PROCEEDINGS OF SCIENCE

Inclusive search for a boosted Higgs boson and observation of the Z boson decaying to charm quarks with the CMS experiment

Andrzej Novák on behalf of the CMS Collaboration[a,*]

[a] *RWTH Aachen University,*
Templergraben 55, Aachen 52064, Germany

E-mail: andrzej.novak@cern.ch

A search for standard model Higgs bosons produced with transverse momentum greater than 450 GeV and decaying to charm quark-antiquark pairs is performed using proton-proton collision data collected by the CMS experiment at the LHC at 13 TeV. The search is inclusive in the Higgs boson production mode. Highly Lorentz-boosted Higgs bosons are reconstructed as single large-radius jets and are identified using a dedicated tagging technique based on a Deep Neural Network. The method is validated with Z to charm quark-antiquark pair decays and this process is observed for the first time in association with jets at a hadron collider.

41st International Conference on High Energy Physics
6-13 June 2022
Bologna, Italy

*Speaker

https://pos.sissa.it/

PoS(ICHEP2022)1131

1. Introduction

Establishing the Higgs boson (H) [1–3] coupling to the second generation of fermions is a key test of the standard model (SM). In recent years, enabled by advances in machine learning and identification methods, considerable progress was made in probing the coupling to charm quarks. In 2018 the ATLAS Collaboration has published its first constraint on the H → $c\bar{c}$ process with 2016 data [4] at 150 × the SM expectation, in the ZH production mode, where the Z boson decays into a pair of charged leptons. The CMS Collaboration has improved upon this measurement, by including the WH production mode as well as ZH production where the Z boson decays into neutrinos [5] and derived a limit of 36 × the SM expectation. Recently both collaborations published results with full LHC Run2 datasets, covering both WH and ZH production, reporting an observed limit of 31 and 7.6 × the SM expectation with ATLAS [6] and CMS [7] experiments respectively.

The CMS Collaboration now reports on a new search [8] for the H → $c\bar{c}$ process in a yet unprobed, gluon fusion (ggH) production mode in Lorentz-boosted topologies.

2. The CMS Experiment, Data and Samples

The CMS apparatus [9] is a multipurpose, nearly hermetic detector, designed to trigger on and identify electrons, muons, photons, and (charged and neutral) hadrons A global “particle-flow" (PF) algorithm [10] aims to reconstruct all individual particles in an event, combining information provided by the all-silicon inner tracker and by the crystal electromagnetic and brass-scintillator hadron calorimeters, operating inside a 3.8T superconducting solenoid, with data from the gas-ionization muon detectors embedded in the flux-return yoke outside the solenoid. The reconstructed particles are used to build τ leptons, jets, and missing transverse momentum (p_T^{miss}).

The search is performed with data collected during the LHC Run2 from 2016 to 2018, amounting to 138 fb^{-1} of integrated luminosity. The simulated samples of signal and background events are produced using various Monte Carlo (MC) event generators. The quantum chromodynamics (QCD) multijet events constitute the dominant background, but are ultimately modelled using a data-driven estimate.

3. Analysis Strategy and Selection

The ggH production mode has the highest cross-section at the LHC, however, When studying the H → $c\bar{c}$ decay, the two-jet final state is difficult to distinguish from the QCD multijet background, which has a production cross section several orders of magnitude larger than the Higgs boson. In order to suppress the background and make this search possible the analysis is constrained to a high transverse momentum (p_T > 450 GeV) regime, where the H decay products are collimated. They can then be captured within a single large radius jet, which can be identified using substructure and flavour tagging algorithms. Specifically N_2 variable [11] is used to select for two-pronged signature within the jet and deep neural network classifier called DeepDoubleX [12] is used to distinguish double-charm decays from double-bottom decays as well as from the QCD background. Events containing leptons or high p_T^{miss} are vetoed.

Events passing the outlined selection are binned in a 2D space of jet soft-drop [13] (SD) mass and jet p_T. Events failing the DDCvL selection, but passing all other cuts, constitute the control

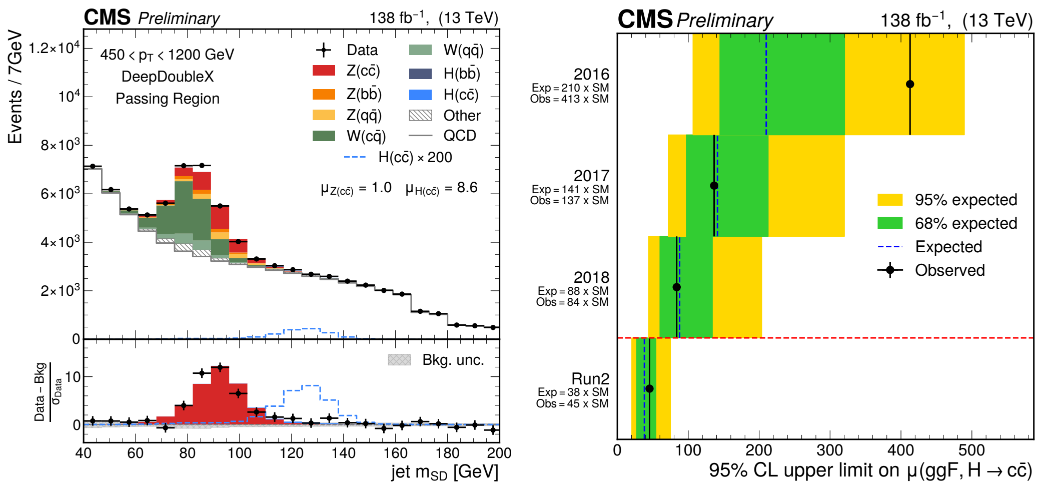

Figure 1: The H candidate jet soft-drop mass distribution in the signal (left) and control (right) regions (left) [8]. The Z → cc̄ peak is clearly visible in red, while the magnified H → cc̄ signal is visualized in dashed blue. The 95% confidence level upper limits on μ_H, split by data-taking years (right). Green and yellow bands indicate the 68 and 95% intervals on the expected limits, respectively.

region. Since the classifier is decorrelated from the two binned variables, it is assumed the QCD background distribution has a sufficiently similar shape between the signal and control regions to use the control region shape and the overall efficiency for a data-driven estimate of the distribution in the passing region. Residual differences between the shapes, originating either from imperfect DDCvL mass decorrelation or imperfect modelling, are parameterized by a fitted 2D surface modulating the nominal shape. The fitted distribution of the jet SD mass in the signal region, summed across the different p_T categories and data-taking years, is shown in Fig. 1 (left).

4. Results

The parameters of interest are the signal strengths μ_H and μ_Z, defined as the ratios of the observed to the SM expected H and Z boson production cross section times the respective branching ratio to charm quarks. The parameters are extracted from a binned maximum likelihood fit. The Z → cc̄ signal strength is measured, in order to validate the analysis method, to be $\mu_Z = 0.91^{+0.18}_{-0.15}$ (syst) ± 0.07 (theo) ± 0.05 (stat). This corresponds to an excess over the null hypothesis with a significance of well over 5 standard deviations under asymptotic assumptions.

For the μ_H measurement the Z signal strength is fixed to its expectation value $\mu_Z \equiv 1$ to constrain the DDCvL signal tagging efficiency. The best fit value for the H → cc̄ signal strength is $\mu_H = 8.6^{+19.9}_{-19.4}$. An observed (expected) upper limit is placed on the signal strength μ_H using the asymptotic $\mathrm{CL_s}$ procedure and found to be 45 (38) at the 95% CL. The observed and expected limits at 68% and 95% CLs, split by data-taking year, are shown in Fig. 1 (right).

5. Conclusion

In this report, the results of a recent search for H → cc̄ by the CMS Collaboration in a newly probed channel at the LHC are presented. The analysis has been enabled by advances in deep learning in the form of the DeepDoubleX classifier. The $Z \to c\bar{c}$ process is observed for the first time in association with jets at a hadron collider. An observed limit of $45 \times$ the SM expectation is derived for the inclusive H production process in the $H \to c\bar{c}$ decay channel. This constraint is a completely orthogonal measurement to previous $H \to c\bar{c}$ searches and will be an important component if we are to hope for evidence of $H \to c\bar{c}$ decays at the HL-LHC.

References

[1] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys. Lett. B* **716** (2012) 1, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.

[2] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett. B* **716** (2012) 30, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.

[3] CMS Collaboration, "Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV", *JHEP* **06** (2013) 081, doi:10.1007/JHEP06(2013)081, arXiv:1303.4571.

[4] ATLAS Collaboration, "Search for the decay of the Higgs boson to charm quarks with the ATLAS Experiment", *Phys. Rev. Lett.* **120** (May, 2018) doi:10.1103/physrevlett.120.211802, arXiv:1802.04329.

[5] CMS Collaboration, "A search for the standard model Higgs boson decaying to charm quarks", *JHEP* **03** (2020) 131, doi:10.1007/JHEP03(2020)131, arXiv:1912.01662.

[6] ATLAS Collaboration, "Direct constraint on the Higgs-charm coupling from a search for Higgs boson decays into charm quarks with the ATLAS detector", arXiv:2201.11428.

[7] CMS Collaboration, "Search for Higgs boson decay to a charm quark-antiquark pair in proton-proton collisions at $\sqrt{s} = 13$ TeV", arXiv:2205.05550.

[8] CMS Collaboration, "Inclusive search for a boosted Higgs boson decaying to a charm quark pairs in proton-proton collisions at $\sqrt{s} = 13$ TeV",.

[9] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[10] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", *JINST* **12** (2017) P10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.

[11] I. Moult, L. Necib, and J. Thaler, "New angles on energy correlation functions", *JHEP* **12** (2016) 153, doi:10.1007/JHEP12(2016)153, arXiv:1609.07483.

[12] CMS Collaboration, "Performance of the mass-decorrelated DeepDoubleX classifier for double-b and double-c large-radius jets with the CMS detector", CMS Detector Performance Note CMS-DP-2022-041, 2022.

[13] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, "Soft drop", *JHEP* **05** (2014) 146, doi:10.1007/JHEP05(2014)146, arXiv:1402.2657.