Searches for exotic Higgs boson decays at CMS

Meng Lu

30 Nov, 2023

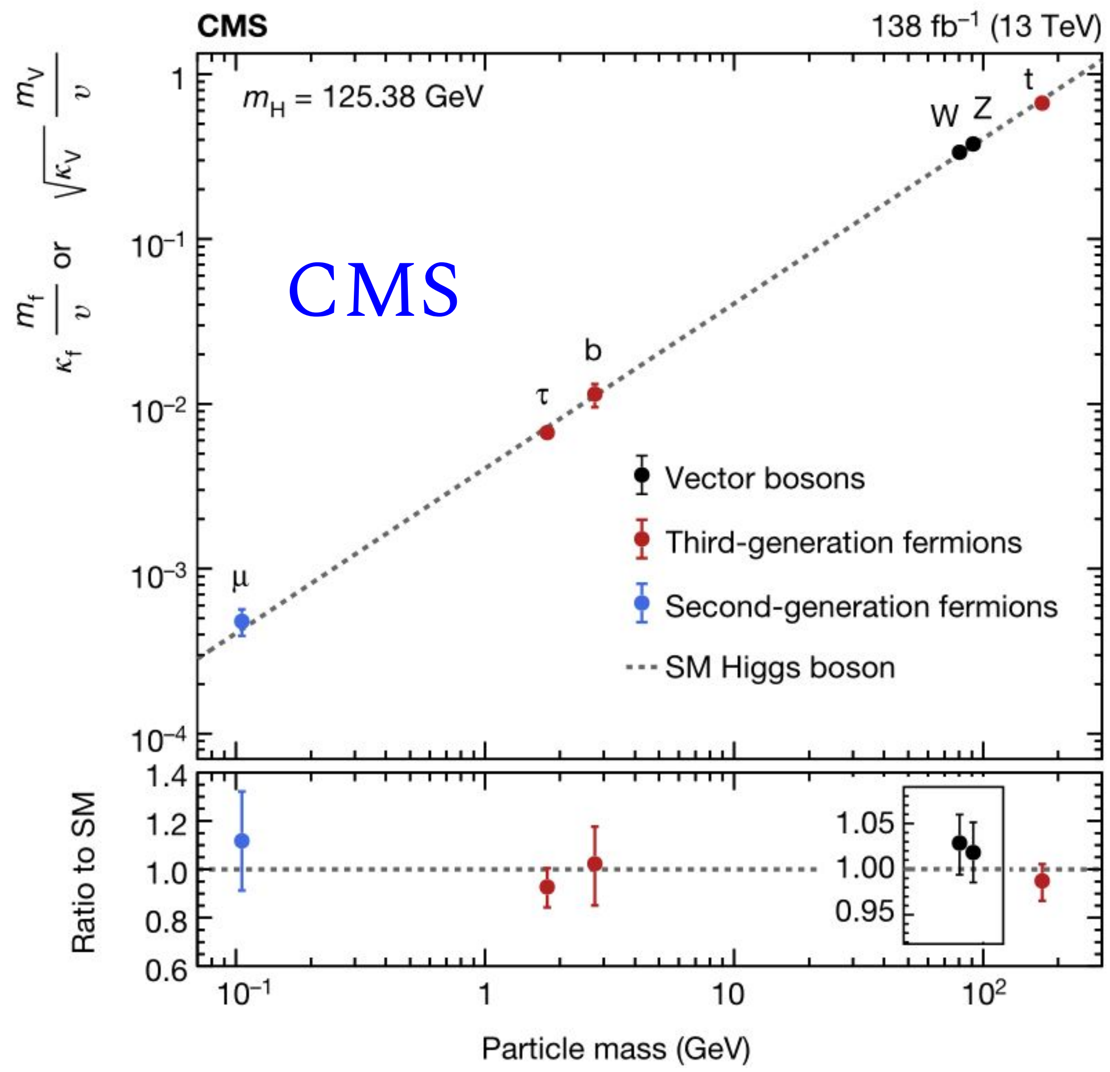
On behalf of the CMS collaboration

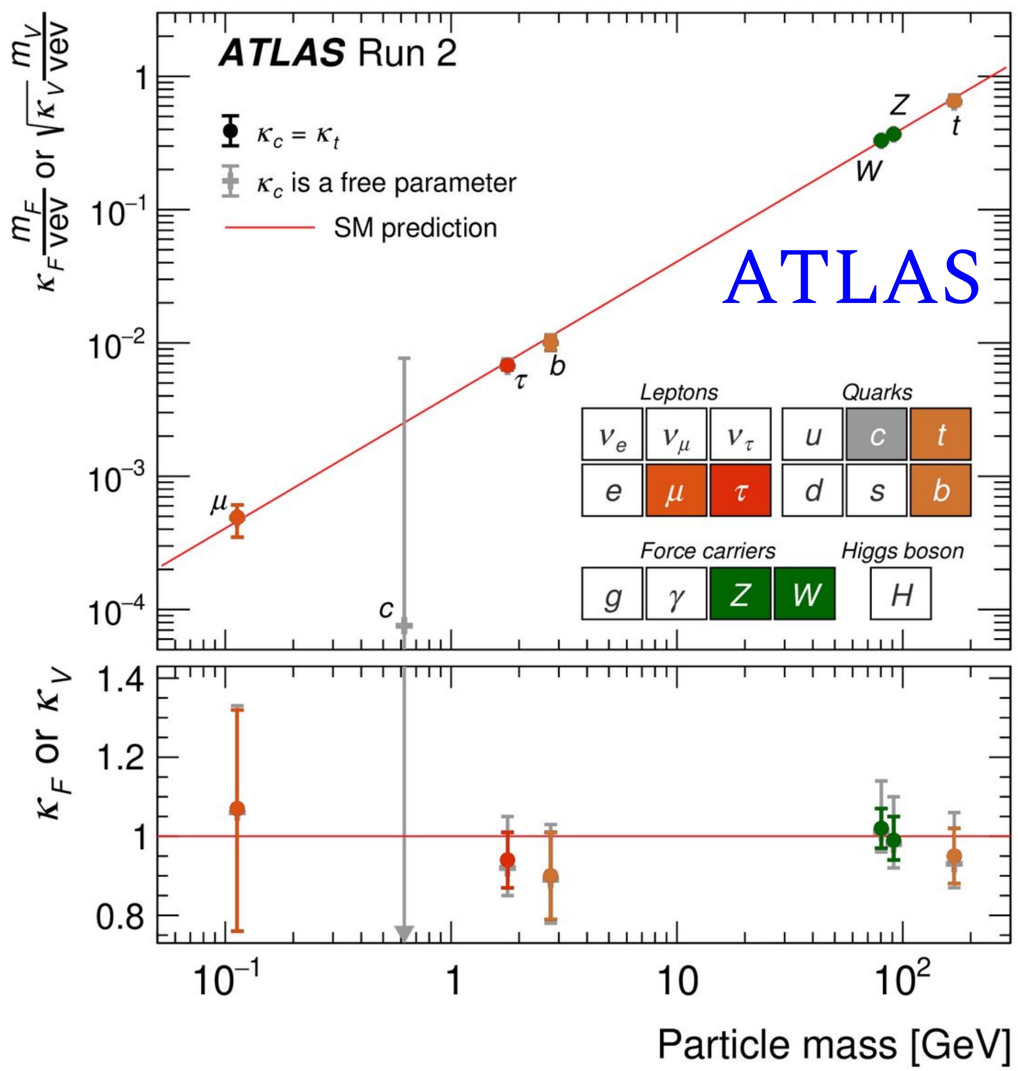
Northeastern University (US)

Introduction

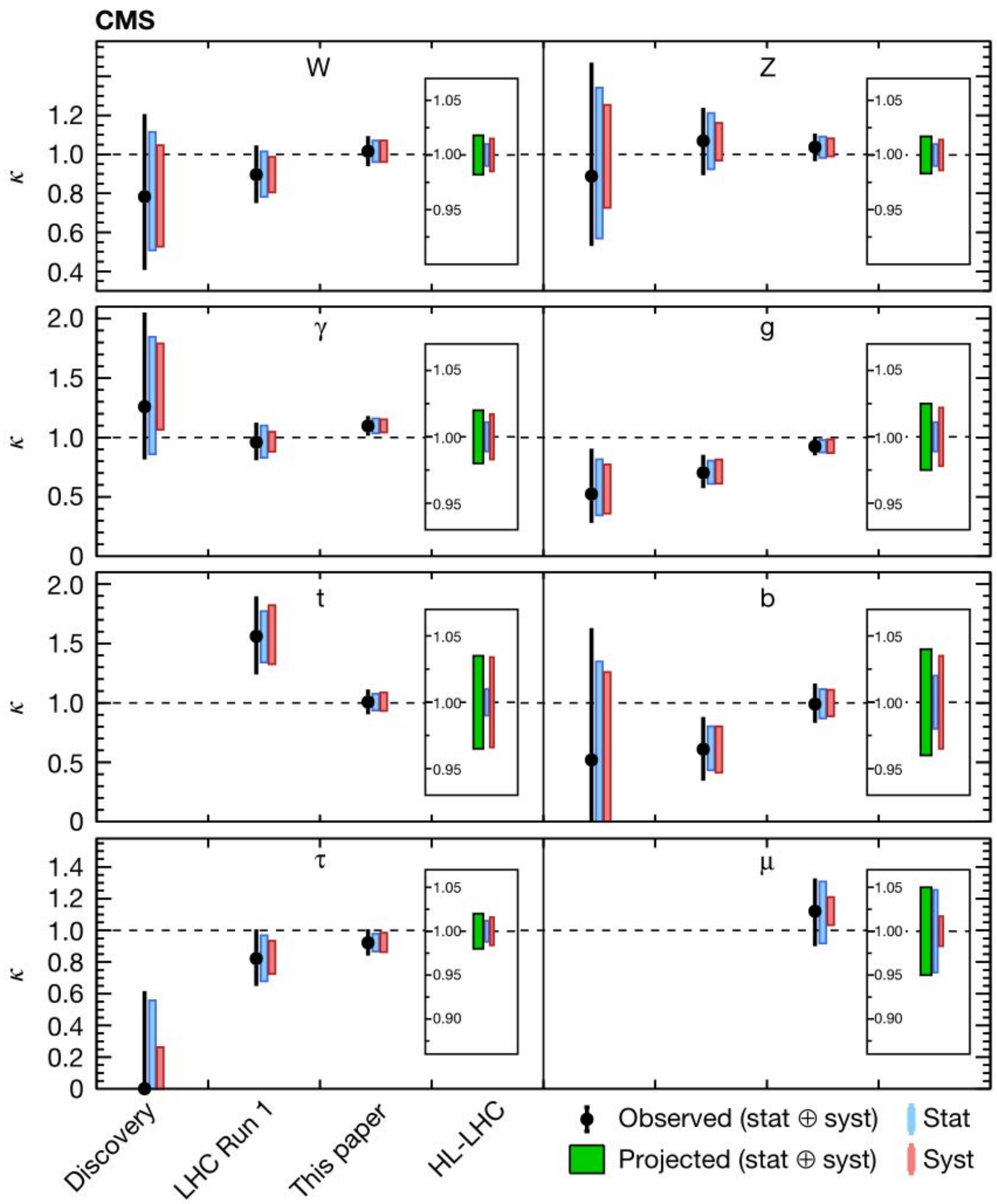
- The most important measurements on Higgs boson
  - Couplings with all SM massive particles
  - Property measurements: mass, Spin, CP, Parity, Width ...

Require Higgs Factory for the precision measurement





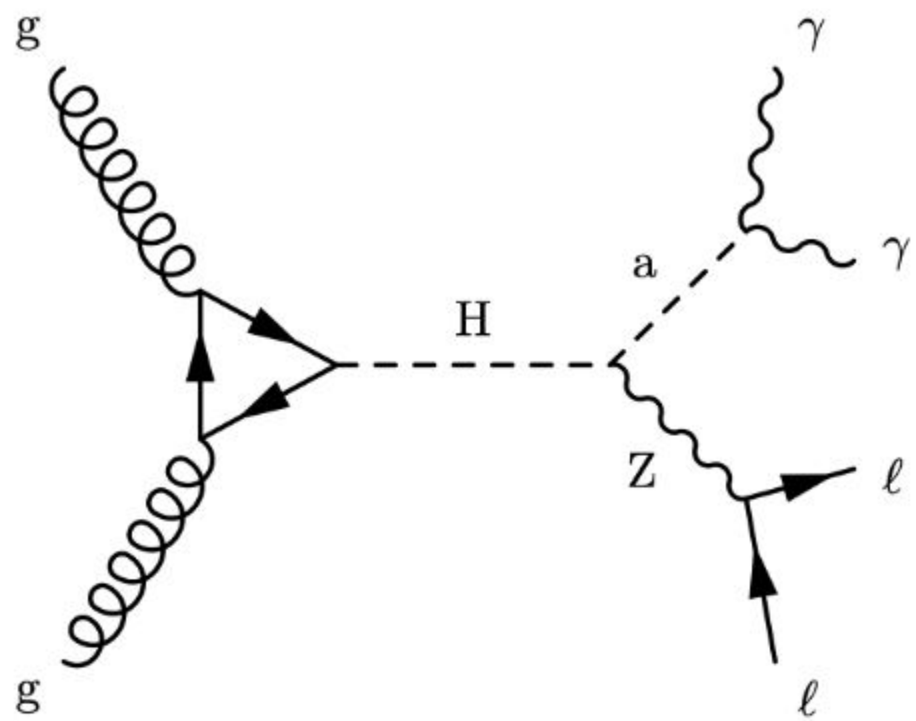
Great success have been achieved in both CMS and ATLAS, good agreement with SM predictions.



- The measured results converge to SM prediction → SM is obviously a great success!
- But on the other hand:
  - Neutrino mass
  - μ g-2
  - ...

Search for something new (new particle/new interactions), but should have small effect on existing SM process. E.g., SM Higgs rare decay or BSM Higgs decay, e.g. **Higgs exotic decay** (**this talk**)

H → Za → 2 lep + 2 Gamma



2311.00130
sub to PLB
138 fb^{-1}
HIG-22-003

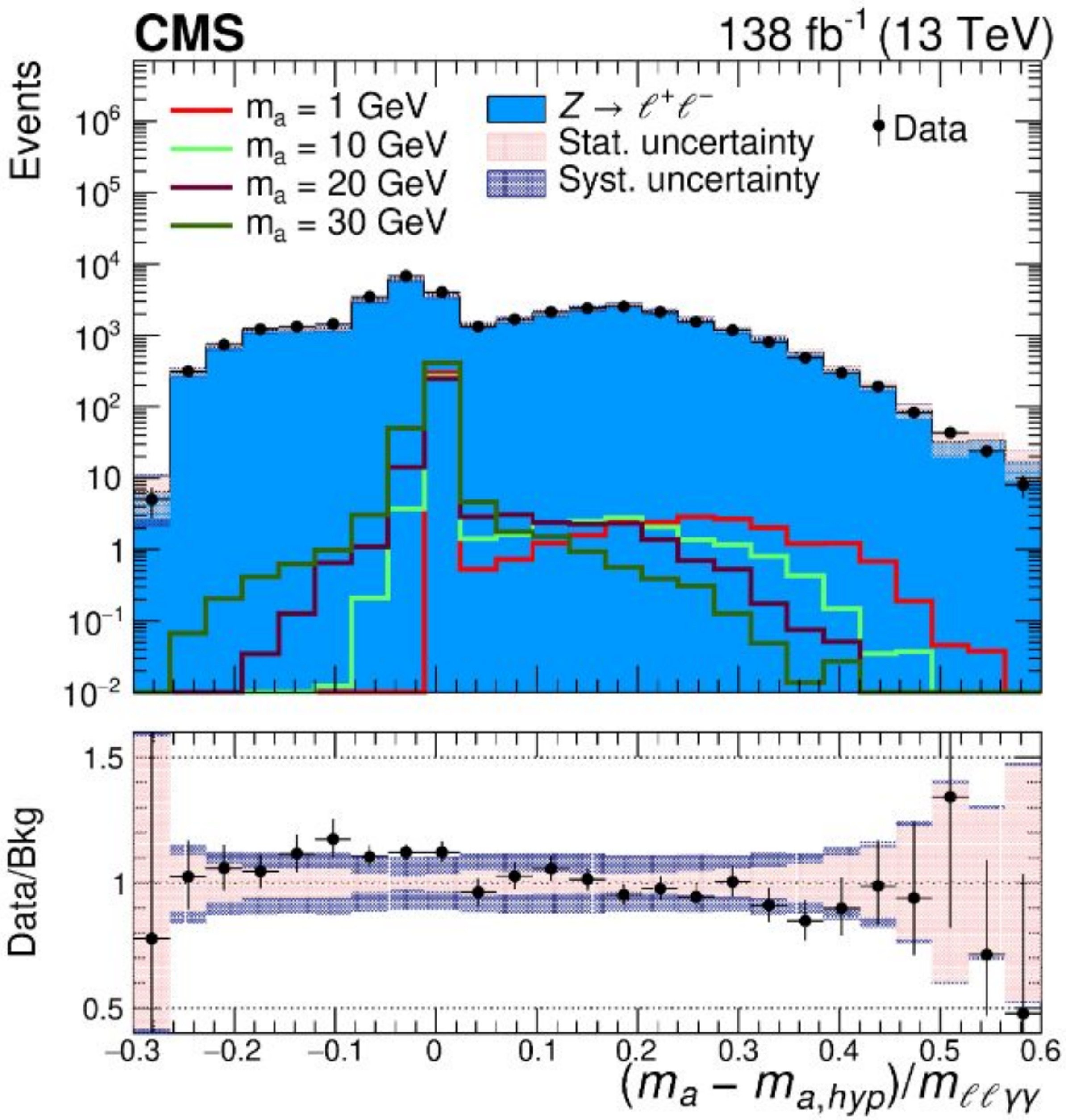
Motivation
- The ALP a can explain the μ g-2
- ALP couples Higgs boson

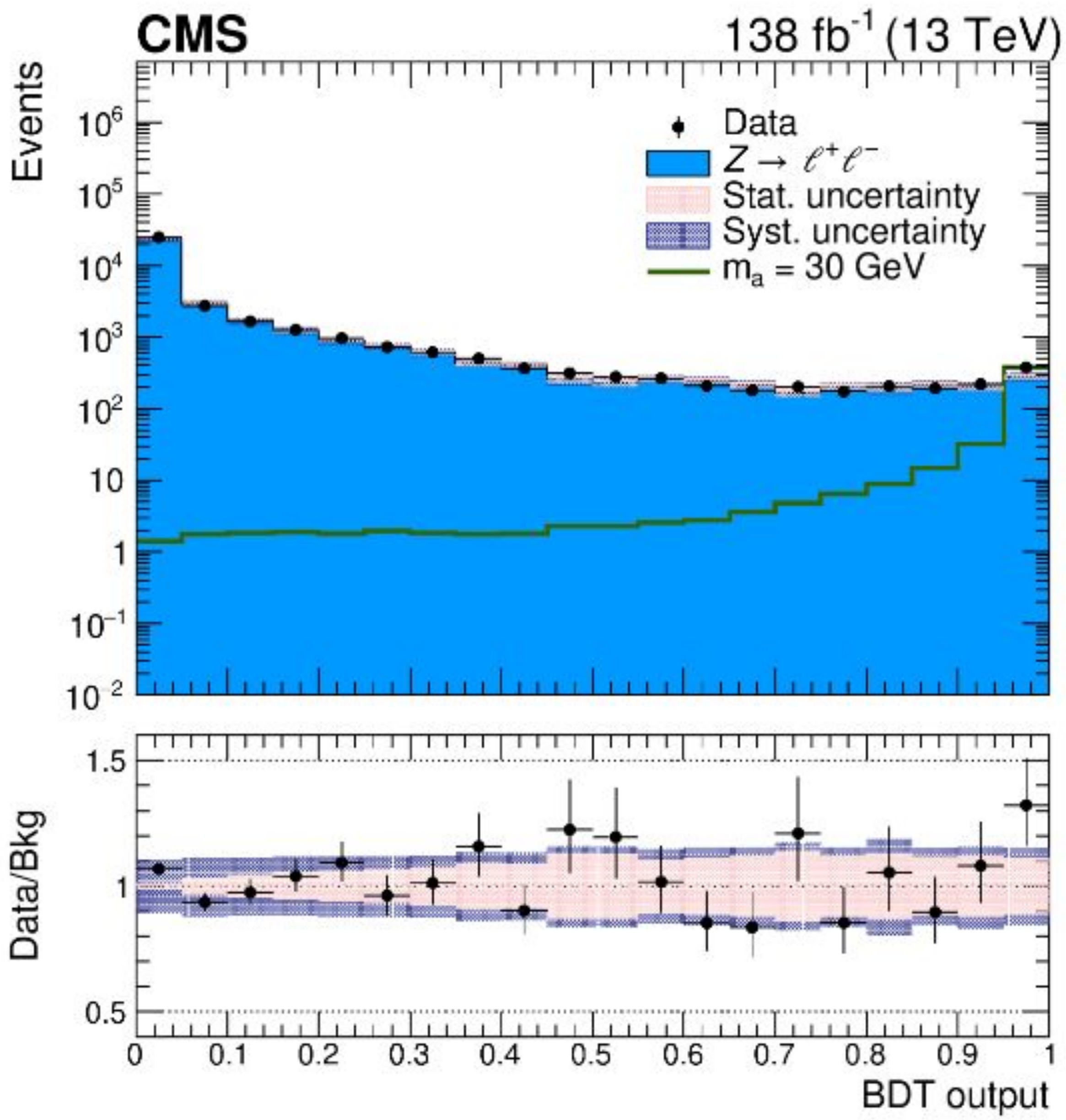
Signal setup
- Only the dominant production mode ggH is considered
- On-shell Higgs boson and Z boson, $m_a = m_H - m_Z$ ~< 30 GeV
- If m_a < 1 GeV, di-photon can not be separated anymore in the detector, m_a = [1,2,3,4,5,6,7,8,9,10,15,20,25,30] GeV

Dominant Bkg:
- DY, estimated using data events

Event selections
- 2 eles/muons with same flavor opposite-sign charged
  - FSR photon recovered → 1-4% improvement on m_H resolution
- 2 photons with pt > 10 GeV

H → Za → 2 lep + 2 Gamma: BDT





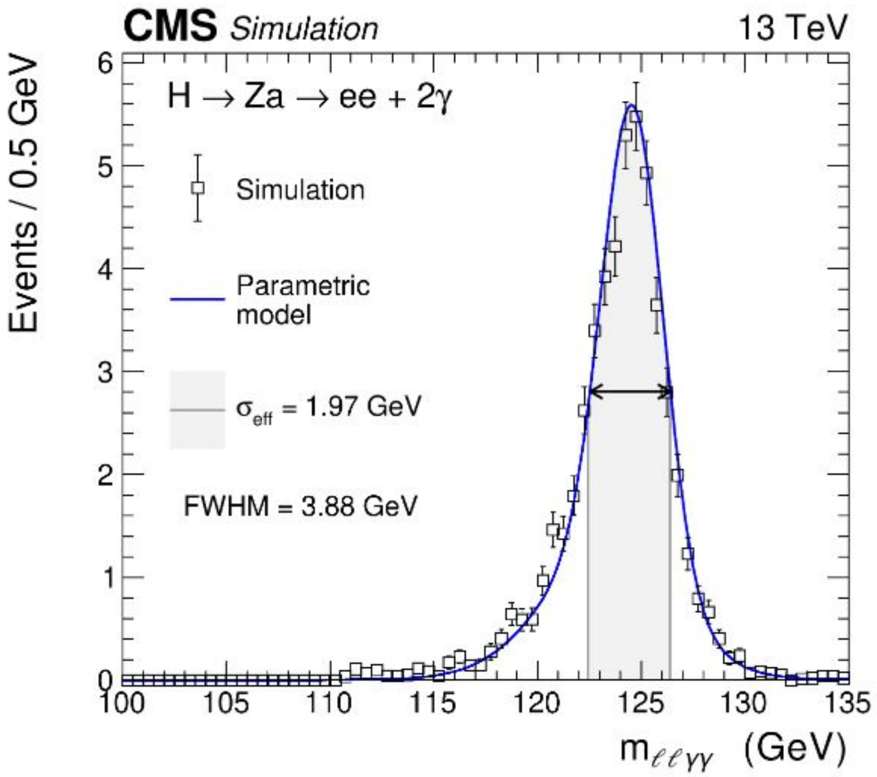
One of the input variable for BDT

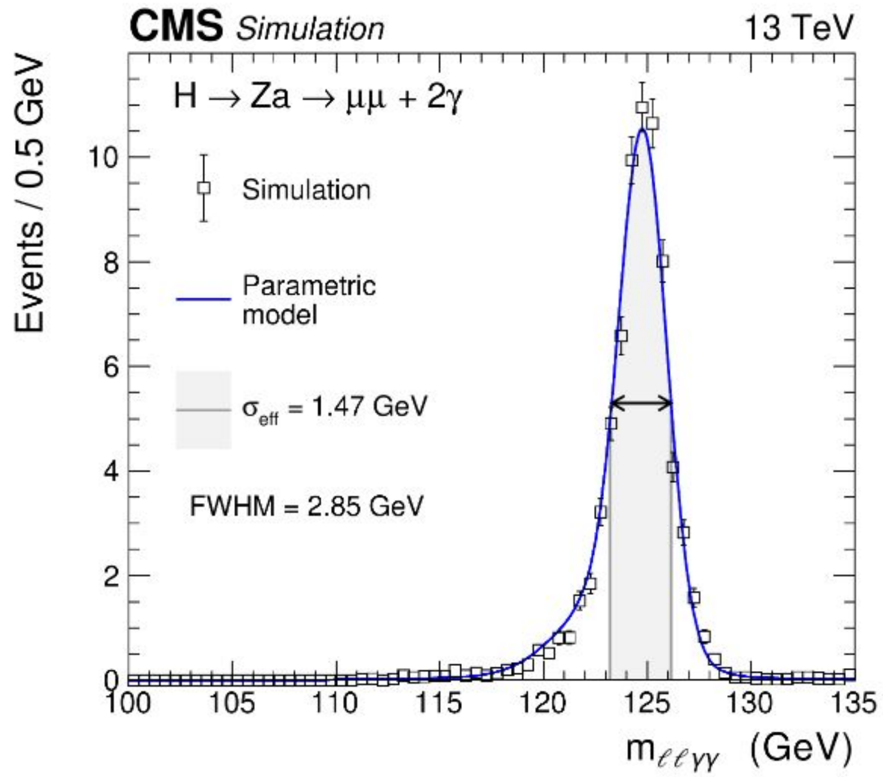
BDT score for m_a=30 GeV

Parametrized BDT is used to separate signal and the bkg DY.

- $(m_a - m_{a\_hyp})/m_{ll\gamma\gamma}$ is the input parameter for the parametrization purpose.
  - m_a: the reconstructed mass of a, i.e., invariant mass of two photons
  - m_{a_hyp}: for each signal sample, it's the mass the a in the MC model. For data and DY, it's a random number from a flat distribution of within the possible value of the search mass range

H → Za → 2 lep + 2 Gamma: shape fit





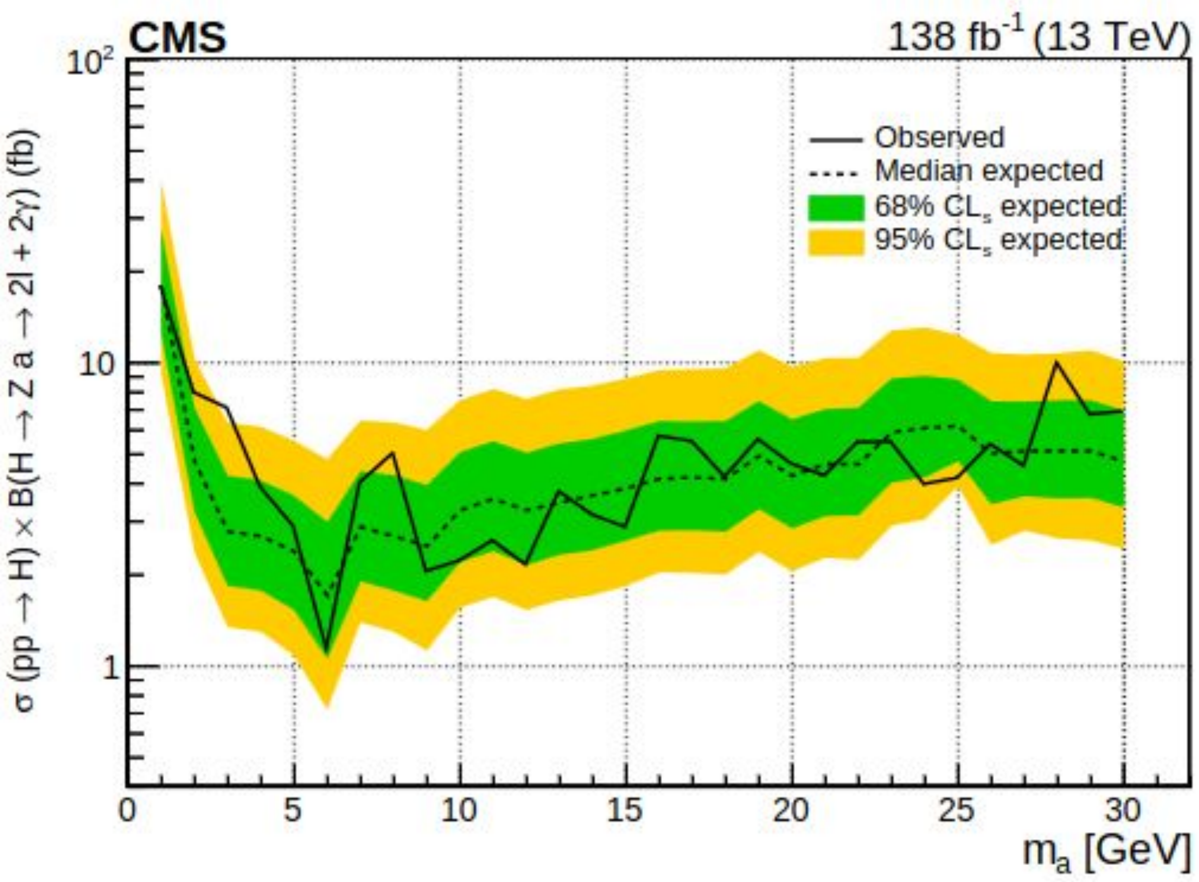
Signal shape:

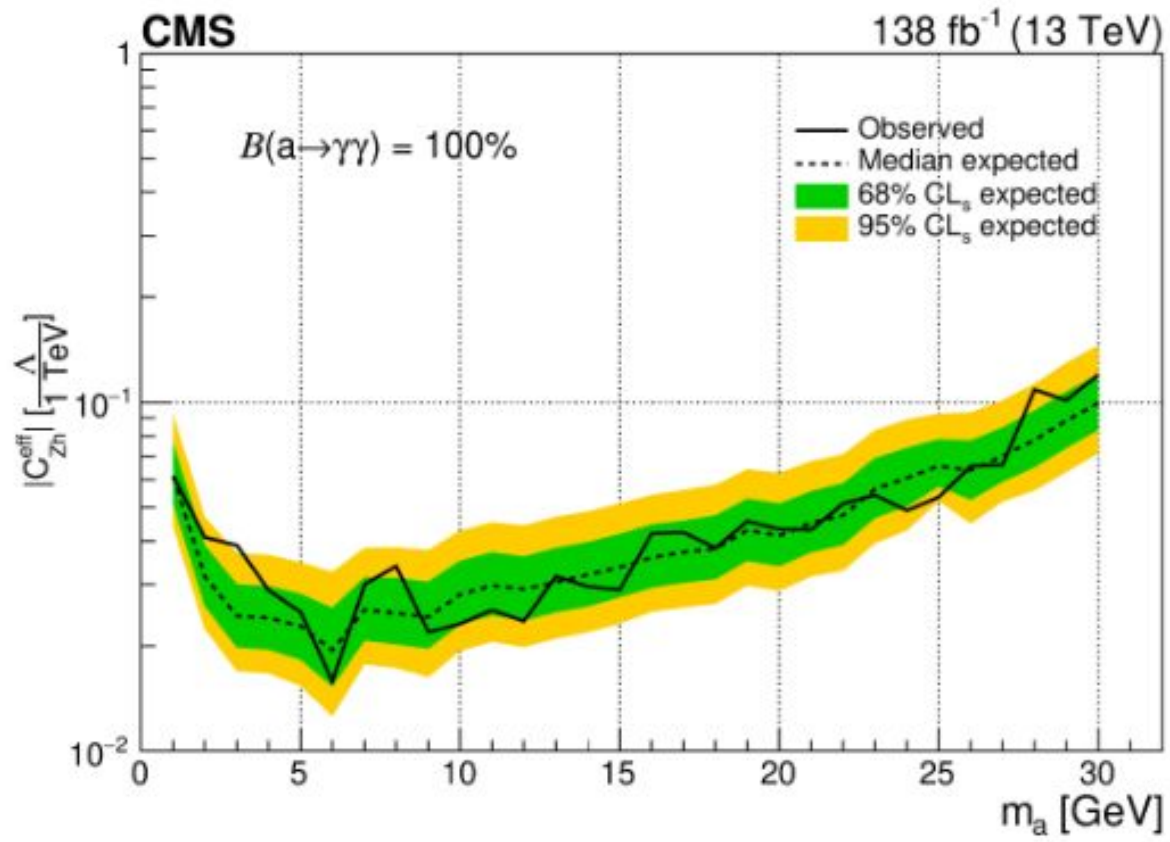
- Fit signal MC using sum of n Gaussian functions (n < 5)
- For each mass point and each year and ele/muon
- Parametrize the normalization for the intermediate mass point between 10-30 GeV

$$\mathcal{F}(m_{\ell\ell\gamma\gamma};\mu,\sigma,s,\vec{\alpha}) := \int \mathcal{N}(\mu,\sigma)(m_{\ell\ell\gamma\gamma}-t)f(t;\vec{\alpha})\Theta(s,t)\,\mathrm{d}t,$$

DY shape:

- Gaussian convoluted with the product of a falling spectrum component and a step function
- Different falling functions are studied, the difference is considered as systematics
- The fit is performed over the range 95 < $m_{ll\gamma\gamma}$ < 180 GeV
- Combine all data to have more events for the fit





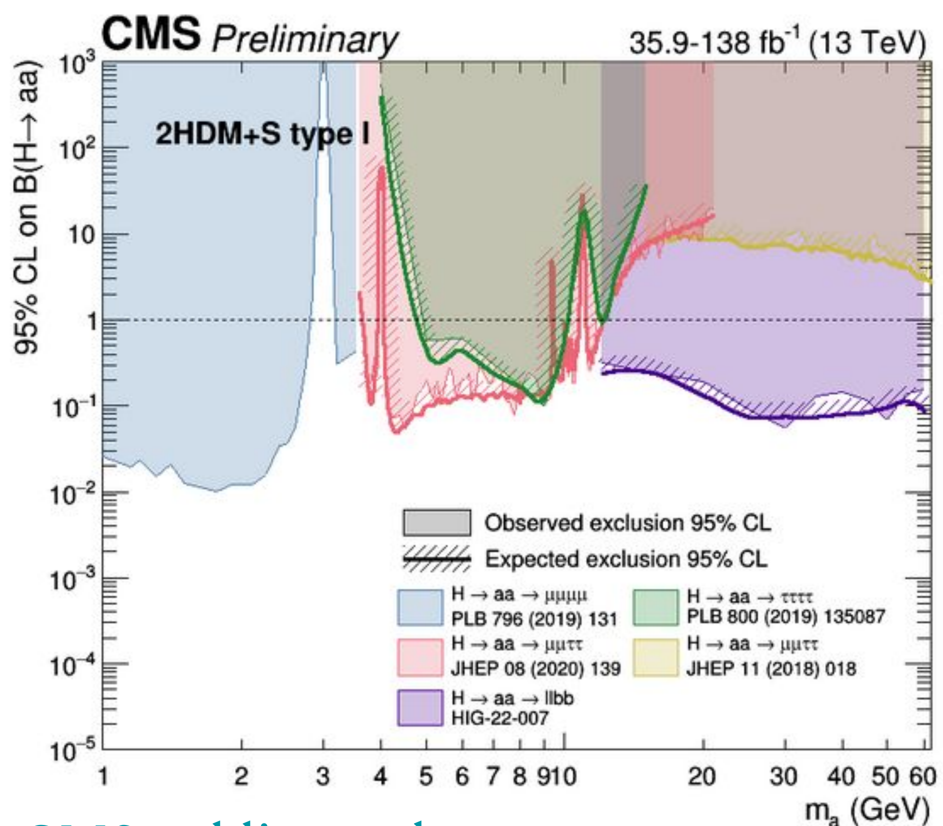
Limits on XS*BR.

An excess of data above the expected SM background with 2.6 (1.3) σ local (global) significance is observed for a mass hypothesis of m_a = 3 GeV

Limits on Effective coupling parameter of the Higgs boson, Z boson and the scalar, Assuming decay branching fraction B(a → γγ) = 100%

Motivation:

- 2HDM+S, two Higgs doubles extended with one scalar → Seven scalar and pseudoscalar, one of them is H(125 GeV), and it can decays to a pair of pseudoscalar a1.
- FCNC condition: require only one of doublets can couple to fermion, then only four types of 2HDM+S.



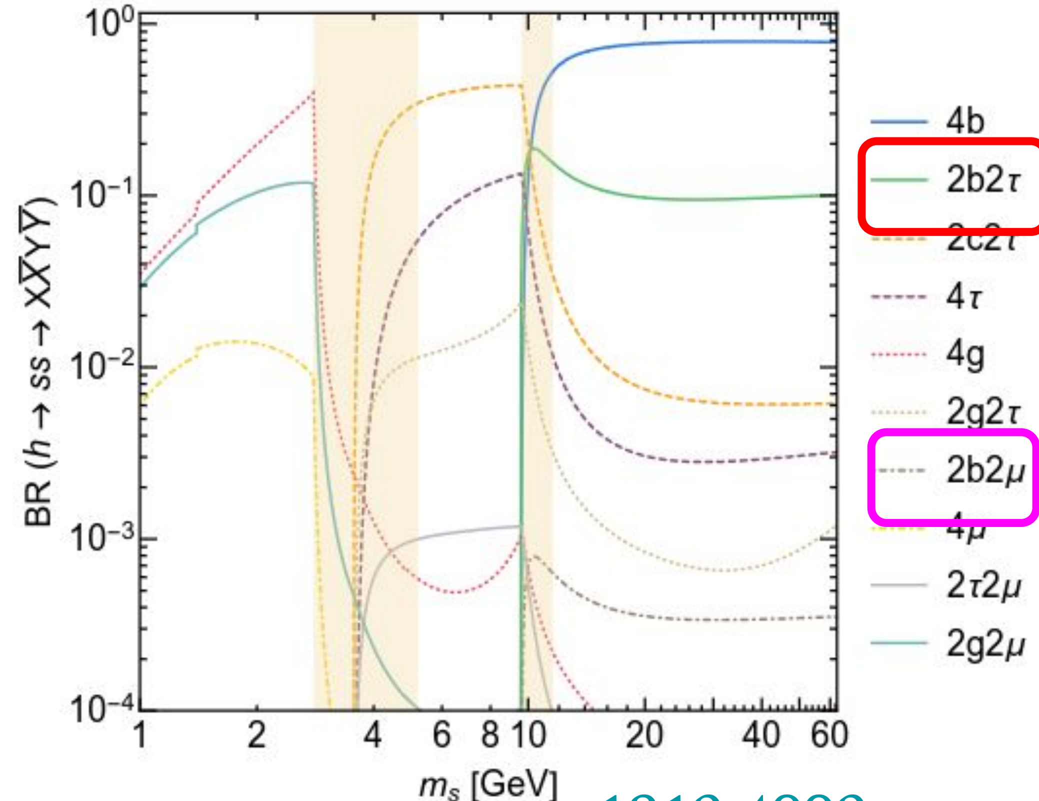
CMS public results

- A set of limits set on four types of 2HDM+S, under different tanβ

Signal model: HIG-22-007 138 fb^{-1}

- BR(H → a1a1) depends on the model type, mass of a1, and the ratio of the vev of the two doublets tanβ
- Bottom plot shows cross section of different decay channel in Type-I 2HDM+S.
  - 2b2τ is considered due to the large cross section and lepton (triggering).
  - 2b2μ channel with low cross section, but provides competitive results with 2b2τ due to the great muon reconstruction
- ggF and VBF Higgs productions are considered
- Model NMSSMHET (a special case of Type-II 2HDM+S) is used for signal MC modeling

- 2b2τ: 11 pseudoscalar masses in [12, 60] GeV are considered
- 2b2μ: 5 GeV step in m_{a1} for ggF, three mass points 20/40/60 GeV for VBF production



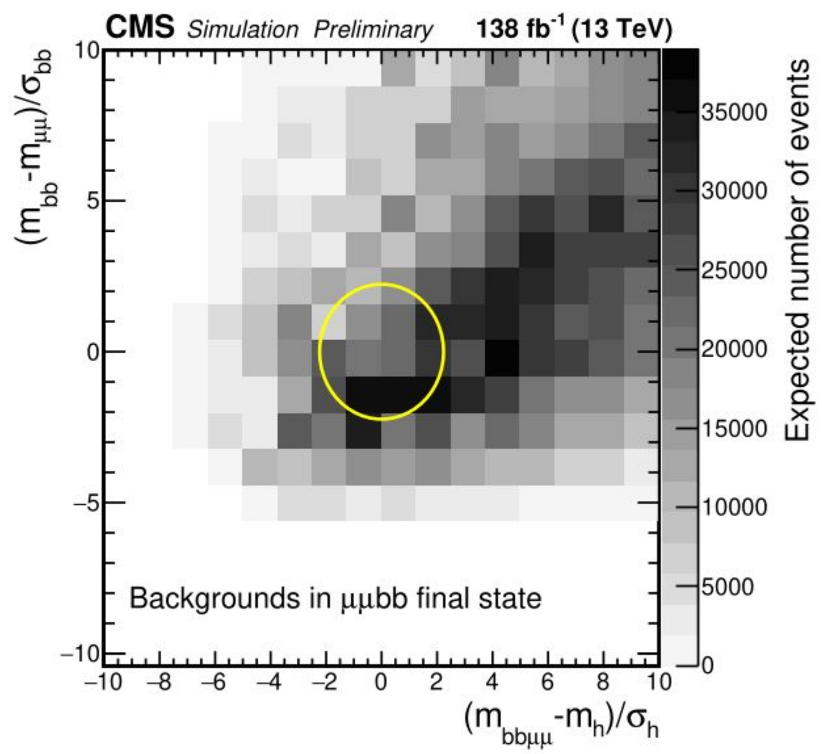
1312.4992

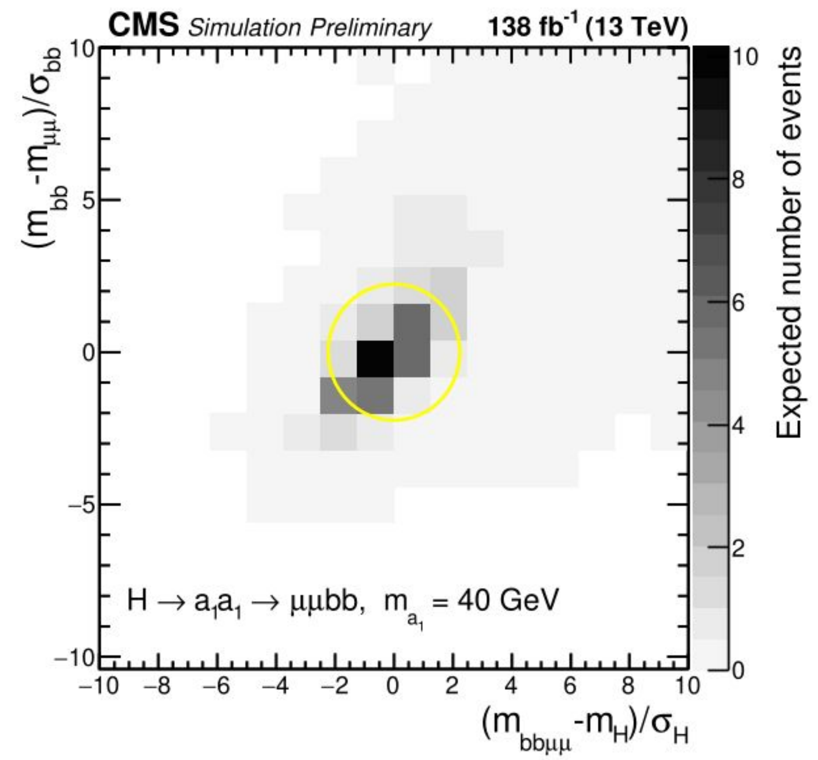
Event selection and category:

- For 2b2μ:
  - Events are selected with at least two b-jets
  - $M_{\mu\mu}$ ~ [14, 70] GeV
  - MET < 60 GeV, to suppress tt~ contribution

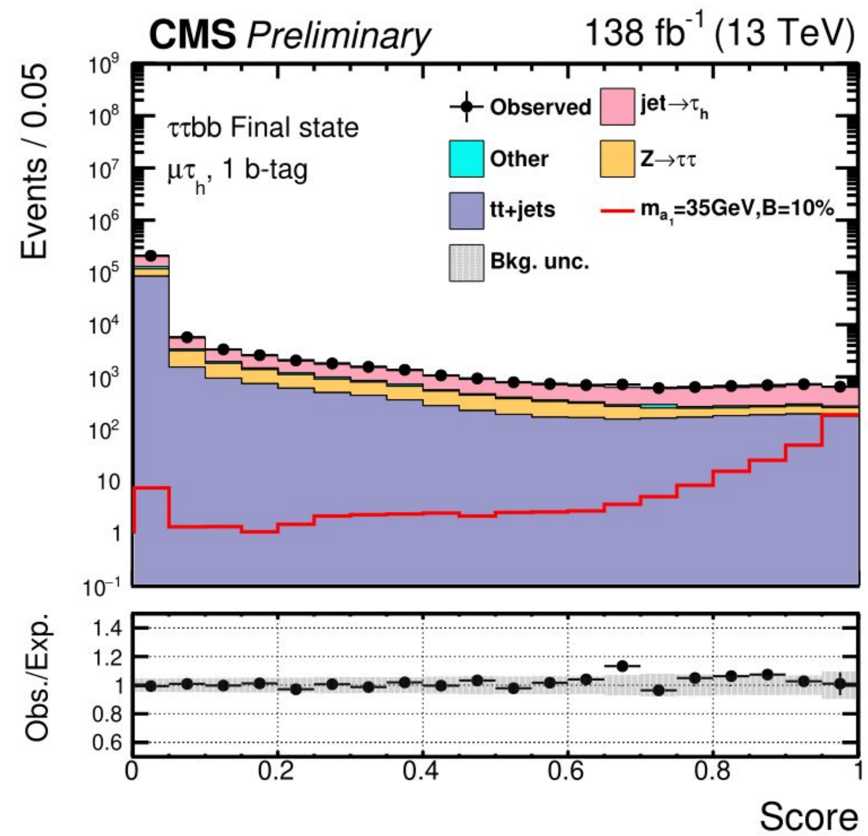
$$\chi_{\rm bb} = \frac{(m_{\rm bb} - m_{\mu\mu})}{\sigma_{\rm bb}}, \quad \text{and} \quad \chi_{\rm H} = \frac{(m_{\mu\mu\rm bb} - 125)}{\sigma_{\rm H}}$$

$$\chi^2_{\rm tot} = \chi^2_{\rm bb} + \chi^2_{\rm H}$$





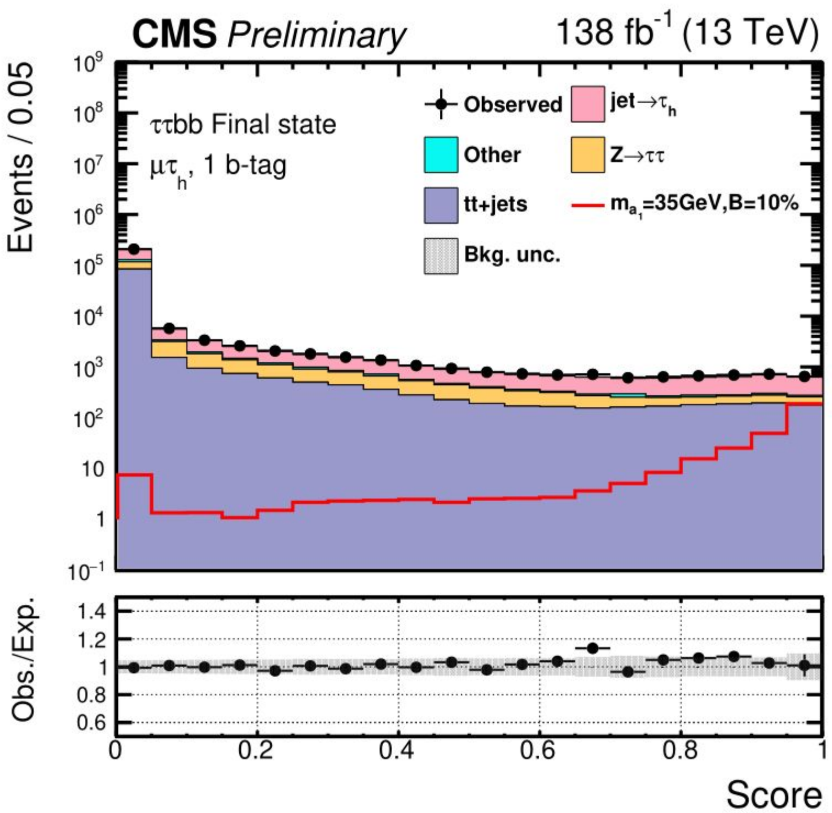
- For 2b2τ:
  - Considering eμ, eτ_h and $\mu\tau_h$ in the final state.
  - For each of them, events are categorized to exactly 1 b-jet and >1 b-jet
  - BDT is used to separate signal and bkg
  - Cut on BDT to suppress further the bkg, then use $m_{\tau\tau}$ to extract signal significance



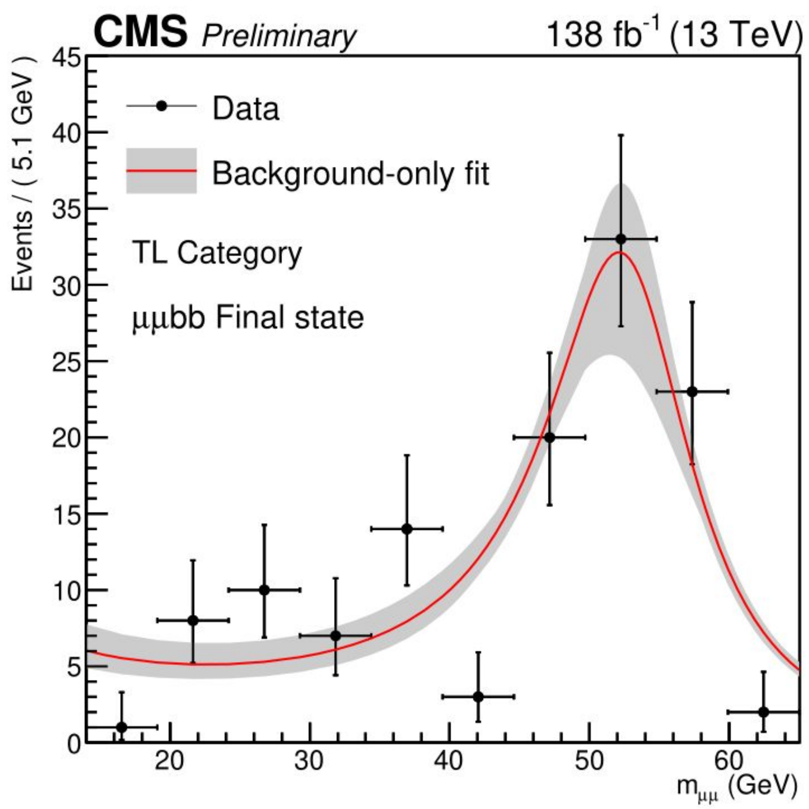
H → aa → 2tau + 2b: bkg estimation

For 2b2τ, dominant bkgs are:

- Z→ 2τ: estimated from data using embedding technique.
- QCD jet mis-identified as e/mu (eμ channel): evaluate the ratio of OS/SS lepton in SR using simulation, scale the SS data events with the ratio
- Jet mis-identified as hadronic τ: calculate the probability (f) of a jet to be identified as the τ_h, apply f/(1-f) on the sideband event
- Remove double counting in ttbar events, e.g., remove eμ events if no matched eμ at generator level



BDT score comparison after backgrounds estimation in 2b2τ channel

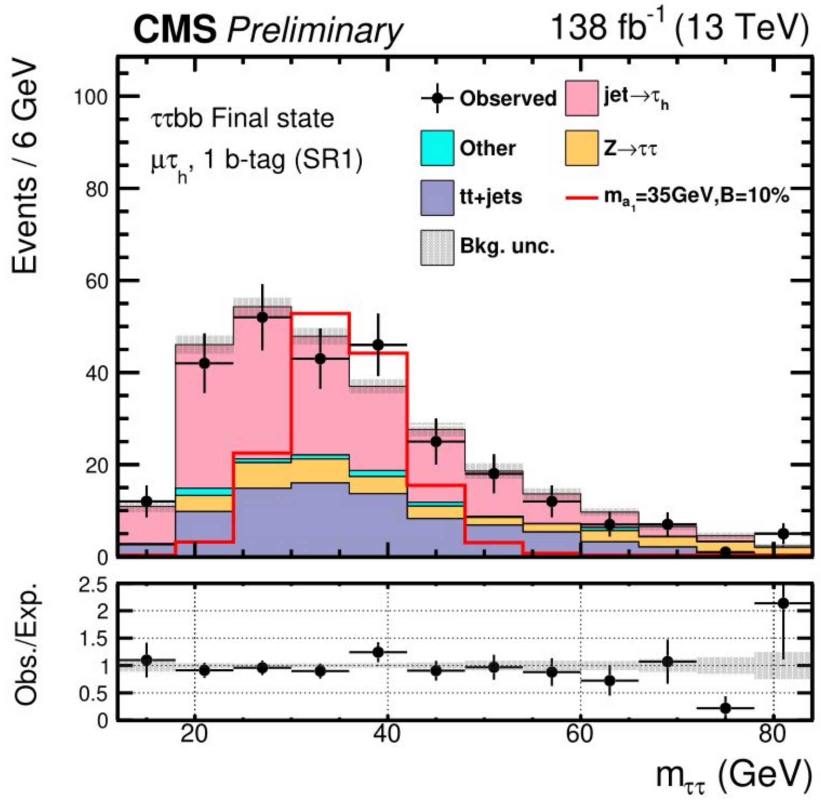


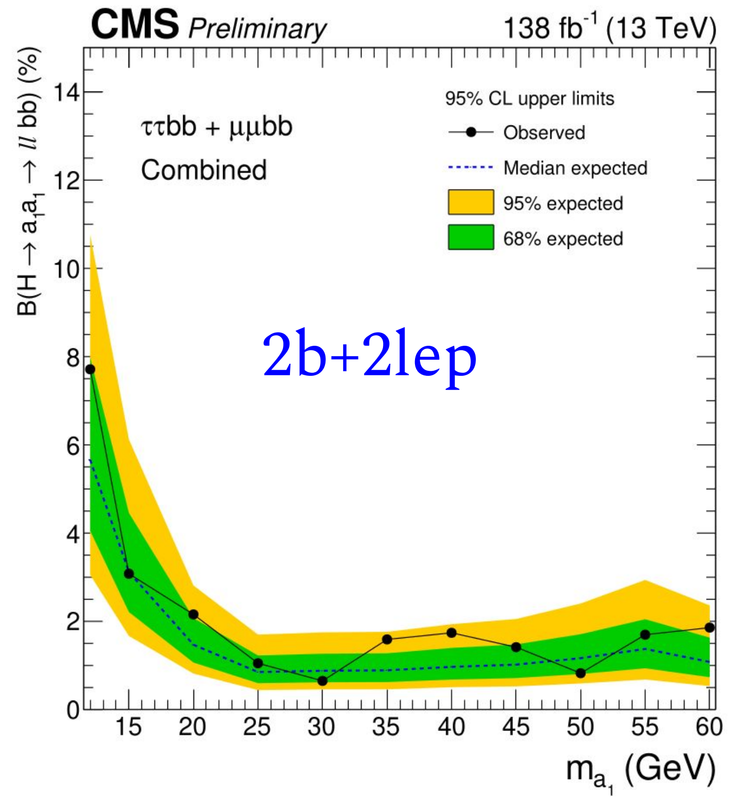
Background fit in 2b2μ channel

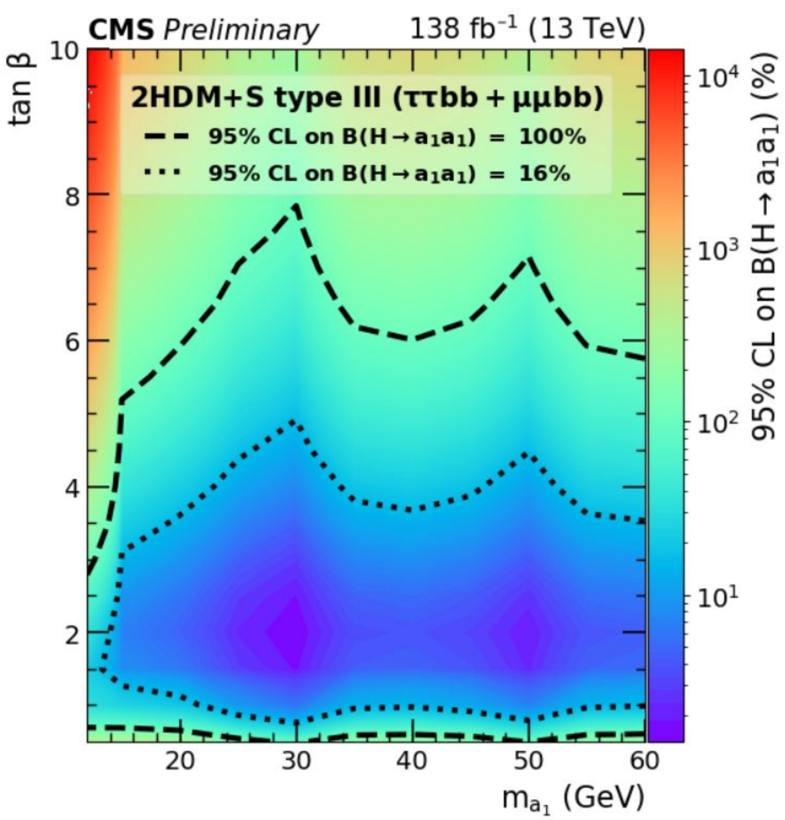
For 2b2μ, bkg is collectively estimated from data. Use polynomials+inverse polynomial to model the bkg for each event category.

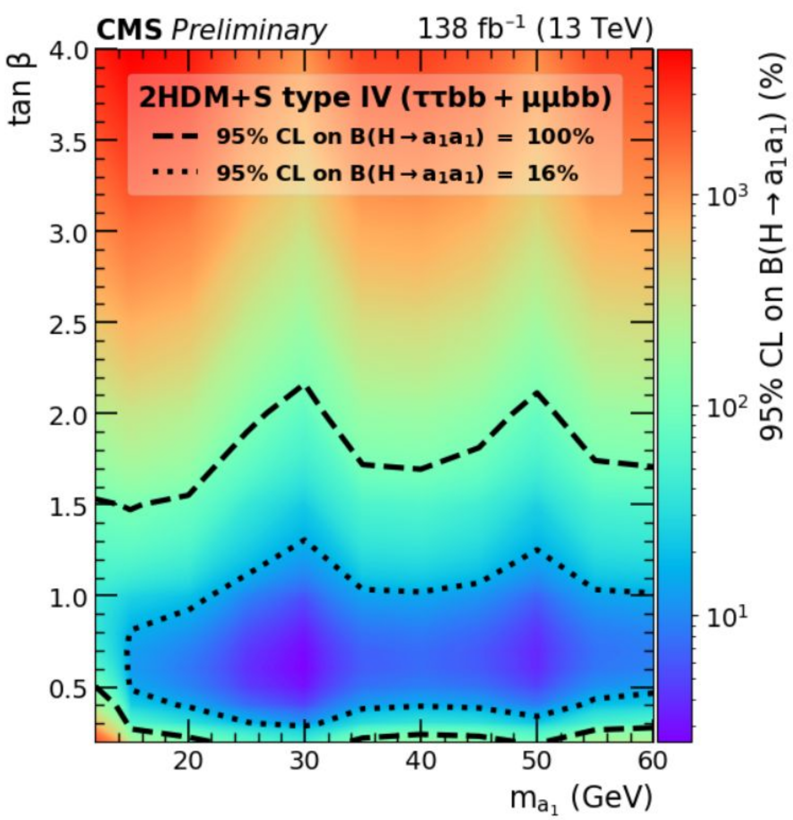
H → aa → 2mu/tau + 2b: fit and results

- For 2b2μ: unbinned fit in SR, signal shape obtained from simulation using weighted sum of a Voigt profile and a Crystal Ball functions
- For 2b2τ:
  - Binned fit using invariant mass of two τ, $m_{\tau\tau}$ (top left plot), obtained using SVFit.









- The brasil plot in the top right shows the 95% CL limit independent of the model type and tanβ
- The 2D limits plots in the left show the exclusion phase space in Type-III and Type-IV w.r.t m_{a1} and tanβ

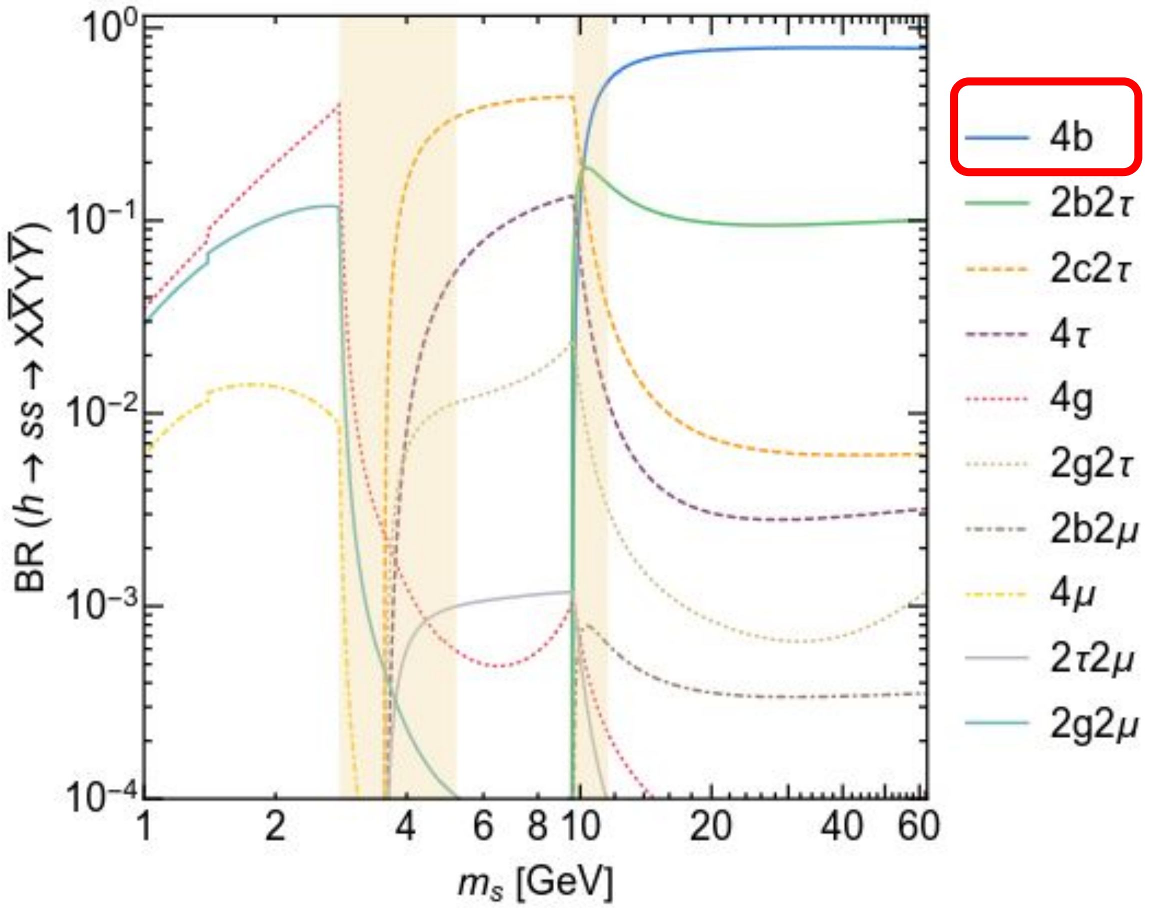
H → aa → 4b

New for Higgs2023

HIG-18-026 138 fb^{-1}

Motivation:

- Study 4b channel under the context of 2HDM+S



- 4b: with largest cross section. Considering WH and ZH production with leptonic decay of V → easy to trigger those events
- mass=[12, 15, 20, 25, 30, 40, 50, 60] GeV

Selections:

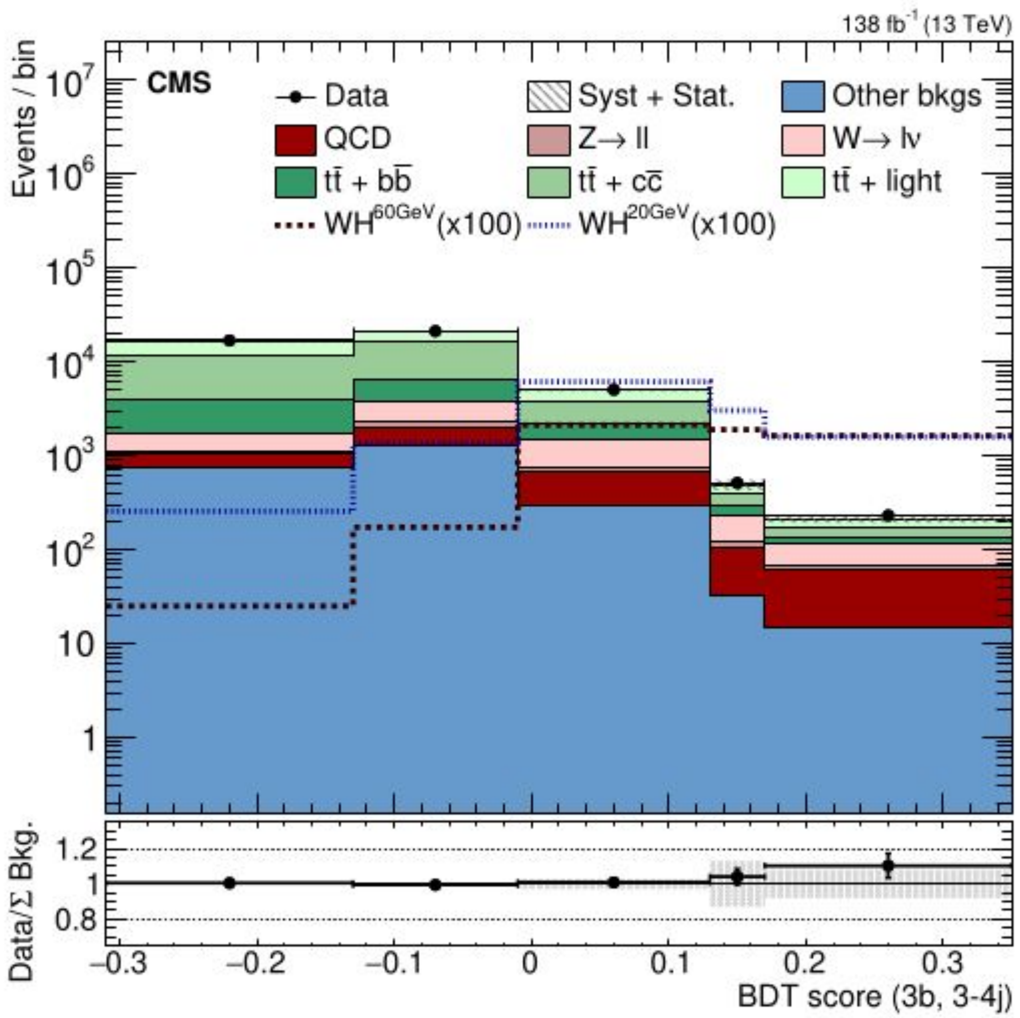
- WH: one electron or one muon, requirement on MET and transverse mass m_T
- ZH: two electrons or two muons with opposite charge
- Both WH and ZH: at least two jets and at least two of them are tagged as B-jets

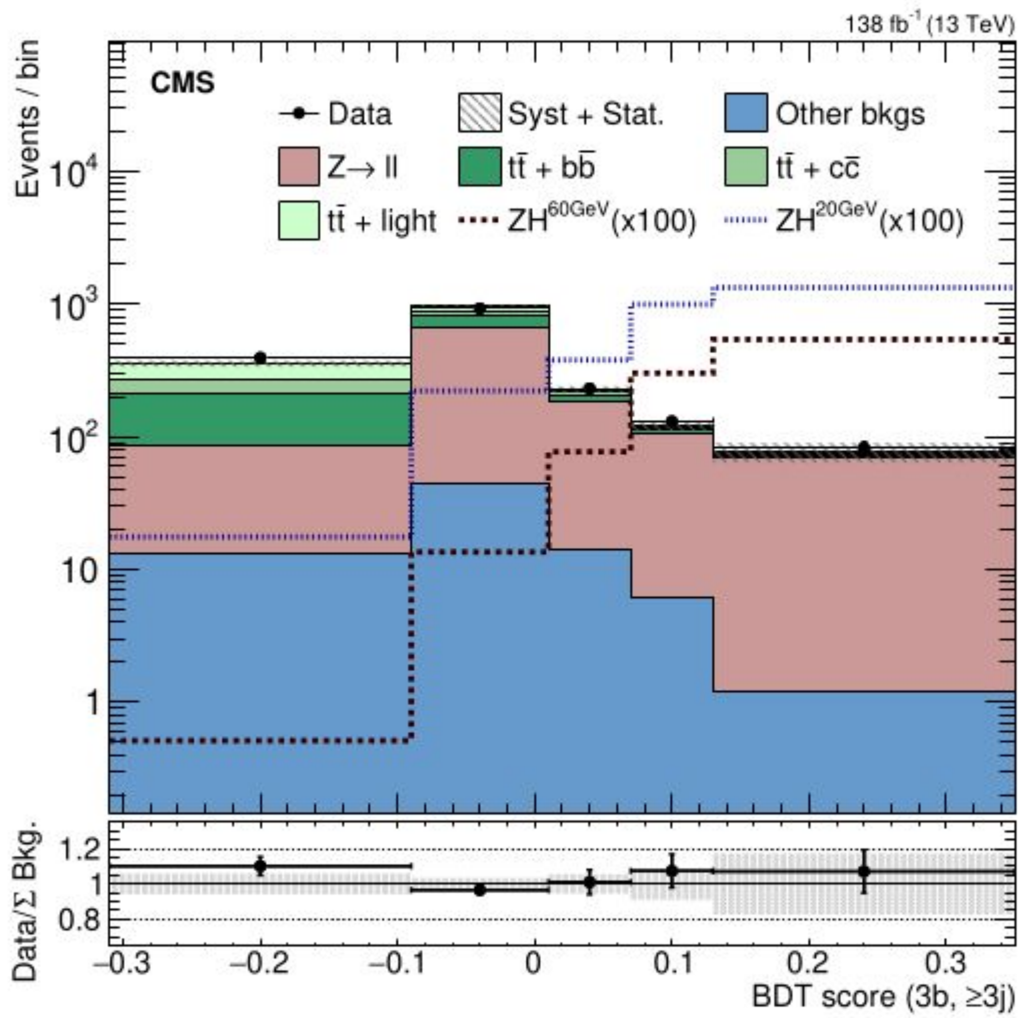
	3b category	4b category
	WH channel	
SR	(3b, 3–4j)	(4b, 3–4j)
CR	(2b, 3j)	(2b, 4j)
	ZH channel	
SR	(3b, ≥3j)	(4b, ≥4j)
CR	(2b, 3j)	(2b, 4j)

New for Higgs2023

Background estimation:

- tt+jet:
  - tt+light flavor (LF): shape and normalization are from simulation
  - tt+ b/c jets (HF): shape and relative rate from simulation, overall normalization from data fit
- V+jet: shapes from simulation, overall normalization from data fit
- Multi-jets in WH channel: ratio method (slide 10) is used

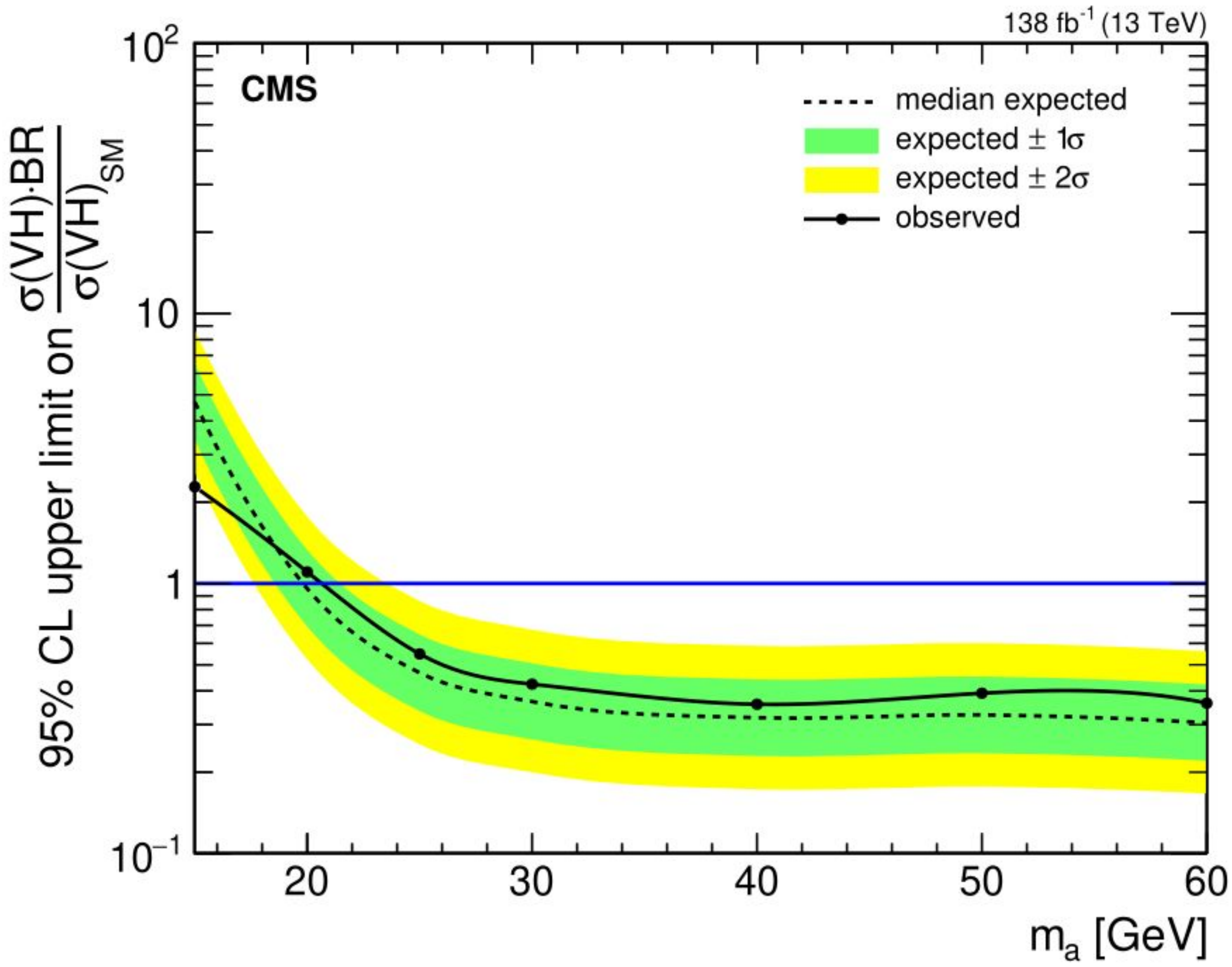




BDT is used to separate signal and backgrounds, good agreements between data and SM predictions

- Left: WH channel
- Right: ZH channel

H → aa → 4b: results



New for Higgs2023

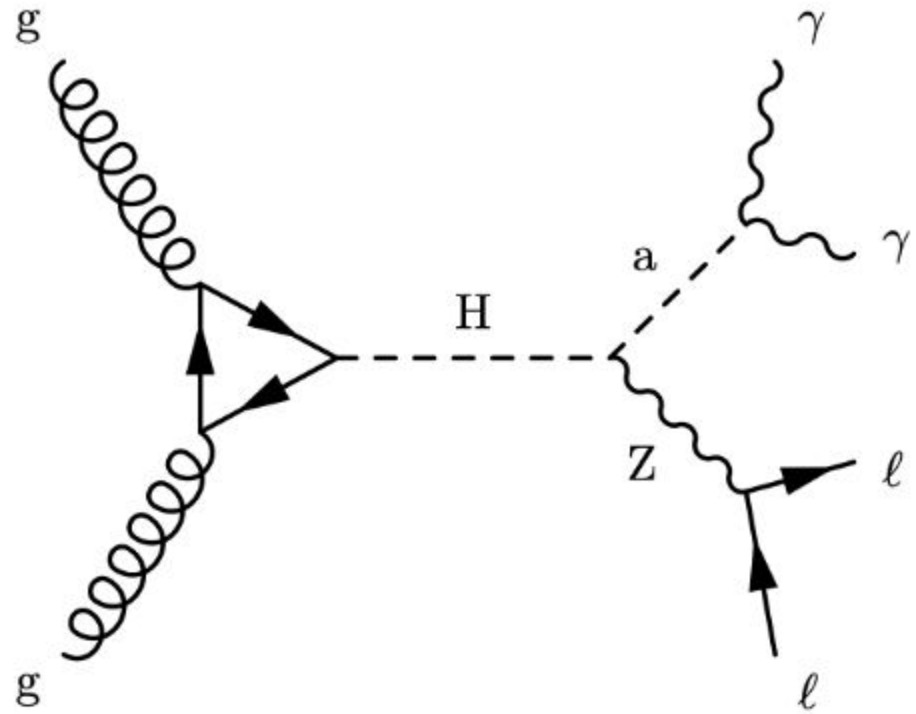
Observed and expected limits on the signal strength μ = σ(VH)B(H → aa →bbbb)/σ(VH)SM with the WH and ZH channels combined. The solid blue line indicates the SM cross section σ(pp → VH) with branching fractions B(H → aa) = 1 and B(a → bb) = 1

Summary:

- A search for H → Za → 2 lep + 2 Gamma is reported, An excess of data above the expected SM background with 2.6 (1.3) σ local (global) significance is observed for a mass hypothesis of ma = 3 GeV
- Two searches for H → aa → 2tau/2mu + 2b and H → aa → 4b are shown, no significant excess is observed
- Exotic decay of Higgs boson still show good agreement with the SM, looking forward to the results at 13.6TeV and the future Higgs factory to bring more clues on BSM

Thanks!

Additional materials



2311.00130
sub to PLB
138 fb^{-1}
HIG-22-003

Motivation

- The ALP a can explain the μ g-2
- ALP couples Higgs boson

Signal setup

- Only the dominant production mode ggH is considered
- On-shell Higgs boson and Z boson, $m_a = m_H - m_Z$ ~< 30 GeV
- If m_a < 1 GeV, di-photon can not be separated anymore in the detector, m_a = [1,2,3,4,5,6,7,8,9,10,15,20,25,30] GeV

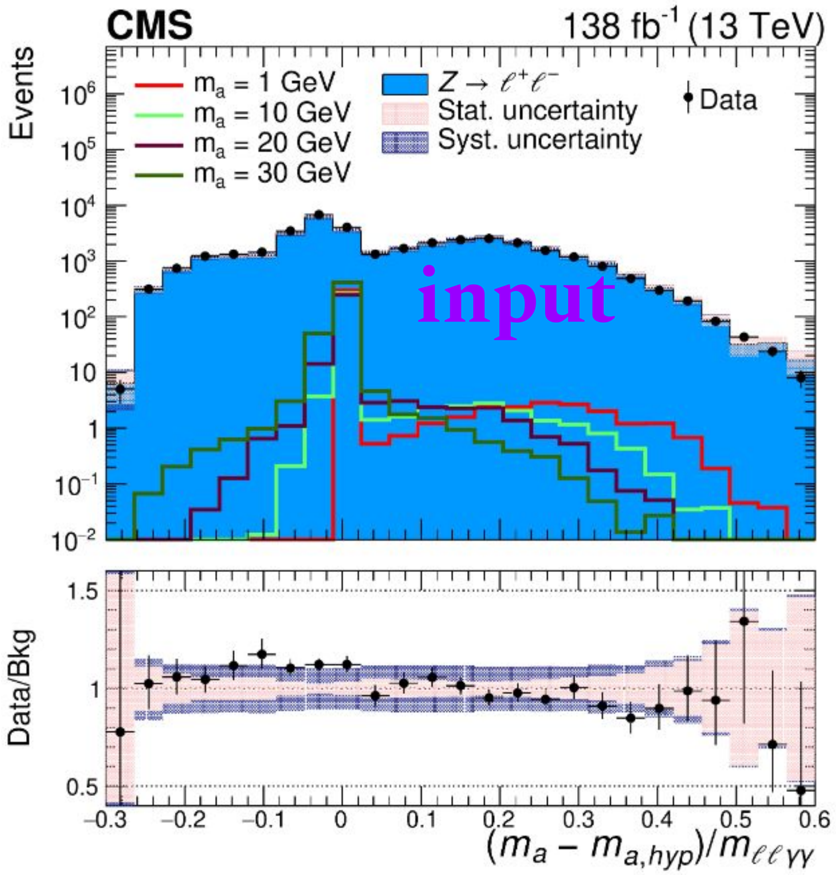
Dominant Bkg:

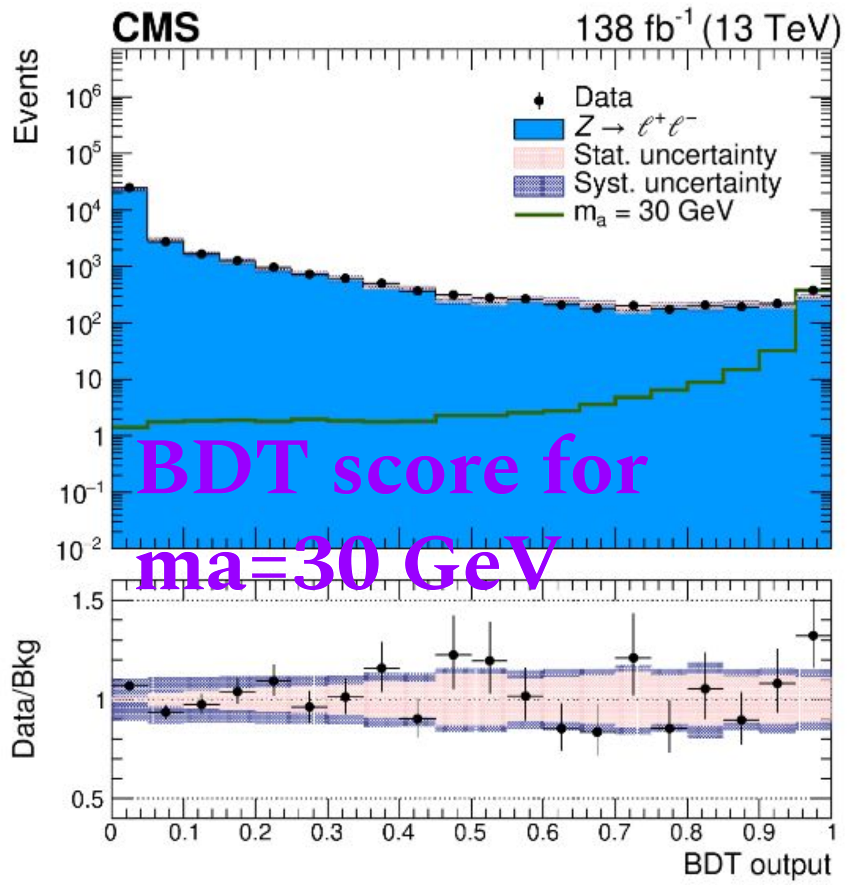
- DY, estimated using data events

Event selections

- 2 eles/muons with same flavor opposite-sign charged
  - FSR photon recovered → 1-4% improvement on mH resolution
- 2 photons with pt > 10 GeV
- $\Delta R(lep, \gamma) > 0.4$
- mll > 50 GeV
- 95 GeV < $mll\gamma\gamma$ < 180 GeV
- 115 GeV < $mll\gamma\gamma$ < 135 GeV → SR

H → Za → 2 lep + 2 Gamma: BDT





Parametrized BDT is used to separate signal and the bkg DY.

- (ma - ma_hyp)/mll$\gamma\gamma$ is the input parameter for the parametrization purpose.
  - ma: the reconstructed mass of a, i.e., invariant mass of two photons
  - ma_hyp: for each signal sample, it's the mass the a in the MC model. For data and DY, it's a random number from a flat distribution of within the possible value of the search mass range
- All the other input parameters are chosen with weak correlation with mll$\gamma\gamma$ the which will be used for signal extraction fit.

$$\mathrm{AMS} = \sqrt{2\left[(S+B)\ln\left(1+\frac{S}{B}\right) - S\right]}$$

Select the BDT score cut to maximize the AMS above in the signal region, for each mass point.

H → Za → 2 lep + 2 Gamma: BDT inputs

- $p_{\mathrm{T}}(\gamma 1)$ and $p_{\mathrm{T}}(\gamma 2)$, where $\gamma 1$ and $\gamma 2$ are the leading and subleading photons, respectively;
- $R_9(\gamma 1)$ and $R_9(\gamma 2)$: the energy sum of the 3×3 crystal array centered around the most energetic crystal in the supercluster, divided by the energy of the supercluster;
- $\sigma_{i\eta i\eta}(\gamma 1)$ and $\sigma_{i\eta i\eta}(\gamma 2)$: the second moment of the log-weighted distribution of crystal energies in η, calculated in the 5×5 matrix around the most energetic crystal in the supercluster and rescaled to units of crystal size;
- $I_\gamma(\gamma 1)$ and $I_\gamma(\gamma 2)$: the isolation variable obtained by summing the p_{T} of photons inside an isolation cone of $\Delta R = 0.3$ with respect to the photon direction, while the impact of another selected photon is also included;
- $I_{\gamma,\mathrm{a}}$, the isolation variable obtained by summing the p_{T} of photons inside an isolation cone of $\Delta R = 0.3$ with respect to the direction of the ALP candidate;
- the angular separation between the Z boson and the diphoton pair, $\Delta R(\mathrm{Z}, \mathrm{a})$;
- the angular separation between the two photons, $\Delta R(\gamma_1, \gamma_2)$;
- the angular separation between the leading photon and the Z boson, $\Delta R(\gamma_1, \mathrm{Z})$;
- the ALP candidate's p_{T} divided by $m_{\ell\ell\gamma\gamma}$;
- the H boson candidate's p_{T};
- the difference between the invariant masses of the ALP candidate and the $m_{\mathrm{a,hyp}}$ parameter divided by $m_{\ell\ell\gamma\gamma}$, $(m_{\mathrm{a}} - m_{\mathrm{a,hyp}})/m_{\ell\ell\gamma\gamma}$.

H → Za → 2 lep + 2 Gamma: BDT

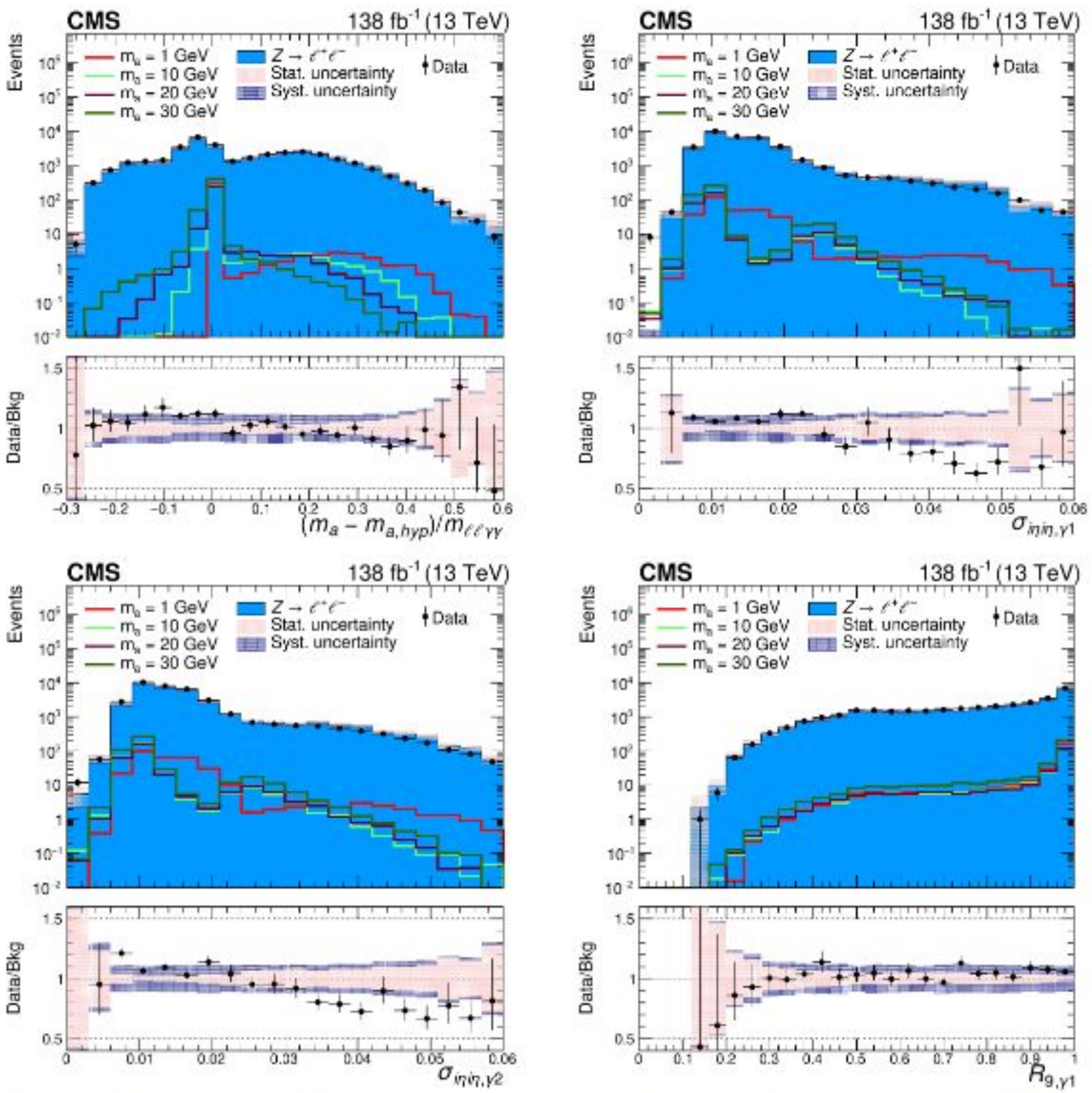


Figure 2: Distributions of the four most discriminating variables used as input to the BDT: $(m_a - m_{a,hyp})/m_{\ell\ell\gamma\gamma}$ (upper left), leading photon's $\sigma_{i\eta i\eta}$ (upper right), subleading photon's $\sigma_{i\eta i\eta}$ (lower left), and leading photon's R_9 (lower right). The events pass the selection criteria described in Section 5. The signal is scaled to a cross section of 0.1 pb and the background sample is normalized to an integrated luminosity of 138 fb^{-1}. The systematic uncertainties included in the shaded band are related to the photon efficiency, lepton efficiency, and pileup modeling. The impact of the remaining disagreement between data and simulation is negligible.

Table 1: Minimum BDT output values used to define the signal region, with the associated signal efficiencies and background yields. The statistical uncertainties is also shown.

m_a (GeV)	Min. BDT output value	Signal efficiency (%)	Drell–Yan background yields
1	0.955	49 ± 3.3	83 ± 27
2	0.980	67 ± 2.7	26 ± 10
3	0.985	76 ± 2.4	7.9 ± 4.9
4	0.980	84 ± 2.1	5.1 ± 4.5
5	0.985	85 ± 2.1	5.1 ± 3.9
6	0.990	82 ± 2.3	2.5 ± 2.2
7	0.985	86 ± 2.1	5.3 ± 4.0
8	0.990	80 ± 2.5	11 ± 4.8
9	0.990	78 ± 2.5	16 ± 5.6
10	0.990	77 ± 2.6	11 ± 4.7
15	0.990	70 ± 2.9	13 ± 5.2
20	0.990	63 ± 3.1	18 ± 6.1
25	0.985	64 ± 2.7	37 ± 11
30	0.980	67 ± 2.2	44 ± 13

$$\mathrm{AMS} = \sqrt{2\left[(S+B)\ln\left(1+\frac{S}{B}\right) - S\right]}$$

In Eq. (2), S and B refer to the number of signal and background (Drell–Yan simulation) events in the signal region

H → Za → 2 lep + 2 Gamma: signal shape

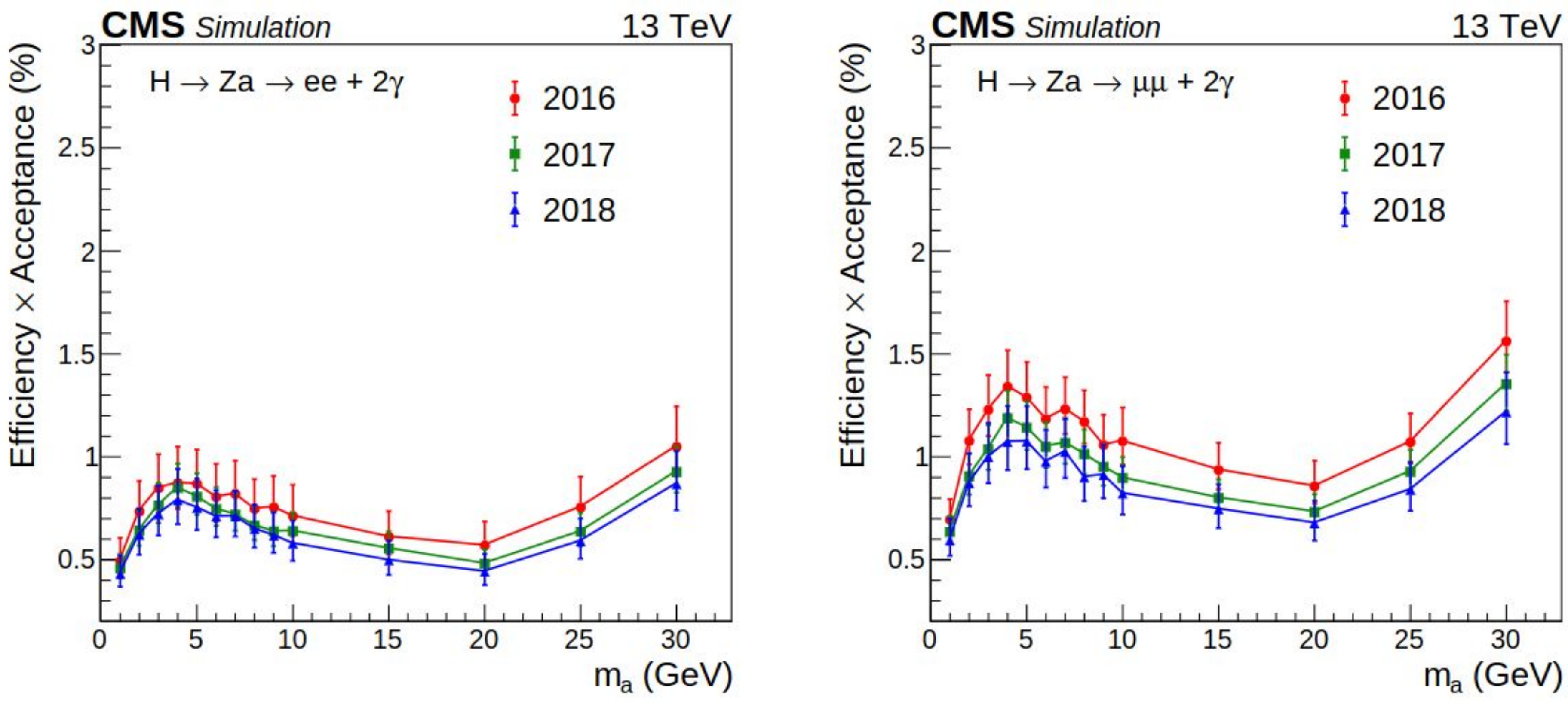


Figure 5: Product of detector efficiency and analysis acceptance for signal samples with various m_a values for the electron (left) and muon channel (right). The error bars include statistical and systematic uncertainties. The photon efficiency, lepton efficiency, and pileup modeling uncertainties are taken into account for the systematic uncertainty.

To build the signal models for the intermediate mass hypotheses in the range 10 < ma < 30 GeV, two factors must be considered: the shape of the $m\ell\ell\gamma\gamma$ distribution and its normalization. Since the shape of the $m\ell\ell\gamma\gamma$ distribution does not significantly depend on ma in the interpolation range, only the normalization of the signal model is parameterize.

For each intermediate point, a signal model is constructed using the $m\ell\ell\gamma\gamma$ shape of the nearest mass hypothesis and the normalization is interpolated from the two nearest mass hypotheses

H → Za → 2 lep + 2 Gamma: fit

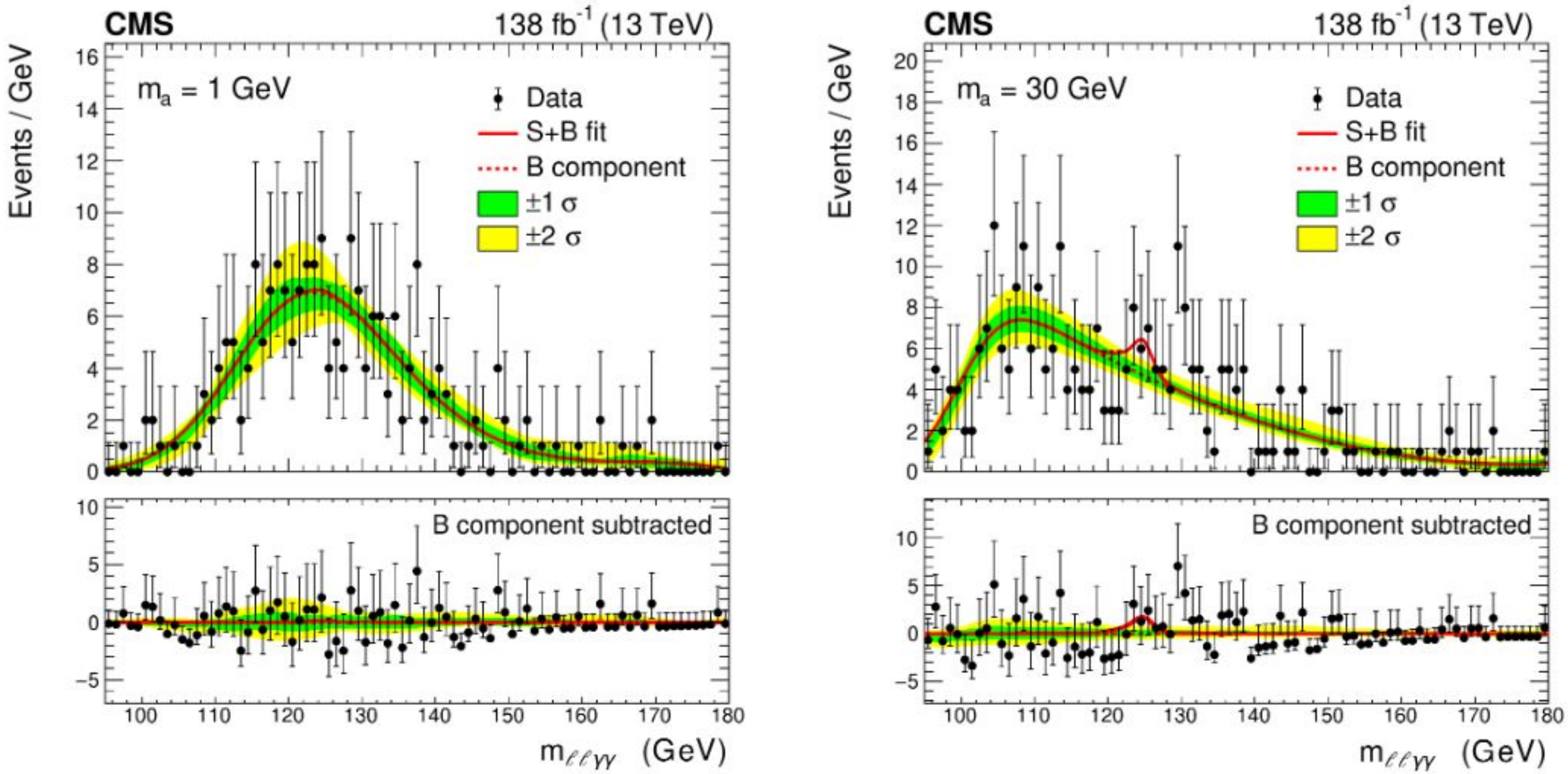


Figure 6: Invariant mass $m_{\ell\ell\gamma\gamma}$ distribution in data (black points). The signal-plus-background model fit is shown for $m_a = 1$ (left) and 30 (right) GeV, where the solid red line shows the total signal-plus-background contribution, and the dashed red line shows the background component only. The lower panels show the residual signal yield after the background subtraction. The one (green, inner) and two (yellow, outer) standard deviation bands show the uncertainties in the fitted background model. These bands include the uncertainty due to the choice of function and the uncertainty in the fitted parameters.

Table 2: Sources of systematic uncertainties and their impact on the signal strength for each data-taking period.

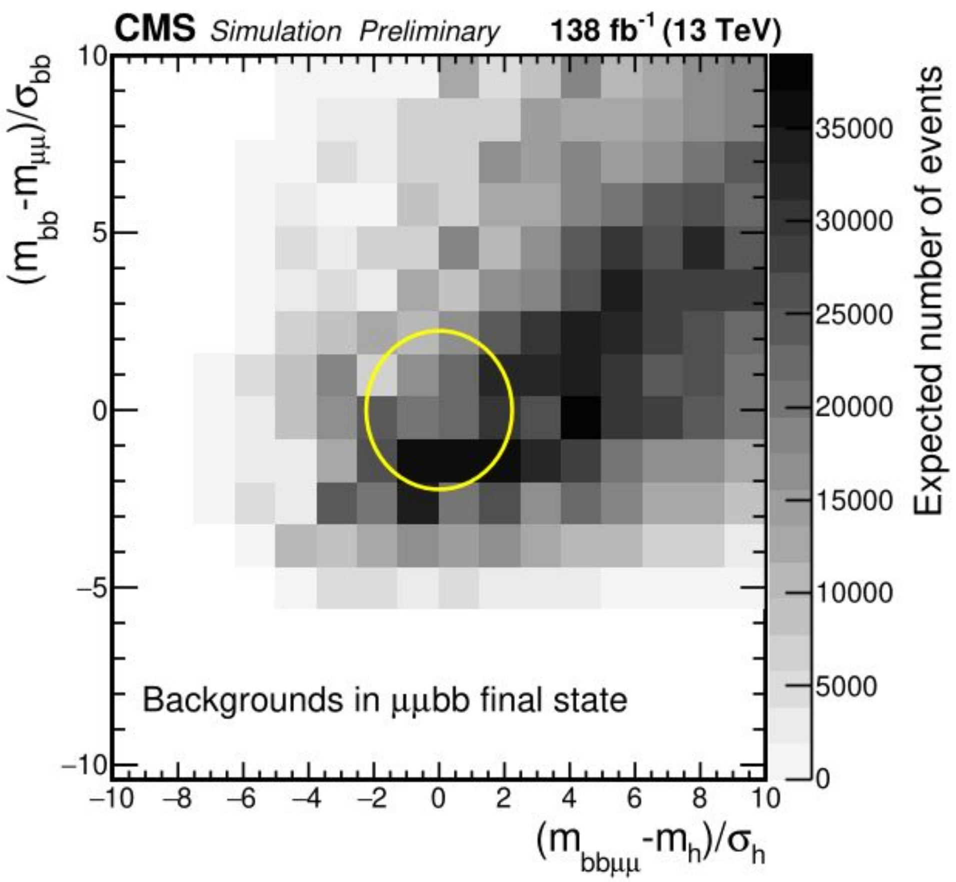
$m_{\ell\ell\gamma\gamma}$ distribution shape	2016	2017	2018
Photon energy scale	<0.10%	<0.10%	<0.10%
Photon energy resolution	5.7%	3.5%	4.5%
Electron energy scale	<0.10%	<0.10%	<0.10%
Electron energy resolution	4.3%	4.2%	4.9%
Muon energy scale	<0.10%	<0.10%	<0.10%
Muon energy resolution	4.9%	4.4%	5.2%
Signal model normalization			
Integrated luminosity	1.2%	2.3%	2.5%
Pileup modeling	2.9%	2.9%	2.5%
Photon efficiency	10%	10%	10%
Electron efficiency	1.7%	1.5%	1.6%
Muon efficiency	0.80%	0.50%	0.50%
BDT uncertainties	<2%	<2%	<2%

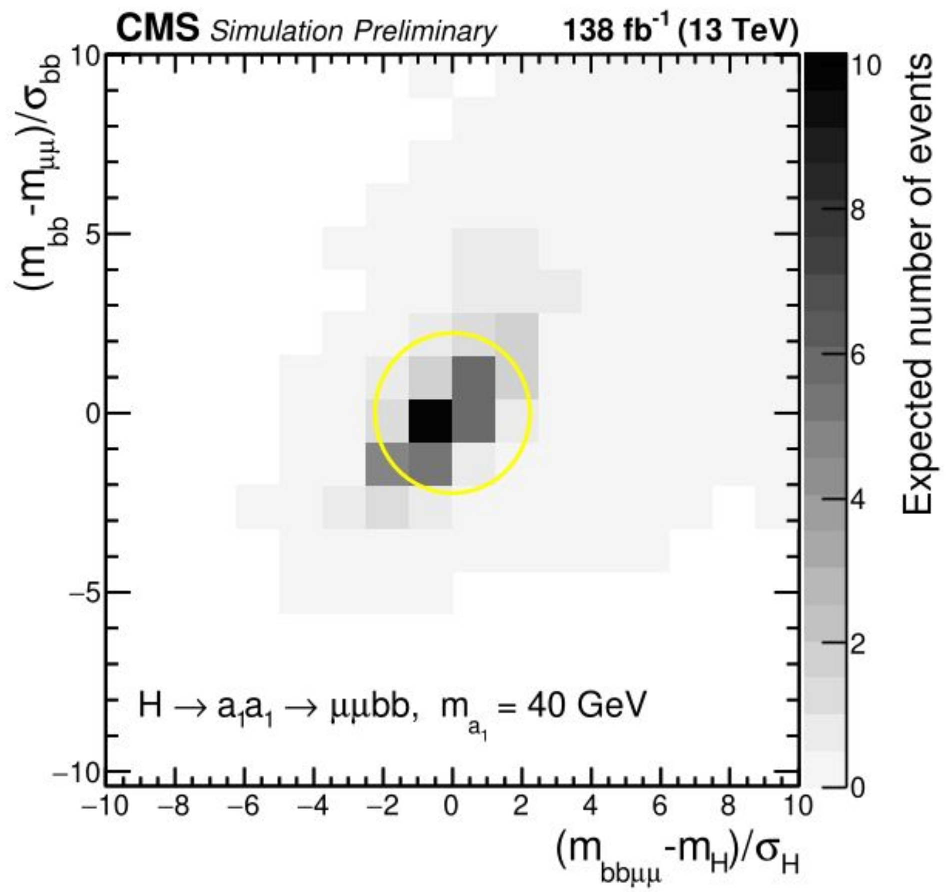
Event selection and category:

- For $2b2\mu$:
  - Events are selected with at least two b-jets
  - $M_{\mu\mu}$ ~ [14, 70] GeV
  - MET < 60 GeV, to suppress tt~ contribution

$$\chi^2_{\mathrm{tot}} = \chi^2_{\mathrm{bb}} + \chi^2_{\mathrm{H}}$$

$$\chi_{\mathrm{bb}} = \frac{(m_{\mathrm{bb}} - m_{\mu\mu})}{\sigma_{\mathrm{bb}}}, \quad \text{and} \quad \chi_{\mathrm{H}} = \frac{(m_{\mu\mu\mathrm{bb}} - 125)}{\sigma_{\mathrm{H}}}$$





$$\begin{pmatrix} \chi_{\mathrm{H}} \\ \chi_{\mathrm{bb}} \end{pmatrix}_{\mathrm{d}} = \begin{pmatrix} \frac{a}{\sqrt{\lambda_1}} & \frac{b}{\sqrt{\lambda_1}} \\ \frac{-b}{\sqrt{\lambda_2}} & \frac{a}{\sqrt{\lambda_2}} \end{pmatrix} \begin{pmatrix} \chi_{\mathrm{H}} \\ \chi_{\mathrm{bb}} \end{pmatrix}_{c},$$

$$\chi^2_{\mathrm{d}} \equiv \chi^2_{\mathrm{H,d}} + \chi^2_{\mathrm{bb,d}}$$

PCA is used to decorrelation, events with χ^2_d<1.5 are selected

Additional event categorization

Low p_T	at least one b-jet with $p_T < 20$ GeV
VBF	two add. jets with $p_T > 30$ GeV, $\lvert\eta\rvert < 4.7$, and $m_{jj} > 250$ GeV

TL	looser b jet passes L but fails M
TM	looser b jet passes M but fails T
TT	looser b jet passes T

	Type-1	Type-2	Type-3 (lepton-specific)	Type-4 (flipped)
right-handed leptons	Φ_1	Φ_2	Φ_2	Φ_1
up-type quarks	Φ_1	Φ_1	Φ_1	Φ_1
down-type quarks	Φ_1	Φ_2	Φ_1	Φ_2

Trigger requirements

p_T thresholds	$\tau\tau$bb			$\mu\mu$bb
	eμ	eτ_h	$\mu\tau_h$	
Electron	23 (leading) 12 (subleading)	25 (2016) 32 and 35 (2017, 2018)	–	–
Muon	23 (leading) 8 (subleading)	–	22 (2016) 24 and 27 (2017, 2018)	17 (leading) 8 (subleading)
τ_h	–	20 (2016) 27 (2017, 2018)		–

Events are further then categorized according to the BDT score

- 1 b-jet category: three SRs
- At least 2 b-jets category: mostly two SRs, but also only one SR due to add more SR won't increase significance

	Exactly one b jet				At least two b jets		
	SR1	SR2	SR3	CR	SR1	SR2	CR
$e\mu$ 2018	> 0.99	$\in [0.95, 0.99]$	$\in [0.85, 0.95]$	< 0.85	> 0.98	$\in [0.94, 0.98]$	< 0.94
$e\mu$ 2017	> 0.985	$\in [0.95, 0.985]$	$\in [0.85, 0.95]$	< 0.85	> 0.97	$\in [0.93, 0.97]$	< 0.93
$e\mu$ 2016	> 0.99	$\in [0.95, 0.99]$	$\in [0.85, 0.95]$	< 0.85	> 0.98	$\in [0.94, 0.98]$	< 0.94

	Exactly one b jet				At least two b jets		
	SR1	SR2	SR3	CR	SR1	SR2	CR
$e\tau_h$ 2018	> 0.97	$\in [0.945, 0.97]$	$\in [0.90, 0.945]$	< 0.90	> 0.96	NA	< 0.96
$e\tau_h$ 2017	> 0.985	$\in [0.965, 0.985]$	$\in [0.93, 0.965]$	< 0.93	> 0.985	NA	< 0.985
$e\tau_h$ 2016	> 0.985	$\in [0.965, 0.985]$	$\in [0.93, 0.965]$	< 0.93	> 0.96	NA	< 0.96

	Exactly one b jet				At least two b jets		
	SR1	SR2	SR3	CR	SR1	SR2	CR
$\mu\tau_h$ 2018	> 0.98	$\in [0.95, 0.98]$	$\in [0.90, 0.95]$	< 0.90	> 0.99	$\in [0.96, 0.99]$	< 0.96
$\mu\tau_h$ 2017	> 0.97	$\in [0.94, 0.97]$	$\in [0.90, 0.94]$	< 0.90	> 0.98	$\in [0.94, 0.98]$	< 0.94
$\mu\tau_h$ 2016	> 0.97	$\in [0.94, 0.97]$	$\in [0.89, 0.94]$	< 0.89	> 0.97	$\in [0.93, 0.97]$	< 0.93

Variables σ_{bb} and σ_H are the mass resolutions of the di-b-jet system and the Higgs boson candidate, respectively. Derived from simulation, σ_H is found to be constant while σ_{bb} increases linearly with m_{a_1}. The latter is modeled as a function of $m_{\mu\mu}$ ($\sigma_{bb} = a \times m_{\mu\mu} + b$), assuming $m_{\mu\mu} = m_{a_1}$. The χ^2_{tot} variable is evaluated on an event-by-event basis. It was shown in the pre-

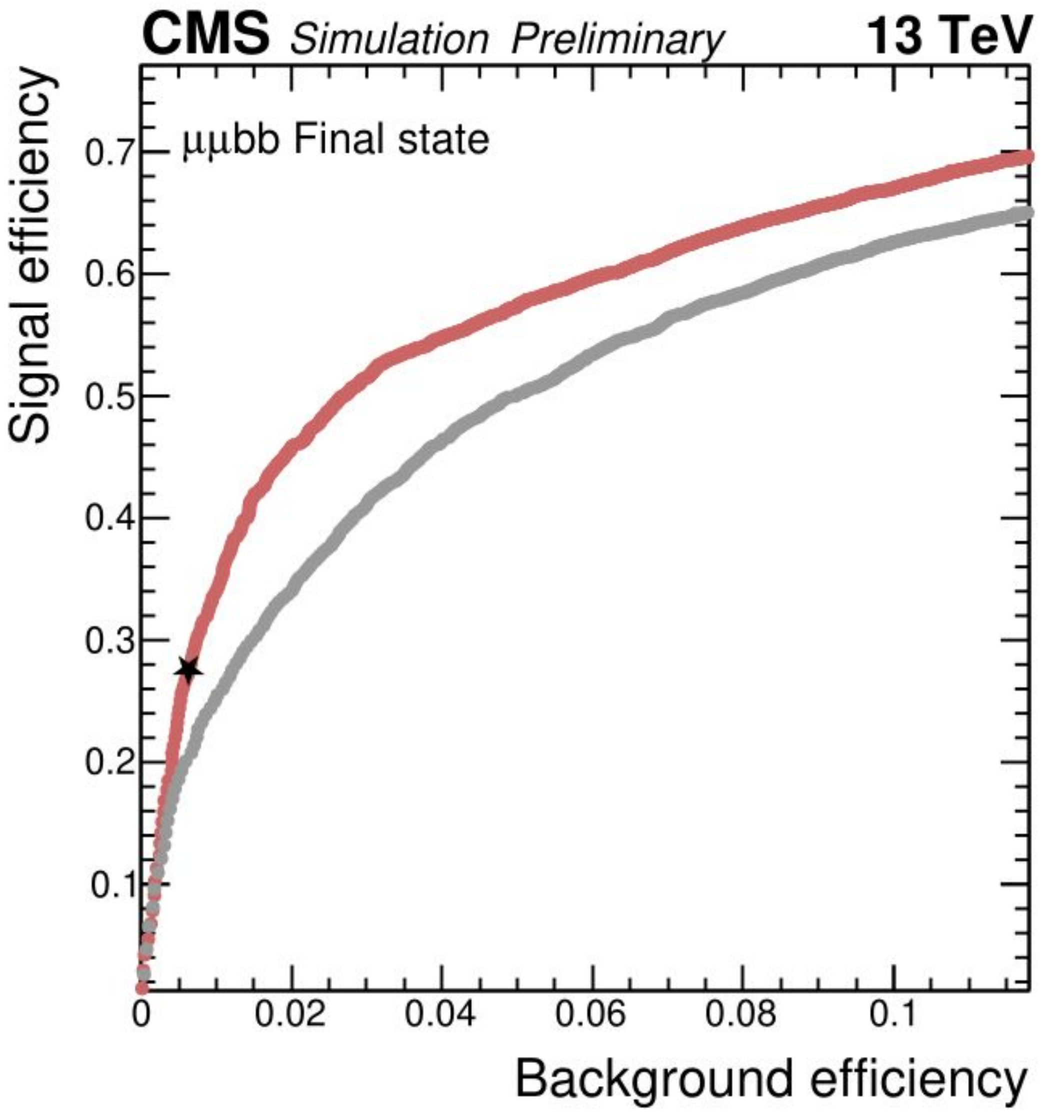


Figure 4: Signal ($m_{a_1} = 40$ GeV) versus background efficiency for different thresholds on χ^2_{tot} (gray) and χ^2_d (red) variables. The black star indicates signal efficiency versus that of background for the optimized χ^2_d requirement.

H → aa → 2mu + 2b: yields

Process	Yield
Top	86.3 ± 2.25
Drell-Yan ($10 < m_{\ell\ell} < 50$)	289.6 ± 89.5
Drell-Yan ($m_{\ell\ell} > 50$)	200.2 ± 31.9
Diboson	1.5 ± 0.9
Single Top	11.4 ± 1.6
Total expected background	589.05 ± 95.09
Data	641

Signal for ggH ($\mu\mu$bb)		
$m_{a_1} = 20\,\text{GeV}$	$m_{a_1} = 40\,\text{GeV}$	$m_{a_1} = 60\,\text{GeV}$
15.4 ± 0.2	18.7 ± 0.2	40.5 ± 0.3

Table 4: The expected yields for backgrounds and different signal hypotheses in each category. The entries are rounded to first decimal place.

Category	Signal for ggH($\mu\mu$bb)			Expected background
	$m_{a_1} = 20\,\text{GeV}$	$m_{a_1} = 40\,\text{GeV}$	$m_{a_1} = 60\,\text{GeV}$	
TL	2.1 ± 0.06	2.8 ± 0.07	6.7 ± 0.11	109 ± 30
TM	2.7 ± 0.06	3.3 ± 0.07	7.7 ± 0.11	27 ± 15
TT	2.8 ± 0.06	4.2 ± 0.08	8.1 ± 0.1	28 ± 11
VBF	0.2 ± 0.02	1.0 ± 0.04	1.1 ± 0.4	5 ± 2
Low p_T	7.4 ± 0.11	7.3 ± 0.11	17 ± 0.17	421 ± 88
Total	15.4 ± 0.2	18.7 ± 0.2	40.5 ± 0.3	589 ± 95

$$m_{\mathrm{T}}(\ell, p_{\mathrm{T}}^{\mathrm{miss}}) \equiv \sqrt{2p_{\mathrm{T}}^{\ell} \cdot p_{\mathrm{T}}^{\mathrm{miss}}\left[1-\cos(\Delta\phi)\right]}, \tag{4}$$

where p_{T}^{ℓ} is the transverse momentum of the lepton and $\Delta\phi$ is the azimuthal angle between the lepton and the $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$. Events from tt and misidentified τ_{h} backgrounds, such as $\mathrm{W}+\mathrm{jets}$, have larger $p_{\mathrm{T}}^{\mathrm{miss}}$, thus result in higher m_T values. Another variable useful in the training is D_{ζ}, defined as

$$D_{\zeta} \equiv p_{\zeta} - 0.85 p_{\zeta}^{\mathrm{vis}} \tag{5}$$

where the bisector of the directions of the visible τ decay products transverse to the beam direction is denoted as the ζ axis. The quantity p_{ζ} is defined as the component of the $p_{\mathrm{T}}^{\mathrm{miss}}$ along the ζ axis, and p_{ζ}^{vis} to be the sum of the components of the lepton transverse momentum along the same direction [55]. The $\mathrm{Z} \to \tau\tau$ background falls in large D_{ζ} values because the $p_{\mathrm{T}}^{\mathrm{miss}}$ is approximately collinear to the $\tau\tau$ system. The tt events tend to have small D_{ζ} values due to a large $p_{\mathrm{T}}^{\mathrm{miss}}$ that is not aligned with the $\tau\tau$ system. The signal has intermediate D_{ζ} values because the $p_{\mathrm{T}}^{\mathrm{miss}}$ is approximately aligned with the $\tau\tau$ system, but its magnitude is small. For events in two b jets category, a variable can be constructed to measure the difference between the invariant mass of the two b jets and the invariant mass of two taus:

$$\Delta m_{\mathrm{a}_1} \equiv (m_{\mathrm{bb}} - m_{\tau\tau})/m_{\tau\tau}. \tag{6}$$

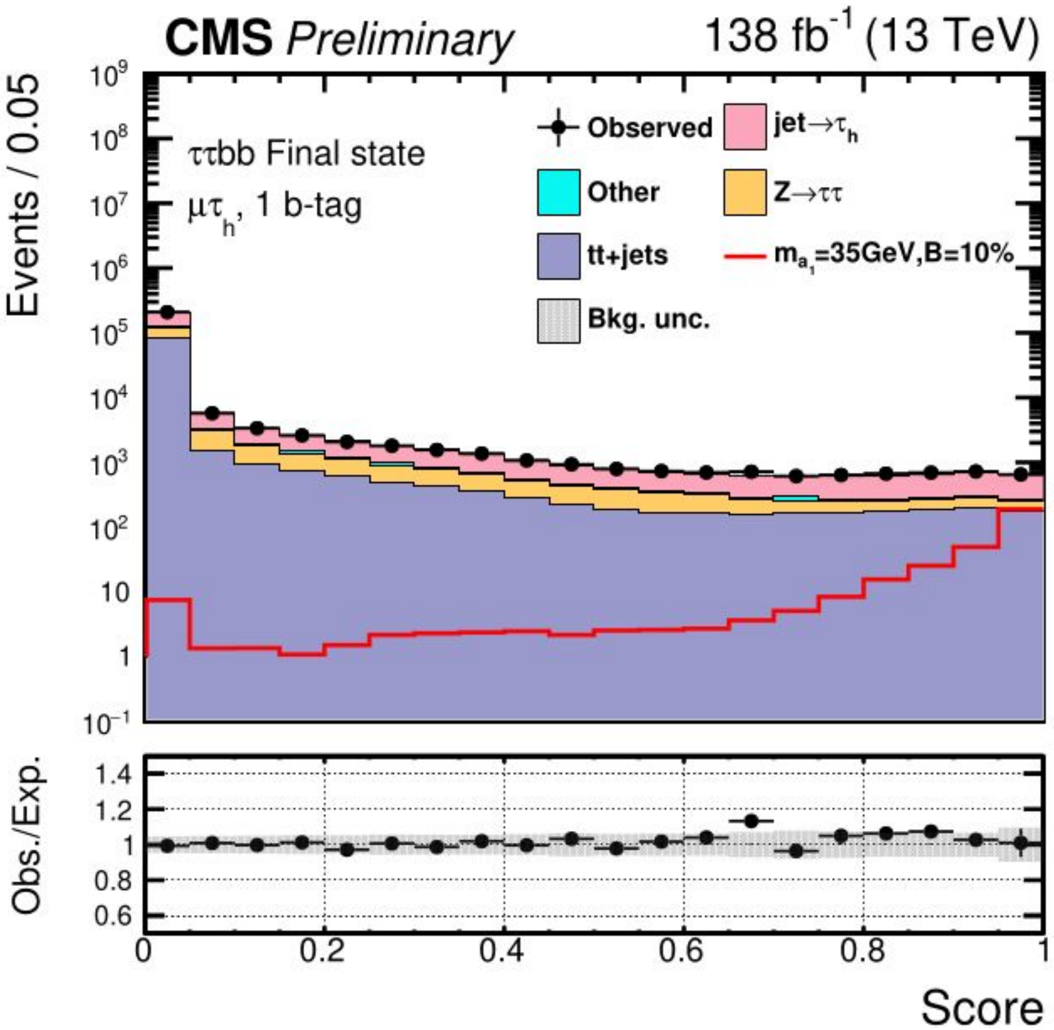
This variable is of particular interest since it peaks at the value 0 for signal events. The di-tau invariant mass ($m_{\tau\tau}$) distribution, reconstructed using the SVfit algorithm [56], is used as the observable for the likelihood fit, and thus is not included as an input to the DNN.

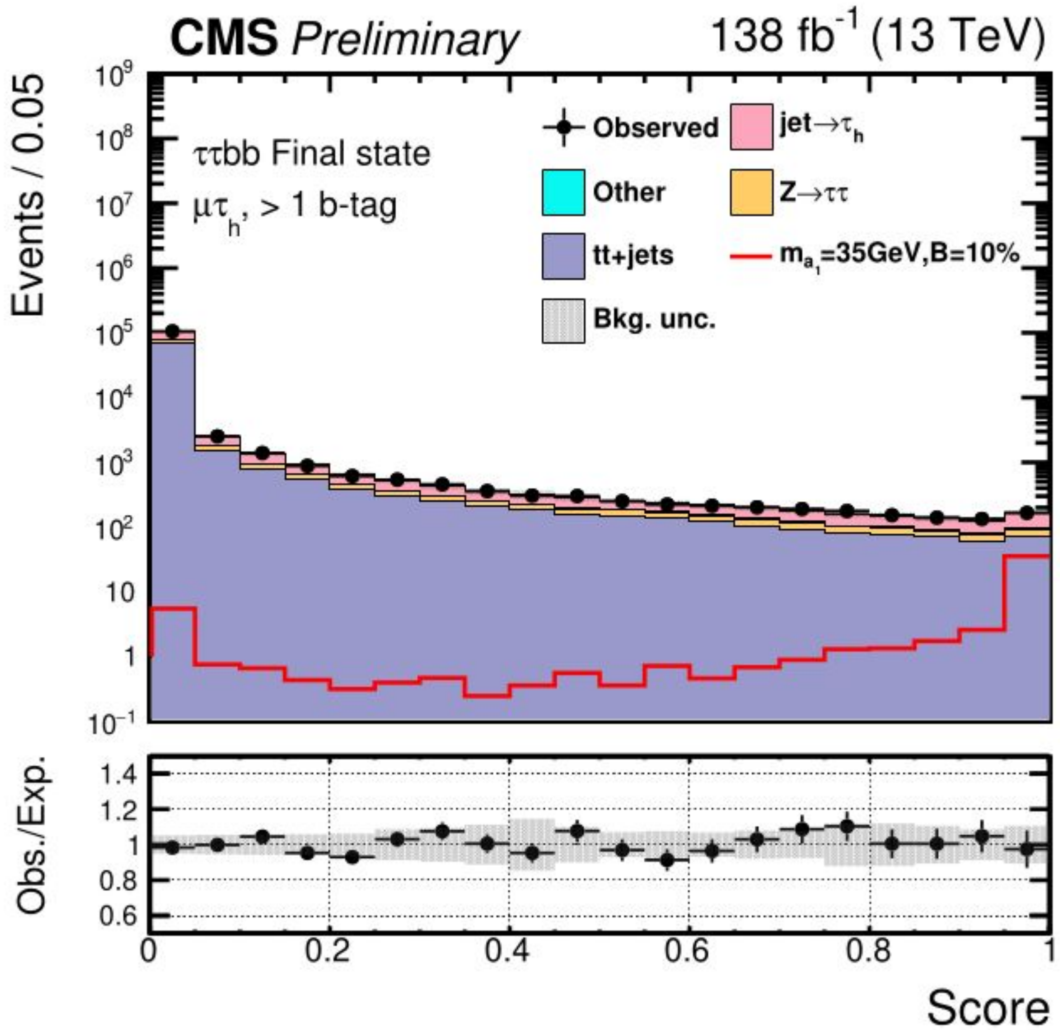
Event selection and category

- For $2b2\tau$:
  - Considering $e\mu$, $e\tau_h$ and $\mu\tau_h$ in the final state.
  - For each of them, events are categorized to exactly 1 b-jet and >1 b-jet

BDT is used to separate signal and bkg.

- Lepton and jet pt, pt and eta of di-τ system, invariant mass and dR of objects





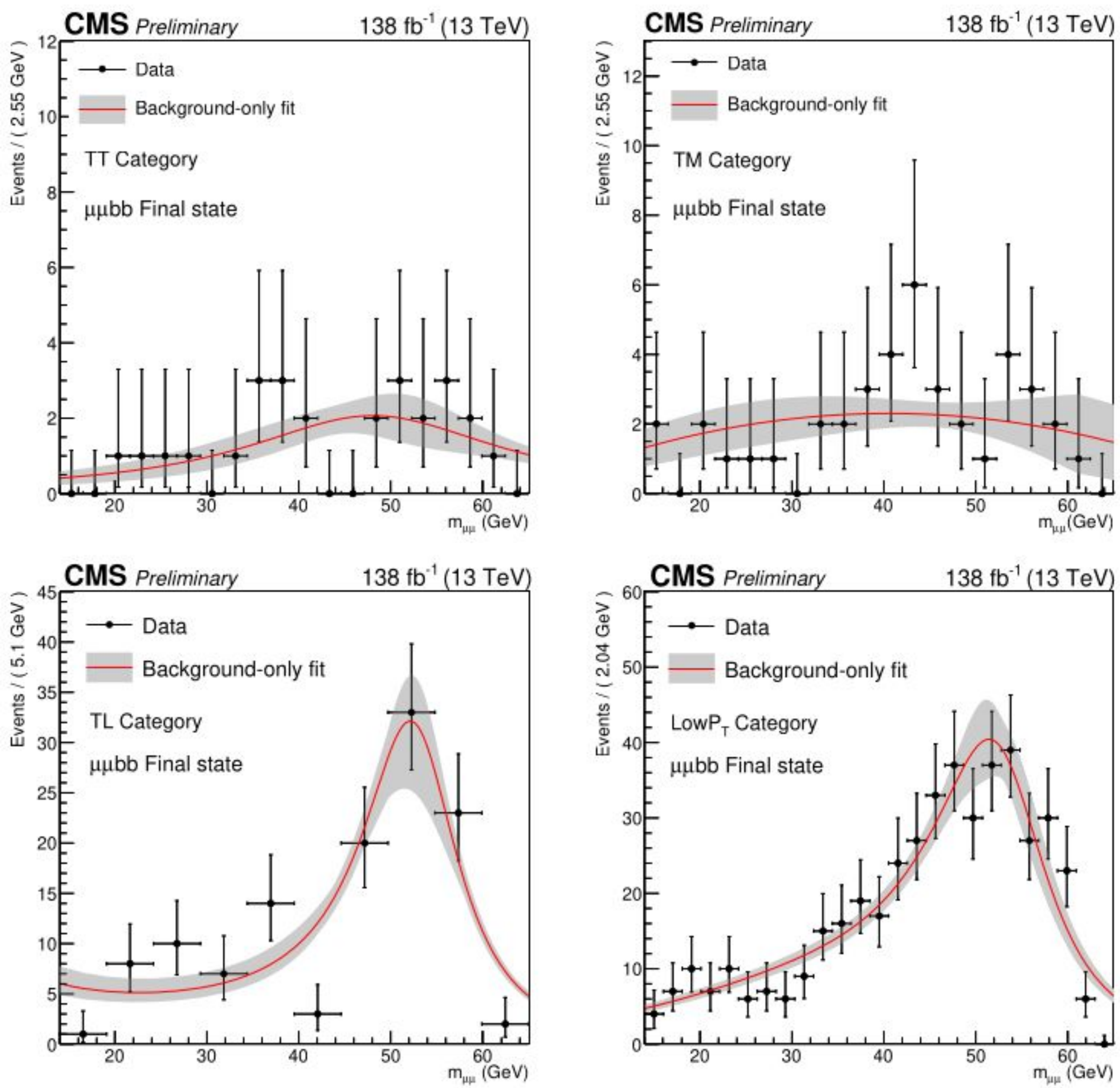
Important BDT inputs:

$$m_T(\ell, p_T^{miss}) \equiv \sqrt{2p_T^{\ell} \cdot p_T^{miss}\,[1-\cos(\Delta\phi)]},$$

$$D_\zeta \equiv p_\zeta - 0.85 p_\zeta^{vis}$$

$$\Delta m_{a_1} \equiv (m_{bb} - m_{\tau\tau})/m_{\tau\tau}$$

2mu+2b: bkg model



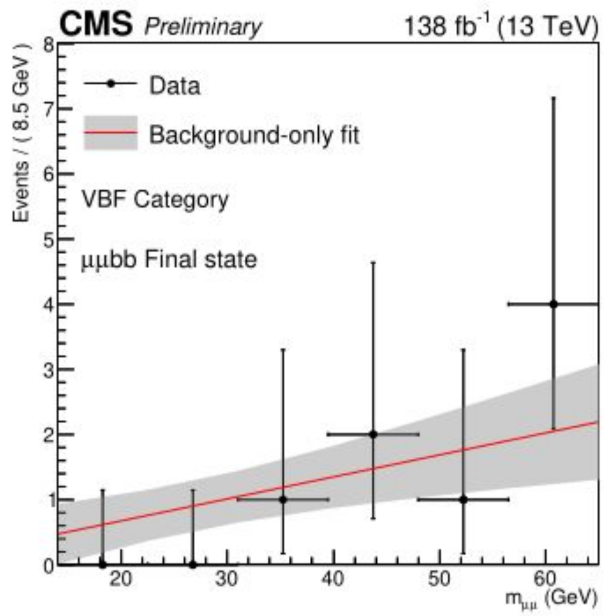


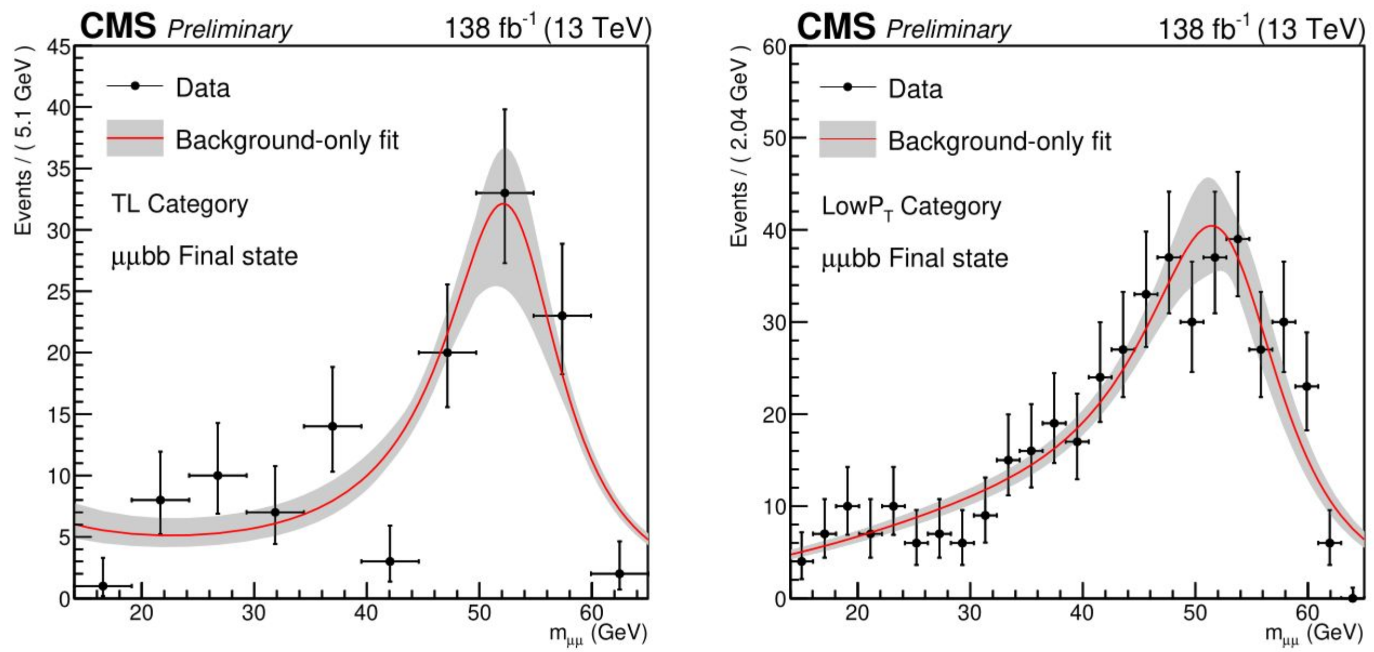
Figure 6: The best-fit background models together with 68% CL uncertainty band from the fit to the data under the background-only hypothesis for the (top left) TT category, (top right) TM, (middle left) TL category, (middle right) Low p_T category, and (bottom) VBF category.

H → aa → 2tau + 2b: bkg estimation

For 2b2τ, dominant bkgs are:

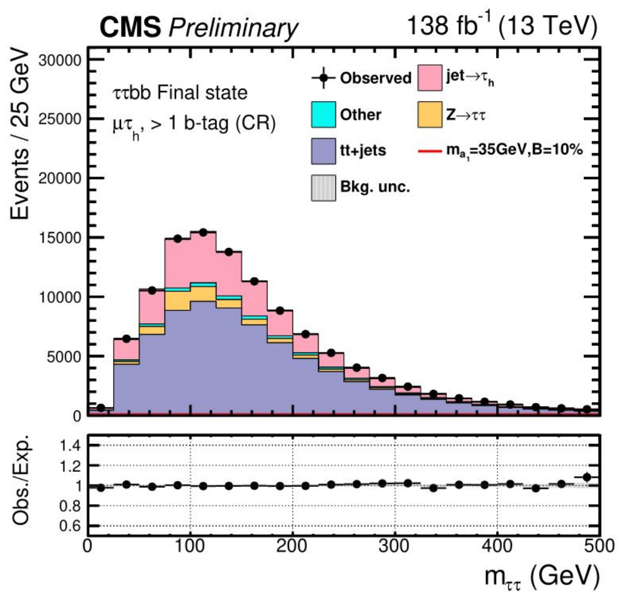
- Z→ 2τ: estimated from data using embedding technique. Replace muons with tau in the Z→$\mu\mu$ events in the data while keeping kinematics the same, in which tau decay is modelled using simulation
- QCD jet mis-identified as e/mu (eμ channel): evaluate the ratio of OS/SS lepton in SR using simulation, scale the SS data events with the ratio
- Jet mis-identified as hadronic τ: calculate the probability (f) of a jet to be identified as the τ_h, apply f/(1-f) on the sideband event
- Remove double counting in ttbar events, e.g., remove eμ events if no matched eμ at generator level

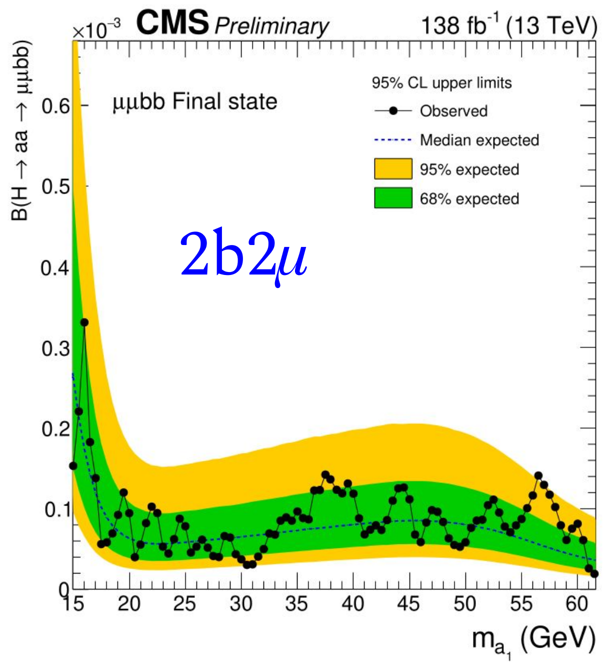
For 2b2μ, bkg is collectively estimated from data. Use polynomials+inverse polynomial to model the bkg for each event category.

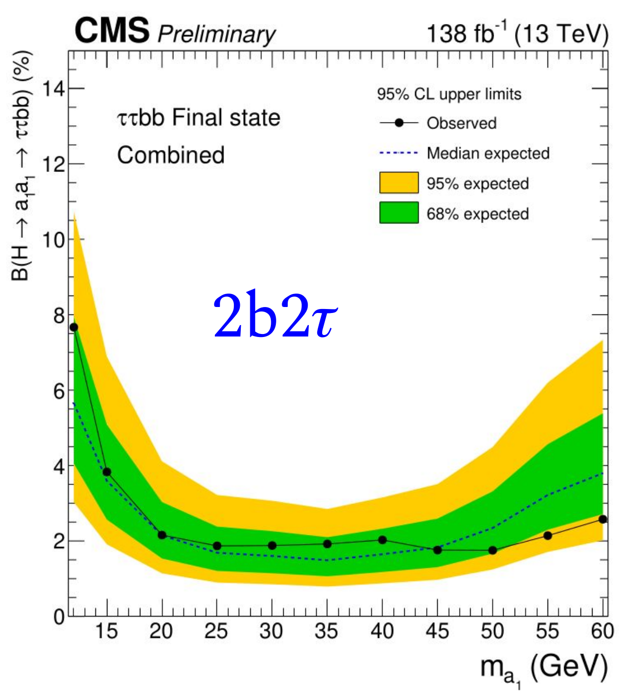


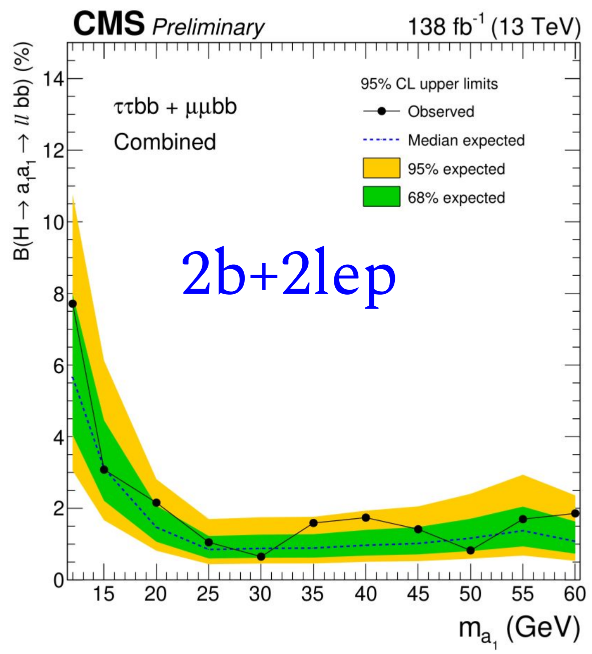
H → aa → 2mu/tau + 2b: fit and results

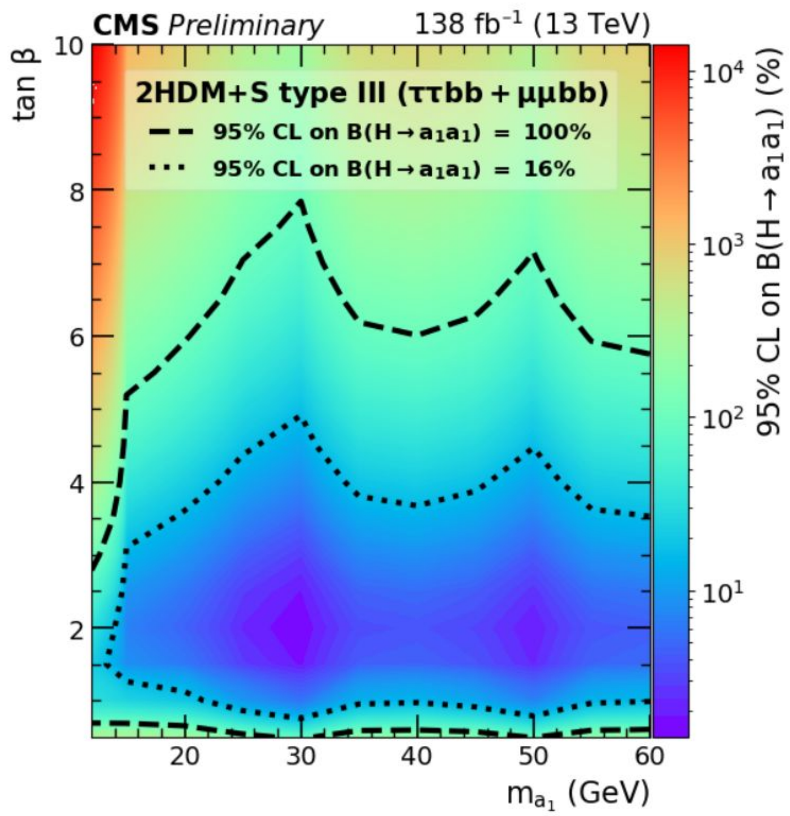
- For $2b2\mu$: unbinned fit in SR, signal shape obtained from simulation using weighted sum of a Voigt profile and a Crystal Ball functions
- For $2b2\tau$:
  - Binned fit using invariant mass of two τ, $m_{\tau\tau}$, obtained using SVFit.

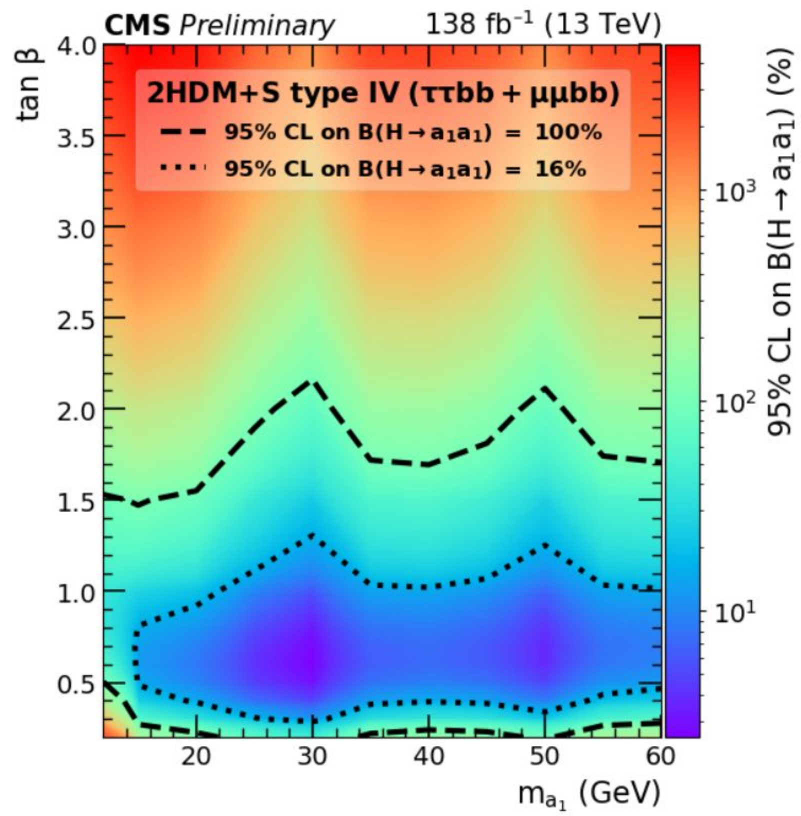




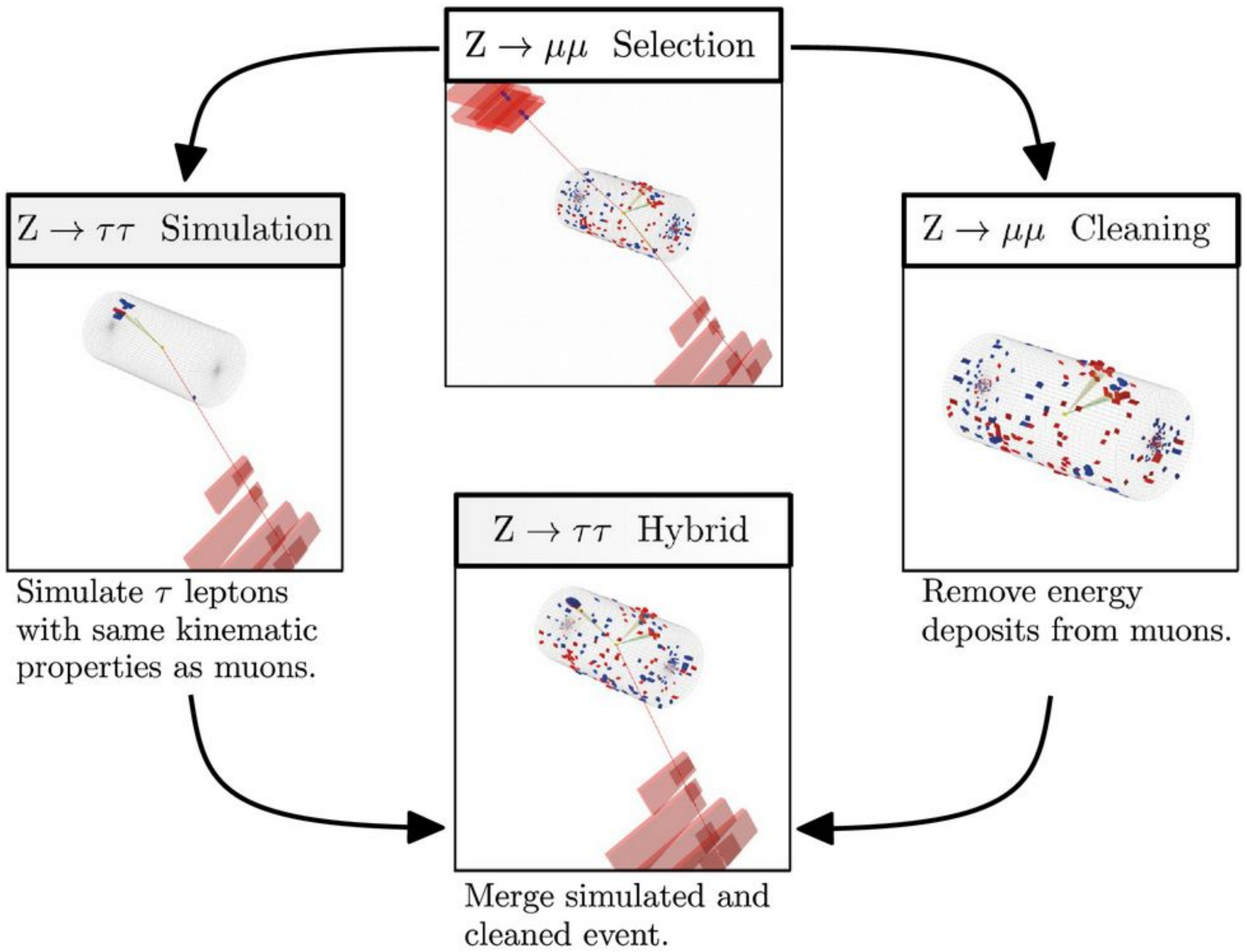




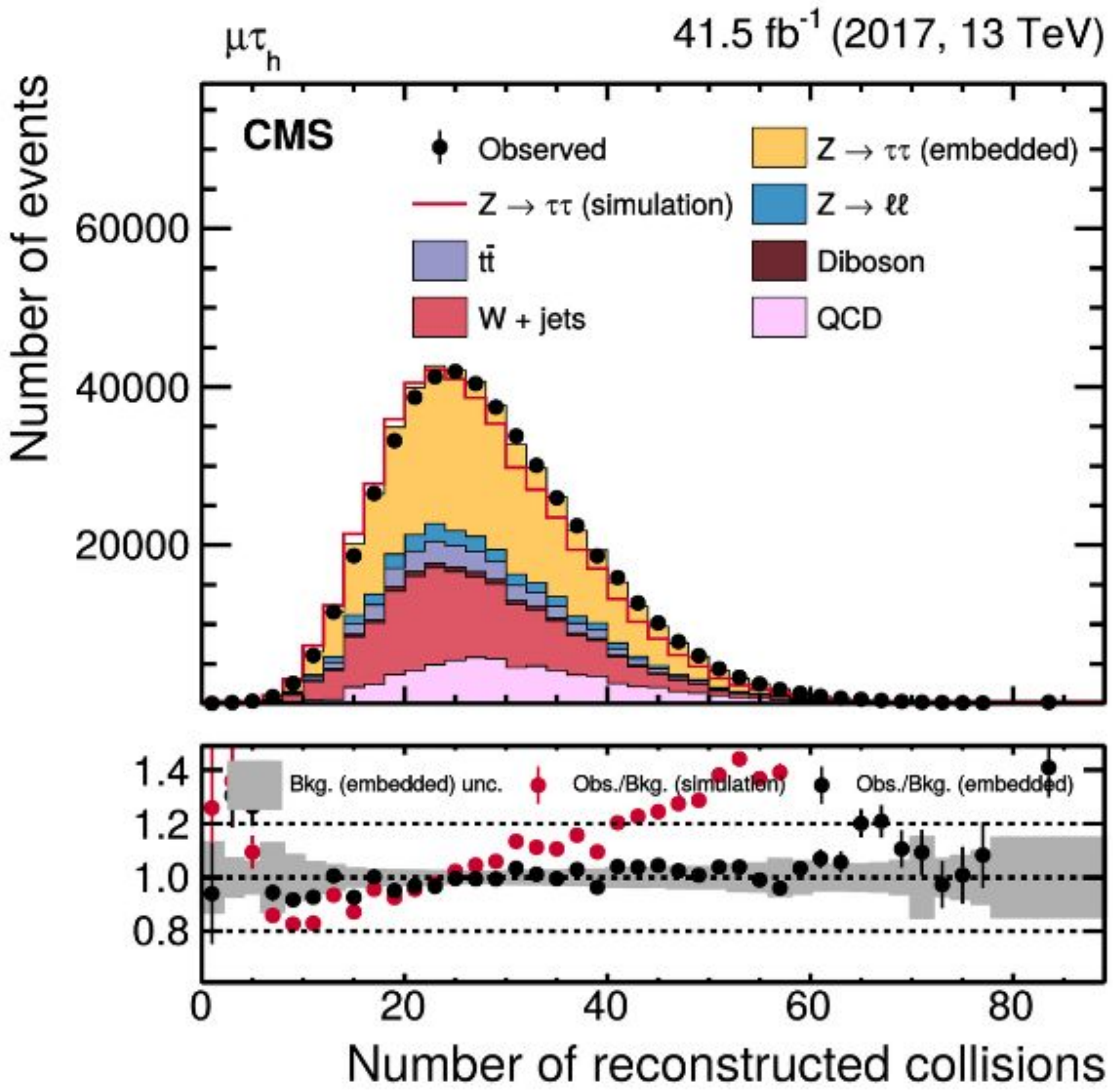




- The brasil plots in the left show the 95% CL limit independent of the model type and tanβ
- The 2D limits plot in the right show the exclusion phase space in Type-III and Type-IV w.r.t ma1 and tanβ



Z → μμ Selection
Z → ττ Simulation
Z → μμ Cleaning
Z → ττ Hybrid
Simulate τ leptons with same kinematic properties as muons.
Remove energy deposits from muons.
Merge simulated and cleaned event.



μτh
41.5 fb-1 (2017, 13 TeV)
CMS
Observed
Z → ττ (simulation)
tt̄
W + jets
Z → ττ (embedded)
Z → ℓℓ
Diboson
QCD
Number of events
60000
40000
20000
Bkg. (embedded) unc.
Obs./Bkg. (simulation)
Obs./Bkg. (embedded)
1.4
1.2
1.0
0.8
0
20
40
60
80
Number of reconstructed collisions

H → aa → 4b: BDT

- p_{T}^{H}: the vector sum of the p_T of the three or four b-tagged jets forming the H boson candidate;
- m_H: the invariant mass of the three or four b-tagged jets forming the H boson candidate;
- p_{T}^{V}: the vector sum of the p_T of the electron or muon and p_{T}^{miss} for the W candidate; the p_T of the dilepton pair for the Z boson candidate;

- H_T: the sum of the scalar p_T of the transverse momenta of the three or four b-tagged jets that define the H boson candidate;
- $\langle \Delta R(b, b') \rangle$: the separation in the η-ϕ plane between any two b-tagged jets in an event, averaged over all such combinations;
- $|\Delta\phi(V, H)|$: the azimuthal angle between the directions of the W or Z boson candidate and the H boson candidate;
- $|\Delta\phi(j, p_{T}^{miss})|^{min}$: the smallest azimuthal angle between $\vec{p}_{T}^{miss}$ and a jet;
- p_{T}^{ℓ}: for WH events, the p_T of the electron or muon; for ZH events, the p_T of the leading electron or muon;
- p_{T}^{miss}: the magnitude of the missing transverse momentum;
- m_T: the transverse mass defined in section 4;
- $\Delta m_{b\bar{b}}^{min} = |m_{b\bar{b},1} - m_{b\bar{b},2}|^{min}$: the minimum difference between two dijet masses formed from all possible combinations with the three or four b-tagged jets. The four (three) b-tagged jets can be grouped in three (two) different ways to form a pair of a $\rightarrow b\bar{b}$ candidates. The $\Delta m_{b\bar{b}}^{min}$ variable represents the grouping that results in the smallest difference between the masses of the two pseudoscalar candidates.

H → aa → 4b: unc

Table 2: Summary of systematic uncertainties for background and signal event yields in the WH channel. Uncertainties that are negligible are indicated with a dash (—).

Source	Uncertainty (%)						
	$t\bar{t}+b\bar{b}$	$t\bar{t}+c\bar{c}$	$t\bar{t}$+LF	W+jets	Z+jets	WH (20 GeV)	WH (60 GeV)
Lepton trigger, identification & isolation	2	2	2	2	2	2	2
Lepton energy scale	—	—	—	—	—	—	—
Jet energy scale	0.5	0.5	0.5	2	2	1	1
Jet energy resolution	0.5	0.5	0.5	4	5	1	0.5
Energy scale of unclustered p_T^{miss}	0.5	0.5	0.5	2	2	1	0.8
b quark tagging	6.0	4.0	4.0	4.5	5.0	9	8.5
c quark mistagging	1.5	6.5	—	4.0	2.5	0.3	0.1
Light-flavor mistagging	1.5	1.0	9.0	7.0	9.5	1	0.5
Limited MC statistical precision	1.0	0.5	0.5	4.5	4.5	1	1
Top p_T mismodeling	2.5	1.3	1.5	—	—	—	—
Pileup	0.2	0.8	2.5	4	4	0.5	0.5

Table 3: Summary of systematic uncertainties for background and signal event yields in the ZH channel. Uncertainties that are negligible are indicated with a dash (—).

Source	Uncertainty (%)					
	$t\bar{t}+b\bar{b}$	$t\bar{t}+c\bar{c}$	$t\bar{t}$+LF	Z+jets	ZH (20 GeV)	ZH (60 GeV)
Lepton trigger, identification & isolation	2	2	2	2	2	2
Lepton energy scale	—	—	—	—	—	—
Jet energy scale	1	—	—	0.5	0.1	0.1
Jet energy resolution	0.1	0.1	0.1	0.1	0.1	0.1
b quark tagging	6.5	4	4	5	9	8
c quark mistagging	0.5	9	—	2.5	0.3	0.1
Light-flavor mistagging	1.0	2	10	6	1	0.5
Limited MC statistical precision	2.5	3	2	1.5	1	1
Pileup	1	1	4	3	0.5	0.5

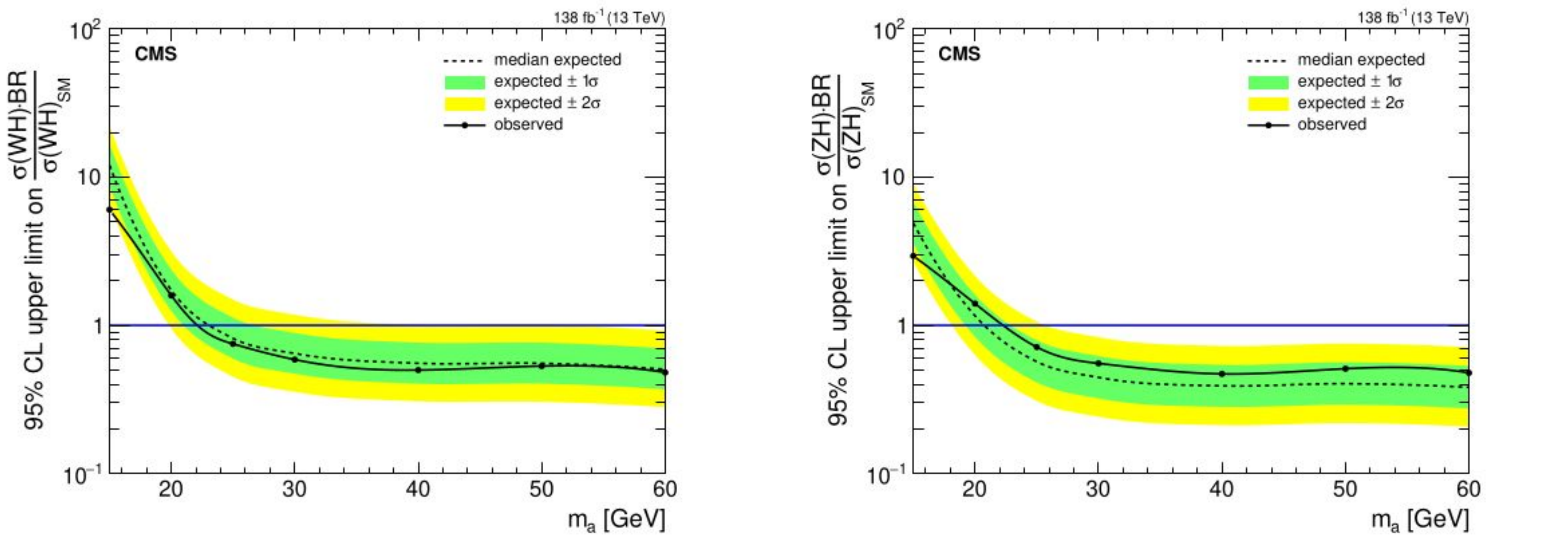


Figure 3: Observed and expected limits on the signal strength $\mu = \sigma(\mathrm{VH})\mathcal{B}(\mathrm{H} \to \mathrm{aa} \to \mathrm{b\bar{b}b\bar{b}})/\sigma(\mathrm{VH})_{\mathrm{SM}}$ in the WH (left) and ZH channel (right). The solid blue line indicates the SM cross section $\sigma(\mathrm{pp} \to \mathrm{VH})$ with branching fractions $\mathcal{B}(\mathrm{H} \to \mathrm{aa}) = 1$ and $\mathcal{B}(\mathrm{a} \to \mathrm{b\bar{b}}) = 1$.