Searching for $t\bar{t}H$ $\left(H \to b\bar{b}\right)$ with ATLAS

RAL Department Seminar

Johnny Raine (Université de Genève)

16th May, 2018

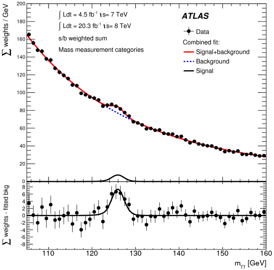
Introduction

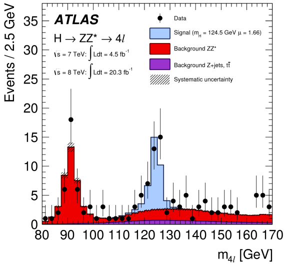
- Aim to give an overview of the search for $t\bar{t}H$ production with the ATLAS detector
  - Focus on the $H \to b\bar{b}$channel
  - Describe the full analysis strategy
  - Present the results and current status
- Analysis presented uses 32.1 fb^{-1} dataset of pp collisions
  - Paper submitted to PRD

ttHbb - PRD

Introduction
Higgs boson

- Discovered in 2012 by ATLAS and CMS collaborations
  - The last piece of the SM to be found
  - Its mass is unconstrained in the SM
- Want to measure as many properties of the particle as possible
  - Mass, charge, width, CP-nature...
  - But also, how it couples to other particles
  - Coupling strengths predicted by SM given Higgs mass
  - Currently all measurements are consistent with SM

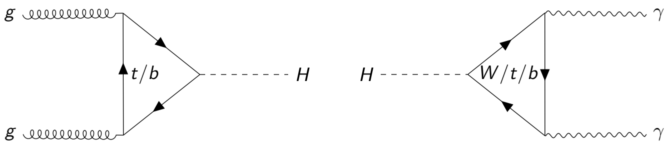




Introduction
Higgs couplings

Bosons

- Higgs coupling to all Bosons has been observed
- Directly to W and Z bosons
- Indirectly to γ and gluons through loop processes
  - $\gamma\gamma$ dominated by W bosons in the loop, top contributes
  - gluon loop dominated by top, small contribution from b-quarks



Introduction
Higgs couplings

Bosons

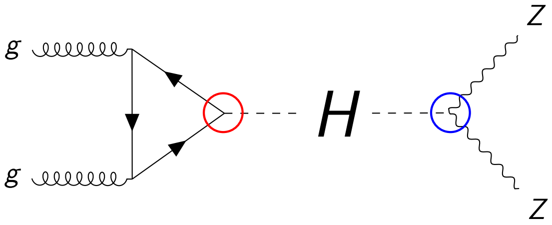
- Higgs coupling to all Bosons has been observed
- Directly to W and Z bosons
- Indirectly to γ and gluons through loop processes
  - $\gamma\gamma$ dominated by W bosons in the loop, top contributes
  - gluon loop dominated by top, small contribution from b-quarks

Fermions

- Less success with measuring the coupling to fermions
- Only direct observation of one Higgs-feriom coupling: $H \to \tau\tau$
- Only evidence for $H \to b\bar{b}$ decay from ATLAS and CMS
- Top-Higgs coupling inferred through loop processes

Introduction
Higgs couplings

- Two areas to probe for Higgs couplings
  - In Higgs production modes and in decay channels
  - Additionally of interest is the Higgs self coupling (di-Higgs searches)

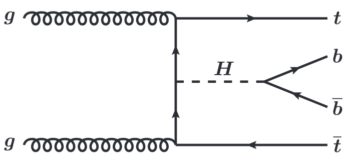


- Shown here is ggF production with $H \to ZZ$ decay
  - Production dominated by coupling to top y_t and b-quarks y_b
  - Decay has direct coupling to Z boson

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

Direct measurement of the two largest Higgs fermion couplings

- $t\bar{t}H$:
  - Direct measurement of y_t
  - y_t can probe scale of new physics
  - One of four main Higgs production mechanisms at LHC
- $H \to b\bar{b}$
  - Largest Higgs BR (58%)
  - In ggF, dominated by multijet background
  - S/B improved by additional final state objects (VH and $t\bar{t}H$ production)
  - Only evidence from $VH(b\bar{b})$



Search for $t\bar{t}H$ $\left(H \rightarrow b\bar{b}\right)$

Analysis Overview

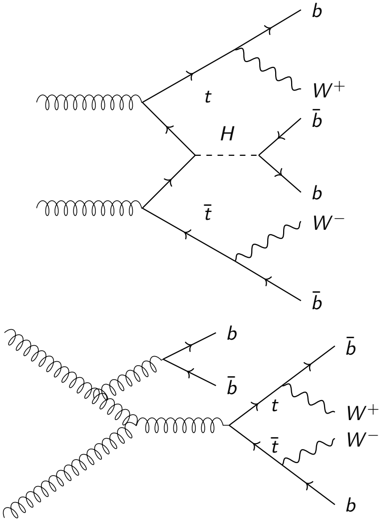
- Large BR but dominated by background from $t\bar{t}$+jets events
  - $t\bar{t}H$ $(H \rightarrow b\bar{b})$ has same final state objects as $t\bar{t} + b\bar{b}$ production
  - Cross section is ∼2 orders magnitude larger than signal ($\sigma_{t\bar{t}H} = 0.5$ pb at $\sqrt{s} = 13$ TeV)
- Large focus of analysis on separating signal from background
  - Event selection and categorisation
  - Multivariate techniques (Reconstruction and Classification)
- Controlling the background modelling and systematic uncertainties
  - The dominant $t\bar{t} + \mathrm{jets}$ background has large systematic uncertainties
  - Perform a simultaneous profile likelihood fit on signal and control regions
- Analysis is split into four steps:
  - Event selection, which split into channels based on lepton number
  - Event categorisation, performed in each channel
  - Reconstruction and Classification
  - Signal strength extraction

Background Modelling

- Search is dominated by $t\bar{t} + \mathrm{jets}$ background
  - Split into $t\bar{t}{+}{\geq}1b$, $t\bar{t}{+}{\geq}1c$ and $t\bar{t} + \mathrm{light}$
  - Defined by matching of b/c-hadrons to additional jets at particle level
- $t\bar{t} + b\bar{b}$ has same final state as $t\bar{t}H\left(H \rightarrow b\bar{b}\right)$ signal
  - $t\bar{t}{+}{\geq}1b$ and $t\bar{t}{+}{\geq}1c$ production not well understood
- Large number of systematic uncertainties cover $t\bar{t} + \mathrm{HF}$ modelling
  - Covering choice of generator, parton shower, 3/4 vs 5FS PDFs
  - Free float normalisation of $t\bar{t}{+}{\geq}1b$ and $t\bar{t}{+}{\geq}1c$ in the fit
- In order to improve the $t\bar{t}{+}{\geq}1b$ modelling, nominal sample has individual $t\bar{t}{+}{\geq}1b$ components adjusted to match dedicated $t\bar{t} + b\bar{b}$ sample produced to NLO precision using 4FS PDF

Event Selection and Categorisation

Event Selection and Categorisation



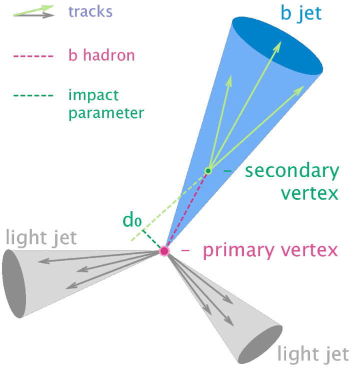
Objects in our Events

- From our signal/background:
  - Four b-quarks
  - Two W-bosons
- $W^{\pm}$ decays to $\ell\nu$ or $q\bar{q}$
  - Always require at least one lepton
  - Provides clean trigger signature[1]
  - Two channels: 1ℓ and 2ℓ
- Detector doesn't see b-quarks
  - Hadronisation and parton shower
  - Collimated shower of particles: Jet reconstruction
  - Attempt to identify jets originating from b-quarks (b-jets)

[1] ATLAS cannot save all events from collisions. Require triggers to decide whether to save an interesting event

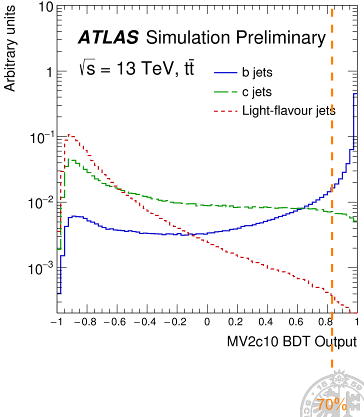
Event Selection and Categorisation
b-tagging

- Jets constructed by grouping energy deposits in the detector (clusters)
  - Use anti-k_t algorithm with $\Delta R = 0.4$
- Exploiting properties of *B*-hadrons to identify *b*-jets
  - Long lifetime $\rightarrow$ flight path in detector
  - Large impact parameter of tracks matched to a secondary vertex
- Three types of algorithms to exploit
  - Impact parameter based
  - Secondary vertex reconstruction
  - Topological decay reconstruction
- Output variables are combined into a single discriminant



Event Selection and Categorisation
b-tagging

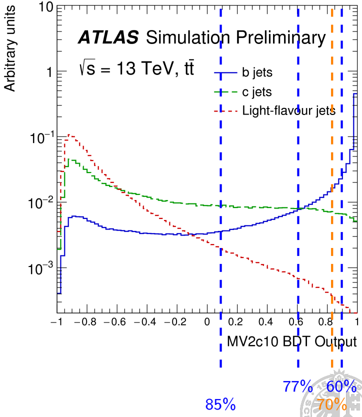
- Boosted Decision Tree to combine multiple input variables
  - Separate *b*-jets from *c*- and light-flavour jets
  - Background is 80% LF, 20% *c*-jet
- Kinematic properties are also included in the training
  - Reweight the distributions to have no kinematic differences
  - But can exploit underlying correlations with other inputs
- Four *b*-tagging efficiency working points
  - 60%, 70%, 77% and 85% *b*-efficiency
  - Define *b*-tagged jets using one WP



Event Selection and Categorisation

b-tagging

- Boosted Decision Tree to combine multiple input variables
  - Separate *b*-jets from *c*- and light-flavour jets
  - Background is 80% LF, 20% *c*-jet
- Kinematic properties are also included in the training
  - Reweight the distributions to have no kinematic differences
  - But can exploit underlying correlations with other inputs
- Four *b*-tagging efficiency working points
  - 60%, 70%, 77% and 85% *b*-efficiency
  - Define *b*-tagged jets using one WP
  - OR use all working points together



Event Selection
Pre-selection

- In the final state expect six (four) b-jets in the 1ℓ (2ℓ) channels
  - Would select four b-jets at the tightest WP
  - Open up the acceptance to take into account detector efficiencies
  - A b-quark could be out of the acceptance of the detector/mistagged
    - Require $\geq$5 ($\geq$3) jets and reduced b-tag requirements in preselection
  - Use jet multiplicity and b-tag working points to define regions
- Additionally: consider a "Boosted" topology in 1ℓ channel
  - High p_T Higgs boson/top quark
  - Jets from decay products have significant overlap, form one fat jat

Event Categorisation

Overview

Overview

- Two methods used to categorise events
  - i Background based categorisation (resolved events)
  - ii Top and Higgs candidate large jet tagging (boosted events)
- Regions are used to define control regions and signal regions
  - Signal regions are enriched in $t\bar{t}H\left(H \rightarrow b\bar{b}\right)$ signal
  - Control regions are all other regions
  - Multivariate techniques are used in signal regions to improve sensitivity of the analysis

Event Categorisation
Overview

Resolved Categorisation

- Start with initial loose preselection of events, consistent with $t\bar{t} + X$
- Use jet info to define regions enriched in different $t\bar{t} + \text{jets}$ composition
  - Split events by jet multiplicity, use *b*-tag WPs of up to four jets

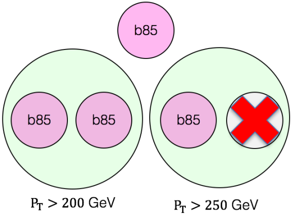
Boosted Categorisation

- Select events with objects corresponding to $t\bar{t}H\,(H \to b\bar{b})$ events with a boosted Higgs and top
  - Require two large jets (anti-k_t with $\Delta R = 1.0$)
    - Tag one as a top candidate
    - Tag the other as a Higgs candidate
- All events go into the boosted region in case of overlap with resolved

Event Categorisation
Boosted

- Only performed in single lepton channel
- Reduces combinatorics of final state objects
- Selection:
  - ≥5 small jets, ≥2 reclustered[1] large jets
  - ≥1 jet tagged@85% WP outside large je
  - ≥1 top candidate, ≥1 Higgs candidate

	Top	Higgs
p_T [GeV]	> 250	> 200
Constituent jets	2	2
Tagged @ 85%	==1	2



[1] Reclustered jets are reconstructed using the anti-k_t algorithm but taking smaller radius jets as inputs instead of clusters

Event Categorisation
Resolved

- Each jet in the event is labelled by the tightest *b*-tag WP it passes
- Separate events into bins of the four jet WPs (jet_1,jet_2,jet_3,jet_4)

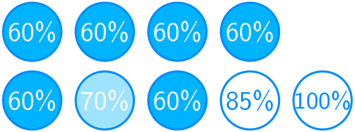
Event Categorisation
Resolved

- Each jet in the event is labelled by the tightest *b*-tag WP it passes
- Separate events into bins of the four jet WPs (jet_1,jet_2,jet_3,jet_4)
  - The tighter the WPs the more signal and $t\bar{t} + b\bar{b}$ enriched the bin

60% 60% 60% 60% $= (60,60,60,60)$ $(t\bar{t}+ \geq 2b)$

Event Categorisation
Resolved

- Each jet in the event is labelled by the tightest *b*-tag WP it passes
- Separate events into bins of the four jet WPs (jet_1,jet_2,jet_3,jet_4)
  - The tighter the WPs the more signal and $t\bar{t} + b\bar{b}$ enriched the bin
  - Bins with looser jets will be enriched in $t\bar{t}+\geq 1c$ and $t\bar{t} + \text{light}$

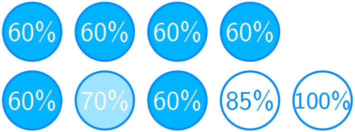


$= (60,60,60,60)$ $(t\bar{t}+ \geq 2b)$

$= (60,60,70,85)$ $(t\bar{t}+\geq 1c)$

Event Categorisation
Resolved

- Each jet in the event is labelled by the tightest b-tag WP it passes
- Separate events into bins of the four jet WPs (jet_1,jet_2,jet_3,jet_4)
  - The tighter the WPs the more signal and $t\bar{t} + b\bar{b}$ enriched the bin
  - Bins with looser jets will be enriched in $t\bar{t}+\geq 1c$ and $t\bar{t} + \text{light}$



$= (60,60,60,60)$ $(t\bar{t}+ \geq 2b)$

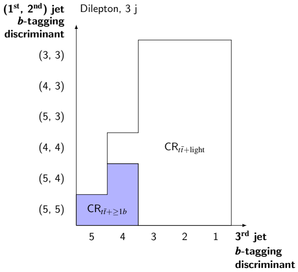
$= (60,60,70,85)$ $(t\bar{t}+\geq 1c)$

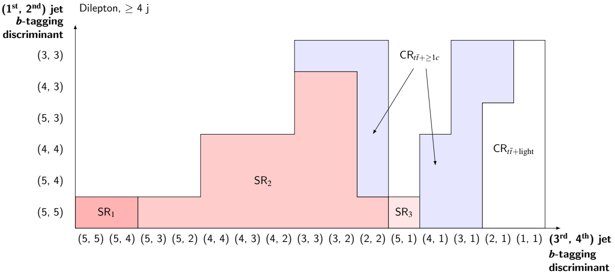
- Each bin will have a different background composition
  - Combine bins with similar backgrounds to form regions
  - Finer $t\bar{t}+\geq 1b$ categorisation using number of additional b-hadrons
  - e.g. Merge all bins with more than 60% $t\bar{t}+ \geq 2b$
    - In 2ℓ events with $\geq$4j: (60,60,60,60) and (60,60,60,70) bins

Event Categorisation
Resolved

- Using this method more freedom to have regions enriched in different backgrounds
  - Help control modelling of individual processes
- Due to shared final state, enriched $t\bar{t} + b\bar{b}$ regions are natural signal regions
- In total have 3 (5) signal regions and 4 (6) control regions in resolved 2ℓ (1ℓ)
- Can represent binning on 2D plot with $y = (jet_1, jet_2)$ and $x = (jet_3, jet_4)$
  - Convention uses b-tag discriminant bin instead of WP
  - 5=60%, 4=70% ... 1=100% (untagged)

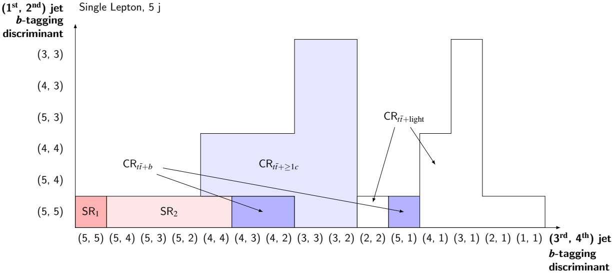
Event Categorisation
Graphical Representation - 2ℓ resolved

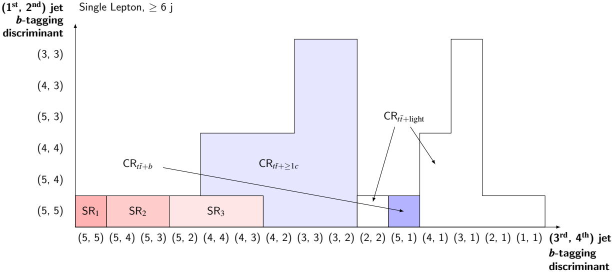




Event Categorisation

Graphical Representation - 1ℓ resolved

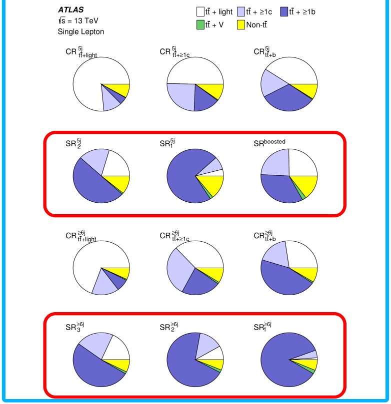




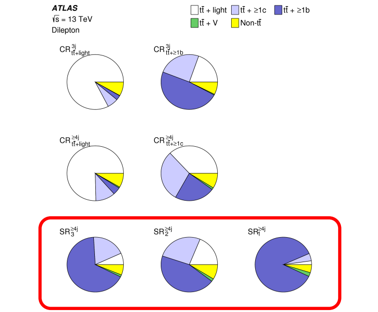
Event Selection and Categorisation
Summary

- 19 regions in total, of which 9 are signal regions
  - Boosted region is classed as a signal region

1ℓ channel



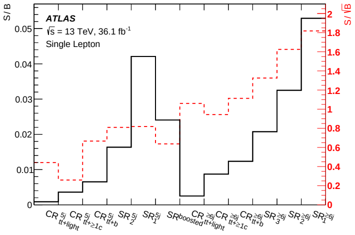
2ℓ channel



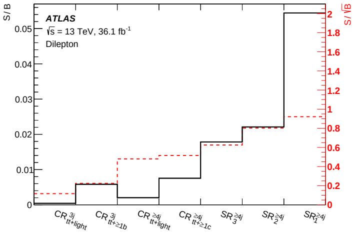
Event Selection and Categorisation
Summary

- 19 regions in total, of which 9 are signal regions
  - Boosted region is classed as a signal region

1ℓ channel



2ℓ channel



Reconstruction and Classification

Analysis Strategy
Overview

- Perform a binned profile likelihood fit simultaneously across all regions
  - In the signal regions want to enhance sensitivity to $t\bar{t}H$ events
  - Use the control regions to help handle the $t\bar{t} + \mathrm{jets}$ background
- Use an MVA discriminant in all signal regions
  - Two stage strategy employed - Reconstruction $\rightarrow$ Classification
- In control regions use either a single bin or scalar sum of jet p_T (H_T^{had})
  - H_T^{had} only used in $t\bar{t} + c$ CRs in 1ℓ regions
  - Required additional control over $t\bar{t}+\geq 1c$ modelling

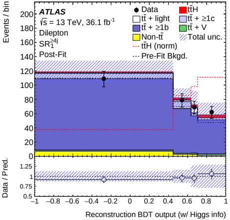
Two Stage MVA
1. Reconstruction

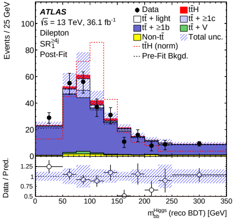
- Solve object combinatorics to reconstruct event hard scatter
  - Match jets/leptons to the partons in $t\bar{t}H\left(H \to b\bar{b}\right)/t\bar{t} + b\bar{b}$
- Three complimentary techniques used
  - i Reconstruction BDT
  - ii Likelihood discriminant
  - iii Matrix Element Method
- From each can construct variables with strong discrimination power
- Note: No explicit reconstruction in the Boosted region
  - Use the tagged Higgs candidate from event selection

Two Stage MVA

1. Reconstruction - Reconstruction BDT

- Train a BDT to assign jets to the partons in $t\bar{t}H\left(H \rightarrow b\bar{b}\right)$ hard scatter
  - Discriminates against combinatoric background
  - Use invariant masses and angular separations of jets/leptons
  - Evaluate on all events to choose jet matching
- Get a most $t\bar{t}H$-like jet-parton matching per event
  - Use BDT output score as a discriminant
  - Signal events more likely to have higher output score
  - Reconstruct object properties from jet assignment - Higgs mass
- Method used in all resolved signal regions





Two Stage MVA
1. Reconstruction

Likelihood Discriminant

- Only used in 1ℓ resolved signal regions
- Probability of an event to be signal or background ($t\bar{t} + b$ or $t\bar{t} + b\bar{b}$)
  - 1D PDFs constructed for inv. masses and angular distributions
  - Probabilities calculated as weighted product of all 1D PDFs
  - Weighted average of all possible combinations per event
- Final discriminant is a likelihood ratio of the sig and bkg probabilities

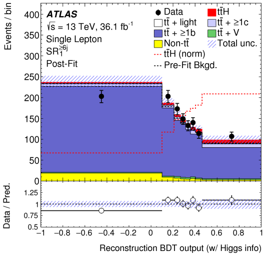
Matrix Element Method

- Only performed in the most signal enriched 1ℓ signal region
- Uses the four vector information of all jets and leptons, and the MET
- Signal and background hypothesis testing performed at parton level
- Final discriminant log of sig and bkg likelihood ratio

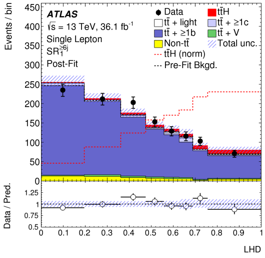
Two Stage MVA

1. Reconstruction

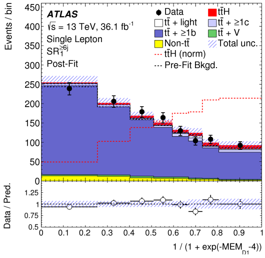
Reco BDT



LHD



MEM



Reco BDT	Exploits correlations within each combination
Reco BDT	Provides jet assignments based on $t\bar{t}H\,(H \rightarrow b\bar{b})$
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

Two Stage MVA
2. Classification

- Contsruct discriminants in each signal region to separate $t\bar{t}H$ from $t\bar{t}$
- Combine multiple variables with moderate separation power
  - Most powerful variables come from the reconstruction methods
- BDT optimised in each signal region
- Cross-validation performed to mitigate problems from overtraining
- Binning optimised for final significance in the fit

Two Stage MVA
2. Classification

Exanmple Input Variables

Reconstruction

- Reco discriminants
- Object properties (i.e. Higgs mass) from reco BDT
- Boosted Higgs/top properties

General Event

- nJets above p_T threshold
- Large Jet substructure

Event Shape

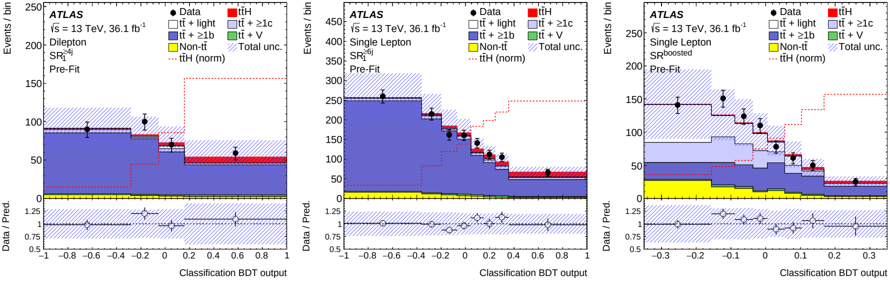
- From event E-$\vec{p}$ tensor (Aplanarity, Sphericity)
- Fox-Wolfram moments

Object Pairs

- Properties of a (b)-jet pair passing criteria
- $\Delta\eta_{bb}^{Max}$, $M_{jj}^{Min p_T}$

Two Stage MVA
Final Discriminant

- Most signal enriched region in each selection (2ℓ, 1ℓ resolved/boosted)
- Regions shown before performing the fit
  - Red is $t\bar{t}H$ assuming SM xsec

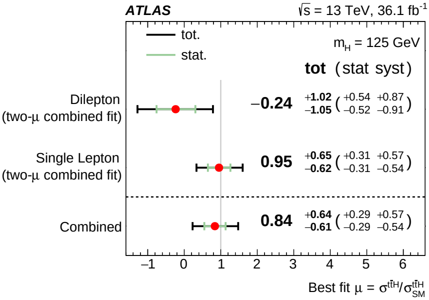


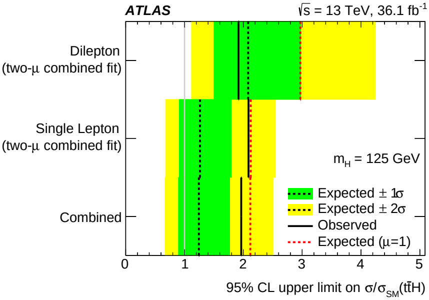
Results

Results
Overview

- Perform binned profile likelihood fit across all bins and regions simultaneously
  - No disctinction made between Signal and Control regions in the fit
- Parameter of interest is $t\bar{t}H$ signal strength $\mu_{t\bar{t}H}$
  - Defined as $\mu_{t\bar{t}H} = \sigma^{obs}_{t\bar{t}H}/\sigma^{SM}_{t\bar{t}H}$
- Large number of nuisance parameters covering modelling and detector systematic uncertainties
  - Free-float $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ normalisation factors

Results

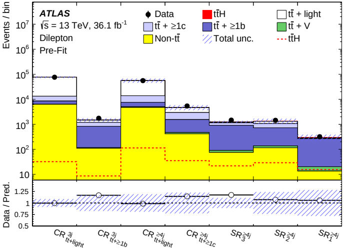


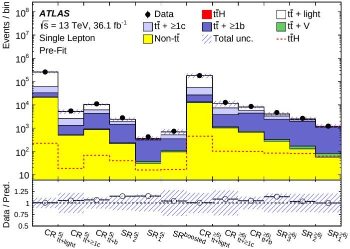


- Individual channel assessed using decorrelated signal strength
  - Still fit all regions simultaneously
- $t\bar{t}H$ has an observed (expected) significance of 1.6σ (1.8σ)
- Exclude $\mu > 2.0$ at 95% CL

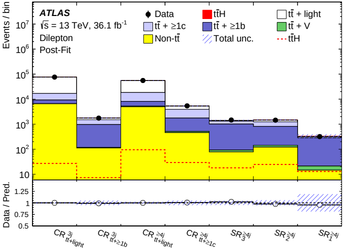
Results
Regions Summary

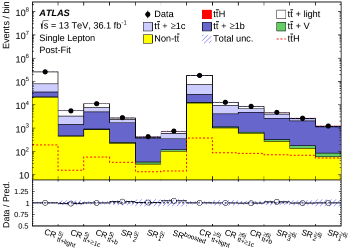
Pre-Fit





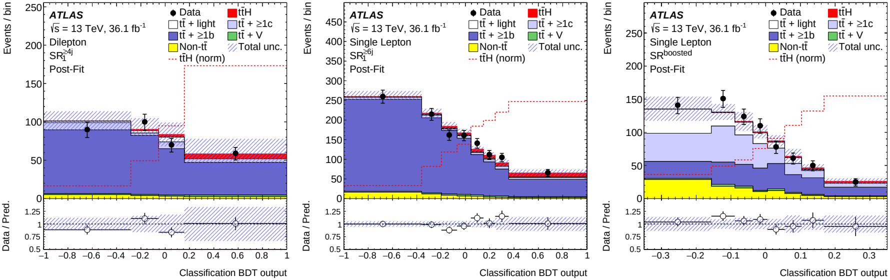
Post-Fit





Results
Final Discriminant

- Lets revisit the three regions shown before
  - $t\bar{t}H$ shown for extracted signal strength $\mu = 0.84^{+0.64}_{-0.61}$



Results

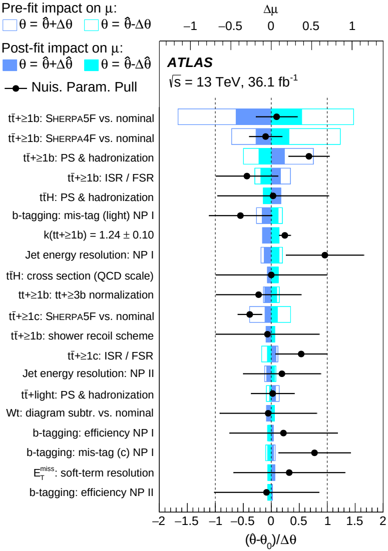
Impact of Systematic Uncertainties

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
$t\bar{t}+\text{light}$ 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 is currently systematically limited
- Largest uncertainties from $t\bar{t}+\mathrm{HF}$ modelling
- Also notable impact:
  - Bkg modelling stats.
  - Flavour tagging
  - Jet energy scale and resolution

Results

Impact of Systematic Uncertainties



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

$t\bar{t}H$ Combination

Results
$t\bar{t}H$ Combination

- $t\bar{t}H$ $(H \to b\bar{b})$ is just one of several searches in ATLAS for $t\bar{t}H$
- Other searches are optimised for other Higgs decay modes
  - $t\bar{t}H$ multileptons: $H \to WW^*/ZZ^*/\tau\tau$
  - $H \to \gamma\gamma$
  - $H \to ZZ^* \to 4\ell$
- All analyses have been performed using same 36.1 fb^{-1} dataset
- A combined fit over all channels has also been performed

ttH ML+comb | $H \to \gamma\gamma$ | $H \to ZZ^* \to 4\ell$

Results
$t\bar{t}H$ Combination

$t\bar{t}H$ multileptons

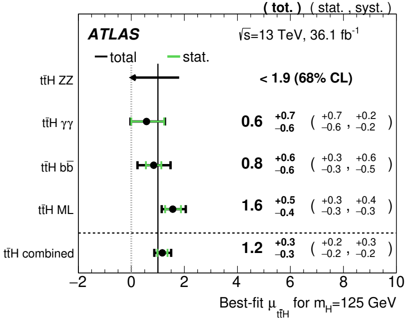
- 8 distinct signal regions targetting different decay modes
- Dominant backgrounds from $t\bar{t} + V$, $t\bar{t}$, fake and non-prompt leptons
  - Use a BDT to suppress non-prompt leptons
  - MVA discriminants used in five signal regions
- Wide range of S/B, from a few percent to >40%

$t\bar{t}H$ resonant searches

- $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ are $t\bar{t}H$ enriched regions in inclusive searches
  - Only use $t\bar{t}H$ enriched regions
- $H \to \gamma\gamma$: Cut based and BDT selections to separate signal from ggF and multijet backgrounds
- $H \to ZZ^* \to 4\ell$: Very pure cut and count, expected < 0.5 events

Results
$t\bar{t}H$ Combination

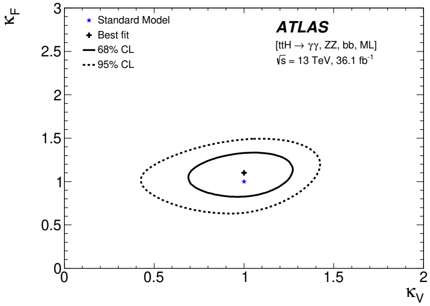
- Combining all $t\bar{t}H$ searches
- Non-$t\bar{t}H$ production modes set to SM values
- Almost all detector and signal and background uncertainties treated as correlated
- Best fit value of
  - $\mu_{t\bar{t}H} = 1.2 \pm 0.19\,(stat)\,^{+0.21}_{-0.23}\,(syst)$
  - $\sigma_{t\bar{t}H} = 590^{+160}_{-150}$ fb^{-1}
- Combined observed (expected) significance of 4.2 $\sigma(3.8\sigma)$



Channel	Best-fit μ Observed	Best-fit μ Expected	Significance Observed	Significance 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

Results
$t\bar{t}H$ Combination - Interpretations

- Wide range of Higgs couplings probed in the combination
- Using the kappa-parameterisation, scale the Higgs-couplings of particles (or groups of particles) by a factor κ_i
  - Look at coupling of Higgs boson to fermions κ_F and vector bosons κ_V
  - Couplings to gluons and photons comes from loop processes



- Consistent with Standard Model

Conclusion

Conclusion

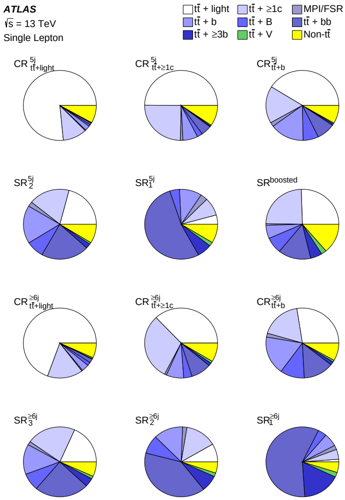
- Latest results from ATLAS for search for $t\bar{t}H$ $(H \to b\bar{b})$ including combination with other channels
- Very challenging analysis with heavy use of multivariate techniques to enhance sensitivity
  - Also makes full use of flavour tagging in region definitions to help control $t\bar{t} + \text{jets}$ backgrounds
- $t\bar{t}H$ $(H \to b\bar{b})$ is currently systematically dominated
  - Observed (expected) significance of 1.6σ (1.8σ)
  - Consistent with SM and B-Only hypotheses
- Evidence for $t\bar{t}H$ with 36.1 fb^{-1} ATLAS Run 2 data in combination

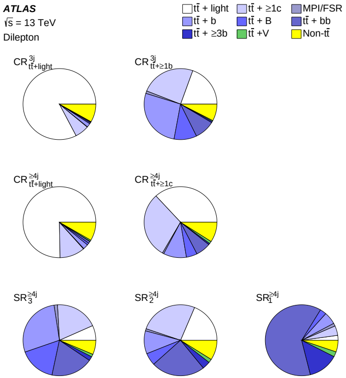
Outlook

- Additional data collected in 2017 could push combination above 5σ
- Potential $H \rightarrow b\bar{b}$ combination to aim for 5σ using 2017 data
  - Combining VH, $t\bar{t}H$ and VBF searches targetting $b\bar{b}$
  - Currently 3σ from $VH(b\bar{b})$ search
- However, further understanding of background modelling required for $t\bar{t}H \left(H \rightarrow b\bar{b}\right)$ search

Backup

Region Composition
Detailed





Region Definitions

Semileptonic Regions			
≥ 6 jets		5 jets	
Region name	Definition	Region name	Definition
$SR_1^{\geq 6j}$	$> 60\%\ t\bar{t} + \geq 2b$	SR_1^{5j}	$> 60\%\ t\bar{t} + \geq 2b$
$SR_2^{\geq 6j}$	$> 45\%\ t\bar{t} + \geq 2b$	SR_2^{5j}	$> 20\%\ t\bar{t} + \geq 2b$
$SR_3^{\geq 6j}$	$> 30\%\ t\bar{t} + \geq 2b$		
$CR_{t\bar{t}+b}^{\geq 6j}$	$> 30\%\ t\bar{t}\ + 1b$	$CR_{t\bar{t}+b}^{5j}$	$> 20\%\ t\bar{t}\ + 1b$
$CR_{t\bar{t}+\geq 1c}^{\geq 6j}$	$> 30\%\ t\bar{t} + \geq 1c$	$CR_{t\bar{t}+\geq 1c}^{5j}$	$> 20\%\ t\bar{t} + \geq 1c$
$CR_{t\bar{t}+\mathrm{light}}^{\geq 6j}$	Remaining events	$CR_{t\bar{t}+\mathrm{light}}^{5j}$	Remaining events
Dilepton Regions			
≥ 4 jets		3 jets	
Region name	Definition	Region name	Definition
$SR_1^{\geq 4j}$	$> 70\%\ t\bar{t} + \geq 2b$		
$SR_2^{\geq 4j}$	$> 1.5\%\ t\bar{t}H$		
$SR_3^{\geq 4j}$	$> 30\%\ t\bar{t}\ +\ 1b$		
$CR_{t\bar{t}+\geq 1c}^{\geq 4j}$	$> 25\%\ t\bar{t} + \geq 1c$	$CR_{t\bar{t}+\geq 1b}^{3j}$	$> 30\%\ t\bar{t} + \geq 1b$
$CR_{t\bar{t}+\mathrm{light}}^{\geq 4j}$	Remaining events	$CR_{t\bar{t}+\mathrm{light}}^{3j}$	Remaining events

Two Stage MVA

Final Discriminant

- Comparing the three signal regions directly
  - $t\bar{t}H$ shown post-fit for extracted signal strength $\mu = 0.84^{+0.64}_{-0.61}$

Pre-Fit

Post-Fit