Yield extraction of a 4th generation signal with the ATLAS detector

Malik Aliev, Sergio Grancagnolo, Heiko Lacker, Rocco Mandrysch, Dennis Wendland

Humboldt-Universität zu Berlin

GK blockcourse
Rathen, 23.03.2010
1 Introduction to a 4th family of fermions
   - Motivation
   - Search signature
   - Particle and event selection

2 Counting analysis
   - Idea
   - Control samples
   - Estimation of control ratios and signal yields

3 Results for b’ (300 GeV)
   - Mixing sample
   - Signal and control regions
   - Result
   - Convergences

4 Summary
Introduction to a 4th family of fermions

- Motivation
- Search signature
- Particle and event selection

Counting analysis

- Idea
- Control samples
- Estimation of control ratios and signal yields

Results for b’ (300 GeV)

- Mixing sample
- Signal and control regions
- Result
- Convergences

Summary
**Motivation**

### Present situation:
- Number of families not fixed by the Standard Model
- Z-resonance (LEP I): $N_\nu = 3$ with $m_\nu < m_Z/2$
- If $\nu_4$ exists: $m_{\nu_4} \gtrsim m_Z/2$

### With a fourth generation:
- Possible way to explain baryogenesis
- Higher Higgs mass (up to 600GeV) possible
  (Kribs et al, Phys.Rev.D76:075016,2007)
Direct limits

Current mass limits on short living particles @ 95% CL:

<table>
<thead>
<tr>
<th>Particle</th>
<th>$m_{min}$ [GeV]</th>
<th>Experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\ell_4$</td>
<td>100</td>
<td>LEP</td>
</tr>
<tr>
<td>$\nu_4$ (Dirac)</td>
<td>90.3</td>
<td>LEP</td>
</tr>
<tr>
<td>$\nu_4$ (Majorana)</td>
<td>80.5</td>
<td>LEP</td>
</tr>
<tr>
<td>$t'$</td>
<td>311</td>
<td>Tevatron</td>
</tr>
<tr>
<td>$b'$</td>
<td>325</td>
<td>Tevatron</td>
</tr>
</tbody>
</table>

Consequence for us:

- Looking for 4th generation quarks with $m \geq 300\text{GeV}$

Tiny mixing angles between 4th and other families

⇒ particles could have long lifetime:

- Mass limit depends on CKM elements and mass of other heavy quark
Introduction to a 4th family of fermions

Assumption: $m(b') - m(t) > m(W)$, $V_{tb'}$ sufficiently large wrt $V_{ub'}$ and $V_{cb'}$

**Final states of $t'$ and $b'$ decay with sizable branching fraction:**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m(t') &gt; m(b')$</td>
<td>$m(t') - m(b') &gt; m(W)$</td>
</tr>
<tr>
<td></td>
<td>$b'\bar{b'} \rightarrow t\bar{t} + 2W \rightarrow b\bar{b}2W^+2W^-$</td>
</tr>
<tr>
<td></td>
<td>$t'\bar{t}' \rightarrow b'\bar{b'} + 2W \rightarrow b\bar{b}3W^+3W^-$</td>
</tr>
<tr>
<td>$m(t') - m(b') &lt; m(W)$</td>
<td>$b'\bar{b'} \rightarrow tW^-\bar{t}W^+ \rightarrow b\bar{b}2W^+2W^-$</td>
</tr>
<tr>
<td></td>
<td>$t'\bar{t}' \rightarrow bW^+\bar{b}W^-$</td>
</tr>
<tr>
<td>$m(b') &gt; m(t')$</td>
<td>$m(b') - m(t') &gt; m(W)$</td>
</tr>
<tr>
<td></td>
<td>$b'\bar{b'} \rightarrow t'W^-\bar{t}'W^+ \rightarrow b\bar{b}2W^+2W^-$</td>
</tr>
<tr>
<td></td>
<td>$t'\bar{t}' \rightarrow bW^+\bar{b}W^-$</td>
</tr>
<tr>
<td>$m(b') - m(t') &lt; m(W)$</td>
<td>$b'\bar{b'} \rightarrow tW^-\bar{t}W^+ \rightarrow b\bar{b}2W^+2W^-$</td>
</tr>
<tr>
<td></td>
<td>$t'\bar{t}' \rightarrow bW^+\bar{b}W^-$</td>
</tr>
</tbody>
</table>
Search scenario

Decay channels:

- $b' \bar{b}' \rightarrow t\bar{t} + W^+ W^- \rightarrow b\bar{b} + 2W^+ 2W^-$
- $t' \bar{t}' \rightarrow b'\bar{b}' + W^+ W^- \rightarrow t\bar{t} + 2W^+ 2W^- \rightarrow b\bar{b} + 3W^+ 3W^-$

⇒ Exclusive reconstruction difficult but inclusive multilepton final state interesting

⇒ $b'$ final states: $\ell^\pm \ell^\pm / 3\ell + \text{jets}$

⇒ $t'$ final states: $\ell^\pm \ell^\pm / 3\ell + \text{jets}$

⇒ rare standard model signatures

- signature also used in "Search for A Fourth Generation b' Quark in tW Final State at CMS in pp Collisions at $\sqrt{s} = 10$ TeV"
  (CMS PAS EXO-09-012)
# Lepton selection

## Leptons

| particle | $p_T^{min}$ [GeV] | $|\eta|^{max}$ | $E_T^{cone20}$ [GeV] |
|----------|------------------|----------------|---------------------|
| $e^\pm$  | 20               | < 2.5          | < 8                 |
| $\mu^\pm$| 20               | < 2.5          | < 10                |

- $E_T^{cone20}$: Energy sum in a cone $R = \sqrt{\Delta \eta^2 + \Delta \phi^2} = 0.2$
- Muons: $\chi^2/NDof < 5$ for matching between tracks in Muon Spectrometer and Inner Detector
Jet selection

Jets

- Cone jet algorithm with $\Delta R < 0.4$
- Remove misidentified jets in a cone $\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2} < 0.2$
  and $|\eta| < 2.5$ with $e^\pm$, $\mu^\pm$, $\gamma$ and $\tau^\pm$
- $E_T^{\text{jet}} > 25\text{GeV}$ in order to suppress jets from underlying event
Event selection

Missing energy:

- Use for calculation:
  - Jets
  - Muons
  - Correction function for dead material (cryostat)
  - Calorimeter cells in $e^{\pm}$, $\tau^{\pm}$, $\gamma$
- $\slashed{E}_T > 20\text{GeV}$

Leptons and jets

- Electrons and muons: $P_T^{\text{leading}} > 35\text{GeV}$
- Jets: $E_T^{\text{leading}} > 85\text{GeV}$
1 Introduction to a 4th family of fermions
   ● Motivation
   ● Search signature
   ● Particle and event selection

2 Counting analysis
   ● Idea
   ● Control samples
   ● Estimation of control ratios and signal yields

3 Results for b' (300 GeV)
   ● Mixing sample
   ● Signal and control regions
   ● Result
   ● Convergences

4 Summary
Motivation

⇒ How to reliably estimate the background on data?
**Idea**

**Procedure**
- Long term: Likelihood fit in $N_{Jets}$
- This talk: Counting analysis ($N_{Jets}$ only used as control distribution)
- Two issues:
  1. Control of background cross sections (included in this talk)
  2. Lepton fake rates (long term)

**Test of procedure using MC pseudo-data**
- Mixing sample is used as pseudo-data
- Selecting randomly events from each MC background and signal sample corresponding to specific luminosity (here: $50\, pb^{-1}$) and assuming specific cross section ($N = \sigma \cdot L$)
- Using the remaining events as the MC prediction
Subtracting background

- Counting of observed entries on data in specific signal region (SR) and subtraction of MC background
- **Signal region**: Lepton category (same sign dileptons)

\[ N_{\text{Sig}} = N_{\text{Data}} - N_{\text{MC}} \]
Lepton categories and dominating background

CR for $W \rightarrow e\nu$

CR for $W \rightarrow \mu\nu$

CR for $t\bar{t}$

CR for $Z \rightarrow ee$

CR for $Z \rightarrow \mu\mu$
Controlling background

- Control main backgrounds in control regions (CR)
- Same event cuts as SR but different lepton categories
- Modifying MC predictions by control ratios

Control region for $t\bar{t}$

Pseudo-data: $\sigma_{t\bar{t}} \rightarrow k_{\sigma} \cdot \sigma_{t\bar{t}}$ (here: $k_{\sigma} = 3$)
Estimation of control ratios and signal yields

Calculation - Iteration 0

\[ N_{Sig}^0 = N_{SR, data} - \sum_i \left( N_{SR, MC, i} \cdot R_{i}^0 \right) - \sum_{i_{other}} N_{SR, MC, i_{other}} \]

\[ R_{i}^0 = \frac{N_{CR, i, data}}{N_{CR, i, MC, i}} \]

- \( N_{MC(data)} \): Number of entries in \( N_{Jets} \) histogram of MC (data) sample
- \( \sum_i \): Sum over data-driven corrected BG samples
- \( \sum_{i_{other}} \): Sum over other, uncorrected BG
- \( R_{i}^0 \): Correction factor for sample \( i \), also calculated for signal in SR
Estimation of control ratios and signal yields

Calculation - Iteration k

\[ R_{i}^{0} = \frac{N_{\text{data}}^{CR,i}}{N_{MC,i}^{CR,i}} \quad \Rightarrow \quad R_{i}^{k} = \frac{N_{\text{data}}^{CR,i} - \sum_{m \neq i} R_{m}^{k-1} \cdot N_{MC,m}^{CR,i}}{N_{MC,i}^{CR,i}} \]

- \( \sum_{m} \): Sum over all BG (corrected and uncorrected) and MC signal sample
- \( R_{m}^{k-1} \): Correction factor of previous iteration for sample \( m \)
  (= 1 for uncorrected BG), also calculated for signal in SR
1 Introduction to a 4th family of fermions
   • Motivation
   • Search signature
   • Particle and event selection

2 Counting analysis
   • Idea
   • Control samples
   • Estimation of control ratios and signal yields

3 Results for $b'$ (300 GeV)
   • Mixing sample
   • Signal and control regions
   • Result
   • Convergences

4 Summary
Mixing sample

Modified cross sections $\sigma$ in mixing sample generation by factor $k_\sigma$

$$\sigma \rightarrow k_\sigma \cdot \sigma$$

<table>
<thead>
<tr>
<th>Modified sample</th>
<th>Modification factor $k_\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b'$ (300 GeV)</td>
<td>1.15</td>
</tr>
<tr>
<td>$W\nu \nu + \text{jets (MC)}$</td>
<td>0.9</td>
</tr>
<tr>
<td>$W\mu \nu + \text{jets (MC)}$</td>
<td>0.9</td>
</tr>
<tr>
<td>$Z\nu \nu + \text{jets (MC)}$</td>
<td>0.9</td>
</tr>
<tr>
<td>$Z\mu \mu + \text{jets (MC)}$</td>
<td>0.9</td>
</tr>
<tr>
<td>$t\bar{t} + \text{jets (MC)}$</td>
<td>1.15</td>
</tr>
<tr>
<td>Dibosons+jets</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Expected result from MC signal sample

Expecting $21.6 \pm 0.3$ signal events
Signal and control regions before correction procedure

Results for $b'$ (300 GeV)

Yield extraction of a 4th generation signal

Dennis Wendland (HU-Berlin)
Results for $b'$ (300 GeV)

Signal and control regions after correction procedure

Dennis Wendland (HU-Berlin)

Yield extraction of a 4th generation signal

23.03.2010 21 / 25
## Results

<table>
<thead>
<tr>
<th>Sample</th>
<th>Events</th>
<th>Final $R_i^k$</th>
<th>$k_\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>data</td>
<td>34 ± 6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$W\nu\bar{\nu} + \text{jets (MC)}$</td>
<td>0.6 ± 0.2</td>
<td>0.91 ± 0.02</td>
<td>0.9</td>
</tr>
<tr>
<td>$W\mu\nu + \text{jets (MC)}$</td>
<td>1.6 ± 0.3</td>
<td>0.90 ± 0.02</td>
<td>0.9</td>
</tr>
<tr>
<td>$Zee + \text{jets (MC)}$</td>
<td>2.6 ± 0.4</td>
<td>0.95 ± 0.06</td>
<td>0.9</td>
</tr>
<tr>
<td>$Z\mu\mu + \text{jets (MC)}$</td>
<td>0.11 ± 0.04</td>
<td>0.90 ± 0.05</td>
<td>0.9</td>
</tr>
<tr>
<td>$t\bar{t} + \text{jets (MC)}$</td>
<td>6.5 ± 0.9</td>
<td>1.0 ± 0.2</td>
<td>1.15</td>
</tr>
<tr>
<td>other (MC)</td>
<td>2.2 ± 0.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Result</td>
<td>20 ± 6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>21.6 ± 0.3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Convergence plots of control ratios

Ratio for $W \rightarrow e\nu$

Ratio for $W \rightarrow \mu\nu$

Ratio for $t\bar{t}$

Ratio for $Z \rightarrow ee$

Ratio for $Z \rightarrow \mu\mu$
1 Introduction to a 4th family of fermions
   - Motivation
   - Search signature
   - Particle and event selection

2 Counting analysis
   - Idea
   - Control samples
   - Estimation of control ratios and signal yields

3 Results for b' (300 GeV)
   - Mixing sample
   - Signal and control regions
   - Result
   - Convergences

4 Summary
\[ \ell^\pm \ell^\pm / 3\ell + \text{jets} \text{ are good search signatures} \]

Data driven analysis for extraction of 4th generation signal developed

Generally applicable on analysis where background and signal pass the same event selection cuts but contribute with different efficiencies in different event categories (here: lepton charges and multiplicities)

Signal yield limited by statistics in SR → Errors on control ratios negligible

Stability: Control ratios converge after few iteration steps

Test with pseudo-data shows no bias for control ratios and signal yield

Not shown: Works also with no signal in pseudo data
Thank you !!!
Backup slides
ST-Ellipse

\[ U = 0 \]

- \( m_h = 115 \text{ GeV} \)
- \( m_h = 200 \text{ GeV} \)
- \( m_h = 300 \text{ GeV} \)
- \( m_h = 1 \text{ TeV} \)
## Particle selections

### Electrons
- Standard (HighPt), tight electrons
- $P_T > 20$ GeV, $|\eta| < 2.5$, $E_T^{\text{cone}20} < 8$ GeV

### Muons
- STACO muons
- $\chi^2/ndf < 5$
- $P_T > 20$ GeV, $|\eta| < 2.5$, $E_T^{\text{cone}20} < 10$ GeV

### Jets
- Cone4H1Tower jets
- Overlap removal with $e, \mu, \tau, \gamma$ within $\Delta R < 0.2$
- $E_T > 25$ GeV, $|\eta| < 2.5$
Background samples

Control regions and the dominating background

- $e^+/e^- : W\nu+\text{jets}$
- $\mu^+/\mu^- : W\mu+\text{jets}$
- $e^+e^- : Zee+\text{jets}$
- $\mu^+\mu^- : Z\mu\mu+\text{jets}$
- $e^+\mu^-/e^-\mu^+ : t\bar{t}+\text{jets}$

Uncorrected samples

- $W\tau\nu+\text{jets}$, $W+bb+\text{jets}$
- $Z\tau\tau+\text{jets}$, $Z+bb+\text{jets}$
- $WW+\text{jets}$, $WZ+\text{jets}$, $ZZ+\text{jets}$
- Single top (s-channel, t-channel, $Wt$-production)
b' (varied $\sigma$) analysis: Signal region

Figure: $l^+l^+ / l^-l^-$ SR
b’ (varied $\sigma$) analysis: $W\nu$ and $W_{\mu\nu}$ control region

Figure: $W\nu$+jets control region

Figure: $W_{\mu\nu}$+jets control region
b' (varied $\sigma$) analysis: Zee and $Z\mu\mu$ control region

Figure: Zee+jets control region

Figure: $Z\mu\mu$+jets control region
b’ (varied $\sigma$) analysis: $t\bar{t}$ control region

Figure: $t\bar{t}$ control region
## Results (b’ 300 GeV)

<table>
<thead>
<tr>
<th>Sample</th>
<th>Events</th>
<th>Final $R_i^k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>data</td>
<td>32 ± 6</td>
<td></td>
</tr>
<tr>
<td>$W\ell\nu$+jets (MC)</td>
<td>0.7 ± 0.2</td>
<td>1.01 ± 0.02</td>
</tr>
<tr>
<td>$W\mu\nu$+jets (MC)</td>
<td>1.8 ± 0.3</td>
<td>1.01 ± 0.02</td>
</tr>
<tr>
<td>$Z$ee+jets (MC)</td>
<td>2.8 ± 0.5</td>
<td>1.03 ± 0.06</td>
</tr>
<tr>
<td>$Z\mu\mu$+jets (MC)</td>
<td>0.11 ± 0.04</td>
<td>0.93 ± 0.05</td>
</tr>
<tr>
<td>$t\bar{t}$+jets (MC)</td>
<td>5.9 ± 0.8</td>
<td>0.9 ± 0.2</td>
</tr>
<tr>
<td>other (MC)</td>
<td>2.2 ± 0.9</td>
<td></td>
</tr>
<tr>
<td>Result</td>
<td>19 ± 6</td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>18.4 ± 0.3</td>
<td></td>
</tr>
</tbody>
</table>
b’ analysis: Signal region

Figure: Signal region
b' analysis: $W\nu$ and $W\mu\nu$ control region

Figure: $W\nu$+jets control region

Figure: $W\mu\nu$+jets control region
b’ analysis: Zee and \(Z_{\mu\mu}\) control region

**Figure:** Zee+jets control region

**Figure:** \(Z_{\mu\mu}+\)jets control region
b’ analysis: $t\bar{t}$ control region

Figure: $t\bar{t}$ control region
b’ analysis: Convergence of b’ signal ratio

Figure: Ratio for b’ signal

ATLAS work in progress
b' analysis: Convergence plots of control ratios

Ratio for $W \rightarrow e\nu$

Ratio for $W \rightarrow \mu\nu$

Ratio for $t\bar{t}$

Ratio for $Z \rightarrow ee$

Ratio for $Z \rightarrow \mu\mu$
b’ analysis: Trilepton categories

Figure: $l^+l^+l^+$/$l^-l^-l^-$

Figure: $l^+l^+l^-/l^-l^-l^+$
### Results (t’ 400 GeV)

<table>
<thead>
<tr>
<th>Sample</th>
<th>Events</th>
<th>Final $R_{i}^{k}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>data</td>
<td>23 ± 5</td>
<td></td>
</tr>
<tr>
<td>$W\ell\nu+$jets (MC)</td>
<td>0.7 ± 0.2</td>
<td>1.00 ± 0.02</td>
</tr>
<tr>
<td>$W\mu\mu+$jets (MC)</td>
<td>1.8 ± 0.3</td>
<td>1.01 ± 0.02</td>
</tr>
<tr>
<td>Zee+jets (MC)</td>
<td>2.8 ± 0.5</td>
<td>1.04 ± 0.06</td>
</tr>
<tr>
<td>$Z\mu\mu+$jets (MC)</td>
<td>0.11 ± 0.04</td>
<td>0.93 ± 0.05</td>
</tr>
<tr>
<td>$t\bar{t}+$jets (MC)</td>
<td>5.9 ± 0.7</td>
<td>0.9 ± 0.2</td>
</tr>
<tr>
<td>other (MC)</td>
<td>2.2 ± 0.9</td>
<td></td>
</tr>
<tr>
<td>Result</td>
<td>10 ± 5</td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>8.39 ± 0.07</td>
<td></td>
</tr>
</tbody>
</table>
Figure: Signal region
## Results @ 50pb\(^{-1}\)

<table>
<thead>
<tr>
<th>Sample</th>
<th>Events</th>
<th>Final (R^k_i)</th>
</tr>
</thead>
<tbody>
<tr>
<td>data</td>
<td>14 ± 4</td>
<td></td>
</tr>
<tr>
<td>(W\ell\nu + \text{jets (MC)})</td>
<td>0.7 ± 0.2</td>
<td>1.00 ± 0.02</td>
</tr>
<tr>
<td>(W\ell\nu + \text{jets (MC)})</td>
<td>1.8 ± 0.3</td>
<td>1.00 ± 0.02</td>
</tr>
<tr>
<td>(Z\ell\ell + \text{jets (MC)})</td>
<td>2.8 ± 0.5</td>
<td>1.03 ± 0.06</td>
</tr>
<tr>
<td>(Z\ell\ell + \text{jets (MC)})</td>
<td>0.11 ± 0.04</td>
<td>0.92 ± 0.05</td>
</tr>
<tr>
<td>(\ell\ell + \text{jets (MC)})</td>
<td>6.3 ± 0.8</td>
<td>1.0 ± 0.2</td>
</tr>
<tr>
<td>other (MC)</td>
<td>2.2 ± 0.9</td>
<td></td>
</tr>
<tr>
<td>Result</td>
<td>0 ± 4</td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
b’ (w/o signal) analysis: Signal region **before** correction
b' (w/o signal) analysis: Signal region after correction

\[ \ell^+ \ell^+ / \ell^- \ell^- \]

ATLAS work in progress

- Mixing sample
- MC t\bar{t}+Jets
- MC Z+Jets
- MC W+Jets
- MC Dibosons+Jets
- MC single-top
- MC Z+bb+Jets
- MC W+bb+Jets
- MC b' (300GeV)
b' (w/o signal) analysis: $t \bar{t}$ CR before/after correction

Figure: Before correction

Figure: After correction
SR with $\not{E}_T > 0$ GeV

Figure: $l^+l^+/l^-l^-$ SR without cut on $\not{E}_T$
SR with $E^\text{LeadingJet}_T < 85$ GeV

Figure: $l^+l^+/l^-l^-$ SR with anti-cut on leading jet $E_T$
SR with $E_{T}^{\text{LeadingJet}} < 85$ GeV and $P_{T}^{\text{LeadingLep}} > 0$ GeV

Figure: $l^{+}l^{+}/l^{-}l^{-}$ SR with anti-cut on leading jet $E_{T}$ and no cut on leading lepton $P_{T}$