Search for H→aa→4γ

Badder Marzocchi[1],
on behalf of the CMS Collaboration

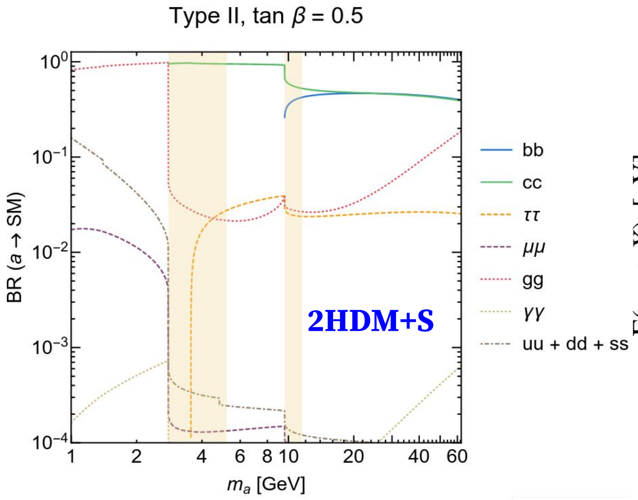
[1]Northeastern University (US)

LHCWG Workshop 2022 - CERN - November 28th-30th
28/11/2022

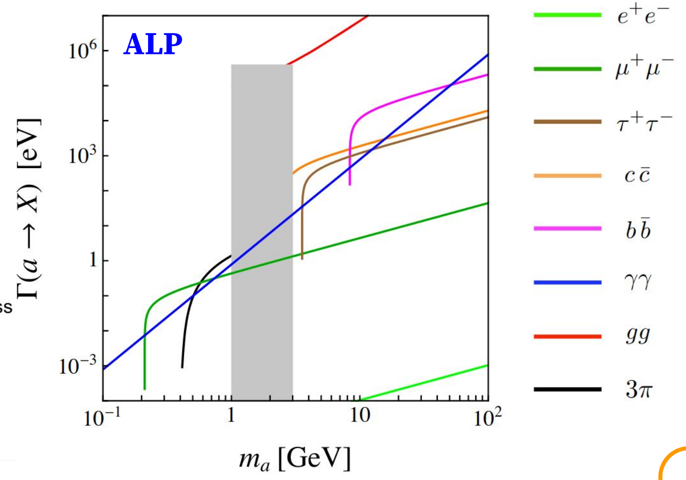
Higgs BSM decays

Higgs decay to BSM particles is possible from Run2 results:

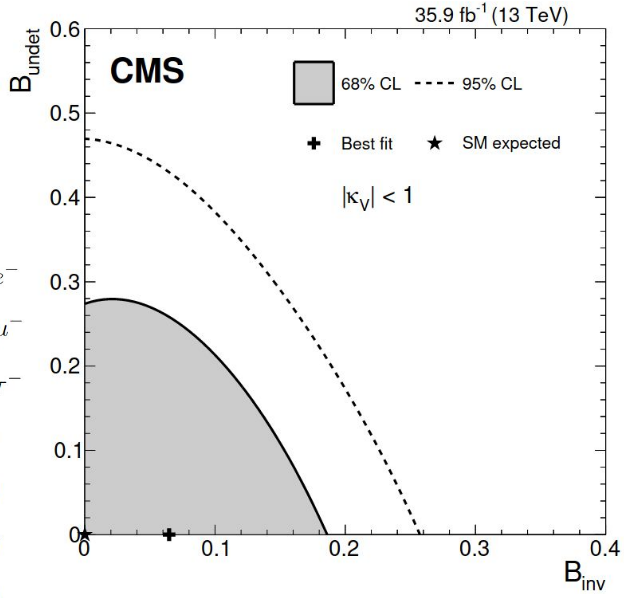
- Br_{BSM} < 20% (95% CL intervals)
- Given the small Γ_H= 4.1 MeV, even small couplings with BSM particles,~O(10^{-2}), yield Br(h→BSM)=10%



[Phys. Rev. D 90, 075004 (2014)]

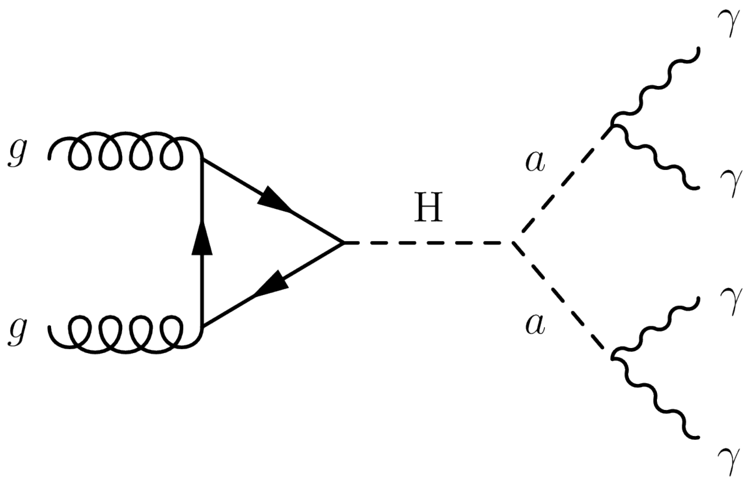


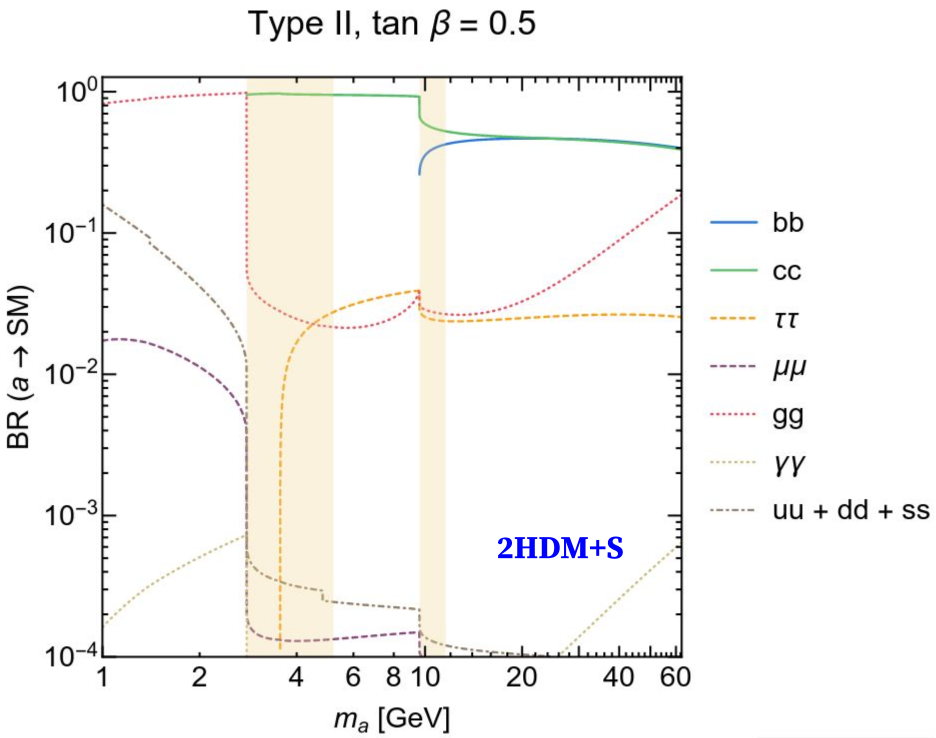
JHEP 12 (2017) 044



[Eur. Phys. J. C 79 (2019) 421]

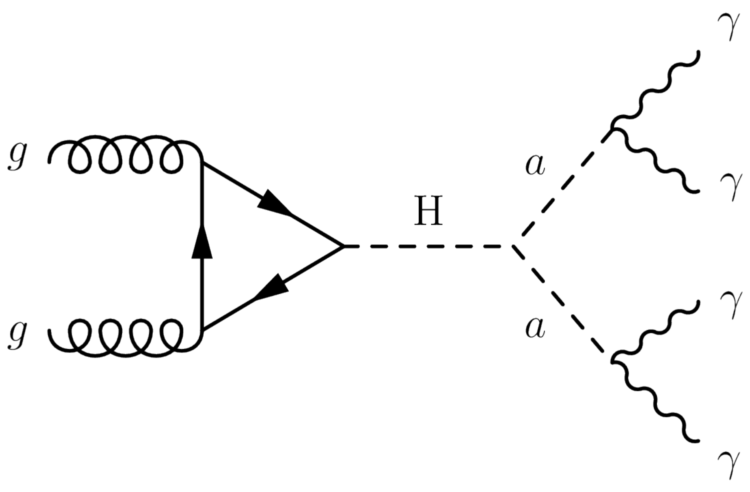
BSM models for H→aa:
- 2HDM+S models
- ALP models

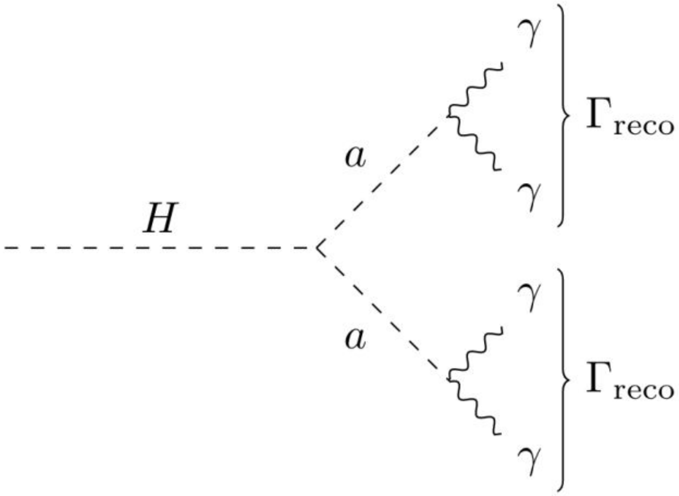




Fully resolved regime: 15 GeV ≤ m_a ≤ 62 GeV

- Low branching ratio: Br(a→γγ) ≈ 10^{-3}
- Clean and efficient reconstruction of isolated photons
- Low background, coming from 4 isolated photons



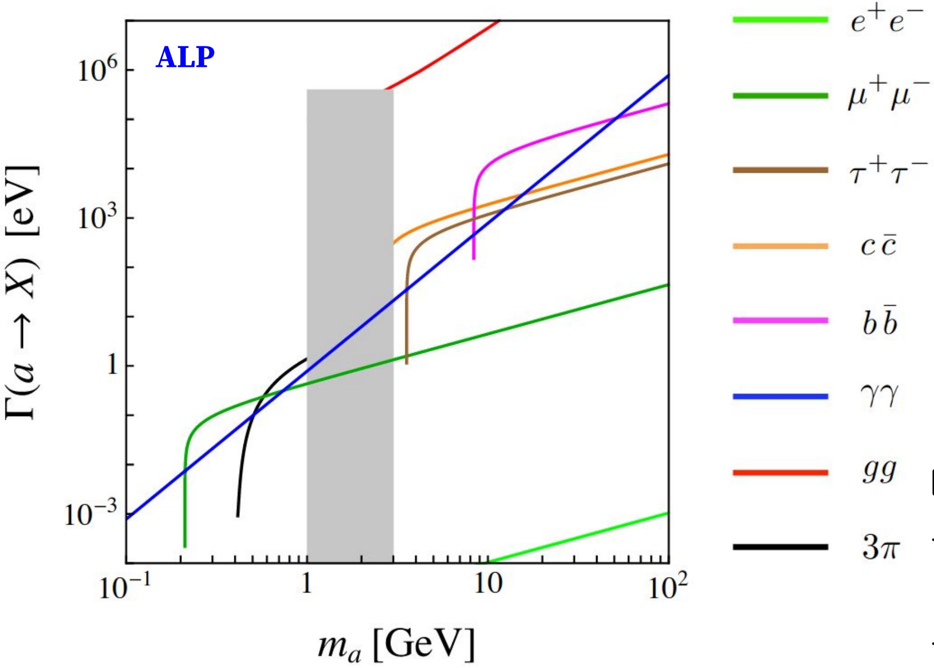


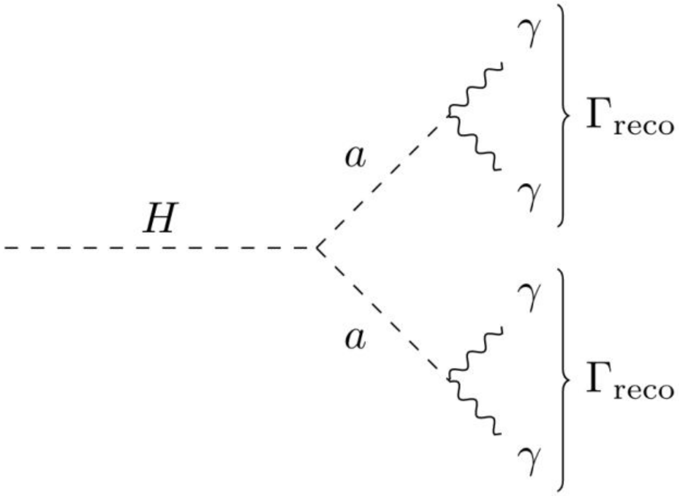
Fully resolved regime: 15 GeV ≤ m_a ≤ 62 GeV

- Low branching ratio: Br(a→γγ) ≈ 10^{-3}
- Clean and efficient reconstruction of isolated photons
- Low background, coming from 4 isolated photons

Fully boosted regime: 100 MeV ≤ m_a ≤ 1.2 GeV

- Below m_a< 1GeV, a→γγ becomes relevant, competing with a→3π for $m_a \gtrsim$ 250 MeV (QCD scale), and a→$\mu^+\mu^-$
- Challenging reconstruction → merged photon pairs (Γ)
- Main background coming from QCD (π^0→γγ)

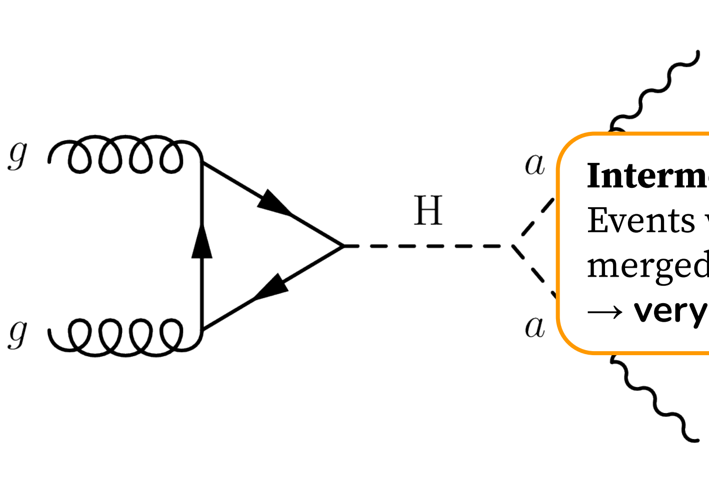




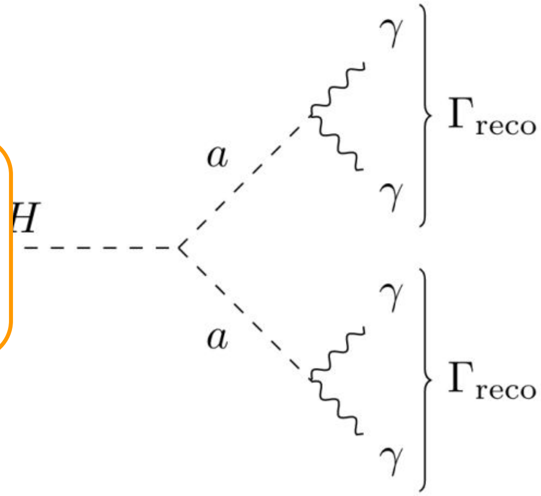
Fully boosted regime: 100 MeV ≤ m_a ≤ 1.2 GeV

- Below m_a< 1GeV, a→γγ becomes relevant, competing with a→3π for $m_a \gtrsim$ 250 MeV (QCD scale), and a→$\mu^+\mu^-$
- Challenging reconstruction → merged photon pairs (Γ)
- Main background coming from QCD (π^0→γγ)

H → aa → γγγγ



Intermediate regimes not addressed:
Events with one resolved pair and one merged pair difficult to reconstruct
→ **very low signal efficiency**



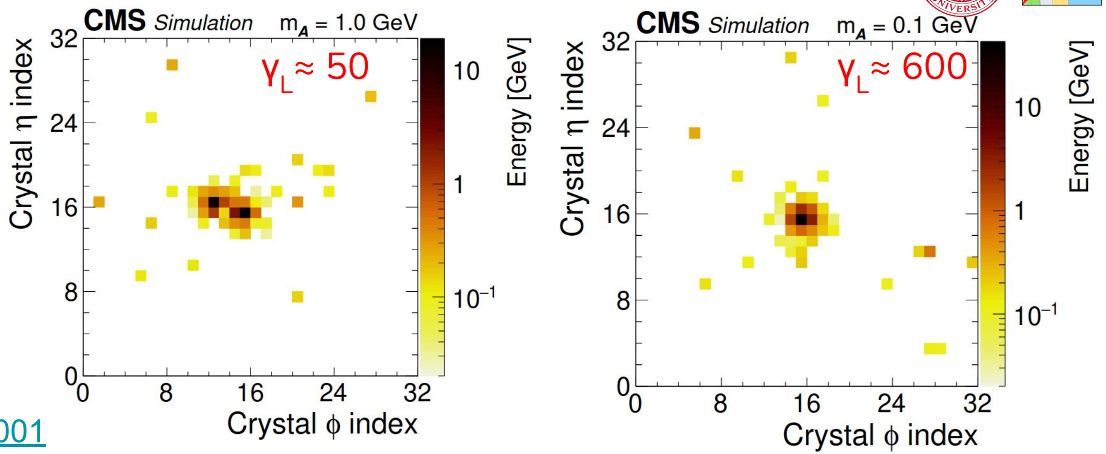
Fully resolved regime: 15 GeV ≤ m_a ≤ 62 GeV
- Low branching ratio: Br(a→γγ) ≈ 10^{-3}
- Clean and efficient reconstruction of isolated photons
- Low background, coming from 4 isolated photons

Fully boosted regime: 100 MeV ≤ m_a ≤ 1.2 GeV
- Below m_a< 1GeV, a→γγ becomes relevant, competing with a→3π for $m_a \gtrsim$ 250 MeV (QCD scale), and a→$\mu^+\mu^-$
- Challenging reconstruction → merged photon pairs (Γ)
- Main background coming from QCD (π^0→γγ)

Boosted Regime

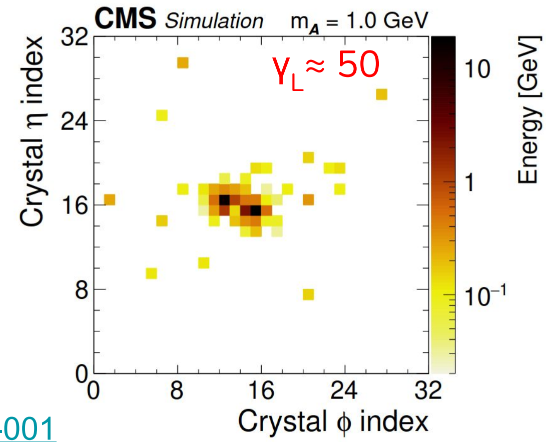
For m_a < 0.4 GeV, Lorentz boost (γ_L) > 150:
$\Delta R(\gamma\gamma)$ < ECAL crystal Molière radius →
γγ reconstructed as one photon (Γ)

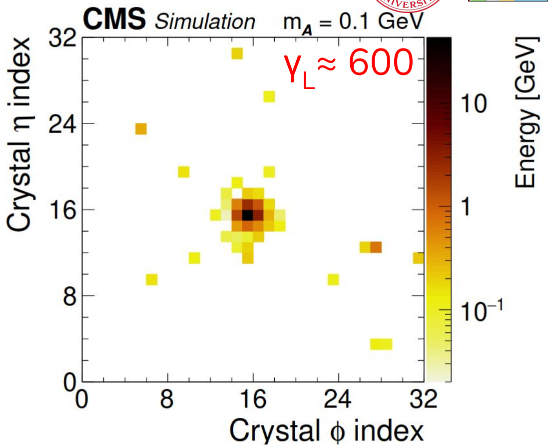
CMS-EGM-20-001



Boosted Regime

For m_a < 0.4 GeV, Lorentz boost (γ_L) > 150:
ΔR(γγ) < ECAL crystal Molière radius →
γγ reconstructed as one photon (Γ)

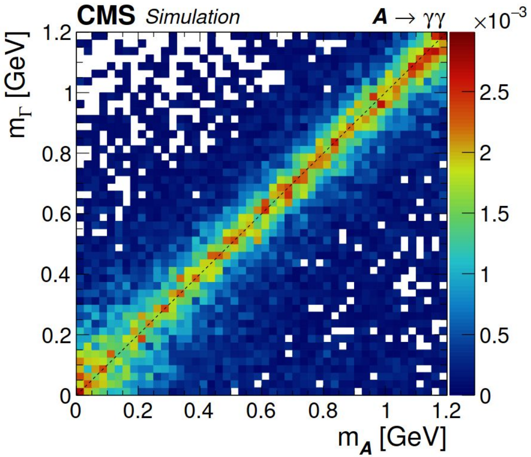


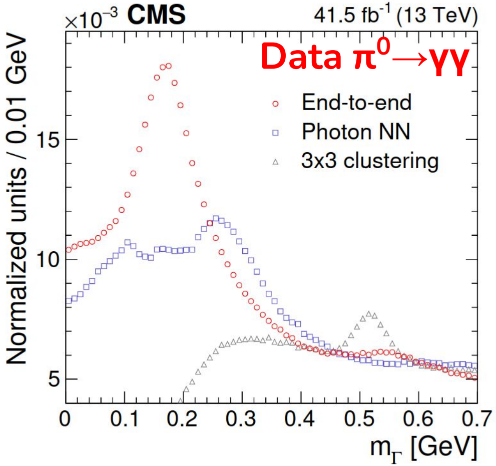


CMS-EGM-20-001

ML end-to-end used to reconstruct Γ:

- CNN architecture: regression algorithm using shower patterns at crystal level
- Training performed using signals A→γγ with p_T=20-100 GeV and m_A=0,1.6 GeV, i.e. 10 < γ_L < 1000





Boosted Analysis: Strategy CMS-HIG-21-016

Data High Level Triggers:

- At least 2 isolated photons with p_T>30 GeV and 18 GeV (2016) or p_T>30 GeV and 22 GeV (2017-2018)
- $m_{\gamma\gamma}$> 90 GeV

Event selections:

- No more than three reconstructed photons
- Exactly 2 photons (Γ_1, Γ_2 merged pair candidates) in barrel ($|\eta_\Gamma|$ < 1.4), passing optimized losse photon identification
- 100 GeV < $m_{\Gamma\Gamma}$< 180 GeV
- $p_T^{\Gamma1}/m_{\Gamma1}$ > ⅓ and$p_T^{\Gamma2}/m_{\Gamma2}$ > ¼

Boosted Analysis: Strategy CMS-HIG-21-016

Data High Level Triggers:

- At least 2 isolated photons with p_T>30 GeV and 18 GeV (2016) or p_T>30 GeV and 22 GeV (2017-2018)
- $m_{\gamma\gamma}$> 90 GeV

Event selections:

- No more than three reconstructed photons
- Exactly 2 photons (Γ_1, Γ_2 merged pair candidates) in barrel ($|\eta_\Gamma| < 1.4$), passing optimized losse photon identification
- 100 GeV < $m_{\Gamma\Gamma}$< 180 GeV
- $p_T^{\Gamma 1}/m_{\Gamma 1}$ > ⅓ and$p_T^{\Gamma 2}/m_{\Gamma 2}$ > ¼

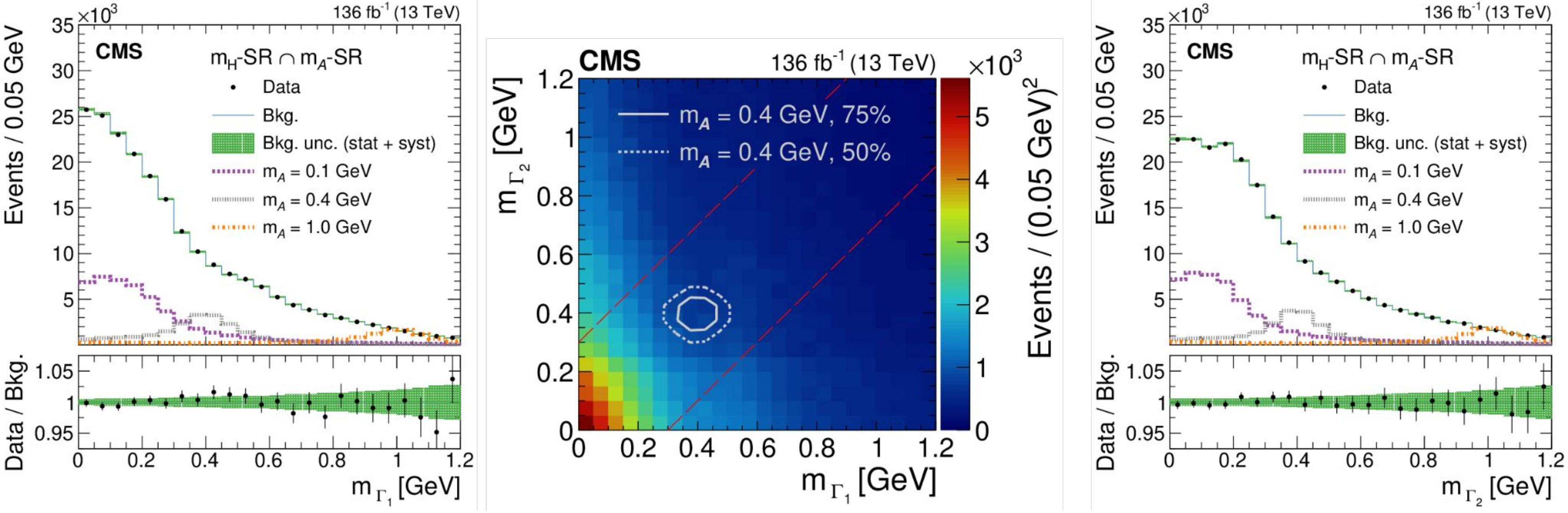
Signal and background model templates:

- Built on the 2D $m_{\Gamma 1}$,$m_{\Gamma 2}$ distribution
- **Signal model from simulation**
- **Background model:**
  - →SM H→γγ simulation
  - →Non resonant: weighted templates from data in the sideband regions

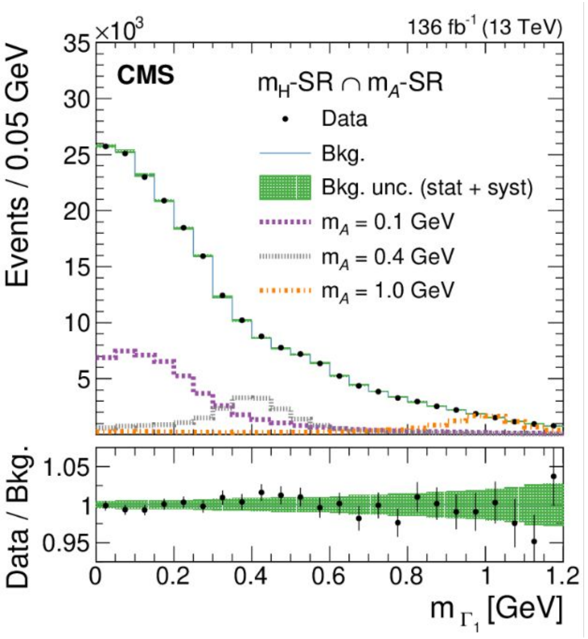
    m_H-SB: 100 GeV < $m_{\Gamma\Gamma}$< 110 GeV or 140 GeV < $m_{\Gamma\Gamma}$< 180 GeV

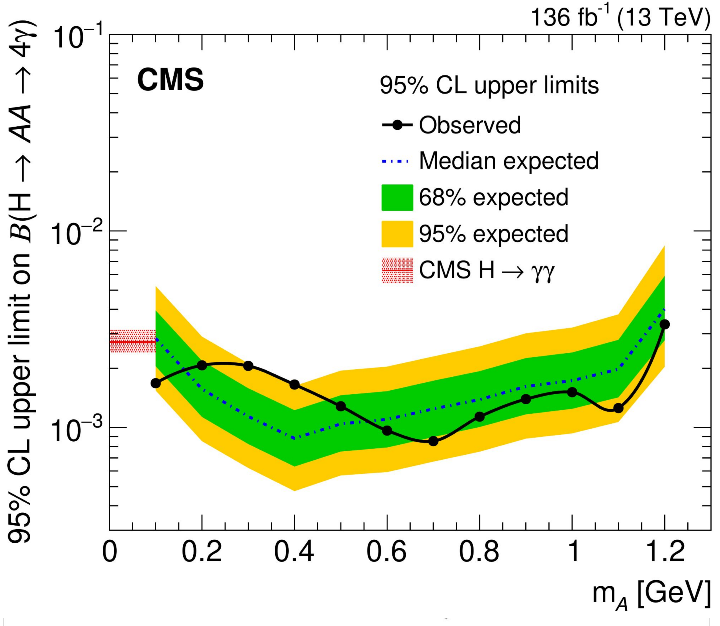
    m_A-SB: $|m_{\Gamma 1}$- $m_{\Gamma 2}|$>0.3 GeV

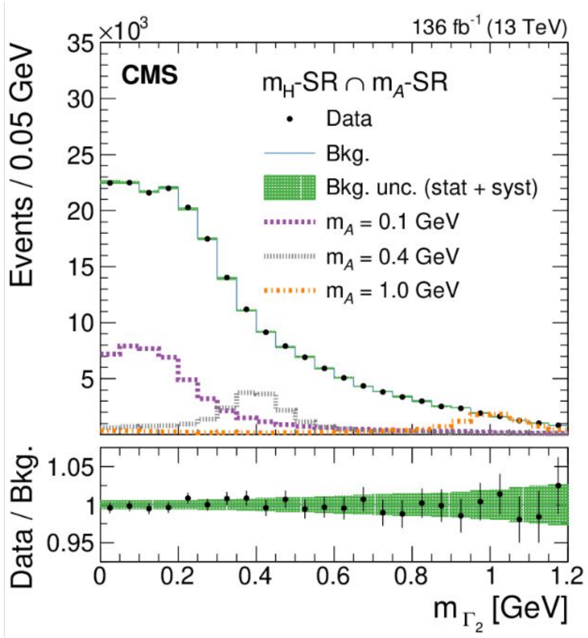
Boosted Analysis: Results



Boosted Analysis: Results







Results:

- Dominated by the statistical uncertainty on background modelling
- No excess and observed limits are in agreement with the expected limits

Resolved Analysis: Strategy

CMS-HIG-21-003

Data High Level Triggers:

- At least 2 isolated photons with p_T>30 GeV and 18 GeV
- $m_{\gamma\gamma}$> 55 GeV (only for 2016-2017)

Event selections:

- Standard $H\rightarrow\gamma\gamma$ preselections on highest-p_T diphoton pair
- At least 4 well isolated photons with:
  $p_T^{\gamma1}$>30 GeV, $p_T^{\gamma2}$>18 GeV (cope with trigger),
  $p_T^{\gamma3}$>15 GeV, $p_T^{\gamma4}$>15 GeV (cope with photon identification)
  $|\eta|$ < 2.5 and passing the electron-veto
- 110 GeV < $m_{\gamma\gamma\gamma\gamma}$< 180 GeV

Data High Level Triggers:

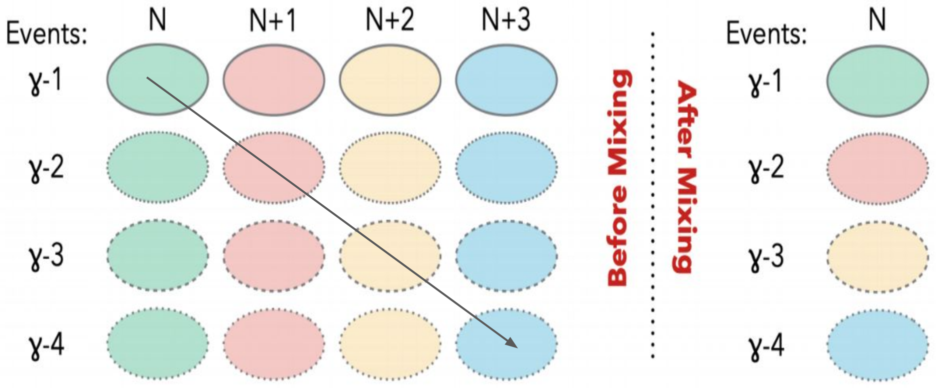
- At least 2 isolated photons with p_T>30 GeV and 18 GeV
- $m_{\gamma\gamma}$> 55 GeV (only for 2016-2017)

Event selections:

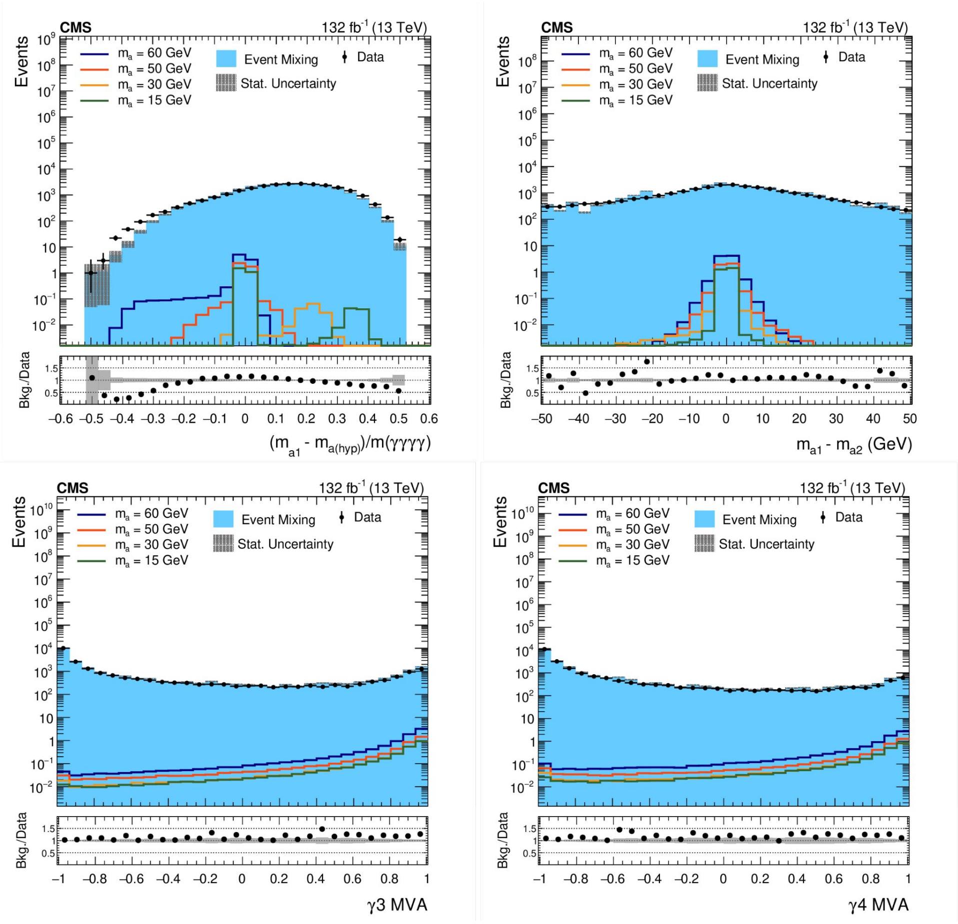
- Standard H→γγ preselections on highest-p_T diphoton pair
- At least 4 well isolated photons with:
  $p_T^{\gamma 1}$>30 GeV, $p_T^{\gamma 2}$>18 GeV (cope with trigger),
  $p_T^{\gamma 3}$>15 GeV, $p_T^{\gamma 4}$>15 GeV (cope with photon identification)
  |η| < 2.5 and passing the electron-veto
- 110 GeV < $m_{\gamma\gamma\gamma\gamma}$< 180 GeV

4γ event classifier:

- Improve signal/background discrimination
- Not enough MC background events →
  Background estimation from data with event-mixing:
  →Second, third, fourth photon in each event replaced with photons from the next three consecutive events
  →Improve data/event-mixing agreement with reweighting from $m_{\gamma\gamma\gamma\gamma}$ sidebands



Resolved Analysis: 4γ event classifier

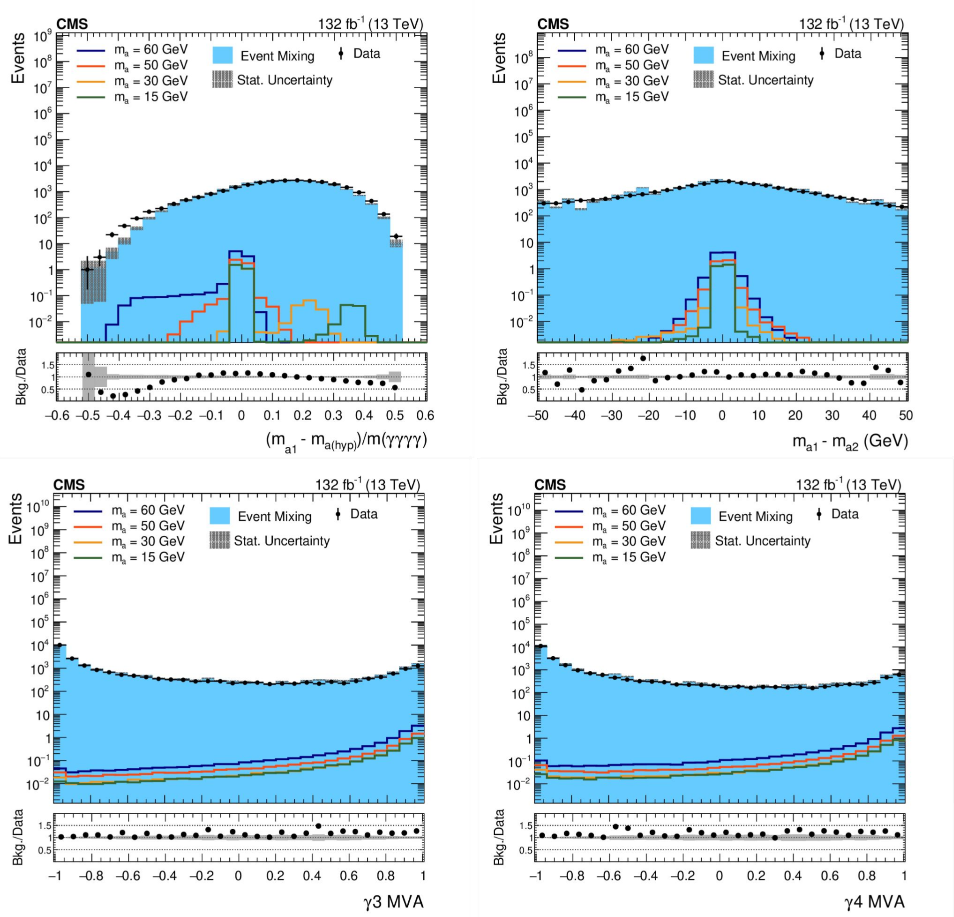


Standard MVA BDT

- **Parameterized training**:
  Train all the signals together, taking as an input signal mass-hypothesis ($m_{a,hyp}$)
  → **Uniform output and sensitive to full m_a range**

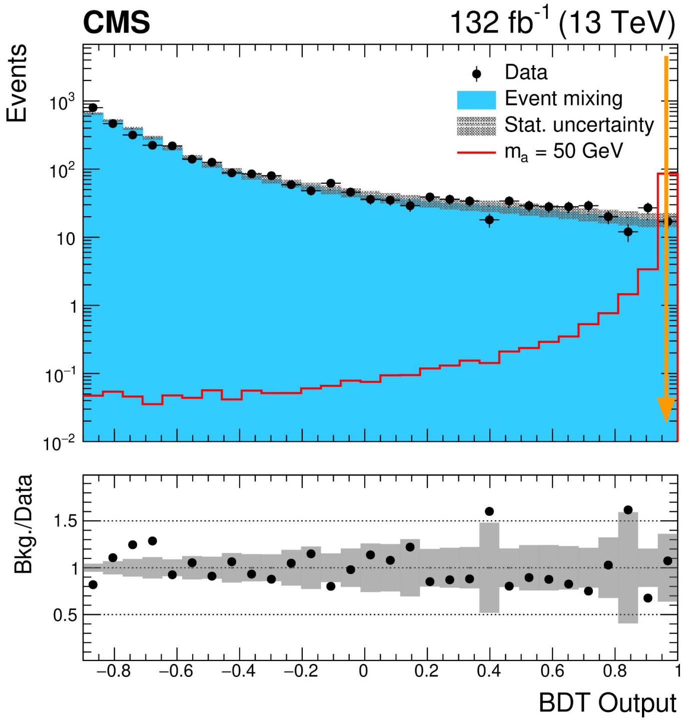
- **Input variables:**
  → Photon Identification MVA, for each photon
  → p_T^{a1} and p_T^{a2}
  → $(m_{a1} - m_{a2})$
  → $(m_{a1} - m_{a,hyp})/m_{\gamma\gamma\gamma\gamma}$
  → $\Delta R_{a1,a2}/m_{\gamma\gamma\gamma\gamma}$
  → $\cos\theta^*_{a,\gamma}$

Resolved Analysis: 4γ event classifier



BDT score selection:

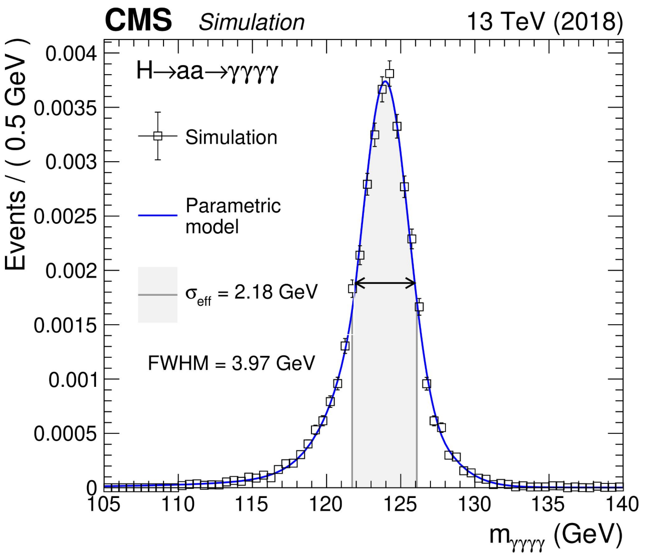
- For each m_a hypothesis, optimized by maximizing $S/\sqrt{B}$

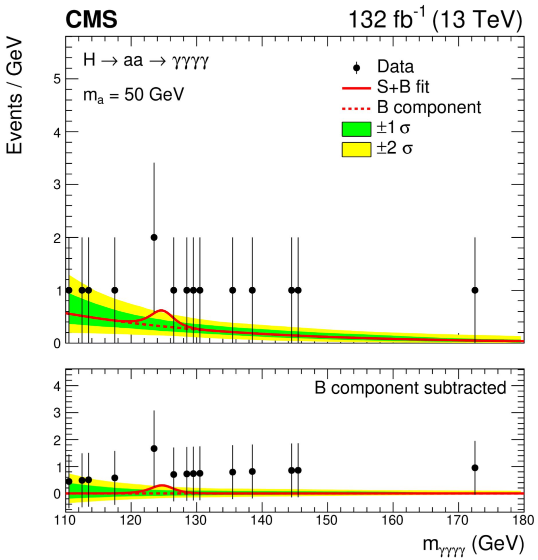


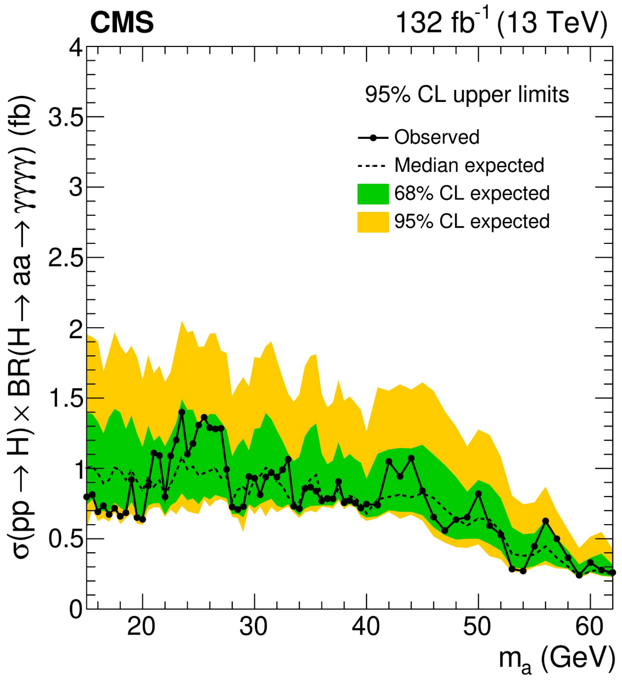
Resolved Analysis: Results

Signal model:
- Built from MC for each nominal m_a
- Modelled using double sided crystal ball function
- Interpolation performed in 1GeV step

Background model:
- Built from selected data (3 years merged) for each nominal m_a
- Modelled using Envelope method





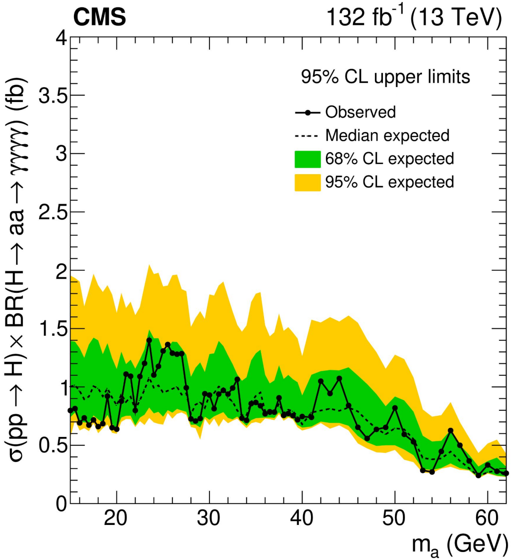


Results:
- Dominated by the statistical uncertainty on background modelling
- No excess and observed limits are in agreement with the expected limits

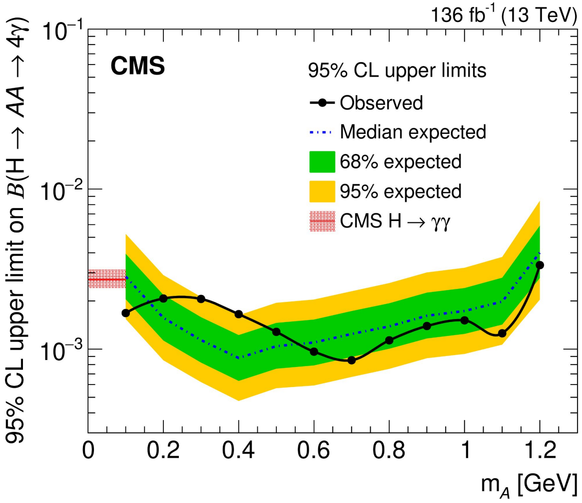
Summary

ATLAS results with 8 TeV data (20.3 fb^{-1}) for 10 GeV < m_a < 62 GeV:
EPJC 76 (2016) 210

CMS Full Run2 results on 13 TeV data:
- **Fully resolved analysis** for 15 GeV < m_a < 62 GeV: CMS-HIG-21-003
- **Fully boosted analysis** for 100 MeV < m_a < 1.2 GeV: CMS-HIG-21-016
  Novel ML end-to-end technique used to reconstruct merged photon pairs
- No excess observed



Novel CNN regression:
- Exploit shower patterns at crystal level
- Proven to be effective to reconstruct merged photon pairs: $A\to\gamma\gamma$, $\pi^0\to\gamma\gamma$, $\eta\to\gamma\gamma$,...
- Improvement under development:
  →Events in the endcaps
  →Use tracker information



Backup

$m_{\gamma\gamma\gamma\gamma}$ distribution shape	2016–2018
Photon energy scale and resolution	0.05–0.15%
Nonlinearity of the photon energy scale	0.10%
Shower shape corrections	0.01–0.15%
Nonuniformity of light collection	0.07–0.25%
Modeling of material in front of the ECAL	0.02–0.05% (EB) and 0.24% (EE)

Signal model normalization	2016	2017	2018
Integrated luminosity	1.20%	2.30%	2.50%
Photon identification	0.25%	0.25%	0.25%
Trigger efficiency	0.50%	1.50%	0.50%
Photon preselections	5.00%	5.00%	5.00%

Boosted Analysis: Systematics Relative uncertainty (%)

Background:

- Bin-by-bin statistical uncertainties, bkg model: 2-13
- Bkg p_T-reweighting: <1
- Fraction of m_H-SB_{low} template: $\lesssim$1
- Fraction of SM $H\rightarrow\gamma\gamma$ template: <1

Signal:

- Bin-by-bin statistical uncertainties, sg model: 1-4
- m_Γ regressor mass scale: $\lesssim$26
- m_Γ regressor mass smearing: <2
- LHC integrated luminosity: $\approx$1
- Photon ID MC scale factors: $\approx$1
- HLT trigger MC scale factors: <1