H→WW measurements with the CMS experiment

Adrián Álvarez Fernández (CIEMAT)
Jónatan Piedra (IFCA)

4th Red LHC Workshop, November 4th

Instituto de Física de Cantabria
GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA E INNOVACIÓN
Ciemat Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas
cfp CIEMAT física de partículas

Content of this talk

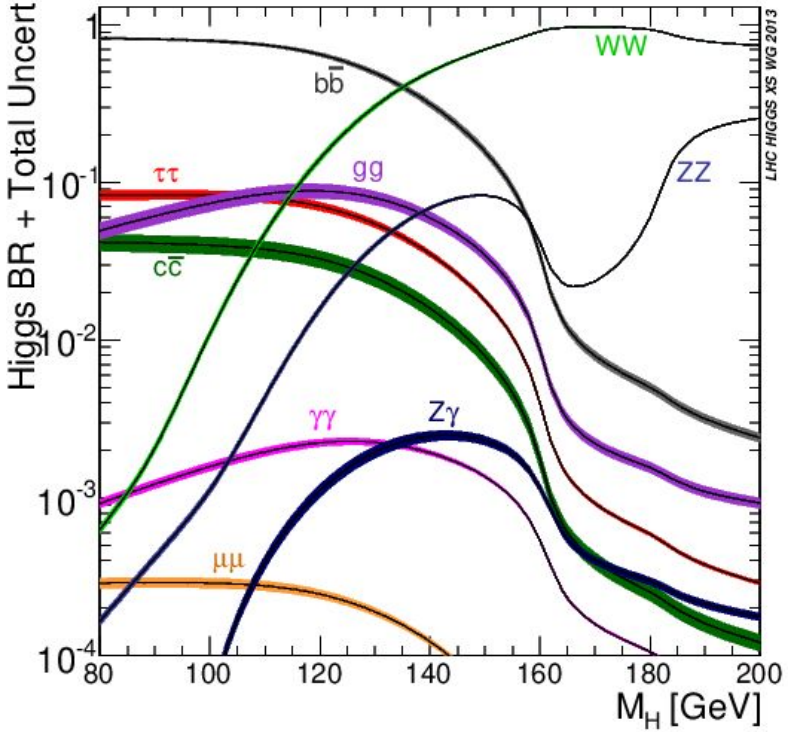
- First measurement of Higgs properties in the $H \rightarrow WW \rightarrow l\nu l\nu$ channel at $\sqrt{s}$=13 TeV. [1]
- Run 2 differential cross section measurement with respect to the Higgs boson transverse momentum and the jet multiplicity using 137 fb^{-1} of collected data. [2]
- Ongoing analysis with full Run 2 data.

[1] **"Measurements of properties of the Higgs boson decaying to a W boson pair in pp collisions at $\sqrt{s}$=13TeV"**. In: Physics Letters B 791 (2019), pp. 96–129. issn: 0370-2693. arxiv.1806.05246

[2] **"Measurement of the inclusive and differential Higgs boson production cross sections in the leptonic WW decay mode at $\sqrt{s}$ = 13 TeV".** arxiv.2007.01984

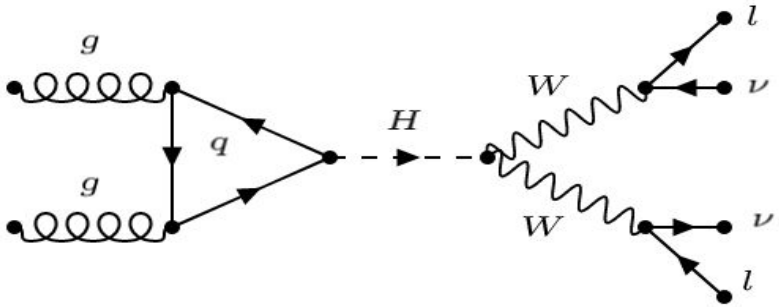
Introduction

- Higgs boson decay into a pair of W bosons has one of the highest branching ratios at m_H=125 GeV.
- Fully leptonic decay ($H \rightarrow WW \rightarrow l\nu l\nu$) has a clean signal from the isolated leptons that allows the study of the Higgs boson in a variety of production modes.
- Good signal sensitivity despite the large background, but low resolution due to neutrinos in the final state.
- Signal strength modifiers and Higgs boson couplings have been studied with 2016 data and now we are exploiting the full Run 2.



HWW event description

- Final state with two leptons (e or μ) and missing energy from the two W bosons.
- Categories with same or different flavour leptons.
  - ggH categories
  - VBF categories (two separated jets)
  - VH categories (2 jets or 3 or 4 leptons)

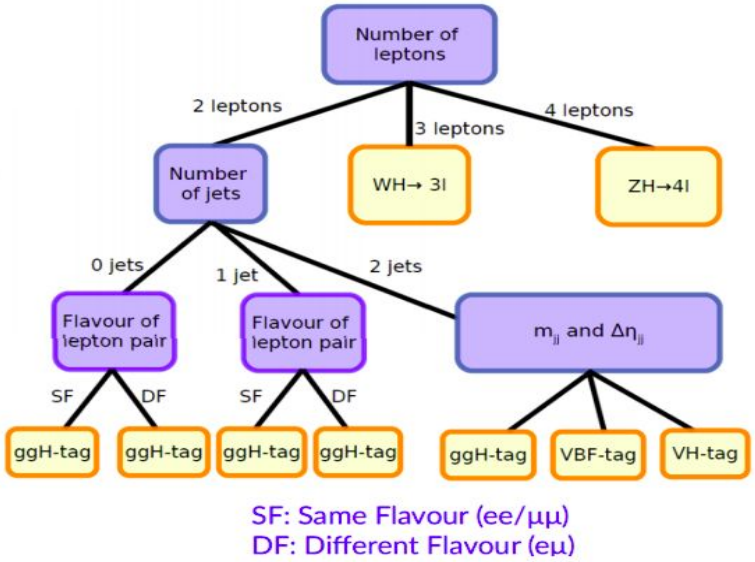


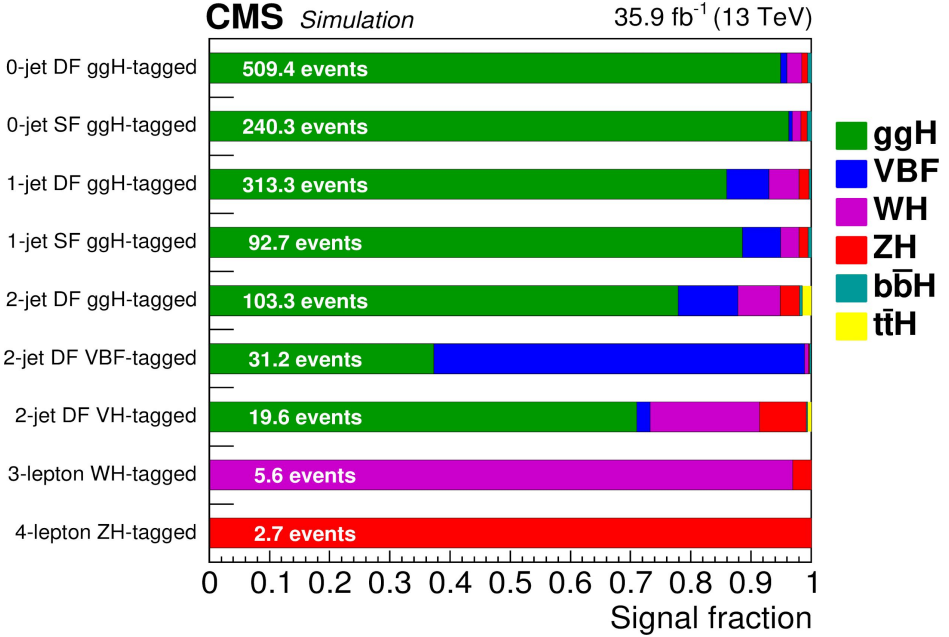
- Main backgrounds:
  - Nonresonant WW production
  - DY ($\tau\tau$ for different-flavour leptons)
  - Top quark production (tW and ttbar)
  - Nonprompt lepton background
    - Misidentified leptons or leptons from heavy-flavor hadron decays
    - Mainly W+jets and ttbar events
    - Validated with a control region of same-sign leptons

Dedicated control regions are used to estimate these two backgrounds

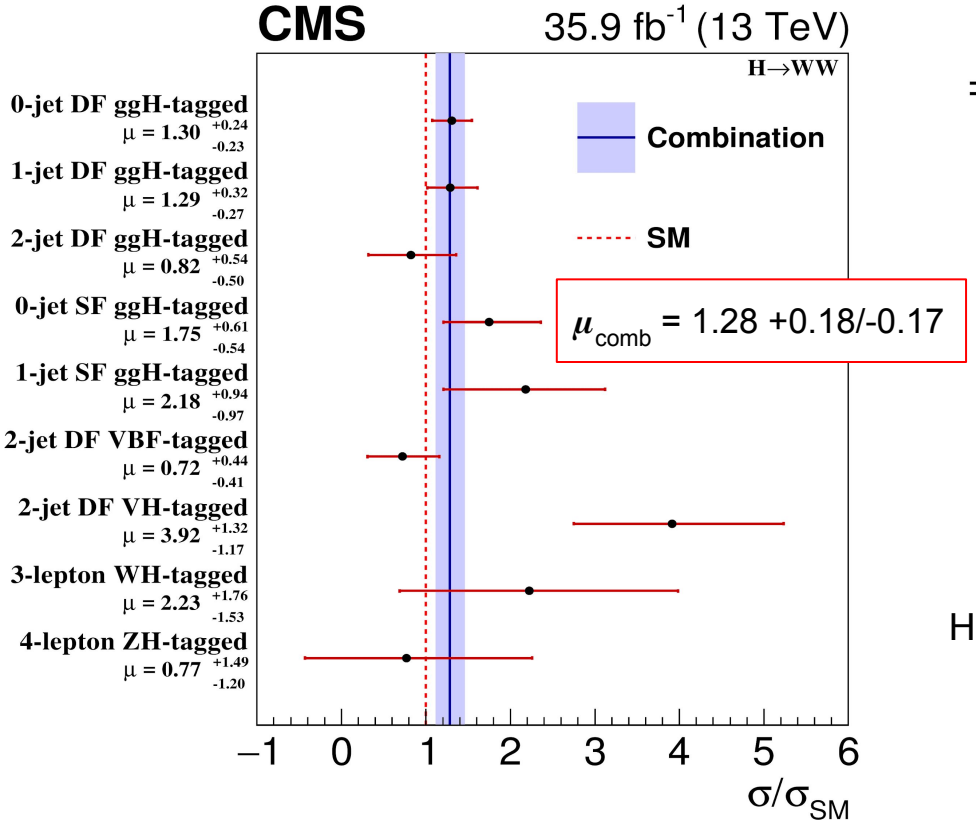
First measurement of HWW with Run 2 data

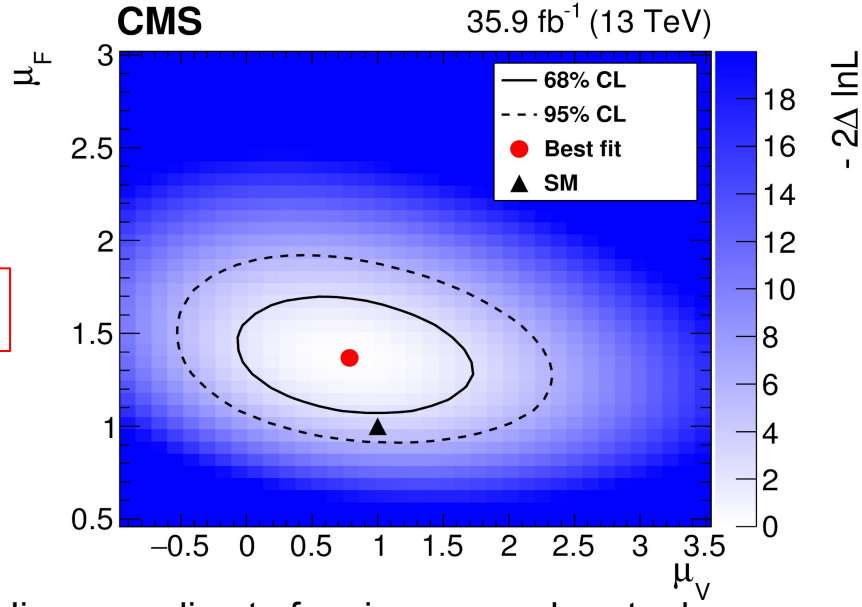
- ggH, VBF and VH production modes.
- Categories according to number of jets and same/different flavour leptons.





Higgs signal strength multipliers and couplings





Higgs coupling to fermions μ_F and vector bosons μ_V

$$\mu_F = 1.37^{+0.21}_{-0.20}\,,\ \mu_V = 0.78^{+0.60}_{-0.57}$$

Fiducial and differential measurements

- Differential cross sections are measured in bins of some observables.
- Fiducial measurements extrapolate to a phase space that matches the experimental selections.
  - Reduce model dependence avoiding the extrapolation to the full phase space.

- We can extract more information from fiducial differential cross sections:
  - Testing SM predictions.
  - Performing measurements in phase spaces more sensitive to BSM effects.

- We use the Higgs boson transverse momentum and the jet multiplicity as observables.
  - p_T^H is a particularly interesting choice, since the SM $d\sigma/dp_T^H$ is computed up to NNLO in QCD and it is known to be sensitive to possible deviations from the SM.
  - Jet multiplicity depends on the relative contribution of the production mechanisms and is useful to probe perturbative QCD radiation effects.

Analysis strategy

Gluon fusion channel is the one with the highest cross section and sensitivity.

Candidate events passing the ggH different-flavour selection are classified according to p_T^H and N_{jet} :

- p_T^H is computed using the sum of the p_T of the two leptons and the p_T^{miss} .
- For N_{jet} , only jets with $p_T > 30$ GeV and $|\eta| < 4.7$ are considered.

Two variables, dilepton invariant mass m_{ll} and the transverse mass of the Higgs boson m_T^H , have strong discrimination power against the background processes. The transverse mass is defined as:

$$m_{\mathrm{T}}^{\mathrm{H}} = \sqrt{2 p_{\mathrm{T}}^{\ell\ell} p_{\mathrm{T}}^{\mathrm{miss}} \left[1 - \cos\Delta\phi(\vec{p}_{\mathrm{T}}^{\ell\ell}, \vec{p}_{\mathrm{T}}^{\mathrm{miss}})\right]},$$

- In each category the signal is extracted from fits to this two-dimensional distribution, using a combination of the background and signal templates.

Event selection

- General selection:
  - μ and e with opposite charge
  - p_{T1} >25 GeV, $|\eta_1|$ <2.5
  - p_{T2} >13 GeV, $|\eta_2|$ <2.5
  - Third lepton veto (p_{T3} <10 GeV)
  - p_{Tll} >30 GeV
  - Missing E_T >20 GeV

- ggH signal region:
  - m_{ll} >12 GeV
  - m_T^H >60 GeV, m_T^{l2} >30 GeV
  - no b-tagged jets with p_T >20 GeV

with $m_T^{l2} = \sqrt{2 p_T^{\ell_2} p_T^{\text{miss}} \left[1 - \cos\Delta\phi(\vec{p}_T^{\ell_2}, \vec{p}_T^{\text{miss}})\right]}$

The fiducial region is defined using the general selection + signal region criteria

Additionally, two control regions are defined to constrain background contributions:

- Top quark control region:
  - m_T^{l2} >30 GeV , m_{ll} >50 GeV
  - 1 b-tagged jet with p_T >30 GeV (p_T >20 GeV in the 0-jet categories)

- DY→$\tau\tau$ control region:
  - m_T^H <60 GeV , 40 GeV <m_{ll} <80 GeV
  - no b-tagged jets with p_T >20 GeV

Analysis categories

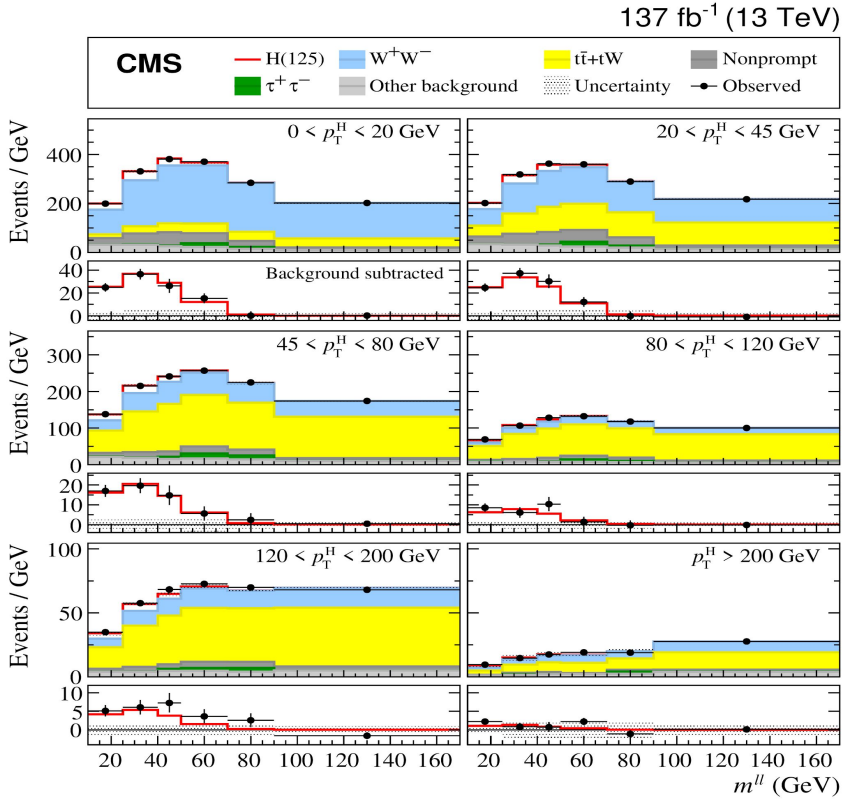
To maximize the sensitivity to the signal, events in each bin of the basis observables p_T^H and N_{jet} are further categorized by the properties of the two lepton candidates:

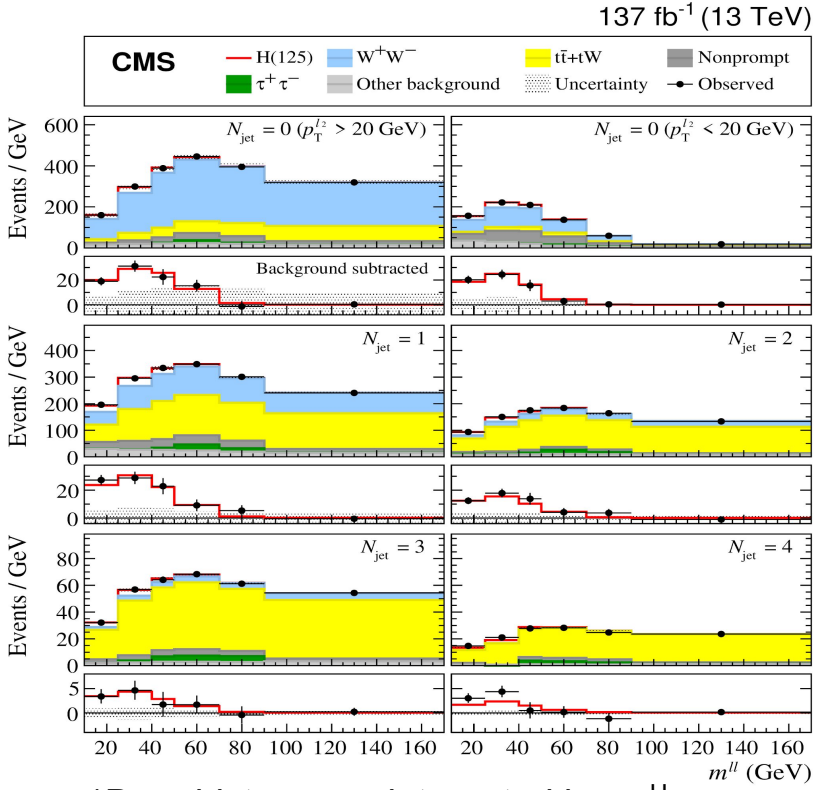
- $e\mu$ / μe , depending on the leading lepton → to further isolate nonprompt lepton background
- $p_{T2} \geq 20$ GeV / $p_{T2} < 20$ GeV → low p_{T2} region has higher signal/background ratio

Bins with lower number of events are not split into 4, but into 3 (no flavour division at $p_{T2} \geq 20$ GeV), 2 (no division in flavour) or 1 (no splitting) instead:

p_T^H (GeV)	0-20	20-45	45-80	80-120	120-200	>200
categories	4	4	4	3	2	2
Njet	0	1	2	3	≥4	
categories	4	4	2	1	1	

Signal region m_{ll} distributions





1D m_{ll} histograms integrated in m_T^H

Unfolding and regularization

- Signal events from one generator-level bin (i) contribute to multiple reconstructed-level bins in the m_{ll} : m_T^H templates, which are all scaled together by the same signal strength multiplier μ_i .

- By performing one simultaneous maximum likelihood fit over all reconstructed-level bin histograms, signal strength modifiers of the generator-level observable bins can be determined.

- The unfolding procedure can be highly sensitive to statistical fluctuations in the observed distributions, and for the p_T^H measurement there are large migrations due to poor p_T^{miss} resolution.
  - To mitigate this effect, a regularization procedure is introduced to obtain the final result.
  - This regularization term acts as a smoothing constraint, reducing unphysical anticorrelations.

Uncertainties

Sources of experimental uncertainties:

- Integrated luminosity
- Trigger efficiency
- Lepton reconstruction and identification efficiencies (1-2%)
- Lepton momentum scale
- Jet energy scale
- E_T^{miss} scale (1-10%)
- b-tagging
- Pile-up reweighting

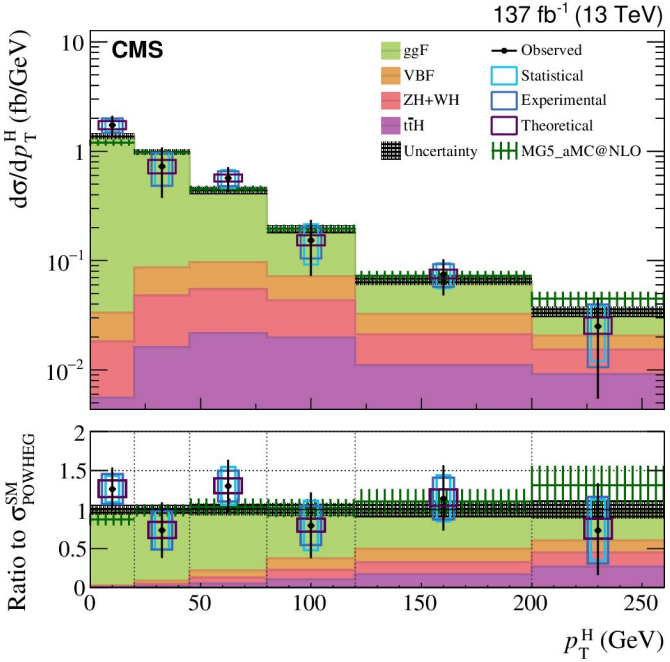
Sources of theoretical uncertainties:

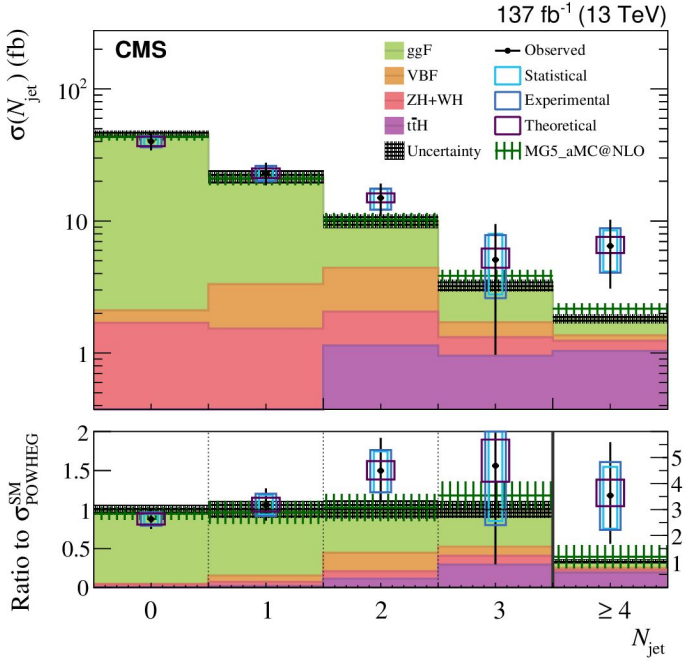
- PDF
- Renormalization and factorization scale
- Parton shower modeling
- Underlying event modeling

Sources of uncertainties in the background modeling:

- Control region/ signal region for top quark and DY
- Top quark p_T reweighting
- Single top / ttbar cross section ratio
- WW NNLL resummation
- Nonprompt lepton background estimation (5–10%)
- WZ and Wγ^* scale factors

Fiducial differential cross section results





Good agreement with theoretical predictions

1.4 σ is the largest deviation, in the ≥4 jet bin

$$\mu^{\text{fid}} = 1.05 \pm 0.12 \left(\pm 0.05\,(\text{stat}) \pm 0.07\,(\text{exp}) \pm 0.01\,(\text{signal}) \pm 0.07\,(\text{bkg}) \pm 0.03\,(\text{lumi}) \right)$$

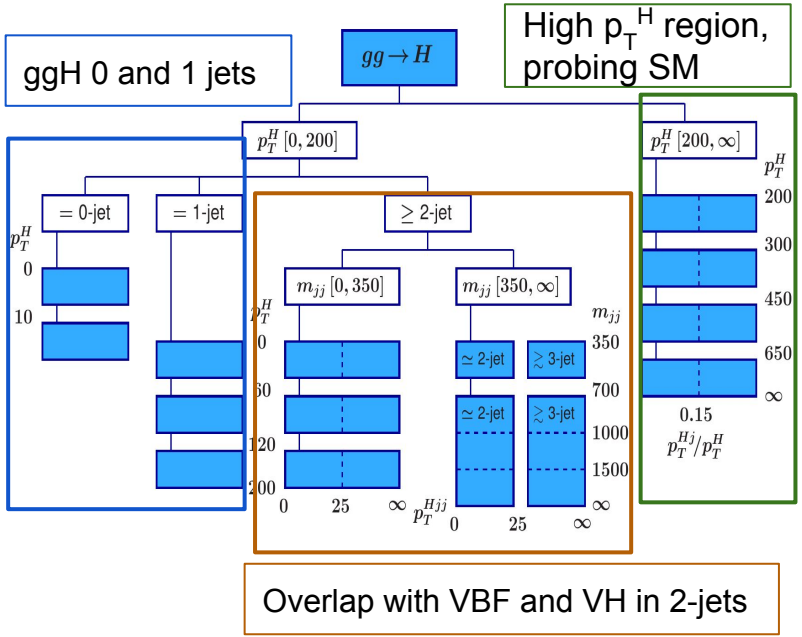
$$\sigma^{\text{fid}} = 86.5 \pm 9.5\,\text{fb}. \qquad \sigma^{\text{SM}} = 82.5 \pm 4.2\,\text{fb}$$

Next results in SM H→WW analyses in CMS

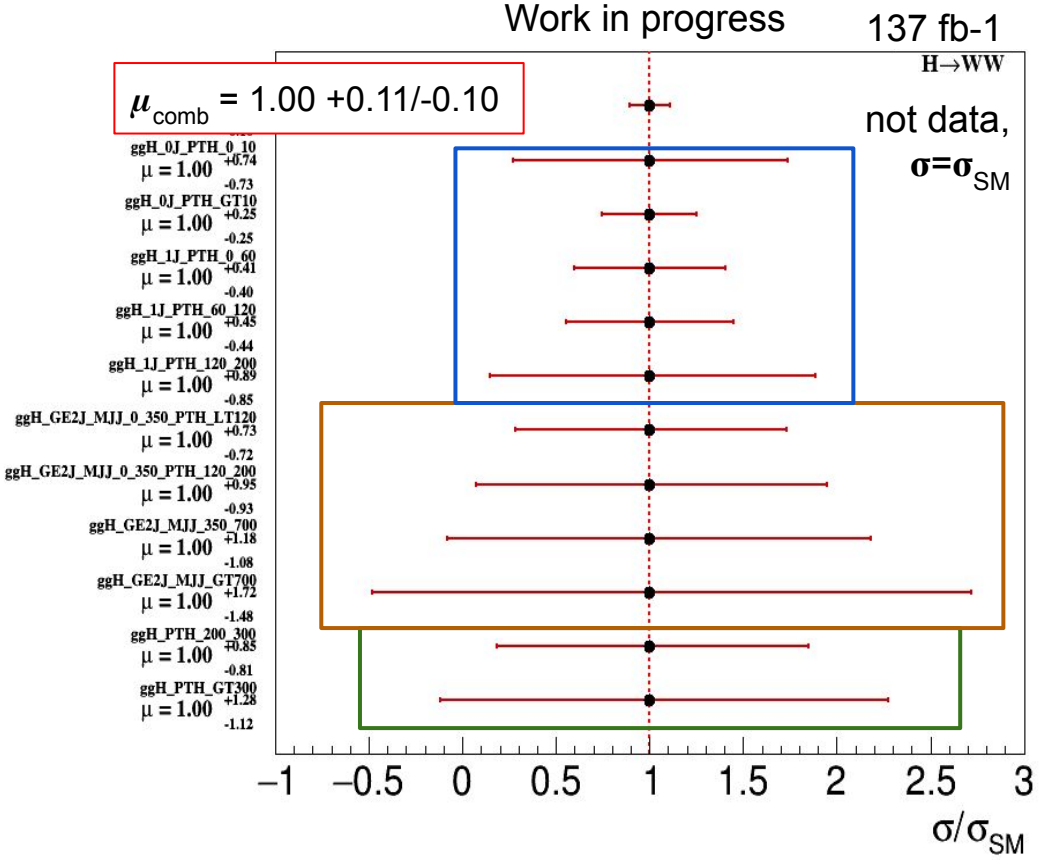
- The H→WW→lνlν group is currently in the latest stages of analyzing the full Run 2 data.

- One new result will be the measurement of stage 1.2 Simplified Template Cross Sections (STXS).
  - Measurement performed in phase space regions, with same definition in CMS and ATLAS.
  - Complementary measurements with respect to differential analyses.
  - Reduce theory uncertainties while maximizing the sensitivity of the analysis.
  - Binning schemes for each production mode, defined in stages of increasing granularity.
  - Allows combination between analyses using different decay modes (e.g. ZZ, γγ, $\tau\tau$, WW).

STXS expected signal strength uncertainties (ggH)



- Uncertainties of the order of 10% for the combination of the bins.
- For regions with uncertainties over 100%, (low statistics) bins can be merged.



Summary and conclusions

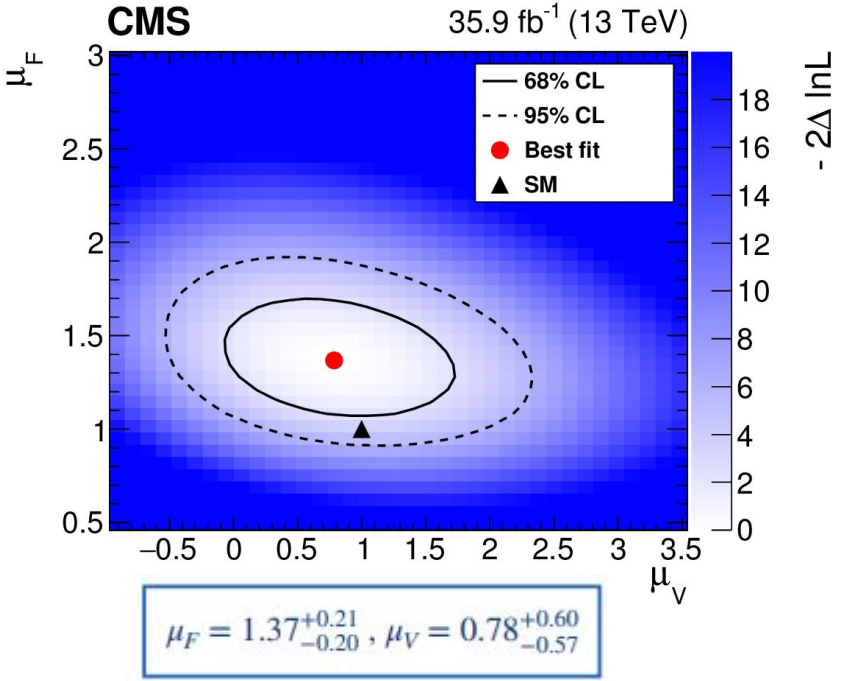
- After the first measurement at 13 TeV of Higgs signal strength multipliers and couplings, the SM H →WW group of CMS has produced a fiducial differential measurement using 137 fb^{-1} of data.
  - Differential measurement with respect to p_T^H and N_{jet}.
  - Events selected targeting the gluon fusion production mode.
  - Final state with two different-flavour leptons and missing energy.
  - Cross section obtained in agreement with the SM theoretical prediction.
- Work in progress: full Run 2 measurements of the ggH, VBF and VH production modes.
  - New results will include a measurement with a divided phase space, using the Simplified Template Cross Section framework.

Thank you for your attention!

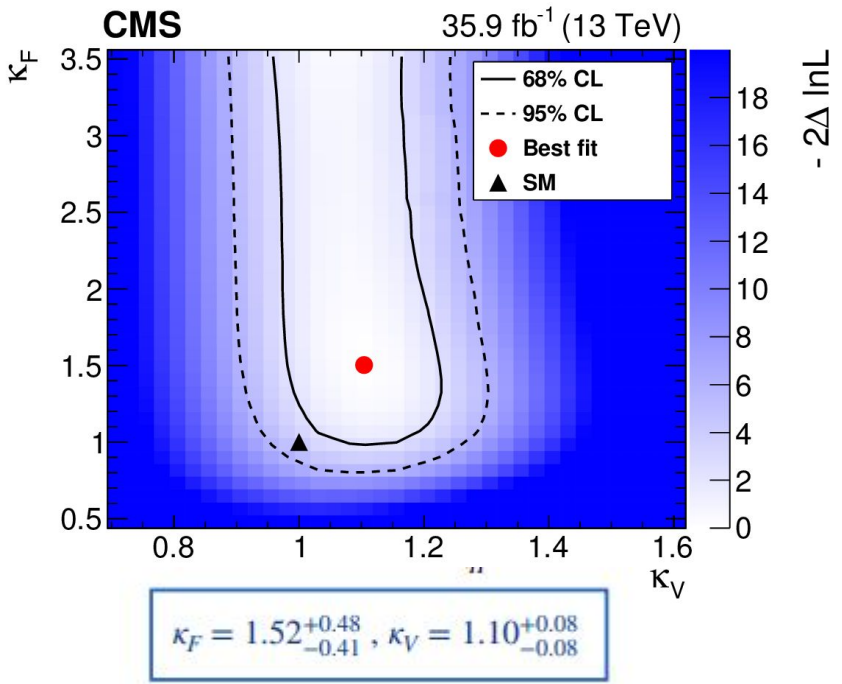
Backup

Higgs couplings in first Run 2 measurement of HWW

kappa framework: $\sigma\mathcal{B}(X \to H \to WW) = \kappa_i^2 \frac{\kappa_V^2}{\kappa_H} \sigma_{SM}\mathcal{B}_{SM}(X \to H \to WW)$



$\mu_F = 1.37^{+0.21}_{-0.20}$, $\mu_V = 0.78^{+0.60}_{-0.57}$



$\kappa_F = 1.52^{+0.48}_{-0.41}$, $\kappa_V = 1.10^{+0.08}_{-0.08}$

Signal extraction

Higgs differential production cross sections are inferred from signal strength modifiers $\mu_i = \sigma_i^{\text{obs}} / \sigma_i^{\text{SM}}$ extracted from a simultaneous maximum likelihood fit to all bins and categories.

This is the likelihood function:

$$\mathcal{L}(\boldsymbol{\mu};\boldsymbol{\theta}) = \prod_j \text{Poisson}\left(n_j; s_j(\boldsymbol{\mu};\boldsymbol{\theta}) + b_j(\boldsymbol{\theta})\right) \cdot \mathcal{N}(\boldsymbol{\theta}) \cdot \mathcal{K}(\boldsymbol{\mu})$$

n data events in bin *j* (m_{ll} and m_T^H bin)

constraints on the systematic uncertainties, taken as nuisance parameters $\boldsymbol{\theta}$

regularization factor that reduces large fluctuations among neighboring bins applied **only in p_T^H measurement**

$$s_j(\boldsymbol{\mu};\boldsymbol{\theta}) = \sum_i \left[R_{ji}(\boldsymbol{\theta}) \mu_i L_j \cdot (\sigma_i^{\text{SM}} + \sigma_i^{\text{SM-out}}) \right]$$

response matrix

luminosity

fiducial cross section

nonfiducial cross section

Regularization procedure

- The unfolding procedure can be highly sensitive to statistical fluctuations in the observed distributions, and for the p_T^H measurement there are large migrations due to poor p_T^{miss} resolution.
- To mitigate this effect, a regularization procedure is introduced to obtain the final result.

$$K(\mu) = \prod_{i=2}^{N-1} \exp\left(\frac{-[(\mu_{i+1} - \mu_i) - (\mu_i - \mu_{i-1})]^2}{2\delta^2} \right)$$

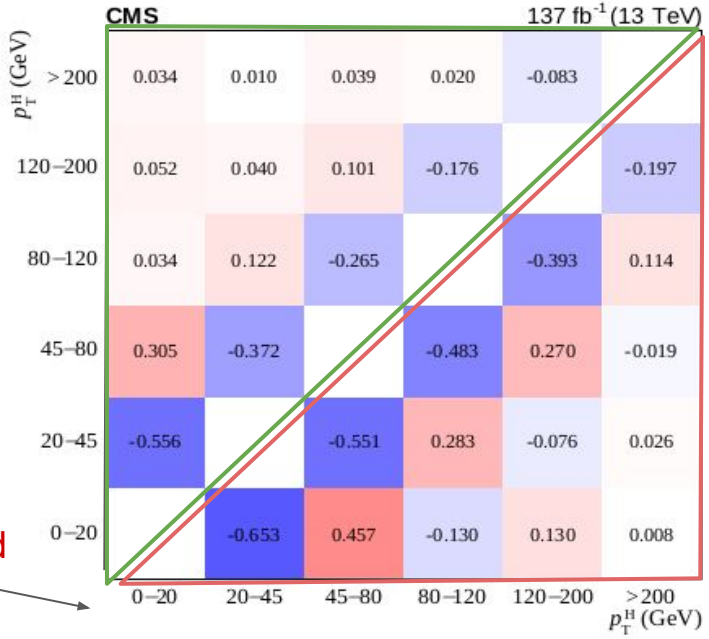
➢ This regularization term acts as a smoothing constraint, reducing unphysical anticorrelations.

➢ δ is optimized by minimizing the mean of the global correlation coefficient.

Correlation among the signal strength modifiers for p_T^H bins

Regularized

Unregularized



Correlation matrix among Njet bins

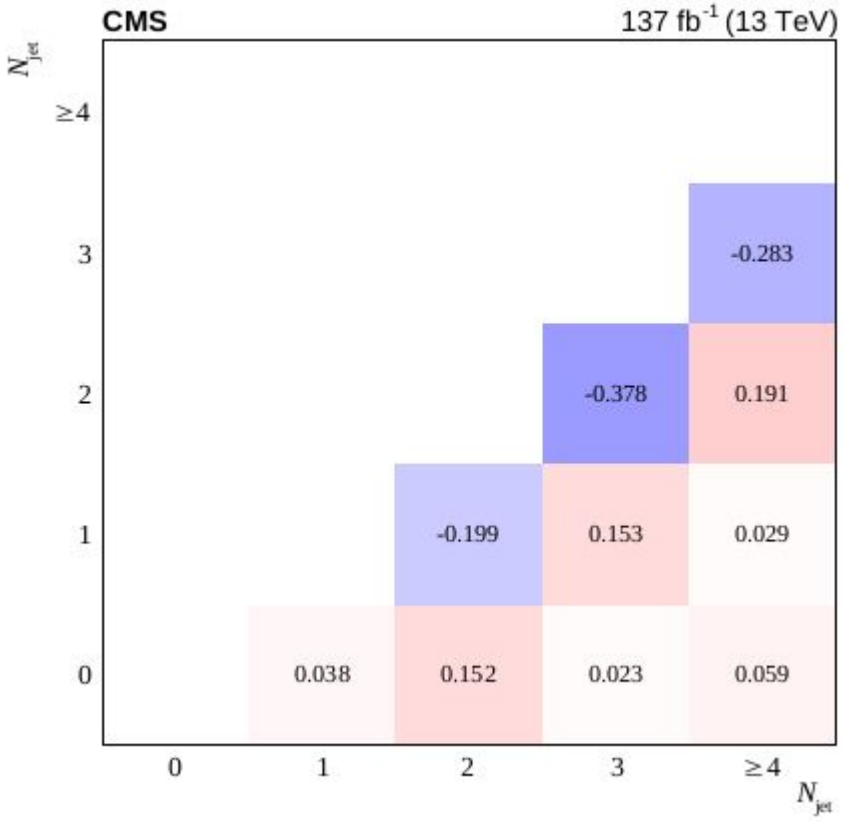


Table of yields

Table 5: Signal and background post-fit (pre-fit) yields in the $\mathcal{RL}$ p_T^H bins.

Process	$\mathcal{RL}$ p_T^H bin					
	[0–20]	[20–45]	[45–80]	[80–120]	[120–200]	> 200
H(125)	1489 ± 81 (1356)	1386 ± 80 (1402)	835 ± 52 (792)	320 ± 36 (344)	217 ± 33 (222)	54 ± 17 (75)
$\tau^+\tau^-$	537 ± 49 (372)	675 ± 43 (585)	684 ± 61 (482)	316 ± 42 (195)	173 ± 24 (219)	104 ± 58 (83)
W^+W^-	26945 ± 213 (22840)	17421 ± 290 (18771)	7444 ± 269 (9048)	2759 ± 250 (3972)	2205 ± 155 (2816)	1037 ± 70 (1637)
$t\bar{t}$ + tW	5571 ± 65 (5492)	14700 ± 176 (14528)	18313 ± 239 (18188)	11482 ± 220 (11624)	6481 ± 137 (6488)	1659 ± 40 (1671)
Nonprompt	3709 ± 127 (5154)	4373 ± 128 (5909)	1822 ± 107 (3143)	1002 ± 80 (1239)	558 ± 52 (749)	197 ± 23 (279)
Other background	2770 ± 102 (2002)	3245 ± 137 (2186)	2160 ± 100 (1431)	1055 ± 64 (778)	737 ± 49 (519)	478 ± 33 (349)

Table 6: Signal and background post-fit (pre-fit) yields in the $\mathcal{RL}$ N_{jet} bins.

Process	$\mathcal{RL}$ N_{jet} bin				
	0	1	2	3	≥ 4
H(125)	2186 ± 92 (2447)	1254 ± 60 (1165)	632 ± 66 (445)	178 ± 48 (109)	98 ± 26 (36)
$\tau^+\tau^-$	740 ± 41 (520)	944 ± 50 (822)	688 ± 99 (301)	255 ± 43 (135)	100 ± 50 (70)
W^+W^-	41058 ± 360 (38437)	13190 ± 252 (15176)	3402 ± 222 (4266)	698 ± 125 (966)	0 ± 0 (240)
$t\bar{t}$ + tW	11125 ± 144 (11870)	20891 ± 179 (21198)	15788 ± 214 (15381)	6853 ± 110 (6510)	3152 ± 52 (3031) top
Nonprompt	6649 ± 188 (8999)	3436 ± 149 (4457)	1066 ± 77 (1792)	480 ± 52 (685)	254 ± 30 (357)
Other background	4513 ± 165 (3394)	3189 ± 139 (2342)	1424 ± 89 (1043)	449 ± 32 (362)	149 ± 12 (124)

Fiducial differential cross section results

p_T^H (GeV)	σ^{SM} (fb)	μ	Regularized μ Value	stat	exp	signal	bkg	lumi	Bias	σ^{obs} (fb)
0–20	27.45	1.37 ± 0.30	1.26 ± 0.27	± 0.17	± 0.19	± 0.01	± 0.10	± 0.03	$+0.00$	34.6 ± 7.5
20–45	24.76	0.52 ± 0.42	0.73 ± 0.36	± 0.24	± 0.25	± 0.01	± 0.10	± 0.03	-0.12	18.2 ± 8.9
45–80	15.28	1.55 ± 0.41	1.30 ± 0.33	± 0.24	± 0.20	± 0.03	± 0.09	± 0.03	-0.03	19.9 ± 5.2
80–120	7.72	0.49 ± 0.52	0.79 ± 0.42	± 0.32	± 0.25	± 0.02	± 0.08	± 0.03	-0.16	6.1 ± 3.3
120–200	5.26	$1.34^{+0.51}_{-0.48}$	1.14 ± 0.41	± 0.29	± 0.27	± 0.04	± 0.08	± 0.03	$+0.11$	6.0 ± 2.2
>200	2.05	$0.64^{+0.63}_{-0.60}$	$0.73^{+0.61}_{-0.57}$	± 0.38	± 0.42	$^{+0.09}_{-0.03}$	± 0.10	± 0.03	$+0.19$	1.5 ± 1.2

N_{jet}	σ^{SM} (fb)	μ Value	stat	exp	signal	bkg	lumi	σ^{obs} (fb)
0	45.70	0.88 ± 0.13	± 0.06	± 0.08	± 0.01	± 0.07	± 0.03	40.1 ± 6.0
1	21.74	1.06 ± 0.20	± 0.12	± 0.14	± 0.01	± 0.08	± 0.03	23.0 ± 4.6
2	9.99	1.50 ± 0.40	$^{+0.25}_{-0.28}$	± 0.28	± 0.04	± 0.11	± 0.03	15.0 ± 4.2
3	3.26	$1.56^{+1.35}_{-1.26}$	$^{+0.89}_{-0.71}$	$^{+0.84}_{-0.76}$	$^{+0.17}_{-0.07}$	$^{+0.29}_{-0.19}$	$^{+0.07}_{-0.04}$	$5.1^{+4.4}_{-4.1}$
≥ 4	1.83	$3.54^{+2.05}_{-1.86}$	$^{+1.10}_{-1.28}$	$^{+1.28}_{-1.32}$	$^{+0.40}_{-0.20}$	$^{+0.38}_{-0.34}$	$^{+0.10}_{-0.07}$	$6.5^{+3.8}_{-3.4}$