Evidence for Higgs boson production in association with a top quark pair

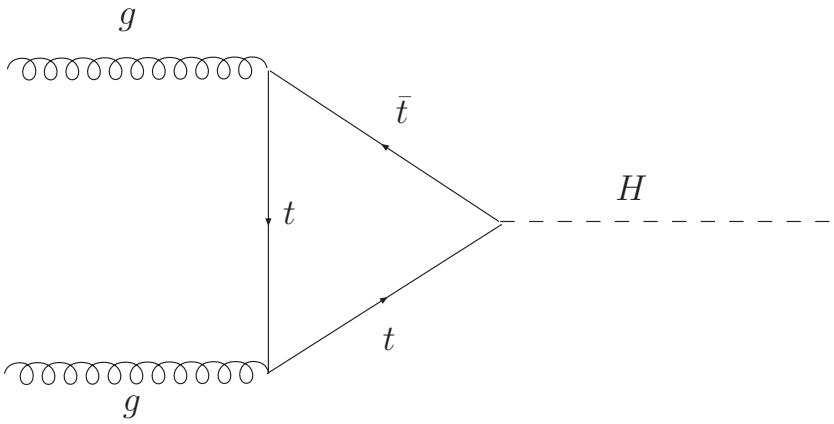
Arthur Chomont (LAL)
On behalf of the ATLAS collaboration

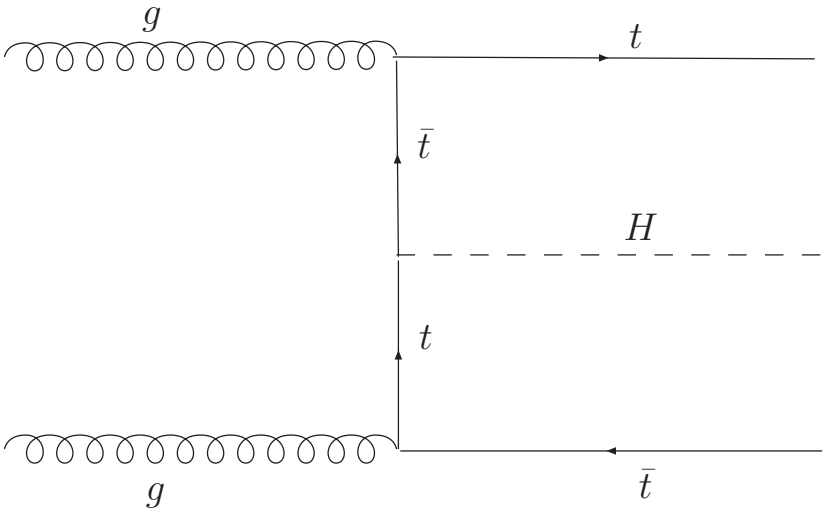
DIS 2018

April 17th 2018

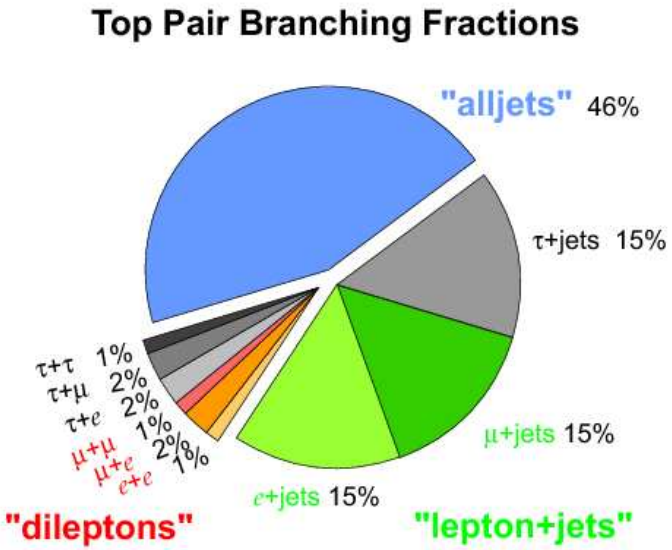
Introduction : $t\bar{t}H$ at LHC

- Top Yukawa coupling is the largest coupling of Higgs boson to fermion and a key parameter of SM
  - $y_t = \sqrt{2} * m_t/v \simeq 1$
  - Sensitive to BSM physics
- Indirectly obtained through measurement of top quark mass
- Indirectly observed through SM Higgs decaying in two photons and production of Higgs by gluon-gluon fusion
- Direct measurement possible through $t\bar{t}H$ production
  - Tree-level process
  - Run 1 results ATLAS+CMS: $\mu_{t\bar{t}H} = \sigma_{meas}/\sigma_{SM} = 2.3^{+0.7}_{-0.6}$

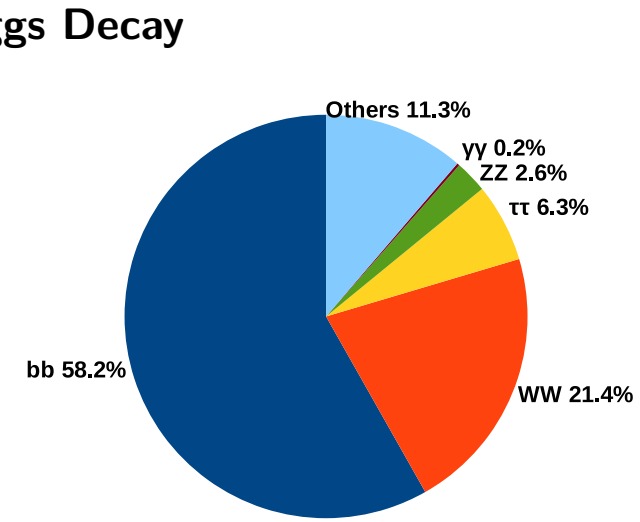




$t\bar{t}H$ experimental signature



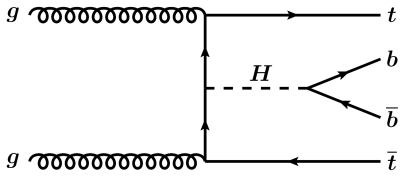
×

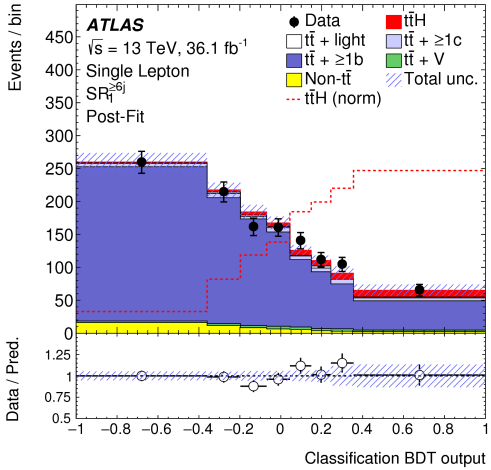


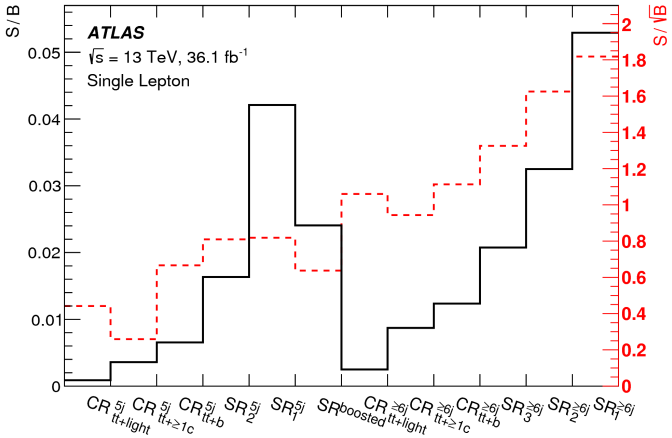
- Possible final states
  - $H \to b\bar{b}$: 4b + 2W $\to$ 1712.08895 (submitted to Phys.Rev.D.)
  - $H \to WW, \tau\tau, ZZ$: 2b + multileptons
    - In case of resonant ($H \to ZZ \to 4\ell$) $\to$ 1712.02304 (submitted to JHEP)
    - Other final states $\to$ 1712.08891 (submitted to Phys.Rev.D.)
  - $H \to \gamma\gamma$: 2b + 2γ $\to$ 1802.04146 (submitted to Phys.Rev.D.)
- The combination of all the final states will also be described $\to$ 1712.08891 (submitted to Phys.Rev.D.)

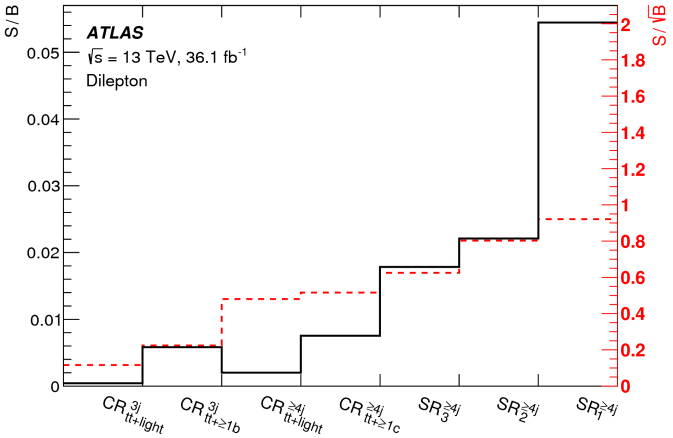
$t\bar{t}H$ ($H \to b\bar{b}$) : Classification and analysis strategy

- Final state with largest BR from Higgs but with very large background contribution ($t\bar{t}$ +jets)
- Categorisation based on number of leptons, jets and b-tagged jets (using 4 b-tagging working points)
  - For both single lepton and dilepton final states → control regions used in the fit to constrain background modelling
  - One boosted category with single lepton included
- Strategy of analysis → Classification BDT using intermediate MVA classifiers in Signal Regions:
  - Reconstruction BDT: reconstruct Higgs and top candidates
  - Likelihood and Matrix Element: classify signal vs background events



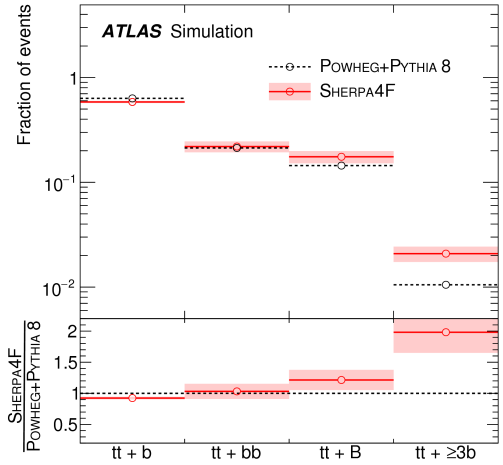


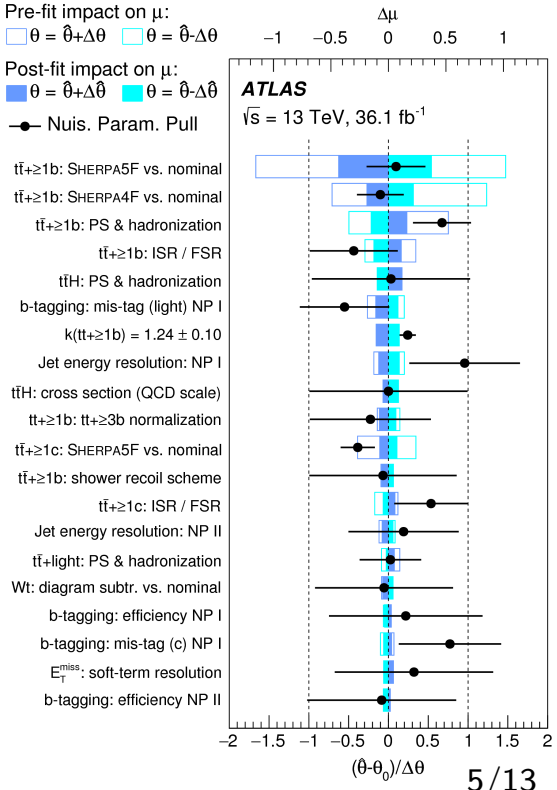


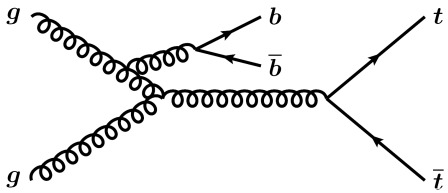


$t\bar{t}H$ ($H \to b\bar{b}$) : Background modelling

- $t\bar{t}$ +jets is the main background
  - 85% to 95% of the background in the different regions
  - Control Regions enriched in either $t\bar{t}$ +$\geq$ 1b, $t\bar{t}$ +$\geq$ 1c, $t\bar{t}$ +light to improve modelling
- Nominal $t\bar{t}$ +jets sample simulated at NLO with Powheg+Pythia8
  - $t\bar{t}$ +$\geq$ 1b: relative contribution of each sub-component reweighted to predictions by Sherpa+OpenLoops (NLO, 4-flavour scheme)
  - Split by number of b and c jets based on additional b/c hadrons in the event
- Normalisation of $t\bar{t}$ +HF free-floating in fit (one normalisation for tt+>=1b and tt+>=1c)
- Systematics on $t\bar{t}$ + $\geq 1b$ from comparisons with alternative simulations quite large



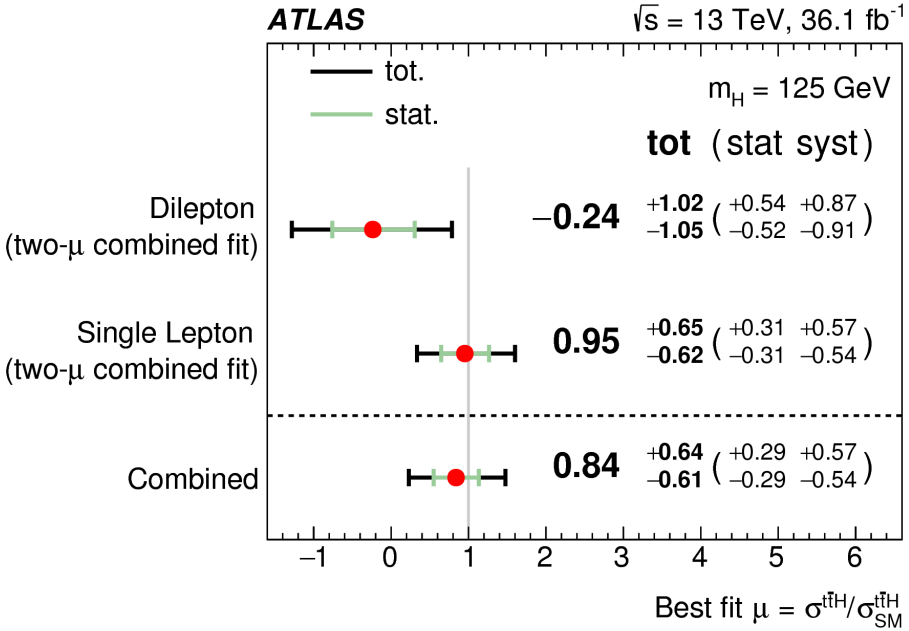




$t\bar{t}H$ ($H \to b\bar{b}$) : Results

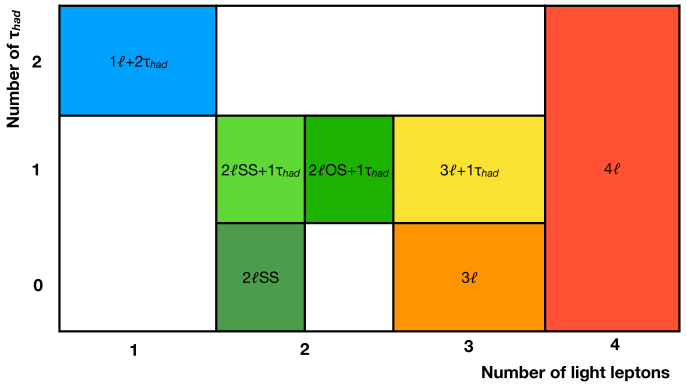
- Final fit done on:
  - classification BDT output in SRs
  - H_T^{had} in $t\bar{t}$ +($\geq 1c$) CRs in 1ℓ channel, event yields in other CRs
- Normalisation factors on $t\bar{t}$ +HF:
  - $t\bar{t}$ +($\geq 1b$): 1.24 ± 0.10
  - $t\bar{t}$ +($\geq 1c$): 1.63 ± 0.23
- Observed (expected) sensitivity at 1.4 (1.6)σ over SM background

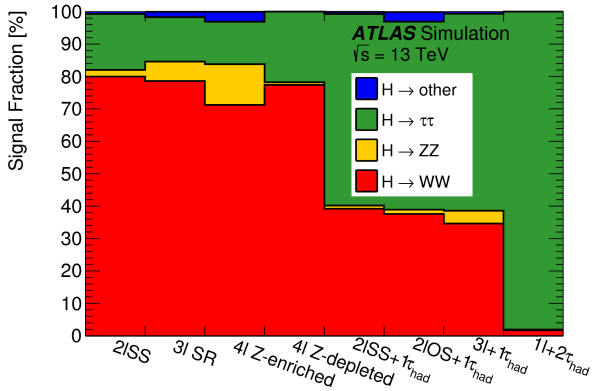
Uncertainty source	$\Delta\mu$	
$t\bar{t}+\geq 1b$ modeling	+0.46	−0.46
Background-model stat. unc.	+0.29	−0.31
b-tagging efficiency and mis-tag rates	+0.16	−0.16
Jet energy scale and resolution	+0.14	−0.14
$t\bar{t}H$ modeling	+0.22	−0.05
$t\bar{t}+\geq 1c$ modeling	+0.09	−0.11
JVT, pileup modeling	+0.03	−0.05
Other background modeling	+0.08	−0.08
$t\bar{t}+$light modeling	+0.06	−0.03
Luminosity	+0.03	−0.02
Light lepton (e,μ) id., isolation, trigger	+0.03	−0.04
Total systematic uncertainty	+0.57	−0.54
$t\bar{t}+\geq 1b$ normalization	+0.09	−0.10
$t\bar{t}+\geq 1c$ normalization	+0.02	−0.03
Intrinsic statistical uncertainty	+0.21	−0.20
Total statistical uncertainty	+0.29	−0.29
Total uncertainty	+0.64	−0.61



$t\bar{t}H(multileptons)$: Classification and analysis strategy

- Target Higgs decays to WW, ZZ and $\tau\tau$ through leptonic final states
  - Less signal than in $t\bar{t}H$ ($H \to bb$) final states but less background contamination
- Categories defined following number, charge and flavour of the charged leptons
  - Target different decays of the Higgs
  - With additional requirements on number of light and b-tagged jets to reject VV and V+jets backgrounds
- Strategy:
  - MVA in lepton definitions to reject non-prompt and charge mis-id backgrounds
  - Event classified in the different regions using MVA (except $3\ell + 1\tau$)

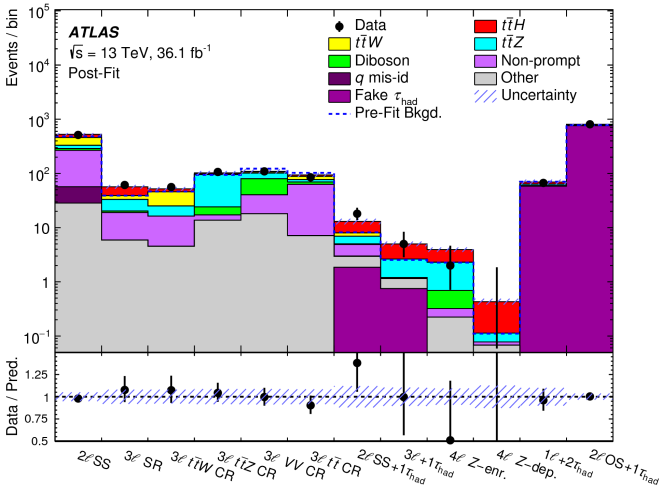


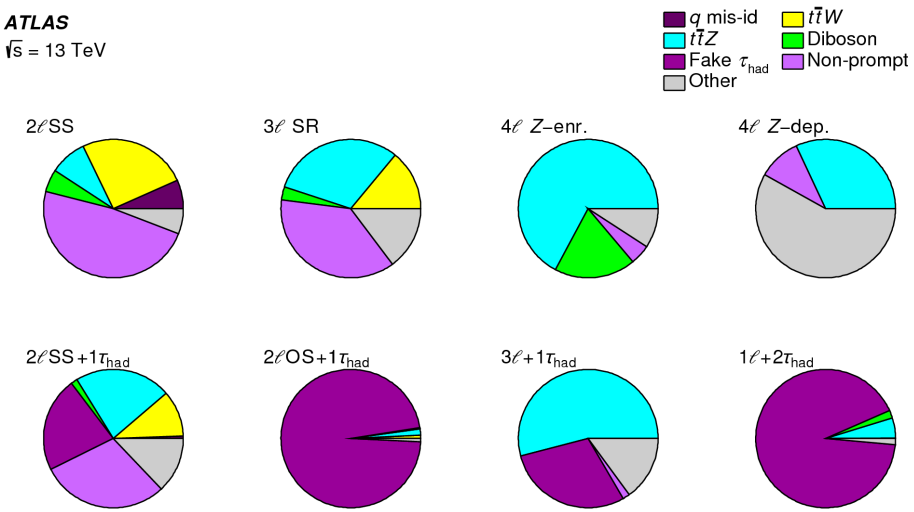


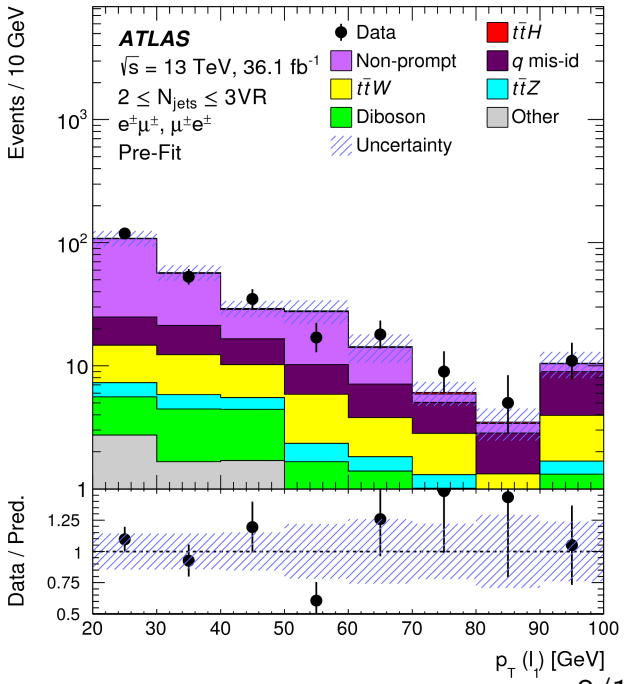
	2ℓSS	3ℓ	4ℓ	1ℓ+2τ_{had}	2ℓSS+1τ_{had}	2ℓOS+1τ_{had}	3ℓ+1τ_{had}
BDT trained against	Fakes and $t\bar{t}V$	$t\bar{t}$, $t\bar{t}W$, $t\bar{t}Z$, VV	$t\bar{t}Z$ / -	$t\bar{t}$	all	$t\bar{t}$	-
Discriminant	2×1D BDT	5D BDT	Event count	BDT	BDT	BDT	Event count
Number of bins	6	5	1 / 1	2	2	10	1
Control regions	-	4	-	-	-	-	-

$t\bar{t}H(multileptons)$: Background estimation

- Irreducible backgrounds: mostly $t\bar{t}V$ with prompt leptons
  - Estimated from MC and validated in CRs (in 3ℓ category)
- Reducible background: $t\bar{t}$ events with mis-reconstructed leptons
  - Estimated through different data-driven methods
  - Non-prompt light leptons: originated from b-hadrons decay or photon conversion
  - q mis-id electrons: due to high p_T electrons or trident process
  - Fake τ_{had}: from light jets and electrons

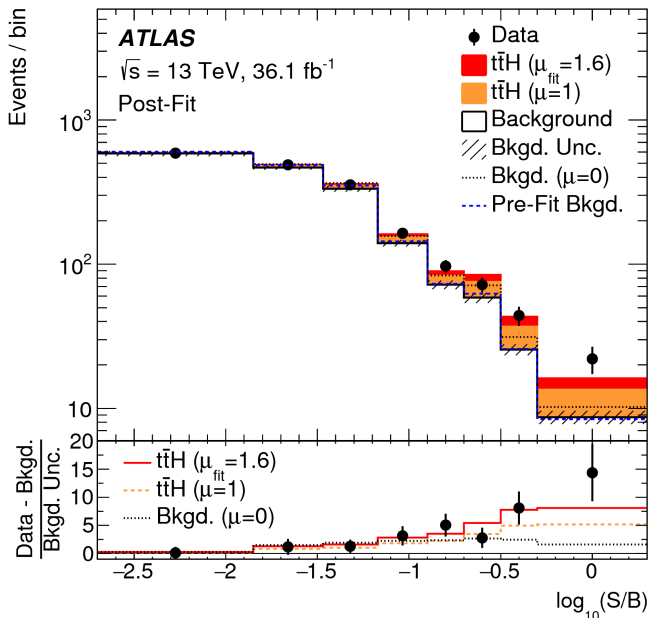




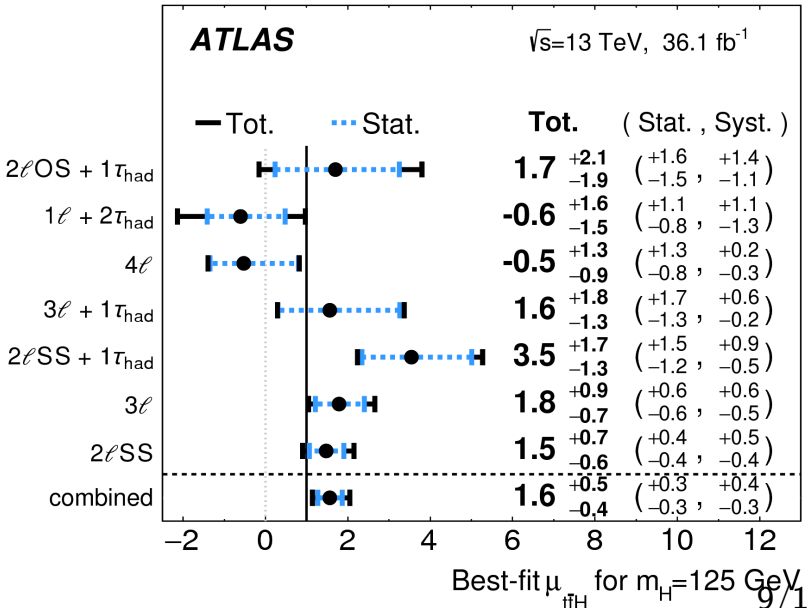


$t\bar{t}H(multileptons)$: Results

- Observed (expected) significance at 4.1 (2.8) σ over SM background
- Alternative fit with free floating normalisation for $t\bar{t}V$
  - Same central μ value for signal (degradation in sensitivity)
  - $\mu_{ttZ} = 1.17^{+0.25}_{-0.22}$ and $\mu_{ttW} = 0.92 \pm 0.32$
- Measured cross-section at $\sigma_{t\bar{t}H} = 790^{+230}_{-210}$fb (expected: 507^{+35}_{-50}fb)

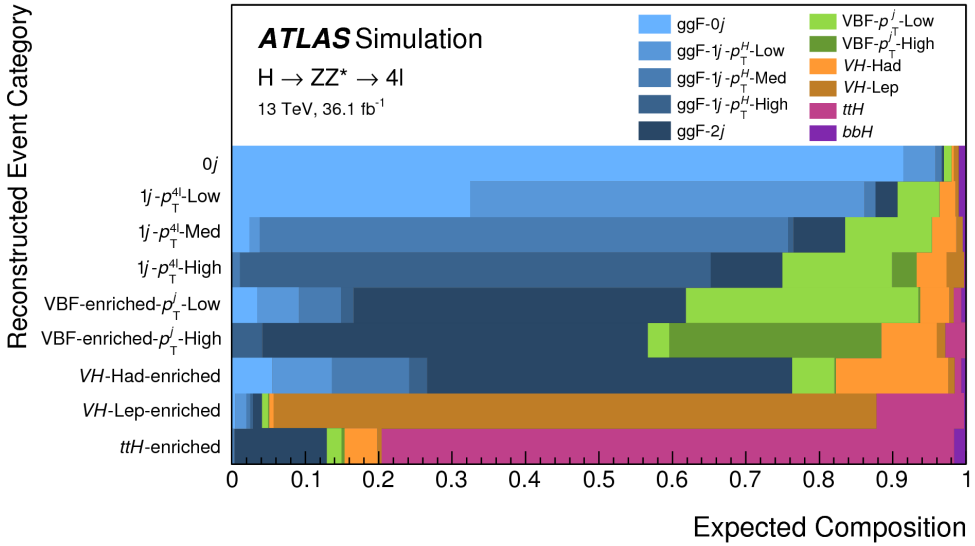


Uncertainty Source	$\Delta\mu$	
$t\bar{t}H$ modeling (cross section)	+0.20	−0.09
Jet energy scale and resolution	+0.18	−0.15
Non-prompt light-lepton estimates	+0.15	−0.13
Jet flavor tagging and τ_{had} identification	+0.11	−0.09
$t\bar{t}W$ modeling	+0.10	−0.09
$t\bar{t}Z$ modeling	+0.08	−0.07
Other background modeling	+0.08	−0.07
Luminosity	+0.08	−0.06
$t\bar{t}H$ modeling (acceptance)	+0.08	−0.04
Fake τ_{had} estimates	+0.07	−0.07
Other experimental uncertainties	+0.05	−0.04
Simulation sample size	+0.04	−0.04
Charge misassignment	+0.01	−0.01
Total systematic uncertainty	+0.39	−0.30



$t\bar{t}H$ ($H \to ZZ* \to 4\ell$)

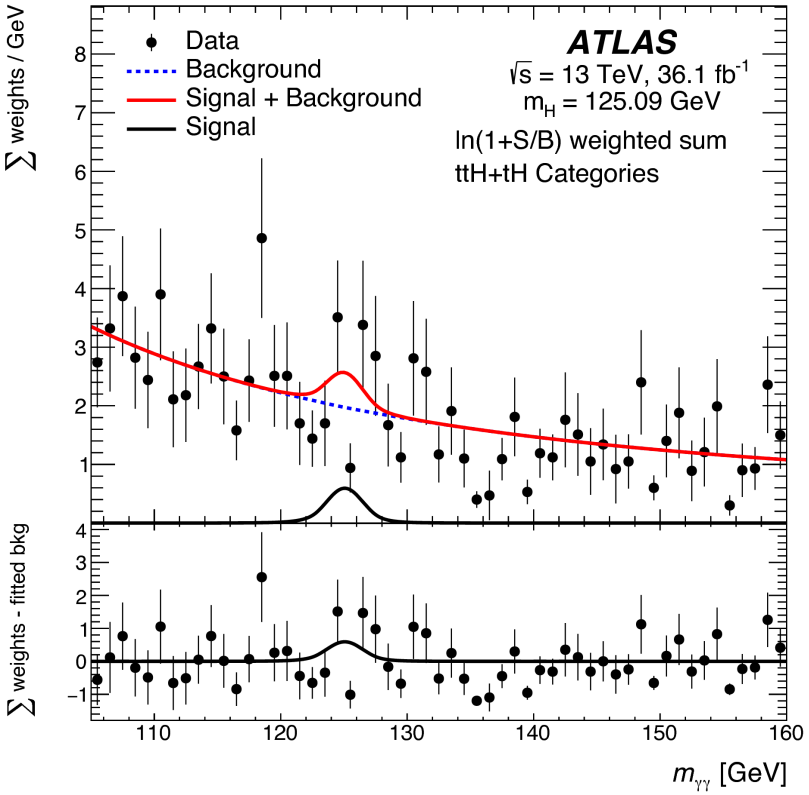
- Rare decay but with very small background
- Higgs boson candidate with 118$< m_{4\ell} <$ 129 GeV
- Additional requirement:
  - $\geq$ 1b-tagged jet
  - $\geq$ 4 jets or $\geq$ 2jet and 1 more lepton
- 0.4 $t\bar{t}H$ event and 0.08 background event expected and 0 data event observed



- $\mu < 1.9$ (68% CL)

$t\bar{t}H$ ($H \to \gamma\gamma$)

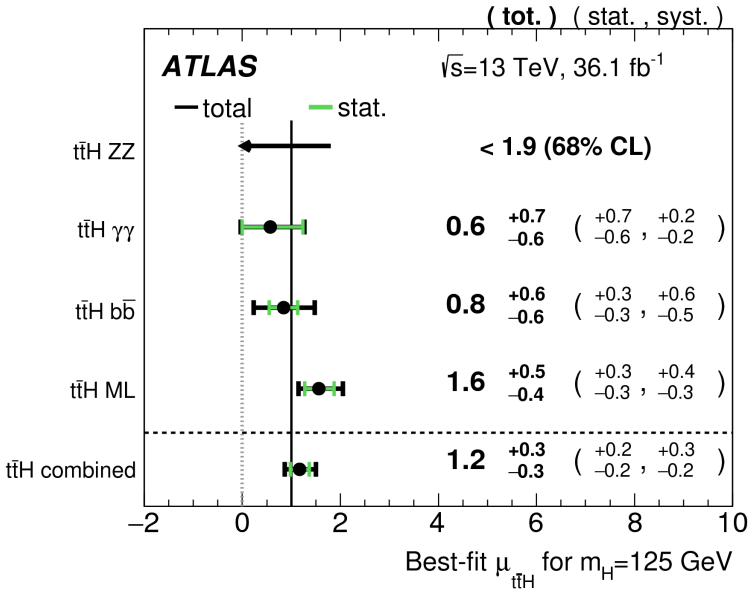
- Very rare decay but very pure signal
- Categories following number of leptons
  - $\geq 1\ell$, at least 2 jets and 1 b-tagged jet
  - 0 lepton, at least 3 jets and 1 b-tagged jet + BDT
- $t\bar{t}H$, tHq and WtH fitted together



- $\mu = 0.5 \pm 0.6$

Combination

- Combination of all previous final states
- Assumptions
  - *tHqb*, *WtH* treated as background (fixed to SM expectations within theory uncertainties)
  - BR of Higgs decay fixed to SM values
- **Cross-section measured: 590^{+160}_{-150}fb (expected: 507^{+35}_{-50}fb)**
- **Observed (expected) significance: 4.2 (3.8)σ**



Channel	Best-fit μ		Significance	
	Observed	Expected	Observed	Expected
Multilepton	$1.6\ ^{+0.5}_{-0.4}$	$1.0\ ^{+0.4}_{-0.4}$	4.1σ	2.8σ
$H \to b\bar{b}$	$0.8\ ^{+0.6}_{-0.6}$	$1.0\ ^{+0.6}_{-0.6}$	1.4σ	1.6σ
$H \to \gamma\gamma$	$0.6\ ^{+0.7}_{-0.6}$	$1.0\ ^{+0.8}_{-0.6}$	0.9σ	1.7σ
$H \to 4\ell$	< 1.9	$1.0\ ^{+3.2}_{-1.0}$	—	0.6σ
Combined	$1.2\ ^{+0.3}_{-0.3}$	$1.0\ ^{+0.3}_{-0.3}$	4.2σ	3.8σ

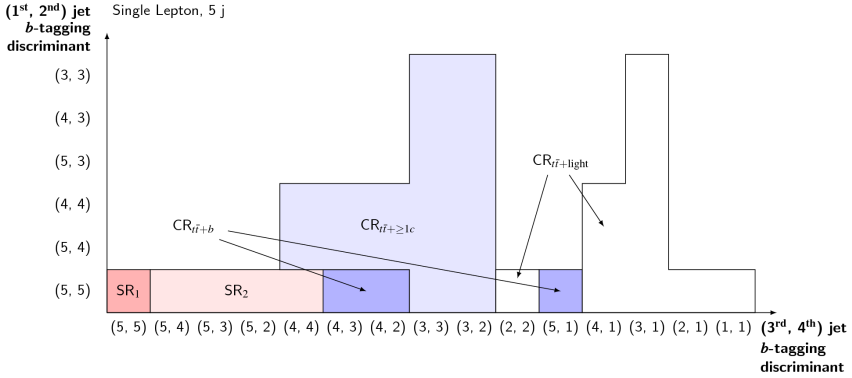
Uncertainty Source	$\Delta\mu$	
$t\bar{t}$ modeling in $H \to b\bar{b}$ analysis	$+0.15$	-0.14
$t\bar{t}H$ modeling (cross section)	$+0.13$	-0.06
Non-prompt light-lepton and fake τ_{had} estimates	$+0.09$	-0.09
Simulation statistics	$+0.08$	-0.08
Jet energy scale and resolution	$+0.08$	-0.07
$t\bar{t}V$ modeling	$+0.07$	-0.07
$t\bar{t}H$ modeling (acceptance)	$+0.07$	-0.04
Other non-Higgs boson backgrounds	$+0.06$	-0.05
Other experimental uncertainties	$+0.05$	-0.05
Luminosity	$+0.05$	-0.04
Jet flavor tagging	$+0.03$	-0.02
Modeling of other Higgs boson production modes	$+0.01$	-0.01
Total systematic uncertainty	$+0.27$	-0.23
Statistical uncertainty	$+0.19$	-0.19
Total uncertainty	$+0.34$	-0.30

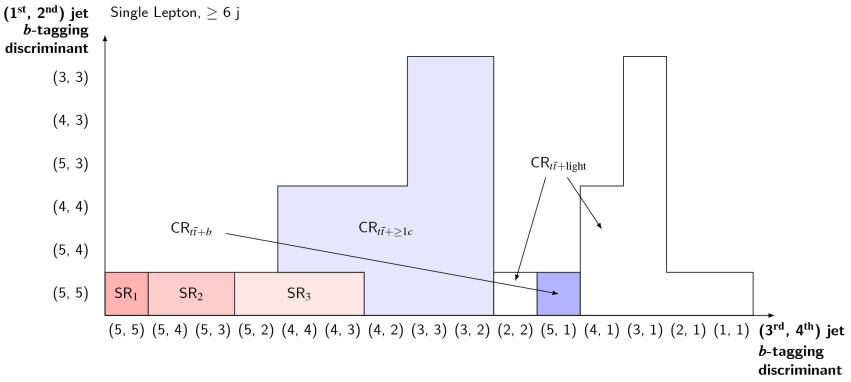
Conclusion

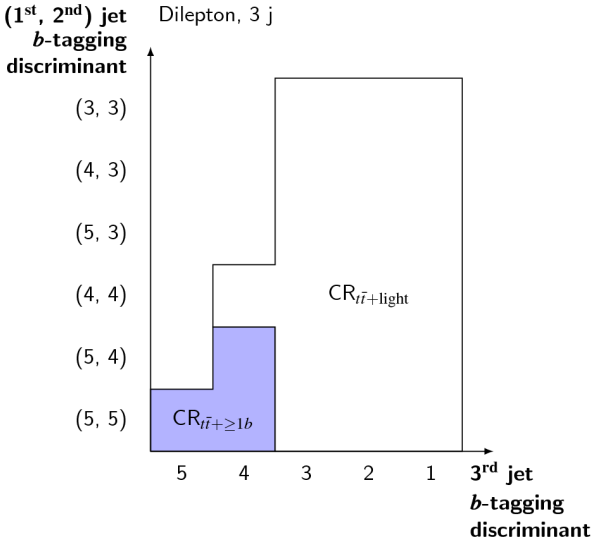
- **Evidence in the $t\bar{t}H$ search at ATLAS with $36.1 fb^{-1}$at 4.2σ (3.8σ expected)**
  - Heavy use of MVA significantly improve sensitivity of the analysis
  - Still large uncertainties on modelling and some irreducible backgrounds such as $t\bar{t}V$ and $t\bar{t}$ +HF
- **Measured cross-section: 590^{+160}_{-150}fb (expected: 507^{+35}_{-50}fb)**
- Still some final states statistically limited such $H \to \gamma\gamma$, $H \to ZZ \to 4\ell$ will significantly improve with 2017 data
- Increase in luminosity will also help in background modelling to reduce systematic uncertainties for the other final states (non-prompt estimates in multilepton channel . . .)

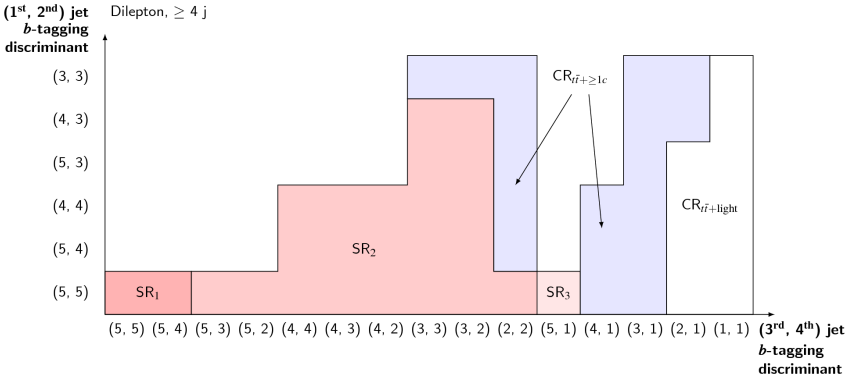
Backup

ttHbb


(1st, 2nd) jet b-tagging discriminant
Single Lepton, 5 j
(3, 3)
(4, 3)
(5, 3)
(4, 4)
(5, 4)
(5, 5)
SR1
SR2
CRtt̄+b
CRtt̄+≥1c
CRtt̄+light
(5, 5) (5, 4) (5, 3) (5, 2) (4, 4) (4, 3) (4, 2) (3, 3) (3, 2) (2, 2) (5, 1) (4, 1) (3, 1) (2, 1) (1, 1)
(3rd, 4th) jet b-tagging discriminant


(1st, 2nd) jet b-tagging discriminant
Single Lepton, ≥ 6 j
(3, 3)
(4, 3)
(5, 3)
(4, 4)
(5, 4)
(5, 5)
SR1
SR2
SR3
CRtt̄+b
CRtt̄+≥1c
CRtt̄+light
(5, 5) (5, 4) (5, 3) (5, 2) (4, 4) (4, 3) (4, 2) (3, 3) (3, 2) (2, 2) (5, 1) (4, 1) (3, 1) (2, 1) (1, 1)
(3rd, 4th) jet b-tagging discriminant


(1st, 2nd) jet b-tagging discriminant
Dilepton, 3 j
(3, 3)
(4, 3)
(5, 3)
(4, 4)
(5, 4)
(5, 5)
CRtt̄+light
CRtt̄+≥1b
5 4 3 2 1
3rd jet b-tagging discriminant

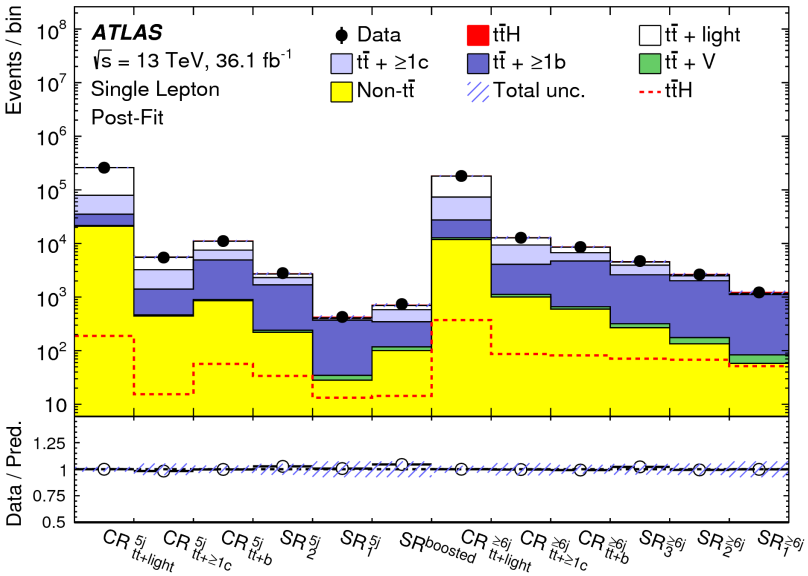

(1st, 2nd) jet b-tagging discriminant
Dilepton, ≥ 4 j
(3, 3)
(4, 3)
(5, 3)
(4, 4)
(5, 4)
(5, 5)
SR1
SR2
SR3
CRtt̄+≥1c
CRtt̄+light
(5, 5) (5, 4) (5, 3) (5, 2) (4, 4) (4, 3) (4, 2) (3, 3) (3, 2) (2, 2) (5, 1) (4, 1) (3, 1) (2, 1) (1, 1)
(3rd, 4th) jet b-tagging discriminant

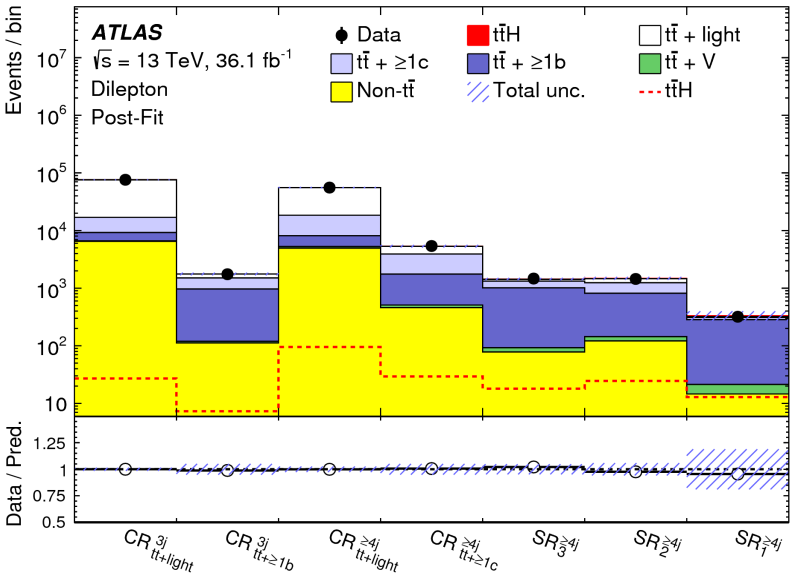
ttHbb

Variable	Definition	$SR^{\geq 6j}_{1,2,3}$	$SR^{5j}_{1,2}$
General kinematic variables			
ΔR^{avg}_{bb}	Average ΔR for all b-tagged jet pairs	✓	✓
$\Delta R^{max\ p_T}_{bb}$	ΔR between the two b-tagged jets with the largest vector sum p_T	✓	–
$\Delta\eta^{max}_{jj}$	Maximum $\Delta\eta$ between any two jets	✓	✓
$m^{min\ \Delta R}_{bb}$	Mass of the combination of two b-tagged jets with the smallest ΔR	✓	–
$m^{min\ \Delta R}_{jj}$	Mass of the combination of any two jets with the smallest ΔR	–	✓
$N^{Higgs\ 30}_{bb}$	Number of b-tagged jet pairs with invariant mass within 30 GeV of the Higgs-boson mass	✓	✓
H^{had}_T	Scalar sum of jet p_T	–	✓
$\Delta R^{min}_{\ell,bb}$	ΔR between the lepton and the combination of the two b-tagged jets with the smallest ΔR	–	✓
Aplanarity	$1.5\lambda_2$, where λ_2 is the second eigenvalue of the momentum tensor [100] built with all jets	✓	✓
H_1	Second Fox–Wolfram moment computed using all jets and the lepton	✓	✓
Variables from reconstruction BDT			
BDT output	Output of the reconstruction BDT	✓*	✓*
m^{Higgs}_{bb}	Higgs candidate mass	✓	✓
$m_{H,b_{lep\ top}}$	Mass of Higgs candidate and b-jet from leptonic top candidate	✓	–
ΔR^{Higgs}_{bb}	ΔR between b-jets from the Higgs candidate	✓	✓
$\Delta R_{H,t\bar{t}}$	ΔR between Higgs candidate and $t\bar{t}$ candidate system	✓*	✓*
$\Delta R_{H,lep\ top}$	ΔR between Higgs candidate and leptonic top candidate	✓	–
$\Delta R_{H,b_{had\ top}}$	ΔR between Higgs candidate and b-jet from hadronic top candidate	–	✓*
Variables from likelihood and matrix element method calculations			
LHD	Likelihood discriminant	✓	✓
MEM_{D1}	Matrix element discriminant (in $SR^{\geq 6j}_1$ only)	✓	–
Variables from b-tagging (not in $SR^{\geq 6j}_1$)			
$w^{Higgs}_{b\text{-tag}}$	Sum of b-tagging discriminants of jets from best Higgs candidate from the reconstruction BDT	✓	✓
B^3_{jet}	3rd largest jet b-tagging discriminant	✓	✓
B^4_{jet}	4th largest jet b-tagging discriminant	✓	✓
B^5_{jet}	5th largest jet b-tagging discriminant	✓	✓

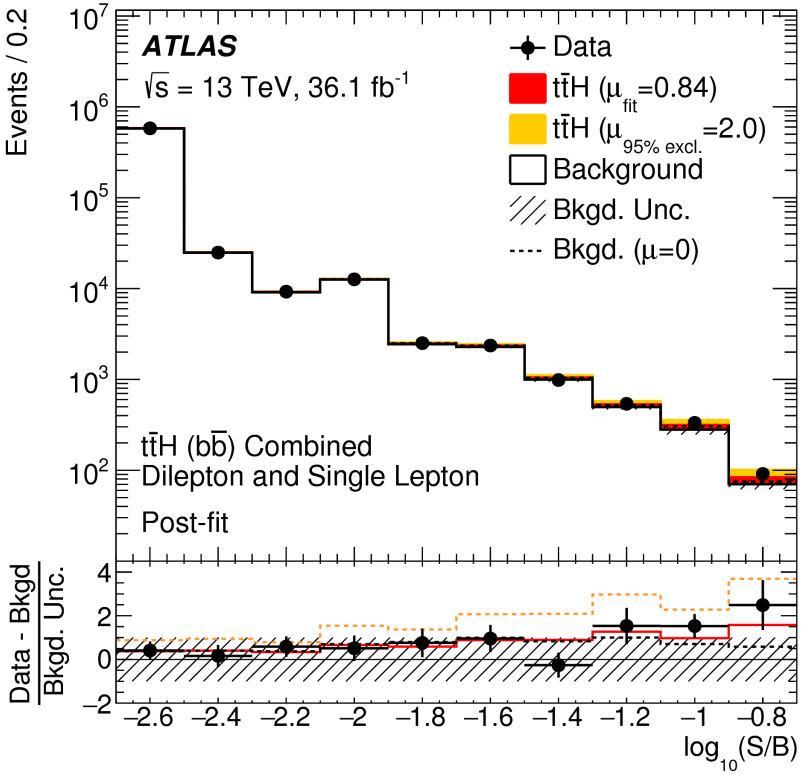
Variable	Definition	$SR^{\geq 4j}_1$	$SR^{\geq 4j}_2$	$SR^{\geq 4j}_3$
General kinematic variables				
m^{min}_{bb}	Minimum invariant mass of a b-tagged jet pair	✓	✓	-
m^{max}_{bb}	Maximum invariant mass of a b-tagged jet pair	-	-	✓
$m^{min\ \Delta R}_{bb}$	Invariant mass of the b-tagged jet pair with minimum ΔR	✓	-	✓
$m^{max\ p_T}_{jj}$	Invariant mass of the jet pair with maximum p_T	✓	-	-
$m^{max\ p_T}_{bb}$	Invariant mass of the b-tagged jet pair with maximum p_T	✓	-	✓
$\Delta\eta^{avg}_{bb}$	Average $\Delta\eta$ for all b-tagged jet pairs	✓	✓	✓
$\Delta\eta^{max}_{\ell,j}$	Maximum $\Delta\eta$ between a jet and a lepton	-	✓	✓
$\Delta R^{max\ p_T}_{bb}$	ΔR between the b-tagged jet pair with maximum p_T	-	✓	✓
$N^{Higgs\ 30}_{bb}$	Number of b-tagged jet pairs with invariant mass within 30 GeV of the Higgs-boson mass	✓	✓	-
$n^{p_T>40}_{jets}$	Number of jets with $p_T > 40$ GeV	-	✓	✓
$Aplanarity_{b\text{-jet}}$	$1.5\lambda_2$, where λ_2 is the second eigenvalue of the momentum tensor [100] built with all b-tagged jets	-	✓	-
H^{all}_T	Scalar sum of p_T of all jets and leptons	-	-	✓
Variables from reconstruction BDT				
BDT output	Output of the reconstruction BDT	✓**	✓**	✓
m^{Higgs}_{bb}	Higgs candidate mass	✓	-	✓
$\Delta R_{H,t\bar{t}}$	ΔR between Higgs candidate and $t\bar{t}$ candidate system	✓*	-	-
$\Delta R^{min}_{H,\ell}$	Minimum ΔR between Higgs candidate and lepton	✓	✓	✓
$\Delta R^{min}_{H,b}$	Minimum ΔR between Higgs candidate and b-jet from top	✓	✓	-
$\Delta R^{max}_{H,b}$	Maximum ΔR between Higgs candidate and b-jet from top	-	✓	-
ΔR^{Higgs}_{bb}	ΔR between the two jets matched to the Higgs candidate	-	✓	-
Variables from b-tagging				
$w^{Higgs}_{b\text{-tag}}$	Sum of b-tagging discriminants of jets from best Higgs candidate from the reconstruction BDT	-	✓	-

ttHbb

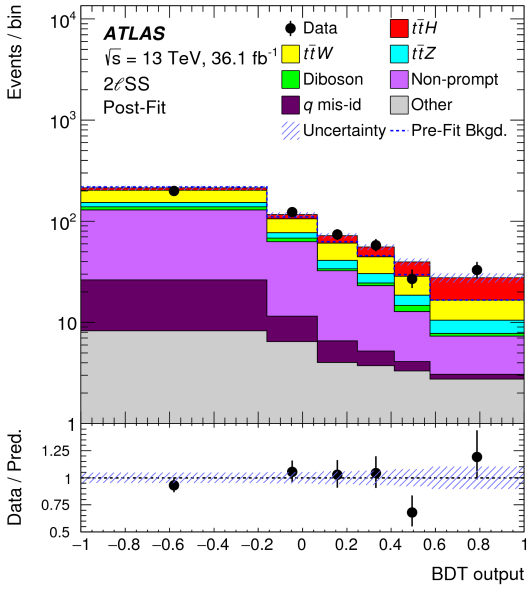

ATLAS
√s = 13 TeV, 36.1 fb-1
Single Lepton
Post-Fit
Data
tt̄H
tt̄ + light
tt̄ + ≥1c
tt̄ + ≥1b
tt̄ + V
Non-tt̄
Total unc.
tt̄H
Events / bin
Data / Pred.
CR 5j tt+light
CR 5j tt+≥1c
CR 5j tt+b
SR 5j 2
SR 5j 1
SR boosted
CR ≥6j tt+light
CR ≥6j tt+≥1c
CR ≥6j tt+b
SR ≥6j 3
SR ≥6j 2
SR ≥6j 1

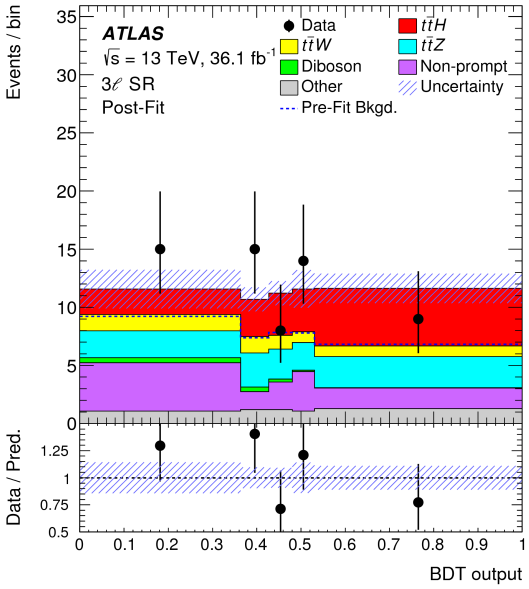

ATLAS
√s = 13 TeV, 36.1 fb-1
Dilepton
Post-Fit
Data
tt̄H
tt̄ + light
tt̄ + ≥1c
tt̄ + ≥1b
tt̄ + V
Non-tt̄
Total unc.
tt̄H
Events / bin
Data / Pred.
CR 3j tt+light
CR 3j tt+≥1b
CR ≥4j tt+light
CR ≥4j tt+≥1c
SR ≥4j 3
SR ≥4j 2
SR ≥4j 1

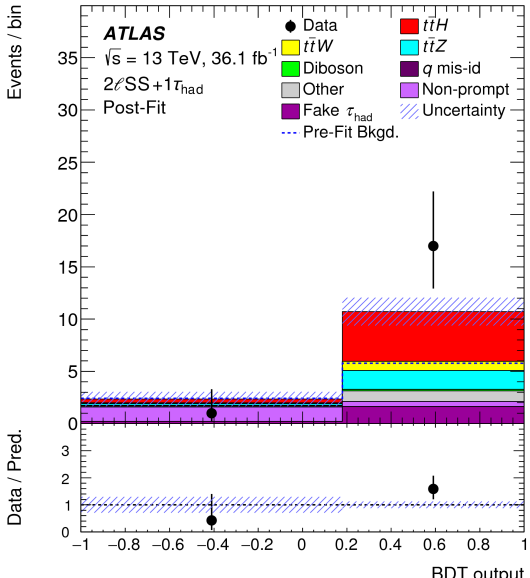
ttHbb

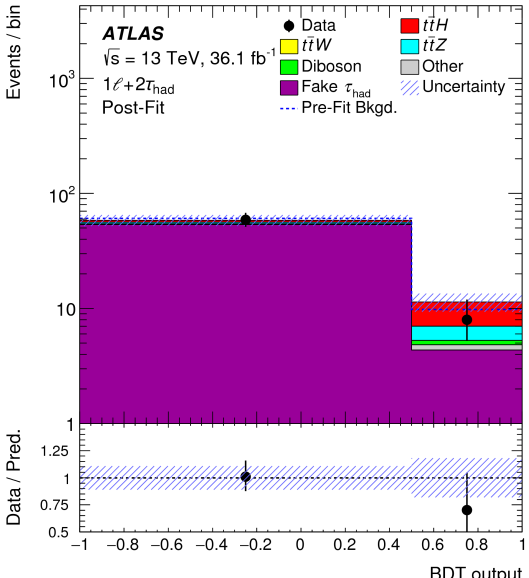

ATLAS
√s = 13 TeV, 36.1 fb-1
Data
tt̄H (μfit=0.84)
tt̄H (μ95% excl.=2.0)
Background
Bkgd. Unc.
Bkgd. (μ=0)
tt̄H (bb̄) Combined
Dilepton and Single Lepton
Post-fit
Events / 0.2
Data - Bkgd / Bkgd. Unc.
log10(S/B)

ttH multilepton: BDT output



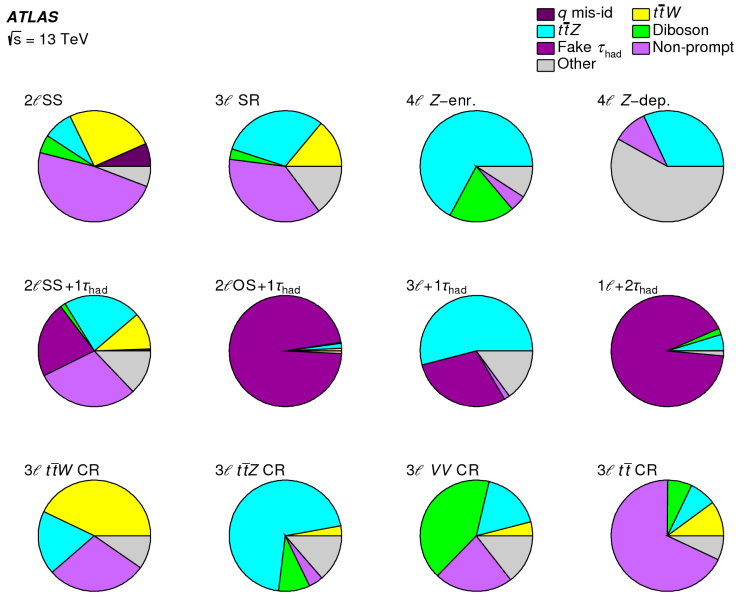




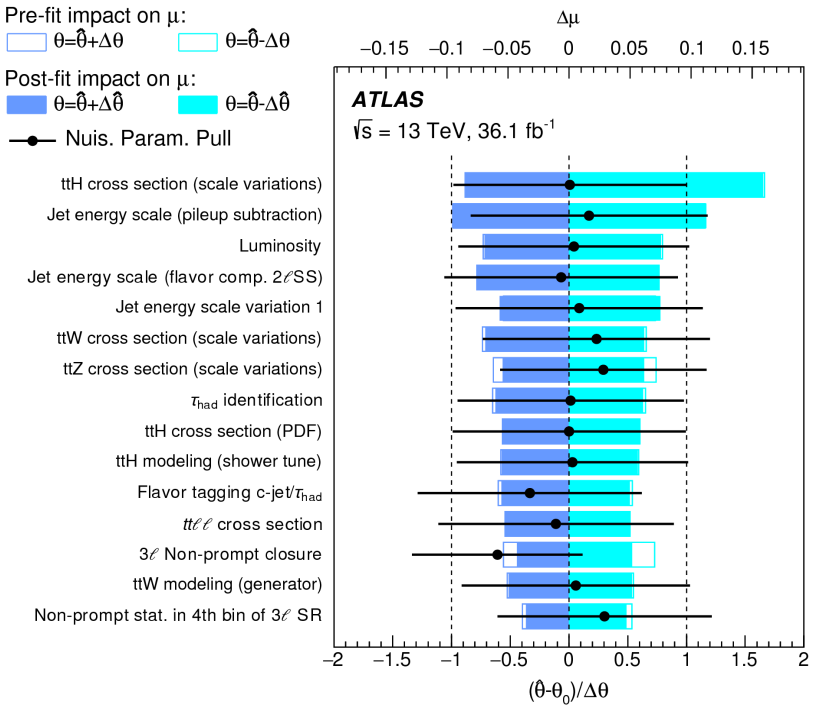


ttH multilepton

Channel	Selection criteria
Common	$N_{\text{jets}} \geq 2$ and $N_{b\text{-jets}} \geq 1$
2ℓSS	Two very tight light leptons with $p_{\text{T}} > 20$ GeV Same-charge light leptons Zero medium τ_{had} candidates $N_{\text{jets}} \geq 4$ and $N_{b\text{-jets}} < 3$
3ℓ	Three light leptons with $p_{\text{T}} > 10$ GeV; sum of light-lepton charges ± 1 Two same-charge leptons must be very tight and have $p_{\text{T}} > 15$ GeV The opposite-charge lepton must be loose, isolated and pass the non-prompt BDT Zero medium τ_{had} candidates $m(\ell^+\ell^-) > 12$ GeV and $\lvert m(\ell^+\ell^-) - 91.2\text{ GeV}\rvert > 10$ GeV for all SFOC pairs $\lvert m(3\ell) - 91.2\text{ GeV}\rvert > 10$ GeV
4ℓ	Four light leptons; sum of light-lepton charges 0 Third and fourth leading leptons must be tight $m(\ell^+\ell^-) > 12$ GeV and $\lvert m(\ell^+\ell^-) - 91.2\text{ GeV}\rvert > 10$ GeV for all SFOC pairs $\lvert m(4\ell) - 125\text{ GeV}\rvert > 5$ GeV Split 2 categories: Z-depleted (0 SFOC pairs) and Z-enriched (2 or 4 SFOC pairs)
1ℓ+$2\tau_{\text{had}}$	One tight light lepton with $p_{\text{T}} > 27$ GeV Two medium τ_{had} candidates of opposite charge, at least one being tight $N_{\text{jets}} \geq 3$
2ℓSS+$1\tau_{\text{had}}$	Two very tight light leptons with $p_{\text{T}} > 15$ GeV Same-charge light leptons One medium τ_{had} candidate, with charge opposite to that of the light leptons $N_{\text{jets}} \geq 4$ $\lvert m(ee) - 91.2\text{ GeV}\rvert > 10$ GeV for ee events
2ℓOS+$1\tau_{\text{had}}$	Two loose and isolated light leptons with $p_{\text{T}} > 25$, 15 GeV One medium τ_{had} candidate Opposite-charge light leptons One medium τ_{had} candidate $m(\ell^+\ell^-) > 12$ GeV and $\lvert m(\ell^+\ell^-) - 91.2\text{ GeV}\rvert > 10$ GeV for the SFOC pair $N_{\text{jets}} \geq 3$
3ℓ+$1\tau_{\text{had}}$	3ℓ selection, except: One medium τ_{had} candidate, with charge opposite to the total charge of the light leptons The two same-charge light leptons must be tight and have $p_{\text{T}} > 10$ GeV The opposite-charge light lepton must be loose and isolated



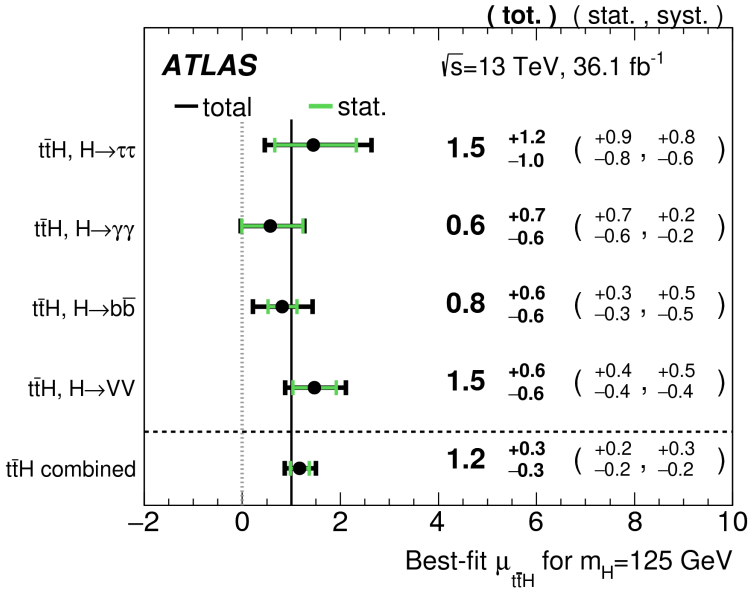
ttH multilepton



Systematic uncertainty	Type	Components
Luminosity	N	1
Pileup reweighting	SN	1
Physics Objects		
Electron	SN	6
Muon	SN	15
τ_{had}	SN	10
Jet energy scale and resolution	SN	28
Jet vertex fraction	SN	1
Jet flavor tagging	SN	126
$E_{\text{T}}^{\text{miss}}$	SN	3
Total (Experimental)	–	191
Data-driven non-prompt/fake leptons and charge misassignment		
Control region statistics	SN	38
Light-lepton efficiencies	SN	22
Non-prompt light-lepton estimates: non-closure	N	5
γ-conversion fraction	N	5
Fake τ_{had} estimates	N/SN	12
Electron charge misassignment	SN	1
Total (Data-driven reducible background)	–	83
$t\bar{t}H$ **modeling**		
Cross section	N	2
Renormalization and factorization scales	S	3
Parton shower and hadronization model	SN	1
Higgs boson branching fraction	N	4
Shower tune	SN	1
$t\bar{t}W$ **modeling**		
Cross section	N	2
Renormalization and factorization scales	S	3
Matrix-element MC event generator	SN	1
Shower tune	SN	1
$t\bar{t}Z$ **modeling**		
Cross section	N	2
Renormalization and factorization scales	S	3
Matrix-element MC event generator	SN	1
Shower tune	SN	1
Other background modeling		
Cross section	N	15
Shower tune	SN	1
Total (Signal and background modeling)	–	41
Total (Overall)	–	315

Combination

- Different acceptances following Hggs decays in mutilepton channels allow for independent μ fit for each Higgs decay



Combination