HIGGS TO 2 PHOTONS IN CMS: RECENT RESULTS

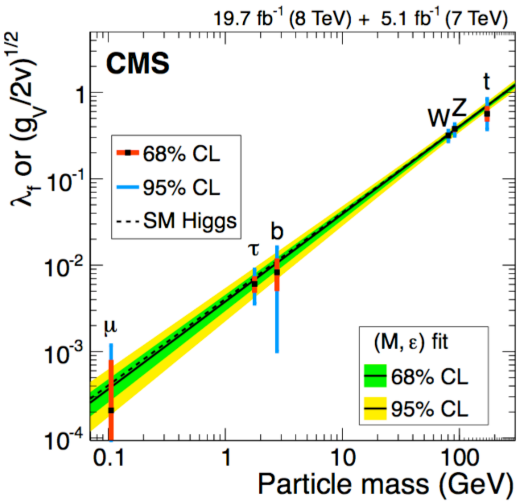
Julie Malclès

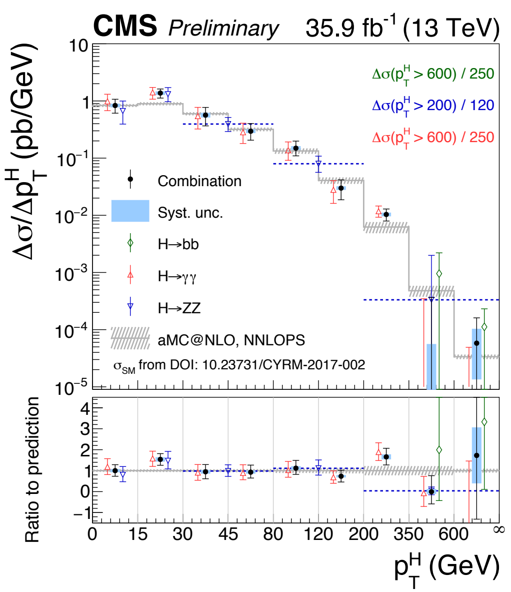
CEA-Saclay, Irfu

On behalf of the CMS collaboration

IRN Terascale, 20-22 May 2019, Annecy

Introduction





After the Higgs discovery during the LHC Run1, the Run2 Higgs analyses aim at measuring precisely its properties: mass, couplings, spin-parity,…

With the increasing statistics, concentrating now:

- on rarer decay or production channels, for example **direct measurements of couplings to fermions**
- **on measurements with increased granularity**

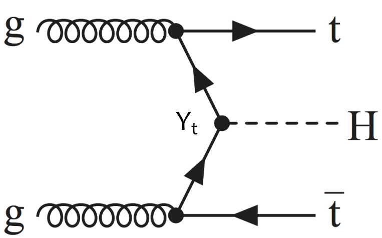
Outline

Two recent results including the 2017 dataset:

Higgs Couplings - November 2018 | PAS-HIG-18-018

2017: 41.5 fb^{-1} + comb with 2016

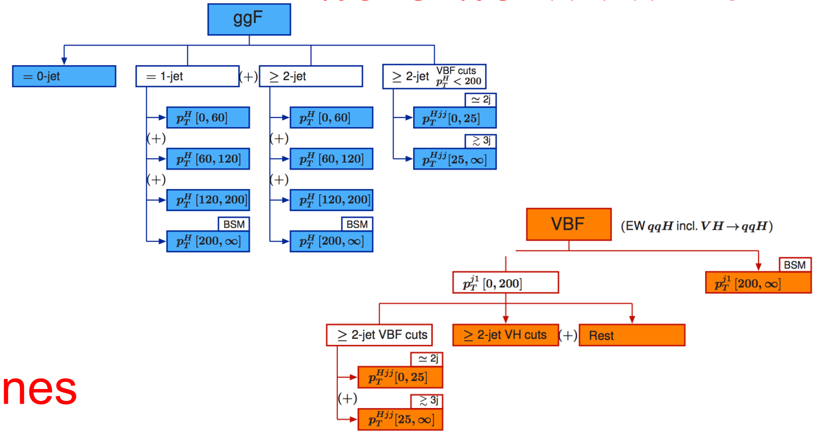
"Measurement of the associated production of a Higgs boson and a pair of top-antitop quarks with the Higgs boson decaying to two photons in proton-proton collisions at $\sqrt{s}$ = 13 TeV"



Moriond EWK – March 2019 | PAS-HIG-18-029

2016+2017: 77.4 fb^{-1}

"Measurements of Higgs boson production via gluon fusion and vector boson fusion in the diphoton decay channel at $\sqrt{s}$ = 13 TeV in the stage 1 simplified template cross-sections framework"

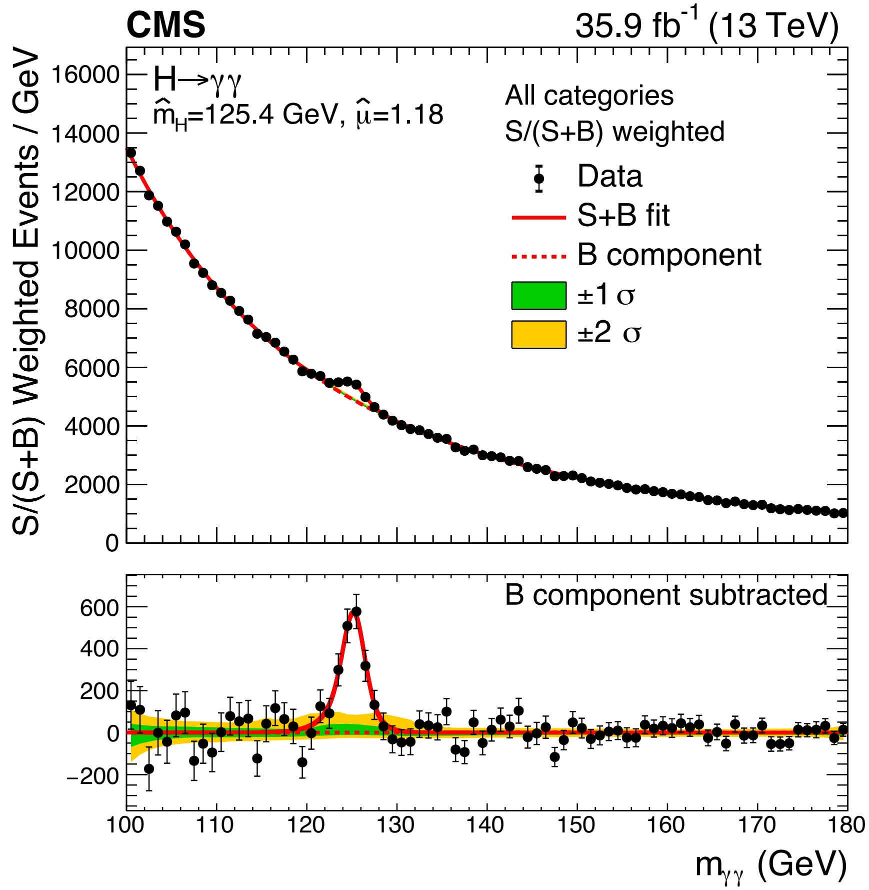


Intermediate results towards full run2 legacy ones

General analysis strategy

- Channel with excellent mass resolution (~1%) with 2 well-reconstructed photons
- Fit small signal peak on top of falling background

$$m_{\gamma\gamma} = \sqrt{2E_1E_2(1-\cos\theta)}$$



1. **Reconstruct photons energies**: excellent energy resolution from lead tungstate crystals in ECAL
2. **Identify vertex and photons** using dedicated MVA discriminating variables
3. **Categorize events** to:
   - target different production modes, with additional objects (jets, leptons,…)
   - improve the sensitivity with dedicated discriminating variables reducing the backgrounds in each category
4. **Perform simultaneous fit to $m_{\gamma\gamma}$** distribution of each category to determine the signal strengths (background fit to the data)
5. Most of systematics uncertainties taken from data/MC comparisons (often using Z)

STXS: introduction

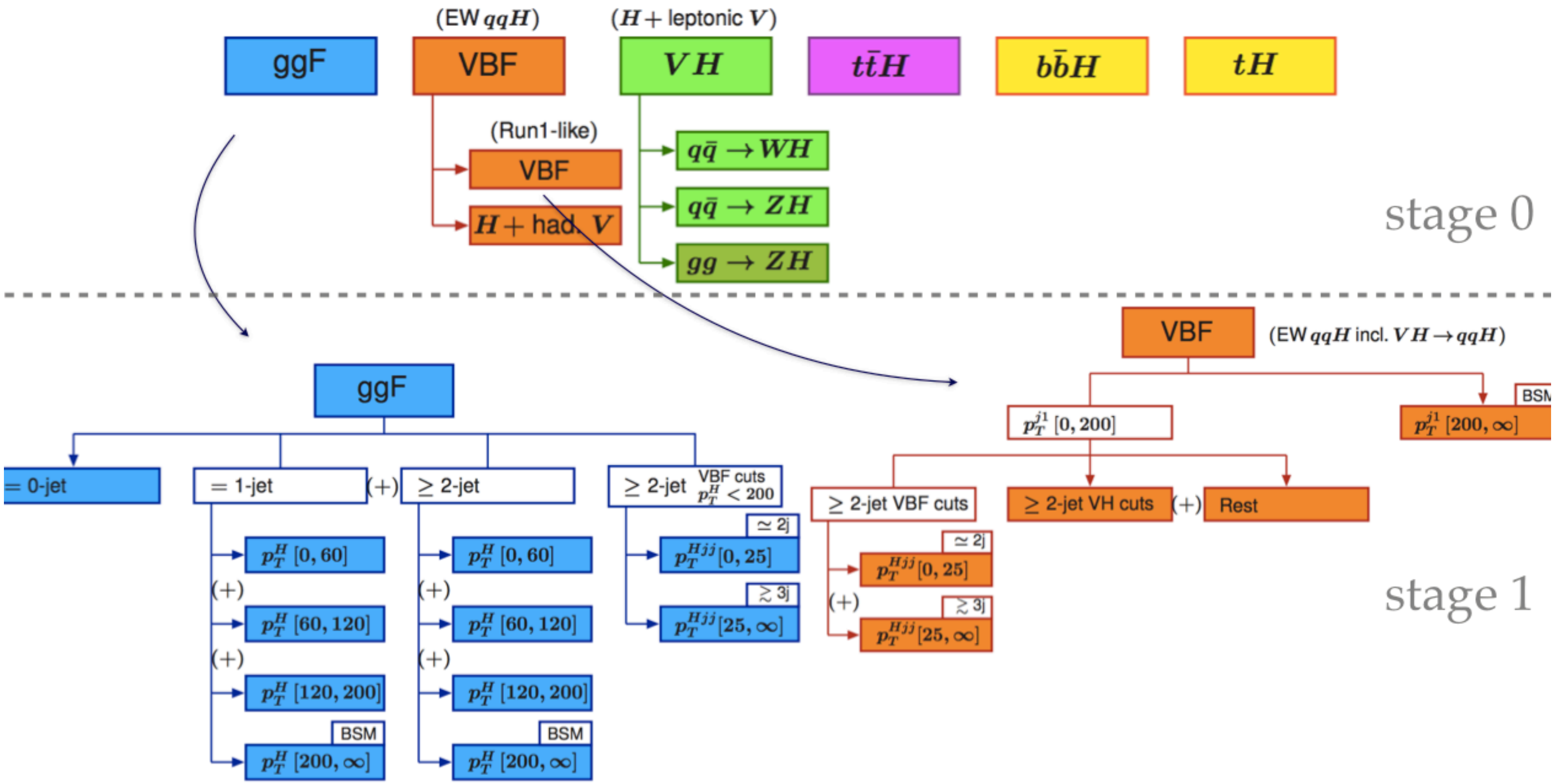
Simplified Template Cross Section (STXS) framework

- STXS framework first outlined in YR4 [1], result of a collaboration between ATLAS, CMS and theorists
- Coherent framework for Higgs measurements, aiming to **maximise experimental sensitivity whilst minimising theory dependence**
- Generator-level kinematic bins based upon the SM production modes are defined, with so-called "stages" increasing in granularity
- Designed to have constant theory uncertainties in each bin, isolating possible BSM effects
- Permits combinations across decay channels and experiments
- The results can be used as inputs to constrain for example EFT parameters

[1] Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector , LHC Higgs Cross Section Working Group, Oct 25, 2016. 849 pp.

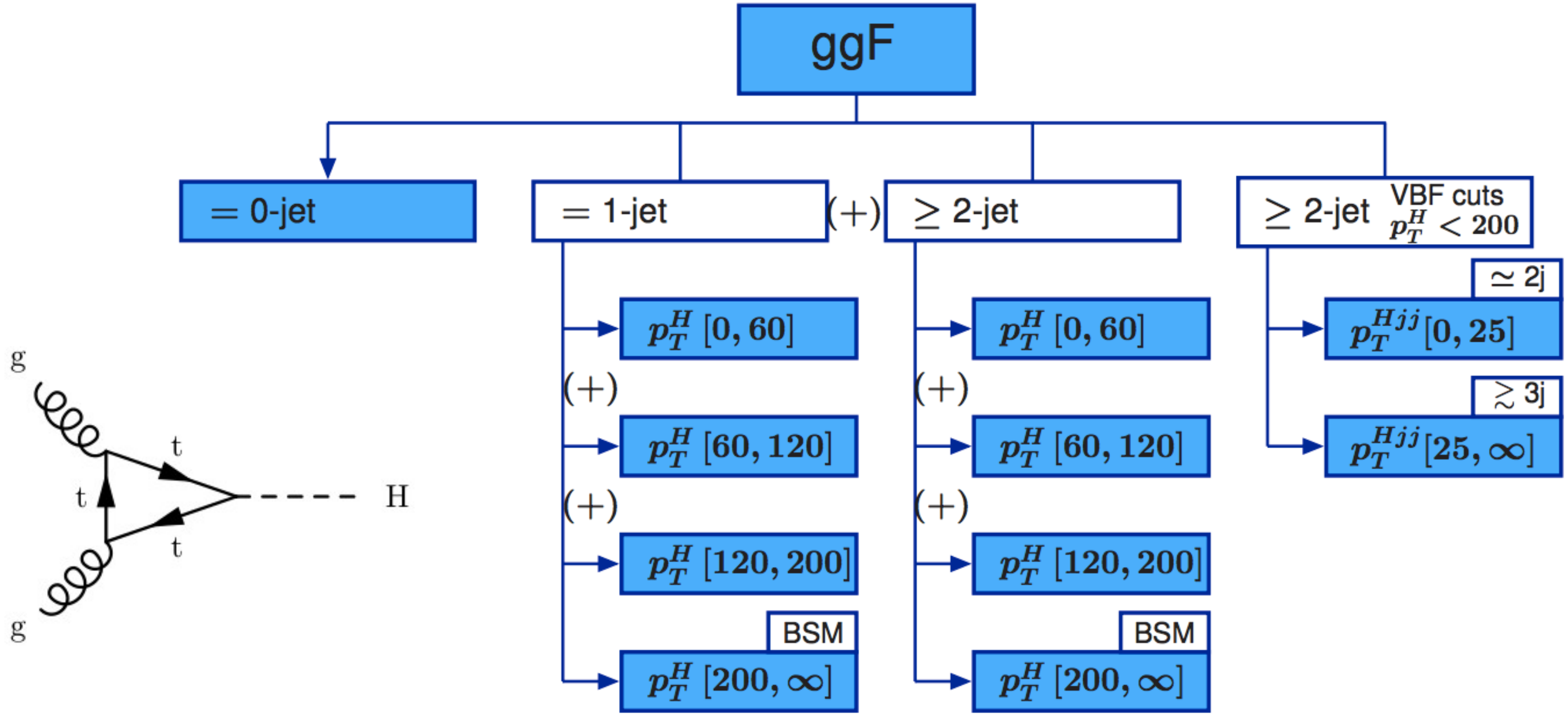
STXS: introduction

- **Stage 0:** **corresponds to the standard SM production processes**



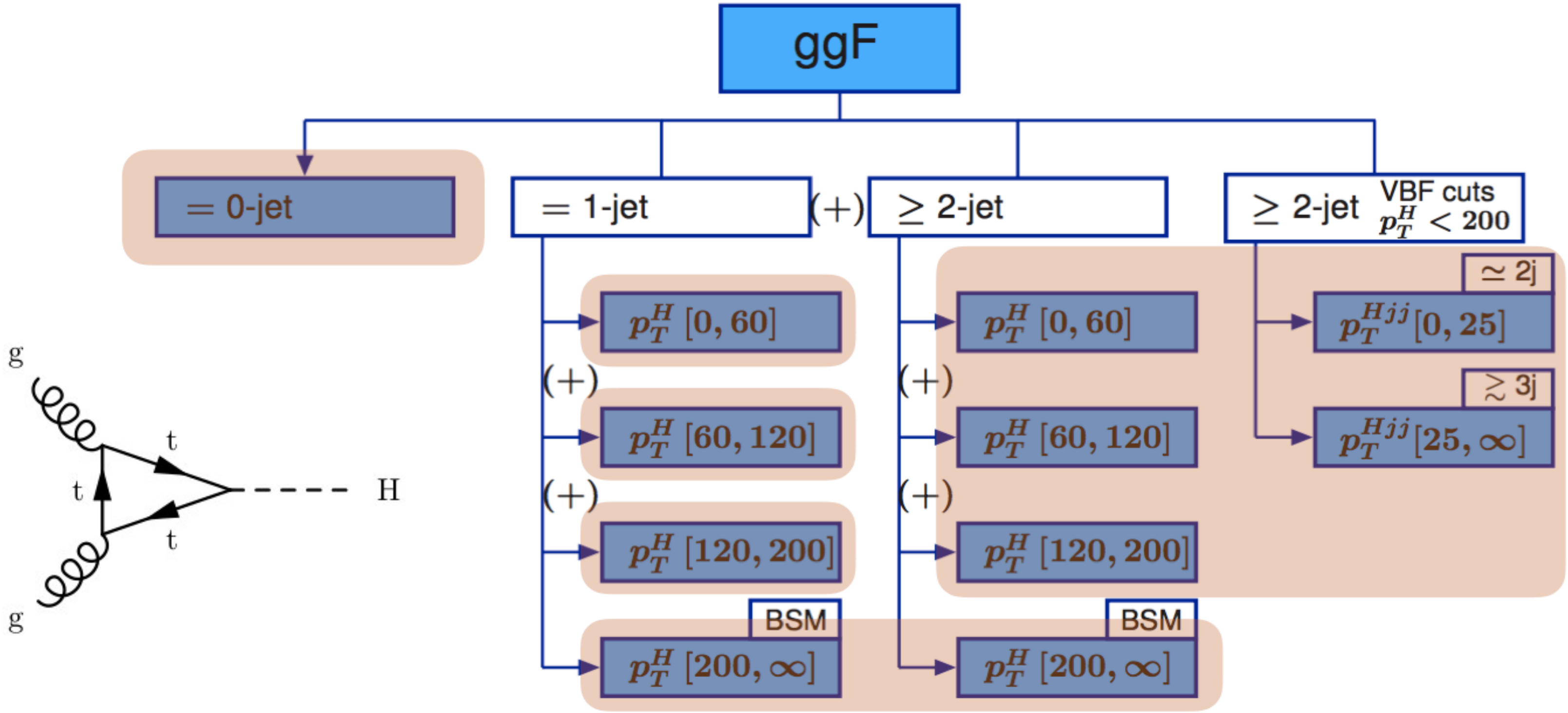
- **Stage 1:** **further splitting based on kinematic properties, e.g. pT(H), nJets**

STXS stage 1: ggH



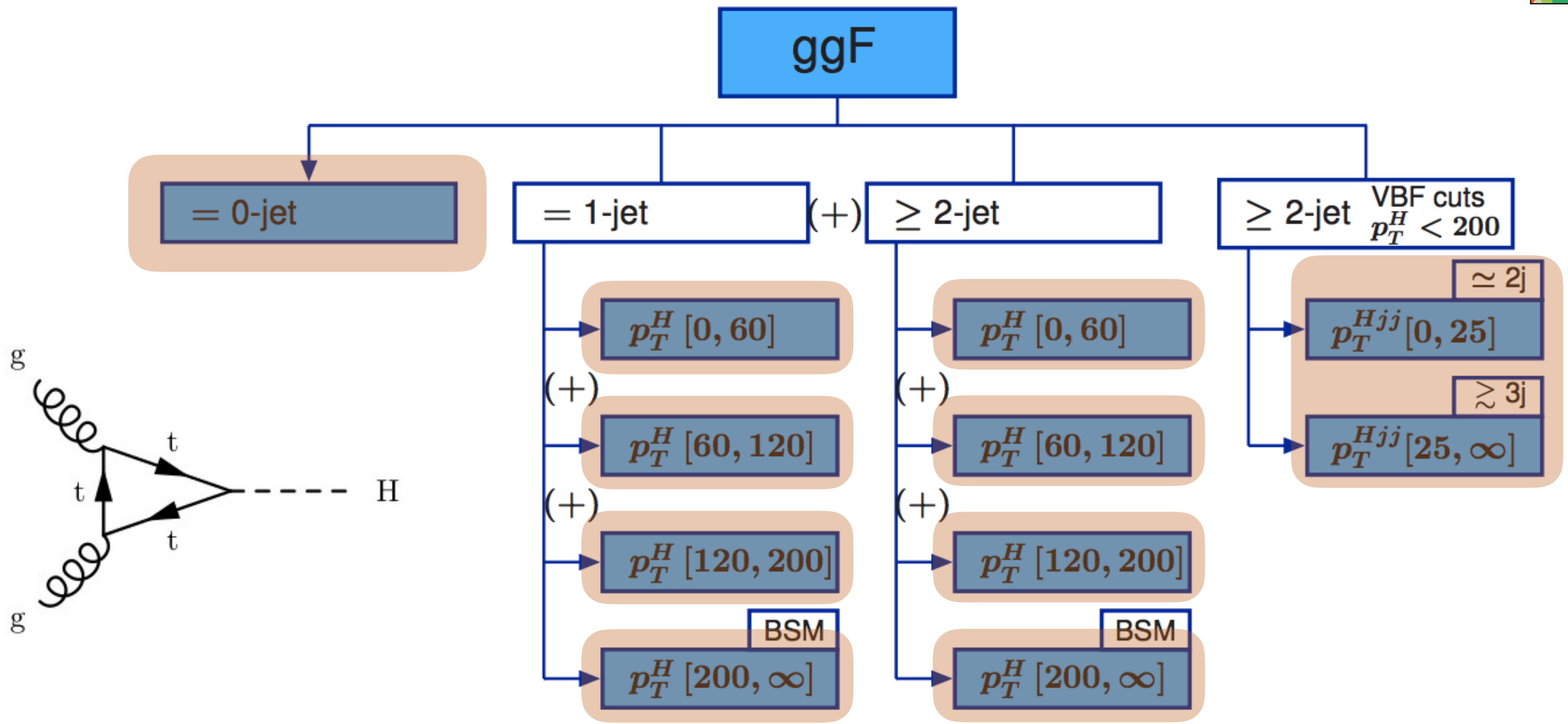
- Eleven generator-level bin definitions
- Split by pT(H) and number of jets (jet pT required to be > 30 GeV)
- Additional VBF-like region with high m_{jj} (> 400 GeV) and Δη (> 2.8)

STXS stage 1: ggH



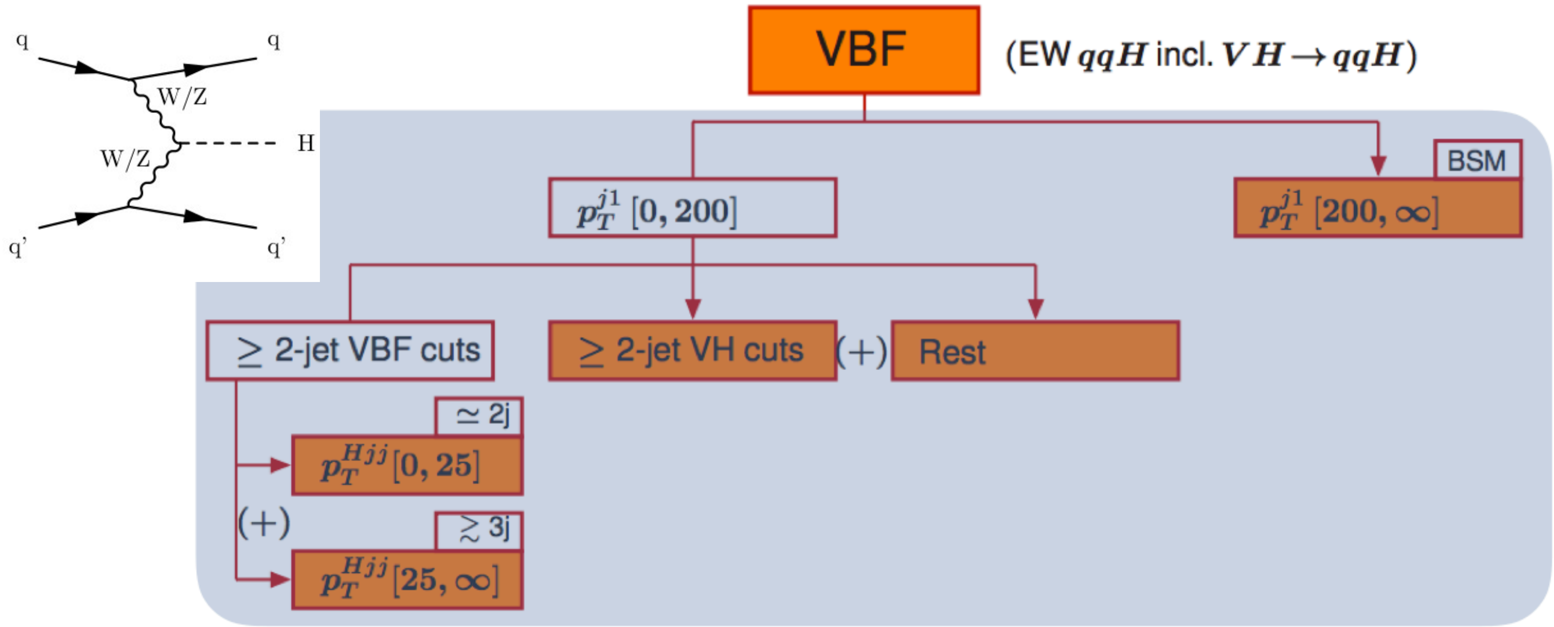
- With the 77.4fb^{-1} (2016+2017) can measure most of the stage 1 bins
- Exceptions: low & medium pT(H) 2J categories and the VBF-like region very difficult to separate the latter from true VBF production
- Statistically limited, in some cases bins need to be merged
- 2 merging scenarios used: **(1) 5 ggH parameters**

STXS stage 1: ggH



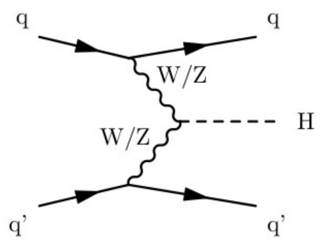
- With the 77.4fb^{-1} (2016+2017) can measure most of the stage 1 bins
- Exceptions: low & medium pT(H) 2J categories and the VBF-like region very difficult to separate the latter from true VBF production
- Statistically limited, in some cases bins need to be merged
- 2 merging scenarios used: **(2) 10 ggH parameters**

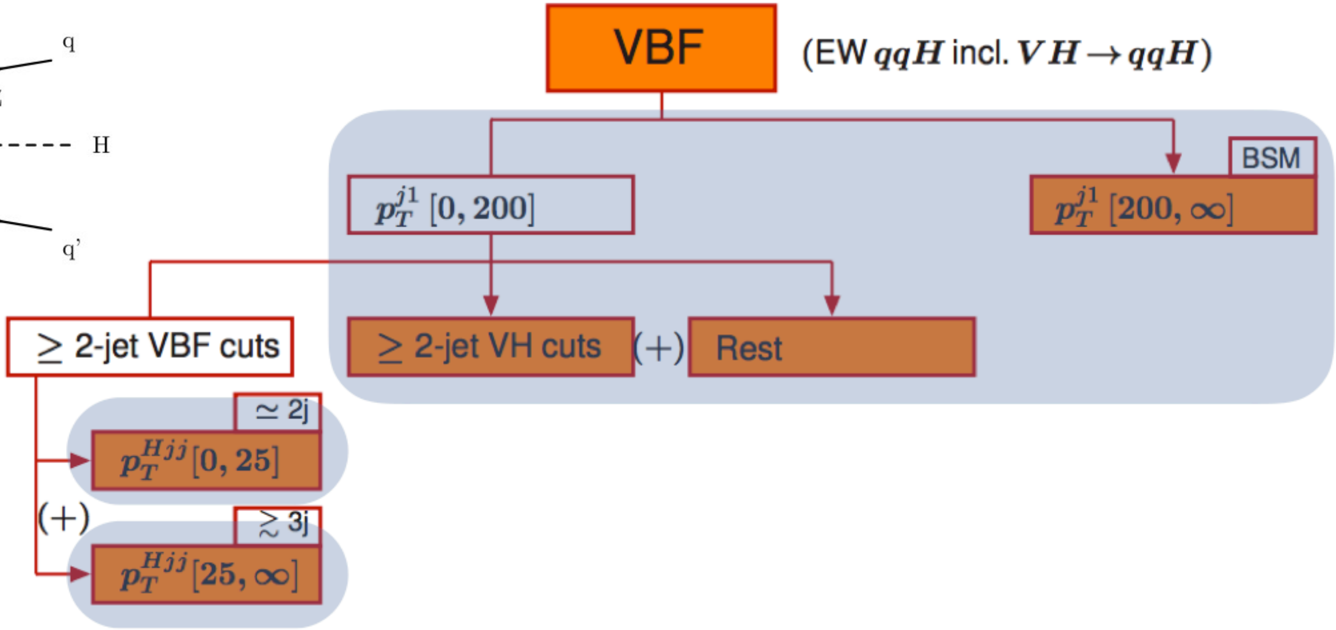
STXS stage 1: VBF



- Two bins defined as the VBF-like bins in the ggH phase space, split into 2J-like and 3J-like with cut on pT(Hjj) (m_{jj} > 400 GeV and Δη > 2.8)
- A BSM bin where lead jet has pT > 200 GeV
- VH bin with 60 < m_{jj} < 120 GeV
- Everything else in "Rest" bin; corresponds to over 60% of signal
- 2 merging scenarios: **(1) 1 single parameter for the 5 bins**

STXS stage 1: VBF





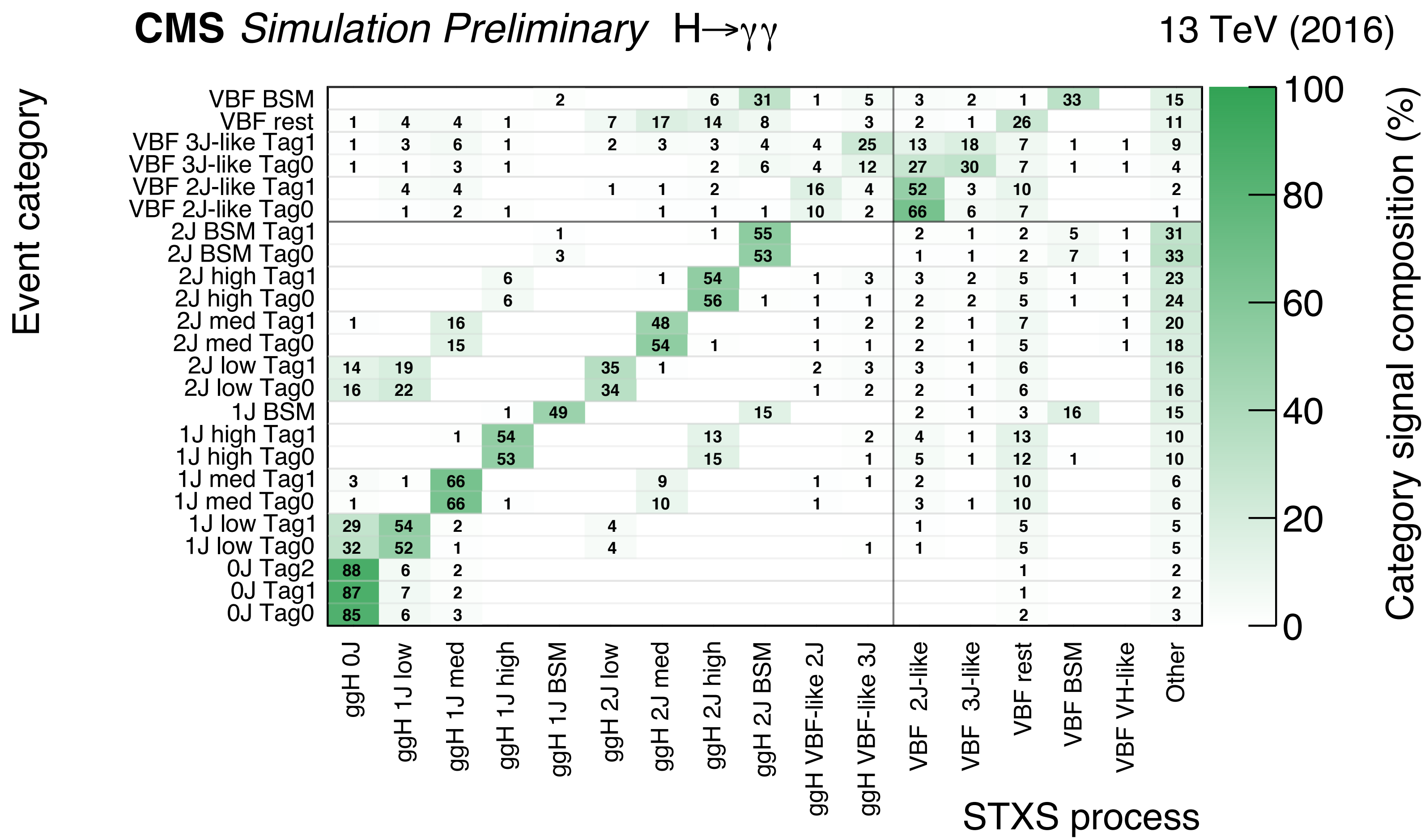
- Two bins defined as the VBF-like bins in the ggH phase space, split into 2J-like and 3J-like with cut on pT(Hjj) (m_{jj} > 400 GeV and $\Delta\eta$ > 2.8)
- A BSM bin where lead jet has pT > 200 GeV
- VH bin with 60 < m_{jj} < 120 GeV
- Everything else in "Rest" bin; corresponds to over 60% of signal
- 2 merging scenarios: **(2) 3 parameters**

STXS: analysis strategy

- Analysis targeting the ggH and VBF phase space regions
- No sensitivity to the stage 1 VH bins, and ttH is not split at stage 1, therefore **do not include ttH or VH dedicated categories**
- Method:
  - define categories targeting the bins, with **cuts on the equivalent reconstructed quantities of the defining generator level variables**
  - reject background using BDTs, with several categories (called “tags”)
- ggH categories:
  - background rejection using the “diphoton BDT” based on **photon kinematics & photon ID BDT**
- VBF categorisation:
  - final categories defined using cuts on “diphoton BDT” and “dijet BDT” **based on jet kinematics (jets pT, m_{jj}, $\Delta\eta_{jj}$, $\Delta\varphi_{jj}$,…)**
  - New: dijet BDT trained on data for backgrounds with non-prompt photons in control regions normalized with appropriate fake factors

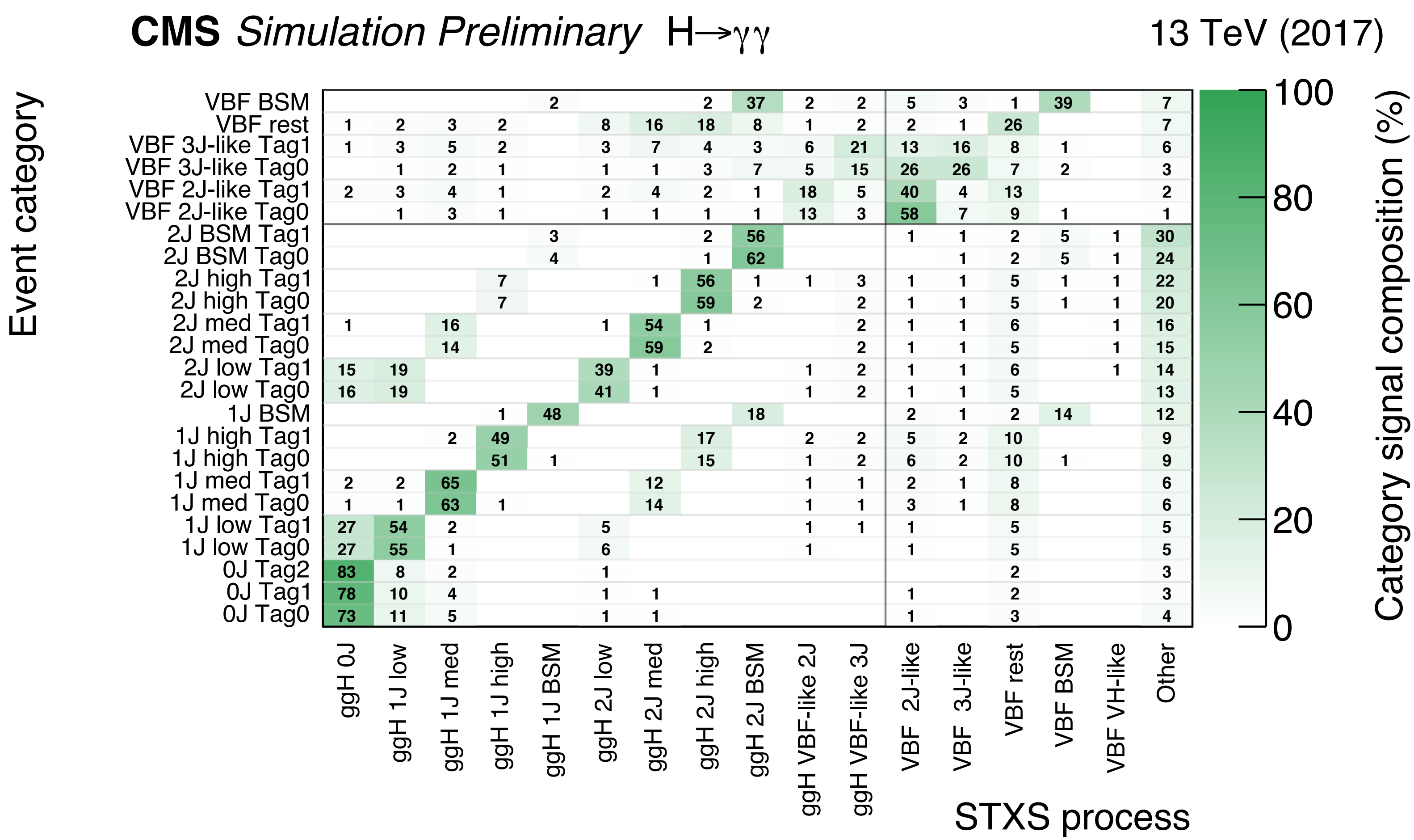
STXS: final categories

Composition of each analysis category in terms of stage 1 bins
Each row sums to 100%

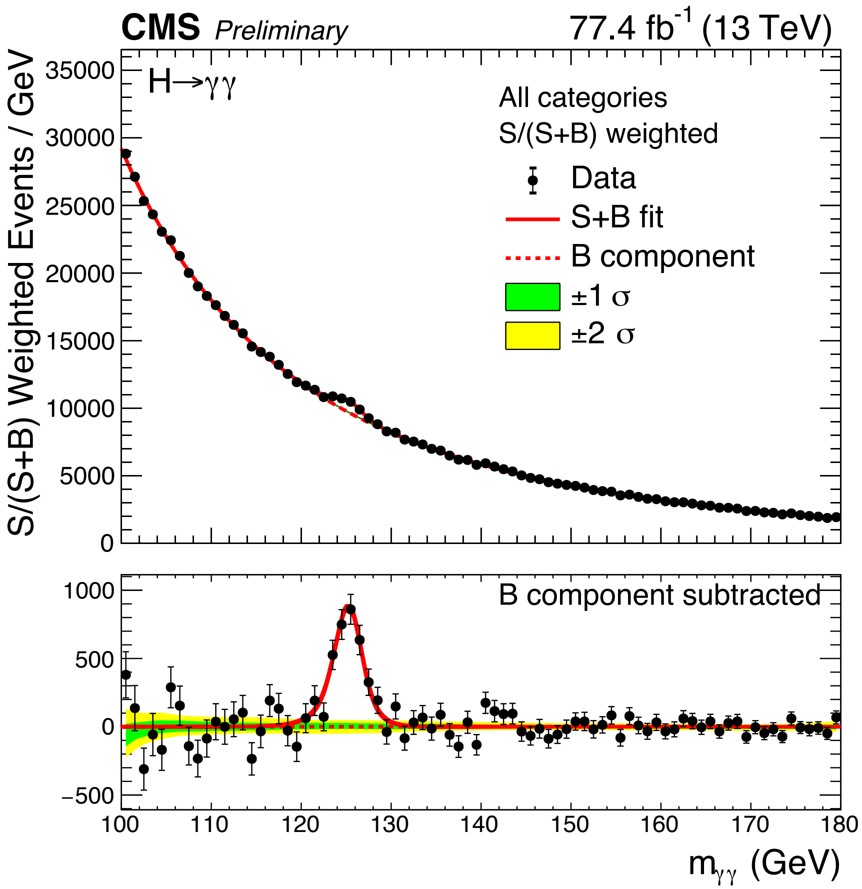


STXS: final categories

Composition of each analysis category in terms of stage 1 bins
Each row sums to 100%



STXS: invariant mass distribution



Diphoton invariant mass fit:

- Simultaneous fit to all categories to determine the free parameters
- Background fit to the data
- Here: all categories included weighted by signal purity
- Very clear peak

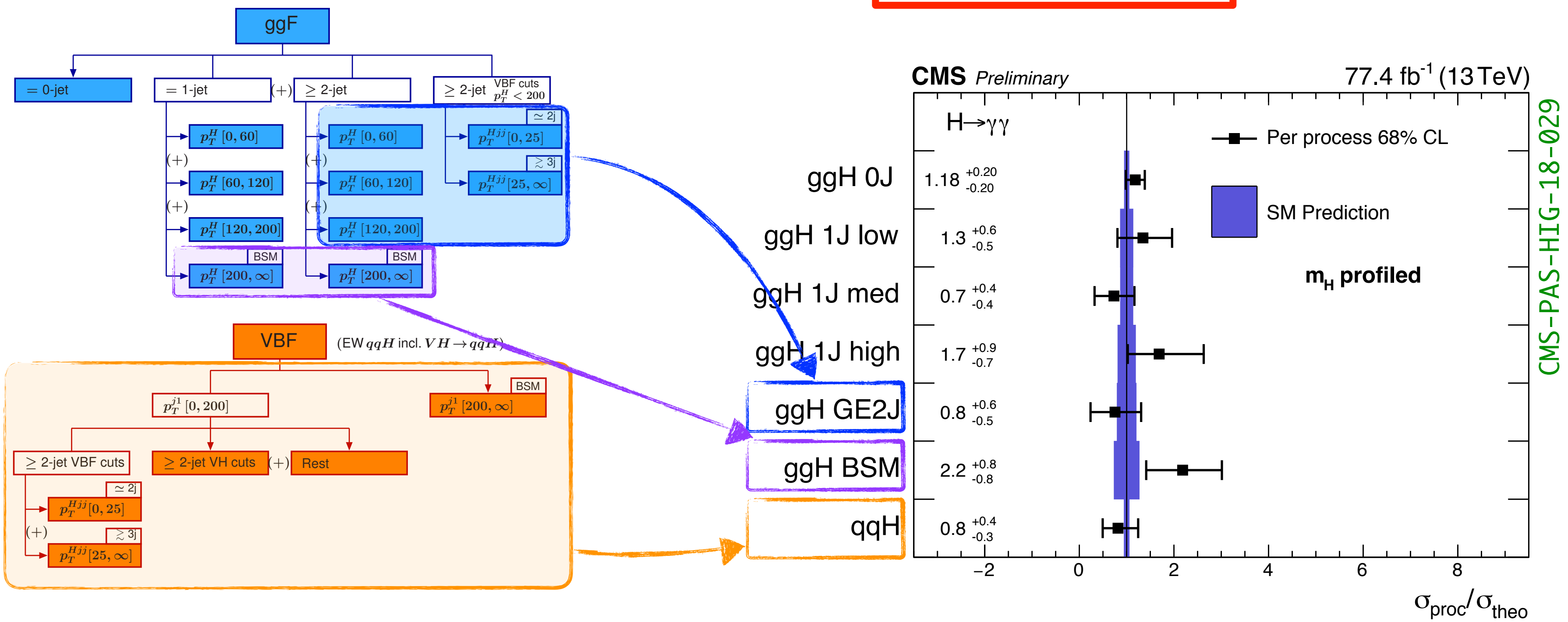
STXS: systematics

- 2 dominant sources of experimental systematics are:
  - Jet energy scale
  - Photon ID systematics

- Theory systematics:
  - **do not include uncertainty on the cross-section itself, considered as an uncertainty on the SM prediction**
  - this differentiates the STXS measurement from a signal strength measurement
  - the effect on the analysis efficiency × acceptance is however included

STXS: results

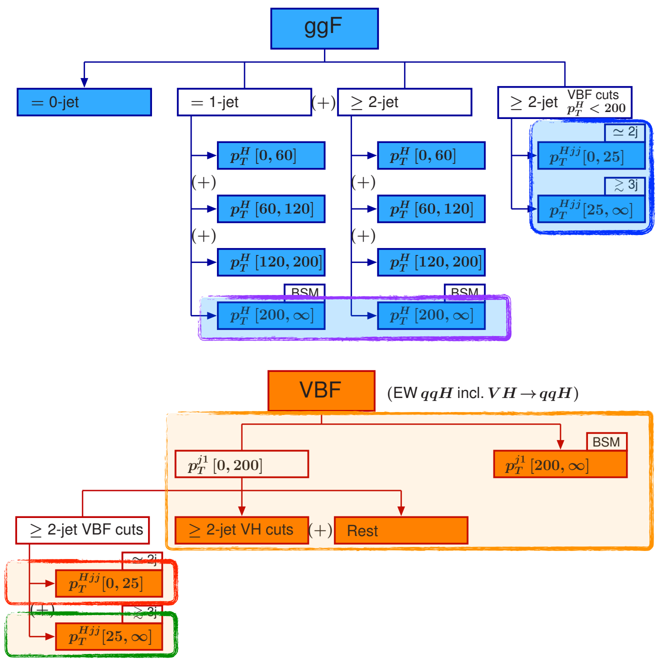
Inclusive σ/σ_{SM}
ggH: 1.15 ± 0.15
VBF: 0.8 $^{+0.4}_{-0.3}$

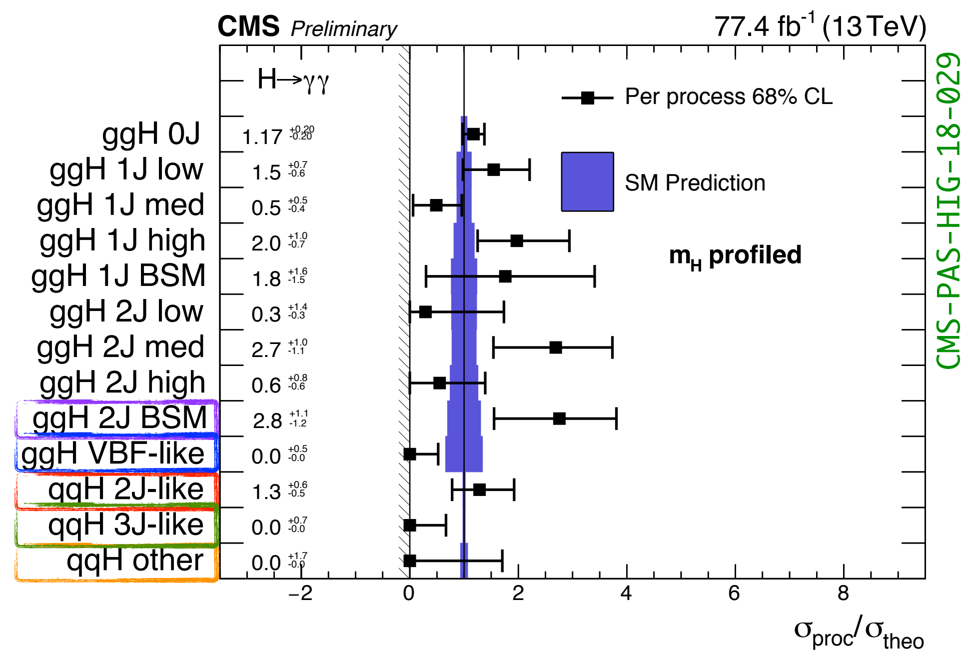


- **First scenario:** 7 parameter fit
- Good agreement with SM prediction, including in BSM bins
- Large statistical uncertainties

STXS: results

Inclusive σ/σ_{SM}
ggH: 1.15 ± 0.15
VBF: $0.8^{+0.4}_{-0.3}$



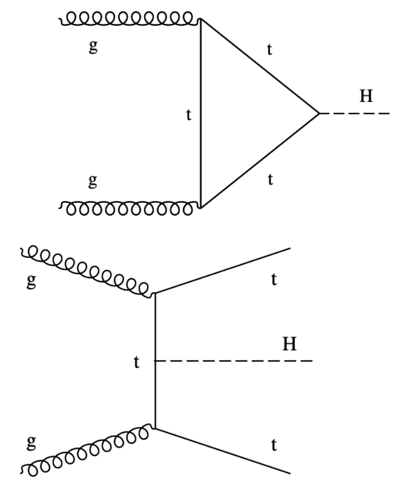


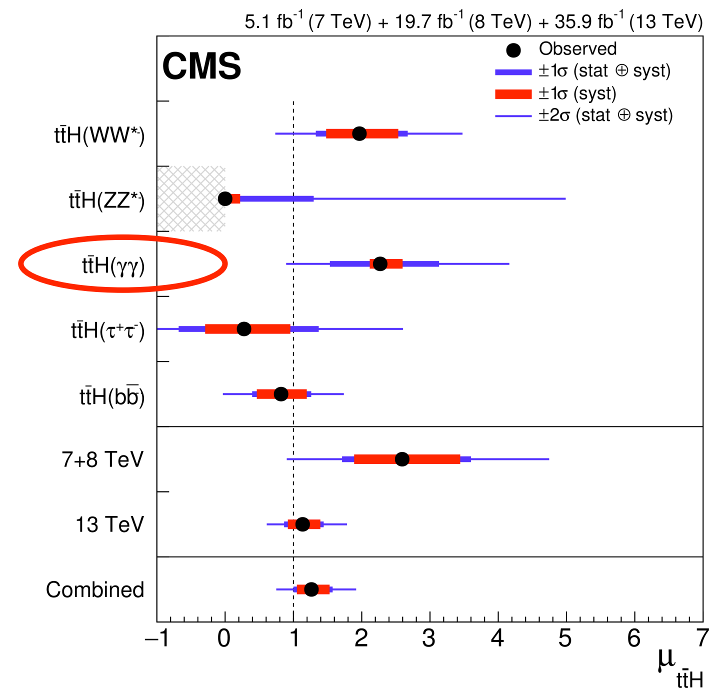
- **Second scenario:** 13 parameter fit

- Good agreement with SM prediction, including in BSM bins
- Very large statistical uncertainties

ttH: introduction

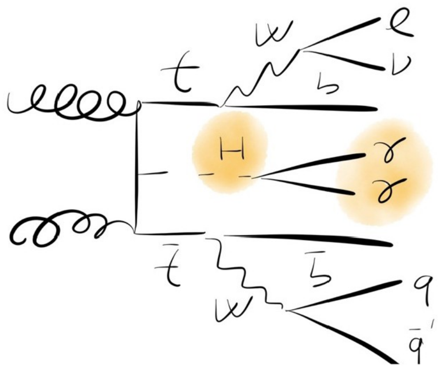
- Since its discovery in 2012: major effort is to study the newly discovered Higgs boson
  - Couplings are one of the fundamental parameters
  - In the SM, Higgs boson's couplings are unambiguously predicted
  - Until recently, only indirect access to H*ff* couplings
- Coupling y_t to top quark is important
  - Top quark has the strongest SM coupling (y_t~1)
  - Can be inferred from contribution to loops in ggH and Hγγ, but deviations from the SM can be masked by other new phenomena
  - Direct ttH production provides direct measurement
- **ttH production discovery** reported by ATLAS and CMS in June 2018. Without 2017 data for the diphoton channel in CMS.
- **2017 diphoton result reported today**



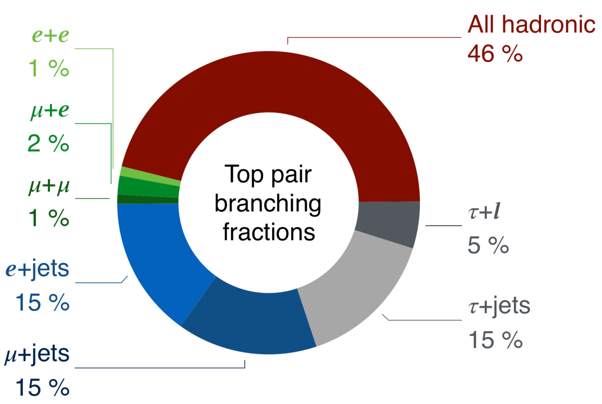


ttH: analysis strategy

- Analysis strategy very similar to STXS one
  - H identified selecting two high pT isolated photons
  - tops identified with decay products



- Exclusive categories depending on $N_{leptons}$
  - Hadronic categories: 0 lepton (e or μ)
  - Leptonic categories: more than 0 lepton



- Background rejected using dedicated BDTs

- Simultaneous fit of mass distributions to determine ttH signal strength

- Combination with existing 2016 result

ttH: improvements

Changes and improvements with regard to 2016 analysis:

- Preselections applied to train in the regions of interest
- Reweighting procedure to improve data/simulation agreement for backgrounds before to train
- MVA hyper-parameters optimized with regard to the expected sensitivity
- Number of categories determined to have improvement at least of a few percents on the sensitivity by adding a category
- **More categories:** 3 hadronic and 2 leptonic (compared to 1 and 1 respectively in 2016)
- ttH hadronic: **more discriminating variables** added (~30 versus 4)
- ttH leptonic: **uses now a BDT (cut-based before)**

Overall the improvement in sensitivity with regard to the 2016 result is **about 40%**

ttH: discriminating variables

Discriminating variables: BDTs including informations from photons, jets, and leptons kinematics, b-tagging information and missing E_T

ttH Hadronic:

- $n_{lepton} = 0$, photon ID > -0.2, $n_{Jets} > 1$
- The following variables are used for the MVA training:
  - **Photon variables:**
    - $p_T/m_{\gamma\gamma}$ of two photons,
    - $|\eta|$ and ϕ of two photons,
    - photon-ID of two photons,
    - pixel seed veto decision of two photons,
    - $p_T/m_{\gamma\gamma}$ of diphoton,
    - the rapidity of diphoton.
  - **Jet variables:**
    - number of jets,
    - p_T, $|\eta|$, and ϕ of four leading jets,
    - H_T.
  - **b-tagged jet variables:**
    - b-tag value of the highest three b tag scored jets,
    - b-tag value of four leading jets.
  - **Missing p_T.**

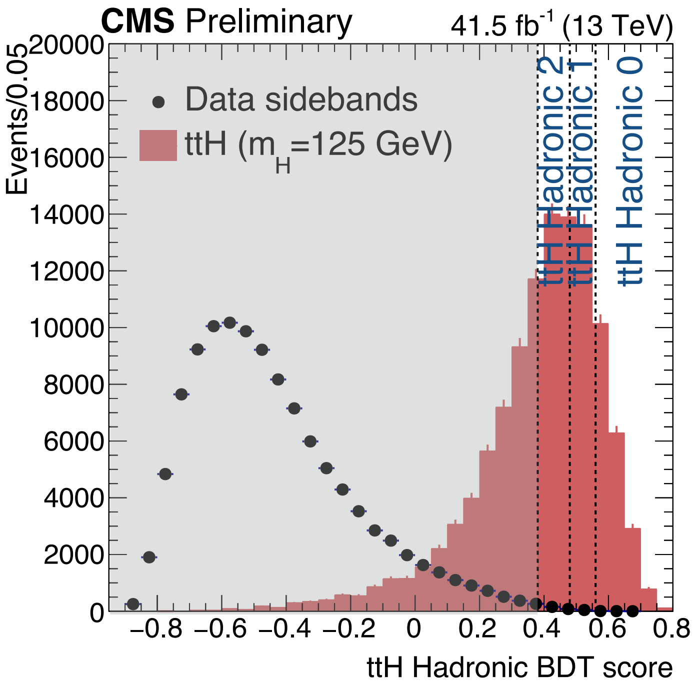
ttH Leptonic:

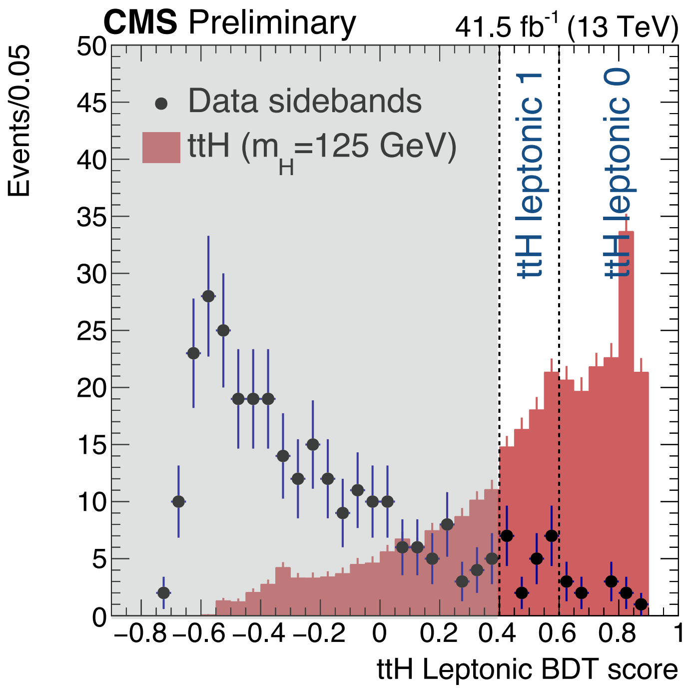
- $n_{lepton} > 0$, photon ID > -0.2, $n_{Jets} > 0$ ($n_{b\ Jets} > 0$)
- The following variables are used for the MVA training:
  - **Photon variables:**
    - $p_T/m_{\gamma\gamma}$ of two photons,
    - $|\eta|$ of two photons
    - $\Delta\phi$ between two photons,
    - photon-ID of two photons,
    - pixel seed veto decision of two photons,
  - **Lepton variables:**
    - p_T and $|\eta|$ of the highest p_T lepton
  - **Jet variables:**
    - number of jets,
    - p_T, $|\eta|$, and ϕ of three leading jets,
  - **b-tagged jet variables:**
    - b-tag value of the highest two b tag scored jets,
    - number of b tagged jets.
  - **Missing p_T.**

ttH: categorisation

Preselection:
$N_{jets} \geq 2$ with b-tagging≥0,
$N_{leptons} = 0$, photonIDs>-0.2

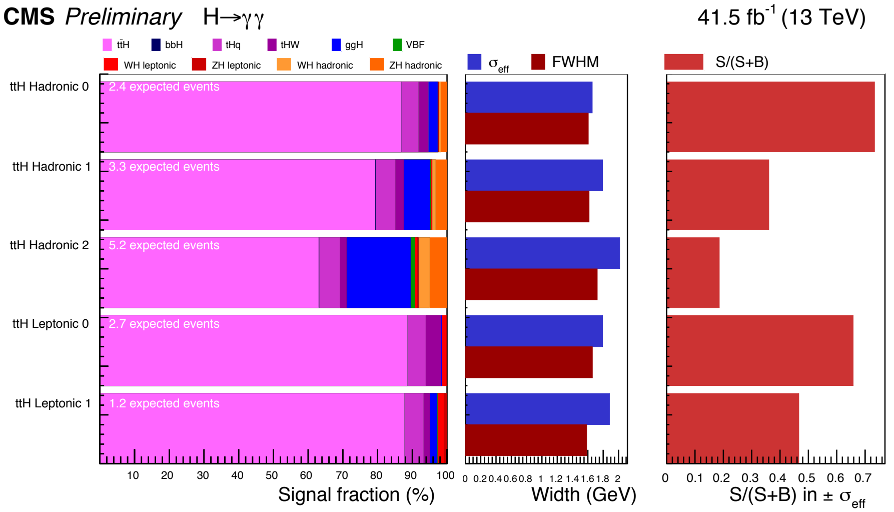
Preselection:
$N_{jets} \geq 1$ with b-tagging≥0,
$N_{leptons} \geq 1$, photonIDs>-0.2

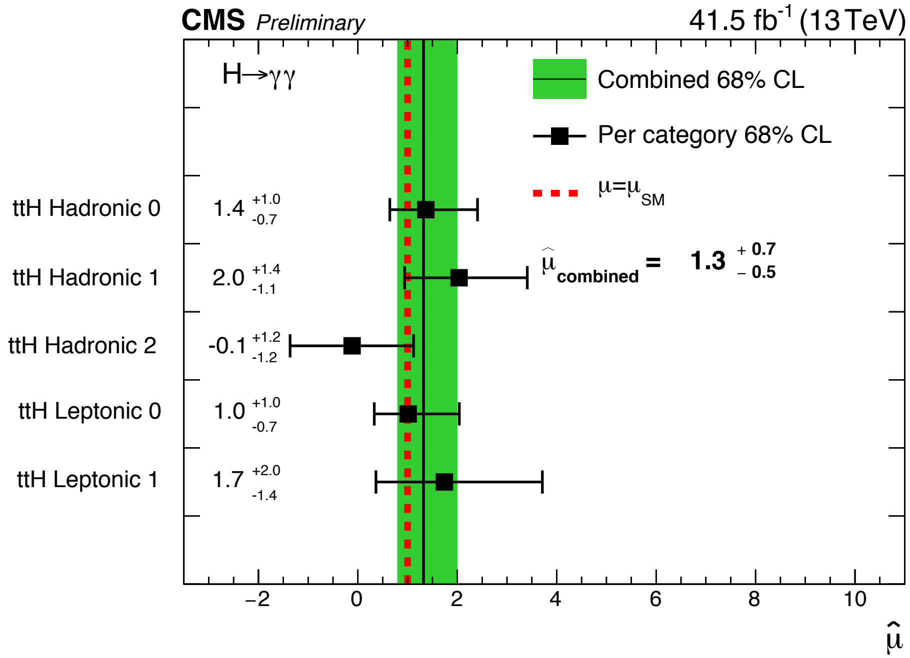




- These discriminating variables are used to define optimal categories
- Boundaries chosen to maximize the expected sensitivity to the ttH production
- Categories driving sensitivity: ttH hadronic 0 and ttH leptonic 0

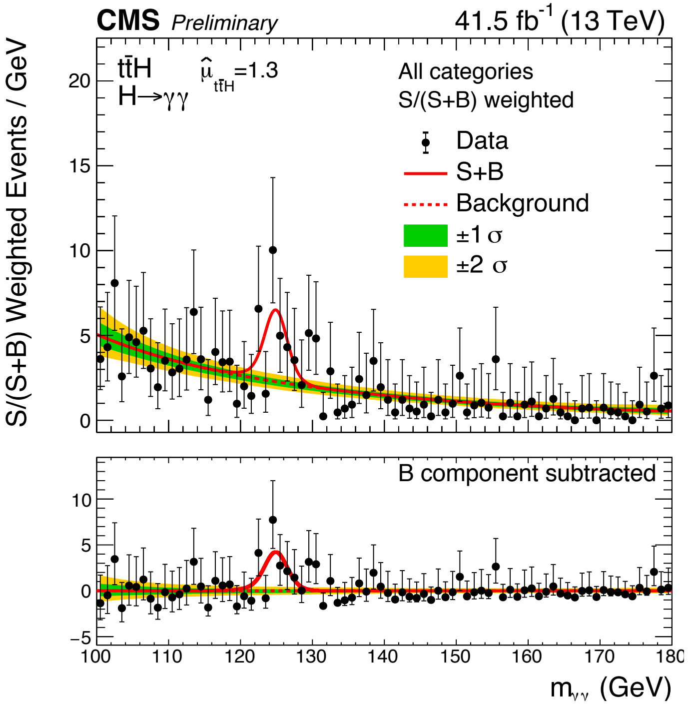
ttH: results on 2017 dataset

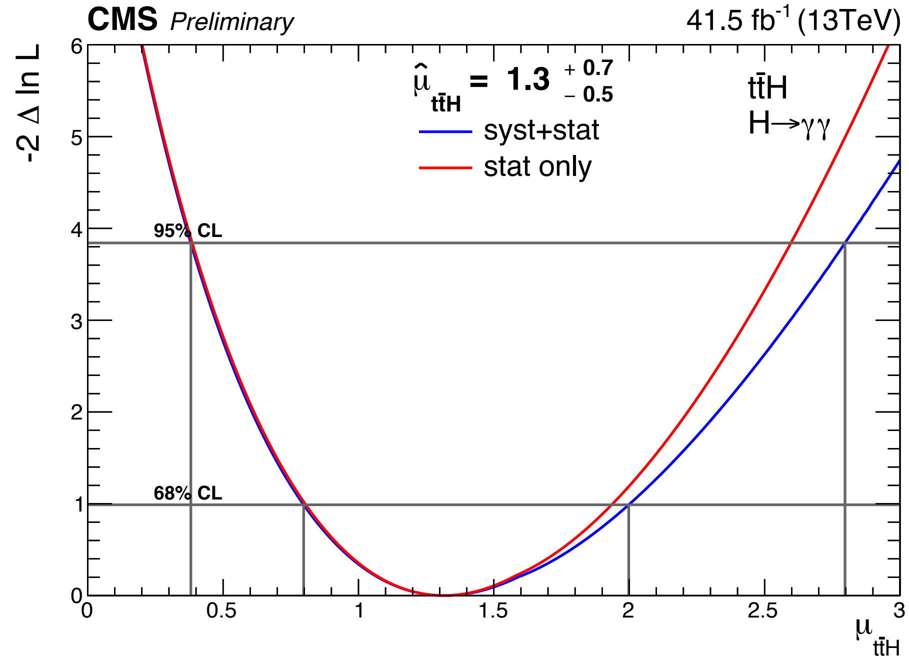




- A handful of events
- Categories driving sensitivity: ttH hadronic 0 and ttH leptonic 0
- Rather pure categories with regard to other production modes
- Large uncertainties dominated by statistical uncertainties
- Good agreement with the SM

ttH: results on 2017 dataset



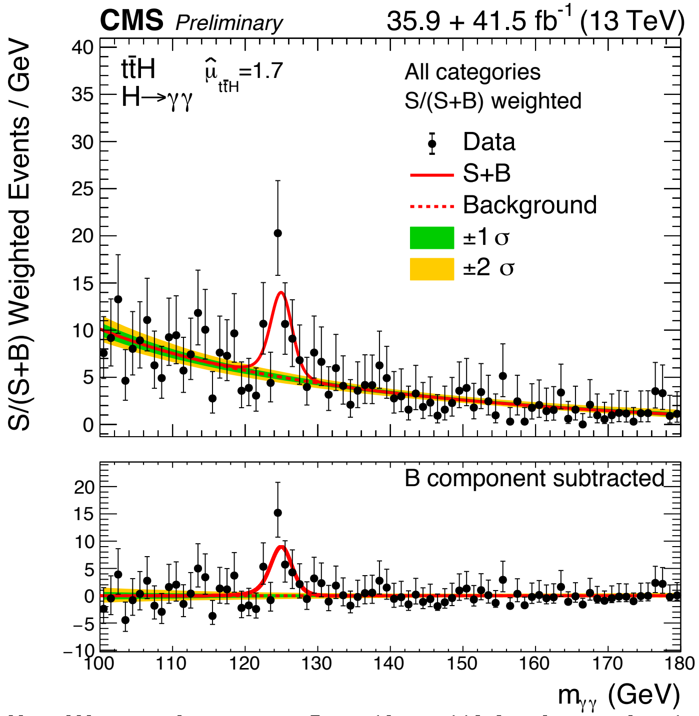


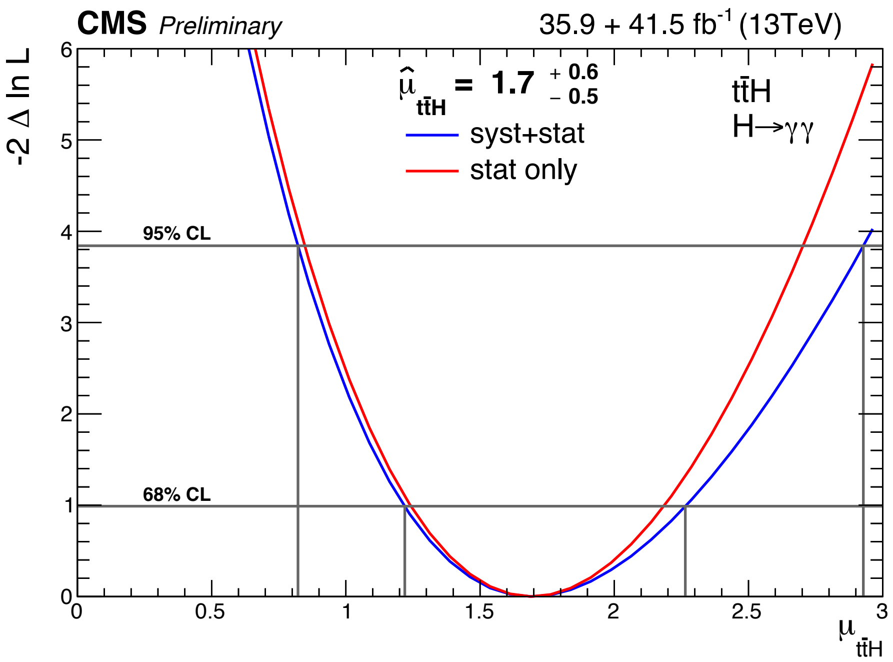
- Likelihood scan for the ttH signal strength where the mass is constrained to the combined run I value (125.09 ± 0.24 GeV)

$$\mu_{ttH} = 1.3^{+0.7}_{-0.5} = 1.3^{+0.6}_{-0.5}(\text{stat.})^{+0.3}_{-0.1}(\text{syst.})$$

- Expected significance: 2.2σ, observed significance: 3.1σ

ttH: combination with 2016





- Likelihood scan for the ttH signal strength where the mass is constrained to the combined run I value (125.09 ± 0.24 GeV)

$$\mu_{ttH} = 1.7^{+0.6}_{-0.5}$$

- Expected significance: 2.7σ, observed significance: 4.1σ

Conclusion

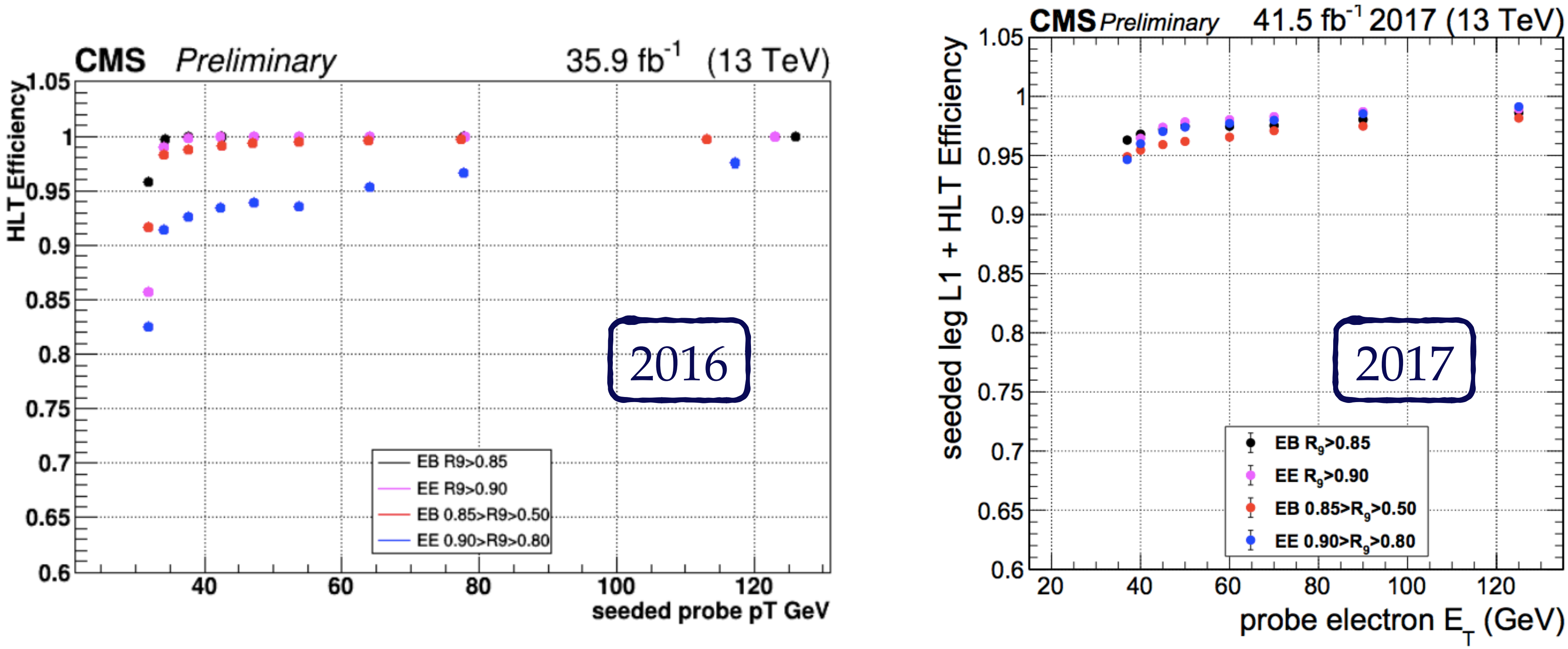
- Intermediate results on 2017 dataset:
  - First CMS result within the STXS stage 1 framework in the diphoton channel
  - Largely improved sensitivity to the ttH production with regard to 2016 results
  - Both results largely statistically limited

- Towards full run2 legacy results:
  - 2018 data analysis on-going
  - Re-analysis of 2016+17 data with several improvements (calibration, strategy)
  - Move to stage 1.1 for the STXS results

Stay tuned!

Backups

Trigger efficiencies



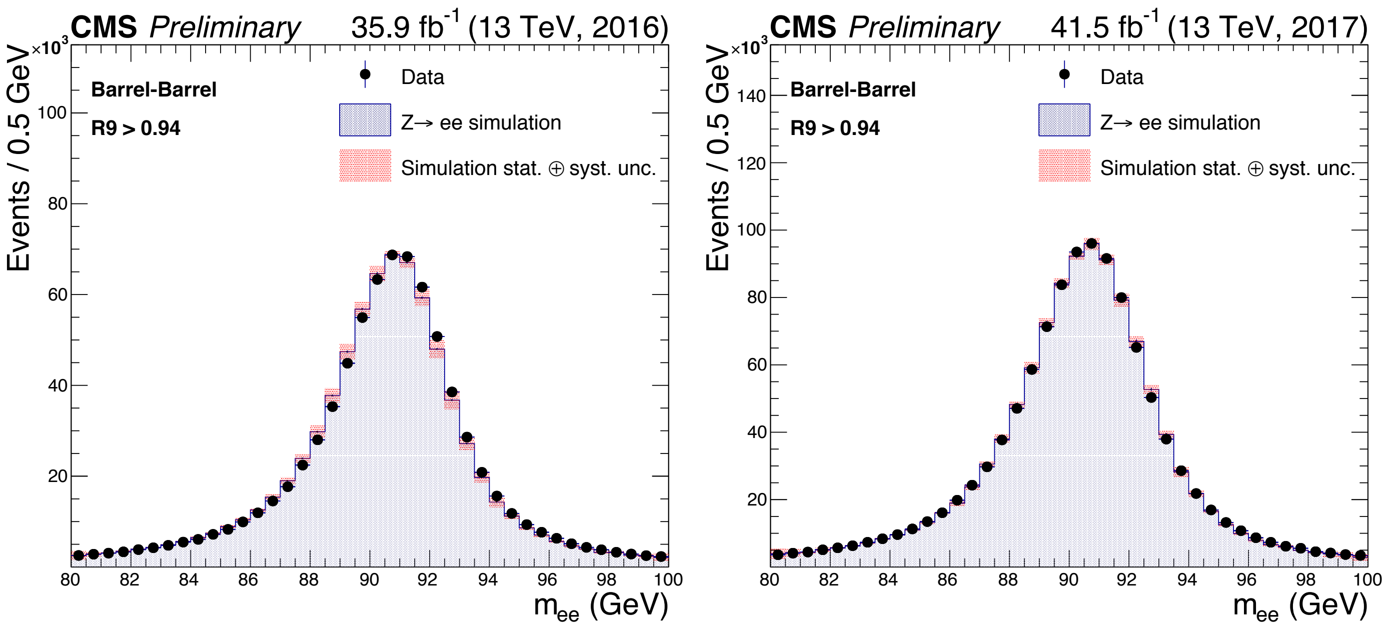
- Diphoton HLT trigger with asymmetric pT thresholds is used
- Trigger efficiency calculated using the tag & probe method on $Z \rightarrow ee$ events

Preselection

- Standard loose selection which is slightly tighter than HLT: photon pair with 100 < $m_{\gamma\gamma}$ < 180 GeV where photons are within ECAL acceptance and not in the gap
- electron veto applied
- absolute and scaled pT cuts also applied
  - 2016: pT > 30 (25) GeV and pT/$m_{\gamma\gamma}$ > 1/3 (1/4)
  - 2017: pT > 35 (25) GeV and pT/$m_{\gamma\gamma}$ > 1/3 (1/4)
- along with shower shape and isolation:

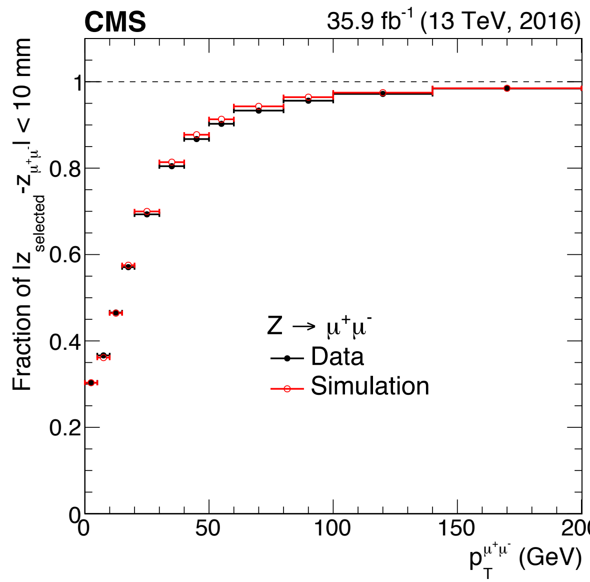
	R_9	H/E	$\sigma_{\eta\eta}$	$\mathcal{I}_{ph}$ (GeV)	$\mathcal{I}_{tk}$ (GeV)
Barrel	> 0.85	<0.08	-	-	-
	[0.50, 0.85]	<0.08	<0.015	< 4.0	< 6.0
Endcaps	> 0.90	<0.08	-	-	-
	[0.80, 0.90]	<0.08	<0.035	< 4.0	< 6.0

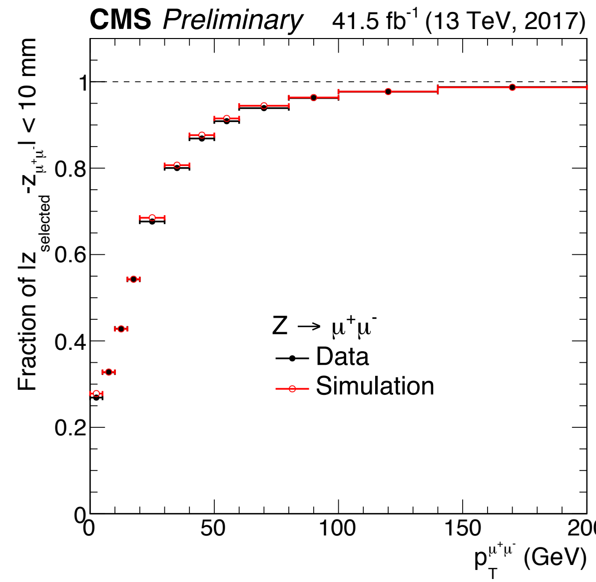
Photon energy



- Corrected for the imperfect containment of the shower and the energy losses from converted photons (MVA regression)
- Data energy scale corrected to match Z peak, in bins of η and R9
- Simulation smearing adjusted to match data using Z

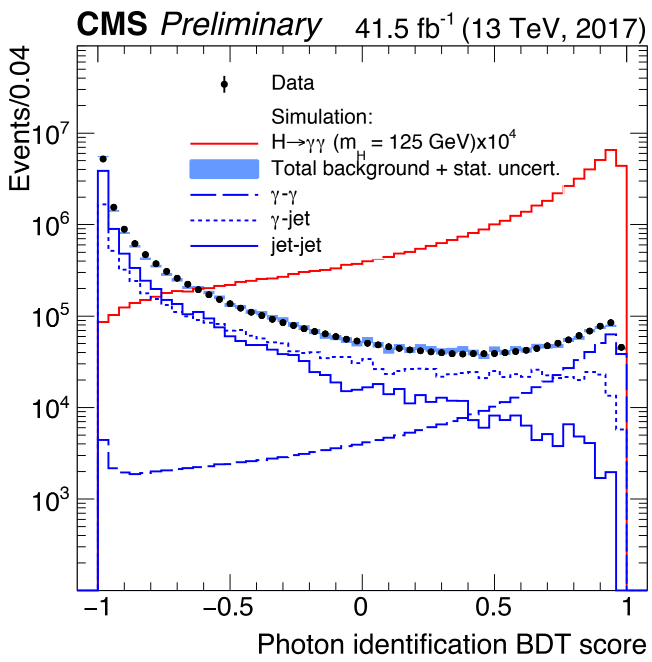
Vertex identification

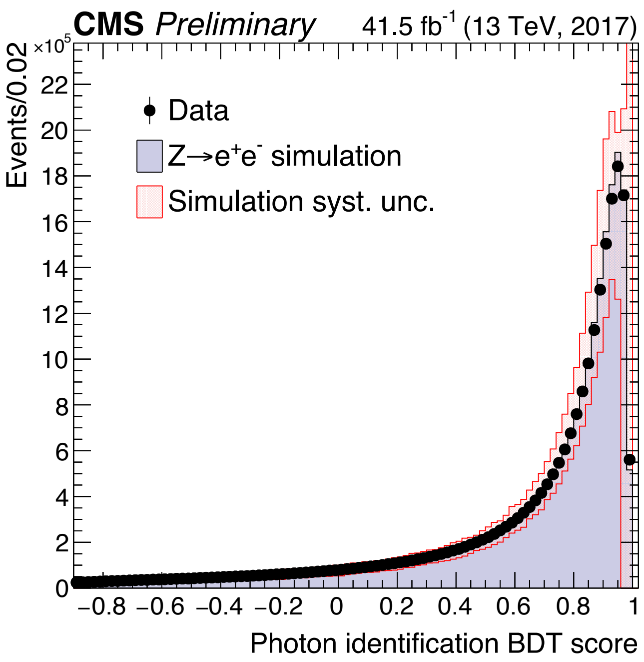




- In absence of tracks, the vertex is unknown
- Important for maintaining the mass resolution, **particularly for ggH events**
- If chosen vertex within ~1cm of the true vertex, negligible impact on resolution
- MVAs to identify the diphoton vertex using recoiling tracks and their balance with the Higgs boson p_T
- Efficiency to choose the vertex within 1cm ~ 80%
- Validated with Z➔μμ events where tracks are removed to mimic a diphoton

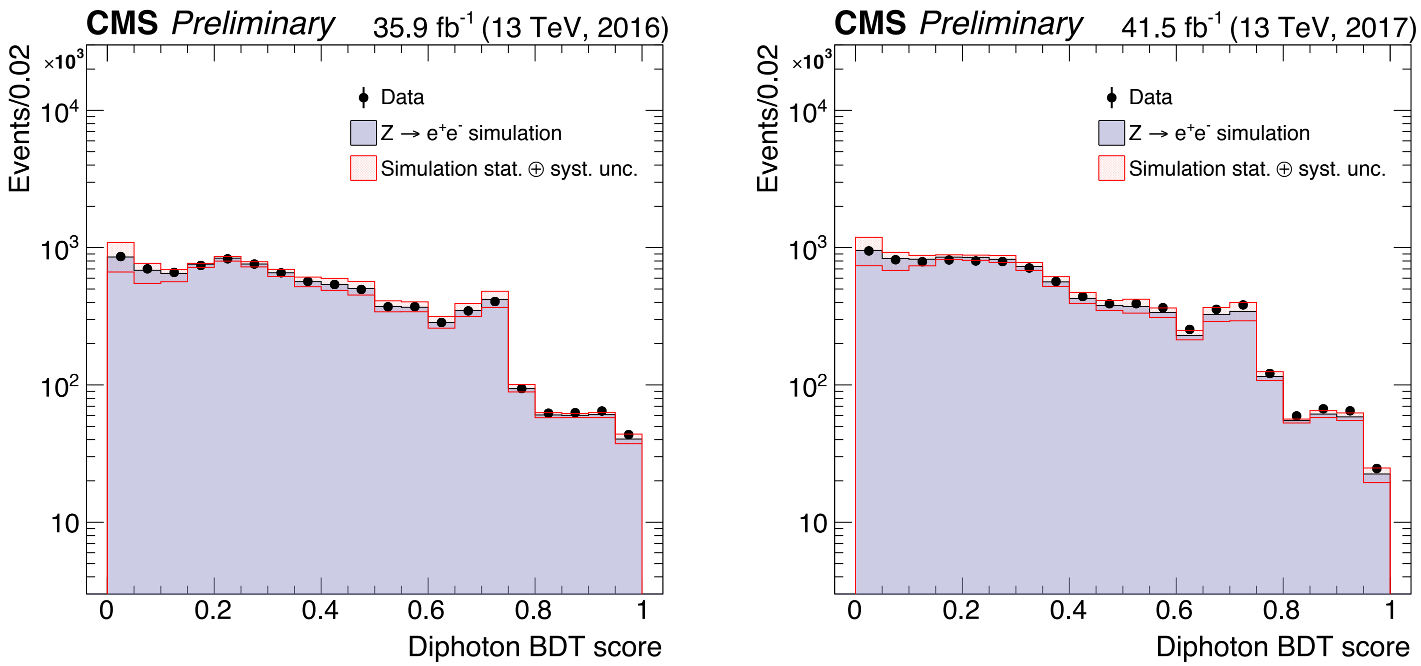
Photon identification





- Photon identification: BDT using shower shape, isolation and kinematic variables to distinguish between prompt and fake photons from neutral hadrons
- Validated in data Z➔ee and Z➔μμγ events
- Systematics assigned to cover the observed discrepancies

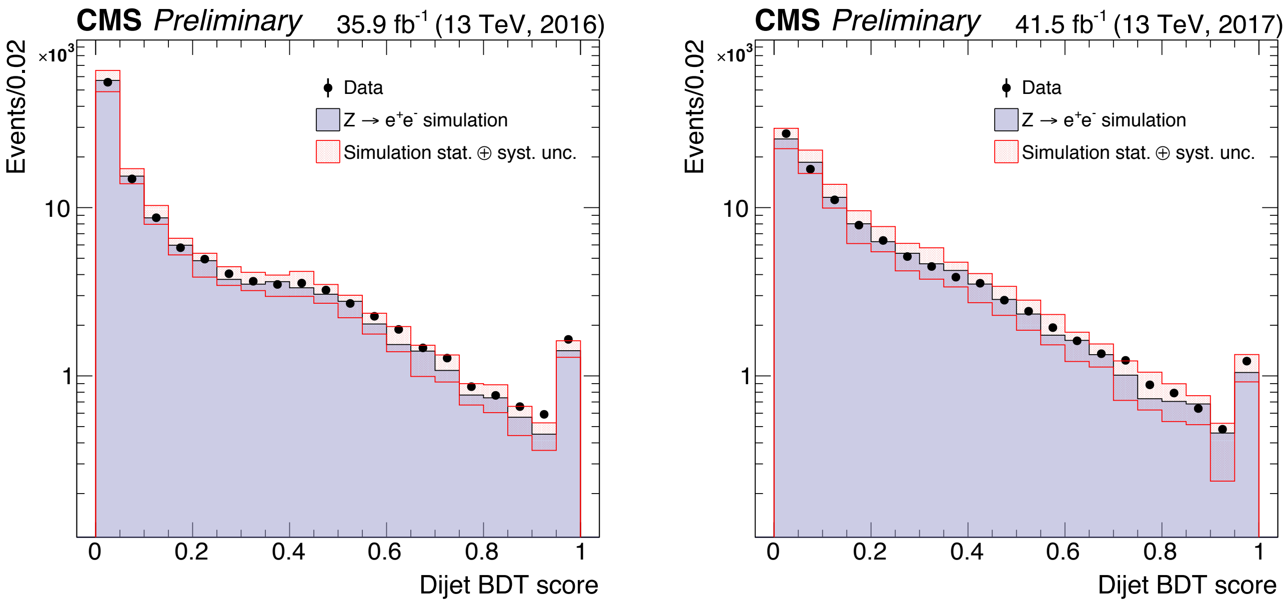
STXS: diphoton BDT



- Reject background using **photon kinematics & photon ID BDTs**
- Classifier uses exact same inputs as in published 2016 result
- Z→ee events used for validation
- Systematics uncertainties related to photon ID and per event energy resolution estimates covers observed discrepancies

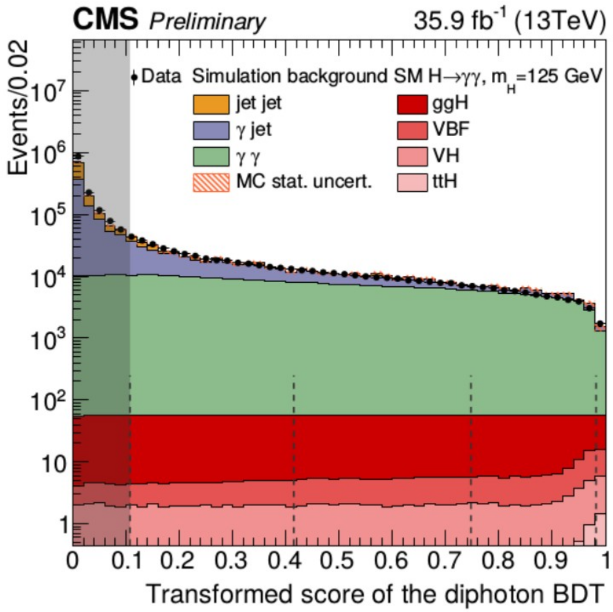
STXS: dijet BDT

- To discriminate between ggH and VBF
- **Inputs mostly related to jet kinematics (jets pT, m_{jj}, $\Delta\eta_{jj}$, $\Delta\varphi_{jj}$,...)**
- New: training on MC (for diphoton background) and data for backgrounds with non-prompt photons in control regions normalized with appropriate fake factors (to increase the statistics for non-prompt backgrounds)
- Validation also performed with Z→ee+jets events
- Agreement is good in both years (JES/JER uncertainties included here)



Diphoton BDT inputs

- the transverse momenta for both photons, rescaled for the diphoton mass, $p_{\mathrm{T}}^{1,(2)}/m_{\gamma\gamma}$;
- the pseudorapidities of both photons, $\eta^{1(2)}$;
- the cosine of the angle between the two photons in the transverse plane, $\cos(\Delta\phi)$;
- the identification BDT score for both photons;
- the per-event relative mass resolution estimate, under the hypothesis that the mass has been reconstructed using the correct primary vertex (σ_{rv});
- the per-event relative mass resolution estimate, under the hypothesis that the mass has been reconstructed using an incorrect primary vertex (σ_{wv});
- the per-event probability estimate that the correct primary vertex has been used to reconstruct the mass, based on the event-level vertex selection MVA as described in Section in [7] (p_{vtx}).



Dijet BDT inputs

- the transverse momenta of the leading and subleading photons divided by the invariant mass of the diphoton candidate: $p_T^{\gamma_1}/m_{\gamma\gamma}$ and $p_T^{\gamma_2}/m_{\gamma\gamma}$
- the transverse momenta of the leading and subleading jets: $p_T^{j_1}$ and $p_T^{j_2}$
- the dijet invariant mass, $m_{j_1 j_2}$
- the difference in pseudo-rapidity between the two jets, $\Delta\eta_{j_1 j_2}$
- the difference in azimuthal angle between the dijet and the diphoton, $\Delta\phi_{(j_1 j_2,\gamma\gamma)}$.
- centrality variable defined as,

$$C_{\gamma\gamma} = \exp\left(-\frac{4}{(\eta_1 - \eta_2)^2}\left(\eta_{\gamma\gamma} - \frac{\eta_1 + \eta_2}{2}\right)^2\right) \quad (7)$$

  where η_1, η_2, and $\eta_{\gamma\gamma}$ are the pseudo-rapidities of the two jets, and the diphoton.
- the difference in azimuthal angle between the two leading jets $\Delta\phi_{jj}$
- the minimum distance between a leading or subleading jet and leading or subleading photon min $\Delta R(\gamma, jet)$.

Samples

- Signal simulation at mH = 120, 125, 130 GeV
  - aMC@NLO dominant processes are ggH, VBF, ttH and VH also include tH, bbH, and ggZH
  - powheg used for MVA training
- Background simulation:
  - Diphoton from Sherpa
  - GJet and QCD from Pythia with EM filter
  - DY to leptons with aMC@NLO for $Z \rightarrow ee$ validation

STXS stage 1: ggH

Region	Definition	Fraction	Cross section (pb)
0J	Exactly zero jets, any p_T^H	60.0%	26.49
1J low	Exactly one jet, $p_T^H < 60$ GeV	15.4%	6.79
1J med	Exactly one jet, 60 GeV $< p_T^H <$ 120 GeV	10.4%	4.61
1J high	Exactly one jet, 120 GeV $< p_T^H <$ 200 GeV	1.7%	0.76
1J BSM	Exactly one jet, $p_T^H > 200$ GeV	0.4%	0.16
2J low	$\geq$ two jets, $p_T^H < 60$ GeV	2.9%	1.26
2J med	$\geq$ two jets, 60 GeV $< p_T^H <$ 120 GeV	4.5%	2.00
2J high	$\geq$ two jets, 120 GeV $< p_T^H <$ 200 GeV	2.3%	1.00
2J BSM	$\geq$ two jets, $p_T^H > 200$ GeV	1.0%	0.43
VBF-like 2J	$\geq$ two jets, $p_T^H < 200$ GeV, $\lvert\Delta\eta\rvert > 2.8$, $m_{jj} > 400$ GeV, $p_T^{Hjj} < 25$ GeV	0.6%	0.27
VBF-like 3J	$\geq$ two jets, $p_T^H < 200$ GeV, $\lvert\Delta\eta\rvert > 2.8$, $m_{jj} > 400$ GeV, $p_T^{Hjj} > 25$ GeV	0.9%	0.38

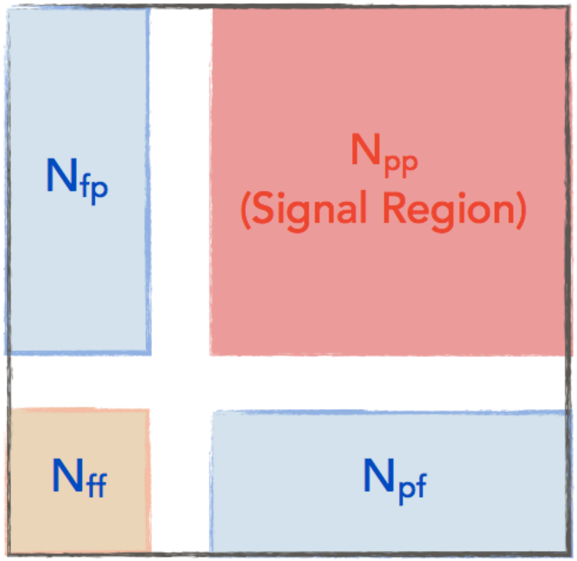
- We have sensitivity to most bins individually; higher pT(H) typically has lower cross-section but similarly lower background
- Exceptions are the low & medium pT(H) 2J categories and the VBF-like region very difficult to separate the latter from true VBF production

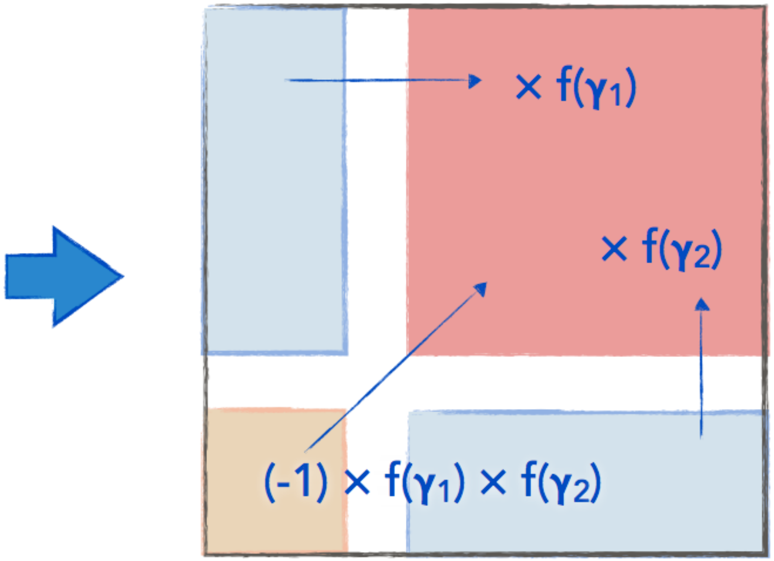
STXS stage 1: VBF

Region	Definition	VBF fraction	VH fraction	Cross section (pb)
BSM	Leading jet $p_{\mathrm{T}} > 200\,\mathrm{GeV}$	4.6%	5.4%	0.23
2J-like	$\geq$ two jets, $\lvert\Delta\eta\rvert > 2.8$, $m_{jj} > 400$ GeV, $p_{\mathrm{T}}^{Hjj} < 25$ GeV	25.8%	0.4%	0.91
3J-like	$\geq$ two jets, $\lvert\Delta\eta\rvert > 2.8$, $m_{jj} > 400$ GeV, $p_{\mathrm{T}}^{Hjj} > 25$ GeV	9.0%	1.7%	0.34
VH-like	$\geq$ two jets, $60 < m_{jj} < 120\,\mathrm{GeV}$	2.3%	34.5%	0.55
Rest	All other VBF events	59.2%	57.9%	2.86

- Two main bins defined in the same way as the VBF-like bins in the ggH phase space, split into 2J-like and 3J-like with cut on pT(H_{jj}) (dijet present with m_{jj} > 400 GeV and $\Delta\eta$ > 2.8)
- A BSM bin where lead jet has pT > 200 GeV
- VH bin with 60 < m_{jj} < 120 GeV
- Everything else in “Rest” bin; corresponds to over 60% of signal

STXS: Dijet BDT data-driven



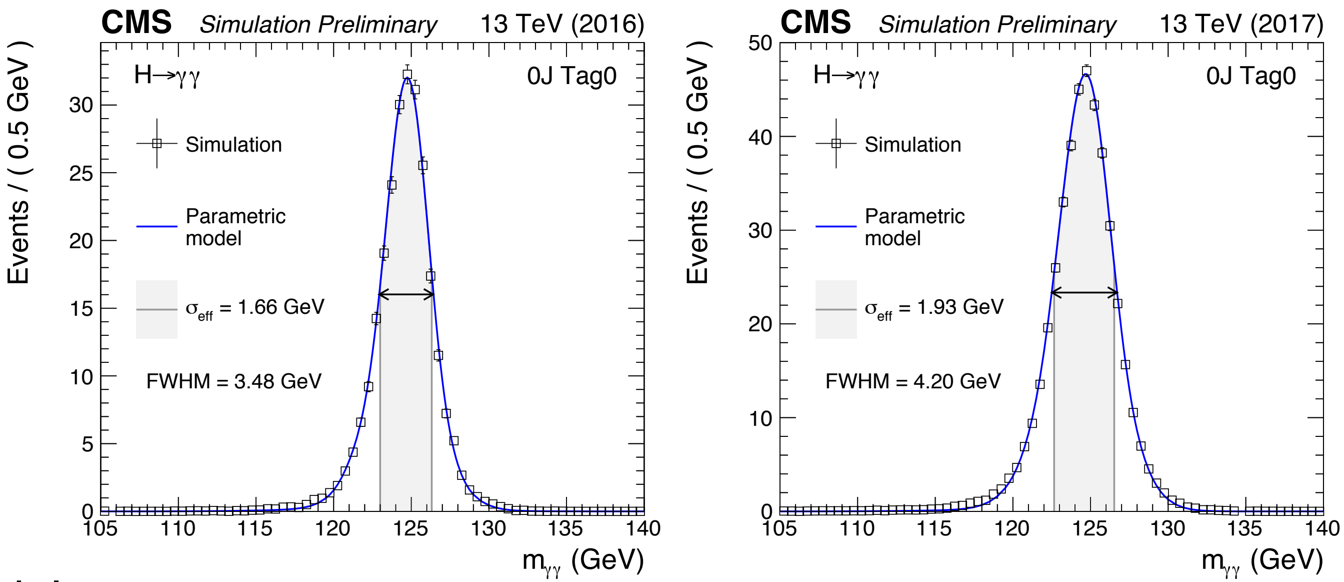


- Fake factors and QCD fraction estimated from MC
  - binned in p_T and η for each photon
- These are applied to data in the control regions, which then replaces the MC in training

$$f(\eta^\gamma, p_T^\gamma) = \left(\frac{N^{SR}(\eta^\gamma, p_T^\gamma)}{N^{CR}(\eta^\gamma, p_T^\gamma)}\right)_{MC}$$

$$p_{QCD}(\eta^\gamma, p_T^\gamma) = \left(\frac{N_{j\gamma}^{CR} + N_{jj}^{CR}}{N_{\gamma\gamma}^{CR} + N_{j\gamma}^{CR} + N_{jj}^{CR}}\right)$$

STXS: signal model



Signal model:

- Parametric signal model with shape parameters linear functions of m_H obtained
    - from simultaneous fit to 120, 125, and 130 GeV mass points including all data/MC corrections to properly reproduce the resolution
    - for each process × category × right/wrong vertex treated separately
- Resolution 10-15% worse in 2017 (calibration issue, will be fixed in re-reco)

Background model: choice of background function being treated as a discrete nuisance parameter

STXS: yields

Event Categories	SM 125 GeV Higgs boson expected signal														Bkg (GeV^{-1})	S/(S+B)
	Total	ggH	VBF	ttH	tHq	tHW	bbH	ggZH	WH lep	WH had	ZH lep	ZH had	σ_{eff}	σ_{HM}		
0J Tag 0	401.1	91.8 %	4.4 %	<0.05 %	<0.05 %	<0.05 %	1.4 %	0.1 %	1.0 %	0.4 %	0.6 %	0.2 %	1.94	1.79	870.3	0.07
0J Tag 1	552.3	93.7 %	3.1 %	<0.05 %	<0.05 %	<0.05 %	1.3 %	<0.05 %	0.7 %	0.4 %	0.4 %	0.2 %	2.42	2.06	2121.9	0.04
0J Tag 2	347.3	95.0 %	2.2 %	<0.05 %	<0.05 %	<0.05 %	1.3 %	<0.05 %	0.5 %	0.4 %	0.3 %	0.2 %	2.72	2.41	3035.8	0.01
1J Low Tag 0	130.8	89.5 %	5.9 %	0.1 %	<0.05 %	<0.05 %	1.1 %	<0.05 %	0.5 %	1.7 %	0.2 %	0.9 %	1.91	1.71	360.2	0.06
1J Low Tag 1	111.5	89.2 %	6.1 %	0.1 %	<0.05 %	<0.05 %	1.1 %	<0.05 %	0.5 %	1.8 %	0.2 %	1.0 %	2.47	2.22	689.4	0.02
1J Medium Tag 0	71.4	81.5 %	12.4 %	0.2 %	0.1 %	<0.05 %	0.5 %	0.2 %	0.9 %	2.5 %	0.4 %	1.3 %	1.85	1.67	110.8	0.11
1J Medium Tag 1	91.1	82.7 %	11.4 %	0.2 %	0.1 %	<0.05 %	0.5 %	0.2 %	0.8 %	2.3 %	0.4 %	1.4 %	2.13	1.91	342.2	0.04
1J High Tag 0	14.7	71.7 %	19.4 %	0.3 %	0.2 %	<0.05 %	0.3 %	1.0 %	2.3 %	2.5 %	1.0 %	1.5 %	1.54	1.51	8.7	0.27
1J High Tag 1	28.2	72.4 %	18.4 %	0.4 %	0.2 %	<0.05 %	0.3 %	0.8 %	2.2 %	2.8 %	0.9 %	1.7 %	1.76	1.77	47.7	0.10
1J BSM	15.5	66.9 %	20.9 %	0.4 %	0.3 %	0.1 %	0.1 %	1.0 %	4.0 %	3.0 %	1.6 %	1.8 %	1.76	1.71	17.5	0.15
2J Low Tag 0	10.9	80.2 %	7.0 %	1.7 %	0.4 %	<0.05 %	1.0 %	0.3 %	0.7 %	4.8 %	0.3 %	3.4 %	1.55	1.52	35.1	0.06
2J Low Tag 1	40.8	77.6 %	8.1 %	3.0 %	0.5 %	<0.05 %	0.8 %	0.3 %	0.7 %	5.4 %	0.3 %	3.1 %	2.06	1.94	249.0	0.03
2J Medium Tag 0	16.8	76.6 %	8.1 %	1.9 %	0.5 %	0.1 %	0.3 %	1.0 %	0.7 %	7.0 %	0.4 %	3.4 %	1.60	1.46	28.9	0.11
2J Medium Tag 1	49.7	74.6 %	9.1 %	3.4 %	0.6 %	0.1 %	0.4 %	0.8 %	0.9 %	6.1 %	0.4 %	3.6 %	2.12	1.86	228.8	0.03
2J High Tag 0	14.0	71.1 %	9.2 %	1.7 %	0.6 %	0.1 %	0.2 %	2.7 %	1.0 %	8.2 %	0.7 %	4.6 %	1.54	1.52	14.2	0.18
2J High Tag 1	24.4	69.1 %	9.4 %	3.7 %	0.8 %	0.2 %	0.2 %	2.3 %	1.1 %	8.2 %	0.5 %	4.7 %	1.42	1.31	64.4	0.08
2J BSM Tag 0	15.8	66.4 %	9.4 %	2.6 %	0.9 %	0.4 %	0.1 %	2.7 %	1.9 %	9.3 %	0.9 %	5.4 %	1.67	1.63	11.1	0.22
2J BSM Tag 1	5.7	60.4 %	9.5 %	9.2 %	1.4 %	0.7 %	0.1 %	2.7 %	1.4 %	9.0 %	1.0 %	4.7 %	1.89	1.82	24.3	0.04
VBF 2J-like Tag 0	13.5	24.8 %	74.4 %	0.1 %	0.1 %	<0.05 %	0.1 %	0.1 %	<0.05 %	0.2 %	<0.05 %	0.2 %	1.90	1.73	5.7	0.30
VBF 2J-like Tag 1	4.8	41.7 %	56.5 %	0.2 %	0.2 %	<0.05 %	0.2 %	0.2 %	0.2 %	0.5 %	<0.05 %	0.3 %	2.28	1.94	9.3	0.07
VBF 3J-like Tag 0	12.7	36.8 %	60.6 %	0.4 %	0.5 %	<0.05 %	0.1 %	0.4 %	0.2 %	0.5 %	0.1 %	0.2 %	1.90	1.69	7.8	0.23
VBF 3J-like Tag 1	7.6	56.0 %	37.8 %	0.8 %	0.9 %	<0.05 %	0.2 %	0.8 %	0.5 %	1.6 %	0.2 %	1.0 %	1.86	1.79	11.1	0.11
VBF Rest	12.9	63.4 %	29.9 %	1.0 %	0.6 %	0.1 %	0.4 %	0.8 %	0.6 %	2.0 %	0.3 %	1.1 %	1.80	1.71	21.3	0.10
VBF BSM	6.5	44.7 %	47.8 %	1.0 %	0.5 %	0.3 %	0.1 %	1.4 %	0.7 %	2.1 %	0.4 %	1.0 %	1.75	1.45	4.5	0.22
Total	1999.8	88.2 %	6.7 %	0.4 %	0.1 %	<0.05 %	1.1 %	0.2 %	0.8 %	1.4 %	0.4 %	0.8 %	2.22	1.98	8320.2	0.04

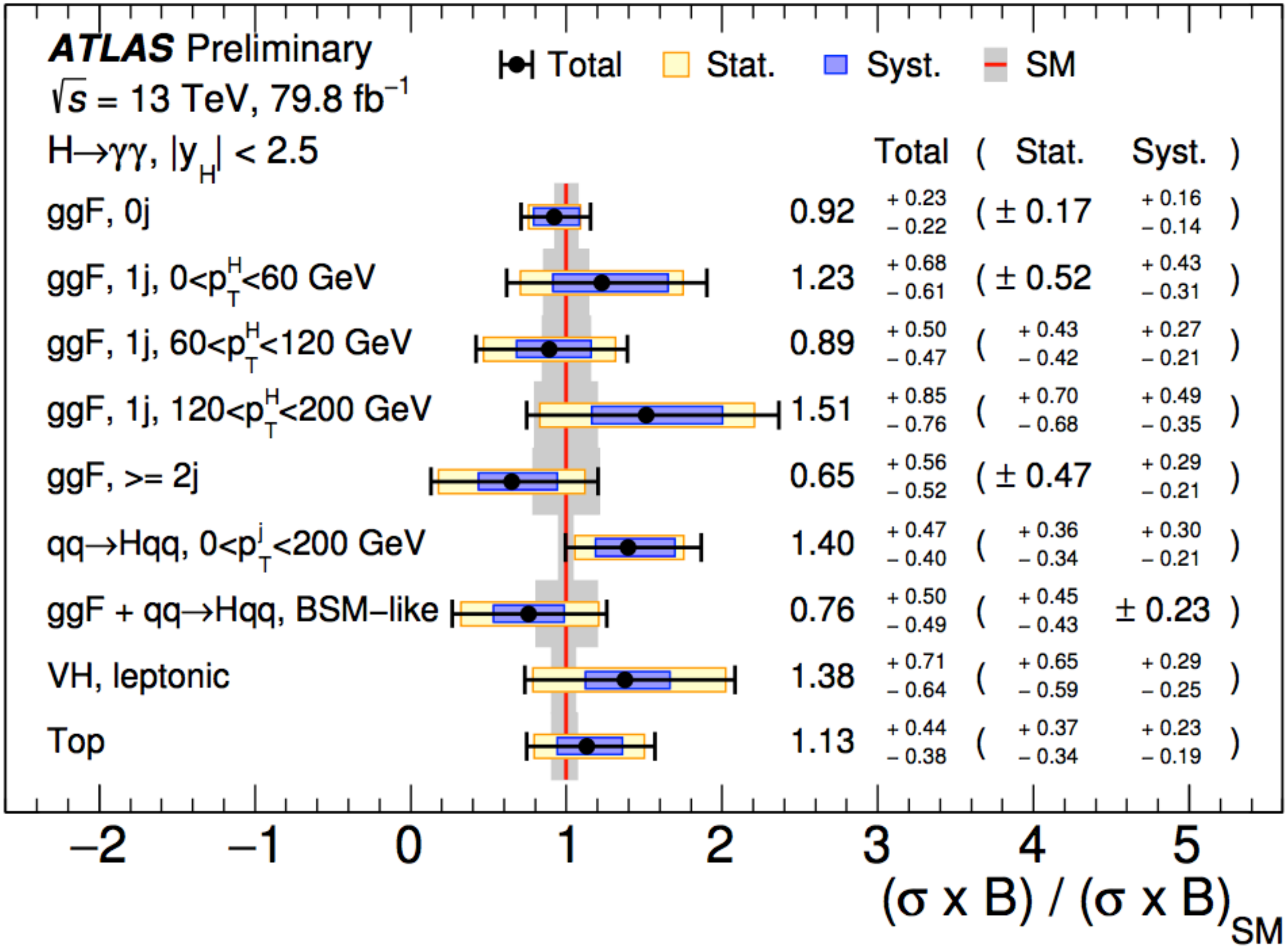
STXS: results

Signal parameter	Cross section (fb) SM pred.	Cross section (fb) Measured	σ/σ_{SM}	Uncertainty on σ/σ_{SM} Total	Stat.	Exp.	Theo.
ggH 0J	61	72	1.17	$^{+0.20}_{-0.20}$	$^{+0.18}_{-0.18}$	$^{+0.08}_{-0.07}$	$^{+0.06}_{-0.04}$
ggH 1J low	15	24	1.5	$^{+0.7}_{-0.6}$	$^{+0.6}_{-0.5}$	$^{+0.2}_{-0.1}$	$^{+0.2}_{-0.1}$
ggH 1J med	10	5.1	0.5	$^{+0.5}_{-0.4}$	$^{+0.4}_{-0.4}$	$^{+0.1}_{-0.1}$	$^{+0.1}_{-0.0}$
ggH 1J high	1.7	3.4	2.0	$^{+1.0}_{-0.7}$	$^{+0.8}_{-0.7}$	$^{+0.3}_{-0.1}$	$^{+0.4}_{-0.2}$
ggH 1J BSM	0.4	0.6	1.8	$^{+1.7}_{-1.5}$	$^{+1.5}_{-1.4}$	$^{+0.3}_{-0.2}$	$^{+0.4}_{-0.1}$
ggH 2J low	2.9	0.8	0.3	$^{+1.5}_{-0.3}$	$^{+1.4}_{-0.3}$	$^{+0.3}_{-0.1}$	$^{+0.3}_{-0.0}$
ggH 2J med	4.6	12	2.6	$^{+1.1}_{-1.1}$	$^{+1.0}_{-1.0}$	$^{+0.3}_{-0.2}$	$^{+0.4}_{-0.3}$
ggH 2J high	2.3	1.3	0.6	$^{+0.8}_{-0.6}$	$^{+0.7}_{-0.6}$	$^{+0.2}_{-0.1}$	$^{+0.3}_{-0.0}$
ggH 2J BSM	1.0	2.7	2.8	$^{+1.1}_{-1.2}$	$^{+0.8}_{-1.0}$	$^{+0.3}_{-0.3}$	$^{+0.5}_{-0.4}$
ggH VBF-like	1.5	0	0.0	$^{+0.5}_{-0.0}$	$^{+0.5}_{-0.0}$	$^{+0.2}_{-0.0}$	$^{+0.1}_{-0.0}$
qqH 2J-like	2.1	2.6	1.3	$^{+0.6}_{-0.5}$	$^{+0.4}_{-0.4}$	$^{+0.4}_{-0.3}$	$^{+0.1}_{-0.1}$
qqH 3J-like	0.8	0	0.0	$^{+0.7}_{-0.0}$	$^{+0.6}_{-0.0}$	$^{+0.2}_{-0.0}$	$^{+0.0}_{-0.0}$
qqH other	8.2	0	0.0	$^{+1.7}_{-0.0}$	$^{+1.6}_{-0.0}$	$^{+0.6}_{-0.0}$	$^{+0.2}_{-0.0}$

STXS: results

Signal parameter	Cross section (fb) SM pred.	Measured	σ/σ_{SM}	Uncertainty on σ/σ_{SM} Total	Stat.	Exp.	Theo.
ggH 0J	61	72	1.18	$^{+0.20}_{-0.20}$	$^{+0.18}_{-0.18}$	$^{+0.10}_{-0.08}$	$^{+0.06}_{-0.05}$
ggH 1J low	15	21	1.3	$^{+0.6}_{-0.5}$	$^{+0.6}_{-0.5}$	$^{+0.2}_{-0.2}$	$^{+0.2}_{-0.1}$
ggH 1J med	10	7.6	0.7	$^{+0.4}_{-0.4}$	$^{+0.4}_{-0.4}$	$^{+0.1}_{-0.1}$	$^{+0.1}_{-0.0}$
ggH 1J high	1.7	2.9	1.7	$^{+1.0}_{-0.7}$	$^{+0.8}_{-0.6}$	$^{+0.3}_{-0.2}$	$^{+0.2}_{-0.1}$
ggH 2J	11	8.4	0.8	$^{+0.6}_{-0.5}$	$^{+0.5}_{-0.5}$	$^{+0.1}_{-0.1}$	$^{+0.3}_{-0.1}$
ggH BSM	1.3	2.9	2.2	$^{+0.8}_{-0.8}$	$^{+0.6}_{-0.6}$	$^{+0.4}_{-0.3}$	$^{+0.3}_{-0.2}$
qqH	11	9.1	0.8	$^{+0.4}_{-0.3}$	$^{+0.4}_{-0.3}$	$^{+0.2}_{-0.1}$	$^{+0.1}_{-0.0}$

STXS: comparison with ATLAS



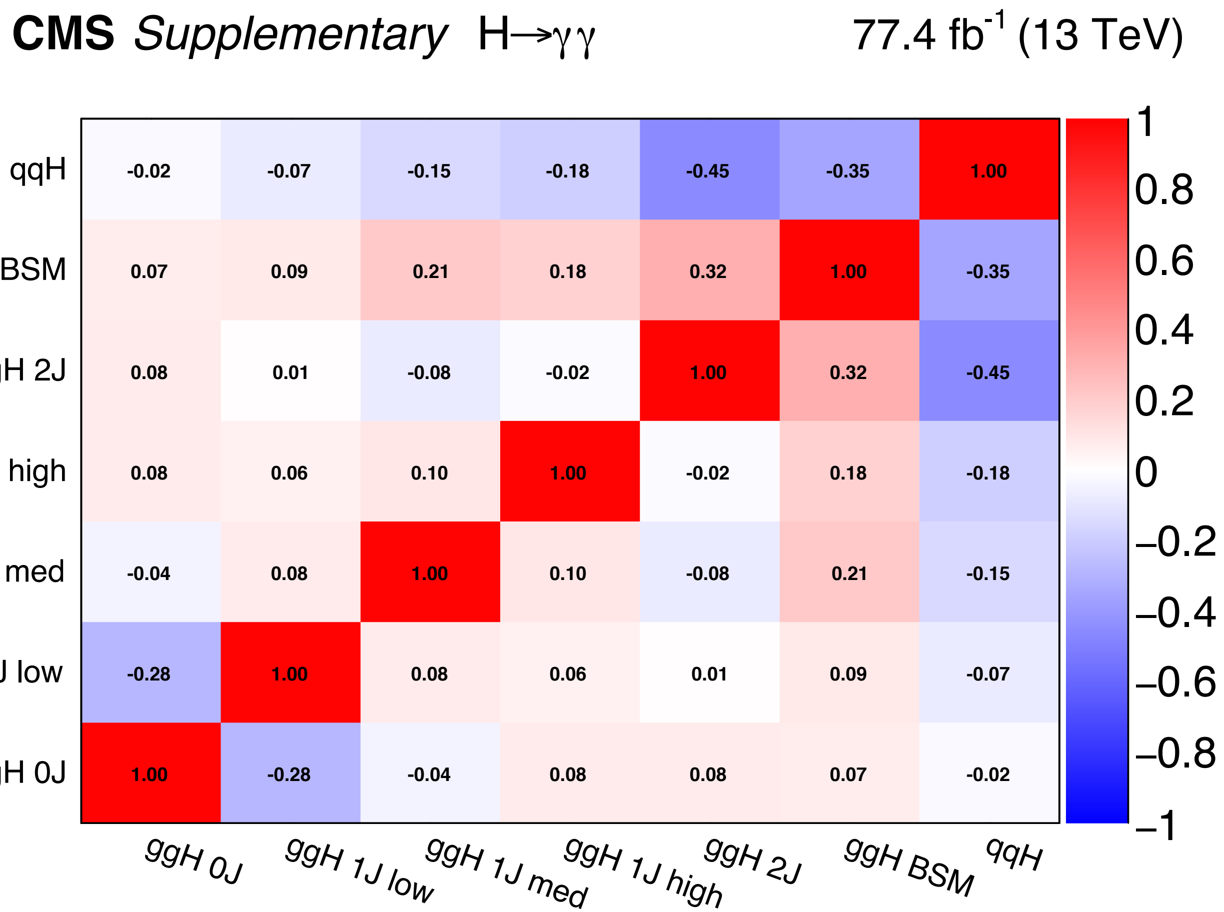
STXS categorisation

- ggH categorisation:
  - Use diphoton BDT to reject background in the categories targeting each ggH stage 1 bin
  - The category definition is a two-step process:
    - **first a target bin is assigned based on the reconstructed pT(H) and nJets**
    - **then the diphoton BDT boundaries are chosen independently for each bin**
  - Limit the maximum number of categories for each bin to three, third category only required for the high stats ggH 0J bin
- VBF categorisation:
  - Six categories are constructed in total:
    1. A single category for the VBF BSM bin, with pT of the leading jet > 200 GeV
    2. Two categories for each pT(Hjj) bin in the "VBF-like" region require the VBF cuts of mjj > 400 GeV and $\Delta\eta > 2.8$ then split at the pT(Hjj) = 25 GeV boundary
    3. A single category for the VBF rest bin, with 120 < mjj < 400 GeV
  - Optimized cuts on both the dijet and diphoton BDTs in each category

STXS: systematics

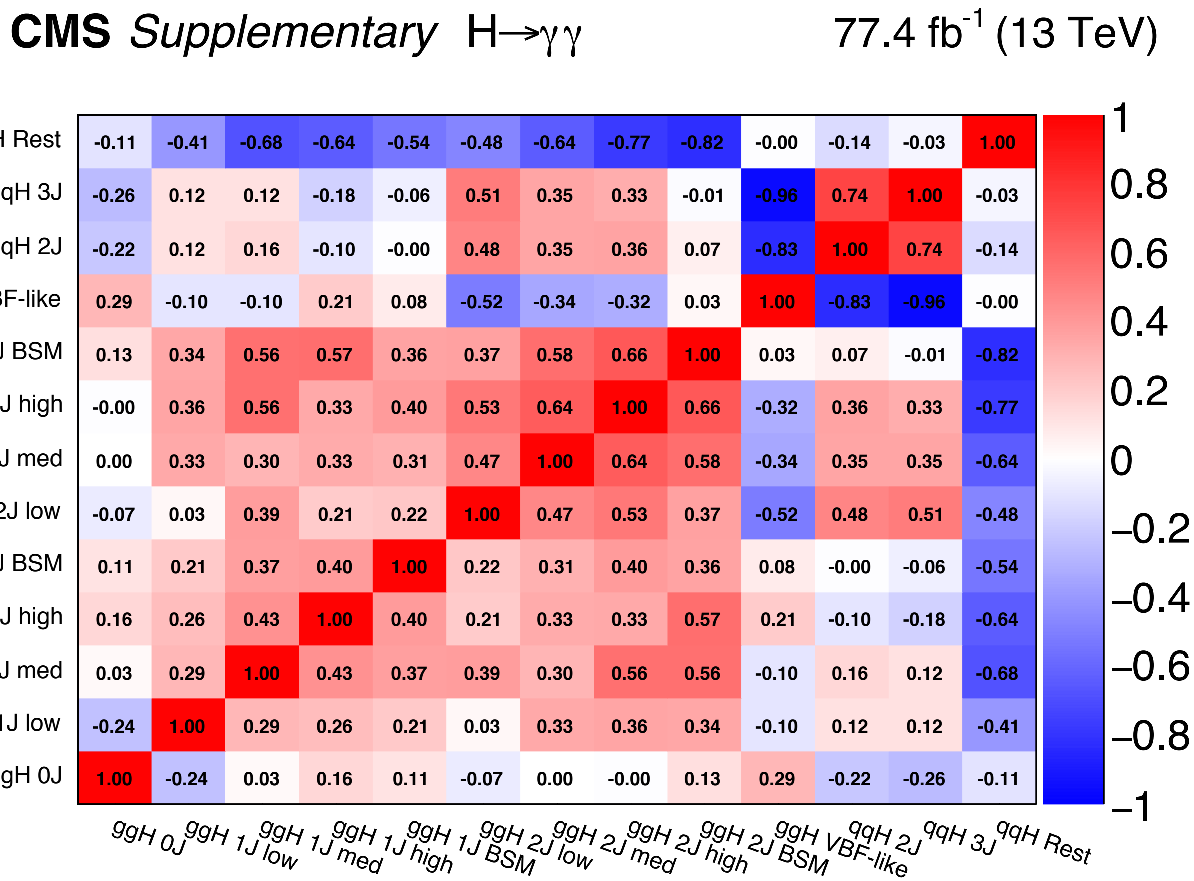
- Jet energy scale is very important for this analysis
- Previously, was implemented as multiple nuisances representing migrations between Untagged and VBF tags, and within VBF tags
- “conservative” approach inherited from Run 1
- However **jets are now also used in the ggH phase space**
  - Single nuisance is standard implementation → try this first checked that these are not highly constrained in the fit ✓
- The other leading experimental systematic is the photon IDMVA
- Theory systematics: **do not include uncertainty on the cross-section itself** this differentiates the STXS measurement from a signal strength
  the effect on the analysis efficiency × acceptance is however included

STXS: correlation matrix



Essential for theory reinterpretations

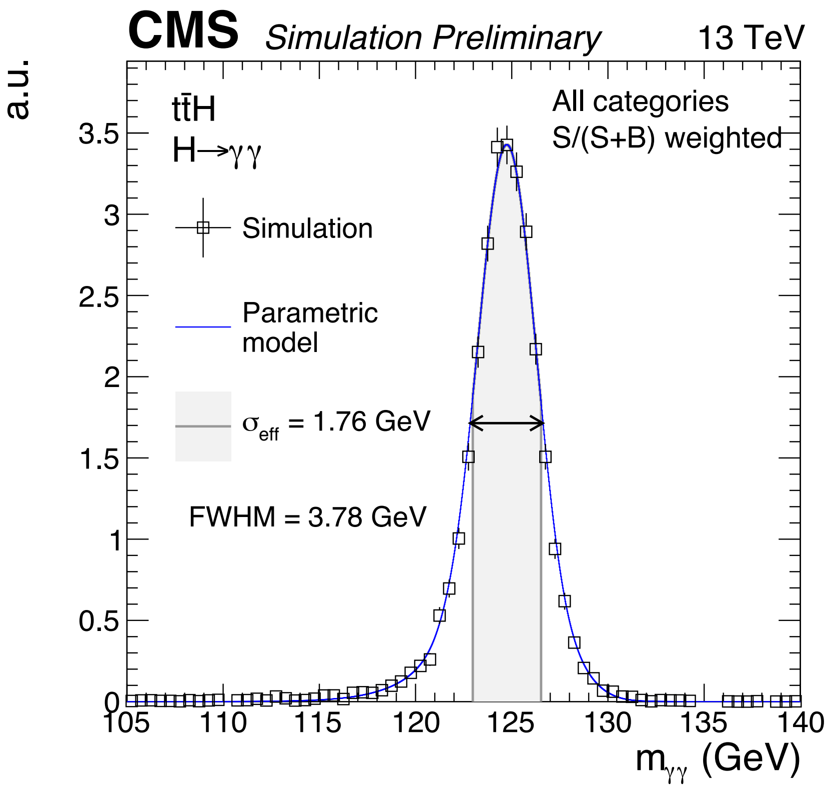
STXS: correlation matrix



	ggH 0J	ggH 1J low	ggH 1J med	ggH 1J high	ggH 1J BSM	ggH 2J low	ggH 2J med	ggH 2J high	ggH 2J BSM	ggH VBF-like	qqH 2J	qqH 3J	qqH Rest
qqH Rest	-0.11	-0.41	-0.68	-0.64	-0.54	-0.48	-0.64	-0.77	-0.82	-0.00	-0.14	-0.03	1.00
qqH 3J	-0.26	0.12	0.12	-0.18	-0.06	0.51	0.35	0.33	-0.01	-0.96	0.74	1.00	-0.03
qqH 2J	-0.22	0.12	0.16	-0.10	-0.00	0.48	0.35	0.36	0.07	-0.83	1.00	0.74	-0.14
ggH VBF-like	0.29	-0.10	-0.10	0.21	0.08	-0.52	-0.34	-0.32	0.03	1.00	-0.83	-0.96	-0.00
ggH 2J BSM	0.13	0.34	0.56	0.57	0.36	0.37	0.58	0.66	1.00	0.03	0.07	-0.01	-0.82
ggH 2J high	-0.00	0.36	0.56	0.33	0.40	0.53	0.64	1.00	0.66	-0.32	0.36	0.33	-0.77
ggH 2J med	0.00	0.33	0.30	0.33	0.31	0.47	1.00	0.64	0.58	-0.34	0.35	0.35	-0.64
ggH 2J low	-0.07	0.03	0.39	0.21	0.22	1.00	0.47	0.53	0.37	-0.52	0.48	0.51	-0.48
ggH 1J BSM	0.11	0.21	0.37	0.40	1.00	0.22	0.31	0.40	0.36	0.08	-0.00	-0.06	-0.54
ggH 1J high	0.16	0.26	0.43	1.00	0.40	0.21	0.33	0.33	0.57	0.21	-0.10	-0.18	-0.64
ggH 1J med	0.03	0.29	1.00	0.43	0.37	0.39	0.30	0.56	0.56	-0.10	0.16	0.12	-0.68
ggH 1J low	-0.24	1.00	0.29	0.26	0.21	0.03	0.33	0.36	0.34	-0.10	0.12	0.12	-0.41
ggH 0J	1.00	-0.24	0.03	0.16	0.11	-0.07	0.00	-0.00	0.13	0.29	-0.22	-0.26	-0.11

Essential for theory reinterpretations

ttH: signal model



ttH: objects definition

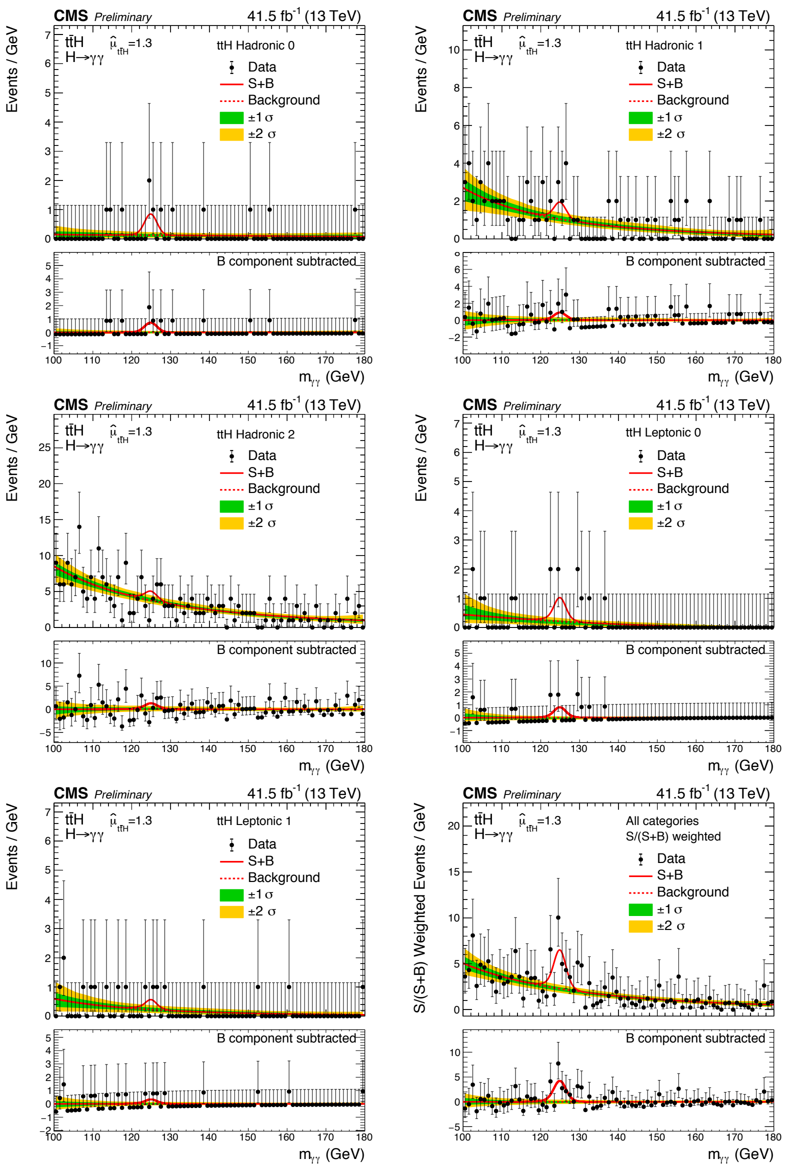
- Jets are reconstructed using anti-k_T algorithm with a radius parameter of 0.4.
  - Jets are selected by requiring tight jetID and $p_T > 25$ GeV in $|\eta| < 2.4$.
- b-jets are tagged using the centrally defined DeepCSV algorithm,
  - medium working point is chosen to quantify the b-jet multiplicity.
- Electrons are identified using the ID provided by e/gamma POG with $p_T > 10$ GeV and $|\eta| < 2.5$, muons are required to have $p_T > 10$ GeV and $|\eta| < 2.4$.
- All leptons are required not to overlap with photons by imposing a $\Delta R > 0.2$. Jets are also required not to overlap with photons and leptons by imposing $\Delta R > 0.4$.
- All scale factors are applied following the POG recommendations.

ttH: yields table

Event categories	SM 125 GeV Higgs boson expected signal													Bkg
	Total	ttH	bbH	tHq	tHW	ggH	VBF	WH lep	ZH lep	WH had	ZH had	σ_{eff}	FWHM	(GeV^{-1})
ttH Hadronic 0	2.4	86.7 %	<0.05 %	5.0 %	2.8 %	2.6 %	0.1 %	0.1 %	0.1 %	0.7 %	1.8 %	1.66	1.61	0.2
ttH Hadronic 1	3.3	79.2 %	0.2 %	5.6 %	2.4 %	7.5 %	0.2 %	0.4 %	0.1 %	1.0 %	3.3 %	1.79	1.62	1.1
ttH Hadronic 2	5.2	62.9 %	0.2 %	5.9 %	1.9 %	18.4 %	1.3 %	0.6 %	0.4 %	3.2 %	5.1 %	2.02	1.72	3.8
ttH Leptonic 0	2.7	88.5 %	<0.05 %	5.2 %	4.4 %	0.2 %	<0.05 %	1.2 %	0.2 %	<0.05 %	0.1 %	1.79	1.66	0.3
ttH Leptonic 1	1.2	87.6 %	<0.05 %	5.5 %	1.8 %	2.0 %	0.2 %	1.9 %	0.8 %	<0.05 %	0.2 %	1.88	1.59	0.3
Total	14.8	77.2 %	0.1 %	5.5 %	2.6 %	8.7 %	0.5 %	0.7 %	0.3 %	1.5 %	2.8 %	1.84	1.65	5.6

Table 2: The expected number of signal events per category and the percentage breakdown per production mode in that category. The σ_{eff}, computed as the smallest interval containing 68.3% of the invariant mass distribution, and FWHM, computed as the width of the distribution at half of its highest point divided by 2.35 are also shown as an estimate of the $m_{\gamma\gamma}$ resolution in that category. The expected number of background events per GeV around 125 GeV is also listed.

ttH:


CMS Preliminary
41.5 fb-1 (13 TeV)
tt̄H
H→γγ
μ̂ttH=1.3
ttH Hadronic 0
ttH Hadronic 1
ttH Hadronic 2
ttH Leptonic 0
ttH Leptonic 1
All categories
S/(S+B) weighted
Data
S+B
Background
±1 σ
±2 σ
B component subtracted
Events / GeV
S/(S+B) Weighted Events / GeV
mγγ (GeV)

ttH: systematics

- The dominant theoretical uncertainties:
  - QCD scale: 9%
  - PDF: 5%
  - Strong coupling constant: 3%
  - $H \rightarrow \gamma\gamma$ branching fraction: 2%
  - ggH contamination: 2%
- The dominant experimental uncertainties
  - Photon identification: 6%
  - Jet energy scale resolution: 4%
  - Shape of the b discriminant: 3%
  - Integrated luminosity: 2.3%