Search for heavy bottom-like quarks
decaying to an electron or muon and jets in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV

We report the most sensitive direct search for pair production of fourth-generation bottom-like chiral quarks \( (b') \) each decaying promptly to \( tW \). We search for an excess of events with an electron or muon, at least five jets (one identified as due to a \( b \) or \( c \) quark) and an imbalance of transverse momentum using data from \( pp \) collisions collected by the CDF II detector at Fermilab with an integrated luminosity of 4.8 fb\(^{-1}\). We observe events consistent with background expectation and exclude upper limits on the \( b' \) pair production cross section \( (\sigma_{bb'} \lesssim 30 \text{ fb for } m_{b'} > 375 \text{ GeV}/c^2) \) and exclude \( m_{b'} < 372 \text{ GeV}/c^2 \) at 95\% confidence level.

PACS numbers: 12.60.-i, 13.85.Rm, 14.65.-q, 14.80.-j

The standard model (SM) of particle physics accommodates three generations of fundamental fermions, but is agnostic on the issue of a fourth generation. Precision measurements in the electroweak sector are not inconsistent with a fourth-generation of fermions if there is a 50-
transverse momentum [16] is reconstructed using fully corrected calorimeter and muon information [12].

Production and decay of b pairs would appear as events with a charged lepton and missing transverse momentum from one leptonically decaying W, and a large number of jets from the two b quarks and the hadronic decays of the other three W bosons. We select events with exactly one electron or muon, at least five jets, and at least 20 GeV/c of missing transverse momentum. At least one of the jets must be identified as due to b quark decay. We find 357 events satisfying these requirements.

We model the production and decay of b pairs with MADGRAPH [17]. Additional radiation, hadronization and showering are described by PYTHIA [18]. The detector response for all simulated samples is modeled by CDFSIM [19]. The signal efficiency for the above requirements is approximately 10%, rising with b mass. There are eight quarks produced in the decay, but the most likely number of reconstructed jets is six, as quarks that are close together are likely to be merged into a single jet, and some of the quarks produce jets which fall below the transverse momentum threshold. Complete mass reconstruction is therefore not possible in the majority of the events; instead, we examine the event $H_T$, the scalar sum of the transverse momentum of the lepton, jets and missing transverse momentum. This is well correlated with the mass of the heavy quark and serves as an approximate mass reconstruction.

The dominant background (80%) is top-quark pair production with additional jets from initial or final state radiation. This background can be distinguished from the signal as it has smaller $H_T$. We model this background using MADGRAPH tt production with $m_t = 172.5$ GeV/c$^2$ in which radiation of up to three additional hard partons (including heavy flavor) are described explicitly using matrix-elements, and additional radiation is described by the parton-shower; the MLM [20] scheme is used to match the matrix-element and parton-shower contributions. This gives a precise description of events with $\leq 7$ jets, where a b$'$ signal would be expected. Events with eight jets and above are described by the parton shower, which has significantly larger systematic uncertainties. We normalize the tt background to the NLO cross section [28], and confirm that it is well modeled by examining $t\bar{t}$-dominated regions in the data.

The second dominant background process ($\approx 10\%$) is the associated production of W boson and jets. Samples of simulated W+jets events with light- and heavy-flavor jets are generated using the ALPGEN [21] program, interfaced with parton-shower model from PYTHIA. The W+jets samples are normalized to the measured W cross section, with an additional multiplicative factor for the relative contribution of heavy- and light-flavor jets, the standard technique in measuring the top-quark pair production cross section [15]. Multi-jet background ($\approx 5\%$), in which a jet is misreconstructed as a lepton, is modeled
using a jet-triggered sample normalized in a background-dominated region at low missing transverse momentum. The remaining backgrounds, single top and diboson production, are modeled using PYTHIA and normalized to next-to-leading order cross sections [22]. The combined background expectation is 365±194 events, including systematic and statistical uncertainties.

A $b'$ signal would be readily separated from the background both in the number of jets and the $H_T$. To take advantage of both of these characteristics, we introduce a variable “Jet-$H_T$”, which equals $H_T$ for events with exactly 5 jets, Jet-$H_T = H_T + 1000 \text{ GeV}$ for events with exactly 6 jets, and Jet-$H_T = H_T + 2000 \text{ GeV}$ for events with at least 7 jets. This is equivalent to a two-dimensional analysis in $N_{jets}$ and $H_T$. Figure 1 shows the distributions of an example $b'$ signal with $m_{b'} = 350 \text{ GeV}/c^2$ and the backgrounds in Jet-$H_T$.

We consider several sources of systematic uncertainty on both the background rates and distributions, as well as on the expectations for the signal. Each affects the expected sensitivity to new physics expressed as an expected cross section upper limit in the no-signal assumption. The dominant systematic uncertainties are the jet energy scale [14], contributions from additional interactions, and descriptions of initial and final state radiation [23]. In each case, we treat the unknown underlying quantity as a nuisance parameter and measure the distortion of the Jet-$H_T$ spectrum for positive and negative fluctuations. Each uncertainty weakens the expected 95% confidence level (C.L.) cross section upper limit by ≈ 60% individually. Additional uncertainty comes from parton distribution functions (PDF) [24, 25], the matching scale used between the matrix-element and the parton shower, overall background normalization, and uncertainties in performance of the $b$-quark identification algorithm. The overall impact on the expected sensitivity is ≈ 100% in the cross section and ≈ 20 GeV$/c^2$ on the expected mass limit.

To validate the description of the backgrounds, we verify that the low $H_T$ region is well-described where there is little signal expected. See Table I. In events with $\geq 7$ jets, the observed $H_T$ is larger than predicted by our background model. The number of observed events with $\geq 7$ jets and $H_T > 500 \text{ GeV}$ is 12 where we expect $3.4 \pm 3.4$. However, the total number of events observed in the low $H_T$ and high $H_T$ regions combined is consistent with expectation. Considering only the number of events in the high $H_T$ regions and taking into account the systematic uncertainties in the background prediction, we see a more significant excess than that observed in the data in 12% of simulated experiments.

The full Jet-$H_T$ spectrum is used in the analysis. Since there is no evidence for the presence of $b'$ events in the data, we calculate 95% C.L. upper limits on the $b'$ production cross section, by performing a binned maximum-likelihood fit in the Jet-$H_T$ variable, allowing for systematic and statistical fluctuations via template morphing [26]. We use the likelihood-ratio ordering prescription [27] to construct classical confidence intervals in the theoretical cross section by generating ensembles of simulated experiments that describe expected fluctuations of statistical and systematic uncertainties on both signal and backgrounds. The observed limits are consistent with expectation in the background-only hypothesis and are given together with theoretical next-to-leading-order (NLO) cross sections [28, 29] in Table II and shown in

![Figure 1: Distributions in jet multiplicity and Jet-$H_T$ (defined in the text). The example $b'$ signal has $m_{b'} = 350 \text{ GeV}/c^2$ and would have 29 ± 4.5 events expected in this sample. Top pane is log scale; bottom pane shows the difference between expected and observed events on a linear scale, as well as the total uncertainty on the expected events.](image-url)
TABLE I: Expected and observed events in a background-dominated control region ($H_T < 400, 450, 500$ for $N_{jet} = 5, 6, \geq 7$, respectively) and in a signal-dominated region ($H_T > 400, 450, 500$ for $N_{jet} = 5, 6, \geq 7$) for our selection (see text). Uncertainties are statistical and systematic, combined in quadrature.

<table>
<thead>
<tr>
<th>Jets</th>
<th>Control Region</th>
<th>Signal Region</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>207 ± 125</td>
<td>199</td>
<td>396</td>
</tr>
<tr>
<td>6</td>
<td>43 ± 31</td>
<td>40</td>
<td>83</td>
</tr>
<tr>
<td>≥ 7</td>
<td>11 ± 3.9</td>
<td>5</td>
<td>16</td>
</tr>
</tbody>
</table>

We convert upper limits on the pair-production cross sections to lower limits on the fermion masses. The relative cross-section uncertainty of $\approx 10\%$ due to scale and PDF uncertainties translates into $\approx 3 \text{ GeV}/c^2$ for the mass lower limits.

In conclusion, we have searched for pair production of $b'$ quarks with subsequent decay to $t\bar{t}W$. Though there are events with larger $H_T$ than expected in the 7-jet event distribution in Fig 1, we do not see evidence of a signal. We calculate upper limits on the $b'$ pair production cross section ($\lesssim 30$ fb for $m_{b'} > 375 \text{ GeV}/c^2$) and set the most restrictive direct lower limit on the mass of a down-type fourth-generation quark, increasing the limit by 34 GeV/$c^2$ to $m_{b'} \geq 372 \text{ GeV}/c^2$ and significantly reducing the allowed mass range.

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the Bundesministerium für Bildung und Forschung, Germany; the World Class University Program, the National Research Foundation of Korea; the Science and Technology Facilities Coun-

TABLE II: Theoretical cross sections ($\sigma_{\text{NLO}}$ in fb [28, 29]), expected $b'$ yield ($N_{exp}$) after selection, median expected 95% C.L. limit ($\sigma_{\text{exp}}$ in fb), and observed 95% C.L. limit ($\sigma_{\text{obs}}$ in fb) for $b'$ at varying masses. $\sigma_{\text{NLO}}$ and $N_{exp}$ have 10% uncertainties.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_{\text{NLO}}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N_{exp}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_{\text{exp}}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_{\text{obs}}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

[7] We specify chiral quarks to distinguish from theories of a vector-like fourth generation.
[11] CDF uses a cylindrical coordinate system with the $z$ axis along the proton beam axis. Pseudorapidity is $\eta \equiv -\ln(\tan(\theta/2))$, where $\theta$ is the polar angle relative to the proton beam direction, and $\phi$ is the azimuthal angle while $p_T = |p| \sin \theta, E_T = E \sin \theta$.
Missing transverse momentum, $E_T$, is defined as the magnitude of the vector $-\sum E_T^i n_i$ where $E_T^i$ are the magnitudes of transverse energy contained in each calorimeter tower $i$, and $n_i$ is the unit vector from the interaction vertex to the tower in the transverse $(x, y)$ plane.


