Search for the Higgs boson decaying to W boson pair with the D0 experiment

Émilien Chapon([1]), on behalf of the DØ Collaboration
([1]) *CEA Saclay / Irfu / SPP, Gif-sur-Yvette, France*

Summary. — We present a search for the Higgs boson in final states with two oppositely charged leptons and large missing transverse energy as expected in $H \to WW \to \ell\nu\ell'\nu'$ decays. The events are selected from the full Run II data sample of 9.7 fb^{-1} of $p\bar{p}$ collisions collected with the D0 detector at the Fermilab Tevatron Collider at $\sqrt{s} = 1.96$ TeV. To validate our search methodology, we measure the non-resonant WW production cross section. In the Higgs boson search, no significant excess above the background expectation is observed. Upper limits at the 95% confidence level on the Higgs boson production cross section are therefore derived, within the Standard Model, but also within a theoretical framework with a fourth generation of fermions, and in the context of fermiophobic Higgs boson couplings.

PACS 14.80.Bn – Higgs.
PACS 13.85.Qk – Drell-Yan.
PACS 13.85.Rm – Limits on production of particles.

1. – Introduction

In the standard model (SM), the $SU(2) \times U(1)$ electroweak symmetry implies that the corresponding vector bosons should be massless. To accommodate the experimental evidence that the W and Z bosons have a mass, this symmetry is broken in the Brout-Englert-Higgs mechanism [1]. It postulates the existence of a single scalar field, doublet of $SU(2)$, which acquires a non-zero vacuum expectation value. Its longitudinal polarizations become the mass of the electroweak boson, and the remaining degree of freedom manifests itself as a single scalar particle, the Higgs boson.

This particle was the last of the SM yet to be observed, and constraints have been set on its mass M_H, a free parameter of the model. The LEP experiments have set a lower bound on M_H at 114.4 GeV [2], and the Tevatron experiments have excluded a mass range around two times the W boson mass [3]. This exclusion range was extended by the ATLAS and CMS experiments at the LHC, and these experiments have reported the observation of a new boson in 2012 [4, 5]. The DØ and CDF experiments also have announced the evidence for a new boson decaying to $b\bar{b}$ [6] at a mass consistent with the LHC discovery.

The phenomenology of the Higgs boson is very rich, because it couples to every massive particle of the SM, with a coupling proportional to its mass. At the Tevatron, the main production mode is gluon fusion ($gg \to H$), but contributions from associated production with a vector boson ($qq' \to V \to VH$, $V = W, Z$) and vector boson fusion ($qq' \to qq'VV \to qq'H$) are also taken into account in this analysis.

The analysis presented in this conference looks for the Higgs boson in final states with two opposite-sign leptons ($\ell\ell' = ee$, $e\mu$ or $\mu\mu$) and missing transverse energy ($\not\!\!E_T$). The main decay mode contributing to this final state is $H \to W^+W^- \to \ell\nu\ell'\nu'$, which is the dominant decay mode for hypothetical Higgs boson masses $M_H > 135\,\text{GeV}$. However any decay mode leading to the same final state is also considered in the analysis, namely $H \to \tau^+\tau^-$ and $H \to ZZ \to \ell\ell + X$ (where $X = \ell\ell$, qq or $\nu\nu$). Hypothetical Higgs boson masses between 100 and 200 GeV are considered in this analysis, with a step of 5 GeV.

Backgrounds to the Higgs signal can be sorted in two categories: instrumental backgrounds (in which one or both leptons are faked by a jet or a photon) and physics backgrounds. Instrumental backgrounds include multijet production, which is completely estimated from data due to large uncertainties on multijet production cross-section and fake lepton identification rates, and W+jets production. For the latter background, one of the two leptons is faked by a photon or a jet, but the presence of a neutrino in the final state (from the W boson decay) can give rather large $\not\!\!E_T$, similar to the expectation from a Higgs signal.

Three processes must be considered as physics backgrounds to this analysis. The dominant one is Drell-Yan (DY) production, $q\bar{q} \to Z/\gamma^* \to \ell\bar{\ell}$. The production cross-section for this process is very large, but it is relatively easy to reject thanks to the absence of neutrino in the final state: there is be little $\not\!\!E_T$, only arising from lepton momentum mismeasurements, and the leptons are emitted back to back. Top pair production gives a final state close to the signal, with a pair of W bosons decaying leptonically, but the additional production of two b jets in this process allows to use b-tagging to reject it. At last, non-resonant diboson production is the most difficult background to reject, in particular WW production which gives the exact same final state as the signal. However the spin 0 of the Higgs boson implies angular correlations that allow some discrimination from non-resonant WW production.

This analysis uses data from the DØ experiment, one of the two general-purpose particle physics experiments (along with CDF) installed at the Tevatron. This accelerator, located next to Chicago, has provided $p\bar{p}$ collisions at $\sqrt{s} = 1.96\,\text{TeV}$ between April 2002 and September 2011, corresponding to $9.7\,\text{fb}^{-1}$ analyzed in this search. This analysis will soon be published [7], and supersedes previous DØ results in the same final states [8, 9]. The data is compared to Monte-Carlo (MC) samples of the signals and backgrounds mentioned above, scaled to the corresponding cross-section from highest order calculations available. The MC samples were generated using PYTHIA or ALPGEN + PYTHIA, followed by a detailed GEANT simulation of the detector.

The innermost part of the DØ detector [10] is composed of a central tracking system with a silicon microstrip tracker (SMT) and a central fiber tracker embedded within a 2 T solenoidal magnet. The tracking system is surrounded by a liquid-argon/uranium calorimeter with electromagnetic, fine, and coarse hadronic sections. A muon spectrometer resides outside the calorimetry and is comprised of drift tubes, scintillation counters, and toroidal magnets.

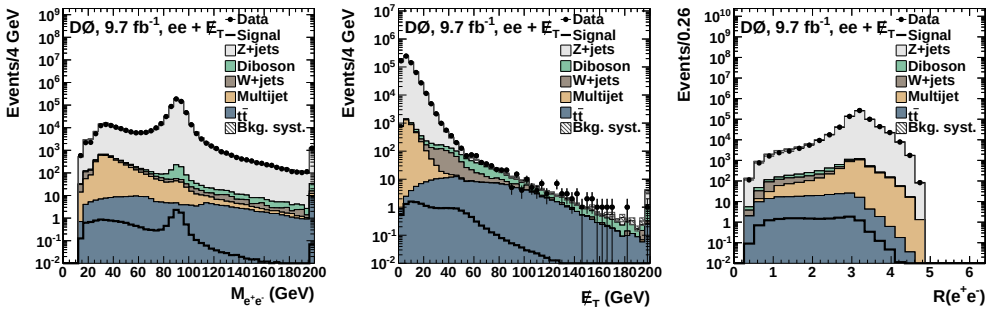

Fig. 1. – Distributions at preselection level in the dielectron channel: (left) invariant mass of the two electrons, (middle) missing transverse energy and (right) angular separation between the electrons $\mathcal{R}(e^+e^-)$

2. – Analysis techniques

In this section we will focus on the ee channel, even though the search strategy is very similar in the $\mu\mu$ and $e\mu$ channels. The analysis is performed in several steps. The first stage is called preselection: we require two opposite-sign leptons, with $p_T > 15\,\text{GeV}$ for the leading lepton and $p_T > 10\,\text{GeV}$, and $M_{ee} > 15\,\text{GeV}$. The dielectron invariant mass, the missing transverse energy and the opening angle between the electrons $\mathcal{R}(e^+e^-) = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ are shown at this stage in Fig. 1. Events are not required to pass any explicit trigger, to maximize the acceptance. However, in order to correct for any possible mismodeling of the lepton trigger and identification efficiencies, and to reduce the impact of the luminosity uncertainty, scale factors are applied to the MC so that it matches the data yield in the Z boson mass peak region.

The events are then classified into several categories. This classification enhances the sensitivity of the analysis for several reasons. First, the phase-space is divided into signal-rich and signal-poor regions, the latter allowing us to constrain the background systematics and the former being the preferred place where to look for an excess in data above background-only expectation (even though all categories are considered). This classification also allows us to identify regions dominated by different signals and backgrounds, and thus to constrain them independently.

Multivariate techniques, in particular Boosted Decision Trees (BDTs), are very efficient in classifying events and are used several times in this analysis. In a context where the signal to background ratio is very low, a single variable is not powerful enough to discriminate signal from backgrounds, and we need to combine several of them into a single discriminant, which in particular takes into account correlations between the input variables. This discriminant is trained using MC and evaluated on all samples, data and MC.

The first classification of events depends on their number of jets (these jets must arise from the same vertex as the leptons and have $p_T > 20\,\text{GeV}$): 0 jet, 1 jet, or 2 or more jets. Then, a BDT (the DY-BDT) is trained against the dominant DY background in each jet multiplicity bin and for each Higgs boson hypothetical mass. A cut is placed on this BDT to reject events that are too DY-like. Input variables to this BDT use the fact that there is no true $\not\!\!E_T$ in DY events. These variables include $\not\!\!E_T$-related variables, such as $\not\!\!E_T$ itself, its direction relative to an electron or a jet, and observables that differentiate

between real and mis-reconstructed $\not\!\!E_T$. Other kinematic variables are used, such as the electron momenta, the azimuthal opening angle between the two electrons, and the dielectron invariant mass.

Events in the 0-jet and 1-jet categories are further categorized depending of the outcome of a BDT (the WW-BDT) trained to discriminate WW production against other SM backgrounds. This allows us to define a WW-depleted and a WW-enriched regions in each of the two jet multiplicity bins, the signal being concentrated in the latter category, together with the WW background. Isolating this background allows us to better constrain the systematics associated to it (in particular its production cross-section), which are among the ones most degrading our sensitivity.

In each of the 5 categories defined so far (and for each hypothetical Higgs boson mass), a final BDT is at last trained to discriminate the signal from all SM backgrounds. The same input variables are used for the WW-BDT, which includes all the input variables of the DY-BDT, plus additional variables useful for rejecting specific backgrounds. For instance the electron quality helps rejecting W+jets events, where one of the two electrons has been misidentified; the output of the b-tagging discriminant for the jets in the events is aimed against $t\bar{t}$ production; and angular variables such as the opening angle between the two electrons in (η, ϕ) space help separate $H \to WW$ and non-resonant WW production.

3. – Results

The results are obtained by comparing the data to both background-only (H_0 hypothesis) and signal+background (H_1 hypothesis) expectations from the MC, in each final BDT output distribution. This comparison is performed by means of the log-likelihood ratio (LLR):

$$\mathrm{LLR} = -2\ln\left(\frac{P(\mathrm{data}|H_0)}{P(\mathrm{data}|H_1)}\right) \tag{1}$$

Systematic uncertainties (signal and background normalizations (cross sections), modeling effects, etc.) are taken into account as nuisance parameters in the fit. Distributions are populated with pseudo-experiments to get an estimate of significance. The observed and expected distributions of LLR are shown in Fig. 2 (left).

The LLR distributions for each hypothetical Higgs boson mass are used to draw an exclusion plot (Fig. 2, middle). This plot gives the upper 95% C.L. limit on Higgs production (cross section times branching ratio), divided by the SM prediction. The data is found to be compatible both with the background-only hypothesis and with expectation from a Higgs boson with $M_H = 125\,\mathrm{GeV}$, as shown on Fig. 2, right. The combination of the three final states (ee, $e\mu$ and $\mu\mu$) leads to the exclusion of a Higgs boson with a mass in the range $159 < M_H < 176\,\mathrm{GeV}$, while the expectation is to exclude the range $156 < M_H < 172\,\mathrm{GeV}$.

The results are also interpreted in the framework of a model with a fourth generation of fermions. If this would be the case, the gluon-fusion cross section would be much enhanced, because the new heavy quarks could be included in the loop, dominated by the top quark in the SM. This cross section could be enhanced by a factor 7 to 9, depending on the masses of the new fermions. Hence the analysis is redone by considering $gg \to H$ signal only (other signals being now negligible) and extending the mass range

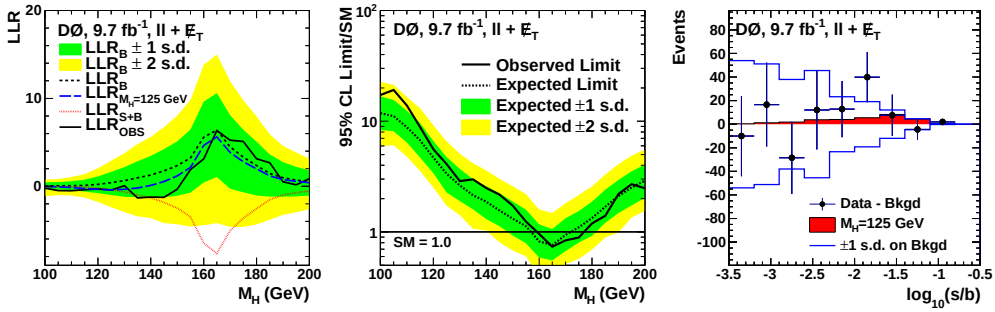

Fig. 2. – (left) The observed LLR as a function of M_H. Also shown are the expected LLRs for the background-only hypothesis, for the signal+background hypothesis, and the expectation in the presence of a signal of $M_H = 125$ GeV. (middle) Excluded cross section, $\sigma(p\bar{p} \to H + X)$, at the 95% C.L. in units of the SM cross section as a function of M_H. In the left and middle panels, the green and yellow shaded bands indicate ± 1 and ± 2 s.d. uncertainties of the expected observation for the background-only hypothesis, respectively. (right) Background-subtracted data distribution for the final discriminants, summed in bins with similar signal to background ratios, for $M_H = 125$ GeV. Also shown is the ± 1 s.d. band on the total background after fitting.

to $100 < M_H < 300\,\text{GeV}$. The exclusion plot in this theoretical framework is shown in Fig. 3.

One can also wonder if the Higgs boson could be fermiophobic. We considered a benchmark scenario, called the fermiophobic Higgs boson model (FHM), in which the couplings of the Higgs boson to other bosons are unchanged, but couplings to fermions are zero at tree level. This directly implies that the gluon fusion production process becomes much suppressed: only associated production and vector boson fusion remain possible. The branching ratio of the Higgs boson to bosons (in particular $H \to WW$) is also enhanced compared to the SM. The analysis has been redone taking these new constrains into account, and the results are presented in Fig. 4.

At last, we have performed a measurement of the non-resonant WW production cross

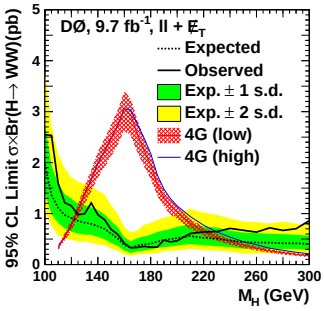

Fig. 3. – Excluded cross section $\sigma(gg \to H) \times BR(H \to WW)$ in pb as a function of M_H using all channels. The red and blue lines correspond to the theoretical prediction for a fourth generation assumption (see text) in two scenarios. The green and yellow shaded bands indicate ± 1 and ± 2 s.d. uncertainties of the background-only hypothesis, respectively.

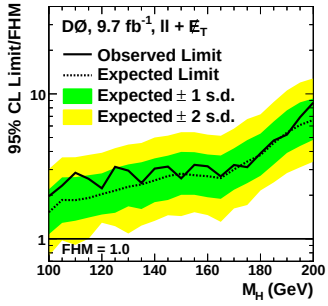

Fig. 4. – Excluded cross section, $\sigma(p\bar{p} \to H + X)$, as a function of M_H using all channels, in units of the Higgs boson production rate expected from the fermiophobic Higgs boson model (FHM).

section, which is an important cross-check of our analysis techniques. Indeed, we want to make sure we are able to model and measure a well-known SM process, our main irreducible background, with the same final state as our signal but a production cross section more than 50 times larger (for $M_H = 125\,\text{GeV}$). The same analysis methods are employed for this cross section measurement, except that the WW-BDT is used instead of the final BDT. We obtain a cross section of $\sigma_{p\bar{p}\to WW} = 11.4 \pm 0.4\,(\text{stat.}) \pm 0.6\,(\text{syst.})\,\text{pb}$, where the theory predicts $\sigma_{p\bar{p}\to WW} = 11.34 \pm 0.7\,\text{pb}$. Fig. 5 shows the combined output distribution of these discriminants, rebinned according to s/b and after the expected backgrounds have been subtracted.

To conclude, we have presented the results of the search for the SM Higgs boson in the channel with two oppositely charged leptons and missing transverse energy at DØ. A SM Higgs boson with a mass in the range $159 < M_H < 176\,\text{GeV}$ is excluded by the analysis. Non-SM models have also been considered: models with a fourth generation of fermions, and the fermiophobic Higgs model. At last, we have also measured the non-resonant WW production cross section to be $\sigma_{p\bar{p}\to WW} = 11.4 \pm 0.4\,(\text{stat.}) \pm 0.6\,(\text{syst.})\,\text{pb}$, in good agreement with the SM prediction.

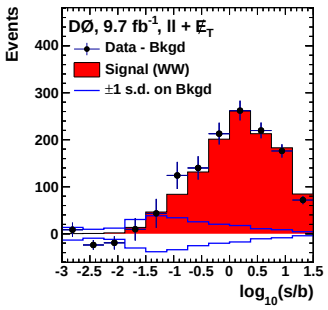

Fig. 5. – The post-fit background-subtracted data distribution for the final discriminant, summed in bins with similar signal to background ratios, for the WW cross section measurement. Also shown is the ± 1 s.d. band on the total background after fitting.

It is worth noticing that a small broad excess is visible in $H \to WW$, compatible with the background-only expectation but also with a Higgs boson with a mass $M_H = 125\,\text{GeV}$. The combination with all other searches for the Higgs boson at the Tevatron leads to a ≈ 3 standard deviations excess [11], compatible with the observation of a new boson at the LHC.

REFERENCES

[1] Higgs P. W., *Phys. Lett.* **12**, (1964) 132; Englert F. and Brout R., *Phys. Rev. Lett.* **13**, (1964) 321; Higgs P. W., *Phys. Rev. Lett.* **13**, (1964) 508; Guralnik G. S., Hagen C. R., and Kibble T. W. B., *Phys. Rev. Lett.* **13**, (1964) 585.

[2] Barate R. *et al.* [LEP Working Group for Higgs boson searches and ALEPH and DELPHI and L3 and OPAL Collaborations], *Phys. Lett. B* **565**, (2003) 61.

[3] Aaltonen T. *et al.* [CDF and D0 Collaborations], *Phys. Rev. Lett.* **104**, (2010) 061802.

[4] Aad G. *et al.* [ATLAS Collaboration], *Phys. Lett. B* **716**, (2012) 1.

[5] Chatrchyan S. *et al.* [CMS Collaboration], *Phys. Lett. B* **716**, (2012) 30.

[6] Aaltonen T. *et al.* [CDF and D0 Collaborations], *Phys. Rev. Lett.* **109**, (2012) 071804.

[7] Abazov V. M. *et al.* [D0 Collaboration], arXiv:1301.1243 [hep-ex].

[8] Abazov V. M. *et al.* [D0 Collaboration], *Phys. Rev. Lett.* **104**, (2010) 061804.

[9] Abazov V. M. *et al.* [D0 Collaboration], *Phys. Rev. D* **86**, (2012) 032010.

[10] Abachi S. *et al.*, [D0 Collaboration], *Nucl. Instrum. Methods Phys. Res. A* **338**, (1994) 185; Abazov V. M. *et al.* [D0 Collaboration], *Nucl. Instrum. Meth. in Phys. Res. A* **565**, (2006) 463.

[11] Tevatron New Physics Higgs Working Group and CDF and D0 Collaborations, arXiv:1207.0449 [hep-ex].