Bc → $\tau\nu$ ($\tau \to 3\pi\nu$)

FCC Physics and Performance meeting

April 19, 2021

C Helsens, D Hill
CERN-EP EPFL

Context

To make progress with the Bc analysis needed a lot of ingredients

- Monte-Carlo
  - Production of exclusive decay modes for signal Bc and background Bu
    - New k4SimDelphes Pythia+EVTGEN interface developed and validated
  - Production of large background samples (3 billions of $Z \to qq$, q=uds,c,b)
    - Where exclusive decay modes are removed from $Z \to bb$
  - Dedicated production for MVA training (0.7 billion events)

- FCCAnalyses
  - Allow preprocessing of very large amount of data with complex code with HTCondor
  - Vertexing from MC perfect seeding
  - Particle Identification (perfect PID assumed here)
  - Implemented *Combination* from awkward C++ (triggered awkward to RDF interest)
  - Developed a lot of code (tau candidate building, analysis utilities…)

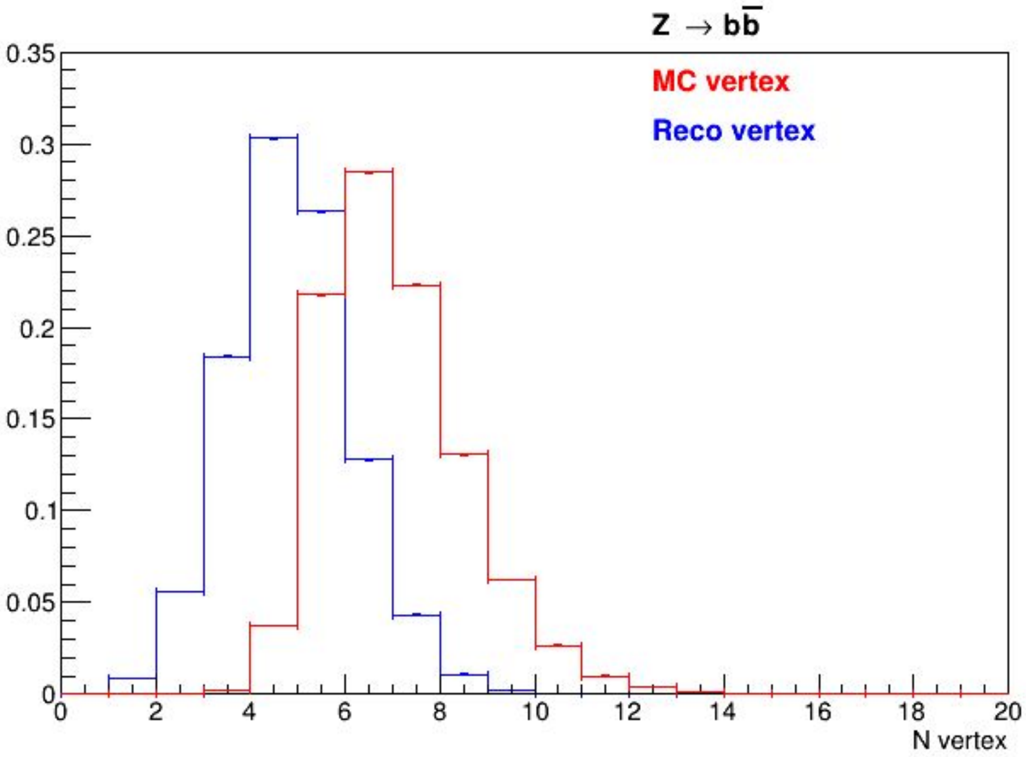
Vertexing

Context

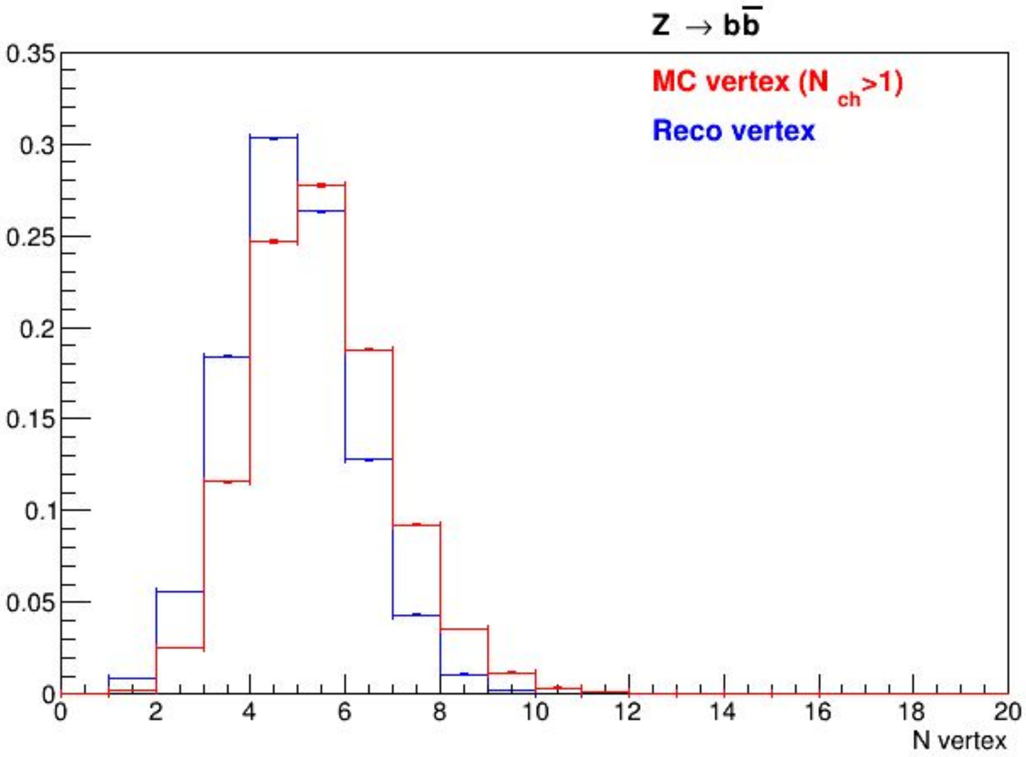
Developed perfect seeding vertex finder

- Procedure:
  - Finds MC vertex with stable charged particles
  - Run vertex fitting using reconstructed tracks associated to the MC ones (if at least 2 tracks found at reco level)
- This is the best vertexing we could dream of as
  - We perfectly seed the vertex fitter
  - We have extremely good displaced vertex reconstruction efficiency
  - But it still takes into account acceptance effect and vertex fit quality
- Possible improvements to make it more realistic would be to
  - merge close by vertex
- Next slides shows some vertexing performance plots

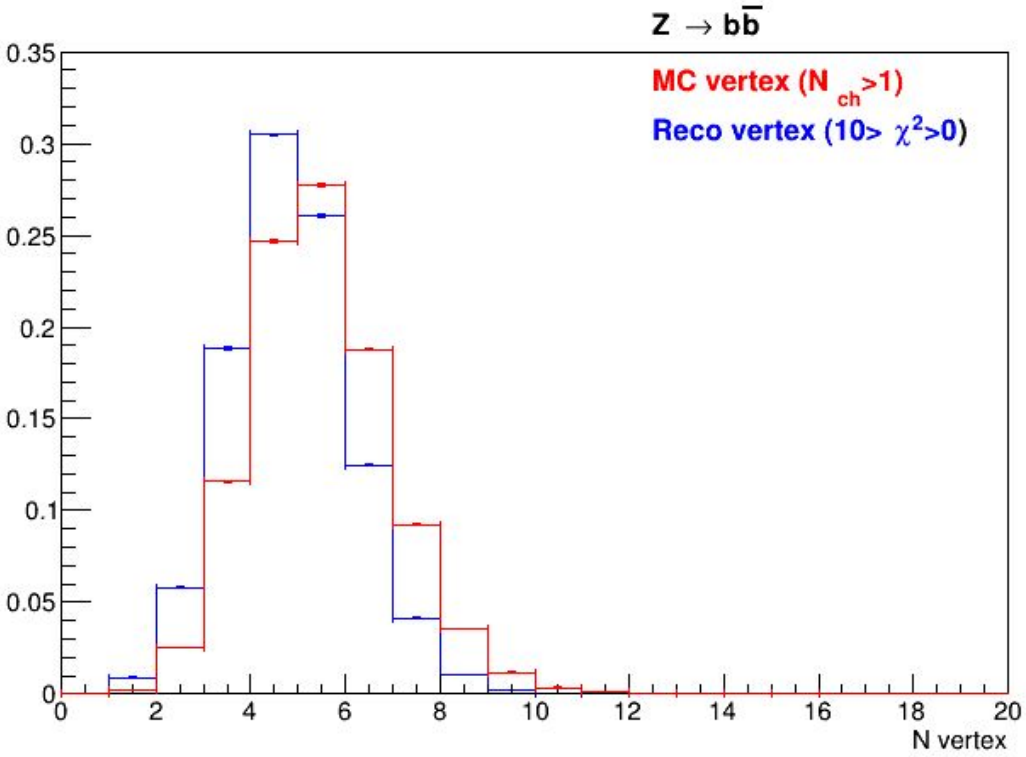
Number of vertex - 1



Number of vertex - 2

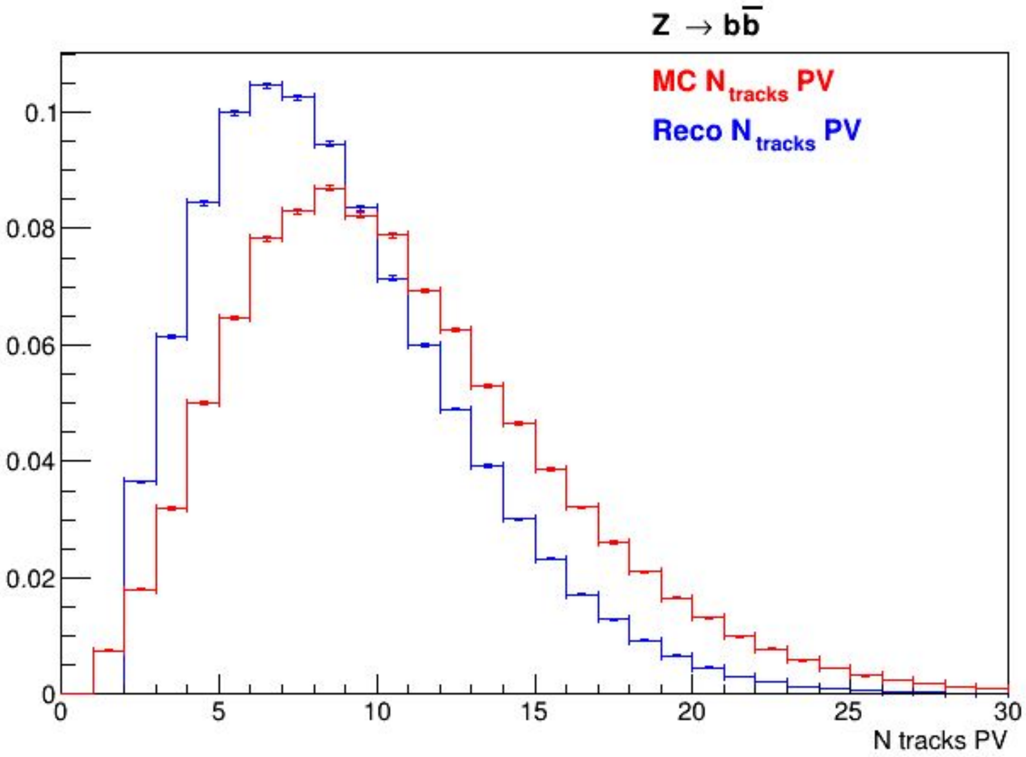


Number of vertex - 3

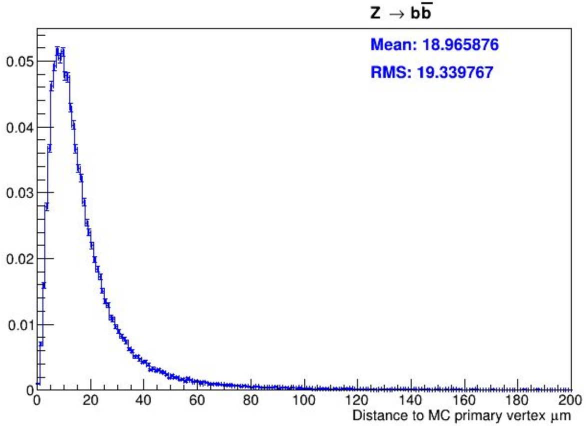


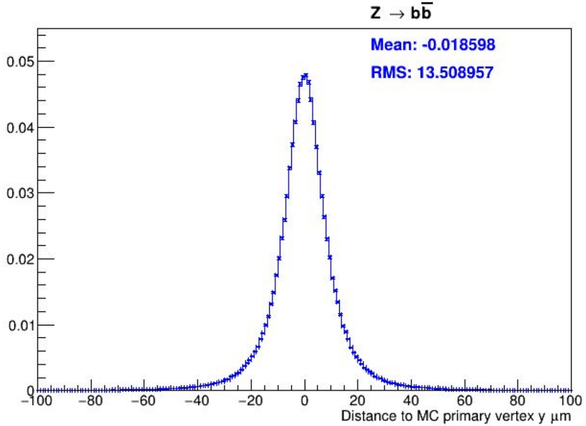
Primary Vertex

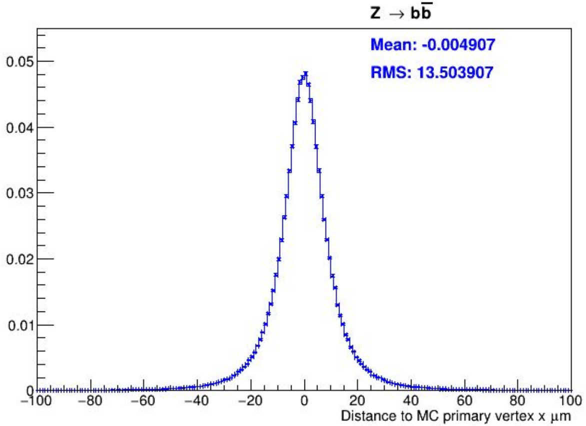
Number of tracks PV

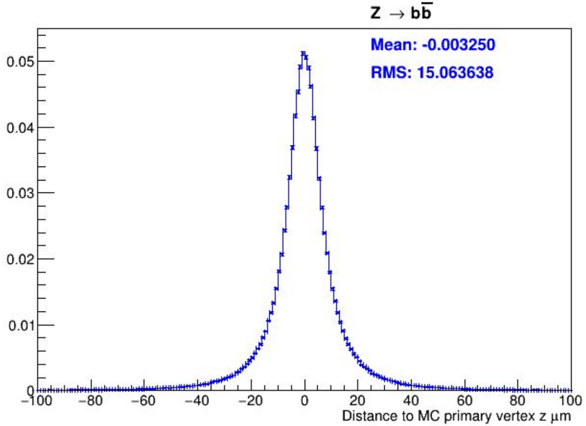


Distance to MC vertex

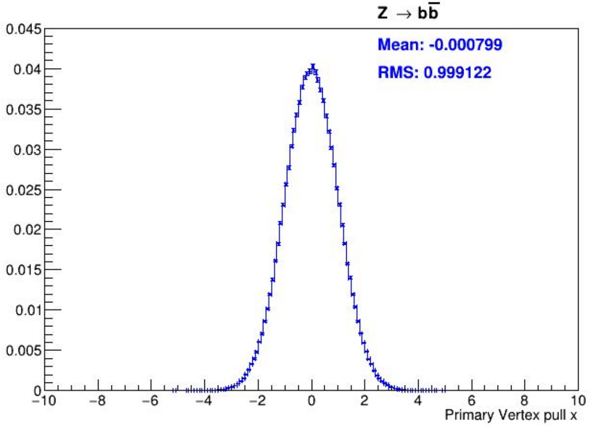




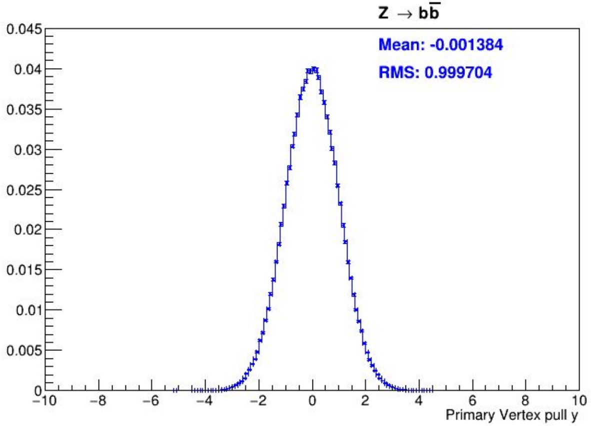




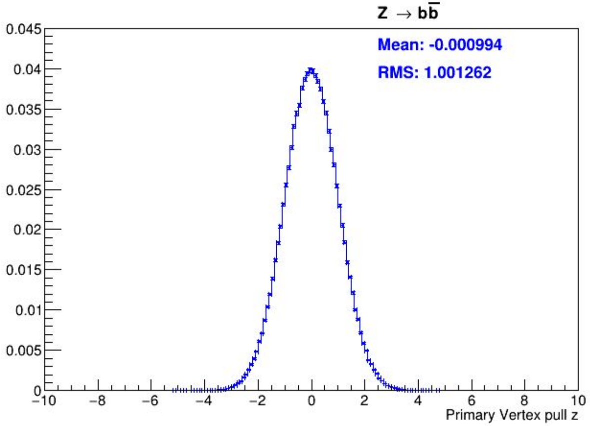
Pull vertex



Z → b$\bar{b}$
Mean: -0.000799
RMS: 0.999122
Primary Vertex pull x



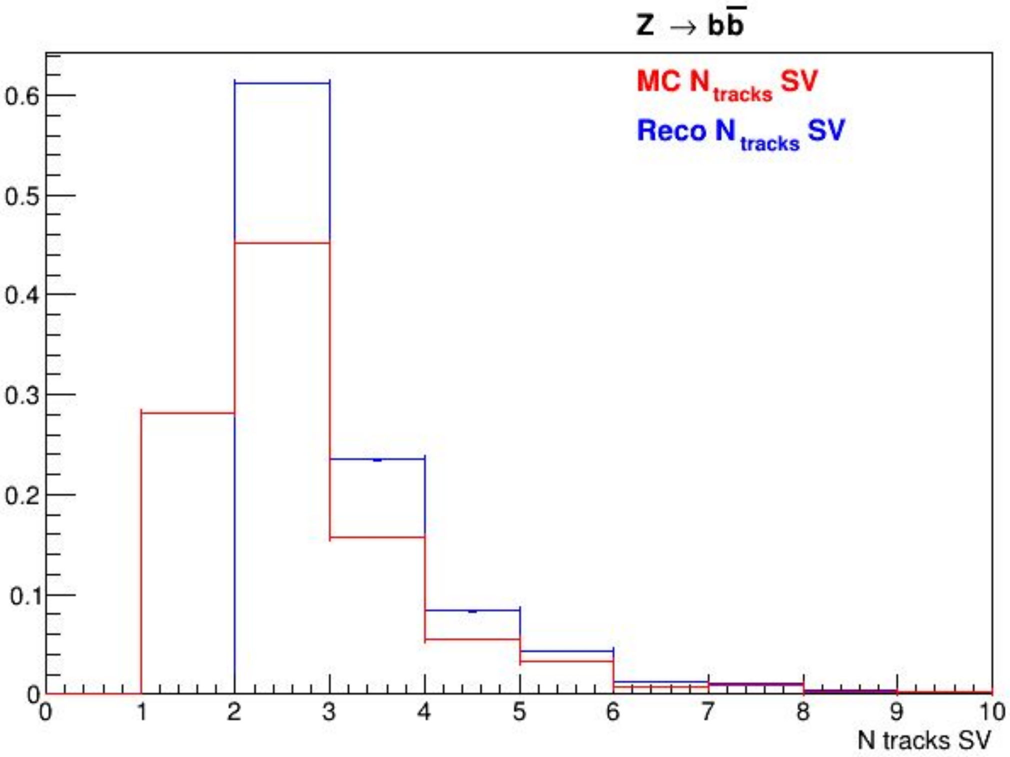
Z → b$\bar{b}$
Mean: -0.001384
RMS: 0.999704
Primary Vertex pull y



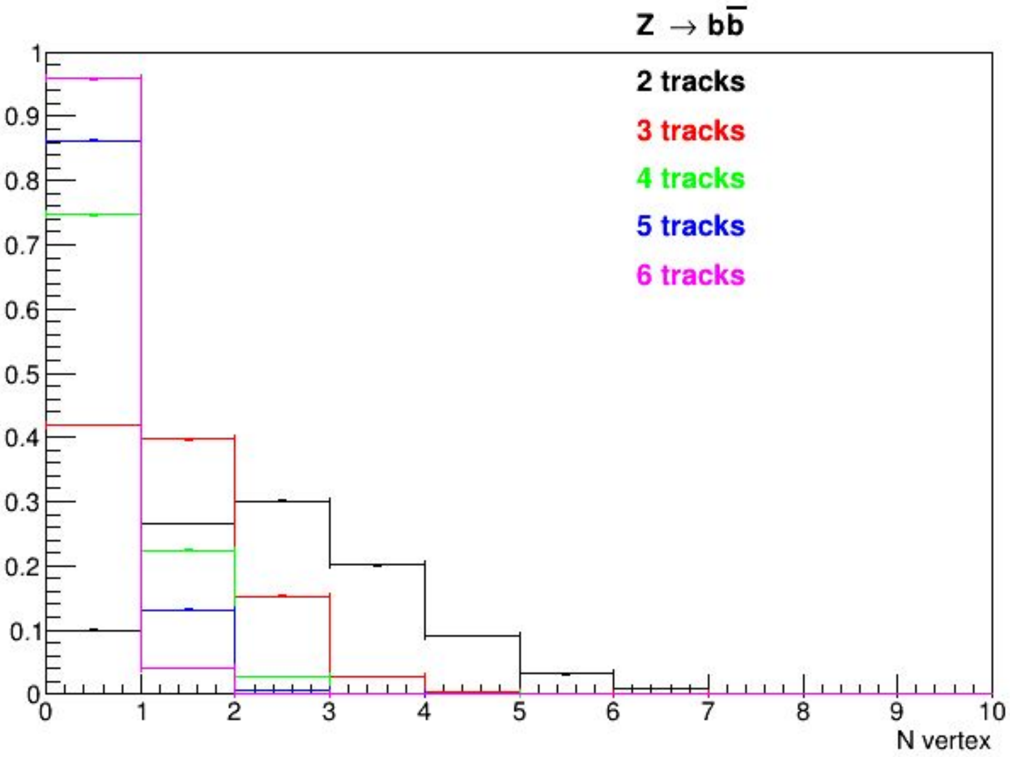
Z → b$\bar{b}$
Mean: -0.000994
RMS: 1.001262
Primary Vertex pull z

Secondary Vertex

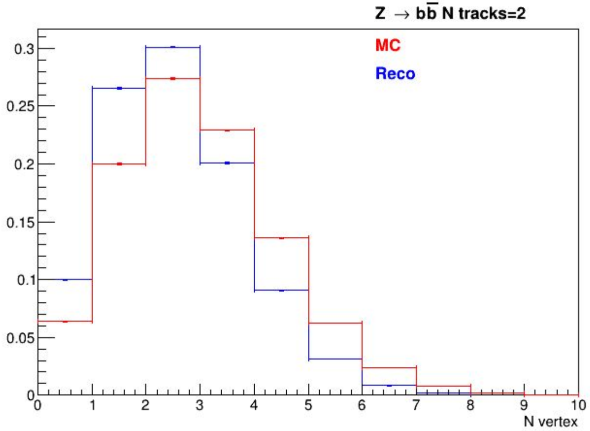
Number of tracks SV

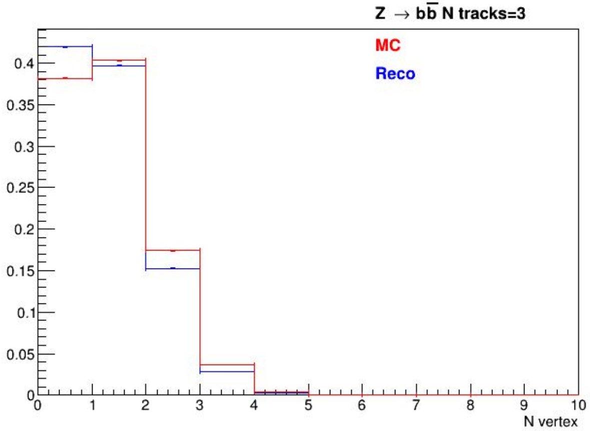


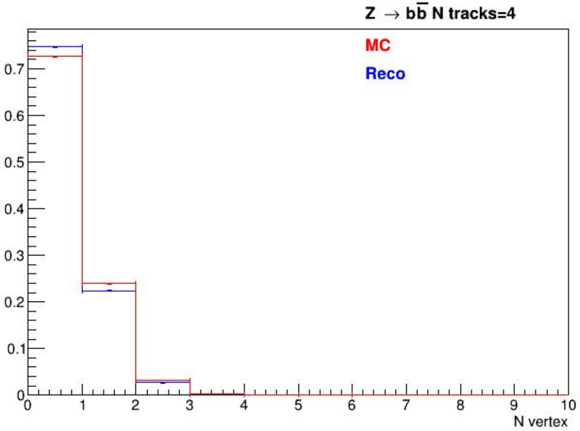
Number reco SV

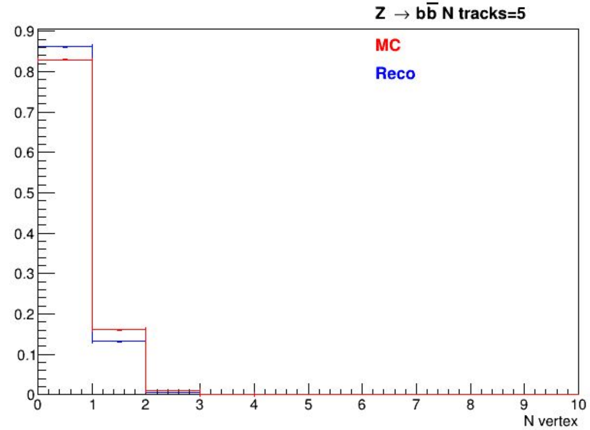


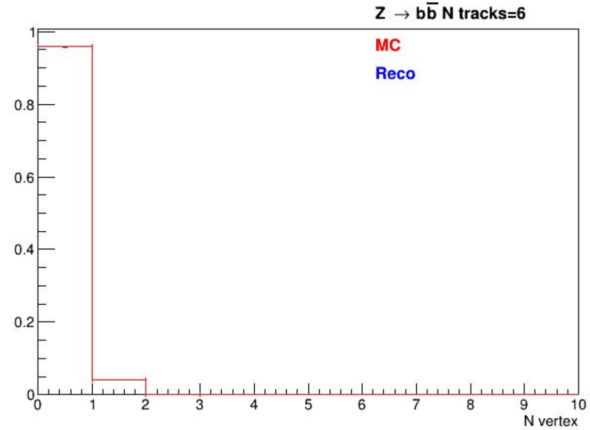
Number SV per track multiplicity



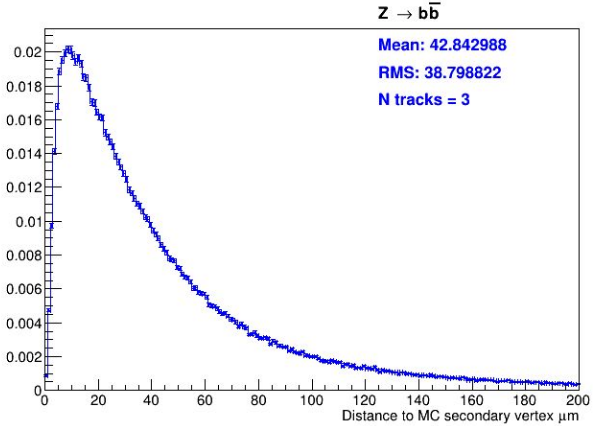


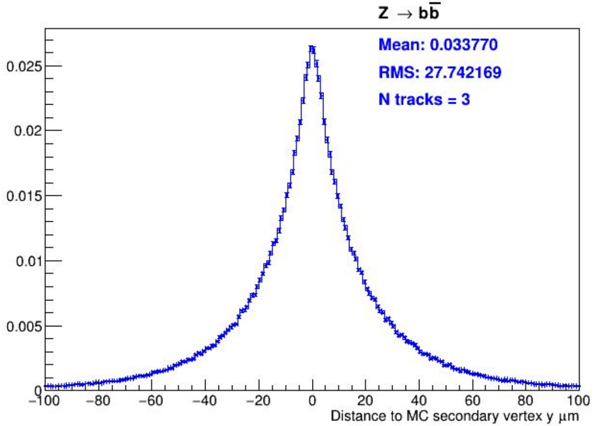


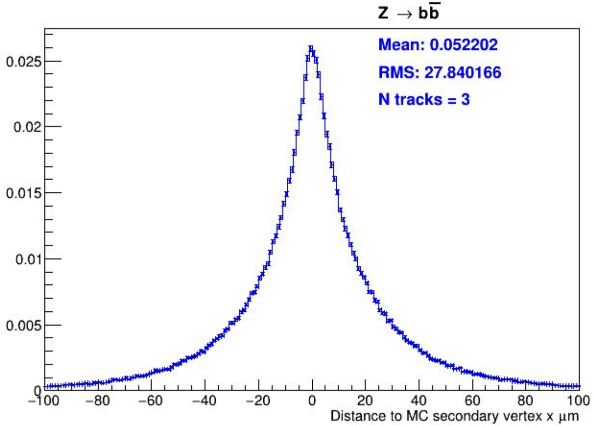


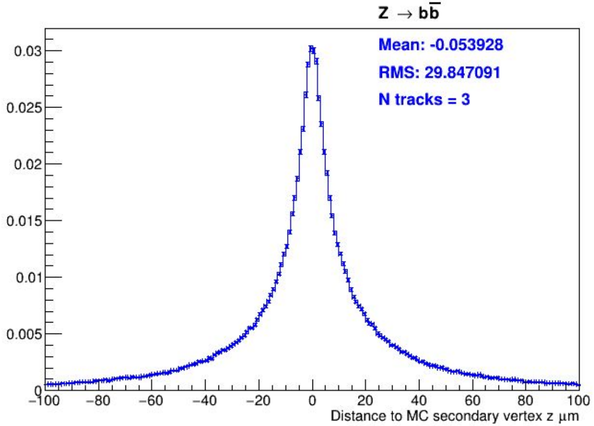


Distance to MC vertex - N tracks = 3



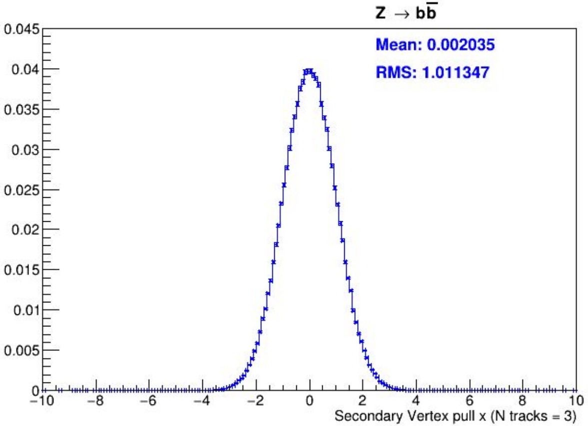


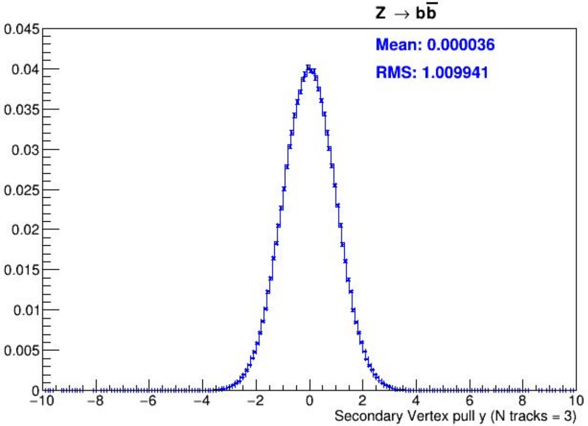




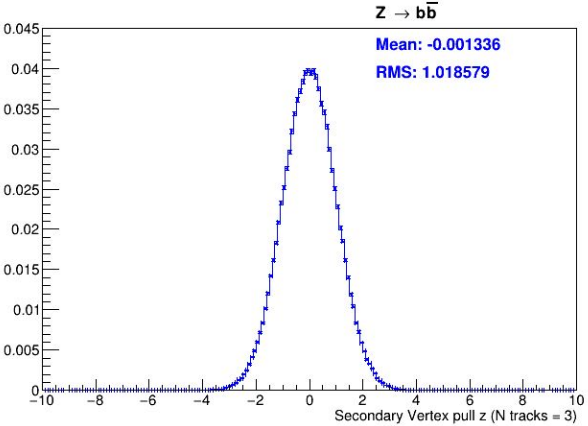
In backup plots for other track multiplicities

Vertex pull - N tracks = 3

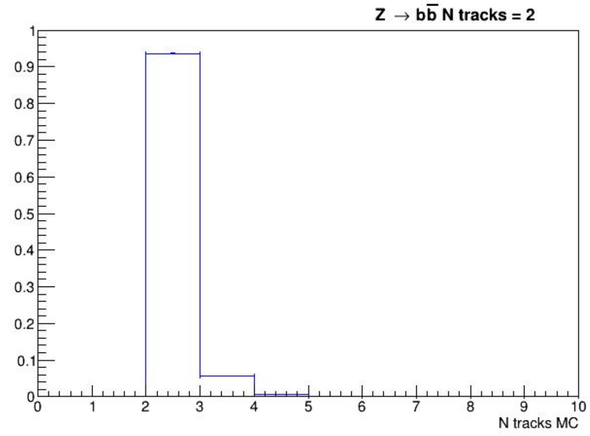


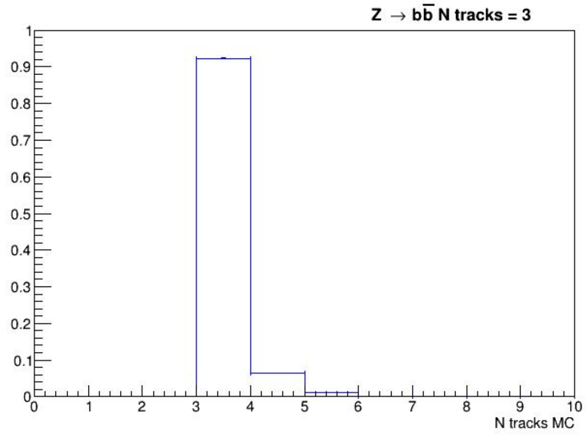


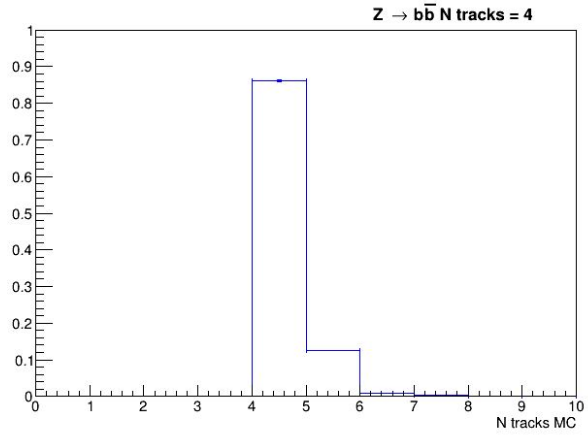
In backup plots for other track multiplicities

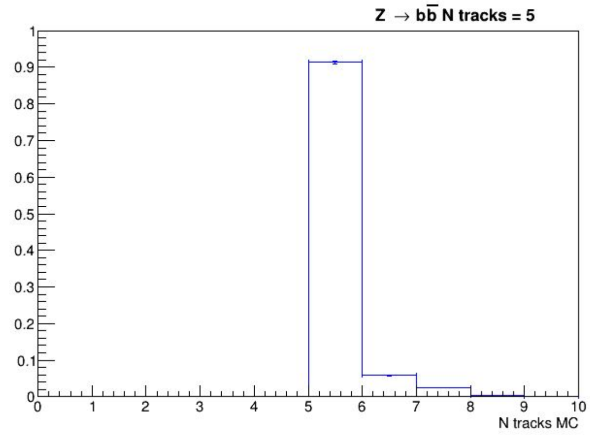


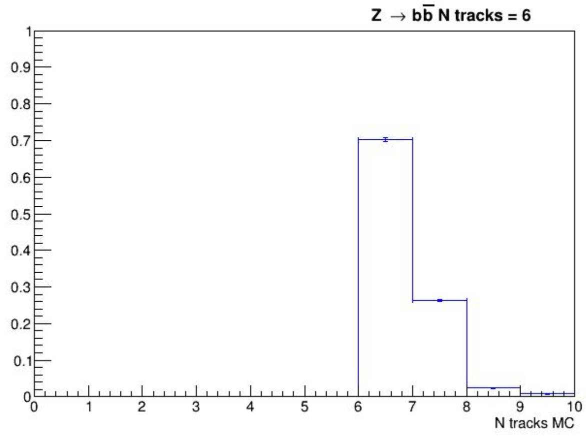
Vertex migration











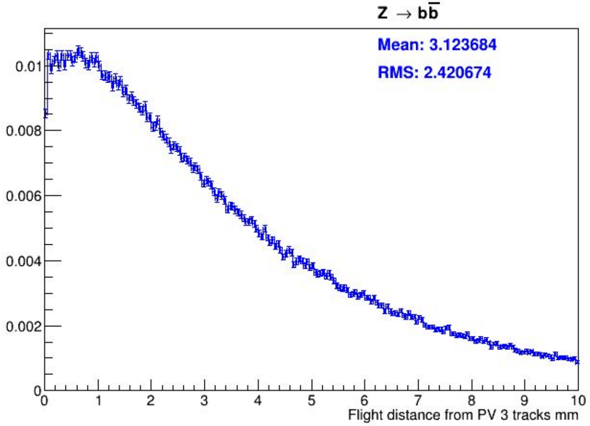
Fraction of reco SV with N tracks versus the number of MC tracks.

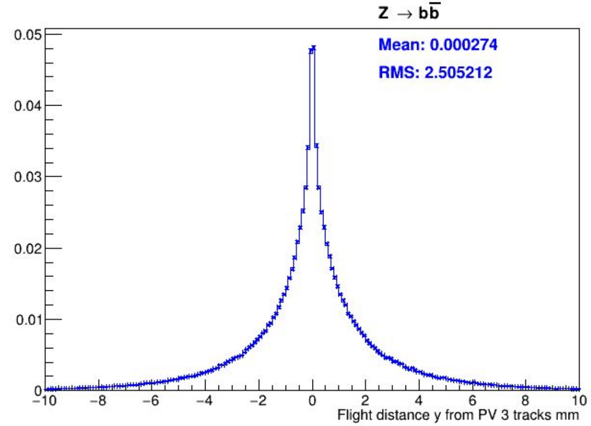
<u>Example:</u>

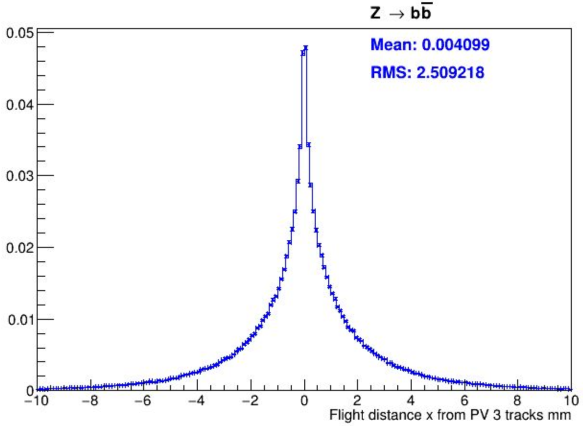
For a 3 tracks reconstructed vertex

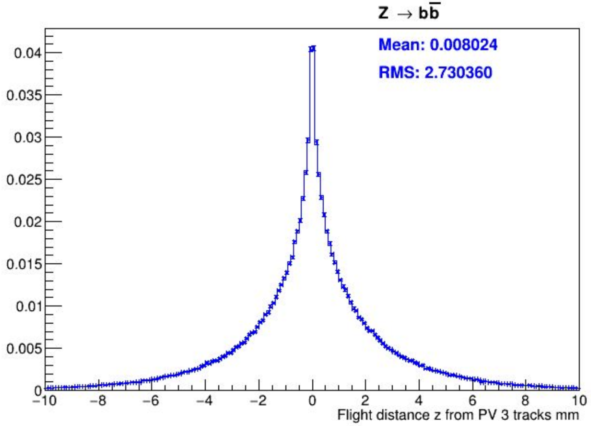
- 92% from 3 tracks MC
- 7% from 4 tracks MC
- 1% from 5 tracks MC

Flight distance - N tracks = 3

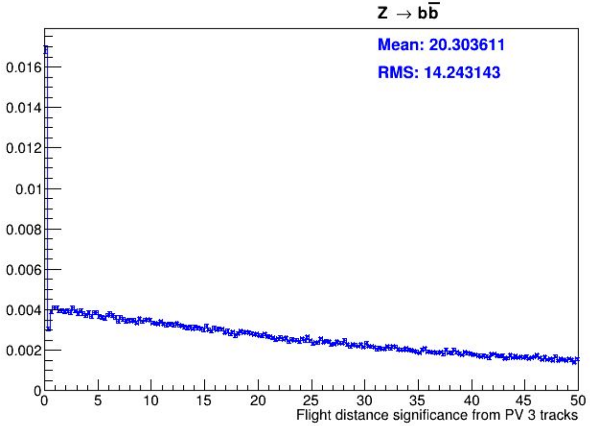


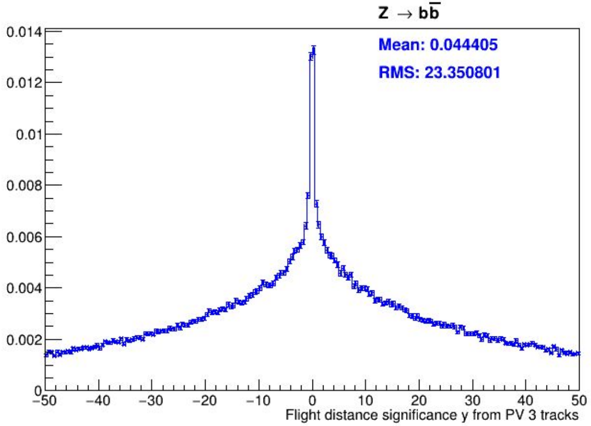




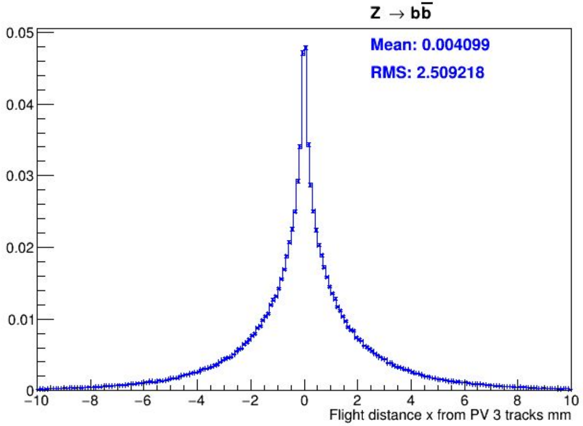


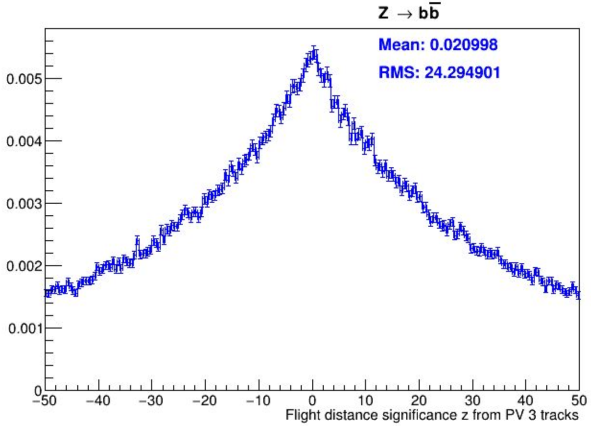
Flight distance significance - N tracks = 3





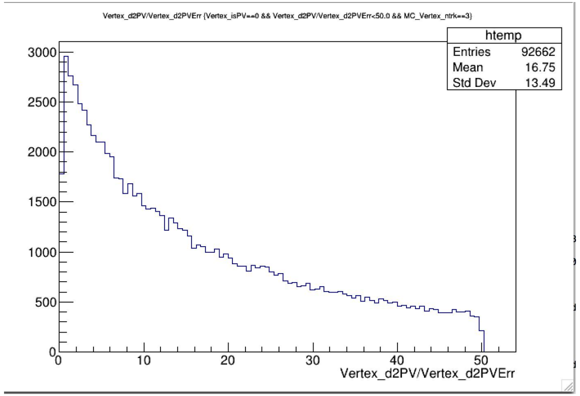
In backup plots for other track multiplicities

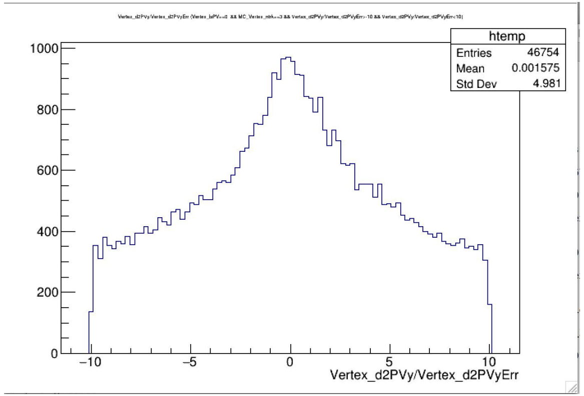


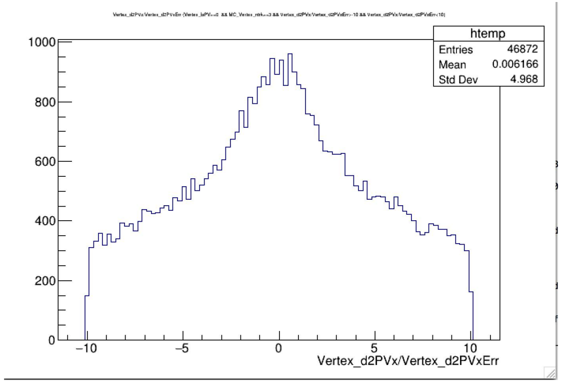


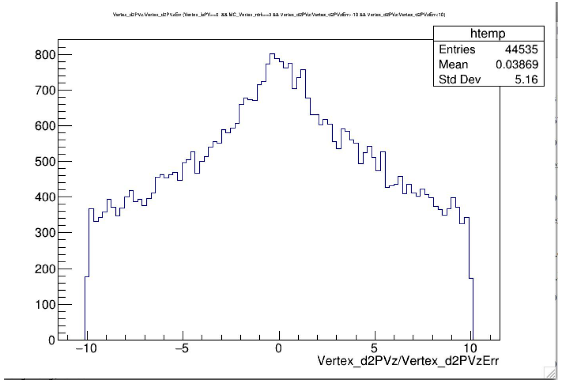
PV is included here, need to redo the plots after the meeting

Flight distance significance - N tracks = 3



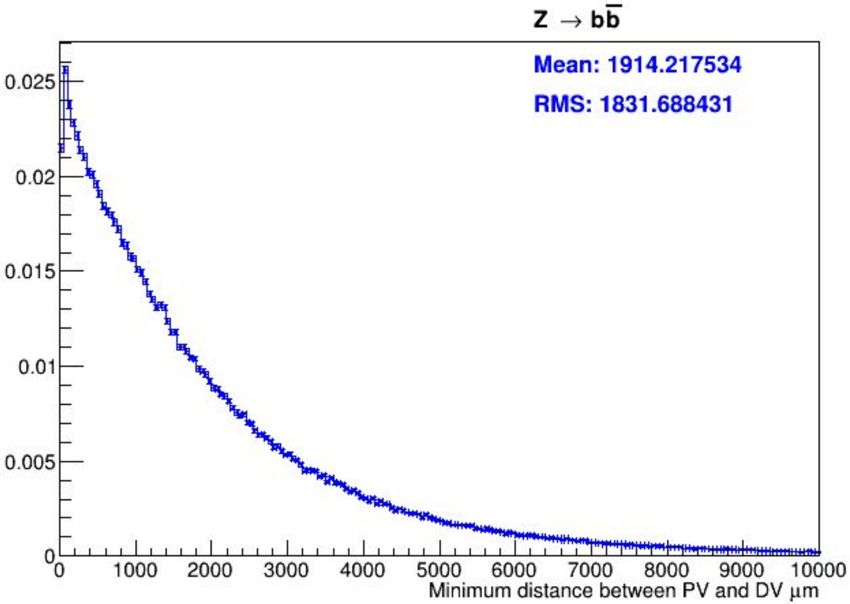


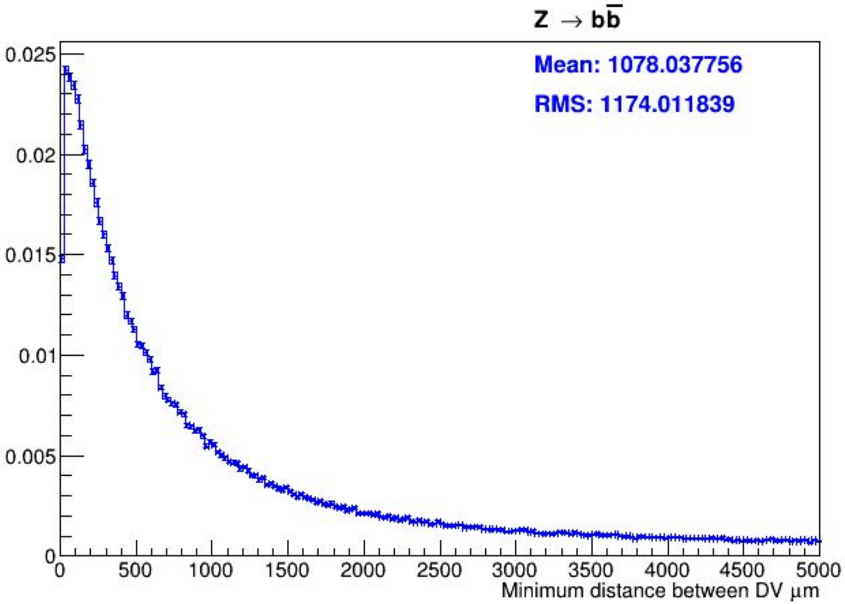




In backup plots for other track multiplicities

Minimum distance PV-DV DV-DV





$B_c \rightarrow \tau\, \nu\ (\tau \rightarrow 3\pi\, \nu)$

Motivation

Leptonic decay of Bc meson - in SM, annihilation of quarks to produce a W

Rate is sensitive to CKM factor |Vcb|, but also to NP Wilson coefficients

Complementary to b → c τ ν modes, since it involves the same vertex factors

- These modes show deviations compared to SM e.g. R(D), R(D*)

Decay is not yet observed: not possible to reconstruct at LHC due to missing energy, and no Bc produced at B factories

Aim to study feasibility of a branching fraction measurement at FCC-ee Z-pole

Existing studies from CEPC

Feasibility study of Bc → $\tau\nu$ (τ → $l\ \nu\ \nu$) at CEPC [arXiv:2007.08234]

Exploits missing energy of signal and hemisphere structure of Z → bb events, using the thrust axis to define hemispheres

No decay vertices reconstructed, so they rely on information from a single electron or muon candidate (like IP to the thrust axis)

Focus on separating signal from Z → qq, cc, bb via MVAs, but also on separating signal from the very similar B+ → $\tau\ \nu$ mode

- B+ mode is CKM suppressed compared to signal (|Vub| vs. |Vcb|), but the production rate is 1000x larger
- Leads to **N(Bc) / N(B+)** = **0.28** ± **0.05 expected in SM**

Our approach

Use the **$\tau \rightarrow 3\pi\,\nu$ decay** (9% branching ratio) to provide a reconstructible τ vertex

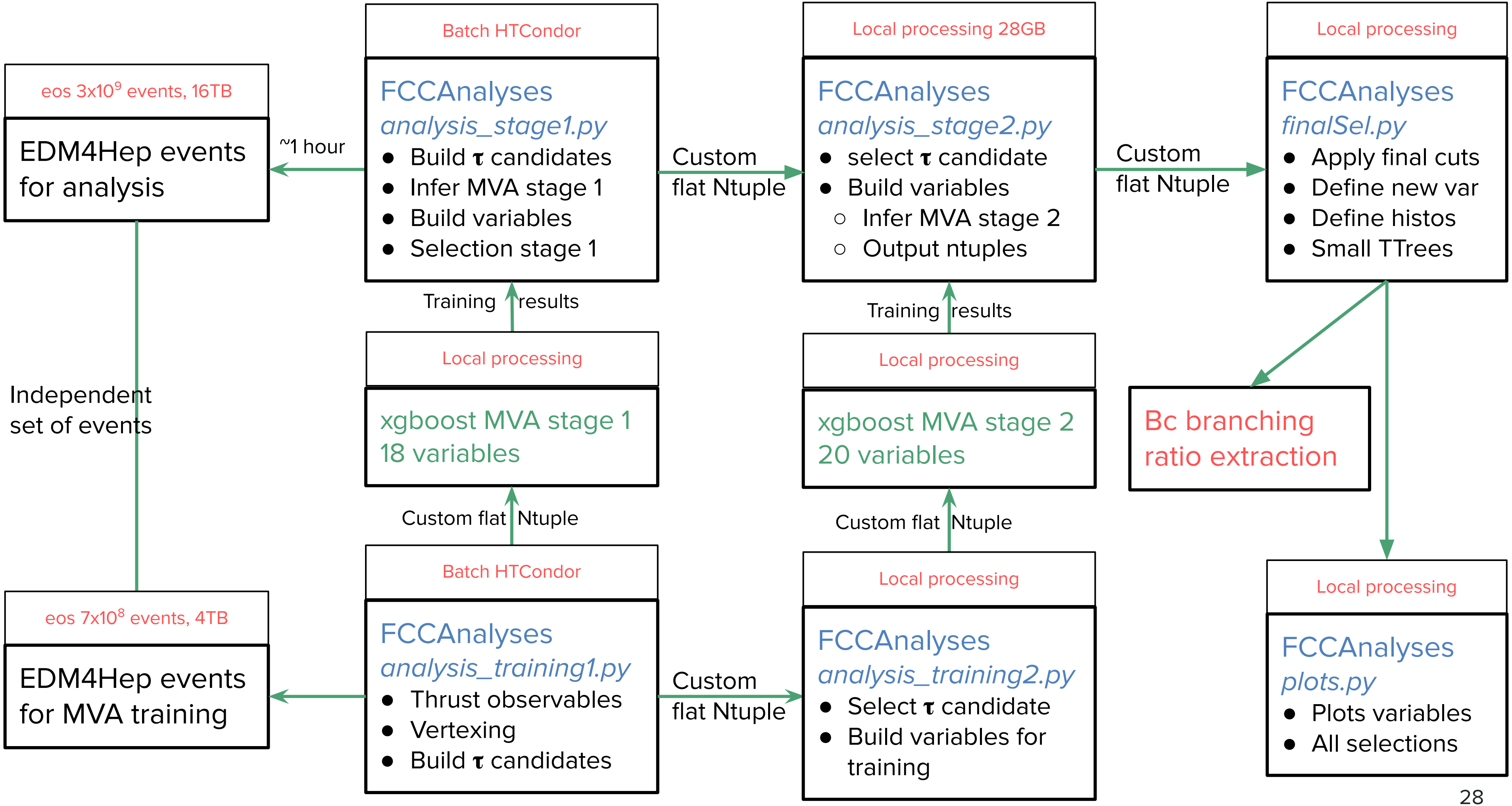
Allows a **precise measure of the combined (Bc + τ) flight**, which can distinguish the signal from B+ $\rightarrow \tau\,\nu$ since the B+ lifetime is 3x larger

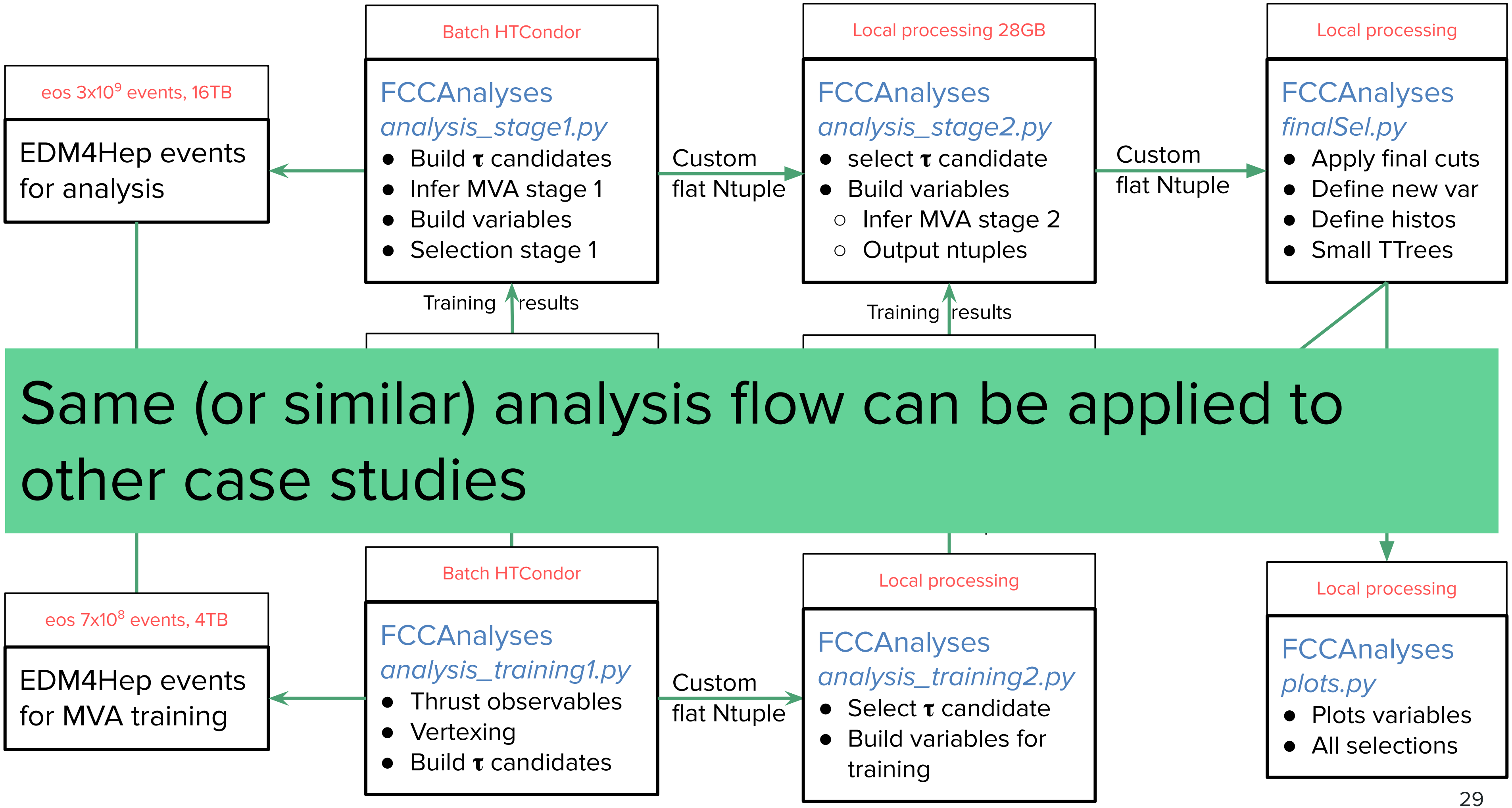
Use an MVA trained on event-level information to suppress Z $\rightarrow$ qq, cc, bb

Use $\tau \rightarrow 3\pi$ decay in signal MC to study vertexing performance - use the reconstructed 3π information to derive a full signal selection

Study the signal purity achievable, and the possible precision of a branching fraction measurement at FCC-ee

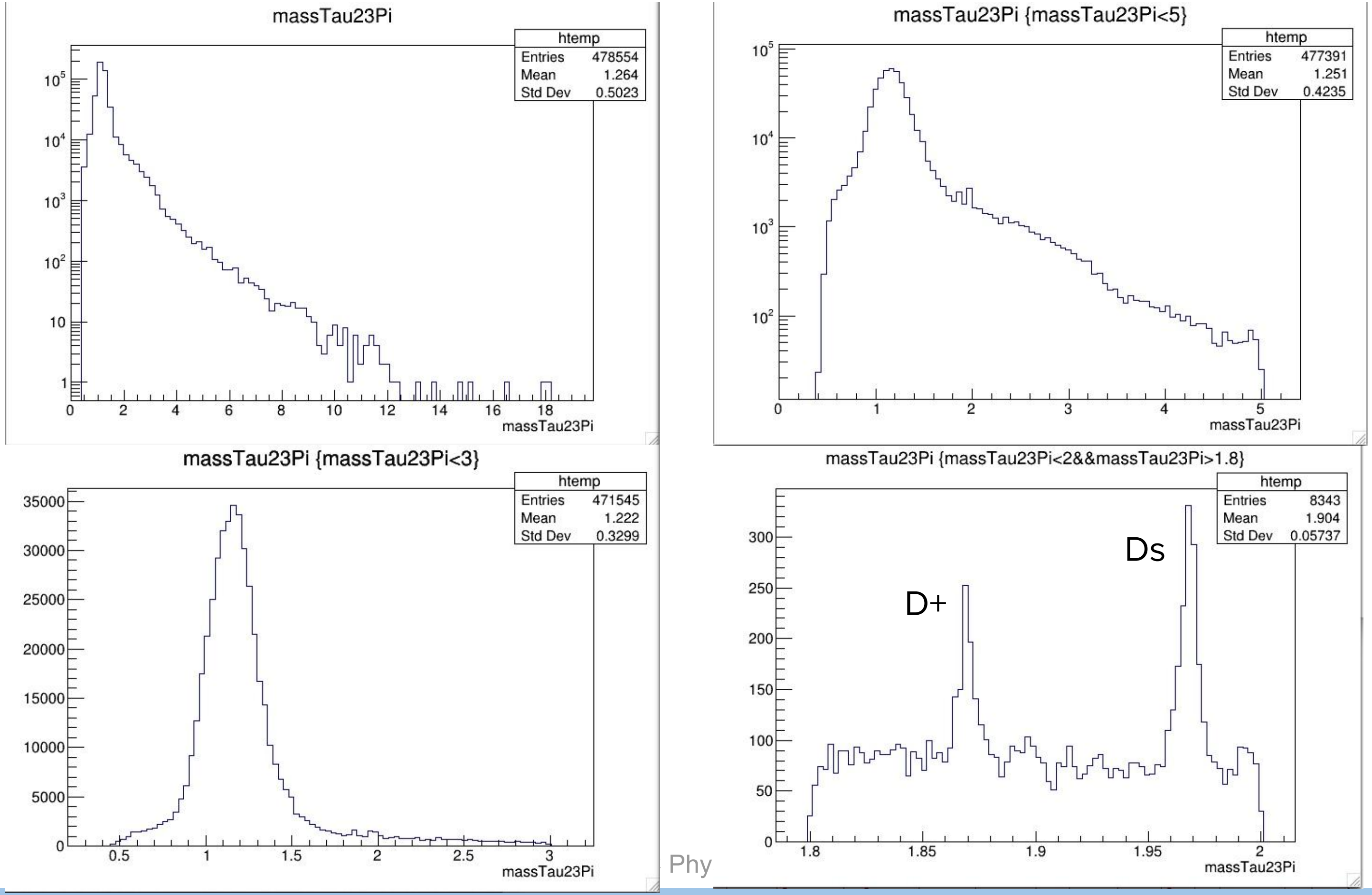
Analysis description


eos $3x10^9$ events, 16TB
EDM4Hep events for analysis
~1 hour
Batch HTCondor
FCCAnalyses
analysis_stage1.py
Build τ candidates
Infer MVA stage 1
Build variables
Selection stage 1
Custom flat Ntuple
Local processing 28GB
FCCAnalyses
analysis_stage2.py
select τ candidate
Build variables
Infer MVA stage 2
Output ntuples
Custom flat Ntuple
Local processing
FCCAnalyses
finalSel.py
Apply final cuts
Define new var
Define histos
Small TTrees
Training results
Local processing
xgboost MVA stage 1
18 variables
Training results
Local processing
xgboost MVA stage 2
20 variables
Bc branching ratio extraction
Independent set of events
Custom flat Ntuple
Custom flat Ntuple
eos $7x10^8$ events, 4TB
EDM4Hep events for MVA training
Batch HTCondor
FCCAnalyses
analysis_training1.py
Thrust observables
Vertexing
Build τ candidates
Custom flat Ntuple
Local processing
FCCAnalyses
analysis_training2.py
Select τ candidate
Build variables for training
Local processing
FCCAnalyses
plots.py
Plots variables
All selections



Same (or similar) analysis flow can be applied to other case studies

3π mass plot (no cuts)



To validate the candidates reconstruction. Everything is doing well

Stage 1 MVA: reject inclusive background

Use hemisphere energy information calculated in FCCAnalyses, based on thrust axis determination (also done in FCCAnalyses)

Train xgboost binary classifier on Bc ➔ (τ ➔ 3π ν) ν and a mixed sample of inclusive decays Z ➔ qq, cc, bb (combined according to Z branching fractions and pre-sel eff)

- Require a PV and also at least one reconstructed 3π candidate as a pre-selection, which reduces Z ➔ qq (100x) / cc (10x) / bb (3x)

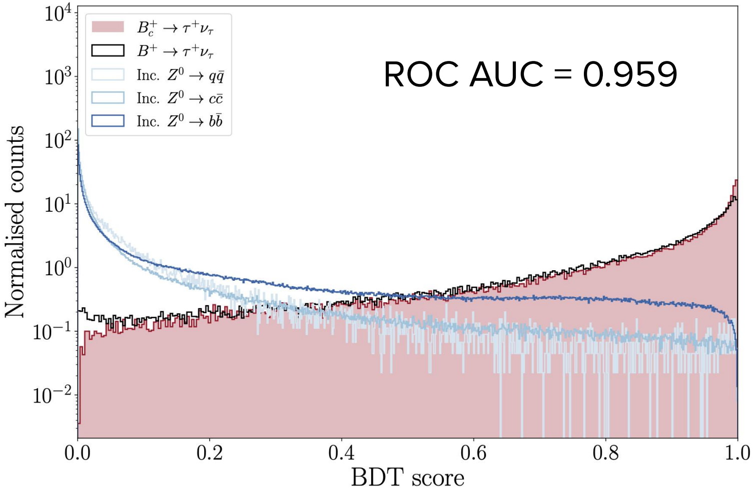
Model training is done in Python, but the trained model is persisted to ROOT file and applied to samples via RDataFrame in FCCAnalyses (cutting edge in ROOT!)

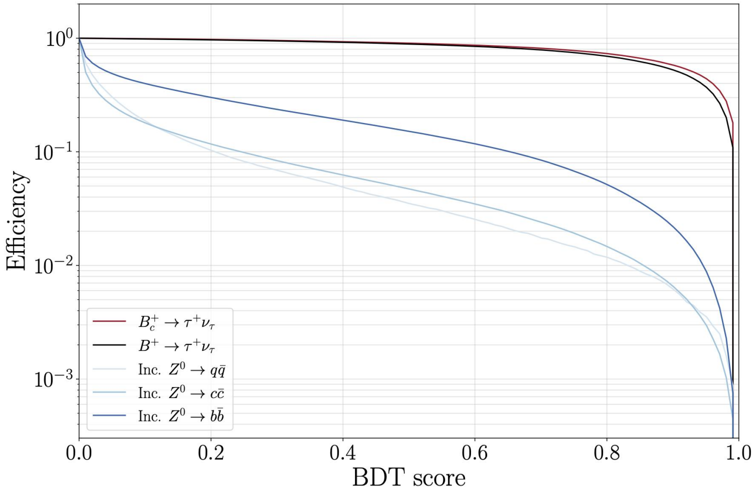
```
TMVA::Experimental::RBDT<> bdt("Bc2TauNu_BDT", "/eos/experiment/fcc/ee/analyses/case-studies/flavour/Bc2TauNu/xgb_bdt_vtx.root");
computeModel = TMVA::Experimental::Compute<18, float>(bdt);
```

Stage 1 MVA: reject inclusive background

High performance on all classes of background, but Z ➔ bb found to be most signal-like (makes sense since b ➔ c ➔ s produces more missing energy)

Signal looks similar to B+ ➔ (τ ➔ 3π ν) ν, which is a prominent background



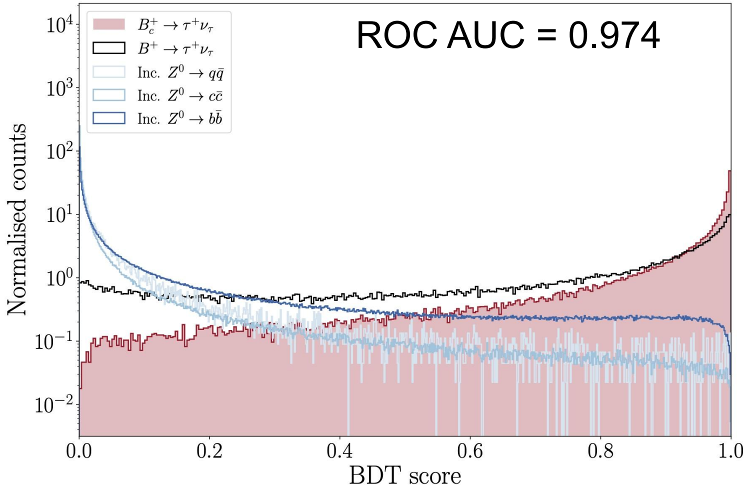


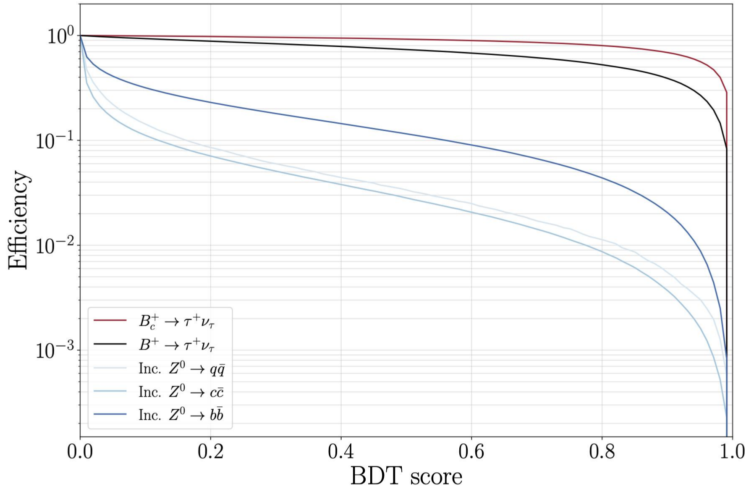
Alternative MVA with additional vertex variables

Also trained the MVA including information on number of reconstructed vertices, number of 3π candidates, PV track multiplicity, and 3π vertex separations from PV

Better performance, and B+ → $\tau\ \nu$ rejected more due to use of vertex distance info

- We proceed to use this MVA, with a **BDT > 0.6 pre-cut** (90% signal-efficient)





Pre-selection efficiency

Cumulative efficiencies

Cut/Process	Bc→$\tau\nu$	Bu→$\tau\nu$	Z→bb	Z→cc	Z→uds
Has PV	0.898	0.984	0.982	0.998	0.9998
N tau cand>0	0.777	0.861	0.316	0.112	0.00332
1^{st} MVA>0.6	0.688	0.590	0.0288	0.00234	8.35e-05

After pre-selection and considering 150 ab^{-1}, the following yields are expected:

N(Bc→$\tau\nu$, τ→3$\pi\nu$)	1 205 000 events
N(Bu→$\tau\nu$, τ→3$\pi\nu$)	5 133 000 events
N(Z→bb)	28 627 500 000 events (28 10^6 MC events left out of 10^9)
N(Z→cc)	1 827 800 000 events (2.3 10^6 MC events left out of 10^9)
N(Z→uds)	233 200 000 events (8.4 10^4 MC events left out 10^9)

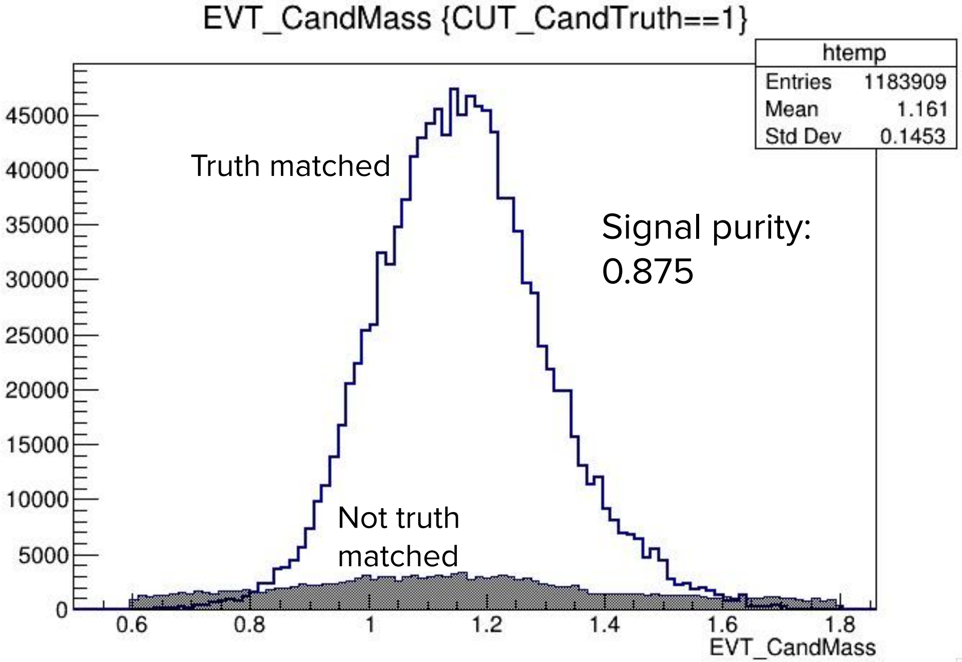
3π candidate selection - 1

Select candidates with:

- $m(3\pi) > 0.6$ GeV and < 1.8 GeV
- Vertex chi2 > 0 and < 10
- If several candidates, chose smallest chi2

Non truth-matched 3π are candidates built in signal events which do not come from the signal **τ**

e.g. a charm meson decay from the non-signal hemisphere



Signal purity: 0.875

3π candidate selection - 2

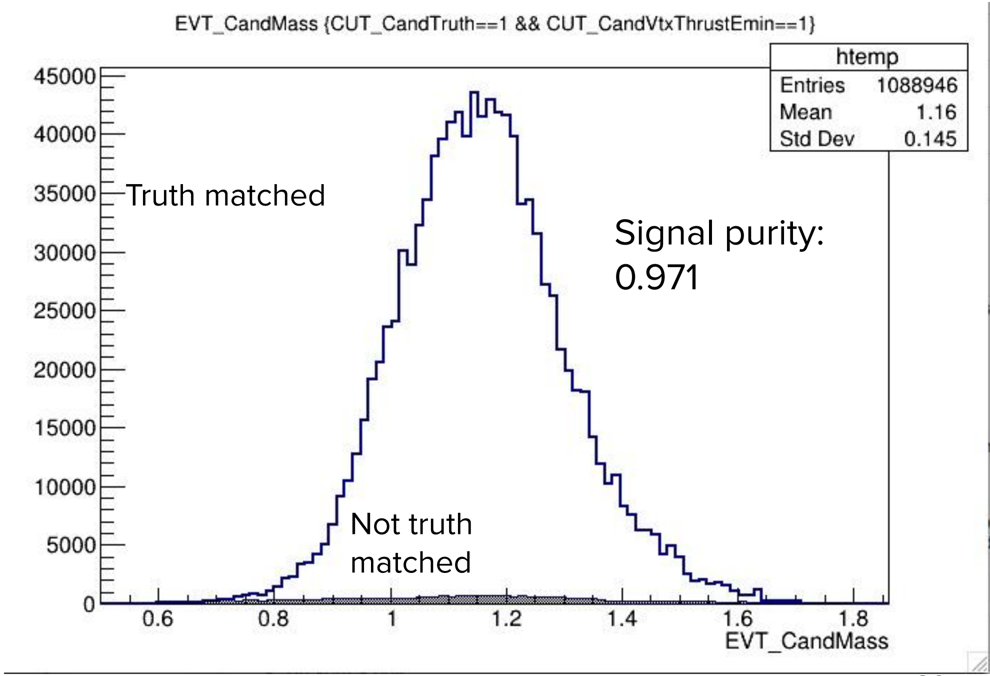
Select candidates with:

- m(3π) > 0.6 GeV and < 1.8 GeV
- Vertex chi2 > 0 and < 10
- If several candidates, chose smallest chi2
- Candidate in minimum E hemisphere

Non truth-matched 3π are candidates built in signal events which do not come from the signal **τ**

e.g. a charm meson decay from the non-signal hemisphere

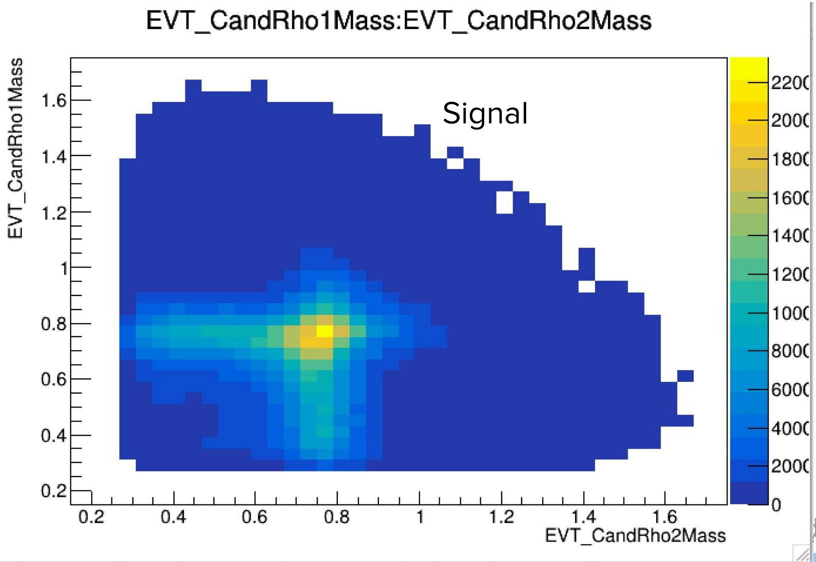
Rate is decreased by requiring the 3π to reside in minimum energy hemisphere

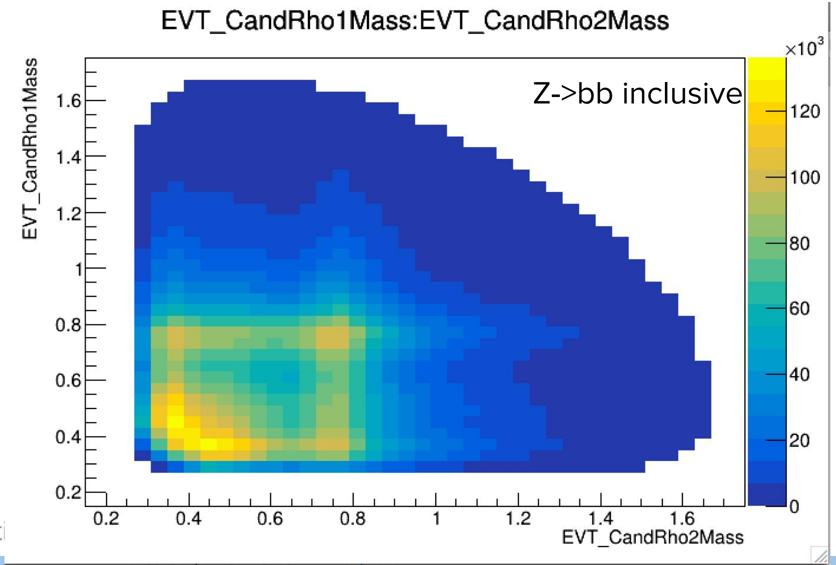


3π candidate selection - 3

Select candidates with a cut on m(ππ) mass:

- At truth level in signal, L-shape seen
- Characteristic of the τ → 3π ν decay
- m(ππ) > 0.6 GeV and < 1GeV
- Select an L shape around signal
- **Reject lower box and outer region**

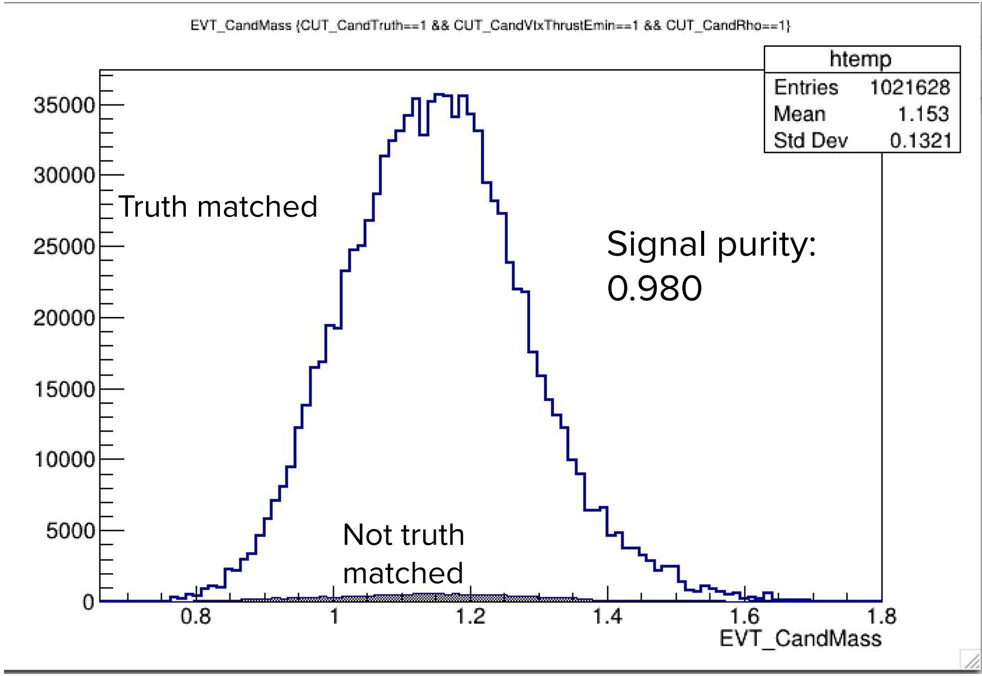




3π candidate selection - 3

Select candidates with a cut on m(πππ) mass:

- At truth level in signal, L-shape seen
- Characteristic of the $\tau \rightarrow 3\pi\,\nu$ decay
- m(πππ) > 0.6 GeV and < 1GeV
- Select an L shape around signal
- **Reject lower box and outer region**



3π candidate selection - summary

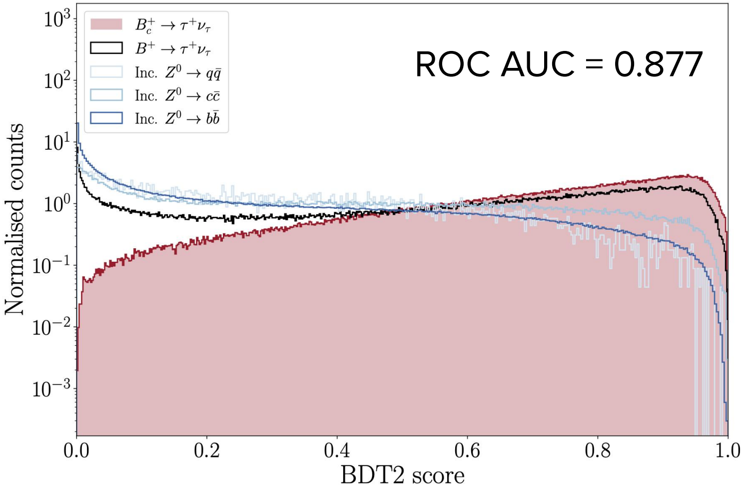
Efficiency cumulative and wrt to
last stage 1 cut

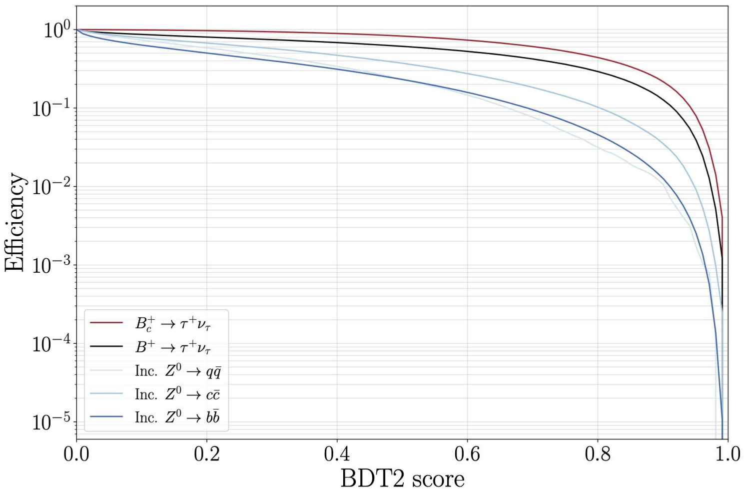
Cut for candidate selection	Efficiency $B_c \to \tau\nu$	Purity $B_c \to \tau\nu$
0.6<mass<1.8GeV 0<chi2<10	0.983	0.875
Candidate in thrust hemisphere with less energy	0.815	0.971
(m1(ππ)<1 && m2(ππ)>0.6 && m2(ππ)<1 GeV) OR (m2(ππ)<1 && m1(ππ)>0.6 && m1(ππ)<1 GeV)	0.758	0.980

Second-stage MVA at 3π candidate level

After best candidate selection per-event, train a second MVA with 3π candidate information such as flight distance, m(3π), m(ππ), momentum, vertex chi2

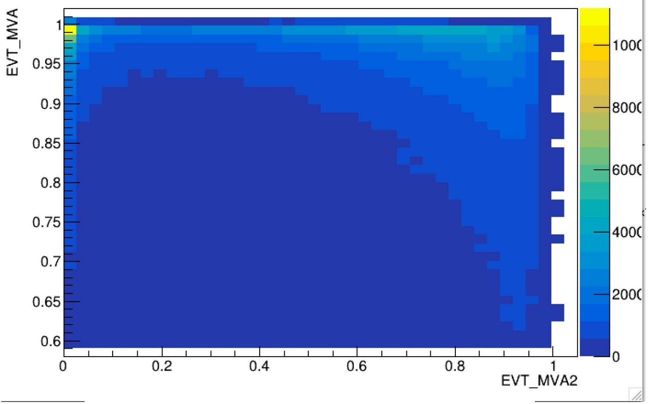
Use truth-matched signal and a combined background sample of Z ➔ qq, cc, bb passing all previous cuts (relative sizes set by branching ratios and efficiencies)



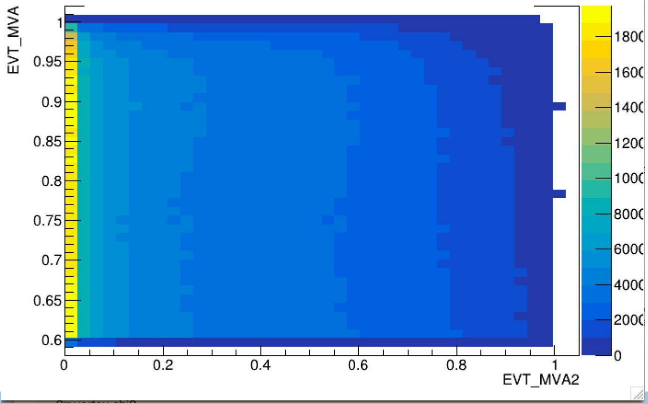


MVAs orthogonality

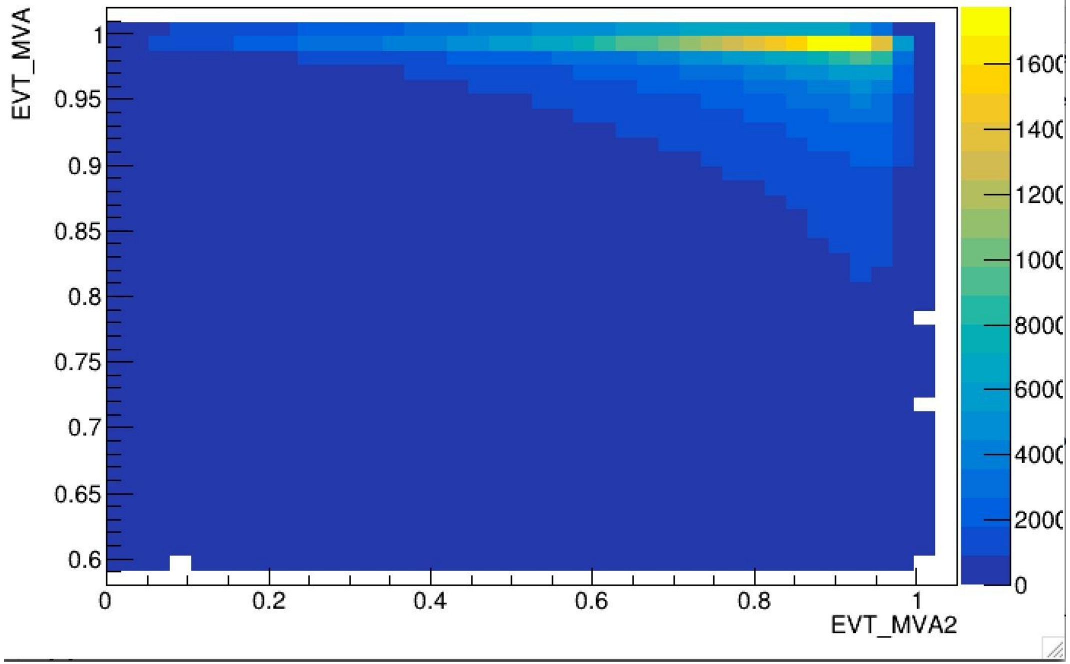
Bu -> tau nu



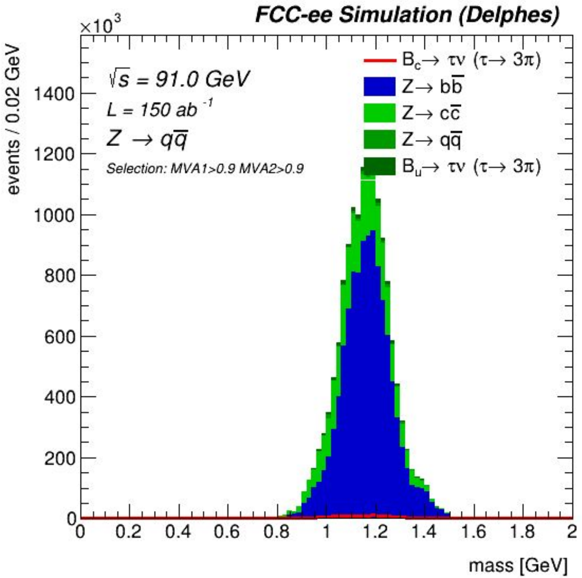
Z->bb inclusive



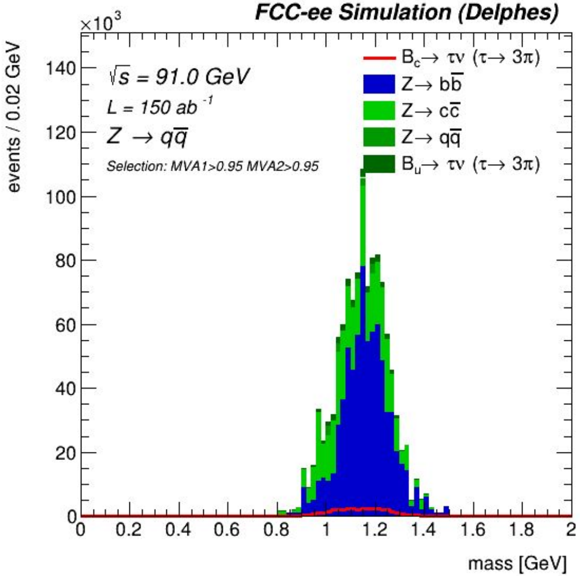
Bc -> tau nu



Tight MVAs selection - 1

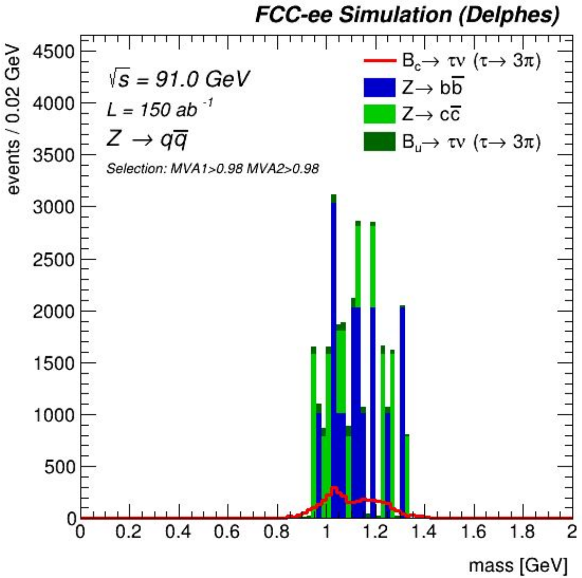


	Events	Raw MC
Bc	142 300	162 441
Bu	222 600	50 920
Zbb	10 973 000	10 840
Zcc	3 264 000	4 130
Zuds	125 600	45

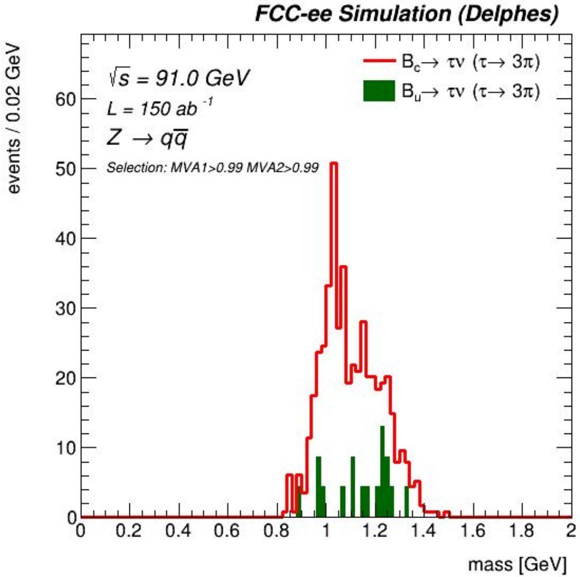


	Events	Raw MC
Bc	37 890	43 240
Bu	33 650	7 698
Zbb	729 880	721
Zcc	355 670	450
Zuds	11 170	4

Tight MVAs selection - 2



	Events	Raw MC
Bc	3 450	3 937
Bu	1 260	288
Zbb	16 200	16
Zcc	11 850	15
Zuds	0	0



	Events	Raw MC
Bc	485	554
Bu	73	17
Zbb	0	0
Zcc	0	0
Zuds	0	0

Other variables to investigate

- Displaced vertex requirements in non-signal hemisphere
  - Require high flight and mass consistent with a B-hadron
  - Would retain events more consistent with $Z \rightarrow bb$
- Candidate angles with respect to thrust axis
  - 3π produced from tertiary vertex, so τ flight direction is not along thrust axis
  - 3π produced directly in B decays will align well with thrust axis
    - e.g. charmless $B+ \rightarrow 3\pi$, large branching ratio $B \rightarrow D\ 3\pi$ decays
- Presence of a vertex consistent with charm in the signal hemisphere
  - Bc contains charm quark, so the other charm quark will produce a charm hadron
  - Look for another displaced vertex consistent with charm flight and mass
  - $c \rightarrow s$ preferentially, so look for displaced kaons (both charged track and $Ks \rightarrow \pi\pi$)

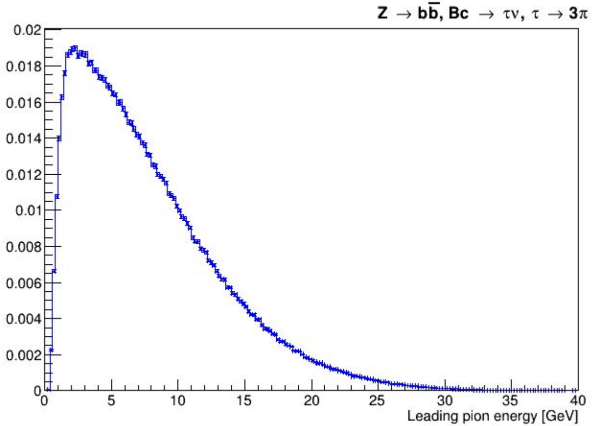
Summary

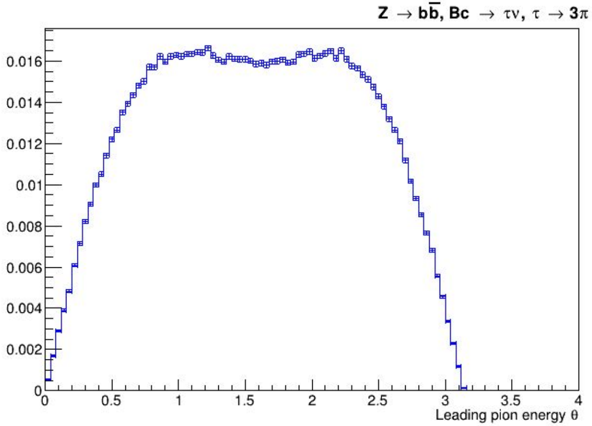
- Close to final analysis flow presented today
  - Perfect PID assumed
  - Perfect vertex seeding assumed
  - Achieved an excellent signal purity selection - enough signal for a 5% yield stat. uncertainty

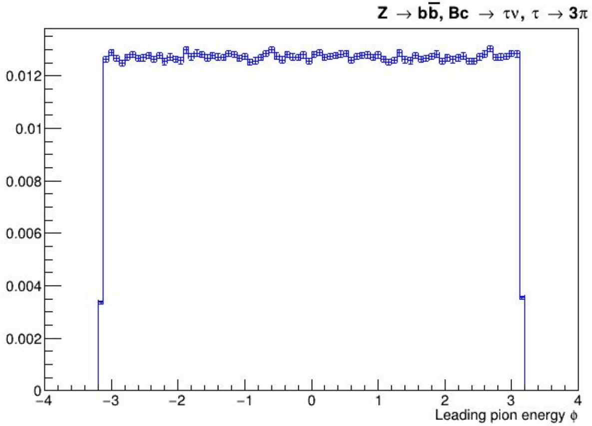
- Next steps
  - Validate the perfect vertex seeding with a procedure that is already in place but not used here
    - Fit any vertex with 3 pions (perfect PID, combinatoric with awkward C++and) choose the best chi2 (as done actually). Should find very similar performances
  - Generate more MC (with official production, BES, etc)
  - Study more discriminating variables
  - Study impact of non perfect PID
  - Extract Bc branching fraction
  - Start to draft a paper

Backup

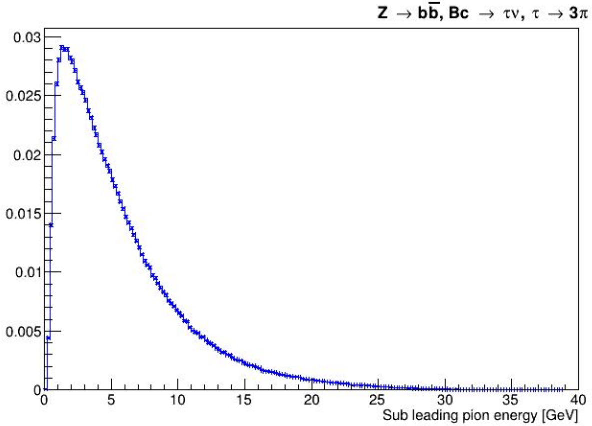
Truth level kinematic leading pion energy

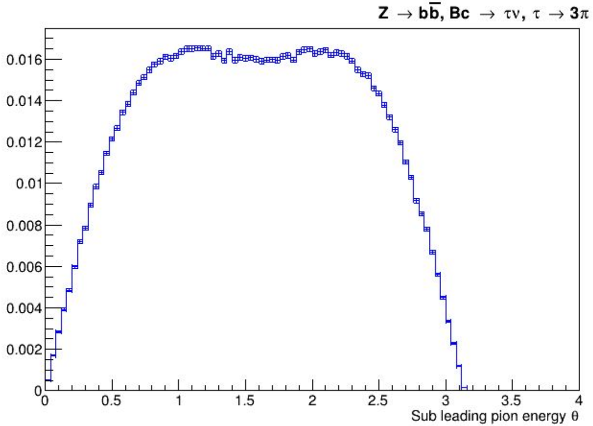


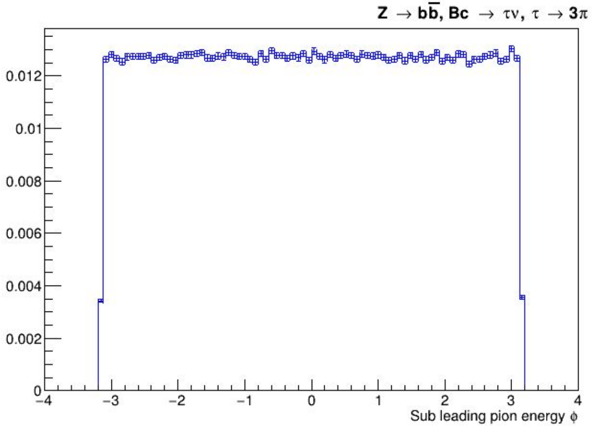




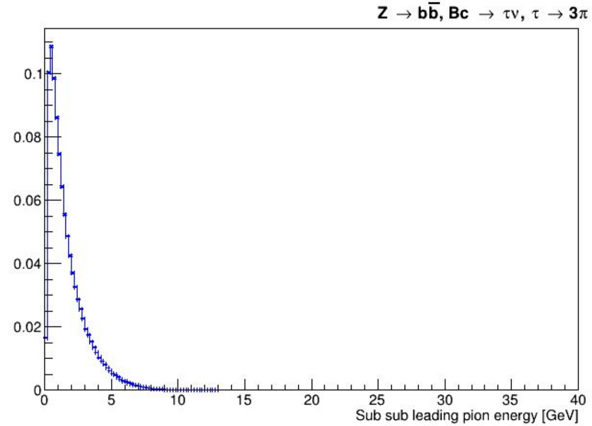
Truth level kinematic sub leading pion energy

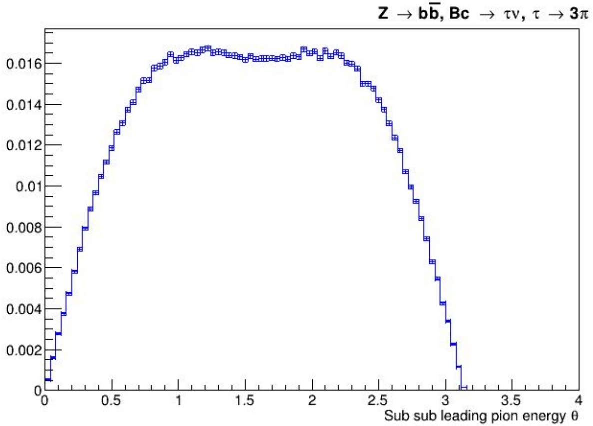


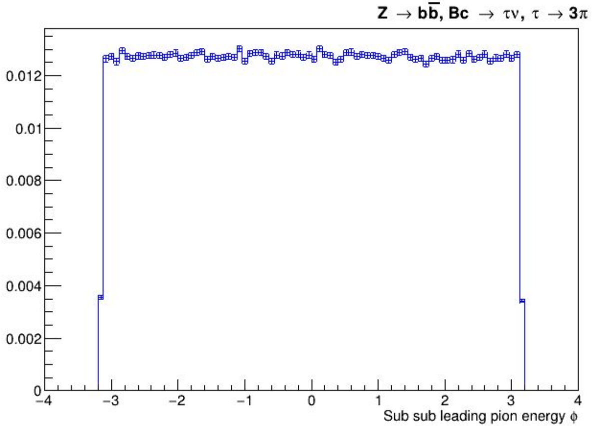




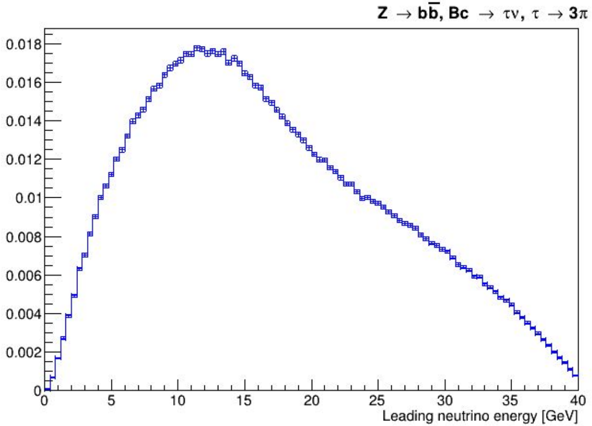
Truth level kinematic sub leading pion energy

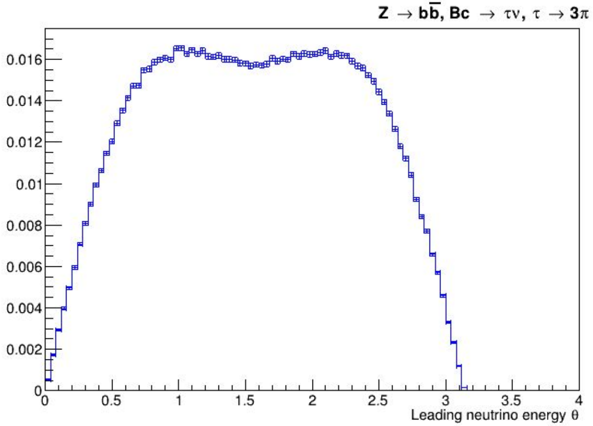


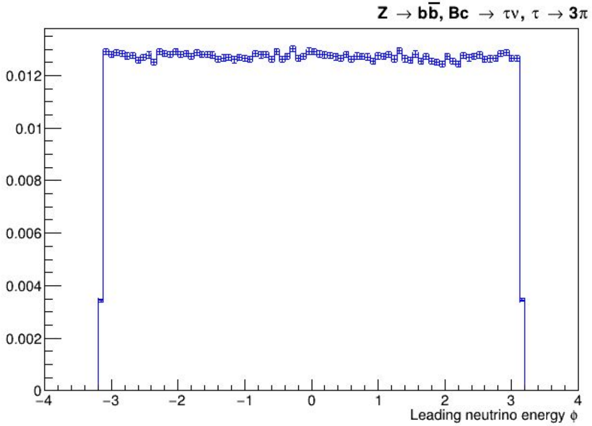




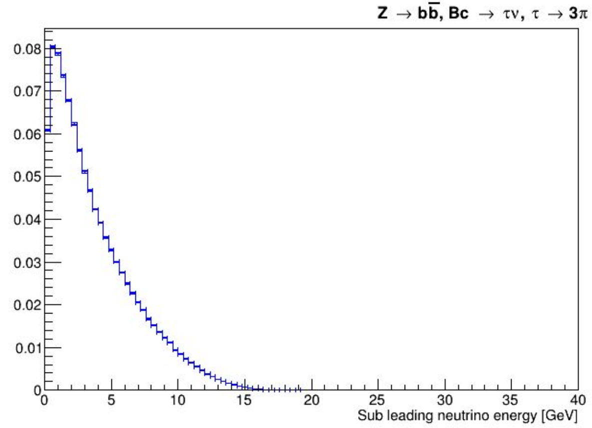
Truth level kinematic leading neutrino energy

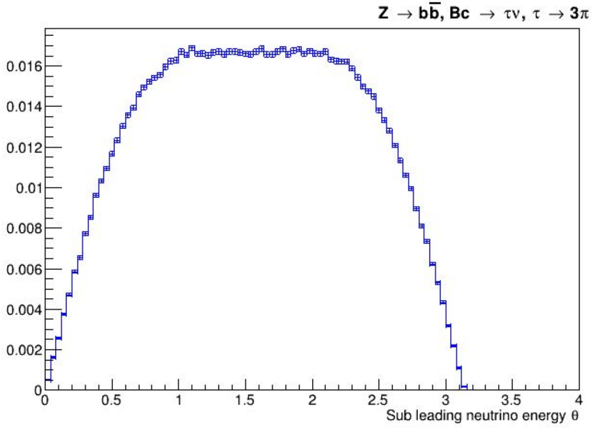


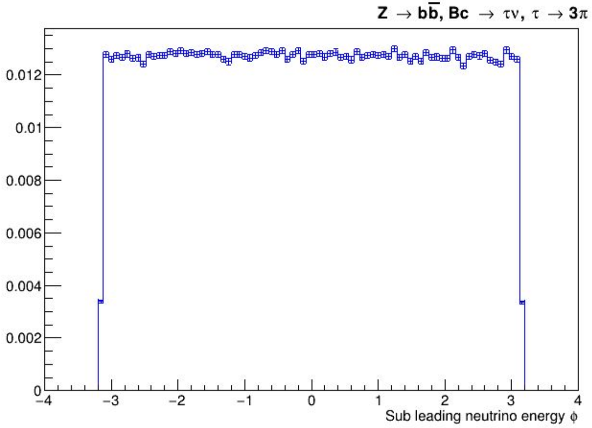




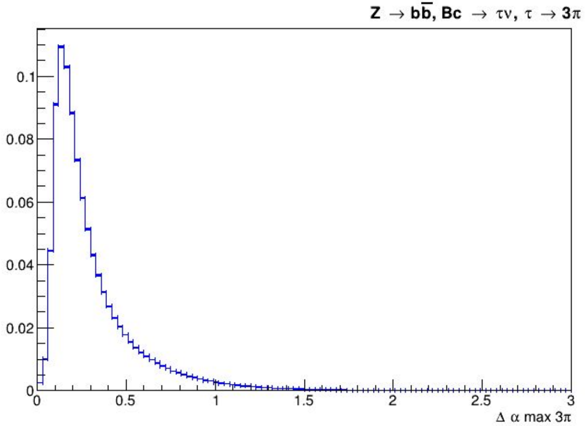
Truth level kinematic sub-leading neutrino energy

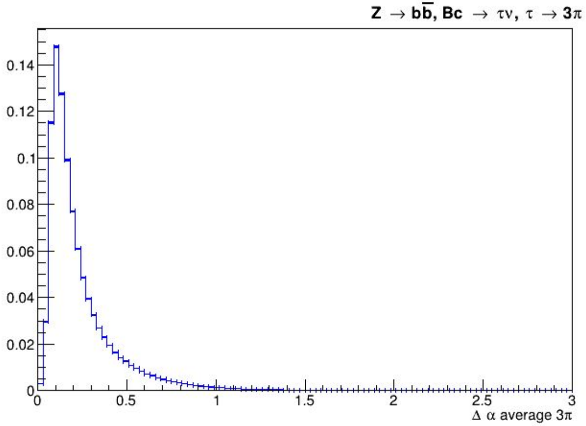


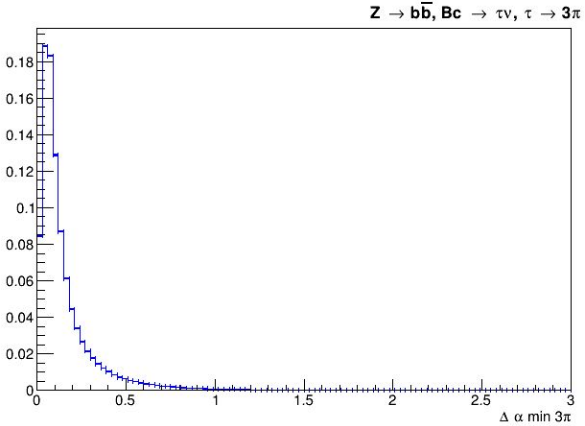




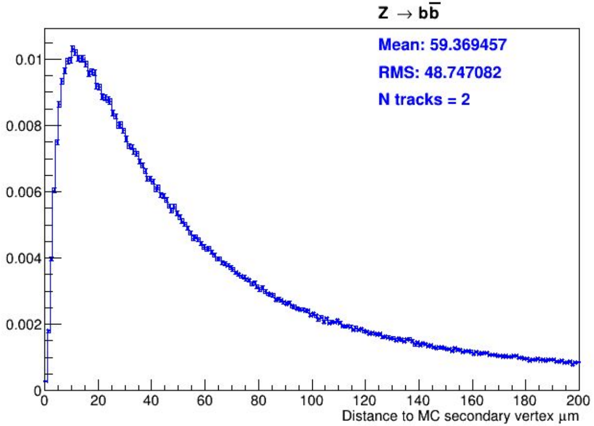
Angular separation between 3 pi reco level

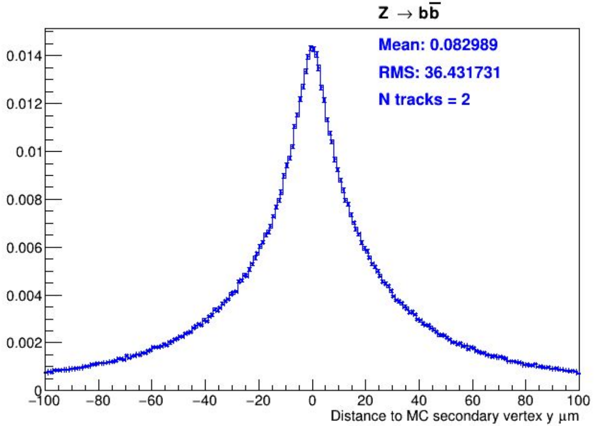


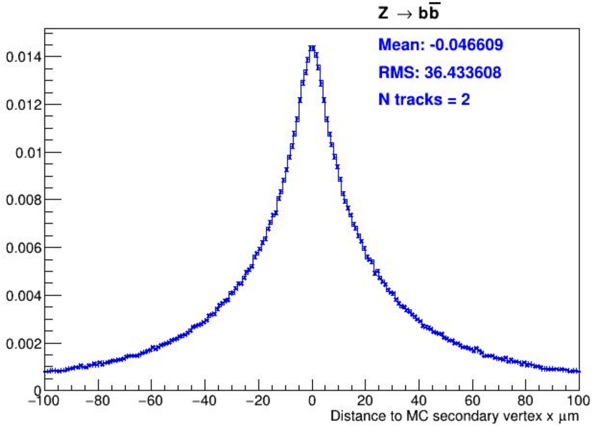


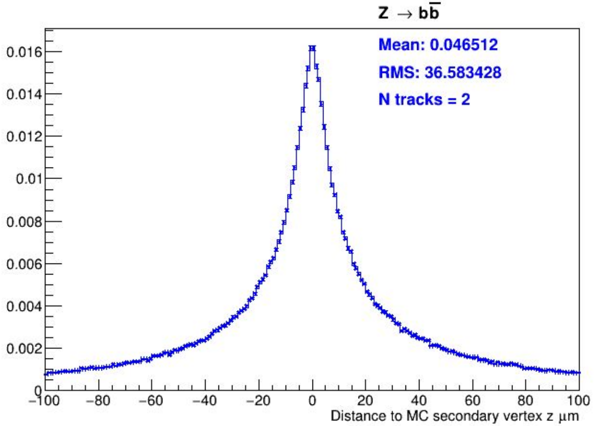


Distance to MC vertex - N tracks = 2

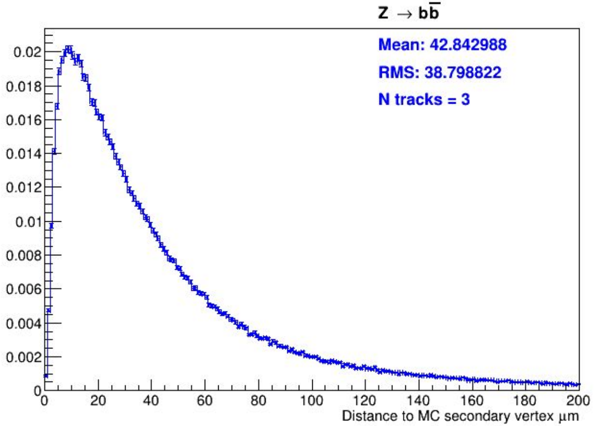


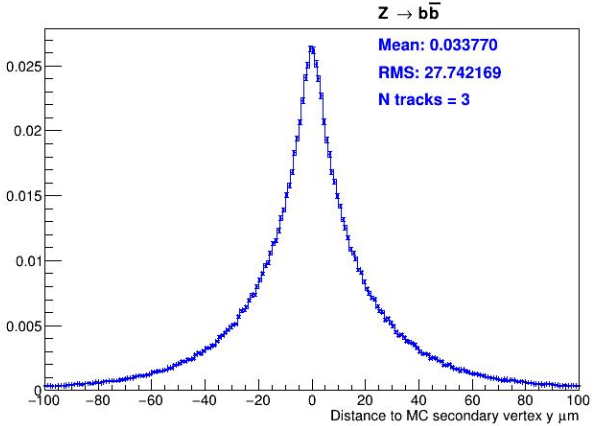


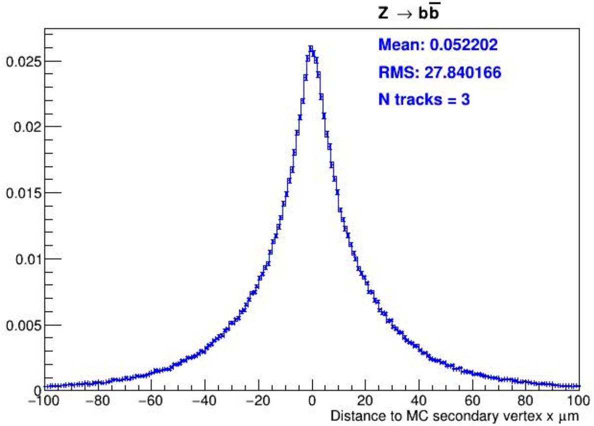


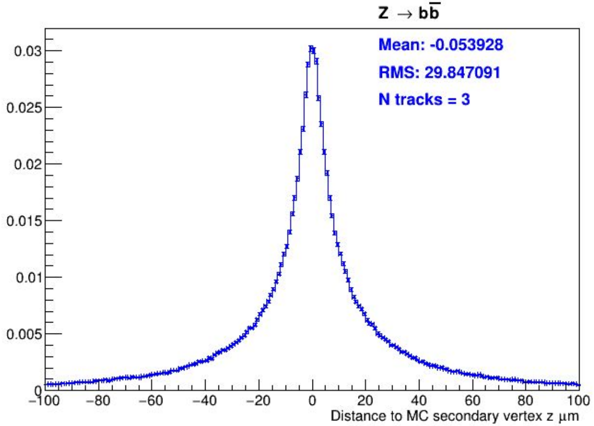


Distance to MC vertex - N tracks = 3

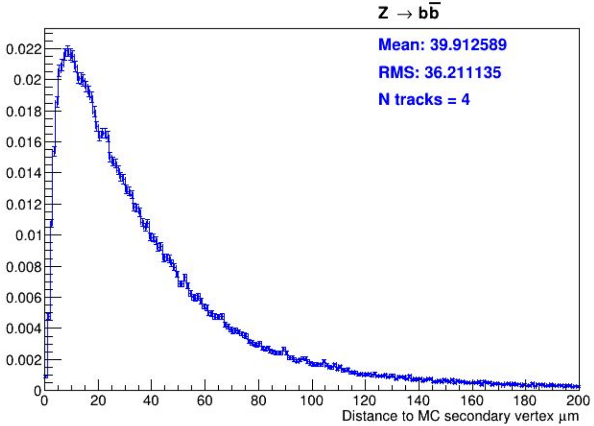


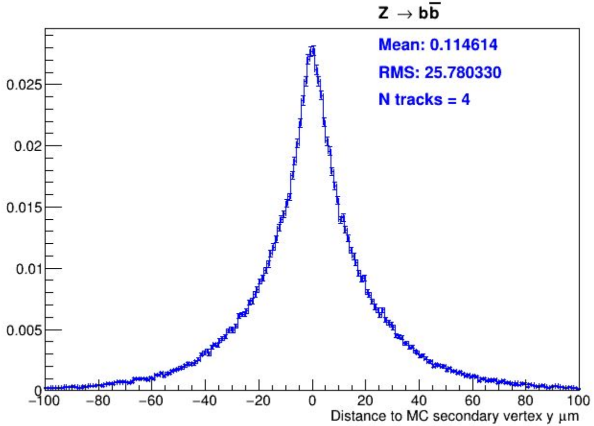


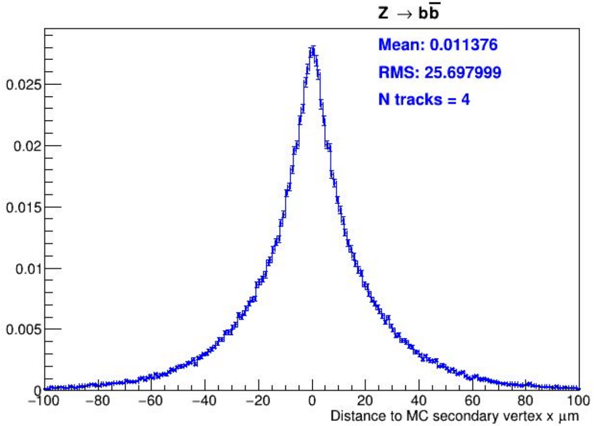


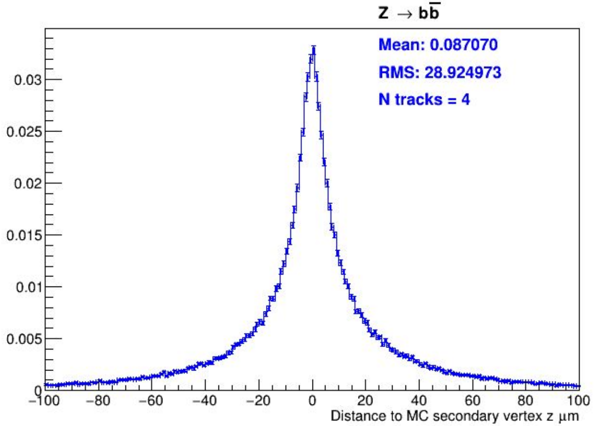


Distance to MC vertex - N tracks = 4

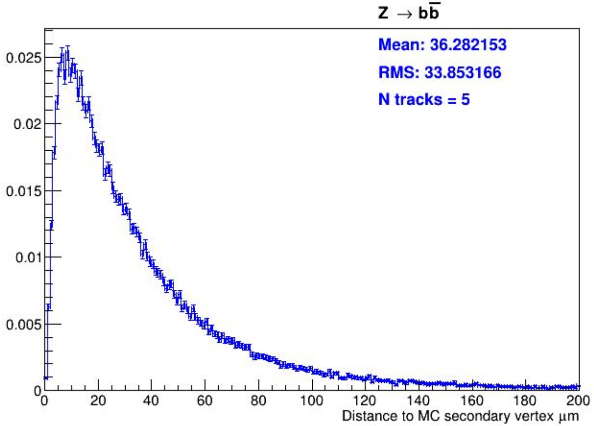


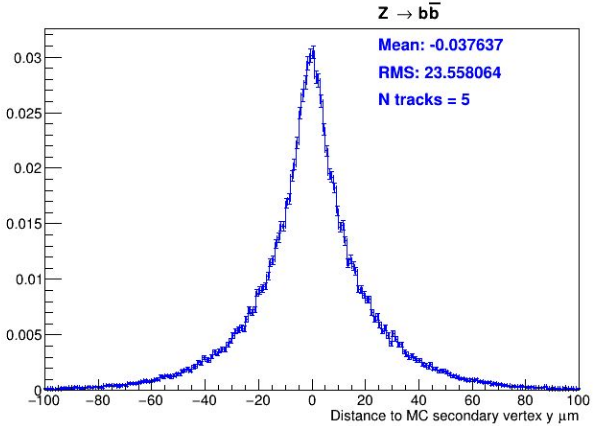


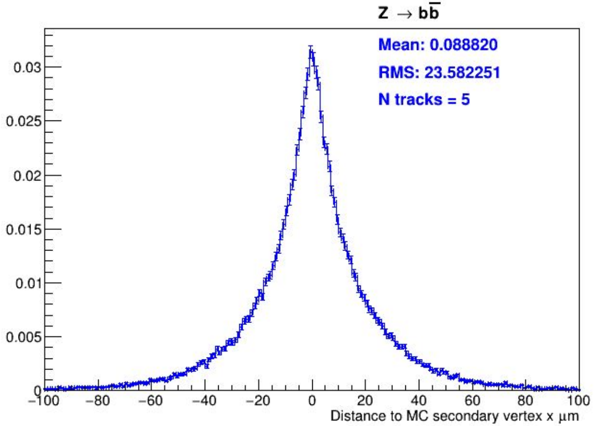


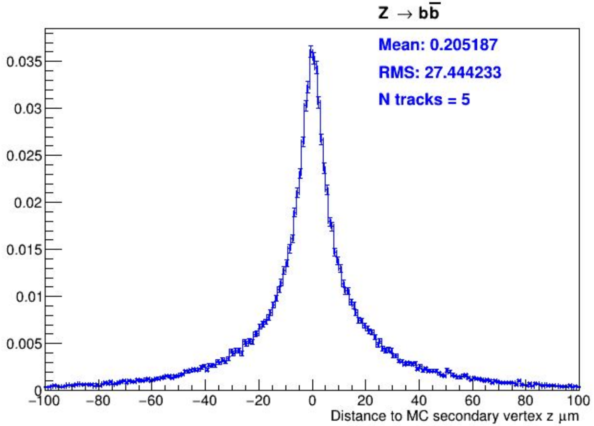


Distance to MC vertex - N tracks = 5

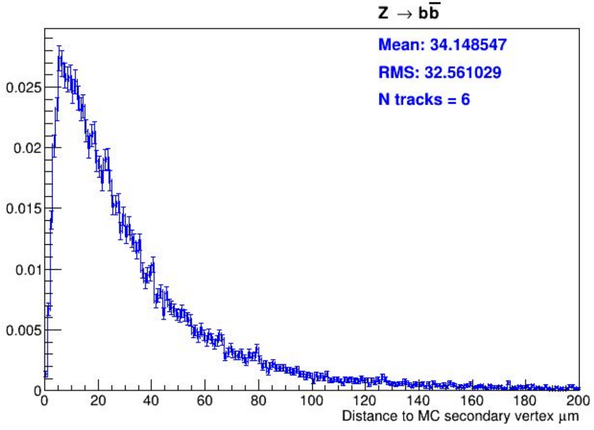


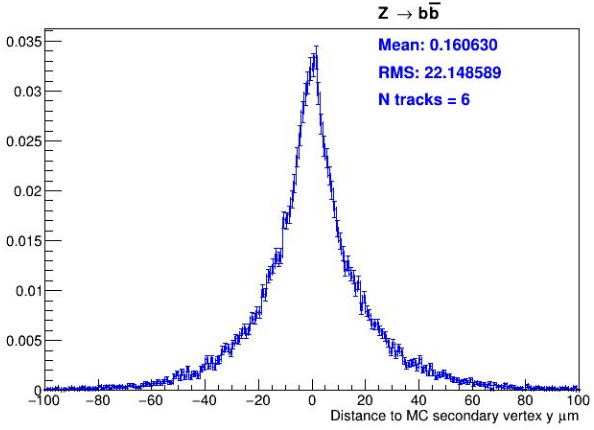


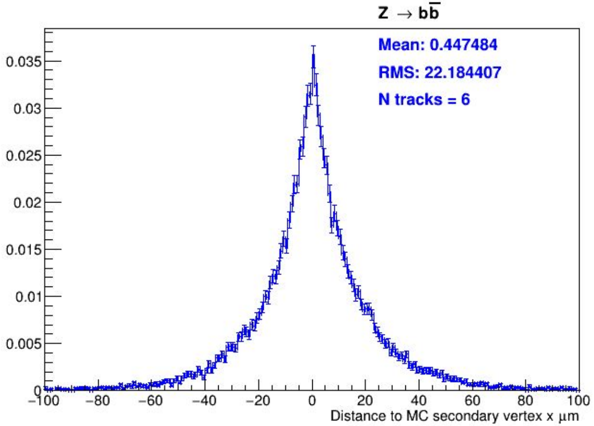


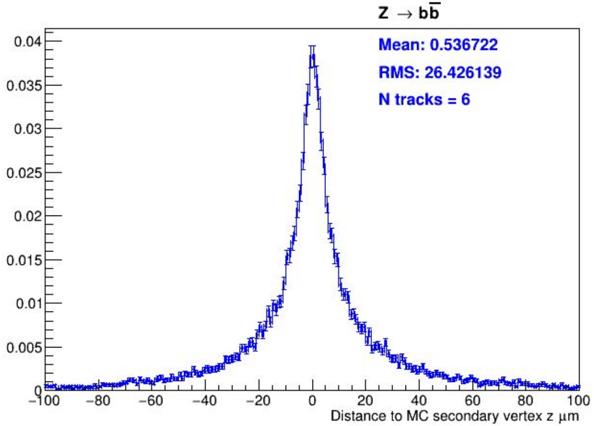


Distance to MC vertex - N tracks = 6

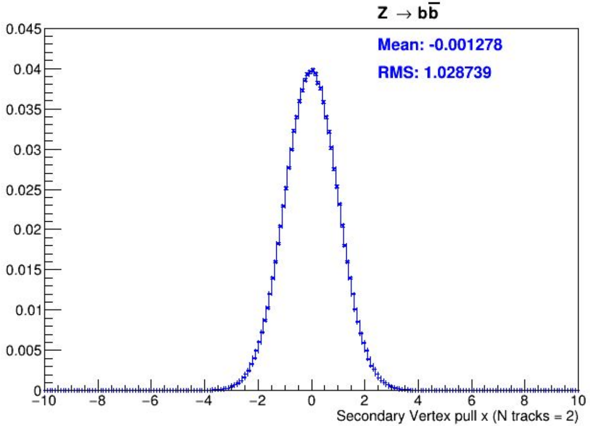


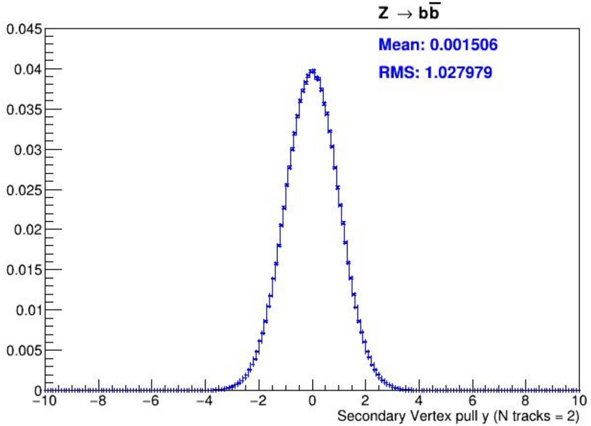


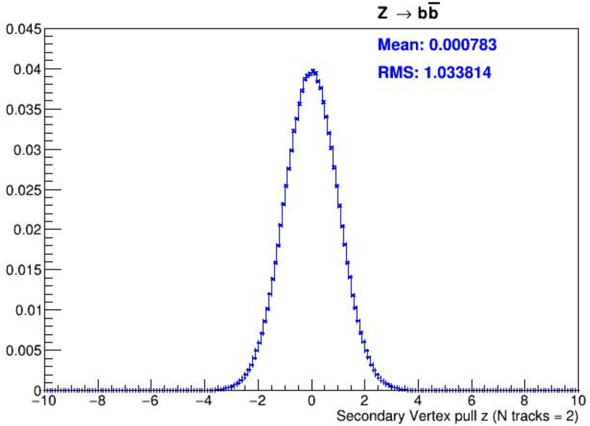




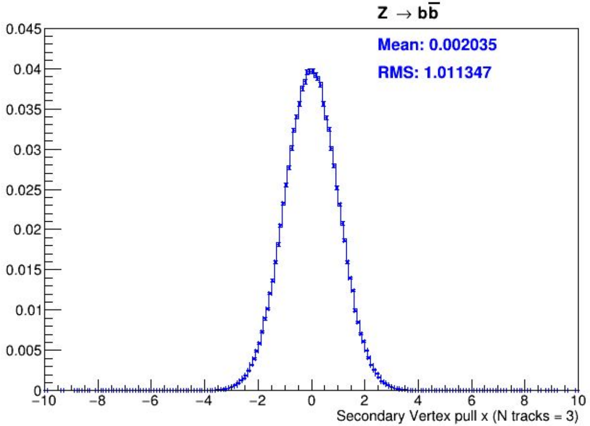
Vertex pull - N tracks = 2

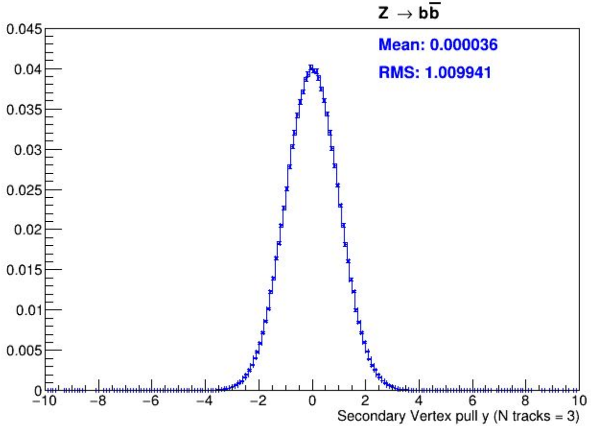


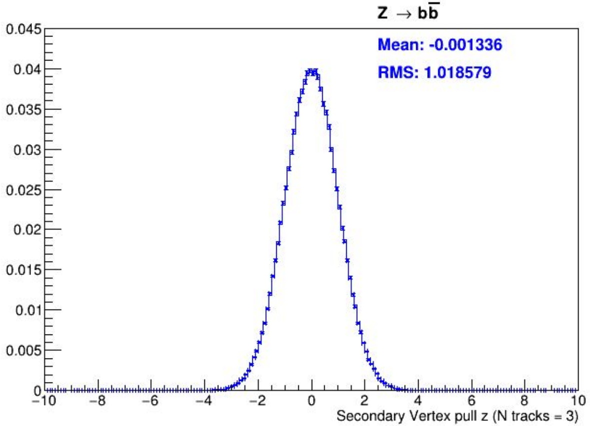




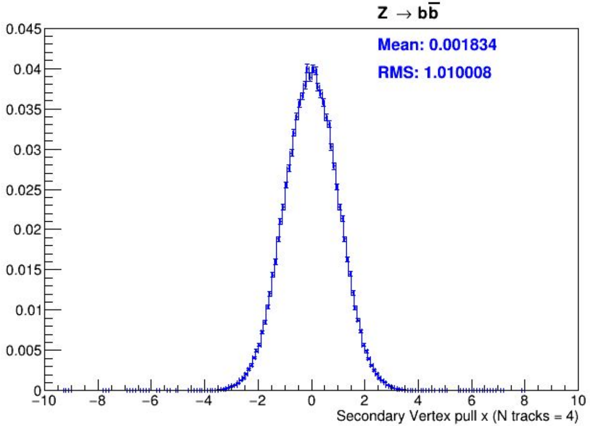
Vertex pull - N tracks = 3

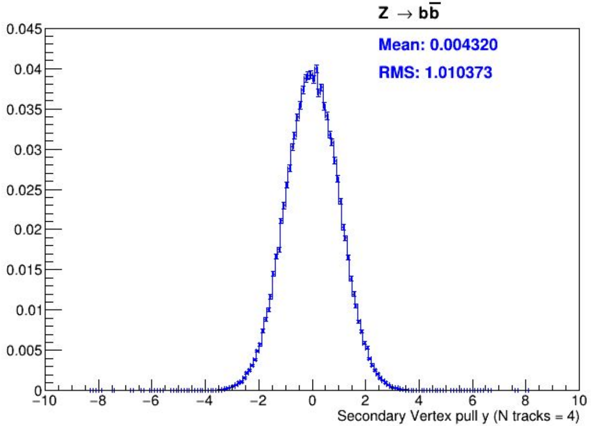


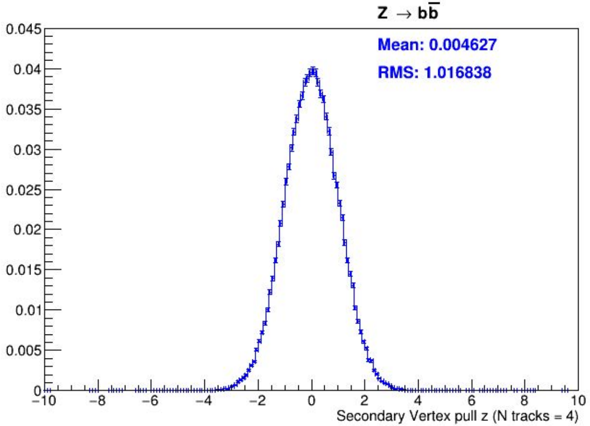




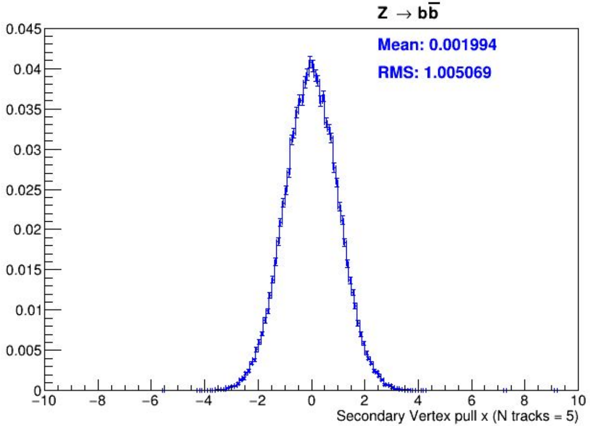
Vertex pull - N tracks = 4

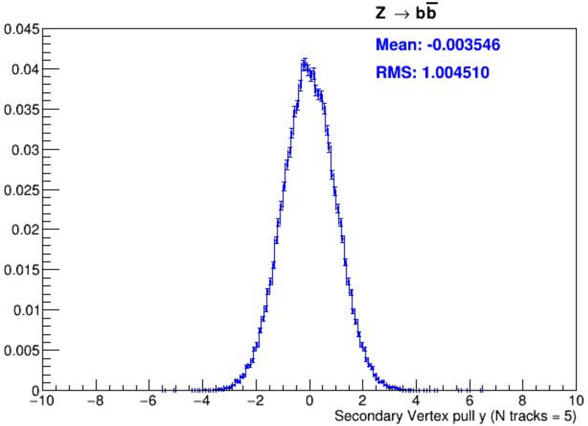


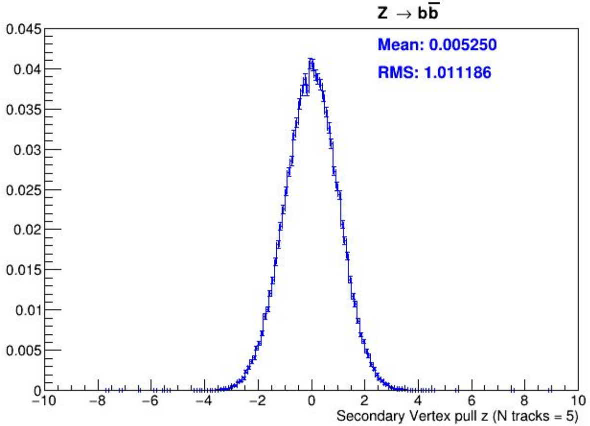




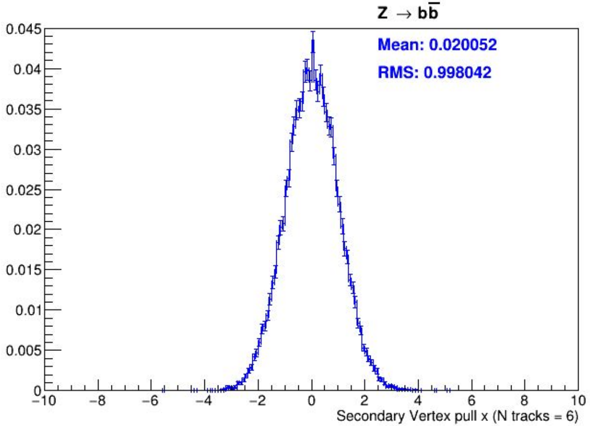
Vertex pull - N tracks = 5

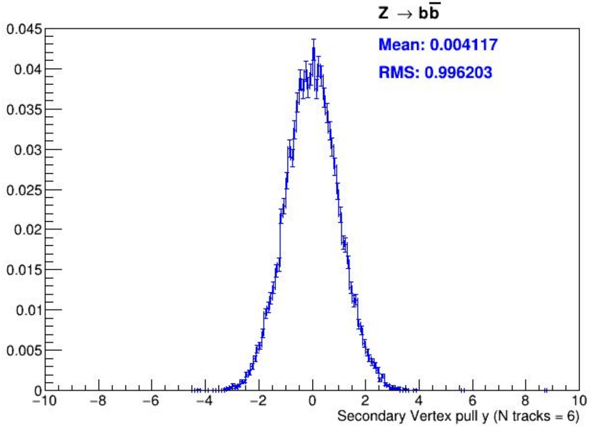


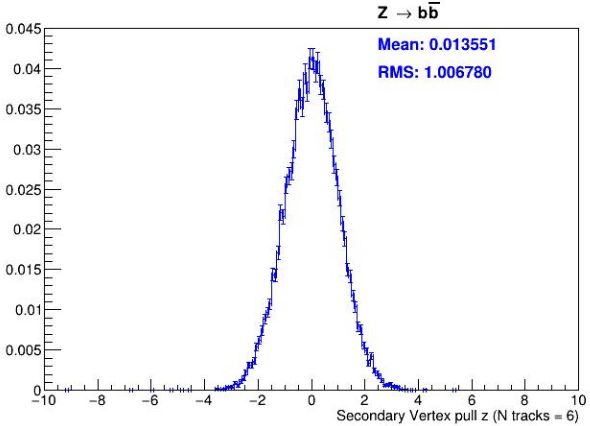




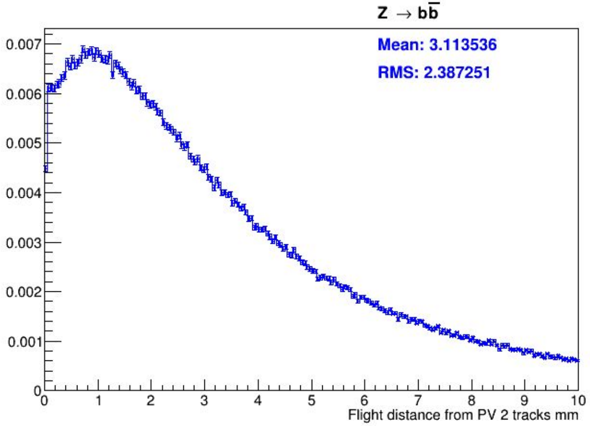
Vertex pull - N tracks = 6

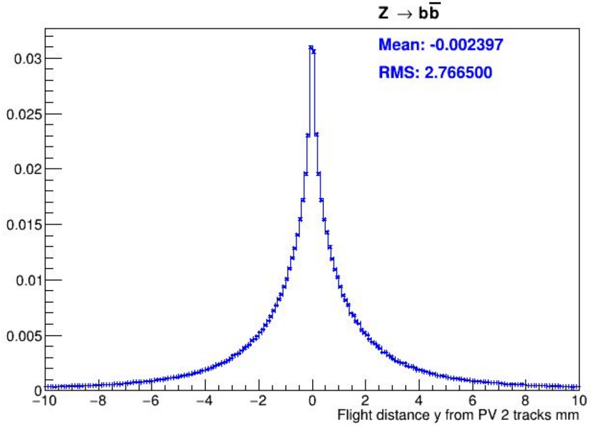


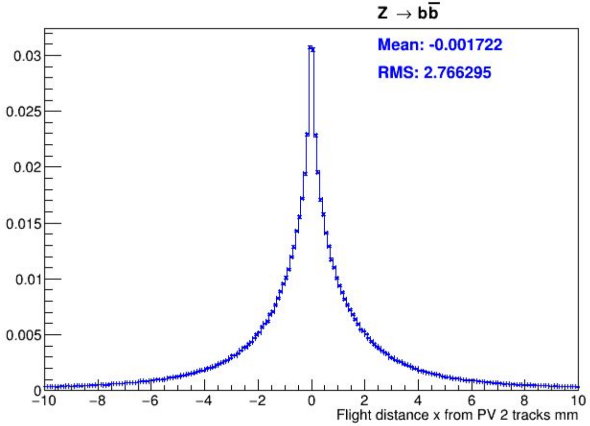


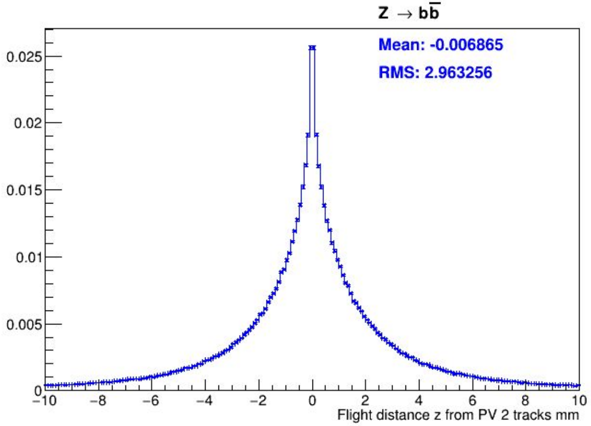


Flight distance - N tracks = 2

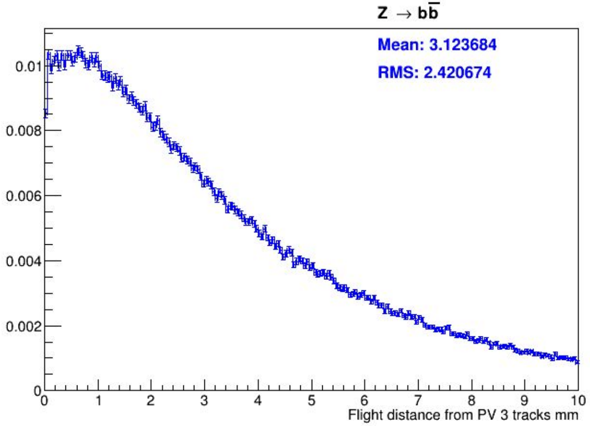


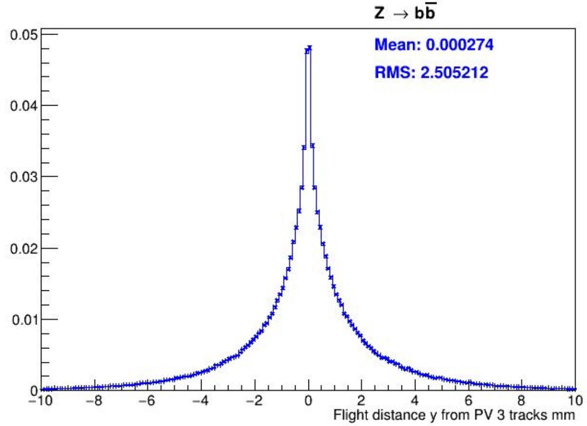


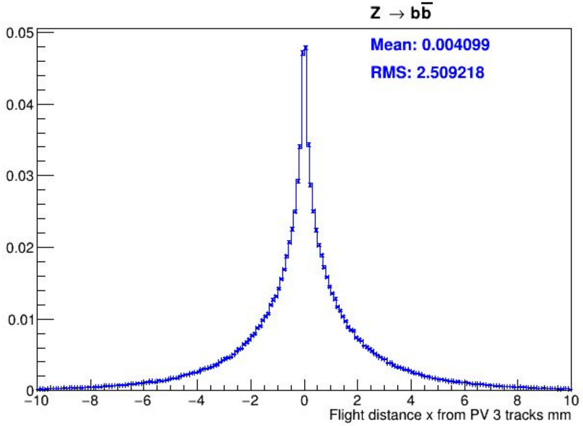


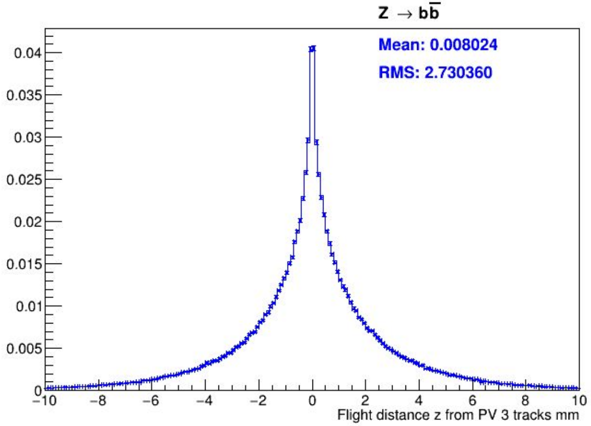


Flight distance - N tracks = 3

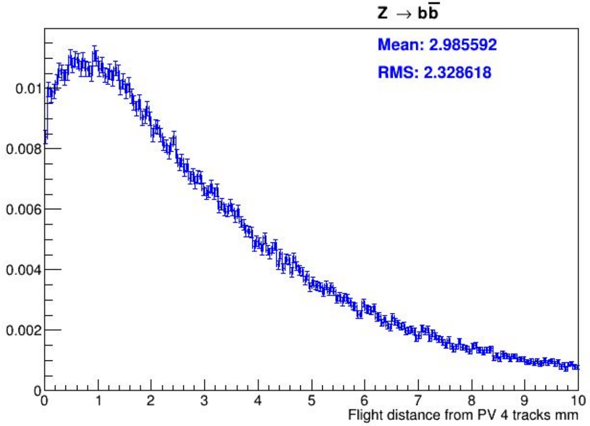


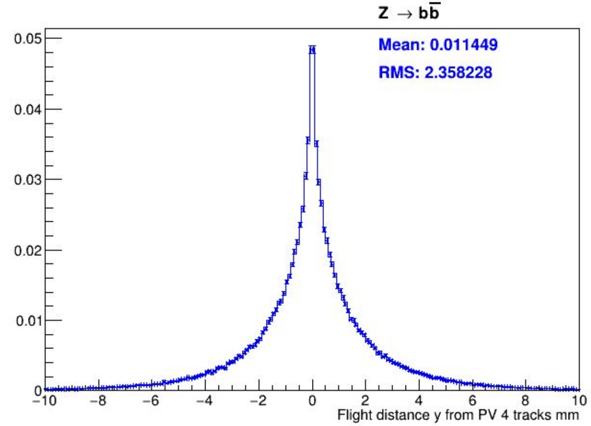


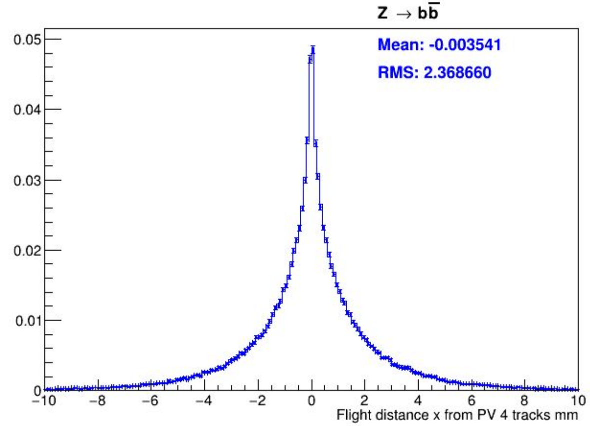


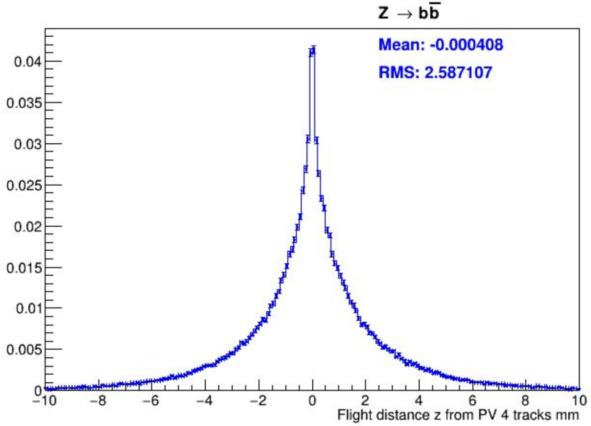


Flight distance - N tracks = 4

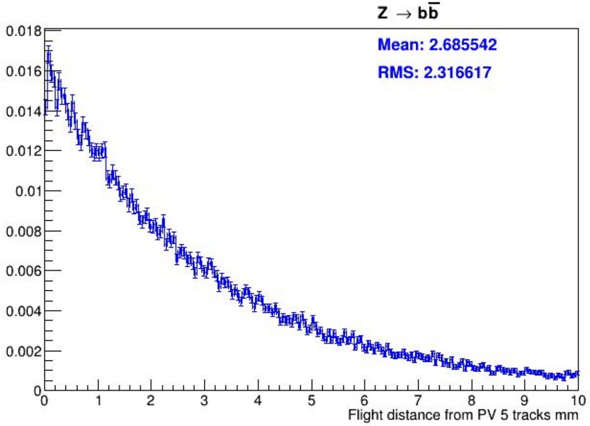


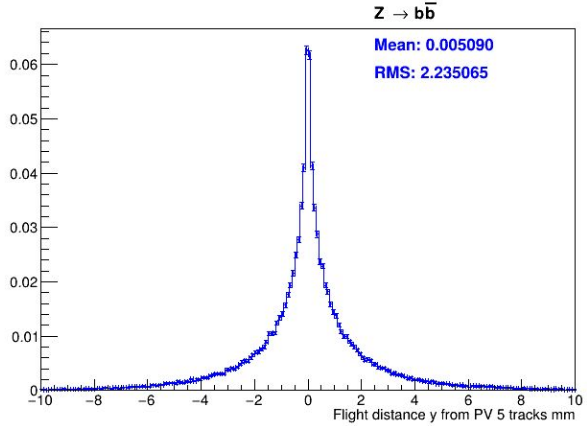


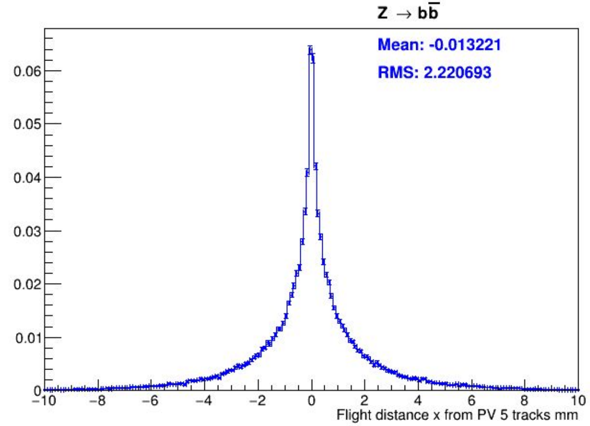


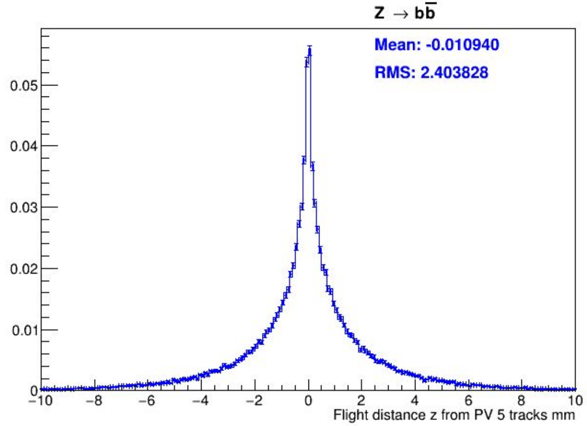


Flight distance - N tracks = 5

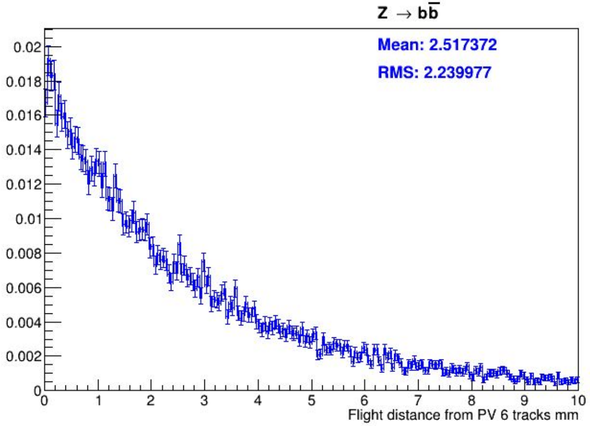


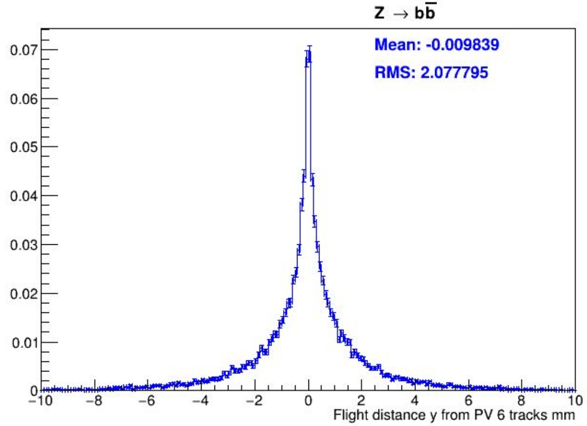


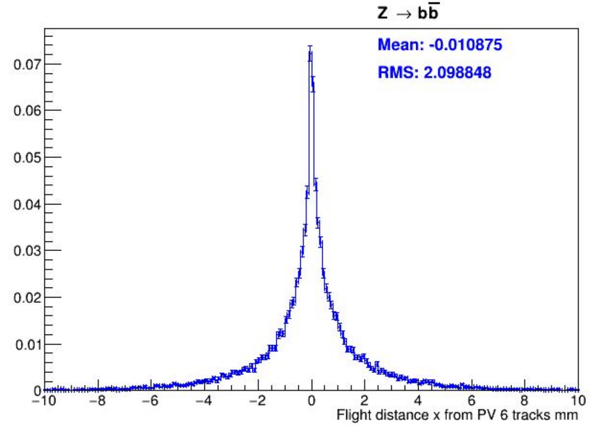


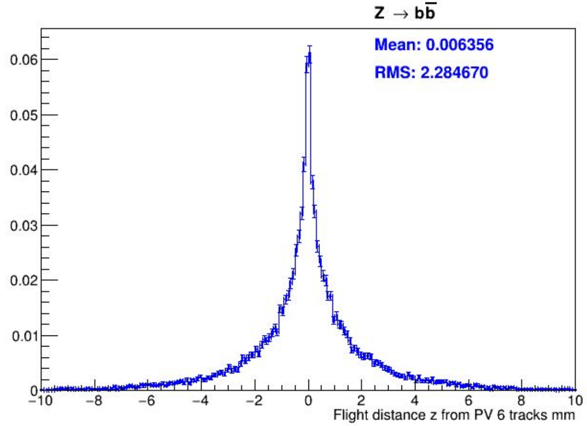


Flight distance - N tracks = 6

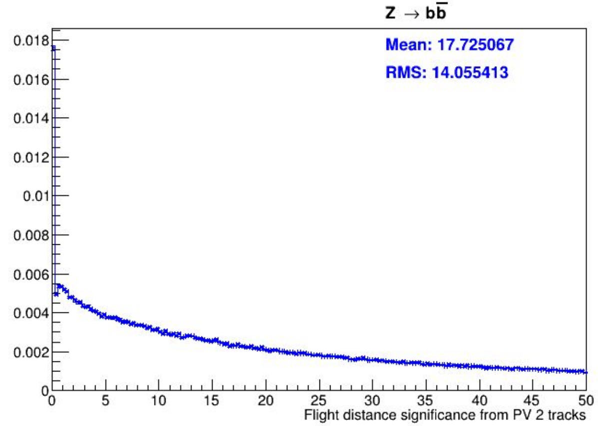


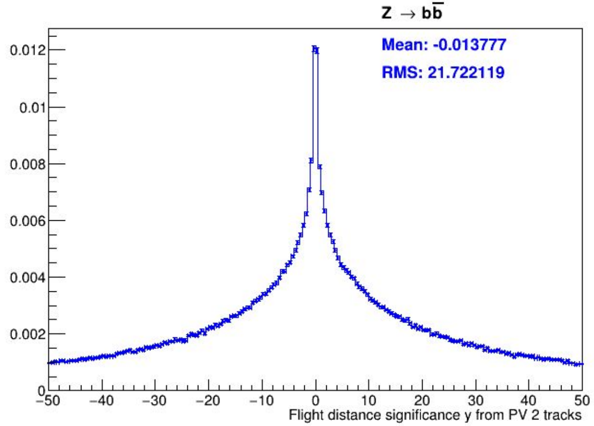


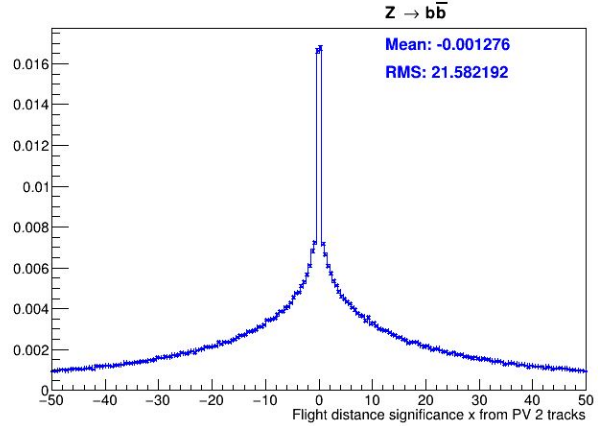


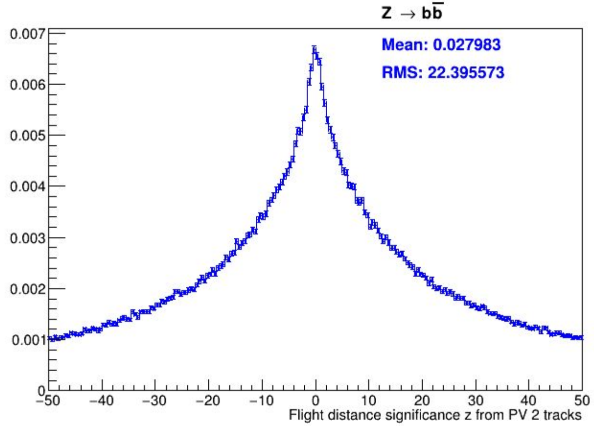


Flight distance significance - N tracks = 2

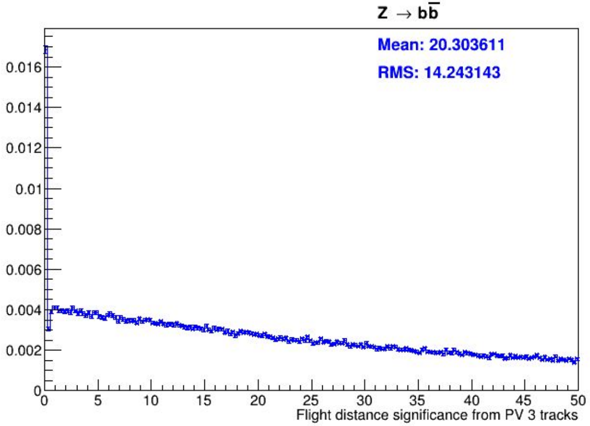


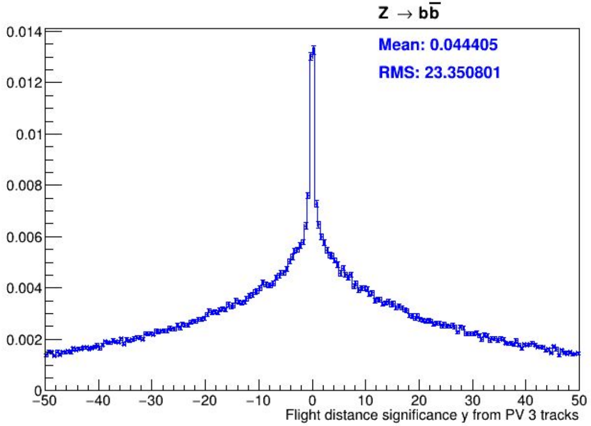


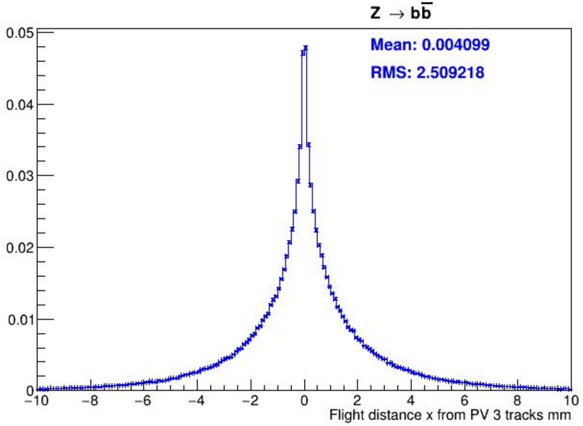


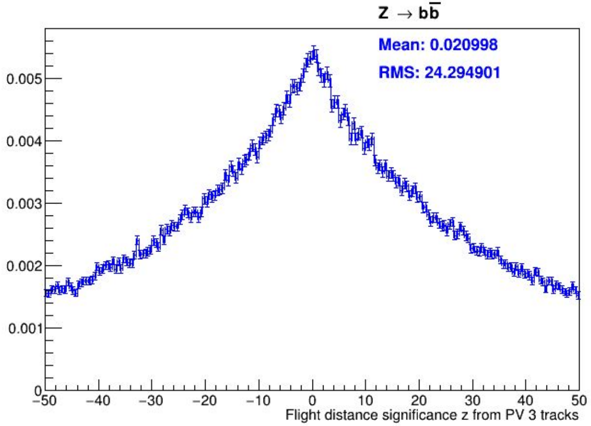


Flight distance significance - N tracks = 3

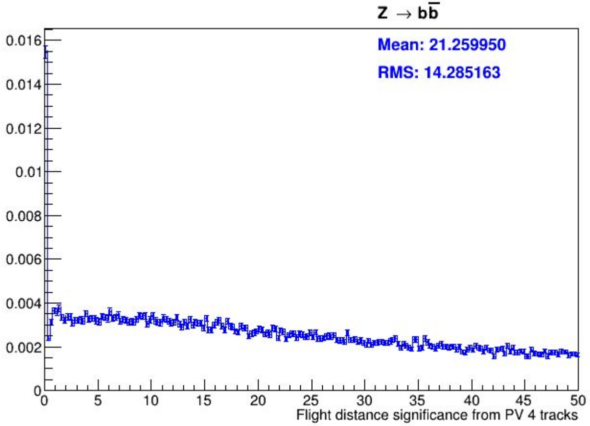


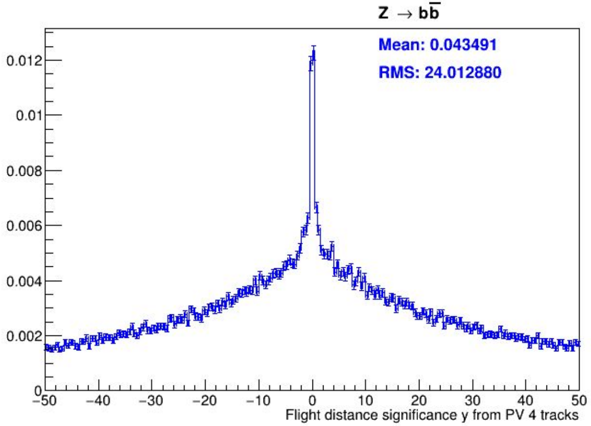


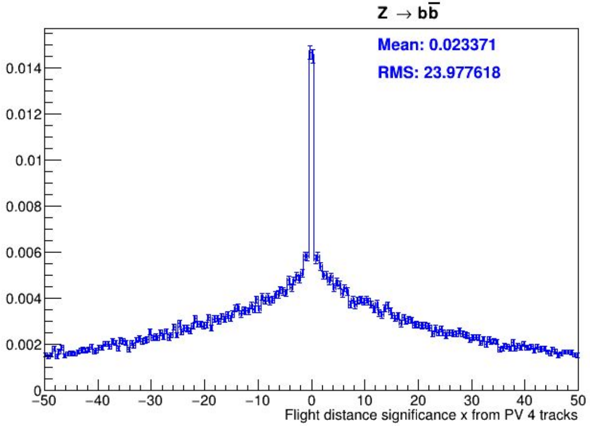


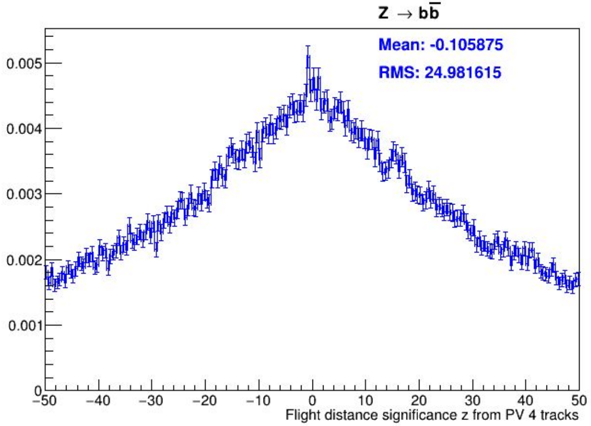


Flight distance significance - N tracks = 4

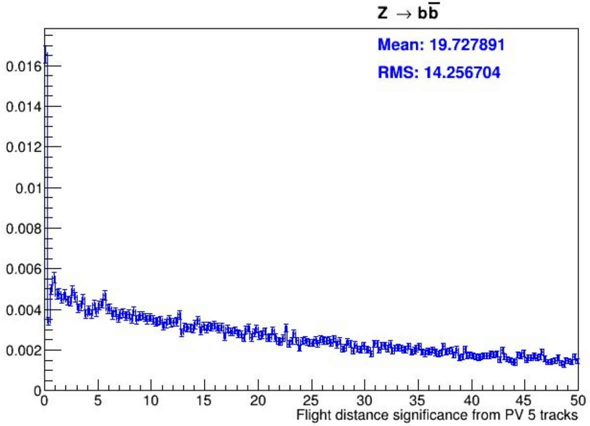


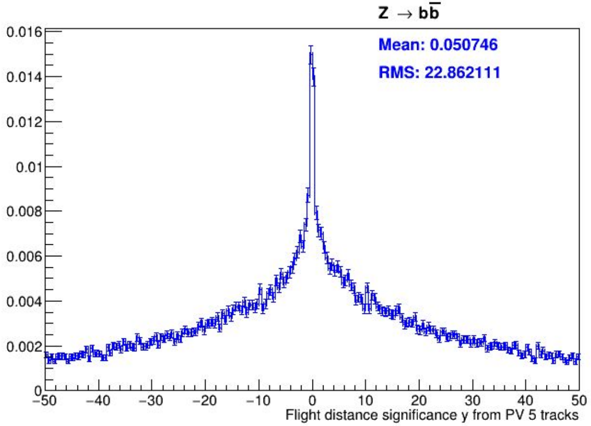


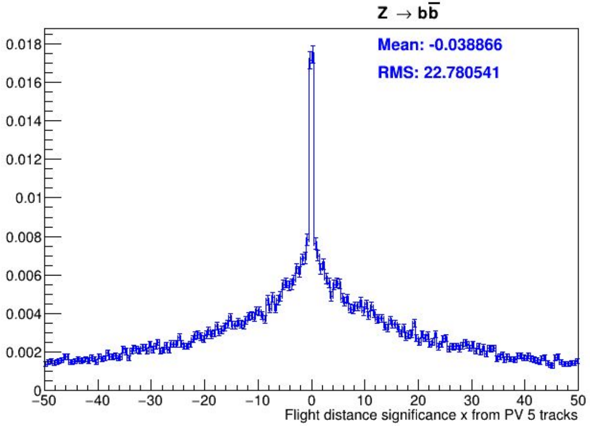


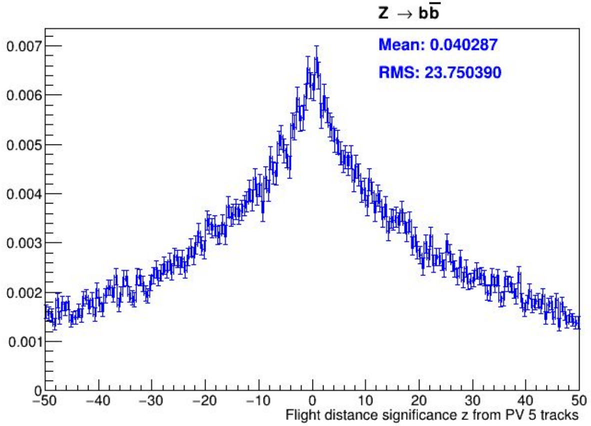


Flight distance significance - N tracks = 5









Flight distance significance - N tracks = 6