LEPTONIC FAKE RATES FROM JETS AT ATLAS

Gemma Wooden
University of Oxford for the ATLAS collaboration

Berkeley workshop on physics opportunities with early LHC data
6th – 8th May 2009
Overview

- Motivation
- Expected electron fake rates from simulation:
  - Cut-based method
  - Multivariate techniques
- Expected muon fake rates from simulation
- Background estimates using data-driven method:
  - In \( W \rightarrow ev \) events
- Conclusions
Motivation

- Important to measure leptonic fake rates since they will affect many interesting physics results:
  - **Electroweak measurements**
    - Top & W mass
  - **Diboson studies**
    - Anomalous triple gauge couplings would enhance production cross-sections
    - Could find evidence for new physics
  - **Higgs searches**
    - Fake leptons will be an important background to $H \rightarrow WW$ and $H \rightarrow ZZ$ channels
Sources of background to leptons

- Electrons are split into 4 categories:
  - Isolated – e from Z, W, t, \( \tau \) or \( \mu \)
  - Non-isolated – e from \( J/\psi, b- \) or c-hadron decays
  - Background – e from photon conversions, \( \pi^0/e \) Dalitz decays or \( u/d/s \)-hadron decays
  - Non-electron – charged hadrons, muons

- Muons are categorised into 2 types:
  - Direct – no quarks in their ancestry
  - Indirect – ancestry includes a heavy quark but not a tau

- There may also be fake muons where charged hadrons are falsely reconstructed as muons

- At design luminosity, there will also be a significant background in the muon chambers from low-energy photons and neutrons (cavern background)
Introduction to cut-based method for electrons

- The PYTHIA event generator was used to generate a large jet sample (8.2 million events)
- This sample, referred to as filtered di-jets, contains all hard-scattering QCD processes with $E_T > 15$ GeV as well as other interesting physics processes, e.g. single W/Z production
- Three standard sets of ATLAS selection cuts for electrons have been used: loose, medium and tight, with cuts on various discriminating variables
- The plot below illustrates one such variable

**Ratio of $E_T$ of electron candidate to sum of $E_T$ and $E_T$ in first layer of hadronic calorimeter**

- Solid line: $Z \rightarrow ee$ sample
- Dotted line: Filtered di-jets sample
Multivariate techniques & comparison with cut-based method

- Several multivariate methods have been investigated as well as the simple cut-based method
- The jet rejection and efficiency has been calculated for one of these, a likelihood discriminant
- The efficiency achieved with the likelihood discriminant method with the rejection fixed at that achieved by the cut-based method for a given electron selection (& vice versa for the rejection) allows a comparison:

<table>
<thead>
<tr>
<th>Cuts</th>
<th>Cut-based method</th>
<th>Likelihood method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Efficiency $\varepsilon_e$ (%)</td>
<td>Rejection $R_j$</td>
</tr>
<tr>
<td>Loose</td>
<td>87.97 ± 0.05</td>
<td>567 ± 1</td>
</tr>
<tr>
<td>Medium</td>
<td>77.29 ± 0.06</td>
<td>2184 ± 7</td>
</tr>
<tr>
<td>Tight (isol)</td>
<td>64.22 ± 0.07</td>
<td>(9.9 ± 0.2) × 10^4</td>
</tr>
<tr>
<td>Tight (TRT)</td>
<td>61.66 ± 0.07</td>
<td>(8.9 ± 0.2) × 10^4</td>
</tr>
</tbody>
</table>

This gives a fake rate of the order $10^{-4}$ or better

This table shows that multivariate methods may improve jet rejection BUT they are only suitable once the detector is well understood – cut-based method will be used for first data!
An inclusive tt\overline{t} sample with one lepton in the final state was produced using the generator MC@NLO.

Also used PYTHIA to generate low- and high-P_T muon samples.

Used standard ATLAS combined muon algorithm (Staco), which statistically combines muon spectrometer tracks with inner detector tracks.

The table shows the efficiency for all found muons (found) & for those which match a true muon (good) as well as the fake rates.

This gives a fake rate of the order 10^{-3} or even better in the high P_T region.

It’s likely that isolation requirements would improve this.
Introduction to data-based background estimation

- Fake electrons from jets provide an important background to W in electron channel:

<table>
<thead>
<tr>
<th>Selection</th>
<th>$W \rightarrow ev$</th>
<th>jets</th>
<th>$W \rightarrow \tau v$</th>
<th>$Z \rightarrow ee$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trigger</td>
<td>37.01±0.09</td>
<td>835±18</td>
<td>1.73±0.02</td>
<td>6.07±0.01</td>
</tr>
<tr>
<td>$E_T &gt; 25$ GeV, $</td>
<td>\eta</td>
<td>&lt; 2.4$</td>
<td>30.84±0.09</td>
<td>383±12</td>
</tr>
<tr>
<td>Electron ID</td>
<td>26.77±0.09</td>
<td>110±6</td>
<td>0.91±0.01</td>
<td>2.95±0.01</td>
</tr>
<tr>
<td>$E_T &gt; 25$ GeV</td>
<td>22.06±0.09</td>
<td>4.6±0.7</td>
<td>0.55±0.01</td>
<td>0.06±0.01</td>
</tr>
<tr>
<td>$M_T &gt; 40$ GeV</td>
<td>21.71±0.08</td>
<td>1.5±0.4</td>
<td>0.43±0.01</td>
<td>0.04±0.01</td>
</tr>
</tbody>
</table>

- Jet production cross-section and fragmentation properties aren’t well known at the LHC
- This introduces significant uncertainty on jet background => can’t rely on Monte Carlo
- Data-driven method desirable
- Idea: Measure normalisation & shape of jet background before missing $E_T$ cut
  - Find signal-free jet sample with $M_T$ distribution & jet multiplicity similar to that of jet background to $W \rightarrow ev$
  - Use this sample to measure rejection of missing $E_T$ cut to estimate jet background in W event selection

Number of events $\times 10^4$ for an integrated luminosity of 50 pb$^{-1}$
W and jet event selections

- Search for events with signature suggestive of W signal:
  - Look for an EM object with a matched track passing a single electron trigger
- Signal sample:
  - Select events with one electron candidate which satisfies medium selection criteria and passes $E_T$ and $\eta$ cuts described on the previous slide
  - Remove events with a second high $P_T$ electromagnetic cluster which, with the first electron, gives an invariant mass close to Z peak
- Jet background control sample:
  - Use single photon trigger with $E_T > 20$ GeV
  - Look for photon cluster satisfying same kinematic & calorimetric cuts as electron candidate
  - To reject events with true electrons, require that there is no inner detector track matching cluster
Simulation indicates that the missing $E_T$ distribution of jet background events is identical (within statistics) to that of the jet background control sample.

Ratio of missing $E_T$ in jet background events to control sample as a function of missing $E_T$ for an integrated luminosity of 50 pb$^{-1}$.
W event results

- Comparison indicates that the estimated jet background from the control sample agrees well with the actual background to W from jet events:

- The estimated background may then be subtracted from the signal sample
- This data-driven method gives a jet background fraction of (0+4-o)% with an uncertainty of $\Delta B = 0.92 \times 10^4$ events
- => significant reduction!
- Further cuts requiring large missing $E_T$ & large $M_T$ may then be applied in order to further select the W signal and remove other sources of background

Comparison of jet background (points) to fitted background for an integrated luminosity of 50 pb$^{-1}$
Conclusions

- Lepton fakes will be a significant contribution to backgrounds to some of the most interesting physics that the LHC hopes to investigate.
- Therefore very important to understand and be able to reduce.
- Results from Monte Carlo suggest fake rates to be of the order $10^{-4}$ for electrons and $10^{-3}$ for muons.
- Can use data-driven techniques to understand fake rates as well as multivariate techniques once detector is well understood.
- Data-driven methods have been investigated in the context of a $W$ cross-section measurement & initial studies indicate that they accurately predict background from jets faking leptons in simulation.
- More work to do – will be able to start trying to understand & characterise jet events with first data.
Backup slides
Samples used for data-driven method

- W sample generated using PYTHIA
- Diboson & ttbar samples generated with MC@NLO

<table>
<thead>
<tr>
<th>Channel</th>
<th>$\sigma \times B_T$</th>
<th>$\epsilon_{filter}$</th>
<th>$N_{ev}$ ($\times 10^3$)</th>
<th>$L$ (pb$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W \rightarrow ev$</td>
<td>20510 pb</td>
<td>0.63</td>
<td>140</td>
<td>11</td>
</tr>
<tr>
<td>$\gamma/Z \rightarrow ee, \sqrt{s} &gt; 60$ GeV</td>
<td>2015 pb</td>
<td>0.86</td>
<td>399</td>
<td>230</td>
</tr>
<tr>
<td>$\gamma/Z \rightarrow ee, \sqrt{s} &lt; 60$ GeV</td>
<td>9220 pb</td>
<td>0.022</td>
<td>197</td>
<td>969</td>
</tr>
<tr>
<td>$W \rightarrow \tau \nu_\tau$</td>
<td>20510 pb</td>
<td>0.20</td>
<td>32</td>
<td>8</td>
</tr>
<tr>
<td>$Z \rightarrow \tau \tau$</td>
<td>2015 pb</td>
<td>0.05</td>
<td>13</td>
<td>129</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>833 pb</td>
<td>0.54</td>
<td>382</td>
<td>850</td>
</tr>
<tr>
<td>Inclusive jets ($p_T &gt; 6$ GeV)</td>
<td>70 mb</td>
<td>0.058</td>
<td>2480</td>
<td>0.0006</td>
</tr>
<tr>
<td>Inclusive jets ($p_T &gt; 17$ GeV)</td>
<td>2333 $\mu$b</td>
<td>0.09</td>
<td>3725</td>
<td>0.02</td>
</tr>
<tr>
<td>$WW \rightarrow (ev)(ev)$</td>
<td>1.275 pb</td>
<td>1.</td>
<td>20</td>
<td>15608</td>
</tr>
<tr>
<td>ZZ</td>
<td>14.8 pb</td>
<td>1.</td>
<td>43</td>
<td>2922</td>
</tr>
<tr>
<td>WZ</td>
<td>29.4 pb</td>
<td>1.</td>
<td>50</td>
<td>1699</td>
</tr>
</tbody>
</table>
Contribution and origin of electron candidates before selection cuts

- From the di-jets sample:

<table>
<thead>
<tr>
<th>Isolated</th>
<th>Non-isolated</th>
<th>Background</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W - 75.0%$</td>
<td>$b$-hadrons - $38.7%$</td>
<td>$\gamma$-conv. - $97.8%$</td>
</tr>
<tr>
<td>$Z - 20.9%$</td>
<td>$c$-hadrons - $60.6%$</td>
<td>Dalitz decays - $1.8%$</td>
</tr>
<tr>
<td>$t - &lt; 0.1%$</td>
<td>$J/\psi - 0.7%$</td>
<td>$u/d/s$-hadrons - $0.4%$</td>
</tr>
<tr>
<td>$\tau - 4.1%$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
ATLAS electron & muon selection criteria

- Electrons must have $|\eta| < 1.37$ or $1.52 < |\eta| < 2.4$ to avoid crack regions in the detector.
- There are three standard sets of identification criteria:
  - Loose: simple shower-shape cuts
  - Medium: further cuts on shower-shape & track quality
  - Tight: tighter cut on track-matching as well as requirements on hits in different regions of the inner detector
  - Tight (isol) applies an additional cut on the isolation energy of the electron.
  - Tight (TRT) doesn’t include any explicit cut on isolation energy but applies tighter cuts on the TRT (Transition Radiation Tracker) information.
- In $W$ events, require one electron with $P_T > 25$ GeV/c satisfying medium criteria.
- Muons must have $|\eta| < 2.5$.
- Low $P_T$ muon sample requires two muons with $P_T > 25$ GeV/c.
- High $P_T$ muon sample has an average muon $P_T = 500$ GeV/c, with a $Z'$ mass of 2 TeV.