CDF/ANAL/ELECTROWEAK/CDFR/6608
Version 1.0
September 23, 2003

Underlying Event Studies and Photon ID Efficiencies in Run II

Doug Benjamin, Alfred Goshaw, Michael Kirby
Duke University, Physics Department, Durham, North Carolina 27708
Beate Heinemann,Helen Hayward
University of Liverpool, Liverpool, UK L69 3BX
Naho Tanimoto
Okayama University, Okayama 700-8530, Japan

Abstract

We present a study measuring the efficiency of several photon ID cuts. Underlying event activity was studied in selected events using random cones. The data sample was selected from $W \rightarrow e\bar{\nu}$, inclusive jets, and a minbias sample. We determined a correction to additional underlying event as a function of the number of vertices. The correction function applies to all objects and is not limited to just photons. The measured efficiency in data was compared with several Monte Carlo samples, and we determined a correction factor for photon ID efficiency in MC.

1 Introduction

Photon identification efficiency plays a large role in many Electroweak, Exotics, and QCD analyses. Normally measuring the efficiency of an object, one selects a completely clean sample in the data, trying not to discriminate on the cuts studied. For electrons, the standard method selects one tight high P_t electron and a loosely selected second electron. The two objects are then required to reconstruct to an invariant mass within the Z boson window. The loose leg is an almost completely clean sample of electrons, and is ideal for the study additional electron ID cuts. Unfortunately, no physical channel exists to cleanly select photons without applying the cuts you intend to study. So, we use alternate methods to study the photon ID efficiency.

While using Z electrons allows a study of the clusters properties within the calorimeter, since the second Z leg has an isolation cut, the photon isolation cuts need to be studied using a different method. To measure the calorimeter and tracking isolation cuts, we performed a random cone study and discuss it in Section 4. In the analysis, random cones are thrown orthogonal to the leading object in an event. The efficiency of this cone to pass the isolation cuts is measured and compared with the detector simulation. Most analyses depend upon the detector simulation for the acceptance and efficiency for identifying photons. Since this is the case, we present the final result of this study as a correction factor to the simulation and a systematic uncertainty on the correction. The correction is summarized in Table 4.7.

2 Dataset

Several different datasets were used in the analysis of photon ID efficiencies. A high P_T electron sample was used for the $W \rightarrow e\nu$ selection, along with a jet 20 and a minbias sample. We generated three different MC samples for comparison to the data. We discuss each of the datasets below.

2.1 High P_T Electron Data

The primary datasets used in this analysis were the bhel08 and bhel09 samples with a high P_t electron prerequisite filter. We selected the sample from the high E_T electron trigger `ELECTRON_CENTRAL_18`, with data collected between March 2002 and the end of May 2003. This sample is the same sample defined by the good run list from the EWK group for Lepton-Photon 2003 and has an integrated luminosity of $128pb^{-1}$. All data samples were processed using offline version 4.10.4, and then ntupled using the Stntuple dev_240. Documentation on the bhel08-hpte sample, and the entire Stntuple catalog, is located at

`http://ncdf41.fnal.gov/murat/Stntuple/Stntuple.html`

This sample is the baseline dataset for all comparisons to Monte Carlo. This dataset includes the events used in Heather Gerberich's study of central photon ID documented in CDF note 6370.

2.2 Additional Data

Additional cross checks for the random cone studies were done using minbias and Jet20 samples. The minbais sample used was copied from Anwar Bhatti and is documented in CDF note 6042. The Stntuple catalog lists information on the Jet20 sample used, which comes directly from the Jet 20 trigger path.

2.3 Inclusive W Monte Carlo

Since the underlying events studies selected W events, we choose the wewk0e Monte Carlo sample for comparison. This sample was generated using Pythia inclusive $W \rightarrow e\nu$ and the Rick Field Tune A for underlying event. As well, a custom detector model with additional material included in the simulation was used to better match the simulation with measurements in data. The simulated events were processed with offline 4.10.4, and Stntuple dev_240.

2.4 $W\gamma$ Monte Carlo

The primary MC sample used for the $W\gamma$ analysis was a sample generated from WGAMMA by Baur and Berger[2]. Both the electron and muon decays of the W were studied and CDFnote 6608 documents the details of the generation of these samples. The samples were processed with the full detector simulation and offline reconstruction in offline version 4.9.1. They were then ntupled using the Stntuple dev_240, offline version 4.10.4. No additional detector material and no multiple interactions were included in the processing of the samples. Additional comparisons were made between the data and these samples to validate the corrections determined from the wewk0e sample.

2.5 Single electron and photon Monte Carlo

Single electron and photon samples from FakeEvt were used as cross checks on the simulation using this same offline presciption as the other Monte Carlo samples.

3 Photon selection cuts

The Photon working group established a set of baseline identification cuts for photon. These selection cuts for central calorimeter identification are listed in Table 3. We choose this cut order for ease in calculating the efficiency. The hadronic to electromagnetic energy ratio cut is first. This ratio cut is the same as the cut applied when constructing electromagnetic clusters in the offline, and so will identically have an efficiency of 1.0 in all studies. The cuts are then ordered: fiducial, isolation, and finally CES cuts. A strong word of warning should be made here; this cut order is different from the standard Photon group cuts, and care must be taken when trying to compare individual cut efficiencies. Since all of the cut quantities are the same and only the order changed, the final complete efficiencies can be compared with other photon ID efficiency measurements.

Central Photon	
E_T	> 7.0 GeV
E_{had}/E_{em}	$< 0.055 + 0.00045 \cdot E\|\| < 0.125 \cdot E$
CES $\|x\|$	< 21.0 cm
CES $\|z\|$	$9.0 < \|z\| < 230.0$ cm
Iso(0.4) (NVertices Corr)	$< 0.1 \cdot E_T$ ($E_T < 20$ GeV)
	$< 2 + 0.02 \cdot (E_T - 20)$ ($E_T > 20$ GeV)
N3D	≤ 1 and $p_t < 1 + 0.005 \cdot E_T$
TrackIso(0.4)	$< 2 + 0.005 \cdot E_T$ GeV
CES $(\chi^2_{wire} + \chi^2_{strip})/2$	< 20
2nd CES Cluster	$< 0.14 \cdot E_T$ ($E_T < 18$)
	$< 2.4 + 0.01 \cdot E_T$ ($E_T > 18$)

Table 1: The list of central photon selection cuts.

4 Photon isolation efficiencies using random cones

4.1 Technique Overview

In order to measure the efficiency of the photon isolation cuts, we applied a random cone technique in the selected events. The necessity for this technique arises since there is no way to gather a pure photon sample without first applying isolation cuts. Therefore to study these cuts, a cone of 0.4 size is constructed in the detector and the isolation cuts applied to the cone. We then calculate the efficiency for all cones thrown. Once an event is selected, the trigger object determines the vector that the cone points toward. After making the W selection cuts listed in Table 4.1 in the high P_T electron sample, the phi of the cone was set to $\phi_e + 90^o$ and the η set to a random value [-1.1,1.1]. The same determination is done in the jet sample, except that the ϕ of the leading jet is used. Since there is no trigger object in the minbias sample, both a random η and ϕ are chosen. Since the cone itself has no E_T, we chose an arbitrary E_T for the cone in order to apply the cuts that are ratios or sliding. For this study, those values were 7, 10, 15, 20, 25, 30, 35, 40, 45, and 50 GeV. The cuts are then applied in the order listed in Table 3 and the efficiency defined as the simple ratio in Equation 1.

$$Eff = \frac{Num^{Passed}_{Cones}}{Num^{Total}_{Cones}} \quad (1)$$

As the fiducial cuts are solely geometric, we assume that they have an efficiency of 100%. There was no requirement that the cone chosen not point at an object. Thus the effect of a jet "clobbering" a photon is measured along with the effect of underlying event.

The isolation cuts have several corrections that are applied to the cone based upon the topology of the event. The largest and most important correction is due to additional underlying event energy from multiple interactions. This correction was determined in Run 1 by Peter Wilson and is documented in Ref [3]. The correction is a linear function and depends on the number of reconstructed vertices in the event. It is then subtracted from the energy found within the cone of 0.4 surrounding the photon. The Run 1 results for the correction is shown in Equation 2.

$$V_{Corr} = 0.2325 \cdot (N_{vertices} - 1) \quad (2)$$

When summing the tracking isolation quantity, the Z_0 of the track is required to be within 5 cm of the event vertex. Finally, all of the energies measured in the calorimeter were calculated adjusting for the event vertex. The event vertex in W events was selected as the Z_0 of the track, while in the other samples as the highest sum P_T Class 12 vertex. For the definition of a Class 12 vertex, see Ref [4]. With these corrections applied, we then measure the efficiency of the photon cuts in data and Monte Carlo.

$W \rightarrow e\nu$ Cuts	
E_T	> 25.0 GeV
$\|\eta\|$	< 1.1
P_T	$> 10.0 GeV/c$
track quality	$N_{Ax} > 2, N_{St} > 2$, BC and LC
E_{had}/E_{em}	$< 0.055 + 0.00045 \cdot E$
E/P	$< 2\|\|E_T > 50.0GeV$
$Iso4/E_T$	$< 0.1GeV$
χ^2_{strip}	< 10.0
$q \cdot \Delta X$	$> -3.0cm\&\& < 1.5cm$
$\|\Delta Z\|$	$< 3cm$
L_{shr}	< 0.2
FIDELE	==1
Missing E_T	$> 25.0GeV$
$M_T(e,\nu)$	$> 30.0\&\& < 120.0GeV/c^2$

Table 2: The list of central photon selection cuts.

4.2 Underlying Event Correction

We first studied the underlying event correction applied for the multiple interactions in order to confirm the results from Run 1. To do this, the total energy in random cones was plotted versus the number of vertices in W events selected from the high E_T electron sample. We chose to take the vertex collection from the "ZVertexCollection", requiring that the vertex be Class 12 or higher. For a small fraction of the W events, no Class 12 vertex is reconstructed containing the high P_T track of the electron. This is due to a cut in the vertexing algorithm on the minimal number of tracks. Looking at the distribution of δZ_0 between the electron track and events with only one Class 12 vertex, shown in Figure 1, the reconstruction resolution appears to be less than 3.0 cm. So for counting vertices, the Z_0 of the electron is one vertex, and then every additional Class 12 vertex more than 3.0 cm away from this Z_0 is an additional vertex. Using this counting convention, the total energy in the cone versus the number of vertices was fitted to a linear function. Figure 2 shows the data for W, Jet20, and minbias events, and Equation 3 gives the final fit for the multiple interaction correction in W events.

$$V_{Corr} = 0.28 \cdot (N_{vertices} - 1) \tag{3}$$

Applying the same technique to jet20 data gave an equivalent fit within errors, ± 0.01. While the fit in the minimum bias data was lower by 0.03. This is expected due to the fact that the minbias data is not luminosity weighted, while the W data is biased towards higher luminosity. So it is recommended that this correction be applied only in high E_T samples, and not be used in minbias samples without further study. The correction shown above was then used in all of the studies that follow.

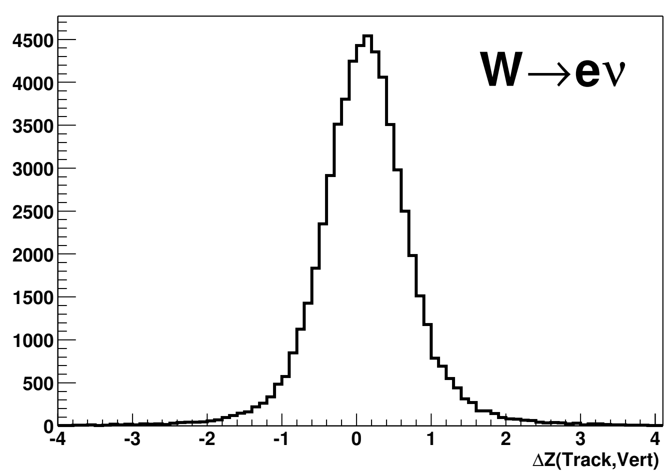

Figure 1: ΔZ distribution between the Z_0 of the W electron and the closest Class 12 vertex from the ZVertexCollection.

4.3 Calorimeter Isolation Efficiency

The efficiency of the calorimeter isolation in data and Monte Carlo are plotted in Figure 3. All events in the data sample are plotted here. If the events are separated by the number of vertices in the event, a strong dependence on the number of vertices is apparent. This is shown in Figure 4. While the correction for underlying event from multiple interactions corrects some of this difference, it is not expected to completely correct it. This effect is due to the fact that with

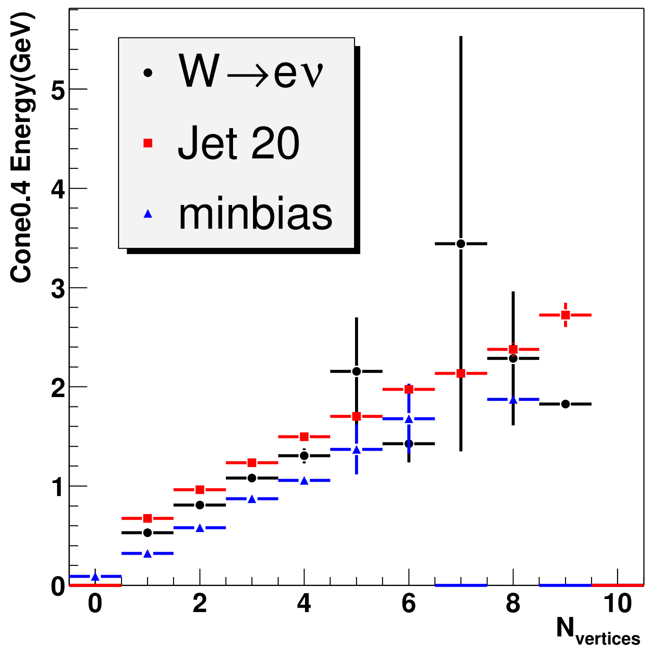

Figure 2: Isolation Energy in a cone of 0.4 versus the number of vertices in the event for $W \to e\nu$, Jet20, and Minbias events.

more than one interaction in an event, there will be more final state objects crossing the detector. A photon will then have less vacant space in the detector to deposit its energy, and thus the photon will be "clobbered" more often. Of course, since there were no additional interactions in the generation or simulation, no such effect occurs in Monte Carlo.

Depending on the instantaneous luminosity during a run, there will be a different rate of multiple interactions. Such changes will affect the efficiency of the calorimeter isolation, and when determining a correction between data and MC, the profile of instantaneous luminosity must be accounted for by weighting the different distributions in Fig 4. But since the efficiency is measured in the data sample itself, for analyses that select W events the weighting is done by taking the single data curve shown in Fig 3. This curve by definition has the appropriate luminosity profile, and therefore the correct weighting of events with multiple interactions. The ratio of the two efficiency curves in Fig 3 was made and fitted with a third order polynomial to give the correction to the MC efficiency as a function of photon E_T. The ratio and fit are shown in Fig 5, with the fitted function given in Eq 4.

A quick note on the error for the correction to the MC efficiency: The number significant digits listed should not be taken as factual. At the same time, since the fit is a third order polynomial small changes in the coefficients does have an effect on the fit. Changes in these parameters based upon the fit error give an error that was determined to be 1% on the calorimeter isolation correction.

$$Eff_{corr} = Eff_{MC}\left(0.883515 + 0.0178008E_T^{\gamma} - 0.00101842E_T^{\gamma 2} + 0.000020308E_T^{\gamma 3}\right) \quad (4)$$

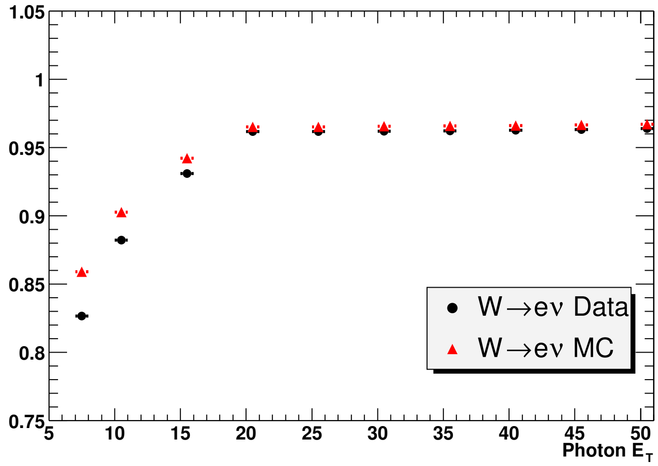

Figure 3: The isolation efficiency is plotted versus the assumed photon energy in $W \to e\nu$ data and Monte Carlo events. The data is for any number of vertices, while the Monte Carlo has only one vertex.

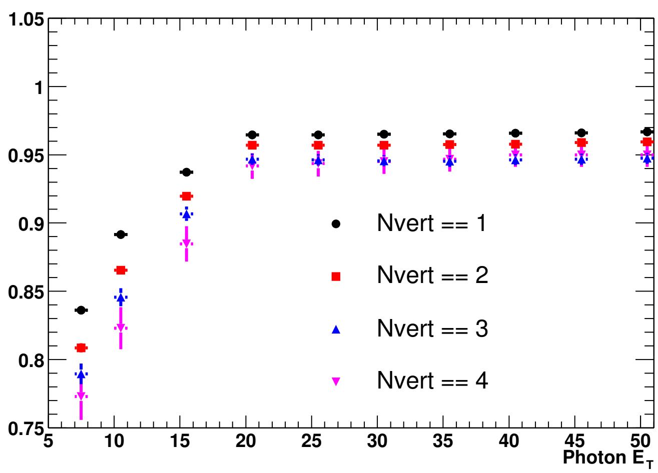

Figure 4: The isolation efficiency is plotted versus the assumed photon energy in $W \to e\nu$ events. The data are separated based upon the number of vertices in the event, and it can be seen that the efficiency falls as the number of multiple interactions increases.

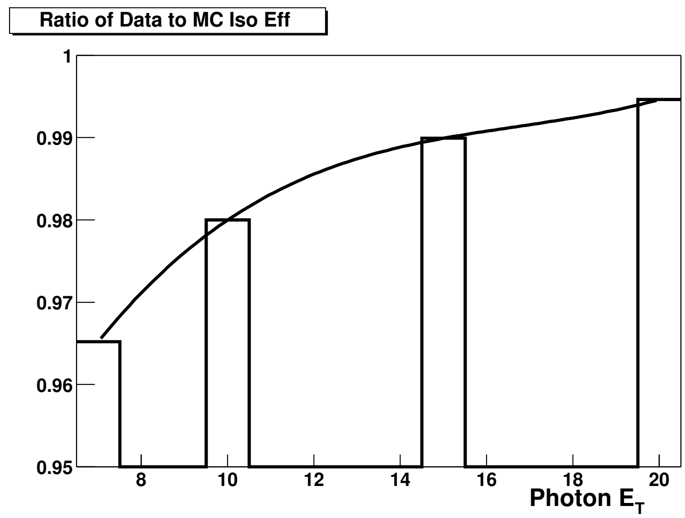

Figure 5: The ratio of the calorimeter isolation efficiency in W events to events in the wewk0e sample. The ratio is fit with a third order polynomial up to $20GeV$, above which the ratio is 99.4%.

4.4 N3D Efficiency

After applying the calorimeter isolation cut to the random cone, the number of 3D tracks pointing at the cone was counted. For a track to be counted, it was required to point at either the central tower, or one of the two shoulder towers in the center of the cone. (These three towers would normally make up the EM Cluster being considered as a photon candidate.) Also, the Z_0 of the track must be within 5 cm of the event vertex. If only one track is matched to the "cluster" towers, then a P_T cut is applied. The efficiency of the cut was found to be very high, and matched between data and MC within errors. Fig 6 shows the efficiency for data and MC. The high efficiency was due to the strong correlation between calorimeter isolation and the number of final state particles traversing that area of the detector. Since the agreement was so good, no correction for the N3D cut was made to the MC efficiency.

4.5 Tracking Isolation Efficiency

After cutting on N3D, a track isolation cut on the random cone is studied. As mentioned previously, for a track to be added to the sum P_T, the Z_0 of the track must be within 5 cm of the event vertex. The sum all of the tracks within the cone of 0.4 is then calculated with no cut on the minimum track P_T. The cut is found to be very efficient and the measured values in data and MC match within errors. Fig 7 shows the efficiency versus photon E_T.

4.6 2nd CES Cluster Efficiency

A cut on the maximum energy of a second CES cluster is applied to photons. Since the random cone is not supposed to be pointing at any object in the calorimeter, there is no first CES cluster. Thus, the highest energy cluster in the wedge was considered as the second cluster using the assumption that the photon will always generate a CES cluster. This assumption seems justified based upon the results of the $Z \rightarrow ee$ study discussed earlier. A second cluster is required to be in the same wedge as the seed tower, and on the same side of the z split of the CES. Since

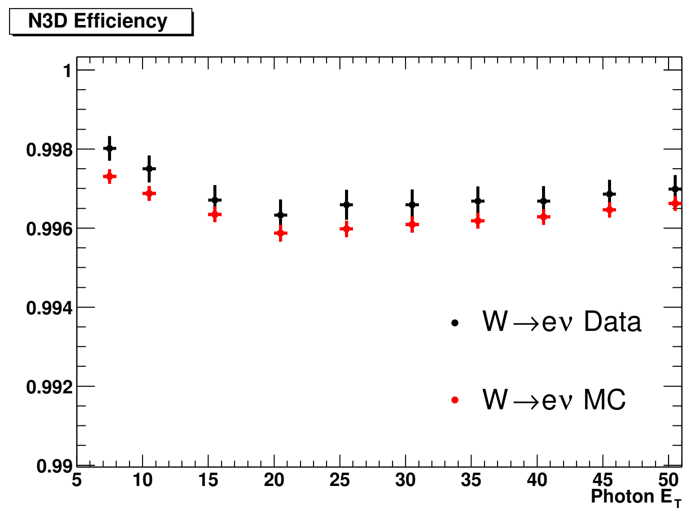

Figure 6: The efficiency of the N3D cut for data and MC versus photon E_T. This plot show the full cut including the sliding cut on P_T if N3D is 1.

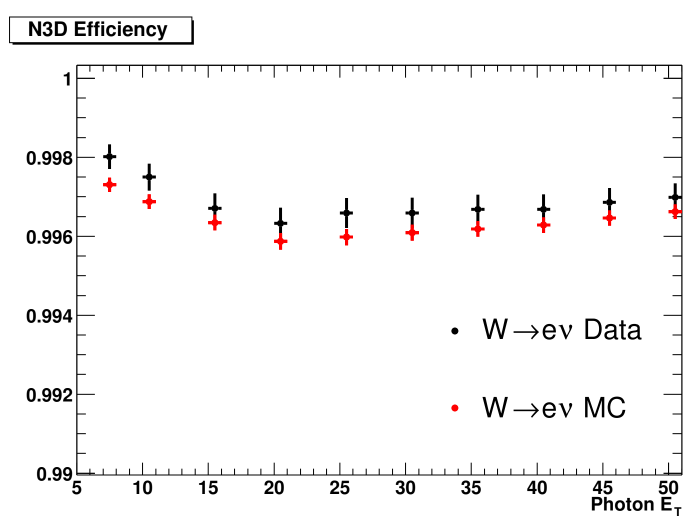

Figure 7: The efficiency of the tracking isolation cut for data and MC versus photon E_T.

the random cone thrown is empty, measuring the efficiency of this cut only considers additional clusters from underlying event. This study has no sensitivity to asymmetric photon conversion after the COT and in the coil. The difference between data a MC is shown in Fig 8. Since there is no dependence on photon E_T, a simple correction of 1% was applied to the MC efficiency.

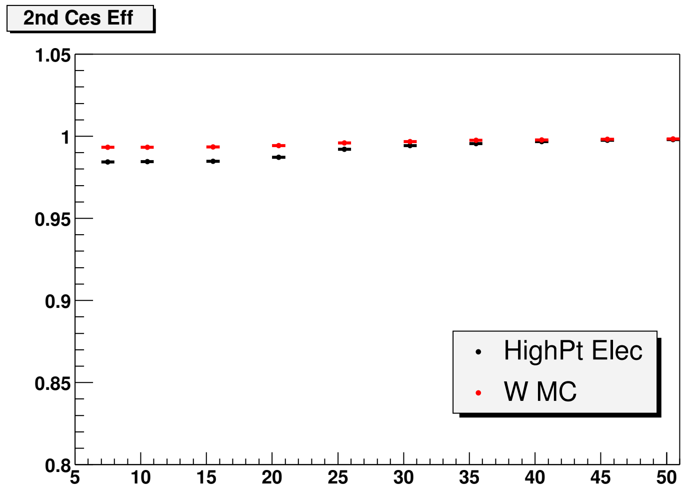

Figure 8: Efficiency of the 2nd CES cluster cut as a function of photon E_T.

4.7 Summary of Random Cone Study

After looking at the isolation cut efficiency in the data and MC, it was determined that only the calorimeter isolation and second CES cluster cut needed to be corrected. Table 4.7 lists the measured efficiency of each cut and the correction, if any. The other cuts matched within errors, and so can be measured directly from the MC. A newly measured correction for the underlying energy in events with multiple interactions was determined and used for the study. The final correction to the MC photon efficiency in W events for isolation cuts is shown in Eqs 5.

$$E_T < 20.0GeV \quad , \quad Eff_{corr} = 0.99 Eff_{MC} \left(0.883515 + 0.0178008 E_T^{\gamma} - 0.00101842 E_T^{\gamma 2} + 0.000020308 E_T^{\gamma 3}\right)$$
$$E_T > 20.0GeV \quad , \quad Eff_{corr} = 0.984 Eff_{MC}$$

Cut	$W \to e\nu$ MC	$W \to e\nu$ Data	Correction	Error
Cal Iso	91.9%	90.3%	F(E_T)	1.0%
N3D & P_t	99.6% 99.7%	None	0.3%	
Track Iso	98.1% 98.0%	None	0.1%	
2nd CES E	93.0%	92.0%	1.0%	1.0%

Table 3: The measured efficiencies of various photon ID cuts in Monte Carlo and Data. The measured correction is listed along with the error.

5 Conclusion

We measured the activity in random cones in the data, and determined corrections for photon selection. The correction to additional underlying event was measured in W, Jet20, and minbias events and found to be significantly different from the Run 1 result. This correction was then applied to random cones to determine the efficiency of photon ID selection. After comparing the measured efficiency in data and monte carlo, we calculated a correction to the photon efficiency. The correction was applied to the $W\gamma$ and $Z\gamma$ Monte Carlo for the $V+\gamma$ electroweak analysis, documented in CDF note 6601 documents these results.

References

[1] ETF and Muon group, "Baseline Analysis Cuts for High P_T Isolated Leptons", `http://www-cdf.fnal.gov/internal/physics/ewk/hipt_lepton_baseline_cuts.html`.

[2] U. Baur and E.L. Berger, "Probing the $WW\gamma$ vertex at the Fermilab Tevatron Collider", Phys. Rev. D 41, 1476 (1990).

[3] Peter Wilson, "Calorimeter Isolation and Lateral Shower Leakage for Photons and Electrons", CDF Note 4170.

[4] J.-F. Arguin, B. Heinemann, A. Yagil, "The z-Vertex Algorithm in Run II", CDF Note 6238.