Search for Lepton-Flavour-Violating Decays of the Higgs Boson

Kieran Amos
on behalf of the ATLAS Collaboration

Higgs 2022

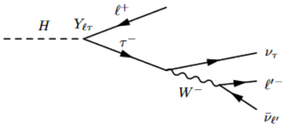
09/11/2022

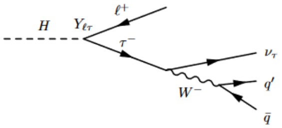
Motivation

- Lepton flavour conserved in the SM.
- Lepton Flavour Violation (LFV) is evident in neutrino oscillations. Can LFV processes also occur in charged lepton sector?
- Higgs boson LFV decays expected in several SM extensions (SUSY, 2HDM, composite Higgs...).
- This talk is on the results of the ATLAS search for $H \to e\tau$ and $H \to \mu\tau$ decays at $138\,\text{fb}^{-1}$. Link to Conf note.

Analysis Introduction

- Search for evidence of Higgs boson decay with charged lepton flavour violation (LFV) with two separate signals, $H \to e\tau$ and $H \to \mu\tau$ and two separate analysis methods, **MC-template** and **Symmetry**.
- All major production modes are considered as signal (ggH, VBF, WH, ZH)





- **Lep-Lep** Final states $e\tau_\mu$ and $\mu\tau_e$ with one electron and one muon.
- Ordering based on p_T of leptons in Higgs frame (new to this analysis).
- Two different methods:
  - **MC-template method**
  - **Symmetry method**

- **Lep-Had** Final states $e\tau_{\text{had}}$ and $\mu\tau_{\text{had}}$ with one lepton and one hadronic tau.
- One analysis method:
  - **MC-template method**

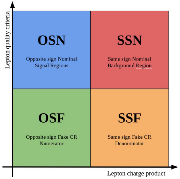
Event selection

Lep-Lep	Lep-Had
Baseline	
1e and 1μ, Opposite sign	1e or 1μ and
No tau-had	1tau-had with Opposite sign
No b-jets	
VBF	
$N_{\text{jets}} \geq 2$	
Non-VBF	
fail VBF selection	

- See Backup for detailed cuts.
- VBF region enhances sensitivity to VBFH production mode using jet kinematics.
- 2 Final states × 2 regions = 4 SRs per analysis method.
- Additional CRs dependent on the analysis (see next slides).

Background Estimation - Lep-Lep MC Template Method

- **Top** and $\boldsymbol{Z \to \tau\tau}$: Normalisation factors (NF) extracted separately for VBF and non-VBF from 1-bin CRs.
- CRs included in fit to constrain background yields.
- 2-POI CRs shared for $e\tau$ and $\mu\tau$.
- $\boldsymbol{Z \to \mu\mu}$: NF from CR at pre-fit level together with norm. unc.
- **Diboson**: Dedicated Validation region.
- **SM** $\boldsymbol{H \to \tau\tau}$ and $\boldsymbol{H \to WW}$ from MC.
- **Fakes**: Data-driven estimate for $j \to \ell$, $\gamma \to e$, $\tau_{\text{had}} \to \ell$.

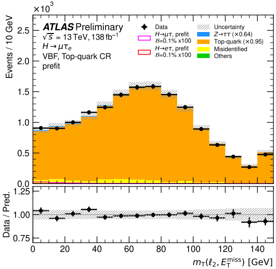


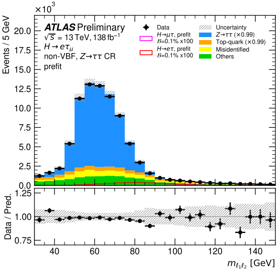
OSN
Opposite sign Nominal
Signal Regions

SSN
Same sign Nominal
Background Region

OSF
Opposite sign Fake CR
Numerator

SSF
Same sign Fake CR
Denominator





Background Estimation - Lep-Had MC Template Method

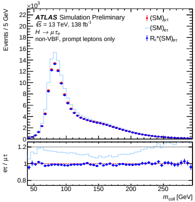
- $\boldsymbol{Z \rightarrow \tau\tau}$: different NF for VBF and non-VBF in the fit.
- **Top**: NF shared with lep-lep for categories where the MC-template is used for lep-lep. Otherwise Top normalization fixed from MC and corresponding theory uncertainties included.
- In 2-POI fit NFs are common for $e\tau_{\text{had}}$ or $\mu\tau_{\text{had}}$.
- $\boldsymbol{Z \rightarrow \mu\mu}$: uncertainties on the normalisation are extracted from dedicated VR.
- **Diboson, SM Higgs, others**: estimated from MC.
- **Fakes**: estimate of $j \rightarrow \tau_{\text{had}}$ with fake-factor. Main sources are W+jets and multijets.

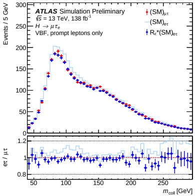
Background Estimation - Symmetry Lep-Lep Method

- Data-driven method where main backgrounds in each channel are estimated using data yields from the other channel.
  - SM processes are symmetric w.r.t. prompt $e \leftrightarrow \mu$ exchange.
  - LFV H decays break this symmetry.
  - Data of each of the two channels ($e\tau_\mu$ and $\mu\tau_e$) can serve as background prediction for the other channel.
  - The Symmetry method measures the difference of LFV signal strengths: $(\mathcal{B}(H \to \mu\tau) - \mathcal{B}(H \to e\tau))$. If one of the signal is assumed to be zero, then it becomes an absolute measurement.

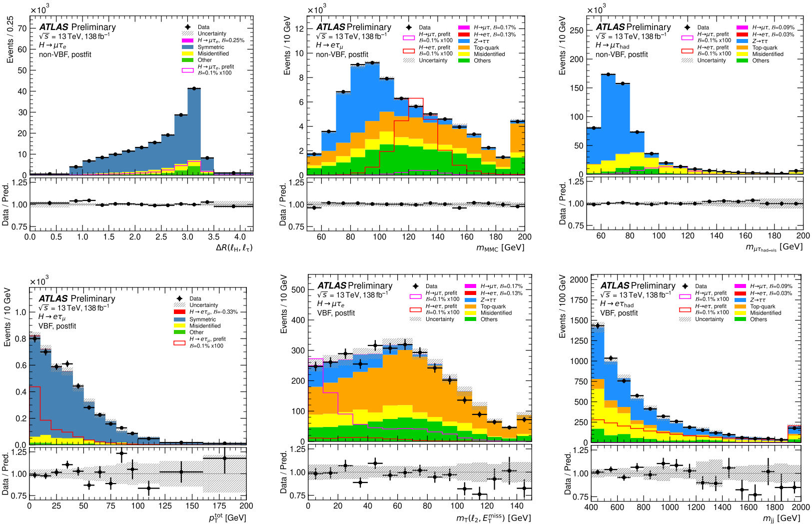
Detector effects induce asymmetries, contribute differently to each channel:

- Misidentified background events
- Different efficiencies for e and μ in reconstruction, identification...





Signal Region



non-VBF

VBF

Symmetry Lep-Lep

MC Lep-Lep

MC Lep-Had

MVA Strategy

- MVAs used to enhance sensitivity. Scores used as final discriminant.
- Symmetry Lep-Lep: NNs trained with Keras.
- MC-Template: BDTs trained with TMVA.
- Separate training for non-VBF and VBF and for $e\tau_{\text{had}}$ and $\mu\tau_{\text{had}}$ but combined across $e\tau_{\mu}$, $\mu\tau_{e}$.

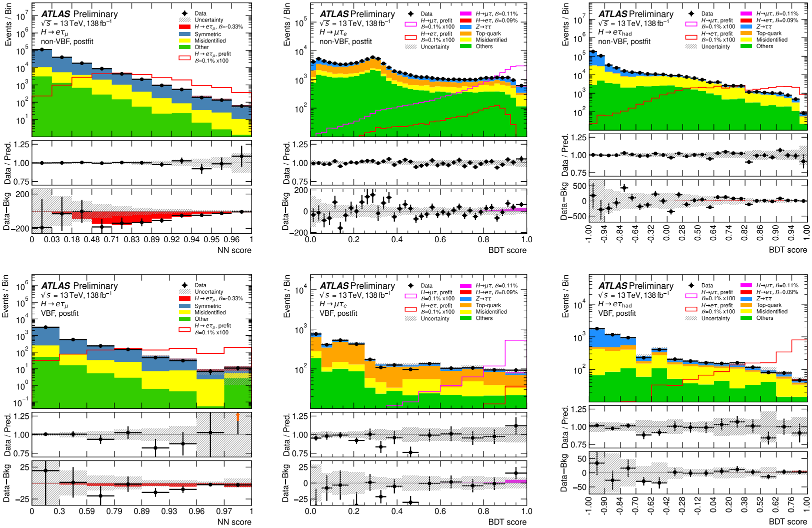
In general: Each MVA linear combination of individual MVAs trained on selected backgrounds.
Example (MC Lep-Lep):

- LFV vs $Z\tau\tau$+$H\tau\tau$+$Z\ell\ell$
- LFV vs Top+VV+HWW
- LFV vs Fakes

Selection of MVA Distributions

non-VBF

VBF



Symmetry Lep-Lep

MC Lep-Lep

MC Lep-Had

Statistical Analysis Overview

- MVA output used as final discriminant for signal strength μ extraction.
- Two different signal parametrisations:

Method	Channel	Category	Region	1 POI fit	2 POI fit
MC-template	$\ell\tau_{\ell'}$	*non-VBF*	SR	✓	✓
			$Z \to \tau\tau$ CR	✓	✓
			Top-quark CR	✓	✓
		VBF	SR		✓
			$Z \to \tau\tau$ CR		✓
			Top-quark CR		✓
MC-template	$\ell\tau_{\mathrm{had}}$	*non-VBF*	SR	✓	✓
		VBF	SR	✓	✓
Symmetry	$\ell\tau_{\ell'}$	*non-VBF*	SR		
		VBF	SR	✓	

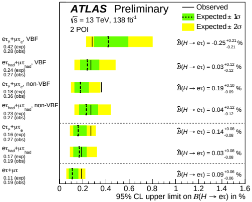
1-POI

- Independent fits in $e\tau$ and $\mu\tau$ channels (for example set $\mathcal{B}(H \to e\tau) = 0$ when measuring $\mathcal{B}(H \to \mu\tau)$)
- non-VBF: MC-LepLep + MC-LepHad
- VBF: Symmetry-LepLep + MC-LepHad

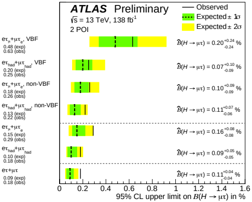
2-POI

- No assumption needed on branching ratios.
- The two signals are fitted simultaneously in the two channels
- non-VBF + VBF: MC-LepLep + MC-LepHad

2-POI results

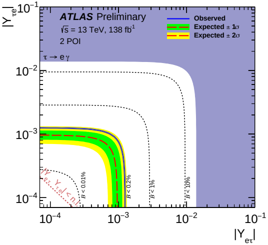


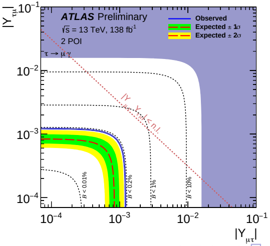
$\mathcal{B}(H \to e\tau)$

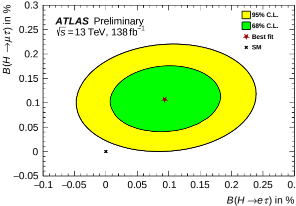


$\mathcal{B}(H \to \mu\tau)$

- Observed (expected) limits at 95% CL are 0.19(0.11)% and 0.18(0.09)%) for $\mathcal{B}(H \to e\tau)$ and $\mathcal{B}(H \to \mu\tau)$ respectively
- $\Rightarrow$ 1.6σ excess for $\mathcal{B}(H \to e\tau)$ and 2.5σ for $\mathcal{B}(H \to \mu\tau)$
- $\Rightarrow$ Compatible with SM branching value of 0 to 2.2σ







Conclusion

- Presented the results of a search for $H \rightarrow e\tau$ and $H \rightarrow \mu\tau$ decays with the ATLAS detector at $138\,\mathrm{fb}^{-1}$
- Two categories, Lep-Lep and Lep-had and two regions, VBF-enhanced and non-VBF
- Two different analysis methods: MC-Template and Symmetry(Lep-Lep only)
- Results quoted as 1-POI fits for independent $\mathcal{B}(H \rightarrow e\tau)$ and $\mathcal{B}(H \rightarrow \mu\tau)$ measurements and combined 2-POI measurement.
- Comparison to previous $36\,\mathrm{fb}^{-1}$ analysis:
  - **Observed Limits** improved by a factor of 2.4(1.5) for $H \rightarrow e\tau (H \rightarrow \mu\tau)$
  - **Expected Limits** improved by a factor of 3.1(4.1) for $H \rightarrow e\tau (H \rightarrow \mu\tau)$

Backup

Detailed Event selection

Selection	$\ell\tau_{\ell'}$	$\ell\tau_{\text{had}}$
Baseline	exactly $1e$ and 1μ, OS	exactly 1ℓ and $1\tau_{\text{had-vis}}$, OS
	τ_{had}-veto	τ_{had} Tight ID
		Medium eBDT ($e\tau_{\text{had}}$)
	b-veto	b-veto
	$p_{\text{T}}^{\ell_1} > 45$ (35) GeV MC-template (Symmetry method)	$p_{\text{T}}^{\ell} > 27.3$ GeV
	$p_{\text{T}}^{\ell_2} > 15$ GeV	$p_{\text{T}}^{\tau_{\text{had-vis}}} > 25$ GeV, $\lvert\eta^{\tau_{\text{had-vis}}}\rvert < 2.4$
	$30\,\text{GeV} < m_{\ell_1\ell_2} < 150\,\text{GeV}$	$\sum\limits_{i=\ell,\tau_{\text{had-vis}}} \cos\Delta\phi(i, E_{\text{T}}^{\text{miss}}) > -0.35$
	$0.2 < p_{\text{T}}^{\text{track}}(\ell_2 = e)/p_{\text{T}}^{\text{cluster}}(\ell_2 = e) < 1.25$ (MC-template)	$\lvert\Delta\eta(\ell, \tau_{\text{had-vis}})\rvert < 2$
	track d_0 significance requirement (see text)	
	$\lvert z_0 \sin\theta\rvert < 0.5$ mm	
VBF	*Baseline*	
	≥ 2 jets, $p_{\text{T}}^{\text{j}_1} > 40$ GeV, $p_{\text{T}}^{\text{j}_2} > 30$ GeV	
	$\lvert\Delta\eta_{\text{jj}}\rvert > 3$, $m_{\text{jj}} > 400$ GeV	
non-VBF	*Baseline* plus fail *VBF* categorisation	
	–	veto events if
	–	$90 < m_{\text{vis}}(e, \tau_{\text{had-vis}}) < 100$ GeV

Uncertainty Sources

1 POI Source of uncertainty	Impact ($\times 10^2$) on observed $\hat{\mathcal{B}}(H \to e\tau)$	 $\hat{\mathcal{B}}(H \to \mu\tau)$
Flavour tagging	0.010	0.003
Misidentified background ($\ell\tau_{\text{had}}$)	0.021	0.015
Misidentified background ($\ell\tau_{\ell'}$)	0.029	0.016
Jet and $E_{\text{T}}^{\text{miss}}$	0.011	0.010
Electrons and muons	0.003	0.005
Luminosity	0.006	0.006
Hadronic τ decays	0.009	0.009
Theory (signal)	0.008	0.006
Theory ($Z+$jets processes)	0.009	0.011
Theory (top-quark processes)	0.003	0.003
$Z \to \ell\ell$ normalisation	0.002	0.006
Symmetric background estimate	0.002	0.001
Background sample size	0.042	0.023
Total systematic uncertainty	0.053	0.038
Data sample size	0.030	0.028
Total	0.061	0.047

2 POI Source of uncertainty	Impact ($\times 10^2$) on observed $\hat{\mathcal{B}}(H \to e\tau)$	 $\hat{\mathcal{B}}(H \to \mu\tau)$
Flavour tagging	0.007	0.003
Misidentified background ($e\tau_{\text{had}}$)	0.021	0.003
Misidentified background ($e\tau_{\mu}$)	0.058	0.003
Misidentified background ($\mu\tau_{\text{had}}$)	0.006	0.015
Misidentified background ($\mu\tau_{e}$)	0.009	0.011
Jet and $E_{\text{T}}^{\text{miss}}$	0.012	0.009
Electrons and muons	0.013	0.005
Luminosity	0.007	0.005
Hadronic τ decays	0.009	0.009
Theory (signal)	0.007	0.007
Theory ($Z+$jets processes)	0.007	0.009
$Z \to \ell\ell$ normalisation ($e\tau$)	< 0.001	< 0.001
$Z \to \ell\ell$ normalisation ($\mu\tau$)	0.002	0.007
Background sample size	0.037	0.023
Total systematic uncertainty	0.051	0.036
Data sample size	0.030	0.027
Total	0.059	0.045

MVA Input Variables

Variable	$\ell\tau_{\text{had}}$		$\ell\tau_{\ell'}$ MC-template		$\ell\tau_{\ell'}$ Symmetry	
	non-VBF	*VBF*	*non-VBF*	*VBF*	*non-VBF*	*VBF*
m_{coll}	✓	✓	✓	✓	✓	✓
m_{vis}	✓	✓	✓	✓	✓	✓
m_{MMC}			✓	✓	✓	✓
$m_{\text{T}}(\tau, E_{\text{T}}^{\text{miss}})$	✓	✓			✓	✓
$m_{\text{T}}(\ell_H, E_{\text{T}}^{\text{miss}})$	✓	✓			✓	✓
$m_{\text{T}}(\ell_1, E_{\text{T}}^{\text{miss}})$			✓	✓		
$m_{\text{T}}(\ell_2, E_{\text{T}}^{\text{miss}})$			✓	✓		
$E_{\text{T}}^{\text{miss}}$	✓	✓	✓	✓	✓	✓
$p_{\text{T}}(\ell_H)$	✓	✓				
$p_{\text{T}}(\tau)$	✓	✓				
$p_{\text{T}}(\ell_\tau)/p_{\text{T}}(\ell_1)$			✓	✓		
$p_{\text{T}}^{\text{rest}}(\ell_H)$			✓	✓	✓	✓
$p_{\text{T}}^{\text{rest}}(\ell_\tau)$			✓	✓	✓	✓
$p_{\text{T}}^{\text{tot}}$				✓	✓	✓
$p_{\text{T}}(\ell_H)/E_{\text{T}}^{\text{miss}}$					✓	✓
$p_{\text{T}}(\tau)/p_{\text{T}}(\ell_H)$					✓	✓
$\sum p_{\text{T}}$					✓	✓
$\Delta R(\ell_H, \tau)$	✓	✓	✓	✓	✓	✓
$\Delta\eta(\ell_H, \tau)$	✓	✓			✓	✓
$\Delta\phi(\ell_H, \tau)$	✓				✓	
$\Delta\phi(\ell_\tau, E_{\text{T}}^{\text{miss}})$			✓	✓	✓	✓
$\lvert\Delta\phi(\ell_H, E_{\text{T}}^{\text{miss}})\rvert - \lvert\Delta\phi(\tau, E_{\text{T}}^{\text{miss}})\rvert$	✓					
$\Delta\alpha$			✓	✓	✓	✓
$\Delta\Phi\left(\ell_H, E_{\text{T}}^{\text{miss}}\right)$		✓	✓	✓	✓	✓
$\Delta d_0\,(\ell_1, \ell_2)$			✓	✓	✓	
$\sigma_{d_0}^{\ell_\tau}$			✓	✓		
$\eta(\tau_{\text{had-vis}})$	✓	✓				
m_{jj}		✓		✓		✓
$N_{\text{jets}(p_{\text{T}}>30GeV)}$		✓				
$\lvert\Delta\eta_{\text{jj}}\rvert$		✓				✓
$\Delta R(\text{j}, \text{j})$				✓		✓
$\lvert\Delta\eta_{\text{jj}}\rvert \cdot \eta_{\text{j}_1} \cdot \eta_{\text{j}_2}$				✓		
$p_{\text{T}}(\text{j}_1)$						✓
$p_{\text{T}}(\text{j}_2)$						✓
$\Delta\phi(\text{j}_1, E_{\text{T}}^{\text{miss}})$				✓		✓
$\Delta\phi(\text{j}_2, E_{\text{T}}^{\text{miss}})$				✓		✓
η-centrality(ℓ_H)				✓		✓
η-centrality(ℓ_τ)				✓		✓

MVA Parameters

Region	Channel	`NTrees`	`MaxDepth`	`MinNodeSize`	`Shrinkage`
non-VBF	$e\tau_{\text{had}}$,$\mu\tau_{\text{had}}$	500	7	1	0.05
VBF	$e\tau_{\text{had}}$	300	10	1	0.01
	$\mu\tau_{\text{had}}$ BDT$_1$	300	8	1	0.009
	$\mu\tau_{\text{had}}$ BDT$_2$	300	6	1	0.0095
non-VBF, *VBF*	$\ell\tau_{\ell'}$	750	8	2.5	0.1

Hyperparameter	Value			
	non-VBF NN	$VBF_{Z\to\tau\tau}$ NN	$VBF_{\text{Top-quark}}$ NN	VBF_{misID} NN
# nodes in 1st layer	512	128	128	128
# hidden layers	2	4	3	4
# output layers	3	2	2	2
L2 weight reg. param.	0.000048	0.000292	0.000094	0.000356
leaky ReLU slope below 0	0.080537	0.019614	0.062515	0.084219
optimiser	SGD	Adam	Adam	Adam
learning rate	0.025810	0.000142	0.000215	0.003507
batch size	128	128	512	1024
epochs	100	100	100	100

MVA Strategy

Symmetry Lep-Lep

- NNs trained with Keras
- Separate training for non-VBF and VBF, but common for $e\tau_{\mu}$ and $\mu\tau_{e}$

non-VBF

1 Multiclassifier NN with 3 output nodes. Signal output node used for fit.

VBF

3 MVAs. Scores combined linearly.

- LFV vs. $Z\tau\tau$+$H\tau\tau$+MCfakes
- LFV vs. Top+VV+HWW
- LFV vs. Fakes

MC-Template Lep-Lep

- BDTs with TMVA
- Separate training for non-VBF and VBF but common for $e\tau_{\mu}$ and $\mu\tau_{e}$

non-VBF and VBF

3 BDTs, combined linearly.

- LFV vs $Z\tau\tau$+$H\tau\tau$+$Z\ell\ell$
- LFV vs Top+VV+HWW
- LFV vs Fakes

MC-Template Lep-Had

- BDTs with TMVA
- Separate training for non-VBF and VBF and for $e\tau_{\text{had}}$ and $\mu\tau_{\text{had}}$

non-VBF $e\tau_{\text{had}}$

3 BDTs, combined linearly.

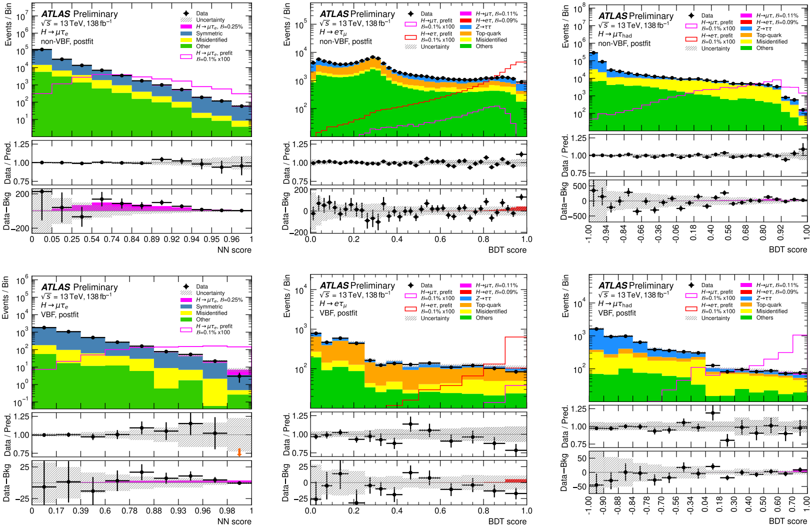
- LFV vs $Z\tau\tau$
- LFV vs Fakes
- LFV vs Rest of bkgs

VBF and non-VBF $\mu\tau_{\text{had}}$

2 BDTs, combined linearly for non-VBF $\mu\tau_{\text{had}}$ and quadraticaly for VBF

- LFV vs $Z\tau\tau$
- LFV vs Rest of bkgs

Additional MVA Distributions



non-VBF

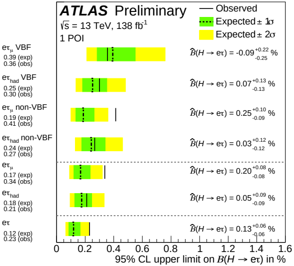
VBF

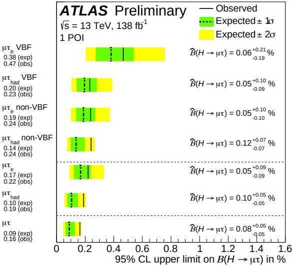
Symmetry Lep-Lep

MC Lep-Lep

MC Lep-Had

1-POI results





- Observed (expected) limits at 95% CL are 0.23(0.12)% and 0.16(0.09)%) for $\mathcal{B}(H \to e\tau)$ and $\mathcal{B}(H \to \mu\tau)$ respectively