Measurement of the $B_s^0$ semileptonic branching ratio to an orbitally excited $D_{s1}^*$ state, $Br(B^0_s \to D_{s1}^{*+}(2536)\mu^+\nu\mu X)$
In a data sample of approximately 1.3 fb$^{-1}$ collected with the D0 detector between 2002 and 2006, the orbitally excited charm state $D_{s1}^{*+}(2536)$ has been observed with a measured mass of $2535.7 \pm 0.6$ (stat) $\pm 0.5$ (syst) MeV/c$^2$ via the decay mode $B^{0}_s \rightarrow D_{s1}^{*+}(2536)\mu^+\nu_X$. A first measurement is
Semileptonic $B^0$ decays into orbitally excited $P$-wave strange-charm mesons ($D^{**}$) are expected to make up a significant fraction of $B^0$ semileptonic decays and are therefore important when comparing inclusive and exclusive decay rates, extracting CKM matrix elements, and using semileptonic decays in $B^0$ mixing analyses. For $B$ meson semileptonic decays to heavier excited charm states, more of the available phase space is near zero recoil, increasing the importance of corrections in heavy-quark effective theory (HQET) [1], effectively tested here.

$D^{**}$ mesons (also denoted $D_{sJ}$) are composed of a heavy charm quark and a lighter strange quark in an $L = 1$ state of orbital momentum. In the heavy-quark limit, the spin $s_Q$ of the heavy quark and the total angular momentum, $j_Q = s_1 + L$ of the light degrees of freedom (quark and gluons), are separately conserved and the latter has possible values of $\frac{1}{2}$ or $\frac{3}{2}$. The surprisingly light masses of the $j_Q = \frac{3}{2}$ states: $D_{s1}^{**}$($2317$) and $D_{s1}(2460)$ [2], plus the observation of new $D_{sj}$ states [3], deepens the need for a better understanding of these $D^{**}$ systems since they may be quark molecular states, a new and very different arrangement of quarks.

In our decay of interest, the $j_Q = \frac{3}{2}$ angular momentum can combine with the heavy quark spin to form the $J^P = 1^+$ ($D_{s1}$) state which must decay through a $D$-wave to conserve angular momentum. $D_{s1}^{**}$($2536$) is expected to decay dominantly into a $D^*$ and $K$ meson to conserve angular momentum.

In this Letter we present the first measurement of semileptonic $B^0$ decay into the narrow $D_{s1}^{**}$($2536$) state. This state is just above the $D^*$ $K_S^0$ mass threshold and has been observed previously [4]. Events compatible with the decay chain $b \rightarrow D_{s1}^{**}$($2536$)$\mu^+\nu_\mu X$, $D_{s1}^{**}$($2536$) $\rightarrow D^* K_S^0$; $D^* \rightarrow D^0 \pi^-$, $K_S^0 \rightarrow \pi^+\pi^-$, $D^0 \rightarrow K^+\pi^-$ are reconstructed. Charge conjugate modes and reactions are always implied in this Letter.

Assuming that $D_{s1}^{**}$($2536$) production in a semileptonic decay is entirely from $B^0$, the branching ratio

$$Br(B^0 \rightarrow D_{s1}^{**}$($2536$)$\mu^+\nu_\mu X)$$

can be determined by normalizing to the known value of the branching fraction $Br(b \rightarrow D^* \mu^+\nu_\mu X) = (2.75 \pm 0.19)\%$ [5] to avoid uncertainties in the $b$-quark production rate. This semileptonic branching ratio includes any decay channel or sequence of channels resulting in a $D^*$ and a lepton (muon in our case), and all $b$ hadrons, and therefore includes the relative production of each $b$ hadron species starting from a $b$ quark. Since the final state of interest, $D_{s1}^{**}$($2536$) $\rightarrow D^* K_S^0$ is reconstructed from a $D^*$ and a $K_S^0$, the selection is broken up into two sections: one to reconstruct the $D^*$ with an associated muon, coming dominantly from $B$ meson decays resulting in a number of candidates, $N_{D^*\mu}$, and then the addition and subsequent formation of a vertex of a $K_S^0$ with the $D^*$ and muon, resulting in $N_{D^*K_S^0}$ candidates. To find the branching ratio, the following formula is used:

$$Br(b \rightarrow B_s^0) \cdot Br(B_s^0 \rightarrow D_{s1}^{**}$($2536$)$\mu^+\nu_\mu X) \cdot \frac{N_{D_{s1}^{**}}}{N_{D^*\mu}} \cdot \frac{1}{\epsilon(b \rightarrow D^*\mu) \cdot \epsilon(K_S^0)}.$$

The input $f(b \rightarrow B_s^0)$ [5] is the fraction of decays where a $b$ quark will hadronize to a $B_s^0$ hadron. $\epsilon(K_S^0)$ is the efficiency in the signal decay channel to reconstruct and make a vertex with a $K_S^0$ to form a $D_{s1}^{**}$($2536$), given that a $D^*$ and a muon have already been reconstructed. Later we will identify the ratio of efficiencies as $R_{D^*K_S^0} = \epsilon(B_s^0 \rightarrow D_{s1}^{**}$($2536$)$\mu^+\nu_\mu X)$/$\epsilon(b \rightarrow D^*\mu)$.

The D0 detector [6] and following analysis [7] are described in more detail elsewhere. The main elements relevant to this analysis are the silicon microstrip tracker (SMT), central fiber tracker (CFT), and muon detector systems.

This measurement uses a large data sample, corresponding to approximately 1.3 fb$^{-1}$ of integrated luminosity collected by the D0 detector between April 2002 and March 2006. Events were reconstructed using the standard D0 software suite. To avoid lifetime biases compared to the MC simulation, the small fraction of events were removed that entered the sample only via triggers that included requirements on impact parameters of tracks.

To evaluate signal mass resolution and efficiencies, Monte Carlo (MC) simulated samples were generated for signal and background. The standard D0 simulation and event reconstruction chain was used. Events were generated with the PYTHIA generator [8] and decay chains of heavy hadrons were simulated with the EVTGEN decay package [9]. The detector response was modeled by GEANT [10]. Two background MC samples were also generated: a $c\bar{c}$ sample, and an inclusive $b$-quark sample containing all $b$ hadron species with forced semileptonic decays to a muon. In both cases, all events containing both a $D^*$ and a muon were retained.

$B$ mesons were first selected using their semileptonic decays, $B \rightarrow D^* \mu^+\nu_X$. At this point in the selection, the $D^* + \mu$ sample is dominated by $B_s^0 \rightarrow D^* \mu^+\nu_\mu X$ decays. For this analysis, muons were required to have hits in more than one muon layer, to have an associated track
in the central tracking system, and to have transverse momentum $p_T > 2$ GeV/c, pseudorapidity $|\eta| < 2$, and total momentum $p^\mu > 3$ GeV/c. Two oppositely charged tracks with $p_T > 0.7$ GeV/c and $|\eta| < 2$ were required to form a common $D^0$ vertex which were then combined with a muon candidate to form a common decay point following the procedure described in Ref. [11]. For each $D^0\mu^+$ candidate, an additional soft pion was searched for with charge opposite to the charge of the muon and $p_T > 0.18$ GeV/c. The $K^-$ and $\pi^+$ from the decay of the $D^0$ were both required to have more than five CFT hits. To reduce the contribution from prompt $c\bar{c}$ production, a requirement was made on the transverse decay length, $L_{xy}$, significance of the $D^*\mu$ vertex of $L_{xy}/\sigma(L_{xy}) > 1$. After these cuts, the total number of $D^*$ candidates in the mass difference, $M(D^*) - M(D^0)$, peak of Fig. 1 is $N_{D^*\mu} = 87506 \pm 496$ (stat).

$D^{\pm}_{sl}(2536)$ candidates were formed by combining a $D^*$ candidate with a $K^{0}_{S}$. $D^*$ candidates were first selected by requiring the mass difference $M(D^*) - M(D^0)$ to be in the range 0.142–0.149 GeV/c². The two tracks from the decay of the $K^{0}_{S}$ were required to have opposite charge and to have more than five hits in the CFT detector. The $p_T$ of the $K^{0}_{S}$ was required to be greater than 1 GeV/c to reduce the contribution of background $K^{0}_{S}$ mesons from fragmentation. A vertex was then formed using the reconstructed $K^{0}_{S}$ and the $D^*$ candidate of the event. The decay length of the $K^{0}_{S}$ was required to be greater than 0.5 cm. To compute the $D^{\pm}_{sl}(2536)$ invariant mass, a mass constraint was applied using the known $D^*$ mass [5] instead of the measured invariant mass of the $K\pi\pi$ system. Finally, the invariant mass of the reconstructed $D^{\pm}_{sl}(2536)$ and muon was required to be less than the mass of the $B^{0}_{s}$ meson [5].

The signal model employed for the fit to the $D^*K^{0}_{S}$ invariant mass spectrum was a relativistic Breit-Wigner convoluted with a Gaussian function, with the resonance width fixed to the value 1.03 ± 0.05 (stat) ± 0.12 (syst) MeV/c² measured by the BaBar Collaboration [12] and a Gaussian width determined to be 2.8 MeV/c² from MC simulation of the signal. The MC width value was scaled up by a factor of 1.10 ± 0.10 to account for differences between data and MC resolution estimates. The unbinned likelihood fit used an exponential function plus a first-order polynomial to model the background with a threshold cutoff of $M(D^*) + M(K^{0}_{S})$. The fit, shown in Fig. 2, gives a central value for the mass peak of 2535.7 ± 0.7 (stat) MeV/c², a yield of $N_{D^*\mu} = 45.9 \pm 9.1$ (stat) events, and a significance of 6.1σ for the background to fluctuate up to or above the observed number of signal events.

The efficiencies used in Eq. 1 are estimated using the MC simulation, after implementing suitable correction factors to ensure proper modeling of the underlying $b$-hadron $p_T$ spectrum, as well as trigger effects. An event-by-event weight, applied as a function of the generated $p_T$ of the $B_s$, was determined by comparing the generated $p_T(B)$ in MC with the $p_T$ distribution of fully reconstructed $B^+ \rightarrow J/\psi K^+$ candidates in data collected primarily with a dimuon trigger [13]. Most events for this analysis were recorded using single muon triggers, and an additional weight was applied as a function of $p_T(\mu)$ to further improve the simulation of trigger effects. Reweighted MC events were used in the determination of efficiencies described below, and indicated uncertainties are due to MC statistics.

Using the MC sample of inclusive $b \rightarrow D^*\mu X$ events, specific major decay modes were identified. Efficiencies for each of these decay modes to pass the $D^*\mu$ selection, including the efficiency to reconstruct the soft pion from the $D^*$, were then determined. The predicted fraction $F_i$ of each channel contributing to the $D^*\mu$ sample before further cuts was found following a
procedure similar to that given in Ref. [14]. The efficiency $\epsilon_i$ for each channel was found and a weighted sum was calculated, giving an estimated total efficiency for reconstruction of $\epsilon(b \rightarrow D^*\mu) = (5.88 \pm 0.80)\%$, where the uncertainty is dominated by the MC statistics used to find $\epsilon_i$ and uncertainties on external inputs [5] used to estimate $F_i$. Applying the same cuts for reconstructing the $D^*\mu$ for the signal channel, the efficiency $\epsilon(B_0^\pm \rightarrow D_{S1}\mu \rightarrow D^*\mu) = (3.20 \pm 0.02)\%$, results in a ratio of efficiencies of $R_{D^*}^{\text{en}} = 0.547 \pm 0.075$.

The signal MC sample was used to determine the efficiency to reconstruct $D_{S1}^{-}(2536) \rightarrow D^+K_S^0$ given a reconstructed $D^*\mu$ as a starting point. This efficiency is hence effectively that of reconstructing a $K_S^0 \rightarrow \pi^+\pi^-$ and forming a vertex with the $D^*\mu$, and includes the branching ratio $Br(K_S^0 \rightarrow \pi^+\pi^-)$ [5] for ease of use in calculating the branching ratio product. The reconstruction efficiency was found to be $\epsilon_{K_S^0} = (10.3 \pm 0.4)\%$ where the uncertainty is due to MC statistics.

The process $c\bar{c} \rightarrow D^+\mu^+\nu\mu X$ can contribute to $N_{D^*\mu}$ since a $D^*$ meson can come from the hadronization of the $c$ quark, and the muon can come from the semileptonic decay of the hadron containing the $c$ quark. To determine the number of events in our signal reconstructed from a prompt $D^*$, a comparison was made of the decay length significance distribution observed in the data with the same distribution predicted by MC for $b \rightarrow D^*\mu X$ and any excess at shorter significances was interpreted as $c\bar{c}$ contribution. For the decay length significance cut used in the analysis, $L_{xy}/\sigma(L_{xy}) > 1$, the fraction of $N_{D^*\mu}$ from $c\bar{c}$ production was estimated to be $(3.9 \pm 2.5)\%$. A check using a prompt $c\bar{c}$ MC sample results in a consistent estimate. The value of $N_{D^*\mu}$ was corrected downward accordingly.

The contribution from $c\bar{c}$ production to $N_{D_{S1}}$ where one charm quark hadronizes directly to a $D_{S1}(2536)$ and the other decays directly to a muon was estimated to be negligible using relative production ratios and spin-counting arguments [15].

Systematic uncertainties for the branching ratio product are summarized in Table I and discussed below. The uncertainty in the normalizing branching ratio [5] $Br(b \rightarrow D^*\mu X)$ was taken as a systematic uncertainty. For determining $N_{D^*\mu}$, the signal and background model parameters were varied in a correlated fashion and a systematic uncertainty was assigned. The estimated $c\bar{c}$ production contribution was varied by the indicated uncertainty. In the determination of $N_{D_{S1}}$, the functional forms of the signal and background models were varied in a number of ways to determine the sensitivity of the candidate yield. In addition, the scaling of the widths was varied by $\pm 10\%$ to check the sensitivity to uncertainty in mass resolution.

By comparing the $p_T(\mu)$ distribution for the signal using the default ISGW2 decay model [16] to the HQET semileptonic decay model [9], a weighting factor was found and applied to the fully simulated signal MC events, and the efficiency determined again. The difference observed was assigned as a contribution to the systematic uncertainty of $\epsilon_{K_S^0}$ and $R_{D^*}^{\text{en}}$.

When estimating $\epsilon_{K_S^0}$, the uncertainty due to modeling of the $b$ hadron $p_T$ spectrum was derived by using an alternate weighting technique. The cuts on the $p_T$ and decay length of the $K_S^0$ were varied and a systematic uncertainty on the efficiency due to this source was also assigned. Discrepancies in track reconstruction efficiencies between data and MC in low-$p_T$ tracks were accounted for by assigning a systematic uncertainty to each of the pion tracks in the $K_S^0$ reconstruction [17, 18].

The uncertainty in $R_{D^*}^{\text{en}}$ is due to a combination of MC statistics and uncertainties in PDG branching ratio values and production fractions, $f(b \rightarrow b\text{hadron})$. The uncorrelated systematic uncertainty is given in Table I.

The estimated systematic uncertainties were added in quadrature to obtain a total estimated systematic uncertainty on the branching ratio product of $16.8\%$. The branching ratio product was determined to be:

$$f(b \rightarrow B_s^0) \cdot Br(B_s^0 \rightarrow D_{S1}\mu^+\nu\mu X) \cdot Br(D_{S1}^{-} \rightarrow D^+K_S^0) = (2.66 \pm 0.52 \text{ (stat)} \pm 0.45 \text{ (syst)} \times 10^{-4}.)$$

<table>
<thead>
<tr>
<th>Source</th>
<th>Systematic uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Br(b \rightarrow D^\mu X)$</td>
<td>6.9%</td>
</tr>
<tr>
<td>$N_{D^*\mu}$</td>
<td>2.9%</td>
</tr>
<tr>
<td>$N_{D_{S1}}$</td>
<td>5.5%</td>
</tr>
<tr>
<td>$\epsilon_{K_S^0}$</td>
<td>11.0%</td>
</tr>
<tr>
<td>$R_{D^*}^{\text{en}}$</td>
<td>8.6%</td>
</tr>
<tr>
<td>Total</td>
<td>16.8%</td>
</tr>
</tbody>
</table>

To assess the systematic uncertainty on the mass measurement, the same variations of the $D_{S1}(2536)$ mass signal model, as well as background functional form, were applied as described above. The mass values used for the mass constraints on the decay products were varied within their PDG uncertainties and were also set to the D0 central fit values. Ensemble tests indicated that the statistical error is correct. From the observed variations, a total systematic mass uncertainty of $0.5\ MeV/c^2$ was taken, for a mass measurement of:

$$m(D_{S1}) = 2535.7 \pm 0.6 \text{ (stat)} \pm 0.5 \text{ (syst)} \text{ MeV/c}^2.$$

This measured mass value is in good agreement with the PDG average value of $2535.34 \pm 0.31 \text{ MeV/c}^2$ [5].

To allow comparison of this measurement to theoretical predictions, the semileptonic branching ratio alone as shown in Table II is extracted by taking the hadronization fraction into $B_s^0$ as $f(b \rightarrow B_s^0) = 0.103 \pm 0.014$ [5] and also assuming that $Br(D_{S1}(2536) \rightarrow D^*K_S^0) =$
0.25 [9]. This is the first experimental measurement of this semileptonic branching ratio and is compared to a number of theoretical predictions [1, 19, 20] of the exclusive rate in Table II. The systematic uncertainty on this quantity is as described earlier, and the error labeled \(\text{"prod. frac."}\) is due to the current uncertainty on \(f(b\to B^0)\). The first two theoretical predictions include relativistic and \(1/m_Q\) corrections, while the third does not. The result is found to be consistent within uncertainties with the first two theoretical predictions, and demonstrates the need for such corrections.

### Table II: Experimental measurement compared with various theoretical predictions.

<table>
<thead>
<tr>
<th>Source</th>
<th>(Br(B_d^0\to D_{s1}^{\pm}(2536)\mu^-\nu_{\mu}X))</th>
</tr>
</thead>
<tbody>
<tr>
<td>This result</td>
<td>(1.03 \pm 0.20, \text{(stat)} \pm 0.17, \text{(syst)} \pm 0.14, \text{(prod. frac.)})%</td>
</tr>
<tr>
<td>Theoretical Predictions</td>
<td>(Br(B_d^0\to D_{s1}^{\pm}(2536)\mu^-\nu_{\mu}))</td>
</tr>
<tr>
<td>ISGW2 [1]</td>
<td>(0.53 \pm 0.27)%</td>
</tr>
<tr>
<td>Relativistic Quark Model &amp; (1/m_Q) corrections [19]</td>
<td>(1.06 \pm 0.16)%</td>
</tr>
<tr>
<td>Non-rel. HQET and ISGW [20]</td>
<td>0.195%</td>
</tr>
</tbody>
</table>

In summary, using 1.3 fb\(^{-1}\) of integrated luminosity collected with the D0 detector, a first measurement of the semileptonic \(B_d^0\) decay into the narrow \(D_{s1}^{\pm}(2536)\) state has been made and compared with theory. In addition, the mass of the \(D_{s1}^{\pm}(2536)\) was measured and found to be in good agreement with the PDG value.

We thank the staffs at Fermilab and collaborating institutions, and acknowledge support from the DOE and NSF (USA); CEA and CNRS/IN2P3 (France); FASI, Rosatom and RFBR (Russia); CAPES, CNPq, FAPERJ, FAPESP and FUNDUNESP (Brazil); DAE and DST (India); Colciencias (Colombia); CONACyT (Mexico); KRF and KOSEF (Korea); CONICET and UBACyT (Argentina); FOM (The Netherlands); Science and Technology Facilities Council (United Kingdom); MSMT and GACR (Czech Republic); CRC Program, CFI, NSERC and WestGrid Project (Canada); BMBF and DFG (Germany); SFI (Ireland); The Swedish Research Council (Sweden); CAS and CNSF (China); Alexander von Humboldt Foundation; and the Marie Curie Program.

---

[a] Visitor from Augustana College, Sioux Falls, SD, USA.
[b] Visitor from The University of Liverpool, Liverpool, UK.
[c] Visitor from ICN-UNAM, Mexico City, Mexico.
[d] Visitor from H. Physikalisches Institut, Georg-August-University, Göttingen, Germany.
[e] Visitor from Helsinki Institute of Physics, Helsinki, Finland.
[f] Visitor from Universität Zürich, Zürich, Switzerland.
[g] Visitor from Lewis University, Romeoville, IL, USA.
[h] Fermilab International Fellow.
[i] Deceased.