Study on PV and SV reconstruction

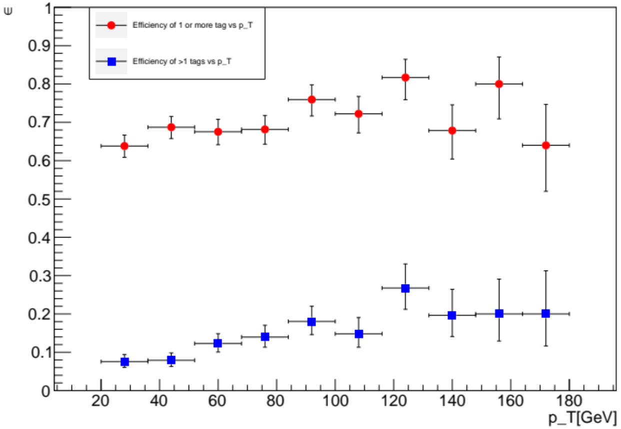
Disclaimer: I am not an expert at all, take everything with a grain of salt

Giacomo Da Molin

Starting motivations

The aim of my thesis is to train a DNN to perform b- and c-jet tagging and apply it to extract the yield of H → bb and H → cc.
In order to do this I performed a preliminary study on the characteristics of such jets and how many can be tagged by a SV. In this study I found out that many jets, contrary to what we expected, are tagged by more than 1 SV. As we believe that this should not happen, so I am performing a study to understand why so there are so many SVs.

Tagging efficiency with 1 (red dots) or more SVs (blue dots)



In the Ntuple too many SVs →

Many jets tagged by more SV, of order 10%

In these presentations I will show data from 2 samples:
*H → bb 3TeV no BIB
*H → bb 3TeV with BIB

Primary vertex study

Let's check if the primary vertex is badly reconstructed as this can impact the rest of the event. SV processor requires PV, and it's used in selecting the pool of the track from which SV can be built, to check if SV is too close to PV (discarded if so), to associateIPTracks.

Primary Vertex collection Tuple

Starts from 3-D point in center of contrain region, add every track with χ^2 fit, then tear-down algorithm: starts from highest χ^2 track and removes it, recomputes everything and then proceed to new highest χ^2 tracks until no tracks have χ^2 above threshold.

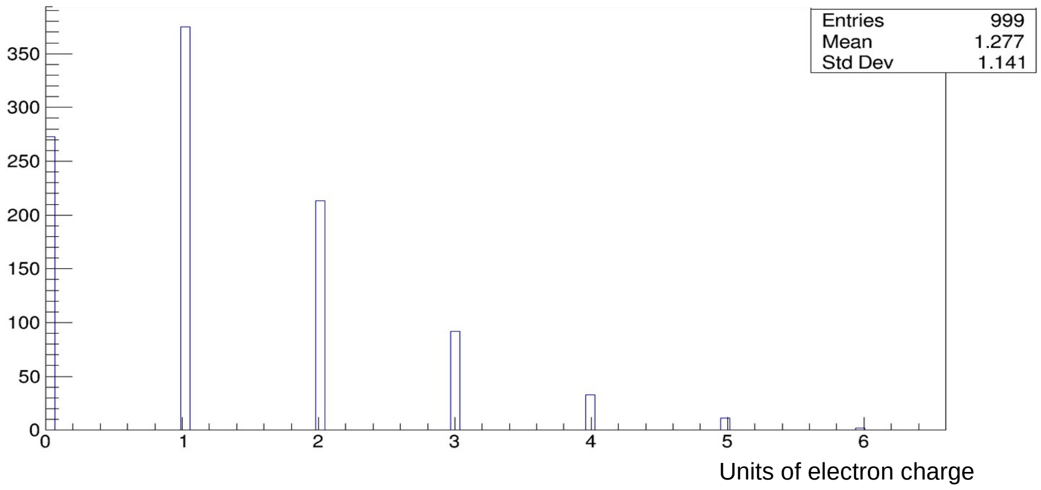
Not many parameters, and except smearing, all on the tracks:

```xml
<parameter name="PrimaryVertexFinder.BeamspotSmearing" type="boolean" value="false" />
<parameter name="PrimaryVertexFinder.TrackMaxD0" type="double" value="0.1" />
<parameter name="PrimaryVertexFinder.TrackMaxZ0" type="double" value="0.1" />
<parameter name="PrimaryVertexFinder.TrackMinVtxFtdHits" type="int" value="4" />
<parameter name="PrimaryVertexFinder.Chi2Threshold" type="double" value="10." />
```

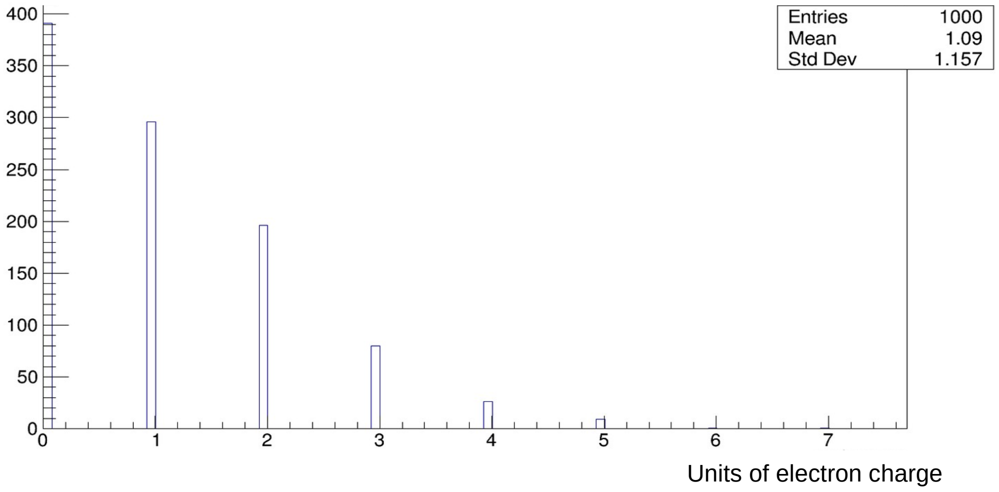
Primary vertex study

Let’s study some variables of the PV

Absolute value of sum of charges of tracks building the PV, no BIB

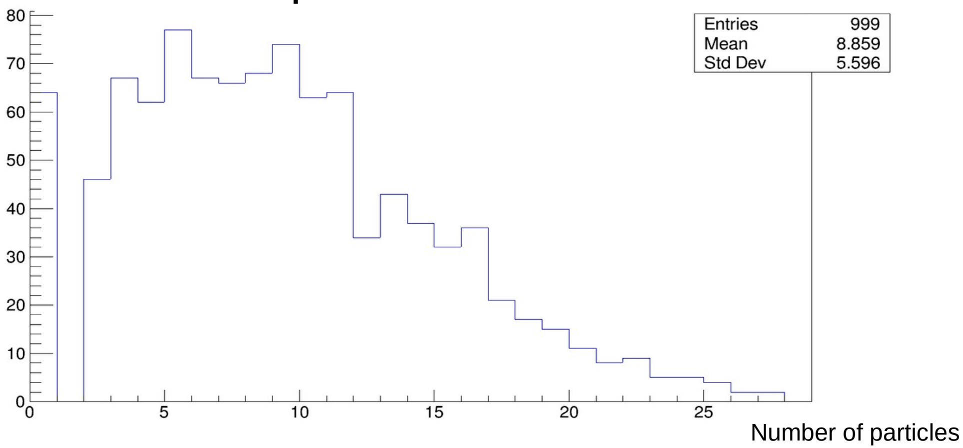


Absolute value of sum of charges of tracks building the PV, with BIB

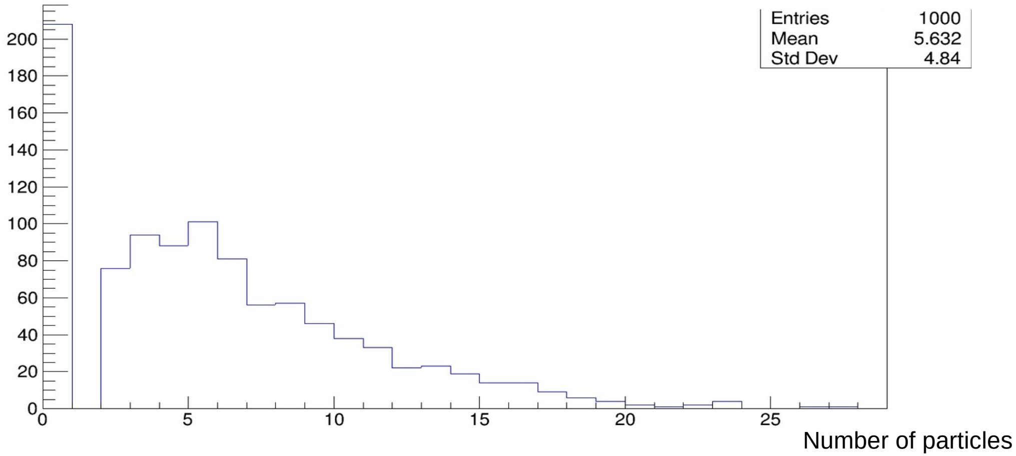


Primary vertex study

Number of particles used for PV, no BIB



Number of particles used for PV, with BIB



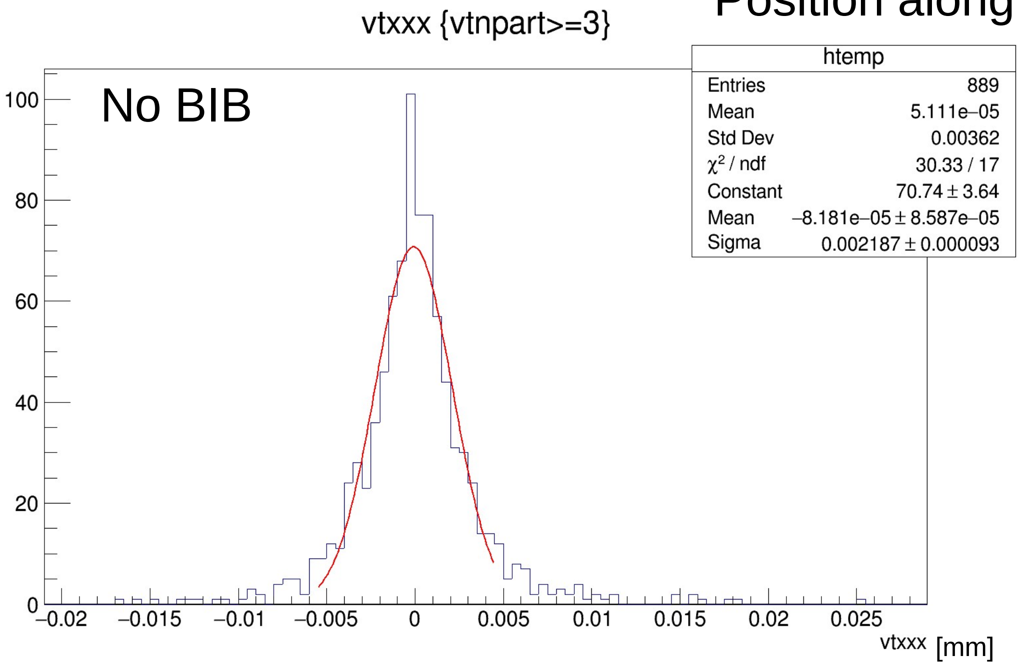
Without at least 3 tracks in the PV, there is no way PV can be well reconstructed and this will have consequences on the rest of the events.

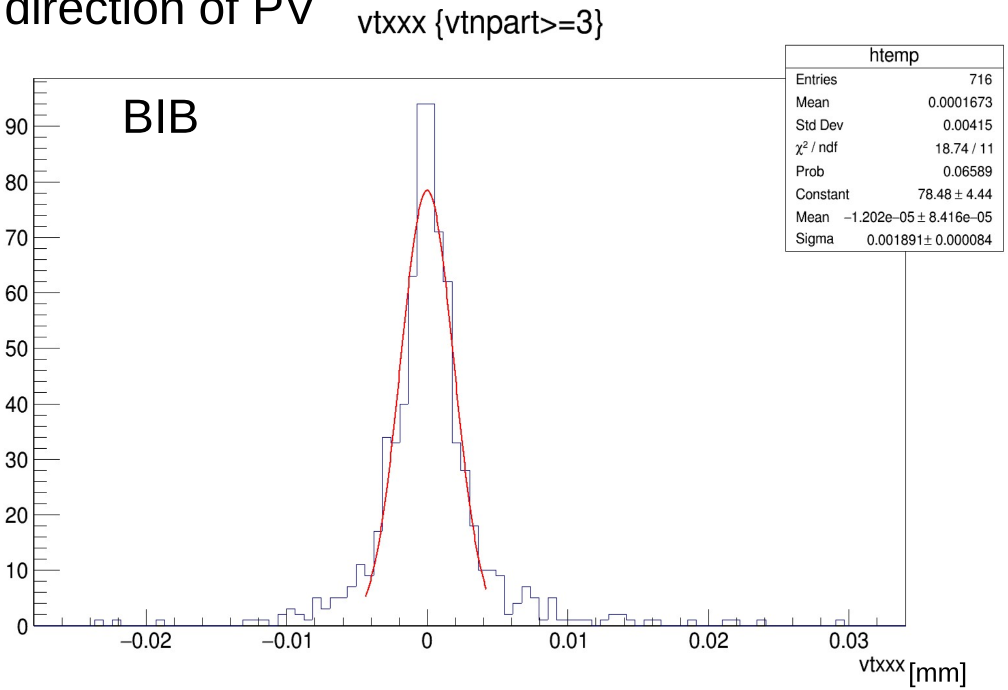
From now on we will only keep events in which the PV has at least 3 tracks associated.

How this affects the SV tagging efficiencies can be seen in the backup.

Primary vertex study

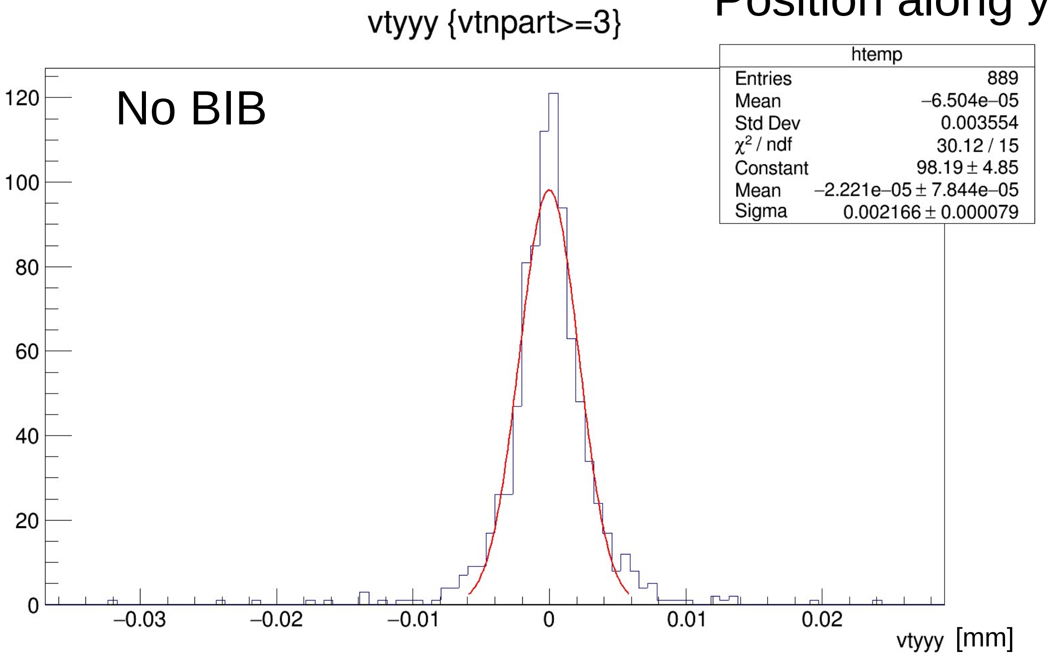
Position along x direction of PV

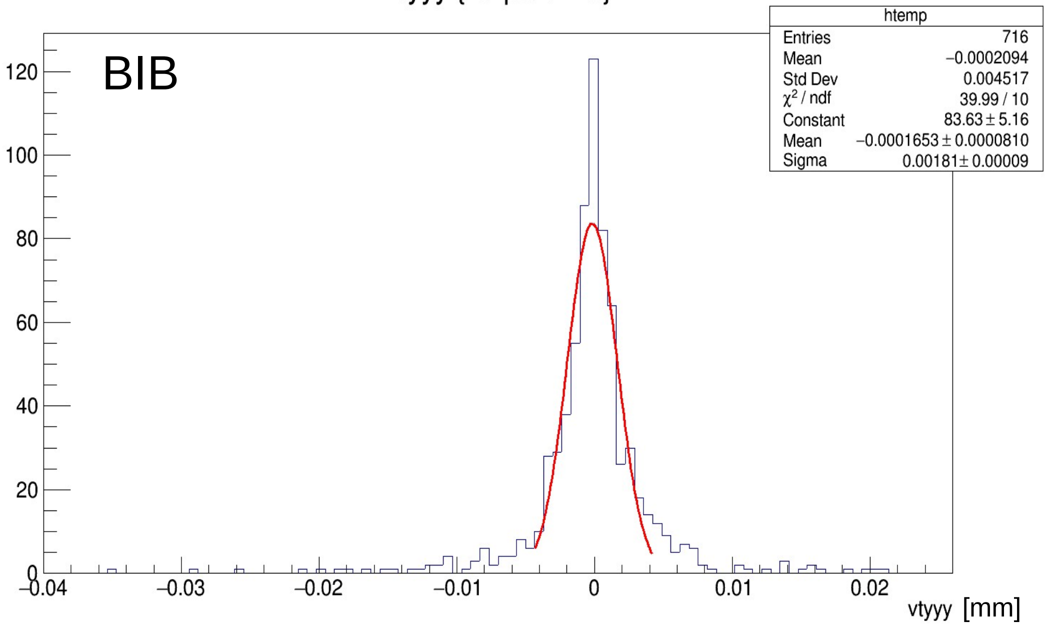




Primary vertex study

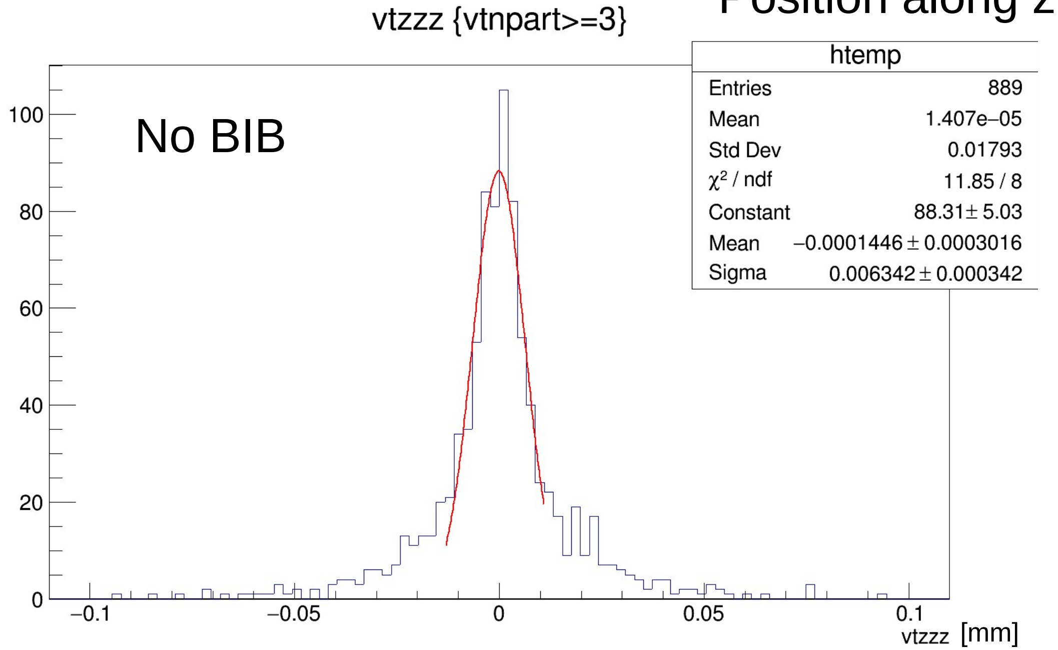
Position along y direction of PV

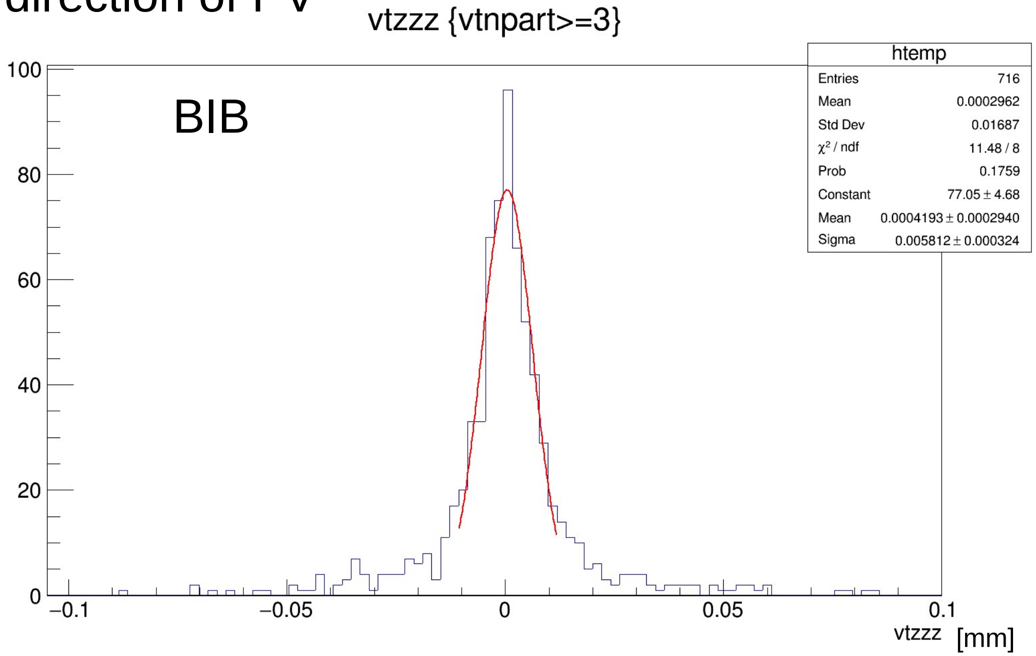




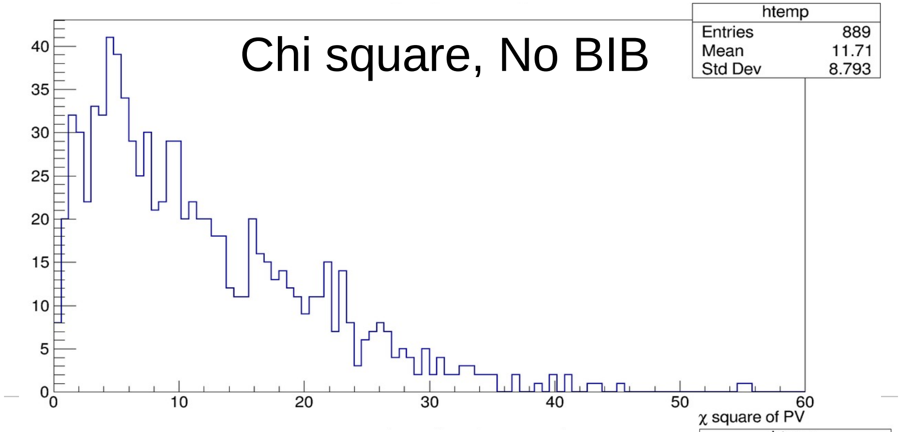
Primary vertex study

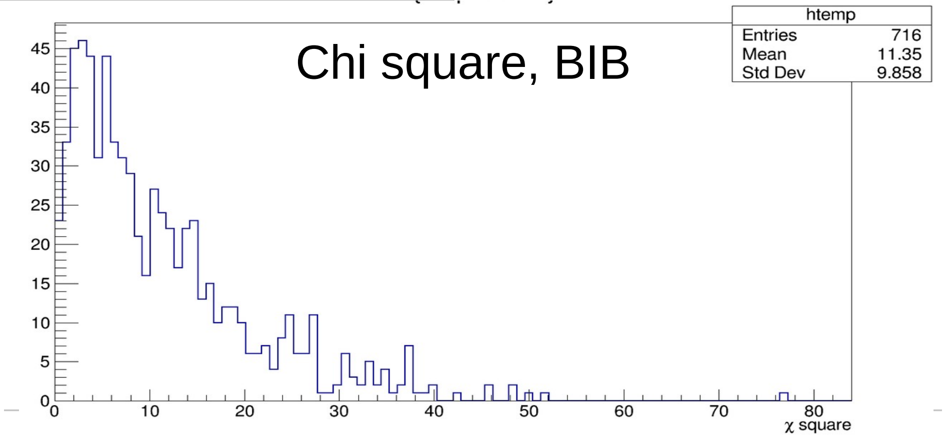
Position along z direction of PV

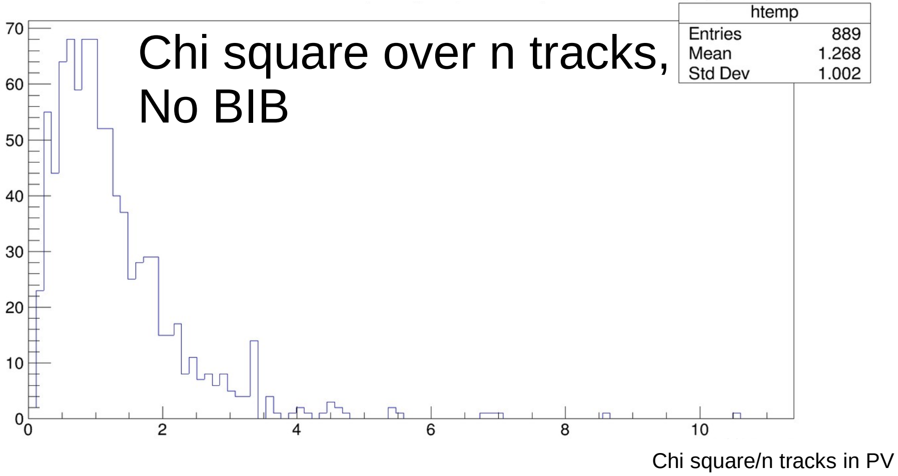


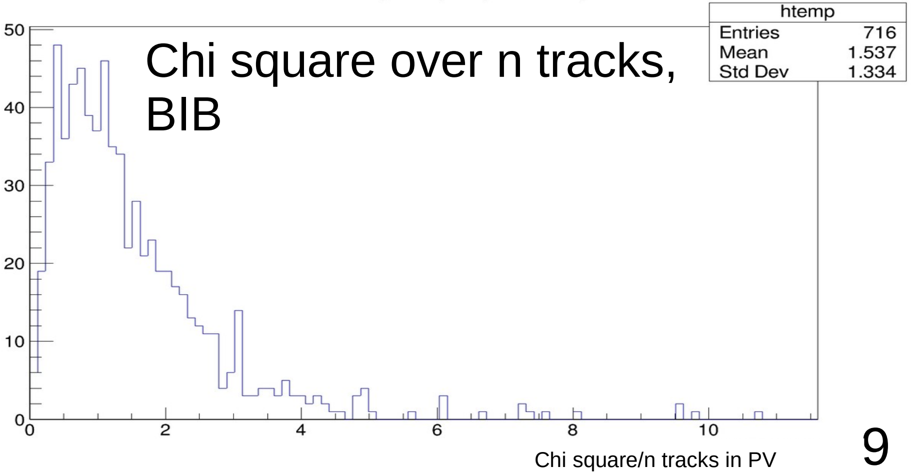


Primary vertex study









Primary vertex study

The distributions of positions and of chi square of the PV after imposing at least 3 tracks to be associated to it is compatible to what we expect, so this is a good cut and will be kept for the rest of the analysis.

Still, it will be important to improve the PV reconstruction in the future, because this cut imply a significant loss in statistics (>25% with BIB).

Secondary Vertex study

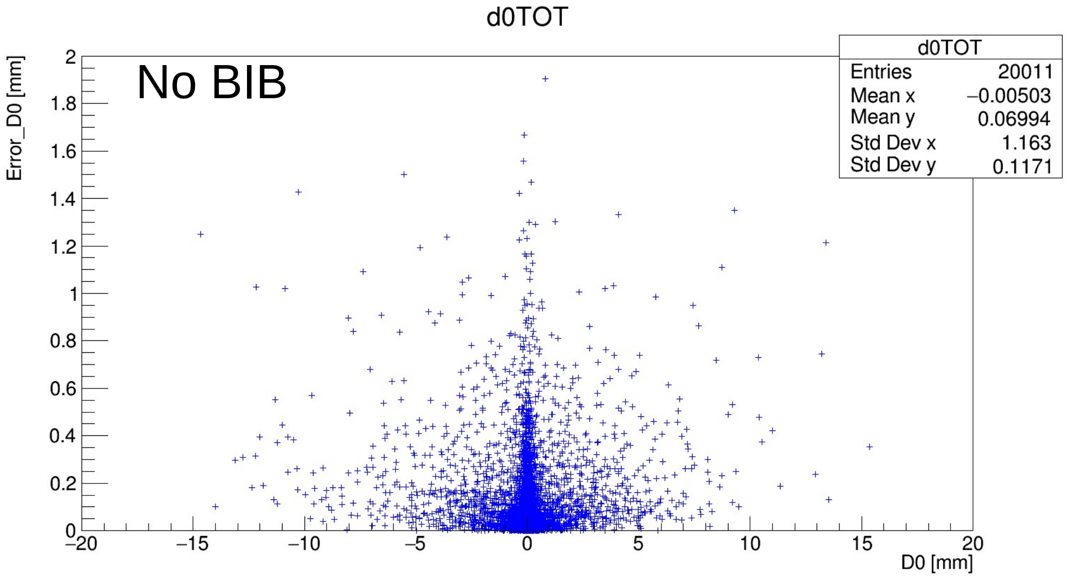
Many more parameters, must perform tracks parameters study. Although by changing some values of the AssocIP things seem to slightly improve.

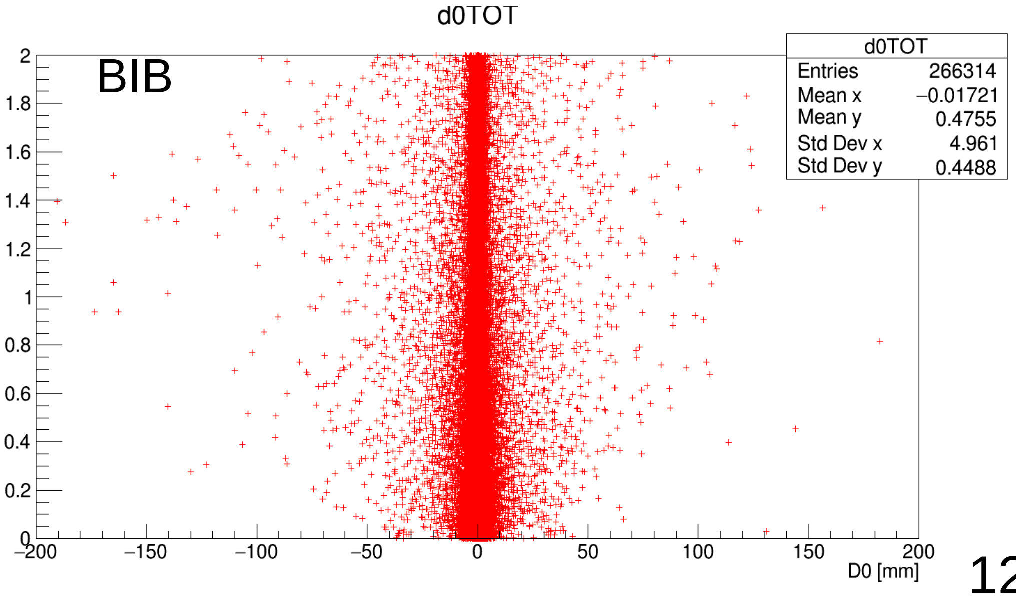
```
<parameter name="BuildUpVertex.TrackMaxD0" type="double" value="5" />
<parameter name="BuildUpVertex.TrackMaxZ0" type="double" value="5" />
<parameter name="BuildUpVertex.TrackMinD0Z0Sig" type="double" value="2" />
<parameter name="BuildUpVertex.TrackMinPt" type="double" value="0.8" />
<parameter name="BuildUpVertex.TrackMaxD0Err" type="double" value="1" />
<parameter name="BuildUpVertex.TrackMaxZ0Err" type="double" value="1" />
<parameter name="BuildUpVertex.TrackMinVxdFtdHits" type="int" value="4" />
<parameter name="BuildUpVertex.PrimaryChi2Threshold" type="double" value="10." />
<parameter name="BuildUpVertex.SecondaryChi2Threshold" type="double" value="5." />
<parameter name="BuildUpVertex.MassThreshold" type="double" value="10." />
<parameter name="BuildUpVertex.MinDistFromIP" type="double" value="0.3" /> TO BE REMOVED
<parameter name="BuildUpVertex.MaxChi2ForDistOrder" type="double" value="1." /> USELESS
<parameter name="BuildUpVertex.AssocIPTracks" type="int" value="1" />
<parameter name="BuildUpVertex.AssocIPTracksMinDist" type="double" value="0." />
<parameter name="BuildUpVertex.AssocIPTracksChi2RatioSecToPri" type="double" value="2" />
<parameter name="BuildUpVertex.UseV0Selection" type="int" value="1" />
<!-- AVF -->
<parameter name="BuildUpVertex.UseAVF" type="boolean" value="true" />
<parameter name="BuildUpVertex.AVFTemperature" type="double" value="1.0" /> AVF USELESS
```

Blue= single tracks parameters
Red= SV parameters
Green= algorithm that moves tracks from PV to SV and parameters

Tracks parameters study

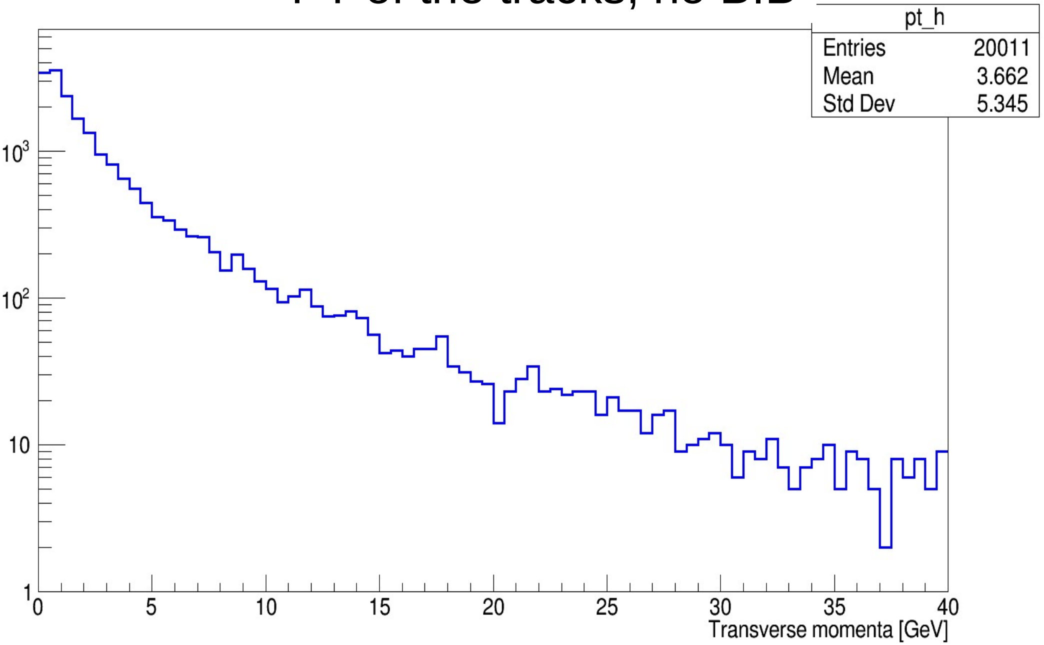
Note: different event samples dimensions for not BIB (blue) vs BIB (red), both for cuts on PV and different dimensions of initial samples (999 vs 848). After PV cuts using 889 events without BIB and 614 with BIB



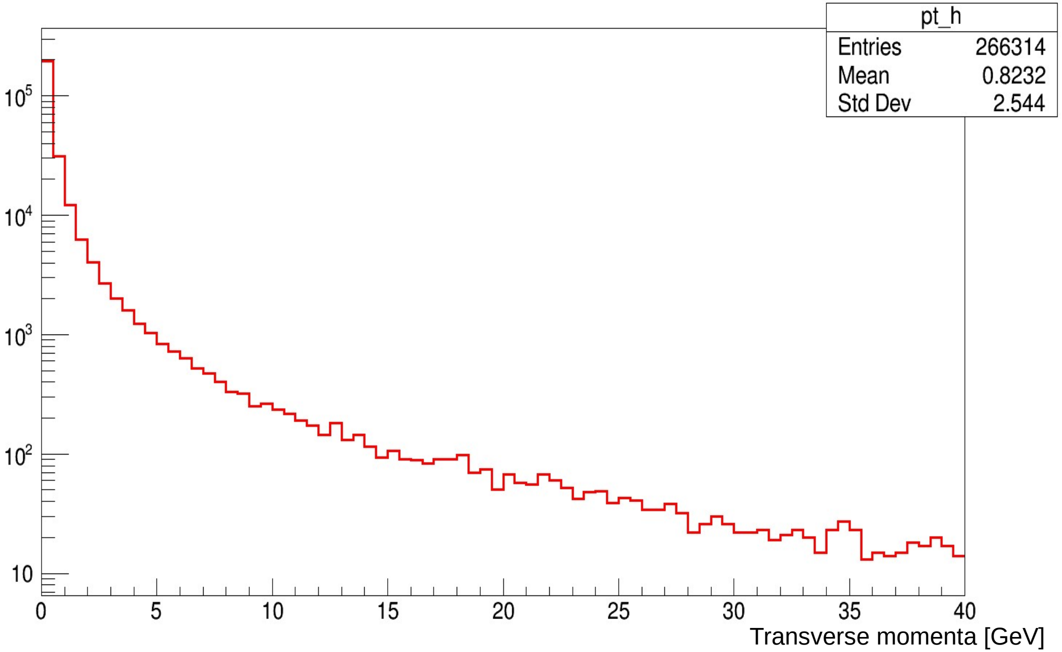


Tracks parameters study

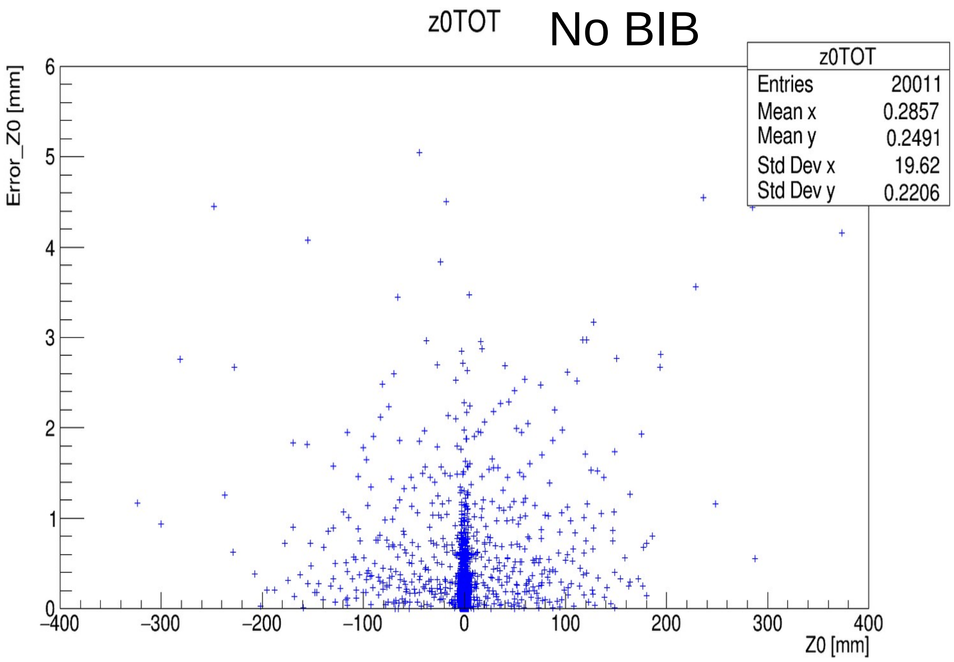
PT of the tracks, no BIB

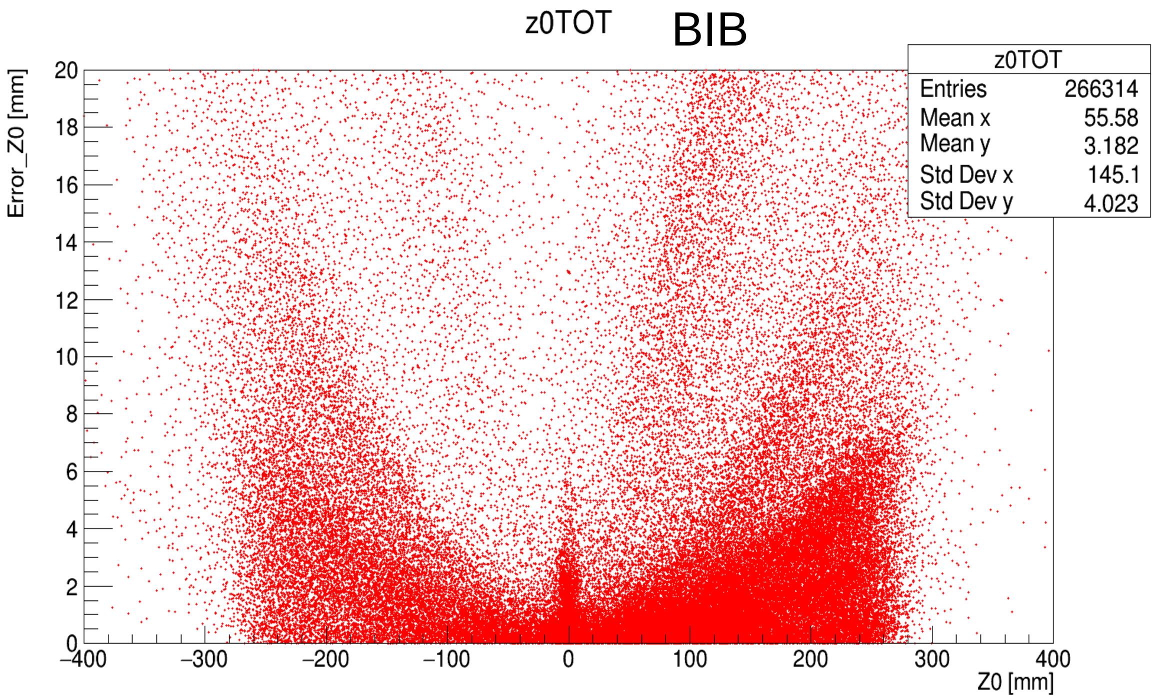


PT of the tracks, with BIB



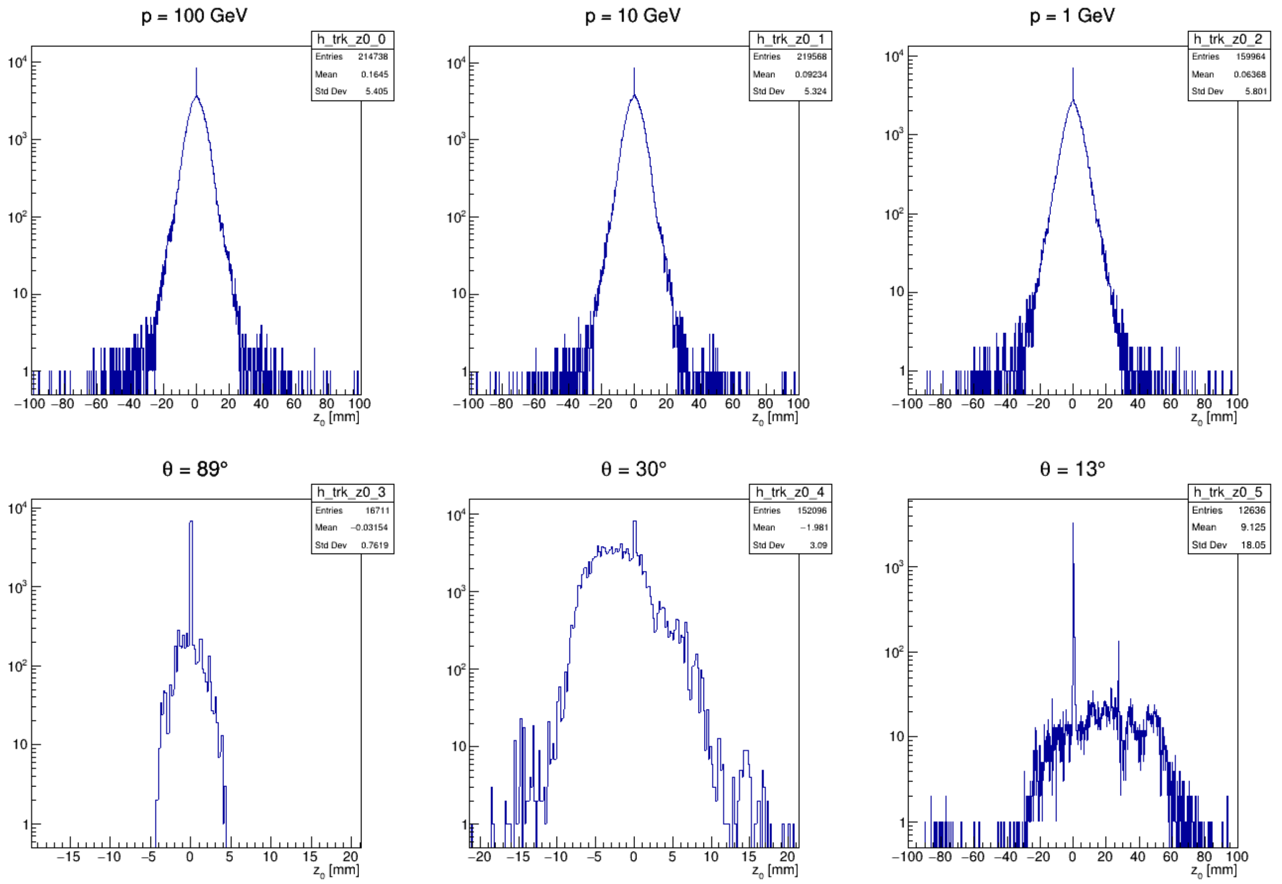
Tracks parameters study





Tracks parameters study

Thanks to Massimo Casarsa

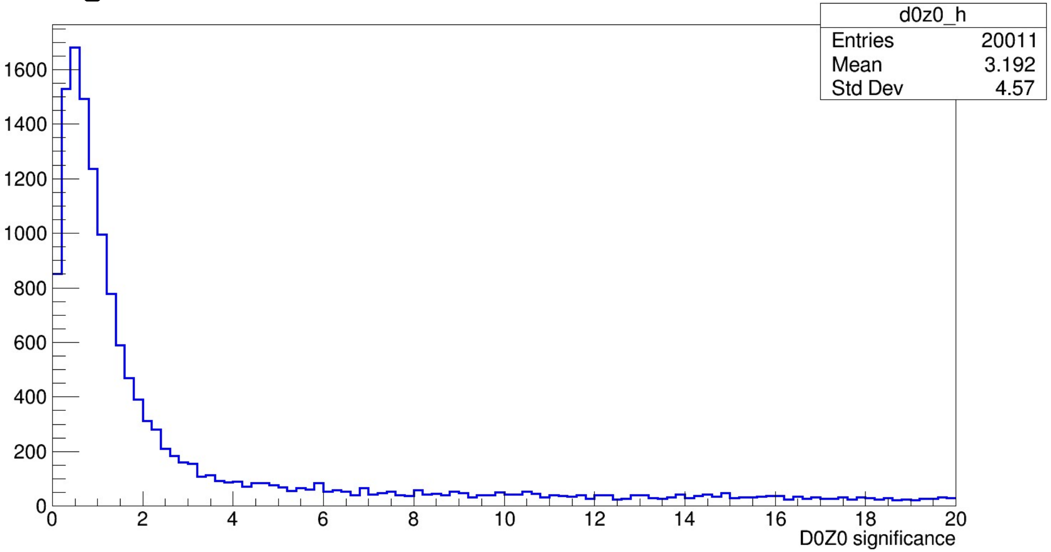


Z0 distributions obtained from a muon gun + BIB

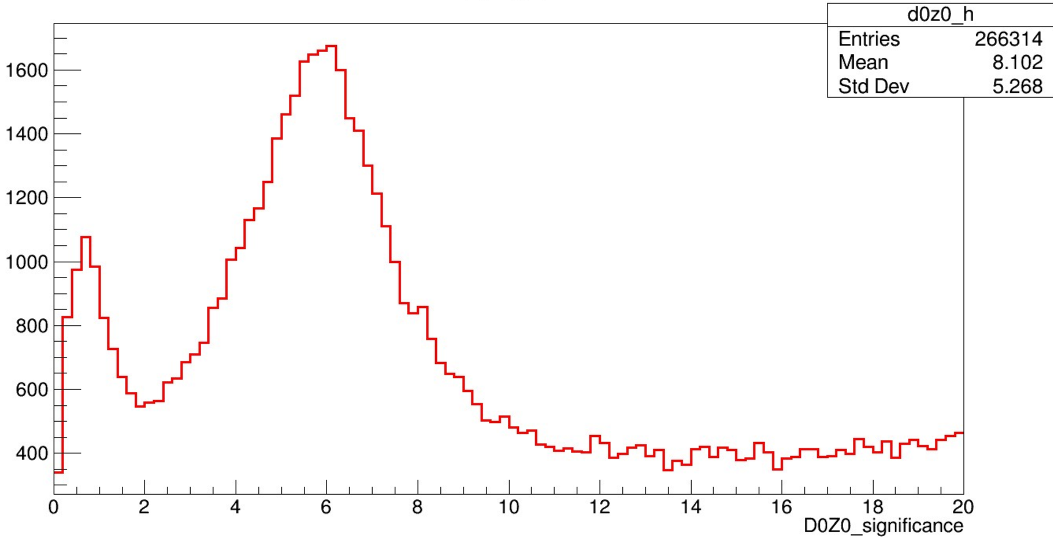
Asimmetry of Z0 is indipendent of p, but has strong dependence on the angle

Tracks parameters study

Significance of D0 and Z0 of the tracks, no BIB

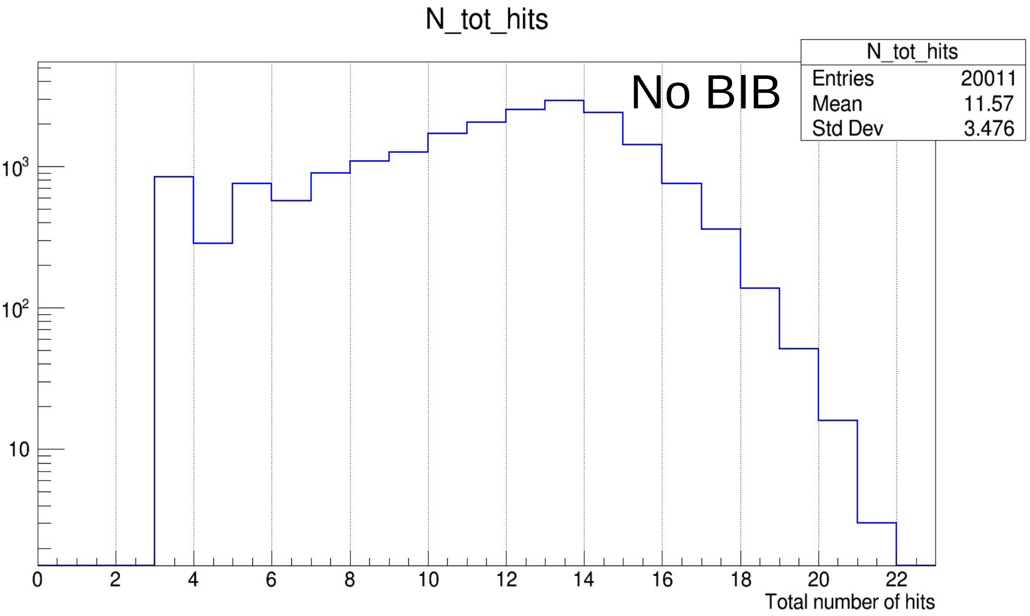


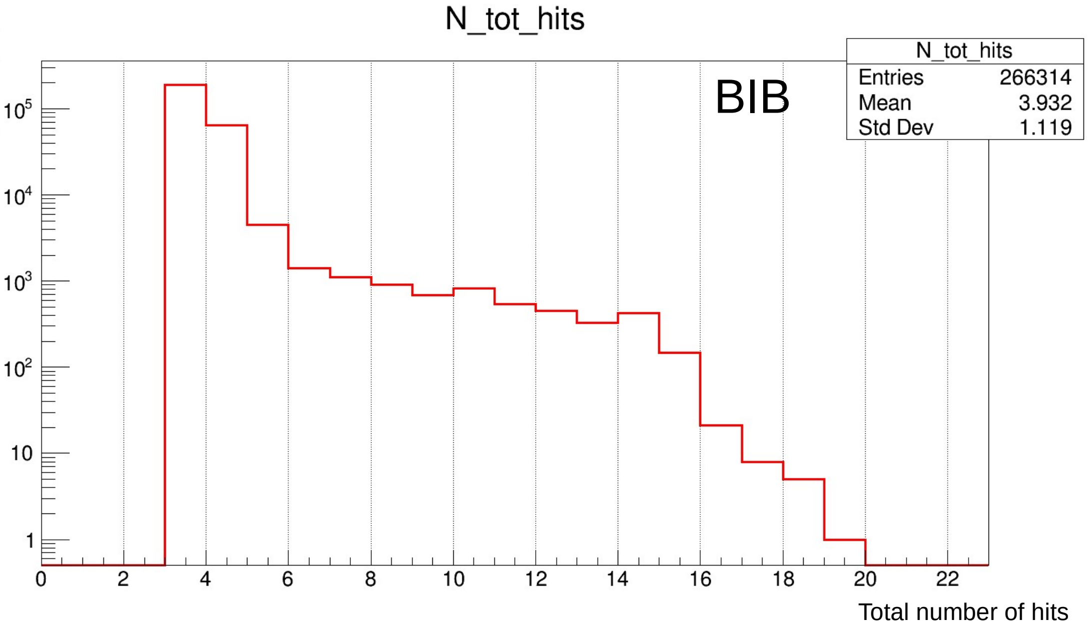
Significance of D0 and Z0 of the tracks, BIB



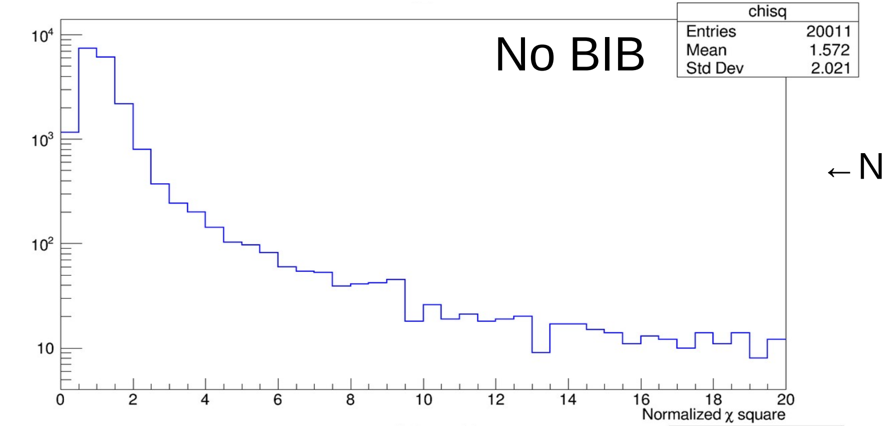
With standard cuts, tracks with significance >2 are used to build SVs

Tracks parameters study

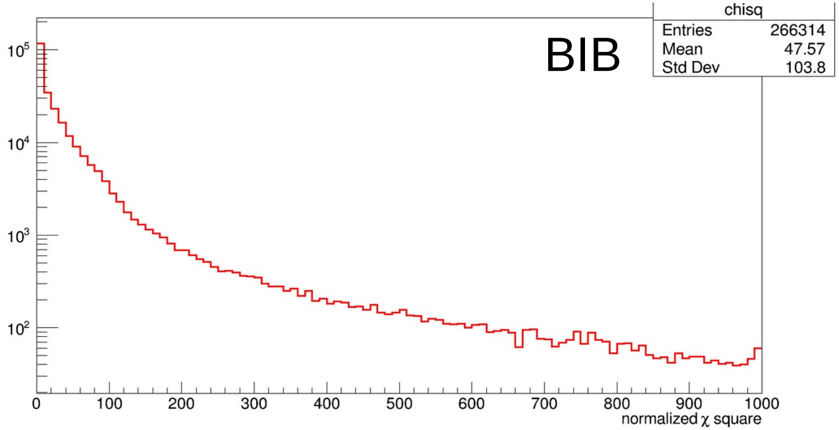


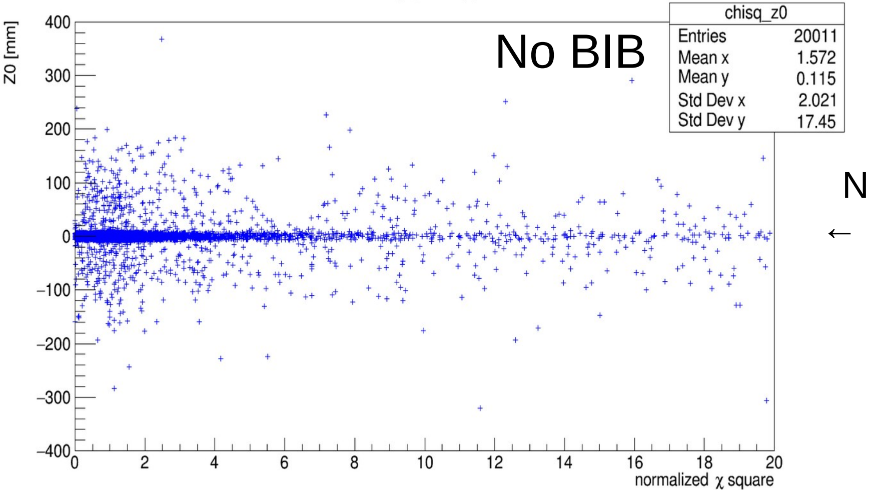


Tracks parameters study

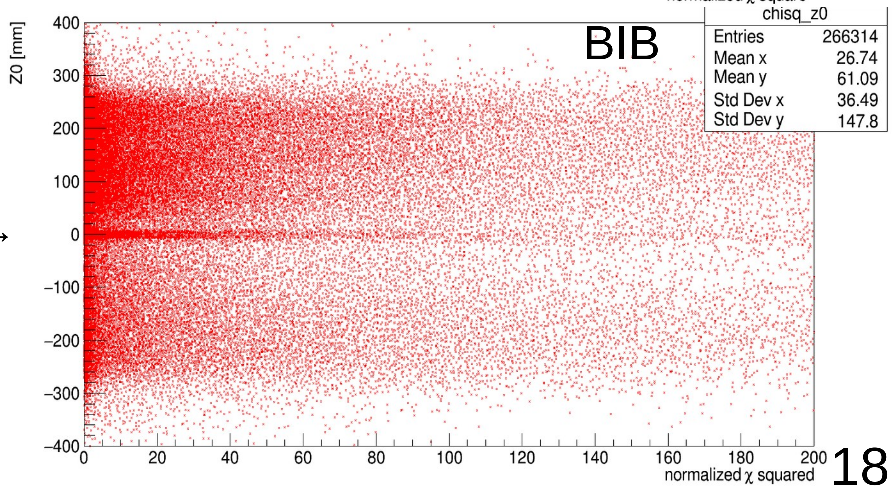


← Normalized chi square →





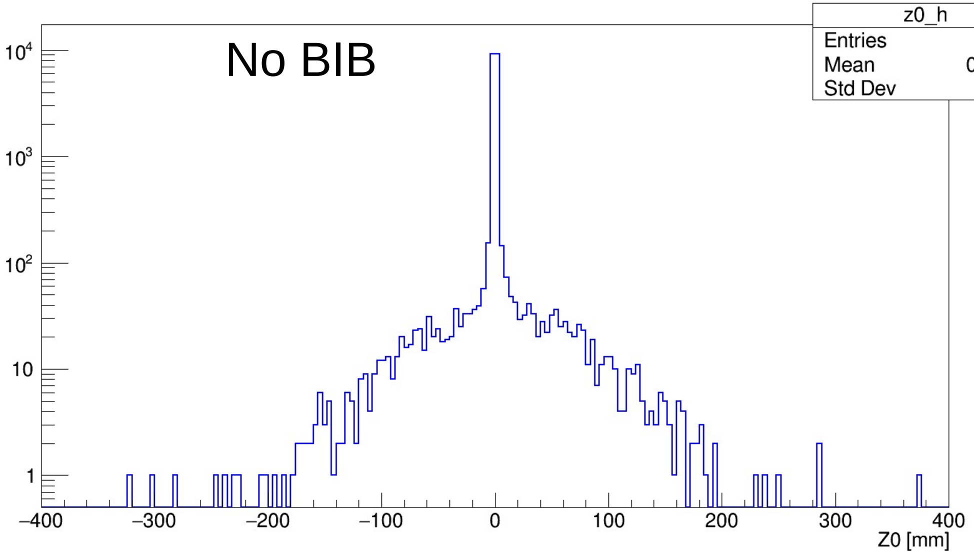
← Normalized chi square vs Z0 →

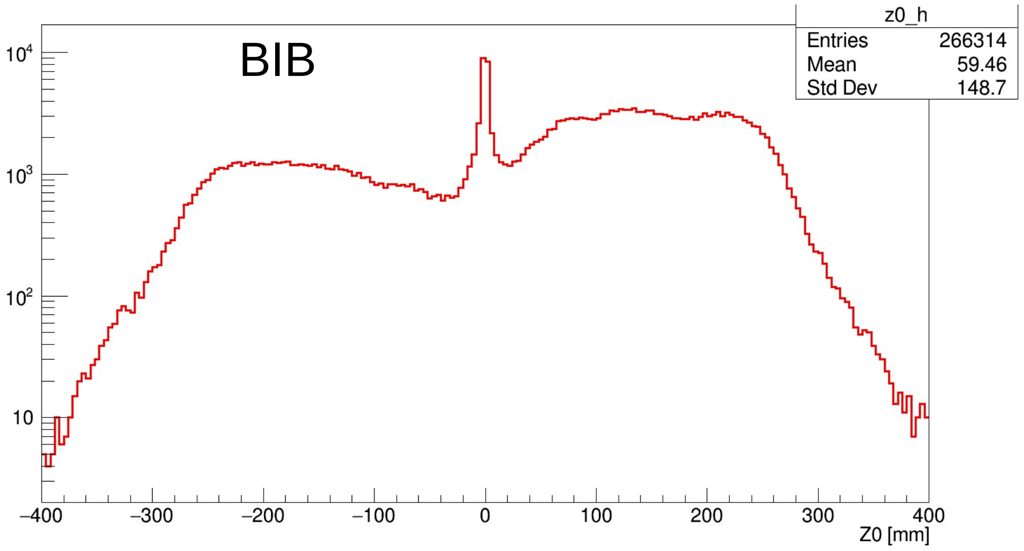


Tracks parameters study

We will now show how Z0 distributions vary imposing several cuts

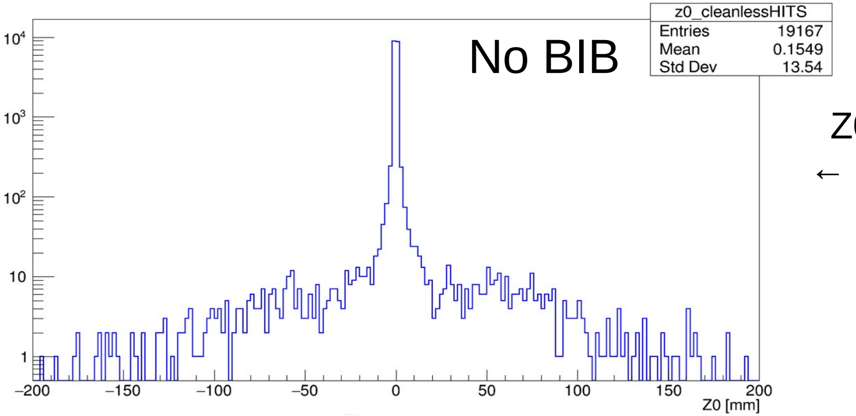
Z0 distributions before any cuts (except discarding events with badly reconstructed PV)



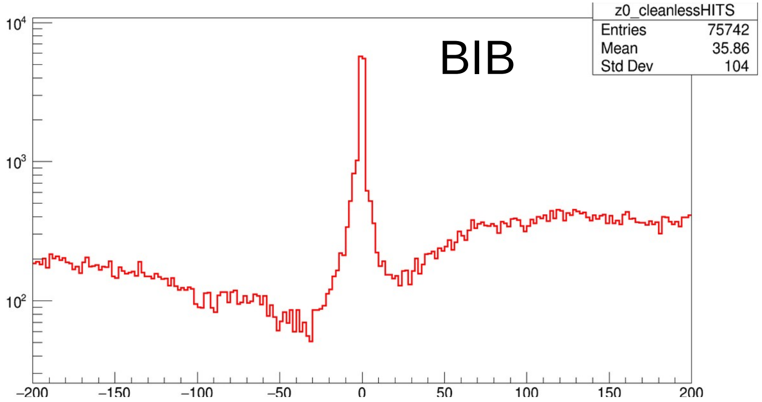


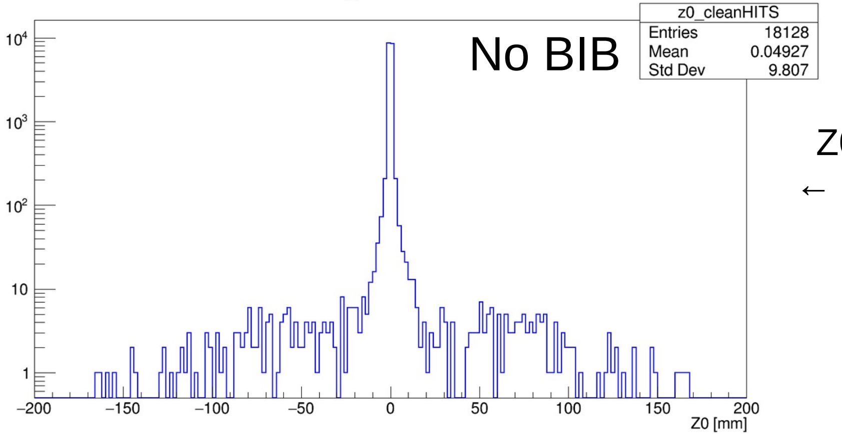
Tracks parameters study

With old cuts on total hits (upper figures, at least 4 hits) and new (bottom figures, at least 6 hits)

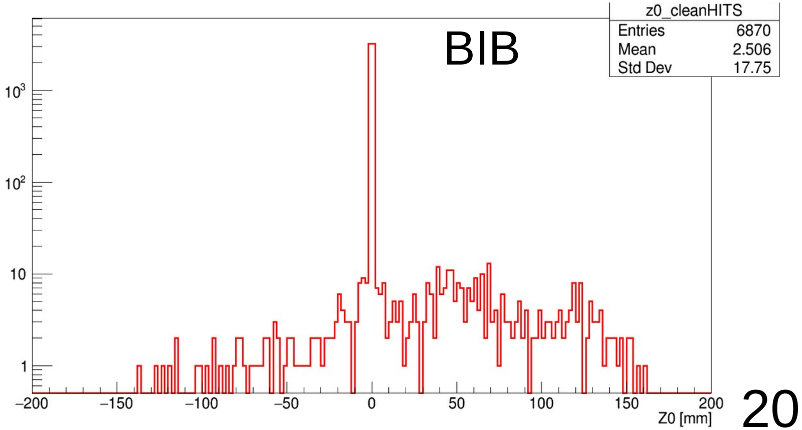


Z0 distributions of tracks
← with at least 4 hits →



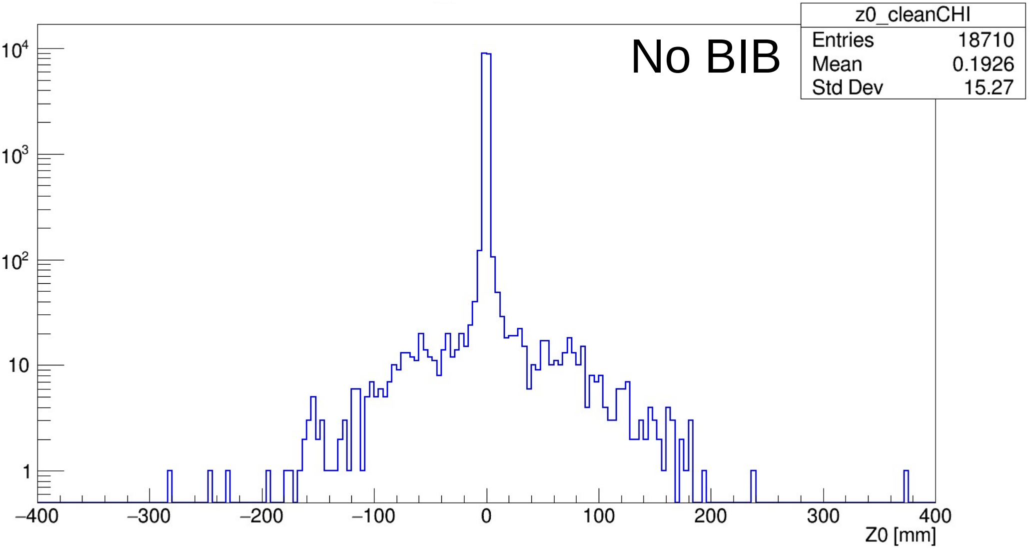


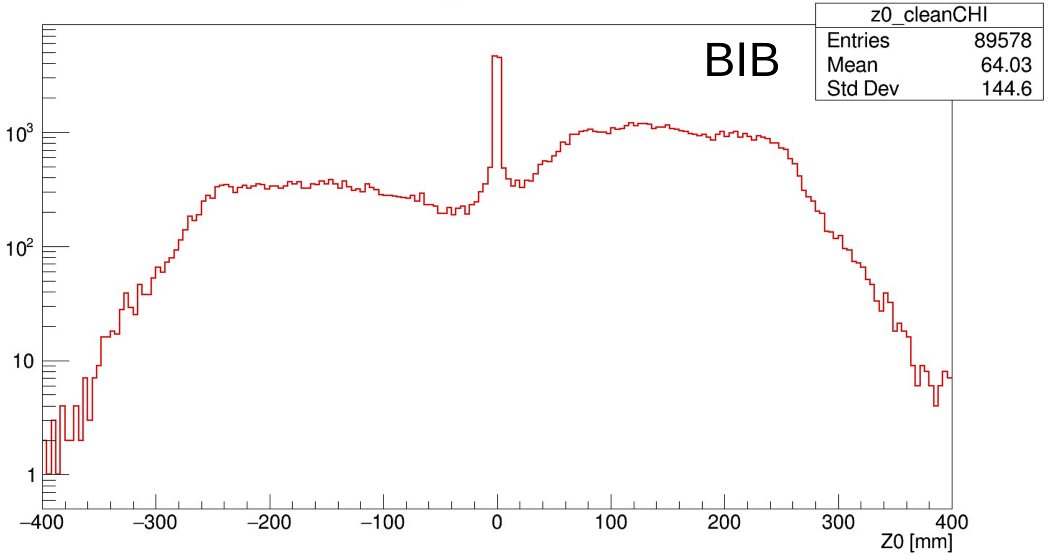
Z0 distributions of tracks
← with at least 6 hits →



Tracks parameters study

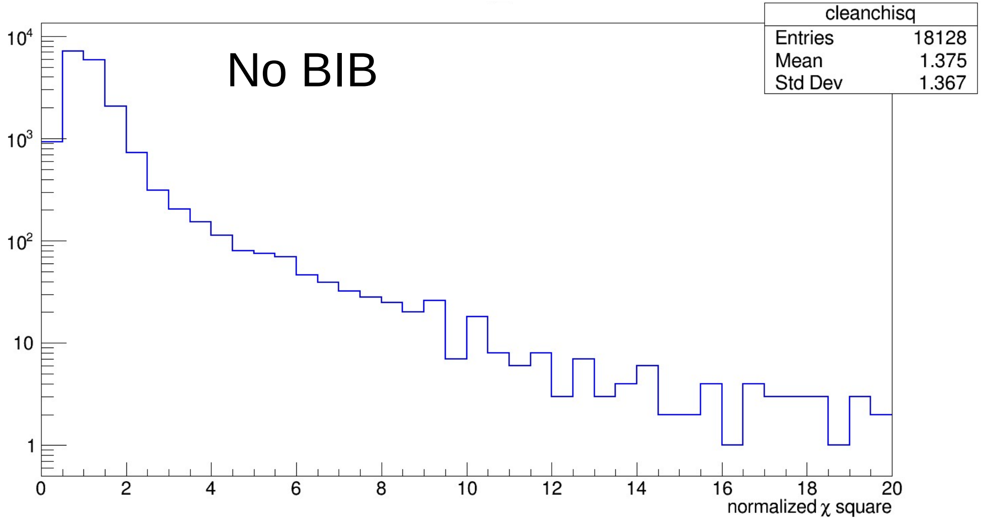
Z0 distributions with ONLY normalized chisquare required < 5

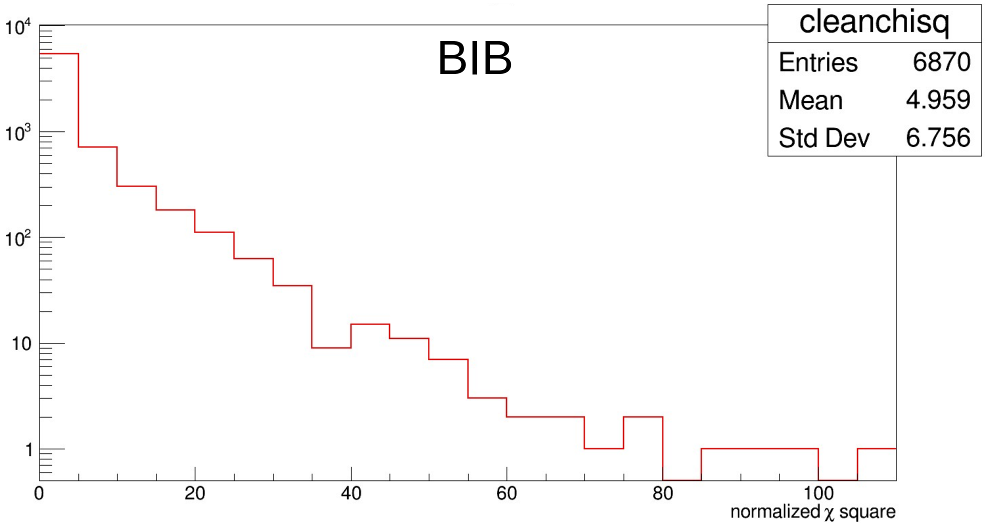




Tracks parameters study

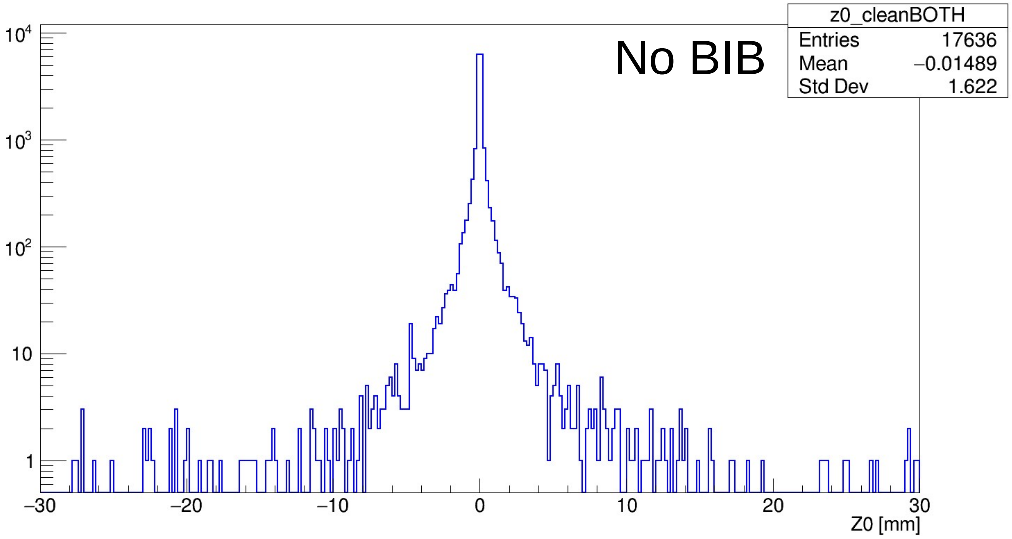
Chi square distributions of tracks with at least 6 hits

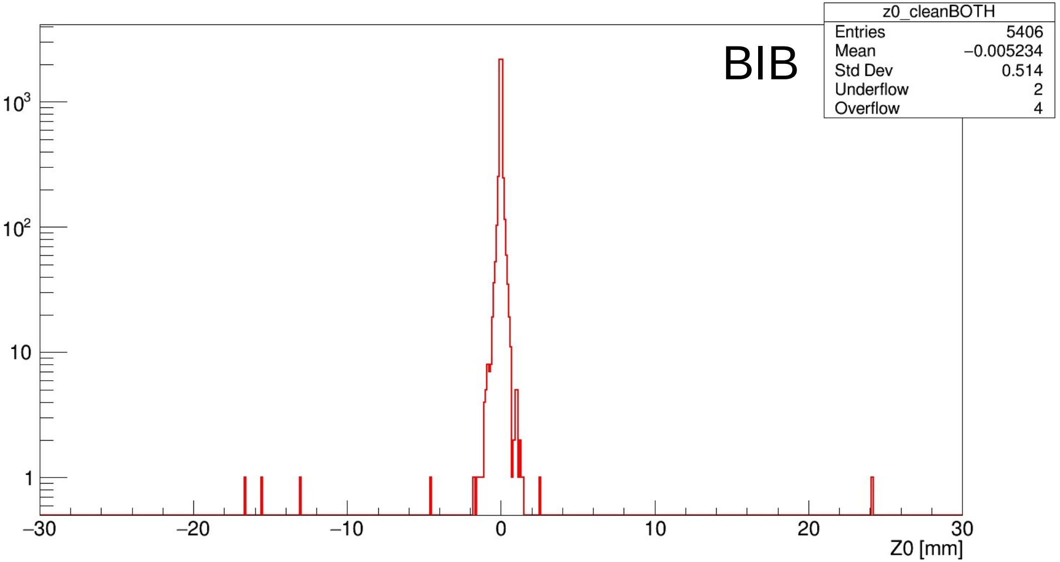




Tracks parameters study

Z0 distributions with tracks BOTH >= 6 hits chisquare and normalized chisquare< 5

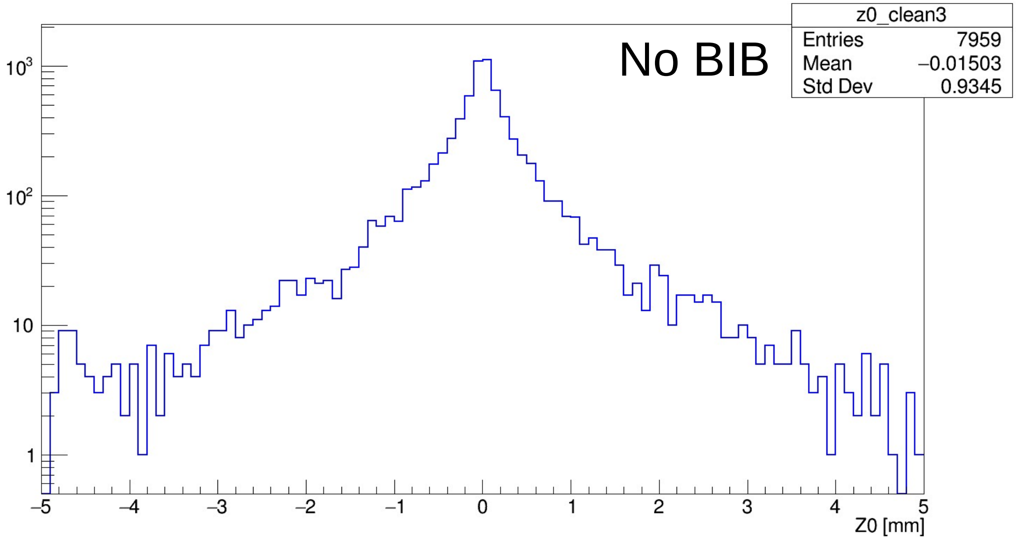


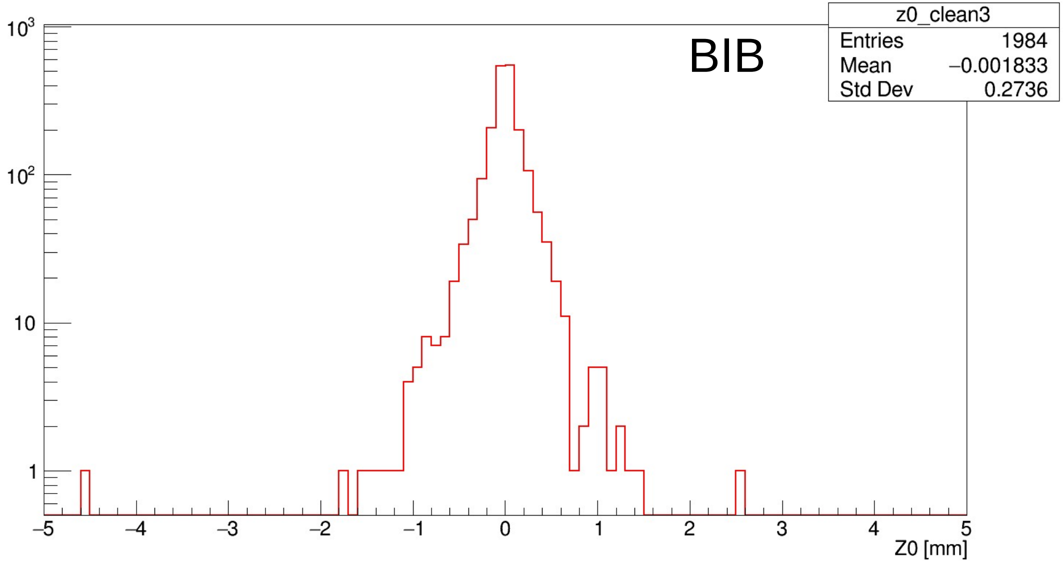


More zoomed in backup

Tracks parameters study

Z0 distribution with both chisquare, 6 hits requirements, and some other default requirements (error of Z0<1, combined significance of Z0 and D0>2)





Conclusions

Track selection and cuts on tracks and vertices parameters need to be optimized for the muon collider studies, the default values were setup for CLIC in completely different background conditions.

The works is in progress following a systematic approach and will be applied to $H \rightarrow b\bar{b}/c\bar{c}$ and HH.

The end

Thank you for you attention!

Backup

Backup

Backup: results after Vtnpart>2 (in H$\rightarrow$ bb 3 TeV no BIB)

Before

Number of vertices already assigned to tagged Jets 236
Total number of vertices out of cone: 66 over a total of 1542.
In percentage: 4.28016
Total tagging efficiency: 0.525201+-0.0102771, while the efficiency of tagging with more than one vertex is: 0.0999576+-0.00617293

Truth vs tagged; there are: 1790 jets recognized as truth.
So the efficiency with ONLY one tag: 0.524581+-0.0118037 and with more tags: 0.121788+-0.00772992
While efficiency with AT LEAST one tag: 0.646369+-0.0113003
So the mistag with ONLY one tag: 0.0407821 and with more tags: 0.00558659
So ONLY one tag is right with prob: 0.757258 while more tags: 0.923729

After

Number of vertices already assigned to tagged Jets 212Total number of vertices out of cone: 59 over a total of 1430. In percentage: 4.12587

Total tagging efficiency: 0.490894+-0.0102885, while the efficiency of tagging with more than one vertex is: 0.0897925+-0.00588359

Truth vs tagged; there are: 1643 jets recognized as truth.

So the efficiency with ONLY one tag: 0.536823+-0.0123018 and with more tags: 0.119903+-0.00801421

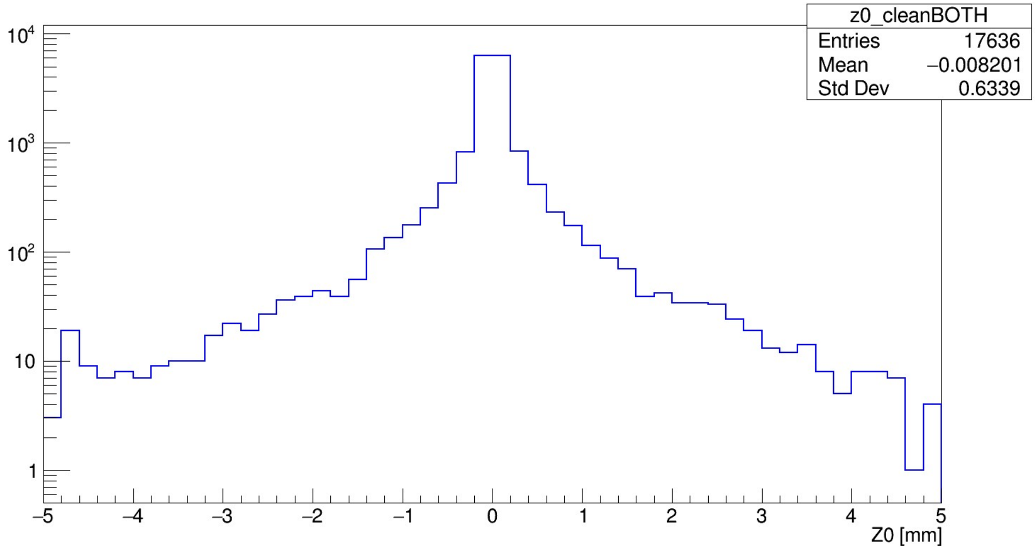
While efficiency with AT LEAST one tag: 0.656726+-0.0117137

So the mistag with ONLY one tag: 0.0432136 and with more tags: 0.00547778

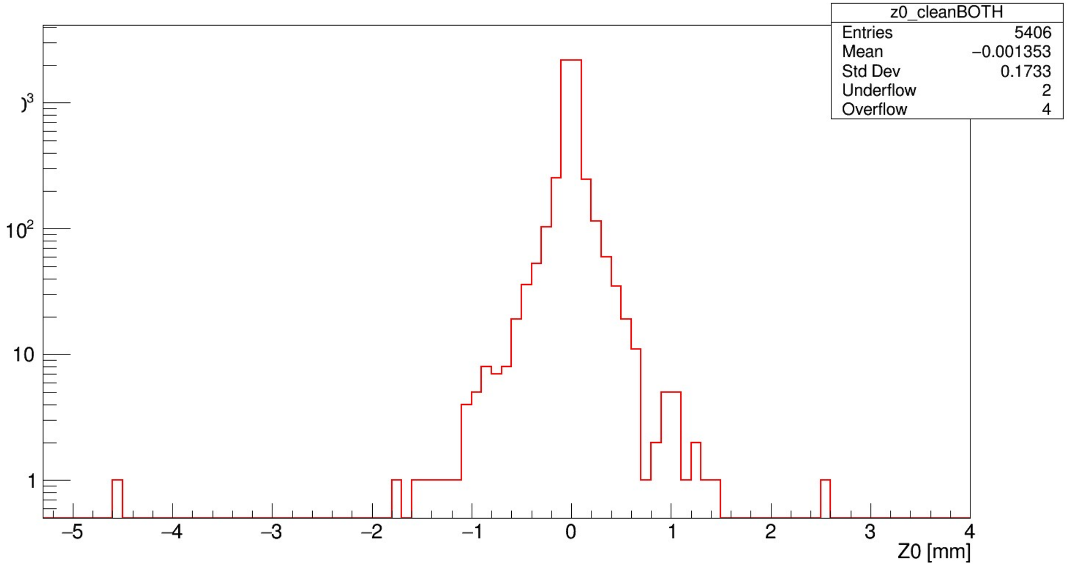
So ONLY one tag is right with prob: 0.761001 while more tags: 0.929245

Number of jets (not events!) not processed because in badly reconstructed event: 199

Zoomed plots

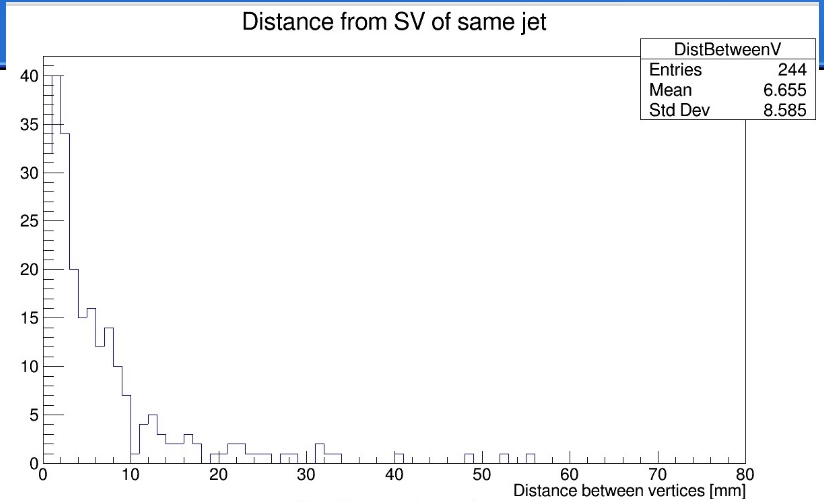


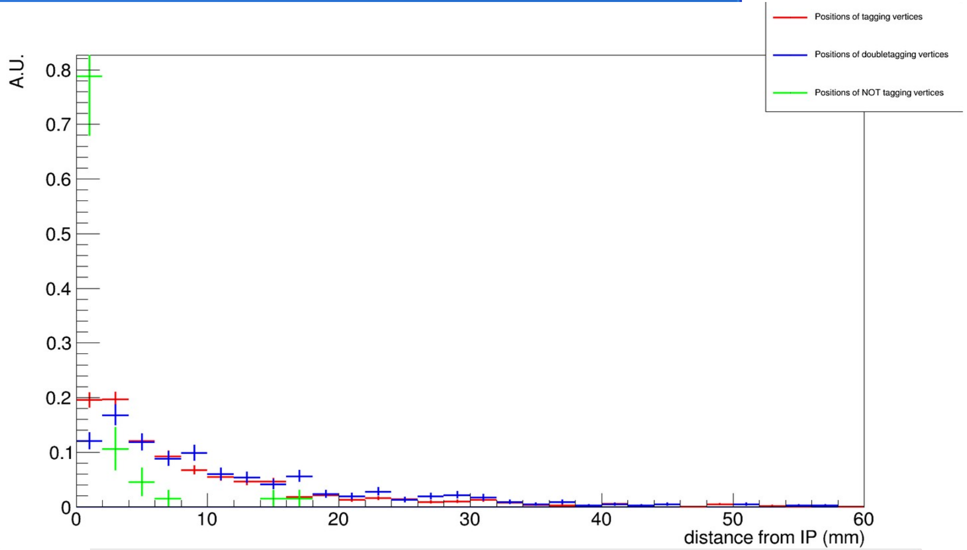
With both chisquare and 6 hits requirements

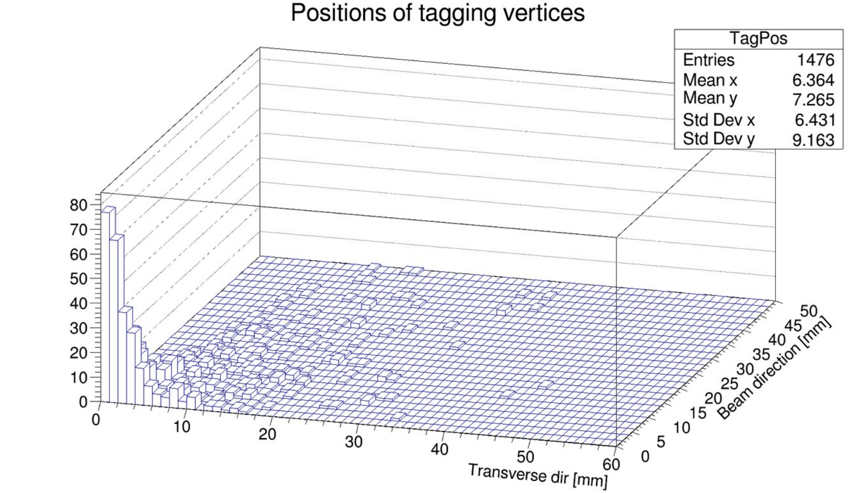


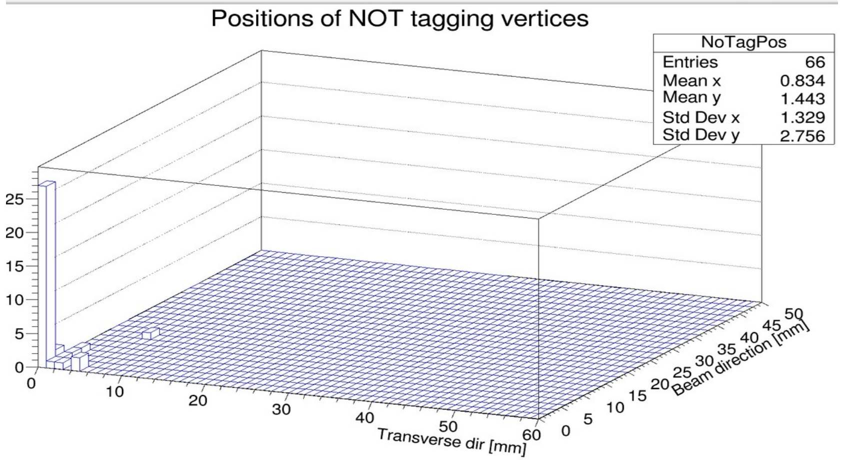
Some of these SV characteristics

All from H→bb 3TeV
No BIB



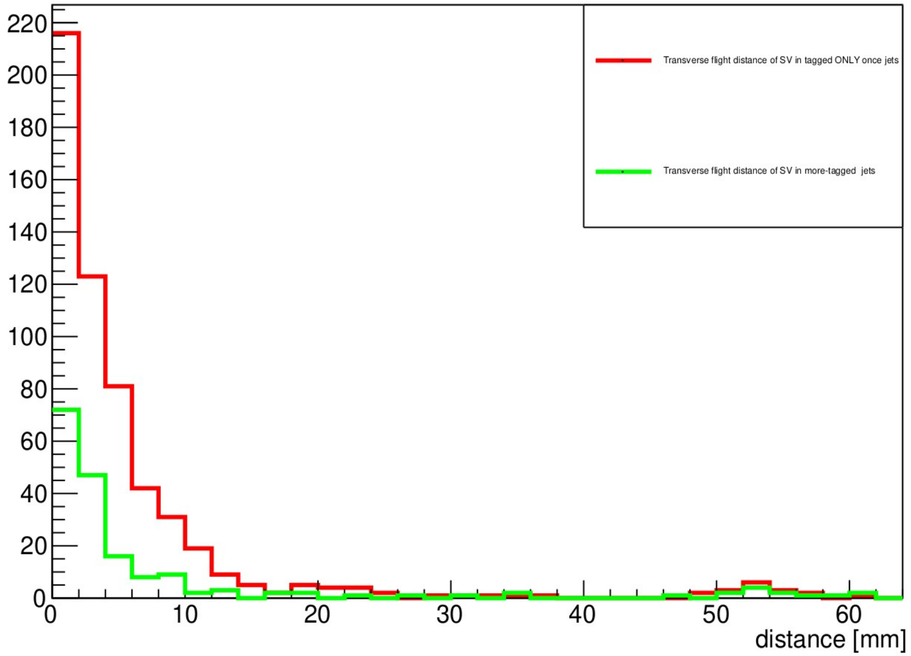






Some of these SV characteristics

Transverse flight distance of SV



With BIB on, similar characteristics, only difference this plot here, where a peak at high distances appear and its more prominent for SV in jets tagged more than once.
In that position lies the second layer of the vertex tracker (barrel).

Without BIB, the peak at 50 mm disappears.

Primary vertex study

First tests if we remove PV search and set it into a default position how do thing change?

Result: impossible to do effectively, even if PV processor is turned off, PV algorithm is set to not compute anything but return *V(0) and so on, it still looks for PV in other ways (i believe it looks in a list saved in Event::Instance(), but not sure).

If one effectively manages to remove PV, the program crashes

Primary vertex study

Other approach: make it impossible for him to find anything (proibitive parameters: max MaxD0=MaxZ0=0 or MaxChiqsquare of the track=0)

All primary Vertices in the final tuple are in 0,0,0 with 0 tracks associated! But since it does not crash nor give any flag like previously did, it’s likely it still can access saved PV, but it does not give him any tracks, which is what we wanted

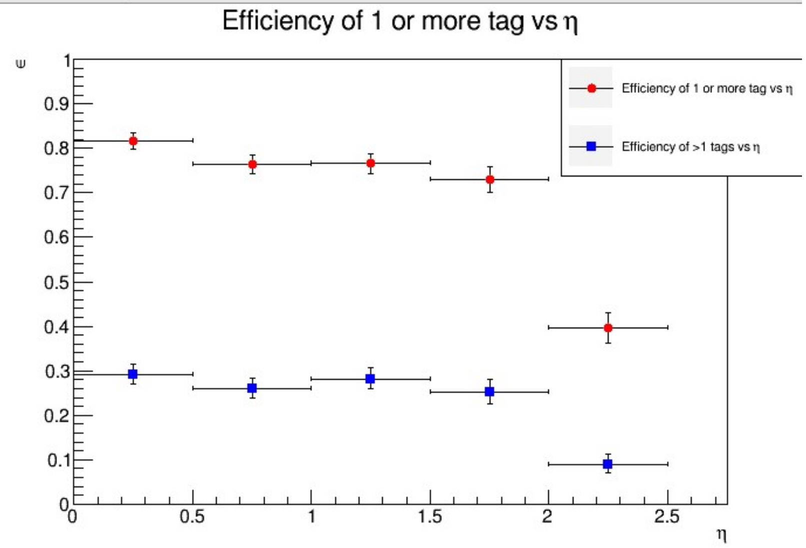
Result test 4: still on H → bb 3 TeV no BIB

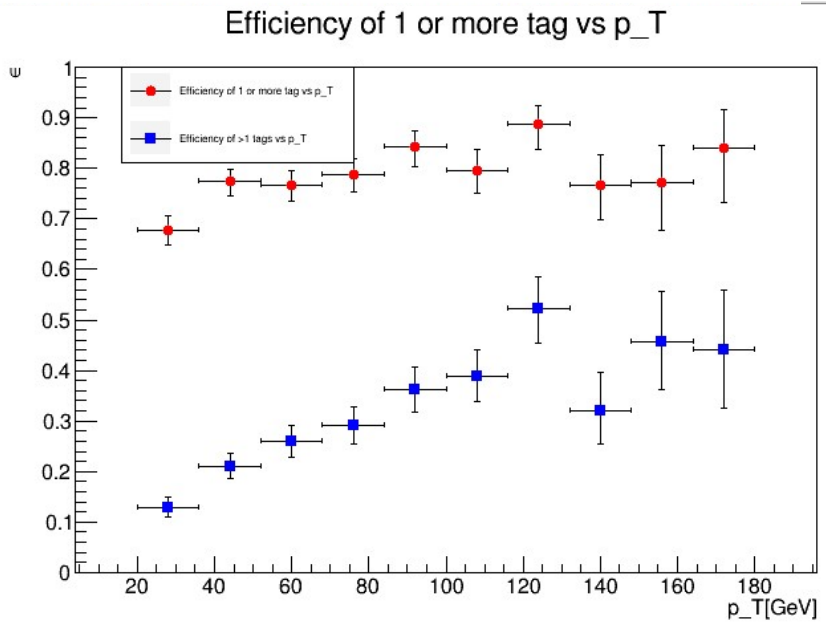
- WITH PRIMARY DISABILITATED
- Number of vertices assigned to already tagged Jets: 569
- Total number of vertices out of cone: 518 over a total of 2524 vertices. In percentage: 20.523.
- **Without truth-matching:**
- Total tagging efficiency: 0.60864+-0.0100443, while the efficiency of tagging with more than one vertex is: 0.241+-0.008802
- **With truth matching:**
- Truth vs tagged; there are: 1790 jets recognized as truth.
- So the efficiency with ONLY one tag: 0.476536+-0.011805 and with more tags: 0.249721+-0.0102309
- While efficiency with AT LEAST one tag: 0.726257+-0.0105388
- So the mistag with ONLY one tag: 0.0692737 and with more tags: 0.00726257
- So ONLY one tag is right with prob: 0.593598 while more tags: 0.785589
-
-
- WITH STANDARD CONFIGURATION (SO IT LOOKS FOR PRIMARY):
- Number of vertices assigned to already tagged Jets: 237
- Total number of vertices out of cone: 72 over a total of 1542 vertices. In percentage: 4.66926
- **Without truth-matching:**
- Total tagging efficiency: 0.522236+-0.01028, while the efficiency of tagging with more than one vertex is: 0.100381+-0.00618454
- **With truth matching:**
- Truth vs tagged; there are: 1790 jets recognized as truth.
- So the efficiency with ONLY one tag: 0.52067+-0.0118079 and with more tags: 0.122346+-0.00774516
- While efficiency with AT LEAST one tag: 0.643017+-0.0113242
- So the mistag with ONLY one tag: 0.0402235 and with more tags: 0.00558659
- So ONLY one tag is right with prob: 0.75588 while more tags: 0.924051

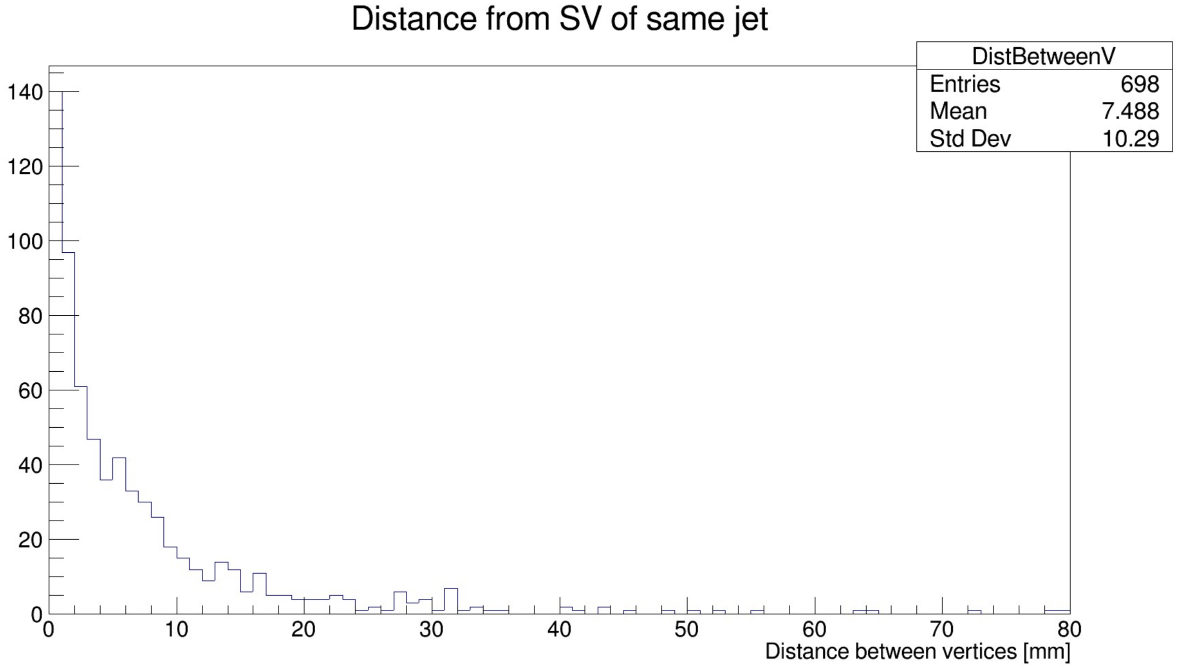
*Lots of more Spurious vertices
*much more likely to tag with more SV than with one
*much more NON b-truth matched jets are tagged (more mistag)

More details in Backup

Results test 4







Results test 4