Associated production of ttH with ATLAS

Silvia Biondi on behalf of the ATLAS Collaboration

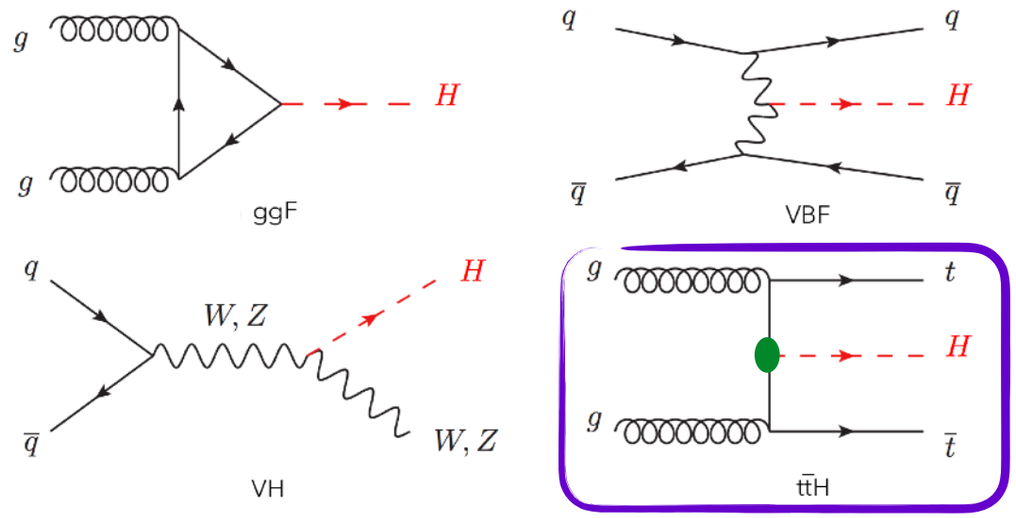
University & INFN of Bologna

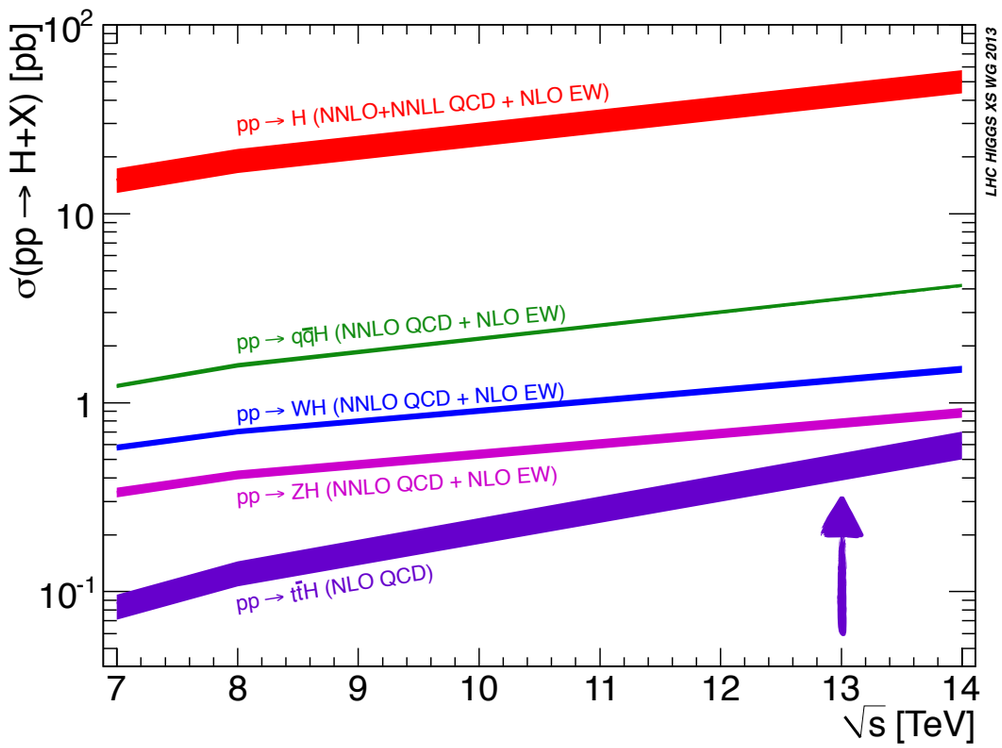
Higgs Hunting conference, Orsay, Paris
24.07.2018

ttH and motivations

Why ttH?

- Due to its short lifetime (10^{-24} s), the top quark **decays before hadronizing**;
  - Unique opportunity to study properties of a bare quark;
- high m_t (173.5 GeV) implies a **large Yukawa coupling with the Higgs boson** (~1), wrt other couplings ($<10^{-2}$).
  - Precise measurement of y_t ($\simeq\sqrt{2}m_{top}/v$) probes BSM contributions in Higgs boson production and decay loops.
- ttH production allows for **direct measurement of top Yukawa**, but very challenging to detect:
  - tiny ttH cross section (O(0.5) pb at 13 TeV), two orders of magnitude smaller than ggF,
  - many and complex final states and large irreducible backgrounds, eg ttbb and ttV.



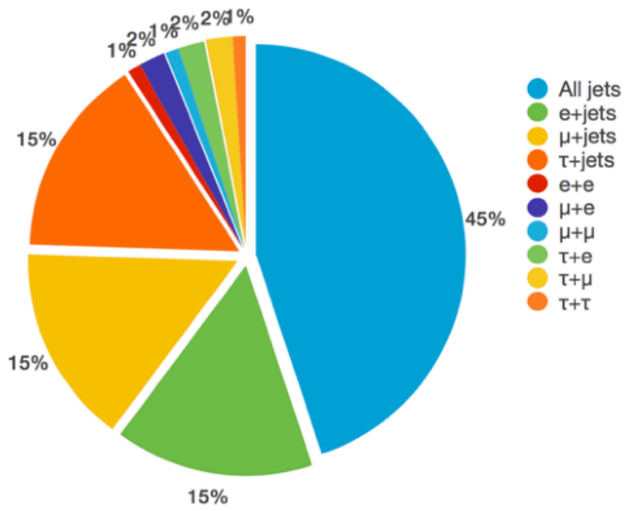


The **highest cross section increase as a function of energy** wrt other production modes

First observation of ttH in ATLAS, in 2018, using Run-2 data!

ttH signatures

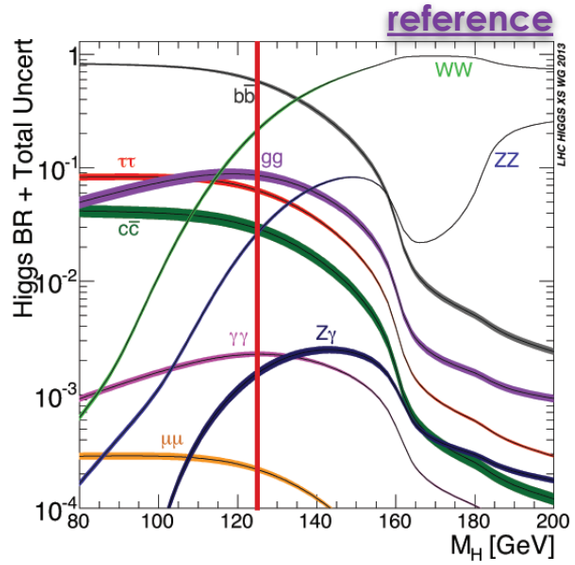
Top pair branching ratio



+

Higgs branching ratio

reference

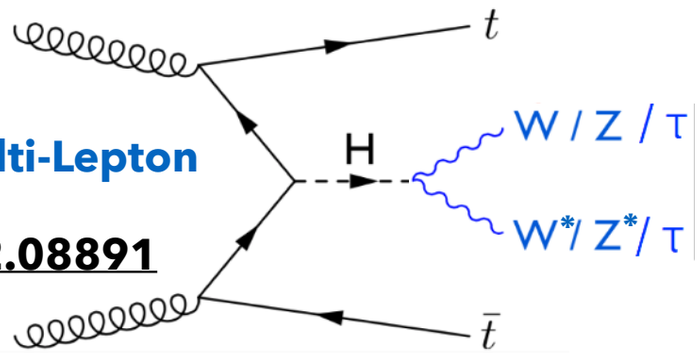


in this presentation...

13 TeV, 36.1 fb^{-1}

ttH Multi-Lepton

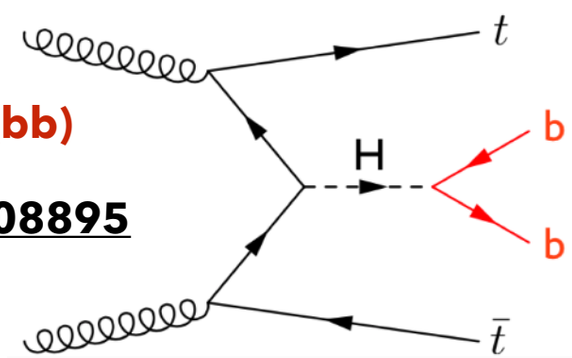
1712.08891



13 TeV, 36.1 fb^{-1}

ttH(bb)

1712.08895



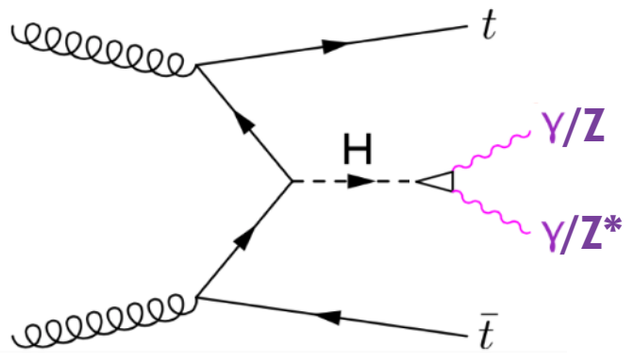
13 TeV, 79.8 fb^{-1}

ttH in H->γγ

1806.00425

ttH(4l) in H->4l

1806.00425



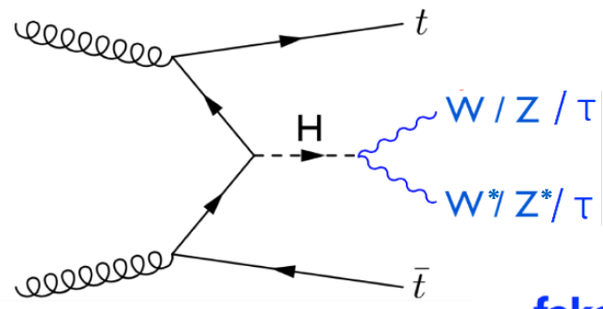
13 TeV, 36.1 - 79.8 fb^{-1}

ttH combination

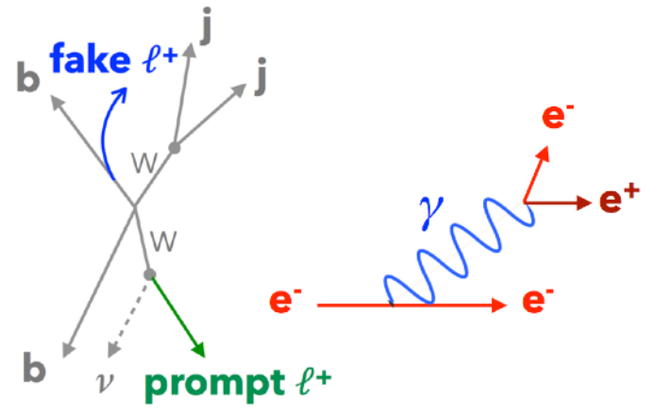
1806.00425

ttH Multi-Lepton: strategy

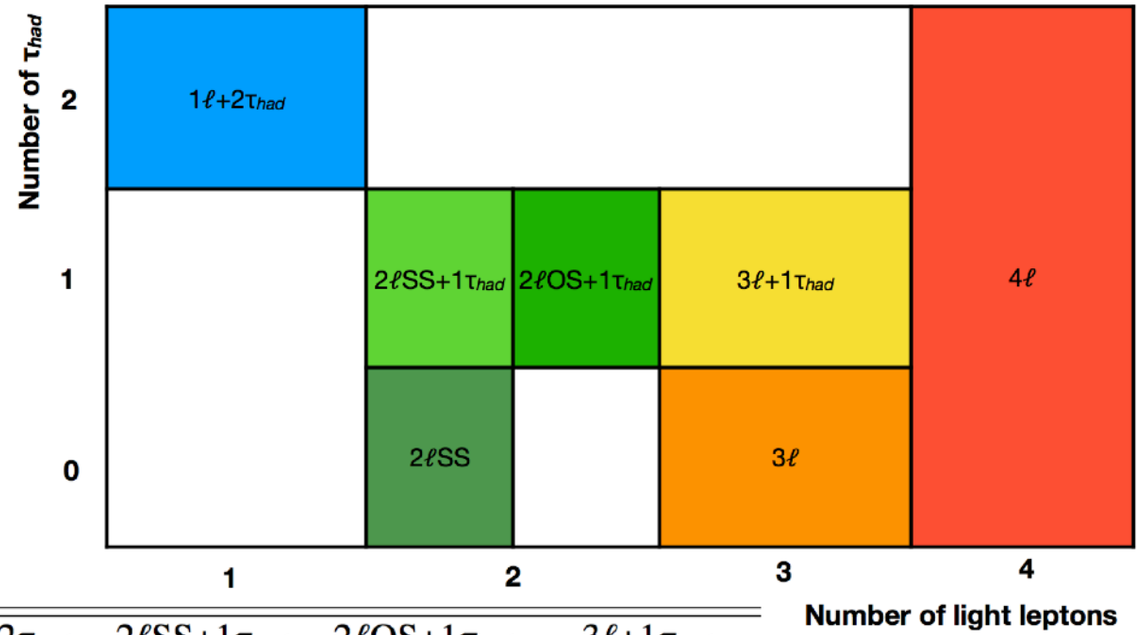
- Targeted Higgs decays:
  WW*(→lνlν,lνqq),ττ,ZZ*(→llνν,llqq)
  - 2 same-sign or ≥3 charged leptons (e,μ,τ_{had}) to reject tt events
- 7 final states categorised in number and flavour of charged leptons
  - additional requirements on (b-)jet multiplicities
- Almost all background events from:
  - **ttW and ttZ**, irreducible: same final state
    - good agreement observed in dedicated control regions
  - **tt**, reducible: mainly non-prompt leptons from Heavy-Flavour decay and charge mis-id
- **MVA lepton selections** to reject non-prompt and charge mis-id backgrounds
- Event classified with **MVA approaches**.



signal



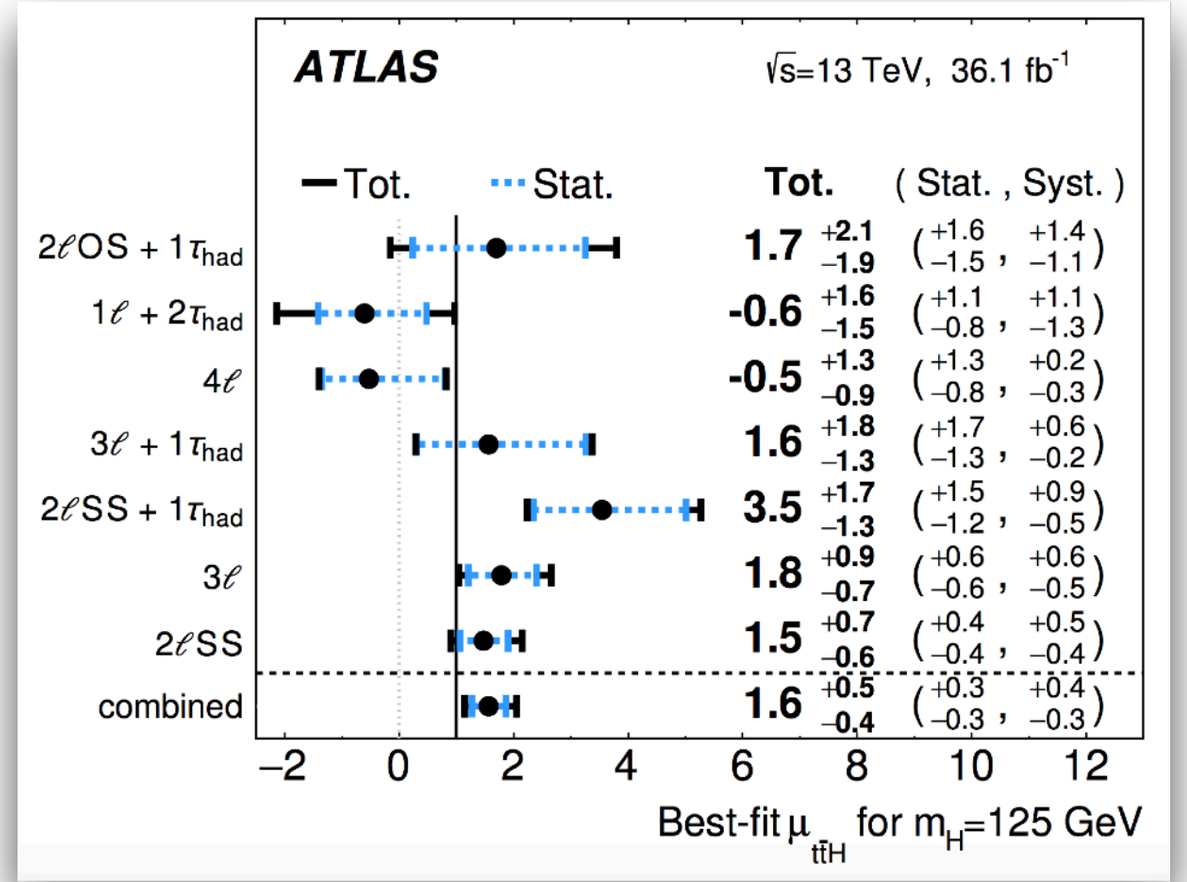
background



	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	-	-	-	-	-

ttH Multi-Lepton: results

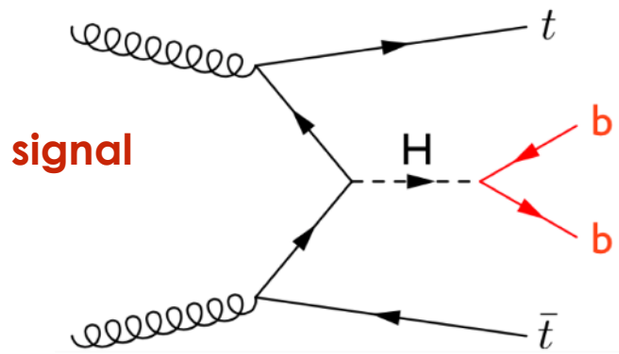
- Obs (exp) excess of **4.1 (2.8)** σ over SM background
- Alternative fit with unconstrained ttW, ttZ normalisations:
  - same best-fit μ_{ttH} with 15% larger errors
  - **$\mu_{ttW} = 0.92 \pm 0.32$, $\mu_{ttZ} = 1.17^{+0.25}_{-0.22}$**
- **Dominant systematics:**
  - ttH and ttV modelling
  - non-prompt background estimate
- Many categories still **statistically limited**

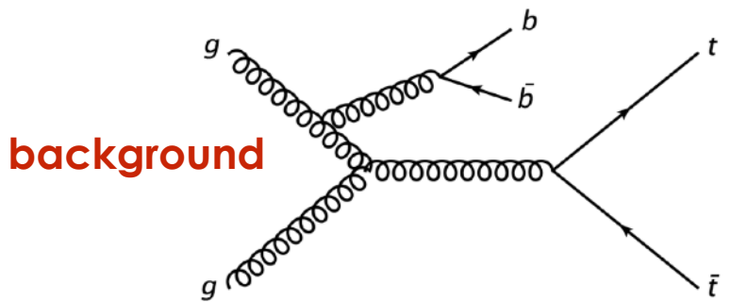


Uncertainty Source	$\Delta\mu$	
$t\bar{t}H$ modeling (cross section)	+0.20	−0.09
Jet energy scale and resolution	+0.18	−0.15
Non-prompt light-lepton estimates	+0.15	−0.13
Jet flavor tagging and τ_{had} identification	+0.11	−0.09
$t\bar{t}W$ modeling	+0.10	−0.09
$t\bar{t}Z$ modeling	+0.08	−0.07
Other 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

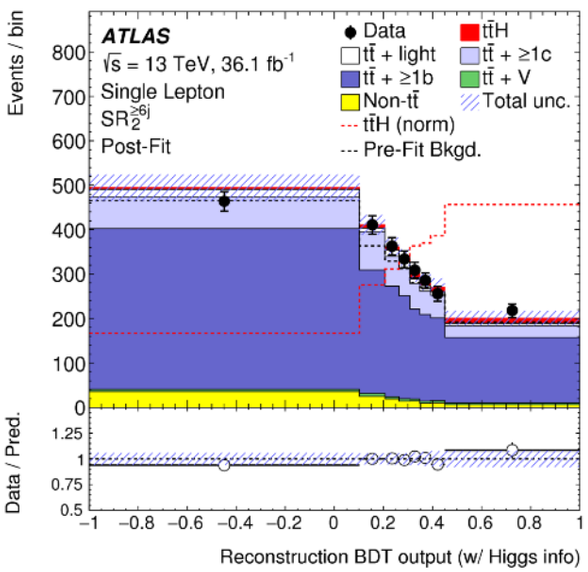
ttH(bb): strategy

- **Single lepton** and **dilepton** channels depending on top decay
- Events categorised by lepton, jets and b-tagged jets multiplicities
  - first time of a single lepton **boosted** category.
- Overall, tt events are 85-96% of total background
  - large irreducible component from **tt + Heavy Flavour (HF)** quarks
- MVA classifiers in Signal Regions (SR):
  - to reconstruct Higgs and top candidates from high combinatorics of (b-)jets (**RecoBDT, LHD, MEM**), only in resolved channels;
  - to classify signal vs backgrounds events (**ClassificationBDT**), both in resolved and boosted channels, included in the fit procedure.

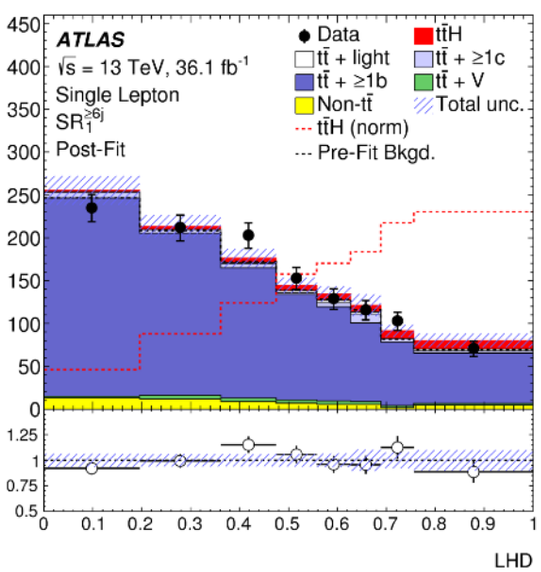




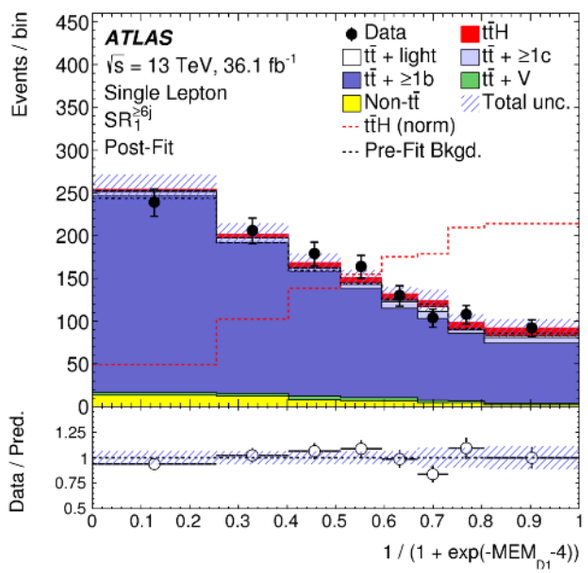
RecoBDT



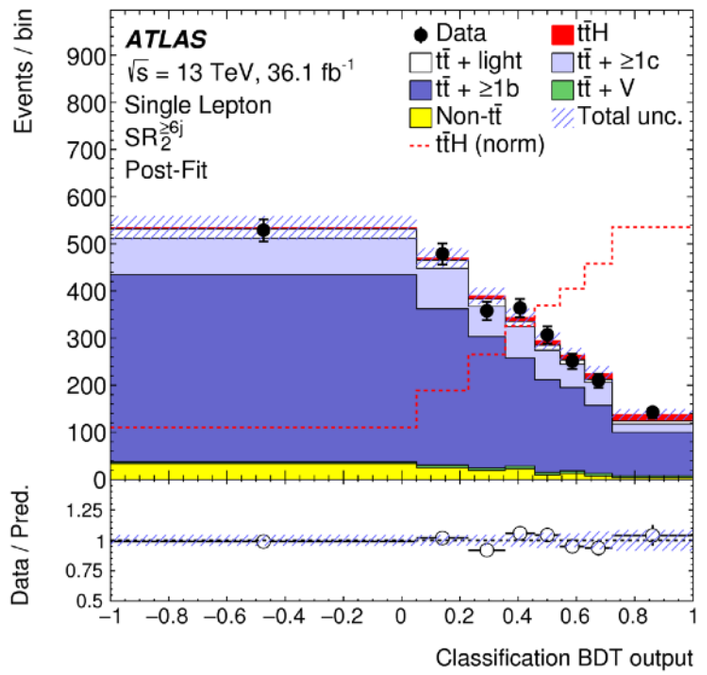
LHD



MEM

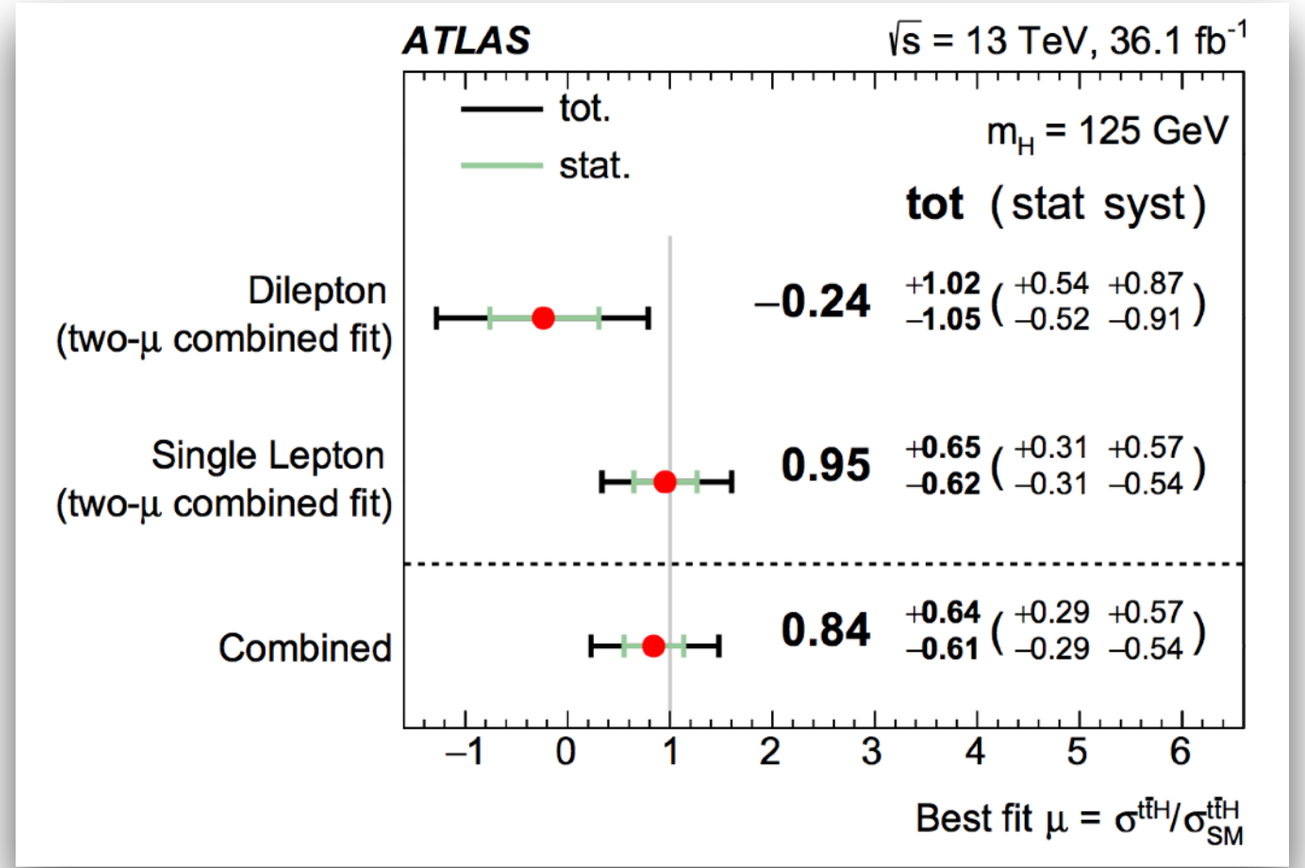


ClassificationBDT



ttH(bb): results

- Combined fit of classification BDT output in SRs and event yields in most of CRs
- Free-floating normalisation factors for tt+HF:
  - **tt+≥1b: 1.24±0.10**
  - **tt+≥1c: 1.63±0.23**
- Obs (exp) excess of **1.4 (1.6)** σ over SM background
- Best-fit $\mu_{ttH} = \sigma_{ttH}/\sigma_{SM} = 0.84^{+0.64}_{-0.61}$
- Precision limited by systematic uncertainty on tt+≥1b simulation.



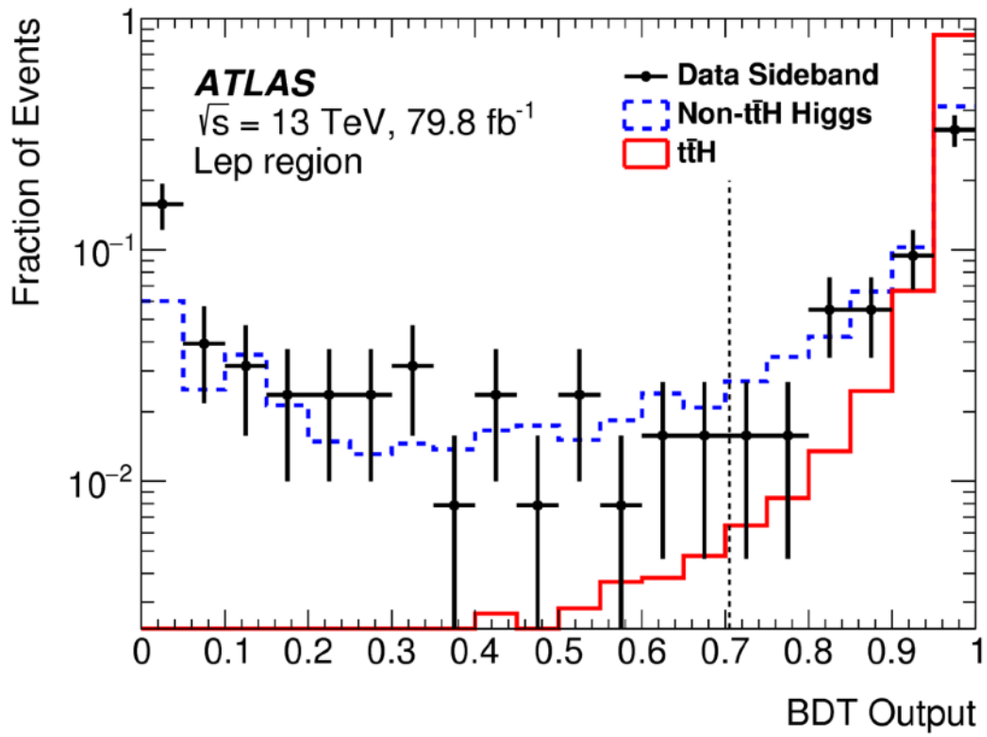
Uncertainty source	$\Delta\mu$	
$t\bar{t}$ + ≥1b modelling	+0.46	−0.46
Background model statistics	+0.29	−0.31
Jet flavour tagging	+0.16	−0.16
Jet energy scale and resolution	+0.14	−0.14
$t\bar{t}H$ modelling	+0.22	−0.05
$t\bar{t}$ + ≥1c modelling	+0.09	−0.11
Jet-vertex association, pileup modelling	+0.03	−0.05
Other background modelling	+0.08	−0.08
$t\bar{t}$ + light modelling	+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
$t\bar{t}$ + ≥1b normalisation	+0.09	−0.10
$t\bar{t}$ + ≥1c normalisation	+0.02	−0.03
Statistical uncertainty	+0.29	−0.29
Total uncertainty	+0.64	−0.61

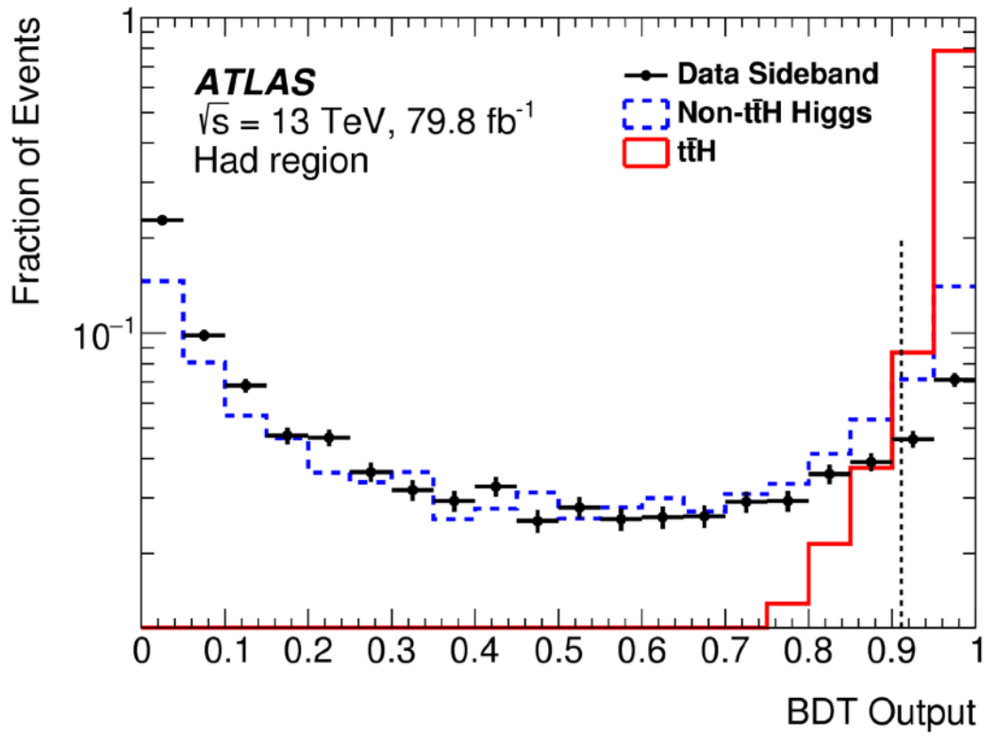
ttH(γγ): strategy

- **2 signal regions** targeting ttH production:
  - Leptonic: ≥1l (semi-leptonic top-quark decay)
  - Hadronic: ≥2jets + 0 isolated leptons (hadronic top-quark decay).

- **Main background** processes:
  - non-resonant diphoton production
  - non-ttH production

- **2 BDTs** trained to discriminate the ttH signal from the main background
  - lep BDT: p_T, η, φ and E of up to four (two) leading jets (leptons) in $p_T + E_T^{miss}$ and photons related observables
  - had BDT: p_T, η, φ, E and b-tagging decision of up to six leading jets + E_T^{miss} and photons related observables





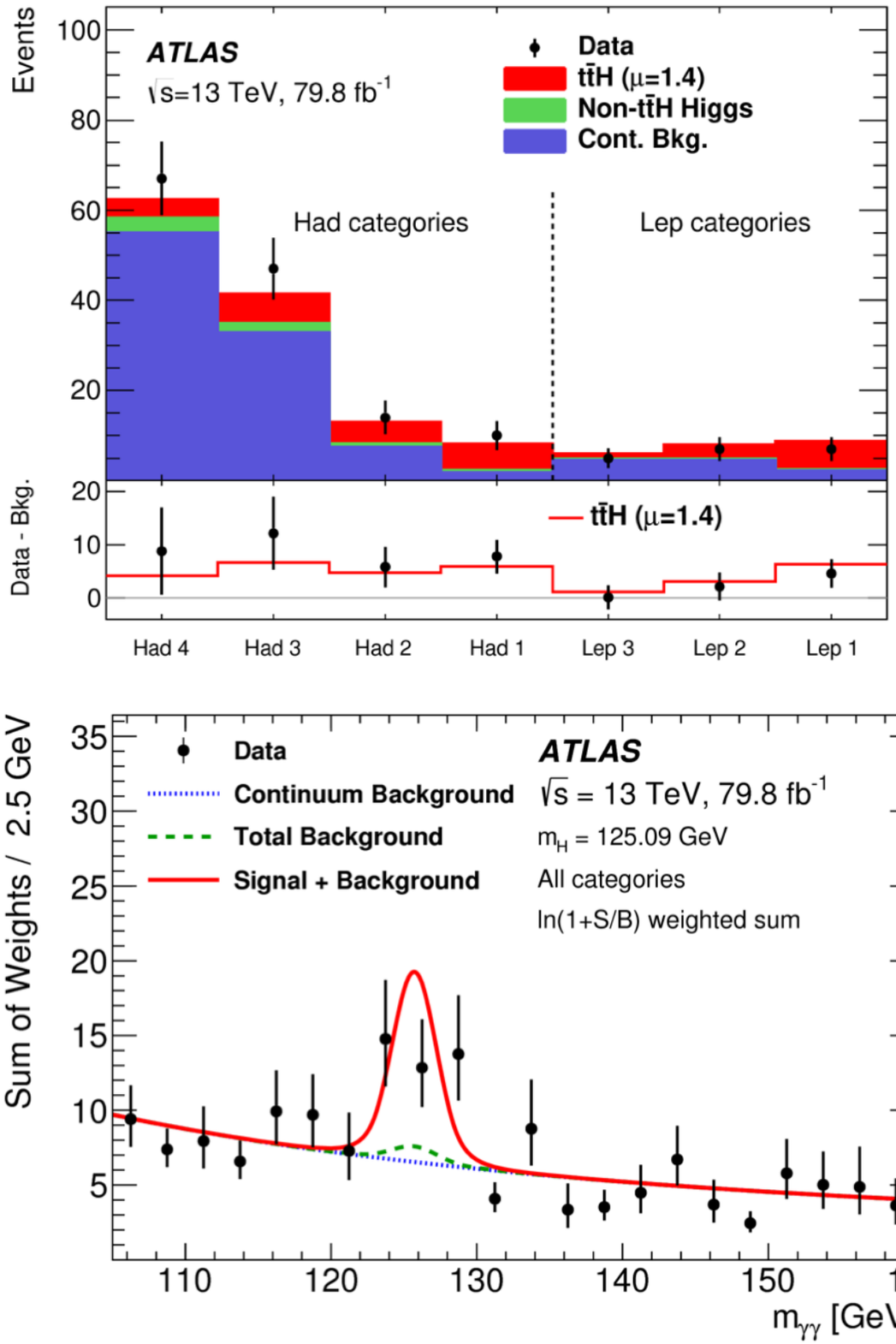
ttH(γγ): results

- 85% (97%) of ttH events are selected and 89% (43%) of background events are rejected in the Had (Lep) region
  - remaining events categorised into four (three) bins in the Had (Lep) region

- ttH signal extracted by a **combined unbinned maximum-likelihood fit** to the diphoton invariant mass spectrum ($105 < m_{\gamma\gamma} < 160$ GeV):
  - functional form to model the Higgs production modes from simulated $m_{\gamma\gamma}$ distributions;
  - functional form to model the background from simulation (Lep) and a dedicated data control region (Had);
  - parameters of the continuum background model left free in the fit.

- **Main theoretical uncertainty:** parton-shower modelling in ttH simulation (P8 vs HW7)

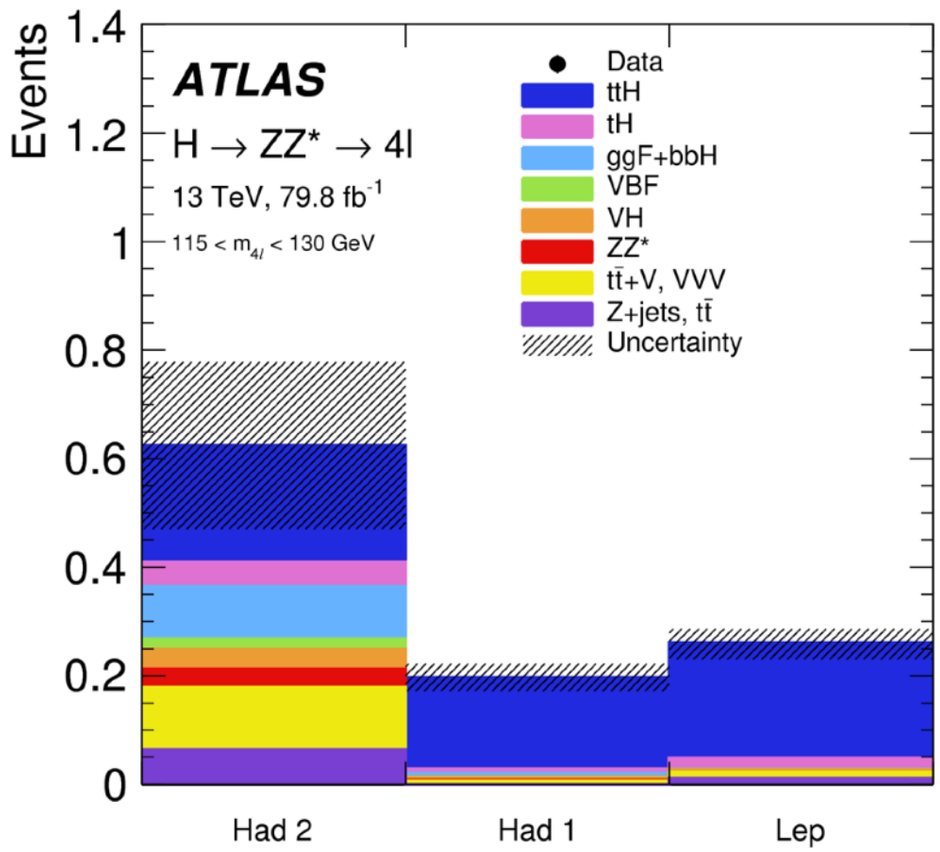
- About **50% more sensitive than the one @36 fb^{-1}**



ttH(4l): strategy and results

- **2 signal regions** targeting ttH production:
  - Leptonic: ≥4l + ≥1jet (≥1 b-jet) +≥1 additional lepton (semi-leptonic top-quark decay)
  - Hadronic: ≥4l + ≥3jets (≥1 b-jet) + 0 additional leptons (hadronic top-quark decay).

- **Main background** processes:
  - ttW, ttZ
  - non-ttH production



- BDT trained in the "Had" region for a **classification in two BDT bins:**
  - 11 observables: invariant mass, dijet p_T, $\Delta\eta$ of the two leading jets, E_T^{miss}, etc…

- Higgs boson candidate with **$115<m_{4l}<130$ GeV**

	Expected				Observed
Bin	$t\bar{t}H$ (signal)	Non-$t\bar{t}H$ Higgs	Non-Higgs	Total	Total
Had 1	0.169(31)	0.021(7)	0.008(8)	0.198(33)	0
Had 2	0.216(32)	0.20(9)	0.22(12)	0.63(16)	0
Lep	0.212(31)	0.0256(23)	0.015(13)	0.253(34)	0

ttH combination

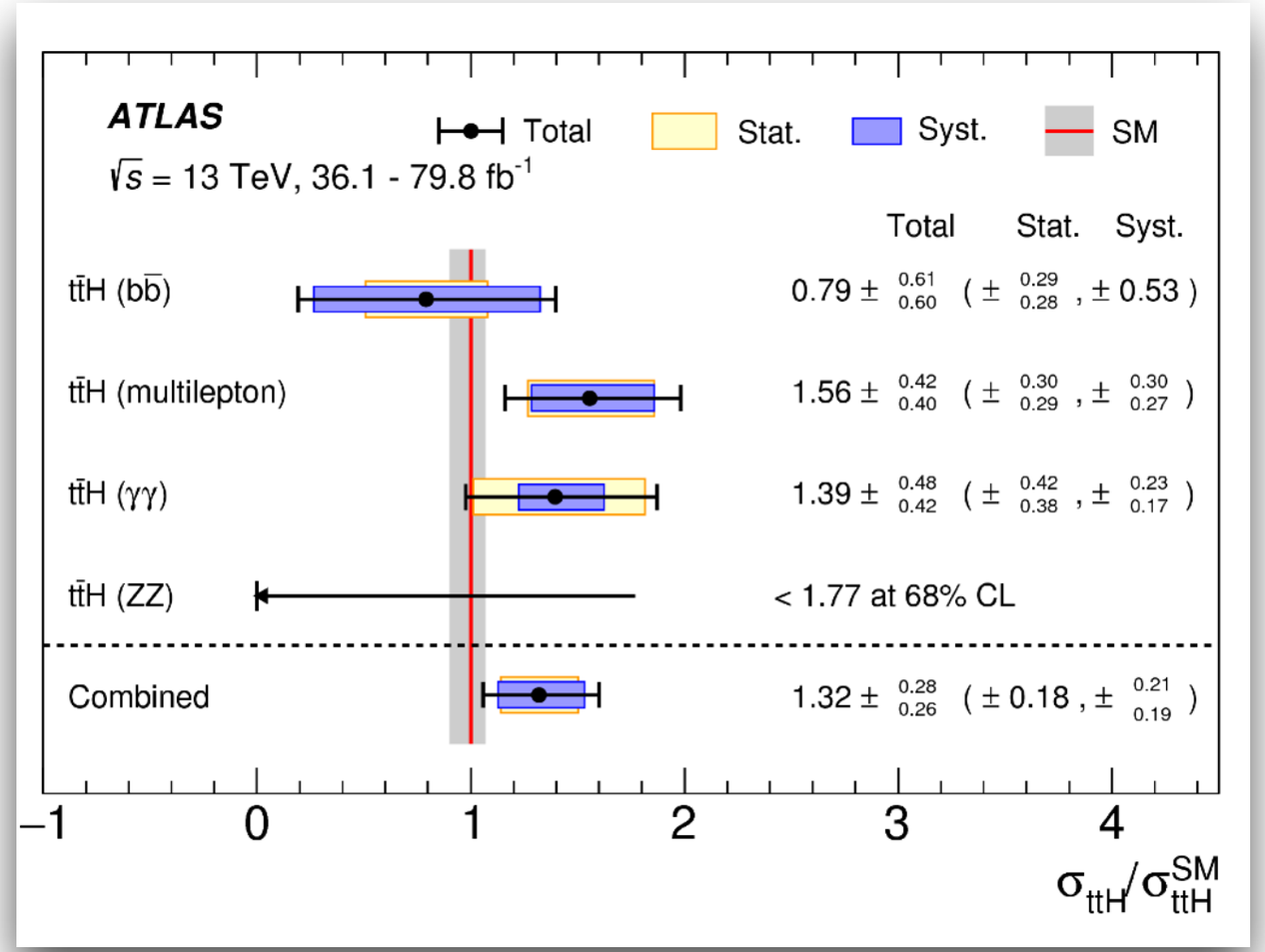
- Combination of ttH searches in **H→ γγ and H→ 4l** (79.8 fb^{-1}) with **H→ bb** and **H→ multi lepton** (36.1 fb^{-1})
- **Profile likelihood method**, based on simultaneous fits to the signal regions and control regions of the individual analyses
- The overlap between the selected events in the different analyses is found to be negligible
- **Correlation scheme** studied in detail

- H → γγ and H → 4l analyses employ improved reconstruction software compared with the H → bb and multi lepton analyses

Analysis	Integrated luminosity [fb^{-1}]	$t\bar{t}H$ cross section [fb]	Obs. sign.	Exp. sign.
$H \to \gamma\gamma$	79.8	$710\ ^{+210}_{-190}$ (stat.) $^{+120}_{-90}$ (syst.)	4.1 σ	3.7 σ
$H \to$ multilepton	36.1	790 ± 150 (stat.) $^{+150}_{-140}$ (syst.)	4.1 σ	2.8 σ
$H \to b\bar{b}$	36.1	$400\ ^{+150}_{-140}$ (stat.) $\pm$ 270 (syst.)	1.4 σ	1.6 σ
$H \to ZZ^* \to 4\ell$	79.8	<900 (68% CL)	0 σ	1.2 σ
Combined (13 TeV)	36.1−79.8	670 ± 90 (stat.) $^{+110}_{-100}$ (syst.)	5.8 σ	4.9 σ
Combined (7, 8, 13 TeV)	4.5, 20.3, 36.1−79.8	−	6.3 σ	5.1 σ

ttH combination: results

- ttH(γγ) and ttH(4l) **statistically limited**;
- ttH(bb) and ttH(ML) **limited by systematic uncertainties**, mostly theoretical uncertainties.
- Main uncertainty sources:
  - signal and background (tt+HF) modelling

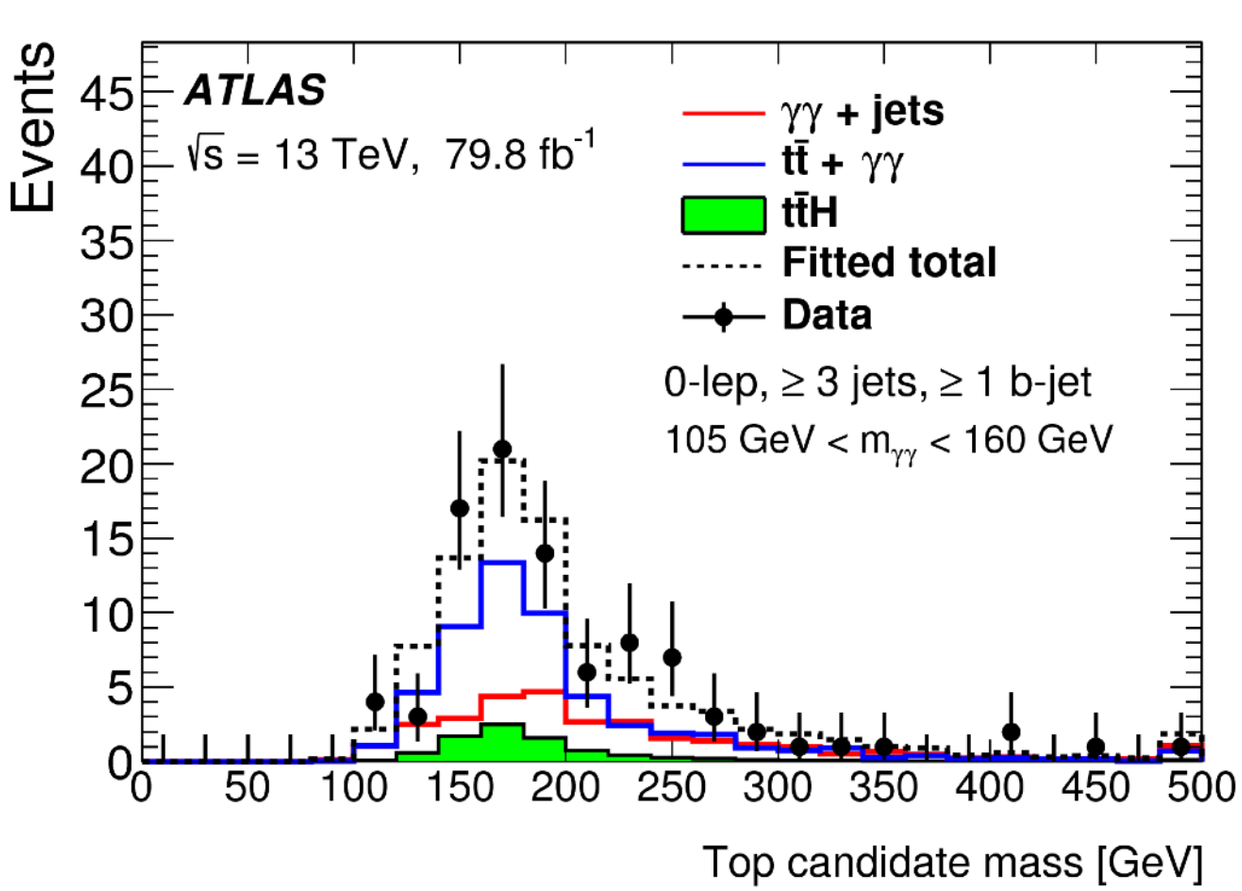


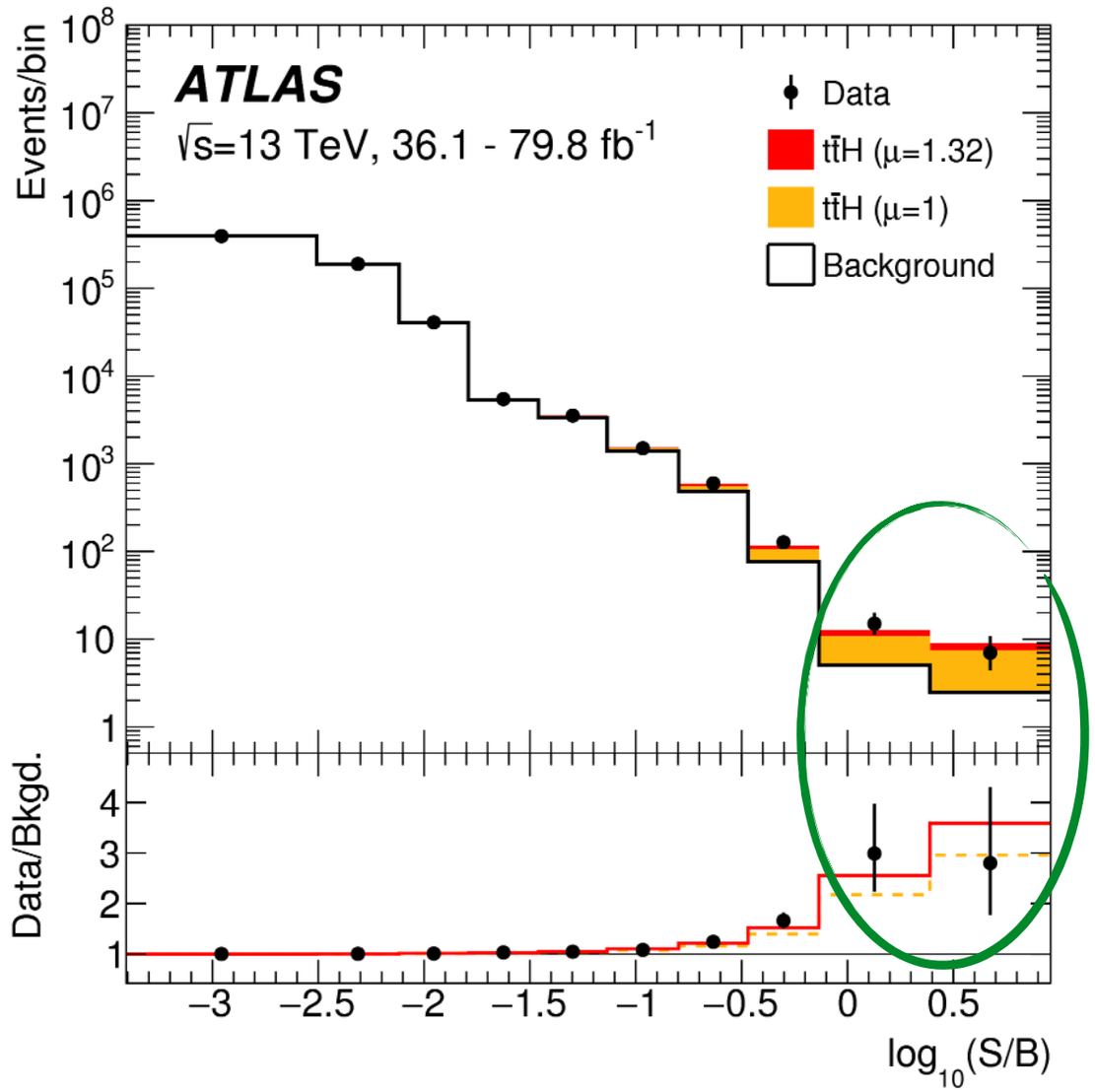
Uncertainty source	$\Delta\sigma_{t\bar{t}H}/\sigma_{t\bar{t}H}$ [%]
Theory uncertainties (modelling)	11.9
$t\bar{t}$ + heavy flavour	9.9
$t\bar{t}H$	6.0
Non-$t\bar{t}H$ Higgs boson production modes	1.5
Other background processes	2.2
Experimental uncertainties	9.3
Fake leptons	5.2
Jets, $E_{\mathrm{T}}^{\mathrm{miss}}$	4.9
Electrons, photons	3.2
Luminosity	3.0
τ-lepton	2.5
Flavour tagging	1.8
MC statistical uncertainties	4.4

- **ttH cross section 1.32 times higher wrt the SM one**
  - still compatible in 20% of the measurement precision
- Significance of ttH production at **5.8σ** (exp 4.9σ) .

ttH combination: results

- ttH(γγ) now can **reconstruct the top quark candidate** for events in the two Had bins with the highest S/B
  - reconstruction from a triplet of jets selected by a dedicated BDT, specifically trained to identify jets originating from the same top quark decays
- **A clear ttH signal-like excess over the background is visible for high $\log_{10}$(S/B).**



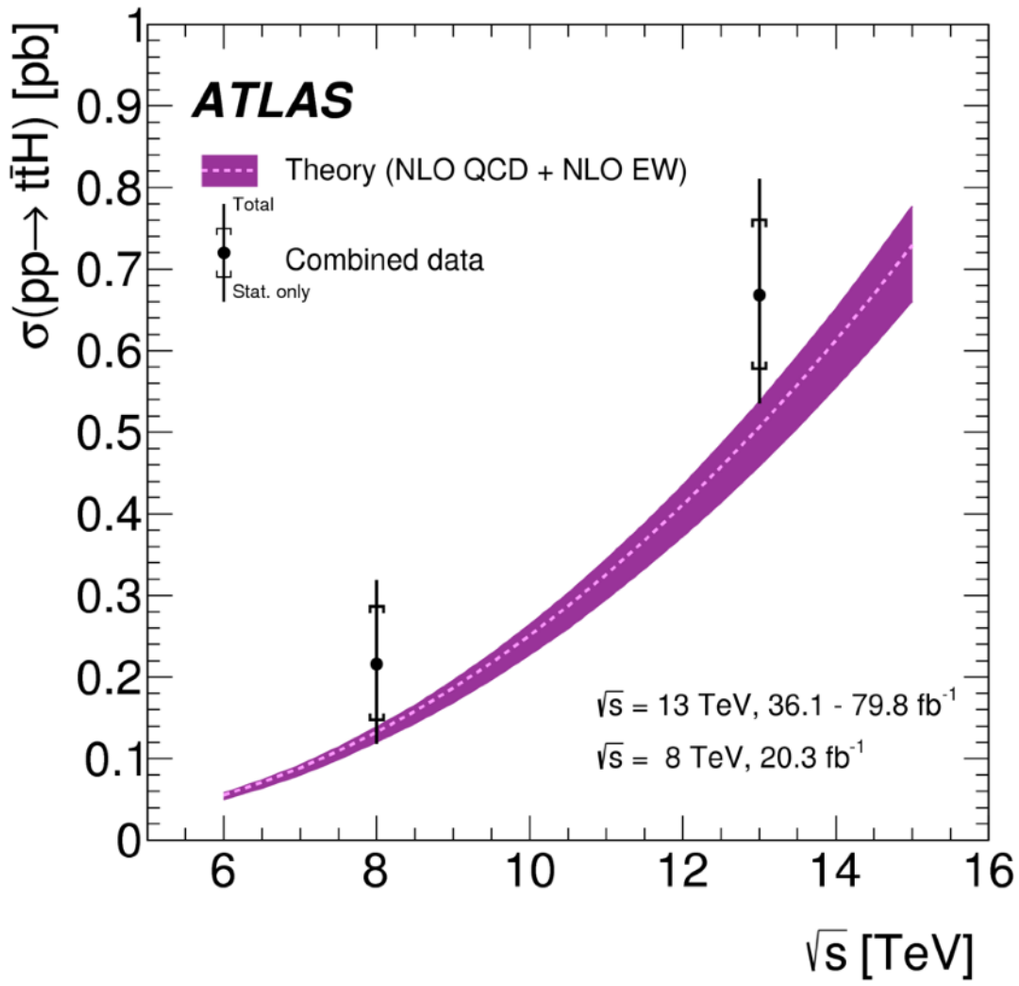


ttH combination with Run-1

Combination with measurements @7TeV (4.5 fb^{-1}) and @8TeV (20.3 fb^{-1})*:

6.3σ (exp 5.1σ)

* Eur. Phys. J. C 76 (2016) 6



Analysis	Integrated luminosity [fb^{-1}]	$t\bar{t}H$ cross section [fb]	Obs. sign.	Exp. sign.
$H \to \gamma\gamma$	79.8	$710\ ^{+210}_{-190}$ (stat.) $^{+120}_{-90}$ (syst.)	4.1 σ	3.7 σ
$H \to$ multilepton	36.1	790 ± 150 (stat.) $^{+150}_{-140}$ (syst.)	4.1 σ	2.8 σ
$H \to b\bar{b}$	36.1	$400\ ^{+150}_{-140}$ (stat.) $\pm$ 270 (syst.)	1.4 σ	1.6 σ
$H \to ZZ^* \to 4\ell$	79.8	<900 (68% CL)	0 σ	1.2 σ
Combined (13 TeV)	36.1–79.8	670 ± 90 (stat.) $^{+110}_{-100}$ (syst.)	5.8 σ	4.9 σ
Combined (7, 8, 13 TeV)	4.5, 20.3, 36.1–79.8	–	6.3 σ	5.1 σ

Conclusions

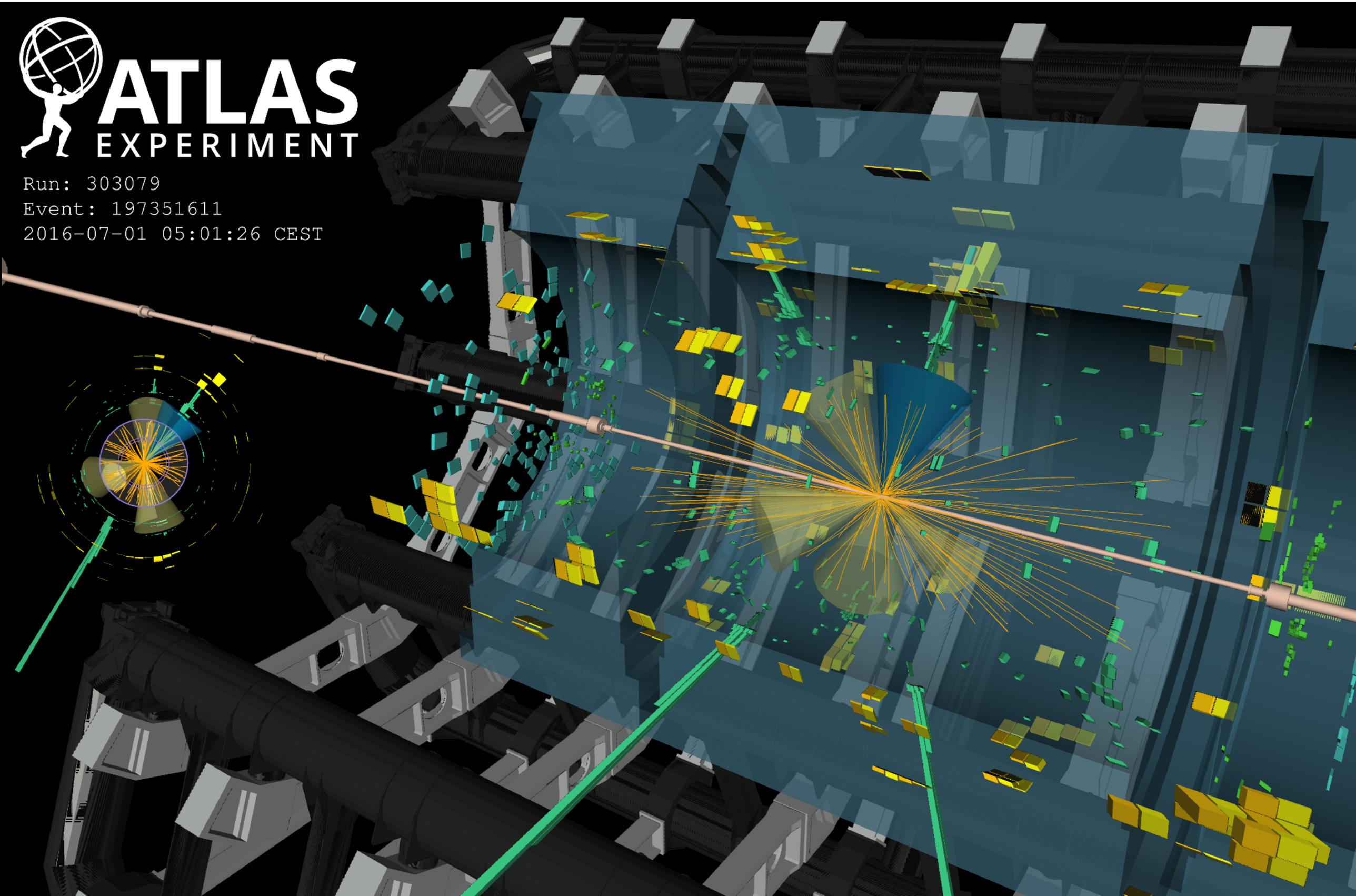
- Search for ttH production performed in ATLAS with **36.1 - 79.8 fb^{-1} of data at 13 TeV**

- Challenging analyses:
  - very **low cross section and high combinatorics** of final state particles
  - heavy use of **MVA techniques** to efficiently discriminate signal from large backgrounds
  - **large systematics uncertainties on modelling** of signal and irreducible backgrounds, ttbb and ttV

- **First ATLAS observation of ttH production at 6.3σ (expected 5.1σ)** →direct observation of Higgs to top Yukawa coupling
  - sensitivity limited by systematic uncertainties
  - some categories still statistically limited though
  - combination with Run-1 increases the (observed) sensitivity from 5.8σ to 6.3σ

- Measured cross section: $\sigma_{ttH} = 670 \pm 90(stat.)^{+110}_{-100}$ **(syst.) fb** (SM $\sigma_{ttH} = 507^{+35}_{-50}$ fb)
  - $\mu_{ttH} = 1.32^{+0.28}_{-0.26}$ compatible with SM.

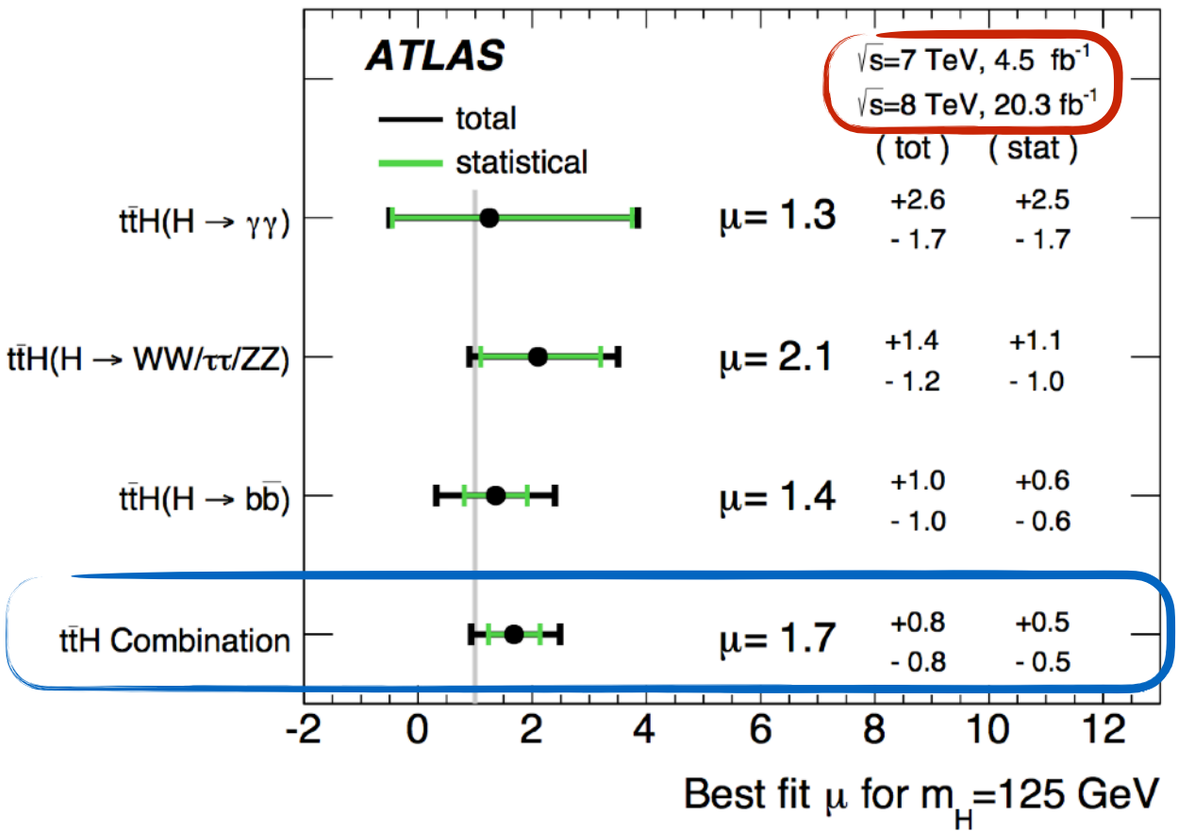
Bright future ahead!

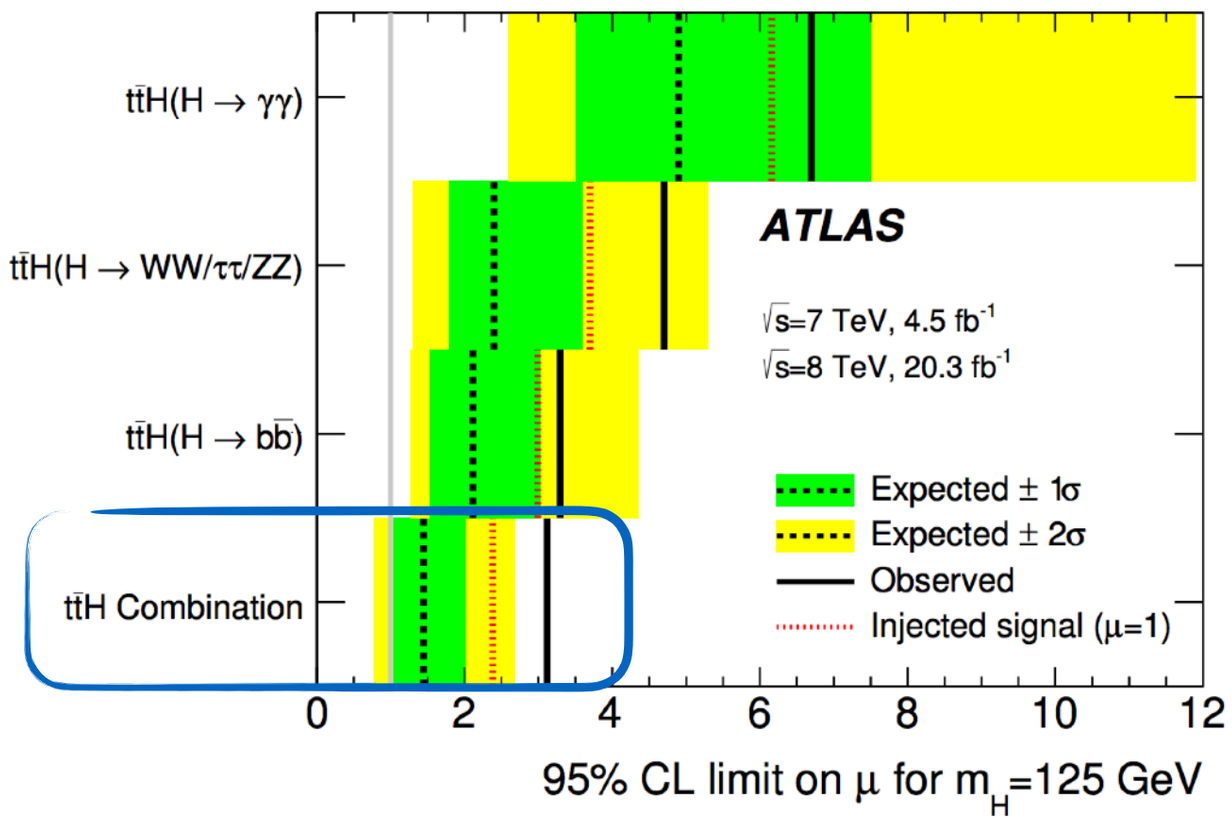


Supporting material

ttH - Run1 results

JHEP05(2016)160



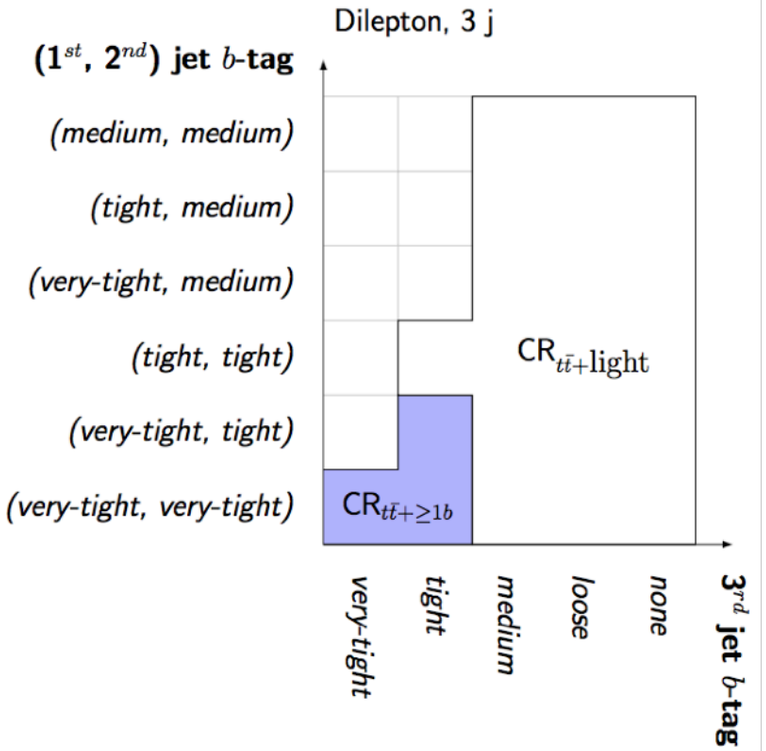


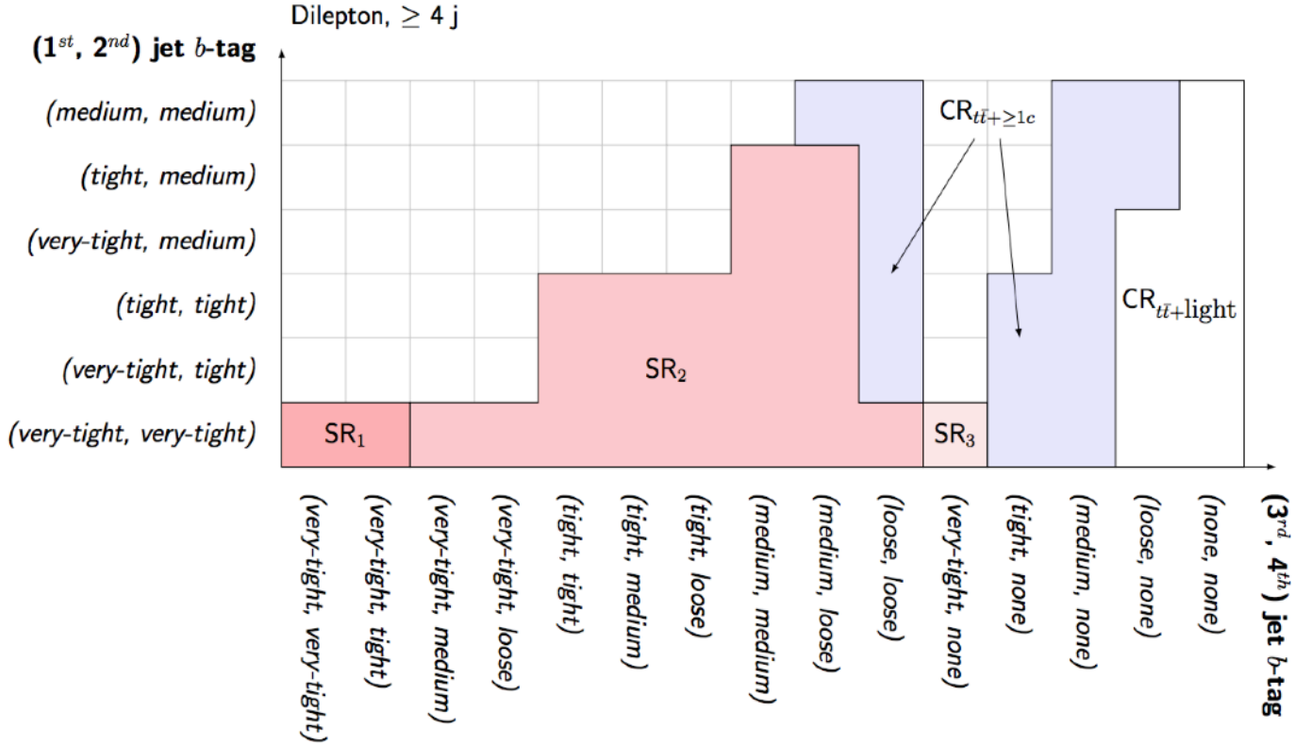
- Many channels sensitive to different final states, depending on the Higgs decay modes: **γγ**, **WW/ZZ/ττ** (multi-lepton) and **bb**;
- a signal strength **μ = 1.7 ± 0.8** has been measured;
- this corresponds to an **observed (expected) significance 2.3 σ (1.5 σ)**;
- **observed (expected) 95% C.L. limit on μ is 3.1 (1.4)**.

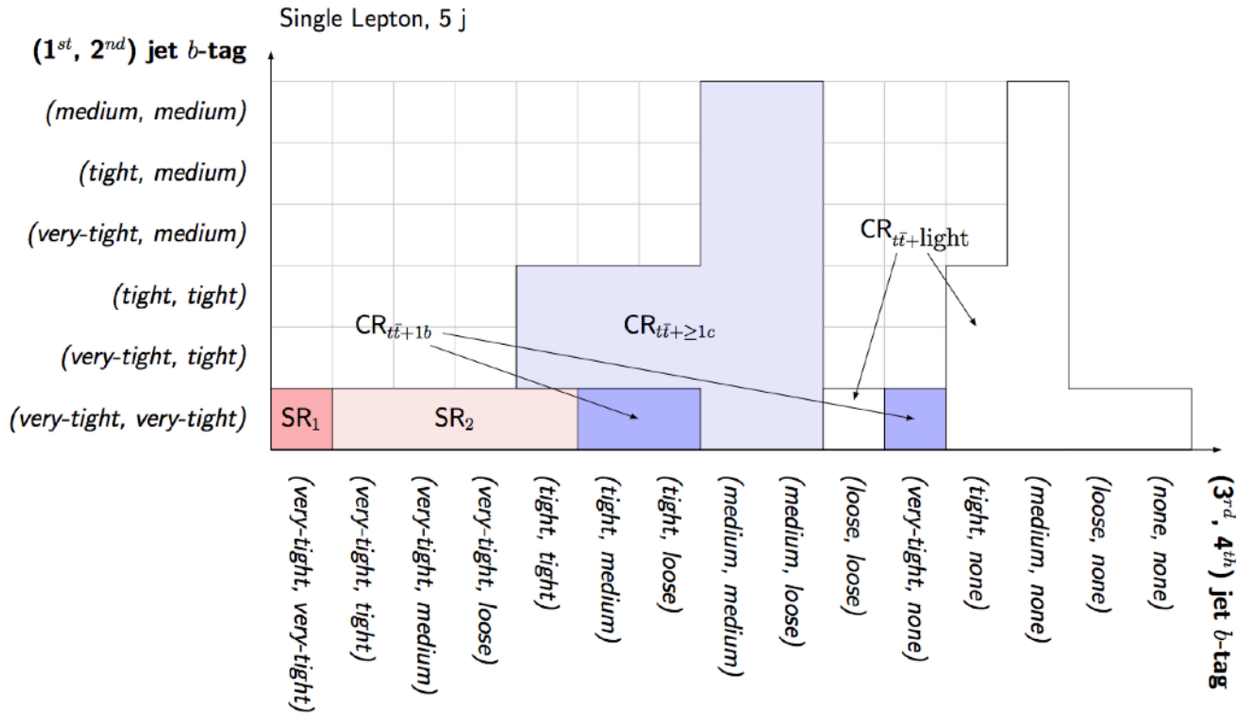
ttH(bb) - selection and classification

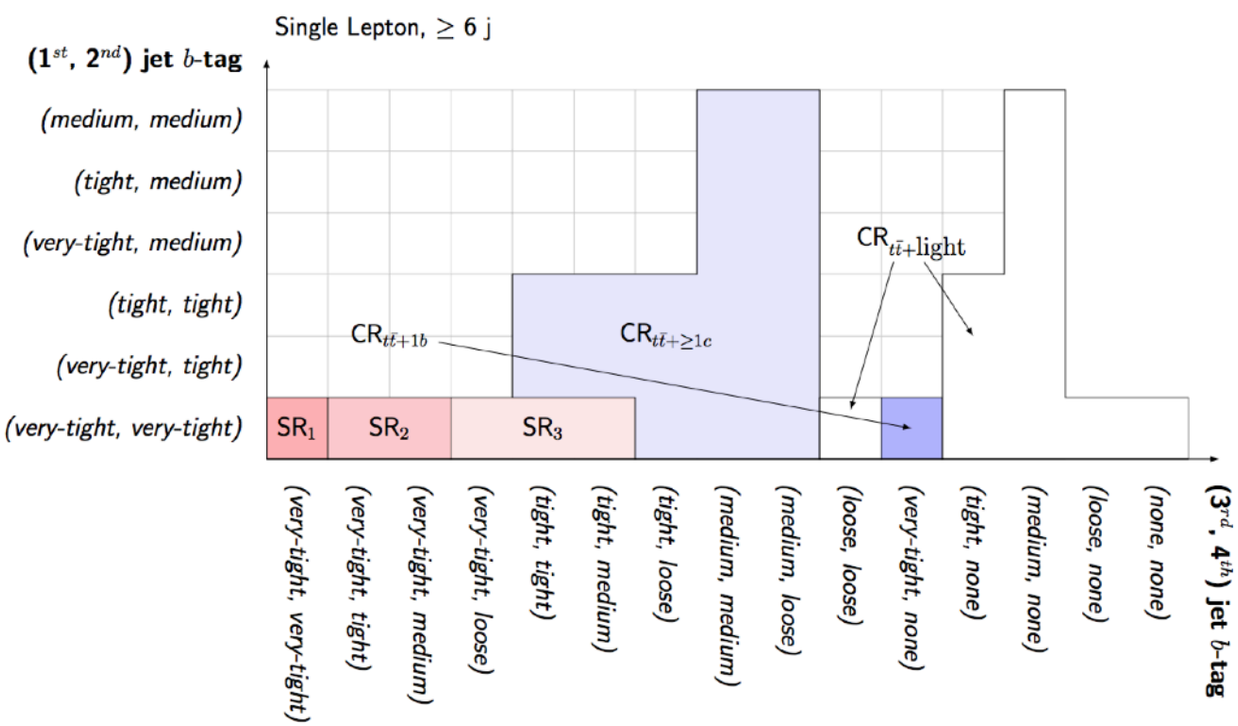
The jets are ordered according to their value of the b-tagging discriminant in descending order.

1 = no b-tagging criteria fulfilled
2 = loose = 85%
3 = medium = 77%
4 = tight = 70%
5 = very tight = 60%



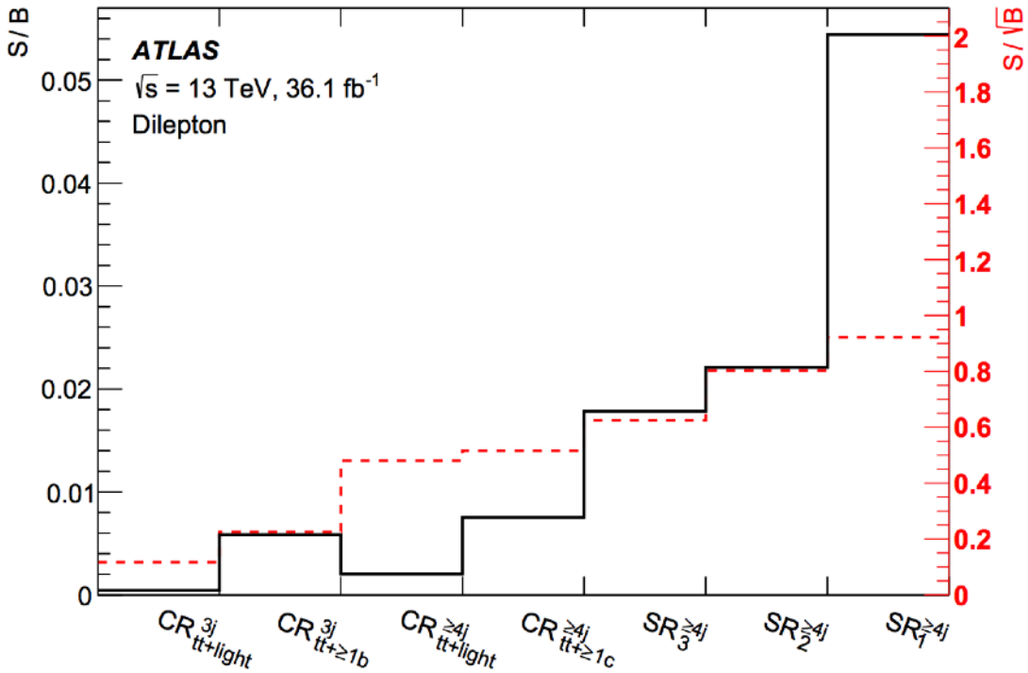


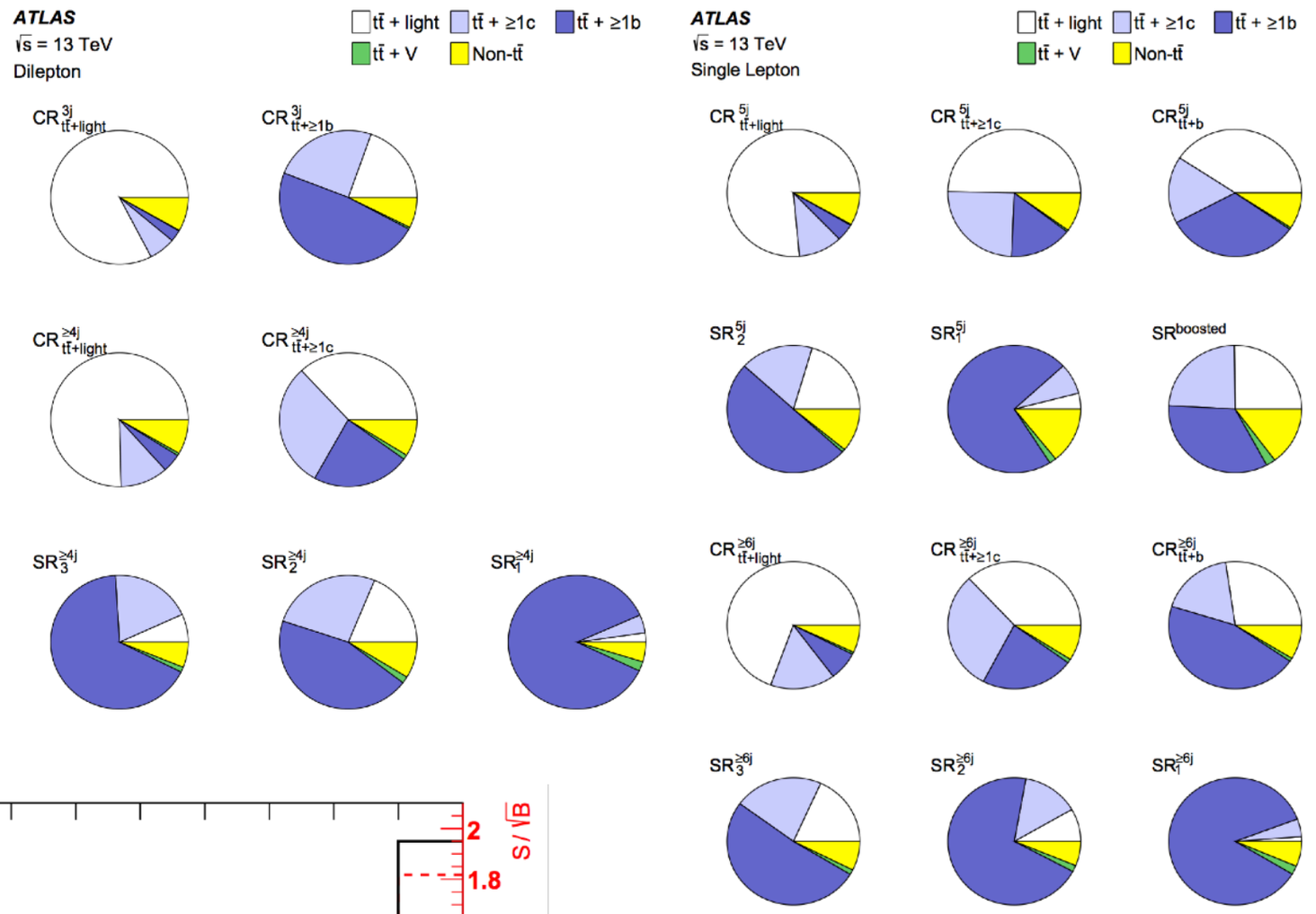


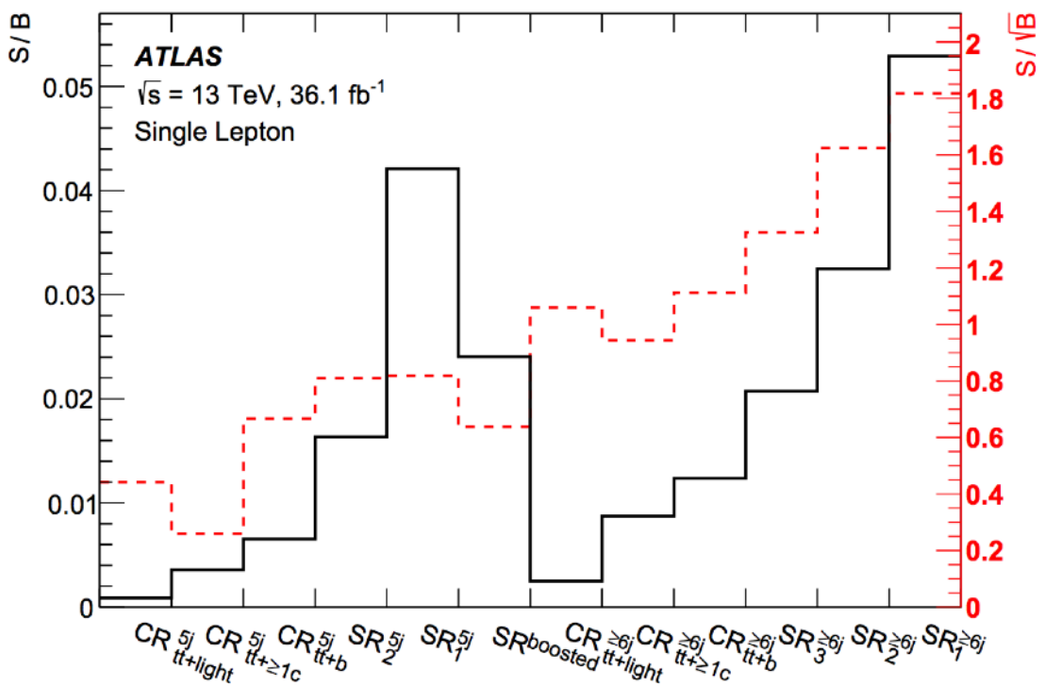


y-axis: values of the b-tagging discriminant for the first two jets, x-axis: values for the third jet or the third and fourth jets.

ttH(bb) - selection and classification





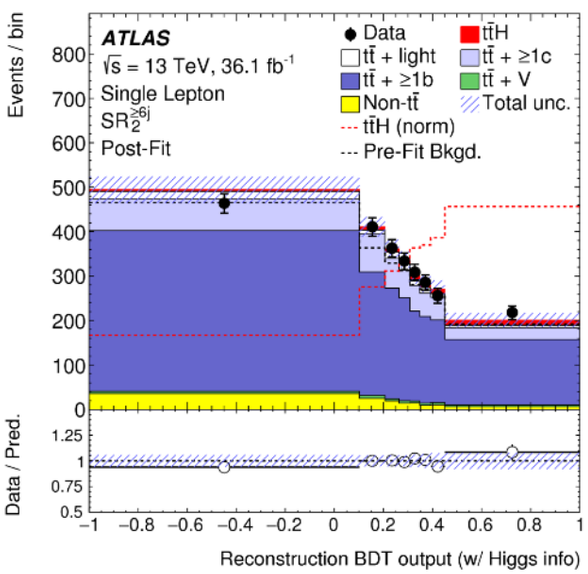


ttH(bb) - Selection and classification

- **Dilepton:** 2 OS leptons with p_T > 27, 15 GeV (m_{ll}~m_Z veto, τ veto), ≥ 3 jets and ≥ 2 medium b-tagged jets
- **Single lepton:** 1 lep with p_T > 27 GeV (veto events with 2τ)
  - Boosted: boosted Higgs (p_T > 200 GeV) or top (p_T > 250 GeV) candidate with large-R jet, and loose b-tagged jet. Higgs candidate: 2 loose b-tagged jet. Top candidate: 1 loose b-tagged jet and 1 jet.
  - Resolved: ≥ 5 jets and ≥ 2 b-tagged jets or ≥ 3 medium b-tagged jets.

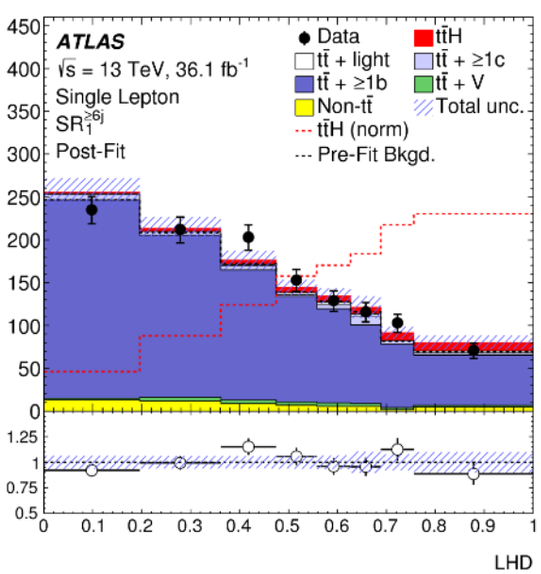
RecoBDT

identify jet assignments to reconstruct top and Higgs candidates



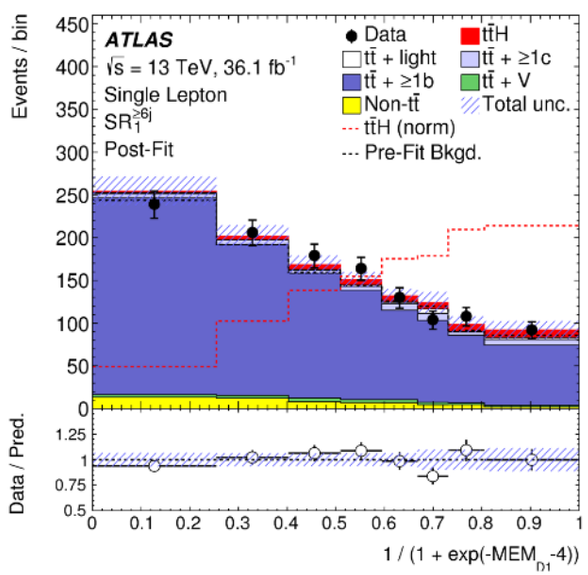
LHD

probability for ttH or tt+≥1b hypotheses with 1D discriminating variable



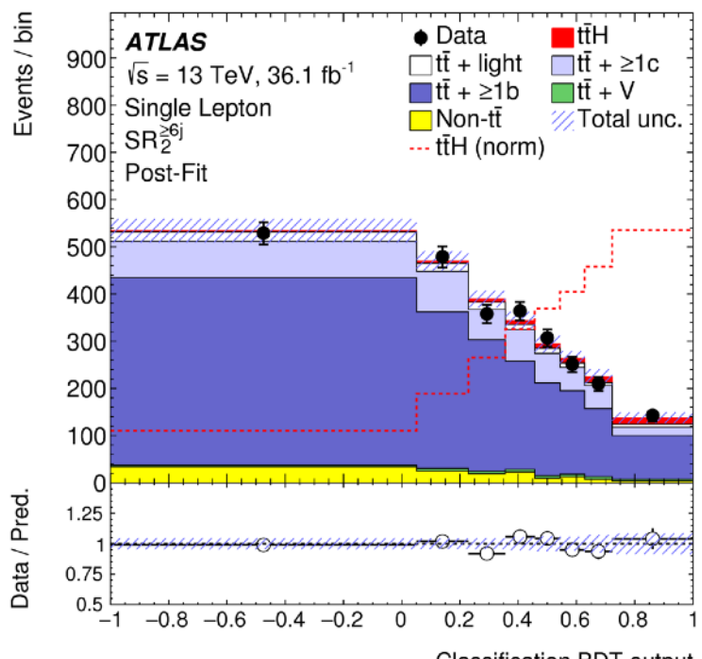
MEM

signal/background probability using ME calculation (only in most sensitive SR)



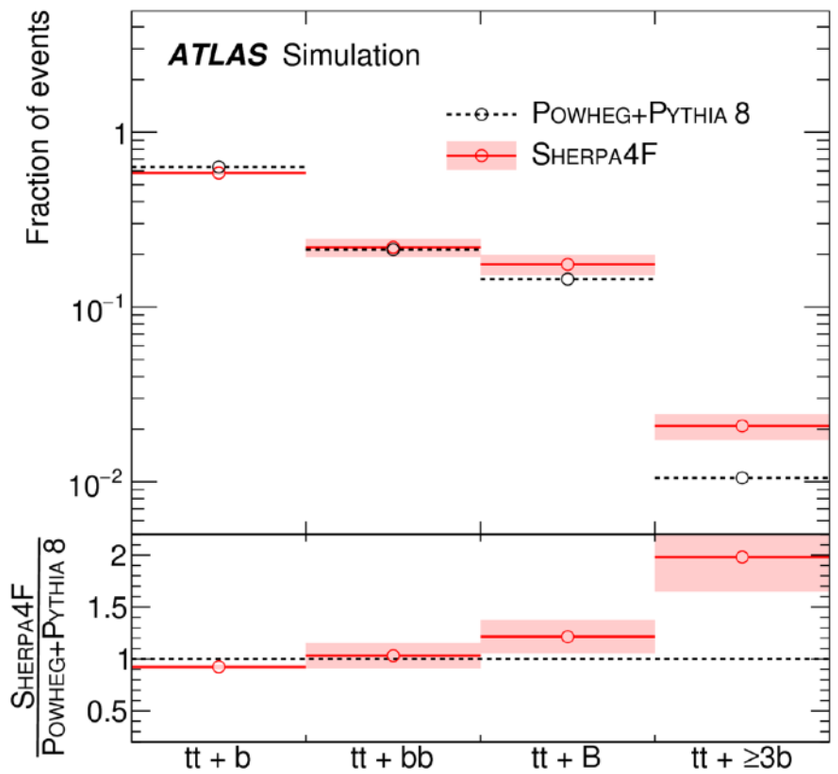
ClassificationBDT

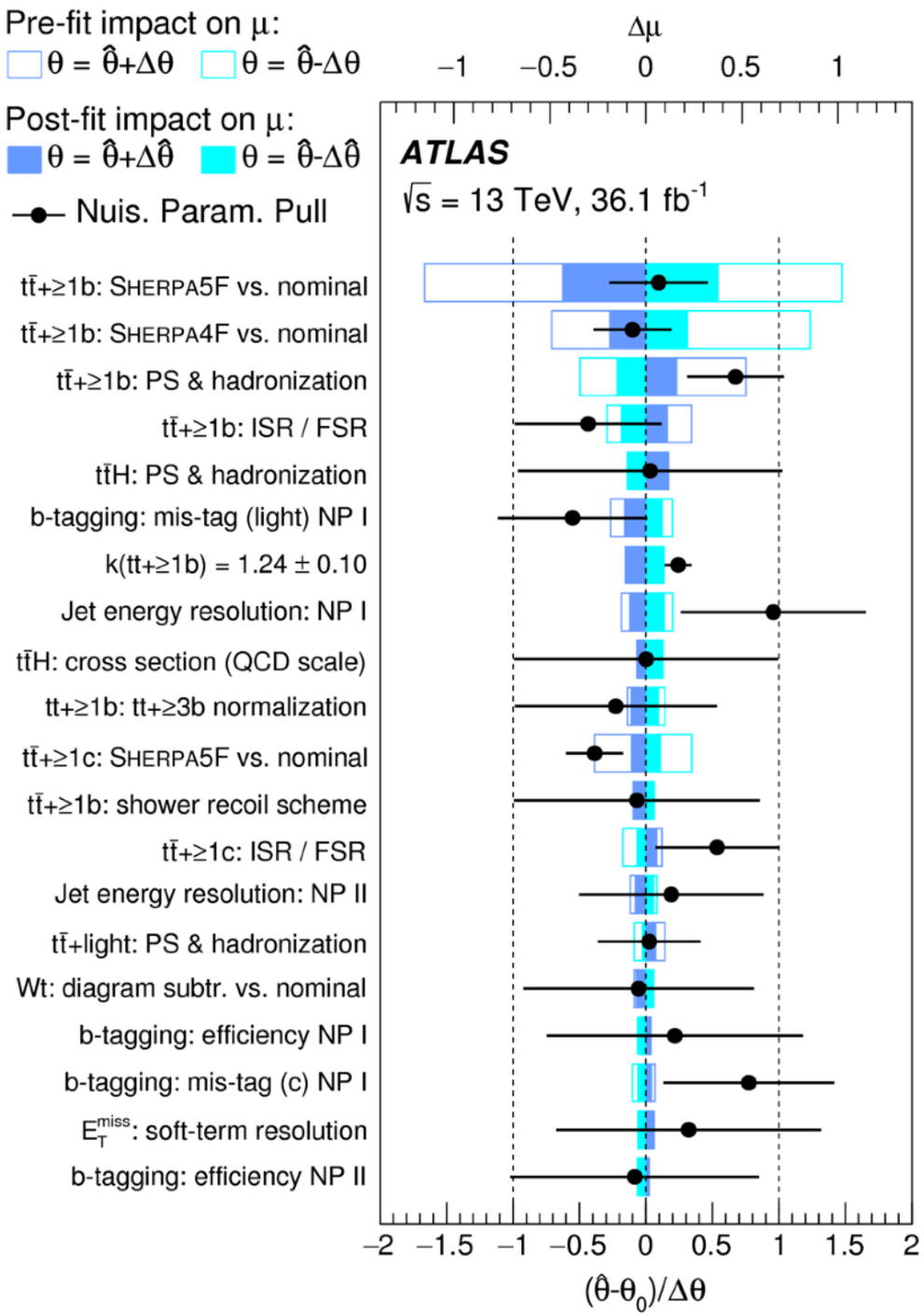
classify signal/background events based on other MVA outputs, kinematic properties of Higgs and top candidates, other event kinematic variables and discrete b-tagging information



ttH(bb): tt modelling and systematics

- Nominal tt sample simulated at NLO with Powheg+Pythia8 and normalised to NNLO+NNLL cross section
  - sample **split by number of b- and c-jets** at parton level
  - **tt+≥1b** contribution reweighted to ttbb predictions by Sherpa+OpenLoops (NLO description of kinematics of the two b-jet, b-quark masses included)
- Normalisations of tt+≥1b and tt+≥1c components are left **free-floating** in fit
- **Large systematics on tt+≥1b** from variations in simulation and from comparison with alternative simulations
  - some of these are constrained in fit to data





ttH(bb) - Selection and classification

Variable	Definition	$SR_1^{\geq 4j}$	$SR_2^{\geq 4j}$	$SR_3^{\geq 4j}$
General kinematic variables				
m_{bb}^{min}	Minimum invariant mass of a b-tagged jet pair	✓	✓	-
m_{bb}^{max}	Maximum invariant mass of a b-tagged jet pair	-	-	✓
$m_{bb}^{\text{min }\Delta R}$	Invariant mass of the b-tagged jet pair with minimum ΔR	✓	-	✓
$m_{\text{jj}}^{\text{max }p_{\text{T}}}$	Invariant mass of the jet pair with maximum p_{T}	✓	-	-
$m_{bb}^{\text{max }p_{\text{T}}}$	Invariant mass of the b-tagged jet pair with maximum p_{T}	✓	-	✓
$\Delta\eta_{bb}^{\text{avg}}$	Average $\Delta\eta$ for all b-tagged jet pairs	✓	✓	✓
$\Delta\eta_{\ell,\text{j}}^{\text{max}}$	Maximum $\Delta\eta$ between a jet and a lepton	-	✓	✓
$\Delta R_{bb}^{\text{max }p_{\text{T}}}$	ΔR between the b-tagged jet pair with maximum p_{T}	-	✓	✓
$N_{bb}^{\text{Higgs 30}}$	Number of b-tagged jet pairs with invariant mass within 30 GeV of the Higgs-boson mass	✓	✓	-
$n_{\text{jets}}^{p_{\text{T}}>40}$	Number of jets with $p_{\text{T}} > 40$ GeV	-	✓	✓
$\text{Aplanarity}_{b\text{-jet}}$	$1.5\lambda_2$, where λ_2 is the second eigenvalue of the momentum tensor [99] built with all b-tagged jets	-	✓	-
$H_{\text{T}}^{\text{all}}$	Scalar sum of p_{T} of all jets and leptons	-	-	✓
Variables from reconstruction BDT				
BDT output	Output of the reconstruction BDT	✓**	✓**	✓
m_{bb}^{Higgs}	Higgs candidate mass	✓	-	✓
$\Delta R_{H,t\bar{t}}$	ΔR between Higgs candidate and $t\bar{t}$ candidate system	✓*	-	-
$\Delta R_{H,\ell}^{\text{min}}$	Minimum ΔR between Higgs candidate and lepton	✓	✓	✓
$\Delta R_{H,b}^{\text{min}}$	Minimum ΔR between Higgs candidate and b-jet from top	✓	✓	-
$\Delta R_{H,b}^{\text{max}}$	Maximum ΔR between Higgs candidate and b-jet from top	-	✓	-
$\Delta R_{bb}^{\text{Higgs}}$	ΔR between the two jets matched to the Higgs candidate	-	✓	-
Variables from b-tagging				
$w_{b\text{-tag}}^{\text{Higgs}}$	Sum of b-tagging discriminants of jets from best Higgs candidate from the reconstruction BDT	-	✓	-

Classification BDT in dilepton channel

ttH(bb) - Selection and classification

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	–	✓
$N^{\text{Higgs 30}}_{bb}$	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 [99] 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	✓	✓
$m_{H,b_{\text{lep top}}}$	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	✓*	✓*
$\Delta R_{H,\text{lep top}}$	ΔR between Higgs candidate and leptonic top candidate	✓	–
$\Delta R_{H,b_{\text{had top}}}$	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	✓	✓

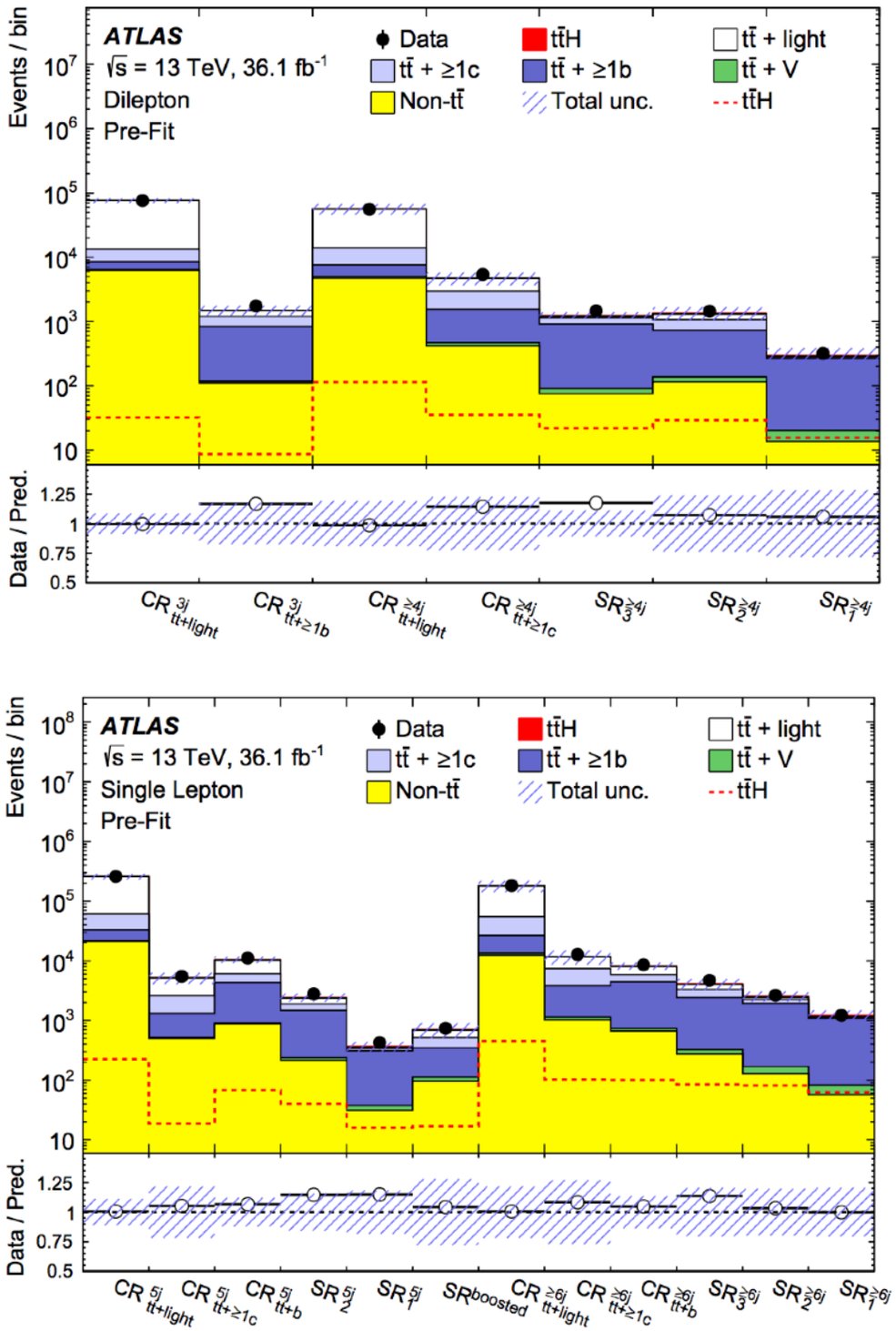
Classification BDT in single lepton channel

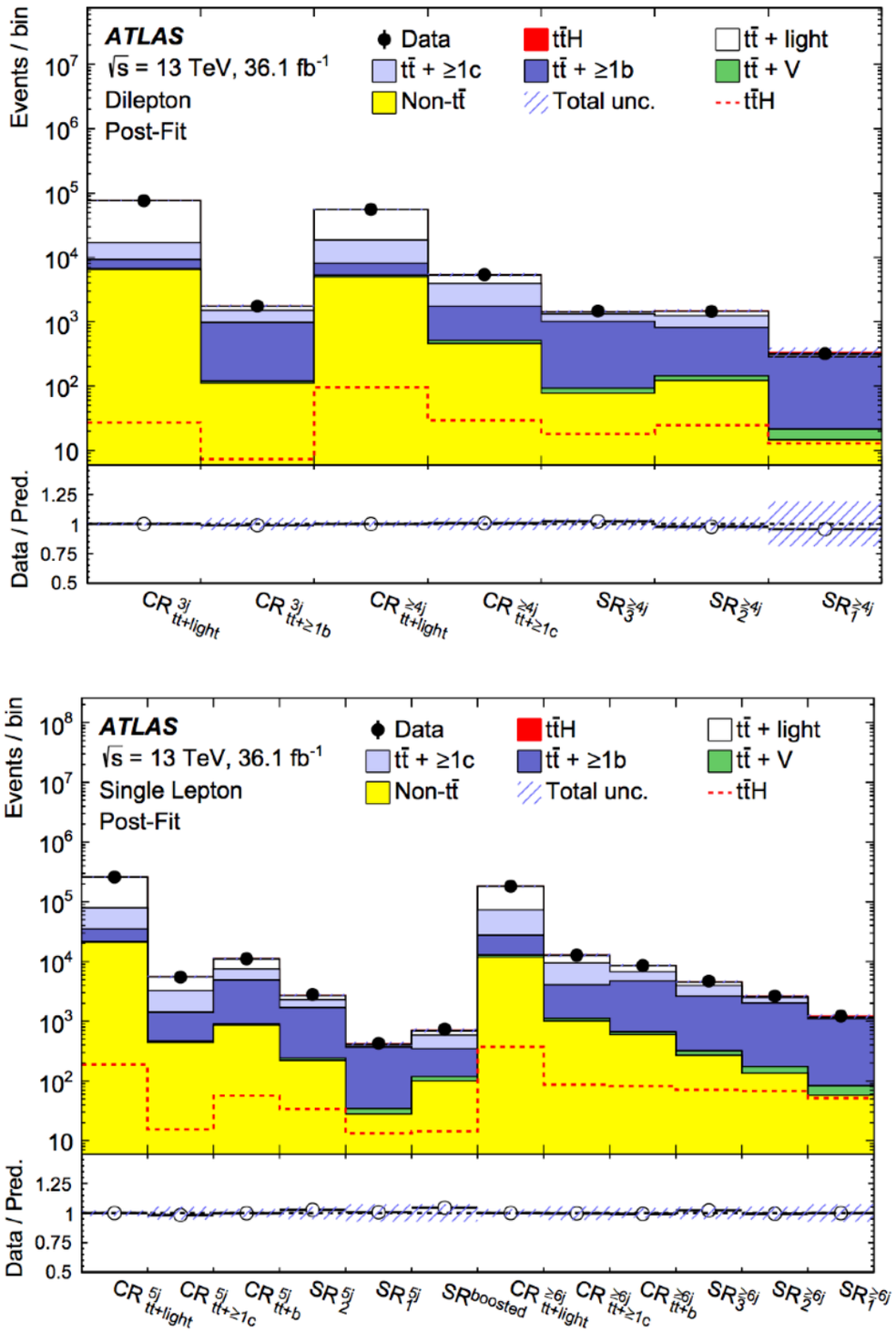
ttH(bb) - Selection and classification

Classification BDT in boosted channel

Variable	Definition
Variables from jet reclustering	
$\Delta R_{H,t}$	ΔR between the Higgs-boson and top-quark candidates
$\Delta R_{t,b^{\text{add}}}$	ΔR between the top-quark candidate and additional b-jet
$\Delta R_{H,b^{\text{add}}}$	ΔR between the Higgs-boson candidate and additional b-jet
$\Delta R_{H,\ell}$	ΔR between the Higgs-boson candidate and lepton
$m_{\text{Higgs candidate}}$	Higgs-boson candidate mass
$\sqrt{d_{12}}$	Top-quark candidate first splitting scale [100]
Variables from b-tagging	
$w_{b\text{-tag}}$	Sum of b-tagging discriminants of all b-jets
$w^{\text{add}}_{b\text{-tag}}/w_{b\text{-tag}}$	Ratio of sum of b-tagging discriminants of additional b-jets to all b-jets

ttH(bb) - Pre and post-fit distributions





ttH(bb) - Systematic uncertainties: tt background

- Nominal samples split in 3 categories depending on the flavour of the additional jets.
- tt+≥1b further divided into 6 sub-components
  - ttb, ttbb, ttB and tt3b scaled to SherpaOL 4FS;
  - ttb MPI/FSR components are not reweighted (absent in SherpaOL 4FS);
  - scaling only changes normalisation.

Systematic source	Description	$t\bar{t}$ categories
$t\bar{t}$ cross-section	Up or down by 6%	All, correlated
$t\bar{t}+\geq 1c$ normalisation	Free parameter $k(t\bar{t}+\geq 1c)$	$t\bar{t}+\geq 1c$
$t\bar{t}+\geq 1b$ normalisation	Free parameter $k(t\bar{t}+\geq 1b)$	$t\bar{t}+\geq 1b$
NLO generator	Sherpa vs. Powheg-Box+Pythia 8	All, uncorrelated
p.s. & hadronisation	Powheg-Box+Herwig 7 vs. Powheg-Box+Pythia 8	All, uncorrelated
ISR / FSR	Variations of μ_R, μ_F, h_{damp} and A14 Var3c parameters	All, uncorrelated
$t\bar{t}+\geq 1c$ 3F vs. 5F	MG5_aMC@NLO+Herwig++ ME prediction vs. incl.	$t\bar{t}+\geq 1c$
$t\bar{t}+\geq 1b$ 4F vs. 5F	SherpaOL vs. Powheg-Box+Pythia 8	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ renorm. scale	Up or down by a factor of two	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ resumm. scale	Vary μ_Q from $H_T/2$ to μ_{CMMPS}	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ global scales	Set μ_Q, μ_R, and μ_F to μ_{CMMPS}	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ shower recoil	Alternative model scheme	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ PDF	MSTW or NNPDF vs. CT10	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ FSR	Prediction from ISR / FSR variation samples vs. nominal	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ MPI	Up or down by 50%	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 3b$ normalisation	Up or down by 50%	$t\bar{t}+\geq 3b$

- Normalisation only
- Shape only
- tt+≥1b fractions systematics

ttH ML - Selection and classification

Table 2: Loose (L), loose and isolated (L†), loose, isolated and passing the non-prompt BDT (L*), tight (T) and very tight (T*) light-lepton definitions. Selections for the tighter leptons are applied in addition to the looser ones. For the muons, the L*, T and T* lepton definitions are identical.

	e					μ		
	L	L†	L*	T	T*	L	L†	L*/T/T*
Isolation	No	Yes				No	Yes	
Non-prompt lepton BDT	No	Yes				No	Yes	
Identification	Loose			Tight		Loose		
Charge misassignment veto BDT	No				Yes	No		
Transverse impact parameter significance, $\lvert d_0\rvert/\sigma_{d_0}$	< 5					< 3		
Longitudinal impact parameter, $\lvert z_0 \sin\theta\rvert$	< 0.5 mm							

Table 4: Summary of the basic characteristics of the seven analysis channels. The lepton selection follows the definition in Table 2 and is labeled as loose (L), loose and isolated (L†), loose, isolated and passing the non-prompt BDT (L*), tight (T) and very tight (T*), respectively. The τ_{had} selection is labeled as medium (M) and tight (T).

	2ℓSS	3ℓ	4ℓ	1ℓ+$2\tau_{\text{had}}$	2ℓSS+$1\tau_{\text{had}}$	2ℓOS+$1\tau_{\text{had}}$	3ℓ+$1\tau_{\text{had}}$
Light lepton	2T*	1L*, 2T*	2L, 2T	1T	2T*	2L†	1L†, 2T
τ_{had}	0M	0M	–	1T, 1M	1M	1M	1M
N_{jets}, $N_{b\text{-jets}}$	$\geq 4, = 1, 2$	$\geq 2, \geq 1$	$\geq 2, \geq 1$	$\geq 3, \geq 1$	$\geq 4, \geq 1$	$\geq 3, \geq 1$	$\geq 2, \geq 1$

ttH ML - Selection and classification

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)
$1\ell{+}2\tau_{\text{had}}$	One tight light lepton with $p_{\text{T}} > 27$ GeV Two medium τ_{had} candidates of opposite charge, at least one being tight $N_{\text{jets}} \geq 3$
$2\ell\text{SS}{+}1\tau_{\text{had}}$	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
$2\ell\text{OS}{+}1\tau_{\text{had}}$	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 $N_{\text{jets}} \geq 3$
$3\ell{+}1\tau_{\text{had}}$	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

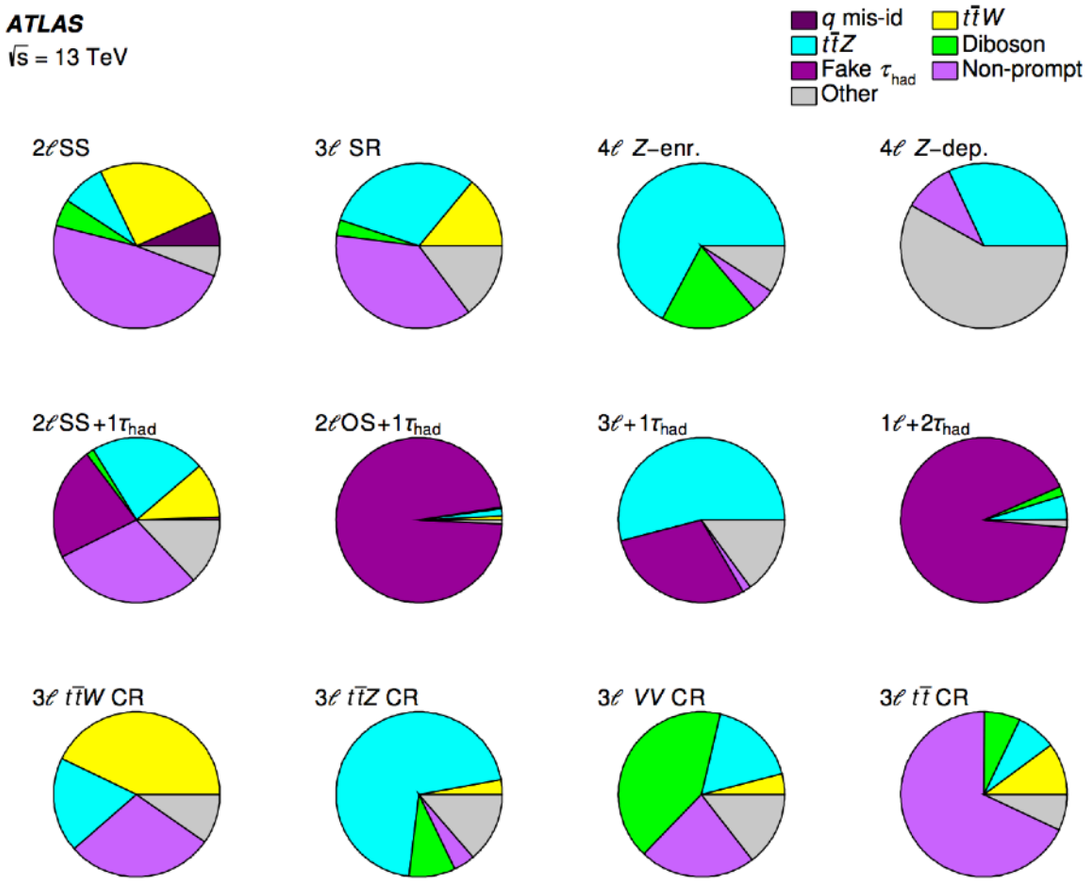
Selection criteria applied in the different channels.
Same-flavour, opposite-charge lepton pairs are referred to as SFOC pairs.
The common selection criteria for all channels are listed in the first line under the title "Common".

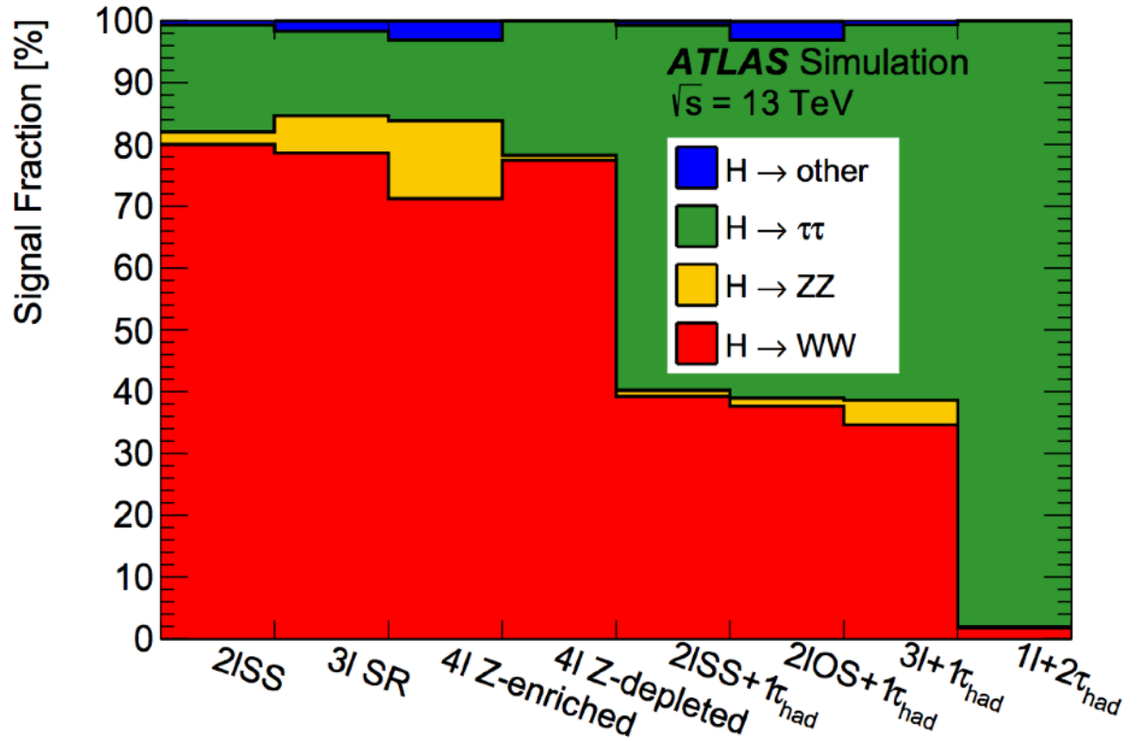
ttH ML - Selection and classification

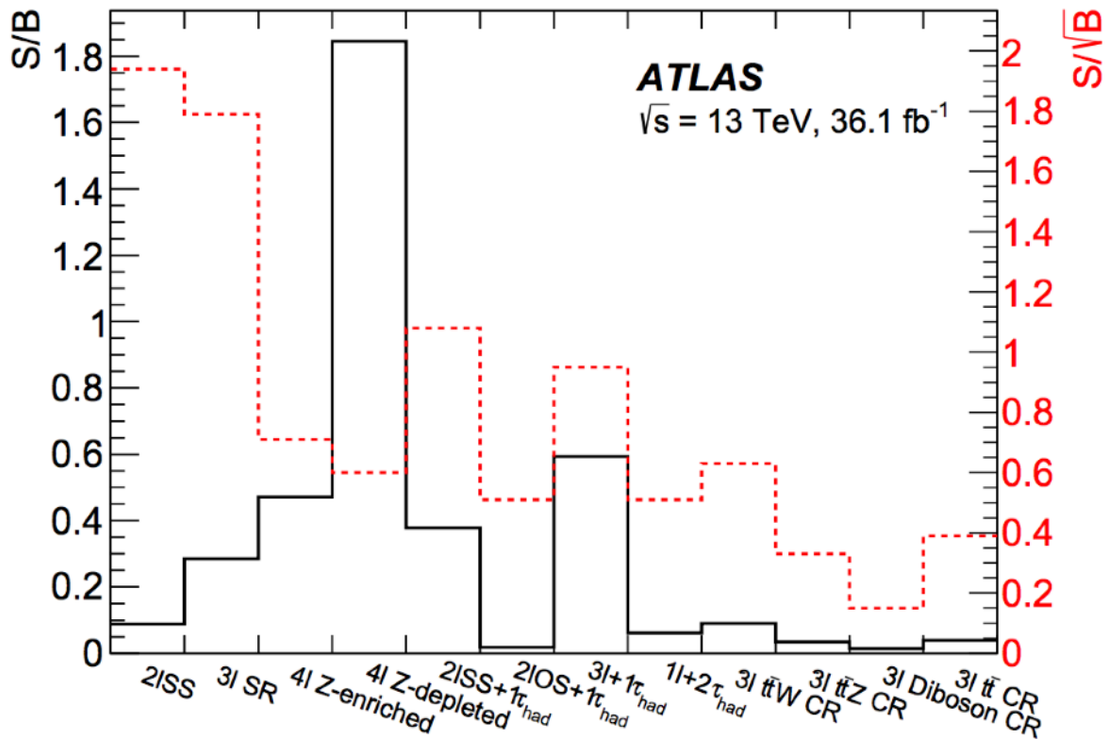
	Variable	2ℓSS	3ℓ	4ℓ	1ℓ+2τ_{had}	2ℓSS+1τ_{had}	2ℓOS+1τ_{had}
Lepton properties	Leading lepton p_T		×				
	Second leading lepton p_T	×	×			×	
	Third lepton p_T		×				
	Dilepton invariant mass (all combinations)	×	×*				×
	Three-lepton invariant mass		×				
	Four-lepton invariant mass			×			
	Best Z-candidate dilepton invariant mass			×			
	Other Z-candidate dilepton invariant mass			×			
	Scalar sum of all leptons p_T			×			×
	Second leading lepton track isolation					×	
	Maximum $\|\eta\|$ (lepton 0, lepton 1)	×				×*	
	Lepton flavor	×*	×*				
	Lepton charge		×				
Jet properties	Number of jets	×*	×*		×	×	×
	Number of b-tagged jets	×*	×*		×	×	×
	Leading jet p_T						×
	Second leading jet p_T		×			×*	
	Leading b-tagged jet p_T		×				
	Scalar sum of all jets p_T		×		×	×	×
	Scalar sum of all b-tagged jets p_T						×
	Has leading jet highest b-tagging weight?		×				
	b-tagging weight of leading jet		×				
	b-tagging weight of second leading jet		×			×	
	b-tagging weight of third leading jet					×	
	Pseudorapidity of fourth leading jet					×	
τ_{had}	Leading τ_{had} p_T				×		×
	Second leading τ_{had} p_T				×		
	Di-τ_{had} invariant mass				×		
	Invariant mass τ_{had}–furthest lepton					×	
Angular distances	ΔR(lepton 0, lepton 1)		×				
	ΔR(lepton 0, lepton 2)		×				
	ΔR(lepton 0, closest jet)	×	×				
	ΔR(lepton 0, leading jet)		×			×	
	ΔR(lepton 0, closest b-jet)		×				
	ΔR(lepton 1, closest jet)	×	×				
	ΔR(lepton 2, closest jet)		×				
	Smallest ΔR(lepton, jet)		×				×
	Smallest ΔR(lepton, b-tagged jet)						×
	Smallest ΔR(non-tagged jet, b-tagged jet)						×
	ΔR(lepton 0, τ_{had})						×
	ΔR(lepton 1, τ_{had})						×
	Minimum ΔR between all jets				×		
	ΔR between two leading jets					×	
$\vec{p}_T^{miss}$	Missing transverse momentum E_T^{miss}	×		×			
	Azimuthal separation $\Delta\phi$(leading jet, $\vec{p}_T^{miss}$)		×				
	Transverse mass leptons (H/Z decay) - $\vec{p}_T^{miss}$			×			
	Pseudo-Matrix-Element			×			

Classification BDT input variables

ttH ML - Selection and classification





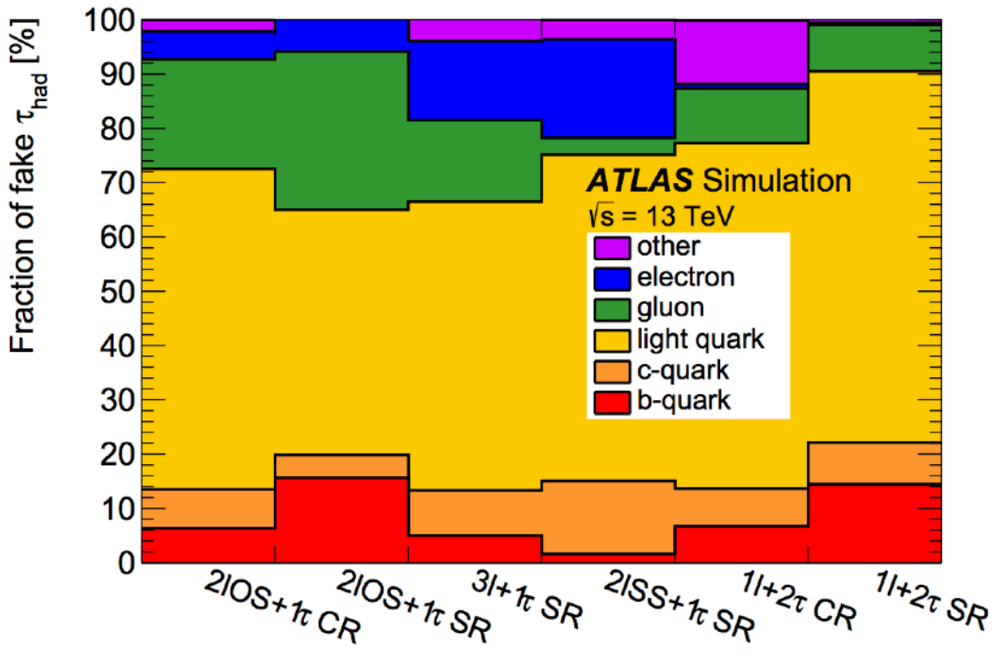


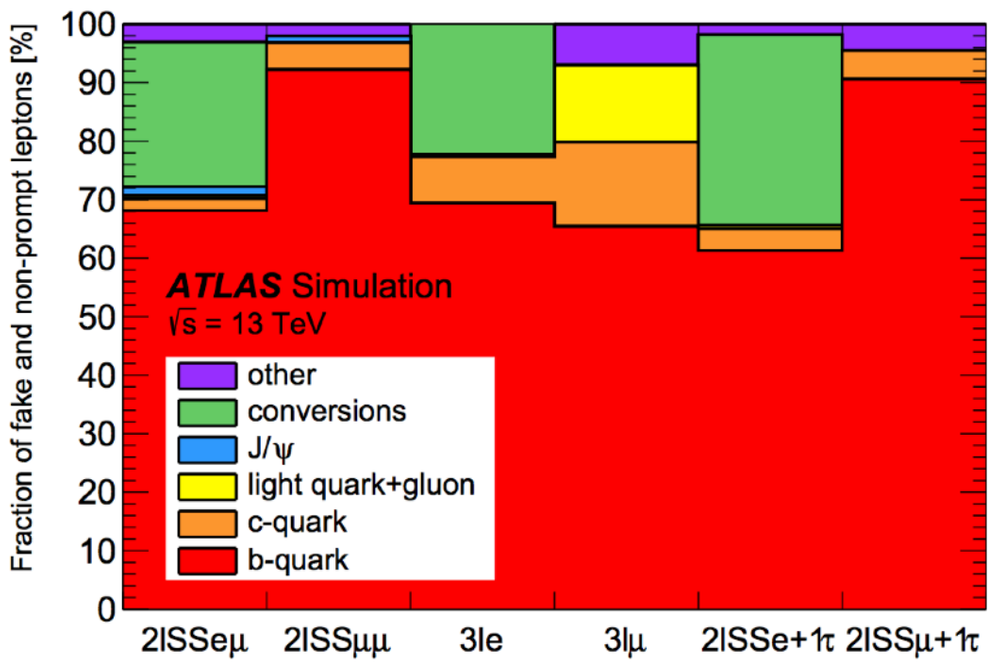
ttH ML - Selection and classification

Channel	Region	Selection criteria
2ℓSS		$2 \leq N_{\text{jets}} \leq 3$ and $N_{b\text{-jets}} \geq 1$
(3ℓ)		One very tight, one loose light lepton with $p_{\text{T}} > 20$ (15) GeV
		Zero τ_{had} candidates
	ϵ_{real}	Opposite charge; opposite flavor
	ϵ_{fake}	Same charge; opposite flavor or $\mu\mu$
4ℓ		$1 \leq N_{\text{jets}} \leq 2$
		Three loose light leptons; sum of light lepton charges ± 1
		Subleading same-charge lepton must be tight
		Veto on 3ℓ selection
	Either	One SFOC pair with $\lvert m(\ell^+\ell^-) - 91.2\text{ GeV}\rvert < 10$ GeV
		$E_{\text{T}}^{\text{miss}} < 50$ GeV, $m_{\text{T}} < 50$ GeV
	or	No SFOC pair
		Subleading jet $p_{\text{T}} > 30$ GeV
2ℓSS+$1\tau_{\text{had}}$		$2 \leq N_{\text{jets}} \leq 3$ and $N_{b\text{-jets}} \geq 1$
		One very tight, one loose light lepton with $p_{\text{T}} > 15$ GeV
		A SFSC pair
		$\lvert m(ee) - 91.2\text{ GeV}\rvert > 10$ GeV
		Zero or one medium τ_{had} candidate, opposite in charge to the light leptons
1ℓ+$2\tau_{\text{had}}$		$N_{\text{jets}} \geq 3$ and $N_{b\text{-jets}} \geq 1$
		One tight light lepton, with $p_{\text{T}} > 27$ GeV
		Two τ_{had} candidates of same charge
		At least one τ_{had} candidate has to satisfy tight identification criteria
2ℓOS+$1\tau_{\text{had}}$		Two loose and isolated light leptons, with $p_{\text{T}} > 25, 15$ GeV
		One loose τ_{had} candidate
		$\lvert m(\ell^+\ell^-) - 91.2\text{ GeV}\rvert > 10$ GeV and $m(\ell^+\ell^-) > 12$ GeV
		$N_{\text{jets}} \geq 3$ and $N_{b\text{-jets}} = 0$

	2ℓSS	3ℓ	4ℓ	1ℓ+$2\tau_{\text{had}}$	2ℓSS+$1\tau_{\text{had}}$	2ℓOS+$1\tau_{\text{had}}$	3ℓ+$1\tau_{\text{had}}$
Non-prompt lepton strategy	DD (MM)	DD (MM)	semi-DD (SF)	MC	DD (FF)	MC	MC
Fake τ_{had} strategy	–	–	–	DD (SS data)	semi-DD (SF)	DD (FF)	semi-DD (SF)
Control Region Selection							
Light lepton	1T*, 1L	3L		1T	1T*, 1L	2L[†]	–
τ_{had}	0M			1T, 1M	$\leq$ 1M	1L	–
N_{jets}	$2 \leq N_{\text{jets}} \leq 3$	$1 \leq N_{\text{jets}} \leq 2$		≥ 3	$2 \leq N_{\text{jets}} \leq 3$	≥ 3	–
$N_{b\text{-jets}}$	≥ 1					$= 0$	–

- Summary of the non-prompt lepton and fake τ_{had} background estimate strategies of the seven analysis channels. DD means data-driven background estimates and the techniques used are the matrix method (MM) and the fake-factor method (FF). The scale factor method (SF), which scales the estimate from simulation by a correction factor measured in data, is partially data-driven.

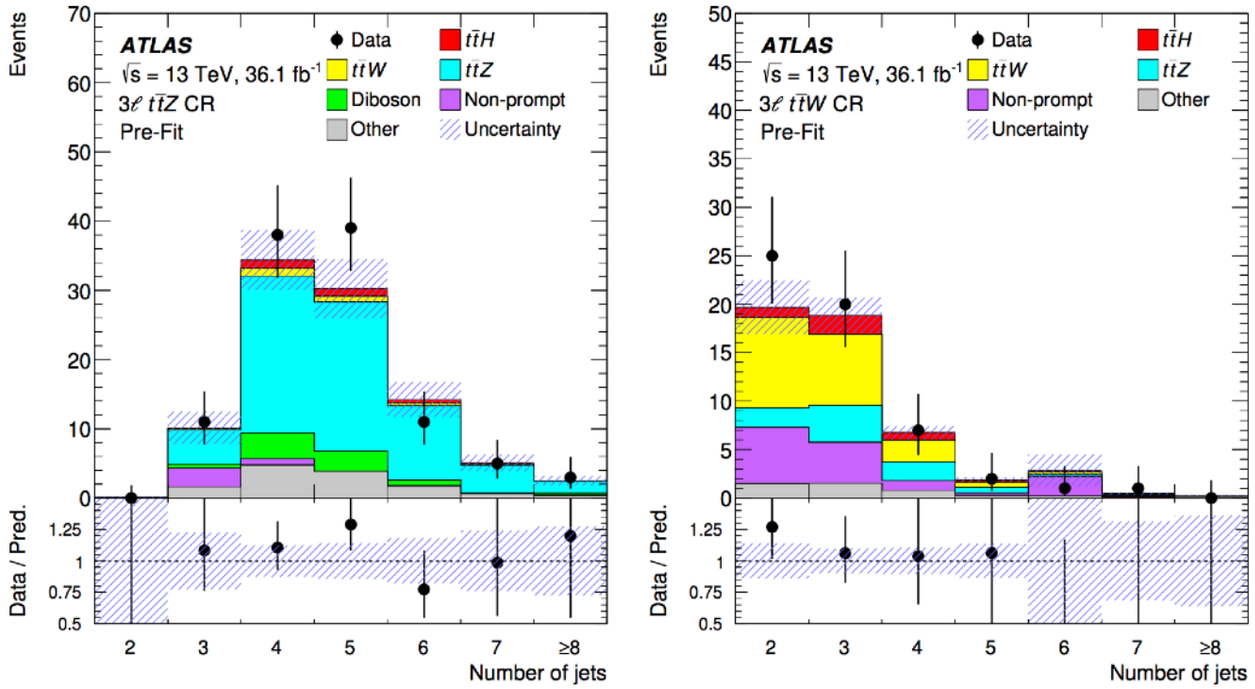




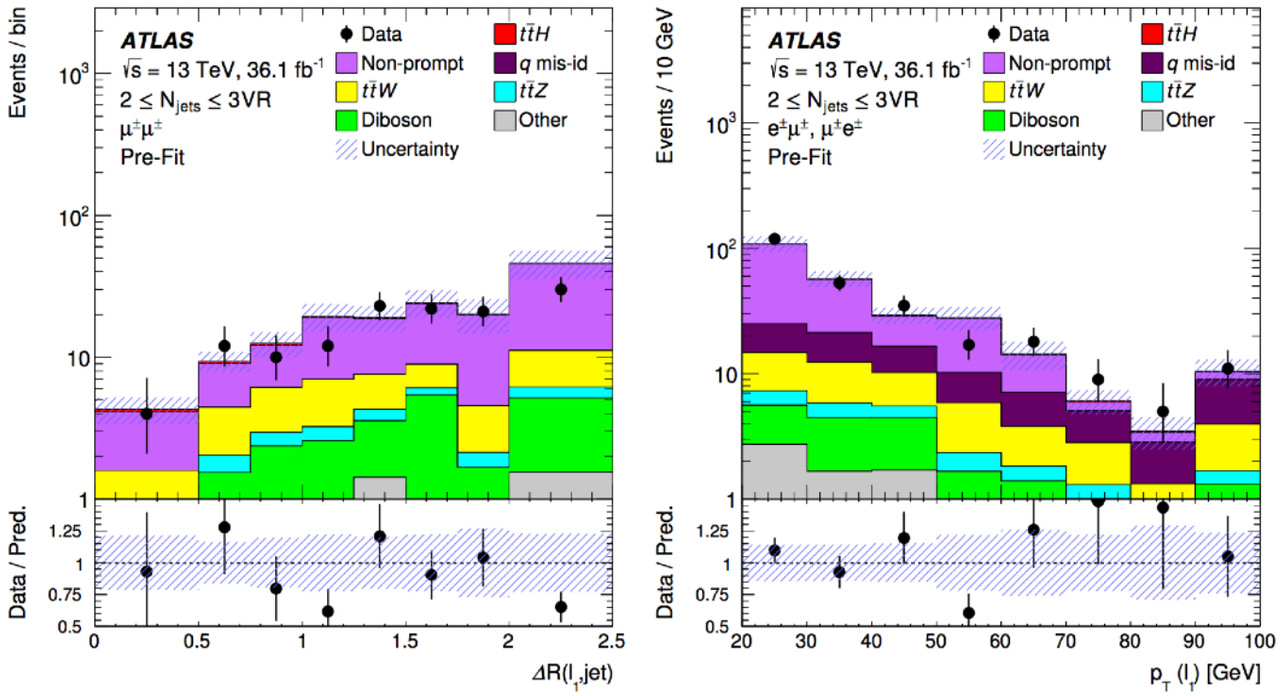
ttH ML: background estimation

- **Irreducible background:** ttW, ttZ with prompt leptons
  - rare SM processes only recently observed (JHEP11(2015)172)
  - estimated from NLO simulation (±12%(QCD scale)±4%(α_s))
- **Reducible background:** mostly tt(+γ) events with mis-reconstructed leptons, difficult to simulate
  - different "data-driven" estimations used based on lepton origin and when available data statistics allows
  - non-prompt leptons from HF decay and conversions (most critical in 2lSS, 3l)
    - data in "sideband" regions with loose leptons to model non-prompt events in SR with tight leptons
    - total uncertainties of 20-30%. Estimated non-prompt background is 1.5-2 times what predicted by simulation
    - large systematic uncertainty on the different loose-to-tight rates for electrons from HF decay and conversions
- **Charge mis-id background** in 2lSS, 2lSS+1τ_{had} estimated with OS events scaled by charge mis-id rate measured in Z→ee data.

ttV control regions with 3l events

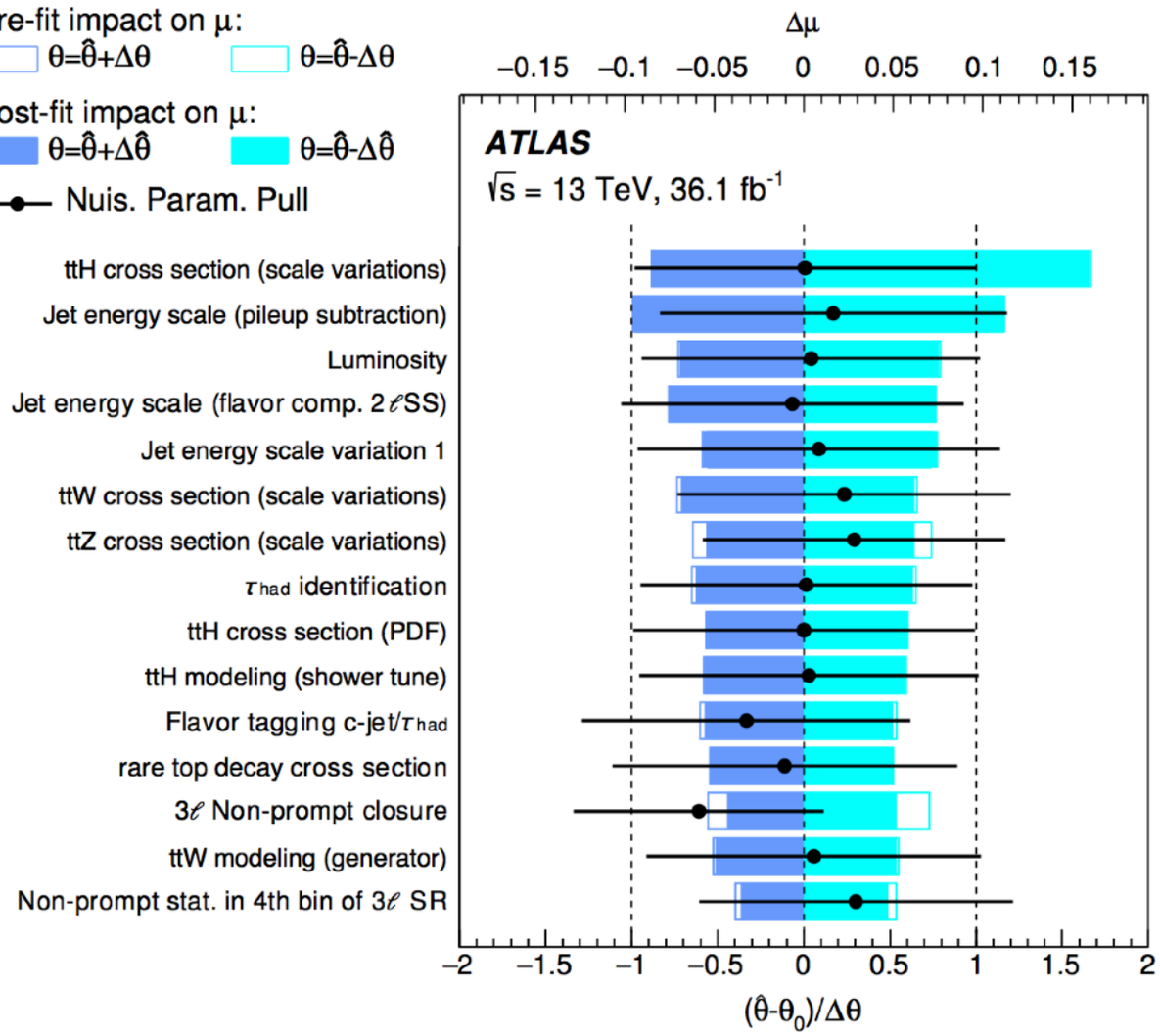


Non-prompt validation regions with 2lSS events

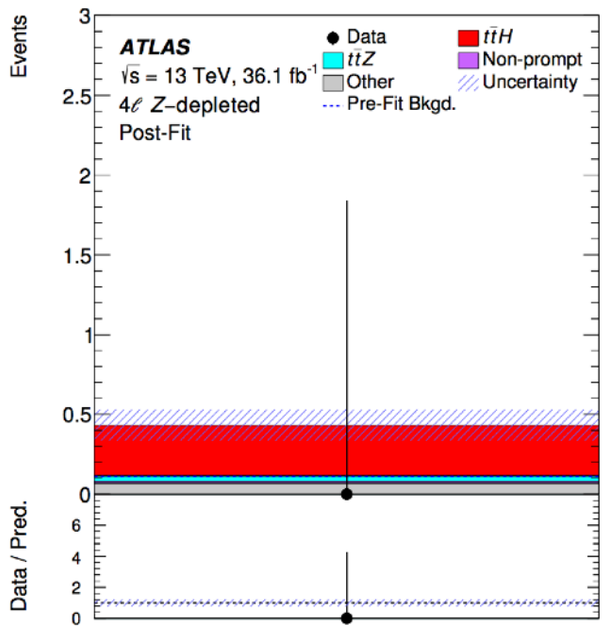


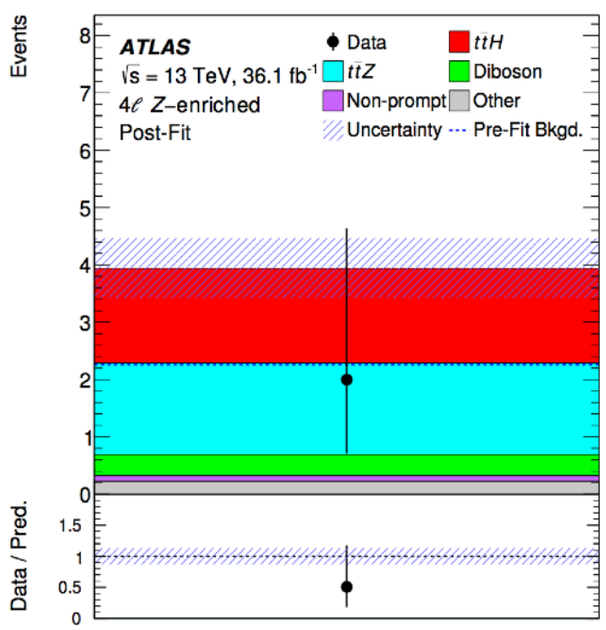
ttH ML - Systematic uncertainties

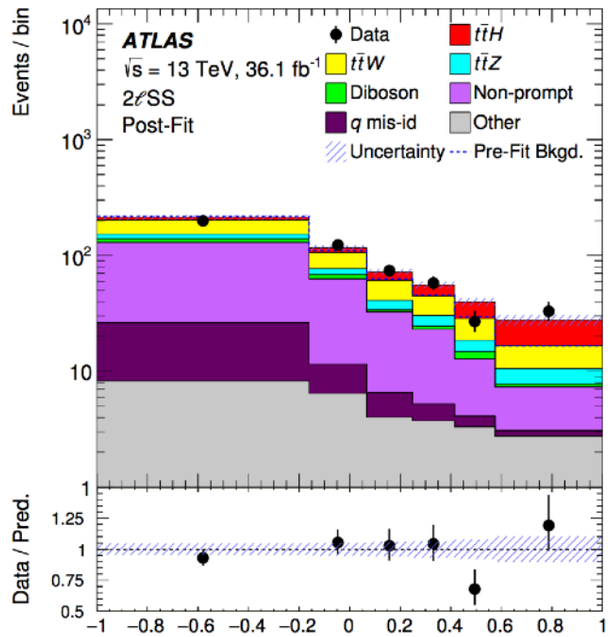
Systematic uncertainty	Type	Components
Luminosity	N	1
Pileup reweighting	SN	1
Physics Objects		
Electron	SN	6
Muon	SN	15
τ_{had}	SN	10
Jet energy scale and resolution	SN	28
Jet vertex fraction	SN	1
Jet flavor tagging	SN	126
E_T^{miss}	SN	3
Total (Experimental)	–	191
Data-driven non-prompt/fake leptons 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
***ttH* 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
***ttW* modeling**		
Cross section	N	2
Renormalization and factorization scales	S	3
Matrix-element MC event generator	SN	1
Shower tune	SN	1
***ttZ* 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

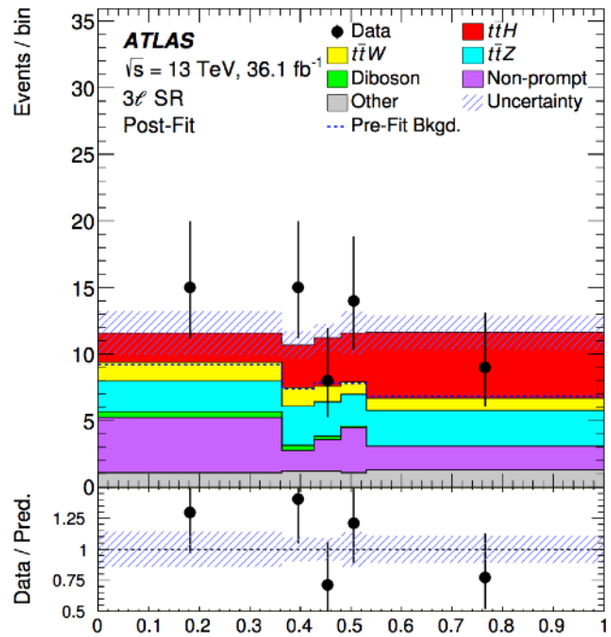


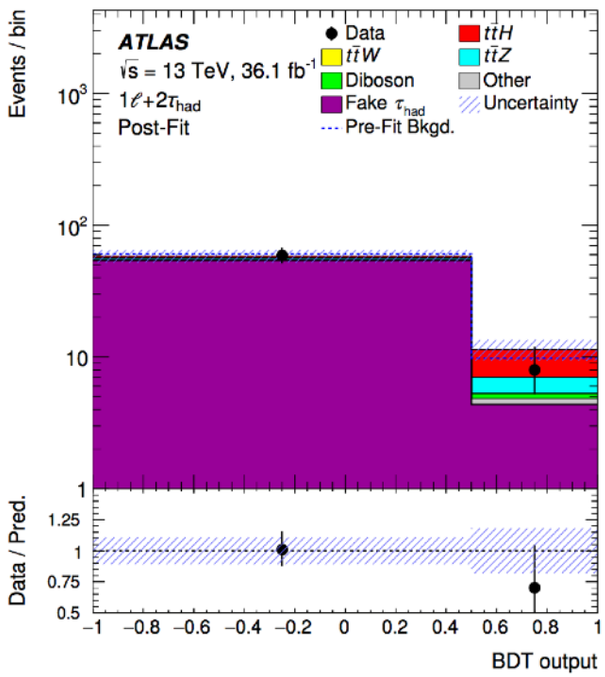
ttH ML - Post-fit yields in SRs

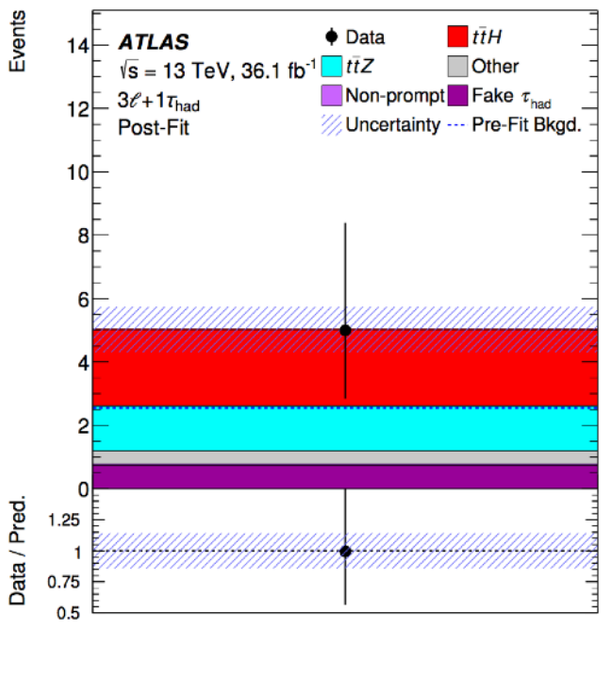


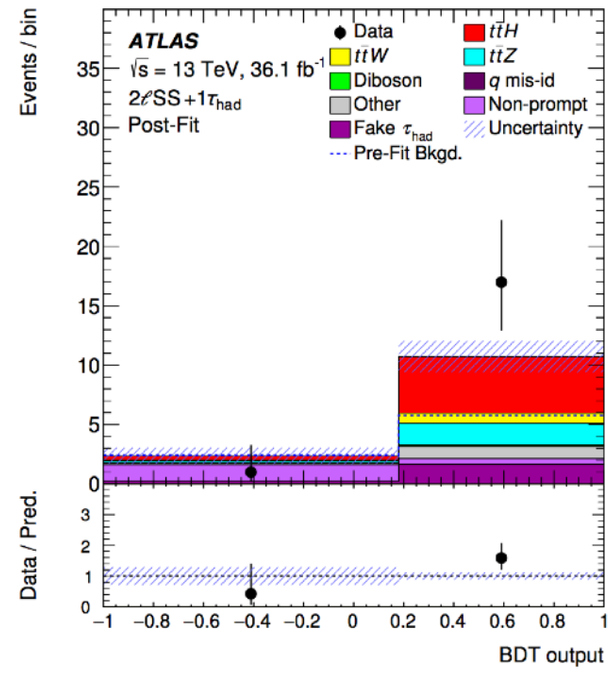


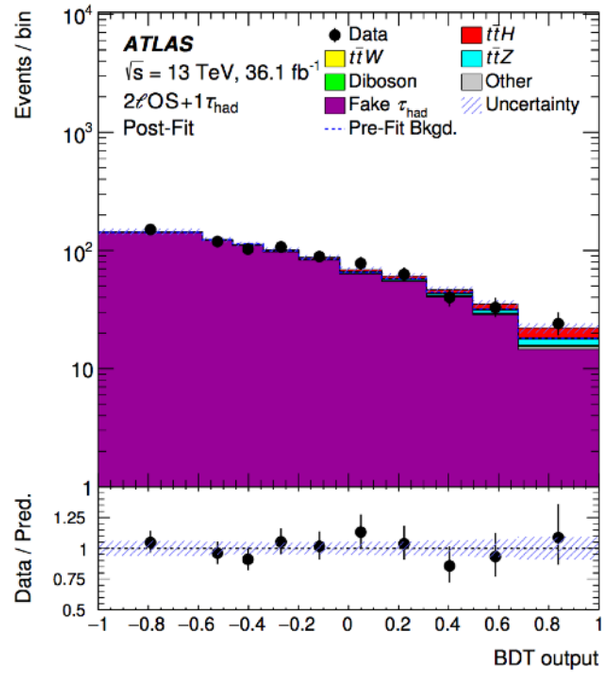




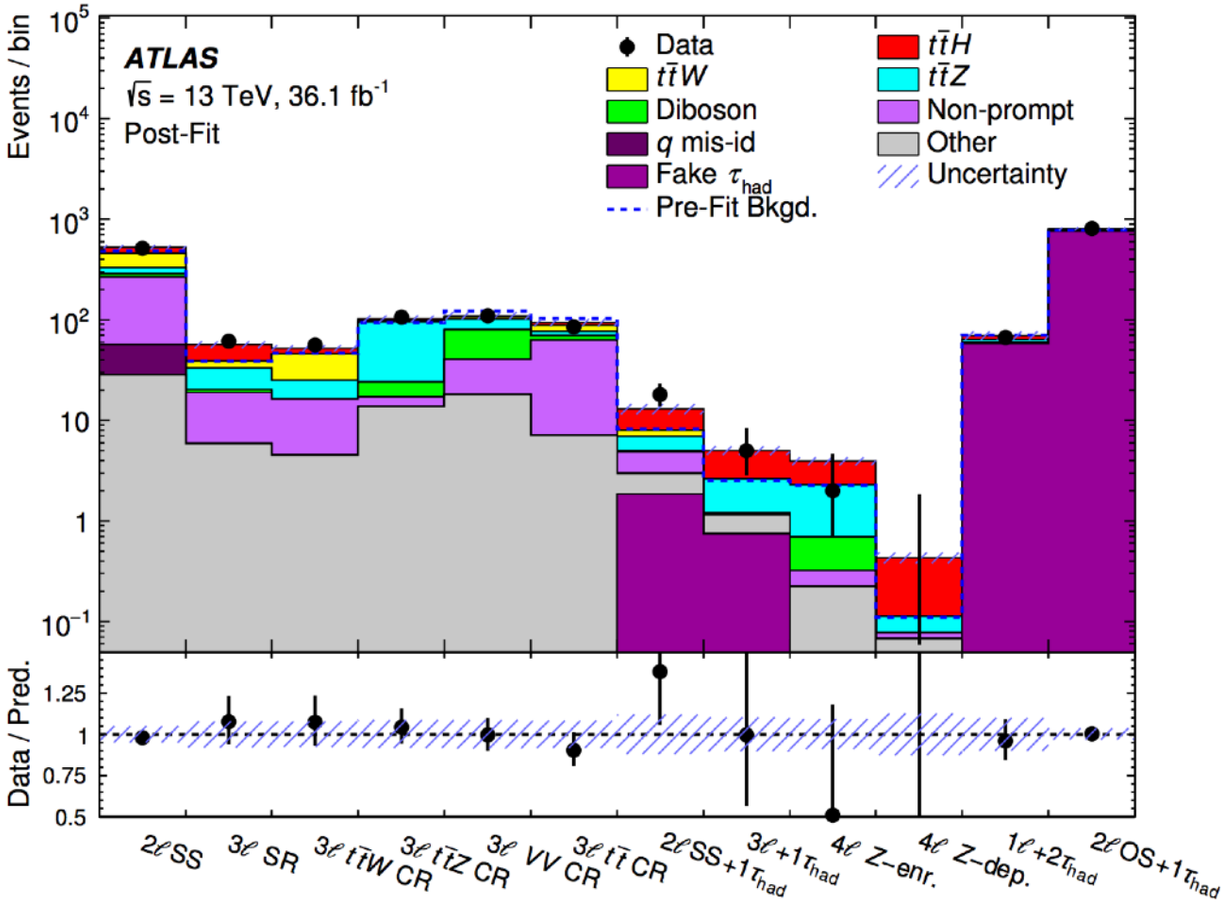


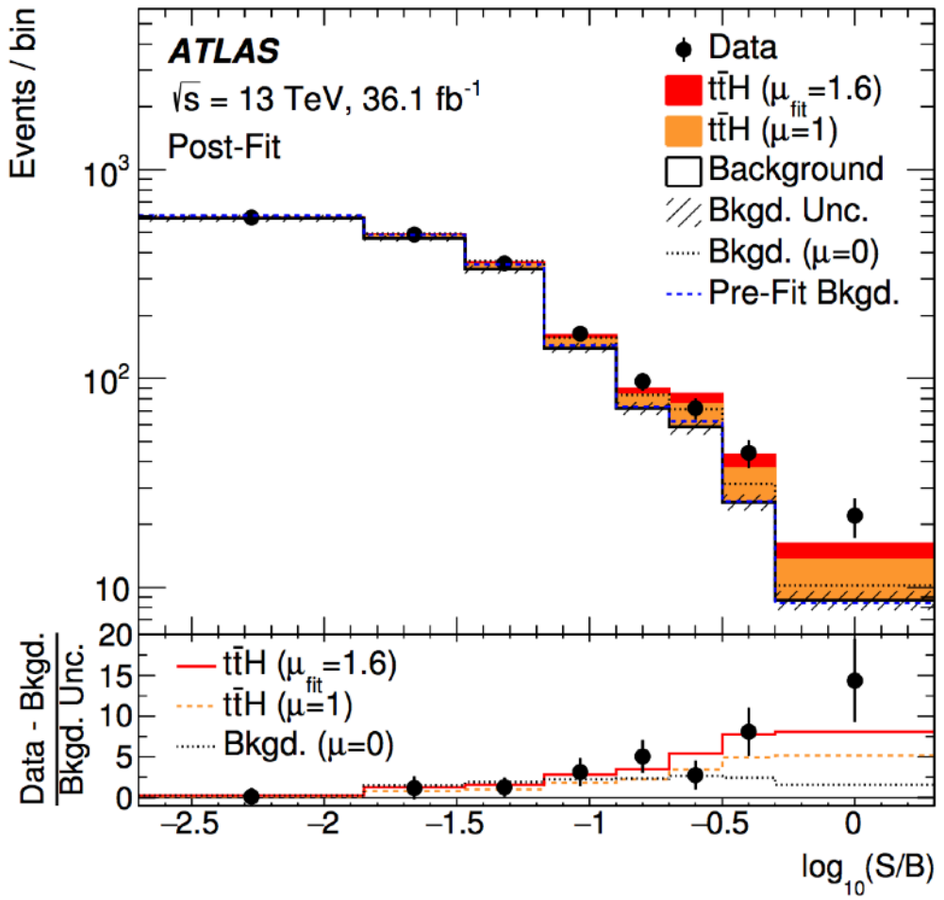






ttH ML - results

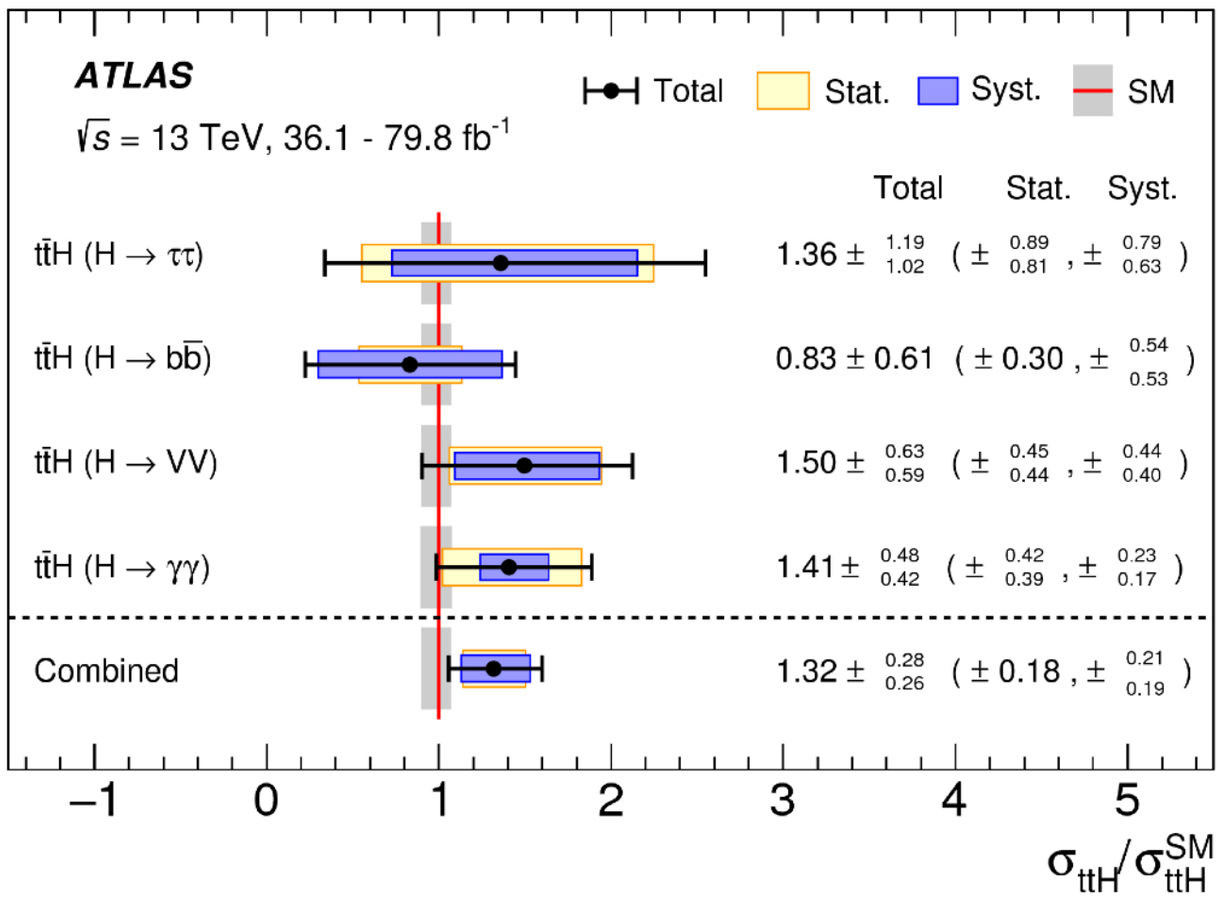




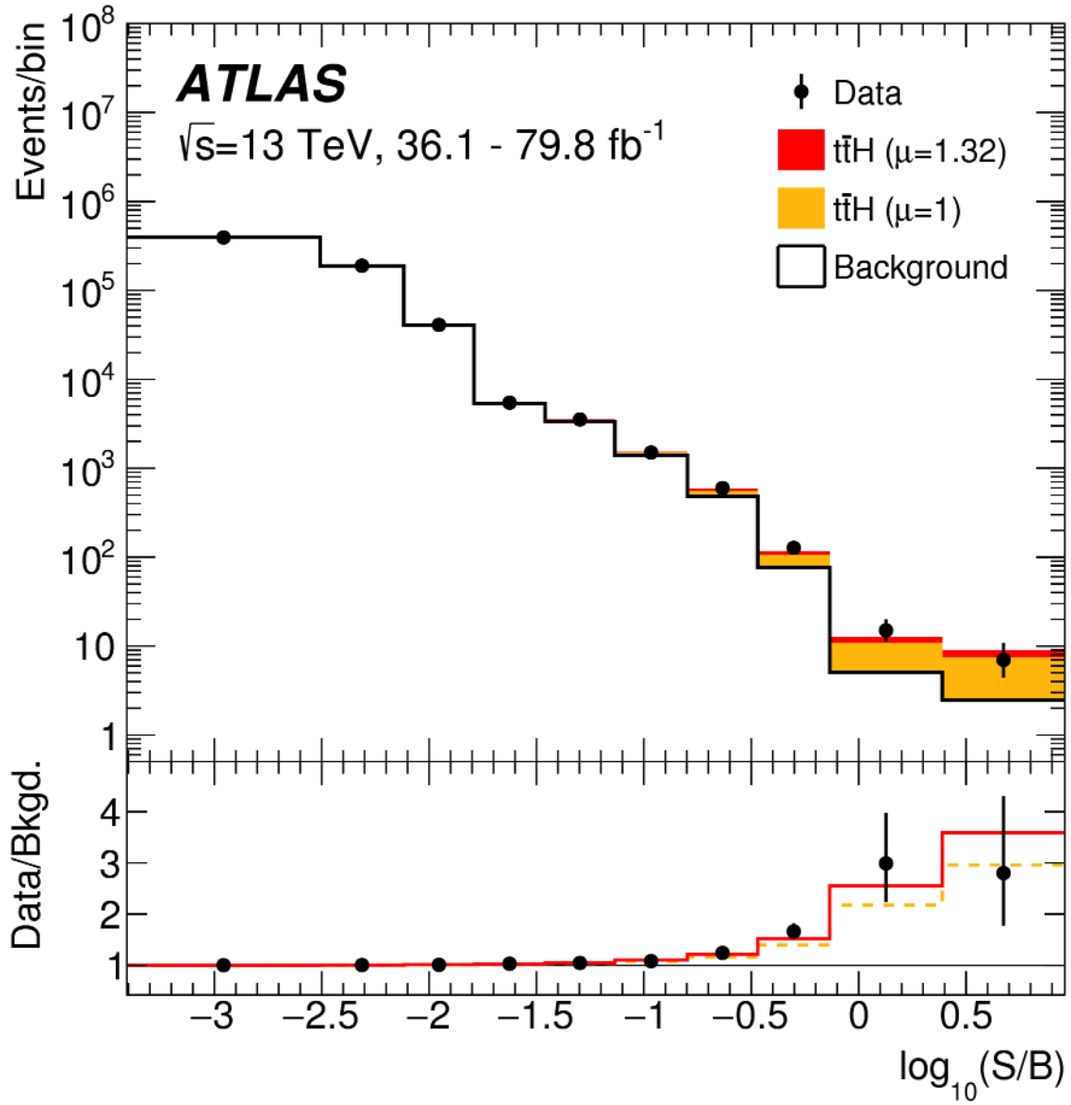
Channel	Best-fit μ		Significance	
	Observed	Expected	Observed	Expected
2ℓOS+1τ_{had}	$1.7\ ^{+1.6}_{-1.5}$ (stat.) $^{+1.4}_{-1.1}$ (syst.)	$1.0\ ^{+1.5}_{-1.4}$ (stat.) $^{+1.2}_{-1.1}$ (syst.)	0.9σ	0.5σ
1ℓ+2τ_{had}	$-0.6\ ^{+1.1}_{-0.8}$ (stat.) $^{+1.1}_{-1.3}$ (syst.)	$1.0\ ^{+1.1}_{-0.9}$ (stat.) $^{+1.2}_{-1.1}$ (syst.)	–	0.6σ
4ℓ	$-0.5\ ^{+1.3}_{-0.8}$ (stat.) $^{+0.2}_{-0.3}$ (syst.)	$1.0\ ^{+1.7}_{-1.2}$ (stat.) $^{+0.4}_{-0.2}$ (syst.)	–	0.8σ
3ℓ+1τ_{had}	$1.6\ ^{+1.7}_{-1.3}$ (stat.) $^{+0.6}_{-0.2}$ (syst.)	$1.0\ ^{+1.5}_{-1.1}$ (stat.) $^{+0.4}_{-0.2}$ (syst.)	1.3σ	0.9σ
2ℓSS+1τ_{had}	$3.5\ ^{+1.5}_{-1.2}$ (stat.) $^{+0.9}_{-0.5}$ (syst.)	$1.0\ ^{+1.1}_{-0.8}$ (stat.) $^{+0.5}_{-0.3}$ (syst.)	3.4σ	1.1σ
3ℓ	$1.8\ ^{+0.6}_{-0.6}$ (stat.) $^{+0.6}_{-0.5}$ (syst.)	$1.0\ ^{+0.6}_{-0.5}$ (stat.) $^{+0.5}_{-0.4}$ (syst.)	2.4σ	1.5σ
2ℓSS	$1.5\ ^{+0.4}_{-0.4}$ (stat.) $^{+0.5}_{-0.4}$ (syst.)	$1.0\ ^{+0.4}_{-0.4}$ (stat.) $^{+0.4}_{-0.4}$ (syst.)	2.7σ	1.9σ
Combined	$1.6\ ^{+0.3}_{-0.3}$ (stat.) $^{+0.4}_{-0.3}$ (syst.)	$1.0\ ^{+0.3}_{-0.3}$ (stat.) $^{+0.3}_{-0.3}$ (syst.)	4.1σ	2.8σ

Observation of ttH production

Combined $t\bar{t}H$ production cross section, as well as cross sections measured in the individual decay channels, divided by the SM cross-section prediction. $H \rightarrow ZZ^*$ and $H \rightarrow WW^*$ are assumed to be scaled by the same factor with respect to the SM prediction, and are shown together as VV.



Observed event yields in all analysis categories in up to 79.8 fb-1 of 13 TeV data. The background yields correspond to the observed fit results, and the signal yields are shown for both the observed results (μ = 1.32) and the SM prediction (μ = 1). The discriminant bins in all categories are ranked by log10(S/B), where S is the signal yield and B the background yield extracted from the fit with freely floating signal, and combined such that log10(S+B) decreases approximately linearly.



ttH(γγ) and ttH(4l): results

☆ ttH→ZZ*→4l

- No events observed
- Analysis statistically limited

☆ ttH→γγ

- about 50% more sensitive than the one @36 fb^{-1}
- Analysis statistically limited

Bin	Expected $t\bar{t}H$ (signal)	Non-$t\bar{t}H$ Higgs	Non-Higgs	Total	Observed Total
		$H \to \gamma\gamma$			
Had 1	4.2(11)	0.49(33)	1.76(55)	6.4(13)	10
Had 2	3.41(74)	0.69(56)	7.5(11)	11.6(15)	14
Had 3	4.70(88)	2.0(17)	32.9(22)	39.6(32)	47
Had 4	3.00(55)	3.2(31)	55.0(28)	61.3(47)	67
Lep 1	4.5(10)	0.25(9)	2.19(59)	6.9(12)	7
Lep 2	2.23(39)	0.27(10)	4.59(91)	7.1(10)	7
Lep 3	0.82(18)	0.30(13)	4.58(91)	5.70(88)	5
		$H \to ZZ^* \to 4\ell$			
Had 1	0.169(31)	0.021(7)	0.008(8)	0.198(33)	0
Had 2	0.216(32)	0.20(9)	0.22(12)	0.63(16)	0
Lep	0.212(31)	0.0256(23)	0.015(13)	0.253(34)	0

ttH(4l): results

☆ ttH→ZZ*→4l

- No events observed
- Analysis statistically limited