Rare decays of the SM Higgs boson to light pseudoscalars in the CMS Experiment

32nd International Symposium on Lepton Photon Interactions at High Energies
Madison, USA
July 25th 2025

Anagha Aravind[1] on behalf of the CMS Collaboration
[1]University of Wisconsin-Madison

Exotic decays of SM Higgs boson to light scalars

Searches for $h \rightarrow aa$

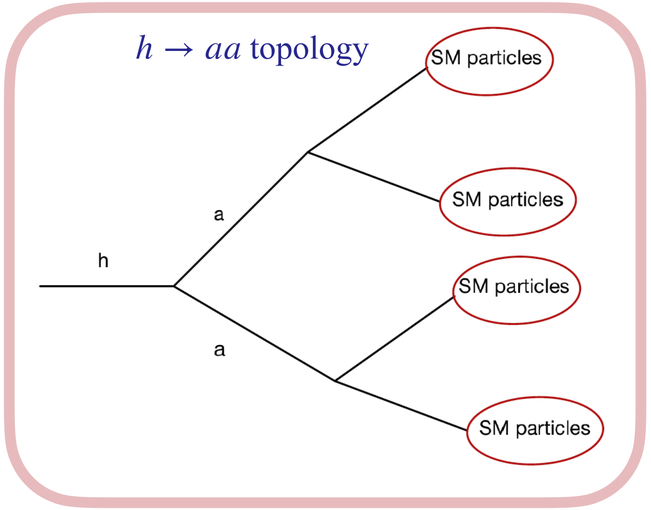
- CMS has put an upper limit of 16% on the B($h \rightarrow BSM$) at 95% CL → searches in exotic final states have potential for discovery

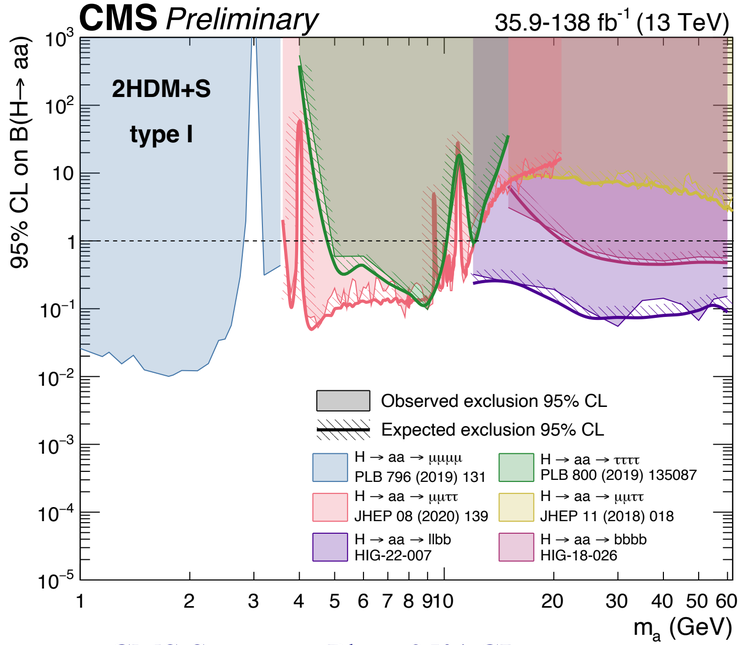
  $\Longrightarrow$Rare Higgs decays may provide window into BSM physics

- SM Higgs decay width is extremely narrow → even BSM particles with low couplings to SM Higgs can have reasonable branching fractions
  - Opportunity for rare Higgs decay searches!

Focus of talk: Searches for 125 GeV Higgs → 2 pseudoscalars (a bosons) → SM particles

- Popular interpretation in 2HDM+S Model: 2 Higgs Doublet+ Singlet extension of SM Higgs
  - Higgs doublets have vevs v_1, v_2, used to define $tan\beta \equiv \frac{v_1}{v_2}$
  - 4 types (Types I, II, III, IV) defined based on interactions with fermions
  - Model parameters set such that BR($a \rightarrow SM$) is close to BR($h \rightarrow SM$)
- Model independent searches explored in talk





CMS Summary Plot : 95% CL on B($h \rightarrow aa$ for 2HDM+S Type I)

Focus of talk: CMS $h \rightarrow aa$ searches

This talk will focus on recent CMS results analyzing Run-2 data:

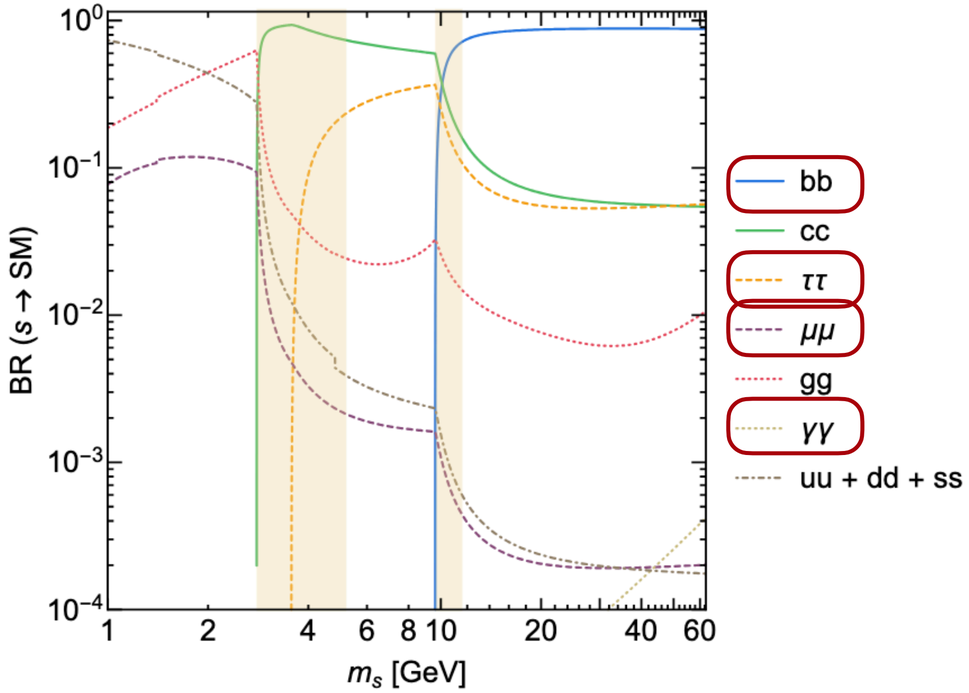
1. $H \rightarrow aa \rightarrow 4e$ (CMS-PAS-EXO-24-031)
   - Interpreted in ALP theories

   $10\ \text{MeV} \leq m_a \leq 100\ \text{MeV}$
2. $H \rightarrow aa \rightarrow 4\gamma$ (CMS-PAS-EXO-24-025)
   - Model independent search

   $1\ \text{GeV} \leq m_a \leq 15\ \text{GeV}$
3. $H \rightarrow aa \rightarrow 4\tau/2\tau 2\mu$ (CMS-PAS-SUS-24-002, arXiv:2508.06947)
   - Both model independent and 2HDM+S interpretations
4. $H \rightarrow aa \rightarrow 2b2\tau$ (CMS-PAG-HIG-22-007, EPJ C 84 (2024) 493)
   - 2HDM+S interpretation
   - Combination with $\mu\mu bb$ final state shown in paper

   $12\ \text{GeV} \leq m_a \leq 60\ \text{GeV}$

B($a \rightarrow SM$) for 2HDM+S Type I



arxiv:1312.4992

- Other recent searches not detailed in talk include $H \rightarrow aa \rightarrow \mu\mu hh$ (CMS-PAS-EXO-24-034) → See LLP talk by Jingyu Luo on 26th August

$H \rightarrow aa \rightarrow 4e$

Overview

- $10 < m_a < 100$ MeV
- Interpreted in axion-like particles theories $\Longrightarrow$ a is DM mediator
- Challenge: Highly collimated final state
- Final state: *eeee*
- Reconstruction of Merged e^-e^+ pairs (MEPs):
  - From GSF tracks in CMS silicon tracker
  - $\Delta R < 0.015$ between the tracks
  - BDT for identification of MEPs

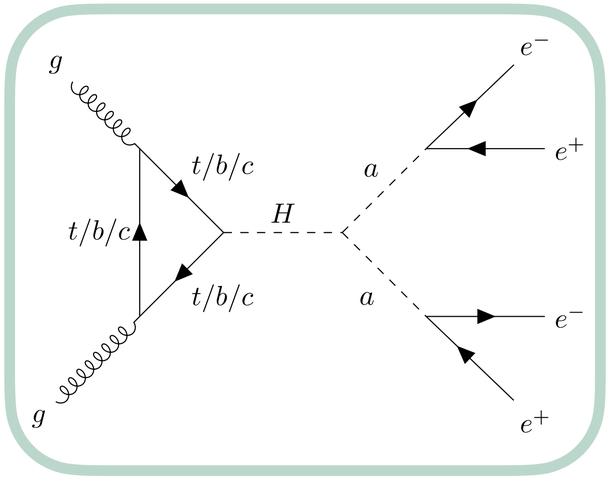
Analysis strategy: Require 2 MEPs

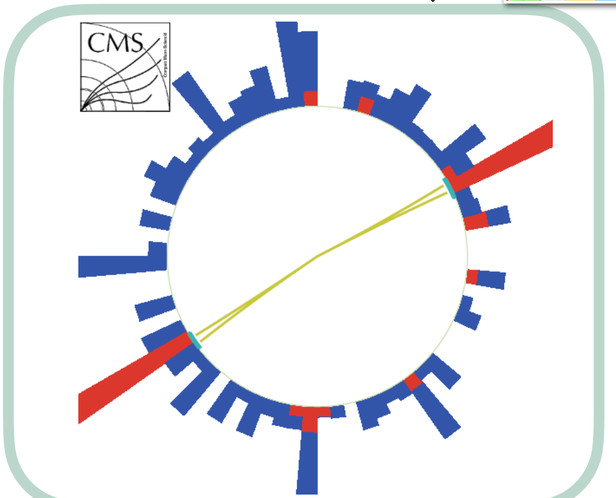
- Signal extraction from m_{4e}

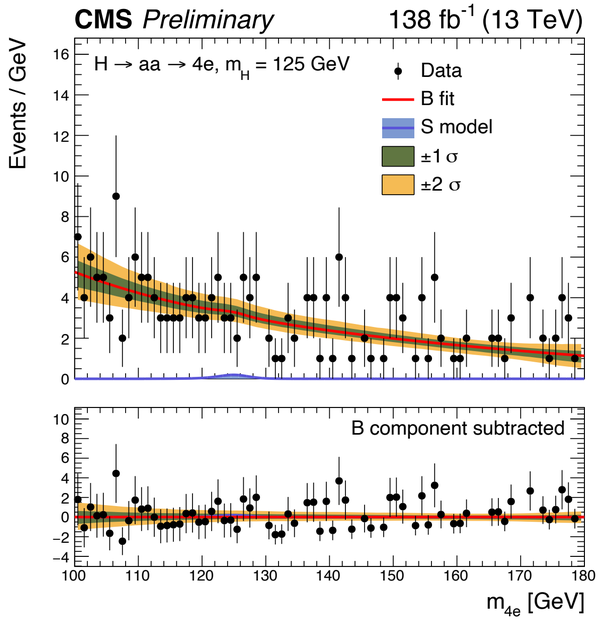
Backgrounds:

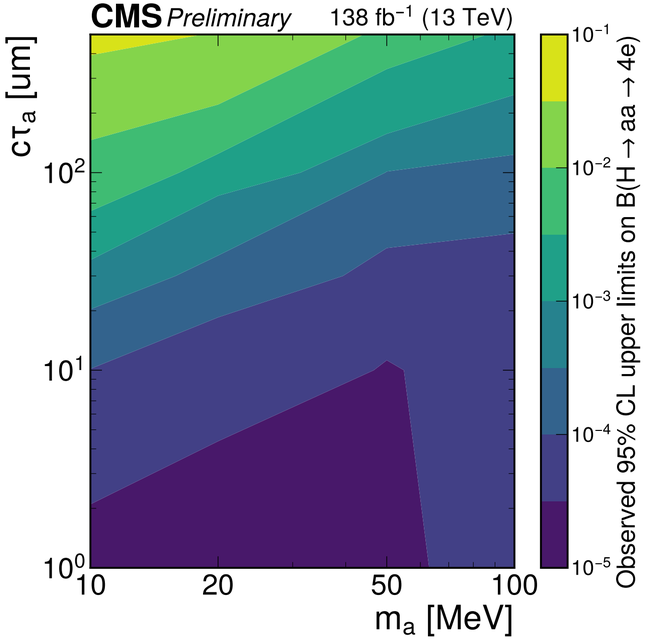
- Prompt and non-prompt photons- From Data in side-band regions
- Other processes- From MC

Results: First of its kind to set limits for $m_a < 100$ MeV







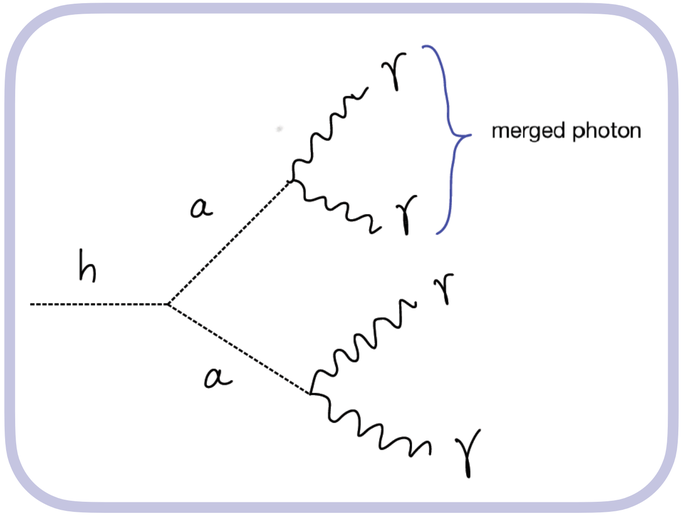


Observed and expected limits at 95% CL for different decay lengths of a

$H \rightarrow aa \rightarrow 4\gamma$

Motivation/Overview

- Model independent search: interpreted in NMSSM, ALP, 2HDM theories
- Final state: 1 merged pair of photons, 1 resolved photon
- Triggers used: diphoton triggers



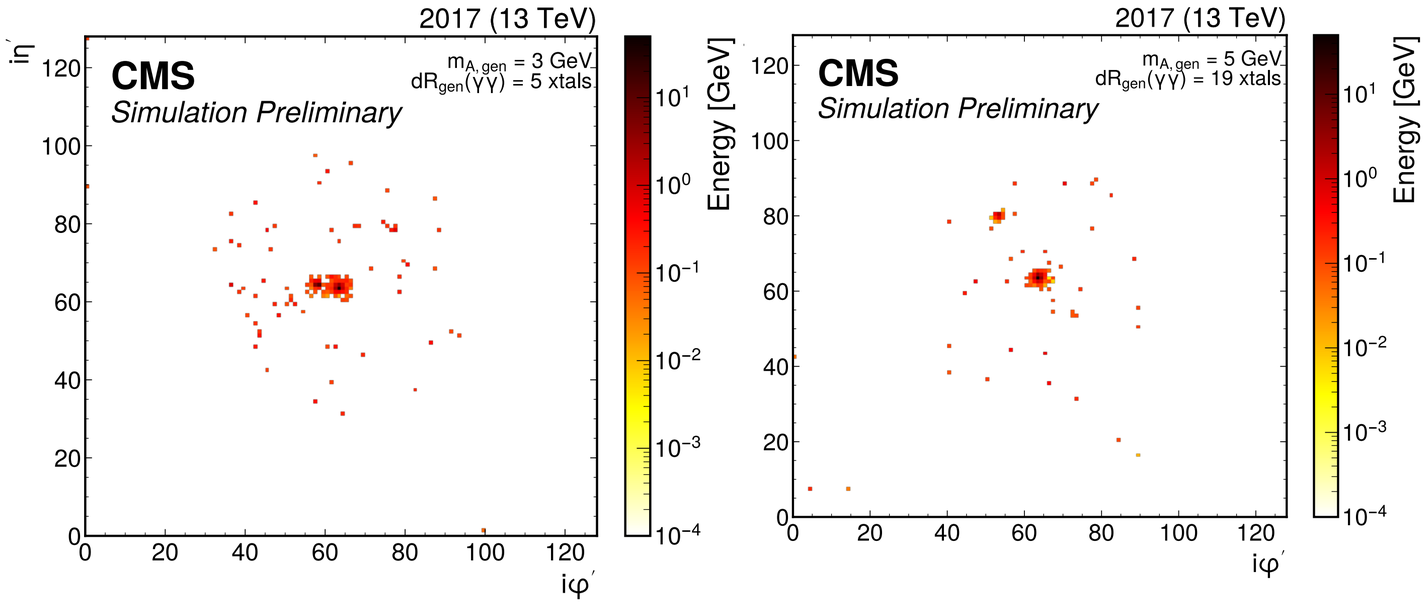
Event selections:

- Require exactly 3 reconstructed photons
- $m_{\gamma\gamma\gamma} > 90$ GeV
- $p_T(\gamma 1) > m_{\gamma\gamma\gamma}/3$ and $p_T(\gamma 2) > m_{\gamma\gamma\gamma}/4$

Background Estimation

1. QCD
2. photon+ jet:
3. Prompt di-photon

Backgrounds are estimated using data driven methods in Side-band region

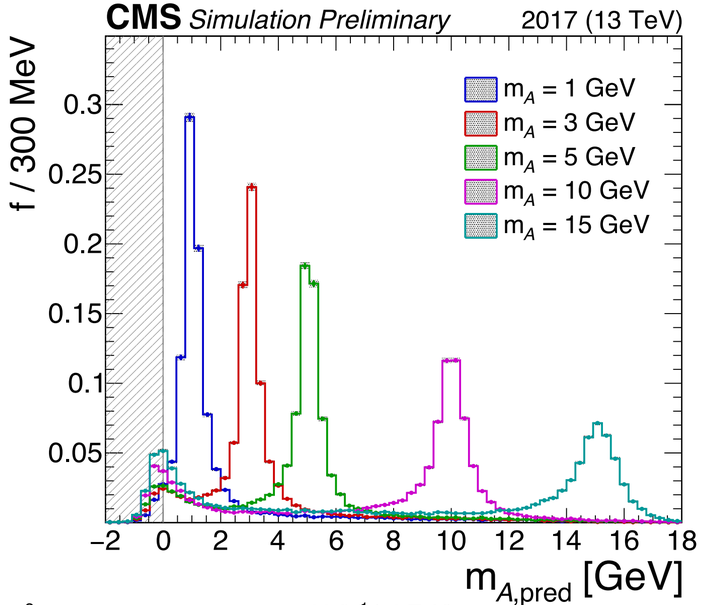


Energy deposit maps in CMS ECAL showing clusters reconstructed as photons

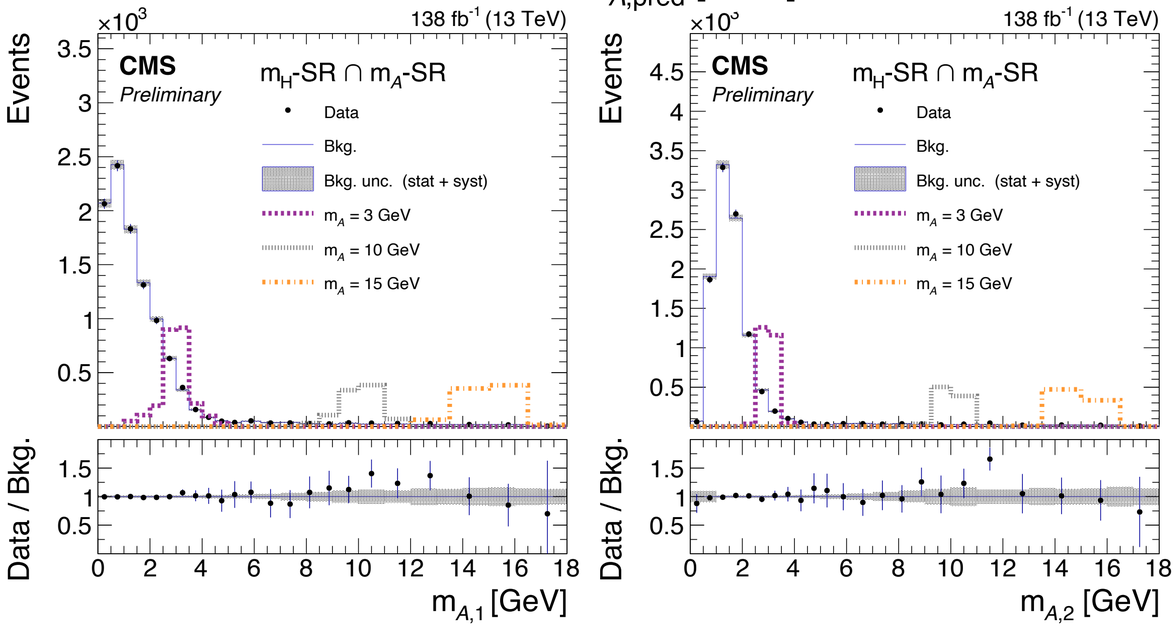
$H \to aa \to 4\gamma$

Analysis Strategy

- For a decaying to fully resolved $\gamma\gamma$: reconstruct p_4 sum of 2 of the closest γ
- For a decaying to merged $\gamma\gamma$: use ML techniques
  - ML regressor fed with ECAL rechit map images
  - GNN trained with privately produced particle guns
- Make 2D distribution of a masses
- Use $\Delta m_a = (m_a(\text{merged leg}) - m_a(\text{resolved leg}))$ and $m_{\gamma\gamma\gamma}$ to define:
  - Signal region (SR): $(\Delta m_a <= 2\text{GeV}) \cap (100 < m_{\gamma\gamma\gamma} < 140\ \text{GeV})$
  - Side band regions (SB): $\Delta m_a > 2\ \text{GeV}$
    - $m_{\gamma\gamma\gamma}$ outside SM Higgs mass window also used to define SBs
    - Used for background estimation



Predicted mass distributions from the ML regressor on the merged leg of $h \to aa \to 4\gamma$



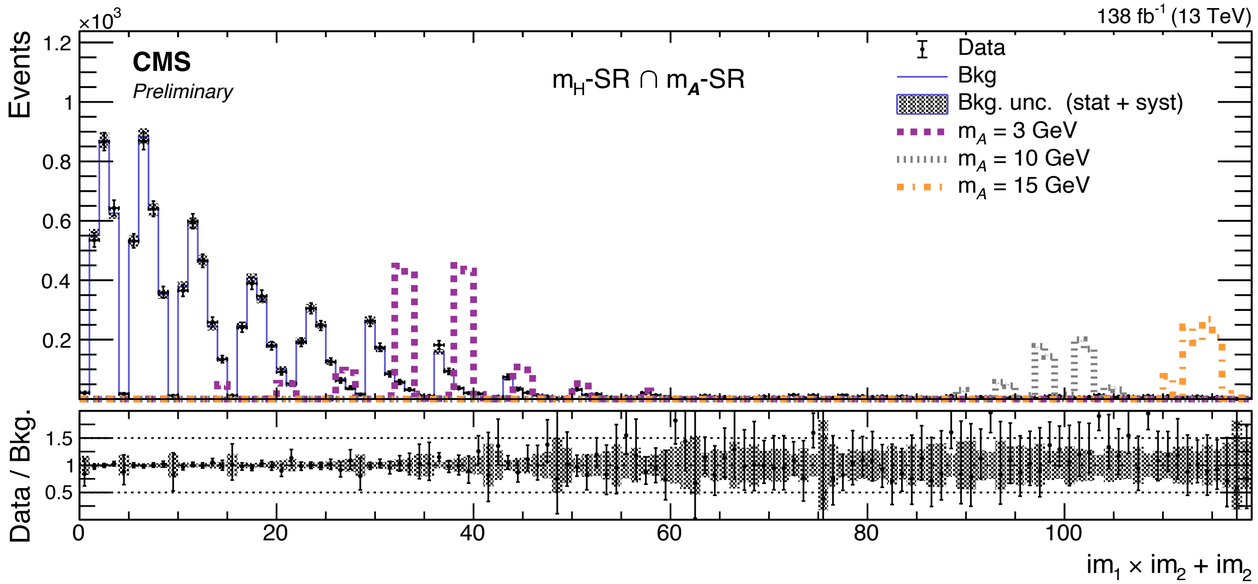
1D projections of the 2D m_a spectra in the SR

$H \rightarrow aa \rightarrow 4\gamma$

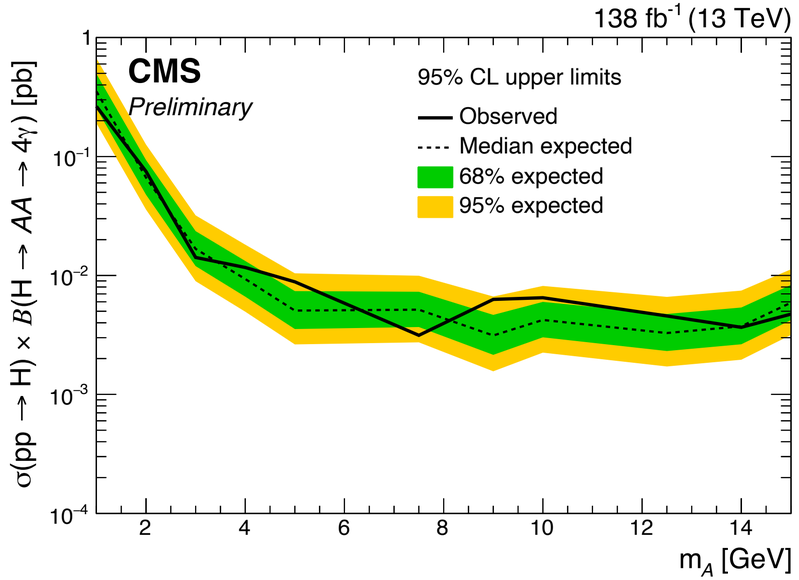
Results

- Signal extracted from simultaneous maximum likelihood fit in 2D m_a distribution

- Analysis sensitive in mass range between previous merged (arXiv:2209.06197) and resolved analyses (arXiv:2208.01469) in CMS



Unrolled distribution of the 2D m_a spectra in the SR

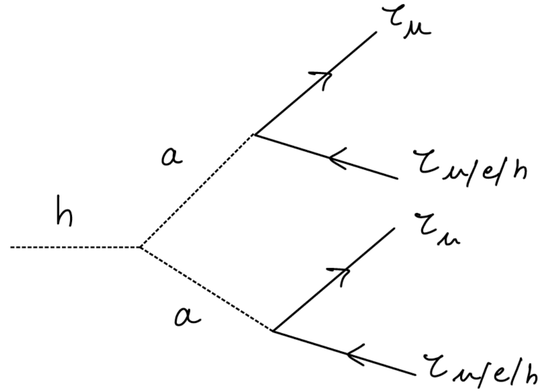


Observed Limits at 95% CL for the exclude signal strength of $h \rightarrow aa \rightarrow 4\gamma$

$H \rightarrow aa \rightarrow 4\tau/2\tau 2\mu$

Motivation/ Overview

- Interpreted in model independent as well as 2HDM+S theories
- Final state: $\tau_{\mu}$and $\tau_{1-prong}$
  - $a \rightarrow \mu\mu$ events also selected but analysis is not optimized for it

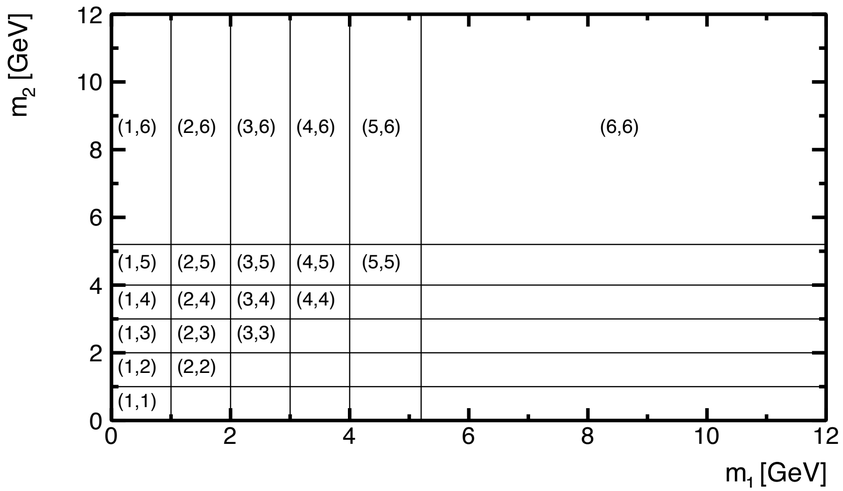


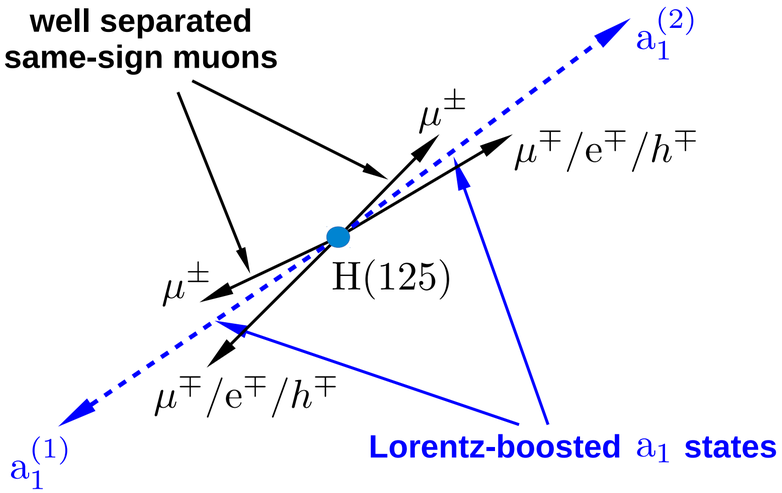
Event Selection:

- Same signed muons, well separated
- Opposite signed charged track nearby
- Triggers: Same signed dimuon triggers

Analysis strategy:

- Signal extraction from 2D map of $m_a(1)$, $m_a(2)$ distribution





Event selections

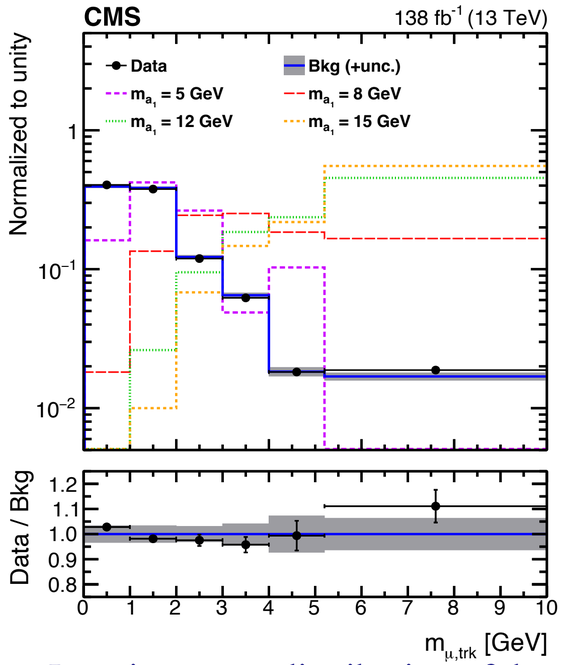
$H \rightarrow aa \rightarrow 4\tau/2\tau 2\mu$

Background estimation: Data driven methods used

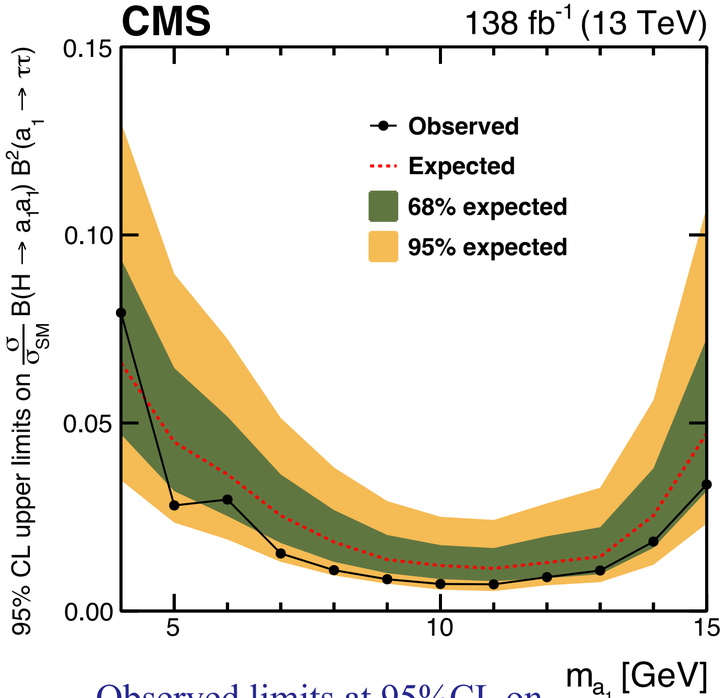
- Major background processes: QCD, Jets faking $\tau_{1-prong}$ (mostly W+jets), $Z \rightarrow \tau\tau$
- Background estimated in CR transferred to SR

Results: Both model independent and 2HDM+S interpretations

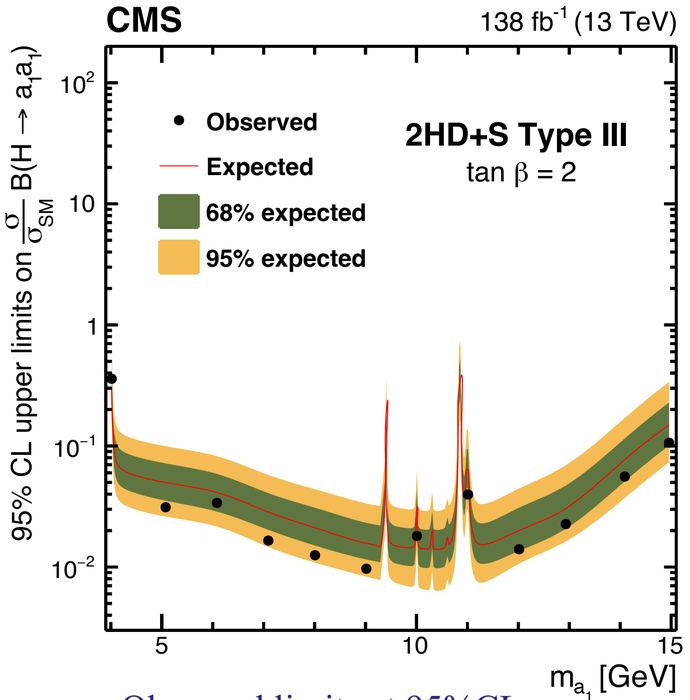
- Most sensitive limits set in 2HDM+S Type III



Invariant mass distribution of the the muon-track system for events passing signal selection



Observed limits at 95%CL on $\sigma(pp \rightarrow H + X)B(H \rightarrow a_1a_1)B^2(a_1 \rightarrow \tau\tau)$ relative to σ_{SM}

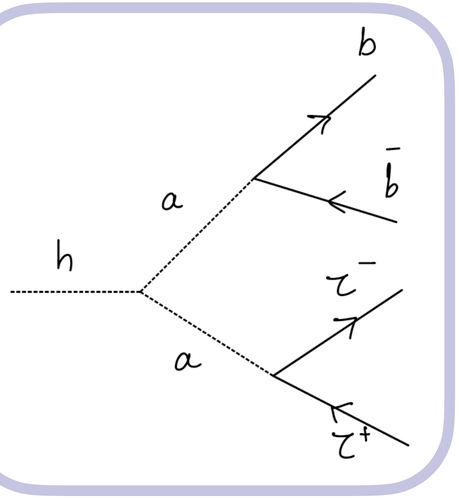


Observed limits at 95%CL on $\sigma(pp \rightarrow H + X)B(H \rightarrow a_1a_1)B^2(a_1 \rightarrow \tau\tau)$ relative to σ_{SM} for 2HDM+S Type III

$H \rightarrow aa \rightarrow 2b2\tau$

Motivation/ Overview

- Interpreted in 2HDM+S theory
- Final state: $bb\tau\tau$
  - favourable branching fractions for higher masses of a boson
- 3 channels: $e\tau_h, \mu\tau_h, e\mu$
- Triggers- single lepton and cross lepton triggers



Event Selection:

- Make $\tau\tau$ pairs: isolated and oppositely charged
- Events require at least one b-tagged jet
- Events categorized into: Exactly 1 b-jet, >1b-jet

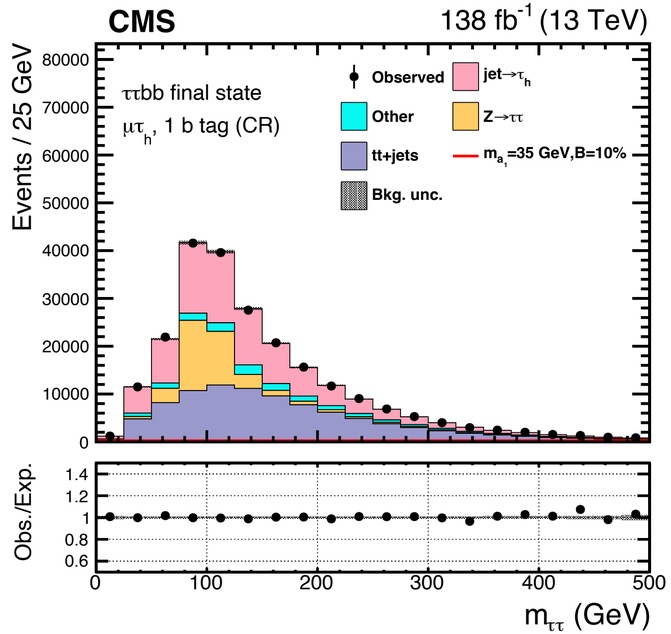
Background Estimation:

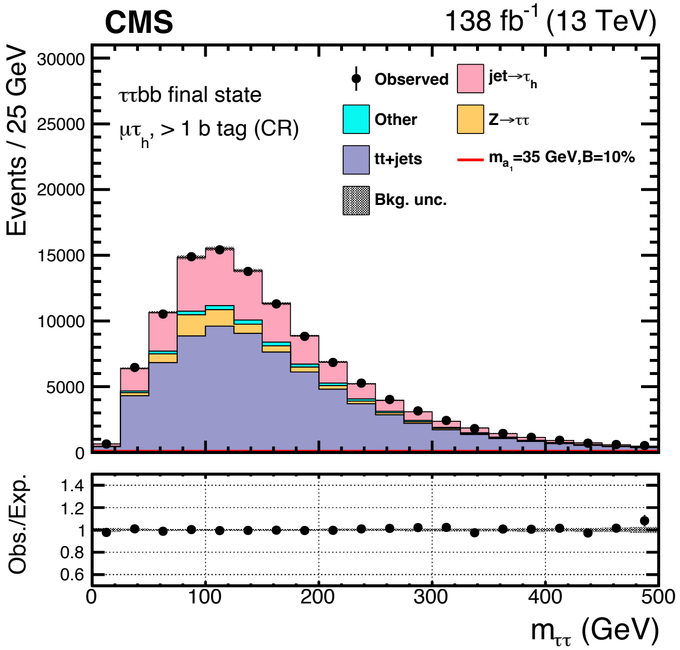
- QCD/ jets faking τ_h: from same signed/ anti iso region
- $Z \rightarrow \tau\tau$: Embedded samples used

From Data

- $t\bar{t}$
- Others: diboson, SM Higgs, single top, W+jets

From MC



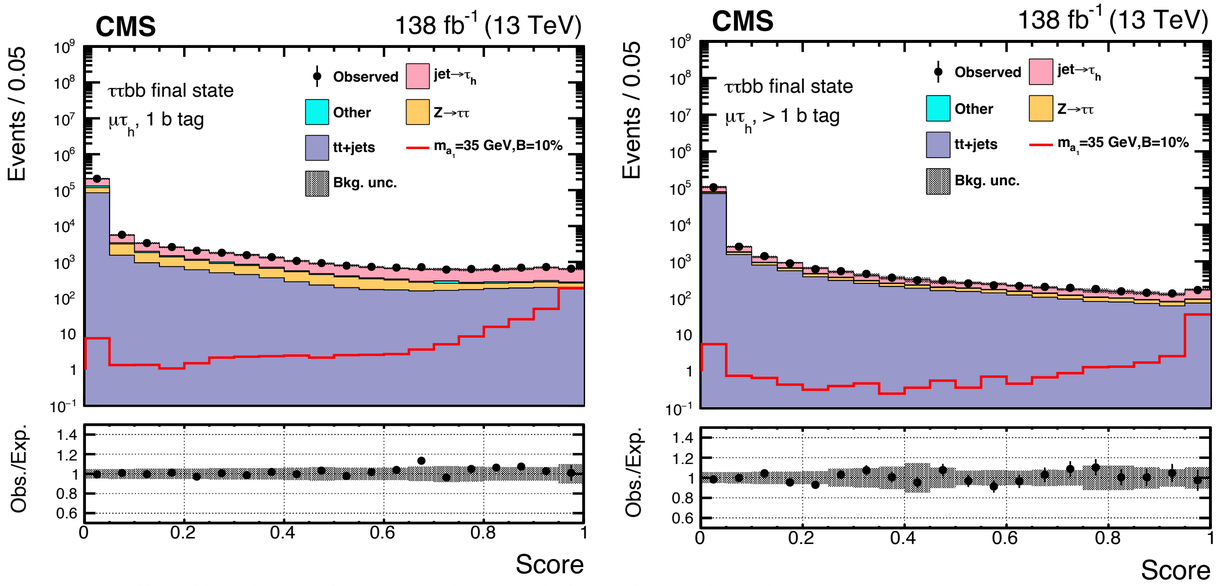


$H \rightarrow aa \rightarrow 2b2\tau$

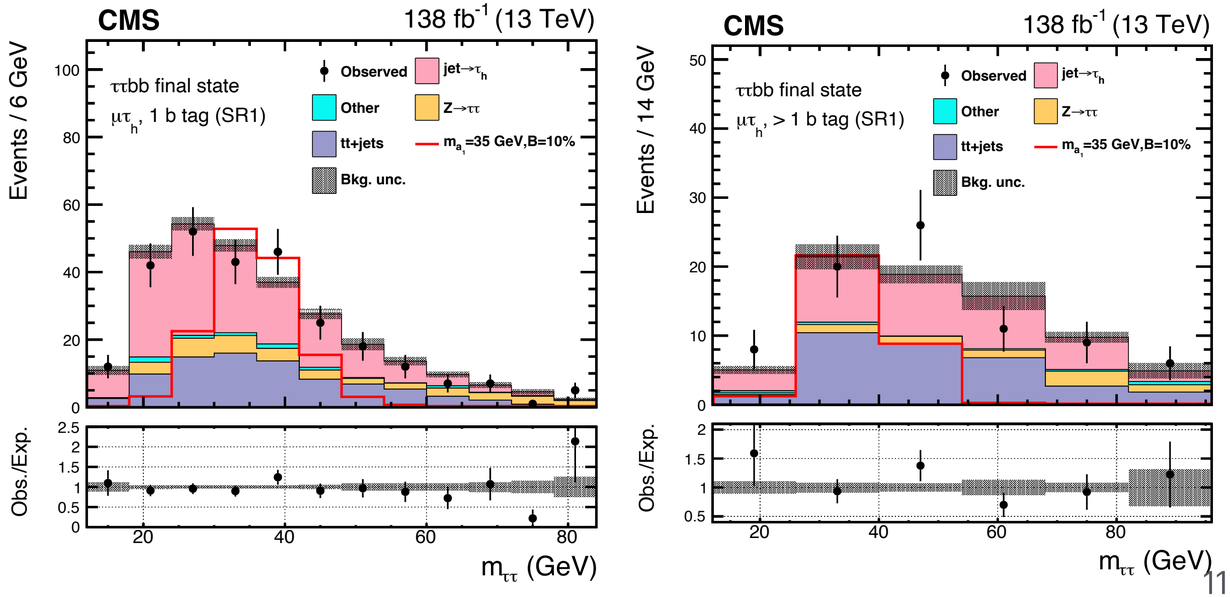
Analysis Strategy

- $m_{\tau\tau}$ reconstructed using SVFIT algorithm
- DNN used for binary event categorization into signal and background
  - DNN for each channel, each b jet category
  - Input features: kinematic variables such as pT, transverse mass of objects, angular separation between objects, Δm_a etc.
  - Cuts on DNN score to define SR
  - Low DNN score: CR
- Maximum likelihood fit on $m_{\tau\tau}$

DNN score distributions for $\mu\tau_h$ channel for 1b and >1b categories



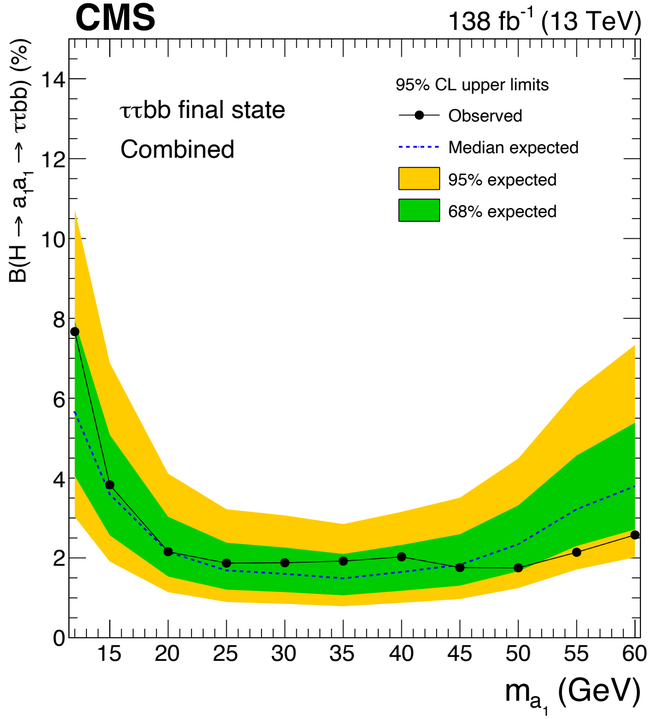
$m_{\tau\tau}$ distributions for $\mu\tau_h$ channel for 1b and >1b categories



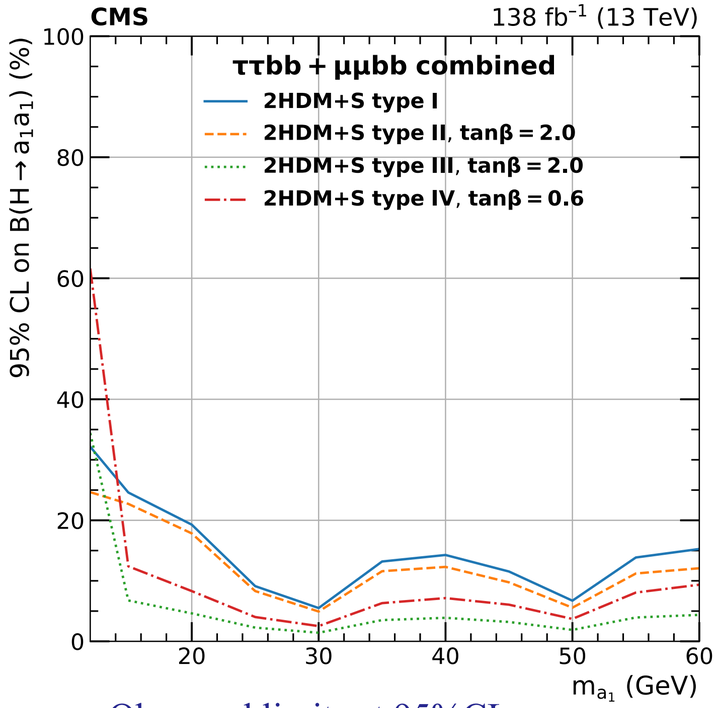
$H \rightarrow aa \rightarrow 2b2\tau$

Results

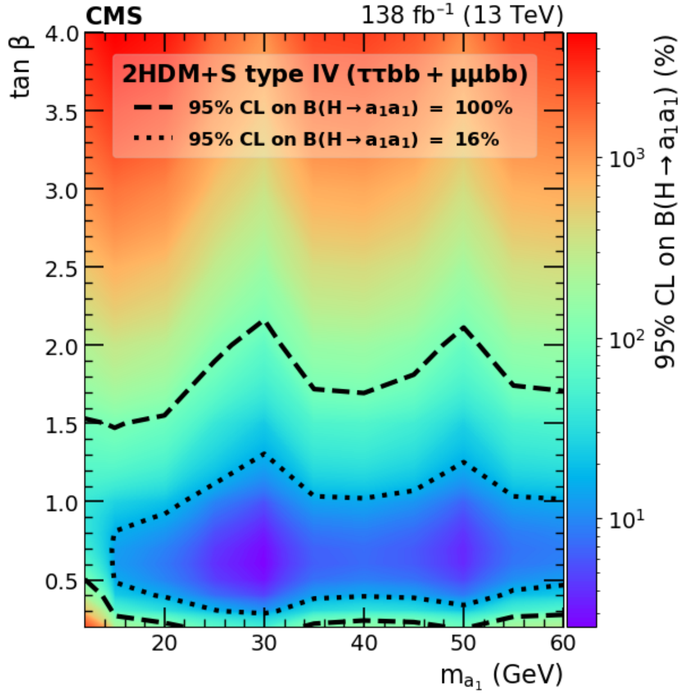
- Results presented for separately $bb\tau\tau$ search and also in combination with $h \rightarrow aa \rightarrow bb\mu\mu$
  - Analyses are mutually exclusive
- Most stringent limits are set in the type III interpretation of 2HDM+S



Observed limits at 95%CL on $B(H \rightarrow a_1 a_1 \rightarrow bb\tau\tau)$



Observed limits at 95%CL on $B(H \rightarrow a_1 a_1)$ obtained from combination of $bb\tau\tau$ and $bb\mu\mu$ final states



Observed limits at 95%CL on $B(H \rightarrow a_1 a_1)$ obtained from combination of $bb\tau\tau$ and $bb\mu\mu$ final states for 2HDM+S

Summary and Outlook

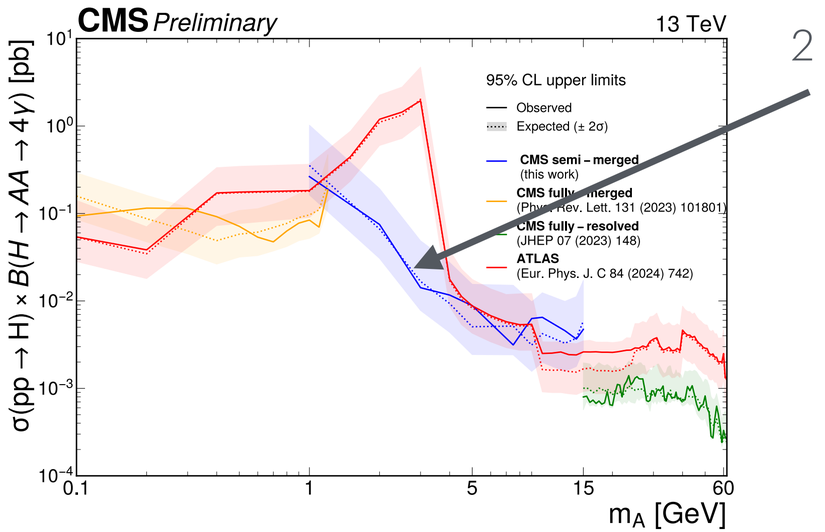
Highlights of selected CMS results:

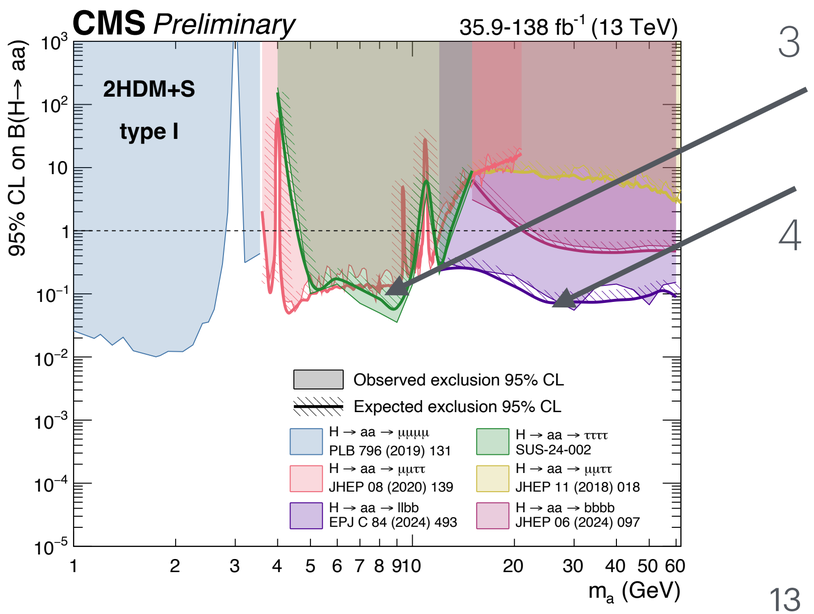
1. $h \to aa \to 4e$:
   - First of its find to set limits for $10 \leq m_a \leq 100$ MeV
2. $h \to aa \to 4\gamma$:
   - Bridges the gap between previous $h \to aa \to 4\gamma$ analysis in CMS
3. $h \to aa \to 2\mu 2\tau / 4\tau$:
   - Significant improvement over previous CMS result by factor of 2
4. $h \to aa \to bb\tau\tau + h \to aa \to bb\mu\mu$ combination:
   - Sets stringent limits for B($h \to aa$) interpretation in 2HDM+S model for $m_a \geq 15$ GeV

Future outlook:

- Potential for other final states to be explored in the future
- Also being probed at CMS: a bosons with different masses, i.e. $h \to a_1 a_2$ decays

Summary plot of observed limits at 95%CL on $\sigma(pp \to H + X)B(H \to a_1 a_1 \to 4\gamma)$
CMS-PAS-EXO-24-025





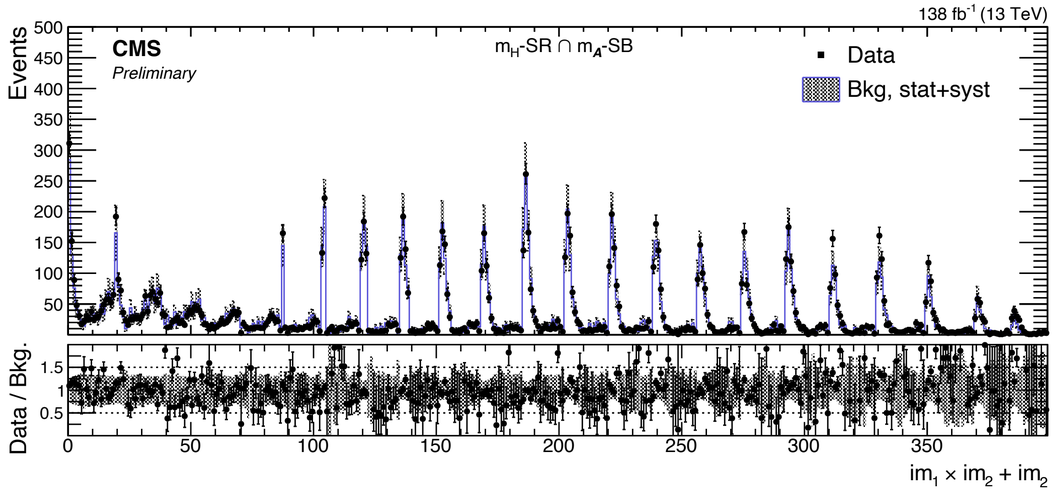
CMS Summary Plot : 95% CL on B($h \to aa$ for 2HDM+S Type I)

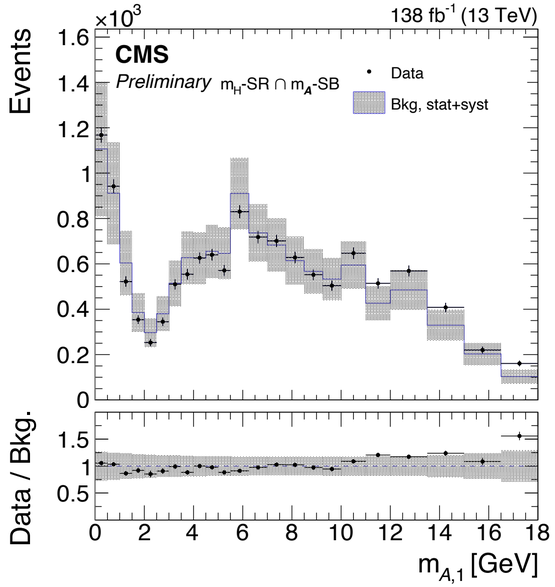
Thank you!

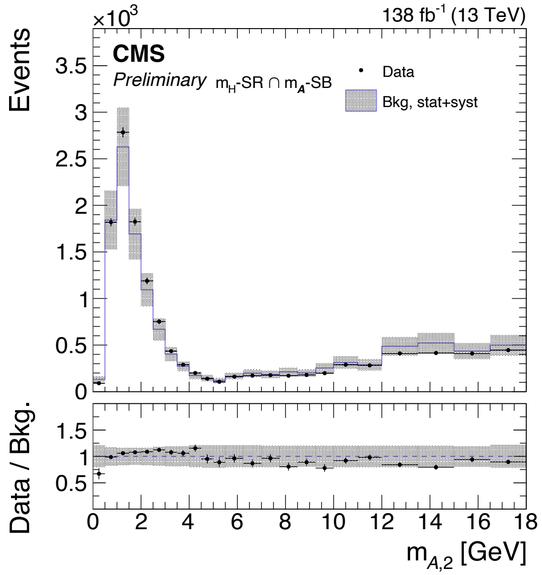
$H \rightarrow aa \rightarrow 4\gamma$

Background Estimation

- Derived from data from the 2D m_a shape template
- Assumptions:
  - shape of background template in SR = linear sum of background templates in SB regions
  - Ratio of events along 2D m_a diagonal derived from m_H SB = Ratio of events in the 2D m_a off diagonal region







$H \rightarrow aa \rightarrow 4\tau/2\tau2\mu$

Background Estimation

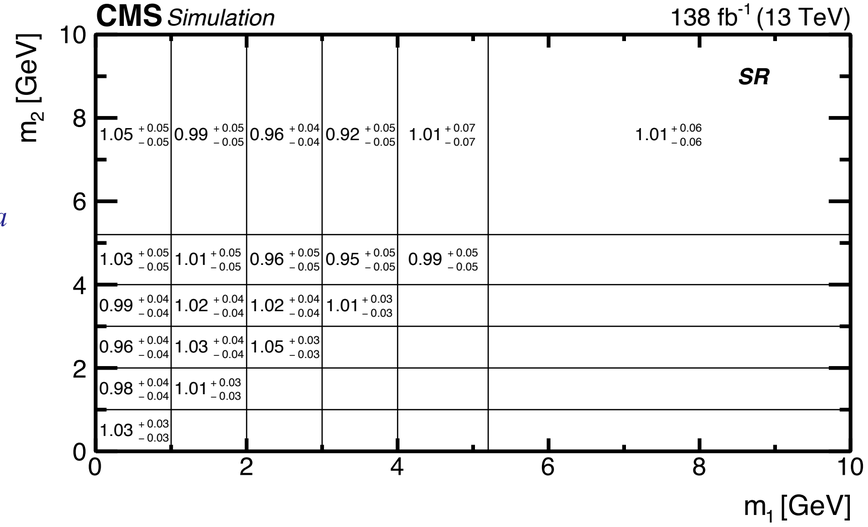
- QCD multijets are major source of background
- Background in SR, $f_{2D}(i,j) = C(i,j)\,(f_{1D}(i)\,f_{1D}(j))$

correlation matrix accounting for correlation of m1 and m2

content of bins in 1-D distributions of muon-track system invariant mass distribution

- $f_{1D}(i)$: Derived in CRs defined by inverting isolation for tracks
- $C(i,j)$: Derived in CR defined by requiring both isolated and non-isolated tracks close to muon

C(i j) in the 2D m_a template in SR



For $f_{1D}(i)$: inv. mass distribution for different isolation for tracks with highest p_T