Measurements of Higgs boson production and properties in the di-photon decay channel using the CMS detector

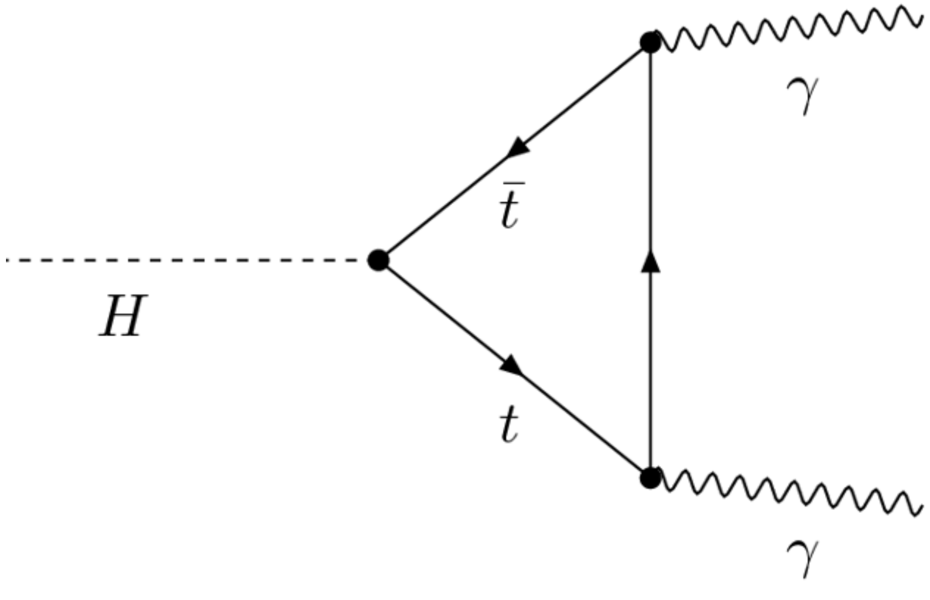
Vittorio Raoul Tavolaro on behalf of the CMS collaboration

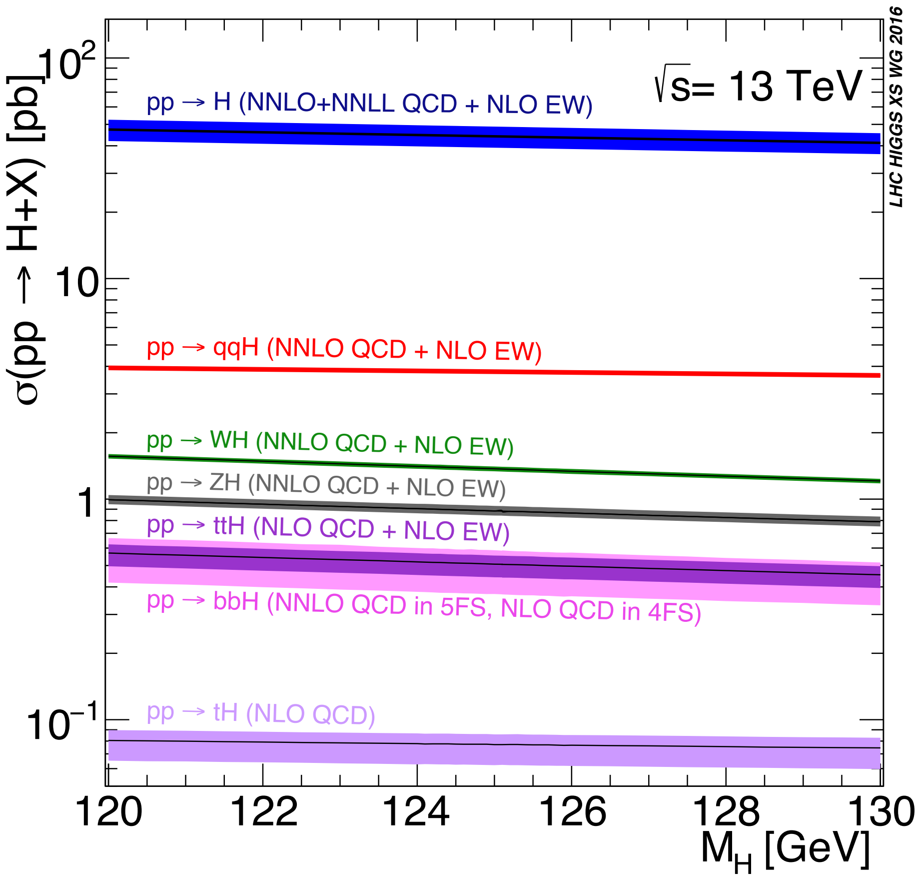
38th International Conference on High Energy Physics, 4 August 2016, Chicago IL

ETH Institute for Particle Physics

Introduction

- Clean final state with two highly **energetic photons**
- Final state **fully reconstructed** with high resolution
- Very small branching fraction (**~0.2%**)



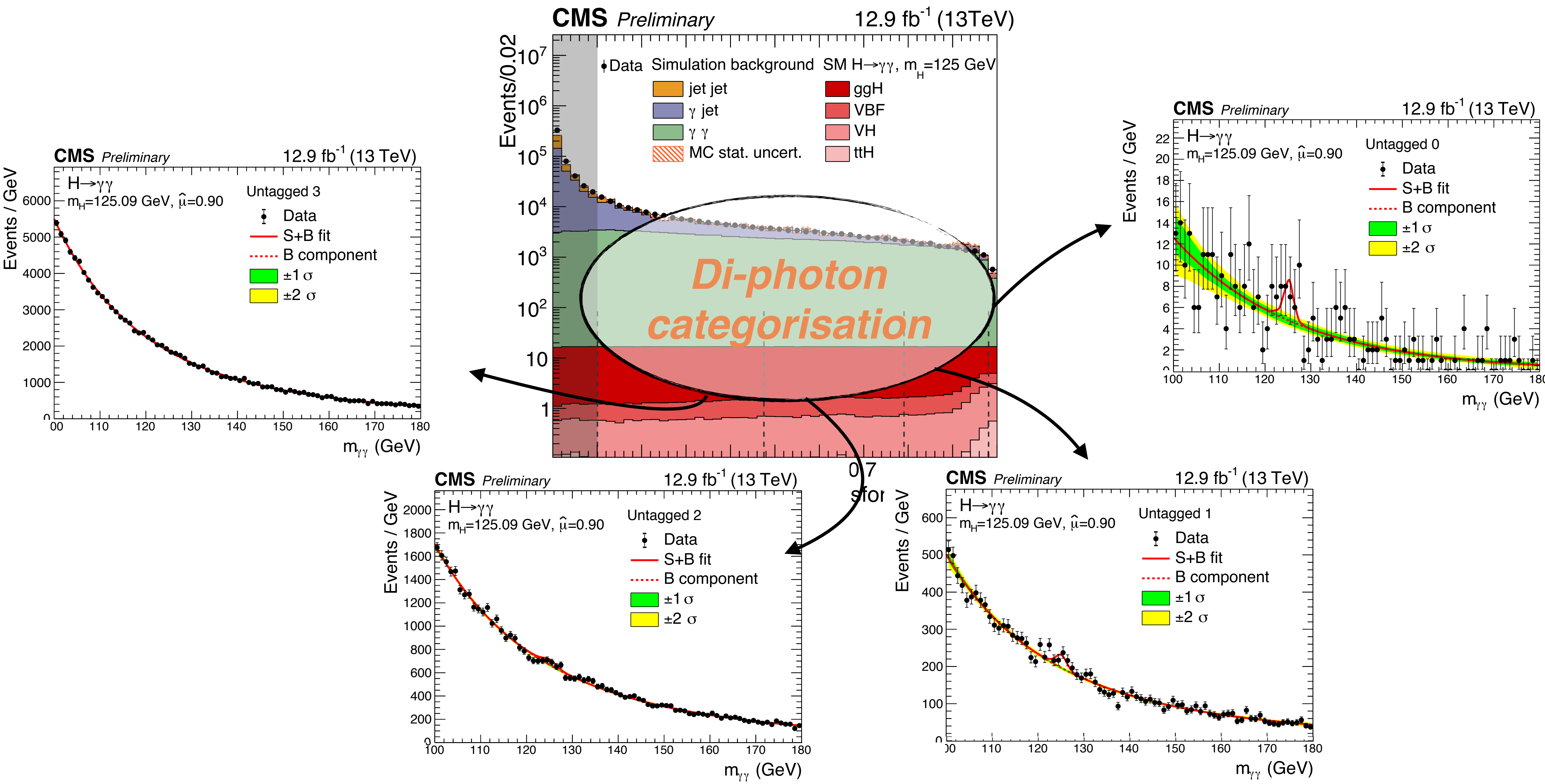


- Large **backgrounds** (ɣɣ, ɣj, jj)
- **Exclusive categories** targeting: gluon-gluon fusion **(ggH),** vector boson fusion **(VBF)** and **ttH** production modes
- Search for a **narrow peak** on a falling background in mass distribution
- **2016** dataset analysed, **12.9/fb** collected at 13 TeV (**HIG-16-020**)

Analysis strategy

ETH Zürich

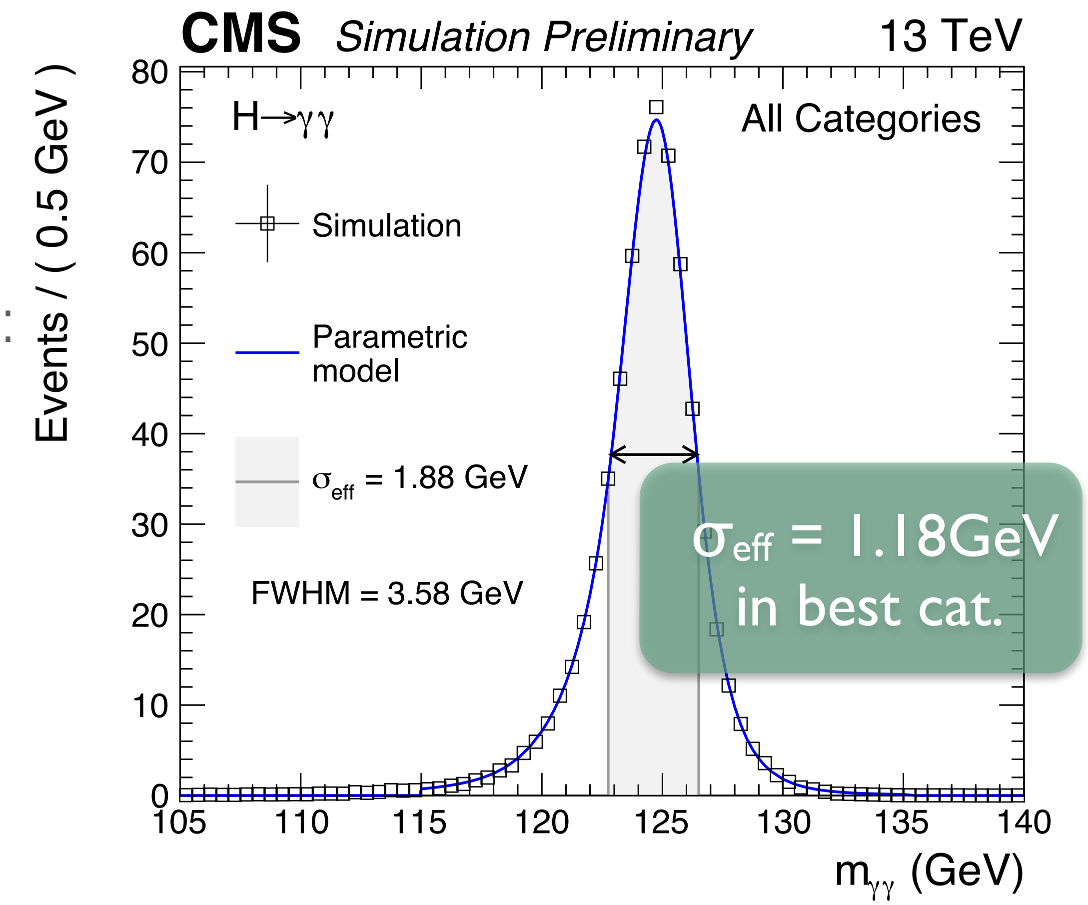
- **Photons** energy reconstructed and corrected, rejection of fake photons
- **Events categorised** into classes (S/B, mass resolution, additional particles) to **improve the analysis** sensitivity
- Extraction of signal through fit of **di-photon invariant mass** spectrum in each category

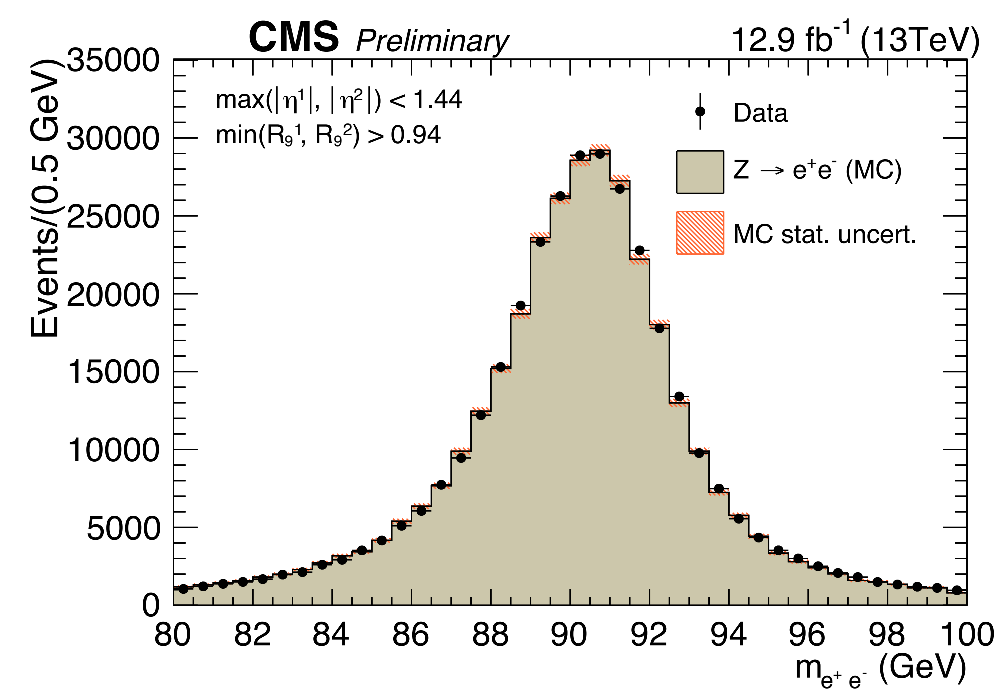


Photon Energy

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

- Electro-magnetic calorimeter (ECAL) response:
  - corrected for **change in time**
  - inter-calibrated to be uniform in η/ϕ
  - adjustment of **absolute scale**
- **Energy and its uncertainty** corrected for local and global shower containment:
  - **regression** targeting E_{true}/E_{reco}
- **Scale** vs time and **resolution** calibration: Z→ee peak used as reference
- **Corrected** energies and resolutions used in the analysis

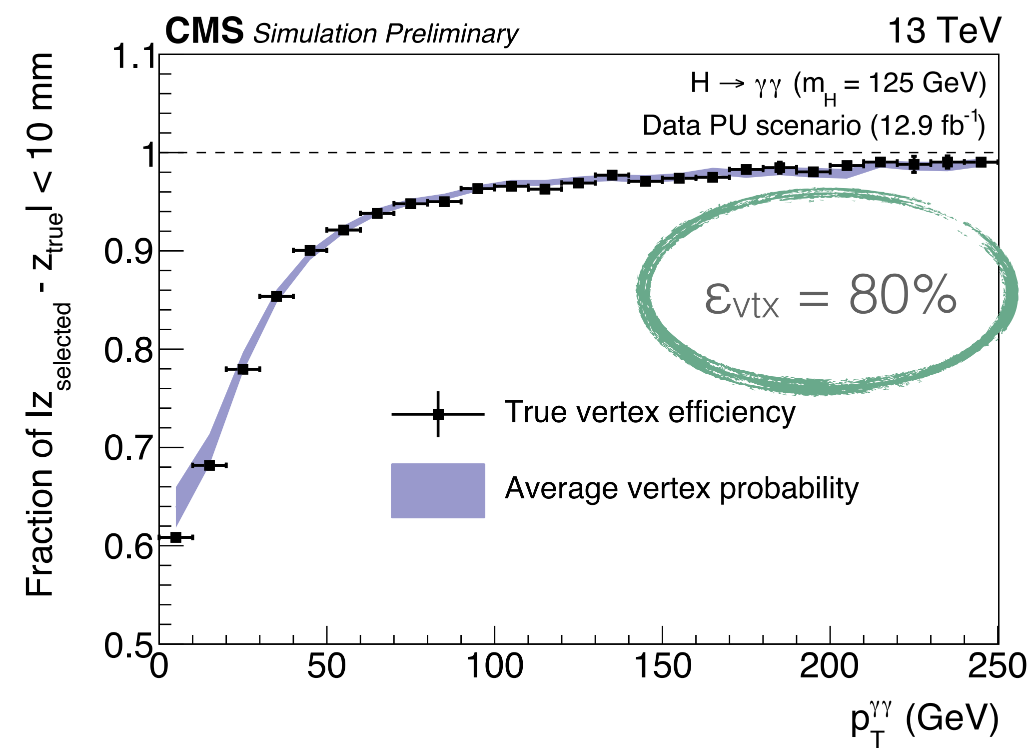


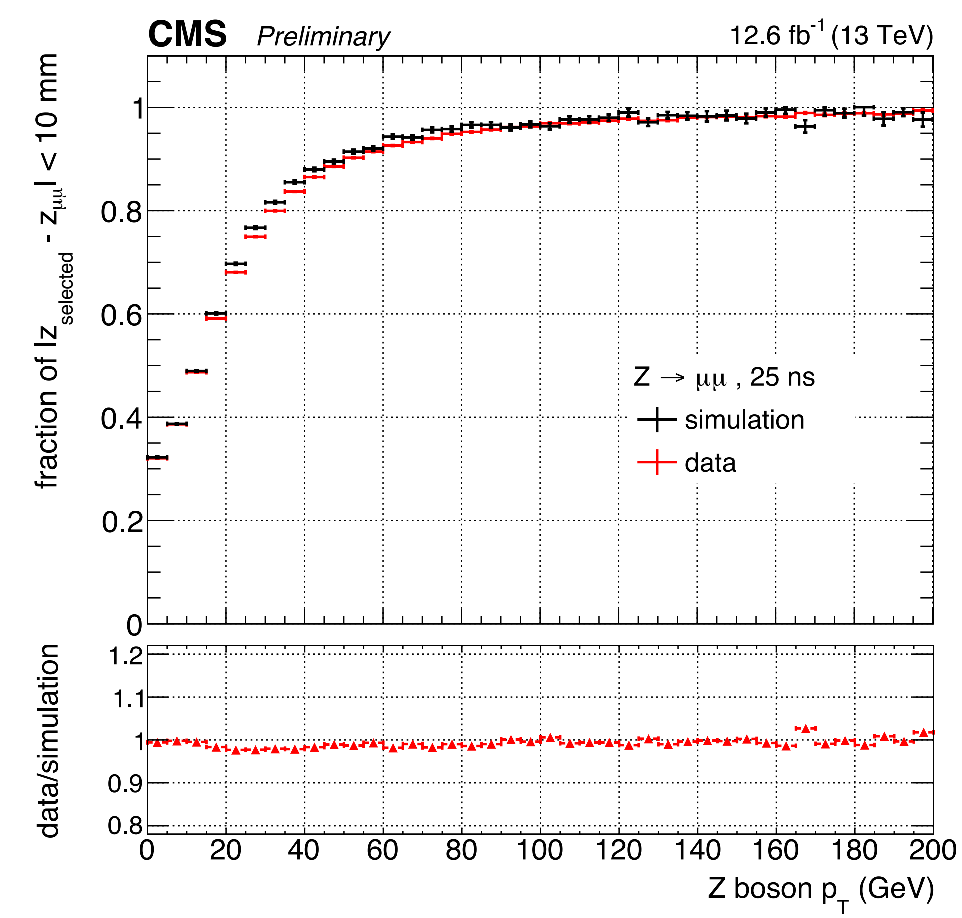


Vertex identification

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

- Vertex assignment correct within **1 cm → negligible** impact on mass resolution
- No ionization in the tracker for photons
- **Multi-variate approach** for vertex identification
  - exploit kinematic correlations and **track** distribution **imbalance**
  - direction of **conversion tracks**, when present
- Second MVA estimates **probability of correct vertex** choice, used for di-photon classification
- Method validated on Z→μμ (γ+j for converted γ) events, where vertex found after removing muon tracks



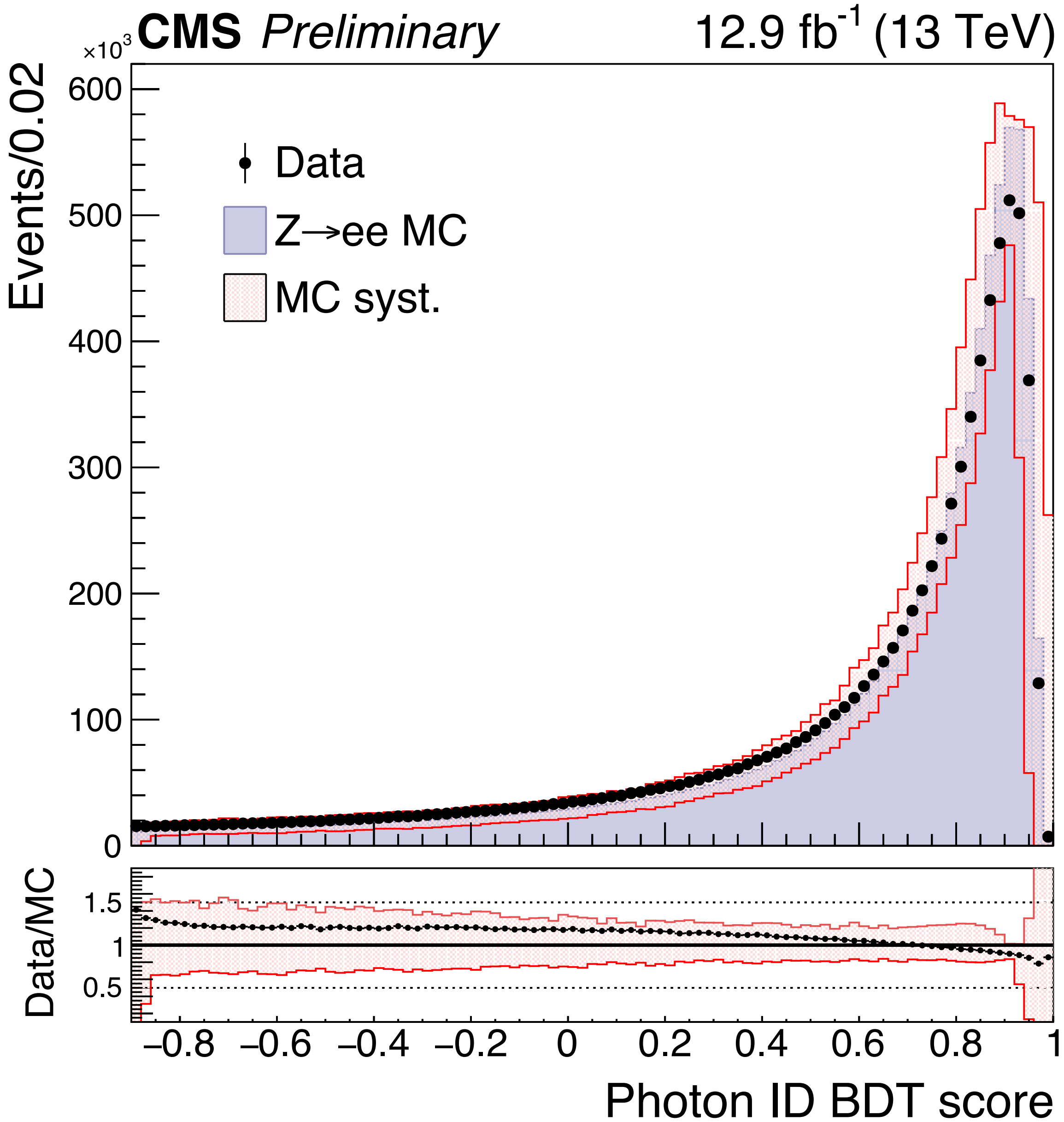


Photon selection

- **Double-photon trigger** selection based on transverse energy, $m_{\gamma\gamma}$, isolation and electromagnetic shower shapes variables

- Minimal **pre-selection**, similar to but tighter than trigger selection

- **MVA based** photon ID classifier:
  - goal: discriminate between **prompt** and **fake** photons
  - inputs are shower shape variables, particle-flow isolations, kinematic of photon, median energy density (ρ)

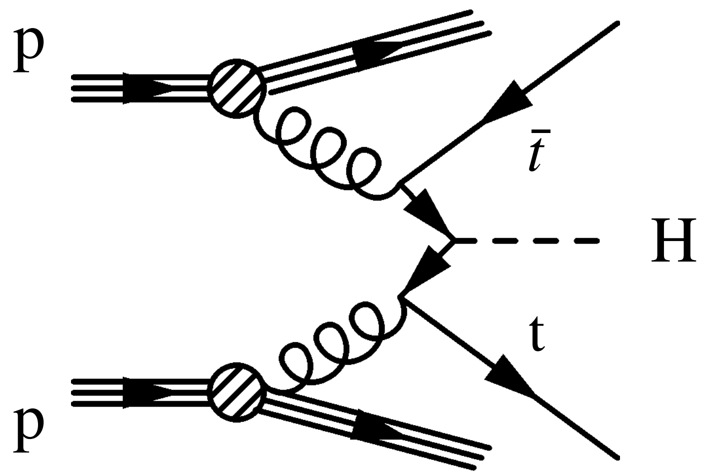


ttH tags

ttH leptonic tag

$t\bar{t} \to bl\nu_l\bar{b}q\bar{q}'$ $t\bar{t} \to bl\nu_l\bar{b}l'\nu_{l'}$

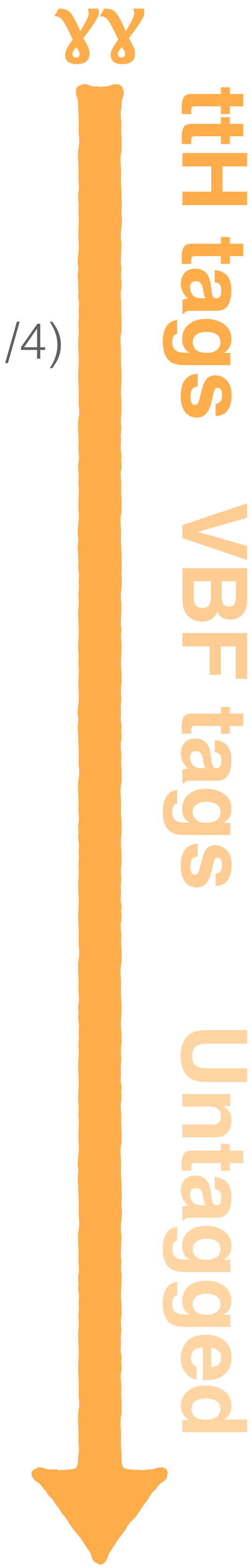
- (sub)lead $p_T/m_{\gamma\gamma} > 1/2(1/4)$
- at least one lepton (ℓ=e,μ), away from Z peak
- ≥ 2 jets
- ≥ 1 b-jet



ttH hadronic tag

$t\bar{t} \to bq\bar{q}'\bar{b}q\bar{q}'$

- (sub)lead $p_T/m_{\gamma\gamma} > 1/2(1/4)$
- 0 leptons
- ≥ 5 jets
- ≥ 1 b-jet

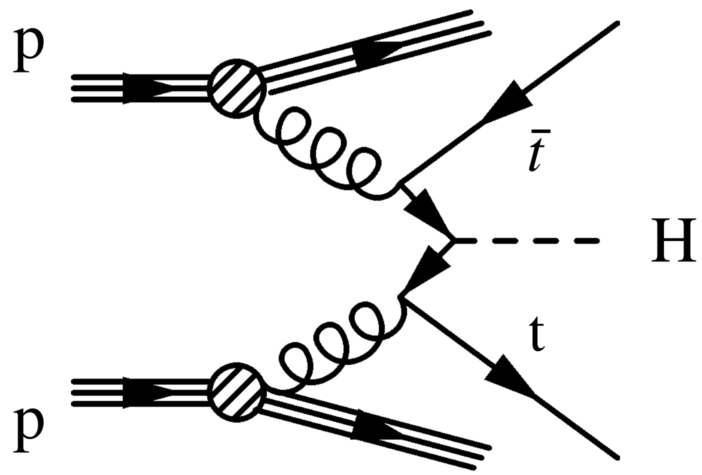


ttH tags

ttH leptonic tag

$t\bar{t} \rightarrow bl\nu_l\bar{b}q\bar{q}^{'}$ $t\bar{t} \rightarrow bl\nu_l\bar{b}l^{'}\nu_{l'}$

- (sub)lead $p_T/m_{\gamma\gamma}$ > 1/2(1/4)
- at least one lepton (ℓ=e,μ), away from Z peak
- ≥ 2 jets
- ≥ 1 b-jet



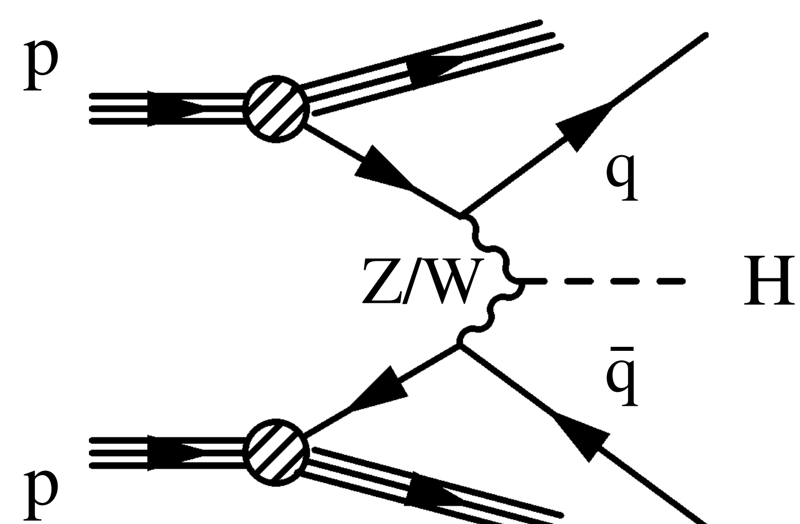
ttH hadronic tag

$t\bar{t} \rightarrow bq\bar{q}^{'}\bar{b}q\bar{q}^{'}$

- (sub)lead $p_T/m_{\gamma\gamma}$ > 1/2(1/4)
- 0 leptons
- ≥ 5 jets
- ≥ 1 b-jet

VBF tags

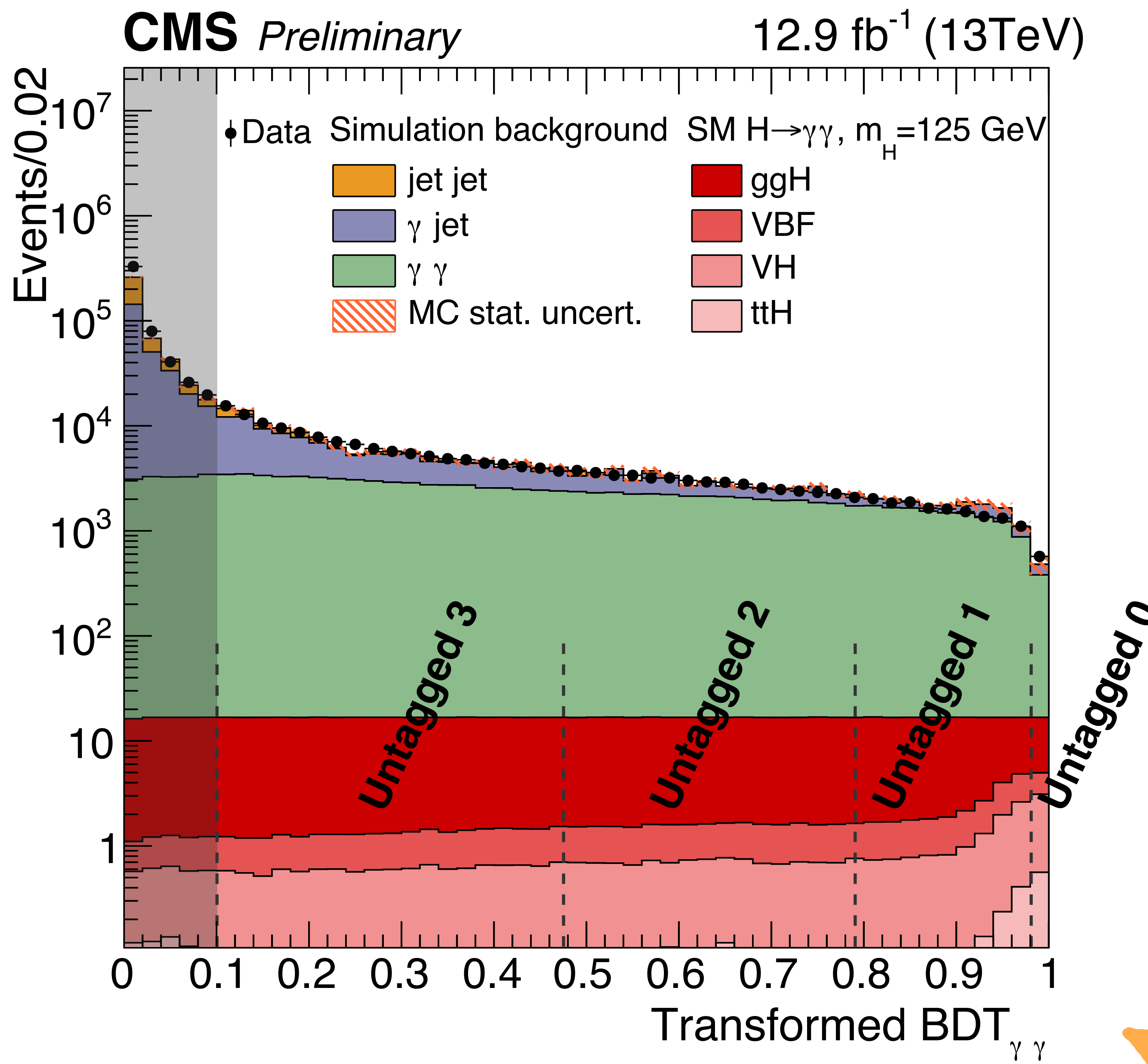
- Identify events with **2 jets** through a MVA
  - inputs: $p_T/m_{\gamma\gamma}$ of both photons, pT of both jets, m_{jj}, $\Delta\eta_{jj}$, Zeppenfeld variable, $\Delta\phi_{\gamma\gamma jj}$
- 2 jets with p_{T1} > 30GeV, p_{T2} >20 GeV, |η| < 4.7, m_{jj} >250 GeV
- **VBF Classification** BDT combines di-jet and diphoton BDT: 2 data categories (**VBF tag 0-1**)

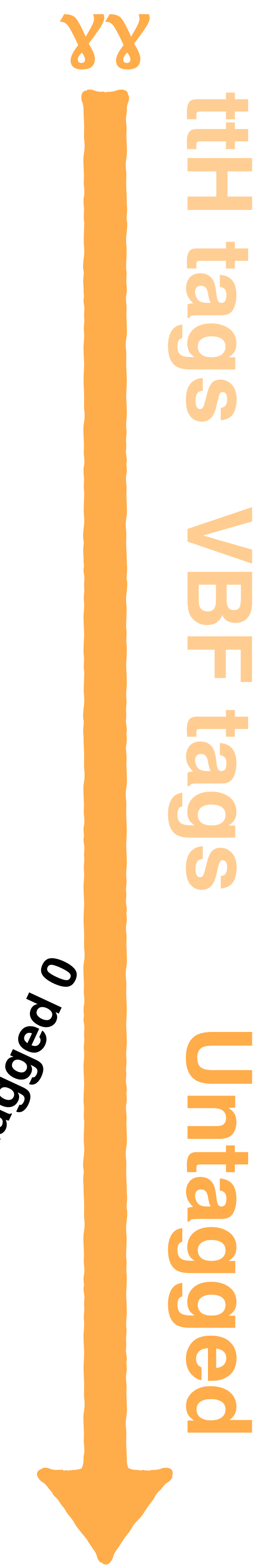


“Untagged” events

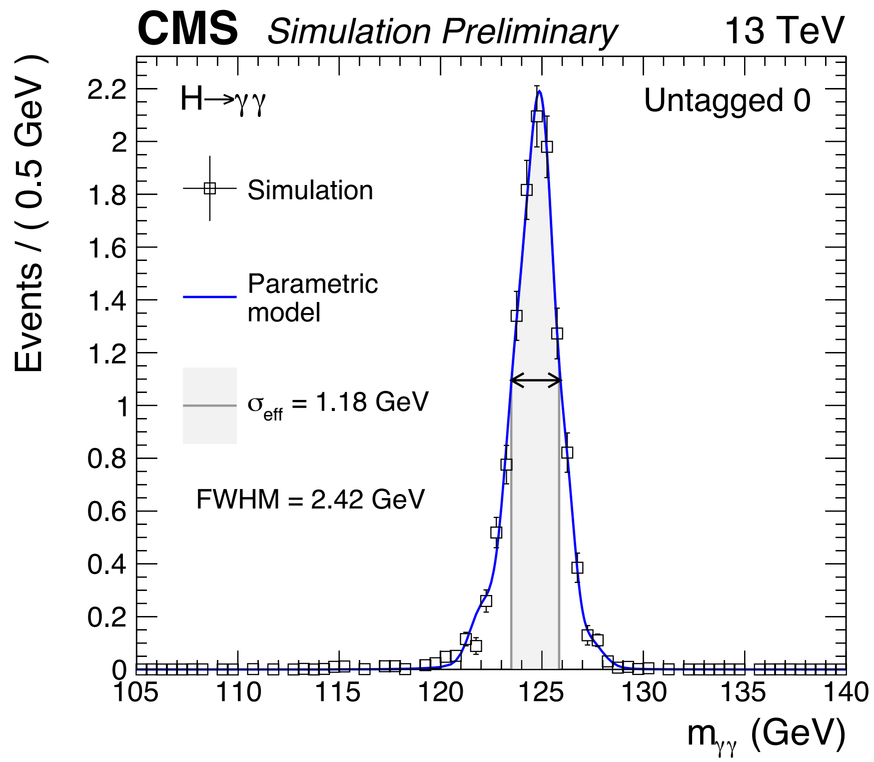
Categorise events into classes (**S/B, resolution**) to improve the analysis sensitivity

- **Multivariate discriminator** is used to separate diphoton pairs with signal-like kinematics, high photon ID scores and good mass resolution from background
- Input variables: pT of the two photons rescaled by pair mass, η, cos(Δϕ), photon ID scores, relative mass resolutions under correct/incorrect vertex selection hypotheses, vertex probability estimate
- 4 event categories (“**Untagged**” events)

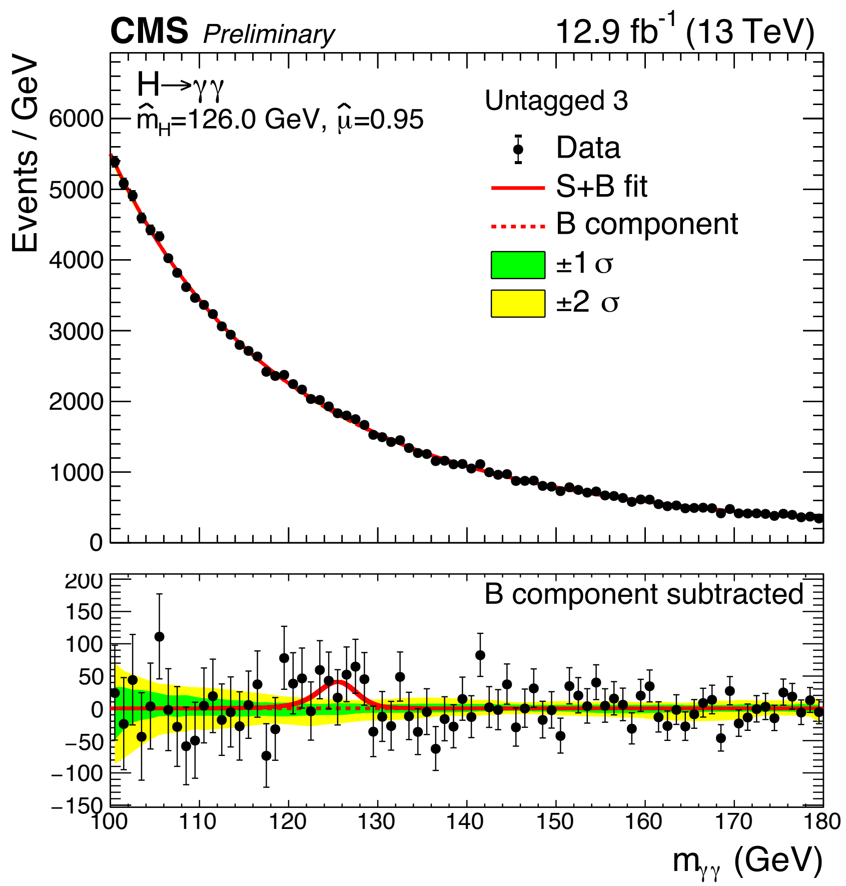




- **Fully parametric** signal model from simulation
    - physical nuisances allowed to float
    - **corrections** and data/MC efficiency **scale factors** applied

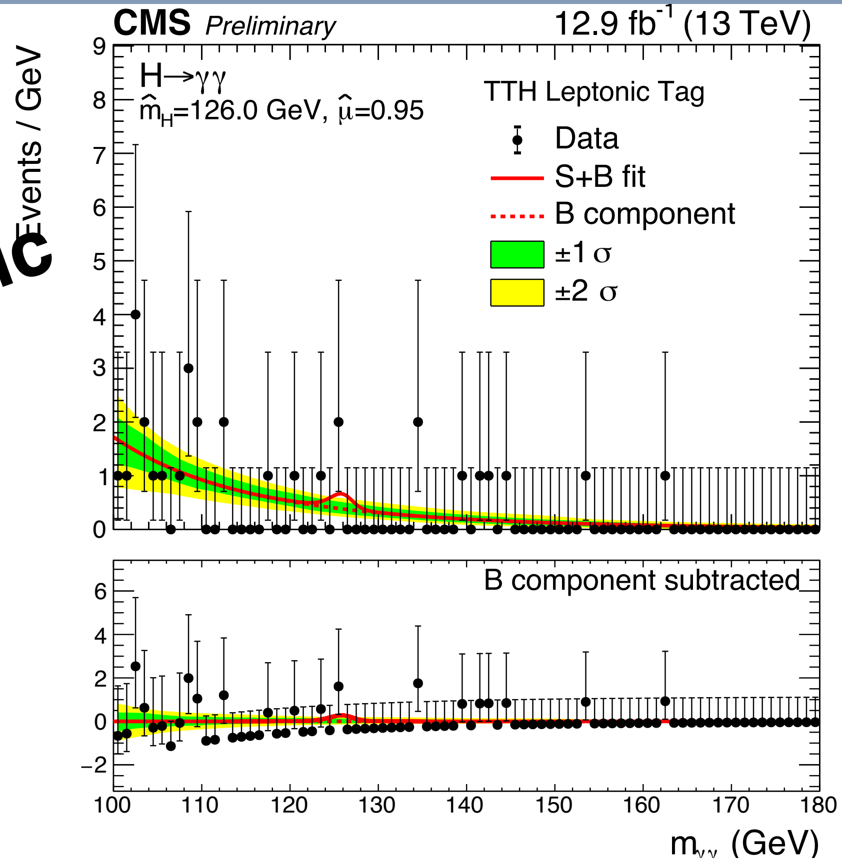


- Background model **data driven**:
    - background functional form treated as **discrete nuisance** parameter
    - for each category, use different functional forms (sums of exponentials, sums of power law terms, Laurent series and polynomials)

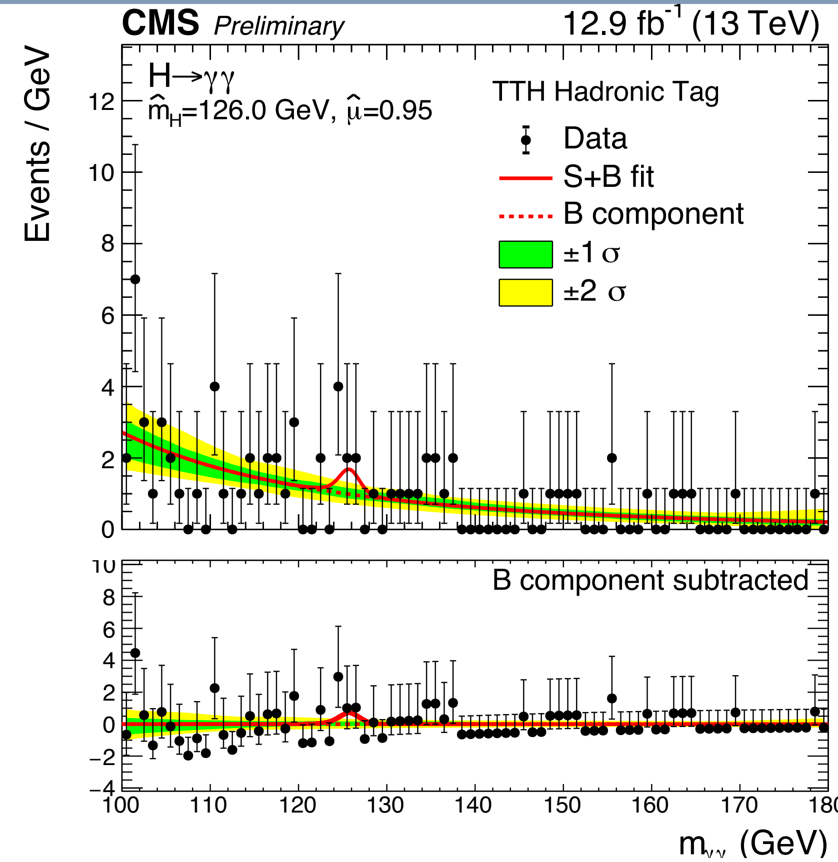


Mass spectra, tagged events

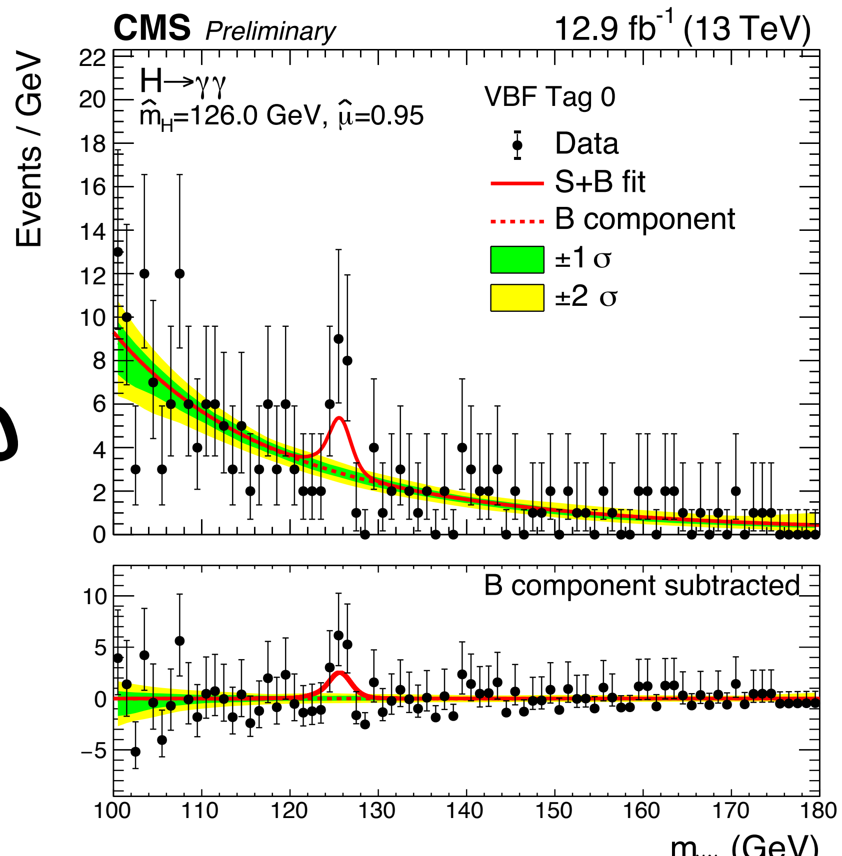
ttH leptonic



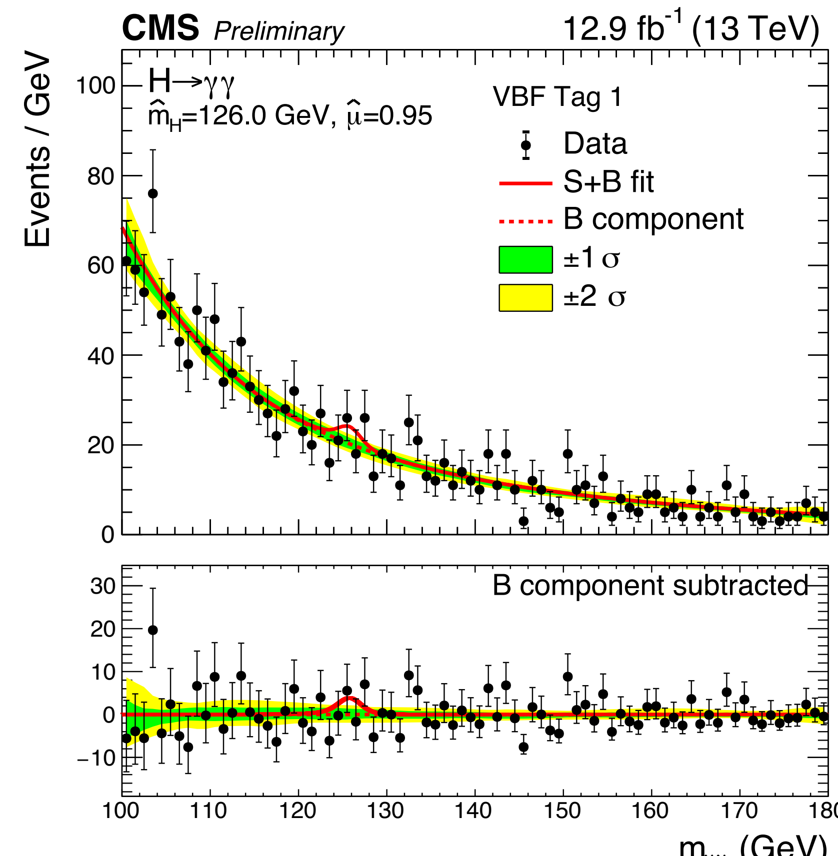
ttH hadronic



VBF tag 0



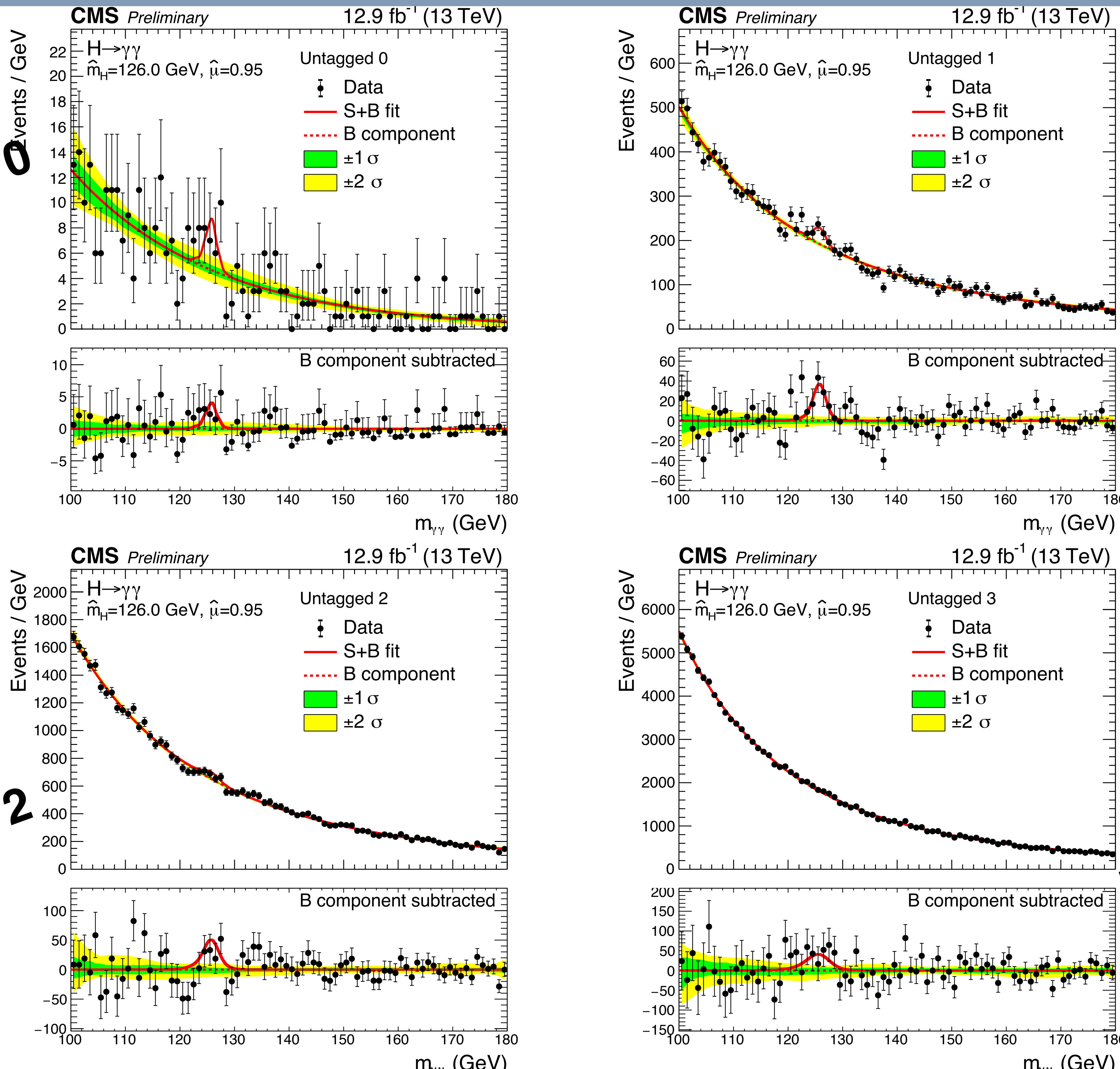
VBF tag 1



Mass spectra, untagged events

Untagged 0

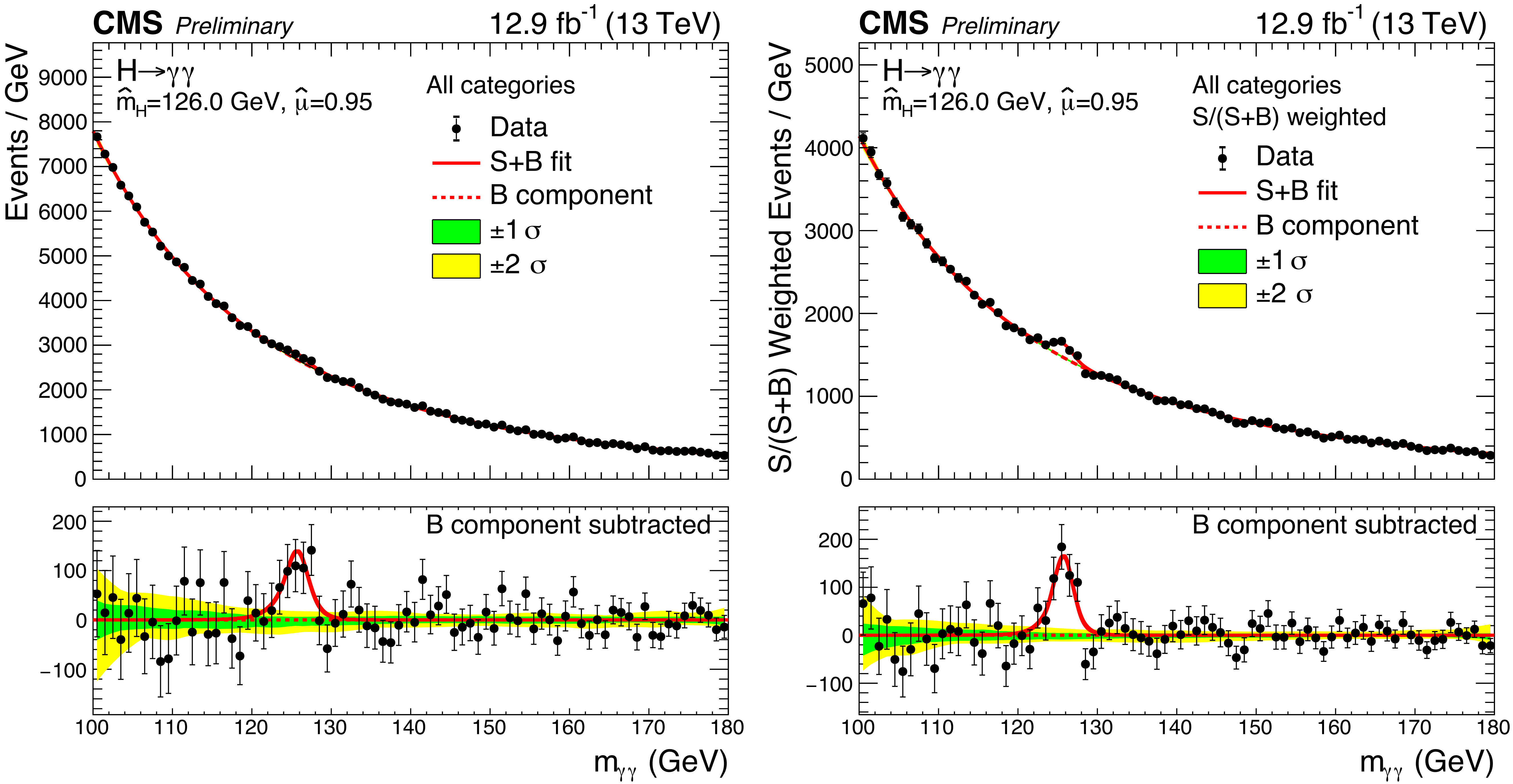
Untagged 1

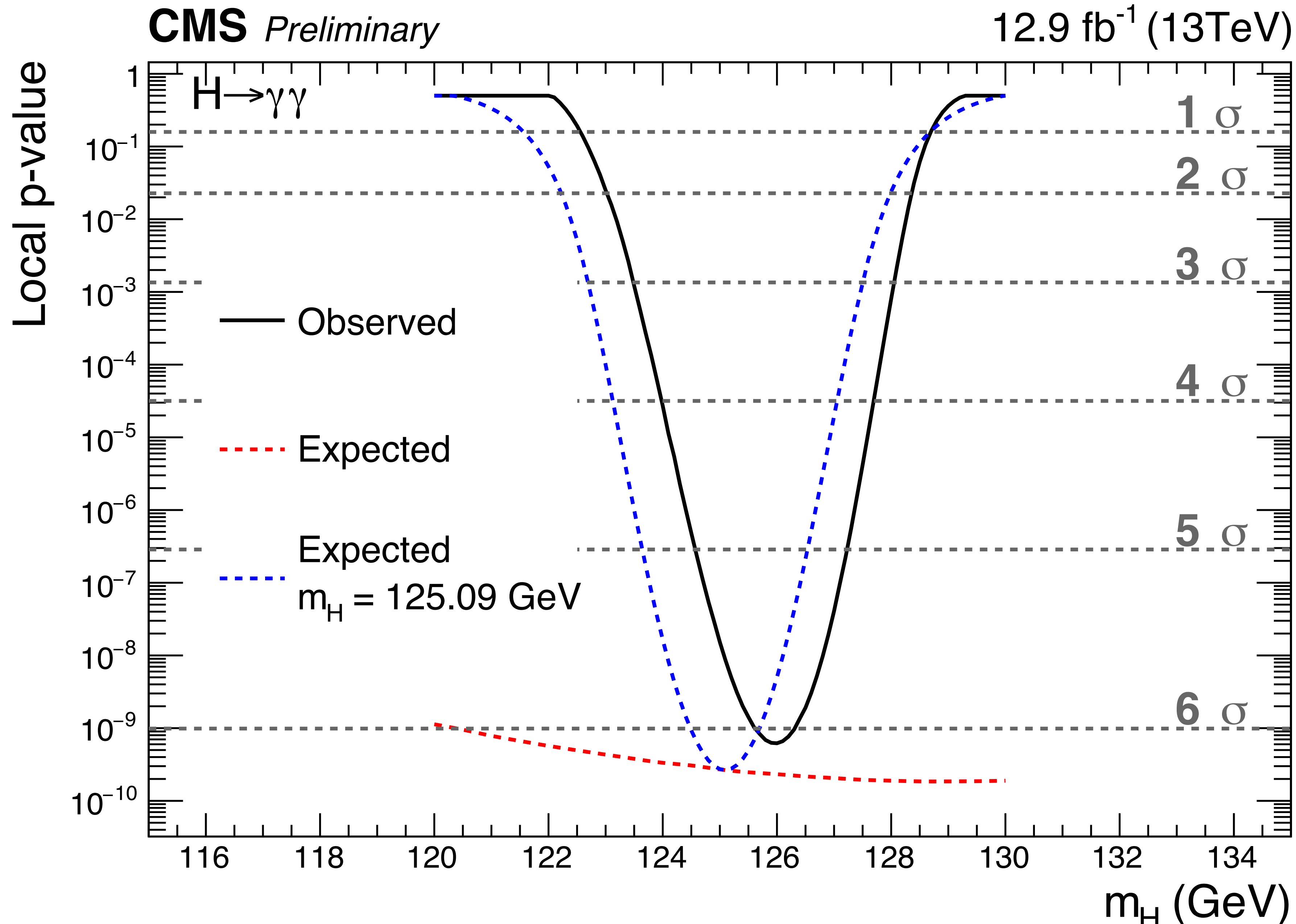


Untagged 2

Untagged 3

Mass spectra, all categories



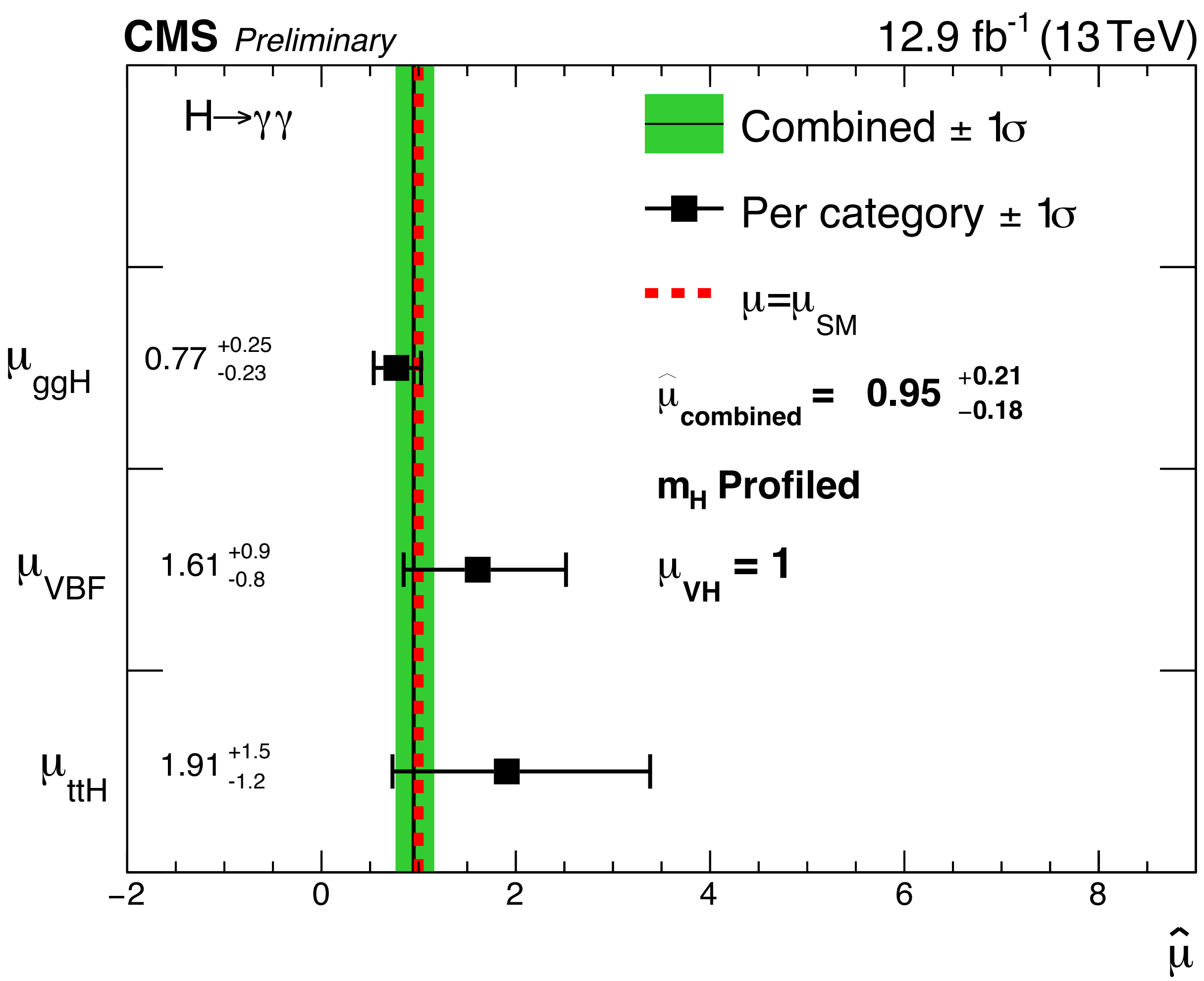


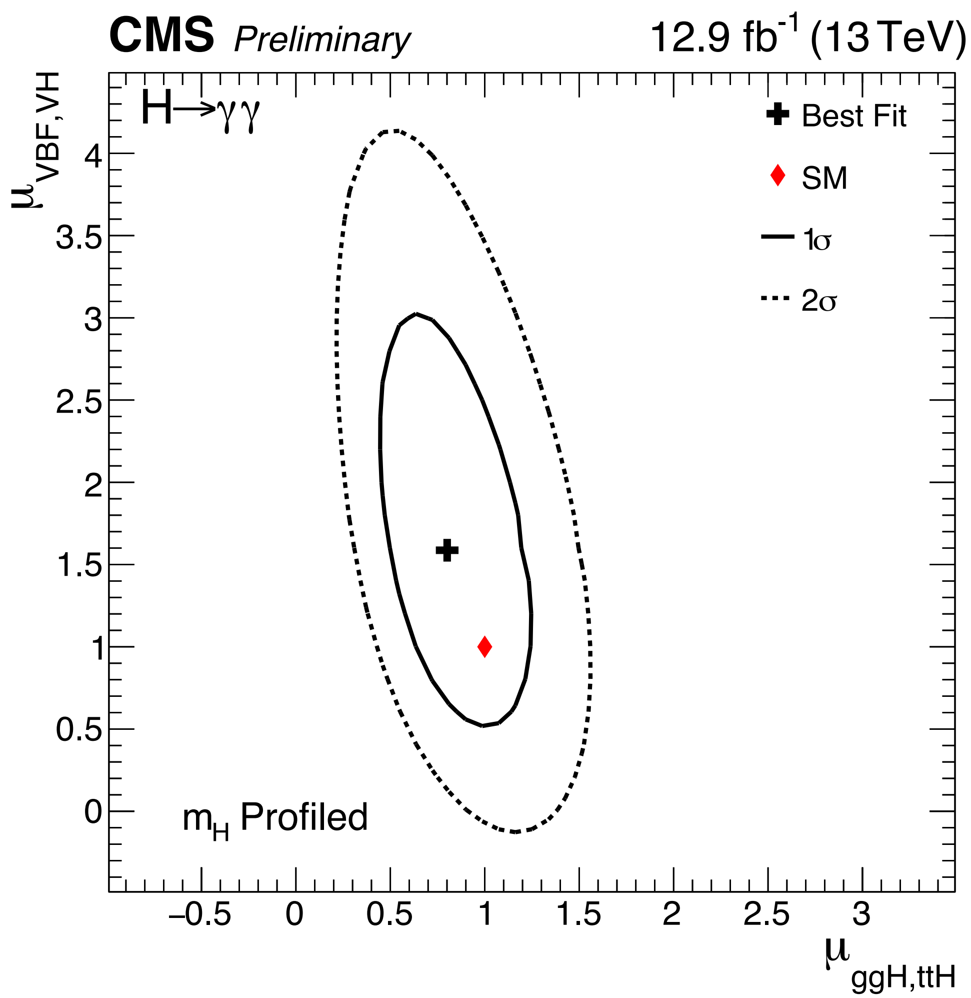
- **Significance** at 125.09 GeV: **5.6σ observed** (6.2σ expected)
- **Maximum observed** significance is **6.1σ** at 126.0 GeV
- Best-fit signal strength $\hat{\sigma}/\sigma_{SM} = 0.95^{+0.21}_{-0.19} = 0.95 \pm 0.17(stat.)^{+0.08}_{-0.05}(theo.)^{+0.10}_{-0.07}(syst.)$

- Signal strength measured is measured in **bosonic and fermionic** components

$$\hat{\mu}_{VBF,VH} = 1.59^{+0.73}_{-0.45}$$

$$\hat{\mu}_{ggH,t\bar{t}H} = 0.80^{+0.14}_{-0.18}$$

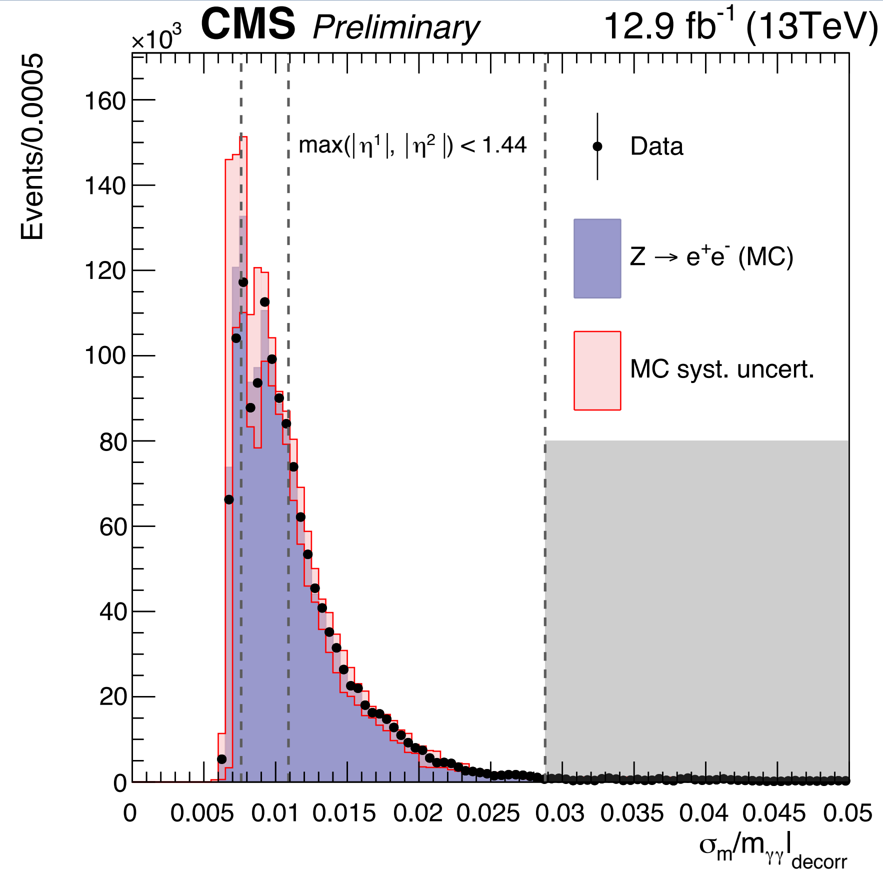




- **Production mechanism** signal strengths measurements **compatible with SM**
- More details on **mass**: see Ulascan Sarica's talk
  "Combined results of the 125 GeV Higgs boson on the mass, tensor structure, and couplings measured by the CMS detector", **today@1:10pm**

Fiducial cross section

- **Different event categorisation**: 3 mass resolution categories

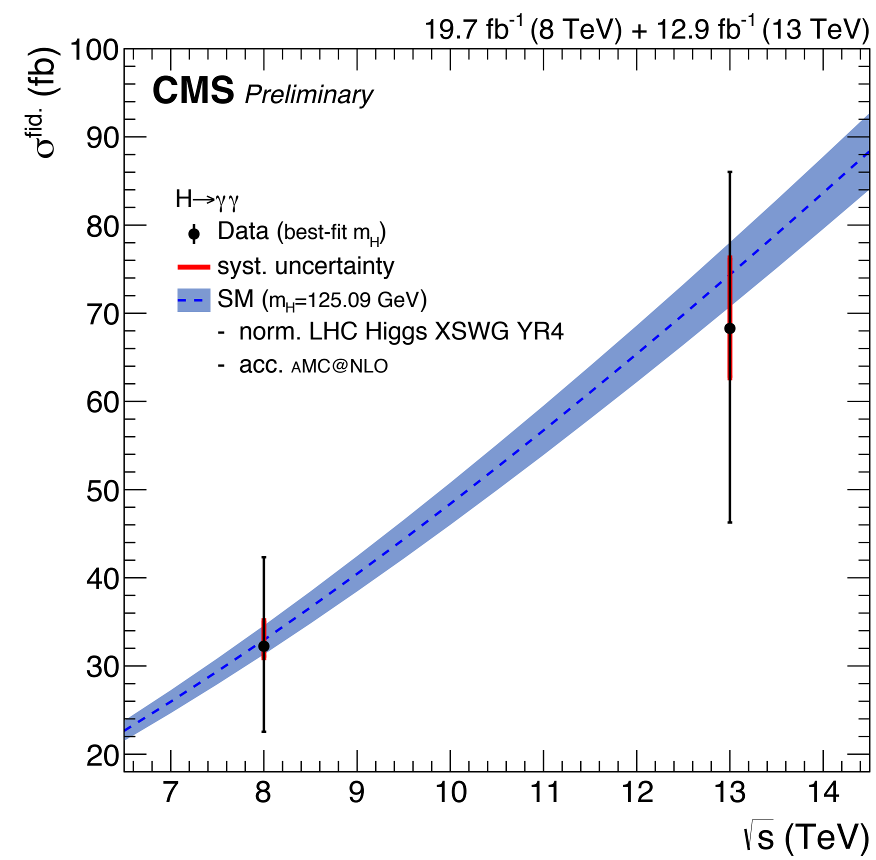


- Fiducial cross section measured profiling m_H:

$$\hat{\sigma}_{fid} = 69^{+16}_{-22}(\text{stat.})^{+8}_{-6}(\text{syst.})\text{fb}$$

- Theoretical prediction for m_H=125.09 GeV

$$\sigma^{th.}_{fid} = 73.8 \pm 3.8\text{fb}$$



Summary

- CMS **H→γγ** results using **12.9/fb of 13 TeV** collision data collected in **2016** have been presented
- Observation (**6.1σ** peak significance) of the Higgs boson in the **di-photon** channel
- Best fit signal strength is $\hat{\sigma}/\sigma_{SM} = 0.95^{+0.21}_{-0.19}$
- **Bosonic and fermionic** components of signal strength are observed

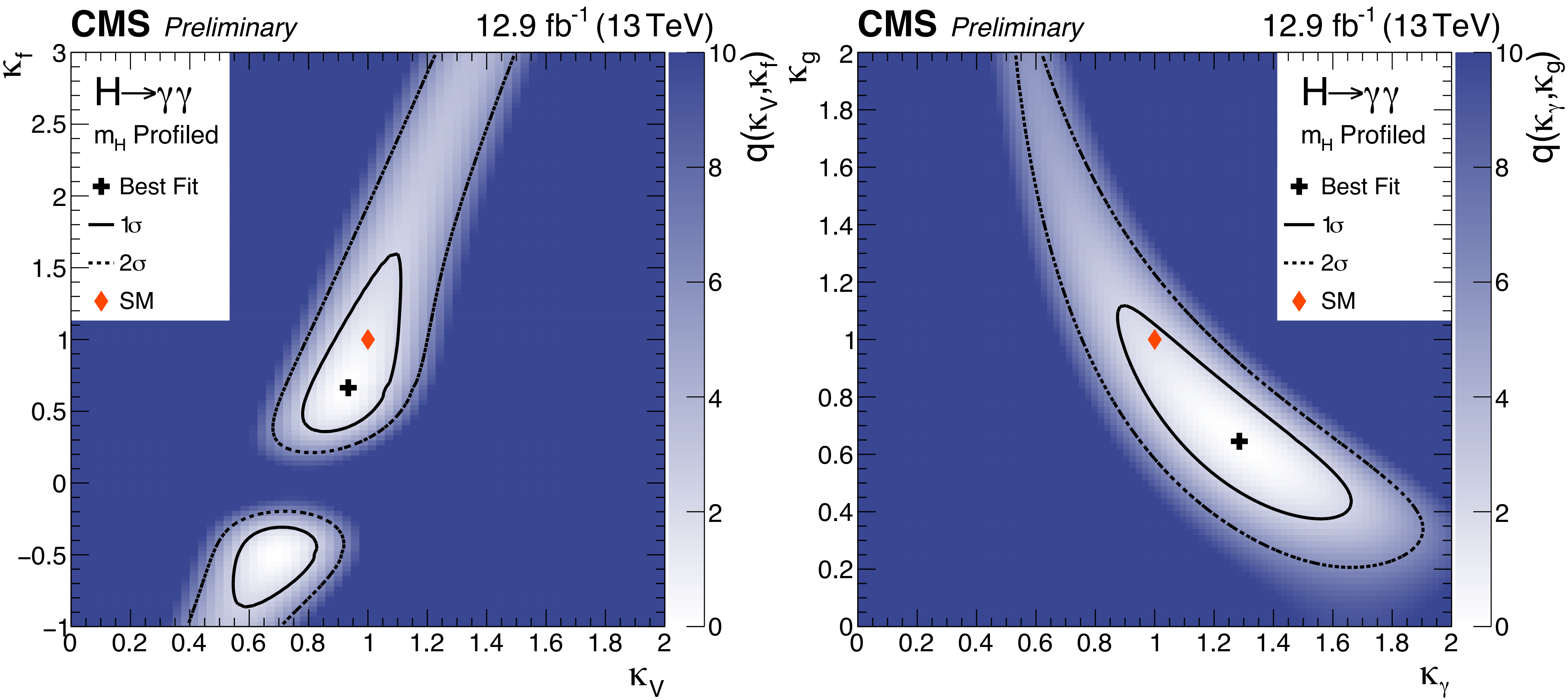
  $\hat{\mu}_{VBF,VH} = 1.59^{+0.73}_{-0.45}$ $\hat{\mu}_{ggH,t\bar{t}H} = 0.80^{+0.14}_{-0.18}$
- **Fiducial cross-section** is measured to be $\hat{\sigma}_{fid} = 69^{+16}_{-22}(\text{stat.})^{+8}_{-6}(\text{syst.})\text{fb}$

BACKUP

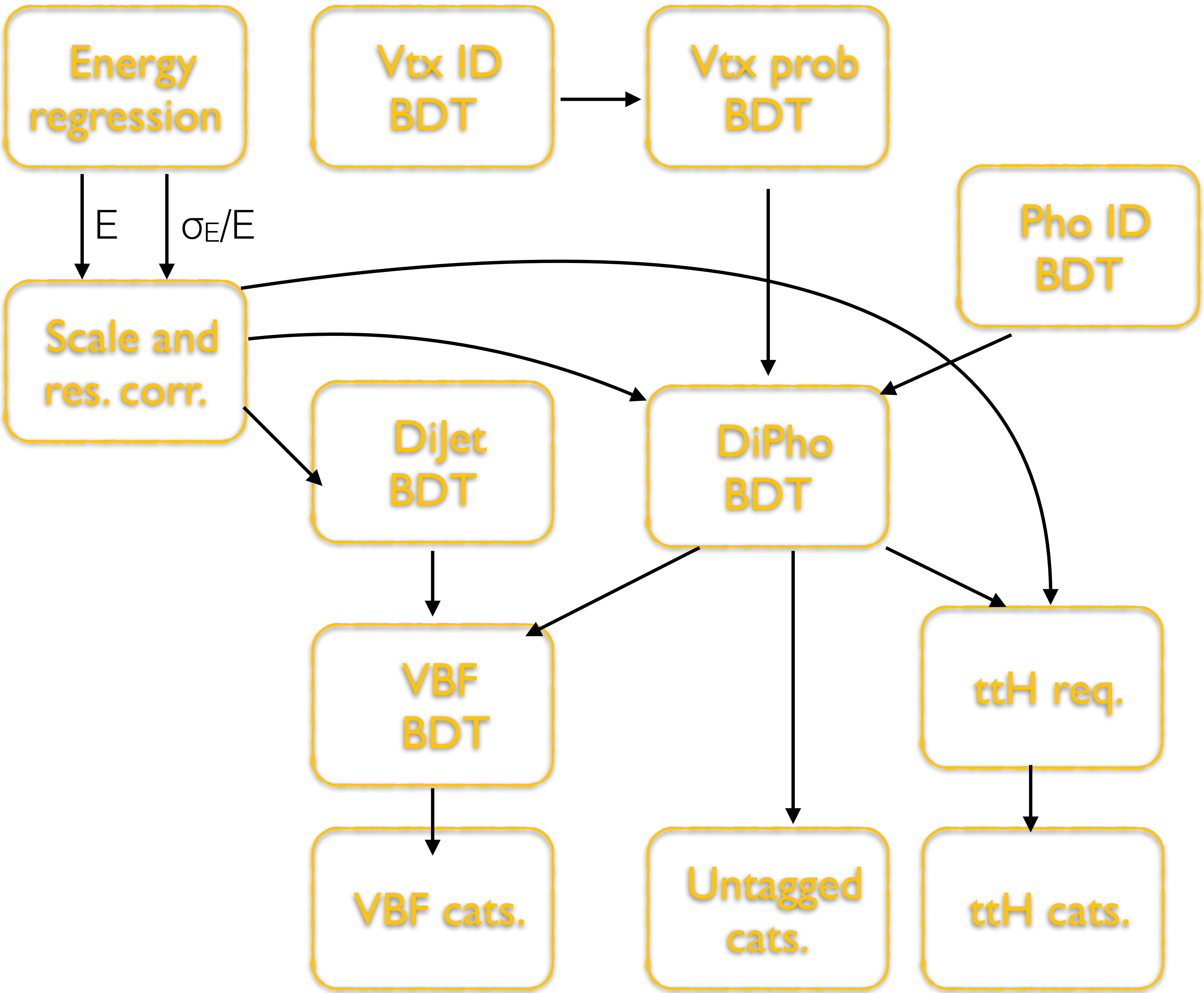
Event yields per category

Event Categories	SM 125GeV Higgs boson expected signal								Bkg (GeV^{-1})
	Total	ggh	vbf	wh	zh	tth	σ_{eff}	σ_{HM}	
Untagged Tag 0	11.92	79.10 %	7.60 %	7.11 %	3.59 %	2.60 %	1.18	1.03	4.98
Untagged Tag 1	128.78	85.98 %	7.38 %	3.70 %	2.12 %	0.82 %	1.35	1.20	199.14
Untagged Tag 2	220.12	91.11 %	5.01 %	2.18 %	1.23 %	0.47 %	1.70	1.47	670.44
Untagged Tag 3	258.50	92.35 %	4.23 %	1.89 %	1.06 %	0.47 %	2.44	2.17	1861.23
VBF Tag 0	9.35	29.47 %	69.97 %	0.29 %	0.07 %	0.20 %	1.60	1.33	3.09
VBF Tag 1	15.55	44.91 %	53.50 %	0.86 %	0.38 %	0.35 %	1.71	1.40	22.22
TTH Hadronic Tag	2.42	16.78 %	1.28 %	2.52 %	2.39 %	77.02 %	1.39	1.21	1.12
TTH Leptonic Tag	1.12	1.09 %	0.08 %	2.43 %	1.06 %	95.34 %	1.61	1.35	0.42
Total	647.77	87.93 %	7.29 %	2.40 %	1.35 %	1.03 %	1.88	1.52	2762.65

- Measurement of coupling modifiers to vector bosons and fermions (k_V, k_f) and to photons and gluons (k_γ,k_g)



Analysis flowchart



Vertex BDTs inputs

- Vertex ID BDT:

$$\sum_i |\vec{p}_T^{\,i}|^2 \text{ , -}\sum_i (\vec{p}_T^{\,i} \cdot \frac{\vec{p}_T^{\,\gamma\gamma}}{|\vec{p}_T^{\,\gamma\gamma}|}) \text{ and } (|\sum_i \vec{p}_T^{\,i}| - p_{\mathrm{T}}^{\gamma\gamma}) / (|\sum_i \vec{p}_T^{\,i}| + p_{\mathrm{T}}^{\gamma\gamma})$$

- if conversions are present

the number of conversions and the pull $|z_{\mathrm{vtx}} - z_{\mathrm{e}}|/\sigma_z$ between the longitudinal position of the reconstructed vertex, z_{vtx}, and the longitudinal position of the vertex estimated using conversion track(s), z_{e}. The variable

- Vertex probability BDT
  - the number of vertices in each event,
  - the values of the $\mathrm{BDT}_{\mathrm{VTX\,ID}}$ output for the three most likely vertices in each event,
  - the distances between the chosen vertex and the second and third choices,
  - the transverse momentum of the diphoton system ($p_{\mathrm{T}}^{\gamma\gamma}$),
  - the number of photons with an associated conversion track.

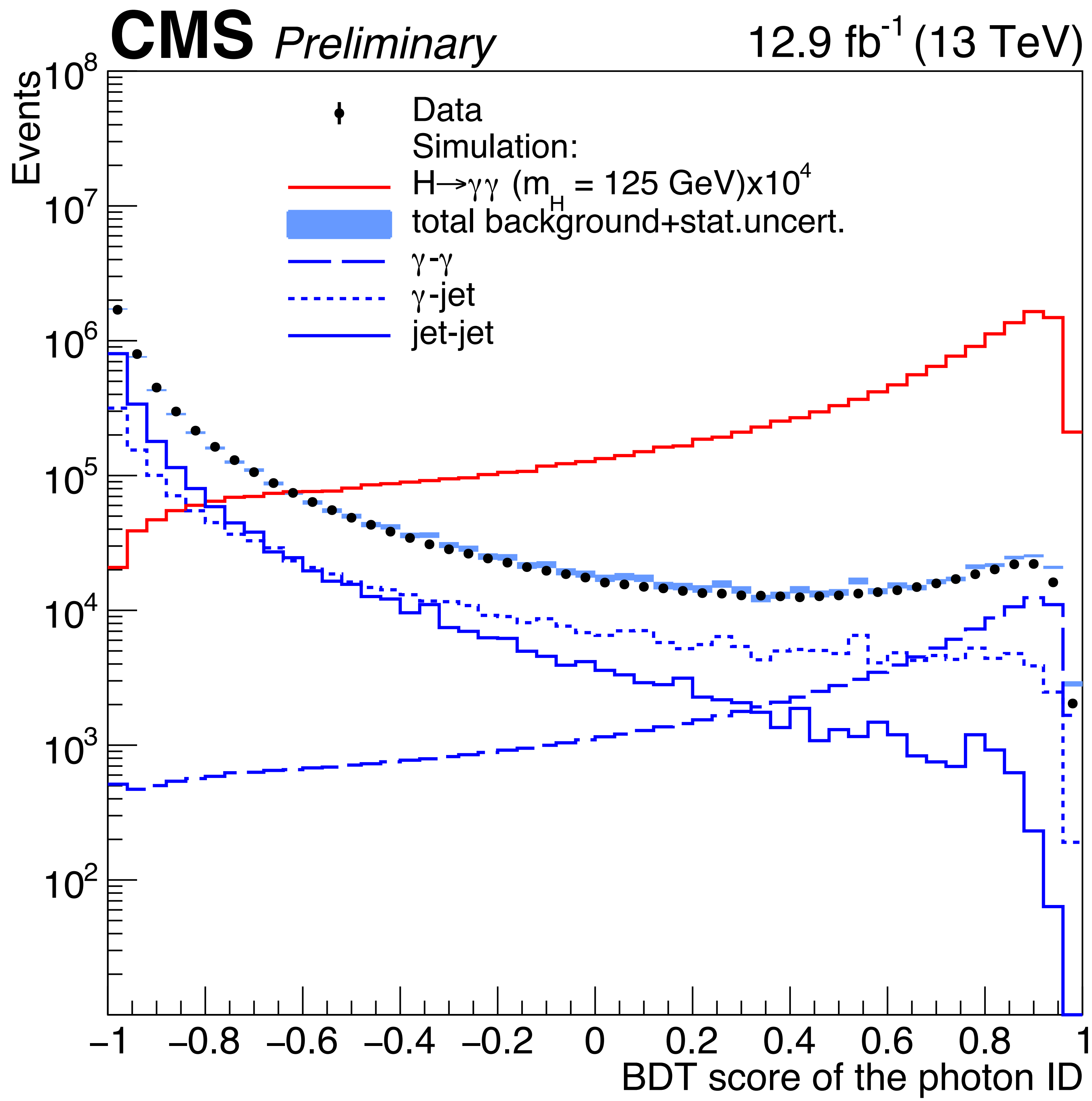
Preselection details

- p_T>30 (20) GeV, $p_T/m_{\gamma\gamma}$ > 1/3 (1/4) for (sub)leading-p_T photon
- $|\eta|$<2.5, removing 1.44<$|\eta|$<1.57, electron veto
- either R_9>0.8, or charged hadron isolation < 20 GeV, or charged hadron isolation relative to p_T <0.3

	H/E	$\sigma_{\eta\eta}$	R_9	photon iso.	tracker iso.
ECAL barrel; $R_9 > 0.85$	<0.08	–	>0.5	–	–
ECAL barrel; $R_9 \leq 0.85$	<0.08	<0.015	>0.5	< 4.0	< 6.0
ECAL endcaps; $R_9 > 0.90$	<0.08	–	>0.8	–	–
ECAL endcaps; $R_9 \leq 0.90$	<0.08	<0.035	>0.8	< 4.0	< 6.0

$BDT_{\gamma\ ID}$ inputs

- shower shape observables;
- isolation variables based on the sums of the p_T of photons and of charged hadrons, within regions of $R < 0.3$ around the candidate. Two such charged hadron isolations are used: one considering hadrons coming from the chosen vertex in the event, and one considering hadrons coming from the vertex with the largest p_T sum among remaining vertices. The latter is effective in rejecting photon candidates originating from mis-identification of jets from a vertex other than the chosen one;
- the median energy density per unit area in the event, ρ, in order for the $BDT_{\gamma\ ID}$ to be independent of pileup;
- photon kinematic observables (pseudorapidity and energy), allowing the $BDT_{\gamma\ ID}$ to account for the dependence of the shower topology and isolation variables on η and p_T.



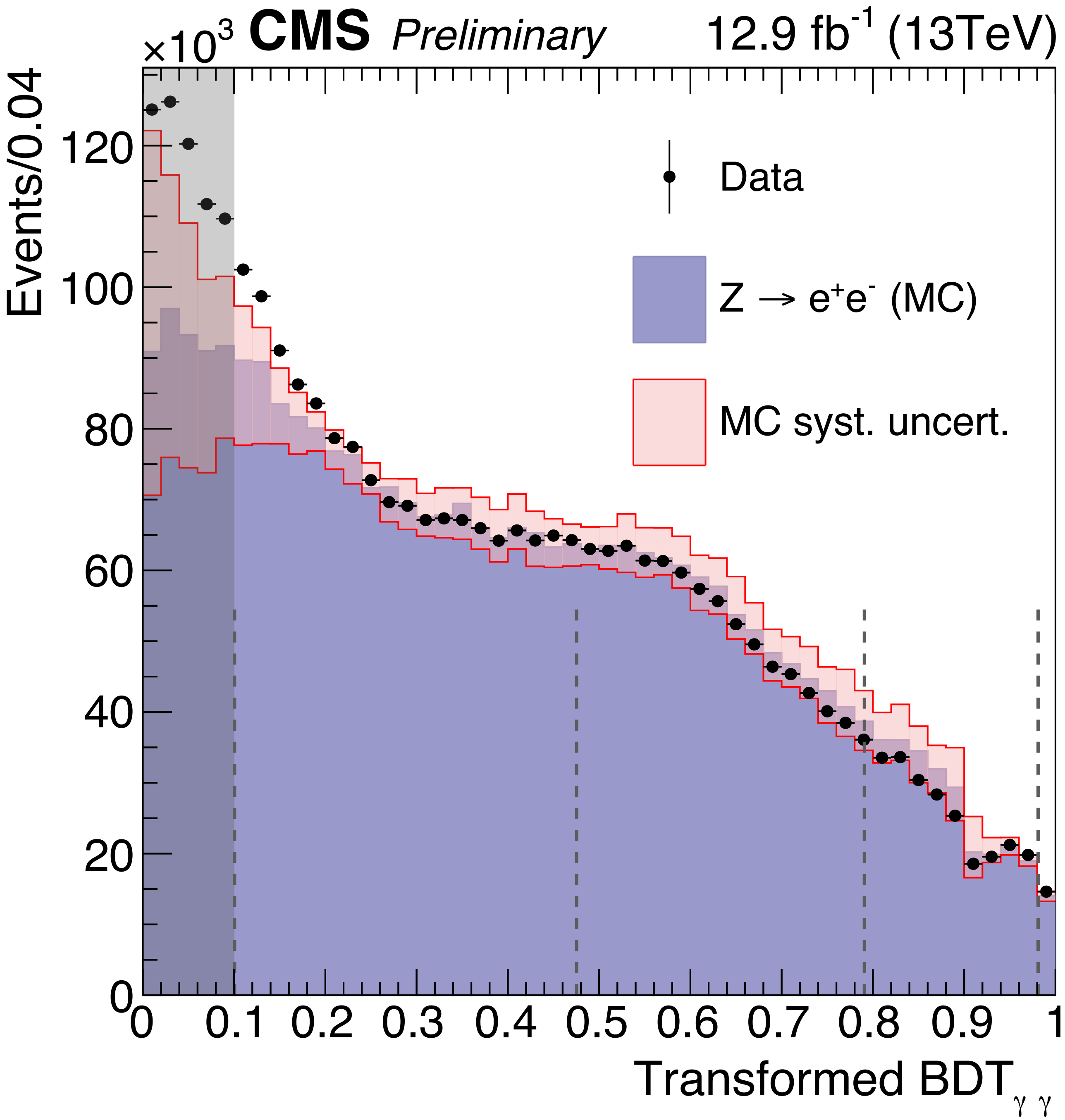
- BDTγ ID of the lowest-scoring photon for data and simulation, after preselection

$BDT_{\gamma\gamma}$ inputs

- the relative transverse momenta of both photons, $p_T^{1(2)}/m_{\gamma\gamma}$;
- the pseudorapidities of both photons, $\eta^{1(2)}$;
- the cosine of the angle between the two photons in the transverse plane, $\cos(\phi_1 - \phi_2)$;
- the relative diphoton mass resolution, under the hypothesis that the mass has been reconstructed using the correct primary vertex, $\sigma_m^{right}/m_{\gamma\gamma}$;
- the relative diphoton mass resolution, under the hypothesis that the mass has been reconstructed using an incorrect primary vertex, $\sigma_m^{wrong}/m_{\gamma\gamma}$;
- the per-event probability estimate that the correct primary vertex has been used to reconstruct the mass, taken from $BDT_{VTX\ PROB}$;
- the $BDT_{\gamma\ ID}$ score for both photons.

BDT$_{\gamma\gamma}$ systematic uncertainties

- Impact on BDT$_{\gamma\gamma}$ of systematic uncertainties associated to:
  - relative energy resolution (±5% relative shift)
  - photon ID BDT (± 0.03 shift plus linearly increasing term)
- Z→ee events with electrons reconstructed as photons



ttH leptonic tag details

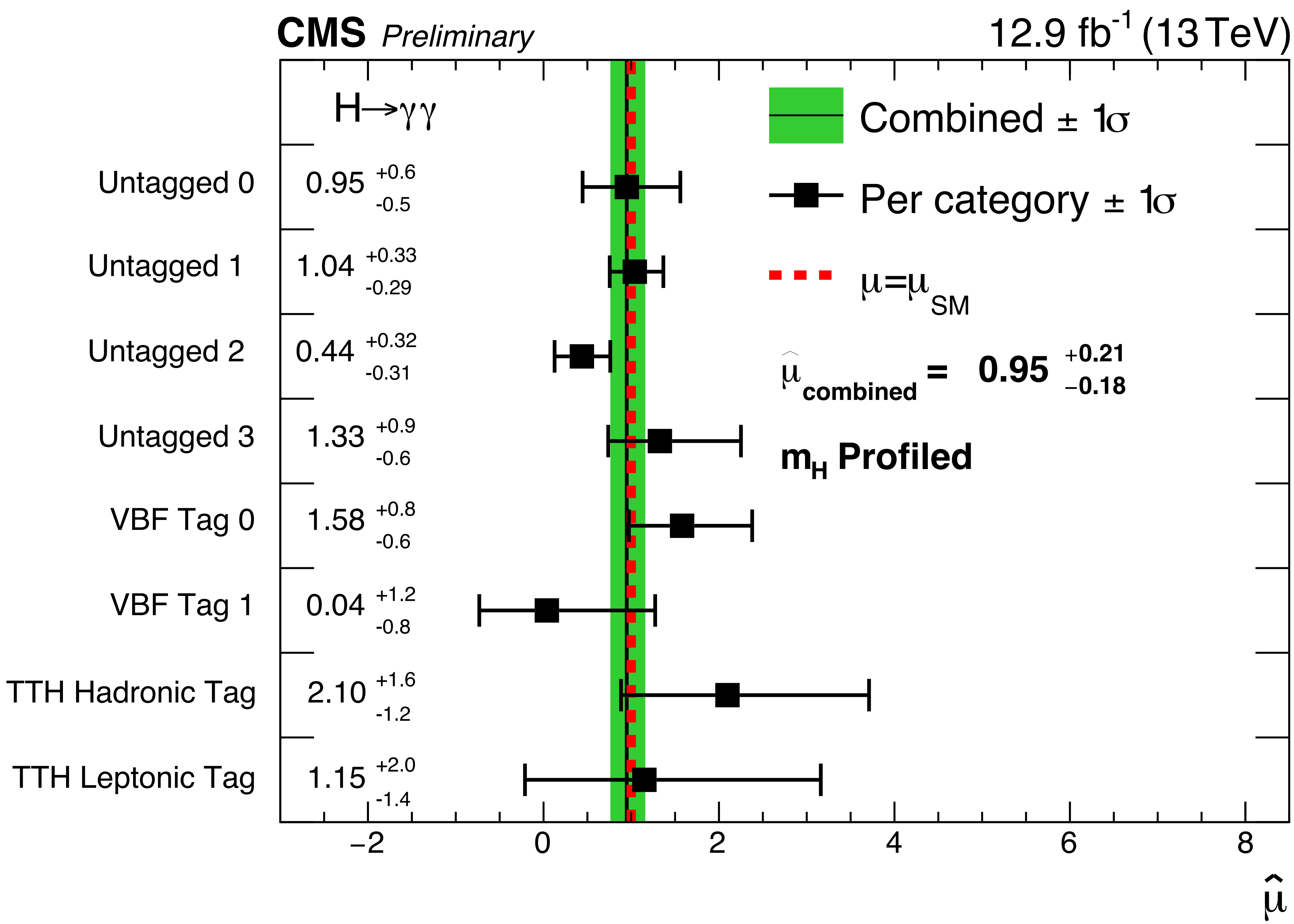
- leading photon $p_T > m_{\gamma\gamma}/2$;
- sub-leading photon $p_T > m_{\gamma\gamma}/4$;
- at least one lepton with $p_T > 20$ GeV: electrons must be within the ECAL fiducial region and pass the recommended criteria for loose requirements on the same observables as described in [27]. In addition the electron should satisfy $|m(e,\gamma) - m_Z| > 10$ GeV, where m_Z refers to the Z boson mass. Muons are required to have $|\eta| < 2.4$ and to pass a tight selection based on the quality of the track, the number of hits in the tracker and muon system, and the longitudinal and transverse impact parameters of the track with respect to the muon vertex; additionally, it has to satisfy a requirement on the relative isolation with pileup correction, based on the transverse momentum of the charged hadrons, transverse energy of the neutral hadrons and photons in a cone of $R = 0.4$ around the muon;
- all selected leptons (ℓ) are required to have $\Delta R(\ell, \gamma) > 0.4$;
- at least two jets in the event with $p_T > 25$ GeV, $|\eta| < 2.4$, and $\Delta R(\text{jet}, \gamma) > 0.4$ and $\Delta R(\text{jet}, \ell) > 0.4$;
- at least one of the jets in the event has to be identified as b jet according to the CSV tagger medium requirement [28].
- $\text{BDT}_{\gamma\gamma}$ output > -0.4. Too few events are available to optimise this selection for significance, so this choice is made simply to remove most of the events with low $\text{BDT}_{\gamma\gamma}$ score.

ttH hadronic tag details

- leading photon $p_{\mathrm{T}} > m_{\gamma\gamma}/2$;
- sub-leading photon $p_{\mathrm{T}} > m_{\gamma\gamma}/4$;
- no leptons defined according to the leptonic tag;
- at least five jets in the event with $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.4$;
- at least one of the jets in the event has to be identified as a b-jet according to the CSV tagger medium requirement [28];
- a minimum value of $\mathrm{BDT}_{\gamma\gamma}$ output. The value is a compromise between significance optimisation and the need of a minimum number of events to fit the background.

Systematic uncertainties

- **Theory uncertainties** (PDFs, α_S, QCD scale, underlying event and parton shower, $H \rightarrow \gamma\gamma$ branching fraction)
- **ggH contamination** in VBF and ttH tagged categories
- **Trigger** efficiency, **integrated luminosity**, **vertex** efficiency, **preselection**
- Non-unformity of light collection, non-linearity, detector simulation, modeling of the material budget, shower shape corrections
- Photon energy **scale and resolution**
- **$BDT_{\gamma\ ID}$** and per-photon **energy resolution**
- **Jet** energy scale and smearings
- **b-tagging** efficiency, **gluon-splitting** fraction, **parton shower**, ID efficiency for **e and μ**


CMS Preliminary
12.9 fb-1 (13 TeV)
H→γγ
Combined ± 1σ
Per category ± 1σ
μ=μSM
μ̂ combined = 0.95 +0.21 −0.18
mH Profiled
Untagged 0
0.95 +0.6 -0.5
Untagged 1
1.04 +0.33 -0.29
Untagged 2
0.44 +0.32 -0.31
Untagged 3
1.33 +0.9 -0.6
VBF Tag 0
1.58 +0.8 -0.6
VBF Tag 1
0.04 +1.2 -0.8
TTH Hadronic Tag
2.10 +1.6 -1.2
TTH Leptonic Tag
1.15 +2.0 -1.4
−2
0
2
4
6
8
μ̂

Fiducial cross section: mass spectra