Available on CMS information server

**CMS AN 2010/323**

# CMS Analysis Note



**October 27, 2010**

# Electron Efficiency Measurements with 2.88pb$^{-1}$ of pp Collision Data at $\sqrt{s} = 7$ TeV

J. Berryhill and K. Mishra

*Fermilab, USA*

G. Daskalakis

*NCSR, Greece*

V. Halyo and J. Werner

*Princeton University, USA*

S. Xie

*MIT, USA*

**Abstract**

We present the first data driven electron efficiency measurements performed at CMS. These measurements were performed as part of the W/Z analysis, and use 2.88pb$^{-1}$ of pp Collision Data at $\sqrt{s} = 7$TeV.

# 1 Motivation

Electron efficiency measurements are a central component of any physics analysis with electrons in the final state. The purpose of this note is to briefly describe the techniques and results of the electron efficiency measurements for the W and Z cross-section measurements of CMS AN-264/2010 [2]. We describe the tag and probe methodology used with two alternative signal shape hypotheses. All relevant systematics are evaluated, and the results are detailed in multiple tables and plots. These efficiency measurements, along with the cross section measurements of CMS AN-264/2010, were performed on a dataset corresponding to an integrated luminosity of $2.88\text{pb}^{-1}$ of pp collisisions at $\sqrt{s} = 7\text{TeV}$. This data was collected at the CMS detector between March 30 - Aug 30, 2010, including runs 132440-144114.

# 2 Introduction to Tag and Probe Methodology

The tag and probe method for computing electron efficiencis leverages Z→ee decays as a high-purity souce of unbiased electrons from which to extract efficiencies. The so called *tag* electron is the control electron to which stringent electron selection criteria is applied. The *probe* electron, on the other hand, is the test electron having selection criteria according to the efficiency under study. Naturally, imposing an invariant mass cut on the tag-probe pair, about the Z→ee peak ensures a high purity sample of tag-probe pairs. This method is well documented, for instance, in CMS AN-2009/111 [1].

After constructing the tag criteria and the probe criteria for the specific efficiency under study, two samples then arise: events with one tag electron and one probe electron passing the particular selection requirements under study (Tag+Pass) and events with a tag electron and a probe failing the selection requirement under study (Tag+Fail). The efficiency calculation then reduces to an estimation of the signal yield in the passing and failing samples. In the present analysis this estimation is carried out via simultaneous maximum likelihood fits to the dielectron invariant mass distributions in the two samples.

For the W and Z cross section measurement, the efficiency calculation is broken down into 3 steps: (1) Super Cluster (SC) to Gaussian Sum Filter track-matched electron (Reco), (2) Reco to the working point 80% selection criteria (WP80), (3) WP80 to the online trigger requirement (HLT). Since the electron selection and background levels are different for electrons in the ECAL barrel and ECAL endcap, efficiencies are measured separately for barrel and endcap. Due to possible correlations between electron selection efficiency and the level of misalignment between the tracker and the ECAL, the efficiency may have residual charge dependance. Therefore the measurement is performed separately for electrons and positrons.

# 3 Efficiency Fit

In order to extract the signal components of the Tag+Pass and Tag+Fail samples, an unbinned maximum likelihood fit is performed in the dielectron invariant mass variable with specific signal models described in sections 4.1 and 4.2, and an exponential background shape. The efficiency enters as an explicit fit parameter, such that correlations are taken into account in the efficiency uncertainties extracted from the fit. The floating parameters of the fit are listed below:

- $N_{\text{Signal}}$ : number of signal Tag+Probe pairs,
- $N_{\text{Bkg}}^{\text{pass}}$ : number of background Tag+Pass pairs,
- $N_{\text{Bkg}}^{\text{fail}}$ : number of background Tag+Fail pairs,
- $\epsilon$ : efficiency,
- $\chi_{\text{pass}}$ : coefficient of the exponential background for the Tag+Pass sample,
- $\chi_{\text{fail}}$ : coefficient of the exponential background for the Tag+Fail sample.

The signal and background are extracted in the Tag+Pass and Tag+Fail samples through the following relations:

$$N^{\text{pass}} = N_{\text{Signal}}(\epsilon) P^{\text{pass}}_{\text{Signal}} + N^{\text{pass}}_{\text{Bkg}} P^{\text{pass}}_{\text{Bkg}} \quad (1)$$

$$N^{\text{fail}} = N_{\text{Signal}}(1-\epsilon) P^{\text{fail}}_{\text{Signal}} + N^{\text{fail}}_{\text{Bkg}} P^{\text{fail}}_{\text{Bkg}} \quad (2)$$

$$(3)$$

where $P^{\text{pass}}_{\text{Signal}}$ and $P^{\text{fail}}_{\text{Signal}}$ are the probability density functions in the dielectron invariant mass variable for the signal in the Tag+Pass and Tag+Fail samples, respectively. $P^{\text{pass}}_{\text{Bkg}}$ and $P^{\text{fail}}_{\text{Bkg}}$ are the probability density functions in the mass variable for the background and given by:

$$P^{\text{pass}}_{\text{Bkg}} \equiv e^{-\chi_{\text{pass}}} \quad (4)$$

$$P^{\text{fail}}_{\text{Bkg}} \equiv e^{-\chi_{\text{fail}}} \quad (5)$$

$$(6)$$

For the current results, sensitivity of the fit to the presence of background in the Tag+Pass sample is low. In order to avoid bias in the fit, we constrain the background in the Tag+Pass sample to be zero.

# 4 Signal Shape Parameterization

## 4.1 Simulation Signal Shape With Additional Gaussian Smearing

In this approach, we construct the Tag+Pass and Tag+Fail mass shapes for the signal from the Monte Carlo simulation using the identical selection requirements that is used for data. In order to account for any effects not included in the simulation, we build the final signal shape probability density function by a convolution of the simulation line shape with an additional Gaussian resolution function. The width and mean of the gaussian is allowed to float in the fit, accounting for an additional smearing to the simulation mass shape and a shift in the Z mass peak due to miscalibration of the energy scale.

## 4.2 Generator-Level Shape Convolved with a Resolution Model

One possible parameterization of the lineshape is to take the signal shape directly from the output of a next to leading order Monte-Carlo generator and then convolve it with some detector resolution model. This is what is done in this section. The generator shape is extracted directly from the POWHEG [3] Monte-Carlo generator. The resolution model is a Crystal-Ball modified to include an extra half Gaussian on the high end tail. This resolution model is purely empirical: Before data taking commmenced, it was observed to describe the detector simulation quite well and was later confirmed to also describe the data well. With 2.88pb$^{-1}$ of data, many of the bins we perform fits in are statistics limited. Therefore at the moment we must take a rather ad-hoc approach to which parameters of the line shape model we allow to float: To summarize, we allow as many parameters to float as possible under the constraints that the fit converges and MINOS [4] is still able to properly calculate the positive and negative uncertainties on each of the floating parameters. In addition to the evaluation of the MINOS uncertainties, detailed pull studies have also been performed so that the statistical features of this fit are thorougly understood to be stable.

# 5 Systematic Uncertainties

## 5.1 Signal Shape Systematics

Due to the large amount of material associated with the tracker and its support structures that an electron must traverse before reaching the electromagnetic calorimeter, a relatively prominent low mass tail is present from bremstrahlung and energy loss. The precise behavior of this tail in the signal mass shape, particularly in the Tag+Fail sample, can significantly affect the results of the efficiency measurement fit. Due to the fact that this efficiency measurement has been performed on the prompt reconstruction data, whose alignment scenario is non-ideal, one is particularly worried about the existence of correlations

between inefficiencies from selection cuts affected by misalignment and the shape of the signal mass in the low mass region. We attempt to constrain the shape of the low mass tail using data driven techniques.

Two different signal shape templates are selected from data, for the Tag+Fail sample. The first template is constructed from events where we require that the tag electron passes the very tight WP60 electron selection, and require that the missing transverse energy is less than 20.0 GeV in order to suppress W+Jet background. The probe is required to fail the nominal WP80 selection. The resulting mass shape is shown in Figure 1, compared with the shape extracted from the Monte Carlo simulation using the same selection. The second template is constructed from events where we require the usual WP80 electron selection for the tag electron, but require that the probe passes the WP95 isolation cuts and fails the remaining WP80 selection cuts. This signal template is shown in Figure 2.

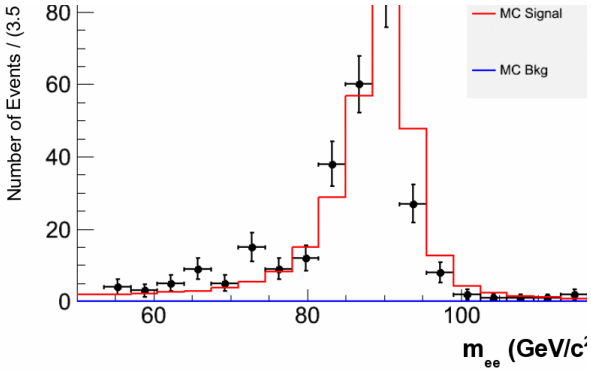



Figure 1: Signal mass shape template extracted from data by selecting events with one electron passing the WP60 electron selection and a probe which fails the WP80 selection, and requiring that the missing transverse energy is less than 20.0 GeV.

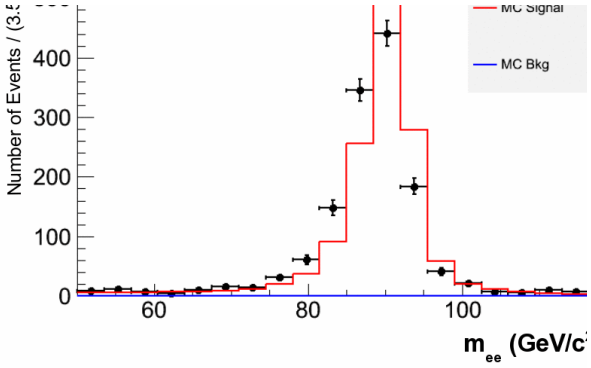



Figure 2: Signal mass shape template extracted from data by selecting events with one electron passing the WP80 electron selection and a probe which passes the WP95 isolation cuts but fails the WP80 selection.

To extract systematic uncertainties due to the signal shape, we perform the efficiency fit using these two fixed templates for the signal shape of the failing sample. The difference between the resulting efficiency measurement and the efficiency measured using the nominal method is taken as a systematic uncertainty. The results of this fit are shown in Figure 3 and 4 for the first and second template respectively. Table 1 summarizes the systematic uncertainties from this procedure. The systematic uncertainties in the endcap are artificially small due to the lack of any events in the mass sidebands in the Tag+Fail sample from tag and probe selection, simply due to the small sample size. A conservative systematic uncertainty of 2% is assigned to both the barrel and the endcap.

## 5.2 Background Shape Systematics

To estimate the systematic due to the background parameterization, large numbers of pseudo-experiments are generated with a background PDF of the form $1/\mathrm{M}^{alpha}$, where $\alpha = 5$. For each pseudoexperiment

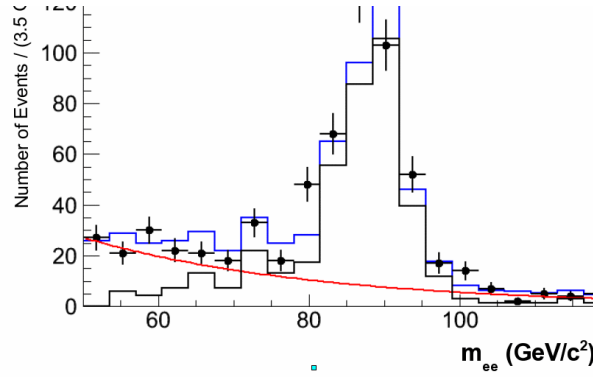



Figure 3: Fit result for the Tag+Fail sample, using the signal template constructed from selecting WP60 Tag electrons and requiring a missing transverse energy cut. The red curve indicated the fitted background, the black histogram is the fitted signal, and the blue histogram is the fit result for the signal plus background.

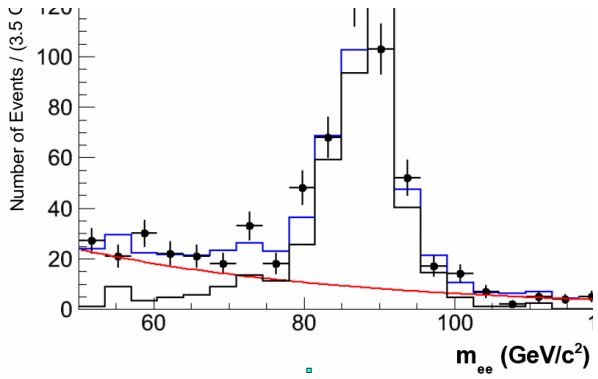



Figure 4: Fit result for the Tag+Fail sample, using the signal template constructed from selecting WP80 Tag electrons and requiring that the probe passes the WP95 isolation cuts. The red curve indicated the fitted background, the black histogram is the fitted signal, and the blue histogram is the fit result for the signal plus background.

| Data Based Signal Template | Barrel | Endcap |
|---|---|---|
| WP60 Tag + MetCut | 1.8% | 0.0% |
| WP80 Tag + Isolated Probe | 1.0% | 0.2% |

Table 1: Systematic uncertainties obtained from the difference between the efficiency measurement using the data driven signal shape templates and the efficiency measured using the nominal signal shape model.

S/B is fixed to the level reported back by the fit to the data for each efficiency step. This pseudo Monte-Carlo is then fit back to the sum of the signal shape from whence it was generated and several background shape hypothesis, including an exponential decay, and a power law. The systematic uncertainty is then quoted as the maximal difference among the remaining background hypotheses. The background systematics are found to be small, and are listed for each efficiency step in table 2.

### 5.3 Energy Scale Systematics

To estimate the systematic due to the uncertainty on the energy scale, the electron energies in the Monte-Carlo are shifted in both the positive and negative directions by the uncertainty on the scale. The maximal difference among the 3 extracted efficiency values (nominal, positive shift, negative shift) is then taken as the systematic for each step. The energy scale systematics are found to be quite small, and are listed for each efficiency step in table 3.

Table 2: The Background Parameterization Systematic Uncertainty for each Efficiency Step

| Step | $\Delta\epsilon$ |
|---|---|
| SC→Reco | 0.0006 |
| Reco→WP80 | 0.002 |
| WP80→HLT | 0.0000 |

Table 3: The Energy Scale Systematic Uncertainty for each Efficiency Step

| Step | $\Delta\epsilon$ |
|---|---|
| SC→Reco | 0.001 |
| Reco→WP80 | 0.002 |
| WP80→HLT | 0.001 |

Table 4: The SC→Reco Efficiencies

| Topology | Monte-Carlo | | Data | | | | Data/Monte-Carlo | Δ(Data/Monte-Carlo) |
|---|---|---|---|---|---|---|---|---|
| | $\epsilon$ | $\Delta\epsilon$ | $\epsilon$ | $\Delta\epsilon$ (stat) | $\Delta\epsilon$ (syst) | $\Delta\epsilon$ (total) | | |
| EB | 0.9851 | 0.0002 | 0.986 | 0.004 | 0.012 | 0.013 | 1.001 | 0.013 |
| EE | 0.9629 | 0.0004 | 0.962 | 0.006 | 0.012 | 0.013 | 0.999 | 0.012 |
| EB $e^-$ | 0.9854 | 0.0002 | 0.985 | 0.003 | 0.012 | 0.012 | 0.999 | 0.012 |
| EE $e^-$ | 0.9624 | 0.0004 | 0.967 | 0.002 | 0.012 | 0.012 | 1.005 | 0.012 |
| EB $e^+$ | 0.9849 | 0.0002 | 0.987 | 0.003 | 0.012 | 0.012 | 1.002 | 0.012 |
| EE $e^+$ | 0.9634 | 0.0004 | 0.951 | 0.002 | 0.012 | 0.012 | 0.987 | 0.012 |

Table 5: The Reco→WP80 Efficiencies

| Topology | Monte-Carlo | | Data | | | | Data/Monte-Carlo | Δ(Data/Monte-Carlo) |
|---|---|---|---|---|---|---|---|---|
| | $\epsilon$ | $\Delta\epsilon$ | $\epsilon$ | $\Delta\epsilon$ (stat) | $\Delta\epsilon$ (syst) | $\Delta\epsilon$ (total) | | |
| EB | 0.8547 | 0.0004 | 0.791 | 0.014 | 0.020 | 0.024 | 0.925 | 0.028 |
| EE | 0.7488 | 0.0006 | 0.692 | 0.016 | 0.020 | 0.026 | 0.924 | 0.035 |
| EB $e^-$ | 0.8545 | 0.0005 | 0.793 | 0.021 | 0.020 | 0.029 | 0.929 | 0.034 |
| EE $e^-$ | 0.7489 | 0.0009 | 0.701 | 0.030 | 0.020 | 0.036 | 0.936 | 0.048 |
| EB $e^+$ | 0.8549 | 0.0005 | 0.780 | 0.024 | 0.020 | 0.031 | 0.913 | 0.036 |
| EE $e^+$ | 0.7487 | 0.0009 | 0.701 | 0.023 | 0.020 | 0.030 | 0.936 | 0.040 |

### 5.4 Results

In this section we provide the efficiency results using the generator-level line shape described in Section 4.2. These results are summarized in Tables 4 - 6, including all statistical and systematic uncertainties. The corresponding fit plots are provided in Appendix B. Results from the Monte-Carlo are also provided, and these were evaluated via simple cut-and-count on the Monte-Carlo truth. For comparison and cross check the results using the alternative simulation based signal shape model are summarized in Appendix A. The two approaches yield efficiencies that are in agreement to within the uncertainties.

## 6 Conclusions

We have presented the first electron efficiency measurements performed at CMS. These measurements were performed as part of the W and Z inclusive cross section measurements, and used 2.88pb$^{-1}$ of pp Collision Data at $\sqrt{s} = 7$ TeV. This note has summarized these meausurements, described all of the associated systematic uncertainties.

Table 6: The WP80→HLT Efficiencies

| Topology | Monte-Carlo | | Data | | | | Data/Monte-Carlo | Δ(Data/Monte-Carlo) |
|---|---|---|---|---|---|---|---|---|
| | $\epsilon$ | $\Delta\epsilon$ | $\epsilon$ | $\Delta\epsilon$ (stat) | $\Delta\epsilon$ (syst) | $\Delta\epsilon$ (total) | | |
| EB | 0.997 | 0.0001 | 0.989 | 0.002 | 0.001 | 0.0022 | 0.992 | 0.002 |
| EE | 0.988 | 0.0003 | 0.992 | 0.004 | 0.001 | 0.0041 | 1.004 | 0.004 |
| EB $e^-$ | 0.997 | 0.0001 | 0.986 | 0.004 | 0.001 | 0.0041 | 0.989 | 0.004 |
| EE $e^-$ | 0.988 | 0.0004 | 0.994 | 0.004 | 0.001 | 0.0041 | 1.006 | 0.004 |
| EB $e^+$ | 0.996 | 0.0001 | 0.994 | 0.003 | 0.001 | 0.0032 | 0.998 | 0.003 |
| EE $e^+$ | 0.989 | 0.0004 | 0.990 | 0.005 | 0.001 | 0.0051 | 1.001 | 0.005 |

# References

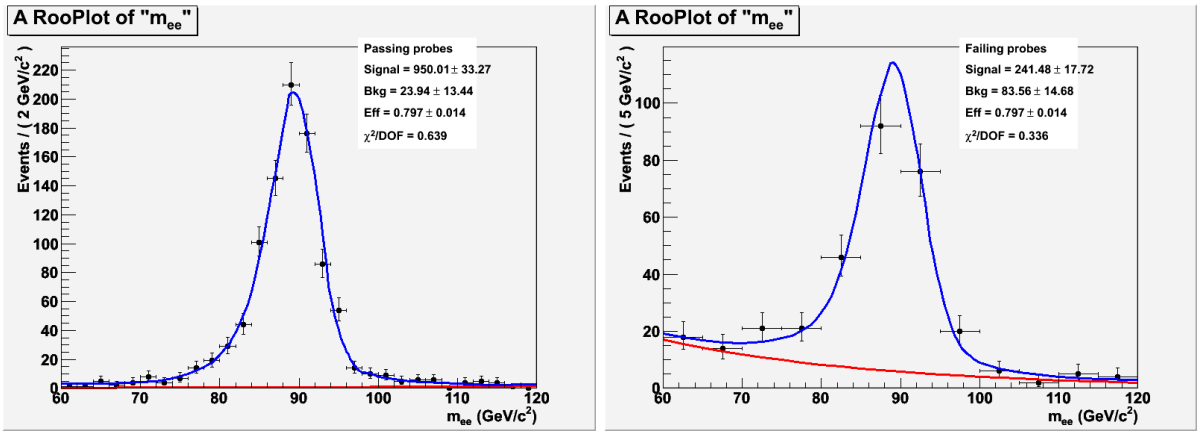

[1] *Generic Tag and Probe Tool for Measuring Efficiency at CMS with Early Data*, CMS AN-2009/111.

[2] *Updated Measurements of Inclusive W and Z Cross Sections at 7 TeV*, CMS AN-2010/264.

[3] *A New Method for Combining NLO QCD with Shower Monte Carlo Algorithms*, P. Nason, JHEP 0411, 040 (2004) [arXiv:hep-ph/0409146].

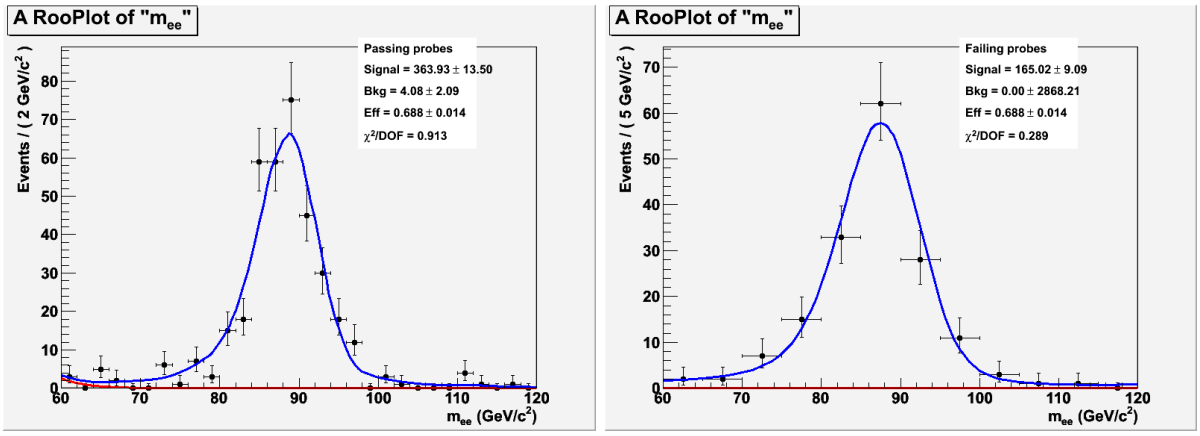
[4] D.Drijard, Statistical Methods in Experimental Physics, North-Holland, 1971.

# A Simulation Based Signal Shape Fit Plots



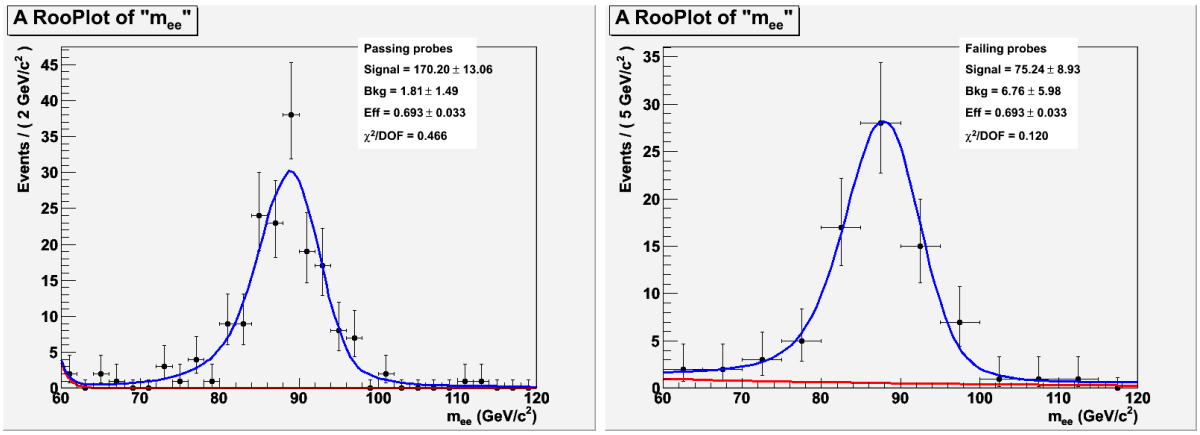

(a) (b)

Figure 5: The passing (a) and failing (b) fits for the Reco→WP80 step in the Ecal Barrel. The data is in black, background fit in red, and signal+background fit in blue.



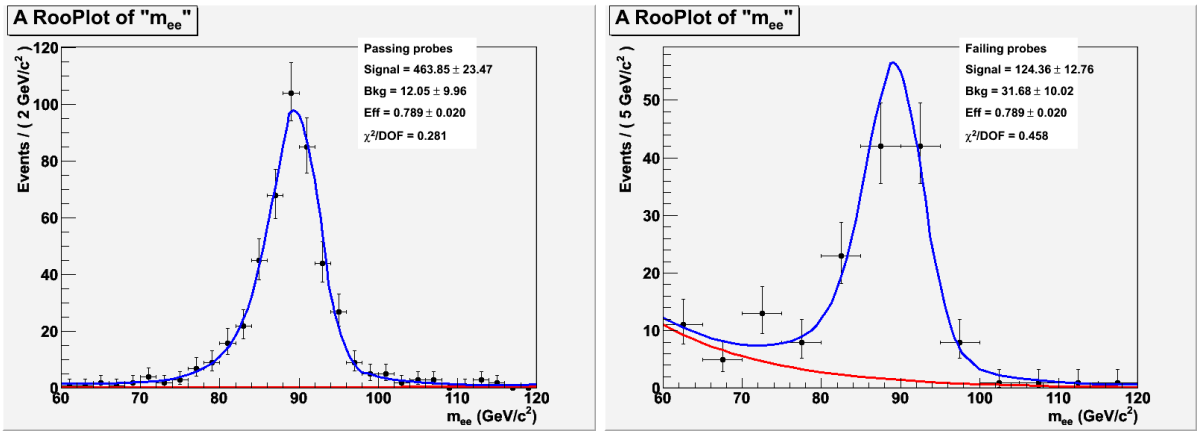

(a) (b)

Figure 6: The passing (a) and failing (b) fits for the Reco→WP80 step in the Ecal Endcap. The data is in black, background fit in red, and signal+background fit in blue.



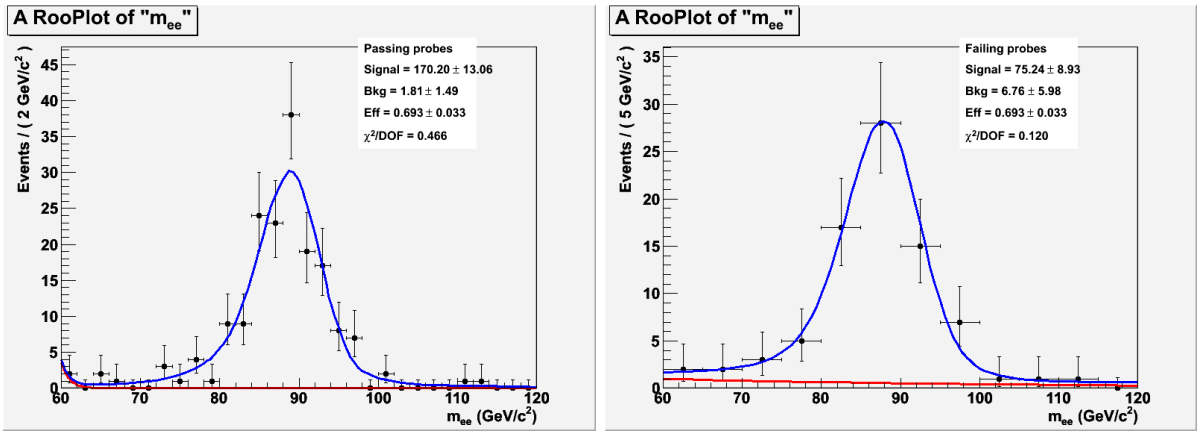

(a) (b)

Figure 7: The passing (a) and failing (b) fits for the Reco→WP80 step in the Ecal Barrel for electrons. The data is in black, background fit in red, and signal+background fit in blue.



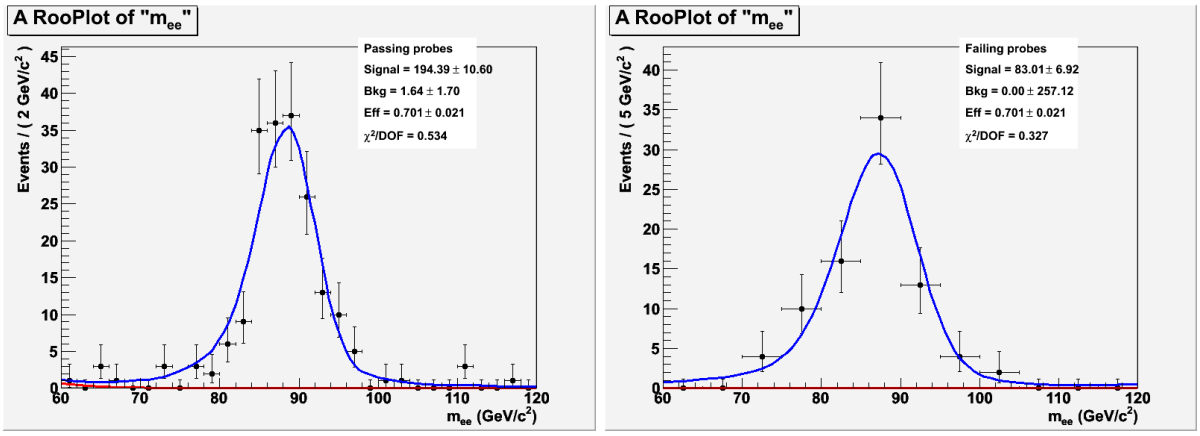

(a) (b)

Figure 8: The passing (a) and failing (b) fits for the Reco→WP80 step in the Ecal Barrel for positrons. The data is in black, background fit in red, and signal+background fit in blue.




(a) (b)

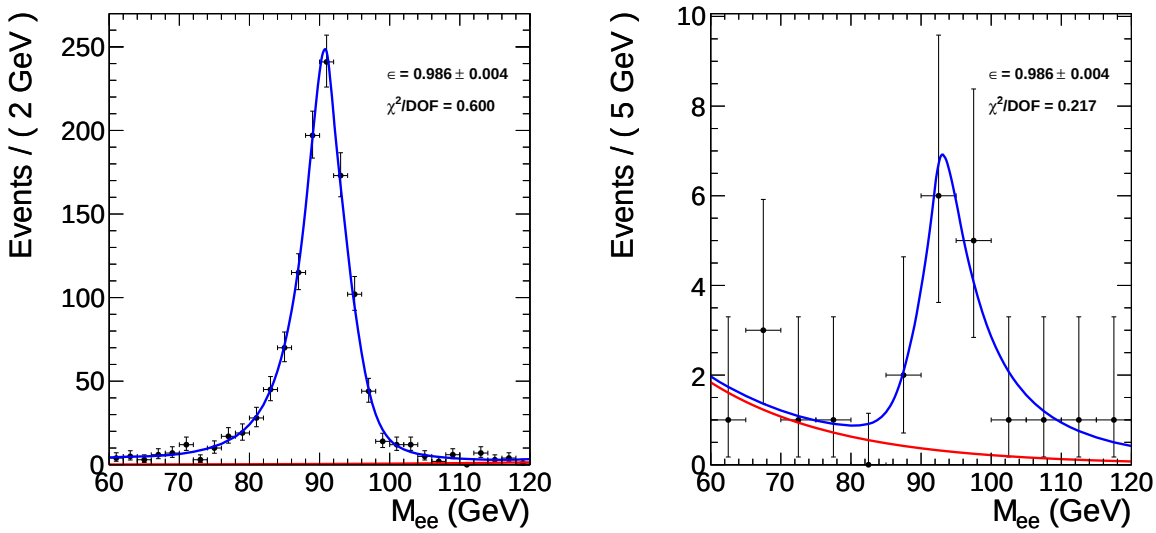
Figure 9: The passing (a) and failing (b) fits for the Reco→WP80 step in the Ecal Endcap for electrons. The data is in black, background fit in red, and signal+background fit in blue.




(a) (b)

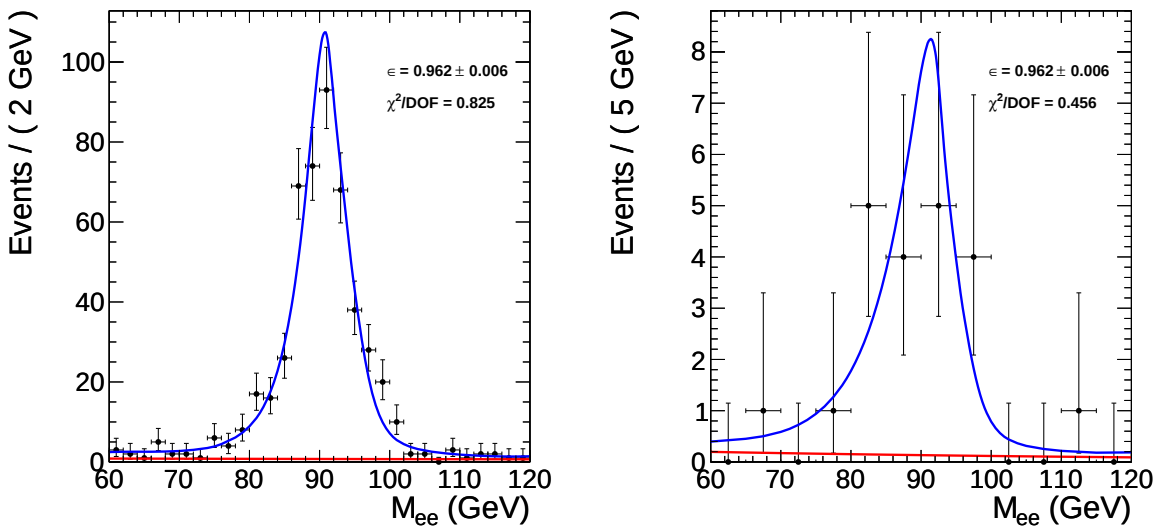
Figure 10: The passing (a) and failing (b) fits for the Reco→WP80 step in the Ecal Endcap for positrons. The data is in black, background fit in red, and signal+background fit in blue.

# B Generator-Level Signal Shape Fit Plots



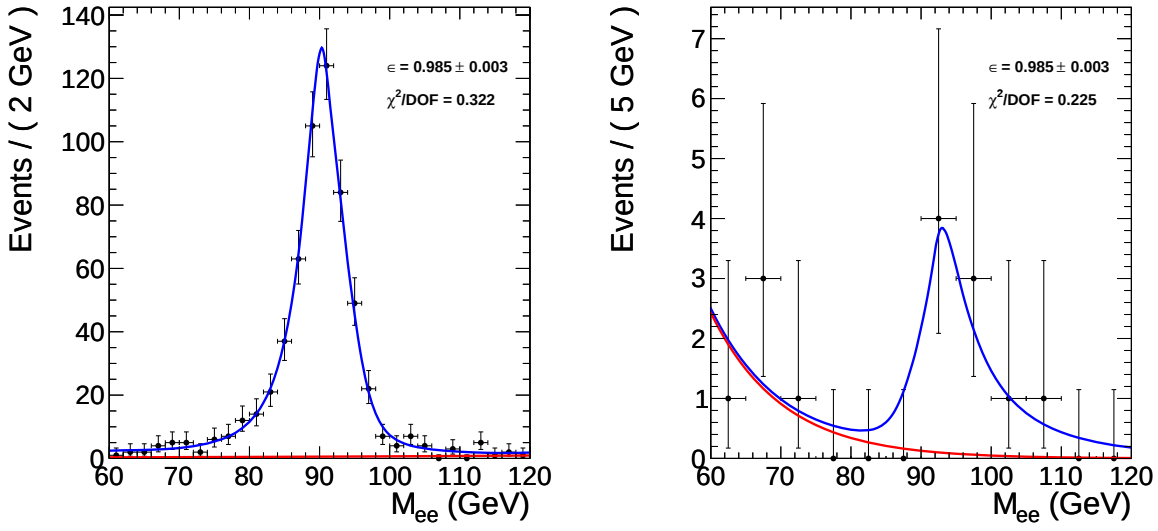

(a) (b)

Figure 11: The passing (a) and failing (b) fits for the SC→Reco step in the Ecal Barrel. The data is in black, background fit in red, and signal+background fit in blue.



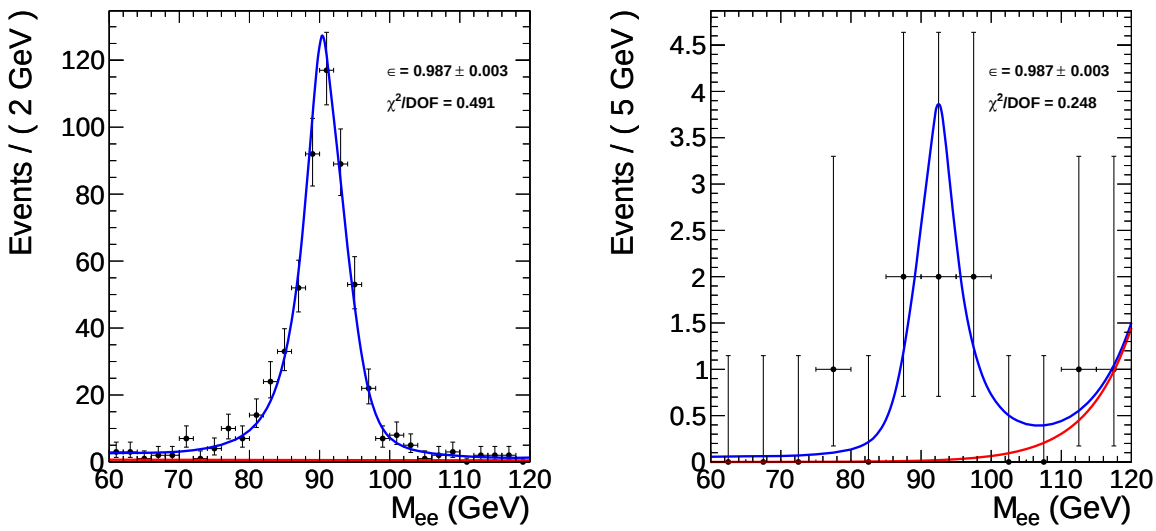

(a) (b)

Figure 12: The passing (a) and failing (b) fits for the SC→Reco step in the Ecal Endcap. The data is in black, background fit in red, and signal+background fit in blu e.



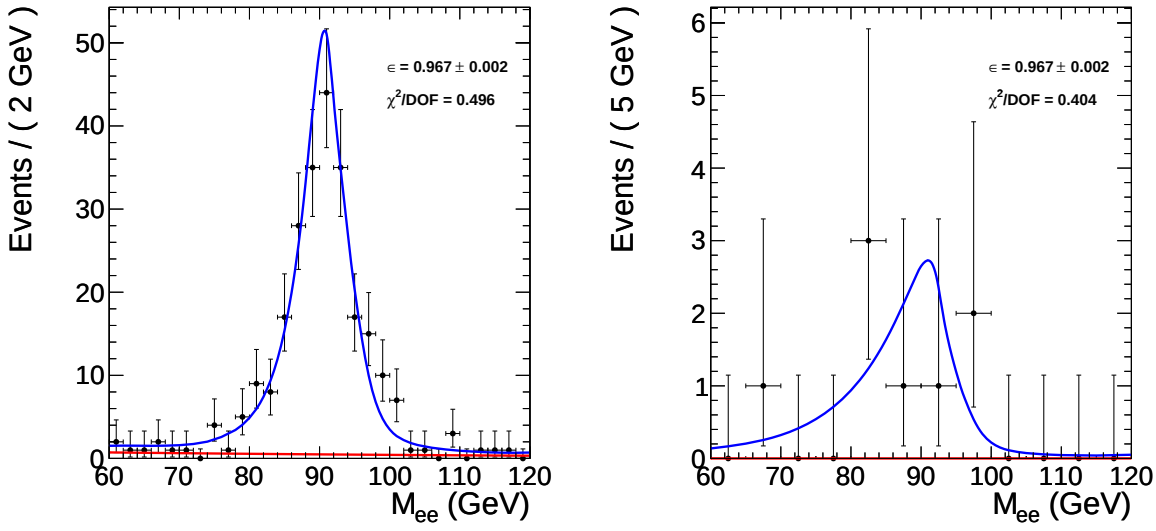

(a) (b)

Figure 13: The passing (a) and failing (b) fits for the SC→Reco step in the Ecal Barrel for electrons. The data is in black, background fit in red, and signal+background fit in blue.



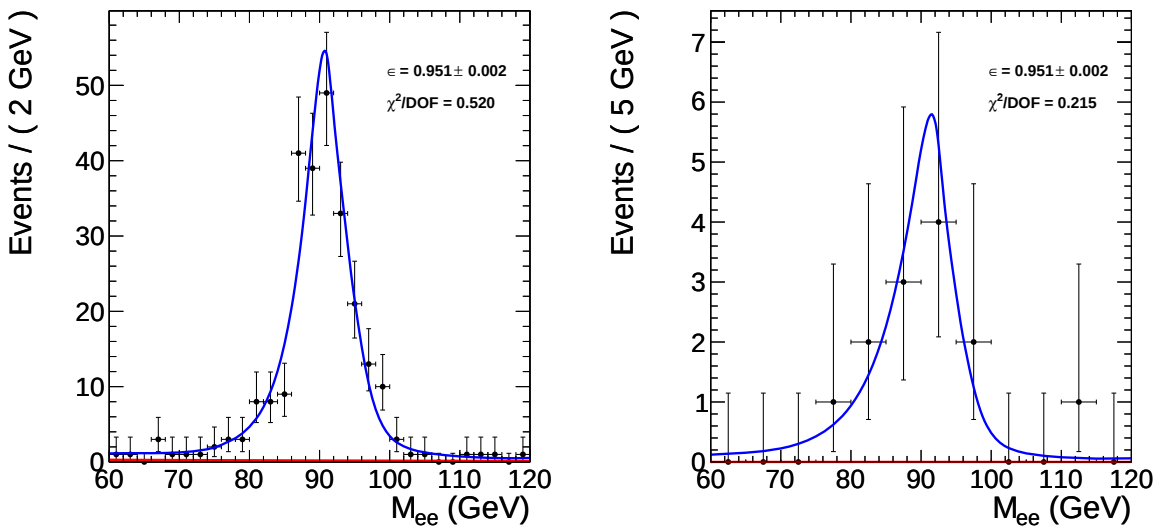

(a) (b)

Figure 14: The passing (a) and failing (b) fits for the SC→Reco step in the Ecal Barrel for positrons. The data is in black, background fit in red, and signal+background fit in blue.



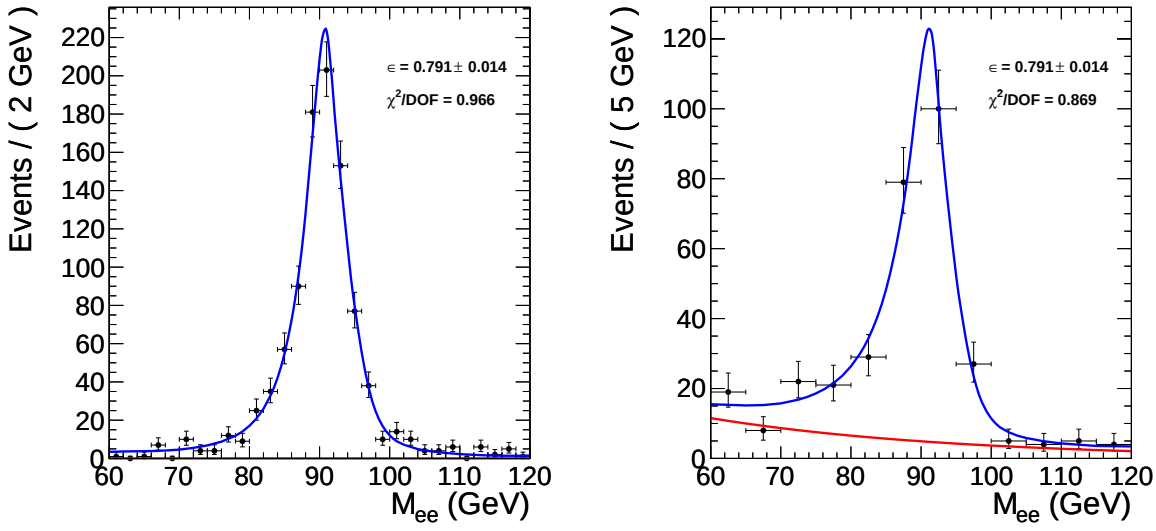

(a) (b)

Figure 15: The passing (a) and failing (b) fits for the SC→Reco step in the Ecal Endcap for electrons. The data is in black, background fit in red, and signal+background fit in blue.



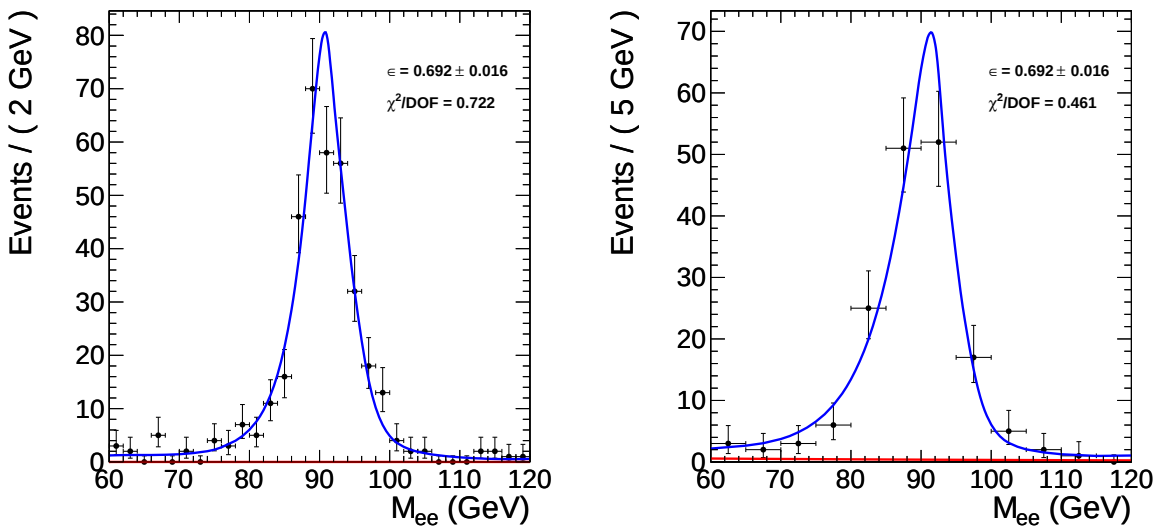

(a) (b)

Figure 16: The passing (a) and failing (b) fits for the SC→Reco step in the Ecal Endcap for positrons. The data is in black, background fit in red, and signal+background fit in blue.



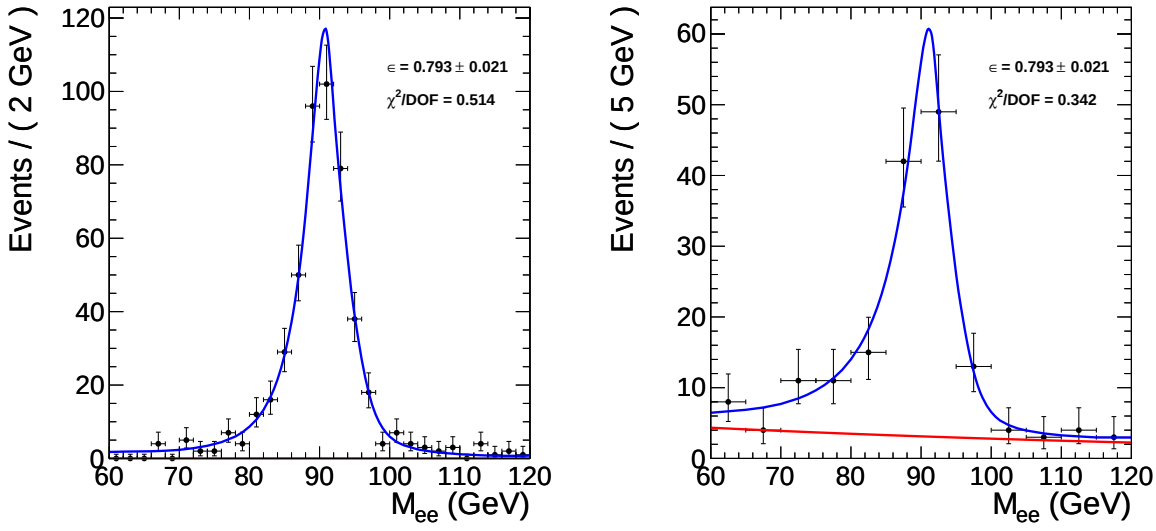

(a) (b)

Figure 17: The passing (a) and failing (b) fits for the Reco→WP80 step in the Ecal Barrel. The data is in black, background fit in red, and signal+background fit in blue.



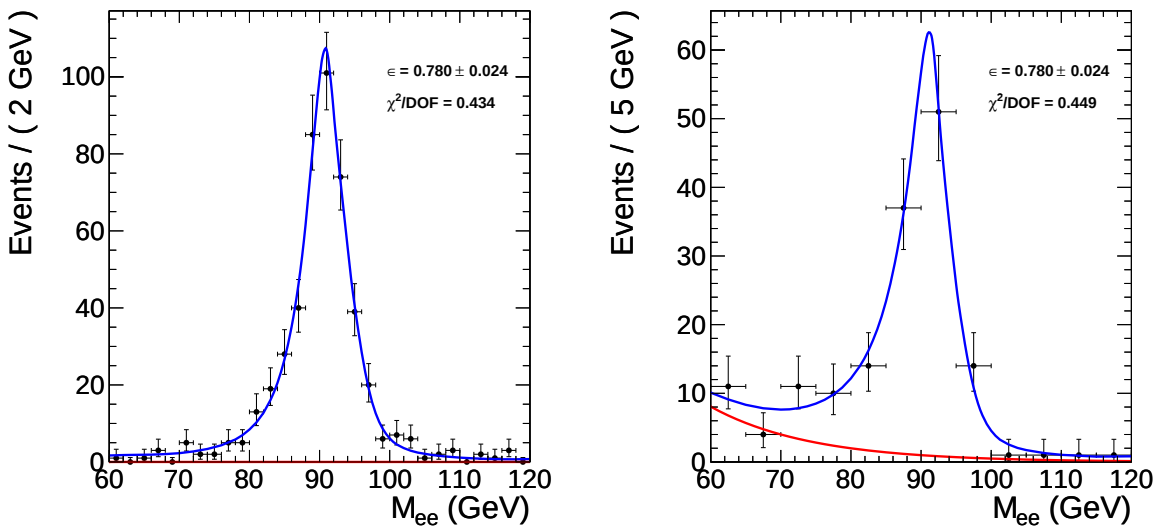

(a) (b)

Figure 18: The passing (a) and failing (b) fits for the Reco→WP80 step in the Ecal Endcap. The data is in black, background fit in red, and signal+background fit in blue.



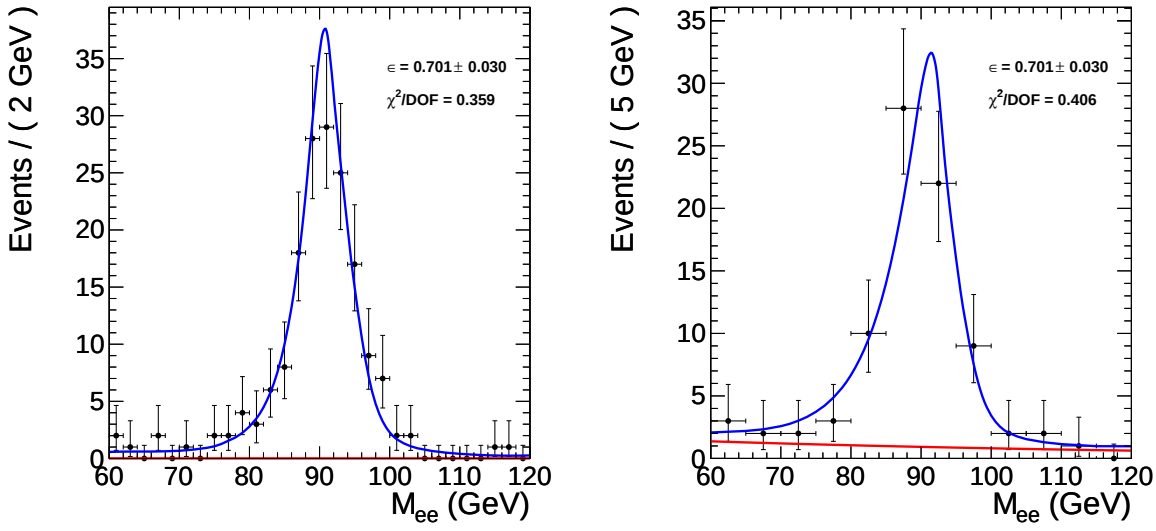

(a) (b)

Figure 19: The passing (a) and failing (b) fits for the Reco→WP80 step in the Ecal Barrel for electrons. The data is in black, background fit in red, and signal+background fit in blue.



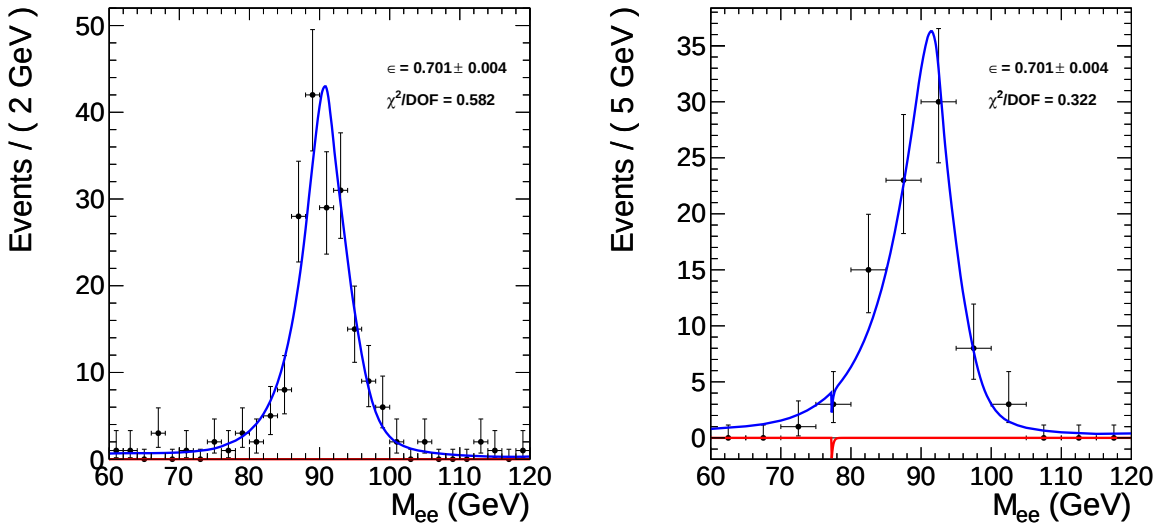

(a) (b)

Figure 20: The passing (a) and failing (b) fits for the Reco→WP80 step in the Ecal Barrel for positrons. The data is in black, background fit in red, and signal+background fit in blue.



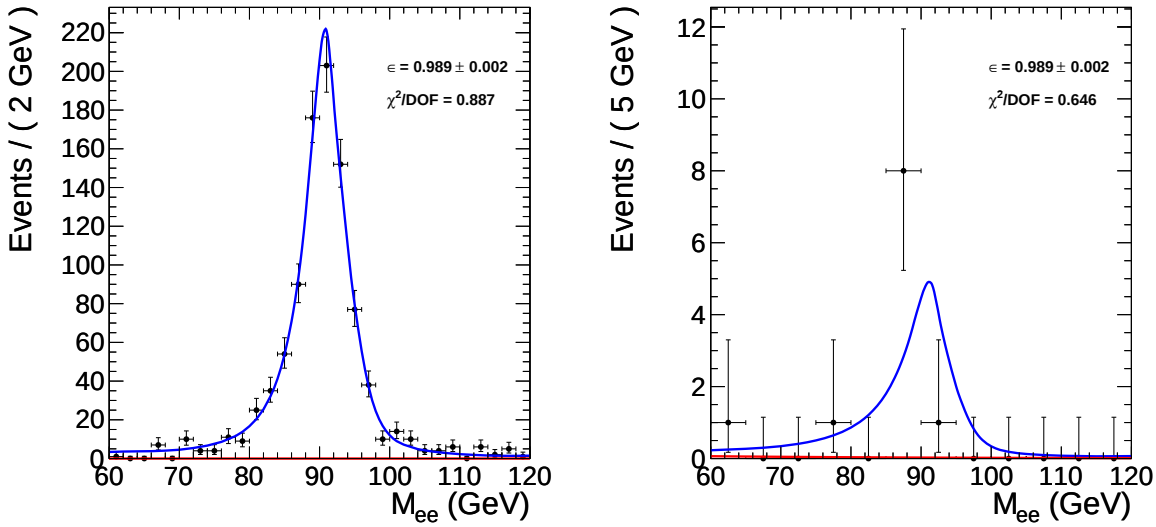

(a) (b)

Figure 21: The passing (a) and failing (b) fits for the Reco→WP80 step in the Ecal Endcap for electrons. The data is in black, background fit in red, and signal+background fit in blue.



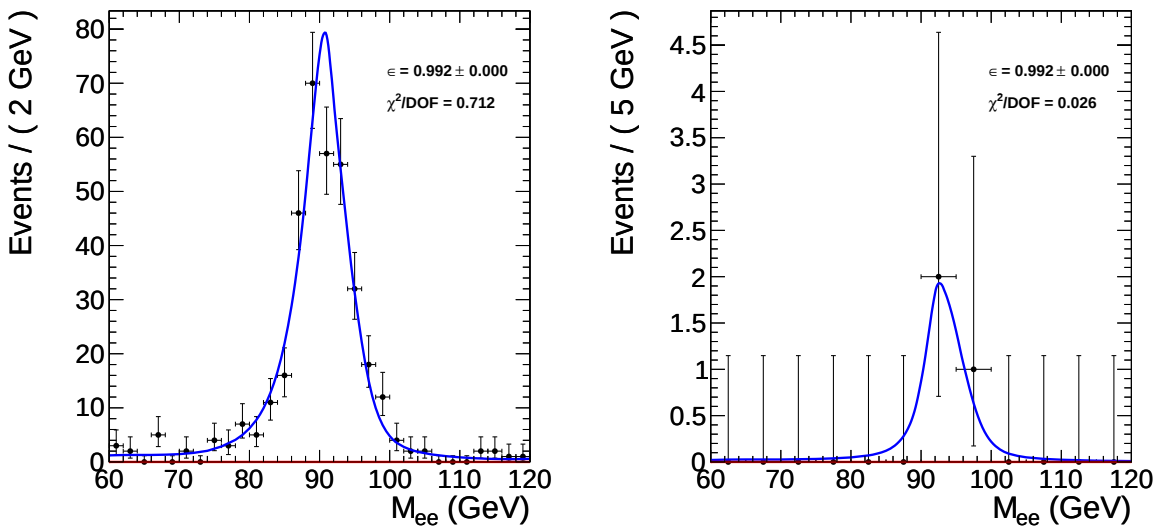

(a) (b)

Figure 22: The passing (a) and failing (b) fits for the Reco→WP80 step in the Ecal Endcap for positrons. The data is in black, background fit in red, and signal+background fit in blue.



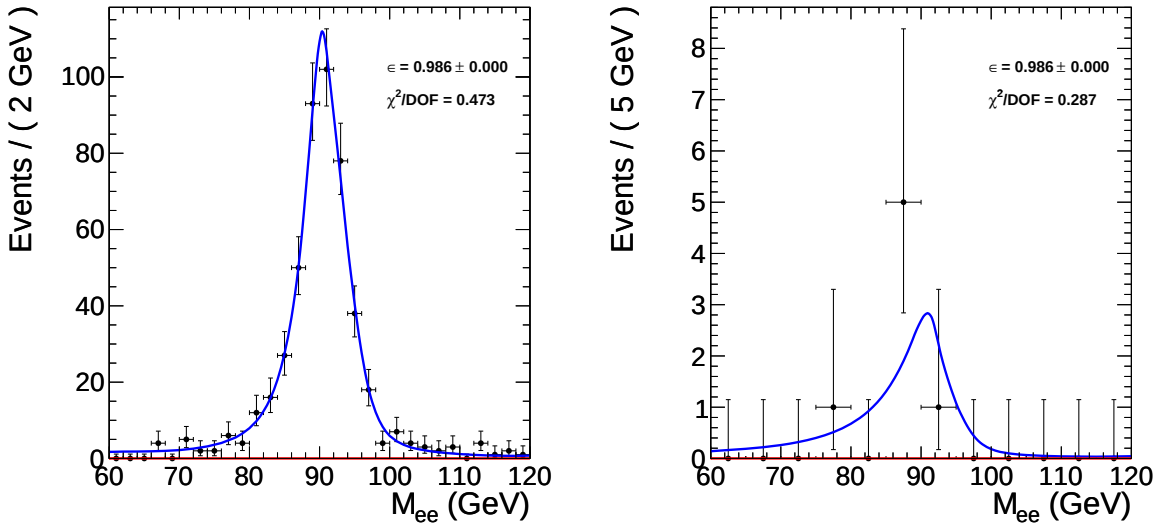

(a) (b)

Figure 23: The passing (a) and failing (b) fits for the WP80→HLT step in the Ecal Barrel. The data is in black, background fit in red, and signal+background fit in blue.



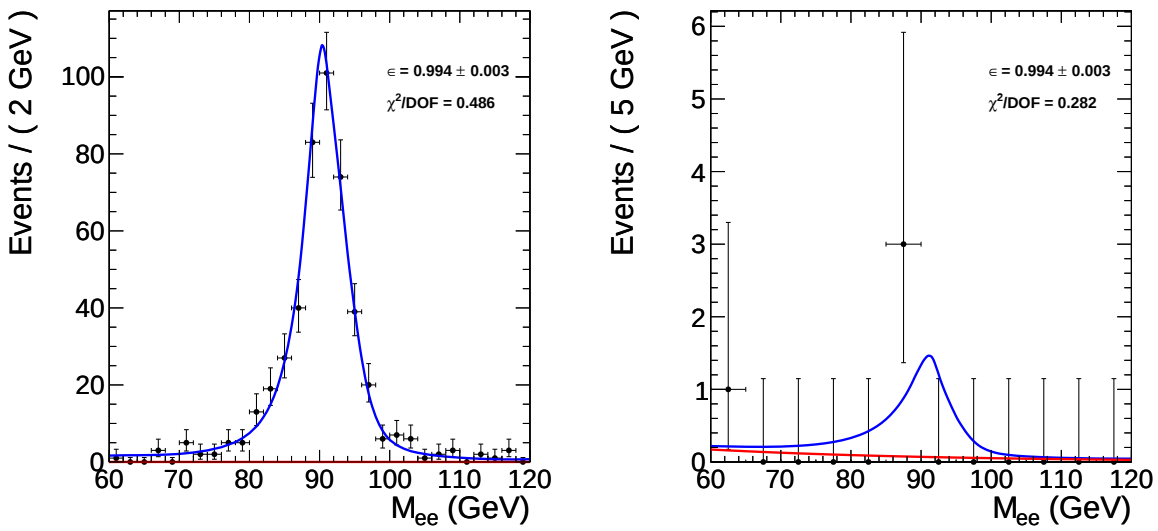

(a) (b)

Figure 24: The passing (a) and failing (b) fits for the WP80→HLT step in the Ecal Endcap. The data is in black, background fit in red, and signal+background fit in blue.



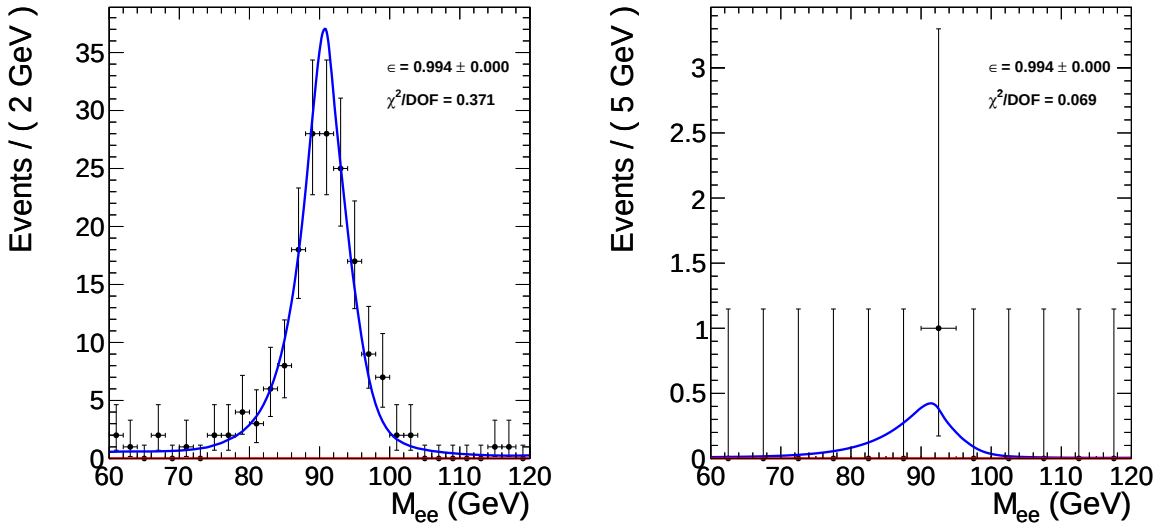

(a) (b)

Figure 25: The passing (a) and failing (b) fits for the WP80→HLT step in the Ecal Barrel for electrons. The data is in black, background fit in red, and signal+background fit in blue.



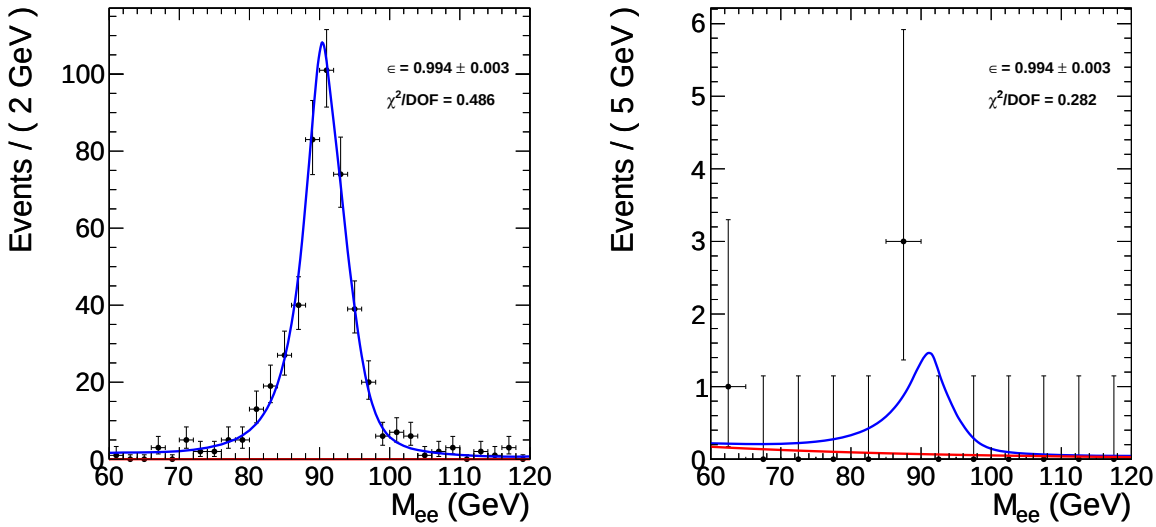

(a) (b)

Figure 26: The passing (a) and failing (b) fits for the WP80→HLT step in the Ecal Barrel for positrons. The data is in black, background fit in red, and signal+background fit in blue.




(a) (b)

Figure 27: The passing (a) and failing (b) fits for the WP80→HLT step in the Ecal Endcap for electrons. The data is in black, background fit in red, and signal+background fit in blue.




(a) (b)

Figure 28: The passing (a) and failing (b) fits for the WP80→HLT step in the Ecal Endcap for positrons. The data is in black, background fit in red, and signal+background fit in blue.