CDF/ANAL/EXOTIC/PUBLIC/8957

Search for Standard Model Higgs Boson Production in Association with $W^{\pm}$ Boson at CDF with 1.7 fb^{-1}

The CDF Collaboration
URL http://www-cdf.fnal.gov
(Dated: August 13, 2007)

We present a search for Standard Model Higgs boson production in association with a $W^{\pm}$ boson. This search uses data corresponding to an integrated luminosity of 1.7 fb^{-1}. We select events matching the W + jets signature and require both jets to be identified as b-quark jets. To further increase discrimination between signal and background, we use kinematic information in an artificial neural network. The number of tagged events and the resulting neural network output distributions are consistent with the Standard Model expectations, and we set an upper limit on the WH production cross section times branching ratio $\sigma(p\bar{p} \to W^{\pm}H) \times BR(H \to b\bar{b}) < 1.4$ to 1.3 pb for Higgs masses from 110 GeV/c^2 to 150 GeV/c^2 at 95% confidence level.

Preliminary Results for Summer 2007 Conferences

I. INTRODUCTION

The success of the Standard Model in explaining and predicting experimental data provides strong motivation for the existence of a neutral Higgs boson. Current electroweak fits combined with direct searches from LEP2 indicate the mass of the Higgs boson is less than $182\,\mathrm{GeV}/c^2$ at 95% confidence level [1, 2].

In proton-antiproton collisions of $\sqrt{s} = 1.96$ TeV at the Tevatron, the Standard Model Higgs boson may be produced in association with a W boson [3]. For low Higgs masses (below $140\,\mathrm{GeV}/c^2$) the dominant decay mode is $H \to b\bar{b}$. The final state from the WH production is therefore $\ell\nu b\bar{b}$, where the high-p_T lepton from the W decay provides an ideal trigger signature.

The previously published WH search from CDF [4] was performed in a dataset with integrated luminosity equivalent to $320\,\mathrm{pb}^{-1}$. This analysis uses more than 5 times of the previous data and employs a neural network to improve discrimination between signal and background. There is also a CDF preliminary result using an integrated luminosity of $955\,\mathrm{pb}^{-1}$ prepared for the Summer 2006 conferences.

II. DATA SAMPLE & EVENT SELECTION

We use data collected through March 2007, corresponding to an integrated luminosity of $1.7\,\mathrm{fb}^{-1}$. The events are collected by the CDF II detector with high-p_T electron or muon triggers which have an 18 GeV threshold. [5]. The electron or muon is further required offline to be isolated with E_T (or p_T) > 20 GeV.

Events having the W+jets signature are confirmed with a missing transverse energy requirement ($\not\!\!E_T > 20$ GeV). The events are classified according to the number of jets having $E_T > 20$ GeV and $|\eta| < 2.0$. Because the Higgs boson decays to $b\bar{b}$ pairs, we employ b-tagging algorithms which relies on the long lifetime and large mass of the b quark.

A. Bottom Quark Tagging Algorithms

To greatly reduce the backgrounds to this Higgs search, we require that both jets in the event be identified as containing b-quarks by one of two b-tagging algorithms. The secondary vertex tagging algorithm identifies b quarks by fitting tracks displaced from the primary vertex. This method has been used in other Higgs searches and top analyses [4, 6]. In addition, we add jet probability tagging algorithm that identifies b quarks by requiring a low probability that all tracks contained in a jet originated from the primary vertex, based on the track impact parameters [7]. To be considered for this analysis, an event is required to have either two secondary vertex tags, or one secondary vertex tag and one jet probability tag.

B. Total WH Acceptance

The signal acceptance is measured in a sample of Monte Carlo events generated with the PYTHIA program [8]. The detection efficiency for signal events is defined as:

$$\epsilon_{WH\to l\nu b\bar{b}} = \epsilon_{Z0} \cdot \epsilon_{trig} \cdot \epsilon_{\mathrm{l}eptonid} \cdot \epsilon_{\mathrm{iso}} \cdot \epsilon^{MC}_{WH\to\ell\nu b\bar{b}} \cdot \left(\sum_{l'=e,\mu,\tau} Br(W \to \ell\nu) \right), \tag{1}$$

where $\epsilon^{MC}_{WH\to\ell\nu b\bar{b}}$ is the fraction of signal events (with $|z_0| < 60$ cm) which pass the kinematic requirements. The effect of the b-tagging scale factor in this fraction is included by randomly selecting tagged jets. The quantity ϵ_{Z0} is the efficiency of the $|z_0| < 60$cm cut; ϵ_{trig} is the trigger efficiency for high p_T leptons; $\epsilon_{\mathrm{l}eptonid}$, is the efficiency to identify a lepton; ϵ_{iso} is efficiency of the energy isolation cut ; and $Br(W \to \ell\nu)$ is the branching ratio for leptonic W decay.

Fig. 1 shows the overall acceptance for the two double-tag categories–including all systematic effects–as a function of Higgs mass. The acceptance for the double secondary vertex tagged category increases from (0.43±0.05)% for a Higgs mass of $110\,\mathrm{GeV}/c^2$ to (0.51±0.05)% for a Higgs mass of $150\,\mathrm{GeV}/c^2$. The acceptance of the secondary vertex plus jet probability category ranges from (0.36±0.04)% to (0.43±0.05)% over the same mass range. These two categories of double-tagged events are defined exclusively, so the total acceptance is given by the sum of the acceptance for the two categories.

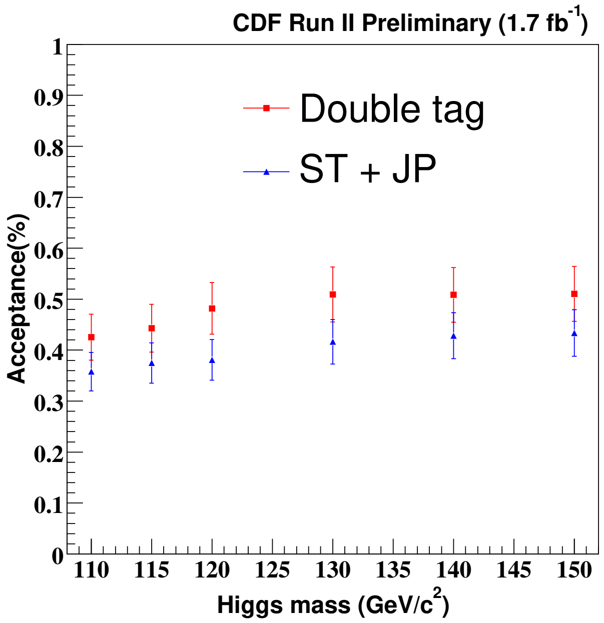

FIG. 1: Calculated WH acceptance for the two b-tagging selection criteria. "Double tag" refers to the double secondary vertex tag category while "ST + JP" refers to the secondary vertex plus jet probability category. These two categories are defined exclusively, so the total acceptance is just the sum of the acceptance for the two categories.

III. BACKGROUNDS

This analysis builds on the method of background estimation detailed in Ref. [6]. In particular, the contributions from the following individual backgrounds are calculated: falsely b-tagged events, W production with heavy flavor quark pairs, QCD events with false W signatures, top quark pair production, and electroweak production (diboson, single top).

We estimate the number of falsely b-tagged events (mistags) by counting the number of negatively-tagged events, that is, events in which the measured displacement of the secondary vertex is opposite the b jet direction. Such negative tags are due to tracking resolution limitations, but they provide a reasonable estimate of the number of false positive tags after a correction for material interactions and long-lived light flavor particles.

The number of events from W + heavy flavor is calculated using information from both data and Monte Carlo programs. We calculate the fraction of W events with associated heavy flavor production in the ALPGEN Monte Carlo program interfaced with the PYTHIA parton shower code [8, 9]. This fraction and the tagging efficiency for such events are applied to the number of events in the original W+jets sample after correcting for the $t\bar{t}$ and electroweak contributions.

We constrain the number of QCD events with false W signatures by assuming the lepton isolation is independent of $\not\!\!E_T$ and measuring the ratio of isolated to non-isolated leptons in a $\not\!\!E_T$ sideband region. The result in the tagged sample can be calculated in two ways: by applying the method directly to the tagged sample, or by estimating the number of non-W QCD events in the pretag sample and applying an average b-tagging rate.

The summary of the background contributions to the double secondary vectex tagged selection is given in Table I, and the summary in the case of secondary vertex plus jet probability tagged events is shown in Table II. Because the expected number of Higgs signal events is small in the 1-,3-, and 4-jet bins, the reasonable agreement between predicted backgrounds and observed data in Fig. 2 gives us confidence in our overall background estimate.

IV. SYSTEMATIC UNCERTAINTIES

The uncertainties on the signal acceptance currently have the largest effect on the Higgs sensitivity. The b-tagging uncertainty is dominated by the uncertainty on the data/MC scale factor $S = 0.95 \pm 0.04$ (stat.+ sys.). The uncertainties due to initial state radiation and final state radiation are estimated by changing the parameters related to

Jet Multiplicity	2-jets	3-jets	>=4-jets
Observed Events (Before b-tagging)	29052	4904	1355
Mistag	3.6± 0.3	2.1±0.2	1.7 ± 0.1
$Wb\bar{b}$	24.5±11.2	9.0±3.6	4.8 ± 2.0
$Wc\bar{c}$	1.9± 0.8	1.0±0.4	0.7 ± 0.3
$t\bar{t}$(6.7pb)	17.3± 2.6	49.5±7.6	85.8 ±13.2
Single top(s-ch)	6.2± 0.9	2.1±0.3	0.6 ± 0.1
Single top(t-ch)	1.5± 0.2	1.3±0.2	0.4 ± 0.1
WW	0.2± 0.02	0.1±0.02	0.1 ± 0.02
WZ	2.2± 0.2	0.6±0.1	0.1 ± 0.02
ZZ	0.1± 0.01	0.1±0.01	0.02± 0.002
$Z->\tau\tau$	0.2± 0.03	0.2±0.03	0.1 ± 0.01
non-W QCD	4.9± 0.9	2.3±0.4	1.0 ± 0.2
Total Background	62.4±12.7	68.0±9.1	95.4 ±13.5
Expected WH signal (120 GeV)	0.9± 0.1	Control region	Control region
Observed Events (Double vertex tagged)	78	76	110

TABLE I: Background summary table for double secondary vertex tag category.

Jet Multiplicity	2-jets	3-jets	>=4-jets
Observed Events (Before b-tagging)	29052	4904	1355
Mistag	11.2± 0.9	8.0 ±0.6	8.6 ± 0.6
$Wb\bar{b}$	20.1± 9.2	7.3 ±3.0	4.3 ± 1.8
$Wc\bar{c}$	5.1± 2.3	2.8 ±1.2	2.0 ± 0.8
$t\bar{t}$(6.7pb)	14.1± 2.2	43.0 ±6.6	72.3 ±11.1
Single top(s-ch)	4.6± 0.7	1.7 ±0.2	0.5 ± 0.1
Single top(t-ch)	1.7± 0.2	1.3 ±0.2	0.4 ± 0.1
WW	0.9± 0.1	0.6 ±0.1	0.4 ± 0.1
WZ	1.7± 0.2	0.5 ±0.1	0.2 ± 0.02
ZZ	0.1± 0.01	0.03±0.004	0.02± 0.002
$Z->\tau\tau$	1.1± 0.2	0.5 ±0.1	0.2 ± 0.03
non-W QCD	8.6± 1.6	4.3 ±0.8	1.7 ± 0.4
Total Background	69.2±12.2	70.1 ±8.4	90.5 ±11.6
Expected WH signal (120 GeV)	0.7± 0.1	Control region	Control region
Observed Events (vertex + jet probability tagged)	81	76	97

TABLE II: Background summary table for the secondary vertex plus jet probability tag category.

ISR and FSR, halving and doubling the default values. The difference from the nominal acceptance is taken as the systematic uncertainty. Other uncertainties on parton distribution functions, trigger efficiencies, or lepton identification contribute to a smaller extent to the overall uncertainty. The summary of these systematic uncertainties on the signal acceptance is given in Table III.

V. ARTIFICIAL NEURAL NETWORK

To further improve signal to background separation we employ an artificial neural network. This neural network combines three kinematic variables into a single function with better discrimination between the Higgs signal and the background processes than any of the variables individually. To train the neural network, JETNET package [10]. The input variables are defined below:

Dijet invariant mass: The invariant mass reconstructed from the two b-tagged jets.

Total System p_T: The vector sum of the transverse momenta of the lepton, the $\not{E}_T$, and the two jets.

p_T **Imbalance:** The scalar sum of the lepton and jet transverse momenta minus the $\not{E}_T$.

The training is defined such that the neural network attempts to produce an output as close to 1.0 as possible for Higgs signal events and as close to 0.0 as possible for background events. For optimal sensitivity, a different neural network is trained for each Higgs mass considered.

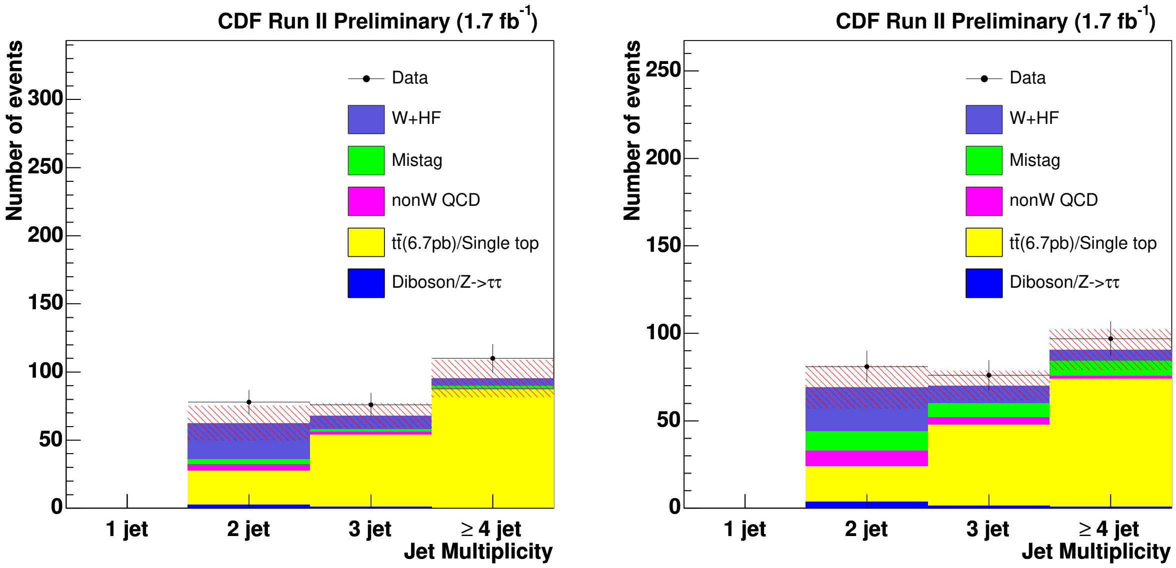

FIG. 2: Predicted and observed W+jet multiplicity with all background contributions. Results are shown for two disjoint selections: double secondary vertex tagged (left) and secondary vertex plus jet probability (right).

Source	Uncertainty (%)	
	ST+ST	ST+JP
Lepton ID	∼2%	∼2%
Trigger	<1%	<1%
ISR/FSR	5.2%	4.0%
PDF	2.1%	1.5%
JES	2.5%	2.8%
b-tagging	8.4%	8.9%
Total	10.6%	10.5%

TABLE III: Systematic uncertainty on the WH acceptance. “ST+ST” refers to double secondary vertex tagged events while “ST+JP” refers to secondary vertex plus jet probability tagged events. Effects of limited Monte Carlo statistics are included in these values.

VI. RESULTS

We perform a direct search for an excess in the signal region of the neural network output distribution from double-tagged W+2 jet events. A binned maximum likelihood technique is used to estimate upper limits on Higgs production by constraining the number of background events to the estimates within uncertainties. For optimal sensitivity, the search is performed simultaneously in the separate double secondary vertex tagged and the secondary vertex plus jet probability tagged samples. The sensitivity in the two double-tagged categories separately and the combined samples is shown in Fig. 3.

For the double secondary vertex tagged events, Fig. 4 shows the output distribution for the neural network trained for a Higgs mass of $120\,\mathrm{GeV}/c^2$ in the data compared to the expectations from background. For reference the dijet invariant mass distribution is also shown. Fig. 5 shows the same distribution for the secondary vertex plus jet probability tagged events. The agreement is reasonable, considering the uncertainties in the background distributions. We set an upper limit on the production cross section times branching ratio as a function of m_H, plotted in Fig. 6. The results are also collected in Table IV.

Higgs Mass	Upper Limit (pb)	
GeV/c^2	Observed	Expected
110	1.4 (8.8)	1.4 (8.6)
115	1.3 (10.1)	1.3 (10.0)
120	1.2 (12.0)	1.2 (11.9)
130	1.2 (18.8)	1.1 (17.4)
140	1.2 (40.0)	1.0 (33.0)
150	1.3 (113.0)	1.0 (80.6)

TABLE IV: Observed 95% C.L. upper limit on $\sigma(p\bar{p} \to WH) \times BR(H \to b\bar{b})$. The number in parenthesis gives the ratio of the upper limit to the SM expectation.

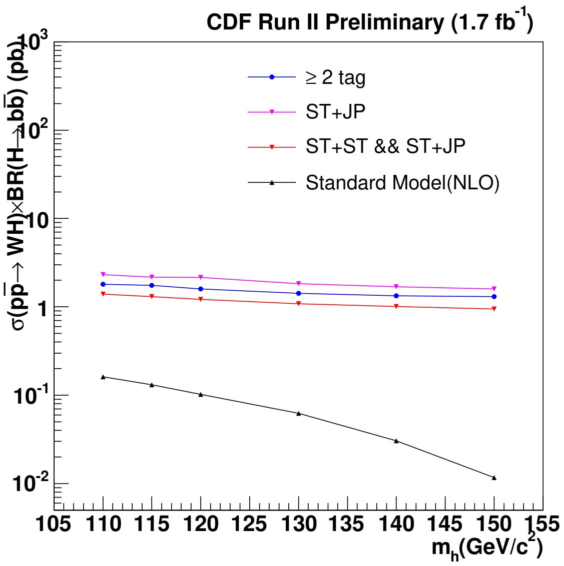

FIG. 3: Expected limits on Higgs production and decay for the separate double-tagged categories and for both categories combined, as a function of the Higgs mass hypothesis. The double secondary vertex tagged category is referred to as "≥ 2 tag, while the secondary vertex plus jet probability category is called "ST+JP." The two categories combined are labeled "ST+ST && ST+JP." The final results combine the two double-tagged categories. The plot on the left shows the expected limit in picobarns. The plot on the right shows the expected limit divided by the SM prediction for Higgs cross section.

VII. CONCLUSIONS

We have searched in a 1.7 fb^{-1} data set for evidence of Standard Model Higgs boson production associated with a W boson. We do not observe any such production in the $H \to b\bar{b}$ mode, and we set upper limits on the production rate times branching ratio. Total rates larger than 1.3 pb are excluded at 95% confidence level for the 115 GeV/c^2 Higgs mass hypothesis, with limits ranging from 1.4 pb to 1.2 pb for other mass values. These limits represent an improvement by a factor of approximately 1.6 over the previous limits obtained using the 955 pb^{-1} data set. The improvement expected only from the increase in luminosity is a factor of 1.3, with the additional 25% coming from improvements in b-tagging, event selection, and background modeling, as well as the addition of a neural network discriminant. Despite these improvements, the result is still limited by the small number of expected Higgs events given the size of the data set and the selection efficiency.

Acknowledgments

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

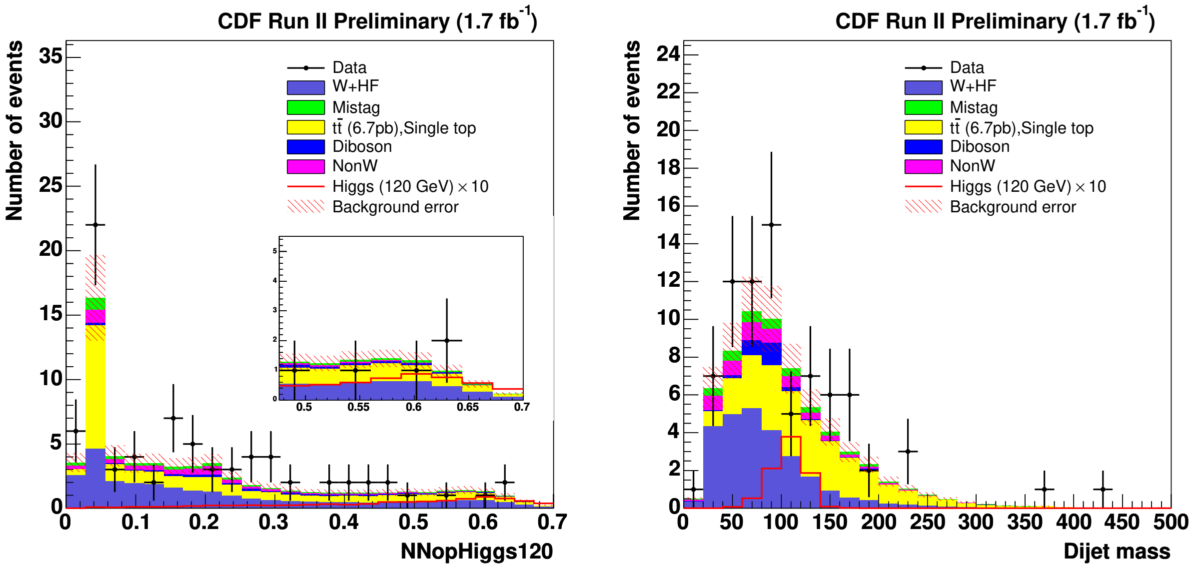

FIG. 4: Predicted and observed output for the neural network trained with a Higgs mass of 120 GeV/c^2 for double secondary vertex tagged events. The output for neural networks trained for other Higgs masses looks similar. For comparison, the dijet mass distribution is also shown.

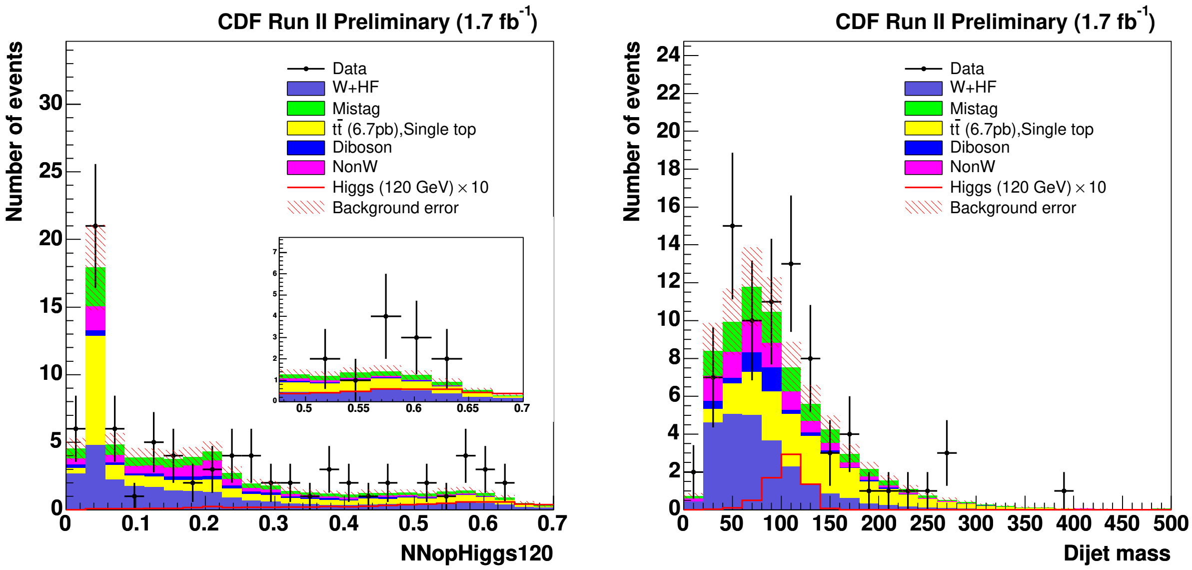

FIG. 5: Predicted and observed output for the neural network trained with a Higgs mass of 120 GeV/c^2 for secondary vertex plus jet probability tagged events. The output for neural networks trained for other Higgs masses looks similar. For comparison, the dijet mass distribution is also shown.

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 Particle Physics and Astronomy Research Council and the Royal Society, UK; the Russian Foundation for Basic Research; the Comision Interministerial de Ciencia y Tecnologia, Spain; and in part by the European Community's

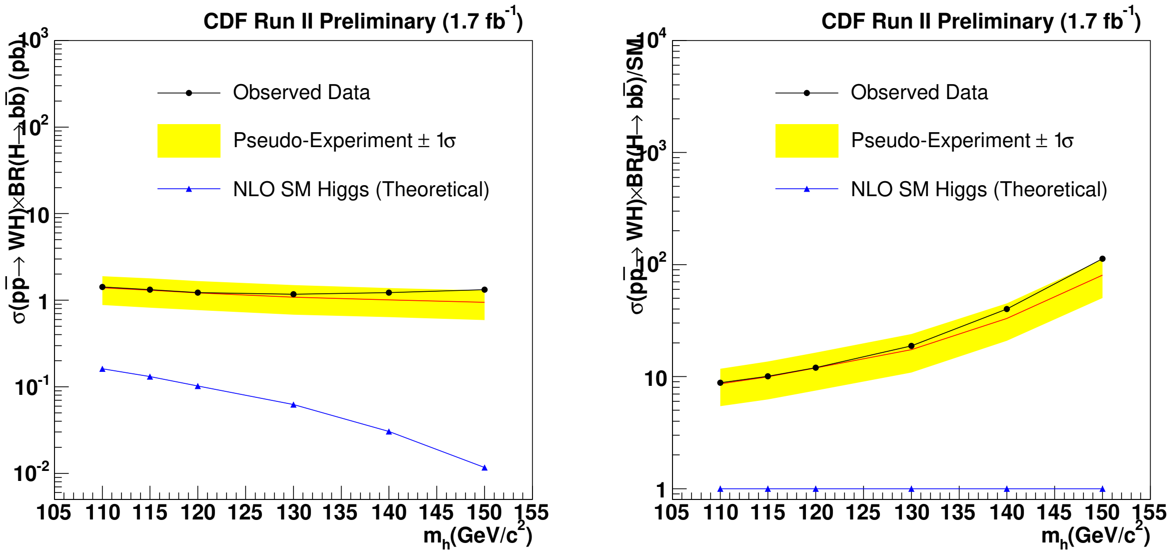

FIG. 6: Observed and predicted rate limits as a function of the Higgs mass hypothesis. These results are based on the combined double tag selections. The plot on the left shows the limits as measured in picobarns. The plot on the right shows the ratio of the limit to the expected SM Higgs cross section.

Human Potential Programme under contract HPRN-CT-20002, Probe for New Physics.

---

[1] LEP Electroweak Working Group, `http://lepewwg.web.cern.ch/LEPEWWG/`
[2] ALEPH, DELPHI, L3, OPAL. The LEP Working Group for Higgs Boson Searches, Phys. Lett **B565** 61 (2003).
[3] T. Han and S. Willenbrock, Phys. Lett. **B273** (1991) 167;
A. Djouadi, J. Kalinowski, and M. Spira, Comp. Phys. Commun. **108 C** (1998) 56.
[4] A. Abulencia *et al.*, Phys. Rev. Lett. **96** 081803 (2006).
[5] F. Abe, *et al.*, Nucl. Instrum. Methods Phys. Res. A **271**, 387 (1988);
D. Amidei, *et al.*, Nucl. Instum. Methods Phys. Res. A **350**, 73 (1994);
F. Abe, *et al.*, Phys. Rev. D **52**, 4784 (1995);
P. Azzi, *et al.*, Nucl. Instrum. Methods Phys. Res. A **360**, 137 (1995);
The CDF II Detector Technical Design Report, Fermilab-Pub-96/390-E.
[6] A. Abulencia *et al.*, Phys. Rev. **D71**, 072005 (2005).
[7] D. Buskulic, *et al.* (ALEPH Collaboration), Phys. Lett. B **313**, 535 (1993);
F. Abe *et al.* (CDF Collaboration), Phys. Rev. D **53**, 1051 (1996);
A. Affolder *et al.* (CDF Collaboration), Phys. Rev. D **64**, 032002 (2001), Erratum-ibid: Phys. Rev. D **67**, 119901 (2003).
[8] P. Edén, C. Friberg, L. Lönnblad, G. Miu, S. Mrenna, E. Norrbin, and T. Sjöstrand, Computer Phys. Commun. **135** (2001) 238.
[9] M.L. Mangano *et al.*, JHEP 0307:001, 2003.
[10] L. Lönnblad, C. Peterson, and T. Rögnvaldsson, Comp. Phys. Comm. 81 (1994) 185.