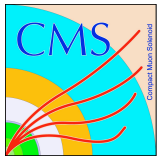
# Search for the Higgs boson decaying to two photons in CMS

Martina Malberti (UCR) on behalf of the CMS Collaboration

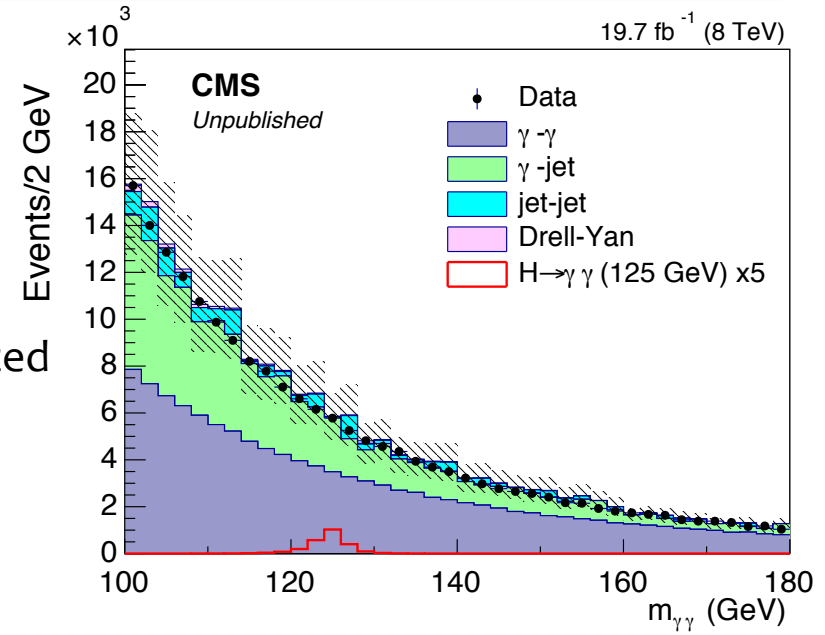



20th Particles and Nuclei International Conference
25-29 August 2014, Hamburg, Germany

UCR

# Introduction and overview

- Small peaking signal on large falling background
- Small BR ~ 0.23%
- Clean signature: two high $p_T$ isolated γ's
  - high precision for mass reconstruction
- Results based on the **full Run I dataset** are presented
  - [arXiv:1407.0558](arXiv:1407.0558)
- Main improvements
  - final set of CMS ECAL calibrations
  - analysis chain completely re-optimized
  - exclusive mode tagging expanded to include *ALL* production modes
  - new method for modeling the background
  - effort into studying the energy scale uncertainties
    - systematic uncertainty on the mass reduced by a factor of 3

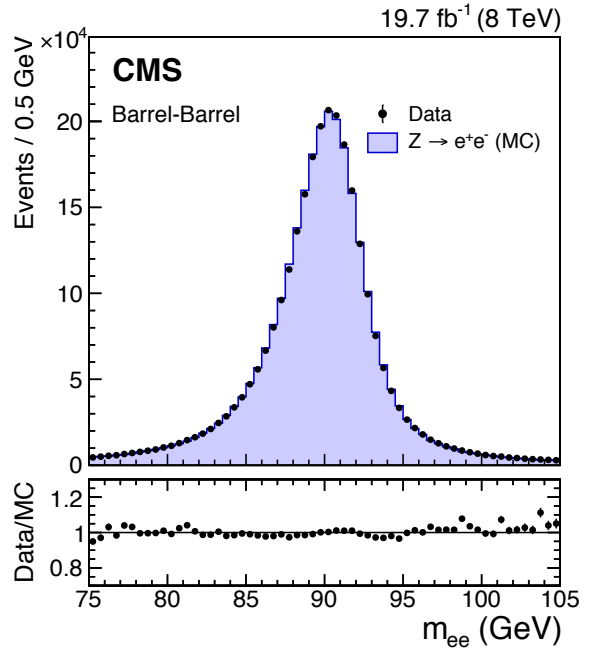


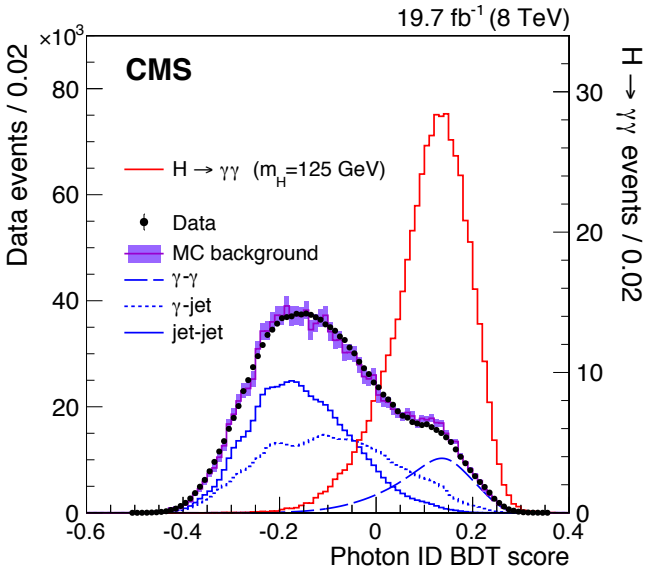

| Change | Improved energy resolution (new calibration+ new regression) | New event selection (re-training + re-categorisation) | Background modelling |
|---|---|---|---|
| Expected improvement over published PAS (Moriond 2013) | ~9% | ~9% | ~7% |

# Photon reconstruction and identification

- **Photon energy**
  - reconstructed from energy deposits in the ECAL
    - single channel response intercalibration and stability vs time ($\pi^0 \rightarrow \gamma\gamma$, $W \rightarrow e\nu$, $Z \rightarrow ee$)
  - **multivariate regression** to correct for shower containment, material and pile-up effects
    - also provides an **estimate of the per photon energy resolution**
  - **additional corrections** on photon MC energy to match resolution/scale in data estimated from data/MC comparison on **Z →ee**

- **Photon identification**
  - use **shower shape and isolation variables** in a multivariate approach (Boosted Decision Trees) to discriminate prompt photons from fakes
  - the BDT provides an **estimate of the photon quality**

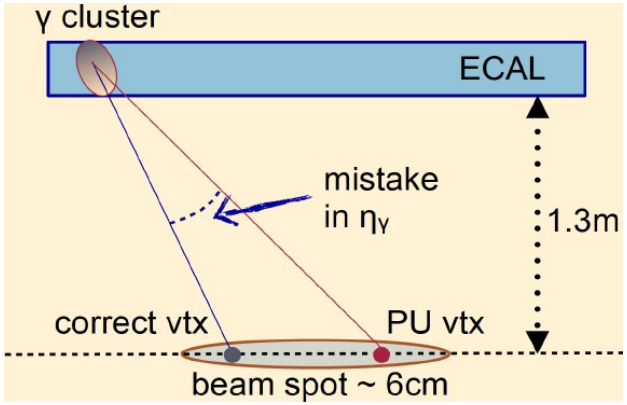


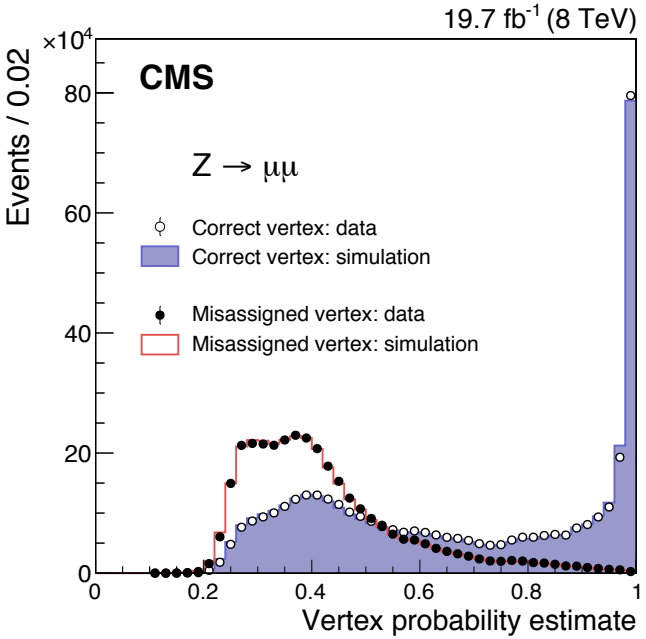



# Vertex assignment

- High **pile-up** conditions
  - 7 TeV: $\langle N_{PU} \rangle = 9$
  - 8 TeV: $\langle N_{PU} \rangle = 21$
- Di-photon invariant mass resolution affected by the vertex choice
  - negligible impact if vertex position known within 10 mm
- Vertex assignment based on a multivariate approach with inputs from:
  - **vertex tracks + di-photon kinematics**
  - **conversions**
- An additional BDT is used to estimate the **per event probability of assigning the correct vertex**
  - ~ 80% for $N_{PU}$ ~20
- Validation on Z →μμ and γ+jet samples

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



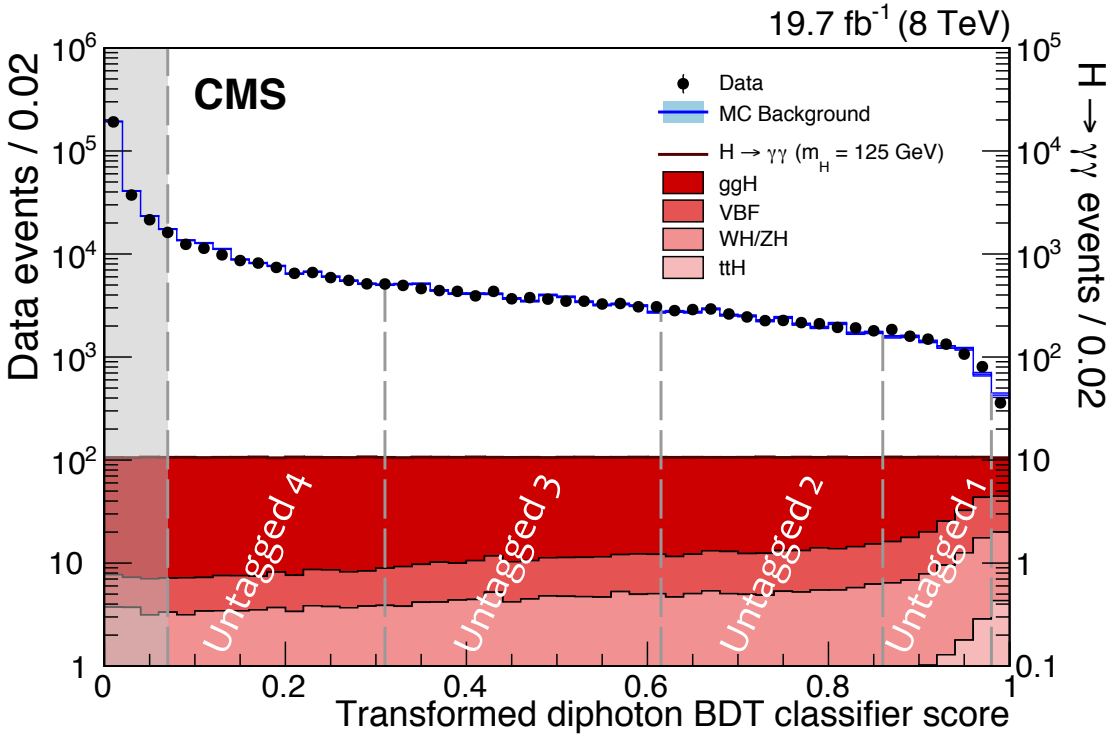
# Events classification

- Split events in categories exploiting different S/B and mass resolution to achieve maximum sensitivity:
  - "**tagged**" categories enriched in VBF, VH, ttH signal production modes improve the analysis reach to measure Higgs couplings
  - "**untagged**" categories have the largest sensitivity (mainly ggH production)
- Events are assigned exclusively to a category following S/B ordering:


- **25 mutually exclusive analysis categories** in total (untagged + tagged)
  - 11 in the 7 TeV dataset
  - 14 in the 8 TeV dataset

# Untagged categories

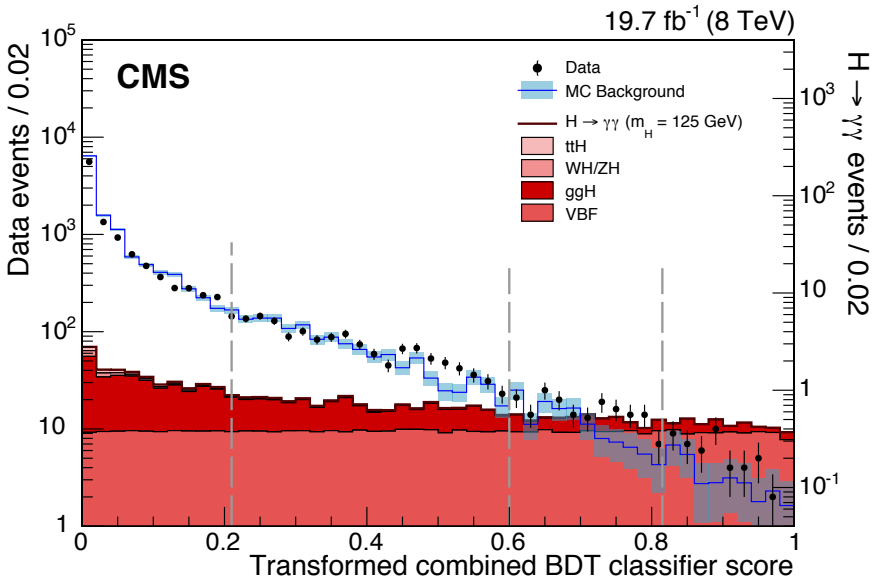
- All the information is combined in a **multivariate event classifier** *(di-photon BDT)*
  - includes kinematics, photon quality, mass resolution, correct vertex probability
  - is mass-independent
  - has high value for events w/ good di-photon mass resolution and high probability of being signal rather than background
  - output of the di-photon BDT used to define event classes (untagged)




S/B increases with the classifier score. The VBF, VH, and ttH processes achieve high scores, due to their harder pTγγ spectrum

# Tagged categories

- Categories "tagged" by the presence of **additional objects in the final state** to target specific production modes
- **VBF di-jet tag**
  - based on the output of a multivariate discriminant built using the di-photon BDT, $p_T^{\gamma\gamma}/m_{\gamma\gamma}$ and variables sensitive to the VBF topology:
    - photons $p_T$, $p_T^{jet1}$ , $p_T^{jet2}$, $m_{j1j2}$, $\Delta\eta_{j1j2}$, $\Delta\Phi(\gamma\gamma,jj)$, $Z=|\eta_{\gamma\gamma}-(\eta_{j1}+\eta_{j2})/2|$
- **WH, ZH tag**
  - one or more charged leptons, missing ET or jets from the W, Z decays
- **ttH tag**
  - b-quarks, leptons, additional jets from top decays

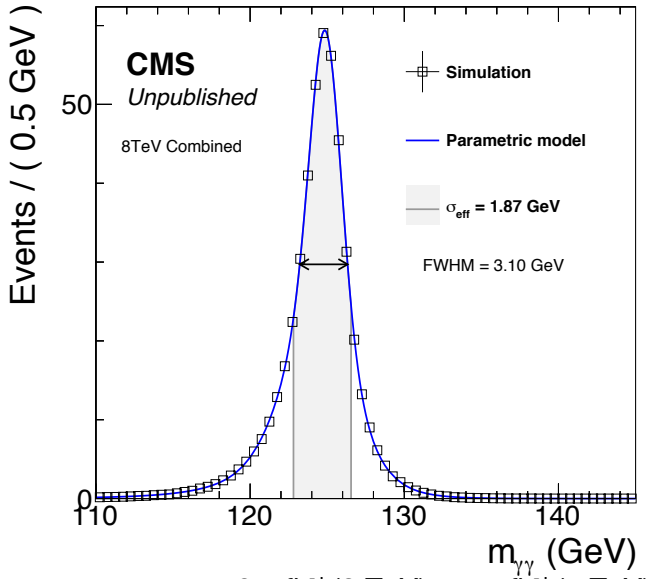


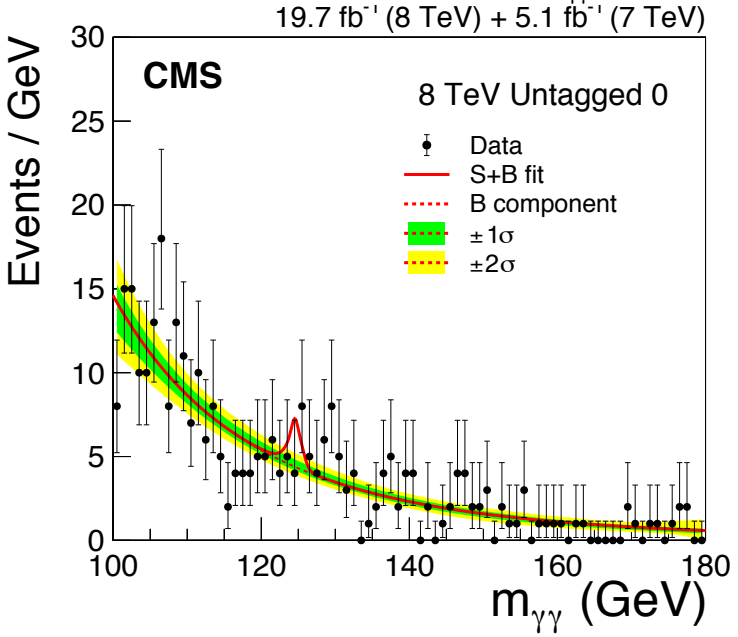

VBF di-jet tagged classes boundaries optimized to minimize the expected uncertainty in the signal strength associated with VBF production

# Signal and background models

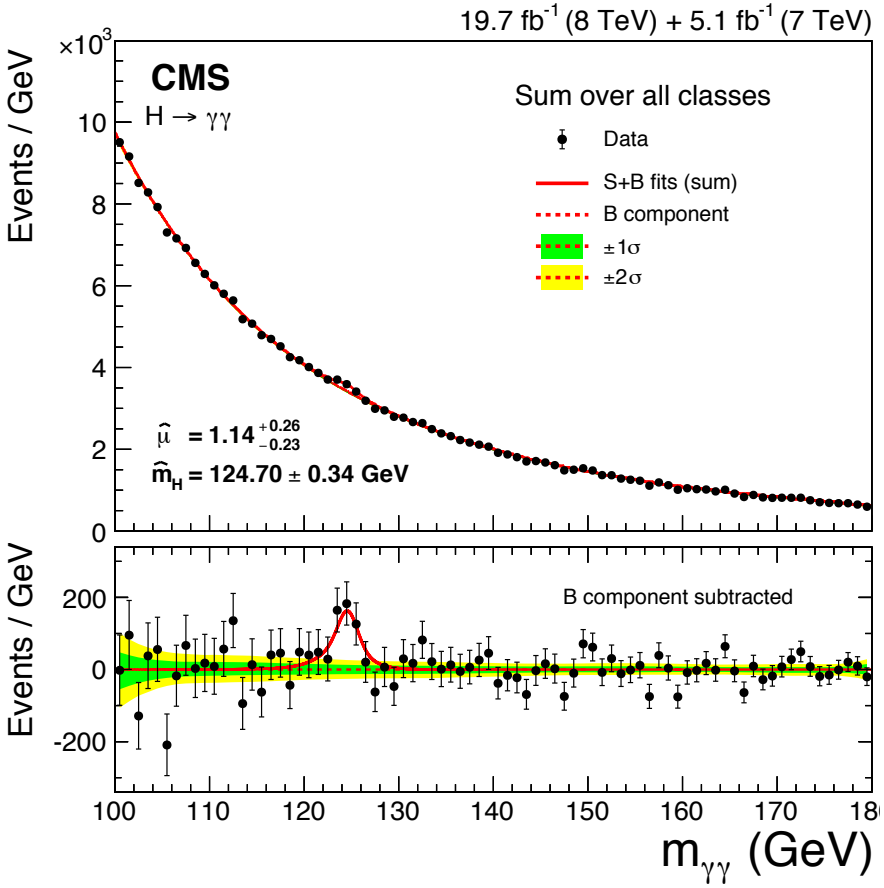
- **Signal model**
  - for each event class, parametric signal model built from fit of the simulated invariant mass after applying corrections derived from data/MC comparisons in Z→ee and Z→μμγ
  - best categories $\sigma_m \sim 1$ GeV

- **Background model fitted from data**
  - smoothly falling background, shape *a priori* unknown
  - use discrete profiling method → **choice of the function included as a discrete nuisance parameter in the likelihood to extract results**
  - all reasonable families of functions considered (expo, power law, polynomials, Laurent series)
  - allow the data to select the one which fits the best
  - subsequent "envelope" around NLL curve of different choices means that the uncertainty will take into account the model assumption

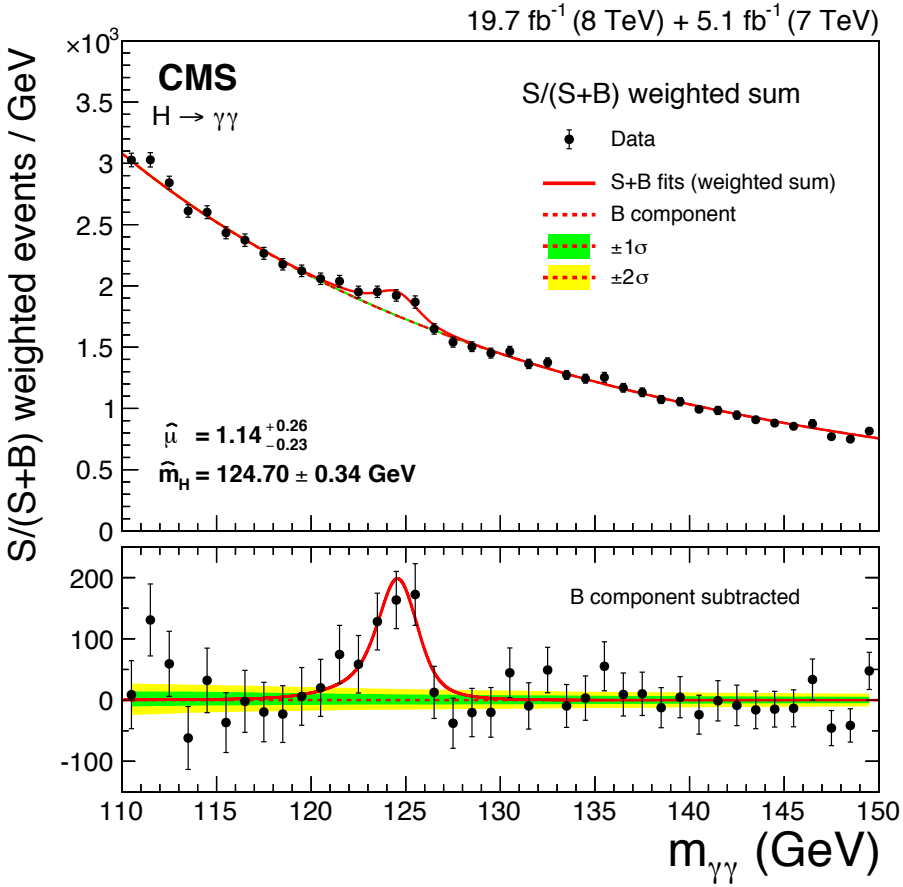


# Mass spectrum




inclusive



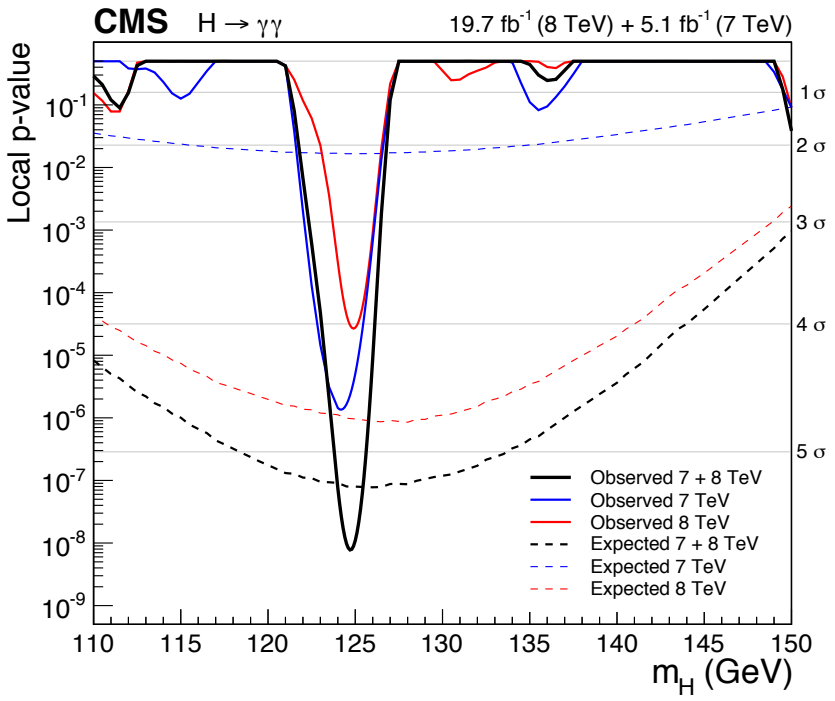

Inclusive – events weighted by sensitivity

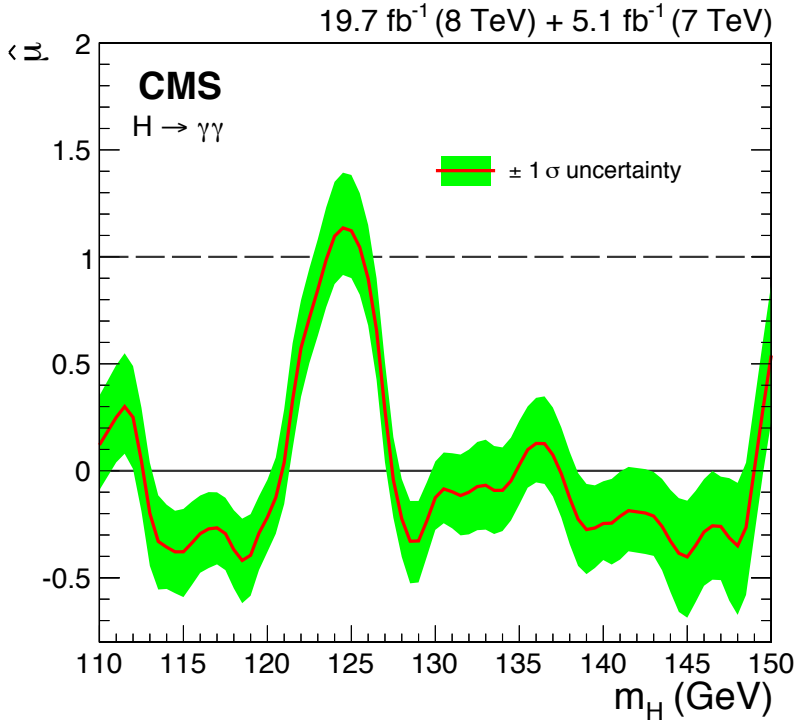
# Signal strength

| Dataset | Observed significance | $\mu = \sigma/\sigma_{SM}$ | $m_H$ (GeV) |
|---|---|---|---|
| 7 TeV | 4.7σ | $2.22^{+0.62}_{-0.55}$ | 124.2 |
| 8 TeV | 4.0σ | $0.90^{+0.26}_{-0.23}$ | 124.9 |
| **7+8 TeV** | **5.7σ** | **$1.14^{+0.26}_{-0.23}$** | **124.7** |

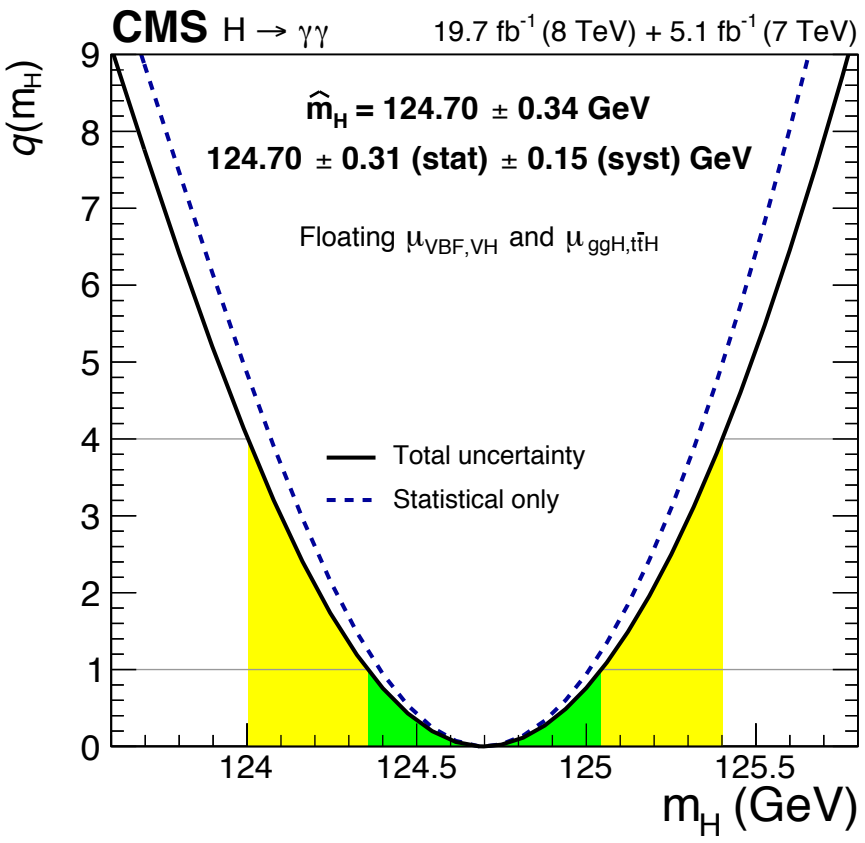
Main systematics affecting the signal yield

| Source of uncertainty | Uncertainty in $\hat{\mu}$ |
|---|---|
| Production cross sect. and branching frac. | 0.11 |
| Shower shape modelling (Section 9) | 0.06 |
| Energy scale and resolution | 0.02 |
| Other | 0.04 |
| All syst. uncert. in the signal model | 0.13 |
| Statistical | 0.21 |
| Total | 0.25 |



# Mass measurement




- 1D mass scan floating relative signal strength to fermions and bosons to make the measurement less model dependent

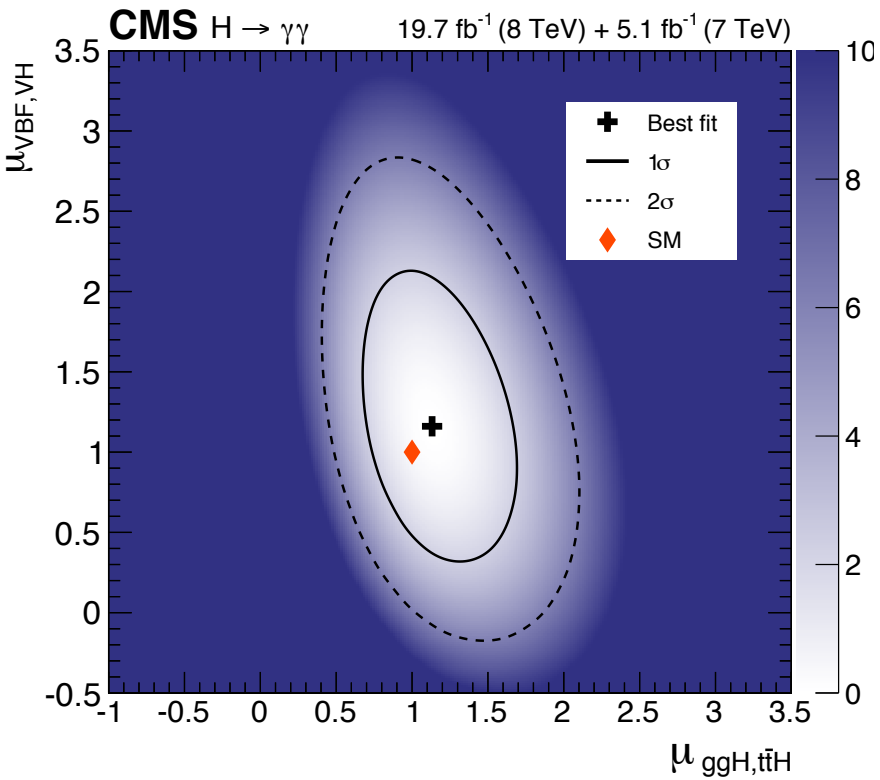
## Mass measurement systematic uncertainties

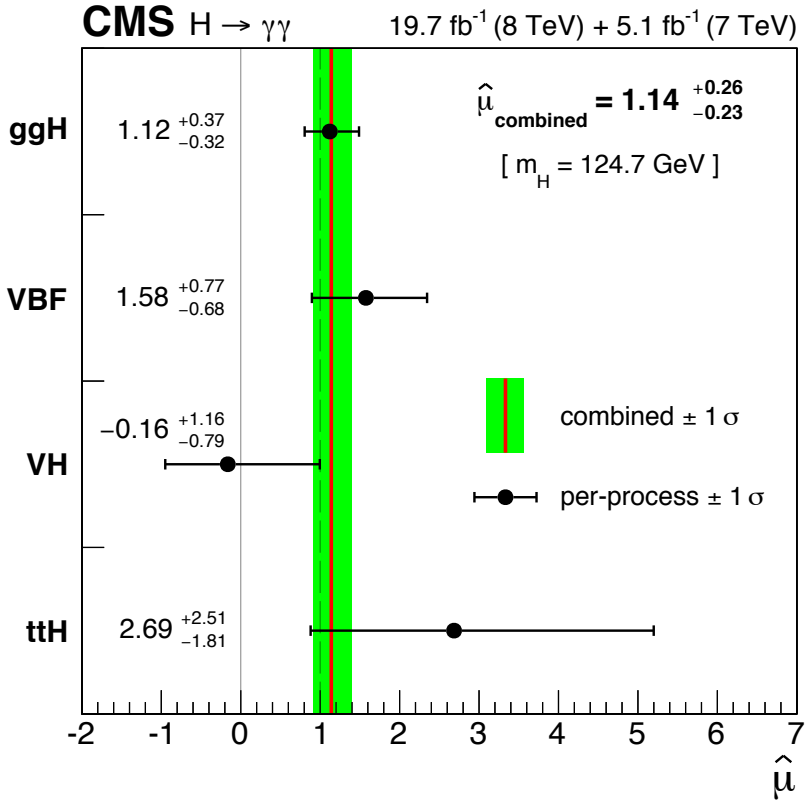
| Uncertainty | Details | Effect (GeV) |
|---|---|---|
| Energy scale calibration and resolution | • Uncertainty on the correction applied<br>• Use ET dependent corrections<br>• Also model stochastic and constant terms in resolution | ± 0.05 |
| Non-linearity in scale extrapolation from $m_Z$ to $m_H$ | • Imperfect modelling in MC of differences between showers from Z→ee at $m_Z$ scale and H→γγ at $m_H$ scale<br>• $E_T$ dependent scale corrections reduce this | ± 0.10 |
| Electron – photon differences not modelled in MC | • Tracker material mis-modelling (inflate tracker material)<br>• Variation in scintillation light peak between electrons and photons<br>• Imperfect EM shower simulation in Geant4<br>• Imperfections in OOT-PU description | ± 0.10 |
| Other | | ± 0.04 |

# Production mechanisms

- Signal strength by production type
- Consistent with SM expectations

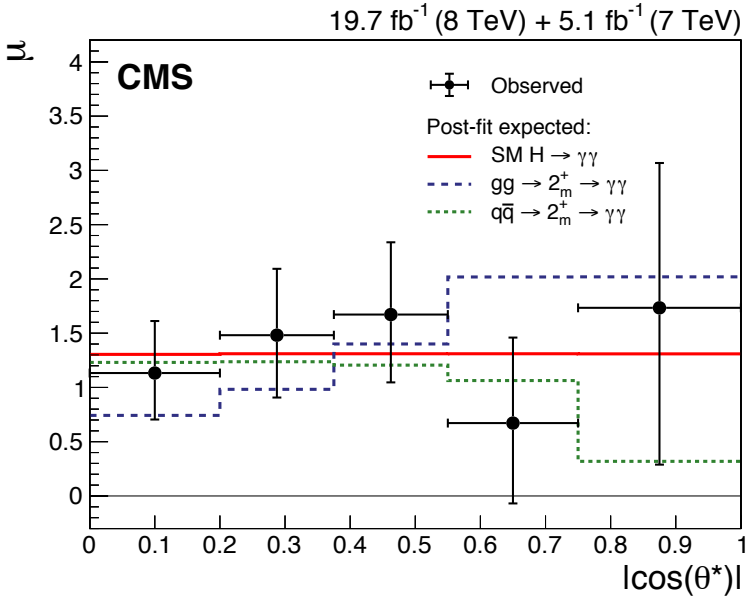
| | Expected | Observed |
|---|---|---|
| $\hat{\mu}_{ggH,t\bar{t}H}$ | $1.00^{+0.34}_{-0.30}$ | $1.13^{+0.37}_{-0.31}$ |
| $\hat{\mu}_{VBF,\,VH}$ | $1.00^{+0.57}_{-0.51}$ | $1.15^{+0.63}_{-0.58}$ |



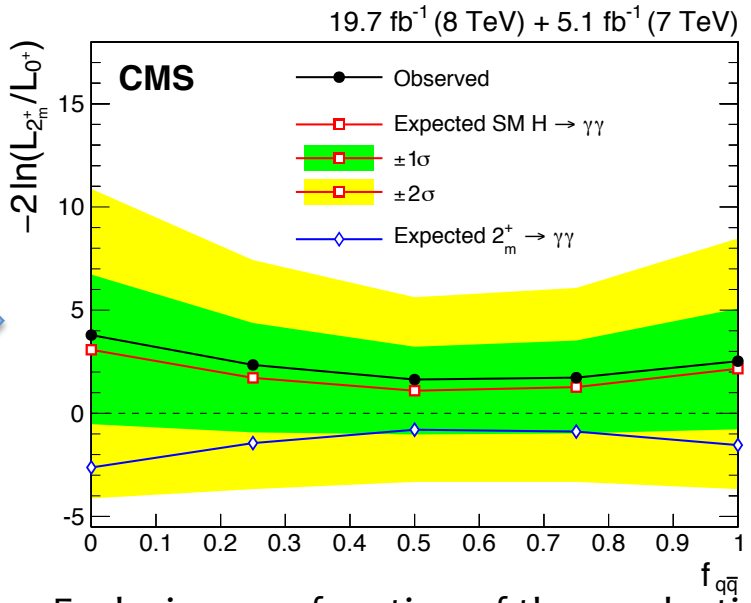



# Testing the spin hypothesis

- Test $0^+$ **vs** $2^+_m$ ("graviton" like)
  - different admixtures of ggH and qqH production considered for $2^+_m$
- Discriminating variable: $\cos\Theta^*_{CS}$ , scattering angle in the Collins-Soper frame
- 4 di-photon categories (η,R9) x 5 categories in $\cos\Theta^*_{CS}$ x 2 (7, 8 TeV) = 40 categories
- Simultaneous S+B fit to classes with signal strength variable in each $\cos\Theta^*_{CS}$ bin and common $m_H$




Flat observed μ vs $\cos\Theta^*_{CS}$ as expected under SM $0^+$ hypothesis




Exclusion as a function of the production mode: $2^+_m$ excluded at 94% C.L. for pure gluon fusion production

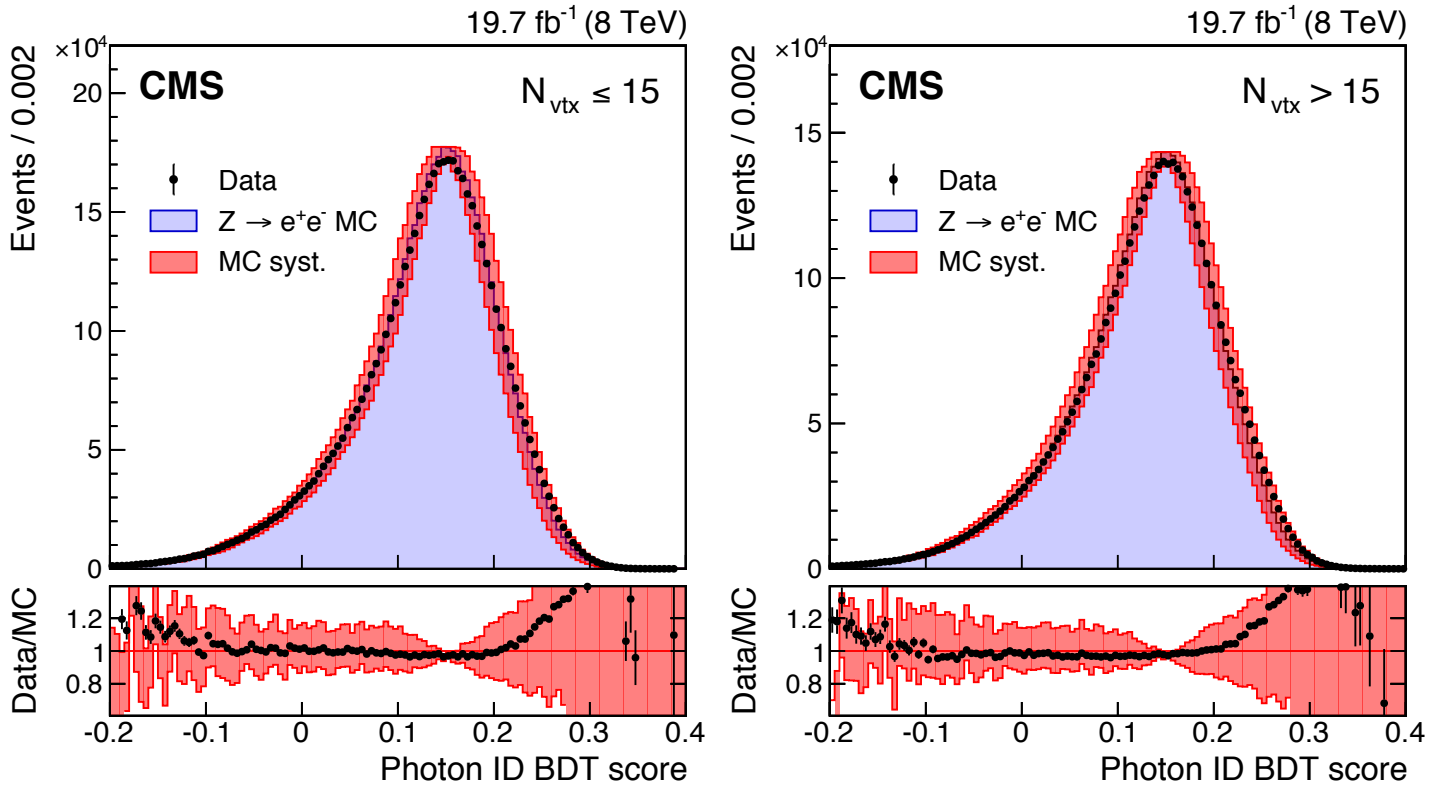
# Conclusions

- The CMS analysis in search for SM H→γγ using the full Run I dataset was presented
- Observed a clear signal with a local significance of 5.7 σ
- Mass of the boson: 124.70 ± 0.31(stat.) ± 0.15(syst.) GeV
- Signal strength: $1.14 \pm 0.31(\text{stat.}) \,^{+0.09}_{-0.05}(\text{syst.}) \,^{+0.13}_{-0.09}(\text{theo.})$
- Properties seem very consistent with a SM Higgs around 125 GeV

# Backup slides

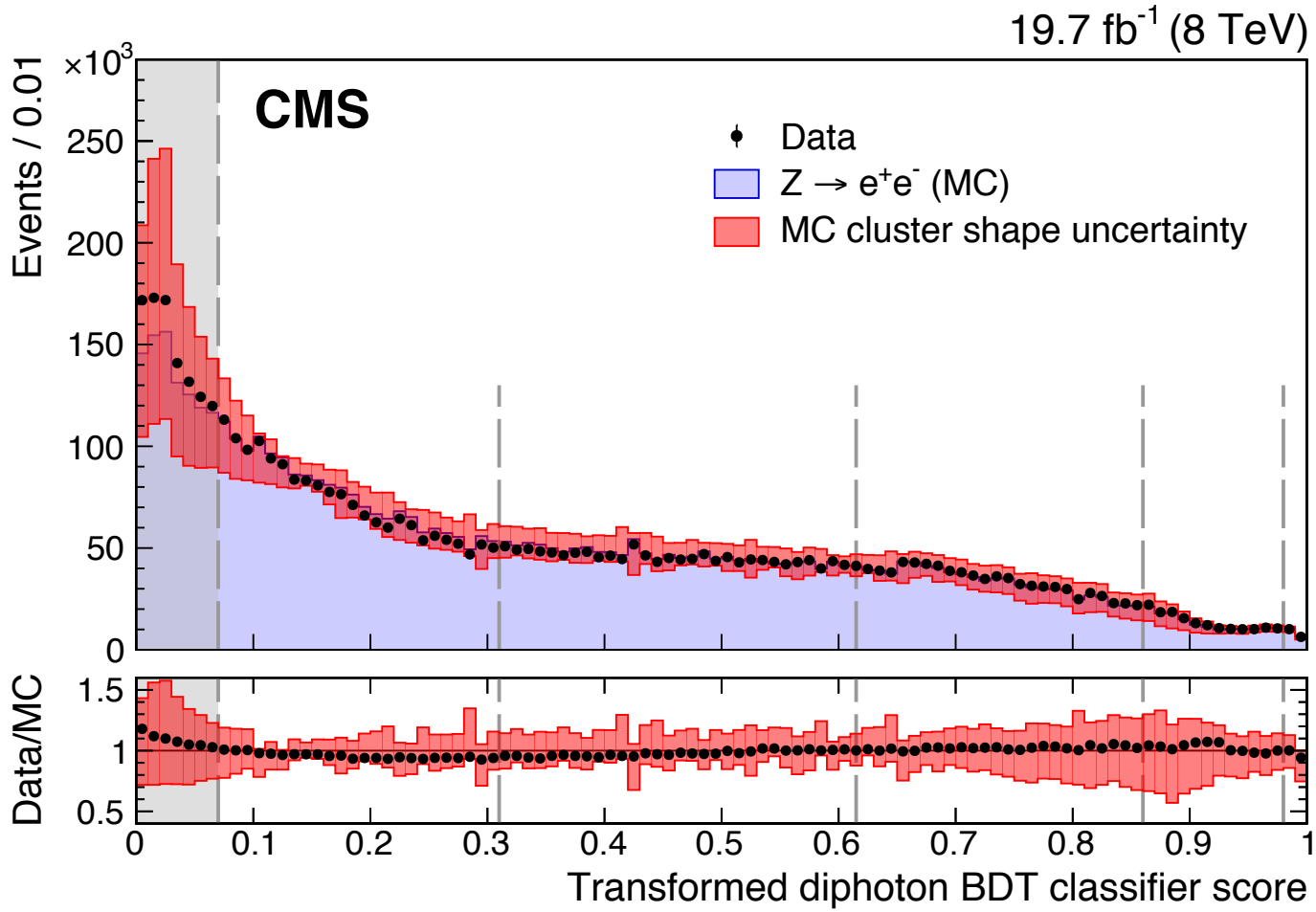
# Photon ID BDT

- Validation photon ID BDT



# Di-photon BDT

- Di-photon BDT classifier score for Z → e+e− events in 8 TeV data, and in MC simulation



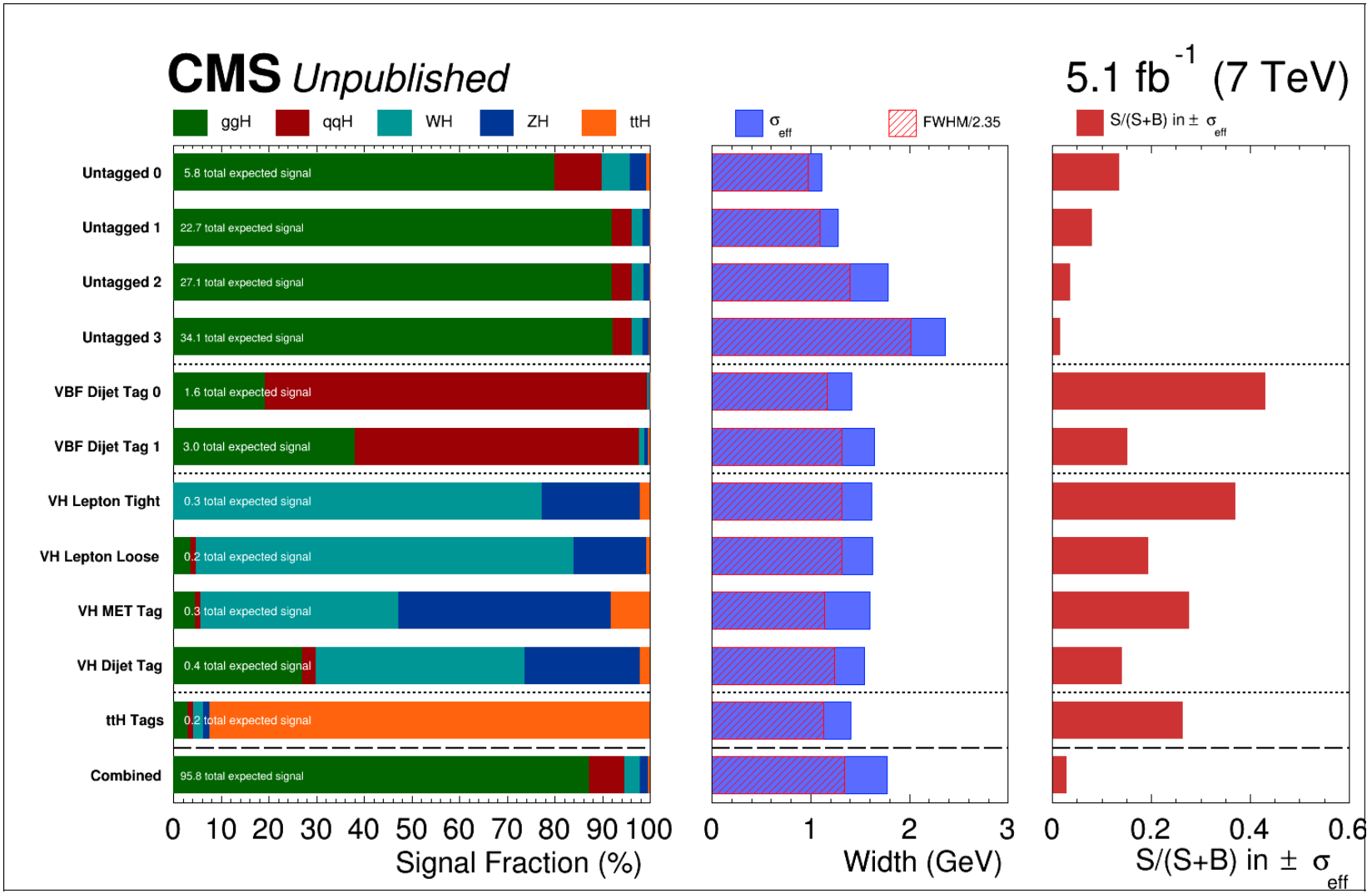
# Event classes details

| Label | No. of classes 7 TeV | 8 TeV | Main requirements |
|---|---|---|---|
| $t\bar{t}H$ lepton tag | $\star$ | 1 | $p_T^{\gamma 1} > m_{\gamma\gamma}/2$<br>1 b-tagged jet + 1 electron or muon |
| VH tight $\ell$ tag | 1 | 1 | $p_T^{\gamma 1} > 3m_{\gamma\gamma}/8$<br>[e or $\mu$, $p_T > 20$ GeV, and $E_T^{miss} > 45$ GeV] or<br>[2e or 2$\mu$, $p_T^{\ell} > 10$ GeV; $70 < m_{\ell\ell} < 110$ GeV] |
| VH loose $\ell$ tag | 1 | 1 | $p_T^{\gamma 1} > 3m_{\gamma\gamma}/8$<br>e or $\mu$, $p_T > 20$ GeV |
| VBF dijet tag 0-2 | 2 | 3 | $p_T^{\gamma 1} > m_{\gamma\gamma}/2$<br>2 jets; classified using combined diphoton-dijet BDT |
| VH $E_T^{miss}$ tag | 1 | 1 | $p_T^{\gamma 1} > 3m_{\gamma\gamma}/8$<br>$E_T^{miss} > 70$ GeV |
| $t\bar{t}H$ multijet tag | $\star$ | 1 | $p_T^{\gamma 1} > m_{\gamma\gamma}/2$<br>1 b-tagged jet + 4 more jets |
| VH dijet tag | 1 | 1 | $p_T^{\gamma 1} > m_{\gamma\gamma}/2$<br>jet pair, $p_T^{j} > 40$ GeV and $60 < m_{jj} < 120$ GeV |
| Untagged 0-4 | 4 | 5 | The remaining events,<br>classified using diphoton BDT |

$\star$ For the 7 TeV dataset, events in the $t\bar{t}H$ lepton tag and multijet tag classes are selected first, and combined to form a single event class.

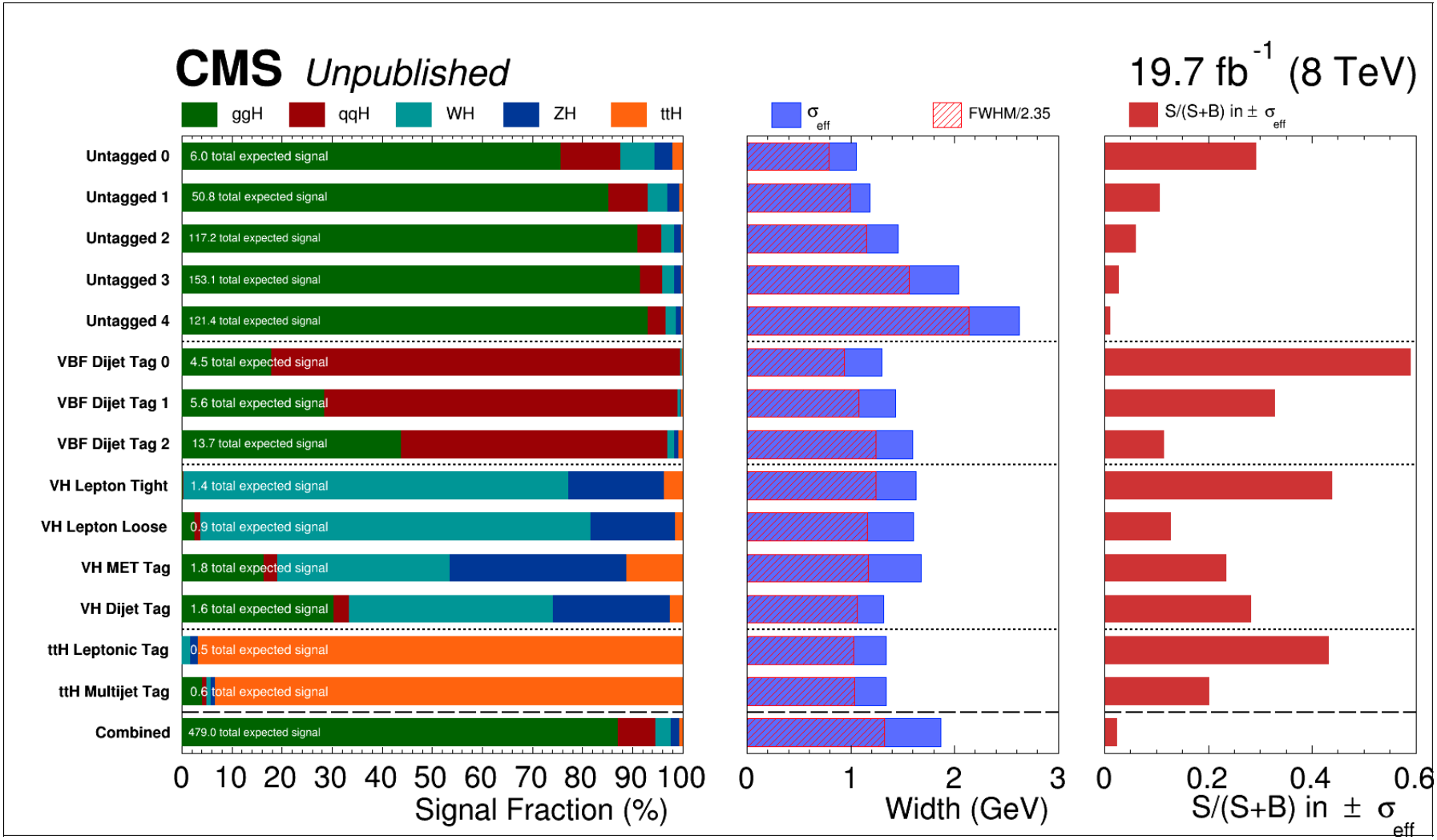
# Event classes composition (7 TeV)

- Event classes composition



# Event classes composition (8 TeV)

- Event classes composition



# Channel compatibility