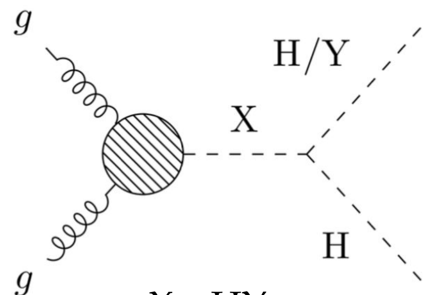
BSM H(125) aspects and searches for new scalar bosons - CMS

Siddhesh Sawant (Baylor University),
on behalf of CMS collaboration

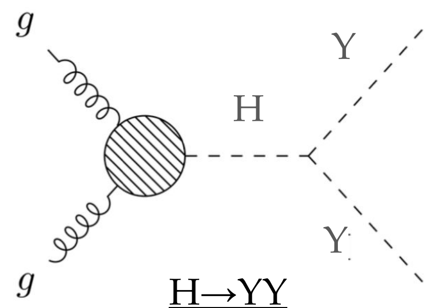
Higgs Hunting 2025 - July 15-17, 2025

Introduction

- The standard model: the most successful so far, but not complete.
  Open questions: the hierarchy problem, origin of dark matter etc
- Many Beyond SM (BSM) theories provides explanation of shortcomings of SM, and postulates new particles
  - 2HDM, MSSM: CP-even h, H, $H^{\pm}$ and CP-odd a
  - 2HDM + singlet, NMSSM: CP-even $h_{1,2,3}$, $H^{\pm}$ and CP-odd A, a
  - Two Real Singlet Models
  - …
- New particles (X, Y) in several appealing BSM theories.
  X, Y: scalar or pseudoscalar
  - X→HH/HY/YY
  - H→YY
- Higher sensitivity when search involves H(125)



X→HY



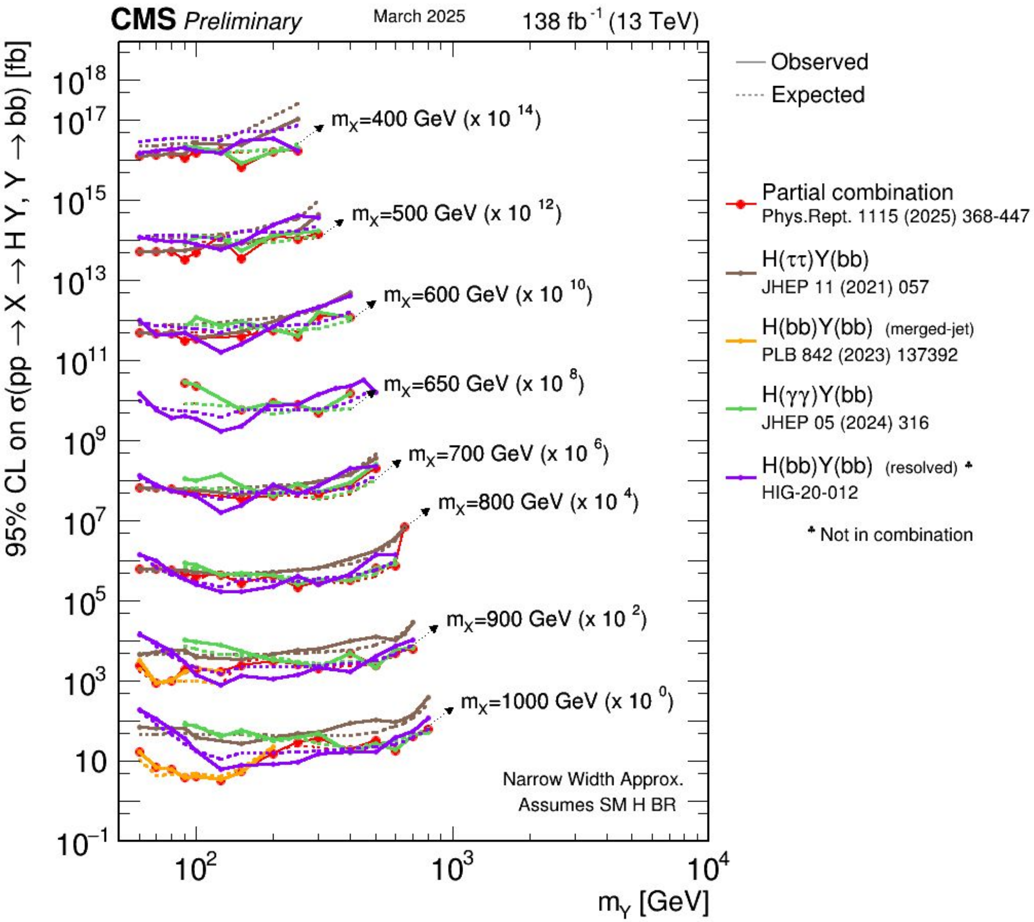
H→YY

Searches for X→H(125)Y at CMS with full Run 2 data

X→H(125)Y

- X→H(bb)Y(bb) [CMS-PAS-HIG-20-012]
- X→H(bb)Y(bb) (boosted) [Phys.Letters.B 842(2023)137392]
- X→H($\tau\tau$)Y(bb) [JHEP11(2021)057]
- X→H($\gamma\gamma$)Y(bb) [JHEP05(2024)316]
- X→H($\gamma\gamma$)Y($\tau\tau$) [CMS-PAS-HIG-22-012]
- X→H(bb)Y($\gamma\gamma$) † [CMS-PAS-B2G-24-001]
- X→H(bb)Y(anomalous) † [CMS-PAS-B2G-24-015]
- X→H(bb)Y(E_T^{miss}) † [CMS-PAS-SUS-24-007]
- X→H(bb)Y(4q) † [CMS-PAS-B2G-23-007]

All X→H(125)Y cover X→H(125)H(125)

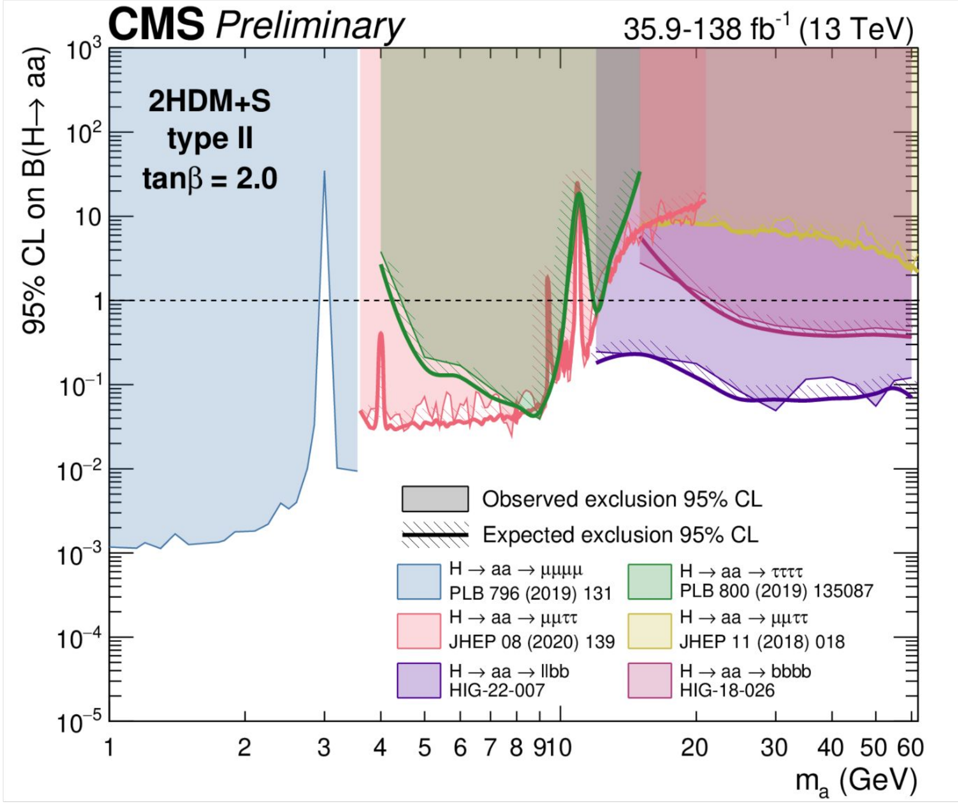


† Scope of this talk. New results since HHunting 2024

Searches for H(125)→aa at CMS with full Run 2 data

H(125)→aa

- H→aa→4γ (boosted) [Phys. Rev. Lett. 131 (2023) 101801]
- H→aa→4γ (resolved) [JHEP 07 (2023) 148]
- H→aa→$\mu\mu\tau\tau$ (boosted) [JHEP11(2021)057]
- H→aa→$\mu\mu\tau\tau$ (resolved) [JHEP11(2018)018]
- H→aa→$\mu\mu$bb/$\tau\tau$bb [Eur. Phys. J. C 84 (2024) 493]
- H→aa→4b [JHEP06(2024)097]
- H→aa→4μ [JHEP12(2024)172]
- H→aa→4τ [CMS-PAS-SUS-24-002]

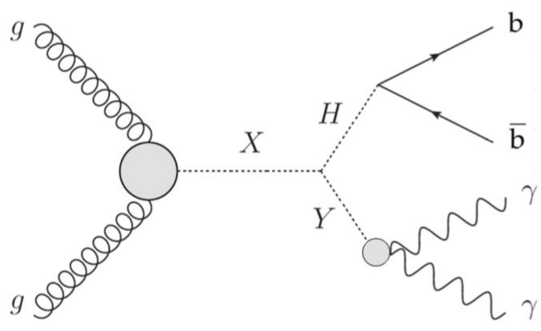


All searches covered in HHunting 2024 or earlier

X → H(bb) Y(γγ)

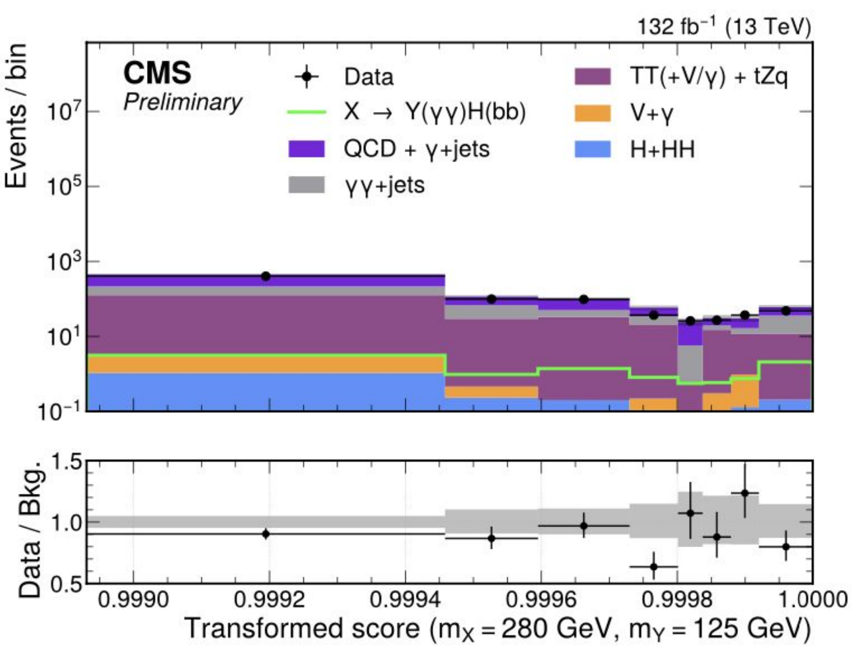
CMS-PAS-B2G-24-001

- Search for X→H(bb) Y(γγ), $m_X \in$ [240, 1000] GeV, $m_Y \in$ [70, 800] GeV,
  Search in model-independent way.
  X,Y: Narrow width scalars.



- Selection conditions (in brief):
  - Events with 2γ (p_T> 30, 18 GeV) and 2 AK4 jets (p_T> 24 GeV) passing b-quark tagging requirements.
  - Veto events with e or μ

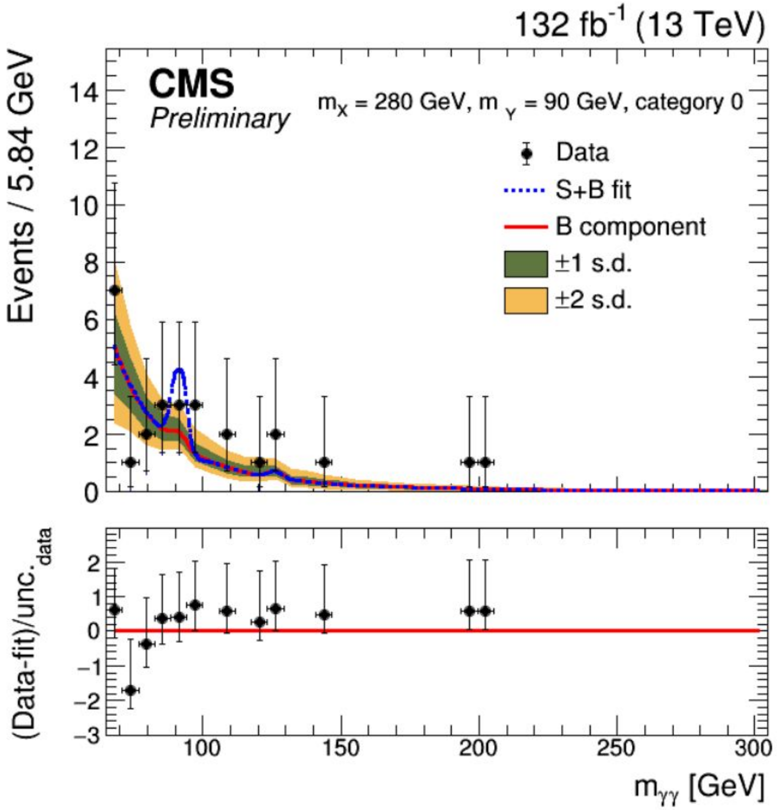
- Trained signal-hypothesis-aware Parametric Neural Network to achieve higher S/B in signal region
  - NN scored is transformed to have flat background



Neural network scored

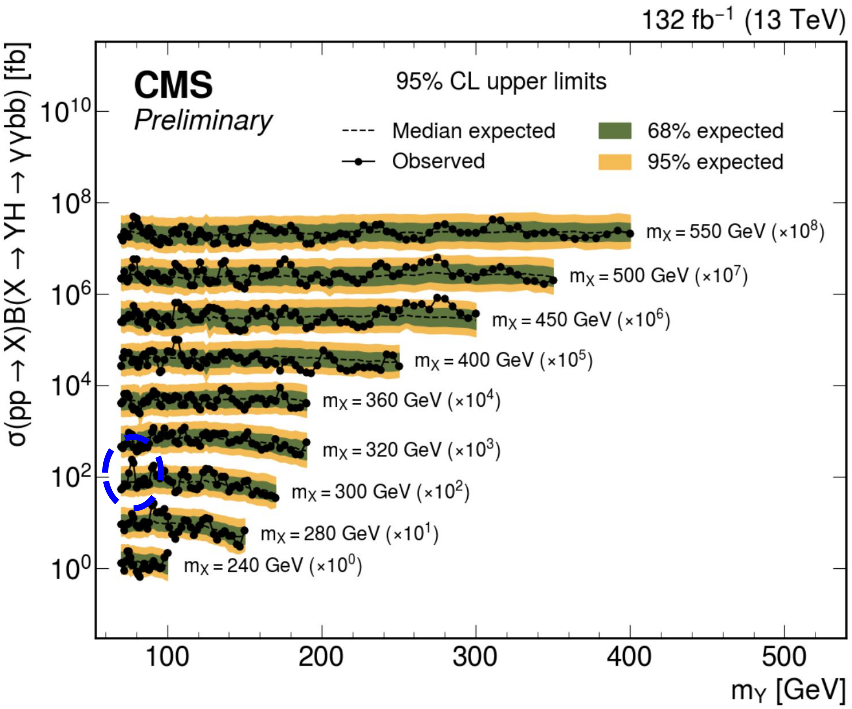
$X \rightarrow H(bb)\ Y(\gamma\gamma)$ (II)

- $m_{\gamma\gamma}$ is used for signal extraction
- Signal: Modeled using Double Crystal-Ball (DCB) function
- Resonant $H \rightarrow \gamma\gamma$ background from H, HH production: Also modeled by DCB function
- Resonant DY background: Derived using a data-driven ABCD method
- Non-resonant background: Modeled by a smooth falling function, selected using discrete profiling procedure.

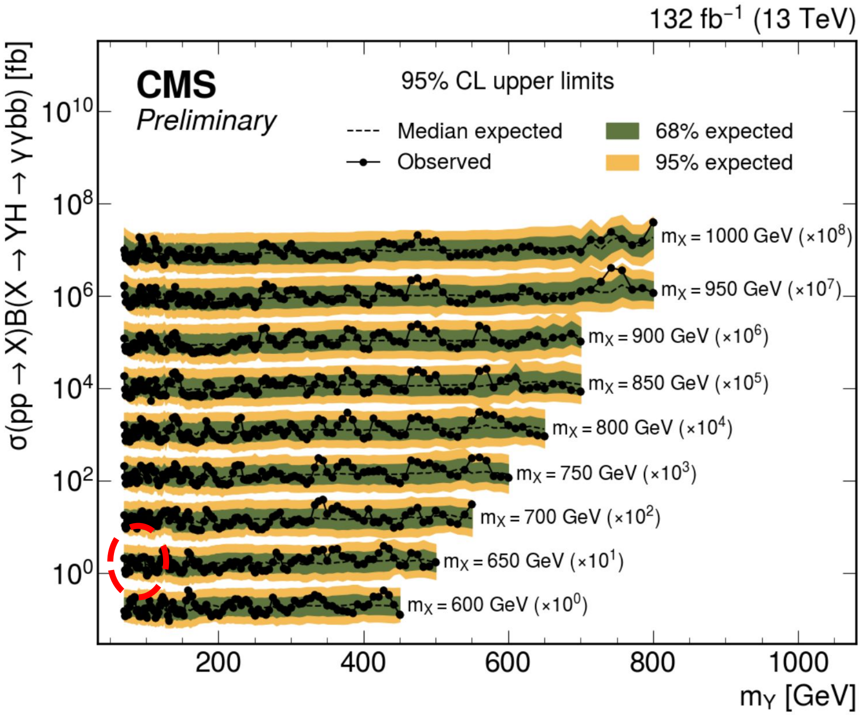


Y candidate mass

$X \rightarrow H(bb)\ Y(\gamma\gamma)$ (III)



Limits on m_Y for $240 \leq m_X \leq 550$ GeV



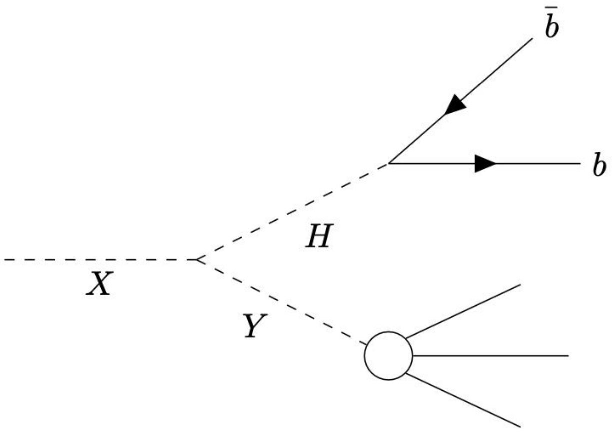
Limits on m_Y for $600 \leq m_X \leq 1000$ GeV

- Upper limit on $X \rightarrow H(bb)\ Y(\gamma\gamma)$: 0.05 to 2.69 fb
- Higher (lower) limits for mass points with lower (higher) m_X-m_Y difference
- The data is compatible with SM.
  Largest local (global) excess of 3.3 (0.6) σ significance at m_X=300 GeV, m_Y=77 GeV
- Local (global) excess of 3.8 (2.6) σ significance at m_X=650 GeV, m_Y=90 GeV reported in JHEP05(2024)316 is not confirmed by the current analysis

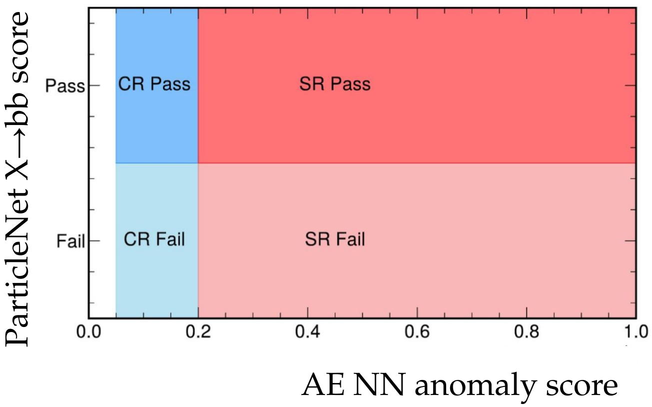
$X \to H(bb)\ Y(\text{anomalous})$

CMS-PAS-B2G-24-015

- Search for X→H(bb) Y(any→jets) in boosted regime,
  $m_X \in [1400, 3000]$ GeV, $m_Y \in [90, 400]$ GeV.



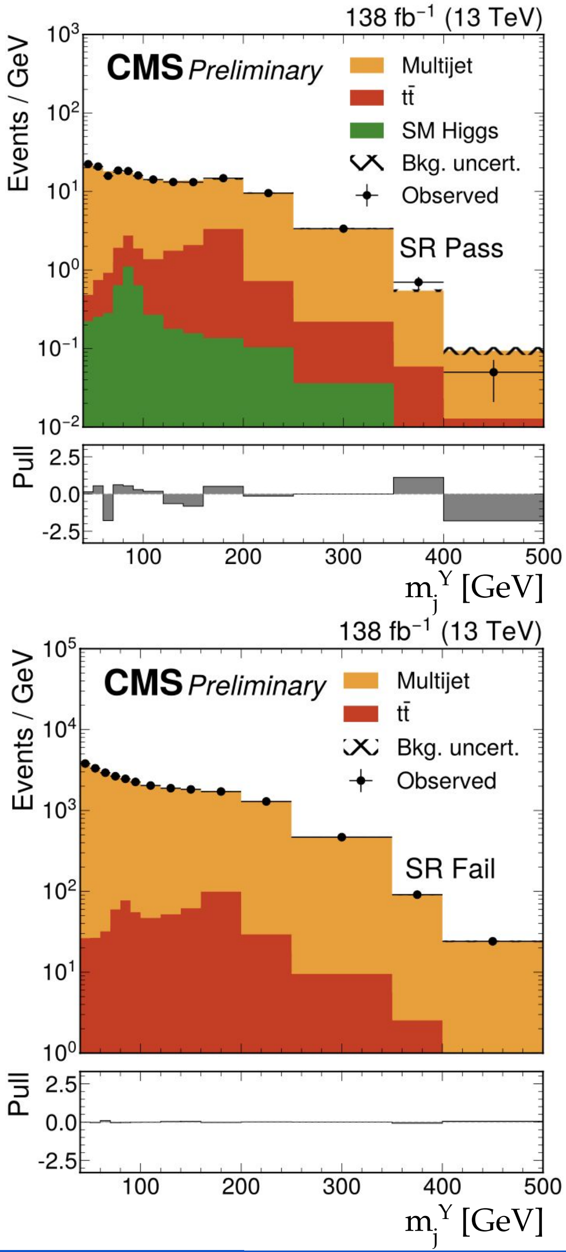
- Selection conditions (in brief):
  - Events with 2 AK8 jets (p_T> 300 GeV) with m_{jj} > 1300 GeV
  - One of the AK8 jet pass ParticleNet X→bb selection and m_j in 100-150 GeV
  - Other AK8 jet pass AutoEncoder Neural Network (AE NN) anomaly score threshold
- AE NN: Trained on QCD jets.
  - Non-QCD jets (single jet, jets from tt etc) gives higher anomaly score.
  - ≈30% mistag rate in signal region (SR)
- SR and control region (CR) defined using AE NN anomaly score.



SRs and CRs

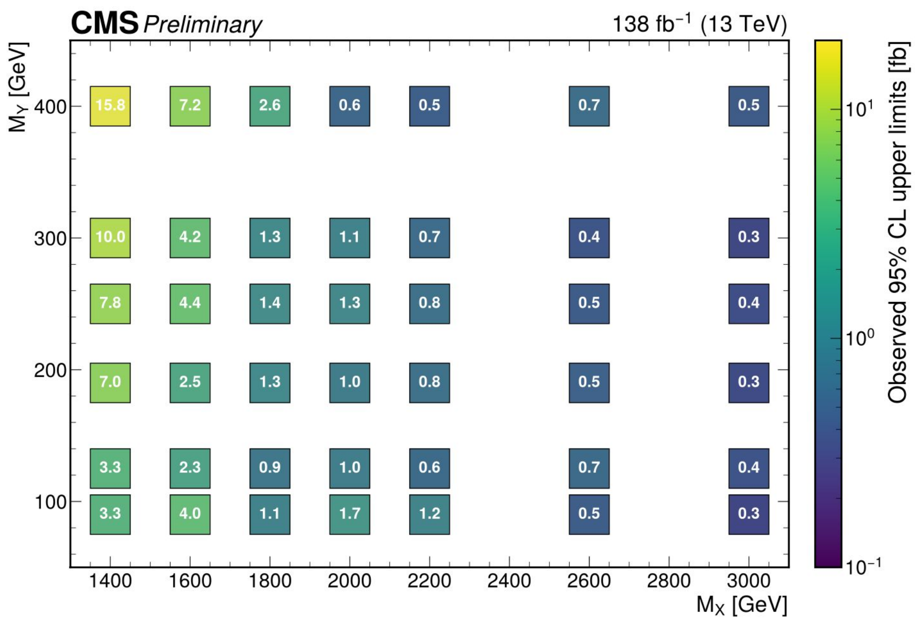
X → H(bb) Y(anomalous) (II)

- ParticleNet X→bb efficiency is calibrated in data using g→bb sample
- AE NN anomaly score cut efficiency is calibrated using the Lund plane reweighting method
- 2D m_{jj}^{X} - m_{j}^{Y} plane is used for signal extraction
  - Signal and tt, Higgs production backgrounds are modeled from MC
  - Nonresonant background is estimated in data from its contribution in 'fail' region scaled by 'pass-to-fail ratio' ($R_{P/F}$)
  - $R_{P/F}$: Polynomial in m_{jj}^{X} and m_{j}^{Y}, Determined during the fit to data

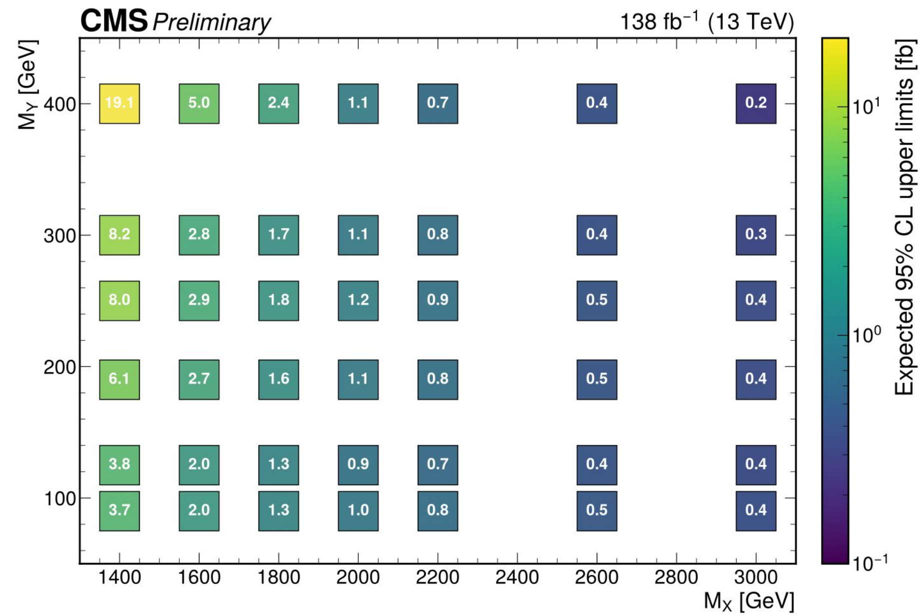


Plots: Y candidate mass in 'SR pass' and 'SR fail'

X → H(bb) Y(anomalous) (III)



Observed limits

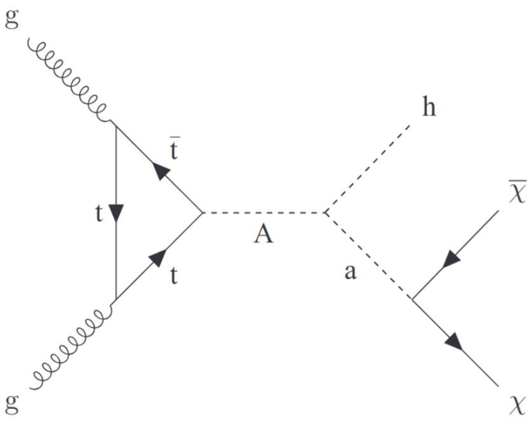


Expected limits

- Upper limits for hadronic Y→WW scenario: 0.3 to 19 fb
- Higher limits when m_X is low and m_Y is high, due to lesser signal reconstructed in the current Lorentz-boosted regime
- The data is compatible with SM.
  Largest local (global) excess of 2.1 (0.1) σ significance at m_X=1600 GeV, m_Y=90 GeV

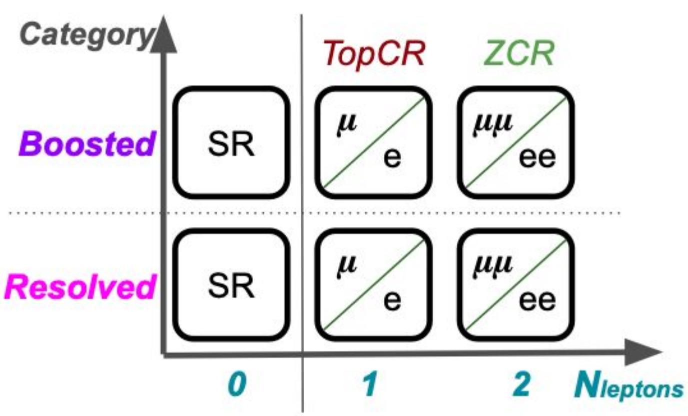
$X \rightarrow H(bb)\ Y(E_T^{miss})$

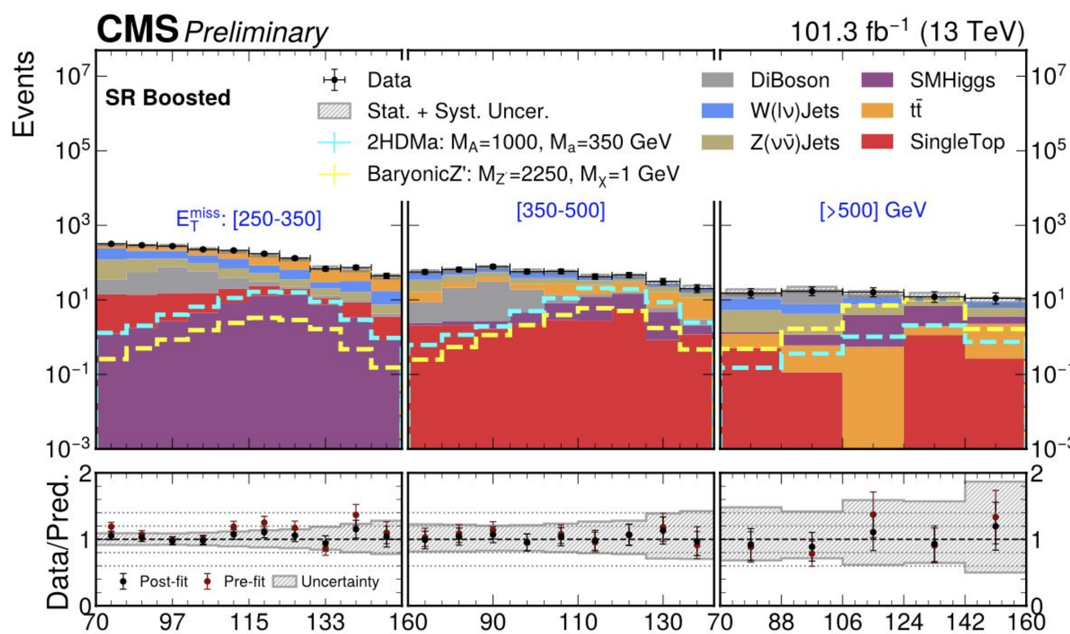
[CMS-PAS-SUS-24-007](CMS-PAS-SUS-24-007)



- Search for X→H(bb) Y(MET).
  2HDM+a model: A→H(bb) a(χχ), m_χ=10 GeV
  Baryonic-Z′ model: Z′→H(bb) Z′(χχ)
- Selection conditions (in brief): Events with
  - Higgs candidate:
    - Boosted cat.: 1 AK8 jets (p_T> 200 GeV, $m_{soft\text{-}drop}$: 70-160 GeV)
    - Resolved cat.: 2 b-tagged AK4 jets (pT> 50, 30 GeV, m_{jj}: 70-160 GeV)
  - E_T^{miss}: 250 (200) GeV for boosted (resolved) cat.

- Top and Z control regions for tt and Z(νν)+jets background estimation

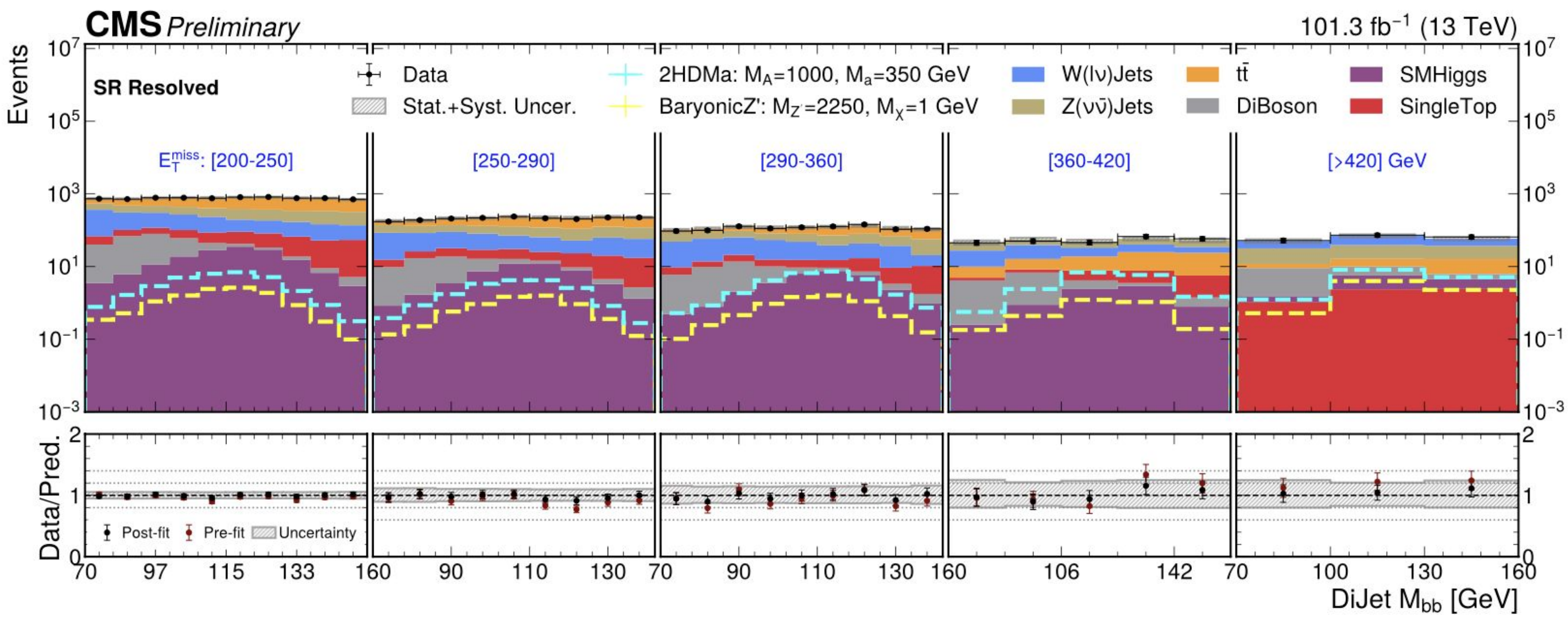




H candidate mass in boosted cat.

$X \rightarrow H(bb)\ Y(E_T^{miss})$ (II)

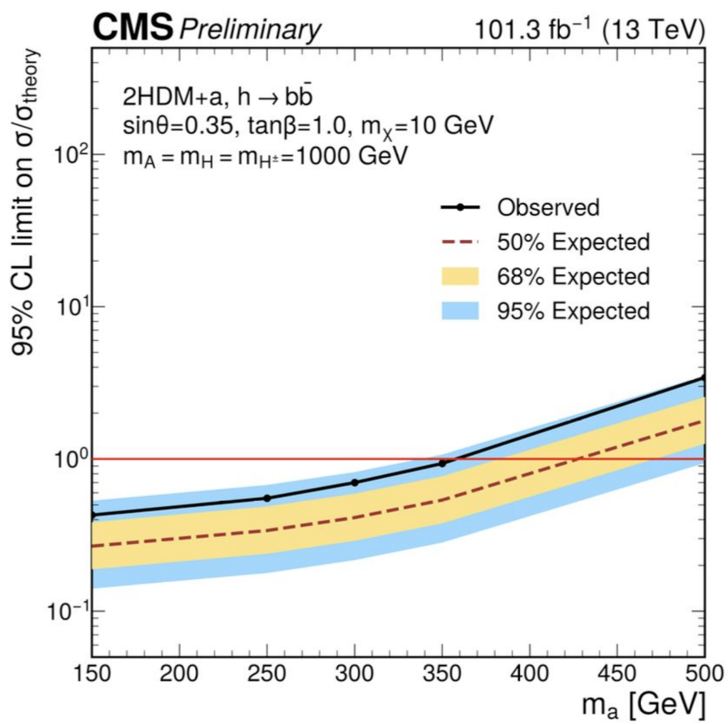
- Singal extraction in m_H in different E_T^{miss} bins
- Simultaneous fit to data in SR and Top, Z CRs in both cats. to constrain ttbar and $Z(\nu\nu)$+jets normalization.
- Other backgrounds and signal are modeled using MC



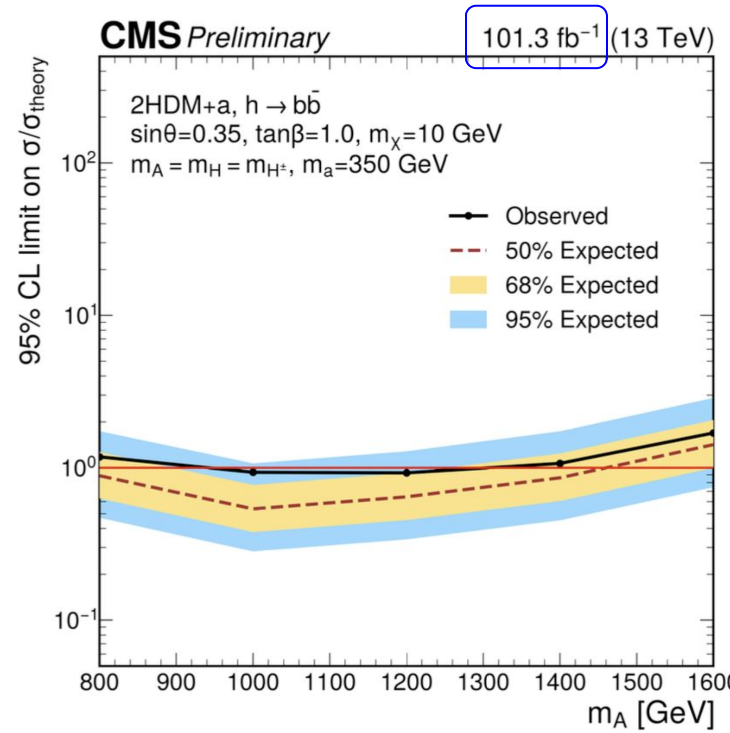
H candidate mass in resolved cat.

$X \rightarrow H(bb)\, Y(E_T^{miss})$ (III)

Results with 2016 under review



Limits on m_a

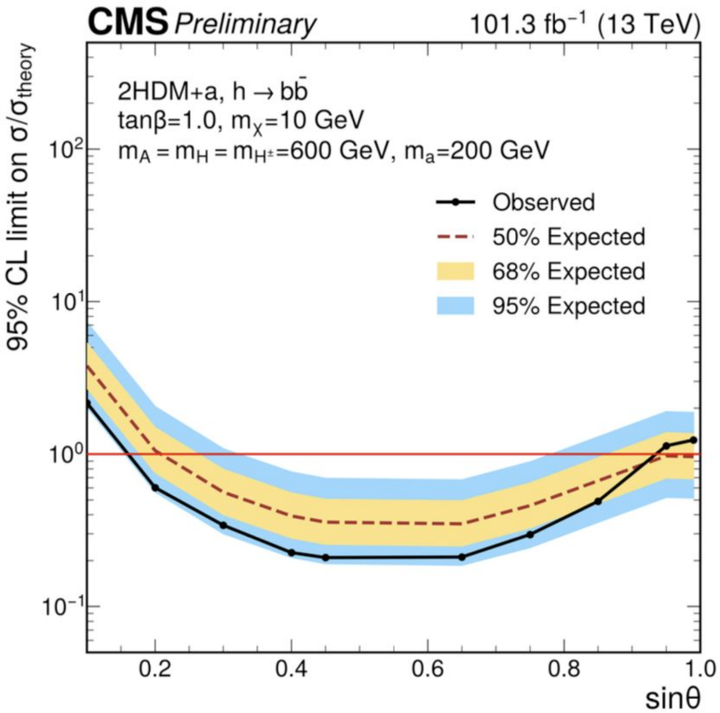


Limits on m_A

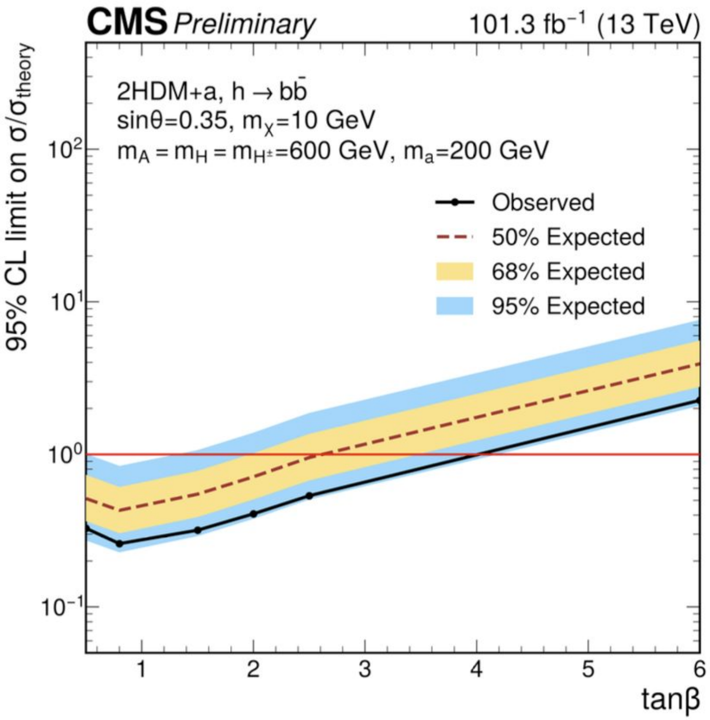
Upper limit on σ/σ_{theory} for 2HDM+a model:

- Exclude $m_a <$ 350 GeV for $m_A =$ 1000 GeV
- Exclude $960 < m_A < 1300$ GeV for $m_a =$ 350 GeV

$X \to H(bb)\ Y(E_T^{miss})$ (IV)



Limits on $\sin\theta$



Limits on $\tan\beta$

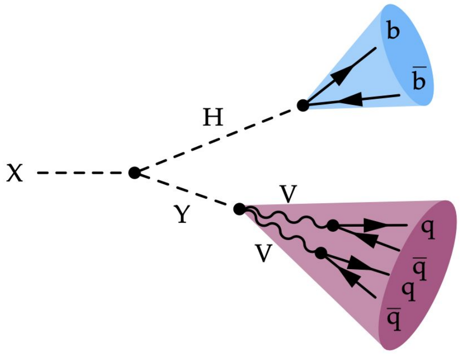
Upper limit on $\sigma/\sigma_{\text{theory}}$ for 2HDM+a model:

- Exclude $m_a <$ 350 GeV for $m_A =$ 1000 GeV
- Exclude 960 $< m_A <$ 1300 GeV for $m_a =$ 350 GeV
- Exclude 0.16 $< \sin\theta <$ 0.93 for $m_A =$ 600 GeV, $m_a =$ 200 GeV
- Exclude $\tan\beta < 4$ for $m_A =$ 600 GeV, $m_a =$ 200 GeV

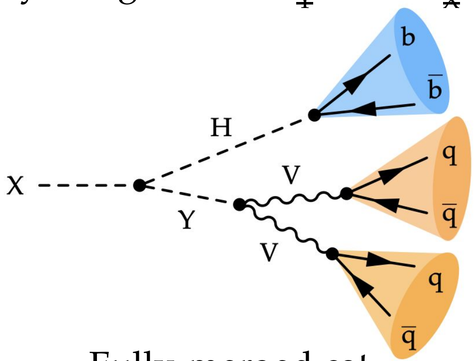
$X \to H(bb)\ Y(VV \to 4q)$

CMS-PAS-B2G-23-007

- Search for $X \to H(bb)\ Y(VV \to 4q)$.
  $m_X \in [900, 4000]$ GeV, $m_Y \in [60, 2800]$ GeV.
  H, V: SM bosons
- Two categories: Higgs candidate reconstructed with large radius AK8 jet
  - Fully merged: $Y \to VV$ candidate reconstructed with single AK8 jet
  - Semi-merged: 2 AK8 jets to reconstruct $V \to qq$ candidates
- ParticleNet jet taggers for $H \to bb$ and $V \to qq$
  - Taggers calibrated in $g \to bb$ and semileptonic tt CRs, respectively
- Particle Transformer jet tagger for $Y \to VV \to 4q$.
  - Lund Plane reweighting method used to calibrate
- Background estimation:
  - QCD estimated from data from its contribution in the taggers 'fail' region scaled by 'pass-to-fail transfer function'
  - Other small background estimated with MC



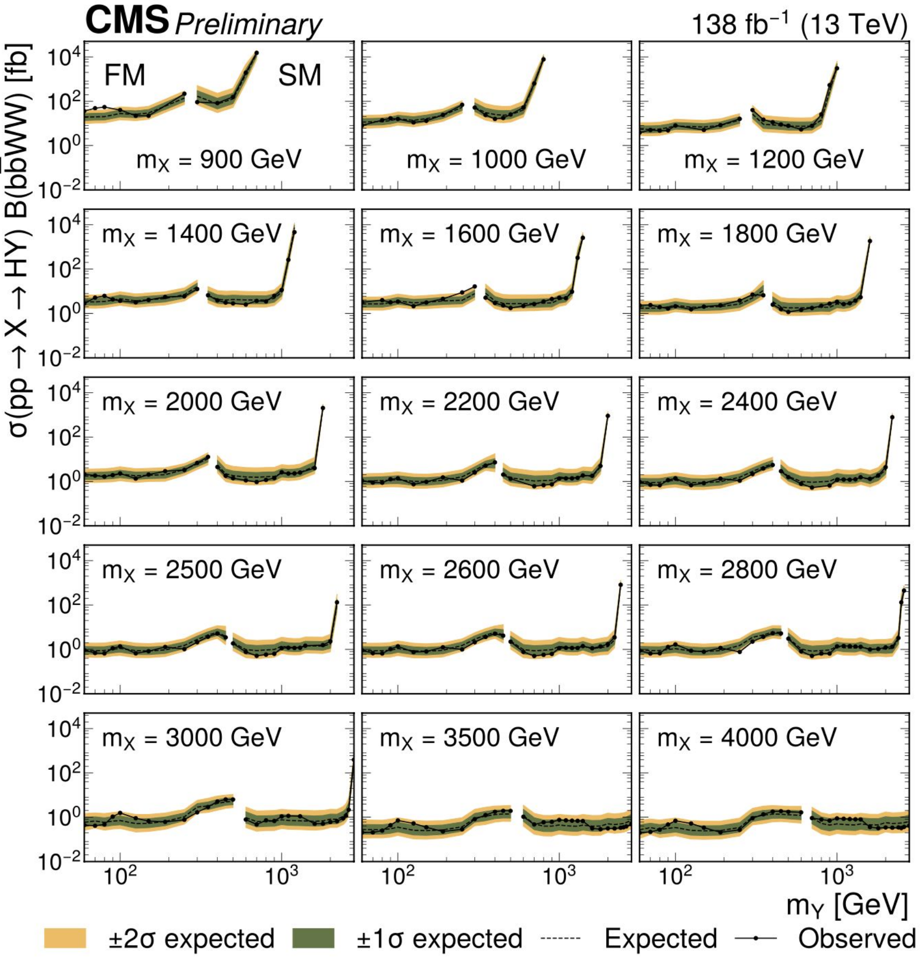
Fully-merged cat., $m_Y < 0.1\ m_X$



Fully-merged cat.,
$0.1\ m_X < 0.1\ m_Y < m_X - m_H$

$X \rightarrow H(bb)\ Y(VV \rightarrow 4q)$ (II)

- Signal extraction in 2D m_X, m_Y plane
- SR: $110 < m_H < 145$ GeV, Validation region: $m_H < 110$ or $m_H > 145$ GeV
- Upper limit on $X \rightarrow H(bb)\ Y(VV)$: ≥ 0.2 fb
- Largest local (global) excess of 3.3 (1) σ significance at m_X=900 GeV, m_Y=80 GeV
- First LHC result in $X \rightarrow H(bb)\ Y(VV)$ in hadronic final state



Are there new scalar particles?

- Search for new scalar particles in different channels with full Run 2 data at CMS
  Discussed a few in the talk

- No significant evidence has been observed so far

- Improving measurements with machine learning techniques for object reconstruction and signal extraction

- Expanding phase space of the searches

- Significant improvements expected with LHC Run 3 and HL-LHC with increased statistics, analysis techniques and modeling

Back up

Motivation

- Measurements of production and SM decays of H(125) are consistent with SM within their uncertainties so far
  - BR(H→invisible) < 16% [1]
    ⇒ BR(H→BSM) ≲ 20% if H is produced with SM stre

- Many beyond SM (BSM) theories predict exotic de
  - For e.g. H→*ss*, H→*aa*, H→*Za*, and *a*/*s*→SM
    where *s* (*a*): (pseudo-)scalar state

- Many well motivated candidates for light (pseudo-
  - Generic single scalar in SM+Singlet
  - Generic singlet (pseudo-)scalar in 2HDM+Singlet
  - Light (pseudo-)scalar of NMSSM
  - Pseudoscalar that mixes with the CP-odd Higgs of (N)MSSM

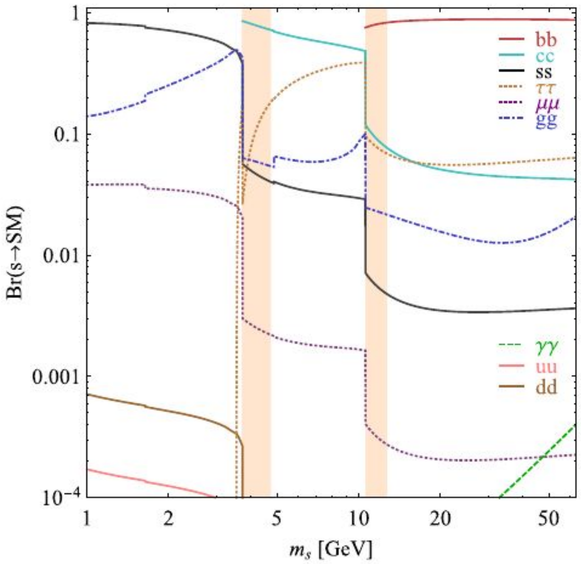


Fig: *s*→SM branching fraction in SM+Singlet model [2]

[1] JHEP 08 (2022) 104
[2] *Phys. Rev. D* 90 (2014) 075004

Motivation (II)

- Searches for light bosons in H decays in CMS:
  - With data collected in 2016 (~36 fb^{-1}):
    - $H \rightarrow aa \rightarrow 4\tau$ [HIG-18-006]
    - $H \rightarrow aa \rightarrow 4\mu$ [HIG-18-003]
    - $H \rightarrow aa \rightarrow 2b\ 2\tau$ [HIG-17-024]
    - $H \rightarrow aa \rightarrow 2b\ 2\mu$ [HIG-18-011]
    - $H \rightarrow aa \rightarrow 2\mu\ 2\tau$ [HIG-17-029, HIG-18-024]
  - With data collected in 2016-2018 (~138 fb^{-1}):
    - **$H \rightarrow aa \rightarrow 4\gamma$ resolved ($m_a \in [15, 60]$ GeV)** [HIG-21-003] ★
    - **$H \rightarrow aa \rightarrow 4\gamma$ boosted ($m_a \in [0.1, 1.2]$ GeV)** [HIG-21-016] ★
    - $H \rightarrow aa \rightarrow 2b\ 2\mu$ ($m_a \in [15, 62.5]$ GeV) [HIG-21-021] ★
    - Model independent searches
    - Searches in other final states are in pipeline

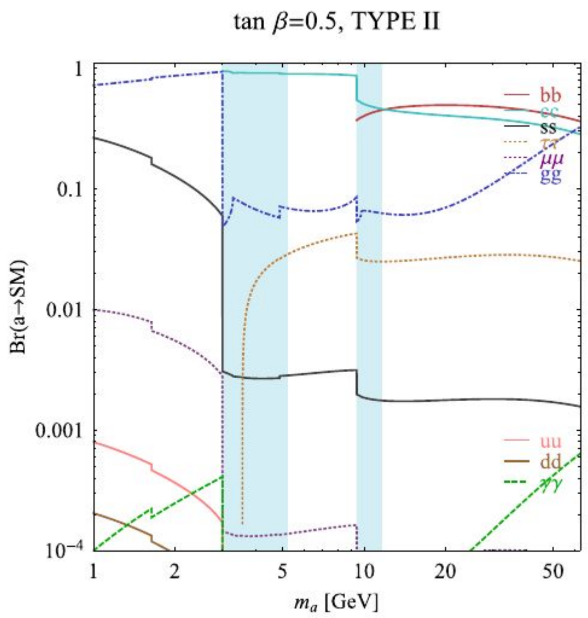
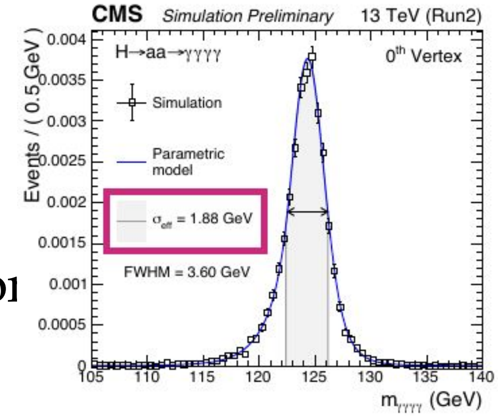


Fig: $a \rightarrow$ SM branching fraction in a specific scenario in 2HDM+Singlet model [1]

★ This presentation
★ Talk by Lakshmi [link]

[1] *Phys. Rev. D* 90 (2014) 075004

H→aa→4γ (resolved) HIG-21-003



Advantage of H→aa→4γ :

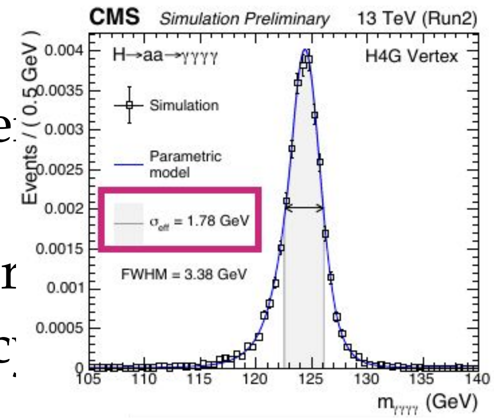
- Relatively low background contribution
- BR(a→2γ) ≈ 100% if a couples at renormalizable level o
  heavy vector-like uncolored matter

Search for H→aa in resolved 4γ in final state for 15 ≤ m_a ≤ 60 GeV

Fig: Standard PV selection (leading $\Sigma\, p_T^2$)

Selection conditions:

- Events with 4γ with p_T > 30, 18, 15, 15 GeV and $|\eta| < 2.4$
- MVA based γ-identification (ID).
  Also required electron veto based on tracker-calorimete
- 110 < $m_{\gamma\gamma\gamma\gamma}$ < 180 GeV
- Used a dedicated BDT to select pp collision primary ver
  ⇒ Improved $m_{\gamma\gamma\gamma\gamma}$ by 3% and PV identification efficienc



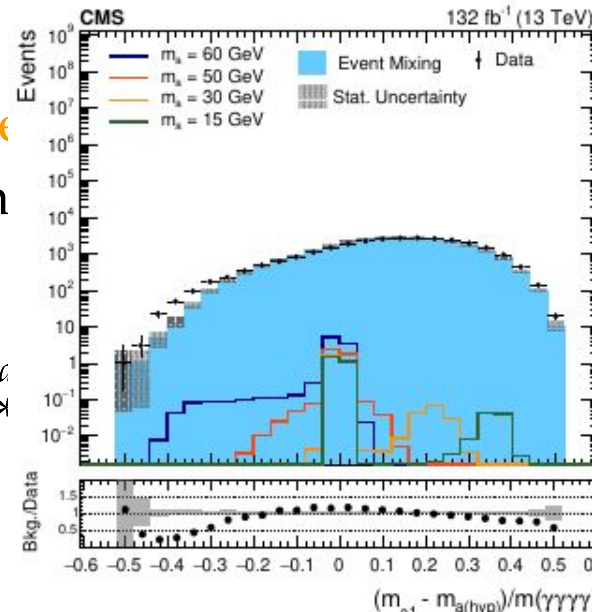
a→2γ **tagging:** Combination of 2γ pairs with the most similar $m_{\gamma\gamma}$ are selected as a→2γ

Fig: BDT based collision vertex selection

H→aa→4γ (resolved) (II)

Signal Region selection:

- BDT (mass-decorrelated) to discriminate signal ove
- Training sample: Signal from MC, background gen model through event mixing.
- Training variables: MVA based γ-ID score, $p_{T,a1}$, $p_{T,a}$ $(m_{a1} - m_{a,\,hyp})/m_{\gamma\gamma\gamma\gamma}$, $(m_{a2} - m_{a,\,hyp})/m_{\gamma\gamma\gamma\gamma}$ and cos θ^*



Background:

- SM γγ + jets, γ + jets and multijet events, in which jets are misidentified as γ.
- Event mixing: 3 out of 4 γ are taken from the next consecutive events from data before preselection.
  - Those mixing γ are required to satisfy all γ selection criteria.
  - Per-event weight, calculated in $m_{\gamma\gamma\gamma\gamma}$ sideband, is applied to improve data - background agreement.

Fig: One of the top ranked variables in BDT training

H→aa→4γ (resolved) (III)

Signal Region (SR) selection (continue):

- Unique BDT output obtained for each m_a hypothesis
- BDT output threshold for SR is decided by maximizing approximate mean significance (AMS), separately for each m_a hypothesis

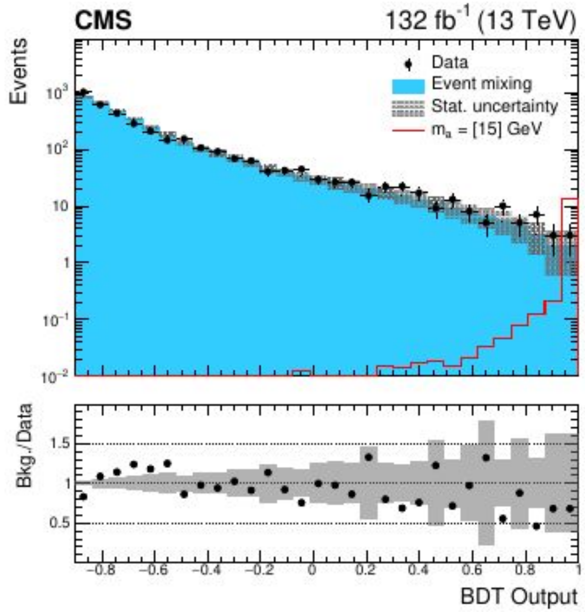


Fig: S vs B BDT for $m_{a,\,\mathrm{hyp}}$ = 15 GeV

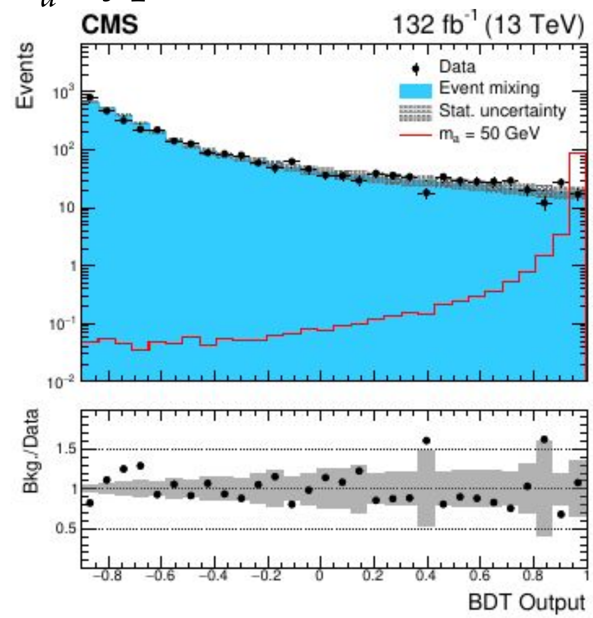
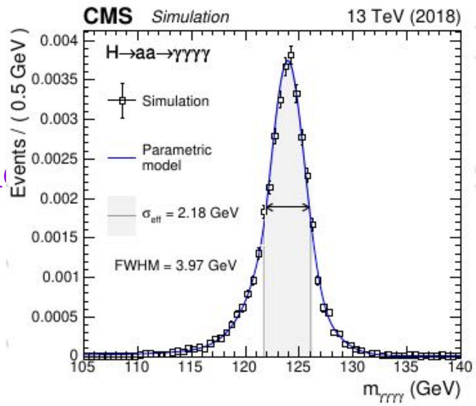


Fig: S vs B BDT for $m_{a,\,\mathrm{hyp}}$ = 50 GeV

H→*aa*→4γ (resolved) (IV)

Signal estimation: Maximum likelihood fit of '$\mu S + B$' fun

μ: signal strength parameter floating in the fit.



Signal modeling (S):

- Signal template is derived by fitting 'Double-sided Crystal Ball' function to $m_{\gamma\gamma\gamma\gamma}$ in signal MC,
- Separately for each m_a hypothesis and for each data taking year

Fig: S profile for $m_{a,\text{ hyp}}$ = 15 GeV

Background modeling (B):

- Background functions: exponentials, Bernstein polyno
  and power law functions.
- Background modeling is performed by likelihood fit of
  to data. Choice of the background function is treated as
  parameter via discrete profiling method [1].

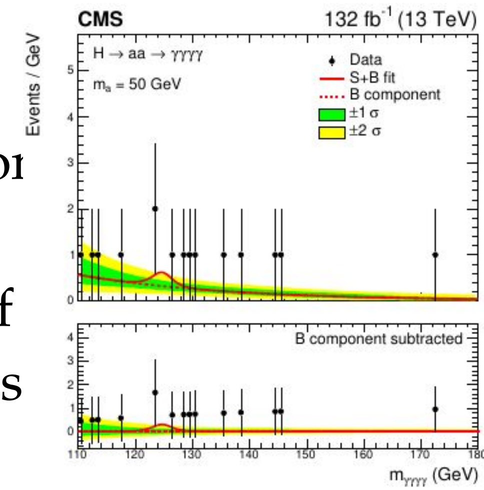


Fig: μS+B fit to data for $m_{a,\text{ hyp}}$ = 50 GeV

[1] *JINST* 10 (2015) P04015

H→aa→4γ (resolved) (V)

- No significant deviation from background-onl
- Observed (expected) upper limits at 95% CL o
  $\sigma_H \times$ BR(H→aa→4γ):
  0.80 (1.00) *fb* for m_a =15 GeV ⇔ BR(H→*aa*
  0.26 (0.24) *fb* for m_a =50 GeV ⇔ BR(H→*aa*

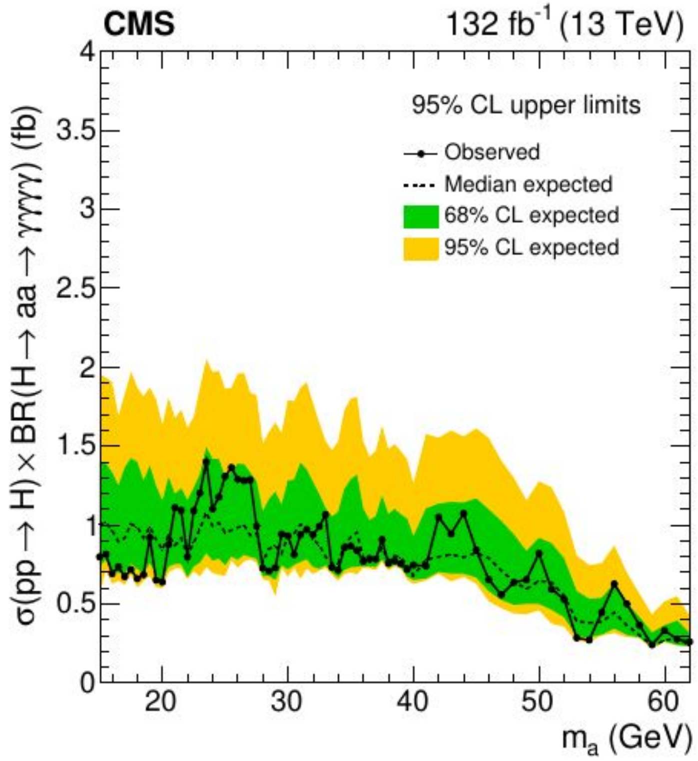


Fig: H→aa→4γ (resolved) upper limit

$H \to aa \to 4\gamma$ (boosted) HIG-21-016

Search for $H \to aa$ in 2 merged γ final for $0.1 \leq m_a \leq 1.2$ GeV

- $BR(a \to 2\gamma)$ enhances for $m_a < m_{2\mu}, m_{2\pi}, m_{J/\psi} = 0.21, 0.28, 3.1$ GeV [1]
- Smaller $m_a \Rightarrow$ larger Lorentz boost $\gamma_L = E_a/m_a$ for the same energy E_a $\Rightarrow a \to 2\gamma$ reconstructed as merged photon object (Γ)
- **End-to-end m_Γ regressor**: Dedicated convolutional neural network to estimate m_a from calorimeter deposits [2]

Selection conditions:

- Events with 2γ with $p_T > 33, 25$ GeV and $|\eta_\Gamma| < 1.4$
- MVA based γ-identification (ID).
  Also required electron veto based on tracker-calorimeter overlap.

[1] *Phys. Rev. D* 90 (2014) 075004
[2] arXiv:2204.12313

H→*aa*→4γ (boosted) (II)

Phase space is divided in different regions:

- m_H-SR: $110 < m_{\Gamma\Gamma} < 140$ GeV
- m_H-SB_{low}: $100 < m_{\Gamma\Gamma} < 110$ GeV; m_H-SB_{high}: $140 < m_{\Gamma\Gamma} < 180$ GeV
- m_A-SR: $|m_{\Gamma 1}-m_{\Gamma 2}| < 0.3$ GeV
- m_A-SB: $|m_{\Gamma 1}-m_{\Gamma 2}| > 0.3$ GeV

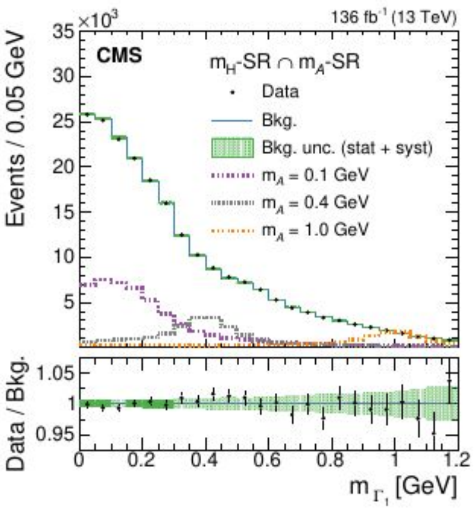


Fig: $m_{\Gamma 1}$

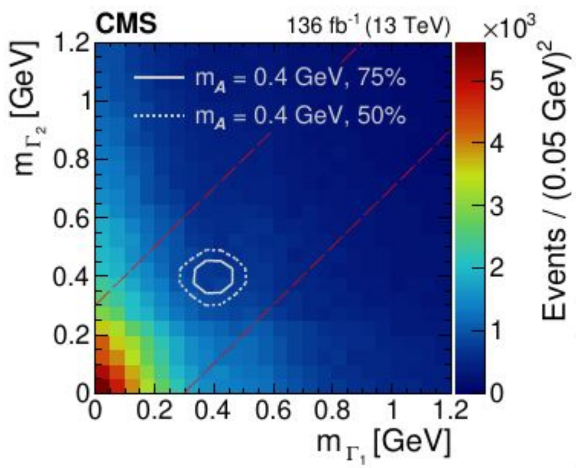


Fig: Observed 2D-m_Γ

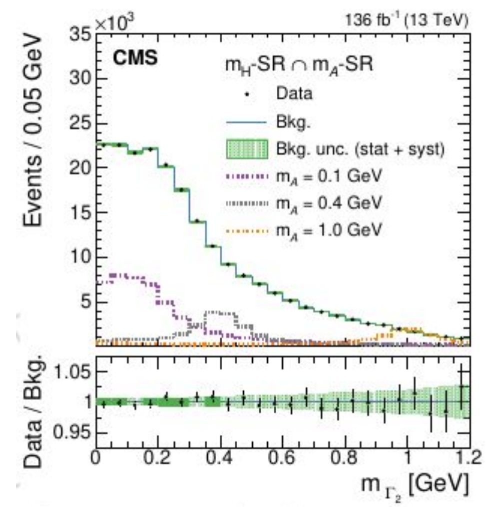


Fig: $m_{\Gamma 2}$

H→aa→4γ (boosted) (III)

Signal model (S): Binned 2D-m_Γ template derived from signal MC

Signal estimation: Maximum likelihood fit of $\mu S + B$ template to data in 2D-m_Γ in m_H-SR ∩ m_A-SR′ region.
μ: signal strength parameter floating in the fit.

Background model (B):

- Components:
  - H→2γ background: 2D-m_Γ template derived from MC, normalize using BR(SM H→2γ)$_{\text{theory}}$
  - Non-resonant backgrounds:
    - Dijet, γ+jet, prompt $\gamma\gamma$
    - Data driven 2D-m_Γ template derived

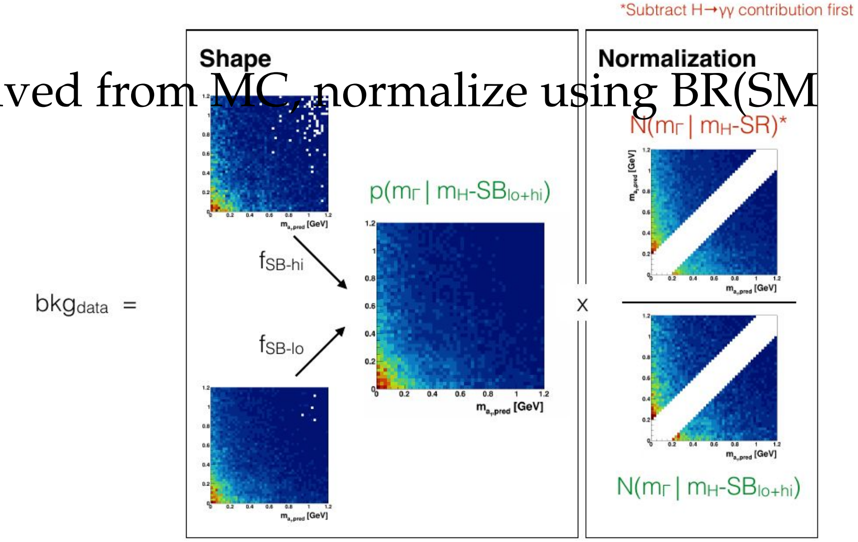


Fig: Data driven estimation of non-resonant background. Plots are for illustrative purpose only

H→aa→4γ (boosted) (IV)

- No significant excess observed over SM-only
- Observed (expected) upper limits at 95% CL
  BR(H→aa→4γ): (0.9 - 3.3)×10^{-3} for m_a (0.1 - 1.
- First CMS H→aa limits below a→2μ threshol

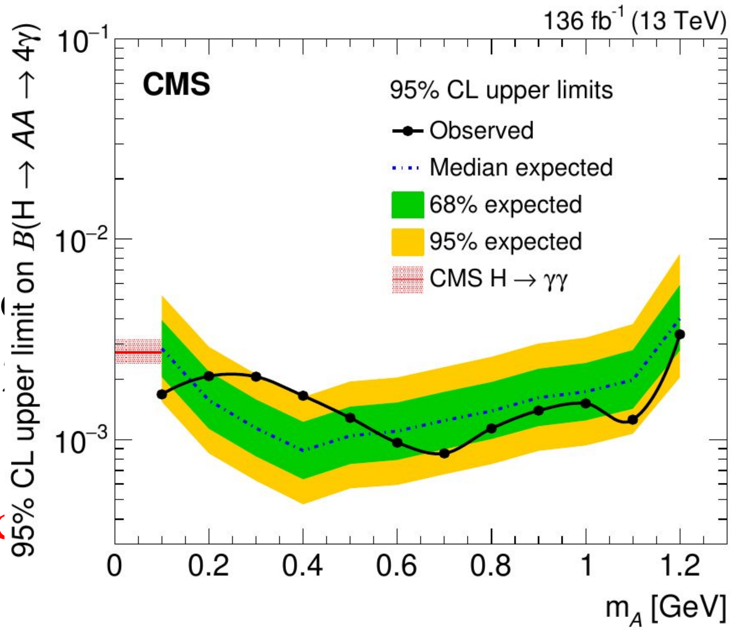


Fig: H→aa→4γ (boosted) upper limit

Summary

- Searches for H→*ss*, H→*aa*, H→*Za* exotic decays provide excellent probe to test many BSM scenarios
- The searches in CMS,
  - With data collected in 2016 (~36 fb^{-1}):
    - H→*aa*→4τ [HIG-18-006]
    - H→*aa*→4μ [HIG-18-003]
    - H→*aa*→2b 2τ [HIG-17-024]
    - H→*aa*→2b 2μ [HIG-18-011]
    - H→*aa*→2μ 2τ [HIG-17-029, HIG-18-024]
  - With data collected in 2016-2018 (~138 fb^{-1}):
    - H→*aa*→4γ resolved [HIG-21-003]
    - H→*aa*→4γ boosted [HIG-21-016]
    - H→*aa*→2*b* 2μ [HIG-21-021]
    - Model independent searches
    - The searches in few more final states are in pipeline. So stay tuned!
  - No signature of H→*aa* or H→*ss* found yet in the searches.
    However improvements from full Run 2 dataset and from state-of-art analysis techniques narrow down the search parameter space for future analyses.

Back up