SEARCH for Heavy 4th generation quarks at CMS

Sadia Khalil
Kansas State University

On behalf of the CMS Collaboration
ICHEP 2012
July 4-12, 2012 - Melbourne, Australia
4th Generation

- A simplest extension of the SM3, adding another fermion generation
- It is not excluded by Electroweak precision measurements
- It allows indirect bounds on the Higgs boson mass to be relaxed
- It can enhance CP violation significantly to explain the matter antimatter asymmetry in the Universe
- If SM4 exists we expect small mass splitting between the $t'$ and $b'$ masses: $|m_{t'} - m_{b'}| < m_W$
  - M. Baak et al., arXiv:1107.0975
4th Gen: Vector-like Quark

• Vector-like fermions (non-chiral fermions) can be found in models like:
  
  ▶ Little Higgs model
  
  ▶ Warped extra dimensions

• These models provide an explanation to the large difference between the Plank and the electroweak scale, the so called hierarchy problem in the SM

• \( T' \to tH, tZ \) or \( B' \to bH, bZ \) -- flavor changing neutral current (FCNC) decays enhance branching fractions
• CMS 4th generation searches @7 TeV

  - t’ pair search
    - dilepton channel
      \[ t't' \rightarrow WbWb \rightarrow \nu b\bar{\nu}b \]
    - lepton+jets channel
      \[ t't' \rightarrow WbW\bar{b} \rightarrow \nu bbq\bar{q} \]
  - b’ search
    - trilepton channel
    - same sign dilepton channel

  - Inclusive t’ and b’ search
    - singly produced
      \[ b'b \rightarrow bWb \]
      \[ b't \rightarrow tWbW \rightarrow bWWbW \]
    - pair produced
      \[ b't' \rightarrow tWbW \rightarrow bWWbW \]
      \[ t't' \rightarrow bWWbW \]
      \[ b'b' \rightarrow tWtW \rightarrow bWWbWW \]

  - T’ vector like pair
    \[ T \bar{T} \rightarrow tZtZ \rightarrow bbWWZZ \]

  - B’ vector-like pair
    \[ B \bar{B} \rightarrow bZbZ \]
Search for $t' \rightarrow bW(l+jets)$

### Selection
- A lepton $e(\mu)$ with $p_T > 35$ GeV
- $\geq 4$ jets of $p_T > 35$, $\geq 1$ b-tagged jet
- Missing $E_T > 20$ GeV

### Discriminating variables

### Strategy
- Apply kinematic fit for mass reconstruction ($M_{fit}$) with constraints
  - $m(l\nu) = m(qq) = M_W$
  - $m(l\nu b) = m(qqb)$
- Look in the $H_T$ and $M_{fit}$ tails for signs of a massive quark decay

$$H_T = p_T^{lepton} + p_T^{miss} + \sum p_T^{jets}$$
Search for $t' \rightarrow bW(l+jets)$

- The 2D $H_T$ vs $M_{fit}$ histograms have few empty or low occupancy bins
- Rebin them to extract the correct statistical inferences
  - Project 2D histograms into 1D profiled with analytic functions
  - Sort by ordering the bins in descending S/B ratio
  - Merge neighboring bins into 1D histogram until a minimum precision in the expected number of background and signal events is achieved

EXO-11-099

4.7 fb$^{-1}$

Friday, July 6, 2012
Search for $t' \to bW(l+\text{jets})$

- Compute the $t'$ pair cross section using CLs method
  - Likelihood ratio is used as a test statistic for an observable $x$, parameter of interest $\sigma$ and nuisance parameters $\alpha$
    \[ t(x|\sigma) = \begin{cases} L(x|\sigma, \hat{\alpha}) / L(x|\hat{\sigma}, \hat{\alpha}) & \text{if } \sigma > \hat{\sigma} \\ 1 & \text{if } \sigma \leq \hat{\sigma}. \end{cases} \]
  - 95% C.L. upper limit corresponds to $CL_s = \frac{CL_{s+b}}{CL_b} = 0.05$.
- The nuisance parameters includes
  - Normalization of electroweak and ttbar backgrounds
  - Jet energy scale
  - Integrated luminosity
  - Lepton efficiency
  - Parton shower matching threshold

Result

![Graph showing the observed and expected limit for $t'$ mass]

Observed limit: $m_{t'} > 560 \text{ GeV/c}^2$ @ 95% CL
Search for inclusive b′/t′ production

- Simplify CKM4 with one free parameter: \( A = |V_{tb}|^2 = |V_{t'b'}|^2 \)
  - \( A > 0.66 \) (|V_{tb}| > 0.81 @ 95% C.L.) from Tevatron

\[
V^{4\times4}_{\text{CKM}} = \begin{pmatrix}
V_{ud} & V_{us} & V_{ub} & V_{ud'} \\
V_{cd} & V_{cs} & V_{cb} & V_{cd'} \\
V_{td} & V_{ts} & V_{tb} & V_{td'} \\
V_{t'd} & V_{t's} & V_{t'b} & V_{t'b'}
\end{pmatrix} = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & \sqrt{A} & \sqrt{1-A} \\
0 & 0 & -\sqrt{1-A} & \sqrt{A}
\end{pmatrix}
\]

- Assume degenerate states: \( m_{t'} = m_{b'} = m_{q'} \)
  - Assume the branching fractions to be \( \sim 100\% \)

- Baseline selection
  - Lepton (e/\( \mu \)) with \( p_T > 40 \) GeV
  - \( \geq 1 \) jet of \( p_T > 30 \) GeV and \( \geq 1 \) b-tagged
  - Missing \( E_T > 40 \) GeV

- Final state topologies contains
  - 1-4 W bosons (\( \geq 1 \) W decay leptonically)
  - 2 b-quarks

- Search is performed
  - Single lepton(e/\( \mu \)) / Same-sign dilepton / Trilepton + jets

- Sensitivity to all signal process
- Sensitivity to signal with 1 or 2 b′ process
Search for inclusive b'/'t' production

Discriminator: Scalar sum of reconstructed objects ($S_T$) and hadronic top mass ($m_{bW}$)

\[ S_T = \not{E}_T + p_T^l + p_T^b + p_T^j + \sum_{i=0}^{N} p_T^{W_{q\bar{q}}} \]

- 2b1W box: single t' (Exactly 2 b-tagged jet)
- 1b2W box (2b2W box): t't' (Exactly 1 (2) b-tagged jet(s), exactly 1 hadronically decaying W)
- 1b3W box (2b3W box): single b' (Exactly 1 (2) b-tagged jet(s), exactly 2 hadronically decaying W)
- 1b4W box (2b4W box): b'b' (Exactly 1 (2) b-tagged jet(s), at least 3 hadronically decaying W)
Search for inclusive $b'/t'$ production

multilepton channel

• Same-sign dilepton channel
  ▸ $\geq 2$ leptons (charge ++ or --), $\geq 4$ jets
  ▸ Backgrounds
    - Wrong-sign lepton (e.g., Z or dilepton events)
    - Fake leptons (single top, semileptonic $t\bar{t}$, $W$)
    - Irreducible background ($WZ$, $ZZ$, $ttV$, $W^{\pm}W^{\pm}$)

• Trilepton channel
  ▸ $\geq 3$ leptons (charge ++- or +--), $\geq 2$ jets
  ▸ Background ($WZ$, $ZZ$, $ttV$) estimate from simulation

• Suppress Z events, $|M_{ll} - M_Z| > 10$ GeV

<table>
<thead>
<tr>
<th>type</th>
<th>2 muons</th>
<th>2 electrons</th>
<th>electron+muon</th>
<th>trilepton</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Background</td>
<td>$0.83 \pm 0.11$</td>
<td>$1.36 \pm 0.19$</td>
<td>$2.27 \pm 0.22$</td>
<td>$0.96 \pm 0.12$</td>
</tr>
<tr>
<td>Signal ($A = 1, m_{q'} = 550$ GeV)</td>
<td>$3.31 \pm 0.15$</td>
<td>$2.03 \pm 0.36$</td>
<td>$5.29 \pm 0.19$</td>
<td>$3.37 \pm 0.16$</td>
</tr>
<tr>
<td>Signal ($A = 0.8, m_{q'} = 550$ GeV)</td>
<td>$3.79 \pm 0.15$</td>
<td>$2.29 \pm 0.36$</td>
<td>$6.00 \pm 0.19$</td>
<td>$3.65 \pm 0.16$</td>
</tr>
</tbody>
</table>

These event counts are used in the limit calculation.
Model-dependent exclusion limit on $m_{t'} = m_{b'}$ as a function of $A$

- $t'b'$ production $\propto A$
- $t'b$ and $t b'$ production $\propto 1 - A$
- $m_{q'} < 685$ GeV excluded at 95% CL

- Effect of mass diff: $m_{t'} - m_{b'} = 25$ GeV is studied
  - Limit shifts about 20 GeV
- The electroweak $t'b'$ process is omitted
  - Less stringent limit for $m_{t'} = m_{b'}$
Search for vector-like $B' \rightarrow bZ(dilepton)$

Event Selection

- Clean signature includes
  - $\geq 2$ jets, $\geq 1$ b-tag ($p_T > 65$ GeV)
  - $Z$ Cand: $(60 < m_{ll} < 120)$ GeV, $p_T > 95$ GeV

Strategy

- Search for resonance peak of mass spectrum of the $bZ$ candidate
  - Reconstruct $B'$ candidate using leading $p_T$ $Z$ boson and the leading $p_T$ b-tagged jet
Search for vector-like $B' \rightarrow bZ$(dilepton)

**Yields**

<table>
<thead>
<tr>
<th></th>
<th>$Z \rightarrow ee$</th>
<th>$Z \rightarrow \mu\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B'(350 \text{ GeV})$</td>
<td>222$\pm$6</td>
<td>345$\pm$9</td>
</tr>
<tr>
<td>Total Pred</td>
<td>648$\pm$15</td>
<td>999$\pm$26</td>
</tr>
<tr>
<td>DATA</td>
<td>604$\pm$24</td>
<td>928$\pm$30</td>
</tr>
</tbody>
</table>

Assuming a branching fraction of 100% $B' \rightarrow bZ$

- With the observed upper limit at 95% CL on the production cross section, we exclude a $B'$ quark with a mass < 550 GeV
- **See details by Kai-Yi Kao in poster session**
## Conclusions

<table>
<thead>
<tr>
<th>Search</th>
<th>Channel</th>
<th>Lower mass limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t' \rightarrow bW ) pair</td>
<td>dileptons</td>
<td>557 GeV/c²</td>
</tr>
<tr>
<td>( t' \rightarrow bW ) pair</td>
<td>lepton+jets</td>
<td>560 GeV/c²</td>
</tr>
<tr>
<td>( b' \rightarrow tW ) pair</td>
<td>trilepton and same-sign dilepton</td>
<td>611 GeV/c²</td>
</tr>
<tr>
<td>( T' \rightarrow tZ ) pair</td>
<td>three leptons</td>
<td>475 GeV/c²</td>
</tr>
<tr>
<td>( B' \rightarrow bZ ) pair</td>
<td>two leptons</td>
<td>550 GeV/c²</td>
</tr>
<tr>
<td>Model-Dependent ( t'/b' )</td>
<td>lepton(s)+jets</td>
<td>685 GeV/c²</td>
</tr>
</tbody>
</table>

- CMS has the most stringent limits on the existence of 4th generation quarks
- We have reached the critical mass of \(~550\) GeV/c² at which fermion’s weak interactions become non-perturbative

Thank you!
Extra material
Particle Flow Algorithm

- Provides a list of observable particles that describe the event
  - muons, electrons, photon, charged and neutral hadrons
- It combines the information from all CMS sub-detectors to achieve this
- This list is used to reconstruct higher level objects like jets, MET
Search for inclusive b’/t’ production

Subsamples with 2 W bosons

- Reconstruct hadronic top mass ($m_{bW}$)
  - Likelihood ratio (LH) with 7 observables (angles, $W$ mass, b-tag discriminator values, $p_T$ of the top quark candidate)
  - Use the jet combination with the largest LR value

- $W \rightarrow qq$ counting procedure
  - remove the 1 or 2 b-tag jets
  - Choose the jet pair that minimizes $|m_{j1j2} - m_{W^{\text{sim}}}|$
  - If $|m_{j1j2} - m_{W^{\text{sim}}}| < \sigma_{W^{\text{sim}}}$, a $W \rightarrow qq$ event is found
  - Remove the jet pair that formed the $W \rightarrow qq$
  - Repeat the procedure until no hadronically decaying $W$ are found
Backgrounds for same-sign dilepton channel

- **Wrong-sign lepton** (e.g., Z or dilepton events)
  - Missing $E_T < 20$ GeV, $M_T < 25$ GeV
  - Require two electrons within a 10 GeV window around Z mass
  - Charge misidentification ratio, $R = N_{SS}/2N_{OS}$
  - Rescale the events passing all selections except the same-sign requirement

- **Non-prompt leptons** (single top, semileptonic ttbar, W)
  - Missing $E_T < 20$ GeV, $M_T < 25$ GeV
  - Veto events with $|M_{ll} - M_z| < 20$ GeV
  - Count # loose ($N_L$) and tight ($N_T$) leptons with $p_T < 35$ GeV
  - Probability that a loose lepton passes the tight cuts: $\epsilon_{fl} = N_T / N_L$
  - Require the events to pass selection criteria except
    - One tight lepton
    - One loose but not tight lepton
  - Scale the data yields by $R_{fl} = \epsilon_{fl}(1 - \epsilon_{fl})$
Search for $t' \rightarrow bW(\text{dilepton})$

**Selection**

- Two opposite sign high $p_T$ leptons (ee, $e\mu, \mu\mu$)
- $Z/\gamma \rightarrow ee/\mu\mu$ veto
- $\geq 2$ jets, $\geq 1$ b-tagged jet
- Missing $E_T > 50$ GeV

**Strategy**

- Invariant mass of the lepton and the b-tagged jets:
  $$M_{lb} = \sqrt{(E_l + E_b)^2 - |p_l + p_b|^2}$$
- Combine the two leptons and two b-tagged jets
  - 4 possible values of $M_{lb} = M_{lb}^{\text{min}}$
  - Choose the minimum
  - For a top decay $M_{lb}^2 < M_t^2 - M_W^2$

**Results and summary**

- Observed limit $8 \text{pb}$
- Expected limit $6 \text{pb}$
- Theoretical cross section $\%$/$\%$

**Exclude $m_{t'} < 557$ GeV @ 95% C.L. on the production cross section**
• Backgrounds (mostly data driven)
  ▪ Category I - b-mistagged jet(s) and prompt leptons
  ▪ Category II - fake lepton(s) and real b-tagged jet(s)
  ▪ Category III - b-mistagged jet(s) and fake lepton(s) (negligible)
  ▪ Category IV - 2 real b-tagged jets and 2 real leptons (obtained from MC)

<table>
<thead>
<tr>
<th>Sample</th>
<th>Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category I (from data)</td>
<td>0.7 ± 0.8</td>
</tr>
<tr>
<td>Category II (from data)</td>
<td>0.0^{+0.4}_{-0.0}</td>
</tr>
<tr>
<td>Category III (simulated)</td>
<td>1.0 ± 0.7</td>
</tr>
<tr>
<td>Total prediction</td>
<td>1.8 ± 1.1</td>
</tr>
<tr>
<td>Data</td>
<td>1</td>
</tr>
</tbody>
</table>

The number of expected events from background processes is 0.7 ± 0.70 and one event is observed in the e\(\mu\) channel. There is thus no evidence for an excess of events above SM expectations. A summary of the observed and predicted yields is presented in Table 1.

The simulated distribution of \(M_{\min}\) from background processes is compared with the data in Fig. 0.5 where the expected distribution for a \(t\bar{t}\) signal with \(M_t = 175\) GeV is also shown. The simulated background yields in the signal region are scaled so that they match the yields estimated from control regions in data, as given in Table 1. Outside the signal region the simulated background yields are taken without rescaling.

Finally, 5 fb\(^{-1}\) confidence level (CL) upper limits on the production cross section of \(t\bar{t}\) as a function of \(t\bar{t}\) mass are set using the CL\(_s\) method, in which a log-normal model of nuisance parameter integration is assumed. The limit calculation is based on the information provided by the observed event count combined with the values and the uncertainties of the luminosity measurement, the background search for \(t^\prime \rightarrow bW(dilepton)\).
Search for $b' \rightarrow tW$ (di/tri-lepton)

Selection

- Leptons ($e/\mu$) with $p_T > 20$ GeV
- Suppress Z events, $|M_{ll} - M_Z| > 10$ GeV
- Reject events with
  - < 4 jets for the same-sign dilepton
  - < 2 jets for trilepton channel
- $\geq 1$ b-tagged jets

Strategy

- Construct scalar quantity, $S_T > 500$ GeV
- $S_T = |E_T^{miss}| + \sum |p_T^l| + \sum |p_T^{jet}|$

$dilepton$  

$trilepton$

CMS $L = 4.9$ fb$^{-1}$ at $\sqrt{s} = 7$ TeV

$M_{b'} < 611$ GeV/c$^2$ is excluded at 95% CL
Search for $b' \rightarrow tW$(di/tri-lepton)

- **Backgrounds (mostly from ttbar)**
  - **Sources for same sign dilepton channel**
    - Type I (data driven) -- Fake lepton
    - Type II (data driven) -- Charge Misidentification
    - Type III (from MC) -- Prompt dileptons
  - **Sources for trilepton channel**
    - Dominated by 3 prompt leptons events (ttW)

<table>
<thead>
<tr>
<th>Sources</th>
<th>Same-charge</th>
<th>Trilepton</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I + Type II</td>
<td>7.8 ± 2.8</td>
<td></td>
</tr>
<tr>
<td>Type III</td>
<td>3.6 ± 0.6</td>
<td>0.78 ± 0.21</td>
</tr>
<tr>
<td>Background sum</td>
<td>11.4 ± 2.9</td>
<td>0.78 ± 0.21</td>
</tr>
<tr>
<td>Observed yield</td>
<td>12</td>
<td>1</td>
</tr>
</tbody>
</table>
Search for vector-like $T' \rightarrow tZ$ (trilepton)

- **Clean signature includes**
  - $\geq 3$ leptons (e/\(\mu\))
  - $\geq 2$ jets
  - $(60 < m_{ll} < 120)$ GeV

**Strategy**

- **Construct the variable $R_T > 80$ GeV**

$$R_T = \left( \sum p_T^{\text{leptons}} + \sum p_T^{\text{jets}} \right) - \left( \sum p_T^{\text{two leading leptons}} + \sum p_T^{\text{two leading jets}} \right)$$

- $\geq 2$ jets & $R_T > 80$ GeV
Search for vector-like $T' \rightarrow tZ$(trilepton)

Yields

<table>
<thead>
<tr>
<th>Component</th>
<th>Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T'(350 \text{ GeV})$</td>
<td>$57.8\pm11.0$</td>
</tr>
<tr>
<td>Background</td>
<td>$4.60\pm1.04$</td>
</tr>
<tr>
<td>DATA</td>
<td>$7$</td>
</tr>
</tbody>
</table>

Assuming a branching fraction of 100% $T' \rightarrow tZ$

- With the observed upper limit at 95% CL on the production cross section, we excludes a $T'$ quark with a mass < 475 GeV

Limit at 95% CL: $M_T > 475 \text{ GeV}/c^2$

Friday, July 6, 2012