Measurement of Higgs boson production in association with top quarks with ATLAS

Rencontres du Vietnam 2018, Quy Nhon

Johnny Raine (Université de Genève)

On behalf of the ATLAS collaboration

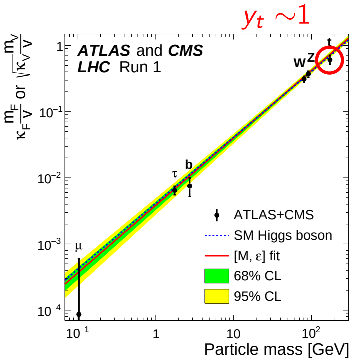
8th August, 2018

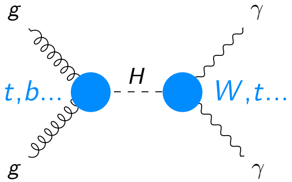
ATLAS EXPERIMENT

Introduction

Higgs Yukawa couplings

- All current measurements of the Higgs boson have been consistent with SM
- Fermions couple with the Higgs boson through *Yukawa interactions*
  - Coupling strength proportional to fermion mass
  - Largest coupling is to the top quark
  - Sensitive to the scale of new physics!
- y_t mainly constrained from loop processes

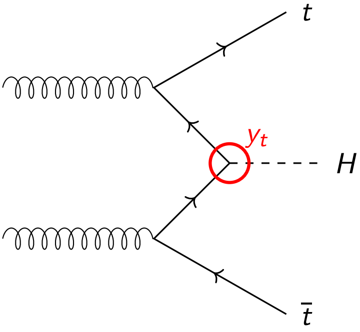




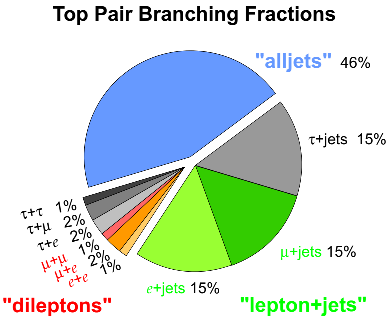
Not model independent, ignores potential BSM contributions

Introduction
$t\bar{t}H$ production

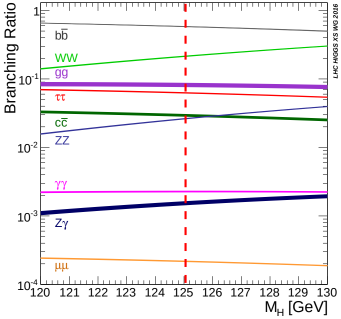
- $t\bar{t}H$: More model independent test of y_t
  - Fourth main Higgs production at LHC
  - Direct measurement of the coupling of Higgs to top quarks
- However, very challenging to measure
  - Small cross section, ∼0.5 pb at 13 TeV
  - Complex final states
  - Large irreducible backgrounds
    - $t\bar{t} + b\bar{b}$, $\mathcal{O}(2)$ magnitudes larger
    - $t\bar{t} + V$, ∼1.5 pb
- Huge efforts to observe $t\bar{t}H$ production during LHC Run 1 and 2



Analysis Strategy



$\otimes$



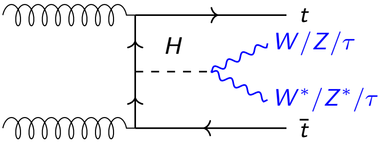
- Wide range of analyses designed to target the various Higgs boson decays
  - Additional considerations to the decay of $t\bar{t}$ pair
  - Final states with many objects: jets, b-jets, e, μ, hadronic τ, photons
  - Huge thanks to the excellent detector performance magnificent effort of ATLAS performance groups

Analysis Strategy

- Four analyses targetting different Higgs decay modes
- Wide range of signal purity and expected yields
- Analysed separately before entering combined analysis

$t\bar{t}H$ multilepton

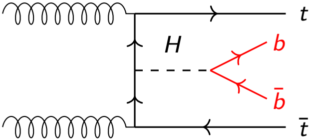
36.1 fb^{-1}, 13 TeV ▸ arXiv



S/B 5–34%

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

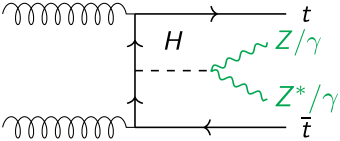
36.1 fb^{-1}, 13 TeV ▸ arXiv



S/B 1.8–5.5%

$t\bar{t}H$ enriched in $H \to \gamma\gamma/4\ell$

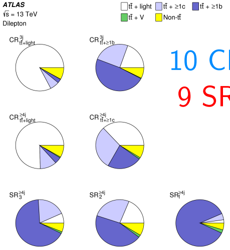
79.8 fb^{-1}, 13 TeV ▸ arXiv



S/B 5–200% ($\gamma\gamma$), 50–500% ($ZZ^* \to 4\ell$)

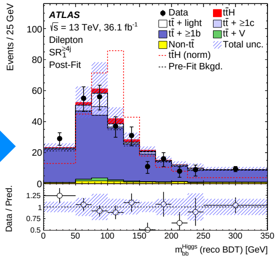
Analysis Strategy
$t\bar{t}H$ $(H \to b\bar{b})$

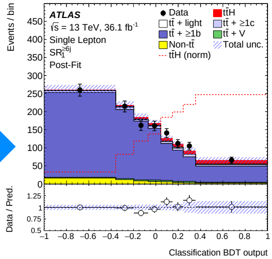
- Benefit from large $H \to b\bar{b}$ BR, selects leptonic top decays
- Large irreducible background from $t\bar{t}+\text{jets}$, especially $t\bar{t}$ +Heavy Flavour
  - Large theory uncertainties, biggest source of systematic uncertainty
- Use of MVA techniques in signal regions to enhance signal sensitivity



10 CRs

9 SRs





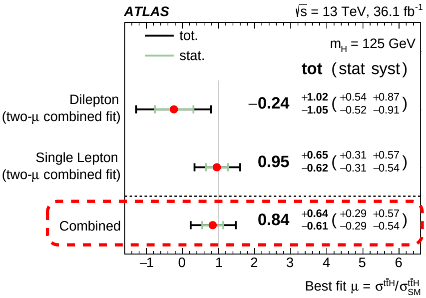
Categorisation
- Use b-tagging of jets and object multiplicities
- Dedicated boosted region targets high p_T top/Higgs

Reconstruction
- Solve object combinatorics to reconstruct final state
- Reco BDT, MEM and Likelihood discriminants

Classification
- BDTs for $t\bar{t}H$ vs $t\bar{t}+\text{jets}$
- Optimised in all SRs
- Reconstruction + event kinematic variables

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

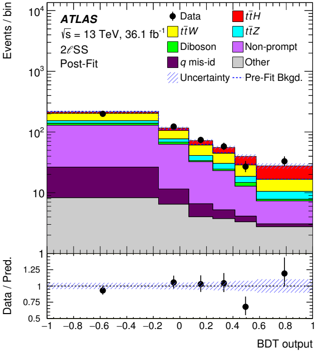
- Binned profile likelihood over all regions
- $t\bar{t}$+$\geq$1b, $t\bar{t}$+$\geq$1c normalisation factors kept free floating
- Significance of 1.4σ (1.6σ expected)
- Systematically limited by modelling of $t\bar{t}$ + HF background

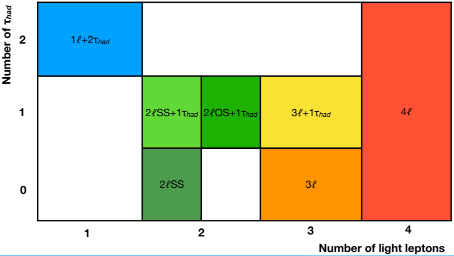


Uncertainty source	$\Delta\mu$	
$t\bar{t}$+$\geq$1b modelling	+0.46	−0.46
Background-model stat. unc.	+0.29	−0.31
b-tagging efficiency and mis-tag rates	+0.16	−0.16
Jet energy scale and resolution	+0.14	−0.14
$t\bar{t}H$ modelling	+0.22	−0.05
$t\bar{t}$+$\geq$1c modelling	+0.09	−0.11
JVT, pileup modelling	+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
$t\bar{t}$+$\geq$1b normalisation	+0.09	−0.10
$t\bar{t}$+$\geq$1c normalisation	+0.02	−0.03
Intrinsic statistical uncertainty	+0.21	−0.20
Total statistical uncertainty	+0.29	−0.29
Total uncertainty	+0.64	−0.61

Analysis Strategy
$t\bar{t}H$ multileptons

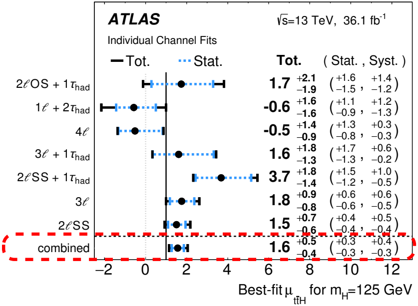
- Target Higgs decays with leptonic final states and leptonic $t\bar{t}$ decays
- Same sign and >3 lepton events reduce $t\bar{t}$ background
  - Requirements on (b-)jet multiplicities
  - Events categorised by number of leptons & hadronic taus
  - Wide range of yields and S/B purity
  - Seven final states in total





- Object level BDTs used to reduce non-prompt leptons and charge mis-ID
- Enhance separation from $t\bar{t}$, $t\bar{t}V$ with BDTs
  - Event count in $3\ell 1\tau_{had}$ and $4l$

Analysis Strategy
$t\bar{t}H$ multileptons Results



Channel	Significance Obs.	Significance Exp.
$2\ell OS + 1\tau_{had}$	0.9σ	0.5σ
$1\ell + 2\tau_{had}$	-	0.6σ
4ℓ	-	0.8σ
$3\ell + 1\tau_{had}$	1.3σ	0.9σ
$2\ell SS + 1\tau_{had}$	3.4σ	1.1σ
3ℓ	2.4σ	1.5σ
$2\ell SS$	2.6σ	1.9σ
Combined	4.1σ	2.8σ

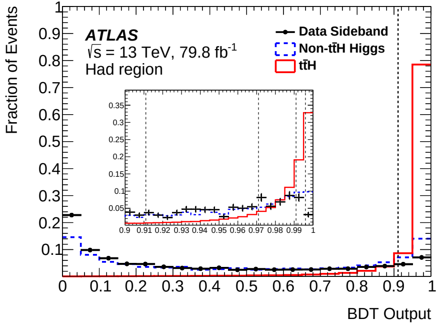
- Binned profile likelihood across all regions
- Observed significance of 4.1σ for $t\bar{t}H$ production (2.8σ exp)
- Additional cut based cross check analysis performed
  - Consistent results with the MVA based approach
  - 15% poorer sensitivity
- Leading systematics from $t\bar{t}H$ and $t\bar{t}V$ modelling, non-prompt lepton estimates and jet energy scale/resolution

Analysis Strategy
$t\bar{t}H\,(\gamma\gamma)$

79.8 fb^{-1}!!!

- Small rate but very signal enriched regions with a continuous background
- Reconstruct Higgs as a narrow peak, use side bands to estimate background
  - Main background from non-resonant $\gamma\gamma$ and non-$t\bar{t}H$ production

- Categorise events by leptonic ($>1\ell$) and hadronic (0ℓ) $t\bar{t}$ decays
- Train BDTs to separate $t\bar{t}H$ from background in lep and had
  - Jet/lepton 4-vector info
  - Photon observables
  - E_T^{miss} and b-tagging
- Cut on BDT distributions to define signal rich regions
  - Seven regions in total



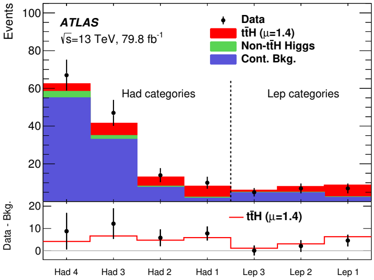
50% improvement in sensitivity

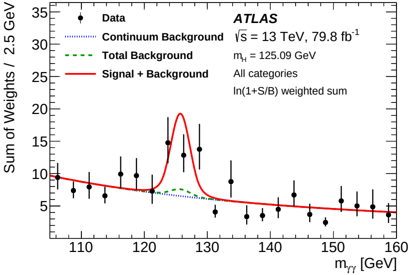
2016 analysis

Analysis Strategy
$t\bar{t}H\,(\gamma\gamma)$ Results

79.8 fb^{-1}!!!

- Unbinned maximum likelihood fit over $m_{\gamma\gamma}$ in range 105 – 160 GeV
  - Non-$t\bar{t}H$ production fixed to SM prediction
  - Function for $\gamma\gamma$ background derived in each regions
    - Leptonic regions: simulation
    - Hadronic regions: data driven from control region
- Observed significance of 4.1σ for $t\bar{t}H$ production (3.7σ expected)
  - Measured signal strength $\mu = 1.39^{+0.48}_{-0.42} = 1.39^{+0.42}_{-0.36}(\text{stat.})^{+0.23}_{-0.17}(\text{syst.})$
- Currently statistically limited ($\sim$29% stat uncertainty)



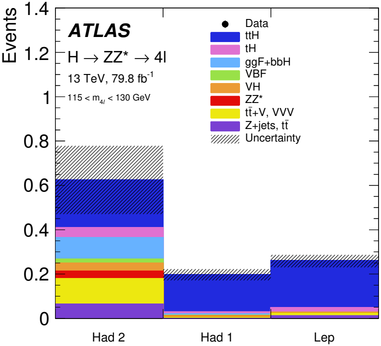


Analysis Strategy
$t\bar{t}H$ (4ℓ)

79.8 fb^{-1}!!!

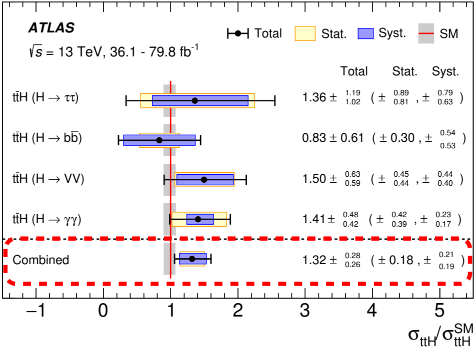
- Extremely low rate but very high signal to background ratio (up to 500%!)
- Look at 4ℓ inv-mass window 115–130 GeV
- Categorise events by $t\bar{t}$ decay: leptonic (1 additional ℓ) and hadronic (0 additional ℓ)
  - Further split hadronic into two bins with BDT to enhance $t\bar{t}H$ purity
- No observed events
  - Fewer than one expected event
- Expected significance of 1.2σ
- Very statistically limited

Region	Expected $t\bar{t}H$	Expected Non-$t\bar{t}H$ Higgs	Expected Other bkg	Observed
Had 2	0.169(31)	0.021(7)	0.008(8)	0
Had 1	0.216(32)	0.20(9)	0.22(12)	0
Lep	0.212(31)	0.0256(23)	0.015(13)	0



Combined Result

- Combination of all four analyses performed using profile likelihood method



Fit details

- Negligible overlap between events in each analysis
- Non-$t\bar{t}H$ Higgs production fixed to SM prediction
- Correlation scheme studied in detail

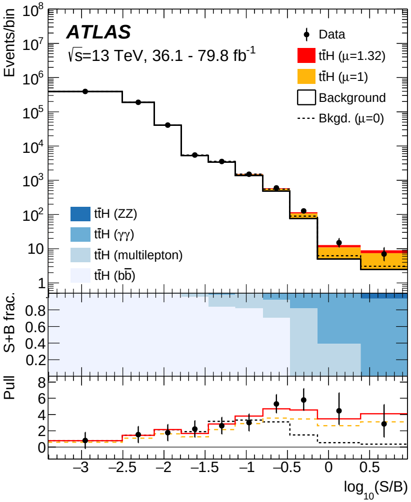
- Observation of $t\bar{t}H$ production at 13 TeV. 5.8σ observed (4.9σ expected)
- Measured $t\bar{t}H$ cross section at $\sqrt{s} = 13$ TeV:

$$\sigma_{t\bar{t}H} = 670 \pm 90(\text{stat})^{+110}_{-100}(\text{syst})\ \text{fb}^{-1}$$

- Cross section 1.32× SM prediction, compatible with SM at around 1σ level

arXiv

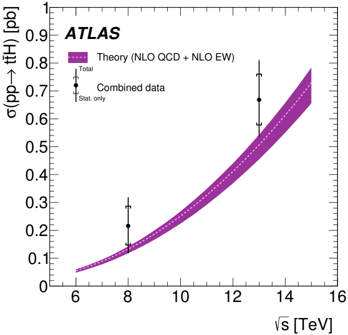
Combined Result



- Some channels still very much limited by statistics
- Modelling uncertainties dominate the systematic uncertainties

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-ttH 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

Combination with Run 1



- Combine the 13 TeV result with the ATLAS Run 1 result ▸ EPJC 76 (2016) 6
  - Additional 4.5 fb^{-1} 7 TeV and 20.3 fb^{-1} 8 TeV data
- 6.3σ observed significance, 5.1σ expected!

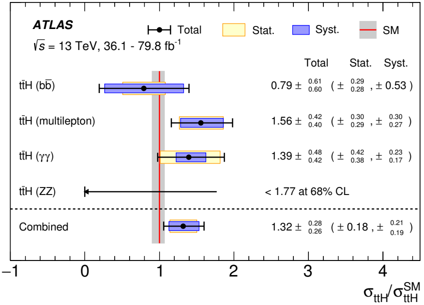
Analysis	Integrated luminosity $[\text{fb}^{-1}]$	$t\bar{t}H$ cross section [fb]	Obs. sign.	Exp. sign.
$H \to \gamma\gamma$	79.8	$710\ ^{+210}_{-190}$ (stat.) $^{+120}_{-90}$ (syst.)	$4.1\ \sigma$	$3.7\ \sigma$
$H \to$ multilepton	36.1	790 ± 150 (stat.) $^{+150}_{-140}$ (syst.)	$4.1\ \sigma$	$2.8\ \sigma$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

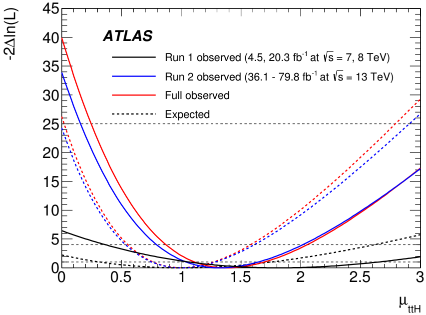
Conclusion

- Search for $t\bar{t}H$ production performed at 13 TeV using 36.1 – 79.8 fb^{-1} data
- Combination of several challenging analyses
  - Extensive use of multivariate techniques to enhance sensitivity
  - Large systematic uncertainties on modelling
  - Some channels statistically limited, will only become more sensitive!
- ATLAS observation of $t\bar{t}H$ with a significance of 6.3σ (5.1σ exp)
  - Direct observation of top Yukawa coupling
  - Measured $\sigma_{t\bar{t}H} = 670 \pm 90(\mathrm{stat})^{+110}_{-100}(\mathrm{syst})$ fb^{-1} at 13 TeV
  - Consistent with SM prediction $\sigma_{t\bar{t}H} = 507^{+35}_{-50}$ fb^{-1}

Backup

Combination

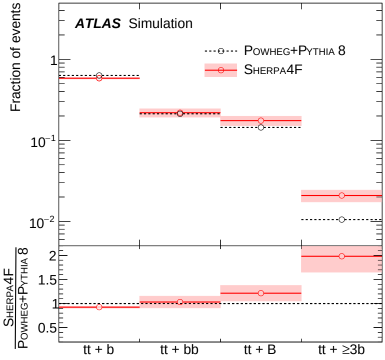




$t\bar{t}H$ $(H \rightarrow b\bar{b})$ - $t\bar{t}$ Background Model

Modelling of $t\bar{t}$ is crucial to the analysis, $t\bar{t} + \mathrm{HF}$ has large theory uncertainty

- Split into $t\bar{t} + \mathrm{light}$, $t\bar{t}+\geq 1c$, $t\bar{t}+\geq 1b$
  - Further split $t\bar{t}+\geq 1b$ by number of additional b-hadrons in jets
- Nominal $t\bar{t}$ sample uses 5FS prediction
  - Use dedicated Sherpa 4FS $t\bar{t} + b\bar{b}$ prediction to improve modelling
    - Both additional b-quarks to NLO precision in QCD
    - Takes b-quark mass into account
  - Reweight relative $t\bar{t}+\geq 1b$ subcomponents to 4FS values

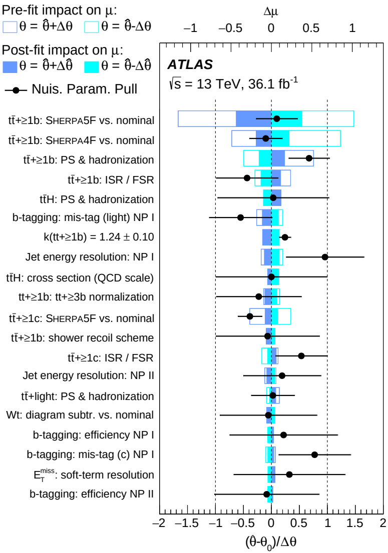


$t\bar{t}H$ $(H \to b\bar{b})$ - $t\bar{t}$ Systematic Model

$t\bar{t}$ modelling is dominant contribution to total systematic uncertainty

Systematic source	Description	$t\bar{t}$ categories
$t\bar{t}$ cross-section	Up or down by 6%	All, correlated
$k(t\bar{t}+\geq 1c)$	Free-floating $t\bar{t}+\geq 1c$ normalization	$t\bar{t}+\geq 1c$
$k(t\bar{t}+\geq 1b)$	Free-floating $t\bar{t}+\geq 1b$ normalization	$t\bar{t}+\geq 1b$
SHERPA5F vs. nominal	Related to the choice of NLO event generator	All, uncorrelated
PS & hadronization	POWHEG+HERWIG 7 vs. POWHEG+PYTHIA 8	All, uncorrelated
ISR / FSR	Variations of μ_{R}, μ_{F}, h_{damp} and A14 Var3c parameters	All, uncorrelated
$t\bar{t}+\geq 1c$ ME vs. inclusive	MG5_aMC@NLO+HERWIG++: ME prediction (3F) vs. incl. (5F)	$t\bar{t}+\geq 1c$
$t\bar{t}+\geq 1b$ SHERPA4F vs. nominal	Comparison of $t\bar{t}+b\bar{b}$ NLO (4F) vs. POWHEG+PYTHIA 8 (5F)	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ 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 $H_{\mathrm{T}}/2$ 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 scheme	Alternative model scheme	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ PDF (MSTW)	MSTW vs. CT10	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ PDF (NNPDF)	NNPDF vs. CT10	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ UE	Alternative set of tuned parameters for the underlying event	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ MPI	Up or down by 50%	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 3b$ normalization	Up or down by 50%	$t\bar{t}+\geq 1b$

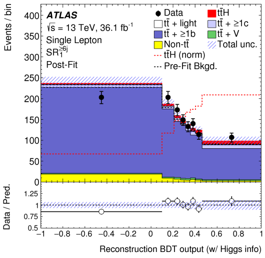
$t\bar{t}H$ $(H \to b\bar{b})$ - Impact of Systematic Uncertainties



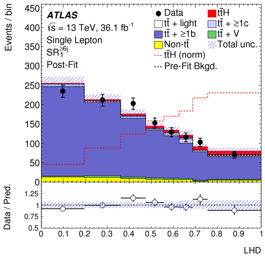
- Analysis is currently systematically limited
- Largest uncertainties from $t\bar{t} + \mathrm{HF}$ modelling
- Also notable impact:
  - Limited MC stats.
  - Flavour tagging
  - Jet energy scale and resolution
- Large number of constrained two-point systematics

$t\bar{t}H\ (H \to b\bar{b})$ - Reconstruction Methods

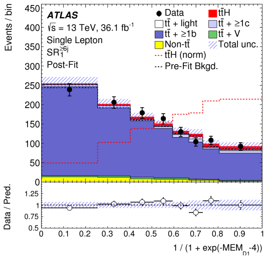
Reco BDT



LHD



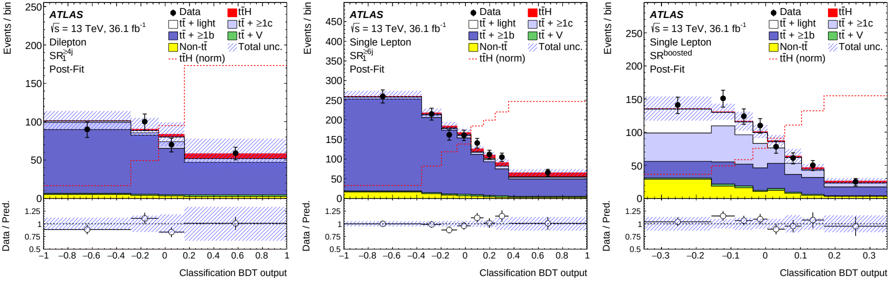
MEM



Reco BDT	Exploits correlations within each combination
Reco BDT	Provides jet assignments based on $t\bar{t}H\ (H \to b\bar{b})$
LHD	Combines all combinations together into one discriminant
LHD+MEM	Calculate both signal and background likelihoods
MEM	Calculates discriminant at parton level using 4-vectors

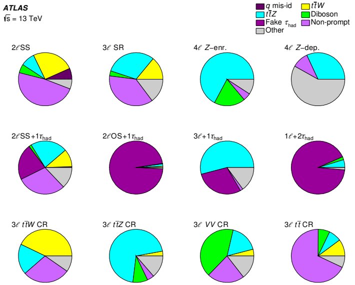
$t\bar{t}H$ $(H \to b\bar{b})$- Final Discriminant

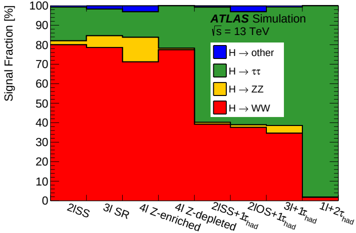
- Looking at three signal regions post fit
  - $t\bar{t}H$ shown for extracted signal strength $\mu = 0.84^{+0.64}_{-0.61}$
  - Showing two most signal enriched regions and boosted signal region
- See good post-fit agreement between data and simulation in all regions



$t\bar{t}H$ ML - Regions

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



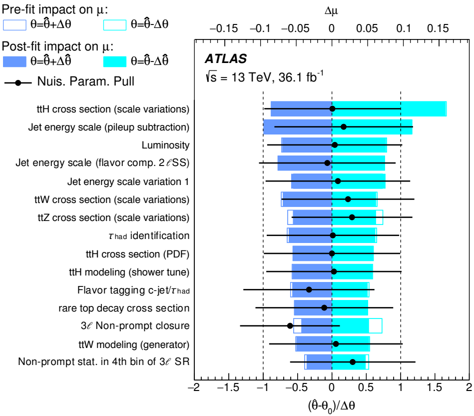


$t\bar{t}H$ ML - BDTs

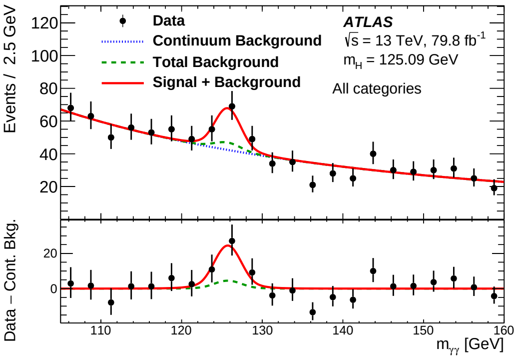
	2ℓSS	3ℓ	4ℓ	1ℓ+$2\tau_{\text{had}}$	2ℓSS+$1\tau_{\text{had}}$	2ℓOS+$1\tau_{\text{had}}$	3ℓ+$1\tau_{\text{had}}$
Light lepton	2T*	1L*, 2T*	2L, 2T	1T	2T*	$2\text{L}^{\dagger}$	$1\text{L}^{\dagger}$, 2T
τ_{had}	0M	0M	–	1T, 1M	1M	1M	1M
N_{jets}, $N_{b\text{-jets}}$	$\geq 4, = 1, 2$	$\geq 2, \geq 1$	$\geq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

$t\bar{t}H$ ML - Systematics

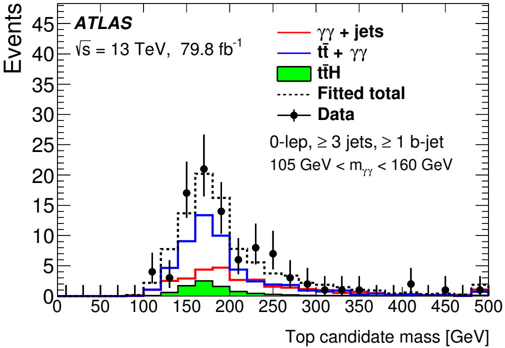
Uncertainty Source	$\Delta\mu$	
$t\bar{t}H$ modeling (cross section)	+0.20	−0.09
Jet energy scale and resolution	+0.18	−0.15
Non-prompt light-lepton estimates	+0.15	−0.13
Jet flavor tagging and τ_{had} identification	+0.11	−0.09
$t\bar{t}W$ modeling	+0.10	−0.09
$t\bar{t}Z$ modeling	+0.08	−0.07
Other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



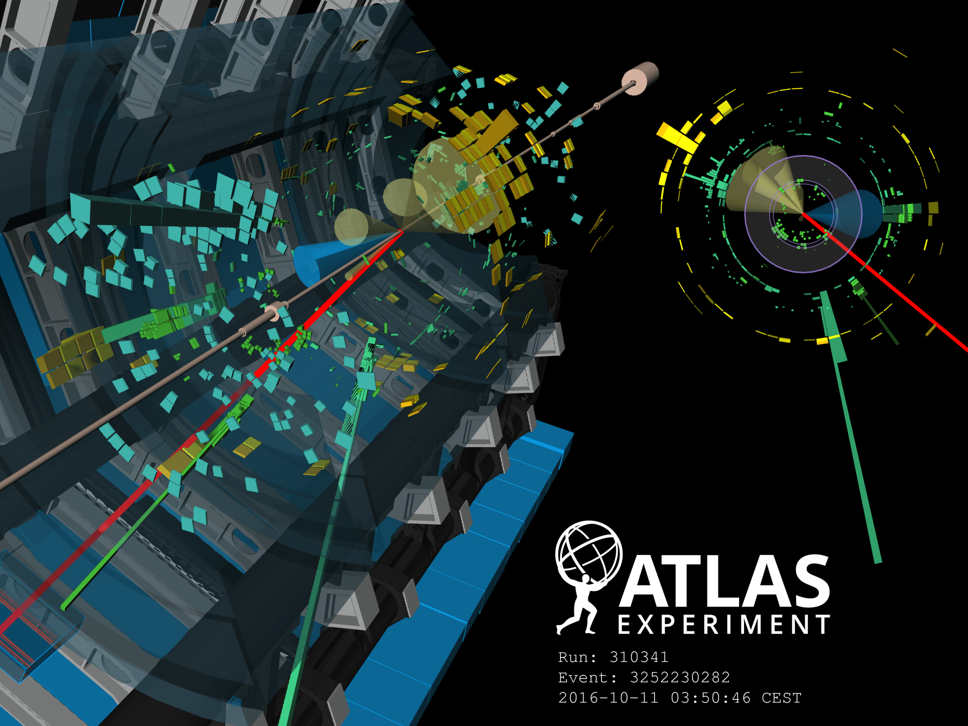
$t\bar{t}H\,(H \to \gamma\gamma)$ - Unweighted $m_{\gamma\gamma}$



$t\bar{t}H\,(H \rightarrow \gamma\gamma)$ - Unweighted $m_{\gamma\gamma}$



- BDT trained to select three jets form hadronic top
- Does not enter the analysis
- Top mass reconstructed in bins with highest S/B
- Excess in events around top mass consistent with $t\bar{t}H$


ATLAS
EXPERIMENT
Run: 310341
Event: 3252230282
2016-10-11 03:50:46 CEST