CDF Note 10217

Search for Standard Model Higgs Boson Production in Association with a W Boson Using Matrix Element Techniques with 5.6 fb^{-1} of CDF Data

The CDF Collaboration
URL http://www-cdf.fnal.gov
(Dated: July 9, 2010)

We present a search for Standard Model Higgs boson production in association with a W boson using 5.6 fb^{-1} of CDF II data collected between 2002 and 2010. This search is performed using a matrix element technique to calculate event probability densities for the signal and background hypotheses.

We observe no evidence for a Higgs boson signal and set 95% confidence level upper limits on the WH production cross section times the branching ratio of the Higgs boson to decay to $b\bar{b}$ pairs of $\sigma(p\bar{p} \to WH) \times BR(H \to b\bar{b})/SM < 2.1$ to 35.3 for Higgs boson masses between $m_H = 100$ GeV/c^2 and $m_H = 150$ GeV/c^2. The expected (median) limit estimated from pseudo-experiments is: $\sigma(p\bar{p} \to WH) \times BR(H \to b\bar{b})/SM < 2.5$ to 27.5 at 95% C.L.

Preliminary Results

INTRODUCTION

Finding evidence for Higgs boson production in association with a W boson is extremely challenging since at best it is rarely produced ($\sigma_{WH} \sim 0.1$ pb) in comparison with other processes with the same final state like $W + b\bar{b}$ and top. The signal to background ratio of the analysis is expected to be tiny, typically on the order of S/B$\sim$1/100.

This note describes the search for WH production using a matrix element technique. The implementation of this method is very similar to the one applied to the single-top search described in detail here [1].

DATA SAMPLE & EVENT SELECTION

The candidate events for this analysis are selected by requiring a W + 2 or 3 jet event topology where the W decays leptonically, $W \to e\nu_e$ or $W \to \mu\nu_\mu$. We require events to contain an isolated electron (CEM or PHX) or muon with offline E_T or p_T $>$ 20 GeV and Missing Transverse Energy, $\not\!\!E_T$ $>$ 20 GeV (25 GeV for forward electrons). Jets are clustered with a cone size of $\Delta R < 0.4$ and are required to have E_T $>$ 20 GeV after correcting the jets for instrumental effects and have $|\eta|$ $<$ 2.0. At least one jet should be identified as a b-jet using the secondary vertex tag requirement. The secondary vertex tagging algorithm identifies tracks associated with the jet originating from a vertex displaced from the primary vertex indicative of decay particles from relatively long lived B mesons. Another tagging algorithm used in this analysis to identify a b-jet is called JetProbability (JP), this algorithm uses tracks associated with a jet to determine the probability for these to come from the primary vertex of the interaction. If two or more jets are tagged by the secondary vertex tagging algorithm (SecVtx), we call the event SVSV; otherwise if only one jet is tagged by SecVtx but another one is tagged by JetProbability, we call it SVJP; and, finally, if only one jet is tagged by SecVtx we call the event SVnoJP.

We veto:

- Dilepton events, events with more than one lepton.
- Z-boson events, events where the invariant mass of the lepton and a second track falls in the Z-boson mass window $76 < m_Z < 106$ GeV/c^2.
- Leptons from photon conversion events: electrons traversing the detector can emit photons due to bremsstrahlung radiation from the interaction with detector material. These photons can then convert to electron-positron pairs which are characterized by two tracks with a small opening angle and a vertex far from the primary interaction point. Events which contain such a conversion are identified and removed from the data sample.
- QCD multi-jet events. Basically, we require:
  - for central electrons (CEM):
    - $m_T(W) > 20$ GeV
    - METsig $\geq$ -0.05$\times m_T(W)$ + 3.5
    - METsig $\geq$ 2.5 - (2.5/0.8)$\times\Delta\phi(\not\!\!E_T$,jet2),
  - for plug electrons (PHX):
    - $\not\!\!E_T > 25$ GeV
    - $m_T(W) > 20$ GeV, METsig > 2
    - $\not\!\!E_T > 45$ - 30$\times\Delta\phi(\not\!\!E_T$,jet1)
    - $\not\!\!E_T > 45$ - 30$\times\Delta\phi(\not\!\!E_T$,jet2)
  - and for muons: $m_T(W) > 10$ GeV

  where METsig is:

$$METsig = \frac{\not\!\!E_T}{\sqrt{\sum_{jets} C_{JES}^2 \; cos^2(\Delta\phi_{jet,\not\!\!E_T}) \; + \; cos^2(\Delta\phi_{vtz,corr})}} \tag{1}$$

  where C$_{JES}$ is the jet energy correction factor, and $\phi_{vtz,corr}$ is the azimuthal angle between corrected and uncorrected missing transverse energy.

BACKGROUND ESTIMATE

Estimating the background contribution after applying the event selection to the WH candidate sample is an elaborate process. Physical backgrounds surviving the selection are $t\bar{t}$, s-channel and t-channel single top, Z + jets, W + heavy-flavor jets, i.e. $W + b\bar{b}$, $W + c\bar{c}$ and $W + c$, and diboson events WW, WZ, and ZZ. Instrumental backgrounds originate from mis-tagged W + jets events (W events with light-flavor jets, i.e. with u, d, s-quark and gluon content, misidentified as heavy-flavor jets) and from non-W + jets events (multi-jet events where one jet is erroneously identified as a lepton). We determine the W + jets normalization from the data and estimate the fraction of the candidate events with heavy-flavor jets using ALPGEN Monte Carlo samples [2]. The heavy-flavor fractions are calibrated in the b-tagged W + 1 jet sample using data distributions which are sensitive to distinguish light-flavor from heavy-flavor jets, e.g. the mass of the secondary-vertex or the output of a Neural Network jet-flavor separator [3]. Based on these studies, the heavy flavor content was corrected by a factor $K_{HF} = 1.4 \pm 0.4$. The probability that a W + light-flavor jet is mis-tagged is parameterized using large statistics generic multi-jet data. The instrumental background contribution from non-W events is estimated using side-band data with low missing transverse energy, devoid of any signal, and we subsequently extrapolate the contribution into the signal region with large missing transverse energy, $\not\!\!E_T$ >20 GeV (25 GeV for forward electrons).

For the rest we use MC based background estimations, we have used the theoretical cross sections shown in Table I. NLO cross section calculations exist for diboson and $t\bar{t}$ production, thereby making the estimation of their contribution a relatively straightforward process. We have added the expected contribution from WH signal using the SM cross sections and branching ratios shown in Table II. The expected signal and background yield in the W + 2 and 3 jet samples are shown in Table III and IV, respectively.

Process	**Theoretical Cross Section (pb)**
s-channel	0.884 ± 0.11
t-channel	1.980 ± 0.25
WW	11.66 ± 0.7
WZ	3.46 ± 0.3
ZZ	1.51 ± 0.25
$t\bar{t}$	6.7 ± 0.8
Z+jets	787.4 ± 85.0

TABLE I: Theoretical crosss sections used for the MC based background estimation.

Higgs Mass (GeV/c^2)	**BR($H \to b\bar{b}$)**	**σ (pb)**	**σ x BR($H \to b\bar{b}$) (pb)**
100	0.812	0.286	0.232
105	0.796	0.253	0.201
110	0.770	0.219	0.169
115	0.732	0.186	0.136
120	0.679	0.153	0.104
125	0.610	0.136	0.083
130	0.527	0.120	0.063
135	0.436	0.103	0.045
140	0.344	0.086	0.030
145	0.256	0.078	0.020
150	0.176	0.070	0.012

TABLE II: SM branching ratios ($H \to b\bar{b}$) and cross sections for all Higgs masses.

Process	**SVSV**	**SVJP**	**SVnoJP**
All Pretag Cands.	121690	121690	107533
WW	0.89 ± 0.20	3.28 ± 1.29	105.75 ± 13.16
WZ	8.26 ± 1.15	6.19 ± 0.98	35.10 ± 3.90
ZZ	0.30 ± 0.05	0.25 ± 0.05	1.37 ± 0.21
$t\bar{t}$ (lepton+jets)	47.00 ± 7.76	37.59 ± 6.84	205.25 ± 28.47
$t\bar{t}$ (dilepton)	28.22 ± 4.60	19.98 ± 3.35	79.62 ± 10.99
Single top (t-channel)	6.31 ± 1.08	6.26 ± 1.32	115.64 ± 16.65
Single top (s-channel)	26.24 ± 4.28	18.37 ± 3.08	65.98 ± 9.14
Z+jets	4.24 ± 0.70	5.13 ± 1.31	79.72 ± 12.17
Total MC	121.47 ± 15.31	97.05 ± 14.80	688.43 ± 59.53
Wbb	142.30 ± 45.73	120.71 ± 39.21	978.43 ± 295.18
Wcc/Wc	13.83 ± 4.73	45.62 ± 16.93	958.46 ± 295.60
Total HF	156.12 ± 50.14	166.32 ± 54.92	1936.89 ± 588.37
Wjj	4.66 ± 1.49	19.32 ± 10.58	945.47 ± 137.97
Non-W	18.95 ± 7.58	29.35 ± 11.74	298.31 ± 119.32
Total Prediction	**301.20 ± 52.99**	**312.05 ± 59.03**	**3869.11 ± 618.87**
WH (m_H = 115 GeV/c^2)	4.06 ± 0.48	2.80 ± 0.35	10.27 ± 0.81
Observed	**282**	**311**	**3878**

TABLE III: Number of expected signal and background events, in the 2 jet bin, in 5.6 fb^{-1} of CDF data, passing all the event selection requirements.

Process	**SVSV**	**SVJP**	**SVnoJP**
All Pretag Cands.	22036	22036	19168
WW	0.99 ± 0.20	2.62 ± 0.93	32.84 ± 3.95
WZ	2.27 ± 0.32	1.94 ± 0.35	9.38 ± 1.05
ZZ	0.19 ± 0.03	0.15 ± 0.03	0.59 ± 0.09
$t\bar{t}$ (lepton+jets)	187.66 ± 30.82	160.87 ± 29.02	503.46 ± 69.66
$t\bar{t}$ (dilepton)	25.43 ± 4.14	18.18 ± 3.08	57.59 ± 7.97
Single top (t-channel)	5.56 ± 0.93	4.95 ± 0.92	26.14 ± 3.68
Single top (s-channel)	8.87 ± 1.45	6.82 ± 1.16	19.48 ± 2.71
Z+jets	2.95 ± 0.50	4.00 ± 1.07	29.65 ± 4.36
Total MC	233.91 ± 34.54	199.55 ± 33.50	679.13 ± 76.30
Wbb	48.63 ± 15.91	47.44 ± 15.78	258.21 ± 78.26
Wcc/Wc	7.11 ± 2.48	22.89 ± 8.59	236.53 ± 72.93
Total HF	55.75 ± 18.16	70.33 ± 23.71	494.74 ± 150.23
Wjj	3.18 ± 1.06	11.33 ± 5.88	254.75 ± 37.58
Non-W	9.63 ± 3.85	21.47 ± 8.59	93.34 ± 37.33
Total Prediction	**302.46 ± 39.22**	**302.68 ± 42.34**	**1521.97 ± 176.63**
WH (m_H = 115 GeV/c^2)	1.16 ± 0.14	0.85 ± 0.12	2.57 ± 0.21
Observed	**318**	**302**	**1491**

TABLE IV: Number of expected signal and background events, in the 3 jet bin, in 5.6 fb^{-1} of CDF data, passing all the event selection requirements.

ANALYSIS METHOD

Matrix Element Method

The matrix element analysis relies on the evaluation of event probability densities for signal and background processes based on the Standard Model differential cross-section calculation.

In general a differential cross-section is given by [4]:

$$d\sigma = \frac{(2\pi)^4 |M|^2}{4\sqrt{(q_1 \cdot q_2)^2 - m_{q_1}^2 m_{q_2}^2}} \, d\Phi_n(q_1 + q_2; p_1, .., p_n) \tag{2}$$

where $|M|$ is the Lorentz invariant matrix element; q_1, q_2 and m_{q_1}, m_{q_2} are the four momenta and masses of the incident particles, and $d\Phi_n$ is the n-body phase space given by [4]:

$$d\Phi_n(q_1 + q_2; p_1, .., p_n) = \delta^4(q_1 + q_2 - \sum_{i=1}^{n} p_i) \prod_{i=1}^{n} \frac{d^3 p_i}{(2\pi)^3 2E_i} \tag{3}$$

The CDF detector would be 'ideal' if we could measure all four momenta of the initial and final state particles perfectly. In this case we could use this formula without modification and normalize it to the total cross section to define the event probability:

$$P_{evt} \sim \frac{d\sigma}{\sigma}$$

However, several effects have to be considered: (1) the initial state interaction is initiated by partons inside the proton and antiproton, (2) neutrinos in the final state are not identified directly, and (3) the energy resolution of the detector can not be ignored. To address the first point, the differential cross section is folded over the parton distribution functions. To address the second and third points, we integrate over all particle momenta which we do not measure (e.g. p_z of the neutrino), or do not measure very well, due to resolution effects (e.g. jet energies). The integration reflects the fact that we want to sum over all possible particle variables (y) leading to the observed set of variables (x) measured with the CDF detector. The mapping between the particle variables (y) and the measured variables (x) is established with the transfer function, $W(y, x)$. After incorporating the effects mentioned above, the event probability takes the form:

$$P(x) = \frac{1}{\sigma} \int d\sigma(y) dq_1 dq_2 f(y_1) f(y_2) W(y, x) \tag{4}$$

where $d\sigma(y)$ is the differential cross section in terms of the particle variables; $f(y_i)$ are the PDFs, with y_i being the fraction of the proton momentum carried by the parton ($y_i = E_{q_i}/E_{beam}$); and $W(y, x)$ is the transfer function. Substituting Equation 2 and 3 into Equation 4, and considering a final state with four particles (n=4), the event probability becomes:

$$P(x) = \frac{1}{\sigma} \int 2\pi^4 |M|^2 \frac{f(y_1)}{|E_{q_1}|} \frac{f(y_2)}{|E_{q_2}|} W(y, x) d\Phi_4 dE_{q_1} dE_{q_2} \tag{5}$$

where the masses and transverse momenta of the initial partons are neglected (i.e. $\sqrt{(q_1 \cdot q_2)^2 - m_{q_1}^2 m_{q_2}^2} \simeq 2E_{q_1} E_{q_2}$).

In this analysis, we calculate event probability densities for the WH signal, as well as for the s-channel and t-channel single top, $t\bar{t}$, $Wb\bar{b}$, $Wc\bar{c}$, Wcj, Mistags (Wgg and Wgj) and diboson background processes. The non-W events are assumed to be represented fairly well by the background probability density.

We calculate the matrix element ($|M|^2$) for the event probability density at leading order perturbation theory by using the HELAS (HELicity Amplitude Subroutines for Feynman Diagram Evaluations) package. The correct subroutines for a given process are automatically generated by the MadGraph program. We use different subroutines for calculating event probabilities for the WH signal and s-channel, t-channel, $Wb\bar{b}$, $Wc\bar{c}$, Wcj, Wgg, diboson and $t\bar{t}$ background hypotheses.

Transfer Function: TF(E_p, E_j)

The transfer function, $W(y, x)$, provides the probability of measuring the set of observable variables (x) that correspond to the set of production variables (y). The set (y) represents all final state particle momenta at the particle level, while the set (x) represents the measured momenta (of the corresponding object) with the CDF detector. In the

case of well-measured objects, $W(y,x)$ is taken as a δ-function (i.e. the measured momenta are used in the differential cross section calculation). When the detector resolution cannot be ignored, $W(y,x)$ is taken as a Gaussian-type function. For unmeasured quantities, like the momenta of the neutrino, the transfer function is unity (the transverse momenta of the neutrino, however, can be inferred from energy and momentum conservation).

The mapping between parton and jet energies is determined by the transfer function $W_{jet}(E_{parton}, E_{jet})$. The jet energy corrections correct the energies of jets in a way that the means of the corrected jet energies and the original parton energies are equal. Such corrections, however, do not account for the shape of the difference in energies: the shape of the $\delta_E = (E_{parton} - E_{jet})$ distribution.

We parameterize the light-jets, gluons and b and c jets transfer functions in a different way due to the different kinematics.

In order to better reproduce the real parton energy, we train 23 Neural Networks (NN) with the Stuttgart Neural Network Simulator (SNNS) using the ROOT interface ROOTSNNSv3.0 [5]. Each NN is trained for a different type of jet (b, c, light and gluons) and a different physic process (WH [100-150] GeV/c^2, $Wb\bar{b}$, $Wc\bar{c}$, Wcg, $t\bar{t}$, s-channel, t-channel, Wjg, Wgg, WW, and WZ) as shown in Table V. All of them based in MC events.

	b jets	**c jets**	**light jets**	**gluons**
WH	X			
$\boldsymbol{Wb\bar{b}}$	X			
$\boldsymbol{Wc\bar{c}}$		X		
$\boldsymbol{t\bar{t}}$	X			
s-channel	X			
t-channel	X		X	
Wcg		X		X
Wjg			X	X
Wgg				X
WW-WZ			X	

TABLE V: Type of jet used to train the different NNs for each process

For all the training we use 7 input variables related to the jet kinematics: the total energy of the jet corrected (E_j), the sum of the transverse momentum of the tracks in the jet (SumE) [9], the transverse momentum of the jet (p_T), the ϕ and η of the jet, the raw (measured) energy of the jet (RawE_j), and the energy of jets clustered with a cone size of $\Delta R < 0.7$, (Ejcone7) [10].

The goal of using a NN to reproduce the parton energy is to use the output of the NN (instead of the jet energy, E_j) as input of the transfer function. Figure 1 shows the difference between the parton energy and the jet energy (black histogram) and between the parton energy and the NN output (red solid histogram) for four different physic processes (WH 115 GeV/c^2, diboson (WW, WZ), $Wb\bar{b}$, and Wgg). It is clear that the NN output is closer to the parton energy than the corrected jet energies and that the distribution is also narrower.

Therefore, since the NN output provides a better jet resolution, using it as input of the transfer function should help to improve the performance of the transfer function.

Event Probability Discriminant

The event probability density makes use of all measured quantities to specify each event. This should provide good discrimination between signal and background. It uses both possible jet combinations in the event so that the right jet-parton association is always included.

We use the event probability densities as ingredients to build an event probability discriminant (EPD), i.e. a distribution which separates signal from background which we can use to fit the data. Perhaps the most intuitive discriminant is the ratio of signal over signal and background probabilities, $EPD = P_s/(P_s + P_b)$. This discriminant is close to zero for ratios dominated by P_b and close to unity for ratios dominated by P_s. The expresion 6 is the event probability discriminant we use in this analysis. We introduce extra non-kinematic information by using the output neural network b-tagger (b) which assigns a probability for each b-tagged jet ($0 < b < 1$).

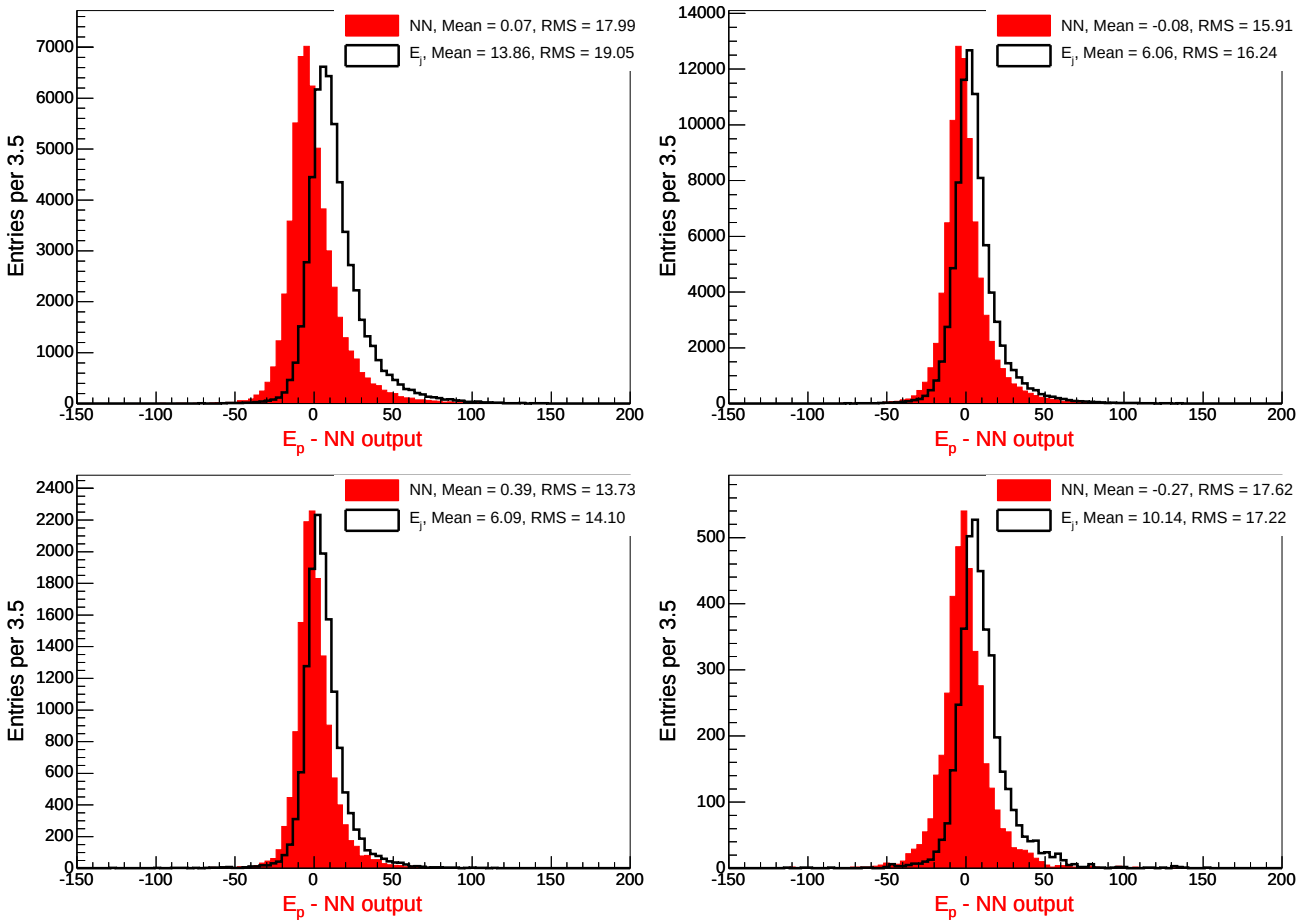

FIG. 1: Difference between the parton energy and the jet energy (black) and the NN output (red). From top left to bottom right: WH (115 GeV) events for b-jets, $Wb\bar{b}$ events for b-jets, diboson events for light jets and Wgg events for gluon.

$$EPD = \frac{b \cdot P_{WH}}{b\left(P_{WH} + P_{singletop} + P_{Wb\bar{b}} + P_{t\bar{t}}\right) + (1-b)\left(P_{Wc\bar{c}} + P_{Wcj} + P_{Wjj} + P_{diboson}\right)} \tag{6}$$

In the search for SM Higgs production, we create separate EPD discriminants for each Higgs mass point, in different b-tagging categories and jet bins:

- 2 jets SVnoJP, 3 jets SVnoJP
- 2 jets SVJP, 3 jets SVJP
- 2 jets SVSV, 3 jets SVSV

This gives us the ability to tune the discriminants independently of the number of jets and b-tagging category.

The level of agreement between data and Monte Carlo simulation is checked for all the input variables in the six signal regions as well as in the control region.

In addition to the validation of the input variables, the discriminants are validated in the control region of 2 and 3 jets with no b-tags, commonly known as the untag region (W + light flavor dominant), the Monte Carlo simulation verus data comparisons are shown in Figure 2. The event probability discriminants in the signal regions for 2 and 3 jet events are shown in Figure 3 and 4, respectively.

We use the EPD discriminants in all the b-tag channels and all the Higgs masses as the final discriminant for a WH search.

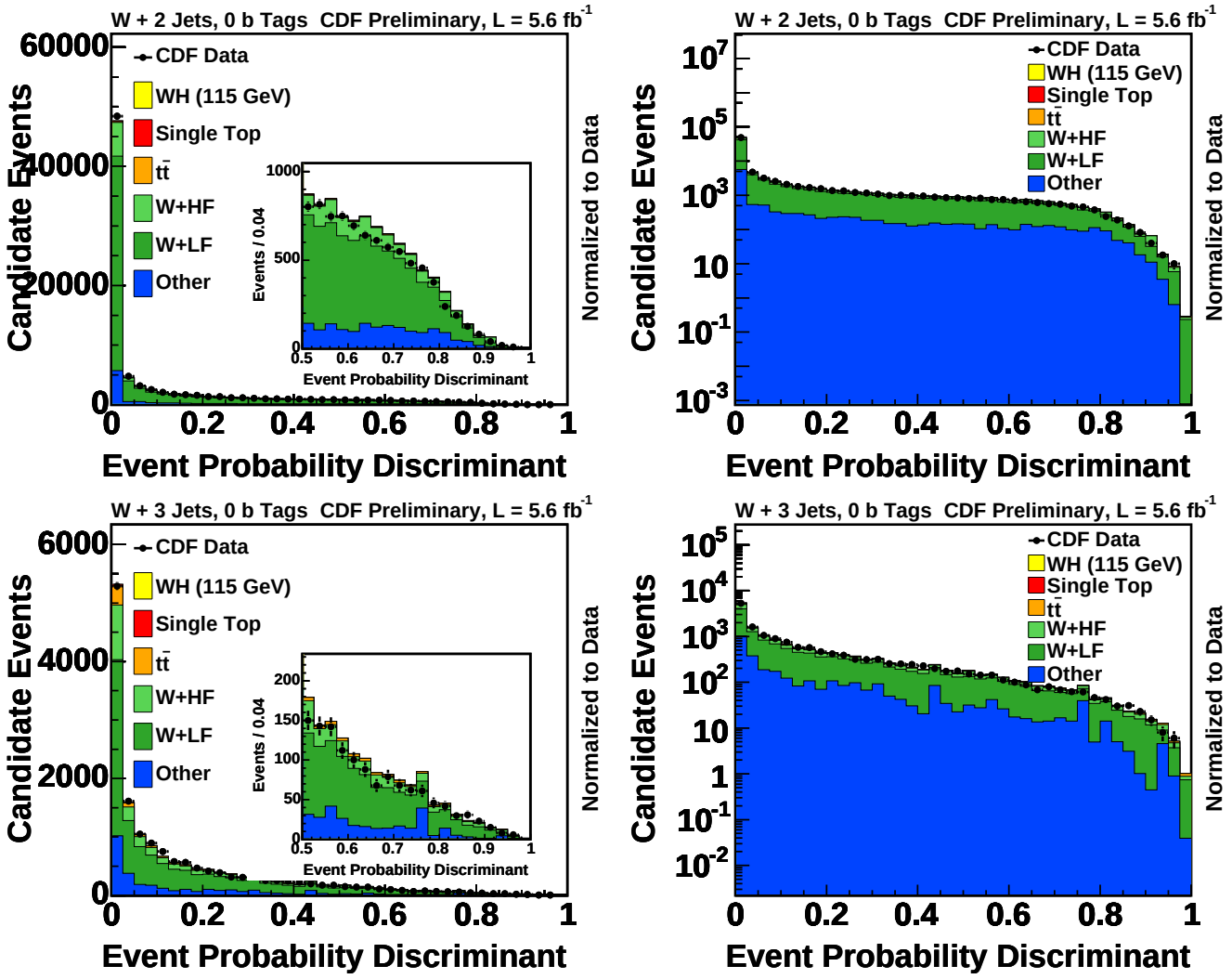

FIG. 2: EPD distributions ($m_H = 115$ GeV/c^2). Top (bottom): EPD applied in the W+2 (3) jets no b-tags sample.

THE LIKELIHOOD FUNCTION

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

We perform a binned likelihood fit to the EPD output. To make it easier to compare the different fit parameters, we define the fit parameter as $\beta_j = \sigma_j^{Fit}/\sigma_j^{SM}$ where β_j is unity when the fit result corresponds to the expected number of events obtained from the independent signal/background estimate:

$$\mathcal{L} = \prod_{j=2}^{6} G_j(\beta_j;\sigma_j) \prod_{k=1}^{B} \frac{e^{-\mu_k}\cdot\mu_k^{n_k}}{n_k!} \tag{7}$$

The Gaussian constraints to the backgrounds are given by:

$$G_j(\beta_j;\sigma_j) = \frac{1}{\sqrt{2\,\pi\cdot\sigma_j^2}} \exp\left[-\frac{1}{2}\cdot\left(\frac{\beta_j - 1.0}{\sigma_j}\right)^2\right] \tag{8}$$

$$\mu_k = \beta_{\rm WH} \cdot T_{jk} + \beta_{\rm single\ top} \cdot T_{jk} + \beta_{\rm W+b\ jets} \cdot T_{jk} + \beta_{\rm W+c\ jets} \cdot T_{jk} + \beta_{\rm mistags} \cdot T_{jk} + \beta_{\rm t\bar{t}} \cdot T_{jk} \quad (9)$$

The index k runs over the bins of the fitted histogram. The template histograms are normalized to the predicted number of events as shown in Table III and IV. This means, $\sum_{k=1}^{B} T_{jk} = N_j^{pred}$.

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

SYSTEMATIC UNCERTAINTIES

Systematic uncertainties can bias the outcome of this analysis and have to be incorporated into the result. We address systematic uncertainties from several different sources: (1) jet energy scale, (2) initial state radiation, (3) final state radiation, (4) parton distribution functions, (5) luminosity, and (6) b-tagging scale factor.

Systematic uncertainties can influence both the expected event yield (normalization) and the shape of the discriminant distribution.

Normalization uncertainties are estimated by recalculating the acceptance using Monte Carlo samples altered due to a specific systematic effect. The WH normalization uncertainty is the difference between the systematically shifted acceptance and the default one and is shown in Table VI.

2 jet					
Channel	**Lepton ID**	**Luminosity**	**$b-$ tagging SF**	**ISR/FSR + PDF**	**JES**
SVnoJP	2%	6%	3.5%	3.1%	2.0%
SVSV-SVJP	2%	6%	8.4%	5.6%	2.0%
3 jet					
Channel	**Lepton ID**	**Luminosity**	**$b-$ tagging SF**	**ISR/FSR + PDF**	**JES**
SVnoJP	2%	6%	3.5%	13.1%	15.8%
SVSV-SVJP	2%	6%	8.4%	21.4%	13.5%

TABLE VI: Rate systematic uncertainties for each channel.

The effect of the uncertainty in the jet energy scale is evaluated by applying jet-energy corrections that describe $\pm 1\sigma$ variations to the default correction factor. Systematic uncertainties due to the modeling of ISR and FSR are obtained from dedicated Monte Carlo samples where the strength of ISR/FSR was increased and decreased in the parton showering to represent $\pm 1\sigma$ variations.

The effect of the b-tagging scale factor and luminosity uncertainty is determined from the background estimate.

For all backgrounds the normalization uncertainties are represented by the uncertainty on the predicted number of background events and are incorporated in the analysis as Gaussian constraints $G(\beta_j|1,\Delta_j)$ in the likelihood function:

$$\mathcal{L}(\beta_1,...,\beta_5;\delta_1,...,\delta_{10}) = \underbrace{\prod_{k=1}^{B} \frac{e^{-\mu_k} \cdot \mu_k^{n_k}}{n_k!}}_{Poisson\ term} \cdot \underbrace{\prod_{j=2}^{5} G(\beta_j|1,\Delta_j)}_{Gauss\ constraints} \cdot \underbrace{\prod_{i=1}^{12} G(\delta_i,0,1)}_{Systematics} \quad (10)$$

$$where,\ \mu_k = \sum_{j=1}^{5} \beta_j \ \cdot \underbrace{\left\{ \prod_{i=1}^{12} [1 + |\delta_i| \cdot (\epsilon_{ji+} H(\delta_i) + \epsilon_{ji-} H(-\delta_i))] \right\}}_{Normalization\ Uncertainty} \quad (11)$$

$$\underbrace{\cdot\ \alpha_{jk}}_{Shape\ P.} \ \cdot \underbrace{\left\{ \prod_{i=1}^{12} (1 + |\delta_i| \cdot (\kappa_{jik+} H(\delta_i) + \kappa_{jik-} H(-\delta_i))) \right\}}_{Shape\ Uncertainty} \quad (12)$$

The systematic normalizations are incorporated into the likelihood as nuisance parameters, conforming with a fully Bayesian treatment [7]. We take the correlation between normalization for a given source into account [8]. The relative strength of a systematic effect due to the source i is parameterized by the nuisance parameter δ_i in the likelihood function, constrained to a unit-width Gaussian (last term in Equation 10). The $\pm 1\sigma$ changes in the normalization of process j due to the i^{th} source of systematic uncertainty are denoted by $\epsilon_{ji}+$ and $\epsilon_{ji}-$ (see Equation part 11). The $\pm 1\sigma$ changes in bin k of the templates for process j due to the i^{th} source of systematic uncertainty are quantified by κ_{jik+}. $H(\delta_i)$ represents the Heaviside function, defined as $H(\delta_i) = 1$ for $\delta_i > 0$ and $H(\delta_i) = 0$ for $\delta_i < 0$. The Heaviside function is used to separate positive and negative systematic shifts (for which we have different normalization). The variable δ_i appears in both the term for the normalization (Equation 11), which is how correlations between both effects are taken into account. We marginalizing the likelihood function by integrating $\mathcal{L}(\beta_1, ..., \beta_N, \delta_1, ..\delta_S)$ over all nuisance parameters for many possible values of the WH cross-section $\beta_1 = \beta_{WH}$. The resulting reduced likelihood $\mathcal{L}(\beta_{WH})$ is a function of the WH cross-section β_{WH} only. We use the MCLIMIT package for our statistical treatment [6].

The event detection efficiency includes uncertainties on the lepton ID, trigger efficiencies and b-tagging scale-factors. The uncertainties on the data derived backgrounds (W+bottom, W+charm, mistags and non-W) are taken from the event yield in Table III and IV.

RESULTS WITH CDF II DATA

We apply the analysis to 5.6 fb^{-1} of CDF Run II data. We compare the EPD output distribution, for a Higgs mass of 115 GeV/c^2, of our candidate events with the sum of predicted WH signal and background distributions as shown in Figures 3 and 4.

In order to extract the most probable WH signal content in the data we perform the maximum likelihood method described before. We perform marginalization using the likelihood function of Equation 10 with all systematic uncertainties included in the likelihood function. The posterior p.d.f is obtained by using Bayes' theorem:

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

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

To set an upper limit on the WH production cross-section, we integrate the posterior probability density to cover 95% [4]. The observed and expected results are shown in Table VII and in Figure 5.

2, 3 jets											
σ / SM	100	105	110	115	120	125	130	135	140	145	150
Expected	2.5	2.7	3.0	3.5	4.4	5.1	6.6	8.7	13.0	17.8	27.5
Observed	2.1	2.6	3.2	3.6	4.6	5.3	8.3	9.2	14.8	18.9	35.3

TABLE VII: Expected and observed upper limit cross sections, in SM units, for different Higgs mass points for 2 and 3 jet events.

SPLITTING THE RESULT IN JET BINS

Table VIII (IX) and Figure 6 left (right) show the expected and observed limits, for each Higgs mass point, for events with 2 (3) jets.

CONCLUSIONS

We have used the matrix element analysis technique in a direct search for Higgs boson production in association with a W boson. To extract the most probable WH content in data, we apply a maximum likelihood

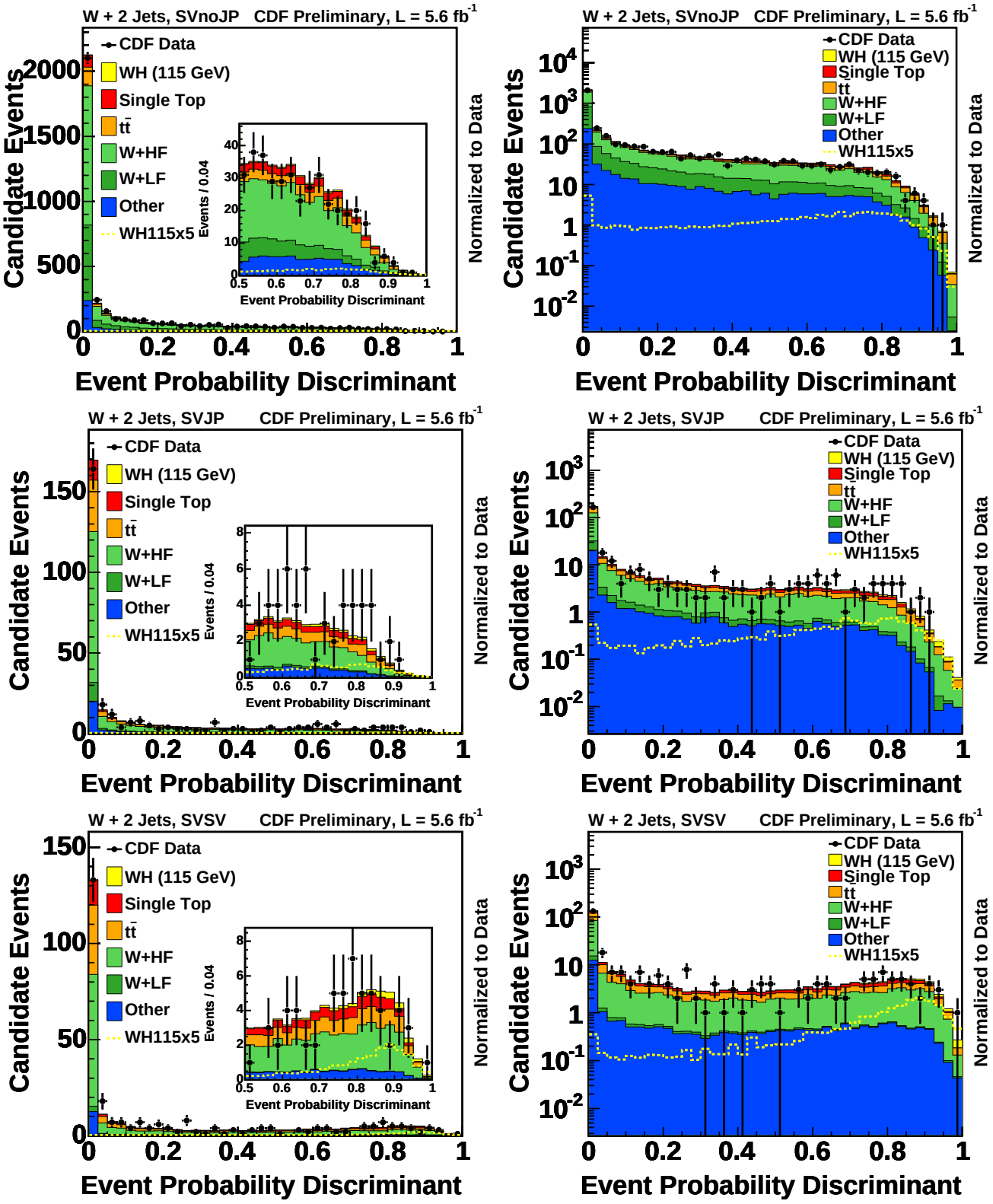

FIG. 3: Comparison of the EPD output for lepton + 2 jets data compared to the Monte Carlo prediction for WH (m_H = 115 GeV/c^2) signal and background. From top to bottom: SVnoJP, SVJP, and SVSV tagged data events, respectively. Note that the signal is twice in these plots, as a stacked plot and as a histo multiplied by 5 (x5).

technique. All sources of systematic rate are included in the likelihood function. We have analyzed 5.6 fb^{-1} of

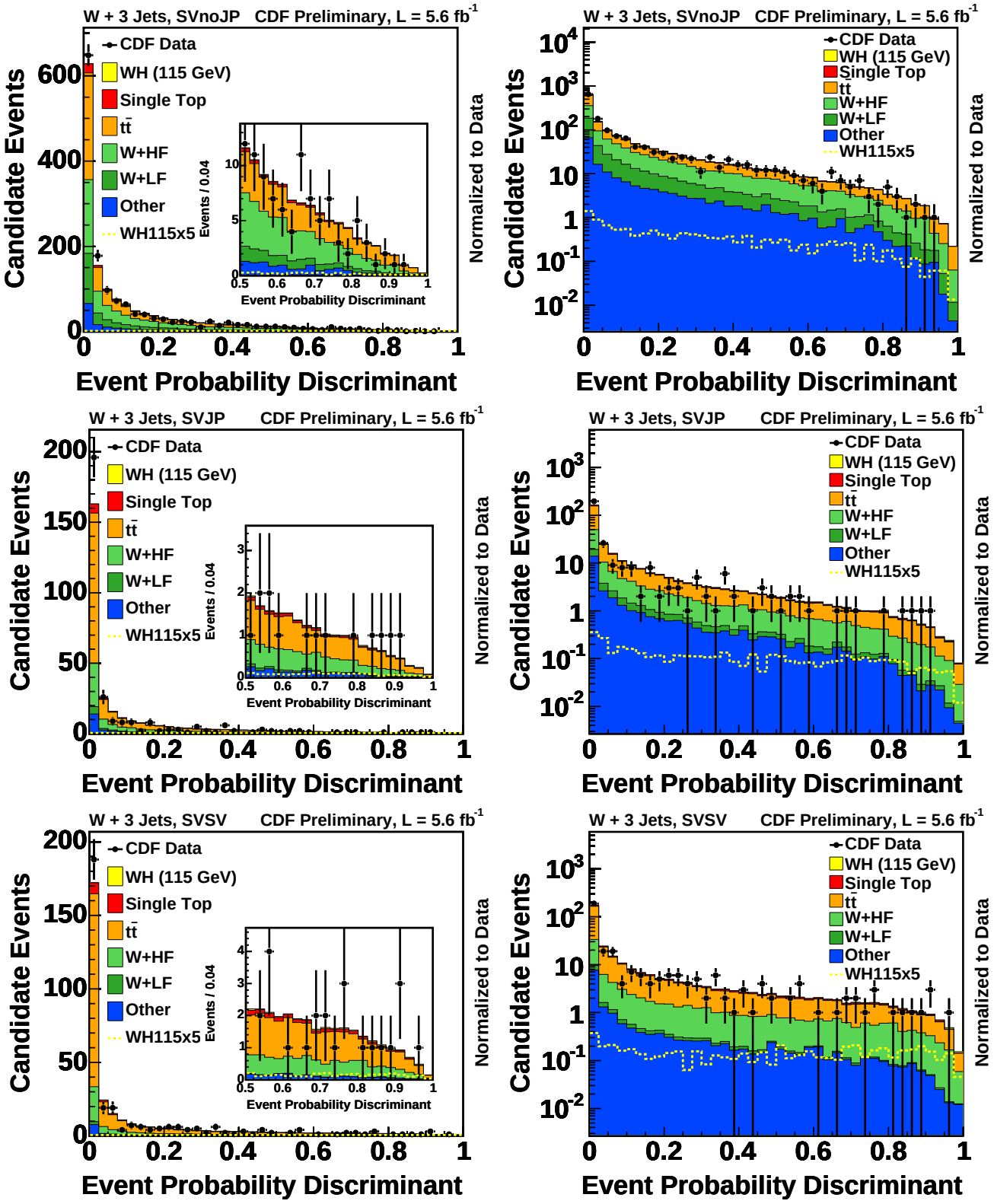

FIG. 4: Comparison of the EPD output for lepton + 3 jets data compared to the Monte Carlo prediction for WH (m_H = 115 GeV/c^2) signal and background. From top to bottom: SVnoJP, SVJP, and SVSV tagged data events, respectively. Note that the signal is twice in these plots, as a stacked plot and as a histo multiplied by 5 (x5).

CDF Run II data. We observe no evidence for a Higgs boson signal and set 95% confidence level upper lim-

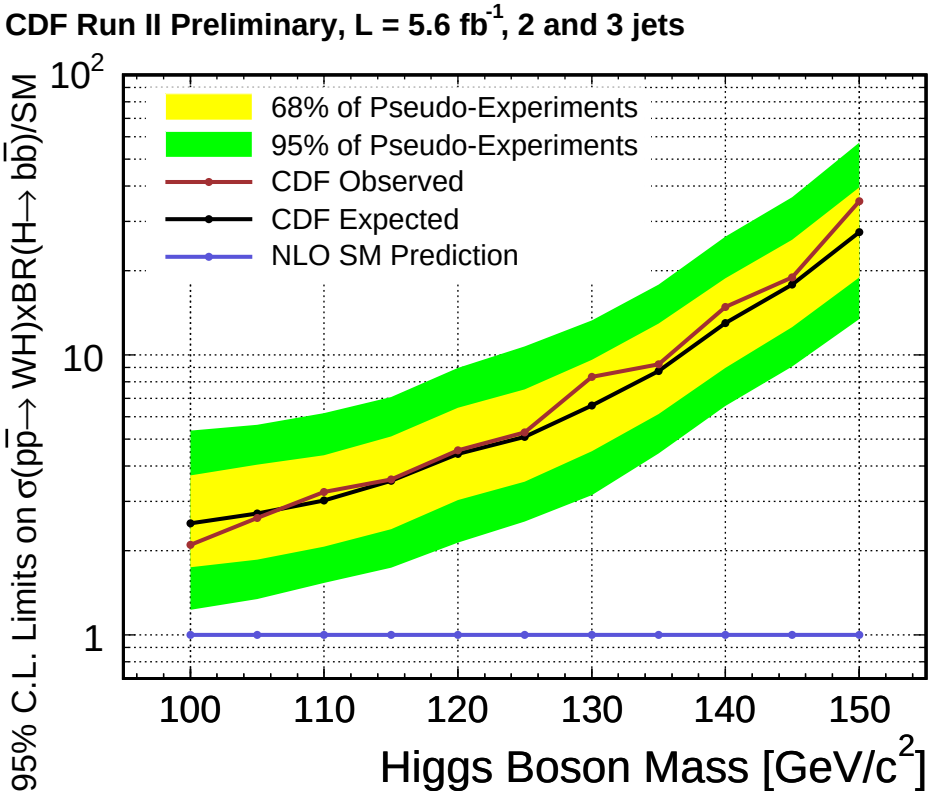

FIG. 5: 95 % C.L. upper limits on the WH production cross-sections times branching ratio for $H \to b\bar{b}$ for Higgs boson masses between $m_H = 100$ GeV/c^2 to $m_H = 150$ GeV/c^2. The plot shows the limit normalized to the predictions from the Standard Model.

2 jets											
σ / SM	100	105	110	115	120	125	130	135	140	145	150
Expected	2.6	2.8	3.2	3.7	4.7	5.5	7.1	9.5	14.2	19.7	30.7
Observed	2.7	3.3	3.7	4.5	5.9	6.8	9.6	12.0	19.3	24.0	43.2

TABLE VIII: Expected and observed upper limit cross sections, in SM units, for different Higgs mass points in the 2 jets channel.

its on the WH production cross section times the branching ratio, in SM units, of the Higgs boson to decay to $b\bar{b}$ pairs of $\sigma(p\bar{p} \to WH) \times BR(H \to b\bar{b})/SM < 2.1$ to 35.3 for Higgs boson masses between $m_H = 100$ GeV/c^2 and $m_H = 150$ GeV/c^2. The expected (median) sensitivity estimated in pseudo experiments is $\sigma(p\bar{p} \to WH) \times BR(H \to b\bar{b})/SM < 2.5$ to 27.5 at 95% C.L.

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the Bundesministerium für Bildung und Forschung, Germany; the Korean Science and Engineering Foundation and the Korean Research Foundation; the Science and Technology Facilities Council and the Royal Society, UK; the Institut National de Physique Nucleaire et Physique des Particules/CNRS; the Russian Foundation for Basic Research; the Ministerio de Ciencia e Innovación-Investigación and Programa Consolider-Ingenio 2010, Spain; the Slovak R&D Agency; and the Academy of Finland.

3 jets											
σ / SM	100	105	110	115	120	125	130	135	140	145	150
Expected	12.2	12.9	13.9	15.8	19.5	23.0	28.1	39.5	56.1	77.9	119.8
Observed	5.1	5.6	8.6	8.5	10.8	12.4	17.3	22.9	33.7	42.5	80.5

TABLE IX: Expected and observed upper limit cross sections, in SM units, for different Higgs mass points in the 3 jets channel.

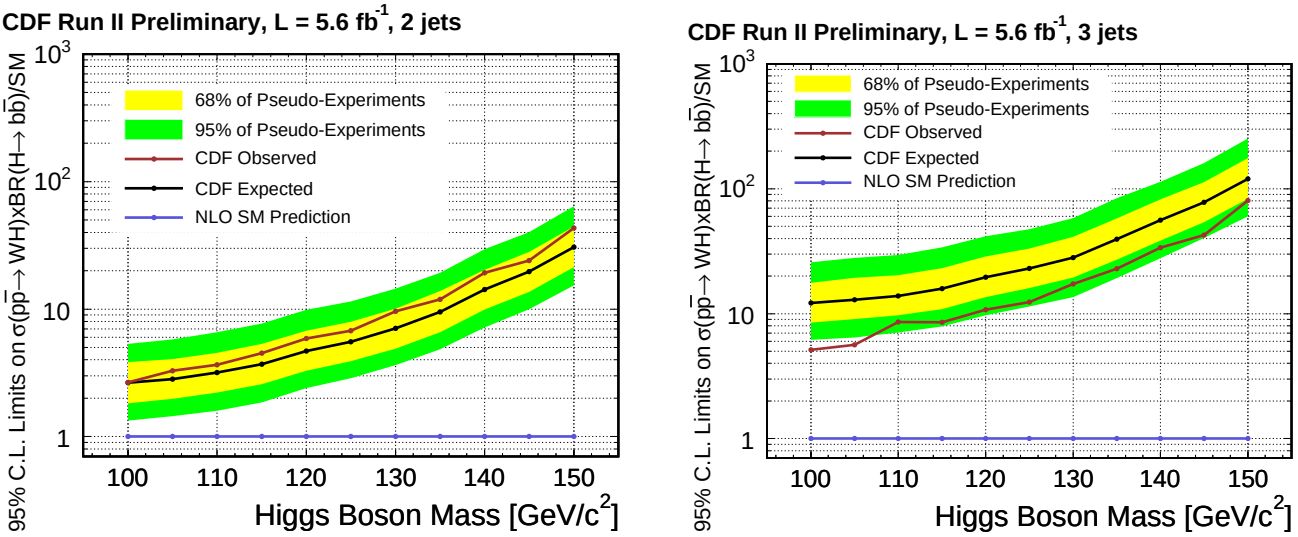

FIG. 6: Left (right): 95 % C.L. upper limits in th 2 (3) jets channel on the WH production cross-sections times branching ratio for $H \to b\bar{b}$ for Higgs boson masses between $m_H = 100$ GeV/c^2 to $m_H = 150$ GeV/c^2. The plot shows the limit normalized to the predictions from the Standard Model.

[1] P. Dong, B. Stelzer, R. Wallny, F. Canelli, *Search for Single Top Quark Production using the Matrix Element Technique*, CDF Note 9117 (2008).

[2] F. Caravaglios *et. al.*, M. L. Mangano *et. al.*, JHEP 0307:001 (2003).

[3] Richter, Svenja, *Search for electroweak single top-quark production with the CDF II experiment*, FERMILAB-THESIS-2007-35, (2007).

[4] Particle Data Group, W.-M. Yao *et al.*, J. Phys. **G33**, 1 (2006).

[5] Konstantin Anikeev, Dmitri Smirnov, *ROOT interface to SNNS*, CDF Note 7999 (2005).

[6] L. Read, J.Phys G **28**, 2693 (2002) and T. Junk, Nucl. Instum. Meth. **434**, 435 (1999). See also P. Bock *et al.* (The LEP Collaborations), CERN-EP-98-046 (1998) and CERN-EP-2000-055 (2000).

[7] L. Demortier, *Bayesian treatments of Systematic Uncertainties*, Proceedings of Advanced Statistical Techniques in Particle Physics, Grey College, Durham, 18 - 22 March 2002.

[8] C. Ciobanu, T. Junk, T. Müller, P. Savard, B. Stelzer, W. Wagner, T. Walter, *Likelihood Function for Single Top Search with 162 pb^{-1}*, CDF Note 7106 (2004).

[9] SumE $= \Sigma \frac{p_{T\ track}}{sin(2 \cdot atan(e^{-\eta_{track}}))}$, $\eta = -\frac{1}{2} \cdot ln(tan\ \theta) \Rightarrow$ SumE $= \Sigma \frac{p_{T\ track}}{sin(\theta)}$.

[10] 16% of the times there are no jets of cone size 0.7 available. In this case, we use the energy of the jet of cone size 0.4.