CDF/DOC/EXOTIC/CDFR/7379
Version 2.1
May 27, 2005

Low $\mathrm{E_T}$ Electron identification efficiency and scale-factors using J/Ψ

Sourabh Dube, John Zhou, Sunil Somalwar

Rutgers University

Abstract

We describe the measurement of electron ID efficiencies for low $\mathrm{E_T}$ electrons using J/Ψ electrons and the scale-factor between data and Monte Carlo. This scale-factor will be used for the SUSY trilepton analysis (low $\mathrm{E_T}$ electrons channel). We measure the isolation efficiency and scale-factor separately using Drell-Yan electrons. We combine the isolation and 'other ID' efficiency and scale factors.

1 Introduction

The electron ID scale-factor between data and Monte-Carlo (MC) is a requisite measurement for the susy trilepton analysis. The signal for the susy trilepton analysis consists of three leptons and $\not{E}_\mathrm{T}$. The lower we drop in electron $\mathrm{E_T}$ the more acceptance we gain for this analysis. Hence a low $\mathrm{E_T}$ efficiency study is vital to obtain a scale-factor between data and MC efficiencies. This scale-factor will be essential to estimate the expected number of signal and background events in data.

We use unbiased electrons from the susy dilepton dataset (edil0d). The method of obtaining the unbiased sample is decribed in section 2. We measure the ID efficiency and isolation efficiency separately using J/Ψ events for the other variables and Drell Yan events for the isolation study as described in sections 4 and 6.

It should be noted that there is an implicit cut $E^{HAD}/E^{EM} \leq 0.125$ while reconstructing a CDF EM object. Thus the efficiency we measure is over and above this E^{HAD}/E^{EM} cut.

In section 4 we describe the process of selecting J/Ψ electrons and measuring efficiencies for the calorimetry and showermax variables using the J/Ψ electrons. In the following section we describe the systematic studies.In section 6, we describe the selection of Drell Yan events in our data and the method for measuring isolation efficiency. The final section has the conclusions and results.

2 SAMPLE

2.1 DATA samples

We use J/Ψ events to measure the electron ID efficiency. To measure the efficiency we require a sample of unbiased electrons. For this we use data sample stripped from the susy dilepton dataset - edil0d. We select events triggered on a single electron by requiring the Level 3 paths for the event to include this path:

- ELECTRON_CENTRAL_4 (This trigger is prescaled by 500)

The method to remove trigger bias due to this selection is described in following subsection.

We use Drell-Yan events selected from data to measure the efficiency of our isolation cut. It is relatively easy to obtain an unbiased sample for this measurement. We use data from the susy dilepton dataset. We require events triggered on this path:

- DIELECTRON_CENTRAL_4

As explained in section 6, the selection of electrons is tighter than the trigger cuts. The trigger also has no requirement on electron isolation. Hence this forms an unbiased sample of electrons to measure the isolation efficiency.

2.1.1 Removing trigger bias for measuring ID efficiency

As mentioned above, we use single-electron triggered data to measure the electron ID efficiency in data. Hence removing trigger bias is an important step as this could lead to artificially high efficiencies for data electrons and an incorrect scale-factor. To remove trigger bias we implement the following steps:

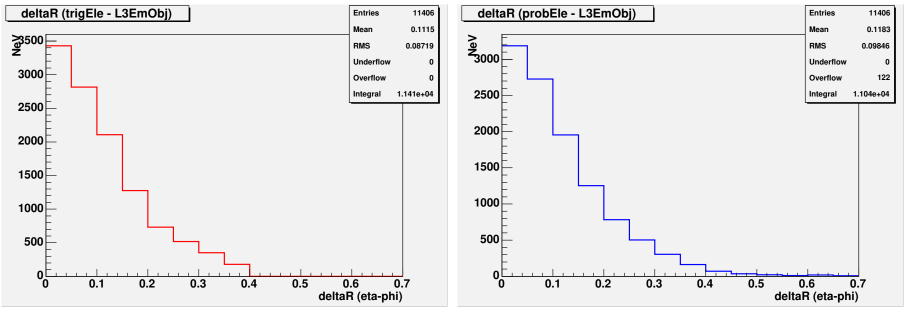


Figure 1: Figure shows the $\triangle R_{\eta\phi}$ between trigger/probe electron and closest L3 object. We require this to be a maximum of 0.4, but it is typically 0.1 or less.

- For each electron in an event, we find a matching Level 3 EmObject. We require $\triangle R_{\eta\phi}$ between electron and matched L3 object to be at most 0.4, but this $\triangle R$ is typically 0.1 or less (Refer Fig 1). We pick the closest L3 object to the electron.

- We then require the matched L3 object to pass the trigger requirements. A match also categorizes the electron under consideration as a 'trigger' electron.

- We require each electron in this event to match to a distinct L3 object. In case we don't find a matched trigger object for any electron in our event we discard this event.

- The actual scheme for selecting an unbiased probe electron is that we consider an electron as a probe electron if and only if its partner (forming J/Ψ mass) is a trigger electron.

It should be noted that for J/Ψ electrons with $E_T \geq 4$ GeV, the two electrons are potentially close in $\triangle R_{\eta\phi}$. Due to this there is the possibility of mismatch in finding the Level 3 EmObj, and also biasing the efficiency due to close electrons. To avoid these effects we require the electrons to be separated by $\triangle R_{\eta\phi} \geq$0.4.

2.2 MONTE-CARLO samples

We use two Monte Carlo samples generated by different generators to assess the scale-factors for electron ID efficiency.

- We use the $b\bar{b}/c\bar{c}$ MC sample generated using Isajet 7.51 and v5.3.3 of CDF software. It has about 600K events .The process for obtaining $b\bar{b}/c\bar{c}$ is gluon splitting. The luminosity for this sample is 8058pb^{-1}.

- We also use the $b\bar{b}$ MC sample generated using Pythia 6.2 and v5.3.3 of CDF software. This sample has about 2M dilepton-filtered events and corresponds to a luminosity of 601pb^{-1}.

We use the following Monte Carlo sample to assess scale-factor for the electron isolation cut efficiency.

- Drell Yan MC sample was generated using Pythia 6.2 and v5.3.3 of CDF software. It has about 700K events. The luminosity for this sample is 1052.6pb^{-1} using a cross section k-factor of 1.34.

Note: The data sample that we use to measure the electron ID efficiency has a $P_T \geq$ 4GeV cut on trigger electron at level 1 and level 2. This is a kinematic effect and needs to be accounted for in MC as well. For MC events, we require that one electron of the pair has $P_T \geq$ 4GeV and then we consider the other as probe.

Refer webpage [1] and analysis note [2] for more details about MC samples.

3 Method

We measure the electron ID efficiency/scale-factor and isolation efficiency/scale-factor separately and then multiply the two to obtain the total scale-factor. We do this by measuring the ID and isolation cut efficiency as 'all-but-one' efficiencies - for the ID efficiency we apply our isolation cut while forming our denominator and then apply ID to obtain the numerator. Thus the ID efficiency we measure is with respect to the isolation.

We repeat the same to measure isolation efficiency viz. we apply ID variables and obtain denominator and then apply isolation and obtain numerator and thus the efficiency.

We assign a systematic error due to this process of measuring the ID and isolation efficiency separately. Refer Sec 5.

4 EFFICIENCY USING J/Ψ

The ID variables for which we will measure efficiency are described in Table 1. The process of selecting good J/Ψ events is described below.

Base cuts
$E_T \geq 4$GeV
trk $\lvert Z_0 \rvert \leq 60cm$
CES Fiducial Ele

ID variables
EoverP ≤ 2.0
HadEm $\leq (0.055 + 0.00045 \times E_T)$
-3.0 cm$\leq$ CES Signed $\triangle$x $\leq$1.5 cm
CES $\lvert \triangle z \rvert \leq 3$ cm
CES $\chi^2_{strip} \leq 10$
Lshr ≤ 0.2

Table 1: The left table shows the baseline cuts on probe leg. The right table shows the ID variables for which we measure efficiency.

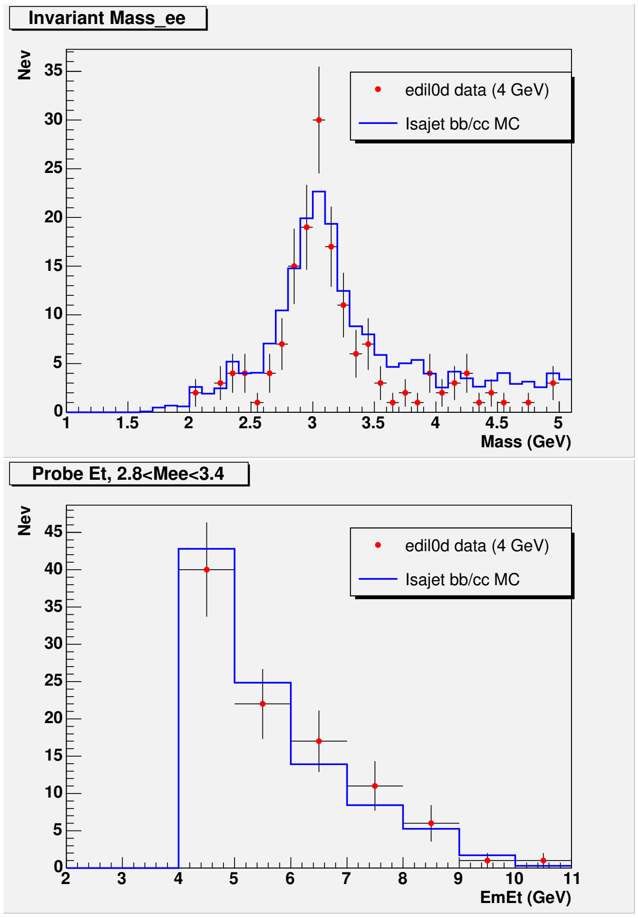


Figure 2: Figure shows the J/Ψ mass in data and MC and the E_T distribution of probe electrons. The mass distributions have been normalized to the number of events from 2.8 to 3.4 GeV. The E_T distributions is normalized to the total number of events.

We remove trigger bias by flagging one leg as trigger leg and other unbiased leg as probe by method described in previous sections. We require both legs to pass the base cuts listed in Table 1. We then apply all ID cuts to the trigger leg. We then apply isolation cut $E_T^{iso}/E_T \leq 0.1$ to the probe leg. We form the invariant mass of the two legs and select J/Ψ events. Refer Fig 2. The

E_T distribution for the probe electrons (for events where $2.8 \leq M_{ee}(GeV) \leq 3.4$) is shown in same plot.

We remove conversions in our data sample by requiring that the two tracks associated with the electrons have separation $S_{xy} \geq 0.25$cm and $\triangle cot(\theta) \geq 0.15$.

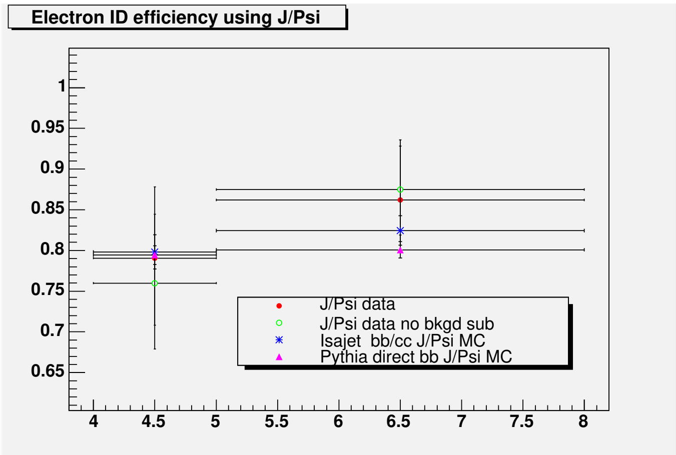


Figure 3: Figure shows the efficiency for variables in Table 1 using the J/Ψ electrons. We show the data efficiency with and without background subtraction. The effect of background subtraction is small. For comparison we also plot the efficiency using the Isajet and Pythia MC samples described in Sec 2.2. The ratio of data efficiency with background subtraction to MC efficiency gives us the scale-factor.

We measure efficiency in two bins of E_T : $4 \leq E_T(\mathrm{GeV}) \leq 5$ and $5 \leq E_T(\mathrm{GeV}) \leq 8$. The results along with isolation efficiency, systematic errors and scale-factors are descibed in Section 7.

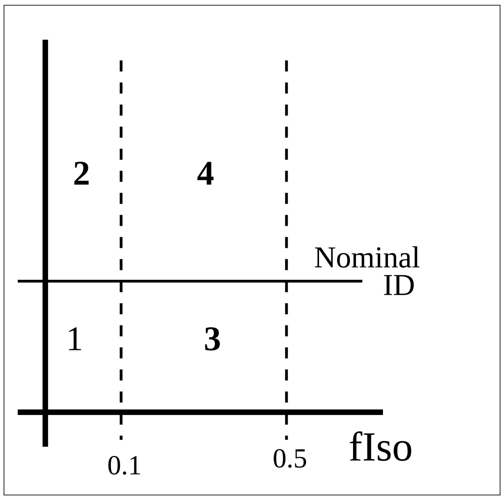
5 SYSTEMATICS

We performed various systematic checks for the measurement of efficiency. These are described below. We define the nominal efficiency as the one with background subtraction, and tight requirements on the tag leg.

- The first systematic check is measuring the ID efficiency using J/Ψ events with and without background subtraction. The shape of the background is determined mostly by the isolation cut on the probe electron. Refer Table 2 for the results. The difference in efficiency due to this is seen to be less than the statistical errors on the efficiency. As a conservative approach, we take the systematic error due to this effect to be the same as statistical error (refer Table 7).

Selection E_T bin	4 - 5 GeV	5 - 8 GeV
Nominal	$79.1^{+8.8}_{-8.3}\%$	$86.2^{+6.6}_{-6.1}\%$
Without bkgd sub	$76.0^{+8.5}_{-8.1}\%$	$87.5^{+6.1}_{-5.6}\%$
loose tag leg	$75.5^{+8.6}_{-8.2}\%$	$88.8^{+5.5}_{-5.0}\%$

Table 2: Efficiencies with different selection requirements on electrons. The errors on the efficiencies are statistical.



Compare

$$\frac{Nev{:}region1}{Nev{:}region1+Nev{:}region2} \sim \frac{Nev{:}region3}{Nev{:}region3+Nev{:}region4}$$

DATA

$0.846\pm0.038 \sim 0.694\pm0.034$

MC

$0.778\pm0.012 \sim 0.696\pm0.009$

- We measure the electron ID and isolation efficiencies separately. Since these cuts will be applied together for the analysis, we performed a systematic study as shown in the figure above. We compared the ID efficiencies of electrons with $\mathrm{E_T^{iso}/E_T} \leq 0.1$ and electrons with $0.1 \leq \mathrm{E_T^{iso}/E_T} \leq 0.5$ for both data and MC. Thus this is a comparison of ID efficiencies for isolated and 'not-so-isolated' electrons.

  Consider the situation where the other ID variables and the isolation variable for an electron are completely uncorrelated. In this case, the two numbers $\frac{1}{1+2}$ and $\frac{3}{3+4}$ would be equal. However there exists some correlation between other ID variables and isolation. Let us call this effect as e_{iso}. Now if e_{iso} was modelled completely accurately in data and MC, then the comparison for two isolation regions would be same in data and MC.

  To obtain the systematic error, we get the scale-factor of Data/MC for isolated electrons from above numbers $= 0.846/0.778 = 1.087$. This number is 1.000 for relatively non-isolated electrons. The difference is 0.087. The nominal scale-factor for first bin with $4 \leq \mathrm{E_T(GeV)} \leq 5$ is 0.9906 . Hence the error on scale-factor for first $\mathrm{E_T}$ bin is
  $(0.087/1.087) \times 0.9906 = 0.0796$.
  The nominal scale factor for second bin with $5 \leq \mathrm{E_T(GeV)} \leq 8$ is 1.0456. Hence the error on scale-factor for second $\mathrm{E_T}$ bin is
  $(0.087/1.087) \times 1.0456 = 0.0840$. We document these error in Table 8.

- As the two electrons are close in $\triangle R$, we also study the effects of applying tight/loose cuts on the tag electrons and how they affect the probe electron efficiency. The difference in efficiency is seen to be less than the statistical errors. As above we inflate this error and take it to be equal to the statistical error (refer Table 7).

6 EFFICIENCY USING Drell Yan

The isolation efficiency for the electrons is measured using Drell Yan (DY) events in our data sample. We use events triggered by the dielectron trigger. This trigger has no isolation requirements, and our offline selection of electrons is tighter than the trigger.

We select Drell Yan events in our data sample in the following way.

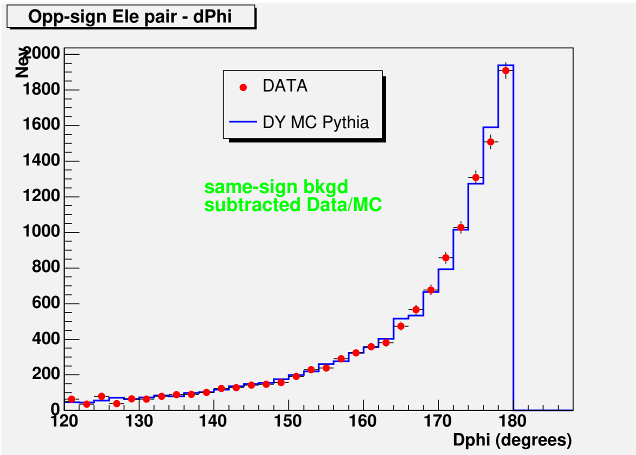


Figure 4: The plot shows the $\triangle\phi$ distribution between pair of electrons for all events after selection cuts have been applied on both legs (ID cuts, $M_{ee} \geq 15GeV$.) We subtract same-sign events to account for the background. We select events with $\triangle\phi \geq 160$ degrees. The distributions have been normalized to the number of events with $\triangle\phi \geq 160$ degrees.

Requirements for Drell Yan events
All cuts in Table 1 on tag leg.
$E_T^{iso}/\mathrm{E_T} \leq 0.1$ on tag leg
All cuts in Table 1 on probe leg.
$\triangle\phi \geq 160^o$
$M_{ee} \geq 15GeV/c^2$
$\triangle R_{\eta\phi} \geq 0.4$ between electron and any jet^*.

Table 3: Selection cuts for DY events. jet^*: We use jets with EmFrac$\leq$0.9 and level 5 corrected $jet\mathrm{E_T} \geq 15$GeV

We select events with two electrons with $\triangle\phi$ between them greater than 160 degrees. We also apply an cut on invariant mass of the two electrons: $M_{ee} \geq 15GeV$. We require both electrons to pass the ID cuts described in Table 1. We then designate one leg as tag leg, one leg as probe and require $E_T^{iso}/\mathrm{E_T} \leq 0.1$ on the tag leg. We require the probe leg to be away from any jets to reduce background. We do this by applying $\Delta R \geq 0.4$ cut between probe electron and any jet with EmFraction$\leq$0.9 and level5 corr$\mathrm{E_T} \geq 15$GeV. This forms our denominator.

We then apply the isolation cut $E_T^{iso}/\mathrm{E_T} \leq 0.1$ on the probe leg, this forms the numerator. We repeat this procedure by switching the tag and probe legs. We measure the efficiency of our isolation cut in bins of 2 GeV in electron $\mathrm{E_T}$, from 4 GeV upto 60 GeV.

Fig 4 shows the $\triangle\phi$ distribution after all cuts on both legs and invariant mass cut have been applied. Fig 5 shows the $\mathrm{E_T}$ distribution of probe leg after all cuts (including $\triangle\phi$) have been applied.

We estimate background from the number of same-sign events and subtract these from the opposite-sign events. We would like to note that this does not take into account all possible backgrounds. The shape of the $\mathrm{E_T}$ distribution in low-$\mathrm{E_T}$ region (Fig 5) suggests a slight excess of background in this region. The shape agrees well with a possible heavy-flavor contribution, but

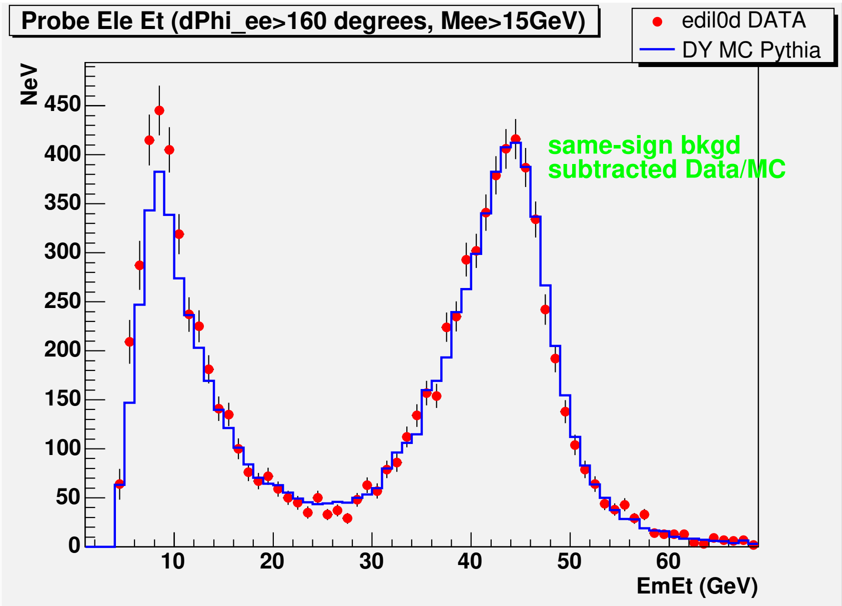


Figure 5: The plot shows the E_T distribution of probe leg after all selection cuts have been applied on both legs (ID cuts, $\triangle\phi \geq 160$ degrees, $M_{ee} \geq 15GeV$.) We subtract same-sign events to account for the background. The distributions have been normalized to the number of events with 30$\leq$ E_T(GeV) $\leq$60.

the heavy-flavor contribution for the same luminosity and event selection is negligible and does not account for this background. This excess background needs further investigation and further updates of this note will include the same.

The measured isolation efficiency is shown in Figure 6.

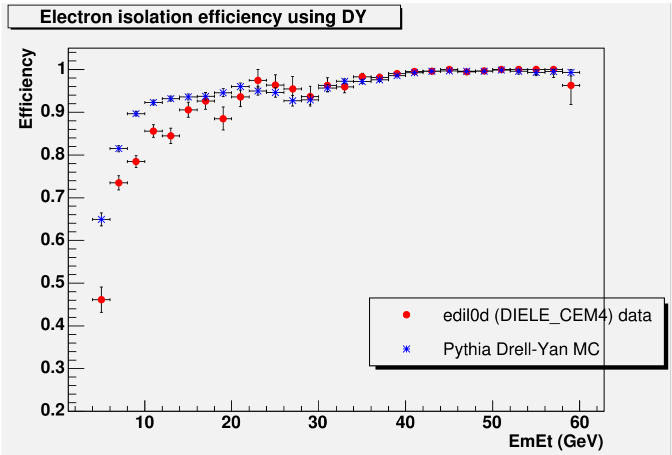


Figure 6: Figure shows the efficiency for isolation cut ($E_T^{iso}/E_T \leq 0.1$) for Drell Yan data and MC electrons.The ratio of data to MC gives us the scale-factor.

7 CONCLUSION and RESULT

Table 4 shows the ID efficiency for data and the ID scale-factors. The systematic errors on the efficiency include the 'loose-trigger' and 'no-background' effects describes in Section 5. The systematic error on the scale-factor includes the systematic error on efficiency and also the isolation-study error described in same section. Table 5 shows the isolation efficiency for the same E_T bins as ID efficiency and the scale-factors. The errors on isolation efficiency are statistical and the errors on isolation scalefactor are obtained from the fit parameters.

E_T bin (GeV)	ID efficiency	Scale factor
4 - 5	$79.1 \pm 8.8(stat) \pm 12.4(sys)\%$	$0.99 \pm 0.11(stat) \pm 0.15(sys)\%$
5 - 8	$86.2 \pm 6.6(stat) \pm 9.4(sys)\%$	$1.05 \pm 0.08(stat) \pm 0.12(sys)\%$

Table 4: ID efficiencies for Data and ID scalefactors. The systematic errors on efficiency include errors from 'loose-trigger' and 'no-background' effects. The systematic error on scale-factor includes these errors and also the isolation study error.

E_T bin (GeV)	Isolation efficiency	Scale factor
4 - 5	$51.6 \pm 6.1(stat)\%$	0.78 ± 0.06
5 - 8	$66.9 \pm 1.6(stat)\%$	0.84 ± 0.05

Table 5: Isolation efficiencies for Data and scalefactors. The errors on the efficiency are statistical, the errors on scale factor are from the fit parameters (Fig 7).

Fig 7 shows the isolation scale-factor as a function of E_T. The fit is
Scale factor $= p_2 - p_1 \cdot e^{-p_0 \cdot E_T}$
We evaluate the scale-factor for the same E_T bins as ID, viz. 4-5 and 5-8 GeV. We use the central value of these bins (4.5 GeV and 6.5 GeV) to obtain the isolation scale-factor for these bins. The results for ID scale-factor, isolation scale-factor and total scale-factor are shown in Table 6. The ID scale-factors have systematic errors from all studies in Section 5. The isolation scale-factors has statistical errors obtained from errors on the fit parameters (Fig 7). The total scale-factor has errors which are all errors combined.

E_T bin (GeV)	ID scale-factor	Isolation scale factor	Total scale-factor
4 - 5	$0.99 \pm 0.11(stat) \pm 0.15(sys)$	$0.78 \pm 0.06(stat)$	0.77 ± 0.19
5 - 8	$1.05 \pm 0.08(stat) \pm 0.12(sys)$	$0.84 \pm 0.05(stat)$	0.88 ± 0.16

Table 6: Scale-factors for electron ID, isolation and total scale-factor. The ID scale-factor has systematic errors from all systematic studies. The isolation scale-factor have statistical errors from the fit parameters (Fig 7) and the total scale-factor error includes all errors.

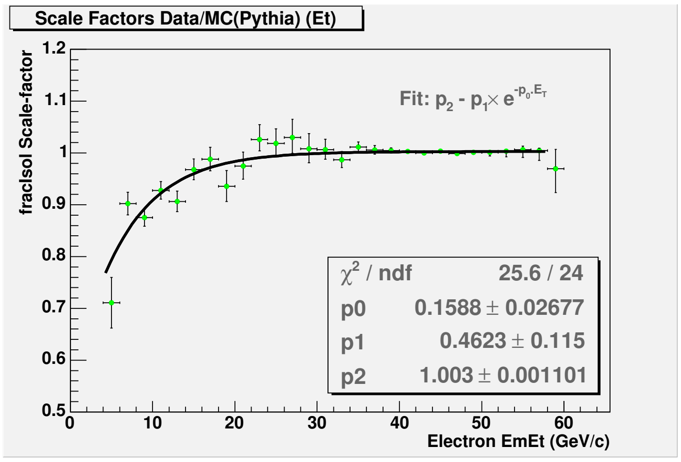


Figure 7: Figure shows the scalefactor (Data/MC) for isolation cut ($E_T^{iso}/\mathrm{E_T} \leq 0.1$) for Drell Yan data and MC electrons. We fit an exponential function with asymptotic value = p_2. Thus the scale factor as a function of $\mathrm{E_T}$ is given by ScaleFactor= $p_2 - p_1 \cdot e^{-p_0 \cdot \mathrm{E_T}}$ with p_0,p_1,p_2 as shown on plot.

$\mathrm{E_T}$ GeV	4 - 5	5 - 8
'loose-trigger'	8.8%	6.6%
'no-background'	8.8%	6.6%

Table 7: Sources of systematic errors on electron ID efficiency. Refer Sec 5

In conclusion we have measured the ID efficiency of electrons using J/Ψ electrons and Drell-Yan electrons. We measure the ID efficiency/scale-factor using J/Ψ electrons and isolation efficiency/scale-factor using DY electrons. We assign a systematic error based on a study we performed due to this separate measurement.

The statistical errors on the ID efficiency measurement are our single highest error, and further studies for low $\mathrm{E_T}$ electrons' ID need to be done with larger data samples.

8 Acknowledgements

We would like to thank the SUSY group convenors Peter Wittich, Teruki Kamon, the Exotic group convenors Beate Heinemann, Song Ming Wang for their comments and suggestions. We would like to thank Alexei Safonov for his valuable insights and effort in reviewing this work.

$\mathrm{E_T}$ GeV	4 - 5	5 - 8
'Isolation study'	0.080	0.084

Table 8: Sources of systematic errors on electron ID scale-factor. Refer Sec 5

References

[1] SUSY-Multilepton group. www-cdf.fnal.gov/internal/physics/exotic/susy/r2a/multilepton/.

[2] John Zhou et al. Searches for Chargino-Neutralino Production in mSUGRA Model in a Dielectron + Track Channel. CDF Note 7478.

A Results for eID with Pythia MC

We measured the scale-factor for eID using the Pythia MC sample as well. Results are summarized in Table 9.

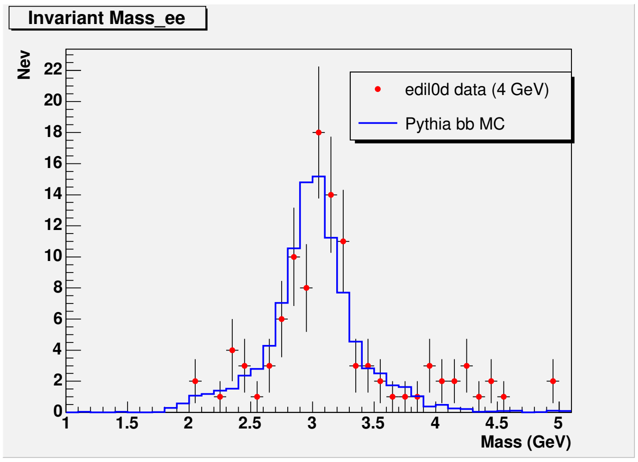


Figure 8: Figure shows the J/Ψ mass in data and MC . The mass distributions have been normalized to the number of events from 2.8 to 3.4 GeV.

E_T(GeV)	Data efficiency	Scale Factor(Data/MC)
4 - 5	$79.1^{+8.8}_{-8.3}\%$	0.9954 ± 0.1081
5 - 8	$86.2^{+6.6}_{-6.1}\%$	1.0764 ± 0.0839

Table 9: Data eID efficiency and Pythia scalefactors.

B N-1 efficiencies and plots for eID

The N-1 efficiencies and distributions are shown for all E_T range for J/Ψ events (Fig 2 bottom plot). Refer Figs. 9 to 14.

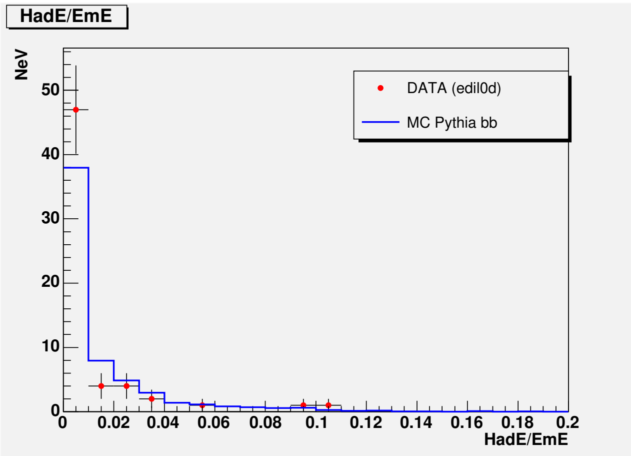


Figure 9: Figures shows N-1 distributions for Data and Pythia MC. MC is normalized to number of events in data.

CUT	Data	Isajet MC	Pythia MC
HadEM	$0.9767^{0.0232}_{0.0293}$	$0.9443^{0.0088}_{0.0086}$	$0.9432^{0.0047}_{0.0047}$
Dz	$0.9767^{0.0232}_{0.0293}$	$0.9984^{0.0015}_{0.0020}$	$0.9959^{0.0014}_{0.0013}$
QDx	$0.9333^{0.0416}_{0.036}$	$0.9984^{0.0015}_{0.0020}$	$0.9766^{0.0032}_{0.003}$
Chi2	$0.9545^{0.0454}_{0.0312}$	$0.9294^{0.0097}_{0.0095}$	$0.9441^{0.0047}_{0.0047}$
Lshr	$0.9999_{0.0263}$	$0.9485^{0.0085}_{0.0083}$	$0.9437^{0.0048}_{0.0047}$
EoP	$0.9999_{0.0263}$	$0.9832^{0.0052}_{0.0049}$	$0.9888^{0.0023}_{0.0022}$
Total	$0.8399^{0.0526}_{0.0499}$	$0.8123^{0.0138}_{0.0137}$	$0.7979^{0.0076}_{0.0076}$

Table 10: N-1 efficiencies for data, Isajet MC and Pythia MC using J/Ψ electrons.

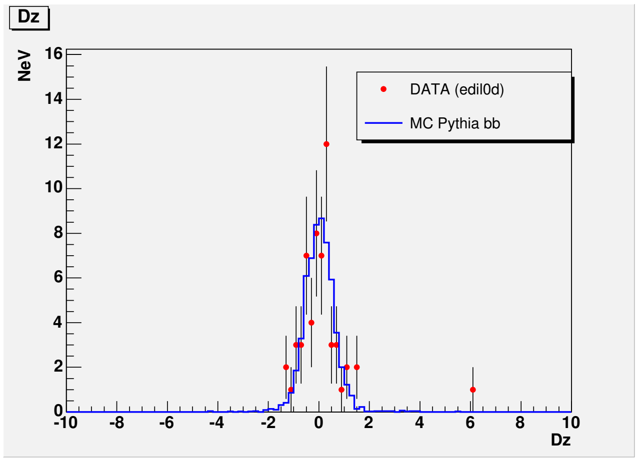


Figure 10: Figures shows N-1 distributions for Data and Pythia MC. MC is normalized to number of events in data.

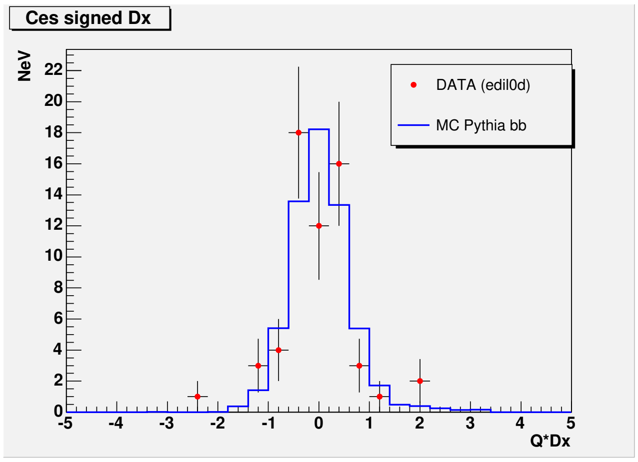


Figure 11: Figures shows N-1 distributions for Data and Pythia MC. MC is normalized to number of events in data.

C Isolation Efficiency Studies

We studied the isolation cut efficiency as a function of number of vertices in the event. Refer Figures 15, 16, 17. These studies were done as checks of our isolation efficiency measurement. We find that our results agree with other studies done within the exotic group.

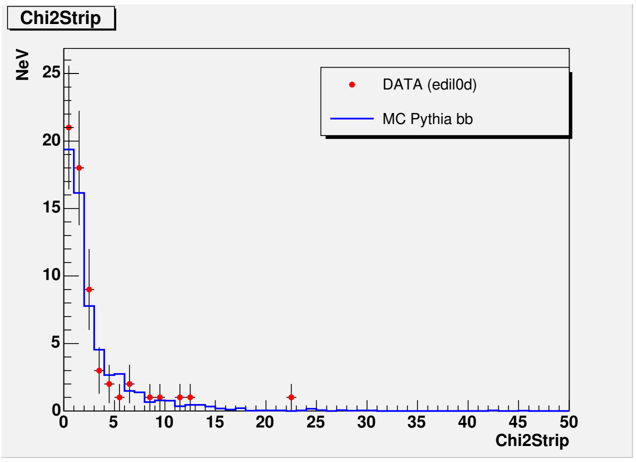


Figure 12: Figures shows N-1 distributions for Data and Pythia MC. MC is normalized to number of events in data.

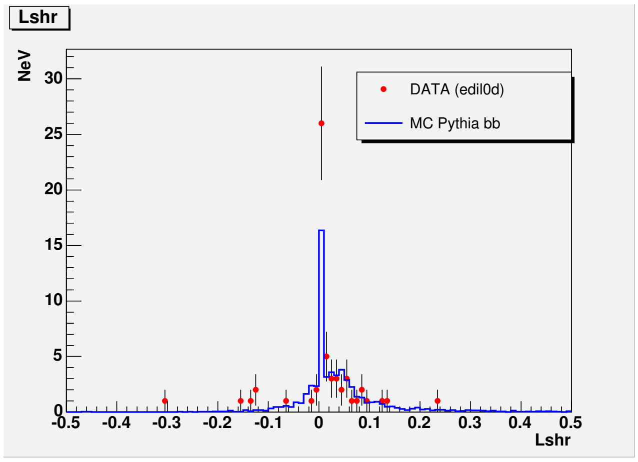


Figure 13: Figures shows N-1 distributions for Data and Pythia MC. MC is normalized to number of events in data.

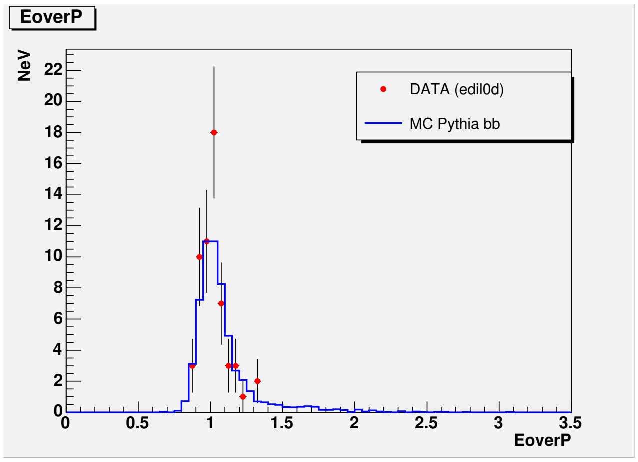


Figure 14: Figures shows N-1 distributions for Data and Pythia MC. MC is normalized to number of events in data.

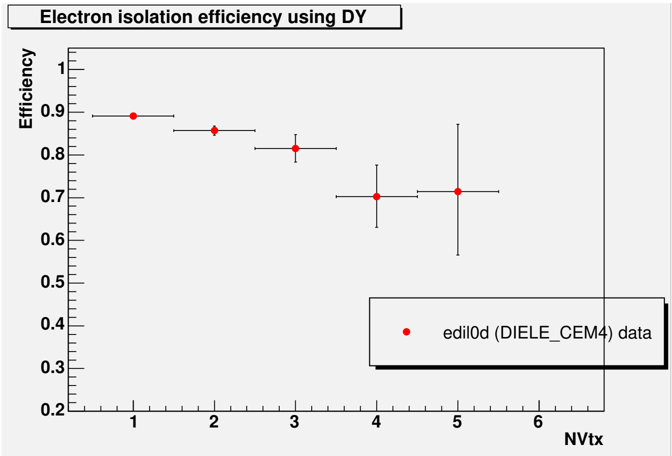


Figure 15: Figures shows isolation efficiency as a function of number of vertices in event for data.

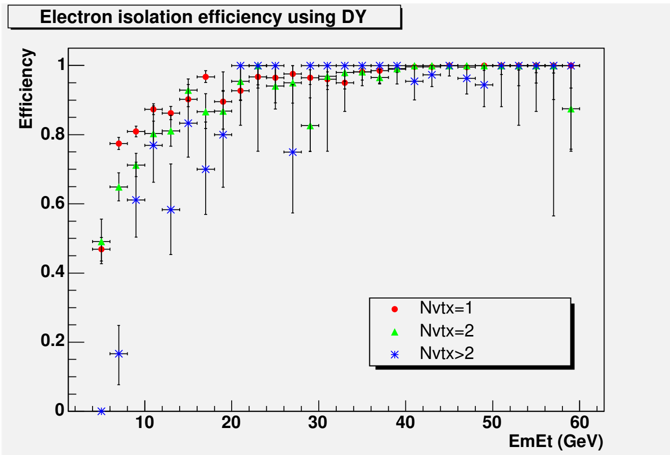


Figure 16: Figures shows isolation efficiency as a function of E_T for Nvtx=1, Nvtx=2 and Nvtx$\geq$3.

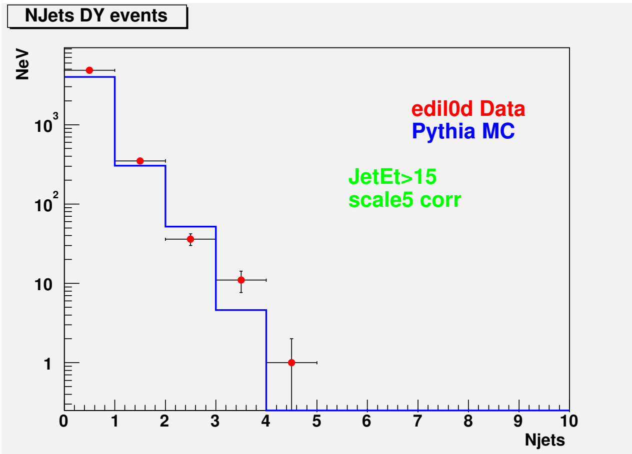


Figure 17: Figures shows jet multiplicity in DY data and MC.