

$H \rightarrow \gamma\gamma$ in CMS with the 2016 dataset

a.k.a.

“Measurements of Higgs boson properties in the diphoton decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV”, arXiv:1804.02716

and

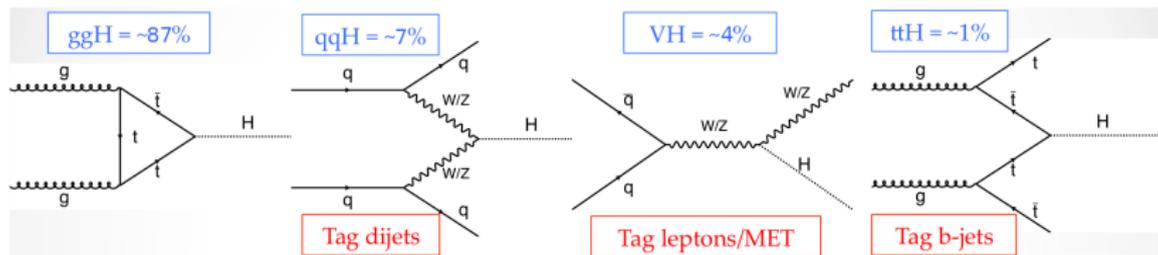
“Measurement of differential fiducial cross sections for Higgs boson production in the diphoton decay channel in pp collisions at $\sqrt{s} = 13$ TeV”, CMS-PAS-HIG-17-015

Federico Ferri, on behalf of the CMS Collaboration

CEA-Saclay Irfu-DPhP

IRN Terascale@Strasbourg, May 30, 2018

Introduction



- Low BR $\approx 0.2\%$, clean signature: narrow resonance of two isolated photons
- Can reconstruct the invariant mass with high precision:

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

- Quest on the detector:
 - excellent **energy** and **angle** resolution
 - resilience against high-luminosity: pileup and irradiation effects
 - excellent **photon identification**

N.B. W.r.t. 7-8 TeV (Run1): harder and more abundant jets in the final state makes the analysis intrinsically more challenging

More data, more measurements

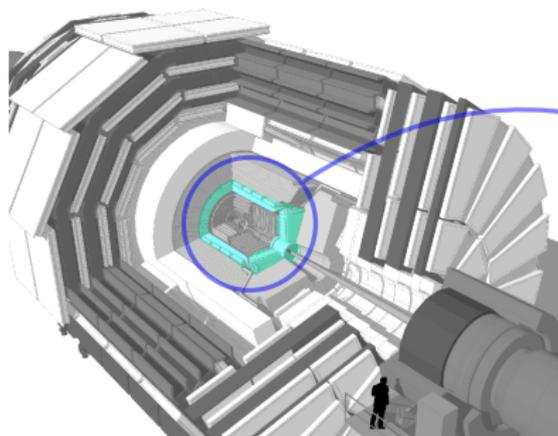
- **Coarse** granularity:
 - **signal strength** relative to the standard model (SM) prediction
 - signal strengths associated with the different Higgs **production mechanisms**
 - **couplings** to bosons and fermions, effective couplings to photons and gluons

- **Finer** granularity:
 - **differential** measurements,
 - **theory-driven selections** (a.k.a. simplified template cross-sections)
 - ...

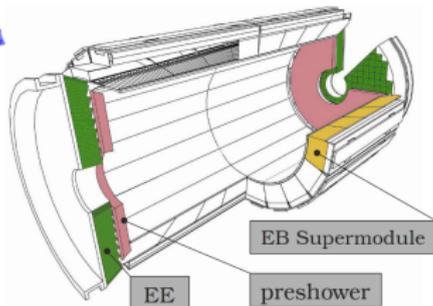
- **Mass measurement**: an optimal compromise between sufficient amount of data and fine assessment of systematic uncertainties

The main CMS actor: the ECAL

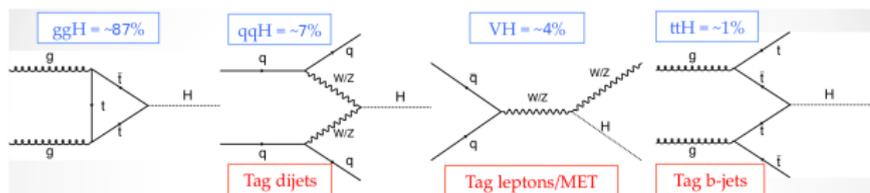
- **Homogeneous, hermetic, high granularity PbWO_4 crystal calorimeter**
 - density of 8.3 g/cm^3 , radiation length 0.89 cm , Molière radius 2.2 cm ,
 $\approx 80\%$ of scintillating light in $\approx 25 \text{ ns}$, refractive index 2.2 , light yield spread among crystals $\approx 10\%$
- **Barrel:** 61200 crystals ($\approx 26X_0$) in 36 super-modules, $|\eta| < 1.48$,
Avalanche Photo-Diode (APD) readout
- **Endcaps:** 14648 crystals ($\approx 25X_0$) in 4-Dees, $1.48 < |\eta| < 3.0$,
Vacuum Photo-Triode (VPT) readout
- **Preshower** (endcaps only): $3X_0$ of Pb/Si strips, $1.65 < |\eta| < 2.6$



- Solenoidal magnetic field: 3.8 T
ECAL fully contained in the coil
- CMS tracker coverage: $|\eta| < 2.5$



Measurement strategy: generalities



Maximize the sensitivity:

- target specific production modes with reduced background contamination, i.e. categorize events with specific features of the final state
- events with no specific features are categorized according to the expected probability to be signal rather than background

Improve the analysis steps **with multivariate techniques**:

- photon reconstruction (E_1 , E_2), identification, vertex ($\cos\vartheta$)
- event classification: mostly to enhance jet identification, in VBF vs ggH, and in ttH vs ggH and background

Simultaneous maximum-likelihood fit to the invariant mass distributions in all event categories:

- model of the **expected signal derived from simulation**
- parametric **background fit to data** (no simulation inputs)

The strategy in 5 steps

1. Photon **energy**

- ECAL clustering independent of particle type (e, γ)
- calibration and corrections
→ estimate per- γ resolution

2. Photon **identification**

- based on shower shapes and isolation variables
→ estimate per- γ quality

4. Event **selection** and analysis **categories**

- All event information in one diphoton BDT: kinematics, resolution, photon quality, vertexing
- + taggers for VBF, VH, ttH, and jet-discriminant BDTs for VBF, ttH

5. **Global fit** (and systematic uncertainties)

★ **Signal model** from simulation

- $Z \rightarrow ee$ to model detector response (i.e. tune MC simulation), γ energy scale, resolution, and efficiencies

3. **Vertex** identification:

- Select vertex among ≈ 20 concurrent collisions
→ estimate per-event probability of correct vertex assignment

Ingredients: photon reconstruction...

- Electrons and photons deposit **energy** over **several crystals** (70% in one 97% in a 3×3 array), **spread in φ** , collected by “**clustering**” algorithms.

$R_9 \equiv$ energy fraction in the 3×3 w.r.t. the cluster
→ distinguish γ conversions

$$E_{e,\gamma} = \mathcal{G} \mathcal{F}_{e,\gamma} \sum_i c_i s_i(t) \mathcal{A}_i$$

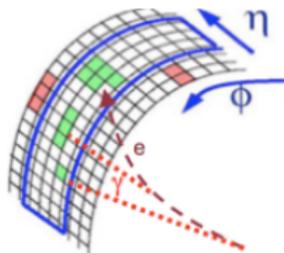
\mathcal{A}_i : single channel amplitude

$s_i(t)$: response **stability** (mostly transparency changes with irradiation)

c_i : response **uniformity**: inter-calibration of the single channel response, using physics: φ - and time-invariance of the energy flow in minimum-bias events, $\pi^0, \eta \rightarrow \gamma\gamma$ and $Z \rightarrow ee$ invariant mass peak, electron E/p

$\mathcal{F}_{e,\gamma}$: **particle energy correction** (geometry, tracker material, clustering, ...)

\mathcal{G} : **global scale** calibration, with $Z \rightarrow ee$ events



Resolution, efficiency and particle ID: $Z \rightarrow ee$

...validation with data ($Z \rightarrow ee$)...

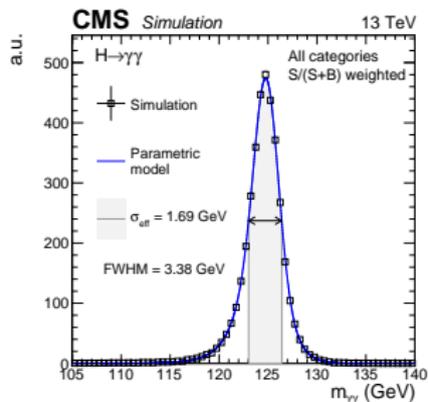
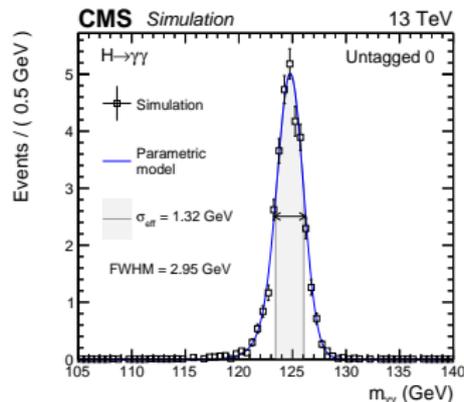
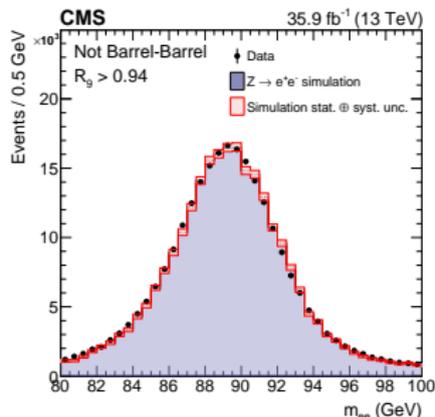
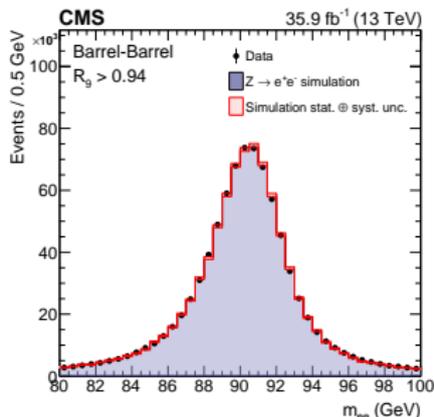
- Electrons reconstructed as photons (track for ID only)

- Energy scale corrections (0.1-0.3% in EB) extracted differentially in time in 2 categories for EB and for EE, 2 categories in R_9

- Additional smearing inclusive for the dataset, same categories in η and R_9 as for the scales (0.1-2.7%) [correlation with R_9 suggests material effects]

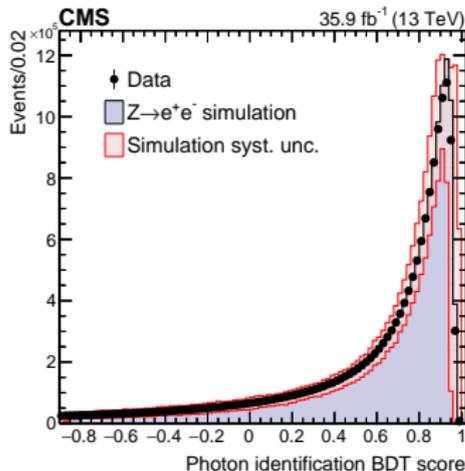
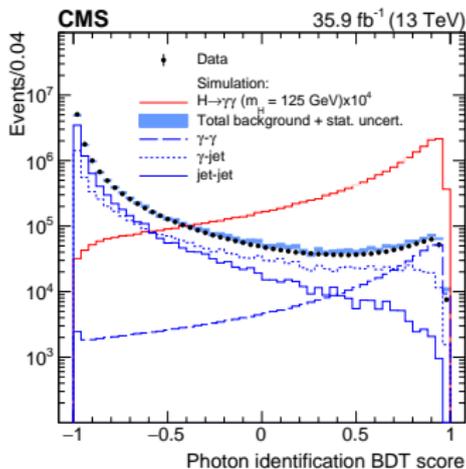


- All incorporated in the signal model \Rightarrow



...and identification

- BDT trained on γ -jet simulation
signal \equiv prompt photons; background \equiv non-prompt photons
- Inputs: shower shapes observables, isolations, photon p_T and η , median energy density of the event (ρ)
- Validated on $Z \rightarrow ee$ events
 - systematic uncertainties applied to cover data-simulation discrepancies



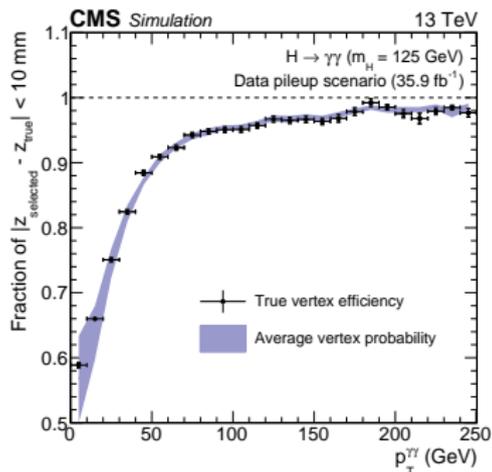
Vertex identification

Resolution unaffected if vertex within 10 mm of the true position

BDT to **identify vertex**:

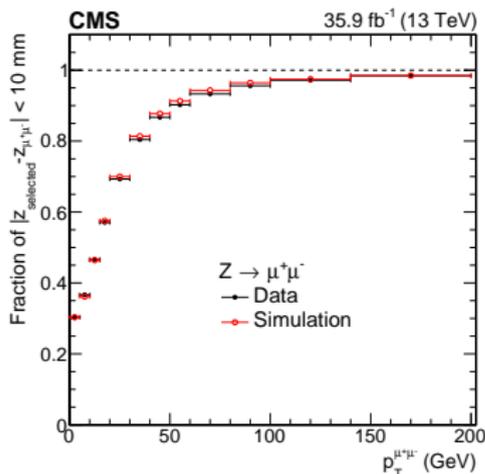
- hardness of interaction p_T
- balance of diphoton system and charged tracks
- γ -conversion information

validation: $Z \rightarrow \mu\mu$ (γ -jet for conversions)



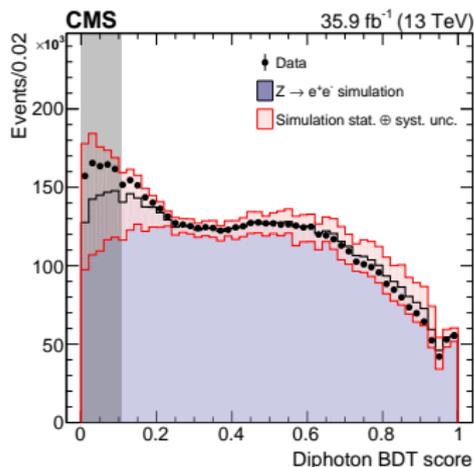
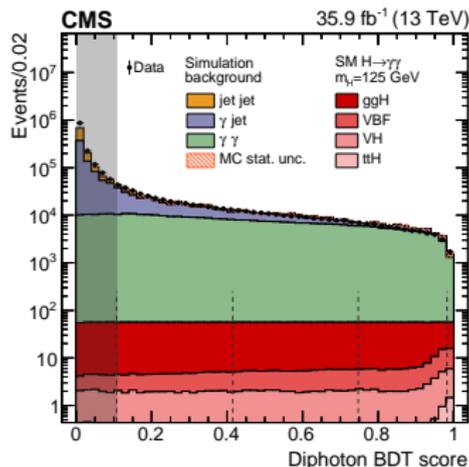
BDT to **assign per-event probability of correct vertex ID**

- BDT score and distance of the three most likely vertices
- number of vertices in the events
- number of conversion tracks
- diphoton p_T



Event selection: diphoton BDT

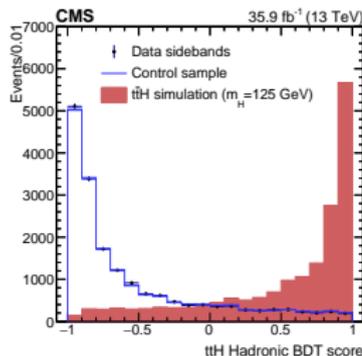
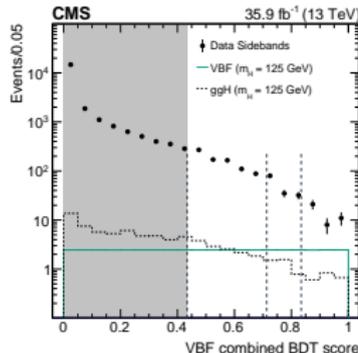
- Combine event information into **one discriminating variable**
 - event kinematics, γ -ID, mass resolution, photon quality
- **High score** for **high resolution** and **high signal-like** topology



- Define cutoff acceptance and set analysis categories if no tagging possible
 - optimized for maximum sensitivity

Tagged categories: VBF, VH, ttH

- **VBF production:** dijet BDT to discriminate against ggH combined with BDT to discriminate between signal and background
- Inputs: jet kinematics for VBF vs. ggH; + diphoton BDT, $p_T^{\gamma\gamma} / \gamma\gamma$ for S vs B
- optimize for sensitivity (3 categories)
- **VH production:** leptons, MET, and jets consistent with decays of W or Z bosons
- **ttH production:** b-jets, leptons, and MET consistent with top pairs
- Tagging of multijet events (hadronic W decay) enhanced by a BDT discriminant



Summary of expected signal yields

(unreadable)

Event categories	Expected SM 125 GeV Higgs boson signal											Bkg		
	Total	ggH	VBF	ttH	bbH	tHq	tHW	WH lep	ZH lep	WH had	ZH had	σ_{eff} (GeV)	σ_{HM} (GeV)	(GeV ⁻¹)
Untagged 0	32.5	72.0 %	16.6 %	2.6 %	0.6 %	0.7 %	0.3 %	0.6 %	0.3 %	4.2 %	2.2 %	1.32	1.26	21.8
Untagged 1	469.3	86.5 %	7.9 %	0.6 %	1.2 %	0.1 %	<0.05 %	0.5 %	0.3 %	1.9 %	1.1 %	1.46	1.32	925.1
Untagged 2	678.3	89.9 %	5.4 %	0.4 %	1.2 %	0.1 %	<0.05 %	0.5 %	0.3 %	1.4 %	0.8 %	1.93	1.67	2391.7
Untagged 3	624.3	91.3 %	4.4 %	0.5 %	1.0 %	0.1 %	<0.05 %	0.5 %	0.3 %	1.2 %	0.7 %	2.61	2.27	4855.1
VBF 0	9.3	15.5 %	83.2 %	0.4 %	0.4 %	0.3 %	<0.05 %	<0.05 %	<0.05 %	0.2 %	<0.05 %	1.52	1.31	1.6
VBF 1	8.0	28.4 %	69.7 %	0.4 %	0.6 %	0.4 %	<0.05 %	0.1 %	<0.05 %	0.3 %	0.1 %	1.66	1.38	3.3
VBF 2	25.2	45.1 %	51.2 %	0.9 %	0.8 %	0.6 %	0.1 %	0.2 %	0.1 %	0.8 %	0.3 %	1.64	1.37	18.9
ttH Hadronic	5.6	7.0 %	0.7 %	81.1 %	2.1 %	4.3 %	2.1 %	0.1 %	0.1 %	0.7 %	1.9 %	1.48	1.30	2.4
ttH Leptonic	3.8	1.5 %	<0.05 %	87.8 %	0.1 %	4.7 %	3.1 %	1.5 %	1.2 %	<0.05 %	<0.05 %	1.60	1.35	1.5
ZH Leptonic	0.5	<0.05 %	<0.05 %	2.6 %	<0.05 %	<0.05 %	0.1 %	<0.05 %	97.3 %	<0.05 %	<0.05 %	1.65	1.43	0.1
WH Leptonic	3.6	1.3 %	0.6 %	5.2 %	0.2 %	3.0 %	0.7 %	84.5 %	4.3 %	0.1 %	0.1 %	1.64	1.43	2.1
VH Leptonic Loose	2.7	8.1 %	2.7 %	2.4 %	0.6 %	1.8 %	0.1 %	64.4 %	19.1 %	0.6 %	0.2 %	1.67	1.56	3.5
VH Hadronic	7.9	47.6 %	4.5 %	4.4 %	0.4 %	1.7 %	0.3 %	0.2 %	0.5 %	25.2 %	15.1 %	1.38	1.30	7.2
VH MET	4.0	18.7 %	2.6 %	15.4 %	0.4 %	2.1 %	1.2 %	26.8 %	30.4 %	1.4 %	0.9 %	1.56	1.39	3.5
Total	1875.0	86.9 %	7.1 %	1.0 %	1.1 %	0.2 %	<0.05 %	0.8 %	0.4 %	1.6 %	0.9 %	1.96	1.62	8237.8

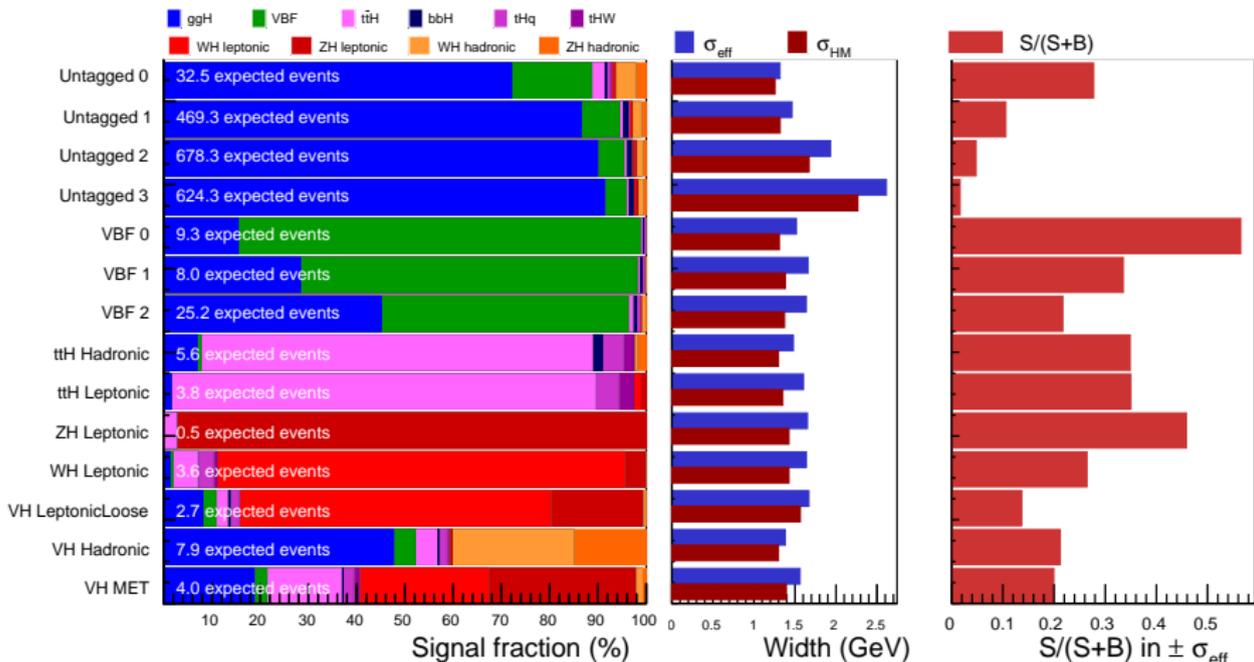
■ Composition of Untagged categories:

- boosted diphoton pairs
- both in the barrel & $R_9 > 0.94$
- both in the barrel
- ...

Summary of expected signal yields

(unwatchable)

CMS Simulation $H \rightarrow \gamma\gamma$



Statistical method

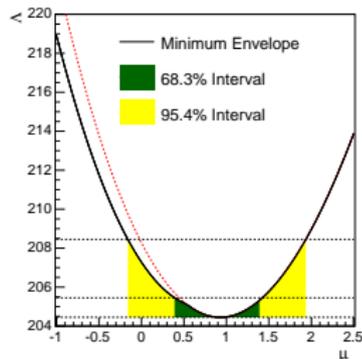
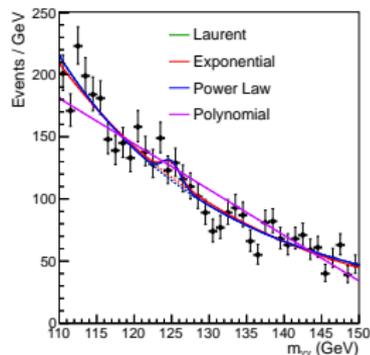
Results extracted by **simultaneous maximum likelihood fit** to the invariant mass distributions of all categories

■ Signal model from simulation

- with systematic uncertainties affecting the shape

■ Background model from data: “discrete profiling” method [JINST 10 (2015) P04015]

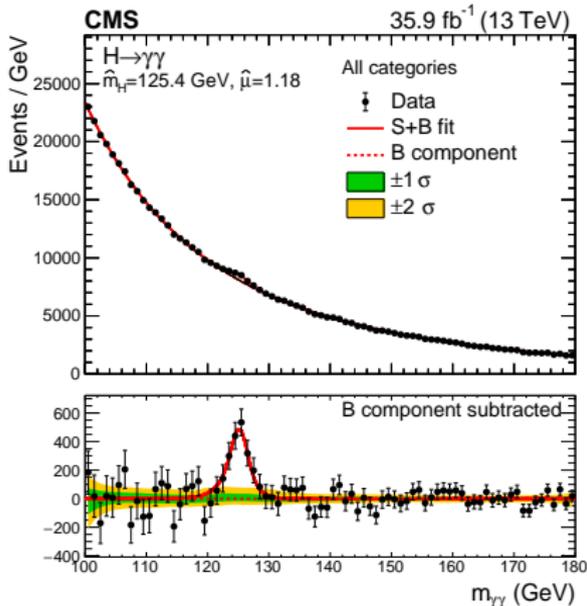
- the max likelihood select the function at the order which fit the best among Bernstein polynomials, Laurent polynomials, power law, exponential families
 - penalty of 1 for each additional parameter of the function
 - condition of fit bias $< 1/5\sigma_{\text{fit}}$ satisfied
- Likelihood “profiled” over the function choice (best fit typically with fewer parameters)



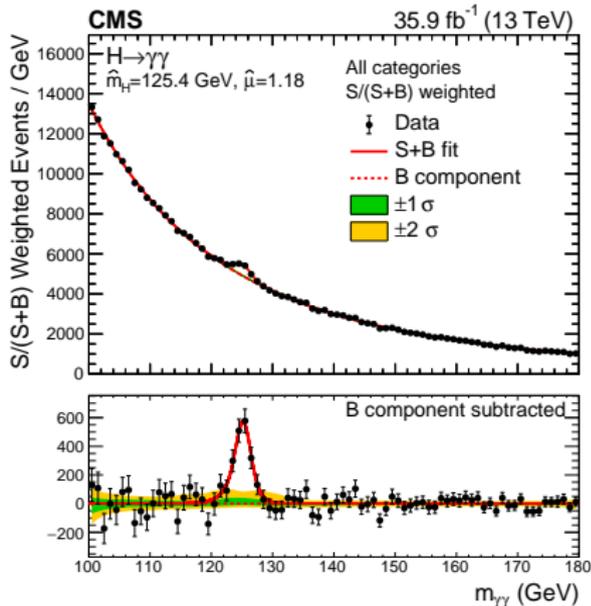
Results

The plots everyone wants to see but nobody really cares about:

All selected events summed

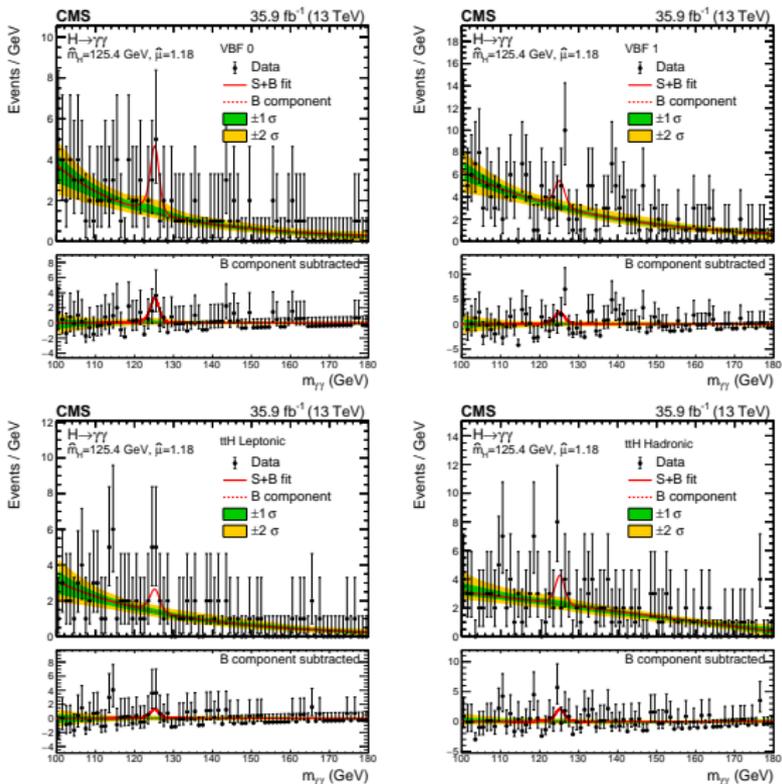


Events weighted by sensitivity



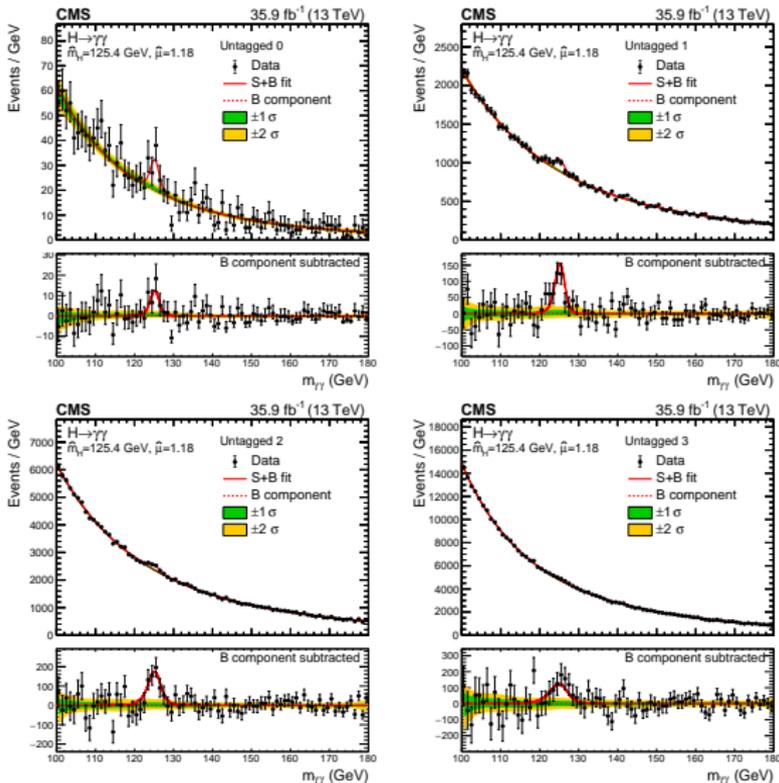
Results: VBF0-1, ttH

The plots nobody wants to see at once, but everybody cares about:



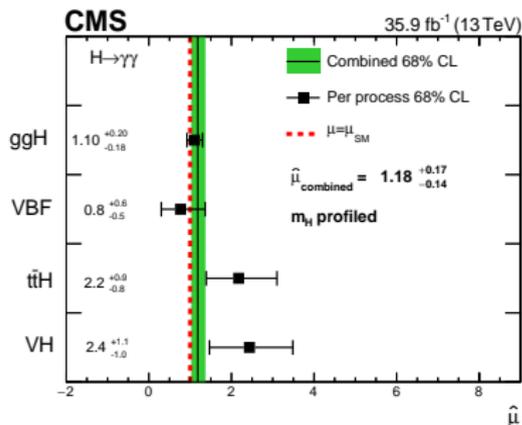
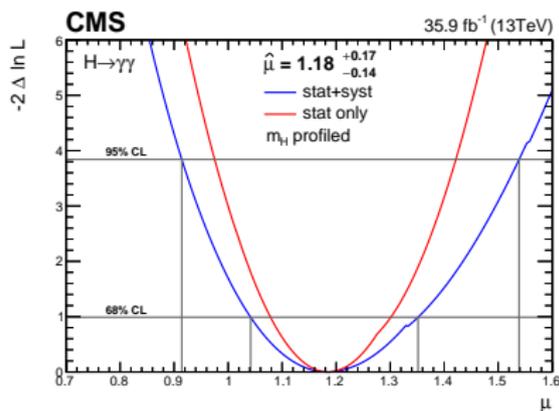
Results: Untagged

The plots nobody wants to see at once, but everybody cares about:



more categories upon request

Results: signal strength modifiers ($\hat{\mu} \cong \sigma/\sigma_{SM}$)



- All categories combined:

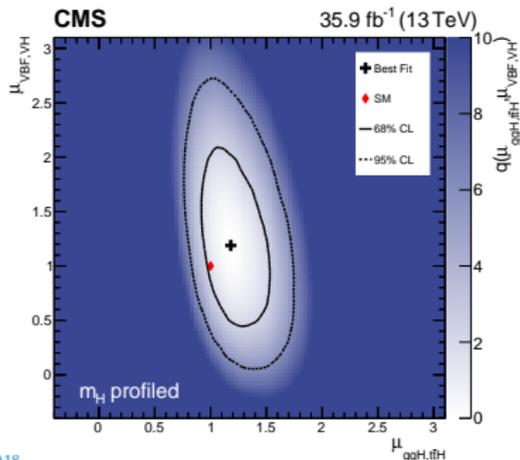
$$\hat{\mu} = 1.18^{+0.12}_{-0.11} (\text{stat})^{+0.09}_{-0.07} (\text{syst})^{+0.07}_{-0.06} (\text{theo})$$

- Fermionic and bosonic production modes:

$$\hat{\mu}_{ggH, ttH} = 1.19^{+0.22}_{-0.18}$$

$$\hat{\mu}_{VBF, VH} = 1.21^{+0.58}_{-0.51}$$

- Mass profiled in the fits



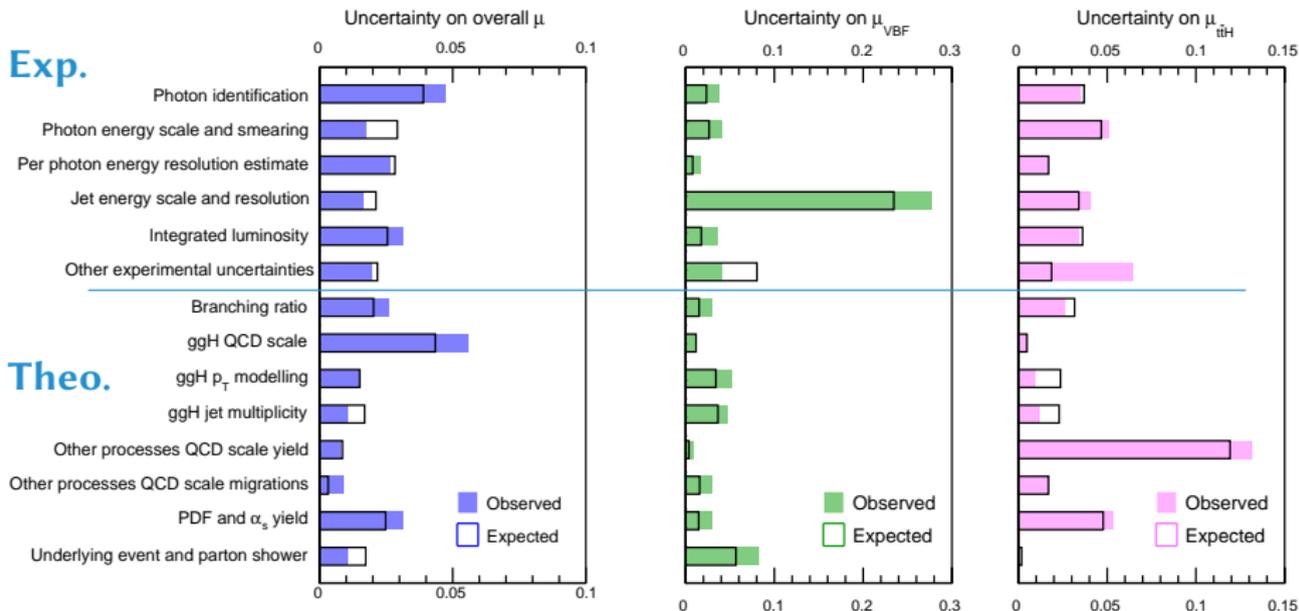
Systematic uncertainties

CMS $H \rightarrow \gamma\gamma$

35.9 fb⁻¹ (13 TeV)

Exp.

Theo.



Experimental uncertainties not far from the theory ones:
compatible by the end of Run2, when $\approx \times 4$ data?

Results: coupling modifiers (κ -framework)

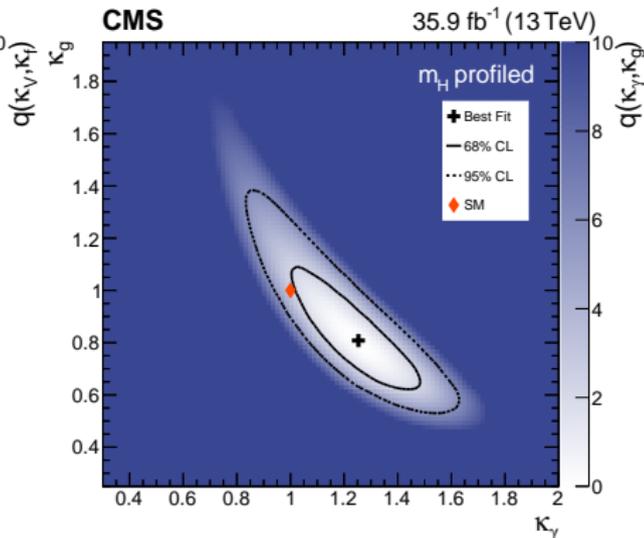
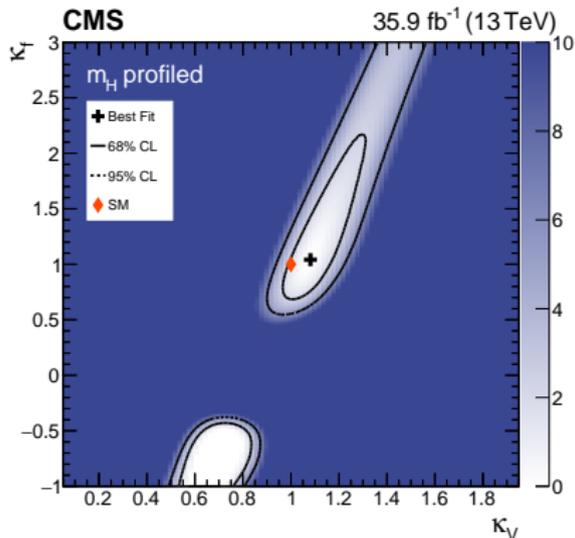
- All fermions scale with κ_f , all vectors scale with κ_V
- No BSM physics in the loops, no invisible decays

⇒ consistency test of the SM



- Allowing BSM physics in the loops, no invisible decays

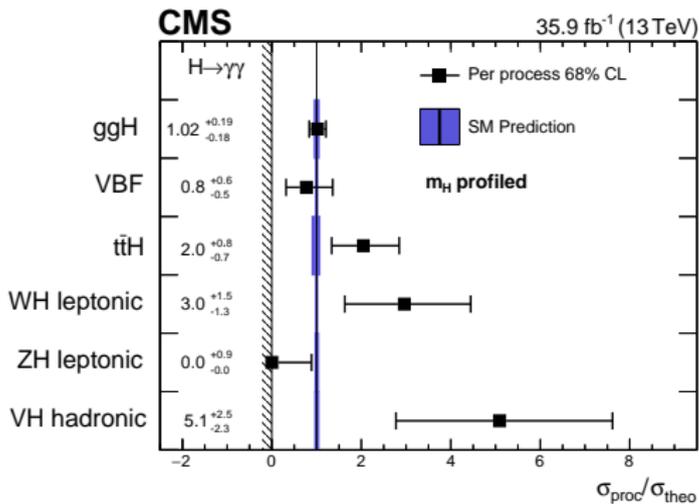
⇒ Probe for new physics



2D likelihood scans, all κ relative to SM expectations, mass profiled in the fit

Simplified template cross-sections (stage 0)

- Evolution of the measurement of the signal strength modifiers
- **Maximize sensitivity** of the measurement
- **Minimize dependence on theory** (and its uncertainties)
- Agreed (among theorists and experimentalists) and simplified phase space division
- **Staged approach**: subsequent stages enabled as sufficient data collected

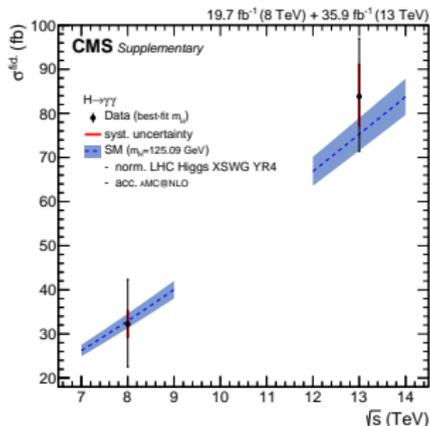
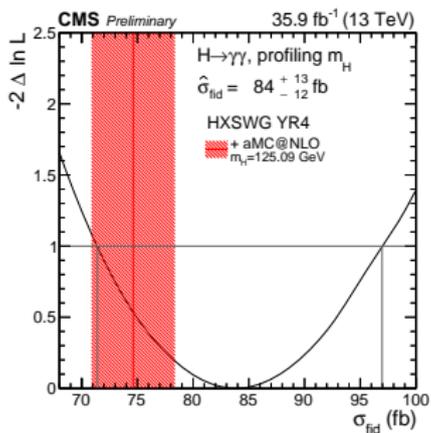


Fiducial cross-sections: inclusive

- Restrict the **phase space as close as possible to the detector acceptance**, with simpler selection and categorization:
 - $p_T^{\gamma^1}/\gamma\gamma > 1/3$; $p_T^{\gamma^2}/\gamma\gamma > 1/4$, $\text{Iso}_{\text{gen},1,2} > 10 \text{ GeV}$, $|\eta_{1,2}| < 2.5$
 - 3 event categories in $\sigma(\gamma\gamma)/\gamma\gamma$, fully decorrelated from $\gamma\gamma$ to prevent shaping of the $\gamma\gamma$ distribution
- Minimal theory dependence and extrapolation**
- Improve longevity of results**, at the price of some sensitivity loss

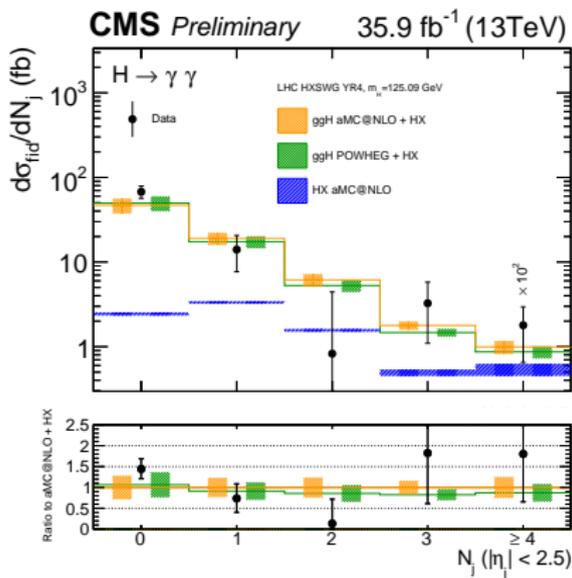
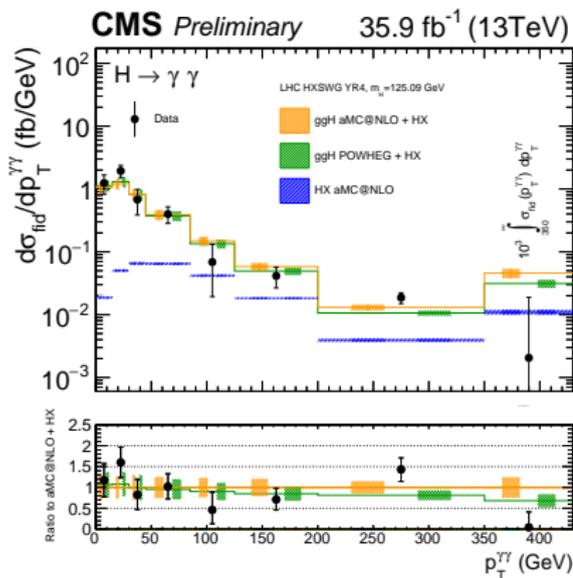
$$\hat{\sigma}_{\text{fiducial}} = 84 \pm 11 \text{ (stat)} \pm 7 \text{ (syst)} \text{ fb}$$

$$\sigma_{\text{fiducial}}^{\text{theory}} = 75_{-4}^{+4} \text{ fb}$$



Fiducial cross-sections: differential

- Diphoton p_T and jet multiplicity differential fiducial cross sections
- General good agreement with SM predictions
- Statistical uncertainty dominant over systematic one



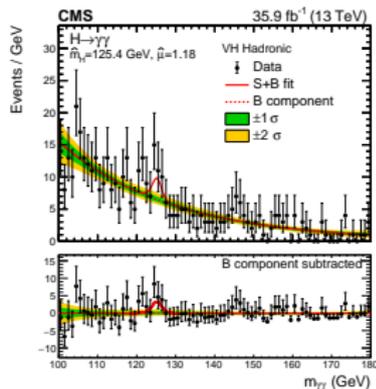
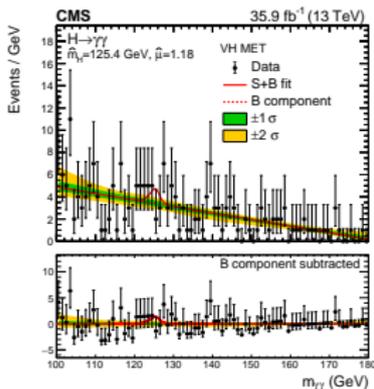
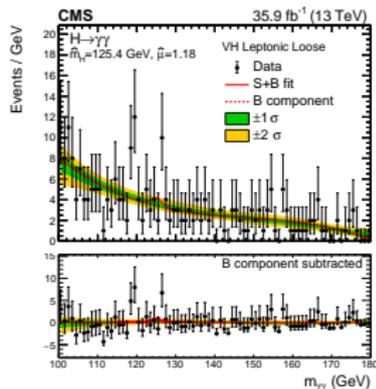
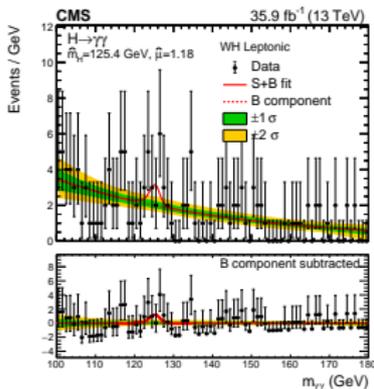
More differential measurements coming out very soon!

Conclusions

- Reported analysis of $H \rightarrow \gamma\gamma$ in CMS
 - signal strengths and coupling modifiers
 - simplified template cross sections (stage 0)
 - fiducial and differential cross sections
- Good compatibility with SM prediction
- LHC delivering a large amount of data (expect $\approx 2016 \times 4$ by the end of this year), which offers a new set of opportunities for **new measurements and stress tests** of the SM

More

Results: VH (but one, next slide)



Results: VBF2, ZH leptonic

