

A New $H \rightarrow WW \rightarrow l\nu jj$ Analysis Adding a New Muon Gap Trigger Using 5.7 fb^{-1} of data

Marco - Andrea Buchmann^{1 7},
 Martina Hurwitz²,
 Florencia Canelli^{2 3},
 Ricardo Eusebi, Craig Group, Enrique Palencia³,
 Bruno Casal⁴,
 Bárbara Álvarez⁵,
 Bernd Stelzer⁶,
 Rainer Wallny⁷,
 Günther Dissertori¹

¹ *ETH Zurich, Switzerland*

² *University of Chicago*

³ *Fermilab*

⁴ *Instituto de Fisica de Cantabria (CSIC-Univ. de Cantabria)*

⁵ *Universidad de Oviedo*

⁶ *Simon Fraser University*

⁷ *University of California, Los Angeles*

Abstract

We present a search for Standard Model Higgs boson production in the $H \rightarrow WW \rightarrow l\nu jj$ channel using 5.7 fb^{-1} of CDF II data. A matrix element technique is used to calculate event probability densities for signal and background hypotheses, which are used to calculate a discriminant. The distributions of this discriminant for signal and background are fit to the data using a binned likelihood approach to search for the Higgs boson signal.

We observe no evidence for a Higgs signal and set 95% confidence level upper limits on the production cross section times the branching ratio for $\sigma(p\bar{p} \rightarrow H) \times BR(H \rightarrow WW)/SM < 12.6$ to 76.0 for Higgs boson masses between $m_H = 150 \text{ GeV}/c^2$ to $m_H = 200 \text{ GeV}/c^2$. The expected (median) limit estimated in pseudo-experiments is $\sigma(p\bar{p} \rightarrow H) \times BR(H \rightarrow WW)/SM < 12.9$ to 42.5 at 95% C.L for Higgs masses between $m_H = 150 \text{ GeV}/c^2$ and $m_H = 200 \text{ GeV}/c^2$. We also describe the first use of the phi-gap trigger for the $W+2$ jet topology.

Contents

1	Introduction	3
2	Event Selection	4
3	Data sample	5
4	Monte Carlo samples and background estimate	6
4.1	Monte Carlo samples	6
4.2	Background estimate	6
4.2.1	Monte Carlo-based estimate	6
4.2.2	Non-W background estimate	8
4.3	Event yields	8
5	New Lepton Categories	10
6	Measurement technique	12
6.1	Templates	13
6.2	Validation of input variables	14
7	The Likelihood Function	28
8	Systematic Uncertainties	29
8.1	Sources of systematic uncertainty	29
8.1.1	Background normalization	29
8.1.2	Signal normalization	29
8.1.3	Jet energy scale	29
8.1.4	PDFs	30
8.1.5	ISR and FSR	30
8.1.6	Q^2 scale	30
8.1.7	Luminosity and lepton ID efficiency	30
8.2	Integration of systematic uncertainties in likelihood fit	33
9	Result with CDF II Data	34
10	Conclusions	36

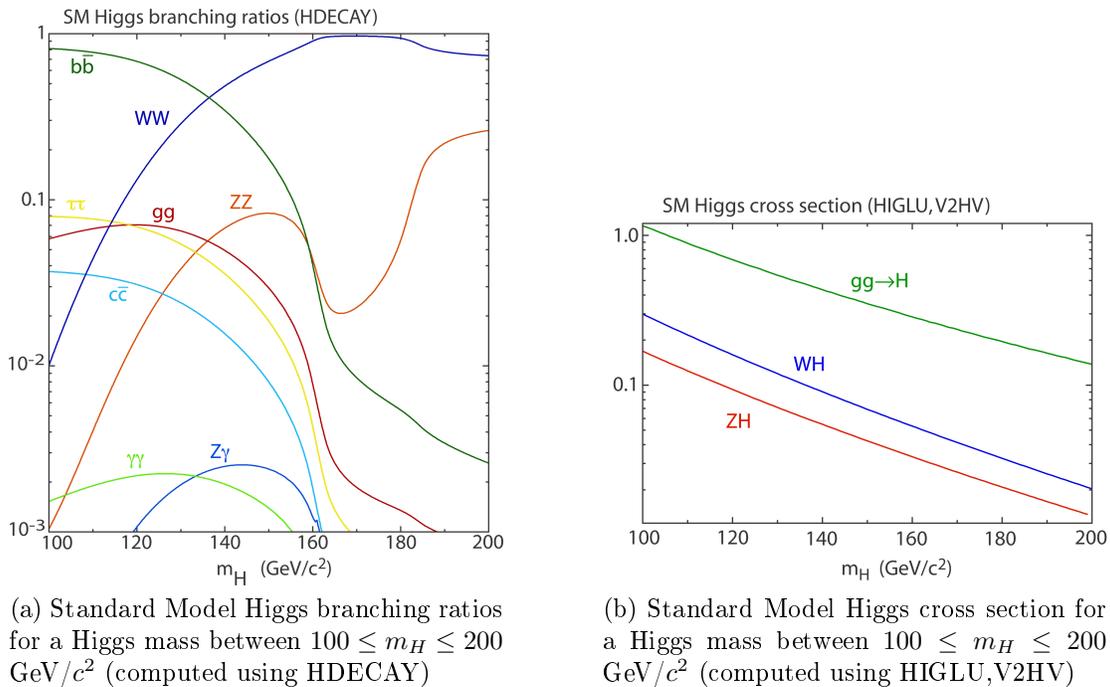


Figure 1: Higgs branching ratios and cross section [14][15]

1 Introduction

We present a search for $H \rightarrow WW$ where one W boson decays leptonically and the second W boson decays hadronically ($WW \rightarrow l\nu jj$).

The Higgs particle's production cross section and branching ratios for masses between $100 \leq m_H \leq 200$ GeV/c^2 are illustrated in Fig. 1a. One can see that up to about 140 GeV/c^2 , the $b\bar{b}$ channel is dominant ("light Higgs") after which the WW channel becomes dominant ("heavy Higgs") and even very dominant in the $160 < m_H < 180$ GeV/c^2 range. The present analysis focuses on the range between 150 and 200 GeV . The W 's branching ratios are given in Table 1; one can see that this channel is more inclusive than the dileptonic one, since the product of the branching ratios, multiplied by the combinatorial factor, is $2 \times \frac{2}{3} \times \frac{2}{9}$ instead of $\frac{1}{3} \times \frac{1}{3}$ for a dileptonic decay. However, this entails a significantly larger background which is dominated by W +jets events. These events in turn represent the major challenge in the analysis. The matrix element technique is used to calculate event probability densities for signal and background hypotheses and to build an event probability discriminant. The analysis is based on the observation and cross section measurement of $WW+WZ$ production with a semileptonic decay described in CDF note 9923 [18]. Much of the event selection and analysis technique is based on the Matrix Element search for WH production described most recently in CDF note 9861 [1] and the Matrix Element analysis used in the

branching ratios for W	
$e \nu_e$	$\frac{1}{9}$
$\mu \nu_\mu$	$\frac{1}{9}$
$\tau \nu_\tau$	$\frac{1}{9}$
hadronic	$\frac{2}{3}$

Table 1: Branching ratios for W

m_H	$\sigma(gg \rightarrow H)$	$\text{BR}(H \rightarrow WW)$
150	0.548	0.682
160	0.439	0.901
170	0.349	0.965
180	0.283	0.935
190	0.231	0.776
200	0.192	0.735

Table 2: Production cross section and branching ratios at the investigated Higgs mass points

observation of single top production described in CDF note 9711 [2]

We have added events from the phi-gap trigger to improve coverage in $\eta - \phi$ space. This addition is explained in section 5.

Section 2 describes the event selection. The data and Monte Carlo samples are described in Sections 3 and 4, respectively. They are followed by section 5 on the new trigger. Next, the matrix element methodology is discussed in Section 6. Section 7 introduces the likelihood function we use, and the treatment of systematic uncertainties is described in Section 8. The results are presented in Section 9.

2 Event Selection

Candidate events for this analysis are selected by requiring a $W + 2$ jet event topology where the W decays leptonically, $W \rightarrow e\nu_e$ or $W \rightarrow \mu\nu_\mu$. Muon events are triggered by the high- p_T CMUP and CMX trigger as well as through MET+2 jets trigger as described in [3]. Electron events are triggered by the high p_T CEM and PEM triggers. Plug (PHX) electrons are not used in this analysis because of difficulty in modeling the non- W background. The threshold of all high- p_T lepton triggers used is 18 GeV.

Events are required to have:

- Exactly one isolated electron or muon with offline E_T or $p_T > 20$ GeV. The electron and muon identification follow the standard tight lepton criteria described in detail in Ref [4].
- $\cancel{E}_T > 20$ GeV in events containing a muon and $\cancel{E}_T > 25$ GeV in events containing an electron. The \cancel{E}_T is corrected for identified muons. The higher \cancel{E}_T requirement for events with an electron reduces the size of the QCD multi-jet background, which is difficult to model.
- Exactly two jets, clustered with a cone size of $\Delta R < 0.4$, and with $|\eta_{detector}| < 2.0$. One jet must have $E_T > 30$ GeV and the other $E_T > 20$ GeV after jet corrections up to Level 5 have been applied.
- We require the dijet mass to be between $65 < M_{jj} < 95$ GeV. This cut was included at the level of method 2 table creation.

The following event vetos are applied:

- Events with a third jet having $E_T > 12$ GeV and $|\eta| < 2.0$ are rejected
- Events with a second lepton are rejected, where the second lepton can be a tight electron or muon or a PHX electron or a non-isolated lepton in the CEM, CMUP, or CMX.
- A Z boson veto is implemented: we search for a second very loosely identified lepton (e.g. an isolated track) with the opposite charge of the tight lepton. If the invariant mass of the two leptons is consistent with the Z mass, i.e. $76 < M_{ll} < 106$ GeV/ c^2 , the event is rejected.
- Conversion veto: events where an opposite-charge track is found close to the lepton track are rejected.
- Cosmic ray veto (in data only).
- QCD veto: we follow the single top implementation, described in Ref. [5]. Events with a muon are required to have $M_T(W) > 10$ GeV and events with an electron are required to have $M_T(W) > 20$ GeV, where $M_T(W)$ is the transverse mass of the \cancel{E}_T -lepton system. Additional cuts are placed on the \cancel{E}_T -significance and the angles between the \cancel{E}_T and the second jet.

3 Data sample

Data from periods 1 through 27 were used. We do not require Silicon for any events. We used the good run list `goodrun_em_mu_nosi_cmignored_v33.list`.

4 Monte Carlo samples and background estimate

4.1 Monte Carlo samples

A list of the Monte Carlo samples used is given in Table 3. Alpgen samples with Pythia parton showering are used to model the W +jets and Z +jets backgrounds. Pythia is used to model the WW and WZ processes.

The QCD multijet (often referred to as non- W) background is modeled using data [6]. Three types of samples are used: the anti-electron sample from the central electron trigger with events that fail two of the five tight electron cuts related to shower shapes; the jet-electron sample from the 20 GeV jet trigger where one jet looks electron-like (this is a very low-statistics sample), and non-isolated muons. In the modeling validation and when building the templates for the fit, the three samples are added together to achieve the maximal possible statistics. Furthermore, the MC is reweighted so that the number of level 12 vertices agrees with our data sample.

4.2 Background estimate

The procedure to estimate background levels is similar to the procedure followed for the pretag sample in the single top search, described in detail in CDF note [7]. The Monte Carlo is used to predict the shape and normalization of backgrounds where the MC modeling is expected to be adequate as well as the shape and normalization of the signals (HWW). A fit to the missing energy spectrum is then performed to determine the QCD (or non- W) fraction. The rest of the events are assumed to come from W +jets.

4.2.1 Monte Carlo-based estimate

The assumed cross sections for the backgrounds are also listed in Table 3. For the W and Z +jets samples, the relative cross sections of the W +np samples are taken from Alpgen. Then the total Z +jets cross section is taken from the CDF inclusive ($Z \rightarrow ll$)+jets measurement [12], and the total W +jets cross section is a free parameter in the final fit.

Process	Monte Carlo	Sample (LL)	Sample (HL)	Cross section
HWW	Pythia	ehht1k (150), ehht1m (160), ehht1o (170), ehht1q (180), ehht1s (190), ehht1u (200)	none	see Table 2
$(W \rightarrow e\nu)+\text{jets}$	Alpgen+Pythia	ptopw0, ptopw1, ptop2w, ptop3w, ptop4w	utop00, utop01, utop02, utop03, utop04	From data fit
$(W \rightarrow \mu\nu)+\text{jets}$	Alpgen+Pythia	ptopw5, ptopw6, ptop7w, ptop8w, ptop9w	utop05, utop06, utop07, utop08, utop09	
$(W \rightarrow \tau\nu)+\text{jets}$	Alpgen+Pythia	utopw0, utop1w, utop2w, utop3w, utop4w	utop10, utop11, utop12, utop13, utop14	
$(Z \rightarrow ee)+\text{jets}$	Alpgen+Pythia	ztopp0, ztopp1, ztop2p, ztop3p, ztop4p	none	787 ± 85 pb
$(Z \rightarrow \mu\mu)+\text{jets}$	Alpgen+Pythia	ztopp5, ztopp6, ztop7p, ztop8p, ztop9p	none	
$(Z \rightarrow \tau\tau)+\text{jets}$	Alpgen+Pythia	ztopt3, ztopt4, ztopt2	none	
Non- W	Data	Anti-electron, jet-electron, non-iso muons	none	From \cancel{E}_T fit
$t\bar{t}$	Pythia	ttop75, ttop75_1fb	none	6.7 ± 0.8 pb
Single top	Madevent + Pythia	stop00, stopm0, stop20, stop2m	none	2.864 pb
WW	Pythia	ihht1a	none	11.66 ± 0.7 pb
WZ	Pythia	jhht1a	none	3.46 ± 0.3 pb
ZZ	Pythia	khht1a	none	1.51 ± 0.2 pb

Table 3: MC samples used.

4.2.2 Non-W background estimate

The \cancel{E}_T spectrum observed in data is fit using a sum of expected shapes of signal and background processes. The normalization of the non- W and W +jets contributions are allowed to float in the fit, so that the expected relative normalization is extracted. The anti-electron sample is used to model the shape of the non- W contribution in the CEM, CMUP, CMX, CMP, and the extended muon categories. Furthermore, the non-iso sample is used for the extended categories (but not the CMP).

The fits are shown in Figure 2. The fit is done before the \cancel{E}_T cut has been imposed. The quoted non- W fractions, however, are calculated after the \cancel{E}_T cut. A conservative 40% uncertainty is imposed on the normalization of the non- W background. The W +jets normalization is also estimated from the missing energy fit; this normalization is used in making the validation plots presented later in the note.

4.3 Event yields

The expected event yields for all processes and the observed number of events are shown in Table 4. The total MC matches the data exactly by construction, because the W +jets yield is the difference between the data yield and the sum of the other backgrounds.

	CEM	CMUP+CMX	CMP	LOOSE
HWW150	7.02 ± 0.55	6.6 ± 0.4	0.62 ± 0.1	0.36 ± 0.05
HWW160	13.85 ± 1.08	12.21 ± 0.8	1.14 ± 0.19	0.36 ± 0.05
HWW170	13.45 ± 1.05	11.46 ± 0.7	1.09 ± 0.18	0.46 ± 0.06
HWW180	10.4 ± 0.81	9.06 ± 0.6	0.87 ± 0.14	0.95 ± 0.12
HWW190	6.99 ± 0.55	6.05 ± 0.4	0.59 ± 0.1	1.24 ± 0.16
HWW200	5.68 ± 0.44	4.92 ± 0.3	0.48 ± 0.08	1.42 ± 0.18
Non-W	341.2 ± 147.7	137.6 ± 66.4	38.5 ± 21.2	10.6 ± 24.1
Diboson	581.8 ± 44.9	508 ± 31.6	48.5 ± 7.2	114.7 ± 13.1
top	47.2 ± 3.4	39.5 ± 2.2	4.1 ± 0.4	15.5 ± 1.2
Z+jets	75.1 ± 9.3	239.6 ± 22.9	32.9 ± 6.2	53.6 ± 7.9
W	5576.3 ± 147.7	4915 ± 66.4	331.3 ± 21.2	731 ± 24.1
Total	6679 ± 213.9	5890 ± 101.7	460 ± 31.5	930 ± 37.4
Data	6679	5890	460	930

Table 4: Expected and observed number of events

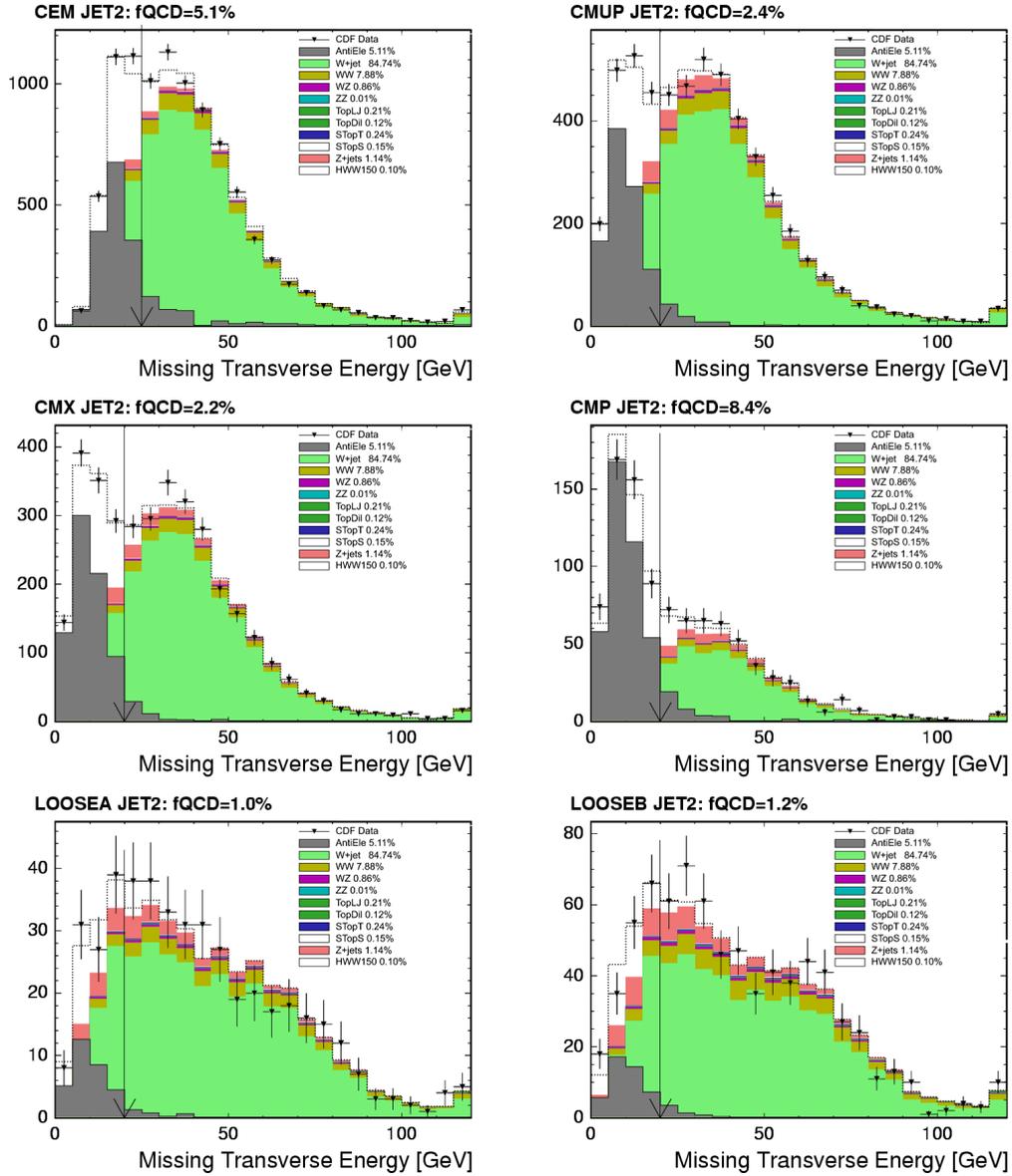


Figure 2: Missing E_T fit to determine non-W normalization in the CEM (top left), CMUP (top right), CMX (second row left), CMUP (second row right), LOOSE A (third row left), and LOOSE B (third row right).

5 New Lepton Categories

Due to the different geometry of the CMP and CMU detectors, some areas in $\eta - \phi$ space are not covered by both detectors. The gaps of the CMU detector can be illustrated by considering muons that have a stub in the CMP but not in the CMU (see Fig. 3). The most evident features are the 2.5° gaps every 15° as well as the central crack which is located where the East and West Calorimeter modules meet. In fact, there are more muons in the 2.5° gaps combined than there are in the central crack[3]. We use the phi-gap (CMP-only) triggers, which are based on tracks pointing to the gaps in the CMU coverage with a “stub” or hit in only the CMP detector and introduce a new category of muons for them, which can be characterized as follows:

CMP category (comes from gap triggers):

- Data: must have fired gap trigger, and must satisfy $\tilde{\phi} < 2$ or $\tilde{\phi} > 13$ where $\tilde{\phi} = \phi \bmod 15^\circ$
- MC: must be a hit in CMP (no trigger turn on is applied), and must satisfy $\tilde{\phi} < 2$ or $\tilde{\phi} > 13$

The cut on $\tilde{\phi}$ just requires that the tracks are within 2% of the gaps in coverage. The addition of this muon category required a modification in the extended muon category to avoid double counting CMP events. We have therefore split up the previous extended muon category (see for instance [16]; we call it the “loose” category, though events have to pass strict tight selection criteria) into two separate categories, henceforth referred to as “loose A” and “loose B”; the main difference between the two is that the former contains loose events in accordance with the usual definition but only from before period 10 whereas any event from period ten or later is part of the “loose B” category, with the exception of CMP events (which have their own category as previously sketched out). More precisely, we have:

LOOSE A:

- Data: must pass MET Jet L2 requirement and be from before period 10
- MC: must pass MET Jet L2 requirement and apply trigger turn on curve

LOOSE B:

- Data: must pass MET Jet L2 requirement and be from period 10 or later
- MC: must pass MET Jet L2 requirement and apply trigger turn on curve
- CMP events are excluded explicitly

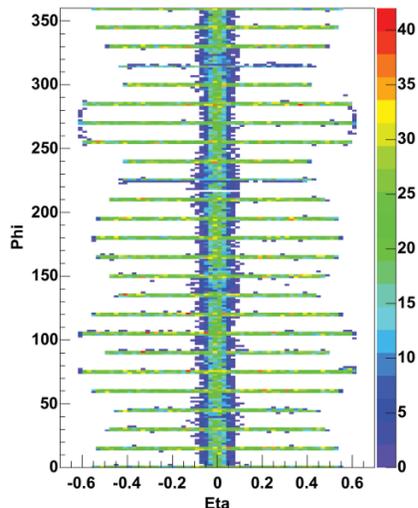


Figure 3: Illustration of the gaps of the CMU detector. In this $\phi - \eta$ scatter plot, only muons are placed that are registered exclusively by the CMP, illustrating the gaps of the CMU detector which lies inside. The 2.5° gaps every 15° are clearly visible as well as the central crack where the East and West Calorimeter modules meet.

selection	CMP (!firedMET)	MET (!firedCMP)	CMUP
MET & QCD	1867 (1191)	1103 (427)	10060
pass L2 jets		570 (214)	-

Table 5: Comparison of the number of events from the MET+jets trigger and the phi gap trigger in periods 10-23. In this run range, 1867 events will be added (corresponding to about 19% of CMUP), whereas we currently have some 570 events from the MET+jets trigger (corresponding to about 6% of CMUP). We thus gain about 1200 events. When cutting on $\tilde{\phi}$, an approximate 6% of the additional muons gets lost.

To estimate the gain achieved by the addition of the CMP data, one can compare the number of events from the MET+jets trigger (for more information about MET plus jet triggers, see [16]) and those from the new trigger; an overview of these numbers is given in Table 5. The CMU gap trigger can be added in the same way; as the data set grows, so will the gain from using these triggers (as they will be included in a larger fraction of the data set).

One can go one step further to compare the potential gain in Higgs events up to period 23, assuming $m_H = 160$ GeV and without cutting on the dijet mass or

vetoing loose jets: While the previous loose category contained an anticipated 2.03 events, the new loose categories contain 1.46 anticipated events, and the CMP contains 2.10 , thus leading to a total of 3.56 anticipated events. The gain of 1.53 events compared to the CMX with 7.91 events corresponds to about 19 % of anticipated CMX Higgs events or approx. 10 % of anticipated CMUP Higgs events. When comparing the gain to the total number of anticipated Higgs events, the improvement is 2.6 %.

6 Measurement technique

A matrix element technique is used to separate signal from background. An event probability is calculated for signal and several background processes using differential cross sections. The details of the methodology are described in CDF note 9404 [8]. The implementation of the matrix element calculation here is identical: the same diagrams, transfer functions, phase space calculations, definition of the event probability density, and numerical integration techniques were used. The Higgs production diagrams are shown in Figure 4.

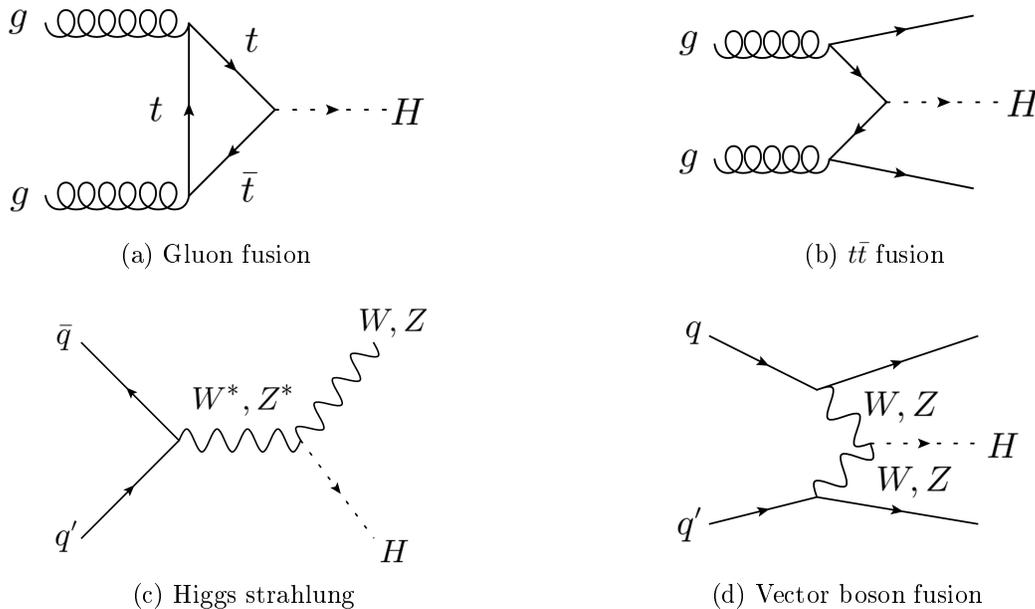


Figure 4: Leading order Feynman diagrams for Higgs production

The event probability densities were calculated for the following processes: WW, WZ, s channel single top (s-chan), t channel single top (t-chan), W plus two gluons (Wgg), W plus one gluon and a quark (Wgj), W plus two c quarks (Wcc), W plus

two b quarks (Wbb), and W plus one c quark and a jet (Wcj). The event probabilities are used to build an event probability discriminant (EPD) which is calculated for each mass point. In general, the EPD has the form $EPD = P_s/(P_s + P_b)$. In our EPD, we have

$$P_b = P_{WW} + P_{WZ} + P_{single\ top} + P_{Wgg} + P_{Wgj} + P_{Wbb} + P_{Wcc} + P_{Wcj} \quad (1)$$

$$P_s = P_{HWW} \quad (2)$$

6.1 Templates

Template histograms of the EPD shape are made for signal and background processes; the data is later fit to a sum of these templates. Although we have six lepton categories, we use only two channels. The first channel consists of central electrons (CEM), CMP muons (CMP), central muons (CMUP), and central muon extension muons (CMX). The second channel contains the two loose categories (loose A and loose B). The templates are chosen to have 40 bins.

The likelihood fit described in section 7 takes into account the bin-by-bin uncertainty due to the statistics in the templates.

The expected shapes of the templates for signal and background processes are shown in Figure 5 added over all lepton categories. Stack histograms with the expected normalizations of the background processes are shown in Figure 6.

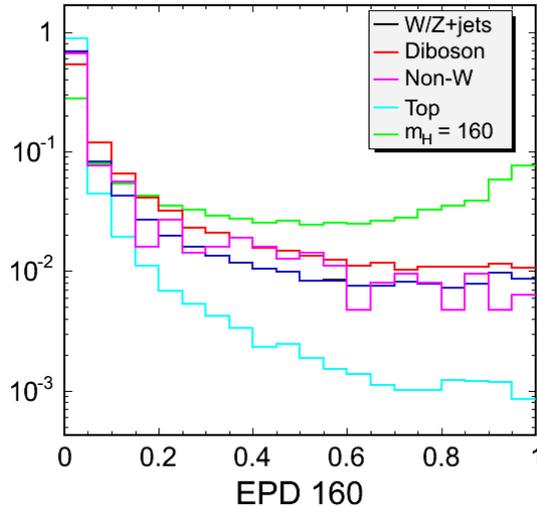


Figure 5: Event probability discriminant (EPD) shape for signal and background processes under the assumption that $m_H = 160$.

6.2 Validation of input variables

The templates for the discriminants used in the fits are extracted from the Monte Carlo described in section 4. We therefore need to validate carefully the modeling of the data, particularly for variables entering the matrix element calculation.

The inputs to the matrix element calculations are the 4-momenta of the lepton and jets, so we validate the E_T or p_T and angular distributions of the lepton and jets. In addition we check the kinematics of the reconstructed W bosons. The inputs to the dijet mass are the jet E_T s and relative angles, so we plot the ΔR between the two jets. We also check the \cancel{E}_T and transverse mass of the leptonic W ($m_T(W)$) to validate our non-W background estimate. Finally, we plot the discriminating variable in regions of dijet mass where little signal is expected.

The plots are made separately for each lepton category as well as for the sum over all categories. The sub-canvases of the plots are: CEM electrons (top left), PHX electrons (top right), CMUP muons (second row left), CMX muons (second row right), extended muons before period 10 (third row left), extended muons after period 10 excluding CMP (third row right), CMP muons (bottom left) and sum of all categories (bottom right). Note that the PHX is plotted but it is not used in this analysis and not included in the plot of the sum of categories.

The KS test is performed using 1000 pseudo-experiments, as described on the top website [13].

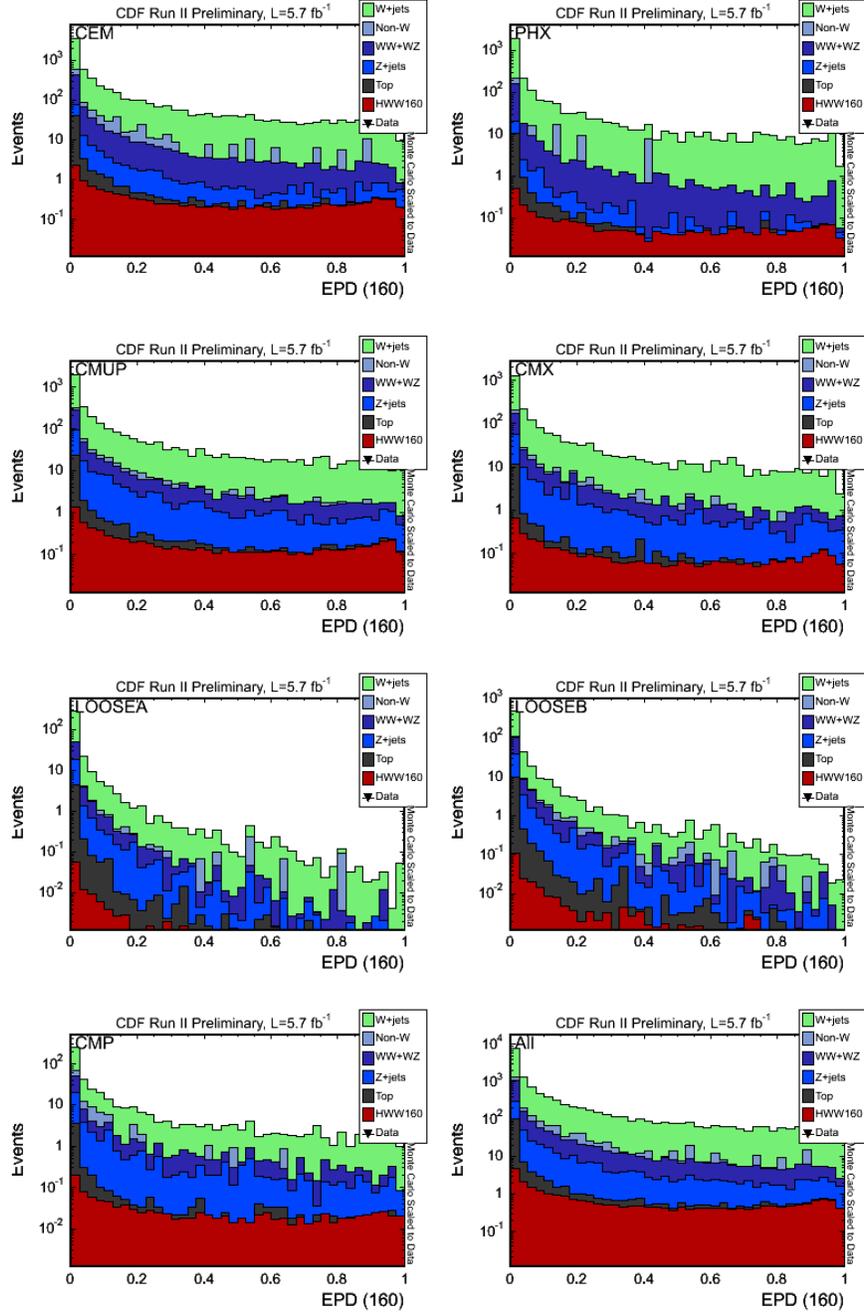
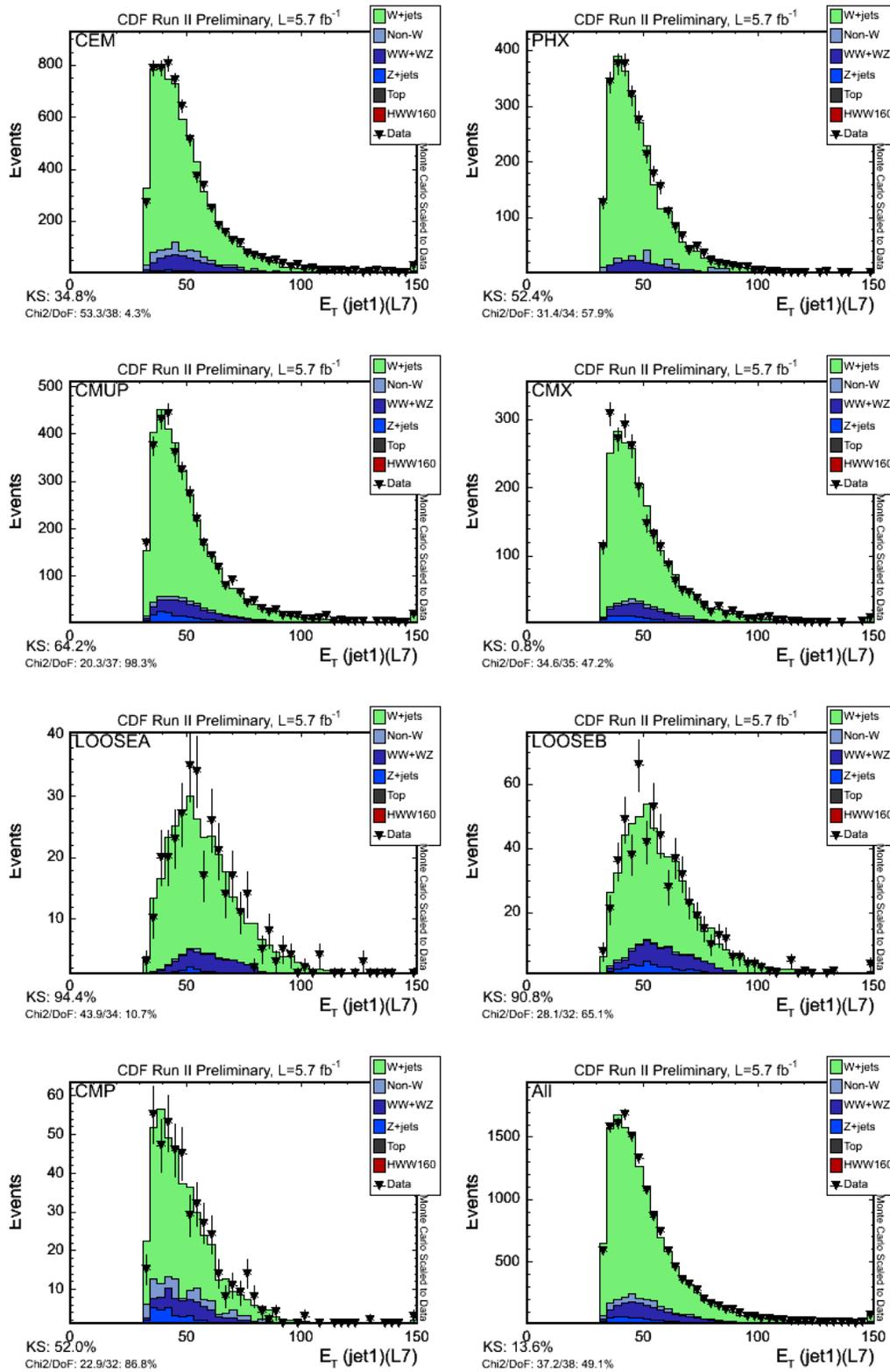
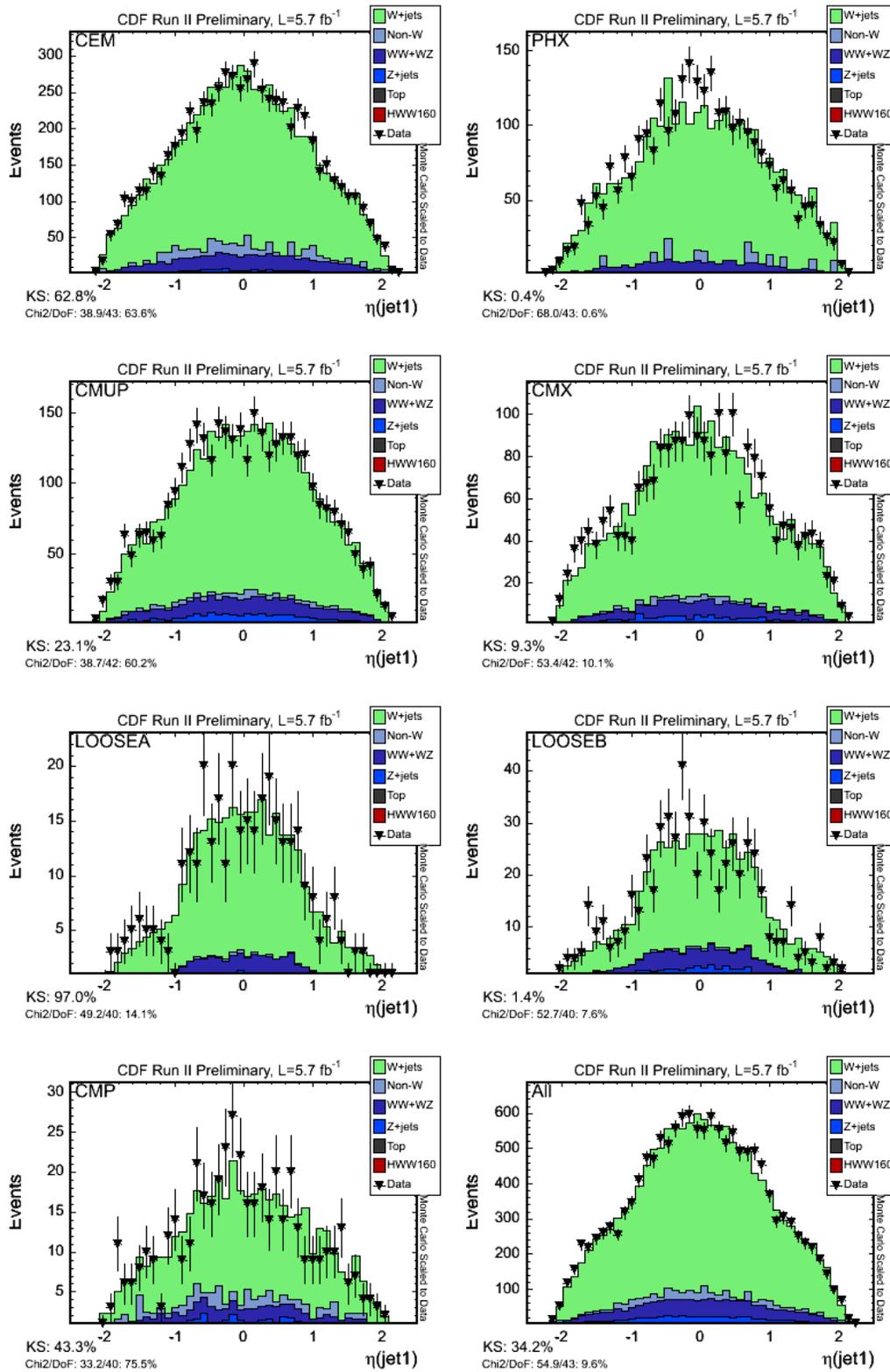
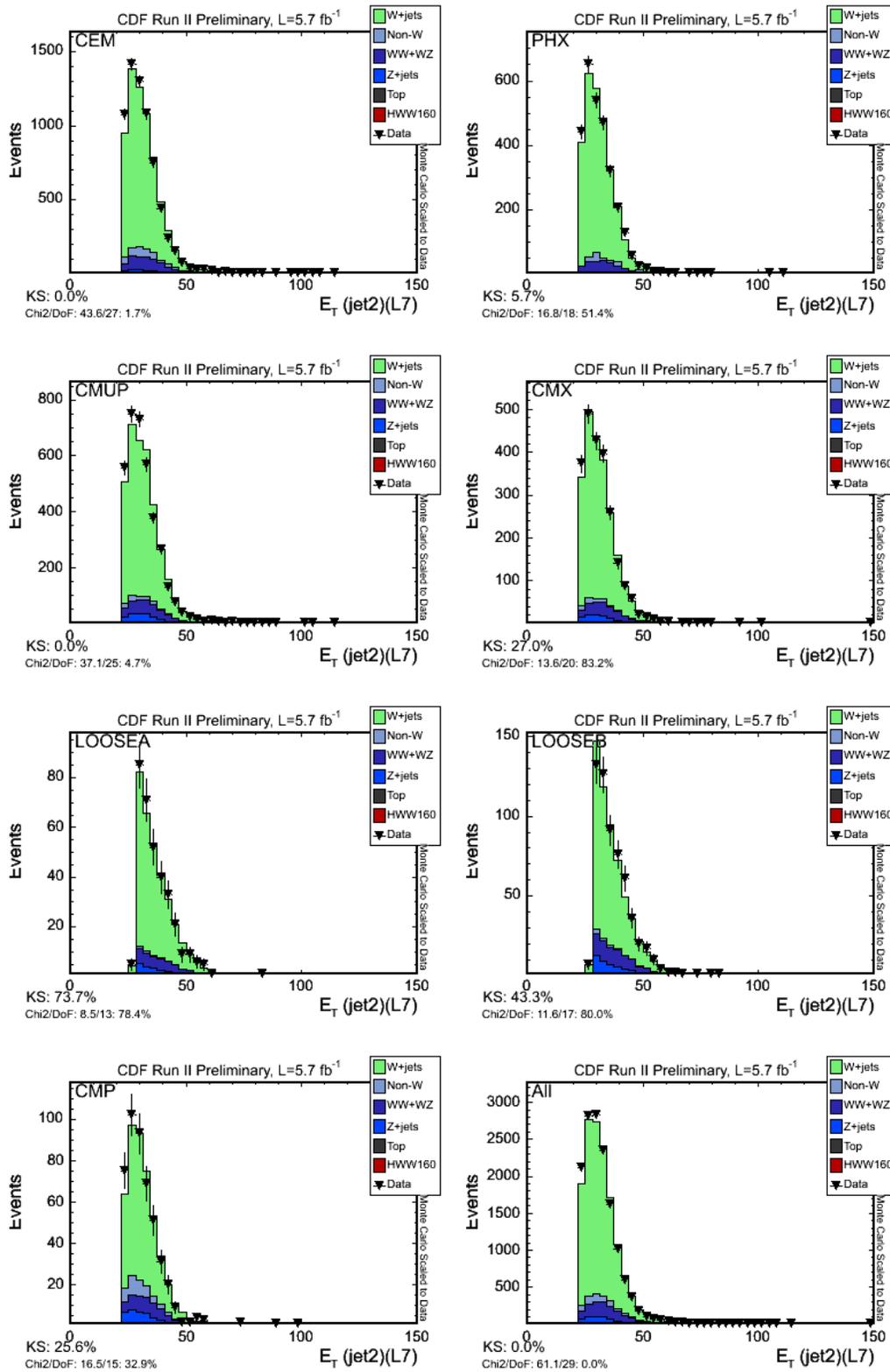
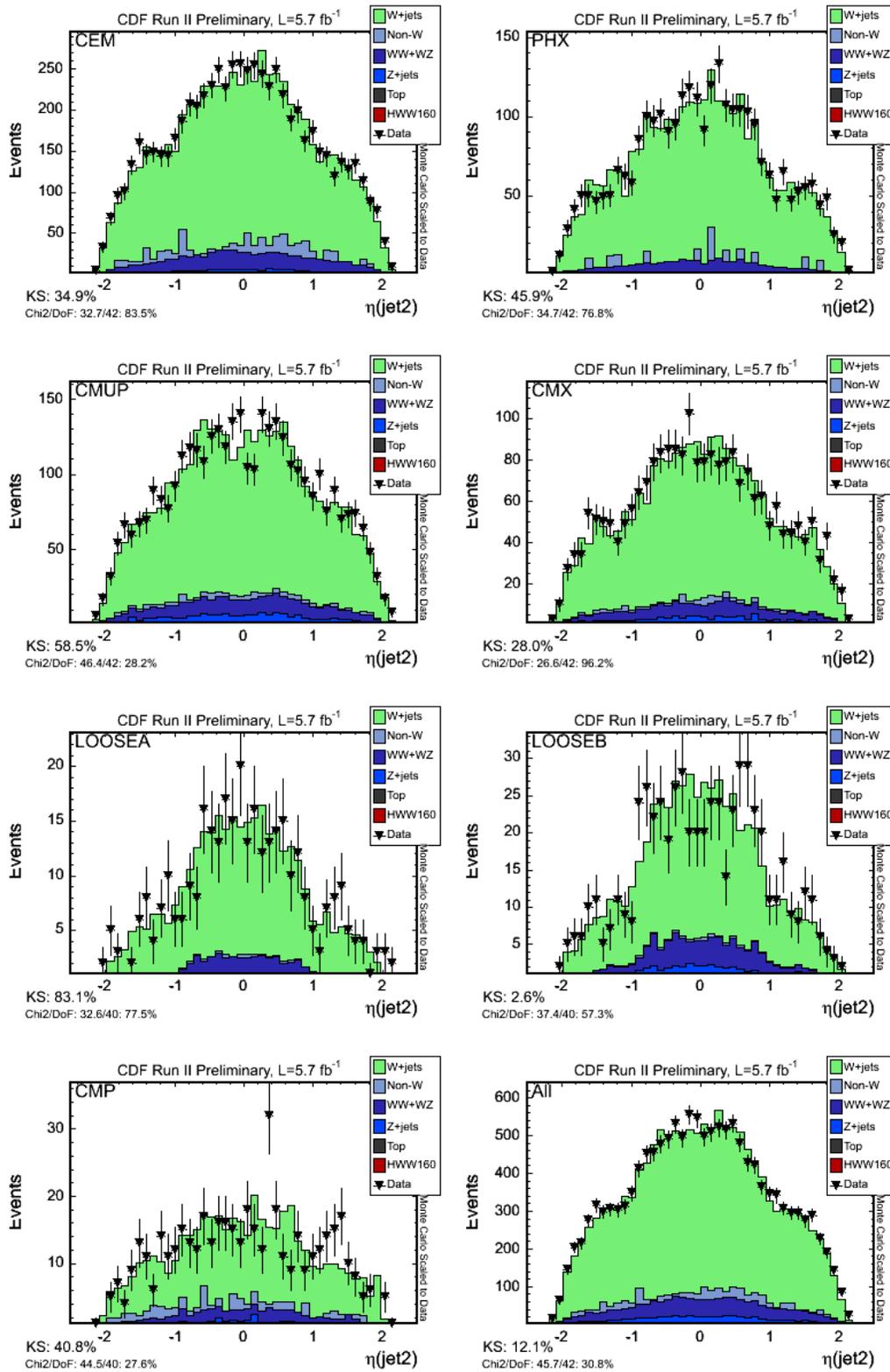


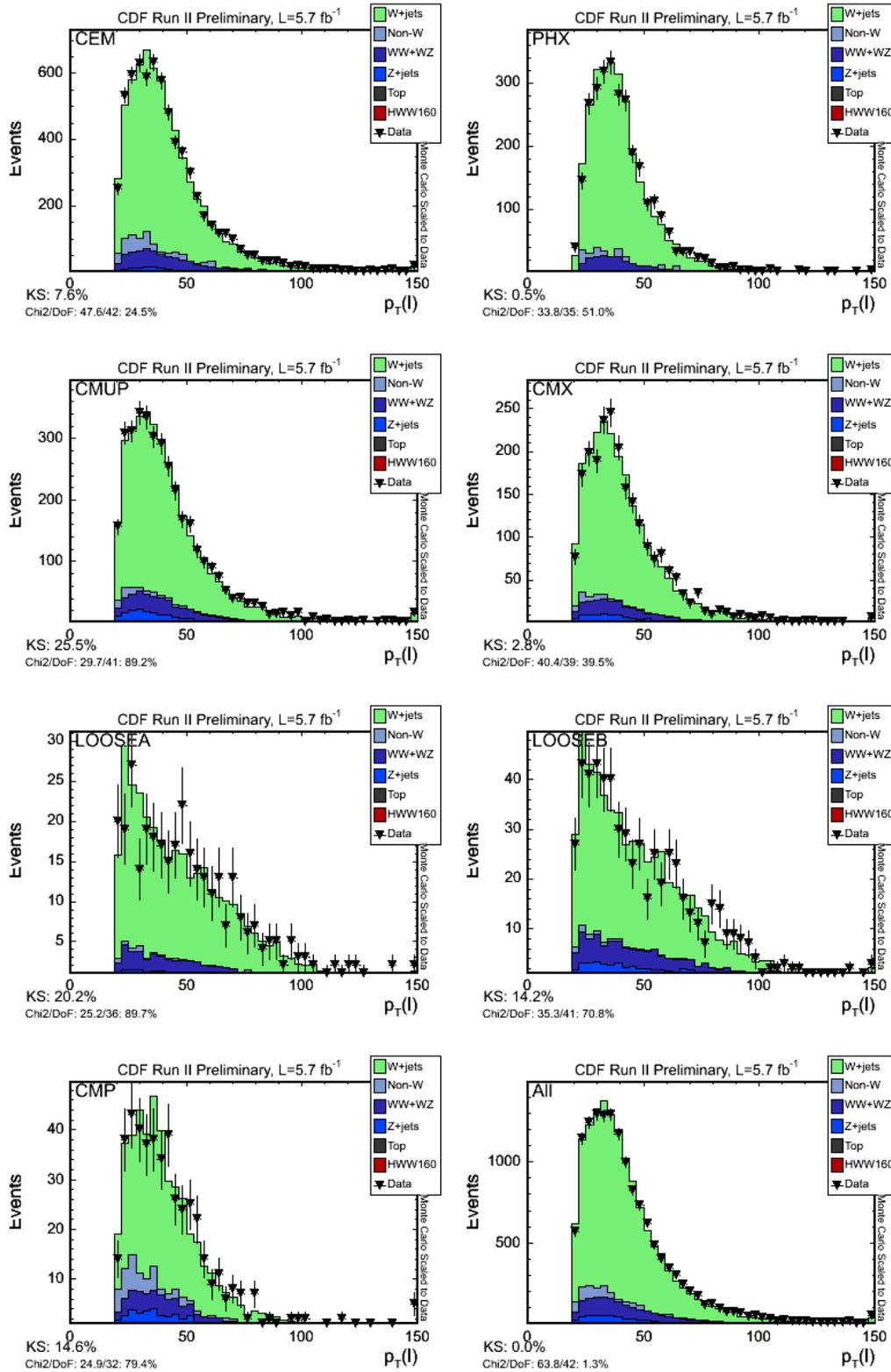
Figure 6: Event probability discriminant (EPD) distribution for the lepton categories used in the analysis and PHX: CEM (top row left), PHX (top row right - not used in analysis), CMUP (second row left), CMX (second row right), LOOSE A and LOOSE B (third row), CMP (bottom row left) and sum of all categories (bottom row right). Note that the scale is adapted for the LOOSE categories and for CMP.

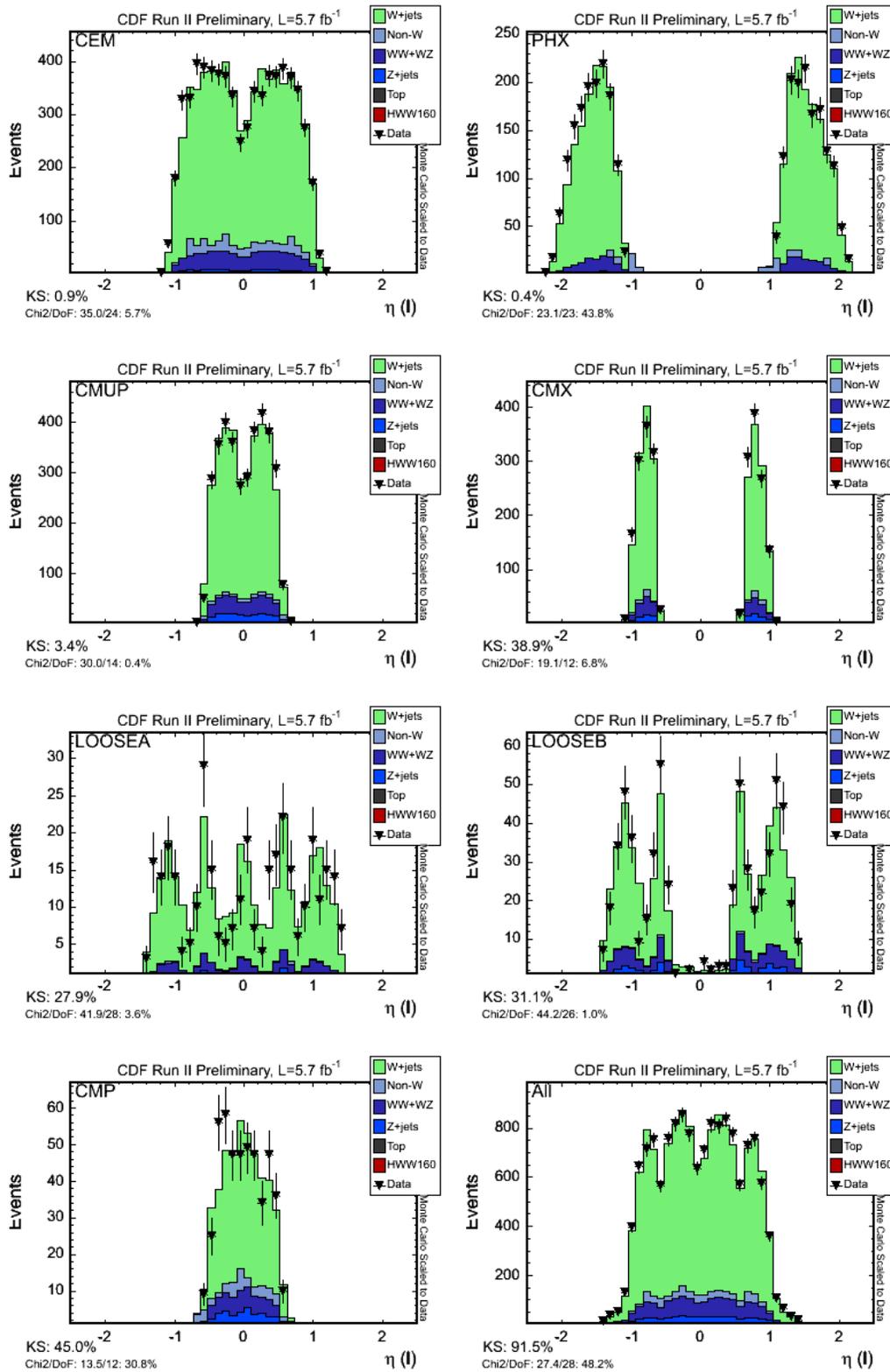
Figure 7: E_T of the higher- E_T jet

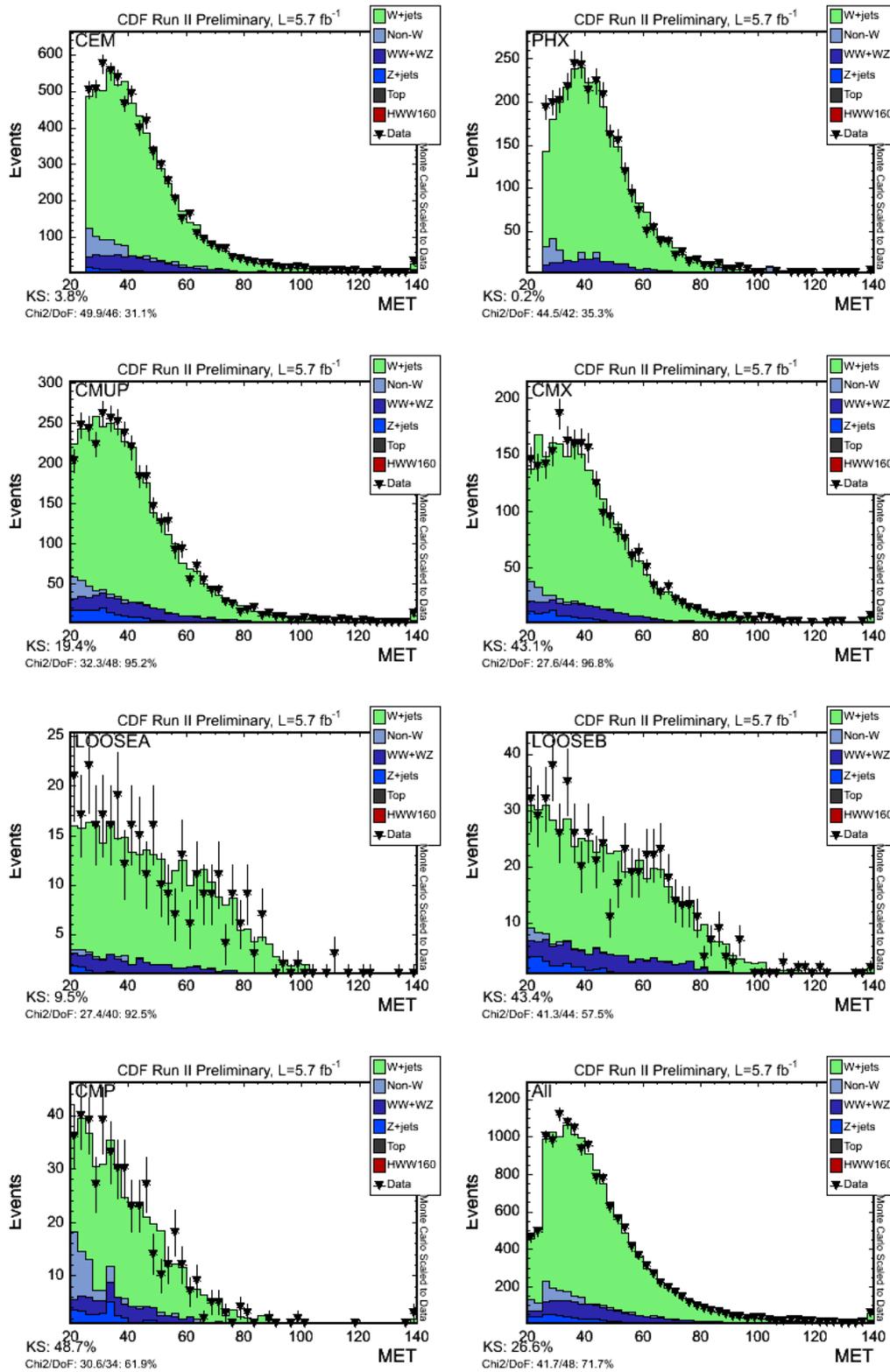
Figure 8: η of the higher- E_T jet

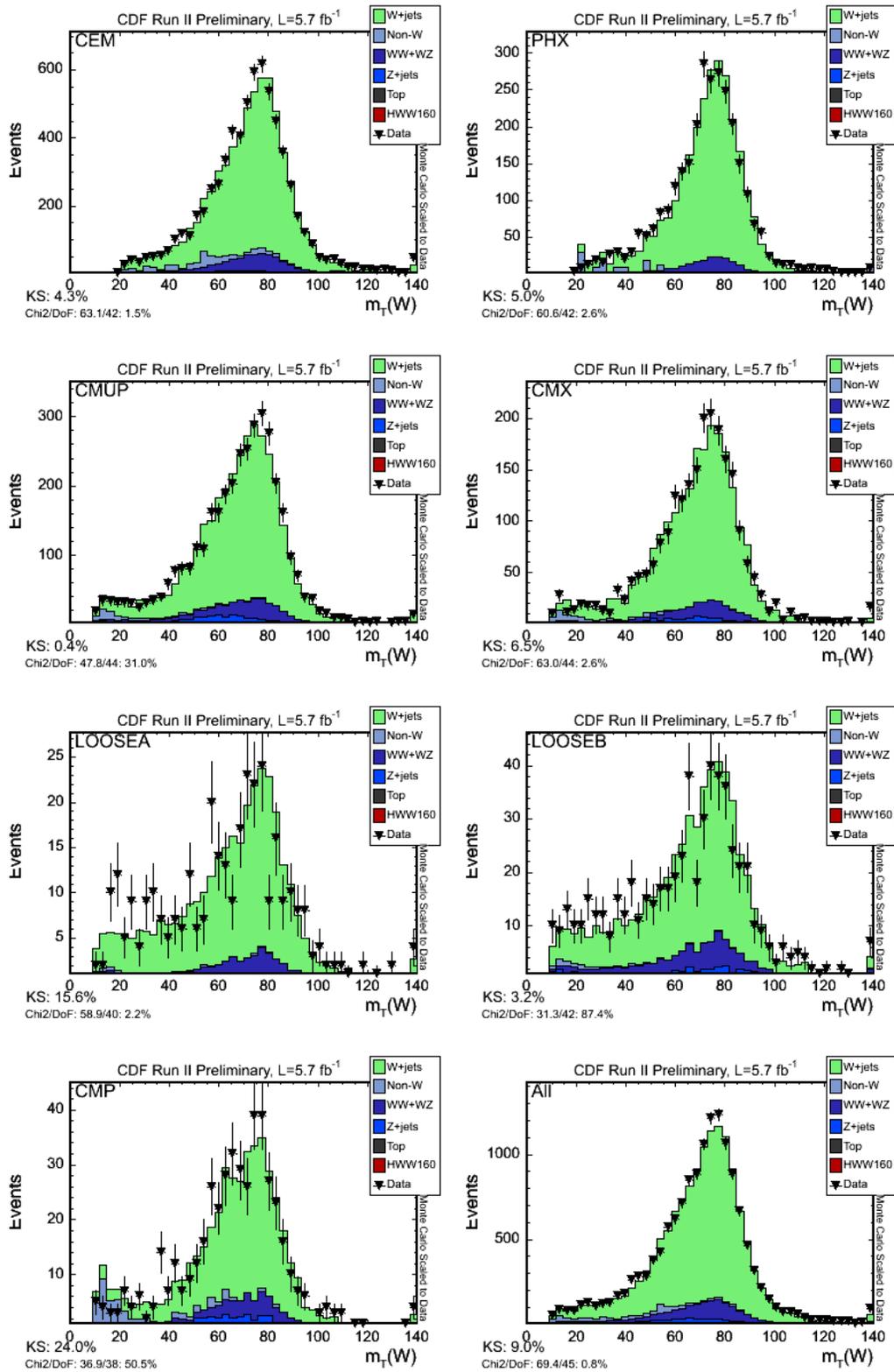
Figure 9: E_T of the lower- E_T jet

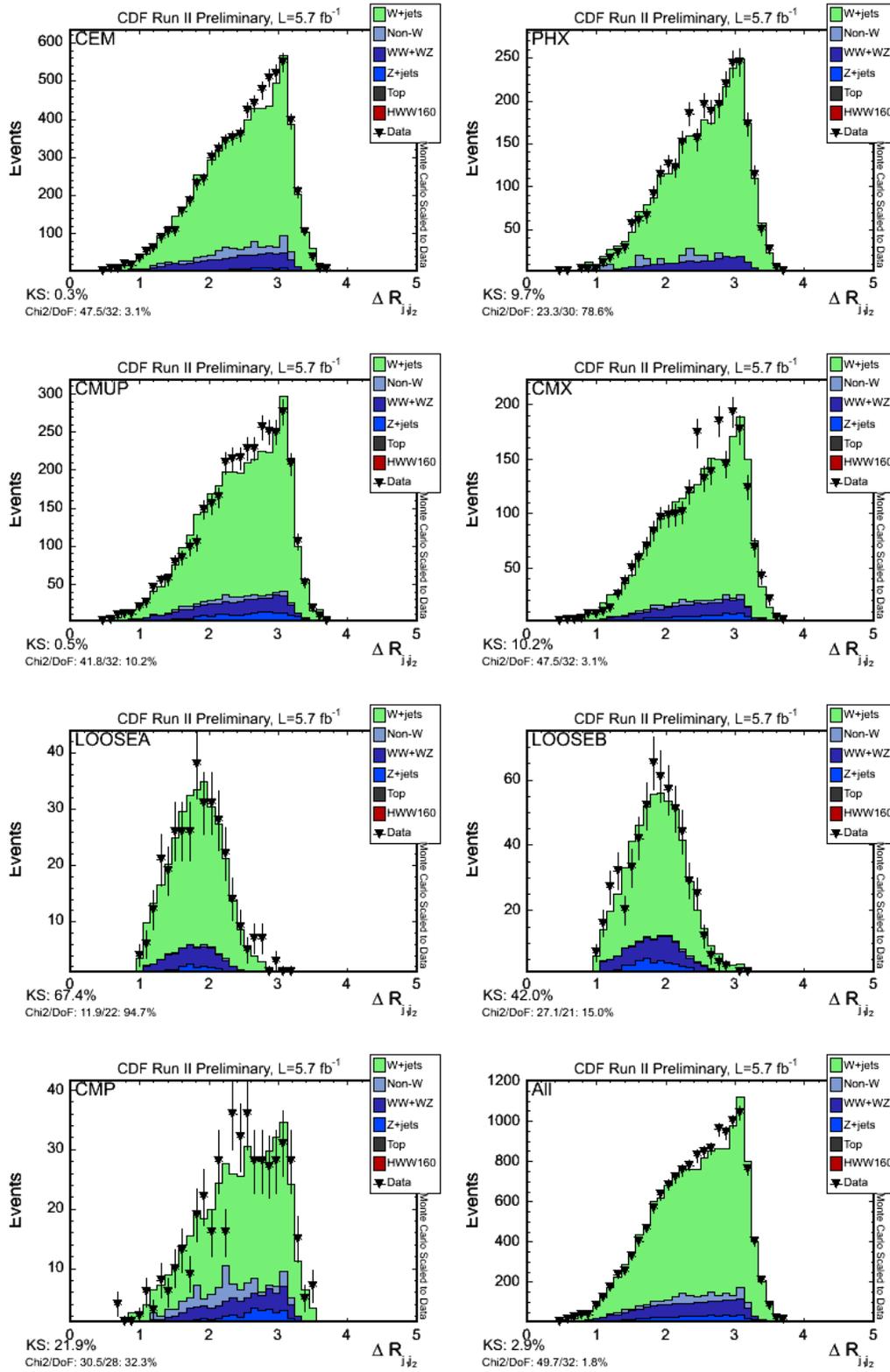
Figure 10: η of the lower- E_T jet

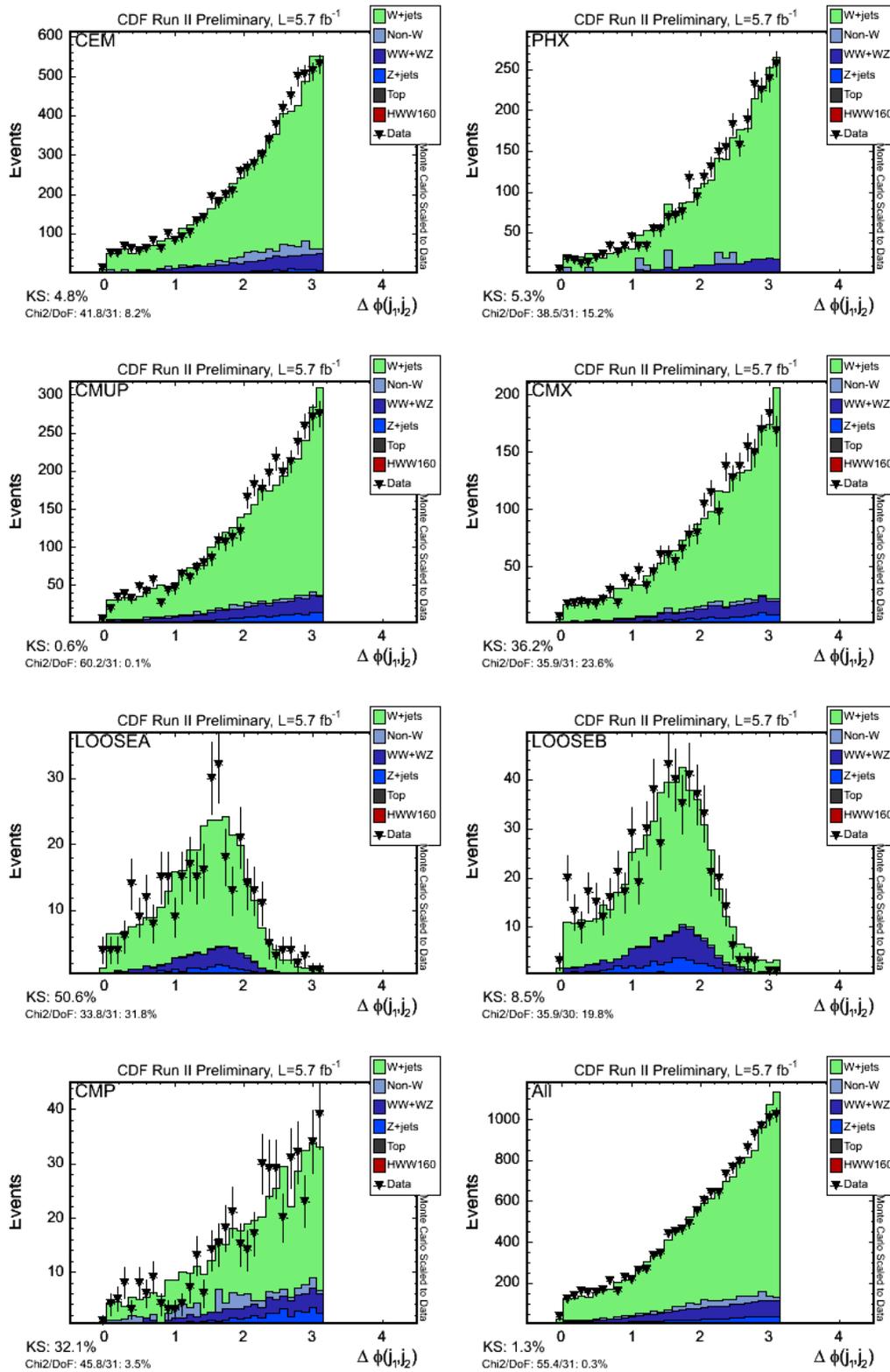
Figure 11: p_T of the lepton.

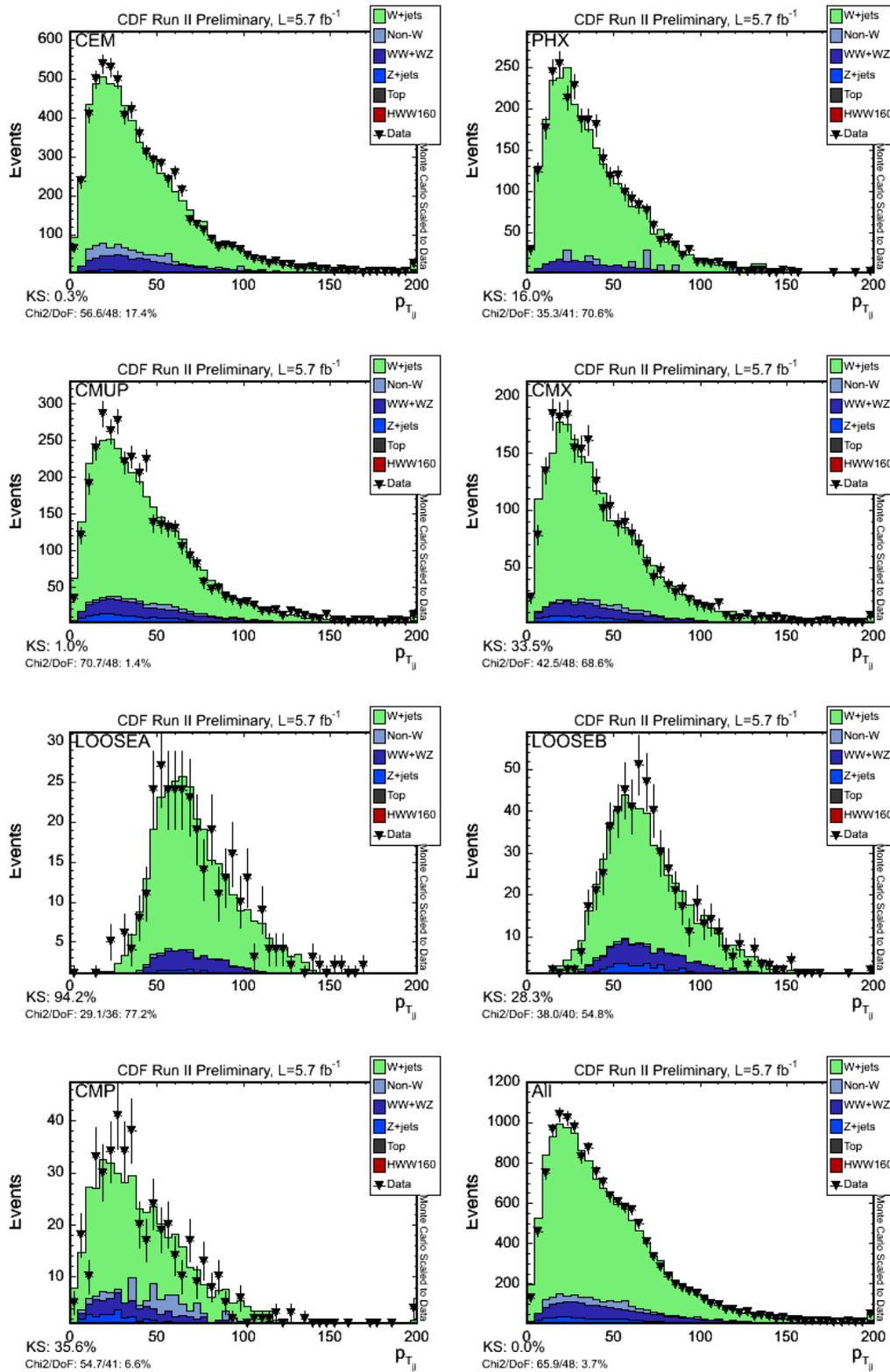
Figure 12: η of the lepton.

Figure 13: \cancel{E}_T

Figure 14: Transverse mass of leptonic W ($\cancel{E}_T + \text{lepton system}$).

Figure 15: ΔR between the two jets.

Figure 16: $\Delta\phi$ between the two jets.

Figure 17: p_T of the dijet system.

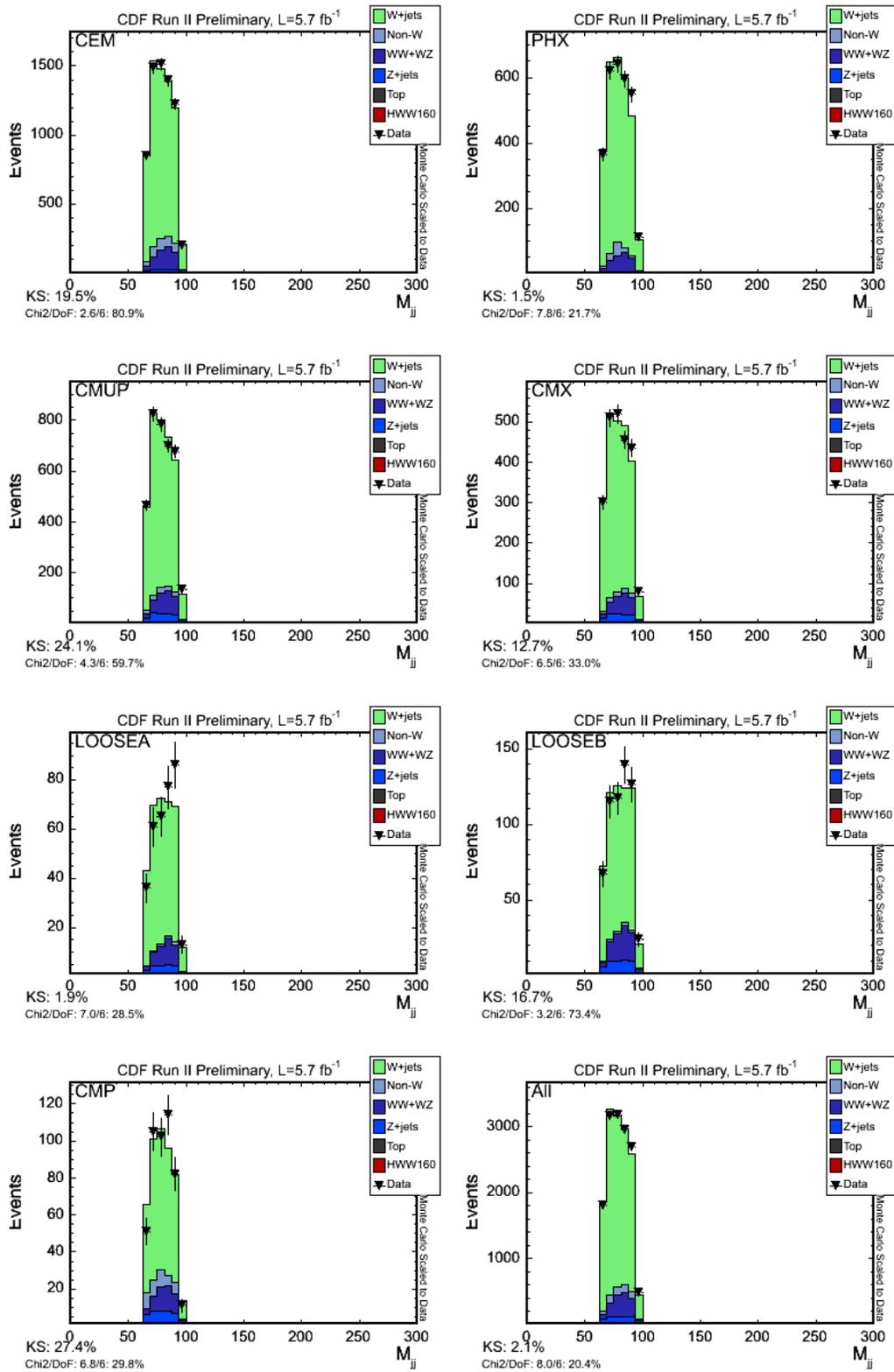


Figure 18: Invariant mass of the dijet system.

7 The Likelihood Function

The likelihood function, \mathcal{L} , is a function of the unknown Poisson means for signal and background and is defined such that it expresses the joint probability of observing the N data events at their respective values of the discriminating variable. The values of the Poisson means at which \mathcal{L} achieves its maximum corresponds to the most probable estimate for the true signal and background content in the data sample.

We perform a binned likelihood fit to the discriminant (M_{jj} or the EPD). To make it easier to compare the different fit parameters, we define the fit parameter as $\beta_j = \sigma_j^{Fit} / \sigma_j^{SM}$ where β_j is unity when the fit result corresponds to the expected number of events obtained from the background normalization estimate. The likelihood is given by:

$$\mathcal{L} = \prod_{j=2}^{N_B+1} G_j(\beta_j; \sigma_j) \prod_{k=1}^B \frac{e^{-\mu_k} \cdot \mu_k^{n_k}}{n_k!}, \quad (3)$$

Where $G_j(\beta_j; \sigma_j)$ is a Gaussian constraint on the background normalization and N_B is the number of background processes.

The Gaussian constraints on the backgrounds are given by:

$$G_j(\beta_j; \sigma_j) = \frac{1}{\sqrt{2\pi \cdot \sigma_j^2}} \exp \left[-\frac{1}{2} \cdot \left(\frac{\beta_j - 1.0}{\sigma_j} \right)^2 \right]. \quad (4)$$

The number of events in a bin is the sum of signal events and background events, so

$$\mu_k = \beta_{HWW} \cdot T_{(HWW)k} + \sum_{j=2}^{N_B+1} \beta_j \cdot T_{jk}, \quad (5)$$

where T_{jk} is the content in bin k of the template for process j . The templates are normalized to the predicted number of events given in table.

The prediction in each bin needs an additional Gaussian uncertainty due to the limitations of Monte Carlo statistics. Each bin is allowed to fluctuate according to the total uncertainty in that bin, which is the sum in quadrature of the weight of each event. This prevents us from overestimating our sensitivity due to a fluctuation in Monte Carlo.

8 Systematic Uncertainties

8.1 Sources of systematic uncertainty

Systematic uncertainties can bias the outcome of this analysis and have to be incorporated into the result. We address systematic uncertainty from several different sources: (1) background normalization¹, (2) jet energy scale, (3) parton distribution functions, (4) ISR and FSR, (5) factorization and renormalization (Q^2) scale in Alpgen, and (6) event selection efficiency and luminosity. All of these uncertainties are described in more detail below.

Systematic uncertainties can influence both the expected event yield (normalization or rate) and the shape of the discriminant distribution for signal or background processes. A summary of all systematic uncertainties is shown in Table 6.

8.1.1 Background normalization

The uncertainties on the background normalizations are taken from the estimates described in Section 4. In the case of Z +jets, $t\bar{t}$, and single top backgrounds they reflect the uncertainty on the cross sections for these processes; they are summarized in Table 6. The normalizations of the single top and $t\bar{t}$ backgrounds are correlated in the fit.

As mentioned before, a 40% uncertainty is assigned to the non- W background normalization. A 20% uncertainty is applied to the W +jets normalization, taken to be significantly more conservative than the 12% uncertainty from the method2 procedure. The likelihood fit ultimately constrains the W +jets normalization to better than 2%, meaning that the 20% uncertainty used is not really a constraint at all and the W +jets normalization is a free parameter in the fit.

8.1.2 Signal normalization

An uncertainty of 10% is used for the signal normalization.

8.1.3 Jet energy scale

The jet energy scale uncertainty is assessed by scaling each jet energy by $\pm 1\sigma$ and re-running the analysis. This is done for the HWW samples.

The signal acceptance changes by +5%/-7% (for $m_H=160$ GeV) as a result of the jet energy scale uncertainty. The change is a function of the Higgs mass, but we use the largest number for all Higgs masses. The matrix elements and EPD

¹The background normalization uncertainties are actually considered part of the statistical uncertainty, but are described in detail here.

are re-calculated, and shape uncertainties are assigned both for the signal process and for the W +jets background. The change in the EPD templates (summed over all lepton types) is shown in Figure 19. These shape uncertainties are smoothed internally in the likelihood fit.

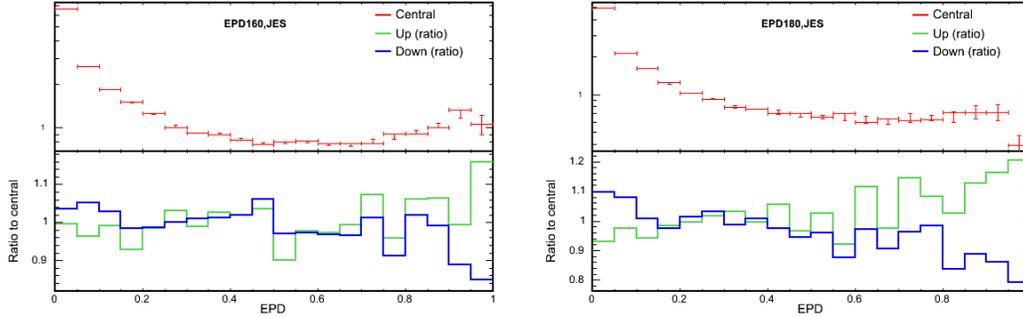


Figure 19: Change in the EPD template for HWW with a mass of $m_H = 160$ GeV/c^2 (left) and $m_H = 180$ GeV/c^2 (right) associated with the JES uncertainty.

8.1.4 PDFs

We have not explicitly calculated the PDF uncertainty for this signal. We apply a 2% rate uncertainty on the signal, which is taken from CDF note 9923 [18].

8.1.5 ISR and FSR

We apply a 6% acceptance uncertainty based on the change observed in the WW acceptance due to ISR/FSR. We haven't explicitly evaluated the uncertainty for HWW .

8.1.6 Q^2 scale

Alpgen W +jets samples were generated with the Q^2 scale at half and double its default values. The resulting change in the modeling is shown in Figure 20.

The change in the W +jets template associated with the Q^2 uncertainty is shown in Figure 21. The shape changes significantly. This uncertainty is currently not truncated at 1σ due to limitations in the fitting framework, but we hope to truncate it in the future.

8.1.7 Luminosity and lepton ID efficiency

A luminosity uncertainty of 6% is applied to all processes whose normalizations are determined by theory and the Monte Carlo ($HWW, WW, WZ, ZZ, \text{single top}, t\bar{t}$,

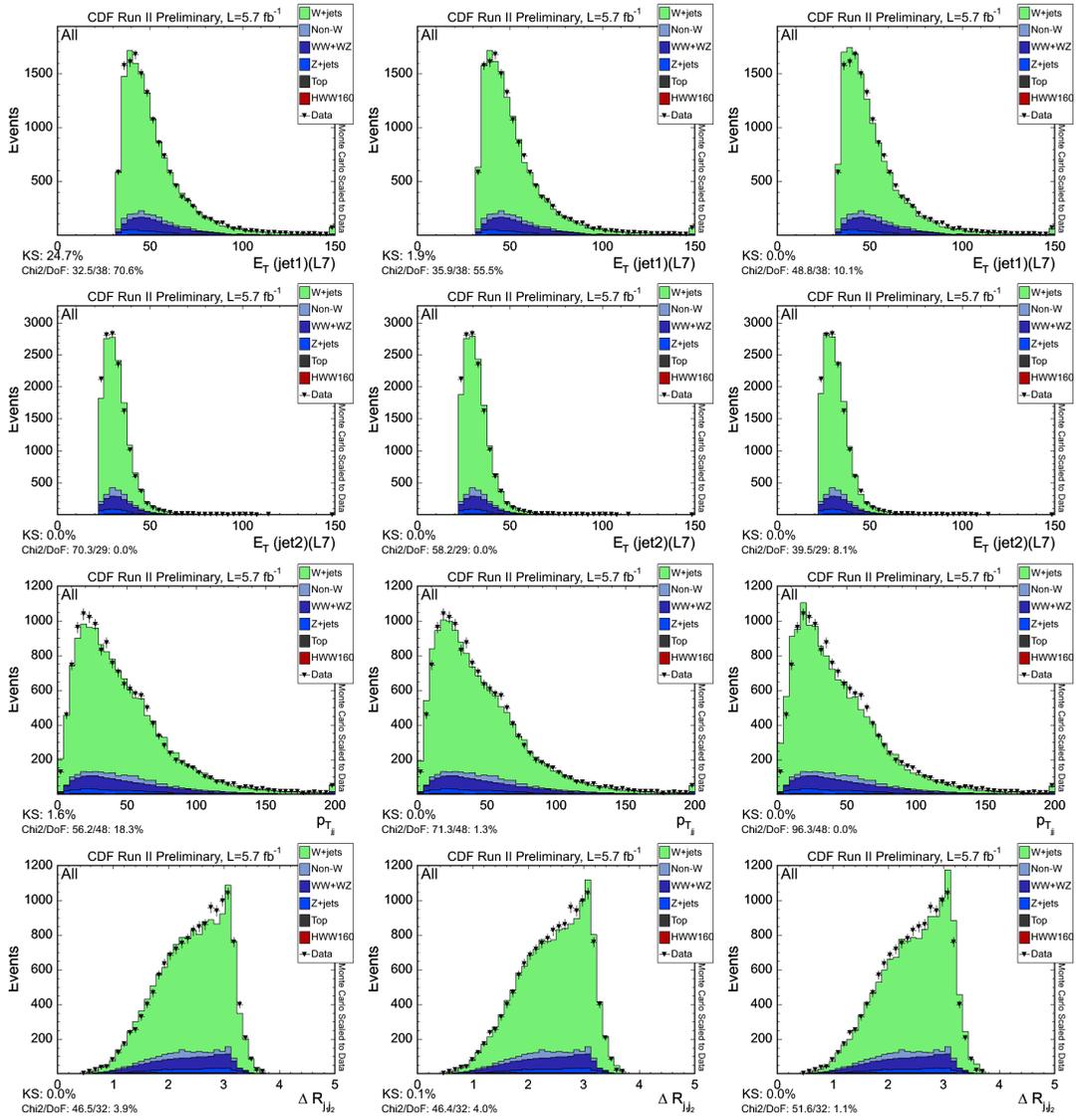


Figure 20: Modeling of various variables with the Q^2 scale decreased by a factor of $\frac{1}{2}$ (left), the central choice of Q^2 scale (center), and the Q^2 scale increased by a factor of 2 (right). The top row shows the E_T of the harder jet, the second the E_T of the second jet, the third the ΔR between the two jets, and the fourth the p_T of the dijet system.

and Z +jets). An additional 2% uncertainty is applied due to the trigger selection and offline lepton identification efficiencies.

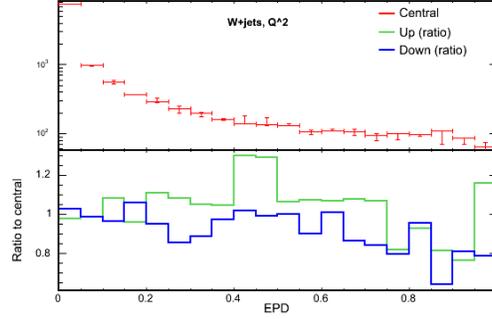


Figure 21: Change in the W +jets template associated with the Q^2 uncertainty for an assumed mass of $m_H = 160 \text{ GeV}/c^2$.

Process	Source	Rate -1σ	Rate $+1\sigma$	Shape?
HWW	Normalization	-10%	10%	no
WW, WZ, ZZ	Normalization	-10%	10%	no
W +jets	Normalization	-20%	20%	no
Z +jets	Normalization	-15%	15%	no
$t\bar{t}$ and single top	Normalization	-12%	12%	no
non- W	Normalization	-40%	40%	no
HWW	JES	-7%	5%	yes
HWW	PDF	-2%	2%	no
HWW	IFSR	-6%	6%	no
W +jets	Q^2	0%	0%	yes
All MC	Event selection efficiency	-2%	2%	no
All MC	Luminosity	-6%	6%	no

Table 6: Summary of systematic uncertainties. All MC refers to all processes except non- W and W +jets, whose normalizations are determined by fit to the data.

8.2 Integration of systematic uncertainties in likelihood fit

All systematic normalization and shape uncertainties are incorporated into the likelihood as nuisance parameters, conforming with a fully Bayesian treatment [9]. The full likelihood function is given by:

$$\mathcal{L}(\beta_1, \dots, \beta_{N_b+1}; \delta_1, \dots, \delta_{N_{sys}}) = \underbrace{\prod_{k=1}^B \frac{e^{-\mu_k} \cdot \mu_k^{n_k}}{n_k!}}_{\text{Poisson term}} \cdot \underbrace{\prod_{j=2}^{N_b+1} G(\beta_j | 1, \Delta_j)}_{\text{Gauss constraints}} \cdot \underbrace{\prod_{i=1}^{N_{sys}} G(\delta_i, 0, 1)}_{\text{Systematics}} \quad (6)$$

where

$$\mu_k = \sum_{j=1}^{N_b+1} \beta_j \cdot \underbrace{\left\{ \prod_{i=1}^{N_{sys}} [1 + |\delta_i| \cdot (\epsilon_{ji+} H(\delta_i) + \epsilon_{ji-} H(-\delta_i))] \right\}}_{\text{Normalization Uncertainty}} \quad (7)$$

$$\cdot \underbrace{T_{jk}}_{\text{Shape P.}} \cdot \underbrace{\left\{ \prod_{i=1}^{N_{sys}} (1 + |\delta_i| \cdot (\kappa_{jik+} H(\delta_i) + \kappa_{jik-} H(-\delta_i))) \right\}}_{\text{Shape Uncertainty}} \quad (8)$$

and N_{sys} is the number of systematic effects included.

We take the correlation between normalization and shape uncertainties for a given source into account [10]. The relative strength of a systematic effect due to the source i is parameterized by the nuisance parameter δ_i in the likelihood function, constrained to a unit-width Gaussian (last term in Equation 6). The $\pm 1\sigma$ changes in the normalization of process j due to the i^{th} source of systematic uncertainty are denoted by ϵ_{ji+} and ϵ_{ji-} (see Equation part 7). The $\pm 1\sigma$ changes in bin k of the templates for process j due to the i^{th} source of systematic uncertainty are quantified by κ_{jik+} and κ_{jik-} (see Equation part 8). $H(\delta_i)$ represents the Heaviside function, defined as $H(\delta_i) = 1$ for $\delta_i > 0$ and $H(\delta_i) = 0$ for $\delta_i < 0$. The Heaviside function is used to separate positive and negative systematic shifts (for which we have different normalization and shape uncertainties). The variable δ_i appears in both the term for the normalization (Equation 7) and the shape uncertainty (Equation 8), which is how correlations between both effects are taken into account.

We marginalize the likelihood function by integrating $\mathcal{L}(\beta_1, \dots, \beta_N, \delta_1, \dots, \delta_S)$ over all nuisance parameters. The resulting reduced likelihood $\mathcal{L}(\beta_H)$ is a function of the H cross-section β_H only. We use the MCLIMIT package for our statistical treatment [11].

σ/SM	150	160	170	180	190	200
Expected	30.025	12.89	19.45	29.72	42.51	39.91
Observed	15.88	12.62	26.55	15.97	43.80	75.98

Table 7: Expected upper limit cross section, in SM units, for different Higgs mass points. The limits are not final since we are still updating with high luminosity Monte Carlo.

9 Result with CDF II Data

We apply the analysis to 5.7 fb^{-1} of CDF Run II data. We compare the EPD output distribution for Higgs masses of 150, 160, 170, 180, 190 and 200 GeV of our candidate events with the sum of predicted $H \rightarrow WW$ signal and background distributions as shown in Fig. 22.

In order to extract the most probable $H \rightarrow WW$ signal content in the data, we perform the maximum likelihood method described previously. We perform marginalization using (6) with all systematic uncertainties included in the likelihood function. The posterior p.d.f. is obtained using Bayes' theorem:

$$p(\beta_1|data) = \frac{\mathcal{L}^*(data|\beta_1)\pi(\beta_{HWW})}{\int \mathcal{L}^*(data|\beta'_{HWW})\pi(\beta'_{HWW})d\beta'_{HWW}} \quad (9)$$

where $\mathcal{L}^*(data|\beta_{HWW})$ is the reduced likelihood and $\pi(\beta_{HWW})$ is the prior p.d.f. for β_{HWW} . We adopt a flat prior, $\pi(\beta_{HWW}) = H(\eta_{HWW})$, in this analysis, with H being the Heaviside step function.

To set an upper limit on the H production cross-section, we integrate the posterior probability to cover 95% [17]. The expected and observed results are shown in Table 7 and in Figure 23.

The upper limits at 95% C.L. on the H production cross section times the branching ratio are between 12.6 and 76.0 times the Standard Model, while the expected limit estimated in pseudo-experiments is between 12.9 and 42.5 times the Standard Model.

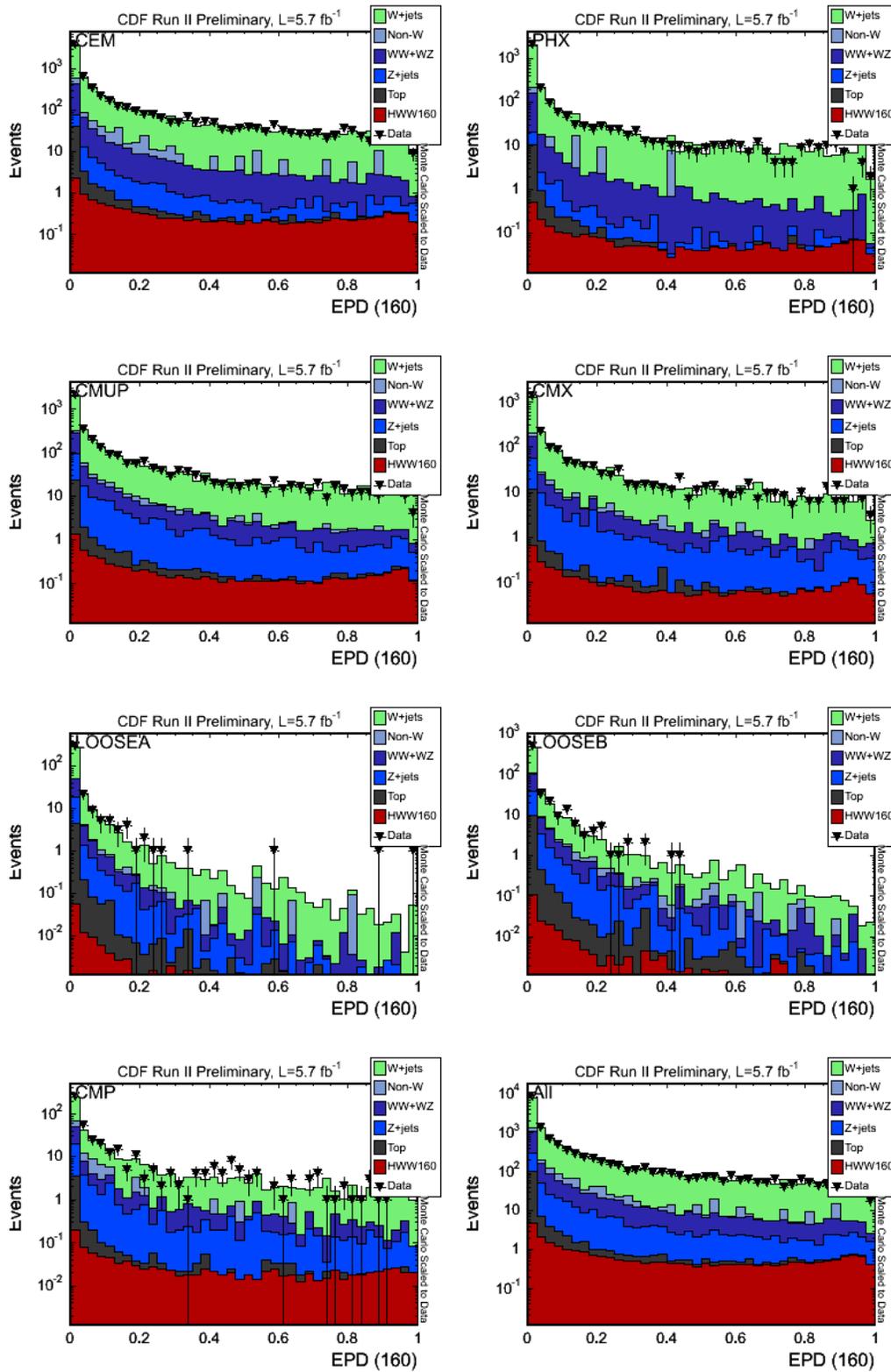


Figure 22: Comparison of the EPD output for data and the Monte Carlo prediction for HWW ($m_H = 160 \text{ GeV}/c^2$) signal and background for the different categories. Note that the scale is adapted for the LOOSE categories and for CMP.

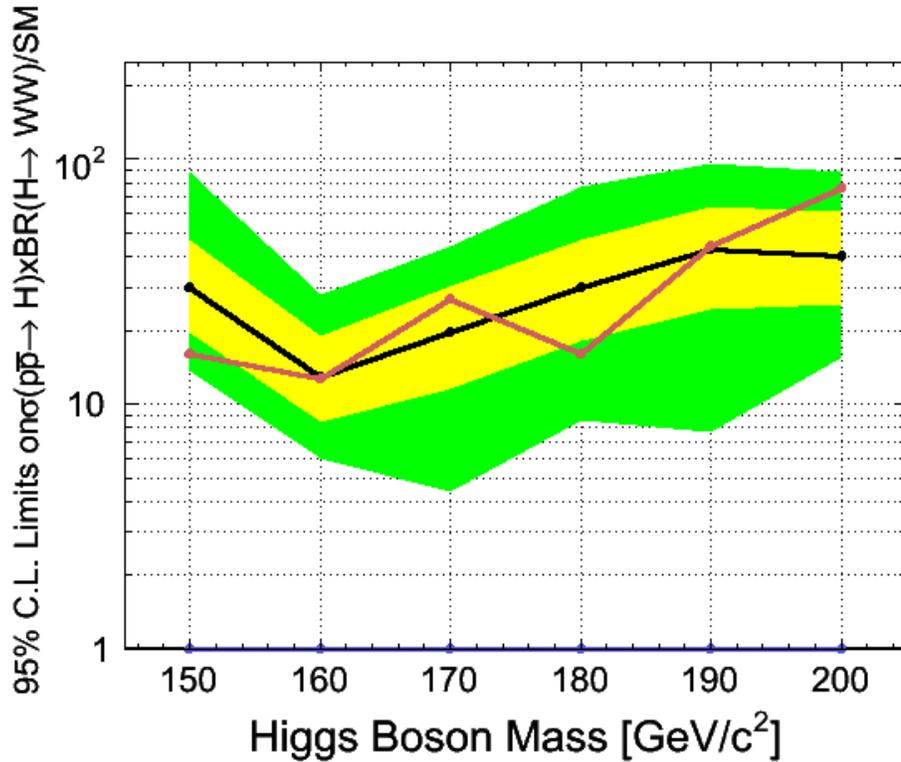


Figure 23: 95% C.L. upper limits on the H cross section multiplied by the $H \rightarrow WW$ branching ratio for Higgs boson masses of 150,160,170,180,190, and 200 GeV/c^2 . The plot shows the limit normalized to the predictions from the Standard Model.

10 Conclusions

We have analyzed 5.7 fb^{-1} of CDF Run II data, employing a matrix element analysis technique in a direct search for Higgs boson production. We observe no evidence for a Higgs signal and set 95% confidence level upper limits on the H production cross section times the branching ratio for $\sigma(p\bar{p} \rightarrow H) \times BR(H \rightarrow WW)/SM < 12.6$ to 76.0 for Higgs boson masses between $m_H = 150 \text{ GeV}/c^2$ to $m_H = 200 \text{ GeV}/c^2$. The expected (median) limit estimated in pseudo-experiments is $\sigma(p\bar{p} \rightarrow H) \times BR(H \rightarrow WW)/SM < 12.9$ to 42.5 at 95% C.L for Higgs masses between $m_H = 150 \text{ GeV}/c^2$ and $m_H = 200 \text{ GeV}/c^2$.

References

- [1] Barbara Alvarez, Florencia Canelli, Bruno Casal, Javier Cuevas, Peter Dong, Ricardo Eusebi, Chiara Ferrazza, Craig Group, Martina Hurwitz, Enrique Palencia, Alberto Ruiz, Bernd Stelzer, Rocio Vilar, Rainer Wallny, Search for Standard Model Higgs Boson Production using ME with 4.3fb⁻¹, CDF note 9861 (2009).
- [2] Florencia Canelli, Bruno Casal Laraña, Peter Dong, Bernd Stelzer, Rainer Wallny, Measurement of Single Top Quark Production in 3.2/fb of CDF Data Using a Matrix Element Technique, CDF note 9711 (2009).
- [3] F. Canelli, B. Casal Larana, C. Group, A. Ruiz, B. Stelzer, R. Vilar, Increasing Muon Acceptance with MET Plus Jet Triggers, CDF Note 9105 (2008).
- [4] S. Budd *et al.*, Event detection efficiency for single-top events and MC based background estimate for Summer 2006, CDF note 8286 (2006).
- [5] T. Peiffer, Search for the Higgs Boson in the WH channel with the CDF II Experiment, FERMILAB Master Thesis, 2008-1, Document Server <http://lss.fnal.gov/archive/masters/fermilab-masters-2008-01.shtml>, (2008).
- [6] Sarah Budd, Matthias Buehler, Catalin Ciobanu, Peter Dong, Thomas Junk, Jan Lueck, Thomas Muller, Svenja Richter, Bernd Stelzer, Wolfgang Wagner, Rainer Wallny, Estimation and Modeling of Non-W Background for Single-Top Searches, CDF Note 8489 (2006).
- [7] J. Adelman *et al.*, Method II For You, CDF Note 9185.
- [8] P. Dong *et al.*, Search for Standard Model Higgs Boson Production using ME+BDT 2.7 1/fb, CDF Note 9404 (2008).
- [9] L. Demortier, Bayesian treatments of Systematic Uncertainties, Proceedings of Advanced Statistical Techniques in Particle Physics, Grey College, Durham, 18 - 22 March 2002, <http://www.ippd.dur.ac.uk/Workshops/02/statistics/proceedings.shtml>.
- [10] C. Ciobanu, T. Junk, T. Müller, P. Savard, B. Stelzer, W. Wagner, T. Walter, Likelihood Function for Single Top Search with 162 pb⁻¹, CDF Note 7106 (2004).
- [11] L. Read, J.Phys G **28**, 2693 (2002) and T. Junk, Nucl. Instum. Meth. **434**, 435 (1999). See also P. Bock *et al.* (The LEP Collaborations), CERN-EP-98-046 (1998) and CERN-EP-2000-055 (2000).

- [12] D. Acosta *et al.*, First Measurements of Inclusive W and Z Cross Sections from Run II of the Fermilab Tevatron Collider, Phys. Rev. Lett. **94**, 091803 (2005).
- [13] http://www-cdf.fnal.gov/internal/physics/top/RunIITopProp/tools/ks_tests.html
- [14] John Conway, Tevatron Run 2 Cross Sections and Branching Ratios, <http://www.physics.ucdavis.edu/~conway/research/higgs/smhiggs-tev.html> (2003).
- [15] A.Djouadi, J.Kalinowski and M.Spira, Comp. Phys. Commun. 108 C (1998) 56, [arXiv:hep-ph/9704448v1](https://arxiv.org/abs/hep-ph/9704448) (1997).
- [16] Florencia Canelli Bruno Casal Larana Craig Group Alberto Ruiz Bernd Stelzer Rocio Vilar , Increasing Muon Acceptance with the MET Plus Jet Triggers, CDF Note 9106 (2007).
- [17] Particle Data Group, The Review of Particle Physics, <http://pdg.lbl.gov> (2006).
- [18] Martina Hurwitz, Florencia Canelli, Ricardo Eusebi, Craig Group, Enrique Palencia, Bruno Casal, Barbara Alvarez, Bernd Stelzer, Measurement of WW+WZ production cross section in $l\nu jj$ channel using 4.6 fb^{-1} , CDF Note 9923 (2009).