

## Search for the SM Higgs boson decaying to $b$ quarks with CMS experiment

S. DONATO(\*) on behalf of the CMS COLLABORATION

*Scuola Normale Superiore - Piazza dei Cavalieri 7 and  
INFN, Sezione di Pisa - Largo B. Pontecorvo 3 - Pisa, Italy*

ricevuto il 31 Luglio 2014

**Summary.** — A search for the Standard Model (SM) Higgs boson decaying to bottom quarks pairs is presented. Two production channels have been analyzed: vector-boson fusion and associated production with a vector boson decaying to leptons. The search is performed on data collected with the CMS detector at LHC during 2011 and 2012, at center-of-mass energies of 7 and 8 TeV, corresponding to integrated luminosities of about  $5.0 \text{ fb}^{-1}$  and  $19.0 \text{ fb}^{-1}$ , respectively. A 95% confidence level upper limit of 1.79 (0.89) times SM Higgs boson cross section has been observed (expected) at a Higgs boson mass of 125 GeV. An excess of events is observed above the expected background with a local significance of 2.2 standard deviations, which is consistent with the expectation from the production of the SM Higgs boson. The signal strength corresponding to this excess, relative to that of the SM Higgs boson, is  $0.97 \pm 0.48$ .

PACS 14.80.Bn – Standard-model Higgs bosons.

### 1. – Introduction

In 2012, a Higgs-boson-like particle has been discovered by CMS [1, 2] and ATLAS Collaborations [3]. The new particle has been observed mainly in the  $\gamma\gamma$ ,  $ZZ$ ,  $WW$  decay channels and its mass is about  $m_H \approx 125 \text{ GeV}$ . In addition, an evidence of the decay of the Higgs boson to a  $\tau$  pair has been recently presented [4, 5]. These measurements set constraint on the coupling of the Higgs boson to the vector bosons and the  $\tau$  lepton, and indirectly to the top quark through the loops involved in the  $\gamma\gamma$  decay and in the gluon fusion production. An observation of the  $b$ -quark decay of the Higgs boson would provide a first measurement of its coupling to the down-type quarks. This paper presents the search for the Standard Model (SM) Higgs boson decaying to bottom quarks, performed by the CMS experiment [6, 7].

(\*) E-mail: [silvio.donato@cern.ch](mailto:silvio.donato@cern.ch)

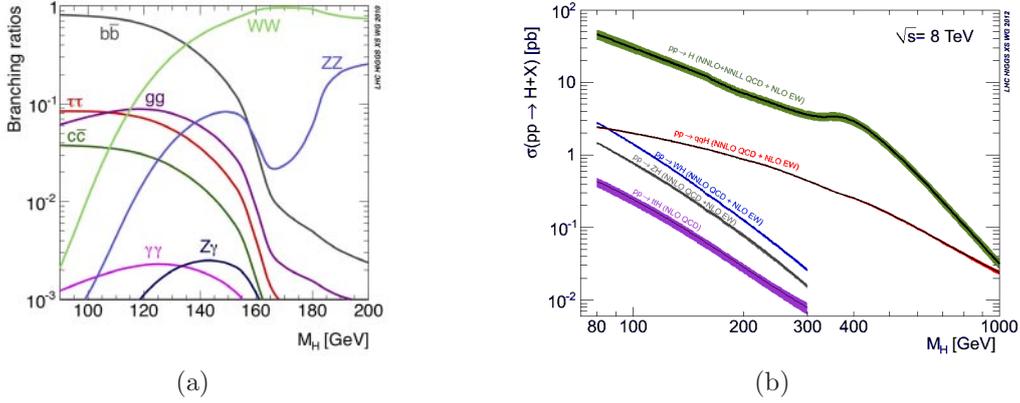


Fig. 1. – (a) The branching ratios of the Higgs boson decay as a function of its mass ( $M_H$ ). (b) The Higgs boson production cross section for proton-proton collision at  $\sqrt{s} = 8$  TeV [8].

Although the Higgs boson decay to  $b$  quark has a high branching ratio, as shown in fig. 1(a), the channel has a low sensitivity due to the large QCD background. Indeed the  $b$ -quark QCD production cross section is some  $10^8$  times larger than the Higgs boson cross section. In order to cope with such a large background the topologies of two distinctive production modes have been exploited: the vector-boson associated production (VH) and the vector-boson fusion (VBF). As shown in fig. 1(b), these production modes have about one tenth of the cross-section of the gluon-gluon fusion dominant mode, but their topologies are useful to reduce the background.

## 2. – Signal topologies

In the VH production, shown in fig. 2(a), only the leptonic decays of the vector bosons are considered. Here, the topology is defined by the presence of two  $b$  jets from the Higgs boson decay, and at least one isolated charged lepton or large missing transverse energy (MET). These requirements make the QCD background negligible.

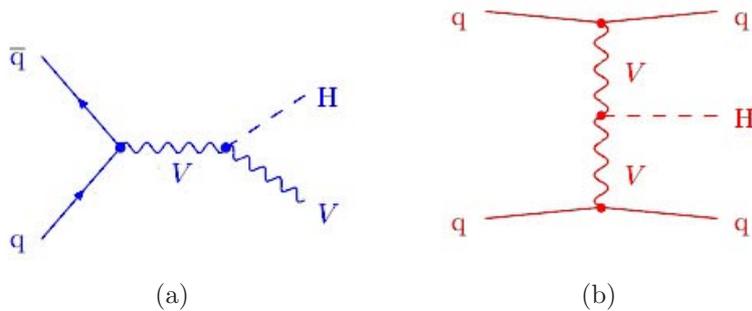


Fig. 2. – (a) The Feynman diagram of the Higgs boson production with vector bosons. (b) The Feynman diagram for the vector-boson-fusion process with the Higgs boson production.

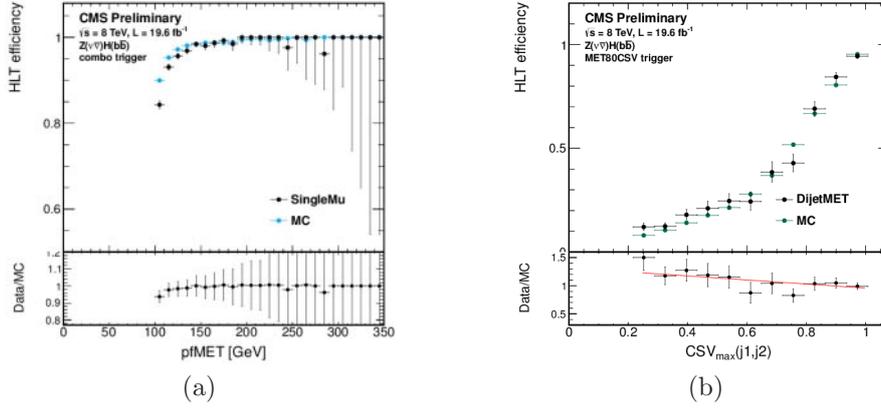


Fig. 3. – (a) The MET trigger efficiency as a function of the offline MET. (b) The MET+ $b$ -tag trigger efficiency as a function of the offline  $b$ -tag discriminator (CSV).

The VBF signal, shown in fig. 2(b), is characterized by the presence of two  $b$  jets from the Higgs boson decay and two energetic quark-jets with a large  $\eta$  separation. No gluons participate in the interaction, so a low additional hadronic activity is expected. In this channel the QCD background remains the main background.

### 3. – Triggers

The VH triggers accepted the events that fulfill at least one of the following requirements:

- i) two isolated electrons;
- ii) one isolated muon;
- iii) one isolated electron and MET;
- iv) MET.

The thresholds of these triggers changed during data taking, depending on the luminosity. In 2012, in order to have a larger acceptance, a MET trigger exploited online  $b$ -tagging. The  $b$ -tagging reduced the background rate, while keeping a high signal efficiency. This allowed to decrease the threshold in online MET from 100 GeV to 80 GeV. Figure 3(a) shows the trigger efficiency as a function of the offline MET. Figure 3(b) shows the efficiency as a function of the offline  $b$ -tag discriminator (CSV): the trigger had a high efficiency when  $b$  jets were presents (CSV  $\sim$  0.95), while rejecting light-jet events (CSV  $\sim$  0.2).

The VBF triggers were based on the presence of four energetic jets, with a pair of jet having a high invariant mass. Some triggers used the online  $b$ -tagging in order to reduce the rate due to gluon/light quark jet background.

### 4. – Signal and background regions

The main processes that can simulate the VH signal topology are:  $W/Z +$  jets and  $t\bar{t}$  production. Their shapes are taken from simulation whereas their normalizations are

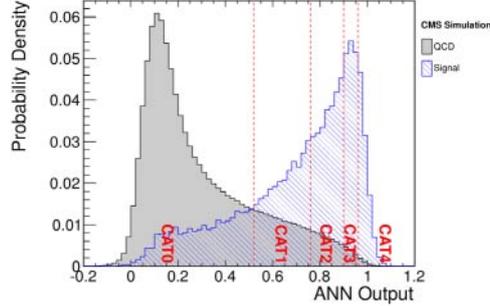


Fig. 4. – Probability distribution of the ANN output for signal and background, for VBF analysis. The vertical dashed lines define the signal regions used in the analysis.

data-driven. The analysis is divided in six sub-channels, according to the vector-boson decay mode:  $W \rightarrow e\nu$ ,  $\mu\nu$ ,  $\tau(1\text{-prong})\nu$  and  $Z \rightarrow ee$ ,  $\mu\mu$ ,  $\nu\nu$ . In addition the channels are binned in two or three vector-boson  $p_T$  bins. In each bin a signal region is defined cutting on: jet kinematic variables,  $b$ -tagging discriminants, lepton momentum and/or MET, number of additional leptons and jets. Inverting some cuts, up to five control regions are defined for  $t\bar{t}$  and  $W/Z$  + heavy/light quark jets backgrounds. They are used to evaluate up to seven scale factors to apply to  $t\bar{t}$  and  $W/Z$  + 0/1/2  $b$ -jets background normalizations.

The main background in the VBF analysis is the multi-jet QCD production. This background is estimated directly by data. Minor backgrounds are  $W/Z$  + jets and  $t\bar{t}$  productions and they are taken from simulation. The signal regions are defined using an Artificial Neural Network (ANN). It is trained to separate the signal from the backgrounds, using simulations. With the exception of the  $b$ -jets kinematic it exploits the most discriminants variables:  $\Delta\eta$  between the most forward/backward jets,  $b$ -tagging discriminants and additional hadronic activity in the event. Five signal regions are defined using the ANN output, as shown in fig. 4. In each region the background and the signal yield are obtained with a fit to the  $b\bar{b}$  jet mass distribution. The QCD background is extrapolated from the side bands using a fifth-degree Bernstein polynomial, while the signal and the minor backgrounds shapes are taken simulations.

## 5. – Jet energy regression

In both analyses a jet energy regression is used in order to improve the  $b\bar{b}$  jet mass resolution. The regression attempts to recalibrate the  $p_T$  to the true  $p_T$  of the particle jet exploiting variables including:

- $b$ -tag discriminator and variables related to the secondary vertex, when it is present;
- jet-energy fractions between charged/neutral hadron and photons, muons and electrons;
- jet kinematic variables ( $p_T$ ,  $\eta$ );
- MET (only for  $ZH \rightarrow \ell\ell b\bar{b}$  and VBF).

As shown in fig. 5, the regression improves the  $b\bar{b}$  jet mass resolution by about 10–20%.

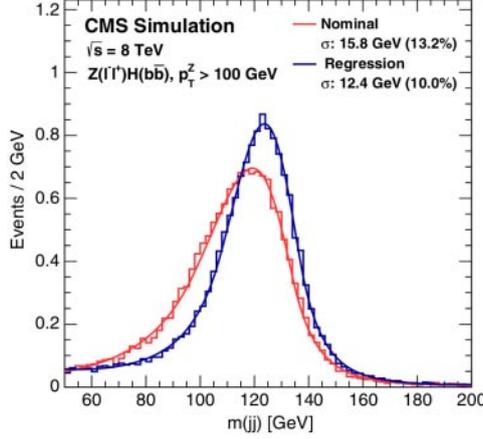


Fig. 5. – Higgs candidate mass distribution in  $Z(\ell\ell)H(bb)$  simulation at  $\sqrt{s} = 8$  TeV, in the bin with  $p_T^{(Z)} > 100$  GeV. The red (blue) histogram shows the distribution before (after) the jet energy regression is applied. Here the regression improves the mass resolution of about 20%.

## 6. – Multi BDT (VH)

In order to reduce the background in VH analysis, three specialized BDT are trained to reject the  $t\bar{t}$ ,  $W/Z$  + jets and  $WW/WZ/ZZ$  backgrounds. The BDT variables are:

- Higgs candidate mass and  $p_T$ ;
- $b$ -tag discriminants;
- lepton momentum and MET;
- number of additional leptons and jets;
- other kinematic variables.

The final multi-BDT score distribution is realized as following. An event rejected by the  $t\bar{t}$  BDT gets a score between  $-1$  and  $-0.5$ . The other events that fail the  $W/Z$  + jets BDT have a score between  $-0.5$  and  $0$ . Again, the other events that fail the  $WW/WZ/ZZ$  BDT have a score between  $0$  and  $0.5$ . A final BDT is applied to reject all backgrounds and it assigns a score between  $0.5$  and  $1$ . In this way the combined multi-BDT is more powerful than the classic one-step BDT. An example of a multi-BDT score distribution is shown in fig. 6.

The last step is the extraction of the signal. In the VH analysis this is obtained with a fit of the multi-BDT score distribution using the shapes from simulation and the data-driven scale factors.

## 7. – A cross-check: the $Z$ peak fit

The  $Z$  boson is produced by the same mechanisms (VBF and of VZ associate production) used in this analysis for the Higgs boson search. The  $Z$  boson decays into  $b\bar{b}$  pairs with large ( $\sim 15\%$ ) branching fraction. A natural cross check for this analysis is to search for a  $Z$  boson peak in the  $b\bar{b}$  jet invariant mass spectrum.

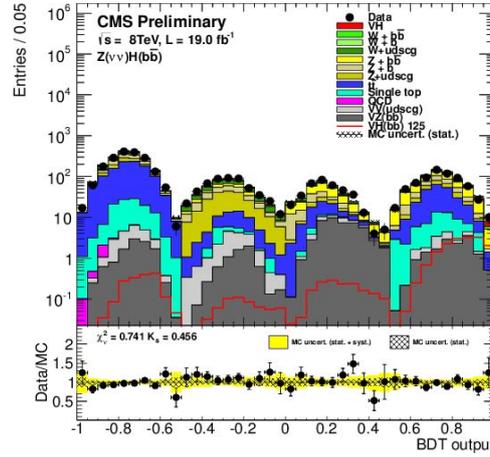


Fig. 6. – The multi-BDT score distribution for the  $Z(\nu\nu)H(bb)$  channel, in the bin with  $MET > 170$  GeV.

For the VH analysis, this search is performed retraining a new multi-BDT, taking the  $Z(bb)$  process as signal. Figures 7(a) and (b) show the final  $Z(bb)$  fits performed for VH and VBF analysis, respectively. The measured  $Z(bb)$  signal strengths,  $\mu = \sigma_{measured}/\sigma_{SM}$ , are:  $\mu = 1.19^{+0.28}_{-0.22}$ , for VH analysis, and  $\mu = 0.99^{+0.12}_{-0.12}$ , for VBF analysis.

## 8. – Results

In the VH analysis an upper limits of 1.89 (0.95) times the SM Higgs boson cross section at 125 GeV with 95% CL has been observed (expected), as shown in fig. 8(a). Corresponding to an excess of events of 2.1 standard deviations and to a signal strength

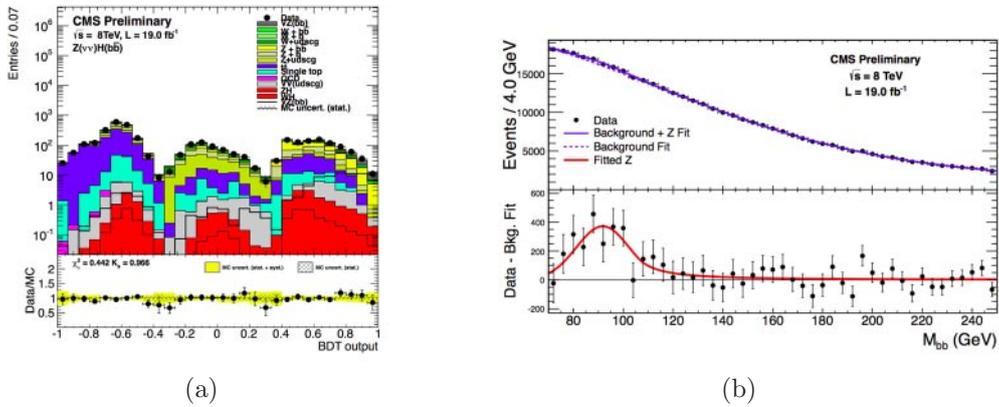


Fig. 7. – (a) The multi-BDT score distribution for the search of the  $Z(bb)$  peak in the VH analysis. (b) The  $b$ -jet pair invariant mass distribution for the search of the  $Z(bb)$  peak in the VBF analysis.

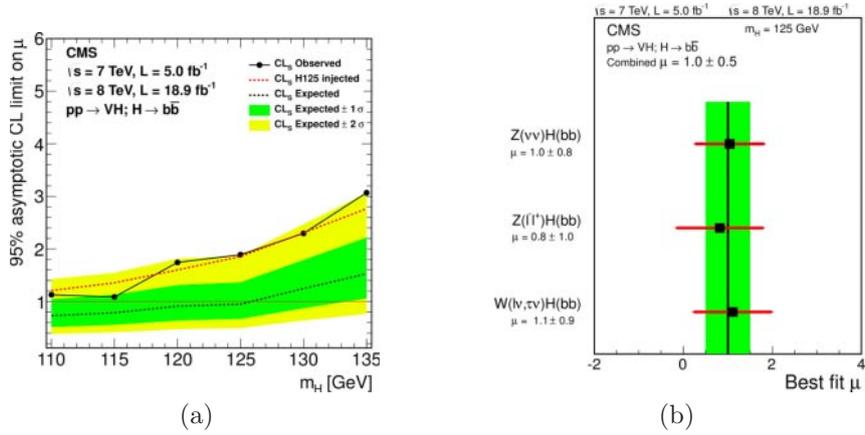


Fig. 8. – (a) The 95% CL upper limits on the signal strength for the SM Higgs boson hypothesis as a function of the Higgs boson mass, for the VH analysis. (b) The measured signal strengths for the VH sub-channels.

of  $\mu = 1.0 \pm 0.5$ . Figure 8(b) shows that compatible results are obtained for all sub-channels. Figure 9 shows the distribution of the  $bb$  dijet invariant mass for the VH analysis, combining all sub-channels.

In the VBF channel an upper limit of 3.6 (3.0) times the SM Higgs boson cross section at 125 GeV with 95% CL has been observed (expected), as shown in fig. 10(a). It corresponds to a signal strength of  $\mu = 0.7 \pm 1.4$ . Figure 10(b) shows the  $bb$  dijet invariant mass distribution in the most sensitive signal region.

A combination of the two analysis gives a signal strength of  $\mu = 0.97 \pm 0.48$ . It corresponds to an excess of 2.2 standard deviations from the expected background, as shown in fig. 11.

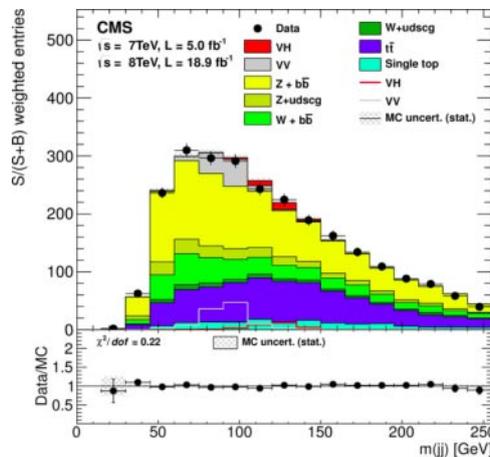


Fig. 9. – The  $b$ -jet pair invariant-mass distribution in the VH analysis.

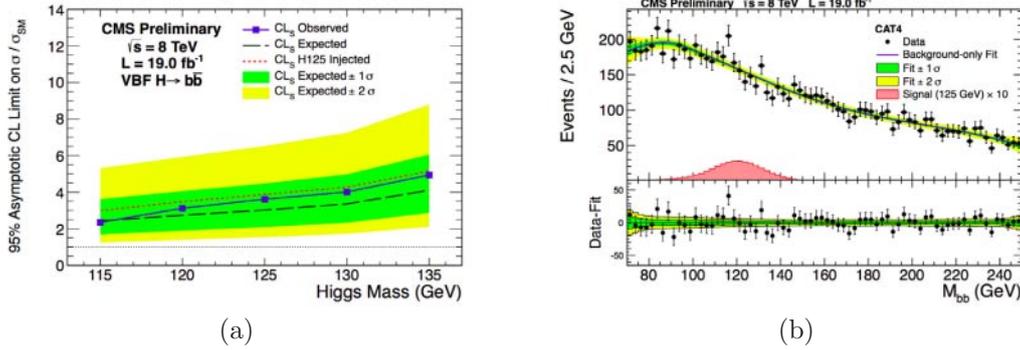


Fig. 10. – (a) The 95% CL upper limits on the signal strength for the SM Higgs boson hypothesis as a function of the Higgs boson mass, for the VBF analysis. (b) The  $b$ -jet pair invariant mass distribution in the VBF analysis.

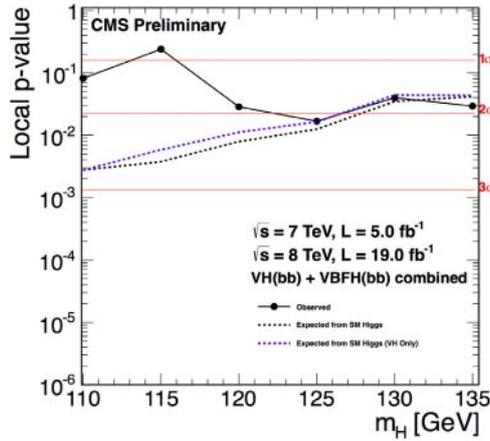


Fig. 11. – The  $p$ -values, as a function of the Higgs mass, of VBF and VH combined analysis. The dashed lines show the expected values in Higgs boson hypothesis, for the combined (black) and VH only (blue) analysis.

### 9. – Conclusions

A search for the standard model Higgs boson decaying to bottom quarks has been presented. Two production channels have been studied: the associated production with vector boson decaying to leptons (VH) and the vector-boson fusion (VBF). An excess of 2.2 standard deviations at a mass of 125 GeV has been reported that corresponds to a signal strength of  $\mu = 0.97 \pm 0.48$ .

## REFERENCES

- [1] CHATRCHYAN S. *et al.* (CMS COLLABORATION), *JINST*, **3** (2008) S08004.
- [2] CHATRCHYAN S. *et al.* (CMS COLLABORATION), *Science*, **338** (2012) 1569.
- [3] AAD G. *et al.* (ATLAS COLLABORATION), *Science*, **338** (2012) 1576.
- [4] CHATRCHYAN S. *et al.* (CMS COLLABORATION), CMS-HIG-13-004, CERN-PH-EP-2014-001.
- [5] AAD G. *et al.* (ATLAS COLLABORATION), ATLAS-CONF-2013-108.
- [6] CHATRCHYAN S. *et al.* (CMS COLLABORATION), *Phys. Rev. D*, **89** (2014) 012003.
- [7] CHATRCHYAN S. *et al.* (CMS COLLABORATION), CMS-PAS-HIG-13-011.
- [8] DITTMAIER S., MARIOTTI C., PASSARINO G. and TANAKA R. (LHC HIGGS CROSS SECTION WORKING GROUP) (Editors), CERN-2012-002 (CERN, Geneva, 2012).