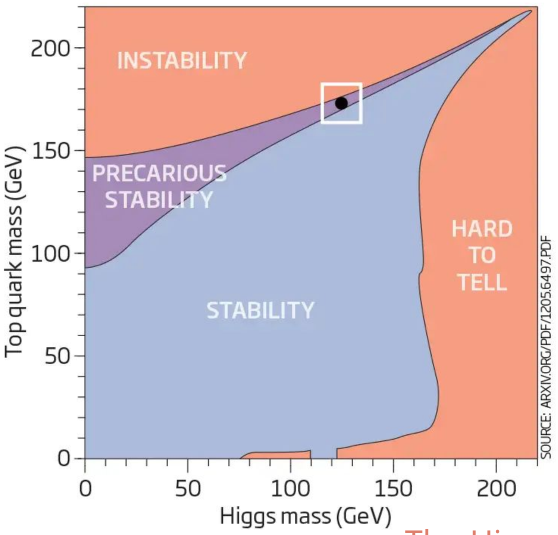
Searches for additional heavy Higgs bosons at CMS

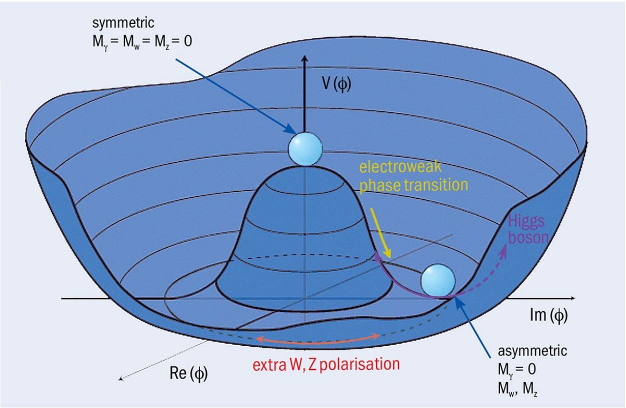
Khawla Jaffel – USTC
on behalf of the CMS Collaboration

Higgs2023: 28 Nov, 2023

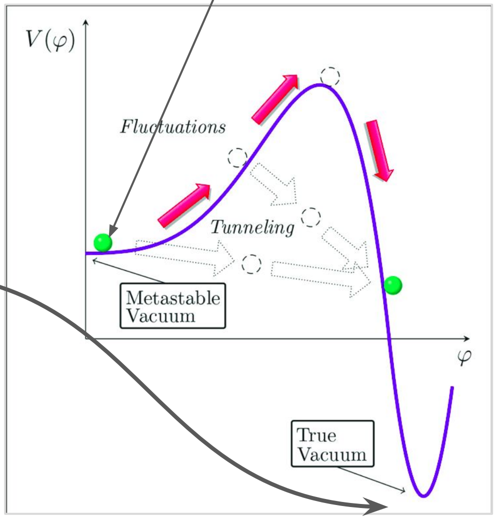
The Higgs bosons connections to new physics

If the top quark were slightly heavier relative to the Higgs boson, the universe would have collapsed long ago





Non zero vacuum
vev =246 GeV
mH = 125 GeV



What would Universe be like in such vacuum?

It's difficult to say for certain.

The Higgs boson can be our gateway to unlock Universe hidden secrets...

BSM Higgs scenarios

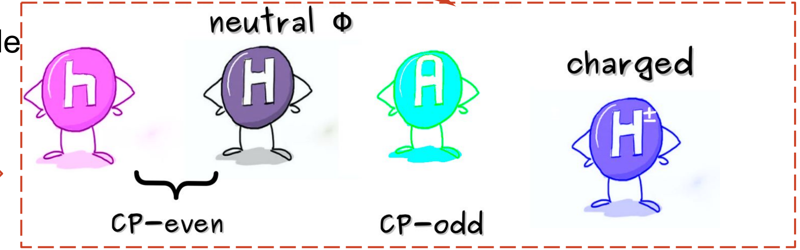
- Well-motivated BSM scenarios designed to address specific issues of the SM. In particular:
    - **Matter – antimatter asymmetry in the Universe** → CP violation.
    - **Flavor problem** by providing mechanisms or symmetries that explain the patterns and hierarchies of masses and mixing angles among different generations of quarks and leptons.
    - **Muon g-2 anomaly** → Offer more parameters to adjust, potentially providing a solution to the anomaly.
    - **Dark matter:** Models with axion-like particles as a potential candidates for dark matter.
        - Extremely light and weakly interacting (outside the scope of this talk).
    - …and more. These scenarios typically induce sizeable modifications of the Higgs couplings.
        - … and involve extensions of the scalar sector, e.g. Two Higgs Doublet Models, N2HDM (2HDM+Singlet, Triplet…), NMSSM, etc.
        - Composite Higgs models.

The Two-Higgs-Doublet Model (2HDM)

5 physical Higgs bosons

- The 2HDM is an effective theory with extra SU(2)L doublet

$$\Phi_i = \begin{pmatrix} \phi_i^+ \\ (v_i + \phi_i^0 + i\varphi_i)/\sqrt{2} \end{pmatrix}$$



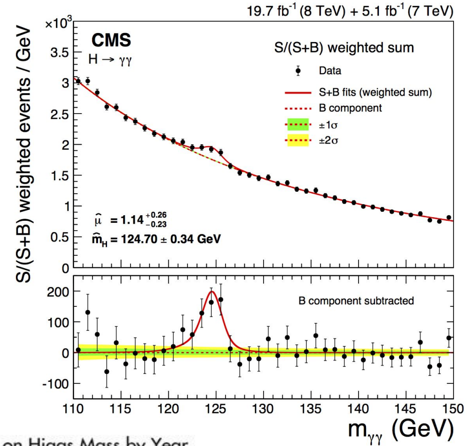
- The free parameters of the model in the **physical Higgs masses basis after EWSB** are:
  - mh, mH , mA and mH± .
  - The mixing angle between the two CP-even Higgses (α).
  - The ratio of the two vacuum expectation values (tanβ = v2/v1).
  - The light CP even h is SM-like for cos(β-α) ~ 0 (alignment limit).
  - Introducing a Z2 symmetry to avoid tree-level **FCNCs** allowed to be softly broken bring an additional parameter m_{12}.

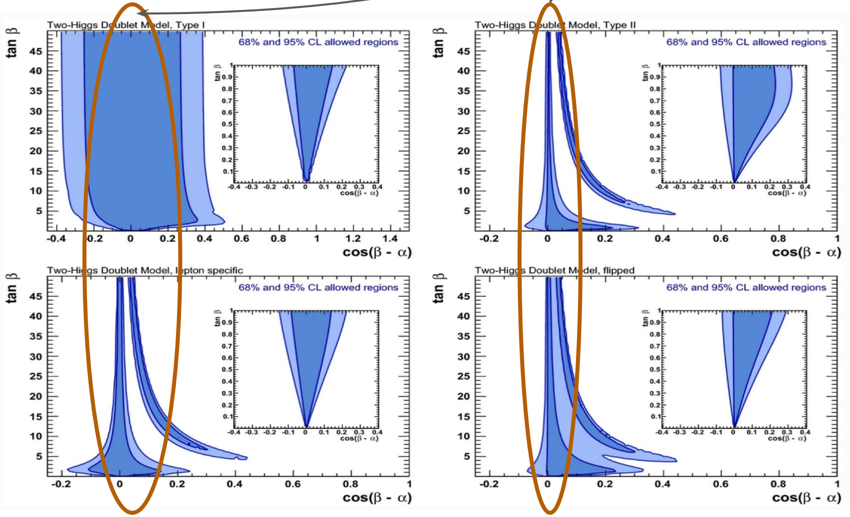
Type	up-type quarks	down-type quarks	leptons
I: Fermiophobic	Φ_2	Φ_2	Φ_2
II: MSSM-like	Φ_2	Φ_1	Φ_1
X: Lepton-specific	Φ_2	Φ_2	Φ_1
Y: Flipped	Φ_2	Φ_1	Φ_2
III: FCNC at tree level	Φ_1,Φ_2	Φ_1,Φ_2	Φ_1,Φ_2
FCNC-free	Φ_1,Φ_2	Φ_1,Φ_2	Φ_1,Φ_2

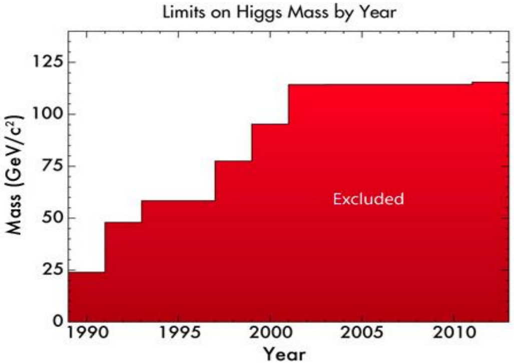
*By convention, Φ2 is the doublet to which up-type quarks couple.

The search for the unknown!

- **In a comparable scenario regarding the Higgs bosons:**
  - Theoretical considerations left a broad range, spanning from 10 GeV to 1 TeV, without a clear indication of where to look.
  - No much clue about the additional heavy Higgs bosons masses (h, A, H, $H^{\pm}$), but extended scalar sector models still favor the alignment limit cos(β-α) ~ 0.

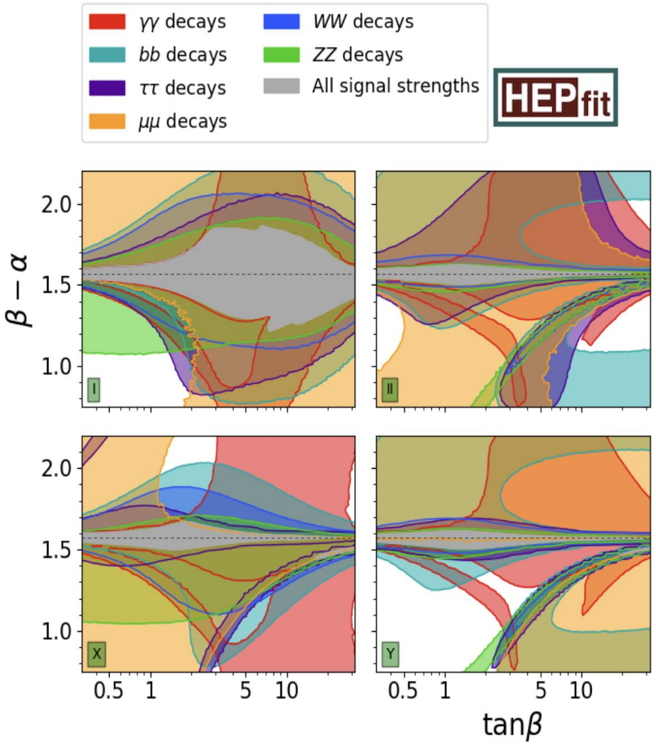






arxiv.1711.02095

The landscape at the start of LHC full run2

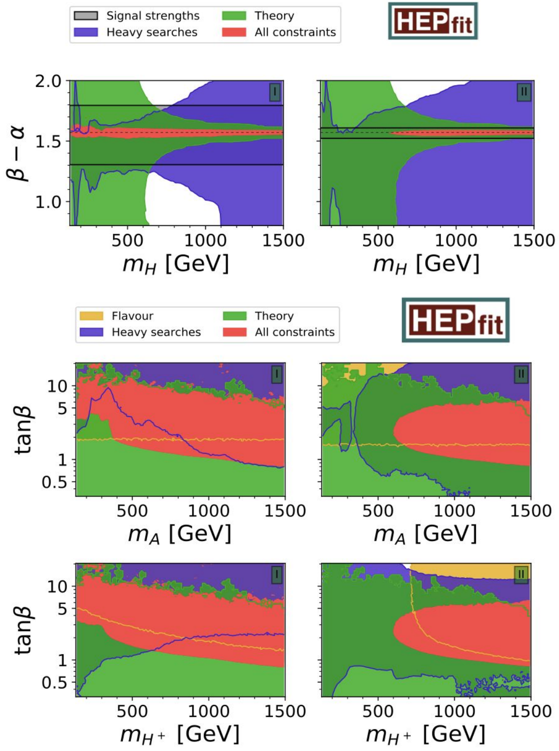


Which parts of 2HDM parameter space are favoured after imposing the latest experimental data from the LHC?

- Constraints from Higgs boson coupling measurements and flavor physics
- Electroweak Precision
- Perturbativity and tree-level unitarity
- etc…

Higgs signal strength push 2HDM toward the alignment limit

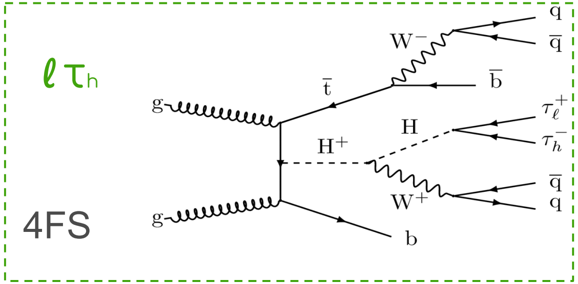
| cos(β-α) | << 1

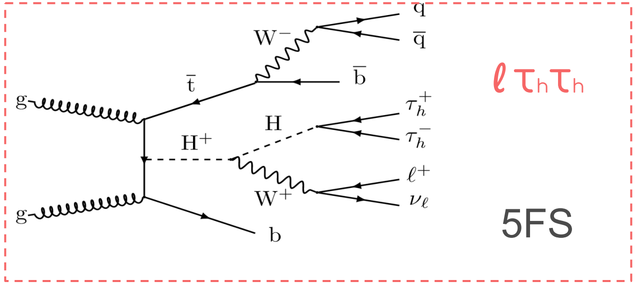


JHEP 09 (2023) 032

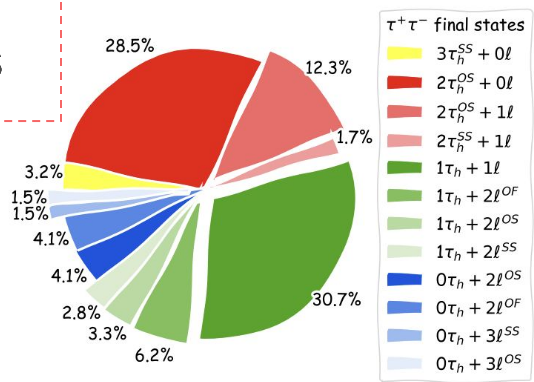
Charged Higgs in WH decays ($H^{\pm} \to W^{\pm}H$)

- $H^{\pm}$ search through the $H^{\pm} \to H W^{\pm}$ and H → ττ decay modes: The H is a heavy CP-even Higgs boson (m_H= 200 GeV) and $H^{\pm}$ in the mass range of 300 to 700 GeV.
- For ($m_{H^{\pm}} > m_t - m_b$) the single-resonant t production dominante.





- Four final states are targeted: eτh, μτh, eτhτh, μτhτh
  - Cover 43% of theoretical BR (30.7% + 12.3%)



- Different fit discriminants are employed, depending on the final state:
  - ℓτh : an MVA boosted decision tree with gradient boost (BDTG)
  - ℓτhτh : the transverse mass of the charged Higgs boson

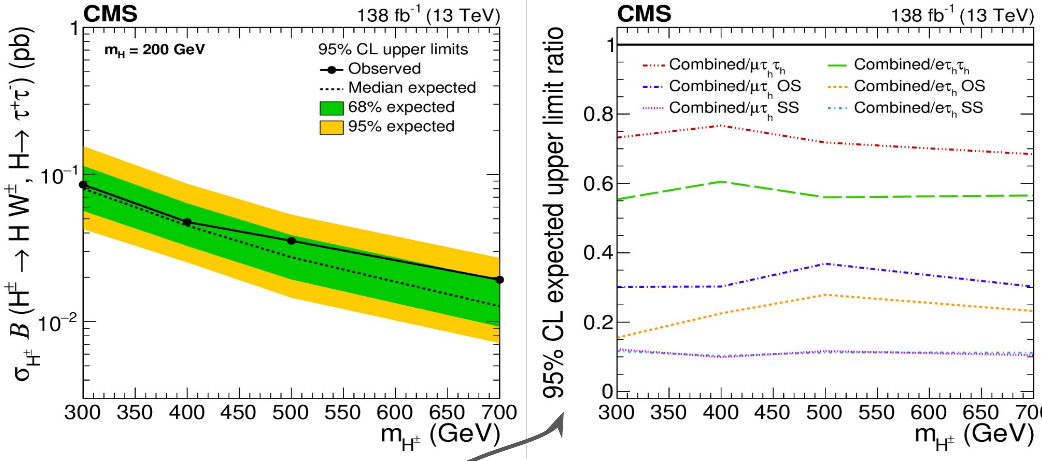
$$m_T^{H^{\pm}} = \sqrt{(E_T^1 + E_T^2 + E_T^W + p_T^{miss})^2 - (\vec{p}_T^{\,1} + \vec{p}_T^{\,2} + \vec{p}_T^{\,W} + \vec{p}_T^{\,miss})^2}$$

JHEP 09 (2023) 032

Charged Higgs in WH decays ($H^{\pm} \rightarrow W^{\pm}H$)

- Four mass points (300, 400, 500 and 700 GeV) are generated for neutral Higgs mass at 200 GeV.
- Observed upper limit between 0.080 pb at 300 GeV to 0.013 pb at 700 GeV.
- Expected sensitivity:
  - $\ell\tau_h\tau_h$ is the most sensitive channel.
  - $\ell\tau_h$ improves sensitivity by 20–35%.

Model Independent Limits



The relative expected contributions of each final state to the overall combination

[JHEP 07 (2023) 073]

MSSM H/A → ττ

- 4 $\tau\tau$ channels: μτh, eτh, τhτh, and eμ
- Production of additional Higgs bosons via ggϕ and bbϕ
- Two search regions: “low-mass” (60–200 GeV), “high-mass” (250–3500 GeV)
- Non-resonant production of $\tau\tau$ by t-channel leptoquark exchange: 5 mass points generated between 1- 5 TeV

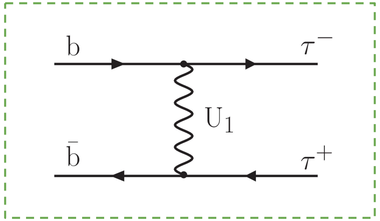
Event categorisation:

- In eτh and μτh channels, split into 2 sub-categories based on:

$$m_{\mathrm{T}}(A,B) = \sqrt{2\, p_{\mathrm{T}}{}^{A}\, p_{\mathrm{T}}{}^{B}\, (1-\cos\Delta\phi^{(A,B)})}$$

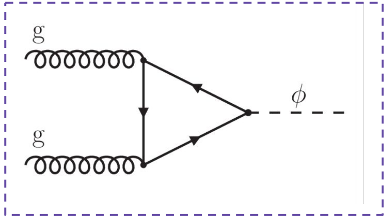
- eμ channel, split into 3 categories based on Dζ:

$$D_\zeta = p_\zeta^{\mathrm{miss}} - 0.85 p_\zeta^{\mathrm{vis}} \Rightarrow \begin{array}{l} p_\zeta^{\mathrm{miss}} = \vec{p}_{\mathrm{T}}^{\,\mathrm{miss}} \cdot \hat{\zeta} \\ p_\zeta^{\mathrm{vis}} = (\vec{p}_{\mathrm{T}}^{\,e} + \vec{p}_{\mathrm{T}}^{\,\mu}) \cdot \hat{\zeta} \end{array}$$

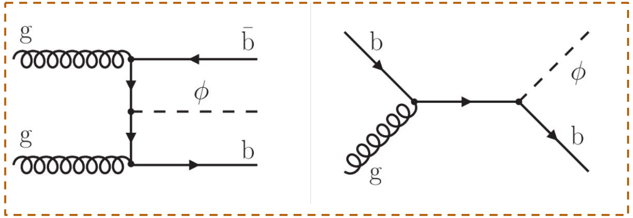


High mass vector-like leptoquarks contribute to $\tau\tau$ events mainly via t-channel exchange.

Vector-like leptoquark

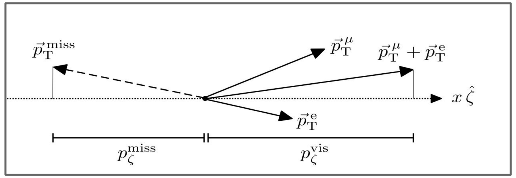


ggφ



bbφ

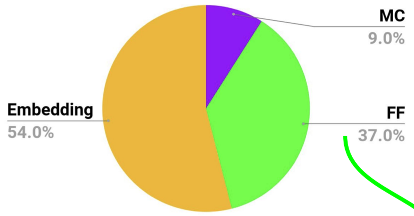
$\hat{\zeta}$ is the vector that bisects $\vec{p}_{\mathrm{T}}^{\,e}$ and $\vec{p}_{\mathrm{T}}^{\,\mu}$



JHEP 07 (2023) 073

MSSM H/A → ττ

- Dominant backgrounds:
  - Genuine di-τ pairs or jet→τh misidentifications
  - Other backgrounds from processes with <2τ e.g. diboson, tt, Z→ $\ell\ell$
  - Di-τ, jet→τh backgrounds, and QCD (eμ) estimated from data.

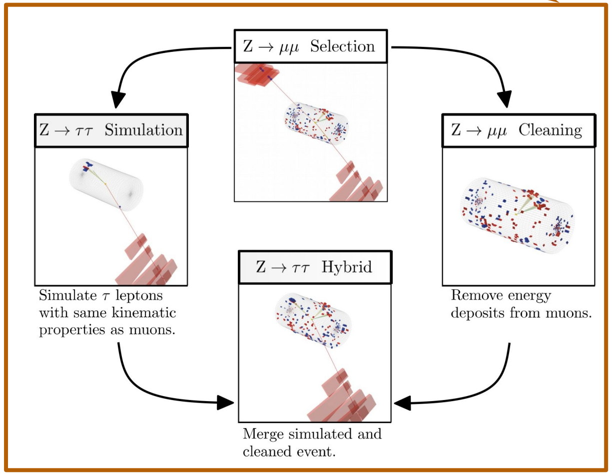


Background modelling:

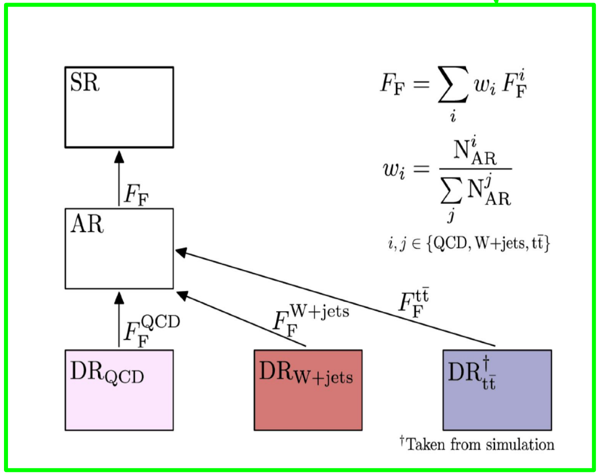
- Embedding method to estimate backgrounds with real di-τ pairs– Z→$\tau\tau$, tt and diboson.

- The “fake factor (FF)” method is used to estimate all backgrounds with jets faking hadronic taus (j→τh)

Scale events by:
FF = (nominal ID) /(relaxed ID)



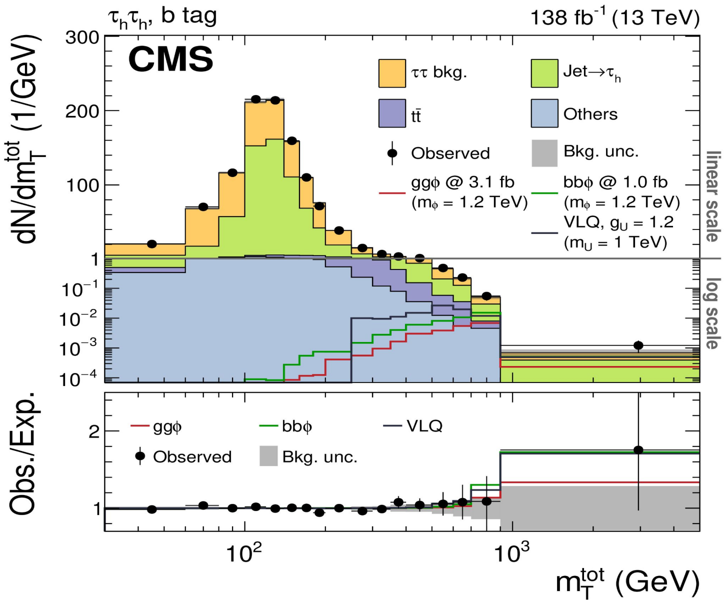
Full details in: JINST 14 (2019) P06032



FF: measured as a function of ρTτh, Njets, and ρTjet/ρTτh

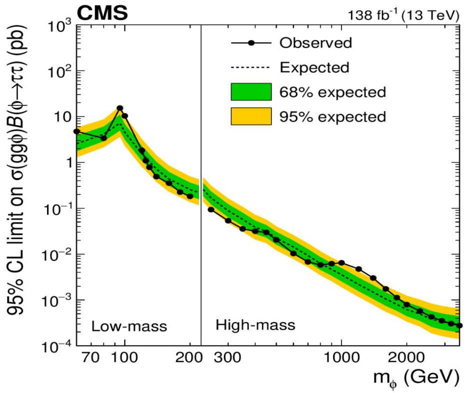
MSSM H/A → ττ

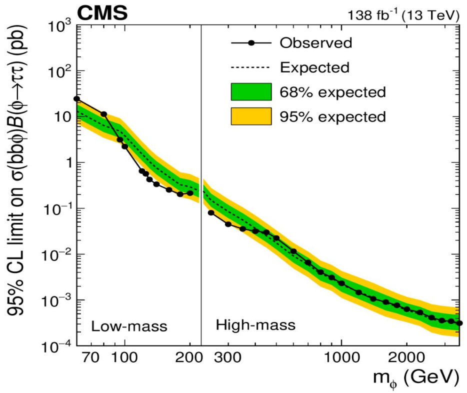
- Two local excesses observed for ggϕ with local (global) significance of
  - 3.1σ (2.7 σ) at 100 GeV
  - 2.8 σ (2.4 σ) at 1.2 TeV



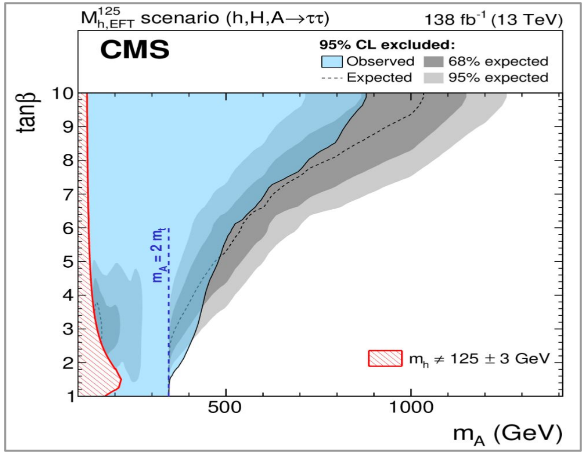
JHEP 07 (2023) 073

Model Independent Limits

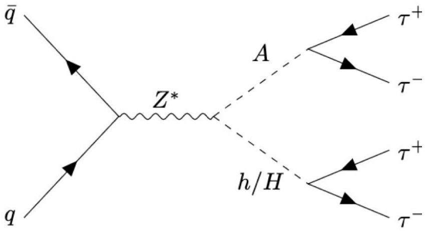




MSSM Model Dependent Limits



Z* → h/H A → 4τ



[CMS-PAS-SUS-23-007]

“ NEW ”

- **Type X (Lepton-Specific) 2HDM:**
  - No suppressed production cross-section.

- **7 decay channels:**
  - Six 4τ final states ⇒ accounting for ≈ 87% of the BR.
  - A 3τ channel to catch events where all τ leptons decay hadronically but one τ_h candidate is lost due to reconstruction inefficiencies.

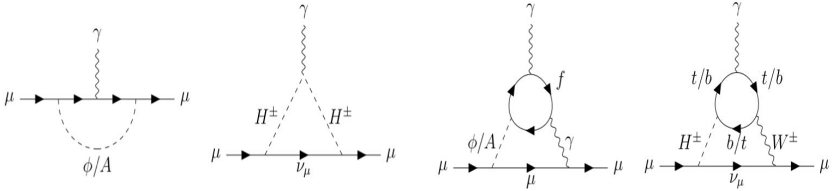
- Improved ML fake factor method to model jets misidentified as hadronic τ.

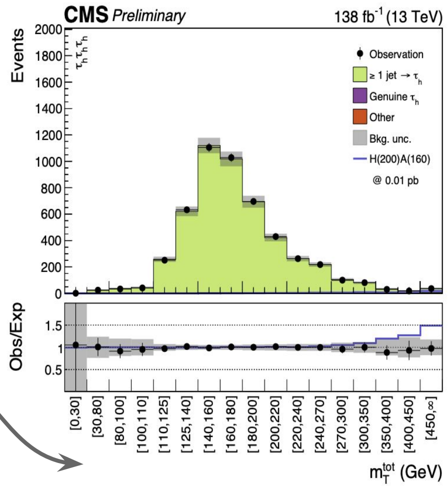
$$m_T^{tot} = \sqrt{\sum_{i=1}^{N_\tau} m_T(\vec{p}_T^{\,\tau_i}, \vec{p}_T^{\,\text{miss}})^2 + \sum_{i,j=1;i\neq j}^{N_\tau} m_T(\vec{p}_T^{\,\tau_i}, \vec{p}_T^{\,\tau_j})^2}$$

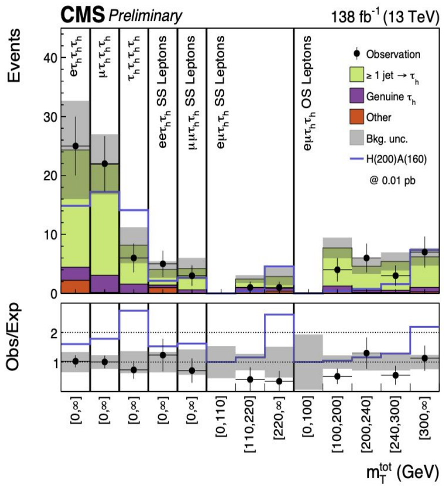
where,

$$m_T(\vec{p}_T^{\,i}, \vec{p}_T^{\,j}) = \sqrt{2p_T^i p_T^j(1 - \cos\Delta\phi)},$$

The additional bosons in the 2HDM can explain the muons g-2 anomaly





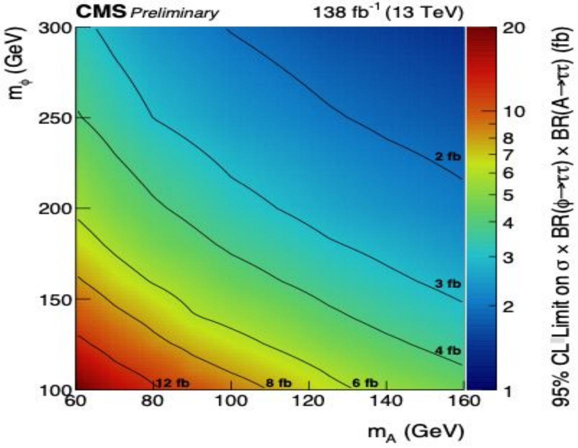


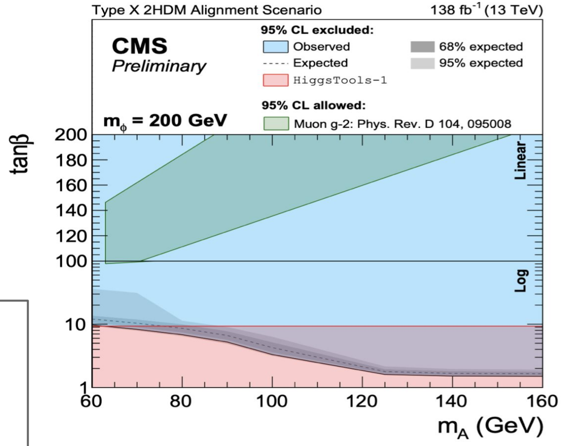
[CMS-PAS-SUS-23-007]

“ NEW ”

Z* → h/H A → 4τ

- $m\varphi >> mZ$: H → ZA decay dominate ⇒ The limit for this search becomes weaker for decreasing mA.
- [MSSM h/H/A → ττ analysis (shown on slide 11)] ⇒ constrain the parameter space to a minimum limit of tan $\beta \approx 10$ ~ an order of magnitude smaller than the area of interest for the g-2 anomaly.
    - For **Z* → h/H A → 4τ** search, as the production cross sections are independent of tan β ⇒ exclude the previously allowed large values of tan β.





TypeX 2HDM as a solution to the g-2 anomaly from [arxiv.2104.10175] is excluded for the whole mass range scanned.

CMS-PAS-HIG-20-002

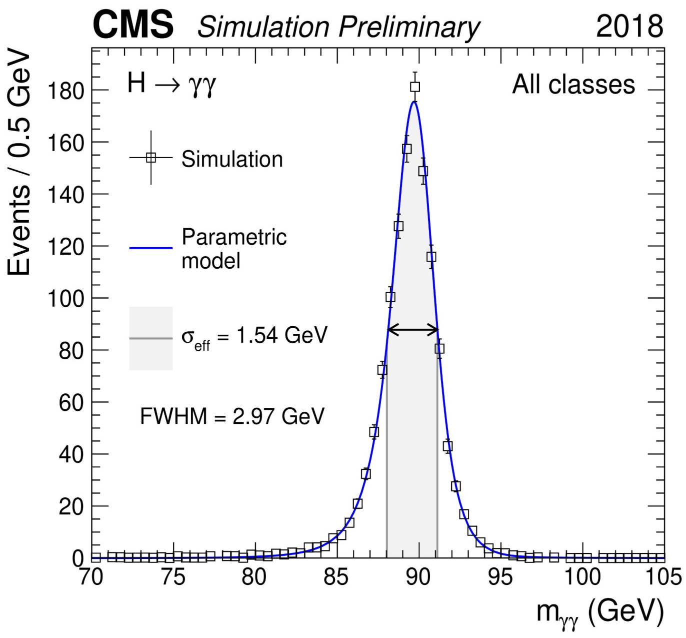
SM-like H → γγ

- **H → γγ for low masses (70 – 110 GeV) ⇒** First search for new resonances in the diphoton final state in this mass range.

- Use of **BDTs** to find diphoton vertex
- Use of the **photonID MVA** to distinguish prompt photons from others.

- Events are then classified according to the output score of the DiphotonBDT distinguishing signal-like from background-like events **⇒ Untagged classification**
- **Dijet and Combined MVA** are trained to distinguish VBF like events **⇒ VBF classification**
- Events that do not pass the VBF class requirement go into **untagged classes.**

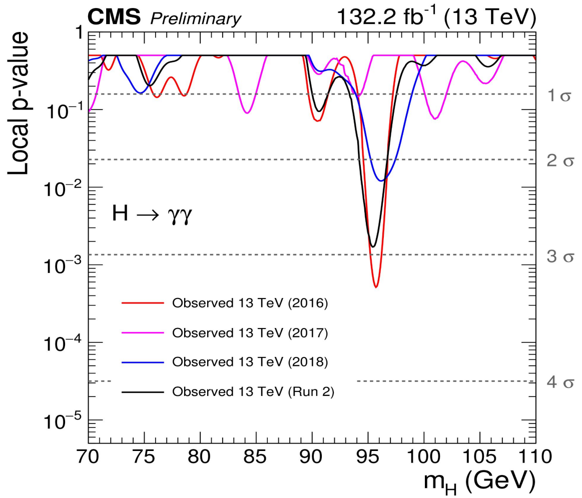
- All production modes (ggH, VBF, WH, ZH, ttH) from 70 GeV to 110 GeV with a 5 GeV granularity are used.

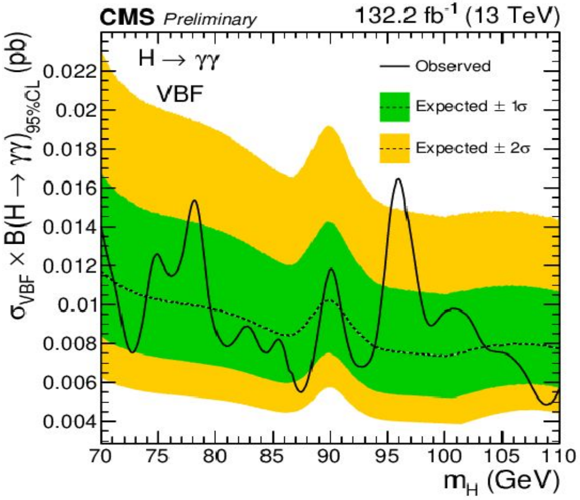


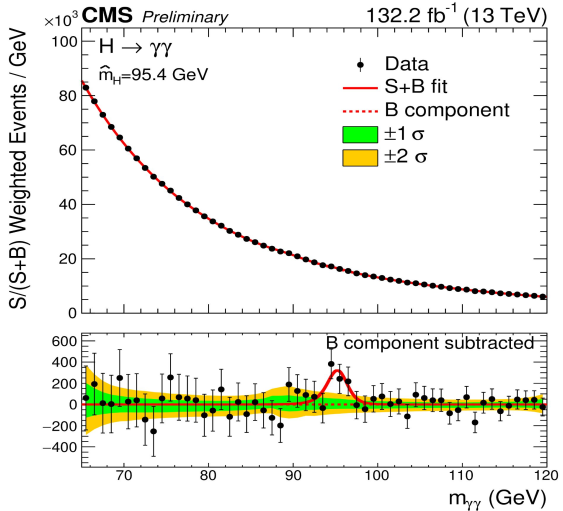
CMS-PAS-HIG-20-002

SM-like H → γγ

- Inflated “f” uncertainty on the DY component normalization, to reduce bias observed at ~90 GeV in some classes*.
    - N*[f*pdf_continuum + (1-f)*pdf_DY] is used to fit the bkg mass (“envelope”). “N” left floating.
- The maximum local significance corresponds to 2.9σ at 95.4 GeV for all production mechanisms and event classes combined (1.3σ global from LEE).
- Signal strengths at 95.4 GeV are compatible among 2016, 2017, 2018 and for all the event classes.







See also talk by Muhammad Aamir Shahzad (Nov 28th, 11:20 AM): here

* see backup slides for event categorisation.

Summary

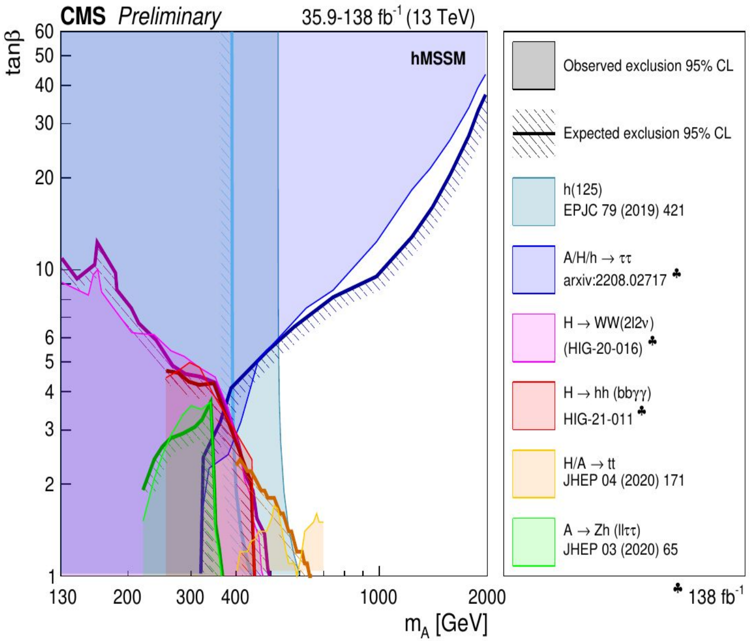
CMSPublic-TWIKI-Summary2HDMSRun2

- CMS has conducted numerous searches for additional heavy Higgs bosons, with many results already published and others still underway…
  - A broad spectrum of signatures, targeting both additional neutral and charged Higgs bosons across various models.
  - With LHC run 3, and the future HL-LHC we are entering the LHC precision era for measurements of the Higgs properties

    ✓ Indirect constraints will also become more relevant.

- Extensive efforts (e.g. search strategies and object reconstruction, including the application of ML techniques, etc.) have resulted in the exclusion of substantial portions of the MSSM parameter space.

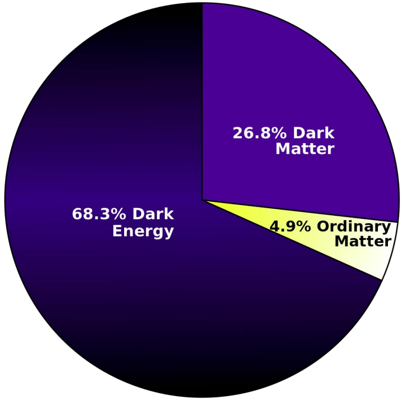
  ⇒ Anticipate further findings and, optimistically, the prospect of groundbreaking discoveries in the near future!



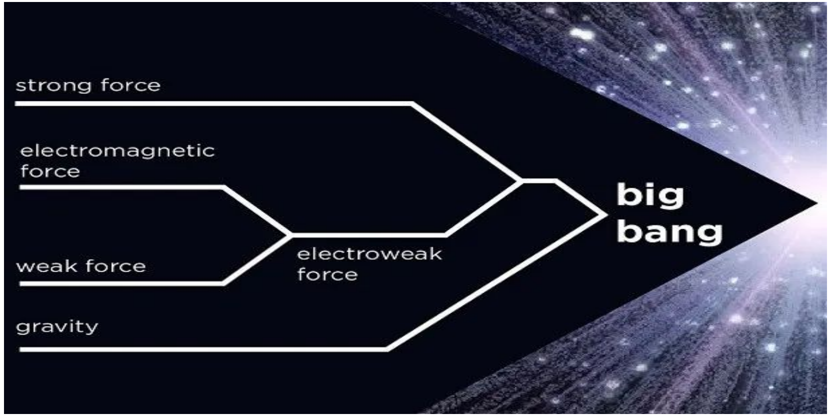
Thanks for listening!

Backup

Problems of the Standard Model (SM)



- Why do neutrinos have mass?
- What is dark matter?
- Why is there so much matter in the Universe?
- Is there a particle associated with the force of gravity?
- **Unification:** Is there a framework that can unify all particles interaction in a so-called Grand Unified Theory (GUT)?



Beyond the Standard Model (BSM) scenarios dealing with these issues tend to:

- Introduce modifications of the Higgs properties, which can be tested by Higgs coupling precision measurements → **Indirect tests of new physics.**
- Introduce new particles in the scalar sector → **Direct searches for new physics.**

arxiv.2004.04172 ; arxiv.1711.02095

Leaving no stone unturned!

Channel	Experiment		Mass range [GeV]	$\mathcal{L}$ [fb^{-1}]
$pp \to H/A \to bb$	CMS	[75]	[0.55;1.2]	2.69
$gg \to H/A \to \tau\tau$	ATLAS	[76]	[0.2;2.25]	36.1
	CMS	[77]	[0.09;3.2]	12.9
$bb \to H/A \to \tau\tau$	ATLAS	[76]	[0.2;2.25]	36.1
	CMS	[77]	[0.09;3.2]	12.9
$pp \to H/A \to \gamma\gamma$	ATLAS	[78]	[0.2;2.7]	36.7
$gg \to H/A \to \gamma\gamma$	CMS	[79]	[0.5;4]	35.9
$gg \to H/A \to Z\gamma[\to (\ell\ell)\gamma]$	ATLAS	[45]	[0.25;2.4]	36.1
$gg \to H/A \to Z\gamma$	CMS	[80]	[0.35;4]	35.9
$gg \to H \to ZZ[\to (\ell\ell)(\ell\ell,\nu\nu)]$	ATLAS	[81]	[0.2;1.2]	36.1
$VV \to H \to ZZ[\to (\ell\ell)(\ell\ell,\nu\nu)]$	ATLAS	[81]	[0.2;1.2]	36.1
$pp \to H \to ZZ[\to (\ell\ell)(\nu\nu)]$	CMS	[82]	[0.6;2.5]	35.9
$gg \to H \to ZZ[\to (\ell\ell)(\nu\nu)]$	CMS	[83]	[0.2;0.6]	2.3
$VV \to H \to ZZ[\to (\ell\ell)(\nu\nu)]$	CMS	[83]	[0.2;0.6]	2.3
$(VV+VH) \to H \to ZZ \to (\ell\ell)(\ell\ell)$	CMS	[84]	[0.13;2.53]	12.9
$pp \to H \to ZZ[\to (\ell\ell)(qq)]$	CMS	[85]	[0.5;2]	12.9
$gg \to H \to ZZ[\to (\ell\ell,\nu\nu)(qq)]$	ATLAS	[86]	[0.3;3]	36.1
$VV \to H \to ZZ[\to (\ell\ell,\nu\nu)(qq)]$	ATLAS	[86]	[0.3;3]	36.1
$gg \to H \to WW[\to (e\nu)(\mu\nu)]$	ATLAS	[87]	[0.25;4]	36.1
$VV \to H \to WW[\to (e\nu)(\mu\nu)]$	ATLAS	[87]	[0.25;3]	36.1
$(gg+VV) \to H \to WW \to (\ell\nu)(\ell\nu)$	CMS	[88]	[0.2;1]	2.3
$gg \to H \to WW[\to (\ell\nu)(qq)]$	ATLAS	[89]	[0.3;3]	36.1
$VV \to H \to WW[\to (\ell\nu)(qq)]$	ATLAS	[89]	[0.3;3]	36.1
$pp \to H \to VV[\to (qq)(qq)]$	ATLAS	[90]	[1.2;3]	36.7
$pp \to H \to hh \to (bb)(bb)$	ATLAS	[91]	[0.3;3]	13.3
	CMS	[92]	[0.26;1.2]	35.9
$gg \to H \to hh \to (bb)(bb)$	CMS	[93]	[1.2;3]	35.9
$pp \to H \to hh[\to (\gamma\gamma)(bb)]$	ATLAS	[94]	[0.275;0.4]	3.2
$pp \to H \to hh \to (\gamma\gamma)(bb)$	CMS	[95]	[0.25;0.9]	35.9
$pp \to H \to hh \to (bb)(\tau\tau)$	CMS	[96]	[0.25;0.9]	35.9
$pp \to H \to hh \to (bb)(VV \to \ell\nu\ell\nu)$	CMS	[97]	[0.26;0.9]	36
$gg \to H \to hh[\to (\gamma\gamma)(WW)]$	ATLAS	[98]	[0.25;0.5]	13.3
$gg \to A \to hZ \to (bb)Z$	ATLAS	[99]	[0.2;2]	36.1
$b\bar{b} \to A \to hZ \to (bb)Z$	ATLAS	[99]	[0.2;2]	36.1

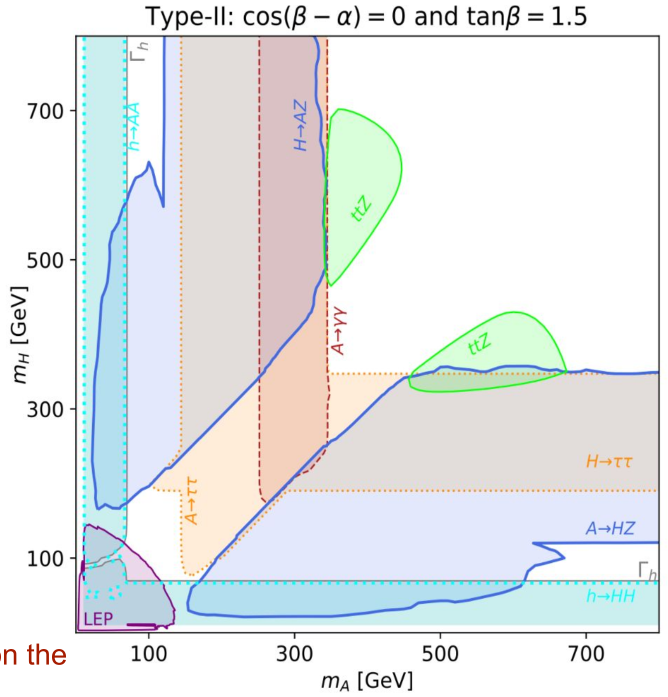
arxiv.2004.04172 ; arxiv.1711.02095

Leaving no stone unturned!

For a specific benchmark of 2HDM free parameter

Constraints from direct searches at 95% C.L. for cos(β-α)= 0 and tanβ = 1.5:

- The combination of all channels cover the majority of the region in which one of the Higgs masses is below the di-top threshold mA, mH < 2 m_top.
- In the gap region: A/H→ ττ, A/H→ γγ are most relevant channels.
- mH (mA) > 2 x m_top, the decay channel A/H→ tt opens up.
- Constraints from the A → Zh and H → V V, hh channels vanish in the alignment limit.



Channel	Experiment		Mass range [GeV]	$\mathcal{L}$ [fb^{-1}]
				2.69
				36.1
				12.9
				36.1
				12.9
				36.7
				35.9
				36.1
				35.9
				36.1
				36.1
				35.9
				2.3
				2.3
				12.9
				12.9
				36.1
				36.1
				36.1
				36.1
				2.3
				36.1
				36.1
				36.7
				13.3
				35.9
				35.9
				3.2
				35.9
				35.9
				36
$gg\to H\to hh[\to(\gamma\gamma)(WW)]$	ATLAS	[98]	[0.25;0.5]	13.3
$gg\to A\to hZ\to(bb)Z$	ATLAS	[99]	[0.2;2]	36.1
$b\bar{b}\to A\to hZ\to(bb)Z$	ATLAS	[99]	[0.2;2]	36.1

⇒ The couplings of the 125 GeV Higgs h are fixed in the SM.
For an extended scalar they are modified and at tree level they depend on the mixing angles cβ−α and tβ.

arxiv.2004.04172 ; arxiv.1711.02095

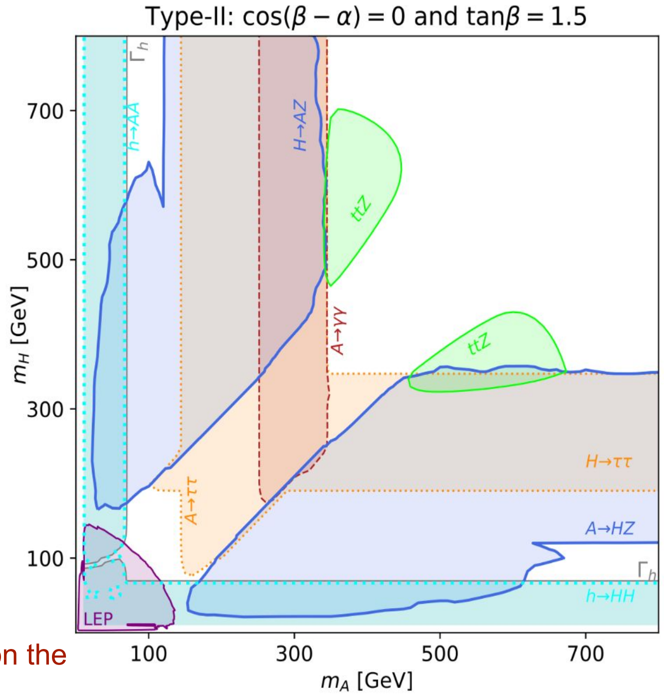
Leaving no stone unturned!

2HDM Type	I	II	III	IV
up-type quarks	ϕ_2	ϕ_2	ϕ_2	ϕ_2
ξ_{huu}	c_α/s_β	c_α/s_β	c_α/s_β	c_α/s_β
ξ_{Huu}	s_α/s_β	s_α/s_β	s_α/s_β	s_α/s_β
ξ_{Auu}	$\tan\beta^{-1}$	$\tan\beta^{-1}$	$\tan\beta^{-1}$	$\tan\beta^{-1}$
down-type quarks	ϕ_2	ϕ_2	ϕ_2	ϕ_1
ξ_{huu}	c_α/s_β	$-s_\alpha/c_\beta$	c_α/s_β	$-s_\alpha/c_\beta$
ξ_{Huu}	s_α/s_β	c_α/s_β	s_α/s_β	c_α/c_β
ξ_{Auu}	$-\tan\beta^{-1}$	$\tan\beta$	$-\tan\beta^{-1}$	$\tan\beta$
lepton	ϕ_2	ϕ_1	ϕ_1	ϕ_2
ξ_{hll}	c_α/s_β	$-s_\alpha/c_\beta$	$-s_\alpha/c_\beta$	c_α/s_β
ξ_{Hll}	s_α/s_β	c_α/c_β	c_α/c_β	s_α/s_β
ξ_{All}	$-\tan\beta^{-1}$	$\tan\beta$	$\tan\beta$	$-\tan\beta^{-1}$

⇒ The couplings of the 125 GeV Higgs h are fixed in the SM.
For an extended scalar they are modified and at tree level they depend on the mixing angles cβ−α and tβ.

Channel	Experiment	Mass range [GeV]	$\mathcal{L}$ [fb^{-1}]
			2.69
			36.1
			12.9
			36.1
			12.9
			36.7
			35.9
			36.1
			35.9
			36.1
			36.1
			35.9
			2.3
			2.3
			12.9
			12.9
			36.1
			36.1
			36.1
			36.1
			2.3
			36.1
			36.1
			36.7
			13.3
			35.9
			35.9
			3.2
			35.9
			35.9
			36
$gg \to H \to hh[\to(\gamma\gamma)(WW)]$	ATLAS [98]	[0.25;0.5]	13.3
$gg \to A \to hZ \to (bb)Z$	ATLAS [99]	[0.2;2]	36.1
$b\bar{b} \to A \to hZ \to (bb)Z$	ATLAS [99]	[0.2;2]	36.1

For a specific benchmark of 2HDM free parameter

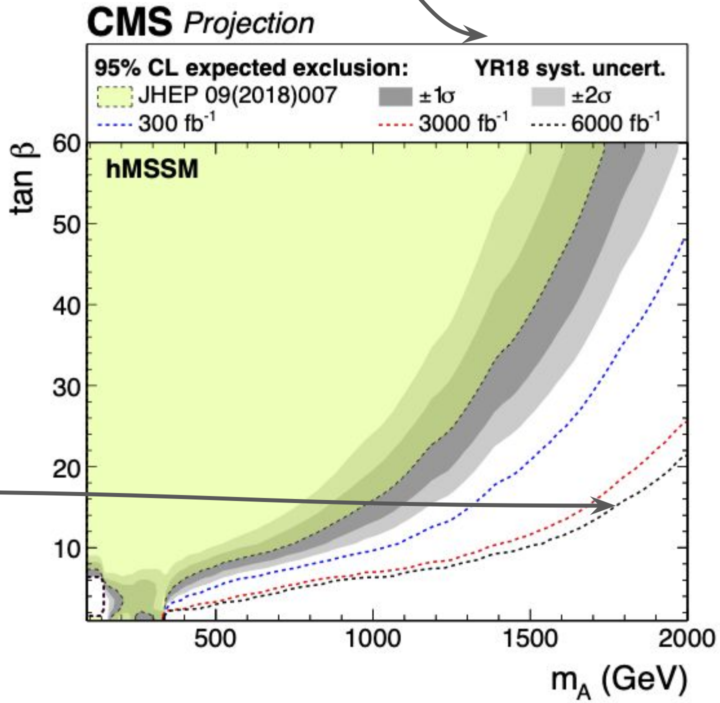


Projection of run2 MSSM H → ττ

YR18 : systematic uncertainties are assumed to decrease with integrated luminosity following a set of assumptions

- For the benchmark scenario $m_{\mathrm{h}}^{\mathrm{mod+}}$ the expected lower limit on the mass of a heavy Higgs boson is extended from 1.25 to 2 TeV for tan β = 36
- For neutral Higgs boson masses above 1 TeV, an improvement by about one order of magnitude is expected in the 95% confidence level upper limits on the cross-section.

Projection of expected MSSM H → ττ 95% CL upper limits based on 2016 data



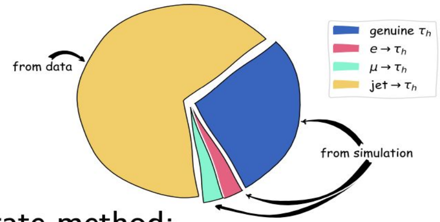
complete HL-LHC:
CMS + ATLAS

JHEP 09 (2023) 032

$H^{\pm} \to W^{\pm}H$: Background Strategy

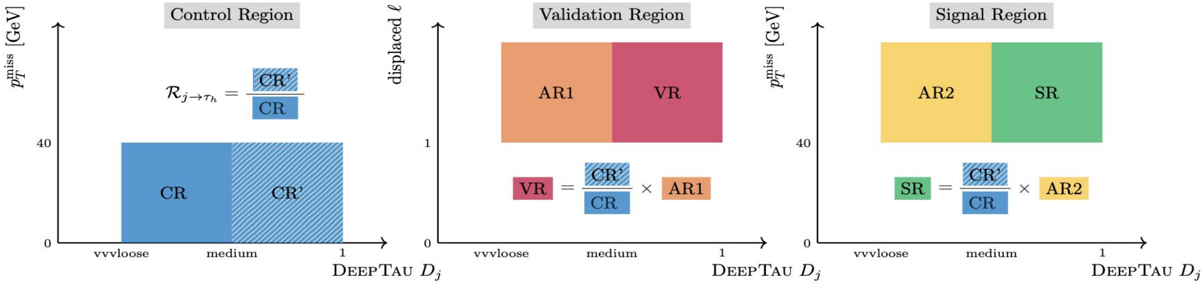
The dominant background ($t\bar{t}$, V+jets) can be decomposed to:

- genuine τ_h
- electron misidentified as τ_h
- muon misidentified as τ_h
- jet misidentified as τ_h



Measure dominant $j \to \tau_h$ from data with fake rate method:

- Estimate τ_h fake rates in control regions (CRs)
- Validate results in validation regions (VRs)



JHEP 09 (2023) 032

$H^{\pm} \rightarrow W^{\pm}H$: BDTG inputs variables

Variable	Description
$\Delta\phi(\tau_h, \vec{p}_T^{miss})$	azimuthal angle between the τ_h and $\vec{p}_T^{miss}$ objects
$\Delta\phi(\ell, \vec{p}_T^{miss})$	azimuthal angle between the ℓ and $\vec{p}_T^{miss}$ objects
$\frac{p_T^{j_1 j_2} - p_T^{H^{\pm}}}{p_T^{j_1 j_2} + p_T^{H^{\pm}}}$	ratio of p_T sums calculated from ℓ, τ_h, j_1, j_2 and $\vec{p}_T^{miss}$
$\frac{p_T^{j_1 j_2}}{H_T}$	ratio of p_T of the first two leading jets and the H_T
$m_T(\ell, \tau_h, j_1, j_2, \vec{p}_T^{miss})$	m_T reconstructed from ℓ, τ_h, j_1, j_2, and $\vec{p}_T^{miss}$
$\frac{p_T^{j_3}}{H_T}$	ratio of the p_T of the third leading jet and the H_T
$m(\ell, \tau_h)$	invariant mass of the ℓ and τ_h objects
$\frac{p_T^{j_1 j_2} + L_T}{H_T}$	ratio of p_T of first two leading jets plus L_T and the H_T
$m_T(\ell, \vec{p}_T^{miss})$	m_T reconstructed from the ℓ and $\vec{p}_T^{miss}$ objects
$p_T^{\tau_h}$	transverse momentum of τ_h object
N_{jets}	number of selected jets in the event
$N_{t^{res}}$	number of selected t^{res} objects in the event

JHEP 07 (2023) 073

MSSM H/A → ττ

High mass: mϕ >= 250 GeV

	No b tag			b tag		
$e\mu$	Low-D_ζ	Medium-D_ζ	High-D_ζ	Low-D_ζ	Medium-D_ζ	High-D_ζ
$e\tau_h$	Loose-m_T	Tight-m_T		Loose-m_T	Tight-m_T	
$\mu\tau_h$	Loose-m_T	Tight-m_T		Loose-m_T	Tight-m_T	
$\tau_h\tau_h$						
$t\bar{t}(e\mu)$				$D_\zeta < -35$ GeV		

Signal region (SR)

Control region

Low mass: mϕ < 250 GeV

	No b tag		b tag	
$e\mu$	Medium-D_ζ: $p_T^{\tau\tau} < 50$ GeV; $50 < p_T^{\tau\tau} < 100$ GeV; $100 < p_T^{\tau\tau} < 200$ GeV; $p_T^{\tau\tau} > 200$ GeV	High-D_ζ: $p_T^{\tau\tau} < 50$ GeV; $50 < p_T^{\tau\tau} < 100$ GeV; $100 < p_T^{\tau\tau} < 200$ GeV; $p_T^{\tau\tau} > 200$ GeV	Medium-D_ζ	High-D_ζ
$e\tau_h$	Tight-m_T: $p_T^{\tau\tau} < 50$ GeV; $50 < p_T^{\tau\tau} < 100$ GeV; $100 < p_T^{\tau\tau} < 200$ GeV; $p_T^{\tau\tau} > 200$ GeV		Tight-m_T	
$\mu\tau_h$	Tight-m_T: $p_T^{\tau\tau} < 50$ GeV; $50 < p_T^{\tau\tau} < 100$ GeV; $100 < p_T^{\tau\tau} < 200$ GeV; $p_T^{\tau\tau} > 200$ GeV		Tight-m_T	
$\tau_h\tau_h$	$p_T^{\tau\tau} < 50$ GeV; $50 < p_T^{\tau\tau} < 100$ GeV; $100 < p_T^{\tau\tau} < 200$ GeV; $p_T^{\tau\tau} > 200$ GeV			
$t\bar{t}(e\mu)$			$D_\zeta < -35$ GeV	

Signal region (SR)

Control region

$$D_\zeta = p_\zeta^{\text{miss}} - 0.85 p_\zeta^{\text{vis}}$$

$$p_\zeta^{\text{miss}} = \vec{p}_T^{\text{miss}} \cdot \hat{\zeta}$$

$$p_\zeta^{\text{vis}} = (\vec{p}_T^{e} + \vec{p}_T^{\mu}) \cdot \hat{\zeta}$$

$\hat{\zeta}$ is the vector that bisects $\vec{p}_T^{e}$ and $\vec{p}_T^{\mu}$

$$m_T(A,B) = \sqrt{2\, p_T{}^A\, p_T{}^B\, (1 - \cos\Delta\phi^{(A,B)})}$$

- **In eτh and μτh channels, split into 2 sub-categories based on:**
  - Tight-mT (mT < 40 GeV)
  - Loose-mT (40 < mT < 70 GeV)
- **In eμ channel, split into 3 categories based on Dζ:**
  - High-Dζ (Dζ > 30 GeV)
  - Medium-Dζ (10 < Dζ <30 GeV)
  - Low-Dζ (-35 < Dζ < 10 GeV)

CMS-PAS-HIG-20-002

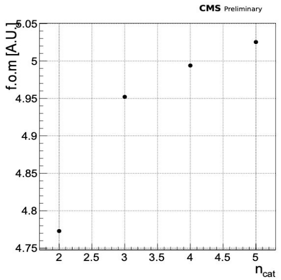
SM-like H → γγ

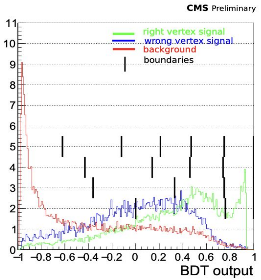
- Event categorization:
  - 2016: 3 untagged (re-categorized) event classes based on LM retrained diphoton BDT
  - 2017 and 2018: 3 untagged event classes based on diphoton BDT, 1 VBF tagged event class based on combined

Event classification: Untagged Classes

Events are classified into untagged classes according to their diphotonBDT output score here in $H \to \gamma\gamma$ on 07/07/2019.

- Simple model using simulated events where class boundaries are adjusted minimizing p-value.
- Enforcing a **minimal width** value for classes to have enough events in each especially for DYMC fitting.
- No significant difference of f.o.m between 3, 4 and 5 classes.
- We choose n_{cat} =3 with boundaries [1.000, 0.753, 0.334, -0.364]



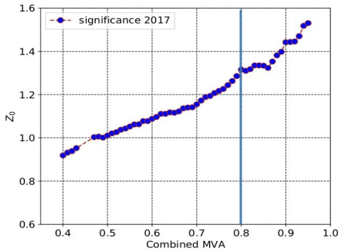


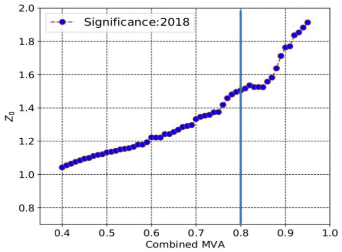
The classification **was redone for 2018** however no significant improvement was found on final results therefore these boundaries are also used for 2018 data analysis.

VBF class: categorisation

Events are classified into the VBF class if they have a Combine MVA output score greater than a defined cut:

- Simple model using simulated events minimizing a dedicated p-value
- Have to make sure to have enough events in the VBF class to perform the related background fits and systematic computation
- due to this the cut value is Combine MVA > 0.8





This cut is also used for 2018 as it has been studied that the obtained signal and background efficiencies with this cut are **similar** wrt to 2017.