DØ/ANALYSIS/3700
Version 1.1
Original Version, Nov. 9th. 1999
January 26, 2000

The $e\mu$ Data Set Used for the Sherlock Analysis

Dave Toback[1], Bruce Knuteson[2], Jim Cochran[3], Harpreet Singh[4]

[1]University of Maryland
[2]University of California at Berkeley
Lawrence Berkeley National Laboratory
[3]Iowa State
[4]University of California at Riverside

Abstract

This note contains a description of the $e\mu$ datasets, background samples and rate estimates used for the Sherlock analysis. These results are based on and are an update to the datasets used for the top group conventional and neural net $t\bar{t} \rightarrow e\mu$ cross section analyses.

1 Introduction

Sherlock is a quasi-model independent prescription for searching for the physics of electroweak symmetry breaking [1]. The prescription can be summarized as follows:

- Defining a final state
- Choosing the relevant variables for that final states
- Using an algorithm which uses those final state variables to detect a possible signature for new physics as a localized excess above expected backgrounds

DØ Note 3701 [1] describes the prescription and the method for choosing the final states and relevant variables. It completely specifies which variables should be used for the general analysis. DØ Note 3715 [2] describes the details the types of variables which are applicable to the Run I DØ detector data. This note details the $e\mu$ dataset for use in the Sherlock prescription, specifically, a description of the $e\mu$ superset, the exclusive datasets, the background samples and rate estimates. While this note can be read stand-alone for the dataset description, it is not intended to stand alone from an analysis point of view. The dataset here is used as input to the Sherlock algorithm [3] with the results of the found in Ref. [4].

The $e\mu$ dataset and the results presented in this note are based on and are an update to the datasets used for the top group conventional [5] and neural net [6] $t\bar{t}\rightarrow e\mu$ cross section analyses. In this note we do the following:

- Define the $e\mu$ superset and exclusive dataset selection criteria and variables
- Identify the Standard Model backgrounds for the final state
- Estimate the number of events for each background channel
- Create datasets of the background and all events in the Run I data passing the selection criteria

The bulk of this note describes the defining, identifying, estimating and creating. We begin with a listing of the final state, the backgrounds and a definition of the exclusive datasets. After a discussion of the data sets used in the analysis, we give the identification and isolation selection criteria. We continue with a description of the efficiencies and correction factors needed to convert from MC samples to event data rates. Finally, we give the expectations and number of observed events.

2 Final State Definition, SM Backgrounds and Datasets

Our final state signature is simply:

- 1 high E_T ($E_T > 15$ GeV) isolated electron
- 1 high P_T ($P_T > 15$ GeV) isolated muon

The dominant Standard Model processes which produce this signature are:

- $t\bar{t} \rightarrow e\mu\nu\bar{\nu}jj$
- $WW \rightarrow e\mu\nu\bar{\nu}$
- $Z^0 \rightarrow \tau^+\tau^- \rightarrow e\mu\nu\nu\bar{\nu}\bar{\nu}$
- Drell-Yan $\rightarrow \tau^+\tau^- \rightarrow e\mu\nu\nu\bar{\nu}\bar{\nu}$
- Fakes $\rightarrow e_{\rm fake}\mu + X$

As described in DØ Note 3701 [1], we use exclusive datasets. Since we can effectively select e's, μ's and jets these become our objects. Since we cannot effectively select events with small $\not\!\!E_{\rm T}$ and since by lepton number conservation there must be at least 2 neutrinos in the event we assume all events have $\not\!\!E_{\rm T}$. We define the following exclusive datasets:

- $e\mu\not\!\!E_{\rm T}$

  Variables: $\rm P_T^e$, $\not\!\!E_{\rm T}$
- $e\mu\not\!\!E_{\rm T}j$

  Variables: $\rm P_T^e$, $\not\!\!E_{\rm T}$, $\rm E_T^{j1}$
- $e\mu\not\!\!E_{\rm T}jj$

  Variables: $\rm P_T^e$, $\not\!\!E_{\rm T}$, $\rm E_T^{j2}$
- $e\mu\not\!\!E_{\rm T}jjj$

  Variables: $\rm P_T^e$, $\not\!\!E_{\rm T}$, $\rm E_T^{j2} + E_T^{j3}$

The number of expected background events becomes very small for other combinations.

3 Data Selection

The data selection process for this sample proceeds through several stages. After a description of the original datasets and the trigger requirements, we discuss some preliminary data reduction and analysis tools used to select the final datasets.

3.1 Preliminary Data Selection

The starting data set for this analysis is the ALL stream μDSTs (149 million events) for Runs 1A, 1B and 1C. From that sample the subset of events which had at least one loose EM object and one very loose muon were stripped and stored (as μDSTs). These requirements are given in Table 1. After stripping, all μDST files were passed through a filter to remove duplicate events. After preliminary selection, all events are required to come in on one of the L2 triggers [7] as given in Table 2.

PELC/PPHO Requirements	PMUO Requirements
$E_T^{EM} \geq 10$ GeV	$P_T \geq 7$ GeV/c
$\lvert\eta^{EM}\rvert \leq 3.3$	$\lvert\eta^{\mu}_{Det}\rvert \leq 2.0$
Iso ≤ 0.3	
$\chi^2 \leq 300$	

Table 1: The preliminary dataset requirements to select $e\mu$ candidates. Each event is required to have at least one PELC/PPHO and one PMUO. Vertex looping was done such that if any PELC/PPHO passed the $E_T^{EM} \geq 10$ GeV cut from any vertex, then that PELC/PPHO passed. No additional particle ID cuts were required for PMUOs.

Run	Triggers Allowed
Run 1A	mu_ele, ele_jet_high or mu_jet_high
Run 1B	mu_ele, ele_jet_high, mu_jet_high or mu_jet_cent
Run 1C	mu_ele_high, ele_jet_higha, mu_jet_cencal, or mu_jet_cent

Table 2: The $e\mu$ dataset trigger requirements for Runs 1A, 1B and 1C.

3.2 Good Runs and Luminosity

All events in runs on the "official" bad run list [8] are removed. In addition, events collected during MRBS-LOSS and MICRO-BLANK gates are rejected [9]. The luminosities for the samples are 13.9 pb^{-1} and 94.4 pb^{-1} for Runs 1A and 1BC respectively. Taking into account the error, the total integrated luminosity for this sample is 108.3±5.7 pb^{-1}. We note that the Prezap fraction is 0.54±0.04 and the Postzap fraction=0.46±0.04 for the Run 1BC sample.

3.3 Event Processing and Ntuple Making

After preselection, the events are processed through a two stage analysis. In the first stage the μDST events are processed using version 7 of the top dilepton ntuple maker [10] and version 5.0 of CAFIX [11]. In the second processing stage, the ntuples are analyzed using a customized standalone package which was used for the $t\bar{t} \rightarrow e\mu$ neural net cross section analysis [6, 12]. The program fills a separate, smaller, ntuple with the subset of events which pass the $e\mu$ selection cuts for each dataset separately. These same packages are used to process the Monte Carlo background samples.

4 Background Samples

In this section we describe the backgrounds and the samples generated to estimate rates and distributions. Each event is processed through DØGEANT [13] using 'showerlib" [14]. In addition each event is run through μsmear [15] with 50% run through Prezap smearing and 50% run through Postzap smearing. Before putting the data into the Ntuples the DØRECO simulation is run. We summarize the primary numbers (the visible cross sections and number of events in the Monte Carlo samples) in Table 3 for later use.

Sample	σ_{visible} (pb)	N_{Total} (Events)
$t\bar{t} \rightarrow e\mu + X$	0.40	101,339
$Z^0 \rightarrow \tau^+\tau^- \rightarrow e\mu + X$	26.6	71,393
$WW \rightarrow e\mu + X$	0.685	101,792
Drell-Yan $\rightarrow \tau^+\tau^- \rightarrow e\mu + X$	5.5	199,990

Table 3: The visible cross sections and number of events in the Monte Carlo samples for use in Section 9. We note that the $t\bar{t}$ sample also includes ee and $\mu\mu$ final states which is reflected in the total visible cross section. Similar statements are true for $Z^0 \rightarrow \tau^+\tau^-$, WW and Drell-Yan $\rightarrow \tau^+\tau^-$.

4.1 $t\bar{t}$

The $t\bar{t}$ sample consists of 101,339 events generated using the HERWIG Monte Carlo generator, version 5.7, with CTEQ3M PDF's. For each event, each W is required to decay to a e, μ or τ. Each τ is required to decay to an e or μ. Using $\text{Br}(\tau \rightarrow eX) = (18.01 \pm 0.18)\%$ and $\text{Br}(\tau \rightarrow \mu X) = (17.65 \pm 0.24)\%$ [16] gives a branching fraction of Br=0.0685. We use a theoretical top cross section [17] of $\sigma_{t\bar{t}} = 5.83$ pb for a total $\sigma_{t\bar{t}}^{\text{visible}} = 0.40 \pm 0.05$ pb.

4.2 $Z^0 \rightarrow \tau^+\tau^-$

The $Z^0 \rightarrow \tau^+\tau^-$ sample consists of 71,393 events generated using the ISAJET Monte Carlo. Here all τ pairs are required to decay to $ee, e\mu$ or $\mu\mu$ for a branching ratio of $\text{Br}(\tau\tau \rightarrow ee, e\mu \text{ or } \mu\mu) = 0.127$. We take the cross section for the $Z^0 \rightarrow \tau^+\tau^-$ to be equal to the published DØ $Z^0 \rightarrow e^+e^-$ cross section [18]. Plugging in we find:

$$\sigma^{\text{visible}}_{Z^0 \rightarrow \tau^+\tau^- \rightarrow e\mu} = (209 \pm 21 \text{ pb}) \times 0.127 = 26.6 \pm 2.66 \text{ pb}. \quad (1)$$

4.3 WW

The WW sample consists of 101,792 events generated using the PYTHIA Monte Carlo. In this sample, as in $t\bar{t}$, we only simulate the leptonic decays. We take a WW cross section of 10±1 pb from Ref. [19] and correct for the branching fraction giving $\sigma^{\text{visible}}_{WW \rightarrow e\mu} = 0.685 \pm 0.69$ pb.

4.4 Drell-Yan $\rightarrow \tau^+\tau^-$

The Drell-Yan $\rightarrow \tau^+\tau^-$ sample consists of 199,990 events generated using the PYTHIA Monte Carlo. In this sample, as in $Z^0 \rightarrow \tau^+\tau^-$, we only simulate the leptonic decays. Due to the exponential rise in the Drell-Yan $\rightarrow \ell^+\ell^-$ cross section as $\text{m}_{\ell\ell} \rightarrow 0$, we use a sample with $m_{\tau^+\tau^-} > 25$ GeV and force the τ's decay to e and μ ($\sigma_{\text{PYTHIA}} = 64$ pb). This sample is then normalized by comparing a second sample from PYTHIA with Drell-Yan $\rightarrow e^+e^-$ for 30 GeV $\leq \text{m}_{ee} \leq 60$ GeV ($\sigma_{\text{PYTHIA}} = 34$ pb) to the 45±7 pb measured in Ref. [20]. Correcting the PYTHIA cross section for the measured Drell-Yan $\rightarrow e^+e^-$ cross section and taking into account the branching ratio we find:

$$\sigma^{\text{visible}}_{\text{Drell-Yan}\to\tau^+\tau^-\to\ell\ell} = (64\text{ pb})\cdot\frac{45\pm 7}{34}\cdot(0.06358) = 5.5\pm 0.9\text{ pb}$$

4.5 Fakes

The $e\mu$ final state can be faked if events which contain jets mimic the electron and/or muon signature. The rate at which hadrons fake the muon signature via punch through is estimated to be $< 10^{-4}$ events [21]. Electrons can be faked from jets at the 10^{-4} level so we only consider them further. Both the number of events containing fake electrons as well as their distributions are taken directly from the data. Specifically, events are selected in the same manner as the final data sample (giving the correct luminosity normalization), but with different, 'bad'-electron ID cuts. The cuts and procedures are described in the following sections.

5 Electron, Muon, Jet and Missing E_T Identification

As previously described [1, 2], we use standard selection criteria and object and variable definitions for electrons, muons, jets and Missing E_T. In this section we describe the identification variables and selection criteria. The efficiencies and corrections used to go from Monte Carlo to data are discussed in the next section.

5.1 Electrons

We use the standard top group "loose" electron definition which is described in Table 4.

$E_T^e \geq 15$ GeV with $\lvert\eta^e\rvert \leq 2.5$
5 dimensional Likelihood ($Ł_5$) ≤ 0.5
Iso ≤ 0.1

Table 4: The standard top group "loose" electron definition which is used in this analysis. Likelihood variables for electrons are defined in Ref. [22].

5.1.1 Electron Fakes

The number of fakes in the final data sample is taken from the data by looking at 'bad' electrons. After all data requirements (good run, trigger, muon ID and iso, etc.) we only require that an electron candidate pass loose EM cuts, given in Table 5. This allows us to use the electron fake rate measurement of Ref. [23]. The fake 'Rate' (given later) is the rate at which jets which pass the cuts in Table 5 also pass the requirements in Table 4[1].

[1] i.e, P(Jet passes Loose|Jet passes Really Loose)

$E_T^e \geq 15$ GeV with $\lvert\eta^e\rvert \leq 2.5$
$\chi^2 < 300$
Isol ≤ 0.3

Table 5: The standard top group fake electron definition which is used in this analysis to determine the fakes sample. Note that the $\chi^2 < 300$ requirement is an additional requirement which uses the CAFIX v5.0 fixed value.

5.2 Muons

We use the standard top group "loose" muon definition as described in Table 6. Note that we use the notation central = CF and endcap = EF, and that EF muons are only accepted in Postzap data.

	1A	1B	1BC	
		Prezap	Postzap	
	CF	CF	CF	EF
P_T^μ	15	15	15	15
IFW4	1	1	1	0
calmip	yes	no	no	no
mtc	no	yes	yes	yes
$\Delta R^{\mu-jet}$	0.5	0.5	0.5	0.5
$\Delta R^{\mu-e/\gamma}$	0.25	0.25	0.25	0.25

Table 6: The standard top group "loose" muon definition. CF$\equiv$ quad $\leq$4, EF$\equiv$ quad $>$ 4 .AND. quad $\leq$ 12. astubs are rejected (the default in top dilep). mtc $\equiv$ (HFrac $\geq$ 1) .OR. (HFrac $\geq$ 0.7 .AND. EFracH1 $>$ 0). calmip $\equiv$ (CD match and 1 GeV in hit cell + 1NN) .OR. (No CD match and 1.5 GeV in hit cell + 1NN). Note that there are no cuts on MUCTAG, XOCT, 3d Impact Param, number of hits, or Bdl. Also note that following the rest of the top group, we exclude 1A and Prezap EF muons. Isolation: $\Delta R^{\mu-jet} > 0.5$ for "all" reconstructed jets (see Section 5.3).

5.3 Jets

We follow the top group standard and consider jets of cone size $\Delta R = 0.5$ which have been corrected with CAFIX 5.0 [11]. In determining whether or not a muon is isolated we require that $\Delta R^{\mu-jet} > 0.5$ for all reconstructed jets with $E_T > 5$ GeV. For calculating H_T and jet counting we consider jets with $E_T > 15$ GeV and $\lvert\eta\rvert < 2.5$.

5.4 Missing E_T

We take the calorimeter missing E_T ($\not\!\!E_T^{\mathrm{cal}}$) directly from the PNUT4 bank.

6 Efficiencies

In this section we discuss the efficiencies, fake rates and correction factors for the electrons and muons. For completeness we note that the efficiency is that rate at which electrons pass all the ID and isolation requirements and the correction factor is used to take Monte Carlo sample rates to data rates.

6.1 Electrons

The efficiency and Monte Carlo correction factors are determined separately for real data and for Monte Carlo (See Ref. [6] and references therein). The data efficiencies have been determined from $Z^0 \to e^+e^-$ data. The Monte Carlo efficiencies are determined without any electron ID requirements (except isolation) and then multiplied by the appropriate electron ID efficiency correction factors determined from the data. The results are given in Table 7.

	CC	EC
Efficiency	0.880±0.016	0.638±0.023
Showerlib Tracking Eff (MC)	0.948±0.006	0.907±0.016
Data Tracking Efficiency ($\geq$ 1 vertex events)	0.827±0.011	0.852±0.010
Total MC Correction factor	$(0.880 \pm 0.016) \cdot \frac{0.827\pm0.011}{0.948\pm0.006}$ $= 0.768 \pm 0.018$	$(0.638 \pm 0.023) \cdot \frac{0.852\pm0.010}{0.907\pm0.016}$ $= 0.599 \pm 0.025$

Table 7: Electron efficiency and Monte Carlo correction factors.

6.2 Fakes

The electron fake rate is determined from a sample of QCD multi-jet events. After loose event selection (see Table 5) the fake 'rate' is then given by the fraction of these events which pass the loose electron selection criteria in Table 4. The results were computed separately for the CC and the EC giving the values listed in Table 8 [23].

Region	Electron Fake 'Rate'
CC	0.0461 ± 0.0008
EC	0.0801 ± 0.0009

Table 8: The rate at which jets which create a PELC bank with Iso<0.3 and $\chi^2 < 300$ also pass the Iso < 0.1 and $L < 0.5$ cuts. Note that the errors given here are statistical only. The full error is 7% for CC and 6% for EC [23].

6.3 Muons

The Monte Carlo correction factors for muons are determined using a combination of data and Monte Carlo simulations and are cross checked using $Z^0 \to \mu^+\mu^-$ data. The correction factors are given in Table 9. Note that differences between Prezap and Postzap data are already incorporated into calculating these correction factors. For more details see Appendix B.

Run	CF Correction Factor	EF Correction Factor
Run 1A	0.865±0.073	-
Run 1B and 1C	0.890±0.062	0.504±0.059

Table 9: The muon correction factor table. Note that differences between Prezap and Postzap data are already incorporated into calculating these correction factors. For more details on how this is done see Appendix B.

6.4 Trigger Efficiencies

The trigger efficiencies have been obtained from the "trig sim" package [24] for both 1A and 1BC triggers. These efficiencies are a function of the detector region for electrons and muons and are given in Table 10. Note that we only use a single efficiency for all three runs, and the efficiency for each combination is obtained from trig sim using list V90. Since no trigger requirement is used in the Monte Carlo the correction factor is the same as the efficiency.

Region	Efficiency
CCCF	0.95 ± 0.05
CCEF	0.93 ± 0.05
ECCF	0.90 ± 0.04
ECEF	0.93 ± 0.05

Table 10: Trigger efficiencies for Run I

6.5 Combined Trigger and Particle Identification Efficiencies

The trigger and particle identification correction factors are combined to give four trigger × ID corrections corresponding to four e and μ classes. These are given by:

- $C_{\mathrm{CCCF}} = C^{\mathrm{Trig}}_{\mathrm{CCEF}} \times C^{eID}_{\mathrm{CC}} \times C^{\mu ID}_{\mathrm{CF}}$
- $C_{\mathrm{CCEF}} = C^{\mathrm{Trig}}_{\mathrm{CCEF}} \times C^{eID}_{\mathrm{CC}} \times C^{\mu ID}_{\mathrm{EF}}$
- $C_{\mathrm{ECCF}} = C^{\mathrm{Trig}}_{\mathrm{ECCF}} \times C^{eID}_{\mathrm{EC}} \times C^{\mu ID}_{\mathrm{CF}}$
- $C_{\mathrm{ECEF}} = C^{\mathrm{Trig}}_{\mathrm{ECEF}} \times C^{eID}_{\mathrm{EC}} \times C^{\mu ID}_{\mathrm{EF}}$

Since we are combining the Run 1A and Run1B datasets and want one number from the Monte Carlos, we do a luminosity weighted correction for the run. This is given by:

$$C_{\mathrm{X}}^{\mathrm{Lum\ Weighted}} = \frac{(\mathrm{Lum}^{1\mathrm{A}} \times C_{\mathrm{X}}^{1\mathrm{A}}) + (\mathrm{Lum}^{1\mathrm{BC}} \times C_{\mathrm{X}}^{1\mathrm{BC}})}{\mathrm{Lum}^{1\mathrm{A}} + \mathrm{Lum}^{1\mathrm{BC}}} \tag{2}$$

with the results given in Table 11.

data set	C_{CCCF}	C_{CCEF}	C_{ECCF}	C_{ECEF}
Run 1A	0.631±0.08	-	0.466±0.06	-
Run 1BC	0.649±0.06	0.360±0.05	0.480±0.05	0.281±0.04
Luminosity Weighted	0.647±0.07	0.314±0.04	0.478±0.05	0.245±0.03

Table 11: Trigger $\times$ Identification correction factors as a function of detector region and running period.

7 Background Event Rate Calculations

In the section we describe the calculations required to estimate the number of events in each sample.

7.1 Monte Carlo Samples

Each Monte Carlo sample must be normalized to the correct cross section, luminosity and acceptance. In addition it must be corrected for luminosity dependent effects.

7.1.1 Luminosity Corrections

The data used in the analysis were taken with a wide range of instantaneous luminosities. At the higher luminosities it is common to have two or three interactions within the detector livetime which has the effect of introducing a luminosity dependence into some of the acceptance corrections. This is automatically built into the fake calculations which are done using real data. For Monte Carlo calculations an additional correction must be determined. The Monte Carlo data were generated using only a single hard scattering per event. This effect is taken into account using three different sets of $t\bar{t}\rightarrow e\mu X$ Monte Carlo events corresponding to one, two and three interactions. Each are generated using the NOISY package [25] with:

- one hard scattering (one interaction)
- one hard scattering plus one minimum bias event (two interactions)
- one hard scattering plus two minimum bias events (three interactions)

Each event is processed through the full DØGEANT and DØRECO and used to calculate results for the selection efficiency as a function of the number of interactions. This effect is studied in detail in Ref. [6] with the final acceptance, after all lepton ID/Isolation, $\not\!\!E_{\mathrm{T}}$, $\mathrm{N}_{\mathrm{Jet}}$ and neural net cuts, listed in Table 12. The acceptance decreases as the number of interactions increase.

Number of Interactions	Acceptance
1	0.35±0.02
2	0.31±0.02
3	0.27±0.02

Table 12: Acceptance after all the cuts for different numbers of interactions. Errors are statistical only.

To determine a correction factor, the acceptance variation was parameterized using a straight line fit as a function of the number of interactions, N, yielding:

$$A(N) = 0.39 - 0.04 \times N \tag{3}$$

For an average of 1.83 interactions per crossing [6], the correction factor, $C^{\text{Mult Int}}$, we take:

$$C^{\text{Mult Int}} = \frac{A(N = 1.83)}{A(N = 1.0)} = 0.905 \tag{4}$$

Because of the kinematical and topological similarities between the $t\bar{t}$ and other processes after the cuts, we apply the same correction factor to the Z/γ and WW channels and ignore differences between the different cuts. While the correction factor is only strictly valid after the neural net cuts, the majority of the correction (7%) is due to lepton ID and kinematic failures and (2%) is due to isolation failures.

7.1.2 Monte Carlo Samples

We calculate the number of events from the different channels using the relation

$$\begin{aligned} N_{\text{Background}} &= \text{Lum} \times \sigma_{\text{visible}} \times C^{\text{Mult Int}} \times \\ &\quad \frac{(N_{\text{CCCF}}C_{\text{CCCF}} + N_{\text{CCEF}}C_{\text{CCEF}} + N_{\text{ECCF}}C_{\text{ECCF}} + N_{\text{ECEF}}C_{\text{ECEF}})}{N_{\text{total}}} \end{aligned} \tag{5}$$

where $N_{\text{CCCF}}, N_{\text{CCEF}}, N_{\text{ECCF}}$ and N_{ECEF} are the number of events passing in each region after all cuts, N_{total} is the starting number of events in the Monte Carlo sample, σ_{visible} is the input visible cross section, and C_{CCCF}, C_{CCEF}, C_{ECCF} and C_{ECEF} are the luminosity weighted, combined e, μ and trigger corrections for the different regions. Here we have assumed that the trigger and identification corrections are the same for all processes once kinematic and isolation requirements have been placed.

7.2 Fakes

To determine the number of fake events in the final event sample, we divide the data set (after all goodrun, trigger and muon requirements) into four regions labeled:

- A: Number of Events with $L < 0.5$, Iso<0.1 [2]
- B: Number of Events with $L > 0.5$, Iso<0.1
- C: Number of Events with $L < 0.5$, 0.1<Iso<0.3
- D: Number of Events with $L > 0.5$, 0.1<Iso<0.3

In addition we define:

- A_E: Electrons in Region A
- A_F: Jets in Region A
- ϵ: Electron efficiency

  (rate at which electrons land in region A)

[2]This is the signal region apart from the additional requirement of $\chi^2_{CAFIX5.0}$ <300.

- R: Electron Fake 'Rate'

  (rate at which jets passing PELC/χ^2/Iso cuts land in region A)

As described in Appendix D the number of fakes in the final data set is given by:

$$A_F = R[\frac{A - \epsilon(A + B + C + D)}{R - \epsilon}] \tag{6}$$

Note that A_F is properly a function of electron region (CC or EC) as given in Table 7.

7.3 Putting it all together

The number of expected events uses Equation 5 with values of Lum $= 108.3, C^{\text{Mult Int}} = 0.905$, σ_{visible} and $\mathrm{N}_{\text{Total}}$ from Table 3, the weighed corrections from Table 11. Note that we have only run the Run 1BC muon selection criteria for use in the Table for simplicity. This introduces a less than 2% error in our rates.

8 Uncertainties

In this section we collect the uncertainties into Table 13 for later use. Some notes are in order. We note that the error on the number of jets in the event, as predicted by the Monte Carlo's, is estimated by comparing the extra jets in $Z^0 \to e^+e^-$ from data and $Z^0 \to \tau^+\tau^-$ events from ISAJET. While numbers agree within error, the uncertainty is on the order of 20%. We add this error in quadrature with all rate estimates for the datasets. While, this is not added to the sum of the datasets, that number is not used anywhere.

Source	Error (%)
CCCF Correction	11%
CCEF Correction	13%
ECCF Correction	11%
ECEF Correction	13%
Jet Fake Prob	7%
Multiple Interactions	7%
Luminosity	5.3%
$\sigma(t\bar{t} \to e\mu + X)$	12%
$\sigma(Z^0 \to \tau^+\tau^- \to e\mu + X)$	10%
$\sigma(WW \to e\mu + X)$	10%
$\sigma(\text{Drell-Yan} \to \tau^+\tau^- \to e\mu + X)$	17%
Monte Carlo Jet Generation	20%

Table 13: Sources of systematic error.

9 Results

The number of events expected for the various samples and datasets is given in Table 14. Figures 1 - 5 show the distributions sent to Sherlock with each distribution weighted to be equal to the number of events expected in the data. Figures 6 through 9 show the background predictions for the relevant variables in the four exclusive datasets.

As a check, we have compared all of these results, as well as those for the data, with the neural net [6] $t\bar{t} \rightarrow e\mu$ cross section analysis. We find excellent agreement, with the results given in Appendix A. In the Appendix we also give a more detailed description of the number of events as a function of the various detector regions.

Dataset	$t\bar{t}$	$Z^0 \rightarrow \tau^+\tau^-$	WW	DY $\rightarrow \tau^+\tau^-$	Fakes	Total
$e\mu\not{E}_T$	0.011±0.003	25.6±6.5	3.9±1.0	0.5±0.2	18.4±1.4	48.5±7.6
$e\mu\not{E}_T j$	0.4±0.1	3±0.8	1.1±0.3	0.1±0.03	8.7±1.0	13.2±1.5
$e\mu\not{E}_T jj$	1.8±0.5	0.5±0.2	0.18±0.05	0.012±0.006	2.7±0.6	5.2±0.8
$e\mu\not{E}_T jjj$	0.7±0.2	0.07±0.04	0.032±0.009	0.005±0.004	0.4±0.2	1.3±0.3
$e\mu\not{E}_T + X$	3.1±0.5	29.2±4.5	5.2±0.8	0.7±0.1	30.2±1.8	68.3±5.7

Table 14: Summary of results. Note that the errors for the $e\mu\not{E}_T + X$ are smaller than the sum of the parts for those backgrounds which come from Monte Carlo as there is an uncertainty on the number of extra jets from initial and final state radiation. Also, note that while we do not use the total number of events in any calculation we have included it here for completeness. We also note that the uncertainty on this number includes the correlation between the number of events from each source.

9.1 Numbers for Sherlock

The results which are used in the Sherlock analysis are output text files and the numbers in Table 14. For the data, fakes as well as all Monte Carlo samples each event is put into the text file with a weight factor as well as the values of the variables shown in Figures 1 - 5. The weighting factor for each event is 1.0 for real data and the luminosity weighted correction factor (see Table 11) depending on the detector location of the electron and muon candidate. For the fakes, we only use the region with $L > 0.5$ and $0.1 <$Iso< 0.3 as it is the purest sample of fakes. We use a weighting factor which only depends on the electron fake rate (see Table 8).

10 Conclusions

We have described the $e\mu$ dataset, background samples and rate estimates. All numbers are updates to intermediate results of the $t\bar{t} \rightarrow e\mu$ cross section analyses using the conventional and neural net analyses. These results represent the state of the art and are intended for use by the Sherlock utility.

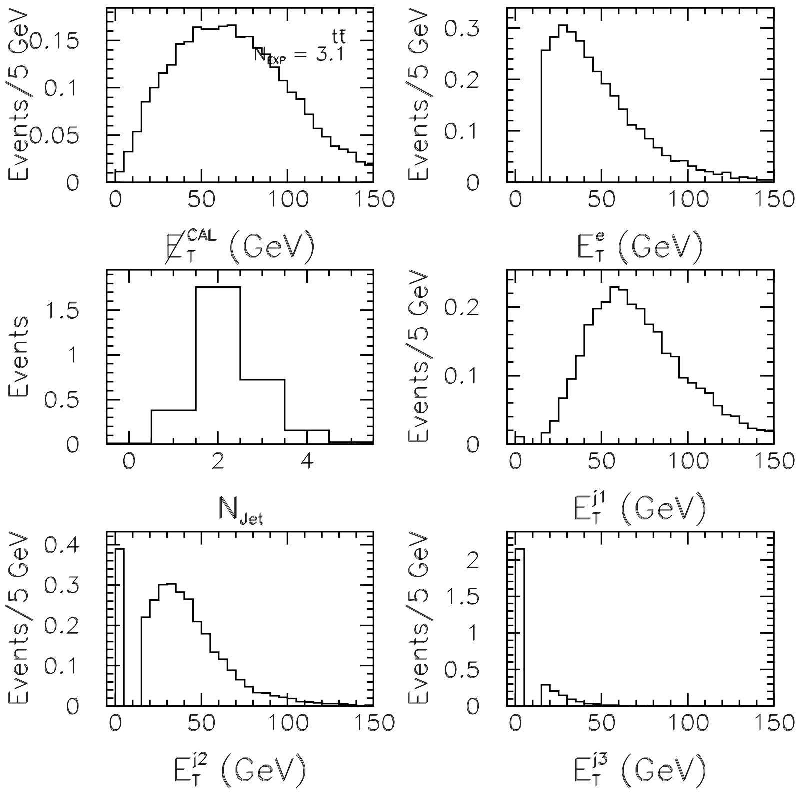

Figure 1: The variables used in Sherlock for the $t\bar{t}$ process. The plot is normalized to the number of events expected in the data.

11 Acknowledgements

The authors would greatly like to thank Boaz Klima, Meena Narain, and Mark Strovink for their input and expertise on the issues discussed in this note.

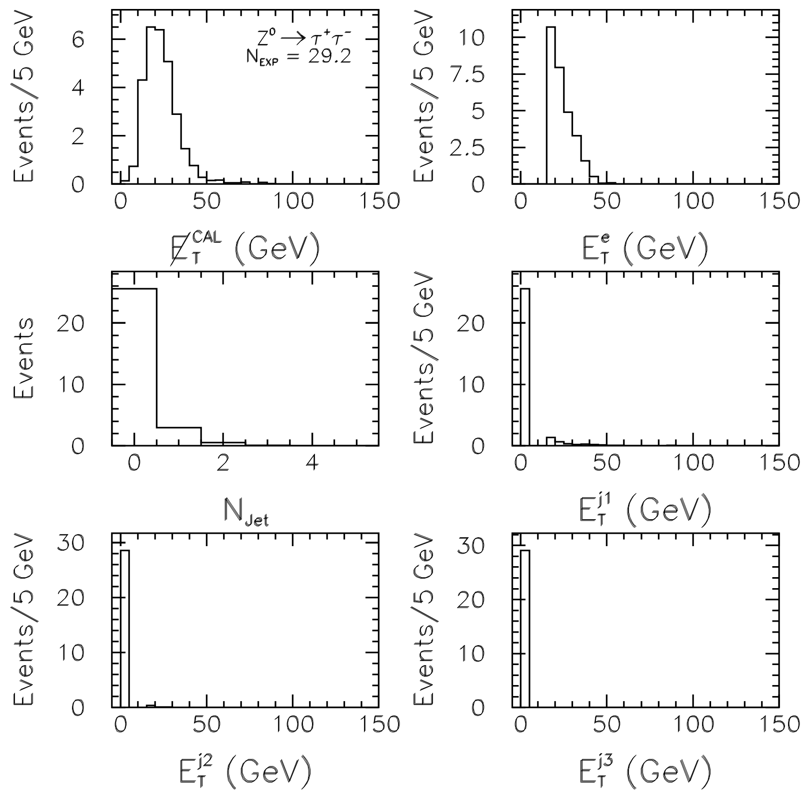

Figure 2: The variables used in Sherlock for the $Z^0 \rightarrow \tau^+\tau^-$ process. The plot is normalized to the number of events expected in the data.

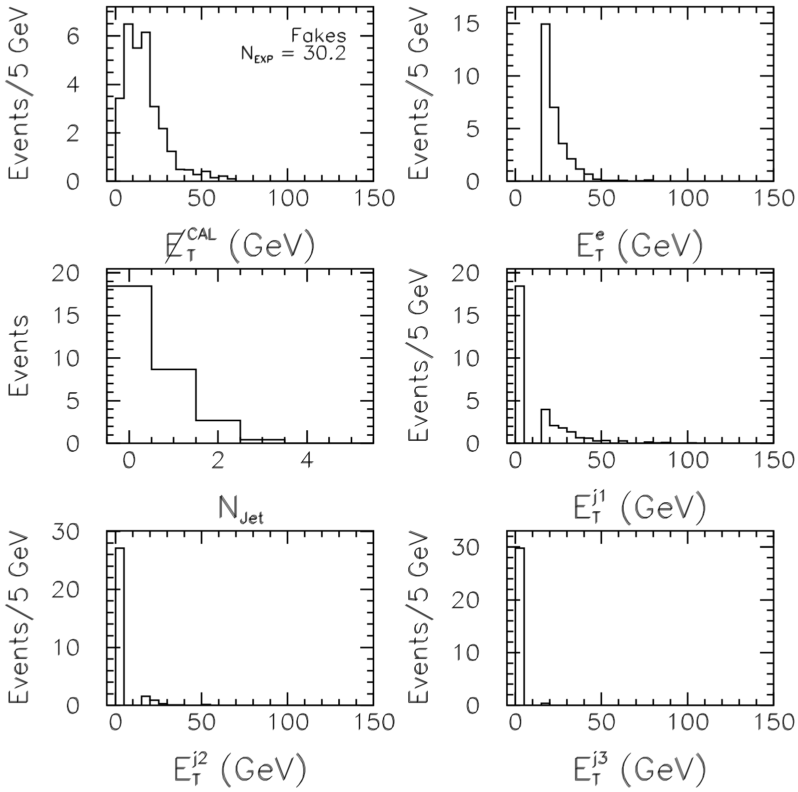

Figure 3: The variables used in Sherlock for the fakes background. The plot is normalized to the number of events expected in the data.

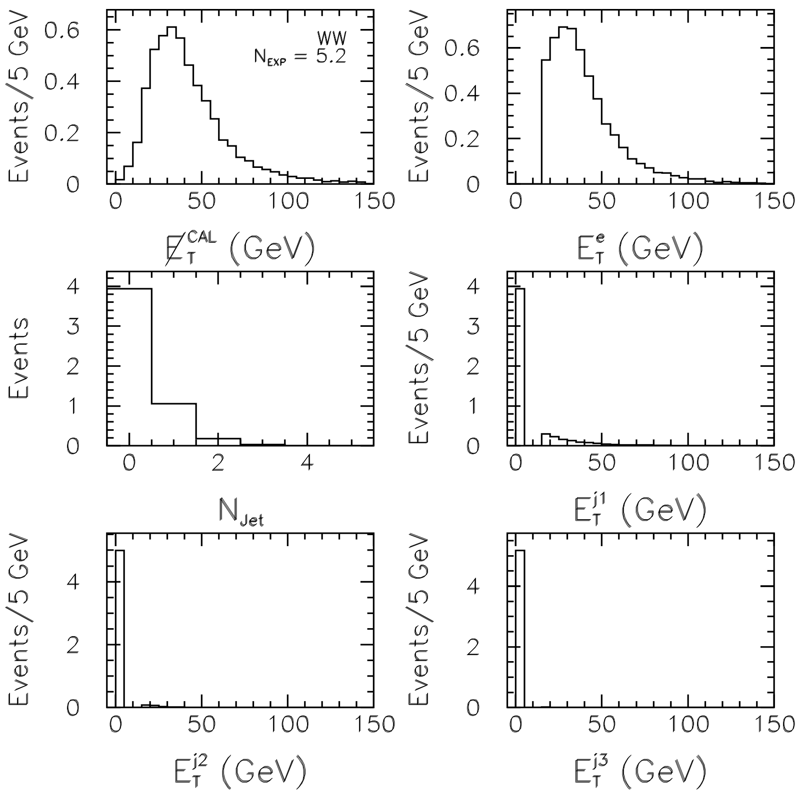

Figure 4: The variables used in Sherlock for the WW background process. The plot is normalized to the number of events expected in the data.

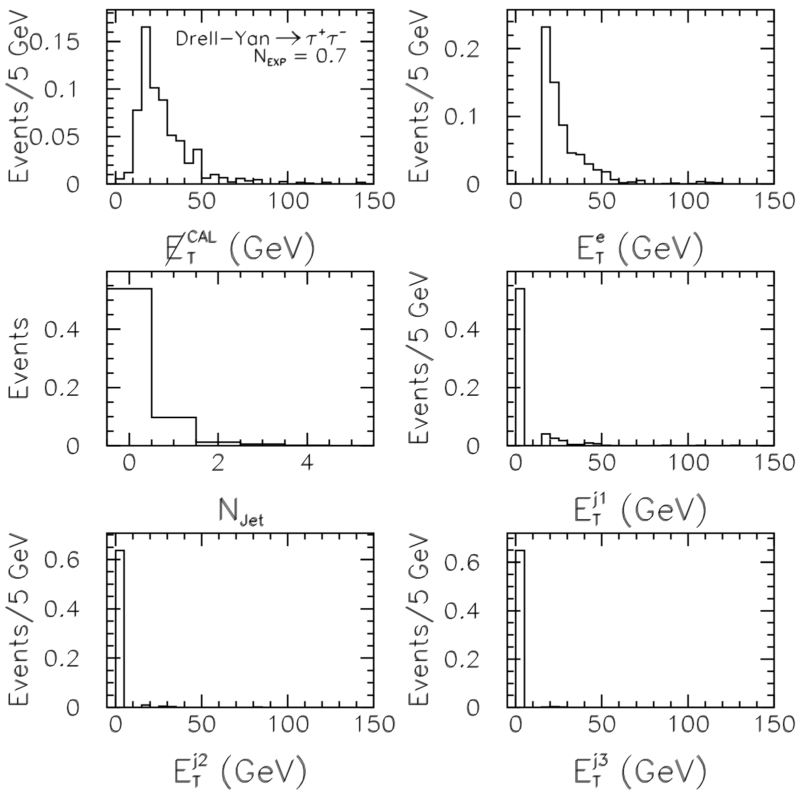

Figure 5: The variables used in Sherlock for the Drell-Yan $\rightarrow \tau^+\tau^-$ process. The plot is normalized to the number of events expected in the data.

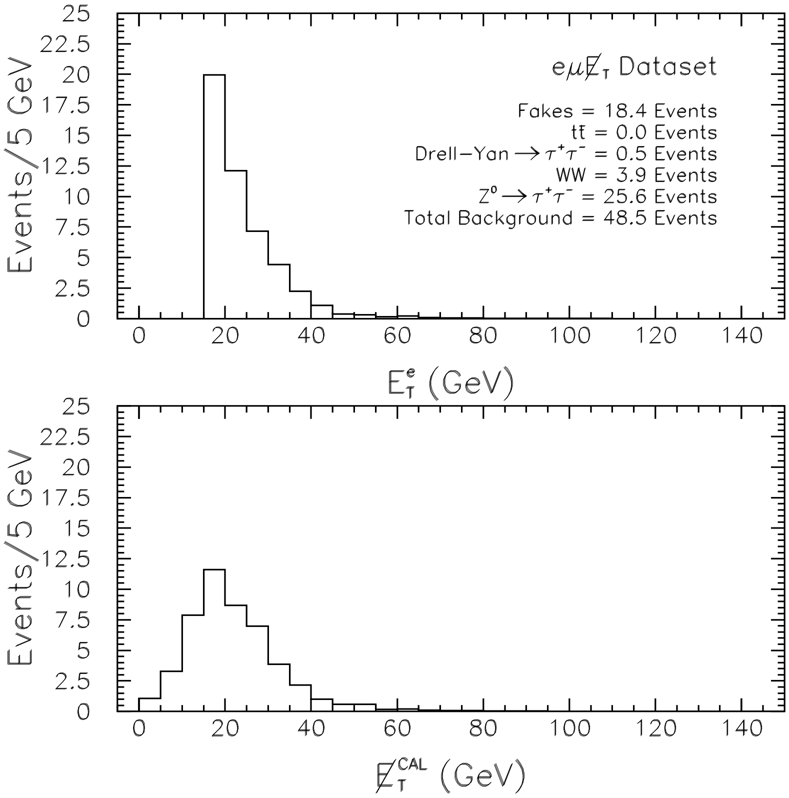

Figure 6: The variables used in Sherlock for the $e\mu\not\!\!E_T$ dataset. The background histogram is normalized to the number of events expected in the data.

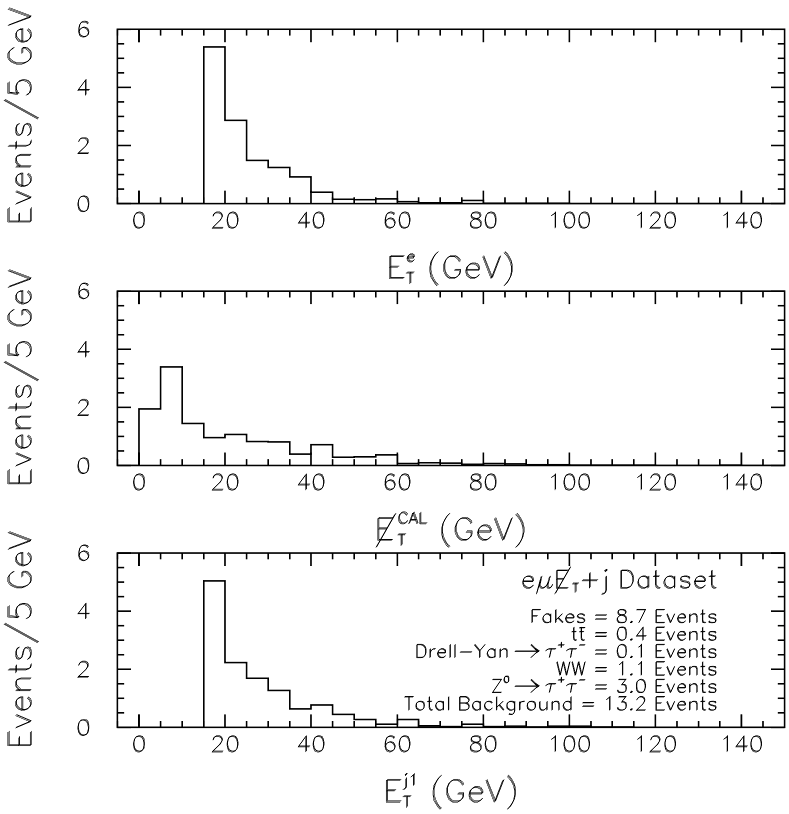

Figure 7: The variables used in Sherlock for the $e\mu\not\!\!E_{\rm T}j$ dataset. The background histogram is normalized to the number of expected events in the data.

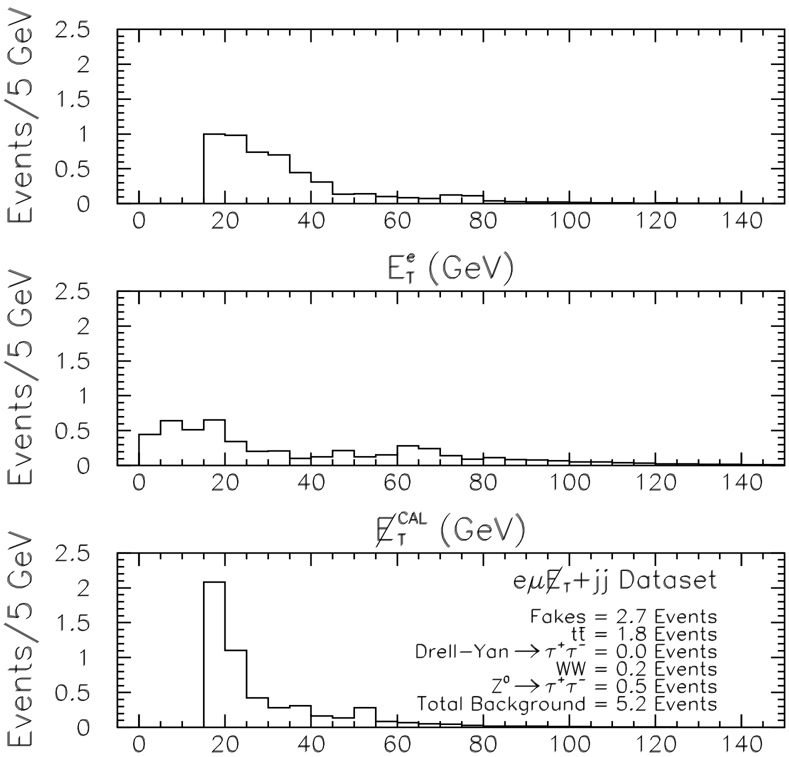

Figure 8: The variables used in Sherlock for the $e\mu\not\!\!E_T jj$ dataset. The background histogram is normalized to the number of expected events in the data.

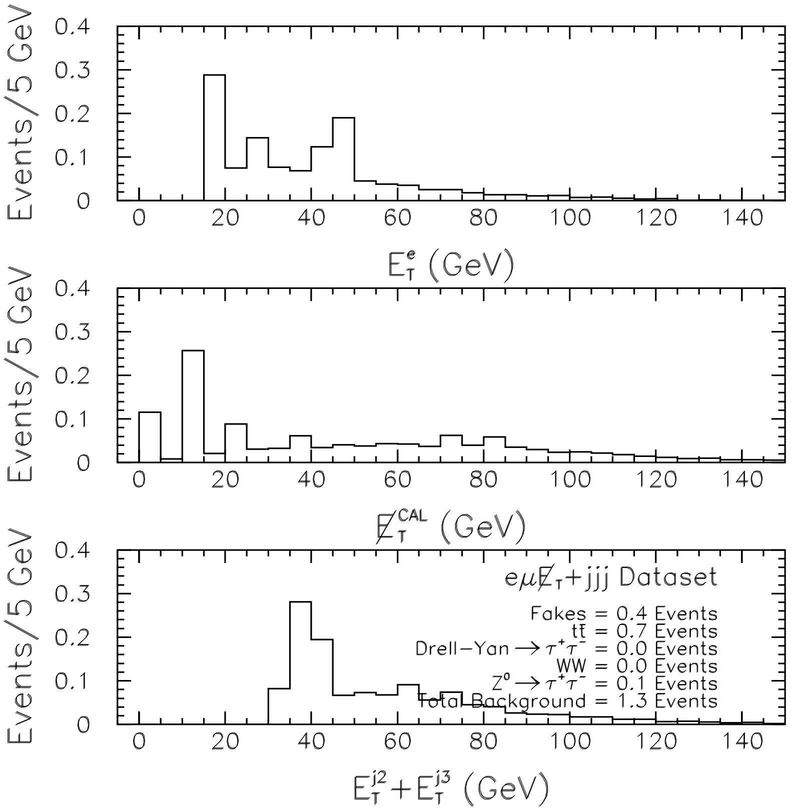

Figure 9: The variables used in Sherlock for the $e\mu\not{E}_T jjj$ dataset. The background histogram is normalized to the number of expected events in the data.

References

[1] B. Knuteson and D. Toback, "The Sherlock Search Strategy: A quasi-model-independent prescription for searching for new physics" DØ Note 3701, Jan. 2000.

[2] B. Knuteson and D. Toback, "Sherlock Search Strategy Supplement: DØ Run I" DØ Note 3715, Jan. 2000.

[3] B. Knuteson and D. Toback, "Sherlock: A multivariate, quasi-model-independent algorithm for finding anomalous events" DØ Note 3686, Jan. 2000.

[4] B. Knuteson and D. Toback, "Results of the Sherlock Analysis with the $e\mu$ Dataset"DØ Note 3718, Jan. 2000.

[5] J. Cochran, "$e\mu$ truths in run 1" February 4, 1997, unposted.

[6] H. Singh, "A Measurement of $t\bar{t}$ Production Cross Section in $p\bar{p}$ Collisions at $\sqrt{s}$ = 1.8 TeV Using Neural Networks, Ph.D. Dissertation, University of California at Riverside, 1999.

[7] http://www-d0.fnal.gov/online/GM/trigger_lists.html

[8] S. J. Wimpenny, B. Kehoe, M. Narain, J. Thompson, and P. Quintas, Bad Run RCP File located at /d0sgiucr2/exports/data0/jc/shappy/topdilep2/bad_run.rcp (Version 3.0), 1996. To be complete, we note that we have NOT used the newer Bantly/Kehoe list to be consistant with Ref. [6]. For the same reasons we do not use top_event_cleanup to remove pathological events.

[9] MICRO_BLANK rejects any event within a 400 ms window following main ring injection, and MRBS_LOSS rejects events when Main Ring bunches occur within 1.6 μs window of the Tevatron beam crossing. The losses for MICRO_BLANK and MRBS_LOSS are 17% and 8% respectively. For more details see J. Butler, "Main Ring Deadtime," DØ Internal Note 1682 (1993)

[10] see d0ucr7$dka0:[prj39.top4.dilepton.top dilep.code.released]top dilep.pbd. S. J. Wimpenny, "The TOP-LEPTONS Analysis Package," DØ Note 2144 (1995)

[11] R. Kehoe, "Resolution Bias in Jet Response Measurement" DØ Note 2052, (1994) R. Kehoe and R. Astur, "Resoution Bias in Jet Response Measurement" DØ Note 2908, (1996)

[12] For those of you which are vaguely familiar with the top dilepton analysis, this is the standalone fortran program (/new.proj13/dilepton/emu/nov96/test1b_nov96.for) which is run over the ntuples and produces ascii text files (and hst4 files). These ascii text files provide input for 'the spreadsheet' which calculates the final efficiencies, backgrounds, etc. and is linked directly to the main top group spreadsheet (aka Herb's spreadsheet). It was this analysis package that was customized by Harpreet in Ref. [6].

[13] R. Brun and F. Carminati, CERN program library long writeup, W5013, (1993)

[14] W. Dharmaratna, R. Raja and C. Stewart, "DØ shower library", DØ Note 1730 (1993)

[15] T. Diehl and P. Quintas, "The MU-SMEAR package documentation", (1994)

[16] Particle Data Book, Review of Particle Physics, Physics Review D54, (1996)

[17] E. Laenen, J. Smith and W. van Neerven, Phys. Lett. B **321**, 254(1994).

[18] G. Gomez, "Extraction of the Width of the W Boson from a Measurement of the Ratio of the W and Z Cross Sections", Ph.D. Dissertation, University of Maryland, 1999.

[19] J. Ohnemus, PRD **44** 1403 (1991)

[20] J. McKinley, Ph.D. Thesis "A Measurement of the Inclusive Drell-Yan e+e- Cross Section in the Invariant Mass Range 30-60 GeV/c^2 from $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV", Michigan State University, (1996).

[21] S. Abachi *et al.*, "Top Quark Search with the DØ 1992-1993 Data Sample" Phys. Rev. D **52** 4877 (1995). V. Bhatnagar and J. B. Singh, DØ Note 1858 (1993). D. Hedin DØ Note 1738 (1993). These notes show that the rate at which punch through hadrons make it through the large number of interaction lengths in the calorimeters and muon toriod system is tiny.

[22] M. Narain and U. Heintz, "A likelihood Test for Electron ID", DØ Note 2386 (1994).

[23] This is the Bob Kehoe reference for the fake rates.

[24] J. Linnemann, "Triggering the DØ Experiment", Proceedings of DPF'92,(1992)

[25] This is the NOISY reference.

[26] S. J. Wimpenny "Muon efficiency corrections for single top analyses", DØ Note 3282 (1997). M. Mason and S. J. Wimpenny, "Studies of Mu-Smear corrections factors for the Run 1b/c single top analysis", DØ Note 3281 (1997)

A Appendix: Comparison of Results with the Neural Net Analysis

As a check, we have compared all of the data and background estimation results with the neural net [6] $t\bar{t} \rightarrow e\mu$ cross section analysis. In this section we enumerate the results for the various samples in Tables 15 - 20. We have included the results for $\not{E}_T^{\rm cal} > 15$ GeV and $\rm N_{Jet} \geq 2$ cuts. Again, the "Expected in Data" column uses Equation 5 with values of Lum $= 108.3, C^{\rm Mult\ Int} = 0.905$, $\sigma_{\rm visible}$ and $\rm N_{Total}$ from Table 3, the weighted corrections from Table 11 and the number of events in each detector region for a given sample in the first 4 columns of the table. Figures 10-14 show the distributions of the variables used in the neural net with each distribution weighted to be equal to the number of events expected in the data. Figures 15 through 20 show the comparison between the data and the background predictions for the various distributions.

We note that there are two differences between this analysis and the NN analysis. We list them here.

A.1 Fakes

The way the fakes were calculated in Ref. [6] is incorrect. It approximates the full formula in Appendix D by $A_F = A * R * \epsilon$. As evidenced from the full formula this is poor approximation and using the approximation reduces the background by about a factor of two. However, since this background is a tiny fraction of the final NN analysis background, and does not contribute significantly to the error, it does not affect the robustness of that result. This correction raises the total number of predicted background events.

A.2 Multiple Interaction Corrections

In Ref. [6] there is a multiple interaction correction. This correction is used to calculate the cross section from the observed number of events minus the background. We have instead used this number to correct the number of predicted top events. We have in addition used this number to correct the number of events from the Monte Carlo samples by assuming that the correction is the same for them. This correction lowers the total number of predicted background events.

Cut	CCCF	CCEF	ECCF	ECEF	Total
Isolated $e\mu$	29	4	22	3	58
$\not{E}_T^{\rm cal}$> 15 GeV	24	3	14	3	44
$\rm N_{Jet} \geq 2$	3	0	2	1	6

Table 15: Results table for data.

Cut	CCCF	CCEF	ECCF	ECEF	Total	Expected in Data
Isolated $e\mu$	9725	2239	1685	395	14044	3.06
$\not{E}_T^{\rm cal}$> 15 GeV	9416	2165	1627	385	13593	2.96
$\rm N_{Jet} \geq 2$	8182	1904	1438	348	11872	2.58

Table 16: Results for $t\bar{t}$.

Cut	CCCF	CCEF	ECCF	ECEF	Total	Expected in Data
Isolated $e\mu$	834	280	287	140	1541	29.2
$\not{E}_T^{\rm cal}$> 15 GeV	702	235	208	121	1266	24.0
$\rm N_{Jet} \geq 2$	18	5	3	1	27	0.54

Table 17: Results Table for $Z^0 \to \tau^+\tau^-$.

Cut	CCCF	CCEF	ECCF	ECEF	Total	Expected in Data
Isolated $e\mu$	8770	2591	2513	870	14744	5.21
$\not\!\!E_T^{cal}$> 15 GeV	8352	2469	2393	826	14040	4.96
N_{Jet} ≥2	359	88	100	26	573	0.21

Table 18: Results table for WW.

Cut	CCCF	CCEF	ECCF	ECEF	Total	Expected in Data
Isolated $e\mu$	243	94	74	44	455	0.66
$\not\!\!E_T^{cal}$> 15 GeV	207	85	61	38	391	0.56
N_{Jet} ≥2	8	2	1	3	14	0.020

Table 19: Results table for Drell-Yan → $\tau^+\tau^-$.

Cut	$L < 0.5$ Iso<0.1 (Region A)	$L > 0.5$ Iso<0.1 (Region B)	$L < 0.5$ 0.1<Iso<0.3 (Region C)	$L > 0.5$ 0.1<Iso<0.3 (Region D)	Total	Expected in Data
Central EM Clusters						
Isolated $e\mu$	32	111	11	208	362	15.8
$\not\!\!E_T^{cal}$> 15 GeV	26	69	10	123	228	9.65
N_{Jet} ≥2	3	3	0	11	17	0.66
EndCap EM Clusters						
Isolated $e\mu$	25	74	5	92	196	14.3
$\not\!\!E_T^{cal}$> 15 GeV	17	41	2	31	91	5.9
N_{Jet} ≥2	3	1	1	6	11	0.58
All EM Clusters						
Isolated $e\mu$	57	185	16	300	558	30.2
$\not\!\!E_T^{cal}$> 15 GeV	43	110	12	154	319	15.5
N_{Jet} ≥2	6	4	1	17	28	1.24

Table 20: Results table for fakes. We note that these numbers are taken directly from the data, but with the additional requirement that $\chi^2_{\mathrm{CAFIX\ v5.0}} < 300$. This is why in Region A for CC+EC there are 57 rather than 58 events.

Cut	$t\bar{t}$	$Z^0 \to \tau^+\tau^-$	WW	DY $\to \tau^+\tau^-$	Fakes	Total	Data
Isolated $e\mu$	3.06	29	5.21	0.66	30	68	58
$\not\!\!E_T^{cal}$> 15 GeV	2.96	24	4.96	0.56	15.5	48	44
N_{Jet} ≥2	2.58	0.54	0.21	0.020	1.2	4.6	6

Table 21: Summary of results.

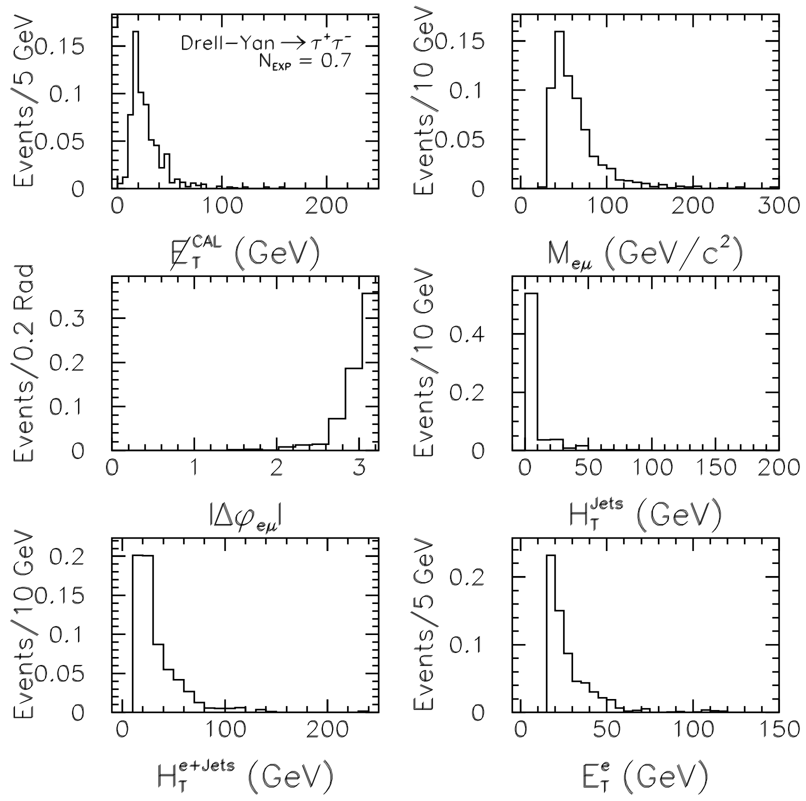

Figure 10: The variables used in Sherlock for the Drell-Yan $\rightarrow \tau^+\tau^-$ process. The plot is normalized to the number of events expected in the data.

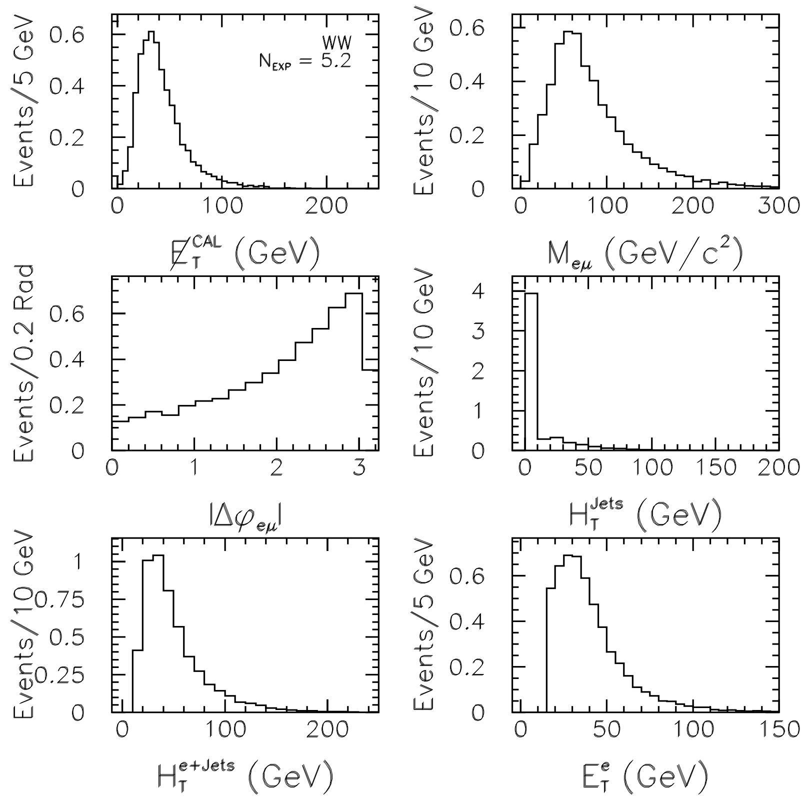

Figure 11: The variables used in Sherlock for the WW process. The plot is normalized to the number of events expected in the data.

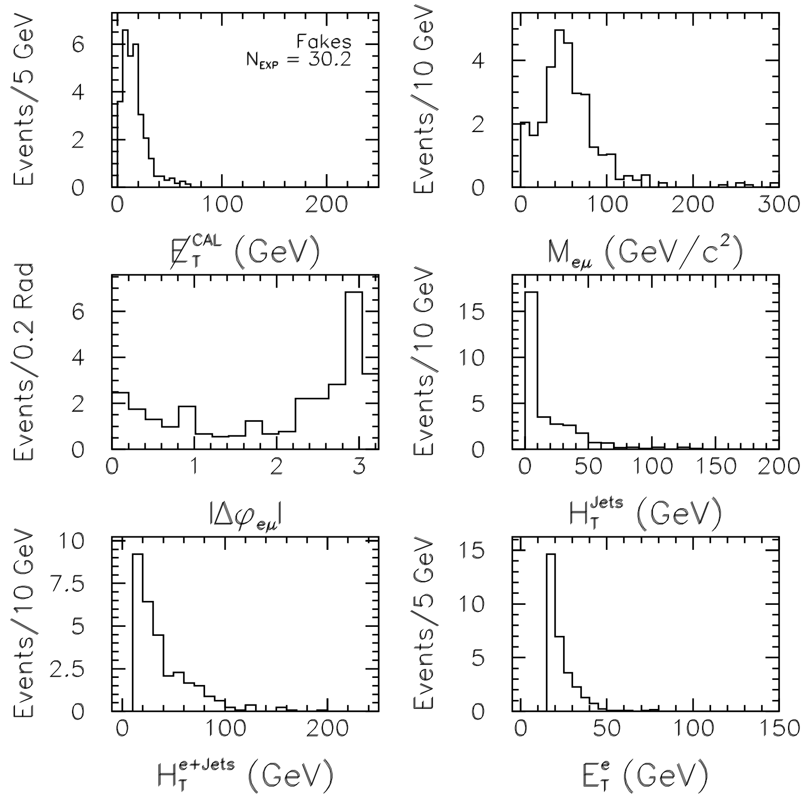

Figure 12: The variables used in Sherlock for the fakes. The plot is normalized to the number of events expected in the data.

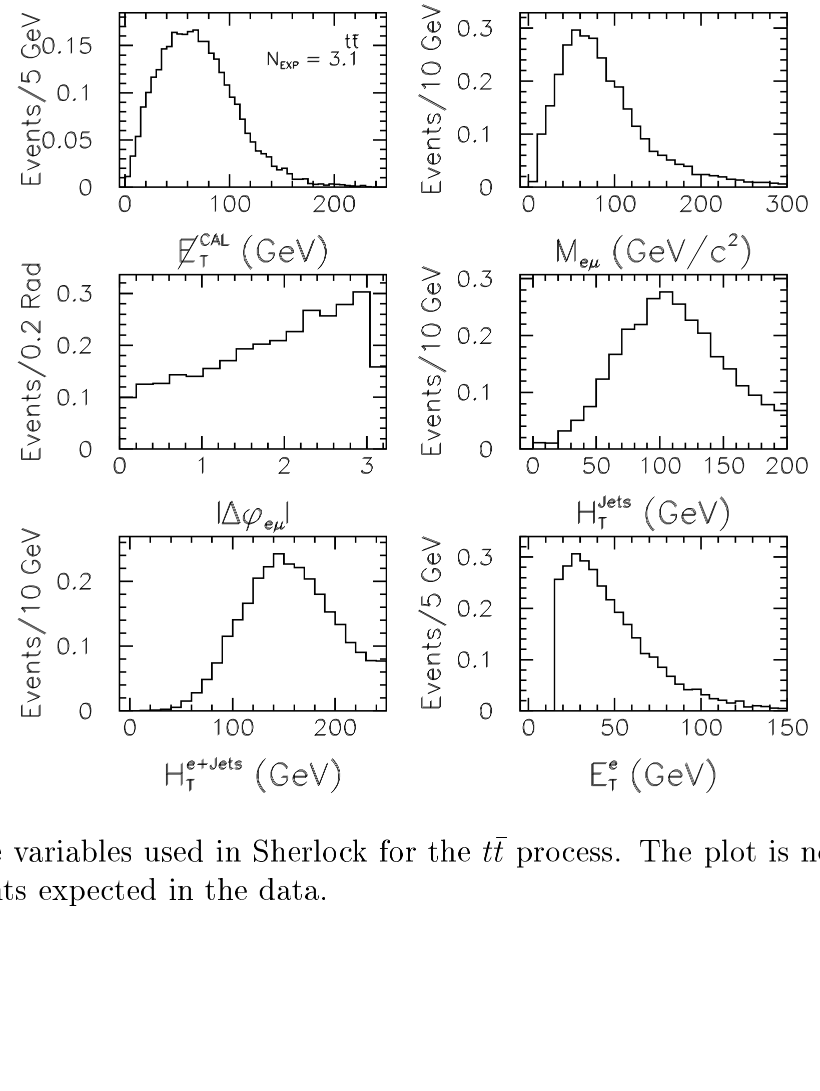

Figure 13: The variables used in Sherlock for the $t\bar{t}$ process. The plot is normalized to the number of events expected in the data.

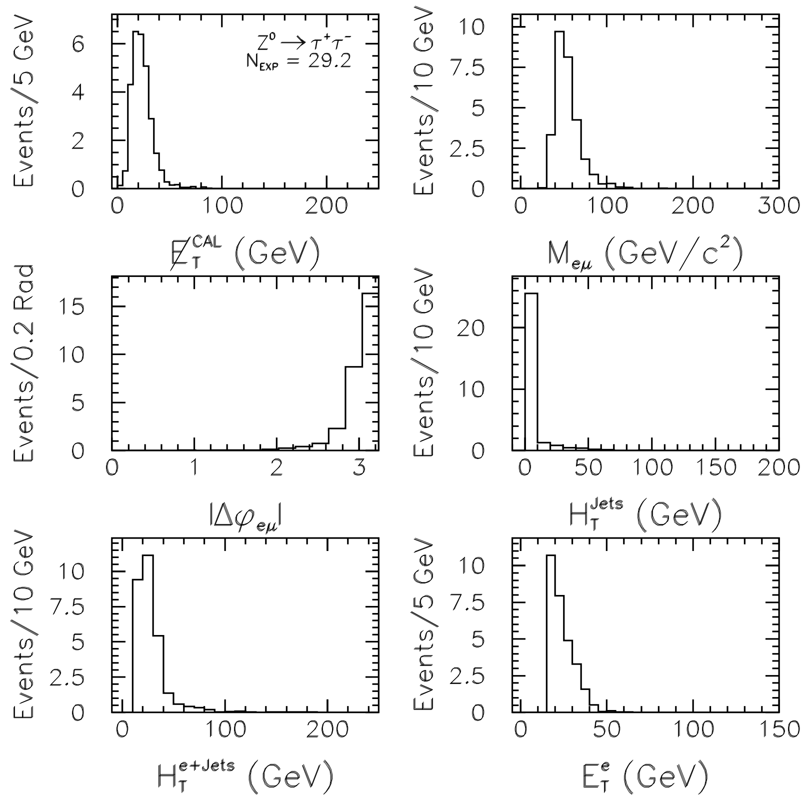

Figure 14: The variables used in Sherlock for the $Z^0 \rightarrow \tau^+\tau^-$ process. The plot is normalized to the number of events expected in the data.

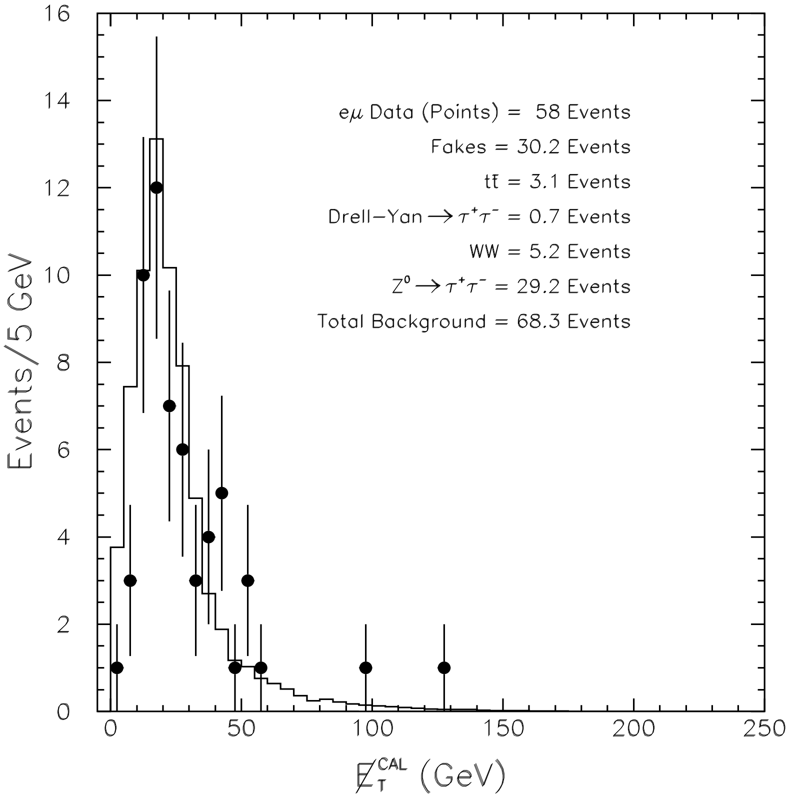

Figure 15: A comparison of the sum of the backgrounds (Drell-Yan $\rightarrow \tau^+\tau^-$, WW, fakes, $t\bar{t}$ and $Z^0 \rightarrow \tau^+\tau^-$) and the data.

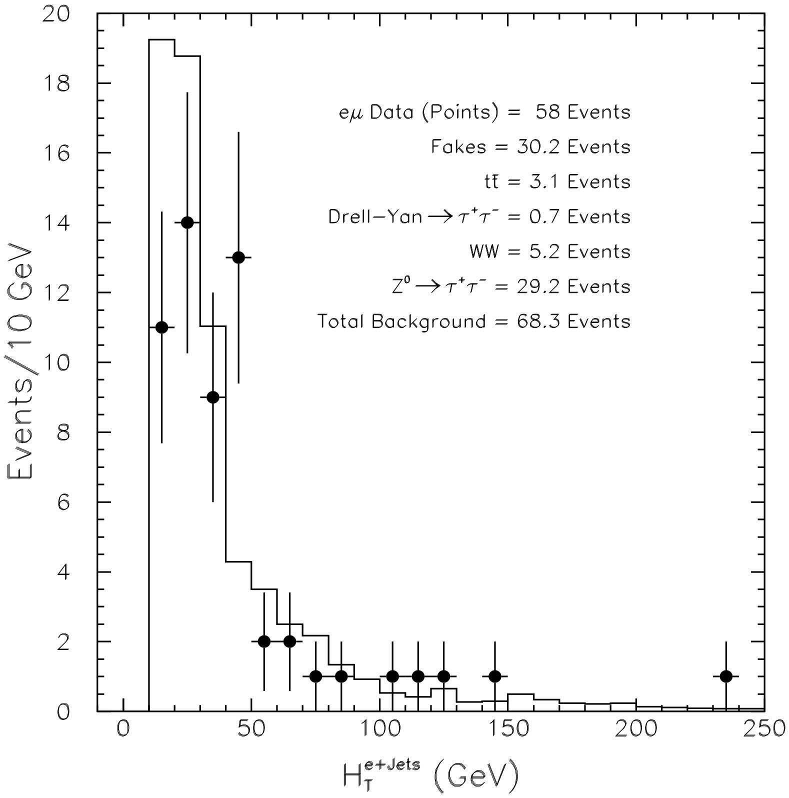

Figure 16: A comparison of the sum of the backgrounds (Drell-Yan $\rightarrow \tau^+\tau^-$, WW, fakes, $t\bar{t}$ and $Z^0 \rightarrow \tau^+\tau^-$) and the data.

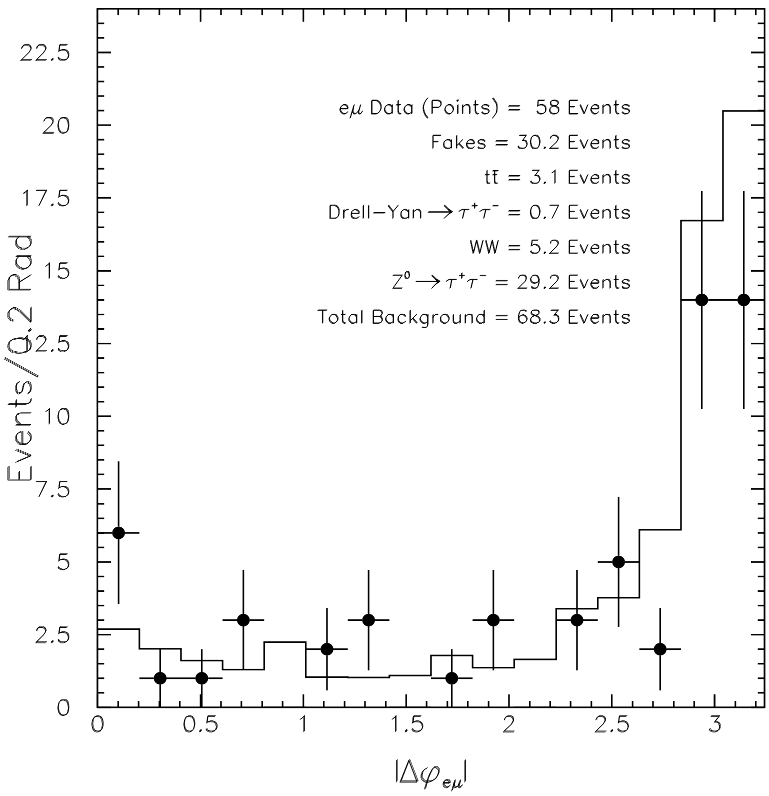

Figure 17: A comparison of the sum of the backgrounds (Drell-Yan $\rightarrow \tau^+\tau^-$, WW, fakes, $t\bar{t}$ and $Z^0 \rightarrow \tau^+\tau^-$) and the data.

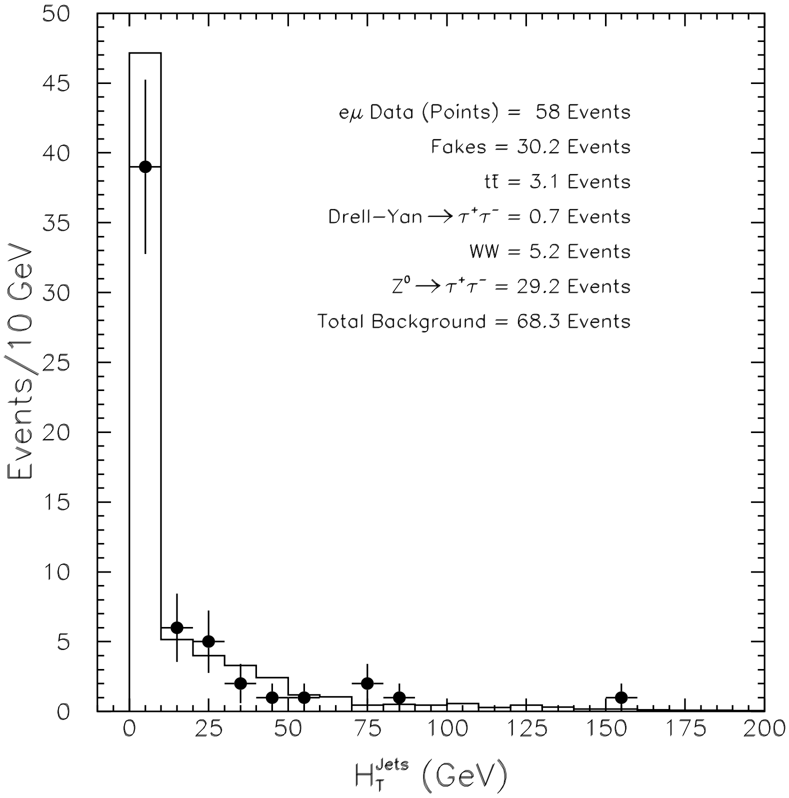

Figure 18: A comparison of the sum of the backgrounds (Drell-Yan $\rightarrow \tau^+\tau^-$, WW, fakes, $t\bar{t}$ and $Z^0 \rightarrow \tau^+\tau^-$) and the data.

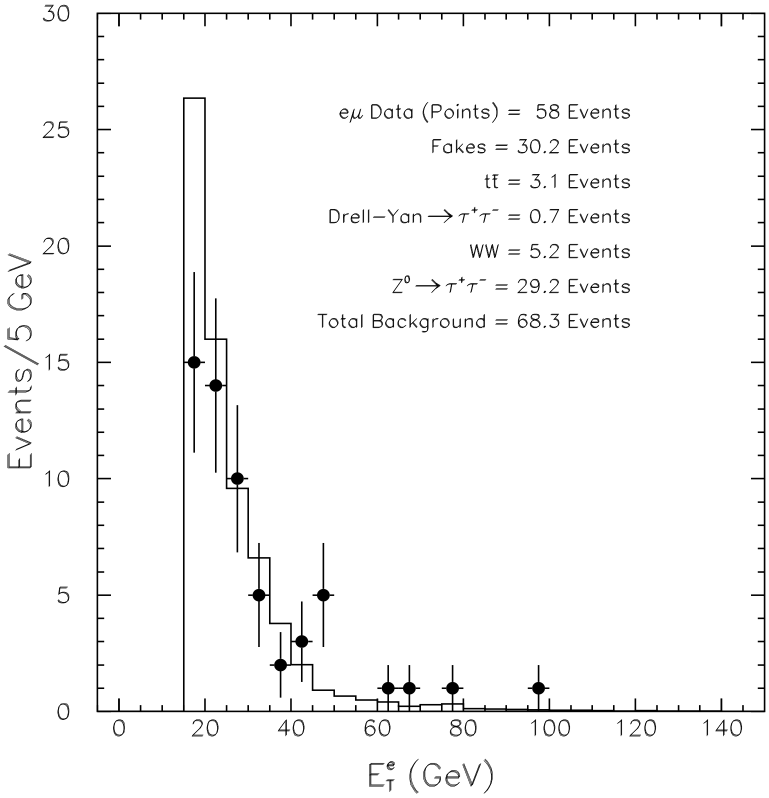

Figure 19: A comparison of the sum of the backgrounds (Drell-Yan $\to \tau^+\tau^-$, WW, fakes, $t\bar{t}$ and $Z^0 \to \tau^+\tau^-$) and the data.

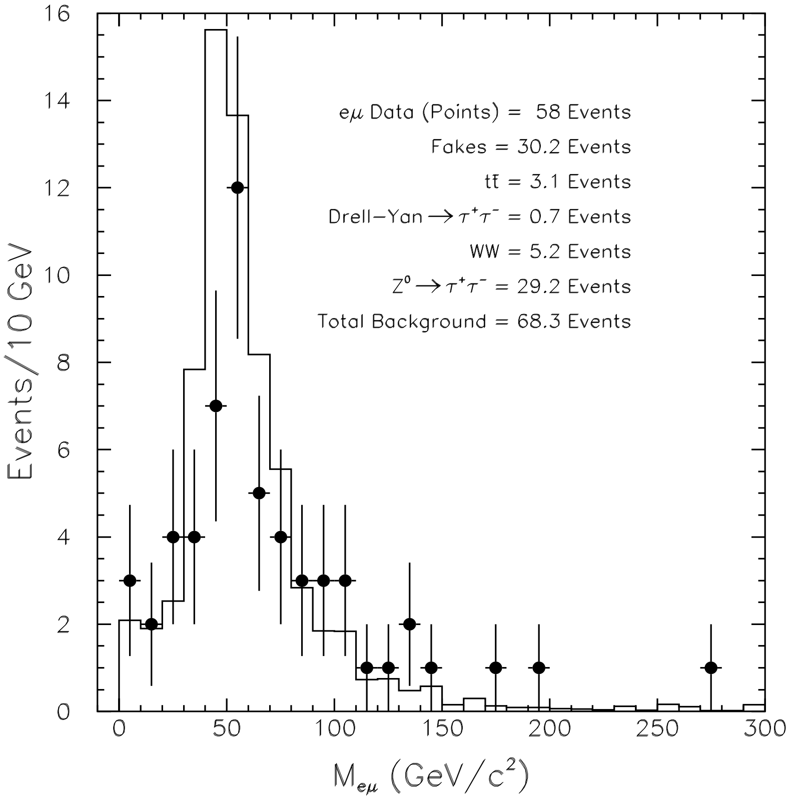

Figure 20: A comparison of the sum of the backgrounds (Drell-Yan $\to \tau^+\tau^-$, WW, fakes, $t\bar{t}$ and $Z^0 \to \tau^+\tau^-$) and the data.

B Appendix: Monte Carlo Muon Correction Factors

There are three correction factors which must be applied to Monte Carlo data for muons to create the numbers in Table 9. The three are collected into a single correction using the following relation:

$$C^{\mu ID} = C_{\mu\text{smear}} \times C_{\phi} \times C_{\text{PMUO}} \tag{7}$$

where C_{PMUO} is the efficiency correction factor for a muon creating PMUO bank in the Monte Carlo, $C_{\mu\text{smear}}$ is the track finding efficiency correction factor in the WAMUS chambers (from the μsmear utility [15]) and C_{ϕ} is a ϕ-dependent correction, "phi hole", which is used to correct the Monte Carlo CF tracking efficiency in the region of the main ring beam pipe [26].

The efficiency corrections are complicated by the fact that the data are broken into three types: Run 1A, Prezap and Postzap. For simplicity the different correction factors for Pre and Postzap are combined and the MC samples are generated in a 50/50 combination of Pre/Postzap μsmearing. To combine Prezap and Postzap data (the above factors being the only differences) a luminosity weighted combination of the two is needed. As given in Section 3.2 the Pre and Post zap fractions are 0.54± 0.04 and 0.46± 0.04 respectively. These are combined using

$$(\text{Prezap}\#) \times (0.54 \pm 0.04) + (\text{Postzap}\#) \times (0.46 \pm 0.04). \tag{8}$$

The results are Taken from Ref. [5] and summarized in Table 22.

	1A CF Prezap	1B CF Prezap	1B CF Postzap	1B EF Postzap
C_{PMUO}	0.941 ± 0.015	0.941 ± 0.015	0.941 ± 0.015	0.911 ± 0.022
C_{ϕ}	0.95 ± 0.05	0.90 ± 0.05	na	na
$C_{\mu\text{smear}}$	0.97 ± 0.06	1.00 ± 0.05	1.00 ± 0.05	1.20 ± 0.10
$C^{\mu ID}_{\text{Unweighted}}$	0.865 ± 0.073	0.85 ± 0.06	0.94 ± 0.05	1.09 ± 0.10

	1A CF Prezap	1BC CF	1B EF
$C^{\mu ID}_{\text{Weighted}}$	0.865 ± 0.073	0.890 ± 0.062	0.504 ± 0.059

Table 22: The muon correction factors for files which are combined Pre+Post Zap (50/50) μsmeared data. The final row reflects the luminosity weighted Pre/Postzap correction factors and are the same numbers as in Table 9. Note that while there are no Run 1A or Prezap EF muons accepted, this is reflected naturally in the correction factors.

C Appendix: Additional Monte Carlo Muon Correction Factors

As described in Appendix B there are a number of correction factors which must be applied to Monte Carlo data for muons. In this Appendix we, for completeness, give the correction factors for other types of Monte Carlo samples. Specifically ‘none’ (no μsmear), ‘Post’ (Postzap μsmear) and ‘both’ which is the combined (50/50) Pre+Post zap μsmear. These are described in Tables 23 through 25 and summarized in Table 26.

	1A CF	Prezap 1B CF	Postzap 1B CF	Postzap 1B EF
pmuo	0.941 ± 0.015	0.941 ± 0.015	0.941 ± 0.015	0.911 ± 0.022
ϕ hole	0.95 ± 0.05	0.90 ± 0.05	na	na
μsmear	0.90 ± 0.05	0.93 ± 0.03	0.93 ± 0.03	0.84 ± 0.05
	0.80 ± 0.06	0.79 ± 0.05	0.88 ± 0.03	0.77 ± 0.05

Table 23: The muon correction factors for files which have not been through μsmear.

	1A CF Prezap	1B CF Postzap	1B CF Postzap	1B EF
pmuo	0.941 ± 0.015	0.941 ± 0.015	0.941 ± 0.015	0.911± 0.022
ϕ hole	0.95 ± 0.05	0.90 ± 0.05	na	na
μsmear	0.97 ± 0.06	1.00 ± 0.05	1.00 ± 0.03	1.00 ± 0.06
	0.87 ± 0.07	0.85 ± 0.06	0.94 ± 0.03	0.91 ± 0.06

Table 24: The muon correction factors for files which have been through Postzap μsmear.

	1A CF Prezap	1B CF Postzap	1B CF Postzap	1B EF
pmuo	0.941 ± 0.015	0.941 ± 0.015	0.941 ± 0.015	0.911 ± 0.022
ϕ hole	0.95 ± 0.05	0.90 ± 0.05	na	na
μsmear	0.97 ± 0.06	1.00 ± 0.05	1.00 ± 0.05	1.20 ± 0.10
	0.87 ± 0.07	0.85 ± 0.06	0.94 ± 0.05	1.09 ± 0.10

Table 25: The muon correction factors for files which are combined Pre+Post Zap (50/50) μsmeared data.

	Run 1A	
data set	CF	EF
Post	0.865 ± 0.073	0.0 ± 0.0
none	0.805 ± 0.063	0.0 ± 0.0
both	0.865 ± 0.073	0.0 ± 0.0
	Run 1B Prezap	
Post	0.85 ± 0.06	0.0 ± 0.0
none	0.79 ± 0.05	0.0 ± 0.0
both	0.85 ± 0.06	0.0 ± 0.0
	Run 1BC Postzap	
Post	0.94 ± 0.03	0.91 ± 0.06
none	0.88 ± 0.03	0.77 ± 0.05
both	0.94 ± 0.05	1.09 ± 0.10

Table 26: Summary of the muon correction factors for files which have been through various stages of μsmear.

D Appendix: Fake Rate Calculations

This Appendix details the fake rate calculation and it's determination. This calculation makes use of the electron efficiency and fake rate numbers of Ref. [23]. We begin by noting that the fake 'rate' is actually the rate at which jets which create PELC banks with $\chi^2 < 300$ and Iso<0.3 ALSO pass the electron ID cuts. This number is denoted R and is given in Table 8 as a function of the calorimeter section. The sample, as described in Section 6.2 has all events from the data which pass all the $e\mu$ identification and isolation requirements except for the electron ID. We remove the L_5 and Iso< 0.1 cuts and replace them with the $\chi^2 < 300$ and Iso<0.3 cuts. This sample can be broken down into four regions which we define A-D:

- A: Number of Events with $L < 0.5$, Iso<0.1 (Signal Region) [3]
- B: Number of Events with $L > 0.5$, Iso<0.1
- C: Number of Events with $L < 0.5$, 0.1<Iso<0.3
- D: Number of Events with $L > 0.5$, 0.1<Iso<0.3

In addition we define:

- E: Total number of electrons in Regions A thru D
- F: Total number of jets in Regions A thru D
- A_E: Electrons in Region A
- A_F: Jets in Region A
- ϵ: Electron efficiency
  (rate at which electrons land in region A)
- R: Electron Fake 'Rate'
  (rate at which jets passing PELC/χ^2/Iso cuts land in region A)

We use the relations:

$$
\begin{aligned}
A + B + C + D &= E + F \\
\epsilon F &= A_E \\
RF &= A_F \\
A &= A_E + A_F
\end{aligned}
$$

Solving for A_F we find:

[3] We note that this is not strictly correct as we have the additional requirement of $\chi^2_{CAFIX5.0}$ <300.

$$
\begin{aligned}
A_F &= A - A_E \\
&= A - \epsilon E \\
&= A - \epsilon(A + B + C + D - F) \\
&= A - \epsilon(A + B + C + D - \frac{A_F}{R}) \\
\Rightarrow A_F(1 - \frac{\epsilon}{R}) &= A - \epsilon(A + B + C + D) \\
\Rightarrow A_F &= R[\frac{A - \epsilon(A + B + C + D)}{R - \epsilon}]
\end{aligned}
$$

E Jet Counting Corrections from the MC

In this section we quickly describe the jet counting correction for the Monte Carlo samples. As stated in Section 8 the number of jets produced in association with $Z^0 \rightarrow e^+e^-$ from data and $Z^0 \rightarrow \tau^+\tau^-$ from ISAJET are the same within errors. We note that the ratio of two jets to no jet requirement is 0.84±0.18 which is completely consistent within errors. We take no correction to the number of jets, and an error of 20% which we add in quadrature with the overall uncertainty with the number of predicted events from Monte Carlo samples. We note that while we have only done this procedure for ISAJET and Z production, we assume the same results for the WW, Drell-Yan $\rightarrow \tau^+\tau^-$, and $t\bar{t}$ samples.

Sample	$Z^0 \rightarrow e^+e^-$ Data Events	$Z^0 \rightarrow \tau^+\tau^-$ ISAJET Events
No Cuts	3517	1541
2 Jets, E_T>15 GeV	87	32
Ratio	0.025±0.003	0.021±0.004
Ratio of Ratios	0.84±0.18	

Table 27: The ratio is the rate at which two jets are produced relative to the full sample.