CDF/PHYS/TOP/GROUP/10837
Version 1.0
December 21, 2012

Search for WZ/ZZ production in events with lepton(s) plus jets plus missing transverse energy

Giorgio Bellettini[1,2] Matteo Cremonesi[1], Giuseppe Latino[1], Vadim Rusu[2], Marco Trovato[1,2,3], Caterina Vernieri[1,2]

Abstract

We describe a search for associated WZ/ZZ production where a W/Z boson decays leptonically, and a Z boson decays to a quark pair. We search for the WZ/ZZ signal in events with lepton(s), large missing E_T and jets. Besides looking at the sample where two exclusive jets are found, we investigate the sample with 3 jets where 33% of the signal lie. In the 3 jets sample the invariant mass of the two E_T-leading jets would normally be chosen to reconstruct the Z-mass. To improve both the mass resolution and the sensitivity of measurement we describe an alternative procedure to reconstruct the Z- mass.

This search uses data collected through period 38, corresponding to an integrated luminosity of $\sim 9/fb$. We attempt to extract our signal through a fit to the Z reconstructed mass spectrum in 4 channels: untagged and tagged 2 jets and 3 jets channels.

[1]INFN-Pisa

[2]University of Pisa

[2]Fermilab

[3]Scuola Normale Superiore, Pisa

Contents

1 Introduction **3**

2 Data Samples and Event Selection **3**
2.1 Triggers . . . 4
2.2 Modeling of the triggers in high-E_T data stream . . . 5
2.2.1 Variable for Trigger Parametrization . . . 5
2.2.2 Trigger Efficiency Parameterization . . . 5
2.3 Personalized Jet Corrections . . . 8
2.4 Event selection . . . 9
2.4.1 QCD Veto . . . 9
2.4.2 $t\bar{t}$ veto . . . 10
2.4.3 Trigger Based Selection Cuts . . . 12
2.4.4 Additional Cuts . . . 12
2.5 Monte Carlo samples . . . 12
2.6 Sample Composition . . . 12
2.7 Electroweak And Top . . . 13
2.8 QCD content . . . 14
2.8.1 Cure of the QCD Model . . . 14
2.9 W+jets content . . . 20
2.9.1 Estimation of QCD and W+jets content - $\not\!\!E_T$ fit . . . 20

3 Fitting Procedure to the Z-Mass Distribution **21**
3.1 Systematic uncertainties . . . 26
3.2 Expected Sensitivity . . . 29
3.3 Fit Results on Data . . . 30

A MC sample names **32**

B Optmization of jet bness cuts **33**

C Electron Identification Cuts **33**

D Muon Identification Cuts **34**

E PHX Identification Cuts **35**

F QCDCalibrations **35**
F.1 PHX Notag2j . . . 35
F.2 TCE Notag3j . . . 35
F.3 PHX Notag3j . . . 35

G Fit results - Bump cuts **35**

Bibliography **35**

1 Introduction

The study of associated WZ boson production with a lepton and a neutrino signalling the W, and a $b\bar{b}$-pair in the final state is important since the event topology of this process is the same as expected for WH associated production of a W and the Standard Model light-Higgs boson ($M_H \lesssim 135$ GeV). Thus, the investigation of the process $W^{\pm}Z \to \ell\nu b\bar{b}$ whose rate can be accurately predicted, allows to calibrate and optimize many of the techniques used in the SM Higgs search and provides a "standard candle" for that crucially important search. In addition, WZ associated production generates a significant background for low mass Higgs Boson searches with H decaying into a $b\bar{b}$ pair.
At the Tevatron, the process $WZ \to Wb\bar{b}$ has an expected $\sigma \cdot BR$ [1] about five times larger than $WH \to Wb\bar{b}$ for $m_H \simeq 120$ GeV/c^2. Therefore, observing this process would be a benchmark for the even more difficult search for the light Higgs in the $WH \to Wb\bar{b}$ process.

Observing associated WZ production at the Tevatron in the channel $WZ \to \ell\nu b\bar{b}$ is extremely difficult for two main reasons.
The event rate is extremely low. A WZ production cross section of ~ 3.22 pb [1] together with a $Z \to b\bar{b}$ branching ratio of $\sim 15\%$ [2] provides about 50 fb in the $WZ \to \ell\nu b\bar{b}$ channel. With a trigger and kinematical selection efficiency of the order of a few %, one expects a handful of events per fb^{-1} of integrated luminosity. This statement remains valid even if the few ZZ events with leptonic decay of one Z that pass the selection cuts are included in the acceptance.
A standard kinematical cut requests exactly two high energy jets (i.e. $E_T > 20$ GeV) in the candidate sample. Simulations show that if a third energetic jet is allowed the signal acceptance is increased by about 1/3. Therefore, it would be very important to be able to extract a $Z \to b\bar{b}$ signal also in events with more than two high energy jets.
A second difficulty is that the signal to background ratio is very poor, due primarily to the contribution of associated production of W and incoherent jets. Optimal dijet mass resolution is of utmost importance for discriminating this background, since a fit to the invariant mass distribution of the two jets, associated to the hadronic decay of Z, is used to disentangle the diboson signal from the backgrounds in the candidate data sample.

In this note, we present a search for WZ/ZZ where a W/Z-boson decays leptonically and a Z-boson decays hadronically. The diboson signal is extracted in events with a high-energy lepton(s), large missing transverse energy and jets. We include ZZ in our signal since $ZZ \to l^+l^-q\bar{q}, l = e, \mu, \tau$ may pass our selection cuts because of fake missing transverse energy (one lepton not identified, or jet mis-measurements). Besides looking at the sample where two exclusive jets are found, we investigate the sample with 3 jets where about the 33% of the signal events lie.

In this analysis, we fit the Z-mass spectrum in W+2 jets selected region with expected signal and background shapes from MC or data (QCD multi-jet background). We make use of a NN jet b-ness tagger[3] to separate heavy from light flavour jets in order to better distinguish our signal from the WW background. In the W+3 jets region rather than building the Z-mass with the two E_T-leading jets we look event-by-event at the scores of four different Neural Networks (NNs) to understand the appropriate jet combination to use (Sec. **??**). However, fits in the W+3 jets are yet to be produced.

2 Data Samples and Event Selection

We select events with a signature of one or more charged leptons, large transverse missing E_T and exactly two or three jets. Considered leptons are listed below

- TCE
- CMUP

[1] BR being the Branching ratios of $Z/H \to b\bar{b}$

- CMX
- PHX
- Extended muon coverage - "EMC" $\equiv$ {CMU, CMP, CMIOCES, CMIOPES, CRKTRK}

2.1 Triggers

We use the dedicated high-P_T lepton triggered data streams (*bhel,bhmu*), the high-$\not\!\!E_T$ data stream (*emet*), and high-P_T plug trigger. The required trigger paths are listed in Table 1.

Lepton	Trigger
TCE	ELECTRON_CENTRAL_18
CMUP	MUON_CMUP, MUON_CMUP_L2_PT15,
CMX	MUON_CMX18, MUON_CMX18_L2_PT15,
	MUON_CMX18_L2_PT15_LUMI_200, MUON_CMX18_LUMI_250,
	MUON_CMX18_DPS
PHX	METPEM
EMC	MET35_&_TWO_JETS, MET35_&_CJET_&_JET,
	MET35_&_CJET_&_JET_LUMI_190,
	MET35_&_CJET_&_JET_DPS,
	MET35_&_CJET25_&_JET10, MET35_&_CJET25_&_JET10_LUMI_265,
	MET35_&_CJET_&_JET10_DPS
	MET_DIJET

Table 1: Trigger path(s) for each lepton category. In black trigger paths used over the whole datasets, in red trigger paths used for periods up to 14, in blue trigger paths used from period 15 to 38.

We use the whole collected dataset up to period 38 [2]. Every event is required to pass all the 6 goodrun lists specified below:

- goodrun_em_nosi.list
- goodrun_cmup_nosi.list
- goodrun_cmx_nosi.list
- goodrun_em_si.list
- goodrun_cmup_si.list
- goodrun_cmx_si.list

Goodrun lists version 45 is used. The integrated luminosity is about $8.9/fb$.

[2] We use reprocessed periods 18-38 with silicon clustering scheme 6.1.6.p+

2.2 Modeling of the triggers in high-E_T data stream

In this section we describe the trigger paths required while analyzing the *emet* data stream (Table 1): we call triggers highlighted in red "MET2J" and the one highlighted in blue "METDI" for simplicity.

As the name suggests, MET2J fires if there is a large missing trasnverse energy in the event as well as two energetic jets. Except for the first version of this trigger (MET35_&_TWO_JETS), at least one jet is required to be central.

Also METDI requires large missing trasnverse energy and two jets. This trigger was designed for Higgs boson searches and is therefore more efficient than MET2J. It was implemented during period 15 and it was never prescaled.

In order to keep things simple we will be requesting MET2J trigger paths for data periods up to 14 and METDI trigger path from period 15 to 38. We are aware that a more optimal way to combine these triggers was developed ([13]). We plan to reproduce that method in the next iteration of this analysis.

2.2.1 Variable for Trigger Parametrization

We parameterize the trigger efficiency turnon curves as a function of the missing transverse energy and the transverse energy of the leading jet. Rather than using a 2D parameterization, we derive two 1D parameterizations, as explained in Sec. 2.2.2.

When we parameterize for the $\not\!\!E_T$, we need this observable to be as close as possible to the trigger $\not\!\!E_T$. The natural choice would be the uncorrected $\not\!\!E_T$, but such a variable is not well modelled in our MC. A better modeling is observed when the $\not\!\!E_T$ is recalculated by using the z position of the primary interaction vertex and is corrected because of the jet corrections. This variable, which we call "jetcorrmet", can be thought as the transverse energy of the W-boson candidate.

2.2.2 Trigger Efficiency Parameterization

The parameterization of the trigger efficiency is performed in two steps:

- derive the jetcorrmet dependence
- derive the leading jet transverse energy ("E_T^1") dependence

When deriving the **jetcorrmet dependence** of the trigger efficiency, we want to reduce as much as possible the effect of trigger biases due to other quantities. In order to do that we select the high-P_T muon sample (*bhmu*) as follows:

- MET2J: Both jets with $E_T > 25\ GeV$, $\Delta R(jet1, jet2) > 0.9$, at least one jet with $|\eta^{det}| < 0.9$, $|\eta^{det}|$ being the pseudo-rapidity computed from the CDF center.
- METDI: Both jets with $E_T > 25\ GeV$ and at least one jet with $E_T > 40\ GeV$

Moreover we require to have at least a well identified muon (CMUP, CMX, CMU, CMP) with $P_T > 20\ GeV$ and $M_T^W > 10\ GeV$.

Jet cuts are imposed since

- MET2J requires at level 2 to have exclusively 1 jet with $E_T^{jet} > 20\ GeV$ and $|\eta| < 1.1$, and at least 2 jets with $E_T^{jet} > 15\ GeV$ and $|\eta| < 3.6$
- METDI requires at level 1 one tower with transverse energy more than 10 GeV

We proceed in deriving the $\mathbf{E_T^1}$ **dependence** of the trigger efficiency. This time, we require $jetcorrmet > 40\ GeV$ besides the aforementioned selection cuts.

Turnon efficiency curves and their parameterizations are shown in Fig. 1. In Fig. 2 it is shown that the modeling is improved by correcting the MC because of the E_T^1-dependent turnon.

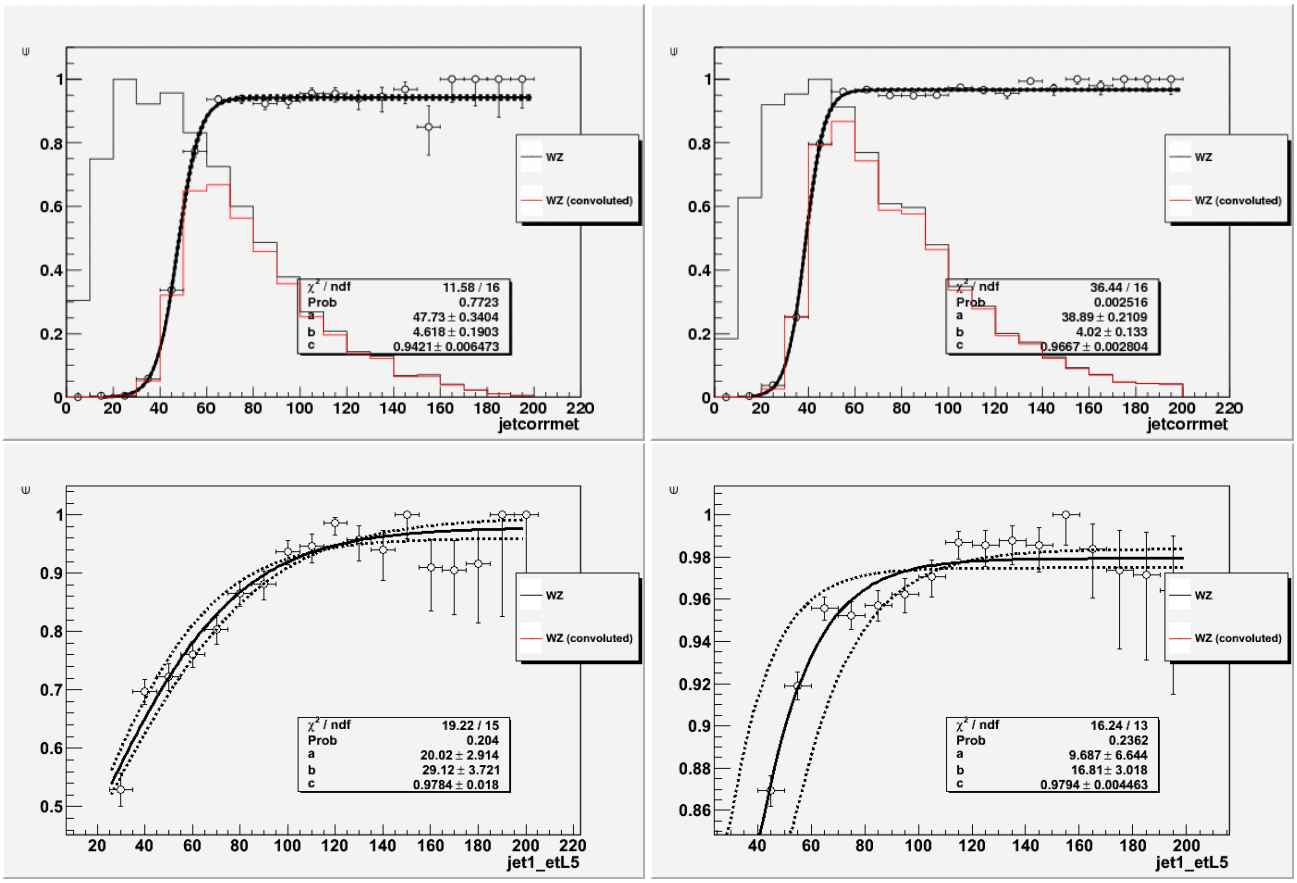

Figure 1: *Turnon efficiency curve parameterizations for MET2J (left) and METDI (right) triggers. Period<15 (left) and Period≥15 data were fitted by using* $\frac{c}{1+e^{\frac{(a-x)}{b}}}$*, a, b, c being free parameters of the fit, x being jetcorrmet (top) or* E_T^1 *(bottom). WZ MC distributions are shown before (black) and after the convolution with the trigger turnon parameterization (red) to get of the effects due to the trigger requirements.*

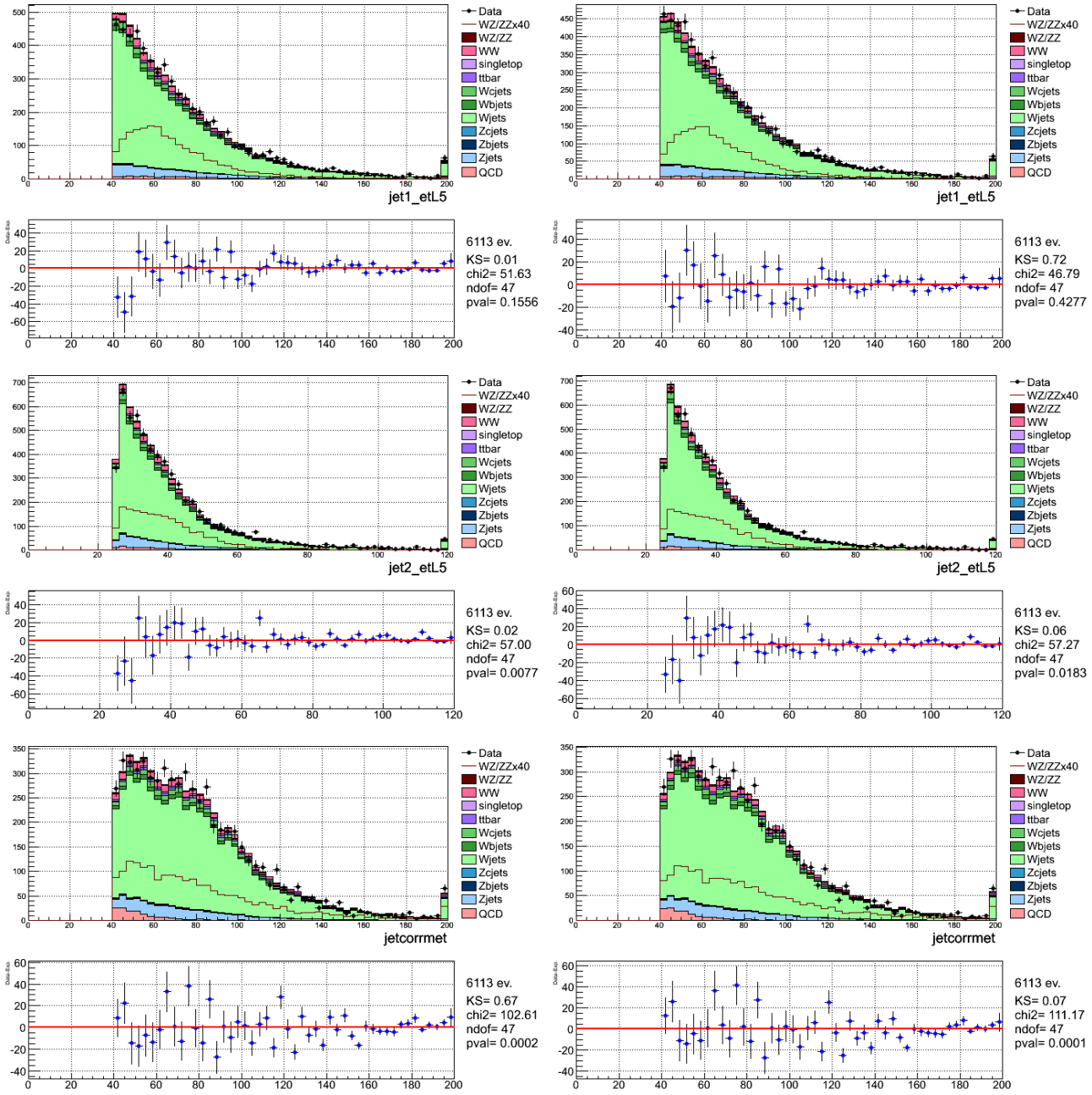

Figure 2: *E_T^{jet1} (top), E_T^{jet2} (center), jetcorrmet (bottom) distributions for EMC notag2j events passing the METDI trigger. MC and QCD templates are superimposed. E_T^1-dependent turnon is/is not applied to MC templates in the right/left plots*

The trigger efficiency parameterizations obtained in this section are going to be used when modeling the $W+jets$ sample. In order not to correct the MC twice for the derived trigger efficiency, the E_T^1 parameterization of the trigger will/will not be used to modify the expected shapes/rates.

2.3 Personalized Jet Corrections

For the measurement described in this note jets are corrected up to level 5 (L5) in both data and MC. In order to improve the modeling of the data in the W+jets sample we apply additional corrections to gluon (K_G) and quark (K_Q) jets in MC [3]. These corrections were derived by studying the balancing of the jets against the γ/Z in γ/Z+1 jet events [4]. These corrections, which we name "L9", have the following values:

- $K_Q = 1.014 \pm 0.027$
- $K_G = 0.921 \mp 0.044$

Therefore, if no stated otherwise, when we refer to jet energy info, we will assume that those energies have been corrected up to L9. We summarize L9 corrections in 1.

$$L9 \equiv \begin{cases} 1 & \text{Data} \\ K_Q & \text{MC, if quark jets} \\ K_G & \text{MC, if gluon jets} \\ 1 & \text{MC, if origin of the jet is unknown} \end{cases} \quad (1)$$

It should be noticed that L9 corrections will not be used to correct the missing trasnverse energy, since the extra energy which we subtract/add to the gluon/quark jets was already used to calculate the uncorrected $\not\!\!E_T$. Fig. 3 shows that the modeling improves by doing so.

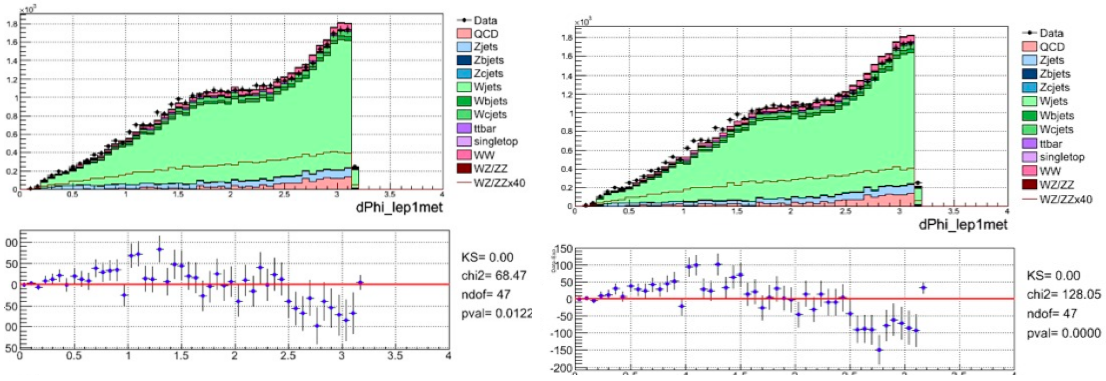

Figure 3: *$\delta\phi(lep1,\not\!\!E_T)$ in the TCE,CMUP,CMX Notag2j sample. $\not\!\!E_T$ is corrected because of L5 corrections (left), L5+L9 corrections. It's striking the mismodeling introduced by correcting the $\not\!\!E_T$ because of L9 corrections.*

[3] We match in direction ($dR < 0.4$) jets and partons in the HEPG bank. The highest-P_T parton in the jet cone determines whether the jet is a quark or gluon jet.

[4] Futher details are given in [14]

2.4 Event selection

We select events consistent with a leptonically decaying W/Z-boson and hadronically decaying Z-boson. In each events we require the presence of at least one lepton. Considered lepton categories are listed in Sec. 2. The highest-E_T lepton ("lep1") must have $E_T > 20$ GeV (electron) or $P_T > 20$ GeV/c (muons). $\not\!\!E_T$, recalculated by using the z position of the primary vertex, corrected for the presence of muons and because of L5 jet corrections [5], is required to be larger than 20 GeV. Jets are clustered by using the JETCLU algorithm with a cone radius of 0.4 and their energy is corrected as explained in Sec. 2.3. Two or three high-energy jets are required in the final sample.

In order to improve the sensitivity of the measurement and the understanding of the sample we divide the preselected sample depending on the following characteristics [6]:

- **lepton category**: we divide the sample based on the lep1 category: TCE, CMUP/CMX, PHX, and EMC.
- **jet flavor**: we isolate the region enriched by heavy flavor jets ("tag" region) from the one highly dominated by light flavor jets ("notag" region). The jet bness is used as b-flavor jet tagger. More details about the reason why we divide the sample in two regions are given in Appendix B, and a description of the jet bness tagger can be found elsewhere ([3],[4]).
- **number of jets**: samples with 2 and 3 reconstructed jets are kept separated.

Therefore, we end up with 16 orthogonal regions[7]. The definition of the regions based on the jet flavor and number of jets is given below.

2 jets: exclusively 2 jets with $E_T > 25$ GeV, $|\eta| < 2$

$\Rightarrow$ if additional jets exist, they must have $E_T < 13$ GeV or $|\eta| > 2$

2 jets - tag ("Tag2j"): $jet1\ bness > 0.75$, $jet2\ bness > -0.2$ [8].

2 jets - notag ("Notag2j"): in order to ensure the orthogonality of the regions, we remove the tag2j region from the 2 jets selected sample.

3 jets: exclusively 3 jets. Two jets with $E_T > 25, 15$ GeV, $|\eta| < 2$, the third jet with $E_T > 13$ GeV, $|\eta| < 3.6$

$\Rightarrow$ if additional jets exist, they must have $E_T < 10$ GeV or $|\eta| > 3.6$

3 jets - tag ("Tag3j"): $jet1\ bness > 0.75$, $jet2\ bness > -0.2$ [8]

3jets - notag ("Notag3j"): we remove the tag3j region from the 3 jets selected sample.

2.4.1 QCD Veto

$\not\!\!E_T$>20 GeV is applied to reduce the QCD contamination in the sample. Then, depending on the lep1 type, the following QCD vetoes are applied:

- CMUP,CMX: $M_T^W > 10$ GeV

[5] All jets stored in the Stntuples are used. However, it should be noticed that jet energy corrections are not applied to jets with raw $E_T < 8\ GeV$

[6] Different sub-samples have a different signal to background ratio and/or different background composition (e.g: multi-jets rate). A better post-fit modeling (See Sec. 3) was achieved for example after splitting the sample upon different lepton categories)

[7] in the notag regions jets are ordered in decreasing E_T, while in the tag regions where heavy flavor jets (HF) dominate, jets are ordered in decreasing bness

[8] bness cuts were optimized on the sensitivity of the measurement. See Appendix B

- TCE: $M_T^W > 30$ GeV
- PHX: $M_T^W > 30$ GeV, $\not\!\!E_T$>25 GeV, $\delta\phi^{min}(jet,\not\!\!E_T) > 0.3$ [9]
- CMU,CMP,CMIOPES: $M_T^W > 10$ GeV
- CMIOCES: $M_T^W > 10$ GeV, $\delta\phi(jet2,\not\!\!E_T) > 0.3$
- CRKTRK: $M_T^W > 20$ GeV, $\delta\phi(jet2,\not\!\!E_T) > 0.3$

It should be noticed that we optimized the QCD-veto cuts based on the signal over background ratio. However, not all the time it was possible to use the most optimal observables for this goal since some of them are heavily mismodeled. An example of mismodeld variables is shown in Fig. 4. It is evident that simulations do not reproduce the data, although a large discrimination power between QCD and signal is expected. CONTROLLARE LA METSIG DOPO LA QCD CURATA. SE BEN MODELLATA RIMUOVERE QUESTO PARAGRAFO E MENZIONARE IL TAGLIO IN METSIG COME POSSIBILE IMPROVEMENT FUTURO

On the other side a variable which shows a good agreement between data and simulation and a large discrimination power between QCD and signal is shown in 5. This variable is used in the QCD veto, as mentioned earlier.

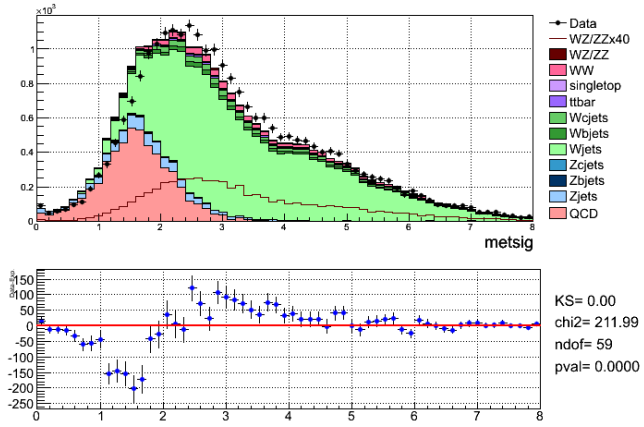

Figure 4: *$\not\!\!E_T$ significance ([5]) in the TCE notag2j sample. Simulations do not reproduce the data.*

2.4.2 $t\bar{t}$ veto

For the Tag2j region we apply the $t\bar{t}$ veto sketched below. More details are available on [16]

1. Events with 2 leptons are rejected if:
   - the first lepton is an electron and the second a muon, and the other way around
   - invariant mass of the 2 leptons outside of a $[60, 120]$ GeV/c^2 window
   - if $NN^{t\bar{t}\ Vs\ WZ} < 0.15$, $NN^{t\bar{t}\ Vs\ WZ}$ being the output of a neural network properly trained on $t\bar{t}$ and WZ simulated samples.

Output of the $NN^{t\bar{t}\ Vs\ WZ}$ and Background rejection Vs Signal efficiency curve are shown in Fig. 6

[9] all jets with $E_T > 5$ GeV are considered

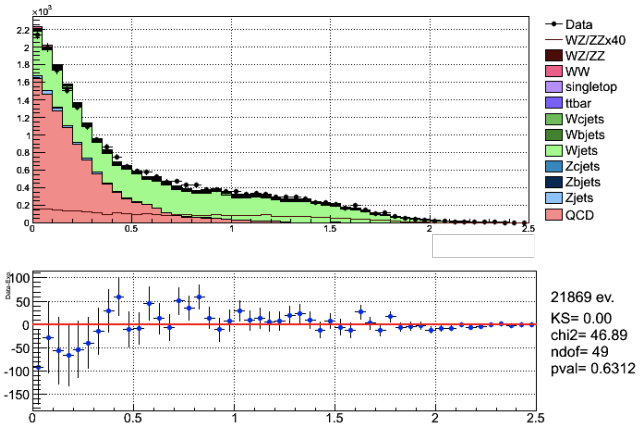

Figure 5: *$\delta\phi^{min}(jet,\not\!\!E_T)$ in the PHX notag2j sample before $\delta\phi^{min}(jet,\not\!\!E_T) > 0.3$ cut is applied. Simulations reproduce very well the data. METTERE PLOT CON LA QCD NUOVA*

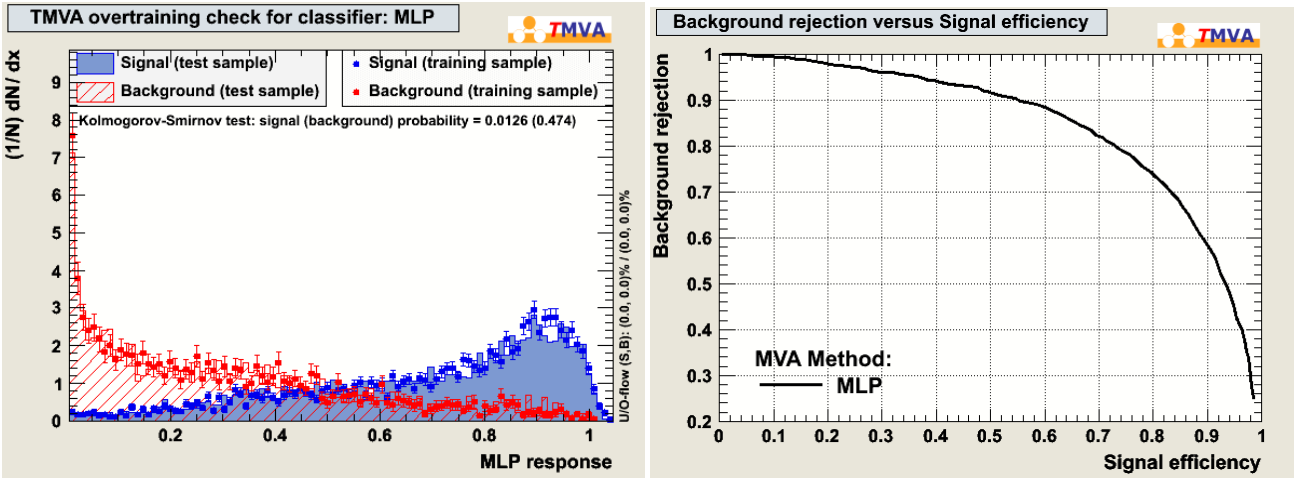

Figure 6: *NN response on signal (WZ) and background ($t\bar{t}$) simulated training and test samples (left). Background rejection Vs Signal efficiency curve (right)*

2.4.3 Trigger Based Selection Cuts

For EMC leptons we further require the trigger-based jet selection cuts described in Sec. 2.2.2

2.4.4 Additional Cuts

In order to enhance the signal over background ratio we further reject events where $\Delta R(jet1, jet2) < 0.7$ or $\Delta R(jet1, jet2) > 3.5$. By looking at Fig. 7 it's obvious that we reject much more background than signal by cutting on $\Delta R(jet1, jet2)$.

These selection cuts were optimized to enhance the WZ signal over the overwhelming background processes. However, as mentioned before, a ZZ event can be accepted in our sample since a lepton can be missed by the CDF detector (in that case the leptonically decaying Z is equivalent to a W from the observed point of view) or fake $\not\!\!E_T$ may arise from calorimetric mis-measurements.

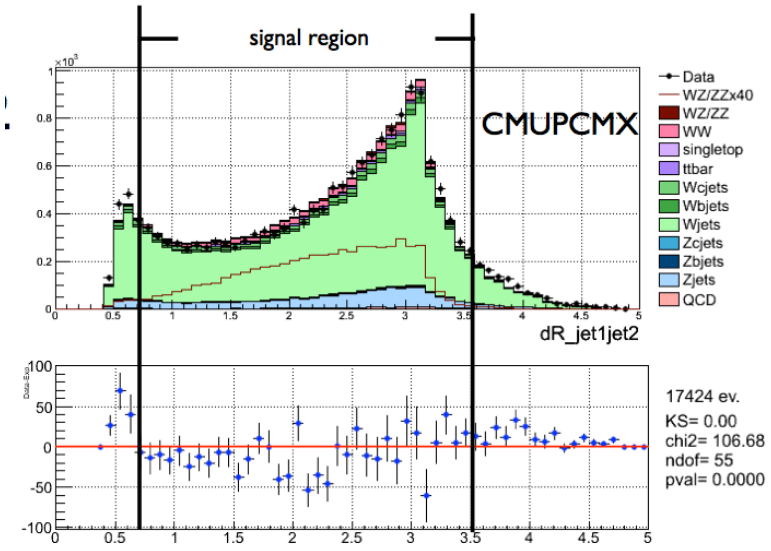

Figure 7: $\Delta R(jet1, jet2)$ *distribution in the Notag2j CMUP,CMX sample*

2.5 Monte Carlo samples

We use both Pythia and Alpgen+Pythia MC dependending on their availability. The dominant W+jets background is modeled by using the top group Alpgen+Pythia samples. These analysis is performed on the Stntuples. In Appendix A we list all the Stnuples catalog names which were used in this analysis.

2.6 Sample Composition

We estimate the expected composition of the selected sample in a sequence of steps by using a method which is similar to Method2, used in several analyses at CDF which exploits heavy flavour tagging algorithms [6]. One of the differences with Method2 is that rather than applying a renormalization ("scale factors") of the b-jets tag efficiency and non-b jets mistag rate in the Monte Carlo, we locate the equivalent cuts in the Monte Carlo that matches the measured efficiency (for processes that have real b jets, like top backgrounds and our signal)

and mistag rates (for processes without real b jets, like WW) in the data. In the same manner, we determine equivalent cuts on bness in the Monte Carlo that match the $\pm 1\sigma$ uncertainty values on the mistag rate and efficiency. These values are summarized in Table 2.

The method assumes that the following processes contribute to the selected data sample:

- **Electroweak and Top**: processes characterized by accurately predicted cross sections (see Sec. 2.7)
- **QCD**: multi-jet production associated to a hadron faking the lepton and $\not\!\!E_T$("fakes", see Sec. 2.8)
- **W+jets**: production of a W-boson associated to hadron jets (see Sec. 2.9)

The steps implemented by the method are the following:

1. The contribution of processes whose cross sections are well known is estimated.
2. The contribution of QCD fakes and of W+jets is estimated with a data-based method.

In this way we avoid using the imprecise theoretical predictions for the production cross section of W-boson with associated jets ([7])

Sample	bness cut in Data	Equivalent MC cut		
		-1σ	Central Value	1σ
Non b-jets	-0.2	-0.425	-0.33	-0.35
Non b-jets	0.75	0.65	0.68	0.72
b-jets	-0.2	-0.06	0.07	0.18
b-jets	0.75	0.79	0.825	0.86

Table 2: Equivalent bness cuts applied to MonteCarlo samples. These cuts are chosen to match the measured mistag rates and b-tag efficiency in data. The uncertainties on these cuts are determined using the calculated uncertainties on the mistag rates and tagging efficiencies ([3])

2.7 Electroweak And Top

The electroweak (EW) and top processes contributing to the selected data sample are WW, WZ, ZZ, Z+jets, $t\bar{t}$, single-*top*. Each of these processes can produce signatures consistent with one or more leptons, large $\not\!\!E_T$, and two or three jets, arising from correctly identified or mis-identified objects in the CDF detector. The contamination of each of the EW processes to the notag and tag samples is estimated as:

$$N_X^{pretag} = \sigma_{\mathrm{p\bar{p}}\to X} \cdot A \cdot \int Ldt \tag{2}$$

$$N_X^{tag} = \sigma_{\mathrm{p\bar{p}}\to X} \cdot A \cdot \epsilon \cdot \int Ldt \tag{3}$$

where

$X = WW, WZ, ZZ, Z + jets, t\bar{t}, single - top$;

$\sigma_{\mathrm{p\bar{p}}\to X}$ is the theoretical cross-section for $\mathrm{p\bar{p}} \to X$ (see Table 3);

A is the MC-derived acceptance [10]

ϵ is the tagged selection efficiency estimated as the number of events in the tag sample divided by the number of events in the notag sample.

In Table 6 we report the number of expected EW and top events in the notag and tag selected data samples.

Process	Cross-Section (pb)
WZ	3.96 ± 0.06
WW	12.4 ± 0.25
ZZ	1.58 ± 0.05
$t\bar{t}$	7.04 ± 0.8
$single - top$	
t-channel	1.98 ± 0.08
s-channel	0.88 ± 0.05
$Z + jets$	787.4 ± 50

Table 3: Theoretical cross sections for Electroweak and Top processes (M_t = 172.5 GeV/c^2) [8].

2.8 QCD content

In this section we present a method to estimate the contamination of multi-jets QCD events in the selected pretag and tag samples.

The background from QCD production arises when a jet is mis-identified as a lepton, and jet energy mis-measurements fake a large missing transverse energy. In particular the semileptonic decays of b-hadrons in jets produce electrons or muons, and neutrinos, altogether faking the signature of W's.

To model the QCD background in the selected samples we use data samples selected as described in Sec. 2.4, except for the identification cuts on the primary lepton. Some of these identification cuts are inverted to ensure that the QCD samples are orthogonal and at the same time kinematically close to the samples that we want to model. In Table 4 we list the lepton types used for modeling the QCD contamination together with the inverted cuts[11]. In Sec. 2.8.1 we describe the calibrations performed to the aforementioned QCD models.

In order to avoid double counting with the other background processes we subtract EW, top, and W+jets expected events from the QCD samples. Theoretical cross-sections are used (A) for the subtraction [12]

2.8.1 Cure of the QCD Model

In this section we present the study aiming at correcting the defects of the QCD model for the TCE sample. Details about the equivalent studies for the PHX sample are reported in Appendix F.

[10] Acceptance is estimated by applying the selection criteria described in Sec. 2.4. Calibration factors accounting for discrepancies in lepton identification efficiencies between data and MC, and trigger efficiencies are applied. All these numbers, except the $MET2J$ and $METDI$ trigger efficiencies (see Sec. 2.2.2) are taken from [20]

[11] For brevity we will often assign the names of these leptons to the QCD templates/models themselves.

[12] We should use the $W + jets$ rate estimated from data (Sec. 2.9.1) rather than the one expected from the theory. However, due to the small amount of $W + jets$ events in the QCD sample, we expect that any difference which may arise is well contained in both QCD rate and shape uncertainties.

lep1 type	QCD model	Inverted Cuts
TCE	Anti-TCE	E_{HAD}/E_{EM}, χ^2, L_{shr}, $Q \cdot \Delta X$, $\lvert\Delta Z\rvert$ *
CMUP,CMX	Non-isolated CMUP,CMX	$Isolation > 0.1$
PHX	Anti-PHX	PEM 3x3 χ^2, E_{HAD}/E_{EM}, PES 5x9 U,V, PEM 3x3 Tower Fit *
EMC	Non-isolated EMC	$Isolation > 0.1$

Table 4: QCD models used per each lep1 type. "*" indicates that two identification cuts are inverted with respect to Sec. C (TCE) or Sec. E (PHX). Further info are at [6]

These studies are performed in sub-sample ("side-bands") which are at the same time qcd enriched, kinematically close, and orthogonal to the signal region. The side-band for TCE sample is defined as summarized in Table 5.

Region	Selection Cuts	frac QCD
Side-band	$\not\!\!E_T/GeV < 20$ or $M_T^W/GeV < 30$	84%
Signal	$\not\!\!E_T/GeV > 20$ and $M_T^W/GeV > 30$	21%

Table 5: Definitions of the region used to correct the QCD template in the TCE notag2j sample. Cuts applied for the signal regions are also reported. All cuts described in Sec. 2 except the $\not\!\!E_T$ and M_T^W ones were previously applied.

The defects of the Anti-TCE model spotted by us are listed below:

- Trigger bias
- Wrong anti-TCE energy scale which mostly affect the $\not\!\!E_T$ modeling [13]

The effect of the **trigger bias** on the QCD models is evident from Fig. 8. One of the cause of such a bias is because of the P_T electon trigger requirerements. The trigger requires at least one seed tower of 8 GeV or a cluster of 7.5 GeV in the electro-magnetic calorimeter, therefore sculpting the Anti-TCE P_T spectrum towards higher values than for the TCE spectrum. Another reason of such discrepancy in the spectrum is due to the amount of energy lying just outside the energetic TCE/Anti-TCE cores. Although, both objetcs are isolated we notice more energy around the Anti-TCE than around the TCE. METTERE PLOT ISOLATION.

The anti − TCE energy scale is studied in QCD MC [14]. The transfer functions, which provides the size of the correction to be applied to the jml rawet in order to obtain the real TCE (anti-TCE) energy scale [15], are defined below:

$$TF^{TCE(anti-TCE)} \equiv P_T^{parton}/P_T^{jml^{TCE\ (anti-TCE)}} \tag{4}$$

where $jml^{TCE\ (anti-TCE)}$ is the jet matched in direction ($dR < 0.4$) to the TCE (anti-TCE) while parton is the highest-P_T HEPG parton in the jml cone. Since our purpose is not to improve the jml energy resolution but to make the anti-TCE energy scale consistent with the anti-TCE, we define the correction to be applied to the anti-TCE jml energy as follows:

$$K \equiv TF^{anti-TCE}/TF^{TCE} \tag{5}$$

[13] Antielectrons jet corretions are used to correct the $\not\!\!E_T$

[14] All MC listed in [23] have been used

[15] We are planning to repeat the same studies by comparing the $anti-TCE$ energy to the energy sum of the HEPG hadrons within the jet cone

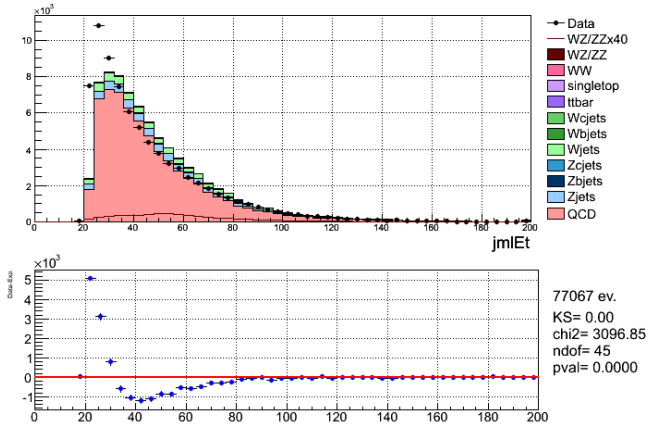

Figure 8: *Transverse energy of the jet matched to TCE (Data dots and MC) or Anti-TCE (QCD).* $\Delta R < 0.4$ *is used as matching criteria. It's evident that the trigger is biasing the QCD shape.*

K dependence on jml raw E_T is parameterized as shown in Fig. 10. Fig 11 shows TCE and anti-TCE TFs before and after the corrections applied to the anti-TCE.

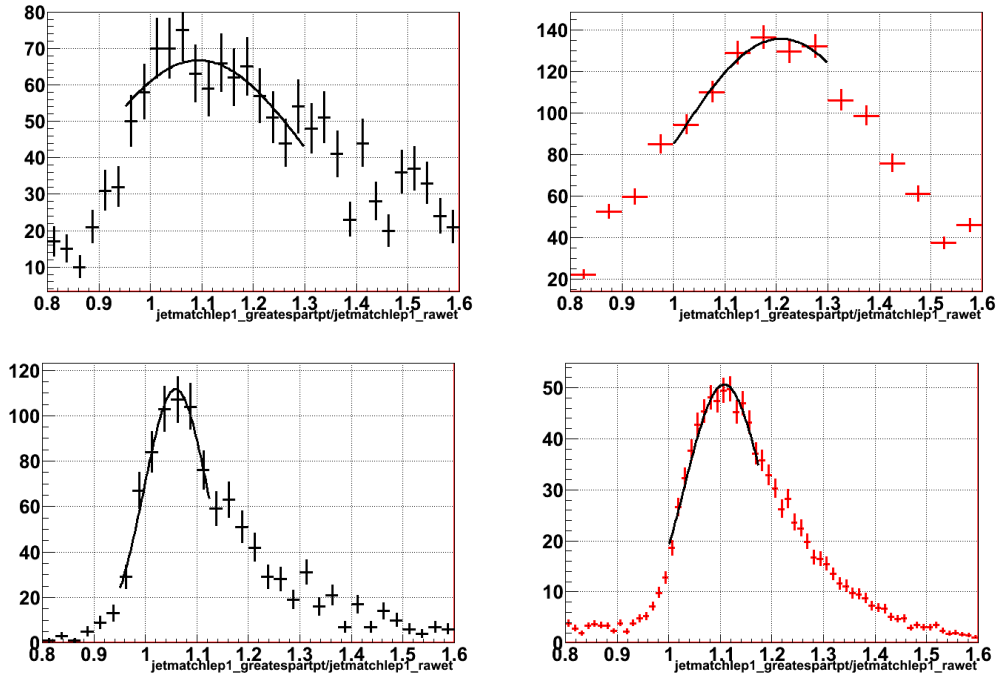

Figure 9: *The most energetic matched parton* P_T *divided by the transverse energy of the jet matched to TCE (left) or* $anti-TCE$ *(left). Two particular jet matched raw* E_T *bins are shown:* $E_T < 50$ *(up),* $100 < E_T < 150$ *(bottom). Matching criteria is* $dR(parton, jet) < 0.4$*. Parton* P_T *within the jet matched to the anti-TCE is reweighted to obtain the same spectrum as the parton* P_T *in the jet matched to the TCE. RIMUOVERE GAUSSIANE.*

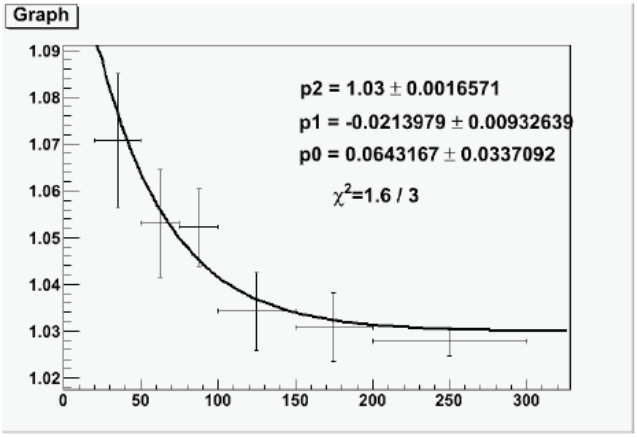

Figure 10: *Corrections to the jet mached to the anti-TCE as function of its raw transverse energy. The exponential functional form* $p0 + exp[p1 * (E_T - 20) + p2]$ *has been used for parameterizing the raw* E_T *dependence. Dots and error bars are derived respectively from the means of the* TF *distributions (Fig 9) and their uncertainties.*

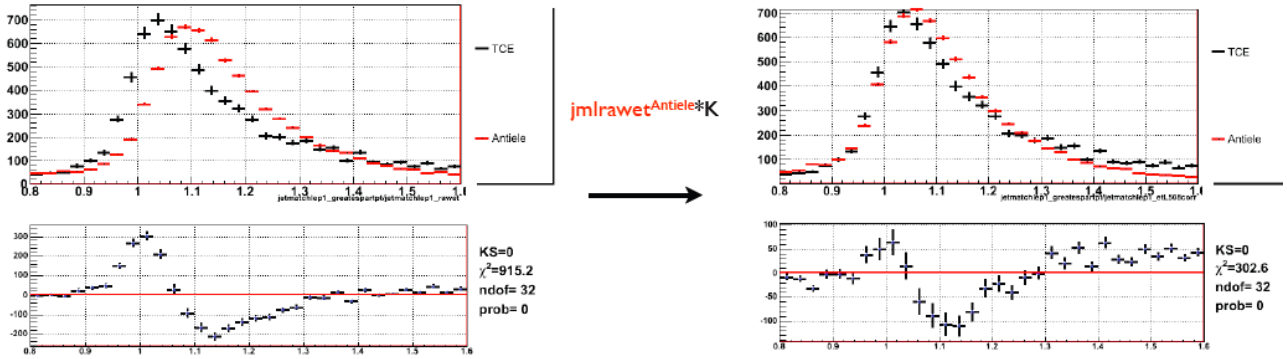

Figure 11: *TCE (black) and anti-TCE (red)* TF*s before (left) and after (right) the anti-TCE jml raw* E_T *hass been corrected with* K *(Fig. 10).*

In summary, in order to cure the aforementioned problems we apply the following procedure:

- we correct jml raw E_T: $(jml^{anti-TCE}$ raw $E_T) \rightarrow K * (jml^{anti-TCE}$ raw $E_T)$
- we reweight the QCD template to make jml corrected E_T shapes in data and QCD+MC concide (see Fig. 12)

Fig. 13-18 compare the modeling of several fundamental distributions prior and after the cure. The QCD templates prior the cure are built without reweighting the anti-TCE events and by applying/not applying the standard jet energy corrections to the $jml^{anti-TCE}$ when its electron-magnetic fraction is below/above 90%. We will refer to this scenario as "out-of-the-box" QCD model, while we will call the QCD model after the cure "cured" QCD model.

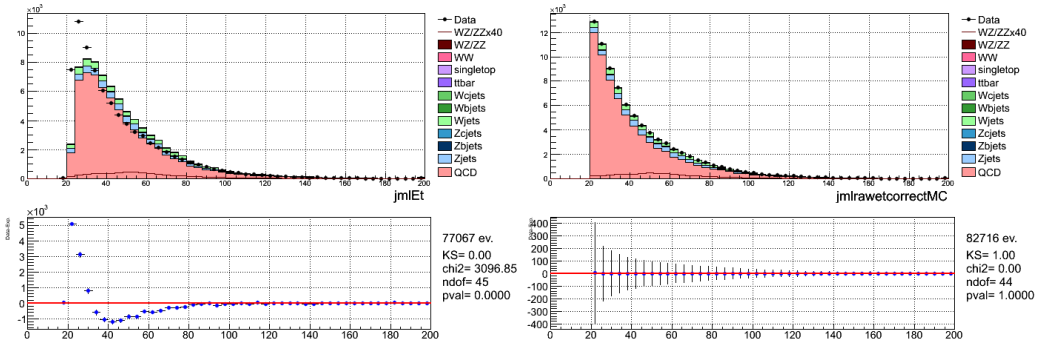

Figure 12: *jet matched raw energy (left) and jet matched corrected energy (K - see Fig 10) once the trigger bias is removed.*

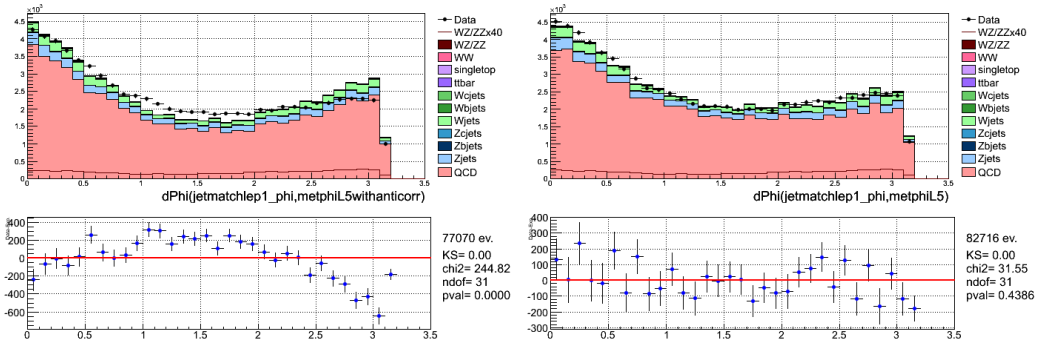

Figure 13: $\Delta\Phi(jml,\not\!\!E_T)$ *with out-of-the-box (left) and cured (right) QCD model.*

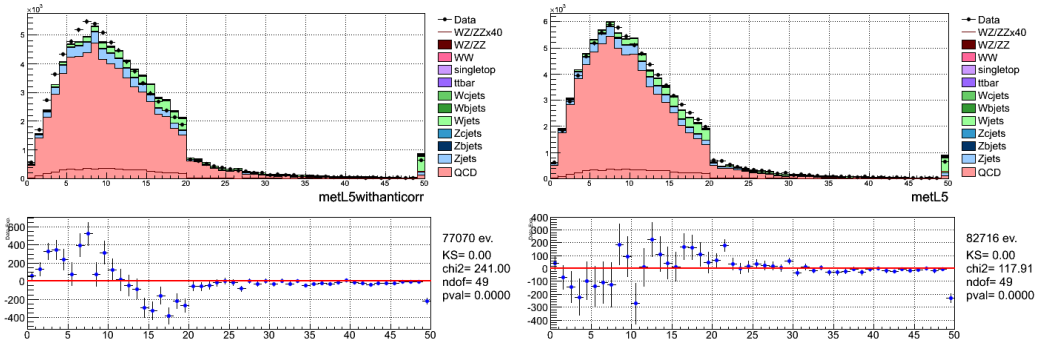

Figure 14: $\not\!\!E_T$ *with out-of-the-box (left) and cured (right) QCD model.*

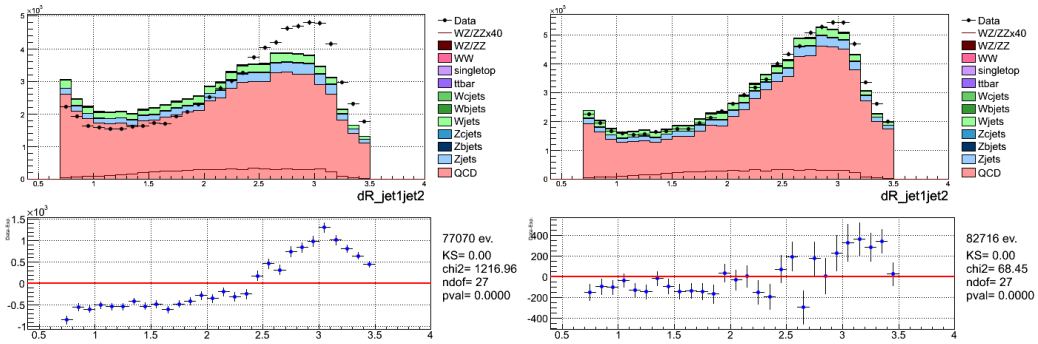

Figure 15: *$dR(jet1, jet2)$ with out-of-the-box (left) and cured (right) QCD model.*

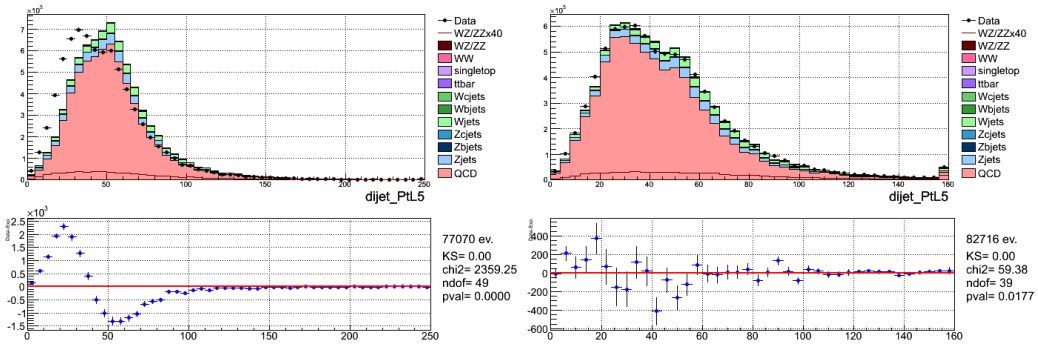

Figure 16: *$P_T(jet1, jet2)$ with out-of-the-box (left) and cured (right) QCD model.*

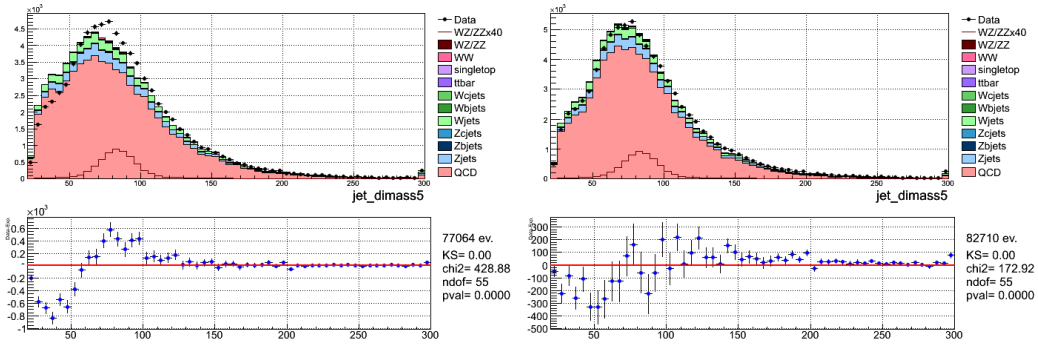

Figure 17: *$M(jet1, jet2)$ with out-of-the-box (left) and cured (right) QCD model.*

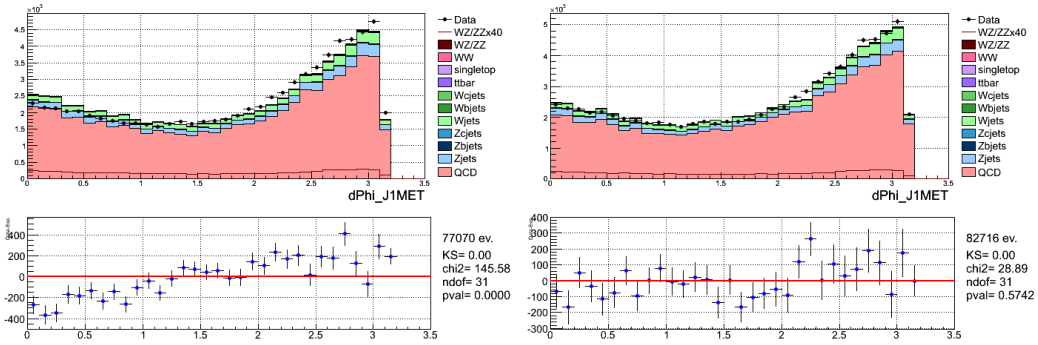

Figure 18: *dPhi(jet1,$\not\!\!E_T$) with out-of-the-box (left) and cured (right) QCD model.*

2.9 W+jets content

Due to the presence of a real lepton and neutrino, the W+jets background is an irreducible background for this measurement. Attempts to reduce such a background have been made. However, no of them showed a consistent improvement in sentivity of the measurement. Therefore we do not include the tools developed to reduce the W+jets background in this iteration of the note. However, we continue working on it.

2.9.1 Estimation of QCD and W+jets content - $\not\!\!E_T$ fit

Depending on lep1 type we perform a fit of the $\not\!\!E_T/M_T^W$ distributions in the notag and tag data samples in order to obtain the expected rates of QCD and W+jets events in those samples. With respect to the selection cuts described in Sec. 2.4, we allow for events with $\not\!\!E_T < 20\ GeV$ for lep1=TCE, $M_T^W < 10$ for lep1=CMUP,CMX,EMC, and $\not\!\!E_T < 25$ for lep1=PHX in order to perform the fit. Three templates are used for the fit:

1. EW: MC-based template built by using EW selected events. The normalization of this template is constrained to the content reported in table 6 (notag sample) or 7 (tag sample)
2. W+jets: MC-based template built by using W+jets selected events. The normalization of this template is a free parameter of the fit.
3. QCD:
   - > Notag samples: data-based template built by using the models described in Sec. 2.8.
   - > Tag samples: we measure in data the tag rate in bins of $\not\!\!E_T$(lep1=TCE,PHX) or M_T^W (lep1=CMUP,CMX,EMC) (Fig. **??**), and we use this rate to weight the pretag QCD template. The resulting distribution is used as QCD template in the tag sample [16].

The fit is performed by using a binned likelihood fit. The results of the fit in the pretag and tag samples are shown in Fig. 23-24. The expected QCD and W+jets rates in the pretag and tag samples are reported in table 6 and 7 respectively.

We need to estimate a **systematic uncertainty** on the QCD rate. There is no need to estimate such uncertainty for the $W + jets$ rate since the $W + jets$ rate will be left free to float in the final fit.

The following assumptions were made while extracting the QCD rate for TCE notag2j sample [17]

[16] The requirements on the b-flavour content of the jets (bness cuts) drastically reduce the number of events of the data-based QCD template, thus making it impossible to build a reasonable QCD template directly in the tag sample

[17] QCD rate systematic uncertainty for PHX Notag2j, TCE and PHX Notag3j is described in Sec. **??**. Due to the limited amount of QCD in the muon sample we didn't bother to estimate an uncertainty for it. Therefore we assume the same QCD uncertainty in the muon and electron samples

1. W+jets, EW, top shapes taken from simulated events
2. EW and top rate derived from theoretical cross-sections
3. a particular model (antiele, antiphx) was chosen to derive the QCD shape
4. the fit was performed by exploiting a particular observables ($\not\!\!E_T$)

We investigate the effect of all the above assumptions on the QCD rate in order to quote a systematic uncertainty on it.

MC shapes are varied by shifting up/down the JES. We assume that the $\not\!\!E_T$uncertainty comes only from the uncertainty on the energy of clustered objects. By doing so we obtain $\Delta QCD/QCD = .9\%$. Results are shown in Fig. 19

EW and top rate is varied within theoretical uncertainties (6%). We obtain $\Delta QCD/QCD = .9\%$. Results are shown in Fig. 20

The uncertainty due to the *chosen QCD template* is investigated as follow. We compare the nominal QCD rate with the number of data events after subtracting the W+jets, EW, Top, expectations in the signal region. We obtain $\Delta QCD/QCD = 4.4\%$. The reference plot is Fig. 21 .

The effect due to the *chosen observable* for the fit is investigated as follows. We perform the QCD fit in $d\phi_{min}(\not\!\!E_T\ , jet)$, where we consider any jet whose $E_T > 5\ GeV$. We obtain a QCD rate of about 16.0% of the data by doing so. We decide to average the QCD rates from the $\not\!\!E_T$ and $d\phi_{min}(\not\!\!E_T\ , jet)$ fit and to use this average as our nominal value: QCD rate = 18.5 % of the data. The difference between the qcd rate obtained from the $\not\!\!E_T$ or $d\phi_{min}(\not\!\!E_T\ , jet)$ is used as systematic uncertainty. Therefore $\Delta QCD/QCD = 13.1\%$. Results are shown in Fig. 22

All the obtained systematic uncertainties on the QCD rate are summed in quadrature to the uncertainty from the $\not\!\!E_T$ fit ($\Delta QCD/QCD = 1.4\%$, see Fig. 23). The total systematic uncertainty on the QCD rate is $\mathbf{\Delta QCD/QCD = 14.1\%}$

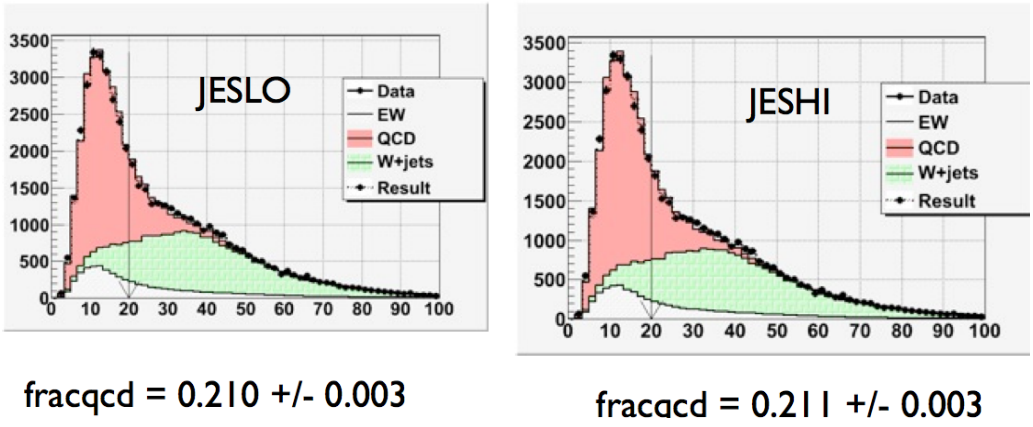

Figure 19: *QCD $\not\!\!E_T$ fits in the notag2j TCE sample. JES has been moved by -1σ (left) and 1σ (right). QCD rate as obtained from the fits are reported. This values have to compared with 23*

3 Fitting Procedure to the Z-Mass Distribution

The WZ/ZZ signal is extracted performing a χ^2 fit of the Z-invariant mass distribution. χ^2 minimizations is perfomed by using MINUIT. *mclimit* package [15] is used for this purpose.

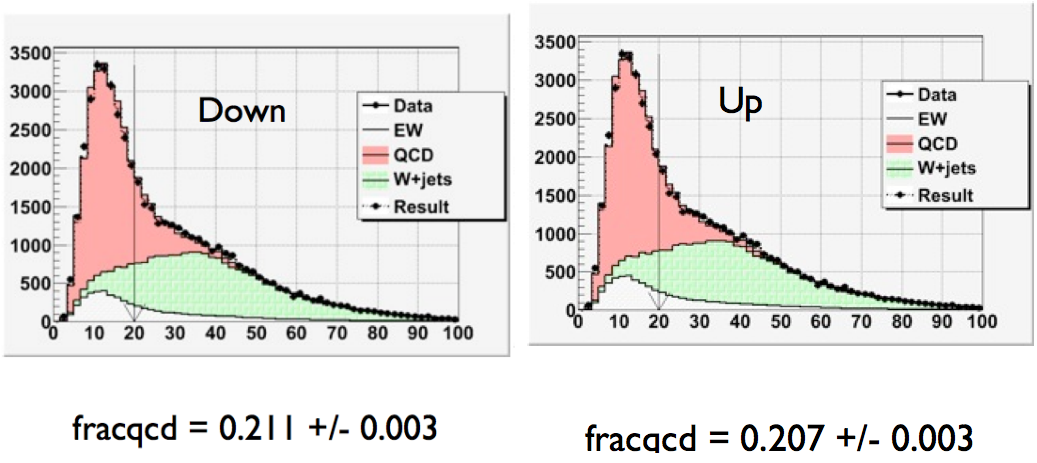

Figure 20: *QCD $\not\!\!E_T$ fits in the notag2j TCE sample. EW+Top rates has been moved by -1σ (left) and 1σ (right), σ being 6%. QCD rate as obtained from the fits are reported. This values have to compared with 23*

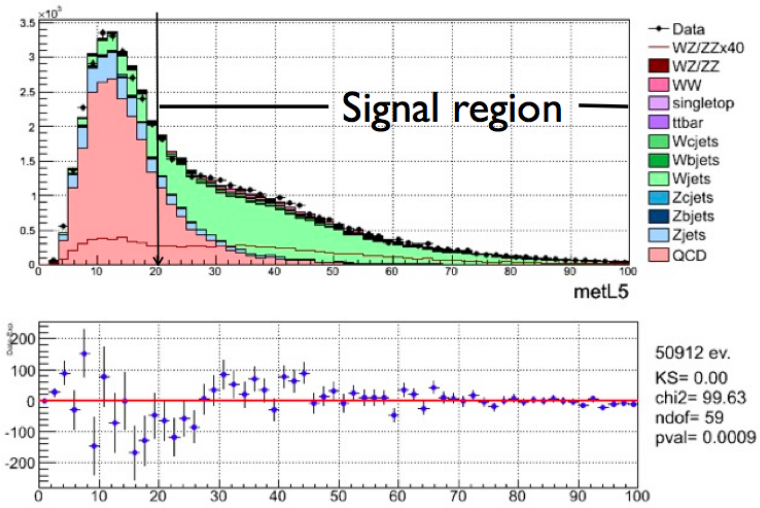

Figure 21: *$\not\!\!E_T$ fit in the notag2j TCE sample.*

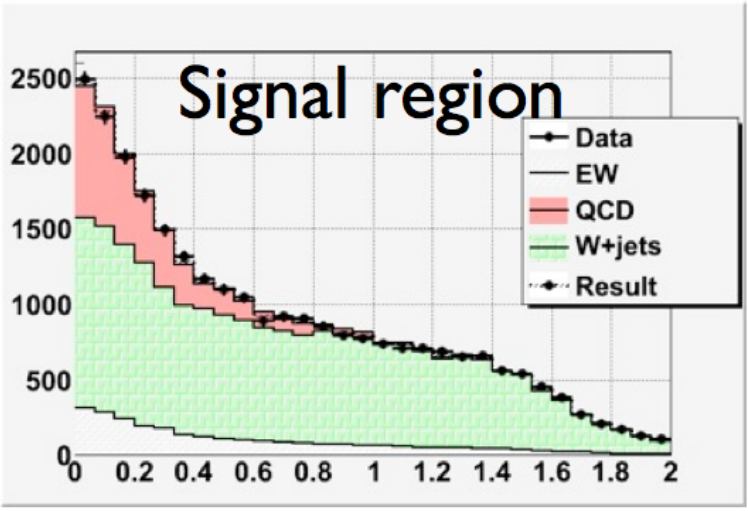

Figure 22: *QCD $d\phi_{min}(\not\!\!E_T$,$jet)$ fit in the notag2j TCE sample.*

Process	Rate (TCE)	Rate (CMUP,CMX)	Rate (PHX)	Rate (EMC)	Rate (all lep)
Signal	161.6 ± 1.6	138.9 ± 1.4	60.7 ± 0.9	63.8 ± 0.8	425.066 ± 2.41582
WW	852 ± 5	589.2 ± 3.7	277.7 ± 2.8	341.2 ± 2.7	2060.05 ± 7.35833
tt	299 ± 1.7	214.2 ± 1.3	52.7 ± 0.7	144.8 ± 1	710.626 ± 2.44998
single-top	134.2 ± 0.6	98.3 ± 0.5	28.6 ± 0.3	71.5 ± 0.4	332.516 ± 0.889941
Zjets	1054.8 ± 6.7	1610.3 ± 6.6	94.1 ± 2.4	646.8 ± 3.7	3406 ± 10.413
Wjets	15606 ± 64.3	12073.1 ± 52	7057.3 ± 47	6426.2 ± 36.6	41162.6 ± 101.895
QCD	4786.5 ± 108.9	126 ± 2.9	2808 ± 89.5	130.7 ± 2.9	7851.12 ± 141
Data	22894 ± 151.3	14850 ± 121.9	10379 ± 101.9	7825 ± 88.5	55948 ± 236.533

Table 6: Predicted and observed number of events of the notag 2jets sample selected according to the requirements described in Sec. 2.4. W+jets and QCD rates are etimated from the $\not\!\!E_T$fit in data (Sec. 2.9.1). The expected rates are separated for the triggered lepton type. We also require the invariant mass of the two E_T-leading jets to be within [20,300] GeV/c^2

. By construction the overall expected rates are the same as the observed ones in each region.

Fits are performed in the $[20, 300]$ GeV/c^2 range with 56 bins in the notag2j regions and $[20, 200]$ GeV/c^2 range with 18 bins in the tag2j regions.

The fit is performed simultaneously in 8 regions (Sec. 2.4):

- Tag2j (4 regions)
- Notag2j (4 regions)

The data is fit to the following templates

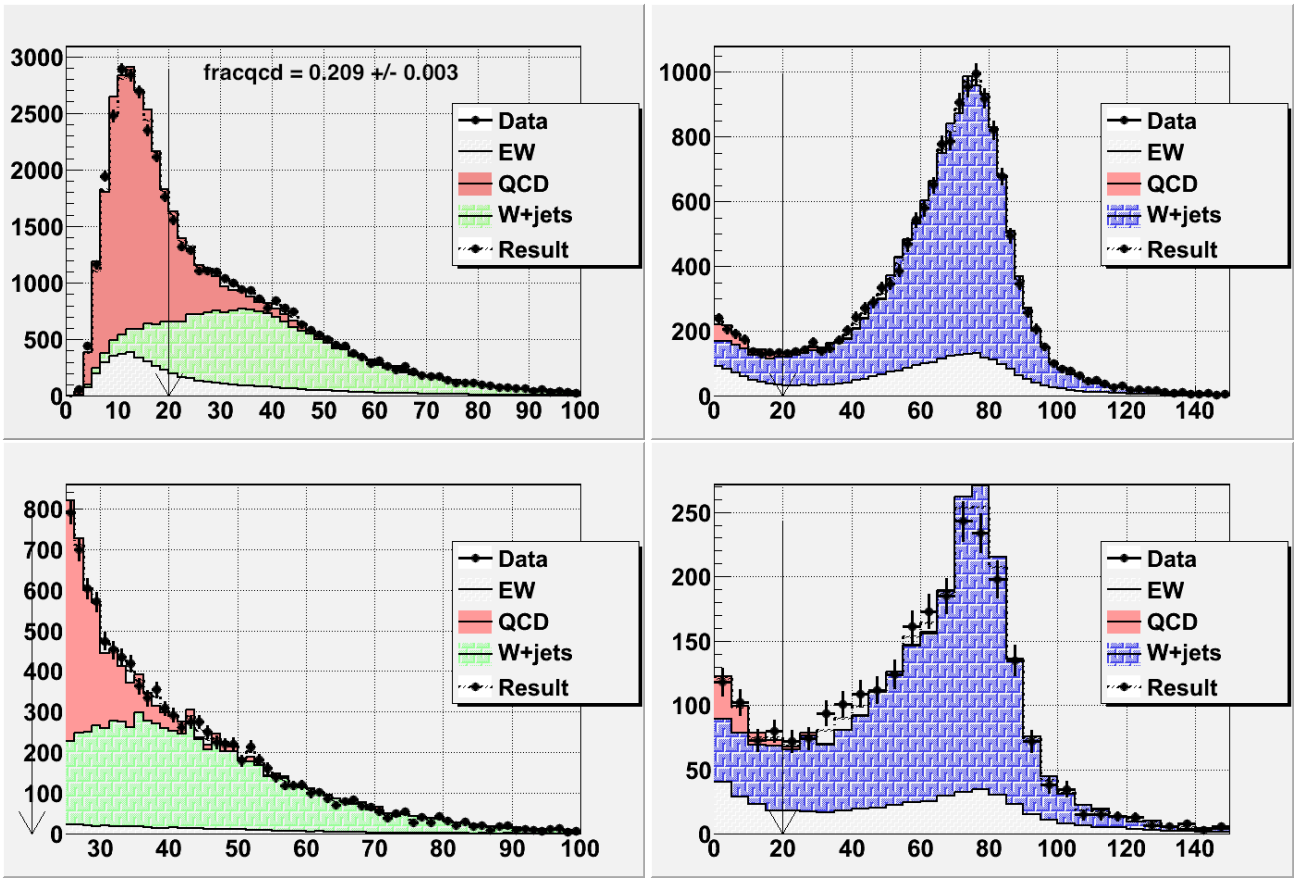

Figure 23: *Fits for the QCD contribution in the notag 2 jets data sample (see Sec.* **??***) when lep1=TCE (upper left), CMUP,CMX (upper right), PHX (lower left), CMU,CMP MET_DIJET trigger (lower right).*

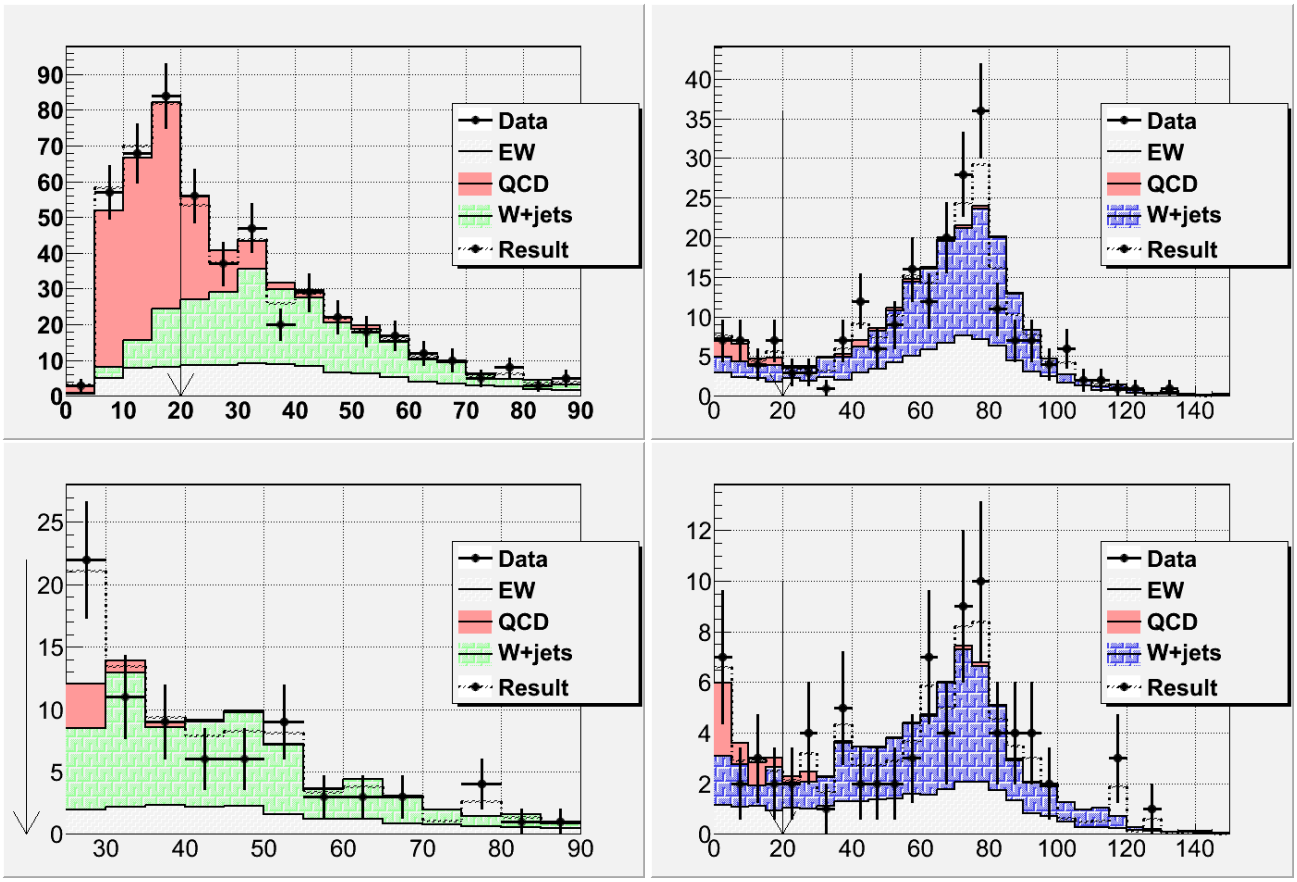

Figure 24: *Fits for the QCD contribution in the tag 2 jets data sample (see Sec.* **??***) when lep1=TCE (upper left), CMUP,CMX (upper right), PHX (lower left), CMU,CMP MET_DIJET trigger (lower right).*

Process	Rate (TCE)	Rate (CMUP,CMX)	Rate (PHX)	Rate (EMC)	Rate (all lep)
Signal	9 ± 0.4	7.6 ± 0.3	3.9 ± 0.2	3.1 ± 0.2	23.7 ± 0.57
WW	5.7 ± 0.4	3.5 ± 0.3	2.1 ± 0.3	1.3 ± 0.1	12.7 ± 0.6
tt	35.5 ± 0.6	27.9 ± 0.5	7.9 ± 0.3	15.4 ± 0.3	86.7 ± 0.9
single-top	22.6 ± 0.2	16.7 ± 0.2	4.5 ± 0.1	8.2 ± 0.1	51.9 ± 0.3
Zjets	7.3 ± 0.7	20.8 ± 0.9	0.7 ± 0.3	6.3 ± 0.4	35.1 ± 1.2
Wjets	162.6 ± 5.4	116.8 ± 4.2	37.7 ± 2.3	57.1 ± 3.6	374.2 ± 8.1
QCD	58.3 ± 0.4	5.7 ± 0	22.1 ± 0.4	4.6 ± 0.2	90.7 ± 0.6
Data	301 ± 17.3	199 ± 14.1	79 ± 8.9	96 ± 9.8	675.0 ± 26.0

Table 7: Predicted and observed number of events of the tag 2jets sample selected according to the requirements described in Sec. 2.4. W+jets and QCD rates are etimated from the $\not\!\!E_T$ fit in data (Sec. 2.9.1). We also require the invariant mass of the two jets with the highest b-ness to be within [20,300] GeV/c^2. By construction the overall expected rates are the same as the observed ones in each region. The expected rates are separated for different triggered lepton type. The calculation of uncertainties on these rates is still on going

- $V + jets$ ($V = W, Z$): We allow the $V + jets$ background rates to float unconstrained in each region. Such rates are varied independently in the tag and notag regions, but they are tied together between the regions with 2 and 3 jets in the final state. The chosen configuration allows us to avoid uncertainties due to the relative fraction of heavy flavor jets in the various channels, but constraints the relative normalization of the regions with different jet mutiplicity to the values expected from the theory.
- Top: We constrain top normalization to the measured cross-sections reported in [21], ($\sigma^{t\bar{t}} = 7.65 \pm 0.42\ pb$, $\sigma^{singletop} = 3.04 \pm^{0.57}_{0.53}\ pb$). Singletop normalization is constrained to the theoretical cross section with an uncertainty of 6%.
- WW: We constrain the WW normalization to its theoretical cross section with an uncertainty of 6%.
- QCD: the normalization is constrained within 14.1% for the TCE, CMUPCMX, EMC samples and 17.9% for the PHX samples (Sec. 2.9.1)
- Signal: Our signal normalization is allowed to float unconstrained in the fit, but unlike the W + jets background, not independently between channels.

Templates normalized to the expected rates (Tables 6-7) and stacked together are shown in Fig. 25-26 along with the data.

3.1 Systematic uncertainties

In performing the fit to the data, we simultaneously fit for systematic uncertainties (also called nuisance parameters), which affect the Z-invariant mass distribution. Besides the rate uncertainties mentioned in Sec 3, we include nuisance parameters for other rate and shape uncertainties.

The shape uncertainties for the $W + jets$, $Z + jets$, and top backgrounds are handled by supplying 2 more templates, corresponding to upward or downward fluctuations of the nuisance parameter, which have been previously symmetrized. We then use the mclimit's vertical morphing technique to morph the shapes between the -1σ and $+1\sigma$ shapes, and beyond. By construction the morphing obtains the central shapes at 0 σ values of all nuisance parameters.

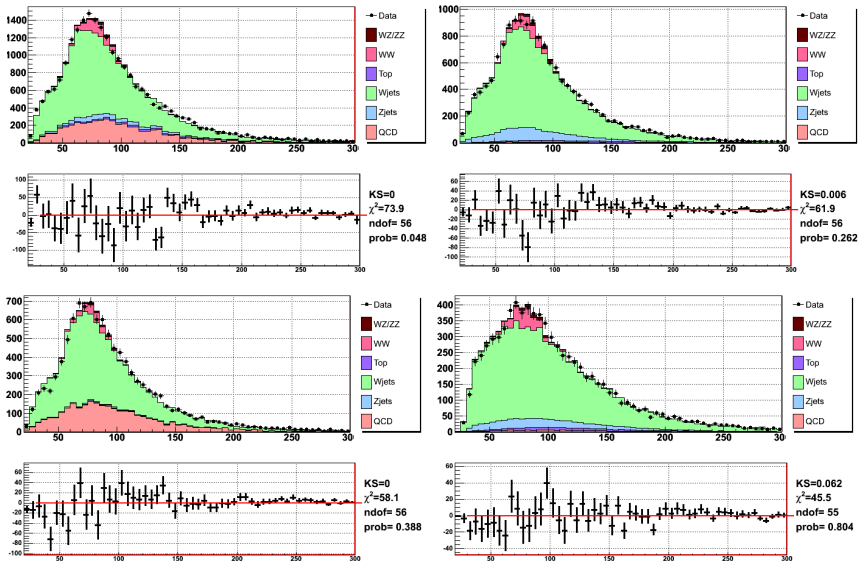

Figure 25: $Mj1j2$-invariant mass in the notag2j samples. From left to the right, from top to the bottom we show TCE, CMUPCMX, PHX, EMC samples. Templates used for the signal extraction are shown in solid colors. Templates have been normalized to the expected rates. Data distributions (dots) are superimposed and differences between predictions and data are plotted in the lower pads

To handle the shape systematics for the WW background and signal we provide an analytic parameterization of the shapes. The parameterization is obtained as follows:

- we fit the central shapes with a Gaussian function on top of a 4^{th}-degree polynomial
- fit the $\pm\sigma$ shapes with the same function. This time the coefficients of the polynomial are constrained to be equal to the ones obtained for the central shapes.
- fit the gaussian parameters (means, standard deviations, and amplitudes) with a linear function. The line is forced to pass through the central value parameters to ensure to have the central shapes at 0 σ.

By doing so, we avoid to use any histogram morphing in mclimit, which can cause kinks in the χ^2 minimization.

The QCD shape uncertainty is handled by allowing for fluctuations of the QCD template within the provide error. More details are available in Sec. 2.8.1.

We list below the considered systematic uncertainties.

- *JES*: JES is raised by $\pm 1\sigma$ [18]. The sign of the uncertainty goes with the change in quark JES: $+1\sigma$ JES implies $+2.7\%$ applied to quark jet energies, and -4.8% applied to gluon jet energies, while -1σ JES implies -2.7% for quark jets, and $+4.8\%$ for gluon jets. Compared to the nominal rates (Table 6-7) different number of events for each processes and different shapes for the templates used in the fit are obtained. See tables 8, for rate differences in regions notag2j and tag2j regions.
- *b−tagger efficiency/mistag rate*: rather than using the $\pm 1\sigma$ uncertainty while fitting for the signal, we vary the bness cuts in MC and we use the cuts which match the $\pm 1\sigma$ uncertainty on the mistag rate and b-tag efficiency. Those cuts are summarized in Table 2. The rate differences from the nominal rates due to the varied bness cuts are used as a nuisance parameter [19]. Rate uncertainties assigned to each process are summarized in tables 8. Shape changes due to uncertainties on the bness cuts are negligible.

[18] We consider the JES uncertainty only due to L9 corrections.

[19] Changing the bness cuts will move events from the no-tag channel to the double-tag channel (or vice-versa)

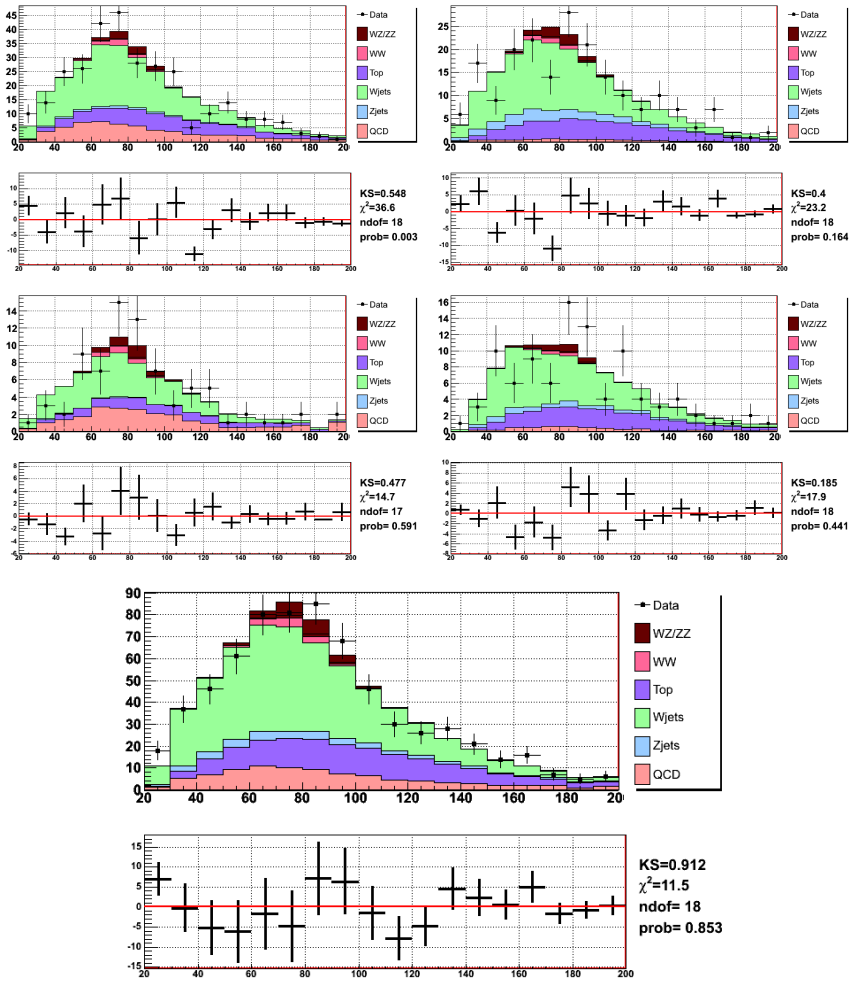

Figure 26: $Mj1j2$-invariant mass in the tag2j samples. From left to the right, from top to the bottom we show TCE, CMUPCMX, PHX, EMC samples. The bottom plot shows all the lepton catories combined together. Templates used for the signal extraction are shown in solid colors. Templates have been normalized to the expected rates. Data distributions (dots) are superimposed and differences between predictions and data are plotted in the lower pads

- *Renormalization and Factorization Scale in the W+jets MC* (Q^2): the `ALPGEN` event generator used for W+jets events requires the renormalization and factorization scale, Q^2, to be set to solve the divergences caused by gluon splitting. As default, the scale is set to $Q^2 = {M^W}^2 + \sum p_T^2$, where M^W is the W boson mass and p_T^2 is the parton transverse energy squared and the sum extends over all the final state partons. This parameter is doubled and halved to create two samples which are used as a shape uncertainty on W+jets template.

In addition to these systematics, we apply the following systematic uncertainties on the acceptance to our signal, and WW processes [20].

- *JER*: Smearing the dijet mass due to energy resolution effects produce an overall normalization difference of 0.7% [3] [21].
- *Luminosity:* The standard 6% systematics is considered.
- *ISRFSR:* Following [3], we assign a 2.5% systematic uncertainty due to more/less ISR and FSR.
- *PDF:* We quote the same systematics used in $WH \to l\nu b\bar{b}$ that was found to be 2.0% [9]
- *Trigger*: We quote 2.2% as systematic uncertainties due to the trigger uncertainty

Systematic	Region	Signal	WW	Top	W+ jets	Z+ jets	QCD
JES shape/rate	Tag2j	yes/$\pm^{1.8\%}_{5.2\%}$	yes/$\pm^{2.7\%}_{1.7\%}$	yes/$\mp^{4.3\%}_{3.7\%}$	yes/no	yes/$\mp^{4.3\%}_{3.7\%}$	no/no
	Notag2j	yes/$\pm^{2.8\%}_{3.5\%}$	yes/$\pm^{3.9\%}_{4.5\%}$	yes/$\pm^{0.2\%}_{1.8\%}$	yes/no	yes/$\mp^{0.0\%}_{0.7\%}$	no/no
bness cuts rate	Notag2j	$\pm^{X.X\%}_{X.X\%}$	$\pm^{X.X\%}_{X.X\%}$	$\pm^{X.X\%}_{X.X\%}$	$\pm^{X.X\%}_{X.X\%}$	$\pm^{X.X\%}_{X.X\%}$	no
	Tag2j	$\pm^{X.X\%}_{X.X\%}$	$\pm^{X.X\%}_{X.X\%}$	$\pm^{X.X\%}_{X.X\%}$	$\pm^{X.X\%}_{X.X\%}$	$\pm^{X.X\%}_{X.X\%}$	no
Q^2	Tag2j	no	no	no	yes	no	no
	Notag2j	no	no	no	yes	no	no
Acceptance	Tag2j	$\pm^{7.1\%}_{7.1\%}$	$\pm^{7.1\%}_{7.1\%}$	$\pm^{7.1\%}_{7.1\%}$ (single-top only)	no	no	no
	Notag2j	$\pm^{7.1\%}_{7.1\%}$	$\pm^{7.1\%}_{7.1\%}$	$\pm^{7.1\%}_{7.1\%}$ (single-top only)	no	no	no

Table 8: List of considered systematic uncertainties for the nota2j and tag2j regions for each of the different templates used in the fit. At this stage we use average systematic uncertainties, without distinguishing among the different lepton categories.

3.2 Expected Sensitivity

Fig. 3.2 shows the $\Delta\chi^2$ between null and test hypothesis fits for pseudo-data generated with and without the WZ/ZZ signal. The median p-value is $\sim (6.61 \pm 0.05)$ %, which translates in $1.505 \pm^{0.004}_{0.004}$ σ

[20] $t\bar{t}$ rate is not affected by the uncertainty on the luminosity since the CDF measured $t\bar{t}$ cross-section already includes this uncertainty. $V + jets$ is a free parameter of the fit

[21] We are working in computing the normalization difference due to JER smearing for this analysis

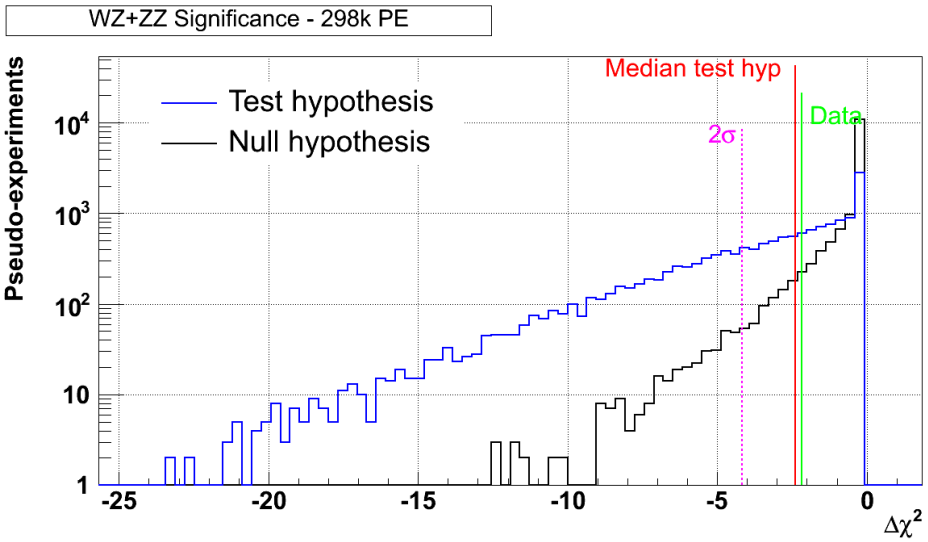

Figure 27: Plot of the $\Delta\chi^2$ distribution for 298000 PEs generated with (blue) and without (black) WZ/ZZ signal. Median of the test hypothesis fit (red line) distribution is shown as well as the measured value (green line). The median p-value, computed as the integral on the null hypothesis distribution from the lower boundary to the median, is $\sim$1.5 σ

3.3 Fit Results on Data

The result of the fit to the data is shown in Table 9. We fit for a signal rate which is consistent within uncertainties with the SM predictions. In Fig. 28 we plot the templates stacked together and adjusted for fitted shape and rate systematic variations. The data distributions is superimosed. A good post-fit modeling of the data is achieved.

To translate our fit to the data to bounds or limits on the true cross section of WZ+ZZ production, we construct Feldman-Cousins bands by analyzing the distribution of fitted (i.e., measured) cross sections in pseudo-experiments generated with a variety of scale factors on the input signal cross section. The set of input cross sections in our pseudo-experiments range from 0.1 to 4.0 times the standard model value with a step size of 0.1.

For each set of pseudo-experiments, we find a range of measured cross-sections which meets a coverage threshold of 68% and 95%. Fig. 29 shows the results of the Feldman-Cousins bands. We find $\sigma_{measured} = (0.9\pm^{0.7}_{0.4}) \times \sigma_{SM}$. Using $\sigma_{SM} = 5.54\ pb$ we calculate $\sigma_{measured} = 5.04\pm^{3.9}_{2.2}$. Since the sensitivity of this measurement is limited, we cannot exclude the possibility of no $WZ + ZZ$ production. Therefore we set a limit of $\sigma_{measured} < 13.30\ pb$ $(2.4 \times \sigma_{SM})$ at 95% CL.

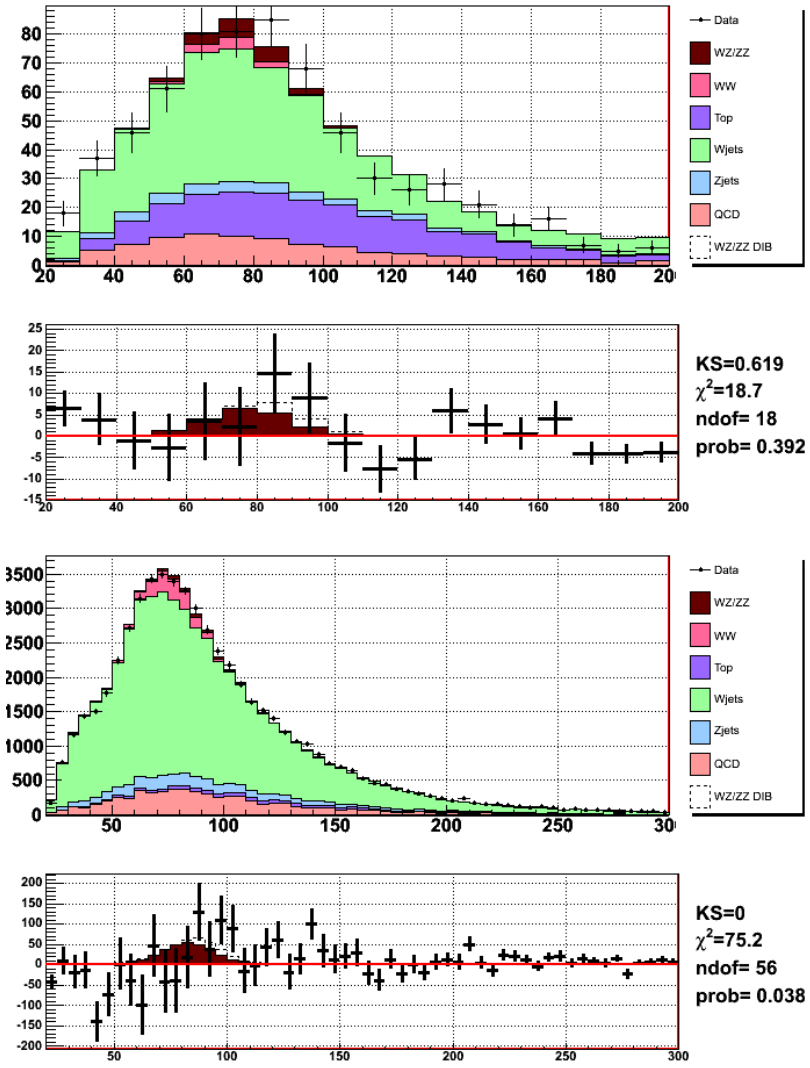

Figure 28: Result of the fit to the data. Fitted $Mj1j2$ invariant mass distributions in the Tag2j (upper), Notag2j (bottom), are superimposed to the data. Both shapes and rates have been adjusted to the results obtained in data. The fitted signal rate is consistent within uncertainties with SM prediction.

Parameter	Fit value (in units of σ or %)
TOP	0.95 ± 0.97
BNESS	0.20 ± 0.37
JES	-1.11 ± 0.35
Q2	-1.52 ± 0.56
WW	-1.08 ± 0.97
Acceptance	-1.11 ± 0.95
Signal	-0.07 ± 0.67
VJETSTag2j	-0.06 ± 0.08
VJETSNotag2j	0.02 ± 0.01
QCDTCETag2j	0.06 ± 0.94
QCDTCENotag2j	-0.26 ± 0.39
QCDCMUPCMXTag2j	0.02 ± 1.00
QCDCMUPCMXNotag2j	-0.09 ± 0.99
QCDPHXTag2j	0.01 ± 0.97
QCDPHXNotag2j	-0.50 ± 0.25
QCDEMCTag2j	-0.02 ± 1.00
QCDEMCNotag2j	-0.02 ± 0.98

Table 9: Fitted nuisance parameters to the data. $VJETSTag2j$, $VJETSNotag2j$ ($V = W, Z$) and $Signal$ values are given in terms of % of the normalization of the template, while the other parameter results are given in units of σ. We fit for a signal rate which is consistent within uncertainties with SM prediction.

A MC sample names

In Table A we list the MC samples used in this analysis.

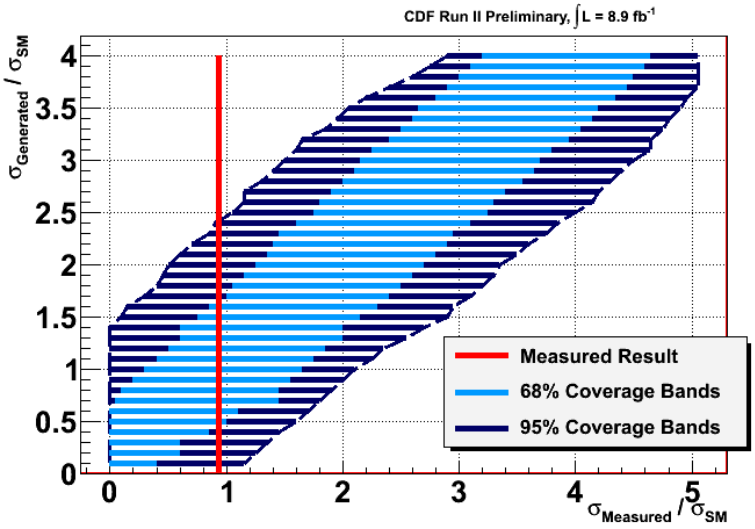

Figure 29: Feldman-Cousins bands for the $M(jet1, jet2)$ mass fit, showing the expected range of measured cross sections as a function of the true cross section, with 68% CL (light blue) and 95% CL (dark blue). Our measured result is $\sigma_{measured} = (0.9 \pm^{0.7}_{0.4}) \times \sigma_{SM}$, and we establish a limit of 1.25 $\times \sigma_{SM}$ with 95% CL.

B Optmization of jet bness cuts

To optimize the bness cuts for this analysis we use the probability of getting 2σ P-value in data. We name it "P2sigma". For a integrated luminosity of about 7 fb^{-}1 P2sigma $\sim$ 14% in the pretag 2jet sample (Sec. 2). We optimize the bness cuts as follows:

1. We order jets in decreasing bness. We consider $WZ \to l\nu b\bar{b}$ as signal and we vary the bness cuts for the two jets in the sample. We estimate the P2sigma for each set of cuts, as shown in Fig. 30. Obtained optimal bness cuts are 0.75,-0.2 for jet1 and jet2 respectively. We call the region selected with such cuts "Reg1".

2. from the pretag sample we subtract Reg1. We decide a bness1 cut, therefore selecting a region which we name Reg2. We estimate the P2sigma to extract the $WZ \to l\nu q\bar{q}$ by fitting in both Reg1 and Reg2. We repeat step 2 by varying bness1 cut. As expected the optimal bness1 cut is -1, which is the same as not assessing any cut.

P2sigma obtained by fitting for $WZ \to l\nu q\bar{q}$ in Reg1 ($bness1 > 0.75, bness2 > -0.2$) and Reg2 (pretag without Reg1) is 18%. Therefore, we improve the sensitivity of the measuremen by about 30% when fitting in regions Reg1 and Reg2 simulataneoulsy rather than the pretag region alone

Lepton Identification Cuts

C Electron Identification Cuts

Electron reconstruction starts when a cluster [22] in the electromagnetic calorimeter is found. A candidate electron is identified if the cluster is matched with at least one track. The type of electrons used in this analysis are required to pass the following identification cuts:

where:

[22] Clusters are made of small, contiguous groups of calorimeter towers with energy deposit.

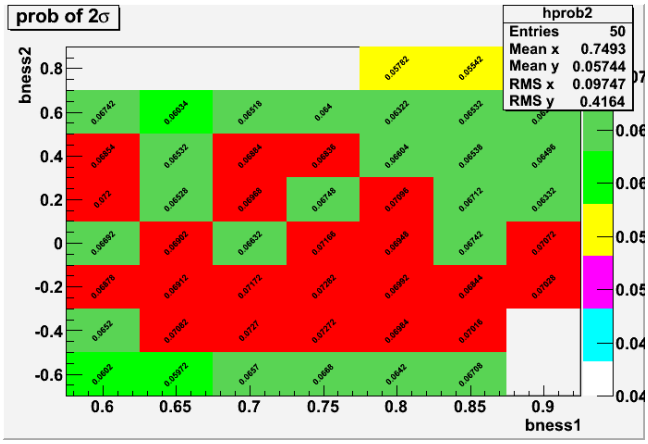

Figure 30: *P2sigma to extract the $WZ \to l\nu b\bar{b}$ as function of bness cuts on the first jet (bness1) and second jet (bness2). WZ MC pretag 2 jet selected sample was used (Sec.* **??** *for details about the selection). Templates are scale to match an integrated luminosity of about 7/fb.*

- $\mathbf{E_{had}/E_{em}}$ is the ratio between the energy deposited in the electromagnetic and hadronic calorimeters,
- **Isolation** is the ratio between the additional transverse energy in a cone of radius $R = 0.4$ around the cluster and the transverse energy of the cluster itself,
- **track $\mathbf{z_0}$** is the position along the beamline of the track at the point of intersection with the beamline itself,
- **track $\mathbf{P_T}$** is the transverse momentum of the track,
- **COT Axial and Stereo Superlayers** are, respectively, the number of axial and stereo superlayers in the COT which have at least 5 hits attached to the track,
- χ^2 of the fit to electron test beam data for shower-maximum profile,
- $\mathbf{L_{shr}}$ is a variable that measures how close the energy in the electron towers adjacent to the cluster seed is to the electron hypothesis,
- **E/p** is the ratio between the cluster energy and the track momentum,
- **Q** is the measured charge of the particle,
- $\mathbf{\Delta X}$ is the signed difference in x between the track and the cluster associated to the electron when the track is extrapolated to the position of the shower max,
- $|\mathbf{\Delta z}|$ is the absolute value of the difference in z position between the cluster and the extrapolated track.

D Muon Identification Cuts

Muons are characterized by a track identified in the tracking system, small energy [23] deposited in the traversed calorimeter, and a short track segment in the muon chambers. The type of muons used in this analysis are required to pass the identification cuts in table D

[23] consistent with a minimum ionizing particle at a given energy

where:

- Track d_0 is the impact parameter of the tracks defined as the distance of the closest approach to the fitted track to the beamline.
- ρ_{COT} is the radius at which the track appears to leave the COT.
- the rest of variable is defined in Sec. C

E PHX Identification Cuts

F QCDCalibrations

F.1 PHX Notag2j

F.2 TCE Notag3j

F.3 PHX Notag3j

G Fit results - Bump cuts

This section is included in this note to mimic what was described in [24].

Compared to selection described in Sec. 2 we changed the following cuts:

- lepton categories: TCE, CMUP, CMX (no PHX, EMC)
- esclusively 2 jets with $E_T > 30\ GeV$ and $|\eta| < 2$
- no $dR(jet1, jet2) < 3.5$ (still keeping $dR(jet1, jet2) > 0.7$)
- $\not\!\!E_T > 25\ GeV$
- $M_T^W > 30\ GeV$ also for the muon sample
- $\Delta\ \phi(\not\!\!E_T, jet1) > 0.4$
- $\Delta\eta(jet1, jet2) < 2.5$

$M(jet1, jet2)$ fits are performed in the same way as described in Sec. 3, except for the differences described below:

- The fit is now performed by using 2 channels (high-P_T electron and muon trigger samples) rather than 8.
- Rather than being unconstrained, WZ/ZZ normalization is constrained to its theoretical cross section with an uncertainty of 6%
- a gaussian distribution centered at 145 GeV with a width of of 14.3 GeV is included as a template of the fit. The normalization of such a distribution is unconstrained.

The result of the fit to the data is shown in Table 14. The fit results are consistent with no resonance at 145 GeV. We then proceed to establish an upper limit of 0.9 pb at 95% CL on the cross-section of such a resonance. Feldman-Cousin bands are shown in Fig. 31

In Fig. 32 we plot the templates stacked together and adjusted for fitted shape and rate systematic variations. The data distributions is superimosed. A good post-fit modeling of the data is achieved.

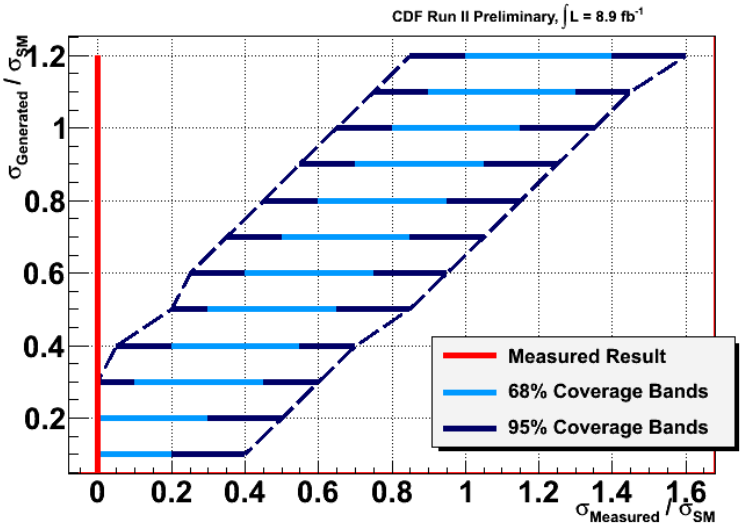

Figure 31: Feldman-Cousins bands for the $M(jet1, jet2)$ mass fit in the sample selected by applying the bump cuts (see text). In the plot we show the expected range of measured cross sections as a function of the true cross section, with 68% CL (light blue) and 95% CL (dark blue). Since no significant excess at 145 GeV is found, we establish a limit of $0.3 \times 3.1\ pb = 0.9\ pb$ with 95% CL.

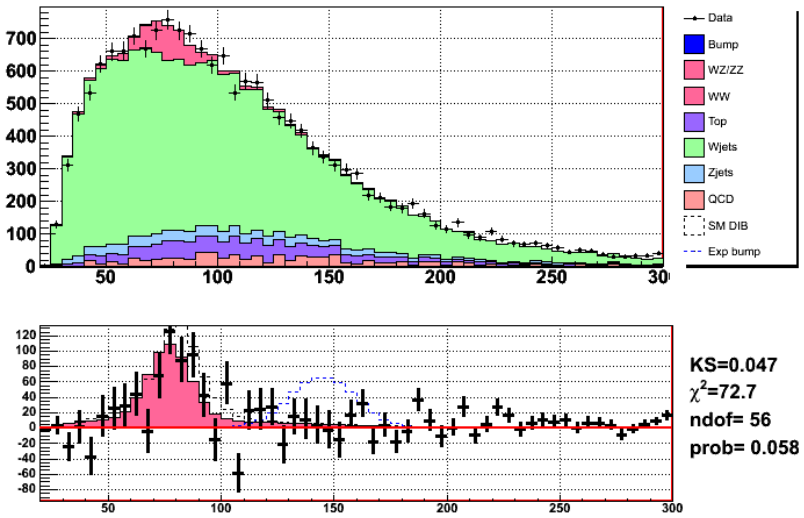

Figure 32: Result of the fit to the data. Both shapes and rates have been adjusted to the results obtained in data. The fitted rate for the resonance at 145 GeV is consistent within uncertainties with SM prediction.

References

[1] J. M. Campbell and R. K. Ellis *Update on Vector Boson Pair Production at Hadron Colliders* Phys. Rev. D **65** (2002) 113007.

[2] K. Nakamura et al. (Particle Data Group) J. Phys. G **37** (2010) 075021.

[3] Search for WZ/ZZ in E/T + $b\bar{b}$ channel, CDF Note 10631.

[4] ketchum, lewis, poprocki, freeman, pronko, rusu, wittich CDF/ANAL/ELECTROWEAK/CDFR/10204 (2010).

[5] R. Culbertson *et al.*, *Search for Anomalous Production of di-photon+MET Events in 2/fb of Data.*, CDF NOTE 9184

[6] J. Adelman et al., "Method II For You", CDF Internal Note 9185, 2008.

[7] Frankiln, *et al.*, *Calibration of Heavy-Flavor Production in QCD Data*, CDF NOTE 8768
Frankiln, *et al.*, *Heavy-Flavor Content of the W+Jets Sample*, CDF NOTE 8765

[8] Tom Junk *et al.* Combined CDF and DZero Upper Limits on Standard Model Higgs Boson Production with up to 6.7 fb^{-1}
CDF NOTE 10241

[9] J. Dittman et al., CDF Note 9401.

[10] W. Ketchum, V. Rusu, Y.K. Kim, CDF Note 10581

[11] Sensitivity, Exclusion and Discovery with Small Signals, Large Backgrounds, and Large Systematic Uncertainties, CDF note CDF/DOC/STATISTICS/PUBLIC/8128 (2007)

[12] G. Bellettini, G. Latino M. Trovato,V. Rusu, G. Velev, C. Vernieri, CDF Note 10642

[13] A. Buzatu,T. Aaltonen,B. Alvarez,G. Chiarelli,J. Dittmann,M. Frank,C. Group,R. Hughes,J. Keung,S. Kim, B.Kilminster,N. Krumnack,K. Lannon,S. Leone,J. Lueck,Y. Nagai,C. Neu,M. Ronzani,F. Sforza, J. Slaunwhite, A. Taffard, E. Thomson,J. Vizan,A. Warburton,B. Winer,H. Wolfe,Z. Wu,W. Yao, CDF Note 10440

[14] W. Ketchum, V. Rusu, M. Trovato, CDF Note 10829

[15] Sensitivity, Exclusion and Discovery with Small Signals, Large Backgrounds, and Large Systematic Uncertainties, CDF note CDF/DOC/STATISTICS/PUBLIC/8128 (2007)

[16] Giorgio Bellettini, Matteo Cremonesi, Giuseppe Latino, Vadim Rusu, Marco Trovato, George Velev, Caterina Vernieri, CDF Note 10838

[17] The CDF Collaboration: First Measurements of Inclusive W and Z Cross Sections from Run II of the Tevatron Collider Phys. Rev. Lett. 94, 091803 (2005)

[18] P.J. Sutton, A.D. Martin, R.G. Roberts and W.J. Stirling, Phys. Rev. **D45**, 2349 (1992); P.J. Rijken and W.L. van Neerven, Phys. Rev. **D51**, 44 (1995); R. Hamberg, W.L. van Neerven and W.B. Kilgore, Nucl. Phys. **B359**, 343 (1991); R.V. Harlander and W.B. Kilgore, Phys. Rev. Lett. **88**, 201801 (2002).

[19] Bachacou, Ferretti, Nielsen, Yao, "Heavy Flavor Contributions to the SECVTX-tagged W + Jets Sample", CDF7007, April 10, 2007; D. Acosta, et al., Phys. Rev. **D71**, 052003 (2005)

[20] $http://www-cdf.fnal.gov/internal/physics/joint_physics/instructions/JPScaleFactor/$

[21] Evelyn Thomson *et. al* Combination of CDF and D0 ttbar cross-sections, CDF note 10916

[22] Zhenbin Wu *et. al* Measurement of Single Top Quark Production in 7.5/fb of CDF Data Using Neural Networks, CDF note 10703

[23] W. Ketchum V. Rusu M. Trovato New Jet Energy Scale Corrections for Quarks and Gluons, CDF note 10829

[24] Phys.Rev.Lett.106:171801,2011

Stnuple Sample	Cross section (pb)	Generator	Process
it0sww (ht0sww)	12.4	PYTHIA	WW
it0swz (ht0swz)	3.7	PYTHIA	WZ
it0szz (ht0szz)	1.38	PYTHIA	ZZ
tt1s25	7.04	PYTHIA	$t\bar{t}$
st0s26 (st0s23)	0.29	MadEvent+PYTHIA	single-top (s-channel)
st0s28 (st0s25)	0.64	MadEvent+PYTHIA	single-top (t-channel)
pt0sw0 (ut0s00)	2520	ALPGEN+PYTHIA	$W(e\nu) + 0p$
pt0sw1 (ut0s01)	315	ALPGEN+PYTHIA	$W(e\nu) + 1p$
pt0s2w (ut0s02)	49.42	ALPGEN+PYTHIA	$W(e\nu) + 2p$
pt0s3w (ut0s03)	7.826	ALPGEN+PYTHIA	$W(e\nu) + 3p$
pt0s4w (ut0s04)	1.442	ALPGEN+PYTHIA	$W(e\nu) + \geq 4p$
bt0s0w (bt0s00)	5.841	ALPGEN+PYTHIA	$W(e\nu) + bb + 0p$
bt0s1w (bt0s01)	1.740	ALPGEN+PYTHIA	$W(e\nu) + bb + 1p$
bt0s2w (bt0s02)	0.5625	ALPGEN+PYTHIA	$W(e\nu) + bb + \geq 2p$
ct0s0w	7.00	ALPGEN+PYTHIA	$W(e\nu) + cc + 0p$
ct0s1w	2.506	ALPGEN+PYTHIA	$W(e\nu) + cc + 1p$
ct0s2w	0.879	ALPGEN+PYTHIA	$W(e\nu) + cc + \geq 2p$
METTERE I MUONI E TAU			
zt0sp0	221.2	ALPGEN+PYTHIA	$Z(ee) + 0p;\ M_Z = [75, 105]\ GeV/c^2$
zt0sp1	30.24	ALPGEN+PYTHIA	$Z(ee) + 1p;\ M_Z = [75, 105]\ GeV/c^2$
zt0szb	4.844	ALPGEN+PYTHIA	$Z(ee) + 2p;\ M_Z = [75, 105]\ GeV/c^2$
zt0s3p	0.77	ALPGEN+PYTHIA	$Z(ee) + 3p;\ M_Z = [75, 105]\ GeV/c^2$
zt0s4p	0.13888	ALPGEN+PYTHIA	$Z(ee) + \geq 4p;\ M_Z = [75, 105]\ GeV/c^2$
xt0s0p	224	ALPGEN+PYTHIA	$Z(ee) + 0p;\ M_Z = [20, 75]\ GeV/c^2$
xt0s1p	11.746	ALPGEN+PYTHIA	$Z(ee) + 1p;\ M_Z = [20, 75]\ GeV/c^2$
xt0s2p	2.254	ALPGEN+PYTHIA	$Z(ee) + 2p;\ M_Z = [20, 75]\ GeV/c^2$
xt0s3p	0.3262	ALPGEN+PYTHIA	$Z(ee) + 3p;\ M_Z = [20, 75]\ GeV/c^2$
xt0s4p	0.05572	ALPGEN+PYTHIA	$Z(ee) + \geq 4p;\ M_Z = [20, 75]\ GeV/c^2$
yt0s0p	5.698	ALPGEN+PYTHIA	$Z(ee) + 0p;\ M_Z = [105, 600]\ GeV/c^2$
yt0s1p	0.9884	ALPGEN+PYTHIA	$Z(ee) + 1p;\ M_Z = [105, 600]\ GeV/c^2$
yt0s2p	0.1638	ALPGEN+PYTHIA	$Z(ee) + 2p;\ M_Z = [105, 600]\ GeV/c^2$
yt0s3p	0.0259	ALPGEN+PYTHIA	$Z(ee) + 3p;\ M_Z = [105, 600]\ GeV/c^2$
zt0sb0	1.4308	ALPGEN+PYTHIA	$Z(ee) + bb + 0p;\ M_Z = [75, 105]\ GeV/c^2$
zt0sb1	0.3752	ALPGEN+PYTHIA	$Z(ee) + bb + 1p;\ M_Z = [75, 105]\ GeV/c^2$
zt0sb2	0.1078	ALPGEN+PYTHIA	$Z(ee) + bb + \geq 2p;\ M_Z = [75, 105]\ GeV/c$
zt0sc0	1.512	ALPGEN+PYTHIA	$Z(ee) + cc + 0p;\ M_Z = [75, 105]\ GeV/c^2$
zt0sc1	0.463	ALPGEN+PYTHIA	$Z(ee) + cc + 1p;\ M_Z = [75, 105]\ GeV/c^2$
zt0sc2	0.150	ALPGEN+PYTHIA	$Z(ee) + cc + \geq 2p;\ M_Z = [75, 105]\ GeV/c$

Identification Cut
Fiducial to CEM
$E_{had}/E_{em} < 0.055 + 0.00045 \cdot E$
Isolation ≤ 0.1
-60 cm $\leq$ Track $z_0 \leq$ 60 cm
Track $P_T \geq 10$ GeV/c
COT Axial Segments ≥ 3
COT Stereo Segments ≥ 3
$L_{shr} \leq 0.2$
$E/p < 2.5 + 0.015 \cdot E_T$
$-3.0 \leq Q \cdot \Delta X \leq 1.5$
$\lvert\Delta z\rvert \leq 3.0\ cm$

Table 11: Central Electron identification cuts.

Identification Cut
$E_{em} \leq 2.0 + max(0, 0.0115(p - 100))\ GeV$
$E_{had} \leq 6.0 + max(0, 0.0280(p - 100))\ GeV$
Isolation ≤ 0.1
-60 cm $\leq$ Track $z_0 \leq$ 60 cm
Track $d_0 \leq 0.2$ cm (0.02 cm w/ silicon hits)
COT Axial Segments ≥ 3
COT Stereo Segments ≥ 2
$\chi^2 \leq 2.3$ (2.75 for particular data periods)
$\rho_{COT} > 140\ cm$

Table 12: Some of the Central Muon identification cuts. More identification cuts are specified in [**?**].

Identification Cut
Fiducial to PEM
$1.13 < \eta_{PES} < 2.8$
PEM 3x3 $\chi^2 < 25$
PES 5x9 U > 0.65
PES 5x9 V > 0.65
$E_{had}/E_{em} < 0.05$
Isolation ≤ 0.1
ΔR between PES and PEM centroids < 3.0 cm
PEM 3x3 Tower Fit > 1.0
60 cm $\leq$ Track $z_0 \leq 60$ cm
Silicon Hits ≥ 3

Table 13: Forward Electron identification cuts.

Parameter	Fit value (in units of σ or %)
TOP	0.53 ± 0.98
BNESS	0.36 ± 0.90
JES	-1.15 ± 0.42
Q2	0.31 ± 0.93
WW+WZ+ZZ	-0.40 ± 0.96
Acceptance	0.37 ± 0.93
VJETS	0.01 ± 0.02
QCDTCE	-0.42 ± 0.71
QCDCMUPCMX	0.00 ± 1.00
Resonance (145 GeV)	-1.00 ± 0.22

Table 14: Fitted nuisance parameters to the data. $VJETS$, and *Resonance* values are given in terms of % of the normalization of the template, while the other parameter results are given in units of σ. We fit for no resonance at 145 GeV.