CDF/PHYS/EXOTIC/CDFR/10086
March 10, 2010
Version 2.0

Search for H $\rightarrow$ WW Production Using 5.3 fb^{-1}

Doug Benjamin, Mark Kruse, Seig Oh, Geumbong Yu

Duke University

Massimo Casarsa, Eric James, Sergo Jindariani

Fermilab

Roman Lysak

IEP SAS, Slovakia

Maria d'Errico, Simone Pagan Griso, Donatella Lucchesi

INFN and University of Padova

Dean Andrew Hidas

Rutgers University

Anadi Canepa

TRIUMF and University of Pennsylvania

Aidan Robson, Rick St. Denis, Peter Bussey

University of Glasgow

Matthew Herndon, Jason Nett, Jennifer Pursley

University of Wisconsin, Madison

Abstract

We present a search for Standard Model (SM) Higgs to WW^* production in dilepton plus $\not\!\!E_T$ channels using approximately 5.3 fb^{-1} of integrated luminosity. The analysis described here relies on almost the same techniques described in CDF Note 9863. The primary differences in the two searches are: the addition of new data (from 4.8 fb^{-1} to 5.3 fb^{-1}); the addition of trilepton search channels; a loosened criteria for likelihood-based electron selection which improves acceptance while doubling the electron fake rate from that of the previous analysis. We utilize multivariate discriminants separately optimized for events containing zero, one, and two or more jets in the opposite-sign dilepton final states. For all jet multiplicites, we take into account four possible Higgs boson production mechanisms: gluon fusion, associated production with a W boson, associated production with a Z boson, and vector boson fusion. We also include an analysis of an opposite-sign dilepton event sample with zero or one jets with a dilepton mass of less than 16 GeV/c^2 and considers only potential signal from gluon fusion. We also include an analysis of a same-sign dilepton event sample which can potentially contain additional signal events originating from the associated Higgs boson production mechanisms. In this update, we also include an analysis of a trilepton event sample which may also potentially contain additional signal events originating from the associated Higgs boson production mechanisms. Based on the dilepton opposite-sign selection criteria defined in CDF Note 9195 and a data sample of about 5.3 fb^{-1}, we observe (summing over all jet bins) a total of 1843 candidate events compared with an expectation of 1840 ± 150 background events and 32 ± 5 signal events for a Standard Model Higgs boson with a mass of 165 GeV/c^2. In our low-dilepton mass opposite-sign event candidate sample we observe 104 candidate events compared with an expectation of 99 ± 8 background events and 0.89 ± 0.15 signal events for a Standard Model Higgs boson with a mass of 165 GeV/c^2. In our same-sign dilepton event candidate sample we observe 88 candidate events compared with an expectation of 81 ± 16 background events and 2.1 ± 0.3 signal events for a Standard Model Higgs boson with a mass of 165 GeV/c^2. In our trilepton event candidate sample we observe 47 candidate events compared with an expectation of 39 ± 5 background events and 1.4 ± 0.2 signal events for a Standard Model Higgs boson with a mass of 165 GeV/c^2. Based on these samples, we determine an observed 95% C.L. upper cross section limit of 1.13 times the SM prediction at NNLL for a Higgs mass of 165 GeV/c^2 to be compared with the value for the median of the expected limit (1.03). Results for eighteen other Higgs mass hypotheses ranging from 110 GeV/c^2 to 200 GeV/c^2 are also presented.

Contents

1 Introduction

In this note, we present a search for Standard Model $H \to WW^* \to l\nu l\nu$ production using approximately 5.3 fb^{-1} of integrated luminosity. This search builds on CDF Note 9863 [1], from which it inherits the most recent improvements of the analysis, most notably the likelihood-based selection of central electrons, and the addition of events containing low-dilepton mass pairs (< 16 GeV/c^2) as a separate search region (CDF Note 9864 [2]).

Many small improvements to the analysis have been incorporated in this update. These are described in detail throughout the text but are summarized here for the reader's convenience:

- A loosened cut on the likelihood-based central electron selection, which improves electron acceptance while approximately doubling the fake rate. This modification is described in detail in Section 2.

- The addition of periods 24 and 25 data, which corresponds to approximately 500 pb^{-1}, or a luminosity increase of 10.6%.

- Inclusion of the reprocessed period 18 data; approximately 100 pb^{-1} of period 18 data which was previously marked bad for silicon data due to a bad silicon calibration has been recovered and is now considered good for silicon.

- Reweighting of the gluon fusion ($gg \to H$) Pythia Monte Carlo samples to match the resummed Higgs p_T spectrum from theoretical calculations.

- Fill in the missing Higgs signal samples on the mass grid, so that a limit is determined for every 5 GeV interval (this involved adding samples for m_H = 115, 125, 135, 185, and 195).

- Incorporate the signal contributions from vector boson fusion and associated production with a W or Z boson in the event sample with opposite-sign dileptons and zero reconstructed jets. Previously this event sample considered only signal contributions from gluon fusion; adding the additional signal production modes corresponds to an increase in expected signal of 16% for a Higgs boson of $m_H = 115$, 6.2% for $m_H = 160$, and 6% for $m_H = 200$ GeV.

- Inclusion of trilepton searches for the signatures $WH \to WWW^* \to l\nu l\nu l\nu$ and $ZH \to ZWW^* \to lll\nu + X$. These searches were separately blessed with 4.8 fb^{-1} as described in CDF Note 10020 [3], and are here updated with the same improvements (where applicable) as the dilepton analyses.

- Improved treatment of many systematic uncertainties, including further investigation of potential shape systematics, described in detail in Section 6.

Another improvement implemented in this version of the analysis is the extension of the $W\gamma$ Monte Carlo samples to include modeling of the later run periods. Previously, the $W\gamma$ samples consisted of one fully run-dependent sample which went through period 13. We now use an additional Monte Carlo sample generated with a higher luminosity profile (based on a specific set of runs from periods 10 to 12) to model data collected from periods 14 onward (periods 14 to 25 for the case of 5.3 fb^{-1}). In order to avoid complications in data-to-Monte-Carlo comparisons, we treat periods 14 to 25 as a single data period with a unique set of scale factors, trigger efficiencies, etc. Monte Carlo events in the higher luminosity samples are each assigned to a specific run number contained within period 15 and a unique event number to ensure that events are not accidentally removed as duplicates. All of the signal samples and most of the other background samples were already updated this way in [4].

2 Updated Lepton Identification

The current Higgs searches are mainly statistically limited; for this reason it's important to maximize the acceptance of the analysis. Since we're looking at both Ws decaying leptonically, one of the key points in maximizing the acceptance is the lepton identification for electrons and muons[1].

Previous $H \to WW$ analyses rely on the lepton identification described in CDF Note 8538 [5], which aims to exploit every region of the CDF detector where an energy measurement can be made to define a set of unique (non-overlapping) categories for electron and muon identification. The muon selections were updated in the most recent $H \to WW$ analysis, described in CDF Note 9697 [4]. The analysis presented in 9697 used the following categories:

- Electrons: TCE, PHX
- Muons: CMUP, CMP, CMU, CMX, CMXMsKs, BMU, CMIOCES, CMIOPES, CrkTrk

The triggers used to collect the data sample for this analysis are based on leptons, and are deeply connected with the lepton definitions. Trigger paths exploited by the analysis in Note 9402 can be summarized in 4 categories (present in the CDF trigger tables in various versions):

- MUON_CMUP18
- MUON_CMP18_PHI_GAP
- MUON_CMX18
- ELECTRON_CENTRAL_18
- MET_PEM

The CMU muons triggered by the MUON_CMU18_ETA_GAP path will be included in a later version of the analysis, as this trigger has only been working properly since period 21 data-taking.

The aim of the study described in this CDF Note is to improve electron identification using a likelihood method. In Section 2.1 we describe in detail this likelihood method. Sections 2.2 and 2.3 describe the lepton ID scale factor and fake rate calculations for each of the categories. Finally, we perform the strongest check of the new electrons' behavior in Section 2.4, measuring the cross section of the Drell-Yan ($p\bar{p} \to Z/\gamma^\star \to l^+l^-$, $l = e, \mu$) process separately for each dilepton type.

[1] τ's are currently considered only if decaying to an electron or muon. Studies are ongoing to include hadronically decaying τ's in the same framework, but they will not be discussed in this note.

2.1 Likelihood Method for Lepton Identification

The standard method for lepton identification uses cuts on many different identification variables (see TCE definition in Tab. 1). Each cut decreases reconstruction efficiency. This way candidates which fail even one cut by a small margin are removed from the sample. Catching each possible lepton candidate is extremely important in such low-statistics searches as looking for the Higgs boson at the Tevatron.

Therefore we developed a method which tries to exploit the identification variables in better way. Instead of cutting on each individual variable, we combine identification variables into one discriminant and cut only on the value of this final discriminant. It is expected that this should increase the efficiency of reconstruction (while keeping the same fake rate).

The method we use is a likelihood method, which calculates the probability of a candidate to be real and "fake" lepton and computes the likelihood discriminant. We apply the method on central electron candidates, so the newly identified central electrons (Likelihood based electrons – LBE) replace in our analysis the up-to-now used TCE.

2.1.1 Likelihood definition for central electrons

For our likelihood method, we choose most of the identification, non-kinematic variables which were used in identification of TCE (for definition of TCE, see Tab. 1). Specifically, these are

- E_{HAD}/E_{EM} – the ratio of the hadronic calorimeter energy to the electromagnetic calorimeter energy associated with the candidate
- E/P – the ratio of the EM cluster transverse energy to the COT track transverse momentum
- $Lshr$ – the lateral shower profile in the transverse plane to the electron direction
- $CalIso$ – The energy E_T in a cone of radius $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \leq 0.4$ around the electron cluster excluding the electron cluster divided by the energy in the electron cluster:
$$CalIso = \frac{E_T^{cone} - E_T^{electron}}{E_T^{electron}}$$
- $TrkIso$ – the same variable as above $CalIso$ but measured using tracks instead of calorimeter
- $Q \times \Delta x_{CES}$ – The distance in the r-ϕ plane between the extrapolated, COT beam constrained track and the best matching CES cluster, times the charge Q of the track.

- Δz_{CES} – The distance in the r-z plane between the extrapolated, COT beam constrained track and the best matching CES cluster.
- $NCotHitsAx$ – number of COT hits on axial layers belonging to track associated to the candidate electron
- $NCotHitsSt$ – number of COT hits on stereo layers belonging to track associated to the candidate electron
- χ^2_{COT} – χ^2 associated with the COT hits belonging to track
- $NSvxHits$ – number of SVX hits belonging to track associated to the candidate electron

These variables are then used in calculation of the likelihood which has the following form:

$$\mathcal{L}(\vec{x}) = \frac{L_{sig}}{L_{sig} + L_{bckg}} = \frac{\prod_{i=1}^{N} P_i^{sig}(x_i)}{\prod_{i=1}^{N} P_i^{sig}(x_i) + \prod_{i=1}^{N} P_i^{bckg}(x_i)} \tag{1}$$

where

- x_i – means the i-th identification variable used in the likelihood. These are the 11 variables described above.
- N - the number of ID variables, in our case 11.
- $P_i^{sig}(x_i), P_i^{bckg}(x_i)$ - these are the functions which give the probability to obtain a certain value x_i for the i-th ID variable. These are obtained using signal and background ("fakes") templates.

By definition, the value of $\mathcal{L}$ is restricted within the range $(0.0, 1.0)$. As can be seen, the essential feature of this method is to obtain templates for ID variables for real electrons and "fakes".

2.1.2 Building of signal and background templates

We use data to build the signal and background templates. This is done by selecting samples dominated by real high p_T electrons (Z boson candidates) for signal templates and dominated by fake electrons (dijet data) for background templates.

Building of the signal templates is done as follows (very similar selection to one use for efficiency calculation, see [5]). We select Z candidates (76 GeV $<$ $M_{\ell\ell}$ $<$ 106 GeV) from high p_T electron data (bhelbj dataset) where one electron is a fully identified central electron (TCE) while the second electron candidate must pass just basic central electron object criteria (must pass "CEM probe" criteria, see definition in Tab. 2). We also require opposite charge between identified electron and "Probe" object ("probe" should

	TCE
Region	central
Fiducial	track fiducial to CES
Track p_T	≥ 10 $(5 if E_T < 20)$
Track $\|z_0\|$	≤ 60 cm
# Ax SL (5hits)	≥ 3
# St SL (5hits)	≥ 2
Conversion	$\neq 1$
E_{HAD}/E_{EM}	$\leq 0.055 + 0.00045 * E$
Iso/E_T	≤ 0.1
Lshr	≤ 0.2
E/P	$< 2.5 + 0.015 * E_T$
signed CES ΔX	$-3 \leq q * \Delta X \leq 1.5$ cm
CES $\|\Delta Z\|$	< 3 cm
Track	Beam constrained

Table 1: Definition of TCE.

Central electron probe
Track or SMX fiducial to CES
Track $p_T \geq 5$ GeV
Track $Z_0 \leq 60$ cm

Table 2: Definition of central electron probe object.

Central fakeable electron
Track fiducial to CES
$E_{HAD}/E_{EM} <= 0.125 + 0.00045 * E$
$CalIso < 0.3$
not a conversion

Table 3: Definition of central electron fakeable object.

also not be a conversion) where both have $E_T > 10$ GeV. Such a sample is completely dominated by Drell-Yan events. The ID variables of these probes are used to make the signal templates. In the case where both legs are identified as TCE, they are both used for template creation.

Building of the templates for "fakes" is done the following way (very similar selection to one used in fake rate calculation, see [5]). We select candidates from jet data (JET20 dataset - gjt1bk) where there is one identified jet ($E_T > 20$ GeV) and another fakeable object ("CEM Fakeable" object, see definition in Tab. 3). We also require

- the "fakeable" object must pass basic kinematic criteria: $E_T > 10\ GeV$, track $p_T > 5\ GeV$, $|z_0| < 60\ cm$.
- the combination of jet and fakeable object fall out of the Z mass peak $(76\ GeV, 106\ GeV)$ to remove contribution from Z events
- fakeable electron is not in the leading E_T jet in order to remove trigger bias (also leading jet must have $E_T^{raw} > 20$ GeV).
- leading jet $|z_0| < 60$ cm and $|\Delta z_0(\text{“}fakeable'' \ object\ and\ leading\ jet)| < 4$ cm.
- reject cosmic events

Since we obtained our templates using data samples, our templates contain background events. In the case of signal, this means "fake" electrons, in case of the templates for "fakes", this means real electrons. Therefore we perform background subtraction to the templates. The background subtraction is done the following way:

- For signal: We plot the distribution of likelihood variables for the probes. As can be seen in Fig. 1, the histogram has a very sharp peak at 1.0 while having a small bump at 0.0 (for the background case, this is opposite). We normalize the likelihood distribution for the "fakes" in such way that it has the same value in the first bin as the value in the 1st bin of the signal likelihood distribution. By obtaining the integral over whole range (0.0,1.0), we get an estimate of the number of "fake" electrons in the signal templates. This turns out to be small ($\sim 3\%$). We then subtract such fraction of background templates from signal templates.
- For fakes: we estimate the expected contribution of "signal" events (real electrons) in the background templates using $W \rightarrow \ell\nu$ and $Z \rightarrow ee$ MC samples. This turns out to be an even smaller effect ($\sim 1\%$), but we again subtract the signal templates (appropriately scaled to 1%) from the background templates.

The signal and background templates for all used electron ID variables are shown in Figs. 2 through 5. Just for comparison, there are shown distributions of variables for data and MC ($Z \rightarrow ee$) in signal case and JET20 and JET100 samples in case of background

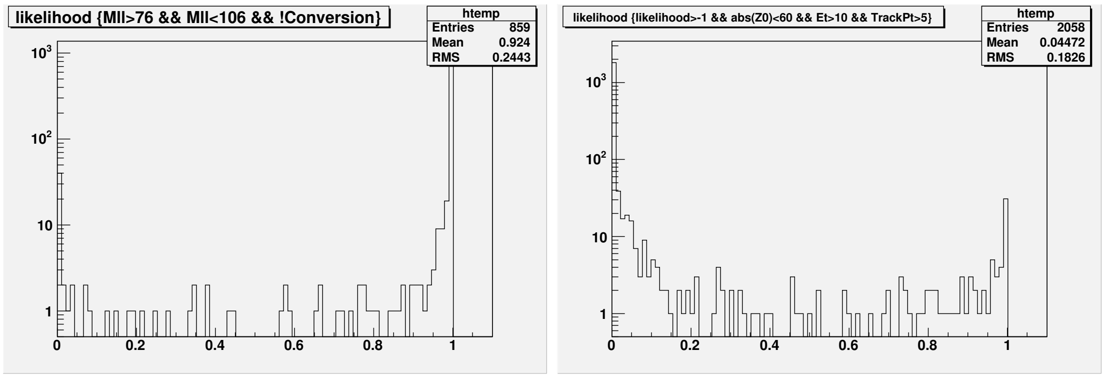

Figure 1: Likelihood distribution for signal probes (left) and fakeable objects (right).

templates. For signal, the data and MC distributions agree quite well for most of the variables. For background, JET20 and JET100 do not agree very well due to the fact that some of the variables are E_T dependent, and the E_T spectrum of these two datasets is quite different. Since the JET20 E_T distribution is closer to what we expect for our signal sample, we use the templates obtained from the JET20 sample.

2.1.3 Likelihood based electron (LBE)

We define a new central electron category, the likelihood based electron (LBE), which we will use instead of TCE electrons. We first apply basic identification and kinematic criteria on electron candidates and then make a cut on the likelihood value at the end. All the requirements used for LBEs are listed below:

- have CEM fiducial track
- $Z_0(track) < 60$ cm
- electron candidate is not a conversion
- $E_{HAD}/E_{EM} < 0.125$ to satisfy the trigger requirement
- $CalIso < 0.3$ - weak cut on the calorimeter isolation to satisfy the requirement on fakeable objects used in the template building procedure
- $p_T(track) > 10$ GeV ($p_T(track) > 5$ GeV if $E_T < 20$ GeV)
- $\mathcal{L} > \mathcal{L}_{cut}$ - value of likelihood must be greater than a cut value, which we set to $\mathcal{L}_{cut} = 0.90$ according to the optimization described in the next section, Sec. 2.1.4. The value of this cut was changed from $\mathcal{L}_{cut} = 0.99$ as compared to the previous version of this analysis described in CDF Note 9863 [1].

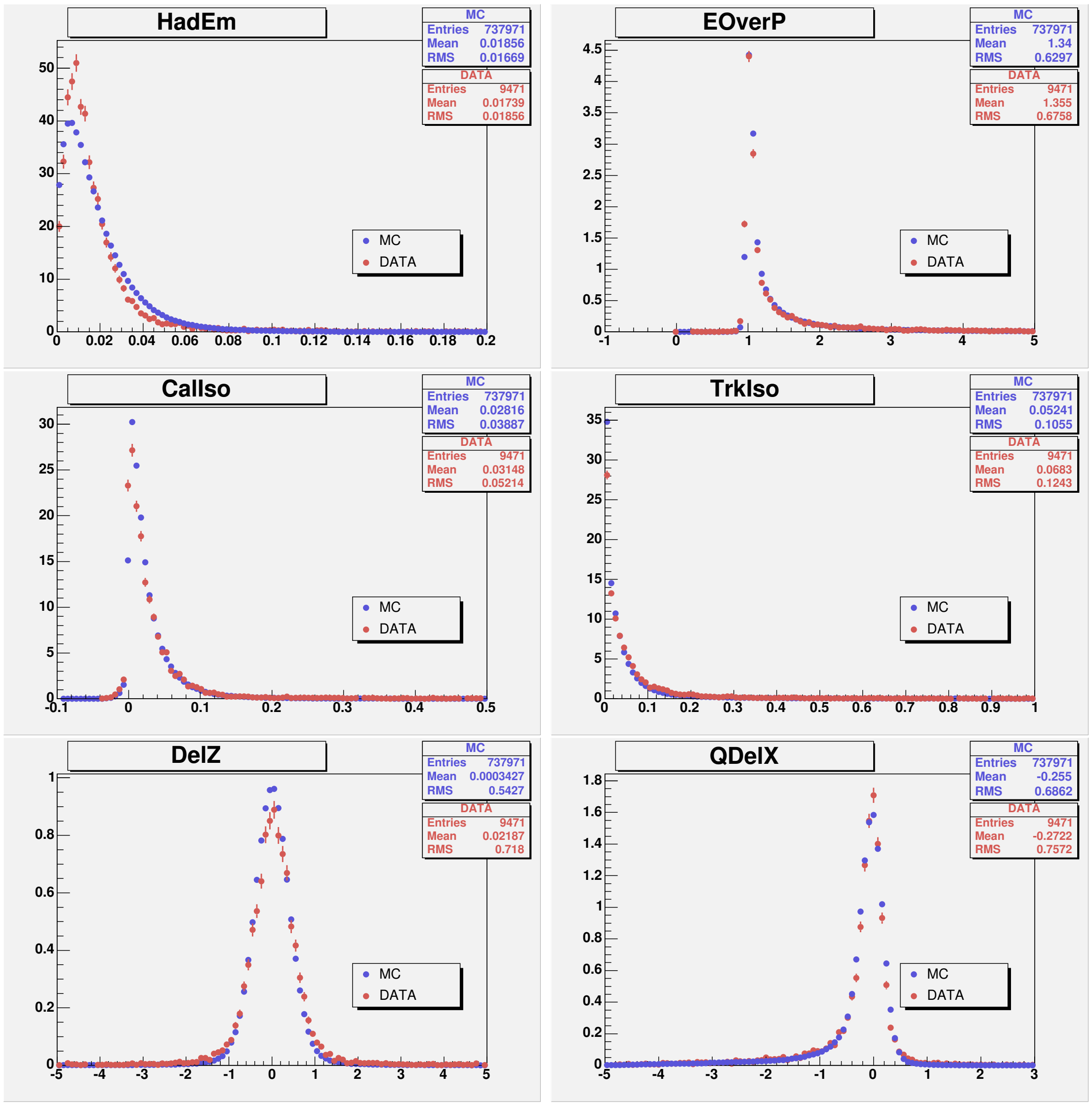

Figure 2: Signal templates for the likelihood method.

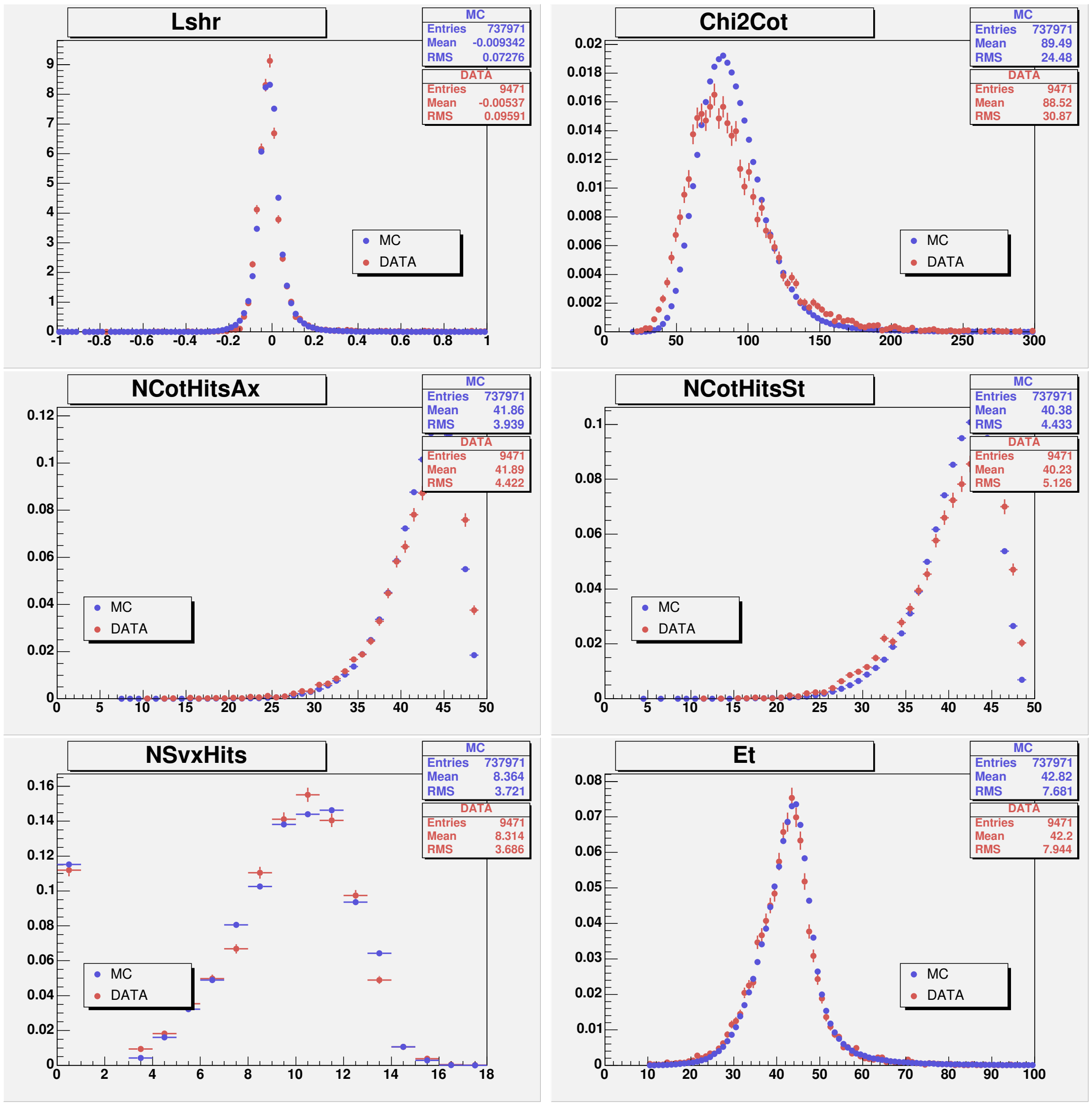

Figure 3: Signal templates for the likelihood method.

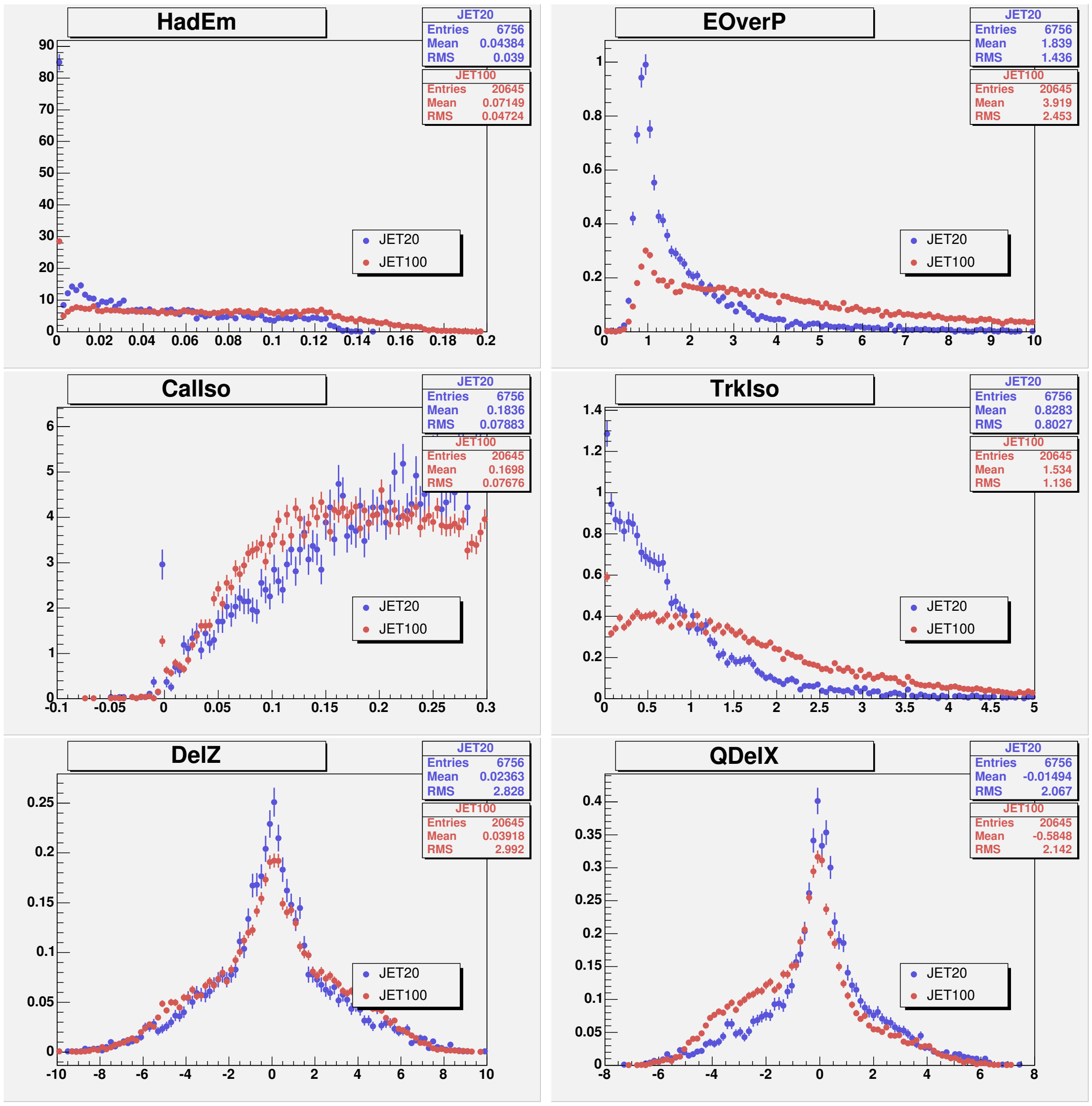

Figure 4: Background templates for the likelihood method.

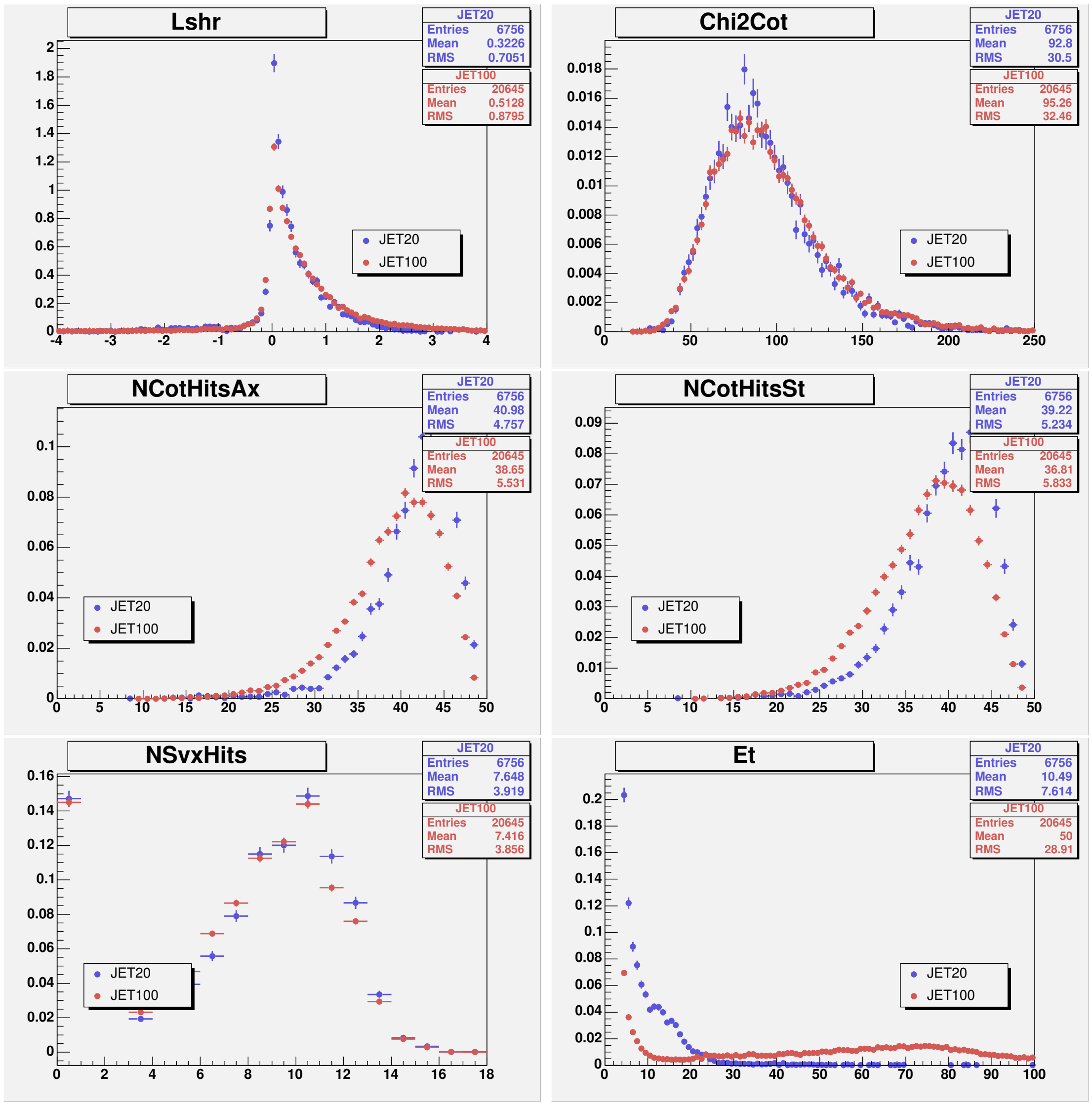

Figure 5: Background templates for the likelihood method.

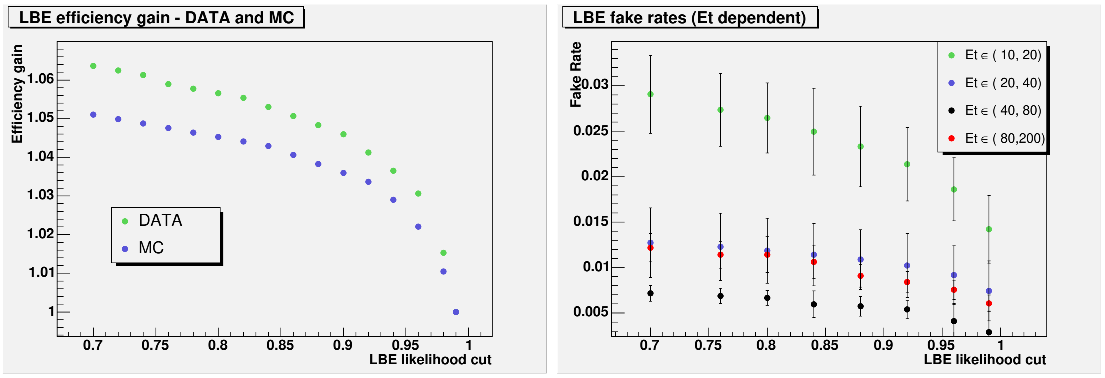

Figure 6: The likelihood ratio cut dependence of reconstruction efficiency (left) and fake rates for LBEs (right). The reconstruction efficiency is shown as a ratio with the efficiency for a LBE > 0.99 cut; the value of 1.05 for LBE > 0.90 indicates a 5% gain over a cut of LBE > 0.99. The fake rates are shown summed over bins of jet E_T; by relaxing the LBE cut from 0.99 to 0.90, the fake rate increases for low E_T jets, but stays nearly the same (within systematic and statistical uncertainties) for jets with $E_T > 20$ GeV.

One assumption which we use for LBE is that it has the same trigger efficiency as TCE. We checked this assumption by selecting $e - \mu$ events from our signal region with 0 jets, with a requirement that the μ fires the trigger. By checking how often TCE and LBE pass the trigger requirement, we obtained the following trigger efficiency for TCE $(78/83 = 0.94)$ and for LBE $(84/88 = 0.95)$. Within our limited statistics, we conclude that the trigger efficiency is the same for LBE and TCE. However, for the future we plan to make more a detailed study of this assumption.

2.1.4 Optimization of LBE likelihood cut

We optimized the cut on the likelihood ratio by calculating the reconstruction efficiency and LBE fake rates for different possible cuts. The reconstruction efficiency is calculated for the Drell-Yan control region using both the data and the Drell-Yan MC samples. The fake rates are calculated from jet-triggered data samples as described in Sec. 2.3. Looking at these dependencies (see Fig. 6), we conclude that moving the cut from 0.99 down to 0.9 will get us most of the available efficiency gain (5% in data) while still keeping the fake rates about the same (within systematic and statistical uncertainties) except for very low E_T jets. Thus we decide to use a cut of 0.90 for LBEs in this updated analysis.

Period	SF
0-9	1.001
10-13	0.945
14-25	0.866

Table 4: Scale Factor evaluated for muons pointing to the bottom 90° of the detector.

2.2 Lepton ID Scale Factors

To measure lepton ID scale factors, we measure the lepton efficiencies in both data and MC simulation and divide the two result to apply a per-event SF to our MC simulations. We again use the *tag & probe* method, selecting Z candidates in the $76 < m(ll) < 106$ GeV dilepton invariant mass window, as described in the CDF Note 8538 [5]. As explained in detail in this reference, we use probe objects for which we assume MC simulation well describes the efficiency, and calculate only the efficiency on top of these requirements. Generic probes are defined for Central and Forward muons and then used with the fiduciality requirement to define the actual probe to be used in the efficiency calculation. Background subtraction is also performed using data from the Z-mass sidebands.

Results are shown in Tables 5-18, grouped by run period ranges as defined in the official webpage [6]. The only exception is what we call period 0k, which groups also data acquired in the 0m dataset, since these data are pretty homogenous. Efficiencies, as well as scale factors, are calculated for both track isolated (which are the ones actually used in the analysis, except in the case of LBE where track isolation is an input to the likelihood) and non-track isolated leptons. Track isolation (< 0.1) is required in a cone of 0.4 around the leptons. We also calculate the electron efficiencies that will be used in the analysis with exactly the same method described in [5]. As also in this reference, we calculate efficiencies for stubless categories as the difference of the one removing vetos and the veto efficiency, in order to properly account for stubbed muon SFs using the same formulas described in that note. For CMIOPES and BMU categories, we apply the PES tracking efficiency scale factor to account for forward tracking reconstruction scale factors between data and MC. This scale factor is evaluated using electrons, since that is the only way to measure it in data, but should also model well what happens for muons.

We observe an excess of central tracks in MC reconstructed near the bottom 90° of the COT. Figure 7 shows the ϕ distribution for central tracks with $\rho_{COT} > 140$ cm in Z events selected with a CMUP tag leg and requiring $86 < m(ll) < 106$ GeV for different run periods (p0-p9, p10-p13, and p14-p23). Data and Monte Carlo distributions are normalized in the region $0° < \phi < 225°$ and $315° < \phi < 360°$. Since the effect is clearly dependent on the run period, we apply a scale factor to all central muon types which point in the $225° < \phi < 315°$ region depending on the run period as shown in Table 4.

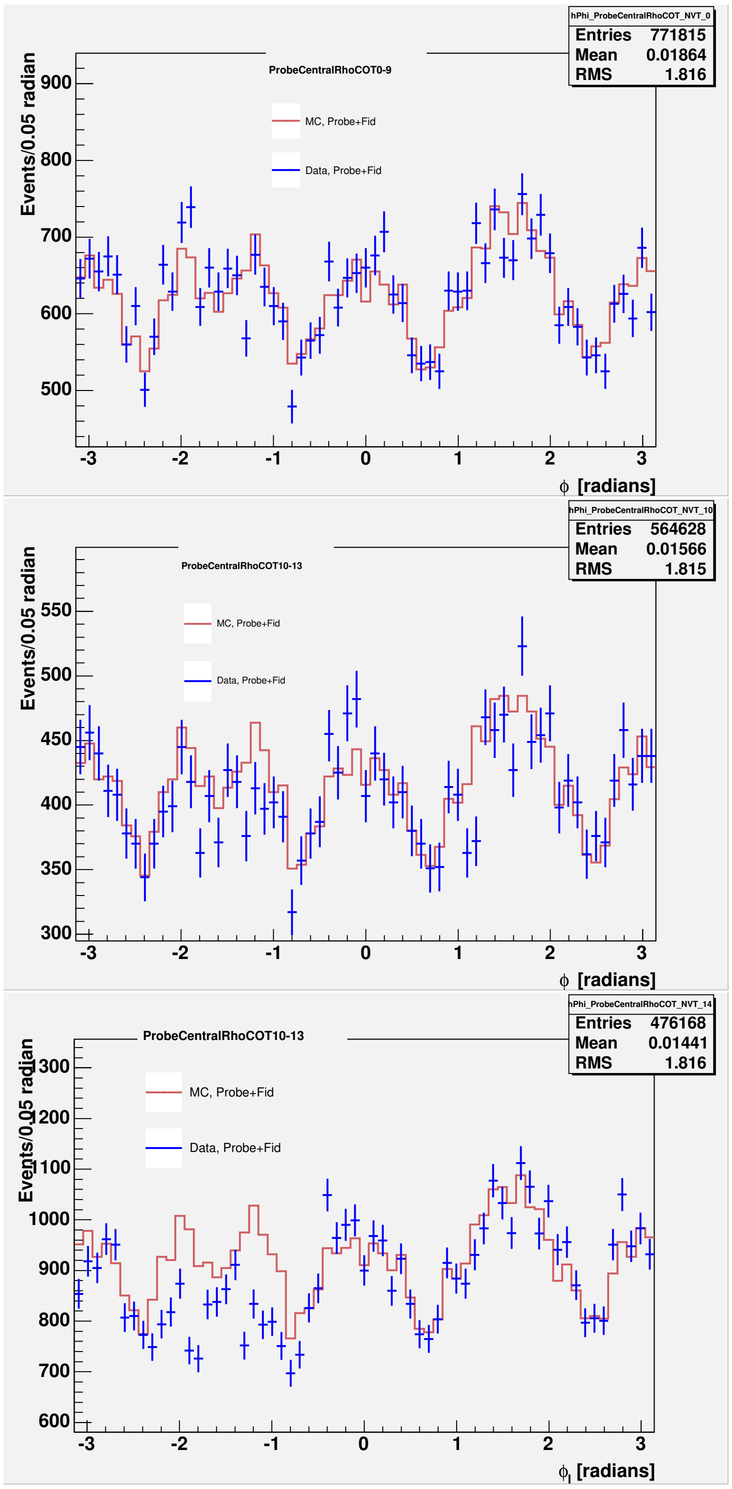

Figure 7: ϕ distribution for central tracks in Z events selected with a CMUP tag. Data periods 0 to 9 shown on top, periods 10 to 13 in the middle, and periods 14 to 23 on the bottom.

Run Range: 0d

	data	MC	Scale Fac
CMUP	0.879 ± 0.011	0.903 ± 0.001	0.973 ± 0.012
CMU	0.000 ± 1.000	0.000 ± 1.000	0.000 ± 1.000
CMP	0.000 ± 1.000	0.000 ± 1.000	0.000 ± 1.000
CMX	0.947 ± 0.014	0.922 ± 0.003	1.027 ± 0.016
CMXMsKs	0.000 ± 1.000	0.000 ± 1.000	0.000 ± 1.000
BMU	0.826 ± 0.023	0.733 ± 0.004	1.127 ± 0.032
CMIOCES	0.367 ± 0.006	0.350 ± 0.002	1.049 ± 0.019
CMIOPES	0.689 ± 0.012	0.689 ± 0.002	1.000 ± 0.018
CrkTrk μ	0.733 ± 0.011	0.765 ± 0.002	0.958 ± 0.015
CMUP Trk Iso	0.869 ± 0.011	0.896 ± 0.001	0.970 ± 0.012
CMU Trk Iso	0.000 ± 1.000	0.000 ± 1.000	0.000 ± 1.000
CMP Trk Iso	0.000 ± 1.000	0.000 ± 1.000	0.000 ± 1.000
CMX Trk Iso	0.927 ± 0.014	0.912 ± 0.003	1.016 ± 0.016
CMXMsKs Trk Iso	0.000 ± 1.000	0.000 ± 1.000	0.000 ± 1.000
BMU Trk Iso	0.816 ± 0.024	0.725 ± 0.004	1.126 ± 0.034
CMIOCES Trk Iso	0.364 ± 0.006	0.347 ± 0.002	1.049 ± 0.019
CMIOPES Trk Iso	0.667 ± 0.012	0.683 ± 0.002	0.977 ± 0.018
CrkTrk μ Trk Iso	0.719 ± 0.011	0.756 ± 0.002	0.951 ± 0.015

Table 5: Muon efficiencies for dataset 0d (period 0) with and without track isolation cut.

Run Range: 0h

	data	MC	Scale Fac
CMUP	0.847 ± 0.008	0.903 ± 0.001	0.938 ± 0.009
CMU	0.000 ± 0.500	0.000 ± 0.500	0.000 ± 0.500
CMP	0.000 ± 0.500	0.000 ± 0.500	0.000 ± 0.500
CMX	0.937 ± 0.015	0.920 ± 0.002	1.020 ± 0.017
CMXMsKs	0.000 ± 0.500	0.000 ± 0.500	0.000 ± 0.500
BMU	0.797 ± 0.018	0.720 ± 0.003	1.107 ± 0.025
CMIOCES	0.364 ± 0.005	0.343 ± 0.001	1.060 ± 0.015
CMIOPES	0.609 ± 0.012	0.606 ± 0.002	1.005 ± 0.020
CrkTrk μ	0.777 ± 0.009	0.795 ± 0.002	0.978 ± 0.012
CMUP Trk Iso	0.834 ± 0.008	0.895 ± 0.001	0.933 ± 0.009
CMU Trk Iso	0.000 ± 0.500	0.000 ± 0.500	0.000 ± 0.500
CMP Trk Iso	0.000 ± 0.500	0.000 ± 0.500	0.000 ± 0.500
CMX Trk Iso	0.906 ± 0.016	0.908 ± 0.002	0.998 ± 0.018
CMXMsKs Trk Iso	0.000 ± 0.500	0.000 ± 0.500	0.000 ± 0.500
BMU Trk Iso	0.767 ± 0.018	0.714 ± 0.003	1.074 ± 0.026
CMIOCES Trk Iso	0.354 ± 0.005	0.340 ± 0.001	1.043 ± 0.015
CMIOPES Trk Iso	0.588 ± 0.012	0.601 ± 0.002	0.979 ± 0.020
CrkTrk μ Trk Iso	0.760 ± 0.009	0.786 ± 0.002	0.967 ± 0.012

Table 6: Muon efficiencies for dataset 0h (periods 1-4) with and without track isolation cut.

Run Range: 0i1

	data	MC	Scale Fac
CMUP	0.840 ± 0.011	0.902 ± 0.001	0.932 ± 0.013
CMU	0.000 ± 0.577	0.000 ± 0.577	0.000 ± 0.577
CMP	0.000 ± 0.577	0.000 ± 0.577	0.000 ± 0.577
CMX	0.943 ± 0.018	0.919 ± 0.003	1.026 ± 0.019
CMXMsKs	0.000 ± 0.577	0.000 ± 0.577	0.000 ± 0.577
BMU	0.776 ± 0.023	0.721 ± 0.004	1.076 ± 0.032
CMIOCES	0.370 ± 0.006	0.341 ± 0.002	1.085 ± 0.018
CMIOPES	0.617 ± 0.015	0.599 ± 0.003	1.029 ± 0.025
CrkTrk μ	0.774 ± 0.012	0.793 ± 0.002	0.976 ± 0.015
CMUP Trk Iso	0.830 ± 0.011	0.892 ± 0.001	0.931 ± 0.013
CMU Trk Iso	0.000 ± 0.577	0.000 ± 0.577	0.000 ± 0.577
CMP Trk Iso	0.000 ± 0.577	0.000 ± 0.577	0.000 ± 0.577
CMX Trk Iso	0.903 ± 0.018	0.905 ± 0.003	0.998 ± 0.020
CMXMsKs Trk Iso	0.000 ± 0.577	0.000 ± 0.577	0.000 ± 0.577
BMU Trk Iso	0.752 ± 0.023	0.715 ± 0.004	1.051 ± 0.033
CMIOCES Trk Iso	0.363 ± 0.006	0.336 ± 0.002	1.082 ± 0.017
CMIOPES Trk Iso	0.600 ± 0.015	0.594 ± 0.003	1.009 ± 0.025
CrkTrk μ Trk Iso	0.758 ± 0.012	0.783 ± 0.002	0.969 ± 0.015

Table 7: Muon efficiencies for dataset 0i1 (periods 5-7) with and without track isolation cut.

Run Range: 0i2

	data	MC	Scale Fac
CMUP	0.857 ± 0.008	0.897 ± 0.001	0.955 ± 0.009
CMU	0.000 ± 0.577	0.000 ± 0.577	0.000 ± 0.577
CMP	0.888 ± 0.029	0.919 ± 0.003	0.965 ± 0.032
CMX	0.920 ± 0.013	0.914 ± 0.002	1.007 ± 0.014
CMXMsKs	0.843 ± 0.033	0.907 ± 0.004	0.930 ± 0.036
BMU	0.799 ± 0.015	0.725 ± 0.003	1.099 ± 0.021
CMIOCES	0.332 ± 0.004	0.305 ± 0.001	1.086 ± 0.014
CMIOPES	0.586 ± 0.011	0.598 ± 0.002	0.980 ± 0.018
CrkTrk μ	0.708 ± 0.009	0.722 ± 0.001	0.973 ± 0.012
CMUP Trk Iso	0.837 ± 0.008	0.887 ± 0.001	0.944 ± 0.009
CMU Trk Iso	0.000 ± 0.577	0.000 ± 0.577	0.000 ± 0.577
CMP Trk Iso	0.852 ± 0.029	0.907 ± 0.003	0.939 ± 0.032
CMX Trk Iso	0.906 ± 0.013	0.900 ± 0.002	1.007 ± 0.014
CMXMsKs Trk Iso	0.792 ± 0.032	0.895 ± 0.004	0.885 ± 0.036
BMU Trk Iso	0.759 ± 0.015	0.719 ± 0.003	1.054 ± 0.022
CMIOCES Trk Iso	0.323 ± 0.004	0.301 ± 0.001	1.069 ± 0.014
CMIOPES Trk Iso	0.559 ± 0.011	0.592 ± 0.002	0.946 ± 0.018
CrkTrk μ Trk Iso	0.681 ± 0.008	0.711 ± 0.001	0.947 ± 0.012

Table 8: Muon efficiencies for dataset 0i2 (periods 8-10) with and without track isolation cut.

Run Range: 0j

	data	MC	Scale Fac
CMUP	0.825 ± 0.010	0.893 ± 0.001	0.924 ± 0.011
CMU	0.000 ± 0.707	0.000 ± 0.707	0.000 ± 0.707
CMP	0.817 ± 0.020	0.916 ± 0.002	0.893 ± 0.022
CMX	0.898 ± 0.016	0.915 ± 0.002	0.981 ± 0.018
CMXMsKs	0.852 ± 0.029	0.911 ± 0.003	0.935 ± 0.032
BMU	0.773 ± 0.020	0.727 ± 0.003	1.064 ± 0.028
CMIOCES	0.303 ± 0.004	0.252 ± 0.002	1.204 ± 0.019
CMIOPES	0.560 ± 0.013	0.586 ± 0.002	0.955 ± 0.023
CrkTrk μ	0.577 ± 0.012	0.583 ± 0.002	0.990 ± 0.020
CMUP Trk Iso	0.796 ± 0.010	0.882 ± 0.001	0.903 ± 0.011
CMU Trk Iso	0.000 ± 0.707	0.000 ± 0.707	0.000 ± 0.707
CMP Trk Iso	0.796 ± 0.020	0.903 ± 0.002	0.881 ± 0.022
CMX Trk Iso	0.867 ± 0.016	0.895 ± 0.002	0.968 ± 0.018
CMXMsKs Trk Iso	0.839 ± 0.028	0.896 ± 0.003	0.937 ± 0.031
BMU Trk Iso	0.748 ± 0.021	0.720 ± 0.004	1.039 ± 0.029
CMIOCES Trk Iso	0.286 ± 0.005	0.248 ± 0.002	1.152 ± 0.021
CMIOPES Trk Iso	0.532 ± 0.013	0.579 ± 0.002	0.919 ± 0.024
CrkTrk μ Trk Iso	0.549 ± 0.011	0.572 ± 0.002	0.961 ± 0.020

Table 9: Muon efficiencies for dataset 0j (period 11-12) with and without track isolation cut.

Run Range: 0j13

	data	MC	Scale Fac
CMUP	0.839 ± 0.010	0.895 ± 0.001	0.937 ± 0.011
CMU	0.000 ± 1.000	0.000 ± 1.000	0.000 ± 1.000
CMP	0.819 ± 0.020	0.913 ± 0.003	0.897 ± 0.022
CMX	0.900 ± 0.018	0.913 ± 0.003	0.986 ± 0.020
CMXMsKs	0.815 ± 0.030	0.916 ± 0.003	0.890 ± 0.033
BMU	0.795 ± 0.025	0.696 ± 0.004	1.142 ± 0.037
CMIOCES	0.300 ± 0.005	0.253 ± 0.002	1.186 ± 0.022
CMIOPES	0.520 ± 0.019	0.521 ± 0.003	0.998 ± 0.037
CrkTrk μ	0.560 ± 0.012	0.588 ± 0.002	0.952 ± 0.021
CMUP Trk Iso	0.815 ± 0.010	0.884 ± 0.001	0.922 ± 0.011
CMU Trk Iso	0.000 ± 1.000	0.000 ± 1.000	0.000 ± 1.000
CMP Trk Iso	0.804 ± 0.020	0.901 ± 0.003	0.892 ± 0.022
CMX Trk Iso	0.886 ± 0.018	0.898 ± 0.003	0.987 ± 0.020
CMXMsKs Trk Iso	0.794 ± 0.029	0.904 ± 0.003	0.878 ± 0.032
BMU Trk Iso	0.773 ± 0.025	0.692 ± 0.004	1.117 ± 0.037
CMIOCES Trk Iso	0.290 ± 0.005	0.249 ± 0.002	1.165 ± 0.022
CMIOPES Trk Iso	0.496 ± 0.018	0.517 ± 0.003	0.959 ± 0.035
CrkTrk μ Trk Iso	0.542 ± 0.012	0.578 ± 0.002	0.938 ± 0.021

Table 10: Muon efficiencies for dataset 0j13 (period 13) with and without track isolation cut.

Run Range: 0k

	data	MC	Scale Fac
CMUP	0.786 ± 0.004	0.889 ± 0.001	0.884 ± 0.004
CMU	0.000 ± 1.000	0.000 ± 1.000	0.000 ± 1.000
CMP	0.798 ± 0.008	0.911 ± 0.002	0.876 ± 0.009
CMX	0.886 ± 0.007	0.906 ± 0.002	0.978 ± 0.008
CMXMsKs	0.827 ± 0.010	0.907 ± 0.002	0.912 ± 0.012
BMU	0.775 ± 0.009	0.705 ± 0.003	1.100 ± 0.013
CMIOCES	0.301 ± 0.002	0.252 ± 0.001	1.196 ± 0.011
CMIOPES	0.565 ± 0.006	0.582 ± 0.002	0.970 ± 0.010
CrkTrk μ	0.565 ± 0.004	0.589 ± 0.001	0.959 ± 0.008
CMUP Trk Iso	0.764 ± 0.004	0.875 ± 0.001	0.873 ± 0.004
CMU Trk Iso	0.000 ± 1.000	0.000 ± 1.000	0.000 ± 1.000
CMP Trk Iso	0.772 ± 0.008	0.897 ± 0.002	0.861 ± 0.009
CMX Trk Iso	0.856 ± 0.007	0.887 ± 0.002	0.965 ± 0.008
CMXMsKs Trk Iso	0.794 ± 0.010	0.891 ± 0.002	0.892 ± 0.012
BMU Trk Iso	0.760 ± 0.009	0.696 ± 0.003	1.092 ± 0.013
CMIOCES Trk Iso	0.291 ± 0.002	0.248 ± 0.001	1.174 ± 0.011
CMIOPES Trk Iso	0.558 ± 0.006	0.574 ± 0.002	0.973 ± 0.011
CrkTrk μ Trk Iso	0.545 ± 0.004	0.577 ± 0.001	0.945 ± 0.008

Table 11: Muon efficiencies for dataset 0k (periods 14-25) with and without track isolation cut.

Run Range: 0d

	data	MC	Scale Fac
LBE ($L > 0.9$)	0.911 ± 0.003	0.900 ± 0.001	1.012 ± 0.004
TCE	0.876 ± 0.005	0.861 ± 0.002	1.017 ± 0.006
LCE	0.042 ± 0.003	0.045 ± 0.001	0.933 ± 0.078
PHXTrk	0.863 ± 0.004	0.865 ± 0.001	0.998 ± 0.005
PHXPEM	0.850 ± 0.005	0.894 ± 0.001	0.951 ± 0.006
PEM	0.808 ± 0.009	0.857 ± 0.001	0.943 ± 0.011
CrkTrk e	0.792 ± 0.013	0.834 ± 0.002	0.950 ± 0.016
PESTrk	0.369 ± 0.005	0.404 ± 0.001	0.913 ± 0.013
LBE ($L > 0.9$) Trk Iso	0.860 ± 0.003	0.858 ± 0.001	1.002 ± 0.004
TCE Trk Iso	0.820 ± 0.005	0.813 ± 0.002	1.009 ± 0.006
LCE Trk Iso	0.032 ± 0.002	0.036 ± 0.001	0.889 ± 0.064
CrkTrk e Trk Iso	0.766 ± 0.012	0.821 ± 0.002	0.933 ± 0.015
PESTrk Trk Iso	0.369 ± 0.005	0.404 ± 0.001	0.913 ± 0.013
PHXTrk Trk Iso	0.863 ± 0.004	0.865 ± 0.001	0.998 ± 0.005
PHXPEM Trk Iso	0.850 ± 0.005	0.894 ± 0.001	0.951 ± 0.006
PEM Trk Iso	0.808 ± 0.009	0.857 ± 0.001	0.943 ± 0.011

Table 12: Electron efficiencies for dataset 0d (period 0) with and without track isolation cut. No track isolation cut is applied to the PHX category.

Run Range: 0h

	data	MC	Scale Fac
LBE ($L > 0.9$)	0.902 ± 0.003	0.901 ± 0.001	1.001 ± 0.003
TCE	0.869 ± 0.004	0.862 ± 0.001	1.008 ± 0.004
LCE	0.039 ± 0.002	0.042 ± 0.001	0.944 ± 0.041
PHXTrk	0.880 ± 0.004	0.874 ± 0.001	1.007 ± 0.004
PHXPEM	0.839 ± 0.004	0.882 ± 0.001	0.953 ± 0.005
PEM	0.762 ± 0.008	0.832 ± 0.001	0.916 ± 0.010
CrkTrk e	0.855 ± 0.013	0.864 ± 0.002	0.989 ± 0.016
PESTrk	0.360 ± 0.005	0.379 ± 0.001	0.949 ± 0.013
LBE ($L > 0.9$) Trk Iso	0.844 ± 0.003	0.855 ± 0.001	0.988 ± 0.004
TCE Trk Iso	0.806 ± 0.004	0.811 ± 0.001	0.994 ± 0.005
LCE Trk Iso	0.029 ± 0.002	0.033 ± 0.001	0.862 ± 0.049
CrkTrk e Trk Iso	0.816 ± 0.011	0.847 ± 0.002	0.963 ± 0.014
PESTrk Trk Iso	0.360 ± 0.005	0.379 ± 0.001	0.949 ± 0.013
PHXTrk Trk Iso	0.880 ± 0.004	0.874 ± 0.001	1.007 ± 0.004
PHXPEM Trk Iso	0.839 ± 0.004	0.882 ± 0.001	0.953 ± 0.005
PEM Trk Iso	0.762 ± 0.008	0.832 ± 0.001	0.916 ± 0.010

Table 13: Electron efficiencies for dataset 0h (periods 1-4) with and without track isolation cut. No track isolation cut is applied to the PHX category.

Run Range: 0i1

	data	MC	Scale Fac
LBE ($L > 0.9$)	0.894 ± 0.004	0.897 ± 0.001	0.996 ± 0.004
TCE	0.861 ± 0.005	0.856 ± 0.001	1.006 ± 0.006
LCE	0.042 ± 0.002	0.043 ± 0.001	0.971 ± 0.052
PHXTrk	0.885 ± 0.004	0.869 ± 0.001	1.018 ± 0.005
PHXPEM	0.826 ± 0.005	0.876 ± 0.001	0.944 ± 0.006
PEM	0.742 ± 0.012	0.813 ± 0.002	0.911 ± 0.015
CrkTrk e	0.821 ± 0.016	0.858 ± 0.002	0.957 ± 0.019
PESTrk	0.364 ± 0.006	0.374 ± 0.001	0.974 ± 0.017
LBE ($L > 0.9$) Trk Iso	0.830 ± 0.004	0.849 ± 0.001	0.978 ± 0.005
TCE Trk Iso	0.794 ± 0.005	0.804 ± 0.001	0.987 ± 0.007
LCE Trk Iso	0.027 ± 0.002	0.033 ± 0.001	0.840 ± 0.080
CrkTrk e Trk Iso	0.788 ± 0.014	0.840 ± 0.003	0.936 ± 0.017
PESTrk Trk Iso	0.364 ± 0.006	0.374 ± 0.001	0.974 ± 0.017
PHXTrk Trk Iso	0.885 ± 0.004	0.869 ± 0.001	1.018 ± 0.005
PHXPEM Trk Iso	0.826 ± 0.005	0.876 ± 0.001	0.944 ± 0.006
PEM Trk Iso	0.742 ± 0.012	0.813 ± 0.002	0.911 ± 0.015

Table 14: Electron efficiencies for dataset 0i1 (periods 5-7) with and without track isolation cut. No track isolation cut is applied to the PHX category.

Run Range: 0i2

	data	MC	Scale Fac
LBE ($L > 0.9$)	0.888 ± 0.003	0.896 ± 0.001	0.992 ± 0.003
TCE	0.858 ± 0.003	0.857 ± 0.001	1.001 ± 0.004
LCE	0.041 ± 0.002	0.043 ± 0.001	0.961 ± 0.044
PHXTrk	0.878 ± 0.003	0.878 ± 0.001	1.001 ± 0.004
PHXPEM	0.812 ± 0.004	0.873 ± 0.001	0.931 ± 0.004
PEM	0.698 ± 0.008	0.799 ± 0.001	0.875 ± 0.010
CrkTrk e	0.809 ± 0.012	0.851 ± 0.001	0.948 ± 0.014
PESTrk	0.352 ± 0.004	0.370 ± 0.001	0.947 ± 0.012
LBE ($L > 0.9$) Trk Iso	0.818 ± 0.003	0.846 ± 0.001	0.966 ± 0.004
TCE Trk Iso	0.782 ± 0.004	0.803 ± 0.001	0.974 ± 0.005
LCE Trk Iso	0.027 ± 0.001	0.033 ± 0.001	0.836 ± 0.046
CrkTrk e Trk Iso	0.762 ± 0.010	0.836 ± 0.002	0.911 ± 0.013
PESTrk Trk Iso	0.352 ± 0.004	0.370 ± 0.001	0.947 ± 0.012
PHXTrk Trk Iso	0.878 ± 0.003	0.878 ± 0.001	1.001 ± 0.004
PHXPEM Trk Iso	0.812 ± 0.004	0.873 ± 0.001	0.931 ± 0.004
PEM Trk Iso	0.698 ± 0.008	0.799 ± 0.001	0.875 ± 0.010

Table 15: Electron efficiencies for dataset 0i2 (periods 8-10) with and without track isolation cut. No track isolation cut is applied to the PHX category.

Run Range: 0j

	data	MC	Scale Fac
LBE ($L > 0.9$)	0.881 ± 0.003	0.888 ± 0.001	0.993 ± 0.004
TCE	0.848 ± 0.004	0.851 ± 0.001	0.997 ± 0.005
LCE	0.045 ± 0.002	0.044 ± 0.001	1.028 ± 0.050
PHXTrk	0.878 ± 0.004	0.879 ± 0.001	0.999 ± 0.004
PHXPEM	0.804 ± 0.005	0.858 ± 0.001	0.939 ± 0.005
PEM	0.660 ± 0.010	0.759 ± 0.001	0.870 ± 0.013
CrkTrk e	0.848 ± 0.018	0.847 ± 0.002	1.002 ± 0.021
PESTrk	0.351 ± 0.005	0.362 ± 0.001	0.966 ± 0.015
LBE ($L > 0.9$) Trk Iso	0.800 ± 0.004	0.834 ± 0.001	0.960 ± 0.005
TCE Trk Iso	0.763 ± 0.005	0.792 ± 0.001	0.965 ± 0.006
LCE Trk Iso	0.029 ± 0.003	0.031 ± 0.001	0.919 ± 0.099
CrkTrk e Trk Iso	0.788 ± 0.014	0.827 ± 0.002	0.953 ± 0.017
PESTrk Trk Iso	0.351 ± 0.005	0.362 ± 0.001	0.966 ± 0.015
PHXTrk Trk Iso	0.878 ± 0.004	0.879 ± 0.001	0.999 ± 0.004
PHXPEM Trk Iso	0.804 ± 0.005	0.858 ± 0.001	0.939 ± 0.005
PEM Trk Iso	0.660 ± 0.010	0.759 ± 0.001	0.870 ± 0.013

Table 16: Electron efficiencies for dataset 0j (period 11-12) with and without track isolation cut. No track isolation cut is applied to the PHX category.

Run Range: 0j13

	data	MC	Scale Fac
LBE ($L > 0.9$)	0.886 ± 0.004	0.891 ± 0.001	0.994 ± 0.005
TCE	0.851 ± 0.005	0.855 ± 0.002	0.995 ± 0.006
LCE	0.047 ± 0.002	0.042 ± 0.001	1.119 ± 0.057
PHXTrk	0.860 ± 0.005	0.857 ± 0.001	1.004 ± 0.006
PHXPEM	0.808 ± 0.006	0.863 ± 0.001	0.936 ± 0.007
PEM	0.688 ± 0.010	0.790 ± 0.002	0.871 ± 0.013
CrkTrk e	0.817 ± 0.018	0.846 ± 0.002	0.966 ± 0.021
PESTrk	0.261 ± 0.006	0.283 ± 0.001	0.922 ± 0.021
LBE ($L > 0.9$) Trk Iso	0.816 ± 0.004	0.840 ± 0.001	0.971 ± 0.005
TCE Trk Iso	0.776 ± 0.005	0.799 ± 0.002	0.971 ± 0.007
LCE Trk Iso	0.034 ± 0.002	0.031 ± 0.001	1.097 ± 0.084
CrkTrk e Trk Iso	0.771 ± 0.015	0.830 ± 0.002	0.929 ± 0.018
PESTrk Trk Iso	0.261 ± 0.006	0.283 ± 0.001	0.922 ± 0.021
PHXTrk Trk Iso	0.860 ± 0.005	0.857 ± 0.001	1.004 ± 0.006
PHXPEM Trk Iso	0.808 ± 0.006	0.863 ± 0.001	0.936 ± 0.007
PEM Trk Iso	0.688 ± 0.010	0.790 ± 0.002	0.871 ± 0.013

Table 17: Electron efficiencies for dataset 0j13 (period 13) with and without track isolation cut. No track isolation cut is applied to the PHX category.

Run Range: 0k

	data	MC	Scale Fac
LBE ($L > 0.9$)	0.875 ± 0.001	0.883 ± 0.001	0.991 ± 0.001
TCE	0.845 ± 0.002	0.847 ± 0.002	0.997 ± 0.003
LCE	0.046 ± 0.001	0.043 ± 0.001	1.066 ± 0.046
PHXTrk	0.888 ± 0.001	0.866 ± 0.001	1.026 ± 0.002
PHXPEM	0.771 ± 0.002	0.847 ± 0.001	0.911 ± 0.002
PEM	0.608 ± 0.004	0.733 ± 0.001	0.829 ± 0.005
CrkTrk e	0.804 ± 0.006	0.834 ± 0.001	0.964 ± 0.007
PESTrk	0.319 ± 0.002	0.352 ± 0.001	0.907 ± 0.006
LBE ($L > 0.9$) Trk Iso	0.797 ± 0.001	0.827 ± 0.001	0.963 ± 0.002
TCE Trk Iso	0.760 ± 0.002	0.786 ± 0.002	0.967 ± 0.003
LCE Trk Iso	0.028 ± 0.001	0.031 ± 0.001	0.916 ± 0.055
CrkTrk e Trk Iso	0.756 ± 0.005	0.814 ± 0.002	0.929 ± 0.007
PESTrk Trk Iso	0.319 ± 0.002	0.352 ± 0.001	0.907 ± 0.006
PHXTrk Trk Iso	0.888 ± 0.001	0.866 ± 0.001	1.026 ± 0.002
PHXPEM Trk Iso	0.771 ± 0.002	0.847 ± 0.001	0.911 ± 0.002
PEM Trk Iso	0.608 ± 0.004	0.733 ± 0.001	0.829 ± 0.005

Table 18: Electron efficiencies for dataset 0k (periods 14-25) with and without track isolation cut. No track isolation cut is applied to the PHX category.

2.3 Fake Rate Measurements

The lepton fake rates (*ie.* the probability for jets to fake a lepton) are measured in the jet samples and corrected for the small contribution of real isolated leptons from electroweak processes, as described in detail in [5]. This probability is measured as a function of the p_T of the object and results are used to derive a data-driven modeling of the W+jets background as described in Section 4.3.

As a brief reminder we repeat here the master formula for the calculation of such fake rates for a generic category i:

$$f_i = \frac{N_i(\text{IdentifiedLeptons}) - \sum_{j \in \text{EWK}} N_{ij}(\text{IdentifiedLeptons})}{N_i(\text{DenominatorObjects}) - \sum_{j \in \text{EWK}} N_{ij}(\text{DenominatorObjects})} \tag{2}$$

As with the scale factor calculation, the procedure relies on generic Central/Forward muon *Denominator Objects* which are then made category specific by applying the fiduciality requirements. In order to perform the electroweak subtraction we normalize the MC simulation to the data effectively analyzed in each of the jet 20, 50, 70, and 100 triggered samples. Table 19 reports the effective (*ie.* corrected for trigger prescale factors) luminosities of such samples.

Sample	Effective Luminosity (nb)
Jet 20	$3154 \cdot 1.019$
Jet 50	$69756 \cdot 1.019$
Jet 70	$592747 \cdot 1.019$
Jet 100	$4738208 \cdot 1.019$

Table 19: Effective luminosity for each of the samples after trigger prescale correction for em_nosi goodrun list. There are used for the electroweak subtraction.

Results of the fake rate calculations are shown in Figures 8 to 9. Fake rate calculations are averaged over all the data available to maximize the statistics available. We assign a systematic based on the scatter of the results between the jet samples: an uncorrelated systematic, α, is added to each jet sample measurement until all the points are consistent within 1σ of the new mean, where the mean and the σ from the mean are calculated using $\sqrt{\text{stat err} + \alpha}$ as the uncertainty on each point. This result is shown as a light gray band in the figures, while the dark band is the result of the statistical error only.

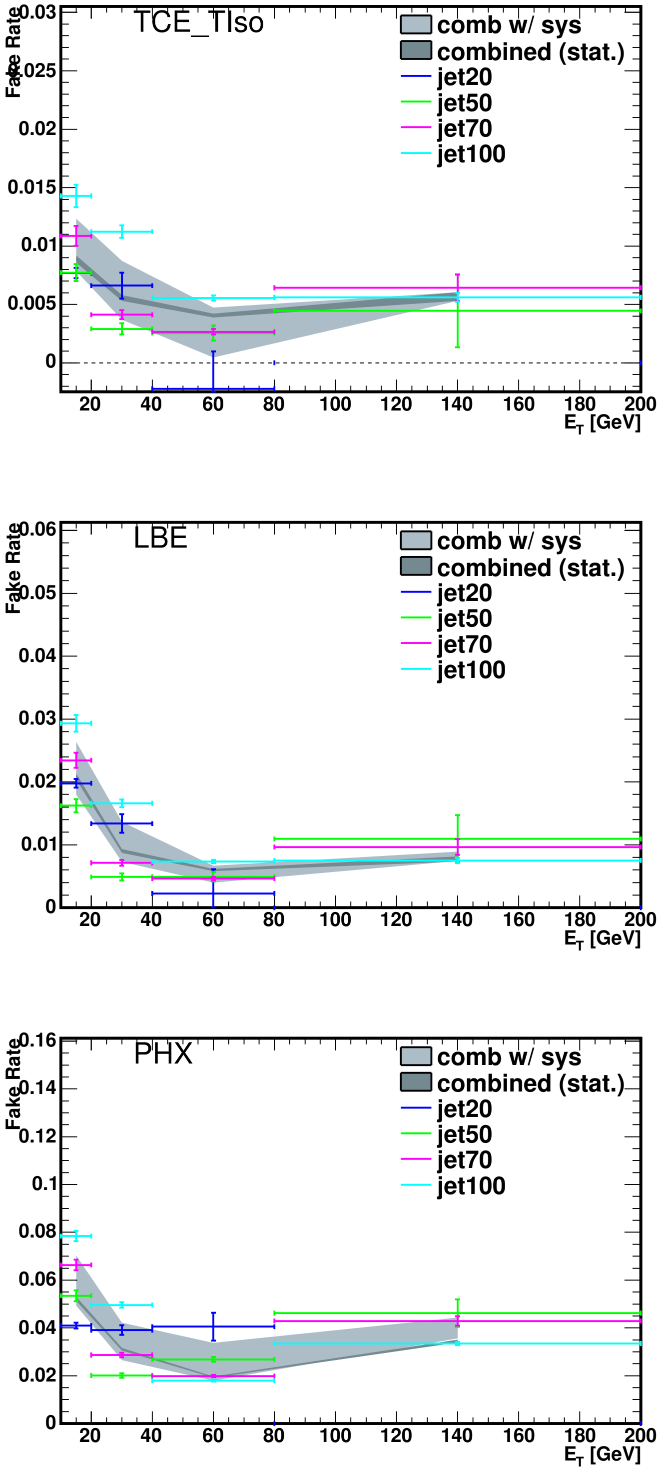

Figure 8: Fake rates for electrons. PHX and LBE have no track isolation requirement.

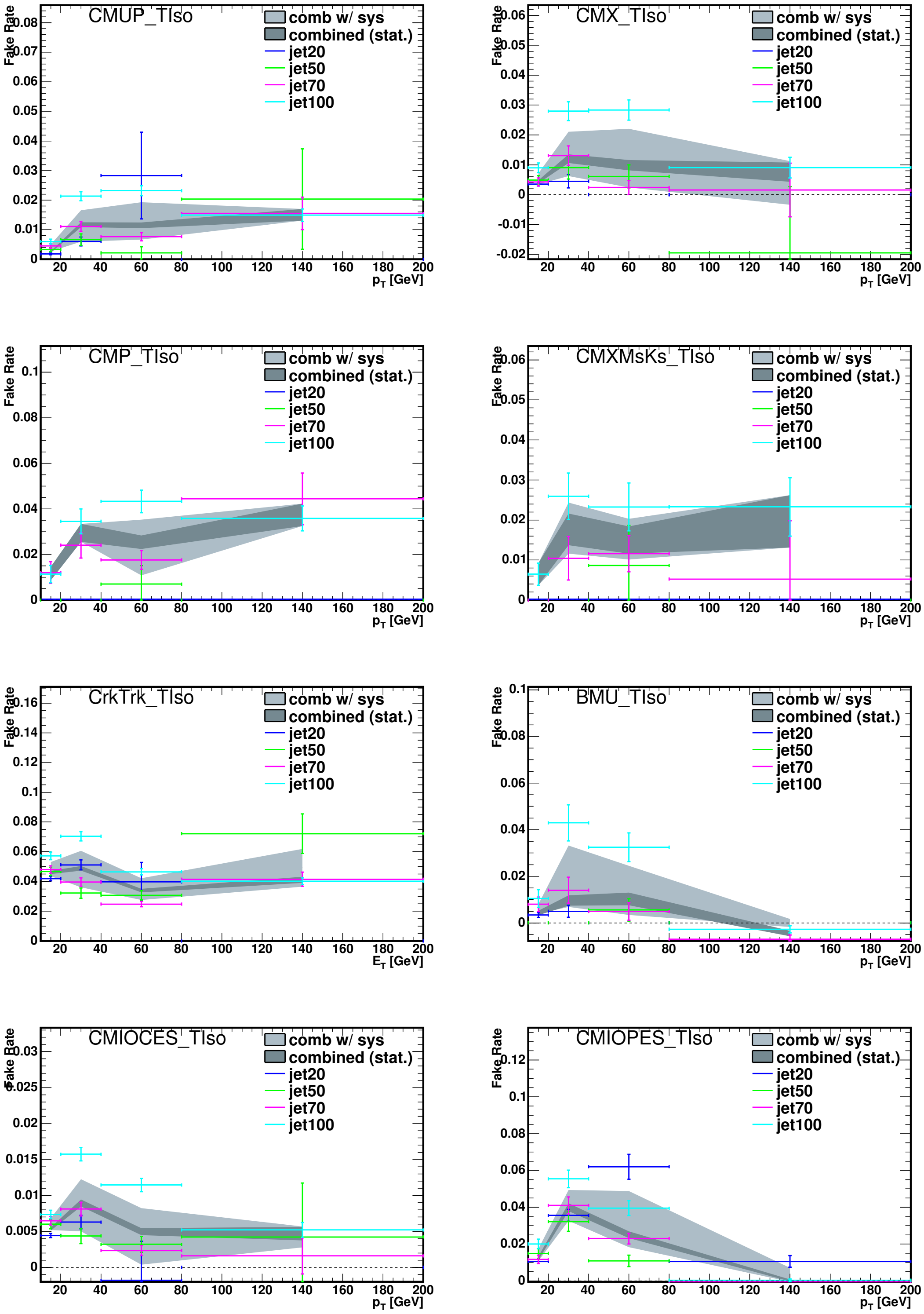

Figure 9: Fake rates for muons.

2.4 DY Cross Section Measurements

In this section we measure the DY cross-section by comparing reconstructed $p\bar{p} \to Z \to l^+l^-(l = e, \mu)$ events in data with Pythia MC simulations. This provides a sensible measurement of the luminosity of the data sample, and is also probably the strongest check of the definition of new categories. Agreement on the cross section value among different dilepton categories tells us that we're correctly modeling their behaviour in our simulation, and the average cross section value of all dilepton categories tells us if the observed yields are in agreement with the luminosity measurement. We select Z candidates requiring two opposite sign leptons (e or μ, belonging to the defined categories) with invariant mass $76 < m(ll) < 106$ GeV, and require $\not{E}_T < 15$ GeV for $e\mu$ events and $\not{E}_T < 25$ GeV for all other combinations. All base analysis requirements are also applied: good run list selection, $\Delta z_0(ll) < 4$ cm, cosmic veto, and requiring a trigger lepton coherent with trigger bits (for data).

In Table 20 we show the following results for each dilepton category: the number of events predicted and observed, their ratio, the resulting DY cross section with its statistical error, and the residual between the measured value and the weighted average of all the categories (systematic uncertainties are not accounted for).

Results show a reasonable agreement for all categories within few percent. The only visible shift is for CMXMsKs, for which the statistics are quite low. Considering that systematic uncertainties (such as trigger efficiencies, SF errors, etc.) are not accounted for in these errors, we think this effect is not worrisome. The PHX category is known to have a lower measurement (and for this reason it's excluded in the calculation of the weighted mean) since in this low $\not{E}_T$ region we do not model correctly the turn-on of the MET_PEM trigger (by which these events are triggered); this effect is not present in the signal region since the $\not{E}_T$ cut puts events in the plateau of the turn-on curve.

Category	DY MC	Data	Data/MC	DY x-sec (pb) $\pm$ stat.	Residual [σ]
LBE LBE	111738.0	110392	0.99	248.28 $\pm$ 0.78	3.8
LBE PHX	104597.5	101760	0.97	244.49 $\pm$ 0.79	-1.1
PHX PHX	2634.5	1010	0.38	96.34 $\pm$ 3.04	-49.1
LBE CMUP	36.9	29	0.79	197.33 $\pm$ 37.65	-1.3
LBE CMP	4.9	5	1.03	257.88 $\pm$ 121.31	0.1
LBE CMX	15.6	28	1.79	450.35 $\pm$ 90.60	2.3
LBE CMXMsKs	3.4	1	0.29	73.44 $\pm$ 74.66	-2.3
LBE BMU	5.1	4	0.78	196.86 $\pm$ 101.24	-0.5
LBE CMIOCES	29.9	27	0.90	226.84 $\pm$ 45.22	-0.4
LBE CMIOPES	12.9	12	0.93	232.98 $\pm$ 69.48	-0.2
PHX CMUP	30.2	34	1.13	283.14 $\pm$ 50.38	0.8
PHX CMP	5.6	5	0.89	224.74 $\pm$ 104.90	-0.2
PHX CMX	13.7	15	1.10	275.31 $\pm$ 73.78	0.4
PHX CMXMsKs	2.6	2	0.76	191.45 $\pm$ 139.08	-0.4
PHX BMU	4.1	6	1.47	369.31 $\pm$ 158.25	0.8
PHX CMIOCES	19.6	31	1.58	396.75 $\pm$ 75.14	2.0
PHX CMIOPES	8.9	7	0.79	197.33 $\pm$ 76.27	-0.6
CMUP CMUP	21664.8	21283	0.98	246.88 $\pm$ 1.74	0.9
CMUP CMP	6092.4	6266	1.03	258.47 $\pm$ 3.39	3.9
CMUP CMX	21580.5	21182	0.98	246.66 $\pm$ 1.75	0.8
CMUP CMXMsKs	4166.0	3839	0.92	231.58 $\pm$ 3.86	-3.6
CMUP BMU	7093.3	6820	0.96	241.62 $\pm$ 3.01	-1.2
CMUP CMIOCES	31772.4	32528	1.02	257.28 $\pm$ 1.48	8.1
CMUP CMIOPES	14571.4	14061	0.96	242.50 $\pm$ 2.10	-1.3
CMX CMP	3087.4	3111	1.01	253.23 $\pm$ 4.73	1.7
CMX CMX	6252.9	5947	0.95	239.01 $\pm$ 3.21	-2.0
CMX CMXMsKs	1543.9	1394	0.90	226.90 $\pm$ 6.29	-2.9
CMX BMU	3628.3	3239	0.89	224.34 $\pm$ 4.06	-5.2
CMX CMIOCES	16420.5	15825	0.96	242.19 $\pm$ 2.00	-1.6
CMX CMIOPES	7212.5	6762	0.94	235.61 $\pm$ 2.95	-3.3
CMXMsKs CMP	906.5	818	0.90	226.77 $\pm$ 8.17	-2.3
CMXMsKs CMXMsKs	356.3	376	1.06	265.18 $\pm$ 14.18	1.4
CMXMsKs BMU	618.5	597	0.97	242.59 $\pm$ 10.22	-0.3
CMXMsKs CMIOCES	2337.2	2850	1.22	306.44 $\pm$ 6.01	10.2
CMXMsKs CMIOPES	983.8	999	1.02	255.20 $\pm$ 8.30	1.2
CMP CMP	694.8	739	1.06	267.30 $\pm$ 10.16	2.2
CMP BMU	757.0	770	1.02	255.61 $\pm$ 9.47	1.1
CMP CMIOCES	3241.5	3583	1.11	277.78 $\pm$ 4.82	6.7
CMP CMIOPES	1530.1	1588	1.04	260.81 $\pm$ 6.72	2.3
LBE CrkTrk	50127.4	46984	0.94	235.55 $\pm$ 1.12	-8.7
PHX CrkTrk	11337.4	9402	0.83	208.41 $\pm$ 2.18	-17.0
CMUP CrkTrk	26224.7	25162	0.96	241.12 $\pm$ 1.56	-2.7
CMP CrkTrk	2698.8	2916	1.08	271.53 $\pm$ 5.18	5.1
CMX CrkTrk	13343.7	12259	0.92	230.88 $\pm$ 2.15	-6.7
CMXMsKs CrkTrk	1971.0	2037	1.03	259.72 $\pm$ 5.94	2.4

Table 20: Drell-Yan cross section measurement for each dilepton category. See text for detailed explanation.

3 Event Selection

In this section, we describe our candidate selection. This includes the triggers, reconstruction of low-level objects, and our high-level cuts. The general strategy is to use cuts to supress the Drell-Yan background and then rely on matrix element calculations and neural networks to discriminate against the other backgrounds.

3.1 Datasets

We use the following trigger paths:

- ELECTRON_CENTRAL_18
- MUON_CMUP18
- MUON_CMX18
- MET_PEM
- MUON_CMP18_PHI_GAP

The corresponding datasets are bhel0d/0h/0i/0j/0k/0m (electrons), bhmu0d/0h/0i/0j/0k/0m (muons), and bpel0d/0h/0i/0j/0k/0m (MET+PEM). We use the following v31 good run lists slightly modified to recover portions of runs that contain bad beamline data (see CDF Note 9195):

- EM_NOSI
- EM_CMUP_NOSI
- EM_MU_NOSI_CMXIGNORED
- EM_SI
- EM_CMUP_SI
- EM_MU_SI_CMXIGNORED
- EM_MU_SI_BMU

We also use separate good run lists for 0n dataset, which consists of the reprocessed silicon data for runs 262865-263615 (from period 18) as documented on http://www-cdf.fnal.gov/internal/dqm/goodrun/p18Sirecovered/goodp18Si.html. The luminosity of each good run list is shown in Table 3.1.

Good run list	$\mathcal{L}$ (fb^{-1})
EM_NOSI	5.3
EM_CMUP_NOSI	5.2
EM_MU_NOSI_CMXIGNORED	5.1
EM_SI	5.1
EM_CMUP_SI	5.0
EM_MU_SI_CMXIGNORED	4.9

Table 21: Luminosity for each of the good run lists (v31) used in this analysis. These have been scaled by the factor of 1.019.

3.2 Low-level Objects

The muon categories introduced in the previous version of this analysis (CMP, CMU, CMXMsKs, and BMU) are described in CDF 9697 [4]. All previous lepton categories are identically defined as in CDF 8719, except for the changes mentioned in CDF 9697. For the CMIOPES category, there are also several additional requirements to those in CDF 8719, designed to veto events with a CMIOPES muon with poorly measured curvature. These cuts were also made in previous versions of the analysis (CDF 9685) and are as follows:

- Number of SVX hits > 3
- Curvature significance > 12 based on error from the covariance matrix

3.3 High-level Cuts

Except for the requirement on the dilepton invariant mass, m_{ll}, we use the same high-level cuts as in CDF 8719 to define a "base region". The m_{ll} cut was lowered from 25 GeV to 16 GeV as in CDF 8700.

4 Sample Modeling

In this section we describe how we model the selected sample. Section 4.1 describes the Monte Carlo samples used, Section 4.2 describes the data based corrections made to the Monte Carlo, and Section 4.3 describes the modeling of the W+jets background.

4.1 Monte Carlo Samples

The acceptance and several of the backgrounds are estimated using Monte Carlo samples. Table 22 lists all the samples used. Here we provide some brief notes on the samples.

WW

As WW is by far the dominant background to this analysis, we take special care using a NLO Monte Carlo generator, MC@NLO. Unfortunately, our current NLO WW sample is only generated with run ranges through period 7. Although new samples have been generated and will be incorporated in the next update, they are not available for the current analysis. As a temporary work-around, we use the NLO WW sample generated through period 7 in our analysis and apply a scale factor determined from the ratio of yields obtained in a region of interest for two Pythia WW samples: one covering all periods from 0 to 25 and another that spans only periods 0 to 7. The resulting scale factors are given in Appendix A, with half of the scale factor assigned as a systematic uncertainty. In both the Pythia and MC@NLO samples both of the W bosons are required to decay into either an electron, muon, or tau.

Heavy dibosons: WZ, ZZ

The WZ and ZZ samples used are all Pythia based. The WZ samples are generated with a forced decay of the Z into a lepton pair (electrons, muons, or taus) while the W is allowed to decay inclusively. The events are filtered using the generator-level information, requiring two leptons with p_T greater than 1 GeV. The ZZ sample is generated having both Z's decay inclusively and then filtering for two leptons with p_T greater than 1 GeV. The ZZ Monte Carlo includes γ^* component with M_{ll} >15 GeV/c^2.

$W\gamma$

The $W\gamma$ Monte Carlo used for this analysis uses the Baur matrix element generator [7] which is passed to Pythia via the LesHouchesModule. Because of indecision over the proper settings for the Pythia QED radiation settings the Gen5 sample has the QED turned off and the Gen6 sample has it turned on. The two settings correspond to slightly over estimating the lepton id efficiency and double counting at order $\alpha_{\rm EM}$ the hard radiation, respectively. Neither of these effects should be significant for this analysis, relative to the conversion finding efficiency and K-factor uncertainties. It should also be noted that the $W\gamma$ cross-section has been measured at CDF [8, 9] and DØ.

$Z\gamma$

The $Z\gamma$ Monte Carlo, which is used only for the trilepton searches, uses the Baur matrix element generator [7] which is passed to Pythia via the LesHouchesModule. The $Z\gamma$ samples are generated with the QED turned on. The $Z\gamma$ cross-section has also been measured at CDF [8, 10] and DØ.

$t\bar{t}$

The $t\bar{t}$ sample is also Pythia based, and is generated for a top-quark mass of 172.5 GeV.

$Z \to \tau\tau$

The $Z \to \tau\tau$ sample is generated with a 10 GeV/c^2 minimum dilepton mass, uses Gen5 for the full run range and has no minbias in the events.

$Z \to ee$ **and** *$Z \to \mu\mu$*

The $Z \to ee$ and $Z \to \mu\mu$ samples are generated with a 20 GeV/c^2 minimum dilepton mass. They use Gen5 for the 0d run range and Gen6 for the 0h/0i/0j/0k/0m run ranges. They have a full luminosity dependent addition of minbias in the events.

4.2 Corrections to the Monte Carlo

Luminosity scaling, data-base corrections, and measured trigger efficiencies are applied to the Monte Carlo samples separately for each of the lepton pair categories. The master formula used is:

$$\frac{\sigma \times \mathcal{B} \times \epsilon_{\text{filter}} \times \epsilon_i^{\text{trg}} \times s_i^{\text{lep}} \times \epsilon_{\text{vtx}} \times \mathcal{L}_i}{N_i^{\text{gen}}(|Z_0| < 60 \text{ cm})} \tag{3}$$

where

mode	Period	Stntuple	$\sigma \times \mathcal{B}$ (pb)	K-factor[a]	Filter Eff
$W\gamma \to e\nu\gamma$	0-25	re0s6d,re0s6h re0s6e,re0s6i	32.38×0.5	1.36	1.0
$W\gamma \to \mu\nu\gamma$	0-25	re0s6f,re0s6j re0s6g,re0s6k	32.38×0.5	1.36	1.0
$W\gamma \to \tau\nu\gamma$	0-11	re0s6a	13.6	1.36	1.0
WW	0-25	we0s5d,we0sbd,we0sgd we0skd,we0snd,we0saf we0sgf	11.66×0.1027	1.0	1.0
$WW@NLO$	0-7	we0sfd,we0sed	11.66×0.1027	1.0	1.0
WZ	0-25	we0s6d,we0scd,we0shd we0sld,we0sod,we0sbf we0shf	3.46×0.101	1.0	0.754
ZZ	0-25	we0s7d,we0sdd,we0sid we0smd,we0spd,we0scf we0sif	1.511	1.0	0.233
$t\bar{t}$	0-25	dt0sh0,dt0sh1	7.884×0.1027	1.0	1.0
$Z \to ee$	0-25	ze1s6d[b],ze1sad,ze0scd ze0sdd,ze0sed,ze0see ze0seh,ze0sej,ze0sei	355[c]	1.38	1.0
$Z \to \mu\mu$	0-25	ze1s6m[b],ze1s9m,ze0sbm ze0scm,ze0sdm,ze0sem ze0sfm,ze0sgm,ze0sim	355[c]	1.38	1.0
$Z \to \tau\tau$	0-25	ze0s8t,ze0sat[d] ze0sbt,ze0sct	355[c]	1.38	1.0
$Z \to ee$	0-11	zx0see	920[e]	1.38	0.0156
$Z \to \mu\mu$	0-11	zx0sem	920[e]	1.38	0.0156
$Z \to \tau\tau$	0-11	zx0set	1272[f]	1.38	0.00713
$Z\gamma \to ee\gamma$	0-11	re0s33	10.33	1.36	1.0
$Z\gamma \to \mu\mu\gamma$	0-11	re0s34	10.33	1.36	1.0
$Z\gamma \to \tau\tau\gamma$	0-11	re0s37	10.33	1.36	1.0

[a] If cross section is NLO, then K-factor is one.
[b] Gen5 tarball for period 0.
[c] $m_{ll} > 20$.
[d] $m_{ll} > 30$ ($\sigma = 237$) for this sample only.
[e] $10 < m_{ll} < 20$.
[f] $m_{ll} > 10$.

Table 22: Monte Carlo samples used in this analysis.

$M_H(GeV^2)$	Period	Stntuple	σ (pb)	BR ($H \to WW$)	BR($WW \to \ell\nu\ell\nu$)
110	0-25	dhgs4a,dhgs6a	1.413	0.0441	0.1027
115	0-25	dhgs4l,dhgs6l	1.240	0.0797	0.1027
120	0-25	dhgs4b,dhgs6b	1.093	0.1320	0.1027
125	0-25	dhgs4m,dhgs6m	0.967	0.2018	0.1027
130	0-25	dhgs4c,dhgs6c	0.858	0.2869	0.1027
135	0-25	dhgs4n,dhgs6n	0.764	0.3828	0.1027
140	0-25	dhgs4d,dhgs6d	0.682	0.4833	0.1027
145	0-25	dhgs4o,dhgs6o	0.611	0.5833	0.1027
150	0-25	dhgs4e,dhgs6e	0.548	0.6817	0.1027
155	0-25	dhgs4p,dhgs6p	0.492	0.7823	0.1027
160	0-25	dhgs4f,dhgs6f	0.439	0.9011	0.1027
165	0-25	dhgs4q,dhgs6q	0.389	0.9610	0.1027
170	0-25	dhgs4g,dhgs6g	0.349	0.9653	0.1027
175	0-25	dhgs4r,dhgs6r	0.314	0.9594	0.1027
180	0-25	dhgs4h,dhgs6h	0.283	0.9345	0.1027
185	0-25	dhgs4s,dhgs6s	0.255	0.8379	0.1027
190	0-25	dhgs4i,dhgs6i	0.231	0.7761	0.1027
195	0-25	dhgs4t,dhgs6t	0.210	0.7495	0.1027
200	0-25	dhgs4j,dhgs6j	0.192	0.7347	0.1027

Table 23: Higgs production via gluon fusion. The production cross section is taken from a NNLL calculation.

$M_H(GeV^2)$	Period	Stntuple	σ (pb)	BR ($H \rightarrow WW$)	BR($WW \rightarrow \ell\nu\ell\nu$)
110	0-25	vhgs1a,vhgs3a,vhgs6a	0.0871	0.0441	0.1027
115	0-25	vhgs2l,vhgs3l,vhgs6l	0.0791	0.0797	0.1027
120	0-25	vhgs1b,vhgs3b,vhgs6b	0.0717	0.1320	0.1027
125	0-25	vhgs2m,vhgs3m,vhgs6m	0.0674	0.2018	0.1027
130	0-25	vhgs1c,vhgs3c,vhgs6c	0.0625	0.2869	0.1027
135	0-25	vhgs2n,vhgs3n,vhgs6n	0.0577	0.3828	0.1027
140	0-25	vhgs1d,vhgs3d,vhgs6d	0.0526	0.4833	0.1027
145	0-25	vhgs2o,vhgs3o,vhgs6o	0.0492	0.5833	0.1027
150	0-25	vhgs1e,vhgs3e,vhgs6e	0.0457	0.6817	0.1027
155	0-25	vhgs2p,vhgs3p,vhgs6p	0.0422	0.7823	0.1027
160	0-25	vhgs1f,vhgs3f,vhgs6f	0.0386	0.9011	0.1027
165	0-25	vhgs2q,vhgs3q,vhgs6q	0.0361	0.9610	0.1027
170	0-25	vhgs1g,vhgs3g,vhgs6g	0.0336	0.9653	0.1027
175	0-25	vhgs2r,vhgs3r,vhgs6r	0.0311	0.9594	0.1027
180	0-25	vhgs1h,vhgs3h,vhgs6h	0.0286	0.9345	0.1027
185	0-25	vhgs2s,vhgs3s,vhgs6s	0.0268	0.8379	0.1027
190	0-25	vhgs1i,vhgs3i,vhgs6i	0.0249	0.7761	0.1027
195	0-25	vhgs2t,vhgs3t,vhgs6t	0.0230	0.7495	0.1027
200	0-25	vhgs1j,vhgs3j,vhgs6j	0.0212	0.7347	0.1027

Table 24: Higgs production via vector boson fusion.

$M_H(GeV^2)$	Period	Stntuple	σ (pb)	BR ($H \to WW$)	Filter Efficiency
110	0-25	fhgs4a,fhgs6a	0.2092	0.0441	0.6880
115	0-25	fhgs4l,fhgs6l	0.1788	0.0797	0.6932
120	0-25	fhgs4b,fhgs6b	0.1529	0.1320	0.6978
125	0-25	fhgs4m,fhgs6m	0.1324	0.2018	0.7004
130	0-25	fhgs4c,fhgs6c	0.1147	0.2869	0.7032
135	0-25	fhgs4n,fhgs6n	0.0993	0.3828	0.7045
140	0-25	fhgs4d,fhgs6d	0.0860	0.4833	0.7065
145	0-25	fhgs4o,fhgs6o	0.0753	0.5833	0.7075
150	0-25	fhgs4e,fhgs6e	0.0660	0.6817	0.7085
155	0-25	fhgs4p,fhgs6p	0.0578	0.7823	0.7098
160	0-25	fhgs4f,fhgs6f	0.0507	0.9011	0.7108
165	0-25	fhgs4q,fhgs6q	0.0444	0.9610	0.7114
170	0-25	fhgs4g,fhgs6g	0.0389	0.9653	0.7125
175	0-25	fhgs4r,fhgs6r	0.0346	0.9594	0.7130
180	0-25	fhgs4h,fhgs6h	0.0307	0.9345	0.7141
185	0-25	fhgs4s,fhgs6s	0.0273	0.8379	0.7149
190	0-25	fhgs4i,fhgs6i	0.0243	0.7761	0.7151
195	0-25	fhgs4t,fhgs6t	0.0217	0.7495	0.7158
200	0-25	fhgs4j,fhgs6j	0.0193	0.7347	0.7165

Table 25: Associated Higgs production with a W boson.

$M_H(GeV^2)$	Period	Stntuple	σ (pb)	BR ($H \to WW$)	Filter Efficiency
110	0-25	uhgs4a,uhgs6a	0.1243	0.0441	0.6930
115	0-25	uhgs4l,uhgs6l	0.1074	0.0797	0.6994
120	0-25	uhgs4b,uhgs6b	0.0927	0.1320	0.7031
125	0-25	uhgs4m,uhgs6m	0.0811	0.2018	0.7061
130	0-25	uhgs4c,uhgs6c	0.0709	0.2869	0.7087
135	0-25	uhgs4n,uhgs6n	0.0620	0.3828	0.7106
140	0-25	uhgs4d,uhgs6d	0.0542	0.4833	0.7122
145	0-25	uhgs4o,uhgs6o	0.0480	0.5833	0.7135
150	0-25	uhgs4e,uhgs6e	0.0425	0.6817	0.7151
155	0-25	uhgs4p,uhgs6p	0.0376	0.7823	0.7155
160	0-25	uhgs4f,uhgs6f	0.0333	0.9011	0.7172
165	0-25	uhgs4q,uhgs6q	0.0295	0.9610	0.7183
170	0-25	uhgs4g,uhgs6g	0.0261	0.9653	0.7184
175	0-25	uhgs4r,uhgs6r	0.0233	0.9594	0.7196
180	0-25	uhgs4h,uhgs6h	0.0208	0.9345	0.7204
185	0-25	uhgs4s,uhgs6s	0.0186	0.8379	0.7212
190	0-25	uhgs4i,uhgs6i	0.0166	0.7761	0.7220
195	0-25	uhgs4t,uhgs6t	0.0150	0.7495	0.7229
200	0-25	uhgs4j,uhgs6j	0.0135	0.7347	0.7239

Table 26: Associated Higgs production with a Z boson.

σ	is the cross-section for the Monte Carlo process (as listed in Tables 22-26)
$\mathcal{B}$	is any branching fraction for the Monte Carlo process (as listed in Tables 22-26)
ϵ_{filter}	is the filter efficiency applied for any filter used in the generation process (as listed in Tables 22-26)
ϵ_i^{trg}	is the effective trigger efficiency for the category (see below)
s_i^{lep}	is the effective lepton id scale factor (see below)
ϵ_{vtx}	is the run dependent efficiency of the z-vertex position requirement ($\|z_0\| < 60$ cm)
$\mathcal{L}_i$	is the luminosity of the dataset in good run list in which the lepton category falls (see Table 3.1)

The trigger efficiencies and livetimes are the same as described in CDF Note 9697 [4]. The lepton ID scale factors for this analysis were described in Section 2.2 and summarized in Tables 5-18.

4.3 W+jets Background

We use a data-driven fake rate proceedure to estimate the backgrounds from fake leptons. The procedure for extracting the fake rates is described in CDF Note 8538. In the dilepton plus MET final state we use both calorimeter and track isolation in our selection criteria for leptons.

Fake rates determined from generic jet samples are applied to a sample of lepton plus denominator object events where the deminator object is treated in the exact same manner as a fully identified lepton. These events are scaled by the rate at which the denominator object "fakes" a specific lepton type (as measured from the generic jet data). For each data event only the highest E_T denominator object is considered, although multiple events can be created in cases where the denominator object has the ability to "fake" more than one lepton type. Both possibilities are considered separately and each generated event is scaled by the appropriate measured fake rate.

A further correction is applied for non-triggerable fake leptons. Such a correction is necessary because events with a single real lepton corresponding to any of our non-triggerable lepton types will not appear in the single lepton samples from which our fake events are generated. The non-triggerable fake rate correction is obtained from our inclusive W Monte Carlo samples. For each dilepton category, we determine a weighted ratio of $W \to \ell\nu$ events with a second lepton originating from a promoted denominator object. The ratio is the weighted number of events in which the triggerable lepton originates from the promoted denominator divided by the weighted number of events in which the real lepton from the W decay is in fact the triggerable lepton ($N^{MC}_{\text{nontrig}}/N^{MC}_{\text{triggerable}}$). The weight of each generated fake event is the simply scaled up by $1 + N^{MC}_{\text{nontrig}}/\ N^{MC}_{\text{triggerable}}$, so that

$$N^{data}_{\text{total}} = N^{data}_{\text{triggerable}} + N^{data}_{\text{triggerable}} \frac{N^{MC}_{\text{nontrig}}}{N^{MC}_{\text{triggerable}}}. \tag{4}$$

The measured ratios of W+jet events with nontriggerable versus triggerable leptons originating from the W decay are shown in Tables 27 and 28 for each dilepton category. Dilepton categories with two trigger lepton types (e.g. CMUP-CMUP) can have a small non-triggerable correction factor due to the fact that one of the two leptons can have a E_T between 10 and 20 GeV/c and hence be non-triggerable by our definition. Non-triggerable fake rate corrections are now obtained separately for each of the different candidate samples.

	Opp-sign(0 jets)	Opp-sign(1 jet)	Opp-sign(2+ jets)	Opp-sign(Low M_{ll})	Same-sign(1+ jets)
e e:					
LBE LBE	0.037	0.038	0.054	0.126	0.000
LBE PHX	0.043	0.050	0.030	0.257	0.000
PHX PHX	0.037	0.017	0.049	0.080	0.000
e μ:					
LBE CMUP	0.043	0.061	0.048	0.167	0.000
LBE CMP	0.066	0.017	0.000	0.329	0.000
LBE CMX	0.051	0.067	0.102	0.187	0.000
LBE CMXMsKs	0.032	0.031	0.000	0.114	0.000
LBE BMU	0.000	0.000	0.000	0.000	0.000
LBE CMIOCES	0.409	0.564	0.559	0.063	1.089
LBE CMIOPES	0.160	0.168	0.179	0.068	0.253
PHX CMUP	0.049	0.050	0.028	0.000	0.000
PHX CMP	0.053	0.021	0.000	0.000	0.000
PHX CMX	0.031	0.065	0.121	0.242	0.000
PHX CMXMsKs	0.027	0.051	0.000	0.232	0.000
PHX BMU	0.000	0.000	0.000	0.000	0.000
PHX CMIOCES	0.635	1.684	1.961	0.748	0.000
PHX CMIOPES	0.246	0.478	0.280	0.255	0.000
μ μ:					
CMUP CMUP	0.067	0.083	0.140	0.232	0.000
CMUP CMP	0.108	0.072	0.000	1.143	0.000
CMUP CMX	0.059	0.060	0.097	0.242	0.000
CMUP CMXMsKs	0.079	0.000	0.000	0.347	0.000
CMUP BMU	4.617	8.871	0.000	0.000	6.066
CMUP CMIOCES	0.618	0.718	1.130	0.425	1.251
CMUP CMIOPES	0.317	0.275	0.926	0.000	0.434
CMX CMP	0.060	0.000	0.000	0.000	0.000
CMX CMX	0.060	0.102	0.000	0.160	0.000
CMX CMXMsKs	0.067	0.000	0.000	0.329	0.000
CMX BMU	3.130	4.003	0.000	0.000	7.333
CMX CMIOCES	0.529	0.844	1.124	0.462	0.983
CMX CMIOPES	0.279	0.189	0.814	0.083	0.355
CMXMsKs CMP	0.137	0.000	0.000	0.000	0.000
CMXMsKs CMXMsKs	0.000	0.000	0.000	2.377	0.000
CMXMsKs BMU	3.404	0.000	0.000	0.000	0.000
CMXMsKs CMIOCES	0.523	0.574	0.000	0.000	1.337
CMXMsKs CMIOPES	0.112	1.180	0.000	0.000	0.675
CMP CMP	0.000	0.000	0.000	0.000	0.000
CMP BMU	2.207	0.000	0.000	0.000	0.000
CMP CMIOCES	1.852	2.101	1.598	0.000	2.754
CMP CMIOPES	0.473	0.805	1.892	0.000	0.948
e Trk:					
LBE CrkTrk	0.112	0.137	0.204	0.026	0.403
PHX CrkTrk	0.199	0.367	0.721	0.088	0.000
μ Trk:					
CMUP CrkTrk	0.235	0.245	0.476	0.084	0.698
CMP CrkTrk	0.493	0.477	0.000	0.976	1.232
CMX CrkTrk	0.177	0.214	0.107	0.081	0.727
CMXMsKs CrkTrk	0.233	0.324	0.334	0.290	0.415

Table 27: Ratio of non-triggerable to triggerable fakes by dilepton category and signal region. Note that the ratio for dilepton categories with two trigger lepton types is expected to be zero in the same-sign region since both leptons are required to have $p_T > 20$ GeV/c.

	Same-sign(Low M_{ll})	Opp-sign LowMetSpec(Low M_{ll})	Same-sign(0 jets)
e e:			
LBE LBE	0.108	0.000	0.000
LBE PHX	0.083	0.150	0.000
PHX PHX	0.115	3.174	0.000
e μ:			
LBE CMUP	0.157	0.157	0.000
LBE CMP	0.147	0.000	0.000
LBE CMX	0.407	0.000	0.000
LBE CMXMsKs	0.169	0.000	0.000
LBE BMU	0.000	0.000	0.000
LBE CMIOCES	0.131	0.000	1.018
LBE CMIOPES	0.000	0.000	0.292
PHX CMUP	0.000	0.000	0.000
PHX CMP	0.000	0.000	0.000
PHX CMX	0.073	0.000	0.000
PHX CMXMsKs	0.000	0.000	0.000
PHX BMU	0.000	0.000	0.000
PHX CMIOCES	0.607	0.000	0.000
PHX CMIOPES	0.118	0.000	0.000
μ μ:			
CMUP CMUP	0.241	1.715	0.000
CMUP CMP	0.239	0.000	0.000
CMUP CMX	0.159	0.000	0.000
CMUP CMXMsKs	0.000	0.000	0.000
CMUP BMU	0.000	0.000	6.481
CMUP CMIOCES	0.607	0.000	1.197
CMUP CMIOPES	0.000	0.000	0.587
CMX CMP	0.000	0.000	0.000
CMX CMX	0.000	0.000	0.000
CMX CMXMsKs	0.000	0.000	0.000
CMX BMU	0.000	0.000	2.979
CMX CMIOCES	0.251	0.000	1.160
CMX CMIOPES	0.292	0.000	0.463
CMXMsKs CMP	0.000	0.000	0.000
CMXMsKs CMXMsKs	0.000	0.000	0.000
CMXMsKs BMU	0.000	0.000	3.374
CMXMsKs CMIOCES	0.000	0.000	1.469
CMXMsKs CMIOPES	1.386	0.000	0.122
CMP CMP	0.000	0.000	0.000
CMP BMU	0.000	0.000	3.046
CMP CMIOCES	0.000	0.000	2.840
CMP CMIOPES	0.000	0.000	0.976
e Trk:			
LBE CrkTrk	0.040	0.000	0.379
PHX CrkTrk	0.097	0.000	0.000
μ Trk:			
CMUP CrkTrk	0.139	0.000	0.638
CMP CrkTrk	0.000	0.000	1.341
CMX CrkTrk	0.065	0.000	0.623
CMXMsKs CrkTrk	0.108	0.000	0.444

Table 28: Ratio of non-triggerable to triggerable fakes by dilepton category and control region. Note that the ratio for dilepton categories with two trigger lepton types is expected to be zero in the same-sign region since both leptons are required to have $p_T > 20$ GeV/c.

5 Control Regions

In this section, we present the control regions in which we have tested our data model. Each control region attempts to isolate or enhance effects that must be modeled. Control regions are generally events obtained using the standard opposite-sign selection criteria from Section 3 but with the cuts on one particular selection requirement reversed. The control regions considered are:

- Events in the Drell-Yan (DY) control region are required to have $15 < \not{E}_{Tspec} < 25$ GeV and a dilepton invariant mass in the Z mass region. Only ee, $\mu\mu$, and e/μ+track events are accepted. This event sample tests our modeling of lepton ID scale factors and trigger efficiencies over all the different lepton categories as well as the luminosity accounting based on our good run lists.

- Events in the same-sign region satisfy an equivalent set of selection criteria to the opposite-sign events contained within our signal regions. These events test our modeling of contributions from photons and jets misidentified as leptons

- Events in the low $\not{E}_{T\mathrm{sig}}$ region test our modeling of tails in the missing E_T distribution for the Drell-Yan process in particular as well as the effect of mismeasured unclustered energy or combinations of leptons on $\not{E}_T$ modeling.

- Events in the Low $\not{E}_{Tspec}$, High $\not{E}_T$ region test the modeling of the effects of mismeasured individual leptons and jets on the missing E_T.

We choose two separate control regions to study $\not{E}_T$ mis-modeling because $\not{E}_T$ can arise from either mismeasurement of single objects (*ie.* a lepton or a jet) or from tails in the response function of the detector as a whole. More detail regarding each of the control regions is given below.

5.1 Dilepton Z-mass (DY) Control Region

Here we consider opposite-sign same-flavor dileptons in the Z mass region which is completely dominated by the Drell-Yan process. We require one trigger lepton and a second lepton with $E_T > 10$ GeV. We also require $15 < \not{E}_{Tspec} < 25$ GeV in this region, and include events with any number of jets.

Table 29 and Figures 10-12 show the agreement between the expected number of events and the observed yield for each dilepton catagory, and for lepton categories combined into ee and $\mu\mu$. We generally find good agreement in the shapes and normalizaton.

5.2 Base Same Sign Control Region

To check modeling of the W+jets and $W\gamma$ fake backgrounds, we look at events that satisfy all signal selection criteria with the exception of the opposite sign requirement which is

Category	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jets	Total	Data
e e	7.2	5.5	5.6	4.6	2961.2	3.0	50.6	3037.7	3065.0
e μ	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
μ μ	5.6	3.7	4.0	3.7	1659.7	0.0	15.2	1691.9	1736.0
e trk	5.4	1.8	1.8	3.3	821.6	0.5	19.3	853.8	789.0
μ trk	3.4	1.1	1.1	2.0	451.4	0.1	9.5	468.6	439.0
Total:	21.5	12.0	12.6	13.6	5893.9	3.6	94.7	6051.9	6029.0

Category	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jets	Total	Data
LBE LBE	4.2	2.9	3.3	3.1	1448.4	0.1	4.5	1466.5	1632.0
LBE PHX	2.6	2.2	2.0	1.4	1299.2	2.5	38.5	1348.3	1290.0
PHX PHX	0.4	0.4	0.3	0.1	213.7	0.3	7.6	222.9	143.0
LBE CMUP	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
LBE CMP	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
LBE CMX	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
LBE CMXMsKs	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
LBE BMU	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
LBE CMIOCES	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
LBE CMIOPES	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PHX CMUP	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PHX CMP	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PHX CMX	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PHX CMXMsKs	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PHX BMU	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PHX CMIOCES	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PHX CMIOPES	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CMUP CMUP	0.8	0.6	0.7	0.7	227.5	0.0	1.3	231.5	260.0
CMUP CMP	0.3	0.2	0.2	0.2	79.8	0.0	0.6	81.2	76.0
CMUP CMX	0.8	0.5	0.6	0.5	227.6	0.0	1.4	231.4	237.0
CMUP CMXMsKs	0.1	0.1	0.1	0.1	42.9	0.0	0.4	43.7	48.0
CMUP BMU	0.2	0.1	0.2	0.1	70.9	0.0	1.3	72.9	65.0
CMUP CMIOCES	1.2	0.8	0.9	0.9	323.3	0.0	1.4	328.5	356.0
CMUP CMIOPES	0.4	0.3	0.3	0.2	149.1	0.0	4.2	154.5	136.0
CMX CMP	0.1	0.1	0.1	0.1	43.8	0.0	0.3	44.5	41.0
CMX CMX	0.2	0.1	0.2	0.1	63.3	0.0	0.5	64.5	71.0
CMX CMXMsKs	0.1	0.0	0.0	0.0	18.6	0.0	0.1	18.9	26.0
CMX BMU	0.2	0.1	0.1	0.1	36.4	0.0	0.2	37.0	34.0
CMX CMIOCES	0.6	0.4	0.4	0.4	170.8	0.0	0.7	173.4	173.0
CMX CMIOPES	0.2	0.1	0.1	0.1	67.5	0.0	1.8	69.8	64.0
CMXMsKs CMP	0.0	0.0	0.0	0.0	11.0	0.0	0.1	11.2	6.0
CMXMsKs CMXMsKs	0.0	0.0	0.0	0.0	3.5	0.0	0.1	3.6	10.0
CMXMsKs BMU	0.0	0.0	0.0	0.0	7.3	0.0	0.0	7.4	2.0
CMXMsKs CMIOCES	0.1	0.1	0.1	0.1	31.5	0.0	0.1	31.9	34.0
CMXMsKs CMIOPES	0.0	0.0	0.0	0.0	11.6	0.0	0.3	12.0	9.0
CMP CMP	0.0	0.0	0.0	0.0	8.2	0.0	0.1	8.4	5.0
CMP BMU	0.0	0.0	0.0	0.0	7.1	0.0	0.0	7.2	8.0
CMP CMIOCES	0.2	0.1	0.1	0.1	41.4	0.0	0.3	42.2	57.0
CMP CMIOPES	0.0	0.0	0.0	0.0	16.4	0.0	0.1	16.7	18.0
LBE CrkTrk	4.2	1.4	1.4	2.8	598.6	0.2	14.4	623.0	629.0
PHX CrkTrk	1.2	0.4	0.4	0.5	223.0	0.3	5.0	230.9	160.0
CMUP CrkTrk	2.0	0.6	0.7	1.3	261.4	0.0	5.7	271.8	261.0
CMP CrkTrk	0.3	0.1	0.1	0.1	31.5	0.0	2.6	34.7	30.0
CMX CrkTrk	1.0	0.3	0.3	0.5	136.6	0.0	2.5	141.3	126.0
CMXMsKs CrkTrk	0.1	0.0	0.0	0.1	21.9	0.0	0.5	22.6	22.0
Total:	21.5	12.0	12.6	13.6	5893.9	3.6	96.9	6054.1	6029.0

Table 29: Predicted and observed yields in the Drell-Yan Region.

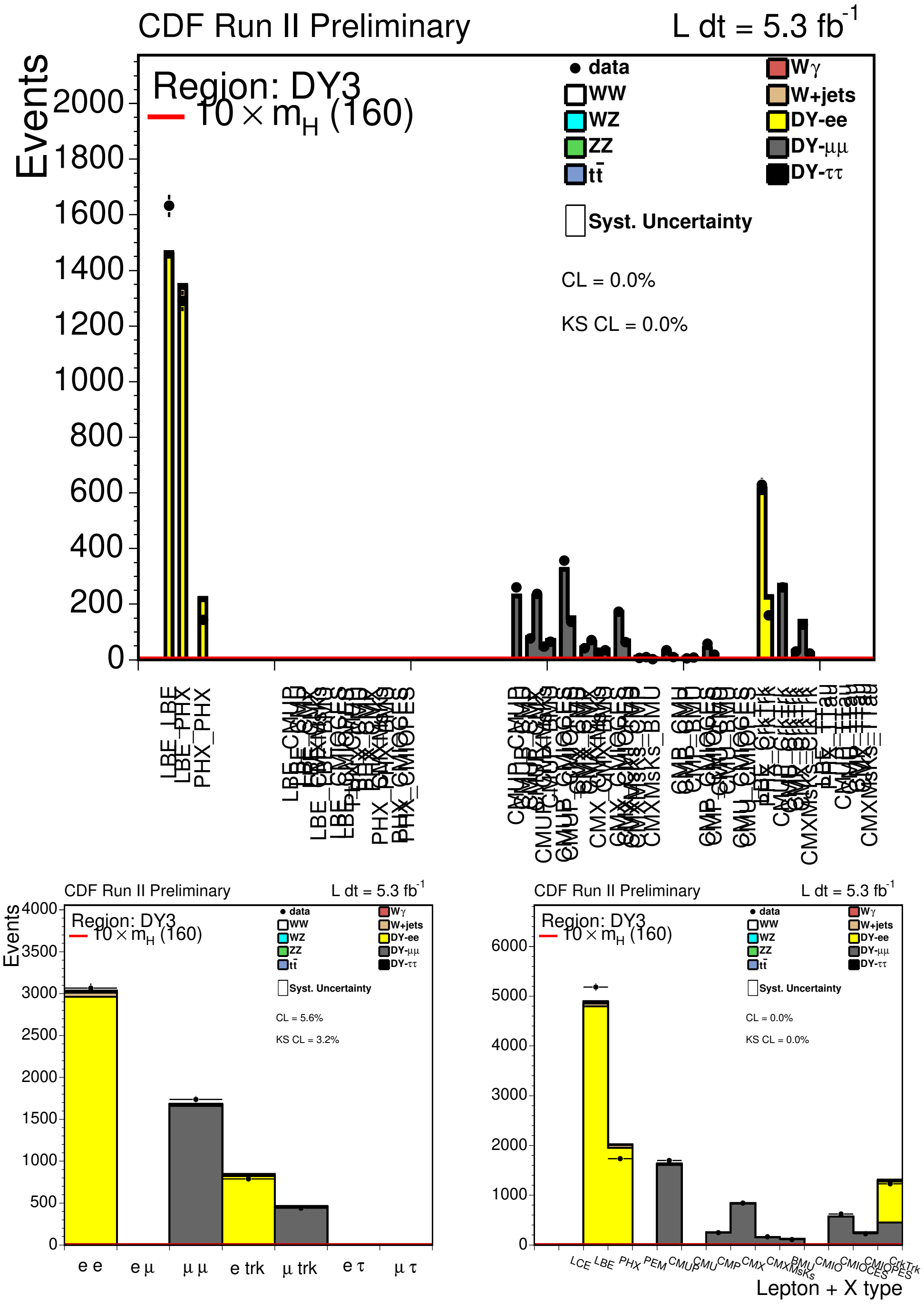

Figure 10: Drell-Yan ($Z \to l^+l^-$) control region.

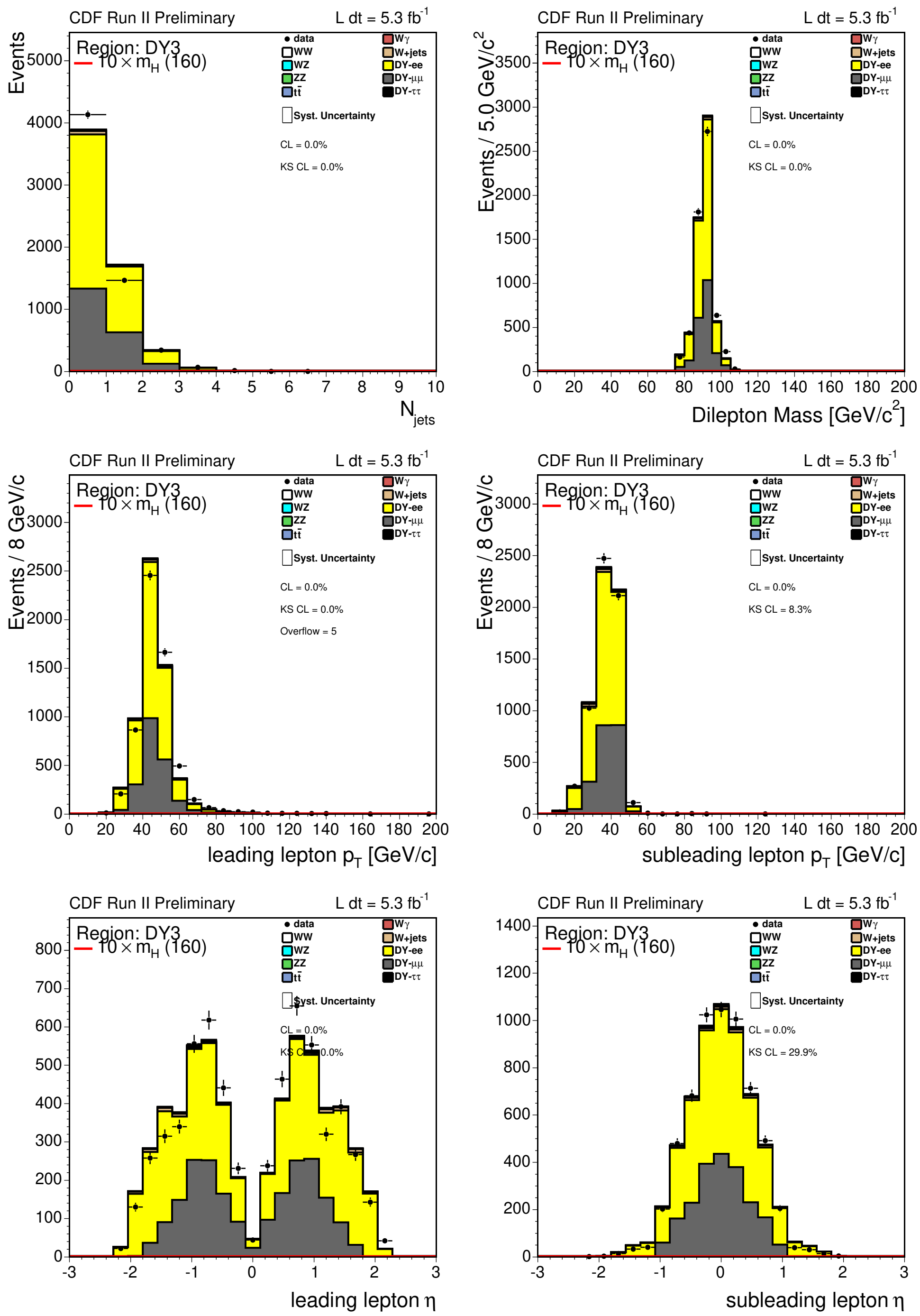

Figure 11: Drell-Yan ($Z \to l^+l^-$) control region.

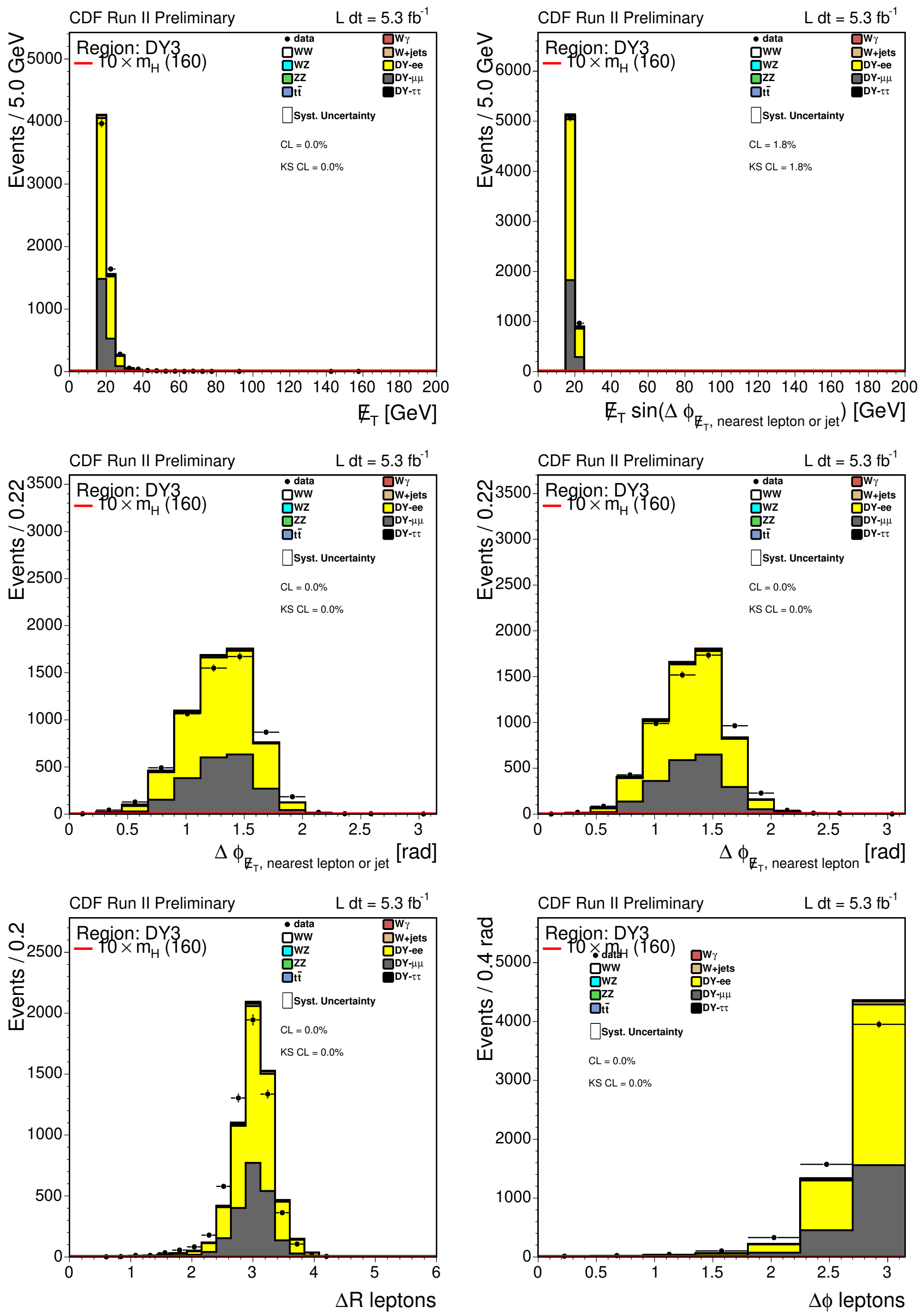

Figure 12: Drell-Yan ($Z \to l^+l^-$) control region.

reversed. We now also include events with any number of jets; previously this control region was limited to events with only zero or one jets. This region is dominated by fake leptons originating either from jets or photons. The PHX categories also have significant contributions from charge mis-identification. The results are shown in Table 30. In this region we observe excellent agreement in the observed and predicted numbers of events as well as in the predicted and observed kinematic distributions as shown in Figures 13-15.

5.3 Low $\not\!\!E_{T\mathrm{sig}}$ Control Region

As a cross check of the $\not\!\!E_T$ simulation and in order to assess a systematic on the Drell-Yan background, we create a control sample using the standard opposite-sign dilepton selection criteria along with an additional requirement of $\not\!\!E_{T\mathrm{sig}} \equiv \not\!\!E_T/\sqrt{\sum E_T} < 2.5$ GeV$^{\frac{1}{2}}$. We now also include events with any number of jets; previously this control region was limited to events with only zero or one jets. The resulting contol sample is populated primarily by Drell-Yan events with large $\not\!\!E_T$ that results from mismeasurement of the unclustered energy or a combination of leptons and jets. Drell-Yan events with $\not\!\!E_T$ due to lost energy from a single lepton or jet are removed by the standard $\not\!\!E_{Tspec}$ cut. In this region , we also find good agreement between the data and our model as illustrated in Table 31 and Figures 16-18.

5.4 Low $\not\!\!E_{Tspec}$ High $\not\!\!E_T$ Control Region

As a secondary check of the $\not\!\!E_T$ modeling, we look in the region of $\not\!\!E_{Tspec}$ between 15 GeV (10 GeV $e\mu$) and the signal region lower requirement of 25 GeV (15 GeV $e\mu$) and also require high $\not\!\!E_T$. We now also include events with any number of jets; previously this control region was limited to events with only zero or one jets. The events in this region are dominated by Drell-Yan with $\not\!\!E_T$ due to mismeasured single leptons or jets. In this region, the predicted and observed numbers of events agree at about the one sigma level as shown in Table 32. However, there are several significant disagreements between the observed and predicted kinematic distributions (in particular an excess of data events above the Z-peak in the dilepton invariant mass distribution) as illustrated in Figures 19-21. We suspect that some of these events might be originating from non-collision sources such as cosmic rays but understanding this tail will require additional studies. In the meantime, we already take a large systematic uncertainty on our modeling of the missing E_T tail in Drell-Yan events and believe that this is sufficient to cover the level of disagreement observed in these distributions.

Category	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jets	Total	Data
e e	7.4	5.4	1.4	3.3	10.0	86.5	96.2	210.2	188.0
e μ	8.7	9.5	0.8	3.8	22.4	75.2	97.3	217.8	195.0
μ μ	0.0	3.0	0.3	0.0	0.3	0.0	7.1	10.7	9.0
e trk	2.5	2.8	0.3	1.1	3.2	17.9	15.8	43.7	55.0
μ trk	0.0	1.5	0.1	0.0	0.5	2.7	7.0	11.8	10.0
Total:	18.6	22.3	3.0	8.3	36.3	182.3	223.3	494.2	457.0

Category	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jets	Total	Data
LBE LBE	0.1	2.0	0.1	0.1	0.2	20.7	15.7	38.9	23.0
LBE PHX	5.5	2.7	1.0	2.7	7.5	46.8	58.7	124.9	117.0
PHX PHX	1.9	0.7	0.3	0.5	2.2	19.0	21.8	46.4	48.0
LBE CMUP	0.0	2.4	0.2	0.1	2.5	11.7	14.7	31.6	23.0
LBE CMP	0.0	0.3	0.0	0.0	0.1	1.3	2.1	4.0	4.0
LBE CMX	0.0	1.2	0.1	0.0	1.7	5.4	7.2	15.6	15.0
LBE CMXMsKs	0.0	0.2	0.0	0.0	0.2	1.0	1.4	2.9	0.0
LBE BMU	0.0	0.3	0.0	0.0	0.1	0.6	0.0	1.0	0.0
LBE CMIOCES	-0.0	1.7	0.1	0.0	0.5	3.7	2.1	8.1	9.0
LBE CMIOPES	0.0	0.6	0.0	0.0	0.2	1.5	7.1	9.5	12.0
PHX CMUP	3.0	1.1	0.1	1.5	6.4	20.9	26.0	59.1	55.0
PHX CMP	0.5	0.1	0.0	0.2	0.4	2.3	3.6	7.3	4.0
PHX CMX	1.4	0.5	0.0	0.6	3.6	12.8	15.2	34.2	23.0
PHX CMXMsKs	0.2	0.1	0.0	0.1	1.1	2.1	2.0	5.6	5.0
PHX BMU	0.4	0.1	0.0	0.1	0.8	1.6	0.0	3.1	4.0
PHX CMIOCES	2.2	0.7	0.1	0.9	2.9	6.9	2.3	16.1	32.0
PHX CMIOPES	0.8	0.2	0.0	0.2	1.7	3.4	5.5	11.8	9.0
CMUP CMUP	0.0	0.4	0.1	0.0	0.0	0.0	0.7	1.3	0.0
CMUP CMP	0.0	0.1	0.0	0.0	0.0	0.0	0.2	0.4	1.0
CMUP CMX	0.0	0.5	0.1	0.0	0.0	0.0	0.6	1.2	1.0
CMUP CMXMsKs	0.0	0.1	0.0	0.0	0.0	0.0	0.1	0.2	0.0
CMUP BMU	0.0	0.1	0.0	0.0	0.0	0.0	0.3	0.4	0.0
CMUP CMIOCES	0.0	0.6	0.1	0.0	0.0	0.0	0.8	1.5	4.0
CMUP CMIOPES	0.0	0.2	0.0	0.0	0.1	0.0	1.8	2.2	1.0
CMX CMP	0.0	0.1	0.0	0.0	0.0	0.0	0.1	0.2	0.0
CMX CMX	0.0	0.1	0.0	0.0	0.0	0.0	0.1	0.2	1.0
CMX CMXMsKs	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.0
CMX BMU	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.2	0.0
CMX CMIOCES	0.0	0.3	0.0	0.0	0.0	0.0	0.4	0.8	1.0
CMX CMIOPES	0.0	0.1	0.0	0.0	0.0	0.0	1.0	1.1	0.0
CMXMsKs CMP	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CMXMsKs CMXMsKs	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CMXMsKs BMU	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CMXMsKs CMIOCES	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.0
CMXMsKs CMIOPES	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.3	0.0
CMP CMP	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CMP BMU	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CMP CMIOCES	0.0	0.1	0.0	0.0	0.0	0.0	0.1	0.2	0.0
CMP CMIOPES	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.4	0.0
LBE CrkTrk	0.0	1.9	0.1	0.0	0.6	8.6	7.9	19.2	18.0
PHX CrkTrk	2.5	0.9	0.3	1.1	2.5	9.3	8.1	24.7	37.0
CMUP CrkTrk	0.0	0.9	0.1	0.0	0.3	1.6	4.3	7.2	2.0
CMP CrkTrk	0.0	0.1	0.0	0.0	0.0	0.2	0.6	1.0	0.0
CMX CrkTrk	0.0	0.5	0.0	0.0	0.2	0.8	1.9	3.3	8.0
CMXMsKs CrkTrk	0.0	0.1	0.0	0.0	0.0	0.1	0.2	0.3	0.0
Total:	18.6	22.3	3.0	8.3	36.3	182.3	215.8	486.6	457.0

Table 30: Expected and observed yields in the Same-Sign region.

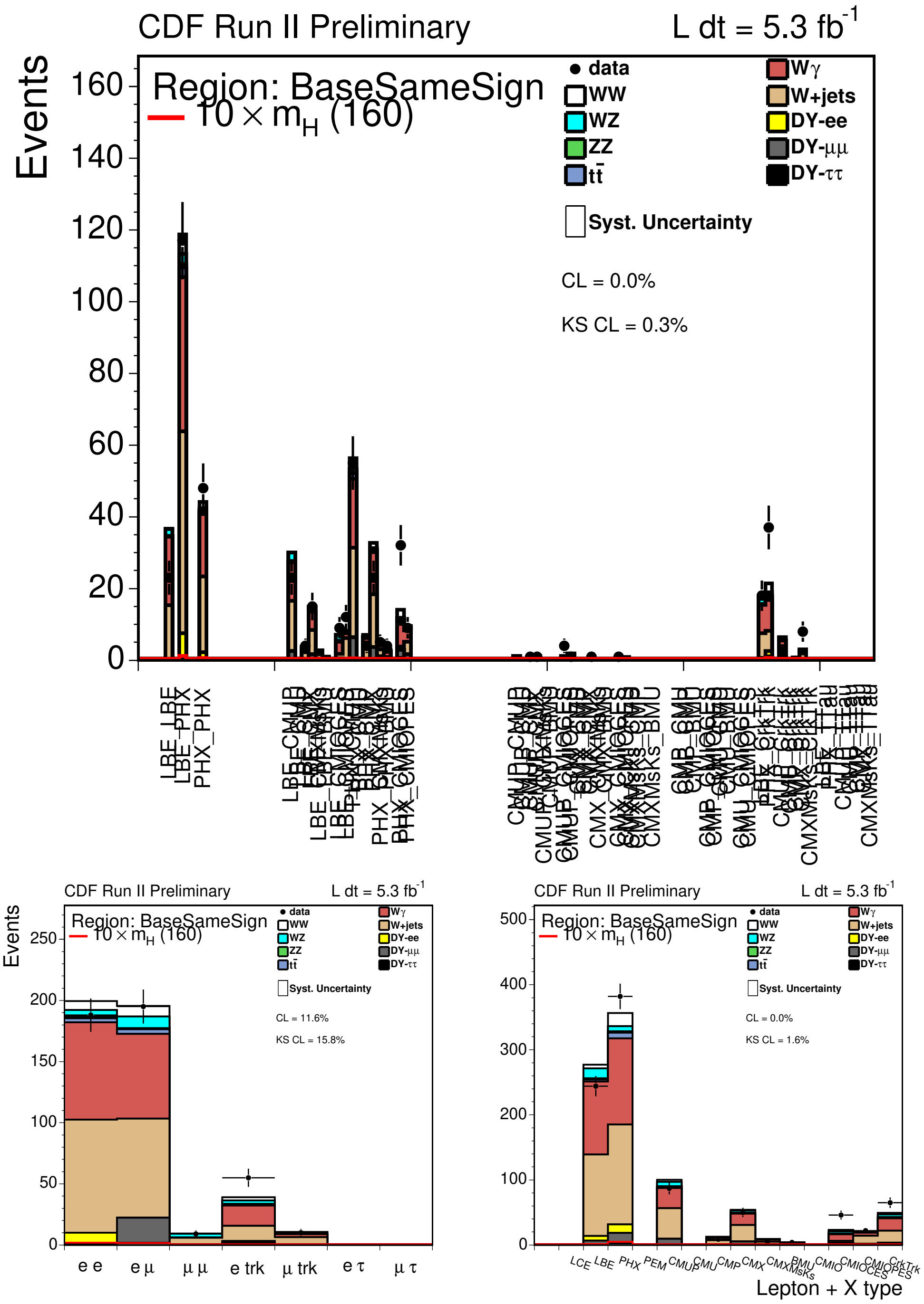

Figure 13: Base Same Sign control region.

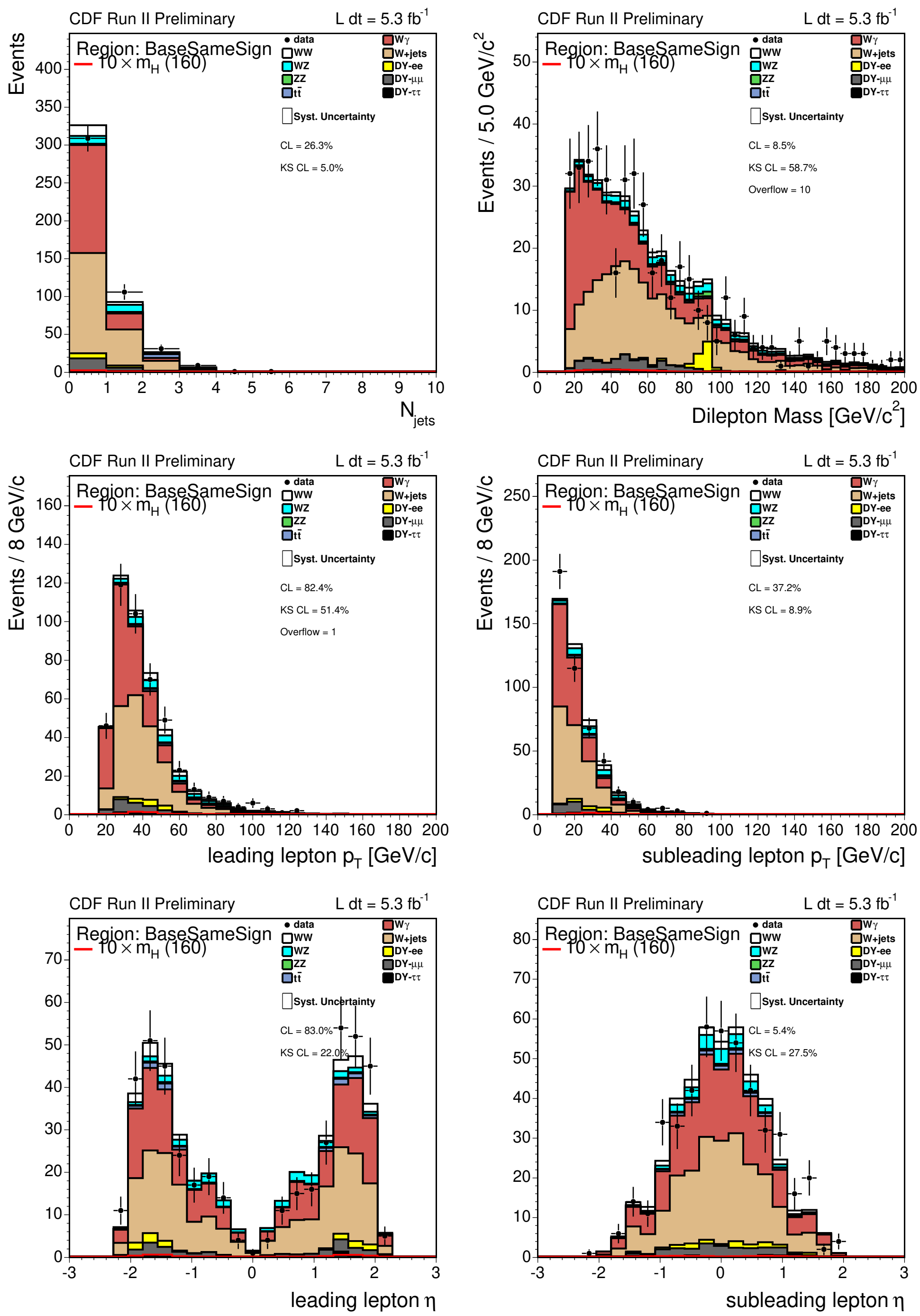

Figure 14: Base Same Sign control region.

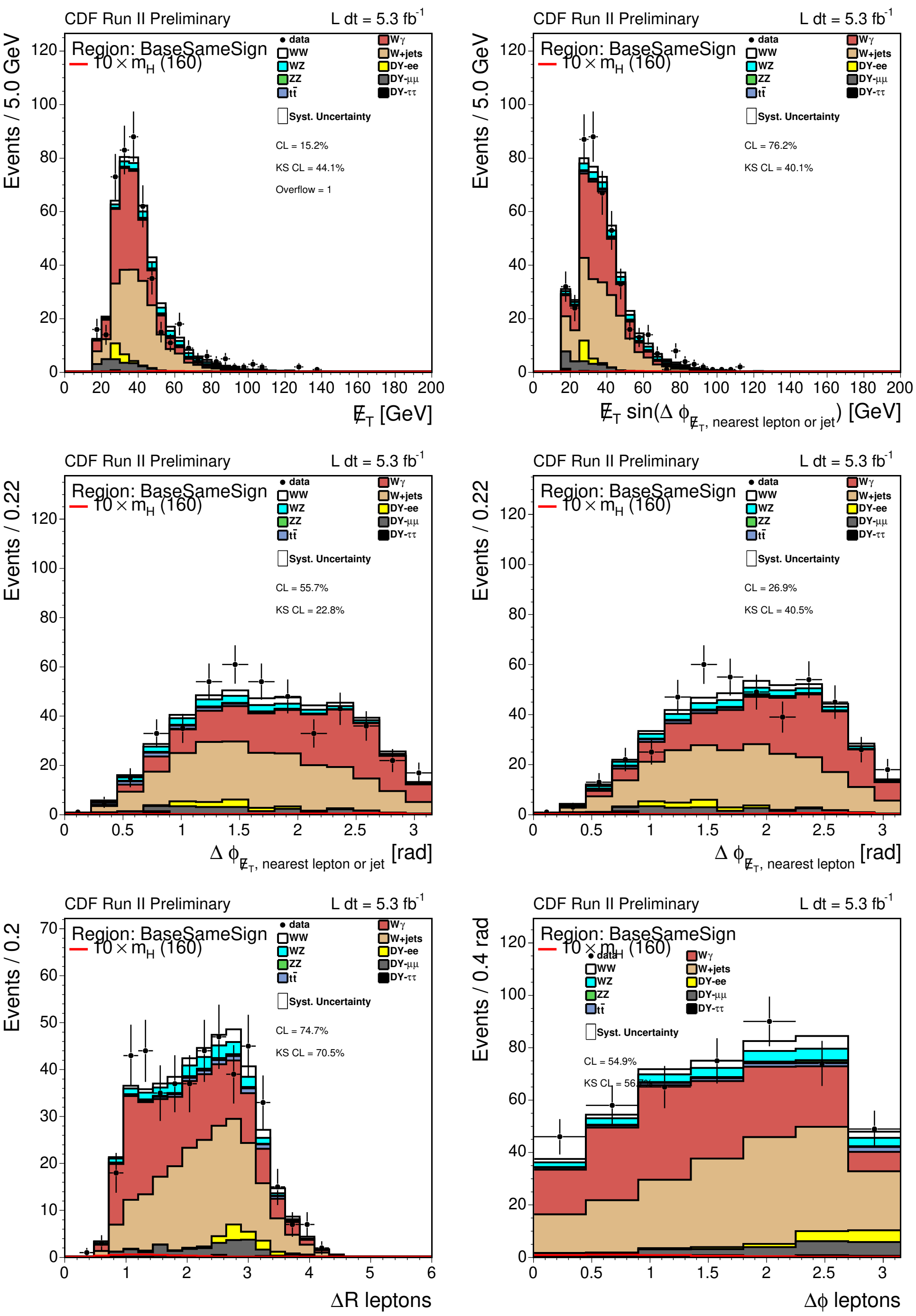

Figure 15: Base Same Sign control region.

Category	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jets	Total	Data
e e	8.5	2.3	1.8	8.1	101.9	13.1	28.2	163.9	201.0
e μ	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
μ μ	6.4	1.5	1.4	6.5	60.2	0.0	4.0	80.0	100.0
e trk	5.9	1.0	0.7	5.6	35.4	2.5	6.3	57.3	57.0
μ trk	3.3	0.5	0.4	3.4	15.2	0.4	3.1	26.3	24.0
Total:	24.1	5.3	4.2	23.6	212.7	16.0	41.6	327.5	382.0

Category	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jets	Total	Data
LBE LBE	5.2	1.3	1.1	5.9	56.5	3.0	5.7	78.6	99.0
LBE PHX	2.8	0.9	0.6	2.0	36.7	7.1	16.0	66.1	80.0
PHX PHX	0.5	0.2	0.1	0.2	8.7	3.0	6.4	19.1	22.0
LBE CMUP	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
LBE CMP	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
LBE CMX	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
LBE CMXMsKs	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
LBE BMU	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
LBE CMIOCES	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
LBE CMIOPES	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PHX CMUP	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PHX CMP	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PHX CMX	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PHX CMXMsKs	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PHX BMU	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PHX CMIOCES	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PHX CMIOPES	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CMUP CMUP	1.1	0.2	0.2	1.3	7.9	0.0	0.4	11.3	14.0
CMUP CMP	0.3	0.1	0.1	0.4	2.3	0.0	0.4	3.6	4.0
CMUP CMX	0.9	0.2	0.2	1.0	8.7	0.0	0.4	11.4	5.0
CMUP CMXMsKs	0.1	0.0	0.0	0.2	1.7	0.0	0.1	2.1	5.0
CMUP BMU	0.2	0.1	0.0	0.2	3.0	0.0	0.0	3.6	5.0
CMUP CMIOCES	1.6	0.3	0.3	1.6	10.3	0.0	0.5	14.7	20.0
CMUP CMIOPES	0.3	0.1	0.1	0.3	5.0	0.0	0.9	6.7	10.0
CMX CMP	0.2	0.0	0.0	0.1	1.6	0.0	0.0	2.0	2.0
CMX CMX	0.2	0.1	0.0	0.2	1.8	0.0	0.1	2.4	4.0
CMX CMXMsKs	0.0	0.0	0.0	0.1	0.7	0.0	0.0	0.9	0.0
CMX BMU	0.1	0.0	0.0	0.1	1.6	0.0	0.0	1.8	6.0
CMX CMIOCES	0.7	0.2	0.1	0.6	6.5	0.0	0.2	8.3	10.0
CMX CMIOPES	0.2	0.0	0.0	0.1	2.6	0.0	0.3	3.2	4.0
CMXMsKs CMP	0.0	0.0	0.0	0.0	0.8	0.0	0.0	0.9	1.0
CMXMsKs CMXMsKs	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.5	0.0
CMXMsKs BMU	0.0	0.0	0.0	0.0	0.6	0.0	0.0	0.7	0.0
CMXMsKs CMIOCES	0.1	0.0	0.0	0.1	1.5	0.0	0.0	1.7	3.0
CMXMsKs CMIOPES	-0.0	0.0	0.0	0.0	0.9	0.0	0.3	1.2	2.0
CMP CMP	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.4	2.0
CMP BMU	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.3	0.0
CMP CMIOCES	0.2	0.0	0.0	0.2	1.1	0.0	0.0	1.6	2.0
CMP CMIOPES	0.0	0.0	0.0	0.0	0.5	0.0	0.4	0.9	1.0
LBE CrkTrk	4.6	0.7	0.5	4.8	26.1	1.0	4.1	41.9	45.0
PHX CrkTrk	1.3	0.2	0.1	0.7	9.3	1.5	2.1	15.4	12.0
CMUP CrkTrk	2.1	0.3	0.2	2.2	9.8	0.2	1.6	16.4	16.0
CMP CrkTrk	0.2	0.0	0.0	0.2	0.9	0.0	0.7	2.1	2.0
CMX CrkTrk	0.9	0.1	0.1	0.8	3.7	0.1	0.9	6.7	3.0
CMXMsKs CrkTrk	0.1	0.0	0.0	0.1	0.8	0.0	0.1	1.2	3.0
Total:	24.1	5.3	4.2	23.6	212.7	16.0	41.8	327.7	382.0

Table 31: Expected and observed yields in the Low $\not\!\!E_{T_{\rm sig}}$ region.

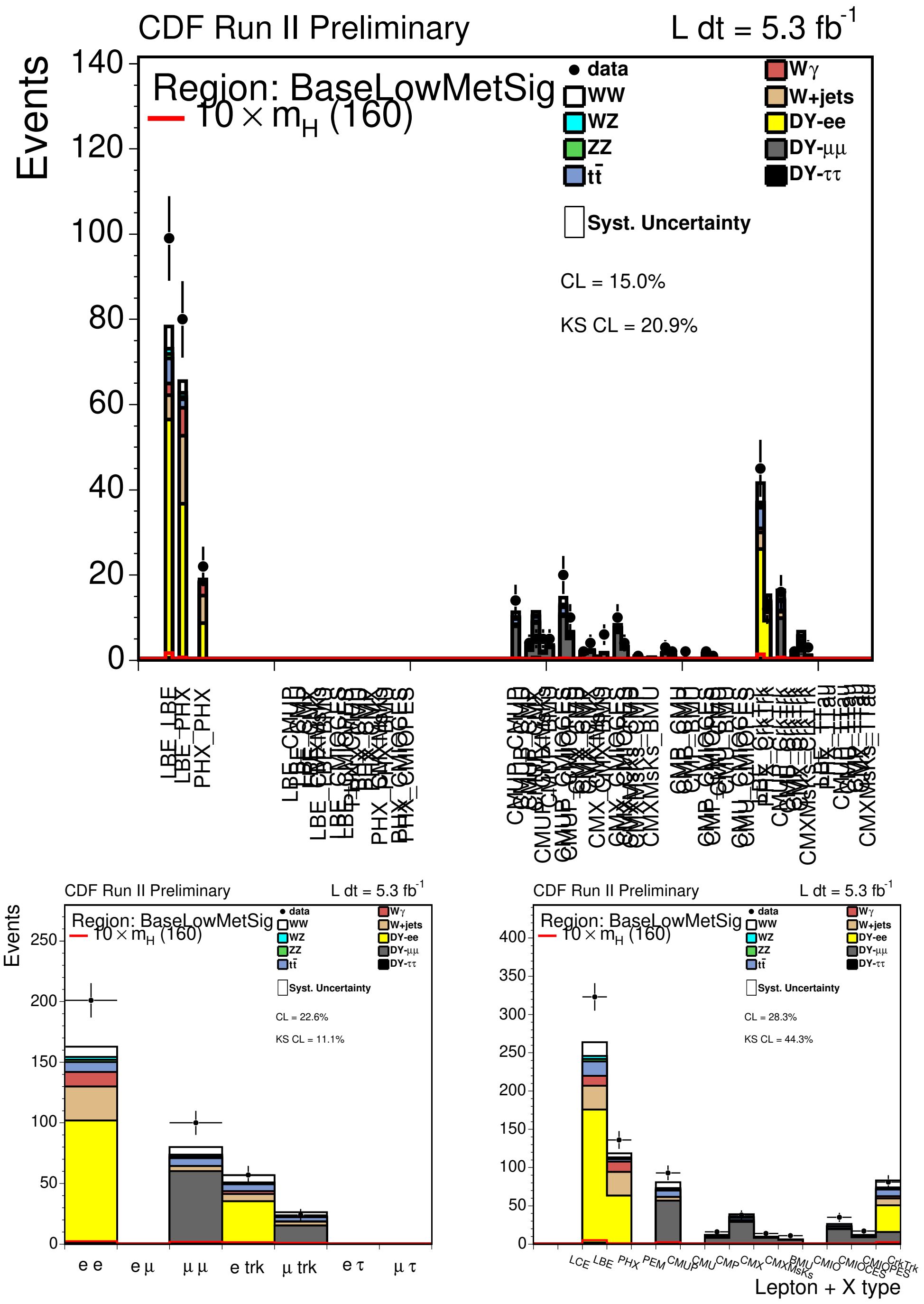

Figure 16: Base Low $\not\!\!E_{T\,\mathrm{sig}}$ control region.

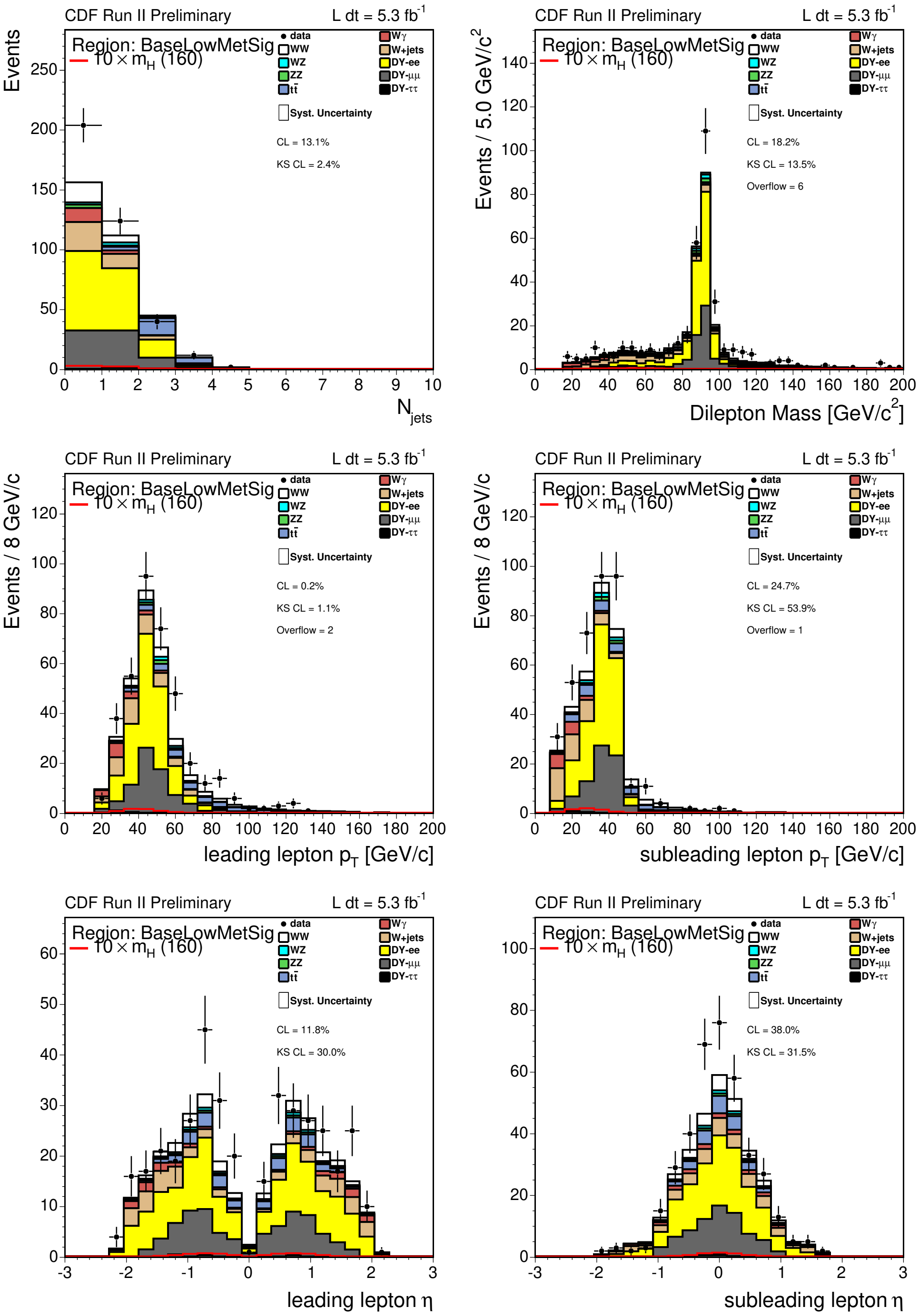

Figure 17: Base Low $\not{E}_{T\,\mathrm{sig}}$ control region.

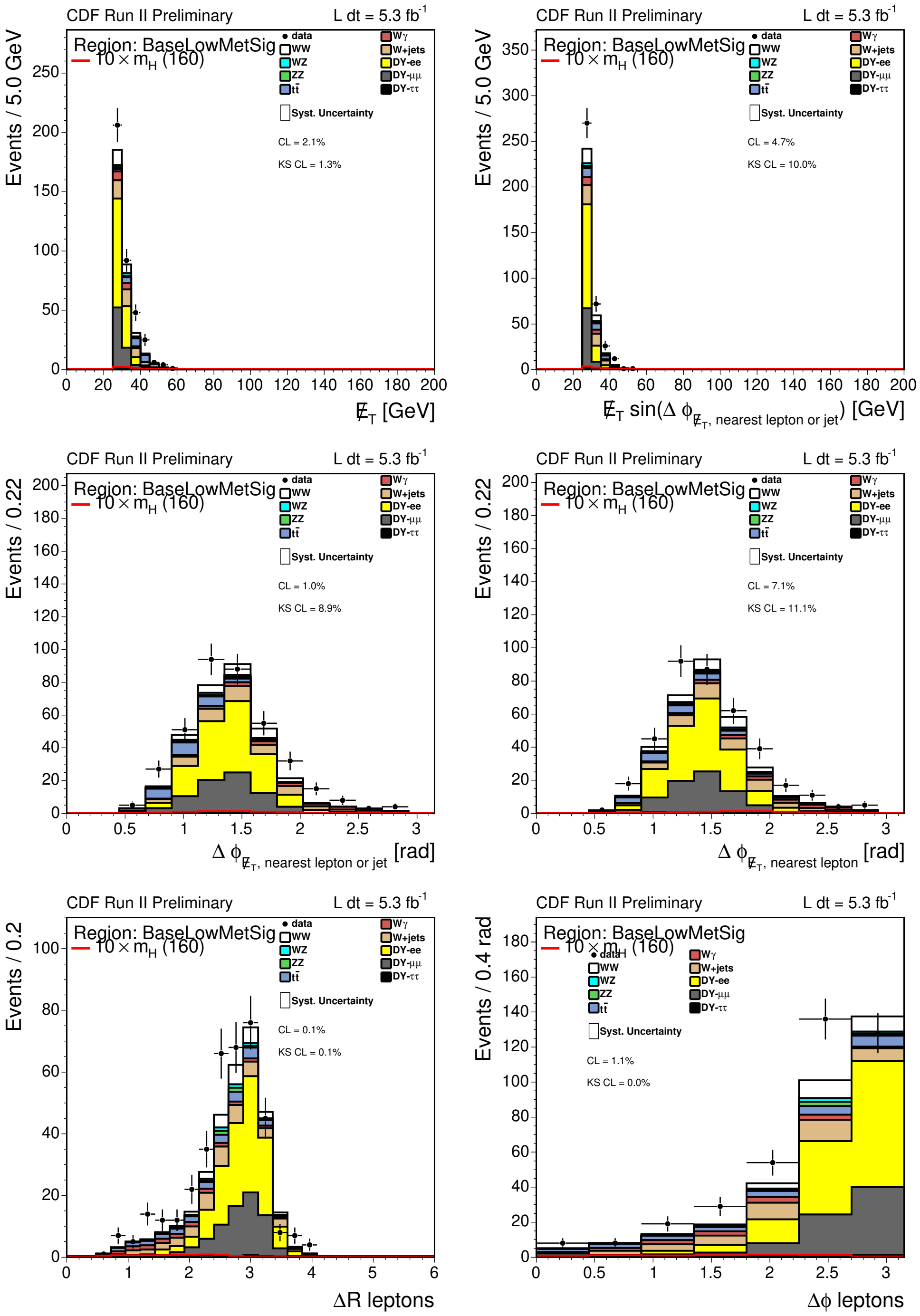

Figure 18: Base Low $\not\!\!E_{T\,\mathrm{sig}}$ control region.

Category	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jets	Total	Data
e e	11.6	4.0	3.4	16.8	209.5	9.6	26.6	281.6	284.0
e μ	17.4	1.0	0.4	20.3	130.6	5.1	24.3	199.2	147.0
μ μ	8.4	2.4	2.4	13.4	144.3	0.0	6.2	177.1	252.0
e trk	8.0	1.7	1.4	12.0	156.9	2.3	9.5	191.9	176.0
μ trk	4.9	0.8	0.7	7.3	32.1	0.3	5.0	50.8	60.0
Total:	50.3	9.9	8.2	69.8	673.4	17.4	71.6	900.6	919.0

Category	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jets	Total	Data
LBE LBE	7.1	2.2	2.1	11.9	112.9	1.9	5.7	143.6	165.0
LBE PHX	4.0	1.5	1.1	4.5	79.8	5.9	14.3	111.1	96.0
PHX PHX	0.6	0.3	0.2	0.5	16.8	1.9	6.6	26.9	23.0
LBE CMUP	4.9	0.3	0.1	7.1	43.0	0.6	4.7	60.7	48.0
LBE CMP	0.8	0.0	0.0	1.0	5.1	0.1	3.3	10.2	4.0
LBE CMX	2.3	0.1	0.1	2.8	17.6	0.3	2.4	25.6	16.0
LBE CMXMsKs	0.3	0.0	0.0	0.5	3.4	0.0	0.5	4.7	4.0
LBE BMU	0.6	0.0	0.0	0.5	3.5	0.1	0.0	4.8	0.0
LBE CMIOCES	3.5	0.2	0.1	4.5	19.0	0.2	0.8	28.2	32.0
LBE CMIOPES	1.1	0.1	0.0	0.9	7.1	0.1	1.2	10.4	3.0
PHX CMUP	1.5	0.1	0.0	1.3	11.2	1.5	5.0	20.6	10.0
PHX CMP	0.2	0.0	0.0	0.2	1.0	0.2	1.6	3.2	3.0
PHX CMX	0.7	0.1	0.0	0.6	8.5	0.9	1.4	12.1	9.0
PHX CMXMsKs	0.1	0.0	0.0	0.1	0.7	0.2	0.4	1.5	1.0
PHX BMU	0.2	0.0	0.0	0.1	1.9	0.1	0.0	2.3	1.0
PHX CMIOCES	1.0	0.1	0.0	0.7	3.9	0.5	0.8	7.0	14.0
PHX CMIOPES	0.4	0.0	0.0	0.1	4.8	0.3	0.6	6.2	2.0
CMUP CMUP	1.5	0.4	0.4	2.7	16.2	0.0	0.5	21.6	32.0
CMUP CMP	0.4	0.1	0.1	0.7	4.7	0.0	0.2	6.2	5.0
CMUP CMX	1.4	0.3	0.3	2.0	16.8	0.0	0.5	21.4	31.0
CMUP CMXMsKs	0.2	0.1	0.1	0.4	3.2	0.0	0.1	3.9	6.0
CMUP BMU	0.3	0.1	0.1	0.4	13.2	0.0	0.1	14.2	24.0
CMUP CMIOCES	1.7	0.5	0.5	3.3	19.6	0.0	0.5	26.1	47.0
CMUP CMIOPES	0.5	0.2	0.1	0.6	18.3	0.0	1.9	21.6	28.0
CMX CMP	0.2	0.0	0.0	0.3	4.1	0.0	0.1	4.9	7.0
CMX CMX	0.3	0.1	0.1	0.4	4.6	0.0	0.1	5.6	7.0
CMX CMXMsKs	0.1	0.0	0.0	0.2	1.8	0.0	0.0	2.1	2.0
CMX BMU	0.2	0.0	0.0	0.2	5.9	0.0	0.2	6.5	6.0
CMX CMIOCES	0.9	0.2	0.2	1.3	12.1	0.0	0.2	15.0	17.0
CMX CMIOPES	0.3	0.1	0.1	0.2	10.1	0.0	0.9	11.7	10.0
CMXMsKs CMP	0.0	0.0	0.0	0.1	0.6	0.0	0.1	0.8	0.0
CMXMsKs CMXMsKs	0.0	0.0	0.0	0.0	0.4	0.0	0.0	0.5	1.0
CMXMsKs BMU	0.0	0.0	0.0	0.0	1.3	0.0	0.0	1.4	2.0
CMXMsKs CMIOCES	0.2	0.0	0.0	0.2	1.7	0.0	0.1	2.1	5.0
CMXMsKs CMIOPES	0.0	0.0	0.0	0.0	1.5	0.0	0.2	1.8	4.0
CMP CMP	0.0	0.0	0.0	0.1	0.8	0.0	0.0	0.9	0.0
CMP BMU	0.0	0.0	0.0	0.0	1.5	0.0	0.0	1.6	3.0
CMP CMIOCES	0.2	0.1	0.1	0.4	3.3	0.0	0.0	4.1	8.0
CMP CMIOPES	0.1	0.0	0.0	0.1	2.6	0.0	0.1	2.9	7.0
LBE CrkTrk	6.5	1.3	1.2	10.3	122.1	1.0	6.2	148.5	136.0
PHX CrkTrk	1.5	0.4	0.3	1.7	34.9	1.3	3.3	43.4	40.0
CMUP CrkTrk	3.1	0.5	0.4	4.7	17.9	0.1	2.9	29.6	43.0
CMP CrkTrk	0.3	0.0	0.0	0.5	1.8	0.0	0.7	3.4	2.0
CMX CrkTrk	1.3	0.2	0.2	1.8	11.4	0.1	1.2	16.3	11.0
CMXMsKs CrkTrk	0.1	0.0	0.0	0.3	1.0	0.0	0.4	1.9	4.0
Total:	50.3	9.9	8.2	69.8	673.4	17.4	69.8	898.8	919.0

Table 32: Expected and observed yields in the Base Low $\not\!\!E_{T spec}$ High $\not\!\!E_T$ region.

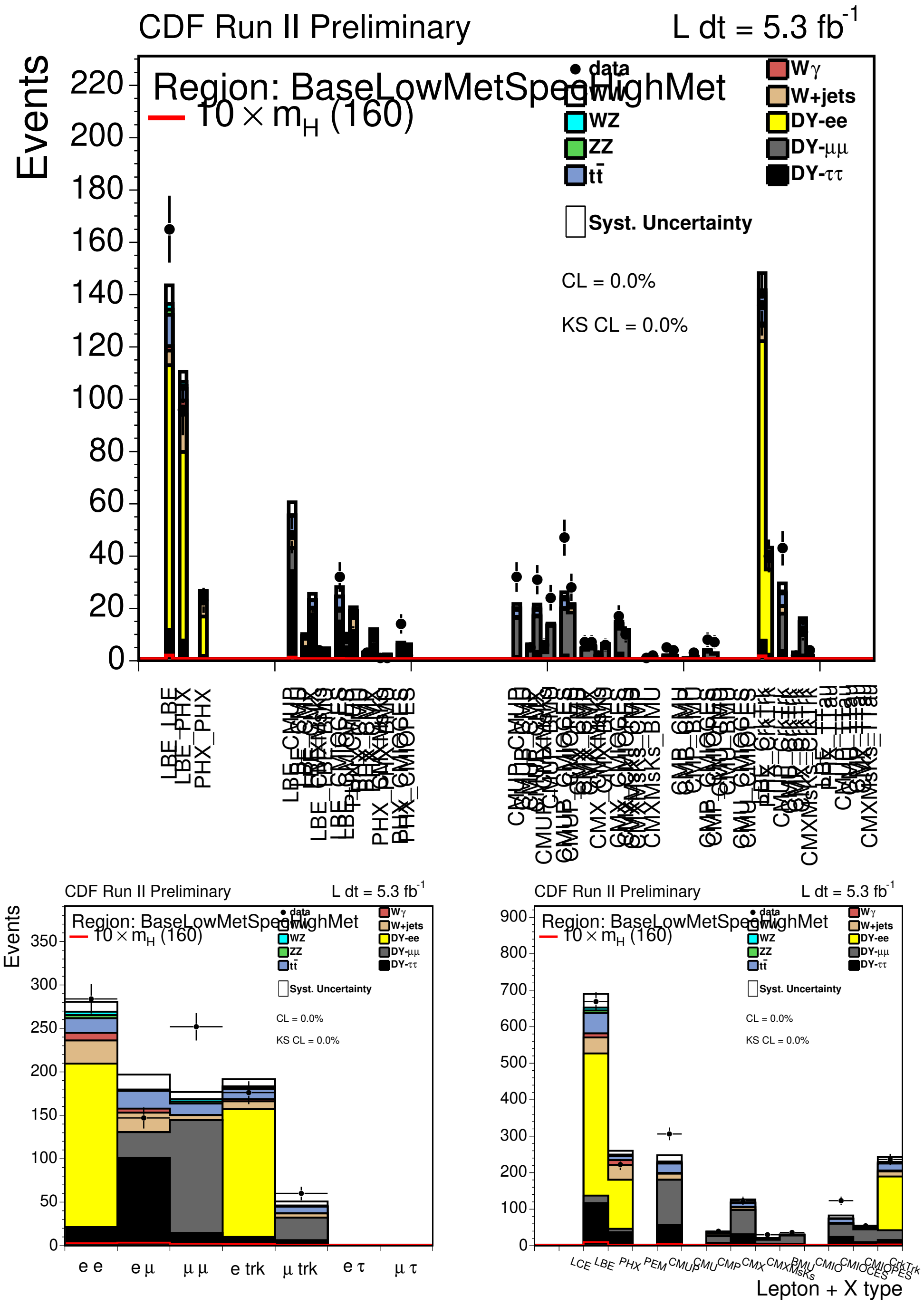

Figure 19: Base Low $\not\!\!E_{T spec}$ High $\not\!\!E_T$ control region.

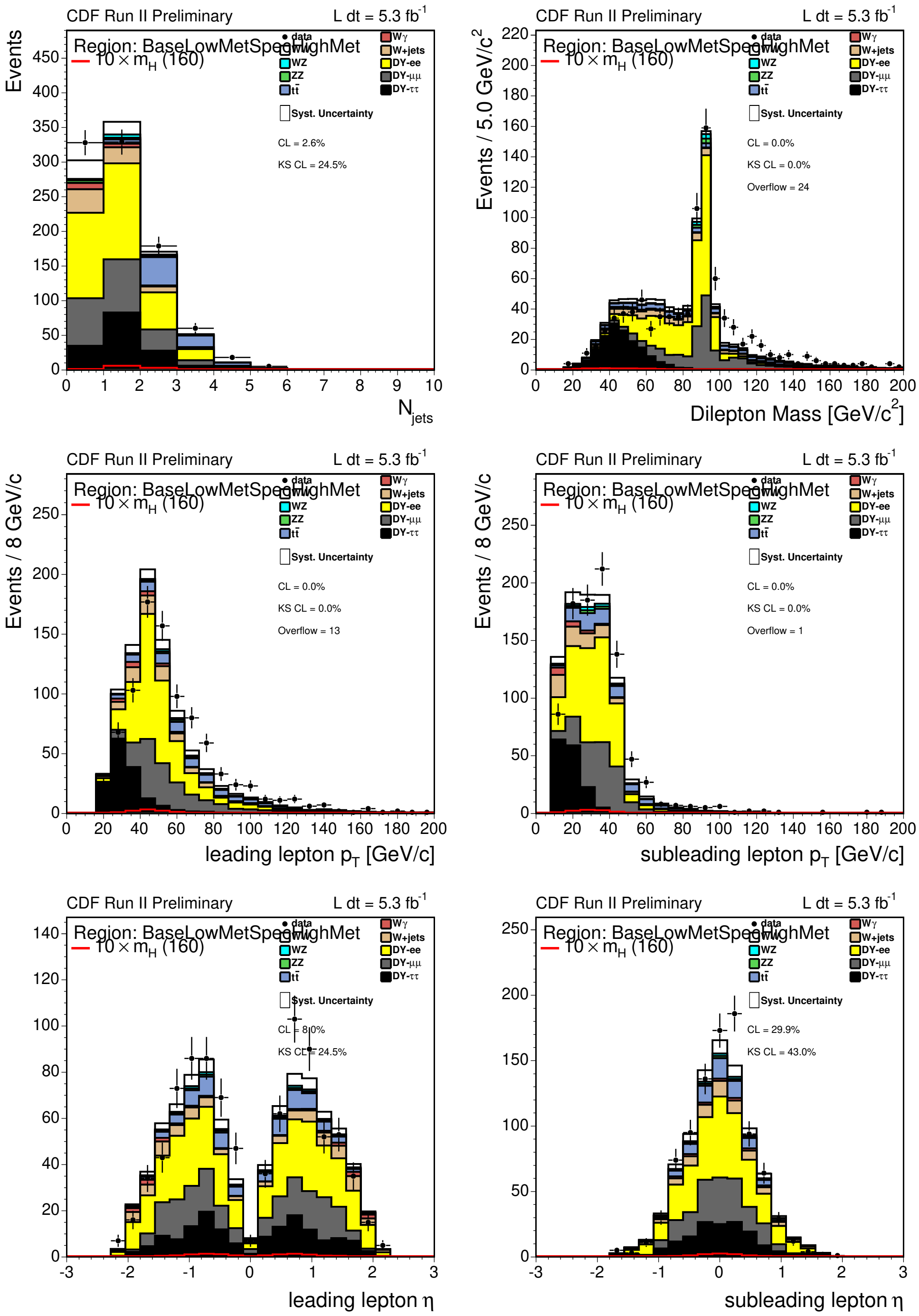

Figure 20: Base Low $\not\!\!E_{T spec}$ High $\not\!\!E_T$ control region.

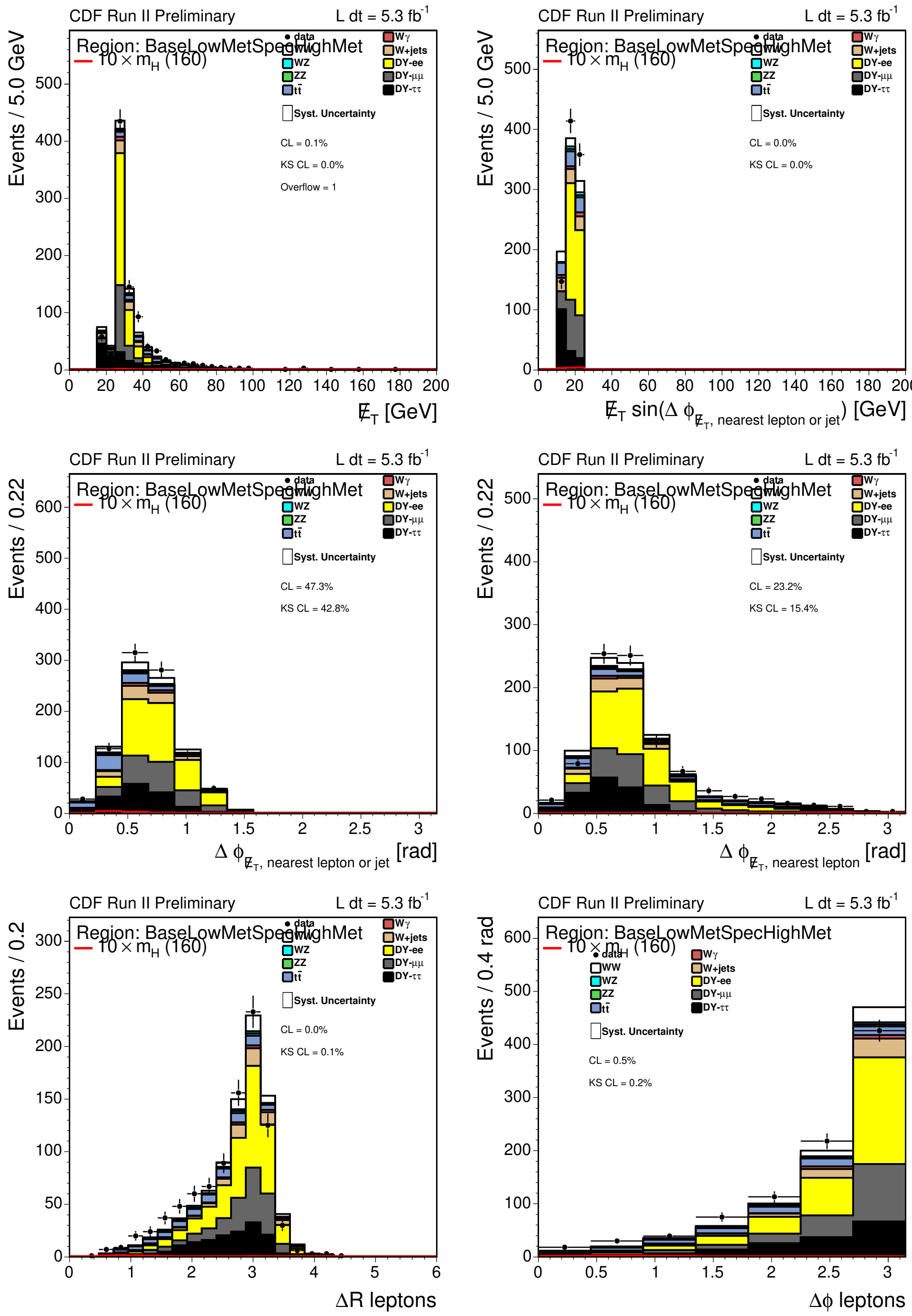

Figure 21: Base Low $\not{E}_{T spec}$ High $\not{E}_T$ control region.

6 Systematics

Sources of systematic uncertainty for this analysis are similar to those for the previous analysis using 4.8 fb^{-1} [1]. The $W\gamma$ and $Z\gamma$ uncertainties are derived in part from the studies of the $M_{ll} < 16$ GeV same sign events as described in Sec. 10.4. The fake rate systematic uncertainties have been re-evaluated for all channels, and jet energy scale systematics have been evaluated for all channels for the first time. These are described in Sections 6.2.1 and 6.2.2.

Systematic uncertainties may effect both the normalization (or "rate") and the shape of the templates used in the limit calculation. As propagating the effects of shape uncertainties is CPU intensive, we only consider shape uncertainties for those expected to be significant. At this time the shape uncertainties considered in order of expected significance are the jet energy scale (JES) and initial state radiation (ISR). These studies are described below in Section 6.5. After evaluating the shapes, we decided not to apply any shape uncertainties at this point.

The rate systematics are determined separately for each opposite-sign jet channel and separately for the same-sign channel. Systematics relating only to the same-sign channel are discussed below in Section 6.4. The summary of all systematics are presented in Tables 40-41 for the 0 jet channel, Tables 46-47 for the 1 jet channel, Tables 53-54 for the 2 or more jets channel, and Tables 74-75 for the same-sign channel.

6.1 Cross Section Systematics

For the background processes, we take the theoretical uncertainties on the cross sections, which are calculated to either next-to-next-to-leading order ($t\bar{t}$ with $m_{top} = 172.4$ GeV [11]), next-to-leading order (WW [12], WZ and ZZ [13]), or leading order with estimated higher order corrections ($W\gamma$ [14], $Z\gamma$, and Drell-Yan). Of the signal processes, VH and VBF are electroweak processes. VH are known to NNLO, so the theoretical uncertainty on these cross sections is small, less than 5% [15]. VBF is known only to NLO, so the residual theoretical uncertainty is estimated to be less than 10% [15].

Gluon fusion is a QCD process, so even though it is known to NNLO, the theoretical uncertainty is larger. Also important, since we divide our events by jet multiplicity, the theoretical uncertainty on the higher multiplicity jet bins is larger than the uncertainty on the inclusive cross section calculation. There exist several programs which compute the theoretical cross section for $gg \to H$ production at the Tevatron to NNLO. The previous zero and one jets version of this analysis used the FEHiP program [16]. For these studies we use the HNNLO program [17, 18, 19], which has the advantage that it makes user-specified histograms of kinematic variables. HNNLO also allows the user to define analysis selection cuts, but for these studies we do not use any selection cuts in HNNLO. Because the Higgs is a pseudo-scalar, we expect the effects on its kinematic variables to be limited to p_{T} and rapidity.

$\mu_r = \mu_f$	0 Jet	1 Jet	2 Jet
320	1.809	0.555	0.061
160	1.992	0.716	0.103
80	2.111	0.896	0.185
40	2.079	0.185	0.355

Table 33: Changes in the value of the cross section calculated by HNNLO when the renormalization and factorization scales (μ_r and μ_f) are varied and divided by jet bins. The exact values of the cross section are dependent on cuts made before calculations; the uncertainty is determined by the percentage change around a central value of either M_H or $M_H/2$.

There are two main independent sources of theoretical uncertainty; other sources were found to be negligible [20]. These are the following: uncalculated higher-order QCD radiative corrections, and knowledge of the parton distributions. To estimate the size of the QCD corrections, we use a method commonly used in perturbative QCD calculations, which is to vary the renormalization and factorization scales (μ_R and μ_F respectively). This was recently done in [21] using the new MSTW2008 PDFs [22] dividing the theoretical cross section uncertainties by jet bins as done in this analysis. We cross-check these results using the HNNLO program with the MSTW2008 PDFs. Comparing NNLO calculations from FEHiP to recent N^{3}LO calculations [23, 24] shows the best value for NNLO predictions is to set $\mu_R = M_H/2$. Another reasonable choice is to set the factorization and normalization scales to be equal. The default scale setting used in [21] is $\mu_R = \mu_F = M_H$, so we calculate results around this central point as well as around $M_H/2$.

To determine the uncertainty on the $gg \to H$ cross section due to higher order QCD corrections, we used $M_H = 160$. The resulting change in the NNLO Higgs p_{T} distribution is shown in Fig. 22. The cross section divided by jet bin is also shown in Fig. 22. The change in the cross section by jet bin is given in Tab. 33. We see similar uncertainties by jet bin to those calculated in [21]. Thus we take the average of the increasing and decreasing scale variations in [21] and apply them by jet bin as the systematic uncertainty on the $gg \to H$ cross section due to scale uncertainty.

The second uncertainty on the cross section is due to our knowledge of the parton distributions. For $gg \to H$ production, we are specifically interested in the gluon luminosities. We previously estimated this uncertainty using NLO PDFs with different gluon luminosities, specifically the NLO MRST99 PDFs with central gg luminosity, higher gg luminosity, and lower gg luminosity [25, 26]. The changes in the NLO Higgs p_{T} and η distributions are shown in Fig. 23. A more detailed study including $gg \to H$ electroweak corrections was done in [27] and found an expected PDF uncertainty of 7.7%. We take this 7.7% uncertainty to be the same in all jet bins.

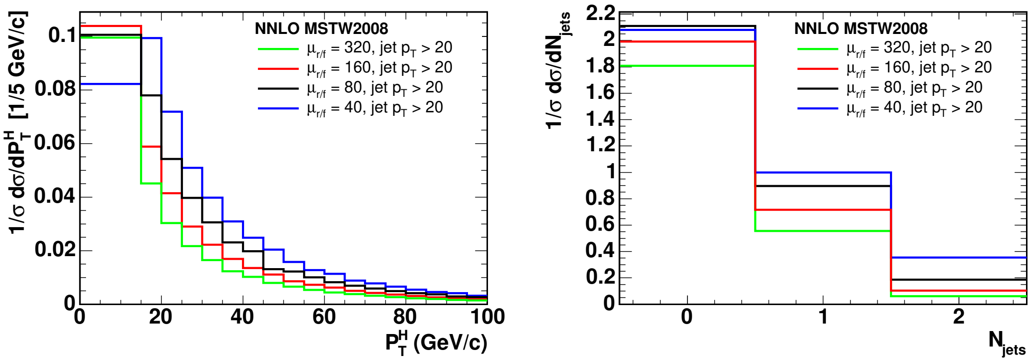

Figure 22: Changes calculated by the HNNLO program when the scales μ_R and μ_F are varied. On the left is the NNLO Higgs p_{T} distribution, while on the right is the cross section divided by jet bins.

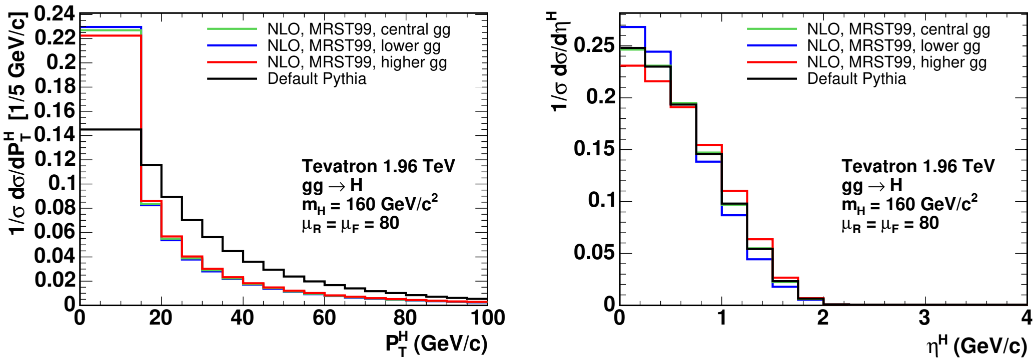

Figure 23: Changes calculated by the HNNLO program for different values of the gluon luminosity. On the left is the NLO Higgs p_{T} distribution, while on the right is the Higgs rapidity distribution. The black lines show the distributions in default $gg \to H$ Pythia with $M_H = 160$.

6.2 Acceptance Systematics

The standard luminosity uncertainty of 5.9% is assigned to all processes except for W+jets, which is determined from data. As previously, the lepton ID and trigger efficiencies are varied coherently up and down to determine their affect on the acceptance. Both uncertainties have a small effect, for the most part around 2%, on the signal and background processes. Lepton ID systematics are determined separately for each jet channel, while the trigger efficiencies are determined inclusively. Except for the $gg \rightarrow H$ signal process, as explained in Section 6.2.3, the PDF model uncertainties are determined using the standard CTEQ6M uncertainties and are the same across different channels. The PDF model systematics on VH and VBF production modes are taken from the studies done for the $H \rightarrow \tau\tau$ analysis, CDF Note 9179 [28].

We also consider the effect of the differences between LO and higher order calculations on the acceptance. We determine the size of higher order effects by comparing a LO Pythia WW sample to the WW MC@NLO sample. We take half of this difference (5%) as a systematic on the WW process, but assign the full difference (10%) to the processes simulated only by LO Pythia. For the WW MC@NLO sample, which has a restricted run range, we assign an additional uncertainty which is determined as described in the WW paragraph of Section 4.1. The $W\gamma$ and $Z\gamma$ processes have an additional 11% systematic uncertainty on the MC modeling of conversions, based upon the studies in Sec. 10.4.

The uncertainties on the modeling of $\not\!\!E_T$ has been shown in previous versions of this analysis to be very small for processes with real $\not\!\!E_T$. However, for the DY process which has no real $\not\!\!E_T$, we look in the DY control region described in Section 5.1 to estimate how well the MC models jets and $\not\!\!E_T$ for a process with no jets at LO. Based on the rate discrepancies seen in this control region, we assign a DY $\not\!\!E_T$ and jet modeling systematic of between 21-32% for each jet channel.

6.2.1 W+jets Acceptance Systematics

The only uncertainty on the W+jets process, which is determined from data, is the uncertainty on the rate at which jets will fake leptons. We estimate the effect of this uncertainty on the acceptance by varying the fake rates within their assigned uncertainties (shown by the solid light grey bands in Figures 8-9). These variations have different effects in each analysis channel and thus are calculated separately for each channel. The final assigned systematic is the average change in acceptance from varying the fake rate up by 1σ and down by 1σ, as the changes are very symmetric for all channels.

6.2.2 Jet Energy Scale Acceptance Systematics

In previous versions of this analysis involving only opposite-sign jet bins, the jet energy scale systematic was found to be a small contribution and thus not included. However, now that more channels are using cuts on the number of jets or even using N_{jets} in the neural networks, we decided to evaluate the jet energy scale systematics for all channels. This was done by varying the jet energy scale within its assigned uncertainty (up and down by 1σ) for all processes modeled by Monte Carlo, and calculating the change in acceptance for each process. This systematic also has different effects in each analysis channel and are calculated separately for each channel. In jet-inclusive channels (such as the trilepton WH channel) the effect is nearly negligible. In the opposite-sign jet bin channels, the effects are anti-correlated (changing the JES moves events between jet bins but does not add or subtract overall acceptance).

6.2.3 $gg \to H$ Acceptance Systematics

The $gg \to H$ acceptance also depends on the cross section. In the previous zero and one jet analysis, the Higgs acceptance systematic was estimated by reweighting the Higgs p_{T} spectrum in Pythia to match an NLO calculation from FEHiP, with $\mu_R = \mu_F = 82.5$. This reweighting led to acceptance changes of less than 10%. With this analysis, we again use the Higgs p_{T} spectrum from Pythia as the default and estimate the acceptance systematic by reweighting Pythia to match the calculations from HNNLO. This time, we reweight in the Higgs p_{T} and also, separately, the Higgs η distribution.

The effect of the Higgs p_{T} reweighting depends on how many jets are in the final state. By reweighting the Higgs p_{T} spectrum, we are reweighting the vector sum of the p_{T} of the two jets which are recoiling against the Higgs. The Pythia Higgs p_{T} spectrum is harder than that generated by HNNLO, so when we reweight Pythia to match HNNLO, events which previously had one or more jets passing selection cuts will move into a lower jet bin. There is more signal in the 0 and 1 jet bins, so the fractional percentage change there is smaller than in the 2+ jet bin. Because this reweighting is effectively moving events back and forth between the 0, 1, and 2+ jet bins, these acceptance systematics (referred to as “Scale (jets)” and “PDF Model (jets)” in the systematics tables) are anticorrelated between the jet bins.

The effect of the Higgs η distribution reweighting also depends on jet bin, again because the fractional percentage change is larger in the 2+ jets bin where there are fewer signal events to begin with. However, this reweighting does not appear to move events between different jet bins. The Higgs η distribution in Pythia is more central than the one generated by HNNLO, so when we reweight Pythia to match HNNLO, we move some of the leptons outside of our forward acceptance. These acceptance systematics, referred to as “Scale (leptons)” and “PDF Model (leptons)” in the systematics tables, are correlated among the jet bins.

For the acceptance systematic from the choice of scale, we use our default NNLO calculation with MRST2004 PDFs and $\mu_R = \mu_F = 80$. We divide the Higgs p_{T} and η distributions from HNNLO by the same distributions from our default $gg \to H$ Pythia sample, at $M_H = 160$. We then reweight the Pythia sample using this ratio histogram. The ratio histograms for the MRST2004 HNNLO sample are shown in Fig. 24. The changes in acceptance after reweighting are given in Tab. 34 (top). As expected, reweighting in η moves events out of our acceptance range, resulting in a decrease of events in all jet bins.

We also use HNNLO to calculate a systematic on the Higgs acceptance due to gluon luminosity/PDF model. This time we use the NLO MRST99 PDFs with $\mu_R = \mu_F = 80$. Again we divide the HNNLO distributions by Pythia and use the ratio histograms to reweight the Pythia. This is an older NLO calculation, so we do not look at the acceptance change from the default Pythia to each of the three gg luminosity models. Instead, we reweight Pythia to match the central gg model and take that as our default for these studies. Then we reweight Pythia to match the higer and lower gg models, and look at the acceptance change from the central gg model. The ratio histograms for all three HNNLO samples are shown in Fig. 25. The changes in acceptance after reweighting are given in Tab. 34. As expected, the total uncertainty is similar for reweighting the gg luminosity up or down. As the final $gg \to H$ acceptance systematic due to the PDF model, we use the average of the absolute value of the changes with higher and lower gluon luminosities.

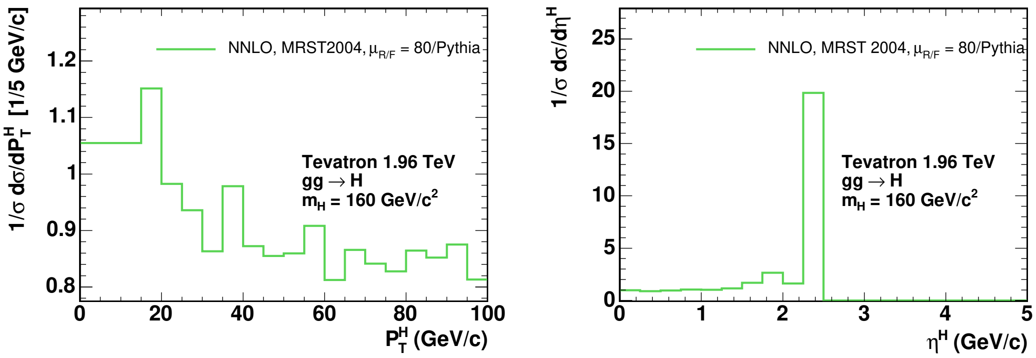

Figure 24: Ratio between default $gg \to H$ Pythia and the calculations of the HNNLO program using the NNLO MRST2004 PDFs with $M_H = 160$ and the scales $\mu_R = \mu_F = 80$. On the left is the ratio used to reweight the Higgs p_{T} distribution, while on the right is the ratio used to reweight the Higgs rapidity distribution.

6.3 2 or More Jets Systematics

As explained in Section 9, the signal region with two or more reconstructed jets employs b-tagging to reduce $t\bar{t}$ background. When we apply b-tagging to our $t\bar{t}$ background model, we use the standard Data/MC efficiency scale factor correction for b-tagging which we take to be 0.94 ± 0.03. This uncertainty is taken as an additional systematic, under the heading b-tag Veto as shown in Table 53.

There is another systematic uncertainty to consider. We do not require data events to be in the silicon good run list. The only issue with this that could arise is that our

	Choice of scale	
	Fractional (%) acceptance change	
Jet bin	Reweight p_{T}	Reweight η
0 jet	+4.60%	−2.46%
1 jet	−5.07%	−2.78%
2+ jets	−8.74%	−3.05%
	Gluon luminosity variations	
	Fractional (%) acceptance change	
Lower gg	Reweight p_{T}	Reweight η
0 jet	+0.76%	+1.57%
1 jet	−1.56%	+1.75%
2+ jets	−2.32%	+2.06%
Higher gg	Reweight p_{T}	Reweight η
0 jet	−1.09%	−1.49%
1 jet	+2.19%	−1.65%
2+ jets	+3.18%	−1.90%
Average Change	Reweight p_{T}	Reweight η
0 jet	0.93%	1.53%
1 jet	1.88%	1.70%
2+ jets	2.75%	1.98%

Table 34: Fractional (%) change in the $gg \to H$ acceptance in different jet bins due to the cross section. The top rows show the change in the acceptance after reweighting the Pythia Higgs p_T and η spectra to match the spectra from HNNLO, using the MRST2004 PDFs and $\mu_R = \mu_F = 80$ for $M_H = 160$. The bottom rows show the change in the acceptance after reweighting the Pythia Higgs p_T spectrum to PDFs with higher or lower gluon luminosities. The **Average Change** is the average of the absolute value of the changes with higher and lower gluon luminosities. The average changes are used for the gluon PDF systematics.

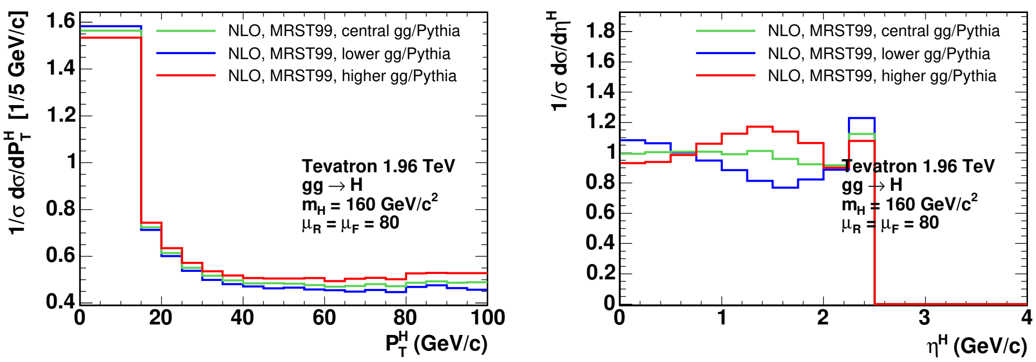

Figure 25: Ratio between default $gg \to H$ Pythia and the calculations of the HNNLO program using the NLO MRST99 PDFs with different gluon luminosities. On the left are the ratios used to reweight the Higgs $p_{\rm T}$ distribution, while on the right are the ratios used to reweight the Higgs rapidity distribution.

Data/MC scale factor may not be properly scaling for events which fail b-tag requirements because they do not have silicon hits; the MC assumes there is always good silicon. To account for this, we look at the number of events in the data sample which come in on each good run list, as shown in Table 35. There are only 2 events not in the silicon good run list which have a b-tagged jet as found by SecVtx.

We then look only at the events in the silicon good run list. In this sample, the ratio of anti-b-tagged events to b-tagged events is 248:182. If the same ratio held true for events not in the silicon good run list (as the MC scale factor assumes it does), we would expect of the 18 events not in the silicon good run list that the ratio of anti-b-tagged events to b-tagged events should be 10.4:7.6. The observed ratio is actually 16:2, as shown in Table 35. Continuing with the thought, if we had only good Si runs we would expect $248 + 10.4 = 258.4$ events in the anti-b-tagged region, while our data show 264 events. In the same way, we expect $184 + 7.6 = 189.6$ events in the b-tagged region but see only 184 events.

We therefore correct our MC prediction in the 2 or more jets signal region by scaling it up by the ratio $264/258.4 = 1.023$, and in the b-tagged $t\bar{t}$ control region by scaling it down by $184/189.6 = 0.970$. We take about half of the change in the signal region as an additional systematic to be added in quadrature with the b-tag Veto uncertainty. Note that this scale factor was reduced by including the reprocessed period 18 data, which is now included on the silicon good run list.

6.4 Same-Sign Systematics

Most of the systematics for the same-sign channel are the same as those in the opposite sign channels. Because we do not make a $\not\!\!E_{\rm T}$ cut, we do not have a systematic due to $\not\!\!E_{\rm T}$

2+ Jets Region	Number of Events
Total	448
Total b-tagged	184
Total anti-b-tagged	264
Total not in EM_SI	18
Total b-tagged not in EM_SI	2
Total anti-b-tagged not in EM_SI	16

Table 35: Summary of events with and without a b-tagged jet in the 2 or more jets signal region and which good run list they are selected in.

modeling. We have one extra systematic based on the ability of Monte Carlo to model charge mismeasurement, as described below.

WW, Drell-Yan, and $t\bar{t}$ contribute to the same-sign analysis only when the charge of a lepton is mismeasured[2]. These backgrounds are estimated using Monte Carlo samples; this assumes that the MC accurately models how often lepton charges are mismeasured in the data.

To test this assumption, we attempt to measure a charge fake rate from the data. We select events in the Drell-Yan control region, described in Sec. 5.1, which should have minimal contamination from other backgrounds such as W+jets, where same-sign events come from QCD rather than charge mismeasurement. We then repeat the procecure of counting the number of opposite-sign events and same-sign events, resulting in charge fake rates taken directly from the data. These CFR are shown in Table 36. Unfortunately the statistics are not large enough to reliably measure a CFR in the muon channels or for PHX electrons.

To determine the uncertainty on MC modeling of the charge fake rate, we compare the LBE-LBE and LBE-CrkTrk charge fake rates measured in data to those measured in MC. We subtract the MC CFR from the data CFR, and take half of that difference as the systematic. For LBE-LBE, this gives a systematic uncertainty of 23.6%, while for LBE-CrkTrk the systematic uncertainty would be 16.1%. We take an uncertainty of 23.6% and apply it to the WW, $t\bar{t}$, and Drell-Yan backgrounds as shown in Table 74.

6.5 Shape Systematics

At this time we consider two sources of systematics which would change the shape of the output NN distribution: the fake rate and the jet energy scale (JES) variations. The only channel in which W+jets is the dominant background contribution is in the same-sign channel. In this channel, the rate uncertainty from varying the fake rates by $\pm 1\sigma$ is 33%.

[2] The case of $t\bar{t}$ where one lepton is missed and one of the b-jets fakes a lepton is included in the W+jets background.

Dilepton type	Opposite-sign events	Same-sign events	Charge fake rate
LBE-LBE	110392	157	1.42e-3
LBE-Central μ	90	6	0.0667
LBE-Forward μ	16	1	0.0625
LBE-CrkTrk	46984	36	7.66e-4
Central μ-Central μ	119741	0	–
Central μ-Forward μ	34836	5	1.44e-4
Central μ-CrkTrk	42374	2	4.72e-5
PHX-PHX	1010	307	0.304
PHX-Central μ	87	57	0.655
PHX-Forward μ	13	12	0.923
PHX-CrkTrk	9402	1288	0.137
LBE-PHX	101760	12770	0.125

Table 36: Calculated charge fake rates (CFR) for all dilepton types in data, using the Drell-Yan control region. The lepton selection is described in the text. "Central μ" includes CMUP, CMP, CMX, CMXMsKs, CMIOCES, and CMP muons, while "Forward μ" includes CMIOPES and BMU muons.

However, as shown in Figure 26, when the neural network output tempates are normalized to the same area, there is no change in shape from varying the fake rate.

The jet energy scale systematic described in Sec. 6.2.2 may also be considered as a shape systematic. The effect is small, as seen when these shape systematics were considered for the trilepton ZH channels (CDF Note 10020 [3]). At this time, for technical reasons the JES variations have not been computed for all signal samples. Thus we do not consider the JES as a shape systematic in this updated analysis, but will consider it for the next version.

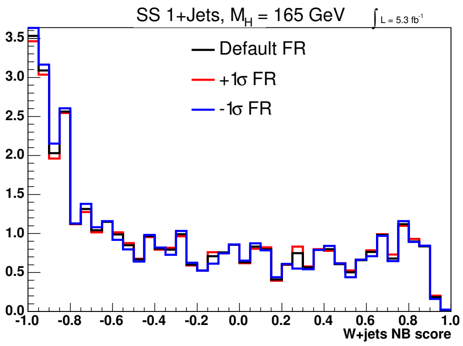

Figure 26: Effect of ± 1 σ variations in the fake rates in the same-sign channel. The default shape is shown in black, while the higher fake rate is shown in red and the lower fake rate in blue. All three templates have been normalized to the area of the default template, which shows there is no visible shape change from modifying the fake rate.

7 Opposite-sign Dilepton Signal Region with 0 Jets

The background expectation and observed yield for the opposite-sign dilepton signal region for events containing no reconstructed jets with $E_T > 15$ GeV are shown in Table 37. The expected and observed kinematic distributions for these events are shown in Figures 27-29. We stack the backgrounds on top of each other, and then for comparison we overlay the contribution from a 160 GeV/c^2 Higgs boson with yields scaled up by a factor of 10. In addition, we calculate a confidence level for data versus background expectations on each plot. The confidence levels shown are for Poisson fluctuations and the KS test.

7.1 Understanding Drell-Yan yield increase

Comparing the yields in Table 37 with the previous version of the analysis in the opposite-sign signal region with 0 jets [1], we see a higher increase in Drell-Yan yields than we expected from adding new data and relaxing the LBE cut to 0.90. We expected to see $\sim 10\%$ increase due to luminosity and another $\sim 5\%$ due to the lower cut on the likelihood value for LBE (ie. $\sim 10\%$ increase for LBE-LBE category). So overall we expected to see an increase of 20% for the LBE-LBE category and 15% for the LBE-PHX category. However, we actually see yield increases of about 36% in the LBE-LBE category and about 20% in the LBE-PHX category.

We tried to understand the reason behind this. The LBE efficiency is measured in a DY dominated region where most of the electrons are pretty well isolated, as shown in see Fig. 30. Looking at the the yield increase from the previous version to the current version of the analysis (i.e. from 4.8 fb^{-1} to 5.3 fb^{-1} analysis) as a function of TrkIso for LBEs in Fig. 31, we see that by relaxing the likelihood LBE cut we gain mostly LBE electrons with larger value of isolation (i.e. less isolated electrons), which is expected.

When we go to the signal region, the DY events are different. Since here we require $\not{E}_{T spec} > 15$ or 25 GeV, DY events pass the selection only by having jets in the final state (reconstructed or not). The isolation of electrons is worse, see Fig. 32. Using this distribution and the expected gain per TrkIso bin as seen in DY region (Fig. 31), we can estimate the expected yield increase of Drell-Yan events in the opposite sign signal region with zero jets. This gives us 10% (12%) efficiency increase for LBE using TrkIso (CalIso) dependence. Overall, we then expect $10\% + 2 * 10(12) = 30(34)\%$ yield increase for LBE-LBE category (20%(22%) for LBE-PHX category) in OS 0Jets signal region which is in very good agreement to what we actually observe.

7.2 Incorporating Additional Signal in the 0 Jet Channel

In CDF Note 9163 it was shown that using Matrix Element (ME) calculations as additional inputs to a NeuroBayes® (NB) neural network improved the overall sensitivity of the analysis. Since the ME calculations provide a model for each process based on the Leading

Category	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jets	Total	Data
e e	99.3	7.1	14.8	0.5	67.4	74.1	76.3	339.5	354.0
e μ	226.3	4.5	0.5	1.0	24.4	63.3	81.3	401.2	396.0
μ μ	73.8	6.1	10.8	0.3	34.7	0.0	11.2	136.9	123.0
e trk	67.6	3.1	5.2	0.3	20.8	14.6	20.6	132.4	162.0
μ trk	37.8	2.1	2.9	0.2	8.3	2.3	12.6	66.3	72.0
Total:	504.9	22.8	34.2	2.3	155.7	154.3	202.0	1076.2	1107.0

Category	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jets	Total	Data
LBE LBE	57.2	3.8	8.9	0.3	37.2	18.4	16.7	142.4	147.0
LBE PHX	35.7	2.8	5.2	0.1	24.8	38.8	44.9	152.4	166.0
PHX PHX	6.4	0.4	0.8	0.0	5.4	16.9	14.8	44.7	41.0
LBE CMUP	64.3	1.1	0.1	0.3	5.5	10.0	16.0	97.3	102.0
LBE CMP	9.2	0.2	0.0	0.1	0.7	1.3	3.6	15.0	12.0
LBE CMX	28.5	0.6	0.1	0.1	2.1	5.0	8.2	44.6	32.0
LBE CMXMsKs	3.4	0.1	0.0	0.0	0.7	0.8	1.3	6.3	4.0
LBE BMU	7.5	0.2	0.0	0.0	0.5	0.8	0.0	9.0	12.0
LBE CMIOCES	45.8	0.8	0.1	0.2	2.4	3.9	3.0	56.2	47.0
LBE CMIOPES	14.2	0.3	0.0	0.0	1.4	1.5	6.9	24.3	18.0
PHX CMUP	19.7	0.5	0.0	0.1	3.9	15.9	18.8	58.9	71.0
PHX CMP	2.7	0.1	0.0	0.0	0.8	1.9	3.0	8.6	5.0
PHX CMX	9.6	0.2	0.0	0.0	3.8	10.2	9.9	33.8	37.0
PHX CMXMsKs	1.2	0.0	0.0	0.0	0.3	1.8	1.9	5.1	4.0
PHX BMU	2.7	0.1	0.0	0.0	0.6	1.5	0.0	4.9	9.0
PHX CMIOCES	12.8	0.3	0.0	0.0	1.5	5.9	1.8	22.4	29.0
PHX CMIOPES	4.7	0.1	0.0	0.0	0.3	2.8	2.7	10.6	14.0
CMUP CMUP	11.7	0.9	1.8	0.1	4.2	0.0	1.2	19.9	19.0
CMUP CMP	3.3	0.3	0.5	0.0	1.3	0.0	0.3	5.6	4.0
CMUP CMX	11.8	0.9	1.6	0.1	3.8	0.0	1.4	19.6	13.0
CMUP CMXMsKs	1.3	0.2	0.3	0.0	0.8	0.0	0.2	2.7	4.0
CMUP BMU	3.0	0.2	0.4	0.0	2.3	0.0	0.3	6.3	8.0
CMUP CMIOCES	17.1	1.3	2.4	0.1	6.4	0.0	1.4	28.7	28.0
CMUP CMIOPES	5.2	0.5	0.7	0.0	4.0	0.0	2.9	13.3	10.0
CMX CMP	1.6	0.1	0.2	0.0	1.2	0.0	0.1	3.2	5.0
CMX CMX	2.3	0.2	0.3	0.0	1.4	0.0	0.2	4.5	2.0
CMX CMXMsKs	0.8	0.1	0.2	0.0	0.2	0.0	0.1	1.3	1.0
CMX BMU	1.1	0.1	0.2	0.0	0.8	0.0	0.1	2.2	3.0
CMX CMIOCES	7.6	0.6	1.1	0.0	3.8	0.0	0.6	13.7	11.0
CMX CMIOPES	2.5	0.2	0.3	0.0	1.4	0.0	1.4	5.8	5.0
CMXMsKs CMP	0.3	0.0	0.1	0.0	0.3	0.0	0.1	0.7	1.0
CMXMsKs CMXMsKs	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CMXMsKs BMU	0.2	0.0	0.0	0.0	0.5	0.0	0.0	0.7	0.0
CMXMsKs CMIOCES	0.6	0.1	0.2	0.0	0.7	0.0	0.1	1.7	2.0
CMXMsKs CMIOPES	0.2	0.0	0.0	0.0	0.4	0.0	0.1	0.8	0.0
CMP CMP	0.3	0.0	0.0	0.0	0.3	0.0	0.1	0.7	2.0
CMP BMU	0.3	0.0	0.0	0.0	0.3	0.0	0.0	0.6	2.0
CMP CMIOCES	2.0	0.1	0.3	0.0	0.6	0.0	0.2	3.2	3.0
CMP CMIOPES	0.6	0.1	0.1	0.0	0.2	0.0	0.5	1.4	0.0
LBE CrkTrk	52.1	2.3	4.1	0.3	14.8	6.9	13.2	93.7	117.0
PHX CrkTrk	15.5	0.8	1.1	0.0	6.0	7.7	7.4	38.6	45.0
CMUP CrkTrk	23.3	1.2	1.8	0.1	5.0	1.3	8.2	41.0	53.0
CMP CrkTrk	2.5	0.1	0.2	0.0	0.6	0.1	1.0	4.5	4.0
CMX CrkTrk	10.8	0.6	0.8	0.1	2.5	0.8	3.0	18.6	10.0
CMXMsKs CrkTrk	1.2	0.1	0.1	0.0	0.3	0.1	0.5	2.3	5.0
Total:	504.9	22.8	34.2	2.3	155.7	154.3	197.9	1072.1	1107.0

Table 37: Yields in opposite-sign dilepton signal region for events with zero jets.

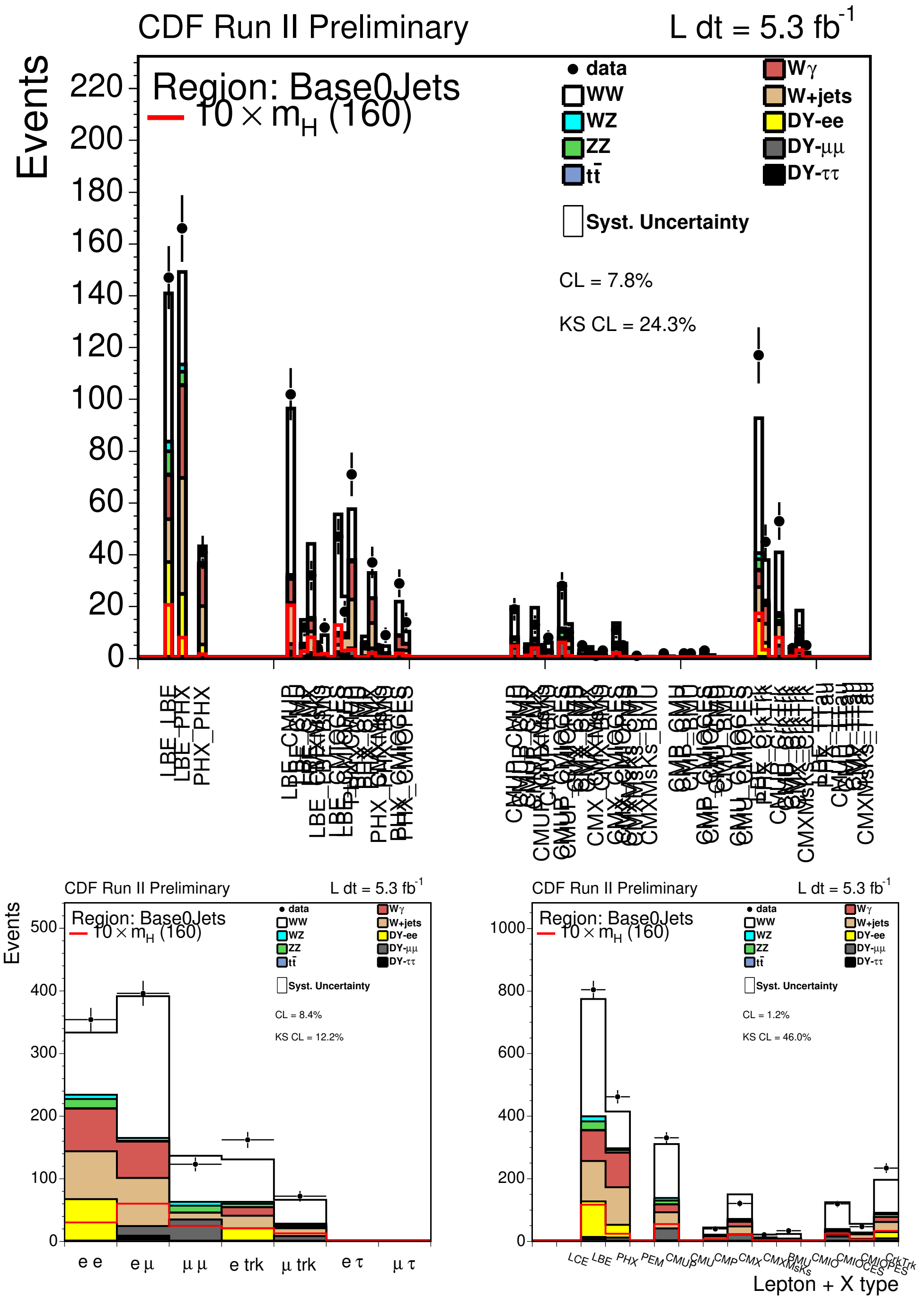

Figure 27: Signal and background in the opposite-sign signal region for events containing no reconstructed jets with $E_T > 15$ GeV.

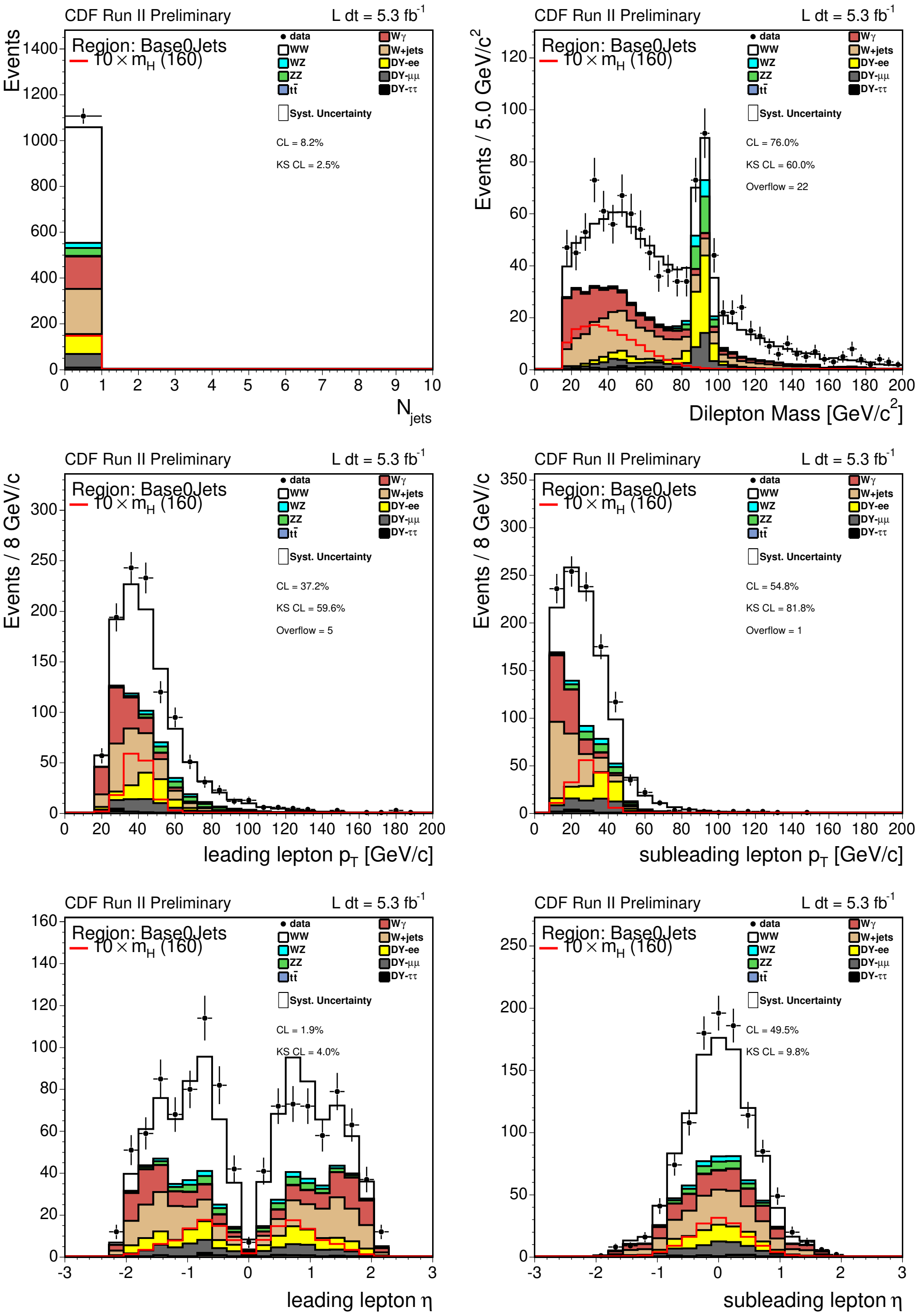

Figure 28: Signal and background in the opposite-sign signal region for events containing no reconstructed jets with $E_T > 15$ GeV.

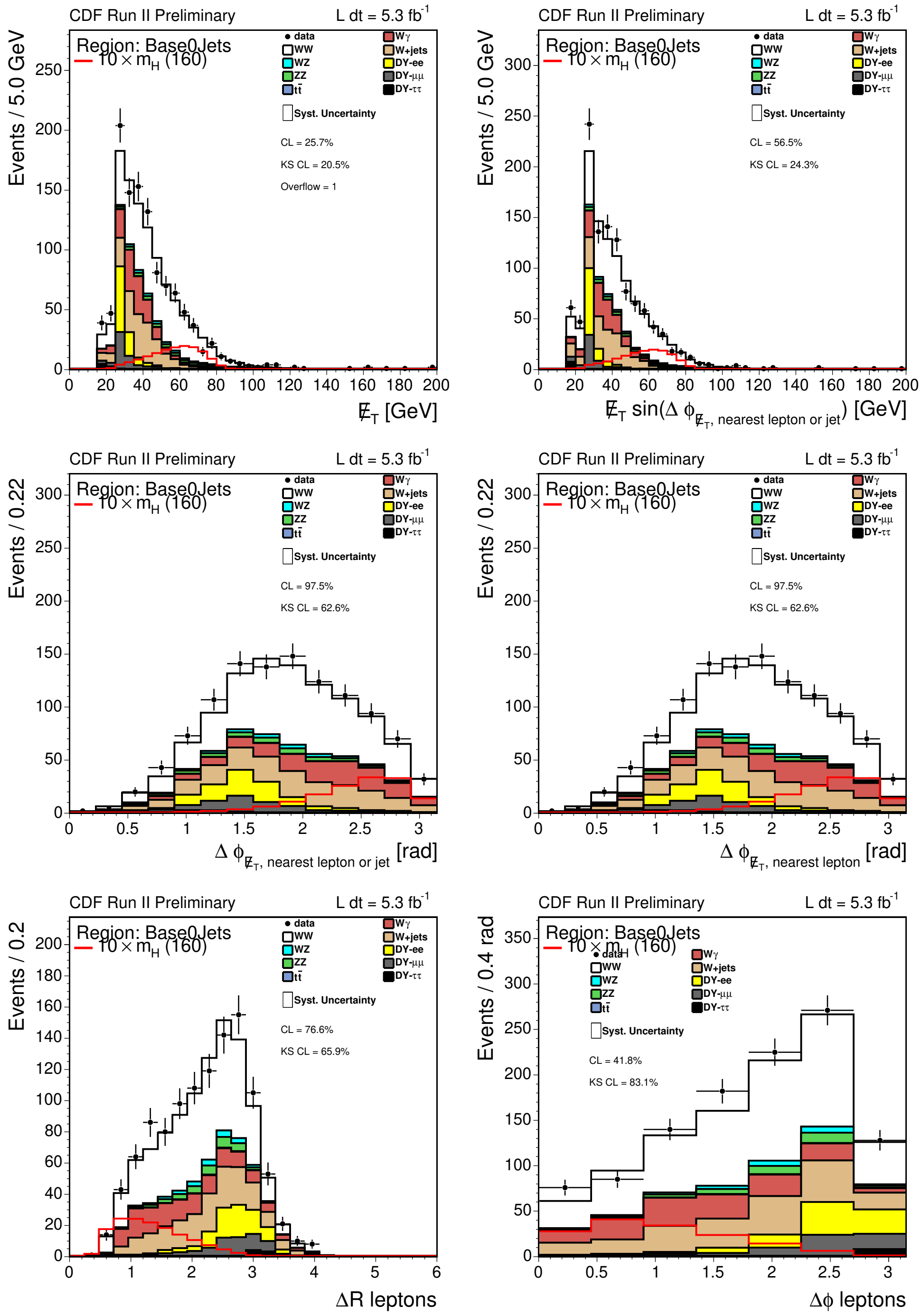

Figure 29: Signal and background in the opposite-sign signal region for events containing no reconstructed jets with $E_T > 15$ GeV.

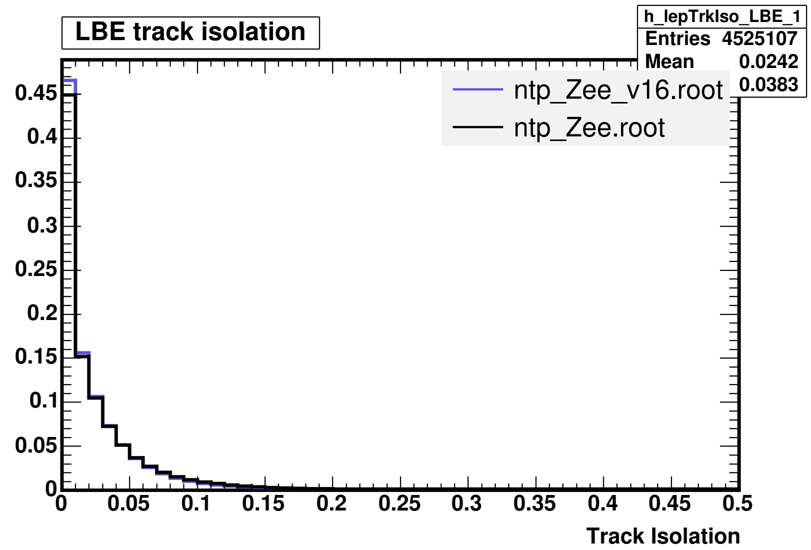

Figure 30: The track isolation distribution for LBE electrons from $Z \to ee$ sample for events passing DY region selection (blue histogram from the v16 or 4.8 fb^{-1} analysis, black histogram from the v17 or 5.3 fb^{-1} analysis).

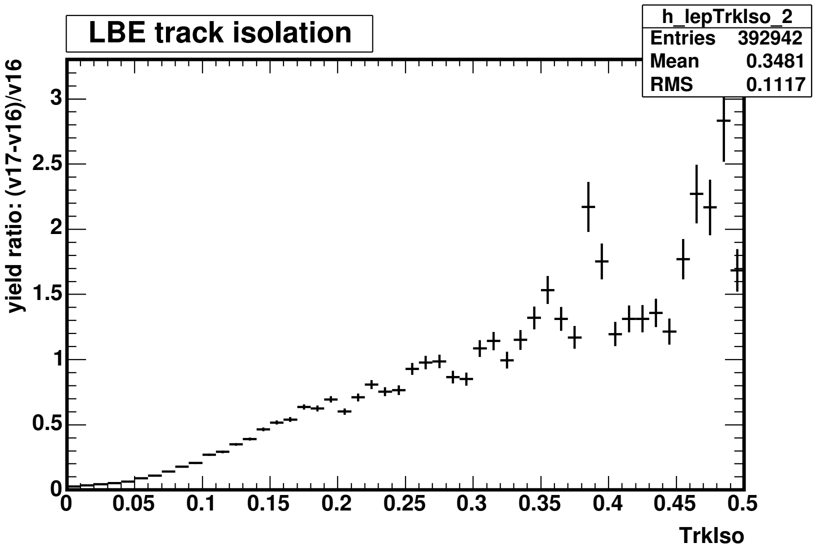

Figure 31: The fractional gain (i.e. ratio = (Yield(v17) - Yield(v16))/Yield(v16)) in yield between v16 and v17 version of analysis shown as dependence on TrkIso value of LBEs (again for LBE electrons from $Z \to ee$ sample for events passing DY region selection).

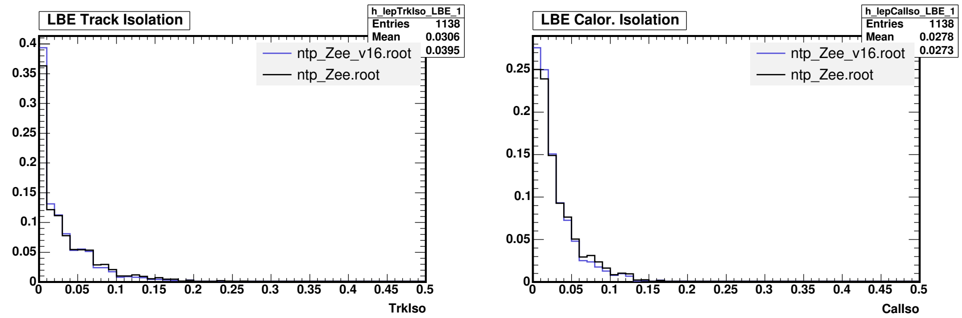

Figure 32: The track (left) and calorimeter (right) isolation distribution for LBE electrons from $Z \to ee$ sample for events passing opposite sign with zero jets signal region selection.

Order (LO) behavior of the differential cross section, we expect that using this information in the neural network works well in cases where the events are LO-like, i.e. when no additional jets are present. We therefore use, for events with 0 jets (0J), ME-based Likelihood Ratios as inputs for a NB neural network in analogy of what has been done previously and described in detail in CDF Note 9163.

In all previous versions of the analysis, the only signal contribution considered was that from gluon fusion production. The primary reason behind this was the considerable time it takes to the run the ME calculations. Incorporating the WH, ZH, and VBF signals at each mass point means running an extra 57 ME calculations. However, the benefit from the extra signal is not negligible, particularly for low Higgs boson mass. At $m_H = 115$ GeV, the associated production and VBF together contribute 16% to the total expected signal. At higher masses the gain is less (6.2% for $m_H = 160$ and 6% for $m_H = 200$ GeV) but still not inconsiderable.

For this update, we have run the default ME calculations for the additional signal production modes. At this time, we have not introduced ME calculations specifically for the new modes (WH, ZH, and VBF), although that could be instituted as an update in the future. We again use the ME-based Likelihood Ratios for WW and $gg \to H \to WW$ as neural network input variables. The neural network training is done on a weighted combination of the four signal production modes; to help discriminate the new signal contributions from the backgrounds, we introduce several additional kinematic variables to the neural network. Previously, the only kinematic variable inputs to the neural network were the ΔR and $\Delta\phi$ between the leptons and the transverse mass H_T. The neural network is now trained including the variables: dilepton mass, the transverse momentum of the leading and subleading leptons, and the transverse mass of the sum of the leptons' 4-momentums and the $\not{E}_T$.

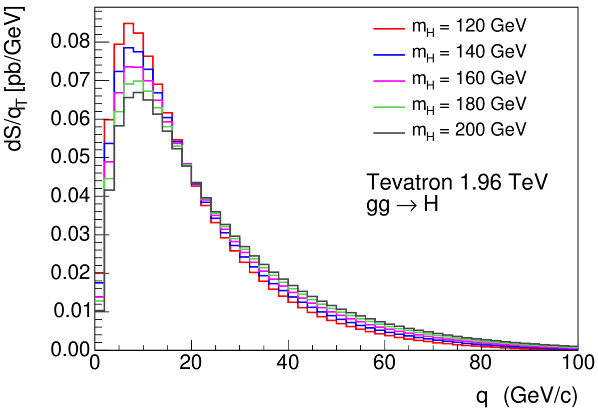

Figure 33: Distributions of $gg \to H$ p_T using the HqT program, plotted for Higgs masses of 120, 140, 160, 180, and 200 GeV. Each Pythia MC sample is reweighted to the appropriate mass-dependent spectrum.

7.3 Reweighting of the Gluon Fusion Signal

Particularly for events with zero reconstructed jets, where the dominant background is direct WW production, the most important discriminating variables is the $\Delta\phi$ between the leptons. In order to predict the $\Delta\phi$ distribution for a given Higgs mass, it is important to correctly model the Higgs boost. Previous versions of this analysis made no corrections to the Higgs p_T spectrum from Pythia, although the HNNLO program was used to derive acceptance uncertainties as described in Section 6.2.3.

Recent theoretical progress has brought us the HqT program [29, 30]. The HNNLO program computes the cross-section for $gg \to$ Higgs production at NNLO in QCD perturbation theory. The HqT program, written by most of the same authors as HNNLO, computes the transverse momentum distribution for $gg \to$ Higgs production at NNLL+NLO accuracy, with the normalization fixed to the total NNLO cross-section. For this version of the analysis, we use HqT to produce transverse momentum distributions for each Higgs mass we consider, and then reweight the Pythia $gg \to H$ transverse momentum distributions.

Several HqT distributions at different Higgs masses are shown in Figure 33. In general, the HqT spectrum peaks at a higher p_T value than the default Pythia spectrum, but HqT also has fewer events at $p_T > 60$ GeV. The net effect of the reweighting is to slightly increase the Higgs acceptance (due to the higher peak value) and to have fewer of the Higgs events with jets (due to having fewer events at high Higgs p_T). Thus the reweighting has the effect of increasing the expected Higgs signal in the 0 jet channel, while slightly reducing the expected Higgs signal in the 1 and 2 or more jet channels.

7.4 Analysis of Opposite-Sign 0 Jet events

To optimize the choice of variables to be used as input to the neural network we begin from a large set of what we think can be discriminant variables; we then evaluate a *ranking* associated to each of them using a NB pre-training functionality [31], which gives an estimation of the discrimination power between signal and background that this variable carries. We then train a set of neural networks starting from the full list of input variables and removing one variable at time, starting from the less significant. We calculate the expected limits for each output template obtained by these networks and finally choose the set of variables which minimize the expected limit; for this study a signal with $m_H = 160$ GeV was used.

The training method used for the neural network is mostly the same as described in CDF Note 9136. We train the net to distinguish the Higgs MC sample from an admixture of each background (both MC-derived and data-driven backgrounds); each background event has a weight according to its relative expectation, which is a function of the cross section, multiple scale factors, trigger efficiencies, and MC statistics. Only for the training, we then reweight all events so that the sum of the weights is equal to the number of signal events. The only significant differences with respect to the previous trainings is the use of a newer version of the NB package (v2.1) and the use of the BFGS algorithm in the minimization. The input variables used for the 0-jet NN are shown in Table 7.4 along with a brief explanation of their meaning; for the 0J net we use 9 input nodes, 11 hidden nodes and 1 output node. The observed distributions of each 0J neural network input parameter compared with the distributions predicted by our background model are shown in Figure 34. For reference, we also include a hypothetical signal component for a Higgs boson with $m_H = 160$ GeV.

Variable	Meaning
	0J Neural Network
LRHWW	Likelihood ratio - $H \to WW$
dRLeptons	ΔR between the leptons
LRWW	Likelihood ratio - WW production
dPhiLeptons	$\Delta\phi$ between the leptons
Ht	Scalar sum of leptons E_T and $\not\!\!E_T$ (no jets in the event)
Mt_H	Transverse mass of the sum of lepton's 4-momenta and ($\not\!\!E_T$, MetX, MetY, 0) where MetX/Y are the X/Y components of the $\vec{\not\!\!E}_T$
dimass	Dilepton invariant mass
lep1_Pt	Transverse momentum of the first lepton
lep2_Pt	Transverse momentum of the second lepton

Table 38: Summary of inputs to 0J neural networks.

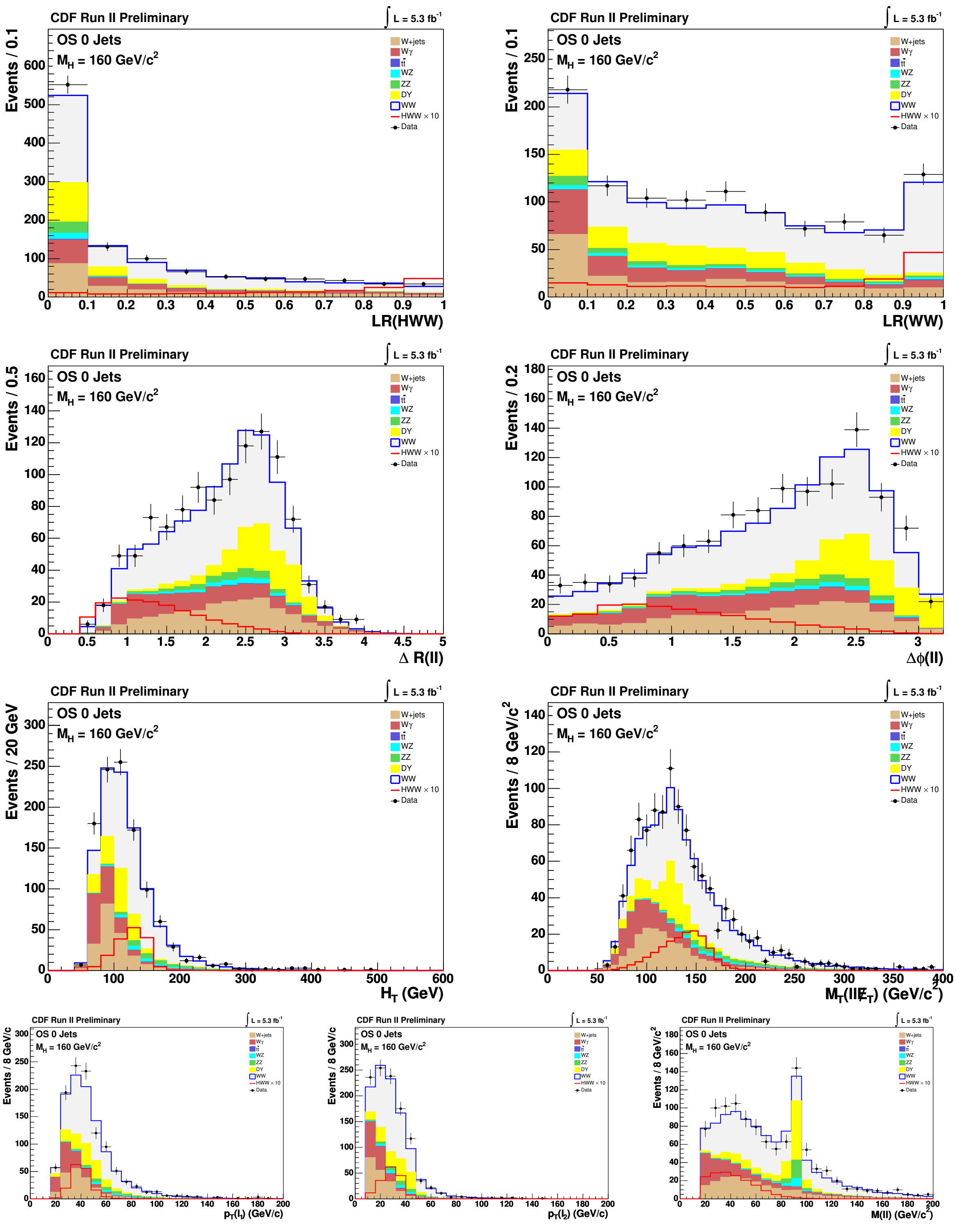

Figure 34: Neural Network input variables for 0J NN (signal for $M_H = 160$ GeV$/c^2$).

Mass (GeV/c^2)	gg→H	WH	ZH	VBF
110	0.874 (0.71/0.164)	0.12 (0.0945/0.0256)	0.0471 (0.0383/0.00878)	0.0044 (0.00367/0.000733)
115	1.71 (1.39/0.322)	0.171 (0.135/0.0359)	0.0763 (0.0631/0.0132)	0.00874 (0.00726/0.00148)
120	2.92 (2.37/0.545)	0.223 (0.177/0.0467)	0.114 (0.0943/0.0198)	0.0158 (0.0131/0.00268)
125	4.54 (3.69/0.852)	0.275 (0.221/0.0542)	0.162 (0.136/0.0267)	0.0257 (0.0213/0.00445)
130	6.33 (5.15/1.17)	0.326 (0.261/0.065)	0.213 (0.177/0.0352)	0.0385 (0.0317/0.00677)
135	8.19 (6.68/1.51)	0.363 (0.292/0.0708)	0.255 (0.212/0.0436)	0.0518 (0.0428/0.00894)
140	10.1 (8.17/1.91)	0.384 (0.309/0.0746)	0.3 (0.251/0.0485)	0.0659 (0.0548/0.0111)
145	11.6 (9.38/2.19)	0.396 (0.318/0.0782)	0.33 (0.277/0.0526)	0.0807 (0.0669/0.0138)
150	12.9 (10.6/2.29)	0.396 (0.322/0.0735)	0.36 (0.304/0.0559)	0.0925 (0.0769/0.0156)
155	14.1 (11.6/2.53)	0.398 (0.324/0.0739)	0.375 (0.314/0.0608)	0.105 (0.0873/0.0178)
160	15.4 (12.6/2.73)	0.391 (0.318/0.0728)	0.389 (0.328/0.0611)	0.121 (0.101/0.02)
165	15.1 (12.4/2.69)	0.364 (0.297/0.0675)	0.373 (0.317/0.0554)	0.126 (0.105/0.0208)
170	13.7 (11.3/2.4)	0.321 (0.261/0.0598)	0.349 (0.295/0.0539)	0.121 (0.101/0.0203)
175	12.3 (10.1/2.24)	0.288 (0.234/0.0534)	0.312 (0.264/0.0483)	0.115 (0.0956/0.0191)
180	10.8 (8.88/1.95)	0.248 (0.203/0.0451)	0.274 (0.232/0.042)	0.105 (0.0869/0.0181)
185	8.78 (7.17/1.61)	0.199 (0.164/0.0348)	0.221 (0.187/0.0344)	0.089 (0.0737/0.0153)
190	7.32 (5.99/1.33)	0.161 (0.132/0.0284)	0.187 (0.157/0.0294)	0.0777 (0.064/0.0137)
195	6.4 (5.2/1.2)	0.139 (0.114/0.0251)	0.162 (0.137/0.0256)	0.0707 (0.0579/0.0128)
200	5.73 (4.63/1.1)	0.12 (0.099/0.0211)	0.143 (0.121/0.0224)	0.0637 (0.0519/0.0117)

Table 39: Expected Higgs Signal yields as a function of the Higgs boson mass in the 0J region for each of the contributing production processes. The numbers inside the parentheses are the expected relative contributions in the High S/B and Low S/B regions.

Templates are created for each neural network and used for calculating 95% CL limits with the MCLimit program [32]. Events are separated into Low S/B and High S/B channels based on the dilepton types and a separate set of templates is created for each of the channels. MCLimit combines the two channels to produce a single set of limits for the 0J region. Output templates for the trained 0J neural network, divided into Low S/B and High S/B, are shown in Figures 35−39.

In the 0J region we now consider potential signal contributions from four Higgs production channels: gluon fusion ($gg \rightarrow H \rightarrow WW \rightarrow ll\nu\nu$), Vector Boson Fusion (VBF), and associated production VH ($q\bar{q} \rightarrow ZH \rightarrow ZWW$ and $q\bar{q} \rightarrow WH \rightarrow WWW$). We evaluate the number of events expected from each production process as a function of the mass of the Higgs. These numbers are shown in Table 39.

In calculating the final limits, we take into account systematic uncertainties on both the signal and background models we use to train our neural networks. Systematic uncertainties for the 0J region are summarized in Tables 40 and 41.

Table 42 and Figure 40 show the expected (based on 10000 pseudo-experiments) and observed limits obtained from the 0J region, incorporating the systematic uncertainties listed in Tables 40 and 41.

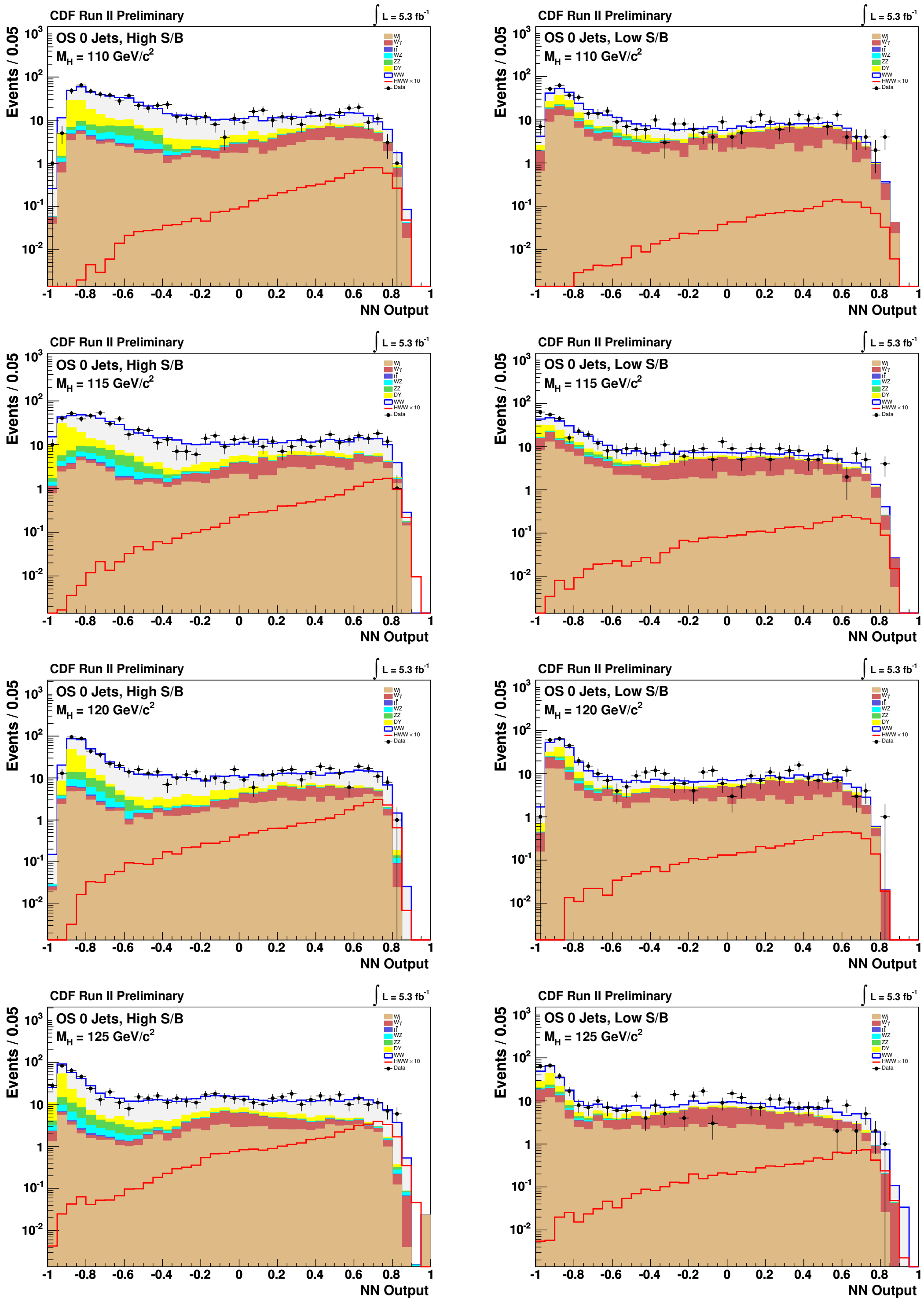

Figure 35: 0J Neural Network output templates for several Higgs masses. On the left for High S/B events and on the right for Low S/B. From the top to bottom: 110 GeV, 115 GeV, 120 GeV, and 125 GeV.

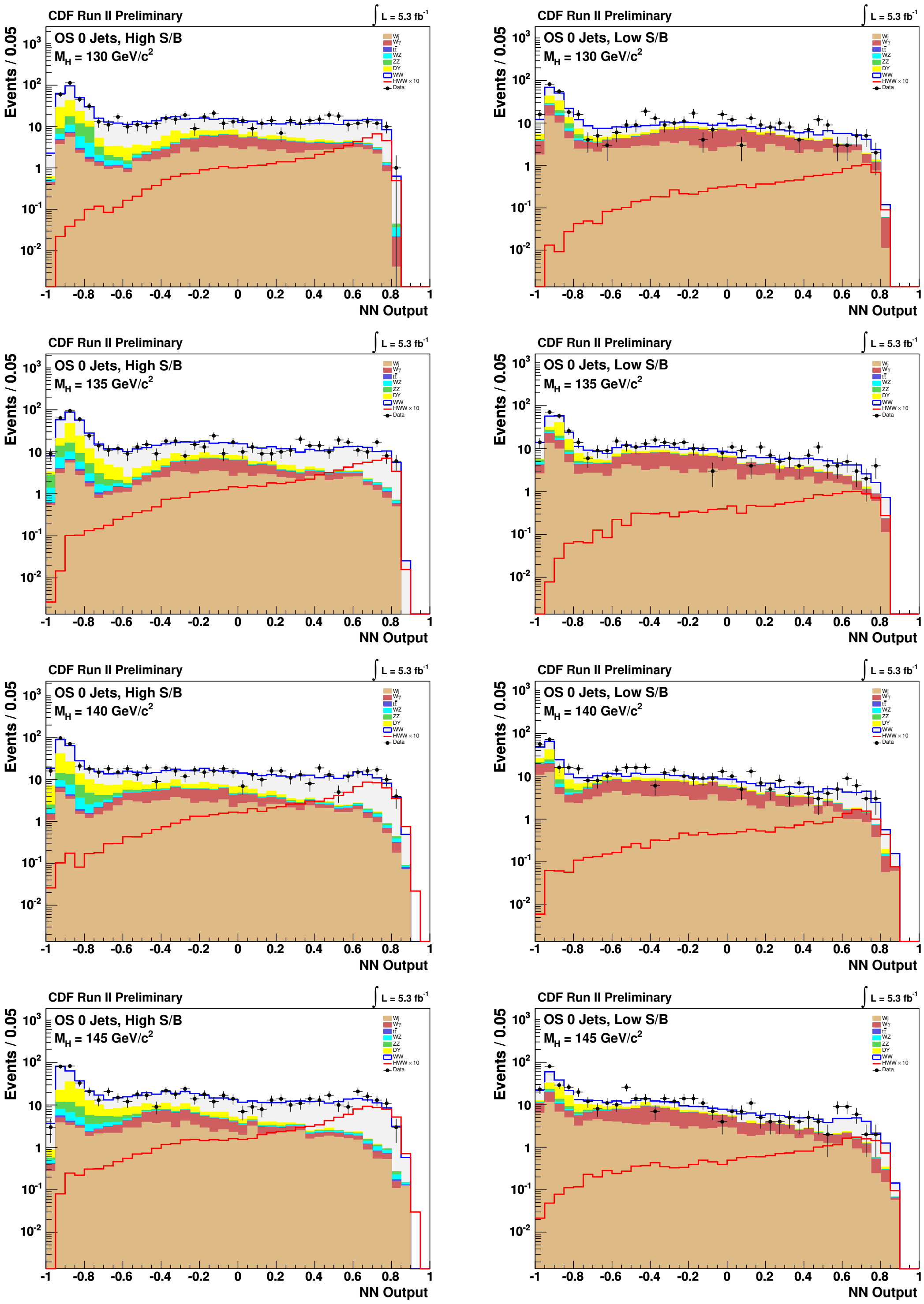

Figure 36: 0J Neural Network output templates for several Higgs masses. On the left for High S/B events and on the right for Low S/B. From the top to bottom: 130 GeV, 135 GeV, 140 GeV, and 145 GeV.

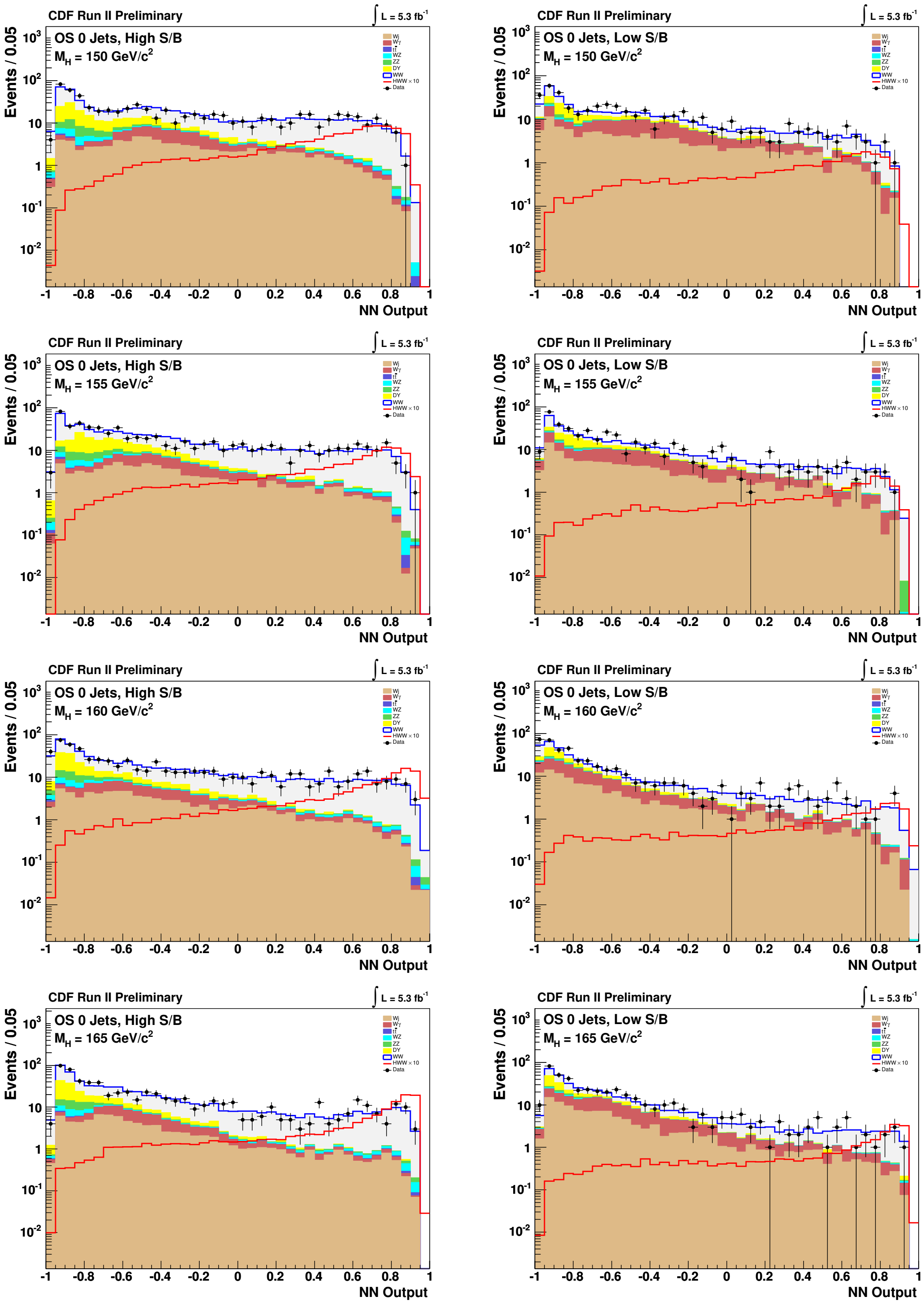

Figure 37: 0J Neural Network output templates for several Higgs masses. On the left for High S/B events and on the right for Low S/B. From the top to bottom: 150 GeV, 155 GeV, 160 GeV, and 165 GeV.

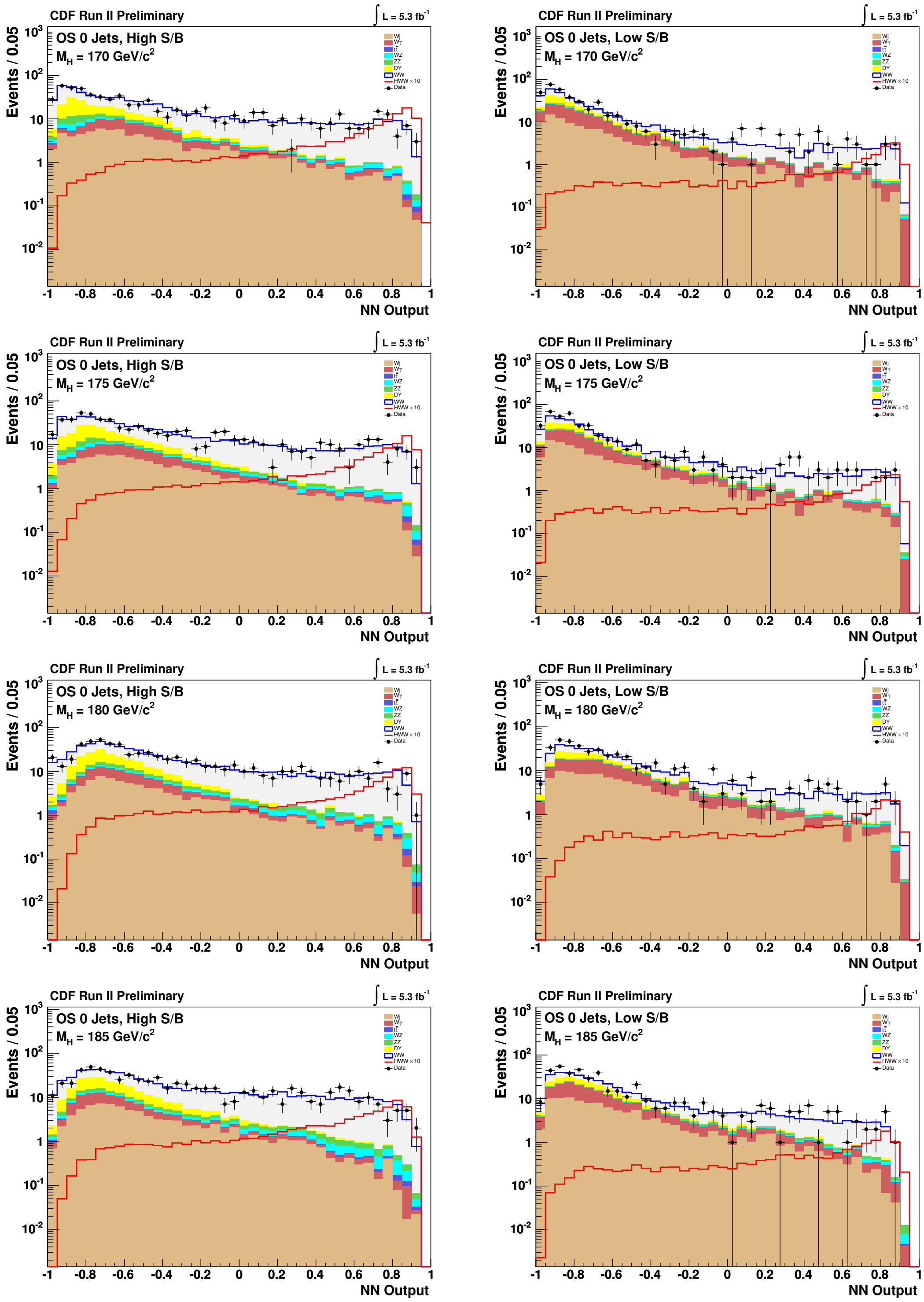

Figure 38: 0J Neural Network output templates for several Higgs masses. On the left for High S/B events and on the right for Low S/B. From the top to bottom: 170 GeV, 175 GeV, 180 GeV, 185 GeV.

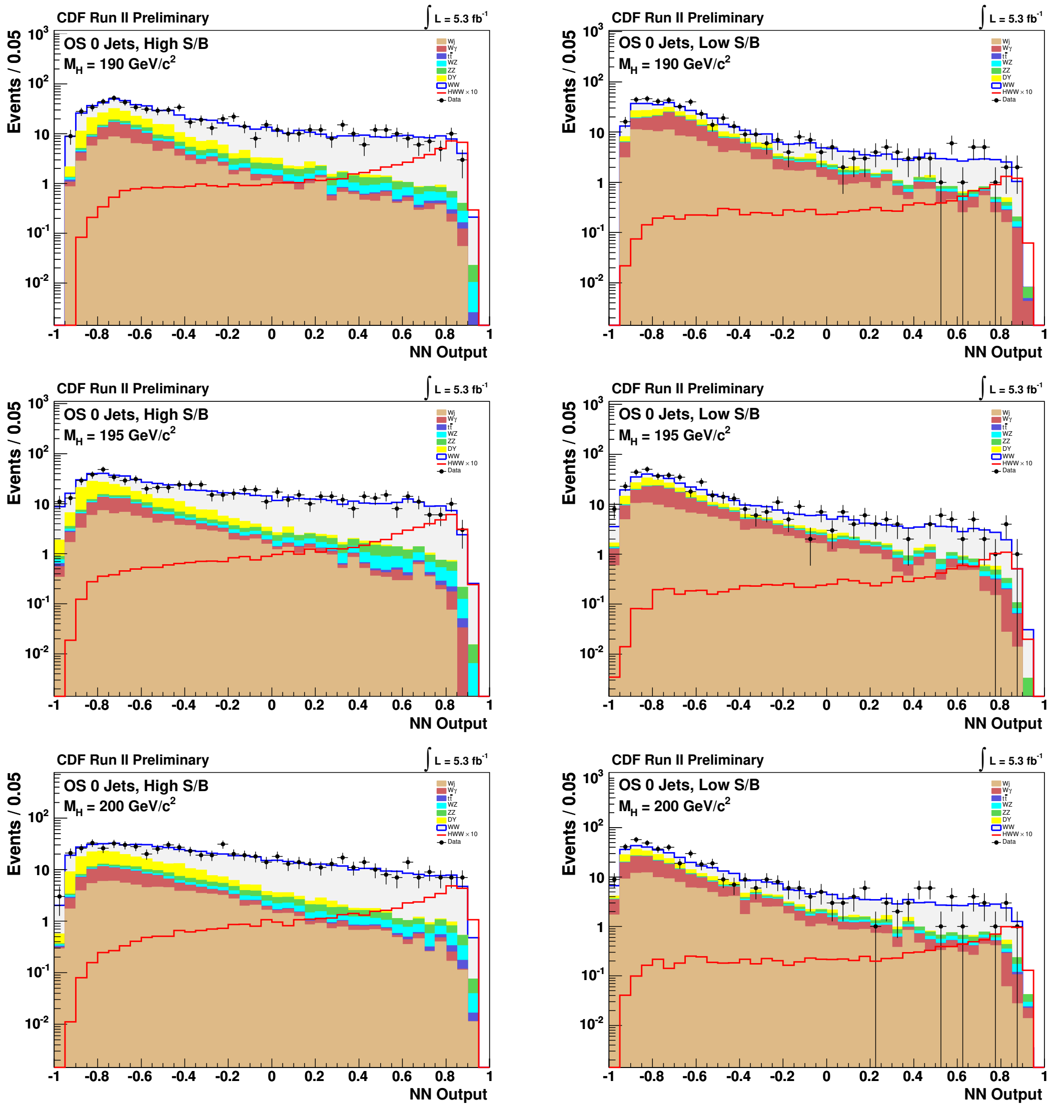

Figure 39: 0J Neural Network output templates for several Higgs masses. On the left for High S/B events and on the right for Low S/B. From the top to bottom: 190 GeV, 195 GeV, and 200 GeV.

Uncertainty Source	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jet
Cross Section							
Total	*6.0%*	*6.0%*	*6.0%*	10.0%	5.0%		
Acceptance							
PDF Model	*1.9%*	*2.7%*	*2.7%*	*2.1%*	*4.1%*		
Higher-order Diagrams	*5.0%*	*10.0%*	*10.0%*	10.0%			
Jet Energy Scale	*2.6%*	*6.1%*	*3.4%*	*26.0%*	*17.5%*	*3.1%*	
Missing E_T Modeling					**20.0%**		
$W\gamma$ Modeling						11.0%	
Conversion Modeling						11.0%	
Jet Fake Rates							
(Low S/B)							22.0%
(High S/B)							24.0%
MC Run Dependence	2.6%						
Lepton ID Efficiencies	*2.0%*	*1.7%*	*2.0%*	*2.0%*	*1.9%*		
Trigger Efficiencies	*2.1%*	*2.1%*	*2.1%*	*2.0%*	*3.4%*		
Luminosity	*5.9%*	*5.9%*	*5.9%*	*5.9%*	*5.9%*		

Table 40: Systematics table for 0J analysis background processes. Cross section systematics are correlated for the diboson processes (shown in italic font) but uncorrelated between all other processes. Jet Energy Scale systematics are anticorrelated between the 0, 1, and 2+ jet bins, as these systematics cause events to fluctuate between jet bins. The DY Met modeling systematic (shown in boldface font) is uncorrelated between channels.

Uncertainty Source	$gg \to H$	WH	ZH	VBF
Cross Section				
Scale	7.0%			
PDF Model	7.7%			
Total		*5.0%*	*5.0%*	10.0%
Acceptance				
Scale (leptons)	2.5%			
Scale (jets)	4.6%			
PDF Model (leptons)	*1.5%*	*1.2%*	*0.9%*	*2.2%*
PDF Model (jets)	0.9%			
EWK Higher-order Diagrams		*10.0%*	*10.0%*	*10.0%*
Jet Energy Scale	*5.0%*	*10.5%*	*5.0%*	*11.5%*
Lepton ID Efficiencies	*1.9%*	*1.9%*	*1.9%*	*1.9%*
Trigger Efficiencies	*3.3%*	*2.1%*	*2.1%*	*3.3%*
Luminosity	*5.9%*	*5.9%*	*5.9%*	*5.9%*

Table 41: Systematics table for 0J analysis signal processes. The Scale (jets) and PDF Model (jets) systematics for $gg \to H$ are anticorrelated between the 0, 1, and 2+ jet bins, as these systematics cause events to fluctuate between jet bins.

OS 0 Jets	110	115	120	125	130	135	140	145	150	155
$-2\sigma/\sigma_{SM}$	20.75	10.60	6.87	4.42	3.45	2.65	2.21	1.88	1.56	1.30
$-1\sigma/\sigma_{SM}$	28.63	14.60	9.45	6.09	4.77	3.68	3.02	2.57	2.16	1.76
Median/$\sigma_{\mathbf{SM}}$	**40.71**	**20.81**	**13.40**	**8.63**	**6.74**	**5.24**	**4.35**	**3.69**	**3.04**	**2.51**
$+1\sigma/\sigma_{SM}$	58.79	29.87	19.21	12.42	9.67	7.49	6.27	5.27	4.39	3.61
$+2\sigma/\sigma_{SM}$	82.74	42.26	26.88	17.57	13.68	10.39	8.60	7.36	6.24	4.99
Observed/$\sigma_{\mathbf{SM}}$	**43.78**	**22.92**	**14.04**	**8.53**	**7.04**	**6.46**	**5.55**	**4.48**	**3.32**	**2.99**
	160	165	170	175	180	185	190	195	200	
$-2\sigma/\sigma_{SM}$	0.90	0.88	0.99	1.22	1.52	2.04	2.48	3.05	3.36	
$-1\sigma/\sigma_{SM}$	1.21	1.19	1.34	1.66	2.07	2.81	3.40	4.27	4.60	
Median/$\sigma_{\mathbf{SM}}$	**1.71**	**1.67**	**1.87**	**2.37**	**2.95**	**4.01**	**4.86**	**6.05**	**6.53**	
$+1\sigma/\sigma_{SM}$	2.47	2.41	2.70	3.38	4.27	5.72	6.98	8.83	9.45	
$+2\sigma/\sigma_{SM}$	3.45	3.37	3.75	4.71	5.99	8.09	9.68	12.24	13.37	
Observed/$\sigma_{\mathbf{SM}}$	**2.05**	**2.01**	**2.47**	**2.83**	**3.23**	**4.86**	**4.41**	**7.04**	**7.82**	

Table 42: $H \to WW$ sensitivity normalized to NNLL σ_{SM} calculation from the combination of the High S/B and Low S/B channels for the 0J region. In this region we consider potential signal contributions from gluon fusion, vector boson fusion, and associated production with either a W or Z boson.

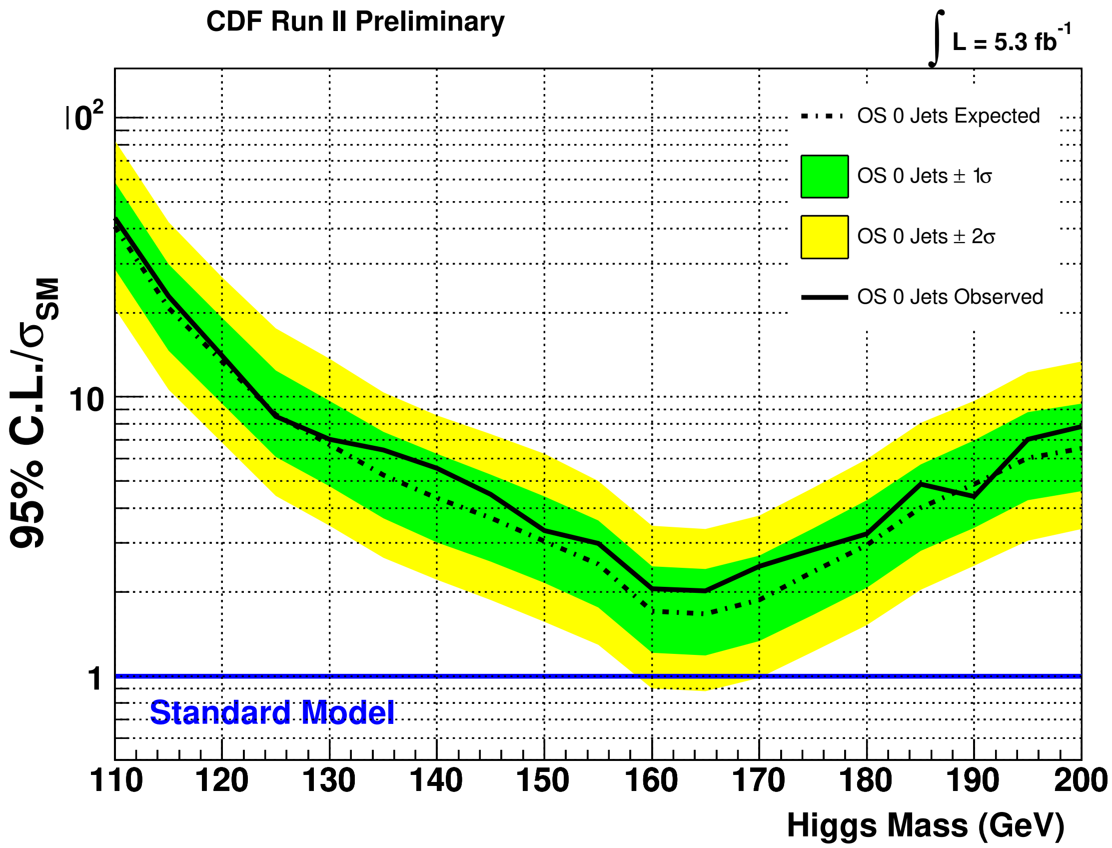

Figure 40: $H \rightarrow WW$ sensitivity normalized to NNLL σ_{SM} calculation from the combination of the High S/B and Low S/B channels in the 0J region. In this region we consider potential signal contributions from gluon fusion, vector boson fusion, and associated production with either a W or Z boson.

8 Opposite-sign Dilepton Signal Region with 1 Jet

The background expectation and observed yield for the opposite-sign dilepton signal region for events containing one reconstructed jet with $E_T > 15$ GeV are shown in Table 43. The expected and observed kinematic distributions for these events are shown in Figures 41-43. We stack the backgrounds on top of each other, and then for comparison we overlay the contribution from a 160 GeV/c^2 Higgs boson with yields scaled up by a factor of 10. In addition, we calculate a confidence level for data versus background expectations on each plot. The confidence levels shown are for Poisson fluctuations and the KS test.

For events with one jet (1J) we train a separate set of NB neural networks, one for each Higgs mass, and do not incorporate any inputs based on the Matrix Element calculations. As discussed previously, the ME calculations are done at LO and therefore are not as reliable for events with an extra jet which presumably originates from an extra quark or gluon radiated within the hard scattering process. Optimization and training of the neural network are done in the same manner as for the 0J case described in Section 7. For the 1J region our neural network uses 8 inputs, 10 hidden nodes and 1 output node. The input variables used for the 1J NN are shown in Table 8. The observed distributions of each 1J neural network input parameter compared with the distributions predicted by our background model are shown in Figure 44. For reference, we also include a hypothetical signal component for a Higgs boson with $m_H = 160$ GeV/c^2.

Templates are created for each neural network and used for calculating 95% CL limits with the MCLimit program [32]. Events are separated into Low S/B and High S/B channels based on the dilepton types and a separate set of templates is created for each of the channels. MCLimit combines the two channels to produce a single set of limits for the 1J region. Output templates for the trained 1J neural network, divided into Low S/B and High S/B, are shown in Figures 45−49.

In the 1J region we consider potential signal contributions from all four Higgs production processes that contribute significantly at high masses: gluon fusion ($gg \to H \to WW \to ll\nu\nu$), Vector Boson Fusion (VBF), and associated production VH ($q\bar{q} \to ZH \to ZWW$ and $q\bar{q} \to WH \to WWW$). The additional processes contribute significantly to the event regions with one or more jets and incorporating their contributions is fairly straight-forward since we do not attempt to use neural network inputs based on Matrix Element calculations in these regions. We evaluate the number of events expected from each production process as a function of the mass of the Higgs. These numbers are shown in Table 45.

In calculating the final limits, we take into account systematic uncertainties on both the signal and background models we use to train our neural networks. Systematic uncertainties for the 1J region are summarized in Tables 46 and 47.

Tables 48 and Figure 50 show the expected (based on 10000 pseudo-experiments) and observed limits obtained from the 1J region, incorporating the systematic uncertainties listed in Tables 46 and 47.

Category	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jets	Total	Data
$e\ e$	27.3	8.6	3.8	10.8	47.8	10.8	28.5	137.6	142.0
$e\ \mu$	61.6	4.1	0.6	24.1	53.6	11.4	35.7	191.0	159.0
$\mu\ \mu$	20.3	4.9	2.7	8.6	33.9	0.0	3.8	74.2	84.0
e trk	18.2	3.5	1.4	7.7	18.3	2.6	5.8	57.4	56.0
μ trk	10.3	1.6	0.8	4.6	8.1	0.4	4.4	30.1	32.0
Total:	137.6	22.6	9.2	55.8	161.7	25.1	78.2	490.3	473.0

Category	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jets	Total	Data
LBE LBE	16.0	4.6	2.3	7.6	25.9	2.2	5.6	64.1	79.0
LBE PHX	9.6	3.5	1.3	2.9	17.0	6.5	16.3	57.2	55.0
PHX PHX	1.7	0.6	0.2	0.3	4.9	2.1	6.6	16.3	8.0
LBE CMUP	18.9	1.0	0.2	8.3	16.3	1.7	6.3	52.7	46.0
LBE CMP	2.6	0.1	0.0	1.2	3.2	0.2	1.8	9.1	2.0
LBE CMX	8.0	0.5	0.1	3.3	6.6	0.9	3.0	22.4	22.0
LBE CMXMsKs	1.1	0.1	0.0	0.6	1.7	0.1	0.7	4.3	2.0
LBE BMU	2.0	0.1	0.0	0.6	0.6	0.1	0.0	3.6	3.0
LBE CMIOCES	12.4	0.7	0.1	5.4	9.9	0.4	1.0	30.0	25.0
LBE CMIOPES	3.5	0.2	0.0	1.0	3.6	0.3	2.4	11.1	8.0
PHX CMUP	5.1	0.4	0.1	1.6	3.6	3.2	7.8	21.8	22.0
PHX CMP	0.7	0.1	0.0	0.2	0.5	0.4	1.3	3.2	5.0
PHX CMX	2.1	0.2	0.0	0.6	1.8	1.8	5.4	11.9	7.0
PHX CMXMsKs	0.3	0.0	0.0	0.1	0.7	0.3	0.6	2.2	2.0
PHX BMU	0.6	0.1	0.0	0.1	0.9	0.2	0.0	1.8	1.0
PHX CMIOCES	3.2	0.3	0.0	0.9	2.6	1.2	1.0	9.3	9.0
PHX CMIOPES	1.0	0.1	0.0	0.2	1.5	0.4	1.7	4.9	5.0
CMUP CMUP	3.6	0.7	0.5	1.7	4.5	0.0	0.3	11.3	16.0
CMUP CMP	0.9	0.2	0.1	0.4	1.3	0.0	0.2	3.2	2.0
CMUP CMX	2.8	0.7	0.4	1.4	4.4	0.0	0.4	10.2	10.0
CMUP CMXMsKs	0.4	0.1	0.1	0.2	0.9	0.0	0.1	1.8	2.0
CMUP BMU	0.8	0.2	0.1	0.3	2.6	0.0	0.6	4.5	2.0
CMUP CMIOCES	4.8	1.1	0.6	2.0	5.0	0.0	0.5	14.0	20.0
CMUP CMIOPES	1.4	0.4	0.2	0.4	3.0	0.0	0.8	6.2	5.0
CMX CMP	0.4	0.1	0.1	0.2	0.3	0.0	0.1	1.2	0.0
CMX CMX	0.7	0.2	0.1	0.3	0.9	0.0	0.1	2.3	2.0
CMX CMXMsKs	0.2	0.1	0.0	0.1	0.5	0.0	0.0	0.9	1.0
CMX BMU	0.2	0.1	0.0	0.1	1.3	0.0	0.0	1.7	3.0
CMX CMIOCES	2.0	0.5	0.3	0.8	3.8	0.0	0.2	7.7	8.0
CMX CMIOPES	0.6	0.2	0.1	0.2	1.8	0.0	0.2	3.0	2.0
CMXMsKs CMP	0.1	0.0	0.0	0.1	0.5	0.0	0.0	0.7	0.0
CMXMsKs CMXMsKs	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.4	0.0
CMXMsKs BMU	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.3	1.0
CMXMsKs CMIOCES	0.2	0.1	0.0	0.1	0.7	0.0	0.0	1.2	4.0
CMXMsKs CMIOPES	0.1	0.0	0.0	0.0	0.5	0.0	0.3	0.9	1.0
CMP CMP	0.1	0.0	0.0	0.0	0.1	0.0	0.0	0.3	1.0
CMP BMU	0.1	0.0	0.0	0.0	0.2	0.0	0.0	0.3	2.0
CMP CMIOCES	0.5	0.1	0.1	0.2	0.6	0.0	0.1	1.7	2.0
CMP CMIOPES	0.2	0.0	0.0	0.0	0.4	0.0	0.1	0.8	0.0
LBE CrkTrk	14.7	2.6	1.1	6.6	13.3	1.1	3.7	43.0	37.0
PHX CrkTrk	3.5	0.9	0.3	1.1	5.0	1.5	2.1	14.3	19.0
CMUP CrkTrk	6.3	1.0	0.5	2.9	4.7	0.2	2.5	18.1	21.0
CMP CrkTrk	0.7	0.1	0.0	0.3	0.8	0.0	0.5	2.5	2.0
CMX CrkTrk	2.9	0.5	0.2	1.2	2.1	0.1	1.3	8.3	8.0
CMXMsKs CrkTrk	0.4	0.1	0.0	0.2	0.5	0.0	0.1	1.3	1.0
Total:	137.6	22.6	9.2	55.8	161.7	25.1	75.9	488.0	473.0

Table 43: Yields in opposite-sign dilepton signal region for events with one jet.

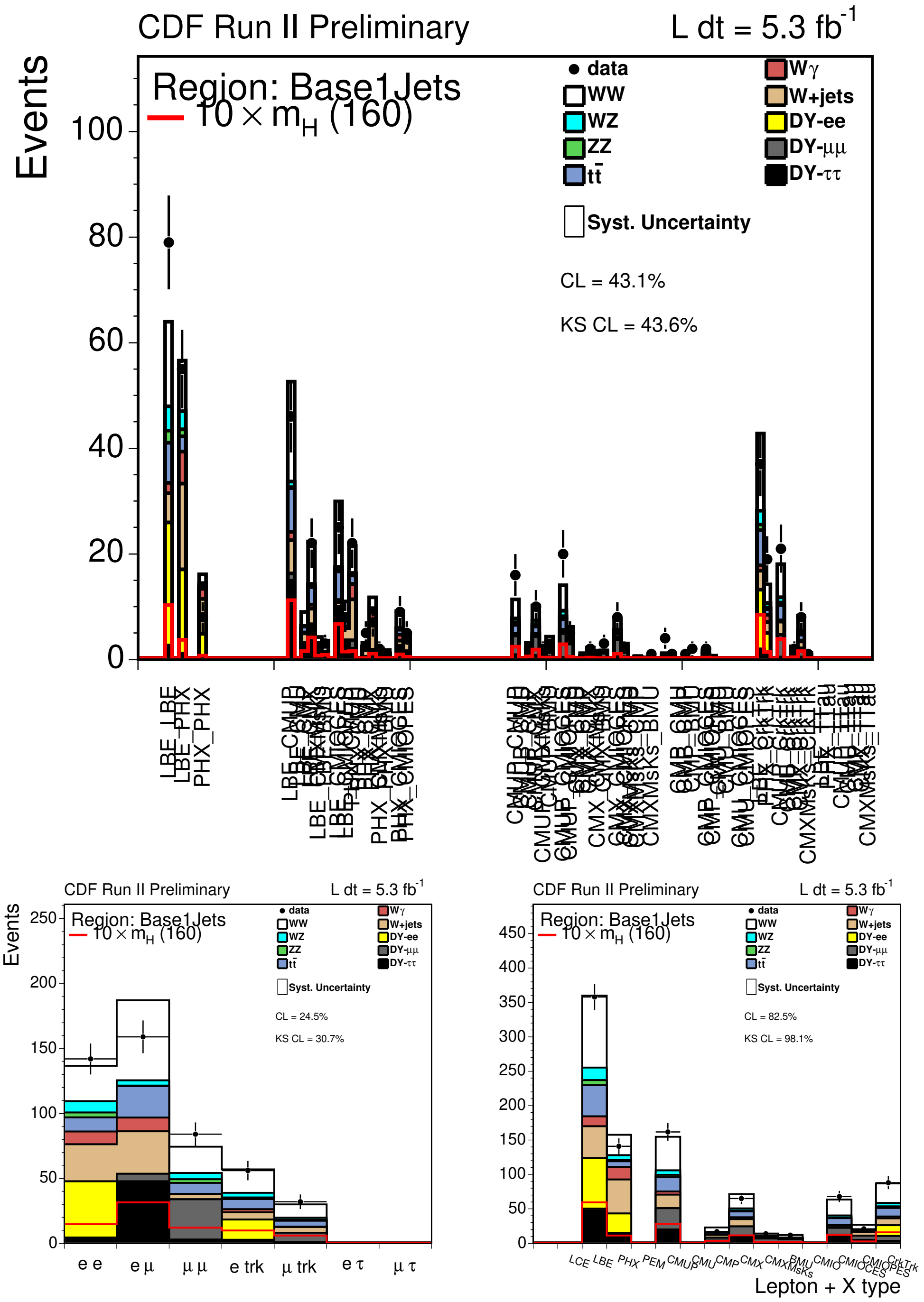

Figure 41: Signal and background in the opposite-sign signal region for events containing one reconstructed jet with $E_T > 15$ GeV.

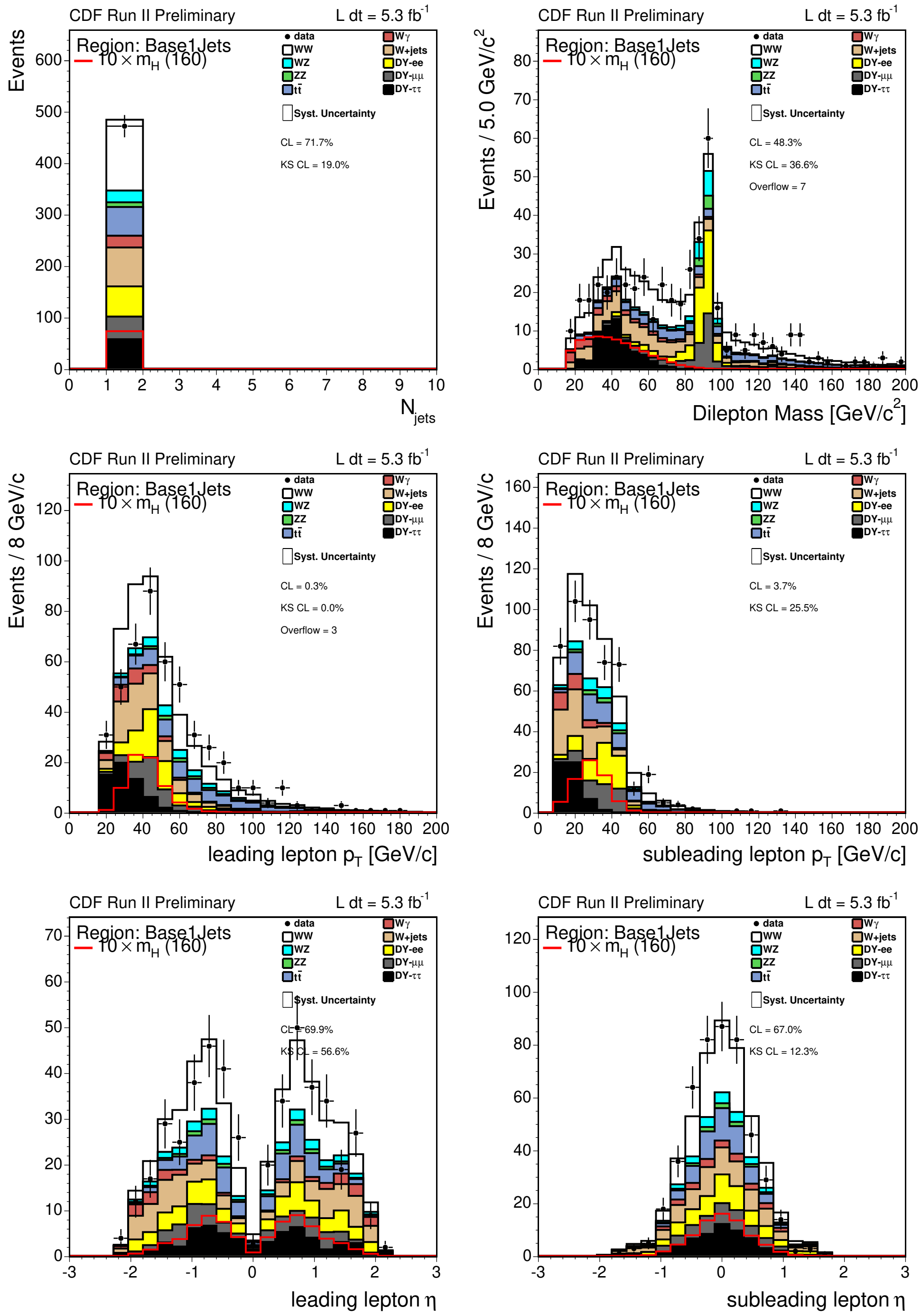

Figure 42: Signal and background in the opposite-sign signal region for events containing one reconstructed jet with $E_T > 15$ GeV.

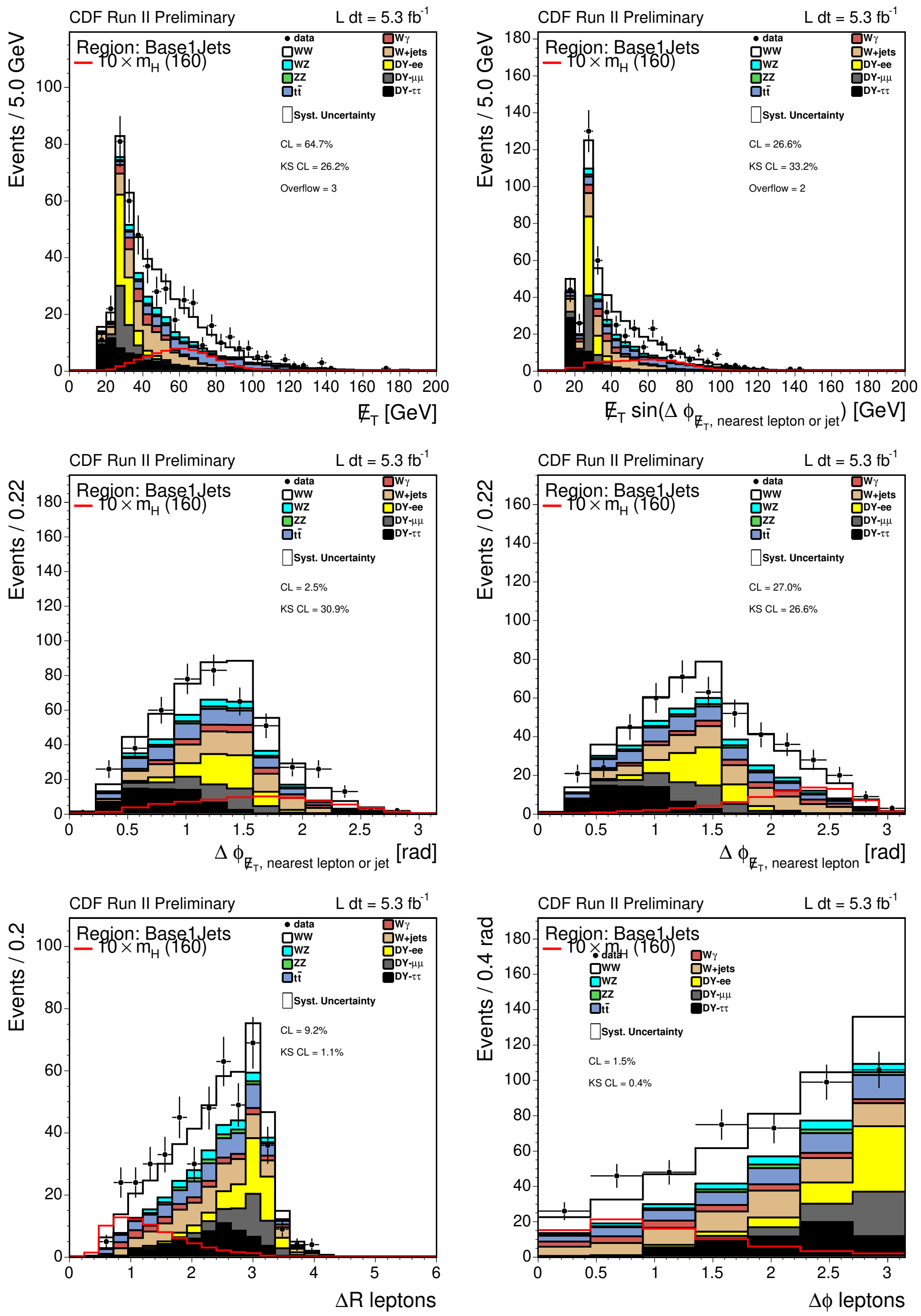

Figure 43: Signal and background in the opposite-sign signal region for events containing one reconstructed jet with $E_T > 15$ GeV.

Variable	Meaning
	1J neural networks
dRLeptons	ΔR between the leptons
Mt_H	Transverse mass of the sum of lepton's 4-momenta and ($\not{E}_T$, MetX, MetY, 0) where MetX/Y are the X/Y components of the $\vec{\not{E}}_T$
dimass	Dilepton invariant mass
MetSpec	$\not{E}_{T spec} = \not{E}_T$ if $\Delta\phi$ ($\vec{\not{E}}_T$, l or j) $> \frac{\pi}{2}$ $\not{E}_{T spec} = \not{E}_T \sin(\Delta\phi$ ($\vec{\not{E}}_T$, l or j)) if $\Delta\phi$ ($\vec{\not{E}}_T$, l or j) $< \frac{\pi}{2}$
lep1_E	Energy of the leading (most energetic) lepton (for muons this is the lepton 3-momentum module)
lep1_Pt	Transverse momentum of the first lepton
lep2_Pt	Transverse momentum of the second lepton
Ht	Scalar sum of Leptons E_T, Jet E_T and $\not{E}_T$

Table 44: Summary of inputs to 1J neural networks.

Mass (GeV/c^2)	gg→H	WH	ZH	VBF
110	0.316 (0.262/0.0535)	0.161 (0.13/0.0305)	0.0536 (0.0427/0.0109)	0.0244 (0.0206/0.00379)
115	0.634 (0.529/0.105)	0.27 (0.218/0.052)	0.0897 (0.0714/0.0183)	0.0479 (0.0405/0.0075)
120	1.06 (0.872/0.186)	0.414 (0.335/0.0789)	0.132 (0.106/0.0259)	0.0838 (0.0707/0.0131)
125	1.68 (1.4/0.289)	0.574 (0.466/0.107)	0.185 (0.149/0.0362)	0.138 (0.116/0.0215)
130	2.48 (2.04/0.438)	0.736 (0.599/0.137)	0.245 (0.2/0.0454)	0.203 (0.17/0.0323)
135	3.3 (2.71/0.589)	0.872 (0.713/0.159)	0.292 (0.237/0.0544)	0.275 (0.23/0.0445)
140	4.09 (3.38/0.712)	0.969 (0.794/0.175)	0.328 (0.27/0.0586)	0.344 (0.289/0.0545)
145	4.87 (4/0.863)	1.05 (0.856/0.189)	0.362 (0.298/0.0647)	0.421 (0.353/0.0676)
150	5.62 (4.64/0.973)	1.07 (0.875/0.195)	0.381 (0.313/0.0685)	0.491 (0.412/0.0785)
155	6.27 (5.19/1.08)	1.07 (0.884/0.189)	0.389 (0.321/0.0684)	0.558 (0.469/0.089)
160	7.09 (5.9/1.2)	1.07 (0.885/0.186)	0.405 (0.336/0.0697)	0.631 (0.533/0.0976)
165	7.1 (5.92/1.18)	1.01 (0.837/0.172)	0.392 (0.328/0.0643)	0.663 (0.561/0.102)
170	6.62 (5.5/1.12)	0.904 (0.753/0.151)	0.352 (0.293/0.0591)	0.638 (0.538/0.1)
175	6.08 (5.04/1.04)	0.818 (0.677/0.141)	0.317 (0.265/0.052)	0.602 (0.506/0.0959)
180	5.45 (4.49/0.962)	0.718 (0.594/0.124)	0.279 (0.233/0.0455)	0.547 (0.459/0.0885)
185	4.45 (3.67/0.779)	0.577 (0.479/0.0977)	0.226 (0.188/0.0383)	0.469 (0.392/0.0767)
190	3.83 (3.15/0.684)	0.482 (0.4/0.0825)	0.187 (0.156/0.0308)	0.41 (0.342/0.0682)
195	3.43 (2.81/0.617)	0.421 (0.349/0.0718)	0.164 (0.137/0.0279)	0.368 (0.306/0.0623)
200	3.1 (2.54/0.561)	0.368 (0.306/0.0621)	0.146 (0.121/0.0248)	0.337 (0.28/0.0575)

Table 45: Expected Higgs Signal yields as a function of the Higgs boson mass in the 1J region for each of the contributing production processes. The numbers inside the parentheses are the expected relative contributions in the High S/B and Low S/B regions.

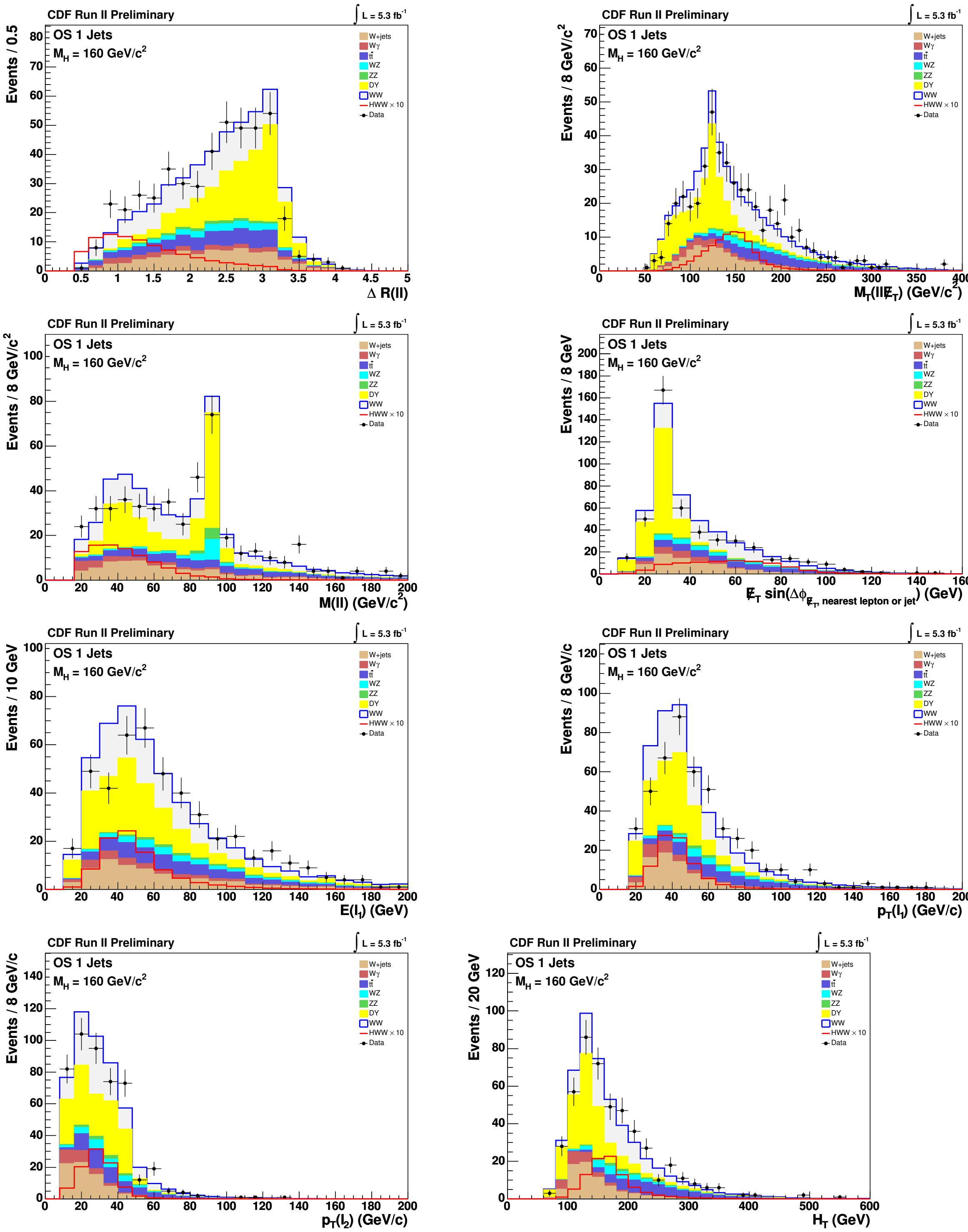

Figure 44: Neural Network input variables for 1J NN (signal for $M_H = 160$ GeV/c^2).

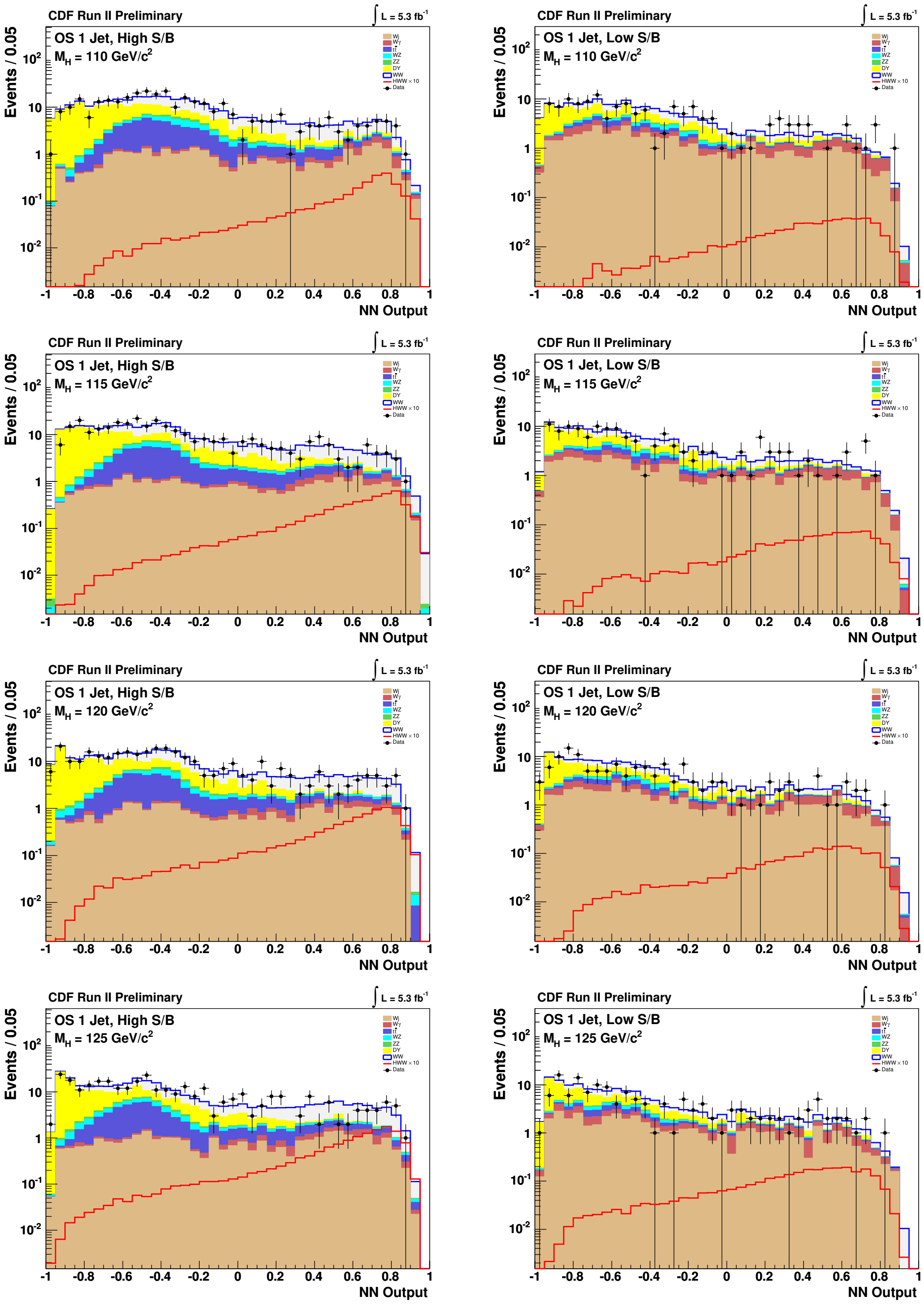

Figure 45: 1J Neural Network output templates for several Higgs masses. On the left for High S/B events and on the right for Low S/B. From the top to bottom: 110 GeV, 115 GeV, 120 GeV, and 125 GeV.

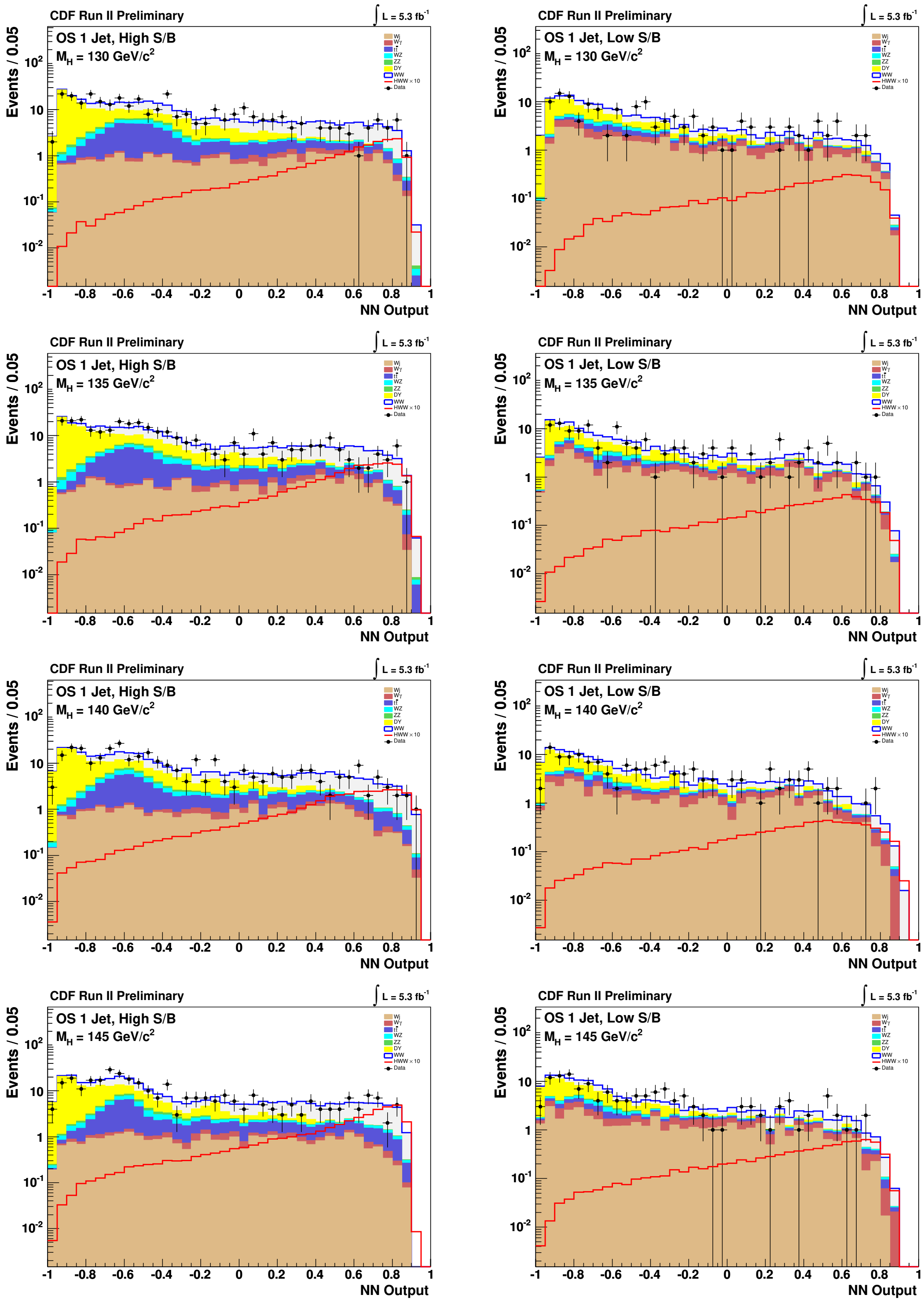

Figure 46: 1J Neural Network output templates for several Higgs masses. On the left for High S/B events and on the right for Low S/B. From the top to bottom: 130 GeV, 135 GeV, 140 GeV, and 145 GeV.

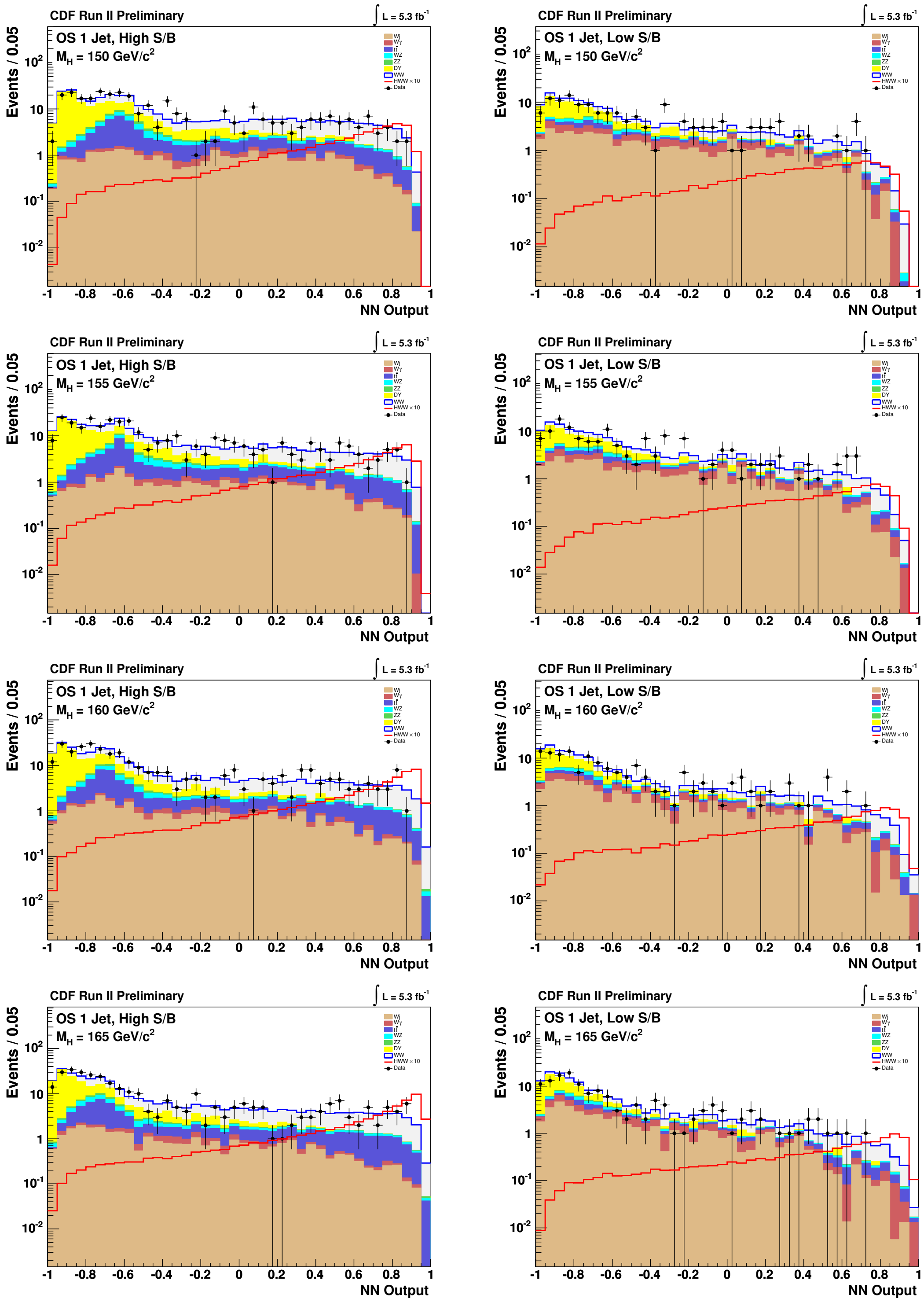

Figure 47: 1J Neural Network output templates for several Higgs masses. On the left for High S/B events and on the right for Low S/B. From the top to bottom: 150 GeV, 155 GeV, 160 GeV, and 165 GeV.

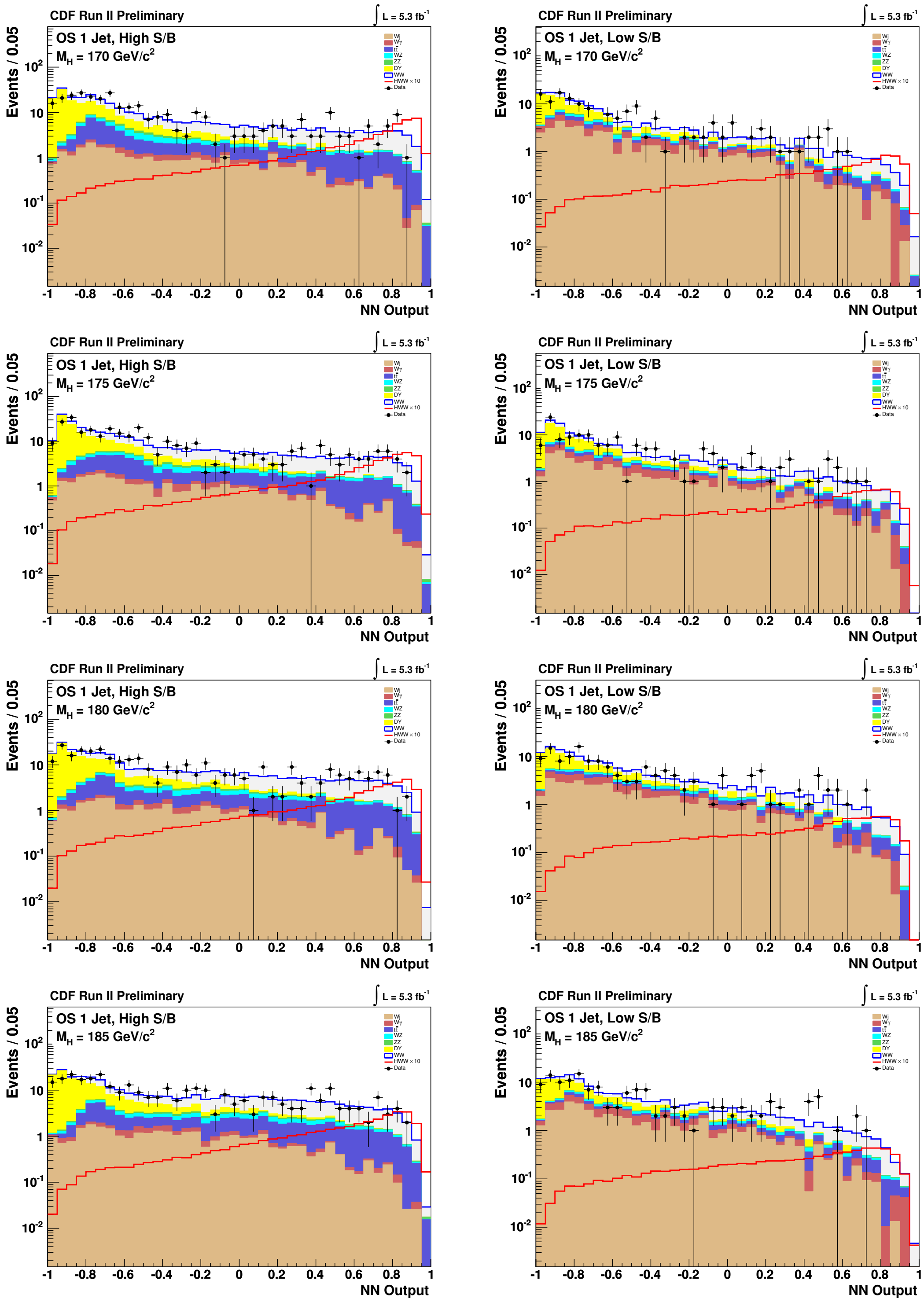

Figure 48: 1J Neural Network output templates for several Higgs masses. On the left for High S/B events and on the right for Low S/B. From the top to bottom: 170 GeV, 175 GeV, 180 GeV, 185 GeV.

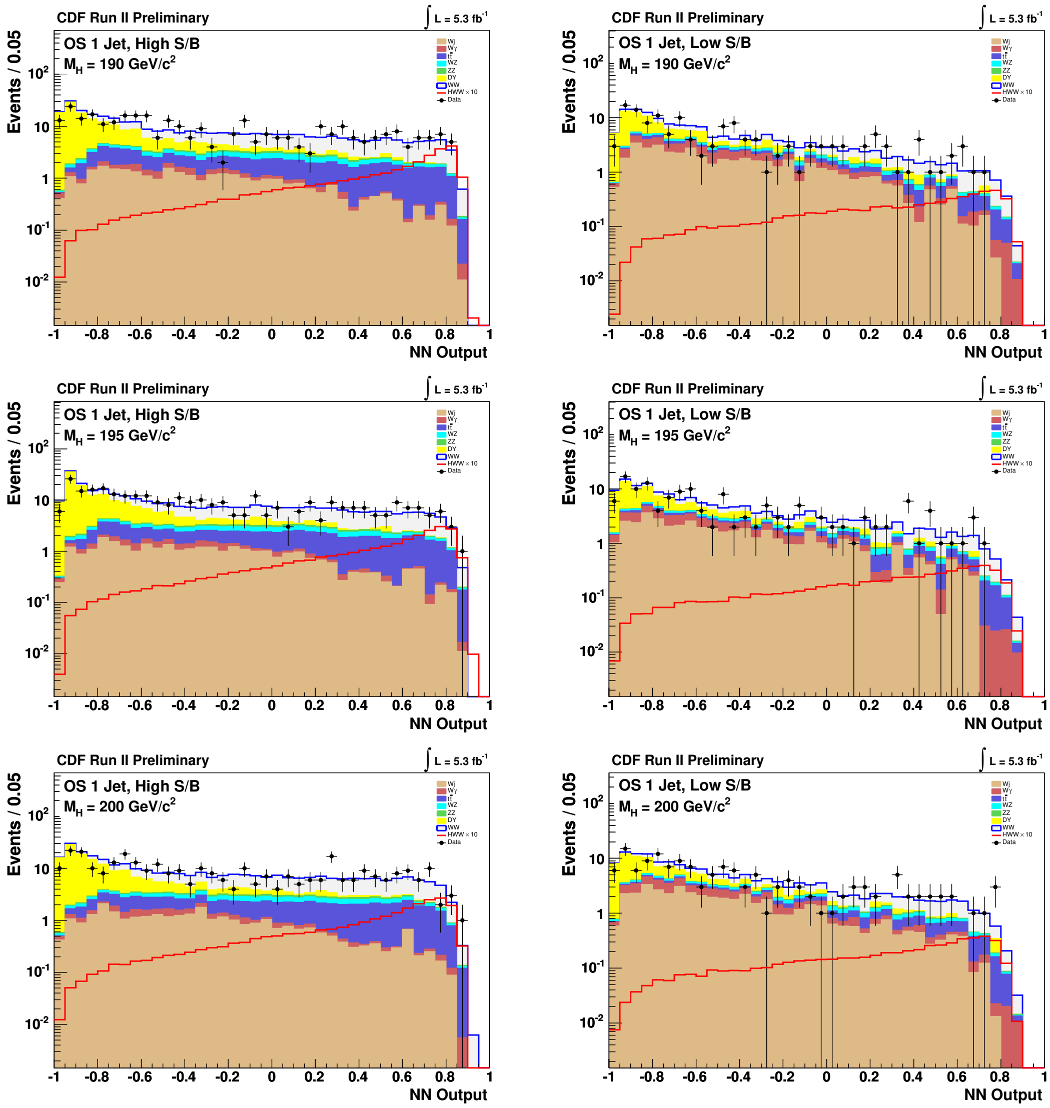

Figure 49: 1J Neural Network output templates for several Higgs masses. On the left for High S/B events and on the right for Low S/B. From the top to bottom: 190 GeV, 195 GeV, and 200 GeV.

Uncertainty Source	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jet
Cross Section							
Total	*6.0%*	*6.0%*	*6.0%*	10.0%	5.0%		
Acceptance							
PDF Model	*1.9%*	*2.7%*	*2.7%*	*2.1%*	*4.1%*		
Higher-order Diagrams	*5.0%*	*10.0%*	*10.0%*	10.0%			
Jet Energy Scale	*-5.5%*	*-1.0%*	*-4.3%*	*-13.0%*	*-6.5%*	*-9.5%*	
Missing E_T Modeling					**20.0%**		
$W\gamma$ Modeling						11.0%	
Conversion Modeling						11.0%	
Jet Fake Rates							
(Low S/B)							22.0%
(High S/B)							26.0%
MC Run Dependence	1.0%						
Lepton ID Efficiencies	*2.0%*	*2.0%*	*2.2%*	*1.8%*	*2.0%*		
Trigger Efficiencies	*2.1%*	*2.1%*	*2.1%*	*2.0%*	*3.4%*		
Luminosity	*5.9%*	*5.9%*	*5.9%*	*5.9%*	*5.9%*		

Table 46: Systematics table for 1J analysis background processes. Cross section systematics are correlated for the diboson processes (shown in italic font) but uncorrelated between all other processes. Jet Energy Scale systematics are anticorrelated between the 0, 1, and 2+ jet bins, as these systematics cause events to fluctuate between jet bins. The DY Met modeling systematic (shown in boldface font) is uncorrelated between channels.

Uncertainty Source	$gg \to H$	WH	ZH	VBF
Cross Section				
Scale	23.5%			
PDF Model	7.7%			
Total		*5.0%*	*5.0%*	10.0%
Acceptance				
Scale (leptons)	2.8%			
Scale (jets)	-5.1%			
PDF Model (leptons)	*1.7%*	*1.2%*	*0.9%*	*2.2%*
PDF Model (jets)	-1.9%			
EWK Higher-order Diagrams		*10.0%*	*10.0%*	*10.0%*
Jet Energy Scale	*-4.0%*	*-8.5%*	*-7.0%*	*-6.5%*
Lepton ID Efficiencies	*1.9%*	*1.9%*	*1.9%*	*1.9%*
Trigger Efficiencies	*3.3%*	*2.1%*	*2.1%*	*3.3%*
Luminosity	*5.9%*	*5.9%*	*5.9%*	*5.9%*

Table 47: Systematics table for 1J analysis signal processes. Cross section systematics are uncorrelated between the signal processes. Higher-order diagrams for the VH and VBF processes, shown in italic font, are correlated among each other but uncorrelated from background processes. The Scale (jets) and PDF Model (jets) systematics for $gg \to H$ are anticorrelated between the 0, 1, and 2+ jet bins, as these systematics cause events to fluctuate between jet bins.

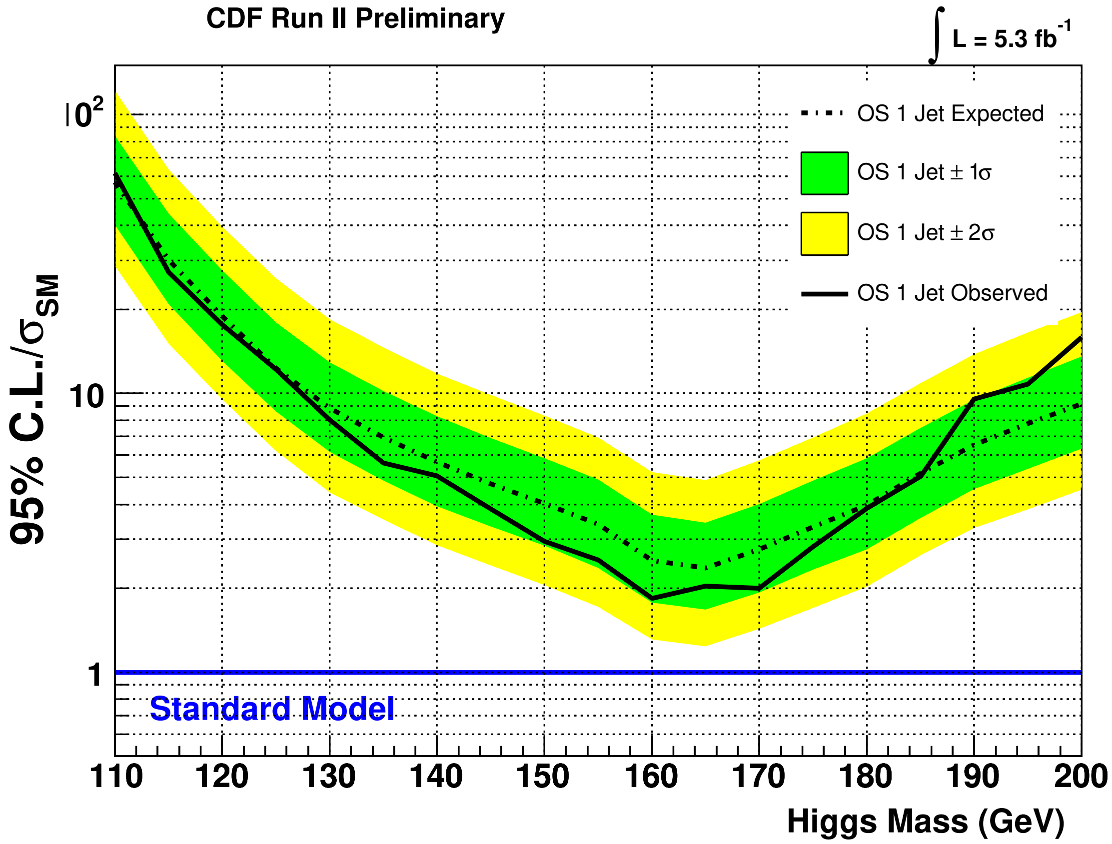

Figure 50: $H \to WW$ sensitivity normalized to NNLL σ_{SM} calculation from the combination of the High S/B and Low S/B channels in the 1J region. In this region we consider potential signal contributions from gluon fusion, vector boson fusion, and associated production with either a W or Z boson.

OS 1 Jet	110	115	120	125	130	135	140	145	150	155
$-2\sigma/\sigma_{SM}$	28.73	15.11	9.53	6.23	4.43	3.53	2.86	2.42	2.05	1.71
$-1\sigma/\sigma_{SM}$	40.20	20.93	13.12	8.61	6.17	4.86	3.94	3.34	2.85	2.36
Median/$\sigma_{\mathbf{SM}}$	**57.31**	**30.11**	**18.97**	**12.30**	**8.87**	**6.96**	**5.67**	**4.75**	**4.03**	**3.39**
$+1\sigma/\sigma_{SM}$	83.81	44.21	27.64	17.96	12.92	10.19	8.28	6.93	5.85	4.92
$+2\sigma/\sigma_{SM}$	121.92	63.59	39.70	25.92	18.48	14.61	11.78	9.88	8.34	6.94
Observed/$\sigma_{\mathbf{SM}}$	**61.61**	**27.24**	**17.70**	**12.19**	**8.05**	**5.62**	**5.06**	**3.85**	**2.95**	**2.53**
	160	165	170	175	180	185	190	195	200	
$-2\sigma/\sigma_{SM}$	1.31	1.24	1.43	1.70	2.03	2.62	3.28	3.84	4.51	
$-1\sigma/\sigma_{SM}$	1.77	1.67	1.93	2.33	2.76	3.58	4.53	5.36	6.34	
Median/$\sigma_{\mathbf{SM}}$	**2.51**	**2.36**	**2.76**	**3.32**	**3.98**	**5.18**	**6.53**	**7.80**	**9.19**	
$+1\sigma/\sigma_{SM}$	3.67	3.44	4.01	4.86	5.86	7.53	9.51	11.40	13.55	
$+2\sigma/\sigma_{SM}$	5.22	4.88	5.74	6.96	8.46	10.89	13.84	16.58	19.60	
Observed/$\sigma_{\mathbf{SM}}$	**1.84**	**2.04**	**2.00**	**2.82**	**3.87**	**5.04**	**9.56**	**10.80**	**15.90**	

Table 48: $H \rightarrow WW$ sensitivity normalized to NNLL σ_{SM} calculation from the combination of the High S/B and Low S/B channels. In this region we consider potential signal contributions from gluon fusion, vector boson fusion, and associated production with either a W or Z boson.

9 Opposite-sign Dilepton Signal Region with 2+ Jets

The event selection for events with two or more jets in the final state (2J) is nearly identical to that for events with zero or one jets. We require two or more jets with $E_{\mathrm{T}} > 15$ GeV and $|\eta| < 2.5$ using Level 5 jet corrections.

The dominant background in this region is from $t\bar{t}$ production where both top quarks decay via $t \to Wb \to \ell\nu b$. Our requirement of two jets in the final state is satisfied by the presence of the two b-quarks in the final state. In order to reduce the $t\bar{t}$ background, we additionally reject events that contain one or more b-tagged jets. For the b-tag, we use the tight SecVtx information in Stntuple [33]. Since we are apply b-tagging to our $t\bar{t}$ background model, we need to use the standard Data/MC efficiency scale factor correction for b-tagging which we take to be 0.94 ± 0.03. There are no real b-jets in the other sources of background, but we expect to occasionally mistag one of these jets as a b-jet. The mistag matrix for tight SecVtx requirements for the first 2.2 fb^{-1} of data is given in [34] and shows that the rate of mistags is very low. Because our other background sources are small contributions and the mistag rate is also very small, we expect the effect of mistags to be negligible and therefore do not attempt to apply any corrections to our models for these effects. An additional scale factor of 1.023 is applied due to not requiring events to be in the silicon good run list, as described in Section 6.3. Even after rejecting events that contain b-tags, $t\bar{t}$ production accounts for more than 50% of the total background in the 2J region.

The background expectation and observed yield for the opposite-sign dilepton signal region for events containing two or more reconstructed jets with $E_T > 15$ GeV are shown in Table 49. The numbers in this table correspond to event totals prior to the removal of events containing b-tags. The expected and observed kinematic distributions for these events are shown in Figures 51-53. We stack the backgrounds on top of each other, and then for comparison we overlay the contribution from a 160 GeV/c^2 Higgs boson with yields scaled up by a factor of 10. In addition, we calculate a confidence level for data versus background expectations on each plot. The confidence levels shown are for Poisson fluctuations and the KS test.

Category	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jets	Total	Data
e e	5.7	2.3	1.1	69.8	15.4	2.0	11.2	107.4	75.0
e μ	13.5	1.3	0.4	162.8	23.4	2.1	14.4	217.9	202.0
μ μ	4.2	1.3	0.8	53.9	11.9	0.0	2.2	74.2	91.0
e trk	3.8	0.9	0.4	47.6	7.6	0.4	2.8	63.5	50.0
μ trk	2.4	0.4	0.2	28.4	3.3	0.0	2.4	37.3	30.0
Total:	29.7	6.2	2.9	362.5	61.5	4.5	32.9	500.2	448.0

Category	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jets	Total	Data
LBE LBE	3.5	1.2	0.7	49.0	8.0	0.5	2.3	65.1	49.0
LBE PHX	1.9	0.9	0.4	18.9	6.2	1.3	6.2	35.7	21.0
PHX PHX	0.3	0.1	0.1	1.9	1.1	0.3	2.7	6.5	5.0
LBE CMUP	4.1	0.4	0.1	56.2	7.3	0.2	3.0	71.3	69.0
LBE CMP	0.5	0.0	0.0	8.0	1.1	0.1	0.7	10.4	11.0
LBE CMX	1.8	0.2	0.1	22.4	3.9	0.1	1.1	29.6	19.0
LBE CMXMsKs	0.2	0.0	0.0	3.9	0.0	0.0	0.3	4.4	9.0
LBE BMU	0.3	0.0	0.0	4.5	0.7	0.0	0.0	5.6	11.0
LBE CMIOCES	2.9	0.2	0.1	35.4	3.3	0.2	0.4	42.5	41.0
LBE CMIOPES	0.7	0.1	0.0	7.1	2.6	0.0	0.9	11.4	11.0
PHX CMUP	1.1	0.1	0.0	10.9	1.9	0.6	3.0	17.5	11.0
PHX CMP	0.1	0.0	0.0	1.4	0.1	0.1	0.4	2.1	2.0
PHX CMX	0.5	0.1	0.0	4.4	1.1	0.3	1.6	8.0	3.0
PHX CMXMsKs	0.1	0.0	0.0	0.8	0.1	0.1	0.4	1.4	3.0
PHX BMU	0.1	0.0	0.0	0.8	0.1	0.1	0.0	1.1	3.0
PHX CMIOCES	0.8	0.1	0.0	5.8	0.6	0.3	0.5	8.0	7.0
PHX CMIOPES	0.2	0.0	0.0	1.2	0.9	0.1	1.2	3.6	2.0
CMUP CMUP	0.6	0.2	0.1	10.3	1.9	0.0	0.3	13.5	15.0
CMUP CMP	0.2	0.1	0.0	2.8	0.3	0.0	0.1	3.5	6.0
CMUP CMX	0.7	0.2	0.1	8.5	2.1	0.0	0.2	11.8	13.0
CMUP CMXMsKs	0.1	0.0	0.0	1.4	0.7	0.0	0.0	2.3	2.0
CMUP BMU	0.1	0.1	0.0	1.7	1.2	0.0	0.0	3.2	4.0
CMUP CMIOCES	1.1	0.3	0.2	13.1	1.9	0.0	0.3	16.8	21.0
CMUP CMIOPES	0.3	0.1	0.0	2.6	0.3	0.0	0.3	3.8	5.0
CMX CMP	0.1	0.0	0.0	1.2	0.3	0.0	0.0	1.6	1.0
CMX CMX	0.2	0.0	0.0	1.6	0.0	0.0	0.0	1.9	3.0
CMX CMXMsKs	0.0	0.0	0.0	0.6	0.2	0.0	0.0	0.9	1.0
CMX BMU	0.1	0.0	0.0	0.7	0.3	0.0	0.0	1.1	2.0
CMX CMIOCES	0.3	0.1	0.1	5.1	1.0	0.0	0.1	6.7	9.0
CMX CMIOPES	0.1	0.0	0.0	1.0	0.9	0.0	0.2	2.3	2.0
CMXMsKs CMP	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.4	1.0
CMXMsKs CMXMsKs	0.0	0.0	0.0	0.1	0.1	0.0	0.0	0.2	0.0
CMXMsKs BMU	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.2	1.0
CMXMsKs CMIOCES	0.1	0.0	0.0	0.7	0.3	0.0	0.0	1.1	2.0
CMXMsKs CMIOPES	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.2	1.0
CMP CMP	0.0	0.0	0.0	0.3	0.1	0.0	0.0	0.4	0.0
CMP BMU	0.0	0.0	0.0	0.2	0.1	0.0	0.0	0.3	0.0
CMP CMIOCES	0.1	0.0	0.0	1.3	0.0	0.0	0.0	1.5	1.0
CMP CMIOPES	0.0	0.0	0.0	0.3	0.1	0.0	0.0	0.4	1.0
LBE CrkTrk	3.0	0.7	0.3	40.8	6.3	0.1	2.0	53.2	37.0
PHX CrkTrk	0.8	0.2	0.1	6.9	1.3	0.2	0.6	10.1	13.0
CMUP CrkTrk	1.5	0.3	0.1	18.2	2.3	0.0	1.6	24.1	20.0
CMP CrkTrk	0.2	0.0	0.0	1.8	0.1	0.0	0.1	2.2	4.0
CMX CrkTrk	0.7	0.1	0.1	7.3	0.8	0.0	0.7	9.6	5.0
CMXMsKs CrkTrk	0.1	0.0	0.0	1.1	0.1	0.0	0.0	1.3	1.0
Total:	29.7	6.2	2.9	362.5	61.5	4.5	31.5	498.8	448.0

Table 49: Yields in opposite-sign dilepton signal region for events with two or more jets, before removing those that contain a b-tagged jet.

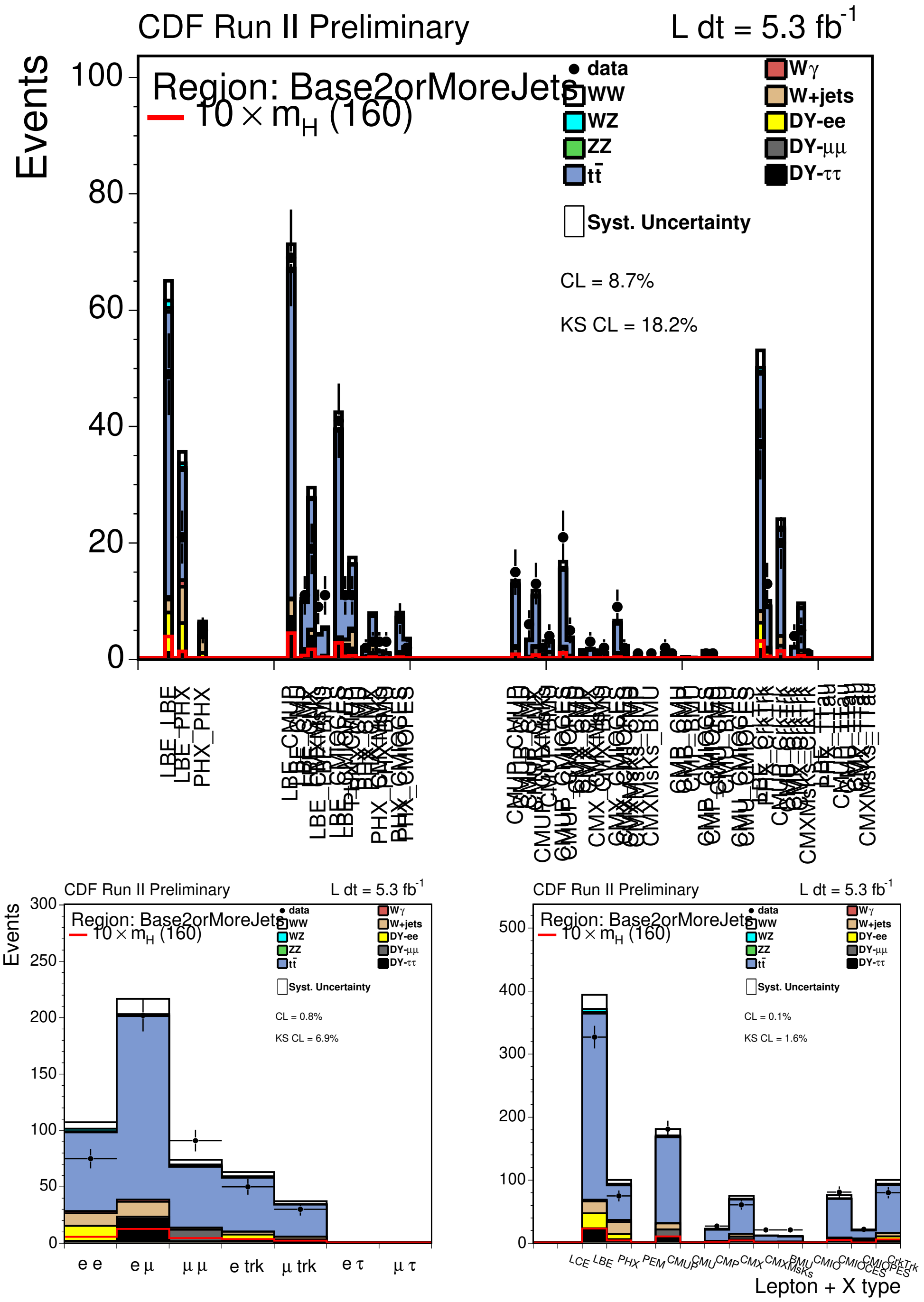

Figure 51: Signal and background in the opposite-sign signal region for events containing two or more reconstructed jets with $E_T > 15$ GeV (prior to removal of events containing a b-tagged jet).

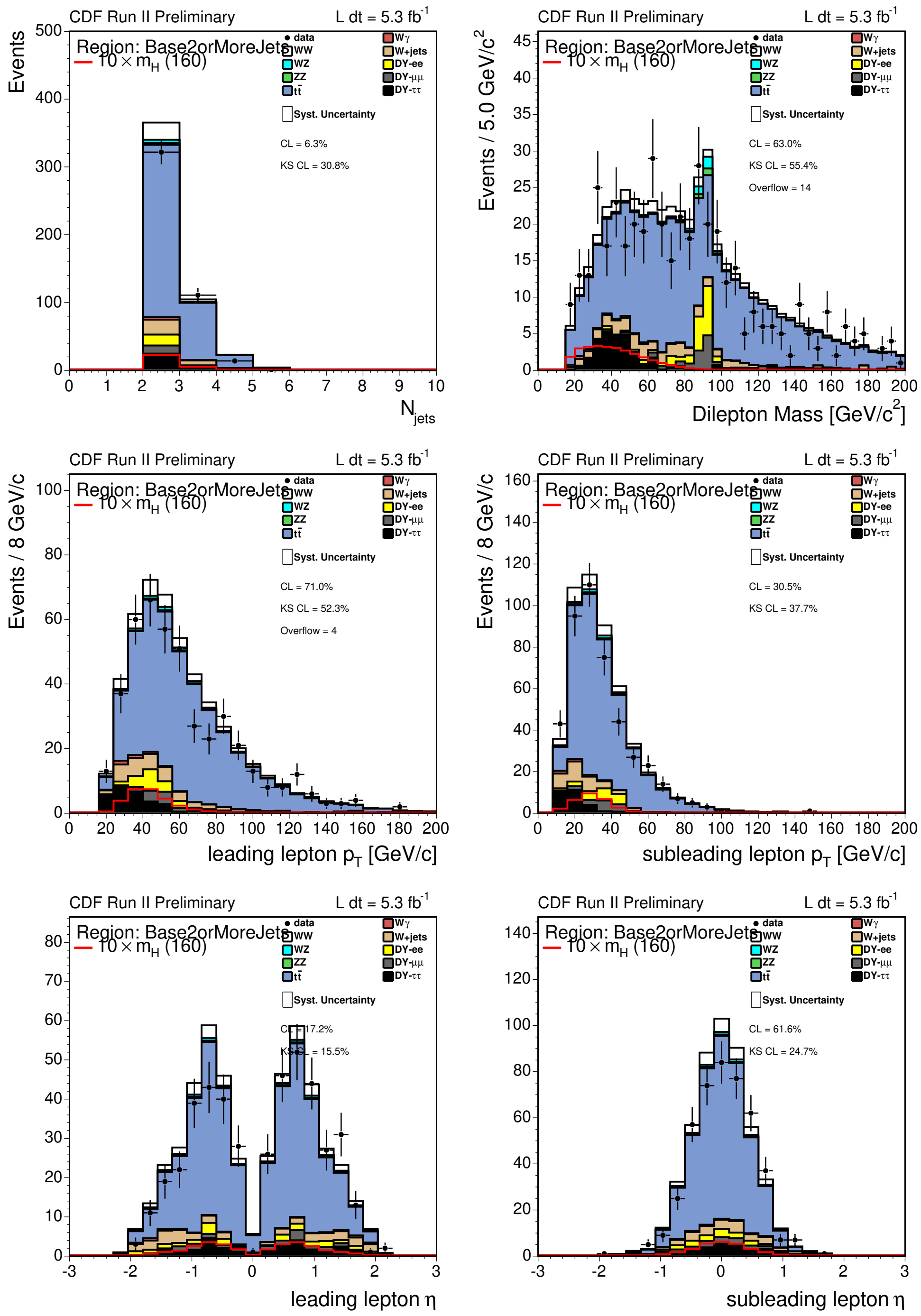

Figure 52: Signal and background in the opposite-sign signal region for events containing two or more reconstructed jets with $E_T > 15$ GeV (prior to removal of events containing a b-tagged jet).

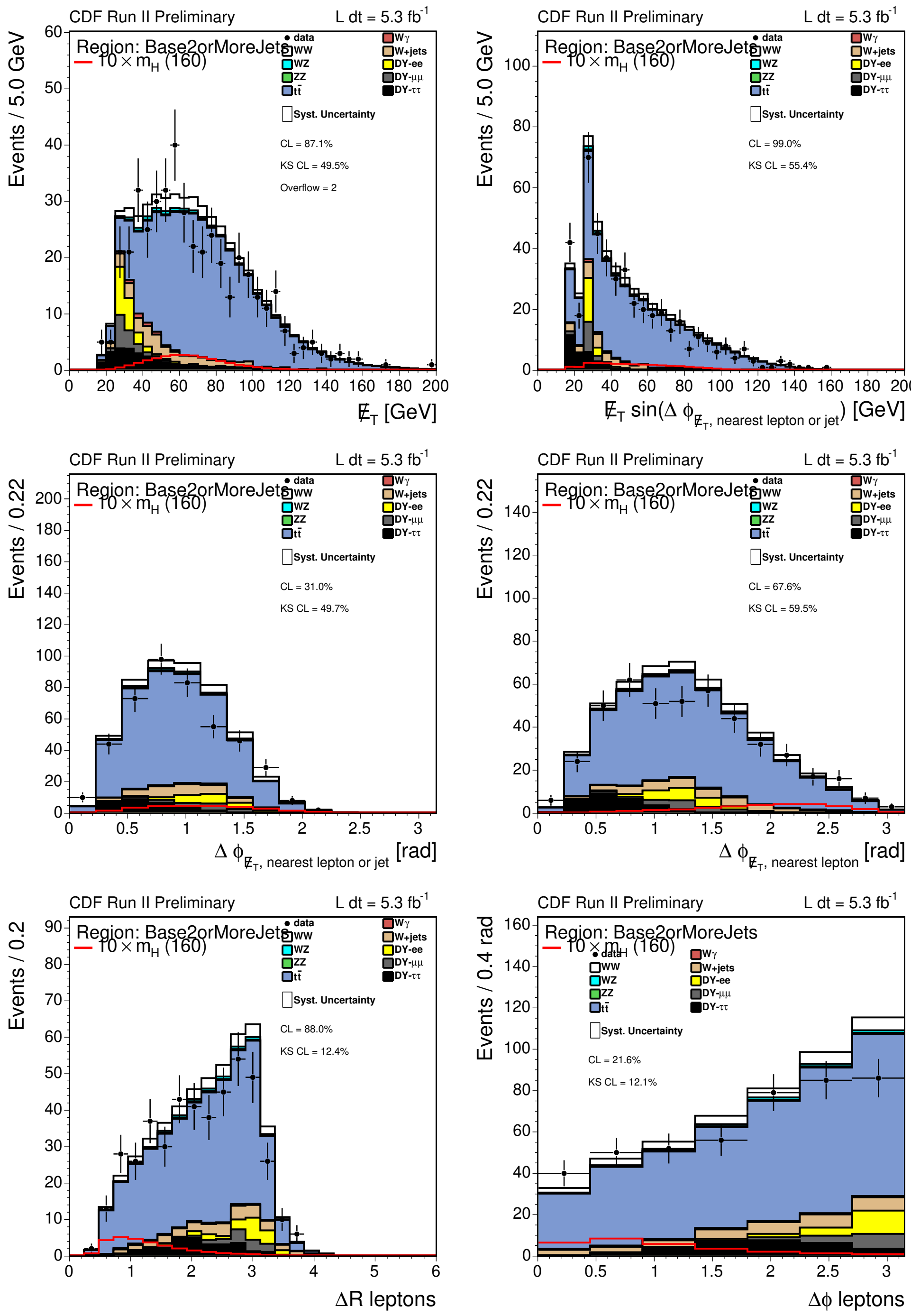

Figure 53: Signal and background in the opposite-sign signal region for events containing two or more reconstructed jets with $E_T > 15$ GeV (prior to removal of events containing a b-tagged jet).

Category	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jets	Total	Data
$e\ e$	0.2	0.1	0.1	37.7	0.4	0.0	1.0	39.5	27.0
$e\ \mu$	0.3	0.0	0.0	86.8	0.7	0.0	1.4	89.2	82.0
$\mu\ \mu$	0.1	0.0	0.0	28.6	0.2	0.0	1.1	30.0	38.0
e trk	0.1	0.0	0.0	25.6	0.3	0.0	0.1	26.1	19.0
μ trk	0.1	0.0	0.0	15.2	0.0	0.0	0.5	15.8	18.0
Total:	0.8	0.1	0.1	193.9	1.6	0.0	4.1	200.6	184.0

Table 50: Dilepton yields in opposite-sign $t\bar{t}$ control region for events with two or more jetss in which at least one jet is tagged as originating from a b-quark.

9.1 $t\bar{t}$ Control Region

Events with at least one b-tagged jet are removed from the signal region and examined as a $t\bar{t}$ control region. The expected and observed dilepton yields are shown in Table 50. Without considering systematic uncertainties on the other background contributions, the 184 observed events give a rough measurement of the $t\bar{t}$ cross-section as 7.2 pb with a systematic uncertainty of about 10%. This is consistent with the theoretical prediction of 7.884 pb and with previous Tevatron measurements. Plots of various kinematic variables shown in Figure 54 show good agreement between the MC prediction and the distributions in data.

9.2 OS 2+ Jets Analysis

As in the case of 1J events, we do not attempt to calculate Matrix Elements for 2J events. Instead, we rely solely on a neural network (NN) to discriminate signal from background using kinematic variables as inputs to the NN. We do not use jet angular variables, as they are not reliably reproduced by Pythia. Also, we find that using jet angular variables in the NN does not significantly improve the sensitivity. Optimization and training of the neural network are done in the same manner as for the 0J case described in Section 7. For the 2J region our neural network uses 8 inputs, 10 hidden nodes and 1 output node. The input variables used for the 2J NN are shown in Table 51. The observed distributions of each 2J neural network input parameter compared with the distributions predicted by our background model are shown in Figure 55. Here, events containing a b-tagged jet have been removed in both the observed and predicted distributions. For reference, we also include a hypothetical signal component for a Higgs boson with $m_H = 160$ GeV/c^2.

Templates are created for each neural network and used for calculating 95% CL limits with the MCLimit program [32]. Due to the low statistics in the 2J region, we do not divide events into High S/B and Low S/B categories as is done for the 0J and 1J regions. Output templates for the trained 2J neural network are shown in Figures 56 through 58.

In the 2J region we also consider potential signal contributions from all four Higgs production processes that contribute significantly at high masses (gluon fusion ($gg \rightarrow$

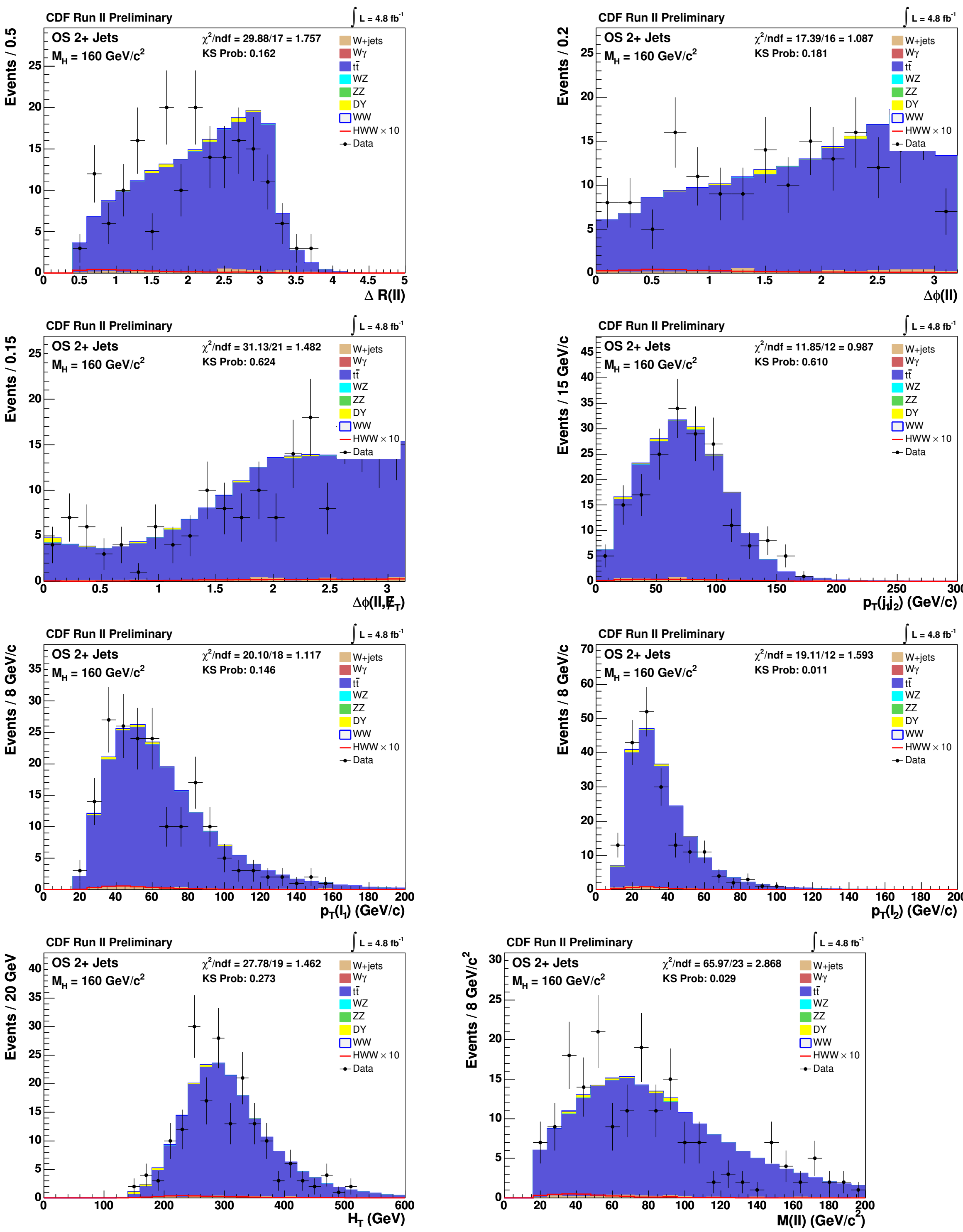

Figure 54: Control region for $t\bar{t}$, events with two or more jets in which at least one jet is tagged as originating from a b-quark (signal for $M_H = 160$ GeV/c^2).

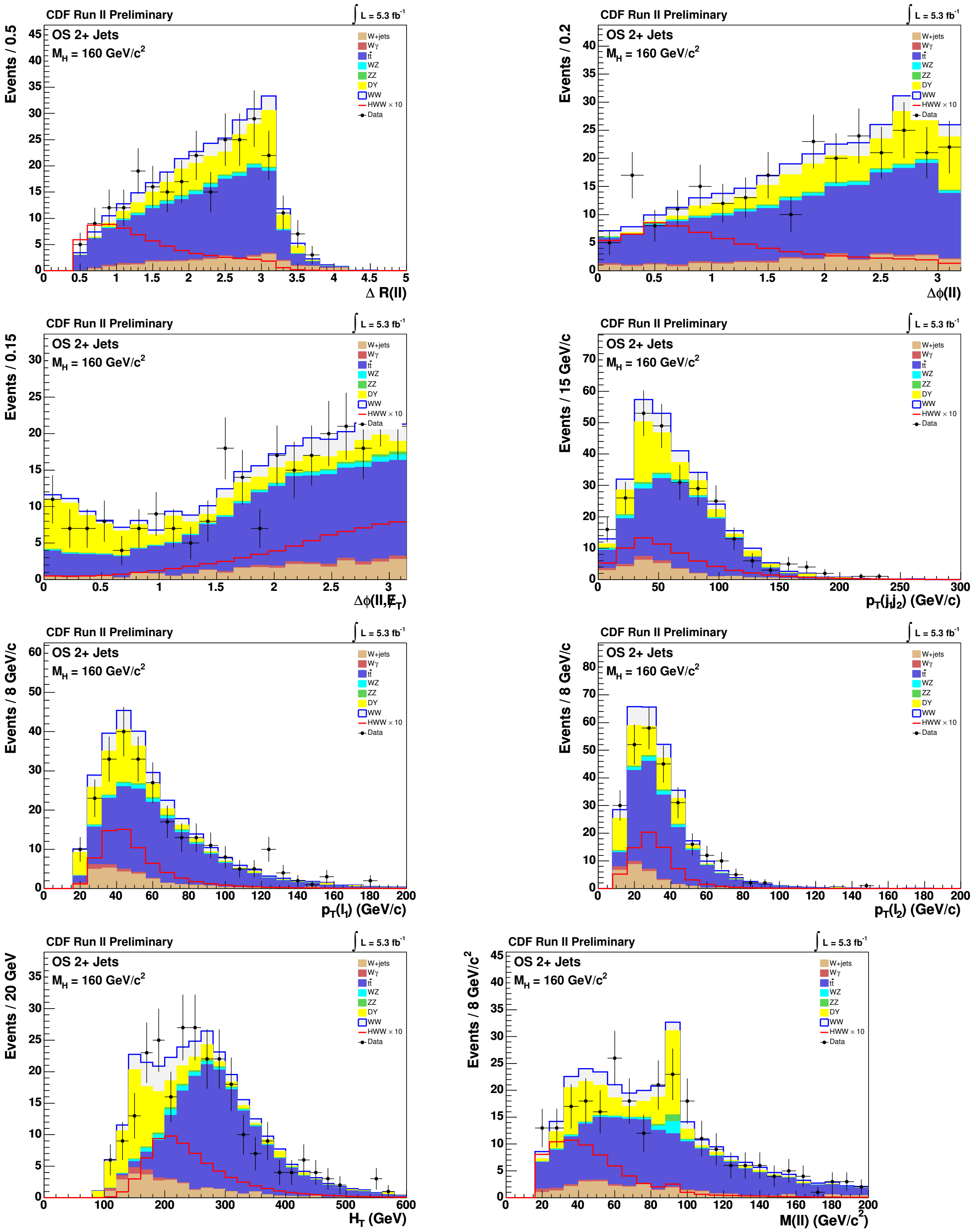

Figure 55: Neural Network input variables for 2J NN (signal for $M_H = 160$ GeV/c^2).

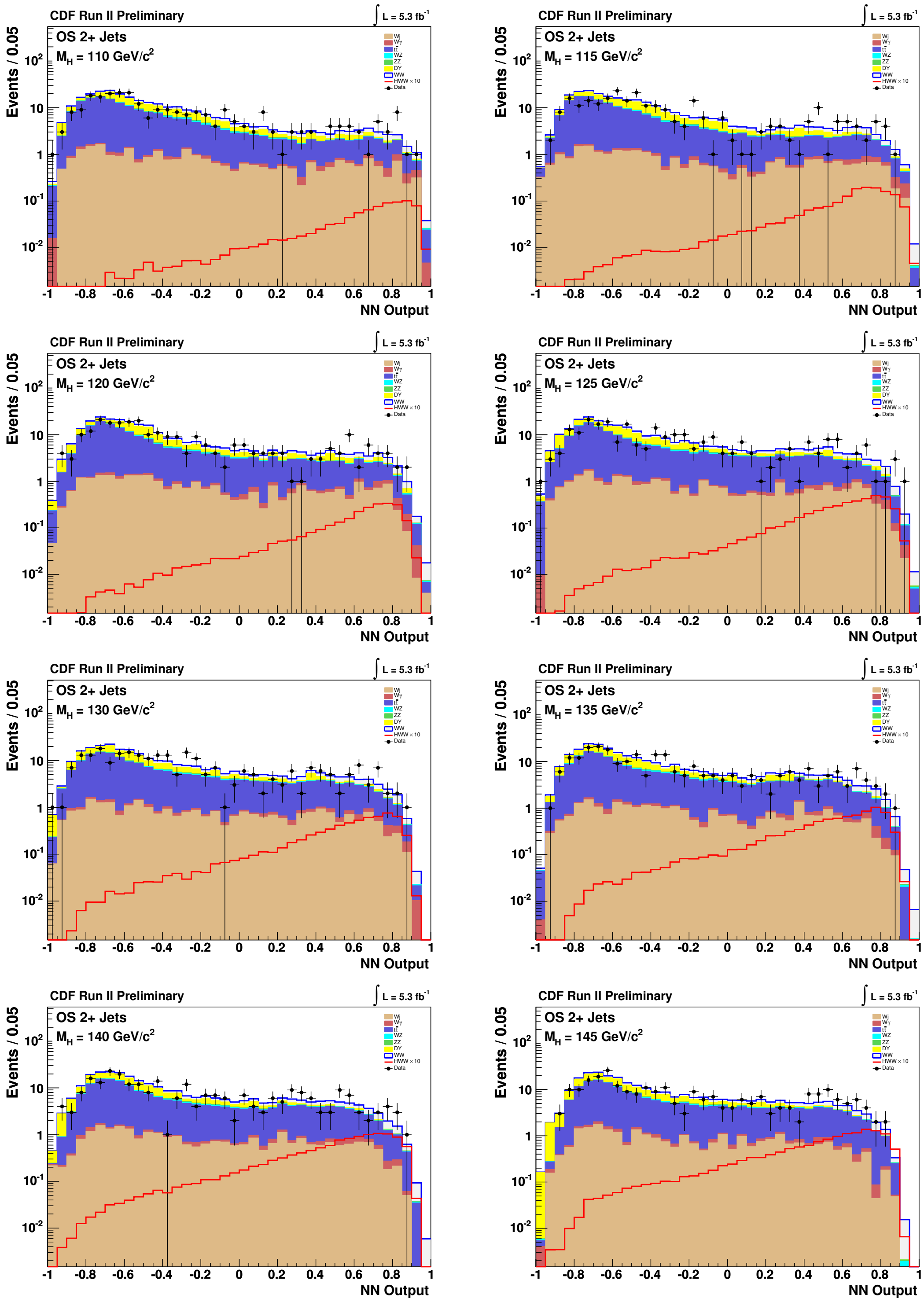

Figure 56: 2J Neural Network output templates for several Higgs masses. From the top left to bottom right: 110 GeV, 115 GeV, 120 GeV, 125 GeV, 130 GeV, 135 GeV, 140 GeV, and 145 GeV.

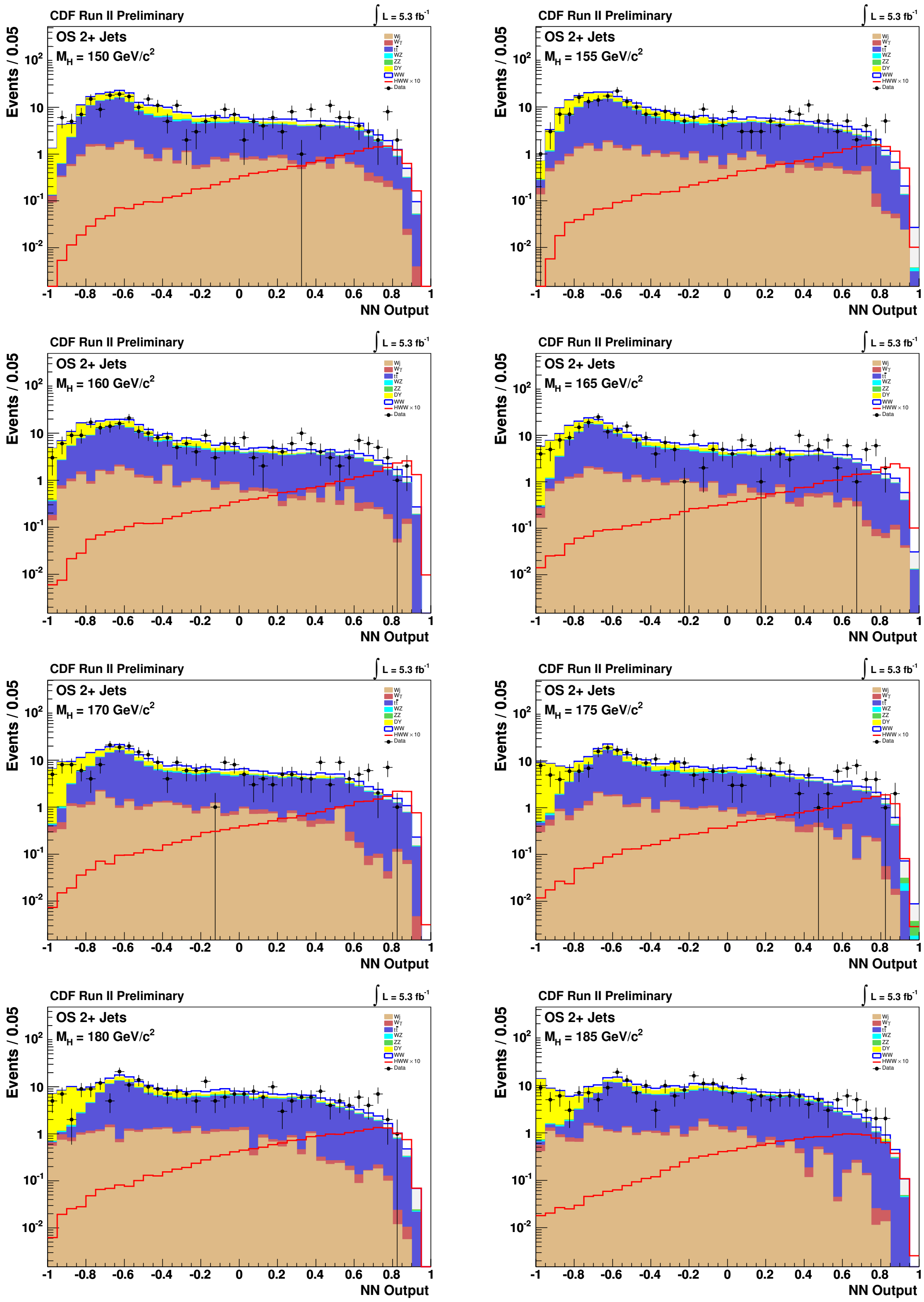

Figure 57: 2J Neural Network output templates for several Higgs masses. From the top left to bottom right : 150 GeV, 155 GeV, 160 GeV, 165 GeV, 170 GeV, 175 GeV, 180 GeV, and 185 GeV.

Variable	Meaning
	2J neural networks
dRLeptons	ΔR between the leptons
dPhiLeptons	$\Delta\phi$ between the leptons
dPhiLepSumMet	$\Delta\phi$ between vector sum of first and second lepton momenta and $\not\!\!E_T$
sumJetPt	Vector sum of first and second jet p_{T}
Lep1Et	Transverse energy E_{T} of the first lepton
Lep2Et	Transverse energy E_{T} of the second lepton
Ht	Transverse mass H_{T} (scalar sum of lepton E_{T}'s, jet E_{T}'s, and $\not\!\!E_T$)
Mll	Dilepton invariant mass M_{ll}

Table 51: Summary of inputs to 2J neural networks.

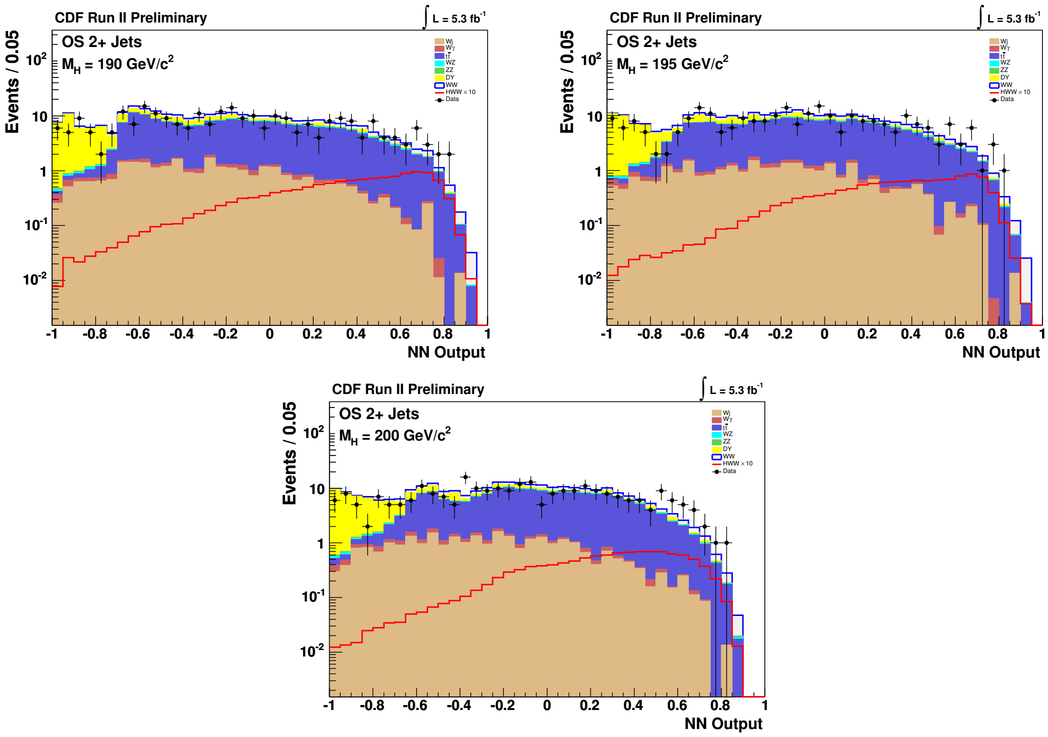

Figure 58: 2J Neural Network output templates for several Higgs masses. From the top left to bottom right : 190 GeV, 195 GeV, and 200 GeV.

Mass (GeV/c^2)	gg→H	WH	ZH	VBF
110	0.0887	0.171	0.0825	0.0463
115	0.176	0.312	0.15	0.0905
120	0.271	0.512	0.25	0.158
125	0.449	0.771	0.372	0.257
130	0.669	1.05	0.514	0.378
135	0.918	1.35	0.663	0.511
140	1.18	1.61	0.786	0.643
145	1.45	1.83	0.902	0.782
150	1.68	2	0.998	0.911
155	1.92	2.14	1.08	1.03
160	2.25	2.3	1.15	1.17
165	2.32	2.23	1.14	1.23
170	2.19	2.05	1.07	1.18
175	2.05	1.89	0.987	1.12
180	1.88	1.71	0.889	1.02
185	1.54	1.4	0.736	0.873
190	1.37	1.18	0.624	0.765
195	1.24	1.05	0.557	0.687
200	1.13	0.941	0.5	0.627

Table 52: Expected Higgs Signal yields as a function of the Higgs boson mass in the 2J region for each of the contributing production processes.

$H \rightarrow WW \rightarrow ll\nu\nu$),Vector Boson Fusion (VBF), and associated production VH ($q\bar{q} \rightarrow ZH \rightarrow ZWW$ and $q\bar{q} \rightarrow WH \rightarrow WWW$). The additional processes contribute significantly to the event regions with one or more jets and incorporating their contributions is fairly straight-forward since we do not attempt to use neural network inputs based on Matrix Element calculations in these regions. We evaluate the number of events expected from each production process as a function of the mass of the Higgs. These numbers are shown in Table 52.

In calculating the final limits, we take into account systematic uncertainties on both the signal and background models we use to train our neural networks. Systematic uncertainties for the 2J region are summarized in Tables 53 and 54.

Tables 55 and Figure 59 show the expected (based on 10000 pseudo-experiments) and observed limits obtained from the 2J region, incorporating the systematic uncertainties listed in Tables 53 and 54.

Uncertainty Source	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jet
Cross Section							
Total	*6.0%*	*6.0%*	*6.0%*	10.0%	5.0%		
Acceptance							
PDF Model	*1.9%*	*2.7%*	*2.7%*	*2.1%*	*4.1%*		
Higher-order Diagrams	*5.0%*	*10.0%*	*10.0%*	10.0%			
Jet Energy Scale	*-14.8%*	*-12.9%*	*-12.1%*	*-1.7%*	*-29.2%*	*-22.0%*	
Missing E_T Modeling					**20.0%**		
$W\gamma$ Modeling						11.0%	
Conversion Modeling						11.0%	
b-tag Veto				3.8%			
Jet Fake Rates							25.0%
MC Run Dependence	2.1%						
Lepton ID Efficiencies	*1.9%*	*2.9%*	*1.9%*	*1.9%*	*1.9%*		
Trigger Efficiencies	*2.1%*	*2.1%*	*2.1%*	*2.0%*	*3.4%*		
Luminosity	*5.9%*	*5.9%*	*5.9%*	*5.9%*	*5.9%*		

Table 53: Systematics table for 2J region background processes. Cross section systematics are correlated for the diboson processes (shown in italic font) but uncorrelated between all other processes. Jet Energy Scale systematics are anticorrelated between the 0, 1, and 2+ jet bins, as these systematics cause events to fluctuate between jet bins. The DY Met modeling systematic (shown in boldface font) is uncorrelated between channels.

Uncertainty Source	$gg \to H$	WH	ZH	VBF
Cross Section				
Scale	67.5%			
PDF Model	7.7%			
Total		*5.0%*	*5.0%*	10.0%
Acceptance				
Scale (leptons)	3.1%			
Scale (jets)	-8.7%			
PDF Model (leptons)	*2.0%*	*1.2%*	*0.9%*	*2.2%*
PDF Model (jets)	-2.8%			
EWK Higher-order Diagrams		*10.0%*	*10.0%*	*10.0%*
Jet Energy Scale	*-17.0%*	*-4.0%*	*-2.3%*	*-4.0%*
Lepton ID Efficiencies	*1.9%*	*1.9%*	*1.9%*	*1.9%*
Trigger Efficiencies	*3.3%*	*2.1%*	*2.1%*	*3.3%*
Luminosity	*5.9%*	*5.9%*	*5.9%*	*5.9%*

Table 54: Systematics table for 2J region signal processes. Cross section systematics are uncorrelated between the signal processes. Higher-order diagrams for the VH and VBF processes, shown in italic font, are correlated among each other but uncorrelated from background processes. The Scale (jets) and PDF Model (jets) systematics for $gg \to H$ are anticorrelated between the 0, 1, and 2+ jet bins, as these systematics cause events to fluctuate between jet bins.

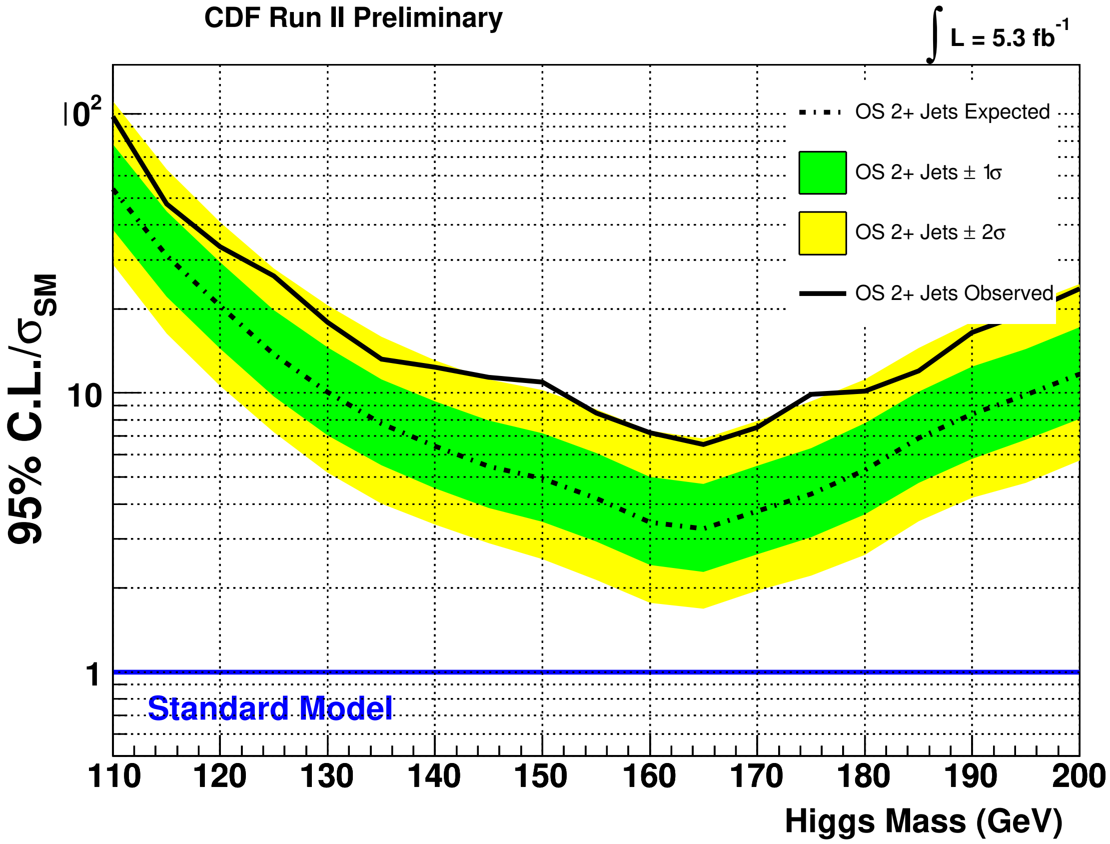

Figure 59: $H \rightarrow WW$ sensitivity normalized to NNLL σ_{SM} calculation in the 2J region. In this region we consider potential signal contributions from gluon fusion, vector boson fusion, and associated production with either a W or Z boson.

OS 2+ Jets	110	115	120	125	130	135	140	145	150	155
$-2\sigma/\sigma_{SM}$	28.95	16.30	10.63	7.25	5.18	4.01	3.38	2.89	2.54	2.13
$-1\sigma/\sigma_{SM}$	38.54	22.13	14.42	9.71	7.04	5.51	4.56	3.88	3.46	2.93
Median/$\sigma_{\mathbf{SM}}$	**53.89**	**31.06**	**20.50**	**13.74**	**10.05**	**7.76**	**6.45**	**5.46**	**4.93**	**4.18**
$+1\sigma/\sigma_{SM}$	77.86	44.80	29.44	19.84	14.57	11.20	9.34	7.96	7.17	6.08
$+2\sigma/\sigma_{SM}$	111.06	63.41	41.03	27.97	20.64	15.92	13.10	11.21	10.23	8.73
Observed/$\sigma_{\mathbf{SM}}$	**97.86**	**47.62**	**33.52**	**26.33**	**17.90**	**13.24**	**12.37**	**11.40**	**10.93**	**8.48**

	160	165	170	175	180	185	190	195	200	
$-2\sigma/\sigma_{SM}$	1.76	1.69	1.95	2.21	2.63	3.47	4.20	4.76	5.72	
$-1\sigma/\sigma_{SM}$	2.42	2.29	2.64	3.04	3.67	4.76	5.80	6.81	8.06	
Median/$\sigma_{\mathbf{SM}}$	**3.44**	**3.26**	**3.78**	**4.34**	**5.29**	**6.90**	**8.38**	**9.84**	**11.69**	
$+1\sigma/\sigma_{SM}$	5.00	4.73	5.48	6.33	7.81	10.12	12.44	14.40	17.25	
$+2\sigma/\sigma_{SM}$	7.34	6.88	7.93	9.33	11.19	14.52	18.03	20.75	24.60	
Observed/$\sigma_{\mathbf{SM}}$	**7.20**	**6.53**	**7.54**	**9.89**	**10.14**	**12.01**	**16.45**	**19.39**	**23.60**	

Table 55: $H \rightarrow WW$ sensitivity normalized to NNLL σ_{SM} calculation in the 2J region. In this region we consider potential signal contributions from gluon fusion, vector boson fusion, and associated production with either a W or Z boson.

10 Low M_{ll} Analysis

In this section we present an analysis of dilepton events having an invariant mass below the cut specified for the opposite-sign zero jet and one jet events, but with all other cuts the same. For this analysis the $M_{\ell\ell}$ cut is reversed to become $M_{\ell\ell} < 16$ GeV/c^2. Two extremely powerful control regions are used for this part of the analysis. One of these regions gives an excellent description of the kinematic parameters in the dominant background and the other provides a control of backgrounds that are more difficult to model. While these regions have been used in the full Higgs to WW analysis, in this analysis there are dramatic changes in the character of the background events of which they are composed. The dominant background is from $W\gamma$ and is easily separated from the Higgs signal. The WW background is much reduced but because of the kinematic restriction much more closely resembles the Higgs signal. The low mass region potentially has Drell-Yan as an important background with contamination from the Υ resonances. For the portion of the Drell Yan and dilepton resonances that are unmodelled, the characteristics are obtained from the data. This region is used to rule out the presence of background from sequential b decays. The unmodelled Drell-Yan background is mostly removed by the $\not{E}_{Tspec}$ cut familiar from the other dilepton analyses. The amount of this background that may remain is estimated by extrapolating the $\not{E}_{Tspec}$ from below the $\not{E}_{Tspec}$ cut to the region above. The shape of the neural network distribution for the unmodelled background is extracted from the BaseMllInvLowMetSpec control region that is defined by events having a $\not{E}_{Tspec}$ value in a band of values just below the $\not{E}_{Tspec}$ cut. Furthermore the same sign control region for the low $M_{\ell\ell}$ analysis is 90% $W\gamma$ events and provides a quantitive measure of the $W\gamma$ rate that is used to correct the background computations in the signal region both in this analysis and in the $M_{\ell\ell} > 16$ GeV/c^2 analysis. The reason that this particular analysis is optimal for measuring this background has been investigated and is understood as follows. The other dominant background comes from jets faking leptons. Fake leptons arise in W plus jet events where the lepton from the W or the fake lepton should satisfy a 20 GeV requirement. The second lepton should have $p_t > 10$ GeV if its real or $p_t > 15$ GeV if it is fake. The invariant mass of these combinations is mostly higher than 16 GeV/c^2. Thus the low $M_{\ell\ell}$ region is very rich in the $W\gamma$ background and provides a measure of that rate.

10.1 Data Sample and Leptons

Except when otherwise noted the data and Monte Carlo samples as well as the lepton selection are those described in Sections 3 and 4. However, events containing leptons with energies above 400 GeV are rejected. For this analysis the nontriggerable fake rates were computed and are shown for the three regions used in Tables 27 and 28; however, it will be shown that an important part of this analysis is the fact that the backgrounds can be measured in the control region. It is shown that W+jets contributes less than 6%

Background	Events	Yields in 5.3 fb^{-1}		
		Total	High S/B	Low S/B
Fake	**501**	**12.6**	**7.0**	**5.6**
W e γ	**6087**	**41.5**	**19**	**23**
W μ γ	**4505**	**26.6**	**18**	**8.3**
W τ γ	162	1.1	0.6	0.5
$t\bar{t}$	599	0.38	0.36	0.02
WZ	5688	0.33	0.30	0.03
ZZ	1531	0.12	0.10	0.02
WW(MC@NLO)	**3291**	**12.6**	**11.5**	**1.1**
DY(ee)	10	0.49	0.41	0.08
DY($\mu\mu$)	47	3.2	1.8	1.4
DY($\tau\tau$)	3	0.1	0.1	0.0
DYLow(ee)	5	0.54	0.23	0.30
DYLow($\mu\mu$)	5	0.44	0.27	0.16
Total		100	60	40
WW(Pythia)	**41436**	**11.4**	**10.4**	**1.0**
Data		104	55	49
Signal(ggH at 160)	**15009**	**1.02**	**0.93**	**0.09**

Table 56: Background composition and number of observed events in the high $\not\!\!E_{Tspec}$ low $M_{\ell\ell}$ region (the BaseMllInv region).

and hence the importance of these nontriggerable fake rates is minor, reflecting a small correction to small differences between dilepton types. Limited statistics lead mostly to the values being near 0.

10.2 High $\not\!\!E_{Tspec}$ Low $M_{\ell\ell}$ Event Composition

This region has been studied using the usual Monte Carlo datasets for the high M_{ll} regions. Table 56 shows the composition of the background event sample in Monte Carlo. The number of Monte Carlo events available is also listed. The dominance of the Wγ events is very striking. The WW background is significantly smaller in this sample as are the Fakes. The rest of the backgrounds are negligible. The number of data events is consistent with the total background. The number of expected signal events in $gg \to H$ at each mass point are given in Table 57.

Since most events are Wγ, it suggests that the same sign control region should provide an excellent measure of this background. Thus an BaseMllInvSameSign region has been defined which simply requires the same sign on the leptons but otherwise passes all the BaseMllInv cuts. This control region is studied in Section 10.3.

M_h $(\mathrm{GeV/c})^2$	Events	Total	High S/B	Low S/B
110	17092	0.185	0.167	0.019
115	17851	0.307	0.278	0.029
120	18073	0.450	0.404	0.045
125	18179	0.614	0.555	0.058
130	17422	0.744	0.675	0.069
135	16893	0.861	0.780	0.081
140	16703	0.949	0.861	0.088
145	15956	0.988	0.899	0.089
150	15289	0.995	0.906	0.089
155	15036	1.000	0.913	0.087
160	15009	1.020	0.926	0.094
165	13855	0.893	0.809	0.084
170	12711	0.738	0.670	0.068
175	11482	0.592	0.536	0.055
180	10017	0.456	0.414	0.042
185	9044	0.331	0.302	0.029
190	7898	0.242	0.221	0.021
195	7118	0.191	0.175	0.016
200	6105	0.147	0.135	0.012

Table 57: Signal event yields in 5.3 fb $^{-1}$ in the BaseMllInv region after correction for the $gg \rightarrow H$ p_T spectrum as described in Section 7.3.

10.3 Control Regions

As was already mentioned, the same sign control region is clearly indicated. Another control region that seems well suited to this analysis is the LowMetSpec region.

The composition of the events in the MllInvSameSign region is shown in Table 58. The total numbers of each background is shown for the sample of MC corresponding to an integrated luminosity of 5.3 fb^{-1}. As expected the $W \to \ell\gamma$ events dominate the sample, contributing 87% while fakes contribute another 11%. The WW events are completely absent as are the rest of the backgrounds. This provides an excellent sample of the dominant background and since the sign of the conversion electron is arbitrary, it also gives a precise prediction of kinematic composition of the main background in this sample.

The breakdown of the number of predicted events by dilepton category is shown for each region in Tables 59, 60 and 61 for the sample of MC corresponding to an integrated luminosity of 5.3 fb^{-1}. For the MllInvSameSign region the behavior is as one would expect. There are no expected Higgs events. There are no events arising in the $\mu\mu$ channel. The $W \to e\gamma$ channel has contributions in the ee channel where a conversion electron has paired with the electron from the W, but not the $e\mu$ channel since there is no source of conversion muons. The crack tracks show a contribution since they are sometimes electrons. For the $W \to \mu\gamma$ channel, one sees no contribution from ee but the $e\mu$ channel dominates since a conversion electron from the photon pairs with the muon from the W decay. The numbers of events are small but data and Monte Carlo show good agreement when divided into the dilepton type categories. The signal region, MllInv, also shows a behavior similar to the MllInvSameSign region for the $W\gamma$ background. The WW and Higgs signal are distributed as one would expect just from lepton counting. The MllInvLowMetSpec is underestimated although Fakes seem the largest contributor to the small prediction. The data populate the dilepton categories where the leptons have the same flavor but there are no $e\mu$ events. One would expect these mixed flavor events if the background were due to sequential b decays. Furthermore it will be shown that what background is predicted in this region comes from the low mass Drell Yan sample.

The variables that will be used in the training of the neural net have been examined in the BaseLowMllSameSign sample and their distributions are shown in Figure 60. The corrections described in Section 10.4 have been applied. It can be seen that there is excellent agreement throughout. The choice of variables and their discriminatory power are described in Section 10.6.

The MllInvLowMetSpec region is compared to the MC prediction in Figure 61. It can be seen that there is almost no MC prediction for this region. This raises concern that there could be contamination of the signal region by processes that are not modelled in the simulation and could be mistaken for a Higgs signal. In order to investigate this possible contamination the $\not\!\!E_{Tspec}$ spectrum is examined in the BaseMllInv region before and after the $\not\!\!E_{Tspec}$ cut is applied. This is shown in Figure 61 and the sample of events is referred to as the BaseMllInvNoMetSpec region which includes the entirety of the BaseMllInv

Background	Total	Hi S/B	Low S/B
Fake	**8.8**	4.0	4.8
W e γ	**42**	19	23
W μ γ	**27**	18	9
W τ γ	1.35	0.76	0.59
$t\bar{t}$	0	0	0
WZ	0.31	0.28	0.03
ZZ	0.02	0.02	0
WW(MC@NLO)	0.12	0	0.12
DY(ee)	0.11	0.06	0.05
DY($\mu\mu$)	1.65	0.74	0.91
DY($\tau\tau$)	0.01	0.01	0
Total	81	43	38.5
Data	81	38	43

Table 58: Composition of background events in the MllInvSameSign in 5.3 fb^{-1} of data. The numbers for W$\ell\gamma$ are shown after correction and hence agreement of Data and Monte Carlo is primarily by design.

region. The rate of these events in the signal region is extracted in Section 10.5. It can be seen that the agreement after the $\not{E}_{Tspec}$ cut is significantly better than before the cut is applied.

The Drell Yan computation was done in two pieces, one with the CKIN parameter that specifies the propagator mass to be above 20 GeV/c^2 and the other, referred to as the Drell Yan low mass sample, with the parameter set to be between 10 and 20 GeV/c^2. The Drell Yan in the BaseMllInvNoMetSpec histogram at low $\not{E}_{Tspec}$ is entirely from this Drell Yan low mass sample. This suggests that ideally a lower cut should be used and that the upsilon resonance region should be modelled. For the analysis at hand the control region is used to estimate this Drell Yan contribution and its potential signal-like behaviour.

The $\not{E}_{Tspec}$ cut was also released for the BaseMllInvSameSign region and the result is shown in Figure 61. This is called the BaseMllInvSameSignNoMetSpec region. As for the BaseMllInv, the BaseMllInvSameSign region is contained entirely with the BaseMllInvSameSignNoMetSpec region. It is clear that there is a significantly smaller contribution from this background in the same sign sample. The BaseMllInvLowMetSpec is used to estimate the dynamics of the background. It is shown in Fig. 68 that the unmodelled events populate low values of the neural net score distribution based on a neural net trained on Monte Carlo events passing the BaseMllInv selection.

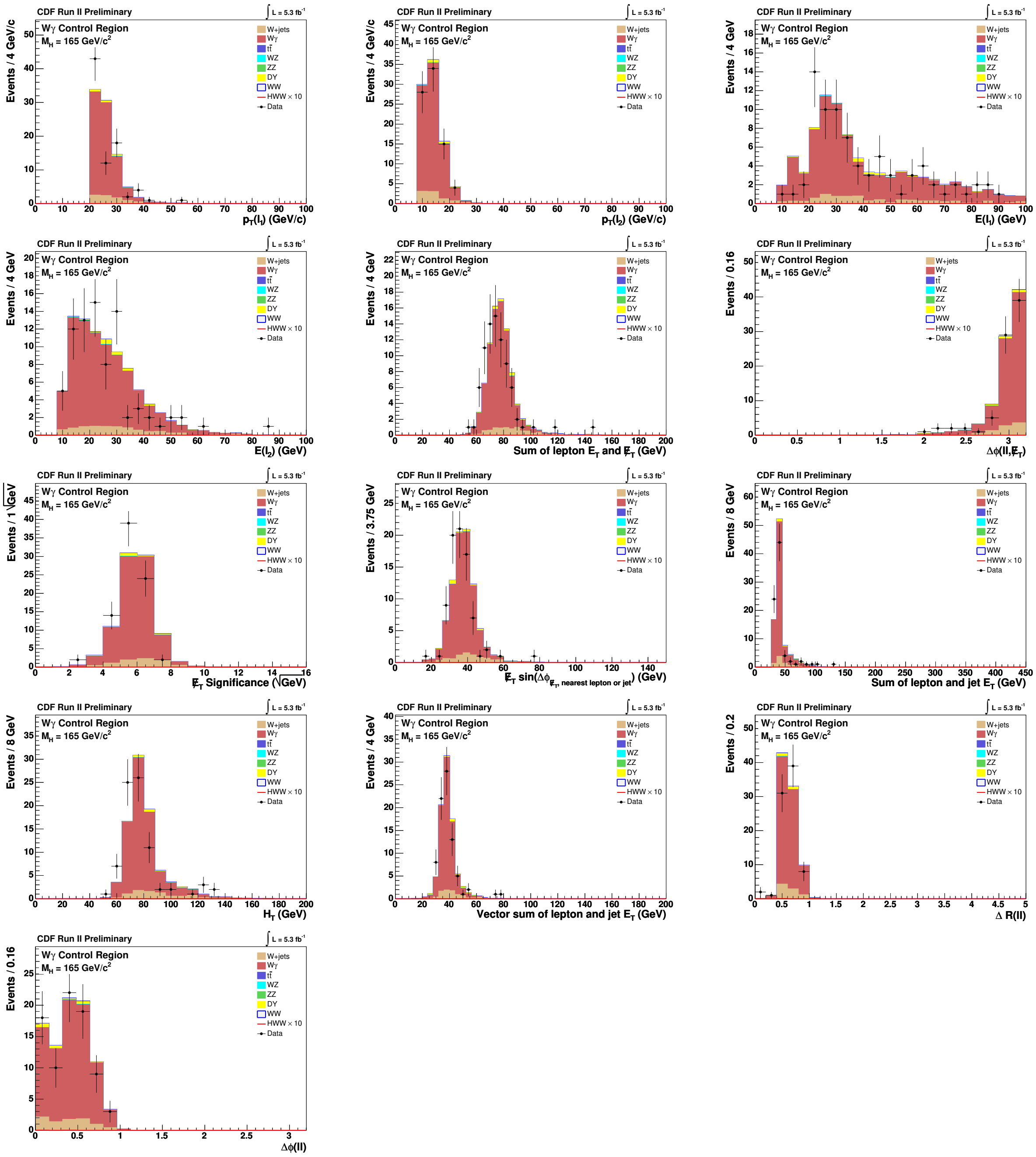

Figure 60: Neural network input variables for the same sign control region of low invariant mass dilepton events with zero or one jets in the final state. The Higgs boson signal is shown for a Higgs of mass $m_H = 160$ GeV$/c^2$ for the $gg \to H$ production mode.

MllInvSameSign						
Background	Total	EE	EMU	MUMU	EKRK	MUKRK
Fake	**8.76**	4.03	2.6	0.27	1.46	0.39
W e γ	**41.75**	**37.39**	0	0	**4.36**	0.0
W μ γ	**26.81**	0	**22.25**	0	**2.7**	**1.86**
W τ γ	1.34	0.65	0.49	0	0.17	0.03
$t\bar{t}$	0	0	0	0	0	0
WZ	0.31	0.06	0.13	0.06	0.03	0.03
ZZ	0.02	0	0.01	0.01	0	0
WW	0.12	0.07	0.03	0	0.01	0
DY(ee)	0.11	0.11	0	0	0	0
DY($\mu\mu$)	1.65	0.0	1.33	0	0.17	0.15
DY($\tau\tau$)	0.01	0	0	0	0.01	0
Total	80.85	42.31	26.84	0.34	8.90	2.46
Data	81	46	24	0	11	0

Table 59: The breakdown by dilepton type of background events in the BaseMllInvSameSign control region in the 5.3 fb^{-1} sample. The numbers for W$\ell\gamma$ are shown after correction and hence agreement of Data and Monte Carlo is primarily by design.

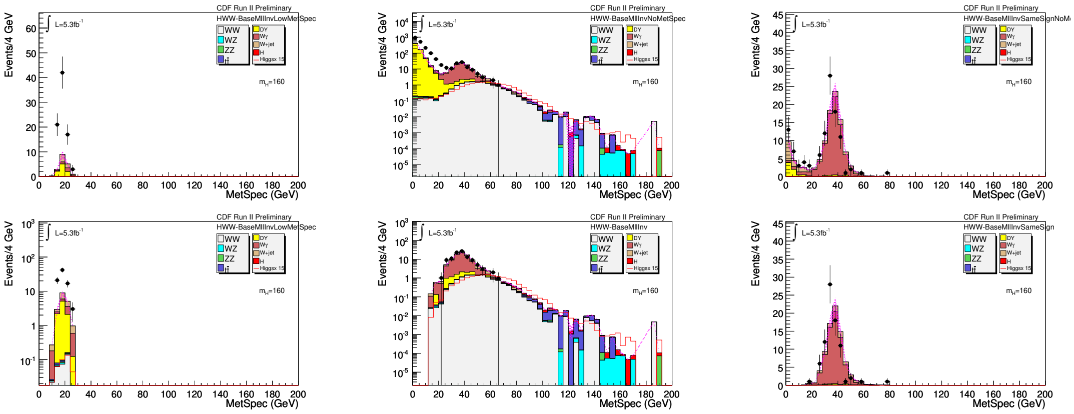

Figure 61: $\not\!\!E_{T spec}$ distribution in the BaseMllInvLowMetSpec sample for data and Monte Carlo on a linear and logarithmic scale (left). $\not\!\!E_{T spec}$ distribution in the BaseMllInv sample for data and Monte Carlo without and with the $\not\!\!E_{T spec}$ cuts (middle). $\not\!\!E_{T spec}$ distribution in the BaseMllInvSameSign sample for data and Monte Carlo without and with the $\not\!\!E_{T spec}$ cuts (right).

MllInv						
Channel	Total	EE	EMU	MUMU	EKRK	MUKRK
W e γ	41.54	**36.74**	0	0	4.79	0
W μ γ	26.61	0	**21.93**	0.01	2.93	1.74
W τ γ	1.15	0.49	0.49	0	0.13	0.04
$t\bar{t}$	0.38	0.08	0.16	0.06	0.05	0.03
WZ	0.33	0.06	0.14	0.07	0.03	0.03
ZZ	0.12	0.05	0.01	0.04	0.02	0.01
WW	12.57	2.49	**5.11**	**2.44**	1.60	0.92
DY(ee)	0.49	0.44	0	0	0.05	0
DY($\mu\mu$)	3.19	0	2.18	0.31	0.41	0.29
DY($\tau\tau$)	0.11	0.04(0.34)	0	0.04	0	0.03
DYLow(ee)	0.54	0.54	0	0	0	0
DYLow($\mu\mu$)	0.44	0	0	0.44	0	0
Total	100.09	46.89	32.94	4.07	12.38	3.81
Data	104	54	33	3	12	2
ggH160	1.02	0.22	0.39	0.17	0.14	0.10

Table 60: The breakdown by dilepton type of background events in the BaseMllInv signal region in the 5.3 fb^{-1} sample.

MllInvLowMetSpec						
Background	Total	EE	EMU	MUMU	EKRK	MUKRK
Fake	5.46	4.27	0.14	0.18	0.46	0.41
W e γ	2.43	2.1	0	0	0.33	0
W μ γ	0.51	0	0.32	0	0.12	0.08
W τ γ	0.23	0.15	0	0	0.07	0.01
$t\bar{t}$	0.03	0.01	0.01	0	0	0
WZ	0.02	0.01	0	0.01	0	0
ZZ	0.01	0.01	0	0	0	0
WW	0.30	0.04	0.06	0.11	0.05	0.05
DY(ee)	0.19	0.19	0	0	0	0
DY($\mu\mu$)	0.10	0	0.02	0.04	0	0.04
DY($\tau\tau$)	0.04	0	0	0	0.03	0
DYLow(ee)	4.96	4.08	0	0	0.88	0
DYLow($\mu\mu$)	3.88	0	0	3.01	0	0.87
Total	18.16	10.86	0.55	3.35	1.94	1.46
Data	83	46	0	25	8	4

Table 61: The breakdown by dilepton type of background events in the BaseMllInvLowMetSpec control region in the 5.3 fb^{-1} sample.

10.4 Corrections to Backgrounds and Final Predictions

A number of corrections have been applied and are summarized in Table 62.

With the higher statistics data sample, a slight discrepancy of data and Monte Carlo is observed. The prediction is in excess of the data and therefore unlikely to be due to a background that is not modeled. If there a deficit of events, then BaseMllInvLowMetSpec region would be used to extrapolate the unmodeled event rate to the signal region; however, there is no indication that this is needed. Indeed it seems there is a slight overestimate of the backgrounds although it is only a 1σ effect.

The number of $W\gamma$ events can be derived from the BaseMllInvSameSign region after subtracting the non-$W\gamma$ contributions. The error on the subtraction is dominated by the fake subtraction but since there are only 8.8 fake events predicted the error, based on the systematic uncertainty on the fakes, is 2.1 fake events. The statistical error is 8.4 events derived from the number of predicted $W\gamma$ events after correction with the BaseMllInvSameSign data and hence the total error on the number of $W\gamma$ events is 8.6 events. The number of predicted $W\gamma$ events is 75.9. There are 81 same sign events measured in the data and subtracting the total predicted number of non-$W\gamma$ events, 11, one has 70 $W\gamma$ events in the BaseMllInvSameSign data. Thus the correction is 70/76 = 0.92 with an error of 8.6/76=0.11. Hence the systematic uncertainty in the $W\gamma$ prediction is 11%. A correction of 0.92 is applied to the number of $W\gamma$ events predicted in the BaseMllInv region.

For the analysis of trileptons in Section 12 the Zγ samples described in Sec. 4.1 are used. These samples are generated in the same manner as the $W\gamma$ samples, and thus we assume the same γ conversion scale factor derived for the $W\gamma$ is applicable to the Zγ samples. However, the Zγ samples have not yet been generated beyond period 11, so an alternate scale factor must be derived using only the period 0-11 $W\gamma$ samples. With these samples, the number of $W\gamma$ candidates without correction is 50.0 W e γ, 32.6 W $\mu\gamma$, and 1.5 W $\tau\gamma$ for a total of 84.1. The total non-$W\gamma$ is the same, 11, and there are 81 events observed, so the $W\gamma$ contribution is 81-11 = 70. This gives a correction of 70/84.1 = 0.83. The statistical error on the $W\gamma$ events is 9.2 and with the error of 2.1 fake events described in the preceeding paragraph this gives 9.4 events which is an error of 9.4/84.1 = 0.11. Hence the systematic uncertainty in the Zγ prediction for the samples used in the trilepton analysis is 11%.

An additional correction is required. The MC@NLO WW prediction is not computed for the full run range and as in Section 4.1, a correction is applied. The value of the correction is derived from a comparison of the Pythia WW generation in the full run range to the Pythia simulation for the same run range as the MC@NLO simulation. Pythia predicts 11.442 events in the full run range 12.631 events in the range of runs from periods 0 to 7. Therefore a factor of 0.906 is applied to the prediction of WW in MC@NLO. This exercise is repeated for each of the control regions and the numbers in Table 62 are obtained.

Item	Correction
BaseMllInv WWNLO	0.906
BaseMllInvSameSign WWNLO	0.971
BaseMllInvLowMetSpec WWNLO	1.012
$W\gamma$	0.92

Table 62: Summary of corrections applied to the ntuples.

The $gg \to H$ signal was also corrected with the p_T weighting described in Section 7.3. With all corrections applied, the final predictions of backgrounds are given in Table 63 for the signal and control regions.

10.5 Unmodelled background rates

It has been shown in Figure 61 that the unmodelled background in the BaseMllInvNoMetSpec sample is significant below the $\not{E}_{Tspec}$ cut used for the BaseMllInv signal sample. It has also been shown in Figure 61 that there is significantly less unmodelled background in the BaseMllInvSameSignNoMetSpec region. Finally, the separation of the data by dilepton category of events in the BaseMllInvLowMetSpec region was shown in Table 61.

The data in Figure 61 were fit to an exponential in the region $0 \leq \not{E}_{Tspec} \leq 25$ GeV. The resulting fit is shown in Figure 62. The fit was repeated with the known background subtracted from the data and this is also shown. The fit was extrapolated (as indicated by the red part of the curve) to the signal region, $\not{E}_{Tspec} > 25$ GeV. The data for this region are shown in Figure 62 both without and with the known background subtracted. The fitted function was evaluated at each bin center and summed to obtain the number of events in the signal region. An estimate of 19 events is obtained.

The exercise was repeated for three other ranges of $\not{E}_{Tspec}$. The variation in the results is summarized in Table 64 and from this we obtain an error on the estimate of 5 events based on half the maximum variation.

10.6 BaseLowMll Analysis

Since the data are exclusive of all the analyses done in the Higgs to WW decay mode thus far, this analysis is treated independently for later combination. The training uses the Neurobayes program [31] and the weighting of events is done as described in the previous analyses [35][36]. TMVA was also used and gives similar results. The input variables chosen were based on those kinematic values used in CDF9685 and are listed in Table 65. Only the $\not{E}_T$ significance of leptons and jets has not been used in the high M_{ll} analyses.

The input distributions are shown in Figure 63. The Higgs signal multiplied by 10 is superimposed as a histogram in red. It can be seen that the shape of the $W\gamma$ background

Background	MllInv	MllInvSameSign	MllInvLowMetSpec
Integrated Lumi (fb^{-1})	5.3	5.3	5.3
Fake	12.6	8.8	5.5
W e γ	41.5	42	2.4
W μ γ	26.6	26	0.51
W τ γ	1.1	1.34	0.23
$t\bar{t}$	0.38	0	0.03
WZ	0.33	0.31	0.02
ZZ	0.12	0.02	0.01
WW(MC@NLO)	12.6	0.12	0.30
DY(ee)	0.49	0.11	0.19
DY($\mu\mu$)	3.2	1.65	0.10
DY($\tau\tau$)	0.1	0.01	0.04
DYLow(ee)	0.5	0	5.0
DYLow($\mu\mu$)	0.4	0	3.9
Total	100	80	18.23
Data	104	81	83
Signal(ggH at 160)	1.02	0	0.02

Table 63: Background composition and number of observed events in the signal and control regions after correction of the $W\gamma$ contributions based on the MllInvSameSign values and a correction of the WW background to account for the full run dependence.

$\not{E}_{Tspec}$ Fit Range (GeV)	Events with $\not{E}_{Tspec} > 25$ GeV/c^2 No Bkg Sub (Events)	Bkg Sub (Events)
0-25	19	18
0-15	23	22
10-25	15	12
5-20	14	13

Table 64: Variation of computation of the nonmodelled background yield with fit range in $\not{E}_{Tspec}$.

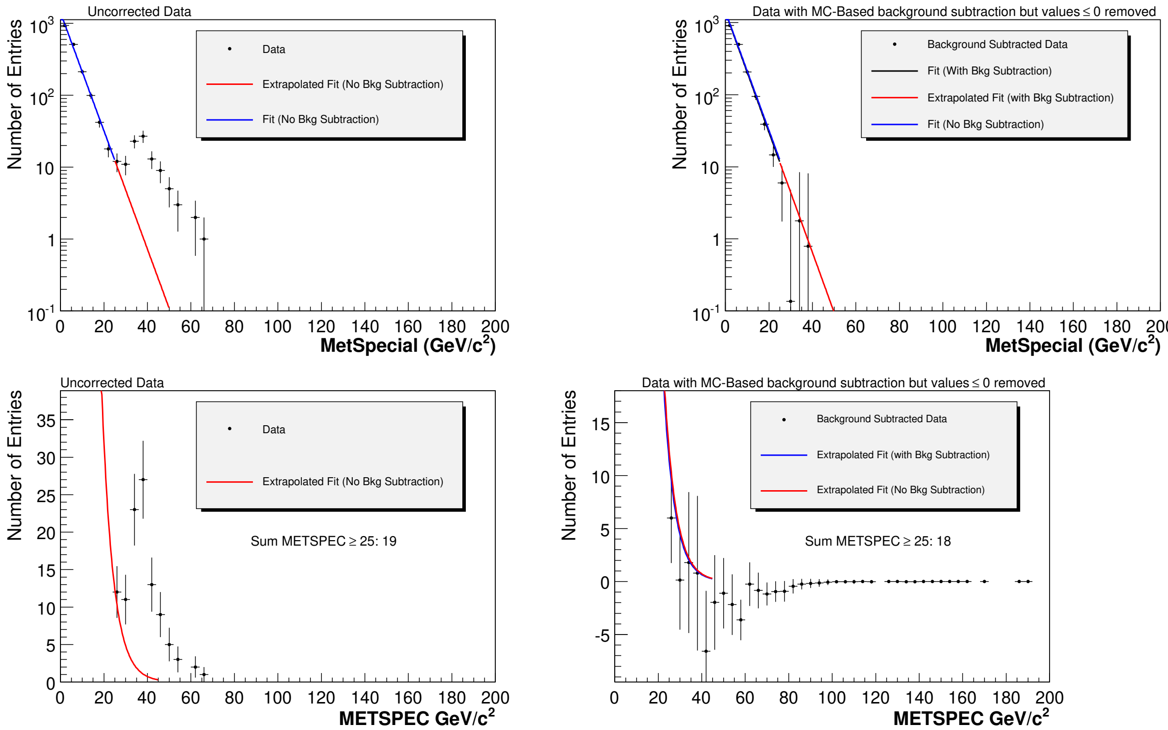

Figure 62: Fit to the $\not\!\!E_{T spec}$ distribution in the BaseMllInvNoMetSpec spectrum. The fitted region in the fit result is show in black and corresponds to $0 \leq \not\!\!E_{T spec} \leq 25$ GeV. The extrapolation of the curve into the BaseMllInv signal region is indicated in red. The data used for the graphs on the left have no subtraction of known backgrounds while those on the right do have a background subtraction. The graphs on the top show the fit on a log scale over the entire BaseMllInvNoMetSpec region. The graphs on the bottom show the fit but only for points having $\not\!\!E_{T spec} > 25$ GeV/c^2. When bins are negative, they are set to zero in order to make the plot on a logarithmic scale.

LepBPt	Lower energy lepton Pt
LepAPt	Higher energy lepton Pt
lep1_E	Higher energy lepton Pt
lep2_E	Lower energy lepton Pt
addEt	Scalar sum of the Et of leptons and $\not{E}_T$
dPhiLepVSumMet	Azimuthal angle between the $\not{E}_T$ and the vector sum of the Lepton Pt
MetSigLeptonJets	$\not{E}_T$ significance with sum of energy over leptons and jets
MetSpec	$\not{E}_T$ multiplied by the sin of the angle between the Met and the closest lepton or jet
dRLeptons	Lepton separation in $\sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$
Ht	Scalar sum of the Et of leptons and jets and the $\not{E}_T$
VSumJetLeptonsPt	Magnitude of the vector sum of Pt of the leptons and jet(s)
SumEtLeptonsJets	Scalar sum of the Lepton and Jet Et
dPhiLeptons	azimuthal angle between leptons

Table 65: Variables used in the neural network training listed in order of the importance as determined by the neural network for a Higgs of mass 160 GeV.

is the same as that in the BaseMllInvSameSign region. Also, the WW and Higgs are distinguished from the $W\gamma$ background but have similar shapes, which results in a poor signal-to-background limit. Further training specifically against the WW background to exploit its differences from the signal could yield a better significance.

10.7 Results

The results of the neural net training and fit are given in Table 66 and plotted in Figure 67. For the fit 10000 pseduexperiments were used with 10000 systematic samplings for each. For the observed limits 500000 systematic samplings were used. The neural net distributions are shown in Figures 64 through 66. Excellent separation of the $W\gamma$ background is seen and the Higgs is pushed with reasonable separation from the WW as well.

10.8 Systematic Uncertainties

Since the sample is dominated by 0 jet events, the systematic uncertainties are derived from those in the opposite-sign zero jet region with the following modifications specific to this analysis. The systematic uncertainties are given in Table 67.

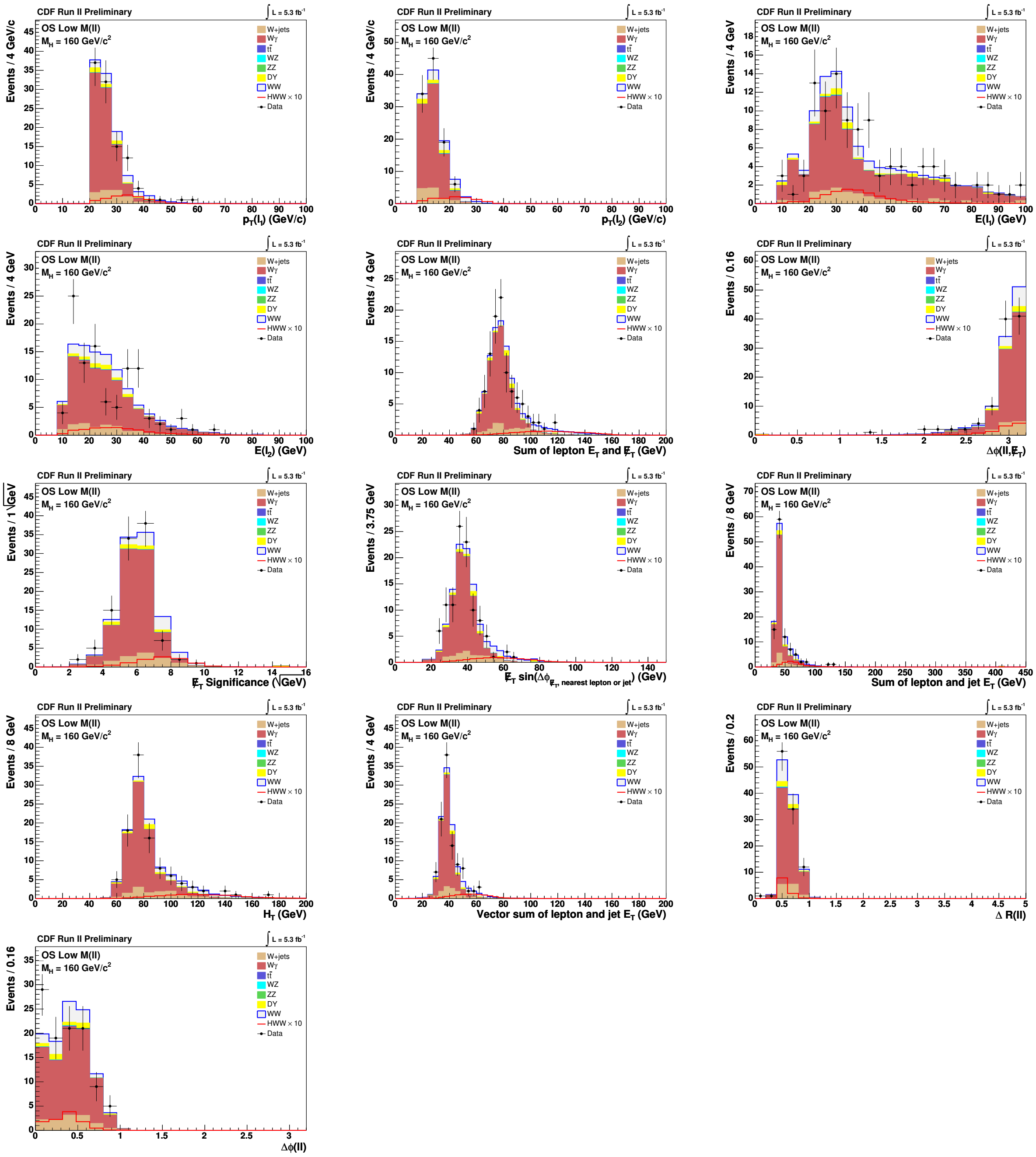

Figure 63: Neural network input variables for the signal region of opposite-sign low invariant mass dilepton events with zero or one jets in the final state. The Higgs boson signal is shown for a Higgs of mass $m_H = 160$ GeV/c^2 for the dominant $gg \rightarrow H$ production modes.

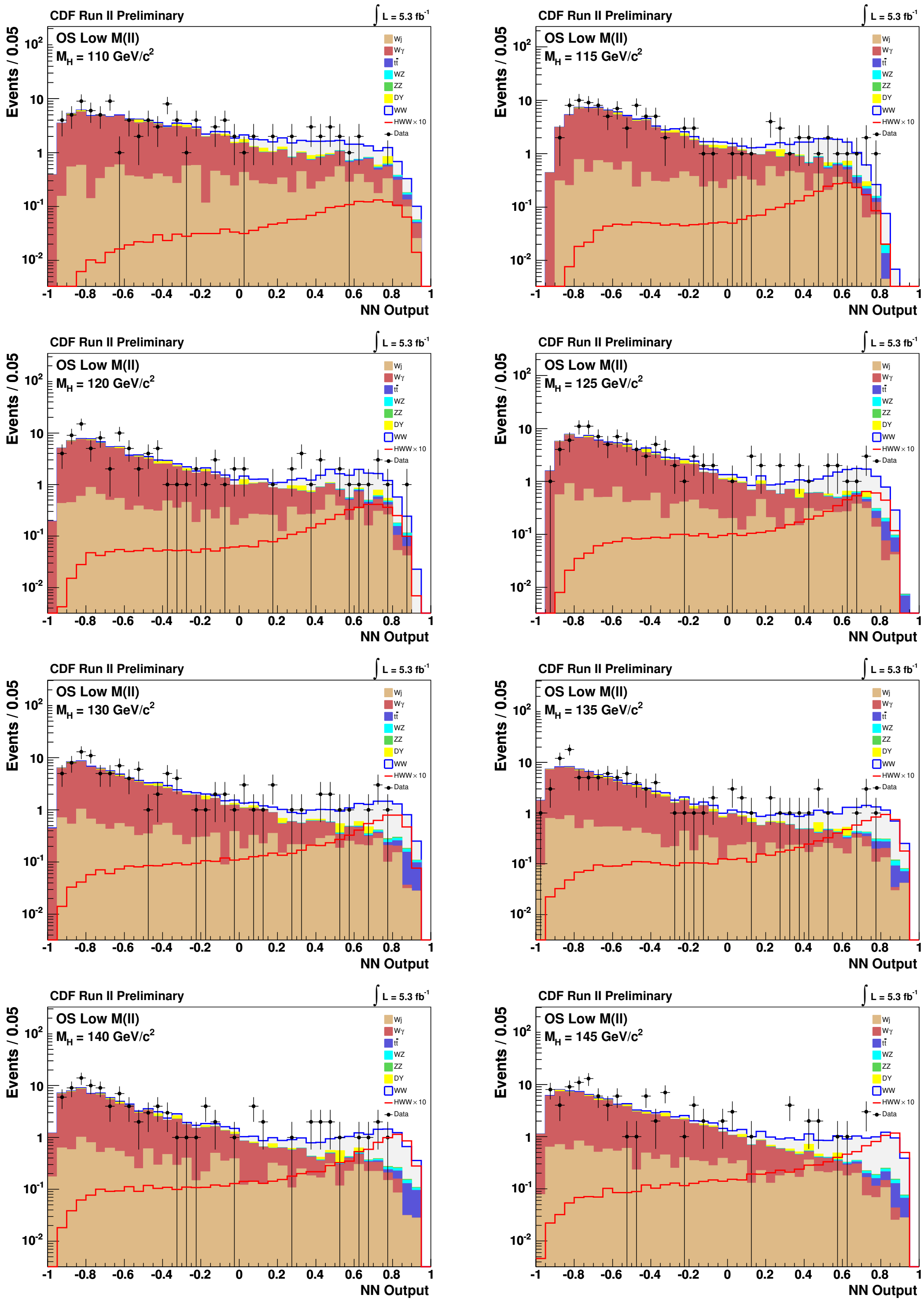

Figure 64: Low M_{ll} Neural Network output templates for several Higgs masses. From the top left to bottom right: 110 GeV, 115 GeV, 120 GeV, 125 GeV, 130 GeV, 135 GeV, 140 GeV, and 145 GeV.

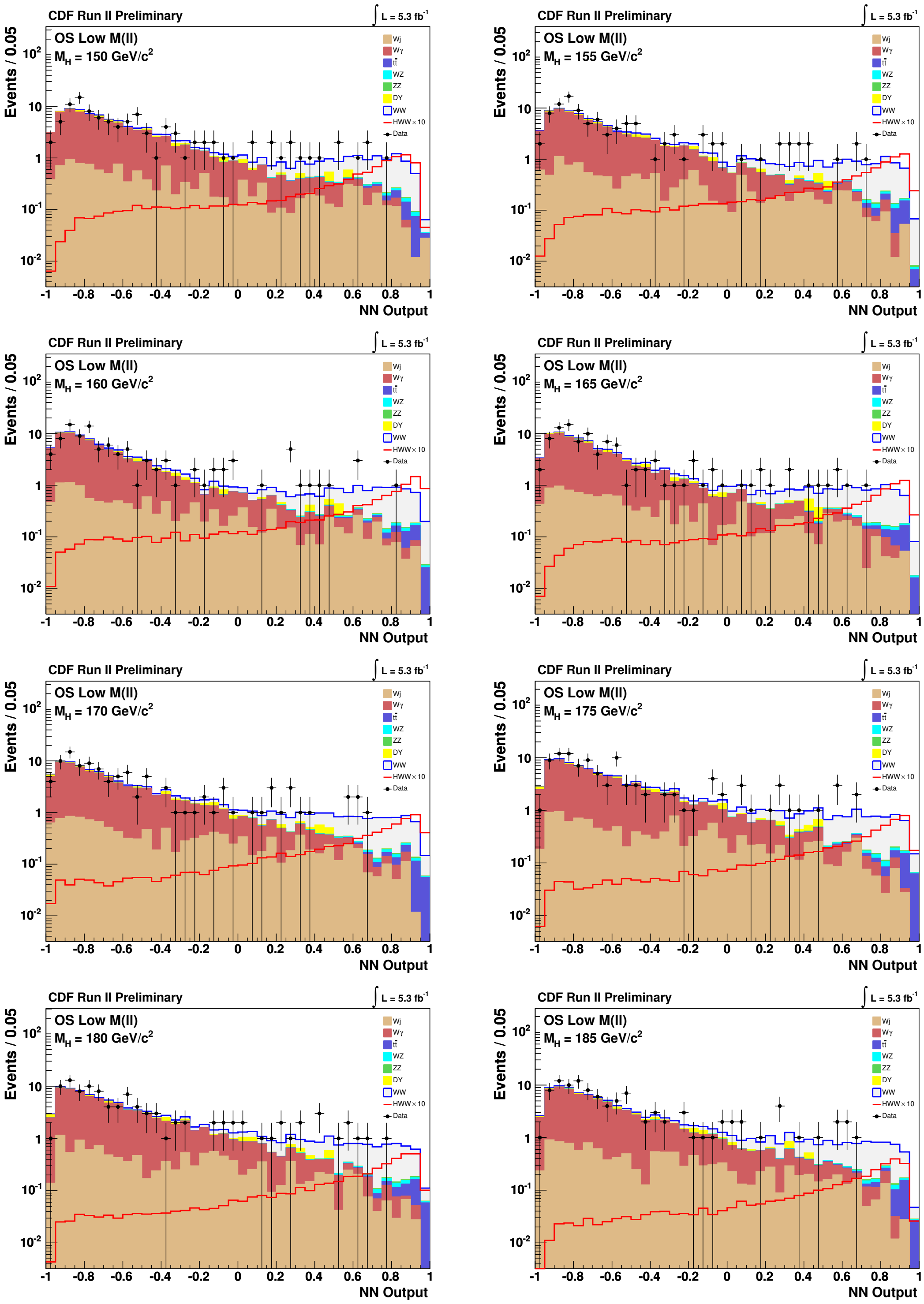

Figure 65: Low M_{ll} Neural Network output templates for several Higgs masses. From the top left to bottom right : 150 GeV, 155 GeV, 160 GeV, 165 GeV, 170 GeV, 175 GeV, 180 GeV, and 185 GeV.

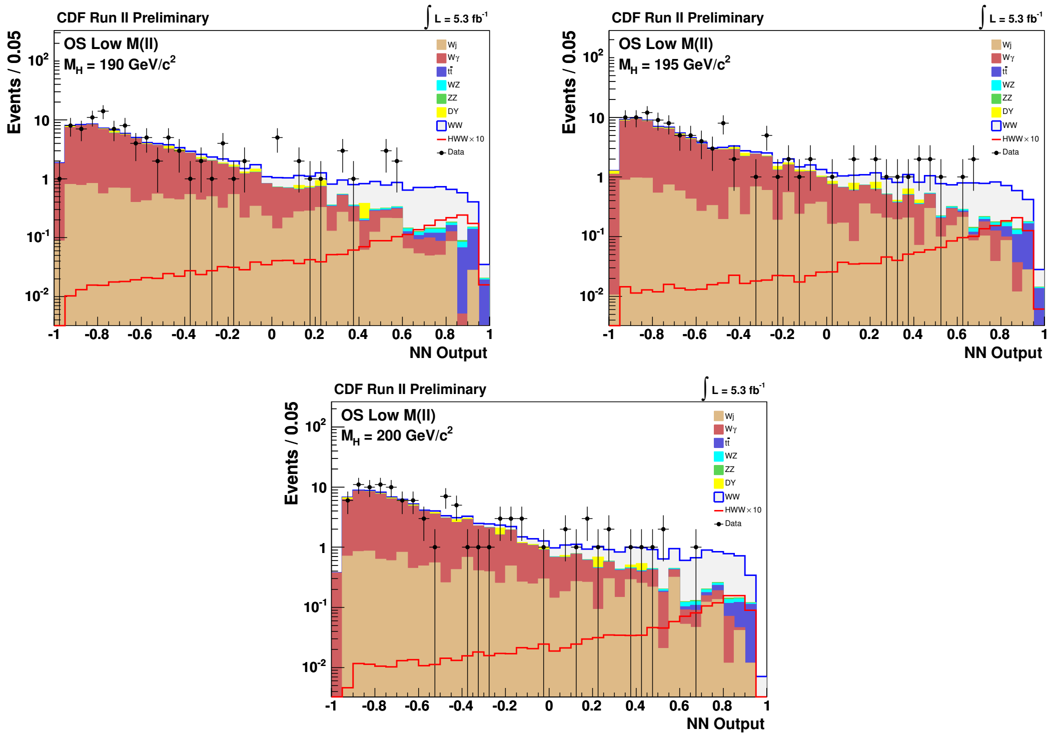

Figure 66: Low M_{ll} Neural Network output templates for several Higgs masses. From the top left to bottom right : 190 GeV, 195 GeV, and 200 GeV.

OS Low M(ll)	110	115	120	125	130	135	140	145	150	155
$-2\sigma/\sigma_{SM}$	42.04	25.37	16.63	12.08	9.63	7.84	7.09	6.59	6.27	6.07
$-1\sigma/\sigma_{SM}$	56.37	33.76	22.19	15.99	12.81	10.39	9.31	8.66	8.23	7.99
Median/$\sigma_{\mathbf{SM}}$	**79.12**	**47.03**	**30.86**	**22.40**	**17.88**	**14.52**	**12.98**	**12.07**	**11.45**	**11.02**
$+1\sigma/\sigma_{SM}$	111.46	66.77	44.22	31.66	25.26	20.88	18.47	17.31	16.36	15.77
$+2\sigma/\sigma_{SM}$	156.76	93.80	61.23	44.65	35.07	28.88	25.90	24.20	22.77	21.93
Observed/$\sigma_{\mathbf{SM}}$	**107.75**	**53.55**	**30.70**	**18.42**	**14.60**	**11.39**	**10.08**	**8.87**	**8.81**	**7.95**

	160	165	170	175	180	185	190	195	200	
$-2\sigma/\sigma_{SM}$	5.61	6.39	7.79	9.74	12.81	18.03	24.28	31.54	40.44	
$-1\sigma/\sigma_{SM}$	7.27	8.38	10.19	12.78	16.73	23.72	31.67	41.01	53.44	
Median/$\sigma_{\mathbf{SM}}$	**10.02**	**11.64**	**14.17**	**17.78**	**23.28**	**32.67**	**44.05**	**55.97**	**74.19**	
$+1\sigma/\sigma_{SM}$	14.18	16.50	20.37	25.30	32.85	46.61	62.53	79.81	104.94	
$+2\sigma/\sigma_{SM}$	19.82	22.85	28.33	35.03	46.55	65.68	87.63	112.31	146.10	
Observed/$\sigma_{\mathbf{SM}}$	**7.96**	**7.83**	**9.68**	**11.55**	**19.89**	**24.75**	**29.45**	**40.78**	**48.56**	

Table 66: $H \to WW$ sensitivity normalized to NNLL σ_{SM} calculation in the low M_{ll} region. In this region we consider potential signal contributions from gluon fusion only.

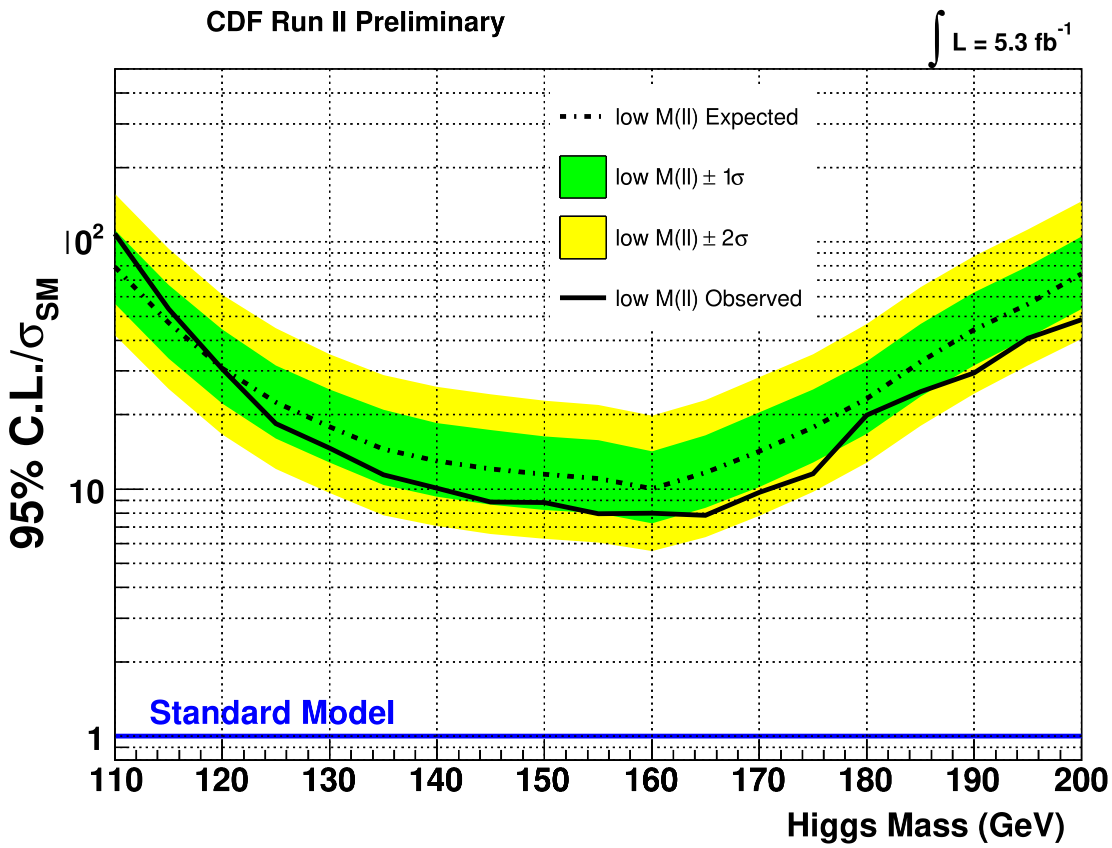

Figure 67: $H \to WW$ sensitivity normalized to NNLL σ_{SM} calculation in the low M_{ll} region. In this region we consider potential signal contributions from gluon fusion only.

- The uncertainty in the $W\gamma$ events is dominated by the statistics with a tiny correction for the subtraction of fakes in the same sign control region as was described in Section 10.4.
- The evaluation of the systematic uncertainty in the WW yield computed with MC@NLO is determined by comparing the full run range and the restricted run range in Pythia as was described in Section 10.4.
- The effect of higher order diagrams is computed by comparing the number of events in WW MC@NLO to those in Pythia for the same run range, periods 0 to 7. From Table 56 WW MC@NLO has 12.6 events. The Pythia prediction for periods 0-7 as indicated in Section 10.4 is 11.4 events. Thus the correction from LO to NLO is 12.6/11.4= 1.11. Following the methods used in the Base0Jet region we assign a systematic uncertainty for NNLO corrections to be half the LO to NLO difference, 5.5%, for WW and the full difference for ZZ and WZ, 11%.
- The rate of events in the unmodelled background has been computed in Section 10.5 and the shape of the unmodelled neural network distribution is computed using events in the BaseLowMetSpec sample as shown in Figure 68. This shape has a similar distribution to that for $W\gamma$ and since it falls at predominantly low values of neural net score is not expected to impact the fit to the signal region. This has been tested by changing the scale factor on $W\gamma$ from 0.92 to 1.116 to account for the potential additional 19 $\pm$ 5 events (27% of events in $W\gamma$) coming from the unmodelled background. At the same time the systematic uncertainty was increased from 11% to 17% to account from the uncertainty in the modelled background rate. The limits and their changes are shown in Table 68. The changes are quite small and no further treatment of the background is considered. This is as expected since the fit will adjust the normalization of the low end and there is very little background in the signal region.

As a check, the distribution of NN scores for events in the BaseMllInvSameSign control region is shown in Figure 69. The neural net is that trained in the signal region. The same sign control region and the background are well-modelled.

Uncertainty Source	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jet(s)	$gg \to H$
Cross Section								
Scale								12.0%
PDF Model								7.7%
Total	*6.0%*	*6.0%*	*6.0%*	10.0%	5.0%			14.3%
Acceptance								
Scale (leptons)								2.6%
Scale (jets)								1.1%
PDF Model (leptons)	*1.9%*	*2.7%*	*2.7%*	*2.1%*	*4.1%*			*1.7%*
PDF Model (jets)								0.3%
Higher-order Diagrams	*5.5%*	*11.0%*	*11.0%*	10.0%				
Jet Energy Scale	*1.0%*	*2.3%*	*2.0%*	*12.9%*	*6.4%*	*1.3%*		*2.4%*
Missing E_T Modeling					**20.0%**			
Conversion Modeling						11.0%		
Jet Fake Rates							18.7%	
MC Run Dependence	4.7%							
Lepton ID Efficiencies	*2.0%*	*1.7%*	*2.0%*	*2.0%*	*1.9%*			*1.9%*
Trigger Efficiencies	*2.1%*	*2.1%*	*2.1%*	*2.0%*	*3.4%*			*3.3%*
Luminosity	*5.9%*	*5.9%*	*5.9%*	*5.9%*	*5.9%*			*5.9%*

Table 67: Systematics table for low M_{ll} analysis background and signal (only $gg \to H$ production is considered) processes. Cross section systematics are correlated for the diboson processes (shown in italic font) but uncorrelated between all other processes. The DY Met modeling systematic (shown in boldface font) is uncorrelated between channels.

M_h	Fit	Sensitivity			Observed		
	Error	Standard	Varied	Diff	Standard	Varied	Diff
110	2.4	76.3	81.8	-5.5	104.6	95.5	9.1
115	1.5	45.7	46.8	-1.1	53.6	48.5	5.1
120	0.9	29.9	31.7	-1.8	30.0	27.0	3.0
125	0.7	21.5	22.9	-1.4	16.7	15.2	1.5
130	0.5	17.2	18.1	-0.9	14.5	13.1	1.4
135	0.4	14.1	15.0	-0.9	10.9	9.9	1.0
140	0.4	12.7	13.2	-0.5	9.8	8.9	0.9
145	0.4	11.6	12.6	-1.0	8.8	8.0	0.8
150	0.4	11.0	11.9	-0.9	8.4	7.7	0.7
155	0.3	10.5	11.2	-0.7	7.4	6.8	0.6
160	0.3	9.5	10.4	-0.9	7.6	7.0	0.6
165	0.3	11.1	11.7	-0.6	7.2	6.8	0.4
170	0.4	13.4	15.2	-1.8	9.3	8.6	0.7
175	0.5	17.0	18.6	-1.6	11.2	10.5	0.7
180	0.6	22.0	23.7	-1.7	18.8	17.0	1.8
185	1.0	31.6	34.3	-2.7	23.6	21.7	1.9
190	1.3	41.6	46.4	-4.8	28.2	26.3	1.9
185	1.7	55.0	58.4	-3.4	39.4	35.9	3.5
200	2.4	71.4	73.2	-1.8	47.0	43.9	0.9

Table 68: Results varying the $W\gamma$ varied by 27% to account for the unmodelled background. Errors on the $W\gamma$ background were also increased from 11% to 17%. High and Low S/B are treated as separate channels. 1000 pseduexperiments were used with 1000 systematic samplings for each. For the observed limits 10000 systematic samplings were used.

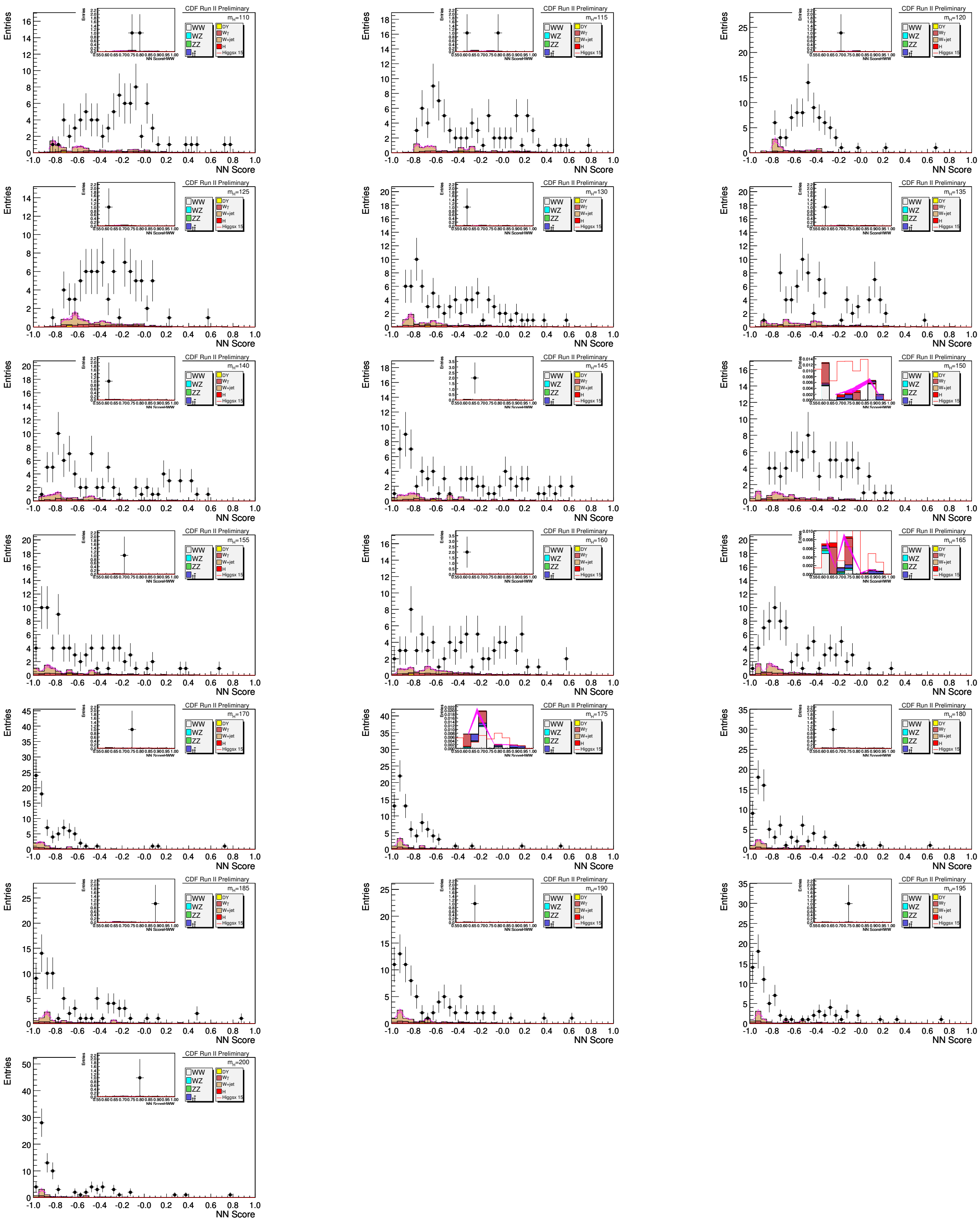

Figure 68: Neural network output templates at each input Higgs boson mass for the BaseMllInvLowMetSpec control region of opposite-sign low invariant mass dilepton events with zero or one jets in the final state.

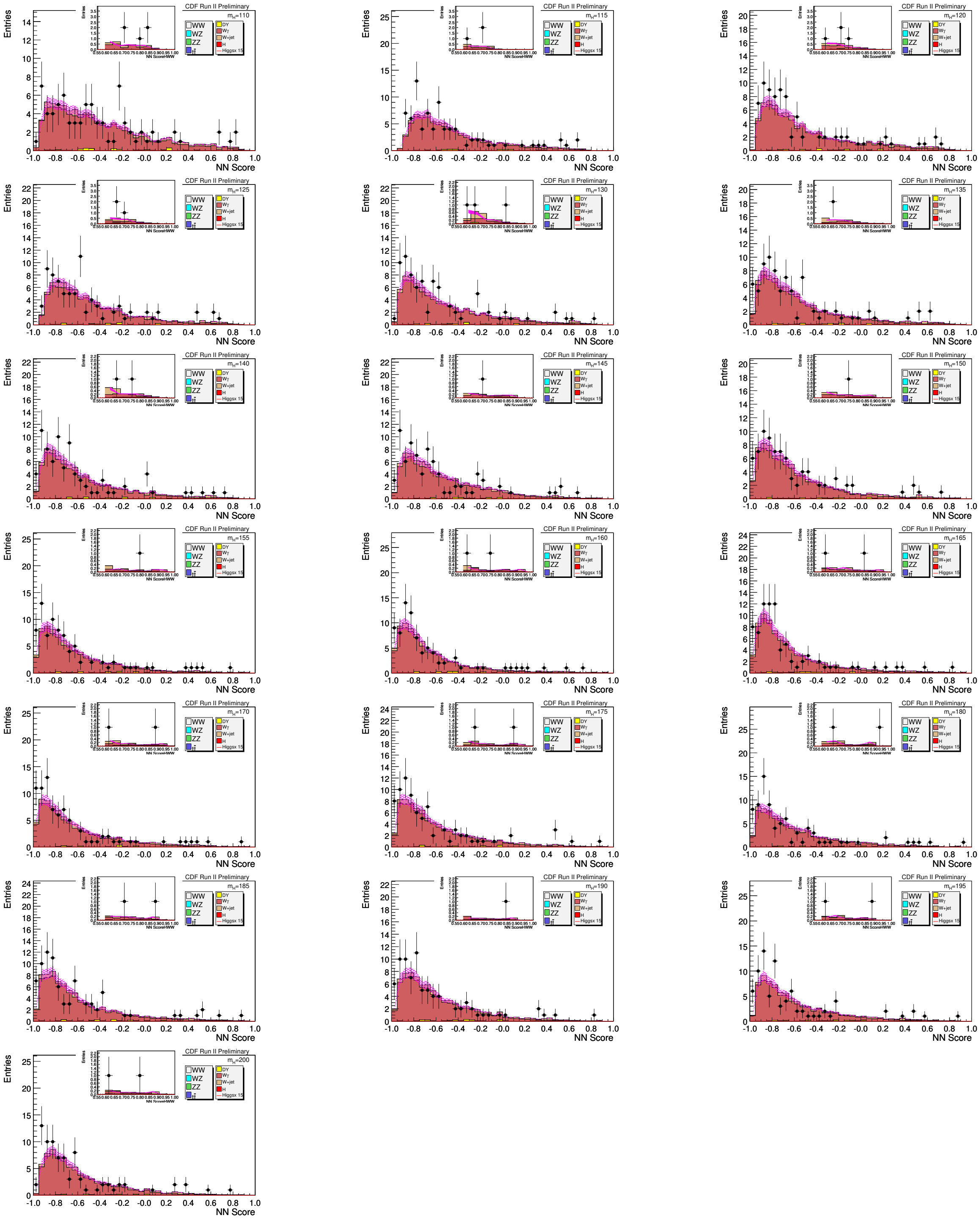

Figure 69: Neural network output templates at each input Higgs boson mass for the BaseMllInvSameSign control region of opposite-sign low invariant mass dilepton events with zero or one jets in the final state.

11 Same-Sign Dilepton Signal Region with 1+ Jets

To gain extra sensitivity, we also search for Higgs signal in like-sign, or same-sign (SS), dileptons. These occur naturally in $ZH \to ZWW$ production, when the Z and one or both W bosons decay leptonically, and $WH \to WWW$ production, when the associated W and one or both W bosons from the Higgs decay leptonically. It is also possible to search for trilepton signatures from these production modes, but this search will be handled as a separate channel to be added in the future.

11.1 Event Selection

The event selection for same sign dilepton events focues on eliminating background from fake and charge misidentified leptons. Because PHX electrons have a high rate of charge mismeasurement, we do not use PHX electrons in the SS analysis. All other lepton types are included. The WH and ZH events tend to have lower values of missing transverse energy because there are three neutrinos, and the leptons are not necessarily back-to-back. Because Drell-Yan contributes to the same-sign region only when a lepton charge is misidentified, it is not as large a background source in the same-sign region as in the opposite-sign. Therefore we remove all $\not\!\!E_T$ and $\not\!\!E_{\mathrm{T}}^{\mathrm{spec}}$ requirements. To reduce the number of fake leptons in the sample, we increase the p_{T} requirement for the second lepton from 10 GeV/c to 20 Gev/c. To reduce the W+jets and $W\gamma$ contributions, we only use events with one or more jets with $E_{\mathrm{T}} > 15$ GeV and $|\eta| < 2.5$ at L5 correction. Same sign events with zero jets are used as a control sample, as described in Section 11.2.

The loosened selection of likelihood-based electrons (Section 2) had a larger impact on the same-sign analysis, where the dominant backgrounds are fake and charge misidentified leptons, than on the opposite-sign analysis. In fact, the rate of charge mismeasurement (as measured in Sec. 11.1) nearly doubles for LBE leptons. With a larger rate of charge mismeasurement, more Drell-Yan events enter the event sample, making Drell-Yan a larger background in the zero jet control region that W+jets. To reduce the now more substantial Drell-Yan background, we introduce a $\not\!\!E_T > 10$ GeV cut in the same-sign region. This removes 49% of the DY background (and nearly 30% of the W+jets), but less than 3% of the expected signal for a Higgs boson of mass 165 GeV.

After all selection cuts, W+jets accounts for about 50% of the background in the same-sign channel. Drell-Yan contributes about 25% of the total background. Smaller background contributions come from WW, WZ, ZZ, $W\gamma$, and $t\bar{t}$ processes. To estimate the backgrounds, we use the same MC samples as used in the opposite-sign HWW analysis. The W+jets contribution is measured from data, as described in Section 4.3.

Non-triggerable Fake Rates

The primary background to the SS analysis is W+jets; thus, it is important to accurately predict the number of W+jets events. The W+jets contribution measured from data must be corrected for non-triggerable fakes as described in Section 4.3. Due to the low statistics in the same-sign regions, it is difficult to measure the non-triggerable fake contributions directly from these regions. Instead, we use the regions with all SS event selections applied, except with opposite-sign leptons instead of same-sign. Because there is no $\not\!\!E_T$ cut (it was not applied at this stage), this region has considerable statistics, primarily due to Drell-Yan events.

The non-triggerable fake rates measured in the opposite-sign regions are then applied to the SS W+jets predictions. The rates, shown in Tables 27 and 28 cause the total number of W+jets events to increase by about 20% in both SS events with 0 jets and SS events with one or more jets.

Charge Fake Rate

WW is expected to be a very small contribution to the same-sign analysis. Even in 5.3 fb^{-1}, only ~ 0.03 WW events are expected. However, it is expected to look very signal-like, so we still want to include it in the neural network training. To model WW we use the WW MC@NLO sample, which has smaller statistics than many of our other MC samples. Running this MC sample through our same-sign event selection, very few MC events pass the selection. There are not enough events to make a NN template from.

To make a smooth NN template, we instead use events which satisfy all same-sign requirements, except the two leptons have opposite signs instead of the same signs. We then multiply the number of events in this template by a charge fake rate, or the likelihood that the charge of one lepton would be misidentified. We measure this charge fake rate (CFR) in the MC as outlined in the following paragraphs.

To measure a charge fake rate for dilepton types, we use the $Z \to ee$, $Z \to \mu\mu$, and WW Pythia samples. To maximize the statistics, we use only very basic selection cuts: lepton ID, trigger lepton requirements, good run list applied, and a ΔZ cut between the two leptons. Using the MC truth information, we compare the reconstructed charge of both leptons to the generated charge of the two leptons. We then count the number of lepton pairs with both charges correctly measured, and the pairs with the charge of one lepton mismeasured. The ratio of these two numbers is the charge fake rate.

Table 69 shows the charge fake rate for all possible dilepton types. Although we do not use PHX electrons, those CFR are included for completeness. The systematic uncertainty due to the modeling of the CFR in MC is discussed in Section 6.4.

Expectation for Base Region

The background expectation and observed yield for the same-sign dilepton signal region for events containing one or more reconstructed jets with $E_T > 15$ GeV are shown in Table 70. The expected and observed kinematic distributions for these events are shown in Figures 70-72. For these plots and tables, the $\not\!\!E_T > 10$ GeV cut has not yet been applied. We stack the backgrounds on top of each other, and then for comparison we

Dilepton type	Opposite-sign events	Same-sign events	Charge fake rate
LBE-LBE	1874600	1404	7.49e-4
LBE-Central μ	810720	373	4.60e-4
LBE-Forward μ	121007	65	5.37e-4
LBE-CrkTrk	121759	633	5.20e-4
Central μ-Central μ	2348216	1	4.26e-7
Central μ-Forward μ	821583	54	6.57e-5
Central μ-CrkTrk	1104521	26	2.35e-5
PHX-PHX	410229	109335	0.267
PHX-Central μ	257241	32070	0.125
PHX-Forward μ	41865	5549	0.133
PHX-CrkTrk	543764	71383	0.131
LBE-PHX	1759455	219633	0.125

Table 69: Calculated charge fake rates (CFR) for all dilepton types, using WW, $Z \to ee$, and $Z \to \mu\mu$ Pythia MC samples. The lepton selection is described in the text. "Central μ" includes CMUP, CMP, CMX, CMXMsKs, CMIOCES, and CMP muons, while "Forward μ" includes CMIOPES and BMU muons.

overlay the contribution from a 160 GeV/c^2 Higgs boson with yields scaled up by a factor of 10. In addition, we calculate a confidence level for data versus background expectations on each plot. The confidence levels shown are for Poisson fluctuations and the KS test.

11.2 Same-Sign Control Region

The primary background in the same-sign signal region is fakes. The best non-overlapping control region to study the agreement of the fakes prediction is in the region of same-sign dileptons with zero jets. Table 71 and Figures 73 through 75 show the agreement between data and the fakes prediction. Again, in these plots and tables the $\not\!\!E_T > 10$ GeV cut has not yet been applied. The agreement is good within statistics for all variables.

Category	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jets	Total	Data
$e\ e$	0.0	2.3	0.5	0.2	30.0	2.2	15.6	50.8	68.0
$e\ \mu$	0.0	4.2	0.9	0.2	4.2	2.8	21.6	34.0	30.0
$\mu\ \mu$	0.0	1.3	0.3	0.0	0.8	0.0	4.7	7.1	7.0
e trk	0.0	2.0	0.4	0.1	10.1	1.9	11.4	25.9	31.0
μ trk	0.0	0.9	0.2	0.0	0.6	0.4	2.7	4.8	3.0
Total:	0.1	10.8	2.2	0.4	45.7	7.4	56.0	122.5	139.0

Category	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jets	Total	Data
LBE LBE	0.0	2.3	0.5	0.2	30.0	2.2	15.6	50.8	68.0
LBE PHX	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PHX PHX	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
LBE CMUP	0.0	1.4	0.3	0.1	1.0	1.0	6.8	10.7	8.0
LBE CMP	-0.0	0.2	0.0	0.0	0.0	0.1	1.7	2.1	3.0
LBE CMX	0.0	0.7	0.1	0.0	0.8	0.5	2.8	5.0	7.0
LBE CMXMsKs	0.0	0.1	0.0	0.0	0.2	0.1	0.6	1.0	1.0
LBE BMU	0.0	0.2	0.0	0.0	0.4	0.1	0.0	0.7	1.0
LBE CMIOCES	0.0	1.1	0.2	0.1	1.4	0.7	2.9	6.3	6.0
LBE CMIOPES	0.0	0.4	0.1	0.0	0.5	0.3	5.6	6.9	4.0
PHX CMUP	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PHX CMP	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PHX CMX	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PHX CMXMsKs	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PHX BMU	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PHX CMIOCES	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PHX CMIOPES	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CMUP CMUP	0.0	0.2	0.0	0.0	0.0	0.0	0.3	0.6	1.0
CMUP CMP	0.0	0.1	0.0	0.0	0.0	0.0	0.2	0.3	0.0
CMUP CMX	0.0	0.2	0.0	0.0	0.0	0.0	0.4	0.6	1.0
CMUP CMXMsKs	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.0
CMUP BMU	0.0	0.1	0.0	0.0	0.0	0.0	0.8	0.9	0.0
CMUP CMIOCES	0.0	0.3	0.1	0.0	0.0	0.0	0.3	0.7	0.0
CMUP CMIOPES	0.0	0.1	0.0	0.0	0.5	0.0	1.0	1.7	3.0
CMX CMP	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.2	0.0
CMX CMX	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.0
CMX CMXMsKs	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CMX BMU	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.3	0.0
CMX CMIOCES	0.0	0.1	0.0	0.0	0.0	0.0	0.3	0.5	0.0
CMX CMIOPES	0.0	0.0	0.0	0.0	0.2	0.0	0.9	1.1	1.0
CMXMsKs CMP	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.0
CMXMsKs CMXMsKs	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CMXMsKs BMU	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CMXMsKs CMIOCES	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.0
CMXMsKs CMIOPES	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.0
CMP CMP	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CMP BMU	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CMP CMIOCES	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.3	0.0
CMP CMIOPES	0.0	0.0	0.0	0.0	0.1	0.0	0.2	0.3	1.0
LBE CrkTrk	0.0	2.0	0.4	0.1	10.1	1.9	11.4	25.9	31.0
PHX CrkTrk	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CMUP CrkTrk	0.0	0.6	0.1	0.0	0.3	0.2	1.5	2.7	2.0
CMP CrkTrk	0.0	0.1	0.0	0.0	0.0	0.0	0.3	0.4	1.0
CMX CrkTrk	0.0	0.3	0.1	0.0	0.3	0.2	0.8	1.6	0.0
CMXMsKs CrkTrk	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.0
Total:	0.1	10.8	2.2	0.4	45.7	7.4	55.6	122.2	139.0

Table 70: Yields in same-sign dilepton signal region for events with one or more jets ($\not\!\!E_T > 10$ GeV cut not yet applied).

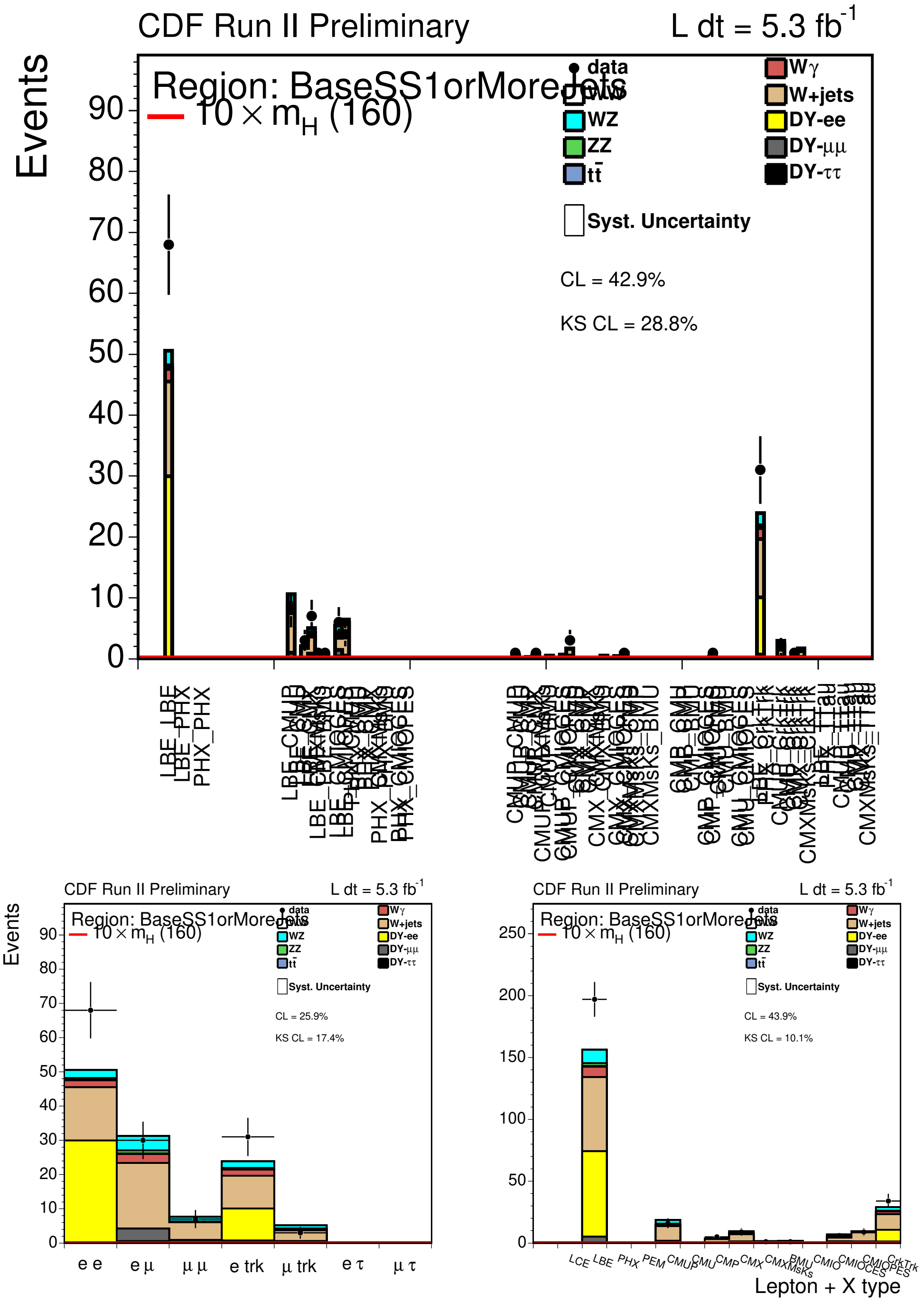

Figure 70: Same-Sign dilepton 1+ jet signal region ($\not\!\!E_T > 10$ GeV cut not yet applied).

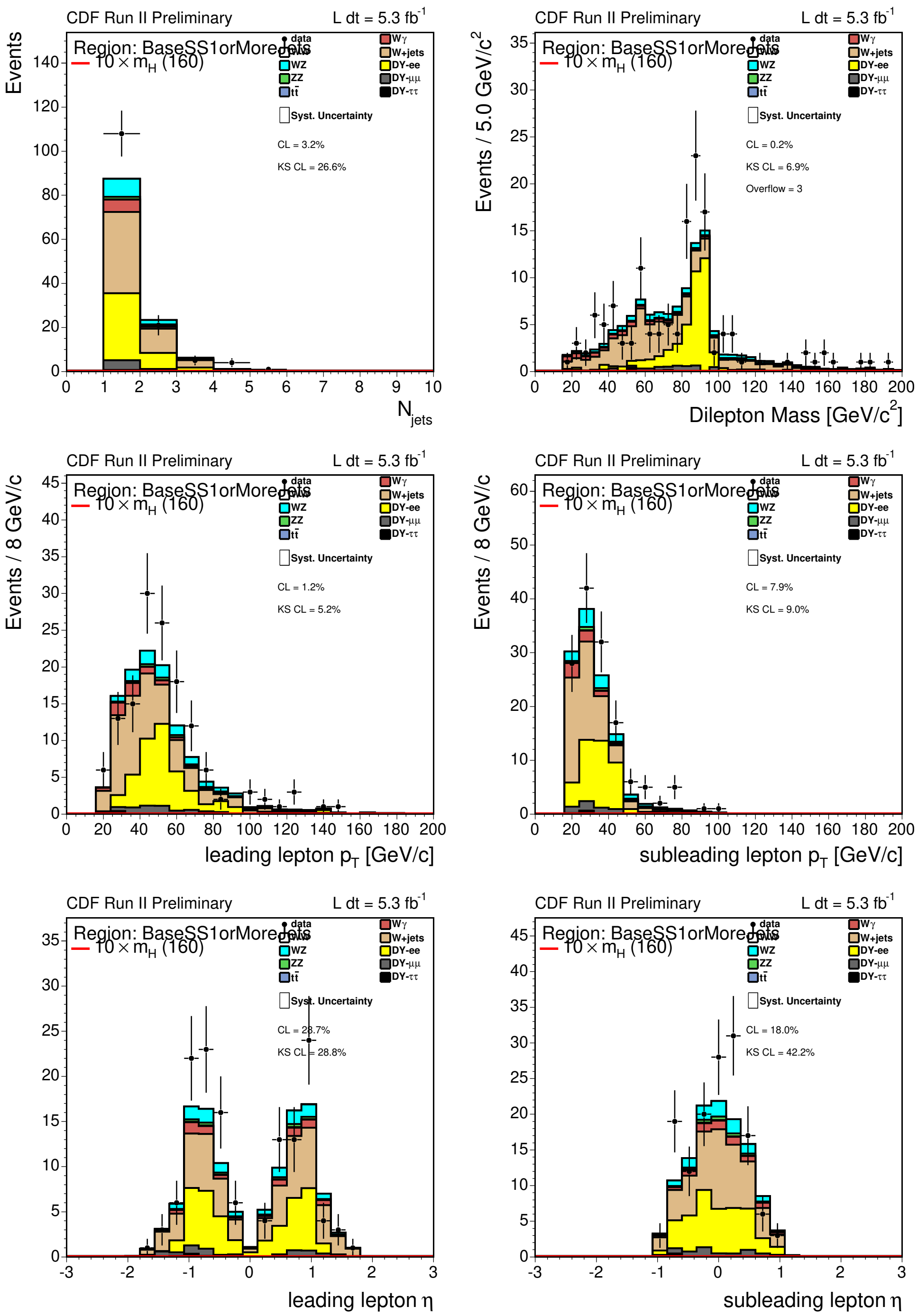

Figure 71: Same-Sign dilepton 1+ jet signal region ($\not\!\!E_T > 10$ GeV cut not yet applied).

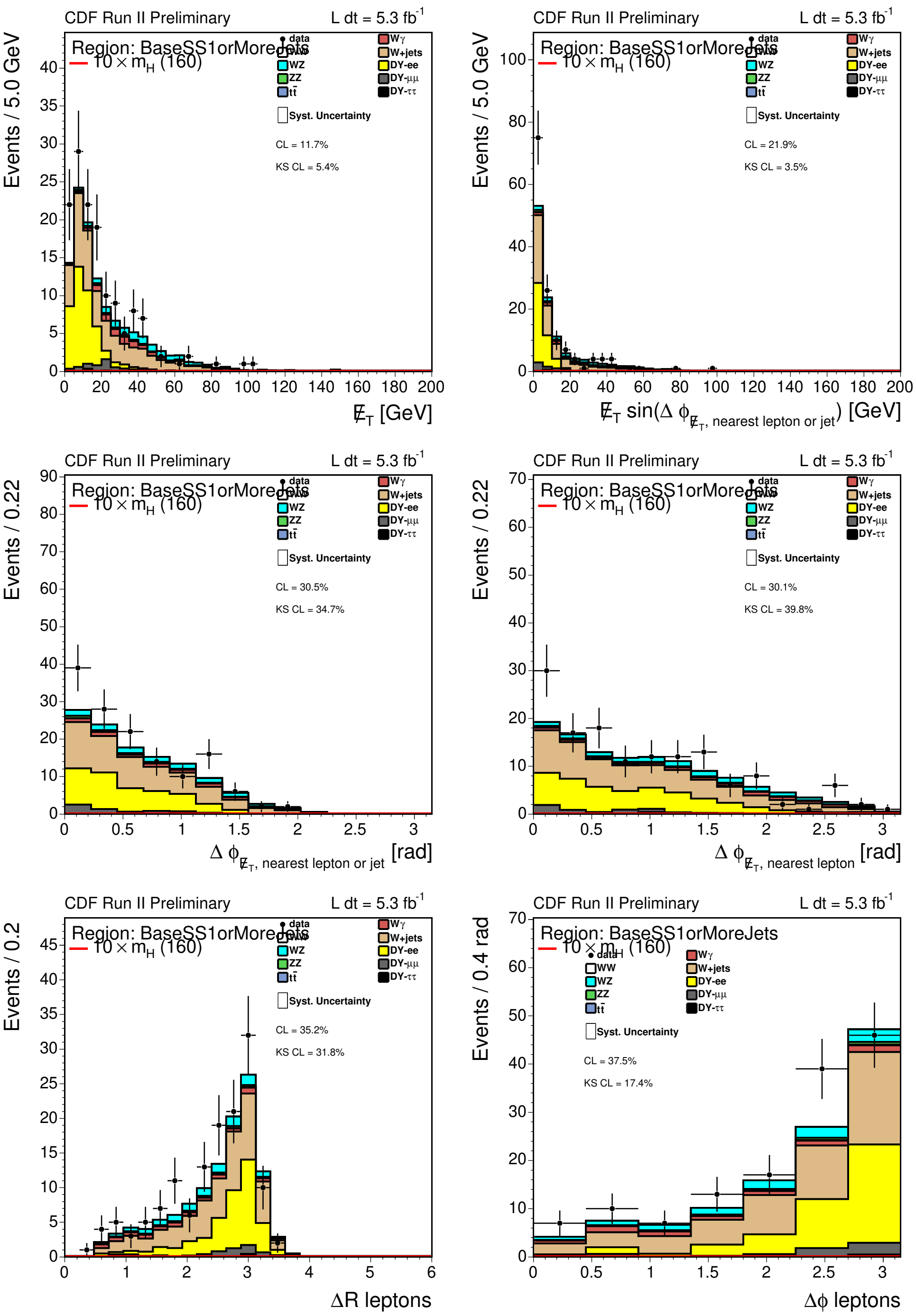

Figure 72: Same-Sign dilepton 1+ jet signal region ($\not\!\!E_T > 10$ GeV cut not yet applied).

Category	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jets	Total	Data
e e	0.1	0.5	0.0	0.0	81.1	5.3	25.1	112.2	146.0
e μ	0.1	3.1	0.1	0.0	9.6	7.3	37.4	57.5	31.0
μ μ	0.0	2.5	0.3	0.0	1.5	0.0	7.8	12.1	18.0
e trk	0.0	0.9	0.0	0.0	25.6	4.2	24.5	55.2	57.0
μ trk	0.0	2.1	0.1	0.0	1.8	1.2	4.8	10.1	5.0
Total:	0.2	9.1	0.6	0.0	119.6	18.0	99.7	247.2	257.0

Category	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jets	Total	Data
LBE LBE	0.1	0.5	0.0	0.0	81.1	5.3	25.1	112.2	146.0
LBE PHX	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PHX PHX	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
LBE CMUP	0.0	1.0	0.1	0.0	4.5	2.6	10.8	19.0	11.0
LBE CMP	0.0	0.1	0.0	0.0	0.5	0.4	2.9	4.0	1.0
LBE CMX	0.0	0.5	0.0	0.0	0.7	1.1	5.6	7.9	2.0
LBE CMXMsKs	0.0	0.1	0.0	0.0	0.4	0.2	1.4	2.1	0.0
LBE BMU	0.0	0.2	0.0	0.0	0.5	0.4	0.0	1.0	1.0
LBE CMIOCES	0.0	0.8	0.0	0.0	2.3	1.9	6.0	11.0	9.0
LBE CMIOPES	0.0	0.3	0.0	0.0	0.7	0.7	9.0	10.8	7.0
PHX CMUP	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PHX CMP	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PHX CMX	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PHX CMXMsKs	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PHX BMU	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PHX CMIOCES	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PHX CMIOPES	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CMUP CMUP	0.0	0.4	0.0	0.0	0.0	0.0	0.7	1.1	1.0
CMUP CMP	0.0	0.1	0.0	0.0	0.0	0.0	0.2	0.3	1.0
CMUP CMX	0.0	0.4	0.0	0.0	0.0	0.0	0.6	1.0	0.0
CMUP CMXMsKs	0.0	0.1	0.0	0.0	0.0	0.0	0.2	0.2	0.0
CMUP BMU	0.0	0.1	0.0	0.0	0.1	0.0	0.3	0.6	1.0
CMUP CMIOCES	0.0	0.5	0.1	0.0	0.0	0.0	1.0	1.6	4.0
CMUP CMIOPES	0.0	0.2	0.0	0.0	0.8	0.0	1.9	2.9	5.0
CMX CMP	0.0	0.1	0.0	0.0	0.0	0.0	0.1	0.2	0.0
CMX CMX	0.0	0.1	0.0	0.0	0.0	0.0	0.2	0.3	1.0
CMX CMXMsKs	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0
CMX BMU	0.0	0.1	0.0	0.0	0.0	0.0	0.1	0.2	2.0
CMX CMIOCES	0.0	0.3	0.0	0.0	0.0	0.0	0.4	0.7	0.0
CMX CMIOPES	0.0	0.1	0.0	0.0	0.4	0.0	1.1	1.6	2.0
CMXMsKs CMP	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CMXMsKs CMXMsKs	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CMXMsKs BMU	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CMXMsKs CMIOCES	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.2	0.0
CMXMsKs CMIOPES	0.0	0.0	0.0	0.0	0.1	0.0	0.3	0.4	1.0
CMP CMP	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.0
CMP BMU	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CMP CMIOCES	0.0	0.1	0.0	0.0	0.0	0.0	0.2	0.3	0.0
CMP CMIOPES	0.0	0.0	0.0	0.0	0.1	0.0	0.5	0.6	0.0
LBE CrkTrk	0.0	0.9	0.0	0.0	25.6	4.2	24.5	55.2	57.0
PHX CrkTrk	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CMUP CrkTrk	0.0	0.7	0.1	0.0	1.1	0.7	2.3	4.8	1.0
CMP CrkTrk	0.0	0.1	0.0	0.0	0.1	0.1	1.7	1.9	1.0
CMX CrkTrk	0.0	0.3	0.0	0.0	0.6	0.4	1.5	3.0	2.0
CMXMsKs CrkTrk	0.0	1.1	0.0	0.0	0.0	0.0	0.3	1.4	1.0
Total:	0.2	9.1	0.6	0.0	119.6	18.0	99.1	246.6	257.0

Table 71: Yields in same-sign dilepton signal region for events with zero jets ($\not\!\!E_T > 10$ GeV cut not yet applied).

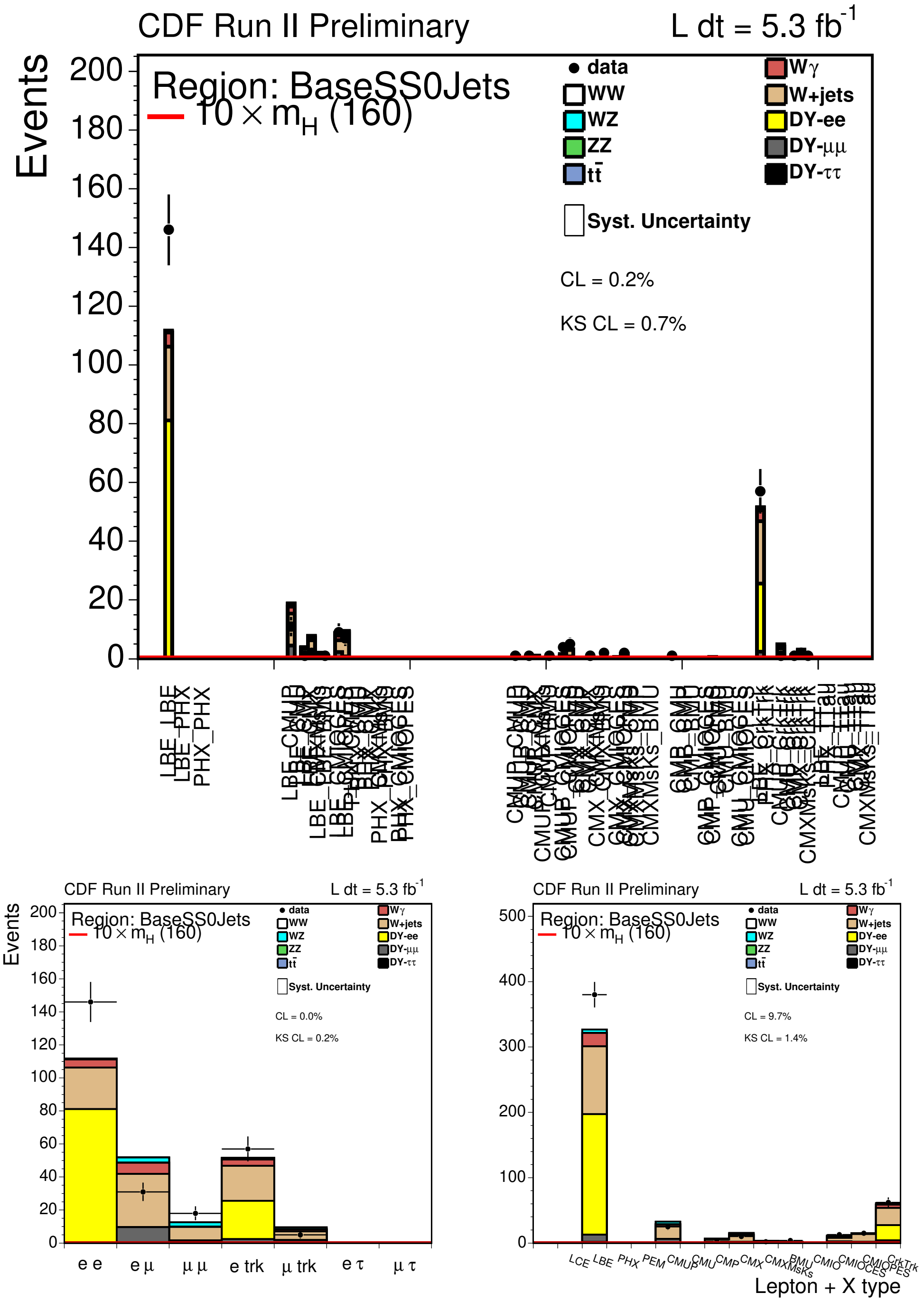

Figure 73: Same-Sign dilepton 0 jet control region ($\not\!\!E_T > 10$ GeV cut not yet applied).

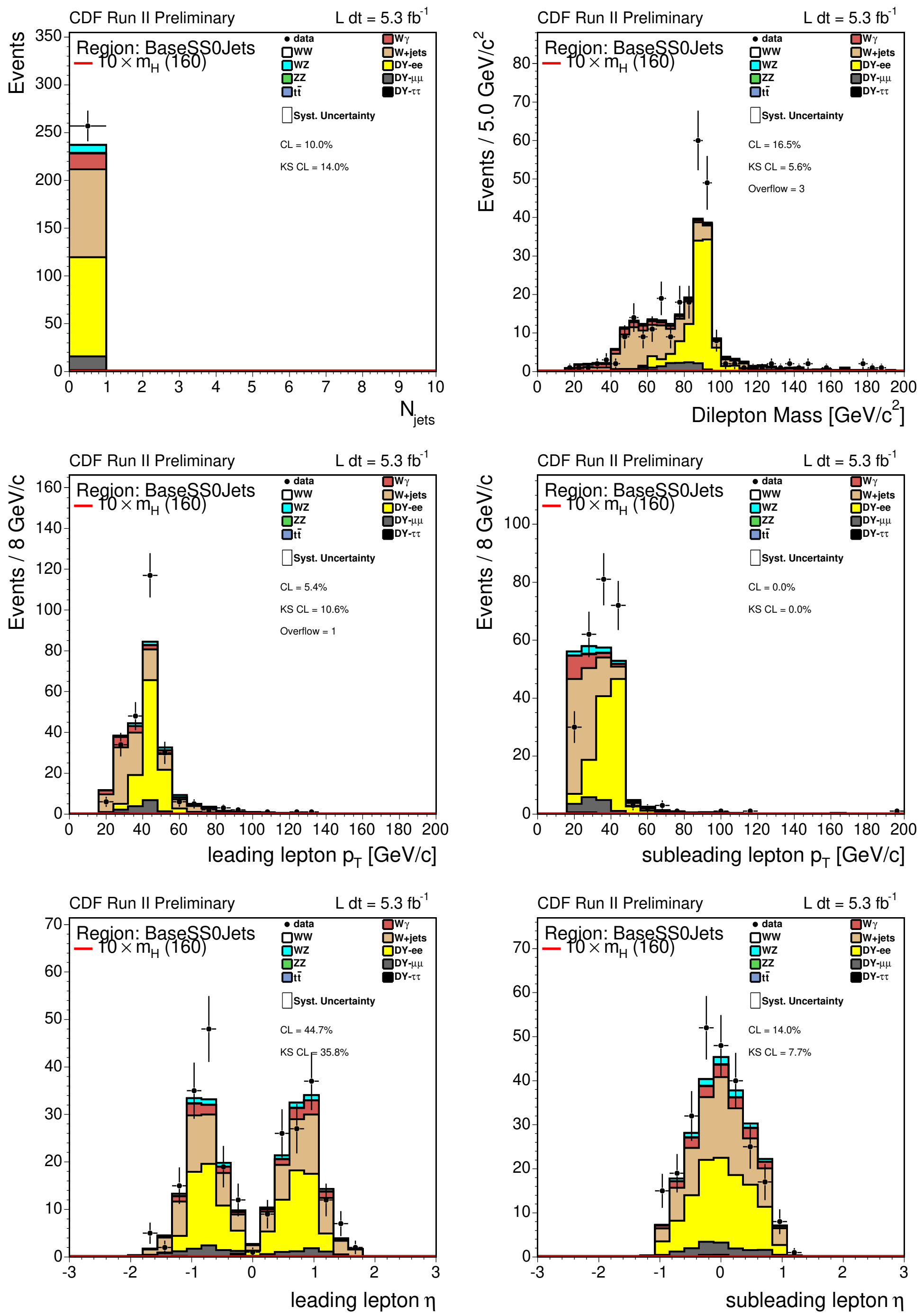

Figure 74: Same-Sign dilepton 0 jet control region ($\not{E}_T > 10$ GeV cut not yet applied).

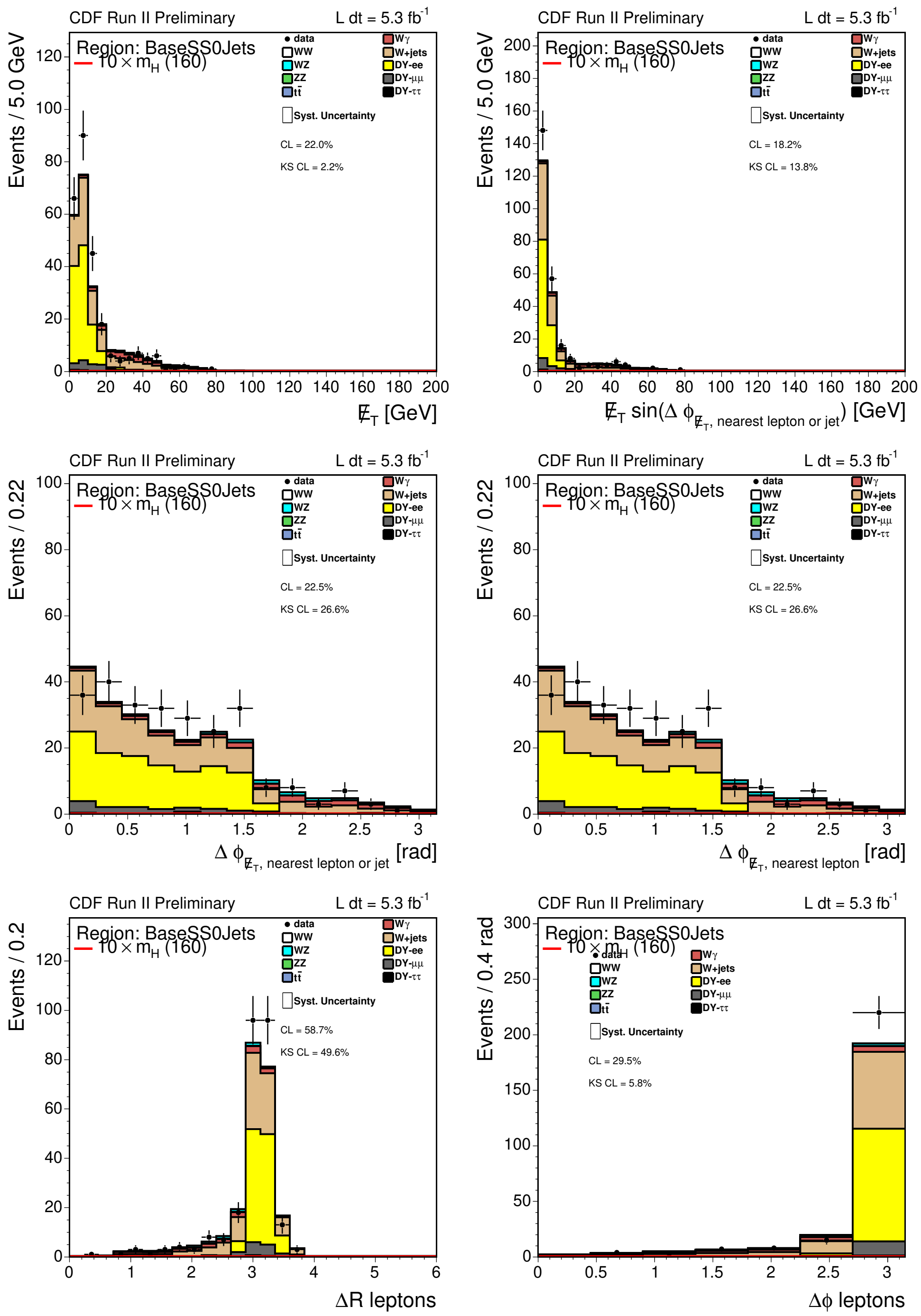

Figure 75: Same-Sign dilepton 0 jet control region ($\not{E}_T > 10$ GeV cut not yet applied).

Variable	Meaning
Met	$\not\!\!E_T$
SumEtJets	Sum E_{T} of all jets
dimass	Dilepton mass M_{ll}
Ht	Transverse mass H_{T}
MetSpec	${\not\!\!E_T}^{\mathrm{spec}}$
jet1Et	E_{T} of leading jet
lep1E	Energy of leading lepton
MetDelPhi	$\Delta\phi$ between $\not\!\!E_T$and nearest lepton or jet
LepSumPt	Vector sum of lepton 1 and 2 p_{T}
dPhiLeptons	$\Delta\phi$ between leptons
MetSig	$\not\!\!E_{T\mathrm{sig}}$
lep2Pt	p_{T} of subleading lepton
lep1Pt	p_{T} of leading lepton

Table 72: Kinematic neural network inputs for the same-sign analysis. Variables are listed in order of decreasing significance to the NN. Plots of the input variables are shown in Figs. 76 through 77.

11.3 Same-Sign Analysis

We do not attempt to calculate Matrix Elements for SS events. Instead, we rely solely on a neural network (NN) to discriminate signal from background using kinematic variables as inputs to the NN. Optimization and training of the neural network are done in the same manner as for the 0J case described in Section 7. For the SS region our neural network uses 13 inputs, 14 hidden nodes and 1 output node. The input variables used for the SS NN are shown in Table 72. The observed distributions of each SS neural network input parameter compared with the distributions predicted by our background model are shown in Figures 76 and 77. For reference, we also include a hypothetical signal component for a Higgs boson with $m_H = 160$ GeV/c^2.

Templates are created for each neural network and used for calculating 95% CL limits with the MCLimit program [32]. Due to the low statistics in the SS region and the fact that we exclude PHX electrons, we do not divide events into High S/B and Low S/B categories as is done for the 0J and 1J regions. Output templates for the trained SS neural network are shown in Figures 78 and 79.

The most significant signal contributions in the SS region are from associated production VH ($q\bar{q} \to ZH \to ZWW$ and $q\bar{q} \to WH \to WWW$). The gluon fusion ($gg \to H \to WW \to ll\nu\nu$) and Vector Boson Fusion (VBF) production methods contribute only when the charge of a lepton is misidentified. Consequently, we only consider WH and ZH production in the SS analysis. We evaluate the number of events expected from each production process as a function of the mass of the Higgs. These numbers are

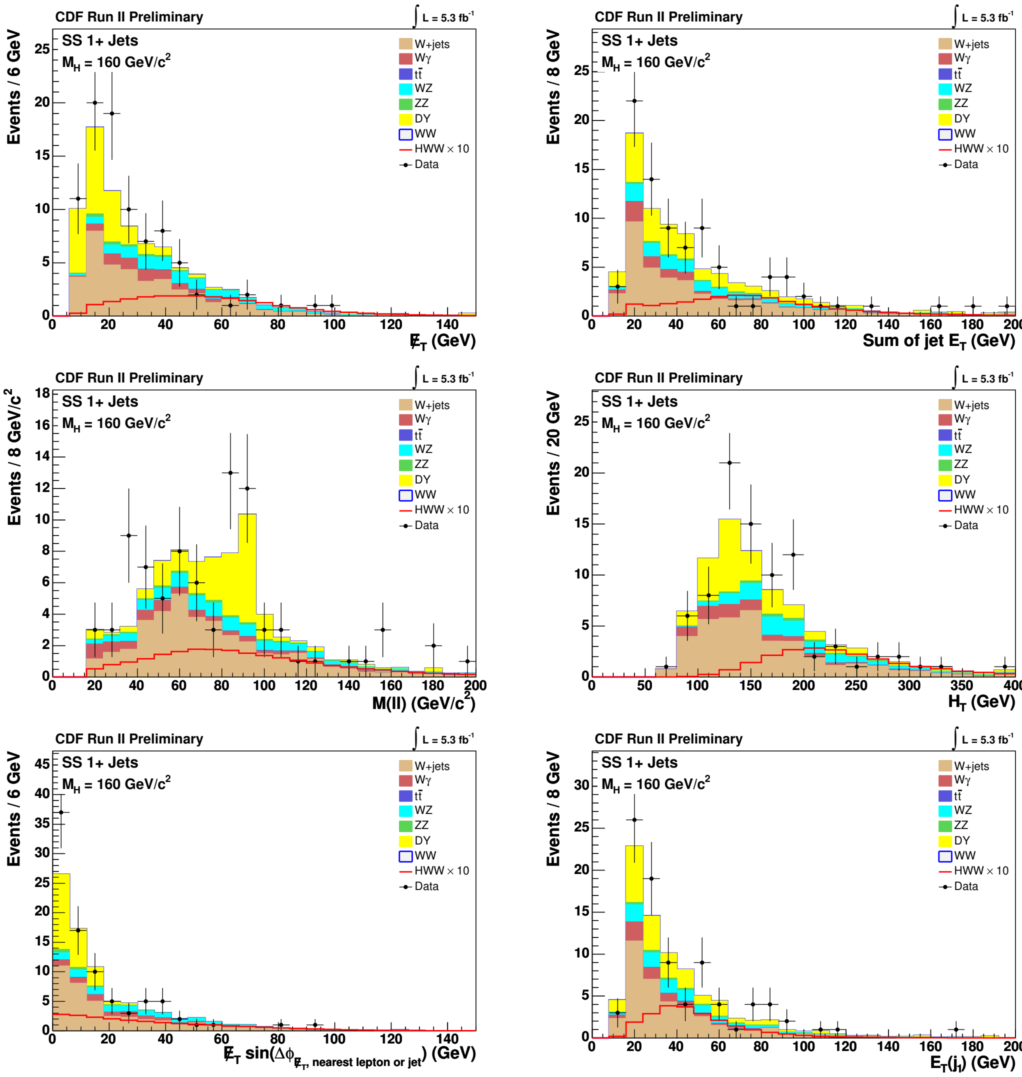

Figure 76: Neural Network input variables for SS NN (signal for $M_H = 160$ GeV/c^2).

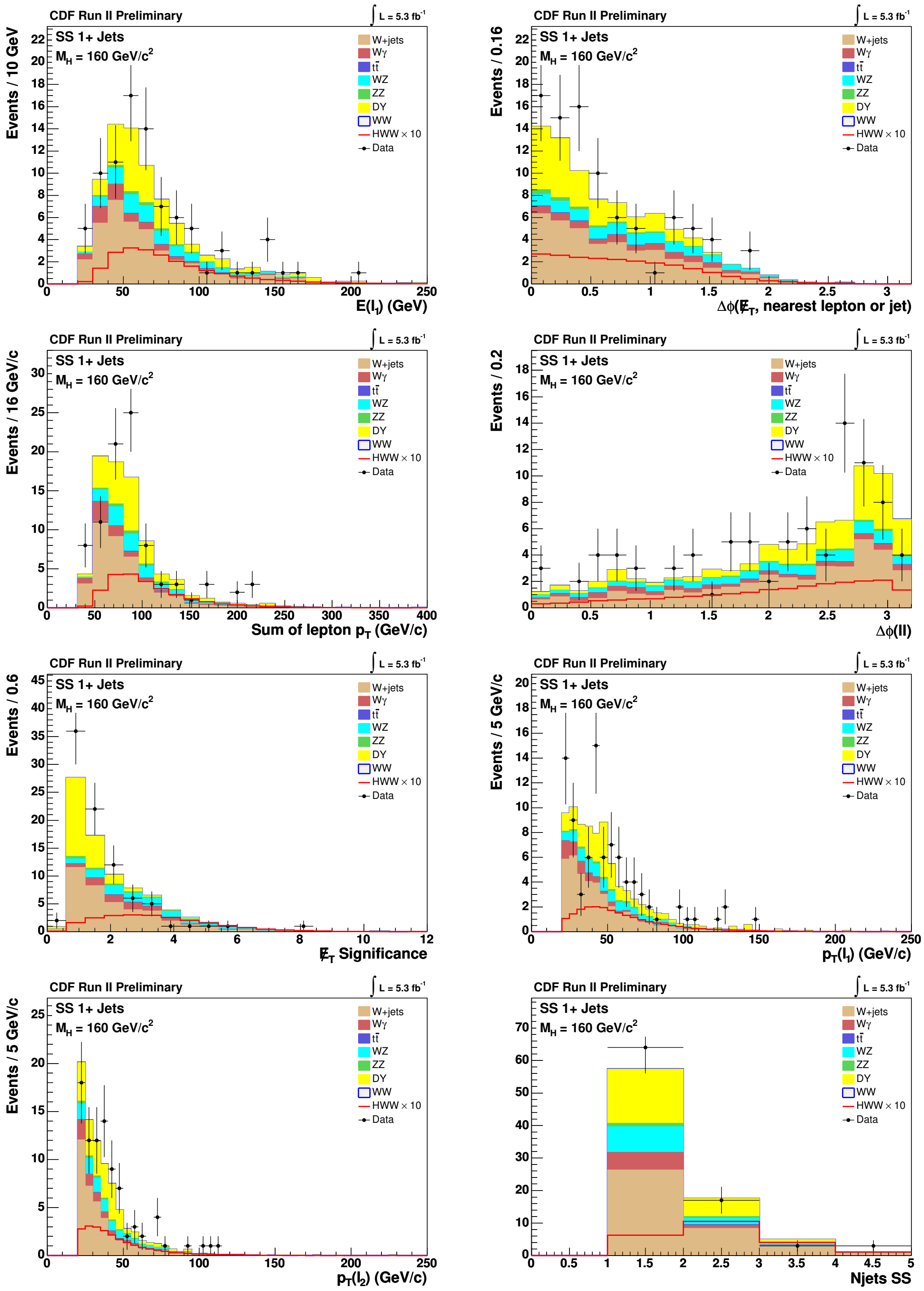

Figure 77: Neural Network input variables for SS NN (signal for M_H = 160 GeV/c^2). N_{jets} is not actually a NN input variable.

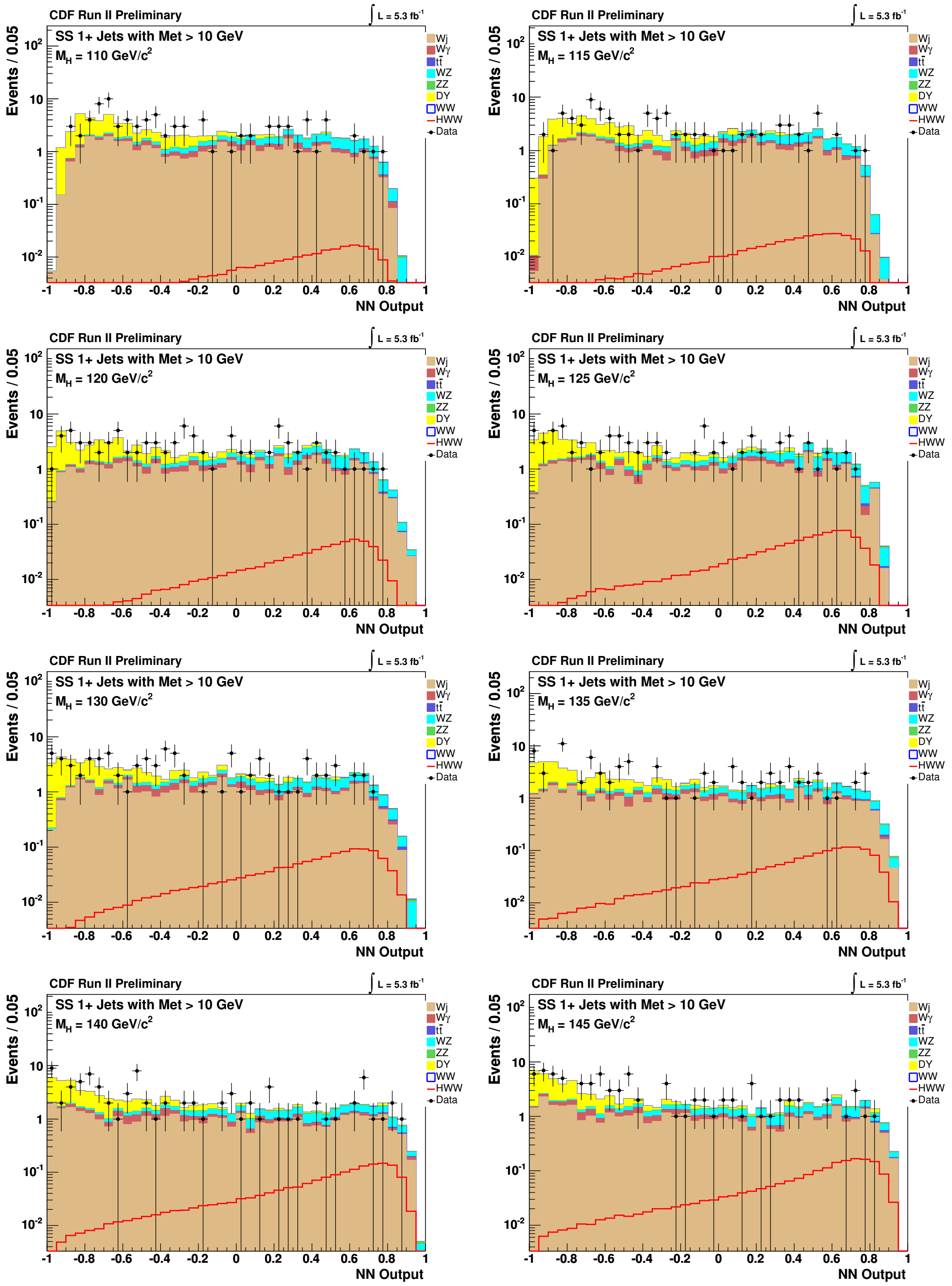

Figure 78: SS Neural Network output templates for several Higgs masses. From the top left to bottom right: 110 GeV, 115 GeV, 120 GeV, 125 GeV, 130 GeV, 135 GeV, 140 GeV, and 145 GeV.

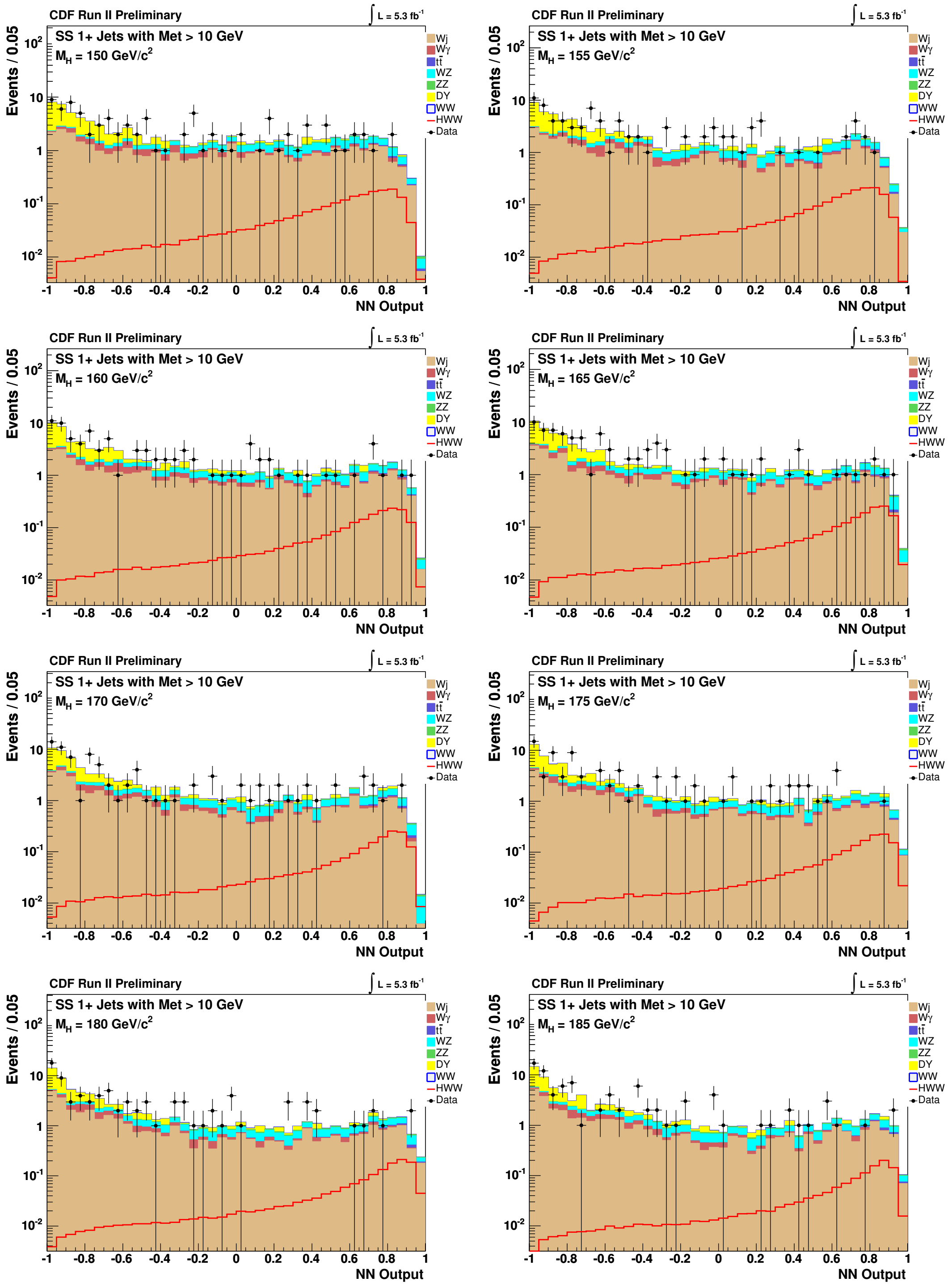

Figure 79: SS Neural Network output templates for several Higgs masses. From the top left to bottom right : 150 GeV, 155 GeV, 160 GeV, 165 GeV, 170 GeV, 175 GeV, 180 GeV, and 185 GeV.

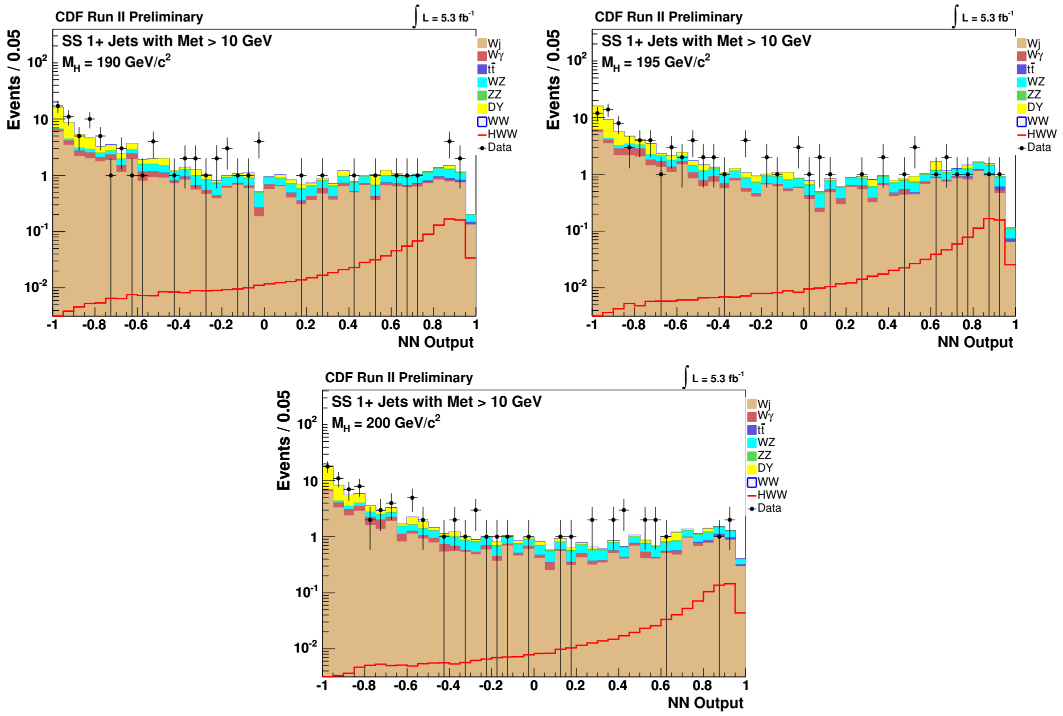

Figure 80: SS Neural Network output templates for several Higgs masses. From the top left to bottom right : 190 GeV, 195 GeV, and 200 GeV.

Mass (GeV/c^2)	gg→H	WH	ZH	VBF
110	-	0.199	0.0325	-
115	-	0.344	0.0557	-
120	-	0.546	0.0853	-
125	-	0.799	0.124	-
130	-	1.07	0.163	-
135	-	1.32	0.2	-
140	-	1.53	0.231	-
145	-	1.69	0.258	-
150	-	1.81	0.277	-
155	-	1.88	0.291	-
160	-	1.92	0.297	-
165	-	1.83	0.292	-
170	-	1.7	0.266	-
175	-	1.57	0.244	-
180	-	1.4	0.223	-
185	-	1.15	0.183	-
190	-	0.99	0.154	-
195	-	0.875	0.138	-
200	-	0.786	0.124	-

Table 73: Expected Higgs Signal yields as a function of the Higgs boson mass in the SS region for each of the contributing production processes.

shown in Table 73.

In calculating the final limits, we take into account systematic uncertainties on both the signal and background models we use to train our neural networks. Systematic uncertainties for the SS region are summarized in Tables 74 and 75.

Table 76 and Figure 81 show the expected (based on 10000 pseudo-experiments) and observed limits obtained from the SS region, incorporating the systematic uncertainties listed in Tables 74 and 75.

Uncertainty Source	WW	WZ	ZZ	$t\bar{t}$	DY	$W\gamma$	W+jet
Cross Section							
Total	*6.0%*	*6.0%*	*6.0%*	10.0%	5.0%		
Acceptance							
PDF Model	*1.9%*	*2.7%*	*2.7%*	*2.1%*	*4.1%*		
Higher-order Diagrams	*5.0%*	*10.0%*	*10.0%*	10.0%			
Jet Energy Scale	*-14.0%*	*-3.9%*	*-2.8%*	*-0.6%*	*-9.3%*	*-7.6%*	
Missing E_T Modeling					**20.0%**		
$W\gamma$ Modeling						11.0%	
Conversion Modeling						11.0%	
Charge Mismeasurement Rate	*23.6%*			*23.6%*	*23.6%*		
Jet Fake Rates							32.9%
MC Run Dependence	3.3%						
Lepton ID Efficiencies	*2.0%*	*2.0%*	*2.0%*	*2.0%*	*2.0%*		
Trigger Efficiencies	*2.1%*	*2.1%*	*2.1%*	*2.0%*	*3.4%*		
Luminosity	*5.9%*	*5.9%*	*5.9%*	*5.9%*	*5.9%*		

Table 74: Systematics table for SS region background processes. Cross section systematics are correlated for the diboson processes (shown in italic font) but uncorrelated between all other processes. The DY Met modeling systematic (shown in boldface font) is uncorrelated between channels.

Uncertainty Source	$gg \to H$	WH	ZH	VBF
Cross Section				
Scale				
PDF Model				
Total		*5.0%*	*5.0%*	
Acceptance				
PDF Model (leptons)		*1.2%*	*0.9%*	
EWK Higher-order Diagrams		*10.0%*	*10.0%*	
Jet Energy Scale		*-1.0%*	*-0.7%*	
Lepton ID Efficiencies		*2.0%*	*2.0%*	
Trigger Efficiencies		*2.1%*	*2.1%*	
Luminosity		*5.9%*	*5.9%*	

Table 75: Systematics table for SS region signal processes. Cross section systematics are uncorrelated between the signal processes. Higher-order diagrams for the VH, shown in italic font, are correlated among each other but uncorrelated from background processes.

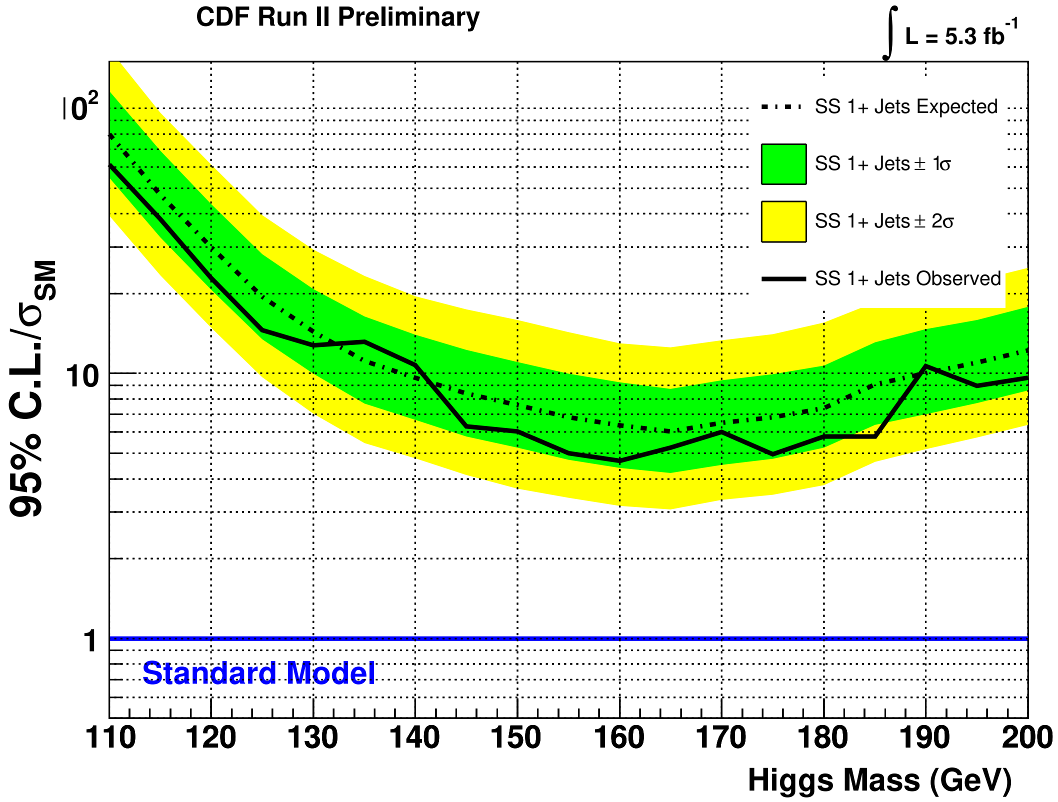

Figure 81: Expected and observed sensitivity for $H \rightarrow WW$ events with two same-sign leptons in the final state.

SS 1+ Jets	110	115	120	125	130	135	140	145	150	155
$-2\sigma/\sigma_{SM}$	39.42	23.40	14.87	9.68	7.05	5.46	4.78	4.13	3.67	3.39
$-1\sigma/\sigma_{SM}$	54.74	32.66	20.75	13.46	9.95	7.70	6.67	5.79	5.24	4.72
Median/$\sigma_{\mathbf{SM}}$	**80.08**	**47.27**	**29.94**	**19.45**	**14.36**	**11.13**	**9.64**	**8.40**	**7.60**	**6.84**
$+1\sigma/\sigma_{SM}$	116.44	69.05	43.49	28.13	20.76	16.37	13.98	12.29	11.05	9.99
$+2\sigma/\sigma_{SM}$	168.01	96.44	61.53	39.61	29.35	23.31	19.56	17.44	15.89	14.29
Observed/$\sigma_{\mathbf{SM}}$	**61.41**	**38.12**	**22.94**	**14.54**	**12.78**	**13.17**	**10.77**	**6.30**	**6.03**	**5.00**
	160	165	170	175	180	185	190	195	200	
$-2\sigma/\sigma_{SM}$	3.16	3.06	3.34	3.48	3.78	4.64	5.17	5.73	6.39	
$-1\sigma/\sigma_{SM}$	4.39	4.20	4.52	4.77	5.24	6.39	7.02	7.73	8.63	
Median/$\sigma_{\mathbf{SM}}$	**6.37**	**6.03**	**6.51**	**6.82**	**7.41**	**9.09**	**10.03**	**10.97**	**12.23**	
$+1\sigma/\sigma_{SM}$	9.26	8.73	9.41	9.91	10.68	13.11	14.70	15.93	17.82	
$+2\sigma/\sigma_{SM}$	13.01	12.52	13.35	14.03	15.54	18.59	20.58	22.38	25.07	
Observed/$\sigma_{\mathbf{SM}}$	**4.67**	**5.24**	**6.00**	**4.96**	**5.78**	**5.78**	**10.64**	**8.97**	**9.62**	

Table 76: $H \rightarrow WW$ sensitivity normalized to NNLL σ_{SM} calculation in the SS region. In this region we consider potential signal contributions only from associated production with either a W or Z boson.

12 Trilepton Signal Regions

The trilepton analysis focuses on the two associated production Higgs production modes WH and ZH. In these modes there are three vector bosons that allow for decays to more than two leptons, whereas the gluon fusion and vector boson fusion production modes do not contribute a real third lepton. Monte Carlo signal simulation does indicate that gluon fusion and vector-boson fusion have negligible contribution to the three-lepton bin. Thus, we are left with two signals to study: a $WH \rightarrow WWW \rightarrow l\nu, l\nu, l\nu$ signal and a $ZH \rightarrow ZWW \rightarrow ll, l\nu,$ jet signal. With two signals we naturally define two new trilepton signal regions attempting to isolate each, ameliorating the effort to discriminate each from background based on their unique characteristics.

Consider the three leptons as ordered by their transverse momentum p_T (or transverse energy E_T for electrons) such that the highest p_T lepton is the 1$^{\text{st}}$ and the lowest p_T lepton is the 3$^{\text{rd}}$. First, we filter trilepton events into an *InZPeak* category if any of the three possible dilepton pairings (that is, pairing the 1$^{\text{st}}$ lepton with the 2$^{\text{nd}}$ lepton; the 1$^{\text{st}}$ lepton with the 3$^{\text{rd}}$ lepton; or the 2$^{\text{nd}}$ lepton with the 3$^{\text{rd}}$ lepton) has an invariant mass value that falls within a 10 GeV window of the Z-boson mass at 91 GeV, have opposite signs, and have same flavor. This *InZPeak* region is chosen to isolate the ZH signal process. The rest of the trileptons events are directed toward the *NoZPeak* region, which focuses on the WH signal process. These regions are new to the $H \rightarrow WW$ analysis group.

Additionally, the WH analysis has a missing energy cut of $\not\!\!E_T > 20$ GeV. This cut drastically reduces the $Z\gamma$ background contribution and also provides a WH control region in $10.0GeV < \not\!\!E_T < 20.0$. Because the $WH \rightarrow WWW \rightarrow l\nu l\nu l\nu$ event topology has three $W \rightarrow l\nu$ decays, the missing energy is relatively large and a negligible amount of signal is lost from moving the $\not\!\!E_T$ cut up to 20.0 GeV from 10.0 GeV.

The $\not\!\!E_T$ distribution for the $ZH \rightarrow ZWW$ trilepton events is somewhat lower than that of the WH analysis because it produces fewer neutrinos ($WWW \rightarrow l\nu, l\nu, l\nu$ has three neutrinos while $ZWW \rightarrow ll, l\nu, jet$ has only one), so defining a control region by a higher $\not\!\!E_T$ cut is less appropriate. The ZH analysis also has somewhat larger backgrounds than the WH region and is topologically similar to the most significant background, WZ. However, for a $ZH \rightarrow ZWW$ event to produce a three-lepton signature we either have one of the W-leptons decaying hadronically or–less frequently–we have a $ZH \rightarrow ZWW \rightarrow llll$ physics event that loses one of its leptons to an area of the detector that is incapable of reconstructing a track (detector holes or too far forward in pseudorapidity, for example) but is still recorded by the calorimeter system. Therefore, ZH trilepton events inherently have a higher number of jets than the backgrounds and very little signal in the NJet= 0 bin. This characteristic of the ZH trilepton signal allows us to create a control region for the ZH analysis in the NJet= 0 bin with very little signal loss, and so NJet= 0 events are not included in the ZH analysis.

Observe in table 12.1 that $\sim 77\%$ of the signal in the *NoZPeak* region is WH, while $\sim 96\%$ of the signal in the *InZPeak* region is ZH. We will see in section 12.3 how this division allows us to focus on the unique characteristic of each signal for discrimination from the background in the NeuroBayes neural net treatment.

12.1 Trilepton Backgrounds

Both regions of this trilepton analysis have five background categories considered: WZ, ZZ, $Z\gamma$ (replacing Drell-Yan), Fakes (data-based WW and Z+jets), and $t\bar{t}$. Each is summarized in table 12.1 along with the predicted signal for a $m_H = 160$ GeV standard model Higgs boson and the data.

CDF Run II Preliminary $\int \mathcal{L} = 5.3$ fb^{-1}

($m_H = 165$ GeV/c^2)	**WH Signal Region**			**ZH Signal Region**		
WZ	7.01	±	0.96_{syst}	9.01	±	1.74_{syst}
ZZ	1.49	±	0.20_{syst}	4.41	±	0.68_{syst}
$Z\gamma$	2.47	±	0.42_{syst}	3.00	±	0.63_{syst}
Fakes (WW,Z+Jets)	3.22	±	0.97_{syst}	7.74	±	2.32_{syst}
$t\bar{t}$	0.18	±	0.07_{syst}	0.03	±	0.01_{syst}
Total Background	14.5	±	1.58_{syst}	24.3	±	3.57_{syst}
WH	0.58	±	0.08_{syst}	0.02	±	0.004_{syst}
ZH	0.18	±	0.02_{syst}	0.58	±	0.08_{syst}
Total Signal	0.76	±	0.10_{syst}	0.60	±	0.08_{syst}
Data		14			33	

Trilepton $H \to WW$ Search

Heavy Dibosons: WZ, ZZ

The WZ and ZZ diboson contributions provide three real leptons, with WZ being the dominant background in both trilepton signal regions. Both samples are Pythia-based, where the W is allowed to decay inclusively and the Z is forced to decay leptonically (electron, muon, or tau pairs). The WW contribution, which can only result in two real leptons, is contained within the data-derived estimate of fakes events, as described below, so we do not use the WW Monte Carlo sample to estimate trilepton backgrounds.

$Z\gamma$

The $Z\gamma$ background in the trilepton analyses replaces the Drell Yan contribution of the dilepton analyses and is created by the Bauer generator. We acquire a third lepton from a Drell Yan process when either an initial or final state radiated photon undergoes a conversion and showers in the calorimeter for the third lepton. As such, the $Z\gamma$ is the restriction of Drell Yan to those events which do radiate a photon for the purpose of working with a larger statistical sample. The estimated $Z\gamma$ background is weighted by 0.83 based on corrections to the photon conversion efficiency, as described in Section 10.4.

Fakes(WW, Z+Jets)

In the dilepton analysis, the Fakes category is measured from single high p_T lepton data (rather than MC) and assumed to have a W+jets event topology, where the one lepton is from the W-boson. From this data sample, events with one-lepton + denominator object are selected and then re-weighted based on the rates at which jets fake a lepton shown in Section 4.3.

Similarly, for the trilepton analyses we are interested in processes that produce two physical leptons + one denominator object from the jets. Two high p_T lepton data is dominated by WW and Z+jets. First note that we do not consider simulated WW background as the dilepton analyses do to avoid double counting the process. Second, because the rate at which a jet is expected to fake a lepton is on the order of $1-5\%$, the rate at which such an event is expected to fake two leptons is drastically lower: $0.01 - 0.25\%$. As such, we consider the contribution of W+jets with one real lepton and two fake leptons to be negligible for the trilepton analyses and so label this category WW,Z+Jets instead of W+Jets.

$t\bar{t}$

The $t\bar{t}$ process is the smallest background, but arguably the most complex. This process decays to two pairs of a b-jet accompanied by a W boson. For the case of trileptons, we consider the case of the two W's decaying leptonically. The third lepton signature is then due to one of the b-jets. As b-quarks may hadronize into B-hadrons which decay leptonically, b-quark jets have a higher probability of containing real, possibly isolated leptons than light quark jets. However, the rate at which a b-quark jet produces an isolated lepton is not well measured experimentally. Because the third lepton is contained in a jet, there will be a contribution from $t\bar{t}$ included in the data-derived fakes background. However, this contribution is scaled by a fake rate derived from a sample of primarily light quark jets, and thus the $t\bar{t}$ contribution to the fakes background is scaled down further than it should be.

To properly account for the $t\bar{t}$ contribution to the trilepton signature, we use the standard $t\bar{t}$ MC sample in a slightly modified way. To avoid double-counting with fakes contributions, leptons reconstructed in the MC samples are required to match to a generator-level lepton or photon (to account for the case of a lepton from photon conversion). For the $t\bar{t}$ in the trilepton analysis, we are also interested in a third lepton whose signature is the result of b-quark hadronization. Thus, the MC generator-level matching is modified to allow leptons matching to a generator-level b-quark. Including leptons from the b-quarks increases the expected $t\bar{t}$ contribution to trilepton events by 46%. Because the overall $t\bar{t}$ contribution is small, this has little effect on the neural network training or shape of the output templates. However, $t\bar{t}$ does give some contribution at high neural network score, so increasing the expected $t\bar{t}$ contribution does have a small (2-3%) effect on the expected sensitivity of the trilepton searches.

There is unavoidably some overlap between the $t\bar{t}$ that in the data-derived fakes background and the modified MC sample which allows matching to b-quarks. We expect the overlap to be very small, due to the tiny fake rates applied to denominator objects in the fakes sample, but the overlap is difficult to account for since we do not know the composition of the data-derived fakes sample. Instead of attempting to subtract the $t\bar{t}$ contribution from the fakes sample, we apply a systematic uncertainty on the $t\bar{t}$ estimate equal to half of the increase (23%) when we allow matching to b-quarks.

12.2 Trilepton Control Regions

The modeling of basic kinematic properties and the discriminating variables in the Monte Carlo simulation is tested by comparing the distributions of these variables in orthogonal "control regions" which are enhanced in specific major backgrounds and contain minimum possible signal contribution. The control regions we choose for the WH and ZH trilepton analyses are:

- WH Analysis Control Region: Same selection as the WH signal region, except requiring $10.0 < \not{E}_T < 20.0$ GeV
- ZH Analysis Control Region: Same selection as the ZH signal region, except requiring the number of jets= 0

Both conrol regions contain minimal signal (see table 12.2) so cutting them out of the analysis regions reduces the background to discriminate against without significantly reducing the signal, while also providing a verification of background modelling.

CDF Run II Preliminary $\int \mathcal{L} = 5.3$ fb^{-1}

($m_H = 165$ GeV/c^2)	**WH Control Region**			**ZH Control Region**		
WZ	0.77	±	0.11_{syst}	32.0	±	6.19_{syst}
ZZ	0.72	±	0.10_{syst}	3.55	±	0.55_{syst}
$Z\gamma$	19.4	±	3.31_{syst}	5.56	±	1.17_{syst}
Fakes (WW,Z+Jets)	7.58	±	2.27_{syst}	9.43	±	2.83_{syst}
$t\bar{t}$	0.01	±	0.002_{syst}	-	±	-
Total Background	28.4	±	4.02_{syst}	50.5	±	7.81_{syst}
WH	0.025	±	0.003_{syst}	0.06	±	0.01_{syst}
ZH	0.014	±	0.002_{syst}	0.06	±	0.01_{syst}
Total Signal	0.038	±	0.005_{syst}	0.12	±	0.02_{syst}
Data		31			49	

Trilepton $H \rightarrow WW$ Search

WH Control Region

The topology of WH associated production in the trilepton channel also contains at least three neutrinos (more if $W \to \tau\nu_\tau \to l_{e,\mu}\nu_{e,\mu}\nu_\tau$ decays are involved), resulting in high missing energy values. The low $\not\!\!E_T$ region is a natural choice for a control region in the WH analysis since it contains negligible signal contribution and is enriched in $Z\gamma$ and fakes backgrounds. Also, using a $\not\!\!E_T > 20$ cut for the WH signal region substantially enhances the signal-to-background ratio in the final signal region. Plots of the kinematic variables used in the WH neural network (described below in Section 12.3) for this control region are shown in Figure 82. The agreement between the expected background predictions and the observed distributions in data is excellent.

ZH Control Region

The topology of ZH associated production lends to a preference for at least one or two jets since one of the two Higgs-W-bosons decays hadronically. Only $\sim 10\%$ of the trilepton ZH signal is present in the NJet$= 0$ bin, but much of its dominant background, WZ, is there. Thus, the NJet$= 0$ bin is a natural choice for the control region of the ZH trilepton analysis. Plots of the kinematic variables used in the ZH neural network (described below in Section 12.3) for this control region are shown in Figure 83. However, there are several inputs to the neural network which rely on the properties of jets in the event; these are always defined for the signal region, but not in the control region which uses only events with zero jets. Therefore, plots are not shown for the following three variables in the control region: N_{jets}, E_T of the leading jet, and ΔR between the W-lepton and the leading jet. The agreement between the expected background predictions and the observed distributions in data is excellent.

12.3 Trilepton Neural Network Training

The trilepton $H \to WW$ analyses rely on the NeuroBayes neural network package to discriminate signal from background, as used by the dilepton analyses. Because the interaction topology under consideration involves three leptons and also because we do not separate the analyses by jet bin (aside from reserving the NJet=0 bin for the *InZPeak* control region of the ZH analysis), the signatures of the signal regions under consideration involve many potentially complex variables whose discriminatory power must be explored. As such, a larger than usual quantity of discriminating variables are used to train the NeuroBayes neural nets; we use 13 input variables for the WH analysis and 16 for the ZH analysis. These variables are listed below (given in order of their importance to the neural network for a Higgs boson mass of 165 GeV), and shown in Figures 84 through 86. The neural network output templates are shown in Figures 87 through 92.

WH Variable Descriptions/Details:

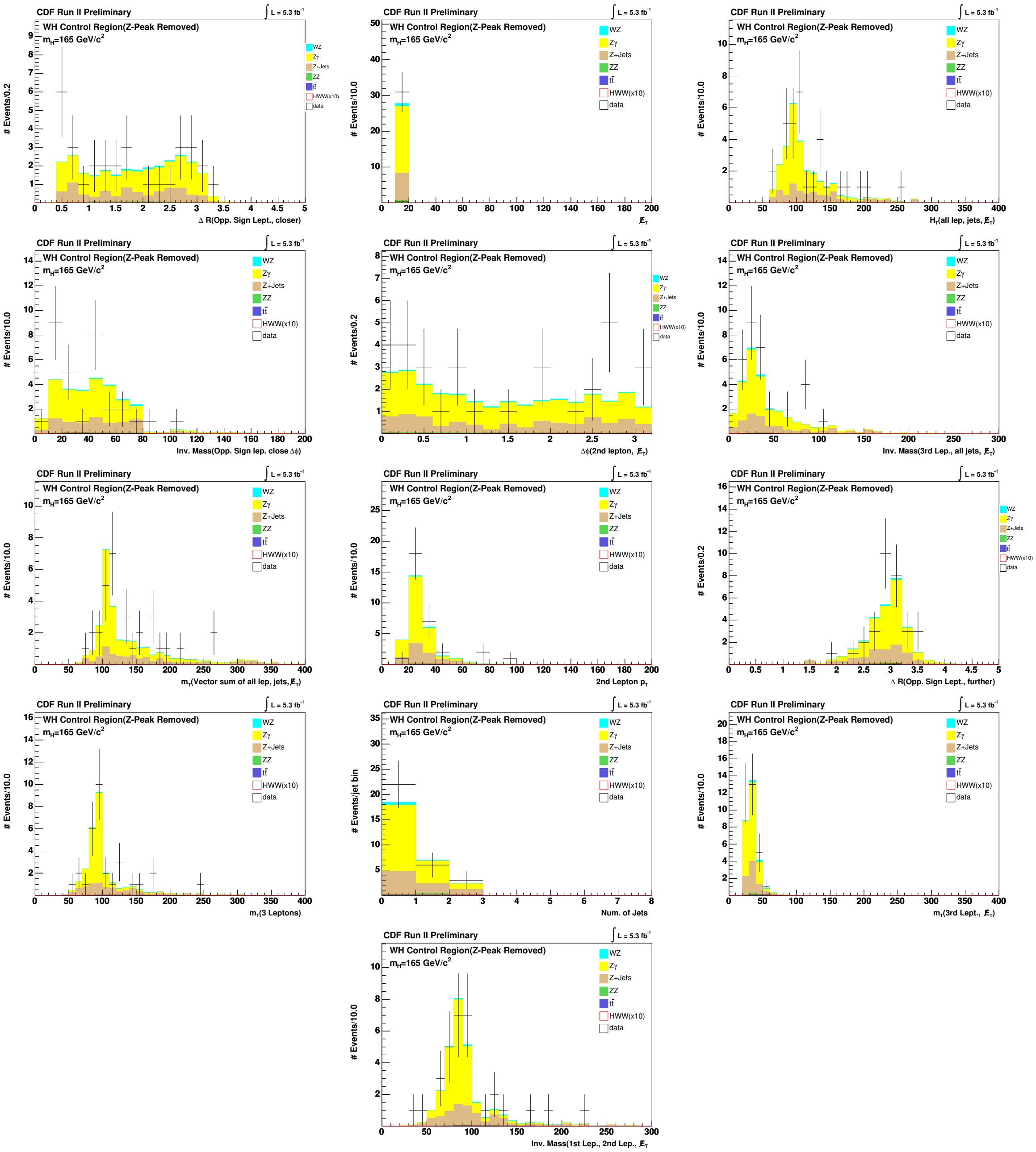

Figure 82: Kinematic variables for Trilepton WH control region (signal for M_H = 160 GeV/c^2).

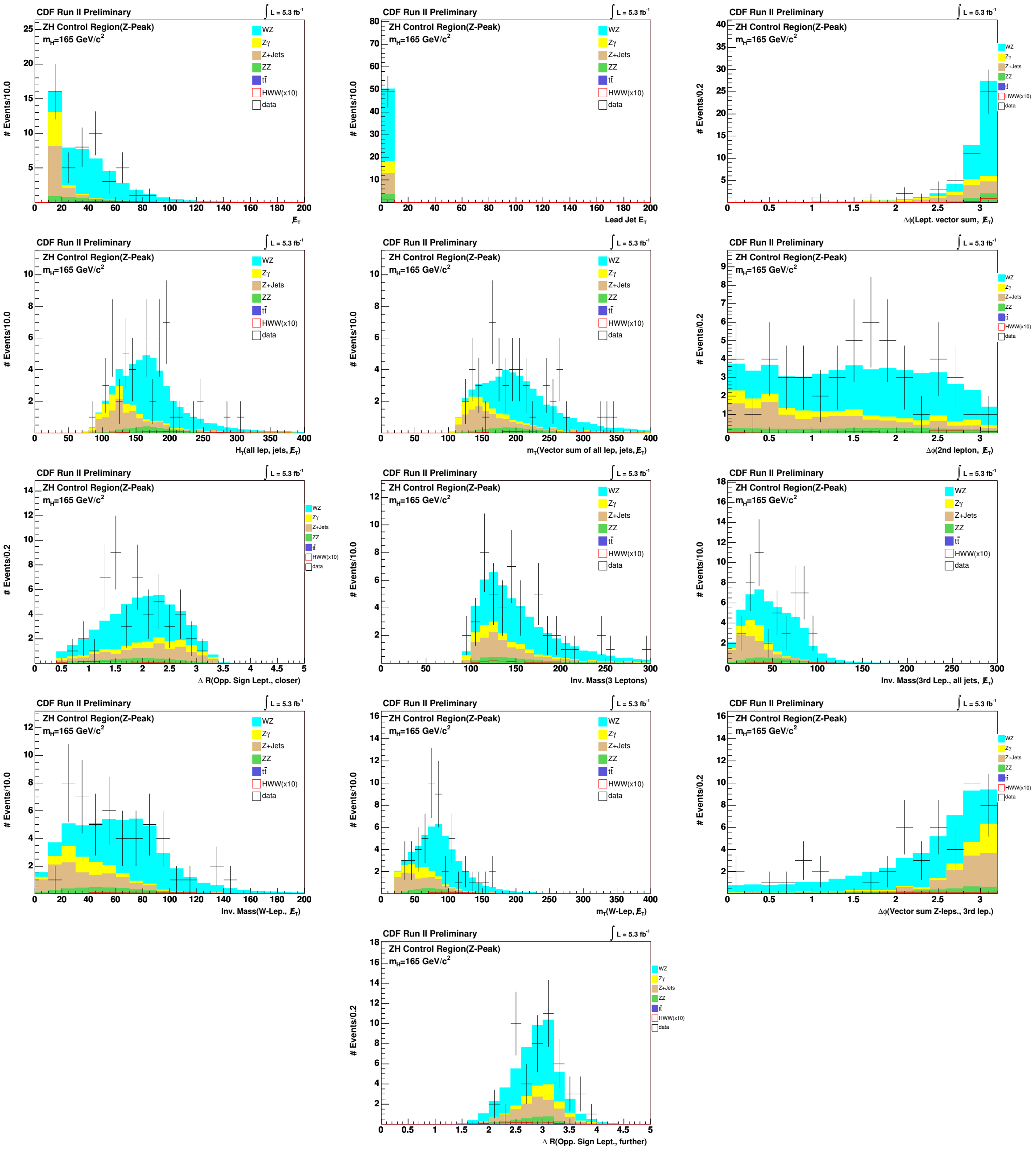

Figure 83: Kinematic variables for Trilepton ZH control region (signal for $M_H = 165$ GeV/c^2).

- ΔR b/w Opp. Sign Close Leptons: With three leptons there are three possible pairings of leptons. Events with all three leptons having the same sign are rejected from this analysis, so every event has two possible pairings of opposite-signed leptons. Of those two pairings, this variable is the ΔR value of the pairing with lower ΔR value.

- $\not\!\!E_T$: Missing Transverse Energy

- H_T: Sum of the transverse energies of all three leptons, the $\not\!\!E_T$, and all jets.

- Dimass b/w Opp. Sign Close by ϕ: Dilepton invariant mass of the opposite-signed lepton pair that is closer in the ϕ coordinate.

- $\Delta\phi$(Lep2,$\not\!\!E_T$): The magnitude of the difference in ϕ between the 2$^{\text{nd}}$ lepton by p_T and the $\not\!\!E_T$.

- Inv. Mass(Lep3,$\not\!\!E_T$,Jets): Invariant mass of the vector sum of the 3$^{\text{rd}}$ lepton, $\not\!\!E_T$, and Jets.

- m_T(Leptons,$\not\!\!E_T$,Jets): Transverse mass of the vector sum of all three leptons, $\not\!\!E_T$, and all jets.

- p_T of the 2$^{\text{nd}}$ lepton by p_T.

- ΔR Opp. Sign Far Leptons: With three leptons there are three possible pairings of leptons. Events with all three leptons having the same sign are rejected from this analysis, so every event has two possible pairings of opposite-signed leptons. Of those two pairings, this variable is the ΔR value of the pairing with higher ΔR value.

- m_T Trilepton Mass: Transverse mass of the vector sum of the three leptons.

- NJet: The number of jets in the event. For this use of NJet, all events with NJet $\geq$ 2 jets are thrown into the NJet = 2 bin.

- m_T (Lep3, $\not\!\!E_T$): Transverse mass of the vector sum of the 3$^{\text{rd}}$ lepton and the $\not\!\!E_T$.

- Inv. Mass(Lep1,Lep2,$\not\!\!E_T$): Invariant mass of the vector sum of the 1$^{\text{st}}$ lepton, 2$^{\text{nd}}$ lepton, and $\not\!\!E_T$.

ZH Variable Descriptions/Details:

- NJet: The number of jets in the event.

- $\not\!\!E_T$: Missing Transverse Energy

- Lead Jet E_T: Transverse energy of the leading jet. Note that the control region for *InZPeak* is NJet $= 0$, so all events in the signal region must have at least one jet by definition. Also, for this use of NJet, all events with NJet ≥ 2 jets are thrown into the NJet $= 2$ bin.

- ΔR(W-Lep, Lead Jet): The *InZPeak* region is defined by having one lepton paring (opposite signed, same flavor) near the Z boson mass. Denote the one other lepton not in this pairing as the W-lepton. This variable is then the ΔR between the W-lepton and the leading jet.

- $\Delta\phi$(Leptons,$\not{E}_T$): $\Delta\phi$ between the vector sum of the three leptons and the $\not{E}_T$.

- H_T(Leptons,$\not{E}_T$,Jets): Sum of E_T of all three leptons, $\not{E}_T$, and all jets.

- m_T(Leptons,$\not{E}_T$,Jets): Transverse mass of the vector sum of all three leptons, $\not{E}_T$, and all jets.

- $\Delta\phi$(Lep2,$\not{E}_T$): The magnitude of the difference in ϕ between the 2$^{\text{nd}}$ lepton by p_T and the $\not{E}_T$.

- ΔR b/w Opp. Sign Close Leptons: With three leptons there are three possible pairings of leptons. Events with all three leptons having the same sign are rejected from this analysis, so every event has two possible pairings of opposite-signed leptons. Of those two pairings, this variable is the ΔR value of the pairing with lower ΔR value.

- Trimass:The invariant mass of the vector sum of the three leptons.

- Inv. Mass(Lep3,$\not{E}_T$,Jets): Invariant mass of the vector sum of the 3$^{\text{rd}}$ lepton, $\not{E}_T$, and Jets.

- Dimass(W-Lep,$\not{E}_T$): The *InZPeak* region is defined by having one lepton paring (opposite signed, same flavor) near the Z boson mass. Denote the one other lepton not in this pairing as the W-lepton. This variable is then the invariant mass of the vector sum of the W-lepton and the $\not{E}_T$.

- m_T Jets: Transverse mass of the vector sum of all jets. Note that the control region for *InZPeak* is NJet $= 0$, so all events in the signal region must have at least one jet by definition.

- m_T(W-Lep,$\not{E}_T$): Transverse mass of the vector sum of the W-lepton and the $\not{E}_T$.

- $\Delta\phi$(Z-Leptons,W-Lepton): ΔR between the vector sum of the two leptons whose dimass is near the Z-boson mass, and the other lepton.

- ΔR Opp. Sign Far Leptons: With three leptons there are three possible pairings of leptons. Events with all three leptons having the same sign are rejected from this analysis, so every event has two possible pairings of opposite-signed leptons. Of those two pairings, this variable is the ΔR value of the pairing with higher ΔR value.

Trilepton Control Regions through Neural Networks

An additional cross-check of the background modeling and neural network training is to run the control region events through a neural network trained on the signal region, and comparing the expected background distribution to that observed in data. This is done using the neural network trained for a Higgs mass of 165 GeV. For the ZH control region, the three input variables related to jets (N_{jets}, E_T of the leading jet, and ΔR between the W-lepton and the leading jet) are removed from the neural network. Additionally, due to limited Monte Carlo statistics, there are no $t\bar{t}$ events enter the ZH control region. Thus, to obtain a neural network result for this ZH control region we had to retrain a neural network on the signal region (NJet ≥ 1) excluding both the three aforementioned discriminating variables and the $t\bar{t}$ background. The neural network output distributions for the two control regions are shown in Figure 93, and there is excellent agreement between the expected and observed distributions.

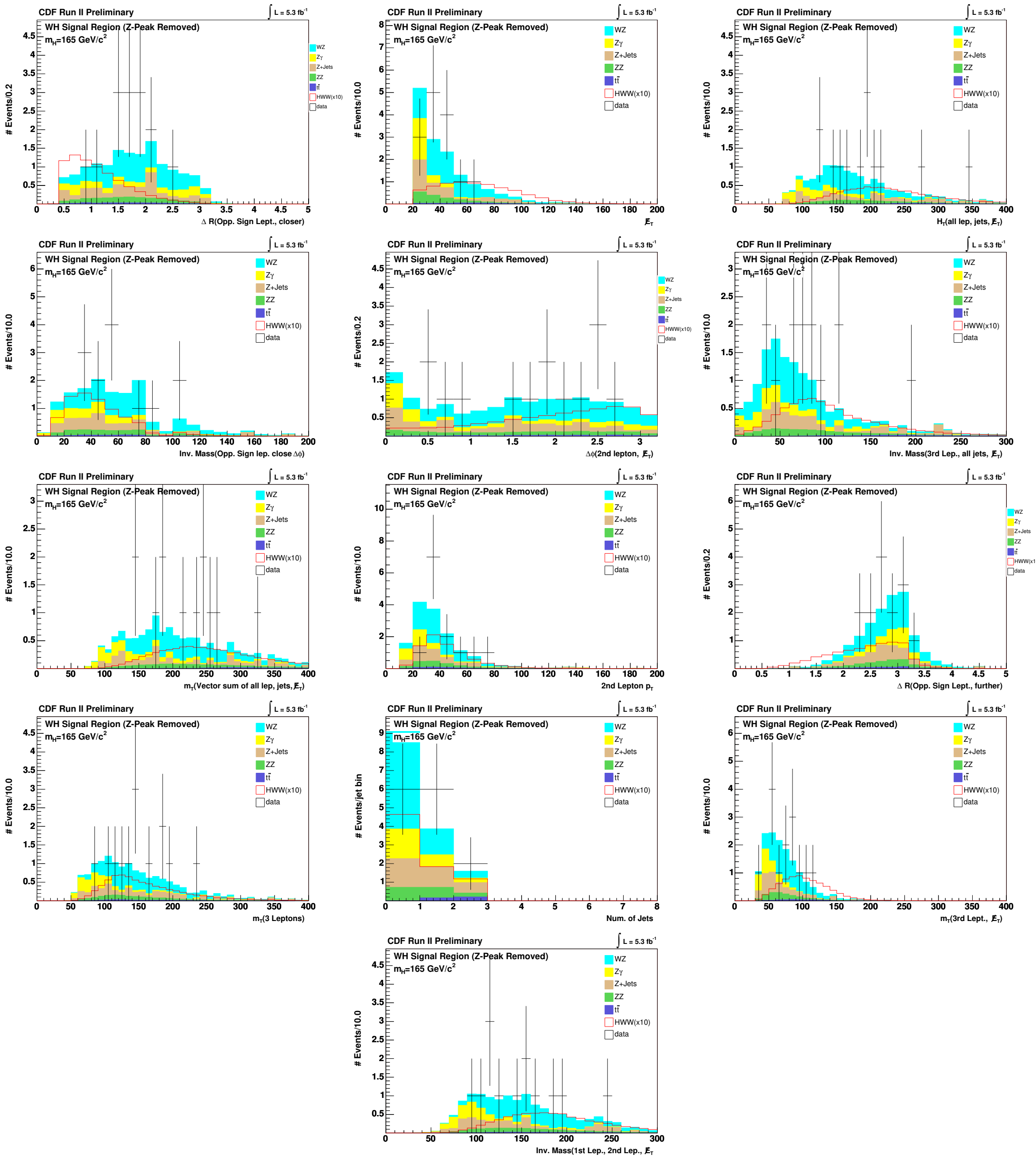

Figure 84: Neural Network input variables for Trilepton WH NN (signal for $M_H = 165$ GeV/c^2).

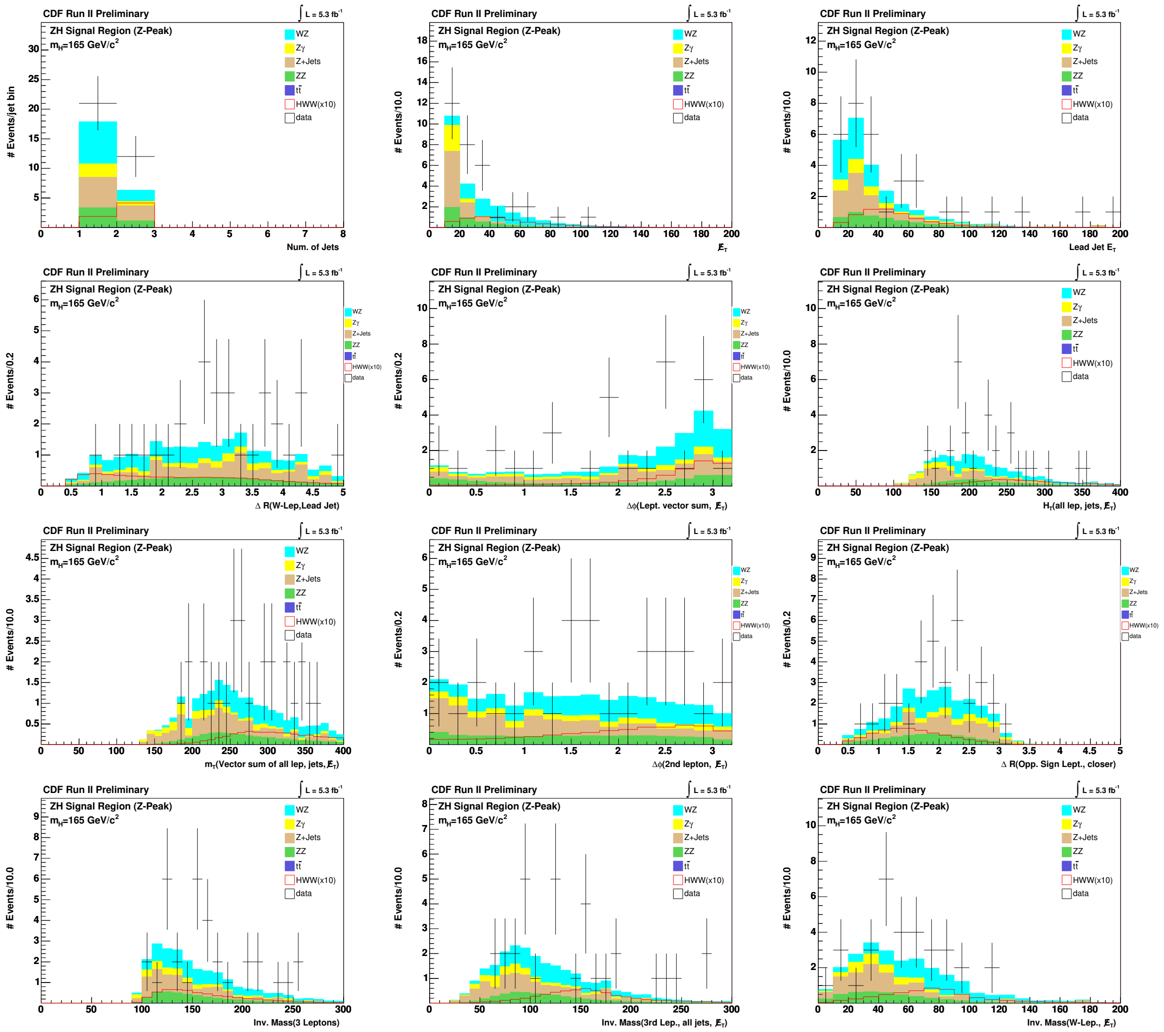

Figure 85: First 12 Neural Network input variables for Trilepton ZH NN (signal for $M_H = 165$ GeV$/c^2$).

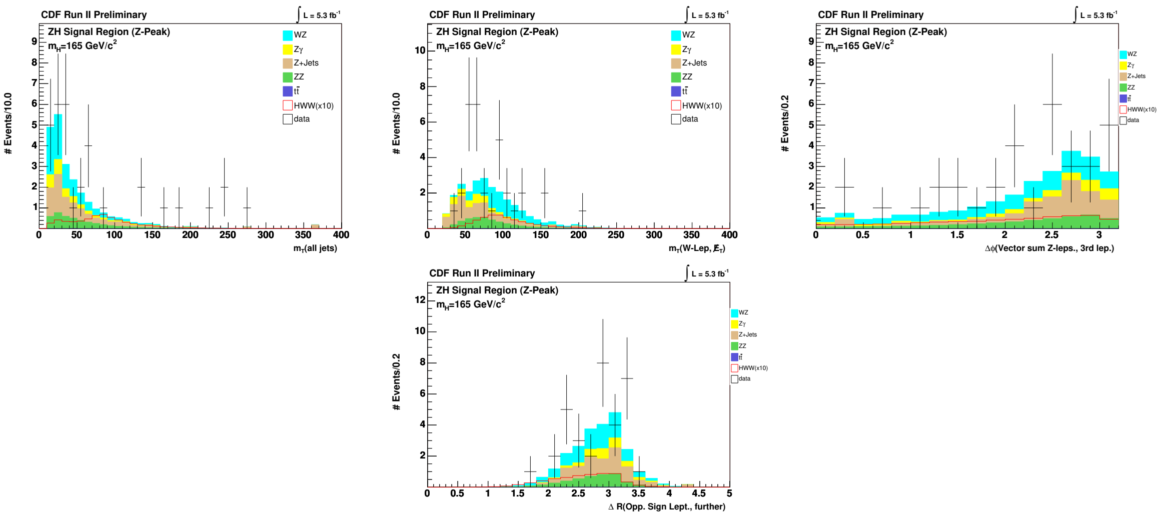

Figure 86: Remaining 4 Neural Network input variables for Trilepton ZH NN (signal for $M_H = 165$ GeV/c^2).

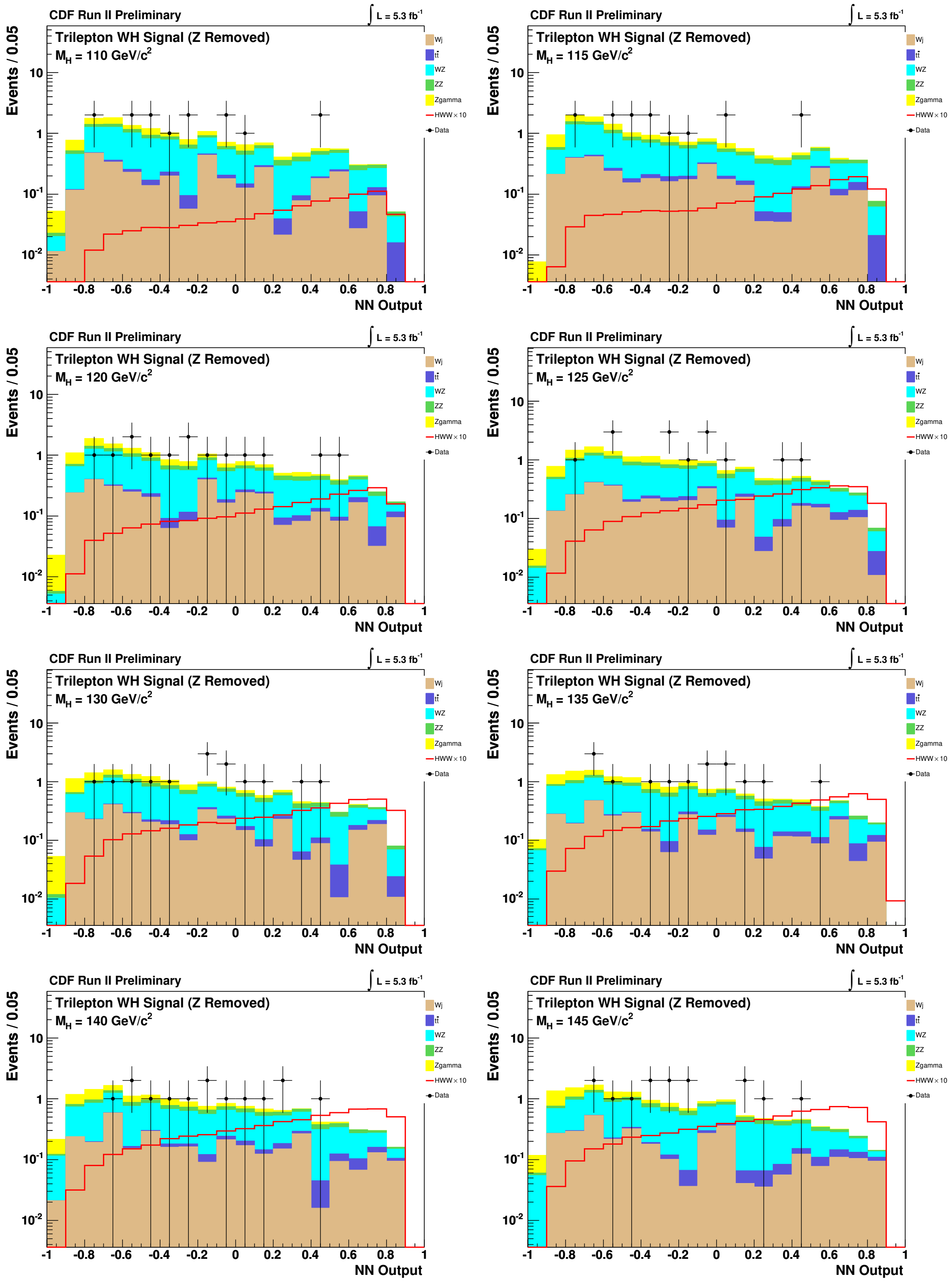

Figure 87: Trilepton WH Neural Network output templates for several Higgs masses. From the top left to bottom right: 110 GeV, 115 GeV, 120 GeV, 125 GeV, 130 GeV, 135 GeV, 140 GeV, and 145 GeV.

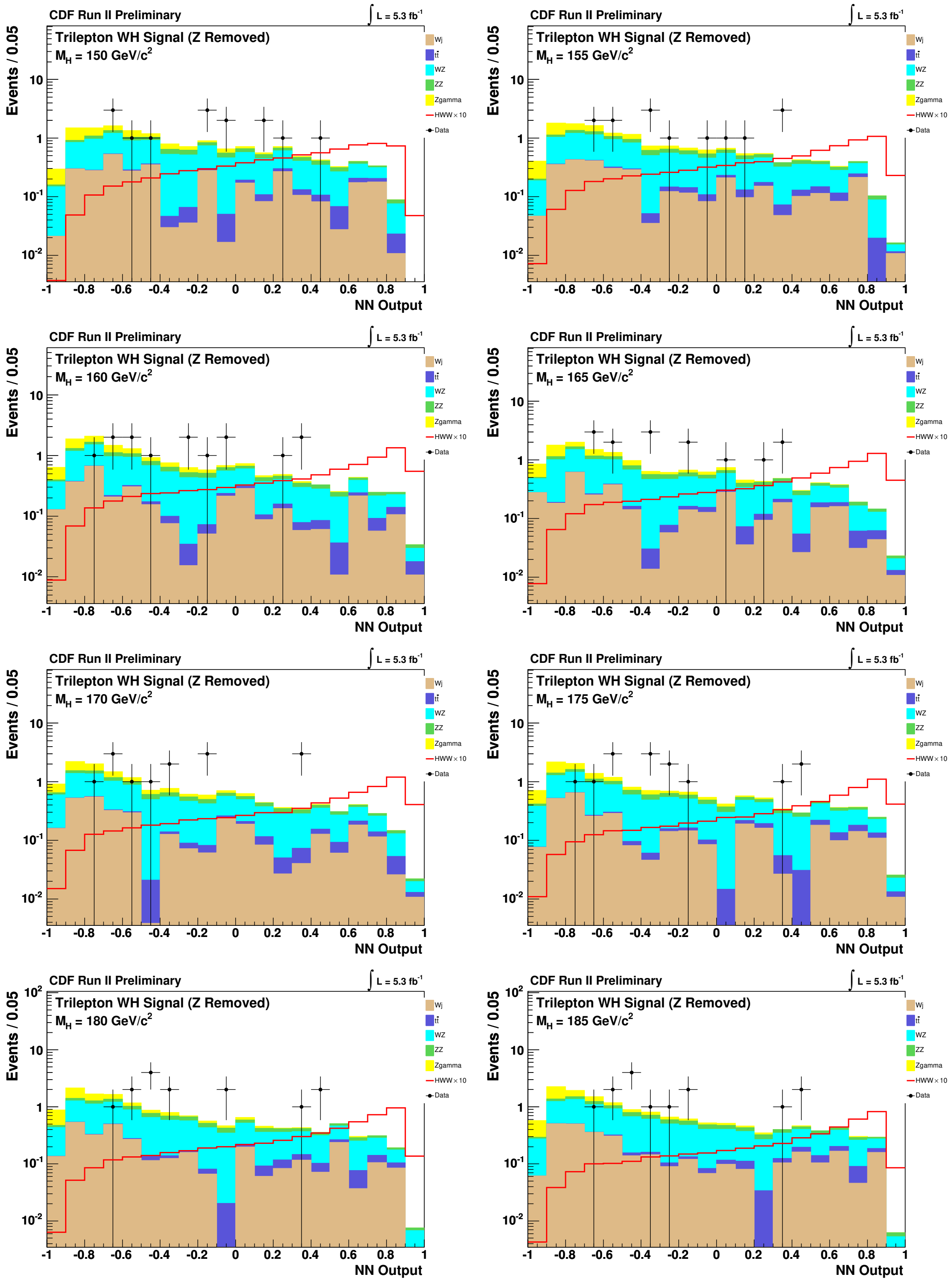

Figure 88: Trilepton WH Neural Network output templates for several Higgs masses. From the top left to bottom right : 150 GeV, 155 GeV, 160 GeV, 165 GeV, 170 GeV, 175 GeV, 180 GeV, and 185 GeV.

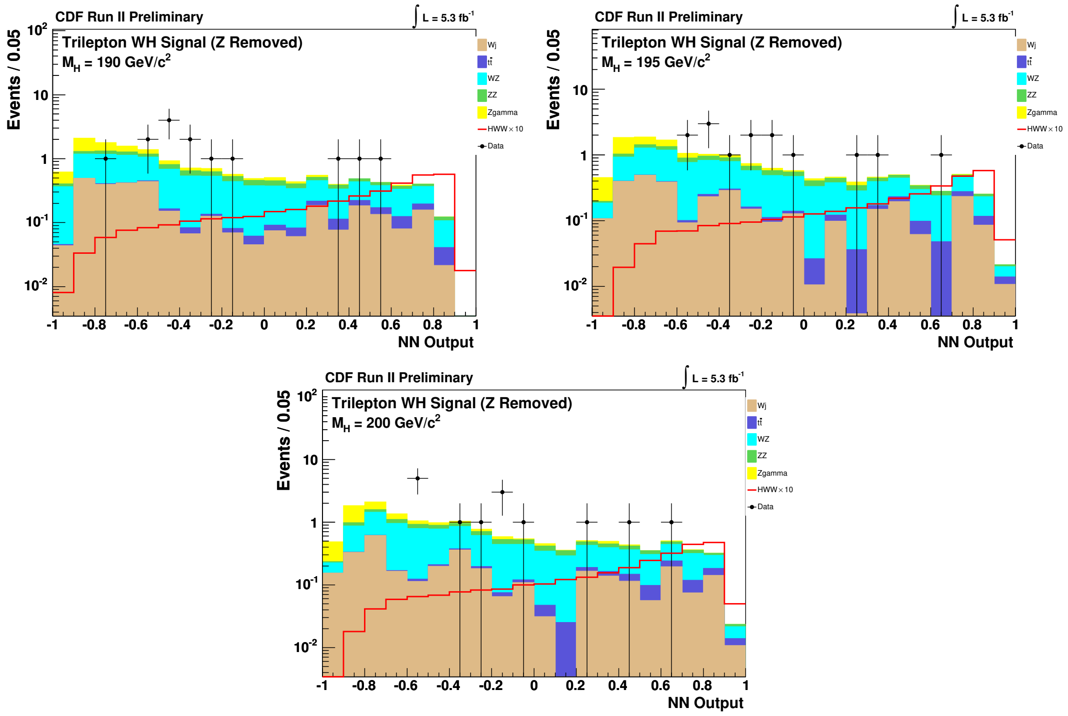

Figure 89: Trilepton WH Neural Network output templates for several Higgs masses. From the top left to bottom right : 190 GeV, 195 GeV, and 200 GeV.

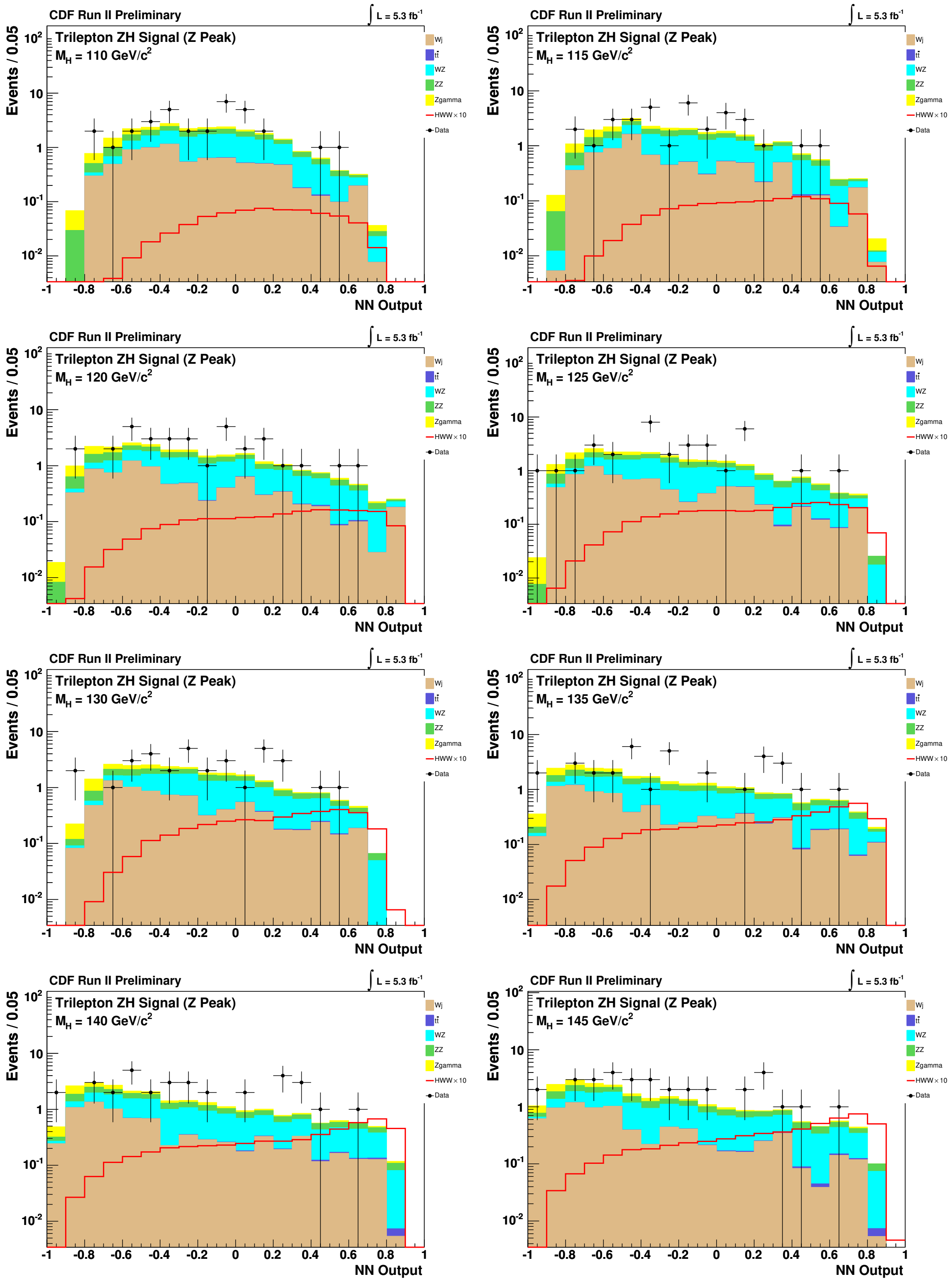

Figure 90: Trilepton ZH Neural Network output templates for several Higgs masses. From the top left to bottom right: 110 GeV, 115 GeV, 120 GeV, 125 GeV, 130 GeV, 135 GeV, 140 GeV, and 145 GeV.

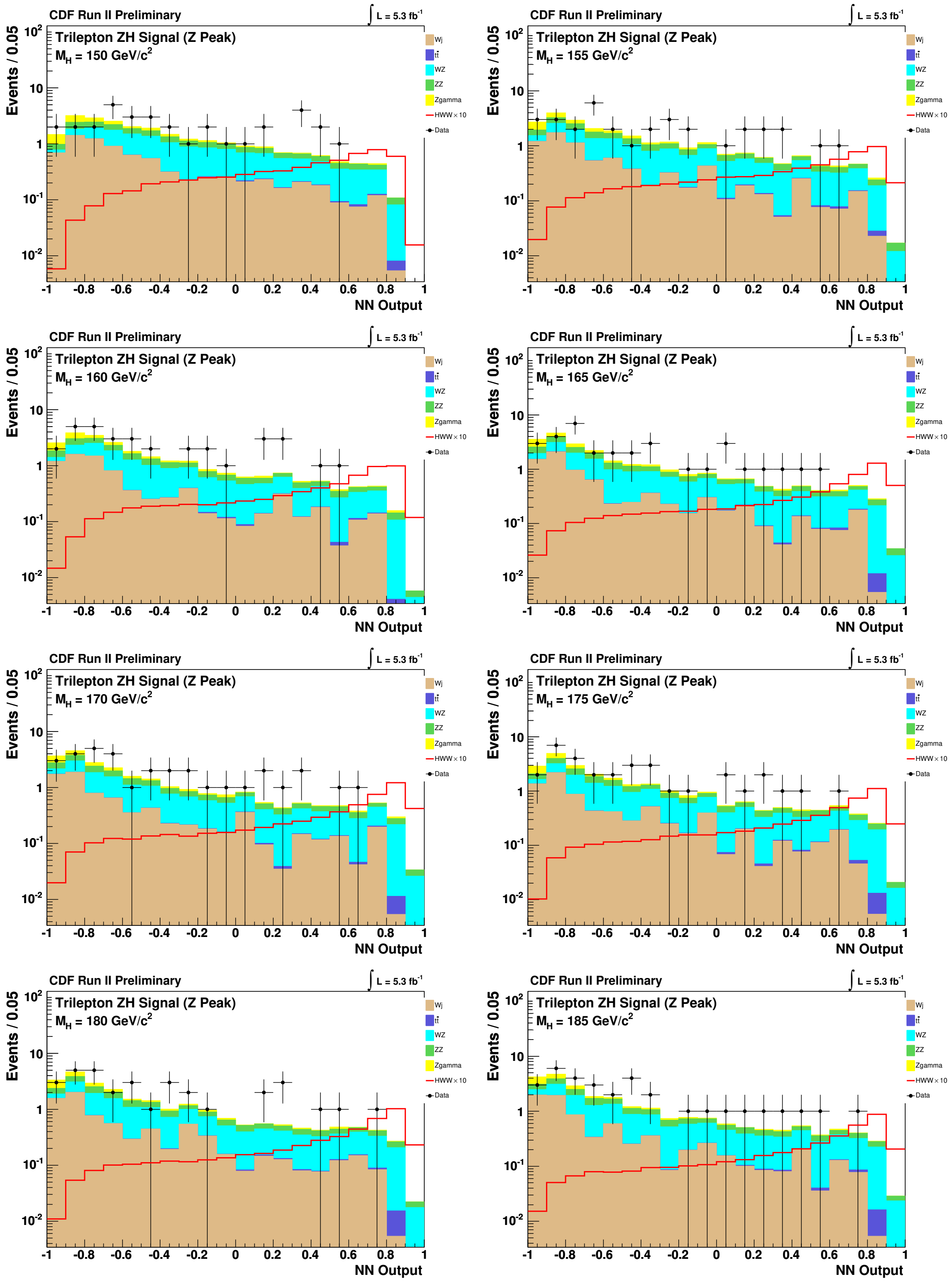

Figure 91: Trilepton ZH Neural Network output templates for several Higgs masses. From the top left to bottom right : 150 GeV, 155 GeV, 160 GeV, 165 GeV, 170 GeV, 175 GeV, 180 GeV, and 185 GeV.

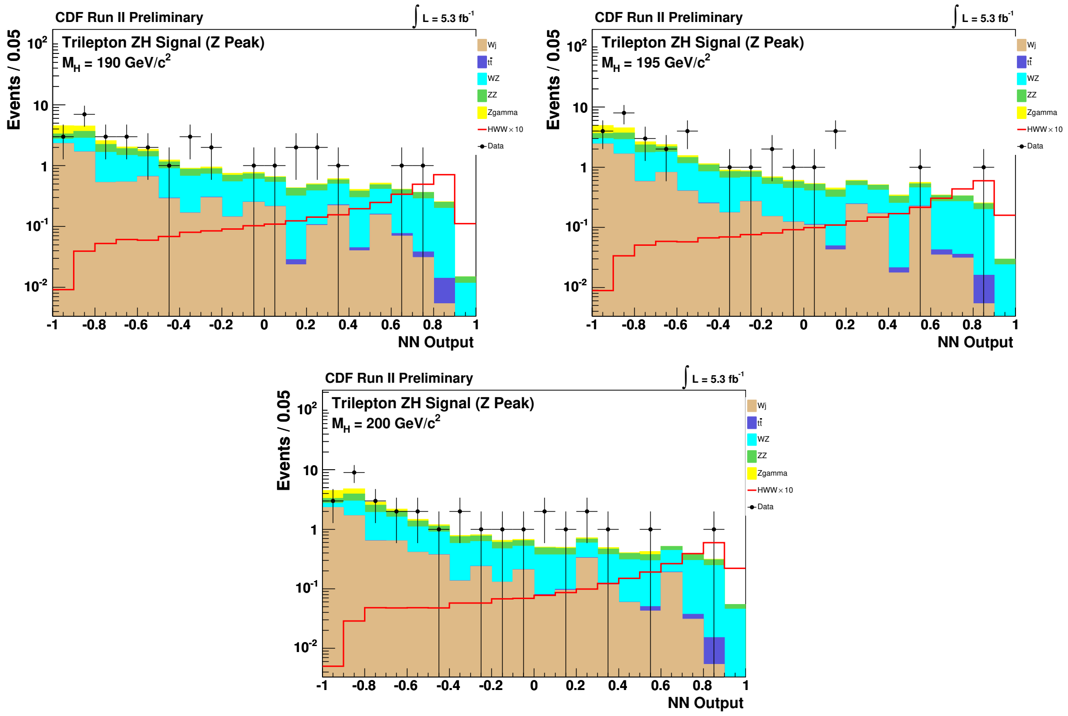

Figure 92: Trilepton ZH Neural Network output templates for several Higgs masses. From the top left to bottom right : 190 GeV, 195 GeV, and 200 GeV.

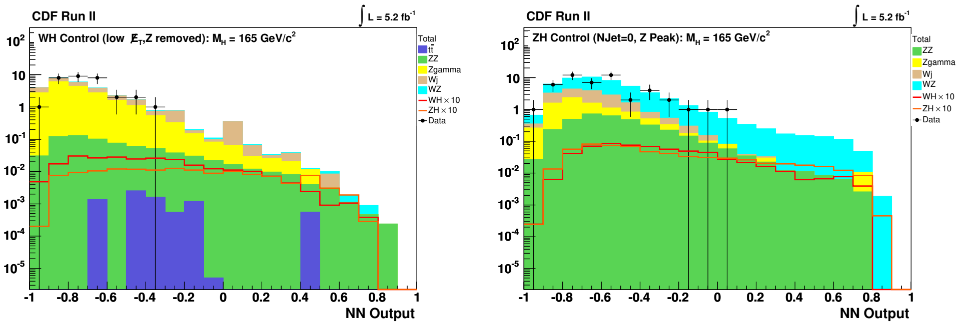

Figure 93: WH Control Region (10.0 GeV $< \not\!\!E_T <$ 20.0 GeV) and ZH Control Region (NJet= 0) neural network results from networks trained on the respective signal regions.

12.4 Trilepton Systematic Errors

The systematic uncertainties are summarized in Tables 77 and 78. Most values used are standard to all $H \to WW$ analyses, but there are several new systematics particular to this analysis which are discussed below.

- *$Z\gamma$ (and $W\gamma$) Scaling*: Note that the $W\gamma$ background is scaled down in dilepton $H \to WW$ analyses due to known mismodelling of photon conversions. We use the same scale factor for the $Z\gamma$ contribution since the same photon conversion affect is assumed. As such, we use the same systematic error associated with this scale factor. Also, we keep this systematic error correlated between the $Z\gamma$ of the trilepton analyses and $W\gamma$ of the dilepton analyses because of the common origin.

- *$Z\gamma$ Higher Order Diagrams*: We have for $W\gamma$ in the dilepton analysis the $W\gamma$ *higher order diagrams* systematic, which accounts for poor MC modeling beyond leading order. Likewise, we assume the same error of 11% for a new $Z\gamma$ *higher order diagrams* systematic since both are modelled by the Bauer MC generator.

- *b-Jet Fake Rate*: As explained in Section 12.1, a systematic uncertainty is assigned on the increase in expected $t\bar{t}$ contribution from allowing generator-level b-quarks to be matched with reconstructed leptons. The size of this systematic is half of the observed difference, or 27.8% for the WH analysis and 28.3% for the ZH analysis.

- *MC Run Dependence*: The $Z\gamma$ stntuples used cover only periods $0 - 11$, so we assign the MC run dependence systematic determined from the WW Pythia sample as described in Section 4.1.

- *Lepton, Trigger ID, Luminosity, Parton Distribution Function Model*: Finally, note that we do not use systematic errors for the lepton, trigger ID, luminosity, and PDF model efficiencies because of the scale factor derived from the $W\gamma$ control region in the dilepton analysis. Since we're measuring the $W\gamma$ normalization directly from data, that systematic should cover these effects. However, to be conservative– especially since we measure the scale factor in a control region with selection cuts that differ from our various signal regions–we keep the systematic uncertainties on the MC that are not related to normalization (higher-order kinematic effects, MC jet modelling).

12.5 Trilepton Analysis Results

Expected and observed limits for the trilepton signal regions are determined in the same manner as for the other channels, using the systematic uncertainties listed in Tables 77 and 78. Table 79 and Figure 94 show the expected (based on 10000 pseudo-experiments) and observed limits obtained from the WH region, while Table 80 and Figure 95 show

Uncertainty Source	WZ	ZZ	$Z\gamma$	$t\bar{t}$	Fakes	WH	ZH
Cross Section							
Total	*6.0%*	*6.0%*	5.0%	10.0%		**5.0%**	**5.0%**
Acceptance							
PDF Model	*2.7%*	*2.7%*		*2.1%*		*1.2%*	*0.9%*
Higher-order Diagrams	*10.0%*	*10.0%*	11.0%	10.0%		**10.0%**	**10.0%**
Jet Energy Scale			*-2.7%*				
Conversion Modeling			11.0%				
Jet Fake Rates					22.2%		
b-Jet Fake Rates				27.8%			
MC Run Dependence			5.0%				
Lepton ID Efficiencies	*2.0%*	*2.0%*		*2.0%*		*2.0%*	*2.0%*
Trigger Efficiencies	*2.1%*	*2.1%*		*2.0%*		*2.1%*	*2.1%*
Luminosity	*5.9%*	*5.9%*		*5.9%*		*5.9%*	*5.9%*

Table 77: Systematics table for Trilepton WH region background and signal processes. Cross section and higher-order diagram systematics are correlated for the diboson processes (shown in italic font) but uncorrelated between all other processes. Cross section and higher-order diagram systematics are correlated for the VH signal processes (shown in boldface font) but uncorrelated between all other processes.

Uncertainty Source	WZ	ZZ	$Z\gamma$	$t\bar{t}$	Fakes	WH	ZH
Cross Section							
Total	*6.0%*	*6.0%*	5.0%	10.0%		**5.0%**	**5.0%**
Acceptance							
PDF Model	*2.7%*	*2.7%*		*2.1%*		*1.2%*	*0.9%*
Higher-order Diagrams	*10.0%*	*10.0%*	11.0%	10.0%		**10.0%**	**10.0%**
Jet Energy Scale	*-9.7%*	*-5.2%*	*-1.0%*	*-9.3%*		*-8.6%*	*-1.1%*
Conversion Modeling			11.0%				
Jet Fake Rates					22.7%		
b-Jet Fake Rates				28.3%			
MC Run Dependence			5.0%				
Lepton ID Efficiencies	*2.0%*	*2.0%*		*2.0%*		*2.0%*	*2.0%*
Trigger Efficiencies	*2.1%*	*2.1%*		*2.0%*		*2.1%*	*2.1%*
Luminosity	*5.9%*	*5.9%*		*5.9%*		*5.9%*	*5.9%*

Table 78: Systematics table for Trilepton ZH region background and signal processes. Cross section and higher-order diagram systematics are correlated for the diboson processes (shown in italic font) but uncorrelated between all other processes. Cross section and higher-order diagram systematics are correlated for the VH signal processes (shown in boldface font) but uncorrelated between all other processes.

the expected (based on 10000 pseudo-experiments) and observed limits obtained from the ZH region.

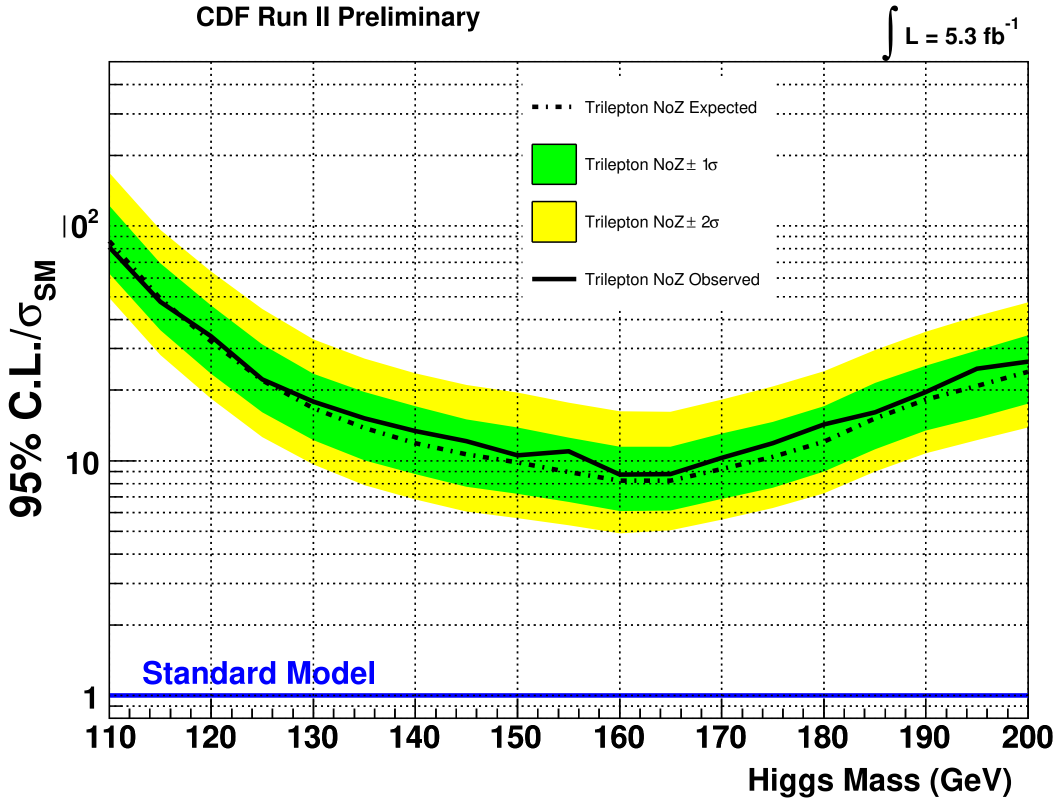

Figure 94: Expected and observed sensitivity for $H \rightarrow WW$ events in the trilepton WH search region.

WH Trileptons	110	115	120	125	130	135	140	145	150	155
$-2\sigma/\sigma_{SM}$	49.34	28.21	18.37	12.59	9.63	7.88	6.84	6.07	5.70	5.31
$-1\sigma/\sigma_{SM}$	62.58	35.91	23.41	16.05	12.24	10.04	8.77	7.76	7.21	6.66
Median/$\sigma_{\mathbf{SM}}$	**86.23**	**48.91**	**32.18**	**21.98**	**16.71**	**13.82**	**11.90**	**10.66**	**9.84**	**8.94**
$+1\sigma/\sigma_{SM}$	121.96	69.15	45.81	31.26	23.50	19.63	17.11	14.97	13.87	12.55
$+2\sigma/\sigma_{SM}$	168.55	96.17	64.13	44.40	32.85	27.29	23.60	21.05	19.57	17.70
Observed/$\sigma_{\mathbf{SM}}$	**81.17**	**47.46**	**33.81**	**22.16**	**17.83**	**15.17**	**13.37**	**12.11**	**10.55**	**10.99**
	160	165	170	175	180	185	190	195	200	
$-2\sigma/\sigma_{SM}$	4.90	5.07	5.61	6.28	7.25	8.99	10.76	12.22	13.87	
$-1\sigma/\sigma_{SM}$	6.09	6.13	6.89	7.67	9.00	11.17	13.44	15.23	17.53	
Median/$\sigma_{\mathbf{SM}}$	**8.20**	**8.23**	**9.23**	**10.40**	**12.09**	**15.16**	**18.20**	**20.79**	**23.94**	
$+1\sigma/\sigma_{SM}$	11.49	11.47	13.04	14.62	17.05	21.46	25.36	29.50	34.22	
$+2\sigma/\sigma_{SM}$	16.24	16.17	18.21	20.69	24.00	29.58	35.35	41.24	47.35	
Observed/$\sigma_{\mathbf{SM}}$	**8.73**	**8.76**	**10.31**	**11.86**	**14.28**	**16.09**	**19.61**	**24.61**	**26.44**	

Table 79: $H \to WW$ sensitivity normalized to NNLL σ_{SM} calculation in the WH trilepton region. In this region we consider potential signal contributions only from associated production with either a W or Z boson.

ZH Trileptons	110	115	120	125	130	135	140	145	150	155
$-2\sigma/\sigma_{SM}$	79.86	46.06	28.33	19.34	14.16	11.29	9.51	8.31	7.61	7.14
$-1\sigma/\sigma_{SM}$	104.84	60.71	36.92	25.34	18.62	14.66	12.26	10.58	9.79	9.05
Median/$\sigma_{\mathbf{SM}}$	**146.78**	**84.54**	**51.10**	**35.16**	**25.79**	**20.22**	**16.80**	**14.52**	**13.30**	**12.29**
$+1\sigma/\sigma_{SM}$	209.09	121.03	73.56	49.93	36.52	29.07	23.83	20.82	18.90	17.40
$+2\sigma/\sigma_{SM}$	295.48	171.45	105.14	70.43	51.18	40.46	33.30	29.34	26.81	24.75
Observed/$\sigma_{\mathbf{SM}}$	**205.34**	**121.36**	**78.49**	**52.97**	**43.11**	**34.25**	**28.58**	**23.18**	**22.98**	**20.83**
	160	165	170	175	180	185	190	195	200	
$-2\sigma/\sigma_{SM}$	6.71	6.76	7.36	7.80	8.57	10.38	12.00	13.26	14.59	
$-1\sigma/\sigma_{SM}$	8.36	8.34	9.01	9.72	10.62	12.79	14.76	16.24	17.91	
Median/$\sigma_{\mathbf{SM}}$	**11.29**	**11.14**	**12.08**	**13.08**	**14.34**	**17.09**	**19.92**	**21.89**	**24.30**	
$+1\sigma/\sigma_{SM}$	15.88	15.57	17.02	18.45	20.22	24.15	28.17	30.73	34.11	
$+2\sigma/\sigma_{SM}$	22.70	22.07	23.52	25.74	28.16	34.03	39.29	42.66	47.75	
Observed/$\sigma_{\mathbf{SM}}$	**15.26**	**13.95**	**18.03**	**17.80**	**21.76**	**25.34**	**32.77**	**33.12**	**35.76**	

Table 80: $H \to WW$ sensitivity normalized to NNLL σ_{SM} calculation in the ZH trilepton region. In this region we consider potential signal contributions only from associated production with either a W or Z boson.

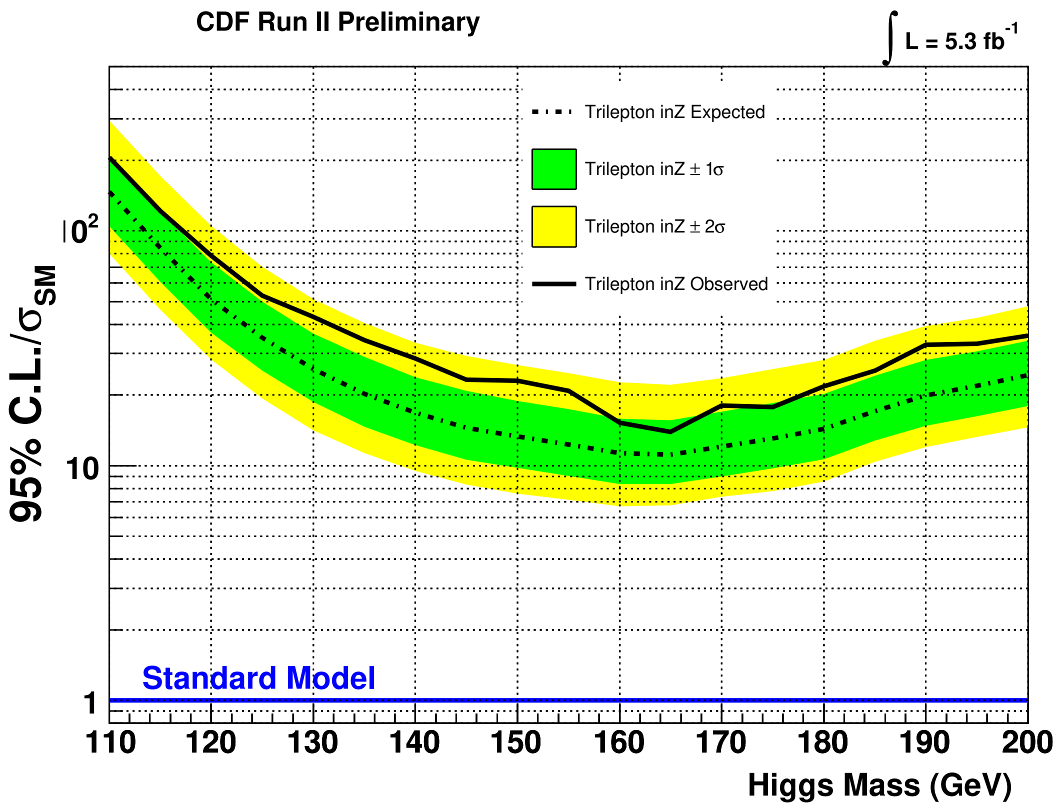

Figure 95: Expected and observed sensitivity for $H \to WW$ events in the trilepton ZH search region.

13 Combination of All High-Mass Search Channels

Because we separate events into High S/B and Low S/B categories in the 0J and 1J regions, there are a total of six opposite-sign channels (including opposite-sign events with low dilepton mass). Each channel has its own unique background composition and signal to background ratio, and therefore will contribute different relative sensitivities to the overall cross section limit. As previously mentioned, in the 2J region, we choose to consider all events as a single channel, primarily to maintain sufficient statistics for some of the Monte-Carlo based templates as well as the W+jets (Fakes) templates which lack statistics for the higher jet multiplicities.

Incorporating the same-sign and trilepton regions leaves us with a total of nine channels from which we produce our combined limits. As in the case of 2 or more-jets region, events in the same-sign channel are grouped as a single channel, primarily to maintain sufficient statistics for some of the Monte-Carlo based templates as well as the $W+$ jets (Fakes) templates which lack statistics for the higher jet multiplicities. In addition, dilepton categories including a PHX electron are removed in the same-sign, and most categories treated as Low S/B in the 0-jet and 1-jet regions contain at least one PHX electron. The sensitivity in the same-sign region comes almost entirely from associated production channels (WH, ZH).

As is the case in each of the individual regions, we use Tom Junk's MCLimit [32] to produce a combined limit from all nine channels. We have opted to set our own statistical errors on each bin and use the CSM_GAUSSIAN_BINERR option in MCLimit. The bin errors are nothing more then the typical statistical errors calculated by ROOT's Sumw2. The current limits from the combination of all nine channels are shown in Table 81 and Figure 96. The updates presented in this note, combining about 500 pb^{-1} of new data with numerous small analysis improvements including the new trilepton searches, leads to an increase in overall expected sensitivity of about 10-15% over the previous analysis presented in CDF 9863 and 9864.

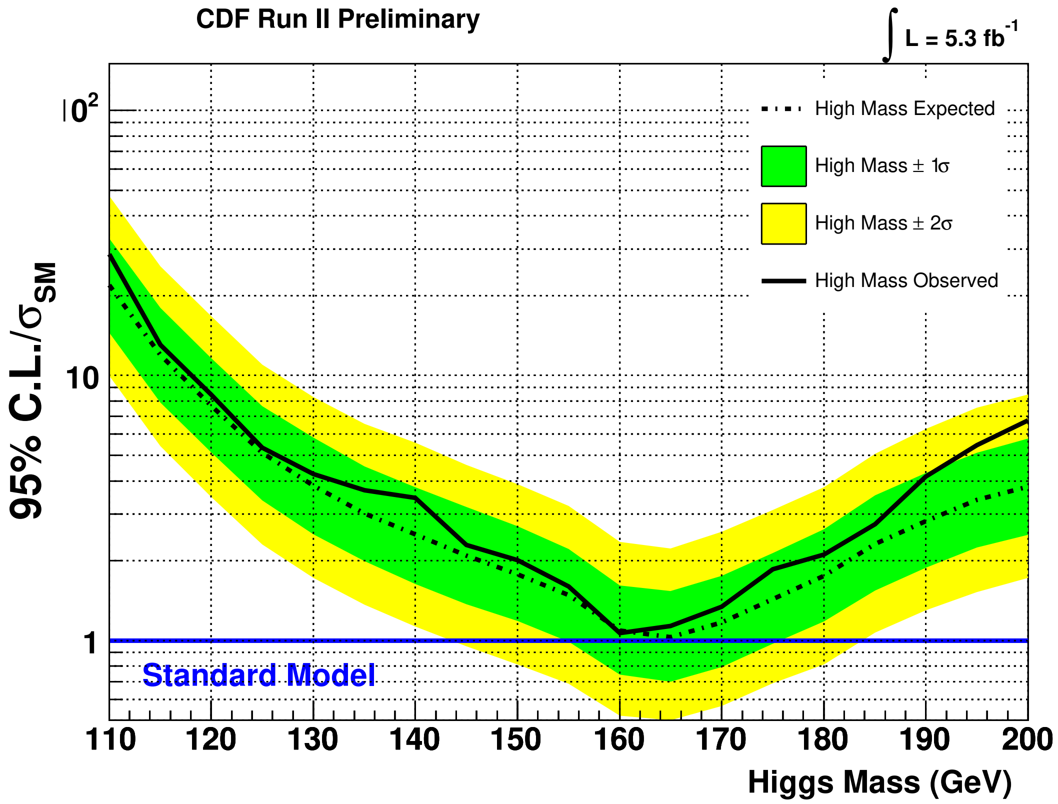

Figure 96: $H \to WW$ sensitivity normalized to NNLL σ_{SM} calculation for combination of all nine channels (0-jet High S/B, 0-jet Low S/B, 1-jet High S/B, 1-jet Low S/B, 2 or more-jets, low-M_{ll}, same-sign, ZH trilepton, and WH trilepton).

$H \to WW^{(*)}$	110	115	120	125	130	135	140	145	150	155
$-2\sigma/\sigma_{SM}$	9.97	5.40	3.48	2.31	1.72	1.37	1.13	0.95	0.81	0.69
$-1\sigma/\sigma_{SM}$	14.50	7.91	5.12	3.37	2.52	1.99	1.63	1.37	1.18	0.99
Median/$\sigma_{\mathbf{SM}}$	**21.88**	**11.93**	**7.70**	**5.11**	**3.83**	**3.02**	**2.51**	**2.09**	**1.78**	**1.48**
$+1\sigma/\sigma_{SM}$	32.92	17.98	11.66	7.67	5.83	4.54	3.79	3.18	2.70	2.22
$+2\sigma/\sigma_{SM}$	47.49	25.75	16.80	11.00	8.30	6.57	5.57	4.59	3.89	3.22
Observed/$\sigma_{\mathbf{SM}}$	**28.72**	**13.03**	**8.47**	**5.33**	**4.25**	**3.69**	**3.45**	**2.30**	**2.01**	**1.60**
	160	165	170	175	180	185	190	195	200	
$-2\sigma/\sigma_{SM}$	0.52	0.50	0.56	0.69	0.81	1.07	1.30	1.52	1.72	
$-1\sigma/\sigma_{SM}$	0.74	0.70	0.80	0.97	1.18	1.54	1.88	2.24	2.51	
Median/$\sigma_{\mathbf{SM}}$	**1.09**	**1.03**	**1.17**	**1.43**	**1.75**	**2.32**	**2.82**	**3.39**	**3.80**	
$+1\sigma/\sigma_{SM}$	1.61	1.54	1.75	2.14	2.63	3.53	4.28	5.11	5.79	
$+2\sigma/\sigma_{SM}$	2.35	2.22	2.57	3.11	3.79	5.06	6.28	7.55	8.47	
Observed/$\sigma_{\mathbf{SM}}$	**1.07**	**1.13**	**1.34**	**1.86**	**2.11**	**2.75**	**4.15**	**5.45**	**6.78**	

Table 81: Limits on Higgs production using all nine channels (0-jet High S/B, 0-jet Low S/B, 1-jet High S/B, 1-jet Low S/B, 2 or more-jets, low-M_{ll}, same-sign, ZH trilepton, and WH trilepton) and 5.3 fb^{-1}.

References

[1] D. Benjamin, D. Hidas, M. Kruse, M. Casarsa, E. James, S. Jindariani, B. Rutherford, R. Lysak, A. Robson, R. St. Denis, et al., *Updated search for h to ww production using 4.8 fb-1* (2009), CDF/ANAL/EXOTIC/CDFR/9863.

[2] D. Benjamin, D. Hidas, M. Kruse, M. Casarsa, E. James, S. Jindariani, B. Rutherford, R. Lysak, A. Robson, R. St. Denis, et al., *Search for higgs to ww production in the low mll region using 4.8 fb-1* (2009), CDF/ANAL/EXOTIC/CDFR/9864.

[3] D. Benjamin, D. Hidas, M. Kruse, E. James, S. Jindariani, B. Rutherford, R. Lysak, J. Nett, A. Robson, R. St. Denis, et al., *Search for vh 2̆192 vww sm higgs production in the trilepton signature* (2009), CDF/PHYS/EXOTIC/CDFR/10020.

[4] D. Benjamin, D. Hidas, M. Kruse, E. James, S. Jindariani, B. Rutherford, R. Lysak, A. Robson, R. St. Denis, P. Bussey, et al., *Updated search for h to ww production using 3.6 fb-1* (2009), CDF/PHYS/EXOTIC/CDFR/9697.

[5] E. Lipeles, M. Neubauer, S.-C. Hsu, and F. Wurthwein, *Lepton id for multilepton diboson analyses* (2007), CDF/PHYS/ELECTROWEAK/CDFR/8538.

[6] *Cdf data processing status, http://www-cdf.fnal.gov/tiki/tiki-index.php?page=ProductionFarm.DataProcessing.*

[7] U. Baur and E. L. Berger, Phys. Rev. **D47**, 4889 (1993).

[8] D. Benjamin, A. Goshaw, M. Kirby, B. Heinemann, H. Hayward, and N. Tanimoto, *Measurement of the wgamma and zgamma production cross sections using 200/pb of run 2 data* (2003), CDF/PHYS/ELECTROWEAK/CDFR/6601.

[9] A. Nagano, J. Deng, A. Goshaw, B. Heinemann, and T. Phillips, *Measurement of the wgamma to e nu gamma cross sections using 1/fb of cdf data* (2007), CDF/ANAL/ELECTROWEAK/CDFR/8756.

[10] J. Deng, M. Goncharov, A. Goshaw, and T. Phillips, *Measurement of z-gamma production in ppbar collisions at sqrt(s) = 1.96 tev* (2009), CDF/PUB/ELECTROWEAK/CDFR/9904.

[11] S. Moch and P. Uwer, Nucl. Phys. Proc. Suppl. **183**, 75 (2008), `arXiv:0807.2794 [hep-ph]`.

[12] J. Campbell and K. Ellis, *Mcfm - monte carlo for femtobarn processes* (2005), http://mcfm.fnal.gov/.

[13] J. M. Campbell and R. K. Ellis, Phys. Rev. **D60**, 113006 (1999), `hep-ph/9905386`.

[14] U. Baur, T. Han, and J. Ohnemus, Phys. Rev. **D57**, 2823 (1998).

[15] *See the* TEV4LHC *working group webpage*, http://maltoni.home.cern.ch/maltoni/TeV4LHC/SM.html.

[16] C. Anastasiou, K. Melnikov, and F. Petriello, Nucl. Phys. **B724**, 197 (2005), `hep-ph/0501130`.

[17] S. Catani and M. Grazzini, Phys. Rev. Lett. **98**, 222002 (2007), `hep-ph/0703012`.

[18] M. Grazzini, JHEP **0802**, 043 (2008), `0801.3232 [hep-ph]`.

[19] HNNLO *program is publically available at*, http://theory.fi.infn.it/grazzini/codes.html.

[20] S. Catani, D. de Florian, M. Grazzini, and P. Nason, JHEP **0307**, 028 (2003), `hep-ph/0306211`.

[21] C. Anastasiou, G. Dissertori, M. Grazzini, F. Stoeckli, and B. R. Webber (2009), `0905.3529 [hep-ph]`.

[22] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt (2009), `0901.0002 [hep-ph]`.

[23] S. Moch and A. Vogt, Phys. Lett. B **631**, 48 (2005), `hep-ph/0508265`.

[24] F. Petriello, *Talk from* CDF *week june 16-20, 2008, with comparisons between* N^{3}LO *and* FEHIP*, day 3 in webtalks*, http://www-cdf.fnal.gov/internal/WebTalks/Archive/0806/080618_collaboration_meeting_day_3/01_080618_collaboration_meeting_day_3_Frank_Petriello_1_roser_Fermi2008.pdf.

[25] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, Eur. Phys. J. C **14**, 133 (2000), `hep-ph/9907231`.

[26] *All* MRST *parton distributions are available from the webpage*, http://durpdg.dur.ac.uk/hepdata/mrs.html.

[27] D. de Florian and M. Grazzini, Phys. Lett. B **674**, 291 (2009), `0901.2427 [hep-ph]`.

[28] K. Yorita and Y.-K. Kim, *Search for the standard model higgs boson using tau* (2008), CDF/ANAL/EXOTIC/CDFR/9179.

[29] G. Bozzi, S. Catani, D. de Florian, and M. Grazzini, Phys. Lett. **B564**, 65 (2003), `hep-ph/0302104`.

[30] G. Bozzi, S. Catani, D. de Florian, and M. Grazzini, Nucl. Phys. **B737**, 73 (2006), `hep-ph/0508068`.

[31] M. Feindt, Technical Report IEKP-KA/01-1 (2001).

[32] T. Junk, *Sensitivity, exclusion and discovery with small signals, large backgrounds, and large systematics* (2006), CDF/DOC/STATISTICS/PUBLIC/8128.

[33] C. Rott, H. Bachacou, K. Hoffman, A. Dominguez, W. Yao, D. Glenzinski, D. Bortoletto, and M. Shochet, *Secvtx optimization for the 2003 winter conferences* (2002), CDF/DOC/SEC VTX/CDFR/6242.

[34] J. Freeman, T. Junk, E. Palencia, and S. Grinstein, *Secvtx mistag matrices for 2.2 fb^{-1} of data up to p13* (2008), CDF/DOC/SEC VTX/CDFR/9280.

[35] D. Benjamin, D. Hidas, M. Kruse, E. James, S. Jindariani, B. Rutherford, R. Lysak, A. Robson, R. St. Denis, P. Bussey, et al., *Search for h to ww production using 3.6 fb-1* (2009), CDF/PHYS/EXOTIC/CDFR/9685.

[36] D. Benjamin, D. Hidas, M. Kruse, V. Necula, S. Burke, E. James, S. Jindariani, R. Lysak, S. Hsu, E. Lipeles, et al., *Search for h to ww production using 3.0 fb-1* (2008), CDF/PHYS/EXOTIC/CDFR/9402.

A Scale factors used in Ntp processing

This section summarizes the scale factors we apply in our our analysis at the level of Ntuple processing. It is meant mainly for our internal purposes.

We apply the following set of the scale factors in this version ("v17") of analysis:

- **WW luminosity rescaling** – the reasoning for this scaling is described in Sec. 4.1. The scale factors are:
  - for 0 jet events: **0.948**
  - for 1 jet events: **0.998**
  - for 2 jet events: **1.041**
- $W\gamma$ **normalization scaling** – **0.92** scale factor is applied to the new $W\gamma$ samples with modeling of later run periods. A **0.83** scale factor is applied to the $Z\gamma$ samples which have run-dependence only through period 11. Both scale factors are determined from the same-sign low-M_{ll} region as described in Section 10.4.
- $t\bar{t}$ **b-tag and anti-btag scale factor** We apply anti-btag scale factor of **1.087** on $t\bar{t}$ events in OS 2+ Jets signal region. Accordingly, we apply b-tag scale factor of **0.94** on events in $t\bar{t}$ control region.
- $t\bar{t}$ **Silicon/No-Silicon good run list scale factor** – we apply a scale factor which to account for difference between Si/No-Si GRL.
  - for anti-b-tagged region (signal region): **1.023**
  - for b-tagged region ($t\bar{t}$ control region): **0.9695**