Measurement of $W+$jets and top quark pair production cross-sections in ATLAS

W. Verkerke (NIKHEF) 
\textit{on behalf of the ATLAS Collaboration}
Introduction

- Measurements of W+jets and ttbar cross-sections

  - Both measurements are important test of perturbative QCD
    - Accuracy of $\sigma$(ttbar) calculation now at 8% level, $\sigma = 164.6 \pm 11.4 \pm 15.7$ pb
    - W/Z boson production good place to study associated jet production

  - Both processes are major background in BSM, Higgs searches

  - New physics can also show in ttbar final state (e.g. $Z' \rightarrow$ttbar)

Production dominated by gluon/gluon mode at LHC

Production dominated quark/anti-quark mode
Detector and data sample
Inner detector

Transition Radiation Tracker & SemiConductor Tracker
(tracking, particle ID, $p_T$ resolution)

Pixel detector: (b-tagging)
Calorimeters

Liquid Argon Calorimeter
(Energy resolution for EM objects)

Tile Calorimeter
(Absorption of hadronic showers)

Jet reconstruction,
Missing Transverse Energy
Muon spectrometer

Particle identification
pT resolution
The 2010 data taking period – data samples

**Data sample for ttbar paper (2.9 pb⁻¹)**

**Data sample for W+jets paper (1.3 pb⁻¹)**

<table>
<thead>
<tr>
<th>Inner Tracking Detectors</th>
<th>Calorimeters</th>
<th>Muon Detectors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pixel</td>
<td>LAr EM</td>
<td>MDT</td>
</tr>
<tr>
<td>SCT</td>
<td>LAr HAD</td>
<td>99.9</td>
</tr>
<tr>
<td>TRT</td>
<td>LAr FWD</td>
<td>99.8</td>
</tr>
<tr>
<td></td>
<td>Tile</td>
<td>99.9</td>
</tr>
</tbody>
</table>

% of good quality data

Wouter Verkerke, NIKHEF
Measurement of the W+jets cross section

Object reconstruction

Muon

Jet

Missing ET

Electron
W+jets – Object definitions

- **Muons** ($p_T > 20$ GeV, $|\eta| < 2.4$)
  - Combined tracks from Inner Detector and Muon Spectrometer
  - Trk Isolation $E_T(R=0.2) < 1.8$ GeV
  - Small impact parameter (to reject cosmics)

- **Electrons** ($p_T > 20$ GeV, $|\eta| < 2.4$)
  - ‘tight’ electrons particle ID requirements
  - Excluding $1.37 < \eta < 1.52$

- **Jets** ($p_T > 20$ GeV and $|\eta| < 2.8$)
  - Anti-$k_T$ Jet algorithm ($r=0.4$) from Topological Calorimeter Clusters
  - $p_T, \eta$ – dependent calibration to hadronic energy scale
  - Must be associated to the primary vertex

- **Missing Transverse energy**
  - Constructed from Topological Cluster in Calorimeter
  - Each cluster is taken at EM or Hadronic scale depending on nature of deposited inferred from cluster shape
  - Cluster(s) associated with muon track removed and substituted with muon track $p_T$
**W+jets – event selection**

- **Event selection criteria**
  - **Muon channel**: Exactly one good muon
  - **Electron channel**: Exactly one good (‘tight’) electron, no extra electrons passing ‘medium’ ID criteria
  - $E_{T}^{\text{miss}} > 25$ GeV, $M_T > 40$ GeV
  - At least one primary vertex with at least 3 tracks

\[
M_T = \sqrt{2 p_T^U p_T^\nu (1 - \cos(\phi^U - \phi^\nu))}
\]
Leptonic backgrounds (estimated from simulation)
- $W \rightarrow \tau \nu$, where $\tau$ decays into electron/muon
- $Z \rightarrow e+e-$/$Z \rightarrow \mu+\mu-$, where one electron/muon is not identified (and hadronic energy is mismeasured $\rightarrow E_T^{miss}$)
- $tt$ in semileptonic decay mode with one electron/muon

QCD multi-jet background (data-driven estimation)
- Component from fake electrons (mis-measured jets) with additional energy mis-measurement resulting in $E_T^{miss}$
- Component from semileptonic $b,c$ quark decays (resulting in a real, but non-prompt electron or muon, and a neutrino generating $E_T^{miss}$)
Data-driven measurement of QCD background

- Measure QCD multi-jet background by fitting ETmiss data distribution to sum of two templates
  1. QCD – μ: from MC, e: from data,
     - Construction of $E_T^{\text{miss}}$ template for QCD multi-jets in the electron channel from sample with looser electron ID cuts (shower-shape cuts) and inverted track-cluster matching requirement.
  2. Signal plus leptonic backgrounds (MC)

- Background estimates for all jet bins:
  \( W \rightarrow e\nu \)
  \( W \rightarrow \mu\nu \)
From yield to cross section

- Correct signal yields to (particle level) cross section
  - Detector and reconstruction efficiency calculated using Alpgen + Atlas detector simulation
  - Define particle-level jets by running jet-clustering on all simulated particles with a lifetime > 10ps (except muons and neutrinos)
  - Define particle-level lepton as lepton after QED radiation plus the energy of all photons within cone with DR=0.1 around lepton
  - Trigger efficiency measured from data (electrons: repeat analysis with $E_T^{\text{miss}}$ trigger, muons: measure efficiency from unbiased muon sample from $Z\rightarrow\mu\mu$ decays)
  - Correction factors expressed as 1-dimensional functions in jet multiplicity, $p_T$ of leading jet, $p_T$ of next-to leading jet
Cross sections vs jet multiplicity

Particle-level cross sections in limited kinematic region

\[ E_T^{\text{jet}}> 20 \text{ GeV}, \ |\eta^{\text{jet}}|<2.8, \]
\[ E_T^e>20 \text{ GeV}, \ |\eta^e|<2.47 \text{ (excl. 1.37}<\eta^e<1.52) \]
\[ p_T^\gamma>25 \text{ GeV}, \ M_T>40 \text{ GeV}, \ \Delta R(lj)>0.5 \]

**Pythia**: Leading-order generator

**Alpgen, Sherpa**: Match N+1 ME to a LL parton shower (rescaled to NNLO inclusive XS)

**MCFM**: NLO prediction at parton level for \( N_{\text{jet}} \leq 2 \), LO for \( N_{\text{jet}} = 3 \)
Cross sections vs (next-to)-leading jet $p_T$

Particle-level cross sections in limited kinematic region

**Pythia**: Leading-order generator

**Alpgen, Sherpa**: Match N+1 ME to a LL parton shower (rescaled to NNLO inclusive XS)

**MCFM**: NLO prediction at parton level for $N_{\text{jet}} \leq 2$, LO for $N_{\text{jet}} = 3$
(N+1)-jet/N-jet cross sections ratios

Particle-level cross sections in limited kinematic region

**Pythia**: Leading-order generator

**Alpgen, Sherpa**: Match N+1 ME to a LL parton shower (rescaled to NNLO inclusive XS)

**MCFM**: NLO prediction at parton level for \( N_{\text{jet}} \leq 2 \), LO for \( N_{\text{jet}} = 3 \)

Wouter Verkerke, NIKHEF
# Systematic uncertainties

- **Leading systematic uncertainties**
  - **Jet Energy Scale uncertainty**: 7-10% \( (p_T, \eta \text{ dependent}) \)
  - **Luminosity uncertainty** (11%)
Summary on W+jets

• First ATLAS measurement of W+jets production as function of jet multiplicity and (next-to) leading jet $p_T$
  ‒ Corrected for all detector effects
  ‒ Quoted in a limited and well-defined range of jet and lepton kinematics
  ‒ As expected, Pythia ($2\rightarrow1$ ME merged with $2\rightarrow2$ ME and LL PS), does not provide good description of data for $N_{jet} > 1$
  ‒ Good agreement with multi-parton ME generators Alpgen,Sherpa, as well as MCFM calculations (NLO for $N_{jet} \leq 2$)

• Full 2010 dataset on tape ~25x larger
  ⇒ Will greatly extend experimental reach in $N_{jets}$
Summary on W+jets

- Uncorrected $N_{\text{jet}}$ plots on 35 pb$^{-1}$

\[ \int L \, dt = 36 \, \text{pb}^{-1} \]

\[ \int L \, dt = 35 \, \text{pb}^{-1} \]

Data 2010 ($\sqrt{s} = 7 \, \text{TeV}$)
- $W \rightarrow e\nu$ (Alpgen)
- QCD
- $Z \rightarrow ee$
- $W \rightarrow \tau\nu$
- $t\bar{t}$

MC normalised to data
Statistical Errors Only
Measurement of the top quark pair production cross section

“Measurement of the top quark-pair production cross section with ATLAS in pp collisions at √s=7 TeV”,
CERN-PH-EP-2010-064, Accepted by EPJC

Calibrating the b-Tag and Mistag Efficiencies of the SV0 bTagging Algorithm in 3 pb⁻¹ of Data with the ATLAS Detector,
ATLAS-CONF-2010-099, cdsweb.cern.ch/record/1299573.
Topology of top quark pair decays

Dilepton channel (2xW→lν), 1/9 of cross section

Lepton+jets channel (1xW→lν), 4/9 of cross section

All Hadronic channel (0xW→lν), 4/9 of cross section

2 leptons large $E_T^{miss}$
2 b-jets

1 lepton large $E_T^{miss}$
2 b-jets
2 light jets

2 b-jets
4 light jets

Wouter Verkerke, NIKHEF
Finding top quark pairs in early data – Analysis strategy

• **Lepton+jets** → Look for events with ≥4 jets, of which ≥1 b-jets, hard lepton, missing ET
  - Can reconstruct also invariant mass of jets from $t\rightarrow W(qq)b$, but simple procedures suitable for early data taking have low efficiency (~20-30%) → Not selection tool for early data analysis
  - Strategy: Evidence for top based on excess of events w.r.t known backgrounds

• **Dilepton** → Look for events with ≥2 jets, two hard leptons, missing ET
  - No easy kinematic signature in events that is characteristic for top
  - Strategy: Evidence for top based on excess of events w.r.t known backgrounds

• Understanding of backgrounds is key
  - Use data driven approach wherever possible

Wouter Verkerke, NIKHEF
Object definitions

- **Muons** ($p_T > 20$ GeV, $|h| < 2.5$)
  - Combined tracks from Inner Detector and Muon Spectrometer
  - Calo, Trk Isolation: $E_T(R=0.4) < 4$ GeV, $p_T(R=0.4) < 4$ GeV
  - Jet isolation: must be DR>0.4 from nearest jet

- **Electrons** ($p_T > 20$ GeV, $|\eta_{\text{cluster}}| < 2.47$)
  - ‘medium’ electron particle PID cuts (excluding $1.37 < |\eta| < 1.52$),
  - Calo Isolation: $E_T(R=0.2) < 4 + 0.023*E_T(\text{el})$ GeV

- **Jets** ($p_T > 25$, $|\eta| < 2.5$)
  - Anti-$k_T$ Jet algorithm ($r=0.4$) from Topological Calorimeter Clusters
  - $p_T, \eta$ - dependent calibration to hadronic energy scale
  - Must be DR>0.2 from electron

- **Missing Transverse Energy**
  - Calculated from topological clusters of calorimeter cells. Cells associated with reconstructed jets are calibrated at the hadronic scale.
  - Cells identified with electrons, muons are removed and substituted with their respective $E_T$ values

Wouter Verkerke, NIKHEF
Identifying jets originating from b-quarks

- Identification of jets originating from b-quarks very important in top physics

B-tagging in ATLAS

- **General concept**: Exploit relatively long lifetime of b-hadrons resulting in flight times of $O(\text{few})$ mm $\rightarrow$ identifiable secondary decay vertex

- Multiple techniques possible – here comparatively simple and robust method exploited: 
  \textit{selection based on decay length significance } $L / \sigma(L)$

- **Alternative concept**: look for soft leptons inside jet (from semileptonic b decay) – low efficiency $\rightarrow$ semileptonic decay method was used to calibrate vertex-based method

Wouter Verkerke, NIKHEF
Decay length significance of secondary vertices

- Consider a jet b-tagged if it contains a secondary vertex and decay length significance that results in 50% b-tag efficiency in simulated ttbar events

- Distribution of $L/\sigma(L)$ for QCD multi-jet events with a secondary vertex

- Performance of b-tagging in data depends critically on many details of detector performance
  - Data-driven estimates of efficiency and mis-tag rate of SV0 tagger
The lepton+jets channel – Event selection

- Two channels: e+jets, μ+jets
- 1,2,3,≥4-jet Pretag samples
  - Exactly one good lepton (p_T>20) GeV
  - E_T^{miss}>20 GeV and (E_T^{miss} + m_T(W))>60 GeV
  - 1,2,3 or at least 4 jets (p_T>25, |η|<2.5)

- All samples dominated by background (W+jets, QCD multi-jet), expect signal to concentrate in ≥4 jet bin

\[ \text{Events} \]
\[ \int L = 2.9 \text{ pb}^{-1} \]

\[ \text{ATLAS} \]
\[ \text{pretag e+jets} \]

- data
- \( tt \)
- single top
- Z + jets
- W + jets
- QCD
- uncertainty

\[ \text{Events} \]
\[ \int L = 2.9 \text{ pb}^{-1} \]

\[ \text{ATLAS} \]
\[ \text{pretag μ+jets} \]

- data
- \( tt \)
- single top
- Z + jets
- W + jets
- QCD
- uncertainty

QCD multijet
estimated from data (see later)

W/Z+jets
from simulation (Alpgen)

tt, single top
from simulation (MC@NLO)

Wouter Verkerke, NIKHEF
The lepton+jets channel – Event selection

- 1, 2, 3, ≥4-jet tagged-samples
  - Each a subset of the corresponding pretag sample
  - At least one of jets is b-tagged

- Signal efficiency ≈ 75% (since 2 b-jets)

- Backgrounds reduced by > O(10), since mostly light jets,
  - Remaining backgrounds dominated by events with b-quark or c-quark jets

QCD multijet estimated from data (see later)

W/Z+jets from simulation (Alpgen)

tt, single top from simulation (MC@NLO)

Wouter Verkerke, NIKHEF
Background estimation

• Next step is to quantify backgrounds in ≥4-jets tagged sample

• Background from $W+(bb/cc/c)+jets$ not surprising
  – Same final state – thus also partly irreducible

• Understanding **QCD multi-jet** background
  – Final states with many jets from gluons and $(b,c,\text{light})$ quarks. *No prompt leptons, no prompt neutrinos.*
  – Small fraction of jets very EM-like $\rightarrow$ Can be misidentified as a reconstructed electron (‘fake’ electron). Event is only background if there is a separate cause of true or fake missing $E_T$.
  – Semileptonically decaying quarks can result in true lepton (‘non-prompt’) plus a neutrino (giving rise to $E_T^{\text{miss}}$). Event is background if resulting lepton is sufficiently isolated

Wouter Verkerke, NIKHEF
Estimating QCD multi-jet background in $\mu$ channel

• Summary: three sources of reconstructed leptons
  1. From prompt true leptons [from $t(t)$, W/Z+jets]
  2. From non-prompt true leptons (from quark decay) [QCD multi-jet]
  3. Fake electrons (from jets, reconstruction mistakes) [QCD multi-jet]

(From here on lump 2 ‘non-prompt’ and 3 ‘fake’ together for notational simplicity)

• Estimating QCD background = estimating fake lepton bkg

• Can make a fully data-driven estimate of fake lepton background
  - Introduce a ‘loose’ lepton selection besides ‘standard’
  - Measure behaviour of real and fake leptons in control regions
  - Apply the ‘matrix method’
The matrix method illustrated

- Given a ‘loose’ and ‘std’ lepton selection the composition of either in term of real and fake leptons is

\[ N_{\text{loose}} = N_{\text{real}}^{\text{loose}} + N_{\text{fake}}^{\text{loose}} \]

This is what we’re after!

\[ N_{\text{std}} = N_{\text{real}}^{\text{std}} + N_{\text{fake}}^{\text{std}} \]

Define fractions \( f, r \) as

\[ f = \frac{\text{loose}}{\text{fake}} \quad r = \frac{\text{std}}{\text{real}} \]

Rewritten using \( f, r \) (‘the matrix’)

\[ N_{\text{loose}} = N_{\text{real}}^{\text{loose}} + N_{\text{fake}}^{\text{loose}} \]

\[ N_{\text{std}} = r \cdot N_{\text{real}}^{\text{loose}} + f \cdot N_{\text{fake}}^{\text{loose}} \]

If we know \( r, f \) externally we can calculate \( N_{\text{fake}}^{\text{std}} \) from \( N_{\text{loose}}, N_{\text{std}} \)

Wouter Verkerke, NIKHEF
Fake lepton background in the $\mu+$jets channel

- Need $N^{\text{loose}}$, $N^{\text{std}}$ (trivial), $r$ and $f$ (from control regions)
- Measure $r = 0.990 \pm 0.003$ from inclusive $Z \to \mu\mu$ events
- Measure $f = 0.339 \pm 0.013\text{(stat.)} \pm 0.061\text{(syst.)}$ from control regions enhanced in fake leptons
  - Key is to find control region that is similar to signal region (in lepton $p_T, \eta$ distribution) for result to be applicable
  - Control region A: $E_T^{\text{miss}} < 10$ GeV, at least 1 jet ($p_T > 25$ GeV)
    $\rightarrow$ Event sample dominated by QCD multijet production
  - Control region B: nominal $E_T^{\text{miss}} (>20$ GeV), at least 1 jet ($p_T > 25$ GeV), lepton with high impact parameter significance (>5)
    $\rightarrow$ Lepton sample dominated by non-prompt leptons
- Apply procedure in every jet bin

### Estimated number of events with fake leptons

<table>
<thead>
<tr>
<th>$\mu+$jets channel</th>
<th>1-jet tagged</th>
<th>2-jet tagged</th>
<th>3-jet tagged</th>
<th>$\geq 4$-jet tagged</th>
<th>3-jet zero-tag</th>
<th>$\geq 4$-jet zero-tag</th>
</tr>
</thead>
<tbody>
<tr>
<td>QCD (DD)</td>
<td>6.1 ± 2.9</td>
<td>3.4 ± 1.8</td>
<td>1.5 ± 0.8</td>
<td>0.8 ± 0.5</td>
<td>4.9 ± 2.3</td>
<td>1.7 ± 1.1</td>
</tr>
</tbody>
</table>

Wouter Verkerke, NIKHEF
Measuring the fake lepton bkg in the e+jets channel

- Application of matrix method in e+jets channel more challenging
  - Two sources of background: non-prompt (true) electrons and fake electrons (reconstruction mistakes)
  - Need to worry about mix of sources being similar in control regions

- Instead apply ‘fitting method’ – *Similar to W+jets analysis*
  - Fit distribution of $E_T^{\text{miss}}$ to sum of 2 templates
  - Template shapes for fake/non-prompt lepton contribution from data control samples
  - Template shape of prompt lepton components (signal, W+jets) from simul.
  - Fit for fraction of fake leptons in side-band ($E_T^{\text{miss}}<20$ GeV) → Extrapolate to fraction of fake leptons in signal region using template shapes
Measuring the fake lepton bkg in the e+jets channel

• Key issue (again) representative control regions to determine shape of fake lepton template
  - ‘jet-electrons’ → Events which have – instead of standard electron – an extra jet with an EM fraction of 80%-95%.
  - ‘non-electrons’ → Standard electron selection modified: candidates must fail track quality cut for innermost detector layers
  - Take difference in results as metric for systematic uncertainty due to choice of control sample

• Apply procedure in every jet bin

<table>
<thead>
<tr>
<th>Estimated number of events with fake leptons</th>
</tr>
</thead>
<tbody>
<tr>
<td>e+jets channel</td>
</tr>
<tr>
<td>1-jet tagged</td>
</tr>
<tr>
<td>QCD (DD)</td>
</tr>
</tbody>
</table>

Wouter Verkerke, NIKHEF
Verifying the QCD multijet background estimate

- Can **verify** performance of QCD multi-jet background estimation on 1,2-jet bins by looking at $m_T$ distributions
  - Here results for 1-jet bin shown (best statistics)

NB: $E_T^{\text{miss}}+m_T(w) > 60$ cut omitted from event selection to avoid suppression of events with low $m_T$
Estimation of W+jets background

- W+jets background in ≥4-jet tagged bin has largely the same particle-level final state as ttbar
  - Thus no easy/obvious handles to exploit to measure the rate from data in same bin (as was done with \( N_{\text{std}}, N_{\text{loose}} \) for QCD)
  - Sample dominated by W+HF+jets final states, and theory predictions not so reliable
  - Instead aim for hybrid data/MC method that exploits relations between jet bins

- Step 1 – Exploit that \( \sigma[W+(N+1)\text{-jet}]/\sigma[W+(N)\text{-jet}] \) is fairly constant (Berends-Giele scaling)
  - NB: Technique only works in pretag-samples
  - Concretely: Use pre-tag W+1,W+2 jet measurements to make a prediction of pre-tag W+≥4 jets rate

\[
W_{\text{pre-tag}}^{\geq 4\text{-jet}} = W_{\text{pre-tag}}^{2\text{-jet}} \cdot \sum_{n=2}^{\infty} (W_{\text{pre-tag}}^{2\text{-jet}}/W_{\text{pre-tag}}^{1\text{-jet}})^n
\]
Estimation of W+jets background

Measure W+jets rate in 1,2-jets pretag bins by

- subtracting data-driven QCD estimate,
- subtracting (small) other backgrounds using simulation estimates

<table>
<thead>
<tr>
<th></th>
<th>1-jet pre-tag $e$</th>
<th>1-jet pre-tag $\mu$</th>
<th>2-jet pre-tag $e$</th>
<th>2-jet pre-tag $\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed</td>
<td>1815</td>
<td>1593</td>
<td>404</td>
<td>370</td>
</tr>
<tr>
<td>QCD multijet (DD)</td>
<td>517 ± 89</td>
<td>65 ± 28</td>
<td>190 ± 43</td>
<td>20.0 ± 9.7</td>
</tr>
<tr>
<td>W(\tau+jets) (MC)</td>
<td>39 ± 10</td>
<td>43 ± 11</td>
<td>11.7 ± 4.4</td>
<td>13.6 ± 5.1</td>
</tr>
<tr>
<td>$Z+$jets (MC)</td>
<td>19.0 ± 9.1</td>
<td>48 ± 12</td>
<td>11.6 ± 5.2</td>
<td>14.0 ± 4.8</td>
</tr>
<tr>
<td>$t\bar{t}$ (MC)</td>
<td>1.7 ± 0.8</td>
<td>1.7 ± 0.8</td>
<td>7.0 ± 3.0</td>
<td>7.7 ± 3.3</td>
</tr>
<tr>
<td>$W$+jets (MC)</td>
<td>4.4 ± 0.7</td>
<td>5.0 ± 0.8</td>
<td>5.2 ± 0.8</td>
<td>5.1 ± 0.8</td>
</tr>
<tr>
<td>diboson (MC)</td>
<td>4.8 ± 4.8</td>
<td>5.7 ± 5.7</td>
<td>3.8 ± 3.8</td>
<td>4.4 ± 4.4</td>
</tr>
<tr>
<td>Total (non W(\nu)+jets)</td>
<td>585 ± 90</td>
<td>168 ± 33</td>
<td>229 ± 44</td>
<td>65 ± 13</td>
</tr>
<tr>
<td>Estimated W(\nu)+jets</td>
<td>1230 ± 100</td>
<td>1425 ± 52</td>
<td>175 ± 49</td>
<td>305 ± 23</td>
</tr>
</tbody>
</table>

\[ W^{\geq 4\text{-jet}}_{\text{pre-tag}} = W^{2\text{-jet}}_{\text{pre-tag}} \cdot \sum_{n=2}^{\infty} (W^{2\text{-jet}}_{\text{pre-tag}} / W^{1\text{-jet}}_{\text{pre-tag}})^n, \]

\[ W^{\geq 4\text{-jet}}_{\text{pre-tag}} = 11.2 \pm 2.2\text{(stat.)} \pm 4.0\text{(syst.)}, \quad e\text{ channel}, \]

\[ W^{\geq 4\text{-jet}}_{\text{pre-tag}} = 18.9 \pm 4.1\text{(stat.)} \pm 5.0\text{(syst.)}, \quad \mu\text{ channel.} \]
Estimation of tagged $W+jets$ background

$W^{\text{tagged} \geq 4\text{jet}} = W^{\text{pretag} \geq 4\text{jet}} \cdot f^{\geq 4-\text{jet}}_{\text{tagged}}$

- Step 2 – Estimate 4-jet tagged fraction
  - Can measure 2-jet tagged fraction
    
    $f_{\text{tagged}}(2\text{-jet}) = \frac{\int L = 2.9 \text{ pb}^{-1}}{f_{\text{tagged}}(2\text{-jet}) = 0.060 \pm 0.018(\text{stat.}) \pm 0.007(\text{syst.})}$
    
    - But flavour composition will be different in 2-jet and 4-jet bin → not directly applicable
      
      Apply simulation-based correction factor

    $f_{\text{tagged}}^{\geq 4\text{-jet}} = f_{\text{tagged}}^{2\text{-jet}} \cdot f_{\text{corr}}^{\geq 4}$

    $W^{\geq 4\text{-jet}}_{\text{tagged}} = 1.9 \pm 0.7(\text{stat.}) \pm 0.9(\text{syst.}), \ e \text{ channel},$
    
    $W^{\geq 4\text{-jet}}_{\text{tagged}} = 3.2 \pm 1.2(\text{stat.}) \pm 1.2(\text{syst.}), \ \mu \text{ channel}.$
Summary on backgrounds – ttbar signal yield

- Obtain ttbar event yield by subtracting all backgrounds from measured event yield
  - Errors on backgrounds include all systematic uncertainties
  - Correlations and anti-correlations between source of systematic uncertainties is taken into account in total background estimate

<table>
<thead>
<tr>
<th></th>
<th>e+≥4jets tagged</th>
<th>μ+≥4jets tagged</th>
</tr>
</thead>
<tbody>
<tr>
<td>QCD (DD)</td>
<td>4.8 ± 3.1</td>
<td>0.8 ± 0.5</td>
</tr>
<tr>
<td>W+jets (DD)</td>
<td>1.9 ± 1.1</td>
<td>3.2 ± 1.7</td>
</tr>
<tr>
<td>Z+jets (MC)</td>
<td>0.2 ± 0.1</td>
<td>0.1 ± 0.1</td>
</tr>
<tr>
<td>single top (MC)</td>
<td>0.7 ± 0.2</td>
<td>0.7 ± 0.2</td>
</tr>
<tr>
<td><strong>Total background</strong></td>
<td><strong>7.5 ± 3.1</strong></td>
<td><strong>4.7 ± 1.7</strong></td>
</tr>
<tr>
<td>Observed</td>
<td>17</td>
<td>20</td>
</tr>
<tr>
<td>Estimated ttbar</td>
<td>9.5 ± 4.1 ± 3.1</td>
<td>15.3 ± 4.4 ± 1.7</td>
</tr>
</tbody>
</table>
Cross section calculation in lepton+jets channel

- Final step is to relate event count to cross-section using the luminosity and \((\text{acceptance} \times \text{efficiency} \times \text{BF})\)

\[
\sigma_{tt} = \frac{(N_{\text{obs}} - N_{\text{bkg}})}{\varepsilon \times \text{BF} \times L}
\]

<table>
<thead>
<tr>
<th>Method</th>
<th>(e+\text{jets})</th>
<th>(\mu+\text{jets})</th>
<th>(e/\mu +\text{jets combined})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Counting (\sigma_{tt}) [pb]</td>
<td>(105 \pm 46^{+45}_{-40})</td>
<td>(168 \pm 49^{+46}_{-38})</td>
<td>(142 \pm 34^{+50}_{-31})</td>
</tr>
</tbody>
</table>

- Systematic uncertainty on cross section reflects
  - Background systematic uncertainties
  - Acceptance systematic uncertainties
  - Luminosity uncertainty

- In total 28 sources of systematic uncertainty are taken into account (in a coherent way across all components)

- Cross-section calculation based on likelihood model for event count in each channel, implementing the above relation
# Breakdown of systematic uncertainties

<table>
<thead>
<tr>
<th>Source</th>
<th>Relative cross-section uncertainty [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
<td>$e$+jets</td>
</tr>
<tr>
<td>Statistical uncertainty</td>
<td>$\pm 43$</td>
</tr>
<tr>
<td><strong>Object selection</strong></td>
<td></td>
</tr>
<tr>
<td>Lepton reconstruction, identification, trigger</td>
<td>$\pm 3$</td>
</tr>
<tr>
<td>Jet energy reconstruction</td>
<td>$\pm 13$</td>
</tr>
<tr>
<td>$b$-tagging</td>
<td>$-10 / +15$</td>
</tr>
<tr>
<td><strong>Background rates</strong></td>
<td></td>
</tr>
<tr>
<td>QCD normalisation</td>
<td>$\pm 30$</td>
</tr>
<tr>
<td>$W$+jets normalisation</td>
<td>$\pm 11$</td>
</tr>
<tr>
<td>Other backgrounds normalisation</td>
<td>$\pm 1$</td>
</tr>
<tr>
<td><strong>Signal simulation</strong></td>
<td></td>
</tr>
<tr>
<td>Initial/final state radiation</td>
<td>$-6 / +13$</td>
</tr>
<tr>
<td>Parton distribution functions</td>
<td>$\pm 2$</td>
</tr>
<tr>
<td>Parton shower and hadronisation</td>
<td>$\pm 1$</td>
</tr>
<tr>
<td>Next-to-leading-order generator</td>
<td>$\pm 4$</td>
</tr>
<tr>
<td>Integrated luminosity</td>
<td>$-11 / +14$</td>
</tr>
<tr>
<td><strong>Total systematic uncertainty</strong></td>
<td>$-38 / +43$</td>
</tr>
<tr>
<td>Statistical + systematic uncertainty</td>
<td>$-58 / +61$</td>
</tr>
</tbody>
</table>

Jet Energy scale $\sim 7\%$

Wouter Verkerke, NIKHEF
Measurement of $b$-tagging efficiency

- Efficiency of $b$-tagging for $b$-jets measured from data using sample of jets with soft muon close to jet axis

Wouter Verkerke, NIKHEF

- Soft muon inside jet (from s.l. $b$ decay)
- Secondary vertex
Measurement of b-tagging efficiency in data

Distribution of $p_T$ of muon relative to jet axis
templates for b,c,light jets

pretag

<table>
<thead>
<tr>
<th>Arbitrary Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.14</td>
</tr>
</tbody>
</table>

$0$ | $0.5$ | $1.0$ | $1.5$ | $2.0$ | $2.5$ |

tagged

<table>
<thead>
<tr>
<th>Arbitrary Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.14</td>
</tr>
</tbody>
</table>

$0$ | $0.5$ | $1.0$ | $1.5$ | $2.0$ | $2.5$ |

Fit sum of b-jet and non-b-jet to data

Efficiency measured vs $p_T, \eta$

for data and simulation

Scale factor (with uncertainty)
is applied to simulation
Kinematic distributions – reconstruction of hadronic top mass

- Mass of top quark can be reconstructed from three-jet invariant mass
  - Empirically identify three jets by picking combination with largest vectorially summed $p_T$
  - Observed distribution agrees with expectation from simulation $\rightarrow$ data consistent with hypothesis that observed excess is $ttbar$
The dilepton channel – event selection

- Three channels reconstructed: \( ee, \mu\mu, e\mu \)
- Event selection
  - Exactly two good opposite charge leptons \((p_T>20 \text{ GeV})\)
  - At least 2 good jets \((p_T>20, |\eta|<2.5)\)
  - \( ee \): \(|m(ll)-m(Z)|>5 \text{ GeV} \)
  - \( \mu\mu \): \(|m(ll)-m(Z)|>10 \text{ GeV} \) (Z veto)
  - \( ee \): \(E_T^{\text{miss}}> 40 \text{ GeV} \)
  - \( \mu\mu \): \(E_T^{\text{miss}}> 30 \text{ GeV} \)
  - \( e\mu \): \(H_T> 150 \text{ GeV} \)
  - Cosmic veto: \( \mu\mu \): \(d0(\mu)<500 \mu m \)

\[ H_T = \text{scalar sum of } E_T \text{ of leptons and selected jets} \]
Background estimation

• Next step is background estimation in ≥2 jets region
  – Backgrounds are generally small (note log-scale on plots)

• Background from Z+(bb/cc/c)+jets not surprising
  – Drell-Yan process also generates lepton pairs off Z-resonance
  – Large cross section and same final particle-level state except for neutrinos (but fake ETmiss can easily be introduced)

• Understanding fake lepton background
  – Mechanism similar to lepton+jets background, but source different
  – Leading contribution from W+jets with one (additional) fake/non-prompt lepton
  – Also small contribution from QCD multi-jets with two fake/non-prompt leptons
Estimation of fake background

- Same principle as \( \mu + \text{jets} \): matrix method
  - Define loose, tight lepton selection \( \rightarrow \) 4 observable states: Loose-Loose, Loose-Tight, Tight-Loose and Tight-Tight
  - Also 4 classes of background
    Fake-Fake, Fake-Real, Real-Fake and Real-Real
  - Matrix method results in a 4x4 equation system

\[
\begin{bmatrix}
rr & rf & fr & ff \\
r(1-r) & r(1-f) & f(1-r) & f(1-f) \\
(1-r)r & (1-r)f & (1-f)r & (1-f)f \\
(1-r)(1-r) & (1-r)(1-f) & (1-f)(1-r) & (1-f)(1-f)
\end{bmatrix}
\]

- Obtain fractions \( f, r \) from control samples and solve equation for \( N_{RF/FR}, N_{FF} \)
- Fraction \( r \) measured from \( Z(\rightarrow \ell\ell) + \text{jets} \) events
- Fraction \( f \) measured from sample with single loose lepton
  (which is dominated by QCD di-jet production – contribution from \( W + \text{jets} \) with real leptons subtracted using simulation prediction)
Estimation of Z+jets background

- Most Drell-Yan background at low missing $E_T$ and close to Z mass
- In 3 pb$^{-1}$ insufficient data for a real data-driven background estimation
  - Instead rescale simulation prediction for signal region with data/simul ratio in normalization region

<table>
<thead>
<tr>
<th></th>
<th>$e^e$</th>
<th>$\mu\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z+jets (Monte-Carlo)</td>
<td>$0.14 \pm 0.03 \pm 0.16$</td>
<td>$0.56 \pm 0.06 \pm 0.39$</td>
</tr>
</tbody>
</table>
Summary on backgrounds – ttbar signal yield

- Obtain ttbar event yield by subtracting all backgrounds from measured event yield
  - Errors on backgrounds in include all systematic uncertainties
  - Correlations and anti-correlations between source of systematic uncertainties taken into account in total background estimate

<table>
<thead>
<tr>
<th>Source</th>
<th>ee</th>
<th>μμ</th>
<th>eμ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z+jets (DD)</td>
<td>0.25 ± 0.18</td>
<td>0.67 ± 0.38</td>
<td>-</td>
</tr>
<tr>
<td>fake leptons (DD)</td>
<td>0.16 ± 0.18</td>
<td>-0.08 ± 0.07</td>
<td>0.47 ± 0.28</td>
</tr>
<tr>
<td>Z(ττ)+jets (MC)</td>
<td>0.08 ± 0.04</td>
<td>0.14 ± 0.07</td>
<td>0.47 ± 0.28</td>
</tr>
<tr>
<td>single top (MC)</td>
<td>0.08 ± 0.02</td>
<td>0.07 ± 0.03</td>
<td>0.22 ± 0.04</td>
</tr>
<tr>
<td>dibosons (MC)</td>
<td>0.04 ± 0.02</td>
<td>0.07 ± 0.03</td>
<td>0.15 ± 0.05</td>
</tr>
<tr>
<td><strong>Total background</strong></td>
<td><strong>0.60 ± 0.27</strong></td>
<td><strong>0.88 ± 0.40</strong></td>
<td><strong>0.97 ± 0.30</strong></td>
</tr>
<tr>
<td>Observed</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

Wouter Verkerke, NIKHEF
Cross section calculation in dilepton channel

- Final step is to divide by signal acceptance x efficiency x BF and luminosity to obtain cross section

\[ \sigma = \frac{(N_{\text{obs}} - N_{\text{bkg}})}{(\epsilon \cdot BF \cdot L)} \]

- Systematic uncertainty on cross section reflects
  - Background systematic uncertainties
  - Acceptance systematic uncertainties
  - Luminosity uncertainty

- In total 32 sources of systematic uncertainty are taken into account (in a coherent way across all components)

- Final cross-section calculation based on likelihood model for event count in each channel, implement the above relation

<table>
<thead>
<tr>
<th>Channel</th>
<th>( \sigma_{\text{ff}} ) [pb]</th>
</tr>
</thead>
<tbody>
<tr>
<td>( ee )</td>
<td>( 193 \pm 243 \pm 84 ) ( -152 -48 )</td>
</tr>
<tr>
<td>( \mu\mu )</td>
<td>( 185 \pm 184 \pm 56 ) ( -124 -47 )</td>
</tr>
<tr>
<td>( e\mu )</td>
<td>( 129 \pm 100 \pm 32 ) ( -72 -18 )</td>
</tr>
<tr>
<td>Combined</td>
<td>( 151 \pm 78 \pm 37 ) ( -62 -24 )</td>
</tr>
</tbody>
</table>

Wouter Verkerke, NIKHEF
# Breakdown of systematic uncertainties

<table>
<thead>
<tr>
<th>Source</th>
<th>Relative cross-section uncertainty [%]</th>
<th>(ee)</th>
<th>(\mu\mu)</th>
<th>(e\mu)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistical uncertainty</td>
<td></td>
<td>-79 / +126</td>
<td>-67 / +100</td>
<td>-56 / +77</td>
</tr>
<tr>
<td>Object selection</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lepton reconstruction, identification, trigger</td>
<td></td>
<td>-2 / +11</td>
<td>-4 / +3</td>
<td>-1 / +3</td>
</tr>
<tr>
<td>Jet energy reconstruction</td>
<td></td>
<td>-7 / +13</td>
<td>-14 / +9</td>
<td>-3 / +5</td>
</tr>
<tr>
<td>Background rates</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fake leptons</td>
<td></td>
<td>-31 / +24</td>
<td>-4 / +1</td>
<td>-15 / +8</td>
</tr>
<tr>
<td>(Z+\text{jets})</td>
<td></td>
<td>-12 / +4</td>
<td>-19 / +5</td>
<td>-2 / +1</td>
</tr>
<tr>
<td>Monte-Carlo simulation statistics</td>
<td></td>
<td>-5 / +3</td>
<td>-3 / +4</td>
<td>± 2</td>
</tr>
<tr>
<td>Theoretical cross-sections</td>
<td></td>
<td>± 3</td>
<td>-5 / +4</td>
<td>± 3</td>
</tr>
<tr>
<td>Signal simulation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial/final state radiation</td>
<td></td>
<td>-4 / +5</td>
<td>-2 / +3</td>
<td>-2 / +3</td>
</tr>
<tr>
<td>Parton distribution functions</td>
<td></td>
<td>-2 / +1</td>
<td>-2 / +3</td>
<td>-2 / +3</td>
</tr>
<tr>
<td>Parton shower and hadronisation</td>
<td></td>
<td>-9 / +14</td>
<td>-6 / +9</td>
<td>± 3</td>
</tr>
<tr>
<td>Next-to-leading order generator</td>
<td></td>
<td>-8 / +11</td>
<td>-11 / +13</td>
<td>-3 / +4</td>
</tr>
<tr>
<td>Integrated luminosity</td>
<td></td>
<td>-11 / +16</td>
<td>-11 / +16</td>
<td>-12 / +14</td>
</tr>
<tr>
<td>Total systematic uncertainty</td>
<td></td>
<td>-25 / +44</td>
<td>-25 / +30</td>
<td>-14 / +25</td>
</tr>
<tr>
<td>Statistical + systematic uncertainty</td>
<td></td>
<td>-83 / +134</td>
<td>-72 / +104</td>
<td>-57 / +81</td>
</tr>
</tbody>
</table>

Jet Energy scale \(\sim 7\%\)

ATLAS Preliminary

\[
\int L = 2.9 \text{ pb}^{-1}
\]

**all channels**
- data
- \(t\bar{t}\)
- single top
- \(Z + \text{jets}\)
- diboson
- fake leptons
- uncertainty

Number of jets
Likelihood model for cross section calculation

- Likelihood for each channel

\[ N^{\text{exp}}(\sigma_{t\bar{t}}, \alpha_j) = L \cdot \epsilon_{t\bar{t}}(\alpha_j) \cdot \sigma_{t\bar{t}} + \sum_{bkg} L \cdot \epsilon_{bkg}(\alpha_j) \cdot \sigma_{bkg}(\alpha_j) + N_{DD}(\alpha_j) \]

\[ L(\sigma_{t\bar{t}}, L, \alpha_j) = \text{Poisson} \left( N^{\text{obs}} \mid N^{\exp}(\sigma_{t\bar{t}}, \alpha_j) \right) \times \text{Gauss}(L_0 \mid L, \delta_L) \times \prod_{j \in \text{syst}} \Gamma_j(\alpha_j). \]

- Profile likelihood fit returns interval on \( \sigma(\text{tt})/\sigma(\text{tt})_{\text{SM}} \)
  - Run with and without systematic uncertainty terms (lumi, other) to extract statistical and total uncertainty
Counting b-tagged jets in dilepton top selection

- B-tagging was not used in dilepton event selection.
  - But expect 2 b-quarks for each top candidate, while (on average) much less than 1 b-jet in background
  - Presence of b-jets in selection support hypothesis that selected candidates contain $t\bar{t}$ events
A 'platinum' event: ttbar $e\mu$ with 2 b-tagged jets

double b-tag
Putting it all together

- Combined lepton+jets and dilepton from joint fit
  - Accounts for all (anti)-correlated systematic uncertainties

<table>
<thead>
<tr>
<th></th>
<th>Cross-section [pb]</th>
<th>Signal significance [$\sigma$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single lepton channels</td>
<td>$142 \pm 34^{+50}_{-31}$</td>
<td>4.0</td>
</tr>
<tr>
<td>Dilepton channels</td>
<td>$151^{+78}<em>{-62}^{+37}</em>{-24}$</td>
<td>2.8</td>
</tr>
<tr>
<td>All channels</td>
<td>$145 \pm 31^{+42}_{-27}$</td>
<td>4.8</td>
</tr>
</tbody>
</table>
Summary & Outlook

- The era of LHC top physics has started
  - First ATLAS measurements of top quark pair production with 3 pb$^{-1}$
  - Agrees with theoretical prediction and CMS measurement

- And (at least) another factor 10x more expected in 2011
- Systematic uncertainties will start to dominate total uncertainties, will enter regime where we can test QCD predictions of cross section
- Other area of top physics (properties, single top) will soon become feasible

Full 2010 dataset / expectations scaled from 2.9 pb$^{-1}$