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

---

**Top Quark Production Cross Section Measurements**

- **ATLAS Preliminary**
  - Run 1,2 $\sqrt{s} = 5,7,8,13$ TeV
  - Data $0.257 \, fb^{-1}$ at $\sqrt{s} = 7$ TeV
  - Data $4.5 \, fb^{-1} - 4.6 \, fb^{-1}$ at $\sqrt{s} = 8$ TeV
  - Data $20.2 \, fb^{-1} - 20.3 \, fb^{-1}$ at $\sqrt{s} = 13$ TeV

**Status**: May 2021
t\bar{t}W production

CMS paper (35.9 $\text{fb}^{-1}$): 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, ...

---

**2$\ell$SS channel in CMS $t\bar{t}H$ ML**


**ATLAS $t\bar{t}t\bar{t}$ ML analysis**

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

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

### CMS 2\ell SS selection

<table>
<thead>
<tr>
<th>Requirement</th>
<th>$2\ell$(p,m)-1b</th>
<th>$3\ell$1b(p,m)</th>
<th>$3\ell &gt; 1b$(p,m)</th>
<th>$&gt; 3$1b(p,m)</th>
<th>$&gt; 3$ &gt; 1b(p,m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n_{jets}$</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>$n_{bjets}$</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>&gt; 1</td>
<td>&gt; 1</td>
</tr>
<tr>
<td>$p_T$ (lead. $\ell$)</td>
<td>&gt; 25 GeV (40 GeV in $ee$ case)</td>
<td>&gt; 25 GeV</td>
<td>&gt; 12 GeV and $</td>
<td>m_{\ell\ell} - m_Z</td>
<td>&gt; 15$ GeV</td>
</tr>
<tr>
<td>$p_T$ (sublead. $\ell$)</td>
<td>&gt; 20 GeV</td>
<td>&gt; 40 GeV</td>
<td>&gt; 40 GeV</td>
<td>&gt; 20 GeV</td>
<td>&gt; 240 GeV</td>
</tr>
<tr>
<td>$m_{\ell\ell}$</td>
<td>&gt; 10 GeV in the $2e$ and $2\mu$ regions</td>
<td>&gt; 10 GeV in the $2e$ and $2\mu$ regions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p_T^{miss}$</td>
<td>&gt; 40 GeV</td>
<td>&gt; 40 GeV</td>
<td>&gt; 40 GeV</td>
<td>&gt; 20 GeV</td>
<td>&gt; 240 GeV</td>
</tr>
</tbody>
</table>

### Variable 3 $\ell$ 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

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

09/06/21 4 / 19
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. $t\bar{t}H$, $WZ$)
  - $3\ell$: 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** $3\ell$ pre-fit agreement
Background estimation and event selection

- **Estimation of charge flip/mis-ID electrons**
  - ATLAS applies data-driven technique using CR with $m_{ll}$ around $m_Z$
  - CMS estimates charge flip/mis-ID probabilities from MC and applies these to $2\ell$OS data

- **BDT-based event selection** in CMS analysis with BDT discriminant $D$
  - Input variables given e.g. as $n_{jets}$, $n_{bjets}$, $H_T$, $p_T^{miss}$, ... (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 $\ell$ CR

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

CMS event-selection BDT

Peter Falke (University of Bonn)
ttX results from ATLAS and CMS

09/06/21 6 / 19
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: $\sigma_{t\bar{t}W} = 0.87 \pm 0.13^{\text{(stat)}} \pm 0.14^{\text{(syst)}} \text{pb}$
  - CMS: $\sigma_{t\bar{t}W} = 0.77 \pm 0.12^{\text{(stat)}} \pm 0.13^{\text{(syst)}} \text{pb}$
  - Theory prediction: $\sigma_{t\bar{t}W} = 0.59^{+0.15}_{-0.10}^{\text{(scale)}} \pm 0.01^{\text{(PDF)}} \text{pb}$ (Eur. Phys. J. C 80 (2020) 428)

- 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: $\sigma_{t\bar{t}W} = 0.87 \pm 0.13^{\text{stat}} \pm 0.14^{\text{syst}} \text{pb}$
  - CMS: $\sigma_{t\bar{t}W} = 0.77^{+0.12}_{-0.11}\text{(stat)} \pm 0.13^{\text{syst}} \text{pb}$
  - Theory prediction: $\sigma_{t\bar{t}W} = 0.59^{+0.15}_{-0.10}\text{(scale)} \pm 0.01\text{(PDF)} \text{pb}$ (Eur. Phys. J. C 80 (2020) 428)

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 $t\bar{t}H$ ML analysis**

**ATLAS $t\bar{t}H$ ML analyses**
(ATALAS-CONF-2019-045)

**CMS $t\bar{t}H$ ML analysis**
t¯tb¯b and t¯tc¯c production

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

ATLAS t¯tbb paper (eµ and ℓ+jets t¯t, 36.1 fb\(^{-1}\)): JHEP 04 (2019) 046

CMS t¯tbb and t¯ttjj paper (dilepton and ℓ+jets t¯t, 35.9 fb\(^{-1}\)): JHEP 07 (2020) 125

CMS t¯tbb paper (all-hadronic t¯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 $\mu_F/\mu_R$ scale choice

Other Uncertainty

Single-{$\ell$} 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 $\mu$, incl. $\tau_{lep}$ decays) with $p_T > 25$ GeV and $|\eta| < 2.4$
  - $n_{jets} \geq 4$ with $p_T > 30$ GeV, $|\eta| < 2.4$ and $\Delta R(\ell, jet) > 0.5$
  - $m_{\ell\ell} > 12$ GeV for all events; $p_T^{miss} > 30$ GeV and $|m_{\ell\ell} - m_Z| > 15$ GeV in $ee$ and $\mu\mu$ cases

- **Matching of jets and partons**
  - Two ($b$-)jets from top decays need to be identified → separate treatment of jets from add. radiation
  - NN trained for jet-parton assignment → use jet kinematics, $b/c$-tagging discriminators, ...
  - Loop over all permutations of jet-parton assignments → score indep. of assignment to top or anti-top quark → 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
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**: $C_{vL} = \frac{P(c)}{P(c)+P(udsg)}$ and $C_{vB} = \frac{P(c)}{P(c)+P(b)+P(bb)}$

- **Dedicated calibration technique** for $C_{vL}$ and $C_{vB}$ scores
  - CRs: $t\bar{t}$ (for $b$-jets), $W + c$ ($c$-jets) and $DY + \text{jets}$ (light jets)
Fit strategy and MVA setup

- **Event-level separation** of \( t\bar{t}bb \), \( t\bar{t}bL \), \( t\bar{t}c\bar{c} \), \( t\bar{t}cL \), \( t\bar{t}LL \) and \( t\bar{t} + \) other classes
  - Multi-class NN using \( \text{CvsL}/\text{CvsB} \) of leading two jets, \( \Delta R(\text{jet,jet}) \) and permutation score
  - **Two discriminators**: \( \Delta_c^b = \frac{P(t\bar{t}c\bar{c})}{P(t\bar{t}c\bar{c}) + P(t\bar{t}bb)} \) and \( \Delta_c^L = \frac{P(t\bar{t}c\bar{c})}{P(t\bar{t}c\bar{c}) + P(t\bar{t}LL)} \)

- Final fit done in **unrolled** \( \Delta_c^L \) vs. \( \Delta_c^b \) distribution
  - \( \Delta_c^L \otimes \Delta_c^b \): \([0, 0.45, 0.6, 0.9, 1.0] \otimes [0, 0.3, 0.45, 0.5, 1.0]\)
Results of cross-section and $R_b/R_c$ measurements

- Measurement of $t\bar{t}b\bar{b}$ (incl. $t\bar{t}bL$), $t\bar{t}c\bar{c}$ (incl. $t\bar{t}cL$) and $t\bar{t}LL$ production as well as $R_b = t\bar{t}b\bar{b}/t\bar{t}jj$ and $R_c = t\bar{t}c\bar{c}/t\bar{t}jj$ (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 → $2^\text{nd}$ $b/c$-jet outside of accep. or merged with $1^\text{st}$ jet
- $t\bar{t}bb$ 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, $\mu_R/\mu_F$ scales in ME calc.

<table>
<thead>
<tr>
<th>Result</th>
<th>POWHEG</th>
<th>MADGRAPH5_aMC@NLO</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_{t\bar{t}c\bar{c}}$ [pb]</td>
<td>$0.165 \pm 0.023 \pm 0.025$</td>
<td>$0.187 \pm 0.038$</td>
</tr>
<tr>
<td>$\sigma_{t\bar{t}b\bar{b}}$ [pb]</td>
<td>$0.119 \pm 0.010 \pm 0.015$</td>
<td>$0.097 \pm 0.021$</td>
</tr>
<tr>
<td>$\sigma_{t\bar{t}LL}$ [pb]</td>
<td>$5.40 \pm 0.11 \pm 0.45$</td>
<td>$5.95 \pm 1.02$</td>
</tr>
<tr>
<td>$R_c$ [%]</td>
<td>$2.42 \pm 0.32 \pm 0.29$</td>
<td>$2.53 \pm 0.18$</td>
</tr>
<tr>
<td>$R_b$ [%]</td>
<td>$1.75 \pm 0.14 \pm 0.18$</td>
<td>$1.31 \pm 0.12$</td>
</tr>
</tbody>
</table>

Fiducial phase space results

<table>
<thead>
<tr>
<th>Result</th>
<th>POWHEG</th>
<th>MADGRAPH5_aMC@NLO</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_{t\bar{t}c\bar{c}}$ [pb]</td>
<td>$8.0 \pm 1.1 \pm 1.3$</td>
<td>$9.1 \pm 1.8$</td>
</tr>
<tr>
<td>$\sigma_{t\bar{t}b\bar{b}}$ [pb]</td>
<td>$4.09 \pm 0.34 \pm 0.55$</td>
<td>$3.34 \pm 0.72$</td>
</tr>
<tr>
<td>$\sigma_{t\bar{t}LL}$ [pb]</td>
<td>$231 \pm 5 \pm 21$</td>
<td>$255 \pm 43$</td>
</tr>
<tr>
<td>$R_c$ [%]</td>
<td>$2.69 \pm 0.36 \pm 0.32$</td>
<td>$2.81 \pm 0.20$</td>
</tr>
<tr>
<td>$R_b$ [%]</td>
<td>$1.37 \pm 0.11 \pm 0.14$</td>
<td>$1.03 \pm 0.08$</td>
</tr>
</tbody>
</table>

Full phase space results

Results from simultaneous fit
Measured $t\bar{t}b\bar{b}$ cross-sections and differential distributions

- **ATLAS $e\mu/\ell+\text{jets}$ 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 $n_{\text{bjets}}$ with clear trend
  → other observables well described by most MC predictions in both channels

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

Unfolded $n_{\text{bjets}}$ 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/\(\ell+\text{jets}\) higher than predicted, while $t\bar{t}jj$ yields show good agreement
  - All-hadronic $t\bar{t}b\bar{b}$ yields in parton-independent (PI) and -based (PB) fiducial volumes → increased $t\bar{t}b\bar{b}$ yield w.r.t. prediction again observed

\begin{align*}
  \sigma_{\text{Fiducial PB}} & = 0.02 \text{ pb} \\
  \sigma_{\text{Total}} & = 0.03 \text{ pb} \\
  \sigma_{\text{Stat. UC}} & = 0.01 \text{ pb} \\
  \sigma_{\text{Total UC}} & = 0.02 \text{ pb}
\end{align*}
Summary

- Presented experimental status of $t\bar{t}W$, $t\bar{t}b\bar{b}$ and $t\bar{t}c\bar{c}$ production

- $t\bar{t}W$ 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 fb$^{-1}$ of data
  - Recent $t\bar{t}H$ and $t\bar{t}t\bar{t}$ analyses also observe high normalisation factors

- $t\bar{t}b\bar{b}$ and $t\bar{t}c\bar{c}$ production
  - Important backgrounds in analyses like $t\bar{t}H(b\bar{b})$
  - CMS provided first dedicated measurement of $t\bar{t}c\bar{c}$ production
  - $t\bar{t}b\bar{b}$ yields slightly higher than predicted, $t\bar{t}c\bar{c}$ and $t\bar{t}LL$ yields slightly lower
  - Generally $t\bar{t}b\bar{b}$ measurements found higher yields than predicted
  - Normalisation factors of $1.2 - 1.7 \times \sigma_{\text{Powheg+Pythia8}}$ (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 $36.1\,\text{fb}^{-1}$
- **Baseline selection** requires $n_{\text{jets}} \geq 2\,(5)$ and $n_{b\text{jets}} \geq 2\,(2)$ in $e\mu\,(\ell+\text{jets})$ channels

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

Baseline data/MC agreement in $\ell+\text{jets}$ 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 $3^{rd}$ ($3^{rd}$ and $4^{th}$) leading jet in $e\mu$ ($\ell$+jets) channels
- Data/MC agreement improved for baseline selection

$e\mu$ correction factors  
$\ell$+jets correction factors

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