GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Investigation of the CP properties of VBF Higgs production using the decay to a pair of tau leptons with the ATLAS detector

ATLAS-CONF-2019-050

Serhat Ördek
on behalf of the ATLAS Collaboration

Georg-August-Universität Göttingen

09.01.2020
Young Scientist Session, Epiphany 2020

GEFÖRDERT VOM
Bundesministerium
für Bildung
und Forschung

ATLAS
EXPERIMENT

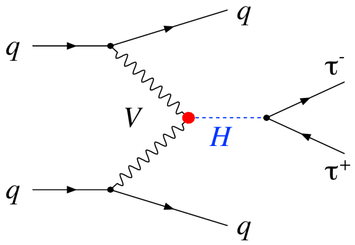
Motivation

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

- Baryon asymmetry observed in the universe
- Sakharov: Need CP violating processes
- But CP violation in Standard Model (SM) presumably not sufficient

$\longrightarrow$ Probe Higgs sector for additional sources of CP violation

- In SM: Higgs couplings are CP-even
- This analysis: Test for CP-odd contribution to otherwise SM-like HVV couplings
- Using VBF Higgs production process
- $H \to \tau\tau$: Good balance between S/B and available statistics



Effective Field Theory

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

- Extend SM Lagrangian with CP-odd dimension-6 operators[1]
- Four new CP-odd couplings after electroweak symmetry breaking:

$$\mathcal{L}_{eff} = \mathcal{L}_{SM} + \tilde{g}_{HAA} H\tilde{A}_{\mu\nu}A^{\mu\nu} + \tilde{g}_{HAZ} H\tilde{A}_{\mu\nu}Z^{\mu\nu} + \tilde{g}_{HZZ} H\tilde{Z}_{\mu\nu}Z^{\mu\nu} + \tilde{g}_{HWW} H\tilde{W}^{+}_{\mu\nu}W^{-\mu\nu}$$

- Two free parameters $\tilde{d}$ and $\tilde{d}_B$ due to $SU(2) \times U(1)$ symmetric ansatz

$$\tilde{g}_{HAA} = \frac{g}{2m_W}(\tilde{d}\sin^2\theta_W + \tilde{d}_B\cos^2\theta_W) \qquad \tilde{g}_{HAZ} = \frac{g}{2m_W}\sin 2\theta_W(\tilde{d} - \tilde{d}_B)$$

$$\tilde{g}_{HZZ} = \frac{g}{2m_W}(\tilde{d}\cos^2\theta_W + \tilde{d}_B\sin^2\theta_W) \qquad \tilde{g}_{HWW} = \frac{g}{m_W}\tilde{d}\,.$$

- Contributions from $\tilde{d}$ and $\tilde{d}_B$ indistinguishable $\Rightarrow$ set $\tilde{d} = \tilde{d}_B$
- Then $\tilde{g}_{HAA} = \tilde{g}_{HZZ} = \frac{1}{2}\tilde{g}_{HWW} = \frac{g}{2m_W}\tilde{d}$ and $\tilde{g}_{HAZ} = 0$

$\Rightarrow$ Strength of CP violation described by single parameter $\tilde{d}$

[1] V. Hankele et al: Phys. Rev. D 74 (2006).

Matrix element & CP test

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

$\tilde{d} \neq 0$ leads to additional term in VBF matrix element

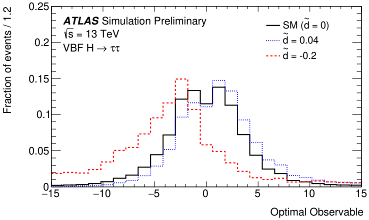
$$\mathcal{M} = \mathcal{M}_{\text{SM}} + \tilde{d}\ \mathcal{M}_{\text{CP-odd}}$$

$$\Rightarrow |\mathcal{M}|^2 = \underbrace{|\mathcal{M}_{\text{SM}}|^2}_{\text{CP-even}} + 2\tilde{d}\ \underbrace{\Re(\mathcal{M}^*_{\text{SM}}\mathcal{M}_{\text{CP-odd}})}_{\text{CP-odd, source of CP violation}} + \tilde{d}^2\ \underbrace{|\mathcal{M}_{\text{CP-odd}}|^2}_{\text{CP-even}}$$

- Measure CP-odd Optimal Observable[2]:

$$OO = \frac{\Re(\mathcal{M}^*_{\text{SM}}\mathcal{M}_{\text{CP-odd}})}{|\mathcal{M}_{\text{SM}}|^2}$$

- CP conserved $\Rightarrow \langle OO \rangle = 0$
  (neglecting possible rescattering effects)



Calculate matrix elements with HAWK[3]:

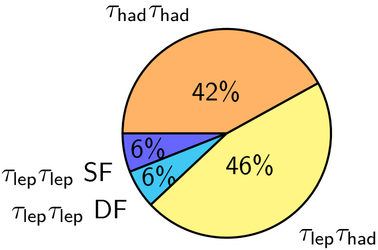
- Needs 4-vectors of 2 jets and Higgs
- For signal prediction with $\tilde{d} \neq 0$: ME-based reweighting $\left(w = \frac{|\mathcal{M}|^2}{|\mathcal{M}_{\text{SM}}|^2}\right)$

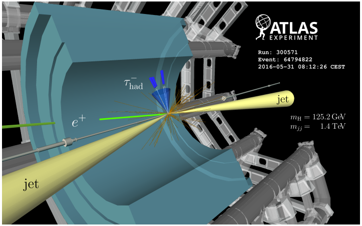
[2] D. Atwood & A.Soni; Phys Lett. D45 (1992), M. Davier et al.; Phys. Lett B306 (1993), M. Diehl & O. Nachtmann; Z. Phys. C62 (1994)

[3] A. Denner et al.; Comput. Phys. Commun. **195** (2015) 161

Analysis Overview

- Goal: Measure $\tilde{d}$ with Optimal Observable
- 2015 + 2016 data, $\int \mathscr{L}\, dt = 36.1\ \text{fb}^{-1}$
- Based on $H \to \tau\tau$ cross-section measurement, Phys. Rev. D **99**, 072001





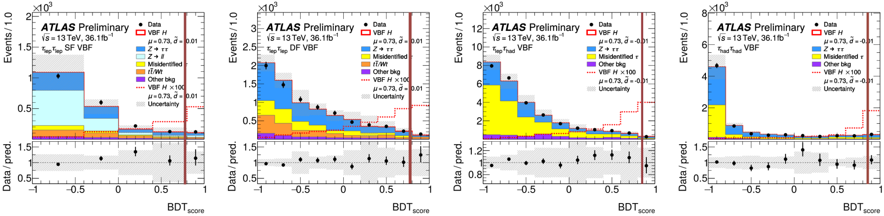
Phys. Rev. D **99**, 072001

VBF selection: Di-τ events with ≥ 2 jets, $m_{jj} > 300$ GeV and $\Delta\eta_{jj} > 3$

	Exp. Signal	Background	Ratio
$\tau_{\text{lep}}\tau_{\text{lep}}$ SF	14.4	1010	1.4%
$\tau_{\text{lep}}\tau_{\text{lep}}$ DF	26.0	1530	1.7%
$\tau_{\text{lep}}\tau_{\text{had}}$	54.0	5350	1.0%
$\tau_{\text{had}}\tau_{\text{had}}$	24.9	1510	1.6%

BDT Selection

Boosted Decision Trees (BDT) to find signal within VBF selection



Define signal region (SR) via cut on BDT score:

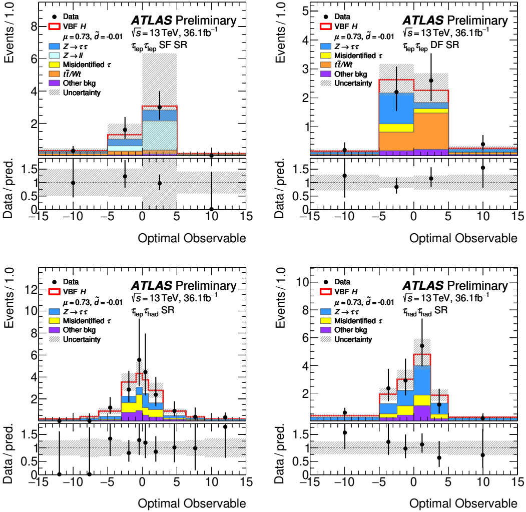
	Exp. Signal	Background	Ratio
$\tau_{lep}\tau_{lep}$ SF	4.5	23	0.2
$\tau_{lep}\tau_{lep}$ DF	6.9	23	0.3
$\tau_{lep}\tau_{had}$	16	19	0.9
$\tau_{had}\tau_{had}$	12	26	0.5

Main backgrounds:

- $Z \to \tau\tau$:
  Taken from simulation
  Normalized to data
- Misidentified τ:
  Estimated via data-driven techniques

Optimal Observable distributions

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN



- *OO* distributions in high-BDT SRs
- From data: model-independent test of CP violation

Channel	⟨Optimal Observable⟩
$\tau_{lep}\tau_{lep}$ SF	-0.54 ± 0.72
$\tau_{lep}\tau_{lep}$ DF	0.71 ± 0.81
$\tau_{lep}\tau_{had}$	0.74 ± 0.78
$\tau_{had}\tau_{had}$	-1.13 ± 0.65

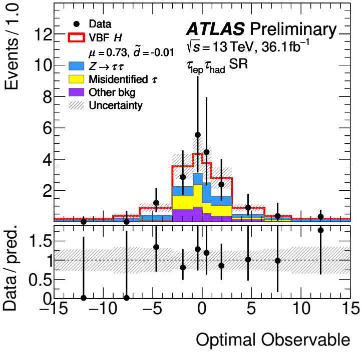
$\Rightarrow$ Consistent with 0

No sign of CP violation

For more quantitative statement on $\tilde{d}$: **Perform likelihood fits**

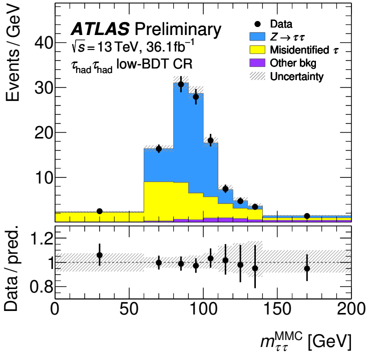
Fit model

High-BDT SR, one per channel:



Fit *OO* distribution
Sensitive to value of $\tilde{d}$

Low-BDT CR, one per channel:



Fit di-τ mass $m_{\tau\tau}^{\mathrm{MMC}}$
Constrain $Z \rightarrow \tau\tau$ normalization

- In $\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$: Fit event yields in additional $Z \rightarrow \ell\ell$ (SF only) and Top CRs
- Scan $\tilde{d}$ values for signal, find $\tilde{d}$ maximizing likelihood

Limits on $\tilde{d}$

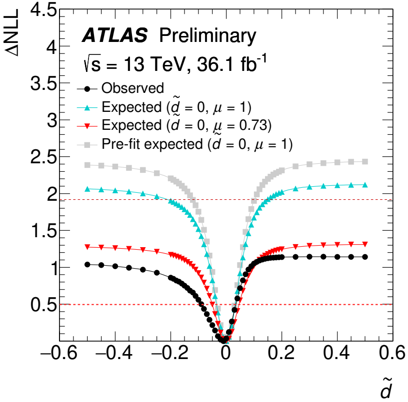
Observed:

- Signal strength $\mu = \frac{\sigma_{\text{Obs}}}{\sigma_{\text{SM}}} = 0.73$
- best-fit $\tilde{d} = -0.01$
- **68% interval: $\tilde{d} \in [-0.090, 0.035]$**

$\Rightarrow$ Fully consistent with SM

Expected with $\mu = 1$:
Based on pseudodata using best-fit uncertainties from CR-only fit

- 68% interval: $\tilde{d} \in [-0.035, 0.033]$
- 95% interval: $\tilde{d} \in [-0.21, 0.15]$

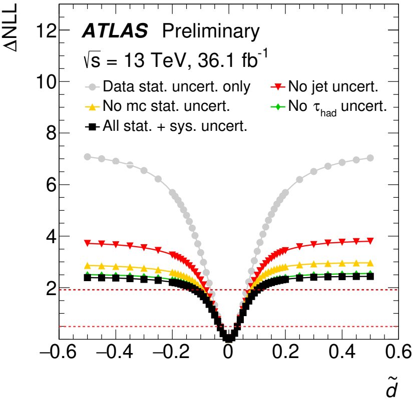


Expected with $\mu = 0.73$: Low event yields reduce sensitivity

Systematic uncertainties

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

- Fit to pseudodata generated with pre-fit uncertainties
- Set groups of uncertainties to constant in fit
- Group with biggest impact: Jet uncertainties
- Effect of ignoring τ_{had} or MC stat. uncertainties smaller
- Other uncertainties almost negligible



Conclusion

- Tested CP invariance of VBF production with Optimal Observable method
- Used VBF Higgs events in di-tau final states

- Means of OO in data consistent with zero
- Expected 95% CI on CP-violation inducing parameter $\tilde{d}$ is $[-0.21, 0.15]$
- Due to low measured signal strength, no observed 95% CI
- Observed (expected) 68% CI: $[-0.090, 0.035]$ $([-0.035, 0.033])$
- All consistent with SM expectation of zero, **no sign of CP violation found**

Additional Material

Construction of “expected” NLL curves

Likelihood fit: Find configuration of your fit parameters $\vec{\theta}$ that maximizes

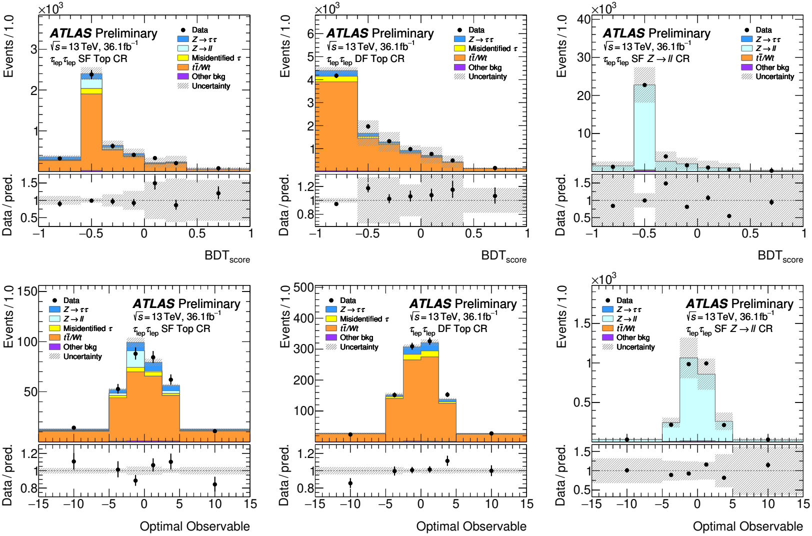
$$\mathcal{L}\left(\vec{\theta}|\vec{x}\right) = \prod_{\text{bins } i} \text{Poisson}\left(x_i|N_i^{\text{exp}}(\vec{\theta})\right) \prod_{j=1}^{\#\text{NPs}} \mathcal{C}_j(\theta_j)$$

given observed data $\vec{x}$ ($\mathcal{C}_j$: constraint fct, usually Gaussian of width 1 around 0)

1. Conduct CR-only fit to data, obtain pulls $\vec{\hat{\theta}}$
2. Construct “Asimov” dataset $\vec{x}_{\text{exp}}$ with NPs set to $\vec{\hat{\theta}}$
3. Construct new likelihood function $\mathcal{L}'$ where center of $\mathcal{C}_j$ is always $\hat{\theta}_j$
4. Maximize $\mathcal{L}'(\vec{\theta}|\vec{x}_{\text{exp}})$, should return pulls $\vec{\hat{\theta}}$ again

BDT and OO in Top and $Z \to \ell\ell$ CRs

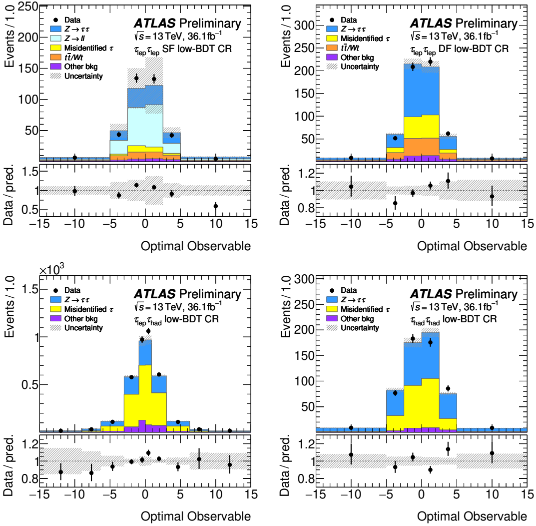
GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN



- Distributions in 1-bin $Z \to \ell\ell$ and $t\bar{t}/Wt$ CRs (not part of fit)

Modelling tests in CRs (2)

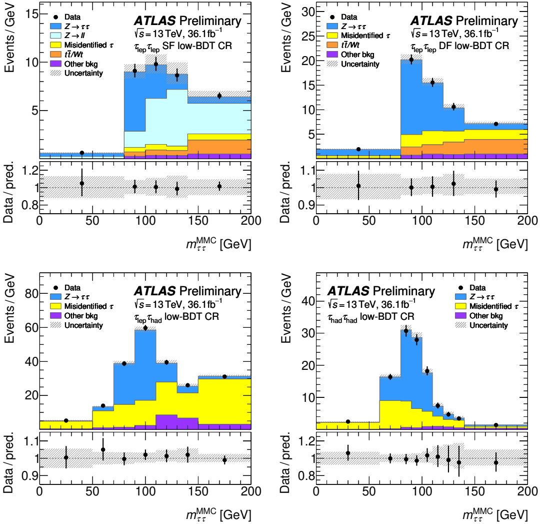
GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN



Optimal Observable modelling in low-BDT CR (not part of fit)

Low-BDT CR $m_{\tau\tau}^{\mathrm{MMC}}$ distributions

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN



MC samples

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

Process	Matrix element (alternative)	PDF set	UEPS model (alternative model)	Prediction order for total cross-section
VBF H	Powheg-Box v2 [55–59]	PDF4LHC15 NLO [60]	Pythia 8 [61] (Herwig 7 [63, 64])	approx. NNLO QCD + NLO EW [43, 44, 62]
ggF H	Powheg-Box v2 NNLOPS [69–71]	PDF4LHC15 NNLO	Pythia 8 (Herwig 7)	N^3LO QCD + NLO EW [65–68]
VH	Powheg-Box v2 [72]	PDF4LHC15 NLO	Pythia 8	$qq/qg \to VH$: NNLO QCD + NLO EW [73, 74] $gg \to ZH$: NLO + NLL QCD [75, 76]
$t\bar{t}H$	MG5_aMC@NLO 2.2.2 [77, 78]	NNPDF3.0LO [79]	Pythia 8	NLO QCD + NLO EW [80–85]
W/Z+jets	Sherpa 2.2.1 [86] (MG5_aMC@NLO 2.2.2)	NNPDF3.0NNLO	Sherpa 2.2.1 [87] (Pythia 8)	NNLO [88, 89]
Electroweak W/Zjj	Sherpa 2.2.1	NNPDF3.0NNLO	Sherpa 2.2.1	LO
$VV/V\gamma^*$	Sherpa 2.2.1	NNPDF3.0NNLO	Sherpa 2.2.1	NLO
$t\bar{t}$	Powheg-Box v2 [90]	CT10 [46]	Pythia 6.428 [91]	NNLO+NNLL [92]
Wt	Powheg-Box v1 [93]	CT10	Pythia 6.428	NLO [93]

Event selection

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

Channel	$\tau_{\text{lep}}\tau_{\text{lep}}$ SF	$\tau_{\text{lep}}\tau_{\text{lep}}$ DF	$\tau_{\text{lep}}\tau_{\text{had}}$	$\tau_{\text{had}}\tau_{\text{had}}$
Preselection	Two isolated τ-lepton decay candidates with opposite electric charge			
	$p_{\text{T}}^{\tau_1} > 19^*/15^*$ GeV (μ/e)	$p_{\text{T}}^{e} > 18$ GeV	$p_{\text{T}}^{\tau_{\text{had}}} > 30$ GeV	$p_{\text{T}}^{\tau_1} > 40$ GeV
	$p_{\text{T}}^{\tau_2} > 10/15^*$ GeV (μ/e)	$p_{\text{T}}^{\mu} > 14$ GeV	$p_{\text{T}}^{\tau_{\text{lep}}} > 21^*$ GeV	$p_{\text{T}}^{\tau_2} > 30$ GeV
	$m_{\tau\tau}^{\text{coll}} > m_Z - 25$ GeV		$m_{\text{T}} < 70$ GeV	$0.8 < \Delta R_{\tau\tau} < 2.5$
	$30 < m_{\ell\ell} < 75$ GeV	$30 < m_{\ell\ell} < 100$ GeV		$\lvert\Delta\eta_{\tau\tau}\rvert < 1.5$
	$E_{\text{T}}^{\text{miss}} > 55$ GeV	$E_{\text{T}}^{\text{miss}} > 20$ GeV		$E_{\text{T}}^{\text{miss}} > 20$ GeV
	$E_{\text{T}}^{\text{miss, hard}} > 55$ GeV			
	$N_{b\text{-jets}} = 0$			
VBF topology	$N_{\text{jets}} \geq 2$, $p_{\text{T}}^{j_2} > 30$ GeV, $m_{jj} > 300$ GeV, $\lvert\Delta\eta_{jj}\rvert > 3$			
	$p_{\text{T}}^{j_1} > 40$ GeV			$p_{\text{T}}^{j_1} > 70$ GeV, $\lvert\eta_{j_1}\rvert < 3.2$
BDT input variables	$m_{\tau\tau}^{\text{MMC}}$, m_{jj}, $\Delta R_{\tau\tau}$, $C_{jj}(\tau_1)$, $C_{jj}(\tau_2)$, $p_{\text{T}}^{\text{tot}}$			
	$m_{\tau\tau}^{\text{vis}}$, $m_{\text{T}}^{\tau_1, E_{\text{T}}^{\text{miss}}}$, $p_{\text{T}}^{j_3}$		$C(\phi^{\text{miss}})/\sqrt{2}$	
	$\Delta\phi_{\tau\tau}$	$E_{\text{T}}^{\text{miss}}/p_{\text{T}}^{\tau_1}$, $E_{\text{T}}^{\text{miss}}/p_{\text{T}}^{\tau_2}$	$m_{\tau\tau}^{\text{vis}}$, $\lvert\Delta\eta_{\tau\tau}\rvert$	$p_{\text{T}}^{\tau\tau E_{\text{T}}^{\text{miss}}}$, $\lvert\Delta\eta_{\tau\tau}\rvert$
Signal region	$\text{BDT}_{\text{score}} > 0.78$		$\text{BDT}_{\text{score}} > 0.86$	$\text{BDT}_{\text{score}} > 0.87$

Post-fit yields

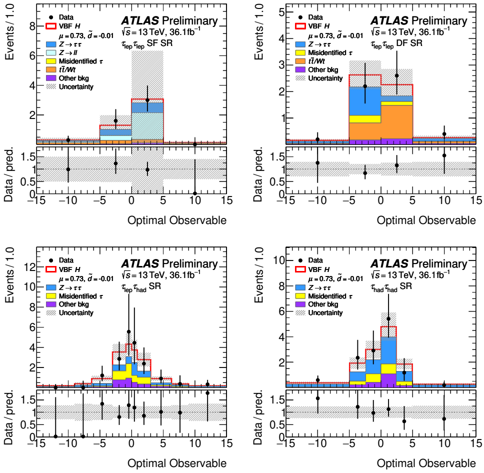
Process	VBFH	$Z \to \tau\tau$	$Z \to \ell\ell$ ($\tau_{\text{lep}}\tau_{\text{lep}}$ SF)	$t\bar{t}/Wt$ ($\tau_{\text{lep}}\tau_{\text{lep}}$)	Fake $\tau_{\text{had-vis}}$ ($\tau_{\text{had}}\tau_{\text{had}}$)
NF	0.73 ± 0.47	0.93 ± 0.08	1.0 ± 0.4	1.16 ± 0.06	0.99 ± 0.09

Process	$\tau_{\text{lep}}\tau_{\text{lep}}$ SF	$\tau_{\text{lep}}\tau_{\text{lep}}$ DF
Data	26	30
VBF $H \to \tau\tau/WW$ ($\mu = 0.73$, $\tilde{d} = -0.01$)	3.3 ± 2.1	5.1 ± 3.1
VBF $H \to \tau\tau/WW$ ($\mu = 1$, $\tilde{d} = 0$)	4.5 ± 2.9	6.9 ± 4.4
$Z \to \tau\tau$	6.6 ± 3.7	8.2 ± 3.8
Fake lepton	0.02 ± 0.20	2.3 ± 0.7
$t\bar{t}$ + single-top	3.8 ± 2.3	10.6 ± 5.5
$Z \to \ell\ell$	11 ± 18	1.8 ± 1.1
Diboson	0.70 ± 0.59	
ggF H / VH / $t\bar{t}H$, $H \to \tau\tau/WW$	0.49 ± 0.48	0.70 ± 0.30
Sum of backgrounds	23 ± 17	23.6 ± 6.1

Process	$\tau_{\text{lep}}\tau_{\text{had}}$	$\tau_{\text{had}}\tau_{\text{had}}$
Data	30	37
VBF $H \to \tau\tau$ ($\mu = 0.73$, $\tilde{d} = -0.01$)	11.8 ± 7.4	8.9 ± 5.6
VBF $H \to \tau\tau$ ($\mu = 1$, $\tilde{d} = 0$)	16 ± 10	12.3 ± 7.7
$Z \to \tau\tau$	7.8 ± 3.5	15.5 ± 5.2
Fake lepton/τ	6.2 ± 1.0	5.4 ± 2.7
ggF H / VH / $t\bar{t}H$, $H \to \tau\tau$	2.1 ± 1.5	2.8 ± 1.4
Other backgrounds	2.8 ± 3.1	2.3 ± 0.8
Sum of backgrounds	19.0 ± 5.5	26.0 ± 6.6

- Channel with highest signal purity: $\tau_{\text{lep}}\tau_{\text{had}}$, followed by $\tau_{\text{had}}\tau_{\text{had}}$

Sub-channel fits



- Using only event yields in all but one SR
- Minima of resulting curves related to high Data - Bkg regions in SR *OO* distributions