Appendix
Earlier Run II HH → $b\bar{b}\tau^+\tau^-$ Searches

This appendix gives an overview of the Run II searches for HH → bb$\tau\tau$ production that were derived in the context of this thesis using the 2015 CMS dataset of 2.7 fb^{-1} and on the first part of the 2016 dataset of 12.9 fb^{-1}. In the following, the two searches are referred to as the "Moriond 2016 search" and the "ICHEP 2016 search" accordingly to the conference in which the corresponding results were presented for the first time. The discussion focuses on the event selection and categorization, and on the results obtained. The selections related to the object quality (compatibility with the primary vertex and identification criteria) are the same discussed in Chap. 5 and thus not repeated here. The MC simulation, the background modelling methods, and the evaluation of the systematic uncertainties also closely follow what has been described in Chap. 6 and are consequently not detailed further.

A.1 Moriond 2016 Search

The Moriond 2016 search constituted the first exploration of HH production at 13 TeV, and was based on an integrated luminosity of 2.7 fb^{-1} collected with the CMS experiment in 2015. The corresponding results are documented in Refs. [1, 2] for the nonresonant and resonant HH production mode, respectively, with additional information provided in the CMS internal analysis notes [3, 4].

A.1.1 Event Selection

The three final states containing at least one semileptonic τ decay ($\tau_\mu\tau_h$, $\tau_e\tau_h$, $\tau_h\tau_h$) are explored. The dataset analysed corresponds to 2.7 fb^{-1} of data collected in 2015. Because of issues experienced with the HF calorimeter, affecting about 0.4 fb^{-1}, inputs from this subdetector are not used in the computation of the $\vec{p}_T^{\,\text{miss}}$ vector. It

L. Cadamuro, *Search for Higgs Boson Pair Production in the $b\bar{b}\tau^+\tau^-$ Decay Channel*, Springer Theses, https://doi.org/10.1007/978-3-030-04055-0

was verified that its magnitude and directions are properly estimated for the selected $bb\tau\tau$ events and well modelled in the MC simulation.

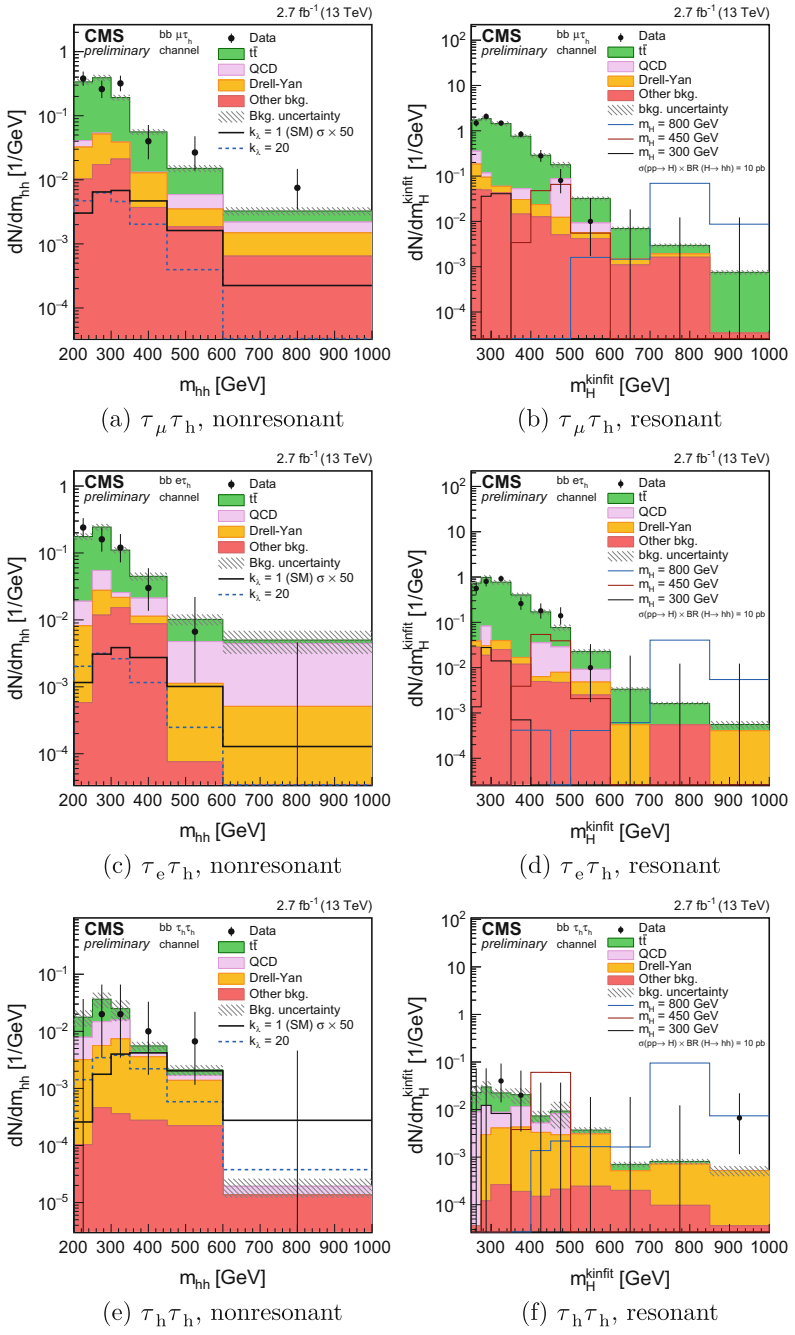
Most of the data in the $\tau_\mu\tau_h$ and $\tau_e\tau_h$ final states are collected with a trigger requiring the presence of, respectively, a single muon or electron in the event. For a first part of the data, corresponding to 76 pb^{-1}, a cross-trigger requiring the additional presence of a τ_h is used. The $\tau_\mu\tau_h$ events thus recorded must contain a muon candidate of $p_T > 19$ GeV and $|\eta| < 2.1$ and a τ_h candidate of $p_T > 20$ GeV and $|\eta| < 2.3$. The selected muon must satisfy the relative isolation requirement $\mathcal{I}^{rel} < 0.1$ while the selected τ_h must have an absolute charged hadron isolation $\sum p_T^{ch} < 3$ GeV; both isolation criteria are described in Chap. 5. Similarly, $\tau_e\tau_h$ events must contain an electron candidate of $p_T > 24$ GeV and $|\eta| < 2.1$, which satisfies $\mathcal{I}^{rel} < 0.1$. The presence of a τ_h candidate that satisfies the same selections as in the $\tau_\mu\tau_h$ final state is also required. Finally, events selected in the $\tau_h\tau_h$ final state are selected with a double-τ_h trigger, and must contain two τ_h candidates reconstructed offline with $p_T > 45$ GeV and $|\eta| < 2.1$, that satisfy $\sum p_T^{ch} < 2$ GeV. In all the three final states, the two leptons that form the $H \to \tau\tau$ candidate are required to have opposite electric charge, and events containing additional isolated leptons are rejected.

The events selected with the previous criteria must also contain two AK4 jets with $p_T > 30$ GeV and $|\eta| < 2.4$. For the nonresonant search, both jets in the $\tau_\mu\tau_h$ and $\tau_e\tau_h$ final state must satisfy the loose b tagging WP corresponding to a misidentification efficiency of gluon and light flavour quark jets of 10%. A multivariate discriminant is applied on these events. The variables used as inputs to the method are a subset of those used for the HM BDT described in Sect. 5.6.2 of Chap. 5. In particular, the $\Delta\varphi(l, p_T^{miss})$, $m_T(\ell)$, and $m_T(\tau_h)$ variables are not used and the method is trained on a combination of nonresonant events corresponding to $k_\lambda = 1$ (SM) and $k_\lambda = 2.45$ (maximal interference of the two diagrams contributing to gluon fusion HH production). For the resonant search, no multivariate method is applied in the $\tau_\mu\tau_h$ and $\tau_e\tau_h$ final states and the two jets must satisfy the medium b tagging WP, corresponding to a misidentification rate of 1%. In the $\tau_h\tau_h$ final state, the two jets must satisfy the loose b tagging WP in both the resonant and nonresonant searches.

An invariant mass requirement on the bb and $\tau\tau$ invariant masses (the latter reconstructed with the SVfit algorithm) is applied on the selected events as $80 < m_{bb}(m_{\tau\tau}) < 160$ GeV.

A.1.2 Results

The m_{HH}^{KinFit} variable described in Sect. 7.2 of Chap. 7 is used for the resonant search, while the four-body mass of the visible $\tau\tau$ decay products and of the selected bb candidates, m_{HH}, is used for the nonresonant search. The corresponding event distributions are shown in Fig. A.1 and the observed and expected number of events is summarized in Table A.1. The 95% CL upper limits on nonresonant production as a function of k_λ are summarized in Fig. A.2a, and the 95% CL upper limits on resonant production as a function of m_X for the spin-0 resonance hypothesis are shown



(a) $\tau_\mu\tau_h$, nonresonant (b) $\tau_\mu\tau_h$, resonant

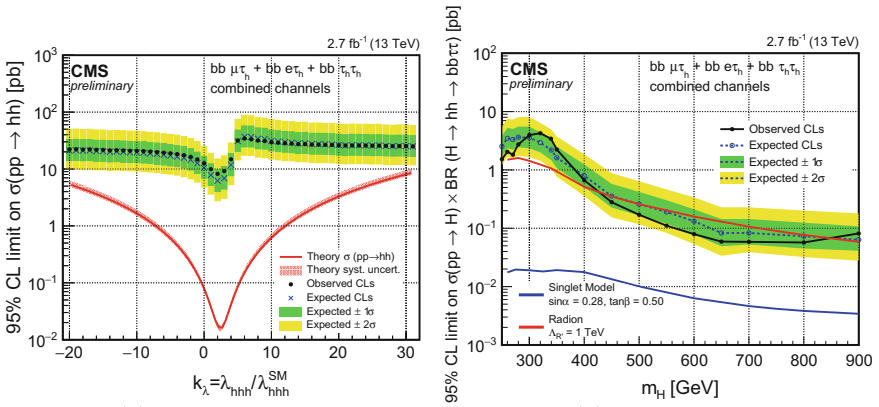
(c) $\tau_e\tau_h$, nonresonant (d) $\tau_e\tau_h$, resonant

(e) $\tau_h\tau_h$, nonresonant (f) $\tau_h\tau_h$, resonant

Fig. A.1 Event distribution of the m_{HH} (left column) and m_{HH}^{KinFit} (right column) variables for the $\tau_\mu\tau_h$ (top row), $\tau_e\tau_h$ (central row), and $\tau_h\tau_h$ (bottom row) final states. The two variables are used for the nonresonant and resonant searches, respectively. The events are selected from the 2.7 fb^{-1} dataset collected in 2015

Table A.1 Observed and expected event yields for the Moriond 2016 HH → bbττ search. Values for the $\tau_\mu\tau_h$ and $\tau_e\tau_h$ final states are separately reported for the resonant and nonresonant searches, that differ for the selections applied, while the same signal region is defined for the $\tau_h\tau_h$ final state. The background values and the errors correspond to the nuisance parameters obtained from a maximum likelihood fit to the observed data under the background-only hypothesis. The expected yields of a few resonant and nonresonant processes are quoted. The former are normalized to $\sigma(pp \to X) \times \mathcal{B}(X \to HH \to bb\tau\tau) = 1$ pb while the latter are normalized to the theoretical cross section

Process	$\tau_\mu\tau_h$		$\tau_e\tau_h$		$\tau_h\tau_h$
	nonres.	res.	nonres.	res.	
$t\bar{t}$	45.5 ± 5.7	203.3 ± 25.4	25.0 ± 3.2	106.8 ± 13.6	2.2 ± 0.3
QCD	1.2 ± 1.0	10.0 ± 3.3	5.2 ± 2.3	4.2 ± 1.8	1.4 ± 1.0
Z+jets	5.2 ± 1.7	7.2 ± 1.8	2.1 ± 0.7	2.1 ± 0.5	1.2 ± 0.3
W+jets	0.9 ± 0.2	1.4 ± 0.1	1.1 ± 0.2	0.5 ± 0.1	–
Single top	2.0 ± 0.2	5.2 ± 0.5	1.1 ± 0.1	3.2 ± 0.3	–
Di-boson	0.4 ± 0.1	0.3 ± 0.1	0.08 ± 0.02	0.11 ± 0.02	0.11 ± 0.02
Tot. exp. bkg.	55.2 ± 6.0	227.4 ± 25.7	34.6 ± 4.0	116.9 ± 13.7	4.9 ± 1.1
$m_X = 300$ GeV		4.0		2.0	0.8
$m_X = 600$ GeV		14.1		8.2	10.9
$m_X = 900$ GeV		18.4		10.5	13.3
$k_\lambda = 1$ ($\times 10^{-2}$)	3.2		1.7		2.1
$k_\lambda = 15$ ($\times 10^{-1}$)	7.8		4.1		5.3
DATA	59	224	30	110	4



(a) Nonresonant search (b) Resonant search

Fig. A.2 95% CL upper limits derived in the Moriond 2016 search for nonresonant production as a function of k_λ (**a**) and as for resonant production as a function of (**b**)

in Fig. A.2b. The observed and expected 95% CL upper limits on $\sigma(\mathrm{gg} \to \mathrm{HH})$ for the SM signal amount to 8.8 and 7.2 pb, respectively. These values correspond to approximately 260 and 215 times the SM prediction.

A.2 ICHEP 2016 Search

The ICHEP 2016 search was based on the first part of the dataset collected in 2016, corresponding to an integrated luminosity of 12.9 fb^{-1}. The results for the nonresonant and resonant search are documented in Refs. [5, 6], respectively, and in the supporting CMS internal analysis notes [7, 8].

A.2.1 Event Selection

As in the Moriond 2016 analysis, events in the $\tau_\mu\tau_\mathrm{h}$ and $\tau_\mathrm{e}\tau_\mathrm{h}$ final states are recorded with a single muon and single lepton trigger, respectively, while events in the $\tau_\mathrm{h}\tau_\mathrm{h}$ final state are recorded with a double-τ_h trigger. Because of the higher instantaneous luminosity experienced in 2016, higher p_T thresholds are applied on muon and electrons and, consequently, they are also increased in the offline selection. Muons in the $\tau_\mu\tau_\mathrm{h}$ final state must have $p_\mathrm{T} > 23$ GeV, and electrons in the $\tau_\mathrm{e}\tau_\mathrm{h}$ final state must have $p_\mathrm{T} > 27$ GeV. In both cases, the leptons must satisfy the additional requirements $|\eta| < 2.1$ and $\mathcal{I}^\mathrm{rel} < 0.1$. In contrast, the same trigger thresholds as in 2015 were maintained for the double-τ_h trigger thanks to the deployment of the upgraded L1 trigger with the τ algorithm described in Chap. 4. The other selections applied on

the τ_h candidates selected in three final states are the same as those of the Moriond 2016 search discussed above.

The nonresonant ICHEP 2016 search follows the same event selection strategy as the Moriond 2016 one, applying the methods previously developed to the larger integrated luminosity analysed. The resonant search is instead improved with the introduction of the bb event categorization. As discussed in Sect. 5.5.2 of Chap. 5, two resolved categories (1b1j and 2b) and one boosted category are defined. The medium WP of the b tagging discriminant is used in the $\tau_\mu\tau_h$ and $\tau_e\tau_h$ final states while the loose one is used in the $\tau_h\tau_h$ final state because of the smaller number of events expected. All events are required to satisfy the invariant mass selection $80 < m_{\tau\tau} < 160\,\text{GeV}$. Events in the two resolved categories are required to satisfy in addition $80 < m_{bb} < 160\,\text{GeV}$ while events in the boosted category must have $90 < m_{AK8} < 160\,\text{GeV}$, where the symbol m_{AK8} denotes the AK8 jet invariant mass estimated with the soft drop jet grooming algorithm.

Results

The distributions of the m_{HH} variable, used to search for the presence of a nonresonant signal, is shown in Fig. A.3, and the corresponding number of events is summarized in Table A.2. The distributions of the m_{HH}^{KinFit} variable used in the resonant search are shown in Fig. A.4 for the different categories of the search, and the expected and observed event yields are shown in Table A.3.

Model independent limits on nonresonant production are set both as a function of k_λ/k_t and for the twelve shape benchmarks, as shown in Fig. A.5a and b, respectively. The results are also used to set 95% CL exclusion limits on anomalous k_λ and k_t couplings as illustrated in Fig. A.5c. The parametric signal modelling detailed in Sect. 6.2 of Chap. 6 was not yet finalized when these results were released. Consequently, the event weighting procedure was applied for a smaller number of points and based on the bidimensional m_{HH} and $|\cos\theta^*|$ distribution obtained from generated events for the (k_λ, k_t) combinations denoted with a circle in the figure. For the SM signal, the observed and expected 95% CL upper limits on $\sigma(gg \to HH) \times \mathcal{B}(HH \to bb\tau\tau)$ amount to 508 and 420 fb, respectively. These values correspond to approximately 200 and 170 times the SM prediction.

The 95% CL upper limits on the resonant production cross section as a function of the mass hypothesis m_X are shown in Fig. A.6.

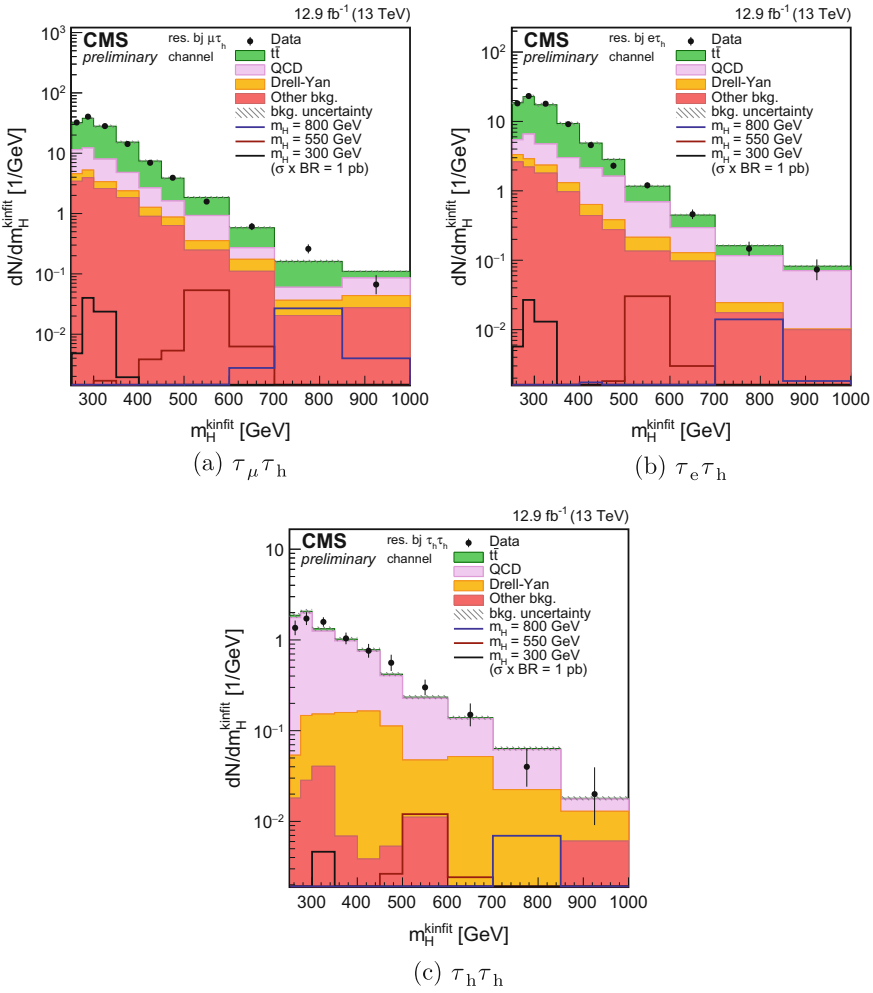
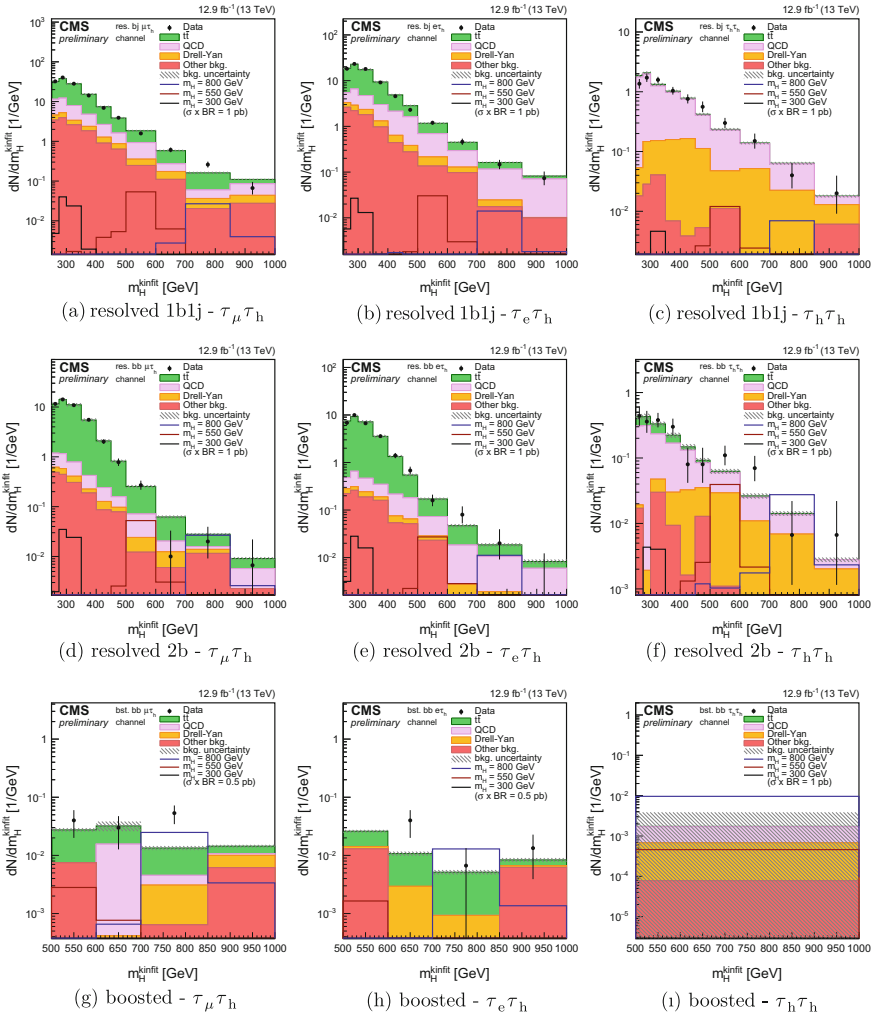


Fig. A.3 m_{HH} distribution for events selected in the nonresonant ICHEP 2016 search in the $\tau_\mu\tau_\mathrm{h}$ (**a**), $\tau_\mathrm{e}\tau_\mathrm{h}$ (**b**), and $\tau_\mathrm{h}\tau_\mathrm{h}$ (**c**) final states

Table A.2 Observed and expected event yields for the ICHEP 2016 nonresonant HH → $b\bar{b}\tau\tau$ search. The background values and the errors correspond to the nuisance parameters obtained from a maximum likelihood fit to the observed data under the background-only hypothesis

Process	$\tau_\mu\tau_h$	$\tau_e\tau_h$	$\tau_h\tau_h$
$t\bar{t}$	368.1 ± 37.2	228.5 ± 23.4	15.3 ± 1.7
Multijet	52.2 ± 6.5	55.7 ± 4.6	45.7 ± 4.1
Z+jets	31.5 ± 3.0	18.7 ± 1.9	10.3 ± 1.1
W+jets	13.0 ± 1.0	11.0 ± 0.9	1.4 ± 0.1
Single top	11.6 ± 1.0	10.7 ± 1.0	1.5 ± 0.2
Di-boson	3.1 ± 0.4	1.4 ± 0.2	0.7 ± 0.1
Total expected background	480.0 ± 37.9	326.0 ± 24.4	74.8 ± 4.6
$k_\lambda = 1$	0.24	0.13	0.12
$k_\lambda = 20$	7.8	4.8	4.1
DATA	464	317	84



(a) resolved 1b1j - $\tau_\mu\tau_h$ (b) resolved 1b1j - $\tau_e\tau_h$ (c) resolved 1b1j - $\tau_h\tau_h$

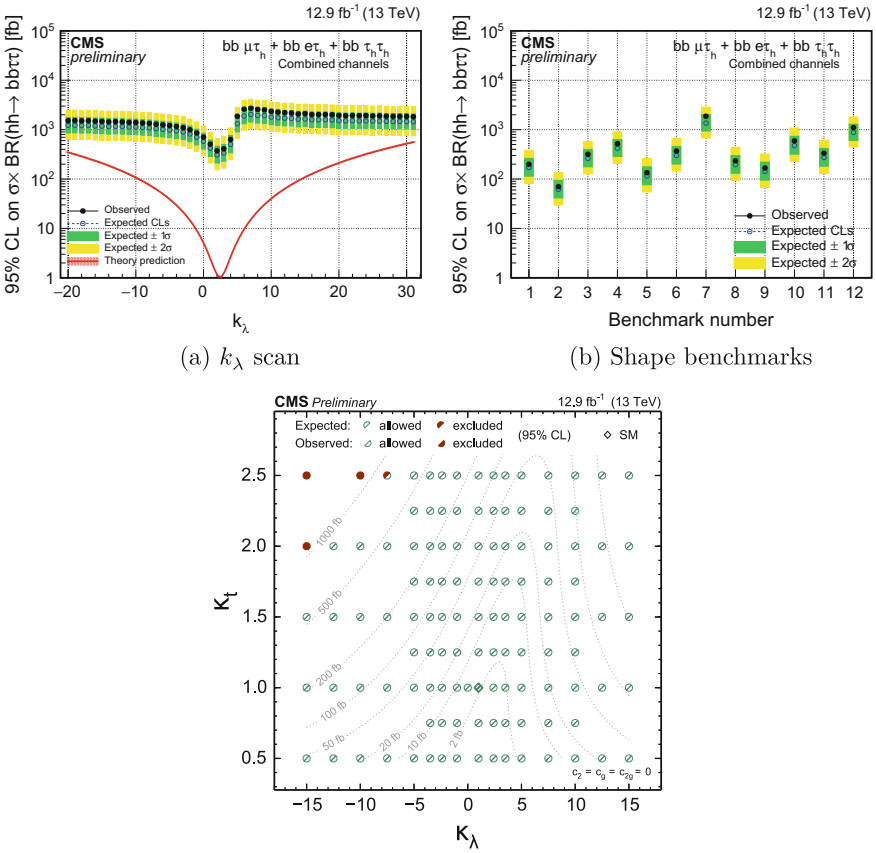
(d) resolved 2b - $\tau_\mu\tau_h$ (e) resolved 2b - $\tau_e\tau_h$ (f) resolved 2b - $\tau_h\tau_h$

(g) boosted - $\tau_\mu\tau_h$ (h) boosted - $\tau_e\tau_h$ (i) boosted - $\tau_h\tau_h$

Fig. A.4 m_{HH}^{KinFit} distribution of events selected in the ICHEP 2016 resonant search. Events are separately shown in the 1b1j (top row), 2b (central row), and boosted category (bottom row), for the $\tau_\mu\tau_h$ final state (left column), $\tau_e\tau_h$ (central column), and $\tau_h\tau_h$ (right column)

Table A.3 Observed and expected event yields for the ICHEP 2016 resonant HH → $bb\tau\tau$ search. The background values and the errors correspond to the nuisance parameters obtained from a maximum likelihood fit to the observed data under the background-only hypothesis. Signal yields are normalized to $\sigma(pp \to X) \to \mathcal{B}(X \to HH) = 1\,pb$

Process	$\tau_\mu\tau_h$	$\tau_e\tau_h$	$\tau_h\tau_h$
Resolved 1b1j category			
$t\bar{t}$	3184.1 ± 358.7	1957.4 ± 226.5	15.2 ± 1.9
Multijet	880.9 ± 30.6	573.3 ± 24.8	261.8 ± 16.5
Z+jets	180.6 ± 17.6	107.7 ± 10.9	44.3 ± 6.0
W+jets	283.5 ± 26.7	162.4 ± 15.9	2.9 ± 0.3
Single top	220.7 ± 22.2	147.3 ± 15.3	1.3 ± 0.2
Di-boson	20.6 ± 2.7	12.6 ± 1.7	2.1 ± 0.3
Total expected background	4770.5 ± 362.1	2960.7 ± 229.2	327.6 ± 17.7
$m_X = 300\,GeV$	34.2	21.5	4.4
$m_X = 600\,GeV$	90.9	47.7	25.8
$m_X = 900\,GeV$	63.4	29.6	21.9
DATA	4755	2938	333
Resolved 2b category			
$t\bar{t}$	1501.6 ± 169.1	975.7 ± 112.9	15.4 ± 1.9
Multijet	72.4 ± 8.7	53.7 ± 7.4	44.2 ± 6.2
Z+jets	18.7 ± 2.4	9.6 ± 1.2	9.8 ± 1.6
W+jets	12.8 ± 1.2	11.4 ± 1.1	1.4 ± 0.1
Single top	45.1 ± 4.5	24.5 ± 2.6	1.5 ± 0.2
Di-boson	2.2 ± 0.3	1.2 ± 0.2	0.53 ± 0.08
Total expected background	1652.8 ± 169.5	1076.2 ± 113.2	72.8 ± 6.7
$m_X = 300\,GeV$	30.1	22.1	4.5
$m_X = 600\,GeV$	84.8	48.0	68.5
$m_X = 900\,GeV$	56.7	24.1	58.3
DATA	1638	1065	82
Boosted category			
$t\bar{t}$	10.7 ± 1.2	6.2 ± 0.6	0.045 ± 0.006
Multijet	2.1 ± 0.8	–	0.52 ± 1.06
Z+jets	1.3 ± 0.1	1.3 ± 0.1	0.30 ± 0.04
W+jets	0.9 ± 0.1	1.0 ± 0.1	–
Single top	0.44 ± 0.04	1.1 ± 0.1	–
Di-boson	0.6 ± 0.1	0.1 ± 0.01	0.08 ± 0.01
Total expected background	16.1 ± 1.5	9.8 ± 0.7	0.94 ± 1.06
$m_X = 300\,GeV$	–	–	–
$m_X = 600\,GeV$	19.5	11.1	9.0
$m_X = 900\,GeV$	148.8	72.9	102.2
DATA	21	11	0



(a) k_λ scan (b) Shape benchmarks

(c) (k_λ, k_t) exclusion

Fig. A.5 95% CL upper limits as a function of the anomalous trilinear coupling k_λ (**a**) and for the twelve shape benchmarks (**b**) obtained from the ICHEP 2016 nonresonant analysis. (**c**) 95% CL excluded regions of the (k_λ, k_t) plane assuming $c_2 = c_g = c_{2g} = 0$. Open green semicircles denote points compatible with the current data while red full semicircles denote points excluded with the current data, with the two halves of the circles denoting the expected and observed exclusion as reported in the legend. The diamond shaped marker indicates to the prediction of the SM. The dotted lines indicate trajectories in the plane with equal HH production cross section, and are labelled with the corresponding value of $\sigma(gg \to HH) \times \mathcal{B}(HH \to bb\tau\tau)$

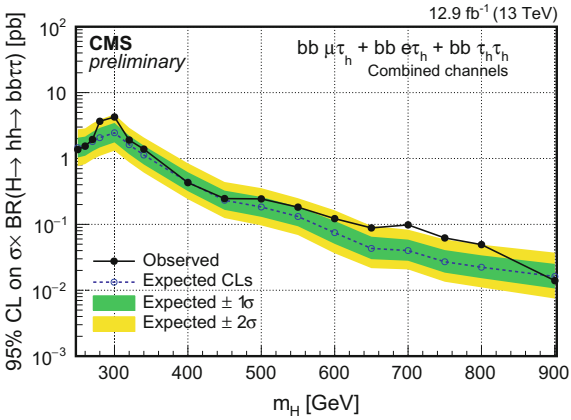


Fig. A.6 95% CL upper limits on $\sigma(gg \to X \to HH) \times \mathcal{B}(HH \to bb\tau\tau)$ as a function of the resonance mass m_X. These results correspond to the resonant ICHEP 2016 search

References

1. CMS Collaboration, Search for non-resonant Higgs boson pair production in the bb̄τ⁺τ⁻ final state, CMS Physics Analysis Summary CMS-PAS-HIG-16-012, CERN (2016), https://cds.cern.ch/record/2139324
2. CMS Collaboration, Search for resonant Higgs boson pair production in the bb̄τ⁺τ⁻ final state, CMS Physics Analysis Summary CMS-PAS-HIG-16-013, CERN (2016), https://cds.cern.ch/record/2139315
3. CMS Collaboration, Search for non-resonant Higgs boson pair production in the bbττ final state, CMS Analysis Note AN-2015/315, CERN (2015), http://cms.cern.ch/iCMS/jsp/db_notes/noteInfo.jsp?cmsnoteid=CMS%20AN-2015/315. Access restricted to CMS members
4. CMS Collaboration, Search for resonant Higgs boson pair production in the bbττ final state, CMS Analysis Note AN-2016/006, CERN (2016), http://cms.cern.ch/iCMS/jsp/db_notes/noteInfo.jsp?cmsnoteid=CMS%20AN-2016/006. Access restricted to CMS members
5. CMS Collaboration, Search for non-resonant Higgs boson pair production in the bbττ final state using 2016 data, CMS Physics Analysis Summary CMS-PAS-HIG-16-028, CERN (2016), https://cds.cern.ch/record/2204934
6. CMS Collaboration, Search for resonant Higgs boson pair production in the bb̄τ⁺τ⁻ final state using 2016 data, CMS Physics Analysis Summary CMS-PAS-HIG-16-029, CERN (2016), https://cds.cern.ch/record/2204936
7. CMS Collaboration, Search for non-resonant Higgs boson pair production in the bbττ final state with 2016 data, CMS Analysis Note AN-2016/232, CERN (2016), http://cms.cern.ch/iCMS/jsp/db_notes/noteInfo.jsp?cmsnoteid=CMS%20AN-2016/232. Access restricted to CMS members
8. CMS Collaboration, Search for resonant Higgs boson pair production in the bbττ final states with 2016 datasets, CMS Analysis Note AN-2016/186, CERN (2016), http://cms.cern.ch/iCMS/jsp/db_notes/noteInfo.jsp?cmsnoteid=CMS%20AN-2016/186. Access restricted to CMS members

About the Author

After completing his studies in particle physics at the University of Milano-Bicocca, Luca Cadamuro obtained his PhD at the University of Paris-Saclay – École polytechnique, developing his research work at the Laboratoire Leprince-Ringuet. His thesis on Higgs boson pair production and trigger algorithms has been recognised by the CMS Collaboration with the 2017 CMS PhD Thesis Award. He is currently continuing the exploration of Higgs boson pair production with the CMS experiment as a Postdoctoral Research Associate at the University of Florida and Distinguished Researcher at the Fermilab LPC.

L. Cadamuro, *Search for Higgs Boson Pair Production in the $b\bar{b}\tau^+\tau^-$ Decay Channel*, Springer Theses, https://doi.org/10.1007/978-3-030-04055-0