First observation of $\bar{B}_s^0 \to D_{s}^{\pm} K^\mp$ and measurement of the ratio of branching fractions $\mathcal{B}(\bar{B}_s^0 \to D_{s}^{\pm} K^\mp)/\mathcal{B}(\bar{B}_s^0 \to D_{s}^+ \pi^-)$

A combined mass and particle identification fit is used to make the first observation of the decay $B^0_s \rightarrow D_s^+ K^-$. This analysis uses 1.2 fb$^{-1}$ integrated luminosity of $pp$ collisions at $\sqrt{s} = 1.96$ TeV collected with the CDF II detector at the Fermilab Tevatron collider. We observe a $B^0_s \rightarrow D_s^+ K^-$ signal with a statistical significance of 8.1σ and measure $B(B^0_s \rightarrow D_s^+ K^-)/B(B^0_s \rightarrow D_s^+ \pi^-) = 0.007 \pm 0.018$ (stat) $\pm 0.009$ (sys).

One of the remaining open questions in flavor physics is the precise value of the angle $\gamma = \arg(-V_{cd}V_{ub}^*/V_{ub}V_{cd}^*)$ of the unitarity triangle. Current measurements use the interference between $b \rightarrow ucs$ and $b \rightarrow cus$ diagrams
in $B^- \rightarrow D^{(*)0} K^{(*)-}$ and $B^- \rightarrow \bar{D}^{(*)0} K^{(*)+}$ decays when $D^0$ and $\bar{D}^0$ are observed in common final states $1 2 3 4 5 6$, but suffer from the large difference between the amplitudes of these decays. With a large sample of hadronic $|B^0|$ decays, it may be possible to determine $\gamma$ from the interference, through $B^0 - \bar{B}^0$ mixing, of the same diagrams in the decay modes $\bar{B}^0_s \rightarrow D^+_s K^-$ and $B^0_s \rightarrow D^+_s K^+ 3 [8]$, which are expected to have a more favorable amplitude ratio; the two decays proceed through color-allowed tree amplitudes whose ratio is suppressed by only a factor $\sim 0.4 [3]$. To determine $\gamma$, a time-dependent measurement of the decay rates of $\bar{B}^0_s \rightarrow D^+_s K^-$ and $B^0_s \rightarrow D^+_s K^+$, $B^0_s \rightarrow D^-_s K^+$, and $B^0 \rightarrow D^+_s K^-$ is required. The first steps in this effort are to observe these decay modes (which we will collectively refer to as $\bar{B}^0_s \rightarrow D^+_s K^-$) and to measure the CP-averaged branching ratio $B(\bar{B}^0_s \rightarrow D^+_s K^+) = \frac{1}{2}[B(\bar{B}^0_s \rightarrow D^+_s K^-) + B(B^0_s \rightarrow D^-_s K^+) + B(B^0_s \rightarrow D^+_s K^-)]$. In this Letter we report the first observation of the $\bar{B}^0_s \rightarrow D^+_s K^+$ decay modes and the first measurement of $B(\bar{B}^0_s \rightarrow D^+_s K^+)/B(\bar{B}^0_s \rightarrow D^-_s \pi^+)$. We measure this branching fraction ratio since many of the systematic uncertainties cancel in the ratio and $B(\bar{B}^0_s \rightarrow D^+_s \pi^-)$ is precisely measured elsewhere [10, 11].

We analyze $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV recorded by the CDF II detector at the Fermilab Tevatron collider with an integrated luminosity of 1.2 fb$^{-1}$. A detailed description of the detector can be found elsewhere [12]. This analysis uses charged particle tracks reconstructed in the pseudorapidity range $|\eta| \leq 1$ from hits in a silicon microstrip vertex detector [14] and a cylindrical drift chamber [13] immersed in a 1.4 T axial magnetic field. The specific ionization energy loss ($dE/dx$) of charged particles in the drift chamber is used for particle identification (PID). A sample rich in bottom hadrons is selected by triggering on events that have at least two tracks, each with $p_T > 2$ GeV/c and large impact parameter; the trigger further requires that these tracks originate from a secondary vertex well displaced from the primary interaction point [16].

We reconstruct $\bar{B}^0_s \rightarrow D^+_s h^\pm$ candidates (where $h = \pi$ or $K$) as follows. First, we identify $D^+_s \rightarrow \phi(\rightarrow K^- K^+) \pi^+$ candidates [17] using the invariant mass requirements $1013 < m(K^- K^+) < 1028$ MeV/c$^2$ and $1948.3 < m(K^- K^+ \pi^+ ) < 1988.3$ MeV/c$^2$. The $D^+_s$ decay tracks must satisfy a three-dimensional vertex fit. No PID requirements are made on the $D^+_s$ decay tracks. We then pair the $D^+_s$ candidates with $h^-$ tracks to define the $\bar{B}^0_s \rightarrow D^+_s h^-$ candidate sample, and require the $D^+_s - h^-$ pair to satisfy a three-dimensional fit to the $\bar{B}^0_s$ decay vertex. No mass constraint is used either on the $\phi$ or on the $D^+_s$ candidate. Finally, we define a mass variable $m(D_s \pi)$ for the $D_s \pi$ hypothesis (i.e., assigning the daughter track $h$ as a pion): $m(D_s \pi)$ is used to provide kinematic separation between the $\bar{B}^0_s \rightarrow D^+_s K^+$ and $\bar{B}^0_s \rightarrow D^+_s \pi^-$ signals.

Further selection requirements are made to reduce combinatorial background. The discriminating variables are the transverse ($|d_0| < 60$ μm) and longitudinal ($|\Delta z/\sigma z| < 3$, where $\sigma z$ is the uncertainty on $z_0$) impact parameter of the $\bar{B}^0_s$ candidate with respect to the primary event vertex; the transverse momentum ($p_T > 5.5$ GeV/c) of the $\bar{B}^0_s$ candidate; the isolation of the $\bar{B}^0_s$ candidate

$$I = \frac{p_T (\bar{B}^0_s)}{p_T (\bar{B}^0_s) + \sum_{\text{tracks}} p_T (\text{track})} > 0.5,$$

where the sum runs over tracks within $\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2} < 1$ around the $\bar{B}^0_s$ direction originating from the same primary vertex; the opening angle $\Delta R(D^+_s, h^-) < 1.5$ between the $D^+_s$ candidate and the track originating from the $\bar{B}^0_s$ decay vertex (the $\bar{B}^0_s$ daughter track); and the projection of the $\bar{B}^0_s$ and $D^+_s$ decay length along the transverse momentum of the $\bar{B}^0_s$ candidate $L_{xy}(\bar{B}^0_s) > 300$ μm, $L_{xy}(\bar{B}^0_s)/\sigma_{L_{xy}}(\bar{B}^0_s) > 8$ (where $\sigma_{L_{xy}}$ is the uncertainty on $L_{xy}$), and $L_{xy}(D^+_s) > L_{xy}(\bar{B}^0_s)$. The $dE/dx$ calibrations are based on large samples of $D^0 \rightarrow K^- \pi^+$ decays taken with the displaced track trigger. To avoid bias, the $\bar{B}^0_s$ daughter tracks are required to pass the same $p_T > 2$ GeV/c trigger requirement as the $D^0 \rightarrow K^- \pi^+$ calibration tracks.

Monte Carlo simulation is used to model signal and background and to determine trigger and reconstruction efficiencies. We generate single $\bar{B}^0_s$ hadrons with BGENERATOR [18, 19] and simulate their decays with EVTGEN [20]. A detailed detector and trigger simulation is then performed.

The greatest challenge in this analysis is to disentangle the various components contributing to the data sample. Apart from the $\bar{B}^0_s \rightarrow D^+_s K^+$ and $\bar{B}^0_s \rightarrow D^+_s \pi^-$ signals, the sample contains partially reconstructed $\bar{B}^0_s$ decays, reflections from decays of other bottom hadron species, and combinatorial background. To separate the components and determine the number of candidates of each component type, we perform a maximum-likelihood fit. The fit variables are the invariant mass $m(D_s \pi)$ of the candidate in the $D_s \pi$ mass hypothesis and the PID variable $Z$, which is the logarithm of the ratio between the measured $dE/dx$ and the expected $dE/dx$ for a pion with the momentum of the $\bar{B}^0_s$ daughter track. The likelihood function is

$$L(f_1, \ldots, f_{M-1}) = \prod_{i=1}^{N} \prod_{j=1}^{M} f_j p_j(m_i) q_j(Z_i),$$

where $f_M = 1 - \sum_{j=1}^{M-1} f_j$. The index $i$ runs over the $N$ candidates, and $j$ runs over the $M$ components: $f_j$ is the fraction of candidates in the $j$th component, to be determined by the fit.

We group $\bar{B}^0_s$ candidates into three categories by source. Combinations where the $D^+_s$ candidate and the track come from a single bottom hadron ($\bar{B}^0_s$, $B^-$, $\bar{B}^0$, $\Lambda_b^0$) are called single-$B$ contributions. Non-bottom contributions where the $D^+_s$ candidate does not come from
a real $D^{+}_s$ are called fake-$D^{+}_s$ combinatorial background. Combinations of a real $D^{+}_s$ with a track coming from fragmentation, the underlying event, or the other bottom hadron in the event are called real-$D^{+}_s$ combinatorial background.

Mass probability density functions (pdf’s) $p_j(m)$ for the single-$B^0$ components are extracted from large simulated samples of $B^0_s \rightarrow D^{*+}X$ decays, where the $D^{*+}$ is forced to decay to $\phi \pi^+$. Separate mass templates are extracted for $B^0_s \rightarrow D^{+}_s \pi^-$ and $B^0_s \rightarrow D^{*+}_s K^\mp$ fully reconstructed decays and for the partially reconstructed modes that overlap in mass with $B^0_s \rightarrow D^{+}_s K^\mp$: $B^0_s \rightarrow D^{+}_s \rho^-$ and $B^0_s \rightarrow D^{+}_s \pi^-$. Partially reconstructed modes missing more than one decay product are collected in one template. Contributions from the $B^0_s$ modes relative to each other are fixed to the values reported in [21]. Rather than parameterizing the mass shapes, which are complicated for most B modes, we use histograms as pdf’s. Sufficiently large Monte Carlo samples (approximately 50,000 candidates after cuts of $B^0_s \rightarrow D^{+}_s \pi^-$ and $B^0_s \rightarrow D^{*+}_s K^\mp$) are generated to make the statistical fluctuations in the pdf’s small.

Special care has to be taken in the treatment of the low-mass tail of the decay mode $B^0_s \rightarrow D^{+}_s \pi^-(n\pi)$, which is dominated by the radiative tail, and which overlaps with $B^0_s \rightarrow D^{+}_s K^\mp$ mass pdf. Improper accounting of the tail can bias both the measurement of the $B^0_s \rightarrow D^{+}_s \pi^-$ yield and the $B^0_s \rightarrow D^{*+}_s K^\mp$ yield by misidentifying a fraction of the $B^0_s \rightarrow D^{+}_s \pi^-$ contribution as part of the (much smaller) $B^0_s \rightarrow D^{*+}_s K^\mp$ contribution. The radiative tail is modeled in EVTGEN by using the PHOTOS algorithm for radiative corrections [22] with a cut-off for photon emission at 10 MeV. We allow the normalization to float in the fit to account for uncertainties in the PHOTOS prediction of the size of the radiative tail. (The radiative tail of $B^0_s \rightarrow D^{*+}_s K^\mp$ does not require special treatment. The kaon radiates less, and any resulting misidentified $B^0_s \rightarrow D^{*+}_s K^\mp$ contribution is easily absorbed by the other fit components, which dominate at $m(D^{*+}_s \pi^-) \lesssim m(D^{*+}_s K^\mp)$ below the $B^0_s \rightarrow D^{*+}_s K^\mp$ peak.)

The mass distribution of the fake-$D^{*+}_s$ background is parameterized with a function of the form $p_{ps}(m) \propto \exp(-\alpha m) + \beta$. The shape parameters $\alpha$ and $\beta$ are determined in an ancillary mass-only fit of $B^0_s$ candidates populating the sidebands of the $D^{+}_s$ mass distribution. To model the real-$D^{*+}_s$ background, we use a sample of same-sign $D^{*+}_s \pi^+$ candidates. A fit analogous to the fake-$D^{+}_s$ case is performed on the $D^{*+}_s \pi^+$ mass distribution. Given to the limited statistics of the signal sample, we cannot separately resolve the real-$D^{*+}_s$ and fake-$D^{*+}_s$ combinatorial backgrounds; in the default fit we therefore combine the two types of background. We assess a systematic uncertainty by allowing the relative size of the two background types to vary.

We determine the $Z$ pdf’s $q_j(Z)$ for pions and kaons from $D^{*+} \rightarrow D^{0}(K^-\pi^+)\pi^{+}$ decays. The flavor of the daughter tracks of the $D^{0}$ in the decay $D^{*+} \rightarrow D^{0}(K^-\pi^+)\pi^{+}$ is tagged by the charge of the soft pion from the $D^{*+}$ decay. Taken together with the large signal-to-background ratio of the $\Delta m = m(K^-\pi^+) - m(K^-\pi^+)$ peak, this charge tagging yields a very clean sample of pions and kaons. We further reduce background contamination by sideband-subtracting in $\Delta m$. The mean values of $Z$ for kaons and pions are separated by approximately 1.4 standard deviations. The $Z$ distributions for both species (shown in Figure 1) have similar widths. Because the data sample contains semileptonic decays, we need to model the $dE/dx$ distributions of muons and electrons as well. For muons, which are a small contribution in the mass region of interest, the pion template can be used without introducing a significant systematic uncertainty. For electrons, we derive a template from a $J/\psi \rightarrow e^+e^-$ sample. The $Z$ pdf for the fake-$D^{*+}_s$ combinatorial background is determined from data by selecting candidates from the sidebands of the $D^{*+}_s$ mass distribution. All $Z$ pdf’s are represented as histograms.

Figures 2 and 3 show the fit projections in mass and $Z$. The yields determined by the fit are $1125 \pm 87$ $B^0_s \rightarrow D^{*+}_s \pi^-$ and $102 \pm 18$ $B^0_s \rightarrow D^{*+}_s K^\mp$ candidates. The branching fraction of $B^0_s \rightarrow D^{*+}_s K^\mp$ relative to $B^0_s \rightarrow D^{+}_s \pi^-$, corrected for the relative reconstruction efficiency $\epsilon_{s}/\epsilon_{K} = 1.071 \pm 0.028$, is $\mathcal{B}(B^0_s \rightarrow D^{*+}_s K^\mp)/\mathcal{B}(B^0_s \rightarrow D^{+}_s \pi^-) = 0.097 \pm 0.018$. A fit performed with the $B^0_s \rightarrow D^{*+}_s K^\mp$ yield set to zero is worse than the default fit by $\Delta \log L = -32.52$; the corresponding statistical significance of the $B^0_s \rightarrow D^{*+}_s K^\mp$ signal is 8.1 standard deviations.

Systematic uncertainties on $\mathcal{B}(B^0_s \rightarrow D^{*+}_s K^\mp)/\mathcal{B}(B^0_s \rightarrow D^{+}_s \pi^-)$ are studied by incorporating each effect in the generation of simulated experiments which are then fitted using the default configuration. The bias on $\mathcal{B}(B^0_s \rightarrow D^{*+}_s K^\mp)/\mathcal{B}(B^0_s \rightarrow D^{+}_s \pi^-)$ is...
Table I: Systematic uncertainties on $B(\overline{B}^0 \rightarrow D^+_s K^\mp)/B(\overline{B}^0 \rightarrow D^+_s \pi^-)$.

<table>
<thead>
<tr>
<th>Source</th>
<th>Systematic uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>$dE/dx$ pdf modeling</td>
<td>0.007</td>
</tr>
<tr>
<td>Mass pdf modeling</td>
<td>0.004</td>
</tr>
<tr>
<td>Combinatorial-background model</td>
<td>0.002</td>
</tr>
<tr>
<td>Fitter bias due to finite statistics</td>
<td>0.001</td>
</tr>
<tr>
<td>Sum in quadrature</td>
<td>0.009</td>
</tr>
</tbody>
</table>

$D^+_s K^\mp)/B(\overline{B}^0 \rightarrow D^+_s K^\mp)$, averaged over 10000 simulated experiments, is used as the systematic uncertainty associated with the effect under study. Table I summarizes the systematic uncertainties. The systematic uncertainties are dominated by the modeling of the $dE/dx$ (0.007), specifically by the differences between the $Z$ distributions of $D^+$ daughter tracks (from which the kaon and pion $Z$ pdf’s are derived) and $\overline{B}^0$ (daughter tracks); these differences arise from effects such as the greater particle abundance in the vicinity of a prompt $D^+$ compared to a $\overline{B}^0$, and hence a higher probability for $D^+$ daughter tracks to contain extraneous hits. Modeling of the mass distributions of the single-$B$ components (0.004), which includes statistical fluctuations in the mass pdf’s, and modeling of the combinatorial-background mass shape (0.002) are comparatively minor contributions. The total systematic uncertainty is obtained by adding the individual systematic uncertainties in quadrature; at 0.009, it is about half as large as the statistical uncertainty.

The analysis procedure was crosschecked in several ways. Most importantly, before performing the measurement on the $\overline{B}^0 \rightarrow D^+_s K^\mp$ signal sample, we verified our method using two control samples, $\overline{B}^0 \rightarrow D^+ X$ and $\overline{B}^0 \rightarrow D^+ X$. In both cases, our results are statistically consistent with world-average values. We measure $B(\overline{B}^0 \rightarrow D^+ K^-)/B(\overline{B}^0 \rightarrow D^+ \pi^-) = 0.086 \pm 0.005$(stat), 1.0 standard deviations from the world average; and $B(\overline{B}^0 \rightarrow D^+ K^-)/B(\overline{B}^0 \rightarrow D^+ \pi^-) = 0.080 \pm 0.008$(stat), 0.3 standard deviations from the world average. [23]. The relative branching fractions $B(\overline{B}^0 \rightarrow D^+ \rho^-)/B(\overline{B}^0 \rightarrow D^+ \pi^-)$, $B(\overline{B}^0 \rightarrow D^+ \pi^-)/B(\overline{B}^0 \rightarrow D^+ \pi^-)$, and $B(\overline{B}^0 \rightarrow D^+ \pi^-)/B(\overline{B}^0 \rightarrow D^+ \pi^-)$ were also found to be consistent with world averages. Finally, the fractional size of the radiative tails of $\overline{B}^0 \rightarrow D^+ \pi^-$, $\overline{B}^0 \rightarrow D^+ \pi^-$, and $\overline{B}^0 \rightarrow D^+_s \pi^-$ are found to be in agreement with each other (and about twice as large as the PHOTOS prediction).

In conclusion, we have presented the first observation of the $\overline{B}^0 \rightarrow D^+_s K^- X$ decay mode with a statistical significance of 8.1 standard deviations. The $\overline{B}^0 \rightarrow D^+_s K^- X$ event yield is 102 ± 18 (statistical uncertainty only). We use this sample to measure $B(\overline{B}^0 \rightarrow D^+_s K^-)/B(\overline{B}^0 \rightarrow D^+_s \pi^-) = 0.097 \pm 0.018$(stat) ± 0.009(sys). This result is consistent with naive expectations based on the branching fraction ratio for the analogous $\overline{B}^0$ and $B^-$ decays, taking into account also the expected contribution from $D^- K^+$ decays.

Figure 2: Mass projection of the likelihood fit. Fit components are stacked. $B \rightarrow DX$ denotes all single-$B$ contributions except $\overline{B}^0 \rightarrow D_s^+ \pi^-$ and $\overline{B}^0 \rightarrow D_s^+ K^\mp$, namely $\overline{B}^0 \rightarrow D^+_s \rho^-$, $\overline{B}^0 \rightarrow D^+_s \pi^-$, partially reconstructed $\overline{B}^0 \rightarrow D^+_s X$ modes missing more than one decay product, $\overline{B}^0 \rightarrow D^+ (K^- \pi^+ \pi^+) X$, $B^- \rightarrow D^+ (K^- \pi^+ \pi^+) X$ and $\Lambda^0 \rightarrow \Lambda^+_c (pK^- \pi^+) X$ reflections, and $B^0 \rightarrow D_s^{(+)} \pi^-$ and $B^0 \rightarrow D_s^{(-)} K^+$; the small peak in the $B \rightarrow DX$ template is due to the $\overline{B}^0 \rightarrow D^+ (K^- \pi^+ \pi^+) \pi^- \pi^-$ reflection. The residual plot at the bottom shows the discrepancy (data minus fit) in units of standard deviation ($\sigma$); for the bins with low statistics, neighboring bins are combined until the predicted number of events is greater than five. The $\chi^2$ of the projection is 79.0 for 72 degrees of freedom.

Acknowledgments

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
Figure 3: $Z$ projection of the likelihood fit in the region of interest for $B_s^0 \rightarrow D_s^+ K^-$ ($5.26 < m(D_s \pi) < 5.35 \text{ GeV}/c^2$). Fit components are stacked. The residual plot at the bottom shows the discrepancy (data minus fit) in units of standard deviation ($\sigma$); for the bins with low statistics, neighboring bins are combined until the predicted number of events is greater than five. The $\chi^2$ of the projection is 30.7 for 14 degrees of freedom.

[13] CDF II uses a right-handed coordinate system with the origin at the center of the detector, in which the $z$ axis is along the proton direction, the $y$ axis points up, $\theta$ and $\phi$ are the polar and azimuthal angles, and $r$ is the radial distance in the $x$-$y$ plane. The pseudorapidity $\eta$ is defined as $-\log \tan(\theta/2)$.
[17] Reference to the charge-conjugate decays is implied here and throughout the text.