

# CDF searches for diboson production in final states with heavy flavor jets

---

**Marco Trovato**<sup>\*†</sup>

*Fermilab and Scuola Normale Superiore - Pisa*

*E-mail:* [mtrovato@FNAL.GOV](mailto:mtrovato@FNAL.GOV)

This article presents the diboson measurements in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV center mass energy with the Tevatron collider. The whole CDF dataset is used. This dataset corresponds to an integrated luminosity of about  $9.5 \text{ fb}^{-1}$ . This paper focuses on final states with heavy-flavor jets as a benchmark for Higgs analysis and as a test of the standard model (SM). A brand new multi-variate tagger is implemented in the described measurements. Assuming the ratio of the  $WZ$  and  $ZZ$  production cross-sections, we measure  $\sigma(WZ + ZZ) = 4.1 \pm 1.4 \text{ pb}$  in agreement with the SM (4.4 pb). The significance of the measurement is about 3.2 standard deviations above the background expectations.

*36th International Conference on High Energy Physics,  
July 4-11, 2012  
Melbourne, Australia*

---

<sup>\*</sup>Speaker.

<sup>†</sup>On behalf of the CDF Collaboration

## 1. Introduction

Studies on the production of  $VZ$  ( $V = W, Z$ ) boson pairs provide an important test for the electroweak sector of the SM. Measuring a significant departure in cross section or deviations in the predicted kinematic distributions would indicate the presence of anomalous gauge boson couplings [1] or new particles in extensions of the SM [2].

Moreover the investigation of diboson processes in final states with heavy flavor jets is important since the event topologies of these processes are the same as the ones of the major SM light-Higgs boson ( $M_H \lesssim 135$  GeV) searches performed at the Tevatron. Thus, the investigation of such processes whose rate can be accurately predicted, allows to calibrate and optimize many of the techniques implemented for the Higgs searches. In addition, diboson production generates a significant background for low mass Higgs Boson searches with  $H$  decaying into a  $b\bar{b}$  pair

In this paper we report the history of CDF diboson measurements in fully leptonic final states or in hadronic final states prior any  $b$  – *tagging* (Sec. 2). Then we introduce the brand new  $b$  – *tagger* (3) and we summarize the three main diboson searches used as Higgs benchmark and their combination (4).

## 2. Diboson Measurements: history

$VV$  ( $= WW, WZ, ZZ$ ) production in the fully leptonic final states has been observed by the CDF collaboration in three decay modes:  $WZ \rightarrow l\nu ll$ ,  $ZZ \rightarrow llll$ ,  $WW \rightarrow l\nu l\nu$ .

CDF also observed the diboson production in semileptonic final states. An analysis exploiting the invariant mass of the two leading jets was able to extract the  $WV \rightarrow l\nu q\bar{q}$  signal in a  $3.9\text{ fb}^{-1}$  dataset. A complementary analysis using matrix-element calculations achieved the similar results with a  $2.7\text{ fb}^{-1}$  dataset. Both samples were selected by requiring the presence of a lepton and large  $\cancel{E}_T$ , besides two energetic jets. The combination of the two methods has a signal significance of 5.4 standard deviations [4]. A cross section  $\sigma(WW + WZ) = 16.0 \pm 3.3\text{ pb}$  is measured. Such a result is in agreement with the SM. Another analysis searching for  $VV = WW + WZ + ZZ$  in a  $3.5\text{ fb}^{-1}$  dataset selected with large  $\cancel{E}_T$  and two energetic jets, measures a cross section of  $18.0 \pm 3.8\text{ pb}$  [5].

## 3. b-tagger

At hadron colliders,  $b$  taggers allow one to distinguish jets containing  $B$  hadrons from other jets. In this paper we briefly describe the Higgs Optimized  $b$  Identification Tagger (HOBIT, [6]). HOBIT employs feed-forward neural networks architectures optimized specifically for finding light Higgs boson decays. HOBIT was developed upon the strengths of previous CDF  $b$ -tagger ([7], [8]) and, by addressing their weakness, achieve great performances. Moreover HOBIT produces a continuous output variable, allowing for efficiency and background rejection to be tuned upon the given search sensitivity.

For two given operation points ("tight", "loose"), tuned to match the light jet rejection ("mistag") rates of previous  $b$ -taggers, HOBIT tags respectively 54%, 59%  $b$ -jets rather the previously achieved 39%, 47% <sup>1</sup>. In order to be used in a physics analysis HOBIT performances must be calibrated.

<sup>1</sup>These efficiencies are estimated from simulated events

$W + 1, \geq 3$  jets events are used to calibrate the  $b$ -tag efficiency and mistag rate. As a cross check we use a sample with di-jet events with one jet containing an electron candidate within its cone. The extracted MC to data  $b$ -tag efficiency and mistag scale factors (SF) are in agreement between the two samples. The combination of the two gives:

- $b$ -tag efficiency SF:  $0.99 \pm 0.03$  (loose),  $0.94 \pm 0.04$  (tight)
- mistag rate SF:  $1.33 \pm 0.13$  (loose),  $1.49 \pm 0.28$  (tight)

Further details are reported at [6]

#### 4. Diboson searches in final states with heavy flavor jets

The use of the  $b$ -tagger is crucial to separate the  $W$  from  $Z$  hadronic peak, which be otherwise undistinguished because of the limited CDF calorimeter resolution.

The diboson signal is searched for in three orthogonal samples:

- 2 identified leptons (" $llb\bar{b}$ ")
- 1 identified lepton and large  $\cancel{E}_T$  (" $lvb\bar{b}$ ")
- no identified leptons and large  $\cancel{E}_T$  (" $vvb\bar{b}$ ")

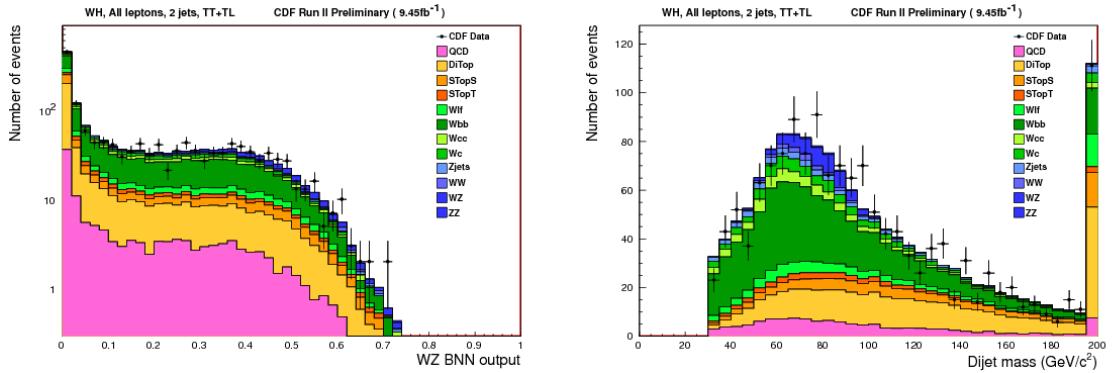
In all the above samples the presence of two or three energetic jets ( $E_T > 20$  GeV) is required. One of the jets must pass the  $b$ -tagger criteria.

In order to improve the sensitivity of the measurement candidate events in  $llb\bar{b}$  and  $lvb\bar{b}$  are also separated based on the lepton flavor (electron or muon) and on tight and loose tagging definitions (Sec. 3). Events with two tight tags, one tight and one loose tag, two loose tags, and a single tight tag are used by both analyses. The  $lvb\bar{b}$  analysis also considers events with a single loose tag. The  $vvb\bar{b}$  analysis makes use of a tight  $b$ -tagging algorithm [7] based on the reconstruction of a displaced secondary vertex and a loose  $b$ -tagging algorithm [8] that assigns a likelihood for the tracks within a jet to have originated from a displaced vertex. Based on the output of these algorithms, events with two tight tags and those with one tight tag and one loose tag are separated into independent analysis channels. The signal in all of the double-tag samples is expected to be primarily composed of events with  $Z \rightarrow b\bar{b}$  decays, with smaller contributions from  $Z \rightarrow c\bar{c}$  and  $W \rightarrow c\bar{s}$  decays. In the single-tag samples the contributions from the three decay modes are comparable.

The primary background is from  $W/Z$ +jets, which is modelled with ALPGEN [9] (matrix element) and PYTHIA [10] (showering and hadronization). The background from multijet production is modelled from control samples in the data. The backgrounds from other SM processes are modelled using PYTHIA Monte Carlo samples. Background rates are normalized either to next-to-leading order (NLO) or higher-order theory calculations or to data control samples. The fraction predicted by ALPGEN of the  $W/Z$ +jets in which the jets arise from heavy quarks (b or c) is corrected based on data control regions. The background from  $t\bar{t}$  events is normalized to the approximate NNLO cross section [11]. The  $s$ -channel and  $t$ -channel cross sections for the production of single-top quarks are derived respectively from approximate NNLO+NNLL calculations [12]

and approximate NNNLO+NLL calculations [13]. The background from WW events is normalized to NLO calculations from MCFM [14].

All the three analyses make use of multi-variate discriminants to extract the signal. These discriminants are based on a full reconstruction of the final state. Compared to the invariant mass of the two leading jets an improved sensitivity in the diboson search is obtained. Although the discriminants have been retrained to extract the VZ, they share the same architecture and input variables of the related low-mass Higgs analyses [15], [16], [17]. An example of discriminant outputs as they appear in data and in the simulations are shown in Fig. 1.



**Figure 1:** Predicted and observed output for the neural network trained to separate the VZ signal from the background used in the  $lvbb$  measurement (left). The invariant mass distribution is also shown (right). Similar plots for the  $llbb$ ,  $vvbb$  measurements are available at [16], [17] respectively.

## 5. Combination of Diboson Searches

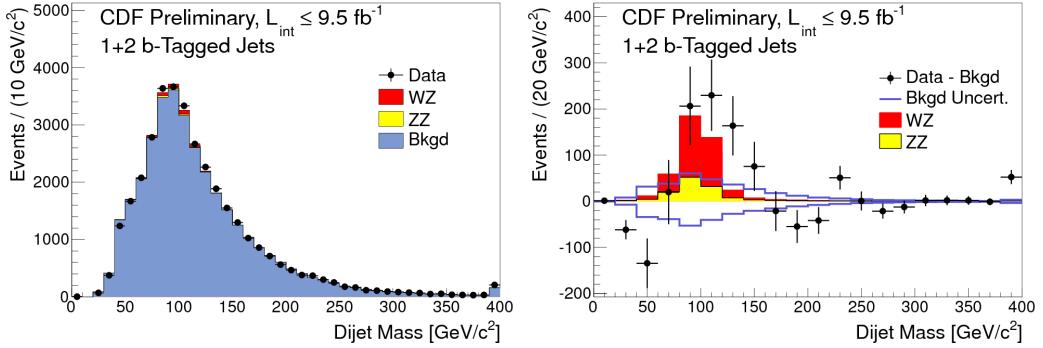
The total VZ cross section is determined from a maximum likelihood fit of the MVA distributions for the background and signal samples from the contributing analyses to the data (Sec. 4). The cross section for the signal is a free parameter in the fit, but the ratio of the  $WZ$  and  $ZZ$  cross sections is fixed to the SM prediction. Events from  $WW$  production are treated as a background and constrained to SM predictions. The fit is performed simultaneously on the distributions in all sub-channels.

Systematic uncertainties are treated as nuisance parameters with truncated Gaussian. Systematic uncertainties affect both the normalizations and the differential distributions of the predicted signal and background templates in correlated ways. They are either treated as fully correlated or uncorrelated. The largest uncertainties on signal arise from integrated luminosity, measured  $b$ -tagging efficiencies, jet energy scale, and other Monte Carlo parameters. The backgrounds with the largest systematic rate uncertainties are in general quite small. Such uncertainties are constrained through fits to the nuisance parameters and do not affect the result significantly. Since normalizations for the  $W/Z$ +heavy flavor backgrounds are obtained from data in the  $lvbb$  and  $vvbb$  analyses, the corresponding rate uncertainties associated with each analysis are treated as uncorrelated.

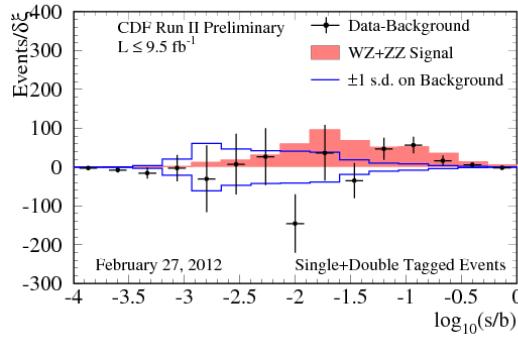
The combined fit for the total VZ cross section yields  $\sigma(WZ + ZZ) = 4.1 \pm 1.4$  pb. This measurement is consistent with the NLO SM prediction of  $\sigma(WZ + ZZ) = 4.4 \pm 0.3$  pb [14], as well as

with the measurement from D0 [18],  $\sigma(WZ + ZZ) = 5.0 \pm 1.6$  pb. Based on the measured central value for the  $VZ$  cross section and its uncertainties, the observed significance is estimated to be 3.2 standard deviations. Figure 2 shows the distributions of the invariant mass of the dijet system, summed over all channels, after adjusting the signal and background predictions according to the results of the fit. It is also shown the background subtracted dijet mass distributions after the fit, demonstrating the presence of a hadronic resonance in the data consistent with the SM expectation.

To visualize the sensitivity of the combined analysis, we calculate the expected signal over background ( $s/b$ ) in each bin of the MVA distributions from the contributing analyses. Bins with similar  $s/b$  are then combined to produce a single distribution, shown in Fig. 3.



**Figure 2:** Comparison of the fitted signal+background to data in the dijet mass distribution (left plot). Comparison of the measured  $WZ$  and  $ZZ$  signals (filled histograms) to background-subtracted data (points) in the dijet mass distribution (left plot). Also shown is the  $\pm 1$  standard deviation uncertainty on the fitted background.



**Figure 3:** Comparison of the measured  $VZ$  signal (filled histogram) to background-subtracted data (points) after the maximum likelihood fit. The distribution is a combination of all final discriminants where the bins are ordered and merged according to their expected signal to background ratio ( $s/b$ ). The x-axis has arbitrary units. Also shown is the  $\pm 1$  standard deviation uncertainty on the fitted background that was subtracted.

## 6. Summary

In summary, we combine analyses in the  $llb\bar{b}$ ,  $lvb\bar{b}$ , and  $vvb\bar{b}$  final states from the CDF experiment to observe, with a significance of 3.2 standard deviations, the production of  $VZ$  ( $V = W$  or  $Z$ ) events. The analyzed samples correspond to  $9.5\text{ fb}^{-1}$  of  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96\text{ TeV}$ . We measure the total cross section for  $VZ$  production to be  $\sigma(VZ + ZZ) = 4.1 \pm 1.4\text{ pb}$ . This result demonstrates the ability of the CDF experiment to measure a SM production process with cross section of the same order magnitude as that expected for the Higgs production. The same set of background-dominated final states containing two heavy-flavor jets are used in both diboson and low mass Higgs searches.

## References

- [1] K. Hagiwara, S. Ishihara, R. Szalapski, and D. Zeppenfeld, Phys. Rev. D **48** (1993).
- [2] J. C. Pati and A. Salam, Phys. Rev. D **10**, 275 (1974); 11 703(E) (1975);  
G. Altarelli, B. Mele, and M. Ruiz-Altaba, Z. Phys. C **45**, 109 (1989); 47, 676(E) (1990);  
L. Randall and R. Sundrum, Phys. Rev. Lett. **83**, 3370 (1999);  
H. Davoudiasl, J. L. Hewett, and T. G. Rizzo, Phys. Rev. D **63**, 075004 (2001);  
H. He *et al.*, Phys. Rev. D **78**, 031701 (2008).
- [3] CDF Collaboration (A. Abulencia *et al.*), Phys Rev Lett **98**, 161801 (2007).
- [4] CDF Collaboration (T. Aaltonen *et al.*), Phys. Rev. Lett. **104**, 101801 (2010).
- [5] CDF Collaboration (T. Aaltonen *et al.*), Phys. Rev. Lett. **103**, 091803 (2009).
- [6] J. Freeman, T. Junk, M. Kirby, Y. Oksuzian, T. J. Phillips, F. D. Snider, M. Trovato and J. Vizan *et al.*, Nucl. Instrum. Meth. A **697**, 64 (2013) [arXiv:1205.1812 [hep-ex]].
- [7] D. Acosta *et al.* (CDF Collaboration), Phys. Rev. D **71**, 052003 (2005).
- [8] A. Abulencia *et al.* (CDF and CDF - Run II Collaborations), Phys. Rev. D **74**, 072006 (2006).
- [9] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau and A. D. Polosa, J. High Energy Phys. 07, 001 (2003).
- [10] T. Sjostrand, L. Lonnblad and S. Mrenna, arXiv:hep-ph/0108264.
- [11] U. Langenfeld, S. Moch and P. Uwer, Phys. Rev. D **80**, 054009 (2009).
- [12] N. Kidonakis, arXiv:1005.3330 [hep-ph] (2010);  
N. Kidonakis, Phys. Rev. D **81**, 054028 (2010).
- [13] N. Kidonakis, Phys. Rev. D **74**, 114012 (2006).
- [14] J. M. Campbell and R. K. Ellis, Phys. Rev. D **60**, 113006 (1999). We used MCFM v6.0. Cross sections are computed using a choice of scale  $\mu_0^2 = M_V^2 + P_T^2(V)$ , where  $V$  is the vector boson, and the MSTW2008 PDF set.
- [15] CDF Collab. (T. Aaltonen *et al.*), Phys. Rev. Lett. **109**, 111804, (2012).
- [16] CDF Collab. (T. Aaltonen *et al.*), Phys. Rev. Lett. **109**, 111803, (2012).
- [17] CDF Collab. (T. Aaltonen *et al.*), Phys. Rev. Lett. **109**, 111805, (2012).
- [18] V. M. Abazov *et al.* (D0 Collaboration), D0 Note 6260-CONF (2011).