Search for the Standard Model Higgs boson produced by vector boson fusion and decaying to bottom quarks

LHCP2015, Saint Petersburg (Russian Federation) | 31 Aug-5 Sep 2015

Giorgia Rauco - University of Zürich, Switzerland

Sara Alderweireldt, Tom Cornelis, Xavier Janssen, Jasper Lauwers, Nick van Remortel - University of Antwerp, Belgium
Paolo Azzurri, Silvio Donato - INFN, Scuola Normale Superiore and University of Pisa, Italy
Konstantinos Kousouris - CERN, Meyrin, Switzerland
Caterina Vernieri - Fermilab, Batavia, IL USA

Abstract

A search for a Standard Model Higgs boson in the vector boson fusion production mechanism with decay to bottom quarks is presented. The search analyzes two data samples of proton-proton collisions at $\sqrt{s}$ = 8 TeV, collected with the CMS detector during 2012, comprising of 19.8 fb^{-1} (prompt) and 18.3 fb^{-1} (parked). Upper limits on the product of the cross section and the branching fraction into a bottom quark pair, as well as the fitted signal strenght relative to the expectation for the standard model Higgs boson, are derived in the Higgs boson mass range from 115 to 135 GeV. In addition, the combination of this result with other CMS searches for the Higgs boson in the same decay channel is reported.

Introduction

At LHC a Standard Model Higgs boson can be produced at through various mechanisms and **vector boson fusion** (VBF) is the second one in order of production cross section (Fig. 1). Additionally, for a Higgs boson mass of 125 GeV, its **dominant decay mode is in a pair of $b\bar{b}$** (Fig. 2).

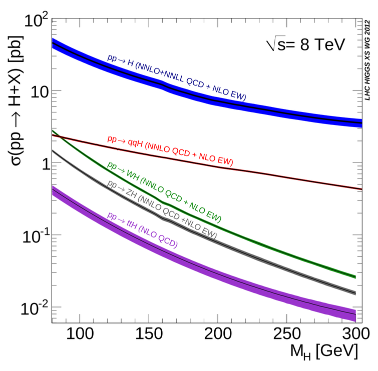


Figure 1 : Standard Model Higgs boson production cross section at $\sqrt{s}$=8TeV.

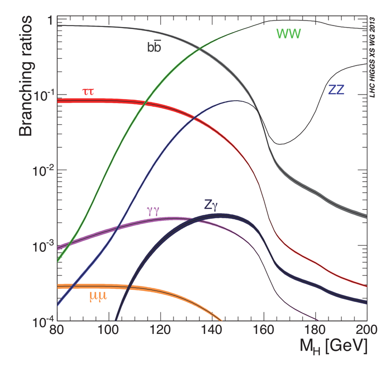


Figure 2 : Standard Model Higgs boson decay branching ratios.

The main challenges of the search in the VBF H→ $b\bar{b}$ channel (Fig. 3) are the **large QCD background** and the implementation of a **dedicated trigger**. However, the search can be performed exploiting the very particular topology of the VBF process.

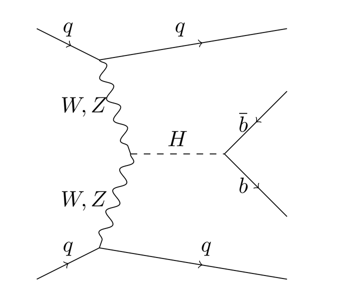
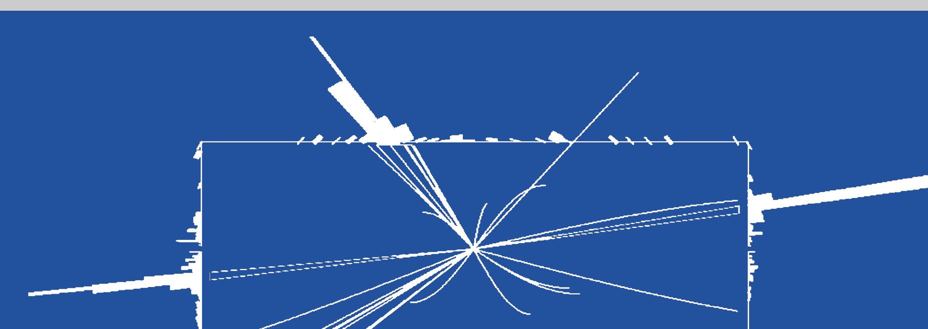


Figure 3 : Feynman diagram for the process VBF H → $b\bar{b}$

- **4-jets final state**:
  - 2 VBF legs (large $\Delta\eta$ and m_{qq})
  - 2 b-jets
- suppressed colour flow between the VBF jets (central rapidity gap)

The search strategy is essentially based on:

1. topological trigger on the signal main properties
2. use multivariate methods to discriminate S/B
3. perform a fit on the $m_{b\bar{b}}$ spectrum



Samples, Triggers and Event selection

Samples:

Set A (Primary data sample)
- 2012 nominal dataset, with 19.8fb^{-1}
- using dedicated triggers

Set B (Additional data sample)
- 2012 parked dataset, with 18.3fb^{-1}
- using general purpose trigger

Triggers:

Dedicated trigger
- L1: p_T cuts for the three leading jets + at least two jets central ($|\eta| \leq 2.6$)
- HLT: cuts on the 4 leading Calo/PF jet p_T + b-tagging + VBF kinematics

General-purpose trigger
- L1: dijet + additional kinematic requirements
- HLT: at least 2 calo jets with $p_T > 35$ GeV + VBF kinematics

Event selection:

- at least four reconstructed PF-jets are requested
- the four p_T-leading ones are considered as the most probable signal jets candidates

	set A	set B
trigger	dedicated	general-purpose
jets p_T	$p_T^{1,2,3,4} > 80, 70, 50, 40$	$p_T^{1,2,3,4} > 30$ $p_T^1 + p_T^2 > 160$
jets $\|\eta\|$	< 4.5	< 4.5
b tag	2 loose b-tags	at least 1 medium and 1 loose b-tags
$\Delta\phi_{bb}$	< 2.0 radians	< 2.0 radians
VBF topology	$m_{qq} > 250$ $\|\Delta\eta_{qq}\| > 2.5$	$m_{qq}, m_{jj}^{trig} > 700$ $\|\Delta\eta_{qq}\|, \|\Delta\eta_{jj}^{trig}\| > 3.5$
veto	none	events that belong to set A

Table 1 : Summary of selection requirements for the two analysis sets.

Event Properties

Certain characteristic properties of the final state allow a significant improvement of the overall sensitivity:

1. b-jet energy regression

- multidimensional calibration targeting the jet p_T at generator level
- provide a corrective factor to the energy of b-jets
- improvement of the dijet invariant mass resolution by ∼ **17%** (Fig. 4)

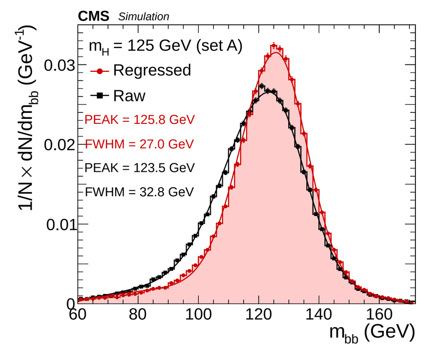


Figure 4 : Improvement of the dijet invariant mass.

2. quark/gluon jets discriminator (QGL)

- QCD background: jets originating from gluons are dominant
- VBF signal electroweak pp interaction: only quark initiated jets
- differences in jet composition and structure are exploited (Fig. 5)

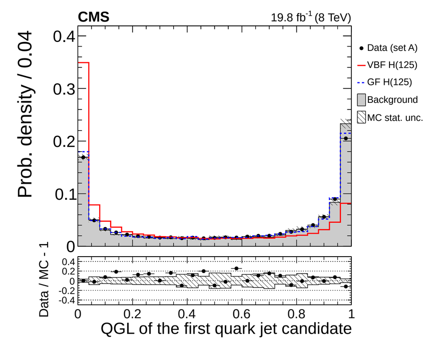


Figure 5 : First quark jet candidate.

3. soft QCD activity

- VBF signal events electroweak production of jets (rapidity gap of suppressed activity between the two VBF tagging jets)
- construction of "soft" hadronic activity observables clustering "additional tracks" from the main interaction vertex (Fig. 6)

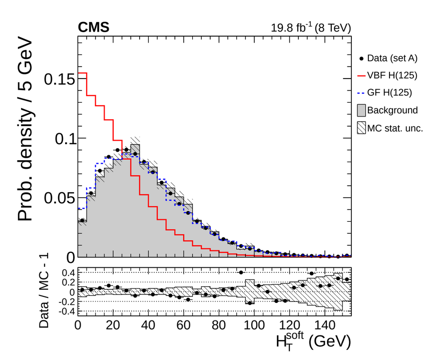


Figure 6 : Soft H_T.

Higgs boson signal vs. background estimation

In order to separate the overwhelmingly large QCD background from the Higgs boson signal, all discriminating variables (and their correlations) have to be used in an optimal way. This is best achieved by using a **multivariate discriminant**. The correlation between the chosen variables and the invariant mass between the two b-jets has to be small.

The used variables are conceptually grouped into five groups:

VBF topology (m_{qq}, $\|\Delta\eta_{qq}\|$, $\|\Delta\eta_{bb}\|$)	**quark-gluon separation** (QGL tags)
b-tagging (2 leading CSV b-tags)	**soft-activity** (H_T^{soft}, N_2^{soft})
angles in CM frame (4 jets CM frame $\theta(qq, bb)$)	

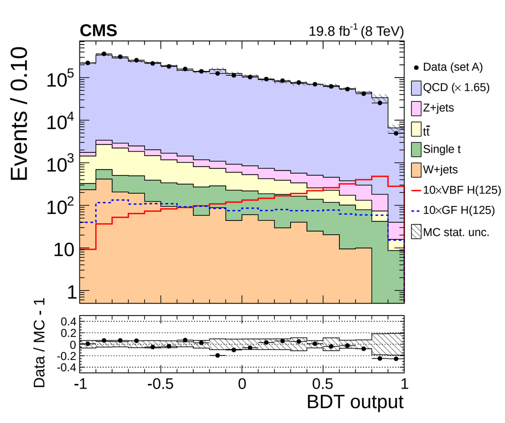


Figure 7 : BDT distribution for Set A, with datapoints overlaid on stacked simulated backgrounds.

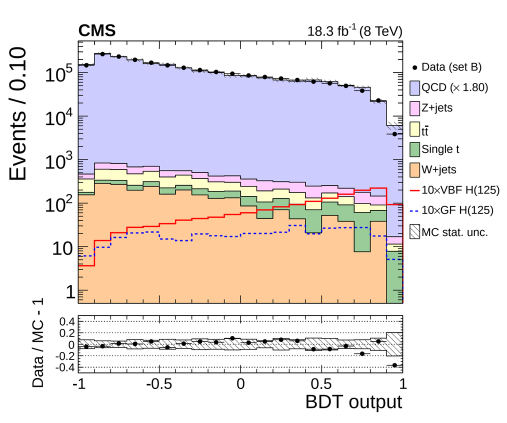


Figure 8 : BDT distribution for Set B, with datapoints overlaid on stacked simulated backgrounds.

In order to **maximize the signal sensitivity** events are divided into **9 categories** depending on the value of the BDT output (Fig. 7-8). Separation values between the categories are chosen in to minimize the expected limit, trying to keep a sufficient statistics. Lower categories are background dominated, while higher ones are signal enriched.

Fit to data

The shape for the **QCD** background is extrapolated with data-driven method from the background dominated event categories. The contributions from the **top** and the **Z+jets** are extracted from simulation. The signal is taken from the simulation and it is parametrized as a Crystal ball function on top of a polynomial background.

The **model is fitted simultaneously in all categories**:

$$f(m_{b\bar{b}}) = \mu_{i,H} \cdot N_{i,H} \cdot H_i(m_{b\bar{b}}; k_{JES}, k_{JER}) + N_{i,Z} \cdot Z_i(m_{b\bar{b}}) + N_{i,top} \cdot T_i(m_{b\bar{b}}) + N_{i,QCD} \cdot R_i(m_{b\bar{b}}) \cdot Q(m_{b\bar{b}}; \vec{p})$$

using transfer functions R_i , floating normalizations N_i for top and Z, and nuissance parameters k_{JES} and k_{JER} and signal strength $\mu_{i,H}$ for the Higgs boson.

The fit procedure is **validated by fitting the known Z resonance** (Fig. 9). The best fitted signal strength is $\mu_Z = \sigma/\sigma_{SM} = 1.28^{+0.50}_{-0.34}$ with an observed (expected) significance of 3.8σ (3.2σ).

A **binned maximum-likelihood fit of the b-quark pair invariant mass distribution** for the Higgs boson signal is performed. All categories of the dataset are simoultaneously fitted (Fig. 10-11).

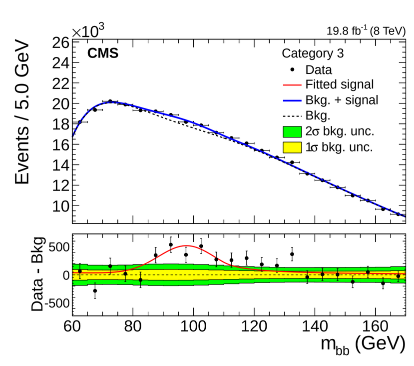


Figure 9 : b-quark pair invariant mass distribution for the Z boson signal, for the signal-enriched event category.

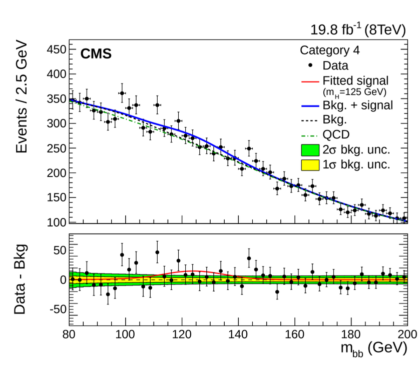


Figure 10 : b-quark pair invariant mass distribution for the Higgs boson signal, for the signal-enriched event category in Set A.

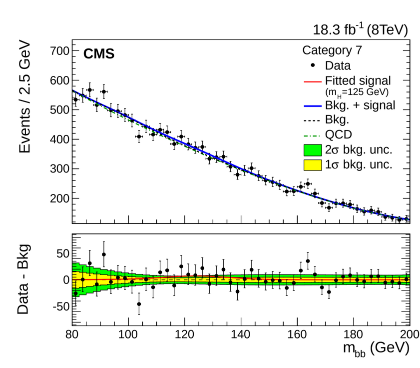


Figure 11 : b-quark pair invariant mass distribution for the Higgs boson signal, for the signal-enriched event category in Set B.

Results

The b-quark pair invariant mass distributions are fitted simulataneously in all categories, and 95**%** asymptotic confidence level limits on the signal strength are derived as a function of the Higgs boson mass (Fig. 12).

- a **signal excess** is observed
- for a Standard Model Higgs boson with mass 125 GeV the observed (expected) significance found is 2.2 (0.8) standard deviations and the fitted signal strength is $\mu = \sigma/\sigma_{SM} = 2.8^{+1.6}_{-1.4}$

Combination with other Standard Model CMS searches for H→ $b\bar{b}$:

The VBF H→ $b\bar{b}$ results have been combined with those of other CMS H→ $b\bar{b}$ searches (Fig. 13).

H → $b\bar{b}$ channel	best-fit (**68%** CL) Observed	Upper Limits (**95%** CL) Observed	Upper Limits (**95%** CL) Expected	Signal significance Observed	Signal significance Expected
VH	0.89 ± 0.43	1.68	0.85	2.08	2.52
ttH	0.7 ± 1.8	4.1	3.5	0.37	0.58
VBF	$2.8^{+1.6}_{-1.4}$	5.5	2.5	2.20	0.83
combined	$1.03^{+0.44}_{-0.42}$	1.77	0.78	2.56	2.70

Table 2 : Observed and expected 95**%**CL limits, best fit values and significance on the signal strength parameter $\mu = \sigma/\sigma_{SM}$ at m_H = 125 GeV, for each H → $b\bar{b}$ channel and combined.

The fitted signal strength of the combination for m_H = 125 GeV is $\mu = 1.03^{+0.44}_{-0.42}$, with a significance of 2.6σ.

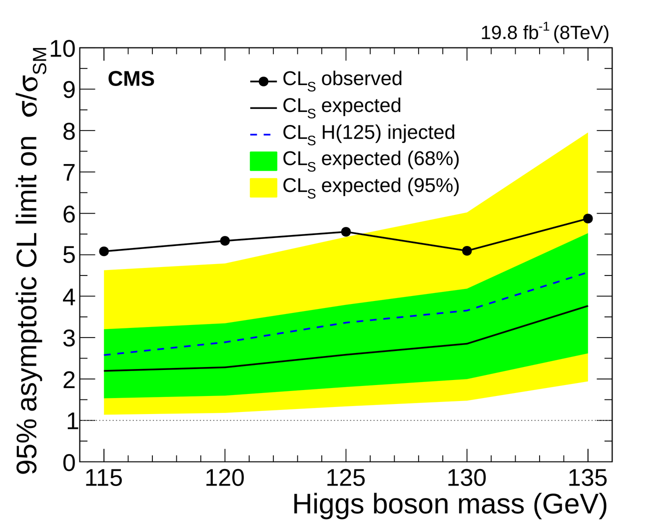


Figure 12 : 95**%** asymptotic confidence level limits on the signal strength in function of the Higgs boson mass.

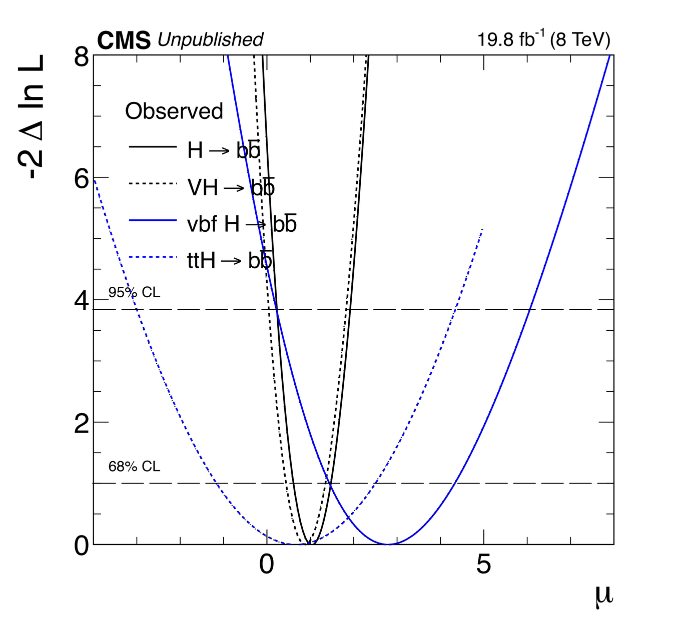


Figure 13 : Observed likelihood for the VH, ttH and VBF production mode, with H → $b\bar{b}$.

Reference

CMS Collaboration, *Search for the standard model Higgs boson produced through vector boson fusion and decaying to $b\bar{b}$*, 2015, CERN-PH-EP-2015-121, CMS-HIG-14-004, arXiv:1506.01010

Giorgia Rauco (giorgia.rauco@cern.ch)