$t\bar{t}X$ results from ATLAS and CMS

Large Hadron Collider Physics (LHCP) 2021 conference
Joint session on Higgs boson and top quark physics

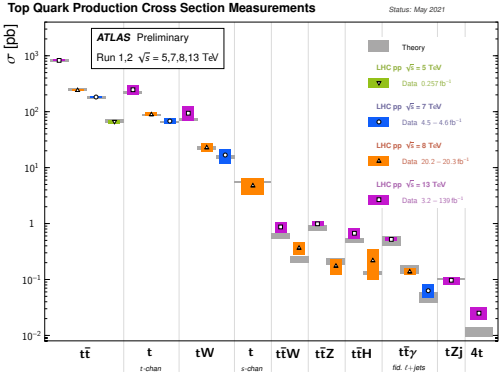
Peter Falke (University of Bonn)
on behalf of the ATLAS and CMS collaborations

09/06/21

Introduction to $t\bar{t}X$ processes

- **Associated production of top-quark pairs**: $t\bar{t}X$ with $X \in \{W, Z, \gamma, H, t\bar{t}, q\bar{q}\}$
  - Allow to probe coupling of top quarks to EW bosons and rare SM processes $\rightarrow t\bar{t}Z/t\bar{t}\gamma$ and $t\bar{t}t\bar{t}$ processes covered in top quark physics session on Monday
  - $t\bar{t}H$ production covered in next talk by Djamel Eddine Boumediene

- This talk: **$t\bar{t}$W**, **$t\bar{t}b\bar{b}$** and **$t\bar{t}c\bar{c}$** production
  - Important backgrounds e.g. for $t\bar{t}H$ analyses
  - $t\bar{t}W$: rare $t\bar{t}X$ process with difficult modelling
  - $t\bar{t} + \mathrm{HF}$: high uncertainties in theoretical calc.
  - **Need precise knowledge of their properties**

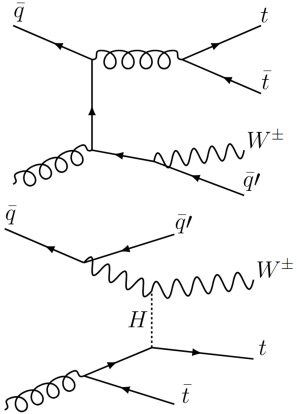


$t\bar{t}W$ production

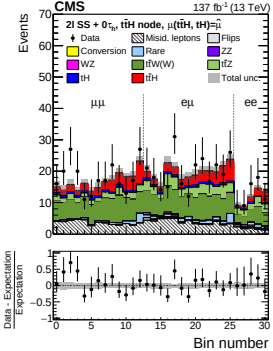
ATLAS paper ($36.1\,\mathrm{fb}^{-1}$): Phys. Rev. D 99 (2019) 072009
CMS paper ($35.9\,\mathrm{fb}^{-1}$): JHEP 08 (2018) 011

Introduction to $t\bar{t}W$ production

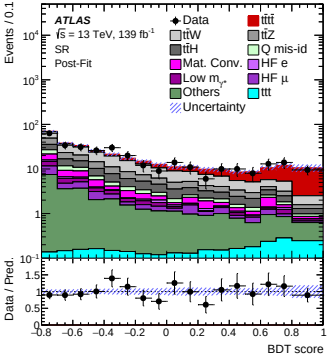
- $t\bar{t}W$ production at LO only from $q\bar{q}'$ initial states
  - **Large NLO corrections** due to additional gq initial states $\rightarrow$ difficult modelling
- Important background for analyses with **multilepton (ML) final states**
  - Examples: **$t\bar{t}H$ and $t\bar{t}t\bar{t}$ production**, searches for new physics, ...



Figures from arXiv:1907.04343



$2\ell SS$ channel in CMS $t\bar{t}H$ ML (Eur. Phys. J. C 81 (2021) 378)



ATLAS $t\bar{t}t\bar{t}$ ML analysis (Eur. Phys. J. C 80 (2020) 1085)

Selection of $t\bar{t}W$ production

- **Two main channels** for $t\bar{t}W$ selection ($\ell \in \{e, \mu\}$ incl. τ_{lep} decays)
  - 2ℓSS with $t\bar{t} \to (\ell^{\pm}\nu b)(q\bar{q}b)$ and $W^{\pm} \to \ell^{\pm}\nu$
  - 3ℓ with $t\bar{t} \to (\ell^{\pm}\nu b)(\ell^{\mp}\nu b)$ and $W^{\pm} \to \ell^{\pm}\nu$
  - ATLAS analysis used both channels, CMS focussed on 2ℓSS channel

CMS 2ℓSS selection

Requirement	$2j$(p,m)	$3j\,1b$(p,m)	$3j > 1b$(p,m)	$> 3j\,1b$(p,m)	$> 3j > 1b$(p,m)
n_{jets}	2	3		> 3	
n_{bjets}	–	1	> 1	1	> 1
p_{T} (lead. ℓ)	$> 25\,\text{GeV}$ (40 GeV in ee case)				
p_{T} (sublead. ℓ)	$> 25\,\text{GeV}$				
$m_{\ell\ell}$	$> 12\,\text{GeV}$ and $\lvert m_{\ell\ell} - m_Z \rvert > 15\,\text{GeV}$				
$p_{\text{T}}^{\text{miss}}$	$> 30\,\text{GeV}$				

Requirement	2ℓ-SS(p,m)-$1b$	$2e$-SS(p,m)-$2b$	$e\mu$-SS(p,m)-$2b$	2μ-SS(p,m)-$2b$
$n_{b\text{-tags}}$	$=1$	≥ 2	≥ 2	≥ 2
$E_{\text{T}}^{\text{miss}}$	> 40 GeV	> 40 GeV	> 40 GeV	> 20 GeV
H_{T}	> 240 GeV			
p_{T} (leading lepton)	> 27 GeV			
p_{T} (subleading lepton)	> 27 GeV			
n_{jets}	≥ 4	≥ 4	≥ 4	≥ 2
Z veto	$\lvert m_{\ell\ell} - m_Z \rvert > 10\,\text{GeV}$ in the $2e$ and 2μ regions			

ATLAS 2ℓSS selection

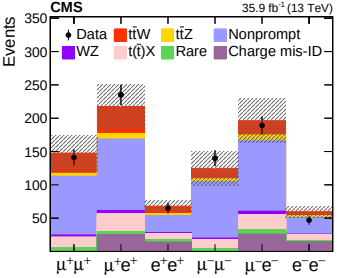
Variable	3ℓp-noZ-2b2j	3ℓm-noZ-2b2j	3ℓp-noZ-1b2j	3ℓm-noZ-1b2j
All leptons	$p_{\text{T}} > 27\,\text{GeV}$			
Z veto (OSSF pair)	$\lvert m_{\ell\ell} - m_Z \rvert > 10\,\text{GeV}$			
n_{jets}	2 or 3			
H_{T}	–		> 240 GeV	
Sum of lepton charges	+1	−1	+1	−1
$n_{b\text{-tags}}$	≥ 2	≥ 2	1	1

ATLAS 3ℓ selection

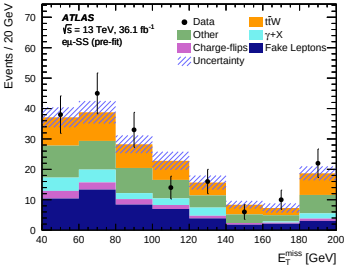
- Both experiments **split channels further** as a function of n_{jets}, n_{bjets} and lepton charge and **veto additional leptons** with looser ID and isolation requirements

Background processes and estimation

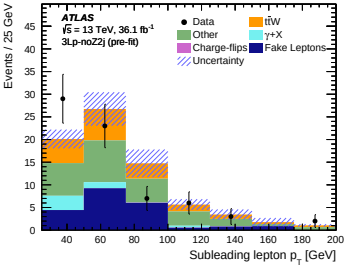
- Main **background processes**
  - $2\ell SS$: fake and non-prompt leptons, charge flip/mis-ID $e^{\pm}$, prompt processes (e.g. $t\bar{t}H$, WZ)
  - 3ℓ: fake and non-prompt leptons, prompt processes (e.g. $t\bar{t}H$ and $t\bar{t}Z$ in "Other" category)
- Estimation of **non-prompt and fake leptons**
  - Main sources: semi-leptonic b-hadron decays, photon conversions, misidentified hadrons, ...
  - Data-driven technique for both experiments with CRs for selection efficiency estimation



CMS $2\ell SS$ pre-fit yields split by lepton flavour/charge



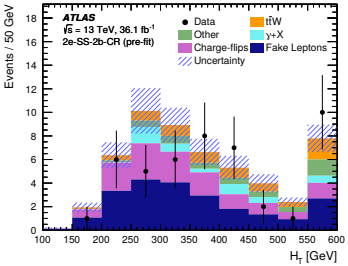
ATLAS $e^{\pm}\mu^{\pm}$ pre-fit agreement



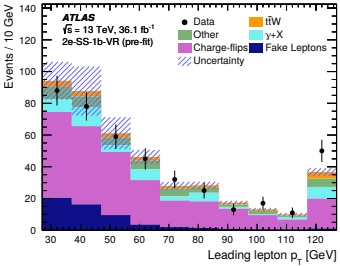
ATLAS 3ℓ pre-fit agreement

Background estimation and event selection

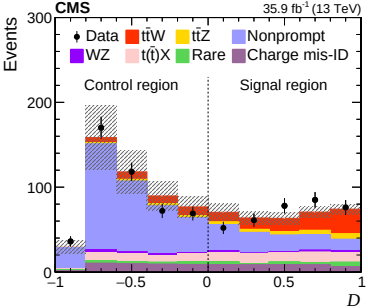
- Estimation of **charge flip/mis-ID electrons**
  - ATLAS applies data-driven technique using CR with $m_{\ell\ell}$ around m_Z
  - CMS estimates charge flip/mis-ID probabilities from MC and applies these to $2\ell\mathrm{OS}$ data
- **BDT-based event selection** in CMS analysis with BDT discriminant D
  - Input variables given e.g. as n_{jets}, n_{bjets}, $H_{\mathrm{T,jets}}$, $p_{\mathrm{T}}^{\mathrm{miss}}$, ... (11 in total)
  - Select only events with $D > 0$, further split into $0 < D < 0.6$ and $D > 0.6$ categories to increase sensitivity



ATLAS fake and non-prompt ℓ CR



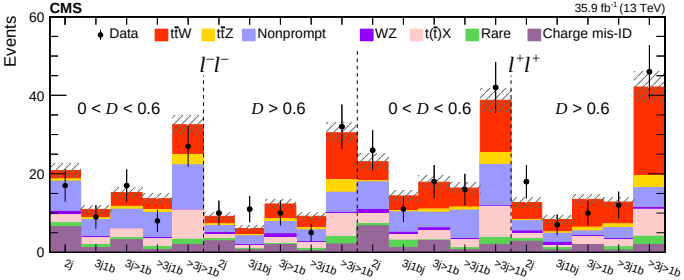
ATLAS charge flip/mis-ID $e^{\pm}$ CR



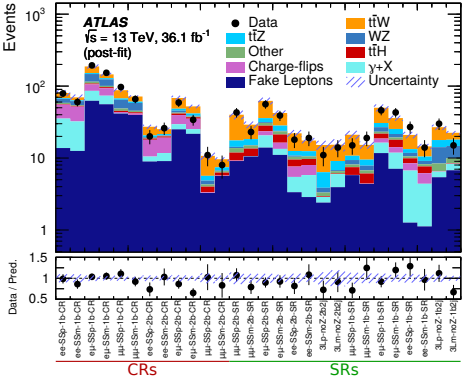
CMS event-selection BDT

Fit strategy and uncertainties

- Both experiments use **profile-likelihood fit** → figures show final data/MC agreement
- Dominant **systematic uncertainties**
  - ATLAS: $t\bar{t}W$ modelling, non-prompt/fake and charge flip/mis-ID estimation (incl. CR data stat.)
  - CMS: Trigger and b-tagging efficiencies, prompt and non-prompt/fake background estimation



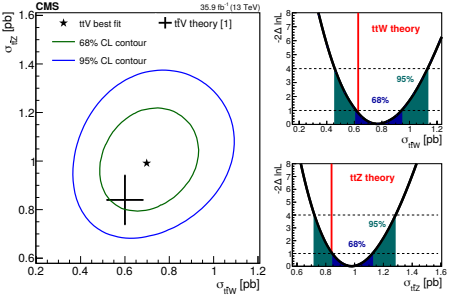
Post-fit agreement in CMS after the BDT selection



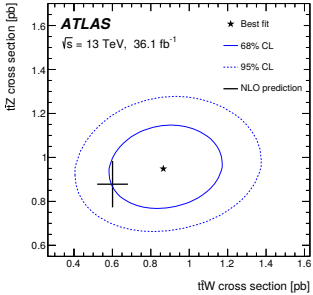
ATLAS post-fit agreement in CRs/SRs

Results on $t\bar{t}W$ production

- Results from **cross-section measurement** (in both cases $t\bar{t}V$ measurement)
  - ATLAS: $\sigma_{t\bar{t}W} = 0.87 \pm 0.13(\mathrm{stat}) \pm 0.14(\mathrm{syst})\,\mathrm{pb}$
  - CMS: $\sigma_{t\bar{t}W} = 0.77 \;^{+0.12}_{-0.11}(\mathrm{stat}) \;^{+0.13}_{-0.12}(\mathrm{syst})\,\mathrm{pb}$
  - Theory prediction: $\sigma_{t\bar{t}W} = 0.59^{+0.15}_{-0.10}(\mathrm{scale}) \pm 0.01(\mathrm{PDF})\,\mathrm{pb}$ (Eur. Phys. J. C 80 (2020) 428)



CMS $t\bar{t}V$ cross-section measurement

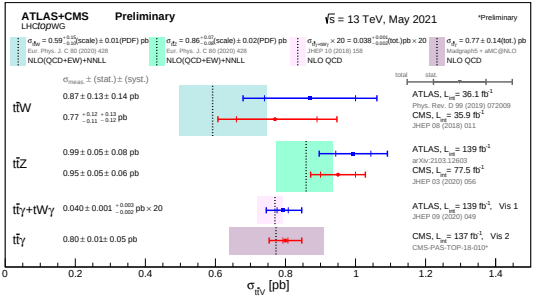


ATLAS $t\bar{t}V$ cross-section measurement

- CMS paper additionally includes **EFT interpretation** that includes $t\bar{t}W$ SRs
  - $t\bar{t}V$ analysis allows to constrain 8 Wilson coefficients, no operators only affect $t\bar{t}W$ alone

Results on $t\bar{t}W$ production

- Results from **cross-section measurement** (in both cases $t\bar{t}V$ measurement)
  - ATLAS: $\sigma_{t\bar{t}W} = 0.87 \pm 0.13(\mathrm{stat}) \pm 0.14(\mathrm{syst})\,\mathrm{pb}$
  - CMS: $\sigma_{t\bar{t}W} = 0.77 \quad ^{+0.12}_{-0.11}(\mathrm{stat}) \quad ^{+0.13}_{-0.12}(\mathrm{syst})\,\mathrm{pb}$
  - Theory prediction: $\sigma_{t\bar{t}W} = 0.59^{+0.15}_{-0.10}(\mathrm{scale}) \pm 0.01(\mathrm{PDF})\,\mathrm{pb}$ (Eur. Phys. J. C 80 (2020) 428)

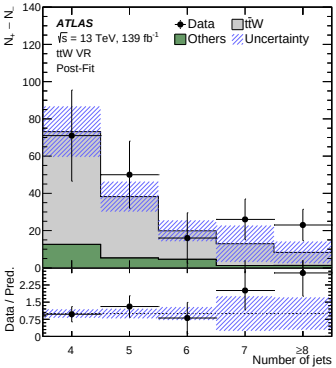


ATLAS + CMS $t\bar{t}V$ summary figure

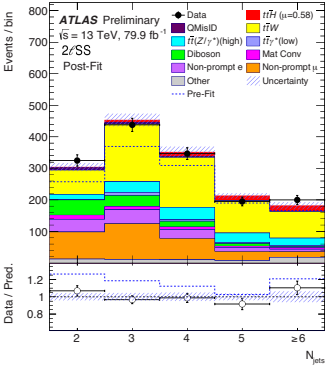
- CMS paper additionally includes **EFT interpretation** that includes $t\bar{t}W$ SRs
  - $t\bar{t}V$ analysis allows to constrain 8 Wilson coefficients, no operators only affect $t\bar{t}W$ alone

$t\bar{t}W$ production in recent $t\bar{t}H$ multilep. and $t\bar{t}t\bar{t}$ analyses

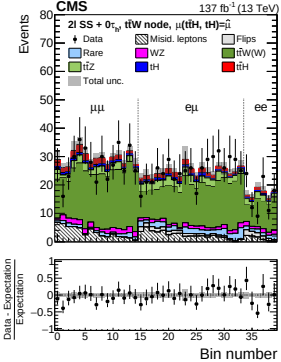
- Several analyses found **increased $t\bar{t}W$ yields w.r.t. theory prediction**
  - Typical normalisation factors in the range of $1.3 - 1.7 \times \sigma_{\text{theory}}$ (post-fit case shown below)
    $\rightarrow$ further mismodelling covered by additional systematics in some analyses
  - Clearly need to **improve theoretical and experimental understanding**



ATLAS $t\bar{t}t\bar{t}$ ML analysis
(Eur. Phys. J. C 80 (2020) 1085)



ATLAS $t\bar{t}H$ ML analysis
(ATLAS-CONF-2019-045)



CMS $t\bar{t}H$ ML analysis
(Eur. Phys. J. C 81 (2021) 378)

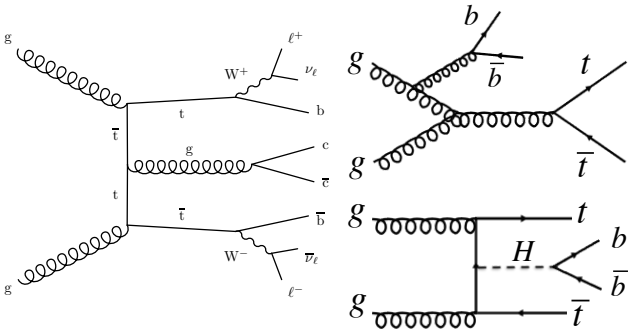
$t\bar{t}b\bar{b}$ and $t\bar{t}c\bar{c}$ production

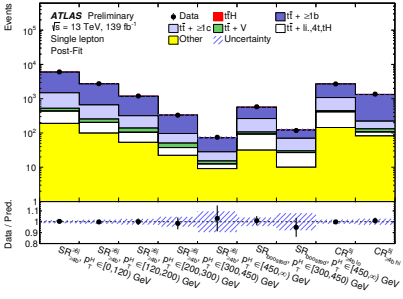
CMS $t\bar{t}c\bar{c}$ paper (dilepton $t\bar{t}$ decays, 41.5 fb^{-1}): submitted to Phys. Lett. B

ATLAS $t\bar{t}b\bar{b}$ paper ($e\mu$ and ℓ+jets $t\bar{t}$, 36.1 fb^{-1}): JHEP 04 (2019) 046
CMS $t\bar{t}b\bar{b}$ and $t\bar{t}j\bar{j}$ paper (dilepton and ℓ+jets $t\bar{t}$, 35.9 fb^{-1}): JHEP 07 (2020) 125
CMS $t\bar{t}b\bar{b}$ paper (all-hadronic $t\bar{t}$, 35.9 fb^{-1}): Phys. Lett. B 803 (2020) 135285

Introduction to $t\bar{t}b\bar{b}$ and $t\bar{t}c\bar{c}$ production

- $t\bar{t}b\bar{b}/t\bar{t}c\bar{c}$ production are **dominant backgrounds e.g. in $t\bar{t}H(b\bar{b})$ measurements**
  - Non-resonant background under peaking Higgs boson contribution
  - **Systematic uncertainties** in $t\bar{t}b\bar{b}$ prediction dominate inclusive $\mu_{t\bar{t}H}$ measurement
- **Theoretical modelling** of $t\bar{t}b\bar{b}/t\bar{t}c\bar{c}$ production is challenging
  - Uncertainties e.g. due to μ_F/μ_R scale choice

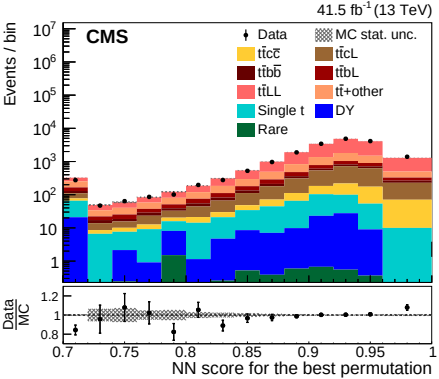




Single-ℓ CRs/SRs of ATLAS $t\bar{t}H(b\bar{b})$ analysis (ATLAS-CONF-2020-058)

$t\bar{t}c\bar{c}$ selection strategy and parton-matching

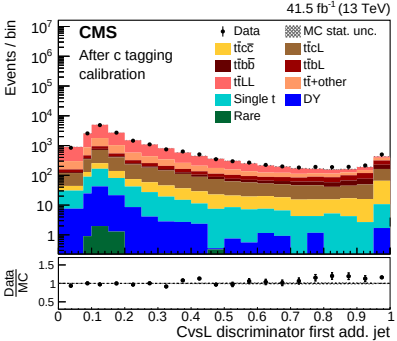
- **CMS $t\bar{t}c\bar{c}$ analysis** presented on the following slides
- **Event selection** with dileptonic $t\bar{t}$ decays
  - Two leptons (e or μ, incl. τ_{lep} decays) with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and $|\eta| < 2.4$
  - $n_{\mathrm{jets}} \geq 4$ with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$, $|\eta| < 2.4$ and $\Delta R(\ell, \mathrm{jet}) > 0.5$
  - $m_{\ell\ell} > 12\,\mathrm{GeV}$ for all events; $p_{\mathrm{T}}^{\mathrm{miss}} > 30\,\mathrm{GeV}$ and $|m_{\ell\ell} - m_Z| > 15\,\mathrm{GeV}$ in ee and $\mu\mu$ cases
- **Matching of jets and partons**
  - Two (b-)jets from top decays need to be identified
    $\rightarrow$ separate treatment of jets from add. radiation
  - NN trained for jet-parton assignment
    $\rightarrow$ use jet kinematics, b/c-tagging discriminators, ...
  - Loop over all permutations of jet-parton assignments
    $\rightarrow$ score indep. of assignment to top or anti-top quark
    $\rightarrow$ jets assigned to top quarks must both be b-tagged
  - Additional pair of **$c\bar{c}$ ($b\bar{b}$) jets identified in 50 (30)% of the cases**



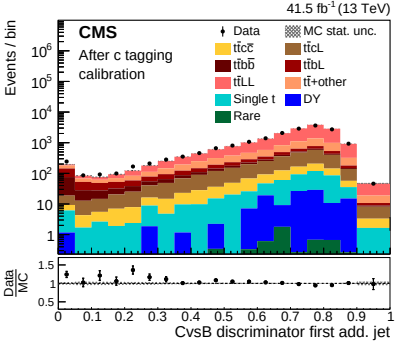
Parton-matching NN score

Identification of charm-jets

- **Tagging of single jets** performed with **multi-class DeepCSV algorithm**
  - Output classes: single b-hadron, two b-hadrons, one or more c-hadrons, light jets ($udsg$)
  - **Analysis uses two combinations**: $\mathrm{CvsL} = \frac{P(c)}{P(c)+P(udsg)}$ and $\mathrm{CvsB} = \frac{P(c)}{P(c)+P(b)+P(bb)}$
- **Dedicated calibration technique** for CvsL and CvsB scores
  - CRs: $t\bar{t}$ (for b-jets), $W + c$ (c-jets) and $\mathrm{DY} + \mathrm{jets}$ (light jets)



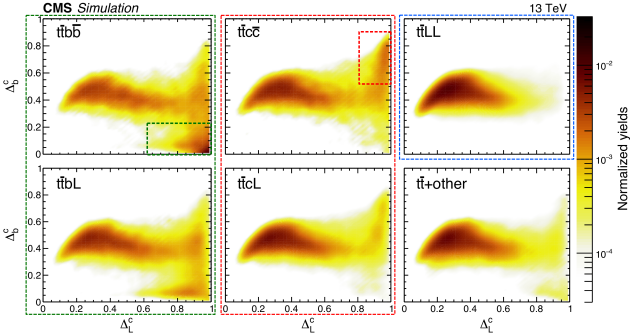
C vs. L discrimination for single jet



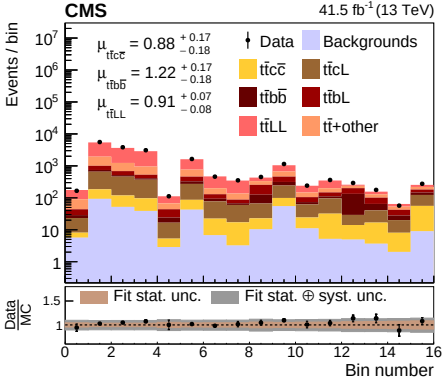
C vs. B discrimination for single jet

Fit strategy and MVA setup

- **Event-level separation** of $t\bar{t}b\bar{b}$, $t\bar{t}bL$, $t\bar{t}c\bar{c}$, $t\bar{t}cL$, $t\bar{t}LL$ and $t\bar{t}+\text{other}$ classes
  - Multi-class NN using CvsL/CvsB of leading two jets, $\Delta R(\text{jet},\text{jet})$ and permutation score
  - **Two discriminators**: $\Delta_b^c = \frac{P(t\bar{t}c\bar{c})}{P(t\bar{t}c\bar{c})+P(t\bar{t}b\bar{b})}$ and $\Delta_L^c = \frac{P(t\bar{t}c\bar{c})}{P(t\bar{t}c\bar{c})+P(t\bar{t}LL)}$
- Final fit done in **unrolled Δ_L^c vs. Δ_b^c distribution**
  - $\Delta_L^c \otimes \Delta_b^c$: $[0, 0.45, 0.6, 0.9, 1.0] \otimes [0, 0.3, 0.45, 0.5, 1.0]$



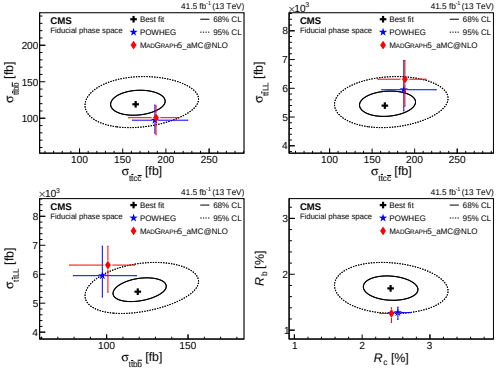
Δ_L^c vs. Δ_b^c 2D distributions



One-dimensional representation of binned Δ_L^c vs. Δ_b^c distribution

Results of cross-section and R_b/R_c measurements

- Measurement of $\mathbf{t\bar{t}b\bar{b}}$ (incl. $t\bar{t}bL$), $\mathbf{t\bar{t}c\bar{c}}$ (incl. $t\bar{t}cL$) and $\mathbf{t\bar{t}LL}$ production as well as $\mathbf{R_b} = t\bar{t}b\bar{b}/t\bar{t}jj$ and $\mathbf{R_c} = t\bar{t}c\bar{c}/t\bar{t}jj$ (separate fit) in **fiducial and full phase space**
  - Merged $t\bar{t}b\bar{b}/t\bar{t}bL$ and $t\bar{t}c\bar{c}/t\bar{t}cL$ categories $\rightarrow 2^{\mathrm{nd}}$ b/c-jet outside of accep. or merged with 1st jet
  - $t\bar{t}b\bar{b}$ yields slightly higher than predicted, $t\bar{t}c\bar{c}$ and $t\bar{t}LL$ yields slightly lower
  - Main uncertainties: c-tagging calib., JES unc., ME/PS matching, $\mu_{\mathrm{R}}/\mu_{\mathrm{F}}$ scales in ME calc.



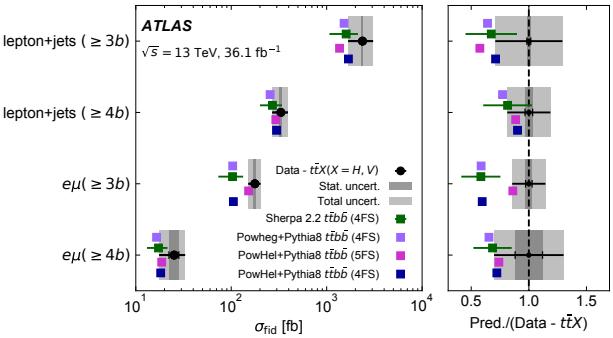
Results from simultaneous fit

	Result	POWHEG	MadGraph5_aMC@NLO
	Fiducial phase space		
$\sigma_{t\bar{t}c\bar{c}}$ [pb]	$0.165 \pm 0.023 \pm 0.025$	0.187 ± 0.038	0.189 ± 0.032
$\sigma_{t\bar{t}b\bar{b}}$ [pb]	$0.119 \pm 0.010 \pm 0.015$	0.097 ± 0.021	0.101 ± 0.023
$\sigma_{t\bar{t}LL}$ [pb]	$5.40 \pm 0.11 \pm 0.45$	5.95 ± 1.02	6.32 ± 0.94
R_c [%]	$2.42 \pm 0.32 \pm 0.29$	2.53 ± 0.18	2.43 ± 0.17
R_b [%]	$1.75 \pm 0.14 \pm 0.18$	1.31 ± 0.12	1.30 ± 0.16
	Full phase space		
$\sigma_{t\bar{t}c\bar{c}}$ [pb]	$8.0 \pm 1.1 \pm 1.3$	9.1 ± 1.8	8.9 ± 1.5
$\sigma_{t\bar{t}b\bar{b}}$ [pb]	$4.09 \pm 0.34 \pm 0.55$	3.34 ± 0.72	3.39 ± 0.66
$\sigma_{t\bar{t}LL}$ [pb]	$231 \pm 5 \pm 21$	255 ± 43	261 ± 37
R_c [%]	$2.69 \pm 0.36 \pm 0.32$	2.81 ± 0.20	2.72 ± 0.19
R_b [%]	$1.37 \pm 0.11 \pm 0.14$	1.03 ± 0.08	1.03 ± 0.09

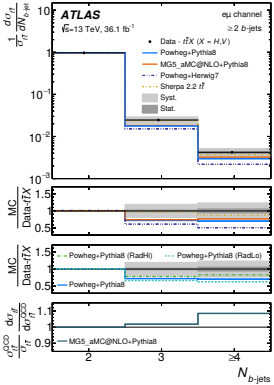
Fiducial and full phase space results

Measured $t\bar{t}b\bar{b}$ cross-sections and differential distributions

- ATLAS $e\mu/\ell$+jets analysis shows higher than predicted $t\bar{t}b\bar{b}$ yields within fid. volume
  - Analysis provides **inclusive and differential cross-section measurements**
  - Differential distribution of $\mathbf{n_{bjets}}$ **with clear trend**
    $\rightarrow$ other observables well described by most MC predictions in both channels



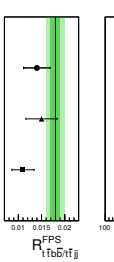
Fiducial measurement of $t\bar{t}b\bar{b}$ cross-sections



Unfolded n_{bjets} distribution for $e\mu$ case

Measured $t\bar{t}b\bar{b}$ cross-sections at CMS

- **Two separate analyses from CMS for $t\bar{t}b\bar{b}$** covering different $t\bar{t}$ decays
  - $t\bar{t}b\bar{b}$ yields in dilepton/ℓ+jets higher than predicted, while $t\bar{t}j\bar{j}$ yields show good agreement
  - All-hadronic $t\bar{t}b\bar{b}$ yields in parton-independent (PI) and -based (PB) fiducial volumes
    $\rightarrow$ increased $t\bar{t}b\bar{b}$ yield w.r.t. prediction again observed



Dilepton $t\bar{t}$ decays | ℓ+jets $t\bar{t}$ decays | All-hadronic $t\bar{t}$ decays

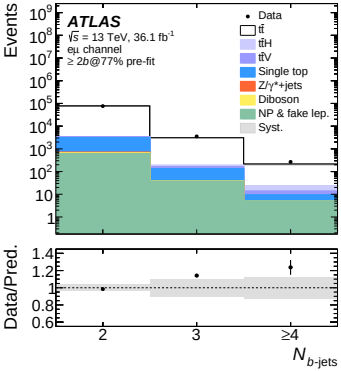
Summary

- Presented experimental status of $\mathbf{t\bar{t}W}$, $\mathbf{t\bar{t}b\bar{b}}$ and $\mathbf{t\bar{t}c\bar{c}}$ production
- $\mathbf{t\bar{t}W}$ production
  - Important background in analyses with **multilepton final states** (e.g. $t\bar{t}H$ and $t\bar{t}t\bar{t}$)
  - Both collaborations observe **higher yields than predicted** with $36\,\mathrm{fb}^{-1}$ of data
  - Recent $t\bar{t}H$ and $t\bar{t}t\bar{t}$ analyses also observe high normalisation factors
- $\mathbf{t\bar{t}b\bar{b}}$ and $\mathbf{t\bar{t}c\bar{c}}$ production
  - Important backgrounds in analyses like $\mathbf{t\bar{t}H(b\bar{b})}$
  - CMS provided **first dedicated measurement of $\mathbf{t\bar{t}c\bar{c}}$** production
    $\rightarrow$ $t\bar{t}b\bar{b}$ yields slightly higher than predicted, $t\bar{t}c\bar{c}$ and $t\bar{t}LL$ yields slightly lower
  - Generally **$\mathbf{t\bar{t}b\bar{b}}$ measurements found higher yields than predicted**
    $\rightarrow$ normalisation factors of $1.2 - 1.7 \times \sigma_{\mathrm{Powheg+Pythia8}}$ (typically 1-2 standard deviations)
- **Stay tuned for full Run 2 results on these processes!**
  - Clearly need to improve theoretical and experimental understanding

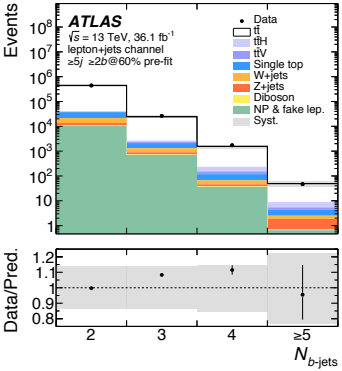
Backup

Event selection in ATLAS $t\bar{t}b\bar{b}$ analysis

- Here: ATLAS $e\mu/\ell$+jets analysis at $36.1\,\mathrm{fb}^{-1}$
- **Baseline selection** requires $n_{\mathrm{jets}} \geq 2\,(5)$ and $n_{\mathrm{bjets}} \geq 2\,(2)$ in $e\mu$ (ℓ+jets) channels



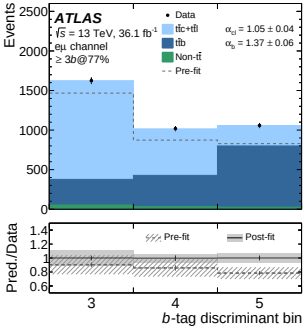
Baseline data/MC agreement in $e\mu$ channel



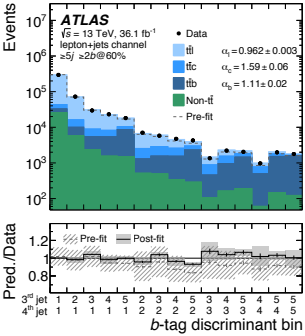
Baseline data/MC agreement in ℓ+jets channel

Correction factors for $t\bar{t}c$ and $t\bar{t}l$ backgrounds

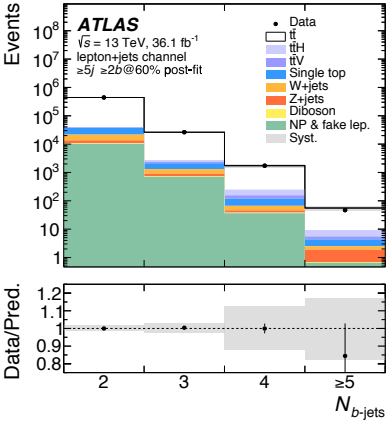
- Normalisation of **$t\bar{t}c$ and $t\bar{t}l$ backgrounds** estimated in template fits
  - Binned vs. b-tagging discriminant of 3rd (3rd and 4th) leading jet in $e\mu$ (ℓ+jets) channels
- Data/MC agreement improved for baseline selection



$e\mu$ correction factors



ℓ+jets correction factors



ℓ+jets data/MC agreement after correction
(impact on $e\mu$ channel very similar)