Four Top Production at the LHC

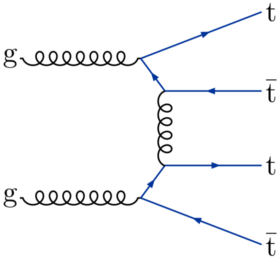
LHCP 2021

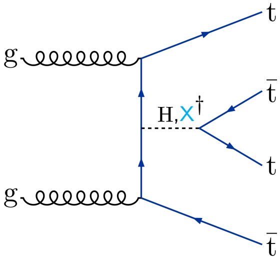
Albert Kong

On behalf of the ATLAS and CMS Collaborations

Monday 7th June, 2021

Introduction





- Top quark has large coupling to SM Higgs Boson and to many predicted particles in BSM theories
- $t\bar{t}t\bar{t}$ cross section can constrain top-Higgs Yukawa coupling, and expected to be sensitive to BSM processes

† for example X could be a pseudoscalar or scalar boson from Type II two-Higgs-doublet models

$t\bar{t}t\bar{t}$ Analyses at the LHC

$t\bar{t}t\bar{t}$ processes grouped into two channels based on final state

2 same-sign leptons, or 3+ leptons (SSML)

- low background
- low combined branching ratio (12-13%[1])
- ATLAS analysis: TOPQ-2018-05[2]
- CMS analysis: TOP-18-003[3]

1 lepton, or 2 opposite-sign leptons (1LOS)

- large irreducible background ($t\bar{t}$+jets)
- higher combined branching ratio (40-57%[1])
- ATLAS analysis: ATLAS-CONF-2021-013[4]
- CMS analysis uses a partial Run 2 dataset: TOP-17-019[5]

[1] differences from whether leptons from tau decays are counted

[2] https://link.springer.com/article/10.1140/epjc/s10052-020-08509-3

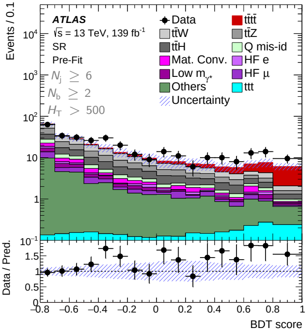
[3] http://dx.doi.org/10.1140/epjc/s10052-019-7593-7

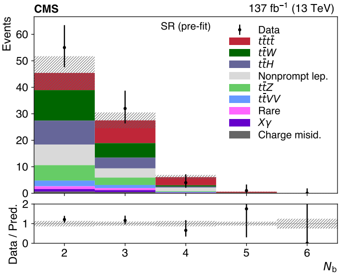
[4] https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2021-013/

[5] http://dx.doi.org/10.1007/JHEP11(2019)082

SSML: Analysis Strategy

- Signal/background separation achieved using a BDT discriminant
- ATLAS BDT used 12 inputs, most discriminating variables:
  - sum of b-tagging scores[1]
  - minimum ΔR between any pair of leptons
- CMS BDT used 19 inputs, most discriminating variables:
  - number of jets N_{jets}
  - number of b-tagged jets N_b
  - number of leptons N_ℓ



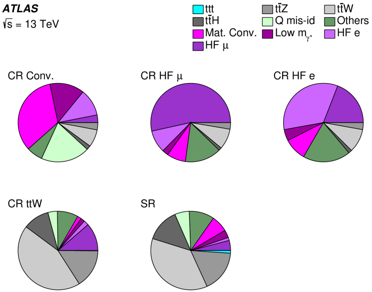


[1] integer based on most stringent b-tagging working point the jet passes

SSML: Background Estimation

TOPQ-2018-05, TOP-18-003

- Dominant background from production of $t\bar{t}X$, where $X = bb, jj, W, Z, H$
- Another significant background from events with a charge-misidentified lepton, or a 'non-prompt' lepton



ATLAS

- Control regions for conversions, heavy flavour (HF) e/μ, and $t\bar{t}W$
- Template method for estimating non-prompt background, data-driven method for estimating charge mis-id

CMS

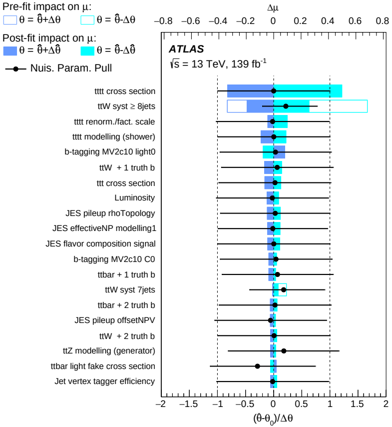
- Control region for $t\bar{t}Z$
- 'Tight-to-loose' ratio method for estimating non-prompt background, data-driven method for charge mis-id

SSML: Systematics

TOPQ-2018-05, TOP-18-003

- Approximately equal contributions from systematic and statistical uncertainties
- Main systematic uncertainties come from $t\bar{t}X$ modelling, then jet and signal[1] modelling

ATLAS



CMS

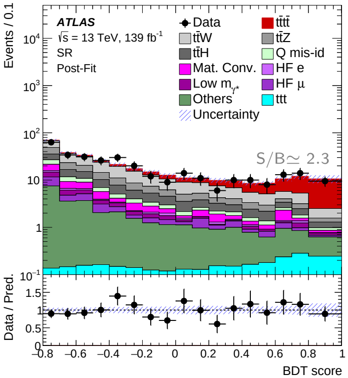
Source	Uncertainty (%)	Impact on $\sigma(t\bar{t}t\bar{t})$ (%)
Integrated luminosity	2.3–2.5	2
Pileup	0–5	1
Trigger efficiency	2–7	2
Lepton selection	2–10	2
Jet energy scale	1–15	9
Jet energy resolution	1–10	6
b tagging	1–15	6
Size of simulated sample	1–25	<1
Scale and PDF variations †	10–15	2
ISR/FSR (signal) †	5–15	2
$t\bar{t}H$ (normalization) †	25	5
Rare, $X\gamma$, $t\bar{t}VV$ (norm.) †	11–20	<1
$t\bar{t}Z$, $t\bar{t}W$ (norm.) †	40	3–4
Charge misidentification †	20	<1
Nonprompt leptons †	30–60	3
$N_{\text{jets}}^{\text{ISR/FSR}}$	1–30	2
$\sigma(t\bar{t}b\bar{b})/\sigma(t\bar{t}jj)$ †	35	11

[1] signal modelling uncertainty only impacts the extracted signal strength μ

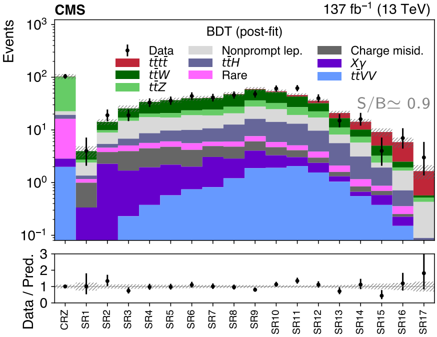
SSML: Results

TOPQ-2018-05, TOP-18-003

ATLAS



CMS



4.3σ (2.4σ) observed (expected) significance

$$\sigma_{t\bar{t}t\bar{t}} = 24^{+7}_{-6} \text{ fb}$$

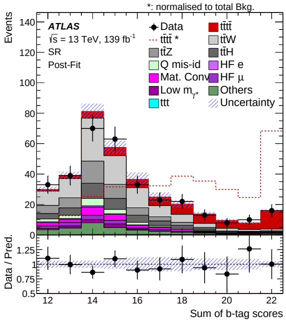
2.6σ (2.7σ) observed (expected) significance

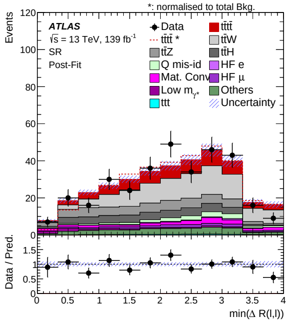
$$\sigma_{t\bar{t}t\bar{t}} = 12.6^{+5.8}_{-5.2} \text{ fb}$$

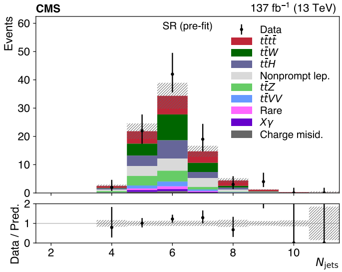
Predicted SM cross section $\sigma(pp \to t\bar{t}t\bar{t}) = 12.0^{+2.2}_{-2.5}$ fb (NLO)

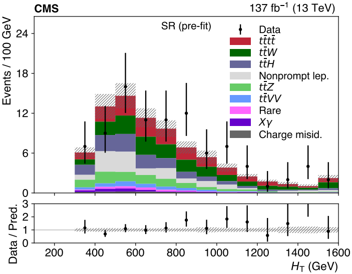
SSML: Kinematic Distributions

TOPQ-2018-05, TOP-18-003



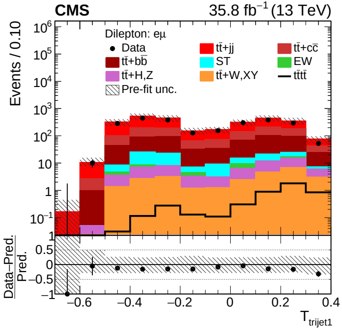


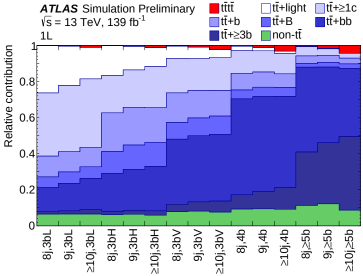




1LOS: Analysis Strategy

TOP-17-019, ATLAS-CONF-2021-013

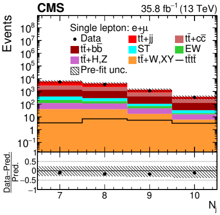


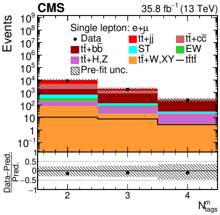


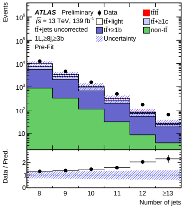
- BDT discriminant used to obtain signal-background separation in signal regions
- ATLAS used 14 input variables, most discriminating were sum of the *b*-tagging scores, N_{jets}, and the minimum ΔR between pairs of *b*-tagged jets
- CMS used a different BDT for each sub-channel, as well as a BDT to distinguish trijets originating from hadronic top quark decays from other sources

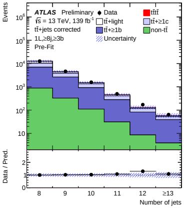
1LOS: Background Estimation

- Background dominated by $t\bar{t}$+jets, with smaller contributions from $t\bar{t}X$ ($< 8\%$), and diboson and rare processes ($< 1\%$)
- Background from mis-reconstructed or non-prompt leptons is negligible in this channel
- Non-$t\bar{t}$ contributions just estimated from simulations
- $t\bar{t}$+jets events are rescaled and reweighted separately in each particle jet flavour category (b, c and light), and constrained by a profile likelihood fit









1LOS: Systematics

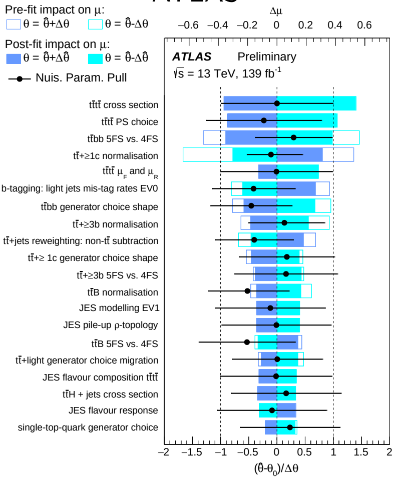
TOP-17-019, ATLAS-CONF-2021-013

CMS

Systematic uncertainty	Normalization	Shape
Integrated luminosity	X	
Pileup modeling	X	X
Lepton reconstruction and identification	X	
Jet energy corrections	X	X
b tagging	X	X
Ren. and fact. scales	X	X
PS scales	X	
ME-PS matching	X	
UE	X	
Jet multiplicity correction	X	
Parton distribution functions	X	X
Top quark p_T reweighting		X
Heavy-flavor reweighting	X	X
Rare process	X	

- Systematically limited
- $t\bar{t}$ modelling uncertainties account for almost all systematic uncertainty
- Signal modelling and jet/b-tagging uncertainties account for the remaining fraction

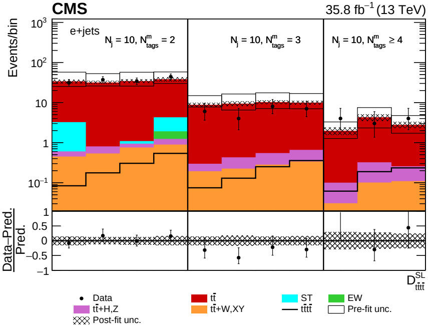
ATLAS



1LOS: Results

TOP-17-019, ATLAS-CONF-2021-013

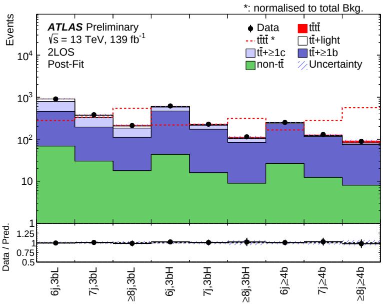
CMS



0.0σ (0.4σ) observed (expected) significance

$\sigma_{t\bar{t}t\bar{t}} = 0^{+20}$ fb

ATLAS



1.9σ (1.0σ) observed (expected) significance

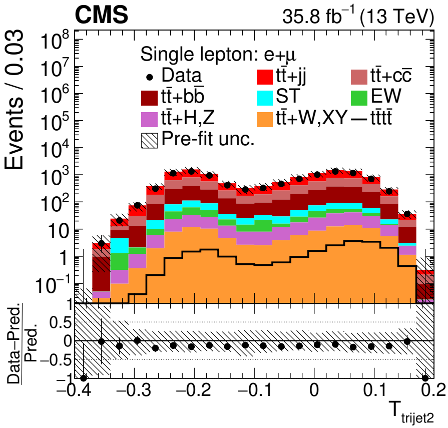
$\sigma_{t\bar{t}t\bar{t}} = 26^{+17}_{-15}$ fb

Predicted SM cross section $\sigma(pp \to t\bar{t}t\bar{t}) = 12.0^{+2.2}_{-2.5}$ fb (NLO)

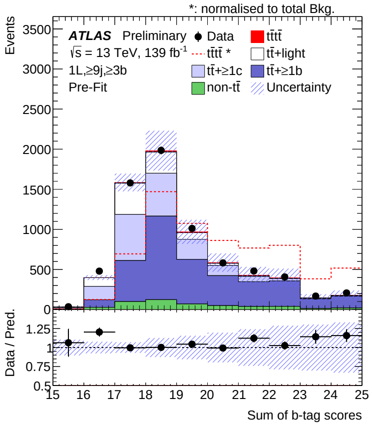
1LOS: Kinematic Distributions

TOP-17-019, ATLAS-CONF-2021-013

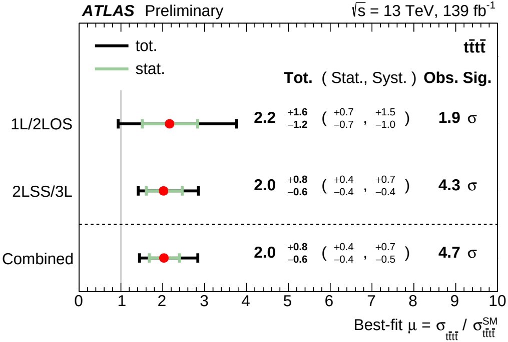
CMS



ATLAS



ATLAS Combination Results



4.7σ (2.6σ) observed (expected) significance

$$\sigma_{t\bar{t}t\bar{t}} = 24^{+7}_{-6}\ \text{fb}$$

Improved observed significance over the SSML-only result from ATLAS

Conclusion and Future Directions

- ATLAS SSML analysis provides first evidence for SM $t\bar{t}t\bar{t}$ production
- CMS SSML and ATLAS 1LOS analyses provide supporting results
- ATLAS SSML+1LOS combination gives improved significance, future runs might reach 5σ observation!
- CMS released interpretations under various BSM frameworks together with results, ATLAS interpretations also on the way![1]

[1] see the dedicated talks for ATLAS and CMS EFT results for more details

Backup Slides

SSML: Tight-to-loose Ratio Method

- Relax tight identification (for electrons) and isolation (electrons+muons) requirements to define control region enriched in non-prompt leptons
- Calculate efficiency $\varepsilon_{\mathsf{TL}}$ of non-prompt leptons that satisfy both loose and tight selections
- Use to estimate non-prompt content in SRs by counting loose-not-tight leptons n_{LNT} in an adjacent 'application region' and reweighting non-prompt events by a factor of $[\varepsilon_{\mathsf{TL}}/(1-\varepsilon_{\mathsf{TL}})]^{n_{\mathsf{LNT}}}$

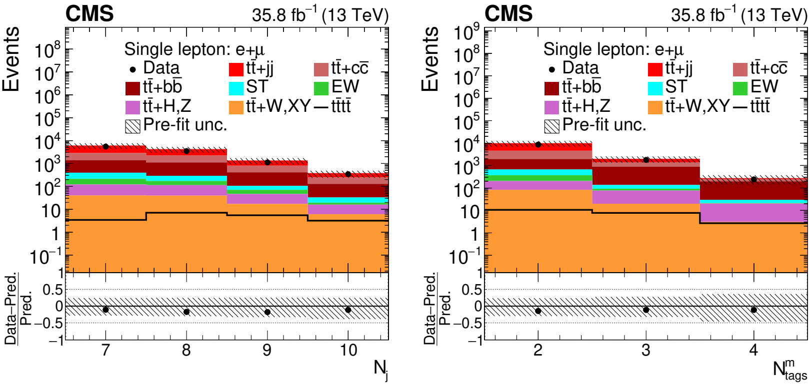
SSML: Data-driven Charge Mis-id Estimation

- Measure charge misassignment rate in data sample of $Z \to ee$ (requiring $|m_{ee} - m_Z| < 10$ GeV)
- Parametrise charge misassignment rate as a function of electron p_{T} and $|\eta|$ (negligible contribution from $\mu\mu$ events)
- Estimate contribution in other data regions by weighting opposite-sign lepton events by misassignment rate (0.002% to 4%)

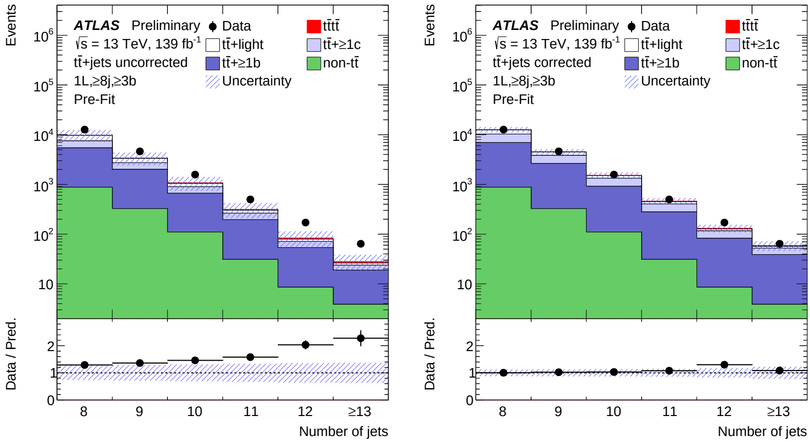
SSML: Extracted Signal Strength

- An ATLAS SUSY analysis ATLAS-CONF-2021-007 conducted in the same phase space also had sensitivity to SM $t\bar{t}t\bar{t}$ production
- Using a completely different background estimation approach (reduced reliance on MC, no BDT), the same central value for $\mu_{t\bar{t}t\bar{t}}$ was obtained

1LOS: $t\bar{t}$ Background Estimation



1LOS: $t\bar{t}$ Background Estimation



1LOS: $t\bar{t}$ Rescaling and Reweighting

- Rescaling corrects yields in the $t\bar{t}+$ light, $t\bar{t}+\geq 1c$ and $t\bar{t}+\geq 1b$ categories separately, using factors from a profile likelihood fit to data
- Reweighting corrects the distributions for N_{jets}, number of large-R jets $N_{\text{LR-jets}}$, $H_{\text{T}}^{\text{all}}$, and average ΔR between any two jets $\Delta R_{\text{avg}}^{\text{jj}}$
- Reweighting performed in three steps, separately for 1L and 2LOS events, such that overall MC prediction matches data in 2b regions
  - First reweight N_{jets} and $N_{\text{LR-jets}}$ as a 2D distribution (one reweighting factor per $(N_{\text{jets}}, N_{\text{LR-jets}})$ bin)
  - Secondly correct the $H_{\text{T}}^{\text{all}}$ spectrum by reweighting the reduced $H_{\text{T}}^{\text{all}}$ variable, defined as $H_{\text{T}}^{\text{all,red.}} = H_{\text{T}}^{\text{all}} - (N_{\text{jets}} - N_{\text{min}}) \times 90$ [GeV], where $N_{\text{min}} = 7$ (5) in the 1L (2LOS) channel ($90 \simeq p_{\text{T}}$ of additional jets)
  - Thirdly reweight the $\Delta R_{\text{avg}}^{\text{jj}}$ distribution, also in $(N_{\text{jets}}, N_{\text{LR-jets}})$ bins