First Observation of $\bar{B}_s^0 \rightarrow D_s^\pm K^\mp$ and Measurement of the Ratio of Branching Fractions $\mathcal{B}(\bar{B}_s^0 \rightarrow D_s^- K^+) / \mathcal{B}(\bar{B}_s^0 \rightarrow D_s^+ \pi^-)$

A combined mass and particle identification fit is used to make the first observation of the decay $B_s^0 \to D_s^+ K^-$ and measure the branching fraction of $B_s^0 \to D_s^+ K^-$ relative to $B_s^0 \to D_s^+ \pi^-$. This analysis uses 1.2 fb$^{-1}$ integrated luminosity of $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV collected with the CDF II detector at the Fermilab Tevatron collider. We observe a $B_s^0 \to D_s^+ K^-$ signal with a statistical significance of 8.1$\sigma$ and measure $\mathcal{B}(B_s^0 \to D_s^+ K^-)/\mathcal{B}(B_s^0 \to D_s^+ \pi^-) = 0.097 \pm 0.018\,(\text{stat}) \pm 0.009\,(\text{syst})$. 

DOI: 10.1103/PhysRevLett.103.191802

PACS numbers: 13.25.Hw, 12.15.Hh
One of the remaining open questions in flavor physics is the precise value of the angle \( \gamma = \arg(-V_{ud}V_{ub}^*/V_{cd}V_{cb}^*) \) of the unitarity triangle [1]. Current measurements use the interference between \( b \to u'c's \) and \( b \to u'c's \) in the \( B^- \to D^{(*)+}K^{(*)-} \) and \( B^- \to \bar{D}^{(*)0}K^{(*)-} \) decays when \( D^0 \) and \( \bar{D}^0 \) are observed in common final states [2–7], 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 \( B^0 \to D_s^+ K^- \) and \( \bar{B}^0 \to D_s^- K^+ \) [8,9], 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 \( \sim 0.4 \) [10]. To determine \( \gamma \), a time-dependent measurement of the decay rates of \( B^0, \bar{B}^0 \) decays and \( B^0 \to D_s^+ K^- \), \( B^0 \to D_s^- K^+ \), \( B^0 \to D_s^+ K^- \), and \( \bar{B}^0 \to D_s^+ K^- \) is required. The first steps in this effort are to observe these decay modes (which we will collectively refer to as \( B^0, \bar{B}^0 \to D_s^+ K^- \)) and to measure the CP-averaged branching ratio \( B(B^0 \to D_s^+ K^-) = \frac{1}{2}(B(B^0 \to D_s^+ K^-) + B(\bar{B}^0 \to D_s^- K^+)) \). In this Letter we report the first observation of the \( B^0 \to D_s^+ K^- \) decay modes and the first measurement of \( B(B^0 \to D_s^+ K^-)/B(\bar{B}^0 \to D_s^+ \pi^-) \). We measure this branching fraction ratio since many of the systematic uncertainties cancel in the ratio and \( B(B^0 \to D_s^+ \pi^-) \) is precisely measured elsewhere [11,12].

We analyze \( p\bar{p} \) collisions at \( \sqrt{s} = 1.96 \text{ TeV} \) recorded by the CDF II detector at the Fermilab Tevatron collider with an integrated luminosity of 1.2 \( \text{fb}^{-1} \). A detailed description of the detector can be found elsewhere [13]. This analysis uses charged particle tracks reconstructed in the pseudorapidity [14] range \( |\eta| \leq 1 \) from hits in a silicon microstrip vertex detector [15] and a cylindrical drift chamber [16] 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 \( p_T > 2 \text{ GeV}/c \) and large impact parameter; the trigger further requires that these tracks originate from a secondary vertex well displaced from the primary interaction point [17].

We reconstruct \( B_s^0 \to D_s^+ h^- \) candidates (where \( h = \pi \) or \( K \)) as follows. First, we identify \( D_s^+ \to b(\to K^- K^0) \pi^+ \) candidates [18] using the invariant mass requirements \( 1013 < m(K^- K^0) < 1028 \text{ MeV}/c^2 \) and \( 1948.3 < m(K^- K^0 \pi^+) < 1988.3 \text{ MeV}/c^2 \). 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 \( B_s^0 \to D_s^+ h^- \) candidate sample. We further require the \( D_s^+ h^- \) pair to satisfy a three-dimensional fit to the \( B_s^0 \) decay vertex. No mass constraint is used either on the \( h^- \) or on the \( D_s^+ \) candidate. Finally, we define a mass variable \( m \) which is the invariant mass of the \( D_s^+ h^- \) combination where the \( h^- \) is assigned the \( \pi^- \) mass; \( m \) is used to provide kinematic separation between the \( B_s^0 \to D_s^+ K^- \) and \( B_s^0 \to D_s^- \pi^- \) signals.

To reduce combinatorial background, further selection requirements are made on the \( B_s^0 \) candidate: transverse impact parameter \( (|d_{xy}| < 60 \mu \text{m}) \); longitudinal impact parameter \( (|z_0|/\sigma_{z_0} < 3 \) where \( \sigma_{z_0} \) is the uncertainty on \( z_0 \) [19]; transverse momentum \( (p_T > 5.5 \text{ GeV}/c) \); isolation \( I = p_T(\bar{B}^0)/[p_T(\bar{B}^0) + \sum_{\text{tracks}} p_T(\text{track})] > 0.5 \), where the sum runs over tracks within \( \Delta R = \sqrt{\Delta \phi^2 + \Delta \eta^2} < 1 \) around the \( B^- \) direction originating from the same primary vertex; the opening angle \( \sqrt{\Delta R(D_s^+, h^-)} < 1.5 \) between the \( D_s^+ \) candidate and the \( h^- \) track; and the projection of the \( B^0 \) and \( D_s^+ \) decay lengths along the transverse momentum of the \( B_s^0 \) candidate \( |L_{xy}(B_s^0) > 300 \mu \text{m}, L_{xy}(\bar{B}^0)/\sigma_{L_{xy}}(\bar{B}^0) > 8 \) (where \( \sigma_{L_{xy}} \) is the uncertainty on \( L_{xy} \)). The \( dE/dx \) calibrations are based on large samples of \( D^0 \to K^- \pi^+ \) decays taken with the displaced-track trigger. To avoid bias, the \( h^- \) track is required to pass the same \( p_T > 2 \text{ GeV}/c \) trigger requirement as the \( D^0 \to K^- \pi^+ \) calibration tracks.

Monte Carlo simulation is used to model signal and background and to determine trigger and reconstruction efficiencies. Single \( B_s^0 \) hadrons are produced with \( \text{GENEVENT} \) [20,21]; their decays are simulated with \( \text{EVGEN} \) [22] and a detailed detector and trigger simulation.

The greatest challenge in this analysis is to disentangle the various components contributing to the data sample. Apart from the \( B_s^0 \to D_s^+ K^- \) and \( B_s^0 \to D_s^- \pi^- \) signals, the sample contains partially reconstructed \( B_s^0 \) 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 \( m \) 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 \( h^- \) track. The likelihood function is \( L(f_1, \ldots, f_{M-1}) = \prod_{i=1}^{M} \sum_{j=1}^{M} f_j \pi_j(m_i) q_i(Z_i) \), 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 the fit components into three categories by source. Combinations where the \( D_s^+ \) candidate and the \( h^- \) track come from a single bottom hadron (\( B_s^0, \bar{B}^0, \bar{B}_s^0, \Lambda_b^0 \)) are called single-\( B \) contributions. Nonbottom 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. The single-\( B \) category is comprised of several independently normalized components, whose normalizations are free
parameters of the fit: the $B^0 \rightarrow D_s^+ K^-$ and $B^0 \rightarrow D_s^+ \pi^-$ signals and the radiative tail of the $B^0 \rightarrow D_s^+ \pi^+$, which will be discussed in more detail below; $B^0 \rightarrow D_s^+ \rho^-$; $B^0 \rightarrow D_s^+ \pi^-$; $B^0 \rightarrow B^+ \rightarrow D^0(K^- \pi^+ \pi^-)X$ and $\Lambda^0_{bg} \rightarrow \Lambda^+ (pK^- \pi^+)X$, which have narrow reflections with masses close to the signal peaks and which have separate fit normalizations; $B^0 \rightarrow D_s^+ h^-$ decays (where $h = \pi, K$), whose relative yields are fixed to the values reported in [23]; and partially reconstructed $B^0$ decay modes missing more than one decay product or a neutrino, which are grouped together in the fit.

Mass probability density functions (PDF's) $p_j(m)$ for the single-$B$ components are extracted from large simulated samples of $B$ decays. Separate mass templates are extracted for each of the components described above. Rather than parameterizing the mass shapes, which are complicated for most of the single-$B$ components, we use histograms as mass PDF's. Sufficiently large Monte Carlo samples (approximately 50 000 candidates after cuts of $B^0 \rightarrow D_s^+ \pi^-$ and $B^0 \rightarrow D_s^+ K^-$) 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 radiative tail of the decay mode $B^0 \rightarrow D_s^+ \pi^-$, which overlaps with the $B^0 \rightarrow D_s^+ K^-$ mass PDF. Improper accounting for the tail can bias the measurement of both the $B^0 \rightarrow D_s^+ \pi^+$ yield and the $B^0 \rightarrow D_s^+ K^-$ yield by misidentifying a fraction of the $B^0 \rightarrow D_s^+ \pi^-$ contribution as part of the (much smaller) $B^0 \rightarrow D_s^+ K^-$ contribution. The radiative tail is modeled in EVTGEN by using the PHOTOS algorithm for radiative corrections [24] with a cutoff for photon emission at 10 MeV. We allow the normalization to float in the fit to account for uncertainties in the PHOTOS predictions of the size of the radiative tail. (The radiative tail of $B^0 \rightarrow D_s^+ K^-$ does not require special treatment. The kaon radiates less, and any resulting misidentified $B^0_s \rightarrow D_s^+ K^-$ contribution is easily absorbed by the other fit components, which dominate at $m$ below the $B^0 \rightarrow D_s^+ K^-$ peak.)

The mass distribution of the fake-$D_s^+$ background is parameterized with a function $p_{bg}(m) \propto \exp(-\alpha m) + \beta$. 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 to the $D_s^+ \pi^+$ mass distribution was performed using the same form for the mass distribution that was used for the fake-$D_s^+$ parameterization. Given 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 $\Delta m = m(K^- \pi^+ \pi^+) - m(K^- \pi^+)$ 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 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 1 and 2 show the fit projections in mass and $Z$. The yields determined by the fit are $1125 \pm 87 \: B_s^0 \rightarrow D_s^+ \pi^-$ and $102 \pm 18 \: B^0 \rightarrow D^+_s K^-$ candidates. The branching fraction of $B^0 \rightarrow D_s^+ K^-$ relative to $B^0 \rightarrow D_s^+ \pi^-$, corrected for the relative reconstruction efficiency, is $B(B^0 \rightarrow D_s^+ K^-)/B(B^0 \rightarrow D_s^+ \pi^-) = 0.097 \pm 0.018$. The reconstruction efficiency differs between the two modes due to the kinematics of the decay and due to the nuclear interaction and decay-in-flight probabilities of kaons and pions [25]. A Monte Carlo simulation of the detector and trigger based on GEANT [26] is used to determine the relative reconstruction efficiency between kaons and pions $\epsilon_{\pi}/\epsilon_K = 1.071 \pm 0.028$, where the uncertainty is due to Monte Carlo statistics; this uncertainty is included in the systematic uncertainty on the ratio of branching fractions. A fit performed with the $B^0 \rightarrow D_s^+ K^-$ yield set to zero is worse than the default fit by $\Delta \log L = -32.52$; the corresponding statistical significance of the $B^0 \rightarrow D_s^+ K^-$ signal is 8.1 standard deviations.

Systematic uncertainties on $B(B^0 \rightarrow D_s^+ K^-)/B(B^0 \rightarrow D_s^+ \pi^-)$ are studied by incorporating each effect in the generation of simulated experiments which are then fitted using the default configuration. The bias on $B(B^0 \rightarrow D_s^+ K^-)/B(B^0 \rightarrow D_s^+ \pi^-)$, averaged over 10 000 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 $B^0_s$ daughter tracks; these differences arise from effects such as the greater particle abundance in the vicinity of a prompt $D_s^+$ compared to a $B^0_s$, and hence a higher probability for $D_s^+$ 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, modeling of the combinatorial-background mass shape (0.002), and uncertainty induced in the Z template by the poorly known particle content for the $B^0$, $B^+$, and $\Lambda_b$ reflections (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.

One of the dominant sources of uncertainty is the $B^0 \to D^+(K^−π^+π^+)X$ reflection, which is strongly anticorrelated with the $B^0 \to D_s^+K^-X$ signal. The normalization of the reflection (like all other background components) is allowed to vary independently of the other contributions in the fit; the uncertainty due to the size of the reflection is therefore accounted for as part of the statistical uncertainty on $B(B_s^0 \to D^+_sK^-)/B(B_s^0 \to D^+_sπ^-)$.

The analysis procedure was crosschecked in several ways. Most importantly, before performing the measurement on the $B^0 \to D^+_sX$ signal sample, we verified our method using two control samples, $B^0 \to D^+X$ and $B^0 \to D^{*+}X$. In both cases, our results are statistically consistent with world-average values. We measure $B(B^0 \to D^+K^-)/B(B^0 \to D^+π^-) = 0.086 \pm 0.005$ (stat), 1.0 standard deviations above the world average; and $B(B^0 \to D^{*+}K^-)/B(B^0 \to D^{*+}π^-) = 0.080 \pm 0.008$ (stat), 0.3 standard deviations above the world average [1]. The relative branching fractions $B(B^0 \to D^{*+}ρ^-)/B(B^0 \to D^{*+}π^-)$, $B(B^0 \to D^{*−}π^-)/B(B^0 \to D^{*−}π^-)$, and $B(B^0 \to D^{*−}ρ^-)/B(B^0 \to D^{*−}π^-)$ were also found to be consistent within 2 standard deviations with world averages. Finally, the fractional sizes of the radiative tails of $B^0 \to D^+π^−$, $B^0 \to D^{*+}π^−$, and $B^0 \to D_s^+π^−$ are

### Table I. Systematic uncertainties on $B(B_s^0 \to D^+_sK^-)/B(B_s^0 \to D^+_sπ^-)$.

<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>Relative $K$ to $π$ efficiency</td>
<td>0.003</td>
</tr>
<tr>
<td>Combinatorial-background model</td>
<td>0.002</td>
</tr>
<tr>
<td>Composition of $B^0$, $B^+$, $Λ_b$ reflections</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>

FIG. 1. Mass projection of the likelihood fit. Fit components are stacked. $B \to DX$ denotes all single-$B$ contributions except $B^0 \to D^+_sπ^−$ and $B^0 \to D^+_sK^- $; the small peak in the $B \to DX$ slice is due to the $B^0 \to D^+(K^−π^+π^+)π^−$ reflection. Although these components are combined in the fit, they are allowed to vary independently in the fit. The residual plot at the bottom shows the discrepancy (data minus fit) in units of standard deviation ($σ$); for the bins with low statistics, neighboring bins are combined until the predicted number of events is greater than five. The $χ^2$ of the projection is 79.0 for 72 degrees of freedom.

FIG. 2. $Z$ projection of the likelihood fit in the region of interest for $B^0 \to D^+_sK^- (5.26 < m < 5.35 \text{ GeV}/c^2)$. Fit components are stacked. The residual plot at the bottom is calculated as in Fig. 1. The $χ^2$ of the projection is 30.7 for 14 degrees of freedom.
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 $B^0_s \to D^+_{s\pi} K^- \pi^+$ decay mode with a statistical significance of 8.1 standard deviations. The $B^0_s \to D^+_{s\pi} K^- \pi^+$ event yield is $102 \pm 18$ (statistical uncertainty only). We use this sample to measure $\mathcal{B}(B^0_s \to D^+_{s\pi} K^-) / \mathcal{B}(B^0 \to D^0 \pi^-) = 0.097 \pm 0.018\text{stat} \pm 0.009\text{syst}$. This result is consistent with naive expectations based on the branching fraction ratio for the analogous $B^0$ and $B^- \to D^- \pi$ decays, taking into account also the expected contribution from $B^0_s \to D^+_{s\pi} K^- \pi^+$ decays [10].

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the Bundesministerium für Bildung und Forschung, Germany; the Korean Science and Engineering Foundation and the Korean Research Foundation; the Science and Technology Facilities Council and the Royal Society, UK; the Institut National de Physique Nucleaire et Physique des Particules/CNRS; the Russian Foundation for Basic Research; the Ministerio de Ciencia e Innovación, Spain; the Slovak R&D Agency; and the Academy of Finland.

---

*Visitor from Texas Tech University, Lubbock, TX 79409, USA
*Visitor from IFIC(CSIC-Universitat de Valencia), 46071 Valencia, Spain.
*Visitor from Royal Society of Edinburgh, Scotland.
*Visitor from Istituto Nazionale di Fisica Nucleare, Sezione di Cagliari, 09040 Monserrato (Cagliari), Italy.

[14] 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 $-\ln \tan(\theta/2)$.
[18] Reference to the charge-conjugate decays is implied here and throughout the text.
[19] The impact parameter is defined as the distance of closest approach of the $B^0_s$ candidate’s trajectory with the event-by-event primary vertex. In events with more than one reconstructed primary vertex, the vertex with the largest scalar sum of track $p_T$ is used.