HIGGS DECAYING INTO PHOTONS MEASUREMENTS @CMS

Yacine Haddad
Imperial College London / IPPP Durham

HiggsTools Final Meeting
12-09-2017 IPPP Durham

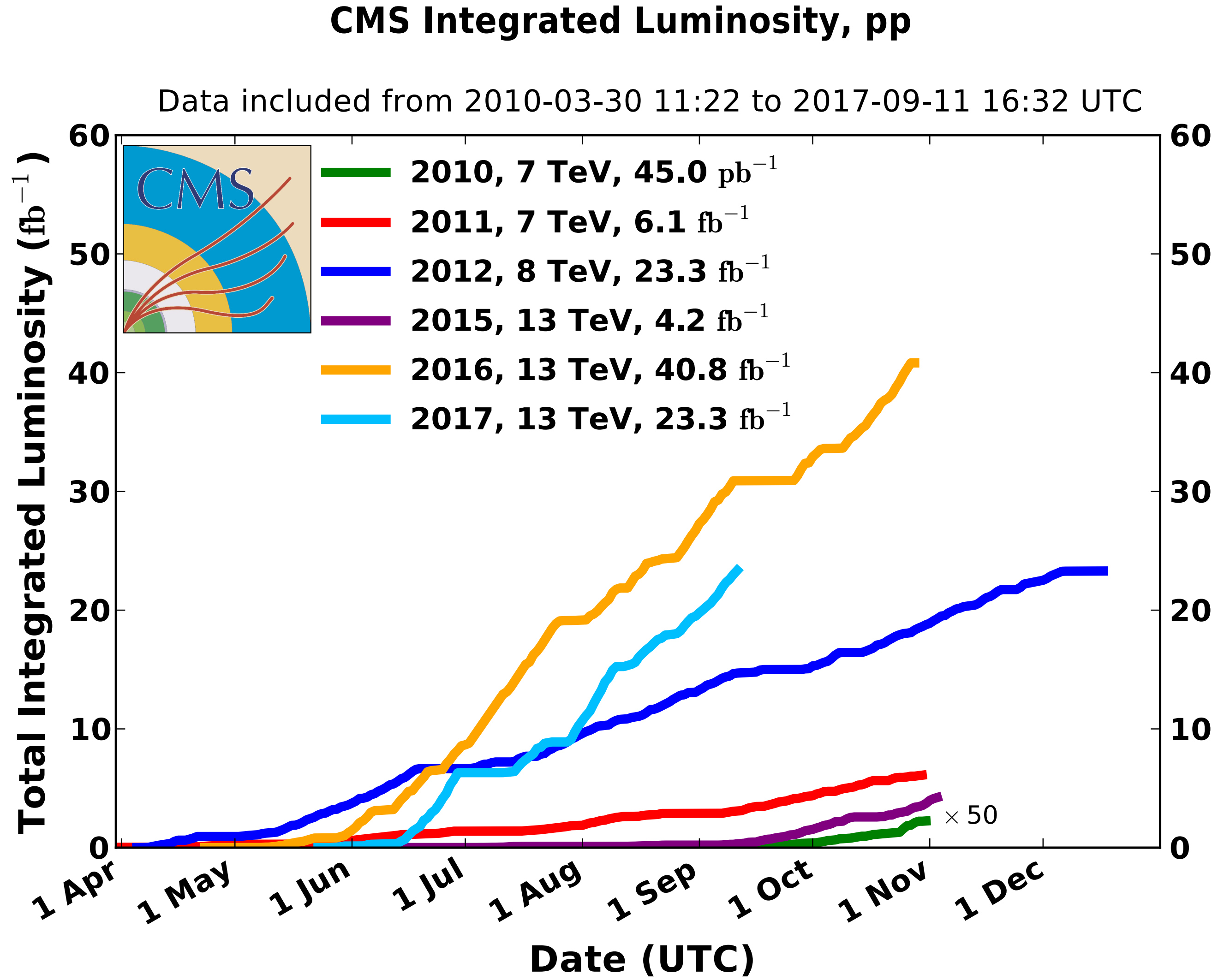
INTRODUCTION

- At the LHC, the $H \rightarrow \gamma\gamma$ channel plays a key role first in the discovery of the Higgs boson, and then in the measurements of Higgs boson properties and also in searches for new physics
  - Good sensitivity: excellent mass resolution, clean signature
  - Sensitive to new physics: loop-induced process
- Today : recent Run 2 @13 TeV CMS results from $H \rightarrow \gamma\gamma$ channel
  - Coupling and mass measurements : **CMS-PAS-HIG-16-040**
  - Fiducial/differential measurements : **CMS-PAS-HIG-17-015**

Not intend to summarise all the important results, apologies if your favourite topic is missing!

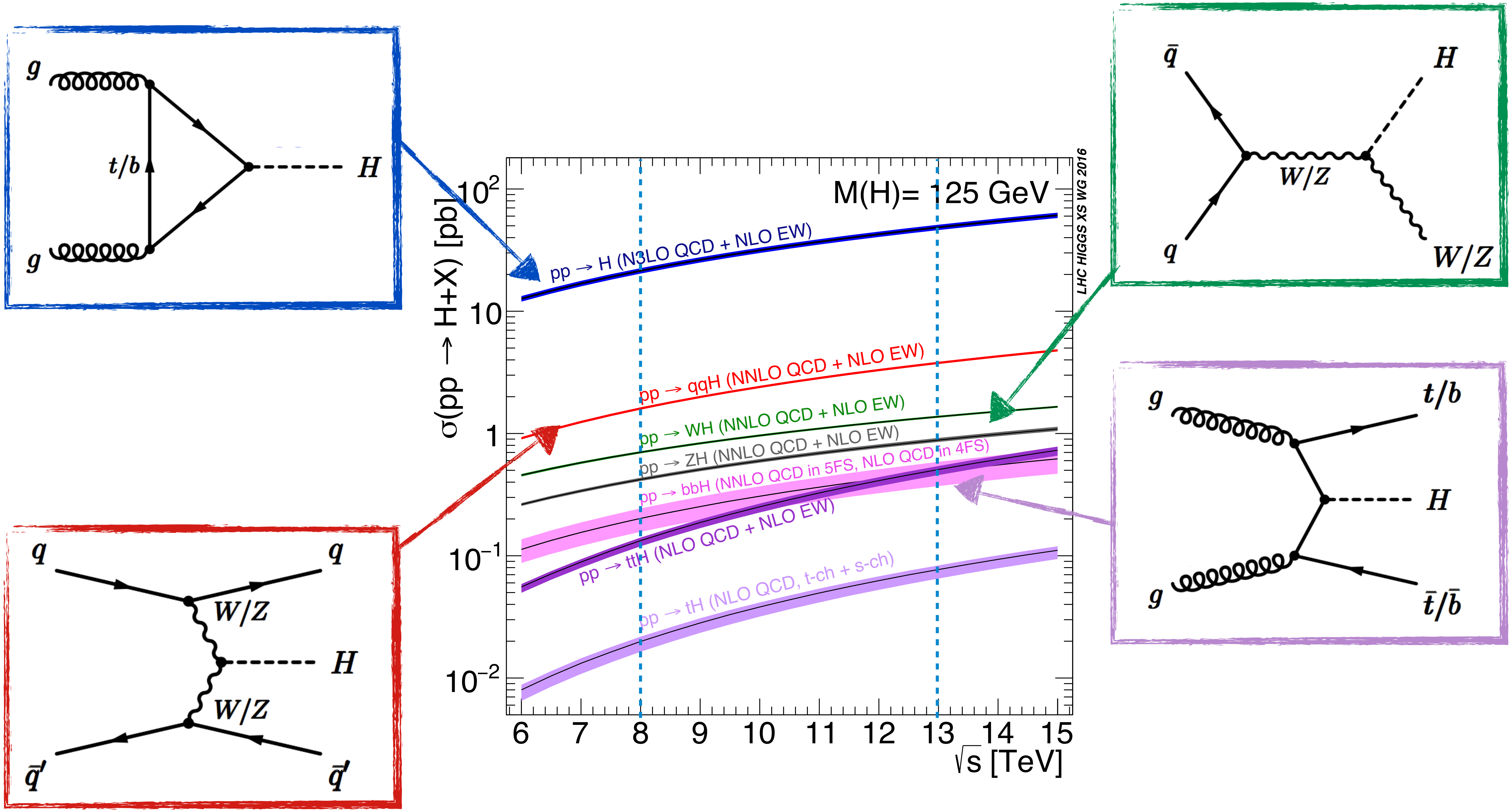
MORE AND MORE DATA

	√s=7 TeV	√s=8 TeV	√s=13 TeV	Ratio 13/8 TeV
ggH	15.3 pb	19.4 pb	44.1 pb	**2.27**
VBF	1.25 pb	1.6 pb	3.8 pb	**2.38**
ttH	88.6 fb	133 fb	507 fb	**3.81**
tt	177 pb	253 pb	832 pb	**3.29**



- Run II of the LHC at 13 TeV is now making a large sample of Higgs boson event available for analysis
  - 2016 was an excellent year for data with 37.82 fb^{-1} collected with CMS 0.6 fb^{-1} in one day ($L=1.4\times10^{34}$ cm^{-2} s^{-1})
- These data are increasing and opening up new channels to study the properties of Higgs boson and the SM particles

SM HIGGS PRODUCTION @LHC



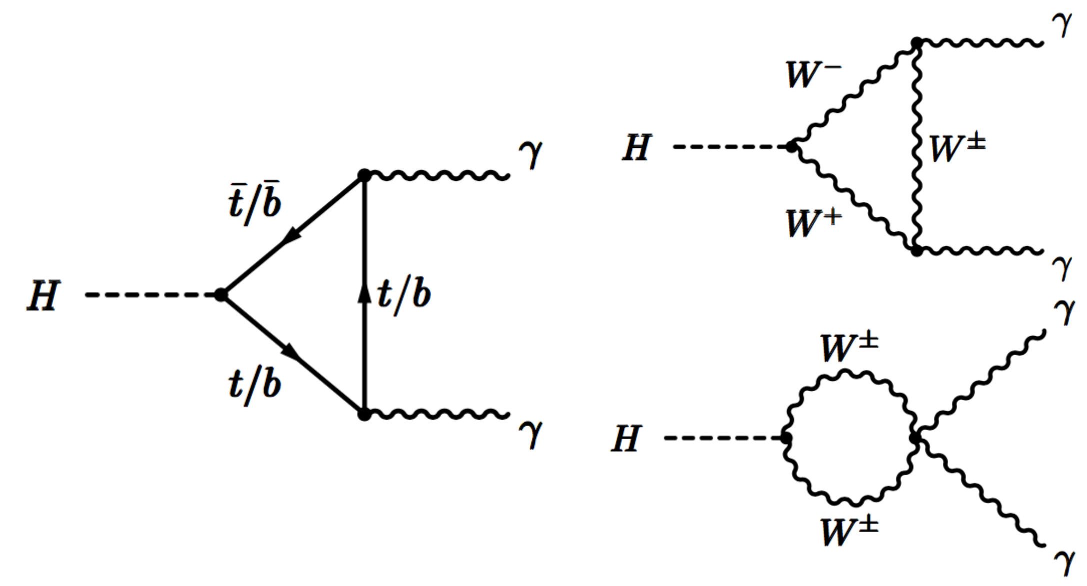
- H→γγ gives access to all the production modes

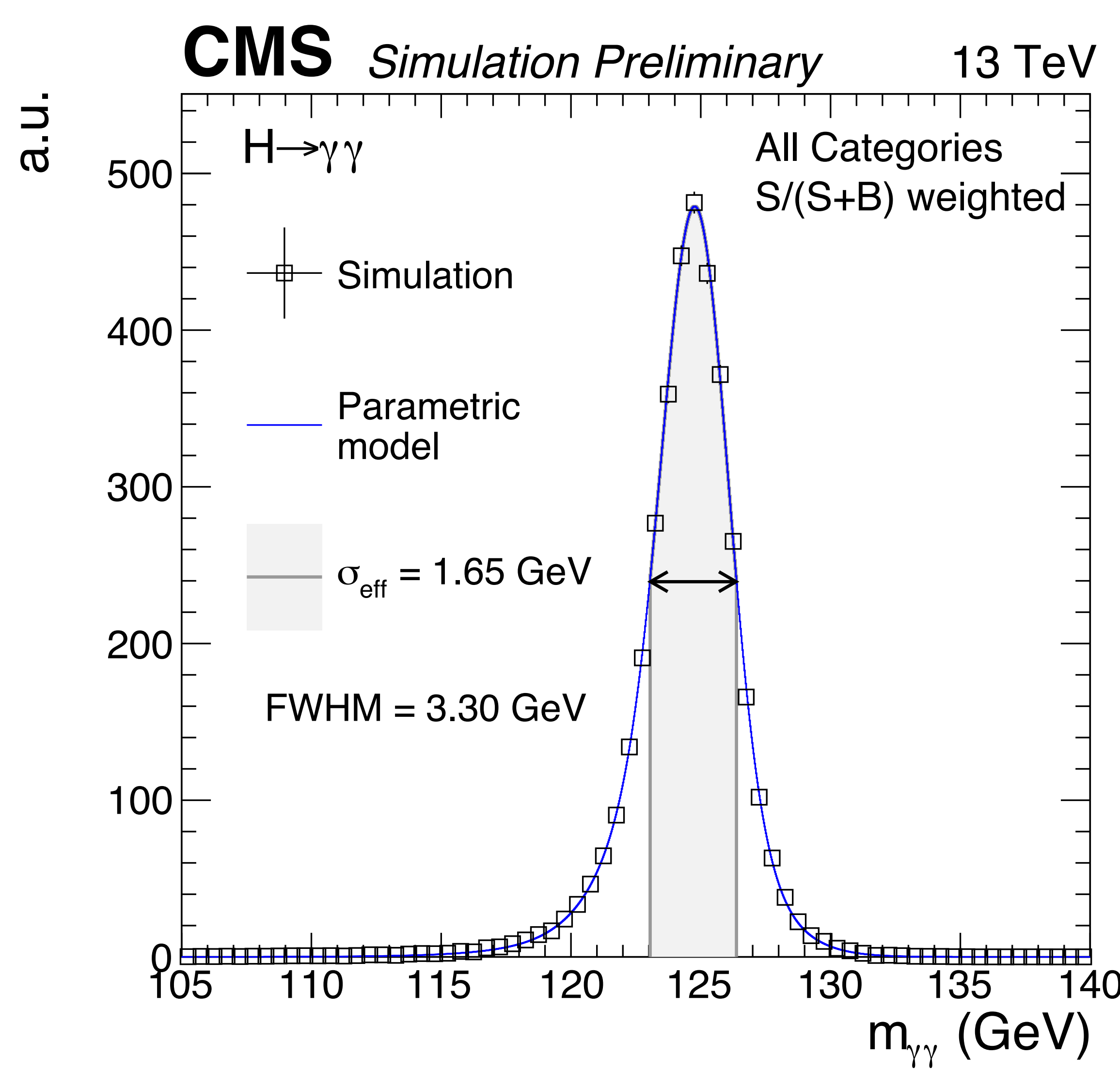
DECAY INTO DIPHOTON : H→γγ

- **Loop-induced decay**
  - Interference helps probe sign of couplings to SM particles
  - New physics could contribute to the loop
- Small branching fraction (0.2%), large background, **excellent mass resolution** (1-2%)
- Mass distribution expected to be determined by detector resolution

$$m^2_{\gamma\gamma} = \boxed{E_{\gamma_1} E_{\gamma_2}}(1 - \boxed{\cos\alpha})$$

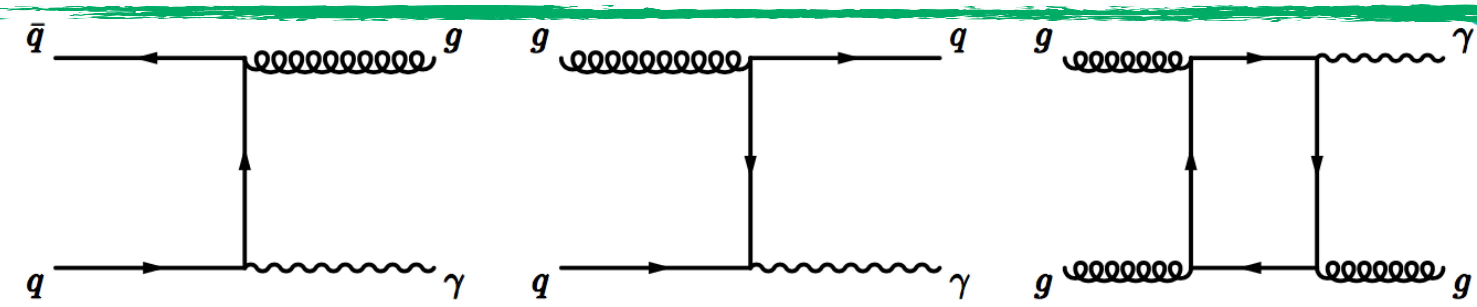
- Requires a good **energy reconstruction** and **correct vertex assignment**



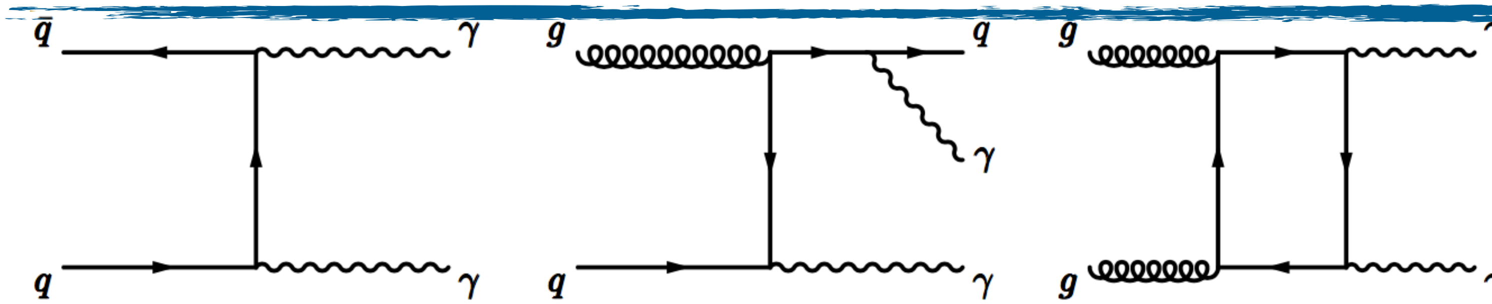


BACKGROUND COMPOSITION

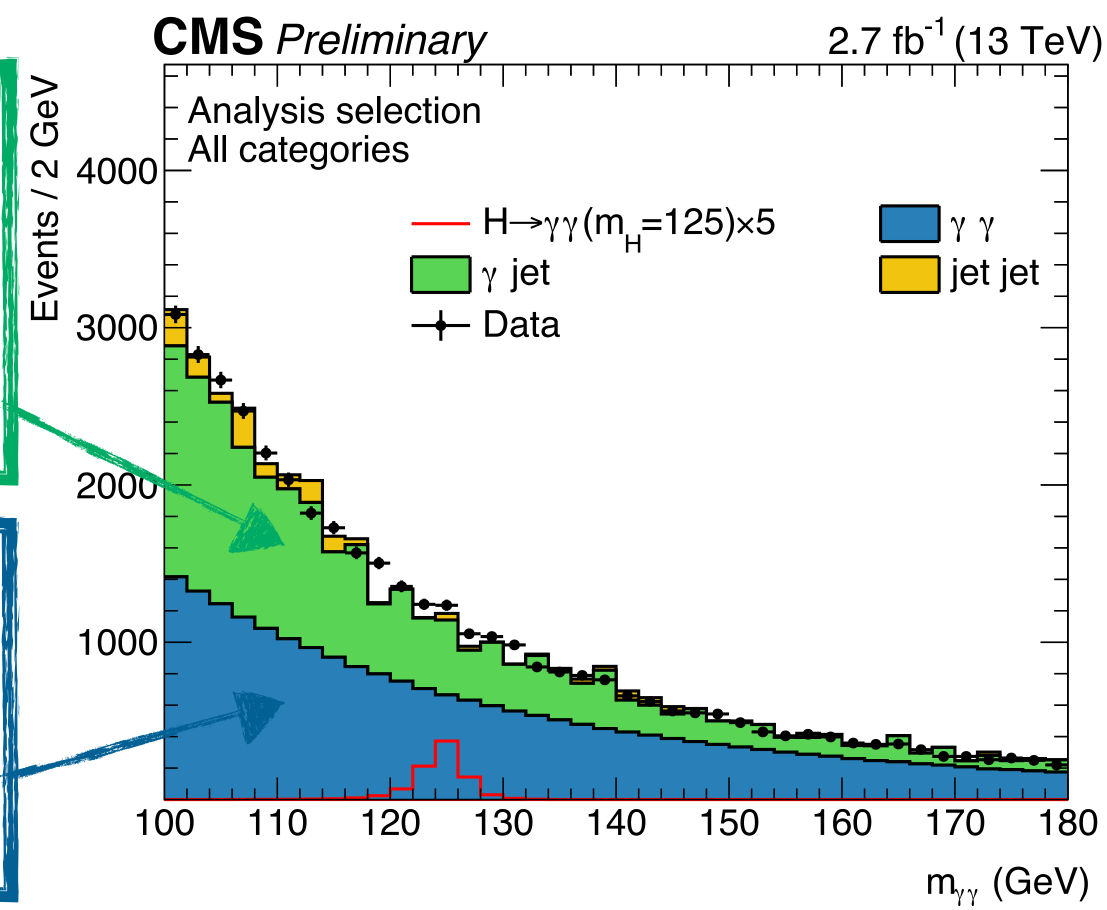
jj: O(500 μb)   *γj*: O(200 nb)   *γγ*: O(30 pb)



Reducible background: SM photon+jet process



Irreducible background: SM diphoton process



- **Small peaking signal on large QCD falling background**
  - irreducible γγ, reducible γ-fake and fake-fake
  - multivariate Photon ID to reject fakes
- Background model **data-driven**:
  - background functional form treated as **discrete nuisance** parameter
- **Event Tagging:** events are sorted into 14 categories depending on Higgs production modes and kinematics, to improve the analysis sensitivity

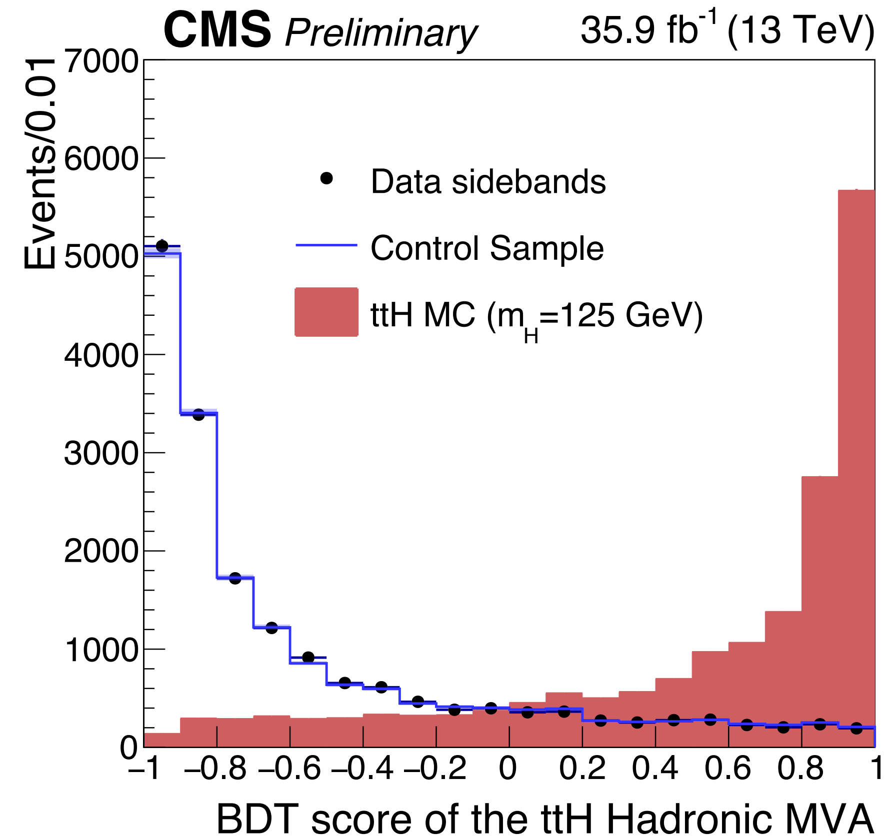
H→γγ : CATEGORISATION

BY PRODUCTION

ttH (leptonic + hadronic)

- (sub-)lead-photon pT/$m_{\gamma\gamma}$ >1/2(1/4), at least one lepton ($\ell=\mu,e$) away from Z peak
- at least two jets with pT>25GeV, |η|<2.5
- at least one of the jet is b-tag

- At least 3 jets + 1 b-jet
- Train an MVA on MC ttH vs MC diphoton using the input variables : N_{jets}, lead b-tag , sub-lead b-tag, lead p_T



VH leptonic	VH Hadronic	VH MET
• W → ℓν or Z→ ℓℓ • (sub-)lead-photon pT/$m_{\gamma\gamma}$ >1/2(1/4) • at least one lepton ($\ell=\mu,e$) away from Z peak • ΔR(γ,μ (e)) > 0.5 (0.2) • diphoton MVA > 0.5 • MET > 45 GEV (WH leptonic)	• W → jj or W → jj • (sub-)lead-photon pT/$m_{\gamma\gamma}$ >1/2(1/4) • at least two jets • p_T > 40 GeV, |η|<2.4, |cosθ*| < 0.5 • 60<m_{jj} < 120 GeV	• W → ℓν or W → νν • MET > 85 GeV • ΔΦ(γγ,MET) > 2.4 • diphoton MVA > 0.79

ttH (LEP)
VH (LEP)
ttH (HAD)
VBF
VH MET
VH (HAD)
Untagged

Imperial College London

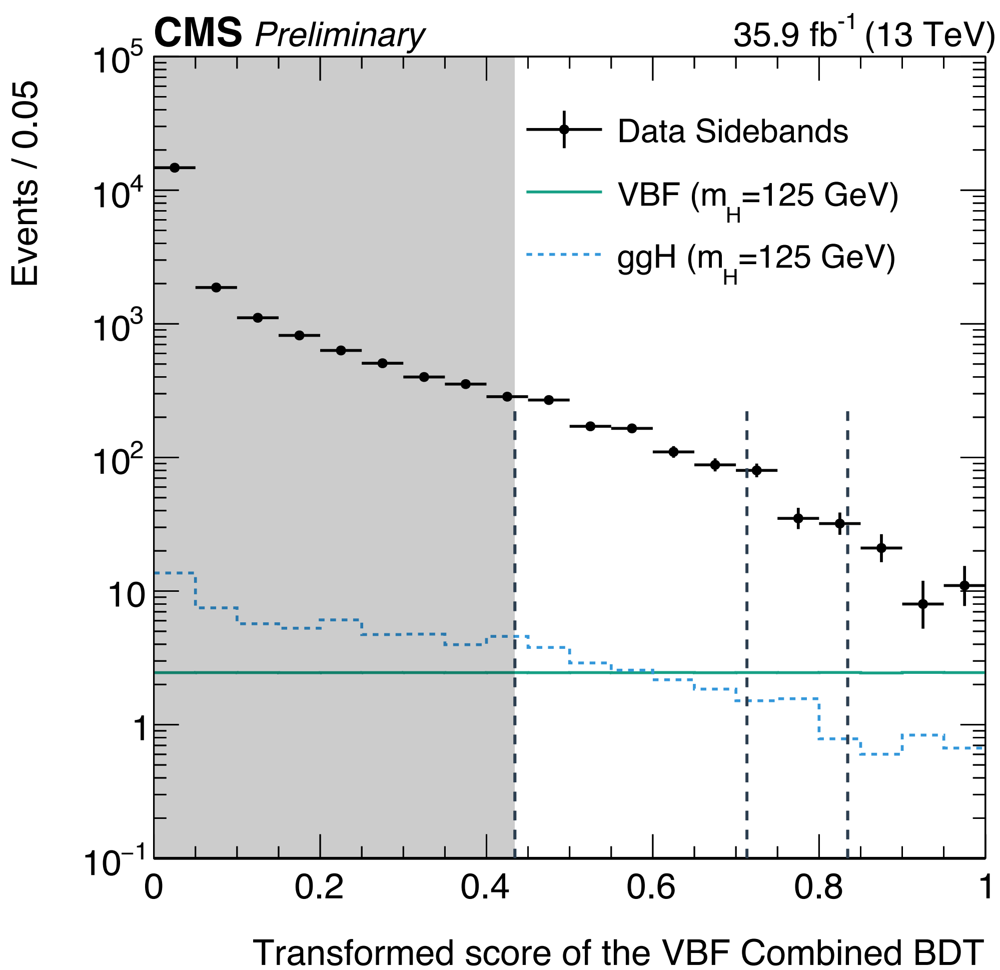
* *Category priority: Tags are prioritised in order of S/B*

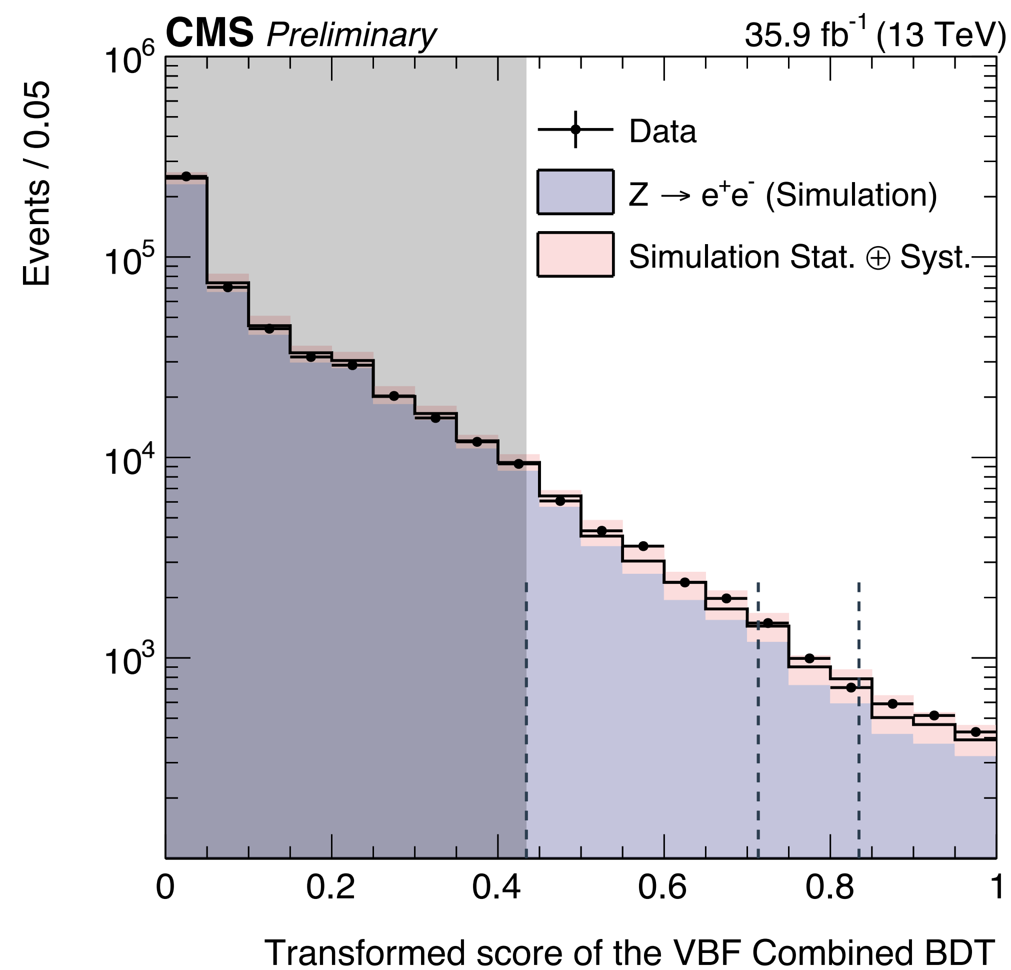
H→γγ : CATEGORISATION

BY PRODUCTION

VBF (0–1)

- Require at least 2 jets with $p_{T1} > 30$GeV, $p_{T2} > 20$ GeV, $|\eta| < 4.7$, $m_{jj} > 250$ GeV
- A diphoton pair with (sub)lead $p_T/m_{\gamma\gamma} > 1/2(1/4)$
- Construct a BDT to identify VBF dijet-like events using:
  - $p_T/m_{\gamma\gamma}$ of both photons, p_T of both jets, m_{jj}, $\Delta\eta_{jj}$, centrality variable, $\Delta\Phi jj,\gamma\gamma$, $\Delta R\gamma j$, $\Delta\Phi jj$
- Final VBF classification combines dijet BDT with BDT estimating diphoton quality (see next slide)
- 3 VBF categories are then defined by sensitivity (VBF tag 0-1-2)

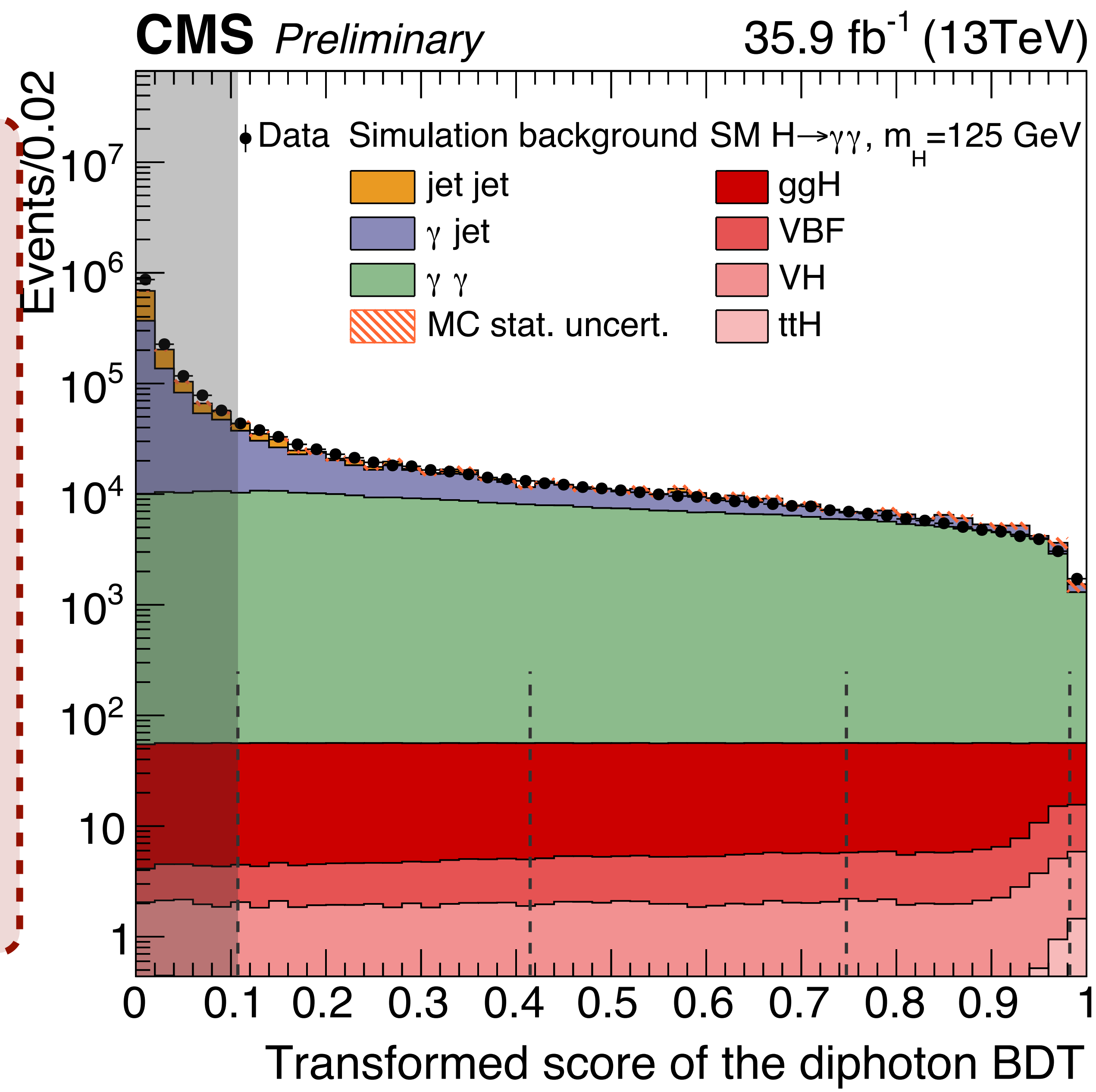


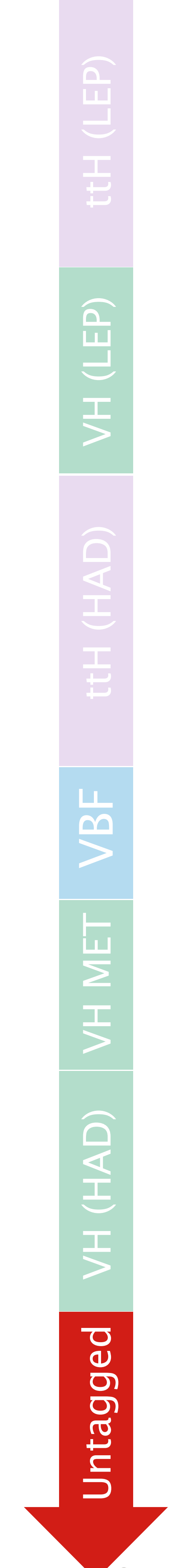


* *Category priority: Tags are prioritised in order of S/B*

H→γγ : CATEGORISATION

- Remaining events fall into the untagged category
- Construct MVA to select diphoton pairs with signal like kinematics, high photon ID score and good mass resolution
- Split events into categories based on output of classifier exploiting S/B ratios and mass resolution
- 4 untagged categories → **14 non-overlapping categories in total**

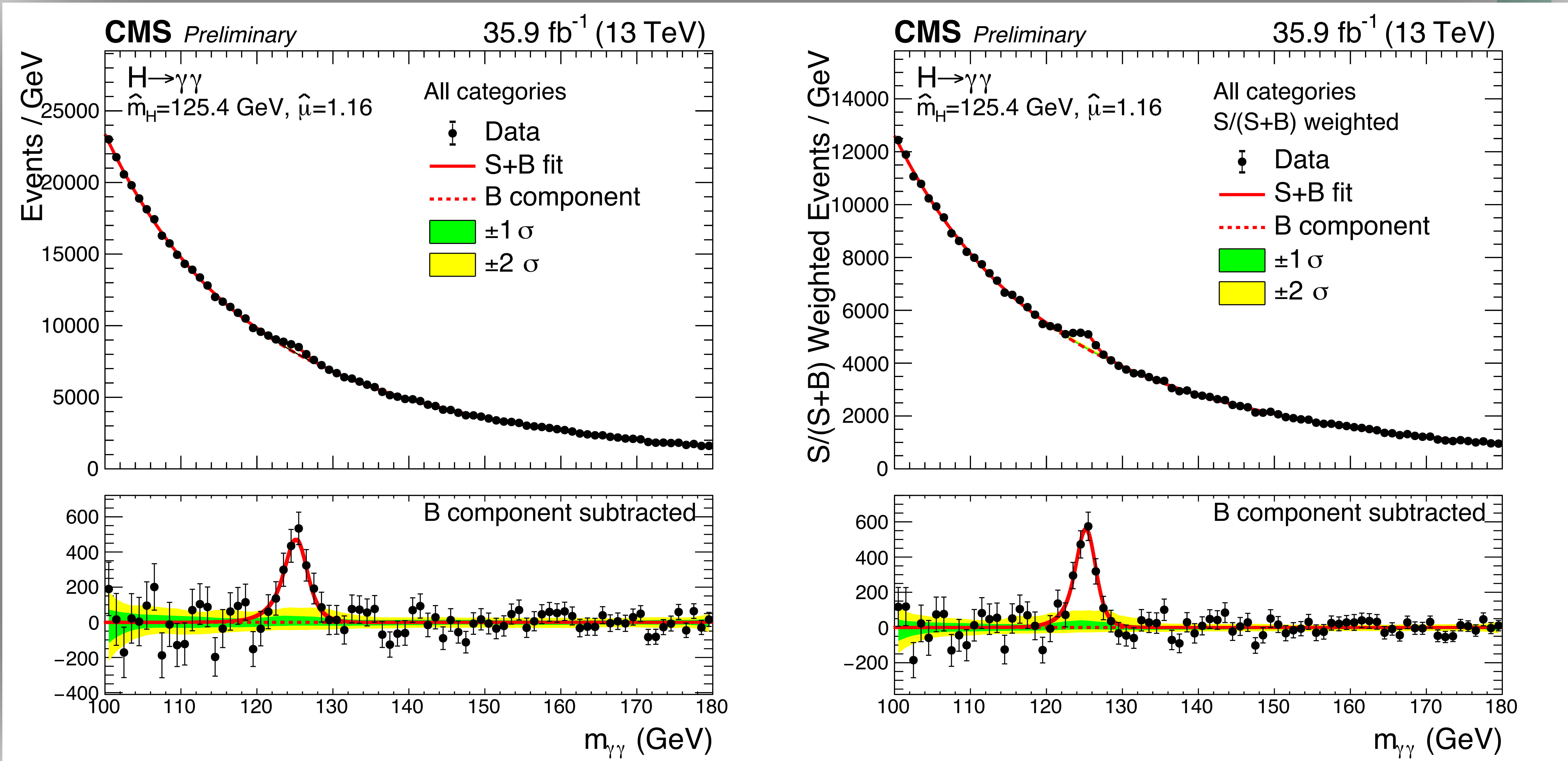




** Category priority: Tags are prioritised in order of S/B*

Imperial College London

H→γγ : CATEGORISATION



ttH (LEP)

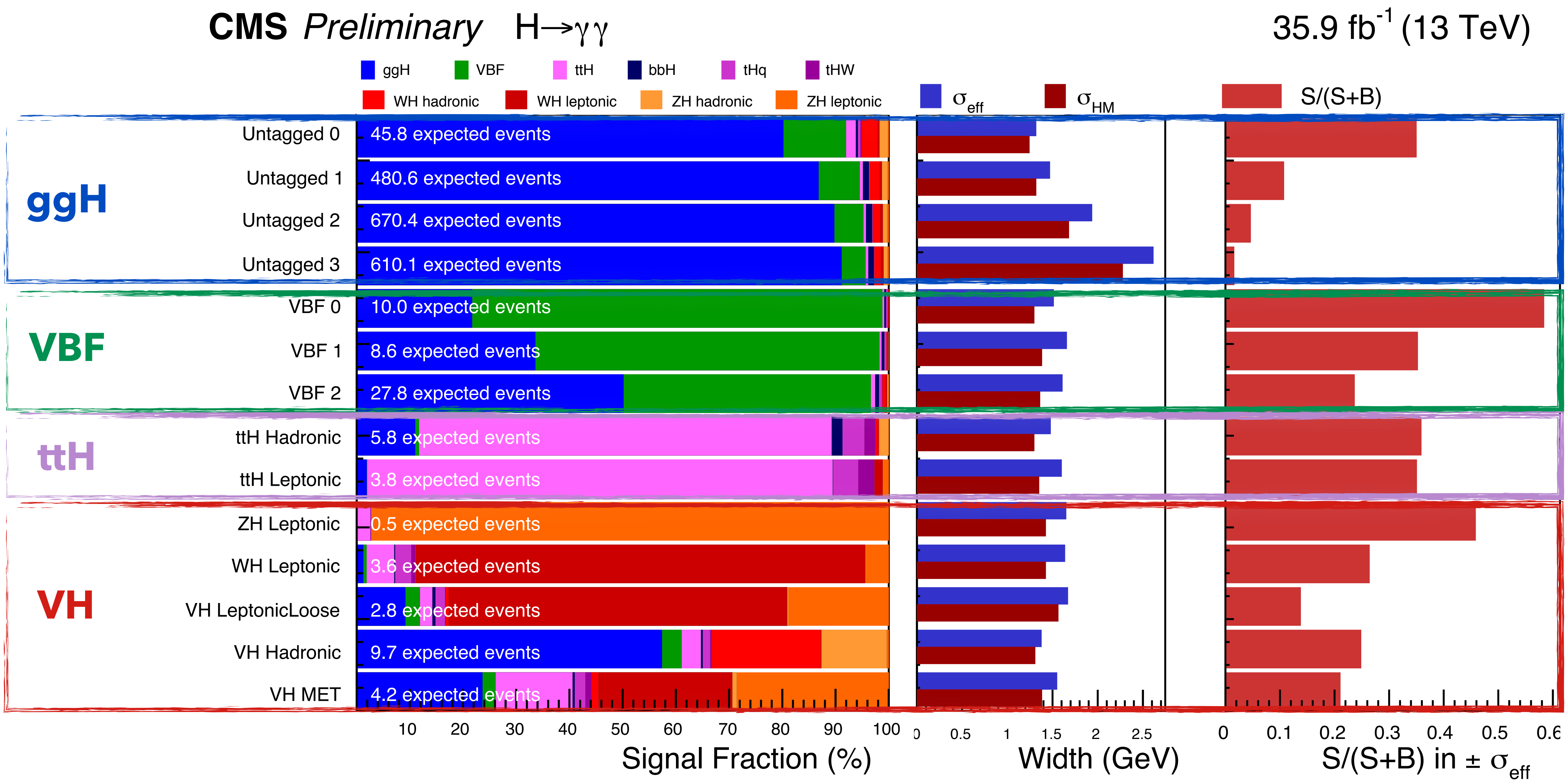
Untagged

Imperial College London

* *Category priority: Tags are prioritised in order of S/B*

MEASUREMENTS

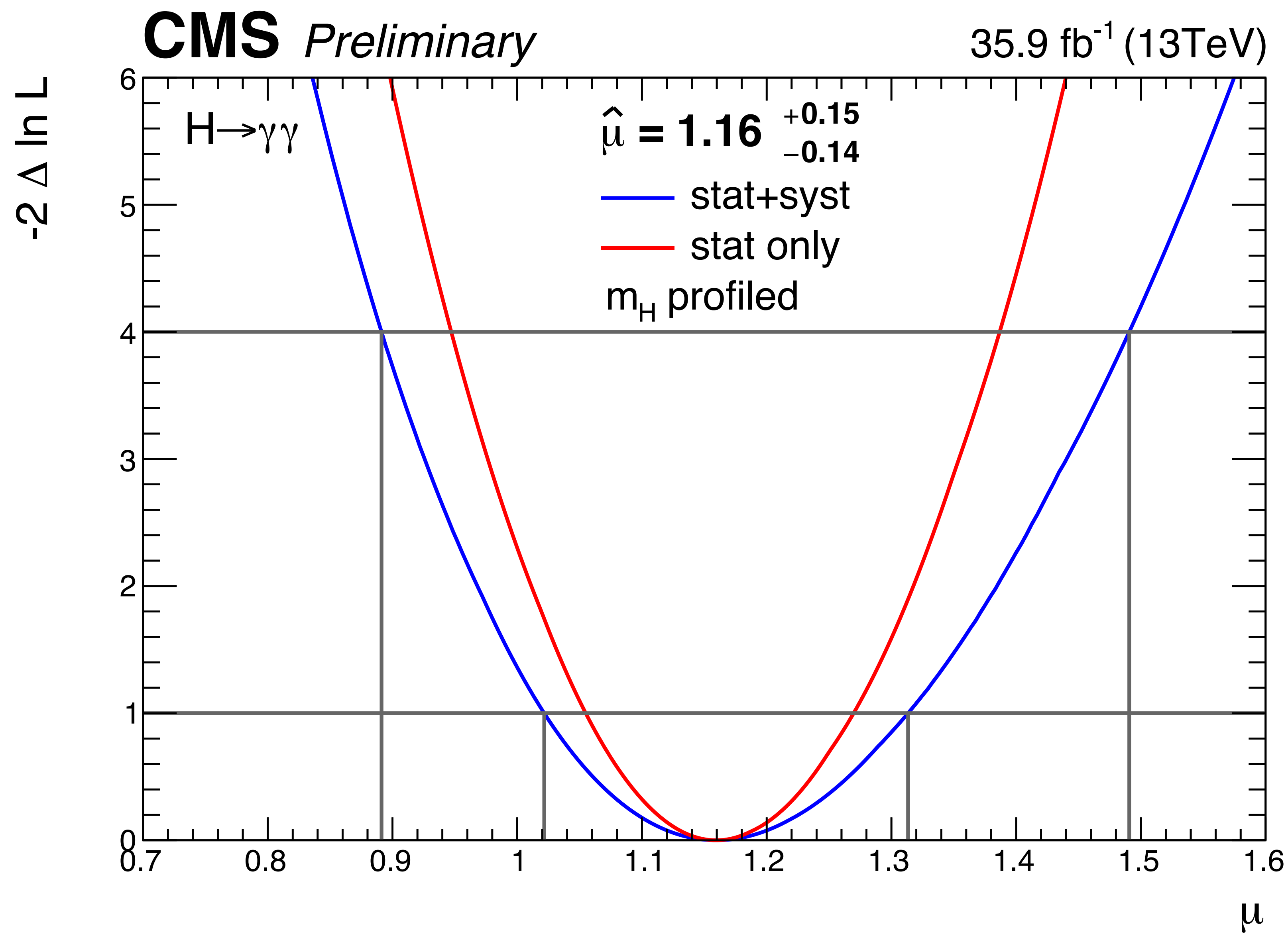
H→γγ : PRODUCTION MODIFIER



- Divide inclusive data samples into categories
  - Isolate different production modes
  - Improve sensitivity by introduction variety in S/B and resolution

MEASURING THE SIGNAL STRENGTH

- **Overall signal strength is consistent with standard model**

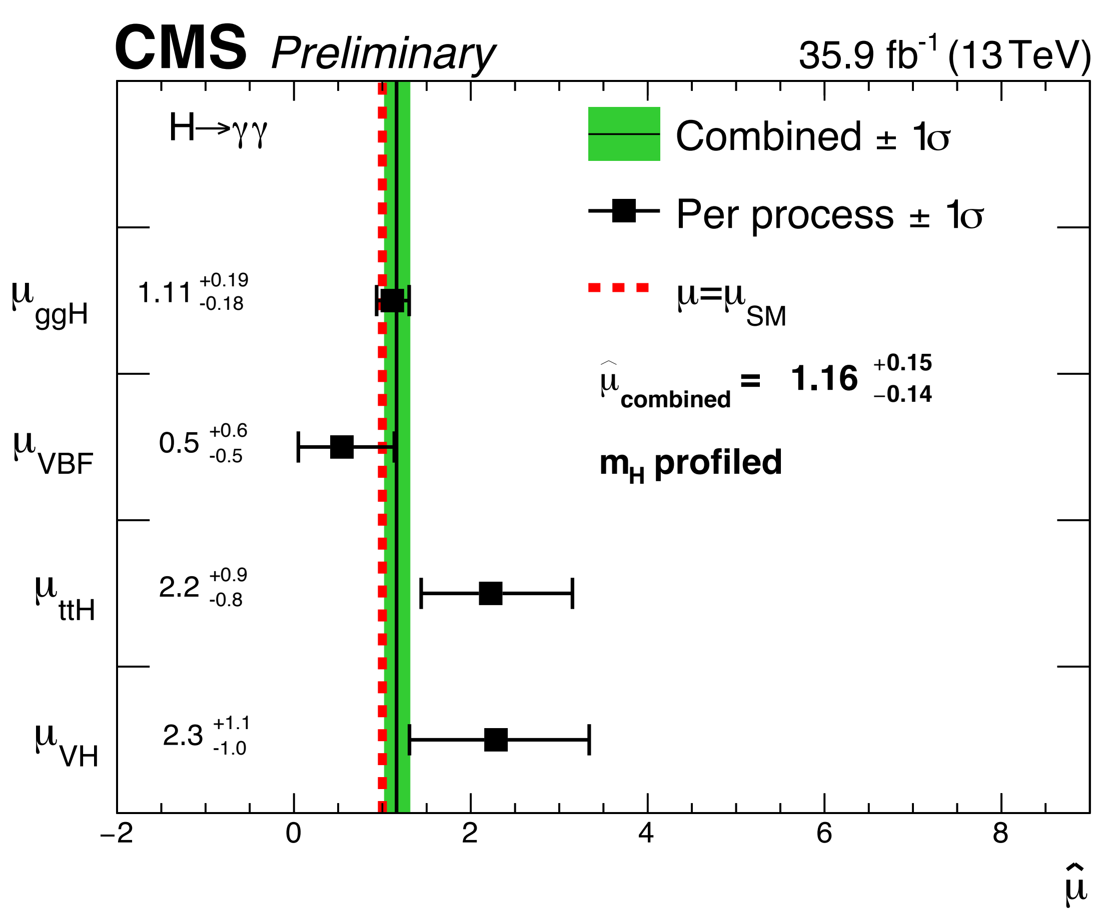


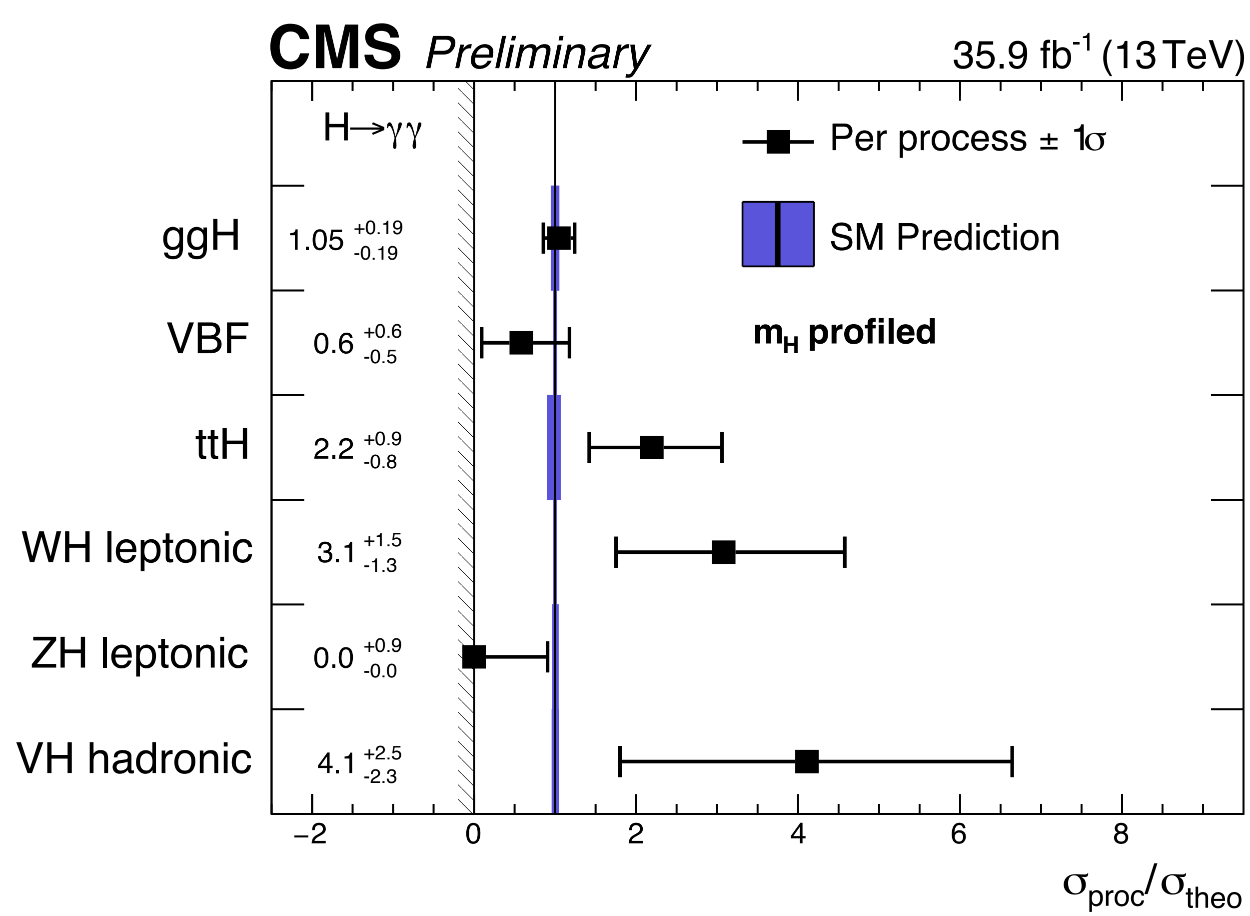
$$\mu = \sigma/\sigma_{SM} = 1.16_{-0.10}^{+0.11}(\text{stat.})_{-0.08}^{+0.09}(\text{syst.})_{-0.05}^{+0.06}(\text{theo.})$$

MEASURING THE SIGNAL STRENGTH

- **A likelihood scan of the signal strength is performed, profiling all other nuisances including the Higgs mass.** In addition, cross section ratios for each process are measured in the Stage 0 Higgs STXS framework

$$\mu = \sigma/\sigma_{SM} = 1.16_{-0.10}^{+0.11}(\text{stat.})_{-0.08}^{+0.09}(\text{syst.})_{-0.05}^{+0.06}(\text{theo.})$$

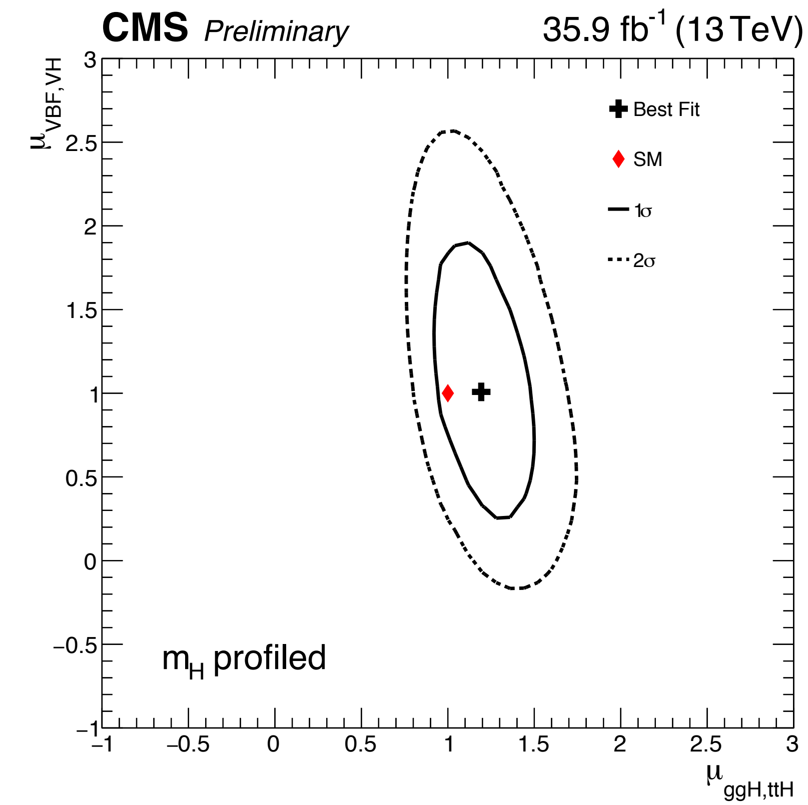


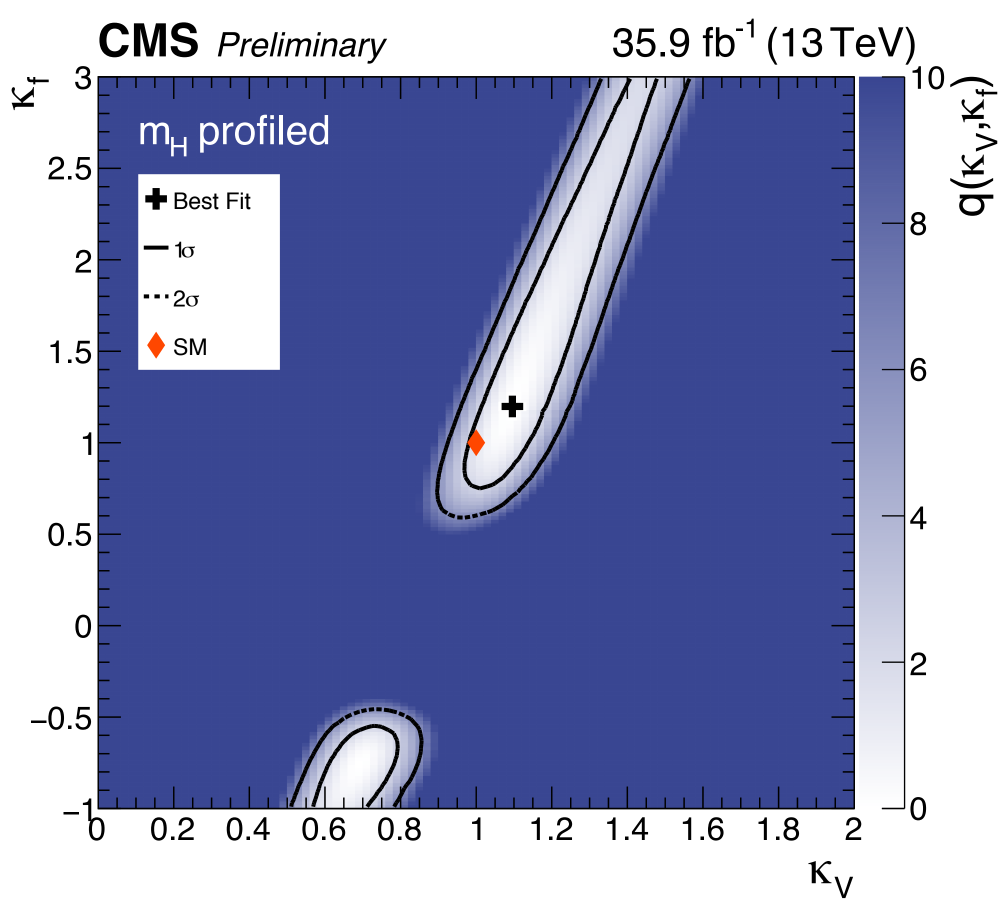


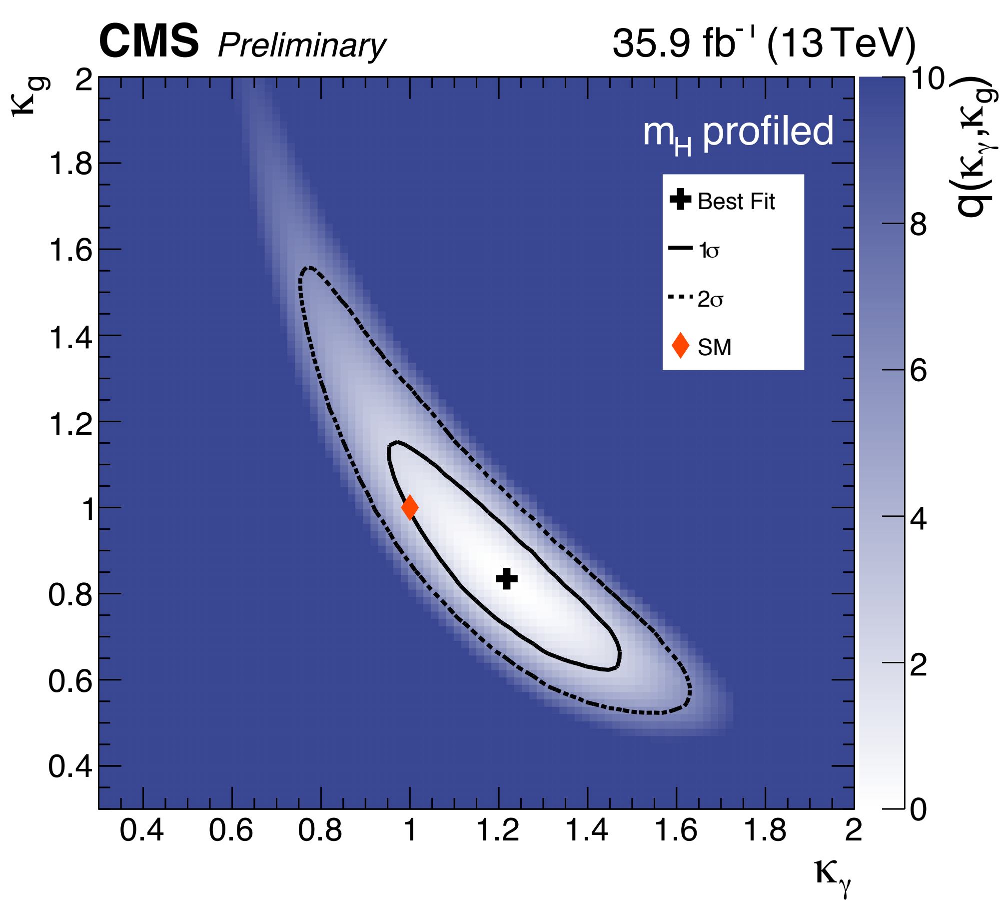
MEASURING THE SIGNAL STRENGTH

2D LIKELIHOOD SCANS

- **A Signal strength** for fermionic production modes (ggH and ttH) vs. signal strength for vector boson production modes (VBF, ZH, WH)
- **Coupling modifiers** to gluons and photons k_g vs. k_γ (left) and coupling modifiers to fermions and bosons k_f vs. k_V (right)







DIFFERENTIAL FIDUCIAL CROSS SECTIONS

INTRODUCTION

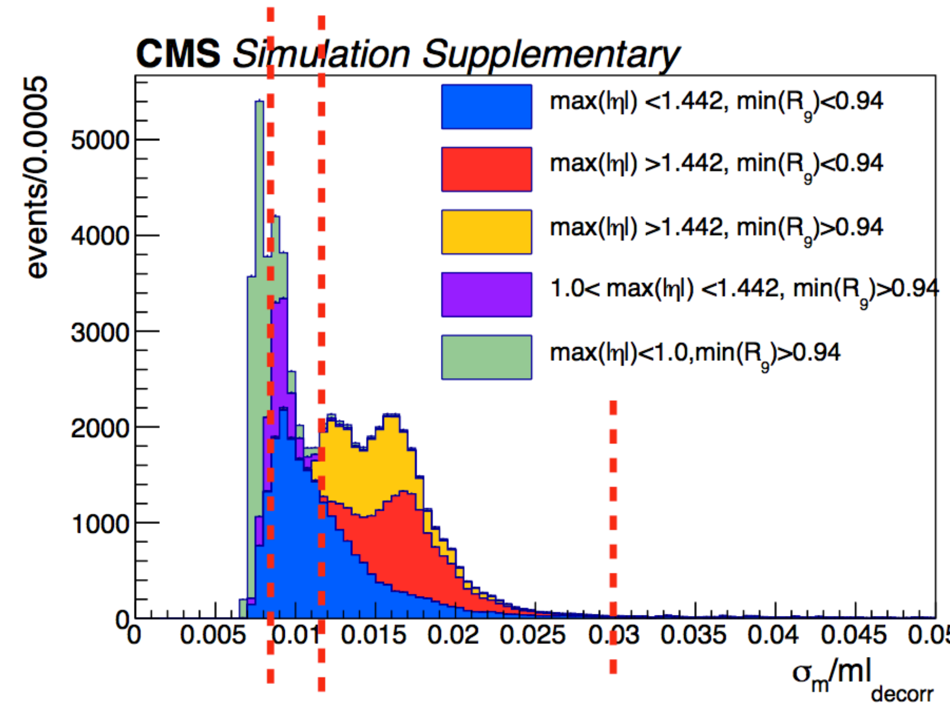
- **A Signal strengths** provide measurement with best sensitivity, but short lifetime:
  - Pre-assumptions on signal + non trivial (re-)interpretation as cross section due to complicate acceptance selection
- **Differentail fiducial XS:**
  - Test SM predictions for full spectrum of observables of interest
  - Probe for hints of BSM physics
  - Different observables are sensitive to different Higgs properties.
  - Restriction to a phase space as close as possible to the detector acceptance, simpler signal selection and categorisation for minimal theory dependence and extrapolation

DIFFERENTIAL FIDUCIAL CROSS SECTIONS

CMS-HIG-16-020

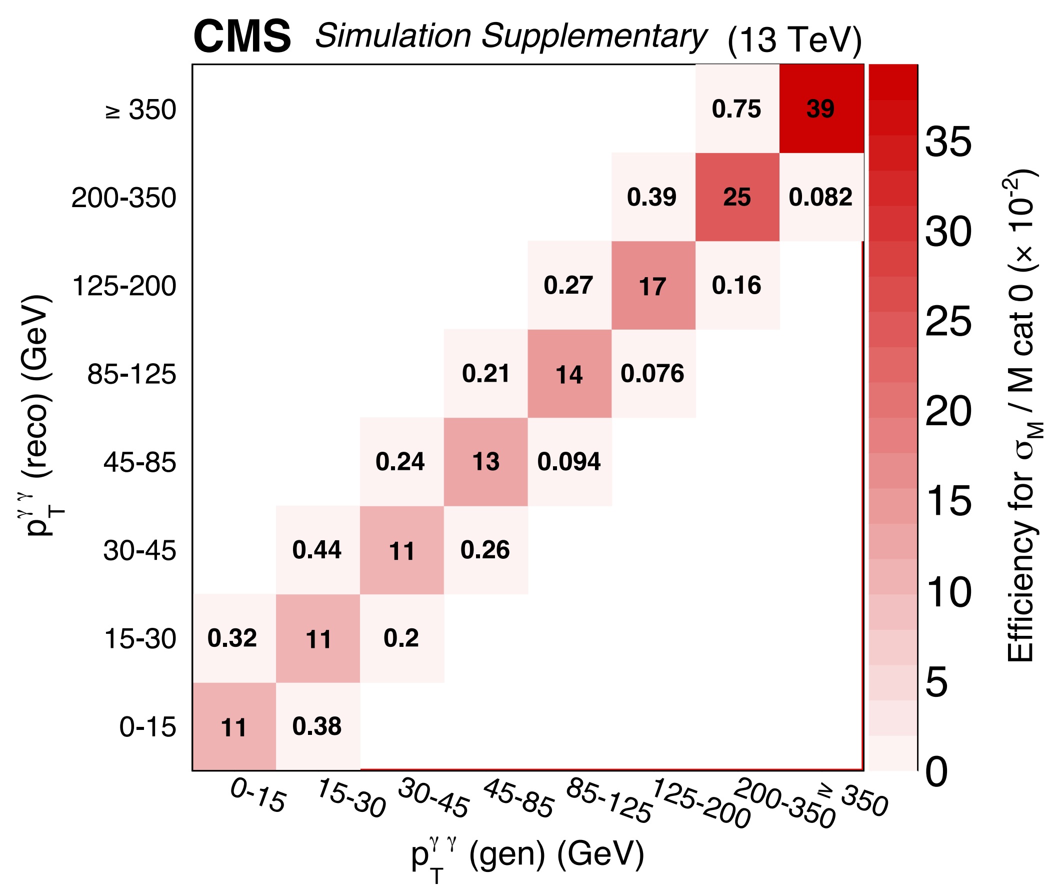
SIGNAL EXTRACTION

- **Two high pT, isolated photons** in the final state
- **3 event categories** defined using a relative mass resolution estimator, fully decorrelated from the mass itself to prevent shaping of $m_{\gamma\gamma}$ distribution
- **Full response matrix:** unfolding simultaneous to signal extraction



Fiducial Volume

$$
\begin{array}{l|l}
p_{T,1}/m_{\gamma\gamma} & > 1/3 \\
p_{T,2}/m_{\gamma\gamma} & > 1/4 \\
Iso_{gen1,2} & < 10\ \mathrm{GeV} \\
|\eta_{1,2}| & < 2.5
\end{array}
$$

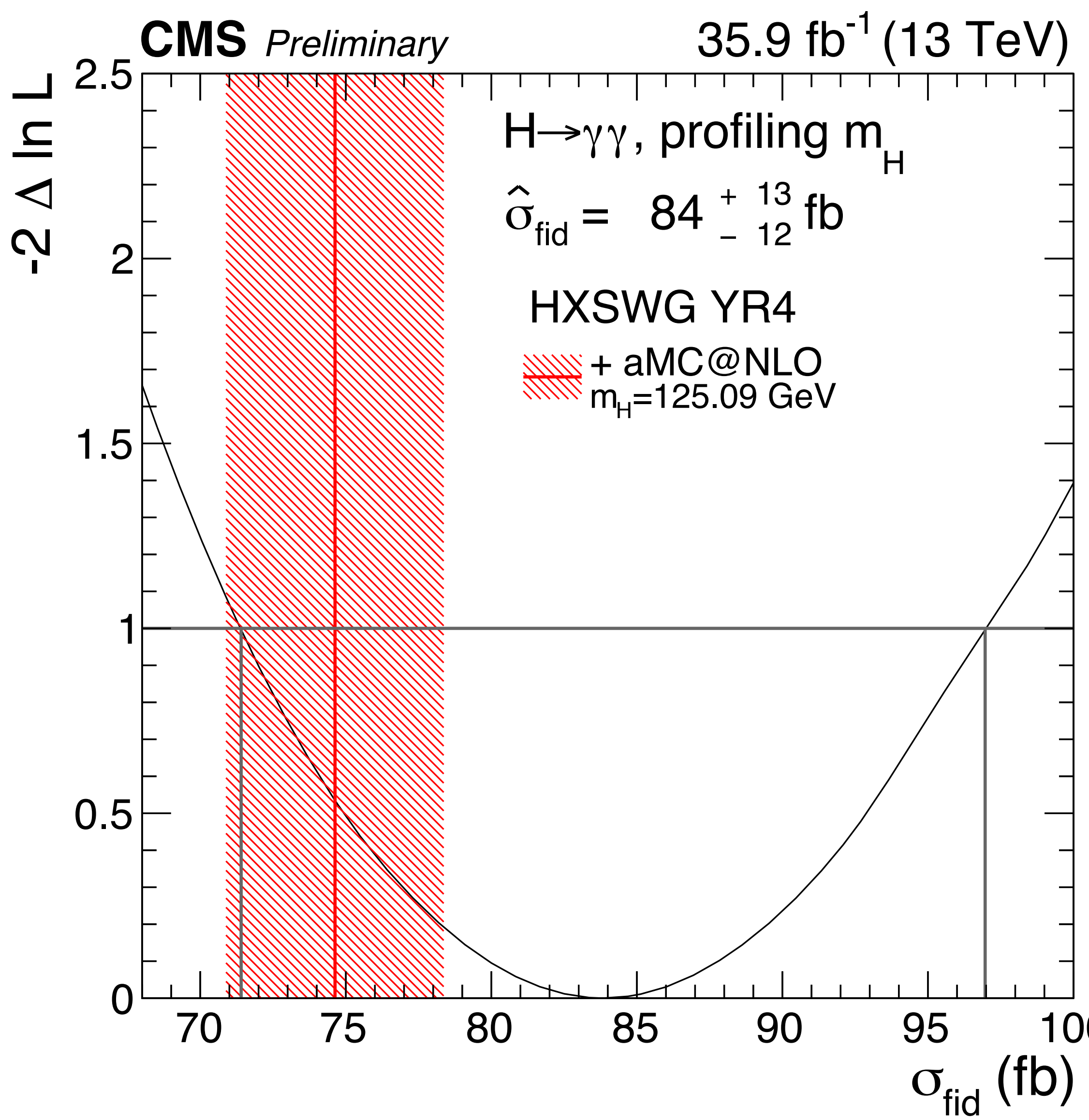


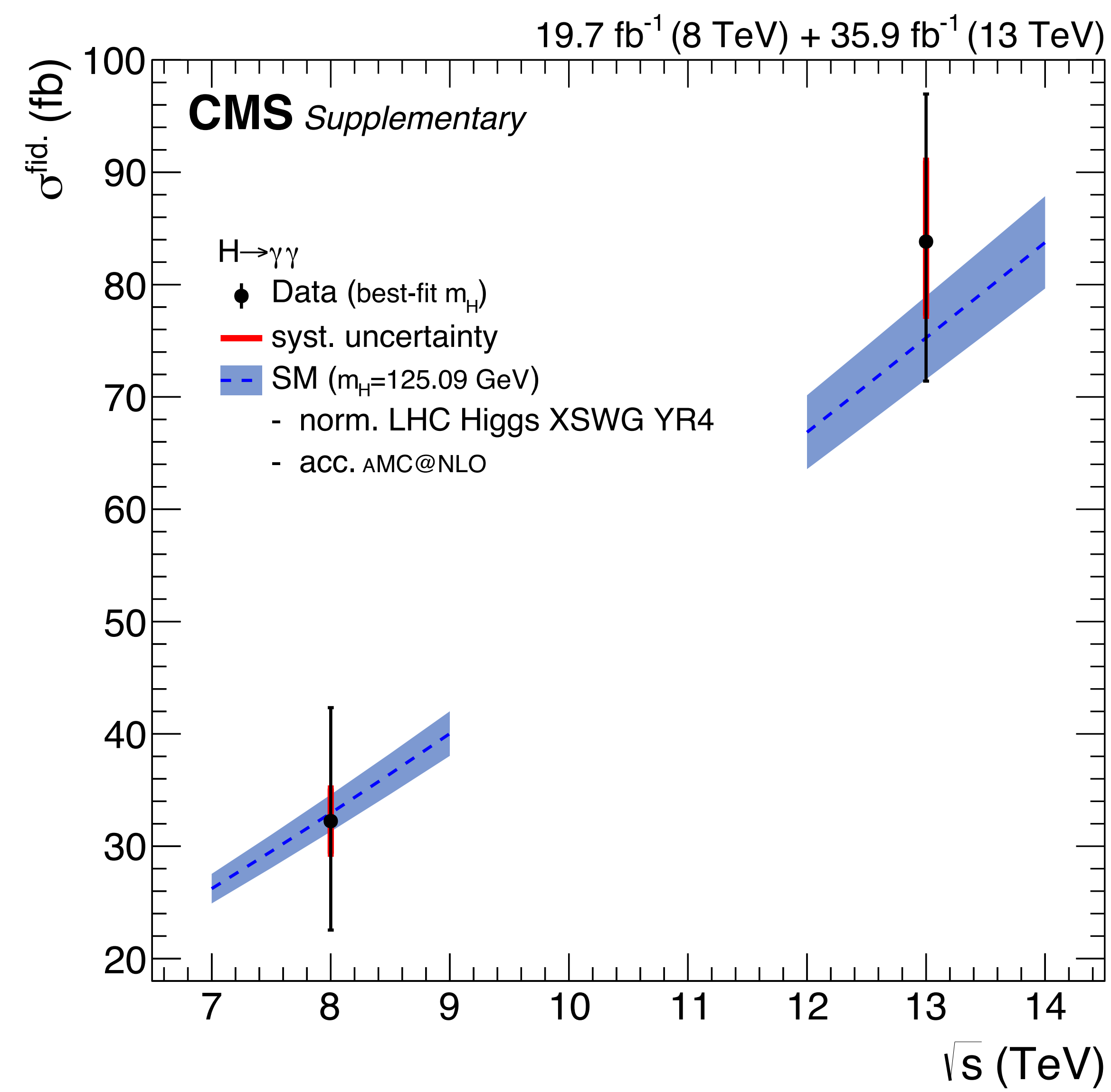
INCLUSIVE FIDUCIAL CROSS SECTIONS

Inclusive fiducial cross section measured to be

σ_{fid} = 84 ± 11 (stat.) ± 7 (syst.) fb = 84_{-12}^{+13} (stat + syst) fb

In agreement with theory prediction : σ_{fid} = 75 ± 4 fb (SM)

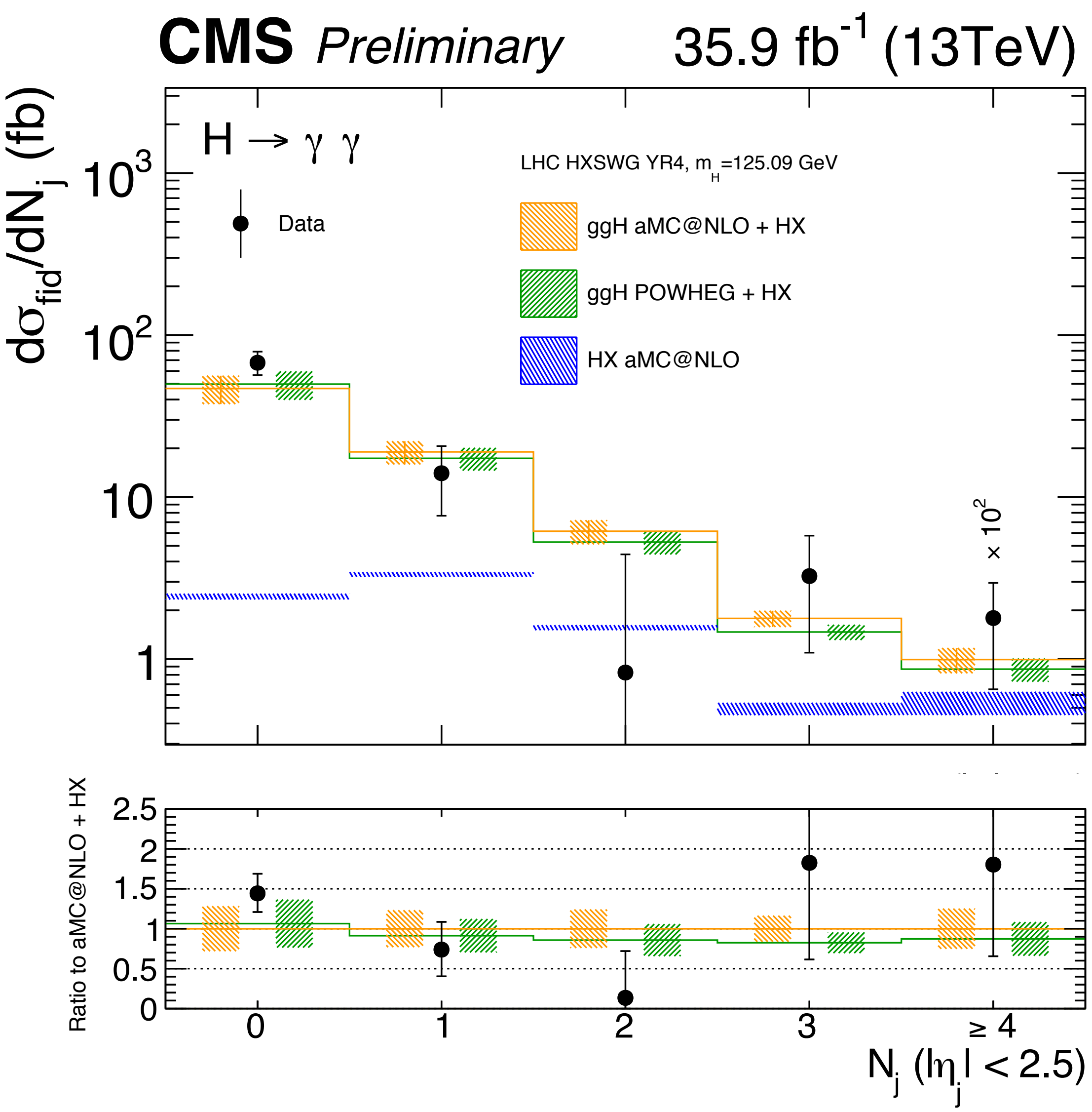


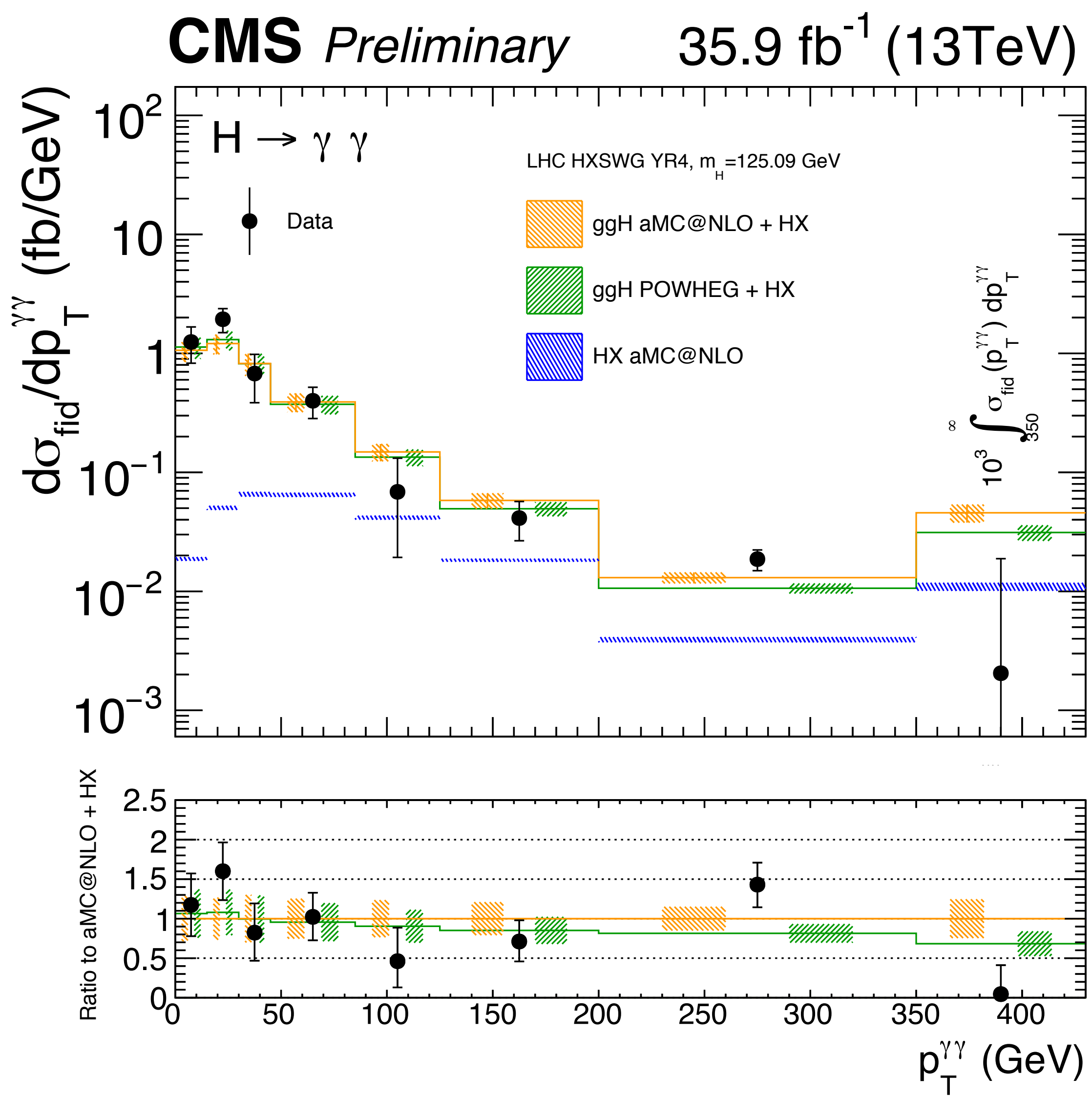


** precision still statistically dominated !*

INCLUSIVE FIDUCIAL CROSS SECTIONS

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



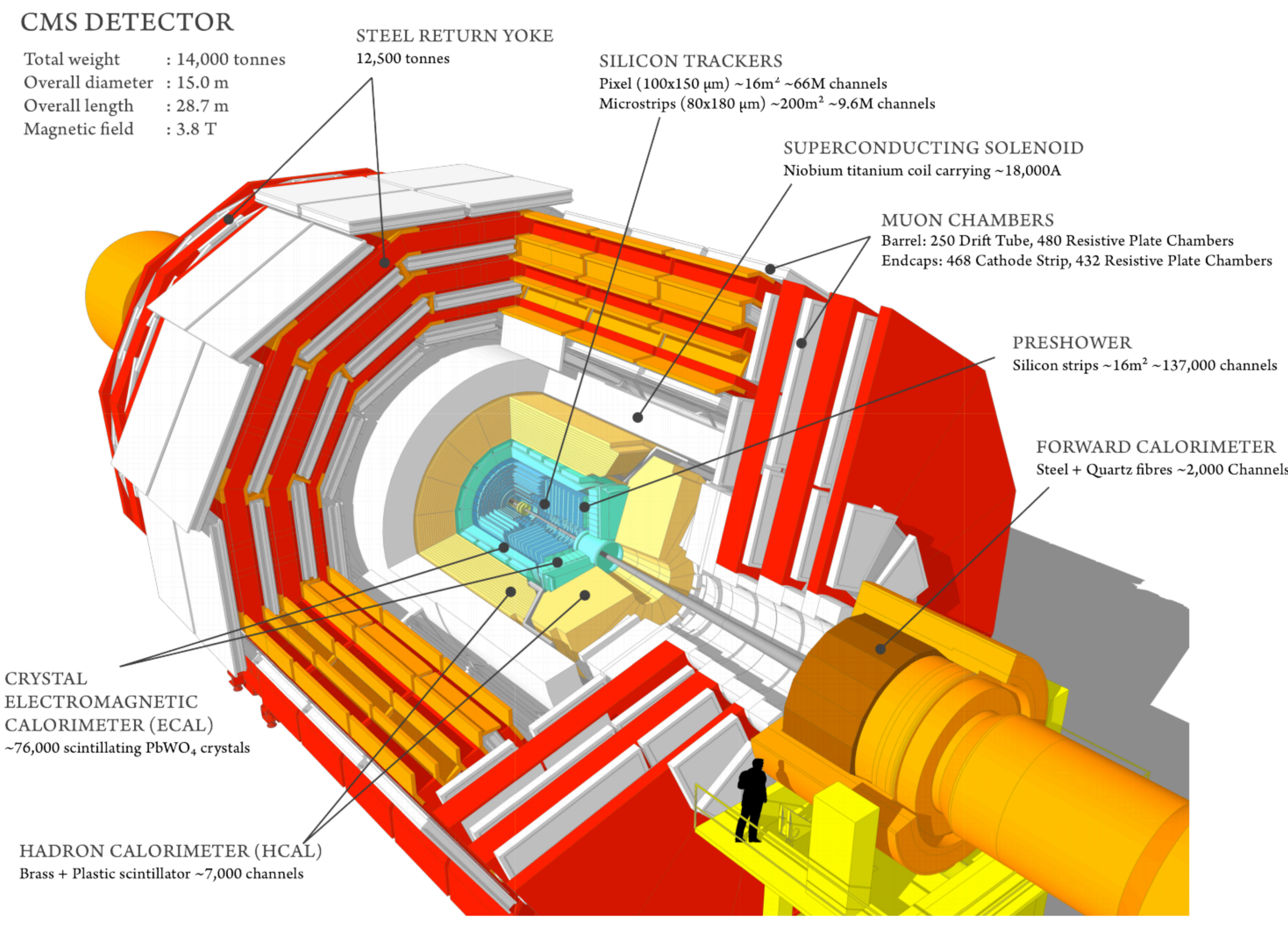


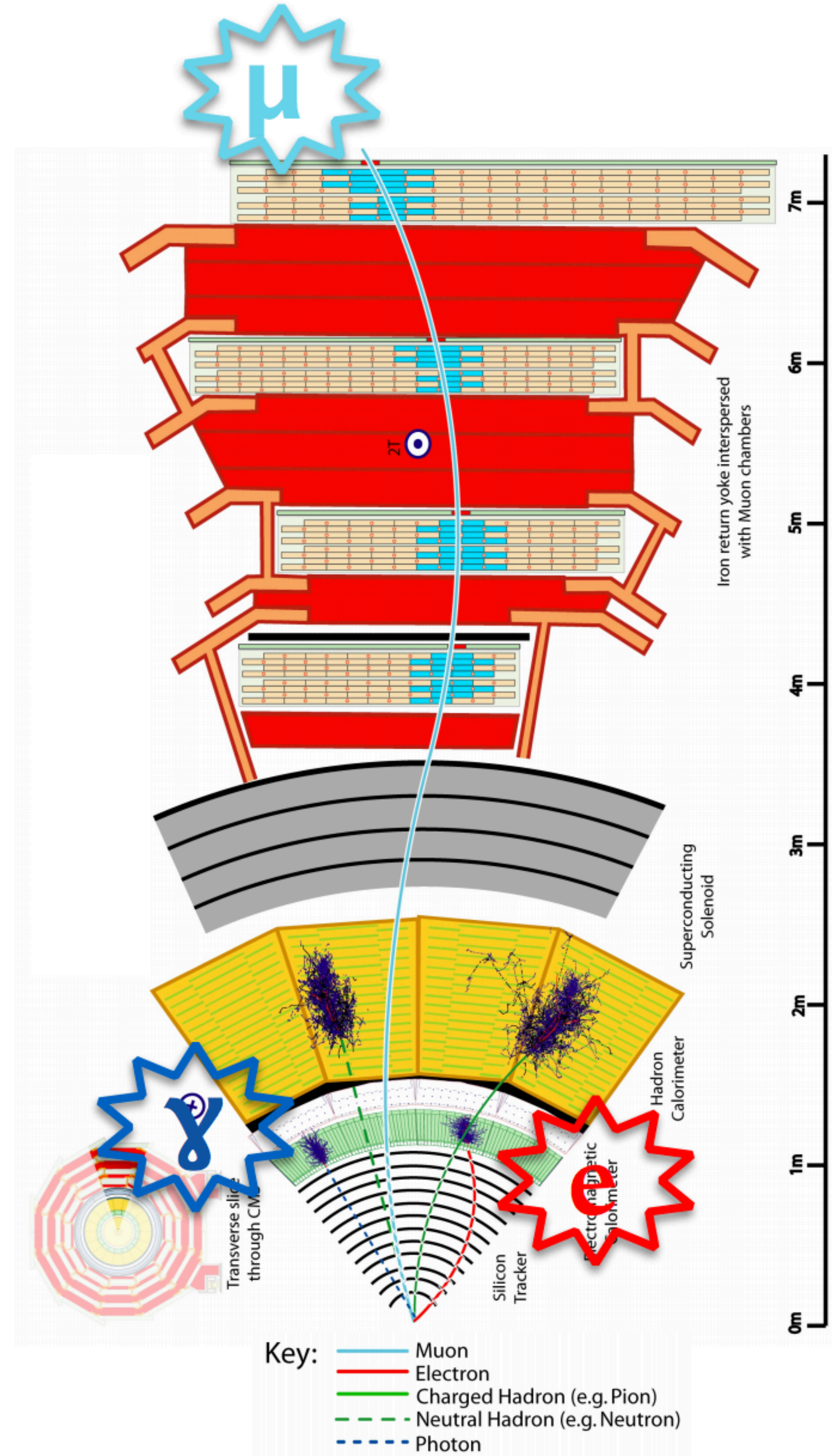
CONCLUSION

- **LHC Run 2 datasets allow study of properties of the Higgs boson**
  - Latest results of $H \to \gamma\gamma$ with 35.9 fb^{-1} collected by CMS at 13TeV
- **Differential fiducial cross sections provide more model-independent, durable measurement for a small sacrifice in sensitivity**
- **Measurements performed using 2016 data show good compatibility with SM prediction, both inclusively and differentially**
  - Results are still statistically limited
  - Expected > 100 fb^{-1} to be delivered by the end of Run2 → Improve precision of the measurements

BACKUP

CMS DECETOR


CMS DETECTOR
Total weight : 14,000 tonnes
Overall diameter : 15.0 m
Overall length : 28.7 m
Magnetic field : 3.8 T
STEEL RETURN YOKE
12,500 tonnes
SILICON TRACKERS
Pixel (100x150 μm) ~16m² ~66M channels
Microstrips (80x180 μm) ~200m² ~9.6M channels
SUPERCONDUCTING SOLENOID
Niobium titanium coil carrying ~18,000A
MUON CHAMBERS
Barrel: 250 Drift Tube, 480 Resistive Plate Chambers
Endcaps: 468 Cathode Strip, 432 Resistive Plate Chambers
PRESHOWER
Silicon strips ~16m² ~137,000 channels
FORWARD CALORIMETER
Steel + Quartz fibres ~2,000 Channels
CRYSTAL ELECTROMAGNETIC CALORIMETER (ECAL)
~76,000 scintillating $PbWO_4$ crystals
HADRON CALORIMETER (HCAL)
Brass + Plastic scintillator ~7,000 channels

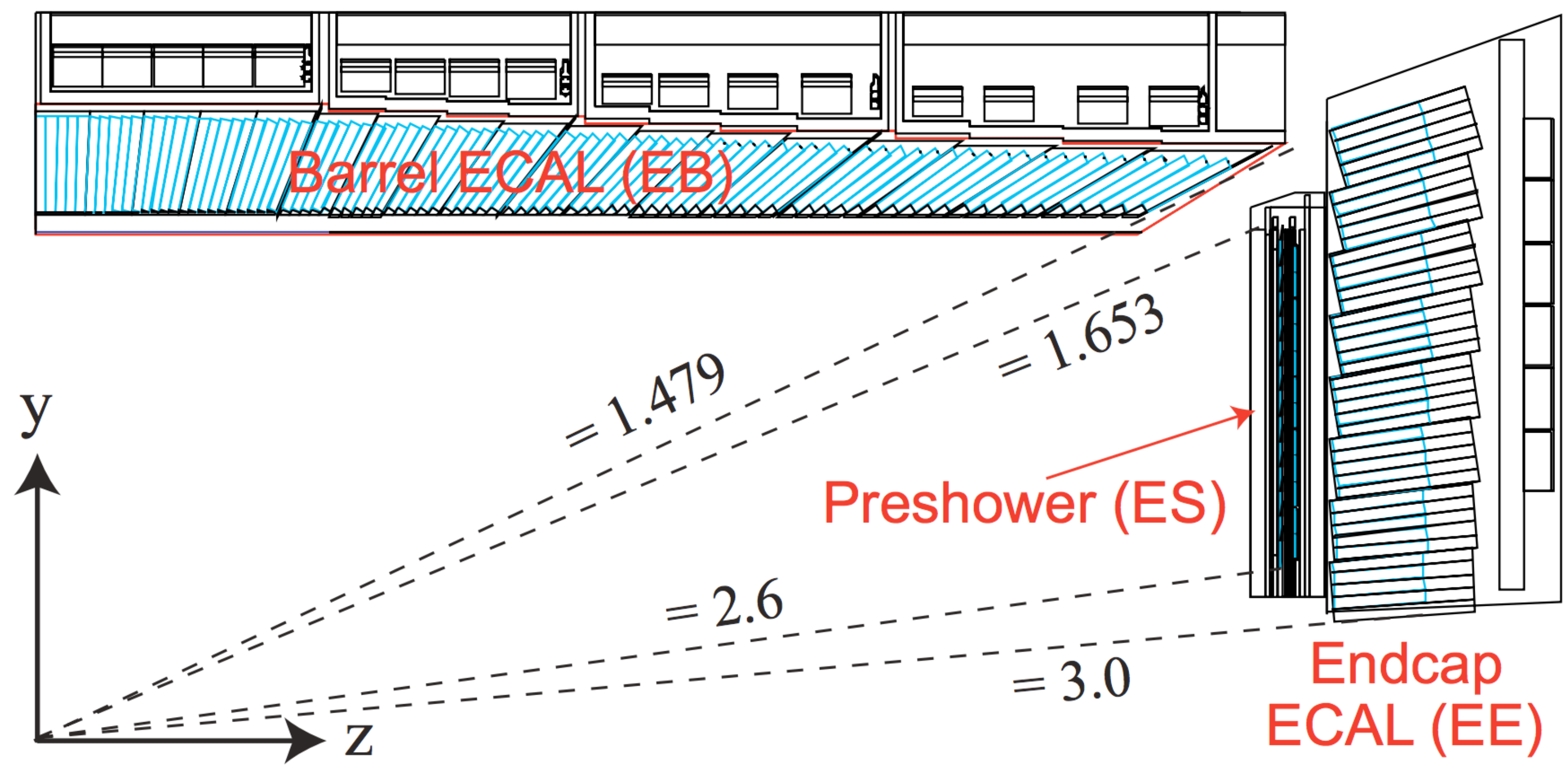

μ
γ
e
7m
6m
5m
4m
3m
2m
1m
0m
2T
Iron return yoke interspersed with Muon chambers
Superconducting Solenoid
Hadron Calorimeter
Silicon Tracker
Key:
Muon
Electron
Charged Hadron (e.g. Pion)
Neutral Hadron (e.g. Neutron)
Photon

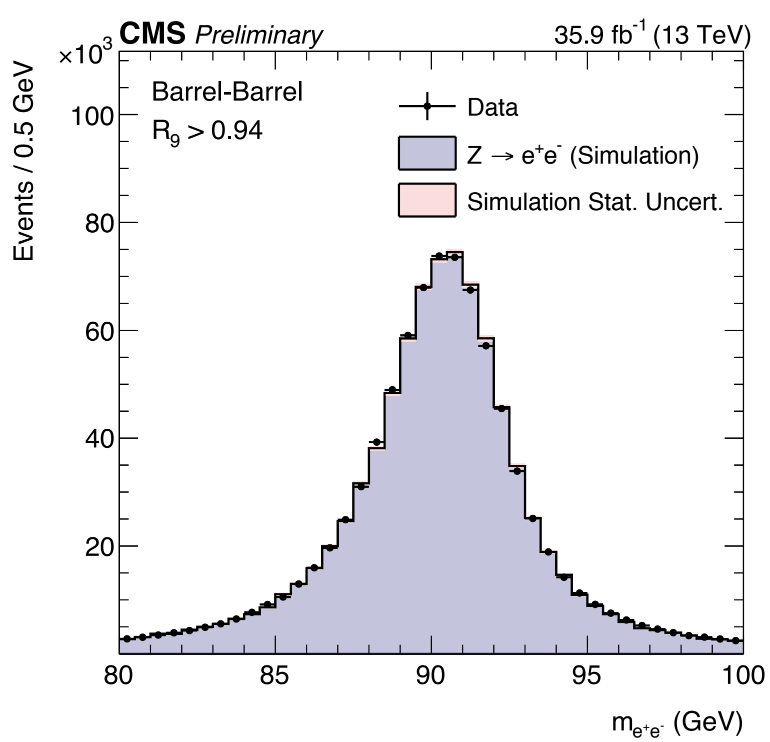
CMS ELECTROMAGNETIC CALORIMETER

PHOTON RECONSTRUCTION

CERN-LHCC-2006-001

- High granularity **PbWO4** crystal layout
- Particle-flow reconstruction :
  - Reconstructs all the particle by combining all the information from all the relevant sub-detectors

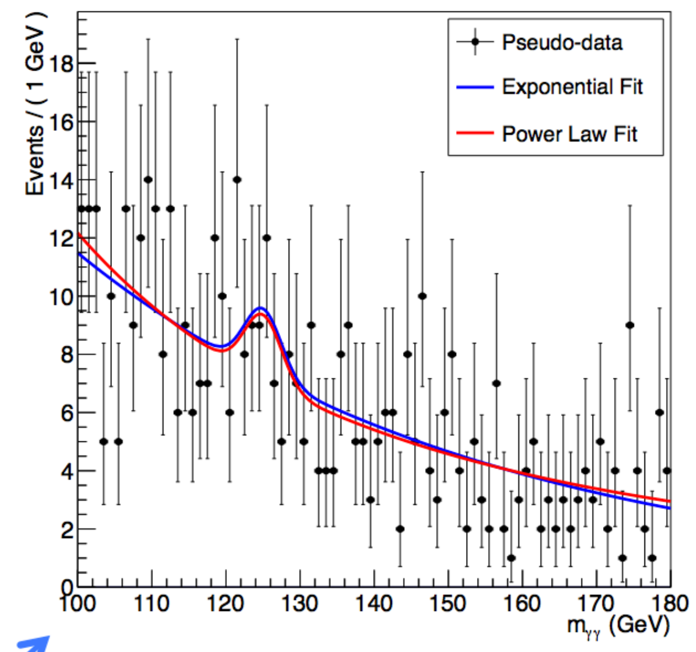


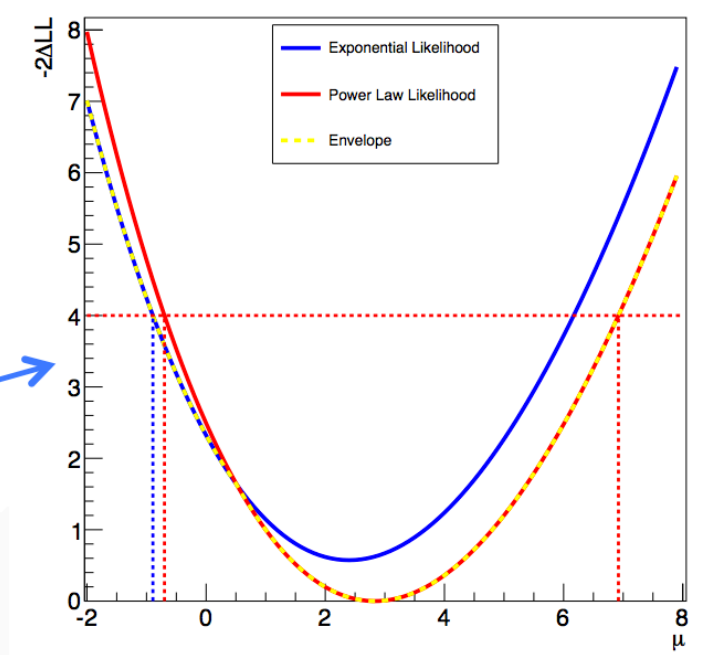


- Photons are identified as ECAL deposits without associated track
- Diphoton vertex determined using recoiling tracks
- MC-based MVA regression for calibration. Using Z→ee events where electron is reconstructed as photon for correcting energy scale in data

BACKGROUND MODELLING

- Employ a new strategy for modelling the background
- Background shape is *a priori* unknown
- The concept is to "profile" over several different function choices (i.e. all are considered in the fit)
- Allow the data to select the one which fits the best
- Subsequent "*envelope*" around NLL curve of different choices means uncertainty will take into account model assumption
- Correct likelihood for functions with different numbers of parameters



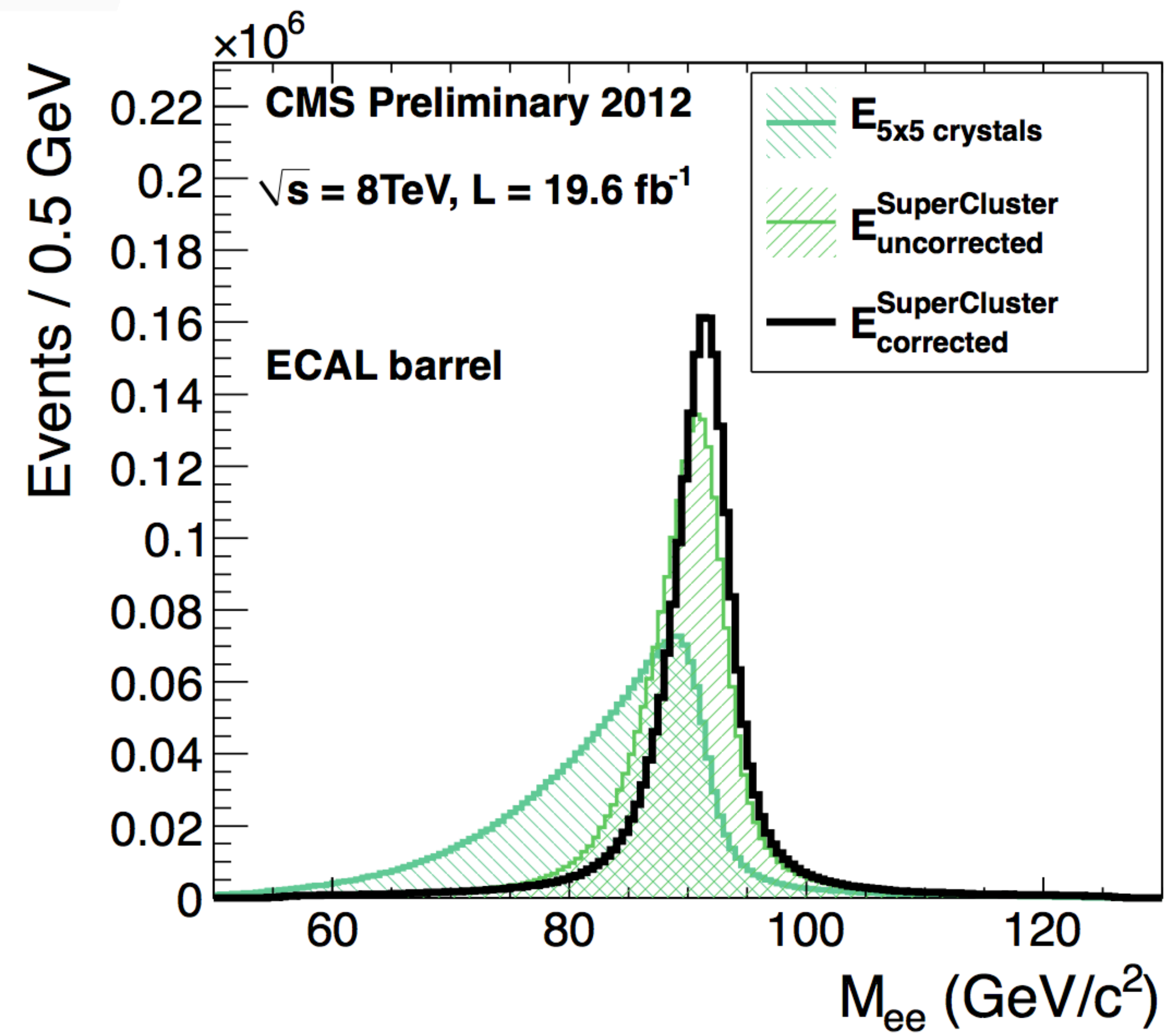


Toy example

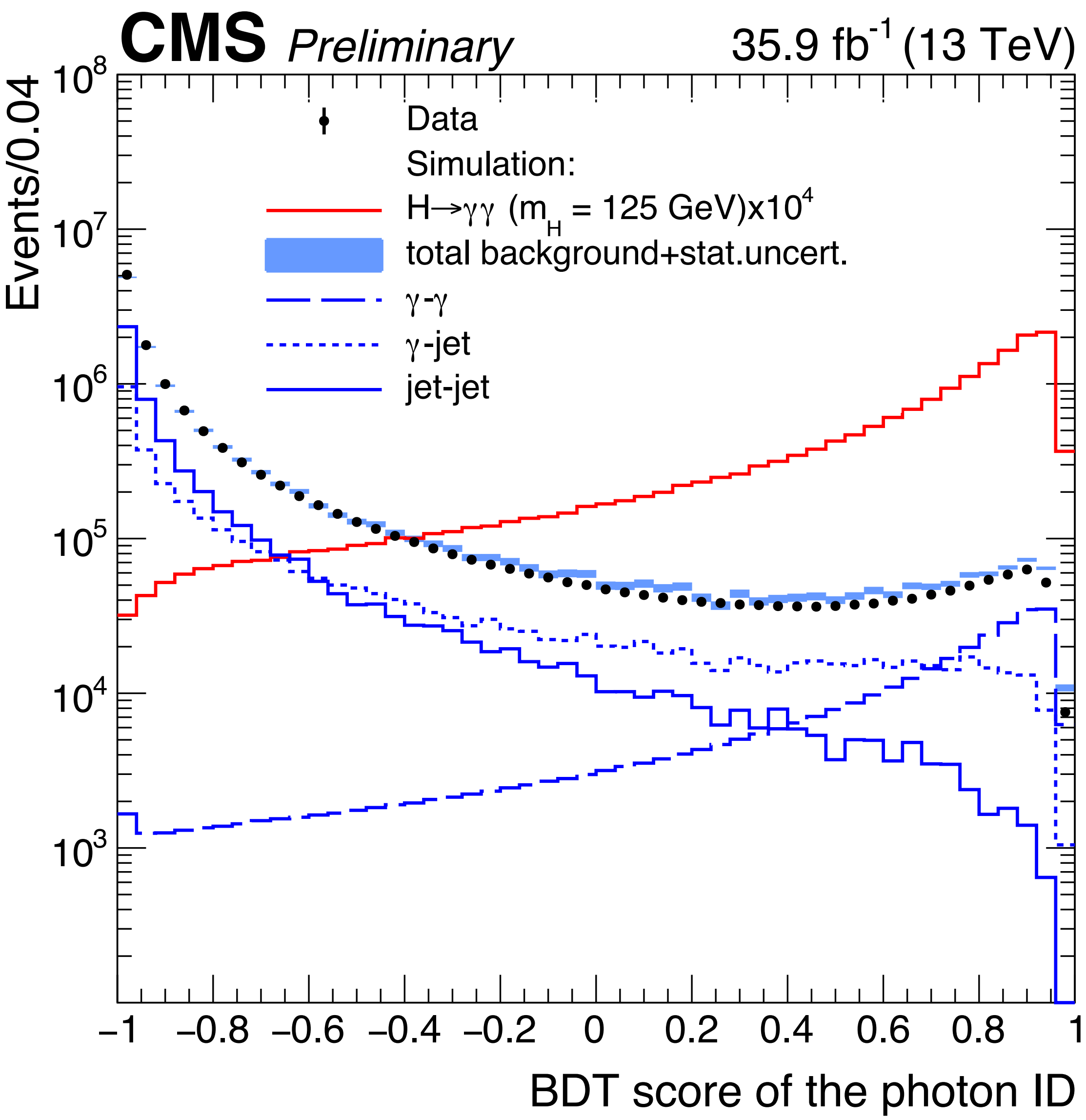
- 1 category, 2 function choices, e^{-px} and x^{-p}
- Profile "envelope" gives best fit with x^{-p}
- 2 sigma error is enlarged by the envelope
- In principle envelope method will increase uncertainty because of different function choices

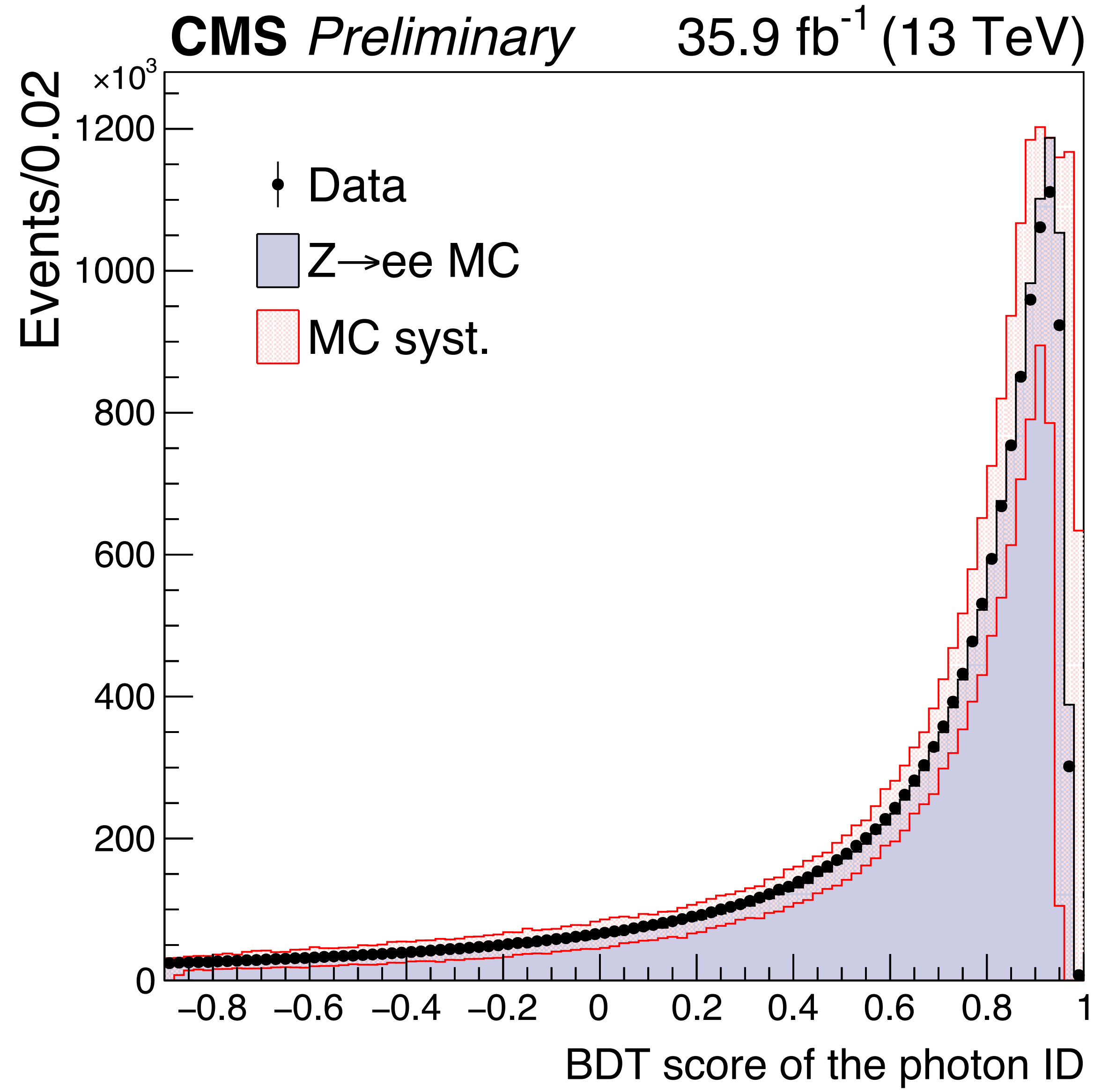
PHOTON ENERGY

- Electromagnetic calorimeter response
  - Corrected through time
  - Uniformity Inter-calibration in eta and phi
  - Transparency correction
- Energy and its uncertainty corrected fro local and global shower containment:
  - Energy regression targeting the true energy (E_{true}/E_{reco})
- Scale versus time and resolution calibration through time:
  - Use Zee peak as a reference



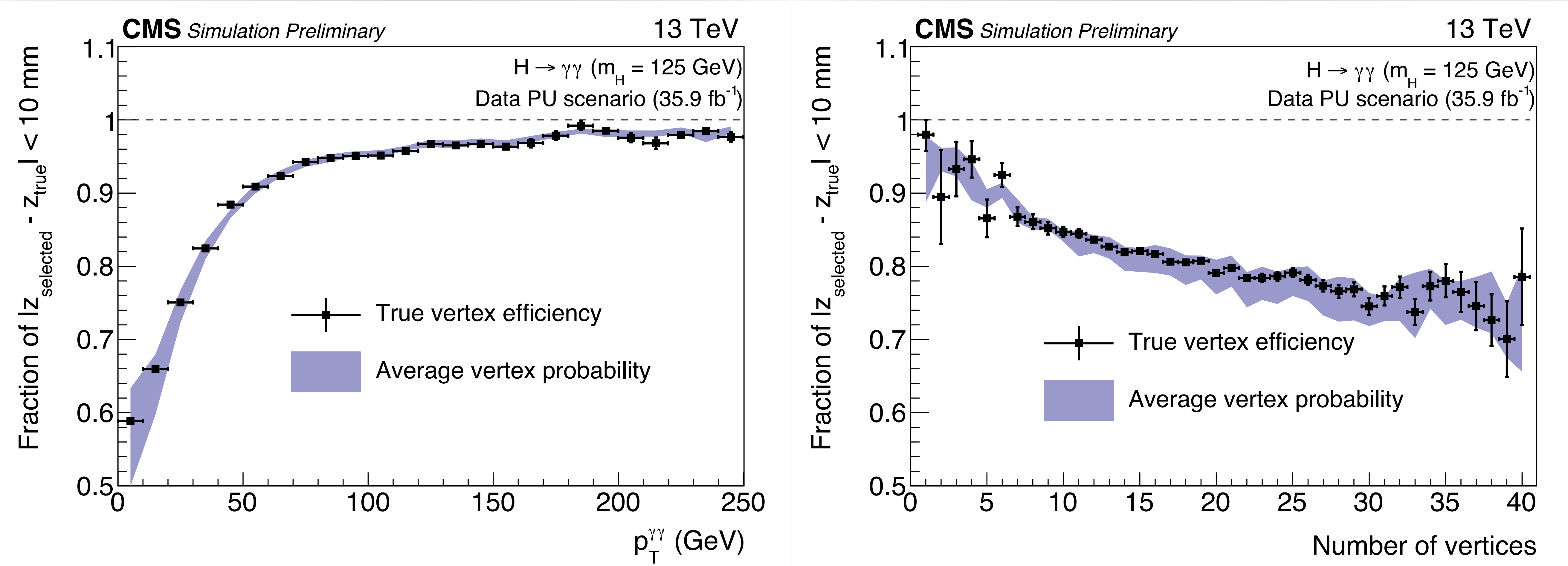
PHOTON IDENTIFICATION





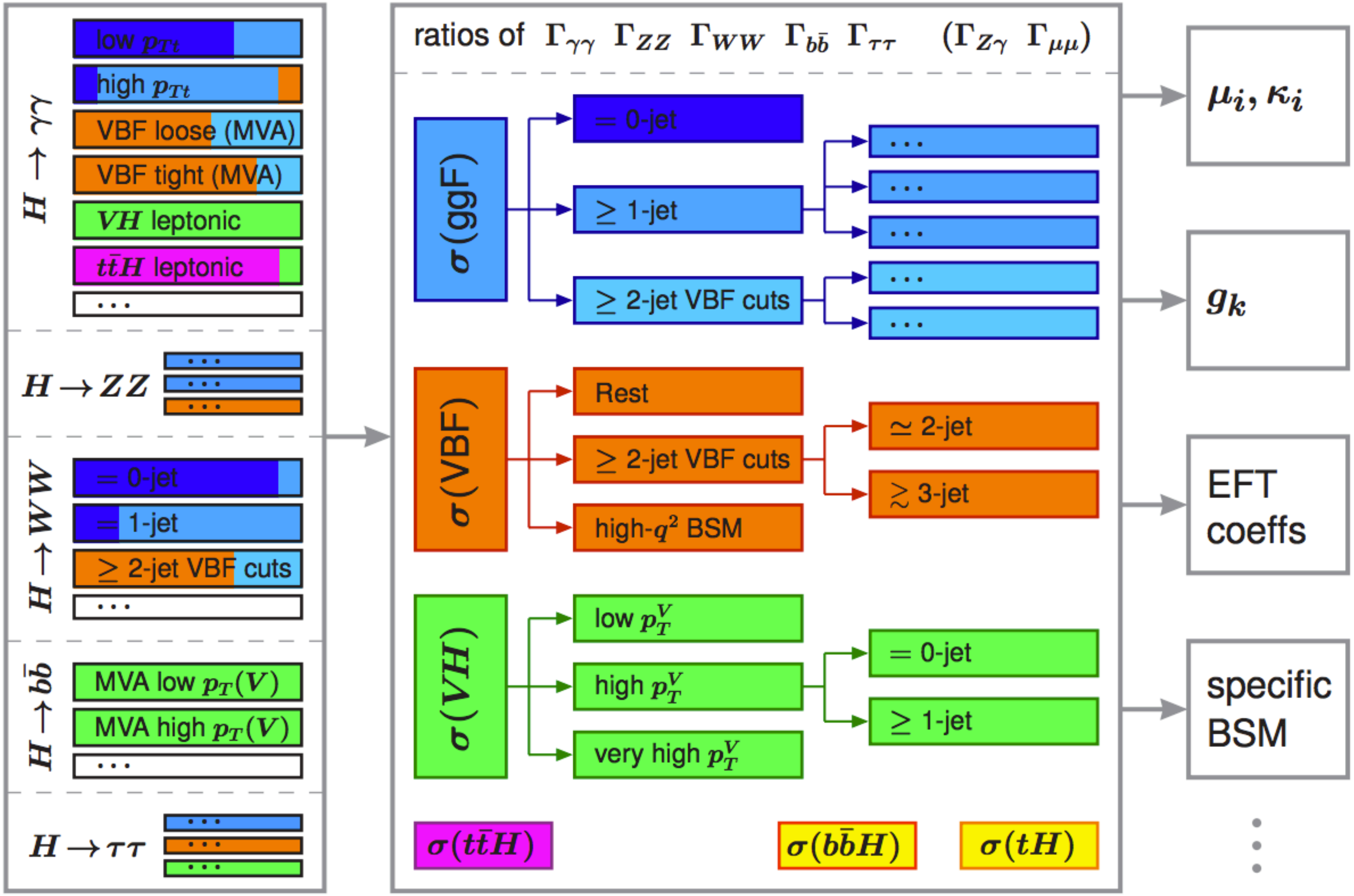
- Photon ID BDT discriminate real and fake photons (from jet fragments)

PRIMARY VERTEX IDENTIFICATION



- BDT using recoiling tracks to assign vertex to diphotons
- Negligible impact on resolution

SIMPLIFIED TEMPLATE CROSS-SECTIONS (STXS)



EXPECTED YIELDS FOR THE FULL 2016 DATASET

Event Categories	SM 125 GeV Higgs boson expected signal								
	Total	ggH	VBF	ttH	bbH	tHq	tHW	WH lep	ZH
Untagged 0	45.83	80.19 %	11.75 %	1.83 %	0.40 %	0.47 %	0.22 %	0.41 %	0.1
Untagged 1	480.56	86.81 %	7.73 %	0.56 %	1.15 %	0.13 %	0.02 %	0.47 %	0.2
Untagged 2	670.45	89.76 %	5.48 %	0.44 %	1.18 %	0.08 %	0.01 %	0.51 %	0.3
Untagged 3	610.07	91.13 %	4.51 %	0.48 %	1.07 %	0.07 %	0.01 %	0.55 %	0.3
VBF 0	10.01	21.69 %	77.09 %	0.34 %	0.35 %	0.29 %	0.03 %	0.03 %	0.0
VBF 1	8.64	33.58 %	64.64 %	0.39 %	0.52 %	0.36 %	0.04 %	0.13 %	0.0
VBF 2	27.76	50.14 %	46.46 %	0.81 %	0.73 %	0.53 %	0.07 %	0.20 %	0.0
ttH Hadronic	5.85	10.99 %	0.70 %	77.54 %	2.02 %	4.13 %	2.02 %	0.09 %	0.0
ttH Leptonic	3.81	1.90 %	0.05 %	87.48 %	0.08 %	4.73 %	3.04 %	1.53 %	1.1
ZH Leptonic	0.49	0.00 %	0.00 %	2.56 %	0.00 %	0.02 %	0.13 %	0.00 %	97.3
WH Leptonic	3.61	1.26 %	0.59 %	5.18 %	0.18 %	3.03 %	0.73 %	84.48 %	4.3
VH LeptonicLoose	2.75	9.16 %	2.70 %	2.34 %	0.57 %	1.81 %	0.13 %	63.62 %	18.8
VH Hadronic	9.69	57.38 %	3.68 %	3.61 %	0.35 %	1.39 %	0.27 %	0.17 %	0.4
VH Met	4.25	23.63 %	2.46 %	14.45 %	0.41 %	2.00 %	1.14 %	25.17 %	28.6
Total	1883.77	86.96 %	7.09 %	1.00 %	1.09 %	0.15 %	0.04 %	0.81 %	0.4

GGH THEORY UNCERTAINTIES

from HXSWG workshop : https://indico.cern.ch/event/595100/contributions/2649194/attachments/1493264/2322052/ggF_Jul17.pdf

Interim 2017 uncertainty scheme

Documentation in preparation: LHCHXSWG-2017-001

9 independent sources of uncertainties

1. QCD scale variation
2. Resummation scale variation
3. 0 □ 1 jet migration
4. 1 □ 2 jets migration
5. VBF phase space
6. VBF phase space with 3rd jet veto
7. Higgs p_T 0-60/60-∞ GeV
8. Higgs p_T 60-120/120-∞ GeV
9. Finite top mass dependence: $p_T > m_t$

Uncertainty on finite top mass effects.
Taken from difference between LO and NLO rescaling.

Jet bins: four sources
Taken from YR4, Tackmann et al.
Cross checked with JVE @N3LO

VBF topology uncertainties
Found consistent central values and uncertainties using Run-1 style, YR3 uncertainties (from MCFM + ST), and using GoSam+Sherpa HJJJ @ NLO

Uncertainties on the Higgs pT shape within a fixed jet multiplicity bin. Taken from QCD scale variations of Powheg NNLOPS

GGH THEORY UNCERTAINTIES

from HXSWG workshop : https://indico.cern.ch/event/595100/contributions/2649194/attachments/1493264/2322052/ggF_Jul17.pdf

Different uncertainties

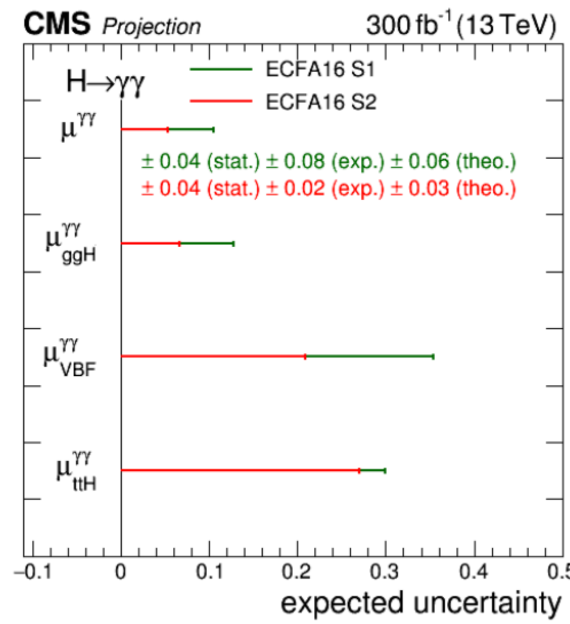
Different phase spaces

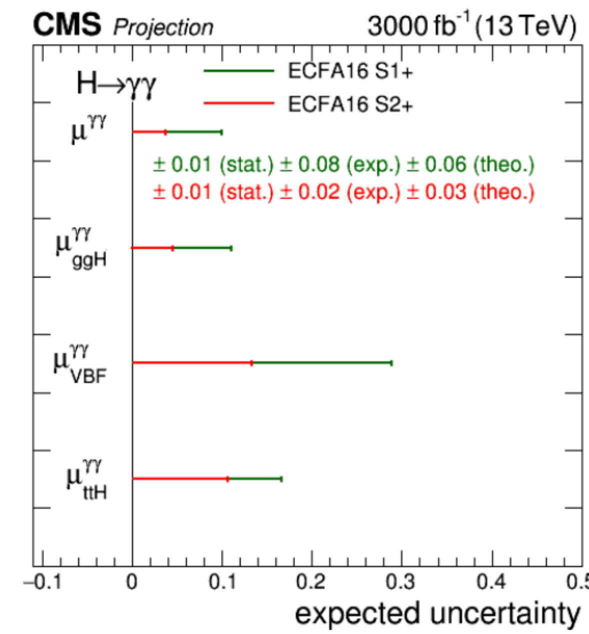
Cross sections and fractional uncertainties

STXS	sig	stat	mu	res	mig01	mig12	VBF2j	VBF3j	pT60	pT120	qm_top	Tot
Incl	48.52	+/- 0.00	+4.6%	+2.1%	-0.0%	-0.0%	+0.3%	-0.0%	+0.0%	+0.2%	+0.2%	+5.1%
FWDH	4.27	+/- 0.01	+4.5%	+1.9%	-0.5%	-0.2%	+0.0%	+0.0%	-0.3%	-0.1%	+0.0%	+4.9%
VBF_J3V	0.27	+/- 0.00	+0.0%	+0.0%	+0.0%	+0.0%	+20.0%	-32.0%	-1.6%	+1.1%	+0.1%	+37.8%
VBF_J3	0.36	+/- 0.00	+0.0%	+0.0%	+0.0%	+0.0%	+20.0%	+23.5%	-0.2%	+2.5%	+0.2%	+31.0%
=0J	27.25	+/- 0.03	+3.8%	+0.1%	-4.1%	+0.0%	+0.0%	+0.0%	+0.0%	+0.0%	+0.0%	+5.6%
=1J_0-60	6.49	+/- 0.01	+5.2%	+4.5%	+7.9%	-6.8%	+0.0%	+0.0%	-4.8%	-1.6%	+0.0%	+13.5%
=1J_60-120	4.50	+/- 0.01	+5.2%	+4.5%	+7.9%	-6.8%	+0.0%	+0.0%	+4.8%	-0.9%	+0.0%	+13.4%
=1J_120-200	0.74	+/- 0.00	+5.2%	+4.5%	+7.9%	-6.8%	+0.0%	+0.0%	+10.0%	+10.1%	+0.5%	+18.9%
=1J_200->	0.15	+/- 0.00	+5.2%	+4.5%	+7.9%	-6.8%	+0.0%	+0.0%	+10.0%	+14.0%	+10.5%	+23.7%
>=2J_0-60	1.22	+/- 0.01	+8.9%	+8.9%	+4.4%	+18.2%	+0.0%	+0.0%	-5.9%	-1.6%	+0.0%	+23.3%
>=2J_60-120	1.86	+/- 0.01	+8.9%	+8.9%	+4.4%	+18.2%	+0.0%	+0.0%	-0.2%	-0.2%	+0.0%	+22.5%
>=2J_120-200	0.99	+/- 0.00	+8.9%	+8.9%	+4.4%	+18.2%	+0.0%	+0.0%	+6.6%	+10.6%	+0.6%	+25.8%
>=2J_200->	0.42	+/- 0.00	+8.9%	+8.9%	+4.4%	+18.2%	+0.0%	+0.0%	+10.0%	+14.0%	+11.8%	+30.7%
=0J	30.12	+/- 0.03	+3.8%	+0.1%	-4.1%	+0.0%	+0.0%	+0.0%	+0.0%	+0.0%	+0.0%	+5.6%
=1J	12.92	+/- 0.02	+5.2%	+4.5%	+7.9%	-6.8%	+0.0%	+0.0%	-0.1%	-0.4%	+0.2%	+12.5%
>=2J	5.47	+/- 0.01	+7.8%	+7.8%	+3.9%	+16.1%	+2.3%	-0.0%	+0.4%	+2.9%	+1.1%	+20.3%
>=1J 60-200	9.09	+/- 0.01	+6.2%	+5.8%	+6.4%	+1.9%	+0.9%	+0.1%	+4.2%	+1.7%	+0.1%	+11.8%
>=1J 120-200	1.96	+/- 0.01	+6.8%	+6.5%	+5.5%	+6.9%	+1.5%	+0.4%	+8.0%	+10.4%	+0.6%	+18.5%
>=1J >200	0.58	+/- 0.00	+7.9%	+7.7%	+5.4%	+11.6%	+0.0%	+0.0%	+10.0%	+14.0%	+11.4%	+26.7%
>=1J >60	9.68	+/- 0.01	+6.3%	+5.9%	+6.3%	+2.5%	+0.8%	+0.1%	+4.6%	+2.5%	+0.8%	+12.2%
>=1J >120	2.54	+/- 0.01	+7.0%	+6.8%	+5.5%	+8.0%	+1.2%	+0.3%	+8.4%	+11.2%	+3.0%	+19.9%
>=1	18.40	+/- 0.02	+6.0%	+5.5%	+6.7%	-0.0%	+0.7%	-0.0%	+0.0%	+0.5%	+0.4%	+10.6%

Higgs beyond run 2

Higgs boson signal strength





- Projections have been obtained by scaling event yields to 300 (3000) fb^{-1} at $\sqrt{s}$ = 14 TeV
  - Use Run 1 Legacy results (7+8 TeV) and assume performance unchanged
- Two scenarios for systematic uncertainties were considered:
  - ➔ Scenario 1: All systematic uncertainties are kept constant with integrated luminosity. The performance of the CMS detector is assumed to be unchanged with respect to the reference analysis.
  - ➔ Scenario 2: All systematic uncertainties are kept constant with integrated luminosity. The effects of higher
  - pileup conditions and detector upgrades on the future performance of CMS are taken into account.

Challenges and tools

- Pile up has reached another frontier in run 2
  - In time pile up (PU) can get as high as 40 in average
  - in 2012 we have collected less than 1% of our luminosity with PU>35
- new techniques are being developed at high level reco using mostly information from Particle Flow objects
  - Charged hadron subtraction
    - Removal of charged particle originated from other than the leading vertex
  - more sophisticated approach (PUPPI) add to the weight many other observables
    - Associate a weight to all particle candidates corresponding to the probability to come from a PU. for PUPPI the weight comes from the distance of the neutral particle w.r.t. to candidates associated to the primary vertex
  - Other tools from run 1 have been refurbished and tuned for the Run 2, pileup jet identification (important specially in the forward region)
- Many other tools has been made and developed for run 2 and currently used in different analyses (QG tagging ,… etc)