UPPSALA UNIVERSITET

ATLAS and CMS $HH \to bb\tau\tau$: comparison of selections and fitted variables

Arnaud Ferrari (Uppsala University) on behalf of the ATLAS and CMS collaborations

HH workshop, Fermilab (USA), 3-8 September 2018

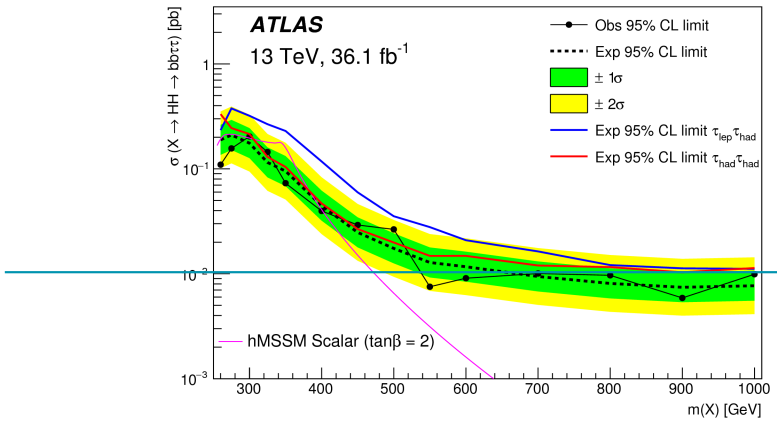
The $HH \to bb\tau\tau$ search channel in two slides

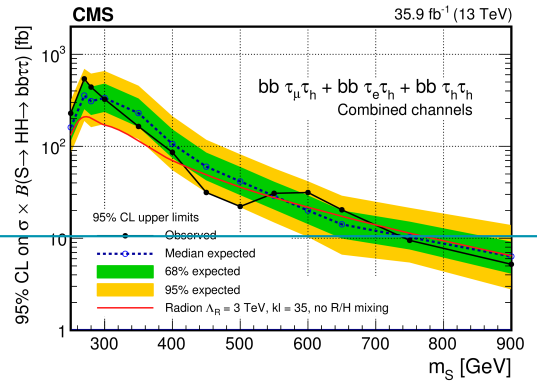
- Third accessible branching fraction (7.4%), but one of the most sensitive channels.
- Same sub-decay channels considered by both ATLAS and CMS: $\tau_\ell\tau_h$ and $\tau_h\tau_h$.
- Main backgrounds: $t\bar{t}$, Z+jets and multi-jet events (see Francesco Brivio's talk for details).
- Results for the non-resonant HH production:
  - ATLAS: Expected (observed) 95% CL limit at 14.8 (12.7) times σ_{SM}: arXiv:1808.00336 [hep-ex].
  - CMS: Expected (observed) 95% CL limit at 25 (30) times σ_{SM}: Phys. Lett. B 778 (2018) 101-127.

Where does the 1.7 times better sensitivity in ATLAS comes from? What can we learn by comparing event selections and fitted kinematic variables?

The $HH \to bb\tau\tau$ search channel in two slides

ATLAS also has a better sensitivity than CMS for the resonant HH production mode ($m \lesssim 1$ TeV):





For instance, for a heavy scalar mass of 400 GeV:

- Combined: ATLAS expected limits are 2.4 times better than CMS (44.5 vs 106 fb).
- $\tau_\ell\tau_h$: ATLAS / CMS $\simeq 1.6$
- $\tau_h\tau_h$: ATLAS / CMS $\simeq 2.7$

Why is ATLAS better sensitivity driven by the $\tau_h\tau_h$ channel?
Why is it more significant in the resonant production mode?

Comparison of trigger strategies (1)

ATLAS Trigger Strategy:

- $\tau_\ell\tau_h$ channel:

* Single-lepton trigger (SLT) $\Rightarrow$
$p_T^\ell > 25$-27 GeV & $p_T^\tau > 20$ GeV.

* If !SLT, lepton-tau trigger (LTT) $\Rightarrow$
$p_T^{e/\mu} > 18/15$ GeV & $p_T^\tau > 30$ GeV.

- $\tau_h\tau_h$ channel:

* Single (STT) or di-τ (DTT) trigger
$\to p_T^{\tau_1} > 40$-180 GeV
$\to p_T^{\tau_2} > 20/30$ GeV for STT/DTT.

- $p_T^{j_1}$ is raised (45$\to$80 GeV) when using LTT or DTT (level-1 trigger jet) while $p_T^{j_2} > 20$ GeV.

CMS Trigger Strategy:

- $\tau_\ell\tau_h$ channel:

* Single-lepton trigger $\Rightarrow$
$p_T^{e/\mu} > 27/23$ GeV & $p_T^\tau > 20$ GeV.

- $\tau_h\tau_h$ channel:

* Di-τ trigger
$\to p_T^{\tau_{1,2}} > 45$ GeV

- $p_T^{j_{1,2}} > 20$ GeV.

[See Agni Bethani's talk]

Comparison of trigger strategies (2)

Did the complicated ATLAS trigger strategy pay off in terms of increased sensitivity?

- $\tau_\ell\tau_h$ channel: the gain in sensitivity from adding the LTT sub-channel is only marginal (3%).

		Observed	-2σ	-1σ	Expected	$+1\sigma$	$+2\sigma$
$\tau_{\text{lep}}\tau_{\text{had}}$ (SLT)	$\sigma(HH \to bb\tau\tau)$ [fb]	52	38.4	52	72	100	134
	$\sigma/\sigma_{\text{SM}}$	21.3	15.7	21.1	29.3	40.8	55
$\tau_{\text{lep}}\tau_{\text{had}}$ (LTT)	$\sigma(HH \to bb\tau\tau)$ [fb]	326	123	165	229	319	428
	$\sigma/\sigma_{\text{SM}}$	134	50	68	94	131	175
$\tau_{\text{lep}}\tau_{\text{had}}$ Combined	$\sigma(HH \to bb\tau\tau)$ [fb]	57	37.2	49.9	69	96	129
	$\sigma/\sigma_{\text{SM}}$	23.5	15.2	20.5	28.4	39.5	53

- $\tau_h\tau_h$ channel: We have not quantified the impact of using an OR of two triggers with respect to DTT only.

Comparison of $\tau_\ell\tau_h$ event selections

ATLAS $\tau_\ell\tau_h$ Event Selections:

- SLT and LTT triggers
- 1 lepton and 1 medium τ_h (opposite-sign)
- p_T^ℓ based on SLT/LTT thresholds
- $p_T^\tau > 20/30$ GeV
- $\geq$2 jets with $p_T^j > 45\text{-}80\ (20)$ GeV

$m_{\tau\tau} > 60$ GeV based on the Missing Mass Calculator.

Only events with 2 b-tagged jets are kept in the SR.

CMS $\tau_\ell\tau_h$ Event Selections:

- SLT triggers
- 1 lepton and 1 medium τ_h (opposite-sign)
- p_T^ℓ based on SLT thresholds
- $p_T^\tau > 20$ GeV
- $\geq$2 jets with $p_T^j > 20$ GeV

$m_{\tau\tau}$ computed with SVfit

$$\frac{(m_{\tau\tau} - 116\,\text{GeV})^2}{(35\,\text{GeV})^2} + \frac{(m_{\text{bb}} - 111\,\text{GeV})^2}{(45\,\text{GeV})^2} < 1$$

Two SR categories: $\geq$2b and 1b1j.

Comparison of $\tau_h\tau_h$ event selections

ATLAS $\tau_h\tau_h$ Event Selections:

- STT or DTT triggers
- 2 medium τ_h's (opposite-sign)
- $p_T^{\tau_1} >$ 40-180 GeV
- $p_T^{\tau_2} > 20/30$ GeV
- $\geq$2 jets with $p_T^j >$ 45-80 (20) GeV

$m_{\tau\tau} > 60$ GeV based on the Missing Mass Calculator.

Only events with 2 b-tagged jets are kept in the SR.

CMS $\tau_h\tau_h$ Event Selections:

- DTT triggers
- 2 medium τ_h's (opposite-sign)
- $p_T^{\tau_1} >$ 45 GeV
- $p_T^{\tau_2} >$ 45 GeV
- $\geq$2 jets with $p_T^j >$ 20 GeV

$m_{\tau\tau}$ computed with SVfit

$$\frac{(m_{\tau\tau} - 116\,\text{GeV})^2}{(35\,\text{GeV})^2} + \frac{(m_{\text{bb}} - 111\,\text{GeV})^2}{(45\,\text{GeV})^2} < 1$$

Two SR categories: $\geq$2b and 1b1j.

Comparison of event selections: summary

1. CMS has also studied boosted topologies for high-mass resonances, while ATLAS has only searched for resolved categories.

   In the CMS boosted category:

   - there is at least one single AK8 jet (instead of two AK4 jets), with $m_J > 30$ GeV, $p_T^J > 170$ GeV and two sub-jets matched to the AK4 jets;
   - $80 < m_{\tau\tau}$ (GeV) < 152 & $90 < m_{bb}$ (GeV) < 160.

2. CMS has an additional 1b1j category: in the mass range where the contribution of boosted topologies is negligible, adding the 1b1j category improves the expected upper limits by 10-20% wrt 2b-only. This is consistent with what ATLAS observed in the 8 TeV analysis.

Analysis strategy for $\tau_\ell\tau_h$ final states (1)

CMS: Use a BDT to reduce the $t\bar{t}$ background, cut on the BDT score, then fit a mass-related variable in remaining events:

- two BDT trainings: one for (non-resonant and resonance masses up to 350 GeV), one for resonances mass above 350 GeV;
- non-resonant HH production $\rightarrow$ fit the stransverse mass m_{T2};
- resonant HH production $\rightarrow$ fit the four-body mass m_{HH}^{KinFit}.

ATLAS: Fit a BDT score distribution in the signal region: trainings performed against $t\bar{t}$ events, separately for the non-resonant HH sample and for each resonance mass.

Can the different analysis strategies explain the better sensitivity for ATLAS in the $\tau_\ell\tau_h$ search channel?

Analysis strategy for $\tau_\ell\tau_h$ final states (2)

CMS approach – BDT to reduce the $t\bar{t}$ background:

Kinematic variable	Non-resonant & $m \leq 350$ GeV	$m > 350$ GeV
$\Delta\phi(H_{bb}, H_{\tau\tau})$	✓	✓
$\Delta\phi(H_{bb}, \vec{p}_T^{miss})$	✓	✓
$\Delta\phi(H_{\tau\tau}, \vec{p}_T^{miss})$	✓	✓
$\Delta\phi(\ell, \vec{p}_T^{miss})$	✓	✓
$m_T(\ell, \vec{p}_T^{miss})$	✓	✓
$m_T(\tau_h, \vec{p}_T^{miss})$	✓	✓
$\Delta R(b, b) \times p_T(H_{bb})$	✓	
$\Delta R(\ell, \tau_h) \times p_T(H_{\tau\tau})$	✓	
$\Delta R(b, b)$		✓
$\Delta R(\ell, \tau_h)$		✓

H_{bb} and $H_{\tau\tau}$ are the two reconstructed Higgs boson candidates, m_T is the transverse mass, ℓ stands for an electron or muon (both $\tau_\ell\tau_h$ final states are merged for BDT training).

Different BDT cuts are applied:

$\to$ $t\bar{t}$ rejection of 90% (high- and low-mass resonant) or 70% (non-resonant), signal efficiencies varying from 65% to 95%.

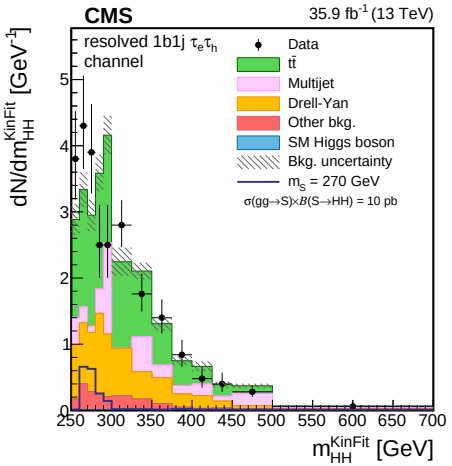
Analysis strategy for $\tau_\ell\tau_h$ final states (3a)

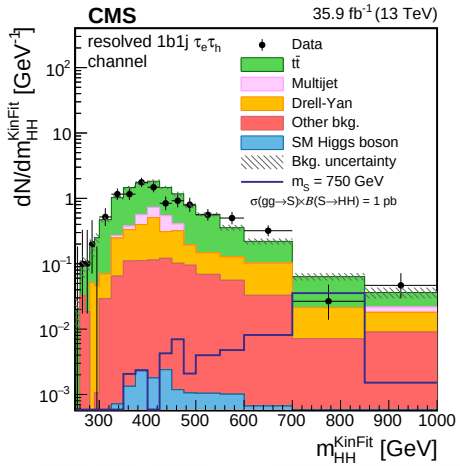
CMS approach – fitted mass variables – resonant

Kinematic fit based on the b, τ and $\vec{p}_T^{miss}$ objects, while assuming two 125 GeV Higgs boson decays.

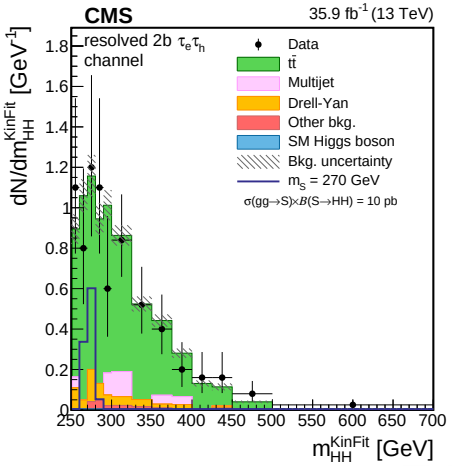
Invariant mass of the visible τ decay products and two b-jets: m_{HH}^{KinFit}.

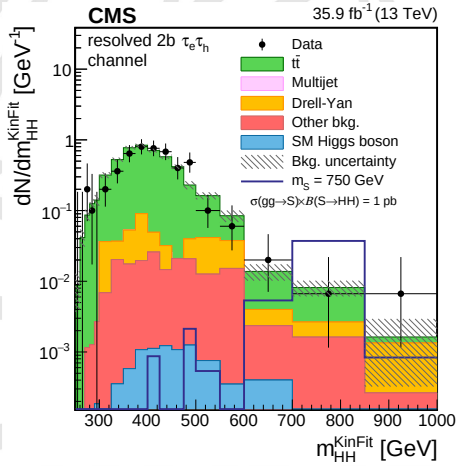
$\tau_e\tau_h$, 1b1j





$\tau_e\tau_h$, 2b





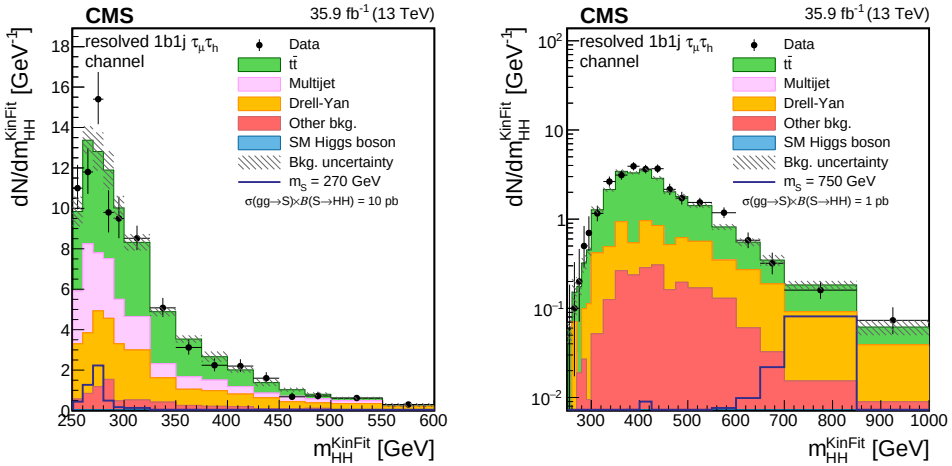
Analysis strategy for $\tau_\ell\tau_h$ final states (3b)

CMS approach – fitted mass variables – resonant

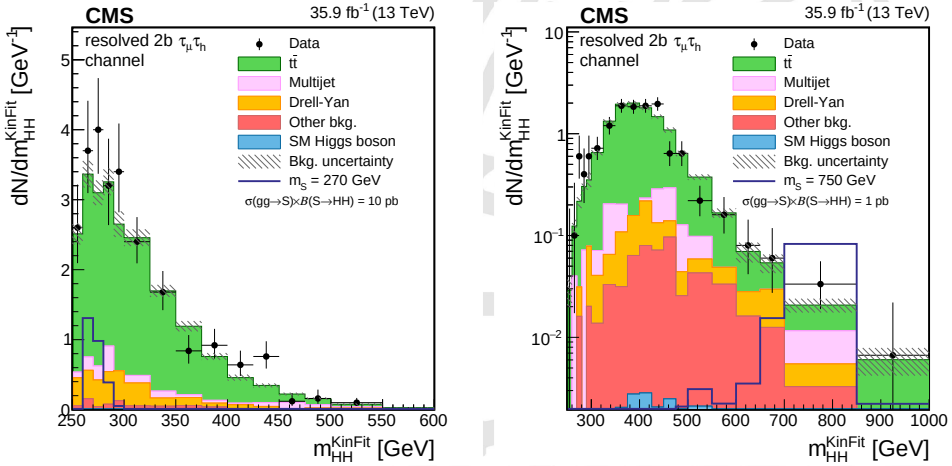
Kinematic fit based on the b, τ and $\vec{p}_T^{miss}$ objects, while assuming two 125 GeV Higgs boson decays.

Invariant mass of the visible τ decay products and two b-jets: m_{HH}^{KinFit}.

$\tau_\mu\tau_h$, 1b1j



$\tau_\mu\tau_h$, 2b



Analysis strategy for $\tau_\ell\tau_h$ final states (3c)

CMS approach – fitted mass variables – non-resonant

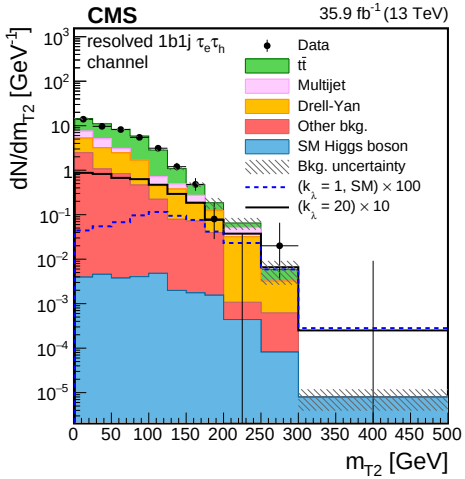
Stranverse mass m_{T2}:

Largest mass of the parent particle that is compatible with the kinematic constraints of the event.

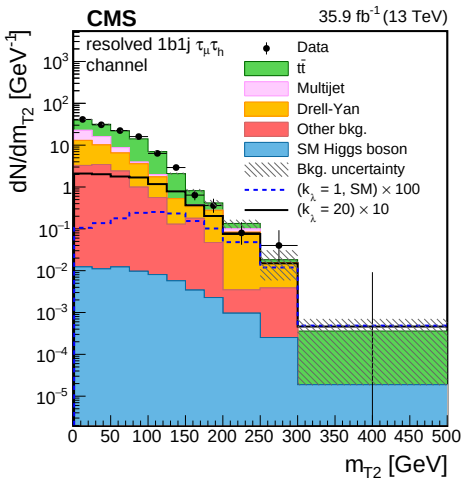
Parent particle: t decaying to b and $W \to \ell/\tau_h + \nu$.

m_{T2} is bounded above by m_t for $t\bar{t}$ events, but not for the HH signal!

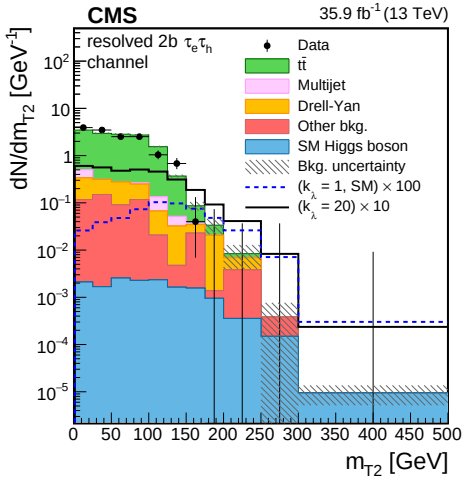
$\tau_e\tau_h$, 1b1j



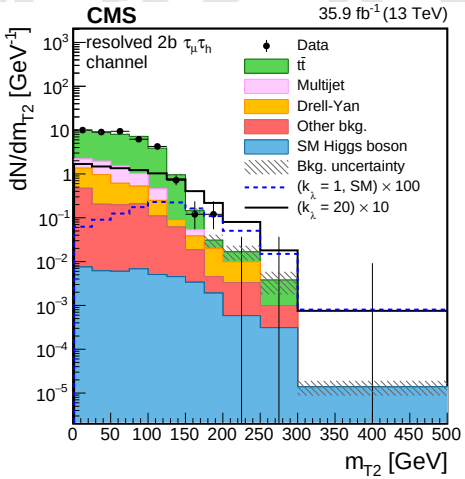
$\tau_\mu\tau_h$, 1b1j



$\tau_e\tau_h$, 2b



$\tau_\mu\tau_h$, 2b



Analysis strategy for $\tau_\ell\tau_h$ final states (4a)

ATLAS approach – BDT input variables

Kinematic variable	SLT resonant	SLT non-resonant & LTT
m_{HH}	✓	✓
$m_{\tau\tau}^{\text{MMC}}$	✓	✓
m_{bb}	✓	✓
$\Delta R(\ell, \tau_h)$	✓	✓
$\Delta R(b, b)$	✓	✓
E_T^{miss}	✓	
E_T^{miss} centrality	✓	
$m_T(\ell, \vec{p}_T^{\,miss})$	✓	✓
$\Delta\phi(H_{bb}, H_{\tau\tau})$	✓	
$\Delta p_T(\ell, \tau_h)$	✓	
Sub-leading b-jet p_T	✓	

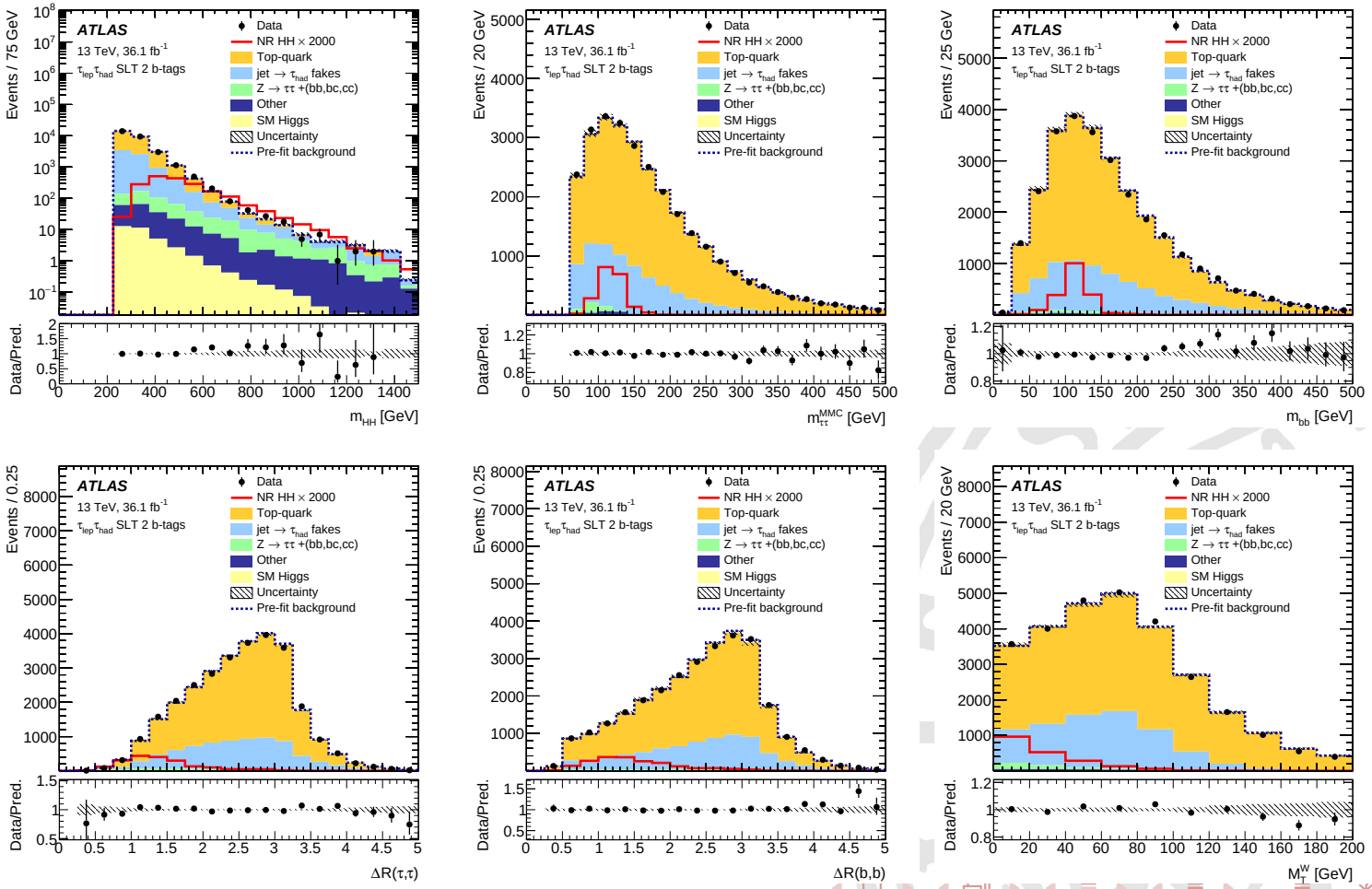
The E_T^{miss} centrality quantifies the angular position of $\vec{p}_T^{\,miss}$ with respect to the visible τ decay products in the transverse plane, i.e. $(A+B)/\sqrt{A^2+B^2}$ where
$A = \sin(\phi_{miss} - \phi_{\tau_h})/\sin(\phi_\ell - \phi_{\tau_h})$,
$B = \sin(\phi_\ell - \phi_{miss})/\sin(\phi_\ell - \phi_{\tau_h})$.

Resonant production: BDTs are trained at each mass point, but the signal model contains the target mass and the two neighbouring ones to ensure sensitivity to masses between simulated points.

Analysis strategy for $\tau_\ell\tau_h$ final states (4b)

ATLAS approach – BDT input variables

Predicted and measured post-fit (background-only) distributions of SLT non-resonant BDT input variables:

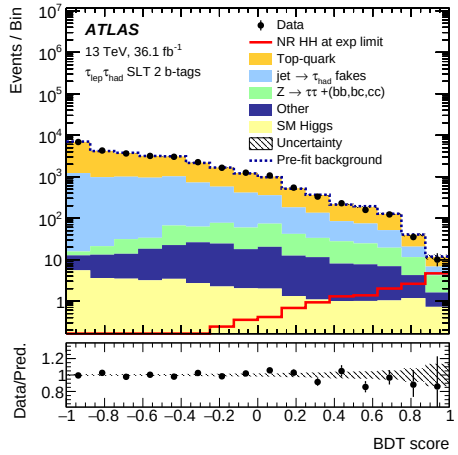


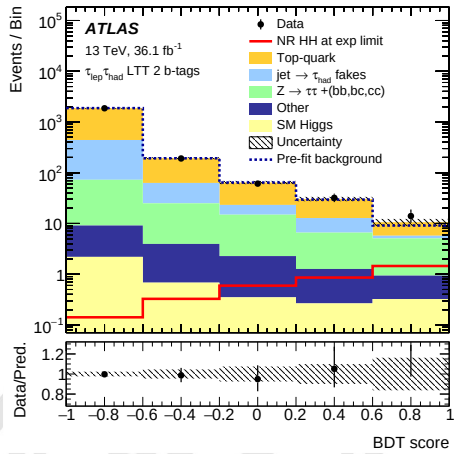
Analysis strategy for $\tau_\ell\tau_h$ final states (4c)

ATLAS approach – BDT score distributions

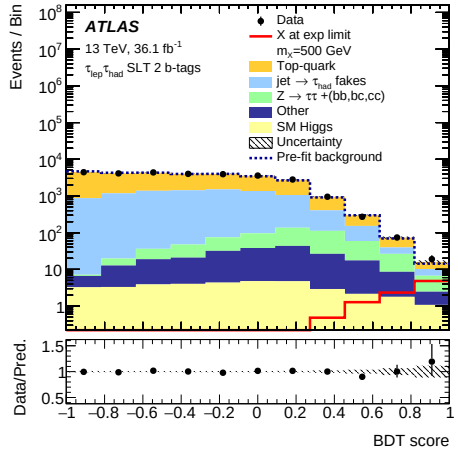
Predicted and measured post-fit (background-only) BDT score distributions in the 2b signal region. The low BDT score region constrains the $t\bar{t}$ normalization.

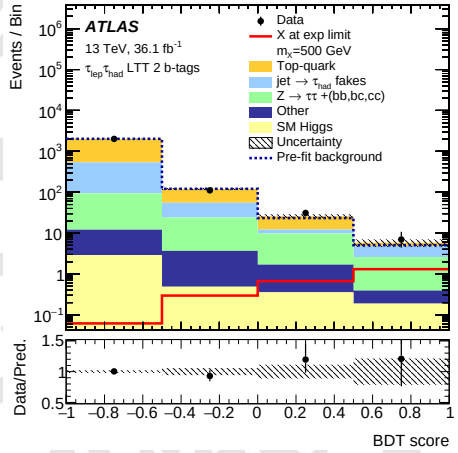
Non-resonant *HH* (SLT, LTT)





Resonant *HH* at 500 GeV (SLT, LTT)





UPPSALA UNIVERSITET

Analysis strategy for $\tau_\ell\tau_h$ final states (5)

Can the different analysis strategies explain the better sensitivity for ATLAS in the $\tau_\ell\tau_h$ search channel?

ATLAS also considered a cut-based analysis for the resonant production mode:

- Event selections;
- Remove m_{HH} from the BDT input variables;
- Cut on the BDT output distribution to keep highest-score bins;
- Fit the m_{HH} distribution.

In the resonant production mode, expected limits with a BDT-cut and m_{HH}-fit were found be worse than for the nominal analysis, over the whole mass range.

Analysis strategy for $\tau_h\tau_h$ final states (1)

CMS: No use of a BDT to reduce the backgrounds, only fit a mass-related variable:

- non-resonant HH production $\to$ stransverse mass m_{T2};
- resonant HH production $\to$ four-body mass m_{HH}^{KinFit}.

ATLAS: Fit a BDT score distribution in the signal region: trainings performed against $t\bar{t}$, $Z \to \tau\tau$, as well as the (data-driven) multi-jet background, separately for the non-resonant HH sample and for each resonance mass.

Can the different analysis strategies explain the better sensitivity for ATLAS in the $\tau_h\tau_h$ search channel?

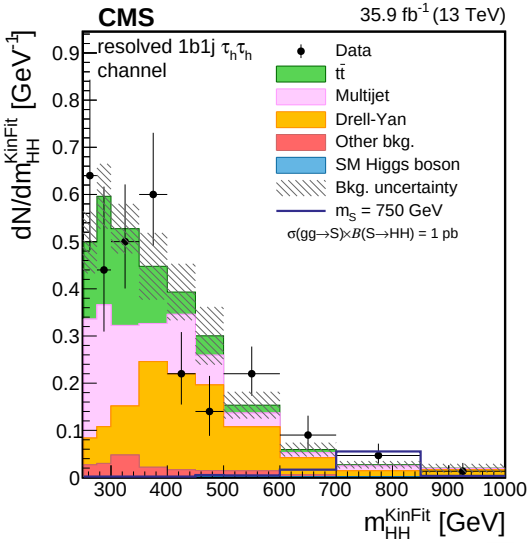
Analysis strategy for $\tau_h\tau_h$ final states (2a)

CMS approach – fitted mass variables – resonant

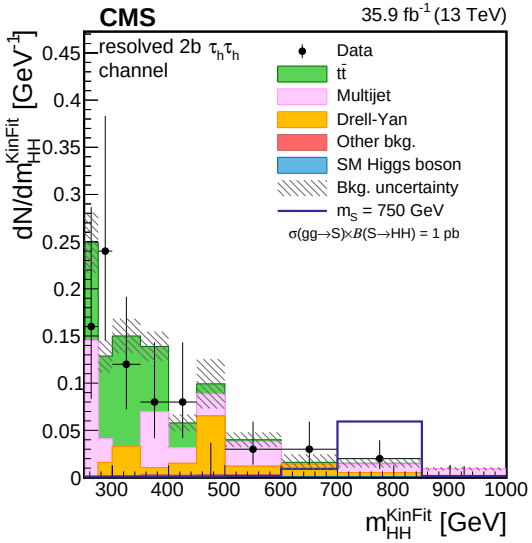
Kinematic fit based on the b, τ and $\vec{p}_T^{miss}$ objects, while assuming two 125 GeV Higgs boson decays.

Invariant mass of the visible τ decay products and two b-jets: m_{HH}^{KinFit}.

$\tau_h\tau_h$, 1b1j



$\tau_h\tau_h$, 2b



Analysis strategy for $\tau_h\tau_h$ final states (2b)

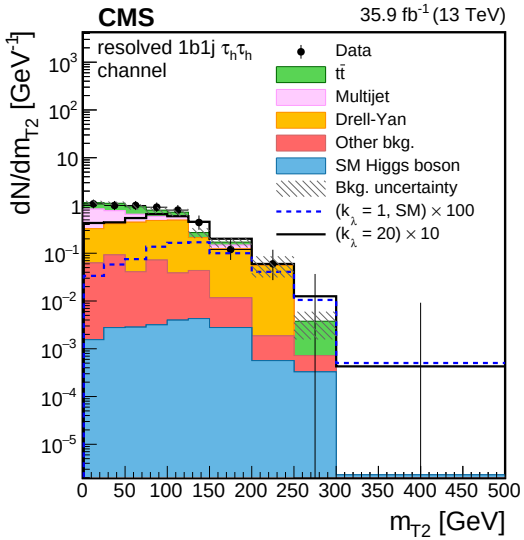
CMS approach – fitted mass variables – non-resonant

Stranverse mass m_{T2}:

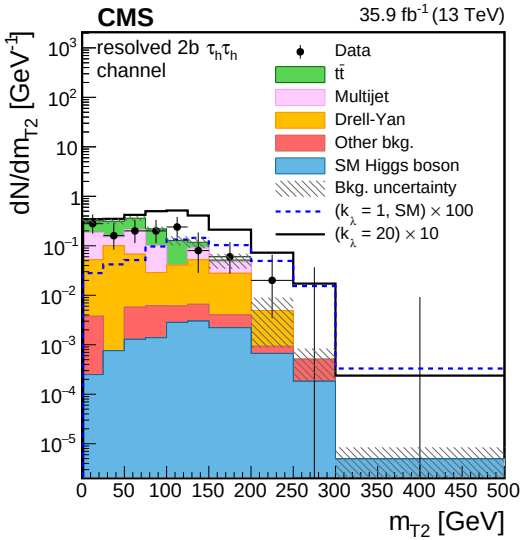
Largest mass of the parent particle that is compatible with the kinematic constraints of the event.

Parent particle: t decaying to b and $W \to \ell/\tau_h + \nu$.

$\tau_h\tau_h$, 1b1j



$\tau_h\tau_h$, 2b



m_{T2} is bounded above by m_t for $t\bar{t}$ events, but not for the HH signal!

Analysis strategy for $\tau_h\tau_h$ final states (3a)

ATLAS approach – BDT input variables

Kinematic variable	$\tau_h\tau_h$
m_{HH}	✓
$m_{\tau\tau}^{\text{MMC}}$	✓
m_{bb}	✓
$\Delta R(\tau_h, \tau_h)$	✓
$\Delta R(b, b)$	✓
E_T^{miss} centrality	✓

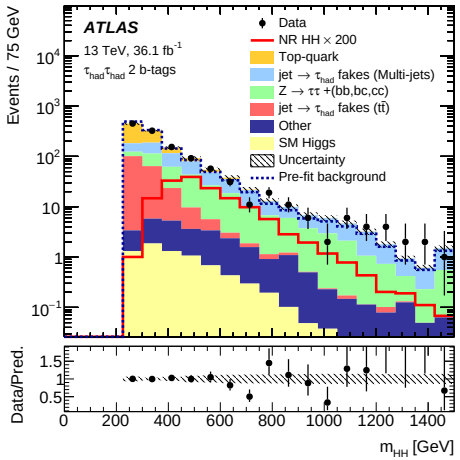
The E_T^{miss} centrality quantifies the angular position of $\vec{p}_T^{miss}$ with respect to the visible τ decay products in the transverse plane, i.e. $(A+B)/\sqrt{A^2+B^2}$ where
$A = \sin(\phi_{miss} - \phi_{\tau_{h2}})/\sin(\phi_{\tau_{h1}} - \phi_{\tau_{h2}})$,
$B = \sin(\phi_{\tau_{h1}} - \phi_{miss})/\sin(\phi_{\tau_{h1}} - \phi_{\tau_{h2}})$.

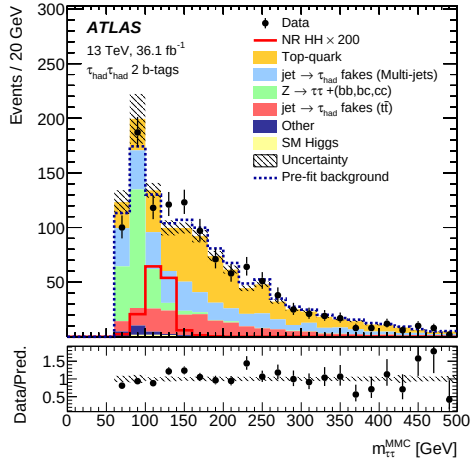
Resonant production: BDTs are trained at each mass point, but the signal model contains the target mass and the two neighbouring ones to ensure sensitivity to masses between simulated points.

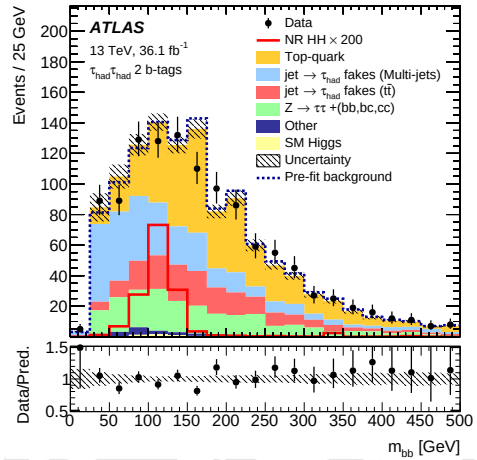
Analysis strategy for $\tau_h\tau_h$ final states (3b)

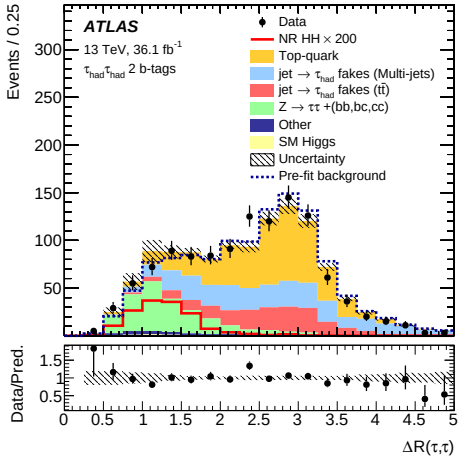
ATLAS approach – BDT input variables

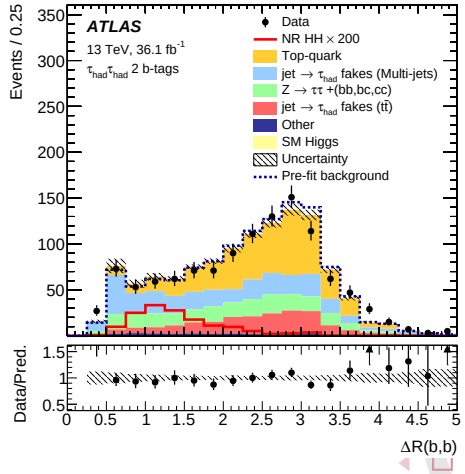
Predicted and measured post-fit (background-only) distributions of BDT input variables:

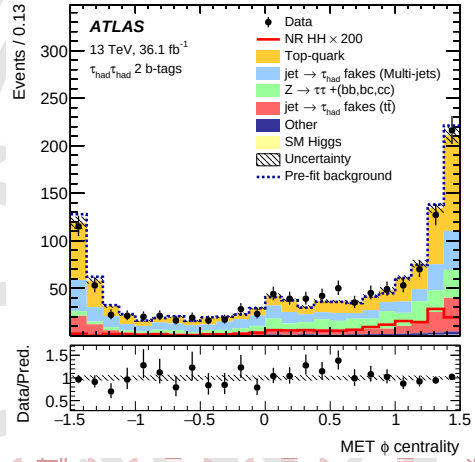










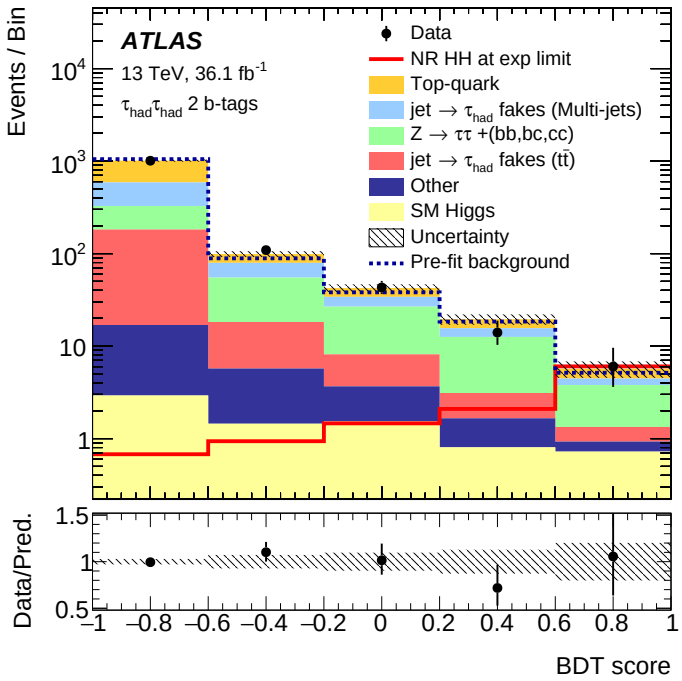


Analysis strategy for $\tau_h\tau_h$ final states (3c)

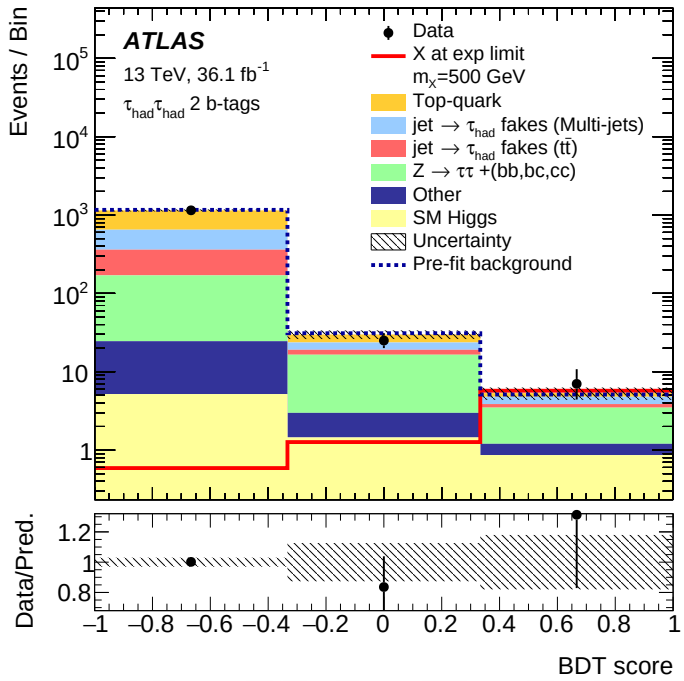
ATLAS approach – BDT score distributions

Predicted and measured post-fit (background-only) BDT score distributions in the 2b signal region.

Non-resonant *HH*



Resonant *HH* at 500 GeV



UPPSALA UNIVERSITET

Analysis strategy for $\tau_h\tau_h$ final states (4)

Can the different analysis strategies explain the better sensitivity for ATLAS in the $\tau_h\tau_h$ search channel?

ATLAS also considered a cut-based analysis for the resonant production mode:

- Event selections;
- Remove m_{HH} from the BDT input variables;
- Cut on the BDT output distribution to keep highest-score bins;
- Fit the m_{HH} distribution.

In the resonant production mode, expected limits with a BDT-cut and m_{HH}-fit were found be worse than for the nominal analysis, over the whole mass range.

Conclusion

1. It seems that fitting a BDT score distribution mostly explain why the ATLAS sensitivity is nearly a factor 2 better than for CMS.
2. Possible improvements towards the end of Run-2:
   - ATLAS: use a 1b signal region to increase a bit the signal acceptance & investigate the use of m_{HH}^{KinFit} and m_{T2} as BDT input variables.
   - CMS: use a BDT as discriminating variable in both $\tau_\ell\tau_h$ and $\tau_h\tau_h$ search channels?
3. General concern: BDTs should be designed to avoid bias towards m_{HH} and λ, or the signal model should cover a wider range of m_{HH} and λ than the target $\Rightarrow$ parameterised NN could help (as in CMS *bbVV*).

UPPSALA UNIVERSITET

Back-up: ATLAS and CMS event yields

ATLAS event yields

In the whole BDT range

	$\tau_{\text{lep}}\tau_{\text{had}}$ channel (SLT)	$\tau_{\text{lep}}\tau_{\text{had}}$ channel (LTT)	$\tau_{\text{had}}\tau_{\text{had}}$ channel
$t\bar{t}$	17800 ± 1100	1475 ± 94	360 ± 100
Single top	1130 ± 110	72.9 ± 7.6	39.7 ± 5.9
Multi-jet fake-τ_{had}	-	-	294 ± 57
$t\bar{t}$ fake-τ_{had}	-	-	160 ± 120
Fake-τ_{had}	9000 ± 1100	475 ± 76	-
$Z \to \tau\tau + (cc, bc, bb)$	416 ± 97	117 ± 28	291 ± 91
Other	197 ± 32	14.5 ± 2.3	22.9 ± 5.9
SM Higgs	38 ± 10	4.1 ± 1.0	8.2 ± 2.1
Total Background	28610 ± 180	2159 ± 46	1178 ± 40
Data	28612	2161	1180
$G_{\text{KK}}(300\,\text{GeV},\ k/\overline{M}_{Pl} = 1)$	23.6 ± 3.7	7.5 ± 1.2	13.1 ± 2.6
$G_{\text{KK}}(500\,\text{GeV},\ k/\overline{M}_{Pl} = 1)$	42.4 ± 6.4	9.9 ± 1.5	36.3 ± 7.0
$G_{\text{KK}}(1000/800(\text{LTT})\,\text{GeV},\ k/\overline{M}_{Pl} = 1)$	2.6 ± 0.4	1.06 ± 0.16	2.11 ± 0.43
$G_{\text{KK}}(300\,\text{GeV},\ k/\overline{M}_{Pl} = 2)$	327 ± 50	82 ± 13	240 ± 46
$G_{\text{KK}}(500\,\text{GeV},\ k/\overline{M}_{Pl} = 2)$	193 ± 29	39.7 ± 6.1	187 ± 36
$G_{\text{KK}}(1000/800(\text{LTT})\,\text{GeV},\ k/\overline{M}_{Pl} = 2)$	8.6 ± 1.3	3.63 ± 0.56	7.9 ± 1.6
$X(300\,\text{GeV})$	39.1 ± 6.3	11.8 ± 1.9	17.9 ± 3.6
$X(500\,\text{GeV})$	3.41 ± 0.52	0.88 ± 0.13	2.84 ± 0.54
$X(1000/800(\text{LTT})\,\text{GeV})$	0.0267 ± 0.0041	0.0228 ± 0.0035	0.0222 ± 0.0044
NR HH	0.99 ± 0.13	0.225 ± 0.033	0.75 ± 0.14

ATLAS event yields

In the last two bins of the BDT

	$\tau_{\text{lep}}\tau_{\text{had}}$ channel (SLT)	$\tau_{\text{lep}}\tau_{\text{had}}$ channel (LTT)	$\tau_{\text{had}}\tau_{\text{had}}$ channel
$t\bar{t}$	18.2 ± 4.2	23.2 ± 1.7	4.5 ± 1.4
Single top	6.4 ± 1.3	3.7 ± 1.2	1.06 ± 0.57
Multi-jet fake-τ_{had}	-	-	3.89 ± 0.87
$t\bar{t}$ fake-τ_{had}	-	-	1.9 ± 1.4
Fake-τ_{had}	12.0 ± 2.3	6.6 ± 1.5	-
$Z \to \tau\tau + (cc, bc, bb)$	10.2 ± 2.6	7.7 ± 3.1	12.6 ± 3.6
Other	3.89 ± 0.69	1.51 ± 0.36	1.09 ± 0.32
SM Higgs	1.94 ± 0.43	0.58 ± 0.14	1.54 ± 0.41
Total Background	52.7 ± 4.5	39.5 ± 3.0	26.7 ± 3.5
Data	45	47	20
NR HH	0.49 ± 0.07	0.16 ± 0.02	0.55 ± 0.10

CMS event yields

$e\tau_h$ from Luca's thesis: https://cds.cern.ch/record/2292733

$\tau_e\tau_h$ final state			
Process	res. 1b1j	res. 2b0j	boosted
$t\bar{t}$	631.8 ± 16.3	311.1 ± 9.3	8.9 ± 0.4
QCD	135.9 ± 11.7	6.7 ± 2.1	6.5 ± 2.1
Z+jets	213.3 ± 7.0	20.2 ± 0.8	2.2 ± 0.1
W+jets	70.2 ± 3.2	0.42 ± 0.02	0.47 ± 0.02
single top	48.9 ± 3.2	10.5 ± 0.8	0.82 ± 0.05
diboson	7.9 ± 0.5	1.1 ± 0.1	0.42 ± 0.03
EWK W/Z	3.3 ± 0.1	0.91 ± 0.03	0.33 ± 0.02
SM Higgs	0.69 ± 0.04	0.41 ± 0.03	0.12 ± 0.01
Tot. exp. bkg.	1112 ± 22	351 ± 10	19.7 ± 2.1
Expected signal			
$k_\lambda = 1$ (SM)	0.16	0.14	0.04
$k_\lambda = 20$	10.28	8.26	0.55
Observed data	1057	355	11

Table 6.6 – Observed and expected event yields in different signal regions of the nonresonant search for the $\tau_e\tau_h$ final state. Quoted uncertainties are obtained after fixing the nuisance parameters to their estimation from a maximum likelihood fit under the background-only hypothesis.

UPPSALA UNIVERSITET

CMS event yields

$\mu\tau_h$ from Luca's thesis: https://cds.cern.ch/record/2292733

$\tau_\mu\tau_h$ final state			
Process	res. 1b1j	res. 2b0j	boosted
$t\bar{t}$	1617.6 ± 38.7	802.2 ± 22.4	20.0 ± 0.9
QCD	443.9 ± 38.2	80.9 ± 7.0	5.6 ± 1.9
Z+jets	629.6 ± 22.3	64.8 ± 2.9	7.1 ± 0.3
W+jets	124.7 ± 6.7	4.9 ± 0.2	0.95 ± 0.04
single top	121.9 ± 7.8	22.0 ± 1.5	2.5 ± 0.2
diboson	18.3 ± 1.2	2.9 ± 0.3	0.89 ± 0.06
EWK W/Z	9.4 ± 0.5	1.2 ± 0.1	0.15 ± 0.01
SM Higgs	1.7 ± 0.1	1.1 ± 0.1	0.18 ± 0.01
Tot. exp. bkg.	2967 ± 60	980 ± 24	38 ± 2
Expected signal			
$k_\lambda = 1$ (SM)	0.38	0.33	0.08
$k_\lambda = 20$	25.75	20.88	1.12
Observed data	3020	996	35

Table 6.5 – Observed and expected event yields in different signal regions of the nonresonant search for the $\tau_\mu\tau_h$ final state. Quoted uncertainties are obtained after fixing the nuisance parameters to their estimation from a maximum likelihood fit under the background-only hypothesis.

CMS event yields

$\tau_h\tau_h$ from Luca's thesis: https://cds.cern.ch/record/2292733

	$\tau_h\tau_h$ final state		
Process	res. 1b1j	res. 2b0j	boosted
$t\bar{t}$	33.6 ± 1.5	16.5 ± 1.1	0.068 ± 0.004
QCD	40.6 ± 7.9	14.5 ± 2.8	0.012 ± 0.012
Z+jets	48.7 ± 6.2	9.1 ± 1.0	2.2 ± 0.1
W+jets	1.11 ± 0.06	-	0.031 ± 0.002
single top	4.2 ± 0.3	0.026 ± 0.002	-
diboson	2.3 ± 0.4	0.57 ± 0.08	0.33 ± 0.03
EWK W/Z	0.78 ± 0.04	-	0.15 ± 0.01
SM Higgs	0.63 ± 0.08	0.38 ± 0.05	0.14 ± 0.01
Tot. exp. bkg.	132 ± 10	41 ± 3	2.9 ± 0.1
Expected signal for $\sigma(gg \to S) \times \mathcal{B}(S \to HH \to bb\tau\tau) = 1$ pb in resonant case			
$m_X = 300$ GeV	20.48	15.03	0.08
$m_X = 600$ GeV	185.27	165.44	40.51
$m_X = 900$ GeV	126.17	105.13	379.10
$k_\lambda = 1$ (SM)	0.24	0.21	0.05
$k_\lambda = 20$	9.20	7.88	0.60
Observed data	140	33	3

Table 6.4 – Observed and expected event yields in different signal regions of the $\tau_h\tau_h$ final state (the definition is the same for both resonant and nonresonant searches). Quoted uncertainties are obtained after fixing the nuisance parameters to their estimation from a maximum likelihood fit under the background-only hypothesis.