CDF/DOC/ANAL/9736
Version 4.0 May 16, 2009

Observation of Diboson Production in $\not\!\!E_T$ + jets channel

G. Flanagan[1] J. Freeman[2], S. Pronko[2], V. Rusu[2]

1 *Purdue University*
2 *Fermilab*

Abstract

Diboson production has been observed at the Tevatron in the lepton channels through the leptonic decays of the electroweak gauge bosons. Doing the same thing with jets is much more difficult due to the large background from QCD multi-jets and V+jets (V=Z,W). Here we attempt to measure the diboson cross-section using events with large $\not\!\!E_T$ and two jets above 25 GeV. Due to limited energy resolution we cannot distinguish between WW, WZ and ZZ events so what we measure is really a sum of all these processes in our selection window. As no cut on number of leptons in the event is performed we are also sensitive to lepton decays of the gauge bosons. The QCD contribution, which is large in this channel, is heavily suppressed through novel algorithms related to $\not\!\!E_T$ significance as described in text. We extract the signal from the background using the invariant mass distribution of the two jets in the event. The extraction of the signal does not use the theoretical calculation of the V+jets integral cross section and its invariant mass shape is cross checked with γ+jets events from the data, hence considerably reducing the systematic uncertainty on the shape of this main background. Using this technique, we find $1,516 \pm 239(stat) \pm 144(sys)$ signal events in our data sample of 44,910 events, corresponding to a measured cross section of $18.0 \pm 2.8(stat) \pm 2.4(syst) \pm 1.1(lumi)$ pb.

1 Event Selection

We start with events that have two cone 0.4 jets corrected to L7 and $\not\!\!E_T > 60\,\text{GeV}$. The jets are required to have $E_T > 25\,\text{GeV}$ and $|\eta| < 2.0$. The jet EM fraction of all jets with raw E_T>10 GeV has to be less than 0.9 such that electrons are not considered in the jet list. Additionally, no other jets above 20 GeV and $|\eta| < 2.0$ can be present in the event[1]. This is mostly to remove events where the additional contribution from ISR/FSR can make the dijet mass calculation ambiguous. We also require two leading jets to be outside of a 0.2 cone in ΔR from the most energetic identified lepton. Our lepton identification is very loose and is described in Ref. [1]. This requirement makes sure that muons and electrons in the regions without EM calorimeter coverage are not mistakenly identified as jets. The $\not\!\!E_T$ is corrected for jets and muons. To suppress the QCD multijet contribution, we require the angle between the $\not\!\!E_T$ vector and any identified jet[2] above 5 GeV to be more than 0.4 radians. Additionally, we require signed $\not\!\!E_T$ significance to be more than 4 (described in Section 1.3). Beam halo is removed by requiring the event EM fraction, E^{EM}/E^{TOT}, to be between 0.3 and 0.85. This event EM fraction is calculated based on energies of all jets with raw E_T>10 GeV. The complete list of cuts is presented in Table 1. We define our signal sample (used in extracting the diboson signal) as events in the 40 GeV<M_{jj}<160 GeV region.

Variable	Cut value
$\not\!\!E_T$	>60 GeV
E_T	>25 GeV
$Jet_{EmFraction}$	< 0.9
$\|\eta\|$	<2
$\not\!\!E_{T\,significance}^{signed}$	>4
$\Delta\phi_{closest}$	>0.4
$\Delta R_{lepton-jet}$	>0.2
$\frac{E^{EM}}{E^{TOT}}$	<0.85
$\frac{E^{EM}}{E^{TOT}}$	>0.3

Table 1: List of cuts applied

1.1 Cosmic removal

Cosmic rays going through the detector can mimic our signature with two clusters on one side of the calorimeter. Obviously this will also produce $\not\!\!E_T$ pointing to the other side of the calorimeter. To remove these events we require the EM timing and the HAD timing of jets to be within 4.5 ns and 15 ns, respectively. The cut is sequentially applied. We reject the

[1] We do not apply any additional restrictions to jets outside $|\eta|$>2.0 region.
[2] All jets in the region $|\eta|$<3.6 are considered.

event if the first jet has an EM time measurement and it is outside the 4.5 ns window. If the first jet does not have an EM time measurement[3], we look at the second jet EM time and reject the event in the case $|T^{EM}_{jet2}|$>4.5 ns. If both jets have no EM timing measurement, we repeat the procedure with the HAD timing. It should be pointed out that we mostly rely on the EM timing as the fraction of events with the EM timing information unavailable is very small[4]. Fig 1 shows the first jet EM timing for the identified $Z \to \mu^+\mu^- + jj$ events in our primary dataset. Based on these events, we derive an efficiency for our timing cuts to be $98.9 \pm 0.2\%$ per event.

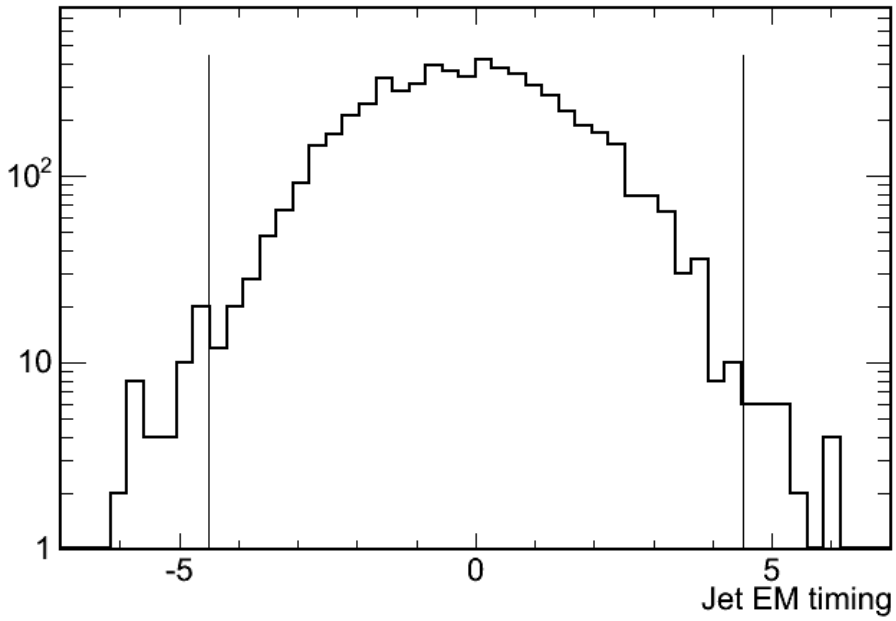

Figure 1: Jet EM timing for Z$\to$ *ll* selected events The 4.5 ns cut is shown.

The remaining background due to cosmics is derived based on a sideband analysis. Fig 2 shows the distributions of EM timing for jets before the cuts have been applied. The shaded red regions (-25,-10) ns and (25,50) ns define our sidebands. Given the total number of events in the sidebands we calculate the expected background which passes our cuts to be 97 ± 6 events. Since our analysis relies solely on shape of the backgrounds it is important to check that there are no features introduced in the dijet mass distribution from cosmic backgrounds. This is indeed the case as shown in Fig 3 where we plot the dijet mass distribution for events in the sideband region. We ignore the cosmic background in the rest of the analysis because it accounts for only 0.2% of our signal sample and it has a smoothly falling M_{jj} spectrum.

[3]Jet must have at least 4 GeV of EM or HAD energy in order to have a good EM or HAD timing measurement, respectively. There is also a small fraction of runs when the EM timing information was not available.

[4]Note the fact that we only use data past run 190697 with the EM timing system already installed

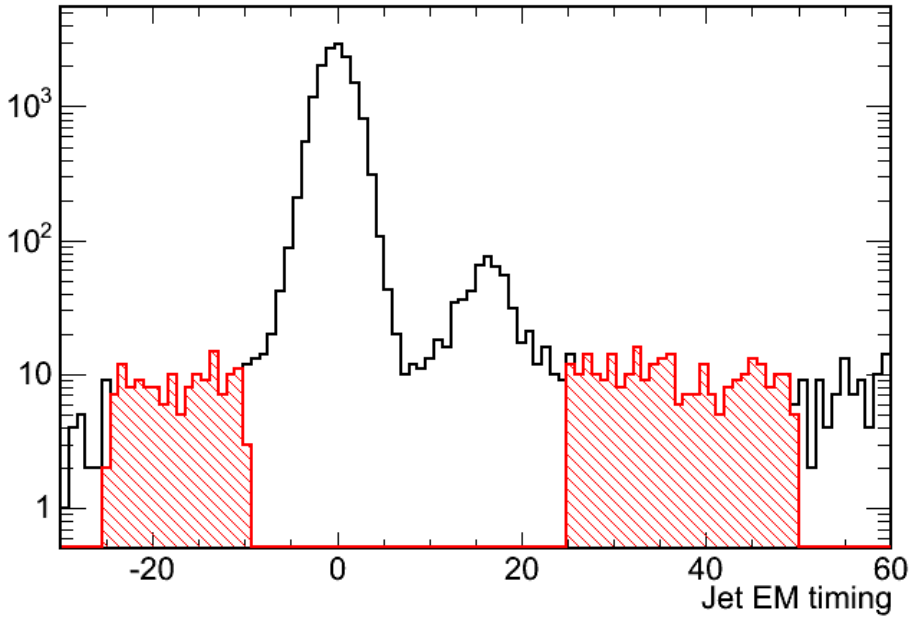

Figure 2: Jet EM timing for events before the timing cuts are applied. The shaded red regions show our sidebands which we use to infer the amount of background inside our analysis region. The second peak at ∼18 ns corresponds to events from satellite bunches (see note [2] for details on different timing shapes).

1.2 Lepton identification

In general we do not explicitly reject events due to the presence of leptons. The exception to this is when an electron from a W decay ends up in the central-plug calorimeter crack and has very little EM deposition (partly because there is poor EM coverage in the wall region []) and therefore can pass our jet EM fraction cut. We want to reject these events for two reasons:

1. Since these electrons will have an energy characteristic of the W decay they tend to be more clustered around 40 GeV. Paired with another jet in the event this will contribute to the signal region more than to the low or high invariant mass regions.

2. We found that in MC, this effect is larger than in the data. This is shown for the MC in Fig. 4. Therefore, we choose to outright reject these events. It turns out that most of these events in this region will be identified as CrackTrack leptons as expected.

In order to reject events where the reconstructed jet was actually a lepton we only count jets that seperated from identified leptons by $\Delta R > 0.2$ (see figure 5). We count as identified leptons:

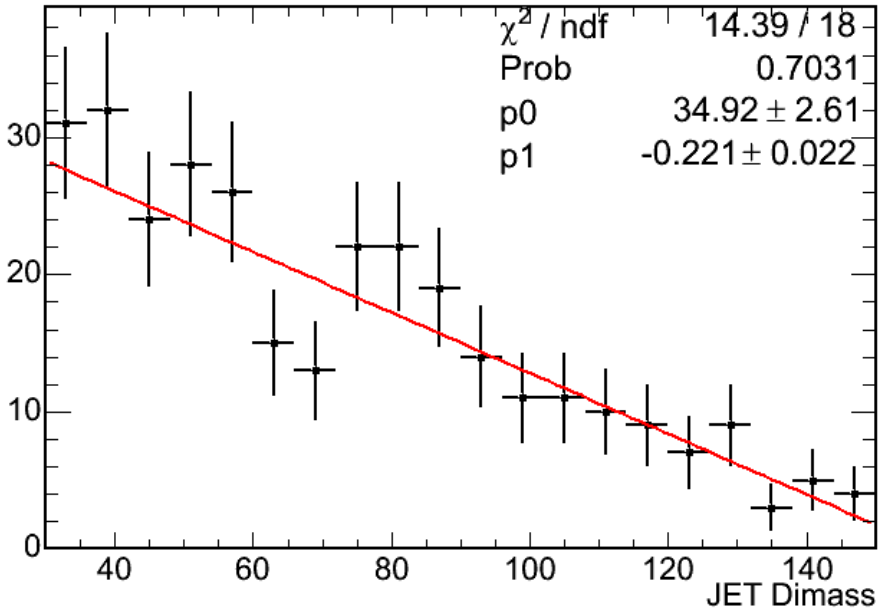

Figure 3: Dijet mass distribution for cosmic events.

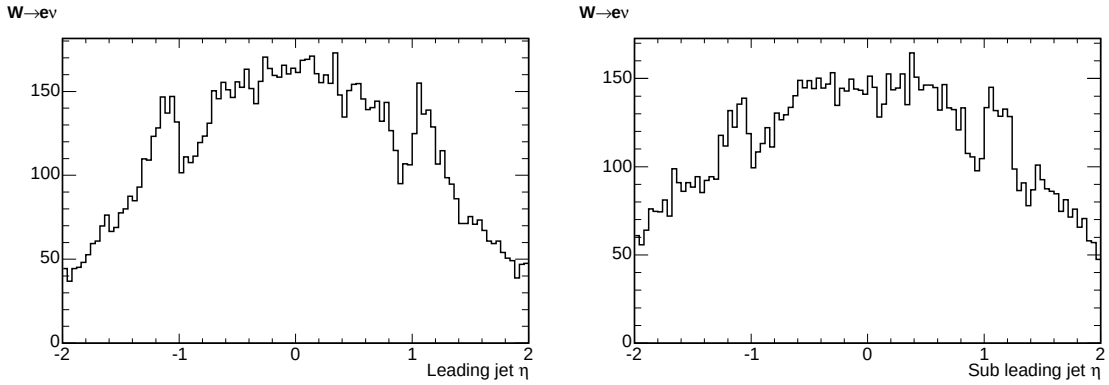

Figure 4: First and second jet eta from W→ $e\nu$ events. The excess in the central-plug crack region is evident.

- TCE - Tight Central Electron
- LCE - Loose Central Electron
- PHX - Phoenix electron
- PEM - Plug electron
- CMUP - Track confirmed in both CMU and CMP chambers

- CMU - Track confirmed in CMU chamber only
- CMP - Track confirmed in CMP chamber only
- CMX - Track confirmed in CMX chamber
- CMXmk - Track confirmed in the Miniskirt and Keystone part of CMX
- CMIOCES - Track consistent with MIP in the central calorimeter
- CMIOPES - Track consistent with MIP in the plug calorimeter
- CRKTRK - Track that points to the crack regions of the calorimeter

All leptons are required to be isolated in the calorimeter.

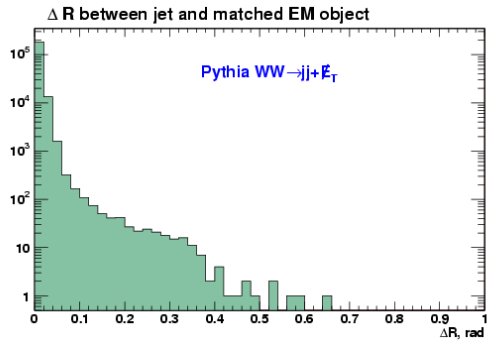

Figure 5: ΔR between jets and any identified leptons in the event (W$\rightarrow e\nu$)

1.3 Rejection of QCD Multijet Background: Met Model

Any measurement or search in the $\not\!\!E_T$+jets final state will suffer from a significant QCD multijet background. QCD multijet production does not typically contain large intrinsic $\not\!\!E_T$, however, when a jet is not reconstructed accurately the event may aquire large $\not\!\!E_T$ and pass the analysis selection criteria. This does not happen often, but because of the high cross section, it can still be a significant background in a $\not\!\!E_T$+jets based analysis.

In order to suppress the QCD background we use a $\not\!\!E_T$ resolution model to distinguish true $\not\!\!E_T$, from undetected neutrinos, and $\not\!\!E_T$ that arises from jets that are not measured accurately. The $\not\!\!E_T$-significance is a dimensionless quantity based on the energy resolution of the jets, the soft unclustered particles, and the event topology, details of model can be found in [3]. When $\not\!\!E_T$ arises from mis-measurement the $\not\!\!E_T$-significance is typically low. In addition to having a small values of $\not\!\!E_T$-significance, the $\not\!\!E_T$ will often be aligned with a jet. To suppress the QCD background we select events that satisfy $\not\!\!E_T$-significance$>$ 4 and the angle between the $\not\!\!E_T$ and the nearest jet ($\Delta\phi^{jet}_{\not\!\!E_T}$) greater that 0.4 radians.

An example of the jet energy resolution (JER) parameterization, $\mathcal{P}(x)$, is shown in Fig. 6. The asymmetric shape (longer tails on the negative side) of JER can be explained by a fact that some fraction of jet's energy can be lost in cracks. The $\not\!\!E_T$-significance is calculated

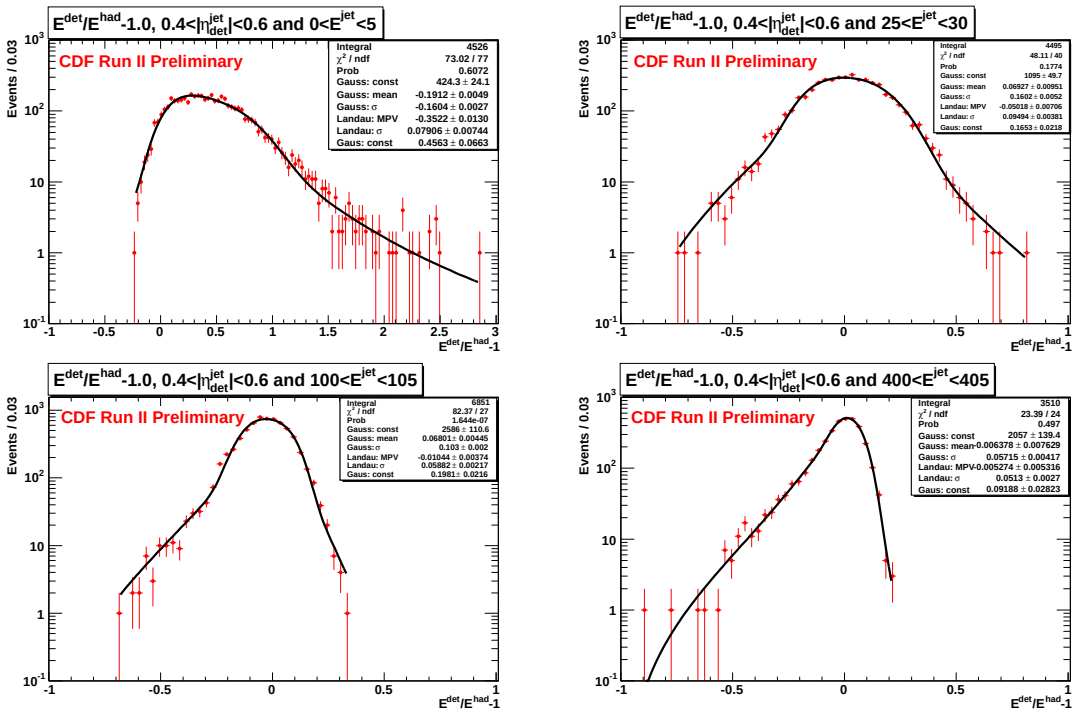

Figure 6: Example of jet energy resolution fits for four different jet energy bins: E^{jet} <5 GeV; 20 GeV< E^{jet} <25 GeV; 100 GeV< E^{jet} <105 GeV; and 400 GeV< E^{jet} <405 GeV.

based on the following formulae:

$$\not\!\!E_T\text{−significance} = -\log_{10}(\widetilde{\mathcal{P}}), \tag{1}$$

$$\widetilde{\mathcal{P}} = \prod \int_{-1}^{y_i} \mathcal{P}_i(x)dx, \quad \text{if} \quad y_i < 0,$$

$$\text{or} \quad \widetilde{\mathcal{P}} = \prod \left(1 - \int_{-1}^{y_i} \mathcal{P}_i(x)dx\right), \quad \text{if} \quad y_i > 0,$$

$$y_i = \not\!\!E_T/(E_T^i cos\Delta\phi_i), \tag{2}$$

where $\mathcal{P}_i(x)$ is the jet energy resolution function (see example at Fig. 6), E_T^i is the energy of a i-th jet, and $\Delta\phi_i$ is the azimuthal angle between that jet and $\not\!\!E_T$. As it was shown in ref. [3], the $\not\!\!E_T$-significance distribution has very different shapes in unbiased events with fake and true $\not\!\!E_T$: exponentially falling and roughly flat shapes[5], respectively. To suppress the QCD multi-jet background, we require events to have $\not\!\!E_T$-significance>4 and minimum angular separation between $\not\!\!E_T$ and any jet with E_T>5 GeV to be min($\Delta\phi$)>0.4 rad. A brief explanation for this choice of cuts is given below. The simple exponential shape of the $\not\!\!E_T$-significance distribution in QCD events allows to choose the desired level of rejection if one knows how many multi-jet events are expected. For the $\not\!\!E_T$>60 GeV and jet cuts that used in this analysis, we may expect $\sim 6\times10^7$ QCD multi-jet events in 3.5 fb^{-1} of data. This ballpark estimate is based on the following assumptions. In order to get at least 60 GeV of fake $\not\!\!E_T$, the original jet must have had E_T>60-70 GeV[6]. The inclusive di-jet x-section for E_T>60-70 GeV is $\sigma_{jj}\sim$80 nb [4]. In addition to an energetic jet which downward energy fluctuation gives rise to fake $\not\!\!E_T$, we must also have an extra (third jet) with E_T>25 GeV to satisfy our N_{jet25}=2 cut. This happens in about 20% of cases: $F_{jet3}=\sigma_{N\geq2}/\sigma_{N\geq3}\sim$0.2. Therefore, the ballpark number of expected QCD events is given by $\sigma_{jj}F_{jet3}\int\mathcal{L}dt\sim6\times10^7$. By requiring the $\not\!\!E_T$-significance>4, we suppress the QCD background by ∼4 orders of magnitude (down to $\sim6\times10^3$ events) while still being efficient for signal. The additional suppression of QCD is achieved by requiring min($\Delta\phi(\not\!\!E_T, jet)$)>0.4 rad. This cut is motivated by an MC study which shows that a fluctuated downward jet often leaves at least a small amount of energy in the calorimeter and the fake $\not\!\!E_T$ is aligned along the direction of that small jet.

The validation of $\not\!\!E_T$-significance and min($\Delta\phi(\not\!\!E_T, jet)$) distributions for events with real $\not\!\!E_T$ is based on the comparison of W+jet events in data and MC. For this purpose, we use ELECTRON_CENTRAL_18 trigger dataset, Pythia inclusive $W\rightarrow e\bar{\nu}$ (we0sej) and $W\rightarrow\tau\bar{\nu}$ (we0sbt) stntuple samples, all three corresponding to period-13. We apply the same selection criteria to data and MC events. We require only one tight central electron with E_T>20 GeV, only one jet with E_T>25 GeV, raw $\not\!\!E_T$>35 GeV, and at least one good vertex (quality 12) with $|z_{vx}|$<60 cm. These selection criteria are "orthogonal" (i.e., no overlap) to the requirements used for the $jj + \not\!\!E_T$ sample. Figure 7 shows the $\not\!\!E_T$-significance distribution

[5]We should point out that any event selection criteria, especially $\not\!\!E_T$-cut, will affect the shapes of $\not\!\!E_T$-significance distribution in events with true and fake $\not\!\!E_T$.

[6]It's much more likely for an energetic jet to completely loose its energy than for a small jet to fluctuate upward by >60 GeV

in data and MC events. One can see that shapes of data and MC distributions agree very well in the region $\not\!\!E_T$-significance>4 (dominated by true W+jet events). Excess of data over the MC predictions in the region of small $\not\!\!E_T$-significance is due to a QCD contamination by multi-jet events where both electron and $\not\!\!E_T$ are fake. This becomes even more apparent if one compares $\not\!\!E_T$ and M_T distributions, Fig. 8, before and after $\not\!\!E_T$-significance>5 cut which results in almost complete removal of the QCD background. The comparison of the $\min(\Delta\phi(\not\!\!E_T, jet))$ distribution before and after $\not\!\!E_T$-significance>5 cut can be found at Fig. 9. This figure clearly demonstrates the effect of the $\not\!\!E_T$-significance>5 cut which removes the excess of data events in the region $\min(\Delta\phi(\not\!\!E_T, jet))$<0.4 (where QCD is expected to dominate).

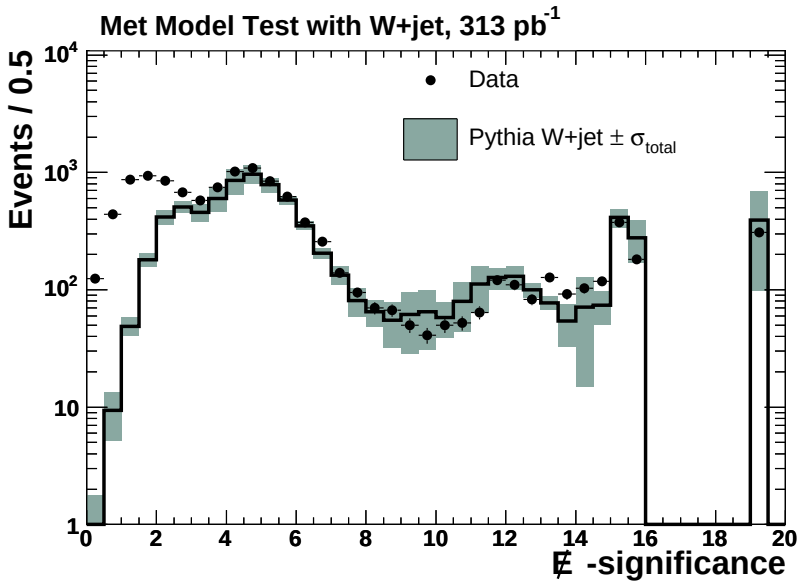

Figure 7: The $\not\!\!E_T$-significance distribution in $W + jet$ data (marker) and MC (histogram) events. The uncertainty (gray band) on MC predictions include: 6% luminosity uncertainty, jet energy scale uncertainty, and uncertainty due to data-MC differences in *Met Model* parameterization.

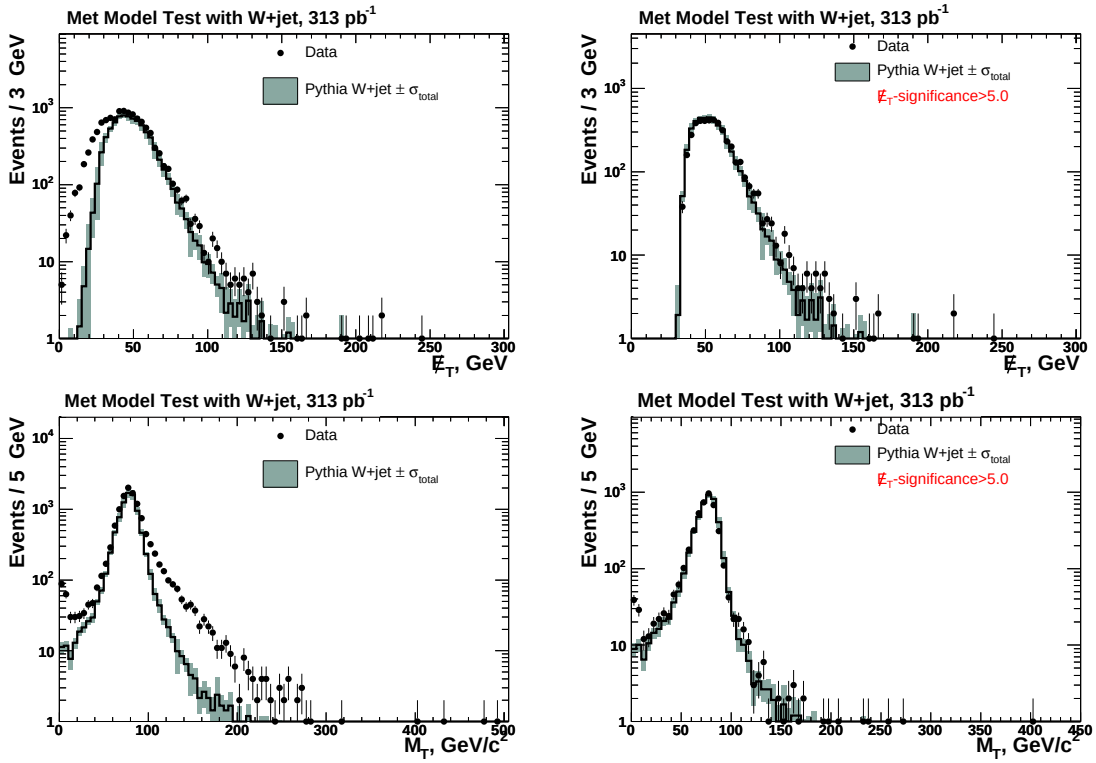

Figure 8: The $\not\!\!E_T$ (top) and M_T (bottom) distributions in $W + jet$ data (marker) and MC (histogram) events before (left) and after (right) the $\not\!\!E_T$-significance>5 cut. The uncertainty (gray band) on MC predictions include: 6% luminosity uncertainty, jet energy scale uncertainty, and uncertainty due to data-MC differences in *Met Model* parameterization.

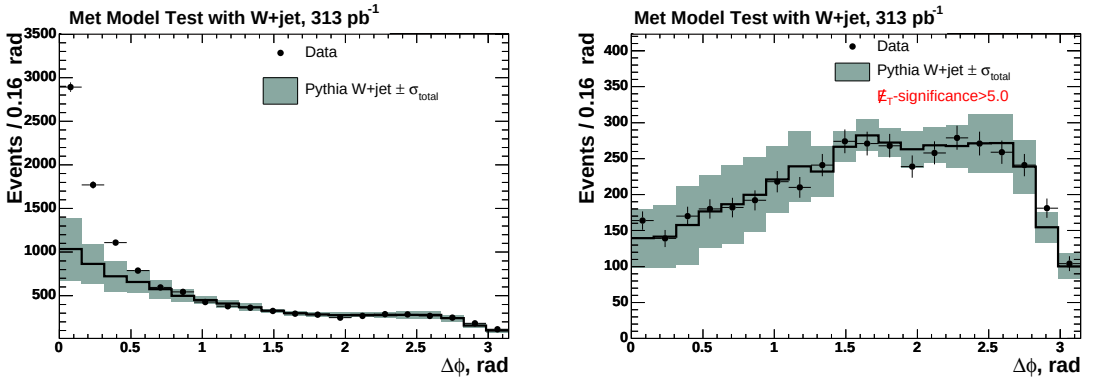

Figure 9: The $\min(\Delta\phi(\not\!\!E_T, jet))$ distributions in $W + jet$ data (marker) and MC (histogram) events before (left) and after (right) the $\not\!\!E_T$-significance>5 cut. The uncertainty (gray band) on MC predictions include: 6% luminosity uncertainty, jet energy scale uncertainty, and uncertainty due to data-MC differences in *Met Model* parameterization.

2 Datasets used

2.1 Data

We use the $\not\!\!E_T$ triggered emet data stream. All triggers which enter this stream are used in the analysis so no specific trigger selection is made. This ensures that we have the largest number of recorded CDF events to extract a signal from. We use data through period 20 with a run range from 190697 to 267718 corresponding to an integrated luminosity of 3.3 fb^{-1}. The good run list applied is good_em_nosi.list v26 which only requires good COT and calorimeter bit and makes no muon or silicon requirements. We find 44,910 events pass our selection cuts and fall in dijet mass range of interest, (40,160) GeV.

2.2 Monte Carlo

We use a combination of Pythia and Alpgen MC depending on availability. For the W+jets background we use the top group Alpgen samples; for Z(ll)+jets and $t\bar{t}$ we use the Pythia EWK samples and for signal we use the top group EWK samples (Pythia); for Z($\nu\nu$)+jets we use our own pythia sample. The list of samples used as they appear in the Stntuple catalog together with their cross sections are listed in Tab. 2.

Since the Z$\rightarrow \nu\nu$ samples are not official samples we include below the tcl file used in generating them:

```
module enable Pythia
module talk Pythia
                PythiaMenu
                   cmEnergy set 1960
# generate gamma*/Z
                   msel set 11
         inListLevel set 0
         evListLevel set 1

          commonMenu

# Setup the decay of the gamma*/Z->mumu :
# decayType=0(1): Channel off(on)
# don't touch decayType = -1!!!!

# Z decay table turning off all allowed decays:
            set_mdme -channelIndex=174 -decayType=0
            set_mdme -channelIndex=175 -decayType=0
            set_mdme -channelIndex=176 -decayType=0
            set_mdme -channelIndex=177 -decayType=0
            set_mdme -channelIndex=178 -decayType=0
```

```
            set_mdme -channelIndex=179 -decayType=0
            set_mdme -channelIndex=182 -decayType=0
            set_mdme -channelIndex=183 -decayType=0
            set_mdme -channelIndex=184 -decayType=0
            set_mdme -channelIndex=185 -decayType=0
            set_mdme -channelIndex=186 -decayType=0
            set_mdme -channelIndex=187 -decayType=0

# Z/gamma turning on neutrino channels:
            set_mdme -channelIndex=183 -decayType=1
            set_mdme -channelIndex=185 -decayType=1
            set_mdme -channelIndex=187 -decayType=1

# Set Minimum Invariant Mass
         set_ckin -index=1 -value=30
    exit
  exit
exit
#--------------------------------------------------------------------
# Pythia tunings
#--------------------------------------------------------------------
source $env(PROJECT_DIR)/mcProduction/tcl/mc_Pythia_WZPt_tune_RunII.tcl
source $env(PROJECT_DIR)/mcProduction/tcl/mc_Pythia_pdf_CTEQ5L.tcl
source $env(PROJECT_DIR)/mcProduction/tcl/mc_Pythia_underlying_event_A.tcl
```

Sample	Cross-section	Description	Process
fhbs1a	2566	Pythia	Z→$\nu\nu$
fhbs3a	2566	Pythia	Z→$\nu\nu$
it0sww	12.4	Pythia	WW
it0swz	3.7	Pythia	WZ
it0szz	3.8	Pythia	ZZ
tt0s75	6.7	Pythia	ttbar
st0s00	0.2864	Madgraph	single top s-channel
st0s01	0.6415	Madgraph	single top t-channel
pt0sw0	1800	Alpgen (W+0p)	W→eν
pt0sw1	225	Alpgen (W+1p)	W→eν
pt0s2w	35.3	Alpgen (W+2p)	W→eν
pt0s3w	5.59	Alpgen (W+3p)	W→eν
pt0s4w	1.03	Alpgen (W+4p)	W→eν
pt0s5w	1800	Alpgen (W+0p)	W→$\mu\nu$
pt0s6w	225	Alpgen (W+1p)	W→$\mu\nu$
pt0s7w	35.3	Alpgen (W+2p)	W→$\mu\nu$
pt0s8w	5.59	Alpgen (W+3p)	W→$\mu\nu$
pt0s9w	1.03	Alpgen (W+4p)	W→$\mu\nu$
ut0sw0	1800	Alpgen (W+0p)	W→ $\tau\nu$
ut0sw1	225	Alpgen (W+1p)	W→ $\tau\nu$
ut0s2w	35.3	Alpgen (W+2p)	W→ $\tau\nu$
ut0s3w	5.59	Alpgen (W+3p)	W→ $\tau\nu$
ze0sat	355	Pythia	Z→ $\tau\tau$
ze0sbt	355	Pythia	Z→ $\tau\tau$
ze1sad	497	Pythia	Z→ee
ze0sdd	497	Pythia	Z→ee
ze0scd	497	Pythia	Z→ee
ze0sed	497	Pythia	Z→ee
ze0see	497	Pythia	Z→ee
ze0seh	497	Pythia	Z→ee
ze1s9m	497	Pythia	Z→$\mu\mu$
ze0sbm	497	Pythia	Z→$\mu\mu$
ze0sdm	497	Pythia	Z→$\mu\mu$
ze0sem	497	Pythia	Z→$\mu\mu$
ze0sfm	497	Pythia	Z→$\mu\mu$
ze0scm	497	Pythia	Z→$\mu\mu$

Table 2: List of MC samples used. The k-factor which accounts for the NLO corrections is included in the cross section whenever necessary (e.g., for all di-boson and Z samples). All Pythia samples with Z's have a virtual photon component as well. There are several instances where we list several samples for one process. This comes about for two reasons; firstly, Alpgen requires several n-parton subsamples to describe a process, and secondly, in the case of some Pythia samples we combine all the available samples. An example of this is where we have two $Z \rightarrow \nu\nu$ samples that we combine to gain statistics.

3 Trigger

The trigger efficiency for the data sample is determined from the high p_T muon data, **bhmu0x**. This sample is enriched with $Z \to \mu\bar{\mu}$ events with "real" $\not\!\!E_T$ in the calorimeter due to muons being MIP at the energies of interest. The high p_T muon trigger is well understood, so it provides a good sample to study the effeciency of our $\not\!\!E_T$ triggers. The trigger efficiency curve is determined for the entire trigger sample by dividing the $M_{jj}(\not\!\!E_T)$ distribution for events that pass the selection cuts from Table 1 and have a Level-3 $\not\!\!E_T$-trigger bit fired by the same distribution for events that just pass the selection cuts. The lower edge of the mass range for the extraction fit is determined by where the efficiency of the trigger combination used plateaus. The trigger efficiency is parameterised by

$$\epsilon = \frac{c}{1 + e^{\frac{a-x}{b}}}.$$

Results for the trigger efficiency as a function of both, $\not\!\!E_T$ and M_{jj} are summarised in the Table 3 and Fig. 10. The slight inefficiency due to saturated towers at L1 is a known feature of the calorimeter trigger before the L2CAL upgrade.

In order to determine the trigger efficiency for our diboson signal, we convolute the $\not\!\!E_T$ efficiency curve with the signal $\not\!\!E_T$ spectrum predicted by the Monte Carlo, Fig. 11. We obtain an integrated efficiency of 96% ± 2% which we will use to correct the final cross section and assign an uncertainty.

The trigger efficiency fits (Fig. 10) show that the trigger efficiency as a function of M_{jj} is constant above 40 GeV, this defines the lower edge of our mass window. The saturation effect mentioned above may make our integral efficiency an overestimate. To account for this, we assume that above 120 GeV the efficiency is actually lower than our fit says by 10%. This results in a 0.98% drop in the total integrated efficiency over the signal spectrum. We will use this as additional uncertainty. Summarizing, the trigger efficiency we use is 96.2% ± 2.2%.

It can be seen from Fig. 12 that the event sample is dominated by a small fraction of the total # of triggers; of the 44,910 events in our sample, 42,162 (or 93.9%) are passed by inclusive $\not\!\!E_T$ triggers (mostly MET40 and MET45), and 43,554 (or 97.0%) by inclusive $\not\!\!E_T$ combined with the $\not\!\!E_T$ plus dijet triggers. The remaining 3% of events are coming from $\not\!\!E_T$+jet, $\not\!\!E_T$+b-jet and $\not\!\!E_T$+jet+track triggers.

Parameters	$M_{jj} - efficiency$	$\not\!\!E_T$-efficiency
a	18.1±4.0	47.8±2.4
b	4.0±2.2	7.3±0.93
b	0.96±0.01	0.99±0.01

Table 3: Parameters of the trigger efficiency fits to $\not\!\!E_T$ and M_{jj} distributions.

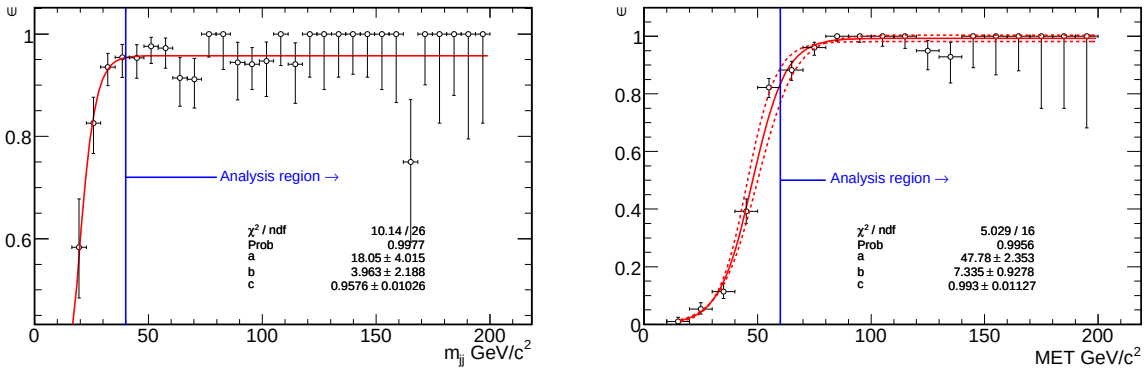

Figure 10: Trigger efficiency as a function of M_{jj} (left) and $\not\!\!E_T$ (right). The dashed lines in the $\not\!\!E_T$ plot show the statistical uncertainty.

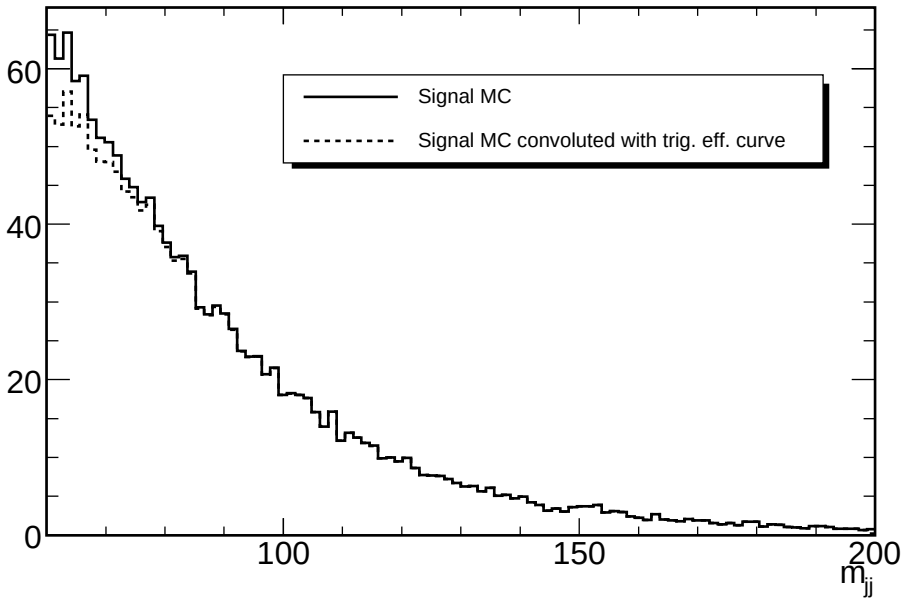

Figure 11: Signal $\not\!\!E_T$ distribution corrected for trigger efficiency curve.

3.1 Luminosity of the sample

We use $Z \rightarrow \mu\mu$ events in the high P_T muon triggered data and $\not\!\!E_T$ triggered data to establish the luminosity of our sample. This effectively does a scaling of the luminosity to a well establish standard candle, the Z cross section.

The number of observed events in the muon triggered data is

$$N = \sigma \cdot A \cdot \epsilon \cdot L_\mu \tag{3}$$

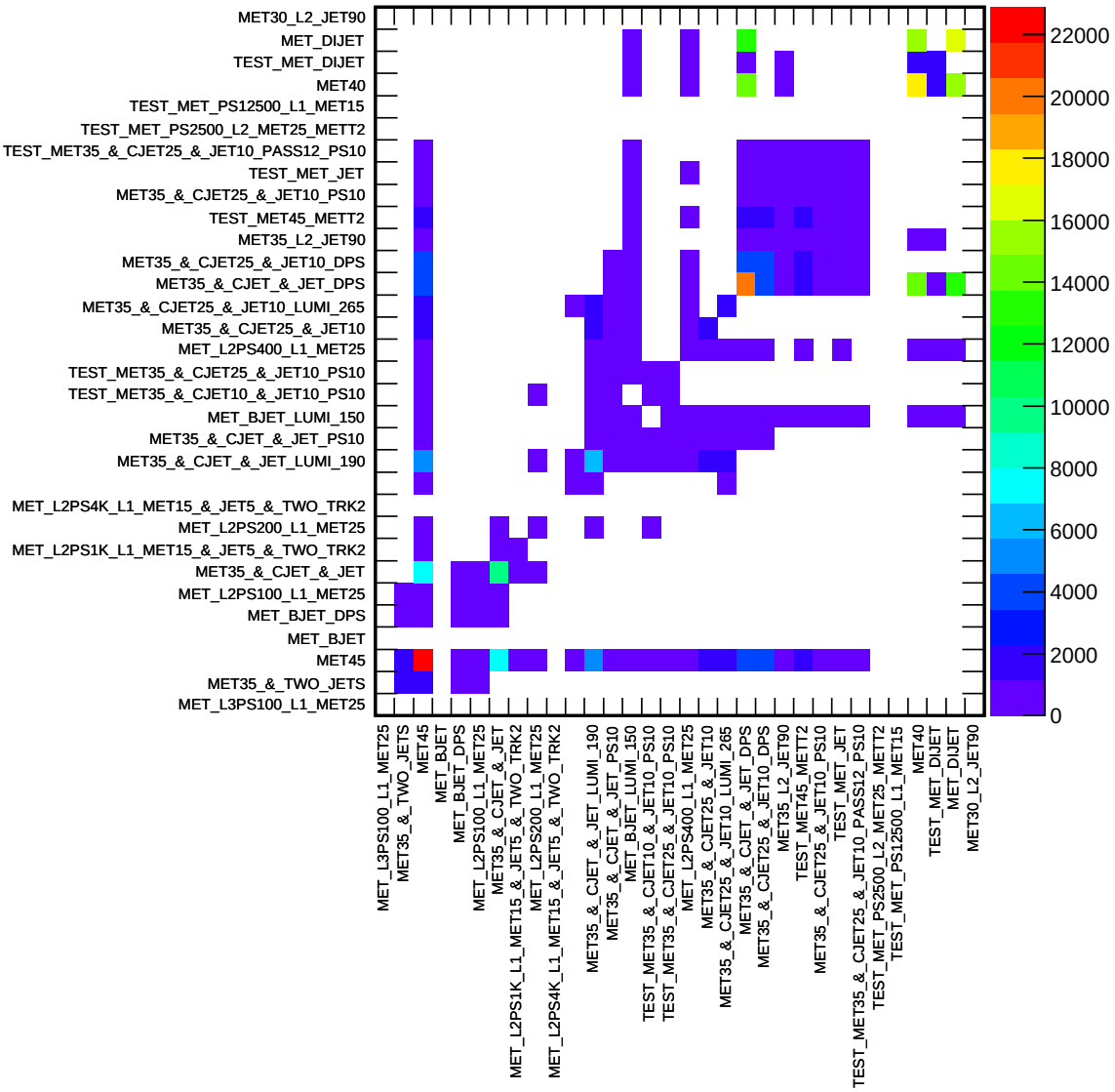

Figure 12: Set of triggers used in the analysis; out of a large table, only a few triggers dominate. Entries in every bin must satisfy both triggers on the x- and y- axes

where σ is the Z cross section (this is nature, has nothing to do with trigger, acceptance etc.), ϵ is the trigger efficiency for two leptons in the high P_T trigger which is taken to be 100% since leptons here means cmup and cmx only. A is the acceptance to this analysis, in other words acceptance to events with $\not\!\!E_T > 60$, two jets above the 25 GeV threshold and the rest of our analysis cuts. On top of all that, two ID'ed muons that come within the Z window and above 20 GeV are required. L_μ is the luminosity exposure of the high P_T trigger.

from the above formula we can calculate:

$$A = \frac{N}{\sigma \cdot L} \tag{4}$$

Using a similar formula to (4) we can calculate the luminosity of our $\not\!\!E_T$ sample:

$$L_{\not\!\!E_T} = \frac{N_{\not\!\!E_T}}{\sigma \cdot A \cdot \epsilon_{\not\!\!E_T}} \tag{5}$$

here, A is the same acceptance because the analysis is the same (still asking two muons as well). $\epsilon_{\not\!\!E_T}$ is the efficiency of the $\not\!\!E_T$ triggers combined and is 96% and we obviously have the same σ here. Putting it all together:

$$L_{\not\!\!E_T} = \frac{N_{\not\!\!E_T}}{N \cdot \epsilon_{\not\!\!E_T}} \cdot L$$

Fig. 13 shows a plot of the lepton dimass in high P_T muon data (black) and $\not\!\!E_T$ triggered data (red).

Finally, we find $N_{\not\!\!E_T} = 339$ and N = 357. The luminosity of the high Pt muon triggers is 3488pb^{-1} which results in $L_{\not\!\!E_T} = 3450\text{pb}^{-1}$. This calculation is still dominated in the end by the uncertainty on the high Pt muon trigger luminosity, therefore we still assign the usual 6% uncertainty on the final cross section.

To check this method we will apply it to the inclusive $\not\!\!E_T$ triggers only (throughout the run runage that was MET40 and MET45). Using the above selection we get $N_{\not\!\!E_T}$=323. As shown before, the inclusive $\not\!\!E_T$ triggers get only 93.9% of the events we select, so we will make the assumption that the efficiency of the inclusive $\not\!\!E_T$ triggers alone is 93.9% of our measured efficiency of 96%. This was we obtain $\epsilon_{\not\!\!E_T} = 90.14\%$. Using (5) we obtain the luminosity of the inclusive $\not\!\!E_T$ triggers to be $L_{\not\!\!E_T} = \frac{323 \cdot 3488}{357 \cdot 0.9014} = 3501\,\text{pb}^{-1}$. If we use the dpslum.sh script from the good run list web page, then the total luminosity that we get for these triggers is $3443\,\text{pb}^{-1}$.

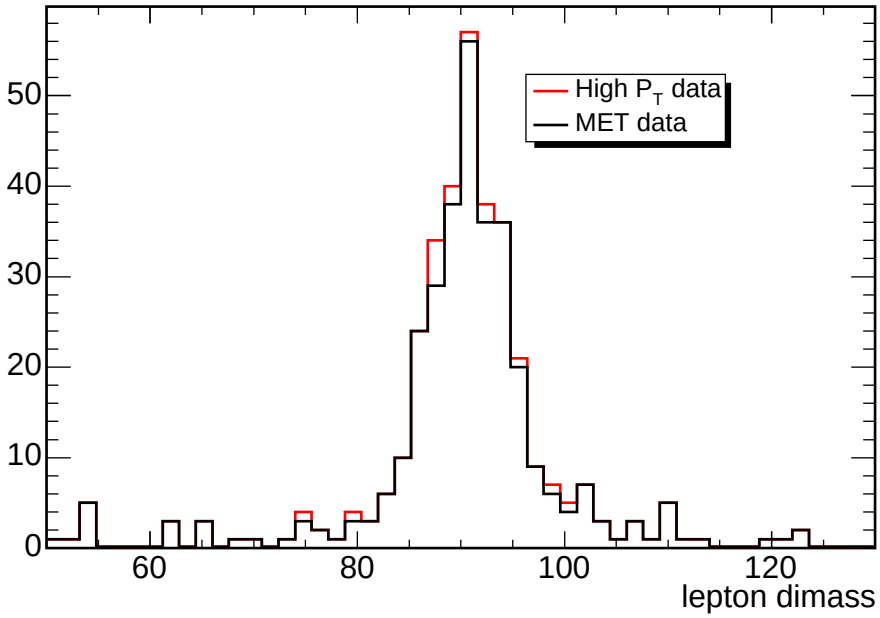

Figure 13: Selected Z events from the high P_T and $\not\!\!E_T$ triggers .

4 Signal and Background Simulation

4.1 Background Simulation/Modelling

The Monte Carlo samples used to model the electroweak and top background are generated with Pythia and Alpgen. Pythia is used for the Z+jets and top backgrounds, the W+jets background samples are generated using Alpgen (see Section 2.2 for details). The background contribution from QCD is derived from the data. The background categories are discussed in more detail below.

4.1.1 Electroweak

This analysis is interested in the case when dibosons (WW,WZ,ZZ) decay to either $\nu\nu jj$ or $l\nu jj$. These events are characterised by low jet multiplicity and large $\not\!\!E_T$. There are several Standard Model processes that produce this (or similar) signature. The contribution from Z+jets (jets from ISR/FSR) events where the Z decays to either, two electrons, two taus or two muons is heavily suppressed by the selection criteria. The fourth decay process for Z+jets (jets from ISR/FSR) is the case where the Z decays to two neutrinos. This background mimics the signal in many ways, however, in contrast to the signal, it has a smoothly falling M_{jj} distribution in the mass window of interest.

Standard Model W+jets (jets from ISR/FSR) production is a significant background to both the $\nu\nu jj$ and $l\nu jj$ decays of the signal. When two jets are produced in conjunction with the W decaying to a lepton and neutrino the event will look like signal.

In both Z+jets and W+jets production the resulting dijet mass distribution looks significantly different from that expected from signal. The overall shape of the electroweak backgrounds is used to form a pdf for the extraction fit. Signal and background Monte Carlo samples are summarised in Section 2.2; their relative amounts, expressed both in terms of the expected # of events as well as their proportion of the (non-QCD) component of our sample, are shown in Tab. 4.

4.1.2 Cosmic

The background contribution from cosmic rays is reduced to a negligible level using electromagnetic and hadronic timing information as discussed in Sec. 1.1.

4.1.3 QCD

QCD multijet production does not typically contain large intrinsic $\not\!\!E_T$, however, when a jet is not reconstructed accurately the event may aquire large $\not\!\!E_T$ and pass the analysis selection criteria. This does not happen often, but because of the high cross section, it can still be a significant background in a $\not\!\!E_T$+jets based analysis. We derive both the initial normalisation and the shape of the QCD background from data. The final measure of the amount of QCD will be determined from the extraction fit.

Sample Description	Expected # of Evts	Expected % of Sample
Z→$\nu\nu$	12804	28.9
Z→ee	5	0.0
Z→$\mu\mu$	300	0.7
Z→ $\tau\tau$	430	1.0
W→eν	6389	14.4
W→ $\mu\nu$	5672	12.8
W→ $\tau\nu$	10697	24.1
$t\bar{t}$	388	1.0
single top	221	0.6
WW	1010	4
WZ	329	1
ZZ	151	0.9

Table 4: Expected contribution of different processes to our total data sample of 44,910 events, expressed in # of events and as a % of the sample in the dijet mass window. QCD is omitted from the list.

The underlying assumption of how QCD enters the analysis is that either jets are mismeasured, or a leading charged track, π^0 or a γ is lost in an uninstrumented region of the detector. We expect the dominant effect to be jet mismeasurement. Most of the QCD multijet background is suppressed by the $\not\!\!E_T$-significance and min($\Delta\phi$) cuts described in detail in Section 1.3. To estimate the remaining QCD contribution, we construct a new variable, $\not\!\!P_T$, to compliment the traditional calorimeter based $\not\!\!E_T$. The $\not\!\!P_T$ is defined as the negative vector sum of tracks (excluding COT stand-alone tracks) with p_T>0.5 GeV/c.

When comparing the azimuth angle (ϕ) for $\not\!\!E_T$ and $\not\!\!P_T$, alignment is expected between the two quantities in the case of true $\not\!\!E_T$ (e.g., for diboson signal and EWK backgrounds). We will call the difference between these two angles as $\Delta\phi_{MET}$. The expected shape (slowly falling spectrum) of the $\Delta\phi_{MET}$ distribution for the QCD background is illustrated at Fig. 14. It was obtained by studying Pythia QCD events that pass our selection cuts. We used large statistics Pythia **q8is01** stntuple sample for these studies. Electroweak backgrounds (and diboson signal) will be present in all regions, but will dominate at low $\Delta\phi_{MET}$ due to the presence of neutrinos. To validate the shape of $\Delta\phi_{MET}$ distribution in events with true $\not\!\!E_T$, we compare $Z \to \mu^+\mu^-$ in data and MC. We select Z+jj events with two high p_T muons forming the Z-peak. We also apply all of our event selection cuts. Of course, we do not apply the muon correction to $\not\!\!E_T$ in the $Z \to \mu^+\mu^-$ sample, so these events have the same topology as the events passing our final signal selection. The $M_{\mu\mu}$ distribution in selected data and MC $Z \to \mu^+\mu^-$ events can be found at Fig. 15 (left plot). One can see from this plot the MC does not perfectly model the data. As the $Z \to \mu^+\mu^-$ events have vertually no QCD contribution, we assume that the observed difference comes from not accurate modeling of resolution effects in MC. To account for this effect, we apply an additional Gaussian smearing

to the $\Delta\phi_{MET}$ distribution in EWK MC backgrounds that is derived from the comparison of $M_{\mu\mu}$ distributions in data and MC. Fig. 15 (right plot) shows $\Delta\phi_{MET}$ distribution before and after this Gaussian smearing is applied. The same smearing procedure is then applied to all of the electroweak backgrounds when comparing the data.

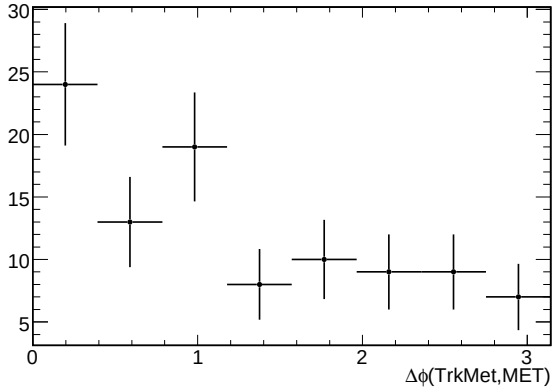

Figure 14: The $\Delta\phi_{MET}$ distribution in Pythia QCD events that pass our selection cuts.

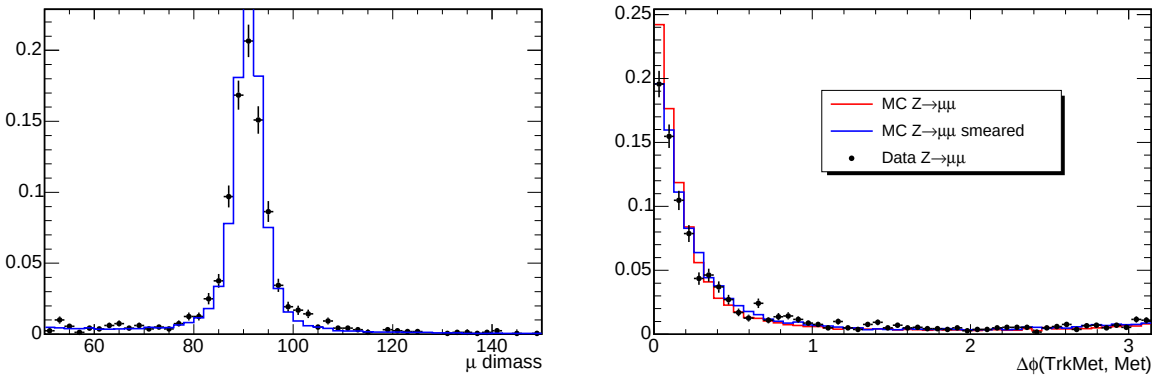

Figure 15: Left plot: invariant mass of two muons in selected events; these events are clearly Z→ $\mu\mu$. Right plot: $\Delta\phi_{MET}$ in $Z \to \mu^+\mu^-$ events before and after the Gaussian smearing of the MC is applied

In order to estimate the amount of remaining QCD background and obtain the template for its M_{jj} distribution, we normalise the EWK MC predictions to the data in the region $\Delta\phi_{MET}$<0.4 rad (see top plot in Fig. 16). Then, we do a bin-by-bin subtraction of the dijet mass distribution in data and electroweak background for events satifying the $\Delta\phi_{MET}$>1.0 rad requirement, i.e. in the region with enhanced QCD contribution. Once the QCD M_{jj} shape is determined (see bottom plot in Fig. 16), it is fit by an exponential: $a \times e^{(b \times m_{jj})}$. Other choices of the fit function were tried, however, the final result was insensitive to the choice. The procedure described above for extracting the QCD contribuition

does not account for the fact that there will be some QCD contamination in the region $\Delta\phi_{MET}$<1.0 rad, Fig. 14. The total number of the QCD events in this M_{jj} template (3167 events) has to be multiplied by a factor of 1.94±0.39 to account for contamination in the peak. Therefore, the integral over the properly normalized QCD M_{jj} template should be equal to 6144 events. We also compare M_{jj} distributions for MC QCD events in two regions: $\Delta\phi_{MET}$<1.0 rad and $\Delta\phi_{MET}$>1.0 rad, 17 . This comparison shows that the slopes of M_{jj} distributions in two regions differ by ∼20%. Although this difference is withing the statistical uncertainties, we concervatively decided to assign a 20% uncertainty on the slope of the M_{jj} distribution for our QCD template (shown at the bottom of Fig. 16). The template normalisation (6144 events) and its uncertainty (∼20%) as well as the uncertainty on slope (20%) will be used to constrain the QCD in the signal extraction fit.

4.1.4 Template Combinations

The background templates described above are used to extract the signal (will be discussed in detail in Section 5) as well as to compare the distributions of important kinematic observables used in the event selection. In Figs. 18-20, we compare data to the EWK+signal predictions (left) and to the total EWK+QCD+signal predictions (right) extracted from the fit. The EWK only plots clearly show where the QCD contribution is expected (e.g., see Fig. 20). Figure 21 shows the comparison of data to the total prediction for H_T, $\Delta\phi = \phi_{jet1} - \phi_{jet2}$, and $|\Delta\eta| = |\eta_{jet1} - \eta_{jet2}|$ distributions. From Figs. 18-21, one can also see that the total predicted shapes (EWK+QCD+signal) agree very well with data. This gives us confidence that our backgrounds are well modeled.

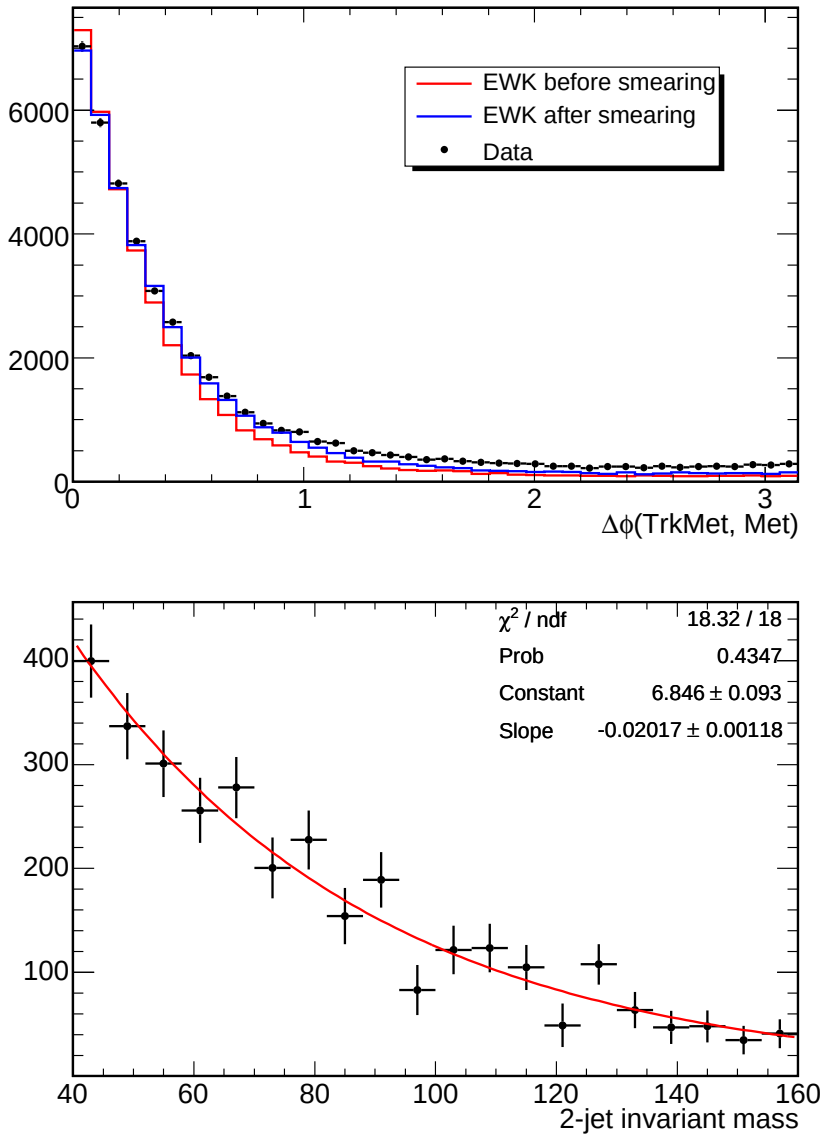

Figure 16: *Top plot*: the comparison of $\Delta\phi_{MET}$ distributions in data and EWK background (Gaussian smearing is applied to MC and it is normalized to data in the region $\Delta\phi_{MET}$<0.4 rad). *Bottom plot*: result of the bin-by-bin subtraction of the data and electroweak dijet mass distributions after the additional requirement of $\Delta\phi_{MET}$>1.0 rad is applied. This M_{jj} distribution will be used as the QCD pdf in the extraction fit.

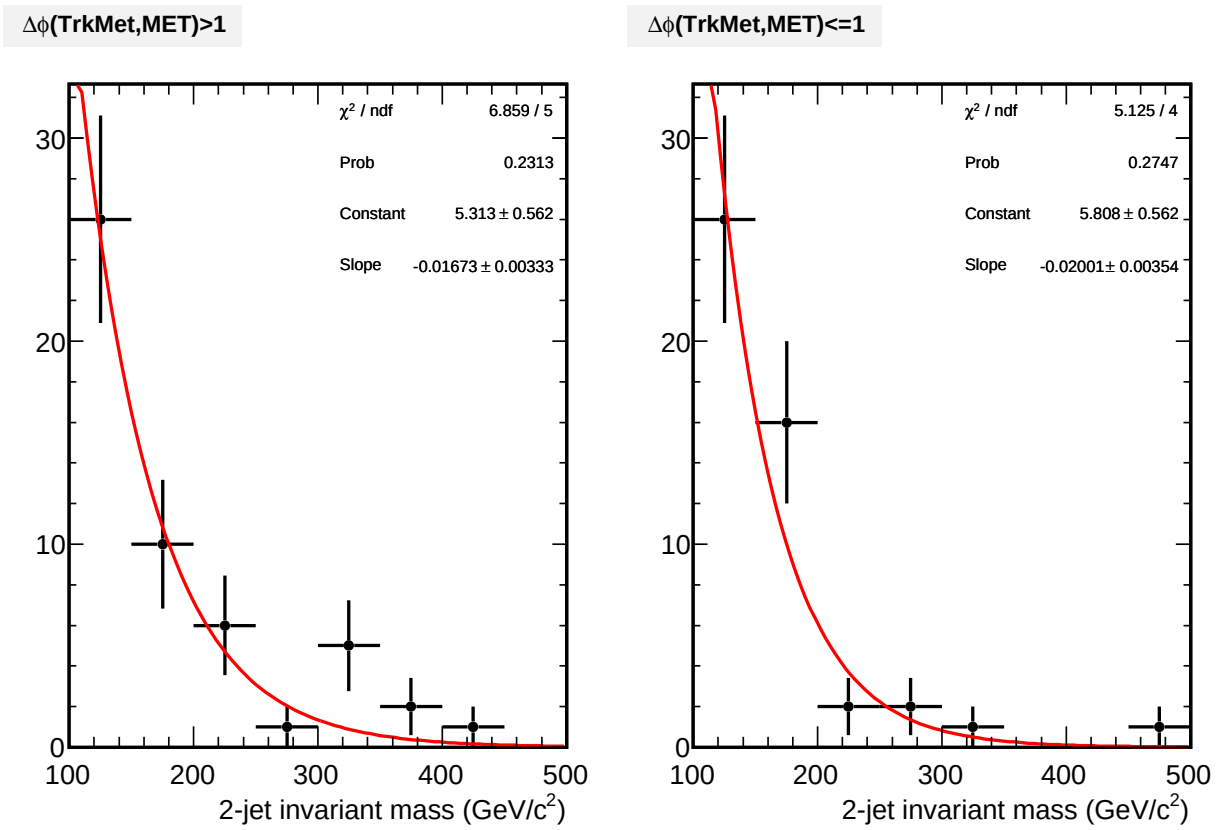

Figure 17: The M_{jj} distribution in Pythia QCD events that pass our selection cuts. The left plot is for events with $\Delta\phi_{MET}$>1.0 rad and the right plot is for events with $\Delta\phi_{MET}$<1.0 rad. Because of the generator level cut on P_T-hat used to produce this sample, we can't go below 100 GeV/c^2.

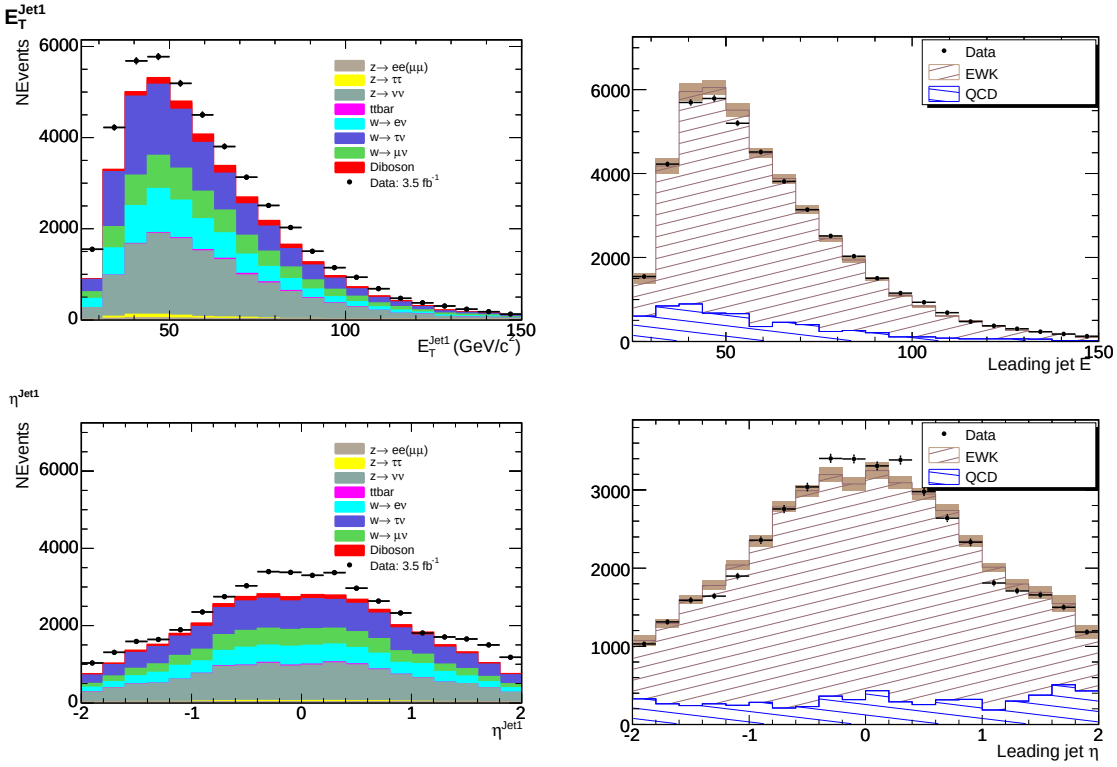

Figure 18: Comparison of data to the EWK+signal only predictions (left) and to the total QCD+EWK+signal predictions (right): E_T (top) and η (bottom) of the first jet. The uncertainty on the total prediction (brown shaded area) includes only 20% uncertainty on the QCD normalization.

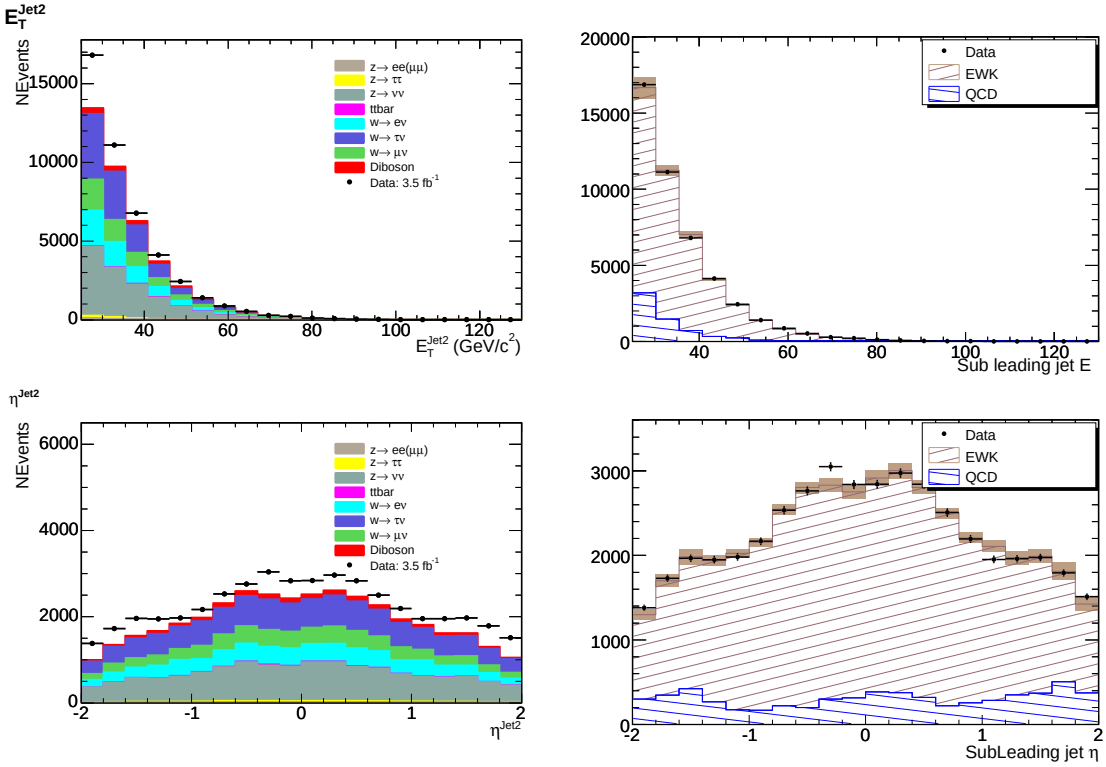

Figure 19: Comparison of data to the EWK+signal only predictions (left) and to the total QCD+EWK+signal predictions (right): E_T (top) and η (bottom) of the second jet. The uncertainty on the total prediction (brown shaded area) includes only 20% uncertainty on the QCD normalization.

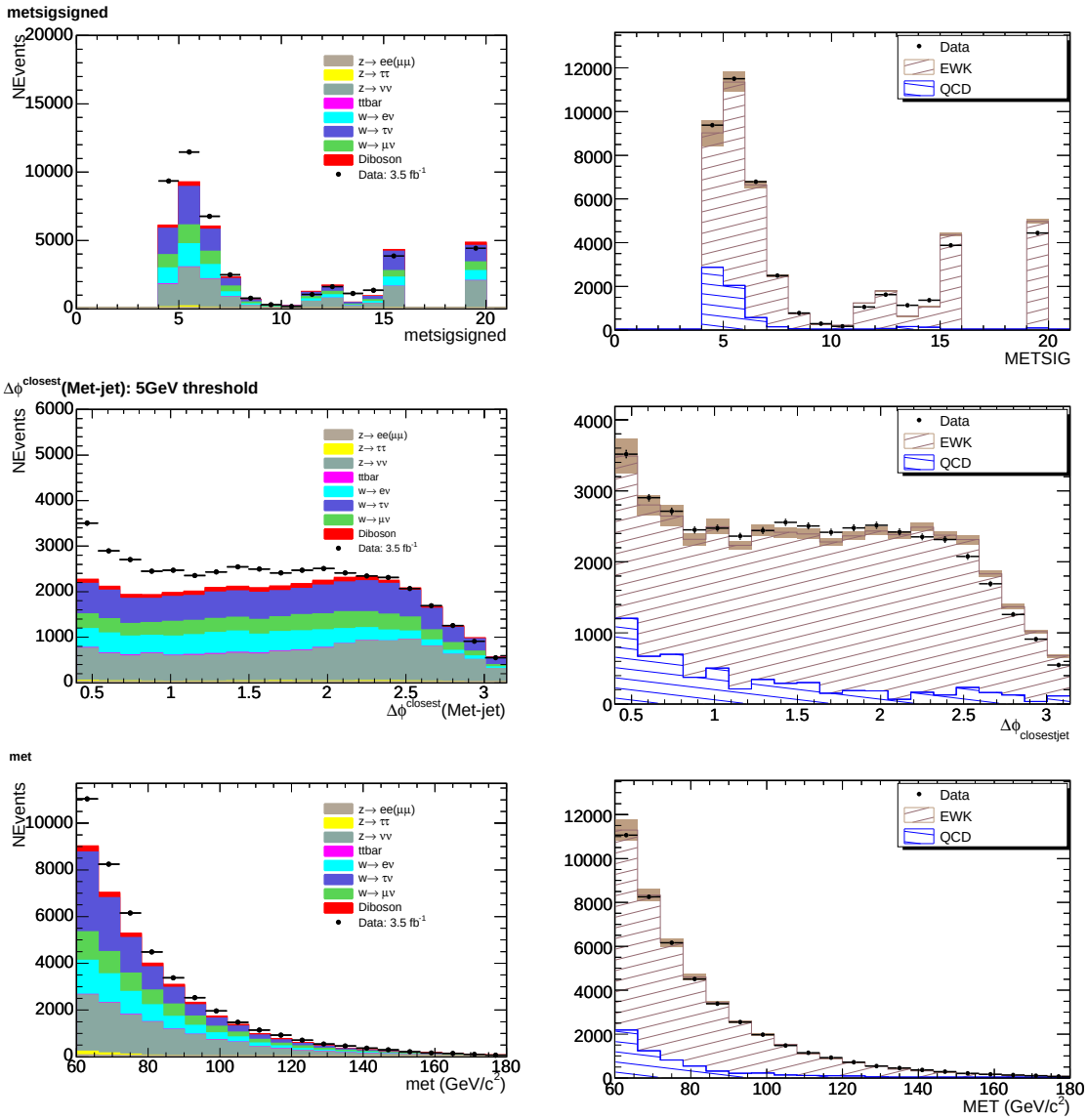

Figure 20: Comparison of data to the EWK+signal only predictions (left) and to the total QCD+EWK+signal predictions (right): $\not\!\!E_T$-significance (top), $\min(\Delta\phi(\not\!\!E_T - jet))$ (middle), and $\not\!\!E_T$ (bottom). The uncertainty on the total prediction (brown shaded area) includes only 20% uncertainty on the QCD normalization.

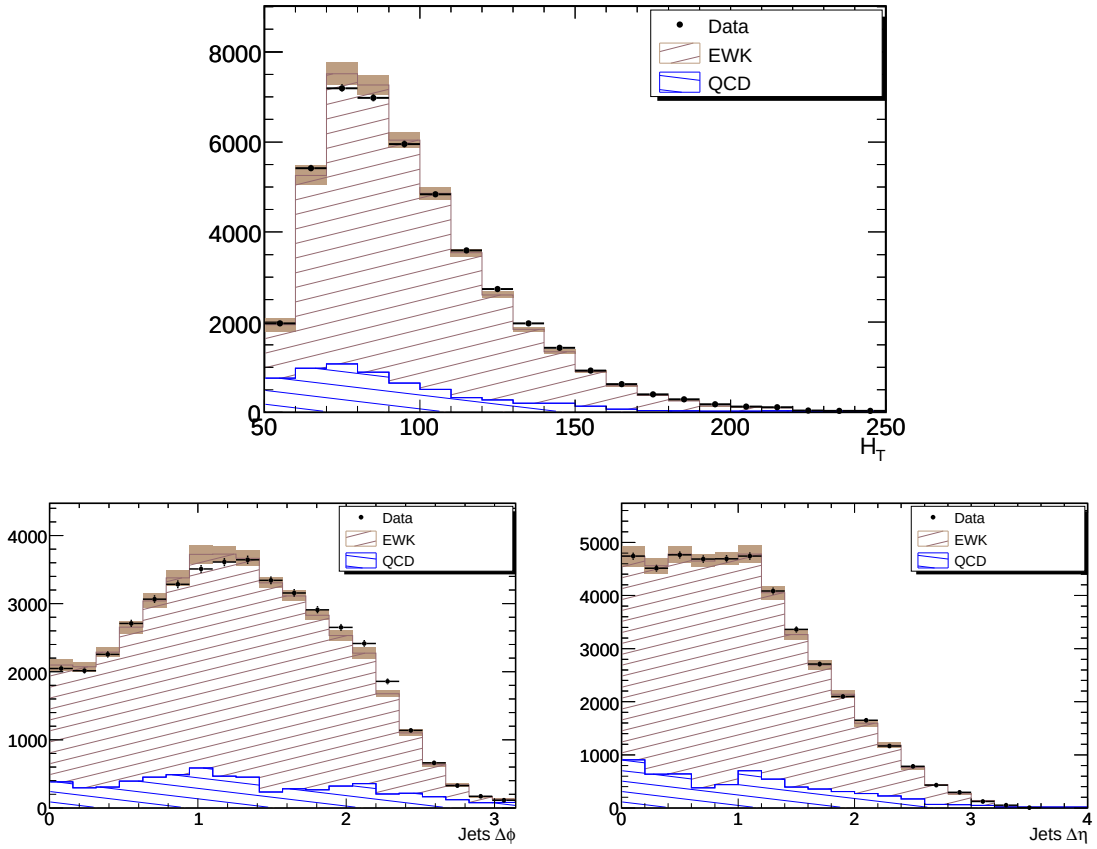

Figure 21: Comparison of data to the total QCD+EWK+signal predictions: H_T (top), $\Delta\phi = \phi_{jet1} - \phi_{jet2}$ (bottom left), and $|\Delta\eta| = |\eta_{jet1} - \eta_{jet2}|$ (bottom right). The uncertainty on the total prediction (brown shaded area) includes only 20% uncertainty on the QCD normalization.

5 Signal Extraction

The signal extraction is performed using a negative log likelihood minimisation within RooFit, the minimisation is delegated to Minuit. The fit is performed for the *a priori* defined dijet mass range of (40,160) GeV/c^2. The triggers used in the analysis are fully efficient for M_{jj}>40 GeV/c^2, as seen from Fig. 10. Three template distributions are used in the fit:

- **Electroweak background distribution:** Based on the shape of the sum of the electroweak backgrounds and (very small) top contribution. The electroweak background is not constrained.

- **QCD distribution:** The QCD template is based on the fit to the dijet mass distribution obtained from the QCD model (shown in Fig. 16). The shape and the normalisation of QCD are treated as nuisence parameters in the fit, they are constrained with gaussian priors to there independantly measured values. The normalisation of the template has a central value of μ=6144 events and σ=1228 events. The slope of the of the exponential fit to M_{jj} has a central value of μ=-0.020 and σ=0.004. Further details can be found in Section 4.1.3).

- **Signal distribution:** The signal distribution is comprised of WW, WZ and ZZ. Due to the jet resolution there is significant overlap between the W and Z peaks, for this reason we do not distinuish the individual contributions. The template for signal is obtained through a fit of a gaussian + polynomial to our WW + WZ + ZZ MC; it is shown in Fig. 22.

The signal extraction fit has five parameters: normalization of the EWK background template, normalization of the QCD background template (with Gaussian constraint), slope of the QCD exponential (with Gaussian constraint), normalization of the signal template, and the jet energy scale. The last one jet is treated simply as a multiplicative factor applied to the mean of the gaussian used to fit the signal template and the width of this gaussian. The change in the invariant mass distribution wth the jet energy scale is shown in Fig. 23. We note that the width of this distribution varies linearly with the jet energy scale, therefore we will use this variation when we float the jet energy scale in the final fit.

The mathematical expression of the likelihood is shown in Eq. (2). From the pseudo-experiment studies, we expect a statistical sensitivity of approximately 6σ. Figure 24 shows the pulls for our fitting procedure as determined from the pseudo-experiments using the expected number of events for signal and background. As expected from a simple 1-dimensional fit, there is no bias and the width of the distribution is ∼1. Figure 25 shows the distribution of extracted signal events from pseudoexperiments generated with 1,516 signal events, 36084 electroweak background events and 7294 QCD events. The expected uncertainty of the measurement is ∼234 events. The significance distribution (i.e., fitted # of signal events divided by its error) from the pseudo-experiments is shown in Fig. 26.

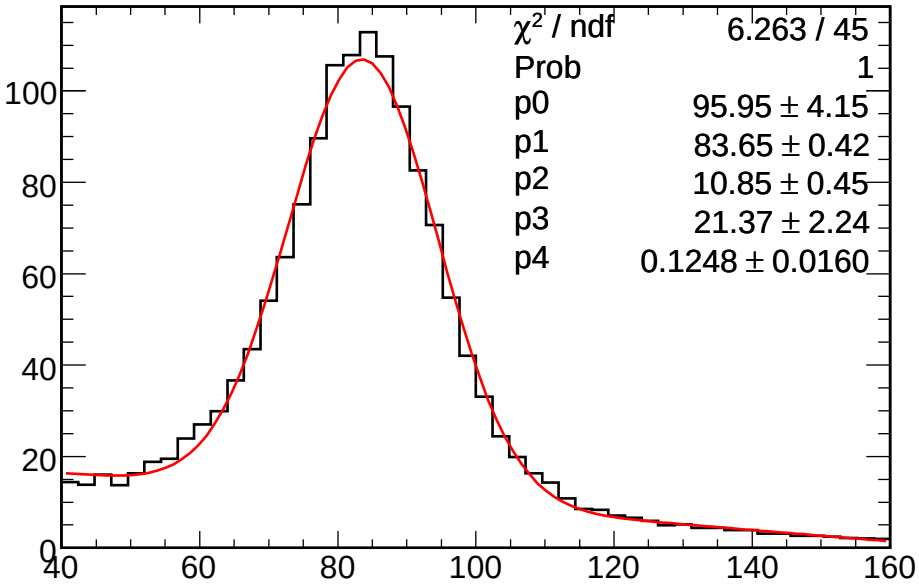

Figure 22: Template used for the signal shape in the fit

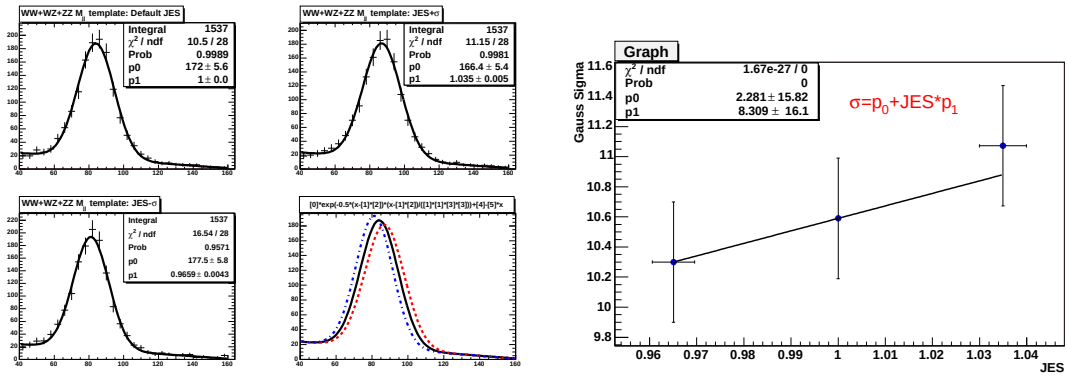

Figure 23: Left: The signal distribution for $\pm 1\sigma$ in jet energy scale. Right: Width of the signal shape vs. jet energy scale variation.

$$L = (N_S * PDF(x|JES) + N_{EWK} * PDF(x) + N_{QCD} * PDF(x|\alpha)) \times G(JES|1, \sigma_{JES}) \times G(\alpha|\mu_\alpha, \sigma_\alpha) \times G(N_{QCD}|M_{QCD}, \sigma_{QCD})$$

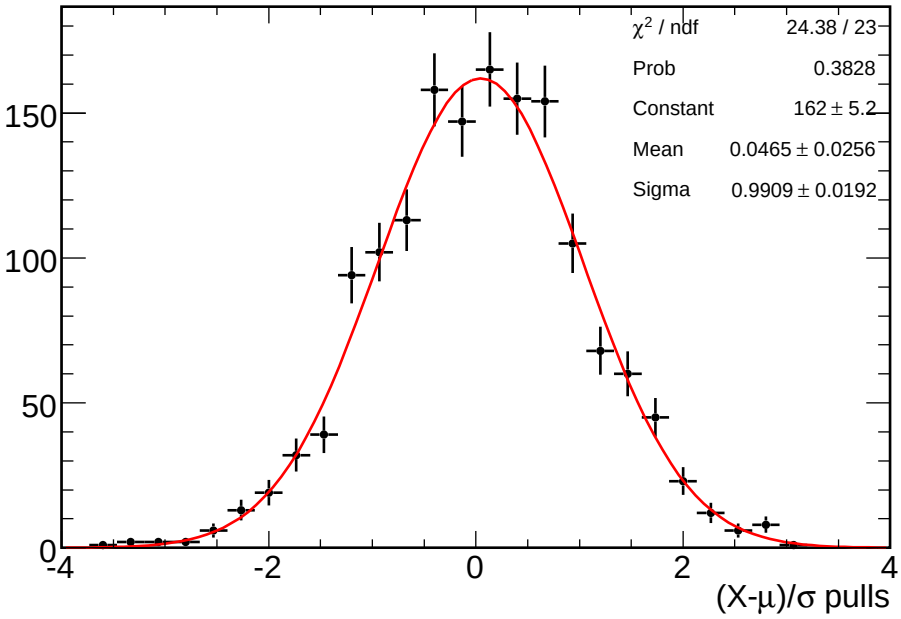

Figure 24: Pulls for the fitting procedure from 1500 pseudo-experiments.

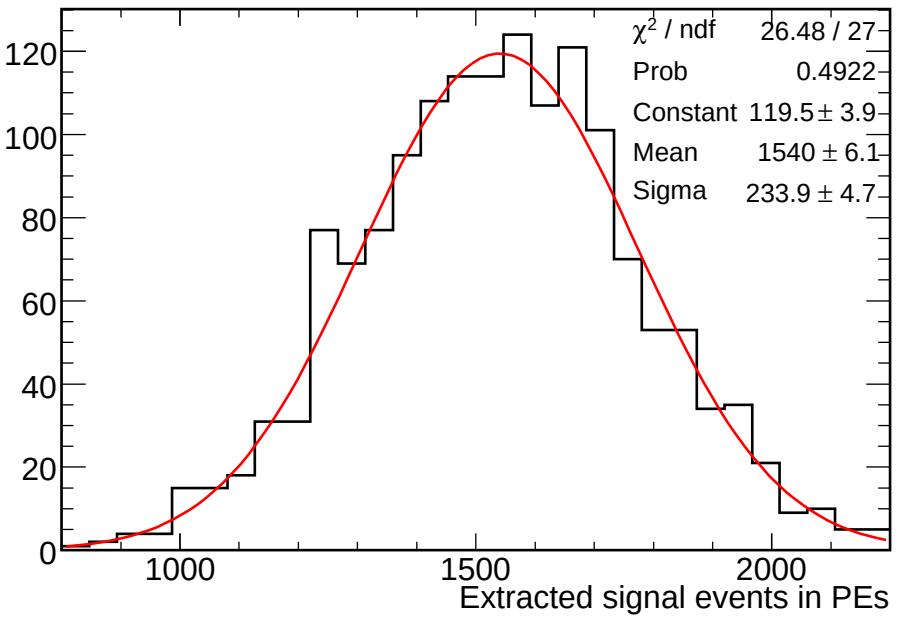

Figure 25: Extracted number of signal events from 1500 pseudo-experiments.

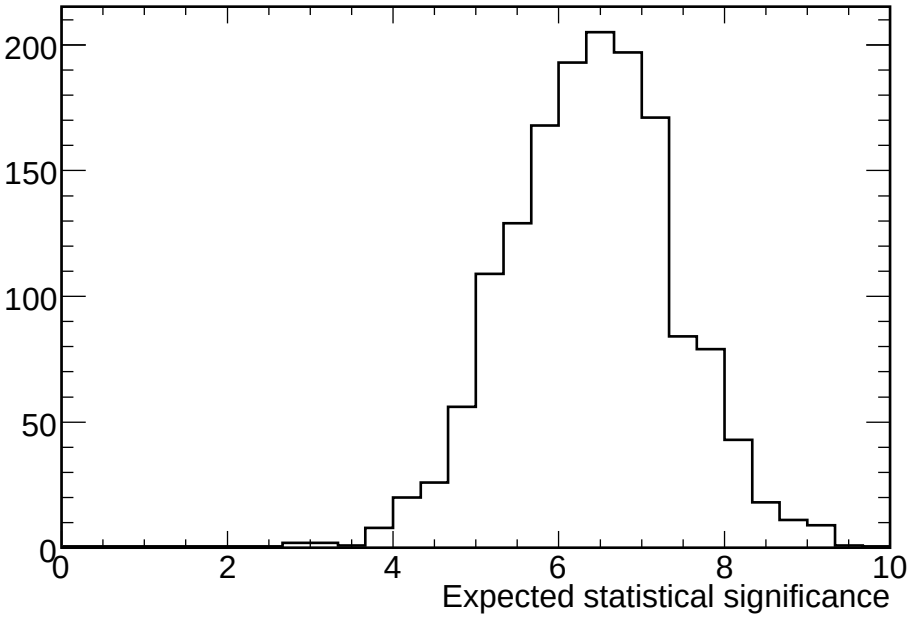

Figure 26: Significance from 1500 pseudo-experiments.

5.1 Results

The procedure described above is applied to the data that passed the selection cuts to extract the diboson signal. The final fit to M_{jj} (top plot) and *data-background* (bottom plot) are shown in Fig. 27. The results of the fit for all free parameters are summarized in Table 5. The correlation matrix for these parameters is shown in Table 6. In summary, we extract 1,516±239 signal events out of 44,910 selected data events. Both the extracted number of signal events and its uncertainty agree very well with the expected numbers, 1398±234 (234 is based on pseudo-experiments). The $\chi^2/n.d.f$ for the final fit is 9.6/9, which corresponds to 38% probability. The $\chi^2/n.d.f$ qouted here is based on figure 27 (bottom) and assumes the degrees of freedom are independent.

As expected from this simple fit, the likelihood profile vs. number of signal events has a well behaved parabolic shape as shown in Fig. 28.

Since the fit returns numbers for the signal and background close to the expectation it is not surprising that all the kinematic distributions are similar to the ones obtained based on expected signal and background. For completeness we show them in Fig. 29 to 32.

6 Systematics

The systematic uncertainties are separated into three categories: *1)* uncertainties affecting the signal extraction; *2)* uncertainties affecting the signal cross section; *3)* and the standard

Floating Parameter	Fitted Value +/- Uncertainty
QCD slope	0.724±0.047
JES	0.985±0.019
# EWK background	36,140±1,230
# QCD background	7,249±1,130
# Diboson signal	1,516±239

Table 5: Results of the fit for all floating parameters. We extact 1,516 signal events from the 44,910 data events which pass our selection cuts.

Fit Parameter	QCD slope	JES	# EWK	# QCD	# Sig
QCD slope	1	0.212	-0.419	0.437	0.062
JES	–	1	-0.010	0.037	-0.116
# EWK	–	–	1	-0.967	-0.382
# QCD	–	–	–	1	0.206
# Sig	–	–	–	–	1

Table 6: Correlation matrix of the fitted parameters.

CDF luminosity uncertainty of 6%.

6.1 Systematics on the Signal Extraction

The shape systematics for the extraction are determined by using appropriately modified distributions in the extraction and looking at the effect these modifications have on the signal extraction. The shape systematics considered are described below; their values are shown in Table 7.

- **Electroweak background shape:** the uncertainty on the shape of the EWK background is determined by replacing the EWK MC pdf with an alternative pdf derived from γ-jet data. This is discussed in detail in Sec. 7. Examples of fits using the default EWK template and γ+2jet template can be found in Fig. 33. The comparison of these two templates can be found in Fig. 34.

- **QCD shape:** The slope of QCD distribution is one of the fit parameters. Therefore, the effect of QCD shape variations is not considered as a separate source of systematic uncertainty because it is folded in the statistical uncertainty on the signal obtained from the fit.

- **jet energy resolution:** The dijet mass template for signal is smeared according to the uncertainty on jet energy resolution (width is increased by ∼11%, see Figs. 42, 44 in Section 8.1), this results in a 5.6% change in the expected number of signal events.

- **Jet energy scale:** the jet energy scale is allowed to float for the signal template in the extraction fit. Given the fact that JES is clearly not an issue for the QCD template, and the fact that the EWK background systematic is calculated relative to a data-based template, this should suffice to cover the effect of JES on the signal extraction. Therefore, the effect of JES uncertainty is folded in the statistical uncertainty on the signal obtained from the fit and is not considered as an independent source of systematic uncertainty.

The individual source of systematics are added in quadrature. The total systematic uncertainty on the signal extraction is 9.5% or 144 events. Therefore, the final result for the diboson signal is 1,516±239(stat)±144(syst) events.

Syst type	% Effect
Ewk template	7.7%
Resolution	5.6%

Table 7: Systematics related to template shapes and, therefore, the signal extraction.

6.2 Systematics on the cross section

In addition to the systematics arising from the background shapes, there are additional systematic uncertainties that effect the cross section measurement. This includes the luminosity uncertainty systematic of 6% ; however it also includes systematics related to the signal acceptance. The sources of uncertainty on the acceptance that were considered:

- **jet energy scale:** the uncertainty on jet energy scale (JES) affects the efficiency of the following event selection cuts: E_T^{jet2}-cut, N_{jet20}-cut, $\not\!\!E_T$-cut, cuts on $\not\!\!E_T$-significance and $\Delta\phi_{closest}$. To estimate a combined effect of JES uncertainty on the acceptance, we shift JES by $\pm\sigma_{JES}$ using the official Jet Corrections package.Then, we apply our selection procedure and take a maximum difference between the number of selected events in the deviated distribution and the default distribution. This has an 8% effect on the expected number of signal events.

- **jet energy resolution:** The dijet mass template for signal is smeared according to the uncertainty on jet energy resolution (width is increased by ∼11%, see Fig. 42 in Section 8.1), this results in a 0.7% uncertainy in the cross section.

- $\not\!\!E_T$ **model:** the $\not\!\!E_T$ model is used to determine $\Delta\phi(closest)$ and the signed $\not\!\!E_T$ significance. There is a 1.0% uncertainty associated with using the model to determine these variables (see Section 8).

- **Trigger efficiency:** We assign a 2.2% uncertainty as described in Sec. 3.

- **ISR/FSR:** Due to lack of signal MC w/ more/less ISR/FSR, we take half the difference of the acceptance shift in the $ZH \to \nu\nu bb$ MC more/less ISR/FSR MC samples; this yields a systematic of 2.5%

- **PDF:** Take this systematic to be 2%, like in the $\not\!\!E_T$+bb Higgs analysis

- **Luminosity:** There is the standard 6% uncertainty on the luminosity.

All systematic uncertainties are summarised in Table 8:

Systematic	% uncert.	# of signal events
Extraction		
EWK shape	7.7	117
Resolution	5.6	85
TOTAL EXTRACTION	9.5	144
Acceptance		
JES	8	121
JER	0.7	11
$\not\!\!E_T$ model	1	15
Trigger inefficiency	2.2	33
ISR/FSR	2.5	38
PDF	2	30
TOTAL ACCEPTANCE	9.0	136
Lumi	6	91
TOTAL SYSTEMATIC	14.4	218

Table 8: Summary of the systematic uncertainties.

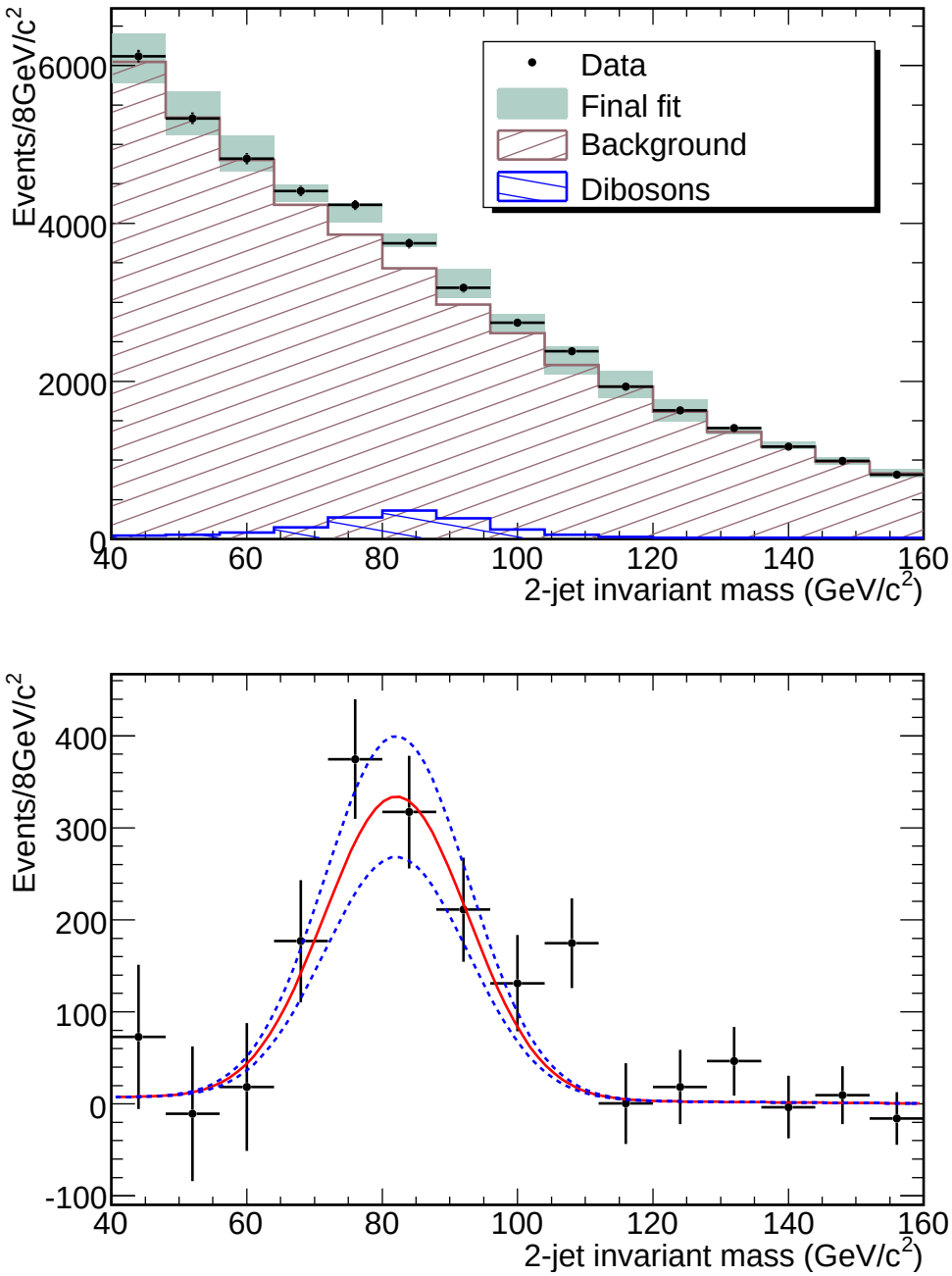

Figure 27: *Top plot:* Signal extraction fit with the jet energy scale allowed to float. The fit uncertainty (brown band) includes systematic uncertainties on shapes of QCD and EWK backgrounds. *Bottom plot:* The difference between data and the total predicted background. The red (solid) curve illustrates the expected signal and the blue (dashed) lines represent $\pm\sigma$ variations due to statistical uncertainty on signal (extracted from fit).

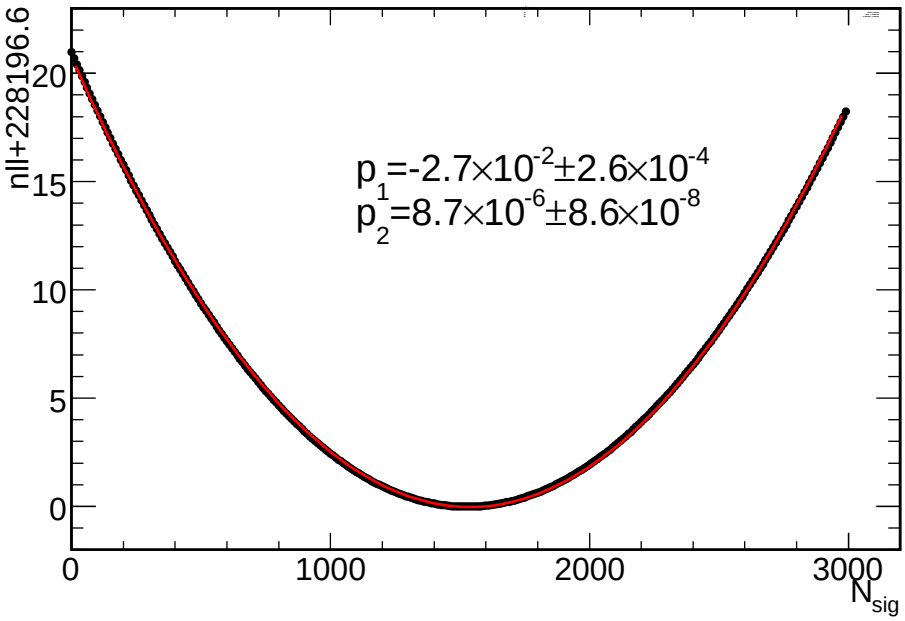

Figure 28: Negative Log Likelihood profile vs. nmber of signal events.

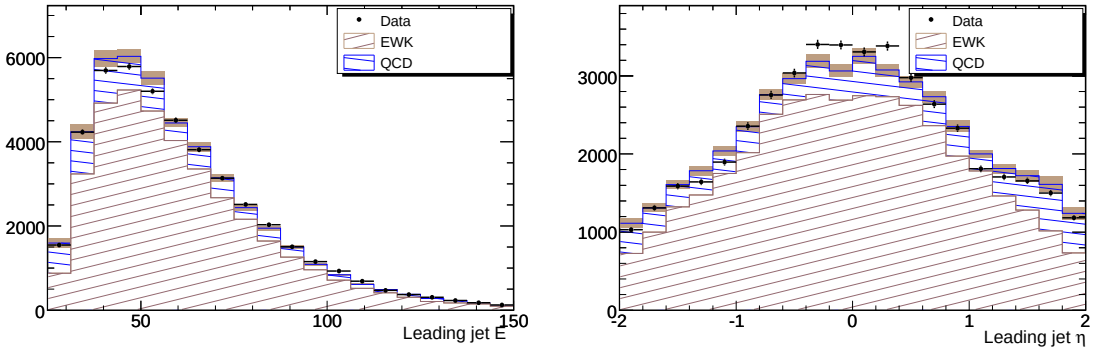

Figure 29: E_T (top) and η (bottom) of the first jet based on the fit results. The uncertainty on the total prediction (brown shaded area) includes only 16% uncertainty on the QCD normalization as returned from the fit.

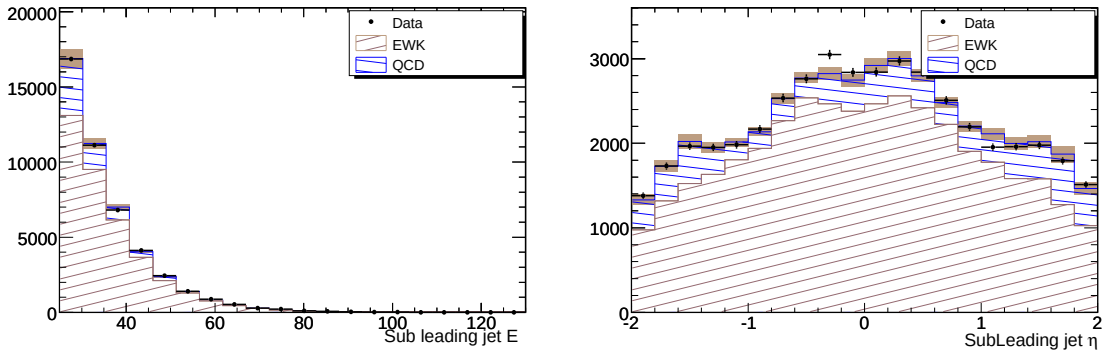

Figure 30: E_T (top) and η (bottom) of the second jet based on the fit results. The uncertainty on the total prediction (brown shaded area) includes only 16% uncertainty on the QCD normalization as returned from the fit.

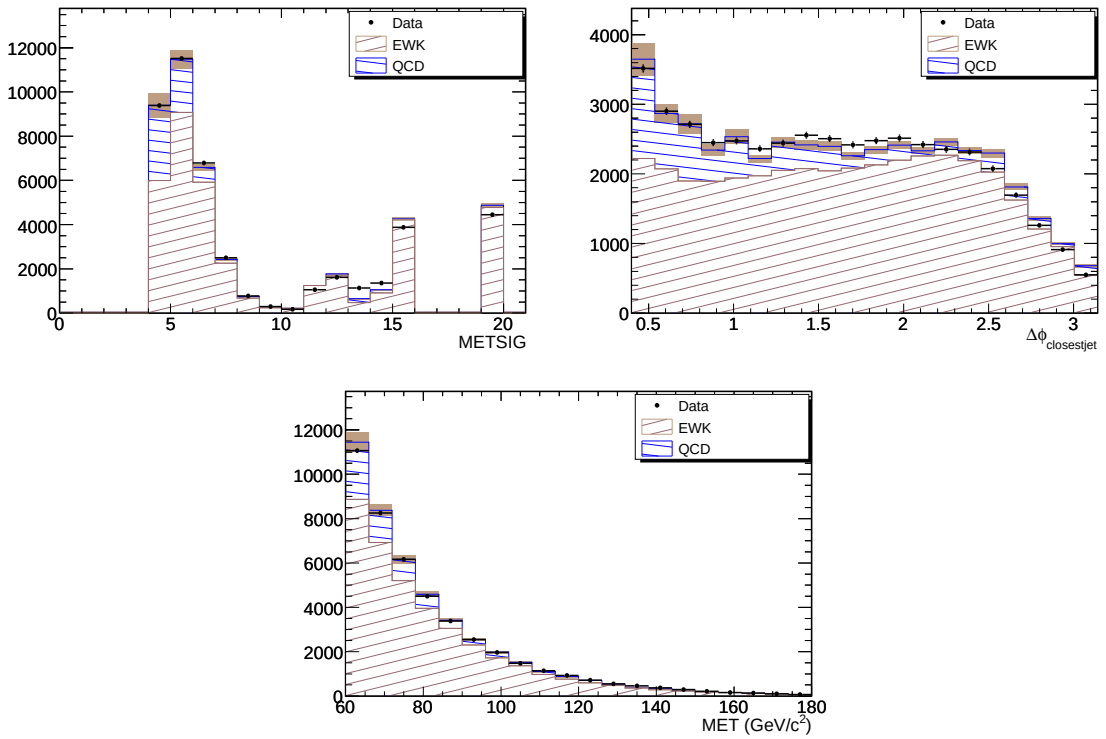

Figure 31: $\not\!\!E_T$-significance (top), $\min(\Delta\phi(\not\!\!E_T - jet))$ (middle), and $\not\!\!E_T$ (bottom) based on the fit results. The uncertainty on the total prediction (brown shaded area) includes only 16% uncertainty on the QCD normalization as returned from the fit.

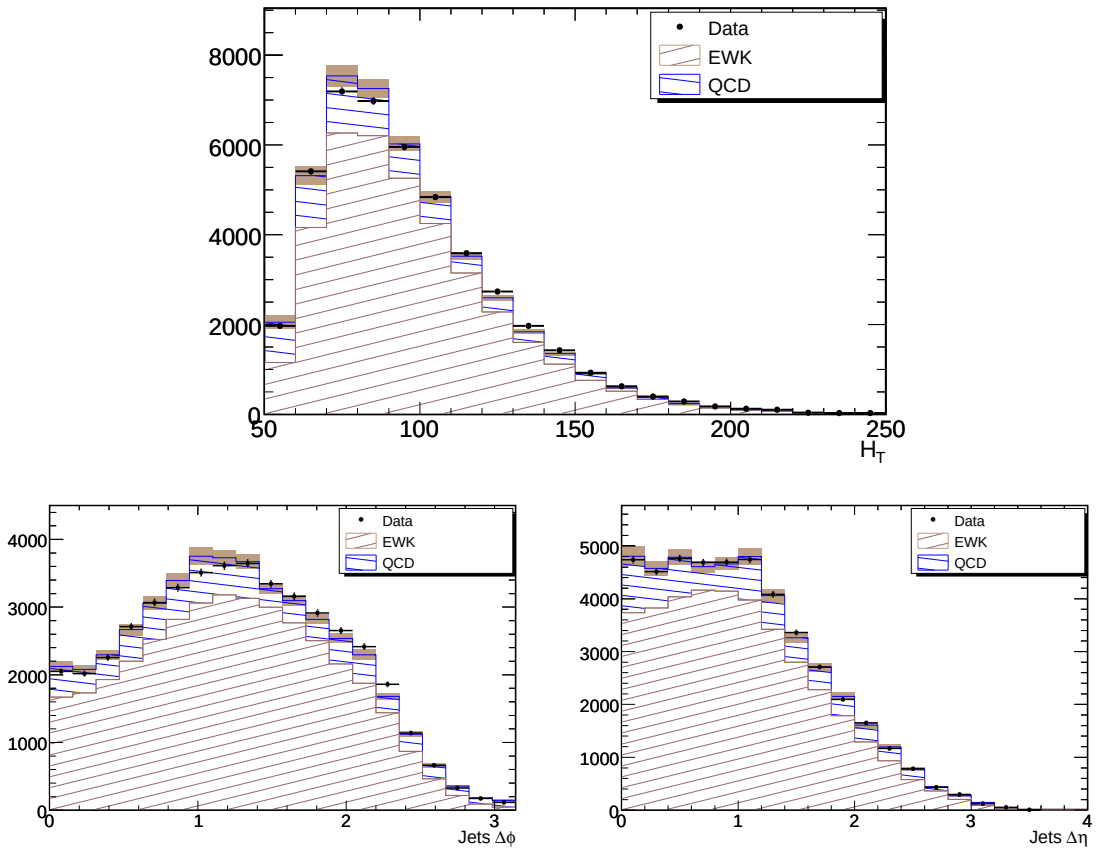

Figure 32: H_T (top), $\Delta\phi = \phi_{jet1} - \phi_{jet2}$ (bottom left), and $|\Delta\eta| = |\eta_{jet1} - \eta_{jet2}|$ (bottom right) based on the fit results. The uncertainty on the total prediction (brown shaded area) includes only 16% uncertainty on the QCD normalization as returned from the fit.

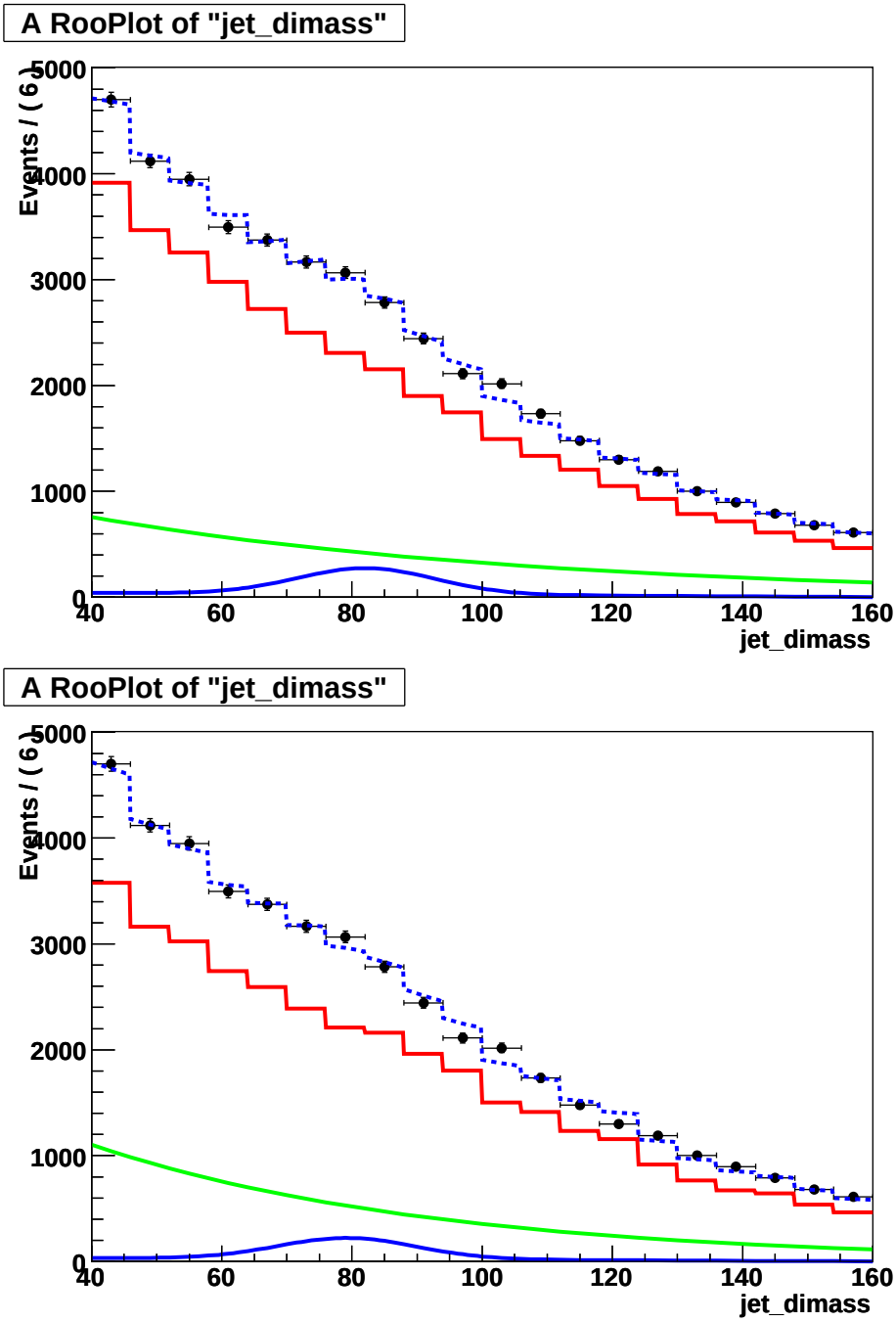

Figure 33: Signal extraction using the default EWK template (top) and the γ+2jets (bottom) template.

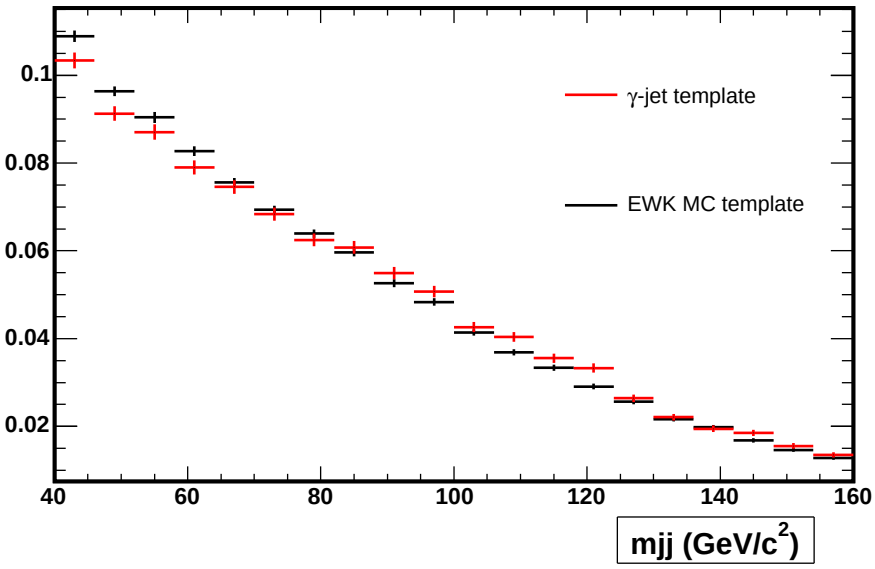

Figure 34: Comparison of the default EWK (black) and the γ+2jets (red) templates.

7 Electroweak Shape Systematic

7.1 γ+ Jets

In fitting our data for the number of signal events, three different templates are used: one for the signal itself, one for the QCD events in the sample, and one for the electroweak events. Systematics need to be assigned to all three; treatments of signal and QCD are given above. As the electroweak background shape is constructed purely out of MC events, all the usual uncertainties come into play: PDFs, generators, jet energy scales and resolutions.

To estimate the effect of these uncertainties, we fit for the number of signal events using a template constructed not from electroweak MC events, but from γ+jets data events from the **cph10x** inclusive photon data samples. Here, the MC uncertainties don't apply. The motivation behind using inclusive photon data is as follows: all our major non-QCD backgrounds consist of a gauge boson (W or Z) plus jets. As a neutral gauge boson, the photon possesses exactly the same interactions as the Z at the matrix element level. While this is not the case with the W due to its charge, this shouldn't cause a great difference in the kinematics between W and γ events. While it is true that there aren't *perfect* parallels between γ+jets and W +jets or Z+jets events – the most glaring difference, clearly, is the mass of the bosons – this is accounted for by a reweighting described in the following section.

Despite similarities at the matrix element level between the processes, when it comes to event reconstruction, clearly the $\not\!\!E_T$ will be smaller on average for γ events, whose energies are deposited in the calorimeter, than for events where a W or Z decays to undetected neutrinos. Very few γ+jets events, therefore, will pass our standard selection cuts. This can be dealt with if we draw a parallel between the γ and the Z decaying to neutrinos or the W decaying to a neutrino and a missed or poorly reconstructed charged lepton. We can take the γ E_T as a stand-in for the E_T of the W or Z, and therefore take the vector sum of the γ E_T and the (purely QCD) $\not\!\!E_T$ in the γ+jets event as a stand-in for the $\not\!\!E_T$ in our data sample, replacing the requirement of $\not\!\!E_T > 60$ GeV with the requirement that it be the vector sum of the γ and $\not\!\!E_T$ to have an $E_T > 60$ GeV. Some other differences exist in the selection cuts applied to γ+jets vs. $\not\!\!E_T$ plus jets data as well; these are summarized in Tab. 9. The jet dimass distribution of the γ+jets data is shown in Fig. 35. Other options were considered; however, a view of Figs. 36 and 37 indicate that the shapes of the γ+jets data distribution are actually quite similar under different selection cuts.

7.2 Data weighting

While it is the case that there are similarities between the γ+jets processes and processes involving heavy gauge bosons plus jets, their jet kinematics aren't identical, and for this reason we don't directly replace the electroweak MC background template with the γ+jets data template shown in Fig. 35. Rather, we reweight the γ+jets data to the electroweak MC templates in the following manner: we take the ratio of dijet mass in our electroweak MC to dijet mass in the γ+jets MC (Pythia sample **pq0sj0**). This is done separately for our five main backgrounds (dijet + $W \to e\nu$, $W \to \mu\nu$, $W \to \tau\nu$, $Z \to \tau\tau$, or $Z \to \nu\nu$), and

$\not\!\!E_T$+jets	γ+jets
$\not\!\!E_T > 60\,\text{GeV}$	$\lvert\vec{\not\!\!E}_T + \vec{E}_{T\,photon}\rvert > 60\,\text{GeV}$
$\Delta\phi_{closest} > 0.4$ wrt $\not\!\!E_T$	$\Delta\phi_{closest} > 0.4$ wrt $\vec{\not\!\!E}_T + \vec{E}_{T\,photon}$
–	photon: tight JP plug/central cuts, iso $E_T < 1\,\text{GeV}$
–	$\Delta R_{photon-jet} > 0.7$
$\not\!\!E_{T\,significance}^{signed} > 5$	–
$\Delta R_{lepton-jet} > 0.2$	–
$\frac{EM}{E_{TOT}} < 0.85$	–
$190697 \leq$ run # ≤ 267718	$141544 \leq$ run # ≤ 266513

Table 9: List of differences between cuts applied to the $\not\!\!E_T$+ jets vs. γ+jets sample

is done using histograms in 10 bins from 40 to 160 GeV in dijet mass, creating five different weightings, as shown in Fig. 38. The other electroweak backgrounds contain negligible event contributions, and are therefore not used in weighting.

These weightings are applied separately to the γ+jets data, in each case creating a weighted dijet mass distribution which is then normalized based on the expected amount of the electroweak background to which it has been weighted. A comparison of these backgrounds with their corresponding weighted γ+jets data is shown in Fig. 39. Additionally, the sum of the electroweak MC backgrounds is shown compared to the sum of the weighted and normalized photon plus jets background in fig. 40. In all cases, the weighted data template has a slightly less steep falloff in dijet mass than the MC template.

7.3 Sample Purity

One concern with using a γ+jets data template as a replacement for background MC is the effect of signal contamination (W/Z $\rightarrow$ jets plus a photon) in the template. This signal contamination in the template will reduce the number of signal events which get fitted when the template is used in the fit. Given 3.4 fb^{-1} of **cph1xx** data, and taking the cross sections and acceptances of our LO Pythia $W/Z \rightarrow$ jets $+\gamma$ MC along with an NLO K-factor of 1.4 [Note 7996], we predict 598 signal events out of the 58416 γ+jets data events. Scaling the number of signal events by the ratio of the number of expected electroweak events (37089) in our $\not\!\!E_T$ plus jets sample to the events in our γ+ jets sample, (58416), means we should correct for a deficit of 378 signal events when fitting with the γ+jets data template rather than the electroweak MC template. Finally, another contamination of the γ+jets data template comes from QCD; it is for this reason that, beyond using tight Joint Physics cuts on the photons in the same, we also require the isolation energy in the cone of 0.4 about the photon to be less than 1 GeV; this is a requirement considerably tighter than the Joint Physics prescription, and will give us a sample with a low expected level of QCD contamination.

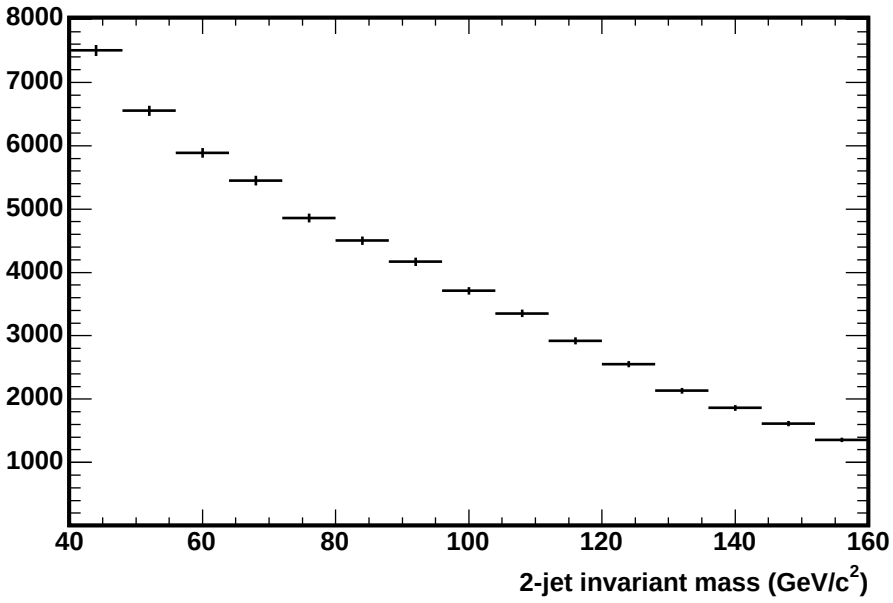

Figure 35: Jet dimass distribution for γ+jets data before weighting is applied

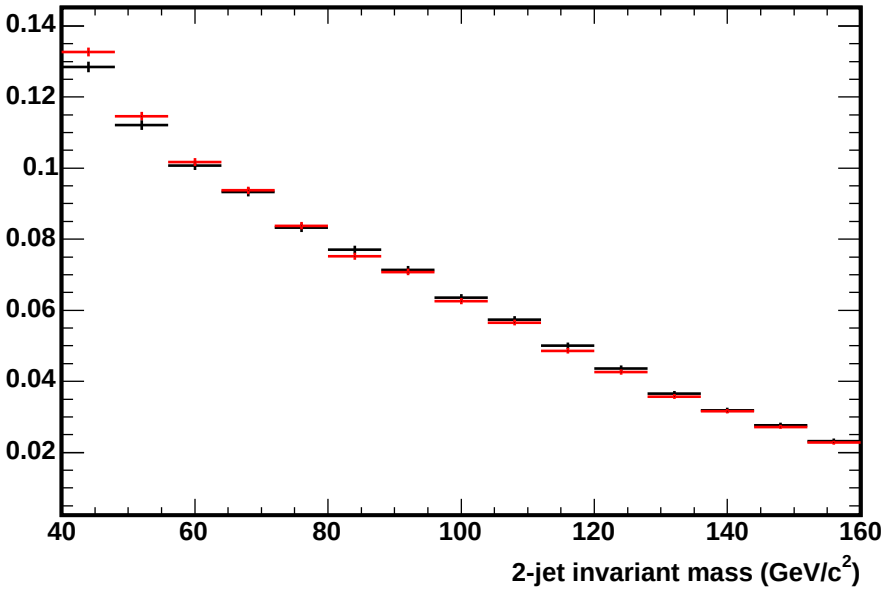

Figure 36: Jet dimass distributions for γ+jets data under normal selection cuts (black) and using the (looser) Joint Physics cut on the photon isolation energy (red)

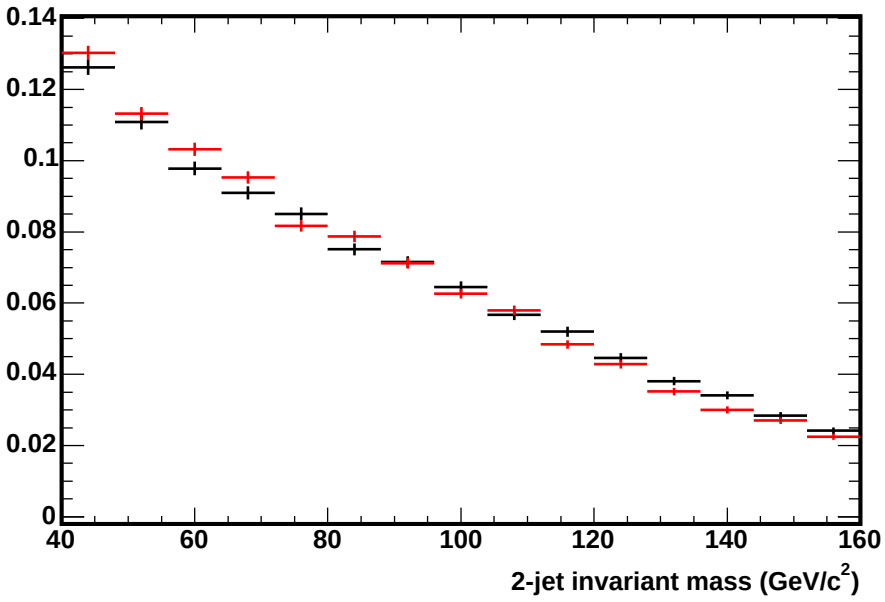

Figure 37: Jet dimass distributions for γ+jets data under normal selection cuts, divided into central photons (black) and plug photons (red)

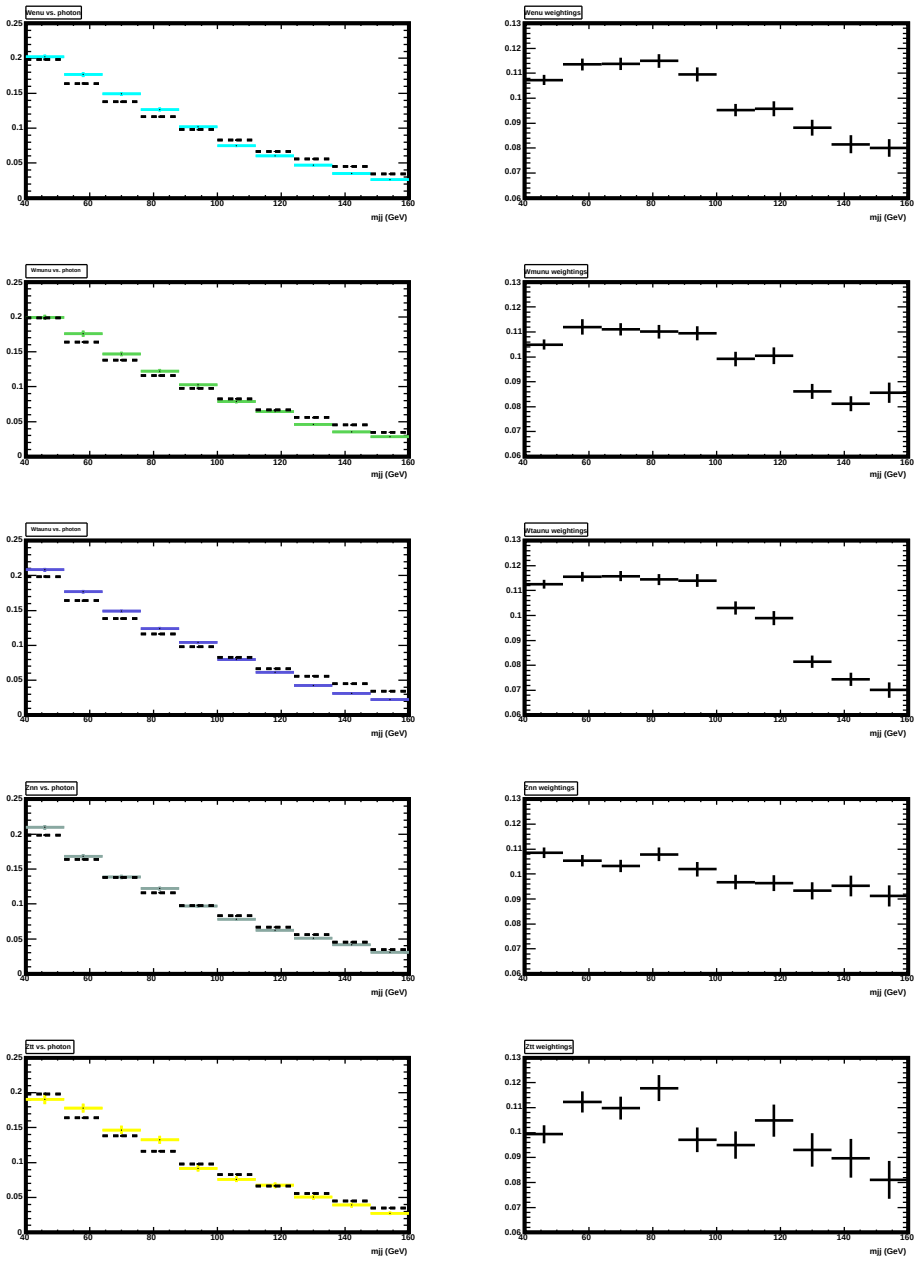

Figure 38: Left column: superposition of jet dimass for γ+jets MC (dashed) and an electroweak background MC (from top: W $\rightarrow e\nu$, W $\rightarrow \mu\nu$, W $\rightarrow \tau\nu$, Z $\rightarrow \nu\nu$, Z $\rightarrow \tau\tau$). Right column: the weightings obtained from the ratio of the corresponding distributions in the left column.

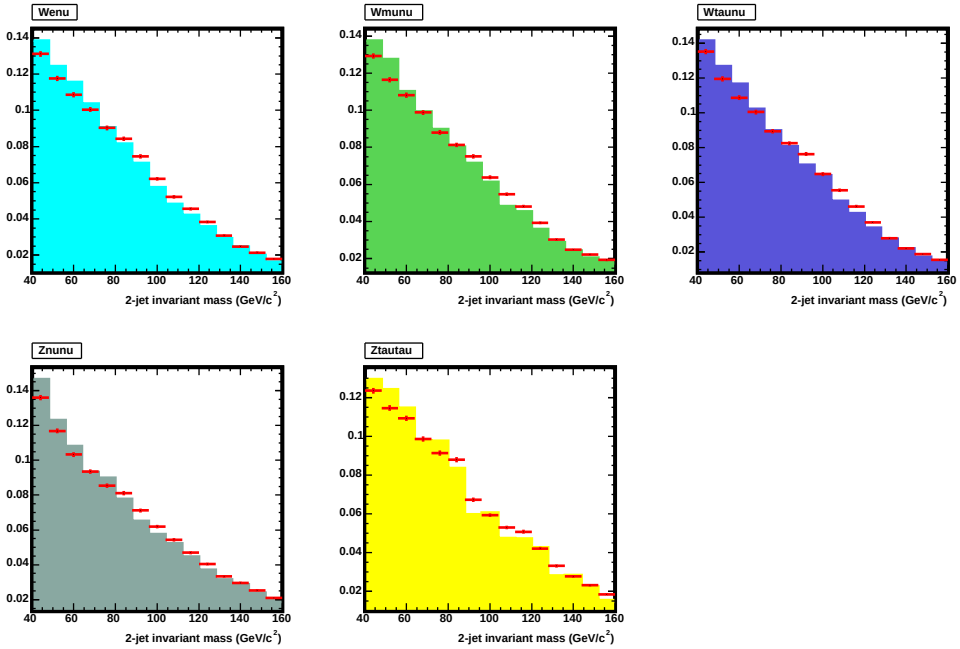

Figure 39: Comparison of the weighted photon plus jets data (markers) with the electroweak background MC samples to which they've been weighted (solid histograms)

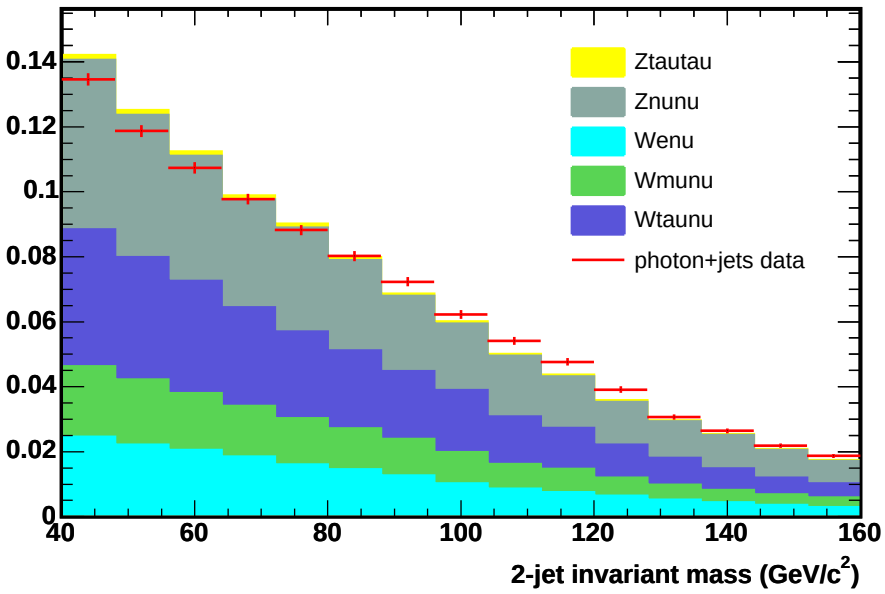

Figure 40: Comparison of the combined electroweak background MC samples, with the combined, weighted $\gamma+$ jets data; it is this latter distribution which we use to fit for our shape systematic

	WW	WZ	ZZ
Energy resolution	0.97%	0.95%	0.82%

Table 10: Systematic uncertainty on the signal acceptance associated with the *Met Model* parametrization. "Energy resolution" is the combined effect of uncertainties in jet and unclustered energy resolution parameterizations on calculation of the $\not\!\!E_T$-significance.

8 Signal Acceptance: Uncertainties Associated with Met Model

We use Pythia inclusive di-boson (WW, WZ, and ZZ) stntuple samples (described in Section 2.2) to study the systematic uncertainties associated with the *Met Model* (due to $\not\!\!E_T$-significance cut). We consider the following two sources of uncertainties which affect only the calculation of $\not\!\!E_T$-significance. The width of JER distributions used in the *Met Model* parameterization (described in note [3]) is varied by ∼5-7% (depending on jet E_T) to account for the data-MC differences in JER. The unclustered energy resolution is obtained from $Z \to e^+e^-$ events in data and MC (described in note [3]). The $\not\!\!E_T$-significance is calculated using both parameterizations and the difference in numbers of selected events is taken as an estimate of the corresponding systematic uncertainty. The summary of the systematic uncertainties discussed above is given in Table 10. We claim 1% total uncertainty on the signal acceptance associated with the $\not\!\!E_T$-significance cut.

8.1 Jet Energy Resolution Systematics

The jet energy resolution (JER) parameterization was obtained by comparing detector and hadron level jets in MC (see note [3] for details), so-called response method. This approach allows to remove physics effects (ISR/FSR and out-of-cone energy) and study only detector effects on JER. To validate JER and study systematics due data-MC differences, we still need to resort to a comparison of JER obtained using the bi-sector method. For this purpose, we study the di-jet (for high E_T) and Z-jet (for low E_T) balance in data and MC events. We adopt the following approach. We fit the di-jet (Z-jet) balance in data by a simple resolution function: $\sqrt{p_0/E_T + p_1/E_T^2 + p_2}$. As one can see from Fig. 41, this fit function has a very good description of data in the entire energy range. We also study systematic uncertainties on JER associated with the following effects: *1)* pile-up (compare JER in events with N_{vx}=1 and N_{vx}>1); *2)* $\Delta\phi_{jj}$ or $\Delta\phi_{Zj}$ cut (default cut of 2.7 rad is varied by ± 0.2 rad); *3)* compare JER in events with no more than 2 (default) or 1 (tight) extra jets with lev-4 energy satifying E_T^{extra}<2.0+0.075∗M_{jj} or E_T^{extra}<2.0+0.075∗M_{Zj}; *4)* compare JER obtained wiht the default and $\sqrt{p_0/E_T + p_1}$ fit functions. These JER uncertainties are illustrated at Fig. 42. Finally, we compare our parametrization of the bi-sector JER in data with the results obtained using Pythia Gen-6 Z-jet events. From Fig. 43, one can see that MC is well described by the data parameterization within quoted systematic uncertainities. Therefore, we claim that the

data-MC difference in the jet energy resolution is within 5-7%, depending on E_T (as shown in Fig. 42).

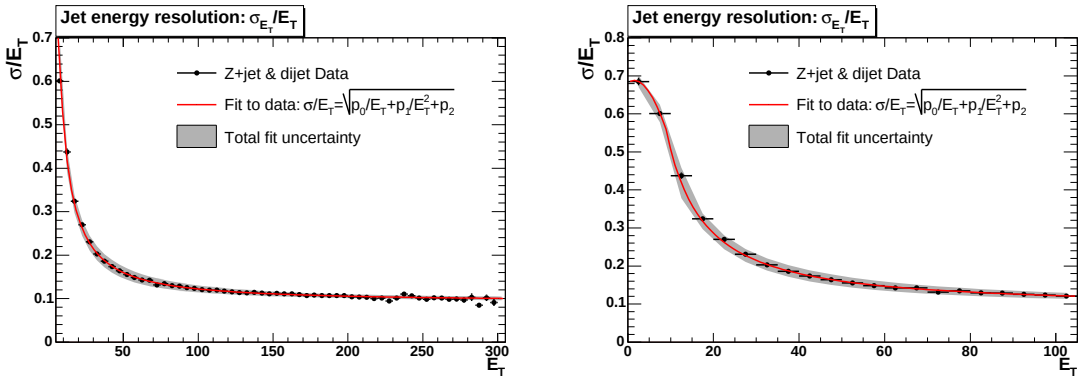

Figure 41: The jet energy resolution obtained with the bi-sector method in data di-jet and Z-jet events. Two plots show JER in the 0-300 GeV E_T-range (left) and a zoom-in in the 0-100 GeV E_T-range (right). Data (points) are fitted with $\sqrt{p_0/E_T + p_1/E_T^2 + p_2}$ function (red line). The grey band illustrates the associated uncertainty. For a better description of the low E_T region (E_T<10 GeV), the original fit function is replaced by a Gaussian.

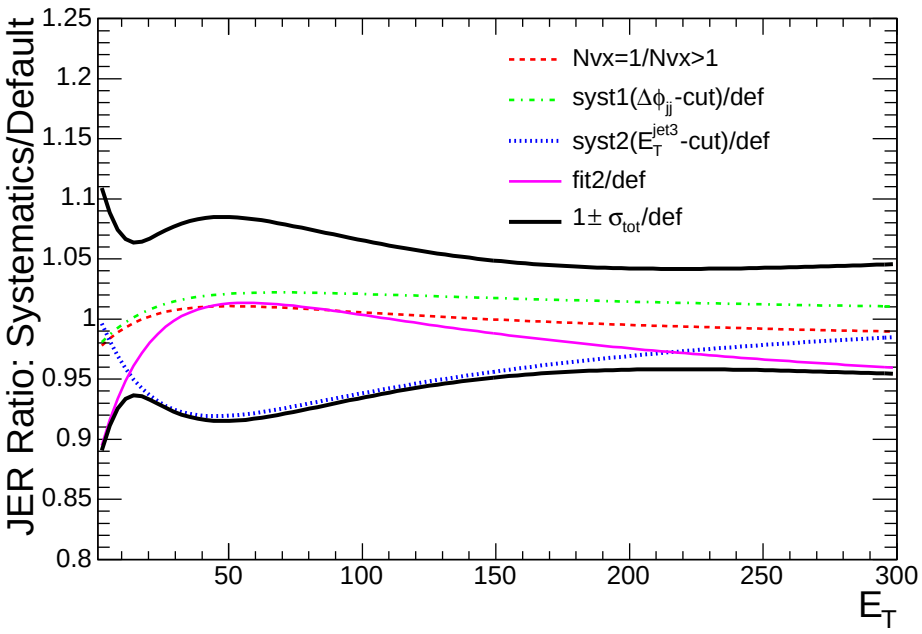

Figure 42: Illustration of the systematic uncertainities on the jet energy resolution obtained with the bi-sector method in data di-jet and Z-jet events. The total uncertainty, default JER$\pm\sigma_{JER}$, is shown by solid black line.

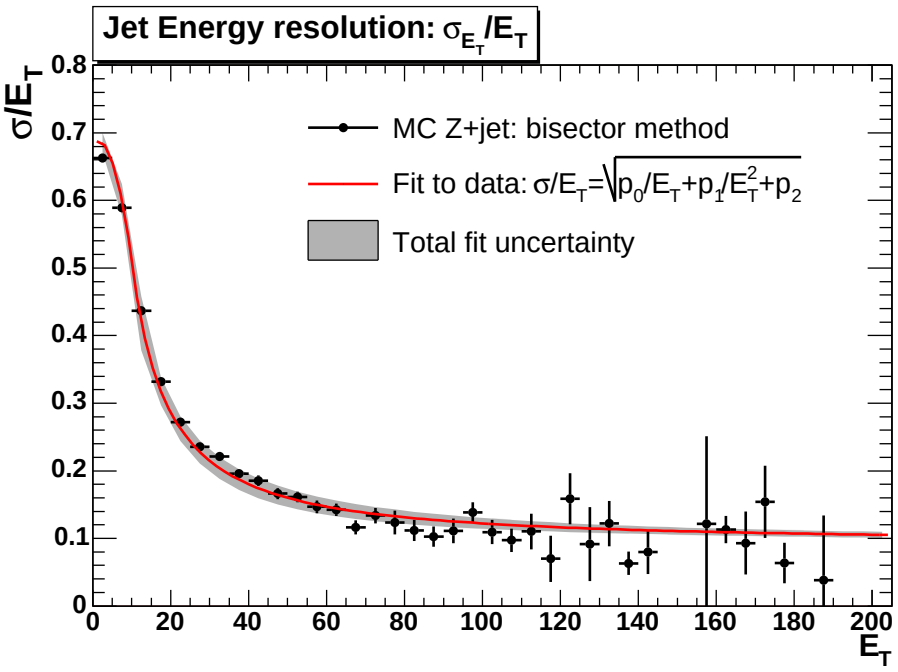

Figure 43: Comparison of the bi-sector JER in Pythia Z-jet events (points) with the JER parameterization obtained in data (red curve). The grey band represents the uncertainty on JER parameterization.

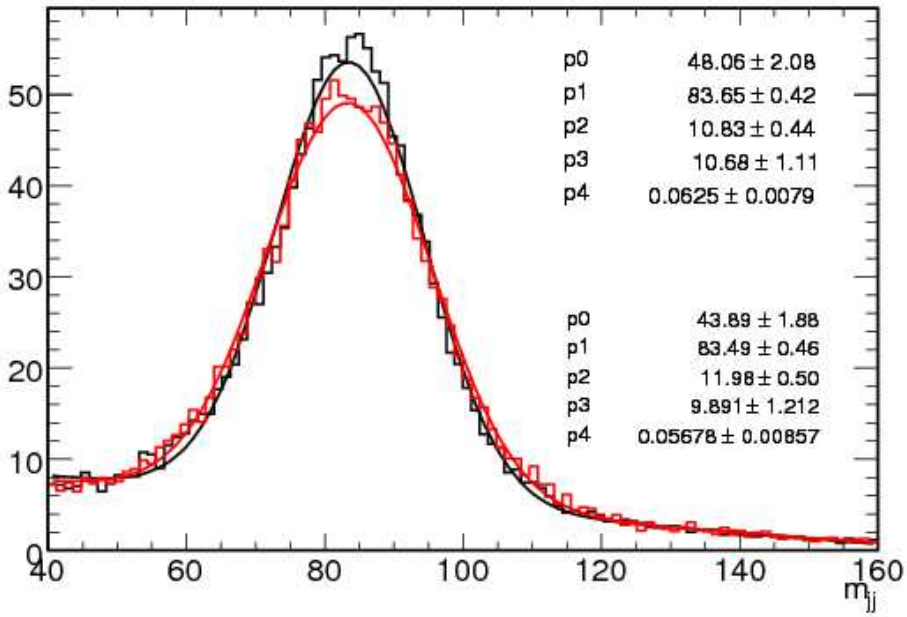

Figure 44: change in signal distribution after resolution function smearing

9 Conclusion

We expect 1416 signal events and find

$$1516 \pm 239(stat) \pm 144(sys),$$

the JES uncertainty, the qcd normalization and the QCD shape are treated as nuisance parameters in the fit (within the gaussian constraints imposed through external calibrations) and the uncertainty is therefore included as part of the statistical uncertainty of the fit. When translated to a cross section measurement, we factor in the 98.9(96.0)% a cosmic(trigger) effiencies; this, combined with the acceptances for our WW, WZ and ZZ signal MC (2.48%, 2.64% and 2.94%, respectively) weighted by their theoretical cross sections (11.7 pb, 3.6 pb, 1.5 pb) as well as the acceptance and luminosity uncertainties, yields a final cross section measurement of $18.0 \pm 2.8(stat) \pm 2.4(syst) \pm 1.1(lumi)$ pb.

References

[1] Z. Alawi, W. Ketchum *et al.*, *Search for diboson production in the ll+jets channel*, CDF note in preparation.

[2] M. Goncharov *et al.*, *Photon Timing Distributions*, CDF note 7960.

[3] R. Culbertson, A. Pronko, and M. Goncharov, *Search for Anomalous Production of* $\gamma\gamma + \not\!\!E_T$ *Events in 2* fb^{-1} *of Data.*, CDF note 9184.

[4] T. Aaltonen *et. al*, Phys. Rev. D **78**, 052006 (2008).

A How to report the cross section

For the signal MC we use the diboson top group samples as listed in 2.2. These are pythia samples and the γ^* component is only generated in the ZZ sample and not in the WZ (evidently, this is not a problem for WW). Furthermore, there is a theshold for the mass of the γ^* at 2 GeV. The electroweak group samples have the same problem except that the threshold is at 15 GeV. My understanding is that this is a PYTHIA issue since there is no Wγ^* in the list of diboson processes in the PYTHIA manual. This poses a presentation problem because we do have some acceptance to low mass events as shown in Fig. 45. In order to be consistent, we will calculate based on the MCFM calculator the cross sections only above 40 GeV. The list of cross sections reported by MCFM v5.4 is listed in Tab. 11 together with the acceptances for each process.

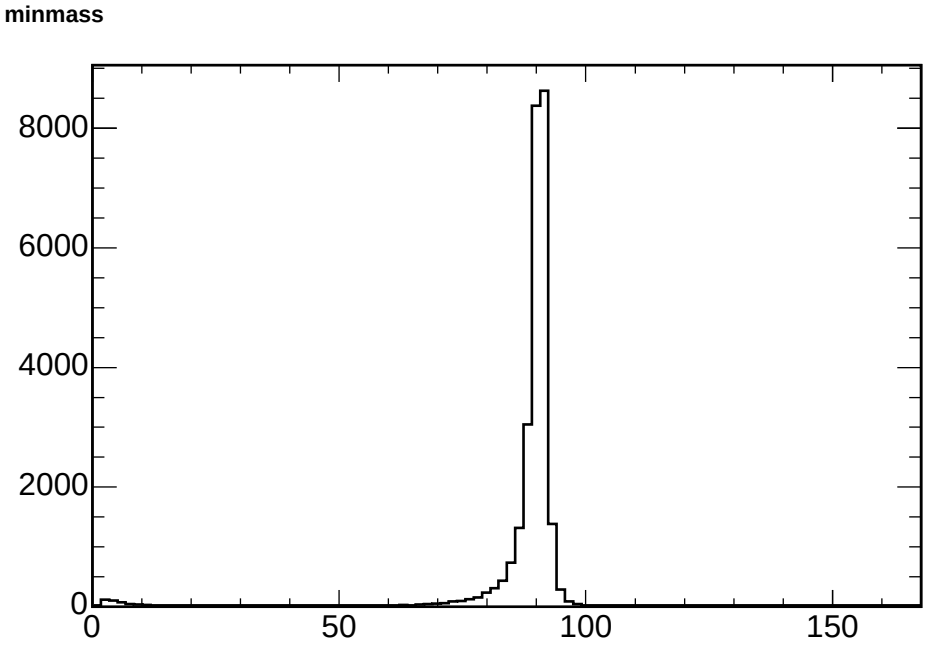

Figure 45: Invariant mass distribution at generated level for accepted events in the ZZ sample.

For completeteness we list at the end of the appendix the input file used for MCFM:

Based on Fig. 45 we know that we have some sensitivity no matter how small to events with the mass of the γ^* below 40 GeV. We can infer based on our PYTHIA MC how many events are due to this effect and turns out to be 0.1% for WW, 1.3% for WZ and 2.3% for ZZ. Given the expected number of events from Tab. 2 that translates into 9 events or a -0.6% correction to the cross section measured.

```
'5.4'[file version number]
```

Process	Cross Section	Acceptance
WW	11.70 ± 0.7	2.48%
WZ	3.60 ± 0.3	2.64%
ZZ	1.49 ± 0.2	2.94%

Table 11: MCFM v5.4 cross sections for our processes with an invariant mass cut at 40 GeV for either boson. We use CTEQ6M native mcfm pdfs and check with LHAPDF

```
[Flags to specify the mode in which MCFM is run]
.false. [evtgen]
.false. [creatent]
.false. [skipnt]
.false. [dswhisto]

[General options to specify the process and execution]
61 [nproc]
'tota'   [part 'lord','real' or 'virt','tota']
'test'['runstring']
1960d0 [sqrts in GeV]
+1 [ih1 =1 for proton and -1 for antiproton]
-1 [ih2 =1 for proton and -1 for antiproton]
120d0 [hmass]
-1d0 [scale:QCD scale choice]
-1d0 [facscale:QCD fac_scale choice]
.false. [dynamicscale]
.false. [zerowidth]
.true. [removebr]
10 [itmx1, number of iterations for pre-conditioning]
20000 [ncall1]
10 [itmx2, number of iterations for final run]
20000 [ncall2]
1089 [ij]
.false. [dryrun]
.true. [Qflag]
.true. [Gflag]

[Pdf selection]
'cteq6_m'[pdlabel]
4 [NGROUP, see PDFLIB]
46 [NSET - see PDFLIB]
```

```
cteq6mE.LHgrid [LHAPDF group]
-1             [LHAPDF set]

[Jet definition and event cuts]
40d0 [m34min]
400d0 [m34max]
40d0 [m56min]
400d0 [m56max]
.true. [inclusive]
'ktal'[algorithm]
0d0 [ptmin_jet]
0d0 [etamin_jet]
1000d0 [etamax_jet]
0.d0 [Rcut_jet]
.false. [makecuts]
0d0 [ptmin_lepton]
1000d0 [etamax_lepton]
0d0 [ptmin_missing]
0d0 [ptmin_lepton(2nd+)]
1000d0 [etamax_lepton(2nd+)]
0.d0 [R(jet,lept)_min]
0.d0 [R(lept,lept)_min]
0d0 [Delta_eta(jet,jet)_min]
.false. [jets_opphem]
0 [lepbtwnjets_scheme]
0d0             [ptmin_bjet]
1000d0            [etamax_bjet]
0d0 [ptmin_photon]
1000d0 [etamax_photon]
0d0 [cone_photon]
0d0 [cone_ptcut]

[Anomalous couplings of the W and Z]
0.0d0 [Delta_g1(Z)]
0.0d0 [Delta_K(Z)]
0.0d0 [Delta_K(gamma)]
0.0d0 [Lambda(Z)]
0.0d0 [Lambda(gamma)]
2.0d0 [Form-factor scale, in TeV]

[How to resume/save a run]
.false. [readin]
```

```
.true. [writeout]
''[ingridfile]
''[outgridfile]
```