Higgs boson measurements in final states with photons at CMS

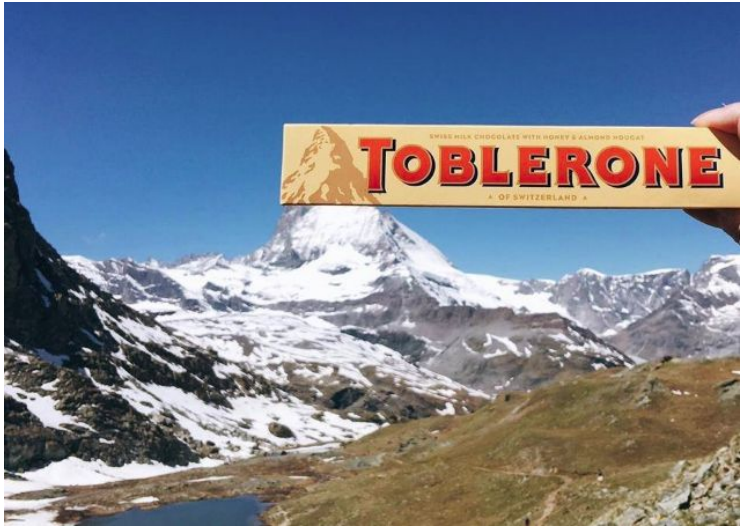
Hualin Mei (UC Santa Barbara)
on behalf of the CMS collaboration

2020/07/30

ICHEP 2020, Prague

H → γγ decay channel

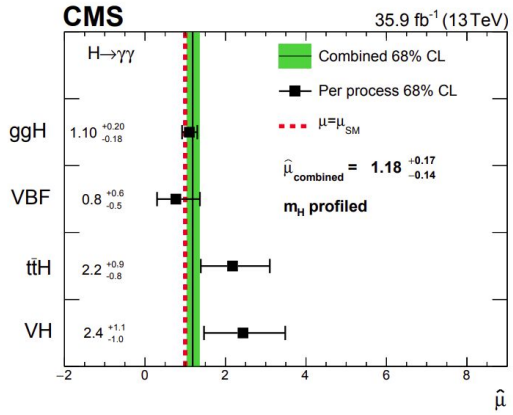
- **Small branching fraction:** ~0.2%
- Clean final state, Higgs decay products (diphoton) **fully reconstructable**
- **Excellent diphoton invariant mass resolution** (1-2%)
- Sensitive to all the principal Higgs boson production modes
  - With data collected by the CMS detector during LHC Run 2 (137fb^{-1})

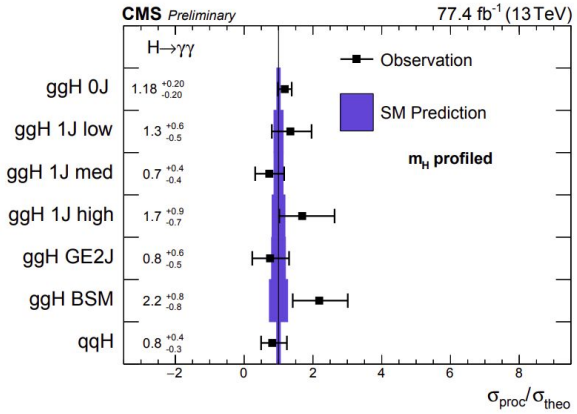


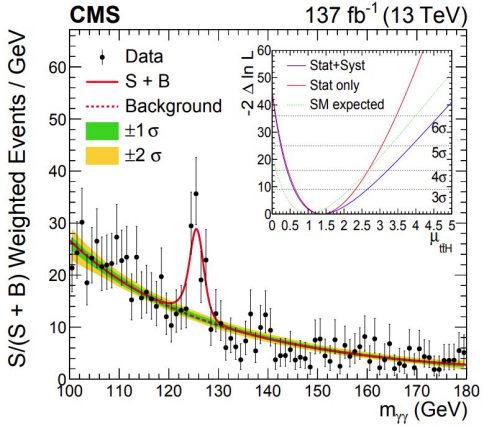
©

One of the most important channels for precision measurements of Higgs boson properties

H → γγ cross section measurements in Run 2







JHEP11(2018)185

Signal strength, STXS stage 0, coupling modifier (2016 data)

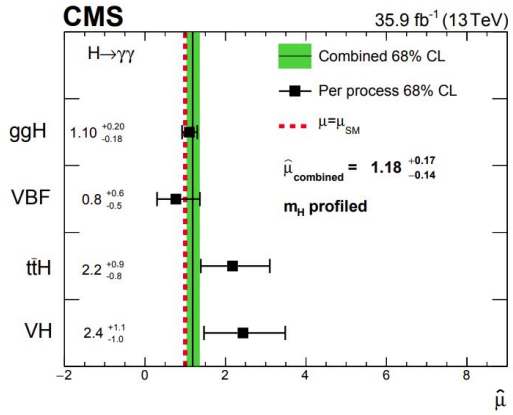
CMS-HIG-18-029

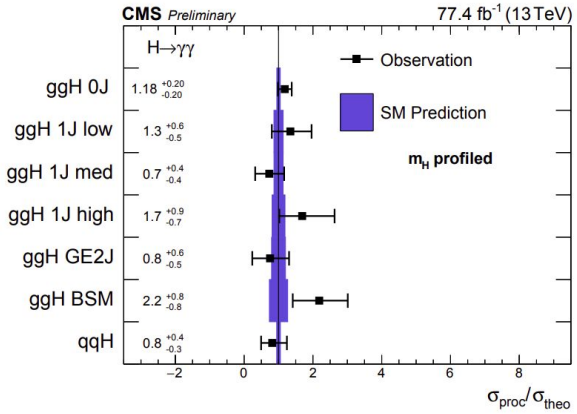
ggH + VBF STXS stage 1.0 (2016 + 2017 data)

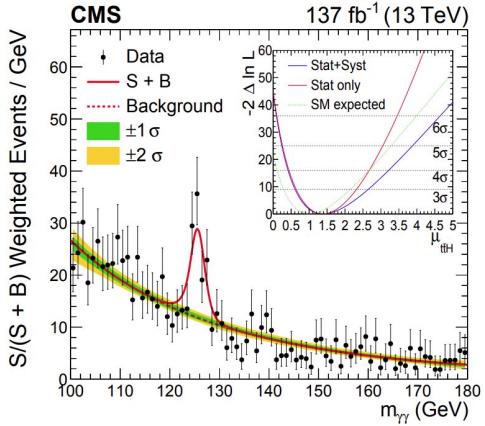
arXiv:2003.10866

ttH production + CP (2016 + 2017 + 2018 data)

H → γγ cross section measurements in Run 2







JHEP11(2018)185

Signal strength, STXS stage 0, coupling modifier (2016 data)

CMS-HIG-18-029

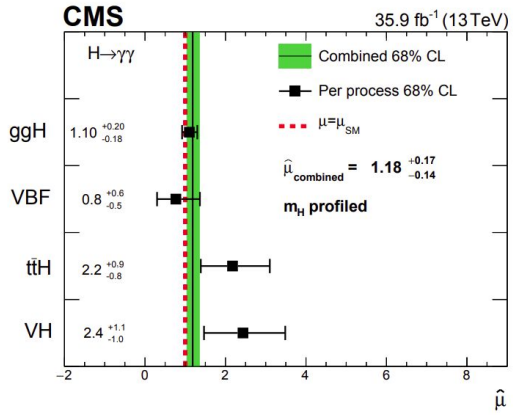
ggH + VBF STXS stage 1.0 (2016 + 2017 data)

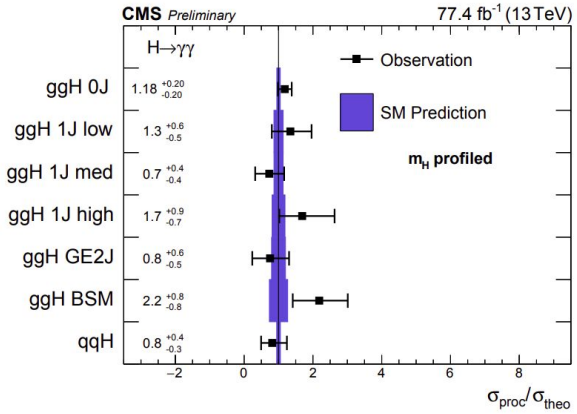
arXiv:2003.10866

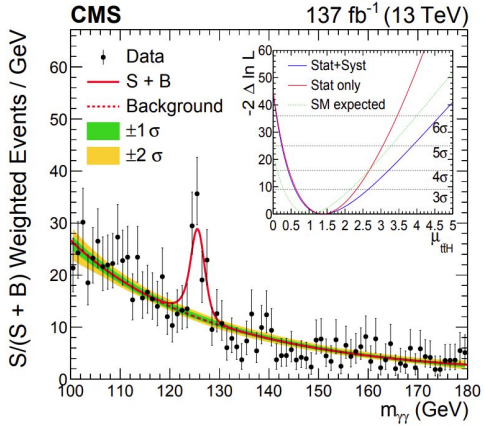
ttH production + CP (2016 + 2017 + 2018 data)

More details in talks by Sergio Sanchez Cruz and Savvas Kyriacou later today

H → γγ cross section measurements in Run 2







JHEP11(2018)185

Signal strength, STXS stage 0, coupling modifier (2016 data)

CMS-HIG-18-029

ggH + VBF STXS stage 1.0 (2016 + 2017 data)

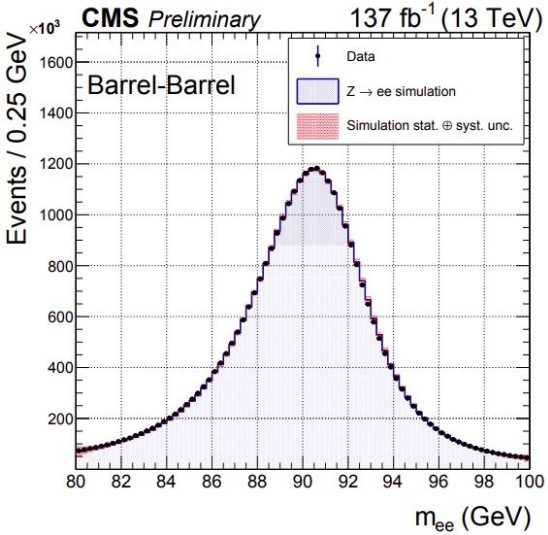
arXiv:2003.10866

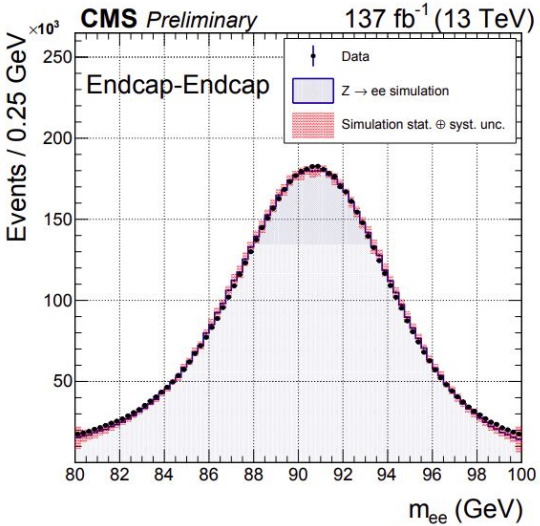
ttH production + CP (2016 + 2017 + 2018 data)

This talk will present a more comprehensive measurement of the Higgs boson properties with H → γγ channel (signal strength, STXS stage 1.2, coupling modifier) with the full Run 2 dataset (2016 + 2017 + 2018)

Photon energy

- Correct photon energy with a multivariate regression technique to account for reconstruction effect
  - Also predicts its uncertainty

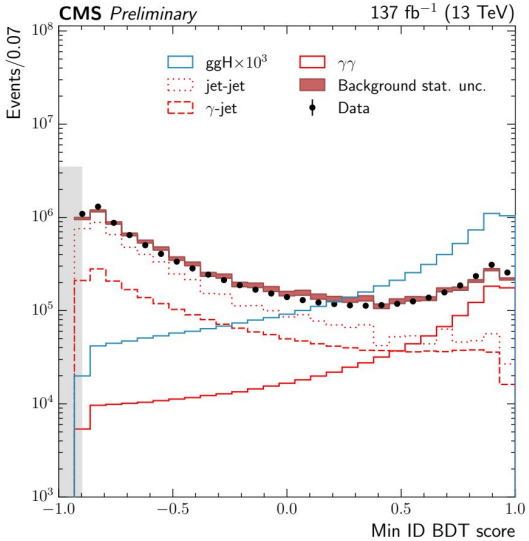


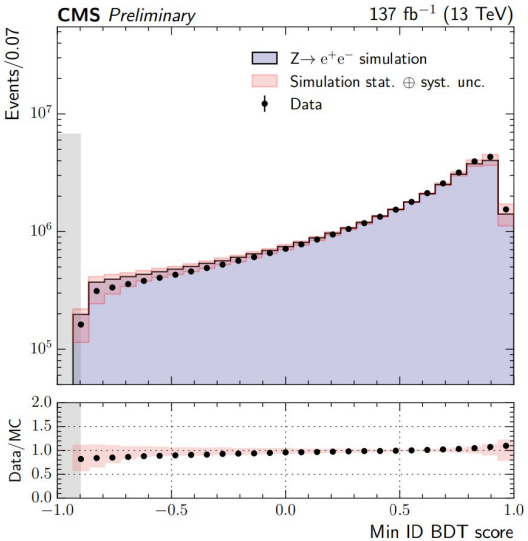


- Use Z → ee events to derive additional corrections (with electrons reconstructed as photons)
  - Correct drift of energy scale in data over time as a function of LHC fill
  - Correct energy scale (resolution) in data (simulation)
    - By aligning dielectron mass distribution between data and simulation

Photon identification

- Use a ID BDT to separate genuine photons from jets mimicking a photon
- Trained on γ+jet simulation, input variables include:
  - Shower shape/isolation variables, photon energy and η, variables sensitive to pileup



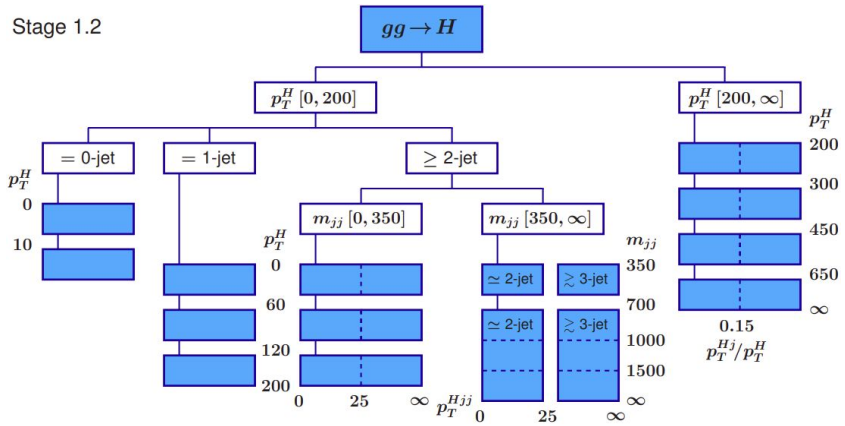


- Correct shower shape and isolation variable's distributions in simulation to agree with data
  - Using a [chained quantile regression method] with electrons from Z → ee
  - Ensure correlations between the corrected variables in simulation are closer to those in data

STXS & Analysis strategy

- STXS framework provides a coherent approach to perform precision Higgs boson measurements

Measure cross sections in mutually exclusive regions of phase space (bins), which are defined based upon Higgs boson production modes

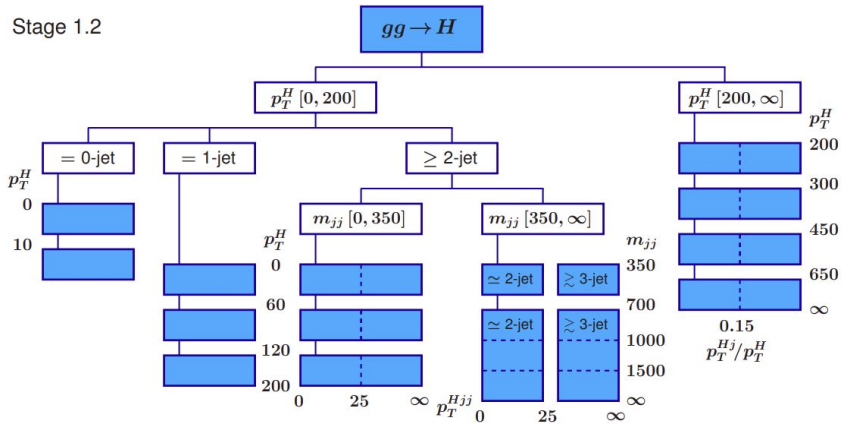


An example of bin definition for ggH at stage 1.2

STXS & Analysis strategy

- STXS framework provides a coherent approach to perform precision Higgs boson measurements

Measure cross sections in mutually exclusive regions of phase space (bins), which are defined based upon Higgs boson production modes



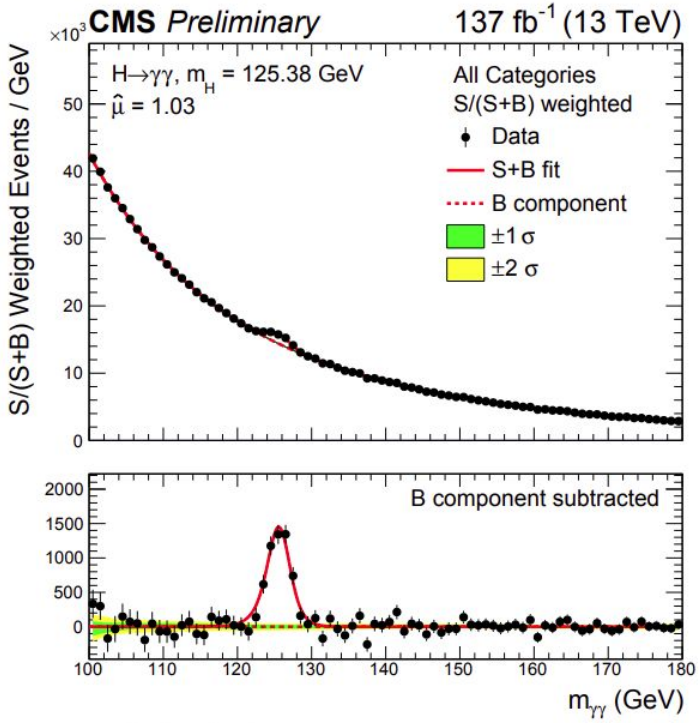
An example of bin definition for ggH at stage 1.2

Analysis strategy

- Build analysis categories targeting as many different stage 1.2 STXS bins as possible
- Allows measurements to be performed across all the major Higgs production modes:
  - ggH, VBF, VH, tH, ttH
- Also provides sensitivity to measure signal strength modifiers (μ) and coupling modifiers (κ) in the κ-framework
- Extract cross section, μ and κ by performing a simultaneous fit to the $m_{\gamma\gamma}$ distribution in each category

STXS & Analysis strategy

- STXS framework provides a coherent approach to perform precision Higgs boson measurements



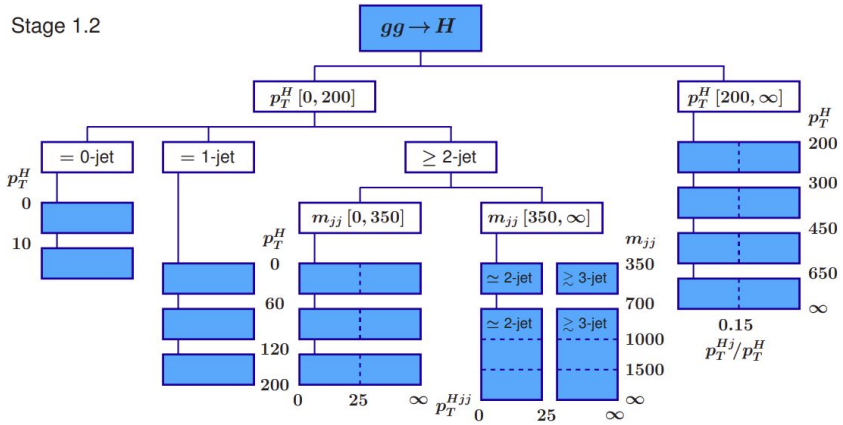
Analysis strategy

- Build analysis categories targeting as many different stage 1.2 STXS bins as possible
- Allows measurements to be performed across all the major Higgs production modes:
  - ggH, VBF, VH, tH, ttH
- Also provides sensitivity to measure signal strength modifiers (μ) and coupling modifiers (κ) in the κ-framework
- Extract cross section, μ and κ by performing a simultaneous fit to the $m_{\gamma\gamma}$ distribution in each category

* All measurements are performed with m_H fixed at 125.38 GeV, current best Higgs mass measurement (Phys. Lett. B 805 (2020) 135425)

Event categorization

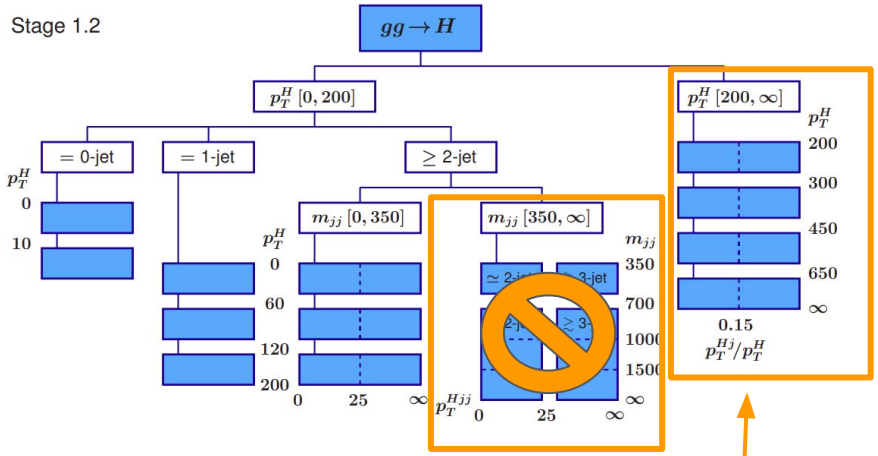
An example of bin definition for ggH at stage 1.2



- Phase spaces that target different Higgs production mode (ggH, VBF, VH, tH, ttH) are further divided to explore various STXS bins

Event categorization

An example of bin definition for ggH at stage 1.2



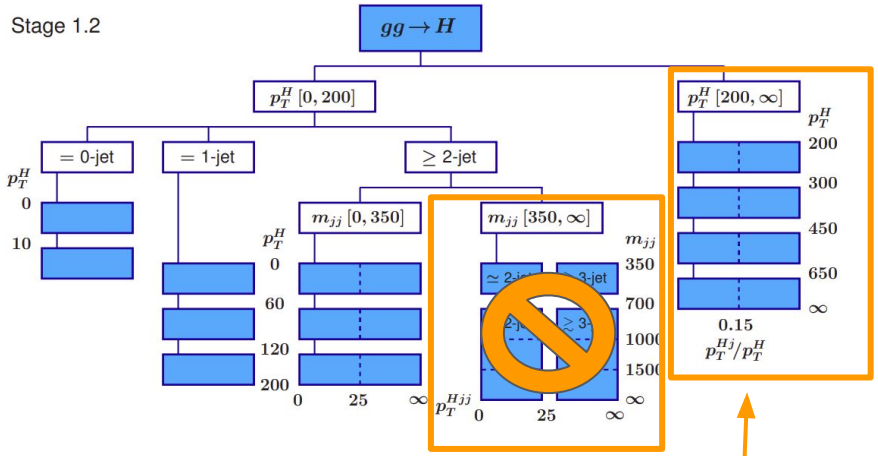
These VBF-like bins are categorized with VBF events

These bins are (partially) merged in final measurement

- Phase spaces that target different Higgs production mode (ggH, VBF, VH, tH, ttH) are further divided to explore various STXS bins
- Certain STXS bins are merged in the analysis due to low statistics with current data

Event categorization

An example of bin definition for ggH at stage 1.2



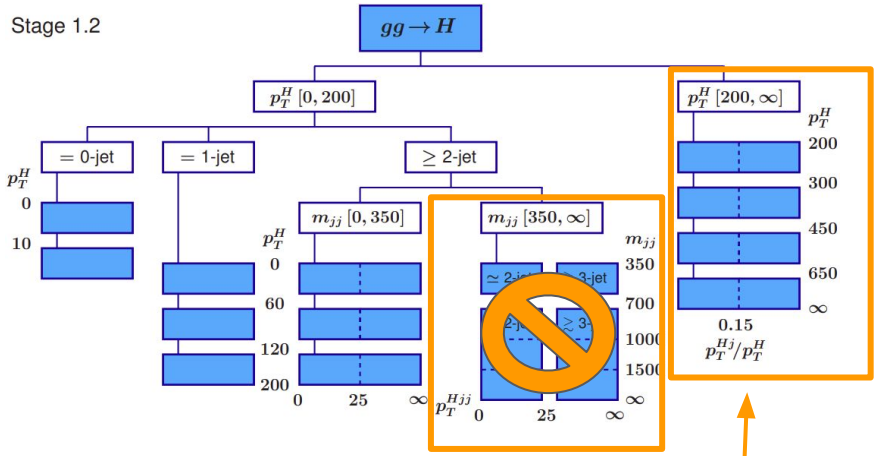
These VBF-like bins are categorized with VBF events

These bins are (partially) merged in final measurement

- Phase spaces that target different Higgs production mode (ggH, VBF, VH, tH, ttH) are further divided to explore various STXS bins
- Certain STXS bins are merged in the analysis due to low statistics with current data
- Each considered STXS region is divided into analysis categories using dedicated BDTs or DNNs
  - To reject background and maximize expected sensitivity

Event categorization

An example of bin definition for ggH at stage 1.2

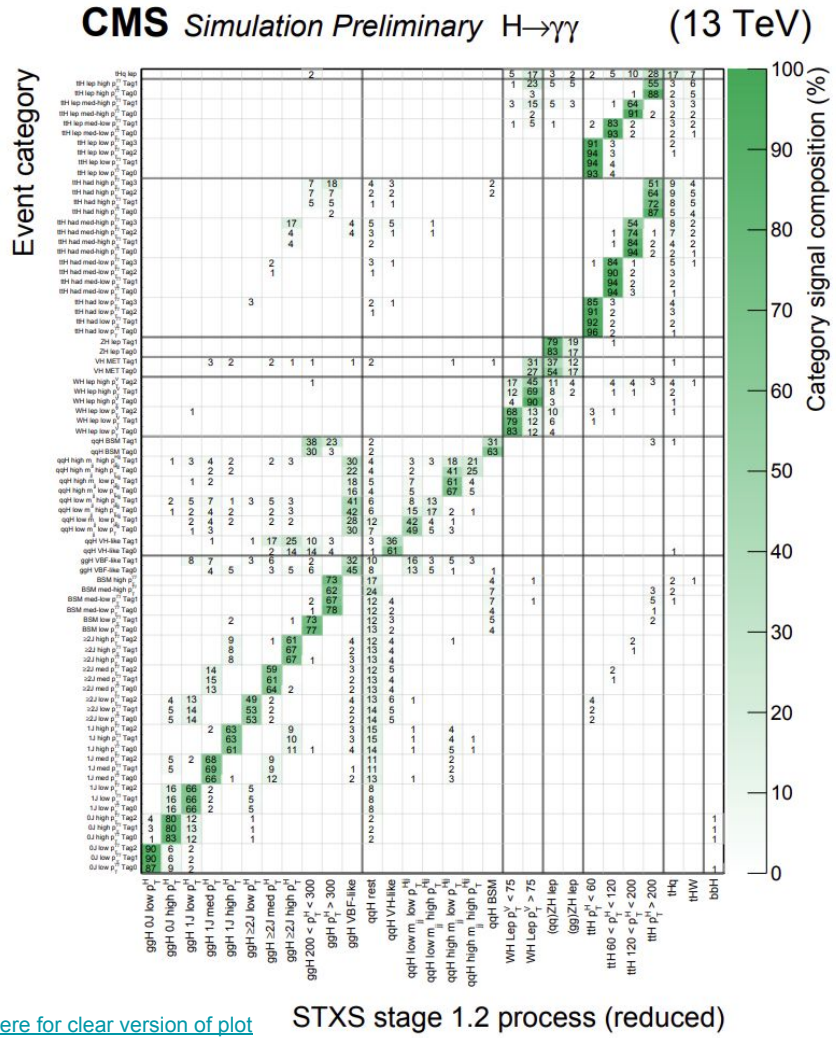


These VBF-like bins are categorized with VBF events

These bins are (partially) merged in final measurement

- Phase spaces that target different Higgs production mode (ggH, VBF, VH, tH, ttH) are further divided to explore various STXS bins
- Certain STXS bins are merged in the analysis due to low statistics with current data
- Each considered STXS region is divided into analysis categories using dedicated BDTs or DNNs
  - To reject background and maximize expected sensitivity

81 analysis categories are considered in this analysis!

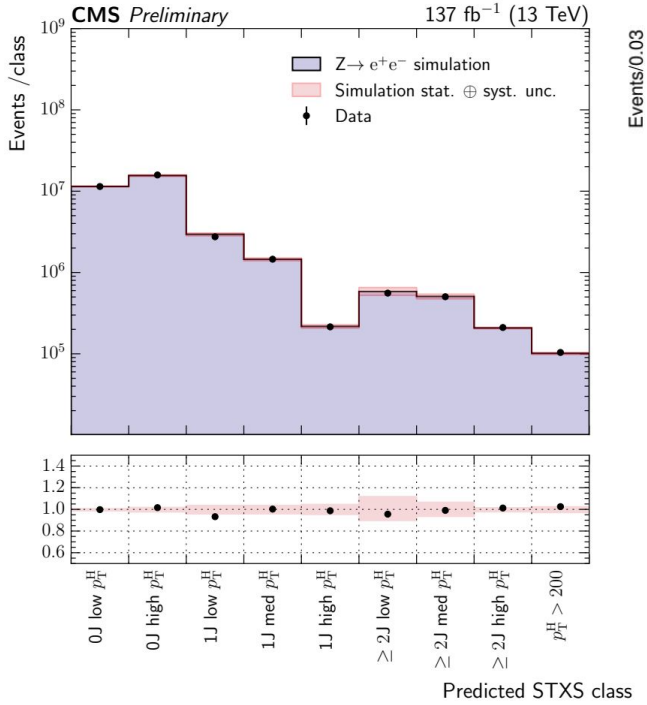


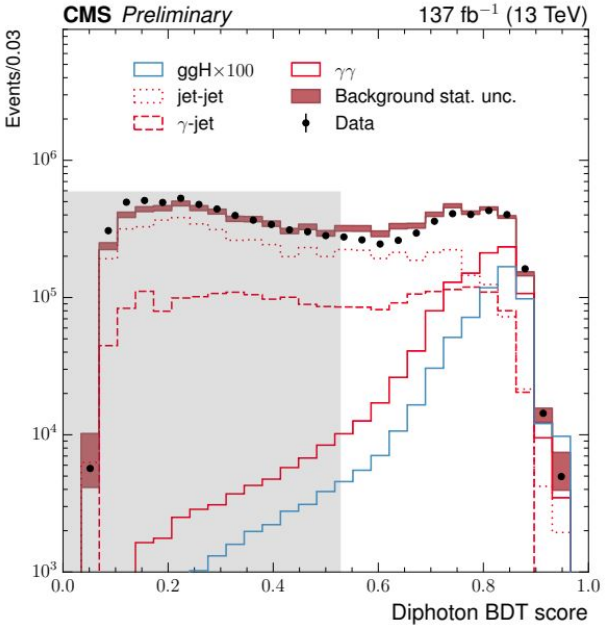
Click here for clear version of plot

- Phase spaces that target different Higgs production mode (ggH, VBF, VH, tH, ttH) are further divided to explore various STXS bins
- Certain STXS bins are merged in the analysis due to low statistics with current data
- Each considered STXS region is divided into analysis categories using dedicated BDTs or DNNs
  - To reject background and maximize expected sensitivity

81 analysis categories are considered in this analysis!

ggH event categorisation as an example

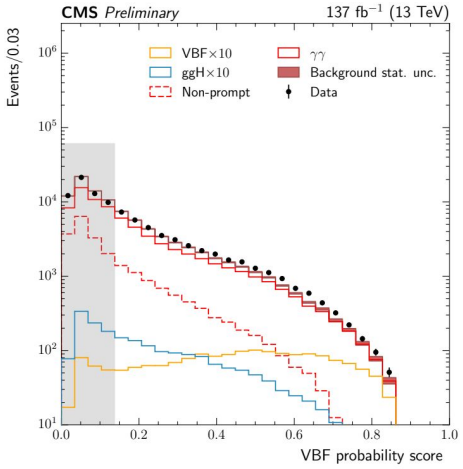




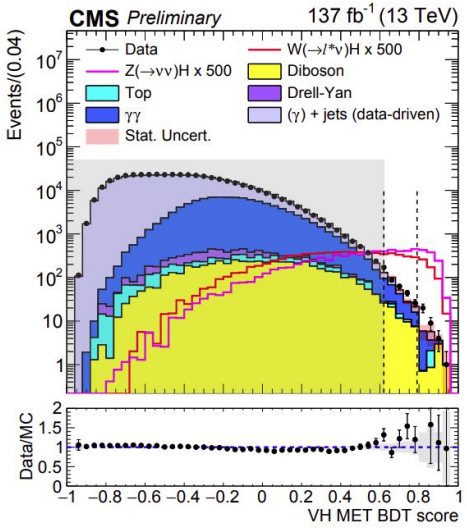
Use a multi-class BDT (9 classes)

- Predicts the probability that an event belongs to a given ggH STXS bin
- 5% improvement on accuracy of bin assignment
- Each class is divided into subcategories (different S/B)
  - With a diphoton BDT
  - Reduces background from SM diphoton production
  - Use photon kinematics in input variables

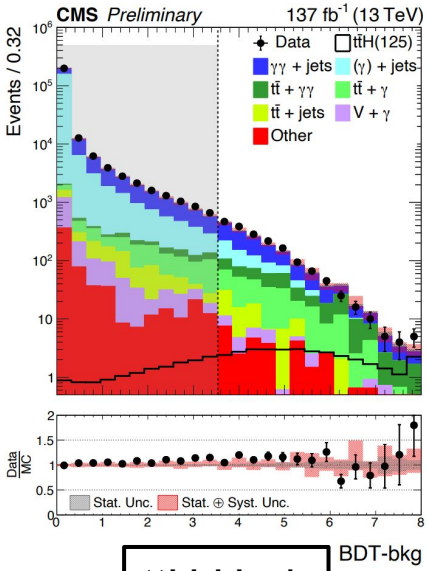
Other production modes (part of) ...



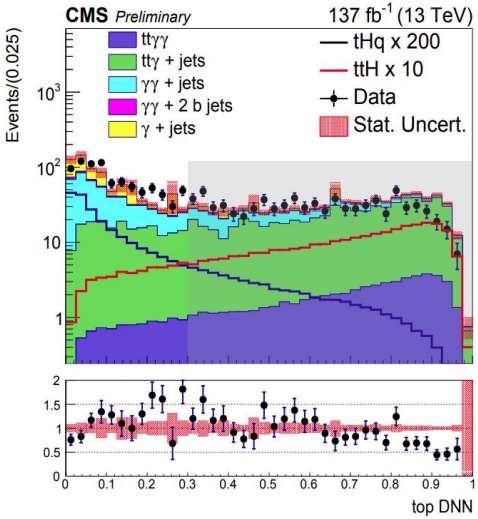
VBF



VH MET



ttH Had

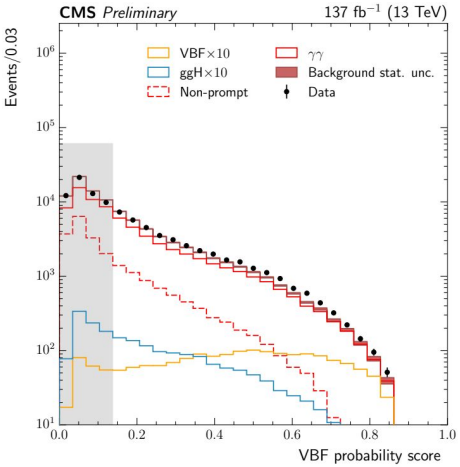


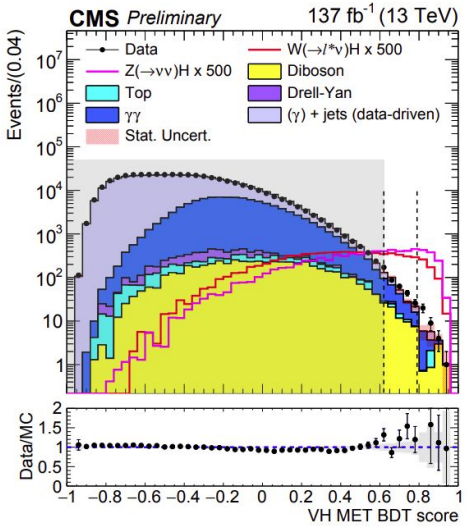
tHq Lep

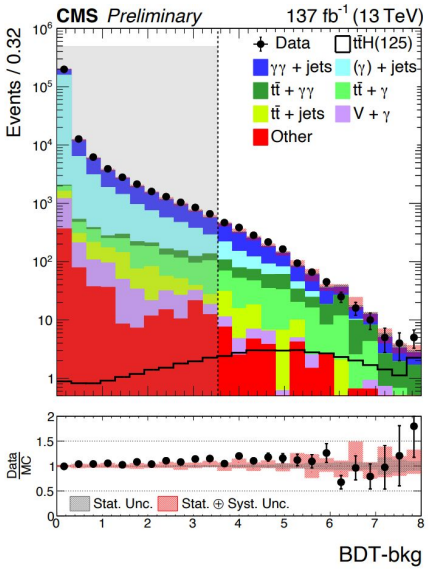
To avoid ambiguity (one event passes selection criteria for > 1 analysis category)

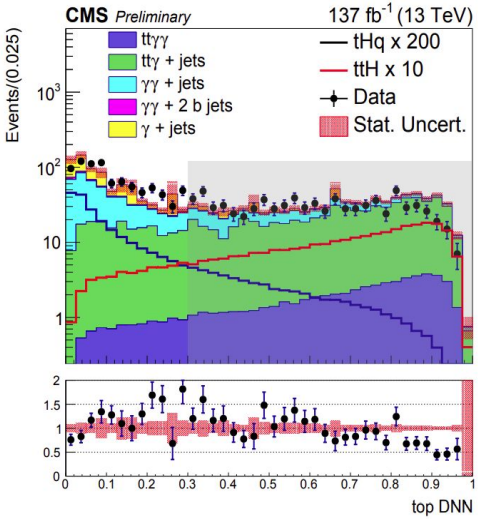
- Define a priority: analysis category with lower expected signal yield has higher priority
- Events that could enter more than one analysis category are assigned to the category with the highest priority

Other production modes (part of) ...





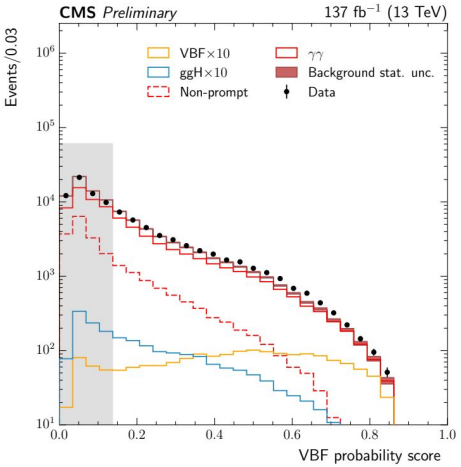


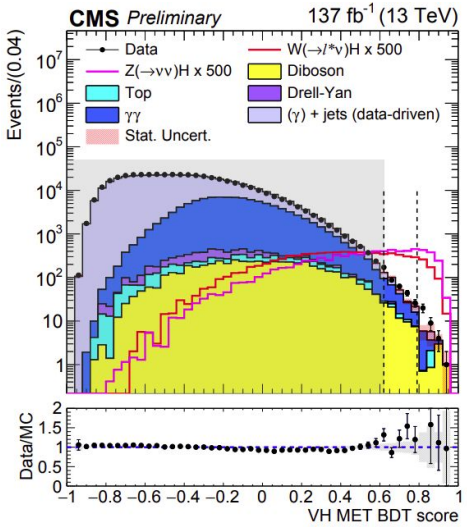


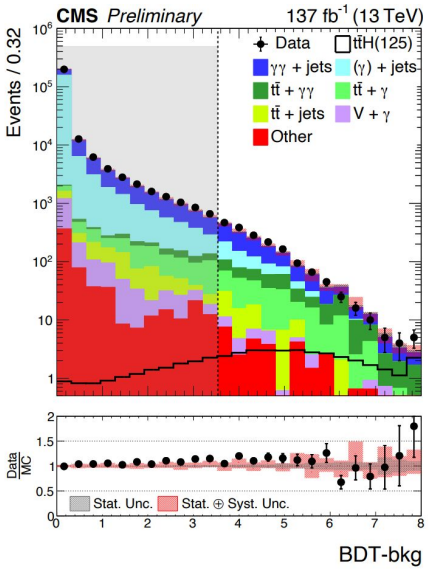
Selected highlights:

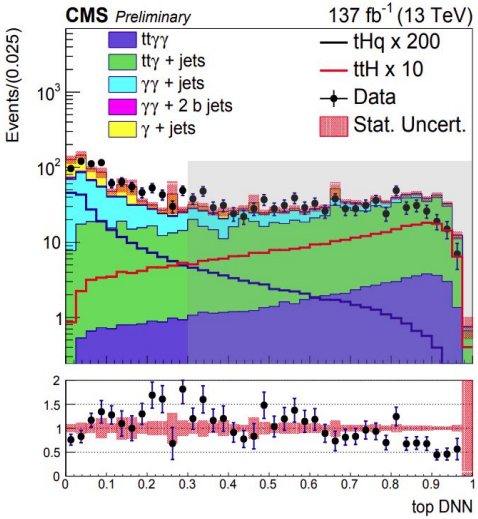
- New comprehensive categorizations for VH events
- Introduce dedicated categorization for tHq events (tHq Leptonic)
- Use data-driven estimation for certain key backgrounds for MVA training
  - VBF, VH MET, ttH Hadronic

Other production modes (part of) ...



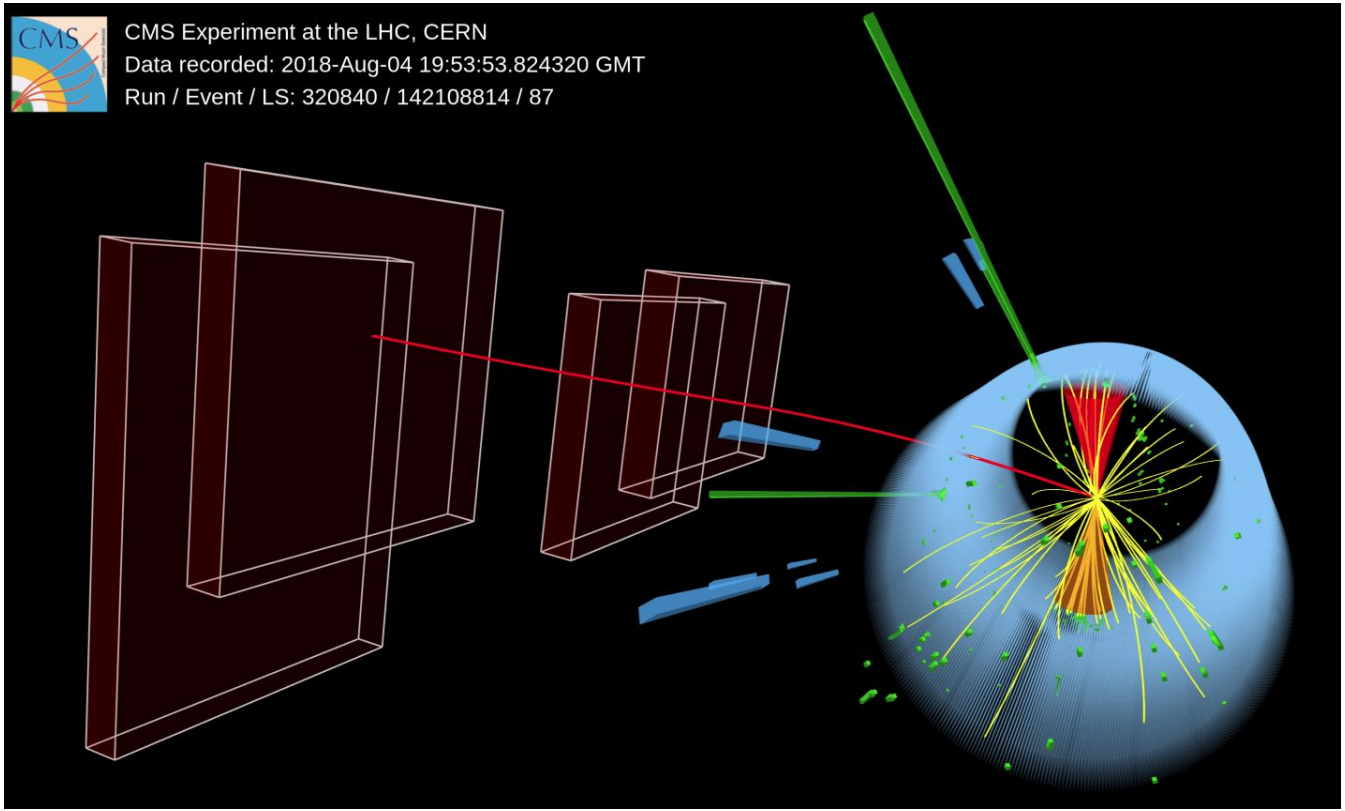




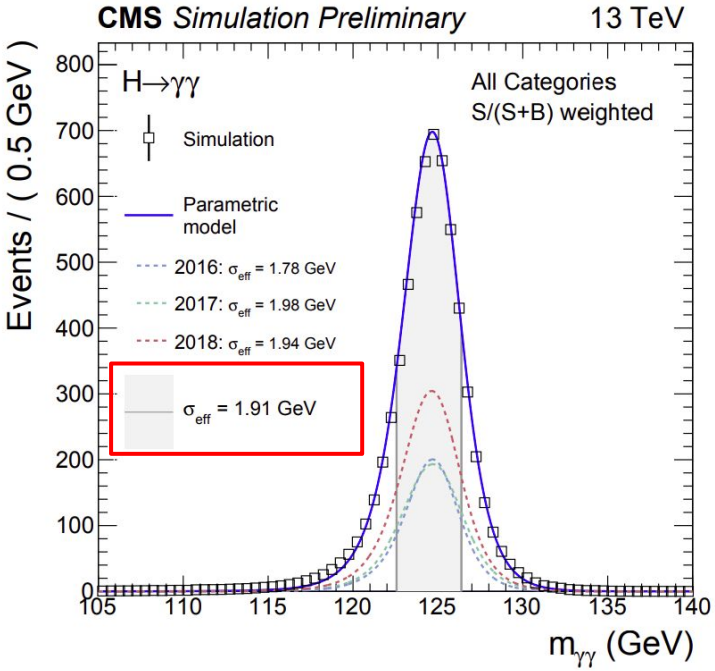


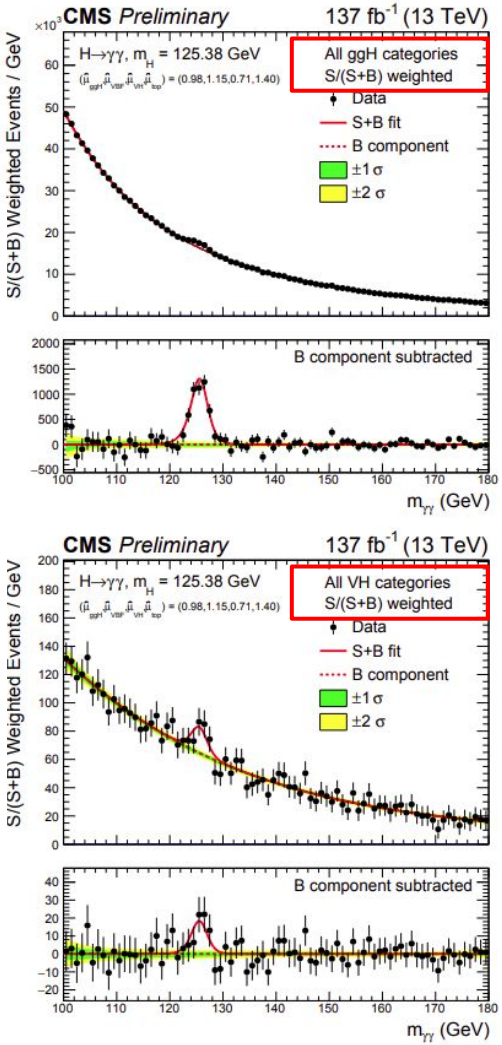
The analysis of each production mode can easily be expanded to a full talk, more details can be found in backup slides and I am happy to explain more details in the discussion session

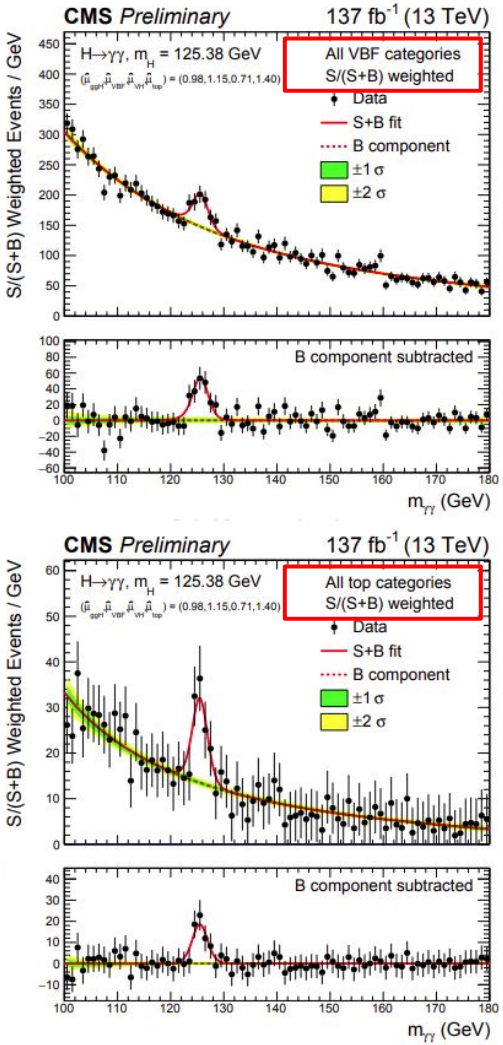
A tHq (H→γγ) candidate



$m_{\gamma\gamma}$ distributions

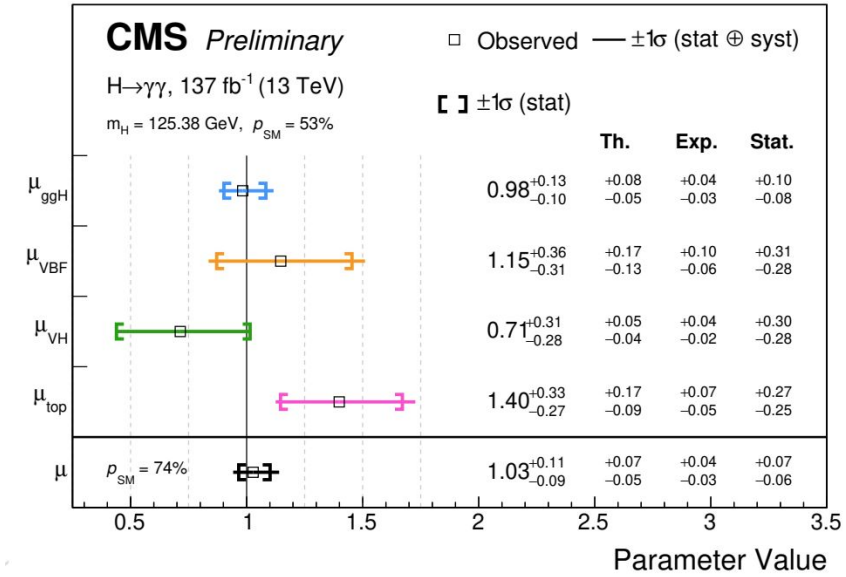






Signal (background) models are derived using simulation (data)

Signal strength modifiers



Common signal strength modifier (μ)

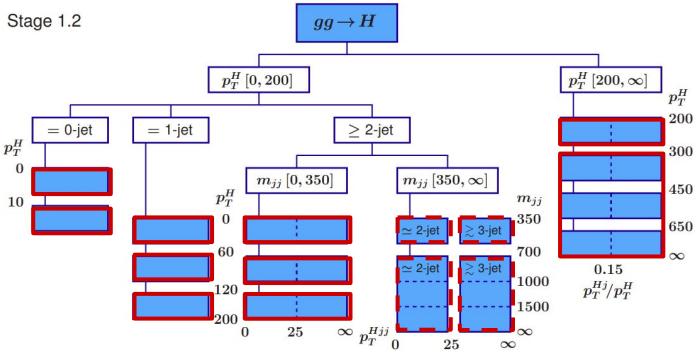
- Ratio of observed (σ_H x diphoton BR) to SM prediction

$$\mu = 1.03^{+0.11}_{-0.09} = 1.03^{+0.07}_{-0.05}\,(\text{theo})^{+0.04}_{-0.03}\,(\text{syst})^{+0.07}_{-0.06}\,(\text{stat}).$$

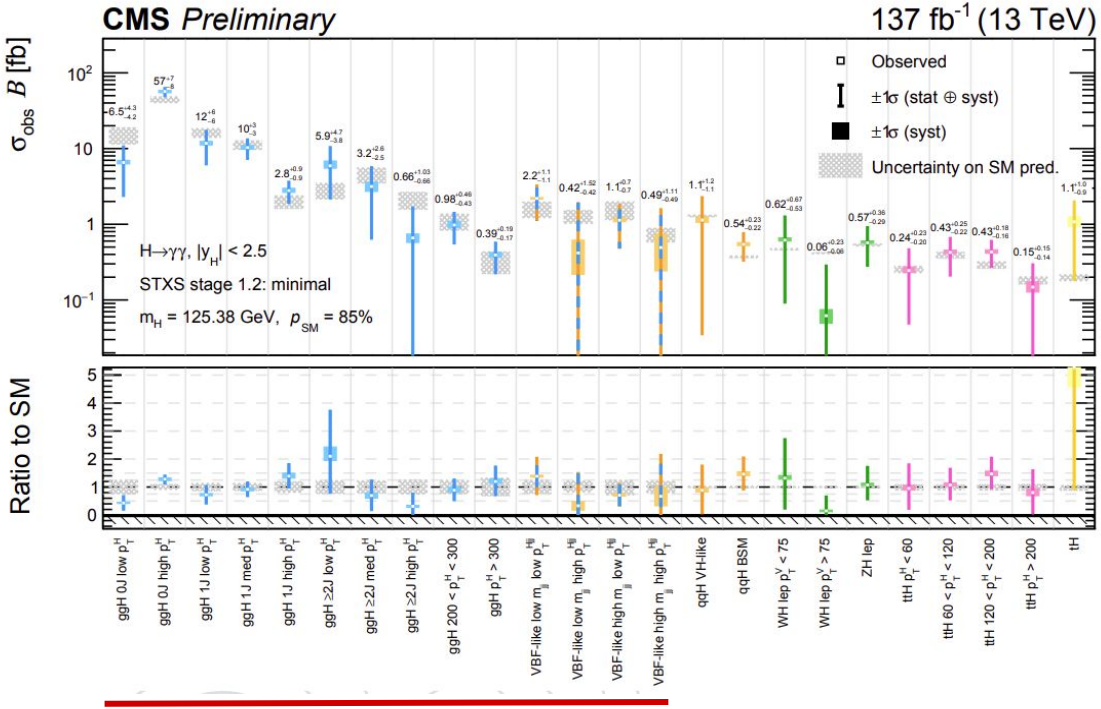
μ per production mode

- μ_{VH} : VH hadronic + VH leptonic
- μ_{VBF} : VBF production
- μ_{top} : ttH + tHq + tHW
- μ_{ggH} : ggH + bbH (starts to be systematic dominated)

STXS minimal merging scheme

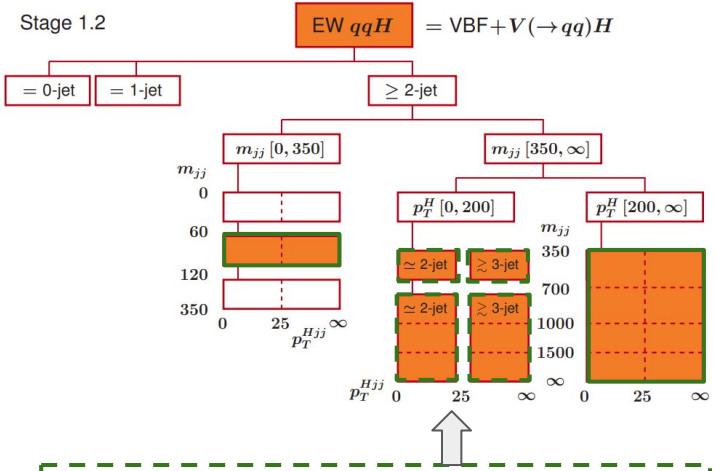


merge with the corresponding bins in qqH scheme to avoid large correlation

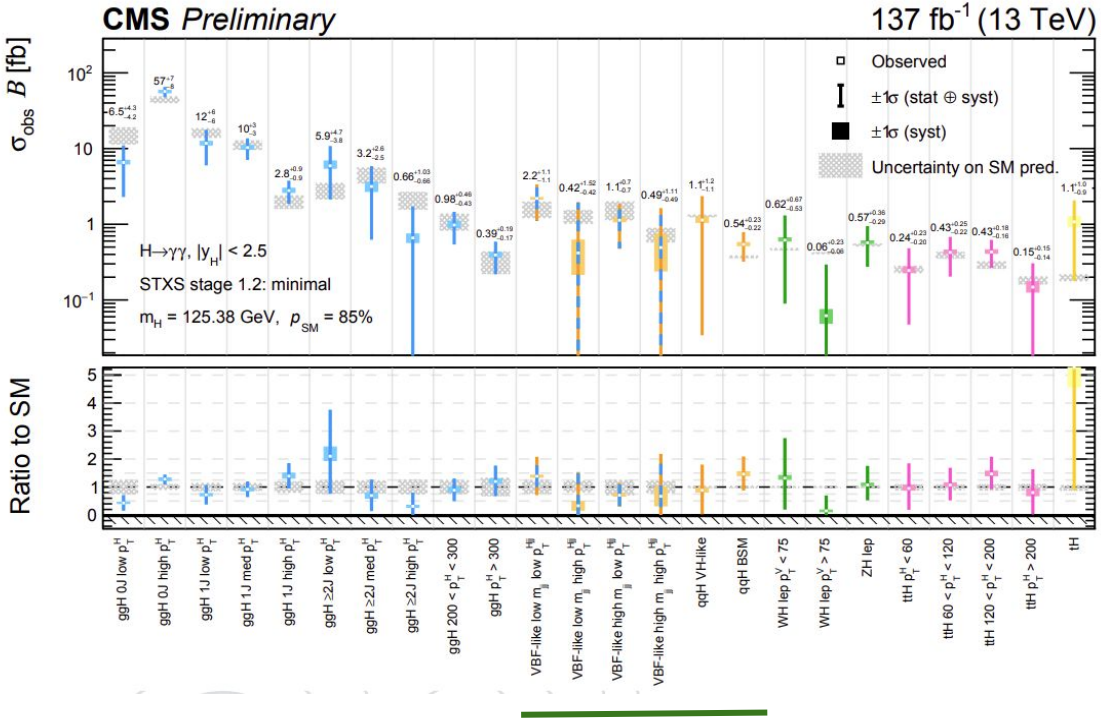


- 24 parameters of interest
- Merge as few STXS bins as possible
- Correlations between bins ≲ 0.75

STXS minimal merging scheme

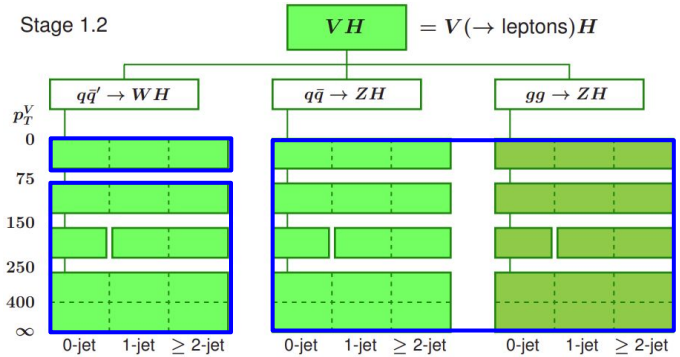


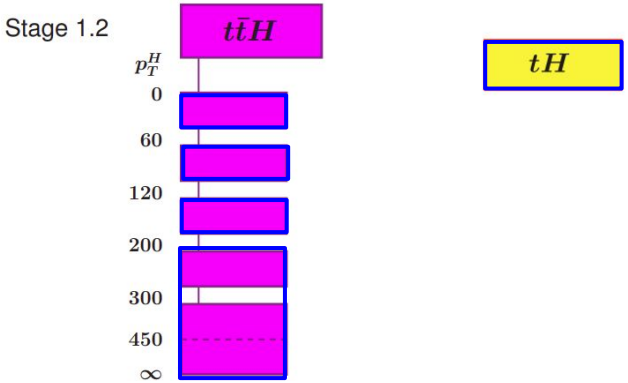
merge with the corresponding bins in ggH scheme to avoid large correlation

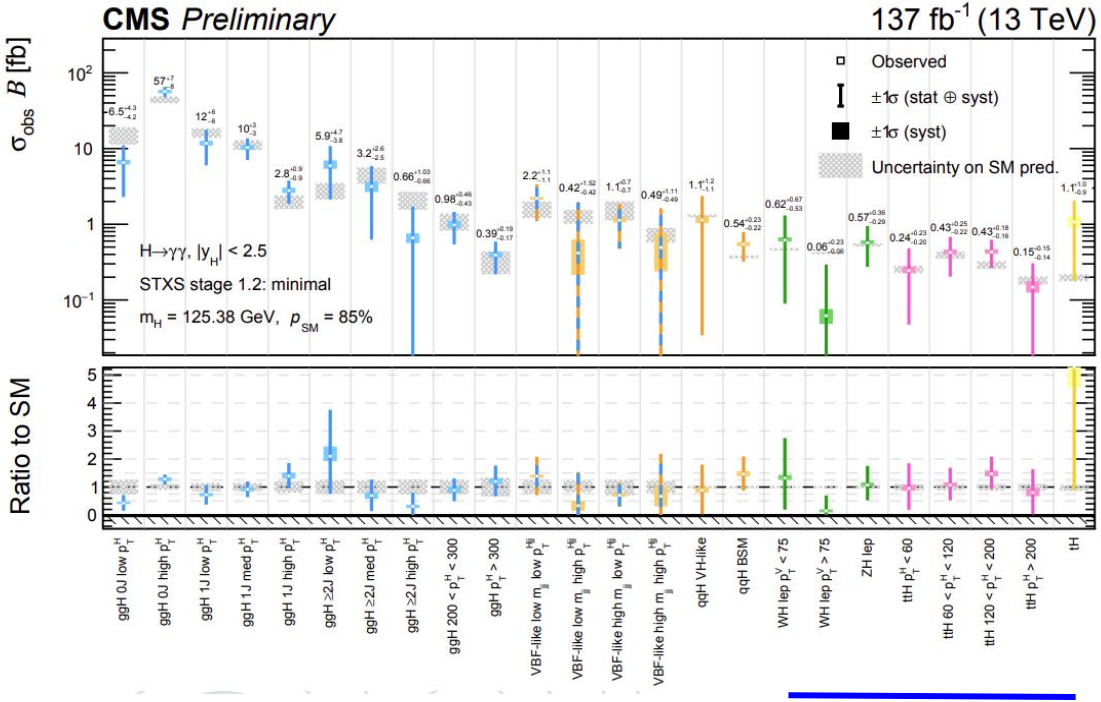


- 24 parameters of interest
- Merge as few STXS bins as possible
- Correlations between bins ≲ 0.75

STXS minimal merging scheme





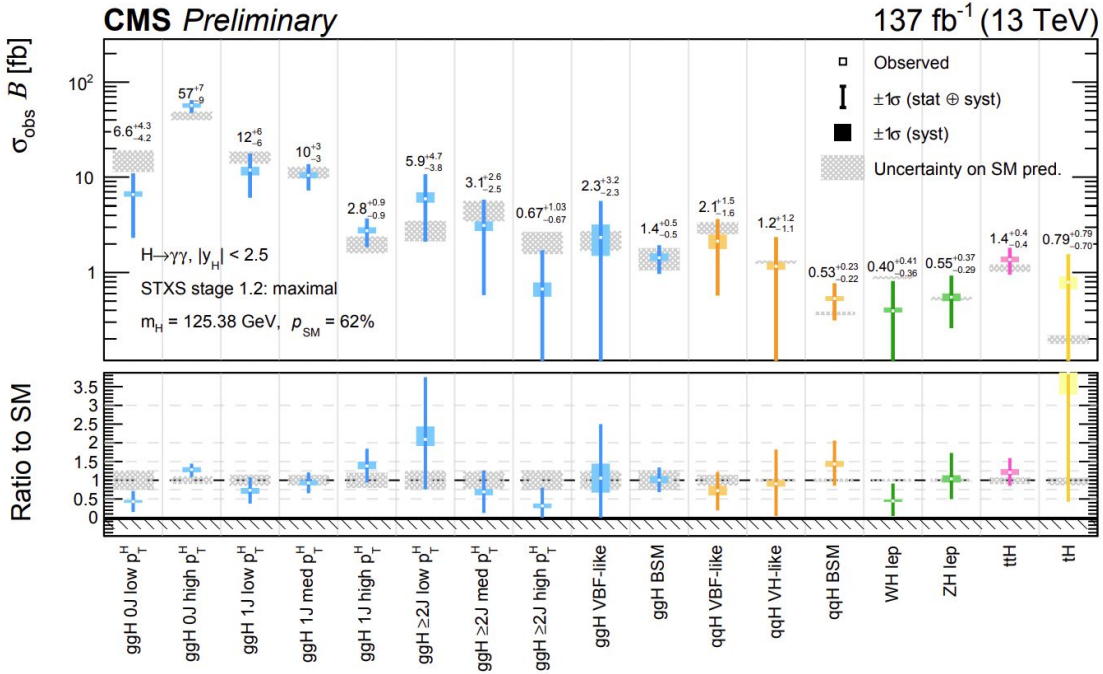


First measurement with additional stage 1.2 splitting (ttH and ggH BSM)

STXS <u>maximal</u> merging scheme

- 17 parameters of interest
- STXS bins merged until expected uncertainty is <150% of SM prediction

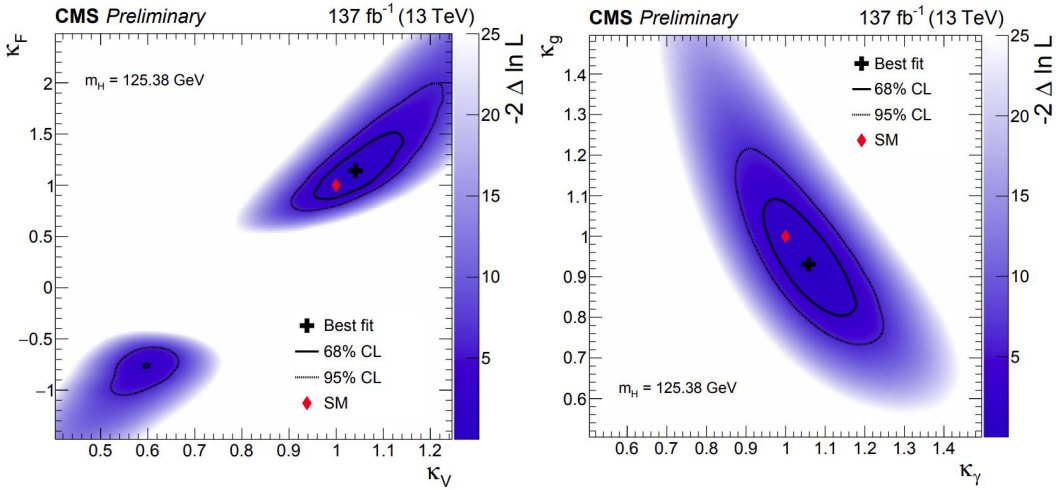
- observed (expected) limit on tH production at 95% CL:
  - 12 (9) times SM prediction
- σ(tH) * BR(H → γγ): 0.8 +0.8/-0.7 fb



Coupling modifiers

κ: parameterize potential deviations in the coupling of Higgs boson to other SM particles

For a given process in SM:
$\kappa^2 = \sigma/\sigma(SM)$ or $\kappa^2 = \Gamma/\Gamma(SM)$



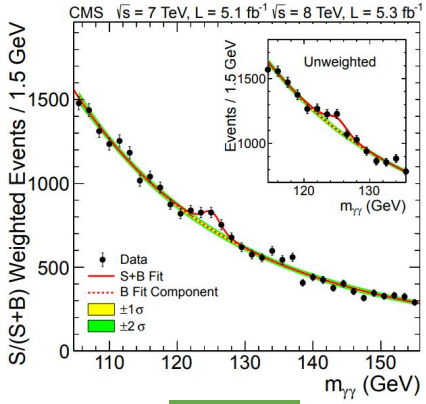
- Resolved κ model: scaling factors of the ggH and H → γγ loops are resolved into their SM components, in terms of the other κ parameters
  - Perform 2D scans of κ_V **and** κ_f : universal coupling modifiers to **vector bosons/fermions**
  - Region with negative κ_f is observed (expected) to be excluded with a significance of 1.0σ (2.3σ)
- Unresolved κ model: ggH and H → γγ loops have their own effective scaling factors (κ_g, κ_γ)

Summary

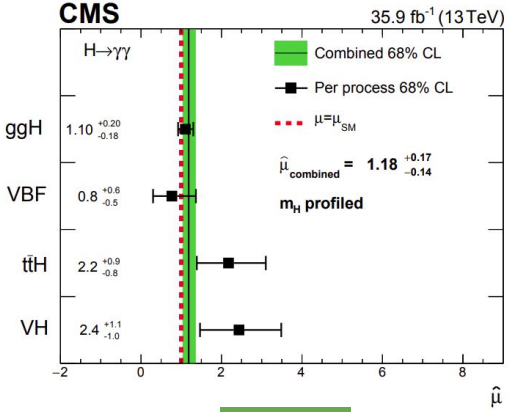
Presented H → γγ analysis based on 137fb^{-1} of data collected by the CMS detector during LHC Run 2 (CMS-PAS-HIG-19-015)

- Inclusive signal strength measurement: 1.03 +0.11/−0.09
- Per-production mode signal strengths and coupling modifiers consistent with SM prediction
- Cross-section measurement in the STXS framework at stage 1.2
  - Two bin merging scenarios considered (17 and 24 cross sections measured respectively)
  - Exclusion limit on tH production at 95% CL is 12 x SM (first time in CMS H→γγ result)
  - First measurement of ttH cross section in 4 bins of the Higgs boson transverse momentum

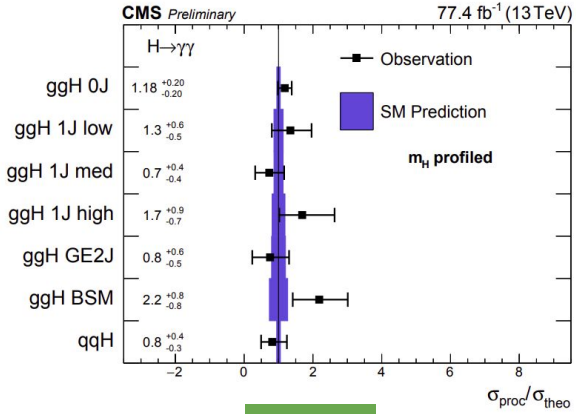
More Higgs property measurements with H → γγ decay are ongoing, stay tuned!



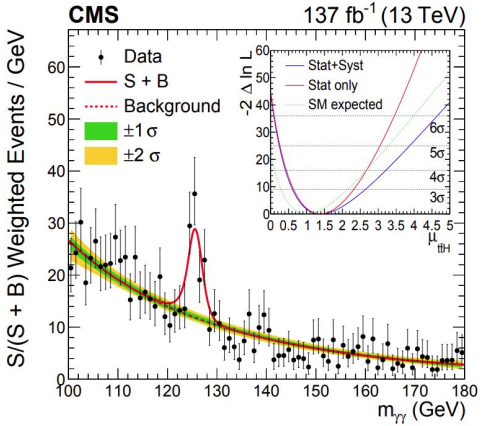
2012



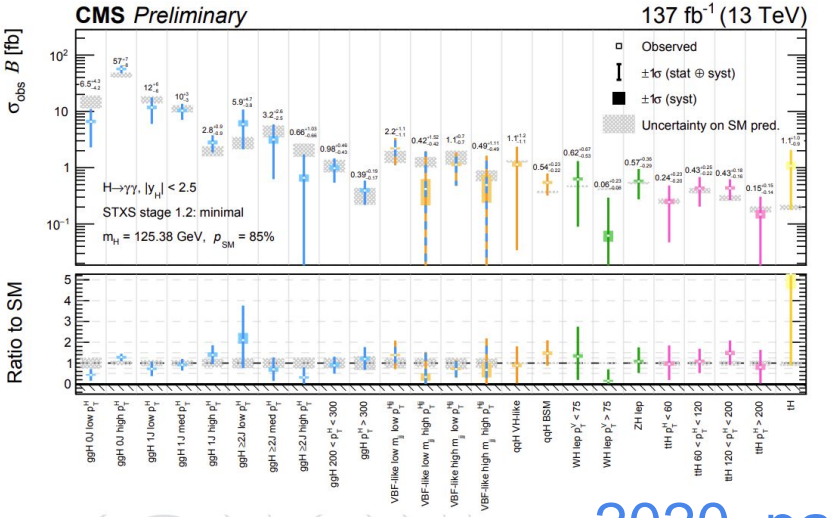
2016



2017



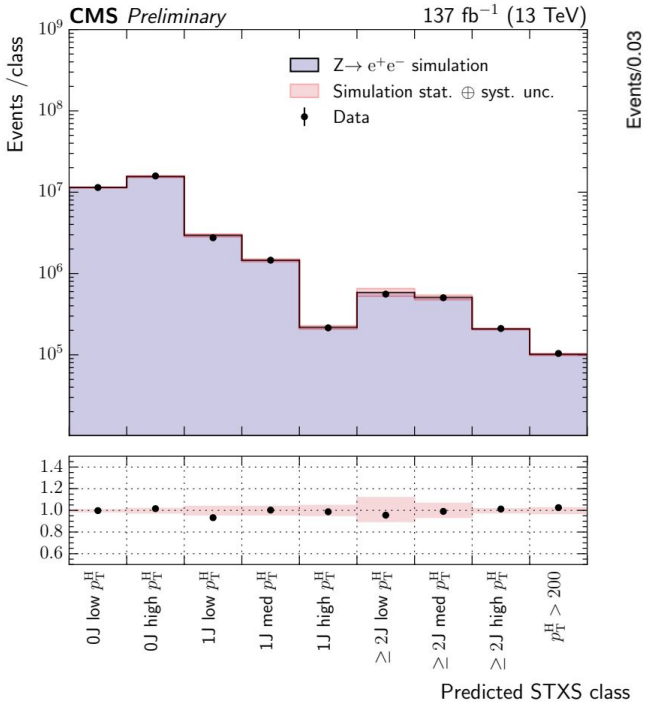
2018

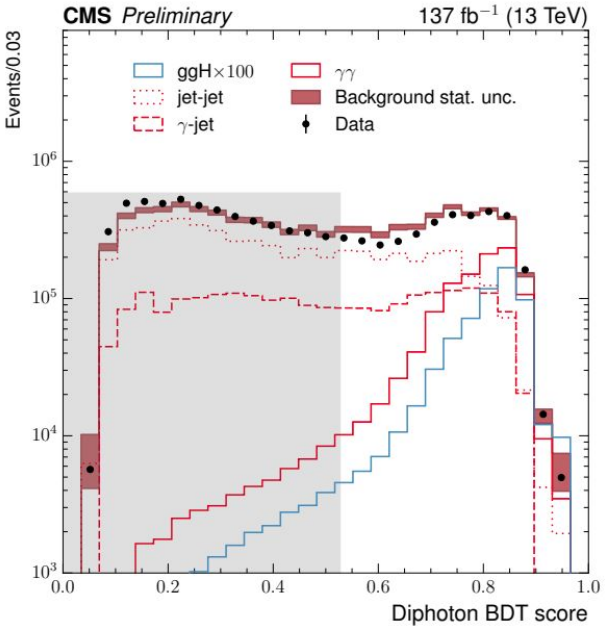


2020, party continues

Backup

Event categorisation for ggH



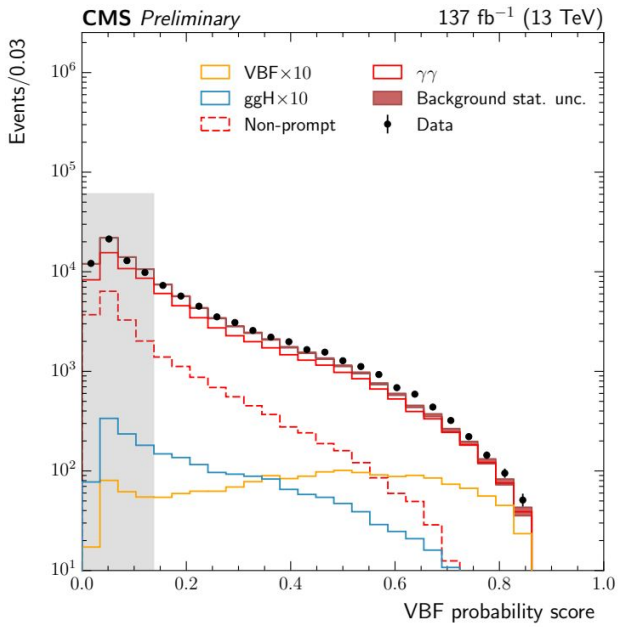


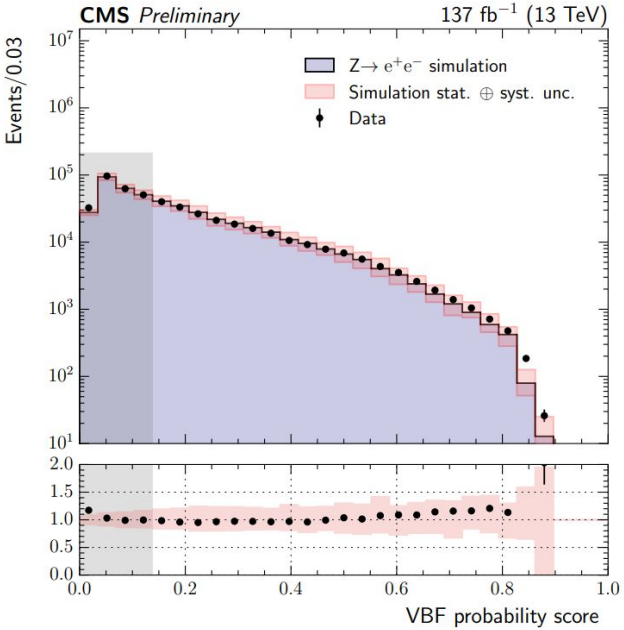
Use a multi-class BDT (9 classes)

- Predicts the probability that an event belongs to a given ggH STXS bin
- 5% improvement on accuracy of bin assignment
- Each class is divided into subcategories (different S/B)
  - With a diphoton BDT
  - Reduces background from SM diphoton production

Consider ggZH events with Z decays hadronically, VBF-like bins are categorised separately

Event categorisation for qqH (VBF)



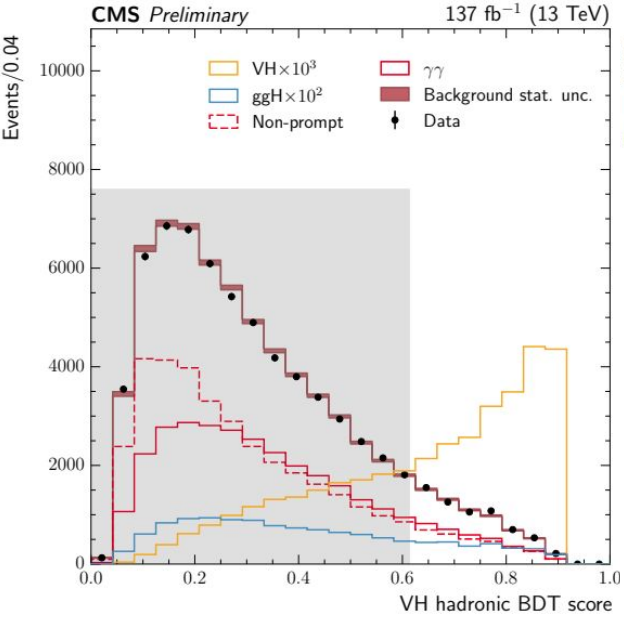


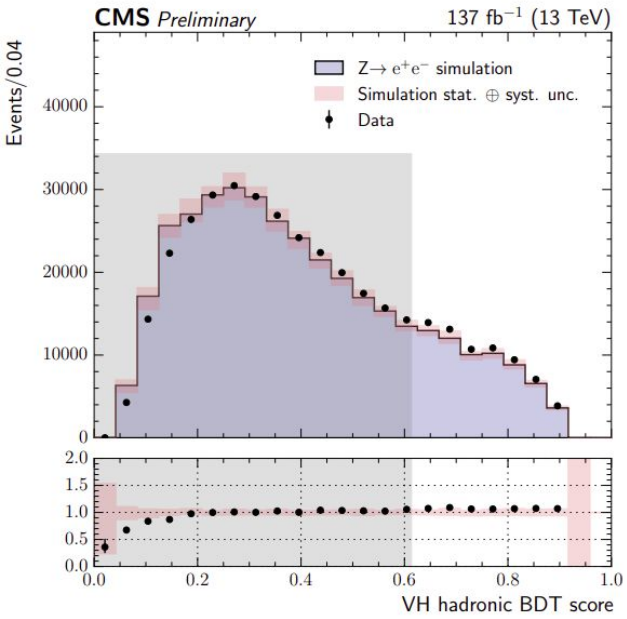
Categorization is based on a di-jet BDT and diphoton BDT (Target 5 STXS bins)

- Di-jet BDT: a 3-class BDT
- Predict the origin of an event (VBF, ggH, or SM diphoton)
  - Diphoton background with fake photon(s) is modeled from data control region
    - Defined by invert photon ID BDT score

Di-jets pre-selection: jet p_T > 40 (30) GeV, $|\eta|$ < 4.7, m_{jj} > 350 GeV

Event categorisation for qqH (VH Hadronic)





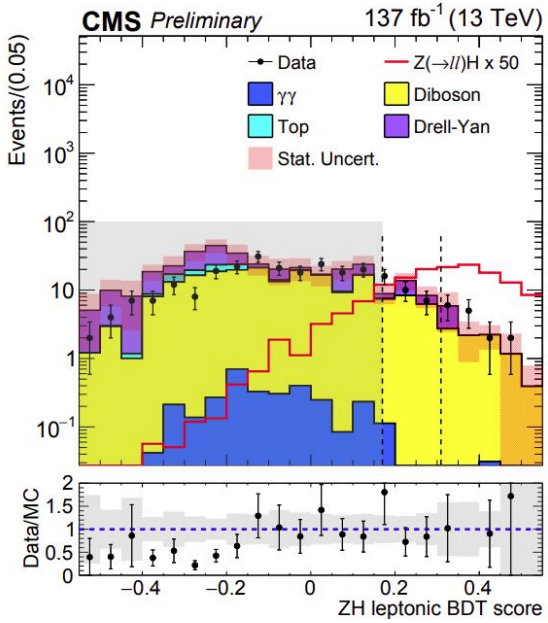
Categorization is based on a VH-Hadronic (VH-Had) BDT and diphoton BDT (Target 1 STXS bin)

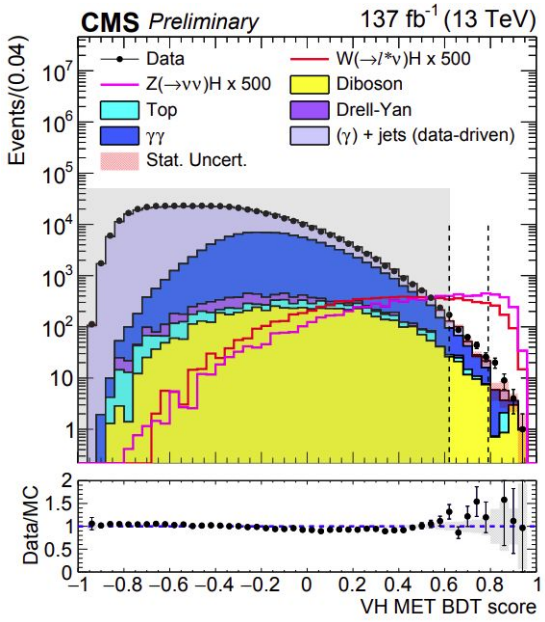
- VH-Had BDT: a 3-class BDT
- Predict the origin of an event (VH, ggH, or SM diphoton)
  - Diphoton background with fake photon(s) is modeled from data control region
    - Same strategy as in VBF

Di-jets pre-selection: jet p_T > 30 GeV, |η| < 2.4, 60 < m_{jj} < 120 GeV

Consider VH events where W/Z decays hadronically

Event categorisation for leptonic VH (2l or 0l)





2 same-flavour lepton

60 < m_{ll} < 120 GeV

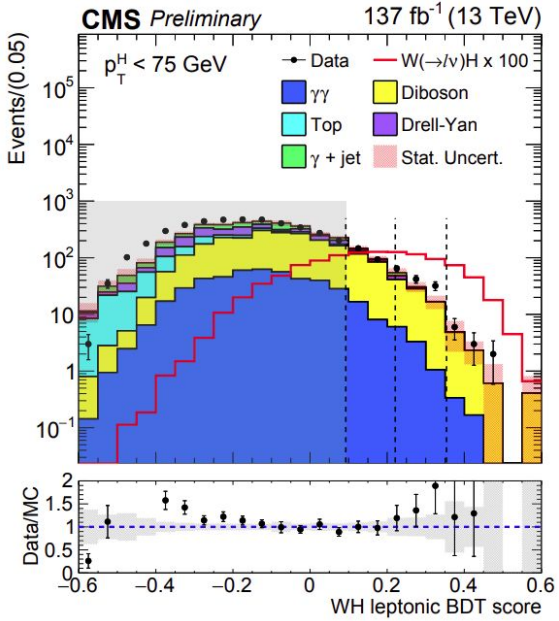
0 selected lepton

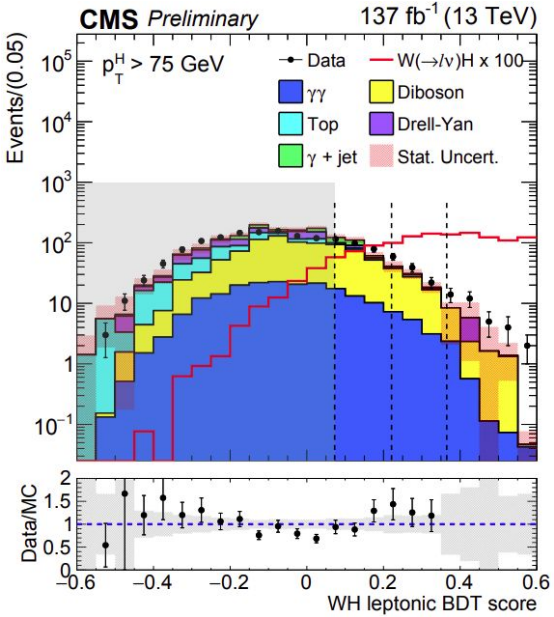
MET > 50 GeV

VH production (W or Z boson decays leptonically), possible final states include 0, 1 or 2 leptons

- For each 3 channels, a dedicated BDT is used to discriminate between VH signal and background
- For events without lepton (VH MET), diphoton background with fake photon(s) is modeled from data control region
  - Same strategy as used in CMS ttH (H→γγ) full Run 2 paper (arXiv:2003.10866)

Event categorisation for leptonic VH (1l)



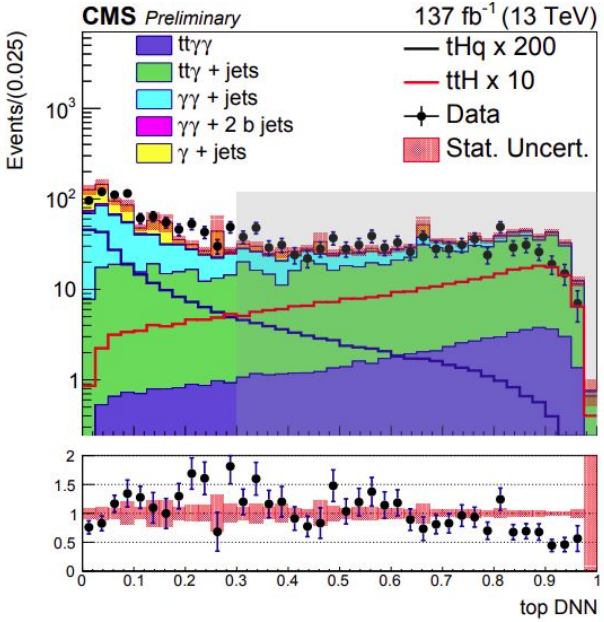


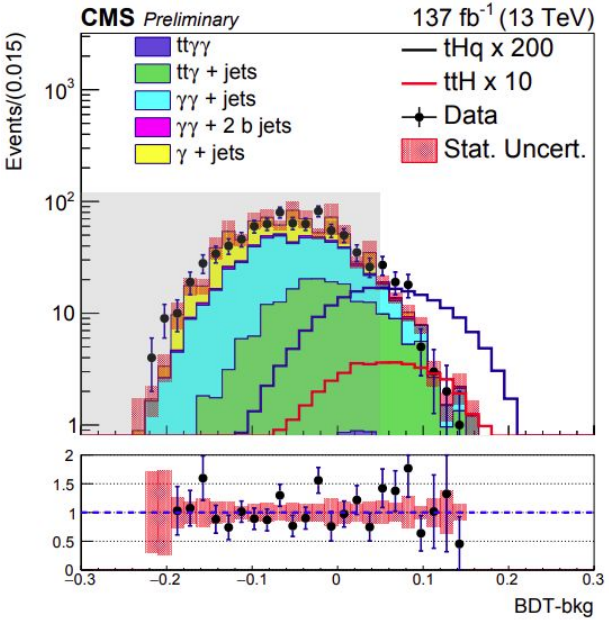
This phase space is sensitive to WH production mode with W decays leptonically

- To probe finer STXS bins splitting, use $p_T(\gamma\gamma)$ = 75 GeV threshold as a proxy of $p_T(V)$ to divide events into two categories
- Due to the presence of neutrino, the W boson itself is not fully reconstructable

Select events with one lepton, with additional requirements on photon ID BDT and events being incompatible with Drell--Yan

Event categorisation for tH



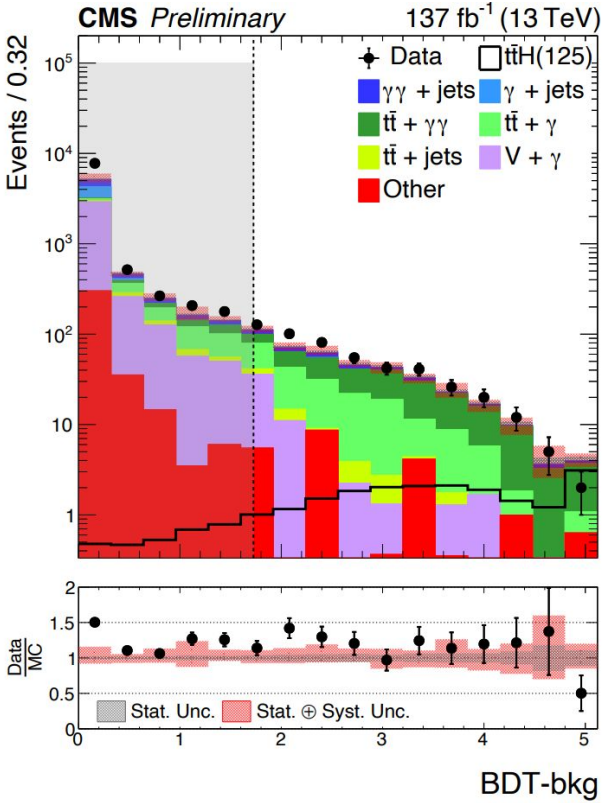


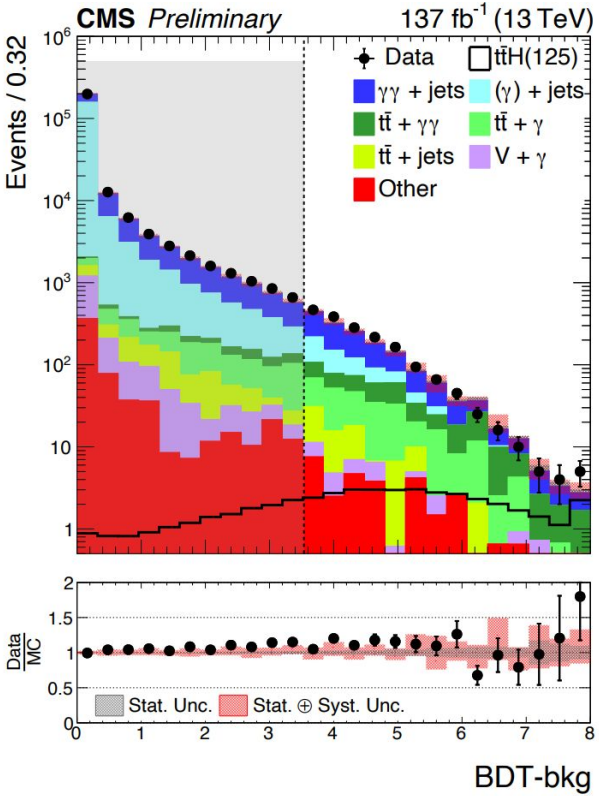
Construct a dedicated category enriched in tHq events where t quark decays leptonically

- Utilizes a dedicated DNN and a tHq leptonic BDT-bkg
- A top DNN is trained to reject ttH events, as final states of tHq and ttH production are very similar
- A BDT-bkg is employed to further reject other non-Higgs SM backgrounds

Pre-selection: ≥ 1 lepton, ≥ 1 b-jet, ≥ 1 additional jet

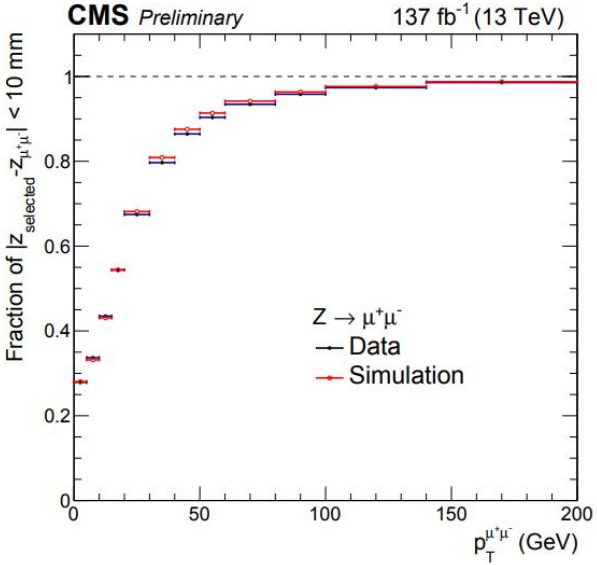
Event categorisation for ttH

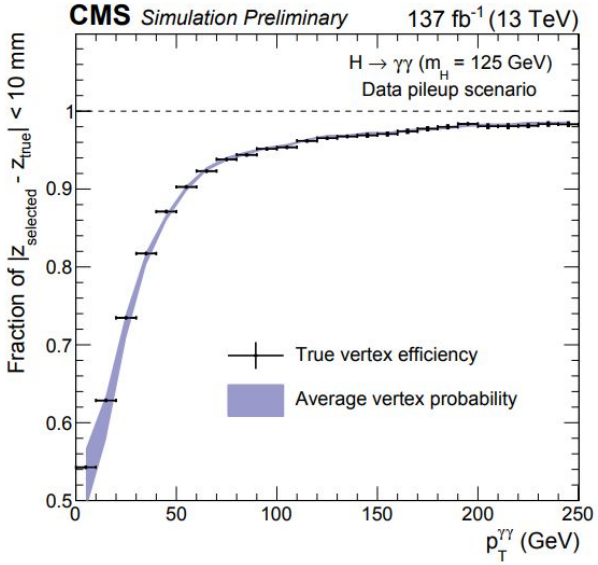




- Same techniques in recent CMS Run 2 ttH (H→γγ) paper (arXiv:2003.10866) for event categorisation
- Further requirement on top DNN score to reduce contamination from tHq
- Construct additional categories to probe individual STXS bins
  - Finer STXS bin splitting using $p_T(\gamma\gamma)$

Vertex identification





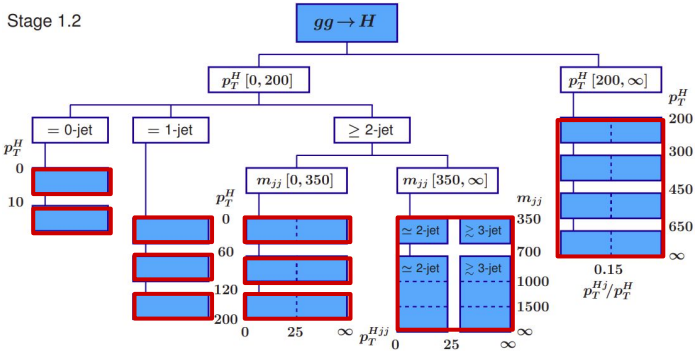
Vertex probability BDT

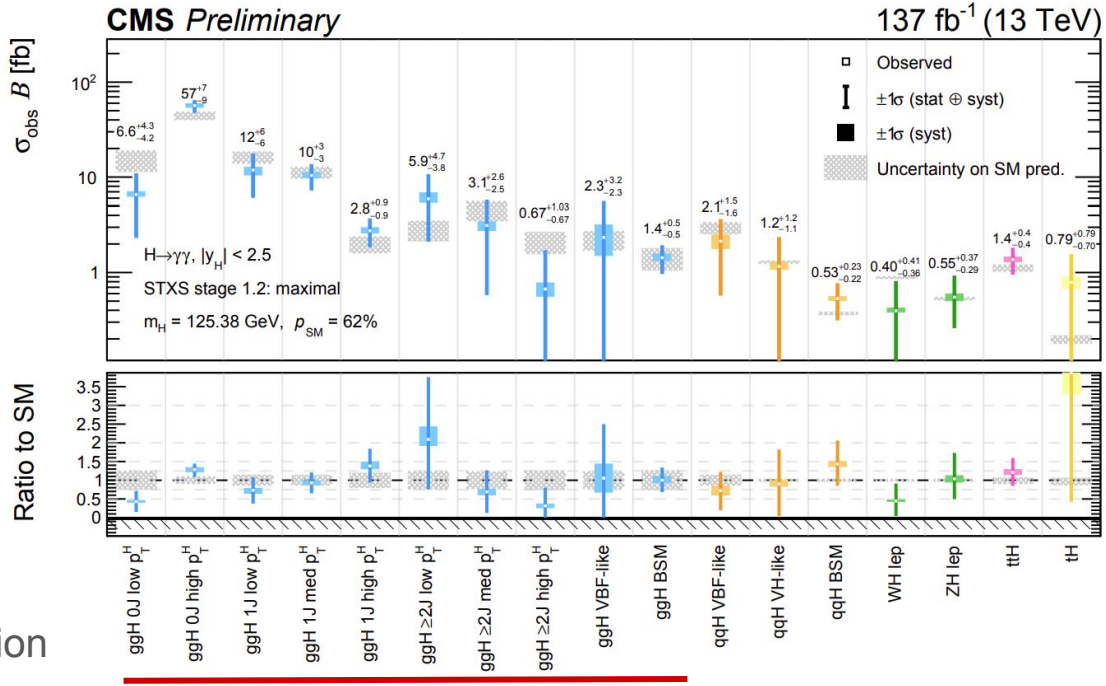
- Estimates probability of assigned vertex position being true within 1 cm
- One of the input to the diphoton BDT

Use a BDT for vertex identification, if vertex position is true within 1 cm then $m_{\gamma\gamma}$ resolution is dominated by photon energy resolution

- true for ~79% events in simulation Vertex
- ID BDT validated with $Z \to \mu\mu$ events

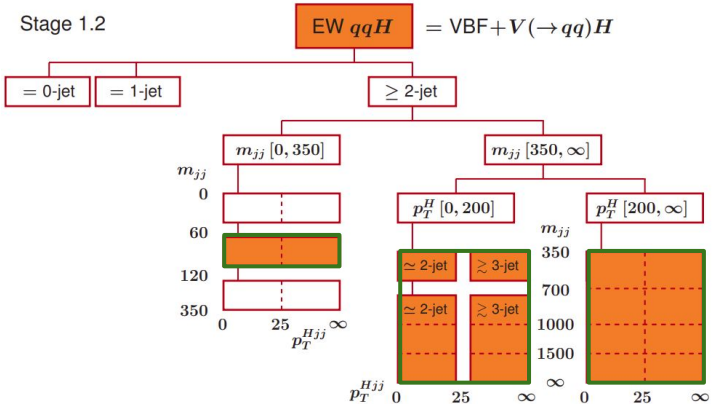
STXS maximal merging scheme

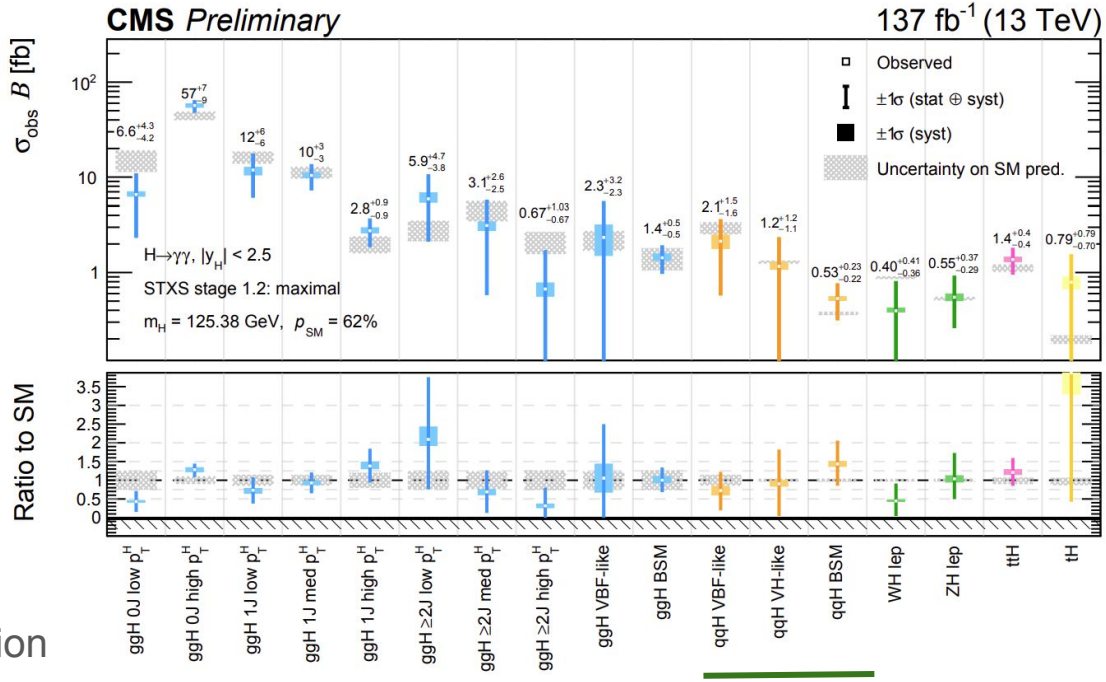




- 17 parameters of interest
- STXS bins merged until expected uncertainty is <150% of SM prediction

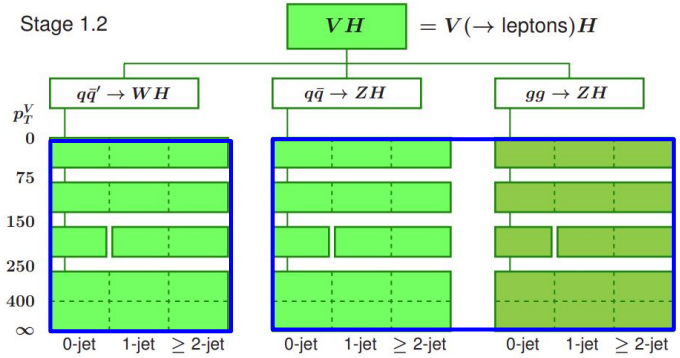
STXS maximal merging scheme

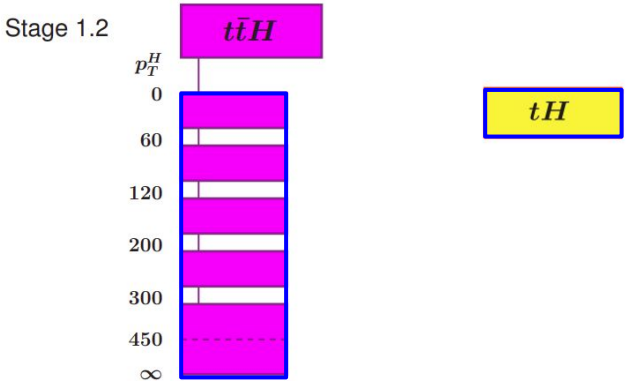


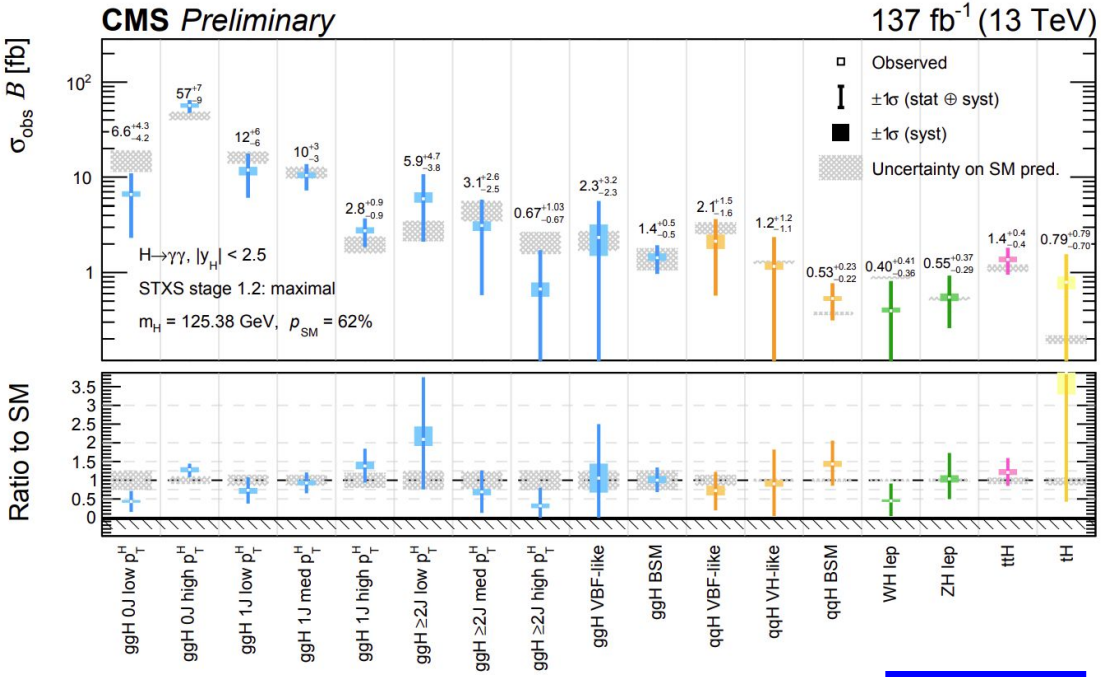


- 17 parameters of interest
- STXS bins merged until expected uncertainty is <150% of SM prediction

STXS maximal merging scheme

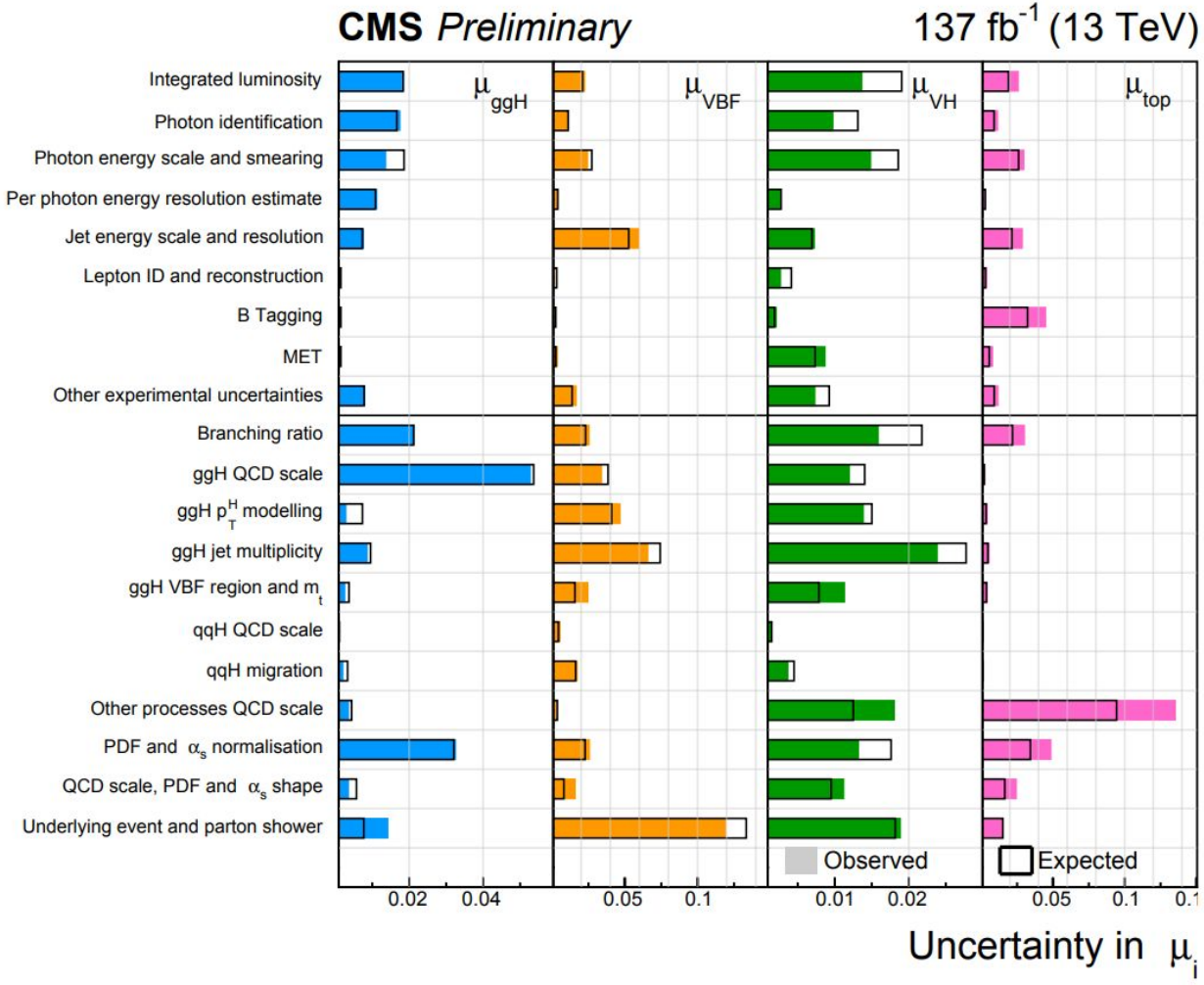




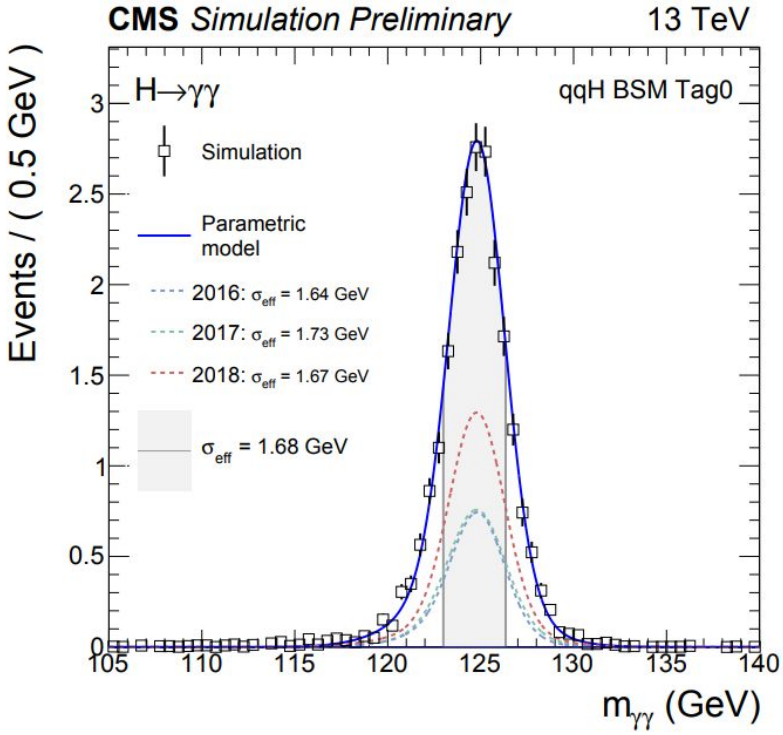


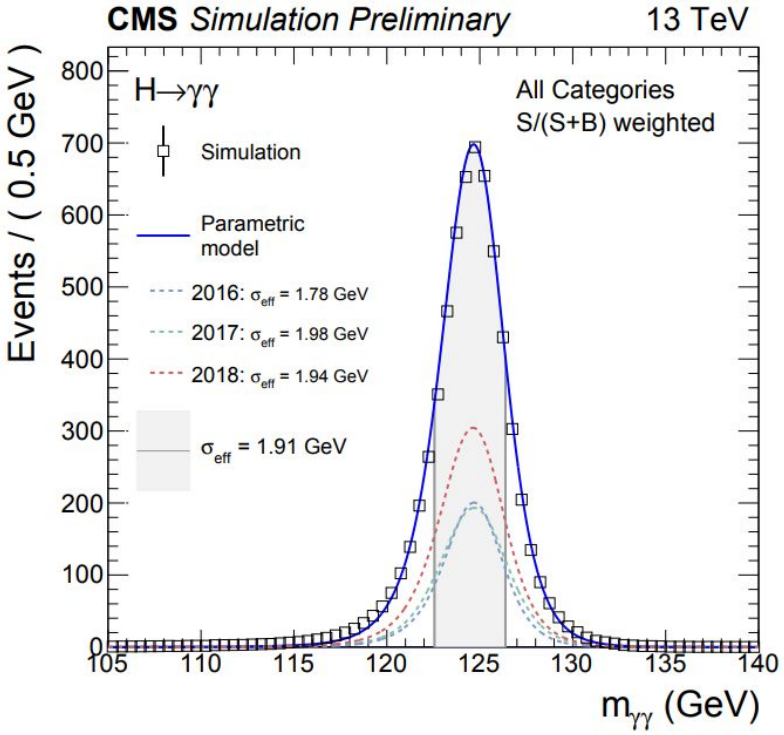
- observed (expected) limit on tH production at 95% CL: 12 (9) tims SM prediction
- σ(tH) * BR(H → γγ) = 0.8 +0.8/-0.7 fb (best to-date)

Uncertainty



Signal models





Correlations

Maximal merging

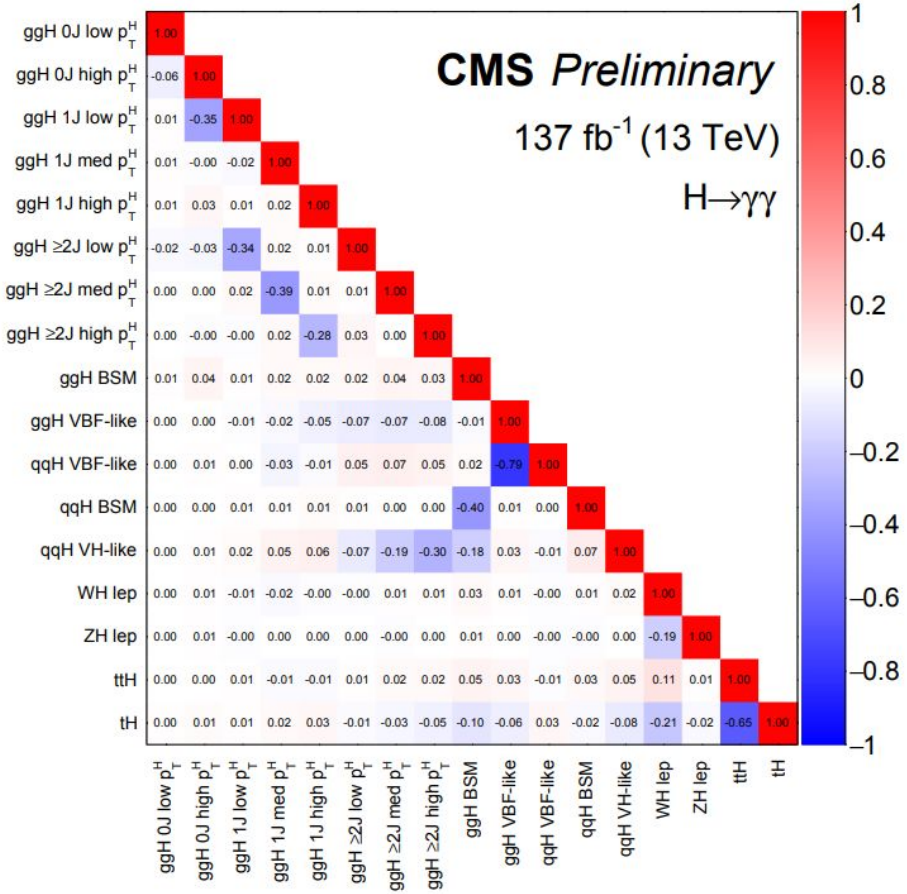


Figure 19: Observed correlations between the 17 parameters considered in the maximal merging fit. The size of the correlations is indicated by the colour scale.

Correlations

Minimal merging

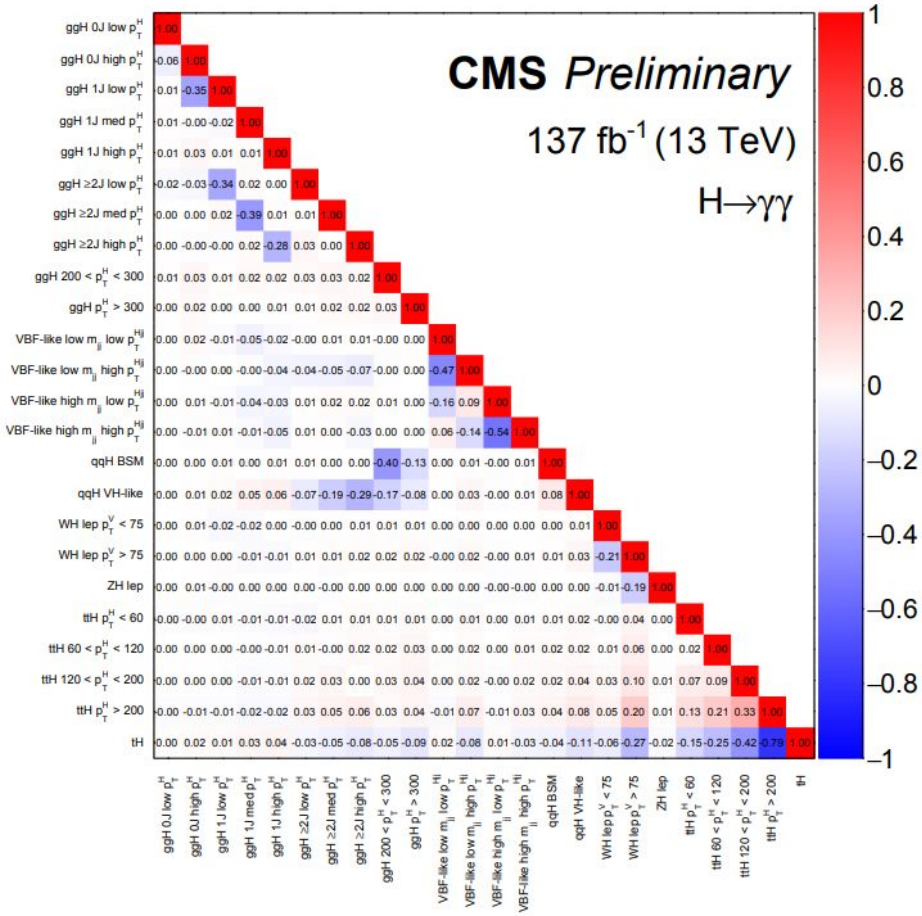
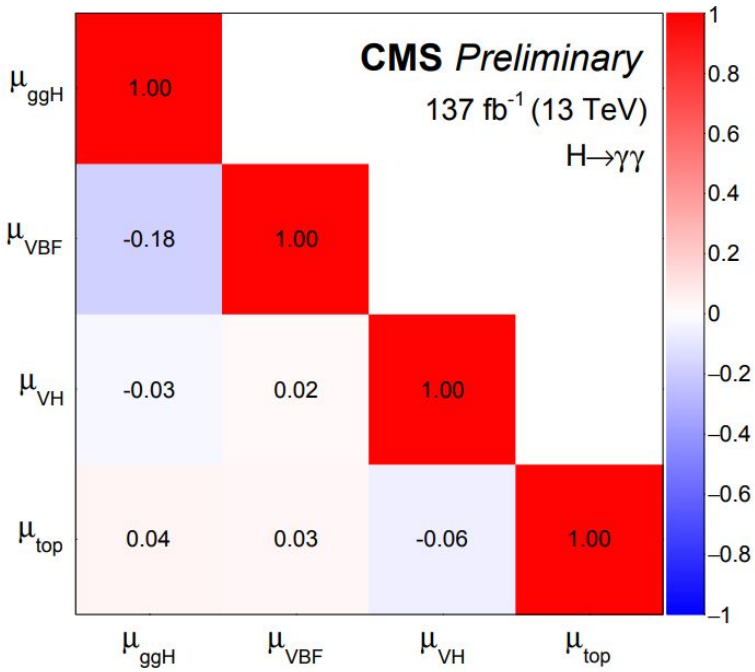


Figure 21: Observed correlations between the 24 parameters considered in the minimal merging fit. The size of the correlations is indicated by the colour scale.

Category sequence

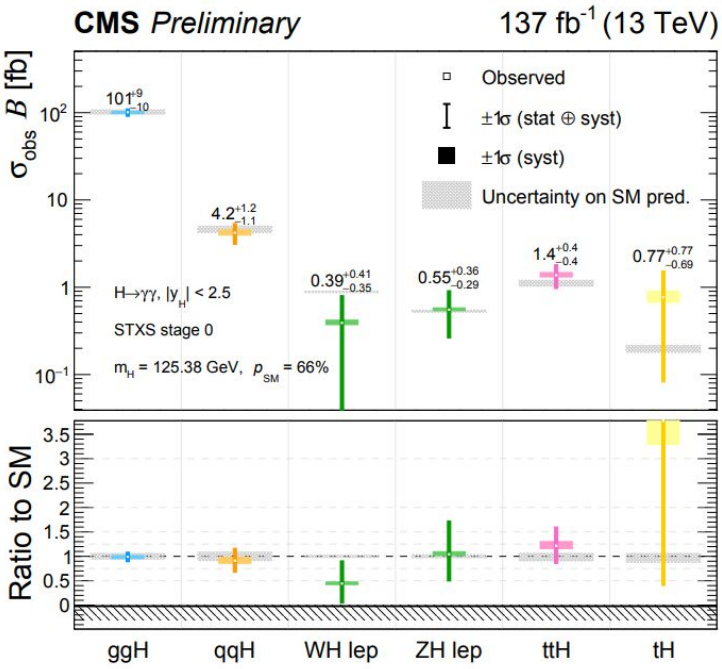
Categorisation region	Particle level STXS bin, (units in GeV)	Number of categories
tHq leptonic	tHq	1
$t\bar{t}H$ leptonic	$t\bar{t}H$ $p_T^H < 60$	4
	$t\bar{t}H$ $60 < p_T^H < 120$	2
	$t\bar{t}H$ $120 < p_T^H < 200$	2
	all $t\bar{t}H$ $p_T^H > 200$ (2 bins total)	2
ZH leptonic	all ZH lep and ggZH lep bins (10 bins total)	2
WH leptonic	WH lep $p_T^V < 75$	3
	all WH lep $p_T^V > 75$ (4 bins total)	3
VH MET	all VH leptonic bins (15 bins total)	2
$t\bar{t}H$ hadronic	$t\bar{t}H$ $p_T^H < 60$	4
	$t\bar{t}H$ $60 < p_T^H < 120$	4
	$t\bar{t}H$ $120 < p_T^H < 200$	4
	all $t\bar{t}H$ $p_T^H > 200$ (2 bins total)	4
VBF	qqH VBF-like low m_{jj} low p_T^{Hjj}	2
	qqH VBF-like low m_{jj} high p_T^{Hjj}	2
	qqH VBF-like high m_{jj} low p_T^{Hjj}	2
	qqH VBF-like high m_{jj} high p_T^{Hjj}	2
	qqH BSM	2
	all ggH VBF-like (4 bins total)	2
VH hadronic	qqH VH-like	2
ggH	ggH 0J low p_T^H	3
	ggH 0J high p_T^H	3
	ggH 1J low p_T^H	3
	ggH 1J med p_T^H	3
	ggH 1J high p_T^H	3
	ggH $\geq$2J low p_T^H	3
	ggH $\geq$2J med p_T^H	3
	ggH $\geq$2J high p_T^H	3
	ggH $200 < p_T^H < 300$	2
	ggH $300 < p_T^H < 450$	2
	ggH $450 < p_T^H < 650$	1
	ggH $p_T^H > 650$	1
No categories	qqH 0J, 1J, m_{jj} <60, qqH 120 <m_{jj} <350, $b\bar{b}H$, tHW	0

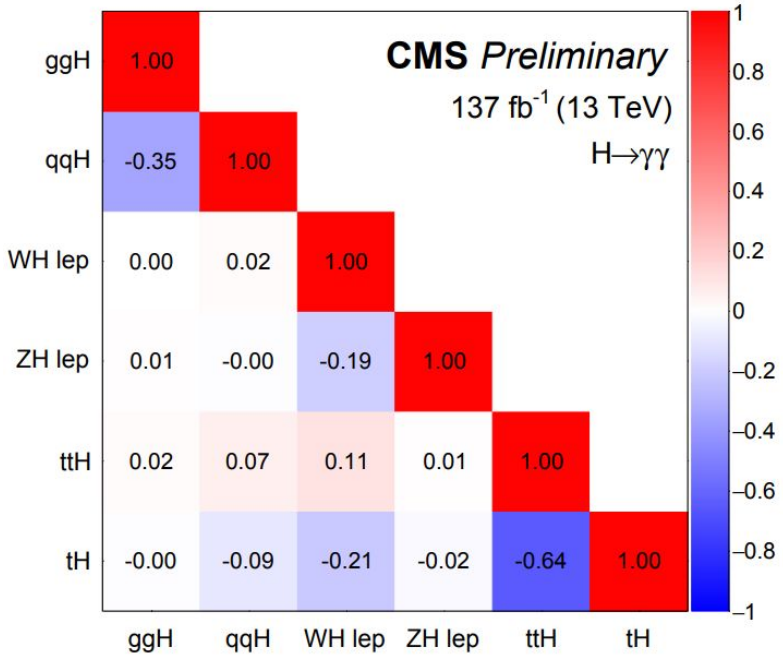
Correlation for per-production mode



Observed correlations between the parameters in the per-production-mode signal strengths fit. The size of the correlations is indicated by the colour scale

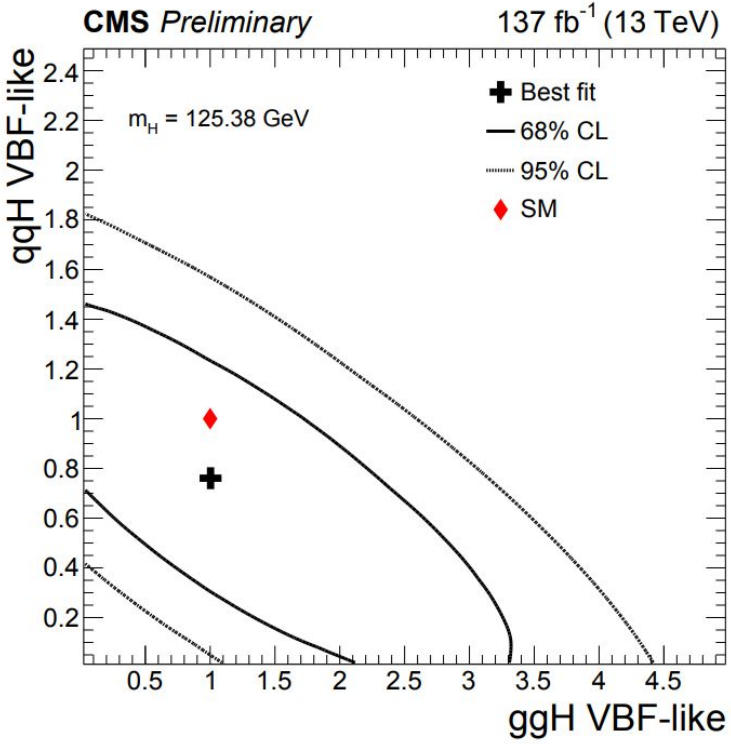
Stage 0 STXS


CMS Preliminary
137 fb-1 (13 TeV)
σobs B [fb]
Observed
±1σ (stat ⊕ syst)
±1σ (syst)
Uncertainty on SM pred.
101+9 −10
4.2+1.2 −1.1
0.39+0.41 −0.35
0.55+0.36 −0.29
1.4+0.4 −0.4
0.77+0.77 −0.69
H→γγ, |yH| < 2.5
STXS stage 0
mH = 125.38 GeV, pSM = 66%
Ratio to SM
ggH
qqH
WH lep
ZH lep
ttH
tH

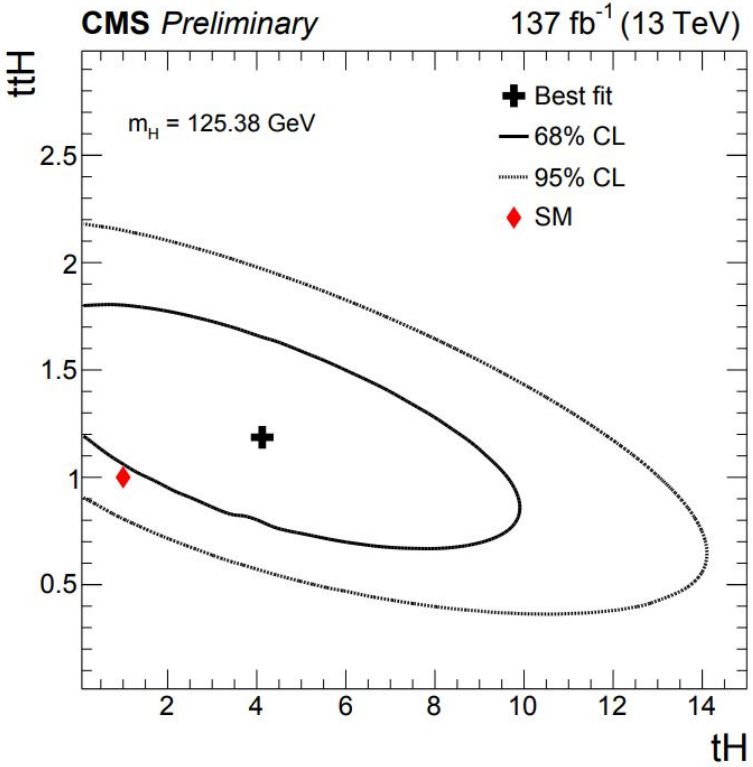

CMS Preliminary
137 fb-1 (13 TeV)
H→γγ
ggH 1.00
qqH -0.35 1.00
WH lep 0.00 0.02 1.00
ZH lep 0.01 -0.00 -0.19 1.00
ttH 0.02 0.07 0.11 0.01 1.00
tH -0.00 -0.09 -0.21 -0.02 -0.64 1.00
ggH qqH WH lep ZH lep ttH tH

Stage 0 STXS

STXS stage 0					
Parameters	$\sigma\mathcal{B}$ [fb]				$\sigma\mathcal{B}/(\sigma\mathcal{B})_{\mathrm{SM}}$
	SM prediction	Observed (Expected)			Observed (Expected)
	(m_H = 125.38 GeV)	Best fit	Stat unc.	Syst unc.	Best fit
ggH	$102.04^{+6.82}_{-6.78}$	$101.12^{+9.49}_{-10.31}\left(^{+10.31}_{-10.00}\right)$	$^{+9.18}_{-9.59}\left(^{+9.59}_{-9.49}\right)$	$^{+2.55}_{-3.78}\left(^{+3.88}_{-3.30}\right)$	$0.99^{+0.09}_{-0.10}\left(^{+0.10}_{-0.10}\right)$
qqH	$4.62^{+0.45}_{-0.44}$	$4.21^{+1.16}_{-1.11}\left(^{+1.22}_{-1.16}\right)$	$^{+1.12}_{-1.08}\left(^{+1.17}_{-1.13}\right)$	$^{+0.32}_{-0.25}\left(^{+0.35}_{-0.27}\right)$	$0.91^{+0.25}_{-0.24}\left(^{+0.26}_{-0.25}\right)$
WH lep	$0.88^{+0.03}_{-0.03}$	$0.39^{+0.41}_{-0.35}\left(^{+0.45}_{-0.40}\right)$	$^{+0.41}_{-0.35}\left(^{+0.45}_{-0.40}\right)$	$^{+0.03}_{-0.02}\left(^{+0.06}_{-0.06}\right)$	$0.44^{+0.47}_{-0.40}\left(^{+0.51}_{-0.45}\right)$
ZH lep	$0.53^{+0.02}_{-0.02}$	$0.55^{+0.36}_{-0.29}\left(^{+0.38}_{-0.33}\right)$	$^{+0.36}_{-0.29}\left(^{+0.38}_{-0.33}\right)$	$^{+0.03}_{-0.01}\left(^{+0.03}_{-0.01}\right)$	$1.04^{+0.68}_{-0.55}\left(^{+0.72}_{-0.61}\right)$
ttH	$1.13^{+0.08}_{-0.11}$	$1.37^{+0.44}_{-0.41}\left(^{+0.41}_{-0.41}\right)$	$^{+0.43}_{-0.40}\left(^{+0.40}_{-0.41}\right)$	$^{+0.10}_{-0.06}\left(^{+0.08}_{-0.05}\right)$	$1.21^{+0.39}_{-0.36}\left(^{+0.36}_{-0.36}\right)$
tH	$0.20^{+0.01}_{-0.03}$	$0.77^{+0.77}_{-0.69}\left(^{+0.72}_{-0.20}\right)$	$^{+0.76}_{-0.68}\left(^{+0.72}_{-0.20}\right)$	$^{+0.14}_{-0.10}\left(^{+0.08}_{-0.08}\right)$	$3.78^{+3.79}_{-3.37}\left(^{+3.53}_{-1.00}\right)$

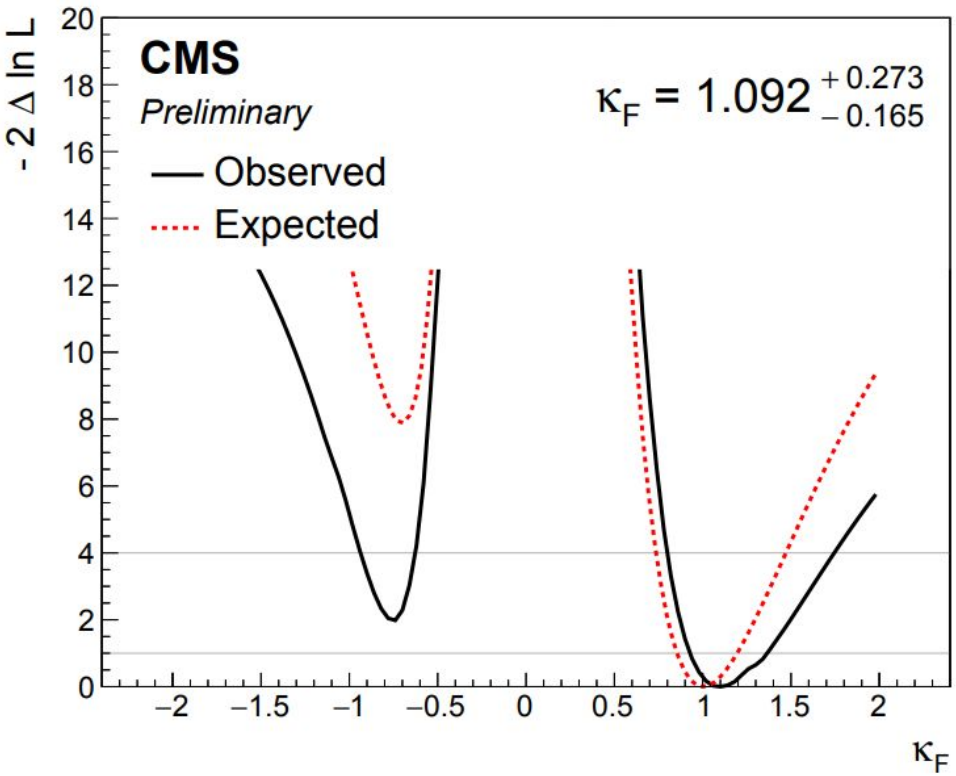


Two dimensional likelihood profile for ggH VBF-like vs qqH VBF-like in the maximal merging scheme with the largest correlations

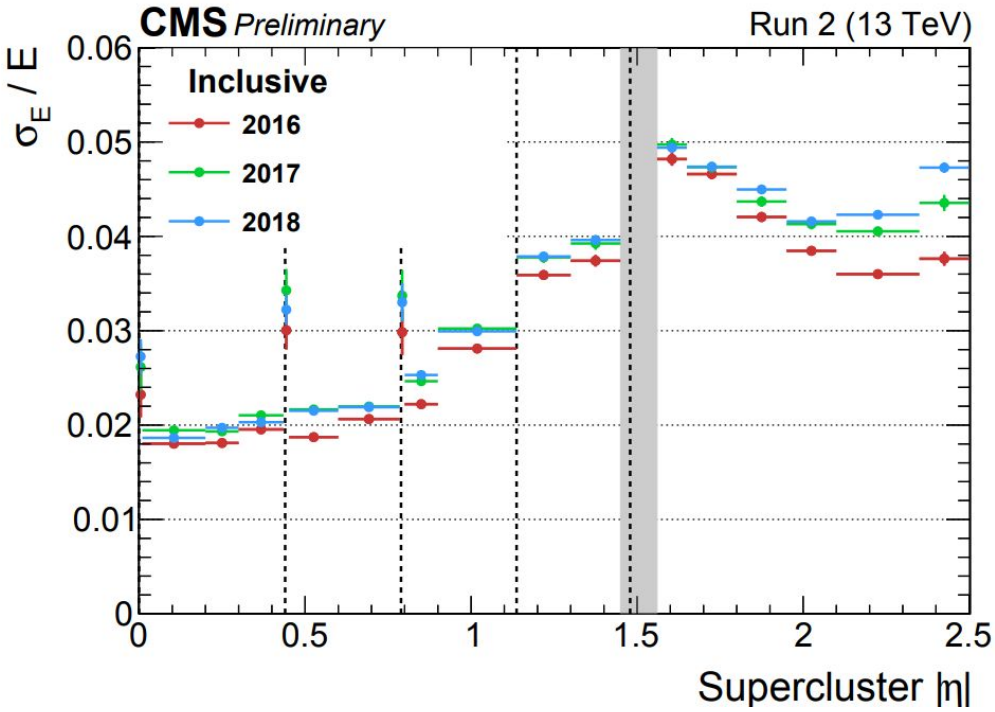


Two dimensional likelihood profile for tH vs ttH in the maximal merging scheme with the largest correlations

κ_F scan



One dimensional likelihood scans in κf, using the resolved κ model and where κV is profiled



Relative electron (ECAL) energy resolution unfolded in bins of pseudorapidity η for the ECAL Barrel and Endcap