Measurement of mass and properties of the SM Higgs boson decaying to two photons using p-p collision data collected by CMS in 2016

Kuntal Mondal

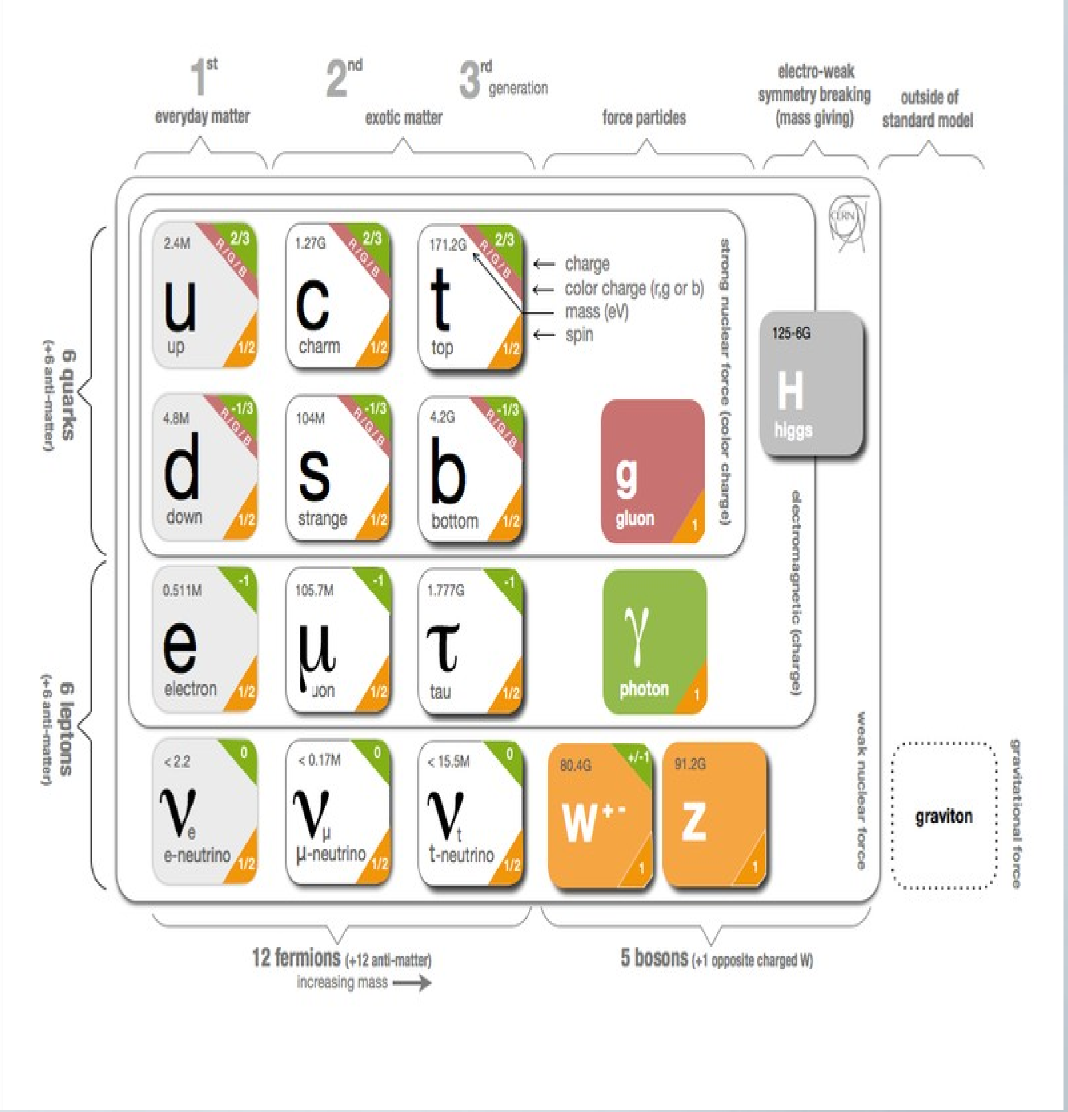
Lunch seminar at CP3, UC Louvain, Belgium on 23-01-2020

Outline

- Theoretical motivation
  - Higgs boson and diphoton decay channel
- Description of CMS detector and trigger system
- Object reconstruction algorithm in CMS
- Higgs to diphoton analysis
- Results
- List of publication

Motivation: Why Higgs boson?

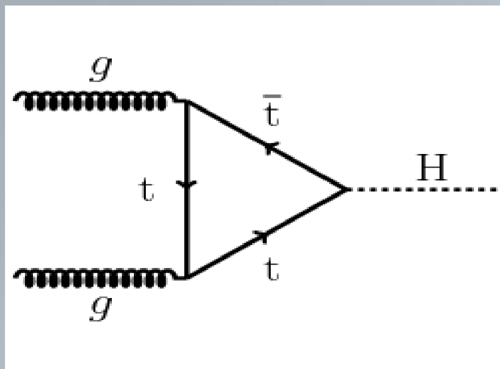
- **Standard Model** is the most tested model of elementary particles
- Almost all the particles predicted by SM were discovered during the last century except one, the **Higgs boson**
- In the year 2012, ATLAS and CMS experiments jointly announced discovery of a new particle of mass 125.09 GeV
  - Resembles the properties of Higgs boson as predicted by SM
- The particle was discovered mainly in diphoton and four lepton decay channels
- Following the discovery, during LHC Run2 (2015 – 2018) period, the emphasis is on
  - Measuring the properties and couplings of the Higgs boson to other particles



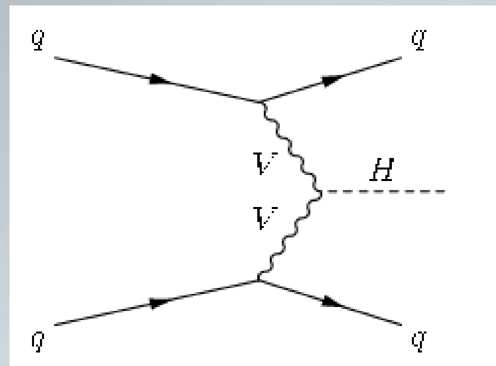
Picture courtesy: CERN

Higgs production mechanisms and decay channels

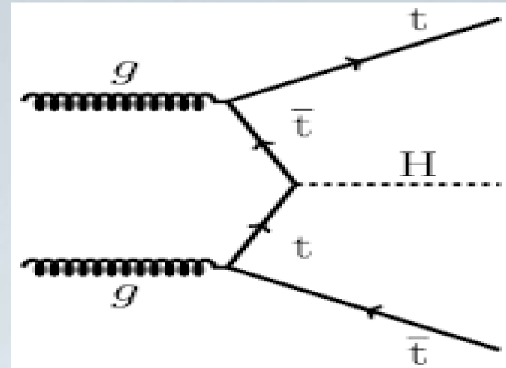
Higgs production mechanisms:



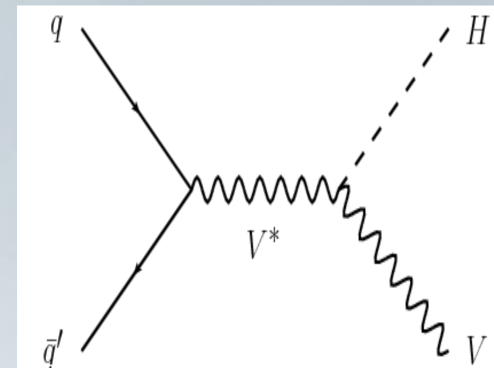
gluon gluon fusion (ggH)



Vector boson fusion (VBF)

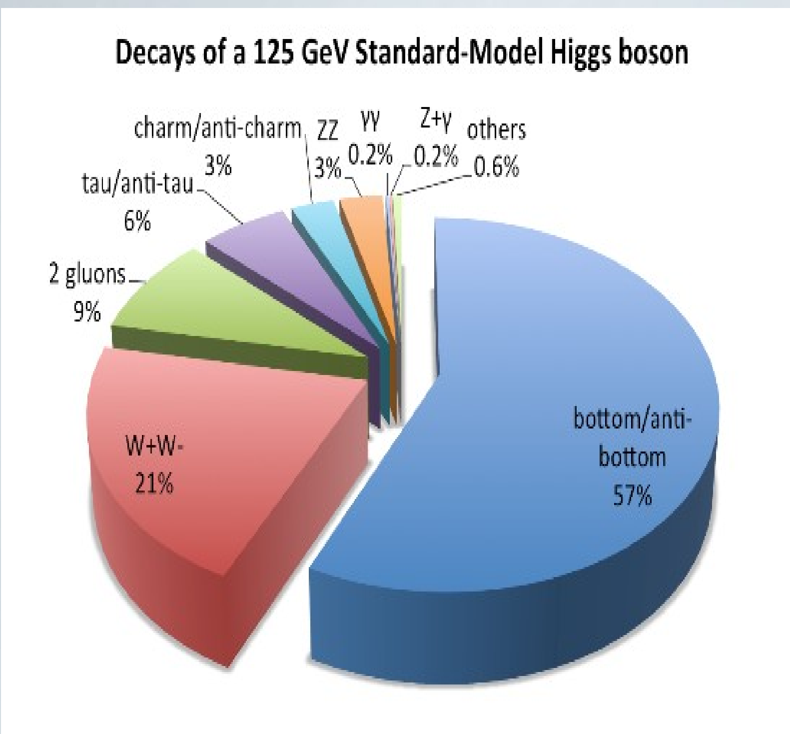


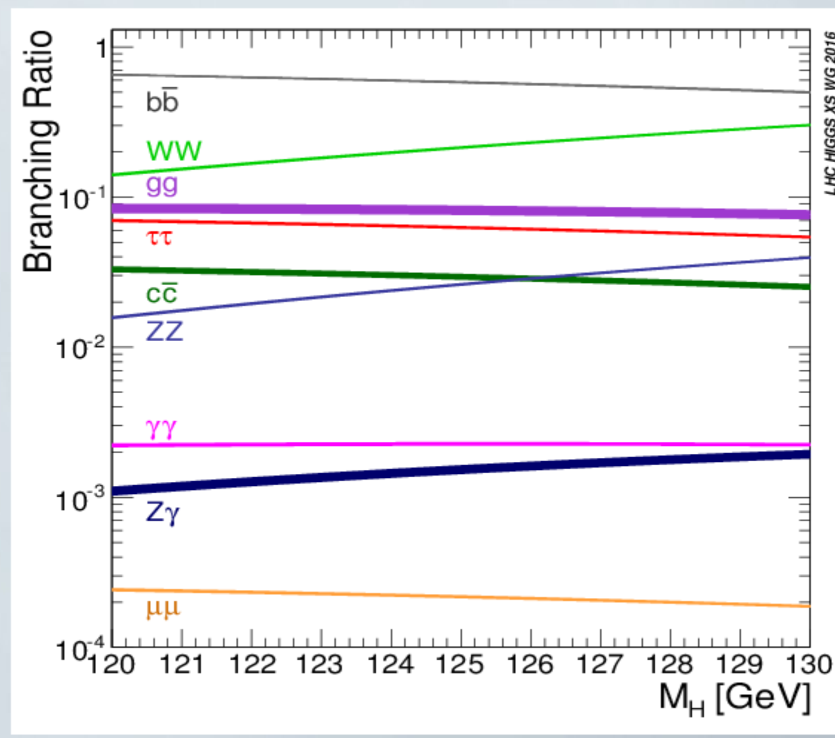
Associate production with top (ttH)



Associate production with vector boson (VH)

Higgs decay channels:



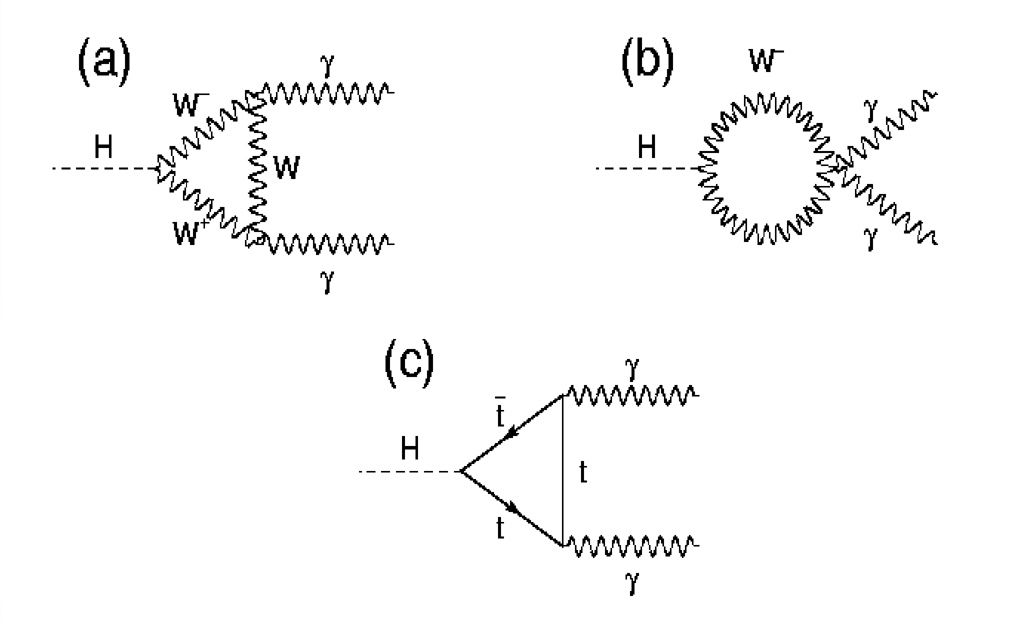


Picture courtesy: CERN

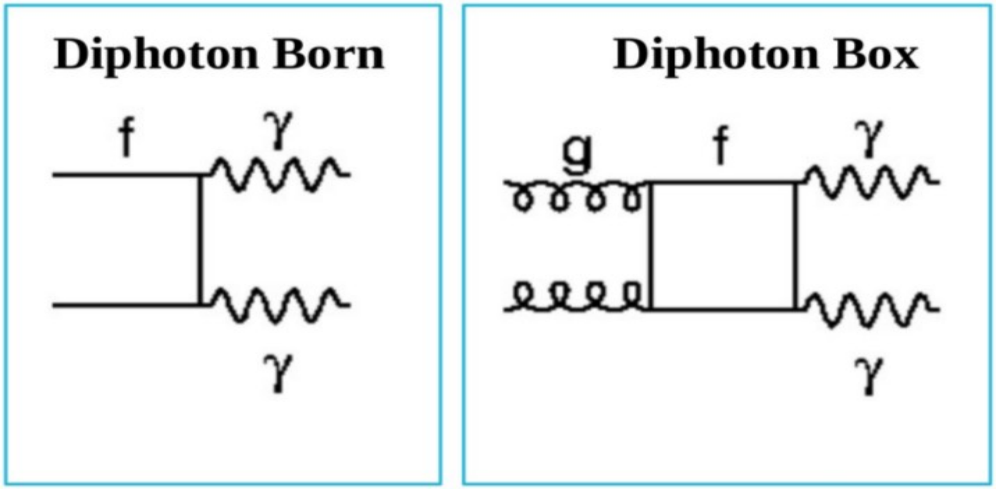
Diphoton decay channel of Higgs

- **Very small branching ratio (~ 0.2 %)**
- **Signture:**
  - **Clean event topology** with two highly energetic photons
  - Narrow width of Higgs boson and high resolution in the reconstruction of diphoton mass gives a clean signal over continuously falling combinatorial background
  - Final state can be reconstructed with high resolution

- **H -> γγ signal:**



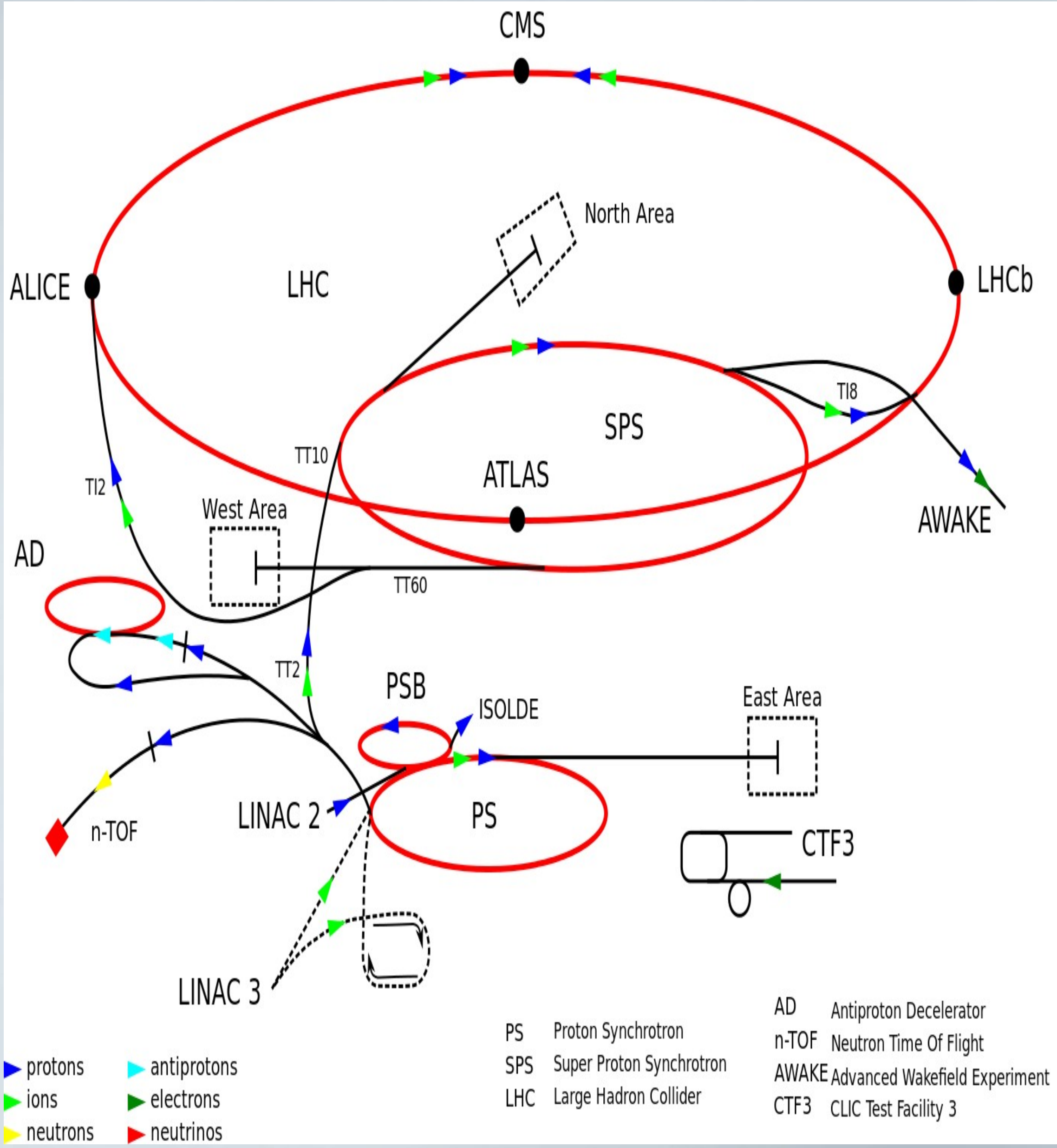
- **Backgrounds:**
  - **Irreducible:** pp -> γ + γ
    - Almost 70% comes from Box Born process



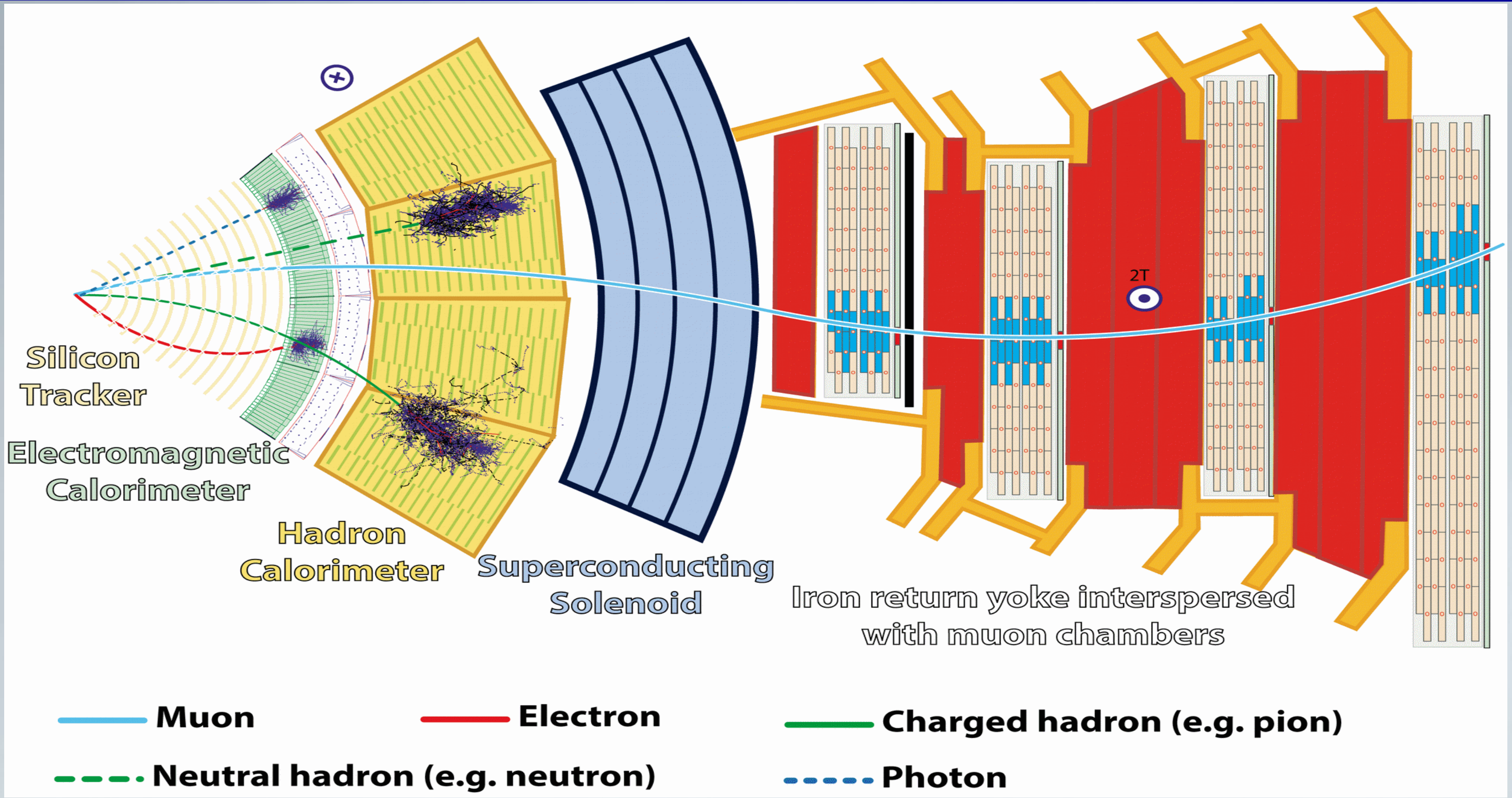
  - **Reducible:** pp -> γ + jet, jet + jet
    - Jets can be misidentified as photons

The Large Hadron Collider

- World's largest and most powerful particle accelerator (27-kilometer ring)
- The beams (p - p, p - Pb, Pb - Pb) travel in opposite directions in separate beam pipes at ultra-high vacuum
- The beams are made to collide at four locations around the accelerator ring, corresponding to the positions of the four detectors
  - ✔ **A** **T**oroidal **L**HC **A**paratu**S** **(ATLAS)**
  - ✔ **C**ompact **M**uon **S**olenoid **(CMS)**
  - ✔ **A** **L**arge **I**on **C**ollider **E**xperiment **(ALICE)**
  - ✔ **LHCb**



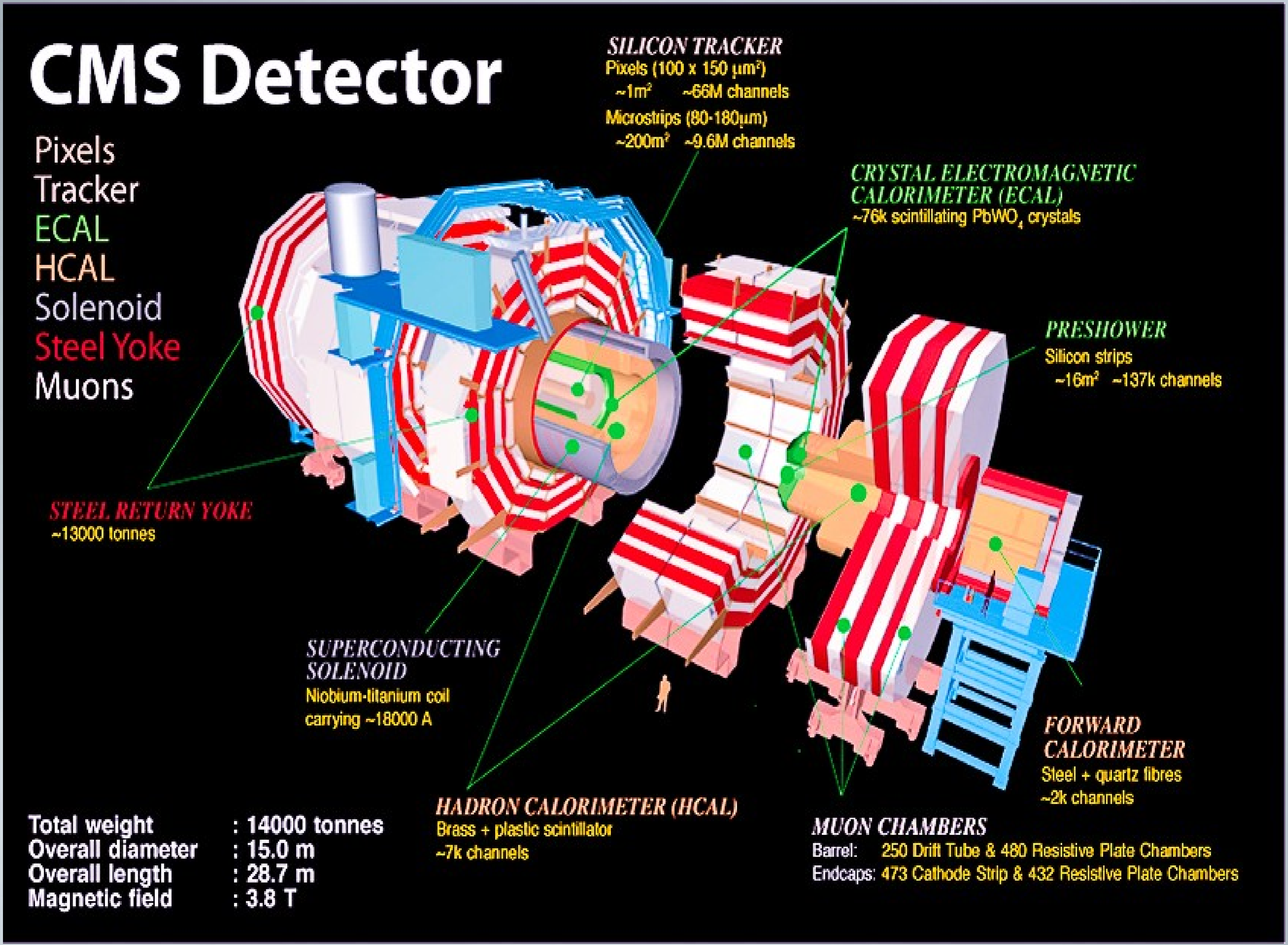
Modern high energy particle detectors



Picture courtesy: CMS, CERN

- **Tracker**
  - Measures momentum of charged particle
- **Electromagnetic Calorimeter**
  - Measures energy of electrons and photons
- **Hadron Calorimeter**
  - Together with ECAL measures energy of hadrons
- **Muon Chamber**
  - Measures momentum of muons

Compact Muon Solenoid detector

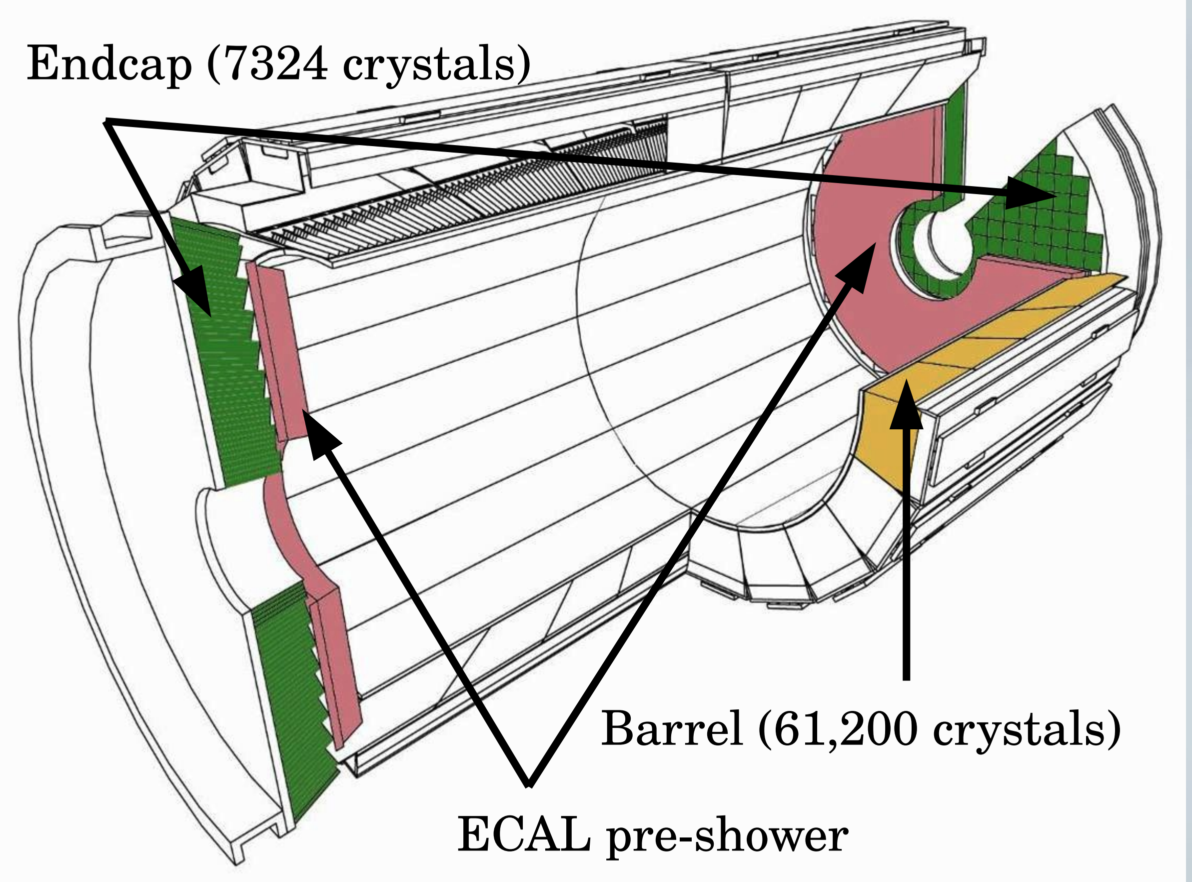


Picture courtesy: CMS, CERN

Electromagnetic Calorimeter (ECAL)

ECAL is the main sub-detector for photon detection

- Made up of lead-tungstate ($PbWO_4$) crystals
- For extra spatial precision, the ECAL also contains "Preshower" detectors that resides infront of the each endcap
- Coverage for ECAL:
  - **Barrel (EB):** $|\eta| < 1.4442$
  - **Endcap (EE):** $1.566 < |\eta| < 2.5$



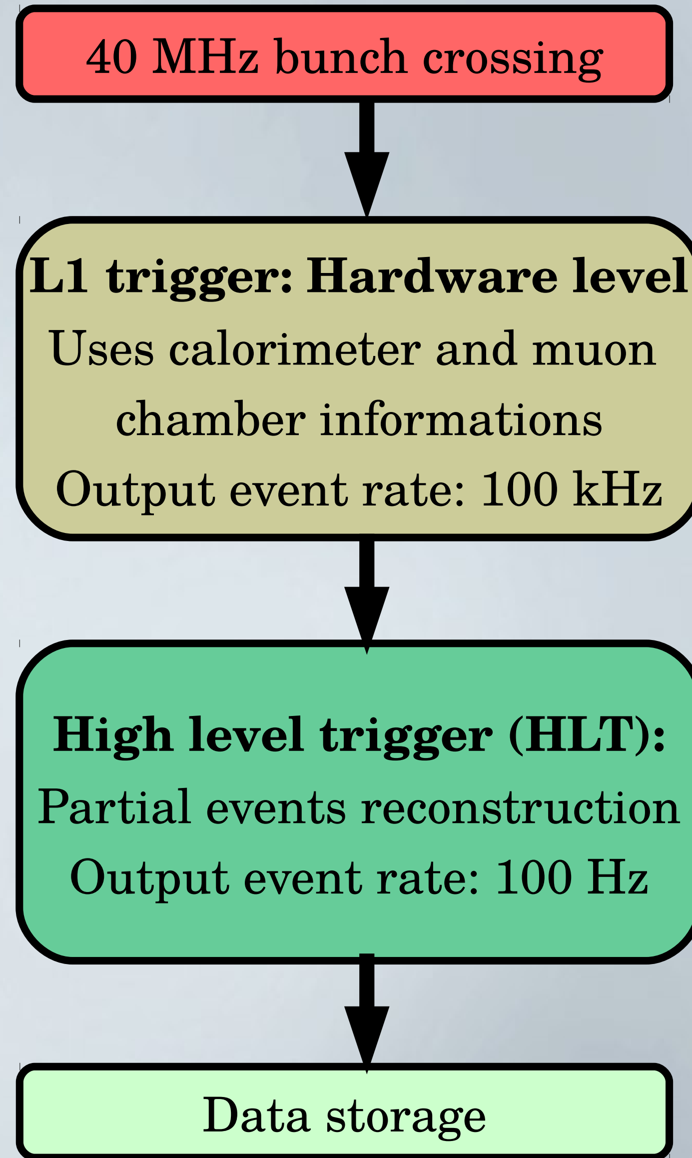
Picture courtesy: CMS, CERN

- Designed to measure energy of "Electromagnetic objects" (electrons and photons) which "**shower**" (deposit) their full energy in the ECAL crystals
- Excellent energy resolution : ~ 0.5 % for a 100 GeV photon

Trigger System

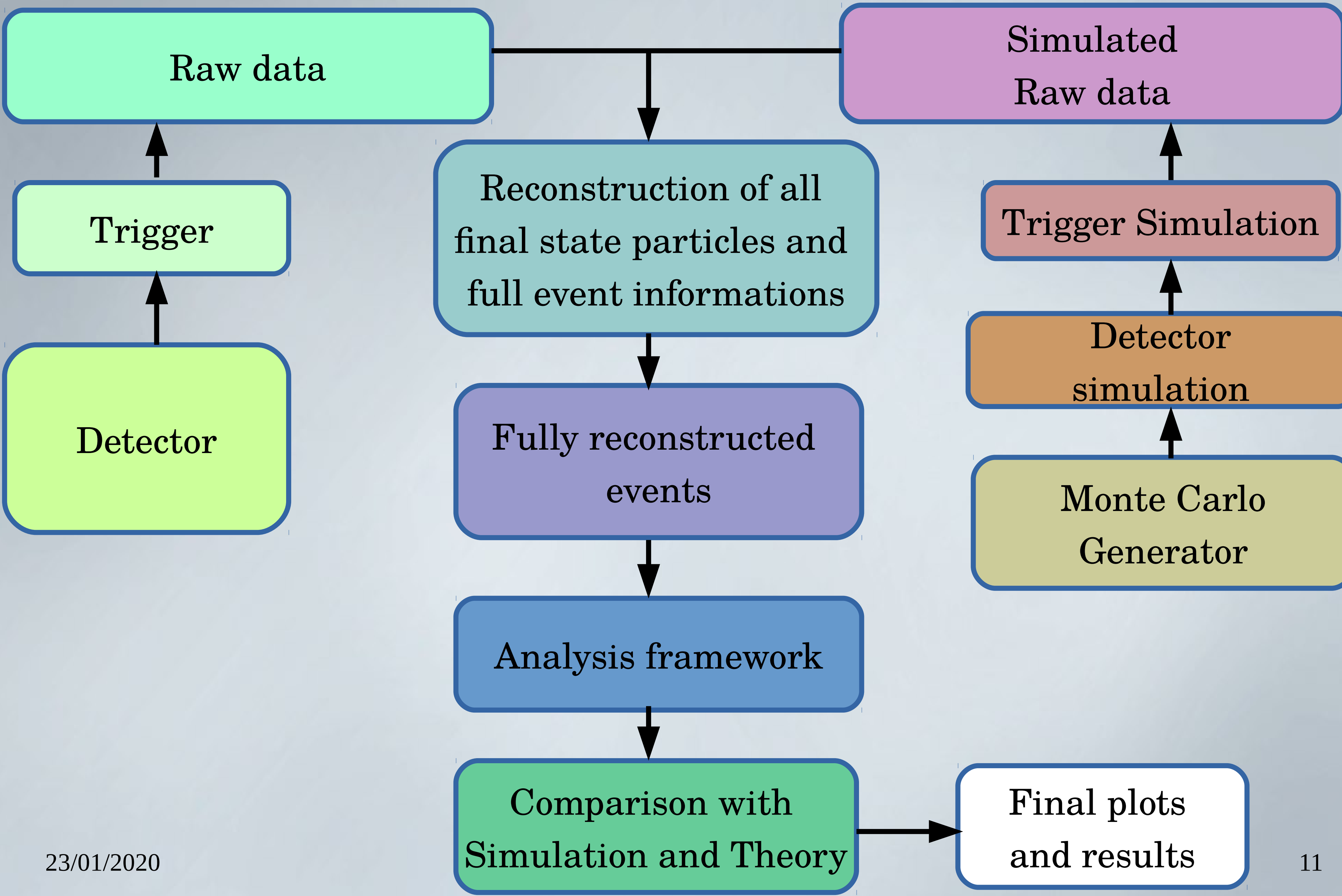
Trigger system allows us to record only interesting events for further analysis

- At designed luminosity of 10^{34} cm^{-2} s^{-1} at LHC with a bunch crossing space of 25 ns p – p collision rate is **40 MHz**
- Move the above box here
- CMS uses a two level trigger system for recording events with interesting physics with great efficiency
  - **Level1 trigger (L1):** Hardware level trigger
    - Reduces event rate to **100 kHz**
  - **High Level Trigger (HLT):** Uses partially reconstructed object information
    - Reduces event rate to **~100 Hz**



Change colour

Analysis flow in High Energy Physics



Object Reconstruction

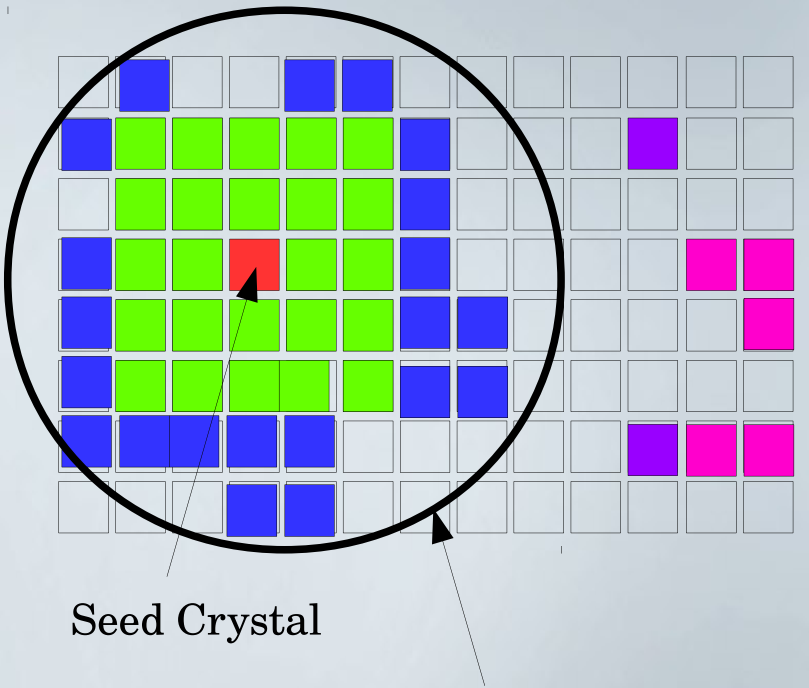
In CMS, all final state particles from p-p collision events are reconstructed using “Particle Flow” (PF) algorithm

- ✔ **Muons** : hits in muon chambers matched with tracks in the silicon tracker
- ✔ **Electrons** : energy showers in the ECAL matched with tracks in the silicon tracker
- ✔ **Photons** : energy showers in the ECAL, but no corresponding track in silicon tracker
- ✔ **Charged hadrons** : energy tower in ECAL and HCAL matched with tracks in the silicon tracker
- ✔ **Neutral hadrons** : energy towers in the ECAL and HCAL not associated to any track in the tracker
- ✔ **Jets**, the experimental analogue of partons, are reconstructed by clustering hadrons within a certain cone originating from same vertex

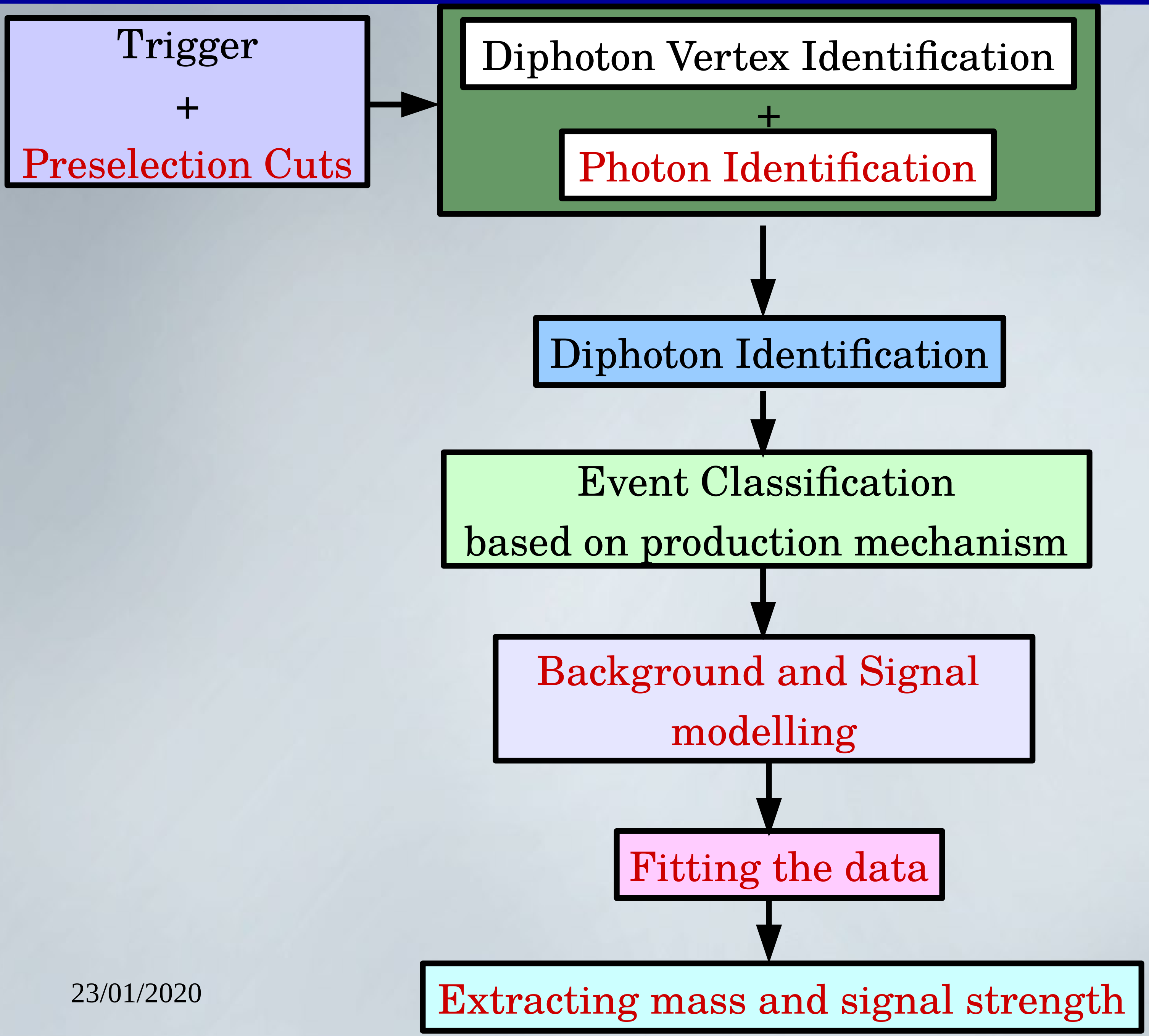
Photon Reconstruction

Shower in Eletromagnetic Calorimeter → No associated charged particle track → The shower may correspond to a photon object

- Photons are reconstructed in three steps:
  - ➢ Local maxima in energy are found among the ECAL crystals above a threshold (called "Seed Crystal")
  - ➢ A cluster topology is built around the seeded crystal
    - ◆ Adding crystals which have
      - ✔ At least one immediate neighbour clustered crystal
      - ✔ Energy above a threshold
    - ◆ The aggregated clusters are called a **"Supercluster"**
  - ➢ Estimation of photon energy is made from the sum of calibrated energy for the super cluster in the ECAL



Analysis Flow



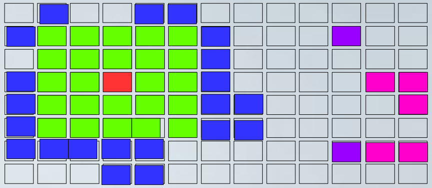
Trigger and Preselection cuts

- **HLT**: HLT_Diphoton30_18_R9Id_OR_IsoCaloId_AND_HE_R9Id_Mass90
  - Based on two L1 photon candidates with transverse momentum (pT) > 22 and 10 GeV
  - Other selection cuts applied on:
    - Transverse energy of the photons, diphoton invariant mass, shower shape and isolation variables of the photons, hadronic energy over electromagnetic energy
- Diphotons from the events passing the HLT are required to pass a set of offline cuts called “preselection cuts”

Coverage in ECAL	EB ($\|\eta\| > 1.4442$), EE ($1.566 < \|\eta\| < 2.5$)
Transverse momentum (pT)	Lead (Sublead) photon > 30 (20) GeV
Invariant mass of diphoton pair ($m_{\gamma\gamma}$)	> 90 GeV
Cuts on shower shape and isolation variables	
Cut on hadronic energy over electromagnetic energy of photons (H/E)	
Electron veto: no track should point to the shower of the photon	

Scale factors for Preselection Cuts

- Scale factor is defined as the ratio of the efficiency of the preselection cuts for data over simulation
  - Need to scale simulated events before comparing with data
- Scale factors for preselection cuts are measured in two different steps
  - Scale factors are measured for Z -> ee events by **"Tag n Probe"** method in four different bins of ECAL coverage and R9 (E_{5x5}/E_{sc}) value for all cuts except "Electron veto"
  - For "Electron veto", scale factor is measured separately in Z -> μμγ channel



- All the scale factors measured
  - Scale factors measured from Z -> ee events

	Data		Simulation		Ratio		
	Eff.	Stat	Eff.	Stat.	Scale	Stat. Unc.	Syst. Unc
Barrel; R_9 >0.85	0.9488	0.0001	0.9499	0.0001	0.9988	0.0001	0.0009
Barrel; R_9 <0.85	0.8471	0.0001	0.8423	0.0002	1.0057	0.0002	0.0010
Endcap; R_9 >0.90	0.9207	0.0004	0.9256	0.0002	0.9947	0.0004	0.0051
Endcap; R_9 <0.90	0.5309	0.0001	0.5622	0.0003	0.9443	0.0005	0.0071

  - Scale factors measured from Z -> μμγ events

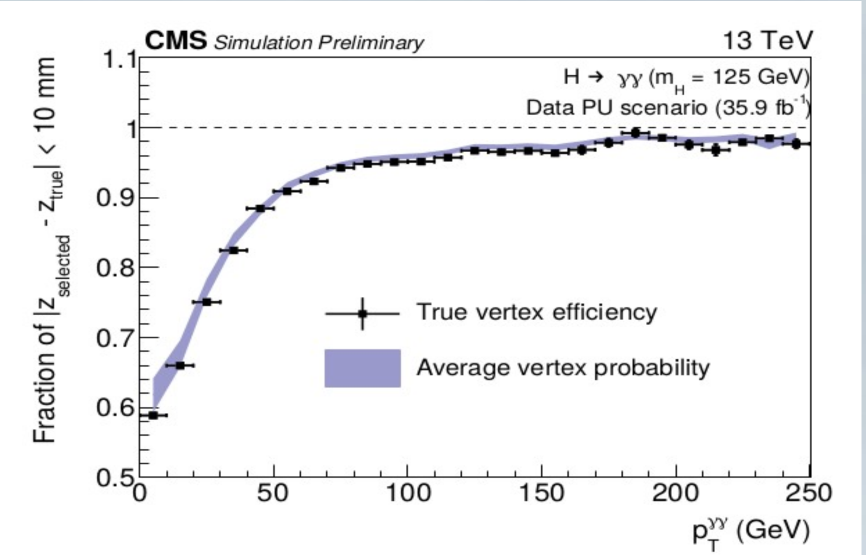
	Data		Simulation		Ratio	
	Eff.	Stat.	Eff.	Stat.	Eff.	Unc.
Barrel; R_9 >0.85	0.9928	0.0003	0.9970	0.0005	0.9958	0.0006
Barrel; R_9 <0.85	0.9741	0.0010	0.9795	0.0020	0.9945	0.0023
Endcap; R_9 >0.90	0.9789	0.0009	0.9863	0.0015	0.9924	0.0017
Endcap; R_9 <0.90	0.9360	0.0033	0.9574	0.0055	0.9777	0.0065

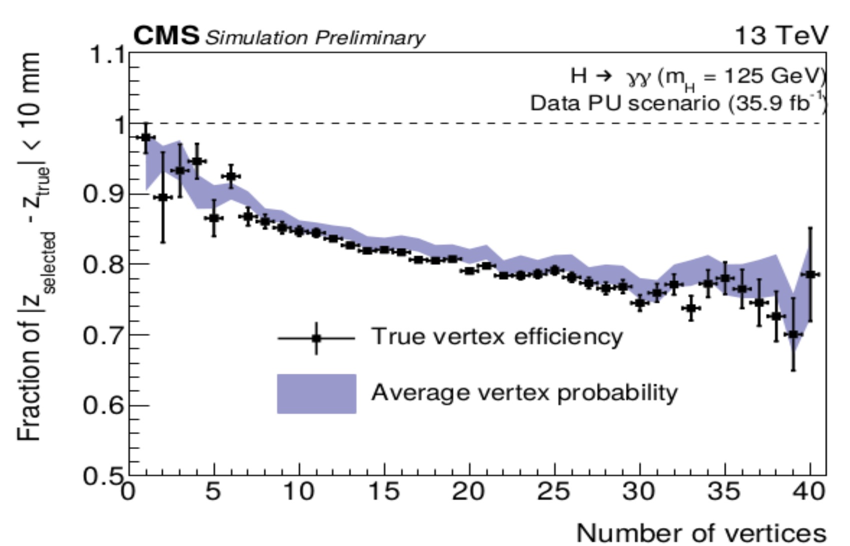
Diphoton Vertex Identification

Diphoton vertex identification is necessary for better mass resolution

- **Multi Variate Analysis (MVA)** is used for vertex identification
  - ➢ Takes as input kinematic variables of photons
  - ➢ Among different available methods **B**oosted **D**ecision **T**ree (**BDT**) is used
    - ✔ Vertex identification method is validated on $Z \rightarrow \mu\mu\gamma$ channel
- A **second MVA** is used to estimate probability of correct vertex assignment

✔ If vertex assignment is **within 1 cm** in z direction of true vertex, negligible effect on mass resolution is observed

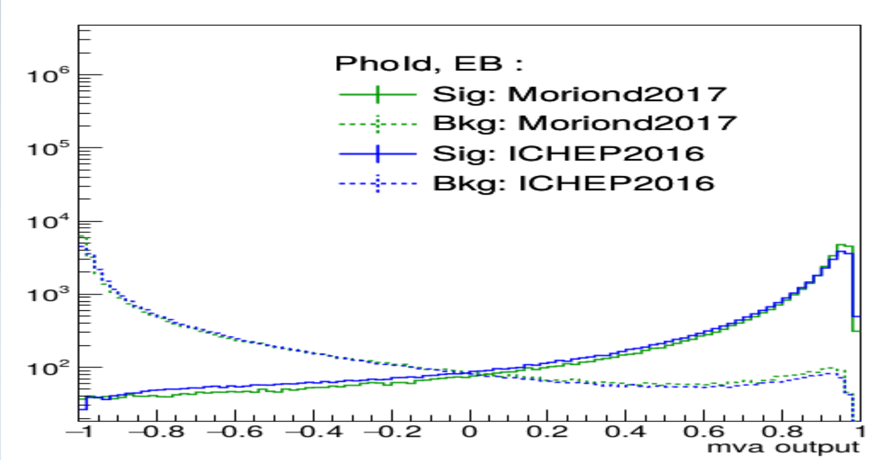


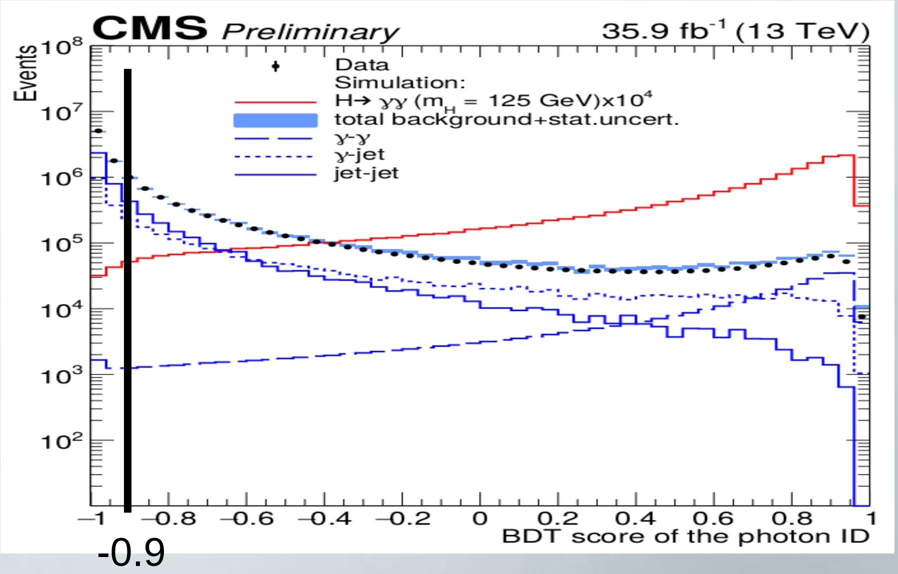


Photon Identification

Two different kind of background components for this analysis considered

- **Irreducible:** diphoton production from QCD events
- **Reducible:** Jets reconstructed as photons (called “fake photons”)
  - Both photon + jet & jet + jet events may fake a H -> γγ event
    - Proper identification of prompt photons crucial
  - To identify photons from jets, a BDT based multivariate discriminator has been used



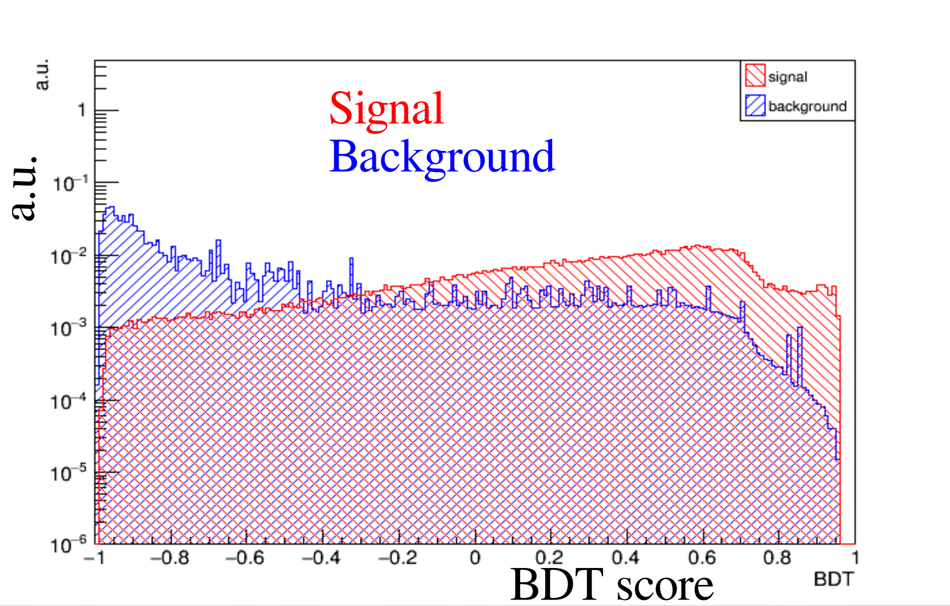


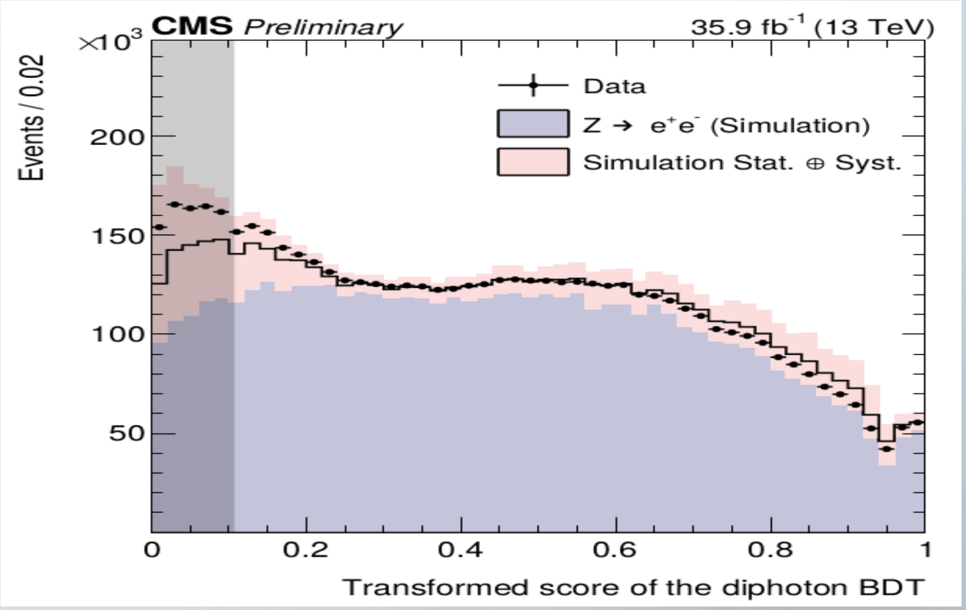
- **Photon ID BDT**: Input variables
  - Shower shape variables
  - Isolation variables
  - Super cluster variables
  - Median energy density per unit area of ECAL per event
  - Generator level information is used to train the BDT to identify prompt photons

Diphoton Selection

Selects signal like diphoton pair from the event after selecting right vertex and all good photons

- The events passing trigger and preselection cuts are further classified using a BDT classifier
- The classifier assigns high BDT score to diphoton pairs if they fulfill
  - Signal like kinematic characteristics
  - Good diphoton mass resolution
  - Photon ID BDT score for both the photons in the "signal-like" region
- The BDT based analysis is validated for Z -> ee events by comparing data and simulation

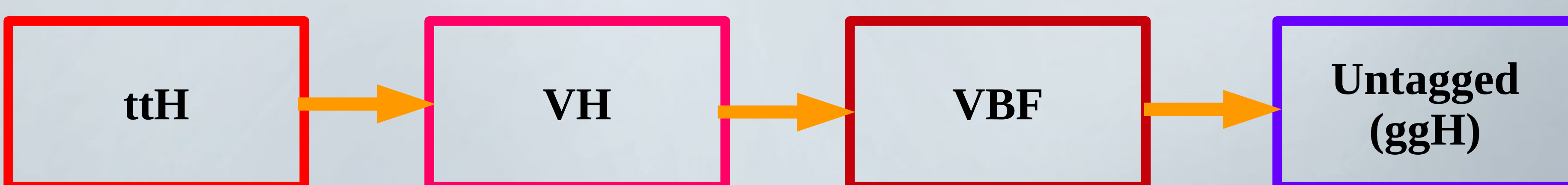




Event Categories

- Events are classified into different categories depending on the production mode of Higgs
  - For measuring different signal strength and coupling of the Higgs boson
  - **Tagged Categories:**
    - Additional selection criteria on extra particles in the final state topology to explore production mode (ttH, VH, VBF)
  - **Untagged Categories:**
    - Mainly ggH events, where splitting is done in terms of different S/B and di-photon invariant mass resolution, optimized to get greater sensitivity

- **Sequence of tagging:**
  - The selected events are required to pass through a tag sequence in the following order for categorization

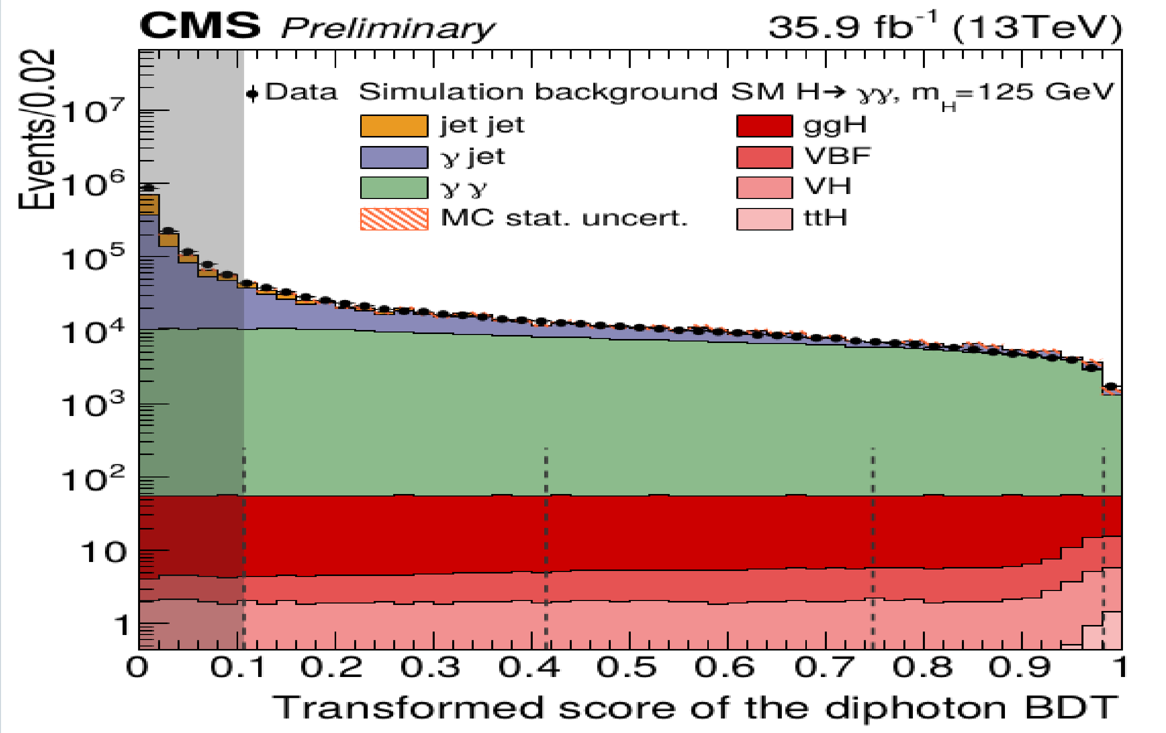


Event Categorization (1)

- Selection for different tags:
  - **ttH Leptonic:**
    - Semi leptonic decay of top
    - At least one lepton required
    - At least two jets. One of the jet has to be b-jet
  - **ttH Hadronic:**
    - Hadronic decay of top
    - No leptons
    - At least three jets
    - At least one of the jets has to be a b-jet
    - Score of the ttH hadronic multivariate discriminant greater than 0.75
  - **VH tag:**
    - **ZHLeptonic**
    - **WHLeptonic**
    - **VHLeptonicLoose**
      - Missing transverse energy < 45 GeV
    - **VHMet**
    - **VHHadronic**

Event Categorization (2)

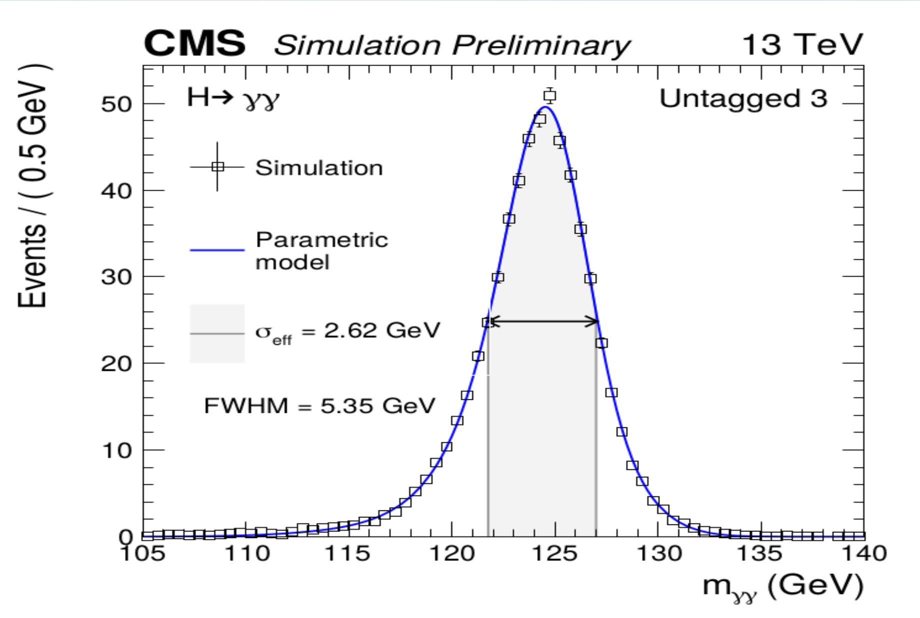
- Selection for different tags:
  - **VBF tag:**
    - Two jets along with two photons
    - Separate MVA based analysis done for the di-jets
    - A final MVA is done taking di-jet MVA and di-photon MVA as input
  - **Untagged (ggH):**
    - The remaining events are from ggH process. They are further classified into sub-categories based on diphoton BDT score

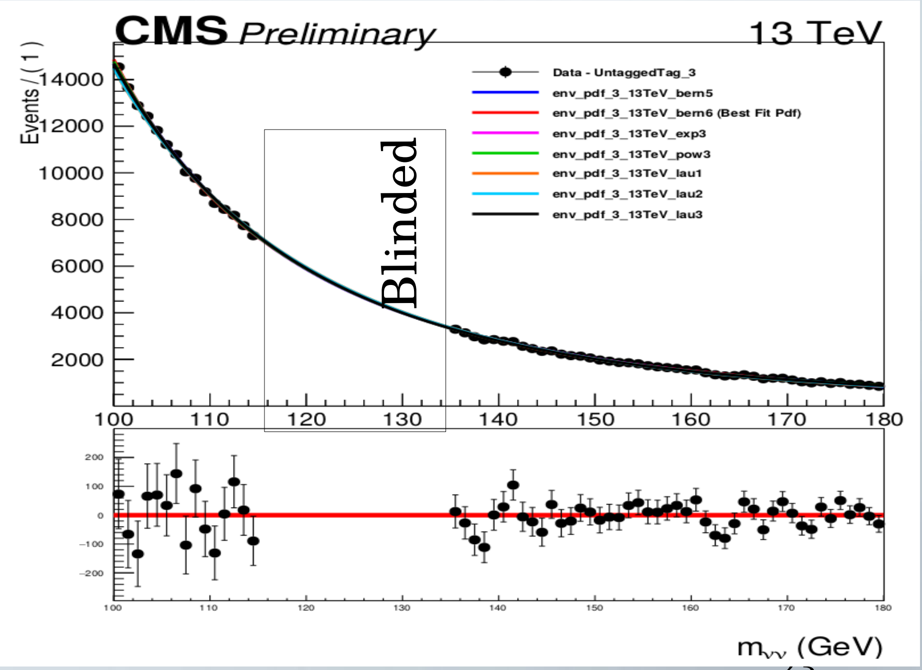


Signal and Background modelling

- **Signal Model:**
  - Should give shape of diphoton mass in each category separately
  - The idea is to describe the signal model by some analytic function
  - Parameters are determined by fitting simulated events for each category using three Higgs mass points (mass)

- **Background Model:**
  - Background model is extracted from data
  - Families of functions considered for
    - ✔ Exponential
    - ✔ Bernstein polynomials
    - ✔ Laurent series
    - ✔ Power law function





Systematic Uncertainties

- All possible source of uncertainties are considered. Some of them are listed below

Theoretical Uncertainties

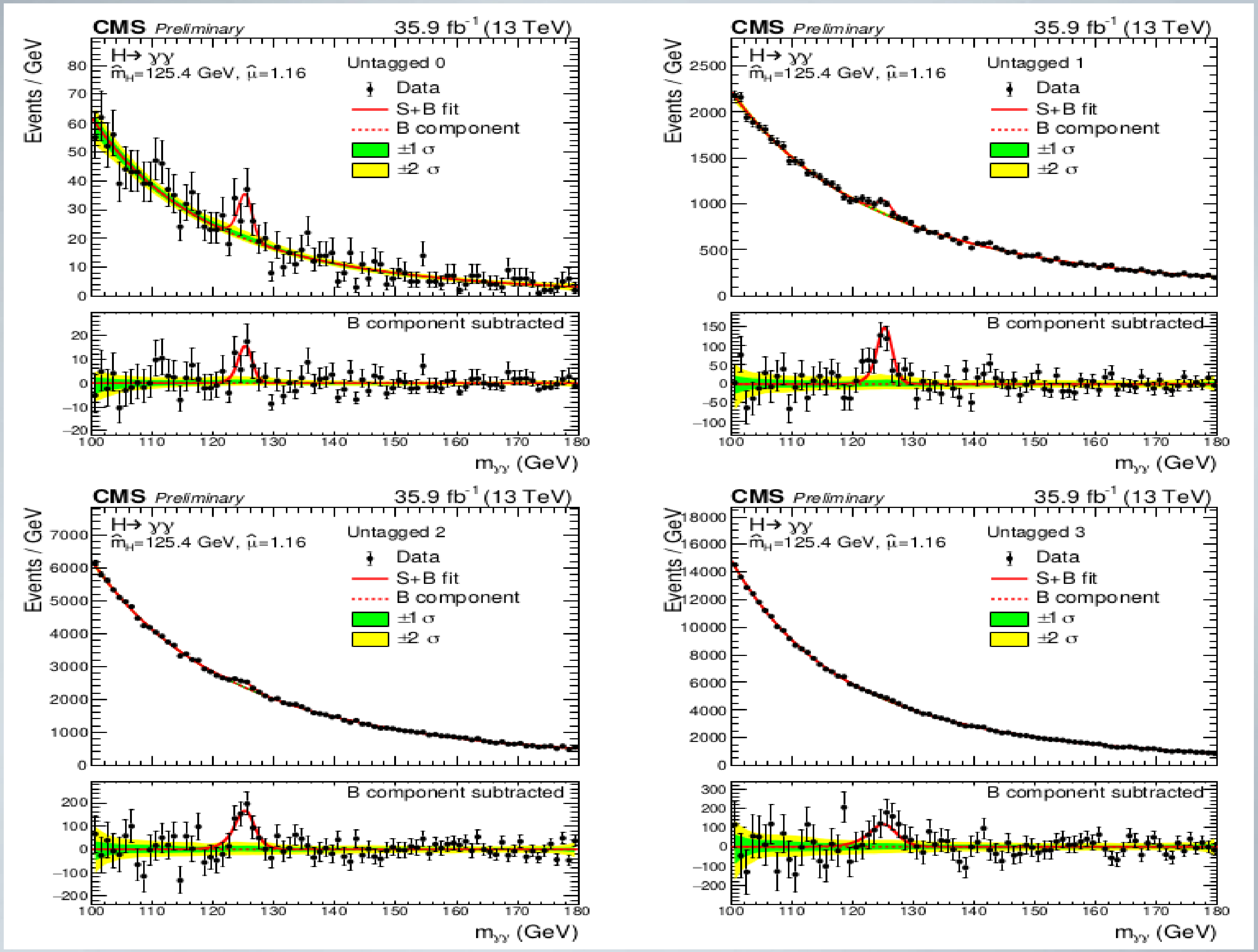
- Parton density function Unc. (2 %)
- Strong coupling constant (α_s) Unc. (3.7 %)
- Parton shower Unc. (7-9 %)
- H->$\gamma\gamma$ branching ratio Unc. (2.08 %)
- Gluon fusion contamination in VBF and ttH (10 %) etc.

Experimental Uncertainties

- Measurement of integrated luminosity (6.2 %)
- Trigger efficiency (0.1 %)
- Photon energy scale and resolution (0.5 %)
- Photon preselection (4 %)
- Vertex finding efficiency (1.5 %)
- Photon ID BDT
- Non uniformity of light collection in ECAL (0.07 %)
- Modelling of material budget infront ECAL (0.17 %)
- Lepton identification (1 %)
- Jet energy scale and smearing (2 %) etc.

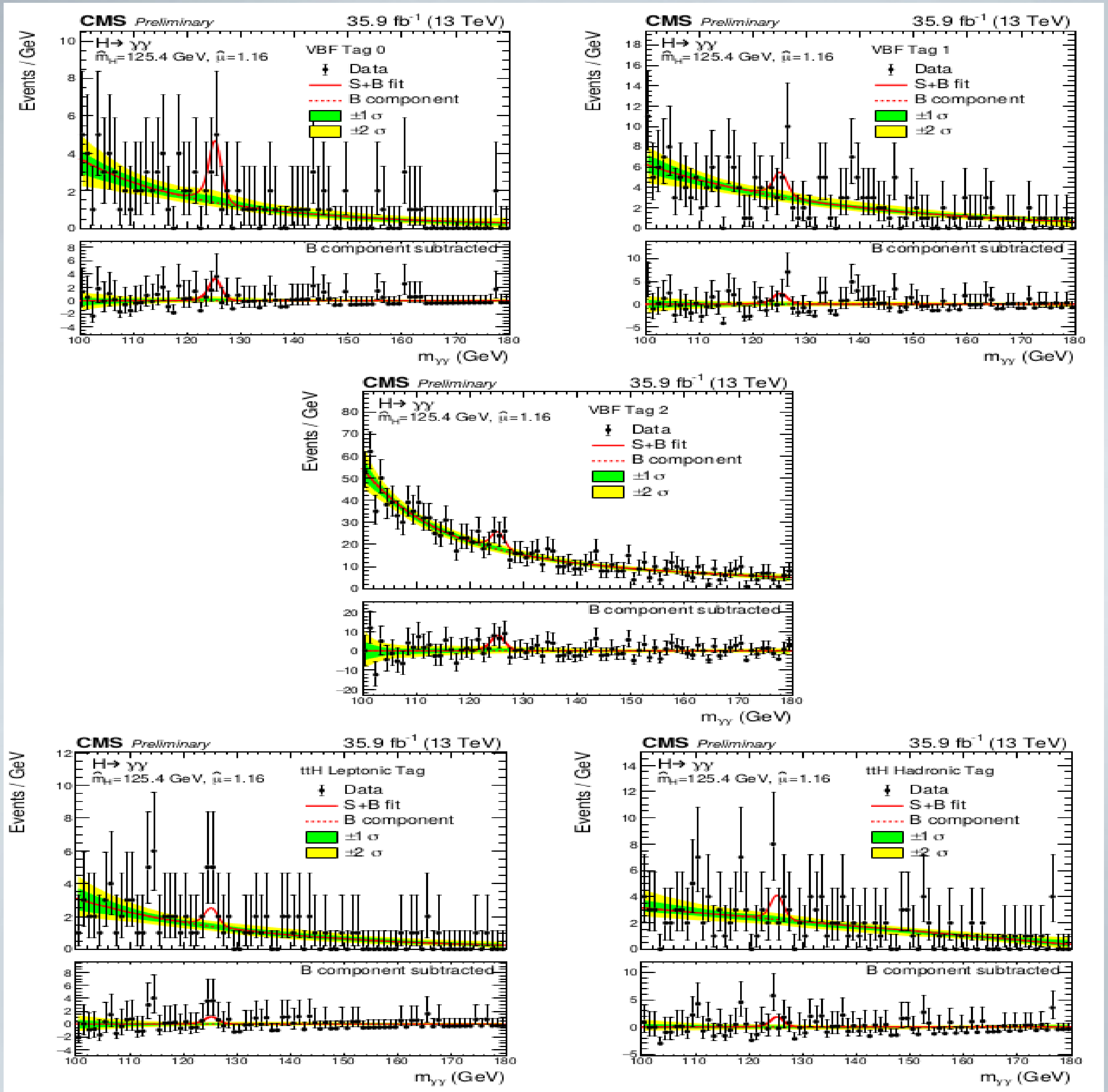
Mass spectra (Untagged categories)

- Mass spectra for all 4 untagged categories (ggH)



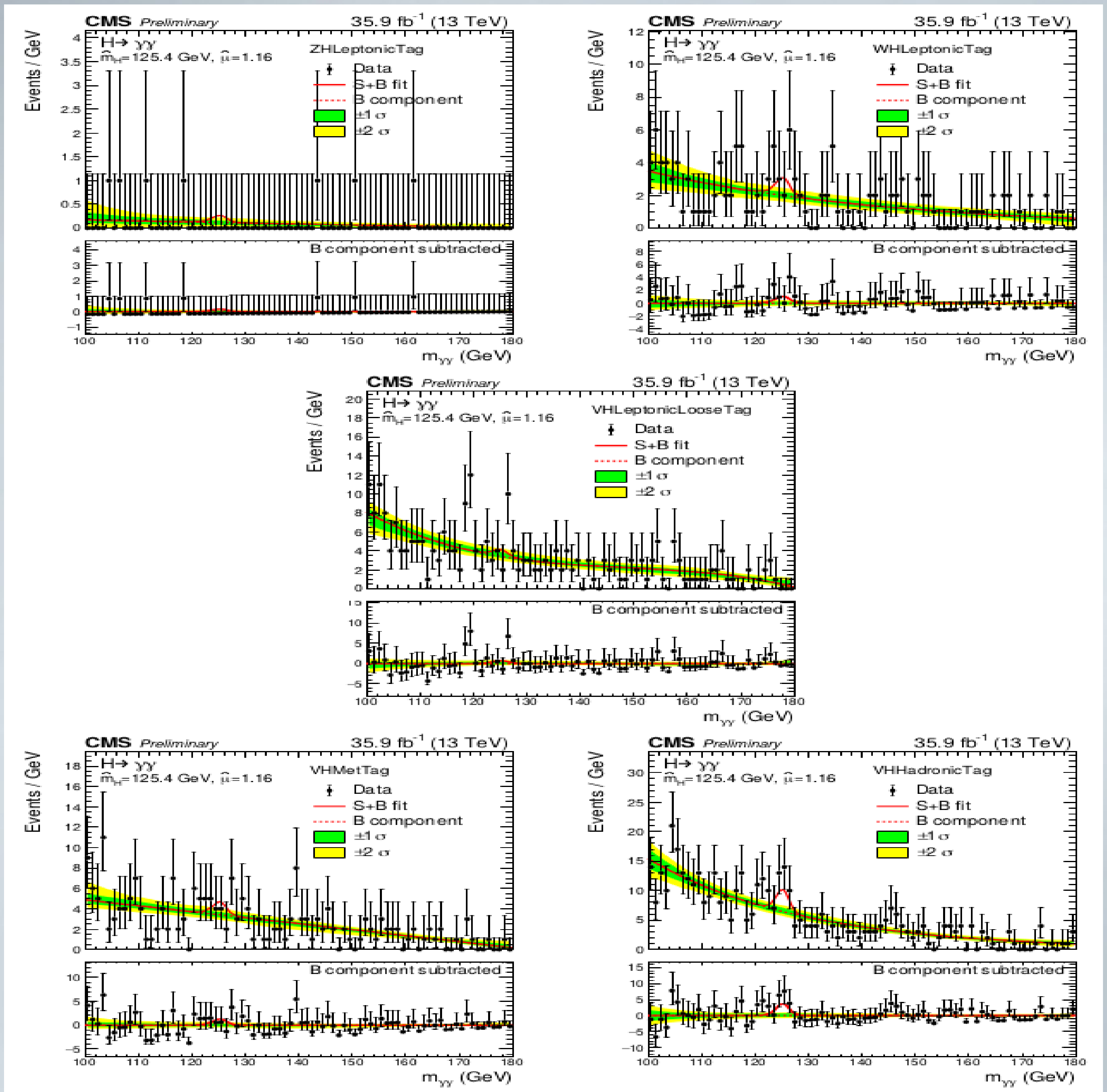
Mass spectra (VBF and ttH categories)

- Mass spectra for all 3 VBF and 2 ttH categories

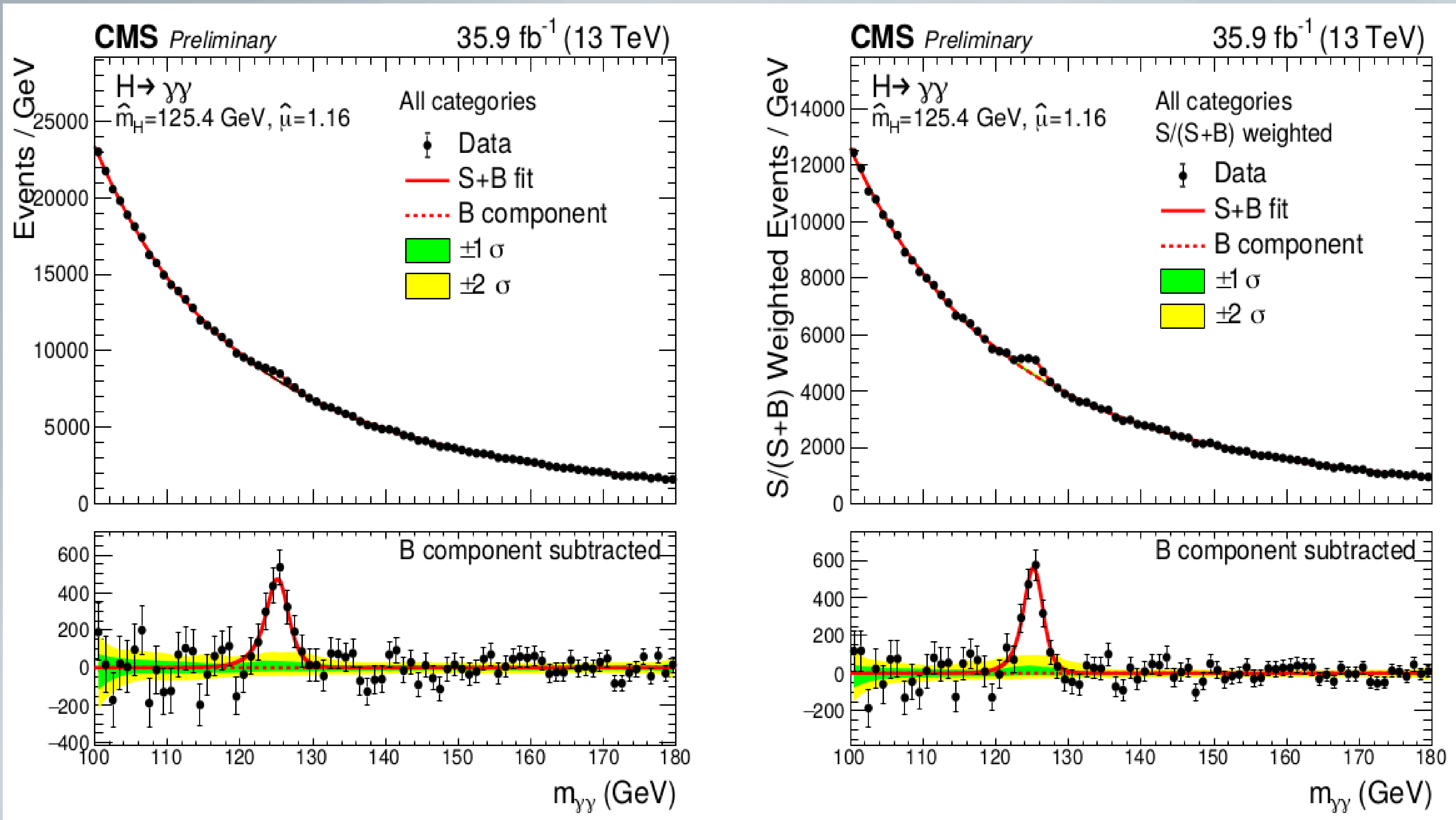


Mass spectra (VH categories)

- Mass spectra for the 5 VH categories



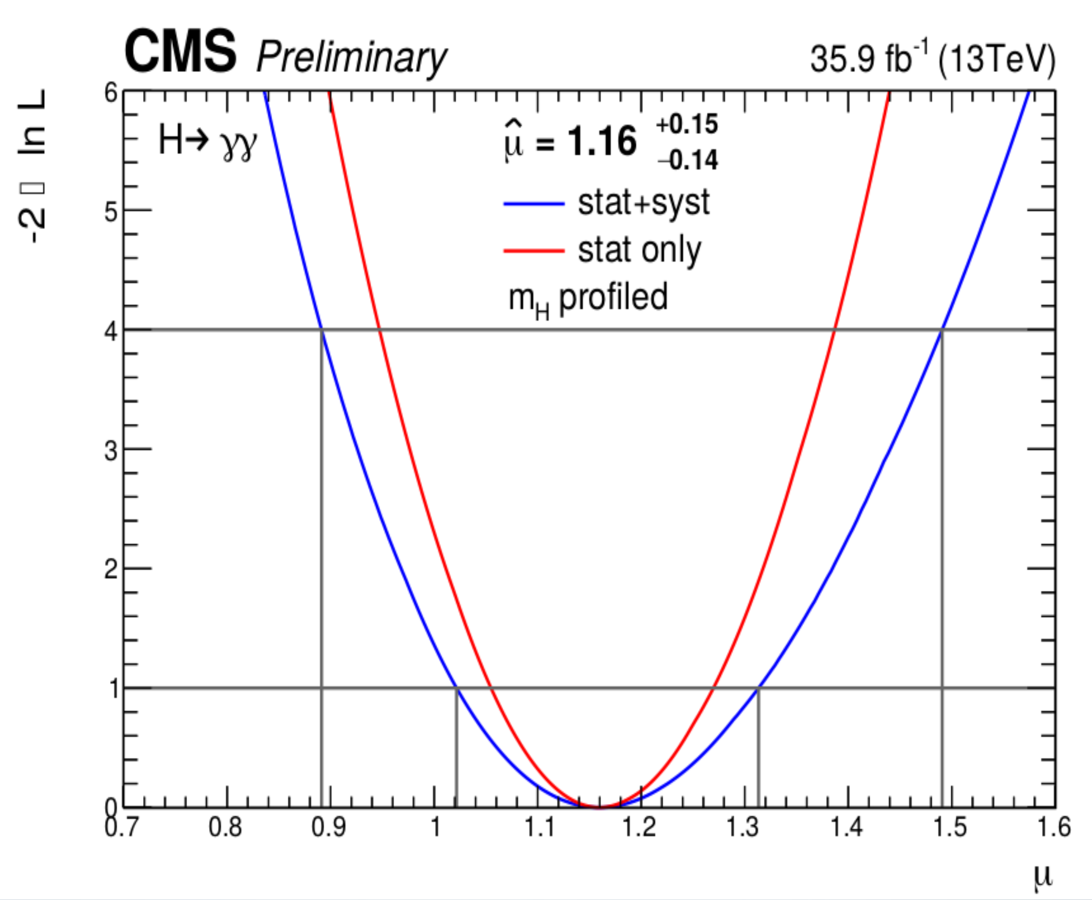
Mass spectra (all categories combined)



Results (1)

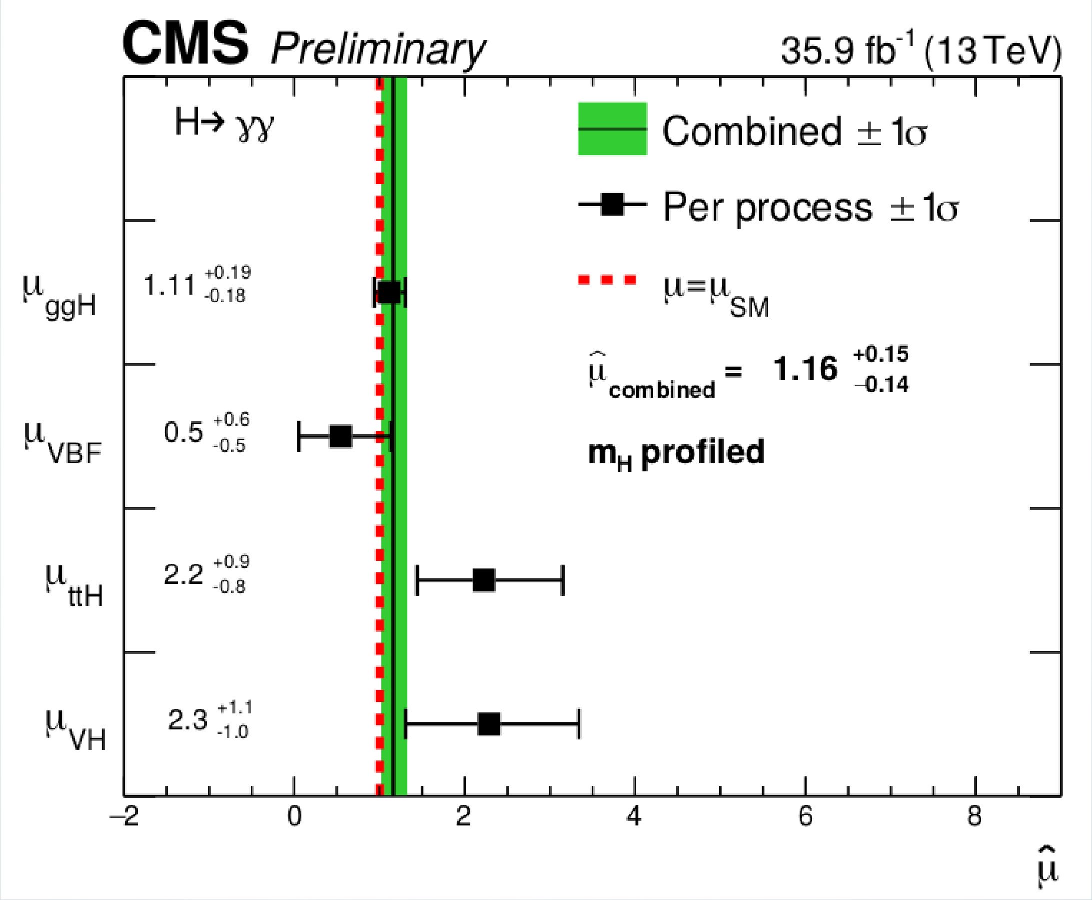
- A likelihood scan of the signal strength is performed, profiling all other nuisances
- The best fit signal strength measured for all categories combined using this method is

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



Results (2)

- Signal strength modifiers measured for each process (black points) for profiled Higgs mass compared to the overall signal strength (green band) and to the SM expectation (dashed red line)

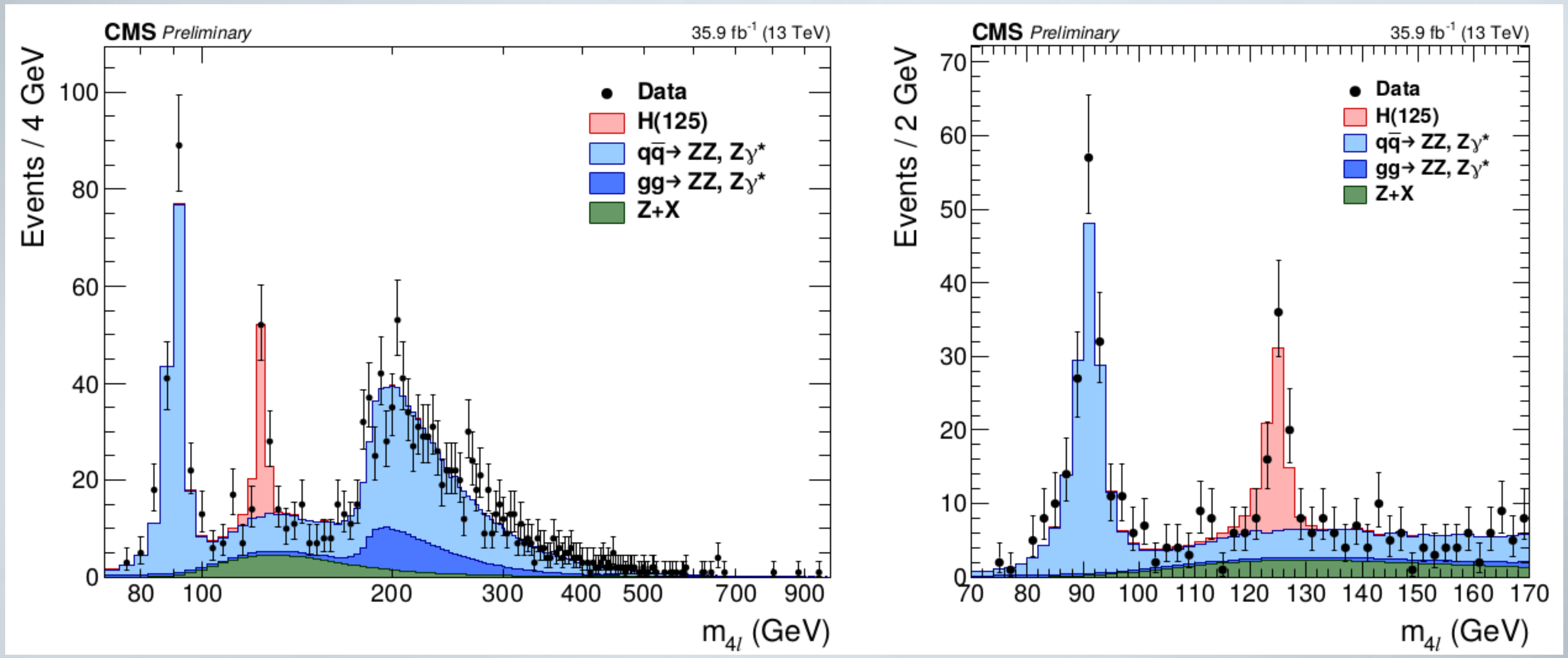


Results (Summary)

- **Best-fit signal strength**
  - $\hat{\mu} = 1.16^{+0.15}_{-0.14} = 1.16^{+0.11}_{-0.10}\,(\text{stat.})\,^{+0.09}_{-0.08}\,(\text{syst.})\,^{+0.06}_{-0.05}\,(\text{theo.})$
- **Signal strength measured in bosonic and fermionic components**
  - $\mu_{VBF,VH} = 1.01^{+0.57}_{-0.51}$
  - $\mu_{ggH,t\bar{t}H} = 1.19^{+0.20}_{-0.18}$
- **Fiducial cross-section measured:**
  - $\hat{\sigma}_{fiducial} = 84 \pm 11(\text{stat}) \pm 7(\text{syst})\ \text{fb} = 84^{+13}_{-12}(\text{stat+syst})\ \text{fb}$
- **Theoretical prediction** for $m_H = 125.09$ GeV
  - $\sigma_{fid} = 73.8 \pm 3.8\ \text{fb}$

Results from other Higgs analysis (1)

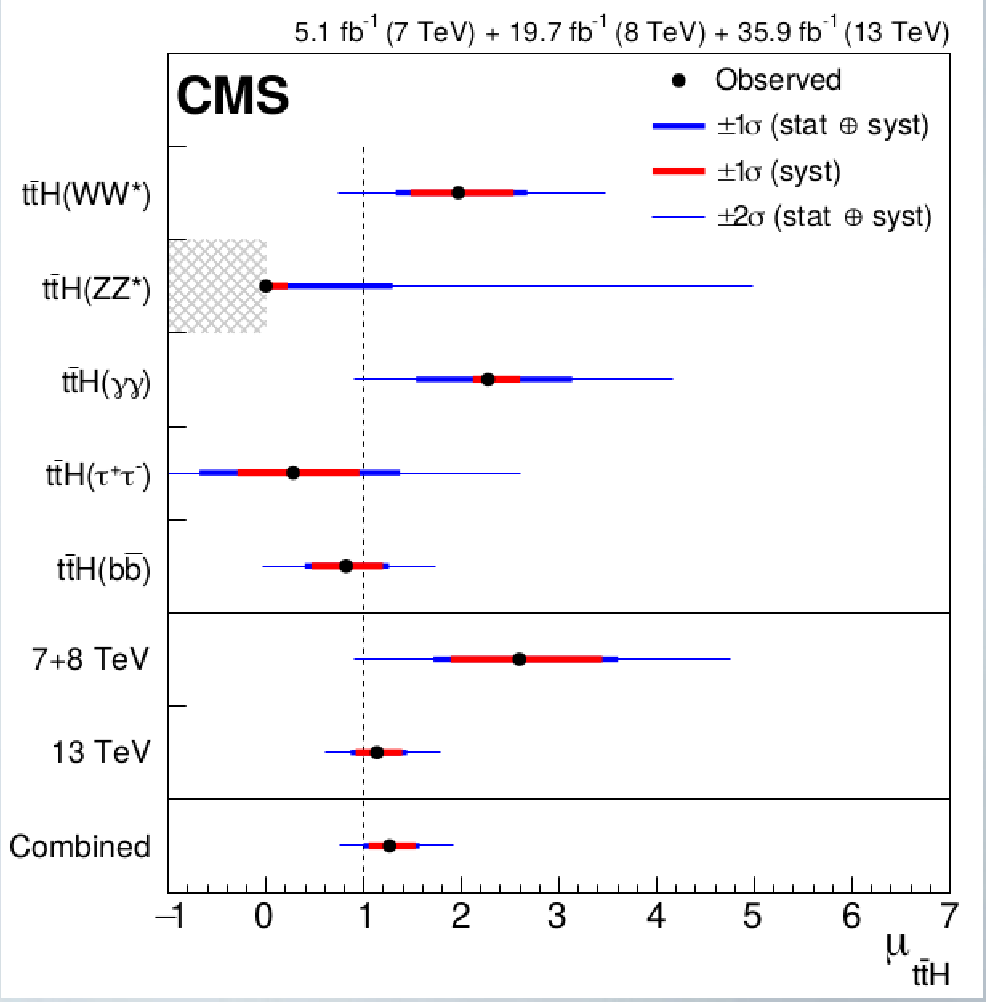
- **H -> ZZ -> 4l with 2016 data (CMS)**
  - ➢ $\mu = 1.05^{+0.19}_{-0.17} = 1.05^{+0.15}_{-0.14}(\text{stat.})^{+0.11}_{-0.09}(\text{sys.})$
- **Signal strength measured in bosonic and fermionic components**
  - $\mu_{\text{VBF,VH}} = 0.00^{+1.37}_{-0.00}$ $\quad \mu_{ggH,ttH} = 1.20^{+0.35}_{-0.31}$



Results from other Higgs analysis (2)

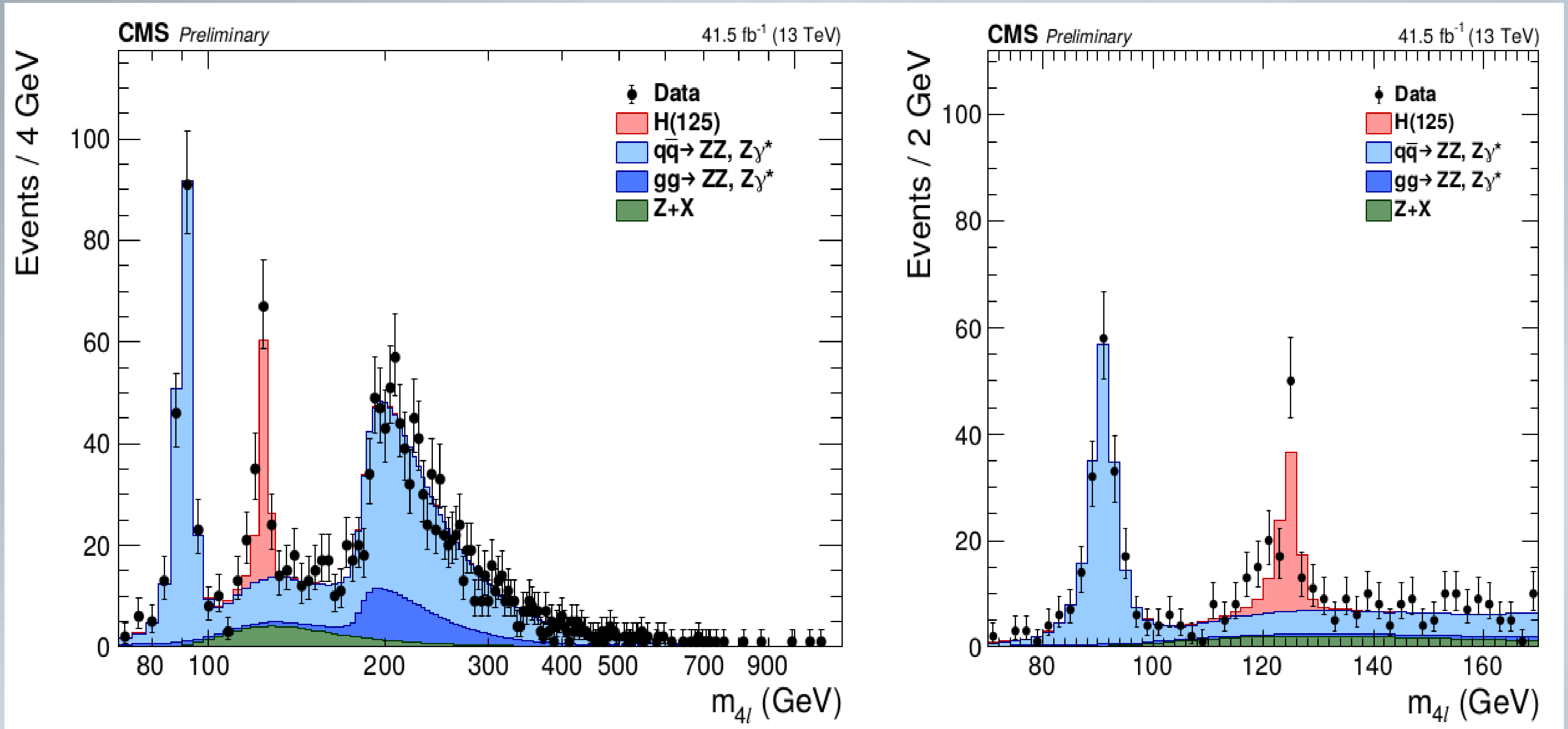
- **ttH signal strength modifier for different decay channels (CMS)**

Parameter	Best fit	Uncertainty: Stat	Uncertainty: Expt	Uncertainty: Thbgd	Uncertainty: Thsig
$\mu^{WW^*}_{t\bar{t}H}$	$1.97^{+0.71}_{-0.64}$	$^{+0.42}_{-0.41}$	$^{+0.46}_{-0.42}$	$^{+0.21}_{-0.21}$	$^{+0.25}_{-0.12}$
	$\left(^{+0.57}_{-0.54}\right)$	$\left(^{+0.39}_{-0.38}\right)$	$\left(^{+0.36}_{-0.34}\right)$	$\left(^{+0.17}_{-0.17}\right)$	$\left(^{+0.12}_{-0.03}\right)$
$\mu^{ZZ^*}_{t\bar{t}H}$	$0.00^{+1.30}_{-0.00}$	$^{+1.28}_{-0.00}$	$^{+0.20}_{-0.00}$	$^{+0.04}_{-0.00}$	$^{+0.09}_{-0.00}$
	$\left(^{+2.89}_{-0.99}\right)$	$\left(^{+2.82}_{-0.99}\right)$	$\left(^{+0.51}_{-0.00}\right)$	$\left(^{+0.15}_{-0.00}\right)$	$\left(^{+0.27}_{-0.00}\right)$
$\mu^{\gamma\gamma}_{t\bar{t}H}$	$2.27^{+0.86}_{-0.74}$	$^{+0.80}_{-0.72}$	$^{+0.15}_{-0.09}$	$^{+0.02}_{-0.01}$	$^{+0.29}_{-0.13}$
	$\left(^{+0.73}_{-0.64}\right)$	$\left(^{+0.71}_{-0.64}\right)$	$\left(^{+0.09}_{-0.04}\right)$	$\left(^{+0.01}_{-0.00}\right)$	$\left(^{+0.13}_{-0.05}\right)$
$\mu^{\tau^+\tau^-}_{t\bar{t}H}$	$0.28^{+1.09}_{-0.96}$	$^{+0.86}_{-0.77}$	$^{+0.64}_{-0.53}$	$^{+0.10}_{-0.09}$	$^{+0.20}_{-0.19}$
	$\left(^{+1.00}_{-0.89}\right)$	$\left(^{+0.83}_{-0.76}\right)$	$\left(^{+0.54}_{-0.47}\right)$	$\left(^{+0.09}_{-0.08}\right)$	$\left(^{+0.14}_{-0.01}\right)$
$\mu^{b\bar{b}}_{t\bar{t}H}$	$0.82^{+0.44}_{-0.42}$	$^{+0.23}_{-0.23}$	$^{+0.24}_{-0.23}$	$^{+0.27}_{-0.27}$	$^{+0.11}_{-0.03}$
	$\left(^{+0.44}_{-0.42}\right)$	$\left(^{+0.23}_{-0.22}\right)$	$\left(^{+0.24}_{-0.23}\right)$	$\left(^{+0.26}_{-0.27}\right)$	$\left(^{+0.11}_{-0.04}\right)$
$\mu^{7+8\,\mathrm{TeV}}_{t\bar{t}H}$	$2.59^{+1.01}_{-0.88}$	$^{+0.54}_{-0.53}$	$^{+0.53}_{-0.49}$	$^{+0.55}_{-0.49}$	$^{+0.37}_{-0.13}$
	$\left(^{+0.87}_{-0.79}\right)$	$\left(^{+0.51}_{-0.49}\right)$	$\left(^{+0.48}_{-0.44}\right)$	$\left(^{+0.50}_{-0.44}\right)$	$\left(^{+0.14}_{-0.02}\right)$
$\mu^{13\,\mathrm{TeV}}_{t\bar{t}H}$	$1.14^{+0.31}_{-0.27}$	$^{+0.17}_{-0.16}$	$^{+0.17}_{-0.17}$	$^{+0.13}_{-0.12}$	$^{+0.14}_{-0.06}$
	$\left(^{+0.29}_{-0.26}\right)$	$\left(^{+0.16}_{-0.16}\right)$	$\left(^{+0.17}_{-0.16}\right)$	$\left(^{+0.13}_{-0.12}\right)$	$\left(^{+0.11}_{-0.05}\right)$
$\mu_{t\bar{t}H}$	$1.26^{+0.31}_{-0.26}$	$^{+0.16}_{-0.16}$	$^{+0.17}_{-0.15}$	$^{+0.14}_{-0.13}$	$^{+0.15}_{-0.07}$
	$\left(^{+0.28}_{-0.25}\right)$	$\left(^{+0.15}_{-0.15}\right)$	$\left(^{+0.16}_{-0.15}\right)$	$\left(^{+0.13}_{-0.12}\right)$	$\left(^{+0.11}_{-0.05}\right)$



Results from other Higgs analysis (3)

- **H -> ZZ -> 4l with 2017 data of 41.5 /fb (CMS)**

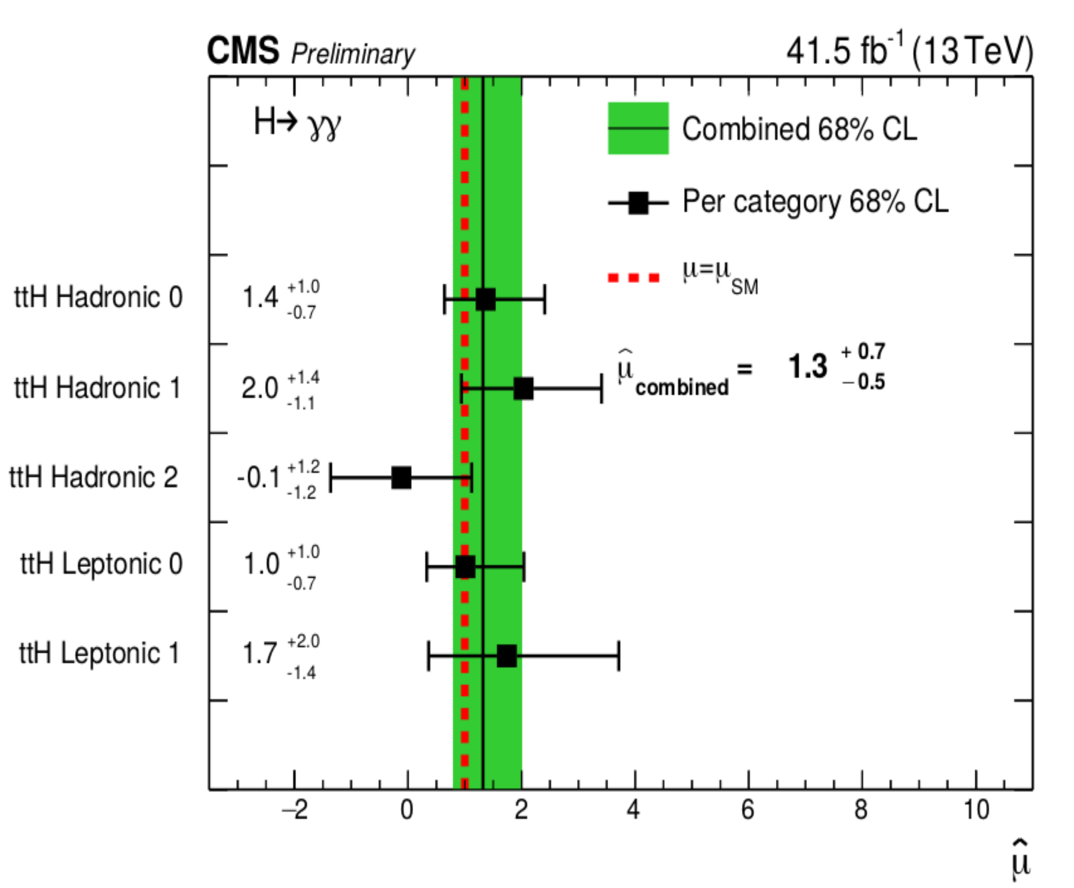


	Inclusive	$\mu_{ggH,b\bar{b}H}$	μ_{VBF}	μ_{VHhad}	μ_{VHlep}	$\mu_{t\bar{t}H,tqH}$
Expected	$1.00^{+0.14}_{-0.13}(\text{stat})^{+0.11}_{-0.09}(\text{syst})$	$1.00^{+0.22}_{-0.20}$	$1.00^{+1.19}_{-0.79}$	$1.00^{+3.24}_{-1.00}$	$1.00^{+3.36}_{-1.00}$	$1.00^{+2.47}_{-1.00}$
Observed	$1.10^{+0.14}_{-0.13}(\text{stat})^{+0.13}_{-0.11}(\text{syst})$	$1.14^{+0.23}_{-0.20}$	$1.12^{+1.19}_{-0.83}$	$0.00^{+1.54}_{-0.00}$	$2.23^{+3.95}_{-2.12}$	$0.00^{+0.93}_{-0.00}$

Results from other Higgs analysis (4)

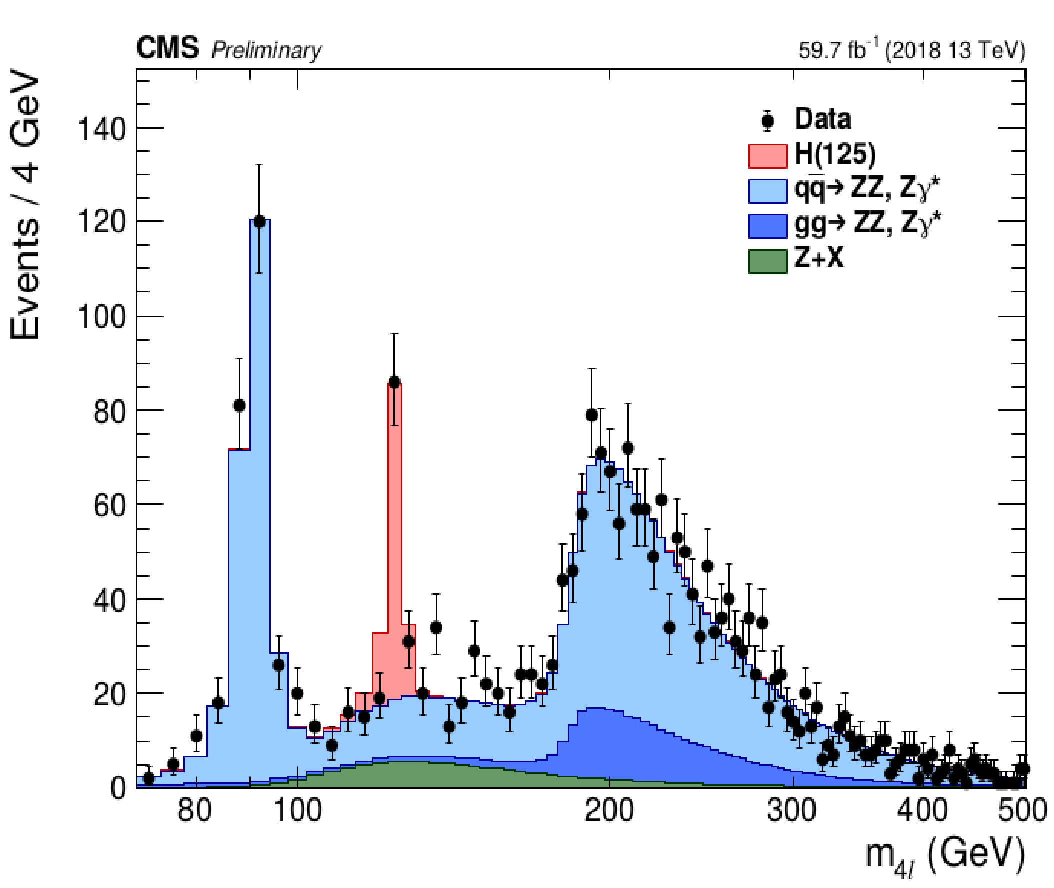
- **ttH analysis only for diphoton decay channel with 2017 data of 41.5 /fb (CMS)**

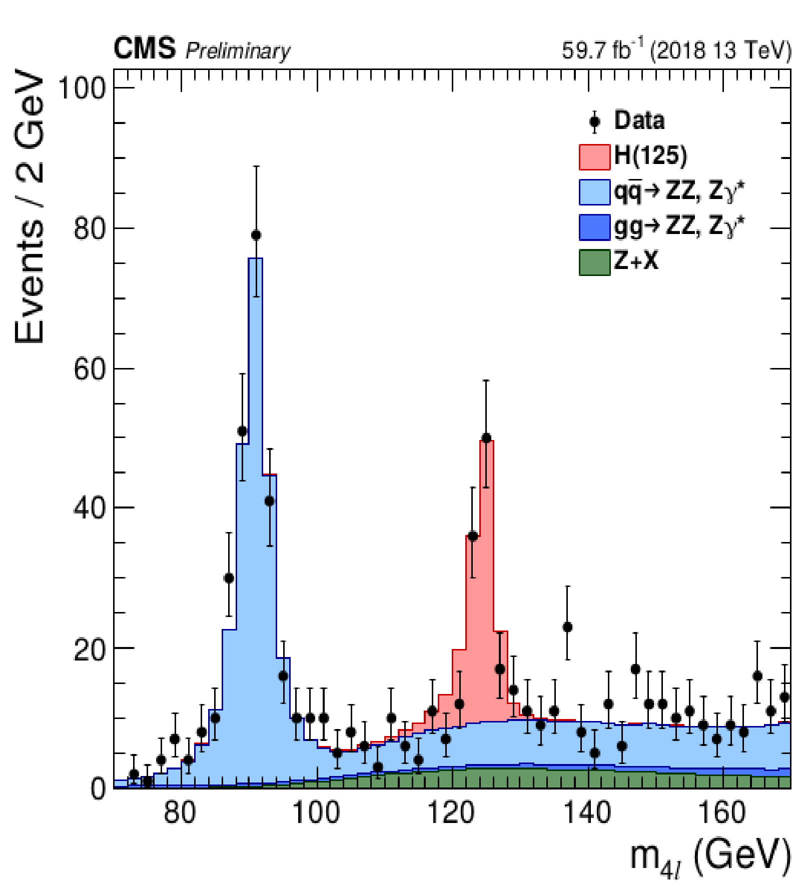
$$\mu_{ttH} = 1.3^{+0.7}_{-0.5} = 1.3^{+0.6}_{-0.5}(\text{stat.})^{+0.3}_{-0.1}(\text{syst.})$$



Results from other Higgs analysis (5)

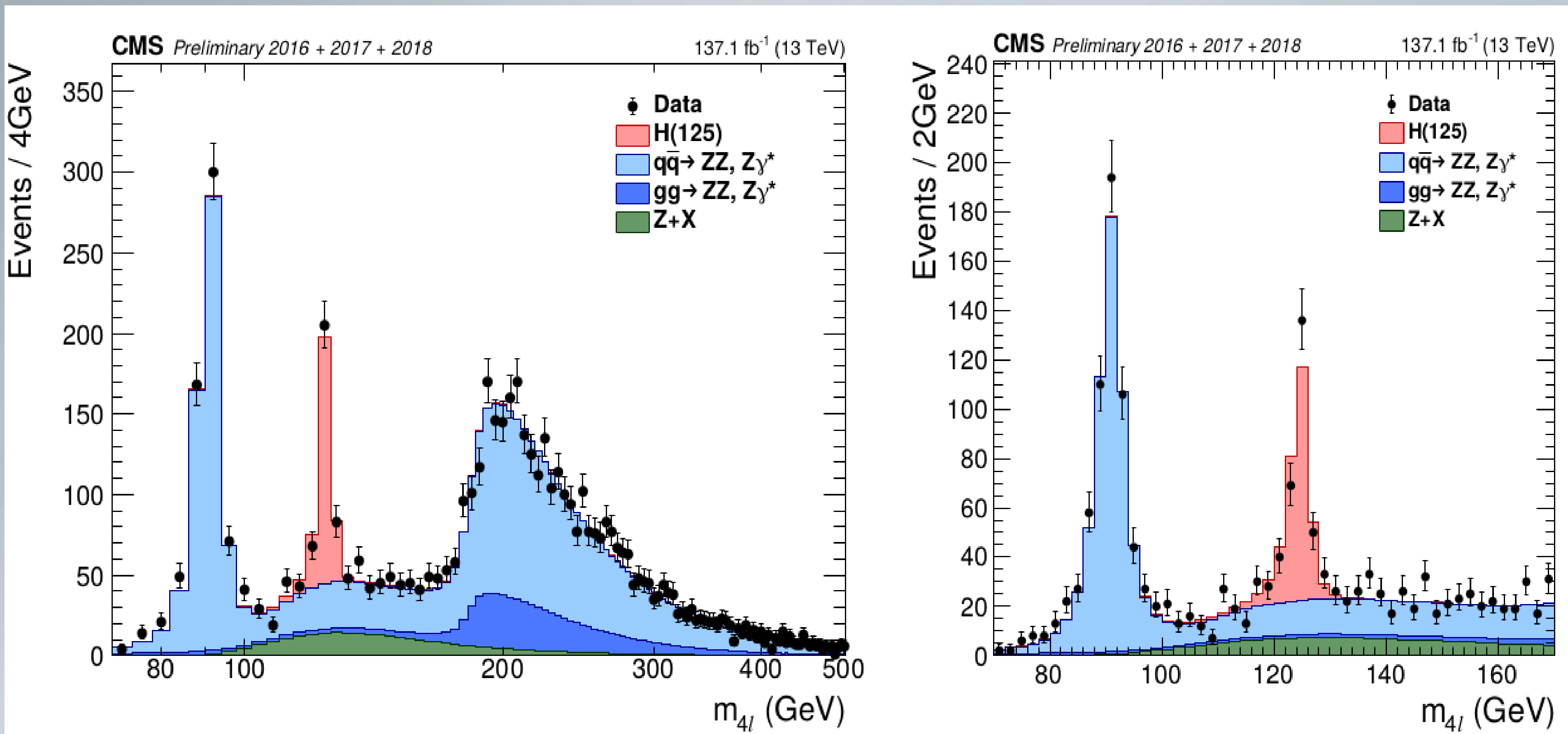
- **H -> ZZ -> 4l with 2018 data of 59.7 /fb (CMS)**





Results from other Higgs analysis (6)

- **H -> ZZ -> 4l for whole Run2 combined together – 137.1 /fb (CMS)**



	Expected	Observed
$\mu_{\text{inclusive}}$	$1.00^{+0.08}_{-0.08}(\text{stat.})^{+0.09}_{-0.07}(\text{syst.})$	$0.94^{+0.07}_{-0.07}(\text{stat.})^{+0.08}_{-0.07}(\text{syst.})$
μ_{ggH}	$1.00^{+0.10}_{-0.10}(\text{stat.})^{+0.09}_{-0.07}(\text{syst.})$	$0.97^{+0.09}_{-0.09}(\text{stat.})^{+0.09}_{-0.07}(\text{syst.})$
μ_{VBF}	$1.00^{+0.54}_{-0.45}(\text{stat.})^{+0.27}_{-0.14}(\text{syst.})$	$0.64^{+0.45}_{-0.36}(\text{stat.})^{+0.16}_{-0.09}(\text{syst.})$
μ_{VH}	$1.00^{+0.91}_{-0.72}(\text{stat.})^{+0.29}_{-0.16}(\text{syst.})$	$1.15^{+0.89}_{-0.72}(\text{stat.})^{+0.26}_{-0.16}(\text{syst.})$
$\mu_{\text{t\bar{t}H,tH}}$	$1.00^{+1.16}_{-0.73}(\text{stat.})^{+0.19}_{-0.04}(\text{syst.})$	$0.13^{+0.92}_{-0.13}(\text{stat.})^{+0.11}_{-0.00}(\text{syst.})$

Summary

- Details of the analysis for Higgs decaying to two photons using data recorded by CMS in the year 2016 are presented here
- Properties measured for the Higgs boson are really close to the SM perdicted one, within the uncertainty
- Results from other CMS analysis for SM Higgs are also shown here
- All of these measurements are comparable to SM prediction taking into account the uncertainties

Thank you All...

Backup Slides... ⬇

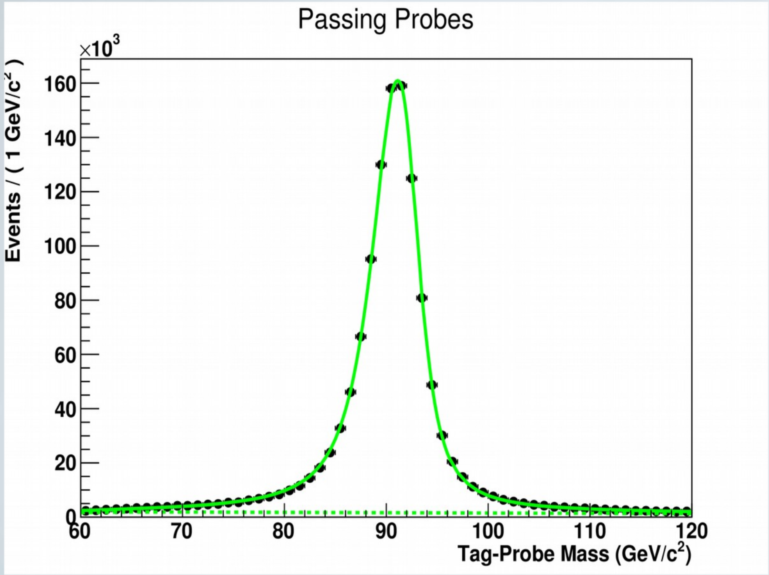
Preselection Cuts

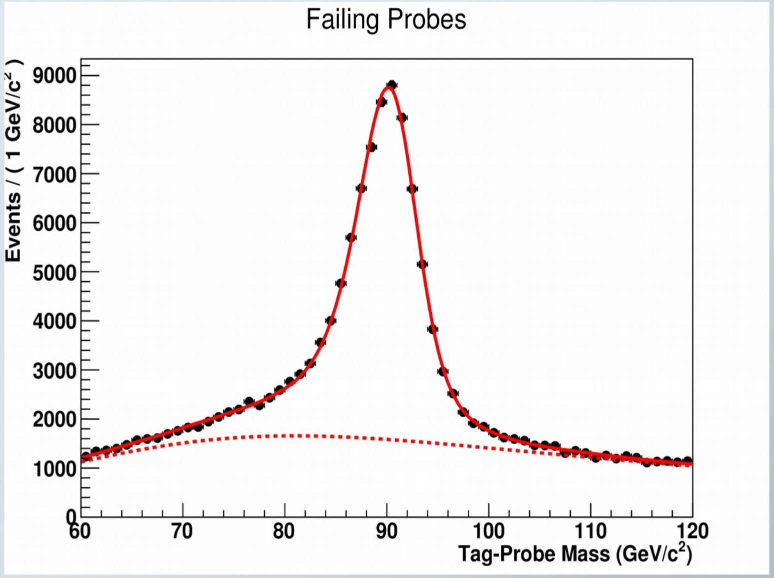
- **HLT:**
  - HLT_Diphoton30_18_R9Id_OR_IsoCaloId_AND_HE_R9Id_Mass90
- **Preselection cuts:**

	H/E	$\sigma_{i\eta i\eta}$ (5x5)	R9 (5x5)	pfPhoIso	TrackerIso
EB; R9>0.85	< 0.08	-	> 0.5	-	-
EB; R9≤0.85	< 0.08	< 0.015	> 0.5	< 4.0	< 6.0
EE; R9>0.9	< 0.08	-	> 0.8	-	-
EE; R9≤0.9	< 0.08	< 0.035	> 0.8	< 4.0	< 6.0

Preselection Scale Factors

- **Tag n Probe method:**
  - "Tag" electron is selected from the event, which passes a set of tight cuts with respect to the preselection cuts
  - Electrons which passes the usual preselection cuts from the same event are "Probe" electrons
  - The invariant mass of the all possible pairs of electrons of the events are fitted to extract amount of signal and background events. The ratio of signal events over all selected events is the effeciency of the preselection cuts





Vertex Identification Variables

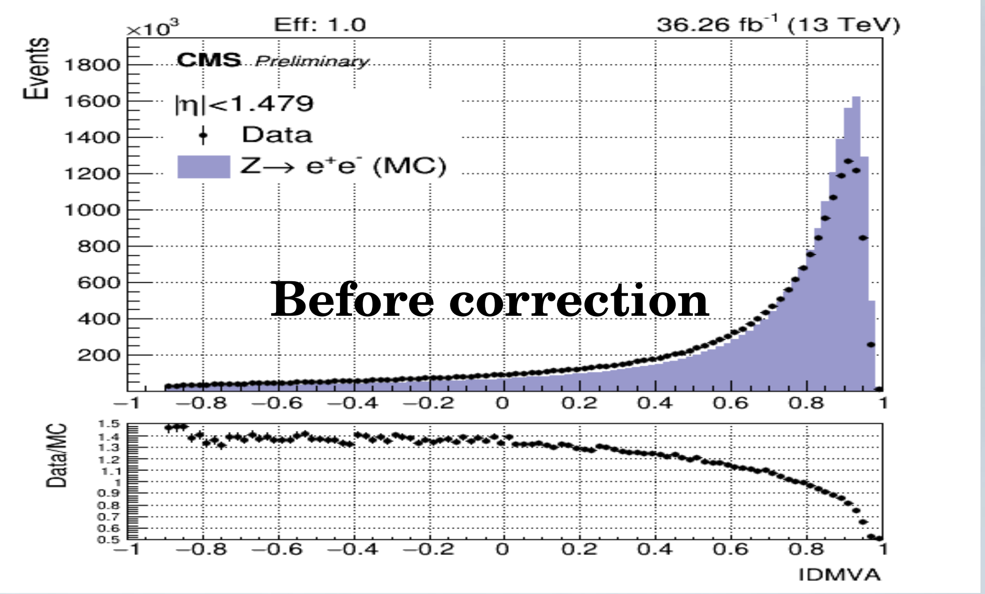
- **Vertex ID BDT:**
  - Sum of pT squared for diphotons
  - pT balance between two photons associated to the vertex
  - Asymmetry of pT between the two photon
  - If one or both of the diphoton is converted, the no. of conversions and pull between longitudinal position of the reconstructed vertex and longitudinal position of the vertex estimated using conversion tracks is used in addition
- **Vertex Probability BDT:**
  - No. of vertices in the event
  - Vertex ID BDT output
  - Distance between chosen vertex and second and third choice
  - Transverse momentum of diphoton system
  - Number of photons with an associated conversion track

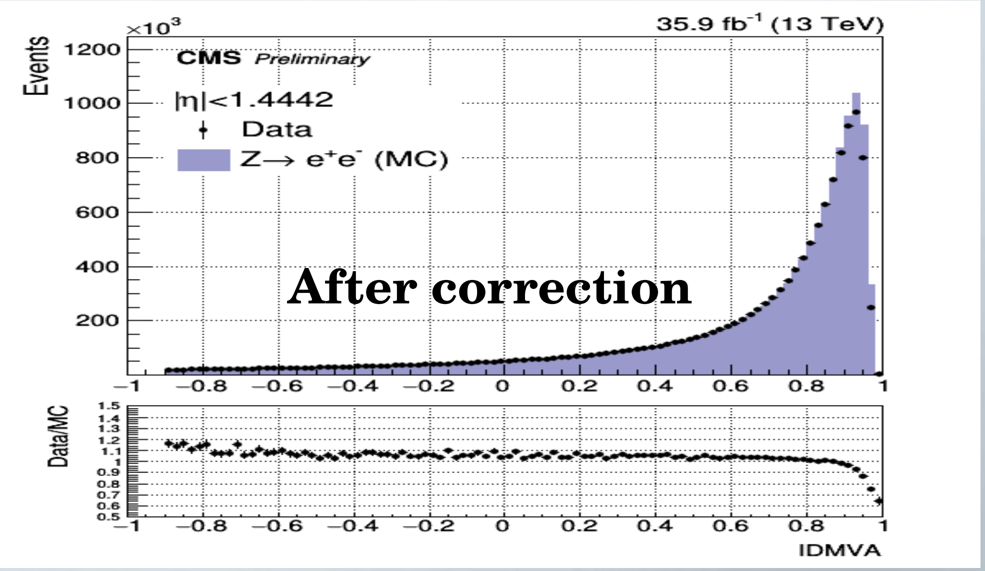
Photon Identification Variables

- **Input variables to Photon ID BDT:**
  - Shower shape variables:
    - R9 (full5x5) : E_{5x5}/E_{sc} where E_{sc} Supercluster energy
    - S4 ratio : E_{2x2}/E_{5x5}
    - σ_{ieie} : The energy weighted standard deviation of single crystal eta within SC
    - Super cluster η width : The lateral extension of EM shower measured in terms of crystal cells in the direction of eta
    - Super cluster φ width : Lateral extension of EM shower in direction of phi
    - Covariance_{ieip} : covariance of single crystal eta and phi within supercluster
    - ES σ_{RR} : Standard deviation of shower spread in the x and y planes of preshower
  - Isolation variables:
    - pf Photon Isolation
    - charged hadron isolation (wrt right vertex)
    - charged hadron isolation (wrt wrong vertex)
  - Super cluster η : η of the supercluster associated with the reconstructed photon
  - Rho (ρ) : the energy median density per unit area in the event
  - E_{RAW}: energy of the super cluster
- **Output:**
  - Photon ID BDT score

Shower Shape Correction

- Input variables for Photon ID BDT and its output are compared between data and simulation for electrons coming from Z decays for validation purpose
- Discrepancies between data and simulation are observed for BDT output
  - Reasons: discrepancies between data and simulation for several input variables of BDT (mainly shower shape variables)
- Corresponding variables are corrected for simulation by reweighting them with respect to data using quantile method of reweight
- Photon ID BDT is retrained after applying these corrections to the simulation

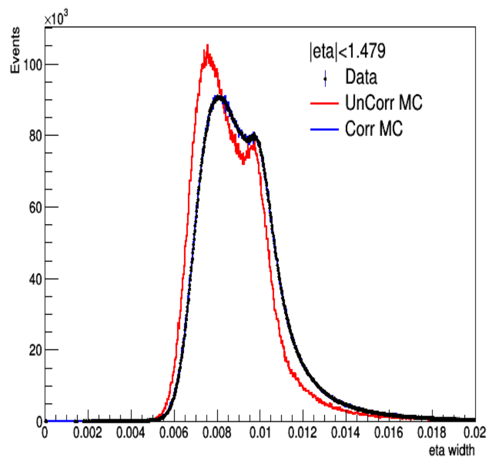


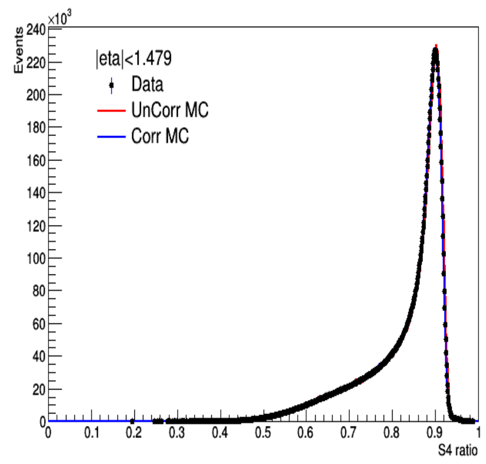


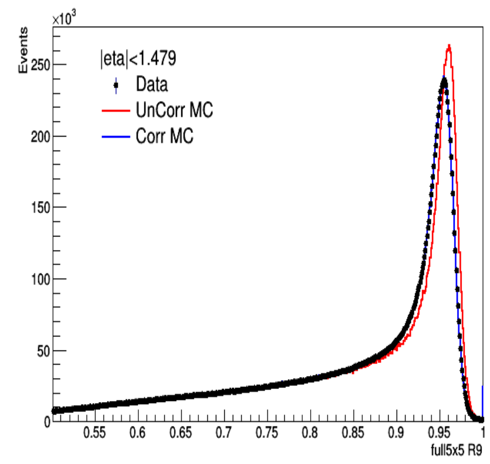
Shower shape correction plots

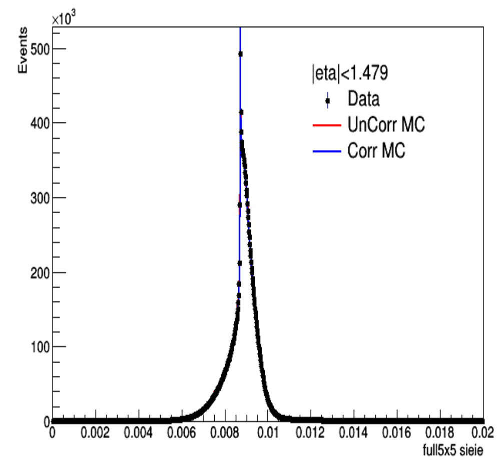
- Following plots are for all variables seperated only in EB and EE

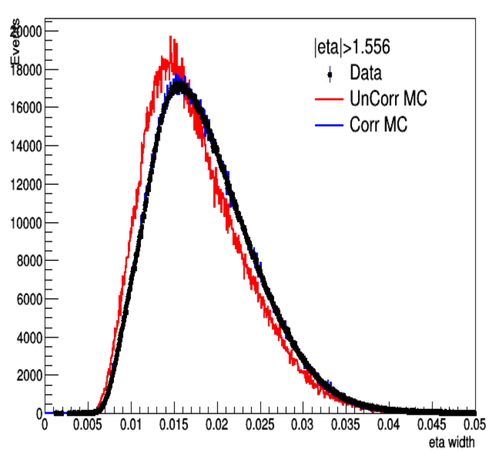
Eta width | **S4 ratio** | **Full5x5 R9** | **Full5x5 sieie**

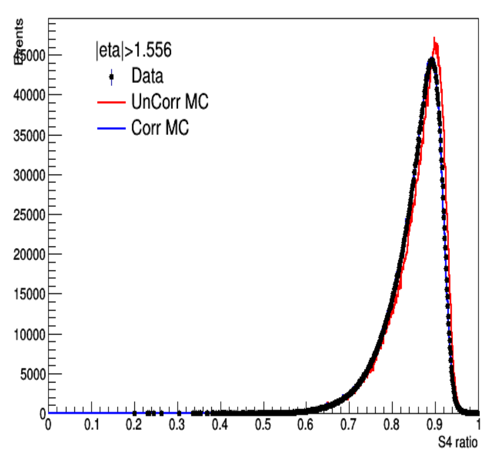


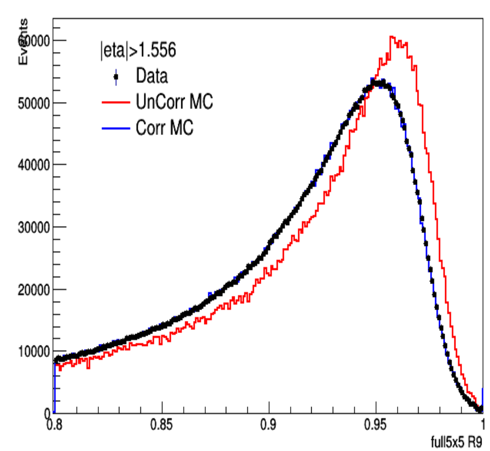


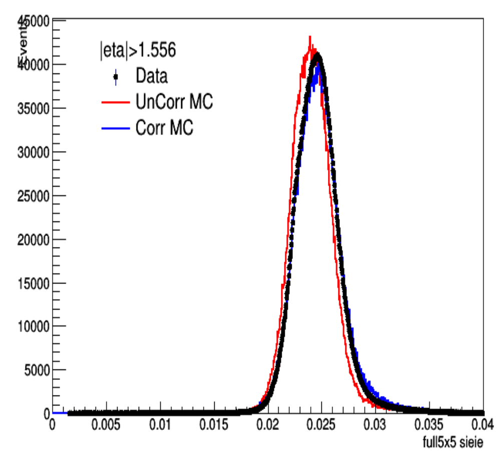












Diphoton BDT Input Variables

- The transverse momentum of both the photons, rescaled for diphoton mass
- Pseudorapidity of both the photons
- Cosine of the angle between the two photons in the transverse plane, cos()
- Photon identification BDT score of both the photons
- Per-event relative mass resolution estimate, under the hypothesis that the mass has been reconstructed using the correct primary vertex
- Per-event relative mass resolution estimate, under the hypothesis that the mass has been reconstructed using an incorrect primary vertex
- Per-event probability estimate that the correct primary vertex has been used to reconstruct the mass, based on the event-level vertex selection MVA

Functions for Background model

- Sums of exponentials:

$$f_N(x) = \sum_{i=1}^{N} p_{2i} e^{p_{2i+1} x},$$

- Sums of polynomials (in the Bernstein basis):

$$f_N(x) = \sum_{i=0}^{N} p_i b_{(i,N)}, \text{ where } b_{(i,N)} := \binom{N}{i} x^i (1-x)^{N-i},$$

- Laurent series:

$$f_N(x) = \sum_{i=1}^{N} p_i x^{-4+\sum_{j=1}^{i}(-1)^j (j-1)},$$

- Sums of power-law functions:

$$f_N(x) = \sum_{i=1}^{N} p_{2i} x^{-p_{2i+1}},$$

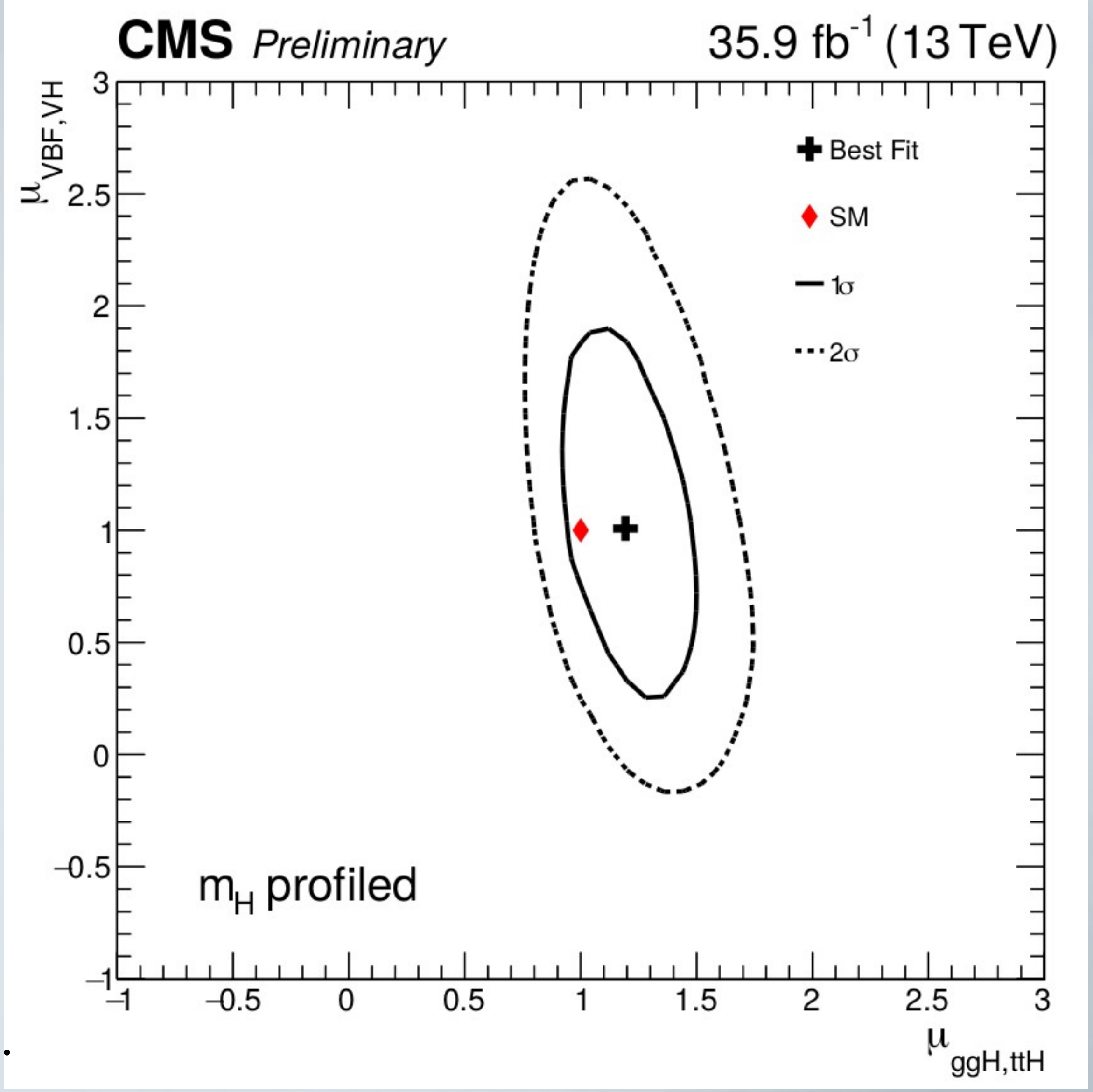
Expected events per category

Event Categories	SM 125 GeV Higgs boson expected signal													Bkg
	Total	ggH	VBF	ttH	bbH	tHq	tHW	WH lep	ZH lep	WH had	ZH had	σ_{eff}	σ_{HM}	(GeV^{-1})
Untagged 0	45.83	80.19 %	11.75 %	1.83 %	0.40 %	0.47 %	0.22 %	0.41 %	0.19 %	2.96 %	1.58 %	1.32	1.24	21.92
Untagged 1	480.56	86.81 %	7.73 %	0.56 %	1.15 %	0.13 %	0.02 %	0.47 %	0.27 %	1.81 %	1.04 %	1.47	1.32	924.21
Untagged 2	670.45	89.76 %	5.48 %	0.44 %	1.18 %	0.08 %	0.01 %	0.51 %	0.34 %	1.40 %	0.81 %	1.94	1.68	2419.53
Untagged 3	610.07	91.13 %	4.51 %	0.48 %	1.07 %	0.07 %	0.01 %	0.55 %	0.30 %	1.21 %	0.69 %	2.62	2.28	4855.00
VBF 0	10.01	21.69 %	77.09 %	0.34 %	0.35 %	0.29 %	0.03 %	0.03 %	0.00 %	0.19 %	-0.01 %	1.51	1.30	1.60
VBF 1	8.64	33.58 %	64.64 %	0.39 %	0.52 %	0.36 %	0.04 %	0.13 %	0.03 %	0.24 %	0.07 %	1.66	1.38	3.25
VBF 2	27.76	50.14 %	46.46 %	0.81 %	0.73 %	0.53 %	0.07 %	0.20 %	0.06 %	0.71 %	0.27 %	1.61	1.36	18.89
ttH Hadronic	5.85	10.99 %	0.70 %	77.54 %	2.02 %	4.13 %	2.02 %	0.09 %	0.05 %	0.63 %	1.82 %	1.48	1.30	2.40
ttH Leptonic	3.81	1.90 %	0.05 %	87.48 %	0.08 %	4.73 %	3.04 %	1.53 %	1.15 %	0.02 %	0.02 %	1.60	1.35	1.50
ZH Leptonic	0.49	0.00 %	0.00 %	2.56 %	0.00 %	0.02 %	0.13 %	0.00 %	97.30 %	0.00 %	0.00 %	1.65	1.43	0.12
WH Leptonic	3.61	1.26 %	0.59 %	5.18 %	0.18 %	3.03 %	0.73 %	84.48 %	4.33 %	0.12 %	0.09 %	1.64	1.43	2.09
VH LeptonicLoose	2.75	9.16 %	2.70 %	2.34 %	0.57 %	1.81 %	0.13 %	63.62 %	18.87 %	0.56 %	0.23 %	1.67	1.56	3.50
VH Hadronic	9.69	57.38 %	3.68 %	3.61 %	0.35 %	1.39 %	0.27 %	0.17 %	0.42 %	20.47 %	12.26 %	1.38	1.31	7.22
VH Met	4.25	23.63 %	2.46 %	14.45 %	0.41 %	2.00 %	1.14 %	25.17 %	28.60 %	1.32 %	0.82 %	1.55	1.38	3.49
Total	1883.77	86.96 %	7.09 %	1.00 %	1.09 %	0.15 %	0.04 %	0.81 %	0.42 %	1.55 %	0.89 %	1.95	1.62	8264.73

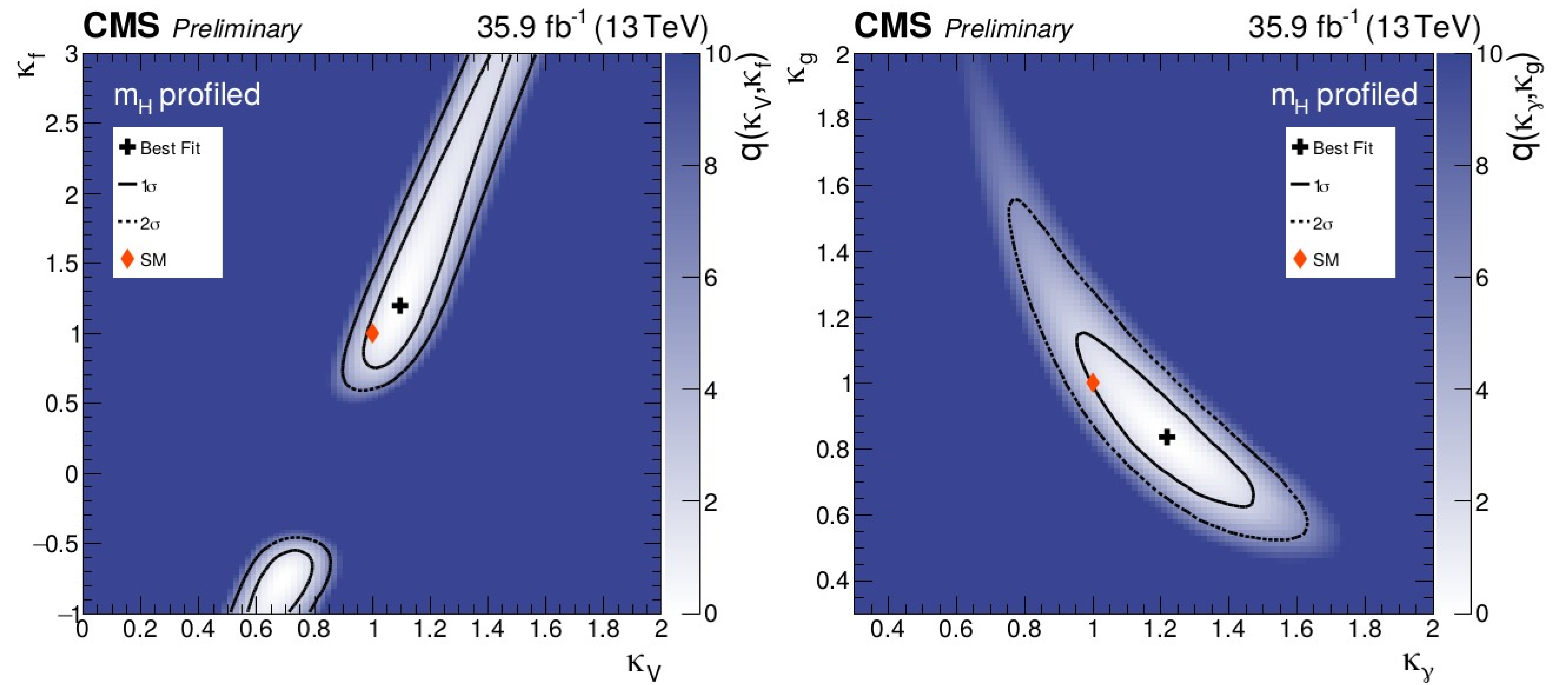
Table 3: The expected number of signal events per category and the percentage breakdown per production mode in that category. The σ_{eff}, computed as the smallest interval containing 68.3% of the invariant mass distribution, and σ_{HM}, computed as the width of the distribution at half of its highest 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.

Results

- The two-dimensional best-fit (black cross) of the signal strengths for fermionic (ggH, tt̄H) and bosonic (VBF, ZH, WH) production modes compared to the SM expectations (red diamond). The Higgs boson mass is profiled in the fit. The solid (dashed) line represents the 1 (2) standard deviation confidence region.



Results



- Two-dimensional likelihood scans of κ-f versus κ-V (left) and κ-g versus κ-γ (right)
- All four variables are expressed relative to the SM expectations
- The mass of the Higgs boson is profiled in the fits
- The crosses indicate the best-fit values, the diamonds indicate the Standard Model expectations.

List of Publication

- “Measurements of Higgs boson properties in the diphoton decay channel in proton-proton collisions at sqrt(s) = 13 TeV”, CMS Collaboration, J. High Energ. Phys. (2018) 2018:185
- CMS public analysis summary:
  - “First measurements of the Higgs boson production in the diphoton decay channel at sqrt(s) = 13 TeV ”- CMS Physics Analysis Summary, CMS PAS HIG-15-005 (public).
  - “Updated measurements of Higgs boson production in the diphoton decay channel at sqrt(s) = 13 TeV in pp collisions at CMS ”- CMS Physics Analysis Summary, CMS PAS HIG-16-020 (public).
  - “Measurements of properties of the Higgs boson in the diphoton decay channel with the full 2016 dataset ”- CMS Physics Analysis Summary, CMS PAS HIG-16-040 (public).