Search for di-Higgs boson production in multi-lepton final states at CMS

Dr. Sandeep Bhowmik
for the CMS collaboration

NICPB Tallinn Estonia

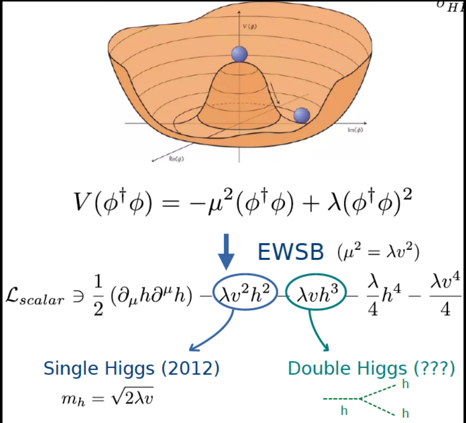
December 12, 2022

IISER Mohali, India

Outline

Motivation

- Since the discovery of the Higgs (H) boson, one important property that remains largely unknown is the H boson self-coupling (λ).
- A precise measurement of this coupling is necessary to determine the shape of the Higgs potential, and thus verify that the mechanism breaking the electroweak gauge symmetry is indeed the Higgs mechanism of the standard model (SM)
- The strength of the trilinear self-coupling, can be determined using measurements of H boson pair (HH) production.



$$V(\phi^\dagger\phi) = -\mu^2(\phi^\dagger\phi) + \lambda(\phi^\dagger\phi)^2$$

EWSB $(\mu^2 = \lambda v^2)$

$$\mathcal{L}_{scalar} \ni \frac{1}{2}(\partial_\mu h \partial^\mu h) - \lambda v^2 h^2 - \lambda v h^3 - \frac{\lambda}{4}h^4 - \frac{\lambda v^4}{4}$$

Single Higgs (2012)

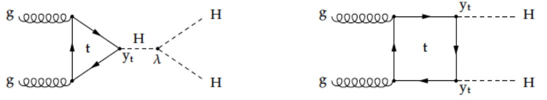
$m_h = \sqrt{2\lambda}v$

Double Higgs (???)

- $\sigma_{HH} \sim \frac{\sigma_H}{1000}$ at $\sqrt{s} = 13$ TeV not so easy to measure
- But there are some good results published by ATLAS and CMS

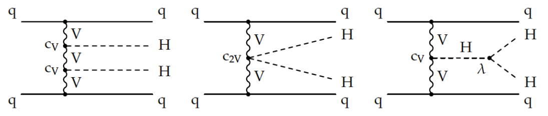
HH production mechanism

The SM gluon fusion (ggHH) process



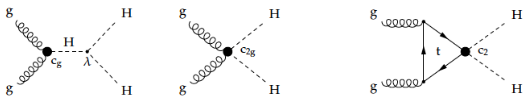
- Cross section $\sigma^{SM}_{ggHH} = 31.1^{+2.1}_{-7.2}$ fb at $\sqrt{s} = 13$ TeV

The SM vector boson fusion (qqHH) process



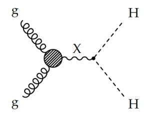
- Cross section $\sigma^{SM}_{qqHH} = 1.73 \pm 0.04$ fb at $\sqrt{s} = 13$ TeV

HH BSM production (EFT)



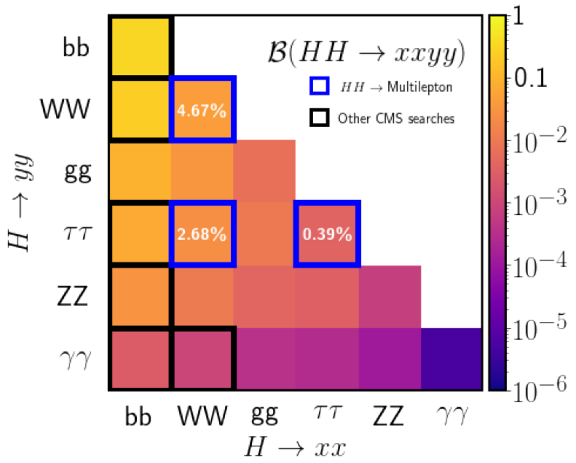
- Various effective field theories (EFT) predict non-SM like Higgs couplings

HH resonance production



- Various BSM studies postulate new heavy particle (X), which can either have spin-0 or spin-2, that decays into a pair of Higgs bosons.
- the X is searched in the HH mass range between 250 GeV to 1 TeV.

HH decays



- No golden channel and each channel comes with its own challenges
- But established HH analyses use decay modes with at least 1 $H \to bb$ decay offering large branching fraction (BR)

$HH \to multilepton$

- The multilepton analysis targets the decay modes
- $HH \to WWWW$
- $HH \to WW\tau\tau$
- $HH \to \tau\tau\tau\tau$
- Covering ∼ 7.7% of the HH decays

Channels

- Events are subdivided into seven mutually exclusive "search categories" based on ℓ(e, μ) and τ_h multiplicity:
- 4ℓ,
- $3\ell + 0\tau_h$,
- $3\ell + 1\tau_h$,
- $2\ell + 2\tau$,
- $2\ell ss + 0/1\tau_h$, (same-sign $\ell\ell$ pair)
- $1\ell + 3\tau_h$,
- $0\ell + 4\tau_h$

Data samples and Monte Carlo (MC) simulation

Data samples

- The analyzed pp collision data correspond to an integrated luminosity of 138 fb^{-1}, collected by the CMS detector over three years: 36 fb^{-1} in 2016, 42 fb^{-1} in 2017, and 60 fb^{-1} in 2018

Signal MC samples

- A variety of HH signal samples were generated at LO and at NLO accuracy in QCD to simulate nonresonant HH production, covering the ggHH and qqHH production processes.
- Separate ggHH samples are produced for SM HH production and for a total of 20 EFT benchmark (BM) scenarios.
- Resonant HH production was simulated at LO for both spin-0 (radion) and spin-2 (graviton) scenarios.
- Each of the H bosons was forced to decay to either WW, ZZ, or $\tau\tau$ in these samples.

Background MC samples

- Background MC samples include processes producing:
- a single W or Z boson,
- two bosons (WW, WZ, ZZ, Wg, and Zg),
- three bosons (WWW, WWZ, WZZ, ZZZ, and WZγ),
- a single H boson (via gluon fusion, vector boson fusion, or associated production with a W or Z boson),
- a single top quark, a top quark-antiquark pair ($t\bar{t}$),
- and top quarks associated with one or more bosons ($t\bar{t}$W, $t\bar{t}$Z, $t\bar{t}$H, tHq, and tHW).

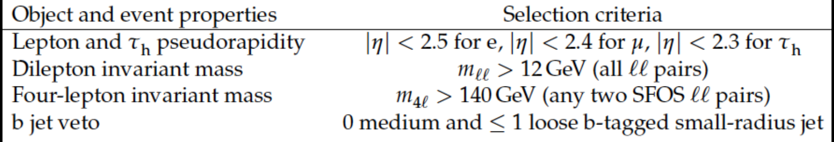
Event reconstruction

- The CMS particle-flow (PF) algorithm aims to reconstruct and identify each individual particle in an event, using an optimized combination of information from the various elements of the CMS detector.
- The particles are subsequently classified into five mutually exclusive types: electrons, muons, photons, and charged and neutral hadrons.
- These particles are then combined to reconstruct hadronic τ decays, jets, and the missing transverse momentum in the event.
- Based on the identification criteria, this analysis makes use of three different levels of lepton selection for both the electrons and the muons: the "loose", "fakeable" and "tight" selections.
- The loose lepton collection is used for overlap cleaning amongst the objects and for the computation of invariant mass quantities used for background rejection.
- The fakeable leptons are used to estimate the fake lepton background from control regions in data, as well as to compute global kinematic properties of the events.
- The tight lepton selection is similar to the fakeable one but with more stringent prompt lepton MVA requirements, and is used for event selection in the signal regions.

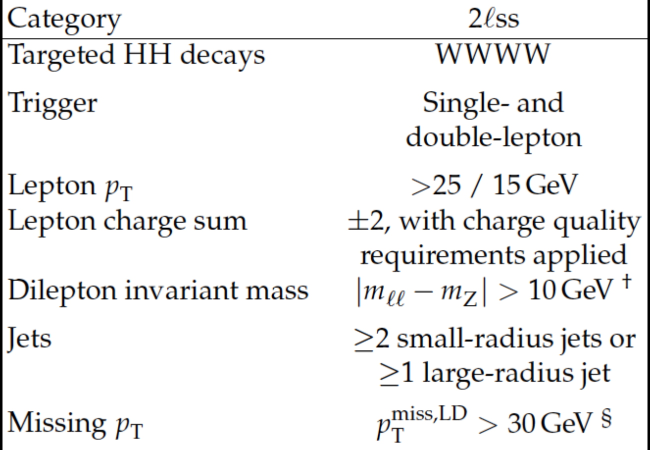
Background estimation

- Lepton/τ_h fake background:
  Non-prompt lepton or jet faking as prompt-lepton/τ_h.
  Estimated in data using fake factor method.
- Lepton charge flip background:
  Background from mis-identication of the charge of leptons.
  Relevant only in $2\ell ss + 0/1\tau_h$ channel.
  Estimated in data.
- Photon conversion background:
  Estimated using MC simulation from events with reconstructed leptons match to generator-level photon ($\Delta R < 0.3$).
- Irreducible background, such as $Z/\gamma^* \rightarrow ee/\mu\mu$, VH, VVV, estimated using MC simulation.
  Exception: WZ, ZZ
- Dedicated control regions (CRs) to estimate WZ, ZZ background.

Event selection

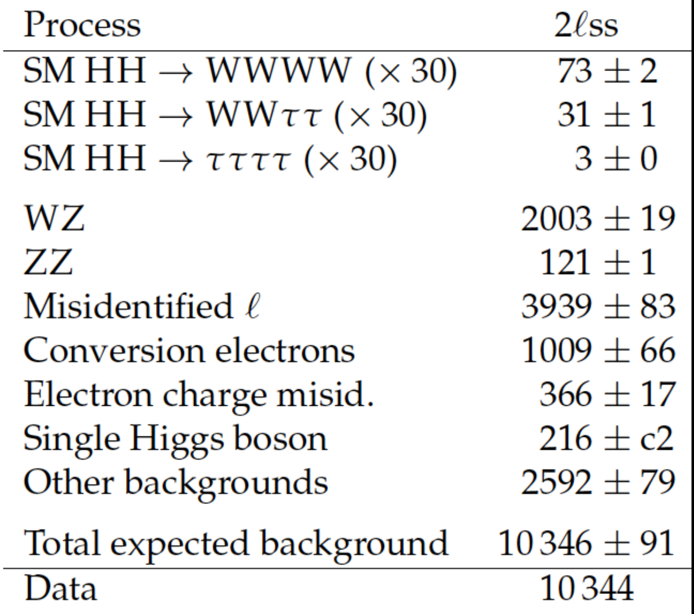


Object and event properties	Selection criteria
Lepton and τ_h pseudorapidity	$\lvert\eta\rvert < 2.5$ for e, $\lvert\eta\rvert < 2.4$ for μ, $\lvert\eta\rvert < 2.3$ for τ_h
Dilepton invariant mass	$m_{\ell\ell} > 12\,\mathrm{GeV}$ (all $\ell\ell$ pairs)
Four-lepton invariant mass	$m_{4\ell} > 140\,\mathrm{GeV}$ (any two SFOS $\ell\ell$ pairs)
b jet veto	0 medium and ≤ 1 loose b-tagged small-radius jet



Category	2ℓss
Targeted HH decays	WWWW
Trigger	Single- and double-lepton
Lepton p_T	>25 / $15\,\mathrm{GeV}$
Lepton charge sum	± 2, with charge quality requirements applied
Dilepton invariant mass	$\lvert m_{\ell\ell} - m_Z\rvert > 10\,\mathrm{GeV}$ †
Jets	$\geq$2 small-radius jets or $\geq$1 large-radius jet
Missing p_T	$p_T^{\mathrm{miss,LD}} > 30\,\mathrm{GeV}$ §

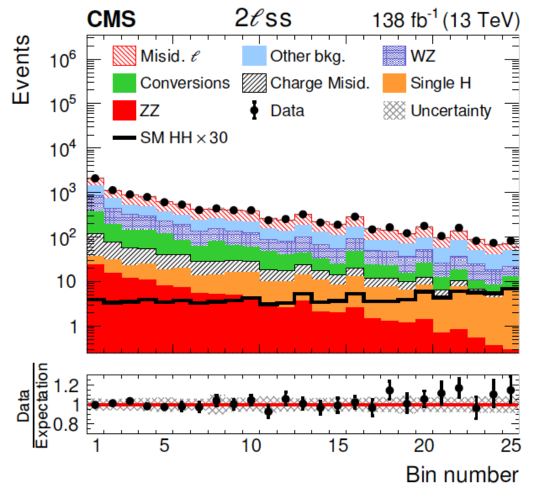
The event yields



Process	$2\ell ss$
SM HH → WWWW (× 30)	73 ± 2
SM HH → WW$\tau\tau$ (× 30)	31 ± 1
SM HH → $\tau\tau\tau\tau$ (× 30)	3 ± 0
WZ	2003 ± 19
ZZ	121 ± 1
Misidentified ℓ	3939 ± 83
Conversion electrons	1009 ± 66
Electron charge misid.	366 ± 17
Single Higgs boson	216 ± c2
Other backgrounds	2592 ± 79
Total expected background	$10\,346 \pm 91$
Data	10 344

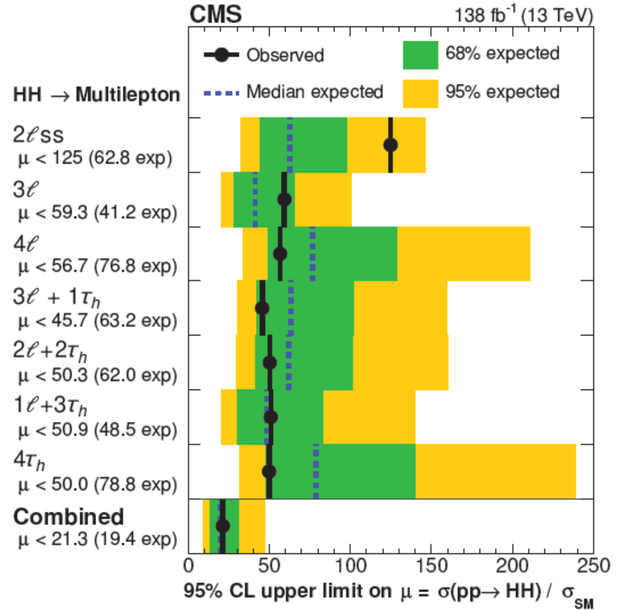
Signal extraction

- The distributions in the output of boosted decision tree (BDT) classifiers, which are trained to discriminate the HH signal from backgrounds.
- The distributions in the BDT outputs for the 7 channels plus the distributions in m_T for the WZ and in m_{4l} for the ZZ control region are used as input to a simultaneous maximum likelihood (combine) fit.
- The rate of the HH signal is extracted through a binned maximum likelihood (ML) fit to the BDT
- Three classifiers are trained for each of the seven search categories, targeting nonresonant HH production and resonant HH production from the decay of heavy particles of spin 0 and of spin 2.



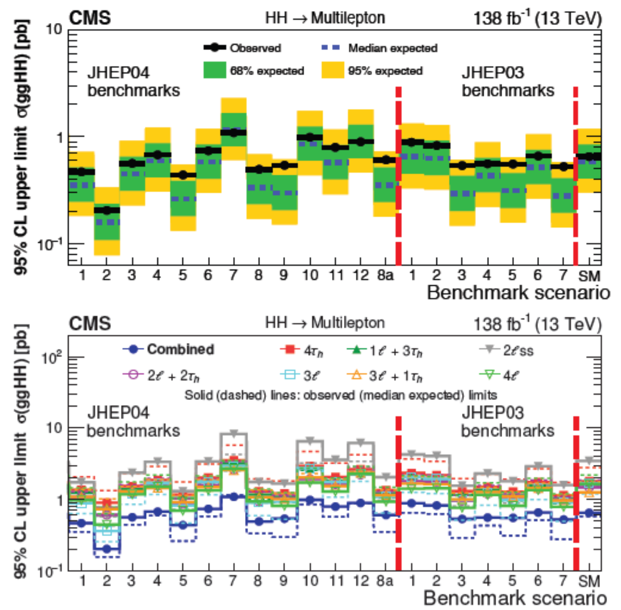
Results: the SM HH production

- Observed and expected 95% CL upper limits on the SM HH production cross section, obtained for both individual search categories and from a simultaneous fit of all seven categories combined.
- The observed (expected) 95% CL upper limit on the cross section for nonresonant HH production is 651 (592) fb.



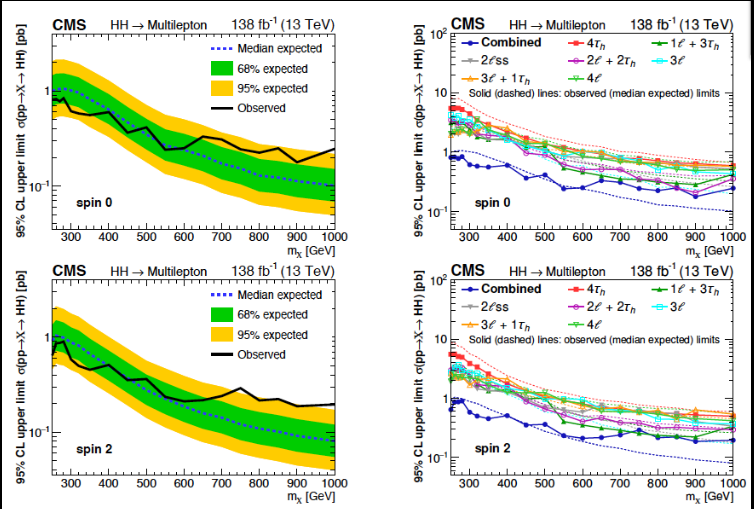
Results: HH BSM production (EFT)

- Observed and expected 95% CL upper limits on the HH production cross section for the twenty benchmark scenarios and for the SM, obtained for both individual search categories and from a simultaneous fit of all seven categories combined.
- The observed (expected) 95% CL upper limits on nonresonant HH production in the different benchmark scenarios range from 0.21 to 1.09 (0.16 to 1.16) pb, depending on the scenario.



Results: HH resonance production

- Observed and expected 95% CL upper limits on the production of new particles X of spin 0 (upper) and spin 2 (lower) and mass m_X in the range 250-1000 GeV, which decay to H boson pairs. The plot on the left shows the result obtained by combining all seven search categories, while the plot on the right shows the limits obtained for each category separately, and the combined limit.
- The observed (expected) 95% CL upper limits on the resonant HH production cross section range from 0.18 to 0.90 (0.08 to 1.06) pb, depending on the mass and spin.



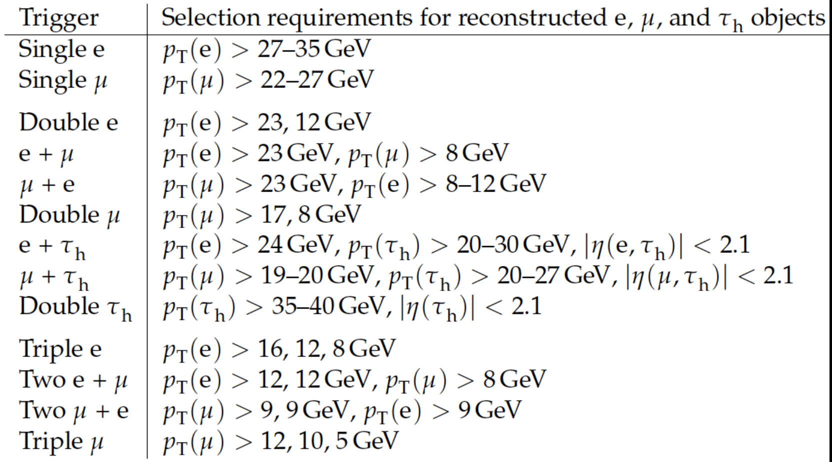
Summary

- The results of a search for nonresonant Higgs boson pair (HH) production in final states with multiple reconstructed leptons, including electrons and muons (ℓ) as well as hadronically decaying tau leptons (τ_h), has been presented.
- The search targets the HH decay modes WWWW, WW$\tau\tau$, and $\tau\tau\tau\tau$, using proton-proton collision data recorded by the CMS experiment at a center-of-mass energy of 13 TeV and corresponding to an integrated luminosity of 138 fb^{-1}
- Seven search categories, distinguished by ℓ and τ_h multiplicity, are included in the analysis: 4ℓ, $3\ell + 0\tau_h$, $3\ell + 1\tau_h$, $2\ell + 2\tau$, $2\ell ss + 0/1\tau_h$, $1\ell + 3\tau_h$, and $0\ell + 4\tau_h$ where ss indicates an $\ell\ell$ pair with the same charge.
- No evidence for a signal is found in the data.
- Upper limits on the cross section for nonresonant HH production is set.
- For nonresonant HH production with event kinematics as predicted by the standard model (SM), the observed (expected) 95% confidence level (CL) upper limit on the HH production rate is 21.3 (19.4) times the rate expected in the SM.
- The observed (expected) limits on the nonresonant HH production cross section in twenty EFT benchmark scenarios range from 0.21 to 1.09 (0.16 to 1.16) pb at 95% CL, depending on the scenario.
- The observed (expected) 95% CL upper limits on the cross section for resonant HH production range from 0.18 to 0.90 (0.08 to 1.06) pb, depending on the mass and spin of the resonance.
- These results are already published in INSPIRE, arXive

Thank You

Backup

Trigger requirements



Trigger	Selection requirements for reconstructed e, μ, and τ_h objects
Single e	$p_\mathrm{T}(\mathrm{e}) > 27$–$35\,\mathrm{GeV}$
Single μ	$p_\mathrm{T}(\mu) > 22$–$27\,\mathrm{GeV}$
Double e	$p_\mathrm{T}(\mathrm{e}) > 23, 12\,\mathrm{GeV}$
e + μ	$p_\mathrm{T}(\mathrm{e}) > 23\,\mathrm{GeV}$, $p_\mathrm{T}(\mu) > 8\,\mathrm{GeV}$
μ + e	$p_\mathrm{T}(\mu) > 23\,\mathrm{GeV}$, $p_\mathrm{T}(\mathrm{e}) > 8$–$12\,\mathrm{GeV}$
Double μ	$p_\mathrm{T}(\mu) > 17, 8\,\mathrm{GeV}$
e + τ_h	$p_\mathrm{T}(\mathrm{e}) > 24\,\mathrm{GeV}$, $p_\mathrm{T}(\tau_\mathrm{h}) > 20$–$30\,\mathrm{GeV}$, $\lvert\eta(\mathrm{e}, \tau_\mathrm{h})\rvert < 2.1$
μ + τ_h	$p_\mathrm{T}(\mu) > 19$–$20\,\mathrm{GeV}$, $p_\mathrm{T}(\tau_\mathrm{h}) > 20$–$27\,\mathrm{GeV}$, $\lvert\eta(\mu, \tau_\mathrm{h})\rvert < 2.1$
Double τ_h	$p_\mathrm{T}(\tau_\mathrm{h}) > 35$–$40\,\mathrm{GeV}$, $\lvert\eta(\tau_\mathrm{h})\rvert < 2.1$
Triple e	$p_\mathrm{T}(\mathrm{e}) > 16, 12, 8\,\mathrm{GeV}$
Two e + μ	$p_\mathrm{T}(\mathrm{e}) > 12, 12\,\mathrm{GeV}$, $p_\mathrm{T}(\mu) > 8\,\mathrm{GeV}$
Two μ + e	$p_\mathrm{T}(\mu) > 9, 9\,\mathrm{GeV}$, $p_\mathrm{T}(\mathrm{e}) > 9\,\mathrm{GeV}$
Triple μ	$p_\mathrm{T}(\mu) > 12, 10, 5\,\mathrm{GeV}$

Trigger	$0\ell + 4\tau_h$	$1\ell + 3\tau_h$	$2\ell ss + 0/1\tau_h$	$2\ell + 2\tau_h$	$3\ell + 0\tau_h$	$3\ell + 1\tau_h$	4ℓ
Double τ_h trigger	✓	✓	X	X	X	X	X
Lepton + τ_h cross trigger	X	✓	X	X	X	X	X
Single lepton trigger	X	✓	✓	✓	✓	✓	✓
Double lepton trigger	X	X	✓	✓	✓	✓	✓
Triple lepton trigger	X	X	X	X	✓	✓	✓

Data to MC corrections

In order to improve the modeling of the data, we apply corrections to simulated events, which we denote as “data-to-Monte Carlo” corrections.

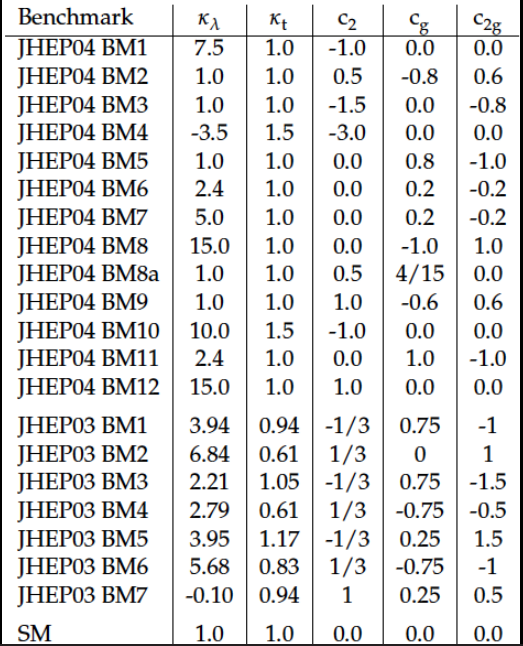
- pileup reweighting
- trigger efficiency
- e and μ identification and isolation efficiency
- τ_h identification efficiency
- τ_h energy scale
- b-tag efficiency
- E_T^{miss} resolution
- prefiring probability of Level-1 ECAL trigger
- reweighting of $t\bar{t}$ events

Systematic uncertainties

- JES, JER and b-tag ID
- Unclustered MET
- Luminosity
- Pileup and top p_T reweighting
- Detector issues (pre-ring, HEM issue)
- Uncertainties in signal and background rates
- Uncertainties in $H \rightarrow \tau\tau$ / WW / ZZ branching fraction
- Dipole recoil scheme for qqHH MC
- Trigger SF uncertainties
- Electron/muon/tau ID and energy SF uncertainties
- Flips background: 30% normalization uncertainty
- Fakes background: 30% normalization and shape uncertainties; w/ additional normalization and shape uncertainties on MC closure
- 30% (50%) normalization uncertainty on conversions (rare) backgrounds

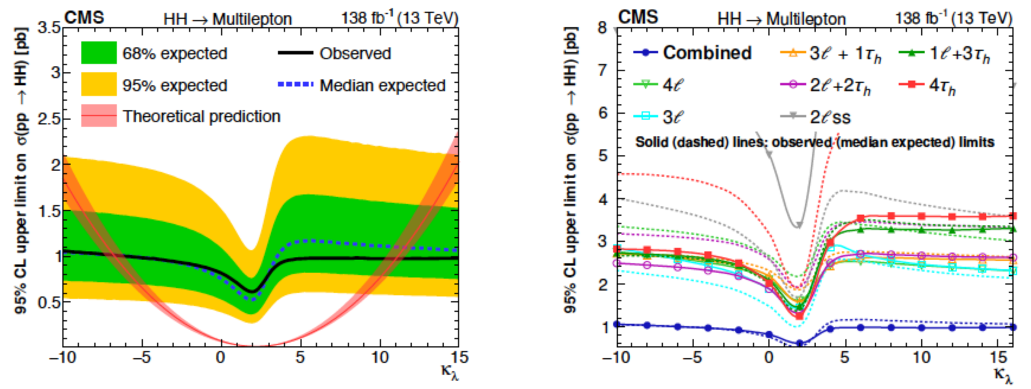
EFT benchmark (BM) scenarios

- Parameter values for k_λ, k_t , c_2, c_g, and c_{2g} in MC samples modeling twenty benchmark scenarios in the EFT approach, plus SM HH production.



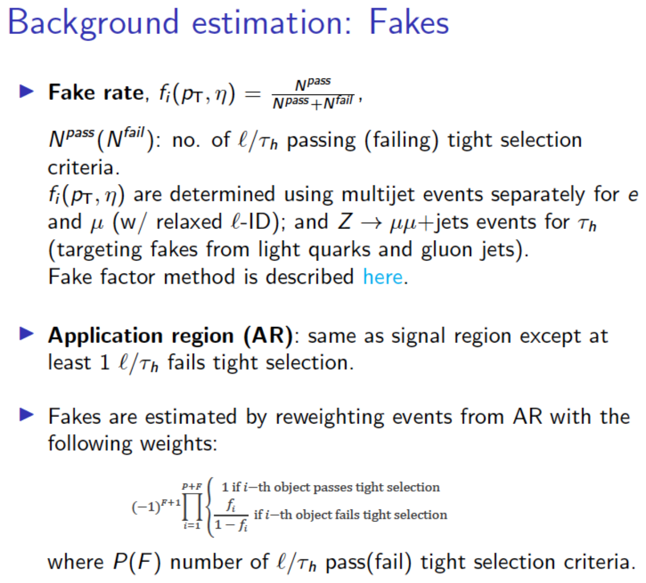
Benchmark	κ_λ	κ_t	c_2	c_g	c_{2g}
JHEP04 BM1	7.5	1.0	-1.0	0.0	0.0
JHEP04 BM2	1.0	1.0	0.5	-0.8	0.6
JHEP04 BM3	1.0	1.0	-1.5	0.0	-0.8
JHEP04 BM4	-3.5	1.5	-3.0	0.0	0.0
JHEP04 BM5	1.0	1.0	0.0	0.8	-1.0
JHEP04 BM6	2.4	1.0	0.0	0.2	-0.2
JHEP04 BM7	5.0	1.0	0.0	0.2	-0.2
JHEP04 BM8	15.0	1.0	0.0	-1.0	1.0
JHEP04 BM8a	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: H boson trilinear self-coupling strength modifier



- Observed (expected) 95% CL interval for the H boson trilinear self-coupling strength modifier is measured to be -6.9 $< k_\lambda <$ 11.1 (-6.9 $< k_\lambda <$ 11.7)

Background estimation



Background estimation: Fakes

- **Fake rate**, $f_i(p_T, \eta) = \frac{N^{pass}}{N^{pass}+N^{fail}}$,

  $N^{pass}(N^{fail})$: no. of ℓ/τ_h passing (failing) tight selection criteria.
  $f_i(p_T, \eta)$ are determined using multijet events separately for e and μ (w/ relaxed ℓ-ID); and $Z \rightarrow \mu\mu$+jets events for τ_h (targeting fakes from light quarks and gluon jets).
  Fake factor method is described here.

- **Application region (AR)**: same as signal region except at least 1 ℓ/τ_h fails tight selection.

- Fakes are estimated by reweighting events from AR with the following weights:

  $$(-1)^{F+1}\prod_{i=1}^{P+F}\begin{cases} 1 \text{ if } i\text{–th object passes tight selection} \\ \frac{f_i}{1-f_i} \text{ if } i\text{–th object fails tight selection} \end{cases}$$

  where $P(F)$ number of ℓ/τ_h pass(fail) tight selection criteria.

Background estimation