PROCEEDINGS OF SCIENCE

Search for the Higgs boson decay into a charm quark-antiquark pair at the CMS experiment and future developments

Angela Zaza[a,*] **on behalf of the CMS Collaboration**

[a] *University of Bari and INFN, Bari, Italy*

E-mail: angela.zaza@cern.ch

In this presentation, the search for $H \rightarrow c\bar{c}$ performed by CMS with Run-2 data is described. So far, the Higgs couplings to third generation fermions and muons have been measured with a good precision, while the assessment of second generation quark and first generation fermions still represents a challenge at the major collider experiments, like CMS and ATLAS. Searches for $H \rightarrow c\bar{c}$ have been carried out by both ATLAS and CMS using data collected during Run-2 at the LHC. The best result was obtained by CMS, by exploiting the VH production mode, with the Higgs boson generated in association with a W or Z boson, which offers a good signal to QCD background discrimination. With this analysis, an upper limit of 14 times the value predicted by the standard model was set on the product of the VH production cross section times the $H \rightarrow c\bar{c}$ branching ratio.

The Eleventh Annual Conference on Large Hadron Collider Physics (LHCP2023)
22-26 May 2023
Belgrade, Serbia

[*] Speaker

https://pos.sissa.it/

PoS(LHCP2023)278

1. Introduction

Since the discovery of the Higgs boson in 2012 [1–3], general-purpose high energy physics (HEP) collider experiments, like CMS [4] and ATLAS [5], have dedicated significant effort in measuring its properties. Establishing the Higgs boson couplings with second generation quarks, still unknown, is currently one of the high priority goals of the CMS Collaboration. This work summarizes the search for the Higgs boson decay in a charm quark-antiquark pair performed by the CMS Collaboration with data collected in proton-proton collisions in 2016-2018 (Run-2) at a center-of-mass energy of $\sqrt{s}$ = 13 TeV, corresponding to an integrated luminosity of 138 fb^{-1}. This search is extremely challenging because of the small branching ratio of the H $\rightarrow$ c$\bar{c}$ decay ($B_{SM} \sim 3\%$) predicted by the standard model (SM), the charm jet identification (c-tagging) complexity and the overwhelming QCD multijet background. In addition to the VH production mechanism, the gluon fusion production channel (ggF) was explored in a high transverse momentum (p_T) phase space. The following sections provide an overview of these analyses.

2. Search for H $\rightarrow$ c$\bar{c}$ in the VH production channel

By targeting the leptonic decays of the vector boson (W or Z) produced in association with the Higgs boson, the QCD multijet background is highly suppressed. Three mutually exclusive analysis categories are defined according to the leptonic decay channels of the W/Z boson: Z $\rightarrow \nu\nu$ (0L), W $\rightarrow l\nu$ (1L) and Z $\rightarrow ll$ (2L), where l can be an electron or a muon. Two complementary approaches are adopted in this search, involving different topologies: the "resolved-jet" and the "merged-jet" topology, used in the regions below and above 300 GeV of the transverse momentum (p_T) of the Higgs boson candidate (H$_{cand}$), which is reconstructed as two separated or as a single large-radius jet, respectively.

2.1 Resolved-jet topology

The H$_{cand}$ is reconstructed from two well-separated and individually resolved c-tagged jets, reconstructed with a distance parameter R=0.4 of the anti-kt algorithm [8]. Heavy flavor tagging is performed, in this topology, by means of the DeepJet algorithm [9]. A boosted decision tree (BDT) algorithm is trained in each category to discriminate the signal against the dominant physics background processes. The signal strength modifier μ, defined as $\frac{(\sigma B)_{obs}}{(\sigma B)_{SM}}$, where σ is the signal production cross section and B is the branching fraction, is extracted from a maximum likelihood fit to data of the BDT output score. Dedicated control regions (CRs) are defined to determine, with fits to data, the normalizations of the V+jets and t$\bar{t}$ backgrounds. Fig. 1 (left) shows the post-fit distribution of the BDT output score in the 2L(ee) category in 2017 data. Analogous figures for the other categories can be found in Ref. [7].

2.2 Merged-jet topology

The hadronization products of the two charm quarks are reconstructed as a single large-radius jet (R=1.5). A novel powerful jet identification algorithm, ParticleNet [10], is used to identify large-radius jets arising from the Higgs boson decay into charm quarks. In order to further discriminate signal from the main backgrounds (V+jets and t$\bar{t}$) a separate BDT algorithm is trained for each category, taking as input kinematic variables not correlated with the H_{cand} mass. As for the resolved-jet topology, background normalizations are obtained by fitting data in dedicated CRs. The signal component is extracted from a fit of the Higgs boson candidate mass in each category. Its combined distribution in all the three categories is shown in Fig. 1 (right).

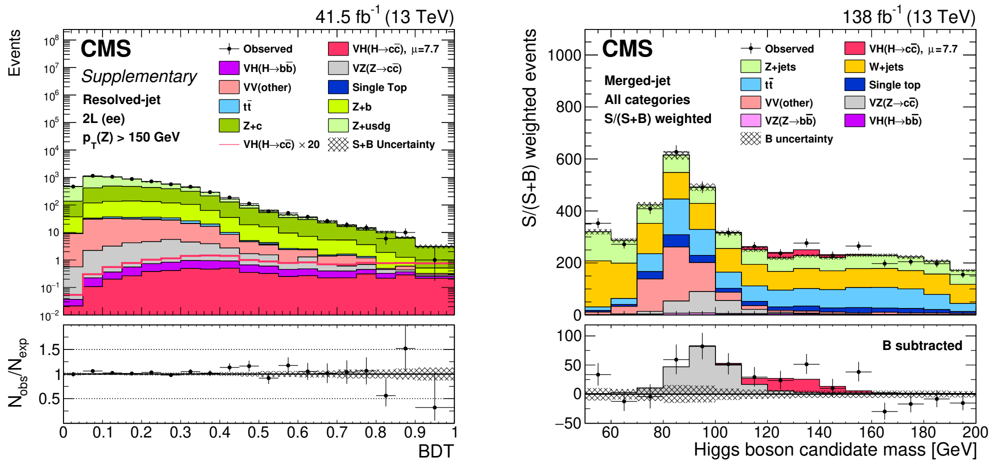

Figure 1: (Left) Post-fit distribution of the BDT discriminant in the high-p_T(V) signal region of the resolved-jet topology in the 2L(ee) channel in 2017 data. The VH(H $\to$ c$\bar{c}$) signal yield is scaled by the best fit signal strength, VH(H $\to$ c$\bar{c}$) = 7.7, in the filled red histogram, and the red line represents the expected signal contribution multiplied by a factor of 20. (Right) Combined H_{cand} mass distribution in all channels of the merged-jet analysis. The lower panel shows data (points) and the fitted VH(H $\to$ c$\bar{c}$) (red) and VZ(H $\to$ c$\bar{c}$) (grey) distributions after subtracting all other processes [7].

2.3 Combined results

In order to improve the sensitivity of the search, a simultaneous fit of the two analysis results is performed. Fig. 2 shows the 95% confidence level (CL) upper limits on the signal strength $\mu_{\mathrm{VH(H\to c\bar{c})}}$ evaluated from the independent analyses and from their combination. The observed (expected) 95% CL upper limit on $\mu_{\mathrm{VH(H\to c\bar{c})}}$ is $14(7.6^{+3.4}_{-2.3})$. Constraints on the Higgs-charm Yukawa coupling modifier k_c are then estimated, by reparametrizing $\mu_{\mathrm{VH(H\to c\bar{c})}}$ in terms of k_c:

$$\mu_{\mathrm{VH(H\to c\bar{c})}} = \frac{k_c^2}{1 + B_{\mathrm{SM}}(\mathrm{H} \to \mathrm{c\bar{c}})(1 - k_c^2)} \tag{1}$$

The observed 95% CL interval is $1.1 < |k_c| < 5.5$, and the corresponding expected constraint is $k_c < 3.4$. The analysis is validated by searching for the Z decay to a charm quark-antiquark pair (Z $\to$ c$\bar{c}$) in VZ production. For the first time at a hadron collider, this process was observed with a significance of 5.7 standard deviations.

3. Search for H $\to$ c$\bar{c}$ in the ggF production channel

A search for the Higgs boson produced with $p_T > 450$ GeV and decaying to a charm quark-antiquark pair was performed by the CMS Collaboration, targeting the ggF production mode [11].

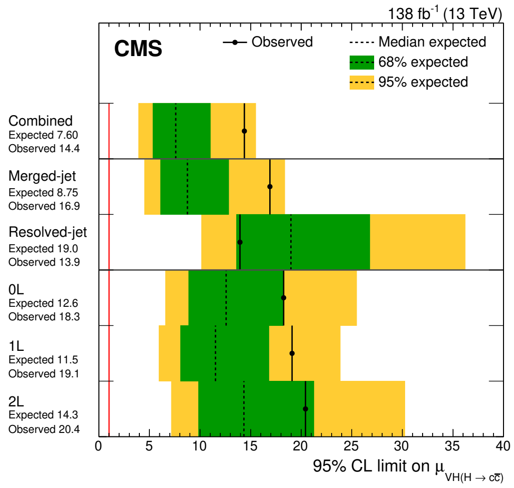

Figure 2: The 95% CL upper limits on $\mu_{\mathrm{VH(H\to c\bar{c})}}$. Green and yellow bands indicate the 68 and 95% intervals on the expected limits, respectively. The vertical red line indicates the SM value $\mu_{\mathrm{VH(H\to c\bar{c})}} = 1$ [7].

The $\mathrm{H_{cand}}$ is reconstructed as a single large-radius jet (R=0.8), tagged with the DeepDoubleX algorithm [12]. The soft-drop (SD) algorithm [13] is applied to the jet mass (m_{SD}) to remove soft and wide-angle radiation. The dominant QCD multijet background is estimated from data by using CRs. The V+jets processes are estimated from simulation, while the $\mathrm{t\bar{t}}$ background is given by simulation and normalized with data in a dedicated CR. The signal strength μ_{H} is extracted from a binned (m_{SD}, p_{T}) maximum likelihood fit to data. The observed (expected) 95% CL upper limit on $\mu_{\mathrm{H\to c\bar{c}}}$ is 47 (39). The search is validated by measuring the $\mathrm{Z} \to \mathrm{c\bar{c}}$ process, observed in association with high-p_{T} jets for the first time in this analysis, with signal strength $\mu = 1.00^{+0.17}_{-0.14}\,(\mathrm{syst}) \pm 0.08\,(\mathrm{theo}) \pm 0.06\,(\mathrm{stat})$.

4. Conclusions and future developments

In this work, the searches for the Higgs boson decay in a charm quark-antiquark pair performed by the CMS Collaboration are summarized. The VH and ggF Higgs production channels have been explored, the former providing the tightest constraint on the Higgs-charm Yukawa coupling modifier to date: $1.1 < k_{\mathrm{c}} < 5.5$ (expected: $|k_{\mathrm{c}}| < 3.4$). During Run-3 of the LHC, new trigger and analysis methods are expected to further improve the constraints on the coupling between the Higgs boson and charm quarks.

References

[1] ATLAS Collaboration, Phys. Lett. ***B*716** (2012) 1, doi:10.1016/j.physletb.2012.08.020

[2] CMS Collaboration, Phys. Lett. ***B*716** (2012) 30, doi:10.1016/j.physletb.2012.08.021

[3] CMS COllaboration, JHEP **06** (2013) 081, doi:10.1007/JHEP06(2013)081

[4] CMS Collaboration, JINST **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004

[5] ATLAS Collaboration, JINST **3** (2008) S08003, doi:10.1088/1748-0221/3/08/S08003

[6] R. L. Workman et al., Review of Particle Physics, PTEP **2022** (2022) 083C01, doi:10.1093/ptep/ptac097

[7] CMS Collaboration, Search for Higgs boson decay to a charm quark-antiquark pair in proton-proton collisions at $\sqrt{s}$ = 13 TeV, Phys. Rev. Lett. **131** (2023) 061801, doi: 10.1103/PhysRevLett.131.061801. Additional figures at: http://cern.ch/go/Hdd9.

[8] M. Cacciari, G. P. Salam and G. Soyez, The anti-kt jet clustering algorithm, JHEP **04** (2008) 063

[9] E. Bols et al., Jet flavour classification using DeepJet, JINST **15** (2020) P12012

[10] H. Qu and L. Gouskos, Jet tagging via particle clouds, Phys. Rev. D **101** (2020) 056019

[11] CMS Collaboration, Search for Higgs Boson and Observation of Z Boson through Their Decay into a Charm Quark-Antiquark Pair in Boosted Topologies in Proton-Proton Collisions at $\sqrt{s}$ = 13 TeV, Phys. Rev. Lett. **131** (2023) 041801

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

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

PoS(LHCP2023)278