Searches for Lepton-flavour-violating Decays of the Higgs Boson into $e\tau$ and $\mu\tau$ using Symmetry Method

Sun Yat-sen University

China LHC Physics Workshop 2023

17/11/023

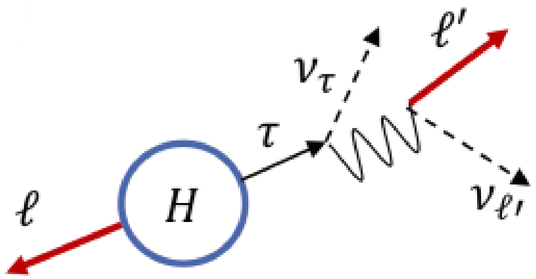
Minlin Wu

Analysis Overview

[JHEP07(2023)166]

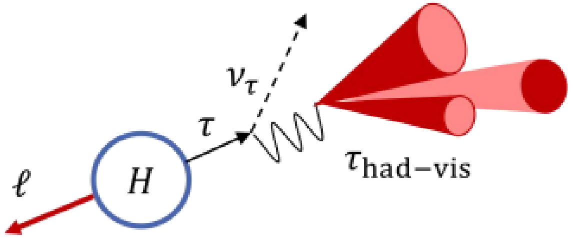
leplep channels

$H \rightarrow \ell\tau_{\ell'}$



lephad channels

$H \rightarrow \ell\tau_{had}$



- Searches using two independent methods:

MC-template Method | leplep channels | lephad channels

dedicated MC-based talk given by Antonio De Maria

Symmetry Method | leplep channels

- Based on the assumption that SM processes are symmetric w.r.t. $e \leftrightarrow \mu$ exchange.
- LFV signals break the symmetry if $\boldsymbol{\mathcal{B}}(H \rightarrow e\tau) \neq \boldsymbol{\mathcal{B}}(H \rightarrow \mu\tau)$.
- Background estimated via data-driven symmetry method.

lephad channels

- Run 2 study in progress, will be involved in R22 analysis.

Analysis Overview

- Considered Higgs production modes: ggH, VBF, VH

- Cut-based signal region categorization: categorized into VBF and nonVBF region, enhance contribution from main Higgs production modes.

- Performed MVA in each signal region to enhance sensitivity.

- Symmetry method:
  - Can measure the branching ratio difference, $\boldsymbol{\mathcal{B}}(H \to e\tau) - \boldsymbol{\mathcal{B}}(H \to \mu\tau)$, or $\boldsymbol{\mathcal{B}}(H \to \mu\tau) - \boldsymbol{\mathcal{B}}(H \to e\tau)$. Search for the Higgs LFV decays by assuming the latter Br is zero:
    - $H \to e\tau$ search: measure $\boldsymbol{\mathcal{B}}(H \to e\tau) - \boldsymbol{\mathcal{B}}(H \to \mu\tau)$.
    - $H \to \mu\tau$ search: measure $\boldsymbol{\mathcal{B}}(H \to \mu\tau) - \boldsymbol{\mathcal{B}}(H \to e\tau)$.

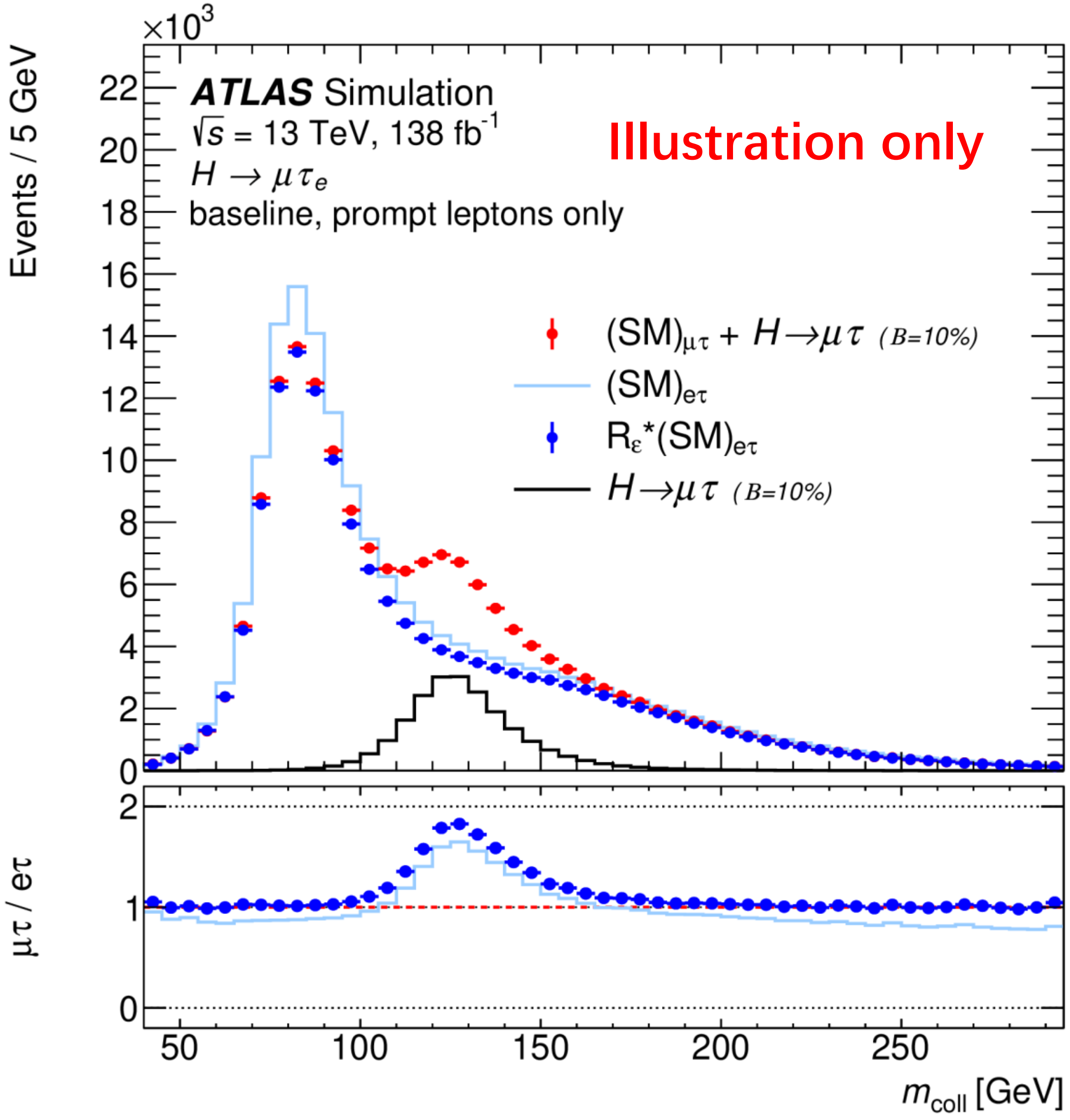
  - Two measurements of the Br difference are compatible, final results interpreted as $\boldsymbol{\mathcal{B}}(H \to \mu\tau) - \boldsymbol{\mathcal{B}}(H \to e\tau)$ (has smaller expected uncertainty).

  - Compare to MC-based method, symmetry method is more sensitive to the Br difference.

Example for $e\mu$ – symmetry Method

- SM processes are symmetric w.r.t. prompt $e \leftrightarrow \mu$ exchange (assume a symmetry between $H \rightarrow e\tau_\mu$ and $H \rightarrow \mu\tau_e$).

Illustration of the $e\mu$-symmetry method showing how the $H \rightarrow \mu\tau$ LFV signal can be discovered by comparing data yields in the $e\tau$ and $\mu\tau$ channels. Based on toy simulated data.



Background Estimation

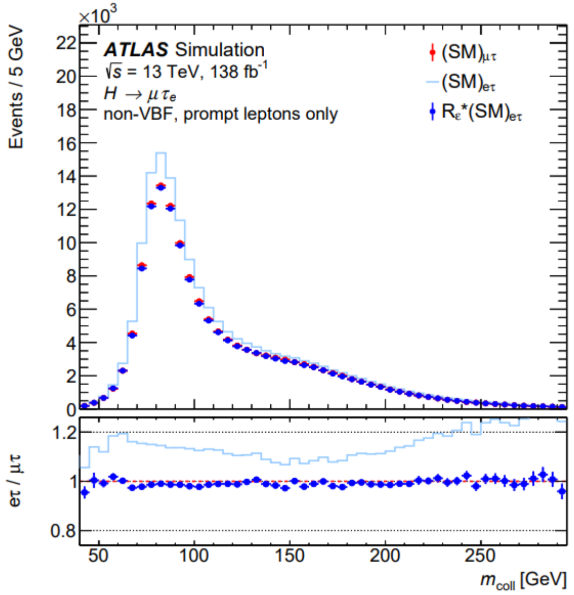
- **Induced asymmetry**:
  - **Data efficiency**: trigger, reconstruction, identification, isolation.
    - Due to the non-perfect detection, the resultant difference in the detection effects give rise to an asymmetry in the number of detected SM events in each channel, which should be corrected via efficiency ratio:

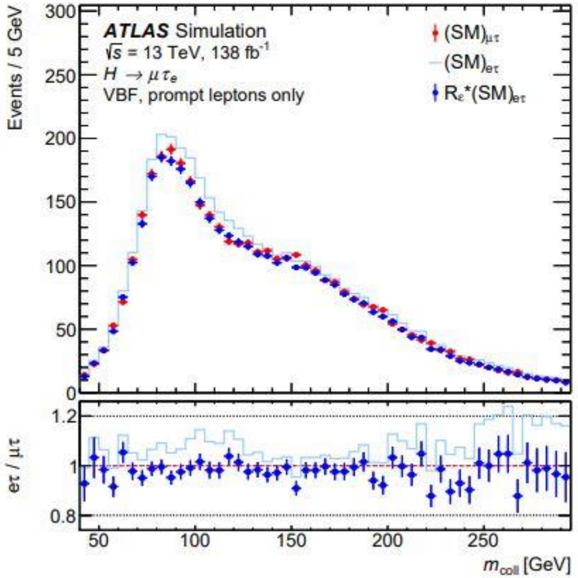
$$\frac{N^{e\tau}}{\epsilon^{e\tau}} = \frac{N^{\mu\tau}}{\epsilon^{\mu\tau}}, \quad N^{\mu\tau} = R^{\epsilon} N^{e\tau}$$

  - **Mis-identified lepton**: Different fake rate for e and μ.

  - Background estimated using the data sample of the other channel:

$$b^{\mu\tau} = R^{\epsilon}\,(N^{e\tau}_{Data} - f^{e\tau}) + f^{\mu\tau}$$





Background Estimation

Symmetric background

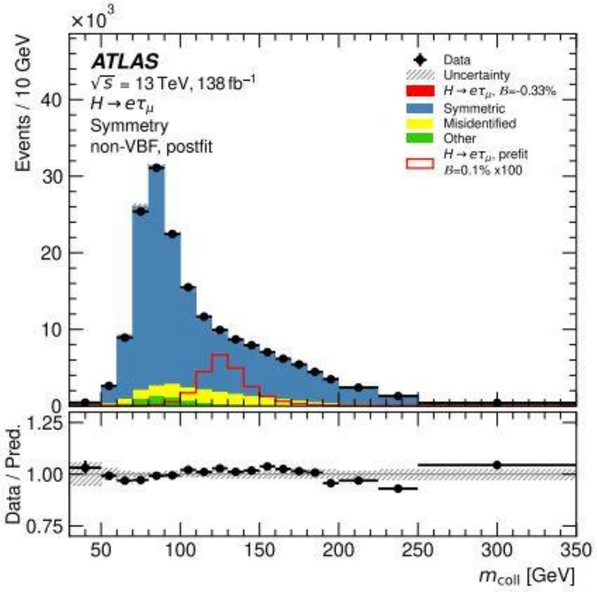
Including $Z \to \tau\tau$, diboson, top-quark processes ($t\bar{t}$ and single top) and SM Higgs boson decays ($H \to \tau\tau$ and $H \to WW$).

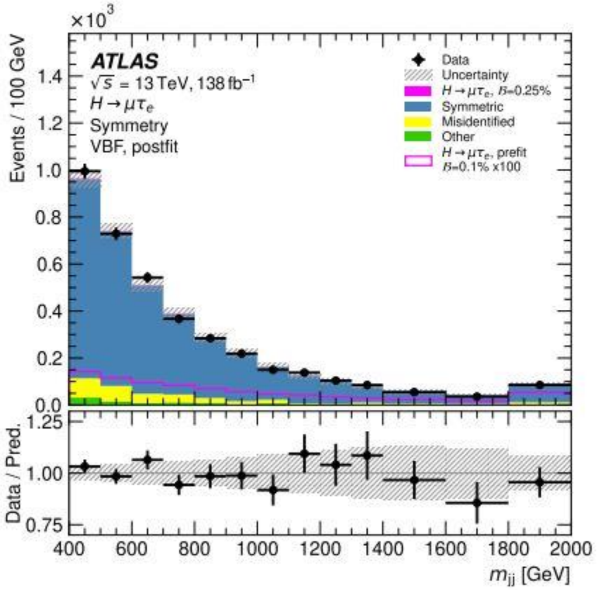
Fake lepton

Fake lepton background from the jets mis-identified as light leptons and light leptons originating from hadronic decays within a jet are estimated by the data-driven fake-factor method. Fake factors estimated in a Z+jets CR.

MC fakes

Fake leptons background from $\tau_{had} \to \ell$, $\mu \to e$, and $\gamma \to e$. Events containing one or more mis-identified light leptons are estimated from MC simulation.





MVA Strategy

Binned significance:

$$Z = \sqrt{\sum_i z_i^2},\ z_i = \sqrt{2\left((s_i + b_i)\log\left(1 + \frac{s_i}{b_i}\right) - s_i\right)}$$

- Exploited either multi-classifier or combined score method.
- **Combined NN score approach**:
  - Two or three NNs are trained focusing on different backgrounds.
  - NN scores are combined linearly or quadratically:

$$S_{linear} = \frac{1}{\sum_i c_i}\left(\sum_i c_i S_i\right),\ S_{quadratic} = \frac{1}{\sqrt{\sum_i c_i^2}}\left(\sum_i c_i^2 S_i^2\right)$$

  - Combinations and coefficients c_i were optimized by binned significance.

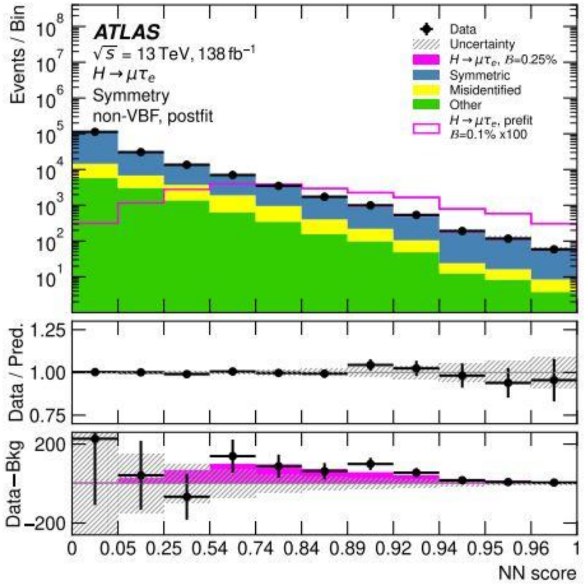
symmetry method | leplep channels

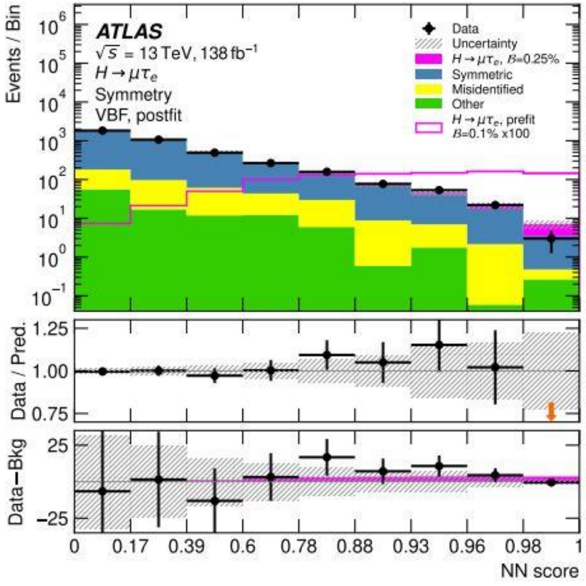
$e\tau$ and $\mu\tau$ trained together

nonVBF: multi-classifier – signal node, symmetric bkg node, fakes node

VBF: 3 NNs

- sig vs $top + diboson + H \rightarrow WW$
- sig vs $jets \rightarrow \ell$ fakes
- sig vs $Z \rightarrow \tau\tau + \mathrm{other} \rightarrow \ell\ \mathrm{fakes} + \mathrm{H} \rightarrow \tau\tau$





Statistical Analysis

- **MVA output** used as final discriminant to extract the signal strength ($\mu_{e\tau}$ and $\mu_{\mu\tau}$) and upper limits at 95% C.L.

1-POI fit

Independent search for the $H \to e\tau$ ($\mu\tau$) signal.
$\boldsymbol{\mathcal{B}}(H \to \mu\tau)$ ($\boldsymbol{\mathcal{B}}(H \to e\tau)$) is set to zero.

	nonVBF	**VBF**
MC-template $\ell\tau_{\ell'}$	✓	
MC-template $\ell\tau_{had}$	✓	✓
Symmetry $\boldsymbol{\ell\tau_{\ell'}}$		✓

2-POI fit

Simultaneous measurement of the $H \to e\tau$ and $H \to \mu\tau$ signals.
No assumption.

	nonVBF	**VBF**
MC-template $\ell\tau_{\ell'}$	✓	✓
MC-template $\ell\tau_{had}$	✓	✓
Symmetry $\boldsymbol{\ell\tau_{\ell'}}$		

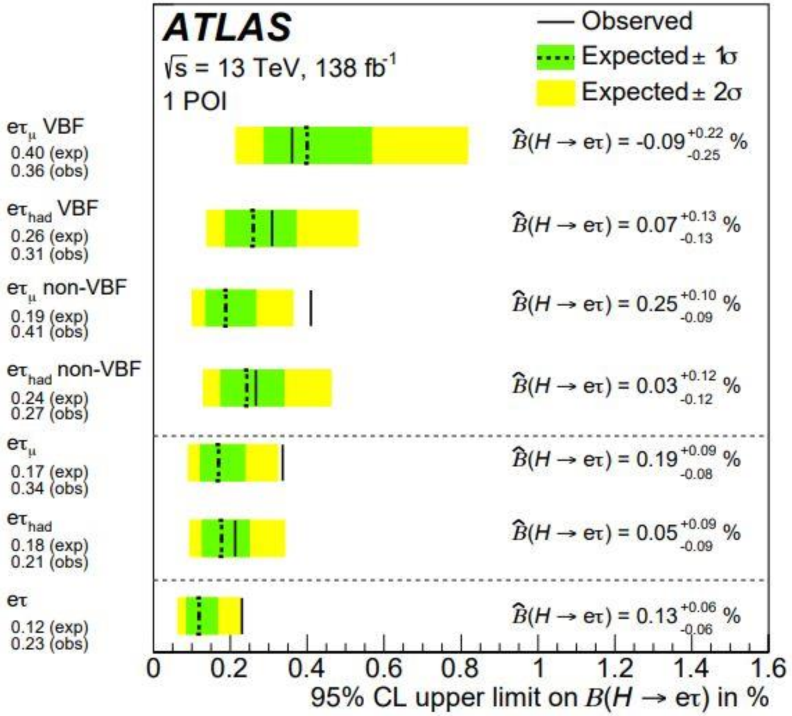
- 1-POI fit: symmetry-based leplep VBF combined with MC-based leplep nonVBF.
- No 2-POI fit for symmetry-based as one channel is required for the background estimation of the other.

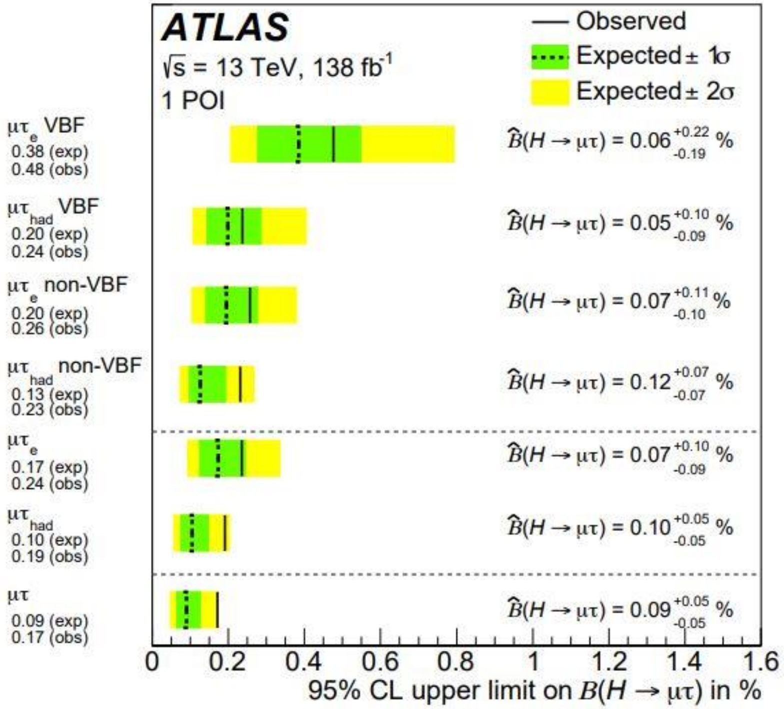
1-POI Fit Results

- 2.2σ excess observed for $\mathcal{B}(H \to e\tau)$,
- 1.9σ excess observed for $\mathcal{B}(H \to \mu\tau)$.
- Largest impacts on the POIs arise from:
  - Symmetric background estimate unc.
  - JET+MET unc.
  - Signal theory unc.

	nonVBF	VBF
MC-template $\ell\tau_{\ell'}$	✓	
MC-template $\ell\tau_{had}$	✓	✓
Symmetry $\ell\tau_{\ell'}$		✓

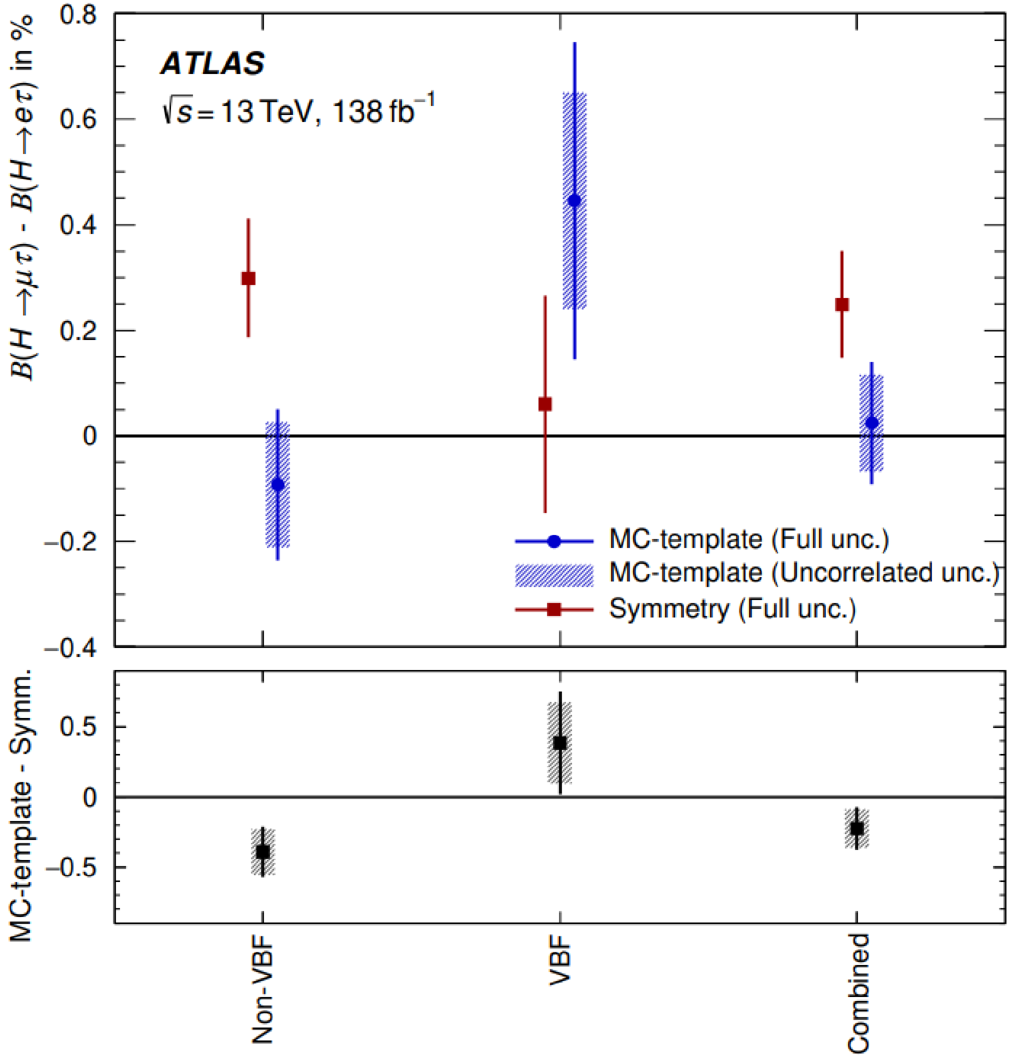
	Observed (expected) 95% CL upper limit [%]
$H \to e\tau$	0.23 (0.12)
$H \to \mu\tau$	0.17 (0.09)





Branching Ratio Difference

- Symmetry method is **more sensitive** to the branching ratio difference $\mathcal{B}(H \to \mu\tau)$ - $\mathcal{B}(H \to e\tau)$.
- The measured value of $\mathcal{B}(H \to \mu\tau)$ - $\mathcal{B}(H \to e\tau)$ is (0.25 ± 0.10)%, favours a larger Br for the $H \to \mu\tau$ signal with 2.5σ significance.
- $\mathcal{B}(H \to \mu\tau)$ - $\mathcal{B}(H \to e\tau)$ is also measured for the MC-template method, the best-fit value is (0.02 ± 0.12)%. The correlations between two methods are taken into account for checking the compatibility (compatible within 2σ):
  - only the data statistical uncertainty and signal uncertainties are correlated.

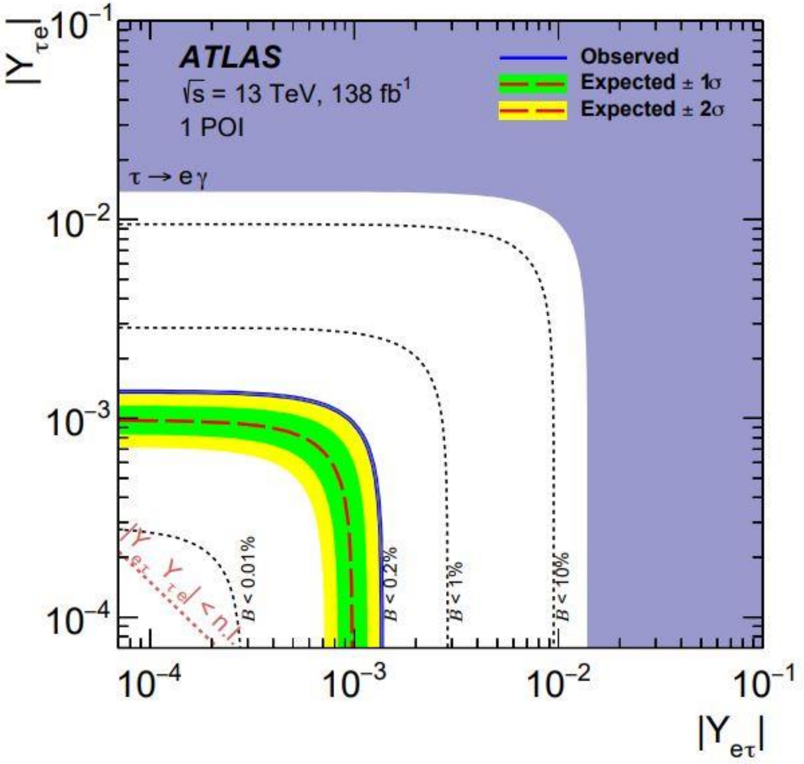


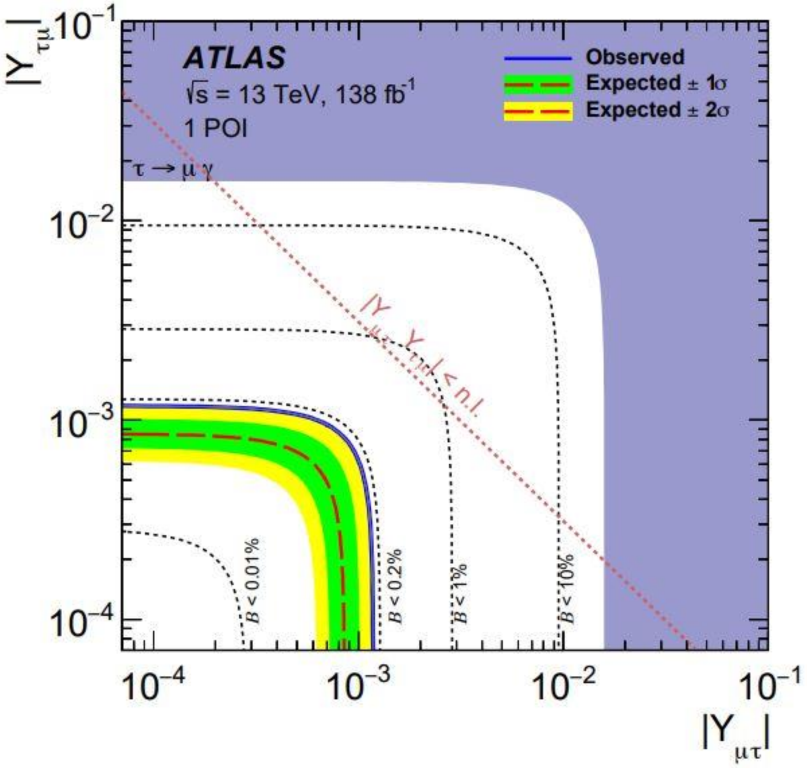
1-POI Fit Results – Yukawa Coupling

- Br values can be related to the non-diagonal Yukawa coupling matrix elements:

$$|Y_{\ell\tau}|^2 + |Y_{\tau\ell}|^2 = \frac{8\pi}{m_H}\frac{\boldsymbol{\mathcal{B}}(H \to \ell\tau)}{1-\boldsymbol{\mathcal{B}}(H \to \ell\tau)}\Gamma_H(SM)$$

- 1-POI fit:
  - $\sqrt{|Y_{e\tau}|^2 + |Y_{\tau e}|^2} < 0.0014$
  - $\sqrt{|Y_{\mu\tau}|^2 + |Y_{\tau\mu}|^2} < 0.0012$

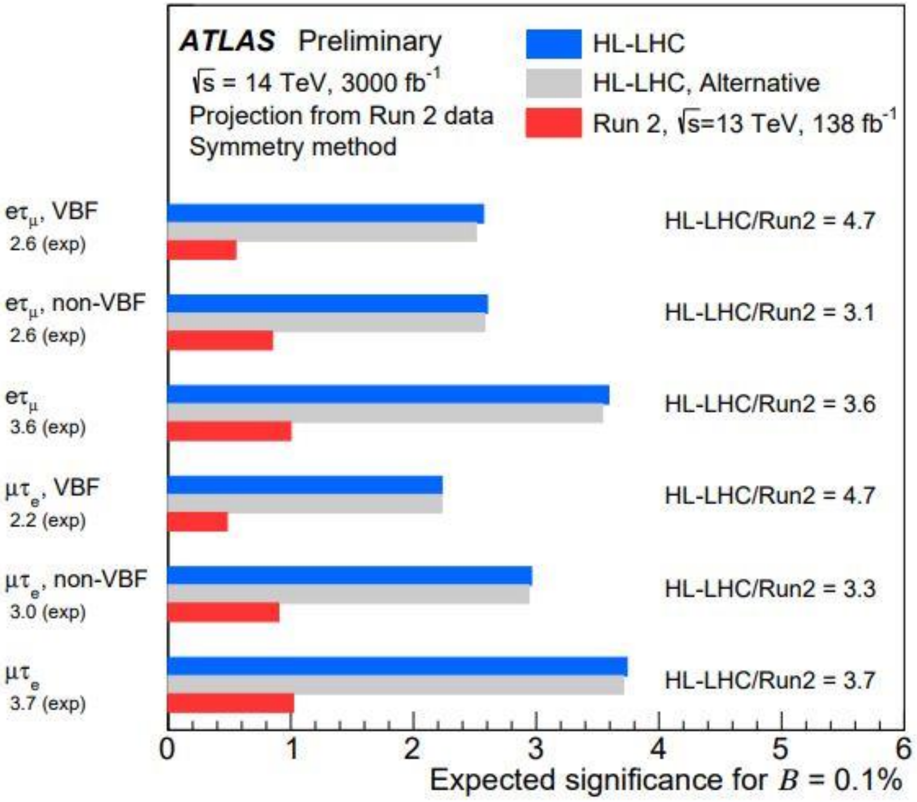


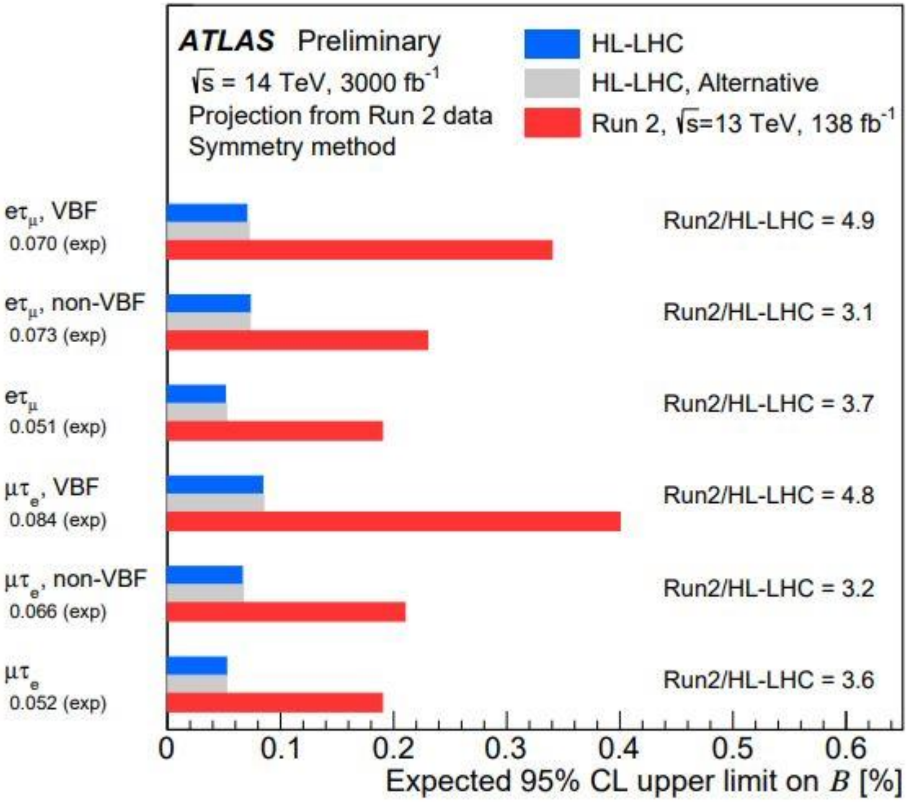


- Compared with the indirect limits from $\tau \to \ell\gamma$ searches, direct limits about one order of magnitude tighter in $Y_{\ell\tau}$.

- In the $H \to \mu\tau$ case, the constraints are tighter than the naturalness limit (n.l.): $\sqrt{Y_{\ell\tau}Y_{\tau\ell}} \lesssim \frac{m_\tau m_\ell}{v^2}$, preventing a non-hierarchical mass spectrum from large off-diagonal terms.

HL-LHC Prospects





- HL-LHC extrapolation [ATL-PHYS-PUB-2022-054]:
  - Run 2 results are extrapolated to HL-LHC scenario.
  - Expect an improvement on the current limits by a factor of 3.6-3.7

Conclusion

- Searches for $H \rightarrow e\tau$ and $H \rightarrow \mu\tau$ using symmetry method.
- The analysis results of leplep channels have been presented, the Run 2 study of $H \rightarrow \ell\tau_{had}$ final states using symmetry method in progress (first time), will be involved in R22 analysis.
- No significant excess is observed, but get hints for Run 3:
  - Symmetry method: $\boldsymbol{\mathcal{B}}(H \rightarrow \mu\tau)$ - $\boldsymbol{\mathcal{B}}(H \rightarrow e\tau)$ is $(0.25 \pm 0.10)\%$, compatible with zero within 2.5σ.
- HL-LHC prospect: expect an improvement on the current limits by a factor of 3.6-3.7.

Thank you!

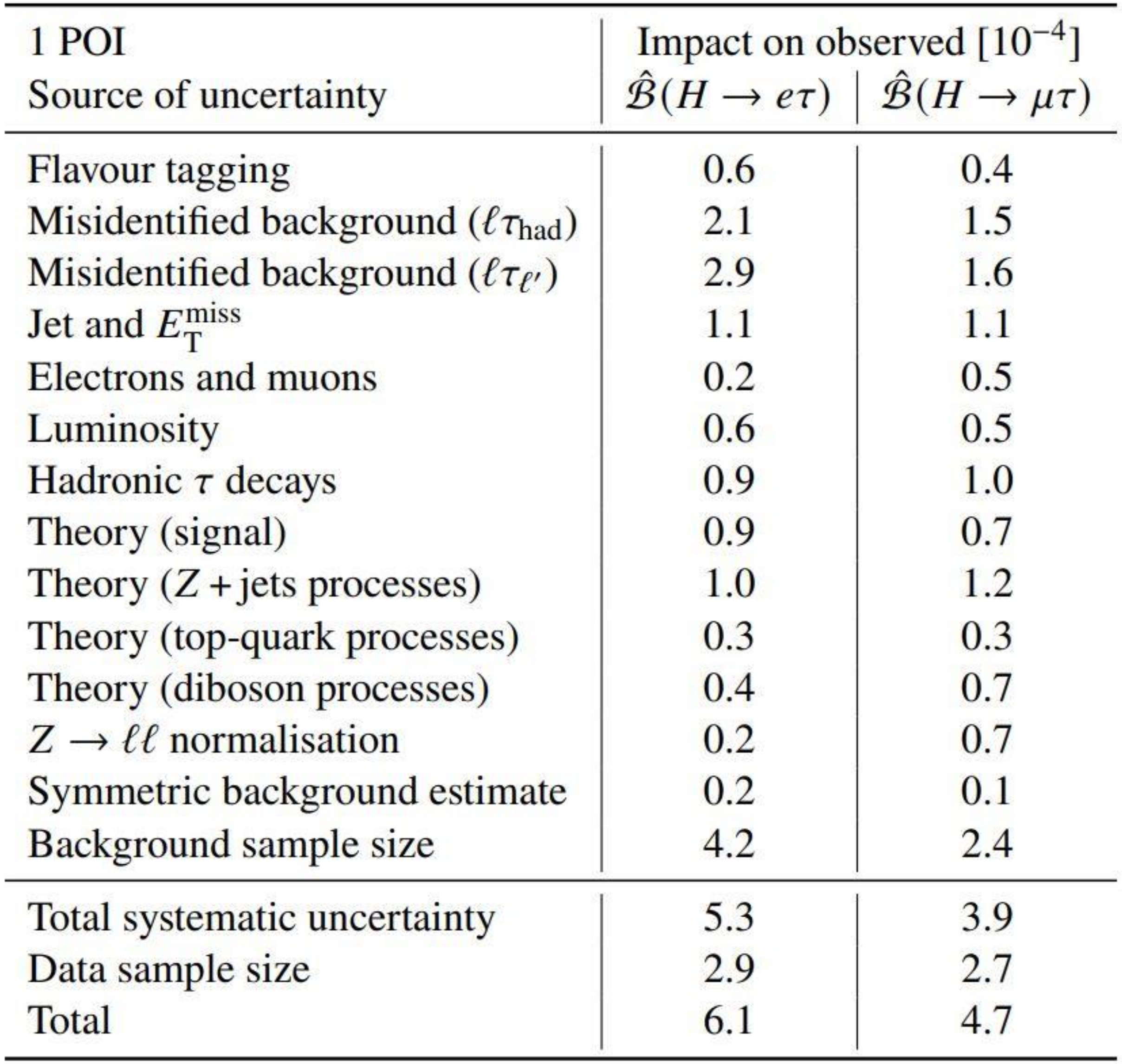
BACKUP

Event Selection

🟥 Symmetry method ⬛ MC-template method

	leplep ($\ell\tau_{\ell'}$)		lephad ($\ell\tau_{had}$)	
	$e\tau_\mu$	$\mu\tau_e$	$e\tau_{had}$	$\mu\tau_{had}$
Baseline	$1e$, 1μ, opposite sign, pass ID, Iso		1ℓ, $1\tau_{had-vis}$, opposite sign, pass ID, Iso	
	Trigger selection			
	Veto b-tagged jets			
	Veto $\tau_{had-vis}$		e-veto to reject e-faking-$\tau_{had-vis}$	**e-veto**
	$p_T^{\ell_1} > 45$ (**35**) GeV, $p_T^{\ell_2} > 15$ GeV		$p_T^{\ell} > 27.3$ (**30**) GeV, $\boldsymbol{|\eta^{\ell}| < 2.4}$	
	$30 < m_{\ell\ell'} < 150$ GeV		$p_T^{\tau_{had-vis}} > 25$ GeV, $|\eta^{\tau_{had-vis}}| < 2.4$	
	If ℓ_2=e, $0.2 < p_T^{track}(\ell_2)/p_T^{cluster}(\ell_2) < 1.25$		$|\Delta\eta(\ell,\tau_{had-vis})| < 2$, to reject QCD multi-jets	
			$\sum \cos\Delta\phi(\ell, E_T^{miss}) > -0.35$, to reject W+jets	
VBF	Pass baseline selection			
	$N_{jets}(p_T > 30) \geq 2$			
	$p_T^{leading\ jet} > 40$ GeV			
	$m_{jj} > 400$ GeV			
	$|\Delta\eta_{jj}| > 3$			
nonVBF	Pass baseline, fail VBF selection			
			veto events if	
			$90 < m_{vis}(e,\tau) < 100$ GeV	$90 < m_{vis}(\mu,\tau) < 100$ GeV

1-POI Fit Uncertainty Breakdown

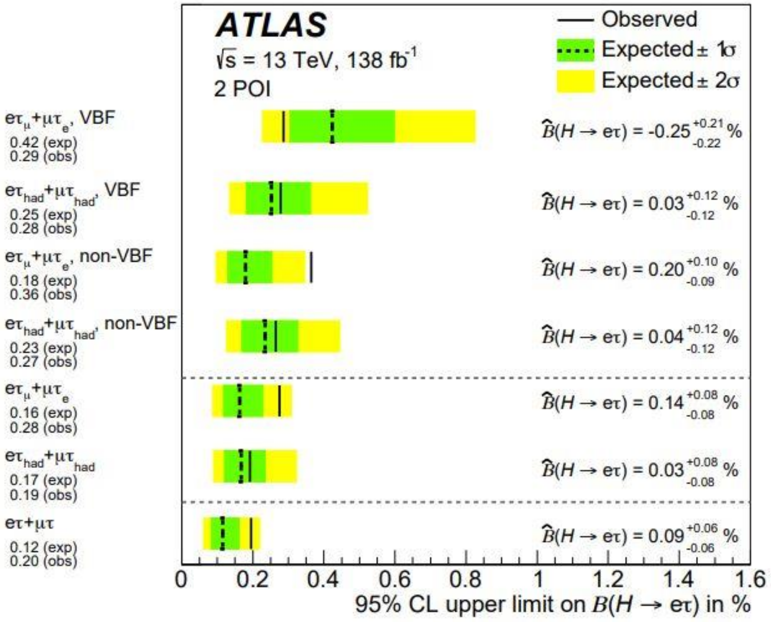


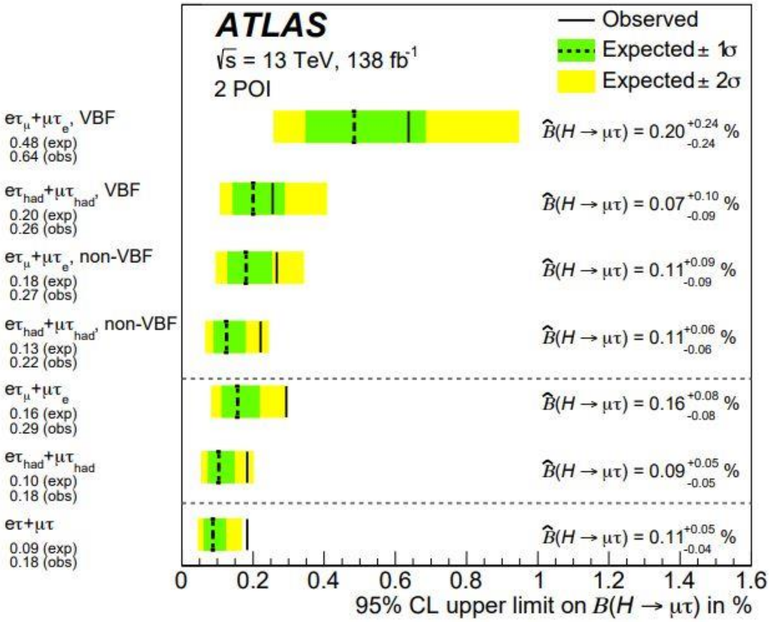
1 POI	Impact on observed $[10^{-4}]$	
Source of uncertainty	$\hat{\mathcal{B}}(H \to e\tau)$	$\hat{\mathcal{B}}(H \to \mu\tau)$
Flavour tagging	0.6	0.4
Misidentified background ($\ell\tau_{\mathrm{had}}$)	2.1	1.5
Misidentified background ($\ell\tau_{\ell'}$)	2.9	1.6
Jet and $E_{\mathrm{T}}^{\mathrm{miss}}$	1.1	1.1
Electrons and muons	0.2	0.5
Luminosity	0.6	0.5
Hadronic τ decays	0.9	1.0
Theory (signal)	0.9	0.7
Theory ($Z+$jets processes)	1.0	1.2
Theory (top-quark processes)	0.3	0.3
Theory (diboson processes)	0.4	0.7
$Z \to \ell\ell$ normalisation	0.2	0.7
Symmetric background estimate	0.2	0.1
Background sample size	4.2	2.4
Total systematic uncertainty	5.3	3.9
Data sample size	2.9	2.7
Total	6.1	4.7

2-POI Fit Results

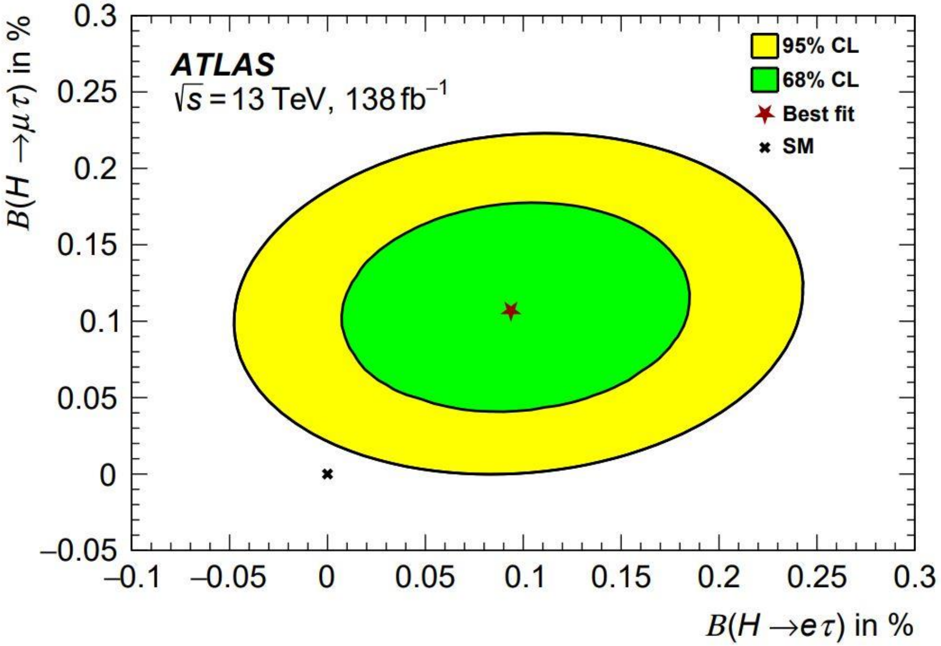
- No significant excess is observed:
  - For $H \to e\tau$ signal, 1.6σ excess.
  - For $H \to \mu\tau$ signal, 2.5σ excess.

- Largest impacts on POIs from:
  - Background statistical uncertainties
  - Mis-identified background uncertainties
  - Jet and MET systematics
  - lepton systematics





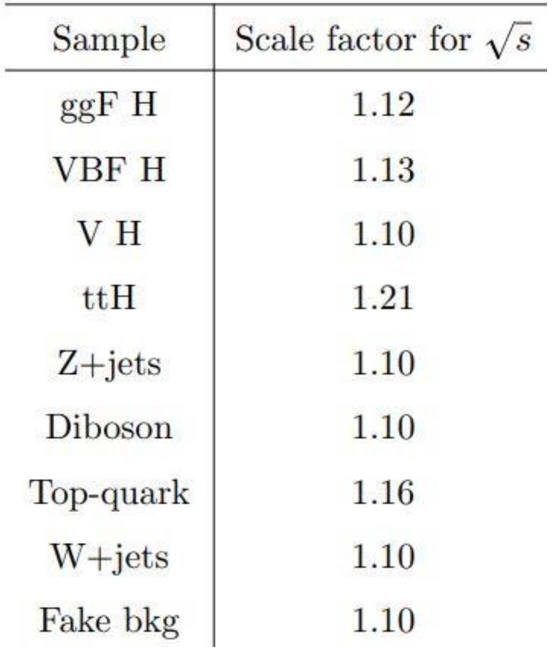
	Obs. (exp.) upper limits at 95% C.L.
$\mathcal{B}(H \to e\tau)$	0.20% ($0.12^{+0.05}_{-0.03}$%)
$\mathcal{B}(H \to \mu\tau)$	0.18% ($0.09^{+0.03}_{-0.02}$%)



- Results is found to be compatible with the SM prediction within 2.1σ

HL-LHC Extrapolation Procedure

- Run-2 analysis samples were used to **extrapolate** to the predictions for the HL-LHC.
- **Scaling** for the differences in:
  - **Integrated luminosity**: scale factor = $3000/138$ = 21.74
  - **Centre-of-mass energy** $\sqrt{s}$: scale factors are applied to different samples according to their expected cross section.
- **Statistical uncertainties:**
  1. Stat. unc. from MC samples are expected to become negligible (scale by 0). Data-driven backgrounds uncertainties scaled by $1/\sqrt{3000/138} = 0.21$.
  2. **Alternative approach**: All samples uncertainties scaled by 0.21.
- **Systematic uncertainties** are scaled by different scale factors according to the sources.
- Performed **statistical analysis** following the same method in Run-2 analysis.
  - Results obtained using **Asimov dataset**.



Sample	Scale factor for $\sqrt{s}$
ggF H	1.12
VBF H	1.13
V H	1.10
ttH	1.21
Z+jets	1.10
Diboson	1.10
Top-quark	1.16
W+jets	1.10
Fake bkg	1.10