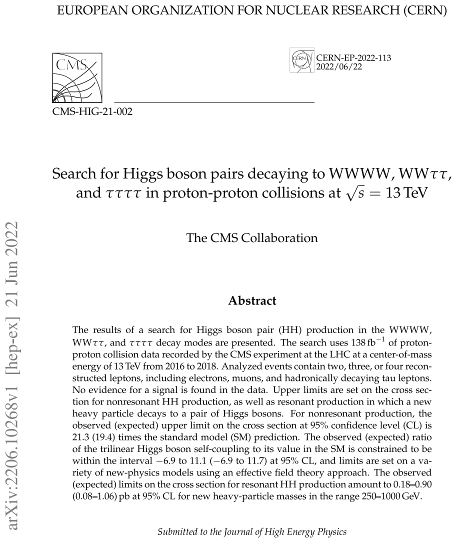
Search for Higgs boson pair production in multilepton final states

Karl Ehatäht (NICPB, EE)
for the CMS collaboration

2nd CERN Baltic Conference,
10–12 October 2022

Table of contents

Based on arXiv:2206.10268
(submitted to JHEP)



EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH (CERN)

CERN-EP-2022-113
2022/06/22

CMS-HIG-21-002

Search for Higgs boson pairs decaying to WWWW, WW$\tau\tau$, and $\tau\tau\tau\tau$ in proton-proton collisions at $\sqrt{s} = 13$ TeV

The CMS Collaboration

Abstract

The results of a search for Higgs boson pair (HH) production in the WWWW, WW$\tau\tau$, and $\tau\tau\tau\tau$ decay modes are presented. The search uses 138 fb^{-1} of proton-proton collision data recorded by the CMS experiment at the LHC at a center-of-mass energy of 13 TeV from 2016 to 2018. Analyzed events contain two, three, or four reconstructed leptons, including electrons, muons, and hadronically decaying tau leptons. No evidence for a signal is found in the data. Upper limits are set on the cross section for nonresonant HH production, as well as resonant production in which a new heavy particle decays to a pair of Higgs bosons. For nonresonant production, the observed (expected) upper limit on the cross section at 95% confidence level (CL) is 21.3 (19.4) times the standard model (SM) prediction. The observed (expected) ratio of the trilinear Higgs boson self-coupling to its value in the SM is constrained to be within the interval −6.9 to 11.1 (−6.9 to 11.7) at 95% CL, and limits are set on a variety of new-physics models using an effective field theory approach. The observed (expected) limits on the cross section for resonant HH production amount to 0.18–0.90 (0.08–1.06) pb at 95% CL for new heavy-particle masses in the range 250–1000 GeV.

Submitted to the Journal of High Energy Physics

Higgs boson pair (HH) production

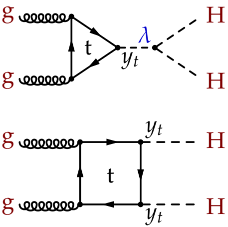
- ▷ Correspondence between Higgs self-coupling λ and $\lambda_{\mathrm{SM}} = m_H^2/(2v^2)$ to be tested

⇒ Direct probe of the Higgs potential $V(\phi) \overset{\mathrm{EWSB}}{\rightarrow} V(h) = \underbrace{\lambda v^2 h^2}_{\text{mass term}} + \underbrace{\lambda v h^3} + \underbrace{\lambda h^4/4}$

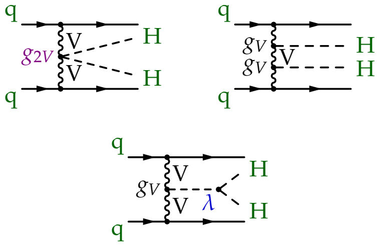
⇒ Further insight into the stability of SM vacuum [1]

- ▷ Not enough sensitivity to indirectly infer from single Higgs measurements [2, 3] + possible BSM in loops
- ▷ HH production accessible, but HHH not (even with the HL-LHC [4])

⇒ Focus on gluon-gluon fusion (ggHH), vector boson fusion (qqHH) production



Kinematics fully described by HH mass (m_{HH}) and angle between them ($\cos\theta^*$) at leading order (LO)



Sensitive to g_{2V}!

HH decay modes

- Challenge: HH production has $\mathcal{O}(10^3)$ smaller cross section compared to single Higgs production
- Assuming SM, only ~ 4300 HH events expected from LHC Run 2 data (billions of events)
- No single „golden" channel $\Rightarrow$ need to combine all viable HH decay modes ☞
- **Our goal**: multilepton channels (leptonic final states of WW^*WW^*, $WW^*\tau\tau$ and $\tau\tau\tau\tau$)

$H\to\mathcal{YY}$ \ $H\to\mathcal{XX}$	$b\bar{b}$	WW^*	gg	$\tau^+\tau^-$	cc	ZZ^*	$\gamma\gamma$	$Z\gamma$	$\mu^+\mu^-$
$b\bar{b}$	✓ 0.3392	✓ 0.2489	0.0954	✓ 0.0731	0.0337	✓ 0.0305	✓ 2.64×10^{-3}	1.79×10^{-3}	2.53×10^{-4}
WW^*		✓ 0.0457	0.0350	✓ 0.0268	0.0124	✓ 0.0112	✓ 9.70×10^{-4}	6.55×10^{-4}	9.30×10^{-5}
gg			6.70×10^{-3}	0.0103	4.73×10^{-3}	4.29×10^{-3}	3.72×10^{-4}	2.51×10^{-4}	3.56×10^{-5}
$\tau^+\tau^-$				✓ 3.93×10^{-3}	3.63×10^{-3}	✓ 3.29×10^{-3}	2.85×10^{-4}	1.92×10^{-4}	2.73×10^{-5}
cc					8.36×10^{-4}	1.51×10^{-3}	1.31×10^{-4}	8.86×10^{-5}	1.26×10^{-5}
ZZ^*						✓ 6.86×10^{-4}	1.19×10^{-4}	8.03×10^{-5}	1.14×10^{-5}
$\gamma\gamma$							5.15×10^{-6}	6.96×10^{-6}	9.88×10^{-7}
$Z\gamma$								2.35×10^{-6}	6.67×10^{-7}
$\mu^+\mu^-$									4.74×10^{-8}

✓ Published LHC Run 2 results

☐ Multilepton: $\sum_i \mathcal{B}_i = 7.64\%$ (9.16%)

$\mathcal{B}_i(HH\to\mathcal{XXYY})$, $m_H = 125$ GeV

Strategy

1) Devise object and event selection criteria targeting $HH \rightarrow WW^*WW^*, WW^*\tau\tau, \tau\tau\tau\tau$ signal with prompt (= from W, Z or τ decays) electrons and muons (ℓ) and hadronic τ decay products (τ_h) in final state
2) Carefully model the signal hypotheses, background processes, uncertainties
3) Develop discriminants for bringing out signal from background
4) Extract upper limits for SM production cross section, constraints on $\kappa_\lambda = \lambda/\lambda^{SM}$ (and on $\kappa_{2V} = g_{2V}/g_{2V}^{SM}$ if possible) using maximum likelihood (ML) fit

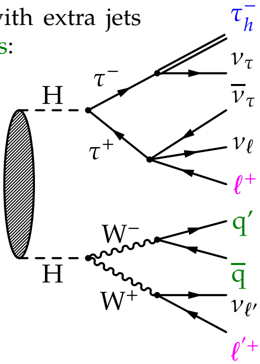
Data: pp collisions at $\sqrt{s} = 13$ TeV recorded by the CMS detector during LHC Run 2 (2016-2018) $\Rightarrow 138\,\text{fb}^{-1}$

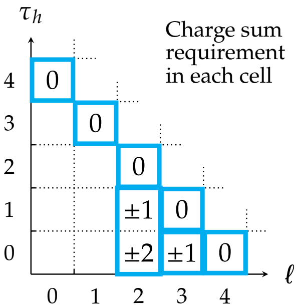
Object and event selection

7 mutually exclusive event categories based on the number of ℓ/τ_h:

- Backgrounds reduced by rejecting events with Z or light meson resonances, with b tagged jets, $H \to ZZ^* \to 4\ell$ events, or Z/γ^* events with fake $\not{E}_T$
- Signal regions (SRs) further refined by requiring fired triggers, low noise levels, fully operational detector
- At least 1 or 2 hadronic jets to catch events with $W \to q\bar{q}'$ in 3ℓ or $2\ell SS/+1\tau_h$

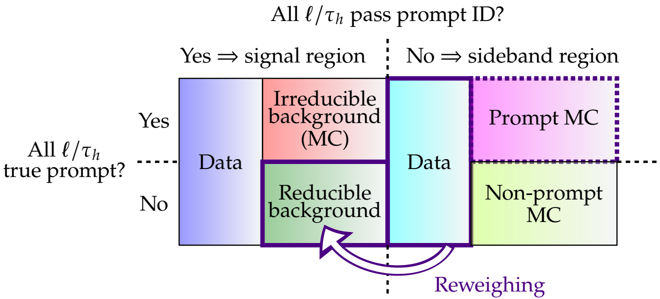
2ℓSS + $1\tau_h$ with extra jets from 2 quarks:





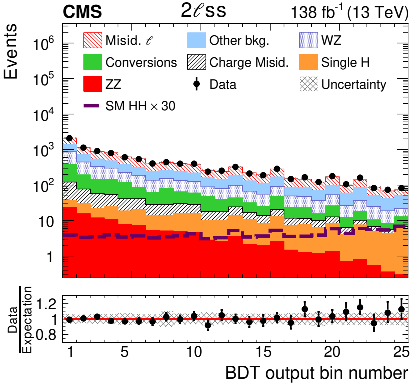
Background modeling

- Irreducible backgrounds, *i.e.* those with genuine prompt ℓ/τ_h estimated with Monte Carlo (MC) simulation (*e.g.*, VV(V), $t\bar{t}$V)
- Reducible backgrounds with no prompt ℓ/τ_h or with incorrect charge estimate have huge cross sections, despite low selection efficiency and tiny phase space
  - Prohibitive to generate the necessary amount of MC events
    $\Rightarrow$ extract from the data
  - Idea: devise additional criteria for ℓ/τ_h (prompt ID, „tight" charge), select only those events in the SR; data events with ℓ/τ_h failing those requirements are extrapolated to the SR by appropriately reweighing them

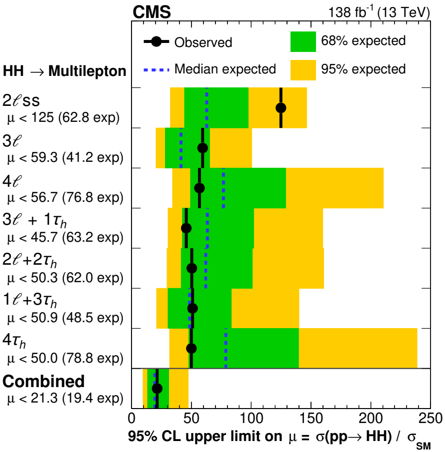


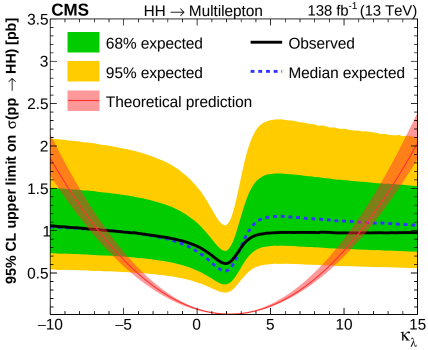
Signal extraction

- Dedicated boosted decision trees (BDTs) [5, 6] for each of the 7 signal categories
  - Floating point score between 0 (background-like) and 1 (signal-like)
  - Hyperparameters optimized using particle swarm algorithm [7] $\Rightarrow \mathcal{O}(10\%)$ gain in sensitivity (see talk by Laurits)
  - Training data maximized for MC statistics
  - Training variables: angular separation and invariant masses between pairs of ℓ, τ_h and jets, transverse momenta, $E\!\!\!/_T$, reconstructed m_{HH}, ...
- All distributions binned into histograms (with uncertainties), input to the ML fit



Results

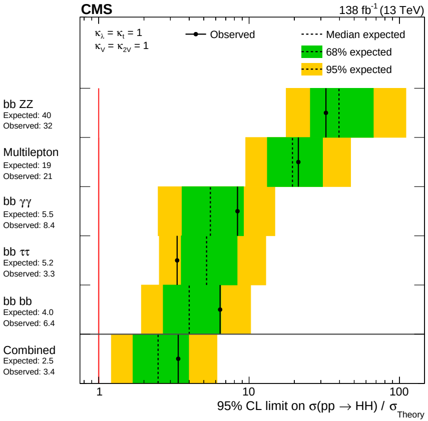


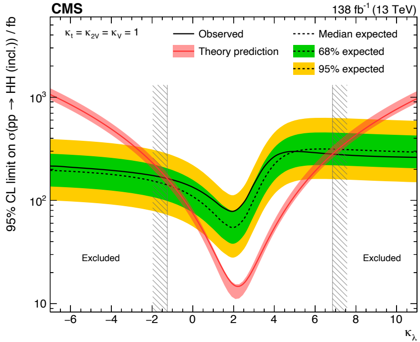


- ▷ Observed (expected): $-6.9 < \kappa_\lambda < 11.1$ ($-6.9 < \kappa_\lambda < 11.7$)
- ▷ Good sensitivity at high $|\kappa_\lambda|$

HH combination

Our analysis results are also featured in an HH combination paper published in Nature for celebrating 10 year anniversary of Higgs boson discovery [8]:





▷ $-1.24 < \kappa_\lambda < 6.49$ ($-1.01 < \kappa_\lambda < 6.31$)

⇒ Every analysis matters

Revisit also Xandra's talk!

Summary & outlook

In conclusion:

- Search for HH signal in multilepton channels based on 138 fb^{-1} of pp collision data produced at the LHC in 2016–2018
- No evidence for signal found
- Observed (expected) upper limit on SM cross section: 21.3 (19.4)
- κ_λ constrained between −6.9 and 11.1 (−6.9 and 11.7)
- Other results discussed in backup (EFT benchmarks and scans, resonant production)

Future plans:

- Based on LHC Run 2 results presented here: combination with more recent HH analyses (non-resonant, resonant)
- For LHC Run 3:
  - Optimize for qqHH signal and κ_{2V} extraction ⇒ see Norman's talk
  - Improve aspects of signal extraction (*e.g.* increase training statistics)

Backup

Signal modeling

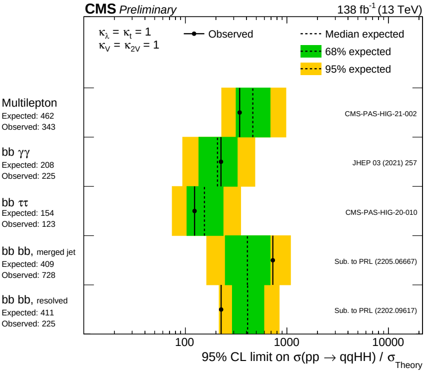
Signal at different values of κ_λ obtained by reweighing next-to-LO (NLO) MC simulation

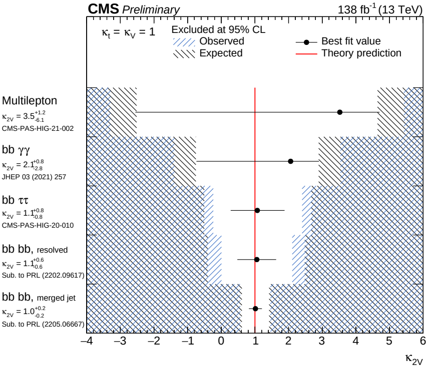
- „Matrix-based" reweighing
- $\mathcal{M}(\kappa_\lambda, \kappa_t) = \kappa_\lambda \kappa_t \mathcal{M}_\triangle + \kappa_t^2 \mathcal{M}_\square \Rightarrow N(\kappa_\lambda) \sim \sigma(\kappa_\lambda) \sim |\mathcal{M}(\kappa_\lambda)|^2 = \sum_{n=0}^{2} A_n \kappa_\lambda^n$
- Pick (at least) three different values of κ_λ (*e.g.* a, b, c), solve for $\{A_n\}_{n=0}^{2}$:

$$\begin{pmatrix} A_0 \\ A_1 \\ A_2 \end{pmatrix} = \begin{pmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{pmatrix}^{-1} \cdot \begin{pmatrix} N(a) \\ N(b) \\ N(c) \end{pmatrix}$$

$\Rightarrow$ By initially having MC estimates for all three values of κ_λ, it is possible to obtain signal estimates for an arbitrary value of κ_λ

Upper limits on qqHH production, exclusion limits on κ_{2V}





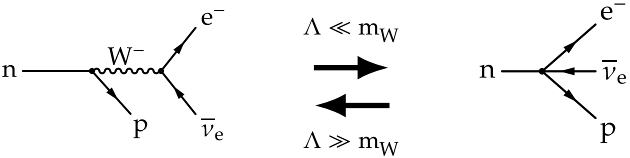
▷ $\kappa_{2V} = 0$ excluded at 6.6σ, assuming SM for all other couplings

Effective field theory (EFT)

Effective field theory (EFT) tells that after integrating out higher modes beyond scale Λ from a more „complete" beyond the SM (BSM) theory, one would be left with the following terms in addition to the SM:

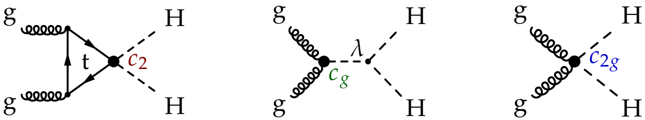
$$\mathcal{L}_{\text{EFT}} = \sum_{d>4} \frac{c_i(\Lambda)}{\Lambda^{d-4}} \mathcal{O}_i \qquad (\text{with } [c_i] = 0 \text{ and } [\mathcal{O}_i] = d)\,.$$

- Non-renormalizable interactions $\mathcal{O}_i$ become more important with increasing energy $\sqrt{s}$ („signs of new physics in tails of distributions") $\Leftrightarrow$ decreasing Λ
- No need to know the UV-completing theory $\Rightarrow$ ideal for initial gauging
- Examples: gravity, β decay (☟)



EFT shape benchmarks

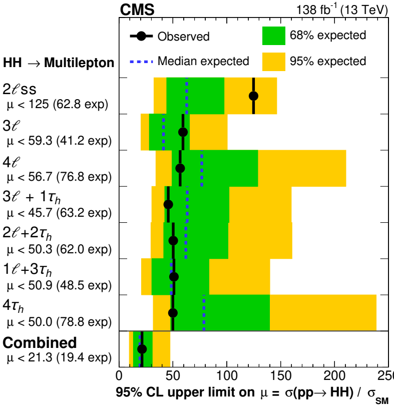
- Based on ggHH (SM: κ_λ, κ_t; BSM: c_2, c_g, c_{2g})
- Small changes in coupling values can drastically alter production rate and event-level kinematics
- Idea: group together coupling scenarios with similar kinematic features
  - Based on similarities of m_{HH} and $\cos\theta^*$ at LO $\Rightarrow$ 12(+1) shape BMs („JHEP04”) [9, 10]
  - Based on the latent space of an autoencoder that was trained on m_{HH} shape of next-to-LO (NLO) signal $\Rightarrow$ 7 shape BMs („JHEP03”) [11]
- Signal modeled by reweighing with a function of $\{\kappa_\lambda, \kappa_t, c_2, c_g, c_{2g}\}$ coupling parameters in bins of $(m_{\mathrm{HH}}, \cos\theta^*)$
- BDT training parametrized by 12 JHEP04 EFT BMs + SM with one-hot encoding, all other coupling scenarios mapped to those

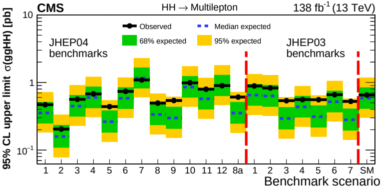


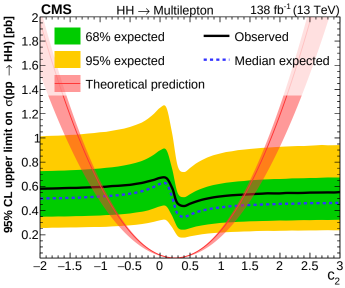
EFT coupling scenarios

Benchmark	κ_λ	κ_t	c_2	c_g	c_{2g}
JHEP04 BM1	7.5	1.0	−1.0	0.0	0.0
JHEP04 BM2	1.0	1.0	0.5	−0.8	0.6
JHEP04 BM3	1.0	1.0	−1.5	0.0	−0.8
JHEP04 BM4	−3.5	1.5	−3.0	0.0	0.0
JHEP04 BM5	1.0	1.0	0.0	0.8	−1.0
JHEP04 BM6	2.4	1.0	0.0	0.2	−0.2
JHEP04 BM7	5.0	1.0	0.0	0.2	−0.2
JHEP04 BM8	15.0	1.0	0.0	−1.0	1.0
JHEP04 BM8a	1.0	1.0	0.5	4/15	0.0
JHEP04 BM9	1.0	1.0	1.0	−0.6	0.6
JHEP04 BM10	10.0	1.5	−1.0	0.0	0.0
JHEP04 BM11	2.4	1.0	0.0	1.0	−1.0
JHEP04 BM12	15.0	1.0	1.0	0.0	0.0
JHEP03 BM1	3.94	0.94	−1/3	0.75	−1
JHEP03 BM2	6.84	0.61	1/3	0	1
JHEP03 BM3	2.21	1.05	−1/3	0.75	−1.5
JHEP03 BM4	2.79	0.61	1/3	−0.75	−0.5
JHEP03 BM5	3.95	1.17	−1/3	0.25	1.5
JHEP03 BM6	5.68	0.83	1/3	−0.75	−1
JHEP03 BM7	−0.10	0.94	1	0.25	0.5
SM	1.0	1.0	0.0	0.0	0.0

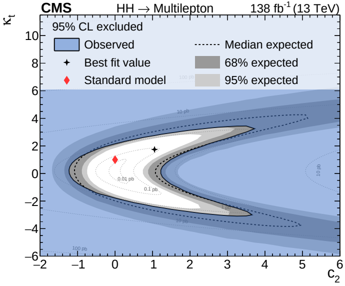
Results: non-resonant BSM production





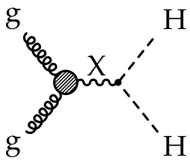


$$-1.05 < c_2 < 1.48 \; (-0.96 < c_2 < 1.37)$$

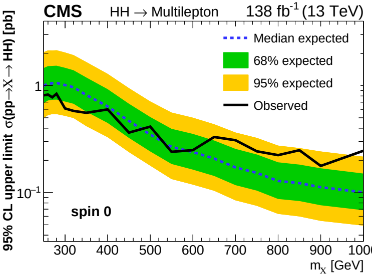


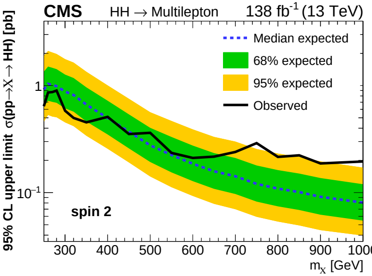
Resonant HH production

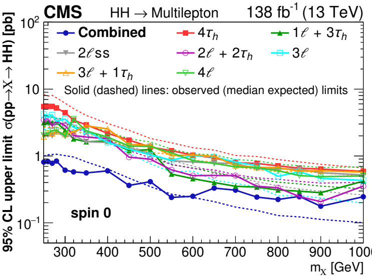
- Narrow-width spin-0 and spin-2 resonance („X")
  - Radion (spin-0) [12] and Kaluza-Klein bulk graviton (spin-2) [13] from Randall-Sundrum model [14–17]
- Dedicated LO signal datasets for both spin hypotheses with resonant mass $m_X \in [250, 1000]$ GeV in steps of 10–100 GeV (19 in total)
- Separate BDT trainings for spin-0 and spin-2, parametrized by m_{HH}

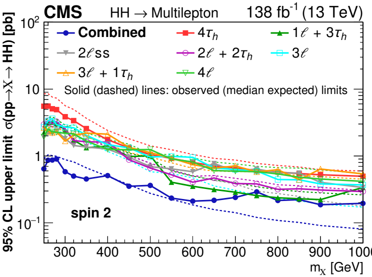


Results: resonant production

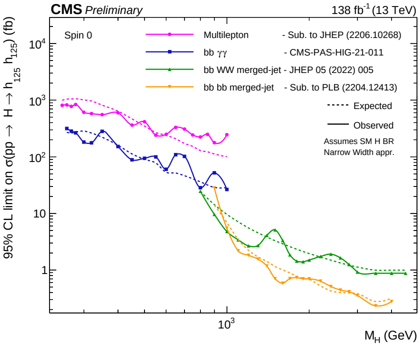


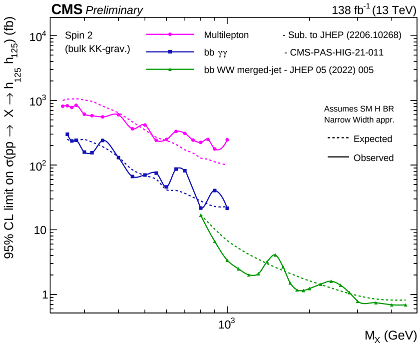






Results: resonant production – all HH analyses in CMS





Bibliography I

[1] Giuseppe Degrassi et al. "Higgs mass and vacuum stability in the Standard Model at NNLO". In: *JHEP* 08 (2012), p. 098. DOI: 10.1007/JHEP08(2012)098. arXiv: 1205.6497 [hep-ph].

[2] ATLAS collaboration. "Constraint of the Higgs boson self-coupling from Higgs boson differential production and decay measurements". In: (2019).

[3] CMS collaboration. "Combined Higgs boson production and decay measurements with up to 137 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV". In: (2020).

[4] I. Zurbano Fernandez et al. "High-Luminosity Large Hadron Collider (HL-LHC): Technical design report". In: 10/2020 (2020). Ed. by I. Béjar Alonso et al. DOI: 10.23731/CYRM-2020-0010.

[5] Tianqi Chen and Carlos Guestrin. "XGBoost: A Scalable Tree Boosting System". In: *CoRR* abs/1603.02754 (2016). arXiv: 1603.02754. URL: http://arxiv.org/abs/1603.02754.

[6] Fabian Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: *CoRR* abs/1201.0490 (2012). arXiv: 1201.0490. URL: http://arxiv.org/abs/1201.0490.

Bibliography II

[7] Laurits Tani et al. "Evolutionary algorithms for hyperparameter optimization in machine learning for application in high energy physics". In: *Eur. Phys. J. C* 81.2 (2021), p. 170. DOI: 10.1140/epjc/s10052-021-08950-y. arXiv: 2011.04434 [hep-ex].

[8] CMS collaboration. "A portrait of the Higgs boson by the CMS experiment ten years after the discovery". In: (2022). DOI: 10.1038/s41586-022-04892-x. arXiv: 2207.00043 [hep-ex].

[9] Alexandra Carvalho et al. "Higgs Pair Production: Choosing Benchmarks With Cluster Analysis". In: *JHEP* 04 (2016), p. 126. DOI: 10.1007/JHEP04(2016)126. arXiv: 1507.02245 [hep-ph].

[10] G. Buchalla et al. "Higgs boson pair production in non-linear Effective Field Theory with full m_t-dependence at NLO QCD". In: *JHEP* 09 (2018), p. 057. DOI: 10.1007/JHEP09(2018)057. arXiv: 1806.05162 [hep-ph].

[11] Matteo Capozi and Gudrun Heinrich. "Exploring anomalous couplings in Higgs boson pair production through shape analysis". In: *JHEP* 03 (2020), p. 091. DOI: 10.1007/JHEP03(2020)091. arXiv: 1908.08923 [hep-ph].

Bibliography III

[12] Walter D. Goldberger and Mark B. Wise. "Modulus stabilization with bulk fields". In: *Phys. Rev. Lett.* 83 (1999), pp. 4922–4925. DOI: 10.1103/PhysRevLett.83.4922. arXiv: hep-ph/9907447.

[13] Walter D. Goldberger and Mark B. Wise. "Bulk fields in the Randall-Sundrum compactification scenario". In: *Phys. Rev. D* 60 (1999), p. 107505. DOI: 10.1103/PhysRevD.60.107505. arXiv: hep-ph/9907218.

[14] Lisa Randall and Raman Sundrum. "A Large mass hierarchy from a small extra dimension". In: *Phys. Rev. Lett.* 83 (1999), pp. 3370–3373. DOI: 10.1103/PhysRevLett.83.3370. arXiv: hep-ph/9905221.

[15] Lisa Randall and Raman Sundrum. "An Alternative to compactification". In: *Phys. Rev. Lett.* 83 (1999), pp. 4690–4693. DOI: 10.1103/PhysRevLett.83.4690. arXiv: hep-th/9906064.

[16] Csaba Csaki, Michael L. Graesser, and Graham D. Kribs. "Radion dynamics and electroweak physics". In: *Phys. Rev. D* 63 (2001), p. 065002. DOI: 10.1103/PhysRevD.63.065002. arXiv: hep-th/0008151.

[17] A. Liam Fitzpatrick et al. "Searching for the Kaluza-Klein Graviton in Bulk RS Models". In: *JHEP* 09 (2007), p. 013. DOI: 10.1088/1126-6708/2007/09/013. arXiv: hep-ph/0701150.