Search for a SM-like low-mass Higgs boson in the $\gamma\gamma$ final state at CMS

Junquan Tao (IHEP/CAS)
陶军全 （中科院高能所）
on behalf of the CMS collaboration

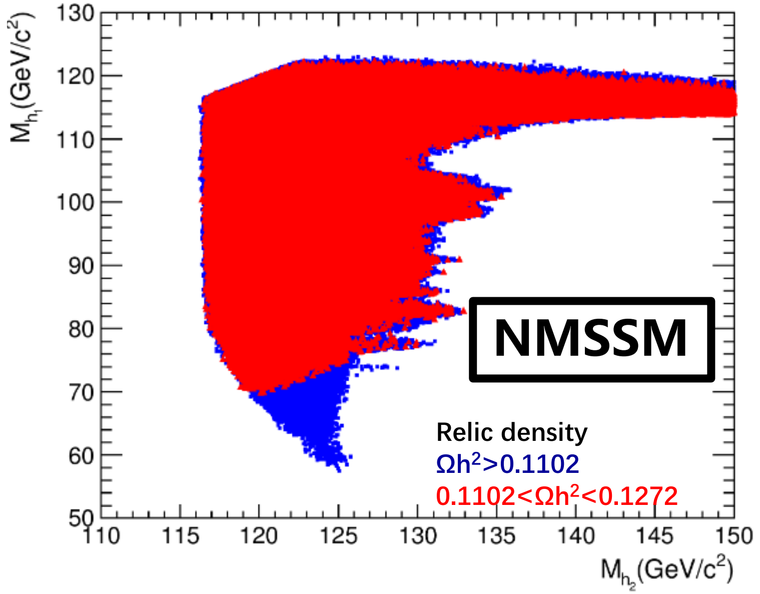
Based on CMS-PAS-HIG-20-002

Institute of High Energy Physics
Chinese Academy of Sciences

Introduction

- Although the Higgs boson discovered at LHC so for is compatible with the SM Higgs boson, there is still **room for extensions of the Standard Model**
- **BSM models :** H(125GeV) + additional Higgs bosons with some of which could have masses below 125 GeV
  - ✓ **2HDM** : 2 CP-even (h, H), 1 CP-odd (a), 2 charged ($H^{\pm}$)
  - ✓ **NMSSM** : 3 CP-even (h_i), 2 CP-odd (a_i), 2 charged ($h^{\pm}$)
  - ✓ Georgi-Machacek(**GM**) model : 2 singlet(h, H), 3 triplet(H_3^0, $H_3^{\pm}$) and 5 fiveplet (H_5^0, $H_5^{\pm}$, $\mathbf{H_5^{\pm\pm}}$) mass eigenstates
  - ✓ **Other extensions** e.g. left-right symmetric models (**LRSMs**), Three Higgs Doublet Model (**3HDM**), ….

J. Fan, JT(*), et al., Chinese Phys. C 38 (2013) 073101

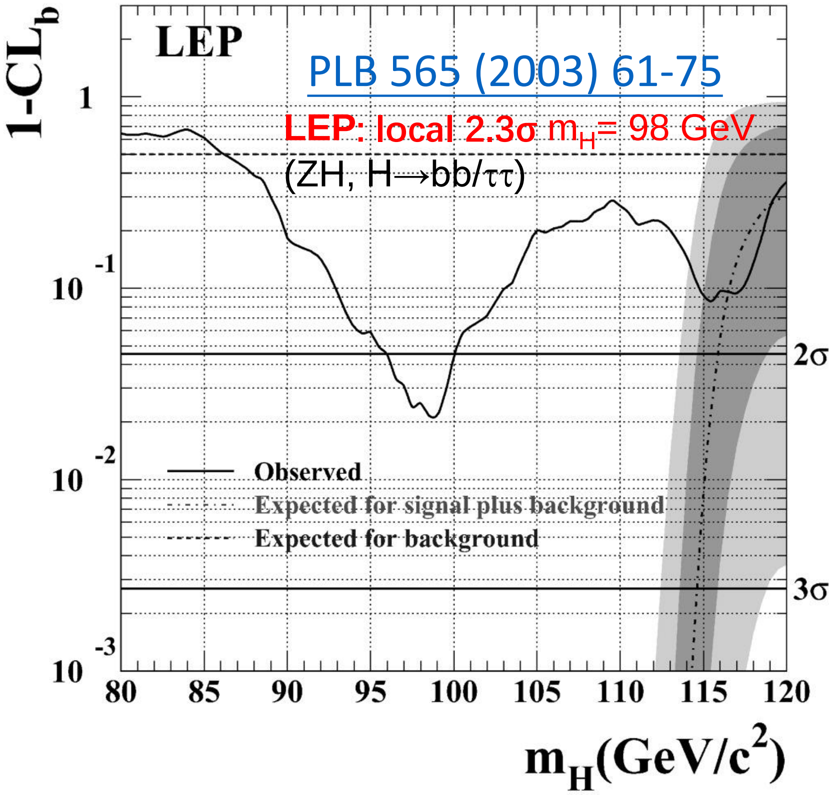


- **Discovery of extra Higgs boson(s) would be an unequivocal sign of new physics**
- LHC is currently the most powerful discovery machine
  - Hope to find hints of BSM

Experimental searches in low-mass region

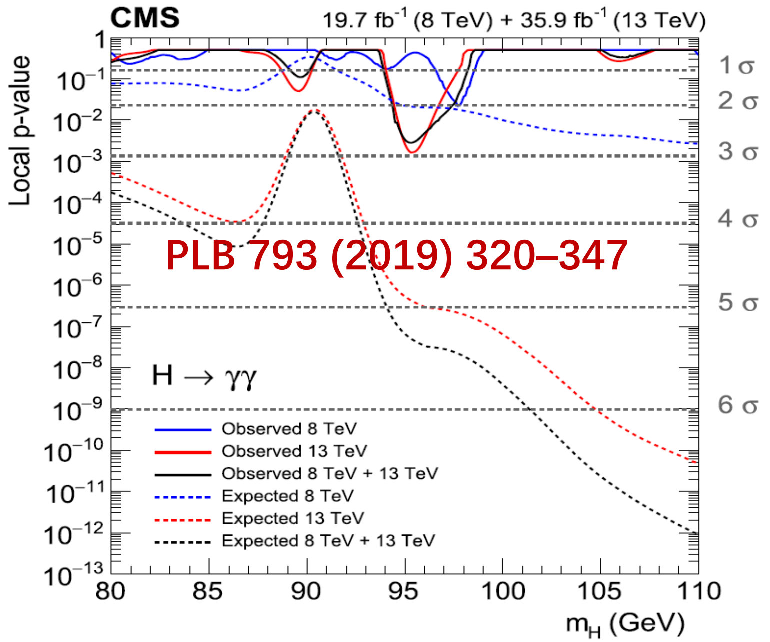
Institute of High Energy Physics, Chinese Academy of Sciences

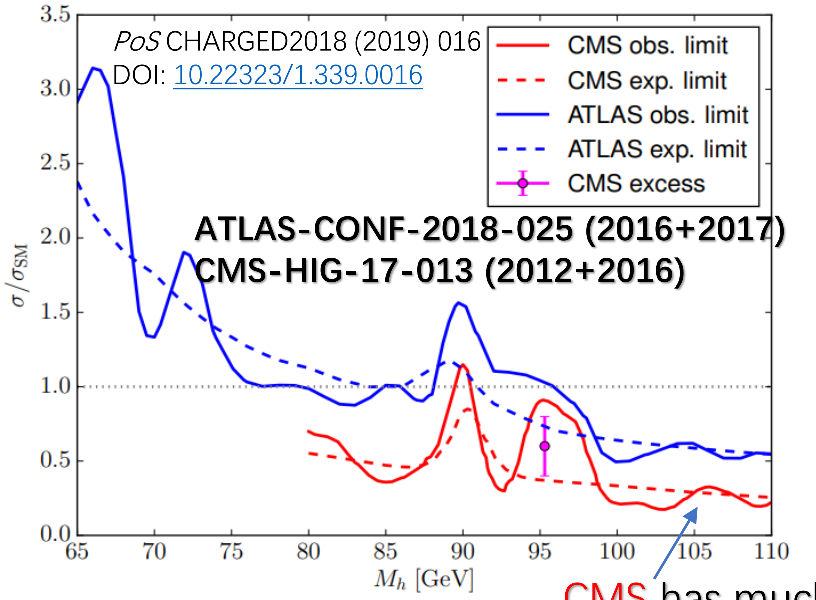
- Final LEP SM Higgs boson search results: >2σ excess at m_H= 98 GeV

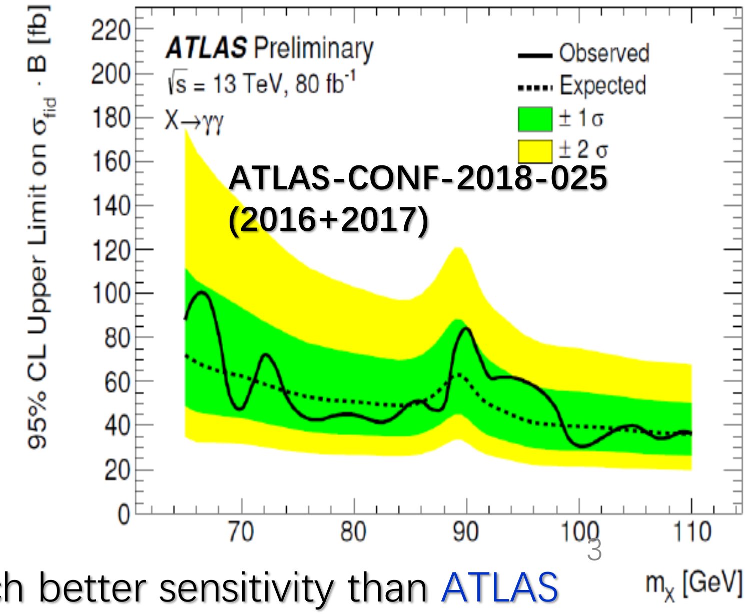


Some experimental results before 2023

- We have performed the LM H→γγ search **since Run1** at CMS
  - ✓ 2012 data (**HIG-14-037**, PAS only) : ~2σ local at 97.5 GeV
  - ✓ 2016 + 2012 data (**HIG-17-013, PLB 793 (2019) 320–347**): 2.8σ local (1.3σ global) significance at 95.3 GeV
- ATLAS LM h→γγ with 80 fb^{-1}: not exclude the CMS observed excess



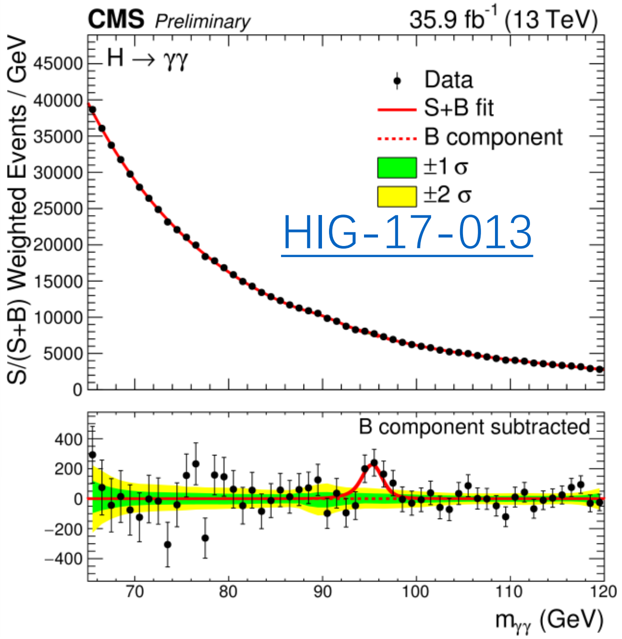


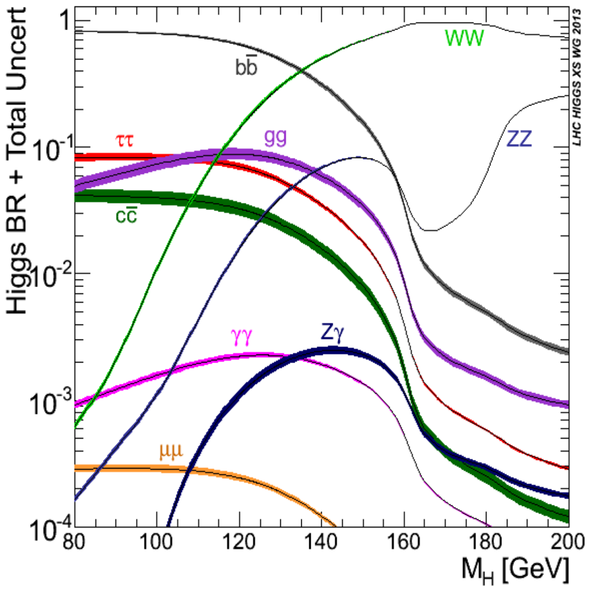


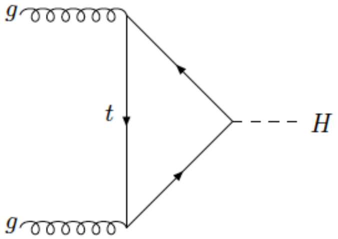
CMS has much better sensitivity than ATLAS

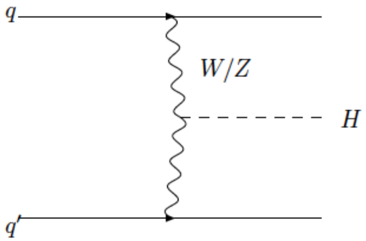
SM-like LM H→γγ search with full Run2

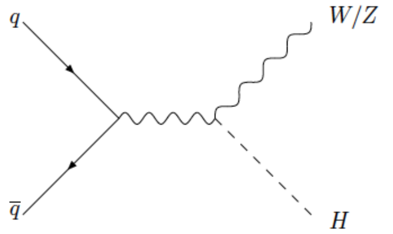
- **H→γγ provides** a clean final-state topology that allows the mass of a Higgs boson to be reconstructed with high precision (1-2%)
  - ✓ **Challenges: Large backgrounds** including continuum **γγ** (irreducible) and fakes from **γ+jet** and **jet+jet** (reducible) and **relic Z→ee bkg**
- **Production modes:** "SM-like", gluon fusion (ggH), vector boson fusion (VBF), in association with a W or Z boson (VH), or with a tt pair (ttH)
- **Cross sections and BR** : LHC Higgs Working Group YR4
- **Data samples**: full Run2 132.2 fb^{-1} data
  - ✓ lost ~5 fb^{-1} since HLT path was not introduced at the beginning of 2018 data-taking
- **Signal/searching region: 70-110 GeV**, backgoud fitting range 65-120 GeV
  - ✓ To avoid the distortion of the diphoton mass spectrum, due to turn-on effects from the **HLT criteria** (**M>55 GeV**)

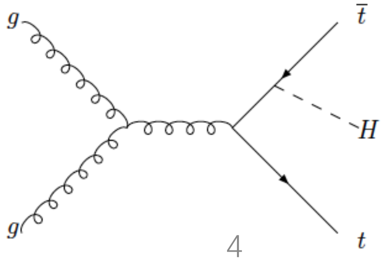








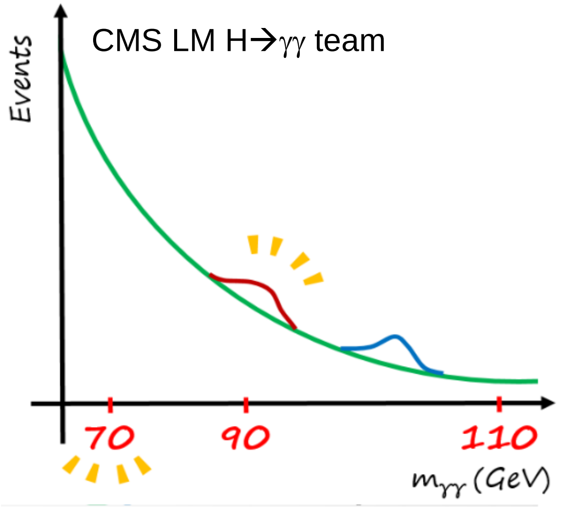




CMS-PAS-HIG-20-002 (approved for Moriond2023): **first LHC full Run 2 result of additional Higgs boson search in this mass range**

Analysis strategy

- Search for **narrow signal peak** over **smoothly-falling background** (direct $\gamma\gamma$, reducible γ + jet, jet+jet processes) except for **relic DY** ($Z \rightarrow ee$), in the diphoton mass spectrum



- Many **elements and techniques** (event reconstruction and calibrations, vertex determination BDT, γ ID techniques with BDT, signal and data-driven background modeling with discrete profiling method ⋯) inherited from **SM H→γγ** analysis

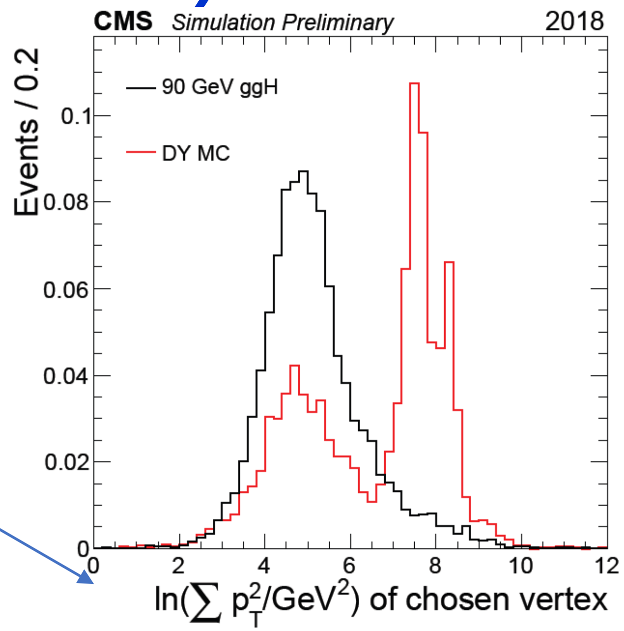
- **Dedicated updates and optimizations,** such as
  - ✓ Dedicated **HLT** paths then event (pre-) **selections** (SFs on MC)
  - ✓ Retrained γ-ID MVA in low-mass phase space, in 2017/2018
  - ✓ Dedicated **DY suppression** strategy (next slides)
  - ✓ **Diphoton BDT** retrained for low-mass case
  - ✓ Optimization of **event categorization (next slides)**

Analysis flow

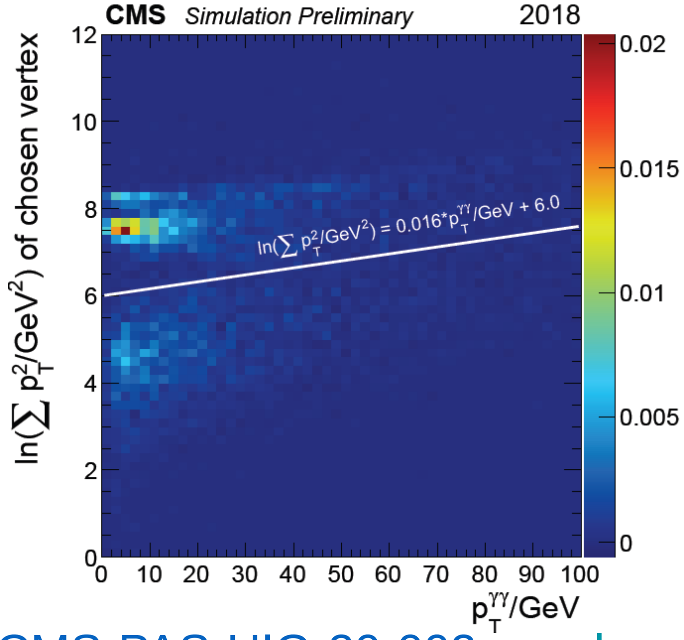
- Data & MC
- Trigger
- Photon reconstruction and energy calibration
- Preselection
- Vertex identification
- Photon identification
- Diphoton BDT/VBF MVA
- **Event categorization**
- **Signal/bkg modeling**
- **Statistical analysis** with "combine"
- **Results**

Analysis strategy (cont.)

- To **suppress DY bkg**, an **updated strategy** employed
  - ✓ **Electron-veto** by asking **no pixel detector hit** (used in previous 2016+2012 paper)
  - ✓ **Rejection** of γ **candidates also reconstructed as electrons**
  - ✓ Maximum value of **ln ($\Sigma p_T^2/GeV^2$)** [tracks in chosen vertex] **as function of $p_T^{\gamma\gamma}$ (GeV)**: optimized cut with ln (Σp_T^2) < 0.016 $p_T^{\gamma\gamma}$ + 6



- **2017/18**: events with additional jets were selected for class targeting VBF process
  - ✓ Di-jet BDT and combined di-jet + diphoton BDT
  - ✓ Validated with $Z \to ee$, diphoton data/MC
  - ✓ An additional VBF tagged event class, with the optimized cut value on combined BDT



DY MC

- **2016**: data reanalyzed with **improved calibration (legacy data)**

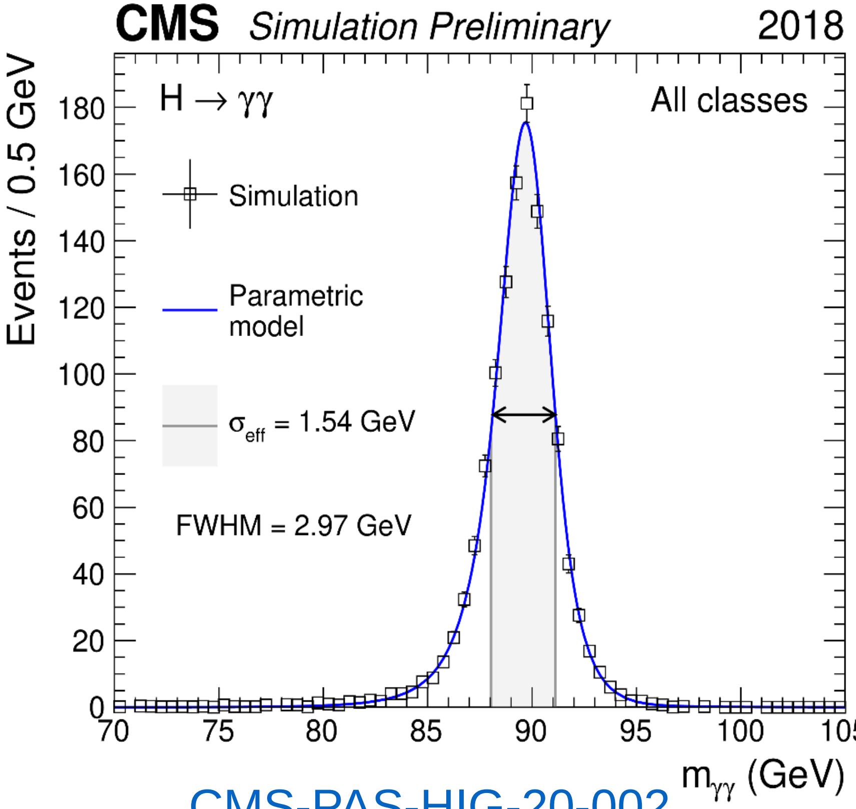
- 3 untagged event classes (targeting ggH) in each of 2016/2017/2018 + 1 additional VBF tagged event class in each of 2017/2018: **11 event classes** in total

CMS-PAS-HIG-20-002 supplementary

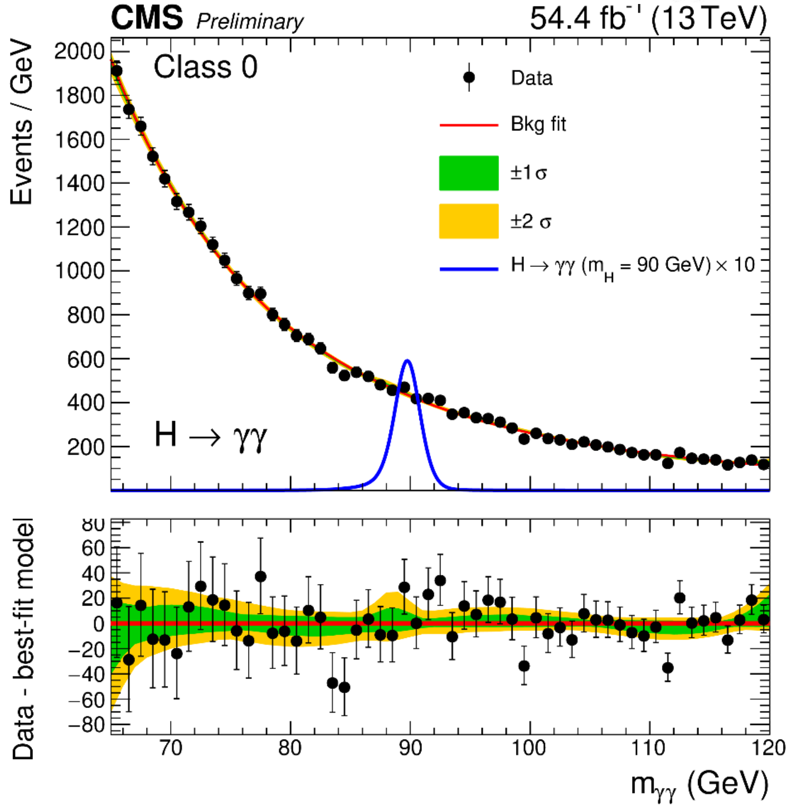
Sig and bkg modeling (parametrization)

- **Signal modeling** (sum of Gaussian functions): event classes x production modes x correct/incorrect vertex
  - ✓ A **simultaneous fit** of all mass point (SSF) from 70-110GeV

- **Background modeling** (discrete profiling): *sums of continuous functions* (different families/orders) for continuum bkg and *"DCB + exponential"* for relic DY (normalization floating)



CMS-PAS-HIG-20-002



- **Background fit, stat. unc. only, 2018**

$$Nf \times Exp(x) + N(1-f)(DCB(x)+exp(x)) = Nf \times \sum_{i=0}^{...} \beta_{2i} e^{\beta_{2i+1}x} + N(1-f)(DCB(x)+exp(x))$$

$$Nf \times Pow(x) + N(1-f)(DCB(x)+exp(x)) = Nf \times \sum_{i=0}^{M} \beta_{2i} x^{\beta_{2i+1}} + N(1-f)(DCB(x)+exp(x))$$

$$Nf \times Ber(x) + N(1-f)(DCB(x)+exp(x)) = Nf \times \sum_{i=0}^{M} \beta_i b_{i,M} + N(1-f)(DCB(x)+exp(x))$$

$$\text{with} \quad b_{i,M} = \binom{M}{i} x^i (1-x)^{M-i}$$

$$Nf \times Lau(x) + N(1-f)(DCB(x)+exp(x)) = Nf \sum_{i=0}^{M} \beta_i x^{-\alpha_i} + N(1-f)(DCB(x)+exp(x))$$

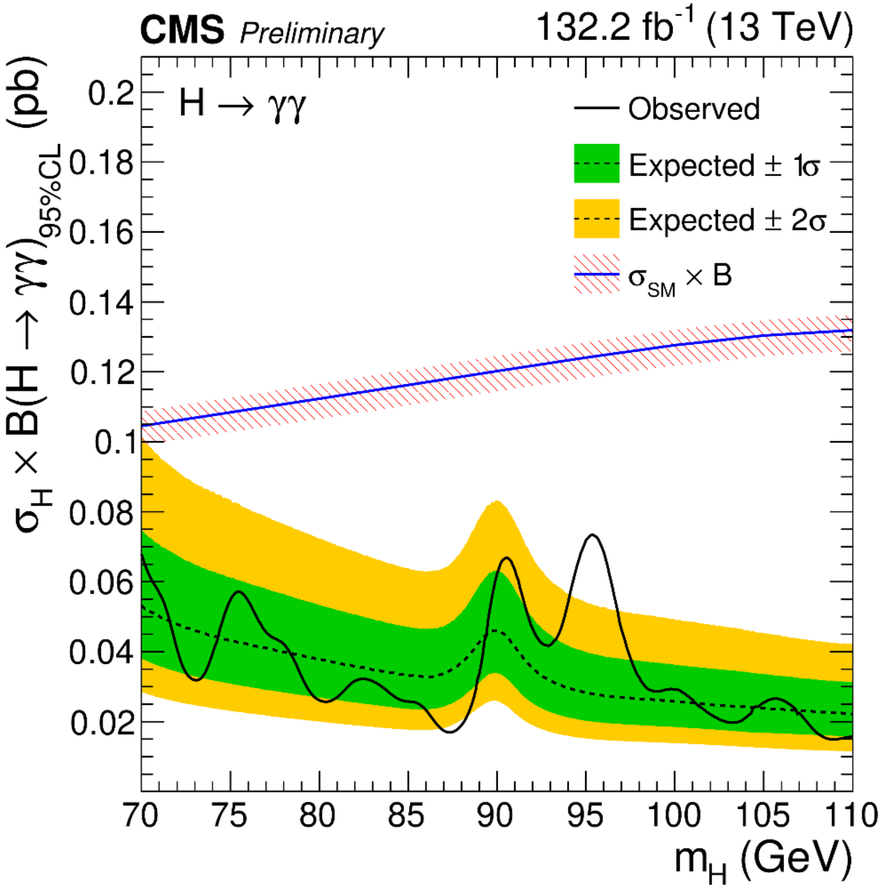
with $\alpha_i = 4,5,3,6,2,7,...$ for $i = 0,1,2,3,4,5,...$

Event class		0	1	2	VBF
2016	Family/Order	Power Law 1	Bernstein 4	Exponential 3	
	DCB + Exp. Fraction (%)	3.0	3.1	3.3	
2017	Family/Order	Bernstein 3	Exponential 3	Bernstein 4	Bernstein 3
	DCB + Exp. Fraction (%)	2.7	1.4	1.9	2.6
2018	Family/Order	Laurent 1	Bernstein 4	Exponential 3	Bernstein 2
	DCB + Exp. Fraction (%)	0.5	4.1	4.8	0.8

Best-fit background functions w/ "DCB + exponential" fractions (0.5% - 4.8%) in [85, 95] GeV

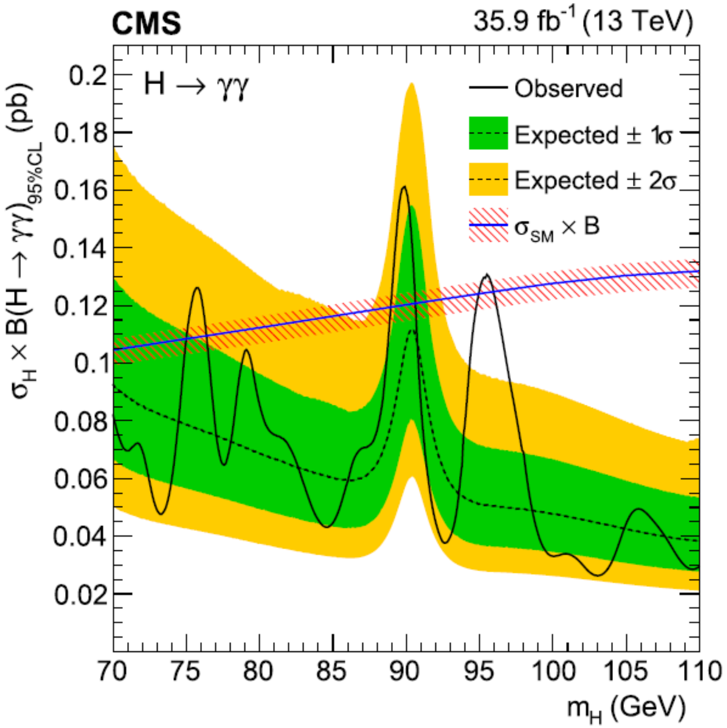
Upper limits on σ X B

- **Observed absolute 95% CL UL on σ X B between 15-73 fb** (*22-53 fb expected*)



CMS-PAS-HIG-20-002

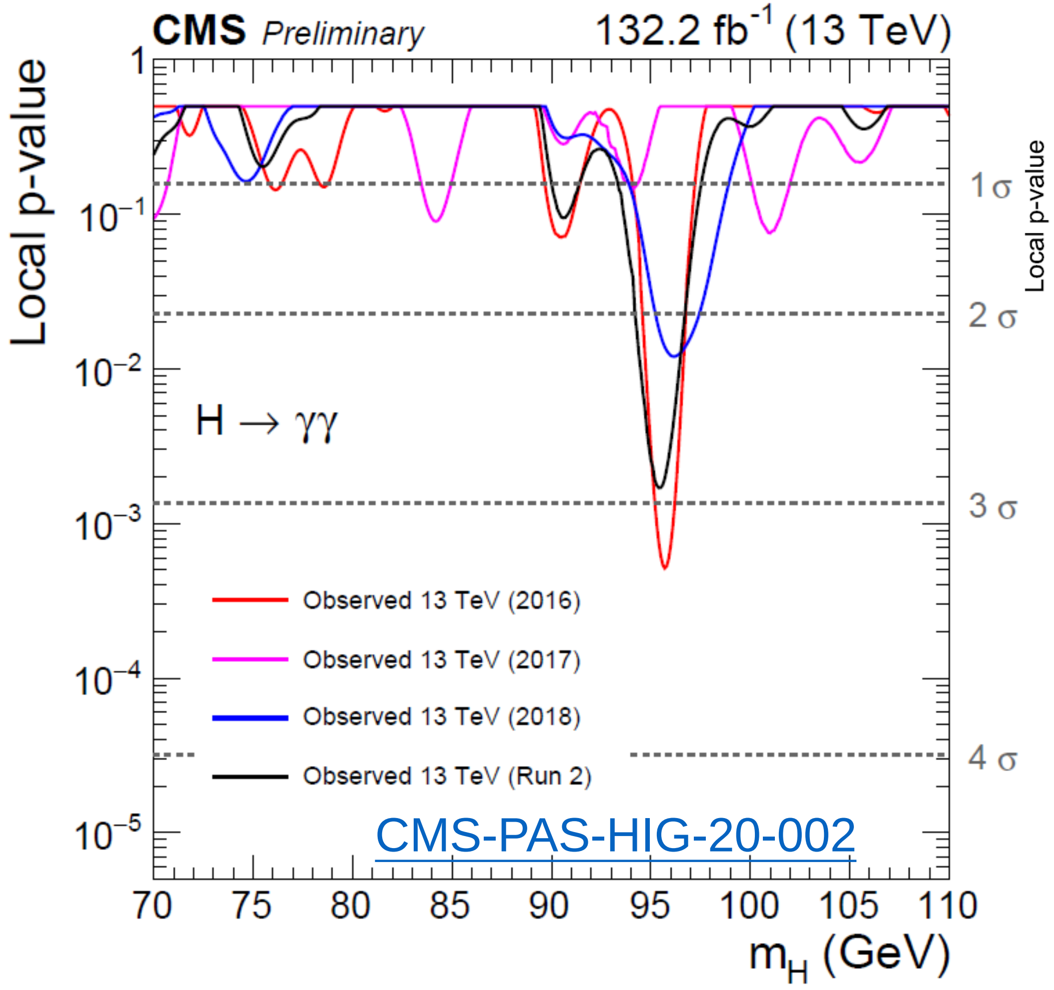
Previous 2016 results (HIG-17-013):
26-161 fb (obs.)
37-110 fb(exp.)

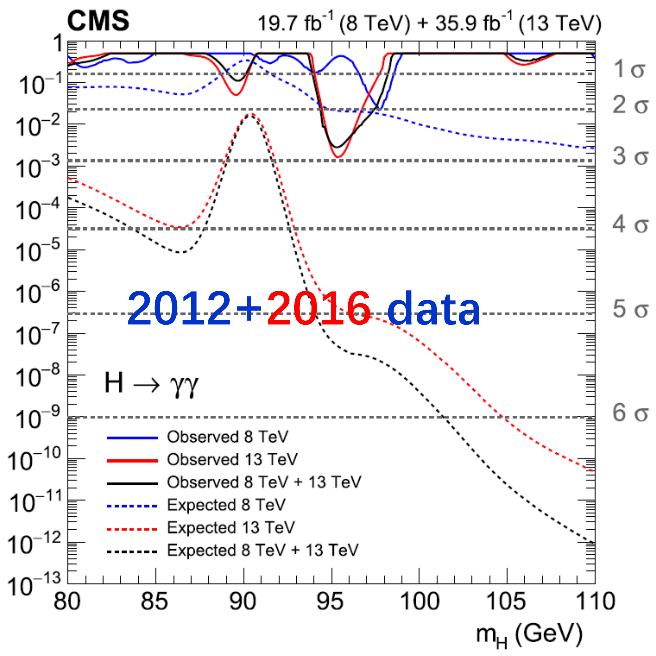


PLB 793 (2019) 320–347

P-values or significances and mass

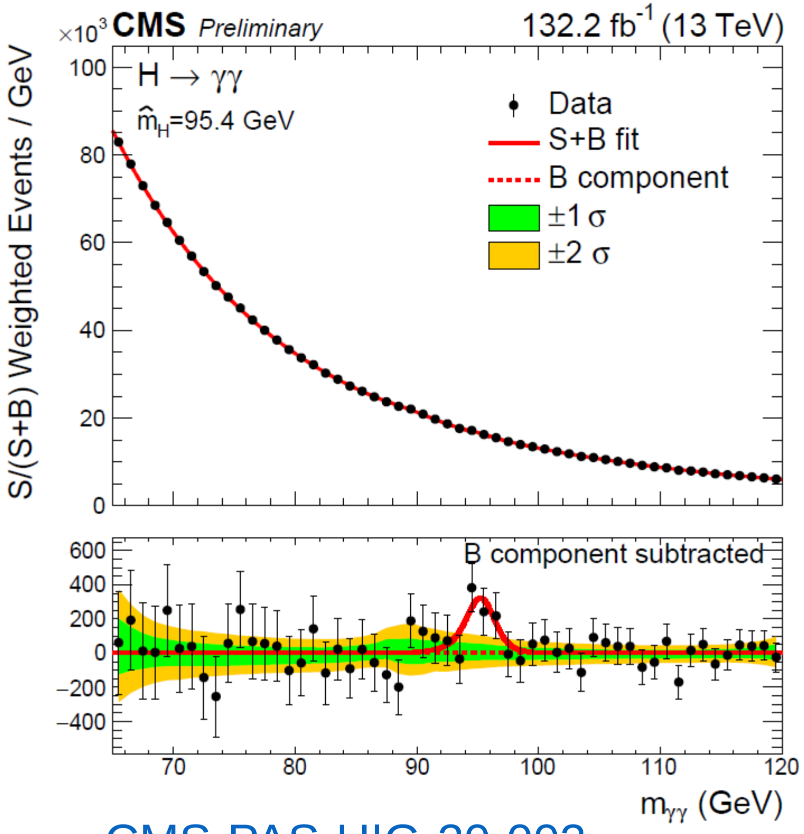
- **Observed local p-values** for 2016, 2017, 2018 and combination





PLB 793 (2019) 320–347

- **S/(S+B)-weighted** $m_{\gamma\gamma}$ distribution with S+B fit for m_H =95.4 GeV

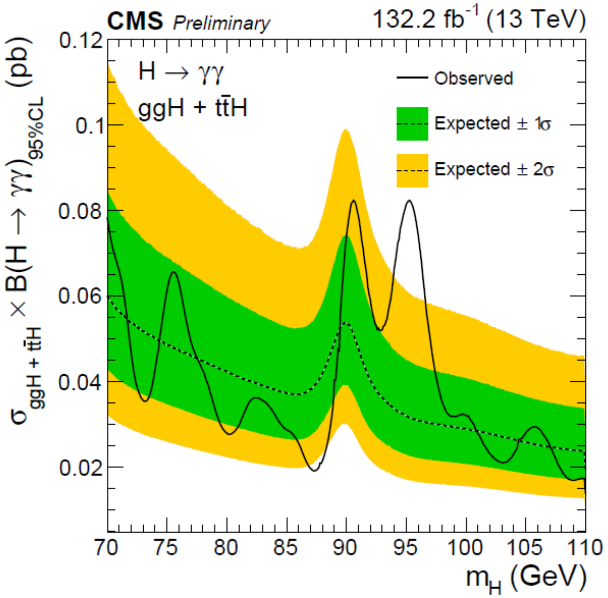


CMS-PAS-HIG-20-002

- Modest excess with ~2.9σ local (1.3σ global) significance at $m_{\gamma\gamma}$=95.4 GeV, more data needed to conclude!

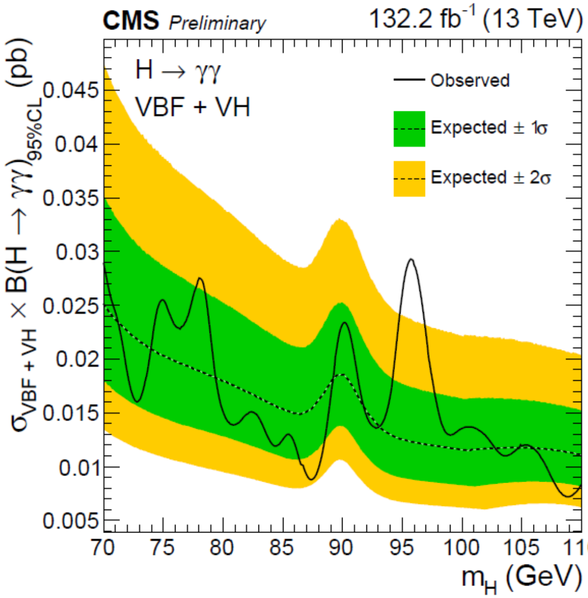
Upper limits by production process

➢ Observed and expected 95% CL limits on σ X B by **production process** (integrated over all event classes)



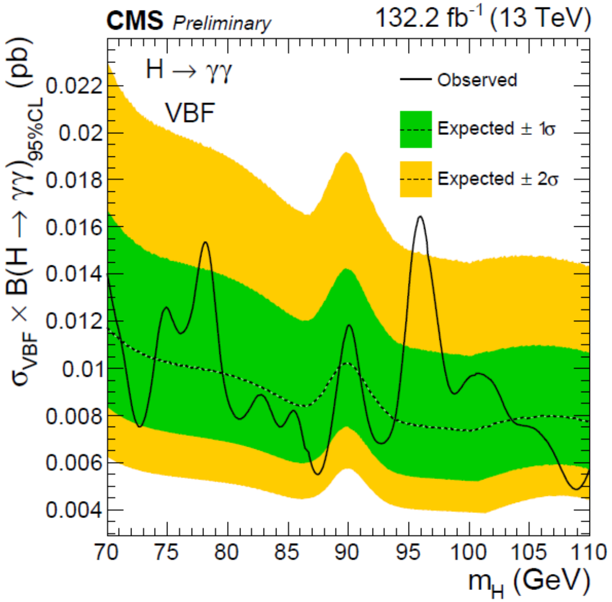
- 100% production via gluon-induced processes (**ggH, ttbarH** in SM proportions)

  17-83 fb observed



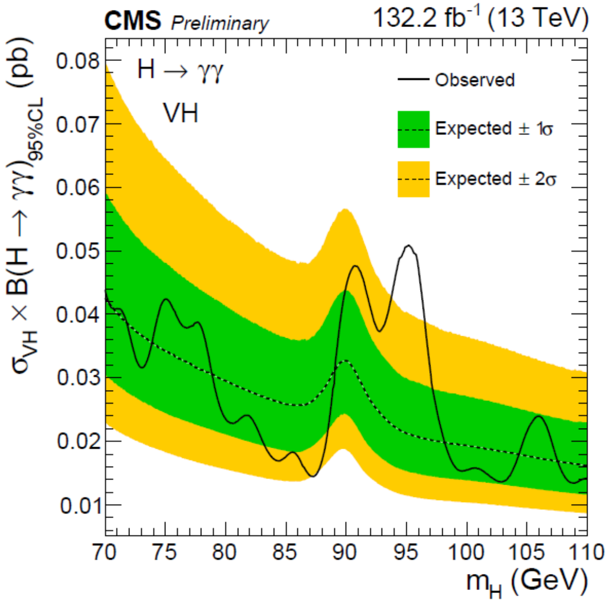
- 100% production via fermion-induced processes (**VBF, VH** in SM proportions)

  7-29 fb observed



- 100% production via **VBF**

  5-17 fb observed



- 100% production via **VH**

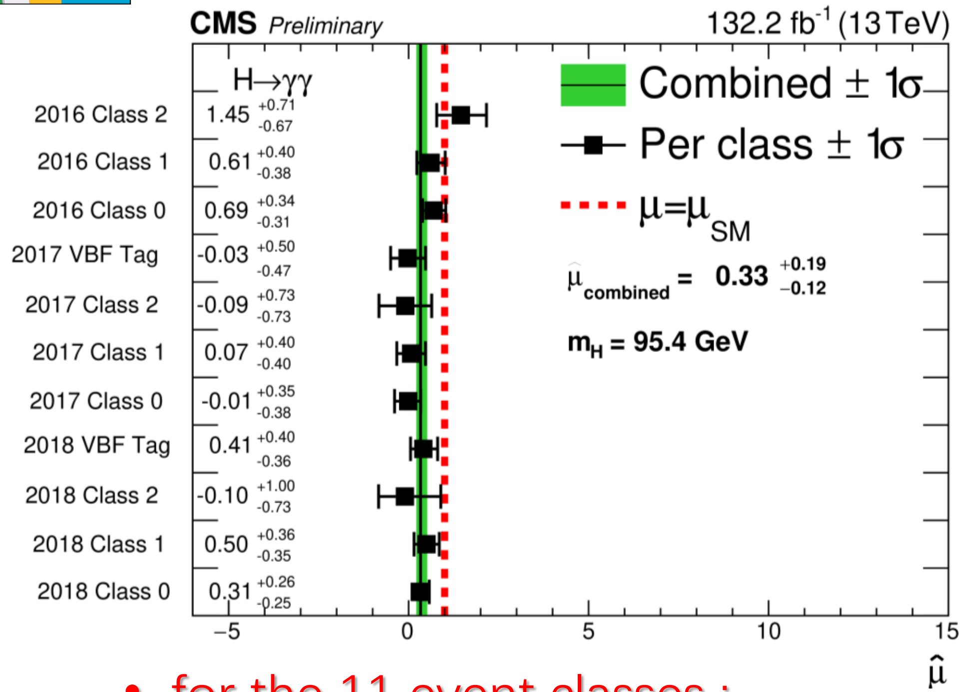
  13-51 fb observed

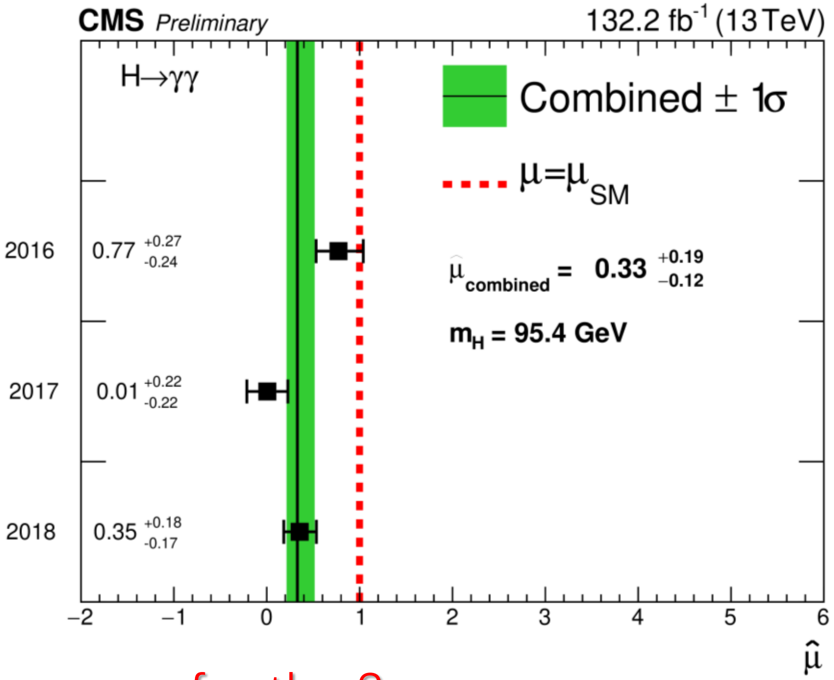
CMS-PAS-HIG-20-002

More interpretation results have been included in the Paper supplementary material : approved by Higgs conveners and will be public in near future together with the paper

‘Signal’ strengths μ : fixing m_H = 95.4 GeV

中国科学院高能物理研究所
Institute of High Energy Physics
Chinese Academy of Sciences



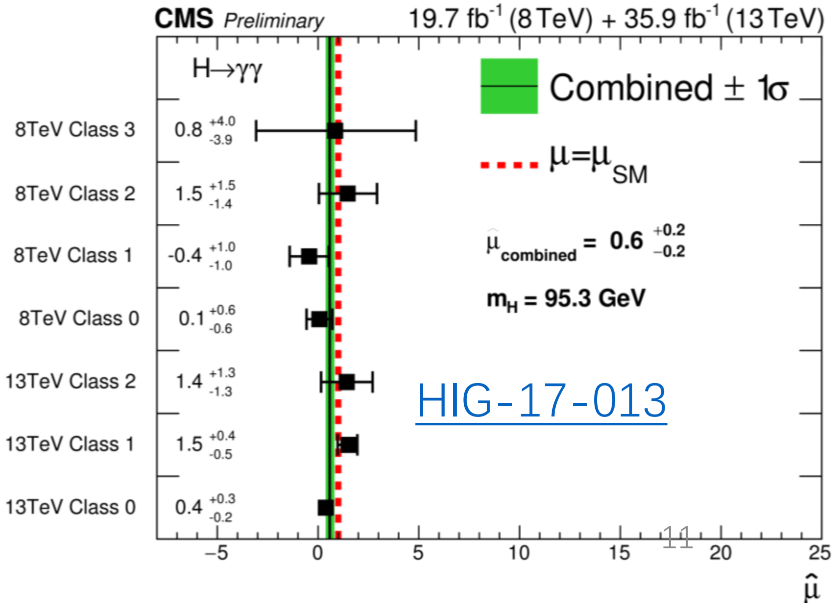


Combined obs. μ is compatible with previous obs. μ in 2012+2016

- for the 11 event classes :
  χ^2 compatibility
  probability: 68%

- for the 3 years
  χ^2 compatibility
  probability: 6%

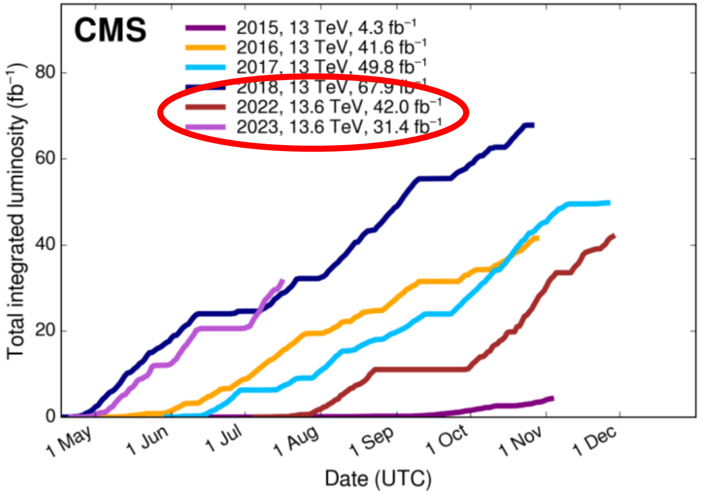
CMS-PAS-HIG-20-002 supplementary

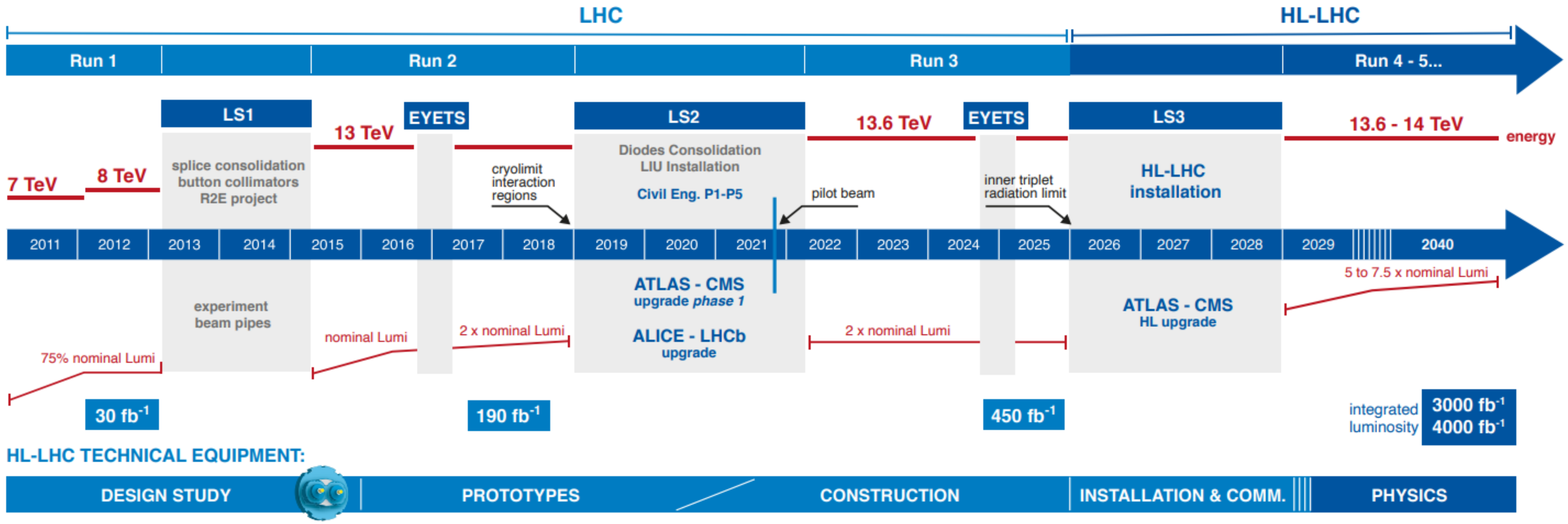


Conclusions and Perspectives

- Presented the **updated results of CMS search for additional low-mass SM-like H→γγ** (70 GeV < m_H < 110 GeV) using full LHC Run 2 data (approved for Moriond2023)
- **First search for new diphoton resonances in this mass range with full LHC Run 2 data !**
- No evidence for the existence of extra Higgs bosons found so far
- Modest excess at $m_{\gamma\gamma}$=95.4 GeV with 2.9σ local (1.3σ global) significance
- More (Run 3) data is needed to conclude on the nature of this excess….and it's on it's way!

2022: ~40fb^{-1}
2023: ~30fb^{-1}



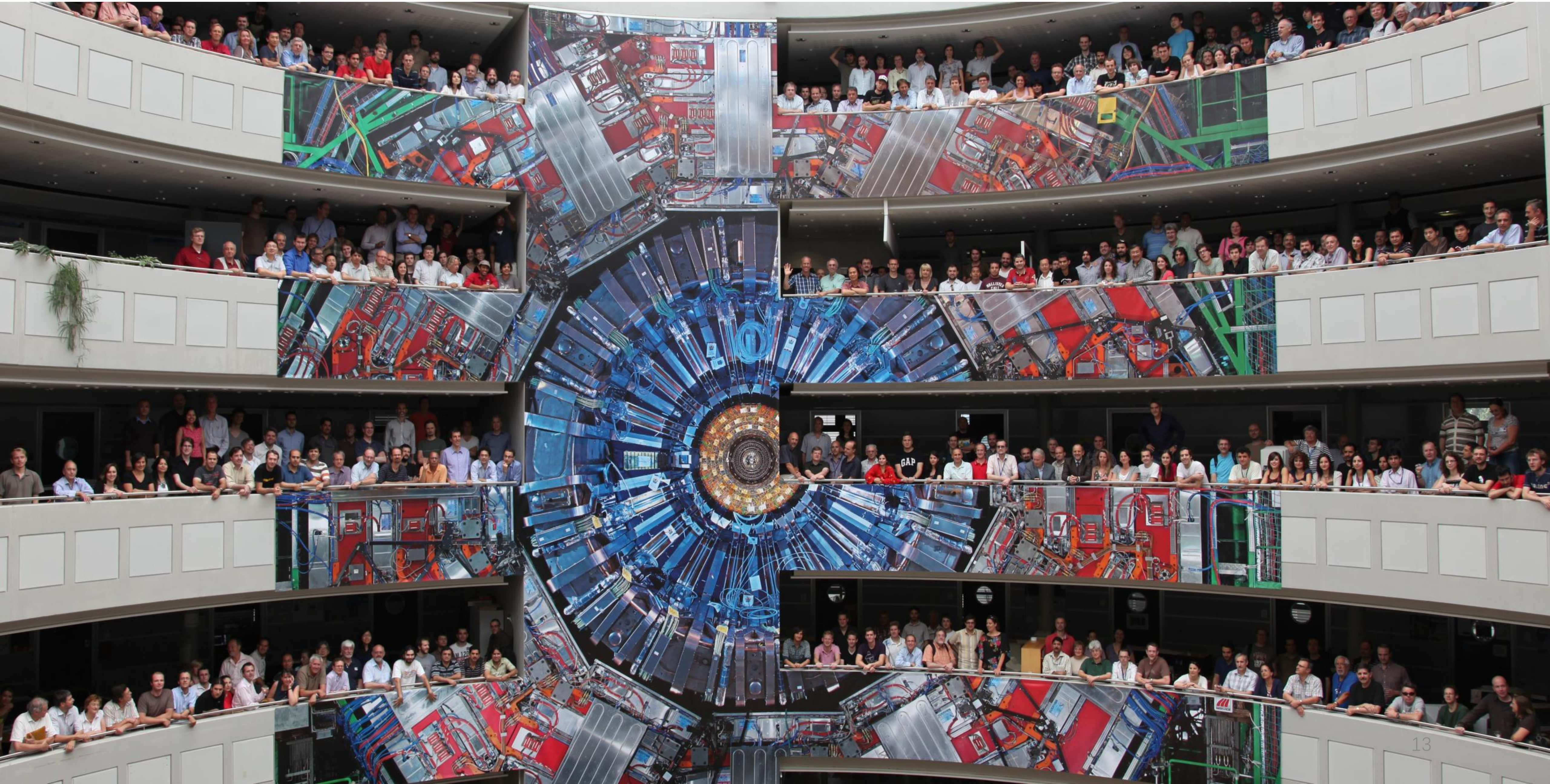


https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults

https://hilumilhc.web.cern.ch/content/hl-lhc-project

(last update February 2022)

Thanks for your attention!

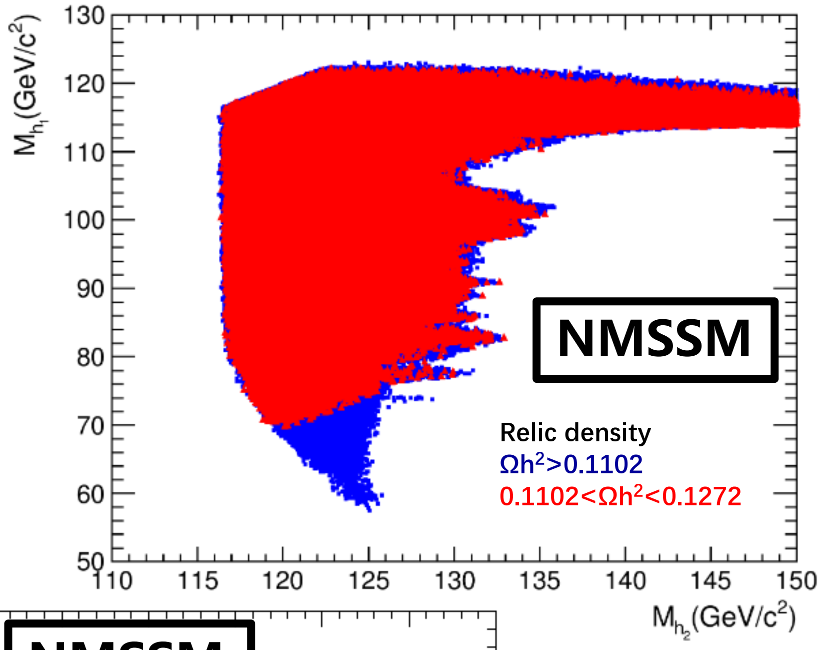


Backup slides

Low-mass diphoton searches

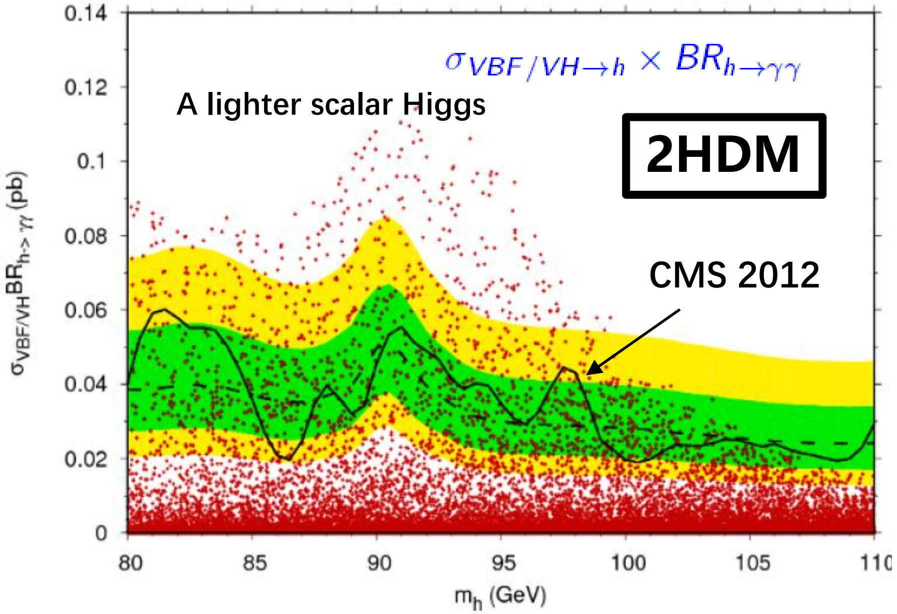
- Many **BSM models** (e.g. NMSSM, 2HDM, Georgi-Machacek model) provide a Higgs boson that is compatible with the LHC observed 125 GeV boson, and additional Higgs bosons with some of which could have masses below 125 GeV
  - ✓ Next-to-minimal supersymmetric model (NMSSM): 3 CP-even (h_i), 2 CP-odd (a_i), 2 charged ($h^{\pm}$)
  - ✓ Two Higgs Doublet Models (2HDM): 2 CP-even (h, H), 1 CP-odd (a), 2 charged ($H^{\pm}$)
  - ✓ Georgi-Machacek(GM) model: 2 singlet(h, H), 3 triplet(H_3^0, $H_3^{\pm}$) and 5 fiveplet(H_5^0, $H_5^{\pm}$, $H_5^{\pm\pm}$) mass eigenstates

J. Fan, **JT(*)**, et al., Chinese Phys. C 38 (**2013**) 073101

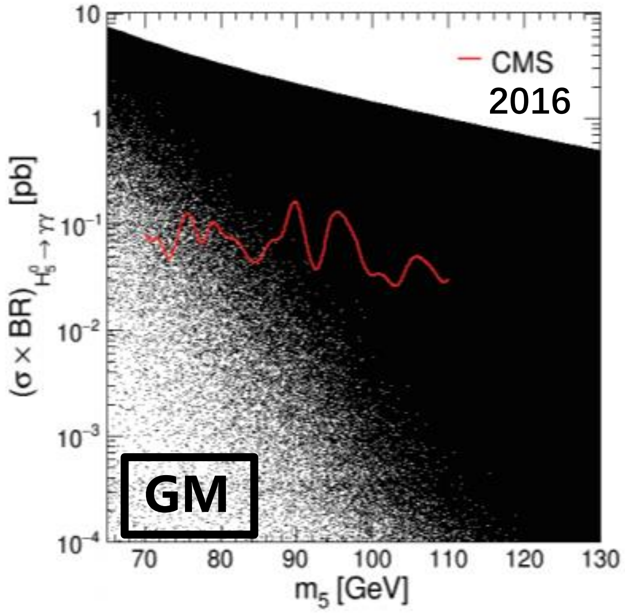


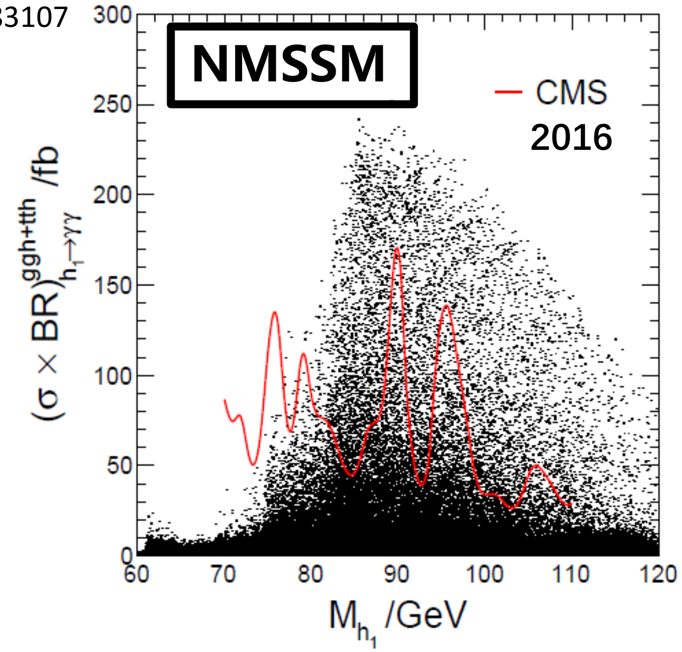
- **Related phenomenology studies**

G.Cacciapaglia, **JT**, et al JHEP12 (**2016**) 068



C. Wang, **JT(*)** et al., Chin. Phys. C 46 (**2022**) 8, 083107





JT et al., Chin. Phys. C 42 (**2018**) no.10, 103107

Analysis strategy

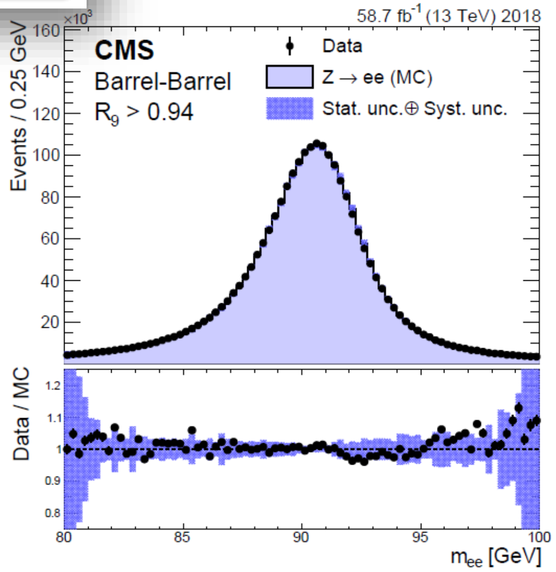
$$m_{\gamma\gamma} = \sqrt{2E_1E_2(1-\cos\theta)}$$

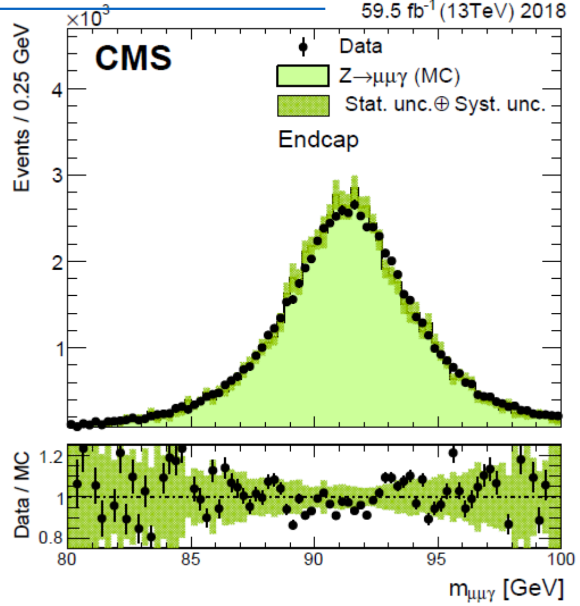
- ➢ **To achieve good mass resolution**
  - ✓ select/reconstruct two **photons** with precise **photon energy MVA *regression*** after calibrations
    -- Validated with both $Z\rightarrow ee$ and $Z\rightarrow\mu\mu\gamma$
  - ✓ A ***MVA BDT*** was trained to select the **primary vertex** of the Higgs decay (details in backup)
    -- *V*alidated on $Z\rightarrow\mu\mu$ after removing muon tracks and γ+j (>=1 converted γ)
  - ✓ Inherited from **SM H→$\gamma\gamma$** analysis

- ➢ **Fake photon suppression**: **photon identification *BDT*** inputs of diphoton BDT after looser cut (>-0.9)
  - ✓ Trained with both prompt and fake photons selected in γ+j MC
  - ✓ Validated with diphoton data/MC, $Z\rightarrow ee$ and $Z\rightarrow\mu\mu\gamma$
  - ✓ 2017/2018: **retrained γ-ID MVA** in low-mass phase space

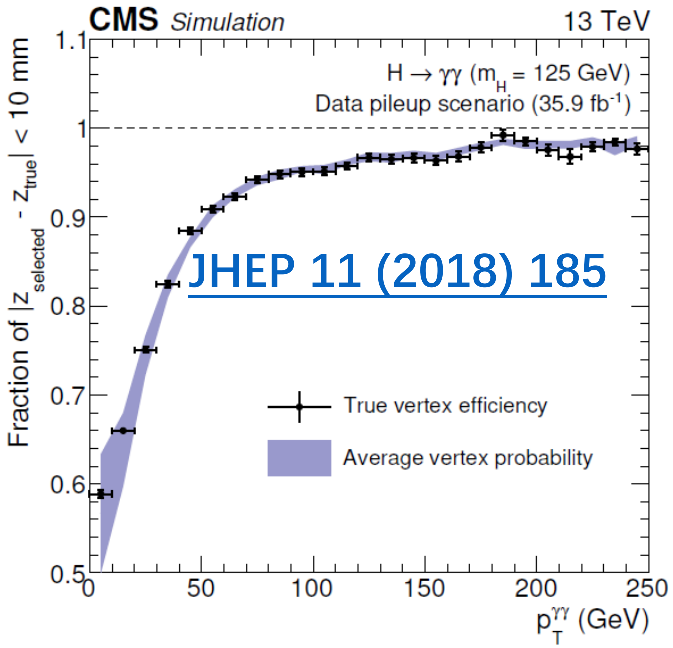
- ➢ A Kinematic **diphoton BDT** ($p_t/m_{\gamma\gamma}$, η , $\cos(\phi_{\gamma1}-\phi_{\gamma2})$, both Photon ID MVA scores, mass resolutions wrt correct and incorrect vertices, vertex probability) for sig and bkg discrimination **retrained and reoptimized** for events categorization, for low-mass case
  - ✓ Trained with signal and bkg MC
  - ✓ Validated with $Z\rightarrow ee$, diphoton data/MC

2021 JINST 16 P05014



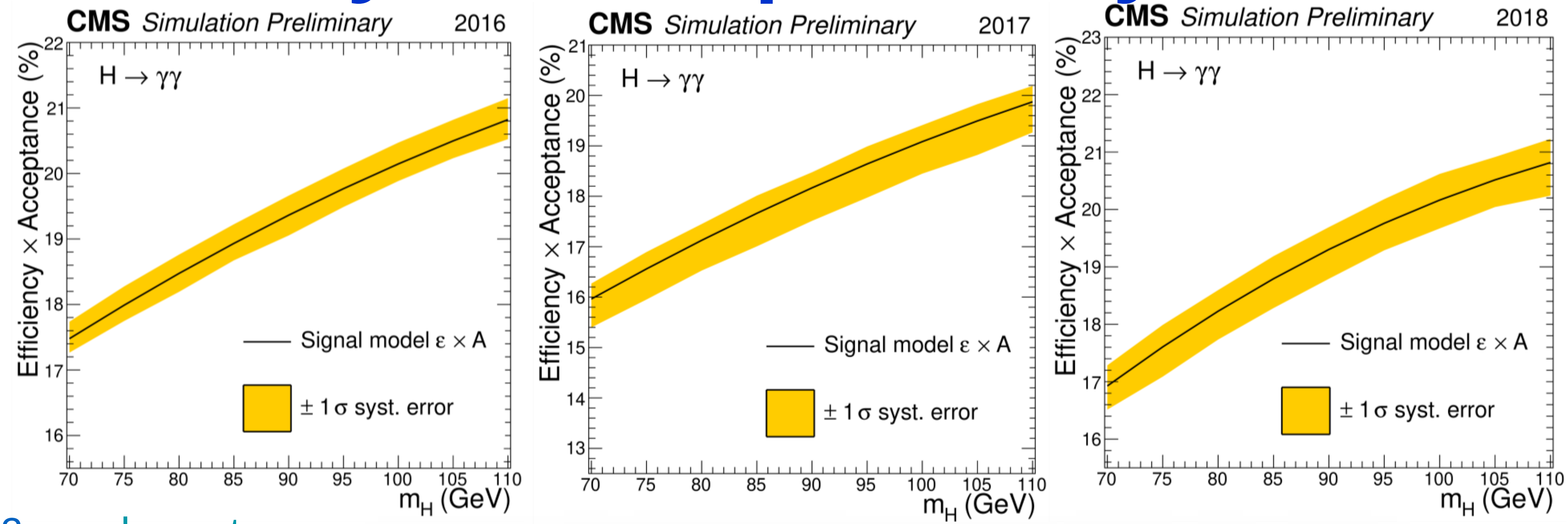


JHEP 11 (2018) 185



Signal efficiency and expected yields

- **Signal efficiency x acceptance** as a function of mass hypothesis
  - ✓ Different selections for 3 years i.e. trigger, offline selections etc
  - ✓ 1σ systematic error band is also shown



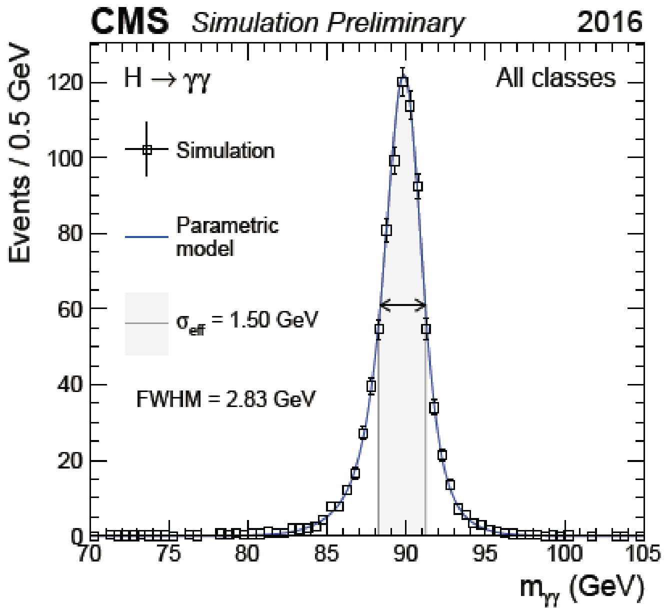
CMS-PAS-HIG-20-002 supplementary

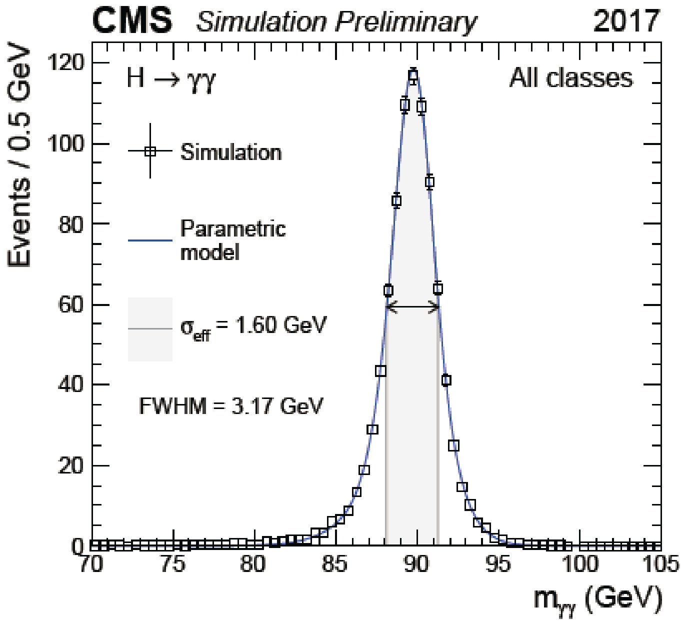
- **Expected number** of SM-like Higgs boson signal events (m_H=90 GeV) per event class and the corresponding percentage breakdown per production process
  - ✓ Also number of background events (``Bkg.'') and Drell--Yan process (``DY Bkg.'') per GeV estimated from the background-only fit to the data, in a σ_{eff} window centered on m_H = 90 GeV

Event classes		Expected SM-like Higgs boson signal yield ($m_H = 90$ GeV)								Bkg.	DY Bkg.
		Total	ggH (%)	VBF (%)	WH (%)	ZH (%)	$t\bar{t}H$ (%)	σ_{eff} (GeV)	σ_{HM} (GeV)	(GeV^{-1})	(GeV^{-1})
2016	0	130	71.9	15.6	6.2	3.6	2.6	1.12	1.00	271	12
36.3 fb^{-1}	1	304	87.4	6.6	3.6	2.1	0.3	1.25	1.07	3093	33
	2	407	94.7	2.5	1.7	1.0	0.1	1.87	1.51	9190	193
	Total	842	88.5	6.0	3.1	1.8	0.6	1.50	1.20	12 554	239
2017	0	104	73.4	11.6	7.5	4.3	3.2	1.27	1.13	248	7
41.5 fb^{-1}	1	347	88.5	5.6	3.5	2.1	0.3	1.40	1.24	3625	83
	2	413	94.4	2.6	1.9	1.1	0.1	1.91	1.64	8169	244
	VBF	26	45.6	51.8	1.0	0.5	1.0	1.33	1.15	29	1
	Total	890	88.2	6.2	3.1	1.8	0.6	1.60	1.35	12 071	338
2018	0	162	75.1	10.2	7.3	4.3	3.0	1.21	1.05	430	3
54.4 fb^{-1}	1	585	90.1	4.8	3.1	1.8	0.2	1.34	1.17	6445	378
	2	473	94.4	2.5	1.9	1.2	0.1	2.01	1.73	10 982	720
	VBF	38	45.4	51.9	1.1	0.6	1.0	1.21	1.03	46	1
	Total	1258	88.4	6.1	3.1	1.8	0.6	1.54	1.27	17 902	1104

Signal parametrization

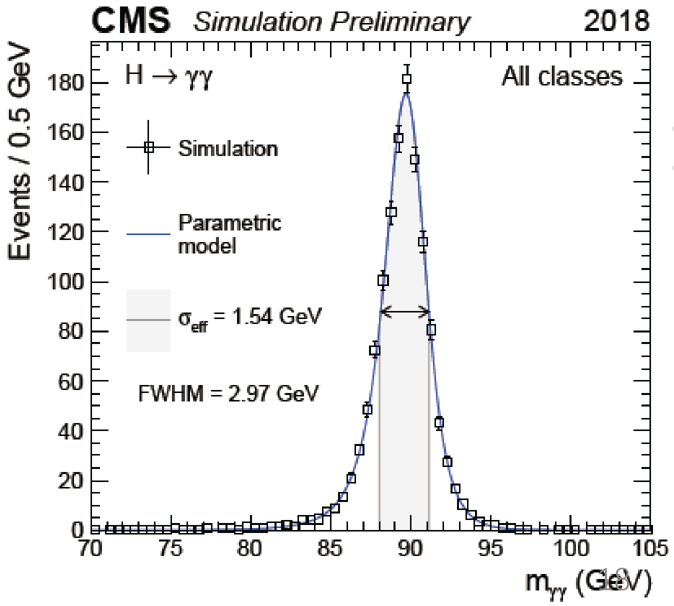
- A parametric model is used to describe the shape of the signal in each event class
- Same as the standard $H \rightarrow \gamma\gamma$ analysis method: Simultaneous Signal Fitting
- All production modes (ggH, VBF, WH, ZH, ttH) from 70 GeV to 110 GeV with a 5 GeV granularity are used
- The final parametrized signal shape for the combination of all production modes for all event classes, is weighted by their SM-like BSM cross sections
- Full parameterized signal shape, integrated over all event classes, in simulated signal events with m_H = 90 GeV





All classes combined

[CMS-PAS-HIG-20-002]



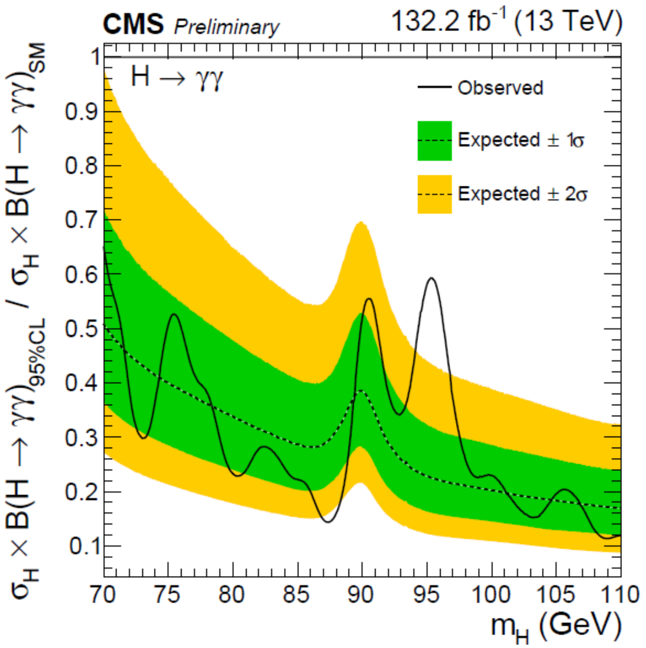
Systematic uncertainties

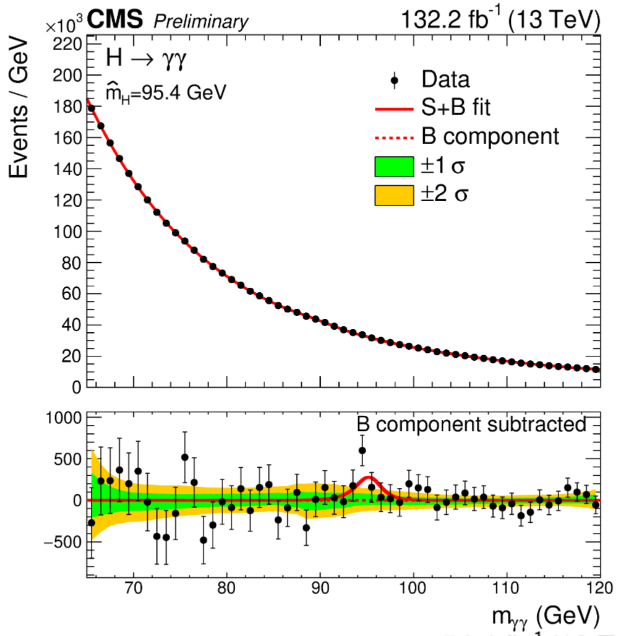
- **Uncertainties evaluated at the per-event level:**
  - ✓ Total integrated luminosity
  - ✓ 2016 and 2017 pre-firing
  - ✓ Underlying event and parton shower
  - ✓ 2018 HEM issue
  - ✓ 2017 and 2018, VBF additional jet radiation issue
  - ✓ Linear cut SF
- **Uncertainties evaluated at the per-photon level:**
  - ✓ Shape of the photon identification BDT distribution
  - ✓ Photon energy scale and resolution
  - ✓ Trigger efficiencies SF
  - ✓ Preselection SF
  - ✓ Electron veto SF and $N_{MatchedEle}=0$ SF
  - ✓ Minimum photon identification BDT
  - ✓ Non-uniformity of light collection (FNUF)
  - ✓ Photon energy scale non-linearity
  - ✓ Vertex selection uncertainty
- **Dedicated systematics have been added for VBF class:**
  - ✓ jet energy correction and resolution,
  - ✓ PUJID
  - ✓ Tight Jet ID
- **Theoretical uncertainties:**
  - ✓ PDF uncertainty
  - ✓ QCD scale and strong coupling strength (α_s) uncertainty
  - ✓ Cross-section uncertainties (for normalized limit and p-value)

- Major systematic uncertainties: per-photon energy resolution <20%, renormalization and factorization scales<14%, UE modeling <27%, PS<16%, JES corrections (VBF class) <16%.

Other related plots

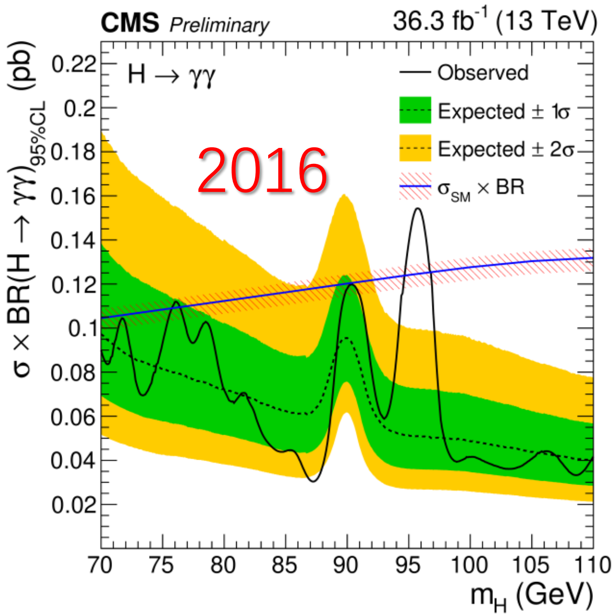
- **Observed and expected 95% CL UL on σ X B** ***relative to SM-like expectation*** **(production processes assumed in SM proportions)**

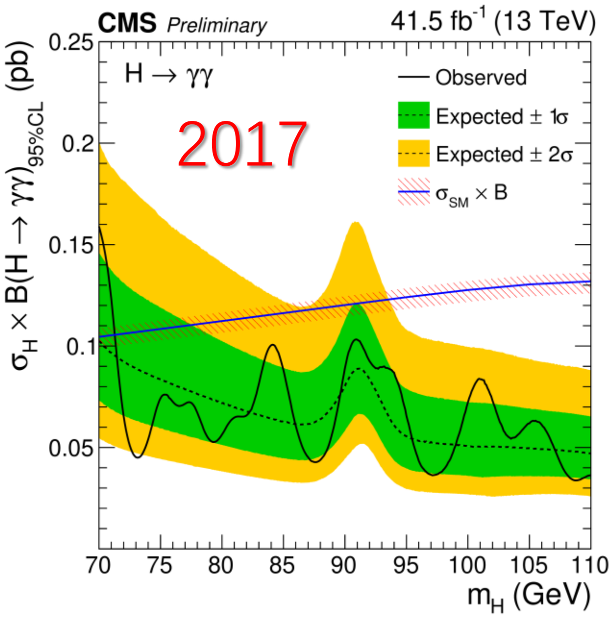


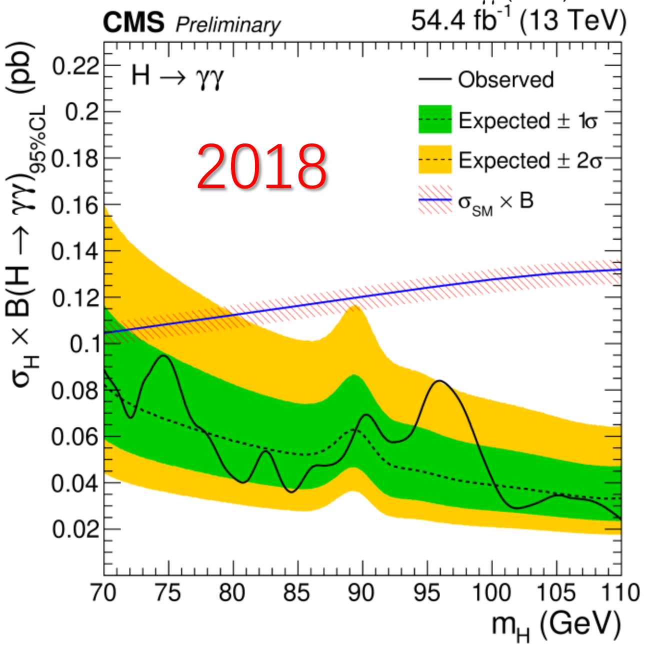


- Events in all classes of the combined 13 TeV data set, S+B fit for m_H =95.4 GeV

CMS-PAS-HIG-20-002 supplementary

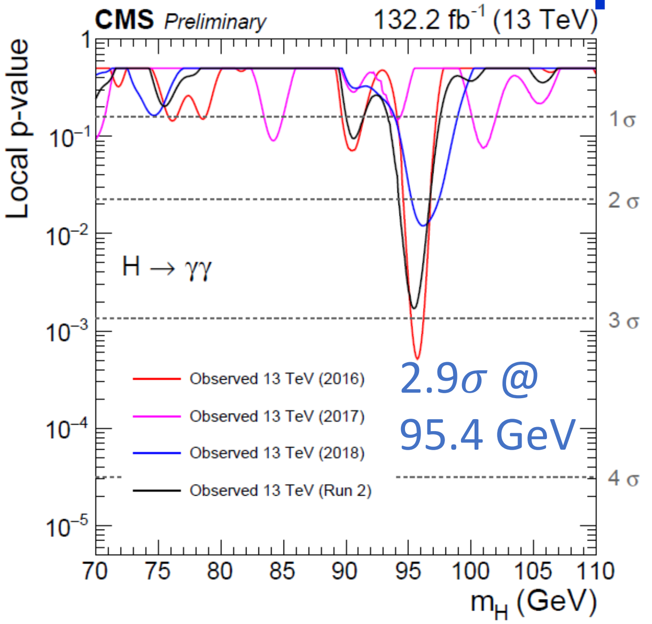


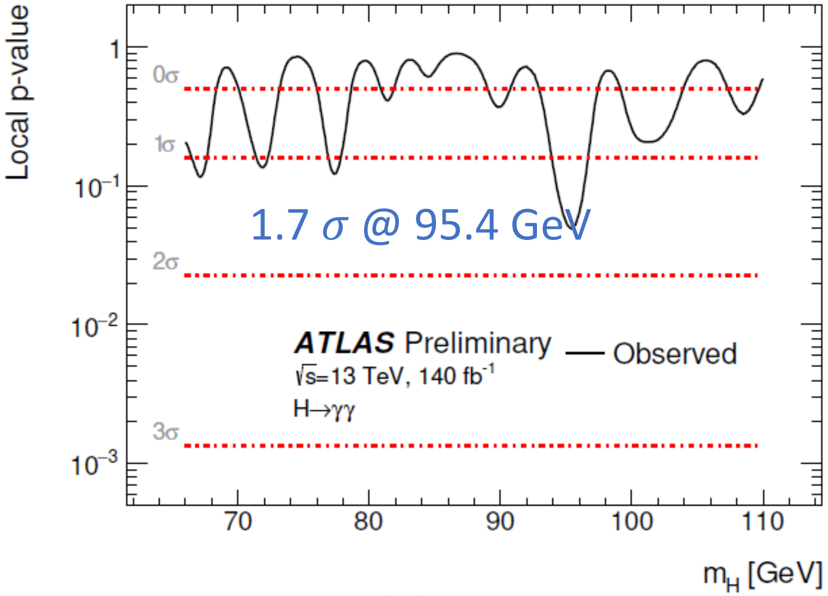




Limits on σ X B in each year

Comparisons with ATLAS

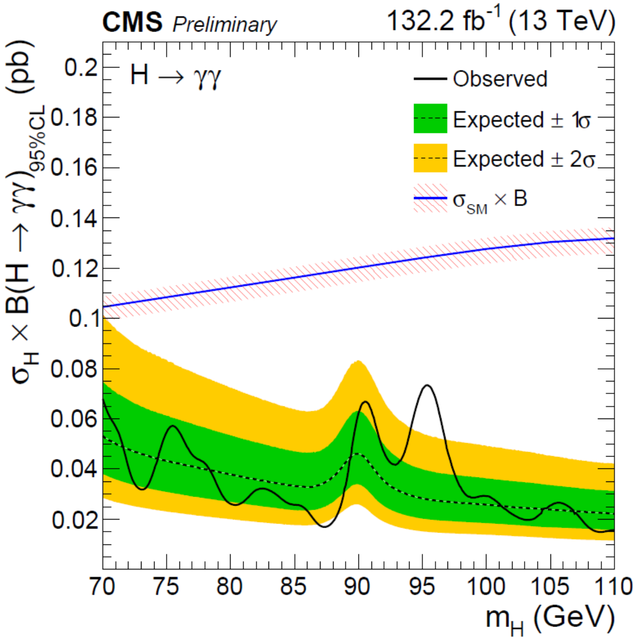


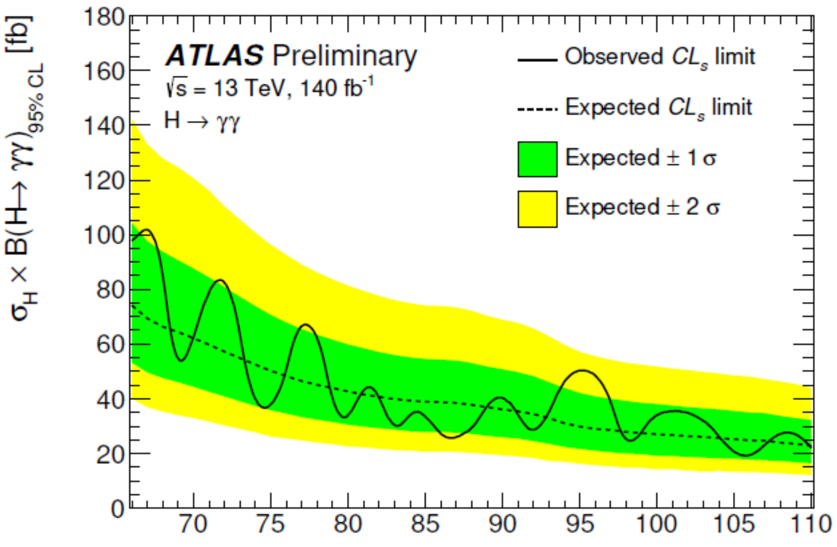


CMS-PAS-HIG-20-002

ATLAS-CONF-2023-035

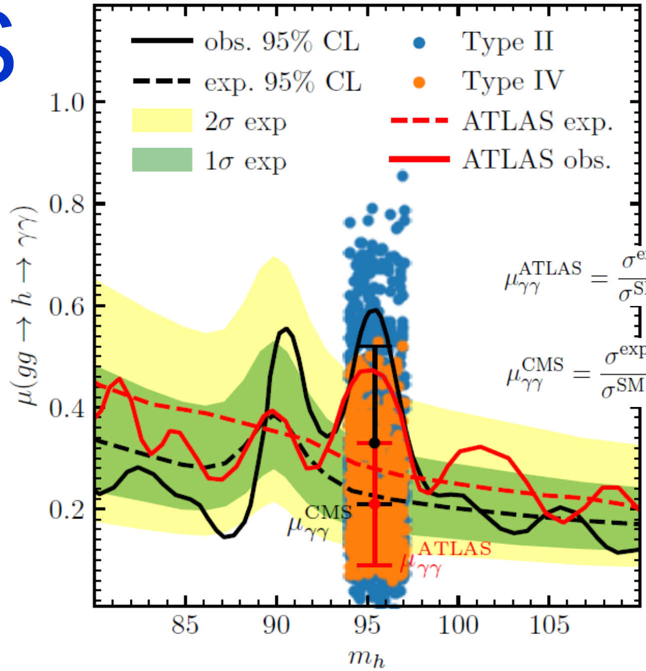
June 6, 2023





Also employing BDT classifier

Extension of the 2HDM by a complex singlet (S2HDM)

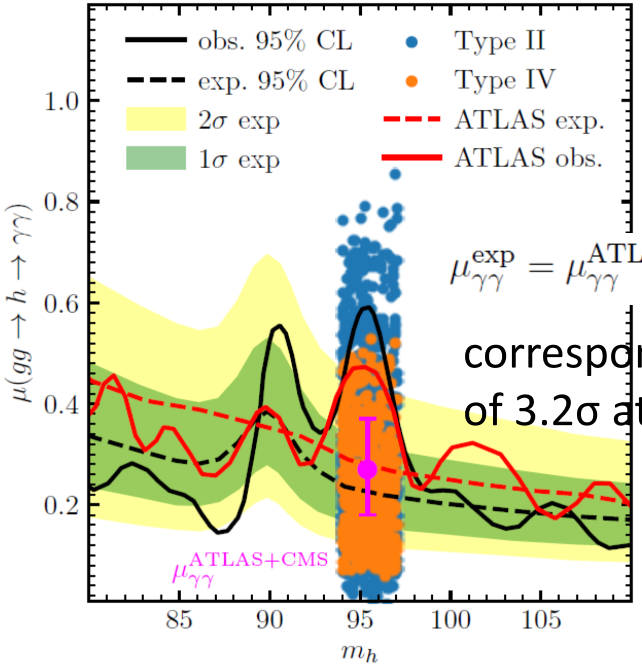


$$\mu^{ATLAS}_{\gamma\gamma} = \frac{\sigma^{exp}(gg \to \phi \to \gamma\gamma)}{\sigma^{SM}(gg \to H \to \gamma\gamma)} = 0.21^{+0.12}_{-0.12}$$

$$\mu^{CMS}_{\gamma\gamma} = \frac{\sigma^{exp}(gg \to \phi \to \gamma\gamma)}{\sigma^{SM}(gg \to H \to \gamma\gamma)} = 0.33^{+0.19}_{-0.12}$$

arXiv:2306.03889

CMS still has better sensitivity (except Z-peak region) than ATLAS, even with less data



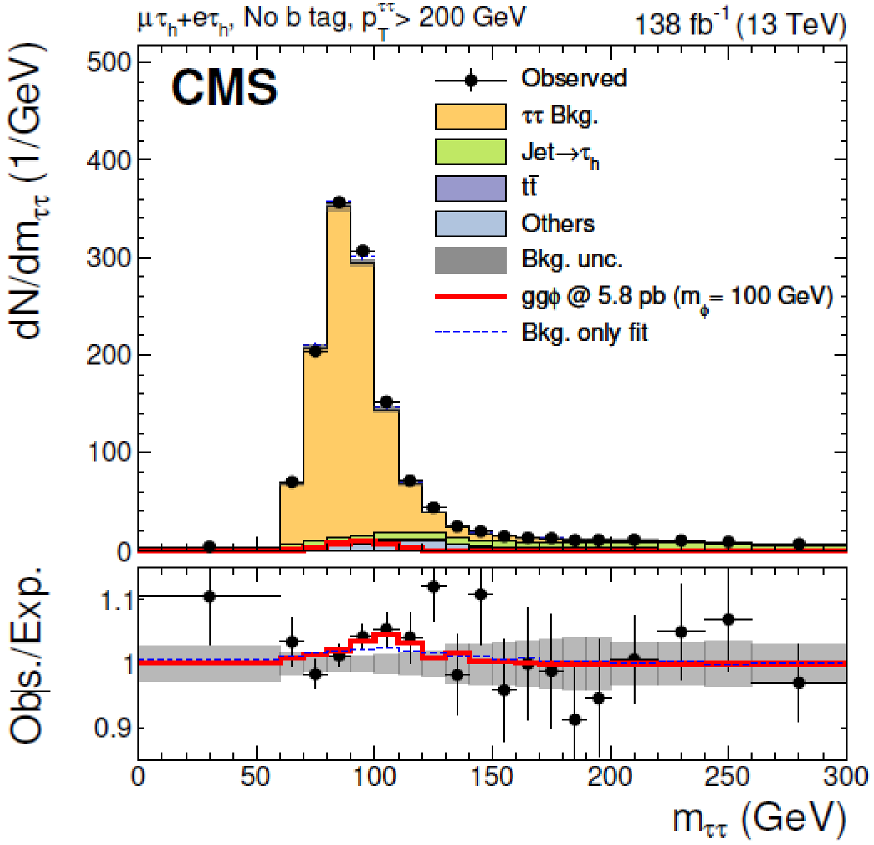
$$\mu^{exp}_{\gamma\gamma} = \mu^{ATLAS+CMS}_{\gamma\gamma} = 0.27^{+0.10}_{-0.09}$$

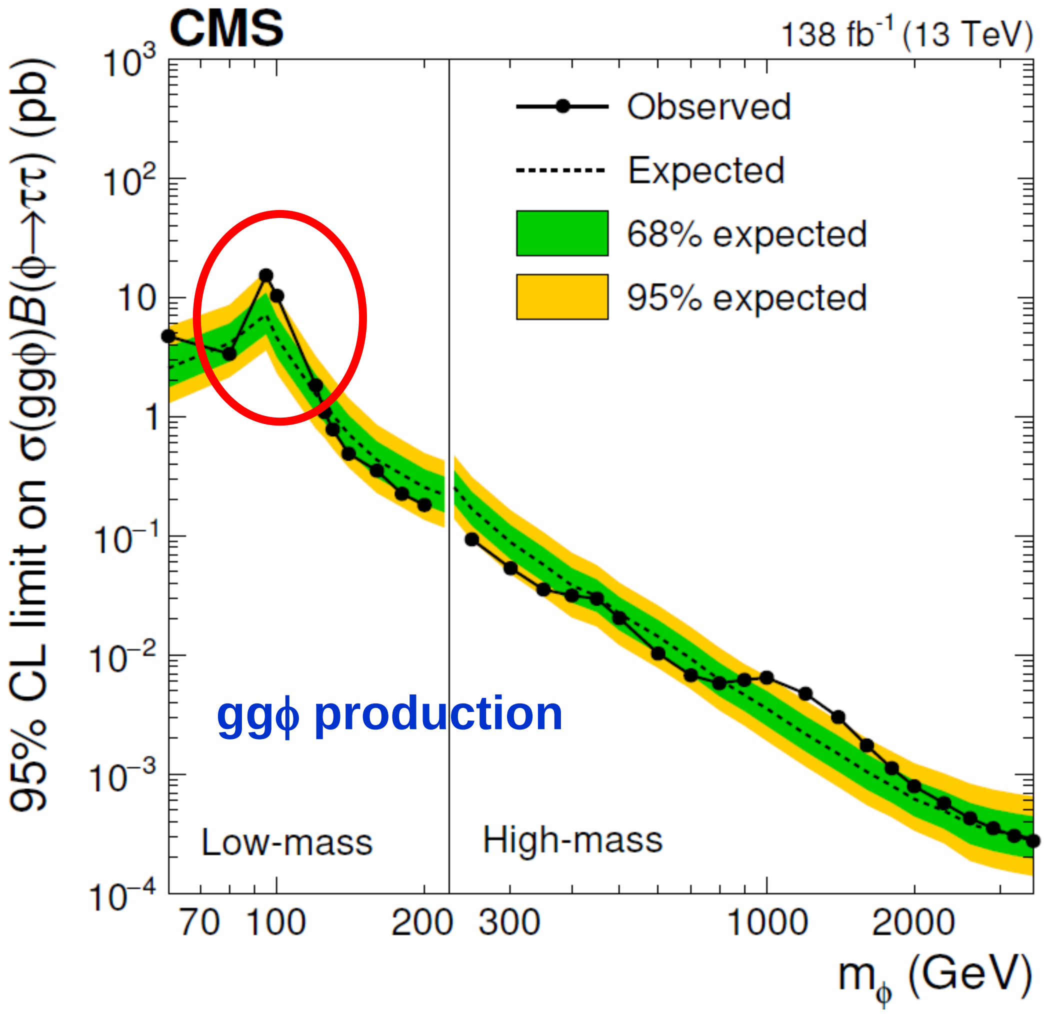
corresponding to an excess of 3.2σ at 95.4 GeV

$\phi \rightarrow \tau\tau$: results

JHEP07(2023)073

- "Low-mass" (60–250 GeV): fitting on $m_{\tau\tau}$ to extract signal





$2.8\sigma(2.2\sigma)$ local (global) @ 1.2 TeV

$3.1\sigma(2.7\sigma)$ local (global) @ 100 GeV

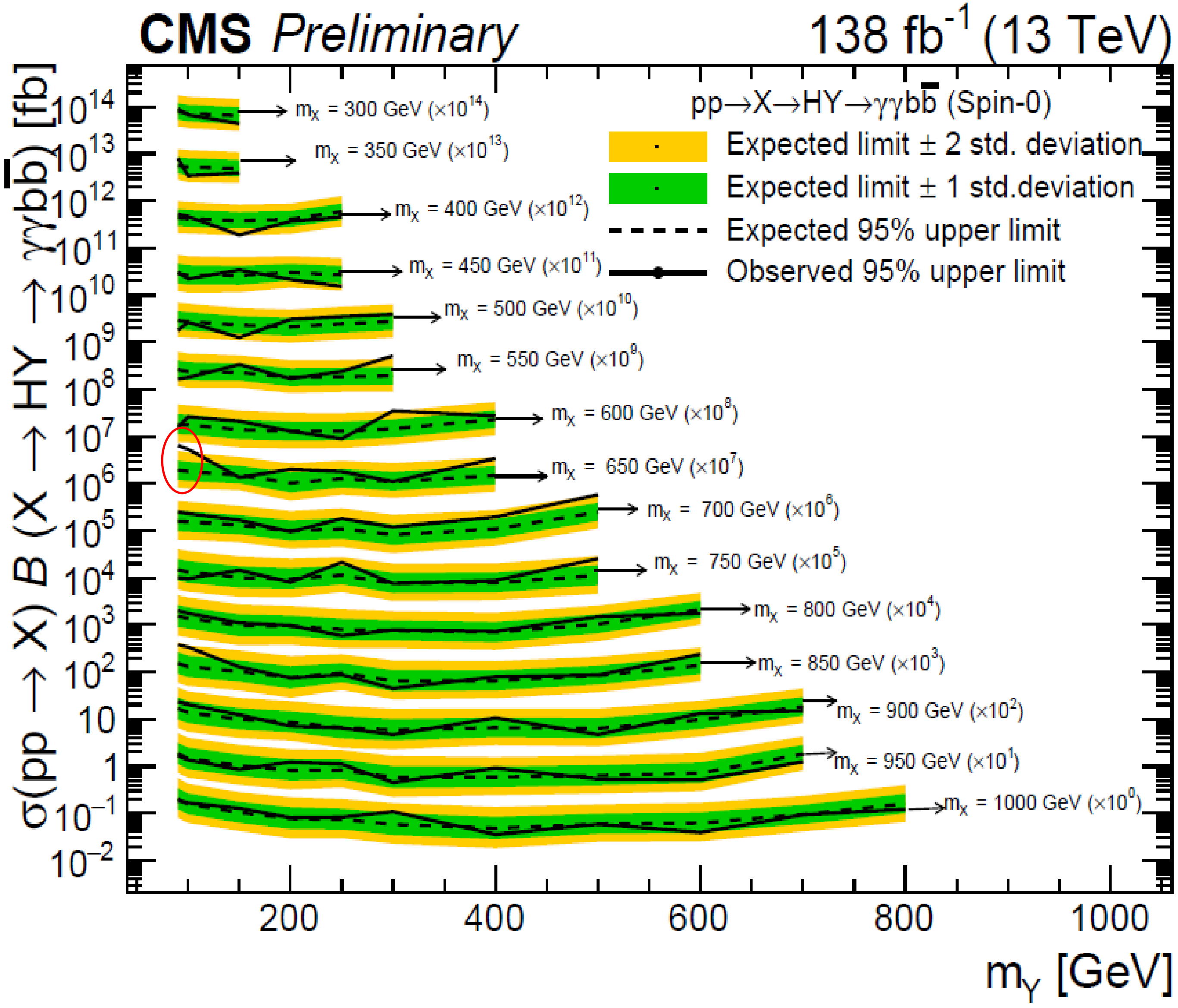
$\mu(gg \rightarrow h \rightarrow \tau\tau) = 1.2 \pm 0.5$

$2.6\sigma(2.3\sigma)$ local (global) @ 95 GeV

$X \rightarrow YH \rightarrow bb\gamma\gamma$

CMS-PAS-HIG-21-011

- BDT (NN) scores to separate signals and non-resonant (resonant) backgrounds
- Six BDT training accounts for different signal m_X-m_Y mass ranges
  - 3 event classes based on BDT output
- A parametric fit in the $(m_{\gamma\gamma}, m_{jj})$ plane is performed for signal extraction for each category



3.8σ local for m_X = 650 GeV and **m_Y = 90 GeV**

3.5σ local for m_X = 650 GeV and **m_Y = 100 GeV**

2.8σ global