FP 93/94 Studying the Z-Boson with the ATLAS Detector at the LHC

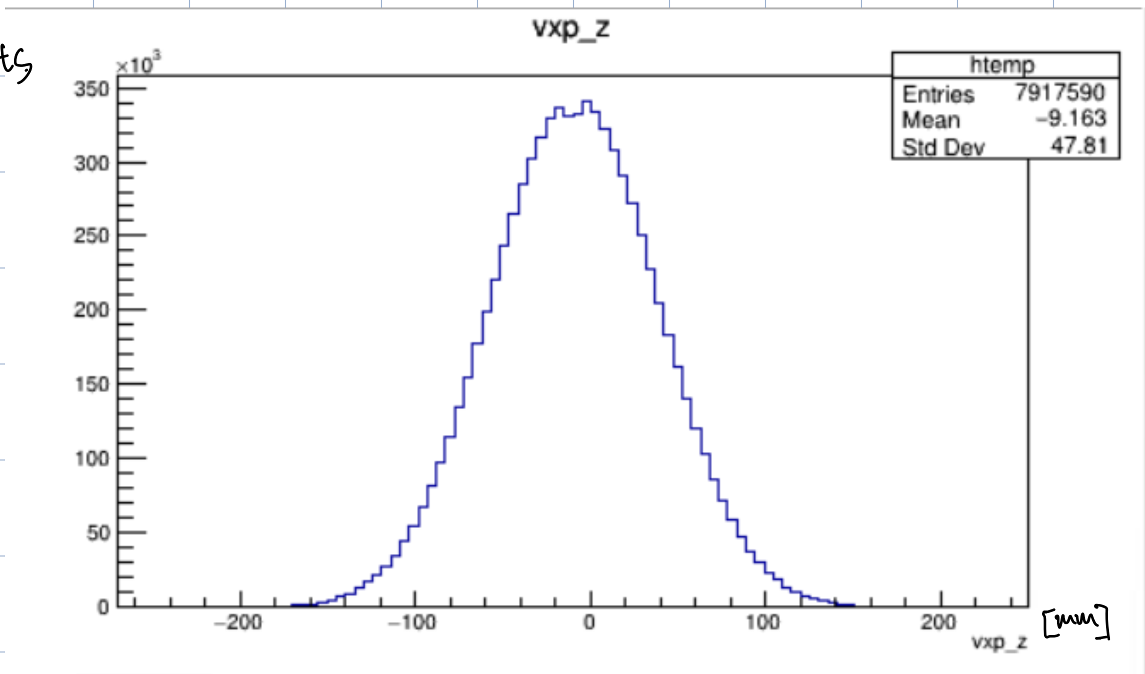
16.09.2024 ~ 17.09.2024 9:00 ~ 17:00

Yuting Shi Yulai Shi

1. Copy Python scripts: 3 Py-Scripts eventloop.py fit.py plot.py are copied into Heibox-Portal

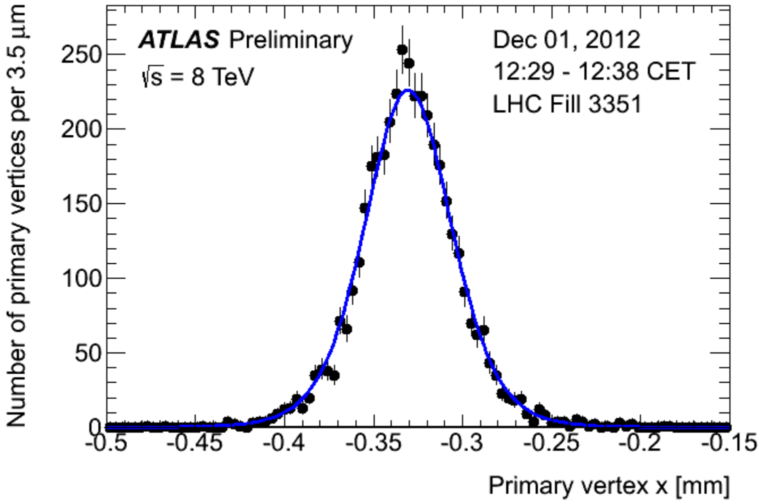
2. Structure of ROOT-Data Files: using TBrowser

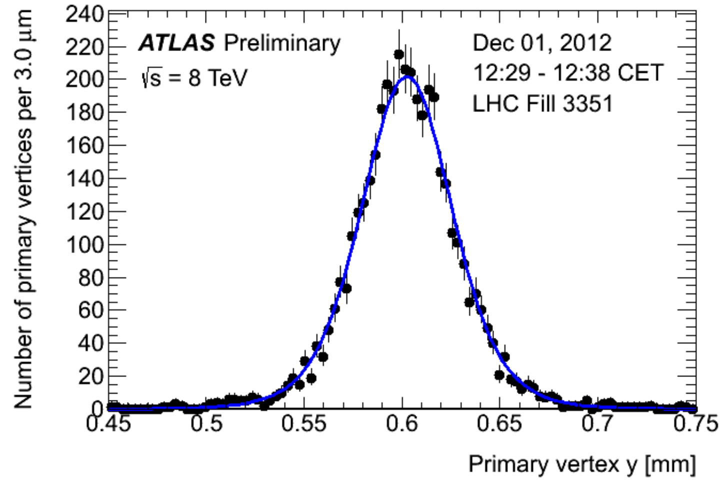
accidents

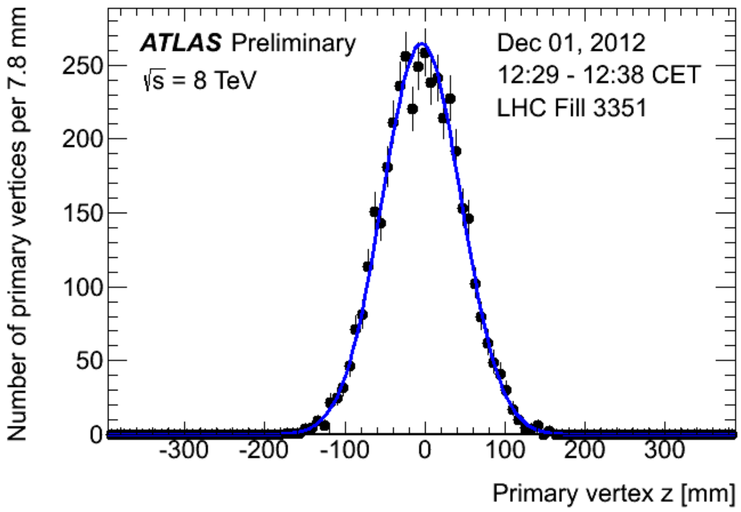


This plot showed us likely a Gaussian distribution, the variable vxp_z refers to the z-position of the primary vertex, the distribution is centered near 0mm on the x-axis, indicating that most interactions occur close to the nominal interaction point at z=0mm. The unit of vxp_z is mm because the width of the distribution is approximately 400, and the detection occurs in the inner detector, which has the radial coverages under 200cm, so the width of the distribution cannot over 200cm and indicates the unit of mm. The y-axis shows the number of events in thousands ($\times 10^3$) for each bin along the z-axis. Statistics Box: Entries (7,917,590) are the total number of events (or vertices) used to construct the histogram. Mean (-9.163) is the average z-coordinate of the vertices is slightly shifted from zero. Standard Deviation (47.81) is the spread of the vertex distribution, indicating that most vertices are within approximately ±48 mm of the mean

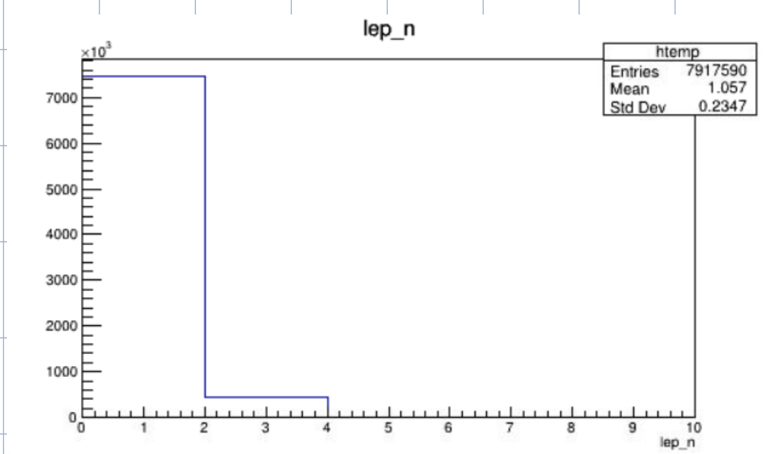
In LHC protons are not continuously injected into the beam but are grouped into "bunches." Each bunch contains a large number of protons, and these bunches are accelerated around the collider, causing collisions at specific interaction points. In the plot above, we observed only a single peak because the proton bunches are not distributed in a perfectly linear fashion but follow a quasi-Gaussian distribution.
To observe multiple peaks, a time-domain plot would be more suitable, as the proton bunches are launched periodically. Consequently, the collisions also occur periodically, allowing multiple peaks to be visible in a time-domain representation.

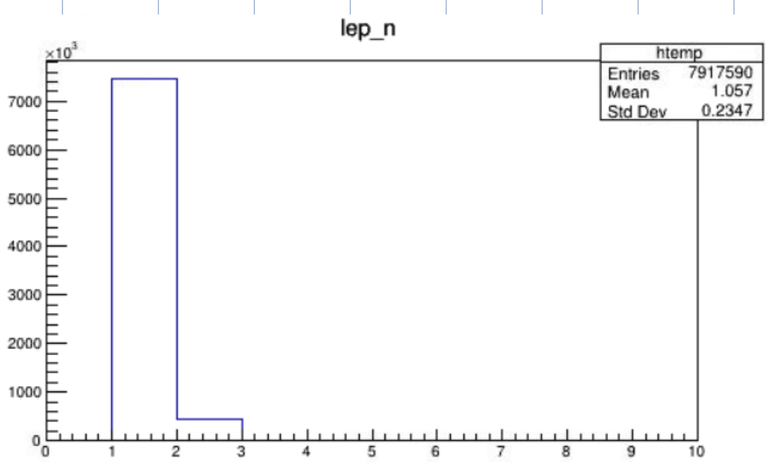






"homogen" (means x, y, z almost same) spatial distribution of Primary vertex, Gauss-like profile opened by Atlas Public





withought ROOT magic, coarse binning

with ROOT magic

Lep_n: Number of preselected Leptons

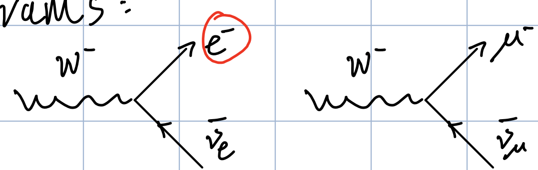
Seen from observation distribution for preselected leptons, 1-Lepton and

2-Lepton final states of Z-decay are same, meanwhile there are also events with more than 2 leptons, however insignificantly.

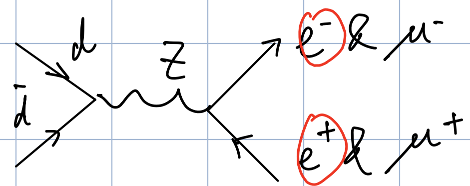
As theoretically predicted, we expect two leptons in the final state of a Z-decay (for instance electron and antielectron). However, events with more than two leptons can occur due to additional radiation or associated production, such as photon emission or Z+W production. Less than two leptons, such as in a 1-lepton final state, can result from incomplete detection or W boson decays, where a neutrino escapes detection

Feynmann-Diagrams:

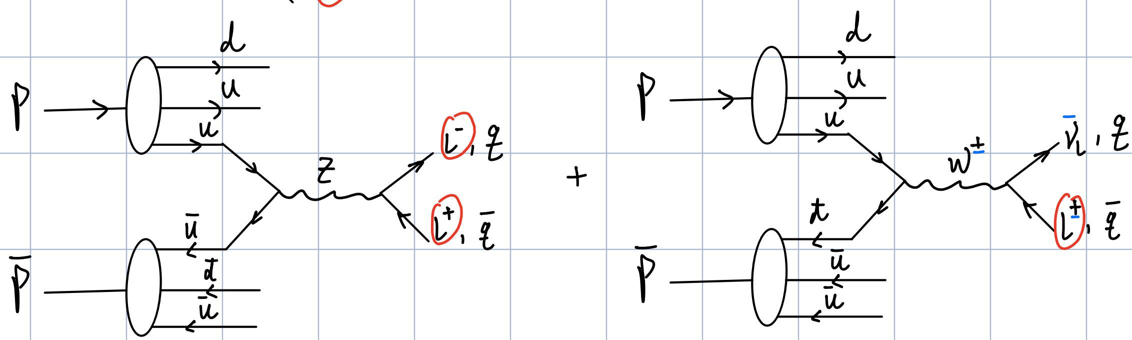
① 1-lepton:



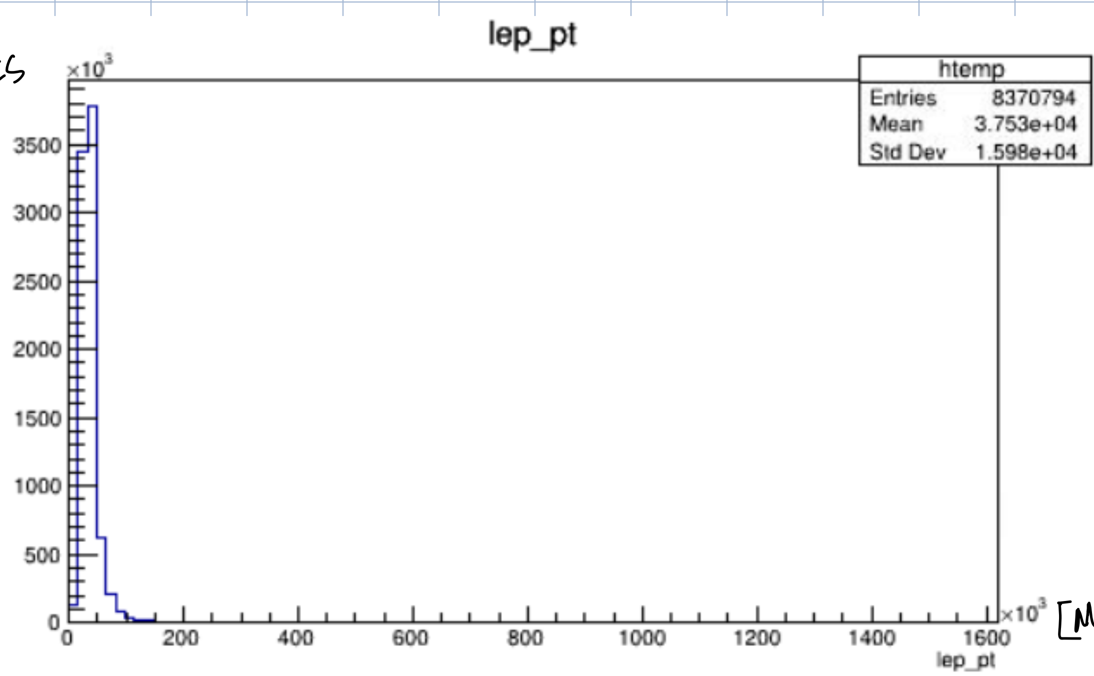
② 2-lepton:



③ 3-lepton:



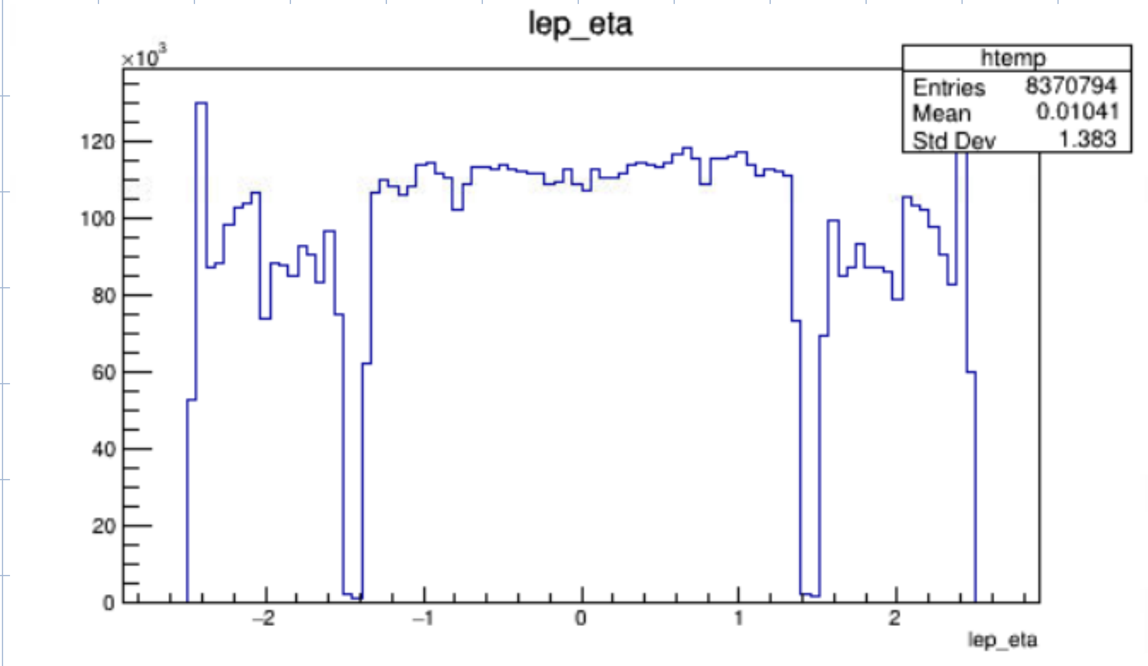
Counts



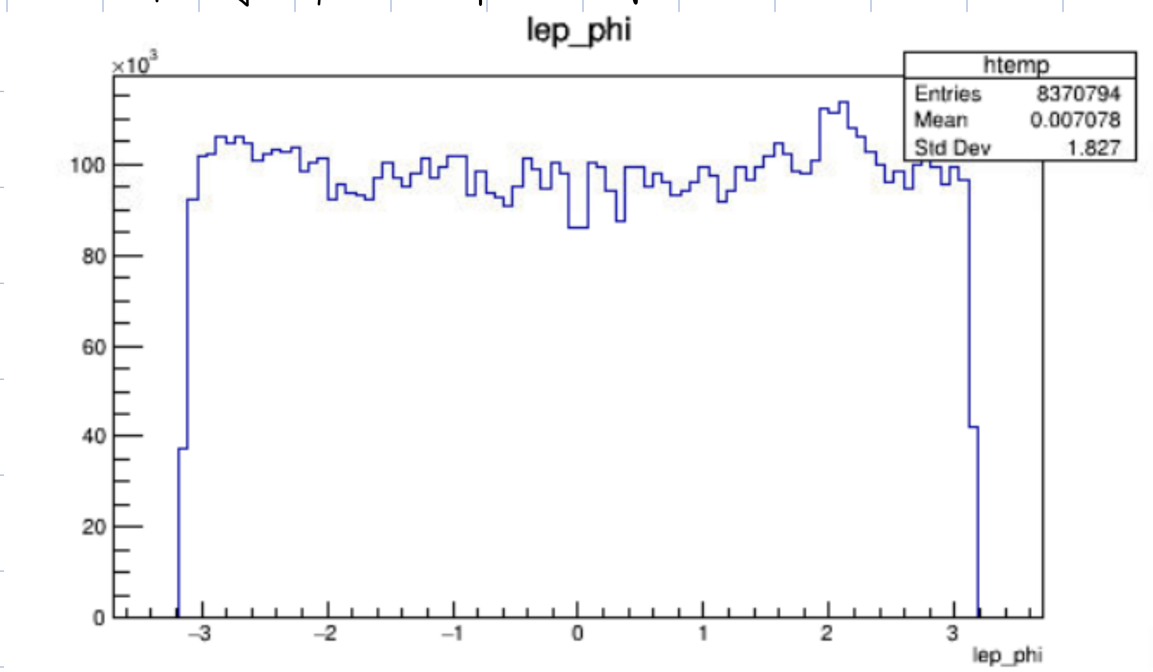
[MeV/c]

lep_pt: Transverse momentum of the lepton

unit: MeV, since the P_T required for Z-Boson production must ≥ 25 GeV/c



lep_eta: Pseudorapidity of the lepton $\eta = -\ln(\tan\frac{\theta}{2})$



lep_phi: Azimuthal angle ϕ of the lepton.

The transverse momentum is distributed within the range of 0 to 200 MeV. Despite the limited number of bins in the plot or an arguably unreasonable zoom range, a Gaussian-like shape can still be observed in the distribution. The mean value is 37.53 GeV, with a standard deviation of 15.98 GeV. The unit of MeV is used because the lepton momenta must exceed 25 GeV.

The distributions of pseudorapidity and azimuthal angle are nearly uniform, exhibiting central symmetry around the origin. However, in the pseudorapidity distribution, two dips are particularly noticeable, and the other parts are approximately constant, as said in script, implying roughly equal number of jets are observed in each interval of pseudorapidity.

There are more entries in this plot compared to the vxp_zv plot because the vxp_z plot only considered entries in one direction. In contrast, this plot involves spherical coordinates, like azimuthal angle and latitude, which represent different orientations within the detector. According to the original data from the LHC accelerator, the expected number of entries is 3564×2808=10,007,712.

An interesting observation is the presence of gaps in the pseudorapidity plot. If we examine the tracker, the active metal (Pb) is distributed as shown in Figure 11 of the lecture. Between values of eta 1.4 and 1.5, there is only dead material, meaning particles hitting this region cannot be detected, causing these gaps to appear.

Additionally, the steep rise around 25 GeV in the P_T spectrum is due to the Gaussian-like distribution. The mass of the Z-boson corresponds precisely to this momentum, which explains this sharp increase

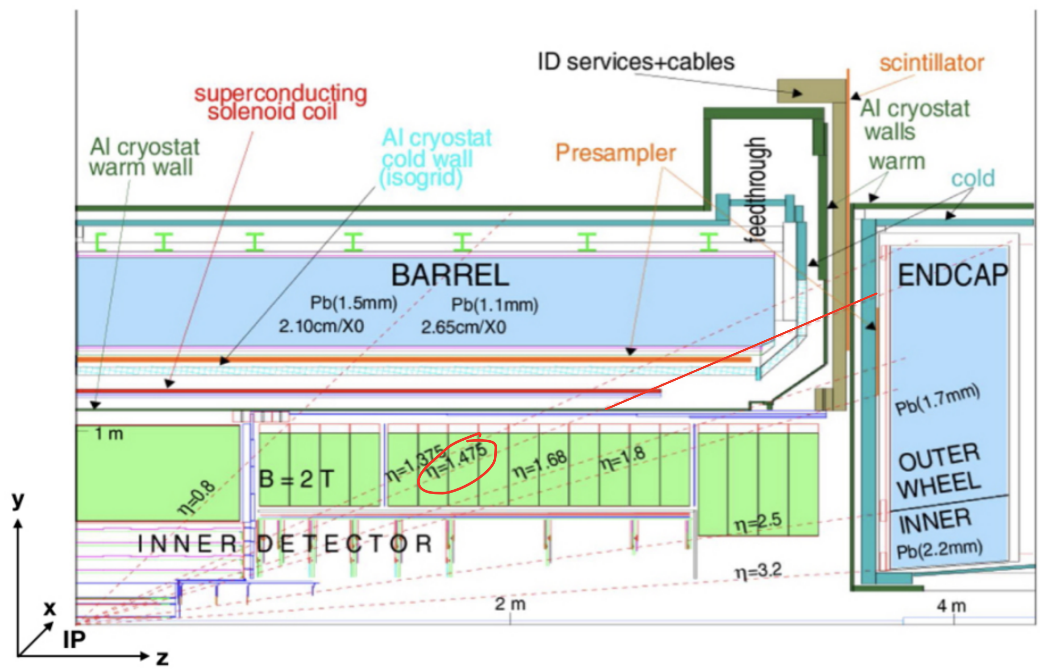


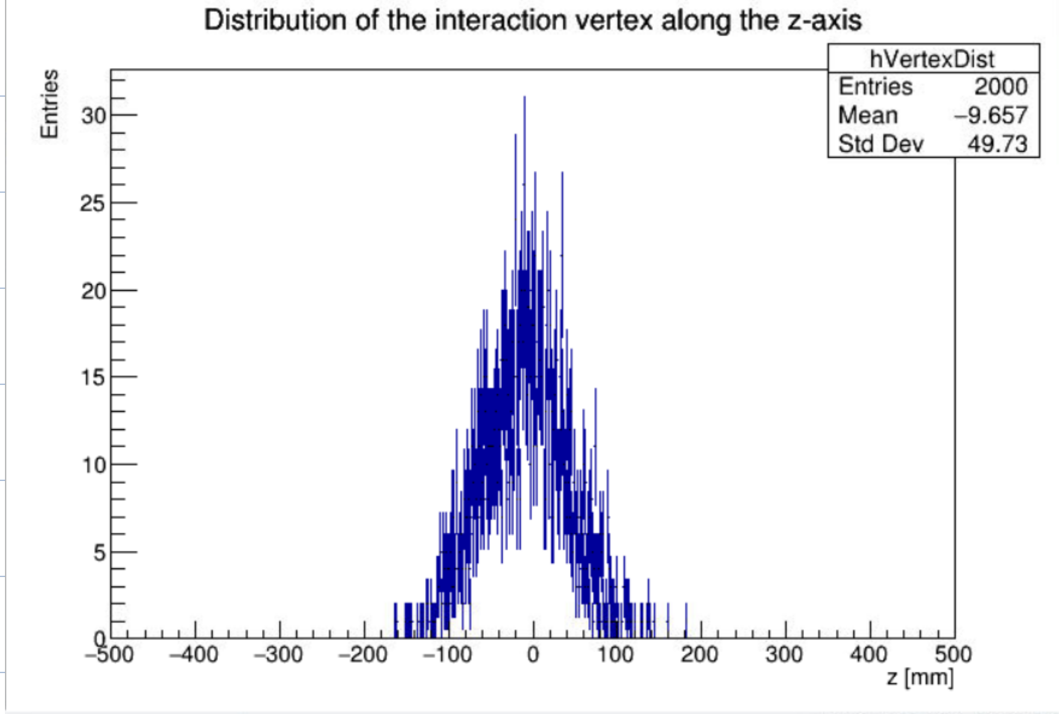
Figure: Longitudal cut-out of the Atlas Liquid Arogon Calorimeter system in y-z plane. Shown is one quarter of the cross section through Calorimeter

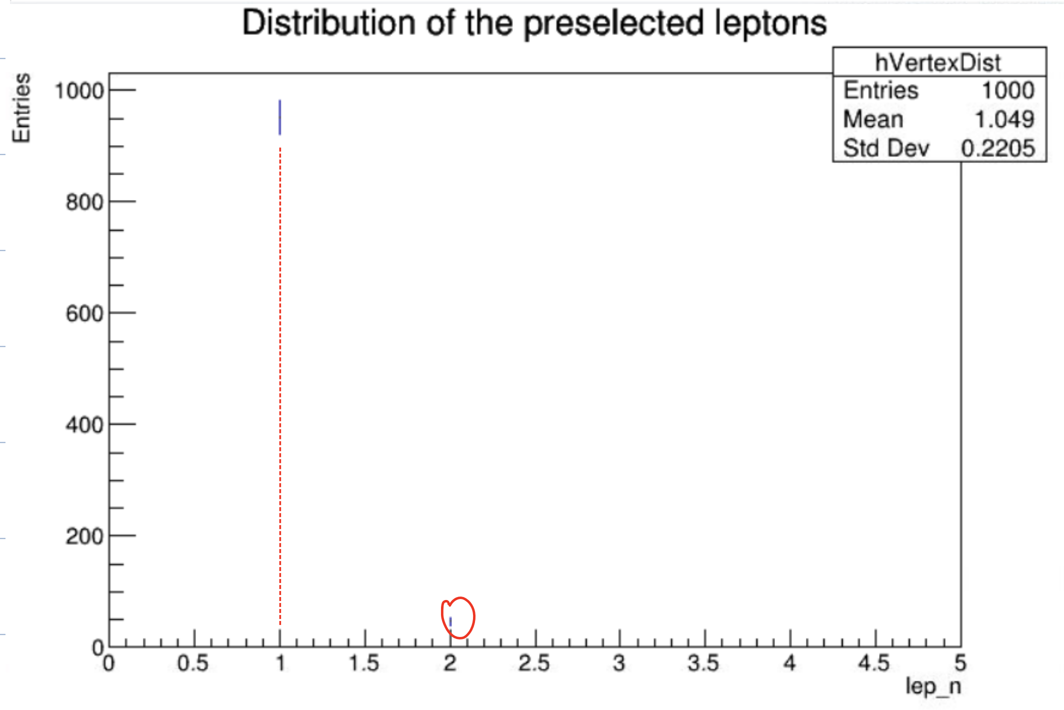
---

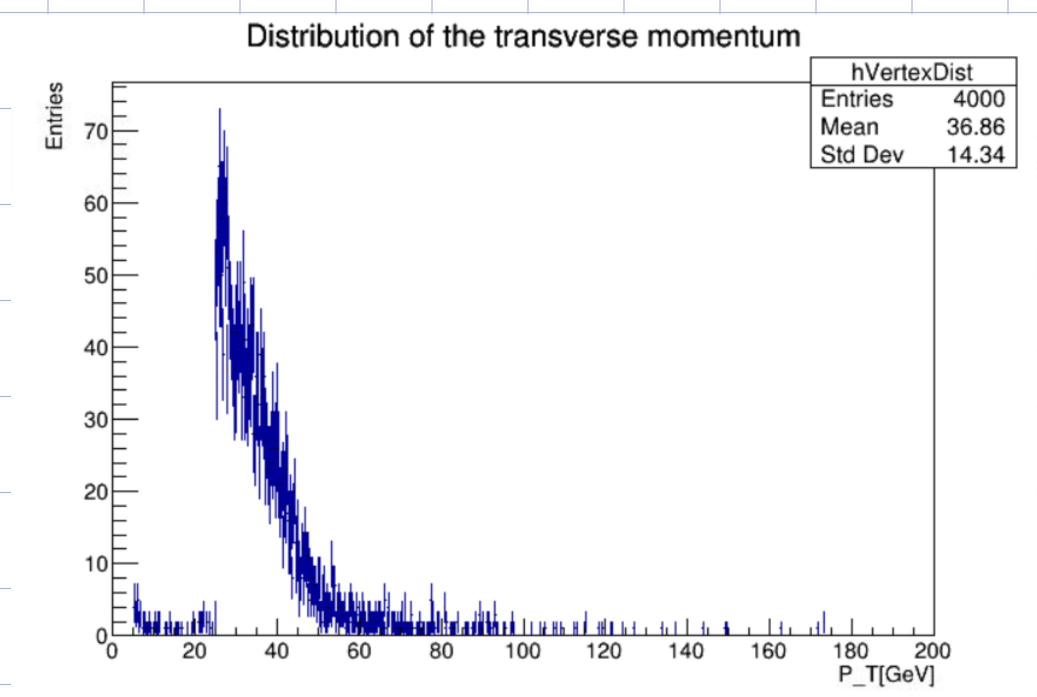
3. Automating Things: using PyROOT to create a loop over all events in the ROOT tree to apply specific operations on certain observables (in form of variables for each event.

python /cipuser/fphys/od300/data/Code/eventloop.py -f /cipuser/fphys/od300/data/Data/

DataEgamma.root -n2000 (Terminal code for running Py. script over Data)





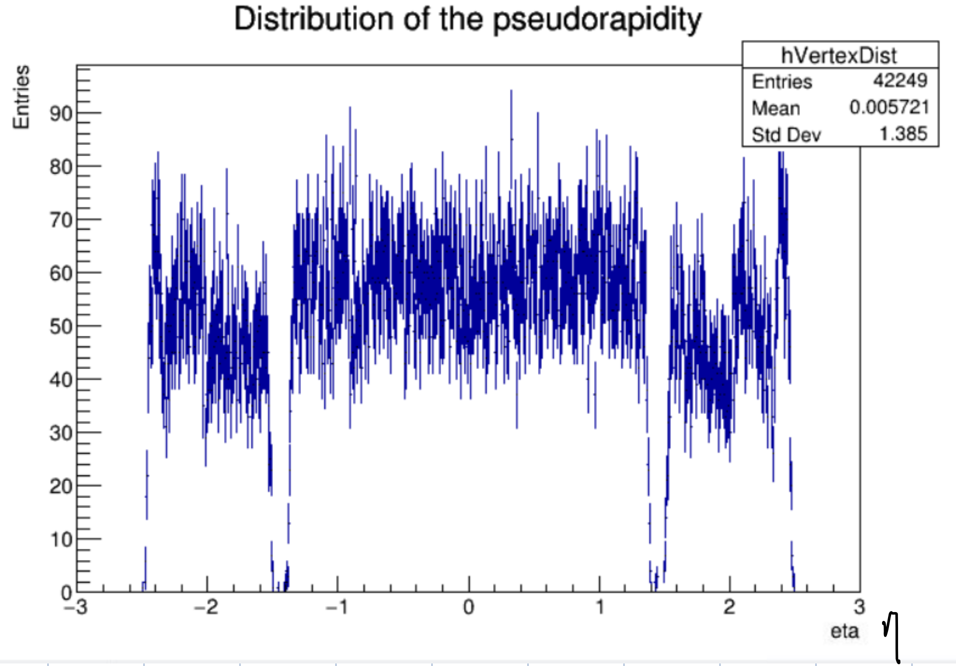


Coding change in script:

for i in range(len(lep_pt)):

lep_pt_gev = lep_pt[i]/1000.0

```
for i in range(len(lep_pt)):  # Loop over the elements in lep_pt
    lep_pt_gev = lep_pt[i] / 1000.0  # Convert each element from MeV to GeV
```



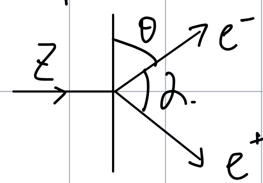
Coding change in the script:

```
d a variable from the input
_eta is likely a collection (vector), so loop over it
ta in myChain.lep_eta:
VertexDist.Fill(eta, weight)
```

Computing the invariant mass of the two leading leptons in the event.

$$s^2 = 4\underbrace{m_e^2}_{=0} + 4|\vec{p}|^2 \sin\frac{\lambda}{2} = 4|\vec{p}|^2 \sin\left(\frac{180^\circ - 2\theta}{2}\right)$$

$$= 4|\vec{p}|^2 \cos\theta$$

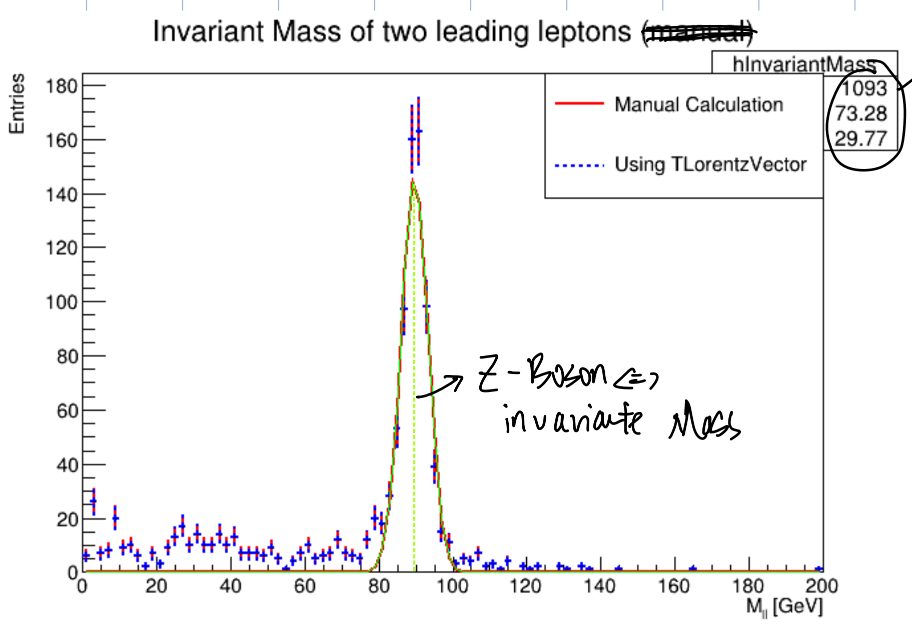


$$\eta = -\ln\left(\tan\frac{\theta}{2}\right) \iff \theta = 2\arctan\left(e^{-\eta}\right)$$

$$|\vec{p}|^2 = p_x^2 + p_y^2 + p_z^2 = p_T^2 + p_T^2 \sinh^2\eta = p_T^2\left(1 + \sinh^2\eta\right)$$

$$p_x = p_T\cos\phi \quad p_y = p_T\sin\phi \quad p_z = p_T\sinh\eta \quad E = p_T\cosh\eta$$

$$\Rightarrow \boxed{s^2 = 4p_T^2\left(1+\sinh^2\eta\right)\cos\theta = E^2 - |\vec{p}|^2}$$



mean and deviation from Data, not the fitted Z-Boson Mass with the Gauss curve

From Gauss-Fitten we got:

$M_{Z^0} = (89.65 \pm 3.63)\ \mathrm{GeV}/c^2$

Error deviation with literature value: $M_Z = (91.1880 \pm 0.0020)\ \mathrm{GeV}/c^2$

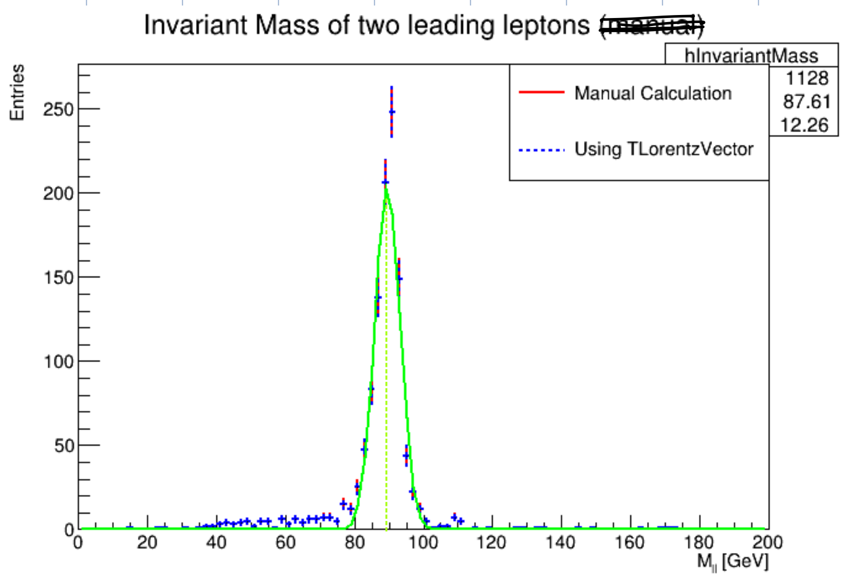
$\Delta \approx 0.43\sigma < 3\sigma$

```
ig304@physik1:~$ python /cipuser/fphys/ig304/Schreibtisch/hahaha/eve
ntloop3.py -f/cipuser/fphys/ig304/Schreibtisch/hahaha/DataEgamma.roo
t -n20000

Info in <TCanvas::MakeDefCanvas>:  created default TCanvas with name
 c1
 FCN=425.737 FROM MIGRAD    STATUS=CONVERGED     283 CALLS         2
84 TOTAL
                     EDM=1.42196e-06    STRATEGY= 1      ERROR MATRI
X ACCURATE
  EXT PARAMETER                                   STEP         FIRST
  NO.   NAME      VALUE            ERROR          SIZE      DERIVATI
VE
   1  Constant     1.46496e+02   8.63753e+00   5.72118e-02  -1.94920
e-04
   2  Mean         8.96539e+01   1.47680e-01   1.41946e-03  -8.66984
e-03
   3  Sigma        3.62870e+00   1.61138e-01   3.20136e-05  -2.62157
e-01
Manual Calculation: Mean = 89.65 GeV, Sigma = 3.63 GeV
 FCN=425.737 FROM MIGRAD    STATUS=CONVERGED     229 CALLS         2
30 TOTAL
```

```
                     EDM=1.05968e-07    STRATEGY= 1  ERROR MATRIX UN
CERTAINTY   2.1 per cent
  EXT PARAMETER                                   STEP         FIRST
  NO.   NAME      VALUE            ERROR          SIZE      DERIVATI
VE
   1  Constant     1.46502e+02   8.67033e+00   6.19223e-02   1.06091
e-04
   2  Mean         8.96540e+01   1.49680e-01  -3.65560e-04  -6.14321
e-04
   3  Sigma        3.62867e+00   1.61832e-01   8.32304e-06   1.44278
e-01
TLorentzVector Calculation: Mean = 89.65 GeV, Sigma = 3.63 GeV
Traceback (most recent call last):
  File "/cipuser/fphys/ig304/Schreibtisch/hahaha/eventloop3.py", lin
e 127, in <module>
    fitFunction.SetLineWidth(0.4)
TypeError: void TAttLine::SetLineWidth(short lwidth) =>
    TypeError: could not convert argument 1 (short int conversion ex
pects an integer object)
ig304@physik1:~$
ig304@physik1:~$ firefox
```

by using Monte-Carlo file mc_14770.Zee.root :



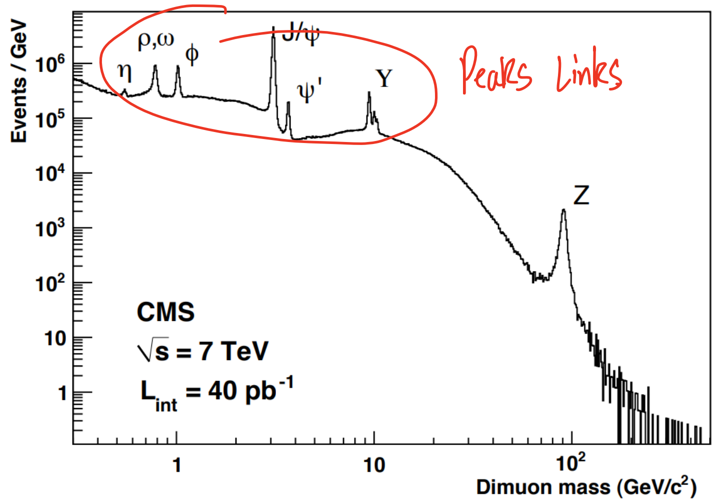

```
  NO.   NAME      VALUE            ERROR          SIZE      DERIVATI
VE
   1  Constant     2.04410e+02   9.92786e+00   4.42795e-02   3.19759
e-06
   2  Mean         8.94764e+01   1.29311e-01   7.96316e-04  -2.04509
e-04
   3  Sigma        3.67573e+00   1.32311e-01   2.72118e-05  -1.01036
e-02
Manual Calculation: Mean = 89.48 GeV, Sigma = 3.68 GeV
 FCN=184.315 FROM MIGRAD    STATUS=CONVERGED     144 CALLS         1
45 TOTAL
                     EDM=3.78494e-09    STRATEGY= 1      ERROR MATRI
X ACCURATE
  EXT PARAMETER                                   STEP         FIRST
  NO.   NAME      VALUE            ERROR          SIZE      DERIVATI
VE
   1  Constant     2.04410e+02   9.92786e+00   4.42796e-02   3.35207
e-06
   2  Mean         8.94764e+01   1.29311e-01   7.96317e-04  -2.09992
e-04
   3  Sigma        3.67573e+00   1.32311e-01   2.72119e-05  -1.04048
e-02
TLorentzVector Calculation: Mean = 89.48 GeV, Sigma = 3.68 GeV
Info in <TCanvas::Print>: file InvariantMassComparison.png has been
created
```

$M_{Z\_MC} = (89.48 \pm 3.68)$ GeV $< M_{Z0}$

Comparision 2-Methods:

The expected distribution of the Z-boson decay should reflect the combination of a natural decay resonance curve and a Gaussian distribution, leading to a single peak. However, if we extend the analysis to include all decay channels instead of just focusing on the 2-lepton final state, the distribution shifts, displaying a left-shifted main peak along with multiple smaller peaks. This shift can be explained by the fact that 1-body decays, which result in lower invariant mass, occur more frequently than 2-body decays. Additionally, other decay processes, like Y-boson decays, can produce systems with lower invariant mass, further contributing to the complexity of the distribution.

Monte Carlo simulations, on the other hand, allow us to observe a smooth distribution without semi-peaks. This is because the data generated by Monte Carlo simulations are based on predefined physics models, excluding external factors such as detector inefficiencies and noise. However, despite the clean distribution, Monte Carlo simulations tend to have lower accuracy compared to real-world data due to the simplifications and idealizations inherent in the models



Peaks Links

figure: CMS invariant-mass spectra of opposite-sign muon pairs using 2010 data

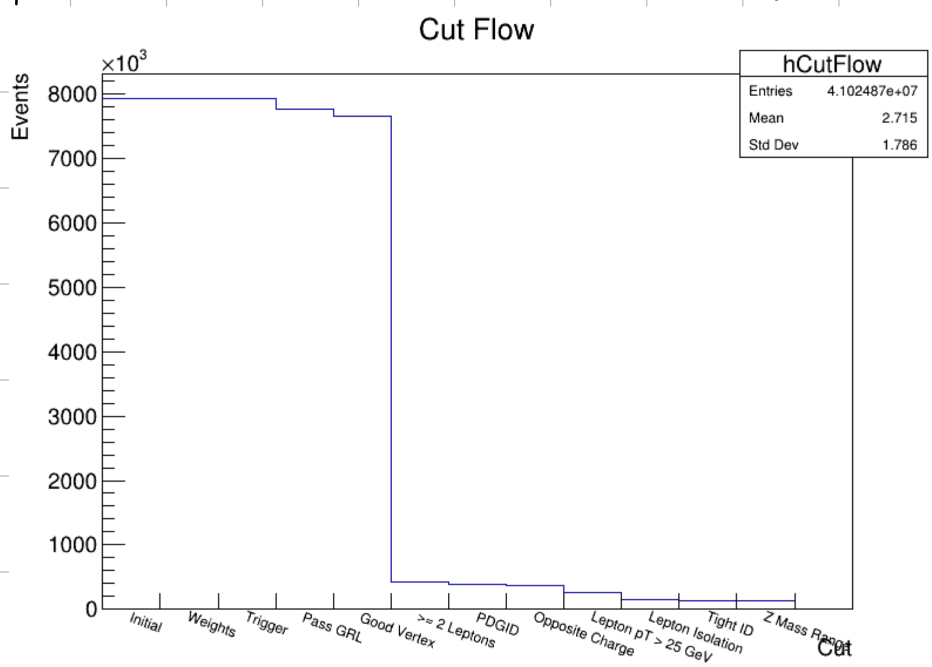
---

4. Selecting Events.

To improve invariant mass distribution and extract a better Z Boson Peak, an event selection is applied using following criteria:

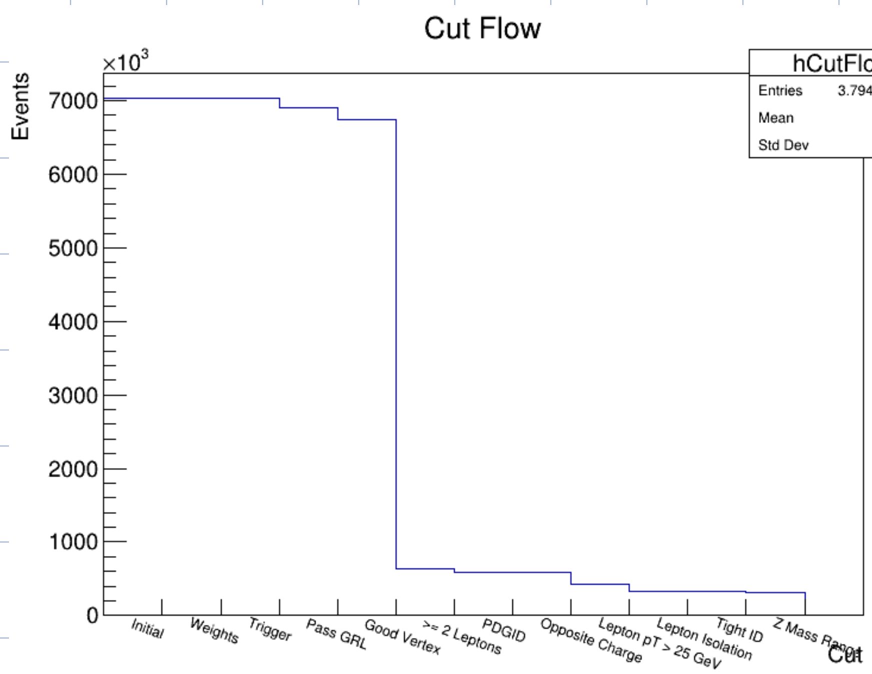
Weights, Trigger, GRL, Vertex, 2 Leptons, PDGID, Charge, P_T Cut, Isolation, Tight ID, Z Mass

After event selection 2 new ROOT files DataEgamma, DataMuons and 3 MC-Files are generated with a cut flow for the selection of Z Boson candidates for the DataEgamma.root file



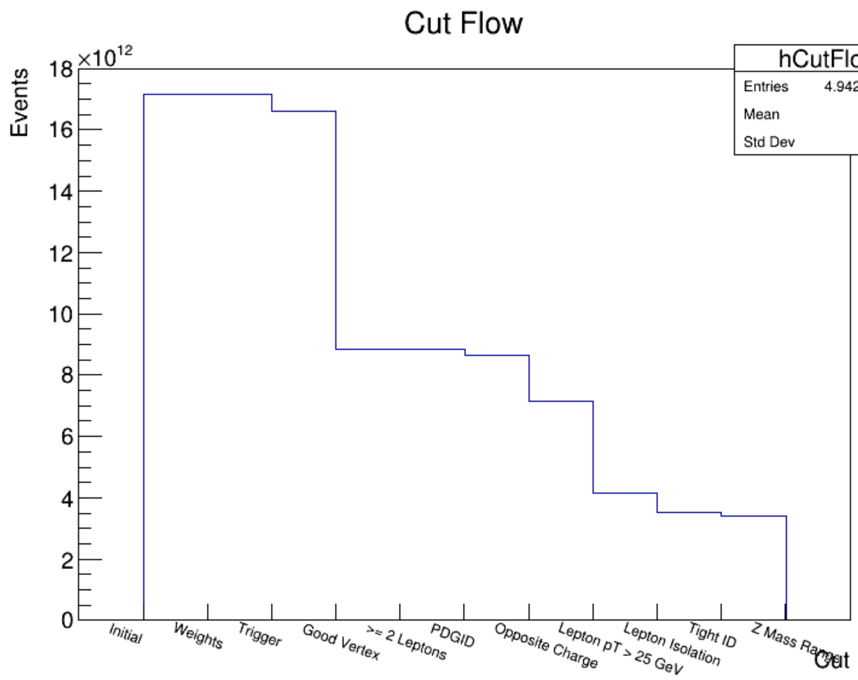
DataEgamma.root

```
Initial: 7917590 events passed
Weights: 7917590 events passed
Trigger: 7917590 events passed
Pass GRL: 7768901 events passed
Good Vertex: 7645544 events passed
>= 2 Leptons: 433583 events passed
PDGID: 386570 events passed
Opposite Charge: 364245 events passed
Lepton pT > 25 GeV: 258989 events passed
Lepton Isolation: 155661 events passed
Tight ID: 131680 events passed
Z Mass Range: 126931 events passed
```



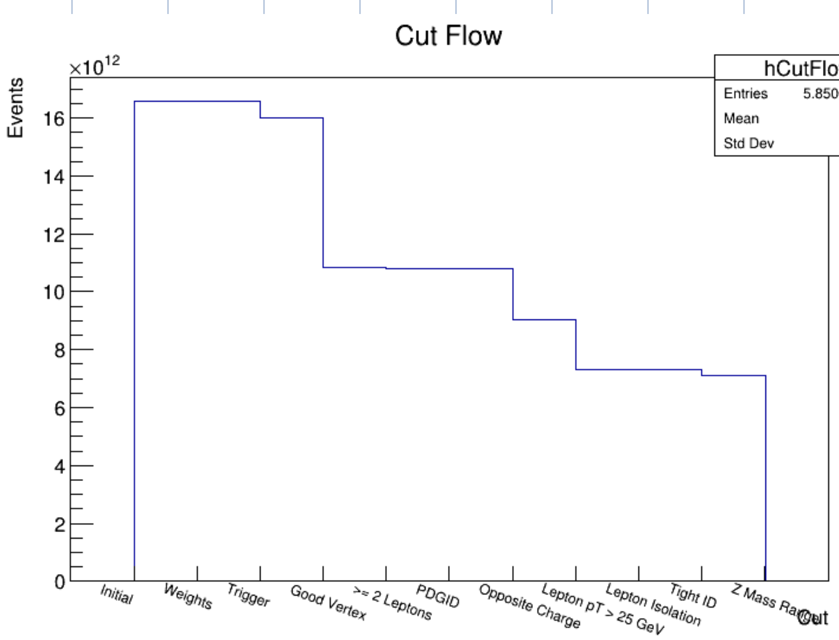
Data Muons.root

```
5Initial: 7028084 events passed
Weights: 7028084 events passed
Trigger: 7028084 events passed
Pass GRL: 6907483 events passed
Good Vertex: 6747211 events passed
>= 2 Leptons: 629846 events passed
PDGID: 593682 events passed
Opposite Charge: 582794 events passed
Lepton pT > 25 GeV: 419377 events passed
Lepton Isolation: 331222 events passed
Tight ID: 331222 events passed
Z Mass Range: 319090 events passed
```



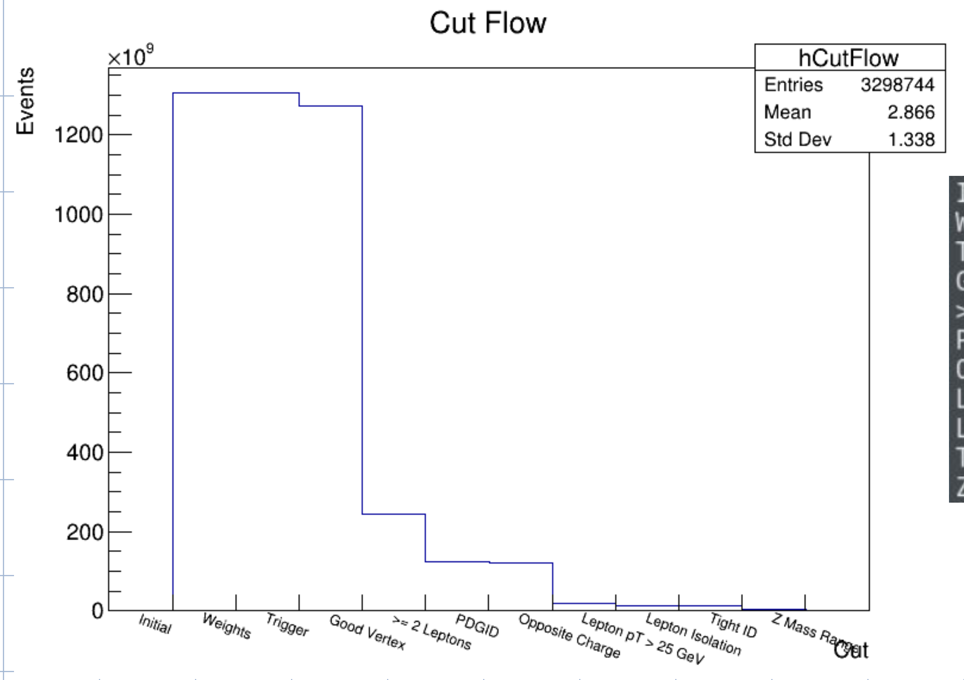
MC_Zee

```
Initial: 7500000 events passed
Weights: 7500000 events passed
Trigger: 7500000 events passed
Good Vertex: 7347094 events passed
>= 2 Leptons: 3999787 events passed
PDGID: 3981464 events passed
Opposite Charge: 3891843 events passed
Lepton pT > 25 GeV: 3060569 events passed
Lepton Isolation: 1741958 events passed
Tight ID: 1473935 events passed
Z Mass Range: 1426230 events passed
```



MC_mumu

```
Initial: 7500000 events passed
Weights: 7500000 events passed
Trigger: 7500000 events passed
Good Vertex: 7323764 events passed
>= 2 Leptons: 5062232 events passed
PDGID: 5039796 events passed
Opposite Charge: 5033762 events passed
Lepton pT > 25 GeV: 4007453 events passed
Lepton Isolation: 3213844 events passed
Tight ID: 3213843 events passed
Z Mass Range: 3111532 events passed
```



MC_tautau

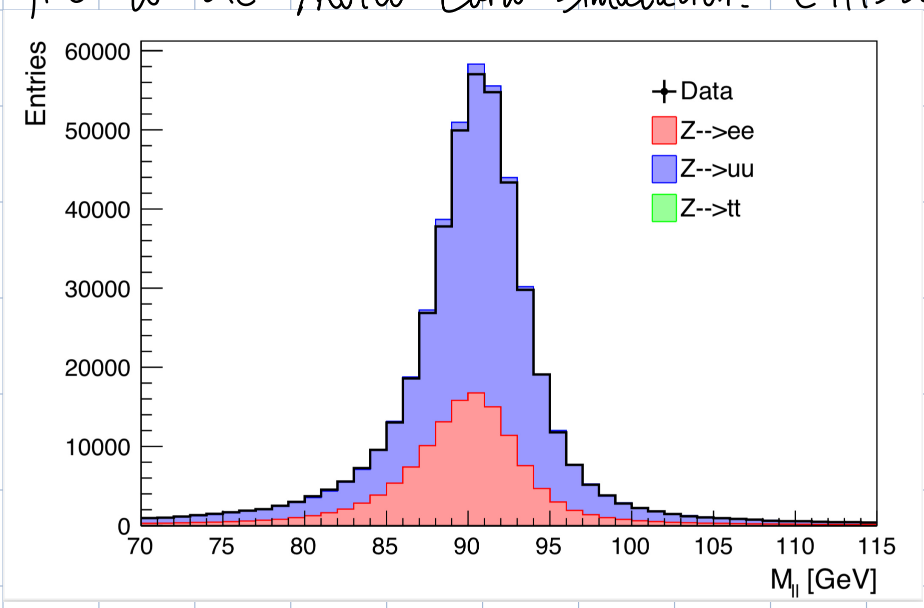
```
Initial: 750000 events passed
Weights: 750000 events passed
Trigger: 750000 events passed
Good Vertex: 738823 events passed
>= 2 Leptons: 142549 events passed
PDGID: 71710 events passed
Opposite Charge: 68826 events passed
Lepton pT > 25 GeV: 10126 events passed
Lepton Isolation: 7192 events passed
Tight ID: 6774 events passed
Z Mass Range: 2744 events passed
```

5. Comparing Data and MC Simulation

To merge ROOT files containing histograms we use hadd:

```
ig304@physik1:~$ hadd /cipuser/fphys/ig304/Schreibtisch/hahaha/resul
t.root /cipuser/fphys/ig304/Schreibtisch/hahaha/selected_DataEgamma.
root /cipuser/fphys/ig304/Schreibtisch/hahaha/selected_DataMuons.roo
t
hadd Target file: /cipuser/fphys/ig304/Schreibtisch/hahaha/result.ro
ot
hadd compression setting for all output: 1
hadd Source file 1: /cipuser/fphys/ig304/Schreibtisch/hahaha/selecte
d_DataEgamma.root
hadd Source file 2: /cipuser/fphys/ig304/Schreibtisch/hahaha/selecte
d_DataMuons.root
hadd Target path: /cipuser/fphys/ig304/Schreibtisch/hahaha/result.ro
ot:/
```

using plot.py script a comparision of final mass distribution of merged data file to the Monto Carlo simulation. (THStack):



$N_l = \mathcal{L}_{int} \cdot \sigma_{Z \to ll}$

$\sigma_{Z \to ll}$: cross section given in script.

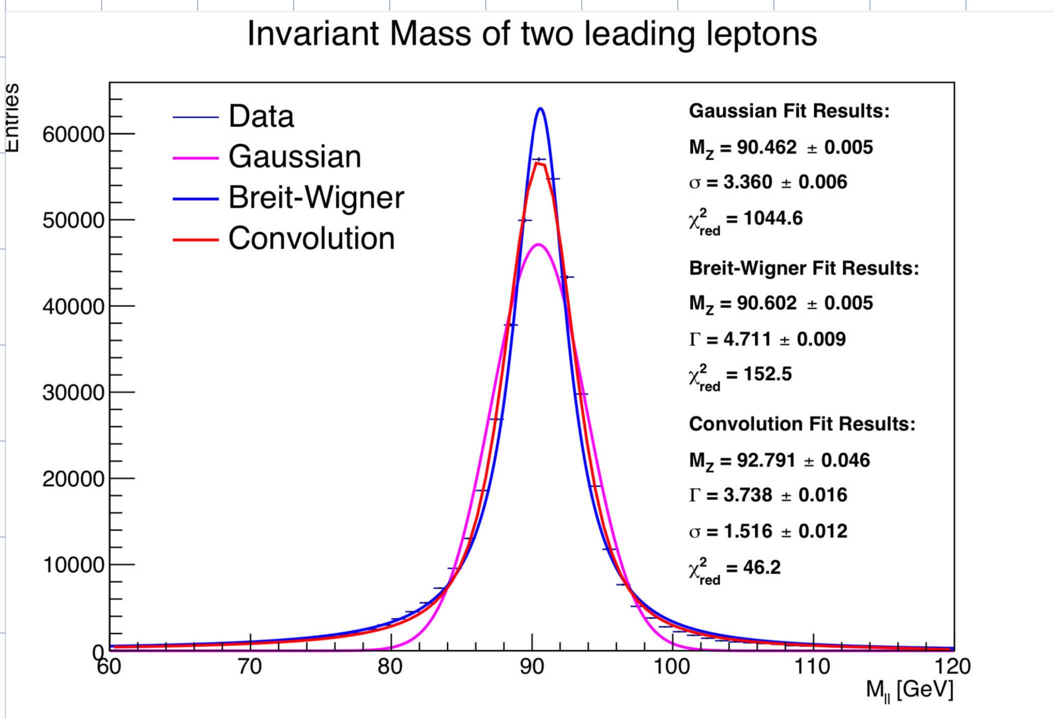
$\mathcal{L}_{int}$: Luminosity

```
ig304@physik1:~$ python /cipuser/fphys/ig304/Schreibtisch/hahaha/plot.py -d/cipuser/fphys/ig304/Schreibtisch/hahaha/result.root -e/cipuser/fphys/ig304/Schreibtisch/hahaha/selected_mc_147770.Zee.root -m/cipuser/fphys/ig304/Schreibtisch/hahaha/selected_mc_147771.Zmumu.root -t/cipuser/fphys/ig304/Schreibtisch/hahaha/selected_mc_147772.Ztautau.root
/cipuser/fphys/ig304/Schreibtisch/hahaha/result.root
/cipuser/fphys/ig304/Schreibtisch/hahaha/selected_mc_147770.Zee.root
/cipuser/fphys/ig304/Schreibtisch/hahaha/selected_mc_147771.Zmumu.root
/cipuser/fphys/ig304/Schreibtisch/hahaha/selected_mc_147772.Ztautau.root
Press any key to continue Info in <TCanvas::Print>: pdf file /cipfile1/fphys/ig304/work_directory/MC.pdf has been created
```

run Py. script

6. Fitting Z Mass

The selected event sample can be used to study properties of Z Boson. using convolution of a Gauss and Breit-Wigner profile for invariance mass distribution fitting, Z-Boson mass M_Z and its decay width can be determined.



When combined (convolved), the resulting shape gives a more realistic representation of the observed invariant mass distribution, capturing both the true natural decay width of the Z boson and the effects of the measurement uncertainties.

Literature value:

$M_Z = (91.1880 \pm 0.0020)\ GeV/c^2$

$\Rightarrow M_{Z\_Gauss} = (90.462 \pm 0.005)\ GeV/c^2$ $\quad M_{Z\_BW} = (90.462 \pm 0.005)\ GeV/c^2$

$M_{Z\_conv} = (92.791 \pm 0.046)\ GeV/c^2$

7. Efficiencies Determination for tight e^- from Z-Final states.

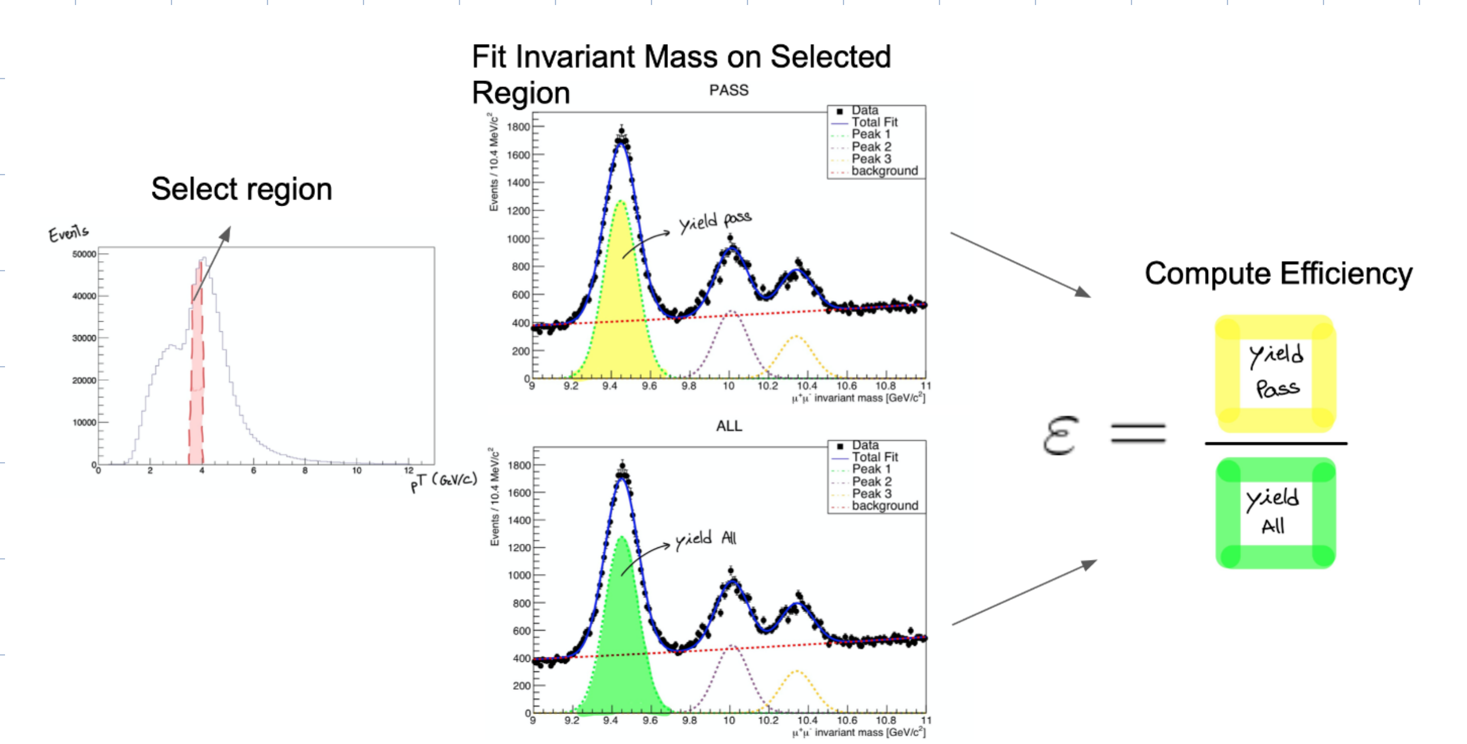
A slightly simplified description of the :

i) resonances are reconstructed as pairs with one leg passing a tight identification (tag) and one passing a loose identification (probe).

ii) "passing probes" are defined according to whatever is the efficiency to measure.

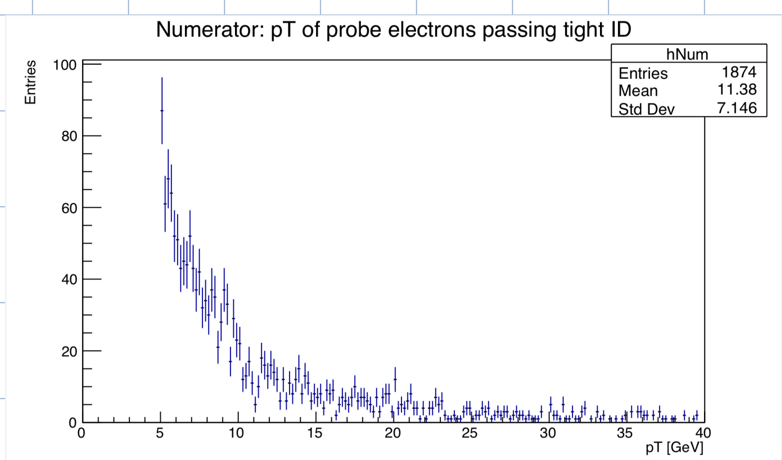
iii) the (tag + passing probe) and (tag + failing probe) lineshapes are fit separately in pT.

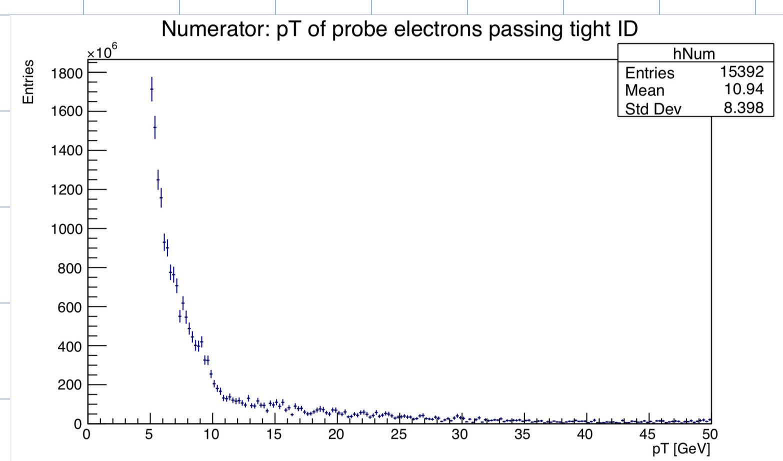
iv) the efficiency is computed from the ratio of the signal yields in the two lineshapes above.

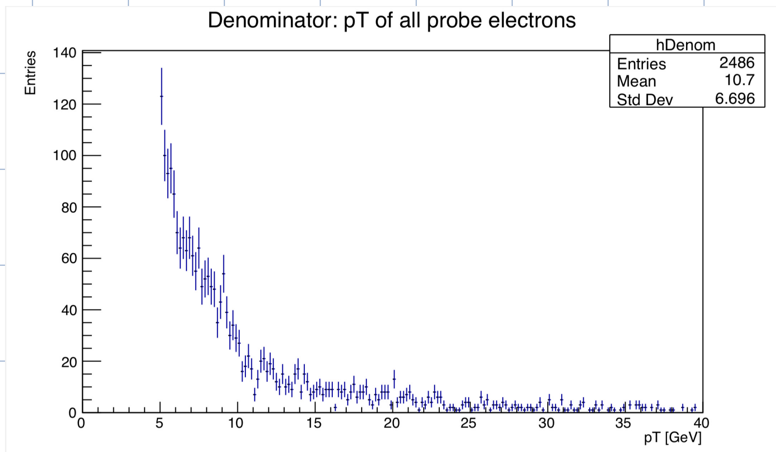


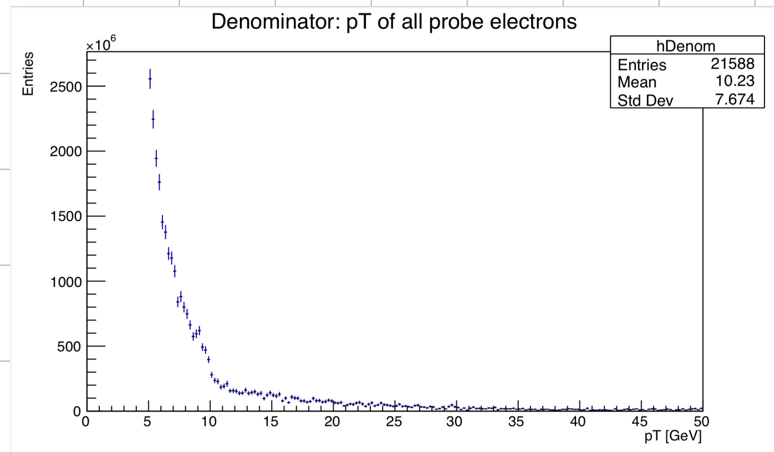
source: CMS-Open-Data Workshop Github

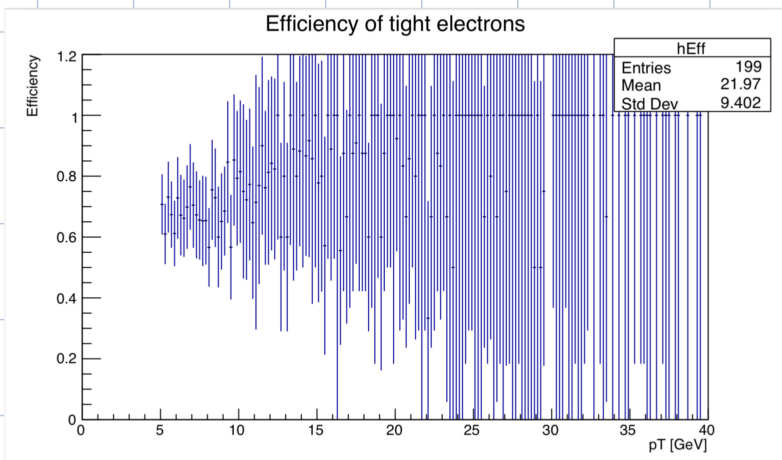
DataEgamma.root Real Data	Monte Carlo Files Z → ee

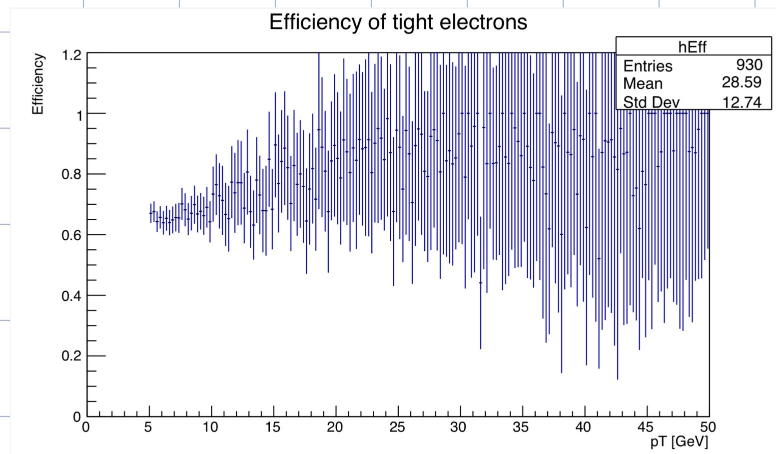












Efficiency values show large fluctuations for lower pT, particularly below 10 GeV. In this range, the uncertainty bars are quite large, indicating high uncertainty in the measurements. As pT increases (around 10–40 GeV), the efficiency stabilizes, though the error bars remain significant, showing some variation in the efficiency values. The efficiency is relatively high (> 0.7) in the mid pT range but appears to decrease and fluctuate more as pT approaches 40 GeV. In the low pTp_TpT range (around 5-10 GeV), efficiency is lower and more erratic, potentially due to limited statistics or challenges in reconstructing tight electrons at low momenta.

---

8. Error Analyzing.

In the determination of the Z boson mass and related measurements, both statistical and systematic errors play roles. Statistical errors arise primarily from the limited event statistics and detector resolution, such as uncertainties in measuring the energy and momentum of particles, which propagate into the final mass calculation. Systematic errors, however, are more complex and stem from uncertainties in the calibration of detectors, the energy scale, and the modeling of the background processes or efficiencies (e.g., electron identification and reconstruction). Systematic effects can also come from theoretical assumptions, such as the choice of the signal and background models used in the fit, because it is also possible that for instance the Z-decay process can not only be described by the natural decay resonance curve.