



**First Measurement of the  $B_s^0$  Semileptonic Branching Ratio to  
an Orbitally Excited  $D_{s1}^{**}$  State,  $Br(B_s^0 \rightarrow D_{s1}^-(2536)\mu^+\nu X)$**

The DØ Collaboration  
URL <http://www-d0.fnal.gov>  
(Dated: March 13, 2006)

In a data sample of approximately  $1.0 \text{ fb}^{-1}$  collected with the DØ detector between 2002 and 2005, the orbitally excited charm state  $D_{s1}^\pm(2536)$  has been observed with a measured mass of  $2535.7 \pm 0.6 \text{ (stat.)} \pm 0.5 \text{ (syst.) MeV}/c^2$  via the decay mode  $B_s^0 \rightarrow D_{s1}^-(2536)\mu^+\nu X$  followed by  $D_{s1}^\pm(2536) \rightarrow D^{*\pm}K_S^0$  with  $D^{*+} \rightarrow D^0\pi^+$ ,  $D^0 \rightarrow K^-\pi^+$  and  $K_S^0 \rightarrow \pi^+\pi^-$ . By normalizing to the known branching ratio  $Br(\bar{b} \rightarrow D^{*-}\mu^+\nu X)$  and to the number of reconstructed  $D^*$  mesons with an associated identified muon, a first ever measurement is made of the product branching ratio  $f(\bar{b} \rightarrow B_s^0) \cdot Br(B_s^0 \rightarrow D_{s1}^-(2536)\mu^+\nu X) \cdot Br(D_{s1}^- \rightarrow D^{*-}K_S^0)$  as well as an extraction of the semileptonic branching ratio  $Br(B_s^0 \rightarrow D_{s1}^-(2536)\mu^+\nu X)$ . Comparisons are made with expectations.

*Preliminary Results for Winter 2006 Conferences*

## I. INTRODUCTION

Semileptonic  $B_s^0$  decays into orbitally excited  $P$ -wave strange-charm mesons ( $D_s^{**}$ ) are of interest for several reasons. They make up a significant fraction of  $B_s$  semileptonic decays and are hence important when comparing inclusive and exclusive decay rates, extracting CKM matrix elements, and using semileptonic decays in  $B_s^0$  mixing analyses. The semileptonic  $B$  decay rate to an excited charm meson is determined by the corresponding matrix elements of the weak axial-vector and vector currents. At zero recoil (where the final excited charm meson is at rest in the rest frame of the initial  $B$  meson), these currents correspond to conserved quantities of the heavy quark spin-flavor symmetry. For  $B$  semileptonic decays to heavier excited charm states, most of the available phase space is near zero recoil, increasing the importance of corrections in heavy quark effective theory (HQET). Measured decay properties can then be compared with theoretical HQET predictions.

$D_s^{**}$  mesons (sometimes denoted  $D_{sJ}$ ) are composed of a charm and strange quark in a  $L = 1$  state of orbital momentum, i.e.,  $P$ -wave. In the limit  $m_c \gg \Lambda_{QCD}$ , the quarks in this state have well defined quantum numbers, with  $L = 1$  and  $S = \frac{1}{2}$ . Hence the total angular momentum (spin + orbital) of the light degrees of freedom (i.e., the “brown muck” of the lighter strange quark plus all the gluons binding the state) can be labeled by  $j_q = \frac{1}{2}$  or  $\frac{3}{2}$  and the spin of the heavy quark can be taken as separately conserved. The  $j_q = \frac{3}{2}$  angular momentum then combines with the heavy quark spin to form two states with  $J^P = 1^+$  ( $D_{s1}$ ) and  $J^P = 2^+$  ( $D_{s2}^*$ ). Being a  $J^P = 1^+$  state, the  $D_{s1}^\pm(2536)$  can decay only into a  $D^*$  ( $J^P = 1^-$ ) and  $K$  meson ( $J^P = 0^-$ ) to conserve angular momentum and parity in a  $D$ -wave decay (relative angular momentum  $L = 2$ ). Due to the angular momentum barrier, these states have narrow widths for decays into a  $D^*$  and a  $K$  meson.

In this note we present the first measurement of semileptonic  $B_s^0$  decay into the narrow charged  $D_{s1}^\pm(2536)$  state, which is just above the  $D^* K_S^0$  mass threshold, has been observed in the past by other collaborations [1]. Events compatible with the decay chain:

$$\bar{b} \rightarrow B_s^0 \rightarrow D_{s1}^-(2536)\mu^+\nu X, D_{s1}^-(2536) \rightarrow D^{*-}K_S^0, D^{*-} \rightarrow D^0\pi^-, D^0 \rightarrow K\pi, K_S^0 \rightarrow \pi^+\pi^-.$$

are reconstructed. Charge conjugate modes and reactions are always implied in this note. Due to its predicted decay channels (see Table I and Ref. [2]) almost no contribution is expected to be seen by the other doublet member,  $D_{s2}^\pm(2573)$  in the same final state channel of  $D^* K_S^0$  plus an associated muon.

TABLE I: Expected decay fractions of  $j_q = 3/2$  doublet [2]. The reference gives the indicated branching ratios to  $K_S^0$ ; these are in error and should be branching ratios to  $K^0$ .

State	Decay Products	Fraction
$D_{s1}^\pm(2536)$	$D^{*+}K^0$	0.5
	$D^{*0}K^+$	0.5
$D_{s2}^\pm(2573)$	$D^{*+}K^0$	0.05
	$D^{*0}K^+$	0.05
	$D^+K^0$	0.43
	$D^0K^+$	0.47

Finally, for  $j_q = \frac{1}{2}$ , there are two states with  $J^P = 0^+$  ( $D_{s0}$ ) and  $J^P = 1^+$  ( $D_{s1}^*$ ). These decay via  $S$ -wave and are normally expected to have wide decay widths. However, the recently discovered [3] particles,  $D_{sJ}(2317)$  and  $D_{sJ}(2460)$ , that are usually assigned to these states are surprisingly light (compared to predictions [4]), observed below the  $DK$  and  $D^*K$  threshold and hence also narrow. Aside from the quark-antiquark interpretation, the  $D_{sJ}(2317)$  has been interpreted as a  $DK$  molecule, a  $D_s\pi$  molecule, or a four-quark state. Although the measurement of the decay angular distribution [5] does increase the likelihood of this particle assignment and decrease the possibility of such exotic states, their light mass, plus the observation (still unconfirmed) of the  $D_{sJ}(2632)$  by SELEX [6], still deepens the need for better understanding of these  $D_s^{**}$  systems.

## II. ANALYSIS OVERVIEW

The branching ratio  $Br(B_s^0 \rightarrow D_{s1}^-(2536)\mu^+\nu X)$  can be determined by normalizing to the known value of the branching fraction  $Br(\bar{b} \rightarrow D^{*-}\mu^+\nu X) = (2.75 \pm 0.19)\%$  [7]. This semileptonic branching ratio includes any decay channel or sequence of channels resulting in a  $D^*$  and a lepton (muon in our case), and is over all  $b$  hadrons, and therefore includes the relative production of each  $b$  hadron species starting from a  $\bar{b}$  quark. Since the final state of

interest,  $D_{s1}^{\pm}(2536) \rightarrow D^{*\pm}K_S^0$ , is taking a reconstructed  $D^*$  and combining it with a reconstructed  $K_S^0$ , the selection is broken up into two sections: one to reconstruct  $D^*$  with an associated  $\mu$ , coming dominantly from  $B$  meson decays, and then the addition and vertexing of a  $K_S^0$  with the  $D^*$  and muon. To find the branching ratio, the following formula is used:

$$f(\bar{b} \rightarrow B_s^0) \cdot Br(B_s^0 \rightarrow D_{s1}^-(2536)\mu^+\nu X) \cdot Br(D_{s1}^-(2536) \rightarrow D^{*-}K_S^0) = \\ Br(\bar{b} \rightarrow D^{*-}\mu^+\nu X) \cdot \frac{N_{D_{s1}(2536)}}{N_{D^*\mu}} \cdot \frac{\epsilon(\bar{b} \rightarrow D^*\mu)}{\epsilon(B_s^0 \rightarrow D_{s1}\mu \rightarrow D^*\mu)} \cdot \frac{1}{\epsilon_{K_S^0}}.$$

The input  $f(\bar{b} \rightarrow B_s^0) = 0.107 \pm 0.011$  [7] is the fraction of time that a  $b$  quark will hadronize to a  $B_s^0$  hadron, and as described above,  $Br(\bar{b} \rightarrow D^{*-}\mu^+\nu X) = (2.75 \pm 0.19)\%$  [7].

$\epsilon_{K_S^0}$  is the efficiency in the signal decay channel to additionally reconstruct and vertex a  $K_S^0$  to form a  $D_{s1}^{\pm}(2536)$  once a  $D^* + \mu$  have already been reconstructed. Finally, we will identify the ratio of efficiencies later as:  $R_{D^*}^{\text{gen}} = \epsilon(B_s^0 \rightarrow D_{s1}\mu \rightarrow D^*\mu)/\epsilon(\bar{b} \rightarrow D^*\mu)$ , i.e., the numerator is the efficiency in the decay channel for reconstructing a  $D^* + \mu$ , while the denominator is the efficiency to reconstruct  $D^* + \mu$  using identical cuts given that a  $b$  quark decays through into a channel or sequence of channels ultimately giving  $D^* + \mu$ . Neither of these efficiencies include any  $K_S^0$  selection requirements. The benefit of a normalization done this way is that the only absolute efficiency needed from Monte Carlo is that of  $\epsilon_{K_S^0}$  which covers just a fraction of the total efficiency for this state. Absolute efficiencies for muon identification,  $D^*$  reconstruction, triggering to get into the single muon sample, etc., are not needed, and uncertainties in the efficiencies tend to cancel in the ratio.

### III. DØ DETECTOR

The DØ detector is described in detail elsewhere [8]. The main elements relevant to this analysis are the central tracking and muon detector systems. The central tracking system consists of a silicon microstrip tracker (SMT) and a central fiber tracker (CFT), both located inside of a 2 T superconducting solenoidal magnet. Located outside the calorimeter, the muon detector [9] consists of a layer of tracking detector and scintillation trigger counters in front of 1.8 T toroidal magnets, and followed by two more similar layers behind the toroids. There is efficient muon detection out to approximately  $|\eta| < 2$ .

## IV. EVENT SAMPLE AND SELECTION

### A. Data Sample

This measurement uses the large preselected single muon data sample corresponding to approximately  $1.0 \text{ fb}^{-1}$  of integrated luminosity collected by the DØ detector between April 2002 and October 2005. Events were reconstructed using the standard DØ software suite [10] without particular trigger selections; however, the majority of the selected events satisfied single muon trigger requirements. Information only from the muon and tracking systems was used in this analysis.

Evidence of  $D_{s1}^{\pm}(2536)$  mesons was found in decays of  $B \rightarrow \mu\nu D_s^{**}X$  as resonances in the  $D^{*+}K_S^0$  invariant mass spectrum.  $D$  mesons were required to decay subsequently to  $D^{*+} \rightarrow D^0\pi^+$ ,  $D^0 \rightarrow K^-\pi^+$  and  $K_S^0 \rightarrow \pi^+\pi^-$ .

The event selections are described below.  $B$  mesons are first selected using their semileptonic decays,  $B \rightarrow \bar{D}^0\mu^+X$ , followed by finding  $D^*$  mesons in  $B \rightarrow D^{*-}\mu^+X$ . This selection is a mostly a standard one, used by the DØ analysis measuring the  $B^+/B_d^0$  lifetime ratio [11] and  $B_d^0$  oscillations [12]. At this point, the  $D^* + \mu$  sample is dominated by  $B_d^0 \rightarrow D^{*-}\mu^+\nu X$  decays before a  $D_{s1}^{\pm}(2536)$  selection is made.

### B. Muon Selection

Muons were identified using standard DØ criteria [13]. For this analysis, muons were required to have hits in more than one muon chamber, to have an associated track in the central tracking system with at least one hit in both SMT and CFT present, and to have transverse momentum  $p_T^\mu > 2 \text{ GeV}/c$ , pseudo-rapidity  $|\eta^\mu| < 2$ , and total momentum  $p^\mu > 3 \text{ GeV}/c$ .

All charged particles in the event were clustered into jets using the DURHAM clustering algorithm [14]. Events with more than one identified muon in the same jet were rejected, as well as the events with an identified  $J/\psi \rightarrow \mu^+\mu^-$  decay.

### C. $D^0$ Selection

The  $\bar{D}^0$  candidate was constructed from two particles of opposite charge included in the same jet as the reconstructed muon. Both particles should have hits in SMT and CFT, transverse momentum  $p_T > 0.7$  GeV/ $c$ , and pseudo-rapidity  $|\eta| < 2$ . They were required to form a common  $D$ -vertex with fit  $\chi_D^2 < 9$ . For each particle, the axial [28]  $\epsilon_T$  and stereo [29]  $\epsilon_L$  projections of track impact parameter with respect to the primary vertex together with the corresponding errors ( $\sigma(\epsilon_T)$ ,  $\sigma(\epsilon_L)$ ) were computed. The combined significance  $\sqrt{(\epsilon_T/\sigma(\epsilon_T))^2 + (\epsilon_L/\sigma(\epsilon_L))^2}$  was required to be greater than 2. The distance  $d_T^D$  between the primary and  $D$  vertex in the axial plane was required to exceed 4 standard deviations:  $d_T^D/\sigma(d_T^D) > 4$ . The angle  $\alpha_T^D$  between the  $\bar{D}^0$  momentum and the direction from the primary to the  $\bar{D}^0$  vertex in the axial plane was required to satisfy the condition:  $\cos(\alpha_T^D) > 0.9$ .

The tracks of muon and  $\bar{D}^0$  candidate were required to form a common  $B$ -vertex with fit  $\chi_B^2 < 9$ . The momentum of the  $B$  candidate was computed as the sum of momenta of the  $\mu$  and  $\bar{D}^0$ . The mass of the  $(\mu^+ \bar{D}^0)$  system was required to fall within  $2.3 < M(\mu^+ \bar{D}^0) < 5.2$  GeV/ $c^2$ . If the distance  $d_T^B$  between the primary and  $B$  vertices in the axial plane exceeded  $4 \cdot \sigma(d_T^B)$ , the angle  $\alpha_T^B$  between the  $B$  momentum and the direction from primary to  $B$  vertex in the axial plane was demanded to satisfy the condition  $\cos(\alpha_T^B) > 0.95$ . The distance  $d_T^B$  was allowed to be greater than  $d_T^D$ , provided that the distance between  $B$  and  $D$  vertices  $d_T^{BD}$  was less than  $3 \cdot \sigma(d_T^{BD})$ .

The masses of the kaon and pion were assigned to particles according to the charge of the muon, requiring  $\mu^+ K^+ \pi^-$  final system. In the following the events falling into the  $K\pi$  invariant mass window between 1.75 and 1.95 GeV/ $c^2$  will be referred to as  $\mu^+ \bar{D}^0$  candidates.

### D. $D^*$ Selection

For each  $\mu^+ \bar{D}^0$  candidate, we search for an additional slow pion ( $\pi^*$ ) with charge opposite to the charge of muon and with  $p_T > 0.18$  GeV/ $c$ . The mass difference  $\Delta M = M(\bar{D}^0 \pi) - M(\bar{D}^0)$  for all such pions when  $1.75 < M(\bar{D}^0) < 1.95$  GeV/ $c^2$  is shown in Fig. 1. To reduce the contribution from  $c\bar{c}$ , particularly from gluon splitting, where one charm quark fragments to  $D^*$  and the other to a meson that subsequently decays to a muon, a requirement was placed on the decay length significance of the  $D^* \mu$  vertex of  $L/\sigma(L) > 1$ . The effect of this requirement is discussed later. The peak corresponding to the production of  $D^* \mu$  is clearly seen. The signal and the background have been modeled by a sum of two Gaussian functions with the same mean and by the sum of exponential and first-order polynomial functions, respectively.

### E. $D_{s1}^\pm(2536)$ Selection

The  $D_{s1}^\pm(2536)$  was reconstructed through the decay channel  $D_{s1}^\pm(2536) \rightarrow D^{*\pm} K_S^0$ .  $D_{s1}^\pm(2536)$  candidates were formed by combining a  $D^*$  candidate with a  $K_S^0$ .  $D^*$  candidates were first selected from a mass difference window of  $0.142 < (M(D^*) - M(D^0)) < 0.149$  GeV/ $c^2$ . The two tracks from the decay of the  $K_S^0$  were required to have opposite charge and to have more than 5 hits in the CFT detector. The  $p_T$  of the  $K_S^0$  was required to be greater than 1 GeV/ $c$  to reduce the contribution of background fragmentation  $K_S^0$  mesons. A vertex was then formed using the reconstructed  $K_S^0$  and the  $D^*$  candidate of the event with a loose requirement of  $\chi^2 < 100$  on the vertex. The decay length of the  $K_S^0$  was required to be greater than 0.5 cm. This cut results in a loss of 19% of the  $K_S^0$  signal, but 52% of the background (mostly arising from the primary vertex) is also eliminated. The  $K^+$  and  $\pi$  from the decay of the  $D^0$  were also both required to have more than 5 CFT hits. Finally, it is required that the invariant mass of the reconstructed  $D_{s1}^\pm(2536)$  and muon is less than the mass of the  $B_s^0$  meson [7]. Combinatoric background, background from  $c\bar{c}$ , and fragmentation  $K_S^0$  mesons will often be at large angles with respect to the muon, pushing this invariant mass above  $M(B_s^0)$ . The invariant mass of  $\pi^+ \pi^-$  candidates in events with reconstructed  $D^* \mu$  candidates and passing the cuts above is shown in Fig. 2, with a fitted yield of  $1817 \pm 54$   $K_S^0$  where  $K_S^0$  candidates are defined as falling inside a mass window of  $0.47 < M(K_S^0) < 0.52$  GeV/ $c^2$ .

To compute the  $D_{s1}^\pm(2536)$  invariant mass, a mass constraint was applied using the PDG value [7] of  $M(D^*) = 2010.0$  MeV for the  $D^*$  mass instead of the invariant mass of the  $K\pi\pi$  system. A mass constraint of  $M(K_S^0) = 497.65$  MeV [7] was also placed on the  $K_S^0$ . The resulting spectrum and a fit to the  $D_{s1}(2536)$  mass peak is shown later in Section VI.

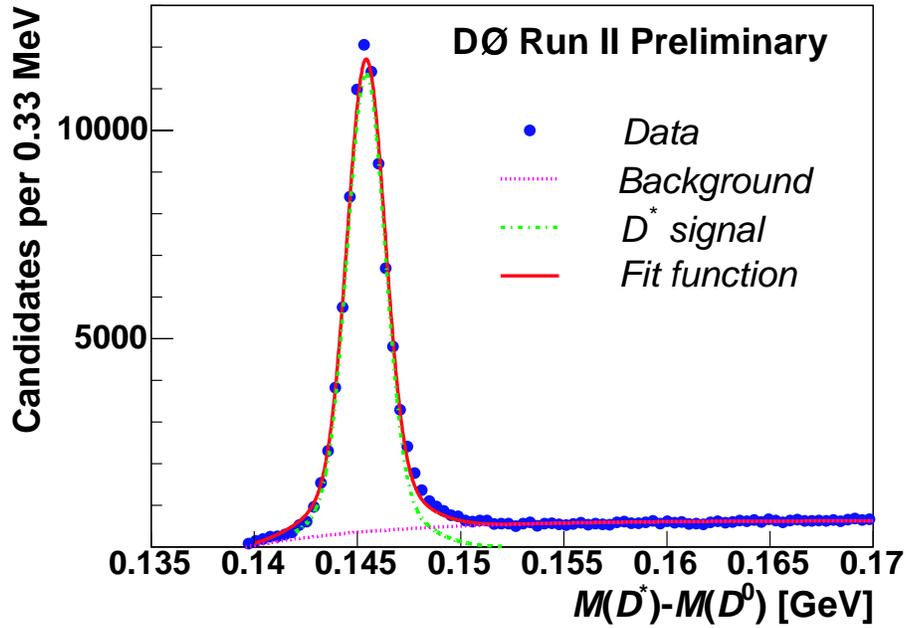


FIG. 1: The mass difference  $M(D^0\pi) - M(D^0)$  for events with  $1.75 < M(D^0) < 1.95 \text{ GeV}/c^2$ . The total number of  $D^*$  candidates is equal to be  $82130 \pm 463$  (stat.) and was defined as the number of signal events in the  $[0.142-0.149 \text{ GeV}]$  mass difference window. In the fit function the signal and the background have been approximated respectively by the sum of two Gaussian functions and by the sum of an exponential and first-order polynomial function.

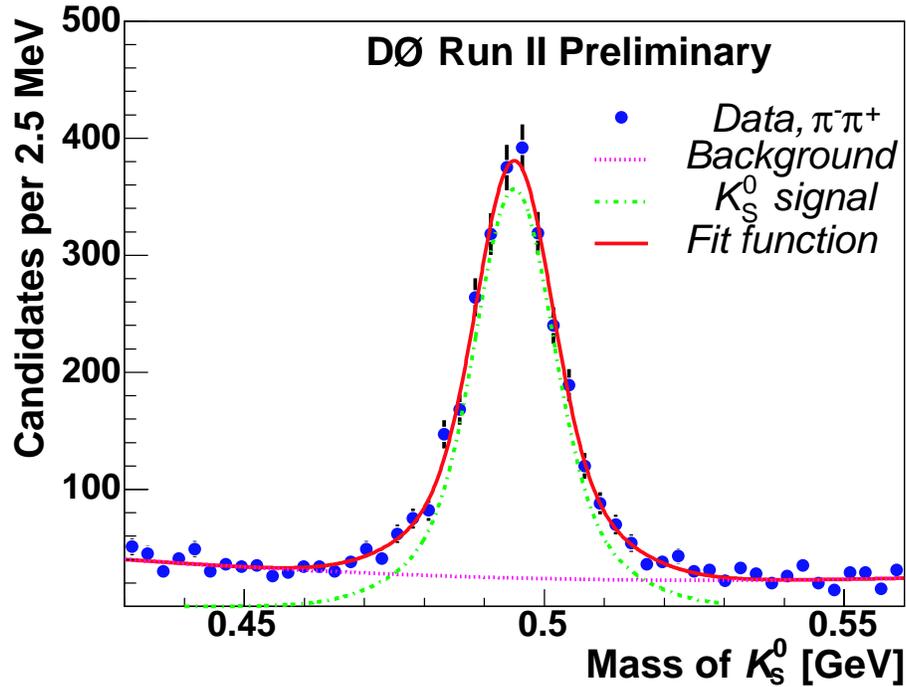


FIG. 2: Mass of  $\pi^+\pi^-$  for events after passing  $D^*$  and  $D^0$  cuts. The mass plot was fitted with a double Gaussian modeling the signal and a second-order polynomial for the background.

## V. MONTE CARLO STUDIES

To evaluate signal mass resolution, efficiencies, and efficiency resolutions, Monte Carlo samples were generated for signal and background. The standard  $D\bar{O}$  simulation chain was used that included the PYTHIA generator [15] interfaced with the EVTGEN decay package [2] followed by full GEANT [16] modeling of the detector response and event reconstruction.

The full decay path of the signal was generated [17] using the default decays of the EVTGEN package, in this case, the ISGW2 semileptonic decay model [18] for the  $B_s^0 \rightarrow D_{s1}(2536)\mu\nu$  decay, and using VVS.PWAVE [2] for the decay  $D_{s1}(2536) \rightarrow D^*K_S^0$ . Applying the same analysis cuts to the signal MC sample, the mass peaks of the intermediate and final candidates are shown in Fig. 3. A fit of a double Gaussian to the  $D_{s1}(2536)$  mass peak results in widths of 2.85 and 5.8 MeV/ $c^2$ , with the narrower Gaussian having a normalization fraction of 29%, and a mean value of  $2535.5 \pm 0.1$  MeV/ $c^2$  compared to the input PDG value of 2535.34 MeV/ $c^2$ .

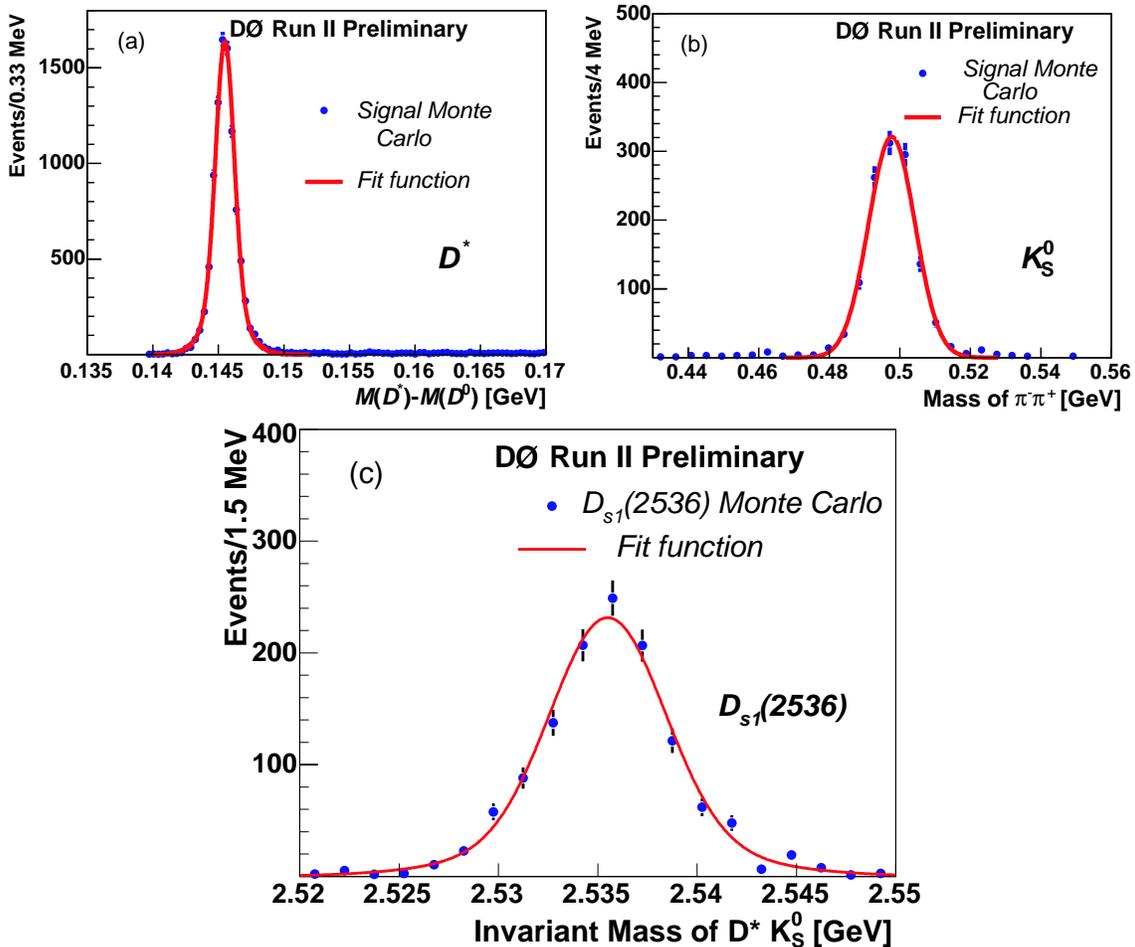


FIG. 3: Mass peaks as reconstructed in the  $B_s^0 \rightarrow D_{s1}(2536)\mu\nu$  signal MC sample showing the (a)  $D^*$ , (b)  $K_S^0$ , and (c)  $D_{s1}(2536)$  mass peaks following analysis cuts.

The background MC sample [19] was an inclusive sample consisting of  $b$  quarks hadronizing to all  $B$  meson species, forcing semileptonic decays to a muon and then retaining all events with decay paths of the  $B$  hadron containing a  $D^*$  meson.

## VI. RESULTS

### A. Number of $D^* + \mu$ Candidates

The total number of  $D^*$  candidates in the peak of Fig. 1 is equal to  $N_{D^*\mu} = 82130 \pm 463$  (stat.), and was defined as number of signal events fit in the  $[0.142\text{--}0.149\text{ GeV}]$  mass difference window.

### B. Number of $D_{s1}^\pm(2536)$ Candidates

The signal model employed for the fit to the  $D^*K_S^0$  invariant mass spectrum was a double Gaussian with width parameters and fraction of each Gaussian determined from MC studies of the previous section. However, mass resolutions predicted by the MC are typically underestimated by 10–20%, and in this case, the MC width values of 2.85 and 5.8 MeV were scaled up by a factor of  $1.10 \pm 0.10$  (with the effect of the variation taken later as a systematic error). The fit used an exponential function plus a second-order polynomial to model the background, and a common threshold cutoff of  $M(D^*) + M(K_S^0)$  was applied. The fit, as shown in Fig. 4, gives a central value for the Gaussian of  $2535.7 \pm 0.7(\text{stat.})\text{ MeV}/c^2$ , a yield of  $N_{D_{s1}(2536)} = 43.8 \pm 8.3$  events, and a calculated significance of  $5.2\sigma$ . The error is statistical only.

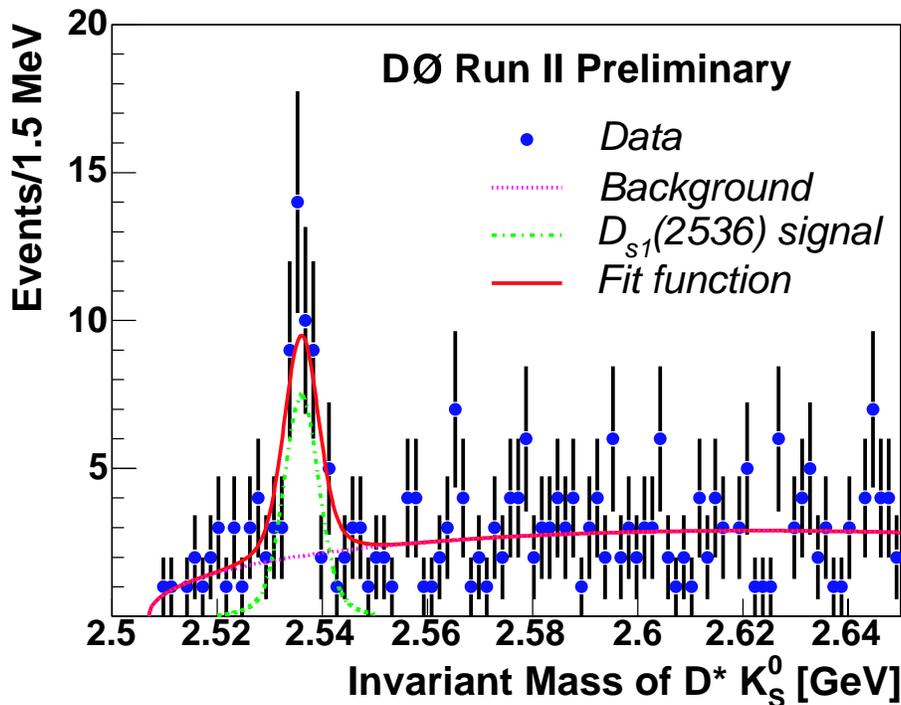


FIG. 4: Invariant mass of  $D^*K_S^0$ . Shown is the result of the fit of the  $D^*K_S^0$  mass with an exponential plus polynomial function with a threshold cutoff at  $M(D^*) + M(K_S^0)$ . The total number of  $D_{s1}(2536)$  candidates in the peak is  $43.8 \pm 8.3$ .

### C. Ratio of $D^* + \mu$ Efficiencies, $R_{D^*}^{\text{gen}}$

It is a known effect that the PYTHIA MC generation of  $b$  production does not model the true  $p_T(b)$  distribution well. In addition, the MC sample has not been passed through a trigger simulation (which has its own deficiencies). A weighting factor [20] as a function of the generated  $p_T$  of the  $b$  hadron was first applied to MC events to match the measured  $b$  hadron  $p_T$  distribution before trigger effects. The data, in this case, the  $D^*\mu$  sample, was then used as the basis for reweighting the MC sample to provide a better description of the data, including trigger effects, particularly

of the single muon triggers. Figure 5(a) shows the data reconstruction of  $p_T(D^*\mu)$  compared with the MC weighted as described above. The disagreement at lower values of  $p_T$  is due to the unsimulated trigger. By dividing the two distributions of Fig. 5(a), and assuming that the trigger efficiency plateaus at higher values of  $p_T$ , the trigger efficiency turn-on curve of Fig. 5(b) is obtained. Weighted MC events, including this trigger efficiency extracted from the data, are included in the determination of efficiencies that follow.

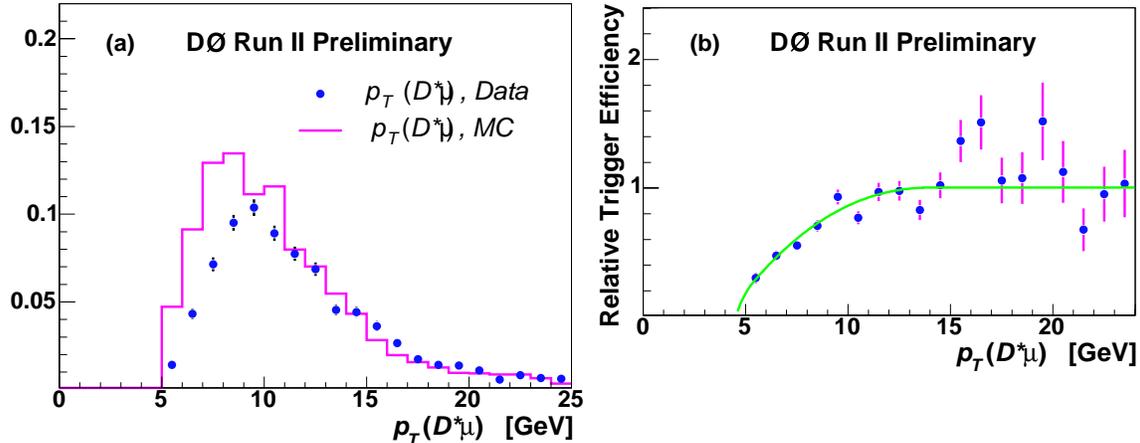


FIG. 5: (a) Generated  $p_T$  distribution of  $D^*\mu$  after weighting in the inclusive  $D^*\mu$  MC sample compared to the  $p_T(D^*\mu)$  distribution from the data, both after application of  $D^*\mu$  requirements. The difference at low  $p_T$  is due to trigger effects in the data. (b) Estimated trigger efficiency turn-on curve by taking the ratio of distributions in (a).

Using the MC sample of inclusive  $b \rightarrow D^*\mu X$  events, specific major decays were identified as listed in Table II. This sample and the MC signal sample were both required to have  $p_T^{\text{gen}}(B) > 4$  GeV in the calculation of efficiencies as a point of normalization, and no MC events generated with  $p_T(B) < 4$  GeV were observed to pass  $D^*\mu$  selection cuts. Efficiencies for generated events to pass the  $D^*\mu$  selection (but none of the  $K_S^0$  requirements) were then determined and shown in Table II. Errors on these efficiencies are due to MC statistics, including the additional statistical uncertainty produced due to the weighting procedure [21]. 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. [12]. The errors indicated on these fractions are dominated by uncertainties in PDG production fraction and branching ratio inputs, and are fully correlated (since they sum to unity by construction).

TABLE II: Efficiencies for reconstructing  $D^*\mu$  and fractions  $F_i$ .

Decay Channel	$\epsilon(b \rightarrow D^*\mu X)$	Fraction, $F_i$
$B_d^0 \rightarrow D^*\mu\nu$	$(6.10 \pm 0.11)\%$	$0.764 \pm 0.032$
$B_d^0 \rightarrow D^{*0}\mu\nu$	$(5.97 \pm 0.61)\%$	$0.070 \pm 0.014$
$B^+ \rightarrow D^{*+}\mu\nu$	$(6.66 \pm 0.53)\%$	$0.149 \pm 0.029$
$B_s^0 \rightarrow D^*\mu\nu$	$(0.096 \pm 0.043)\%$	$0.018 \pm 0.015$
$\sum \epsilon_i F_i$	$(6.08 \pm 0.50)\%$	

Applying the same cuts for reconstructing  $D^*\mu$  for the signal channel, the efficiency  $\epsilon(B_s^0 \rightarrow D_{s1}\mu \rightarrow D^*\mu) = (3.64 \pm 0.02)\%$  (MC statistical error only), resulting in the ratio of efficiencies of  $R_{D^*}^{\text{gen}} = 0.600 \pm 0.049$ .

#### D. Efficiency to Reconstruct $D_{s1}^\pm(2536)$

The signal MC sample was used to determine  $\epsilon_{K_S^0} = (\text{No. of } D^*\mu \text{ events passing additional } K_S^0 \text{ requirements}) / (\text{No. of } D^*\mu \text{ events})$ , i.e., the efficiency to reconstruct  $D_{s1}^\pm(2536) \rightarrow D^*K_S^0$  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 vertexing it with the  $D^*\mu$ , and already includes the branching ratio  $Br(K_S^0 \rightarrow \pi^+\pi^-) = 0.6895$  [7] for ease of use in calculating the product branching ratio.

A  $p_T$ -dependent weight factor as described above was applied to the signal MC to force agreement of the initial  $p_T(B_s^0)$  distribution. Figure 6 compares the true  $p_T$  of the  $B_s^0$  in the signal MC with that found from  $D^*\mu$  in the data after weighting.

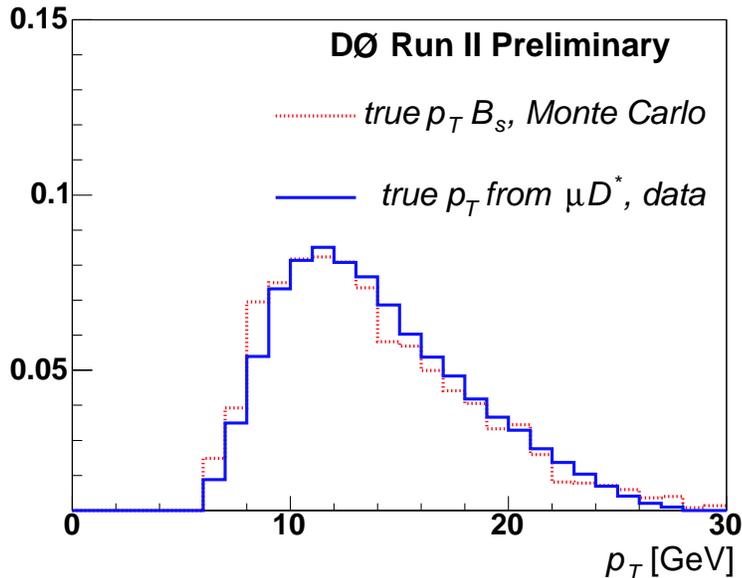


FIG. 6: Comparison of the true  $p_T$  of the  $B_s^0$  in the signal MC after weighting described in the text to the estimated  $p_T(B)$  in the data, both after application of  $D^*\mu$  requirements.

After this weighting, the efficiency  $\epsilon_{K_S^0}$  was found to be  $(11.1 \pm 0.3)\%$ . This error is only from MC statistics plus the statistical fluctuations of the weights [21]; additional errors due to the uncertainty in the determination of the weights and the procedure will be considered later.

### E. $c\bar{c}$ Contribution

The process  $c\bar{c} \rightarrow D^{*-}\mu^+\nu X$  can contribute to  $N_{D^*\mu}$ . The analysis described in both Ref. [12] (before flavor tagging that reduces the  $c\bar{c}$  contribution) and Ref. [22] use a similar selection for  $D^*\mu$  before the decay length significance cut on the  $D^*\mu$  vertex. From such previous studies, without the decay length significance cut, the fractional contribution for  $c\bar{c}$  contamination was estimated to be  $(9 \pm 3)\%$ .

The  $D^*\mu$  decay length significance cut was introduced in this analysis to reduce the  $c\bar{c}$  contamination in the  $D^*\mu$  sample since these products from direct charm production will typically have shorter decay lengths than if they arise as products of  $B$  meson decay. Estimating the fraction of  $c\bar{c}$  using MC studies can be difficult since much will arise from  $c\bar{c}$  production via gluon splitting where the charm quarks are close in phase space, with one decaying to  $D^*$  and the other to a muon. Instead, the decay length significance distribution observed in the data, compared to the decay length significance distribution predicted by MC for  $b \rightarrow D^*\mu X$  was used to estimate the fraction of  $c\bar{c}$  events in the  $D^*\mu$  sample.

Both the inclusive  $b \rightarrow D^*\mu$  and signal  $B_s^0 \rightarrow D_{s1}(2536)\mu\nu$  MC samples were used to determine the expected shape of the decay length significance distribution for  $B$  decays. At a large value of significance greater than 5, where the charm contribution should be negligible, these MC distributions were scaled to give the same statistics as the data distribution beyond this value. For smaller values of the significance cuts, the excess of  $D^*\mu$  candidates above that predicted by the MC samples cutting at different significance values was attributed as coming from charm. The average between the MC samples was taken, and the uncertainty assigned as the difference between the two MC predictions added in quadrature with the statistical error.

The value found when no decay length significance cut is applied is consistent with the  $(9 \pm 3)\%$  estimated by other techniques [12, 22, 23]. As the cut is tightened, the charm fraction drops as expected, until consistent with zero, albeit with significant uncertainty. For the cut value used in the analysis, the fraction of charm in  $N_{D^*\mu}$  was estimated to be  $(3.9 \pm 2.5)\%$ , and the value of  $N_{D^*\mu}$  was scaled down appropriately.

## F. Product Branching Fraction

Using the equation of Section II, the product branching ratio is found:

$$f(\bar{b} \rightarrow B_s^0) \cdot Br(B_s^0 \rightarrow D_{s1}^-(2536)\mu^+\nu X) \cdot Br(D_{s1}^- \rightarrow D^{*-}K_S^0) = (2.29 \pm 0.43 \text{ (stat.)}) \times 10^{-4},$$

i.e., this is the value for  $Br(\bar{b} \rightarrow D_{s1}^-(2536)\mu^+\nu X) \cdot Br(D_{s1}^- \rightarrow D^{*-}K_S^0)$ .

## VII. SYSTEMATIC UNCERTAINTIES

Systematic uncertainties are estimated for the product branching ratio and the mass measurement.

### A. Product Branching Ratio

The uncertainty in the normalizing branching ratio:  $Br(\bar{b} \rightarrow D^{*-}\mu^+\nu X) = (2.75 \pm 0.19)\%$  [7] was taken as a systematic error.

For determining  $N_{D^{*}\mu}$ , uncertainties in modeling the signal and background were studied. A triple Gaussian was used instead of a double Gaussian for the signal, and the background was fit using both an exponential function alone and an exponential function plus a square root function. The maximum variation in each case was taken as the estimated systematic error due to fit modeling. The estimated  $c\bar{c}$  contribution of  $(3.9 \pm 2.5)\%$  was varied by the indicated uncertainty.

In the determination of  $N_{D_{s1}(2536)}$ , the signal model was varied in a number of ways to determine the sensitivity of the candidate yield. The signal was fit with a single Gaussian instead of a double Gaussian. When using the predicted mass shape determined using Monte Carlo, the scaling of the widths was varied from 1.0 to 1.20 from the default value of 1.10 to check the sensitivity to uncertainty in mass resolution. The fraction of the narrow Gaussian was varied within its uncertainty from the MC fit. An unbinned likelihood fit was made to the invariant mass distribution instead of a binned fit. The maximum variation from the default fit over these variations was taken as the systematic error due to this source. The background model was changed to an exponential plus a square root function to determine this systematic error. Examination of MC events passing all cuts did not show any peaking backgrounds. No evidence of the other doublet member,  $D_{s2}^{*\pm}(2573)$ , decaying into the same channel appears in the data. The branching ratio of this state into  $D^*K_S^0$  is expected to be low, and even if a signal appeared, given expected mass resolution, it would not contaminate the  $D_{s1}(2536)$  mass peak.

Regarding the possibility of residual  $c\bar{c}$  contamination in  $N_{D_{s1}(2536)}$ , the fraction of  $c$  quarks fragmenting into  $D_{s1}(2536)$  was estimated to be approximately 32 times smaller than the fraction of  $c$  quarks fragmenting into  $D^*$  from relative production ratios [24] and spin-counting arguments [25]. When the decay length significance cut on the  $D^*\mu$  vertex was added, reducing the charm content from 11.2% to 3.9%, the resulting small drop in  $N_{D_{s1}(2536)}$  was completely consistent with the small decrease in the efficiency for signal due the addition of this requirement. Distributions of  $D_{s1}(2536)$  decay length, decay length significance, and  $D_{s1}(2536) - \mu$  invariant mass in the signal mass window, after sideband subtraction, were consistent between data and signal MC (see Fig. 7), with no significant discrepancies that may indicate the presence of  $c\bar{c}$  contamination, and no further correction was made.

When finding  $\epsilon_{K_S^0}$ , the uncertainty in  $p_T$  weighting was found by using different weightings techniques, i.e., weighting directly to the data without including trigger effects before weighting the true  $p_T(B)$  distribution, and using  $k$ -factors as in lifetime analyses to estimate  $p_T(B)$  from  $p_T(D^*\mu)$ . Reweighting was performed in each case, and the maximum variation in this efficiency was taken as the systematic error due to uncertainty in the weighting procedure.

By comparing the  $p_T(\mu)$  distribution for the signal using the default ISGW2 decay model to the HQET semileptonic decay model [2], 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 the systematic error due to uncertainty on the decay model.

To assess the effects of differences between data and MC on the modeling of  $K_S^0$  kinematics and decay length, the  $p_T$  cut on the  $K_S^0$  was varied in steps from its nominal value of greater than 1.0 GeV down to 0.75 GeV and up to 1.50 GeV. The cut on the decay length of the  $K_S^0$  was varied in steps from its nominal value of 0.5 cm down to 0.25 cm and up to 1.5 cm. The resulting variation of the fitted signal divided by the new MC efficiency determined in each case was found and the RMS spread of these “number produced” taken as a systematic error.

The uncertainty in  $R_{D^*}^{\text{gen}} = 0.600 \pm 0.049$  was due to a combination of MC statistics, and uncertainties in PDG branching ratio values as well as uncertainties in production fractions,  $f(\bar{b} \rightarrow b \text{ hadron})$ . Systematic effects due to decay modeling uncertainties as well as weighting factor uncertainties were tested as above, applying different

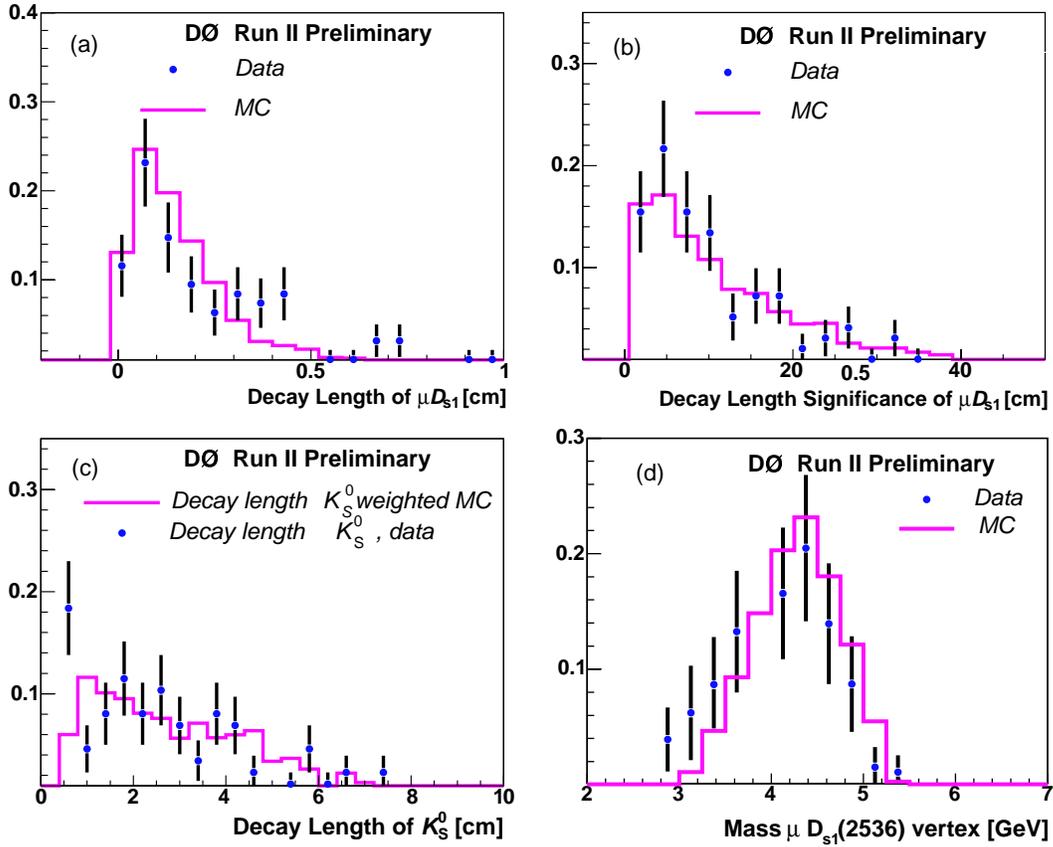


FIG. 7: Comparison of data in the mass window  $2.52 < M(D^* K_S^0) < 2.55$  GeV, after sideband subtraction, to signal MC for (a) decay length of the  $D_{s1}(2536)\text{-}\mu$  vertex; (b) decay length significance of the same vertex; (c) decay length of the  $K_S^0$  vertex; (d) invariant mass of  $D_{s1}(2536)\text{-}\mu$ .

weighting to both the  $D^* \mu$  inclusive sample and the signal MC sample. To assess the uncertainty of the trigger efficiency turn-on curve in the weighting, the MC was weighted to agree directly with  $p_T(D^* \mu)$  without the turn-on curve, as well as varying the turn-on curve within statistical errors, and the difference was taken as a systematic error. The necessity of weighting as a function of  $\eta$  was explored by finding and then using a separate trigger turn-on curves in each of the two regions  $|\eta| < 1$  and  $|\eta| > 1$ . The observed difference is 2.7% in the ratio of efficiencies. It was estimated that approximately 10% of the events were collected with an impact parameter trigger. To check for any possible biases, an offline requirement similar to the trigger requirement was made and applied to signal tracks before the calculation of efficiency for 10% of the candidates. This resulted in a 2.6% change of  $R_{D^*}^{\text{gen}}$  with the conclusion that these variations are already covered by assigned systematic uncertainties.

The estimated systematic errors on the product branching ratio are summarized in Table III and added in quadrature to obtain a total estimated systematic error on the product branching ratio of 15.5%.

Including the systematic error, the product branching ratio is determined to be:

$$f(\bar{b} \rightarrow B_s^0) \cdot Br(B_s^0 \rightarrow D_{s1}^-(2536)\mu^+\nu X) \cdot Br(D_{s1}^- \rightarrow D^{*-} K_S^0) = (2.29 \pm 0.43 \text{ (stat.)} \pm 0.36 \text{ (syst.)}) \times 10^{-4}.$$

## B. Mass Measurement

The same variations of the  $D_{s1}(2536)$  mass signal model, as well as background functional form were made as described above, i.e., the number of Gaussian functions fit, the mass resolution variations, and the shape of the MC predicted peak, etc. The mass values used for the mass constraints on the decay products were varied within their PDG uncertainties, and also set to the DØ central fit values. A new central mass value found in each case. The maximum variation observed was 0.4 MeV/ $c^2$ . These tests were also repeated on the higher statistics of the signal

TABLE III: Estimated systematic errors.

Source	Systematic Error
Normalizing $Br$	$Br(b \rightarrow D^* \mu X)$ 6.9%
$N_{D^* \mu}$	Signal Modeling 0.5%
	Background Modeling 1.3%
	$c\bar{c}$ Contribution 2.7%
$N_{D_{s1}(2536)}$	Signal Modeling 3.0%
	Background Modeling 4.6%
$\epsilon_{K_S^0}$	MC Statistics 2.8%
	Semileptonic Decay Model 1.2%
	Weighting Procedure 2.4%
$R_{D^*}^{\text{gen}}$	Detector Modeling 4.0%
	MC Statistics, PDG $Br$ and $f$ Uncertainties 8.2%
	Weighting Procedure 7.4%
	Semileptonic Decay Model 0.9%
<b>Total</b>	<b>15.5%</b>

MC with smaller variations found. The signal MC was broken up into 50 ensembles, each with statistics close to the data, and the mass found in each case. The pull and width of the pulls indicated that the statistical error is consistent. The mass value found tends to remain stable due to the mass constraints on the decay products, as well as the peak location close to threshold, despite the larger variations observed in other typical mass peaks due to momentum scale uncertainties. The difference between the mass fit in the large signal MC sample and the input PDG mass value was  $0.16 \pm 0.10$  MeV/ $c^2$ . To check for momentum scale shifts for the signal, the fitted value of the mass difference  $M(D^*) - M(D^0)$  in the data, signal MC, and inclusive  $D^* \mu$  MC were compared to the PDG value, with a maximum observed difference of 0.2 MeV. An unbinned likelihood fit was made to the invariant mass distribution with minimal difference in fitted mass value. A total estimated systematic mass error of 0.5 MeV/ $c^2$  was taken, for a mass measurement of  $2535.7 \pm 0.6$  (stat.)  $\pm 0.5$  (syst.) MeV/ $c^2$ .

### VIII. DISCUSSION AND CONCLUSION

To allow comparison of this measurement to theoretical predictions, the semileptonic branching ratio alone is extracted by taking the hadronization fraction into  $B_s^0$  as  $f(\bar{b} \rightarrow B_s^0) = 0.107 \pm 0.011$  [7] and also assuming that  $Br(D_{s1}(2536) \rightarrow D^* K_S^0) = 0.25$  [2]. The first experimental measurement of this value is compared to a number of theoretical predictions [18, 26, 27] in Table IV. The systematic error on this quantity is as described earlier, and the error labeled “(prod. frac.)” is due to the current uncertainty on  $f(\bar{b} \rightarrow B_s^0)$ .

TABLE IV: Experimental measurement compared with various theoretical predictions.

Source	$Br(B_s^0 \rightarrow D_{s1}(2536) \mu \nu X)$
This Result	$(0.86 \pm 0.16$ (stat.) $\pm 0.13$ (syst.) $\pm 0.09$ (prod.frac.))%
	$Br(B_s^0 \rightarrow D_{s1}(2536) \mu \nu)$
ISGW2 [18]	0.53%
RQM [26]	0.39%
HQET & QCD sum rules [27]	0.195%

The measured mass value of the  $D_{s1}(2536)$  of  $2535.7 \pm 0.6$  (stat.)  $\pm 0.5$  (syst.) MeV/ $c^2$  can be compared to the PDG average value of  $2535.34 \pm 0.31$  MeV/ $c^2$  [7].

- 
- [1] ALEPH Collab., A.H. Heister *et al.*, *Phys. Lett.* **B526** (2002) 34;  
 OPAL Collab., *Z. Phys.* **C76** (1997) 425;  
 CLEO Collab., J. Alexander *et al.*, *Phys. Lett.* **B303** (1993) 377;  
 ARGUS Collab., *Phys. Lett.* **B297** 425.

- [2] EVTGEN package home page <http://hep.ucsb.edu/people/lange/EvtGen/> . Note that the branching ratio  $Br(D_{s1}(2536) \rightarrow D^*K_S^0)$  indicated in the decay table is 0.5; by isospin arguments, this should be the value for decays into  $D^*K^0$ , with the  $K^0$  being 50%  $K_S^0$  and 50%  $K_L^0$ .
- [3] BaBar Collab., B. Aubert *et al.*, *Phys. Rev. Lett.* **90** (2003) 242001;  
CLEO Collab., D. Besson *et al.*, *Phys. Rev.* **D68** (2003) 032002;  
BELLE Collab., P. Krokovny *et al.*, *Phys. Rev. Lett.* **91** (2003) 262002.
- [4] S. Godfrey, N. Isgur, *Phys. Rev.* **D32** (1985) 037502;  
S. Godfrey, R. Kokoski, *Phys. Rev.* **D43** (1991) 1679;  
M. Di Pierro, E. Eichten, *Phys. Rev.* **D64** (2001) 114004.
- [5] E. Robutti [BABAR Collaboration], *Recent BaBar results on hadron spectroscopy*, *Acta Phys. Polon. B* **36** (2005) 2315.
- [6] A. V. Evdokimov *et al.* [SELEX Collaboration], *Phys. Rev. Lett.* **93** (2004) 242001 [arXiv:hep-ex/0406045].
- [7] S. Eidelman *et al.*, *Phys. Lett.* **B592** (2004) 1, and 2005 partial update for edition 2006.
- [8] V.M. Abazov *et al.*, *The Upgraded D0 Detector*, submitted to *Nucl. Instr. and Methods*, Fermilab-Pub-05/341-E [arXiv:hep-physics/0507191].
- [9] V.M. Abazov *et al.*, *Nucl. Instr. and Meth. A* **552** (2005) 372-398.
- [10] p17 release.
- [11] V. M. Abazov *et al.* [D0 Collaboration], *Phys. Rev. Lett.* **94** (2005) 182001 [arXiv:hep-ex/0410052].
- [12] DØ Note 5029-CONF ,  $B_d^0$  mixing measurement using Opposite-side Flavor Tagging, G. Borissov *et al.*, Feb. 23, 2006.
- [13] Muon ID web page  
[http://www-d0.fnal.gov/computing/algorithms/muon/muon\\_algo.html](http://www-d0.fnal.gov/computing/algorithms/muon/muon_algo.html)
- [14] S. Catani, Yu.L. Dokshitzer, M. Olsson, G. Turnock, B.R. Webber, *Phys. Lett.* **B269** (1991) 432.
- [15] T. Sjöstrand *et al.*, *Comp. Phys. Commun.* **135** (2001) 238.
- [16] R. Brun *et al.*, CERN Report No. DD/EE/84-1, 1984.
- [17] Private generation, J. Rieger.
- [18] Ref. [27], for input using the results of D. Scora and N. Isgur, *Phys. Rev.* **D52** (1995) 2783.
- [19] MC Request ID 11685,  $b \rightarrow D^*\mu X$ .
- [20] V.M. Abazov *et al.*, [DØ Collaboration] *Phys. Rev. Lett.* **94** (2005) 071802; private communication, R. Bernhard.
- [21] CDF Memo/Statistics/Public/7168, C. Blocker, Aug. 4, 2004.
- [22] DØ Note 4639, C. Ay *et al.*, Jan. 26, 2006.
- [23] G. Borissov, S. Burdin, A. Nomerotski, “Observation of semileptonic B decays to narrow  $D^{**}$  mesons”, DØ note 4324 (2004).
- [24] D. Acosta *et al.*, [CDF Collaboration] *Phys. Rev. Lett.* **91** (2003) 241804.
- [25] A. F. Falk and M. E. Peskin, *Phys. Rev.* **D49** (1994) 3320 [arXiv:hep-ph/9308241].
- [26] Ref. [27], for input using the results of D. Ebert, R.N. Faustov and V.O. Galkin, *Phys. Lett.* **B434** (1998) 365.
- [27] H. B. Mayorga, A. Moreno Briceno and J. H. Munoz, *J. Phys. G* **29** (2003) 2059.
- [28] In the plane perpendicular to the beam direction.
- [29] In the plane parallel to the beam direction.