# Comparison of DIS and Drell-Yan electron distributions from ZEUS and CMS using a common NanoAOD-like format

Ihor Melnyk, Taras Shevchenko National University of Kyiv, Ukraine

Supervisors: Achim Geiser, Yewon Yang

September 7, 2022

**Abstract**

The aim of this project is to be able to combine physics analysis with the ZEUS and CMS data. In order to do that it was necessary to convert the data coming from the process of deep inelastic scattering in the ZEUS experiment to the NanoAOD format common in the CMS. Also, in this work, a comparison of initial CMS variables for Drell-Yan electron and converted ZEUS variables for DIS electrons was made.

# Contents


**1 Introduction** 3

**2 Theory** 3

**3 Conversion** 5

3.1 Converted variables . . . 5

3.1.1 Electron_cutBased . . . 5

3.1.2 Point of the closest approach to the beamline . . . 5

3.1.3 Electron_phi . . . 6

3.1.4 Electron_SCeta . . . 6

3.1.5 Electron_eta . . . 6

3.2 Additional variables . . . 7

3.2.1 Emxda_ZEUS . . . 7

3.2.2 Emxel_ZEUS . . . 7

3.2.3 Emxda_cell_ZEUS . . . 7

3.2.4 Emyda_ZEUS . . . 7

3.2.5 Emyel_ZEUS . . . 7

3.2.6 Emyda_cell_ZEUS . . . 7

3.2.7 Emq2da_ZEUS . . . 7

3.2.8 Emq2el_ZEUS . . . 7

3.2.9 Emq2da_cell_ZEUS . . . 7

3.2.10 Electron_energy_ZEUS . . . 8

3.2.11 Electron_theta_ZEUS . . . 8

**4 Electron selection** 8

4.1 ZEUS: DIS electrons . . . 8

4.2 CMS: Drell-Yan electrons . . . 9

**5 Comparison** 9

5.1 ZEUS only variables . . . 10

5.2 Common variables . . . 13

5.3 $Z^0$ invariant mass distribution . . . 22

5.4 CMS only variables . . . 22

**6 Appendix** 25

6.1 Tables of variables . . . 25

6.2 Further investigation . . . 26

# 1 Introduction

This comparison's actual goal is to analyze data gathered in both ZEUS and CMS experiments. The ZEUS experiment was running from 1992 to 2007 years at the HERA accelerator at DESY. HERA is an electron-proton collider that collides particles on center-of-mass $\sqrt{s} = \sqrt{4E_pE_e} \approx 318$ GeV. While the Compact Muon Solenoid (CMS) experiment is one of the fourth experiments that has been running on the Large Hadron Collider (LHC) since November 2009. CMS works on center-of-mass energy equals 0.9–13 TeV and the daughter particles are produced there as a result of the proton-proton collision. Both experiments study a wide range of physical processes, that are different compared to each other, and run on different machines with diverse properties and reconstruction algorithms. Thereby, these experiments produced different types of data, with divergent variables' names and properties. There is something in the common, despite distinct collision processes, the secondary particles might take part in the same or similar further interactions. In order to analyze data taken from both colliders, it is crucial to reduce data to one format. This is the ***NanoAOD*** format which substituted the previous ***Analysis Object Data (AOD)*** format from CMS Run 1. It has to be noticed that each experiment has its own set of variables, according to the detector properties and conventions regarding variables' names. Not every CMS variable can be found in ZEUS n-tuples and the inverted statement is also true. Therefore, there are some ZEUS variables that are supposed to be just renamed according to the CMS convention, however, plenty of CMS variables are needed to be calculated using existing ZEUS variables. For conversion were used data_07p ZEUS sample and Run2011A CMS sample. In order to validate conversion is needed to make a comparison of the converted ZEUS variables and initial CMS variables. Also, it is useful to perform an analysis of the $Z \to e^+e^-$ Drell-Yan process and reconstruct the $Z$-boson invariant mass in CMS.

# 2 Theory

As it is mentioned in the Introduction, the ZEUS and CMS collision processes are completely different. Therefore, the produced particles, in particular, electrons have very different properties. For instance, electron-proton scattering is occurring either via the intermediate virtual $\gamma$, or $Z^0$- or, $W^{\pm}$-bosons. This process is referred to as ***deep-inelastic scattering*** (DIS). While, in the case of $pp$-scattering in CMS experiment, a quark of one hadron and an antiquark of another hadron annihilate, creating a virtual photon or $Z$-boson which then decays into a pair of oppositely-charged leptons. The energy of the quark-antiquark pair annihilation can be almost entirely transformed into the momentum of new particles. This process is called ***Drell-Yan process***. The DIS and Drall-Yan interactions have similarities and, moreover, they are closely related to each other. The Feynman diagram of the Drell–Yan process looks like the diagram of DIS rotated by 90°. In Figure 1 and Figure 2, it can be seen both DIS and Drell-Yan diagrams, respectively. Combining the electrons that come from $Z$ in the Drell-Yan process in CMS the $Z$ invariant mass can be reconstructed.

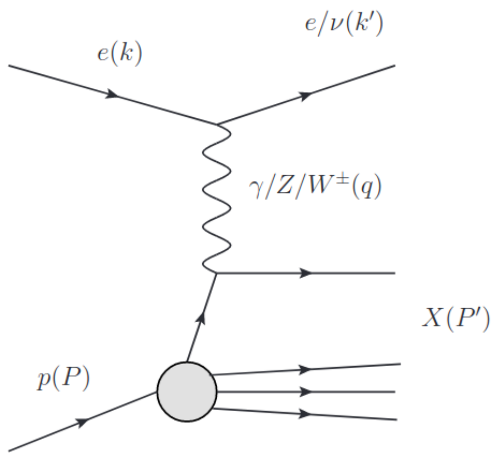



Figure 1: The Feynman diagram for DIS at HERA.
A quark from the incoming hadron interacts with the virtual photon

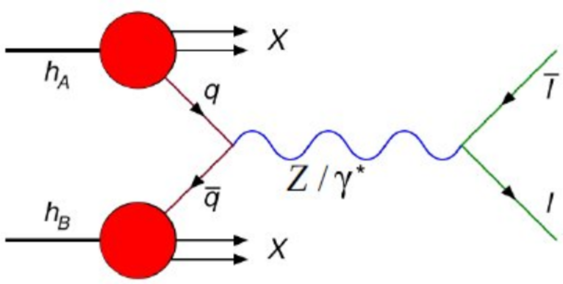



Figure 2: The Feynman diagram for Drell-Yan process.
After hadron collision both quarks annihilate and produce the virtual boson or photon

# 3 Conversion

Before conversion, it was needed to match ZEUS variables with appropriate CMS variables if they exist. ZEUS variables are already stored in form of n-tuples. There are two types of ZEUS electron variables: EM and SIRA. I had chosen the EM algorithm. Not all the ZEUS variables have analogs in CMS data, and not every CMS variable can be found in ZEUS n-tuples. Therefore, I have added some new custom variables in order to substitute missed ZEUS variables that do not have the same or at least similar data in CMS. In the meantime, not every CMS variable has been matched with a ZEUS variable, because of a disability to connect physics or some difficulties in the conversion. The following subsections describe the variables that were calculated. All the other variables can be found in the Tables of variables in the Appendix.

## 3.1 Converted variables

There are two types of electrons in ZEUS depending on the part of the detector where the certain electron was detected. These are the electrons that have an association with the tracking part of the ZEUS detector, and the electrons that were registered only in the calorimeter. The Emtrknr variable taken from the ZEUS n-tuples indicates whether the electron has this association or not, and matches the electron's track with the track in the total tracking block. If these variable equals zero the electron was not detected by the tracking system. The Trk_ntracks ZEUS variable points out the total number of tracks in the tracking block. Also, every track from the tracking block has its own track ID described by Trk_id variables.

### 3.1.1 Electron_cutBased

The integer value of Electron_cutBased depends on a quality of an electron. The "Electron Grand Probability" called Emprob is responsible for the electron quality in the ZEUS n-tuples. In case if Emprob $< 0.001$, Electron_cutBased $= 0$, otherwise Electron_cutBased equals either 4, or 2, depending on whether the electron belongs to the track or not, respectively.

### 3.1.2 Point of the closest approach to the beamline

The point of the closest approach to the beamline in the CMS has three components: Electron_x, Electron_y, and Electron_z.
In the ZEUS tuples, we have a 3-dimensional variable for the position of the point of the closest approach called Trk_pca. In order to convert this variable to the CMS format, it was enough to select only those electrons that have an association with the tracking system:

- Electron_x := Trk_pca[0],
- Electron_y := Trk_pca[1],

- Electron_z := Trk_pca[2],

where the tracking ID Trk_id and electron's track number Emtrknr are the same.

### 3.1.3 Electron_phi

The Electron_phi variable is the $\phi$ angle of an electron. There are two variables in the ZEUS n-tuples that describe the azimuthal angle. The Emph variable is the $\phi$ angle measured from the calorimeter, and Emtrkph is taken from the beginning of the track. Therefore, if a certain electron has an association with the track we chose Emtrkph, and if not – Emph. It is also necessary to take into account that this variable must be in the range from $-\pi$ pi to $\pi$.

### 3.1.4 Electron_SCeta

The Electron_SCeta is an electron's pseudorapidity taken from the Super Cluster (calorimeter part of the detector). In order to calculate in we have to follow the next steps:

- Calculate every component of Electron_SCeta:

  Electron_SCeta_x = Emcalpos[0] − Xvtx := $\eta_x$,

  Electron_SCeta_y = Emcalpos[1] − Yvtx := $\eta_y$,

  Electron_SCeta_z = Emcalpos[2] − Zvtx := $\eta_z$,
  where Emcalpos is an electron position in the calorimeter, and (Xvtx, Yvtx, Zvtx) is the vertex.

- Calculate the absolute value of the radius-vector of the position in the horizontal plane:

  $R = \sqrt{\eta_x^2 + \eta_y^2}$

- Calculate the polar angle $\theta = \arctan(R/\eta_z)$

- Finally, calculate pseudorapidity:

  $\boxed{\eta = -\ln(\tan\frac{\theta}{2})}$

### 3.1.5 Electron_eta

The Electron_eta variable is a pseudorapidity of the electron. In the ZEUS data, we have the Emtrkth variable that describes electron $\theta$ taken from the track block. It means if a certain electron has an association with the track, we use the well-known formula Electron_eta $= -\ln(\tan(\text{Emtrkth}/2))$, otherwise, we fill this variable as the appropriate Electron_SCeta value.

## 3.2 Additional variables

### 3.2.1 Emxda_ZEUS

Emxda_ZEUS was taken from ZEUS n-tuples and just renamed. This variable is x Bjorken calculated with the double-angle method based on zufos.

### 3.2.2 Emxel_ZEUS

The same as Emxda_ZEUS, but calculated with the electron method.

### 3.2.3 Emxda_cell_ZEUS

$x$ Bjorken calculated with the double-angle method based on cells.

### 3.2.4 Emyda_ZEUS

Emyda_ZEUS is also renamed ZEUS variable. It means inelasticity y calculated with the double-angle method based on zufos.

### 3.2.5 Emyel_ZEUS

Inelasticity $y$ calculated with the electron method.

### 3.2.6 Emyda_cell_ZEUS

Inelasticity $y$ calculated with the double-angle method based on cells.

### 3.2.7 Emq2da_ZEUS

The virtuality Q2 calculated with the double-angle method based on zufos has been named Emq2da_ZEUS. This variable means the squared 4-momentum transferred from the electron to the proton:

$$Q^2 = -(k - k')^2$$

### 3.2.8 Emq2el_ZEUS

Virtuality $Q^2$ calculated with the electron method.

### 3.2.9 Emq2da_cell_ZEUS

Virtuality $Q^2$ calculated with the double-angle method based on cells.

### 3.2.10 Electron_energy_ZEUS

Electron_energy_ZEUS variable represents the energy $E_{e'}$ of produced after the collision electron, taken from ZEUS data. This variable was calculated using the following formula:

$$E_{e'} = p_T \cdot \cosh(\eta),$$

where $p_T$ is named as Electron_pt and $\eta$ as Electron_eta.

### 3.2.11 Electron_theta_ZEUS

Electron_theta_ZEUS describes the $\theta$ angle of an electron. This variable was also calculated, according to the following equality:

$$\theta = 2 \cdot \arctan(\exp(-\eta)),$$

where pseudorapidity $\eta$ is already mentioned above.

# 4 Electron selection

In order to compare the electrons from different sources and obtained in diverse processes in different accelerators, it was needed to select only "good electrons". It means applying some kinematic constraints on both ZEUS and CMS variables. The distributions in each case can be different, have different origins, and have different physics behind them, so the restrictions will be different.

## 4.1 ZEUS: DIS electrons

In the case of ZEUS n-tuples, all the electrons have been reconstructed by the EM reconstruction algorithm, but not all of them satisfy "good" electron conditions. In order to select well-reconstructed electrons in ZEUS, the following constraints were applied:

- Electron_cutBased = 4,
- Selected only the first candidate which has the greatest Emprob value,
- 125 < Emq2da_ZEUS < 20000 GeV$^2$,
- 0.2 < Emyda_ZEUS < 0.6,
- -30 < PV_z < 30 cm,
- Electron_energy > 10 GeV,

## 4.2 CMS: Drell-Yan electrons

If we work with CMS electrons that come from Drell-Yan processes, particularly $Z^0 \to e^+e^-$ decay, we are interested in reproducing $Z$-boson invariant mass distribution. $Z$-boson daughter products should also meet requirements for well-reconstructed electrons. Therefore, the following restrictions were applied to CMS electron variables:

- Electron_pt $> 7$ GeV,
- (Electron_deltaEtaSC + Electron_eta) $< 2.5$,
- Electron_lostHits $\leq 1$ and $|$Electron_sip3d$| < 4$,
- $|$Electron_dxy$| < 0.5$ and $|$Electron_dz$| < 1$ cm,
- $80 <$ M($Z \to ee$) $< 100$ GeV,

where *Electron_pt* is electron's transverse momentum,
*Electron_deltaEtaSC* – Super Cluster pseudorapidity,
*Electron_eta* – electron pseudorapidity,
*Electron_lostHits* – NanoAODplus extension,
*Electron_sip3d* – 3D impact parameter significance wrt first PV,
*Electron_dxy* – $dxy$ (with sign) wrt first PV,
*Electron_dz* – $d_z$ (with sign) wrt the first PV.

# 5 Comparison

In this section, a reader can find the output histograms only for some converted variables. The histogram on the top shows the initial CMS variable and below it is seen as the appropriate converted ZEUS variable.

## 5.1 ZEUS only variables

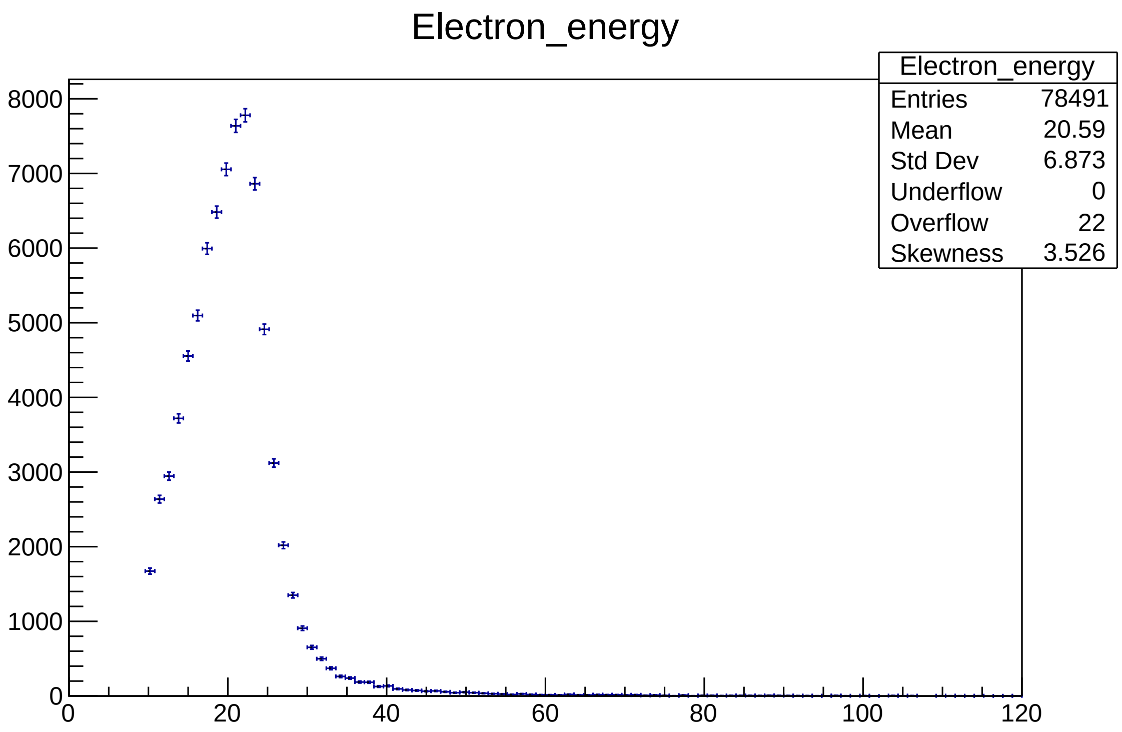



Figure 3: Calculated $E_{e'}$ in GeV

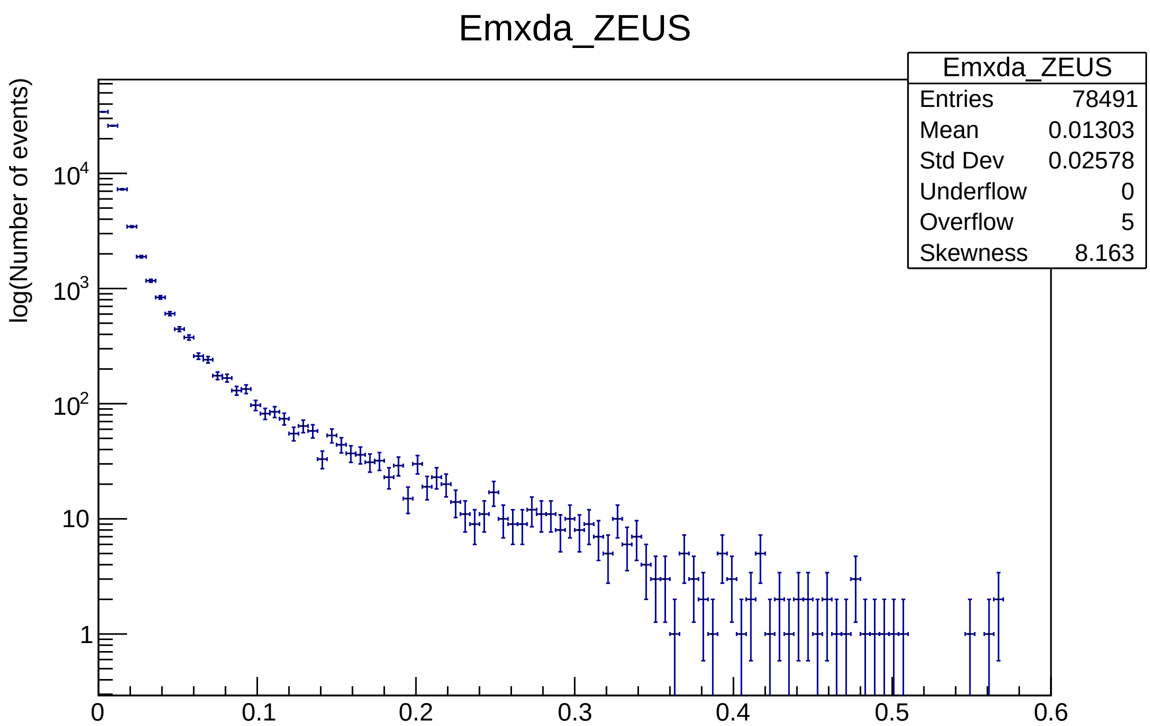



Figure 4: $x$ Bjorken calculated with double-angle method

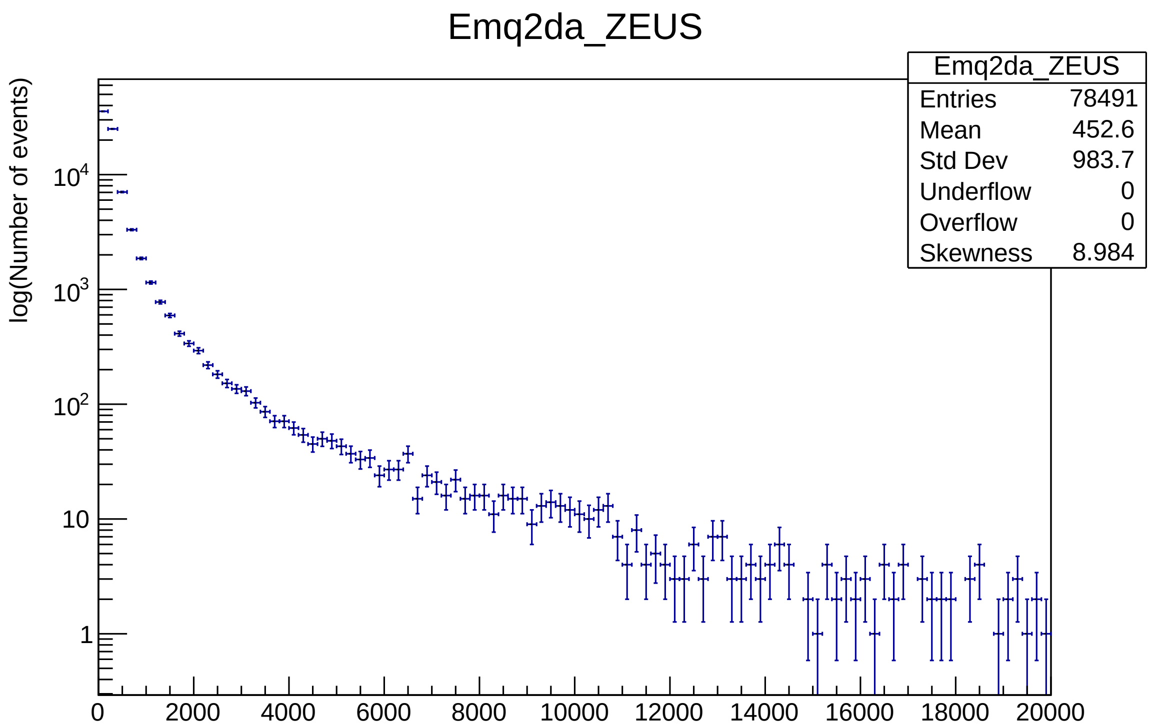



Figure 5: $Q^2$ in GeV$^2$

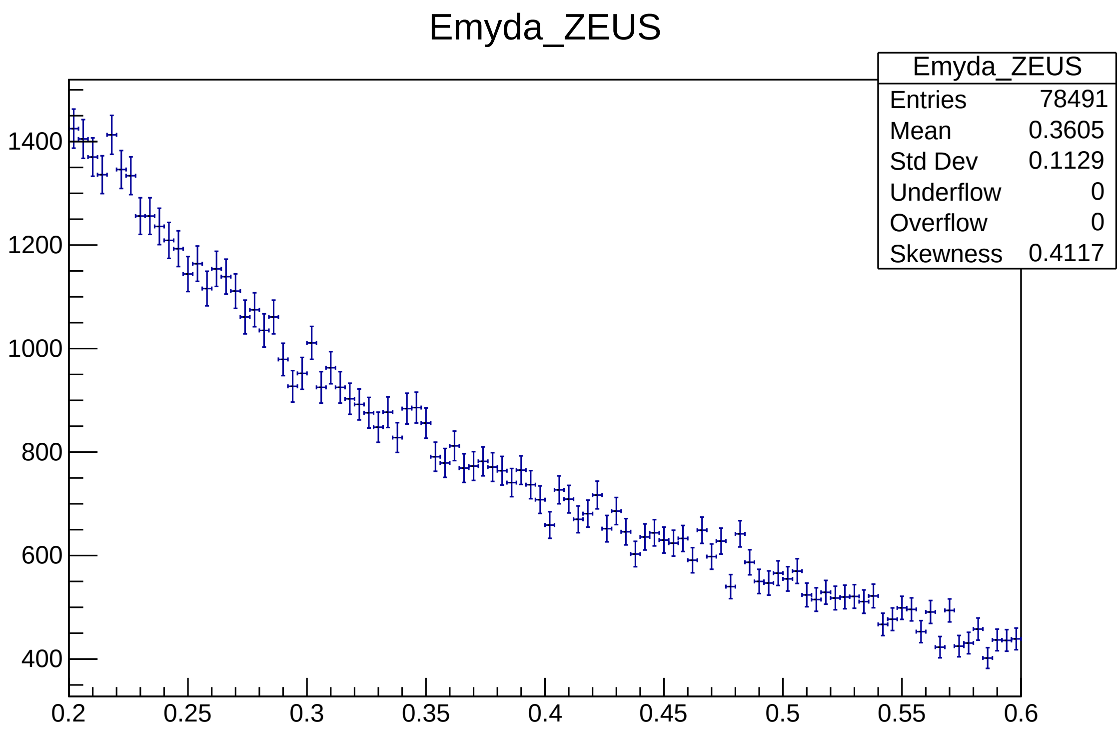



Figure 6: inelasticity $y$ calculated with double-angle method

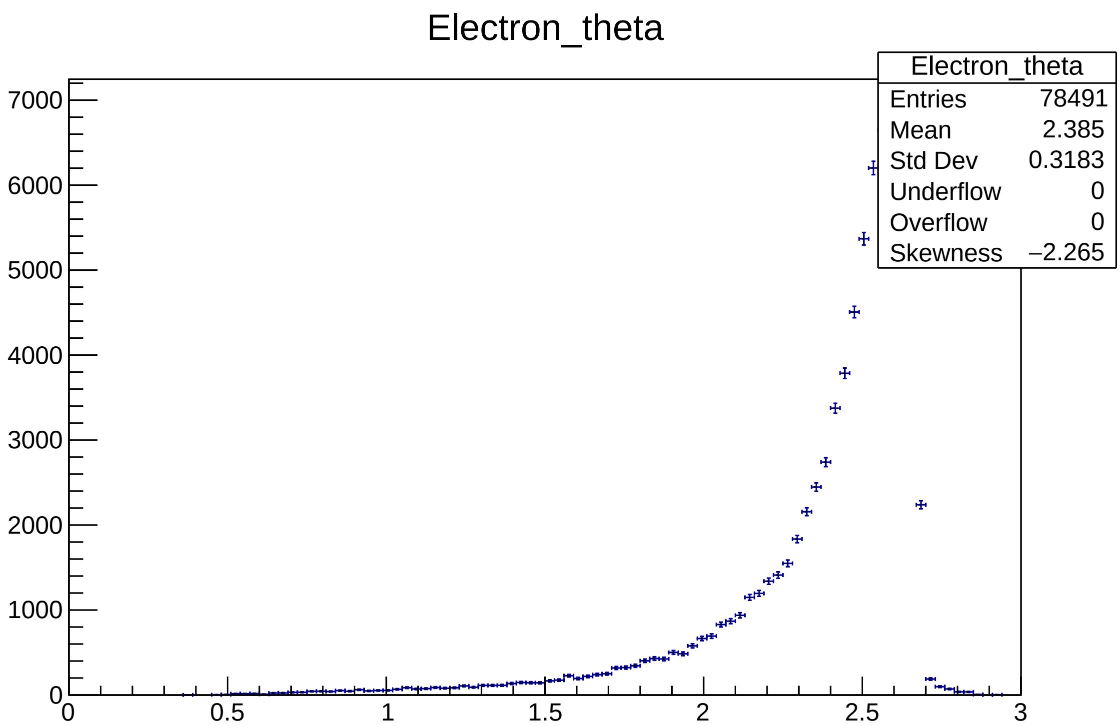



Figure 7: Electron polar angle $\theta$

## 5.2 Common variables

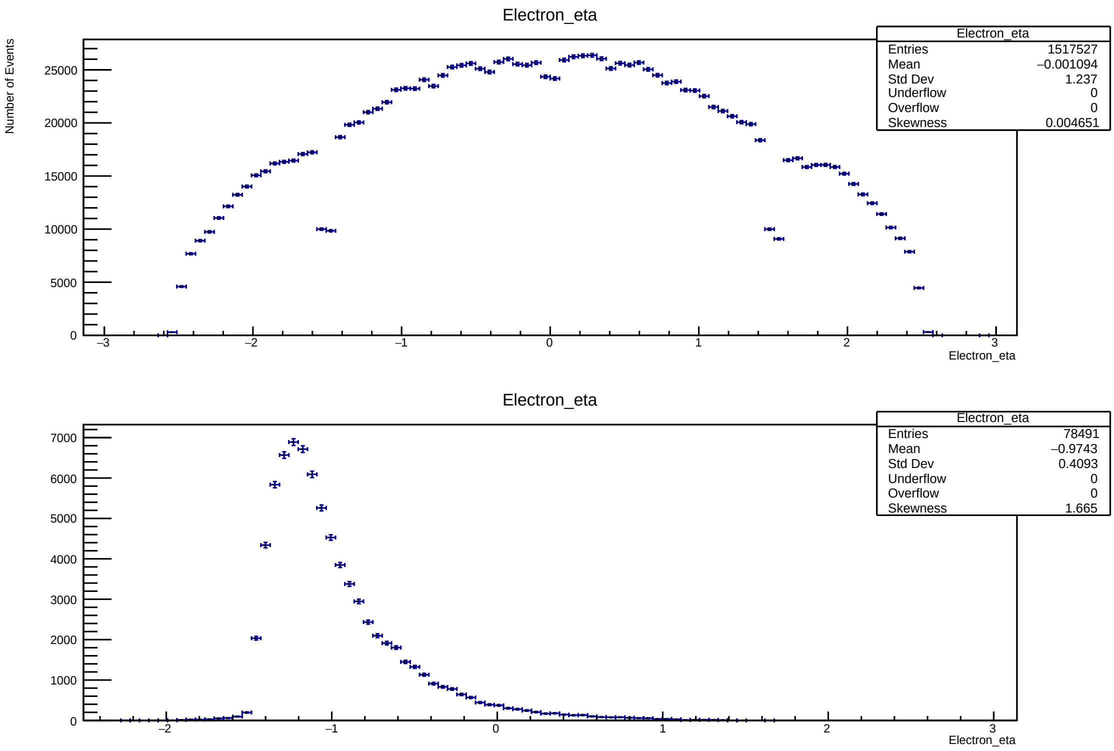



Figure 8: Electron $\eta$. CMS on the top and converted ZEUS below.

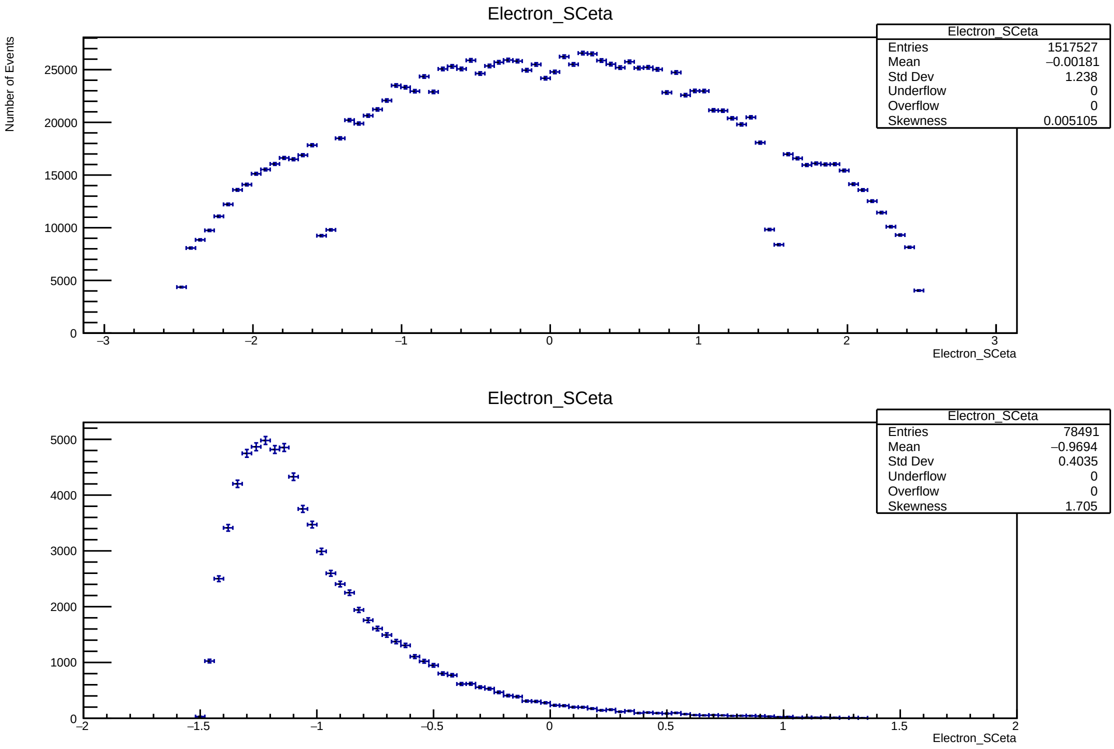



Figure 9: Electron $\eta$ taken from Super Cluster. CMS on the top and converted ZEUS below.

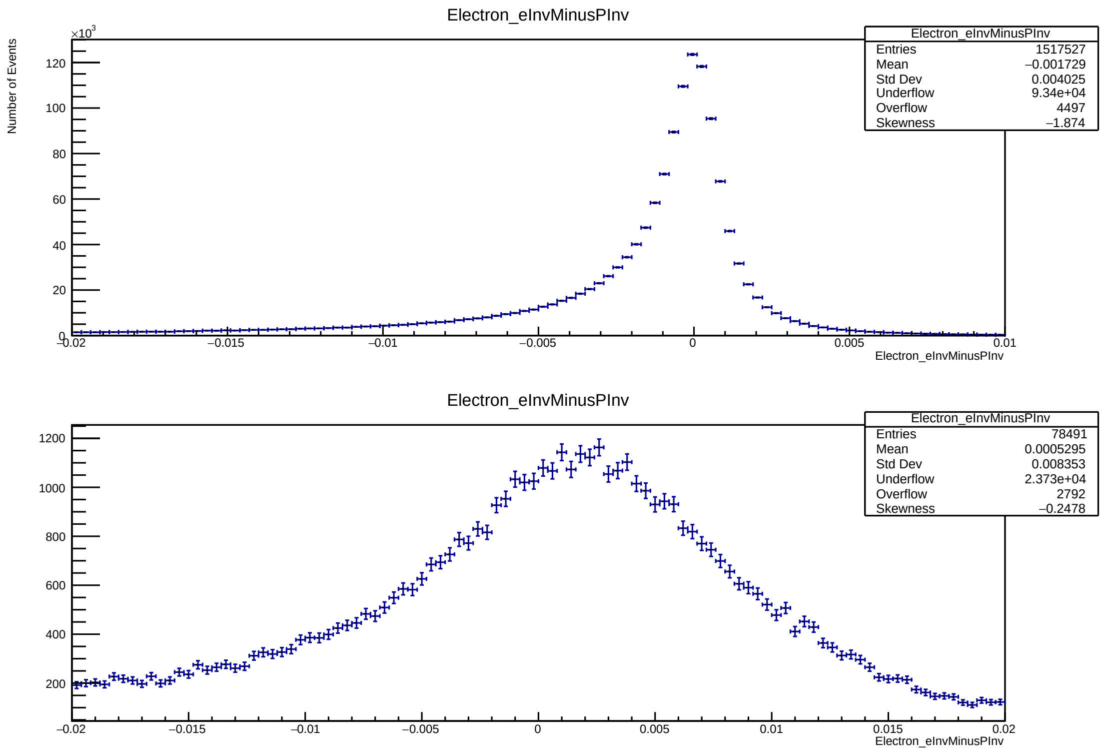



Figure 10: Electron_eInvMinusPInv. CMS on the top and converted ZEUS below.

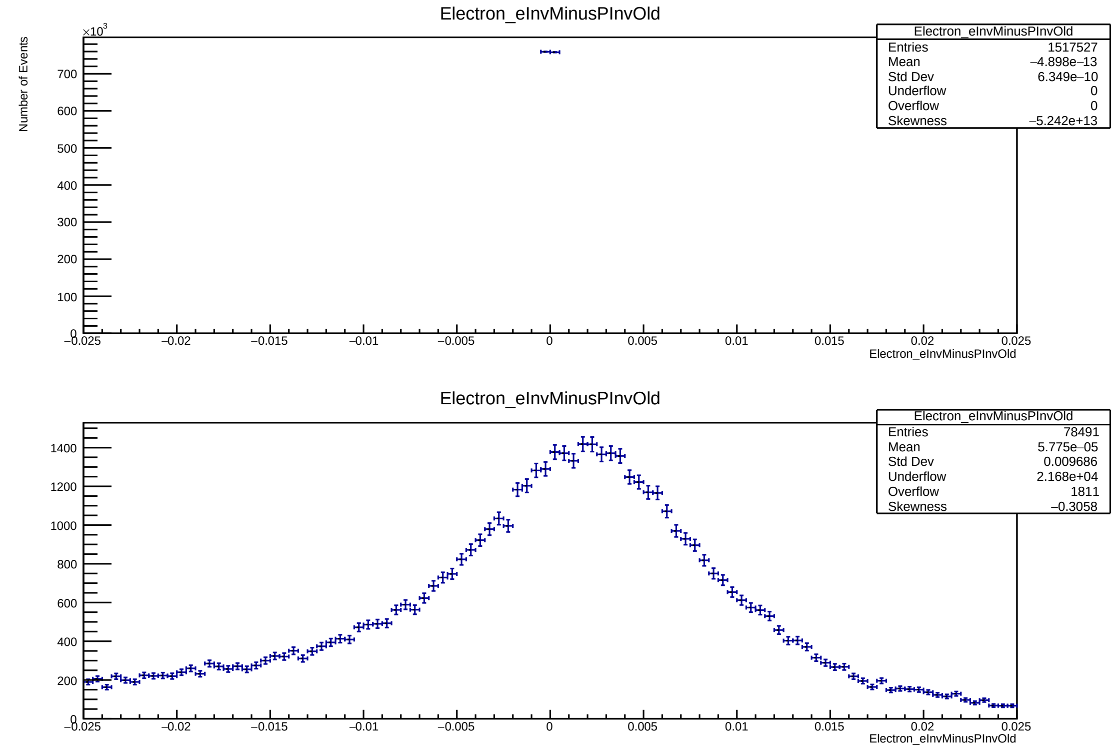



Figure 11: Electron_eInvMinusPInv. CMS on the top and converted ZEUS below.

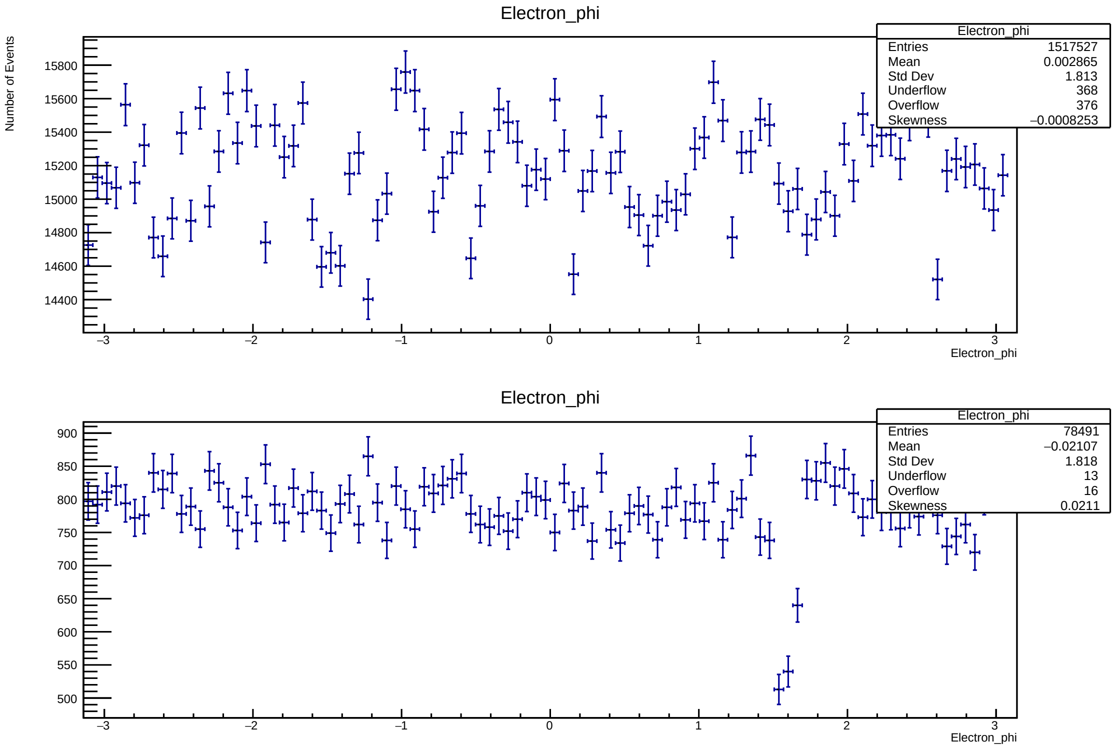



Figure 12: Electron angle $\phi$. CMS on the top and converted ZEUS below.

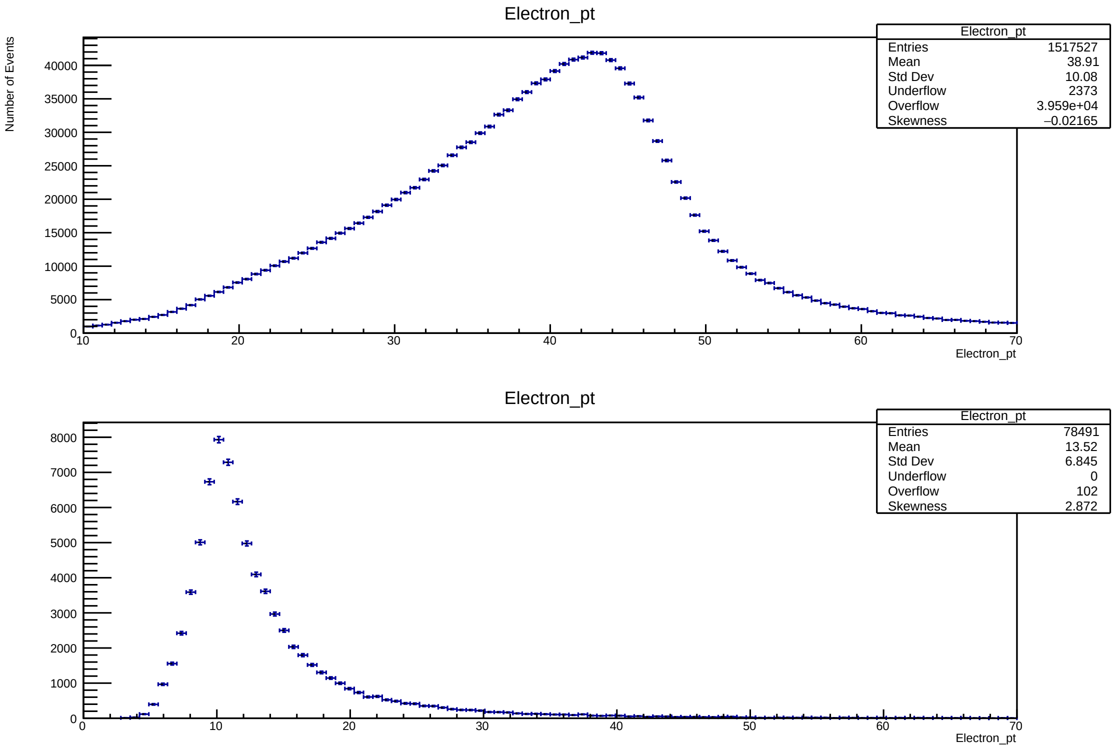



Figure 13: Electron $p_T$. CMS on the top and converted ZEUS below.

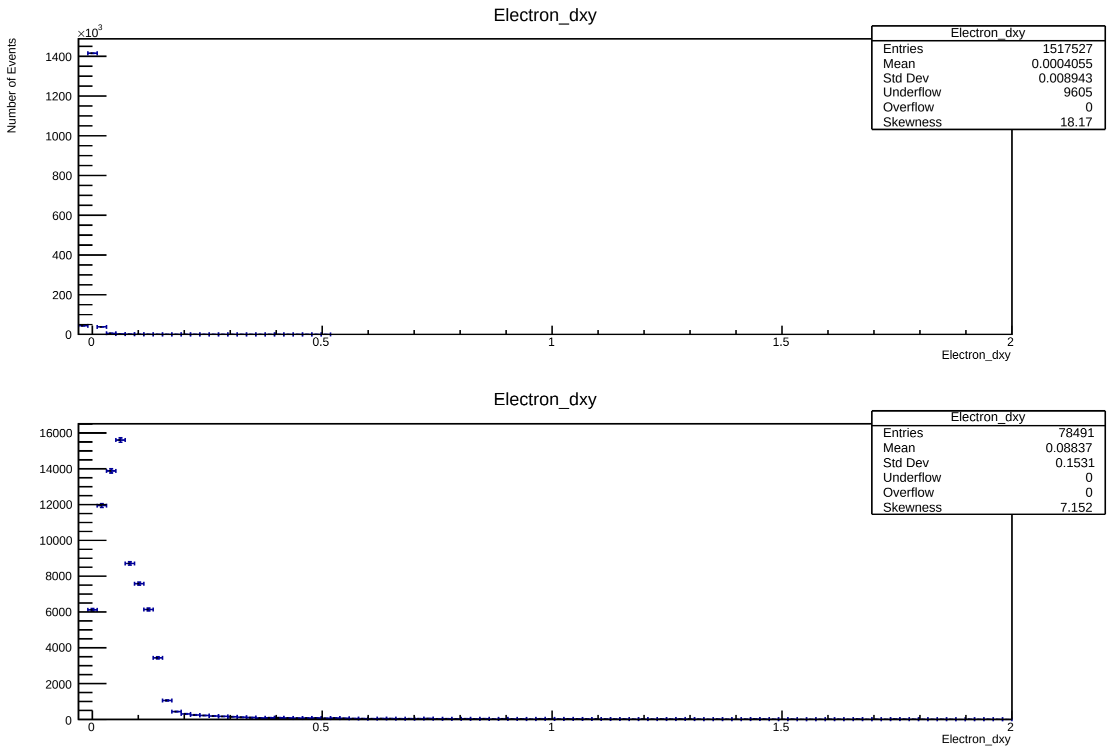



Figure 14: Electron $dxy$. CMS on the top and converted ZEUS below.

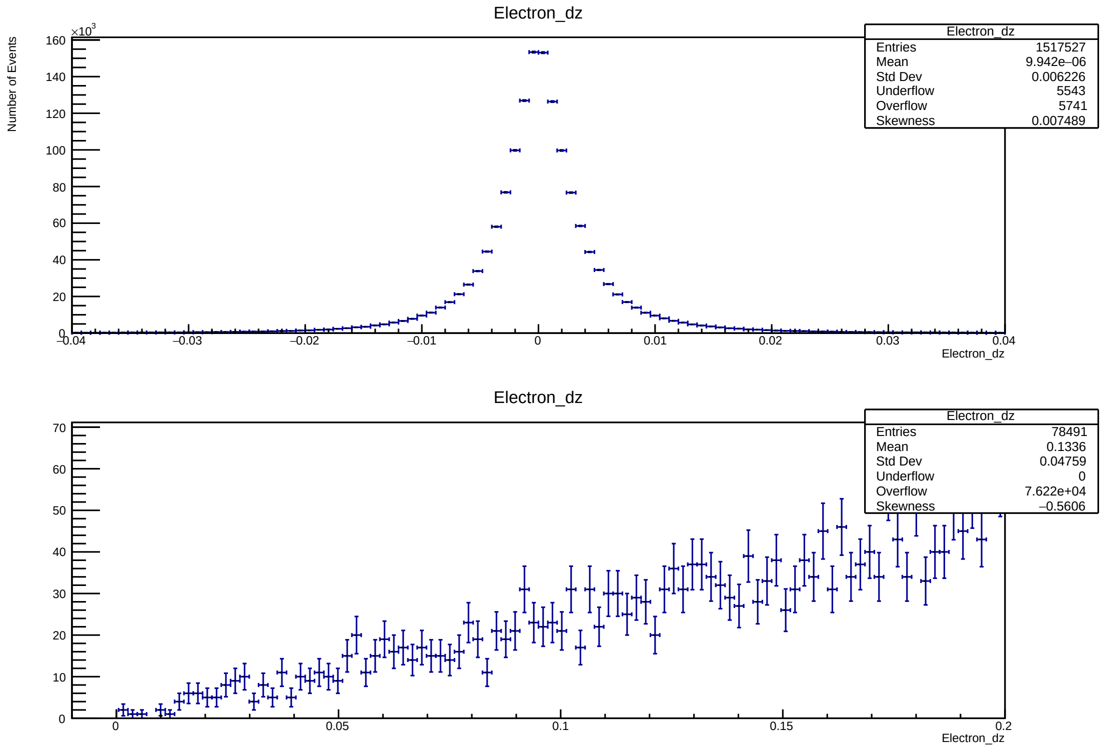



Figure 15: Electron $d_z$. CMS on the top and converted ZEUS below.

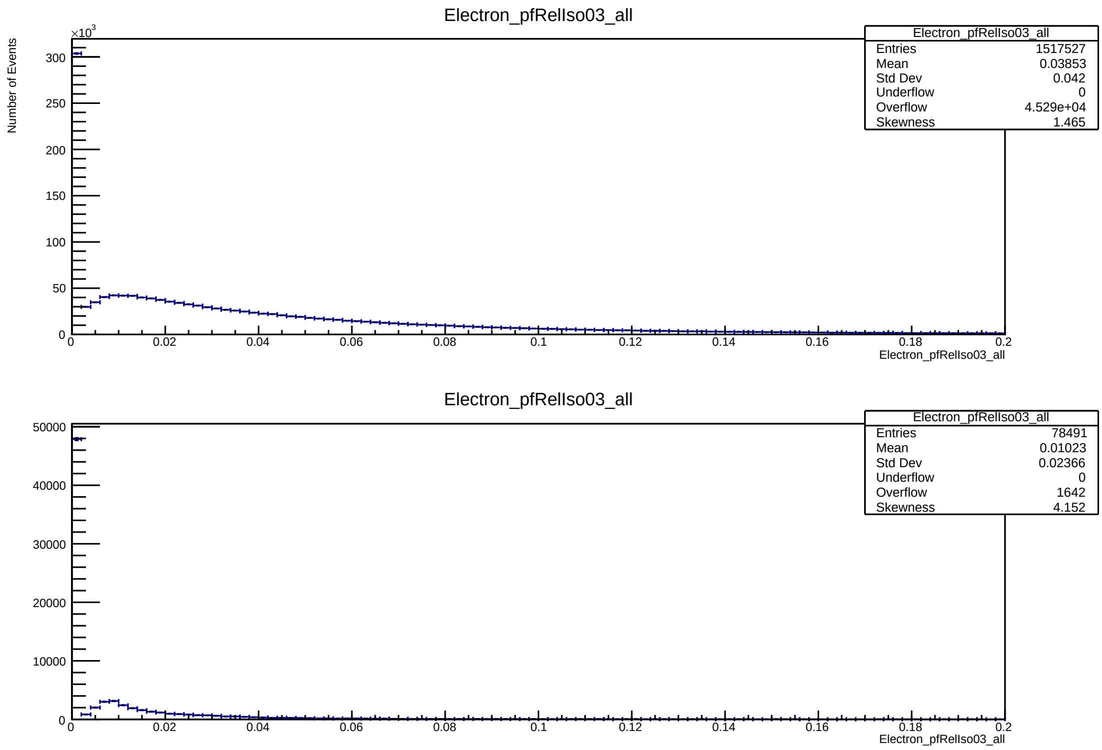



Figure 16: PF relative isolation $dR = 0.3$. CMS on the top and converted ZEUS below.

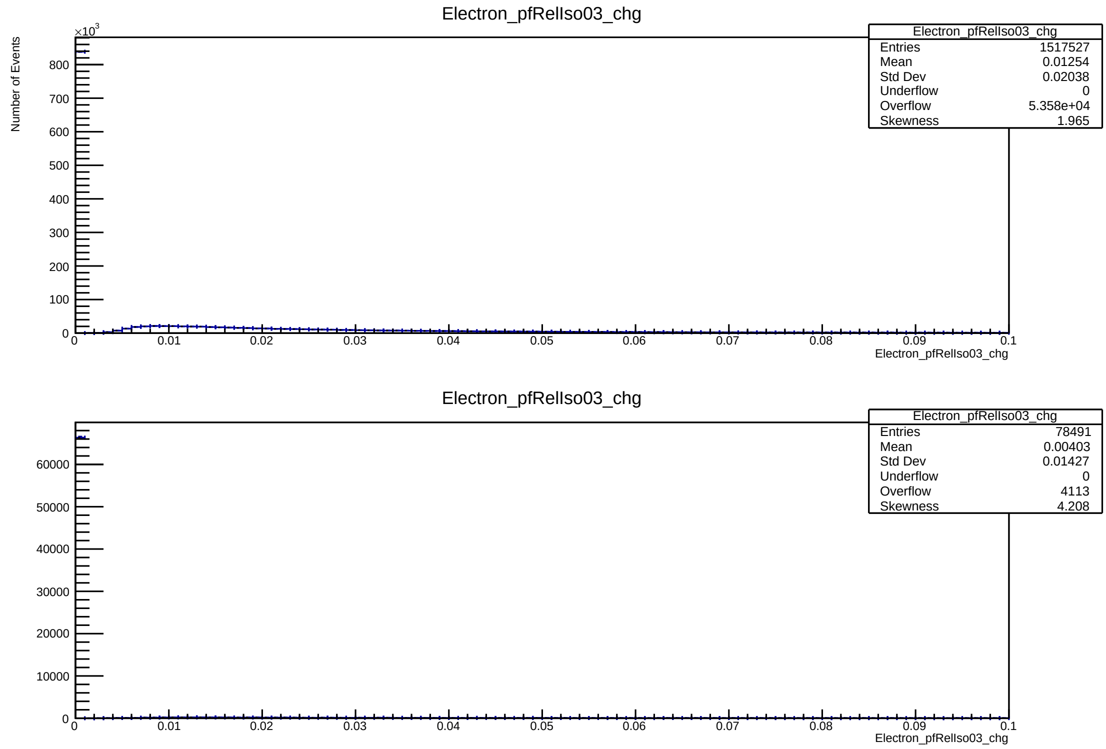



Figure 17: PF relative isolation $dR = 0.3$, charged component. CMS on the top and converted ZEUS below.

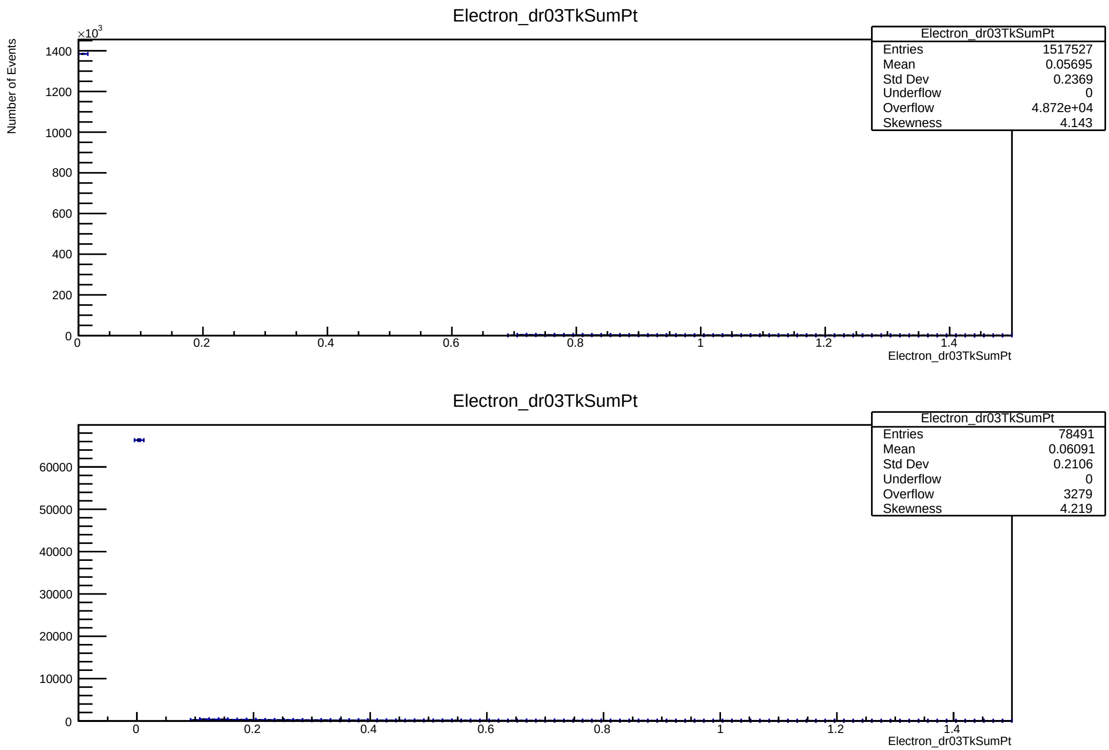



Figure 18: Non-PF track isolation within a delta R cone of 0.3 with electron $p_T > 35$ GeV. CMS on the top and converted ZEUS below.

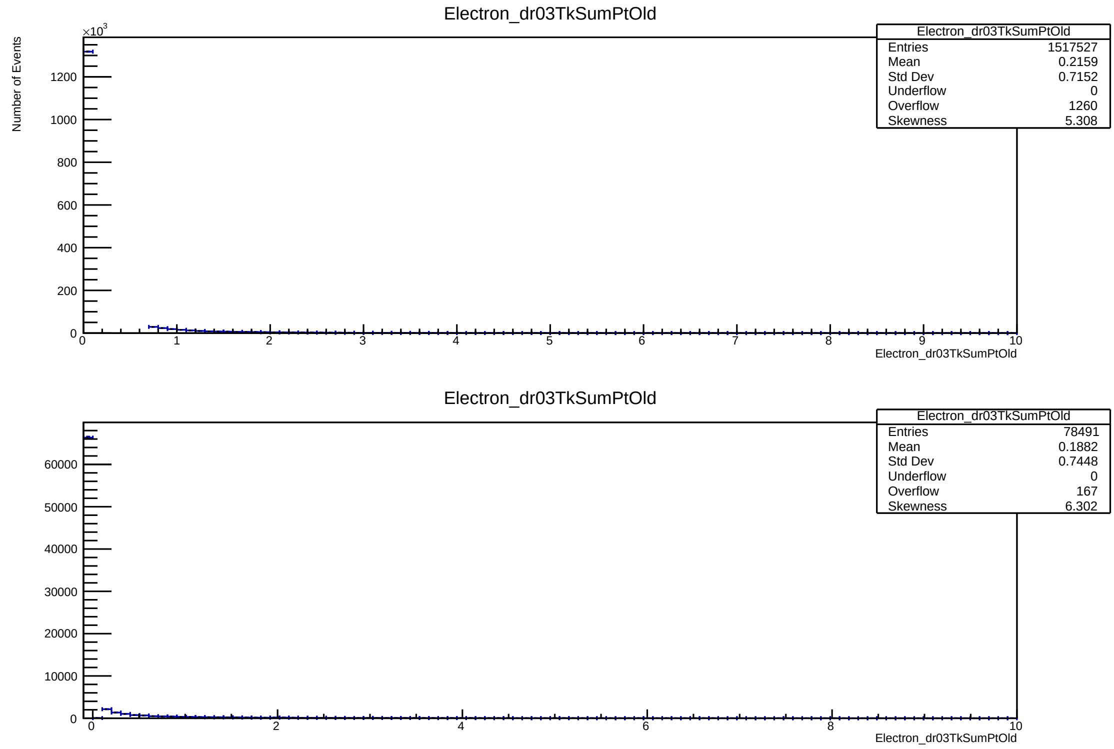



Figure 19: old: Non-PF track isolation within a delta R cone of 0.3. CMS on the top and converted ZEUS below.

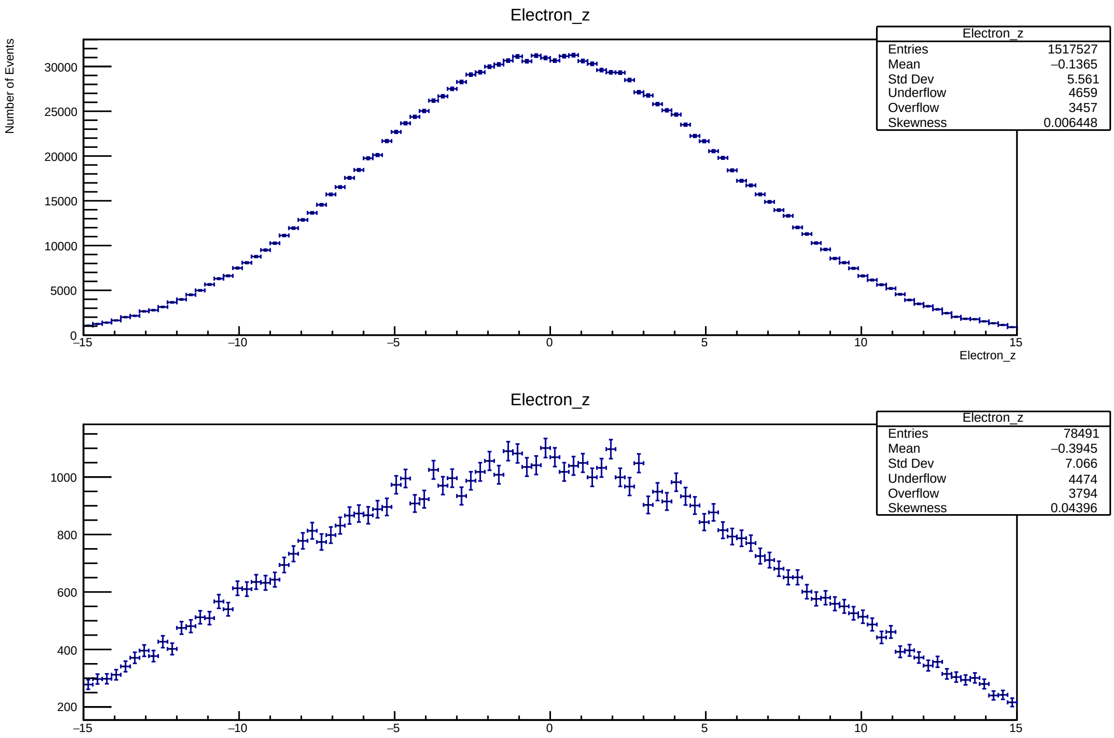
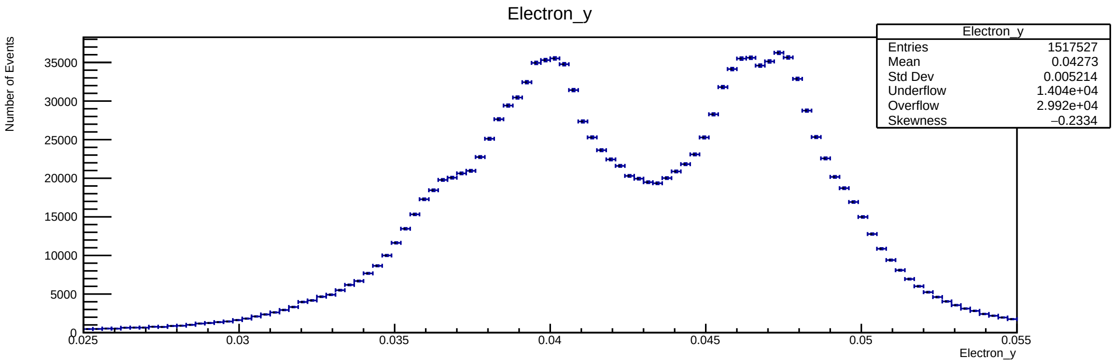



Figure 20: $z$ point of closest approach to the beamline, in cm. CMS on the top and converted ZEUS below.



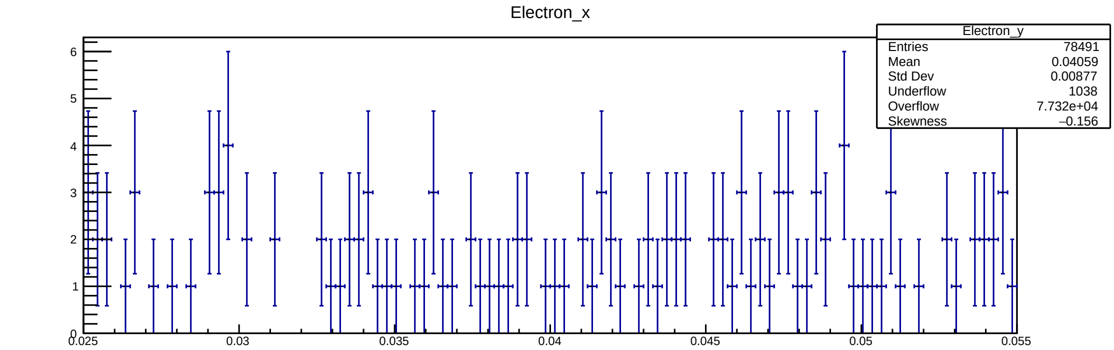
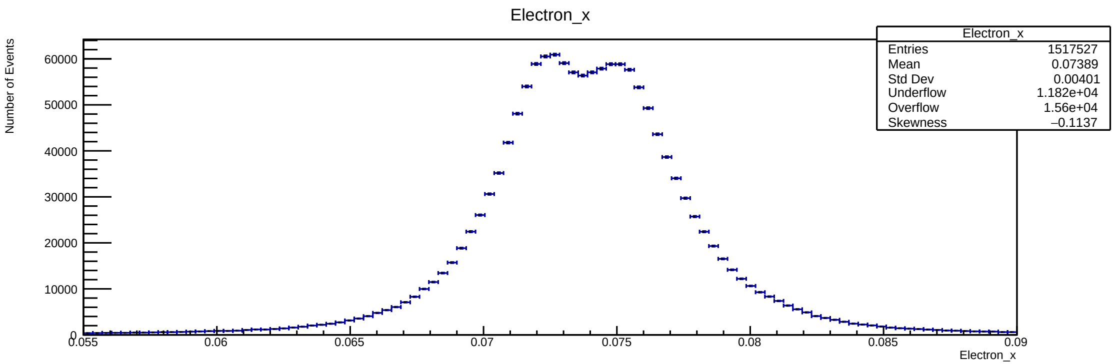




Figure 21: $y$ point of closest approach to the beamline, in cm. CMS on the top and converted ZEUS below.



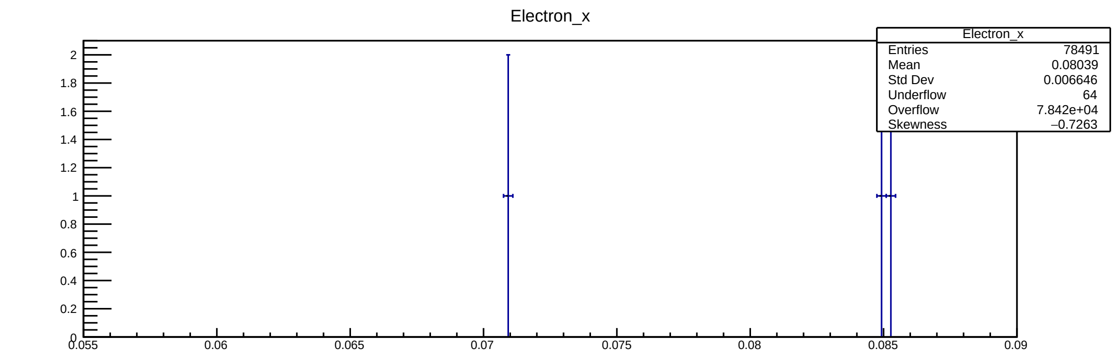




Figure 22: $x$ point of closest approach to the beamline, in cm. CMS on the top and converted ZEUS below.

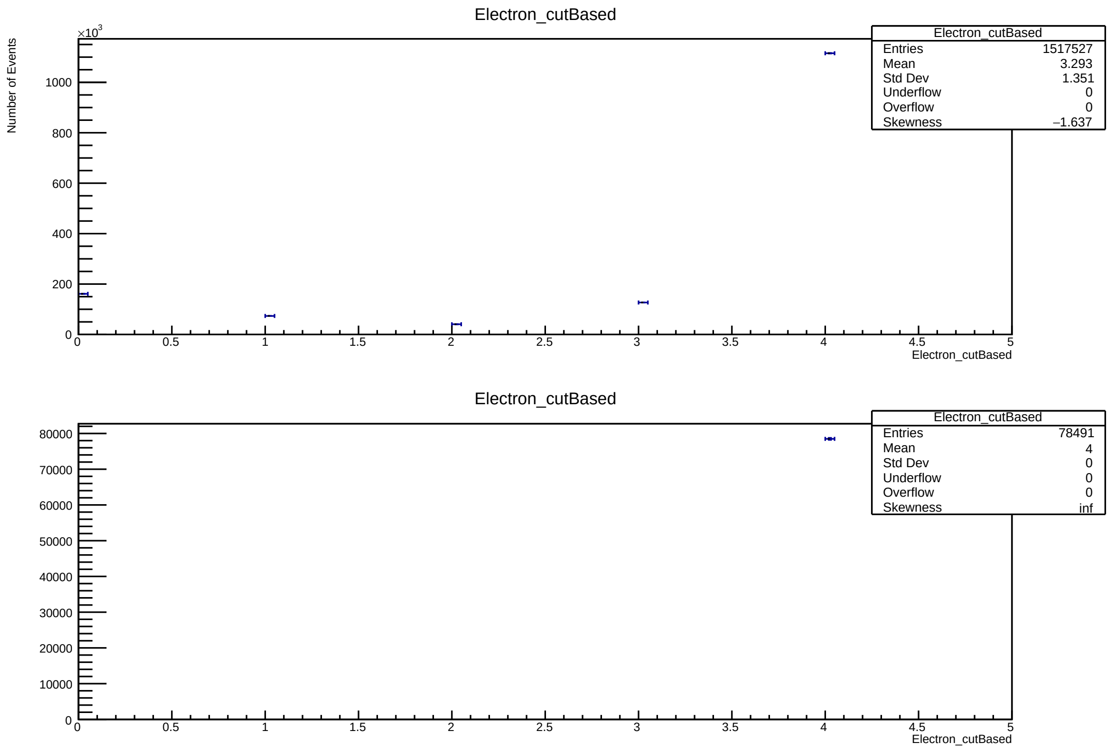



Figure 23: cut-based ID. CMS on the top and converted ZEUS below.

## 5.3 $Z^0$ invariant mass distribution

The $Z^0 \to e^+e^-$ invariant mass was reconstructed using the CMS data.

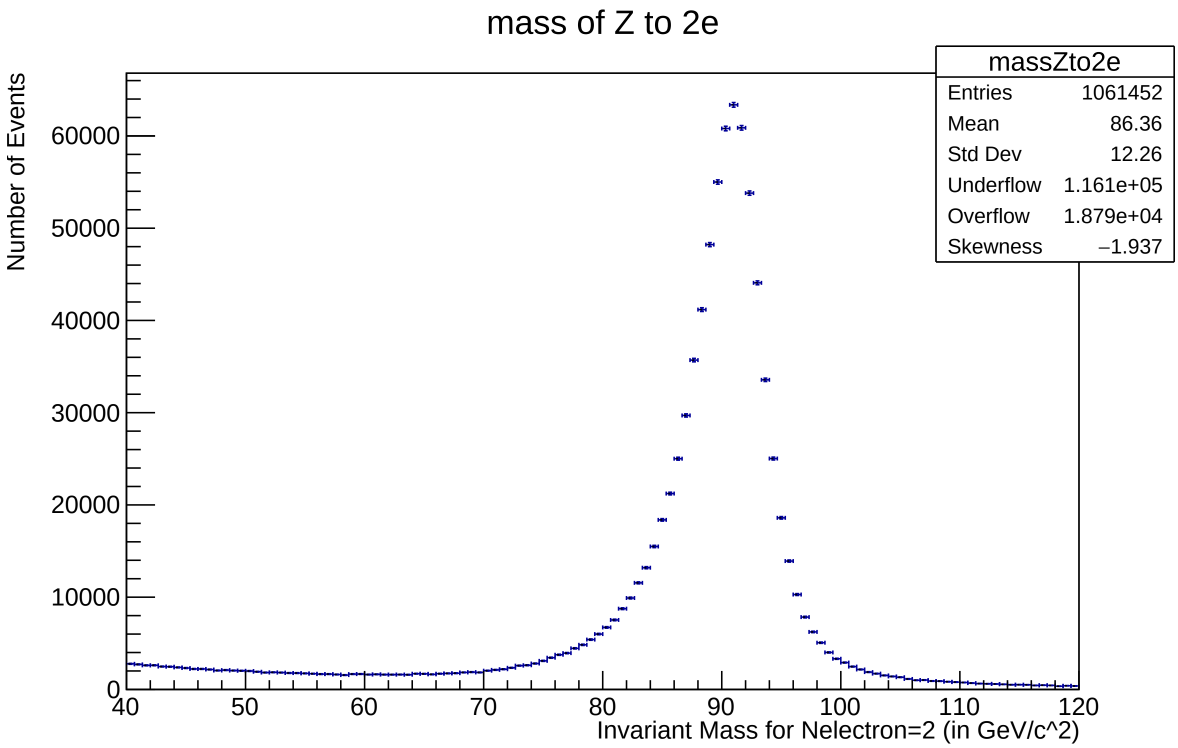



Figure 24: $Z^0 \to e^+e^-$ invariant mass distribution in GeV

## 5.4 CMS only variables

Some histograms of CMS variables that do not have appropriate ZEUS variables can be seen on the following pages. Also, the complete list of missed CMS variables can be found in the Further investigation section in Appendix.

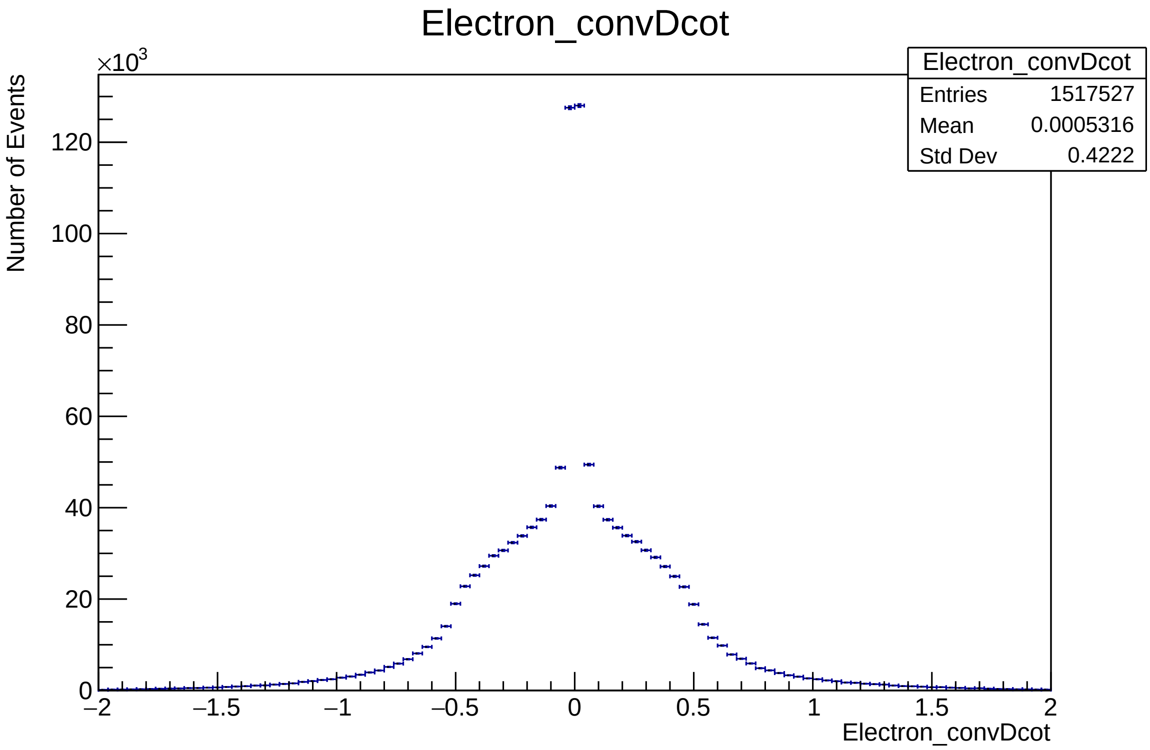



Figure 25: NanoAODplus extension. CMS on the top and converted ZEUS below.

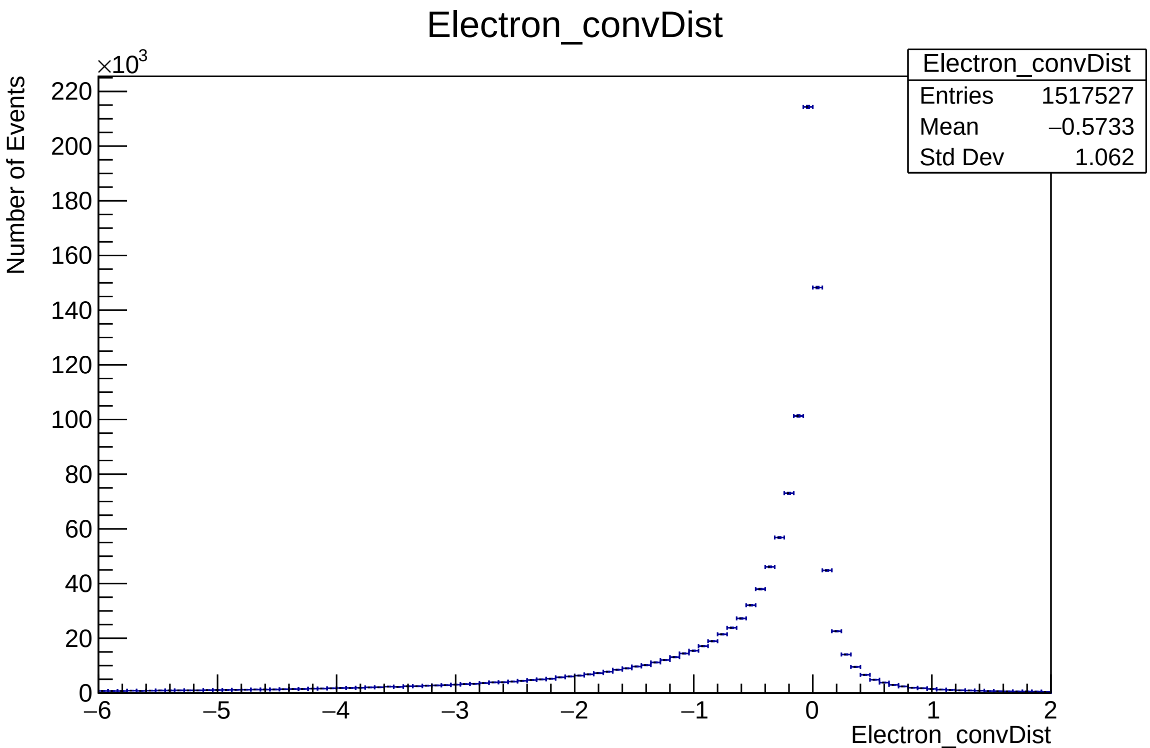



Figure 26: NanoAODplus extension. CMS on the top and converted ZEUS below.

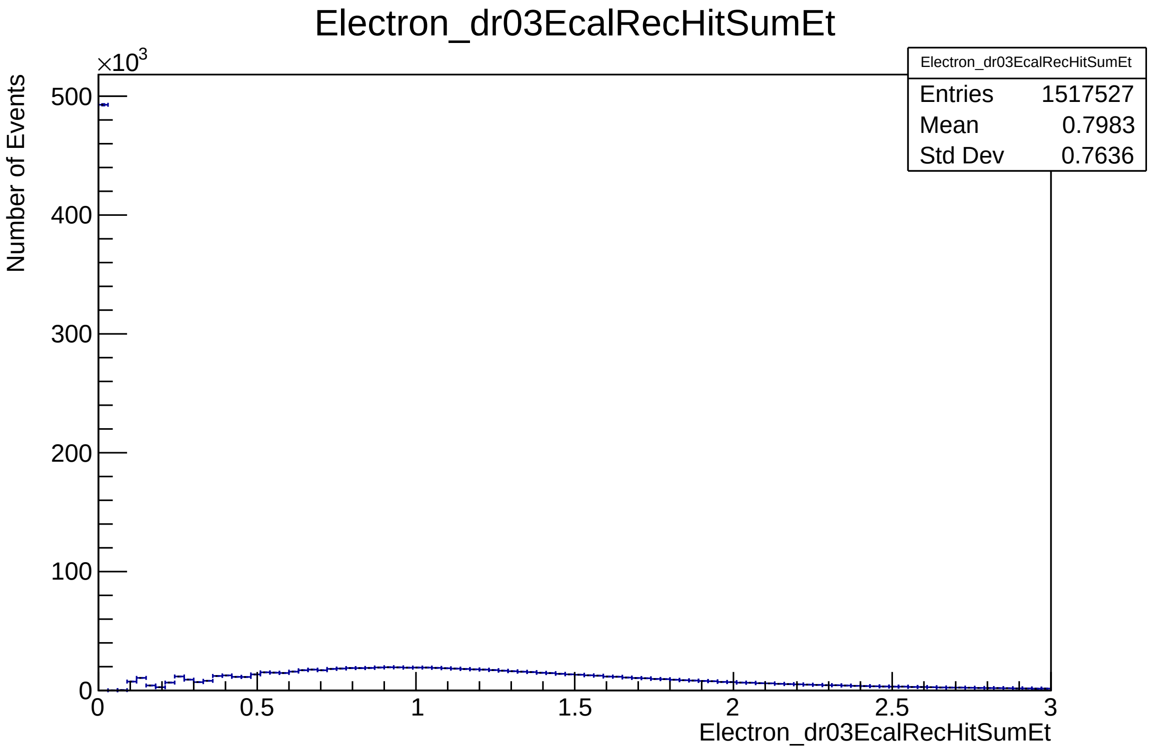



Figure 27: Non-PF Ecal isolation within a delta R cone of 0.3 with electron pt ¿ 35 GeV. CMS on the top and converted ZEUS below.

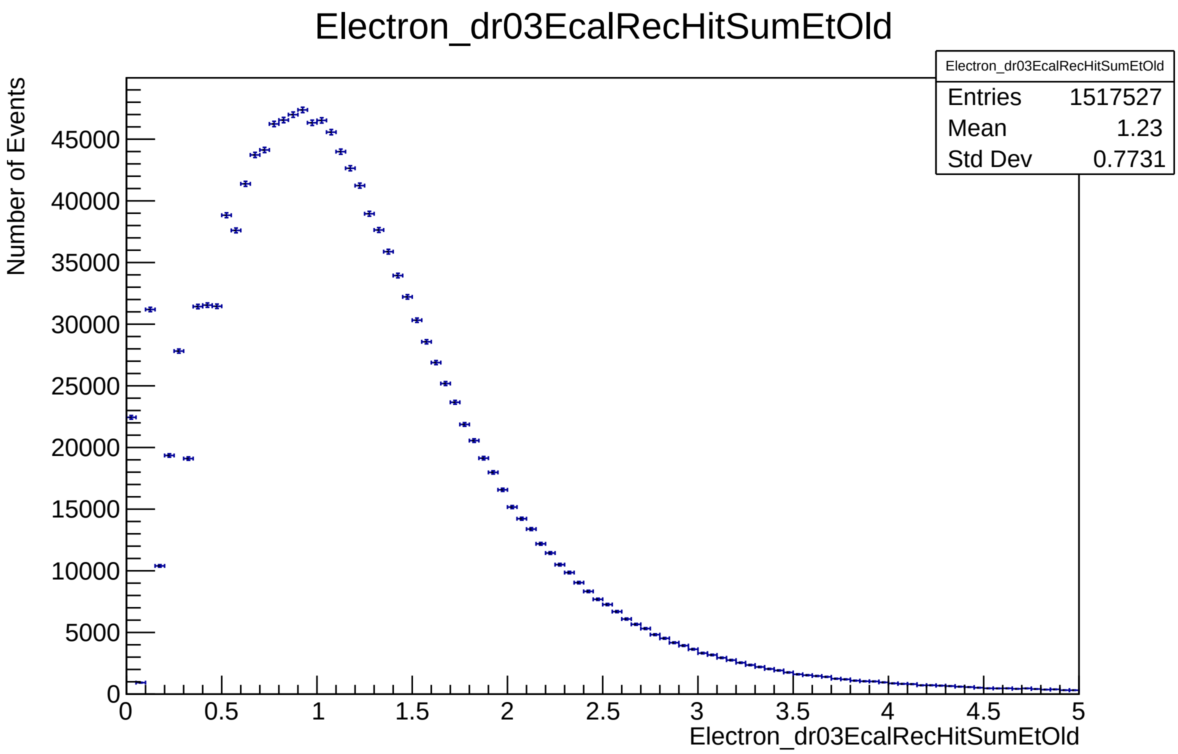



Figure 28: Non-PF Ecal isolation within a delta R cone of 0.3. CMS on the top and converted ZEUS below.

# 6 Appendix

## 6.1 Tables of variables

| **CMS variable** | **Conversion Formula** |
|---|---|
| Electron_charge | Emtrkq |
| Electron_dr03TkSumPt | Emetneartrk[1] |
| Electron_dr03TkSumPtOld | Emetneartrk[1] |
| Electron_dxy | Emdcabeam |
| Electron_dz | sqrt( Emdca$\hat{2}$ - Emdcabeam$\hat{2}$ ) |
| Electron_eInvMinusPInv | 1/Emcalene - 1/Emtrkp |
| Electron_eta | $-$ ln ( tan ( Emtrkth / 2 ) ) |
| Electron_genPartIdx | -1 |
| Electron_hoe | (1/Emfemc) - 1 |
| Electron_ip3d | Emdca |
| Electron_isEB | -120 $<$ Empos[2] $<$ 200 |
| Electron_isEE | !(-120 $<$ Empos[2] $<$ 200) |
| Electron_isNano | true for all |
| Electron_isPFcand | true for all |
| Electron_mass | take it from PDG |
| Electron_nNano | Emncand |
| Electron_pfRelIso03_all | Emenin / Empt |
| Electron_pfRelIso03_chg | Emetneartrk[1] / Empt |
| Electron_phi | Emtrkph |
| Electron_pt | Empt |
| Electron_simId | -1 |
| Electron_vtxIdx | -1 for each candidate |
| Electron_x | Trk_pca[0] |
| Electron_y | Trk_pca[1] |
| Electron_z | Trk_pca[2] |
| nElectron | Emncand |

## 6.2 Further investigation

In this subsection, the variables that have not been converted are listed. These are:

- Electron_convDcot
- Electron_convDist
- Electron_convVeto
- Electron_cutBased
- Electron_deltaEtaSC
- Electron_deltaEtaSCtr
- Electron_deltaPhiSC
- Electron_deltaPhiSCtr
- Electron_dr03EcalRecHitSumEt
- Electron_dr03EcalRecHitSumEtOld
- Electron_dr03HcalDepth1TowerSumEt
- Electron_dr03HcalDepth1TowerSumEtOld
- Electron_dr03HcalTowerSumEt
- Electron_eInvMinusPInvOld
- Electron_sieie
- Electron_sieieR1
- Electron_sip3d
- Electron_tightCharge

These variables are needed further investigations that are beyond the frame of the Summer Program.

# References

[1] will be completed in the final version ... *Author name*