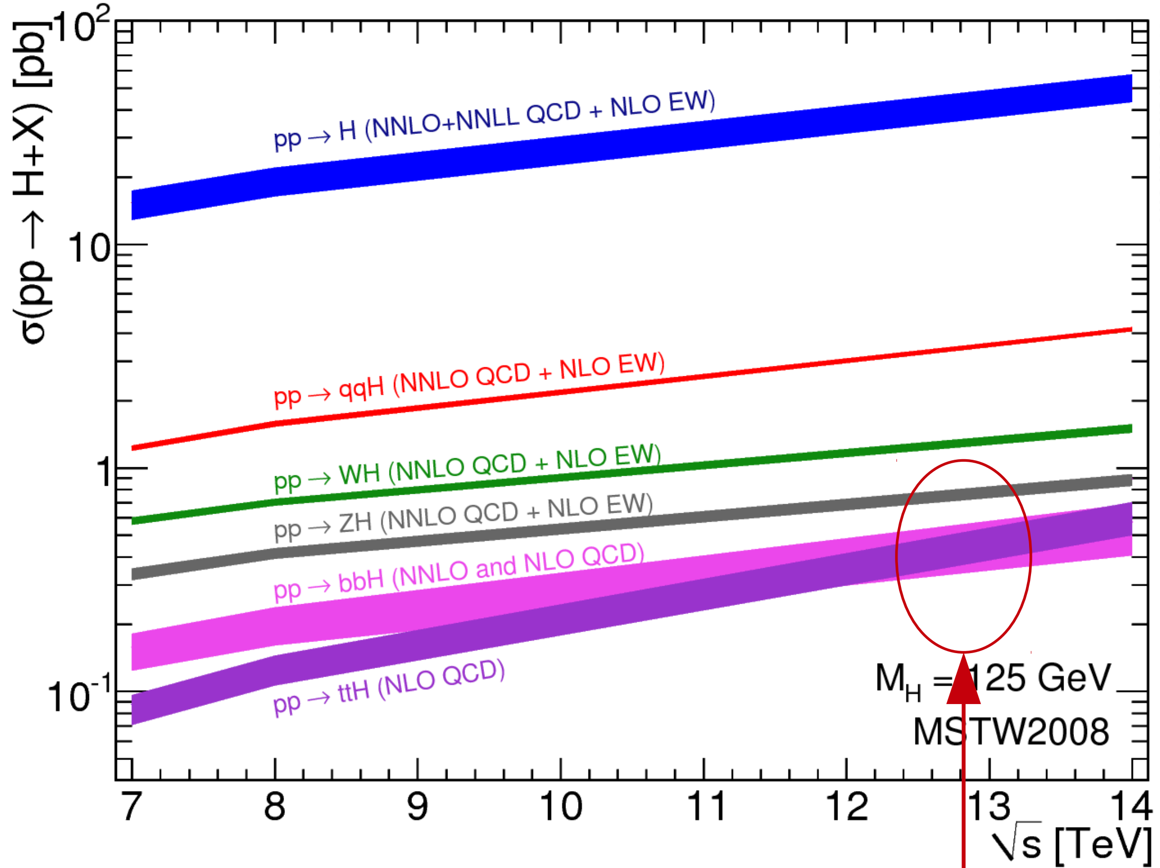
Latest CMS results on
Higgs → γγ

Photon 2017
22-26 June

Hamed BAKHSHIANSOHI
Université catholique de Louvain

Higgs Boson
Production @ 13 TeV

Production cross section increases by a factor of **~2** in 13 TeV in **~all** production mechanisms

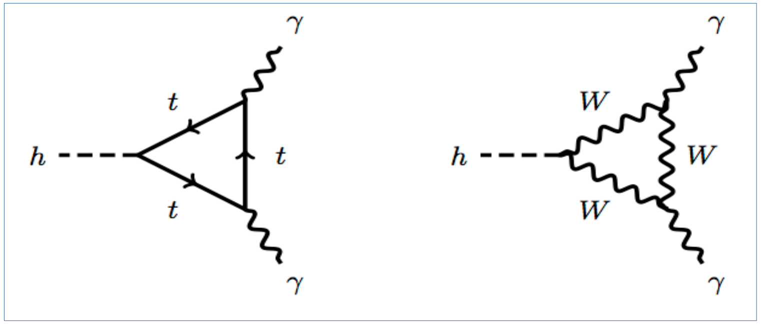


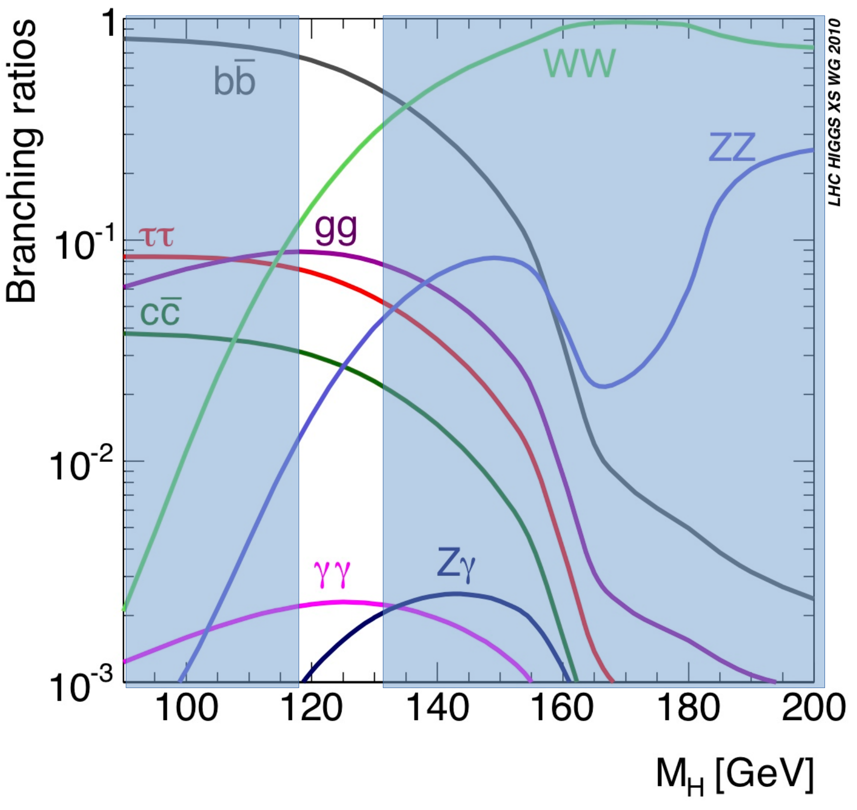
$t\bar{t}H$ production increases by a factor of ~4

Higgs Boson

Decay to photons

- No direct Higgs-Photon interaction
- Higgs decays to two photons via loops
- The Higgs → γγ branching-fraction is ~0.2%
  - Almost the least branching fraction for Higgs(MH=125GeV)



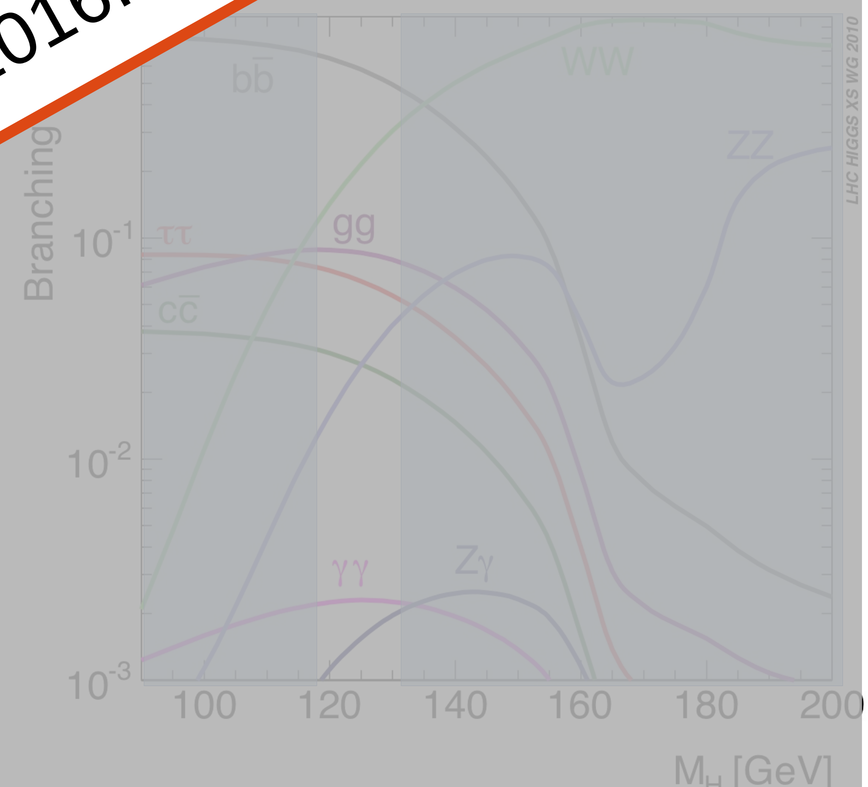


Higgs Boson

Decay to photons

- No direct Higgs-Photon interaction
- Higgs decays to two photo… via loops
- The Higgs → γγ … fraction is ~…
  - Almost … branching fraction … Higgs(MH=125GeV)

Total number of $H \rightarrow \gamma\gamma$ in 35.9 fb^{-1} of pp collisions at CMS in 2016: ~4.2K



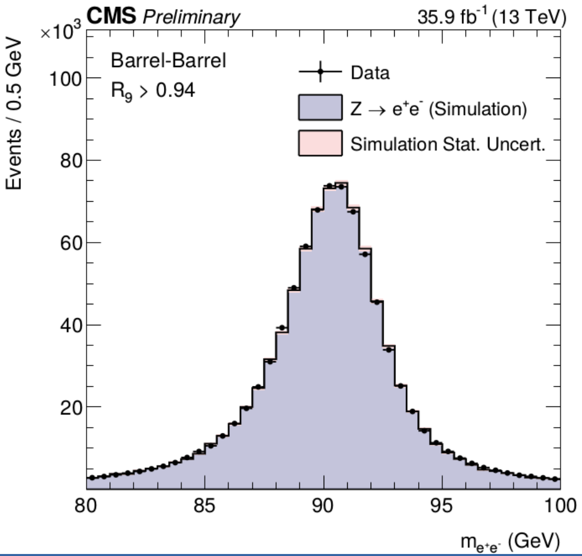
Data and Simulation

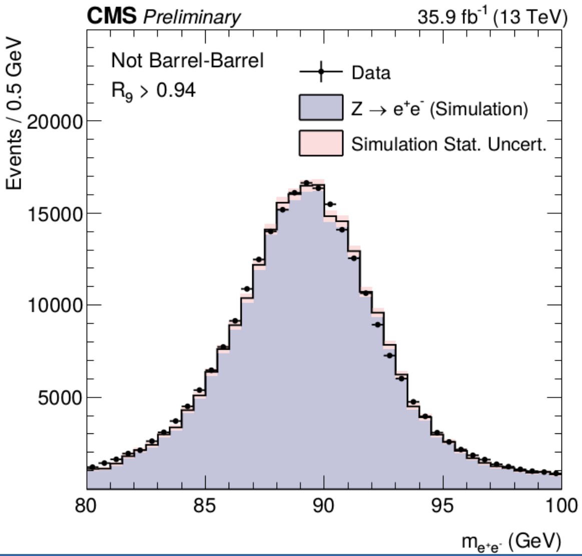
- **35.9 fb^{-1}** of pp collision recorded by CMS
- **Online selection:**
  - Events with two photon candidate
  - Loose isolation and shower shape criteria
  - $p_T^1 > 30$ GeV and $p_T^2 > 18$ GeV
- **Simulation:**
  - Higgs signals are generated at **NLO** using **MG5_aMC@NLO**
  - Showering and hadronization in **Pythia8**
- **Documented in**
  - CMS-HIG-16-040: Inclusive
  - CMS-HIG-17-015: Differential and Fiducial

Photon reconstruction in CMS

- ECAL energy clusters not linked to the extrapolation of any charged-particle trajectory to the ECAL
  - No matched pixel detector hit
- Collection of complete energy of the photon in clustering
  - Contain the energy of converting photons
- The energy is corrected for the EM showers and energy loss of converted photons
  - Regression method trained in simulation
- Photon energy scale applied in data
  - Simulation is smeared to reproduce the energy resolution of data

- Validated in Z→ ee events, where electron showers are reconstructed as photons





Photon Selection

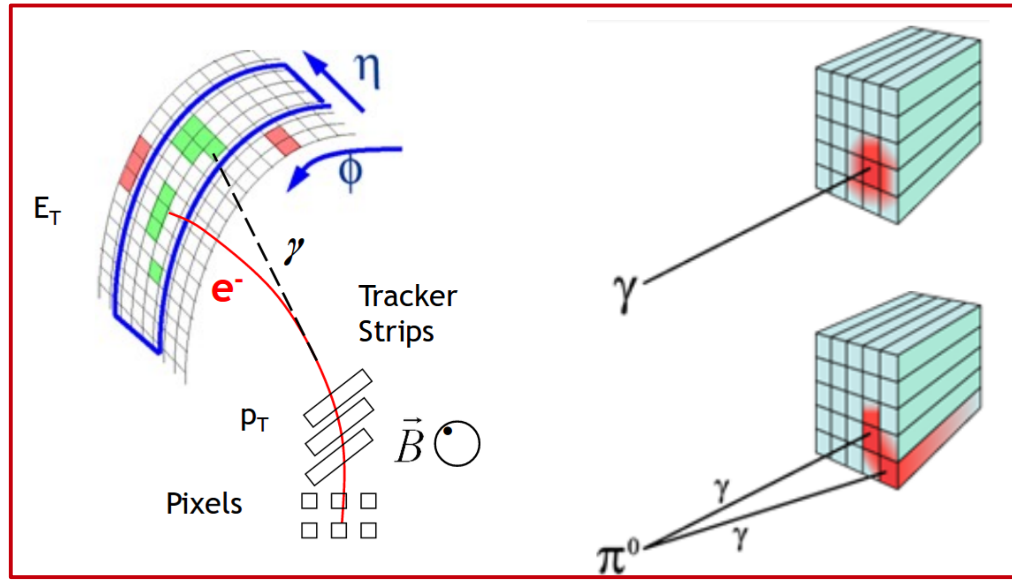
Pre-selection

- $p_T^{1(2)} > 30(20)$ GeV
- R9 and $\sigma_{\eta\eta}$
- H/E energy
- Electron Veto
- Isolation (photons, tracks, charged hadrons)

Efficiencies from $Z \rightarrow ee$ events in data and MC

For electron-veto: $Z \rightarrow \mu\mu$

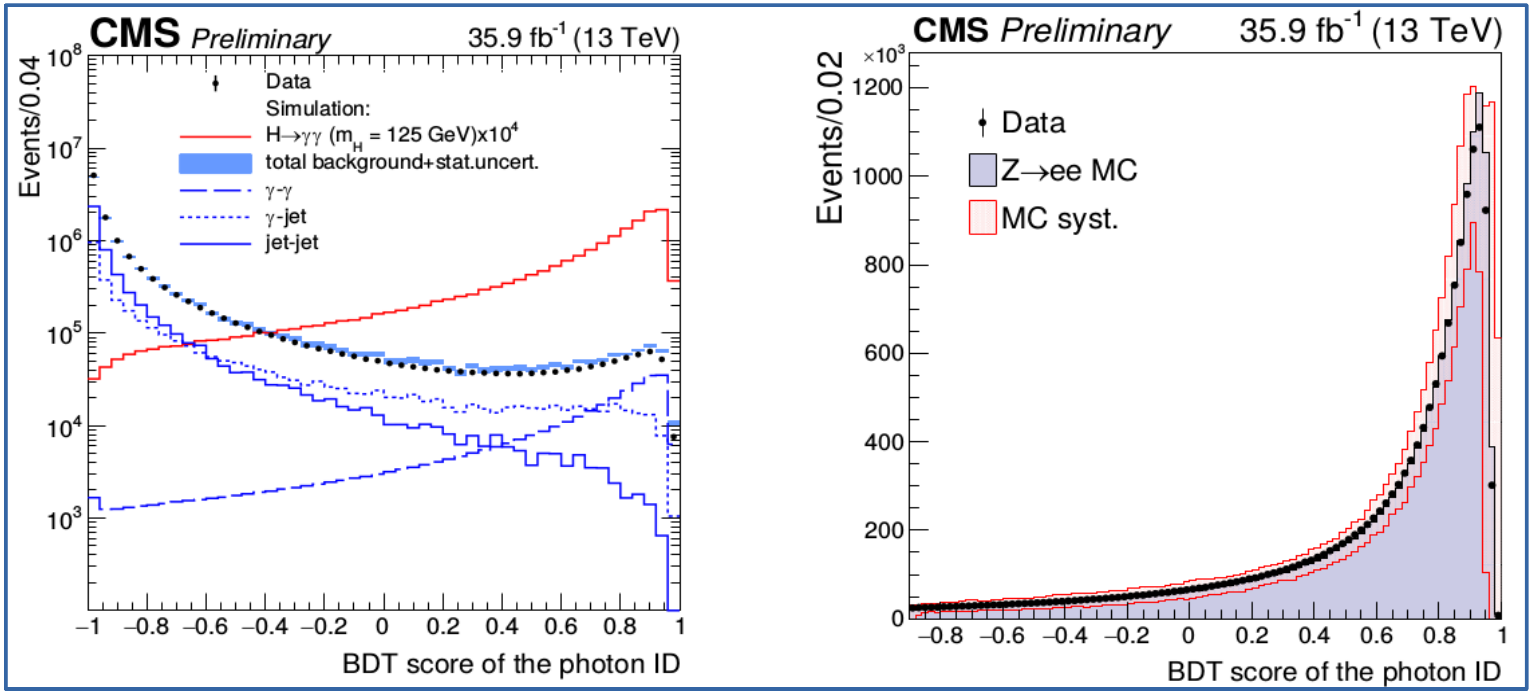
	R_9	H/E	$\sigma_{\eta\eta}$	$\mathcal{I}_{ph}$	$\mathcal{I}_{tk}$
Barrel	$[0.5, 0.85]$	< 0.08	< 0.015	< 4.0	< 6.0
	> 0.85	< 0.08	–	–	–
Endcaps	$[0.8, 0.90]$	< 0.08	< 0.035	< 4.0	< 6.0
	> 0.90	< 0.08	–	–	–



Preselection category	ϵ_{data} (%)	ϵ_{MC} (%)	$\epsilon_{data}/\epsilon_{MC}$
Barrel; R_9 >0.85	94.2 ± 0.9	94.7 ± 0.9	0.995 ± 0.001
Barrel; R_9 <0.85	82.5 ± 0.7	82.5 ± 0.7	1.000 ± 0.003
Endcap; R_9 >0.90	90.1 ± 0.2	91.3 ± 0.1	0.987 ± 0.005
Endcap; R_9 <0.90	49.7 ± 1.4	53.8 ± 1.5	0.923 ± 0.010

Photon Identification

- Boosted Decision Tree (BDT) trained using γ+jet simulated events
- Input variables
  - All the variables used for selection
  - the median energy density per unit area in the event



Diphoton Vertex

$\Delta z_{vertex} < 1cm$ → Mass resolution dominated by photon energy resolution

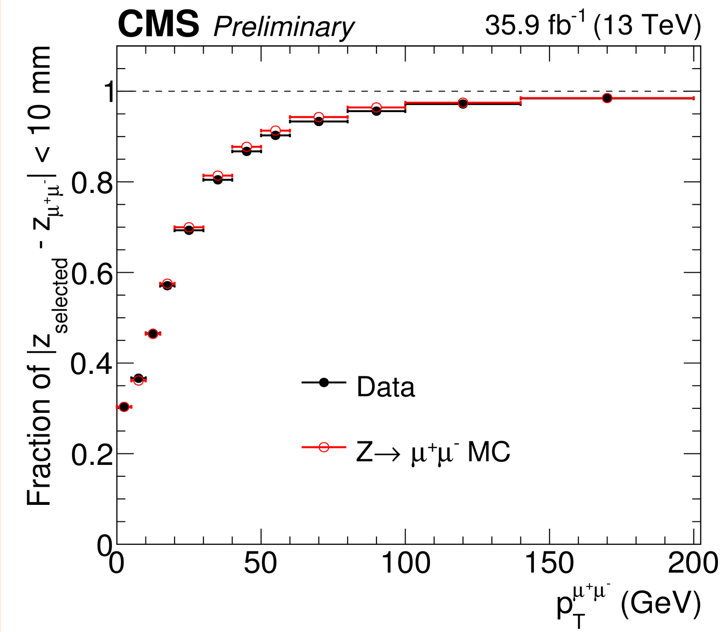
- **BDT** is trained to assign the photon-pair to the right vertex

- **Input variables:**
  - Observables related to tracks recoiling against the diphoton system
  - For photons with conversion tracks, information of the vertex of those tracks are also included

- $\sum_i |\vec{p}_T^{\,i}|^2$,
- $-\sum_i (\vec{p}_T^{\,i} \cdot \frac{\vec{p}_T^{\,\gamma\gamma}}{|\vec{p}_T^{\,\gamma\gamma}|})$,
- $(|\sum_i \vec{p}_T^{\,i}| - p_T^{\gamma\gamma})/(|\sum_i \vec{p}_T^{\,i}| + p_T^{\gamma\gamma})$,

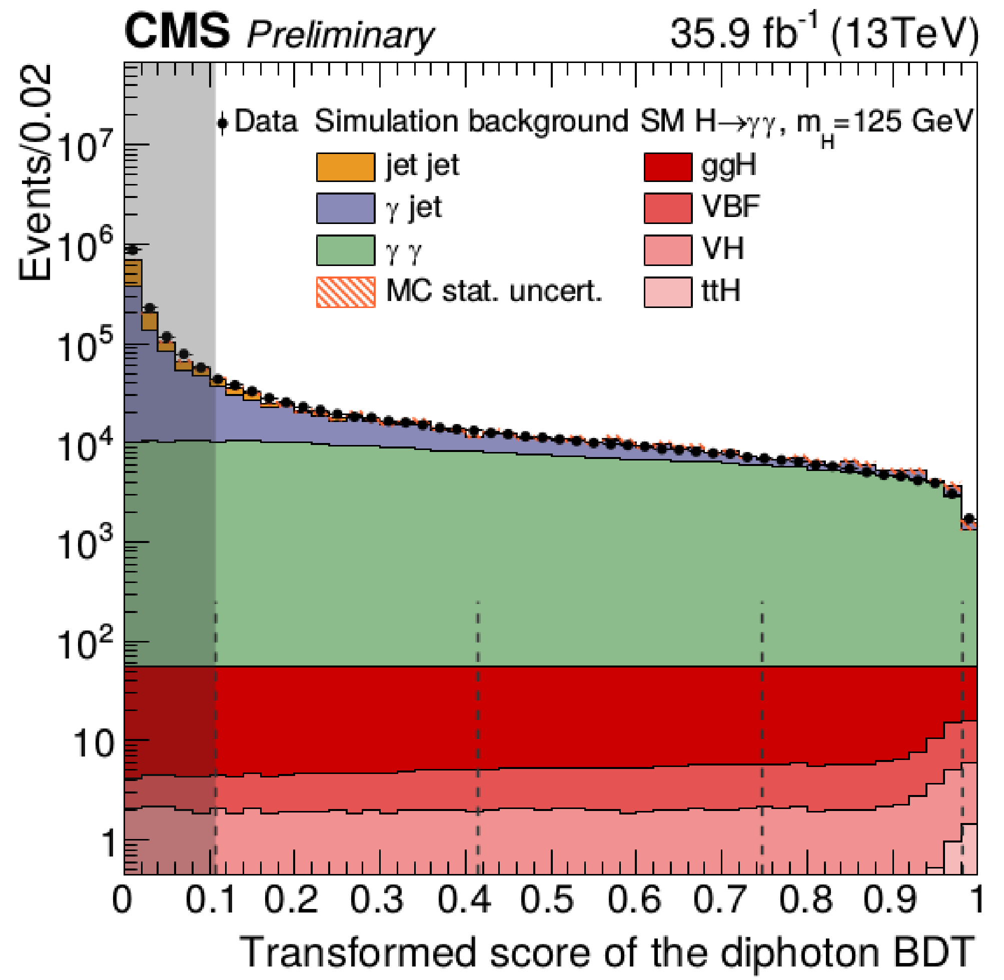
- Vertex assignment is validated using $Z \rightarrow \mu\mu$ events

- The assigned vertex is within 1cm of the real one in 81% of cases



Diphoton identification

- **BDT** Trained to evaluate the di-photon mass resolution per-event
  - And select signal-like events
- **Input variables:**
  - PT / Mass and η of each photon
  - Δφ of the photons
  - Photon ID BDT scores
  - Mass resolution estimates
  - Probability of the correct vertex assignment
- **Validated** in Z→ ee events
- Used for **categorization** of “untagged” events



Event Categorization

ttH

- A dedicated BDT is trained
  - Cut optimized along with γγ BDT ID
- Leptonic
  - At least one lepton, far enough from photons
  - At least 2 jets, one is tagged as b-jet
- Hadronic
  - No lepton
  - At least 3 jets, one is tagged as b-jet

VBF

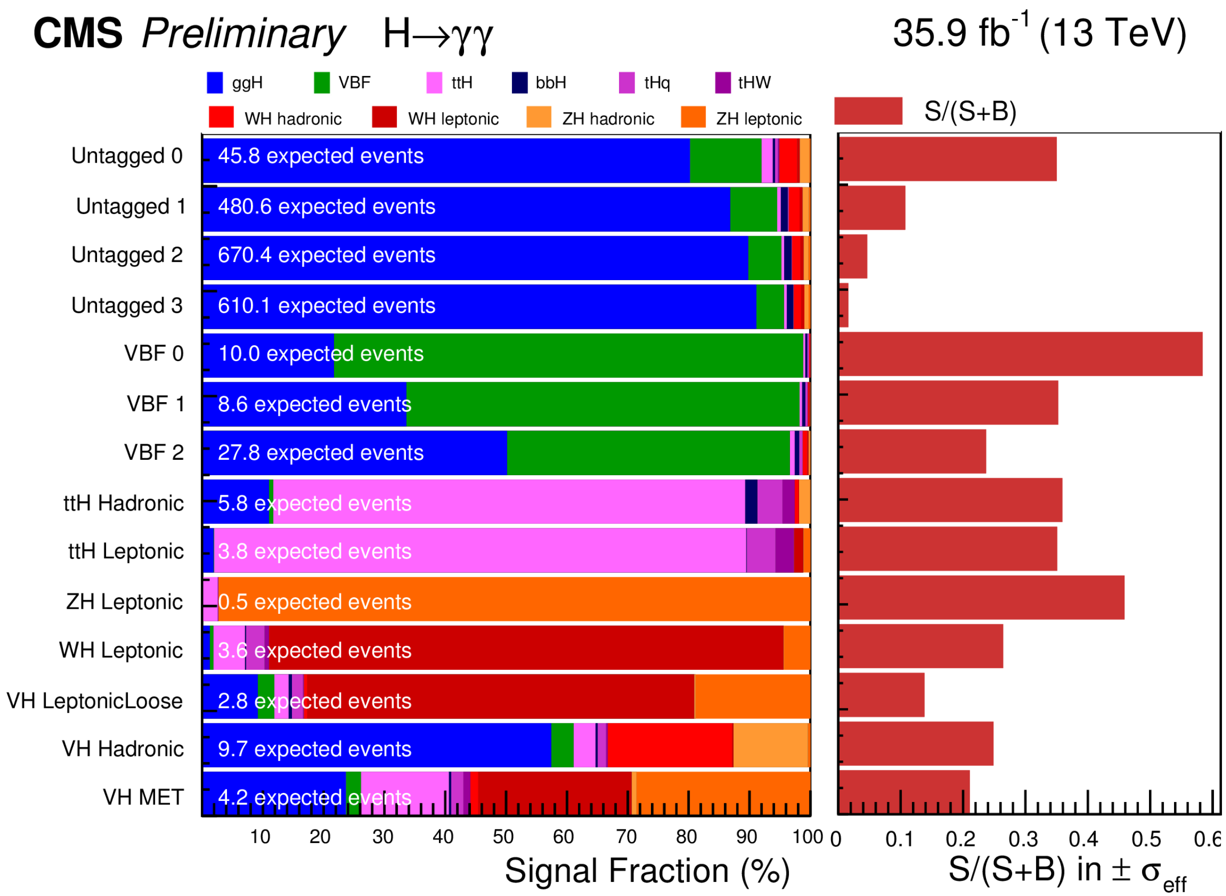
- A dedicated BDT is trained
  - Forward jet kinematics
  - Optimize exp. significance in 3 categories
- Two forward jets, m(jj) > 250

V (Z,W) H

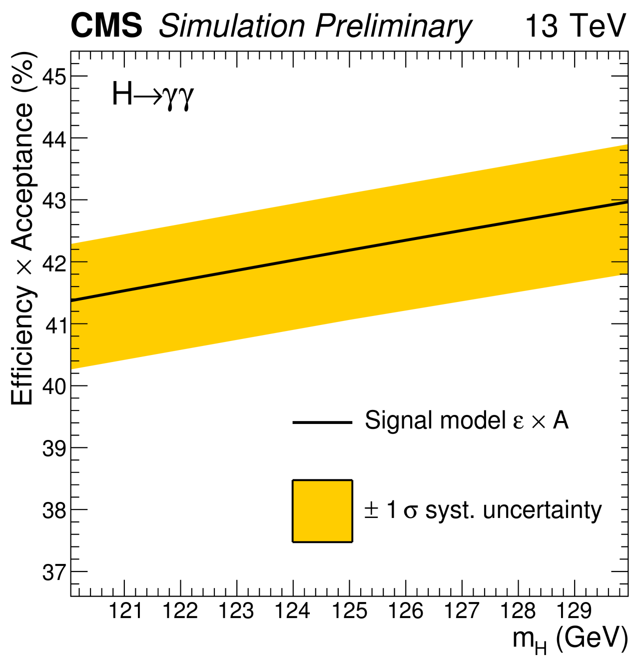
- ZH-Leptonic
  - Two leptons with 70<M<110
- WH-Leptonic
  - At least one lepton, MET > 45 GeV
  - nJets < 3
- VH-Leptonic Loose
  - Same as WHLeptonic, MET < 45 GeV
- VH-MET
  - MET > 85 GeV
  - $\Delta\varphi$(MET,γγ) > 2.4
- VH-Hadronic
  - At least two jets, 60 < m(jj) < 120 GeV

TTH Leptonic → 3 VH Leptonic → TTH Hadronic → VBF 0-2 → VH MET → VH Hadronic → Untagged 0-3

Selection Results



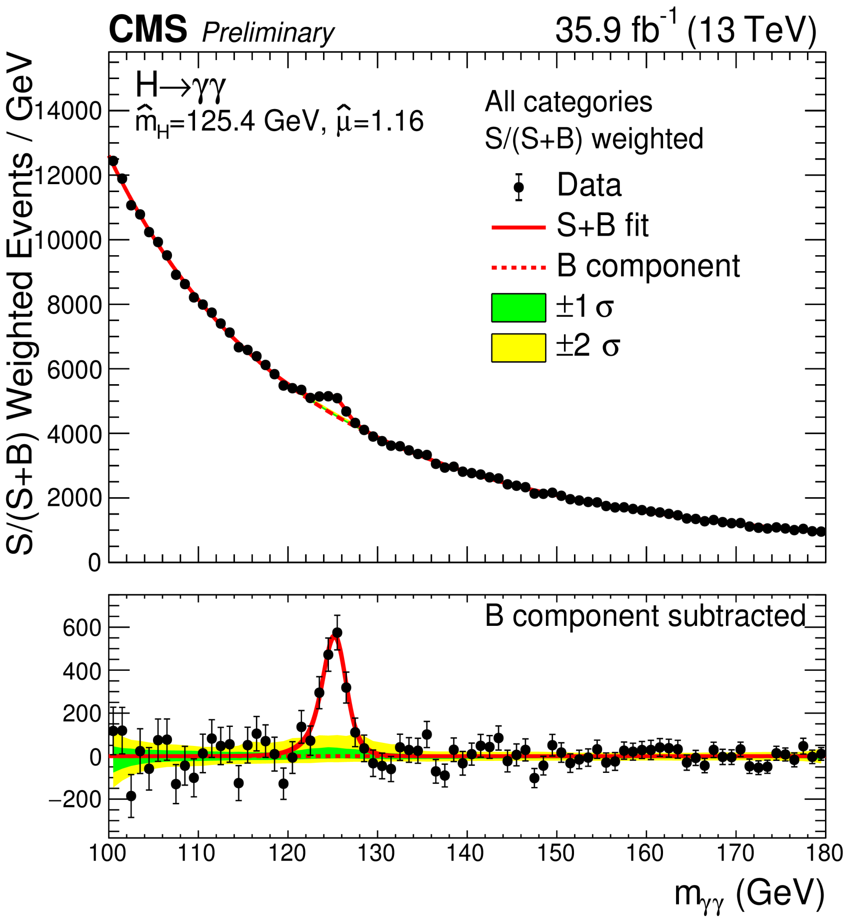
Yields in different categories for different Higgs production modes / Backgrounds



ε × Acc. vs. Higgs Mass

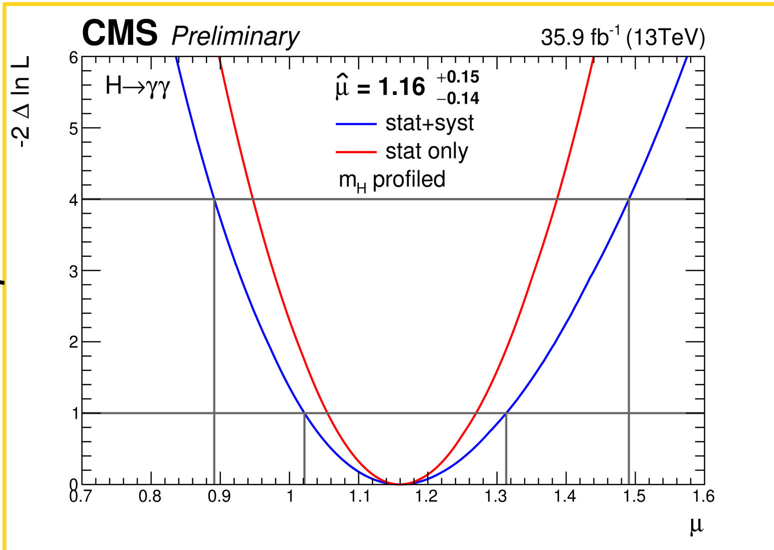
Search for the signal

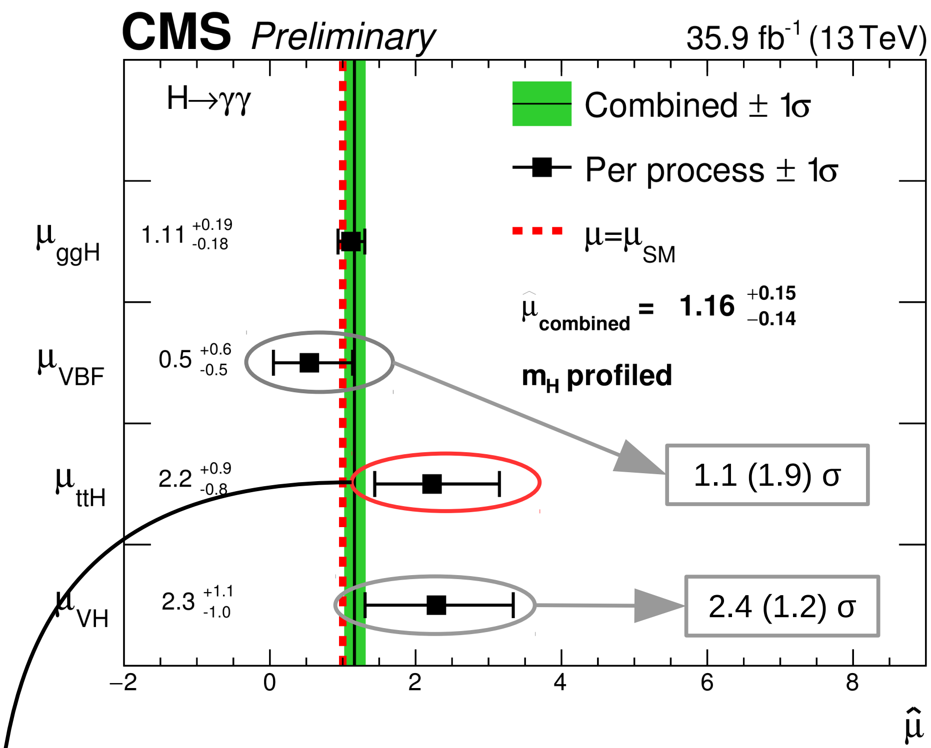
- **Strategy:**
  - Search for the Higgs bump in the falling background
- **Modeling the signal shape:**
  - Separate for Wrong-Vertex and Right-Vertex events
  - Sum of n-Gaussians (n < 5)
- **Modeling the falling background:**
  - Fully data-driven, fitting across the full range
  - A large set of functions are fitted:
    - *Exponentials, Bernstein polynomials, Laurent and power law functions*
  - F-Test to determine the maximum order to be used



Results

$$\hat{\mu} = 1.16^{+0.15}_{-0.14} = 1.16^{+0.11}_{-0.10}\ \text{(stat.)}\ ^{+0.09}_{-0.08}\ \text{(syst.)}\ ^{+0.06}_{-0.05}\ \text{(theo.)}$$

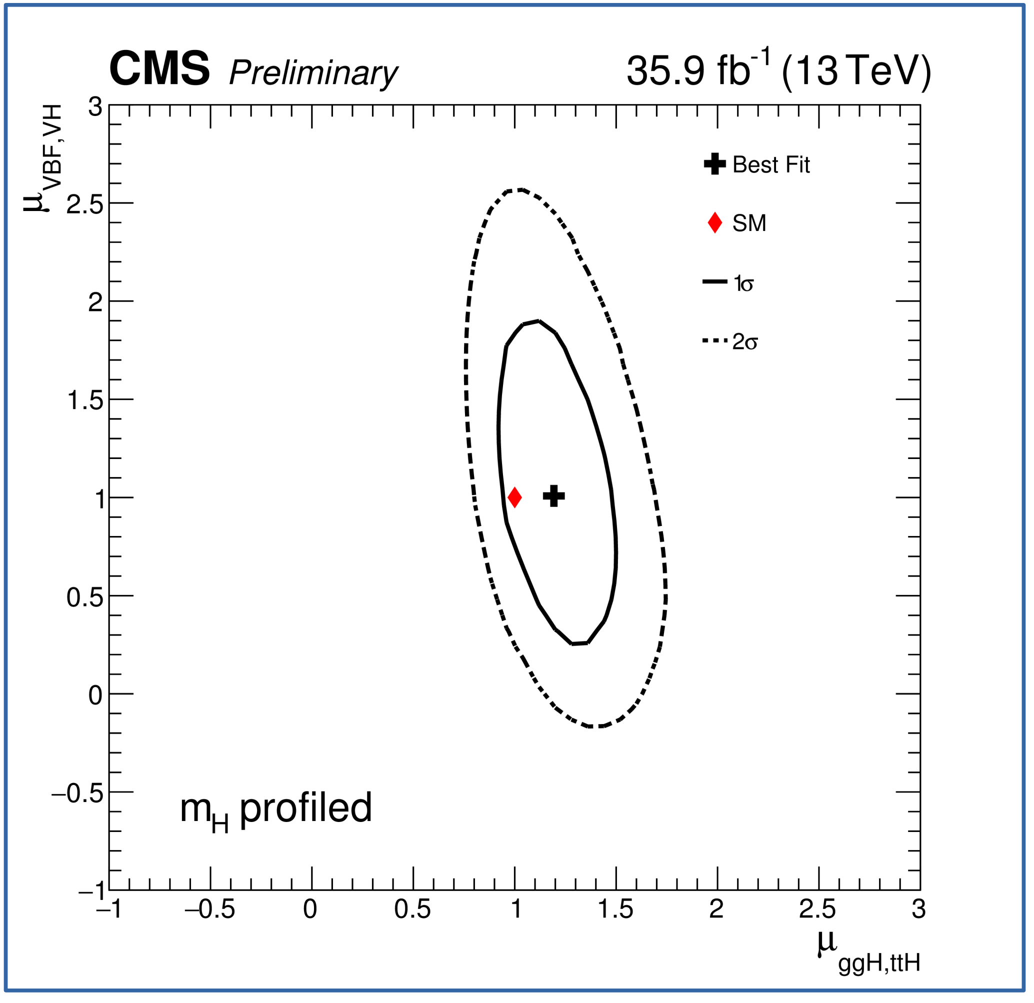


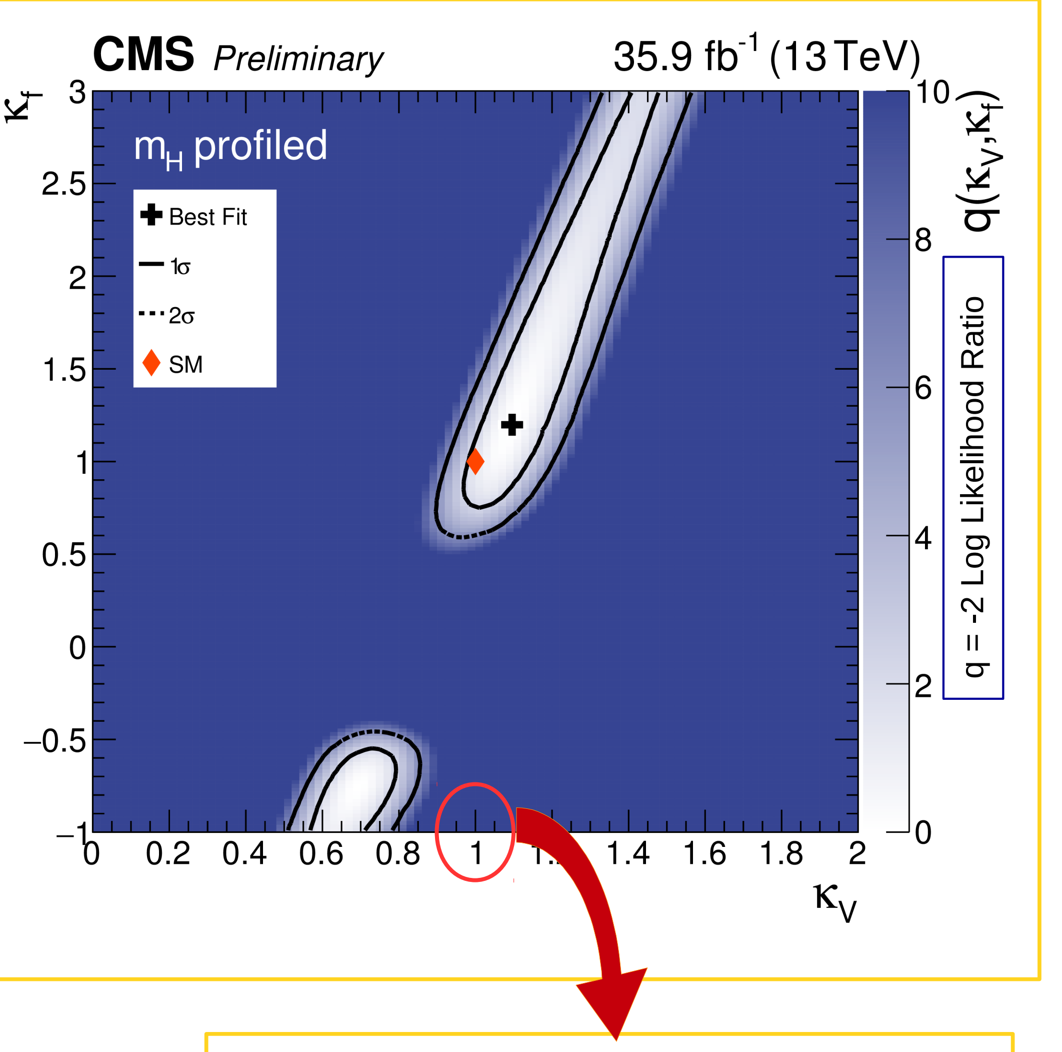


Best Fit Mass:
M_H= 125.4 (GeV) ± 0.15 (stat.) ± ~0.3 (syst.)

Observed (Expected) Signal Strength for $t\bar{t}H$: 3.3 (1.5) σ

Results on Higgs Couplings





q = 35.6 ~ 5.6σ away from best fit

Differential/Fiducial cross section measurement

HIG-17-015

Categorization

Based on “Uncorrelated mass resolution” estimator

$$y(\sigma_m/m|m) = \int_0^{\sigma_m/m} f(\sigma_m/m'|m)\, d\sigma_m/m'$$

uniformly distributed, thus not correlated with mass

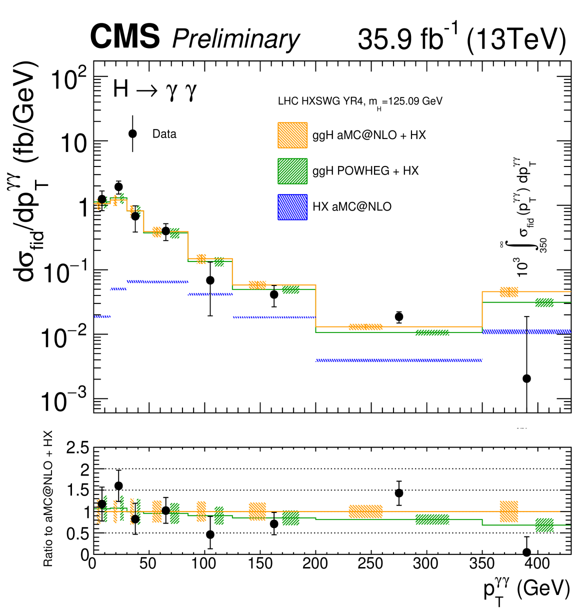
Make cumulative distribution functions

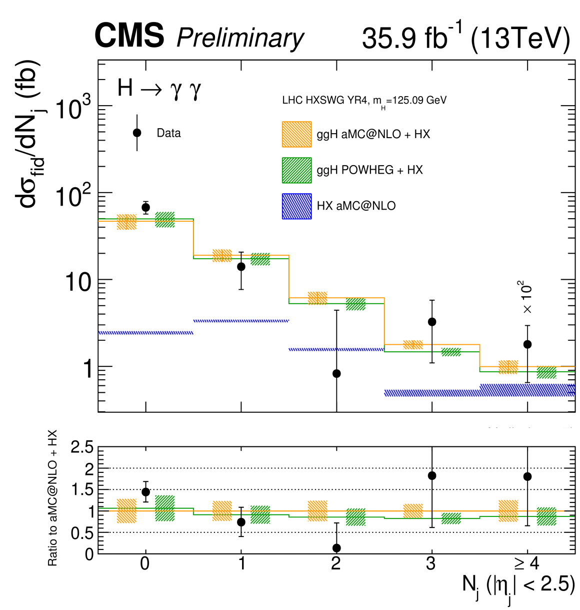
Constructed by sorting values of σ_m/m in bins of m

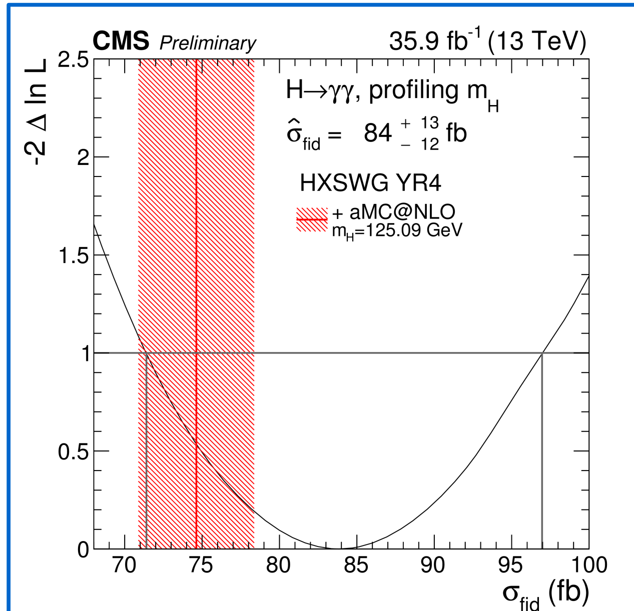
- Verified using Z→ ee events, with electrons reconstructed As photons
- 3 Categories enough to saturate the maximum achievable sensitivity

Results

- Data is unfolded to particle-level using a likelihood fit
- Total fiducial cross section in $p_T^{1(2)}$=m/3(4) and |η|<2.5 is measured
- Differential fiducial cross section in $p_T^{\gamma\gamma}$ and nJets bins are reported







Summary

- Higgs → $\gamma\gamma$ is (one of) the most precise channel to study the Higgs properties

- CMS has developed dedicated selection to tag photon pairs coming from Higgs decay
  - overall efficiency ~45%

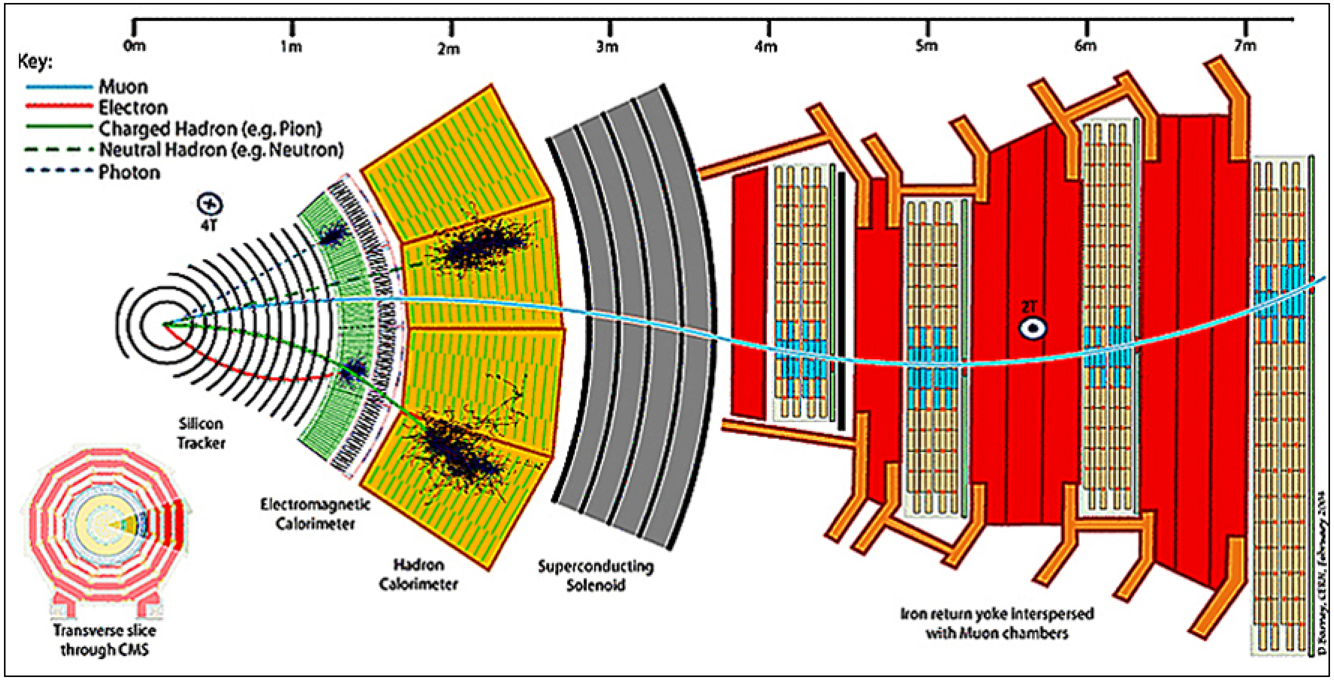
- 2016 dataset has been analyzed by CMS and first results were presented
  - Some evidence of ttH production mode
  - Diff. results up to p_T^{Higgs}~2×m_{top} and n_{Jets}>3 were presented

- Stay tuned for the most precise mass measurement in this channel using the 2016 dataset

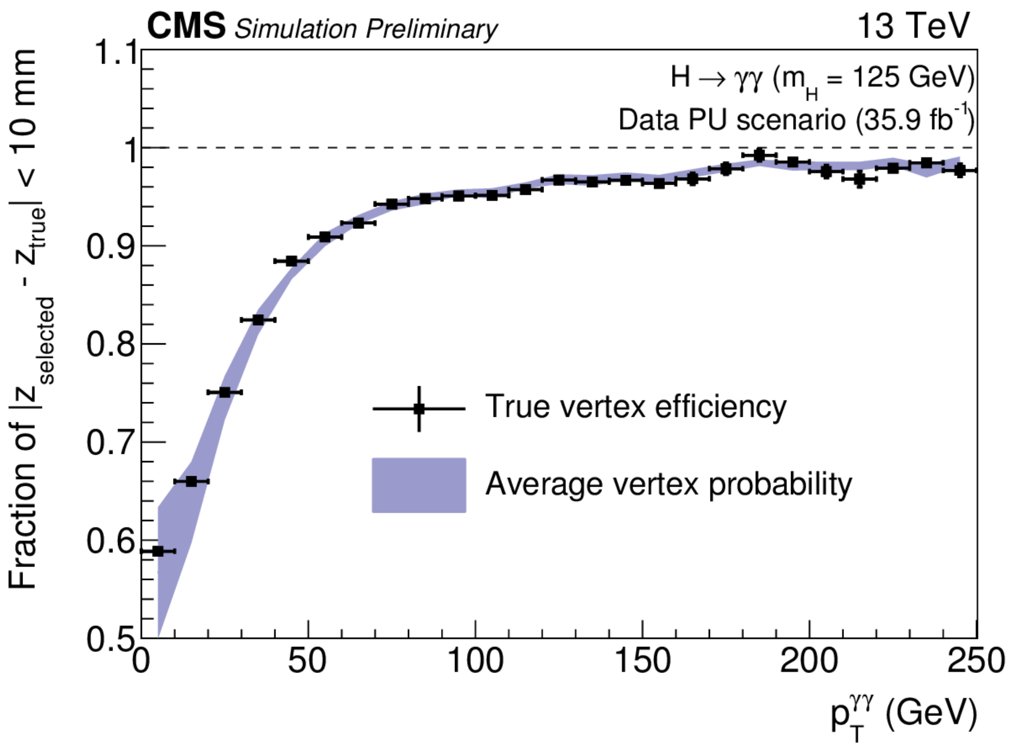
BACKUP

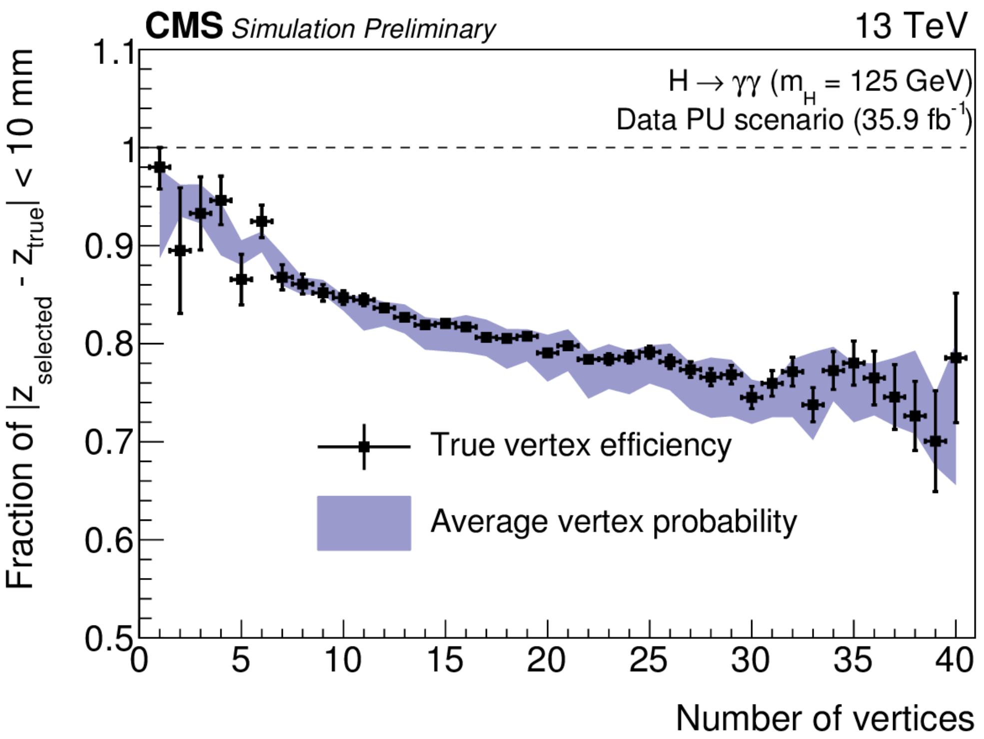
The CMS detector

- Silicon pixel and strips trackers
  - Covering |η|<2.5
- lead-tungstate crystal electromagnetic calorimeter (ECAL)
  - Barrel extends to |η|<1.48
  - Endcap : 1.48<|η|<3.0
- A lead/silicon-strip preshower detector in front of the ECAL-Endcap

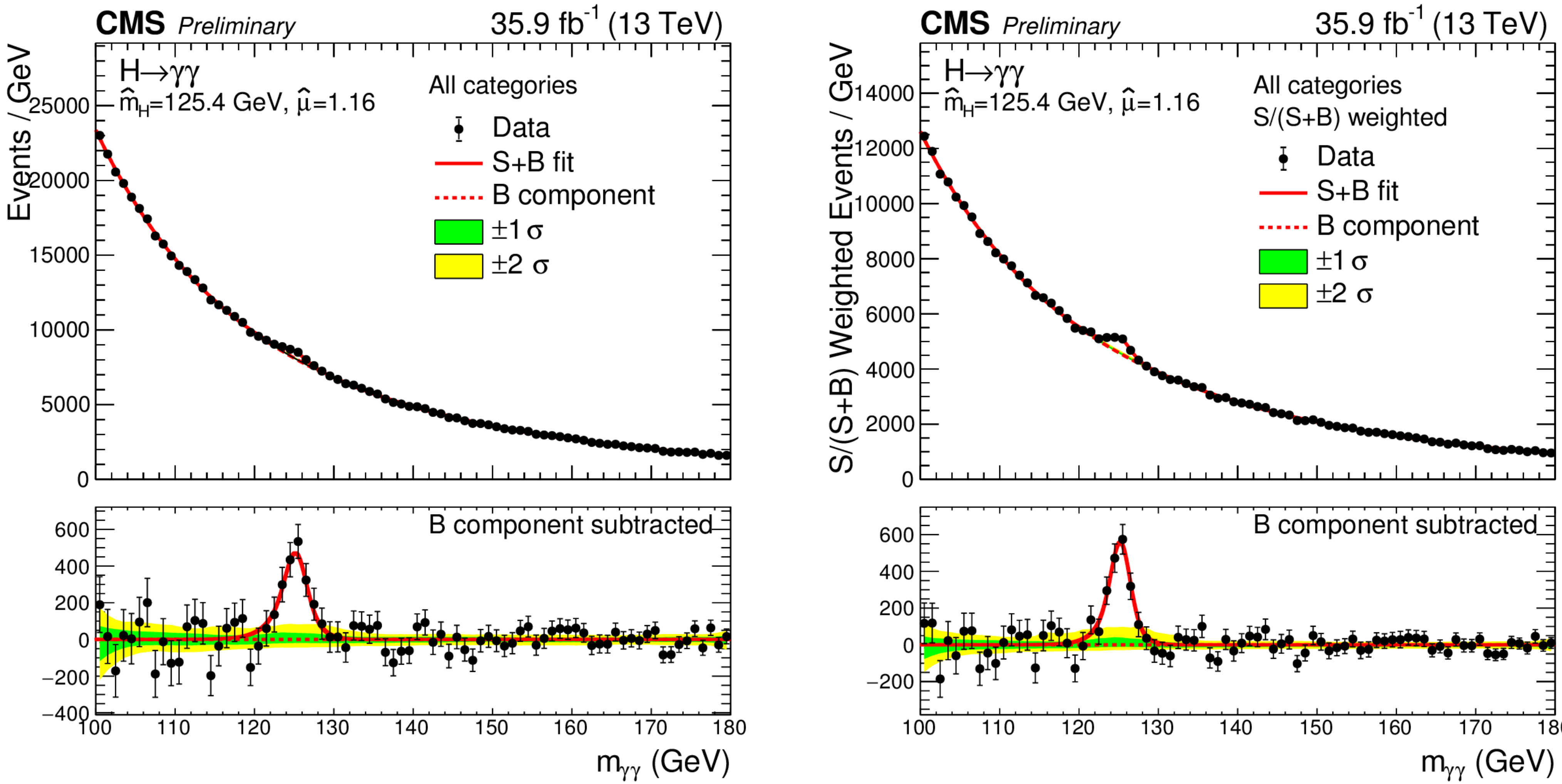


Vertex assignment efficiency

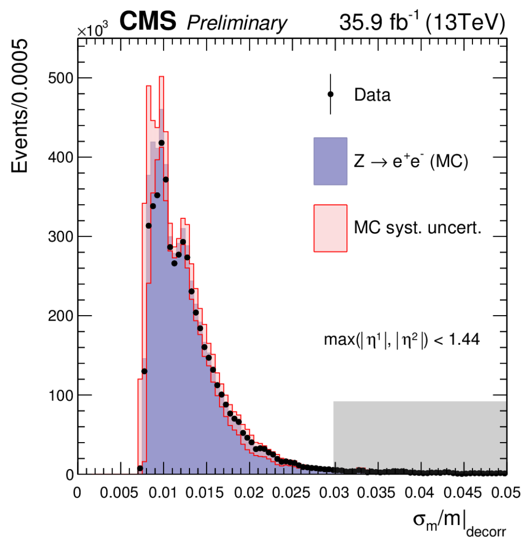




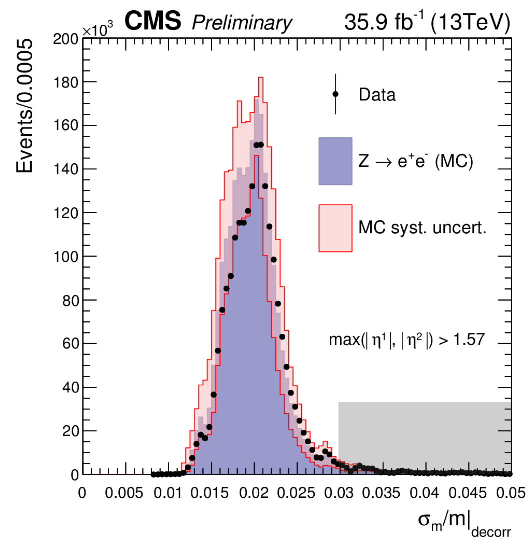
Signal, Background, data all categories summed



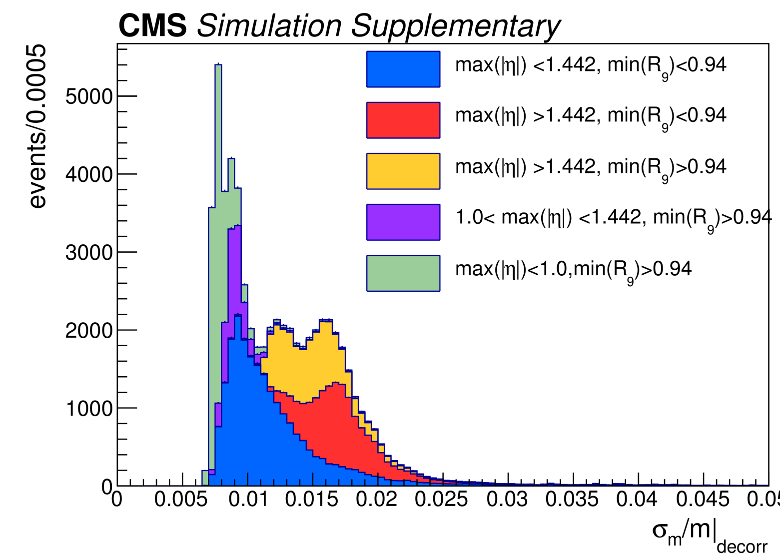
Differential : uncorrelated mass resolution



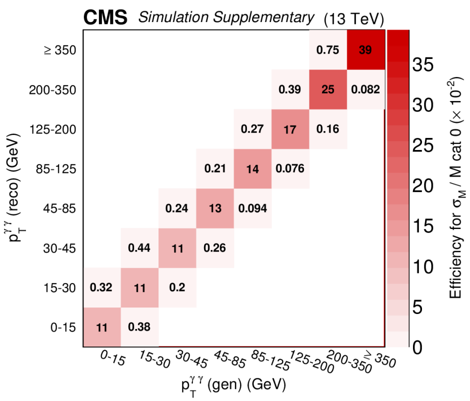
(a)

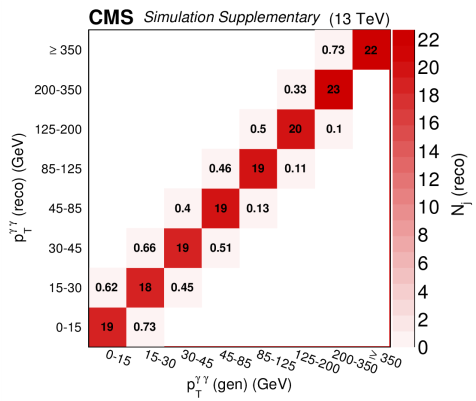


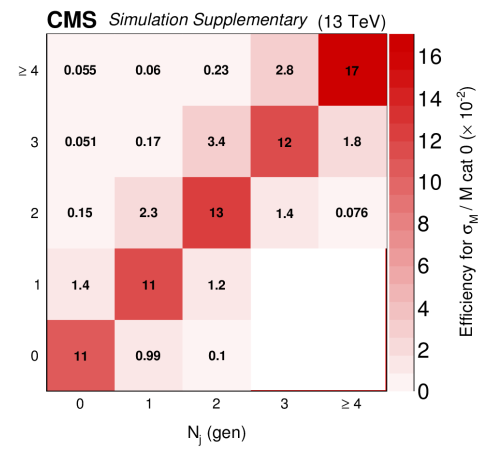
(b)

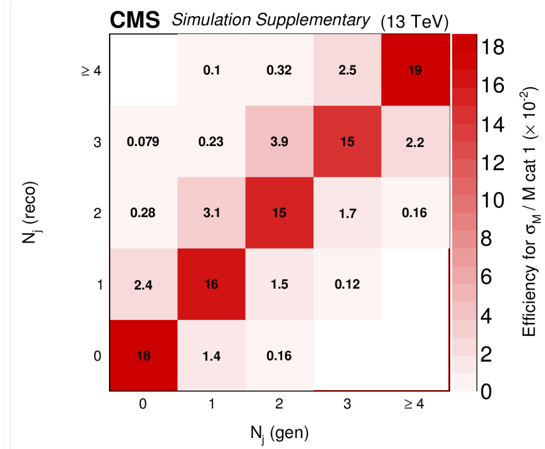


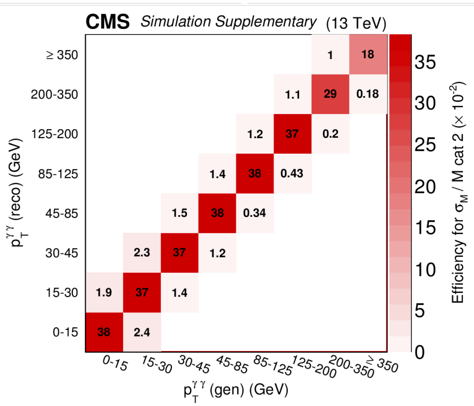
Differential : Response matrices

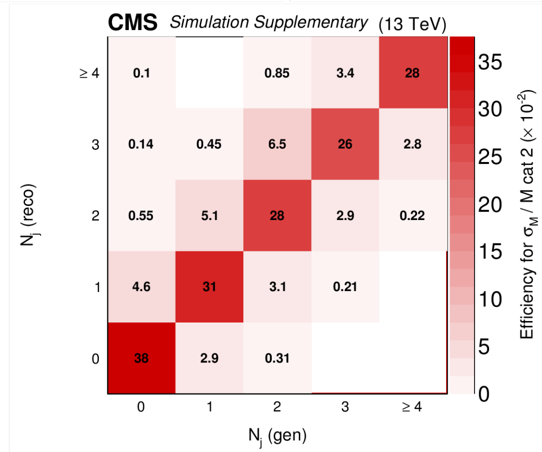












Differential : Unfolding

$$\mathcal{L}(\text{data}|\vec{\Delta\sigma}, \vec{n_{bkg}}, \vec{\theta_S}, \vec{\theta_B}) =$$

$$\prod_{i=1}^{n_{cat}} \prod_{j=1}^{n_b} \prod_{l=1}^{n_{m_{\gamma\gamma}}} \left(\sum_{k=1}^{n_b} L\Delta\sigma_k K_k^{ij}(\vec{\theta_S}) S_k^{ij}(m_{\gamma\gamma}^l|\vec{\theta_S}) + n_{OOA}^{ij} S_{OOA}^{ij}(m_{\gamma\gamma}^l|\vec{\theta_S}) + n_{bkg}^{ij} B^{ij}(m_{\gamma\gamma}^l|\vec{\theta_B}) \right)^{n_{ev}^{lij}} \cdot$$

$$\text{Pois}(n_{ev}^{ij}|n_{sig}^{ij} + n_{bkg}^{ij}) \cdot Pdf(\vec{\theta_S}) Pdf(\vec{\theta_B})$$

- L is the total integrated luminosity analyzed;
- $\vec{\Delta\sigma} = (\Delta\sigma_1 ... \Delta\sigma_{n_b})$ is the vector of fiducial cross sections being measured;
- K_k^{ij} are the response matrices, which represent the probability that an event in the k-th kinematic bin at generator level is reconstructed in the ij-th reconstruction-level category (with the index i running on the $\sigma_\mathrm{m}/\mathrm{m}|_\mathrm{decorr}$ categories and the index j running on the kinematic bin).
- the functions S_k^{ij} and B^{ij} are the signal and background distributions in $m_{\gamma\gamma}$, which will be described in the next section;
- n_{ev}^{ij}, n_{sig}^{ij}, n_{bkg}^{ij} are the number of observed, signal and background events in the ij-th reconstruction-level category respectively;
- the terms $n_{OOA}^{ij} S_{OOA}^{ij}$ represent the contributions of the Higgs boson signal originating outside of the fiducial region. The contribution of the out of acceptance Higgs boson signal is estimated from simulations to be roughly 1% of the total expected SM signal;
- the parameters $\vec{\theta_S}$ and $\vec{\theta_B}$ are the nuisance parameters associated with the signal and background model respectively.