Search for Anomalous Heavy-Flavor Quark Production in Association with W Bosons
We search for anomalous production of heavy-flavor quark jets in association with a $W$ boson at the Fermilab Tevatron $p\bar{p}$ Collider in final states in which the heavy-flavor quark content is enhanced by requiring at least one tagged jet in an event. Jets are tagged using one algorithm based on semileptonic decays of $b/c$ hadrons, and another on their lifetimes. We compare jets ($164 \text{ pb}^{-1}$) and jets ($145 \text{ pb}^{-1}$) channels collected with the D0 detector at $\sqrt{s} = 1.96 \text{ TeV}$ to expectations from the standard model and set upper limits on anomalous production of such events.

DOI: 10.1103/PhysRevLett.94.152002 PACS numbers: 13.85.Qk, 13.85.Ni
At the Tevatron, the primary SM contributions to $W+\text{HF}$ quarks in the final state are expected from $t\bar{t}$, $Wb\bar{b}/c\bar{c}$ (where the $bb$ or $c\bar{c}$ pairs arise from gluon splitting), and $Wc$ final states, with additional contributions arising from single top quark or $WZ$ (with $Z \rightarrow bb/c\bar{c}$) production. The production of $W$ bosons accompanied by light quarks or gluons (referred to as $W+$ jets in this Letter) contributes to the background when the light-quark or gluon jets are misidentified as jets from HF quarks. Since $W$ bosons are identified through their $W \rightarrow e\nu$ and $W \rightarrow \mu\nu$ decays, background can arise from $Zb\bar{b}$, $Z\gamma$ (with one $Z \rightarrow bb/c\bar{c}$), and $Z+$ jets production when one of the leptons from the $Z \rightarrow \ell^+\ell^-$ decay is not detected. The main instrumental background arises from multijet processes in which a jet is misidentified as a lepton, and an imbalance in transverse momentum ($\vec{E}_T$) is generated through a mis-measurement of the jets or a lepton. To be accepted, such events must also contain tagged HF jets or misidentified non-HF jets.

The data were collected with the D0 detector [2] in Run II of the Fermilab Tevatron in $p\bar{p}$ collisions at $\sqrt{s}=1.96$ TeV. The detector components used in this analysis include the central tracker, calorimeter, and muon detectors. The central tracker consists of a silicon microstrip tracker (SMT) and a central fiber tracker (CFT), both located within a 2 T superconducting solenoidal magnet, covering the regions $|\eta|<3.0$ and $|\eta|<2.0$, respectively. The uranium/liquid-argon calorimeter consists of three sections, each housed in a separate cryostat [3]. The central calorimeter covers pseudorapidity $|\eta|\leq 1.1$, while the two end calorimeters extend the coverage to $|\eta|=4.0$. The muon system is located outside the calorimeters and consists of layers of tracking detectors and scintillation trigger counters, one within and one beyond 1.8 T iron toroids.

The $W \rightarrow e\nu$ and $W \rightarrow \mu\nu$ candidates are selected by triggering on electrons and muons. The mean trigger efficiency for electrons with transverse momentum $p_T>20$ GeV/c and $|\eta|<1.1$ is $(97.0 \pm 0.3)\%$, while for muons with $p_T>20$ GeV/c and $|\eta|<1.6$ it is $(62.1 \pm 3.4)\%$. The integrated luminosities are $164 \pm 11$ pb$^{-1}$ for the electron and $145 \pm 9$ pb$^{-1}$ for the muon samples.

Candidates for $W \rightarrow e\nu$ are restricted by requiring just one isolated electron with $p_T>20$ GeV/c and $|\eta|<1.1$, defined relative to the geometrical center of the detector. Lepton isolation requires a separation in $\eta$ and azimuth ($\phi$) of $R=\sqrt{\Delta \eta^2 + (\Delta \phi)^2}>0.5$ from all jets. Electrons are defined using a cone algorithm, and by the energies deposited in calorimeter towers within a radius of $R=0.2$ of the electron axis, with at least 90% required to be within the electromagnetic portion of the calorimeter, and by the total energy in a cone of $R=0.4$ centered on the same axis, which must not exceed the electron’s energy by more than 15%. In addition, the shower shape must be compatible with that expected for an electron.

Candidates for $W \rightarrow \mu\nu$ must contain just one isolated muon with $p_T>20$ GeV/c and $|\eta|<1.6$, defined relative to the geometrical center of the detector. Muons must satisfy two additional isolation criteria: (i) the transverse energy deposited in the calorimeter in the annular region of $0.1 < R < 0.4$ around the muon’s path must be smaller than 2.5 GeV and (ii) the vector sum of the $p_T$ of all tracks within $R=0.5$ of the muon’s trajectory must be less than 2.5 GeV/c (excluding the track matched to the muon).

Lepton identification is refined by requiring the trajectory of a track reconstructed in the SMT and CFT to match either the position of the electron energy cluster in the calorimeter or the position of hits in the muon detector. To complete the selection, all events are also required to have $\vec{E}_T>20$ GeV, and the azimuthal angle between the lepton and the direction of the $\vec{E}_T$ must be greater than $22.5^\circ$. To eliminate poorly reconstructed events, the primary vertex (PV) of the event must contain at least three tracks, and its $z$ position (along the beam) has to be closer than 60 cm from the center of the detector. Finally, to reject a multijet background, the reconstructed transverse mass of $(\ell, \vec{E}_T)$ must be consistent with that of the $W$ boson, $40 < M_{\ell\nu} < 120$ GeV/c$^2$ (i.e., we assume that $\vec{E}_T$ corresponds to the transverse energy of the neutrino).

After selecting $W$-boson candidates, we evaluate the HF-quark content of each event. Jets are defined using an iterative seed-based cone algorithm, clustering calorimeter energy within $R=0.5$ [4]. This is subsequently corrected for the jet energy scale (JES), based on momentum balance in photon + jet events. We consider only jets with $p_T>25$ GeV and $|\eta|<2.5$ and evaluate them using two HF-tagging algorithms, as described below.

The soft-lepton tagging algorithm is based on low-$p_T$ muons that arise from semileptonic decays of HF quarks produced in a jet nearby in $(\eta, \phi)$. Only muons with $p_T>4$ GeV/c and $|\eta|<2.0$ are considered. To reject a $Z \rightarrow \mu\mu$ background, we require $p_T<15$ GeV/c for the muon. Jets with a muon within $R=0.5$ of the jet axis are deemed tagged. Typical SLT efficiencies for $b$-quark jets are $11\%$, $3\%$ for $c$-quark jets, and $0.4\%$ for light-quark jets. The muon present in SLT events enhances the average single-muon trigger efficiency from $(62.1 \pm 3.4)\%$ to $(68.4 \pm 3.5)\%$.

Secondary-vertex tagging is used to identify displaced secondary vertices (SV) of long-lived particles. To form these SV, charged tracks are selected on the basis of the significance of their distance-of-closest-approach (dca) to the PV. Tracks are first grouped in $p_T>1$ GeV/c and $dca/\sigma_{dca}>3.5$, where $\sigma_{dca}$ is the uncertainty on the track’s dca. Protovertices are formed by adding tracks to the initial grouping, provided their contribution to the $\chi^2$ of the vertex fit is small. Secondary vertices are selected by requiring the transverse distance from the SV to the beam direction, $L_{xy}$, to be less than 2.6 cm, and the decay-length significance, $L_{xy}/\sigma_{L_{xy}}$, to be greater than 7, where $\sigma_{L_{xy}}$ is the estimated uncertainty on $L_{xy}$ calculated from the error matrices of the tracks in the vertex. Jets are considered tagged by this algorithm when a
SV lies within $R = 0.5$ of the jet axis. This SVT algorithm exhibits a typical tagging rate for $b$-quark jets of $\approx 25\%$, 8% for $c$-quark jets, and 0.25% for light-quark jets.

To predict SM rates, Monte Carlo (MC) events are generated for the processes mentioned above, with the exception of multijet production, which is estimated from data (see below). $W/Z +$ jets (both HF and light-quark jets), $t\bar{t}$, and diboson processes are simulated with ALPGEN [5]. Single top-quark processes are simulated using COMPHEP [6]. All events are generated with $m_{\text{top}} = 175$ GeV/$c^2$. Hadronization and showering of these events is based on PYTHIA [7]. The exceptions are $W/Z + b$ processes, where $Zb$ is simulated using PYTHIA, and the contribution from $Wb$ is estimated from the parametrized MCFM MC [8] and used to calculate a cross section relative to $Wb\bar{b}$ production assuming the jet-$p_T$ spectrum from PYTHIA for inclusive $W$ boson production. All MC events are generated at $\sqrt{s} = 1.96$ TeV, using CTEQ5L [9] parton distribution functions and a detailed detector simulation based on GEANT [10]. To simulate the effect of multiple interactions in beam crossings, a Poisson-distributed minimum-bias event overlay, with an average of 0.8 events, is included for all events. To avoid an incorrect combination of cross sections among simulated $W/Z +$ jets samples, only events with the same number of reconstructed jets as the number of initial partons are retained. The background from multijet events, in which a jet is misidentified as a lepton, is evaluated using the “matrix method” as follows. Defining a “tight” $W +$ jets candidate sample that has the lepton-identification criteria described above, and a “loose” sample in which some of these identification criteria are relaxed, a comparison of the probabilities for true leptons to be identified as loose and jets as tight leptons (as determined from independent studies of samples of pure leptons and pure jets) yields the fractions of true leptons and of misidentified jets in the tight and loose samples.

After the $W$-boson and jet selections, we apply the two HF-tagging algorithms to the jets. The MC samples are normalized to the appropriate luminosity and corrected for differences in HF-tagging and lepton-identification efficiencies relative to data. Also, data collection inefficiencies from trigger requirements are introduced into the simulated samples based on independent data measurements for each set of selections.

In the following, we combine the $e +$ jets and $\mu +$ jets samples, and Figs. 1 and 2 show the exclusive number of jets in events with at least one SLT-tagged jet and at least one SVT-tagged jet, respectively. The transverse mass for $W$-boson candidates containing at least one SLT- or one SVT-tagged jet, shown in Fig. 3, agrees well with SM expectation. The distribution for events with at least one jet tagged with both algorithms is shown in Fig. 4.

The dominant sources of experimental uncertainty are as follows: (i) a 6.5% uncertainty on integrated luminosity, (ii) a 6% per jet uncertainty from JES corrections and jet identification, (iii) a 10% per jet uncertainty from the HF-tagging algorithms, and (iv) a 10%–18% uncertainty on the predicted MC cross sections (depending on the
sample). The total systematic uncertainties on $t\bar{t}$, single top-quark, and $W/Z +$ jets backgrounds are 16%, 21%, and 22%, respectively.

No excess is observed in the “doubly-tagged” jet sample. We therefore set a limit on the rate of anomalous HF-quark production in association with a $W$ boson. Because we do not propose a model for such production, we do not base this limit on any specific efficiency or jet spectrum. We quote limits on the number of events beyond SM expectation per exclusive jet bin. The 95% confidence level (C.L.) upper limits for additional event production in each bin are shown in Table I. These limits are calculated using a modified Frequentist (CL$_s$) method [11].

Assuming that anomalous HF production has the same event topology as certain SM processes, the above limits can be translated into limits on cross sections. To this end, we consider two benchmark scenarios:

(1) “$Wb\bar{b}$-like” production in which two $b$ quarks are produced in association with a $W$ boson. In this scenario, additional light quarks or gluons can be produced and thereby shift the event topology to more than two jets. Jets not within the acceptance of the detector can also cause the event topology to drop to less than two jets. We model this production using efficiencies for SM $W/Z + b\bar{b}$ production.

(2) “Toplike” production in which a heavy particle is produced and decays to a $W$ boson and a $b$ quark. An event can contain two such heavy particles (“$t\bar{t}$-like”) or one heavy quark (“single-top-like”), with additional light or heavy quarks and gluons possible for both cases. We model this scenario using the combined cross-section weighted efficiencies for SM $t\bar{t}$ and single top-quark production.

To calculate a limit on exclusive jet production for each case, we first ignore the probability for reconstructing the predicted number of jets, providing a model-independent comparison of processes with specific jet topologies. The remaining efficiency represents the effect of $W$-boson selection and HF tagging, and limits for specific models can be extracted by multiplying this value by the efficiency to reconstruct the number of jets in each exclusive jet bin. This is given in Table II.

To evaluate an upper cross-section limit on inclusive jet production for each scenario, we reintroduce the efficiency for reconstructing the predicted jets and sum the relevant exclusive jet bins as shown in Table III. For inclusive $Wb\bar{b}$-like anomalous production, the contribution from higher bins being negligible, we sum the first two $W +$ jets bins. For toplike anomalous production, we sum all $W +$ jets bins, except $n = 1$, where the contribution is again negligible. Table III shows the 95% C.L. event limits for the combinations of bins for these two hypotheses, and also the corresponding anomalous HF production limits. The jet reconstruction efficiency is included in the calculations, and the limits contain the expected efficiencies for the specified SM processes.

In summary, we observe no excess beyond the SM prediction for heavy-flavor quark production in association with $W$ bosons in 164 pb$^{-1}$ of data in $e +$ jets and 145 pb$^{-1}$ in $\mu +$ jets channels. Using a sample of events containing at least one jet tagged with both the SLT and SVT algorithms, we derive 95% C.L. limits on anomalous heavy-flavor production (Table I). Using benchmark SM processes, we also derive limits of 26.4 pb for $Wb\bar{b}$-like and 14.9 pb for toplike anomalous scenarios. For comparison, the D0 Collaboration has published a similar study in

![FIG. 4. Exclusive jet multiplicity of $W$-boson candidate events with at least one jet tagged with both the SVT and SLT algorithms. The fourth bin represents the sum of events containing four or more jets.](image-url)

TABLE I. The numbers of observed and predicted $W$-boson events with at least one jet tagged by both the SLT and SVT algorithms, as a function of exclusive jet multiplicity. Also shown are the 95% C.L. limits in the form of additional events.

<table>
<thead>
<tr>
<th>Source</th>
<th>1 jet</th>
<th>2 jets</th>
<th>3 jets</th>
<th>$\geq 4$ jets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data observation</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>SM prediction</td>
<td>5.0 ± 1.2</td>
<td>2.0 ± 0.5</td>
<td>1.0 ± 0.2</td>
<td>0.4 ± 0.1</td>
</tr>
<tr>
<td>95% C.L. limit (excess events)</td>
<td>6.7</td>
<td>3.9</td>
<td>4.1</td>
<td>3.0</td>
</tr>
</tbody>
</table>

TABLE II. Cross-section upper limits in pb, based on the hypotheses of $Wb\bar{b}$-like and toplike anomalous production of exclusive jets. Each value must still be corrected for the efficiency of reconstructing the predicted number of jets.

<table>
<thead>
<tr>
<th>Model</th>
<th>1 jet</th>
<th>2 jets</th>
<th>3 jets</th>
<th>$\geq 4$ jets</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Wb\bar{b}$-like</td>
<td>35.0</td>
<td>9.2</td>
<td>6.0</td>
<td>4.5</td>
</tr>
<tr>
<td>Toplike</td>
<td>12.6</td>
<td>8.0</td>
<td>11.3</td>
<td>15.4</td>
</tr>
</tbody>
</table>

TABLE III. Cross-section limits on heavy-flavor production of $W$ bosons in $e +$ jets and $\mu +$ jets channels. Using a sample of events containing at least one jet tagged with both the SLT and SVT algorithms, we derive 95% C.L. limits on anomalous heavy-flavor production (Table I). Using benchmark SM processes, we also derive limits of 26.4 pb for $Wb\bar{b}$-like and 14.9 pb for toplike anomalous scenarios. For comparison, the D0 Collaboration has published a similar study in
the form of a search for $Wb\bar{b}$ production [12]. Based on the two-jet topology, with both jets HF tagged, that study sets a 95% C.L. upper cross-section limit of 6.6 pb on $Wb\bar{b}$ production for jets with $p_T > 20$ GeV/c and $R(b, \bar{b}) > 0.75$.

We thank the staffs at Fermilab and collaborating institutions and acknowledge support from the Department of Energy and National Science Foundation (USA), Commissariat à l’Énergie Atomique and CNRS/Institut National de Physique Nucléaire et de Physique des Particules (France), Ministry of Education and Science, Agency for Atomic Energy and RF President Grants Program (Russia), CAPES, CNPq, FAPERJ, FAPESP, and FUNDUNESP (Brazil), Departments of Atomic Energy and Science and Technology (India), Colciencias (Colombia), CONACyT (Mexico), KRF (Korea), CONICET and UBACyT (Argentina), The Foundation for Fundamental Research on Matter (The Netherlands), PPARC (United Kingdom), Ministry of Education (Czech Republic), Canada Research Chairs Program, CFI, Natural Sciences and Engineering Research Council and WestGrid Project (Canada), BMBF and DFG (Germany), A. P. Sloan Foundation, Research Corporation, Texas Advanced Research Program, and the Alexander von Humboldt Foundation.

*Visitor from University of Zurich, Zurich, Switzerland.


<table>
<thead>
<tr>
<th>Source</th>
<th>1,2 jets</th>
<th>≥ 2 jets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data observation</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>SM prediction</td>
<td>6.9 ± 1.2</td>
<td>3.3 ± 0.5</td>
</tr>
<tr>
<td>95% C.L. limit</td>
<td>6.6</td>
<td>4.4</td>
</tr>
</tbody>
</table>

TABLE III. 95% C.L. limits for the number of events summed over the indicated jet bins. Also shown are cross-section limits based on the hypotheses of $Wb\bar{b}$-like and toplike anomalous production for the selected number of jets.