

Combined search for the standard model Higgs boson with the D0 experiment

Émilien Chapon^{1,a}, on behalf of the D0 collaboration

¹Service de Physique des Particules, CEA Saclay, F91191 Gif-sur-Yvette cedex, France

Abstract. We present searches for the standard model Higgs boson using up to 9.7 fb^{-1} of $p\bar{p}$ collisions collected with the D0 detector at Fermilab. The analyses consist in a series of distinct final states and are sensitive to Higgs boson masses (M_H) ranging from 90 GeV to 200 GeV. These analyses are combined and allow to exclude a standard model Higgs boson at 95% Confidence Level (C.L.) in the range $90 \text{ GeV} < M_H < 101 \text{ GeV}$ and $157 \text{ GeV} < M_H < 178 \text{ GeV}$, with an expected exclusion in the range $155 \text{ GeV} < M_H < 175 \text{ GeV}$. An excess of data of about two Gaussian standard deviations is seen in the range $120 \text{ GeV} < M_H < 145 \text{ GeV}$, consistent with the observation of a boson with a mass of 125 GeV at the Large Hadron Collider and with the evidence for a particle decaying to $b\bar{b}$ at the Tevatron.

The easiest way to give a mass to the W and Z electroweak vector bosons is via the well-established $SU(2) \times U(1)$ electroweak symmetry breaking mechanism. In the standard model (SM), one introduces a single elementary scalar field, doublet of $SU(2)$, which acquires a nonzero vacuum expectation value. Once accounting for the longitudinal polarizations, i.e., for the mass of the electroweak vector bosons, one degree of freedom remains, which manifests itself as a single scalar particle, the Higgs boson. Its mass is a free parameter of the model but is constrained from direct searches from LEP [1] and Tevatron [2]. Its properties are also consistent with the evidence for a new particle decaying to $b\bar{b}$ at the Tevatron [3], and with the discovery of a new boson with a mass $M_H \approx 125 \text{ GeV}$ at the LHC [4, 5].

The phenomenology of the SM Higgs boson is very rich: because it couples to all massive particles at tree level, it can be produced and decays in a number of different ways. At the Tevatron, the dominant production modes are gluon fusion (GGF, $gg \rightarrow H$), associated production with a vector boson (VH, $qq' \rightarrow WH, ZH$) and vector boson fusion (VBF, $qq' \rightarrow qq'H$). The dominant decay modes are $H \rightarrow b\bar{b}$ and $H \rightarrow W^+W^-$, but some analyzes are also sensitive to $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ$ or $H \rightarrow \tau^+\tau^-$. In addition, W and Z bosons and τ leptons further decay to particles long-lived enough to be detected (and / or to neutrinos), thus leading to a large range of final states to look for.

All analyzes searching for a SM Higgs boson follow a similar scheme. They start with a first set of selection criteria (triggers, lepton quality and transverse momentum p_T , etc.) as loose as possible, in order to maximize the acceptance. Then a certain number of multivariate analysis techniques are used (e.g. Boosted Decision Trees,

BDTs). These techniques combine several discriminating variables between some signal and some background into one discriminant. They make the most of correlations between the input variables. They can be trained against specific backgrounds, e.g. $t\bar{t}$ or multijet background.

Most of the time, the analyzes are split into sub-categories (depending, for instance, on the lepton flavor or on the number of jets in the event). This allows to take advantage of different signal and background compositions in each category, as well as different signal to background ratios, and to better constrain systematic uncertainties associated to specific background processes. The background modeling is constrained in control regions, and can further be checked with a diboson cross-section measurement (in the $H \rightarrow b\bar{b}$ and $H \rightarrow W^+W^-$ analyzes), as can be seen on Fig. 1 for the $WZ + ZZ$ production cross section measurement. At last, a final discriminant is used to look for an excess in the data compared to the background-only expectation.

Analyzes presented here use up to 9.7 fb^{-1} of $p\bar{p}$ collisions collected with the D0 detector at Fermilab. We will briefly describe the following analyzes: $WH \rightarrow l\nu b\bar{b}$, $ZH \rightarrow \ell^+ \ell^- b\bar{b}$, $ZH \rightarrow \nu\bar{\nu} b\bar{b}$ and $H \rightarrow W^+W^- \rightarrow \ell^+ \ell^- \nu\bar{\nu}$. The following analyzes also enter the D0 combination [6]: $H+X \rightarrow W^+W^- \rightarrow \mu^\pm \tau_h^\mp + \leq 1 \text{ jet}$, $H \rightarrow W^+W^- \rightarrow \ell\nu q' \bar{q}$, $VH \rightarrow ee\mu/\mu\mu e+X$, $VH \rightarrow e^\pm \mu^\pm +X$, $VH \rightarrow \ell\nu q' \bar{q} q' \bar{q}$, $VH \rightarrow \tau_h \tau_h \mu + X$, $H+X \rightarrow \ell\tau_h jj$, and $H \rightarrow \gamma\gamma$.

1 $H \rightarrow b\bar{b}$

The Higgs boson predominantly decays to a pair of b quarks for $M_H < 135 \text{ GeV}$. However it is not possible to use the GGF production mode to look for the Higgs boson in this decay mode, because of the overwhelming multijet background at a hadron collider like the Tevatron, and analyzers focus on the VH production mode. This results

^ae-mail: emilien.chapon@cern.ch

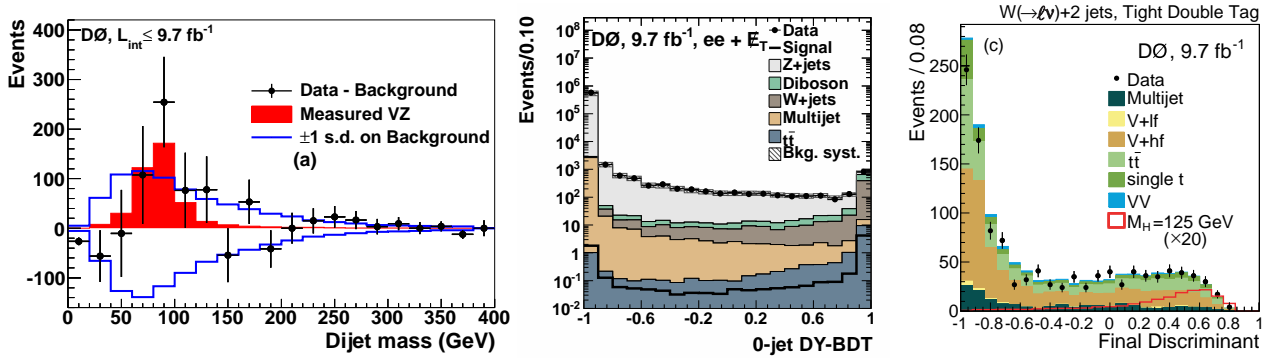


Figure 1. Left: Background-subtracted distribution of the reconstructed dijet mass m_{jj} for the $WH \rightarrow \ell\nu b\bar{b}$, $ZH \rightarrow \nu\bar{\nu} b\bar{b}$, and $ZH \rightarrow \ell^+\ell^- b\bar{b}$ searches, used in the measurement of the $WZ + ZZ$ production cross section. Middle: Distribution of one of the BDTs used to discriminate signal from the Z/γ^* background in the $H \rightarrow W^+W^- \rightarrow e^+e^-\bar{\nu}$ analysis. Right: output of one of the final discriminants used in the $WH \rightarrow \ell\nu b\bar{b}$ analysis to look for a data excess over the background-only expectation and to set limits on SM Higgs boson production.

in three distinct analyzes: $WH \rightarrow \ell^\pm\nu b\bar{b}$, $ZH \rightarrow \ell^+\ell^- b\bar{b}$ and $ZH \rightarrow \nu\bar{\nu} b\bar{b}$. The presence of a vector boson in the final state allows to reduce and constrain the backgrounds in these analyzes.

It is crucial for these analyzes to be built on an excellent b -tagging algorithm, to disentangle jets originating from a b quark (b -jets) from those originating from a light quark. Such an algorithm uses the fact that B hadrons have a relatively long lifetime and travel in the detector before they decay. Variables gauging these characteristics are combined into a multivariate technique to maximize the tagging efficiency and minimize the fake rate. By adjusting the minimum requirement on the output of the b -tagger, a range of signal efficiencies and purities is achieved.

The $WH \rightarrow \ell^\pm\nu b\bar{b}$ analysis [7] is divided into categories depending on the lepton flavor (electron or muon) and on the number and quality of b -tagged jets. It also uses a BDT trained against the multijet background, which is used as an input to the final discriminants (one for each category of events).

The advantage of the $ZH \rightarrow \ell^+\ell^- b\bar{b}$ analysis [8] is that it provides a fully reconstructed final state, without any missing energy from neutrinos as opposed to other Higgs analyzes. This allows to constrain the kinematics of the events, using a kinematic fit. The first step in this analysis is to select a $Z \rightarrow \ell^+\ell^-$ candidate with $70 \text{ GeV} < M_{\ell\ell} < 110 \text{ GeV}$. A dedicated BDT is trained against the $t\bar{t}$ background, which allows to split the analysis into $t\bar{t}$ depleted and enriched regions. A final discriminant is at last trained for each event category.

The main challenge in the $ZH \rightarrow \nu\bar{\nu} b\bar{b}$ analysis [9] is to model and reject the dominant multijet background, and this is done combining several tools. First, analyzers cut on a variable called E_T significance, which gauges if the missing transverse energy (E_T) is likely to have been mis-measured in the event. To further reject the multijet background, a dedicated BDT is trained against this background and a requirement is placed on its output. At last, the output L_b of the b -tagging discriminant for the two jets

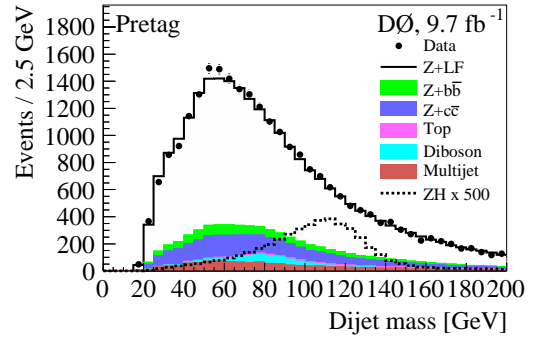


Figure 2. Invariant mass of the two jets in the $ZH \rightarrow \ell^+\ell^- b\bar{b}$ analysis, before b -tagging requirement.

is summed to obtain the variable $L_{bb} = L_b(\text{jet}_1) + L_b(\text{jet}_2)$. Events with a low L_{bb} are rejected, and the remaining events are categorized into medium / high L_{bb} regions. A final discriminant is trained in each of these two event categories.

2 $H \rightarrow W^+W^- \rightarrow \ell\nu\ell\nu$

The $H \rightarrow W^+W^- \rightarrow \ell\nu\ell\nu$ analysis [11] is the most sensitive channel for $M_H > 135 \text{ GeV}$, but it has also a good sensitivity at lower masses. Even if the branching ratio of $W^+W^- \rightarrow \ell\nu\ell\nu$ is relatively low ($\sim 6.4\%$) compared to semi- or full-hadronic decay modes of a W boson pair, the dileptonic decay mode benefits from a clean experimental signature, with two high- p_T leptons (e^+e^- , $\mu^+\mu^-$ or $e^\pm\mu^\mp$) and large E_T . In order to maximize the acceptance, no explicit cut on E_T is placed. Instead, a BDT is trained in the e^+e^- and $\mu^+\mu^-$ channels to reject the Z/γ^* background, while the $e^\pm\mu^\mp$ channel places cuts on E_T -related variables.

Then, the analysis is split into 0 jet, 1 jet and ≥ 2 jets multiplicity bins. Events are further categorized into WW -depleted and WW -enriched regions, using a BDT in the 0

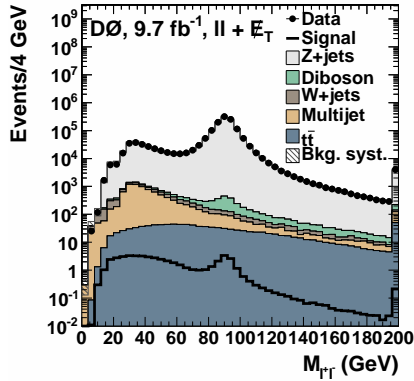


Figure 3. Invariant mass of the two leptons in the $H \rightarrow W^+W^- \rightarrow \ell\nu\ell\nu$ analysis, before rejection of the Z/γ^* background.

and 1 jet bins (e^+e^- , $\mu^+\mu^-$) or using the lepton quality in the 0 jet bin ($e^\pm\mu^\mp$).

At last, a final BDT is trained in each sub-samples against the remaining backgrounds. Angular variables such as the opening angle $\mathcal{R}(\ell\ell) = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ between the two leptons are of primary importance, because we look for a standard model Higgs boson which has a spin 0. The final discriminant is used to look for any data excess in a signal-like region and to set limits on Higgs boson production.

3 D0 combined limits on standard model Higgs boson production

The modified frequentist CL_s method [12] is used to gauge the compatibility of the data with the background-only hypothesis (B) or signal-plus-background hypothesis (S+B) and to set limits on SM Higgs boson production, where the test statistic is a log-likelihood ratio (LLR) for the (B) and (S+B) hypotheses. The degrading effect of systematic uncertainties is reduced by fitting individual background contributions to the data by maximizing a profile likelihood function for the (B) and (S+B) hypotheses separately, appropriately taking into account all correlations between the systematic uncertainties [13].

In order to achieve the best possible sensitivity to Higgs boson production, analyzers have combined their results together. This allows to take advantage of the specificities of each analysis: for instance $H \rightarrow b\bar{b}$ channels are sensitive to the coupling of the Higgs boson to fermions and are more sensitive to low M_H , while $H \rightarrow W^+W^-$ channels are sensitive to the coupling of the Higgs boson to vector bosons and are more sensitive at high M_H . The value of the combined LLR as a function of the hypothetical Higgs boson mass M_H is shown on Fig. 4.

When combining all analyzes available, the D0 collaboration is able to exclude a SM Higgs boson with a mass in the range $90 \text{ GeV} < M_H < 101 \text{ GeV}$ and $157 \text{ GeV} < M_H < 178 \text{ GeV}$ at 95 % confidence level (C.L.) [6], while the expectation is to exclude the range $155 \text{ GeV} < M_H <$

175 GeV (see Fig. 4, right). A data excess over expectation of up to 2 standard deviations (s.d.) is however seen in the range $120 \text{ GeV} < M_H < 145 \text{ GeV}$, which is seen from the the background-only p -value as a function of M_H (see Fig. 4). As a comparison, the combination with the CDF Tevatron experiment [2] excludes the ranges $90 \text{ GeV} < M_H < 109 \text{ GeV}$ and $149 \text{ GeV} < M_H < 182 \text{ GeV}$, and there is a data excess corresponding to a local significance of 3.0 s.d. for $M_H = 125 \text{ GeV}$. This excess is compatible with the observation of a new boson by the ATLAS and CMS collaborations [4, 5].

We want to study this excess into more detail. First, we compute separate limits combining only analyzes sensitive to $H \rightarrow b\bar{b}$ or $H \rightarrow W^+W^-$ (see Fig. 5), and we find that both $H \rightarrow b\bar{b}$ and $H \rightarrow W^+W^-$ combined limits report a broad excess of up to 1.5 s.d. We also compute the best fit signal strength ($\sigma_H \cdot \mathcal{B}/(\sigma_H \cdot \mathcal{B})_{\text{SM}}$) for a hypothetical Higgs boson with a mass $M_H = 125 \text{ GeV}$, as discovered at the LHC (see Fig. 5), for the combination as well as for the different Higgs boson decay channels considered. We find that each of the four main Higgs boson decay modes, $H \rightarrow b\bar{b}$, $H \rightarrow W^+W^-$, $H \rightarrow \tau^+\tau^-$ and $H \rightarrow \gamma\gamma$, contribute to the observed excess.

4 Conclusion

We have reported selected searches for the SM Higgs boson from the D0 collaboration, focusing on the $H \rightarrow b\bar{b}$ and $H \rightarrow W^+W^-$ decay channels. These analyzes are very complementary, both because they probe different Higgs boson mass ranges and are sensitive to different classes of couplings of the Higgs boson. We have also reported the results of the combined search for a Higgs boson by the D0 Collaboration (also including more analyses, in particular in the $H \rightarrow \tau^+\tau^-$ and $H \rightarrow \gamma\gamma$ channels), which exclude a SM Higgs boson with a mass in the range $90 \text{ GeV} < M_H < 101 \text{ GeV}$ and $157 \text{ GeV} < M_H < 178 \text{ GeV}$ at 95 % C.L. (expected: $155 \text{ GeV} < M_H < 175 \text{ GeV}$). An excess of data of up to 2 s.d. over background expectation is seen in the range $120 \text{ GeV} < M_H < 145 \text{ GeV}$, compatible with the evidence for a new particle decaying to $b\bar{b}$ from the Tevatron and the discovery of a new boson from the LHC. The analyses combined here also provide inputs to the overall Tevatron combination, which shows evidence for a SM-like Higgs boson [2].

Acknowledgments

We thank the staffs at Fermilab and collaborating institutions, and acknowledge support from the DOE and NSF (USA); CEA and CNRS/IN2P3 (France); MON, NRC KI and RFBR (Russia); CNPq, FAPERJ, FAPESP and FUN-DUNESP (Brazil); DAE and DST (India); Colciencias (Colombia); CONACyT (Mexico); NRF (Korea); FOM (The Netherlands); STFC and the Royal Society (United Kingdom); MSMT and GACR (Czech Republic); BMBF and DFG (Germany); SFI (Ireland); The Swedish Research Council (Sweden); and CAS and CNSF (China). The author would also like to thank the organizers of the conference for offering him financial support.

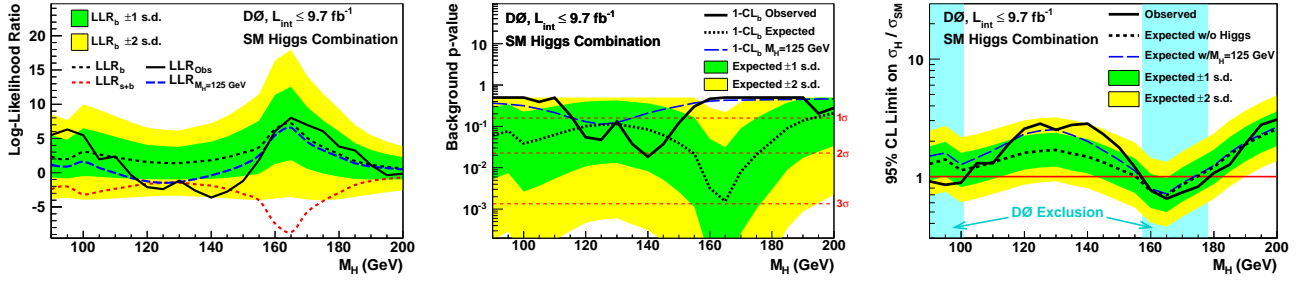


Figure 4. Left: Observed and expected LLR as a function of the Higgs boson mass. Middle: distribution of CL_b (background-only p -value) as a function of the Higgs boson mass. Right: 95% C.L. limits on SM Higgs boson production. On each of the three plots, the observed value in data is shown as a plain black line, the expectation from background only is shown as a dotted black line, and the expectation from a standard model Higgs boson with $M_H = 125$ GeV is shown as a dashed blue line.

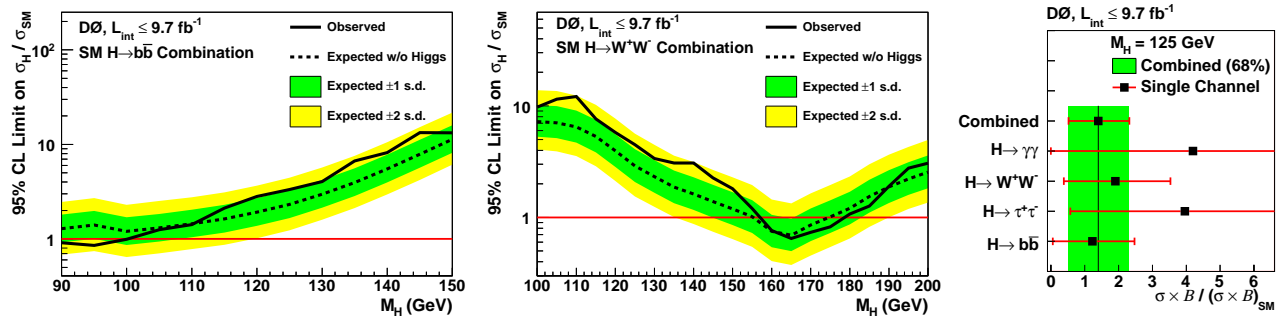


Figure 5. Left: 95% C.L. limits on SM Higgs boson production, when only analyzes sensitive to $H \rightarrow b\bar{b}$ are considered. Middle: 95% C.L. limits on SM Higgs boson production, when only analyzes sensitive to $H \rightarrow W^+W^-$ are considered. Right: best fit signal strength for various Higgs boson decays and for the combination, assuming $M_H = 125$ GeV.

References

- [1] R. Barate *et al.* [LEP Working Group for Higgs boson searches and ALEPH and DELPHI and L3 and OPAL Collaborations], Phys. Lett. B **565**, 61 (2003).
- [2] T. Aaltonen *et al.* [CDF and D0 Collaborations], arXiv:1303.6346 [hep-ex].
- [3] T. Aaltonen *et al.* [CDF and D0 Collaborations], Phys. Rev. Lett. **109**, 071804 (2012).
- [4] G. Aad *et al.* [ATLAS Collaboration], Phys. Lett. B **716**, 1 (2012).
- [5] S. Chatrchyan *et al.* [CMS Collaboration], Phys. Lett. B **716**, 30 (2012).
- [6] V. M. Abazov *et al.* [D0 Collaboration], arXiv:1303.0823 [hep-ex].
- [7] V. M. Abazov *et al.* [D0 Collaboration], arXiv:1301.6122 [hep-ex].
- [8] V. M. Abazov *et al.* [D0 Collaboration], arXiv:1303.3276 [hep-ex].
- [9] V. M. Abazov *et al.* [D0 Collaboration], Phys. Lett. B **716**, 285 (2012).
- [10] V. M. Abazov *et al.* [D0 Collaboration], arXiv:1301.5358 [hep-ex].
- [11] V. M. Abazov *et al.* [D0 Collaboration], arXiv:1301.1243 [hep-ex].
- [12] T. Junk, Nucl. Instrum. Meth. A **434**, 435 (1999); A. Read, J. Phys. G **28**, 2693 (2002).
- [13] W. Fisher [D0 Collaboration], FERMILAB-TM-2386-E.
- [14] V. M. Abazov *et al.* [D0 Collaboration], Phys. Rev. Lett. **109**, 121802 (2012).