Winter LHCC meeting, 27 Feb 2019, CERN, Geneva (Switzerland)

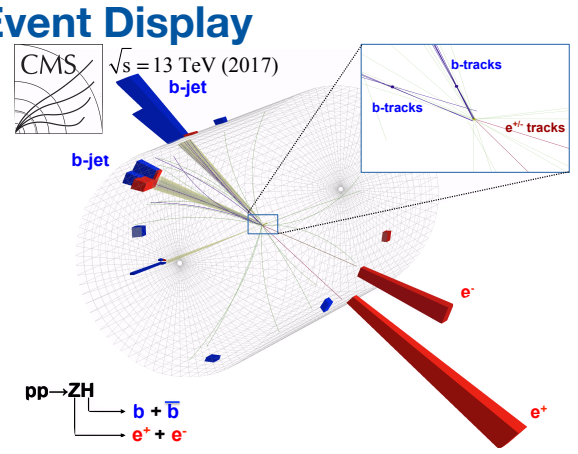
Observation of the Standard Model Higgs Boson Produced in Association with a Z or W Boson, and Decaying to Bottom Quarks.

Gaël L. Perrin, on behalf of the CMS Collaboration

Motivations

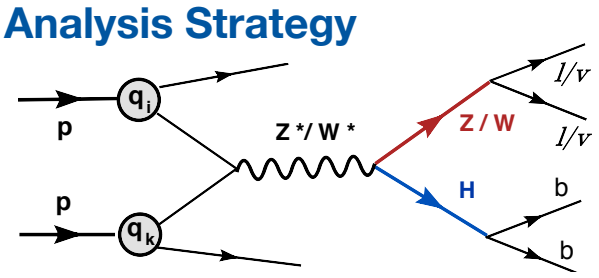
- The **Higgs boson** was discovered in 2012 by the CMS and ATLAS collaborations at CERN, using the decay channels (H→γγ, H→ZZ*). A measurement of the **Yuakawa coupling between the Higgs and the b quark** is an important test of the Standard Model of particle physics (SM).
- In the SM, a Higgs boson with a mass 125 GeV decays to bottom quarks with the **largest BR: 58%**.

Event Display



An event candidate for the production of a Z boson in conjunction with a Higgs boson in the CMS detector.

Analysis Strategy



The search is performed on **Higgs decaying to two b quarks produced with an associated vector boson (VH)** decaying in **three different channels**:

- **Z→ ll** (e^+e^- or $\mu^+\mu^-$)
- **Z→ νν**
- **W→ lν** (eν or μν)

The **Higgs and vector boson requires to be boosted**. This selection reduces large backgrounds from W and Z production in association with jets and top quark production, makes accessible the Z(νν)Hbb channel via large missing transfers energy and improves mass resolution of the Higgs candidates.

The dominant backgrounds are coming from:

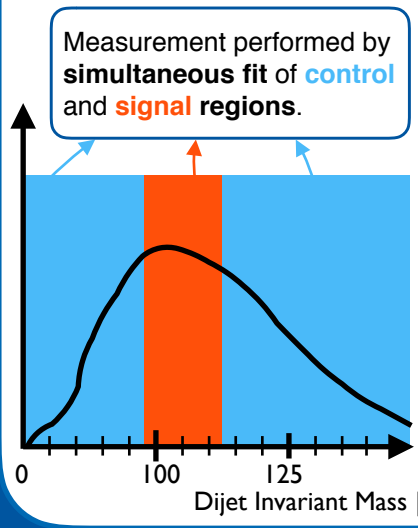
- production of a Z/W boson associated with jets
- production of vector boson pairs
- production of top quark pair or a single top quark.

Signal and Control Regions

Signal Regions (SR): selection defined to optimise the signal purity.

Control Regions (CR): selection defined to optimise the purity of specific backgrounds, used to study the modelling of the simulation.

Scale Factors (SF): correct differences in normalisation and shape between data and MC and are derived by performing a simultaneous fit on the SR and CR.

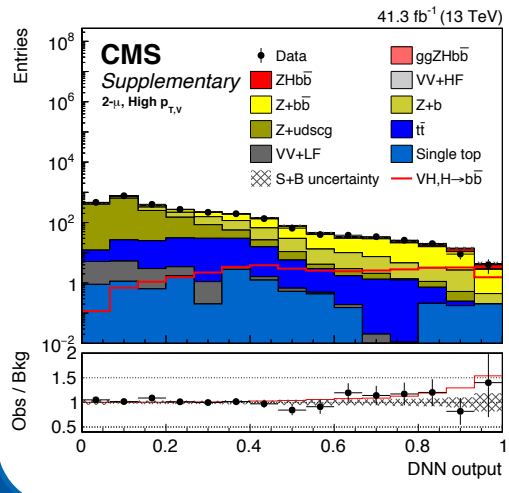


The dijet invariant mass (the reconstructed invariant mass of the two b jets) is one of the most discriminating variable and peaks around 125 GeV for a SM Higgs boson.

The signal regions are defined around the 125 GeV mass peak, the control regions in the mass sidebands.

Deep Neural Network

In the **signal regions**, the main method to discriminate signal from background events is a **Deep Neutral Network** (DNN), a multivariate analysis technique trained separately for each decay channel on MC using a set of discriminating variables, as the dijet mass, to separate signal from background events. A **maximum likelihood fit** is then applied on the DNN output to **extract** the observed **signal contribution**.

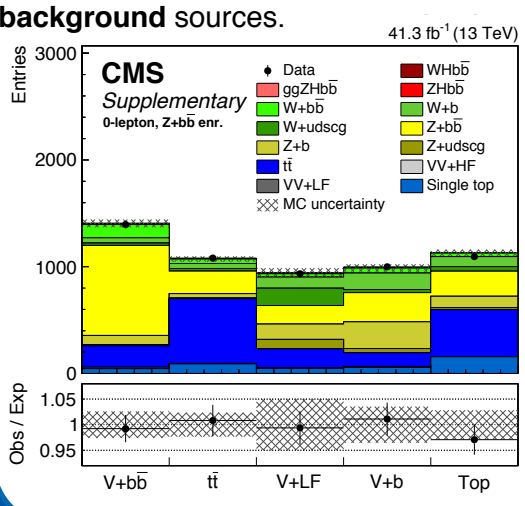


DNN output trained on events with Z decaying to two muons. Events with low (large) DNN output values are more likely to be background (signal) events. The data correspond to the 2017 dataset.

Multiclassifier

In the **control regions**, the discrimination between the background sources is performed by a **multi-output DNN (mutliclassifier)**.

The multiclassifier is trained on MC using the same set of discriminating variables as the signal region DNN and returns 5 probabilities per event, one per background category. It is **included in the maximum likelihood fit** to **constrain** the **normalisation uncertainties** of the **main background** sources.

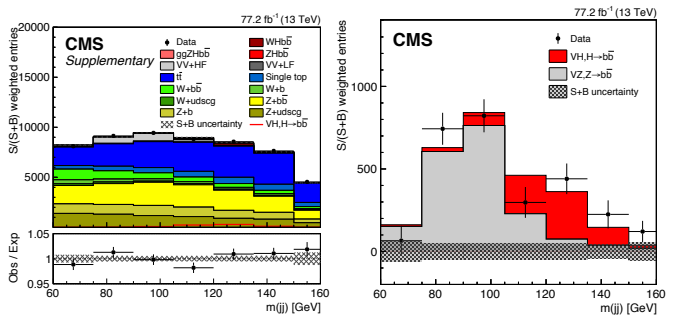


Example of multiclassifier distributions on the 2017 dataset in the Z→ νν channel, in a control regions enriched in Z boson associated with two jets.

Mass Analysis

The **mass analysis** is performed as a **cross-check** to the DNN-based analysis. Events are separated in four categories based on a massless DNN, similar to the main analyses DNN except the variables are not correlated to the dijet mass **(mjj)**. This avoids a sculpting of the background shapes.

The dijet mass distributions are fitted simultaneously in the four categories in the signal regions, together with the multiclassifier in the control regions.

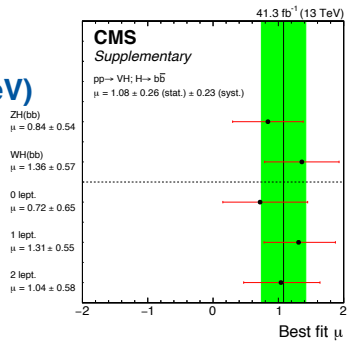


Combined invariant dijet mass distributions for all categories and channels. Data and MC events are both reweighed by a S/(S+B) from their corresponding DNN category.
left: all background included.
right: subtract all backgrounds but VZ on data and MC.

Results

Using 2017 dataset (41.3fb^{-1} at 13 TeV)

- Best fit of signal strength for the combination of the three channels: **μ = 1.08 ± 0.34**
- **observed** (expected) **significance**: **3.3 σ** (3.1 σ)



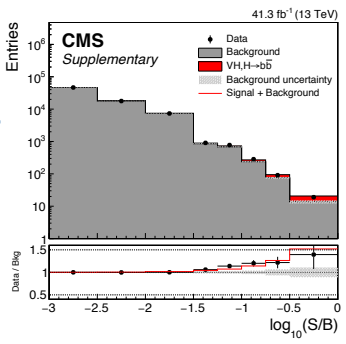
The best-fit signal strength and uncertainty per-channel and for the WH and ZH processes, extracted from a simultaneous fit of all channels for the 2017 analysis.

Using the Run1 + Run2 dataset

- **μ = 1.01 ± 0.23**
- **obs.** (exp.) **sig.** : **4.8 σ** (4.9 σ)

Mass analysis using 2016 + 2017 dataset (77.2fb^{-1} at 13 TeV)

- **μ = 1.16 ± 0.29**
- **obs.** (exp.) **sig.** : **3.8 σ** (3.3 σ)



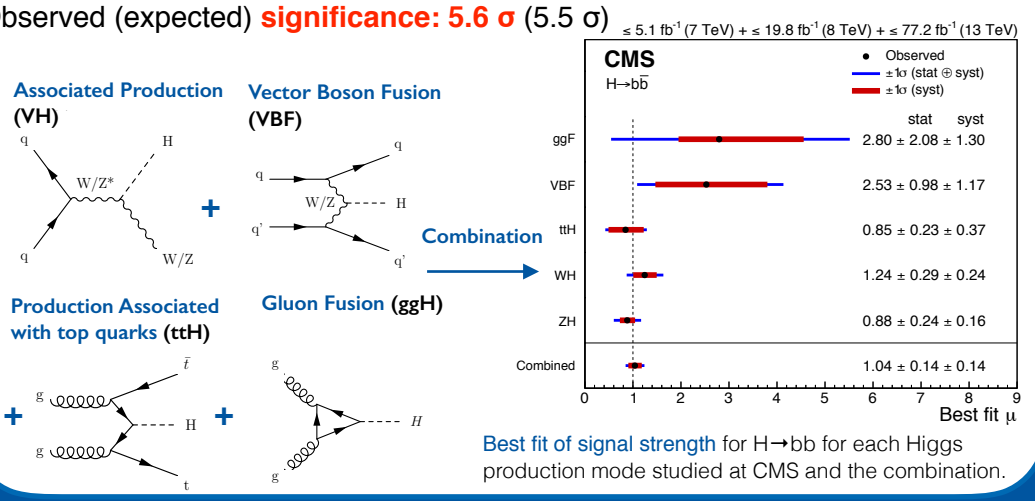
Distributions of signal, background, and data event yields sorted into bins of similar signal-to-background ratio, as given by the result of the fit to their corresponding DNN. All events in the 2017 signal regions are included.

Combination with other Production Channels

To study the **decay H→bb**, this **VH** study is **combined with** CMS results for **other Higgs production** modes: vector boson fusion, production associated with top quarks and gluon fusion.

Result: μ = 1.08 ± 014 (stat) ± 014 (sys)

Observed (expected) **significance: 5.6 σ** (5.5 σ)



Best fit of signal strength for H→bb for each Higgs production mode studied at CMS and the combination.

References:

• Observation of Higgs Boson Decay to Bottom Quarks, CMS Collaboration, Phys. Rev. Lett. **121**, 121801