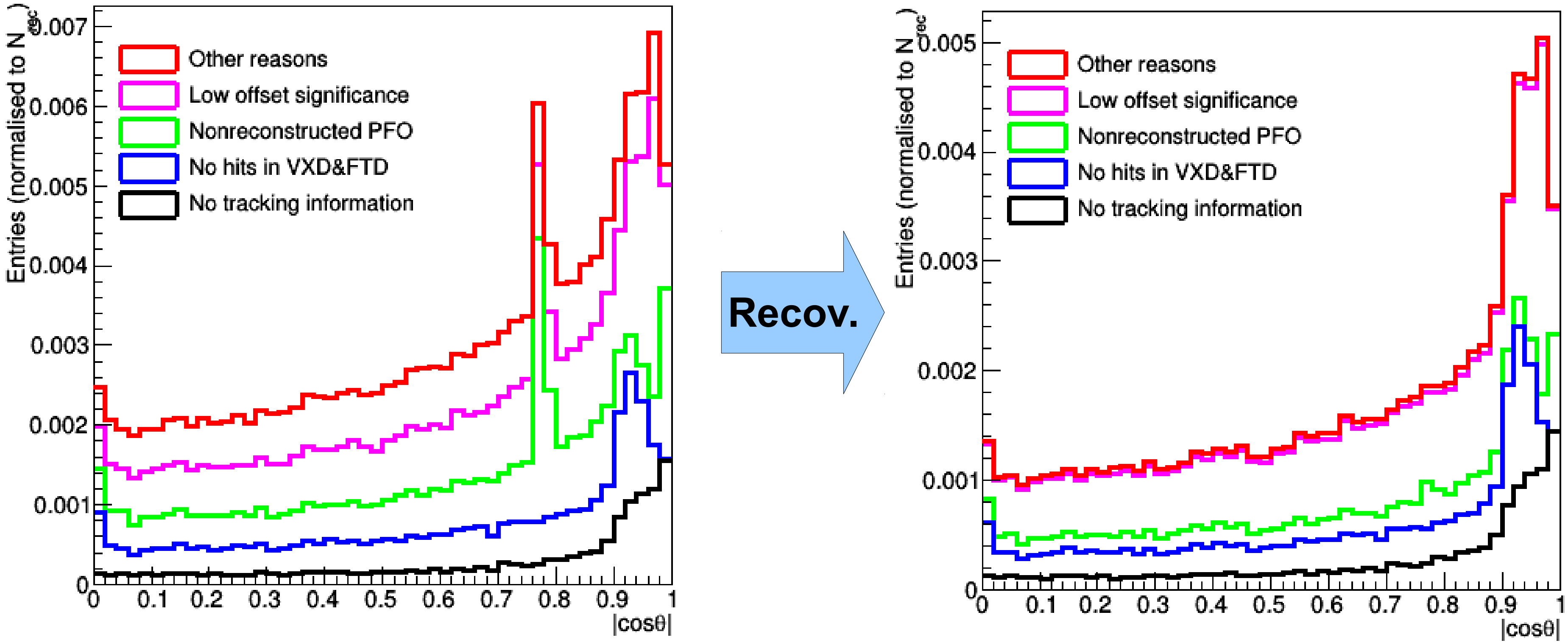
- **ttbar**



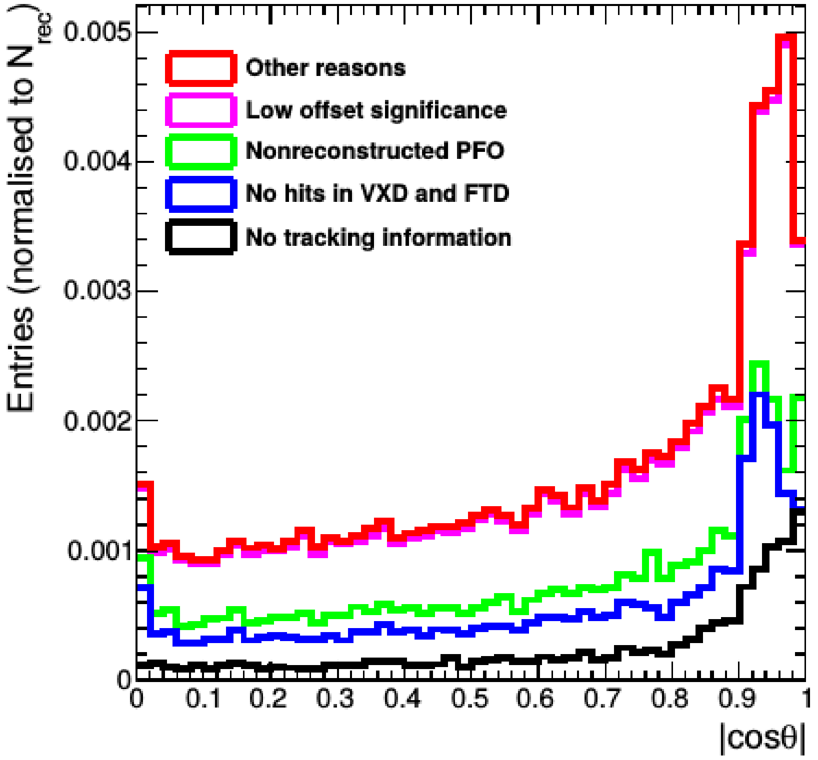
- **In both cases, if alpha<0.001, the track is recovered always.**

Track/Vertex Recovery

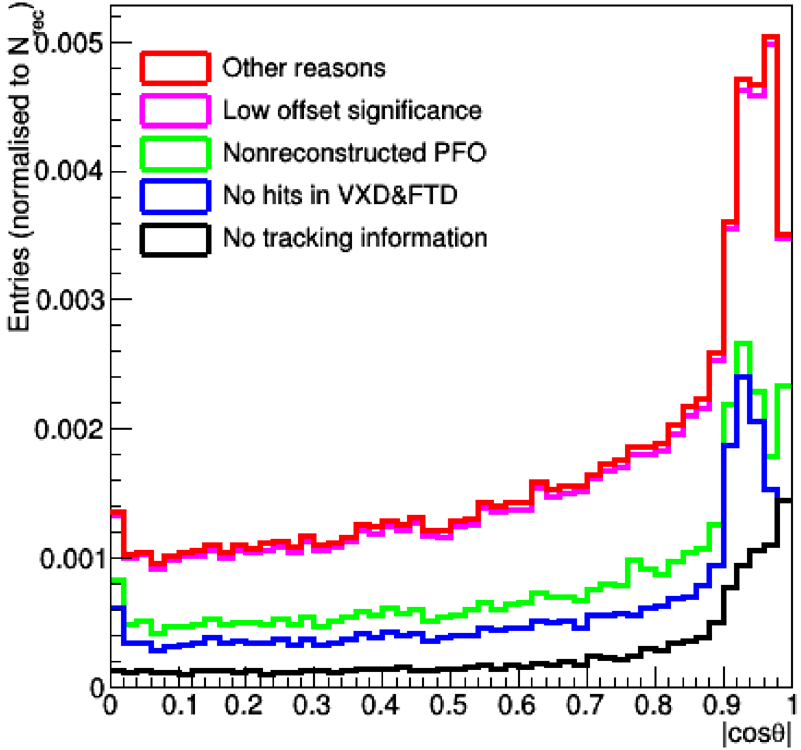
- **A. Irles (bbbar 250GeV, DBD)**



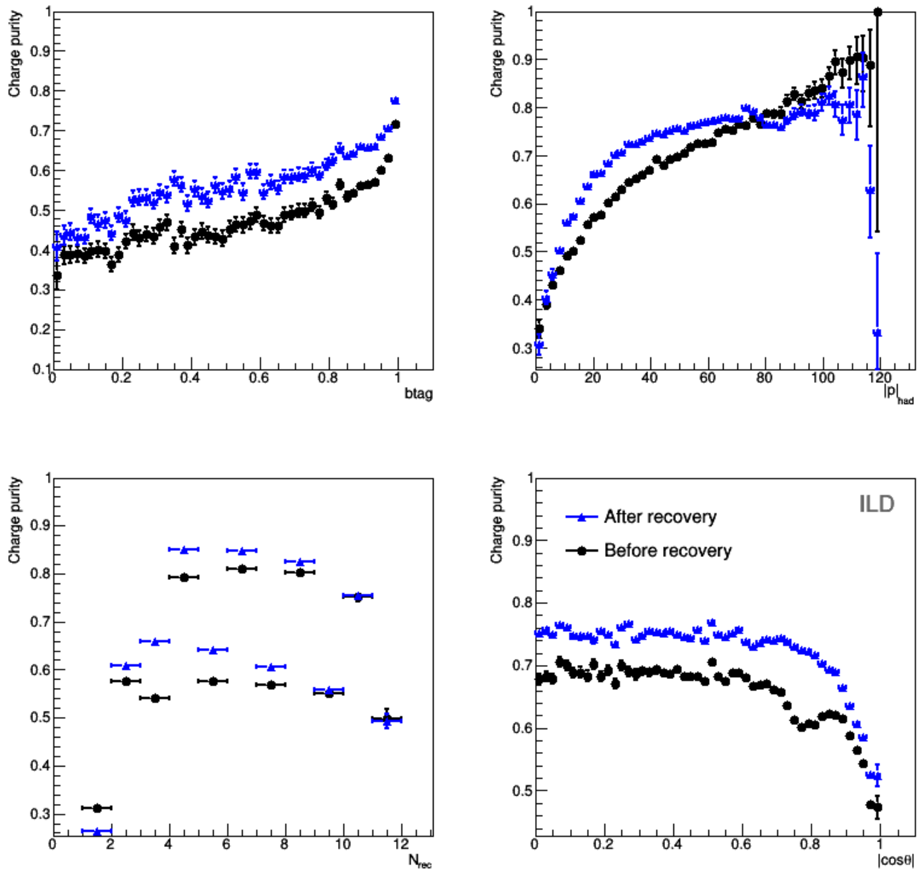
- **ttbar**


Entries (normalised to N_rec)
0.005
0.004
0.003
0.002
0.001
0
Other reasons
Low offset significance
Nonreconstructed PFO
No hits in VXD and FTD
No tracking information
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
|cosθ|

- **bb**

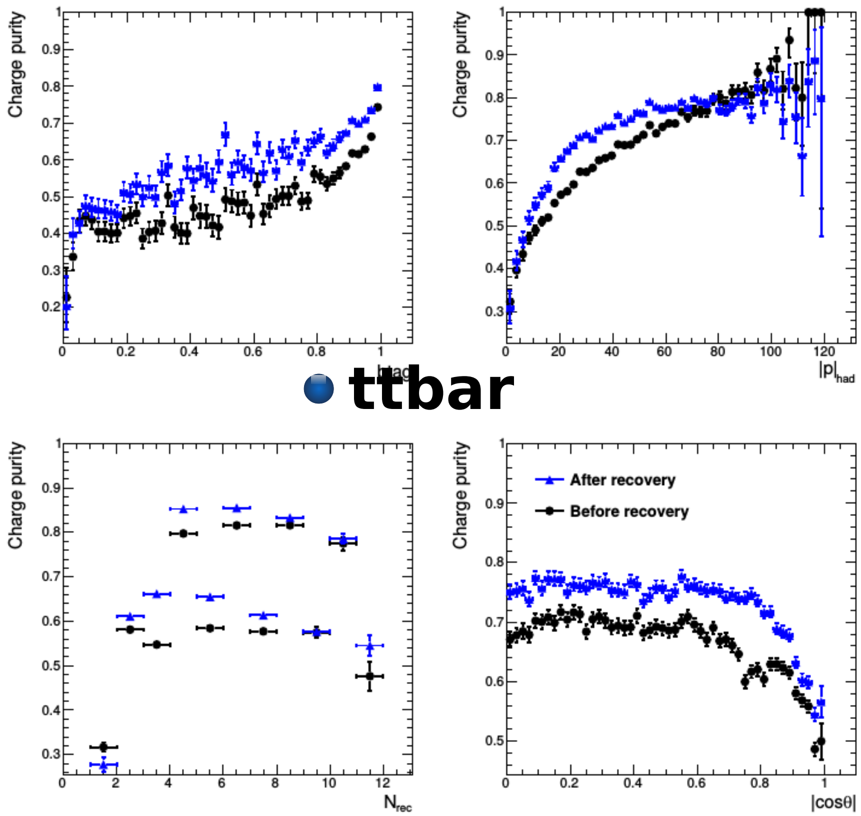

Entries (normalised to N_rec)
0.005
0.004
0.003
0.002
0.001
0
Other reasons
Low offset significance
Nonreconstructed PFO
No hits in VXD&FTD
No tracking information
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
|cosθ|

- **A. Irles (bbbar 250GeV, DBD)**

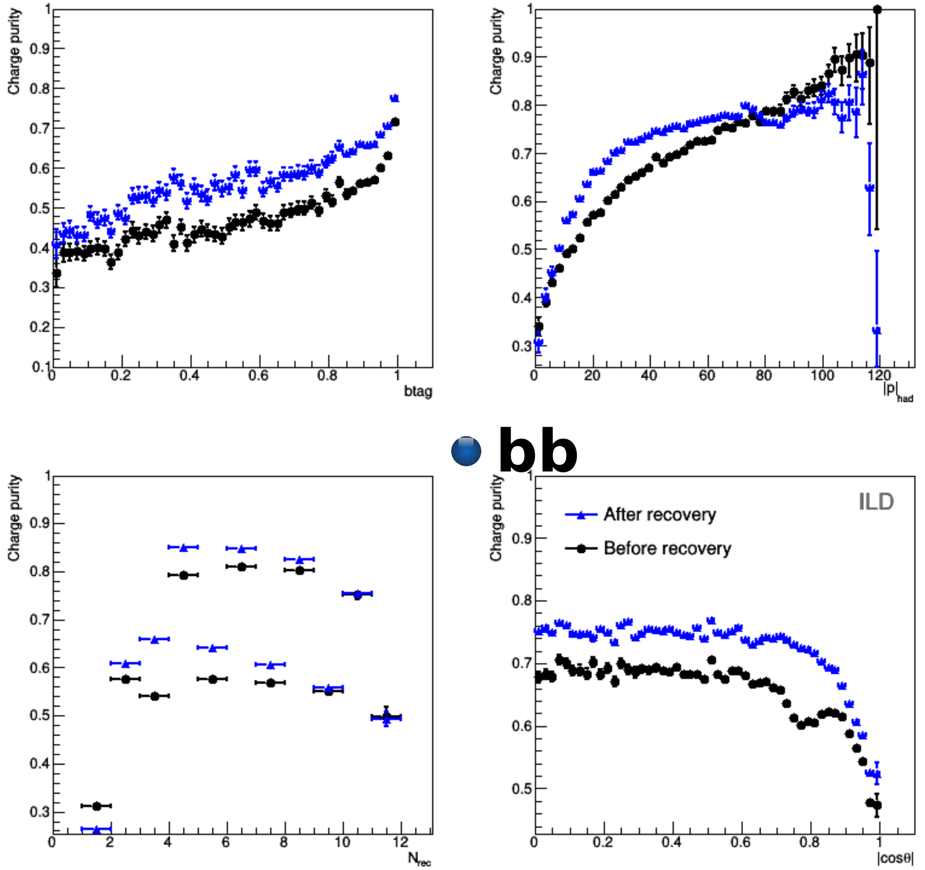


Track/Vertex Recovery

- **ttbar**

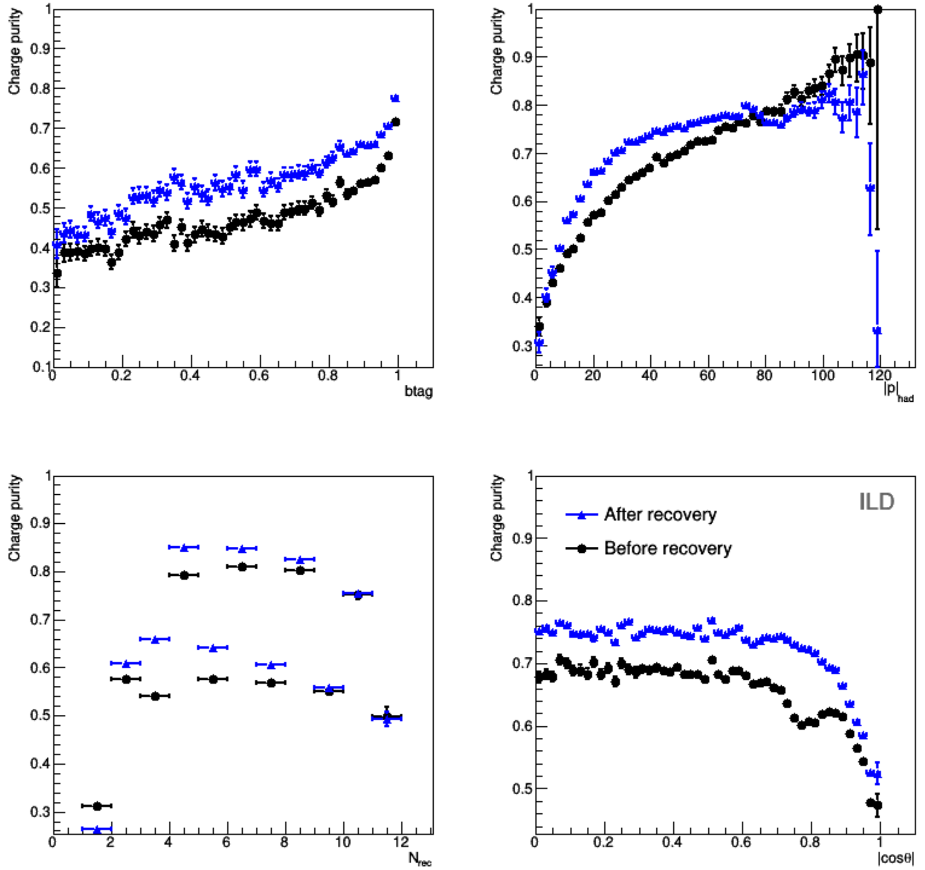


- **bb**

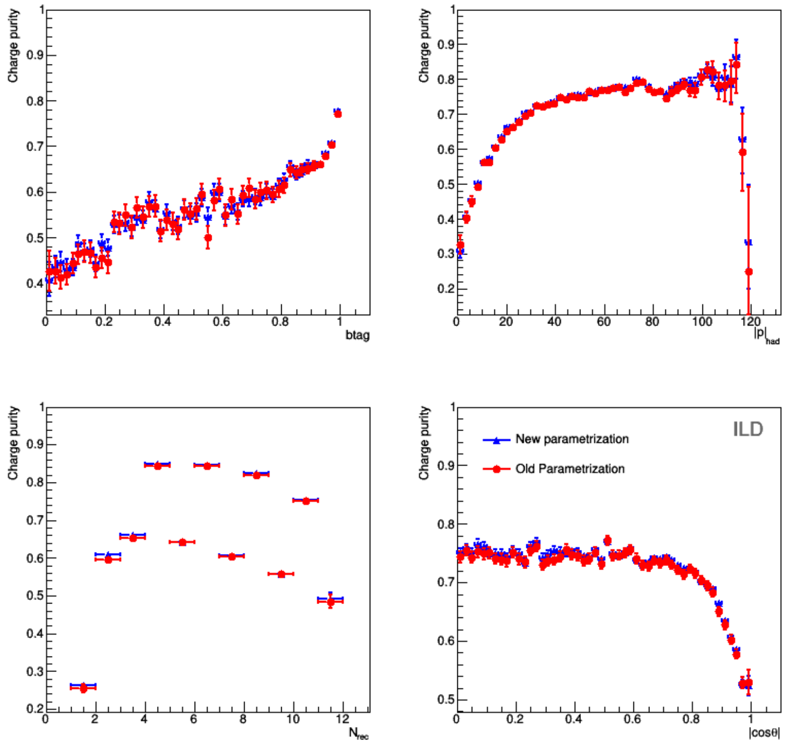


Track/Vertex Recovery

- **A. Irles (bbbar 250GeV, DBD)**

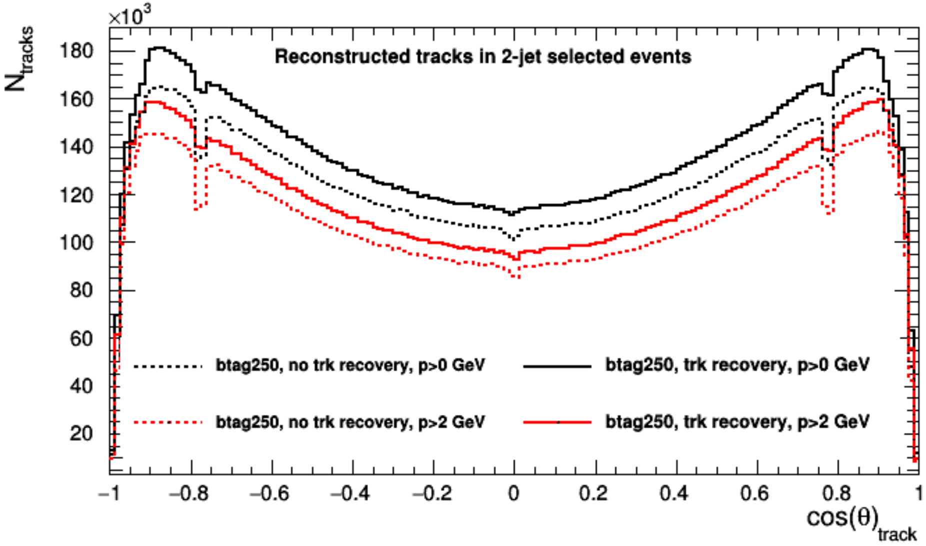


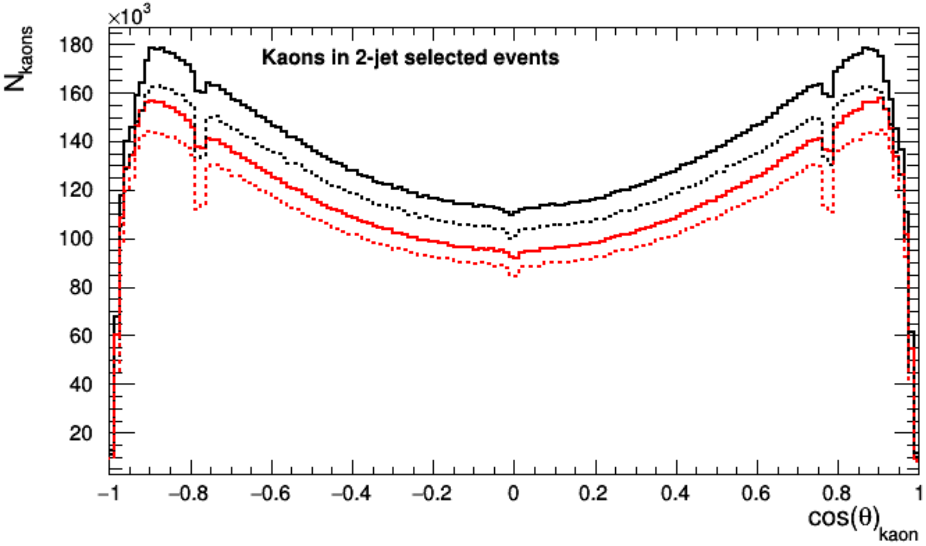
- **We are now (almost) sure that we use the method properly.**
- **Still some issues to understand:**
  - Why the arbitrary "observable" variable in the recovery? Using it or not makes no difference for bbbar.
  - Is the processor doing the same when is running in Test mode than when is not? The test mode is needed for the production of the alpha plots to tune the parametrization.
  - The impact of the new vs old parametrization is minimum
- **Any new parametrization has to be hardcoded, for the meanwhile**

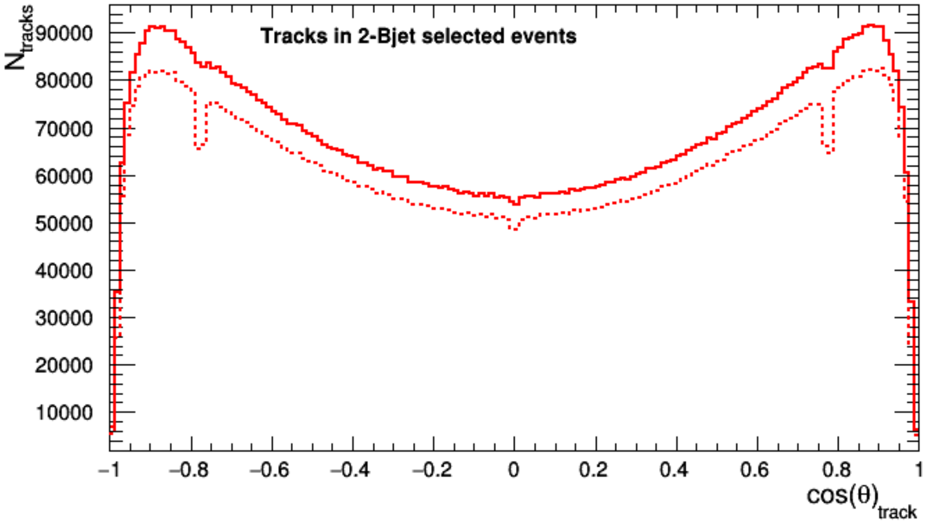


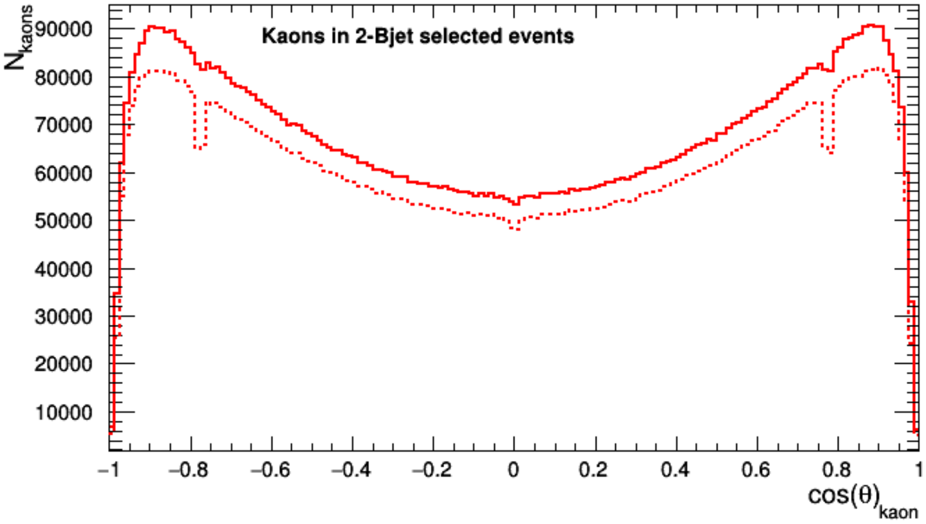
Track/Vertex Recovery: impact on the observable

- **Left polarization: total efficiency is slightly increased (+0.7%)**
- The “horns” are mostly recovered. Next slide where I show the cos(theta) of all tracks in all the simulated events and all of them but having two b-tags jets
- The purity in the charge calculation increases

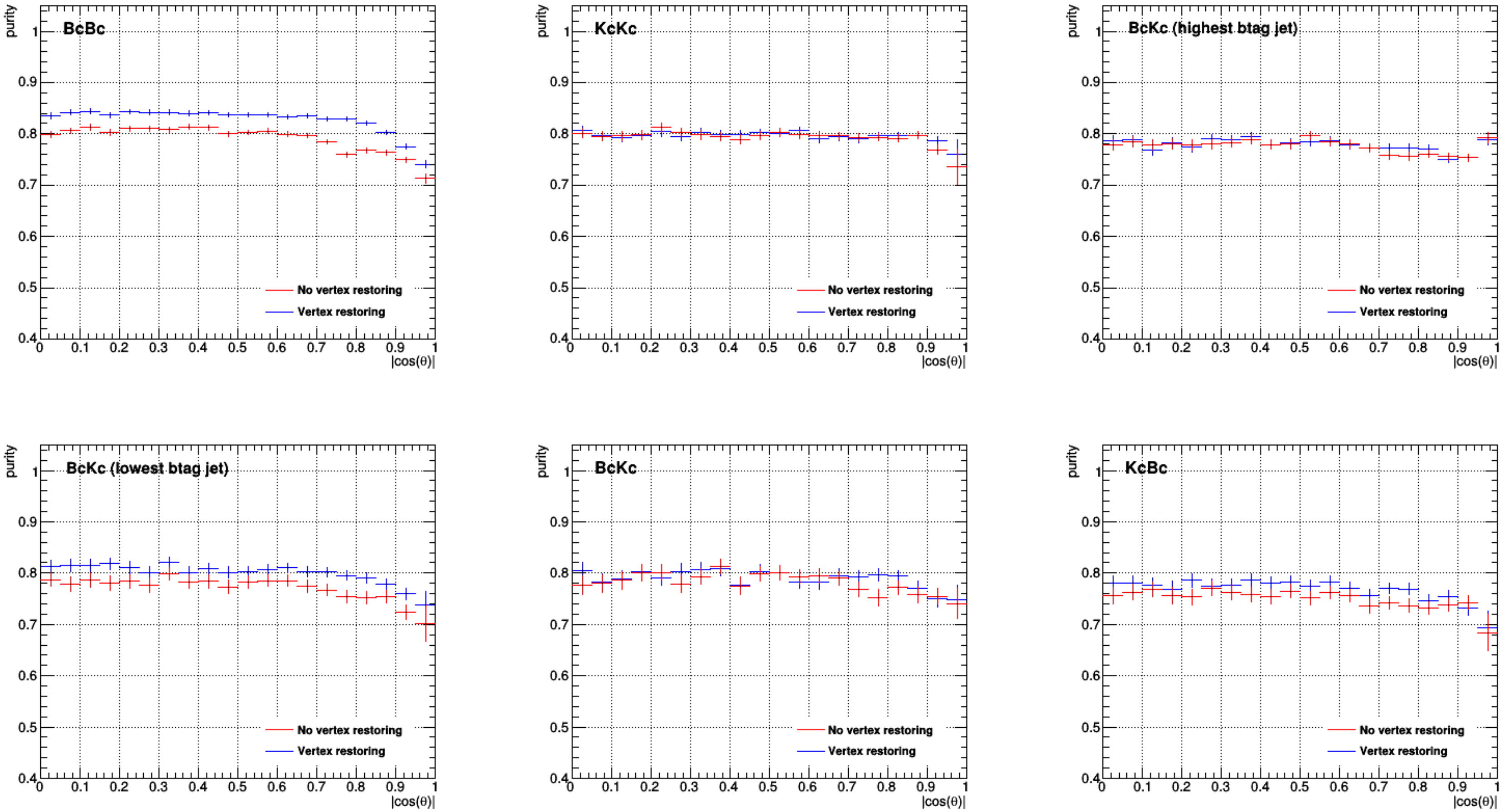

Reconstructed tracks in 2-jet selected events
N_{tracks}
$\times 10^3$
180
160
140
120
100
80
60
40
20
btag250, no trk recovery, p>0 GeV
btag250, trk recovery, p>0 GeV
btag250, no trk recovery, p>2 GeV
btag250, trk recovery, p>2 GeV
−1 −0.8 −0.6 −0.4 −0.2 0 0.2 0.4 0.6 0.8 1
$\cos(\theta)_{track}$


Kaons in 2-jet selected events
N_{kaons}
$\times 10^3$
180
160
140
120
100
80
60
40
20
−1 −0.8 −0.6 −0.4 −0.2 0 0.2 0.4 0.6 0.8 1
$\cos(\theta)_{kaon}$


Tracks in 2-Bjet selected events
N_{tracks}
90000
80000
70000
60000
50000
40000
30000
20000
10000
−1 −0.8 −0.6 −0.4 −0.2 0 0.2 0.4 0.6 0.8 1
$\cos(\theta)_{track}$

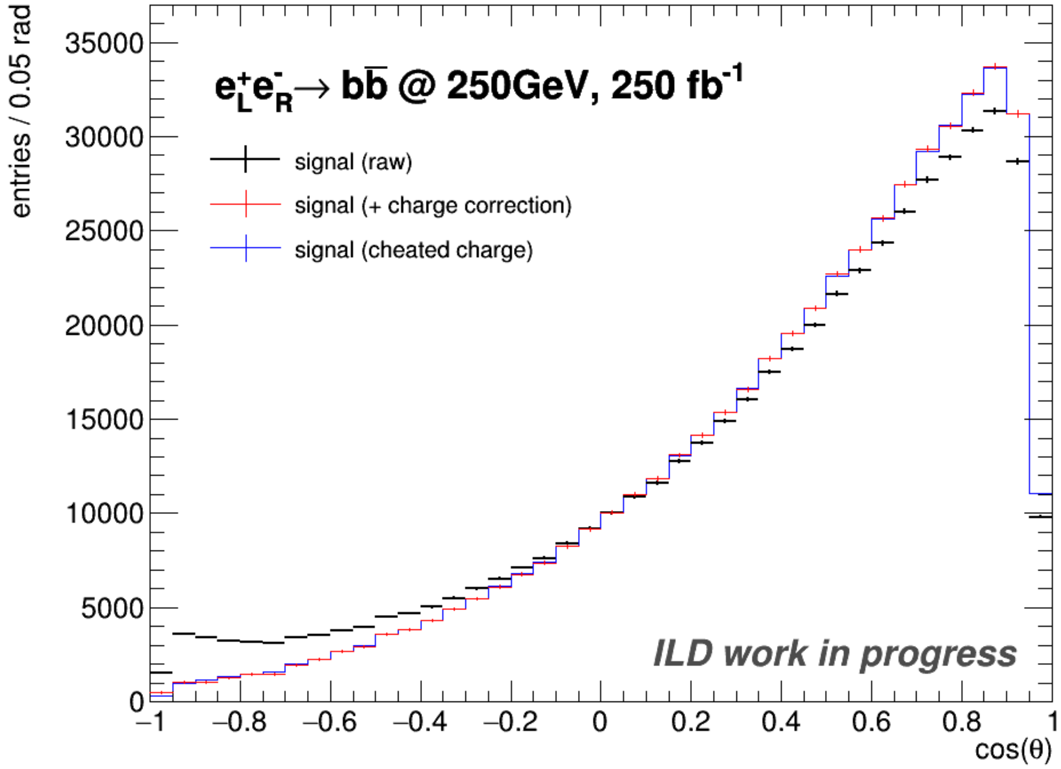

Kaons in 2-Bjet selected events
N_{kaons}
90000
80000
70000
60000
50000
40000
30000
20000
10000
−1 −0.8 −0.6 −0.4 −0.2 0 0.2 0.4 0.6 0.8 1
$\cos(\theta)_{kaon}$

Track/Vertex Recovery: impact on the measured purity



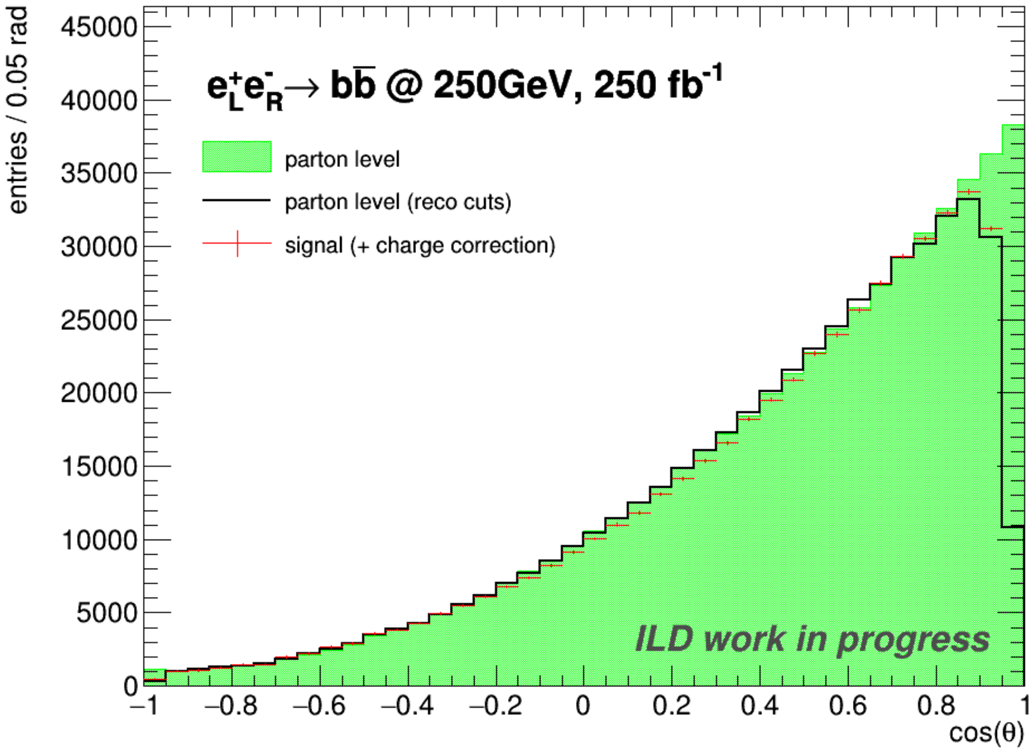
Detector Correction issue

Detector Correction issue: Intro



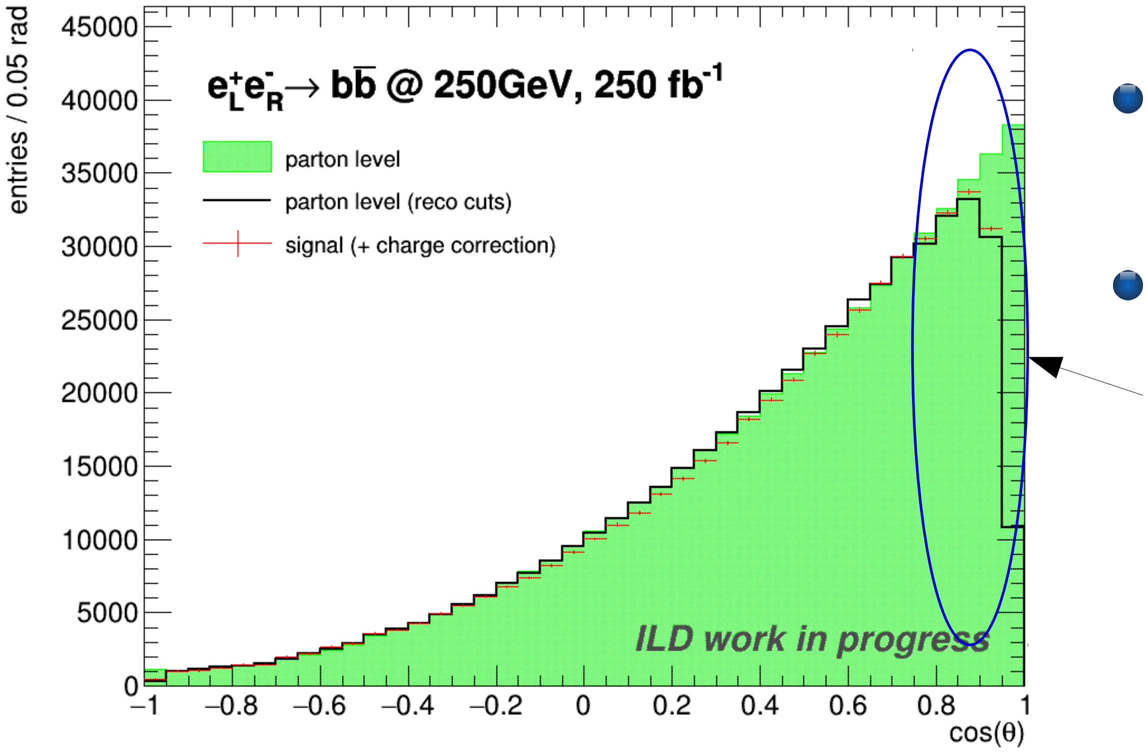
- Left plot: comparison of the result of the p-q method (red) and the reconstructed distribution with cheated charge (blue)
- **The p-q method works perfectly**
  - But… (next slide)

Detector Correction issue: Intro



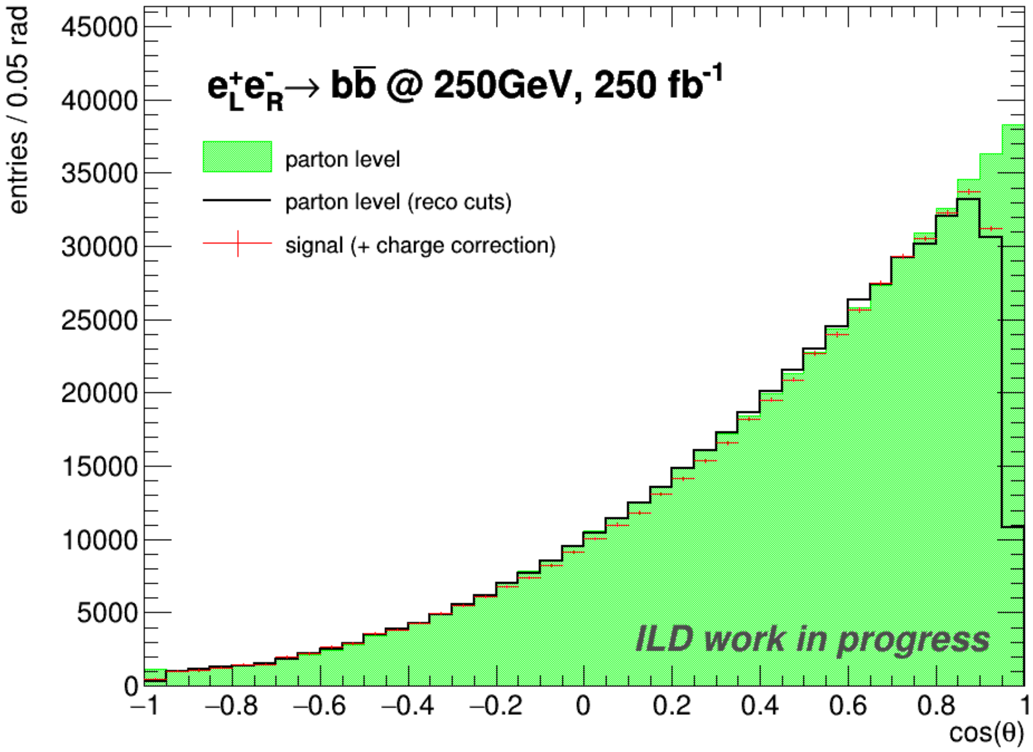
- When comparing the charge-corrected distribution with the parton level, we see a large disagreement in the barrel and in the forward regions.
- What is the origin?

Detector Correction issue: Intro



- When comparing the charge-corrected distribution with the parton level, we see a large disagreement in the barrel and in the forward regions.
- What is the origin?
  - Detector acceptance issues (due to geometry)

Detector Correction issue: Intro

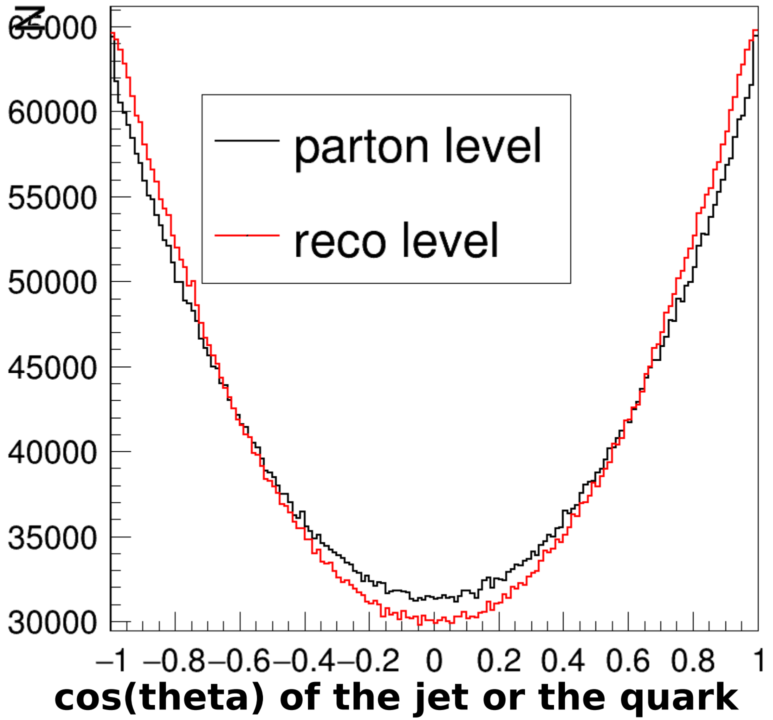


- What is happening in the barrel?
  - We thought that was an acceptance issue due to the lost tracks (together with a problem in the normalization)
  - But, when we compare with the parton level distribution for only the reconstructed events, we see the same pattern.
  - Plus a residual acceptance issue (distance between black and green)

- **All this hints for migration effects due to detector resolution in the measured angle.**

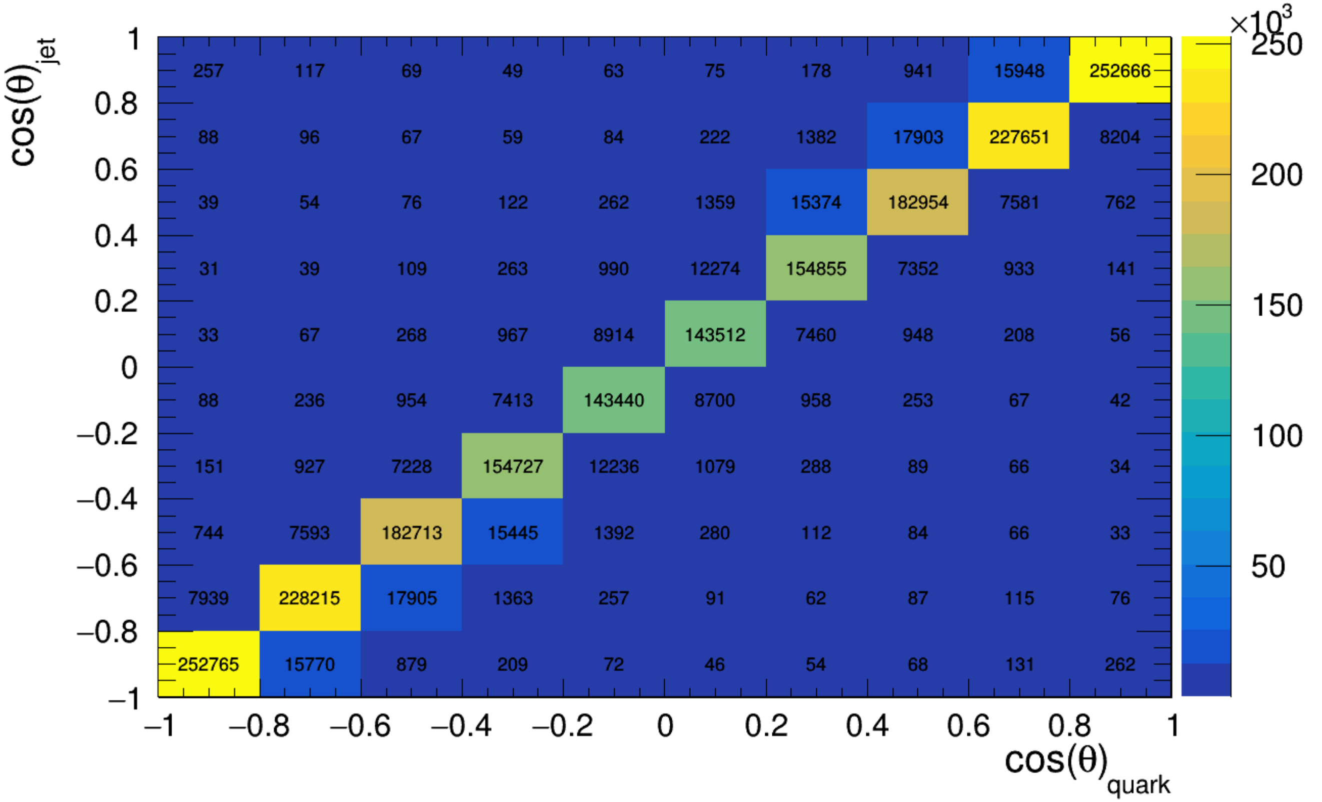
Detector Correction issue: going back to the jet level

2 entries per event (all events)



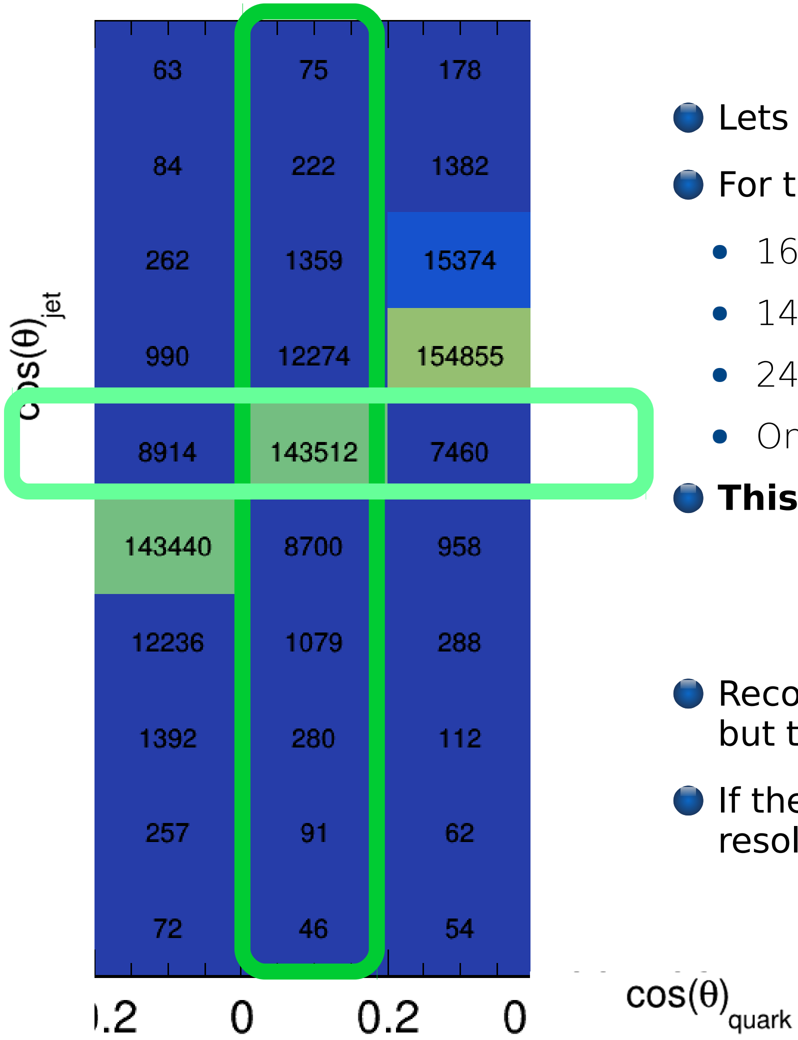
- First look at the resolution effects.
- Left plot: jet/quark angular distribution for ALL simulated EVENTS.
  - We have exactly the same number of events in both plots but a different shape.

Detector Correction issue: going back to the jet level



- Jets for final selection cuts.
- The resolution matrix is asymmetric (near the diagonal)

Detector Correction issue: going back to the jet level



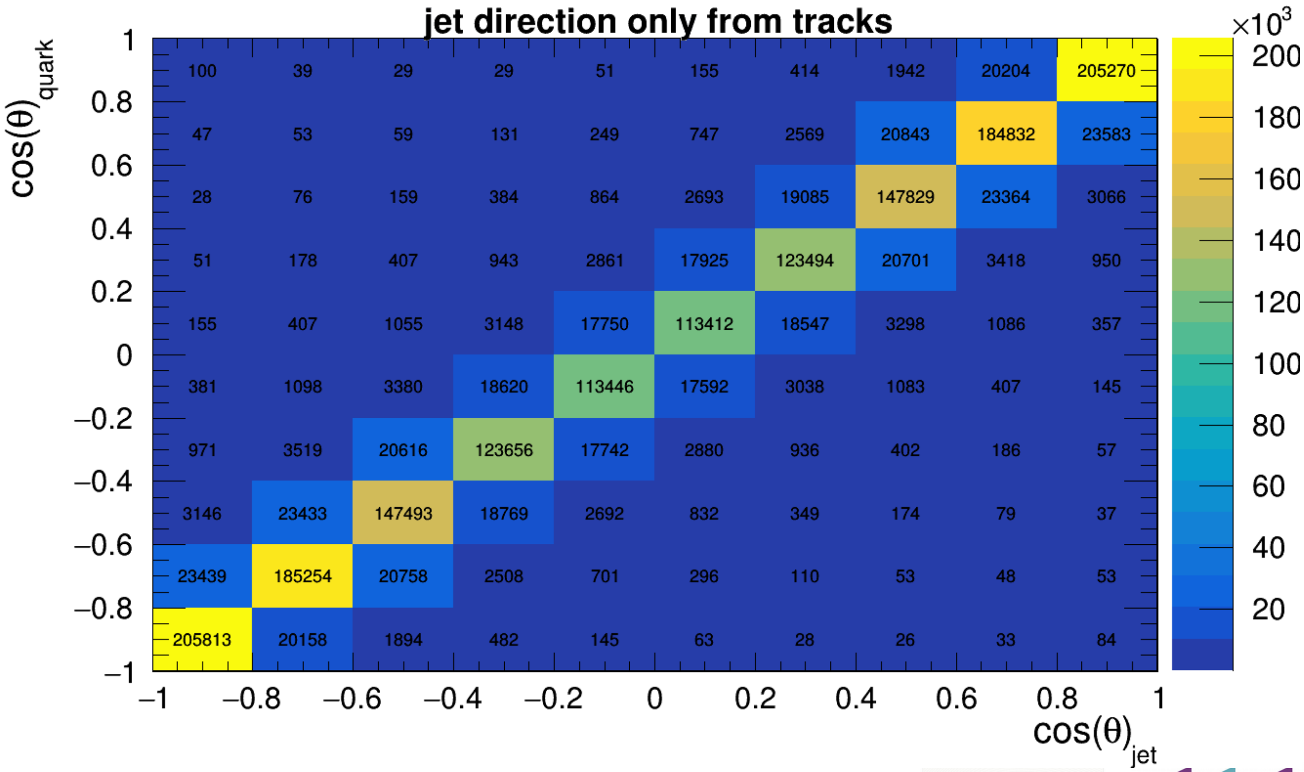
- Lets look at only a couple of bins.
- For the 0-0.1
  - 167638 events are generated
  - 143512 (85.6%) are reconstructed in the correct bin.
  - 24126 go to different bins.
  - Only 18921 events are migrated in from different bins.
- **This will be seen as a 3% of acceptance loss !!**

- Recovering tracks can slightly improve the selection after btagging but the impact of the migrations is still large.
- If the tracks would have been added to the jet... would the resolution matrix improve?

Detector Correction issue: going back to the jet level

- **Exercise**, try to define the jet direction using only tracks from secondary vertexes. **AFTER VERTEX RECOVERY.**
- The matrix is more diagonal, although the migrations are larger.

- Repeat the same exercise for the bin 0-0.1
  - 156595 events are generated
  - 113412 (72,4%) are reconstructed in the correct bin.
  - 43183 go to different bins.
  - and 45803 events are migrated in from different bins.
  - → **Acceptance increase of 1.6%**

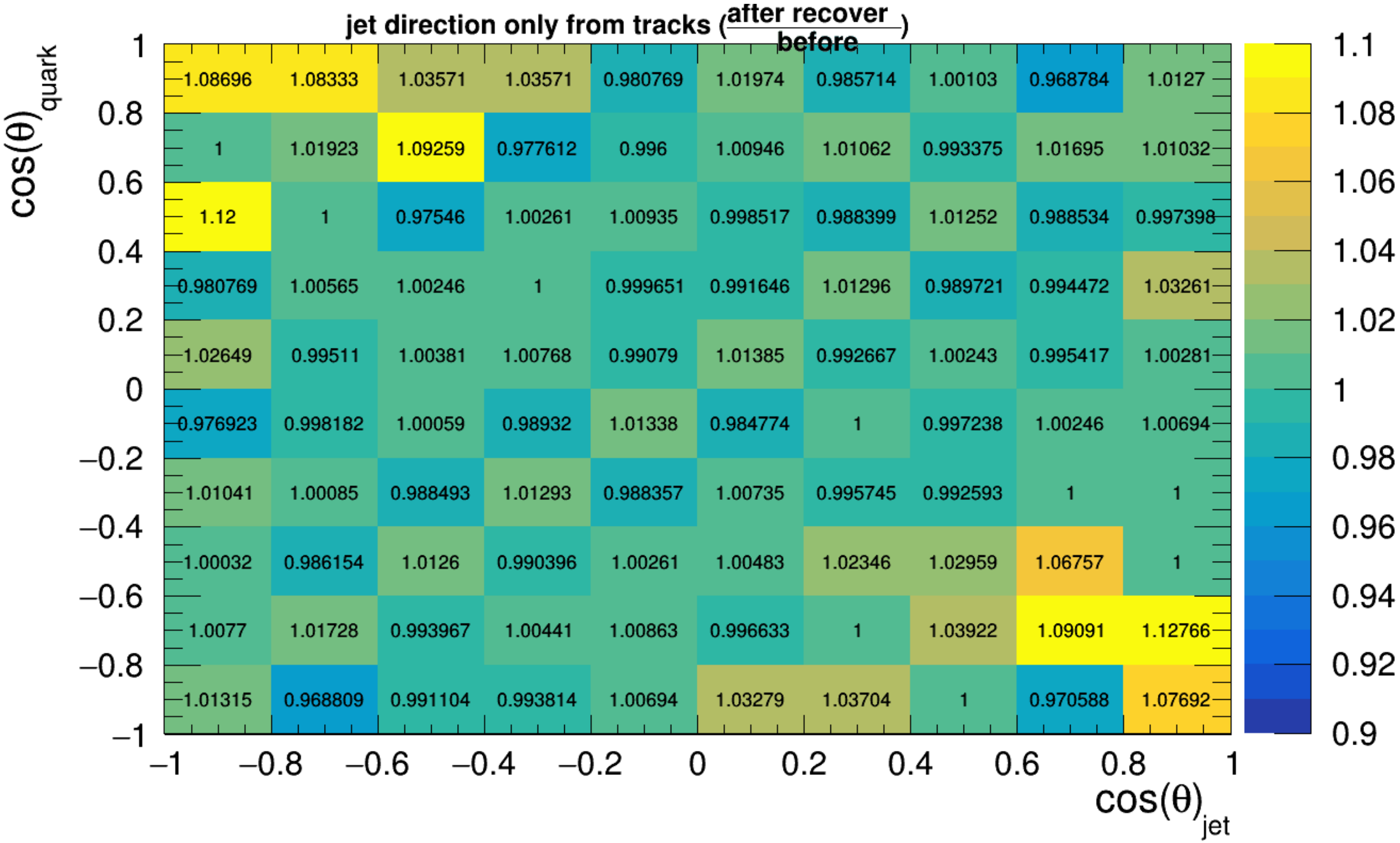


PRESTIGE

LABORATOIRE DE L'ACCÉLÉRATEUR LINÉAIRE

Detector Correction issue: going back to the jet level

- Comparison **before and after recovery: the matrix systematically improves (in ~1-2%)**

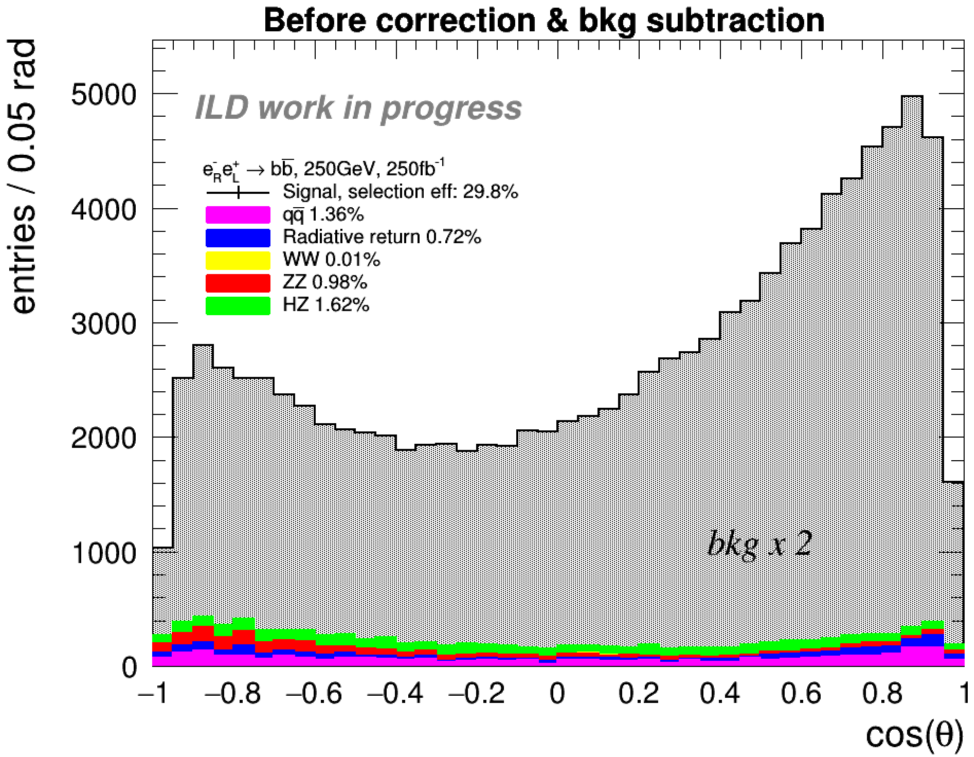


Detector Correction issue: going back to the jet level

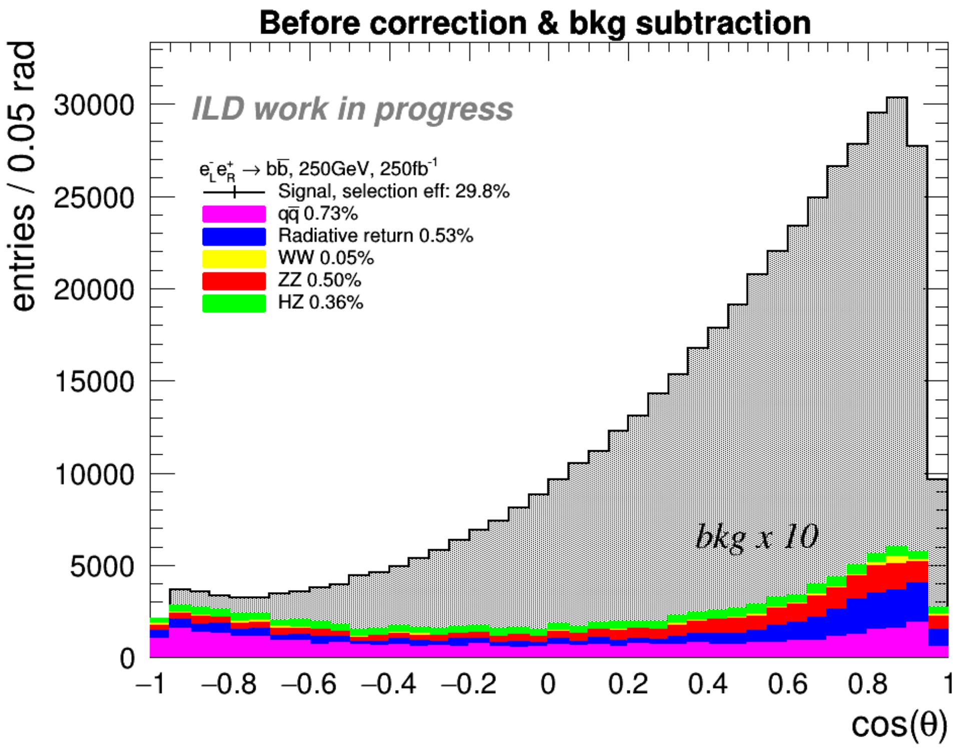
- **Conclusions**:
- The resolution effects are seen as acceptance effects. If the resolution is well treated, the acceptance effects should be smaller (at least far from the forward regions).
- Defining the jet direction using only secondary tracks is not optimal but:
  - If we properly use the recovered tracks for the recalculation of the kinematics of the jets will improve the matrix (including all the rest of PFOs)
- **Short term plan: apply correction factors for these detector effects.**
  - Next slides.
- **Longer term plans:**
  - Try to reduce these corrections to the minimum.
  - For that we should study this more deeply with the new software (and maybe with newly generated samples at 250GeV).

Results

Results



Before correction & bkg subtraction
ILD work in progress
$e^-_R e^+_L \to b\bar{b}$, 250GeV, 250fb^{-1}
Signal, selection eff: 29.8%
$q\bar{q}$ 1.36%
Radiative return 0.72%
WW 0.01%
ZZ 0.98%
HZ 1.62%
bkg x 2
entries / 0.05 rad
cos(θ)

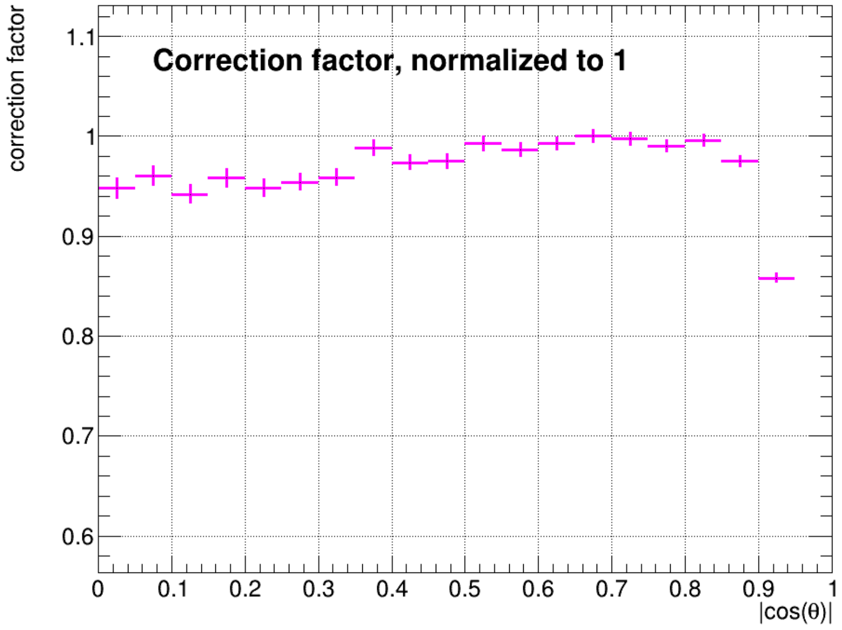


Before correction & bkg subtraction
ILD work in progress
$e^-_L e^+_R \to b\bar{b}$, 250GeV, 250fb^{-1}
Signal, selection eff: 29.8%
$q\bar{q}$ 0.73%
Radiative return 0.53%
WW 0.05%
ZZ 0.50%
HZ 0.36%
bkg x 10
entries / 0.05 rad
cos(θ)

How to correct for detector effects?

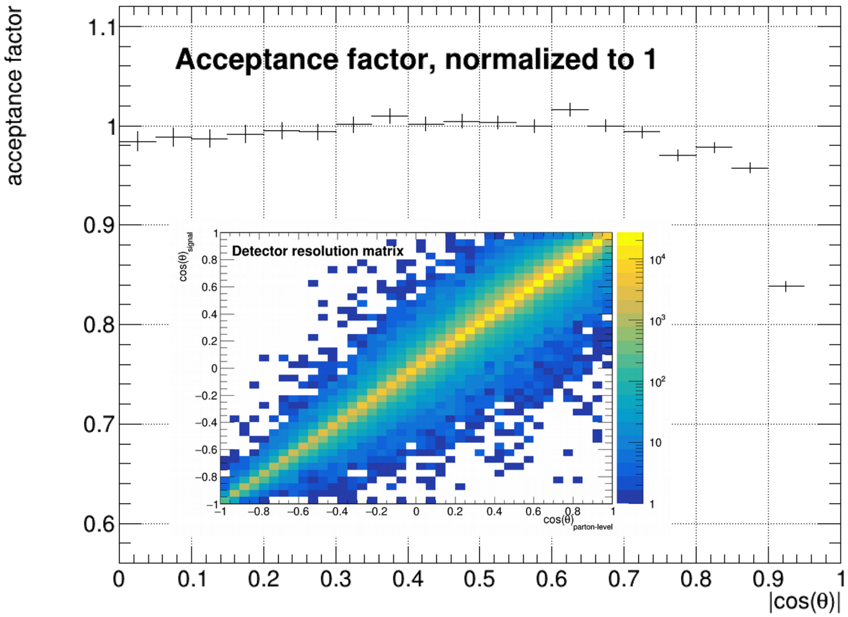
Option a)

- Correction factor = $(\text{parton/truthreco})^{-1}$



Option b)

- Corrected = M^{-1} x Reco x $(\text{Acceptance})^{-1}$ (atPartonLevel)

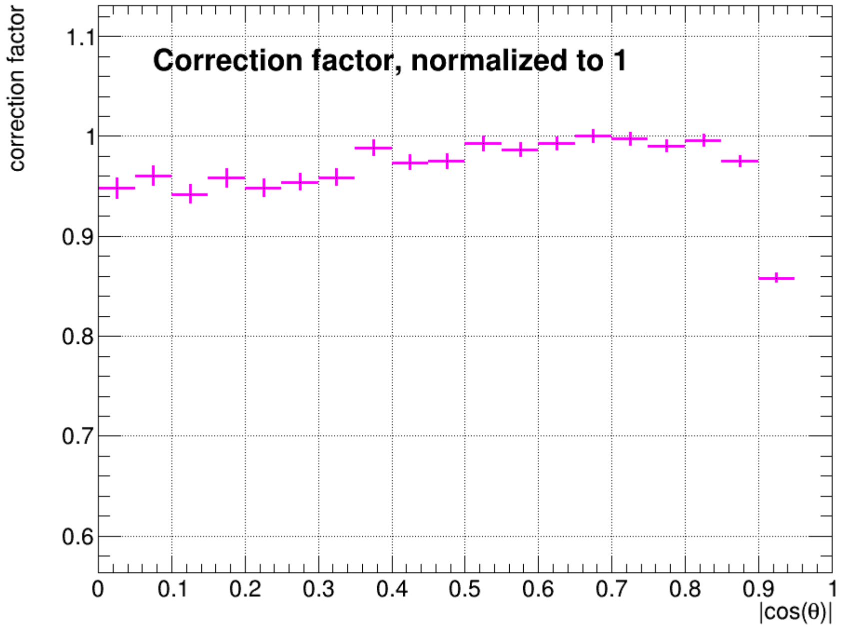


This factor includes correction for normalization, **acceptance and resolution effects.**

How to correct for detector effects?

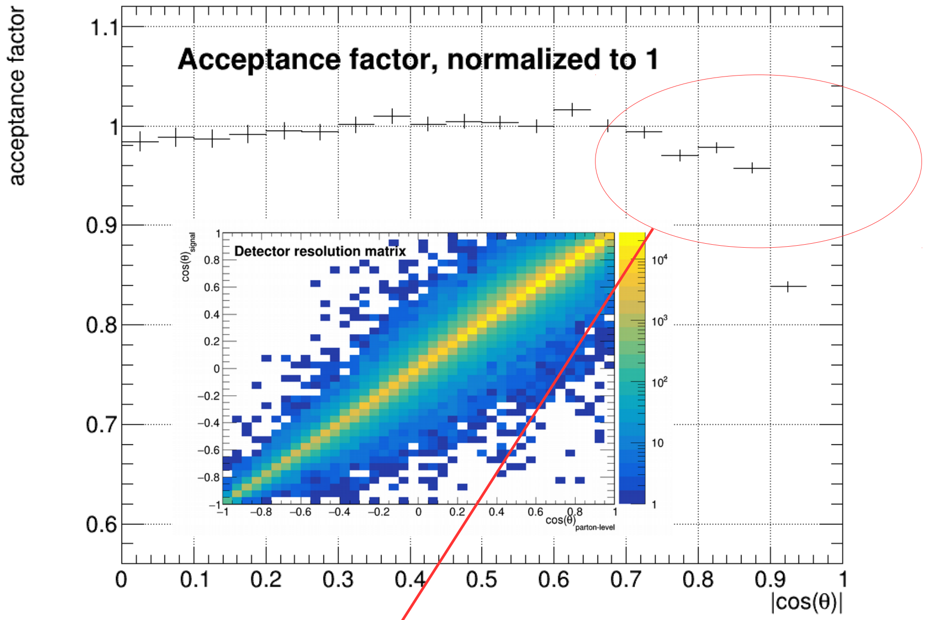
Option a)

- Correction factor = (parton/truthreco)$^{-1}$



Option b)

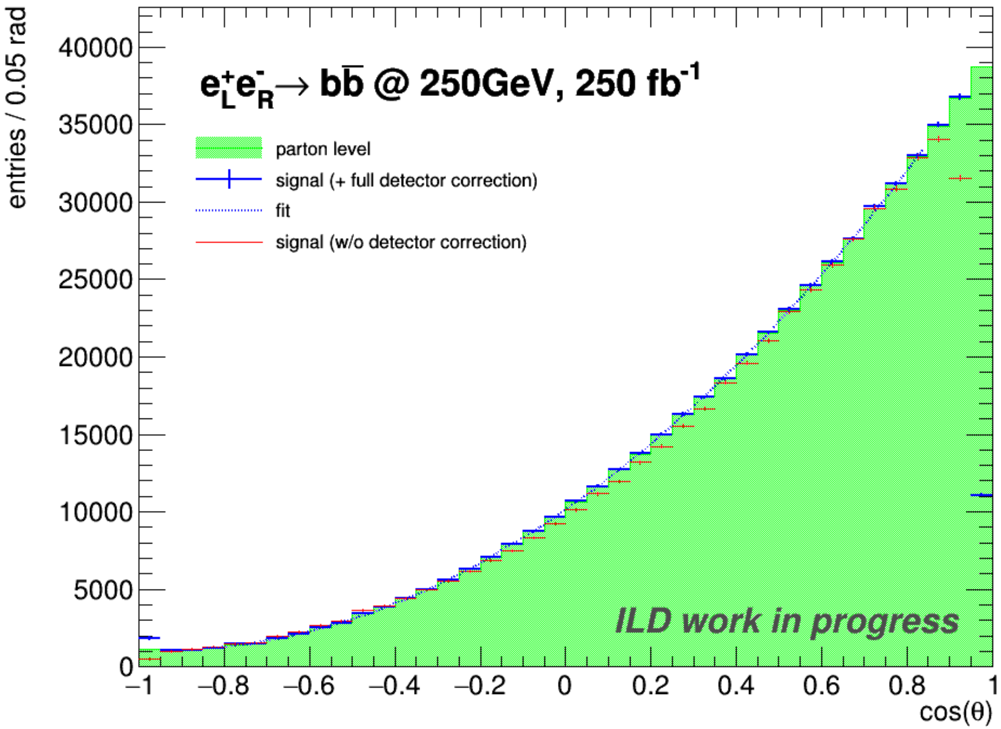
- Corrected = M^{-1} x Reco x (Acceptance)$^{-1}$ (atPartonLevel)



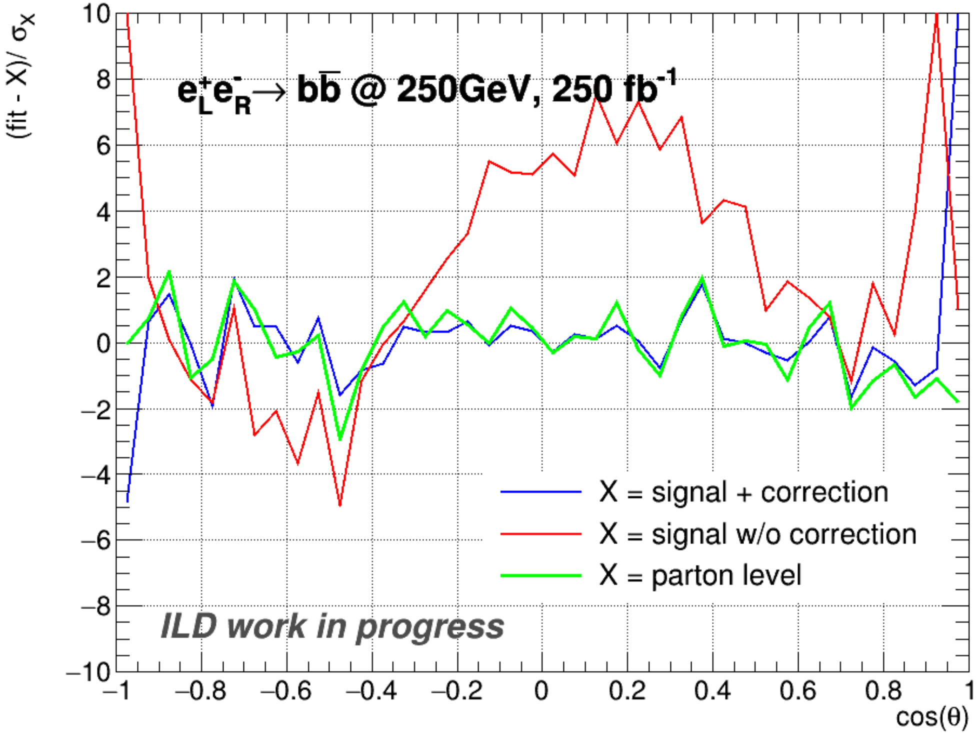
- **Is the acceptance correction starting to be large already a bit before than cos(theta)=0.8 ??**

PRESTIGE POSTDOCTORAL RESEARCH FELLOWSHIPS IN FRANCE
LABORATOIRE DE L'ACCÉLÉRATEUR LINÉAIRE

Option a), eL

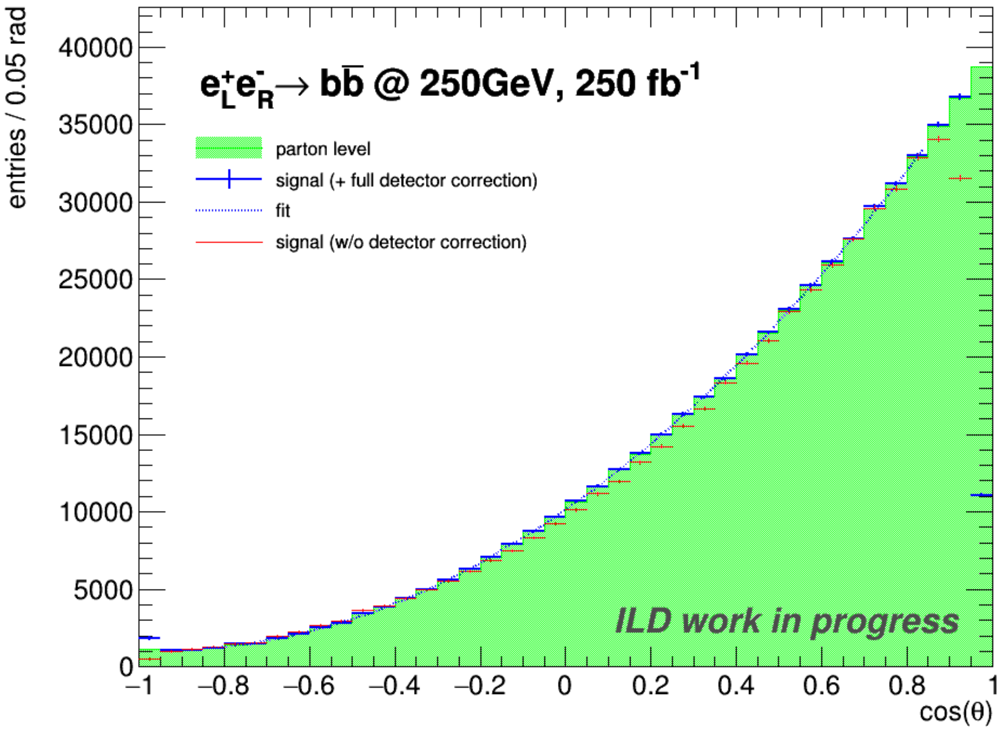


$e^+_L e^-_R \rightarrow b\bar{b}$ @ 250GeV, 250 fb^{-1}
parton level
signal (+ full detector correction)
fit
signal (w/o detector correction)
ILD work in progress
entries / 0.05 rad
cos(θ)



$e^+_L e^-_R \rightarrow b\bar{b}$ @ 250GeV, 250 fb^{-1}
X = signal + correction
X = signal w/o correction
X = parton level
ILD work in progress
(fit - X)/ σ_X
cos(θ)

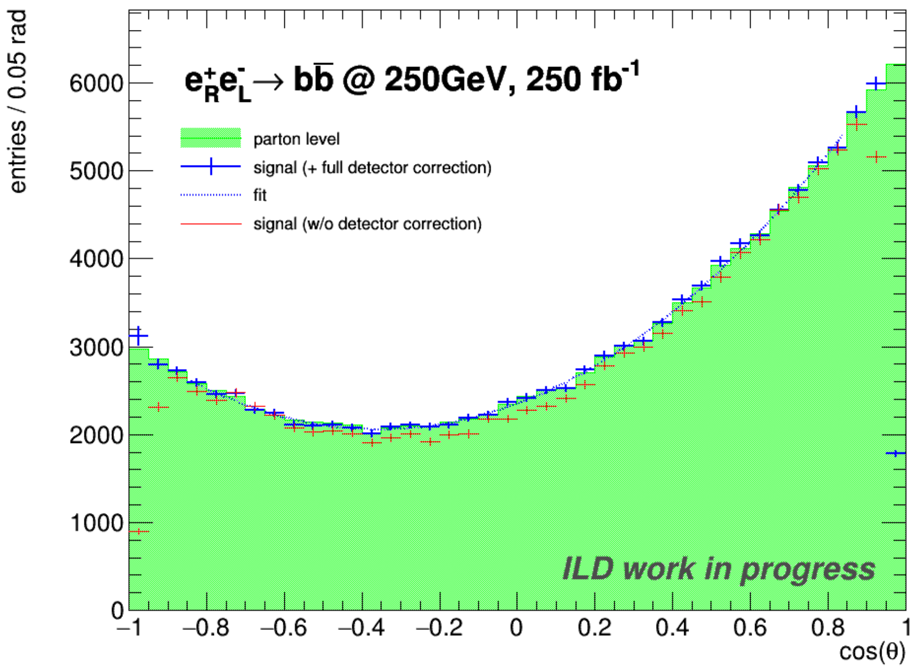
Option a), eL

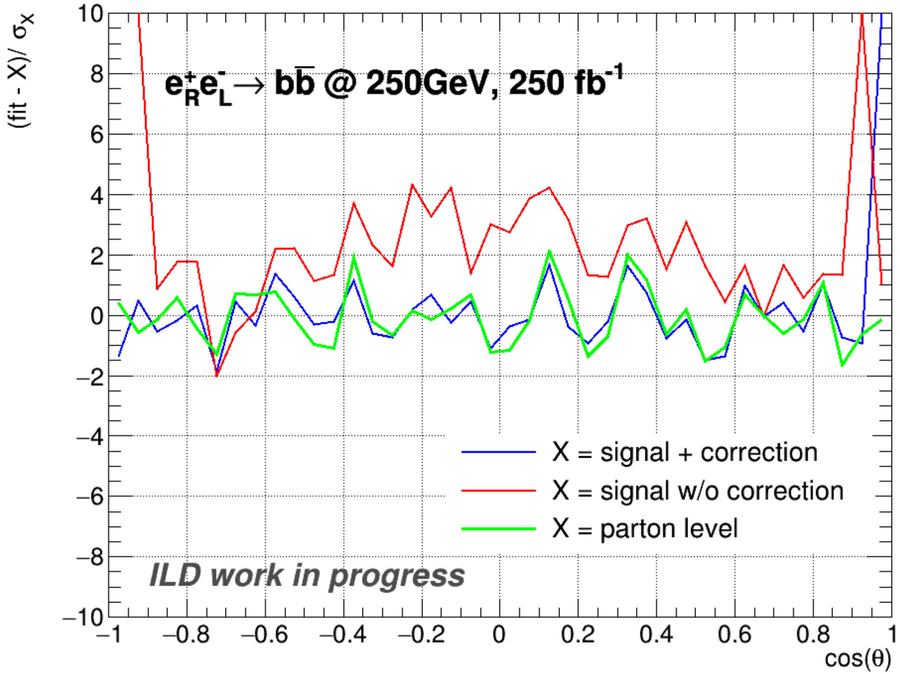



```
####### LEFT POLARIZATION
 -------------------------
250fb-1, parton level distribution
Afb = 0.707516 +- 0.00137244  ( 0.19398 %)
S = 10156.7 +- 9.03981  ( 0.089003 %)
A = 19162.8 +- 20.9342  ( 0.109244 %)
Correl = 0.814717
Chi2 = 35.5937, NDF =32 Chi2/NDF= 1.1123

250fb-1, reco + fully corrected distribution
Afb = 0.709079 +- 0.00216884  ( 0.305868 %)
S = 10158.2 +- 14.4238  ( 0.141992 %)
A = 19207.9 +- 35.1474  ( 0.182984 %)
Correl = 0.694899
Chi2 = 22.2375, NDF =32 Chi2/NDF= 0.694921
```

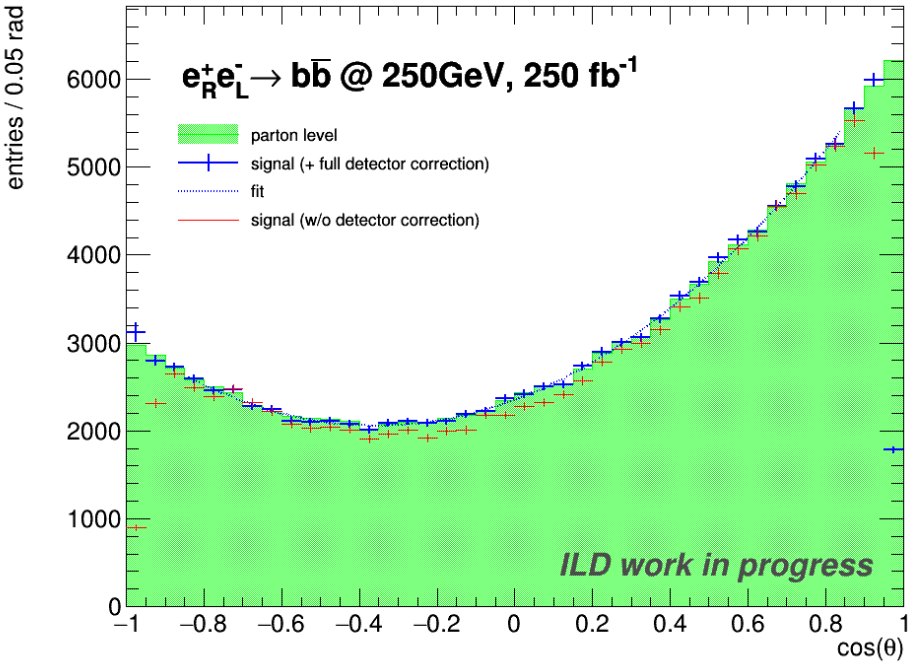
- **The chi2 is a bit small…** maybe because the correction and the fit are done with the same data.

Option a), eR


$e^+_R e^-_L \to b\bar{b}$ @ 250GeV, 250 fb^{-1}
parton level
signal (+ full detector correction)
fit
signal (w/o detector correction)
ILD work in progress
entries / 0.05 rad
cos(θ)


$e^+_R e^-_L \to b\bar{b}$ @ 250GeV, 250 fb^{-1}
X = signal + correction
X = signal w/o correction
X = parton level
ILD work in progress
(fit - X)/ σ_X
cos(θ)

Option a), eR

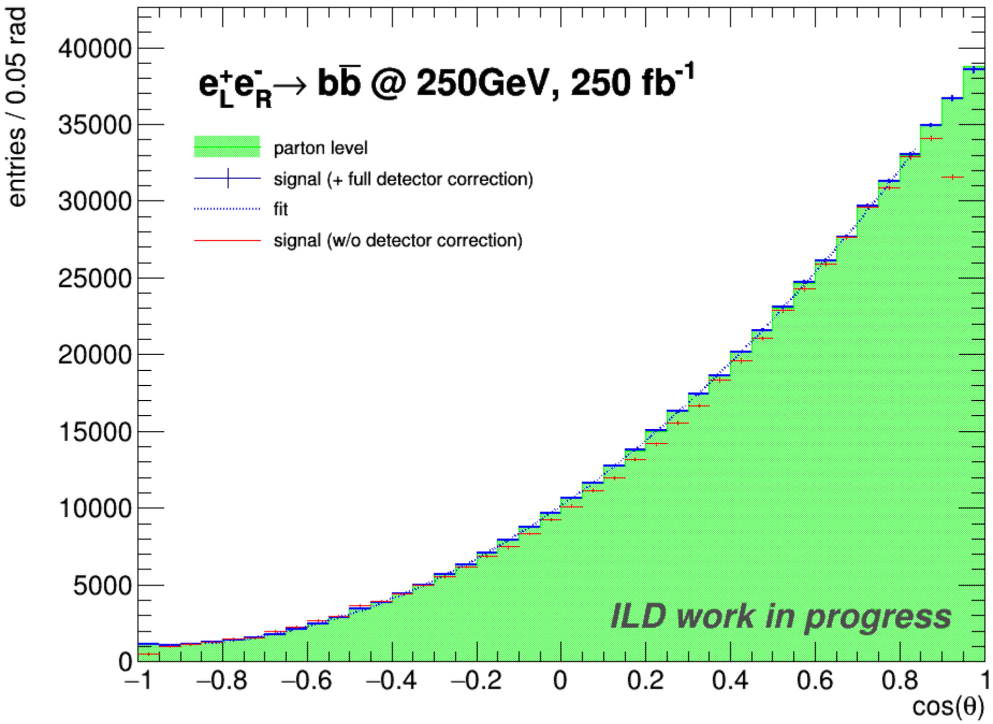


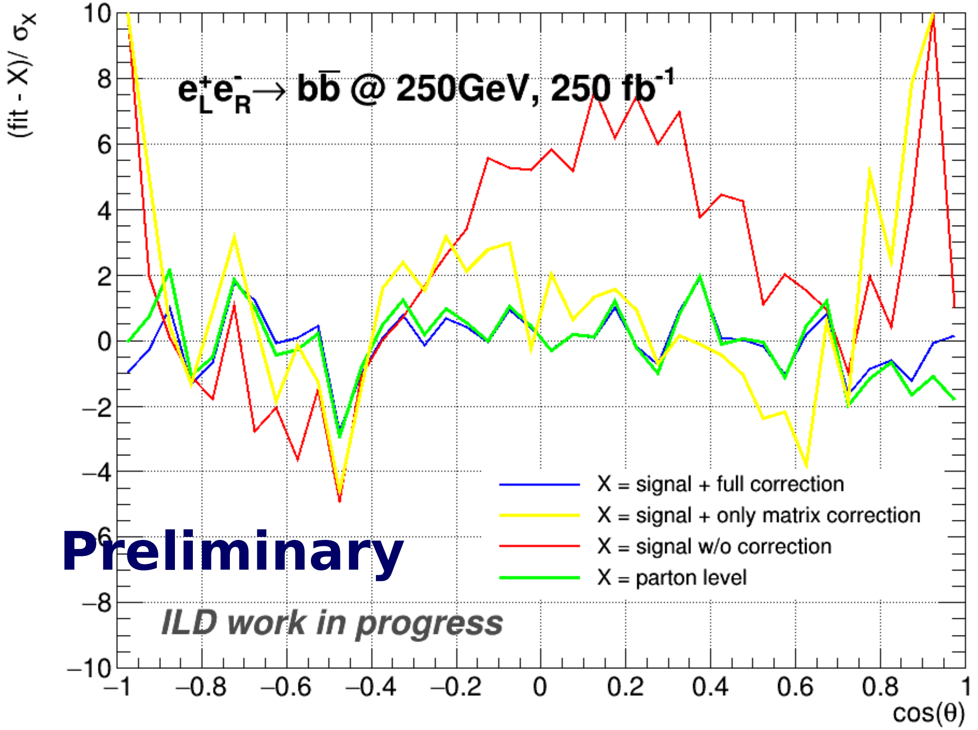

```
####### RIGHT POLARIZATION
 -------------------------
250fb-1, parton level distribution
Afb = 0.262686 +- 0.00215606  ( 0.820777 %)
S = 2354.91 +- 4.37572  ( 0.185813 %)
A = 1649.6 +- 11.6312  ( 0.705092 %)
Correl = 0.26353
Chi2 = 31.7237, NDF =32 Chi2/NDF= 0.991366

250fb-1, reco + fully corrected distribution
Afb = 0.265943 +- 0.0034225  ( 1.28693 %)
S = 2354.31 +- 7.03023  ( 0.298611 %)
A = 1669.63 +- 18.6822  ( 1.11895 %)
Correl = 0.231132
Chi2 = 24.7313, NDF =32 Chi2/NDF= 0.772854
```

Option b)

- A apply a numerical method for the matrix inversion and correction.
- Singular Value Decomposition SVU. I will show the details in a future meeting.
  - hep-ph/9509307

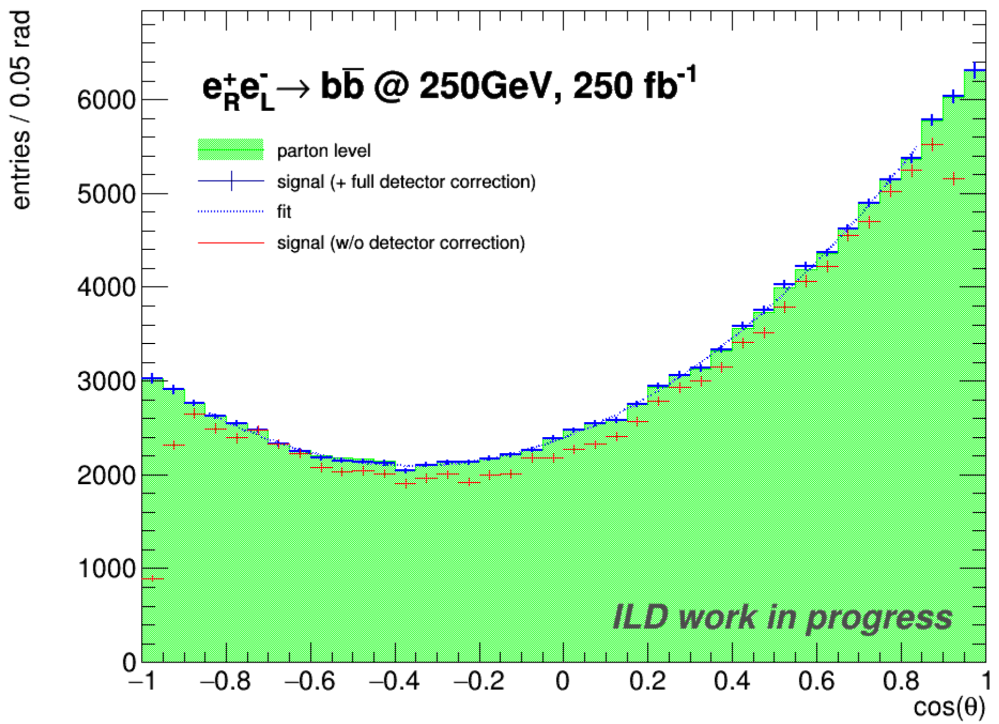
Option b), eL

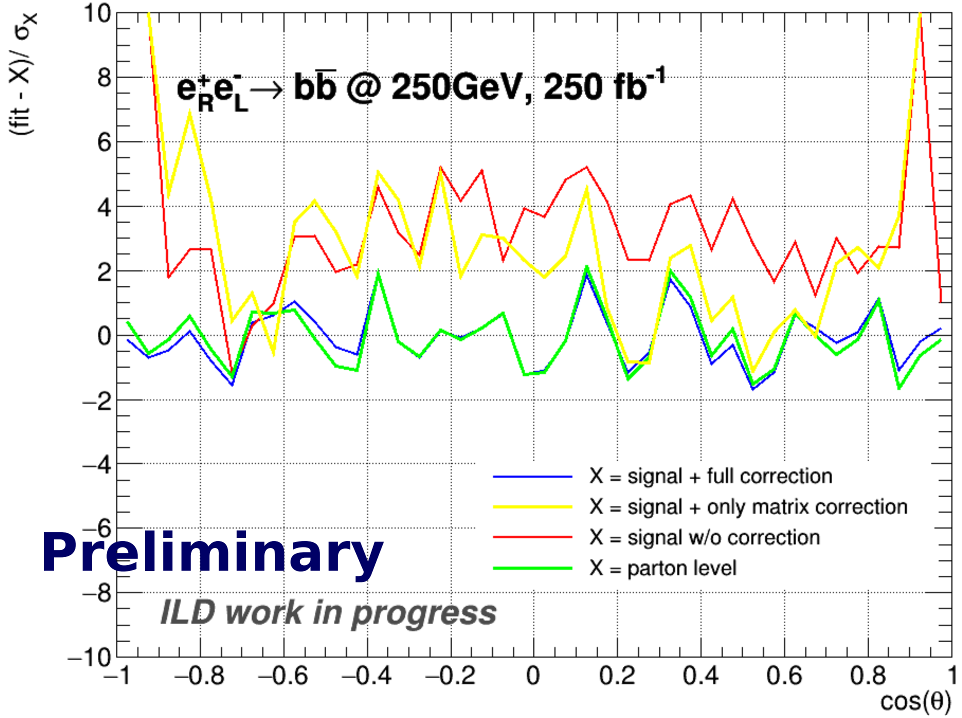




- The acceptance correction(yellow) is now half size than before (red). The resolution correction takes care of most of the correction.

Option b), eR





Preliminary

- Case a)   Case b)

Case a):

```
 ####### LEFT POLARIZATION
 ------------------------
250fb-1, parton level distribution
Afb = 0.707516 +- 0.00137244  ( 0.19398 %)
S = 10156.7 +- 9.03981  ( 0.089003 %)
A = 19162.8 +- 20.9342  ( 0.109244 %)
Correl = 0.814717
Chi2 = 35.5937, NDF =32 Chi2/NDF= 1.1123

250fb-1, reco + fully corrected distribution
Afb = 0.709079 +- 0.00216884  ( 0.305868 %)
S = 10158.2 +- 14.4238  ( 0.141992 %)
A = 19207.9 +- 35.1474  ( 0.182984 %)
Correl = 0.694899
Chi2 = 22.2375, NDF =32 Chi2/NDF= 0.694921
```

Case b):

```
250fb-1, reco + fully corrected distribution
Afb = 0.709093 +- 0.00167844  ( 0.236702 %)
S = 10167.5 +- 10.9382  ( 0.10758 %)
A = 19225.9 +- 25.3811  ( 0.132016 %)
Correl = 0.851728
Chi2 = 29.8717, NDF =32 Chi2/NDF= 0.933491
```

- Correlation gets better (closer to parton level). Chi2 closer to 1.
- Uncertainties are smaller.
- **Warning**: the matrix correction creates a covariance matrix (statistics correlations) that is still not accounted.

- Case a)

```
####### RIGHT POLARIZATION
-------------------------
250fb-1, parton level distribution
Afb = 0.262686 +- 0.00215606  ( 0.820777 %)
S = 2354.91 +- 4.37572  ( 0.185813 %)
A = 1649.6 +- 11.6312  ( 0.705092 %)
Correl = 0.26353
Chi2 = 31.7237, NDF =32 Chi2/NDF= 0.991366

250fb-1, reco + fully corrected distribution
Afb = 0.265943 +- 0.0034225  ( 1.28693 %)
S = 2354.31 +- 7.03023  ( 0.298611 %)
A = 1669.63 +- 18.6822  ( 1.11895 %)
Correl = 0.231132
Chi2 = 24.7313, NDF =32 Chi2/NDF= 0.772854
```

Case b)

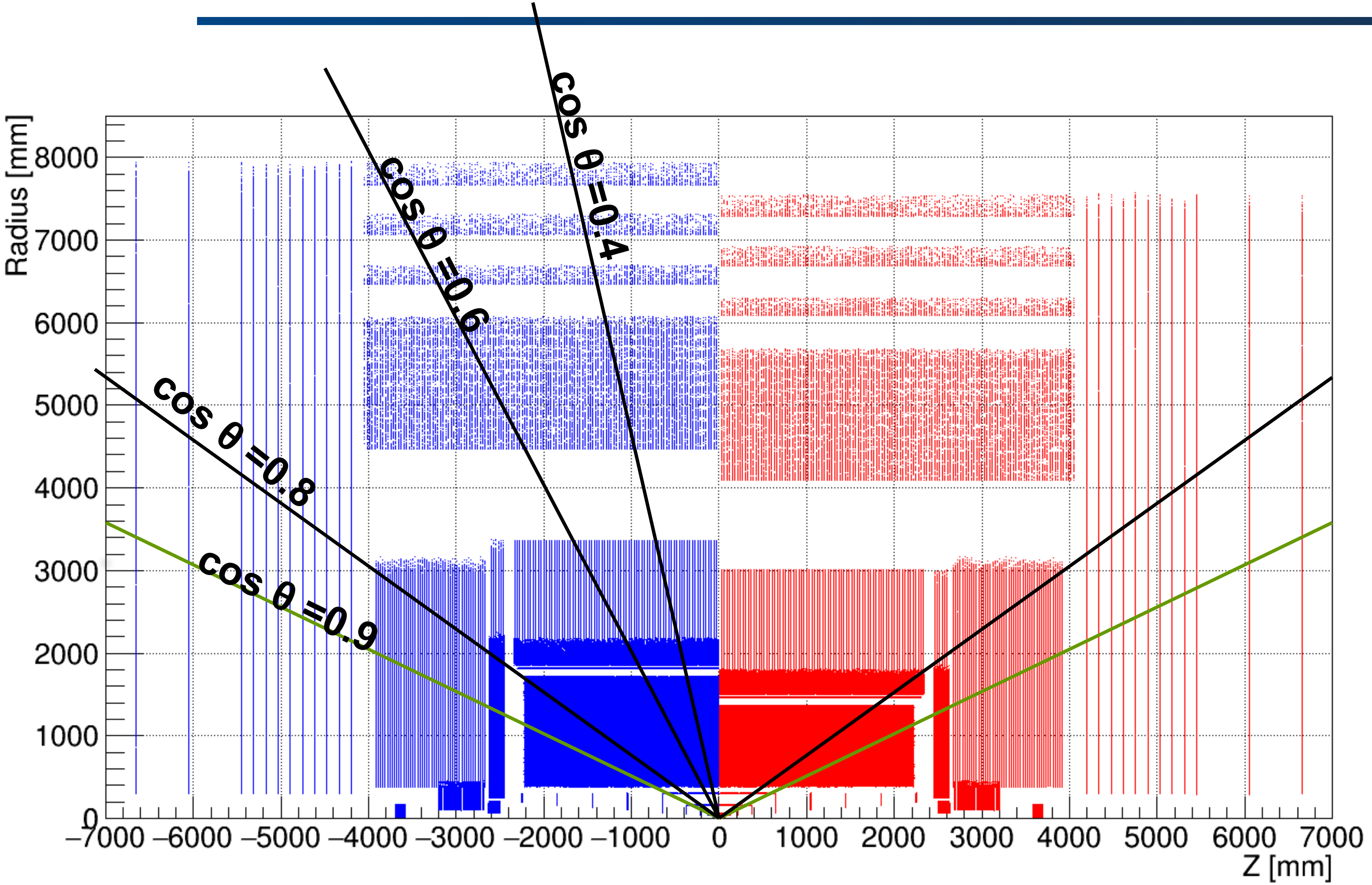
```
250fb-1, reco + fully corrected distributior
Afb = 0.265984 +- 0.00248509  ( 0.934301 %)
S = 2394.8 +- 4.99581  ( 0.20861 %)
A = 1698.61 +- 13.3305  ( 0.784789 %)
Correl = 0.324502
Chi2 = 28.6213, NDF =32 Chi2/NDF= 0.894417
```

- Correlation is a bit worst.
- Uncertainties are smaller.
- **Warning**: the matrix correction creates a covariance matrix (statistics correlations) that is still not accounted.

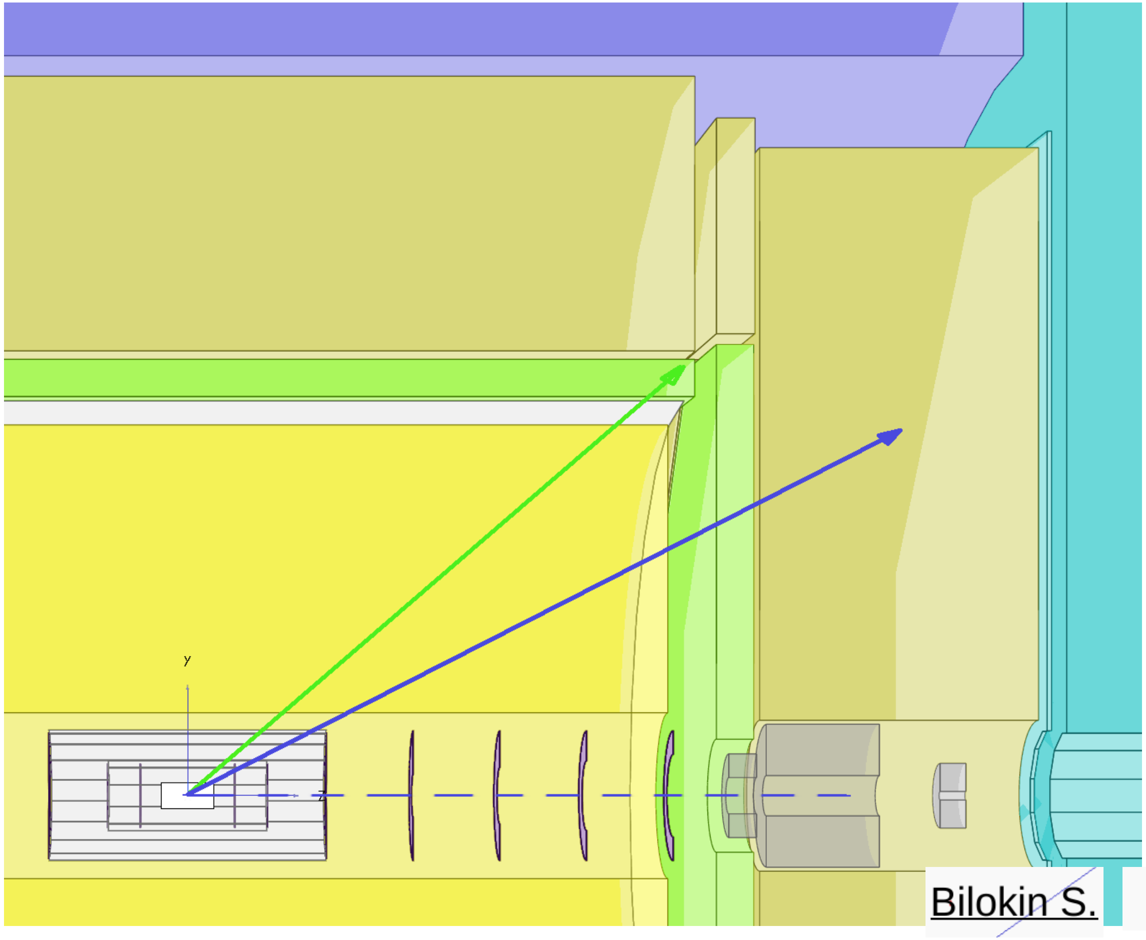
Summary

- **The analysis for 250GeV is now finished except :**
  - Add the knowledge of the covariant matrix to the fit.
  - Estimate the impact of the backgrounds.
  - More systematics?
- **I am now in position to apply all the experience to 500GeV**
  - Except that I still need some help to run LCFIPlus. After Arlington I got, from Ryo, some instructions to run it but, are they still valid? (same version? Same steering file? Same calibration files?)
  - If I run it today I can produce some results for Wednesday… what results?
- **To study the correction issue on detail and try to minimize it (by improving the jet resolution) we may need to run the analysis for 250GeV bbbar samples.**

Impact of using the vertex restorer: efficiency



ILD geometry



Bilokin S. ECFA LC Workshop 04/06/16