Fermilab FERMILAB-FN-0723 August 2002

Update study of low mass Higgs using $pp \to qqH$ at CMS

N. Akchurin[a], D. Green[b], S. Kunori[c], R. Vidal[b], W. Wu[b*], M. T. Zeyrek[d]

[a]*Texas Tech University, Lubbock, Texas, U.S.A.*
[b]*Fermi National Acceleration Laboratory, Batavia, Illinois, U.S.A.*
[c]*University of Maryland, College Park, Maryland, U.S.A.*
[d]*Middle East Technical University, Ankara, Turkey*

Abstract

In our previous note, we reported on a preliminary study of the Standard Model Higgs Boson ($m_H = 120\,\mathrm{GeV}$) produced via the vector boson fusion channel. In this study, the Higgs boson decays to a W^+W^- pair, with the subsequent decay of the two W's to $l^+l^-\nu\bar{\nu}$. The production of a Higgs boson via the vector boson fusion channel is characterized by two jets at large rapidity. Thus, we looked for two jets at large angles, two electrons or two muons, and missing transverse momentum in the final state. Furthermore, to reduce contributions from background processes, we rejected events with jets in the central region.

In this note, we continue this study and report some new additional information. First, we have evaluated qqZ as a another possible background to the signal qqH. Second, we show how to use qqZ to calibrate the measurement of the cross-section of the signal qqH. And last, we also include the τ in the decay of the two W's, with the τ's decaying into electron or muon.

*) send comments/questions to weimin@fnal.gov

1 Introduction

One of the main objectives of the CMS experiment is to discover or to rule out the Standard Model Higgs particle. The LEP II experiments have set a lower bound of about 113 GeV. Since a low Higgs mass is preferred by supersymmetry, the low mass Higgs search is presently the focus of intensive theoretical and experimental studies at the LHC. The Weak Boson Fusion (WBF) process has emerged as potentially one channel where a light or medium mass Higgs particle might be first discovered. In addition, this production mechanism depends only on the defining HWW coupling [2] [3], unlike the gluon fusion process which depends to a large extent on the uncertain Higgs-top coupling.

In this and previous studies [1], we investigate the Higgs decaying to a W^+W^- pair, and the subsequent decay of these W's to $l^+l^-\nu\bar{\nu}$. For now, we consider only a light Higgs ($m_H = 120$ GeV). The production and decay are:

$$pp \to jjH \quad (qq \to q\prime q\prime H), \quad H \to W^{(*)}W \to l^+l^-\nu\bar{\nu} \tag{1}$$

where the leptons, $l^\pm$, are either electrons or muons produced, either directly, or via τ leptonic decay. The distinct feature of this process is that the forward and backward jets tend to preserve the initial parton direction, since color exchange is absence in the t-channel. There are energetic forward jets with significant transverse momentum ($p_T \approx m_W/2$) and suppressed hadronic activity in the central region. We refer to these jets as the forward “tagging jets”. [1].

The largest background for the WBF channel is from $t\bar{t}$ production [4]. Since we are looking for a final state with $l^+l^-\nu\bar{\nu} + jj$, production of $t\bar{t} + jets$, where each top decays to $l\nu b$, might possibly mimic the WBF Higgs production. Suppression of the $t\bar{t} + jets$ background is achieved mainly by requiring jets in the opposite forward regions, and vetoing events with extra jet activity in the central region.

In addition to $t\bar{t}j$, other backgrounds processes with equally complex topologies require close attention. Chiefly, they are QCD production of $WWjj$, electro-weak production of $WWjj$, and qqZ. As noted in references[4], the contribution from QCD and electro-weak production of $WWjj$ are only about half of the $t\bar{t} + jets$ background. In this note, we postpone inclusion of this background, and focus instead on the qqZ case. In the qqZ case, the Z decay can leptonically (including the important τ channel), so the qqH and qqZ can have a very similar event topology.

A study of the qqZ background has additional importance because it can be used as a calibration of the signal. The absence of a sharp mass peak means that we rely on a good knowledge of the cross section normalization. The production mechanism of qqZ is very similar to the production process for qqH. Since WZ events will be plentiful at the LHC, and since the cross section depends only on the WWZ coupling, we anticipate that this coupling will be well measured at the LHC, and the production cross-section for qqZ reliably predicted. Even though our analysis selection procedure may introduce some systematic errors, we expect that by correcting for relative the acceptances of qqH and qqZ events after the analysis cuts, the ratio of observed events can be used to calibrate the qqH cross section from the expected qqZ cross section. In section 4 of this note, we investigate this calibration procedure.

In our previous study [1], we omitted the τ channel because it contributes only a small amount to the signal and the backgrounds studied. Only one ninth of the W's decay to tau's, and the tau branching fraction to electrons and muons is only $17 - 18\%$. Also, the electrons and muons from tau decays usually have smaller E_T, compared with the electrons and muons coming directly from a W or Z decay, and are more likely to fall below our lepton cuts. In the qqZ case, many of the events where the Z decays directly to an electron or muon pair will fail our previous cuts. Specifically, many are removed by the di-lepton mass cut. Since the di-lepton mass cut is also needed to suppress other backgrounds, to simulate the correct number of qqZ events passing our cuts, we must include the tau decays in our analysis. A significant fraction of the tau decays will survive the di-lepton mass cut. Therefore, we now include the tau decays in both qqH and qqZ in our study.

2 Event Generation and Reconstruction

The signal and background events are generated using the CompHEP [5] parton level matrix elements. The complete matrix element calculation method is used in the generation of signal and background events because this method represents the three body final state in the process $qq \to qqH$ correctly.

CompHEP produces cross sections with the proper phase space weighting. This information is stored in a special data base, called PEVLIB. CompHEP generated events are then interfaced to PYTHIA to produce detectable final states through hadronization and decay, with Initial and Final State Radiation (ISR and FSR) turned on. Thus, we expect that radiation from the external lines of the Feynman diagrams of CompHEP has been treated properly.

The qqH, qqZ and some of the $t\bar{t}j$ events were generated by Slava Ilyin, using CompHEP41.10 and interface46 to Pythia6.1. We have used a filter to select the events with the final state $l^+l^-\nu\bar{\nu} + jj$ from the unconstrained W decays. The remaining $t\bar{t}j$ event were generated by Shuichi Kunori, using the "patriot" framework, which interfaces to Pythia6.2. In the "patriot" frame, the ISR and FSR parameters were set by using the CompHEP file (except that the user must set the scale according to the physics process). After careful checking of various distributions, we concluded that there is no significant differences between the two $t\bar{t}j$ samples. Therefore, we combined them in the analysis with proper normalization.

Table 1 shows the number of events and cross sections before analysis cuts. The number of events refers to events containing ee, $\mu\mu$ and $e\mu$ final states from the W decays. The cross section σ are taken from the calculation by CompHEP, and are not the full inclusive cross sections. The CompHEP cross sections include an initial set of jet cuts at the generation level: $p_T > 15$ GeV, $|\eta| < 5$ and $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} > 0.5$ separation between jets. These cuts represent weak detection criteria for reconstructible jets within the CMS angular acceptance. The value for $\sigma \cdot B$ represents the charged lepton ee, $\mu\mu$ and $e\mu$ final states. For the signal, the branching ratio of $H \to WW$ is also included.

Table 1: Cross sections and the event generation

	Signal	$t\bar{t}j$	qqZ
No. of events	6537	56736	25786
σ	2104 fb	788000 fb	3910 fb
$\sigma \cdot B$	16.9 fb	52796 fb	273 fb

3 Basic event selection

The basic event selection criteria came from the cuts previously proposed in [4], and have been described [1] in our previous note. We will not repeat the details here, nor show the specific figures and distributions which lead us to choose these cuts for event selection and background suppression.

The event selection followed six steps and is summarized below:

- The selection criteria for the forward tagged jets, j_1 and j_2, are given in (2) and (3). We require that each jet have a minimum transverse momentum of $20\ \mathrm{GeV}$ and a pseudo-rapidity well within the CMS calorimeter. In addition, we require a minimum ΔR between any two jets, a large separation in pseudo-rapidity between the two tagged jets, and that the two jets be in opposite hemispheres. We require:

$$p_{Tj} \geq 20\,\mathrm{GeV}, \quad |\eta_j| \leq 4.5, \quad \Delta R_{jj} \geq 0.6 \tag{2}$$

$$\Delta\eta_{jets} = |\eta_{j1} - \eta_{j2}| \geq 4.2, \quad \eta_{j1}.\eta_{j2} < 0 \tag{3}$$

- The invariant mass of the two forward jets is large. We require:

$$m_{jj} > 600\,\mathrm{GeV} \tag{4}$$

- The final state W's are produced almost at rest in the decay of a 120 GeV Higgs. Since the Higgs is a scalar and the W is a spin-one boson, the two charged leptons often are emitted back-to-back to the two neutrinos in the rest frame of the Higgs. The invariant mass of the two leptons and the invariant mass of the two neutrinos are almost equal and kinematically can not exceed $M_H/2$. The invariant mass of the two charged leptons is thus small. We require:

$$m_{ll} < 60\,\mathrm{GeV} \tag{5}$$

- Many background events that survive previous cuts have considerable extra jet activity in the central region, In most of these background events, one of the forward jets originated from light quarks or gluons and the other came from a b-quark from t-decay. Usually, the second b-quark from the other t-decay ends up in the central rapidity region. For the signal, any additional jet activity from QCD is usually FSR along the forward jets.

  Since we did not simulate the CMS tracking systems or CMS b-tagger in this study because of limited computer resources and time, we identified central b-quarks in the $t\bar{t}j$ background events by matching the generated b-quarks in the central region ($p_T > 20$ GeV) with the jets found in the simulation, excluding the two forward tagged jets. The ΔR is calculated between those jets and the generated b-quarks, and a cut of $\Delta R < 0.4$ is applied to identify these central b-quarks. Events are removed that have a central b jet. Since the efficiency of the CMS b-tagger for b-quark jets with $p_T > 20$ GeV is expected to be quite good, this should be sufficient for this analysis, although the rejection is better than we could obtain with a realistic CMS b-tagger.

  There may still be some central jet activity from QCD radiation. From color coherence between the initial and final state quarks in the signal, we expect most of the gluon radiation that does occur to be in the forward direction. In contrast, for $t\bar{t}j$ background events we expect most of the gluon radiation to be in the central region. Removing central b jets is therefore very powerful in suppressing the $t\bar{t}j$ background.

- Because the longitudinal momenta of the two neutrinos from the W decays cannot be determined, the invariant mass of the Higgs boson cannot be reconstructed uniquely. A transverse mass of the WW can be reconstructed using:

$$m_T(WW) = \sqrt{(E_T^{miss} + E_{T,ll})^2 - (\vec{p}_T^{\,miss} + \vec{p}_{T,ll})^2}. \tag{6}$$

  The parameters $\vec{p}_T^{\,miss}$ and $\vec{p}_{T,ll}$ are the transverse momentum vectors of the missing momentum and the two charged leptons, respectively. For the $t\bar{t}j$ background, after the forward jet and lepton selection (2,3,4 and 5), this transverse mass reconstruction gives a broad distribution that peaks around the top quark mass, distinctively different from the signal. We limit the transverse mass for the WW to be near the Higgs mass.

$$50\,\mathrm{GeV} < m_T(WW) < 140\,\mathrm{GeV} \tag{7}$$

4 Calibration of qqH Cross Section using qqZ

As we all know, the number of events that survive after detection and analysis is given by:

$$N = f.\sigma \tag{8}$$

The quantity f is the detector efficiency including the trigger, selection cuts, systematics error etc.. The cross section is σ which here includes the branching ratio. N is the number of events remaining after complete detection and analysis procedure. If we assume that the detection efficiency "f" is about the same for the qqZ and qqH events, then the observed number of events and predicted cross section for qqZ can be used to determine the absolute cross section qqH. However, in reality, there are differences in detection efficiency for qqZ and qqH. For example, qqZ events more likely will pass the cut on m_{ll} only if the Z first decays to τ, which then decays leptonically. The plot of m_{ll} for Z's decaying to τ is showing Fig.1. Thus, including the separate detector efficiencies for qqH and qqZ, we get:

$$\sigma^{qqH} = (A^{qqZ}.N^{qqH}/A^{qqH}.N^{qqZ}).\sigma^{qqZ} \tag{9}$$

here, A^{qqZ} is the acceptance for qqZ, A^{qqH} is the acceptance for qqH, N^{qqH} is the number of events surviving for qqH after all cuts, while N^{qqZ} is the number of events surviving for qqZ after all the same cuts.

From Table 2, we show how the event reduction after each cut. The overall acceptance for the qqH is 18.9% (1235/6537), while for the qqZ, it is 3.34% (862/25786). The integrated luminosity corresponding to these events for the qqH is 386.8 fb^{-1}, and for qqZ is 94.45 fb^{-1}. As a check, if we normalize the number of events to the same integrated luminosity of the signal for qqZ, the number of events surviving detection and analysis for qqZ is 3530. Thus, we find:

$$\sigma^{qqH} \quad = \quad (1235 * 0.0334/3530 * 0.189) * 273\ fb \quad = \quad 16.8\ fb$$

compared to 16.9 fb found in the Table 1. Note that we can find qqZ events in the dilepton triggers with similar cuts except for the m_{ll} cut. These events will allows us to predict the qqZ background in the qqH search.

5 Analysis results

After the first year of operation of the LHC (assuming low luminosity running conditions) an accumulated luminosity of 60 fb^{-1} at CMS would results in 193 signal events (1235*60/386) and 550 qqZ background events (862*60/94), and 391 $t\bar{t}j$ background events (7*60/1.075). Thus the signal to background fluctuation $\sqrt{\mathrm{B}}$ is

$$\mathrm{S}/\sqrt{\mathrm{B}} = 6.3.$$

Note that a good understanding of the detection efficiency is necessary in this search. The detection of qqZ in dilepton triggers with a Z mass peak serves to normalize the qqZ events which appear as background in the low m_{ll} region arising from τ decays. The events reduction factor due to the analysis cuts are very similar compare to our previous study. For qqH, the acceptance was 17.6 $\%$ versus 18.9 $\%$ found here. For ttj events, the acceptance was 0.019 $\%$ versus 0.013 $\%$ here.

Table 2: Cross sections and selection efficiencies

	Signal	$t\bar{t}j$	qqZ
$Number of events$	6537	56736	25786
Forward Jet Tagging (Eq. 2,3)	5595	4466	21158
m_{jj} (Eq. 4)	2570	706	16566
m_{ll} cuts (Eq. 5)	2449	315	2338
b jet (from top decay) veto	2242	22	2145
extra-jet veto $p_T > 20$ GeV	1910	15	1569
$M_T(WW)$ cut (Eq. 7)	1235	7	862

6 Conclusions

With forward tagged jets, Higgs production through the WBF process gives a very clear signal topology. In addition, exploiting the unique kinematics of the isolated leptons between the jets further enhances the experimental signature of this channel. Further improvement in the signal to background can be achieved by the suppression of events with additional central jets.

Since it is near the lowest mass excluded experimentally, we chose to study a low mass Higgs, $m_H = 120$ GeV, even though it is the most difficult case. The calibration of qqH cross section using the qqZ will allow us to compare to theory which depends only on HWW coupling.

As we mentioned earlier, the suppression of $t\bar{t}j$ background is enhanced greatly with efficient b jet vetoing. The method described here to reject the central b-jets that come from t-decay uses the correlation of extra jets with b-quarks instead of tagging b's, and it is not conclusive. Good tracking and good lepton identification efficiency are also essential. Until we are able to use realistic lepton and b-tagging efficiencies, the results of this study are not compelling. Nevertheless, this WBF channel, with its distinct characteristics is one of the most promising channels for detecting a low mass Standard Model Higgs particle with CMS at the LHC. In particular, the WBF process with subsequent WW Higgs decay is theoretically very well understood and predicted.

7 Acknowledgments

We would like to thank V. A. Ilyin for his CompHEP event generation. We also thank D. Rainwater and Steve Mrenna for their valuable help during this study.

References

[1] Study of low mass Higgs using $pp \to qqH$ at CMS, FERMILAB-FN-714, Feb., 2001

[2] CDF and D0, Higgs Search at the Tevatron. FERMILAB-Conf-99/053-E

[3] Esther Ferrer Ribas, Standard Model Higgs at LEP. LAL 00-50, Sep. 2000

[4] N. Kauer, T. Plehn, D. Rainwater, D. Zeppenfeld, Physics Letters B 503 (2001) 113-120.

[5] A. Pukhov et. al., Preprint INP MSU 98-41/541, hep-ph/9908288.

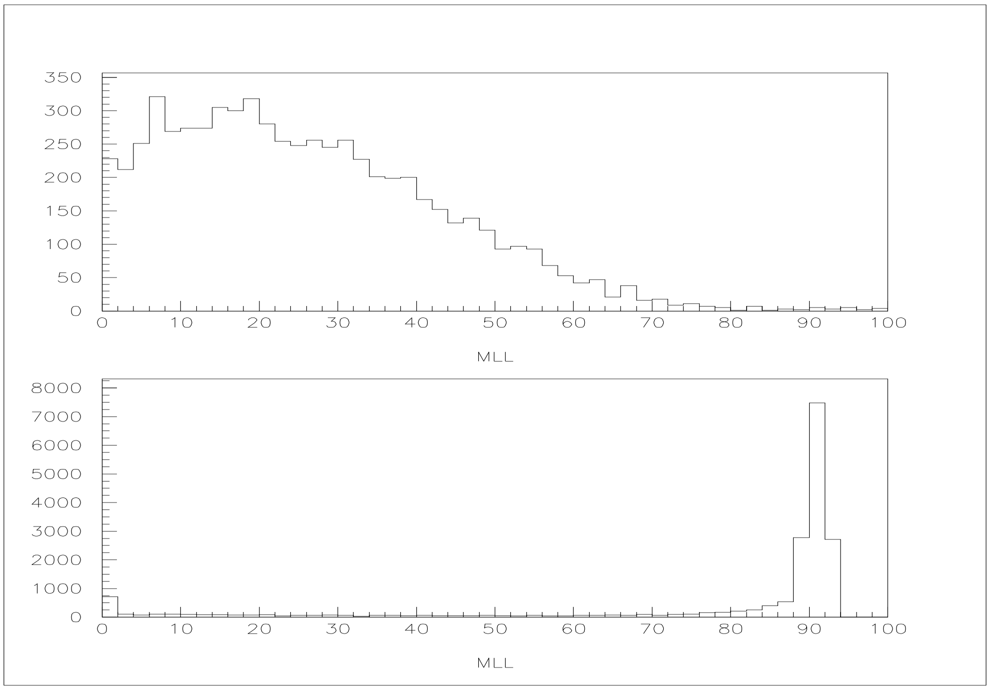

Figure 1: Invariant mass distribution of charged leptons after jet cuts (Eqs. 2-4) for the signal qqH(up) and the qqZ(down). Because we include the τ decay, so there is long tail in lower side for qqZ. The major contribution is the electron and muon directly from the Z, so the peak is well site at the Z mass, low side tail reflects the contribution mainly from the τ