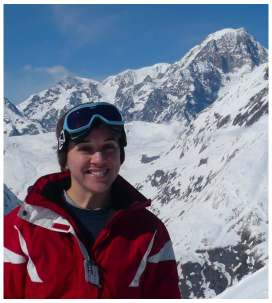
SEARCH FOR SM HIGGS BOSON IN THE $\tau^+\tau^-$+ 2 JETS FINAL STATE

Kathryn Tschann-Grimm
Department of Physics and Astronomy, SUNY Stony Brook,
Stony Brook, NY 11794-3800, USA



This note reviews the search for the Standard Model Higgs boson at the DØ experiment with the final state containing two τ's and at least two jets. Data from Run 2b of the DØ experiment are used with an integrated luminosity of 4.3 fb^{-1}. This final state is sensitive to the production mechanisms gluon gluon fusion, vector boson fusion, and associated Higgs production with a W or Z, for Higgs masses from 100 to 200 GeV. No evidence for the Higgs boson is yet observed, so upper limits are placed on the cross section of the SM Higgs production. Including a previous DØ measurement with 1.0 fb^{-1} of data, we set a 95% CL. limit on the measured H cross sections for M_H = 110, 130 and 160 GeV that are factors of 20, 24 and 11 larger than expectations from the standard model, respectively.

1 Introduction

We present a search for the standard model (SM) Higgs boson (H) in final states with a lepton (e or μ), a candidate for the decay of $\tau \to$ hadrons + ν_τ, and two jets. We refer these two final state signatures as the $e\tau jj$ or $\mu\tau jj$ channels, and analyze such events for contributions from $q\overline{q} \to H(\to b\bar{b})Z(\to \tau\tau)(\text{denoted } HZ)$, or from $q\overline{q} \to ZH(\text{denoted } ZH)$, or from $q\overline{q} \to WH$, or from $gg \to H + (\geq 2 \text{ jets})(\text{gluon gluon fusion, GGF})$, or from $q\overline{q'} \to q\overline{q'}H(\text{virtual vector boson fusion, VBF})$.

The ZH, WH, GGF and VBF production processes are sought both through the $H \to \tau\tau$ and $H \to W^+W^-$ decays, denoted with subscripts $X_{\tau\tau}$, or X_{WW}, respectively. For the VH_{WW} subprocesses ($V = W$ or Z), the lepton can be produced either directly from $W \to \ell\nu$ or $Z \to \ell\ell$ with one ℓ not detected, or through V decays to τ lepton states with subsequent decay $\tau \to \ell\nu\overline{\nu}$.

The backgrounds to the $\tau(\ell)$ τ(hadronic) jet jet signatures are from $t\bar{t}$, $W+$ jets, $Z+$ jets, multijets and diboson ($WW/WZ/ZZ$) production.

We use 4.3 fb^{-1} of data collected with the upgraded DØ detector [2].

2 Data and Monte Carlo event samples

2.1 Monte Carlo samples

Monte Carlo samples for the Higgs signals and for the backgrounds are generated using a combination of PYTHIA [3] and ALPGEN [4]. The signals and backgrounds are processed using the standard DØ GEANT3 [5] detector simulation, digitization and event-reconstruction programs.

2.2 Event preselection

A description of object identification is given in the longer conference note [6]. We select a sample of candidate events with the following requirements: One isolated lepton, e or μ; At least one $\tau \rightarrow$ hadrons candidate with opposite electric charge as the e / μ ; At least two jets. Leptons, taus and jets are required to be separated from each other by $\Delta\mathcal{R} > 0.5$, where $\mathcal{R} = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$. To assure orthogonality to other H searches, we require no additional electron and no additional muon for the $\mu\tau jj$ ($e\tau jj$) channels, respectively.

For the $e\tau jj$ analysis, there is a substantial background contribution from $Z(\rightarrow ee)+$ jets. Therefore we impose additional requirements on τ's in this channel including removal of τ candidates falling in regions of the detector with degraded EM energy response and use of a Neural Net trained to specifically distinguish τ's and electrons.

2.3 Estimation of background from multijet events

Multijet (MJ) events in which jets mimic electrons, muons or τ's are not reliably simulated in our MC, and are estimated with data-driven techniques. We select a baseline MJ background sample in which both the τ and lepton pass significantly relaxed selection cuts but also fail the tight selection criteria used for the signal sample.

The MJ background sample is normalized by comparing the ratio of the opposite-sign lepton-tau pairs to the same-sign pairs in the signal sample and the MJ background sample.

2.4 Yields

The estimated number of events from all background sources and the number of observed data in our signal selection are given in Table 1. For a Higgs boson mass of 160 the total number of expected signal events is 1.36 for $\mu\tau jj$ and 0.34 for $e\tau jj$.

Table 1: The number of background events expected from SM processes, MJ background, and observed data after preselection in the $\mu\tau jj$ and $e\tau jj$ analyses. "DB" stands for di-boson processes.

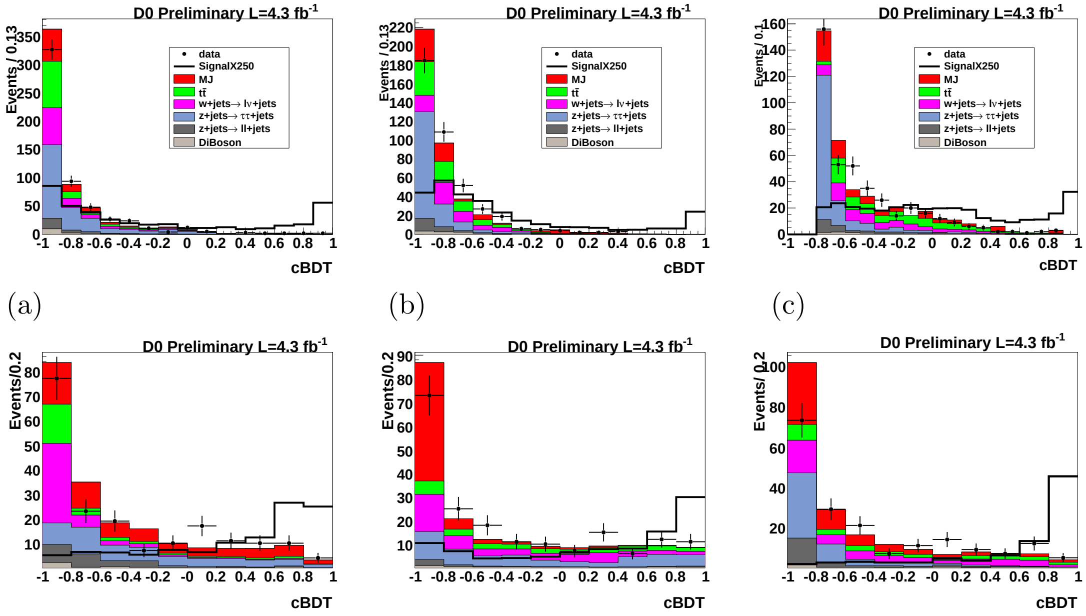
	$t\bar{t}$	W+jets	$Z_{\mu\mu}$+jets	$Z_{\tau\tau}$+jets	DB	MJ	ΣBkgd	Data
$\mu\tau jj$	82.0	64.8	22.3	154.5	11.3	70.0	404.9	414
$e\tau jj$	24.4	42.6	19.8	48.6	3.6	59.2	198.3	188

3 Multivariate analysis

As no single set of selections on kinematic variables suffices to discriminate signal from the background, we turn to multivariate techniques to attain better separation. We use stochastic gradient boosted decision trees (BDT) [7] as implemented in TMVA [8] for this purpose.

To train the BDT's we choose a set of 17 well-modelled kinematic variables, for which the distributions of at least some signal and some background are different.

There are three rather distinct regions of Higgs mass in this analysis, $M_H < 125$, $125 \leq M_H \leq 135$ GeV and $M_H > 135$ GeV, in which the dominant production and decay processes are different. The most dominant signals in the three regions are $\mathrm{GGF}_{\tau\tau}$, $VH_{\tau\tau}$, and $\mathrm{VBF}_{\tau\tau}$ in the low mass region, $\mathrm{GGF}_{\tau\tau}$, GGF_{WW}, $VH_{\tau\tau}$, and VH_{WW} in the intermediate mass region, and GGF_{WW}, VH_{WW}, and VBF_{WW} in the high mass region. The BDT's are trained separately in the three mass regions against the backgrounds $Z+$ jets, MJ, and $t\bar{t}$ plus $W+$ jets. This gives a total of 30 separate BDT's. The BDT's successfully separate the signals and backgrounds they are trained against, but sometimes the signals (backgrounds) that are not used for the training fall in the low (high) regions of BDT output where backgrounds (signals) are expected to dominate. We therefore construct a final combined BDT (cBDT) in each mass region, using the individual BDT outputs in that region as inputs into the cBDT. The task of the cBDT is to weigh conflicting information, *e.g.* whether a particular event is more like one of the signals than any of the backgrounds. The final cBDT distributions are shown in Fig. 1.



(a) (b) (c)

(d) (e) (f)

Figure 1: Combined BDTs for $\mu\tau jj$ (top) and $e\tau jj$ (bottom) for the: low M_H region (a), (d); intermediate M_H region (b), (e); high M_H region (c), (f) . Signals are multiplied by a factor of 250.

4 Systematic uncertainties

Systematic uncertainties for each factor that influences the final cBDT distributions are estimated by changing the relevant factor by ± 1 standard deviation from its nominal value and propagating the change to the cBDT distribution. The largest systematic uncertainties arise from jet identification and reconstruction (≈20%), from jet resolution (≈15%) and jet energy scale (≈15%) and from the estimation of the MJ background (≈15%).

5 Limit calculation

The upper limits on the production cross section of Higgs bosons assuming SM Higgs decay branching ratios are calculated using the modified frequentist method [9]. The test statistic is the negative of a binned Poisson log-likelihood ratio (LLR) computed at each of the assumed

Higgs mass values from 105 to 200 GeV in 5 GeV steps. The LLR for different hypotheses (*e.g.* background-only, LLR_b, or signal+background, LLR_{s+b}) are used to compute the confidence levels $\mathrm{CL_b}$ and $\mathrm{CL_{s+b}}$ that give the probability that the LLR value from a set of 50,000 simulated pseudo-experiments is less likely than that observed, at a given confidence level.

The cross sections of the hypothesized Higgs signal at a given M_H are then scaled up from their SM values until the value of $\mathrm{CL_s} = \mathrm{CL_{s+b}}/\mathrm{CL_b}$ reaches 0.05 which defines the limit cross sections at 95% CL. Figure 2 shows the expected and observed limits and the expected LLRs for the $\mu\tau jj$ and $e\tau jj$ channels combined (including the previous $\mu\tau jj$ limit [1]).

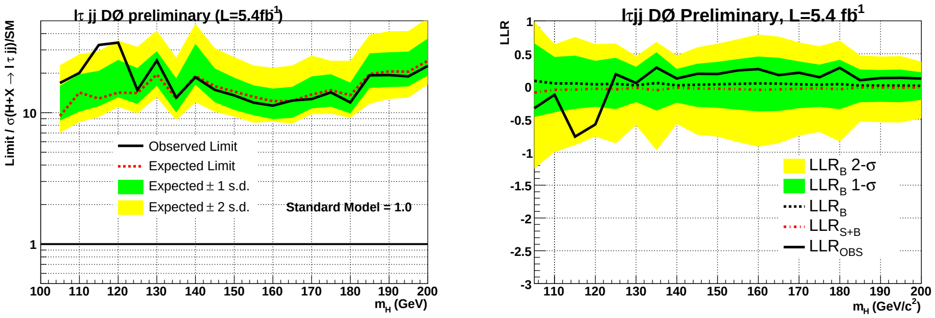


Figure 2: For the combined 4.3 fb^{-1} $\mu\tau jj$ and $e\tau jj$ analyses and the 1.0 fb^{-1} $\mu\tau jj$: the ratio of the 95% upper C.L. limits to the SM cross section (left) and the LLR as functions of Higgs boson mass (right).

6 Conclusion

We have searched for SM Higgs boson production in final states containing an electron or muon, a hadronically decaying τ plus two jets. Several different Higgs production processes contribute to this final state. At $M_H = 110$, 130 and 160 GeV, after combining the results in this analysis with that of a previous publication [1], we set a final combined limit on SM Higgs boson production that is a factor of 20, 24, and 11 times larger than the cross section predicted in the SM, to be compared to the expected ratios of 14, 20 and 12.

References

1. V.M. Abazov *et al.*, (DØ Collaboration), Phys. Rev. Lett **102**, 251801 (2009).
2. V.M. Abazov *et al.*, (DØ Collaboration), Nucl. Instrum. Methods Phys. Res A **565**, 463 (2006).
3. T. Sjöstrand *et al.*, Computer Phys. Commun. **135**, 238 (2001); T. Sjöstrand *et al.*, arXiv:(hep-ph) 0108264 (2001).
4. M. Mangano *et al.*, J. High Energy Physics **0307**, 001 (2003).
5. R. Brun and F. Carminati, CERN Program Library Long Writeup W5013, 1993 (unpublished), version V3.21.
6. The DØ Collaboration, ”Search for the SM Higgs boson in the tau + tau + 2 jets final state” http://www-d0.fnal.gov/Run2Physics/www/results/prelim/HIGGS/H101/H101.pdf
7. L. Breiman *et al.*, *Classification and Regression Trees* (Wadsworth, Stamford CT, 1984); D. Bowser-Chao and D.L. Dzialo, Phys. Rev D **47**, 1900 (1993).
8. TMVA Users Guide, CERN-OPEN-2007-007 and arXiv:physics/0703039.
9. A. Read, J. Phys. G: Nucl. Part. Phys. **28**, 2693 (2002); T. Junk, Nucl. Instrum. Methods A **434**, 435 (1999).