



$$p\bar{p} \rightarrow t\bar{t}H:$$

## A Discovery mode for the Higgs boson at the Tevatron

J. Goldstein<sup>1</sup>, C. S. Hill<sup>2</sup>, J. Incandela<sup>1</sup>, Stephen Parke<sup>3</sup>, D. Rainwater<sup>3</sup> and D. Stuart<sup>1</sup>

<sup>1</sup>*Particle Physics Division, Fermi National Accelerator Laboratory, Batavia, IL 60510, U.S.A.*

<sup>2</sup>*Department of Physics, University of California, Davis, CA 95616, U.S.A.*

<sup>3</sup>*Theoretical Physics Department, Fermi National Accelerator Laboratory,  
Batavia, IL 60510, U.S.A.*

### Abstract

The production of a Standard Model Higgs boson in association with a top quark pair at the upcoming high luminosity run ( $15 \text{ fb}^{-1}$  integrated luminosity) of the Fermilab Tevatron ( $\sqrt{s} = 2.0 \text{ TeV}$ ) is revisited. For Higgs masses below 140 GeV we demonstrate that the production cross section times branching ratio for  $H \rightarrow b\bar{b}$  decays yields a significant number of events and that this mode is competitive and complementary to the searches using  $p\bar{p} \rightarrow WH, ZH$  associated production. For higher mass Higgs bosons the  $H \rightarrow W^+W^-$  decays are more difficult but have the potential to provide a few spectacular events.

One of the untested features of the elegant and highly successful Standard Model (SM) of electroweak interactions is the Higgs mechanism for spontaneous electroweak symmetry breaking. The Higgs sector of the SM consists of one doublet of complex scalar fields which is used to break

the  $SU(2)\times U(1)$  electro-weak gauge symmetry to  $U(1)$  electromagnetism. This mechanism gives rise to the masses of the  $W$  and  $Z$  bosons and is also used to give masses to the charged leptons and quarks. An additional physical neutral scalar particle is also produced, the undiscovered Higgs boson. All properties of this boson are predicted by the SM except for its mass. Most extensions to the SM's electroweak symmetry breaking mechanism predict more than one physical scalar state, however one of these states tends to have properties very similar to the SM Higgs boson.

Direct collider searches have so far ruled out a SM Higgs boson with a mass less than  $\approx 107$  GeV [1] at the 95% confidence level. The final run of the Large Electron Positron collider (LEP) at CERN this year should be able to extend the exclusion limit or provide a  $5\sigma$  discovery up to about 114 GeV. The next CERN collider, after LEP, is the 14 TeV proton-proton Large Hadron Collider (LHC), which will have the capability to discover a SM Higgs boson of nearly any mass<sup>1</sup> in multiple search channels. However, the LHC is not scheduled to begin taking data until 2005 at the earliest. In the meantime, the upcoming Run II of Fermilab's 2 TeV proton-antiproton collider, the Tevatron, will have an opportunity to discover a Higgs boson above the LEP exclusion limit and below 180 GeV. In this Letter we will concentrate on the SM case over this mass range.

The SM Higgs boson couples at lowest order to all SM particles with a strength proportional to their mass. For the massless photon and gluon a coupling is induced at higher order via loops of heavy particles. Therefore there are many ways for the Higgs boson to decay and also many ways to produce it at a hadron collider. For the mass range of interest here,  $100 \lesssim M_H \lesssim 200$  GeV, the Higgs boson decays predominantly to either a pair of  $b$ -quarks ( $b\bar{b}$ ) for a mass less than 135 GeV, or to four fermions via real or virtual  $W$  gauge bosons for a mass above 135 GeV. These decays are referred to as  $H \rightarrow b\bar{b}$  and  $H \rightarrow W^+W^-$  respectively, and at 135 GeV they have approximately equal branching ratios (43% to 40% respectively). At this mass, the remainder of the decay branching ratio is primarily into  $\tau$  lepton,  $c$  quark and gluon pairs.

---

<sup>1</sup>In the SM  $M_H \gtrsim 1$  TeV is ruled out by unitarity constraints.

Cross sections for various Higgs boson production modes at the Tevatron are shown in Fig. 1(a) [2]. The dominant mode is via gluon-gluon fusion,  $gg \rightarrow H$ , which proceeds via a top quark loop. The next largest mode is Higgs boson associated production with a  $W$  or  $Z$  boson. At a still lower cross-section is Higgs boson production via bremsstrahlung off a top quark pair. The current searches planned for the Tevatron Run II concentrate on the dominant production modes:  $WH, ZH$  associated production with decays  $H \rightarrow b\bar{b}$  [3], and gluon fusion production  $gg \rightarrow H$  with subsequent decay  $H \rightarrow W^+W^-$  [4]. While detailed studies with detector simulation and neural net analysis show promise, they do indicate the need for large integrated luminosity, and even with  $30 \text{ fb}^{-1}$  these searches have difficulty in the mass regions 130-150 GeV and above 180 GeV. In this letter we investigate the feasibility of a search in  $t\bar{t}H$  associated production, first proposed in Refs. [5] in the context of the Superconducting Supercollider. Decays to both  $H \rightarrow b\bar{b}$  and  $H \rightarrow W^+W^-$  are considered, with a varying number of tagged  $b$  quarks and additional jets in the final state and at least one charged lepton<sup>2</sup> to help discriminate against the large QCD backgrounds. We believe that, in spite of the low cross section, a search for the SM Higgs boson in association with a top quark pair in the Tevatron Run II data set ( $15 \text{ fb}^{-1}$  integrated luminosity) is likely to be comparable to the earlier search and discovery of the top quark in the Run I data set ( $0.1 \text{ fb}^{-1}$  integrated luminosity).

We have calculated the signal at the parton level for  $p\bar{p}$  collisions at  $\sqrt{s} = 2.0 \text{ TeV}$  using exact tree-level matrix elements generated by MADGRAPH [6] and COMPHEP [7], and NLO corrected decay rates of the Higgs boson via HDECAY [8]. We take the K-factor to be 1.33: the ratio of the NLO  $t\bar{t}$  cross section to the leading order value. CTEQ4L [9] parton distribution functions

---

<sup>2</sup>Top quarks decay predominantly to a  $W$ -boson and a  $b$ -quark. The  $W$ -boson subsequently decays 2/3 of the time to a pair of quarks (jets) and 1/9 of the time into each of  $e\nu_e, \mu\nu_\mu$  and  $\tau\nu_\tau$ . The neutrinos are not observed and the  $\tau$ -lepton further decays 2/3 of the time to a jet plus neutrinos and 1/3 of the time to  $e$  or  $\mu$  and more neutrinos. Thus, about 26% of the time a  $W$ -boson decays into a taggable charged lepton ( $e$  or  $\mu$ ).

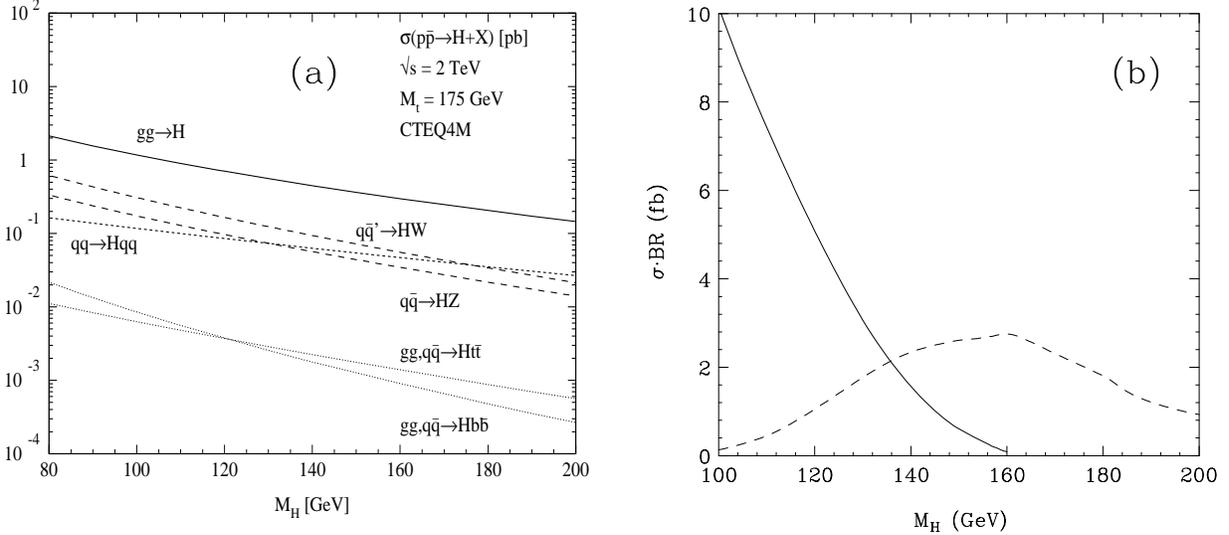


FIG. 1. For  $p\bar{p}$  collisions at  $\sqrt{s} = 2.0$  TeV: (a) cross section (pb) for various Higgs boson production modes [2]; (b) the  $t\bar{t}H$  production cross section times branching ratio (fb) for  $H \rightarrow b\bar{b}$  (solid) and  $H \rightarrow W^+W^-$  (dashed).

are used, and both the factorization and renormalization scales are taken as the top quark mass,  $m_t$ . The cross section times Higgs branching ratio are shown in Fig. 1(b). For  $H \rightarrow b\bar{b}$  the cross section times branching ratio falls steeply and becomes smaller than 1 fb for Higgs masses above 140 GeV. In contrast, the cross section times branching ratio for  $H \rightarrow W^+W^-$  is a broad bump above 1 fb between 125 and 190 GeV. Thus, a search strategy developed for a 160 GeV Higgs would likely work over a much wider mass range, covering some of the unobservable regions of the current searches. These results will be used later as normalization of the generator-level simulations.

For lower Higgs boson masses ( $< 140$  GeV), the most interesting decay mode of the Higgs is  $H \rightarrow b\bar{b}$  and the final state of the hard interaction is  $W^+W^-b\bar{b}b\bar{b}$ . For higher mass Higgs bosons,  $H \rightarrow W^+W^-$  is the dominant decay mode and the final state signature is instead  $W^+W^-W^+W^-b\bar{b}$ . For both modes, each W-boson will decay either hadronically into a quark and anti-quark or leptonically into a charged lepton and neutrino. Backgrounds to these modes consist of resonant production, e.g. irreducible  $t\bar{t}Z$  events with  $Z \rightarrow b\bar{b}$  or  $Z \rightarrow \ell^+\ell^-$ ; continuum  $t\bar{t}b\bar{b}$  or  $t\bar{t}c\bar{c}$  production; and reducible backgrounds such as  $t\bar{t} + jets$ , where the additional jets

TABLE I. Inclusive cross sections for the major physics and reducible backgrounds for  $p\bar{p}$  collisions at  $\sqrt{s} = 2.0$  TeV. Continuum production of final states involving heavy quarks or weak bosons plus additional jets include a  $p_T \geq 20$  GeV cut on the additional jets, but no cut on the decays of heavy states. Typical uncertainties on these estimates are on the order of  $\pm 50\%$ .

backgrounds to $H \rightarrow b\bar{b}$	cross section (fb)	backgrounds to $H \rightarrow W^+W^-$	cross section (fb)
$t\bar{t} + jj$ ( $\Delta R(jj) > 0.4$ )	1030	$t\bar{t} + jj$	1030
$t\bar{t} + b\bar{b}$	27	$t\bar{t} + W$	17
$t\bar{t} + Z, Z \rightarrow b\bar{b}$	1.5	$t\bar{t} + Z, Z \rightarrow \ell^+\ell^-$	0.9
$WZ + jj, Z \rightarrow b\bar{b}, W \rightarrow e\nu, \mu\nu$	10		

are light quarks or gluons but may be misidentified as  $b$  quarks. Table I itemizes the major backgrounds and gives their cross sections. All other SM backgrounds are negligible. Clearly, the largest background arises from  $t\bar{t} + jets$ . Thus, the search for  $t\bar{t}H$  will critically depend on the detector ability to tag  $b$  quark jets and simultaneously to suppress mistagging of non- $b$  jets.

We have calculated the parton level cross section for  $t\bar{t} + jj$  events using the exact tree-level matrix elements [7,10], which contain both collinear and soft singularities for the additional jets. The collinear singularities are removed by imposing the requirement that the final-state partons be well-separated in space,  $\Delta R(jj) > 0.4$ . This corresponds to the experimental requirement that the observed jets be similarly separated.<sup>3</sup> The severity of the soft singularity for a given  $p_T$  cut may be estimated by examining the ratio of the cross sections for  $t\bar{t} + jets$  production to that for  $t\bar{t}$ , the latter of which may be regarded as an inclusive rate; the overall rate can later be normalized to NLO calculations. For the Tevatron we find  $\sigma_{t\bar{t}jj}/\sigma_{t\bar{t}} \approx 1/7$ , including a requirement of  $p_T > 20$  GeV for the additional jets, indicating that the calculation is largely perturbative. This is in line with general expectations that each additional gluon jet of  $p_T \gtrsim 20$  GeV would multiply

---

<sup>3</sup> $\Delta R$  is a separation in the space of detector coordinates azimuth and pseudorapidity,  $\Delta R(jj) = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ .

the inclusive rate by a factor  $\approx 1/2 - 1/3$ . One may also use an exponentiation approximation for the soft gluons [11], which yields a result 30% lower than the value in Table I.

We have compared the results of our matrix element calculation with Pythia 6.115 [12], and while there are differences in cross section as a function of jet  $p_T^{min}$ , the differences at Tevatron energies are largely within the uncertainties of the matrix element calculations. Ultimately, any uncertainty in the rate for additional jets in  $t\bar{t}$  events is not a concern as the  $t\bar{t} + jets$  sample in Run II will be large enough to calibrate the Monte Carlos. Furthermore, by retaining the larger estimate we make this analysis more conservative.

These uncertainties aside, it is clear that a major background could come from the mistagging of non- $b$  jets. Both the efficiency for  $b$  jet tagging and non- $b$  jet suppression will thus be extremely important in this search. Our studies of the upgraded Tevatron collider detectors predict an average  $b$  jet tagging efficiency of about 60% and a  $c$  quark mistag rate of about 25%, while at the same time providing at least a factor of 500 suppression of light quark and gluon jet mistags. Implementation of 3-D vertexing algorithms may improve the tagging efficiency to as high as 70% and decrease the  $c$  mistagging rates to as low as 10%, but we use the more conservative values here.

For a Higgs boson of mass  $M_H \lesssim 140$  GeV, where the  $W^+W^-b\bar{b}b\bar{b}$  signature is preferred, a 60%  $b$  tagging efficiency translates to an 82% probability to tag  $\geq 2$   $b$  jets, 48% for  $\geq 3$  tags, and 13% to tag all four. As an example, consider the case of  $M_H = 120$  GeV, with one  $W$  decaying leptonically and one decaying hadronically, which has much better known backgrounds than the larger all-hadronic  $W$  decay channel. The cross section times branching ratio is 2.0 fb. For three  $b$  tags,  $15 \text{ fb}^{-1}$  of data would contain  $\sim 14$  signal events and of order 110 background events. Requiring four  $b$  tags would yield a signal sample of approximately 4 events, on top of a background of about 25 events. These counts are before any top reconstruction or Higgs boson mass binning.

To extract these few signal events we study their reconstruction by performing a generator level Monte Carlo simulation using Pythia 6.115, which is in good agreement with our parton-level calculations for  $t\bar{t}b\bar{b}$  production using a jet separation corresponding to a parton-level cone

cut of  $\Delta R > 0.7$ . Event selection is determined using a parameterized detector simulation. For reasonable rejection of the  $t\bar{t} + jets$  background, we require at least one isolated lepton and 4 jets with  $p_T > 20$  GeV, at least 3 of which are  $b$  tagged, and two additional jets with  $p_T > 15$  GeV as well as a missing  $E_T > 15$  GeV. Before top quark or Higgs boson reconstruction, kinematically the  $t\bar{t}H$  signal and  $t\bar{t} + jets$  background events are similar. To separate them, the best strategy appears to be a mass peak search in pairs of jets, where at least one jet is  $b$  tagged. Not surprisingly, unless we can re-construct at least one of the top quarks to remove its decay products from consideration in the combinatorics there is no significant excess. However if we assume we can reconstruct the top quark which decays hadronically then there is a pronounced shoulder in the Higgs mass distribution<sup>4</sup>, see Fig. 2(a). Reconstructing both top quarks gives Fig. 2(b). Here the signal exhibits a significant peak above the background. For  $M_H = 120$  GeV, this would correspond to a  $4\sigma$  observation using Poisson statistics converted to a Gaussian equivalent. Of course this assumes 100% reconstruction efficiency for the top quarks. In reality the top quark reconstruction efficiency will be less than 100% but we expect this efficiency to be quite high and reasonably well known early on in Run II. The actual sensitivity to a Higgs mass peak will be somewhere between that shown in Figs. 2(a) and (b) and thus provide substantial sensitivity at 120 GeV.

For higher Higgs boson masses,  $M_H \geq 140$  GeV, the dominant signature is  $W^+W^-W^+W^-b\bar{b}$ . One possible strategy is to demand  $\geq 3$  charged leptons, which is one of the smaller components of this decay mode at  $\sim 6\%$  of the total rate. For  $M_H = 160$  GeV and  $15 \text{ fb}^{-1}$  of data for one experiment, this translates to only 2.2 events of the 36 total  $t\bar{t}H$  events produced. Adding a requirement to see  $\geq 1$   $b$  tag and minimal cuts, the signal sample is only 1 event. The expected background, however, is much less than one event, so there is the potential to observe one or two spectacular events per experiment<sup>5</sup>. This estimate holds approximately over the Higgs boson

---

<sup>4</sup>We are using a conservative 17% dijet mass resolution.

<sup>5</sup>Any background from  $t\bar{t}W$  may be eliminated by requiring the observation of extra hadronic activity

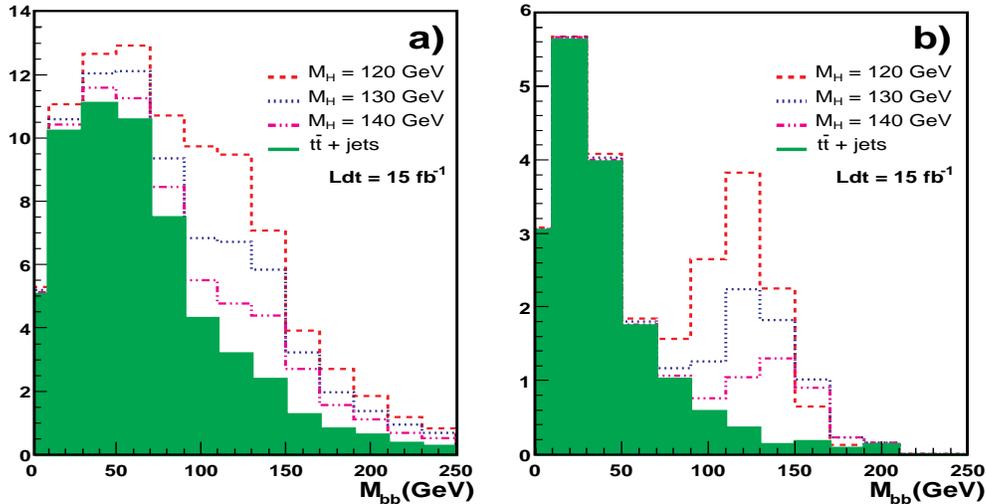


FIG. 2. The  $m_{bb}$  distribution for candidate  $b$ -jet combinations in  $t\bar{t} + jets$  and  $t\bar{t}H$  events for  $p\bar{p}$  collisions at  $\sqrt{s} = 2.0$  TeV using the cuts and selection criteria described in the text: (a) assuming the removal of the decay products from the top quark which decays hadronically, (b) assuming the removal of the decay products from both top quarks. Events are selected using a parameterized detector simulation and correspond to  $15 \text{ fb}^{-1}$  integrated luminosity.

mass range 140-170 GeV.

The signal sample containing exactly two charged leptons is larger, by a factor of 3.6, but there is much more background. Requiring the two leptons to have the same charge reduces the background dramatically, to of order 1-2 events if additional jets are required to be observed. Further mass constraints may make this much smaller than one. However, only 1/3 of the two lepton signal sample will have same-sign leptons, one expects another 2-3 spectacular events per experiment.

Despite the low signal rate for  $t\bar{t}H$  associated production events, there is considerable potential in this channel to observe a SM Higgs boson, or to set exclusion limits on its mass. We have not examined cases of non-SM Higgs bosons, but the outlook is optimistic as production cross sections times branching ratio are often enhanced relative to the SM, e.g. in SUSY or

---

from the fourth signal  $W$ -boson.

topcolor models. Our analysis suggests an optimistic search strategy for Higgs bosons up to about 140 GeV in the  $H \rightarrow b\bar{b}$  channel, and some promise in the  $H \rightarrow W^+W^-$  channel for the mass range 140-170 GeV. A thorough understanding of and improvements to the top quark reconstruction efficiency and the identification of  $b$ ,  $c$  and light quark jets are very important for this search. Since the backgrounds to these signals are very small, increasing the integrated luminosity of the Tevatron by a factor of 2 would increase the significance of this search to more than  $5\sigma$ .

### ACKNOWLEDGMENTS

We would like to thank U. Baur, E. Eichten, K. Ellis, C. T. Hill, M. Mangano and C. Quigg for useful discussions. Fermilab is operated by URA under DOE contract No. DE-AC02-76CH03000.

## REFERENCES

- [1] Aleph/Delphi/L3/Opal combined results note, ALEPH 2000-028, CONF 2000-023.
- [2] J. D. Hobbs, hep-ph/9903494.
- [3] A. Stange, W. Marciano and S. Willenbrock, Phys. Rev. **D49**, 1354 (1994) [hep-ph/9309294].
- [4] T. Han and R. Zhang, Phys. Rev. Lett. **82**, 25 (1999) [hep-ph/9807424].
- [5] W. J. Marciano and F. E. Paige, Phys. Rev. Lett. **66**, 2433 (1991); J. F. Gunion, Phys. Lett. **B261**, 510 (1991).
- [6] T. Stelzer and W. F. Long, Comp. Phys. Comm. **81**, 357 (1994).
- [7] COMPHEP collaboration, A. Pukhov *et al.*, hep-ph/9908288.
- [8] A. Djouadi, J. Kalinowski and M. Spira, Comput. Phys. Commun. **108**, 56 (1998).
- [9] H. L. Lai *et al.*, Phys. Rev. **D55**, 1280 (1997).
- [10] A. Stange, private communication.
- [11] V. Barger, R.J.N. Phillips, and D. Zeppenfeld, Phys. Lett. **B346**, 106 (1995); D. Rainwater, R. Szalapski, and D. Zeppenfeld, Phys. Rev. **D54**, 6680 (1996); D. Rainwater, D. Summers, and D. Zeppenfeld, Phys. Rev. **D55**, 5681 (1997).
- [12] T. Sjostrand, Comput. Phys. Commun. **82**, 74 (1994); and references therein.