Observation of \( Bs_0J/K^*(892)0 \) [\( B_{s0} J / \psi K^*(892) \) superscript 0] and \( Bs_0J/KS0 \) [\( B_{s0} J / \psi K_{s0} \) superscript 0] decays
Observation of $B^0_s \rightarrow J/\psi K^0(892)^0$ and $B^0_s \rightarrow J/\psi K^0$ decays


(CDF Collaboration)

1Institute of Physics, Academia Sinica, Taipei, Taiwan 11529, Republic of China
2Institute of Physics, Academia Sinica, Taipei, Taiwan 11529, Republic of China
3Argonne National Laboratory, Argonne, Illinois 60439, USA
4University of Athens, 157 71 Athens, Greece
5Instituto de Física d’Altes Energies, ICN, Universitat Autònoma de Barcelona, E-08193, Bellaterra (Barcelona), Spain
6Baylor University, Waco, Texas 76798, USA
6aIstituto Nazionale di Fisica Nucleare Bologna, I-40127 Bologna, Italy
6bUniversity of Bologna, I-40127 Bologna, Italy
7University of California, Davis, Davis, California 95616, USA
8University of California, Los Angeles, Los Angeles, California 90024, USA
9Instituto de Física de Cantabria, CSIC, University of Cantabria, 39005 Santander, Spain
10Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA
11Enrico Fermi Institute, University of Chicago, Chicago, Illinois 60637, USA
12Comenius University, 842 48 Bratislava, Slovakia, and Institute of Experimental Physics, 040 01 Kosice, Slovakia
13Joint Institute for Nuclear Research, RU-141980 Dubna, Russia
14Duke University, Durham, North Carolina 27708, USA
15Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA
16University of Florida, Gainesville, Florida 32611, USA
17Laboratori Nazionali di Frascati, Istituto Nazionale di Fisica Nucleare, I-00044 Frascati, Italy
18University of Geneva, CH-1211 Geneva 4, Switzerland
19Glasgow University, Glasgow G12 8QQ, United Kingdom
20Harvard University, Cambridge, Massachusetts 02138, USA
21Division of High Energy Physics, Department of Physics, University of Helsinki and Helsinki Institute of Physics, FIN-00014, Helsinki, Finland
22University of Illinois, Urbana, Illinois 61801, USA
23The Johns Hopkins University, Baltimore, Maryland 21218, USA
24Institut für Experimentelle Kernphysik, Karlsruhe Institute of Technology, D-76131 Karlsruhe, Germany
25Center for High Energy Physics, Kyungpook National University, Daegu 702-701, Korea; Seoul National University, Seoul 151-742, Korea; Sungkyunkwan University, Suwon 440-746, Korea; Korea Institute of Science and Technology Information, Daejeon 305-806, Korea; and Chonnam National University, Gwangju 500-757, Korea; Chonbuk National University, Jeonju 561-756, Korea
26University of Kentucky, Lexington, Kentucky 40506, USA
27University of Liverpool, Liverpool L69 7ZE, United Kingdom
28University College London, London WC1E 6BT, United Kingdom
29Centro de Investigaciones Energéticas Medioambientales y Tecnologicas, E-28040 Madrid, Spain
30Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
31Institute of Particle Physics, McGill University, Montreal, Quebec, Canada H3A 2T8; Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6; University of Toronto, Toronto, Ontario, Canada M5S 1A7; and TRIUMF, Vancouver, British Columbia, Canada V6T 2A3
I. INTRODUCTION

This paper presents the first observation of the Cabibbo-suppressed decay modes $B^0_d \to J/\psi K^{(892)^0}$ and $B^0_s \to J/\psi K_S^{0}$ (and the corresponding charge conjugate decays) using a sample derived from an integrated luminosity of 5.9 fb$^{-1}$ of proton-antiproton collisions at $\sqrt{s} = 1.96$ TeV produced at the Fermilab Tevatron. In addition to isolating these signals, we normalize the observed yields to the corresponding Cabibbo-favored $B^0$ decay modes ($B^0 \to J/\psi K^{*0}$, where $K^{*0}$ refers to $K^{(892)^0}$), and $B^0 \to J/\psi K^0_S$) to extract the branching ratios for these newly observed $B^0$ decay modes.

With the establishment of the decay modes presented here, future measurements can be considered that will further aid our experimental investigation into the physics of the $B^0$ system. The success of the Cabibbo-Kobayashi-Maskawa (CKM) three-generation description of charge conjugation-parity inversion (CP) violation [1] in the bottom and kaon sectors has continued to motivate additional, more precise tests of CP violation in the flavor sector. In recent years, attention has turned to the $B^0$ meson as new territory to explore the possibility of nonstandard-model contributions, specifically in the CKM matrix element $V_{ts}$. Precise measurement of the frequency of $B^0$ flavor oscillations [2] has significantly limited the magnitude of new physics amplitudes. However, possible large new physics phases remain poorly constrained.

Cabibbo-suppressed $B^0_d$ modes could provide complementary information on the $B^0_s$ mixing phase and on the width difference $\DeltaGamma_{B^0_d} = \Gamma_{B^0_d}^{*} - \Gamma_{B^0_d}^{0}$, where $\Gamma_{B^0_d}^{*}$ ($\Gamma_{B^0_d}^{0}$) is the width of the light, even (heavy, odd) $B^0_d$ CP eigenstate [3]. The decay $B^0_d \to J/\psi K^{(892)^0}$ is a pseudoscalar to vector-vector transition and can be used to help disentangle penguin contributions in $B^0_s \to J/\psi \phi$ [4]. With a sufficiently large data sample, it would be possible to measure $\DeltaGamma_{B^0_d}$ and the polarization amplitudes. Furthermore, the Cabibbo-suppressed decay $B^0_s \to J/\psi K_S^{0}$ is a CP-odd final state (ignoring CP violation in the kaon system), and therefore a measurement of the lifetime in this decay mode is a direct measure of $\Gamma_{B^0_s}^{0} = 1/\tau_{B^0_s}^{0}$. With a larger data sample, a tagged CP asymmetry analysis of the $B^0_s \to J/\psi K_S^{0}$ mode, in conjunction with our precise knowledge of CP violation in $B^0 \to J/\psi K_S^{0}$, can yield information on the angle $\gamma$ of the unitarity triangle [5].

In the naïve spectator model, the ratio of branching ratios is given by the ratio of the squares of the CKM elements,

$$\frac{\mathcal{B}(B^0_s \to J/\psi K_S)}{\mathcal{B}(B^0_s \to J/\psi K^*)} = \frac{|V_{td}|^2}{|V_{ts}|^2} \approx 0.051 \pm 0.006, \quad (1)$$

where $K$ represents $K^0_S$ or $K^{*0}$. The numerical value is derived from $|V_{td}| = 0.230 \pm 0.011$ and $|V_{ts}| = 1.023 \pm 0.036$ [3].

Experimentally, we extract the relative branching ratios using the relation

$$\frac{\mathcal{B}(B^0_s \to J/\psi K_S)}{\mathcal{B}(B^0_s \to J/\psi K^*)} = A_{\text{rel}} \frac{f_d}{f_s} \frac{N(B^0_s \to J/\psi K_S)}{N(B^0_s \to J/\psi K^*)}, \quad (2)$$

where $A_{\text{rel}}$ is the relative acceptance, $f_s/f_d$ is the ratio of fragmentation fractions, and $N(B^0_s \to J/\psi K_S)/N(B^0_s \to J/\psi K^*)$ is the measured ratio of yields.

We can use the result from Eq. (1) to estimate the relative yield in the spectator model. The value for $f_s/f_d$ is extracted from the most recent Collider Detector at Fermilab Tevatron (CDF) measurement [6] of $f_s/(f_s + f_d) \times \mathcal{B}(D_s \to \phi \pi)$ and $f_d/(f_s + f_d)$, along with the current world-average value [3] for $\mathcal{B}(D_s \to \phi \pi)$. Combining the value $f_s/f_d = 0.269 \pm 0.033$ with the assumption that $A_{\text{rel}} = 1$, Eq. (2) yields

$$\frac{N(B^0_s \to J/\psi K_S)}{N(B^0_s \to J/\psi K^*)} = \frac{\mathcal{B}(B^0_s \to J/\psi K_S) f_s}{\mathcal{B}(B^0_s \to J/\psi K^*) f_d A_{\text{rel}}} = 0.014 \pm 0.002. \quad (3)$$

While the result holds only in the simple spectator case, it provides useful guidance that we might expect one to two Cabibbo-suppressed $B^0_s \to J/\psi K$ events for every 100 Cabibbo-favored $B^0 \to J/\psi K$ events.

After a description of the detector, data sample, and simulated samples utilized here, we describe the
\(B^0 \rightarrow J/\psi K^*(892)^0\) analysis in Sec. III, followed by the \(B^0_s \rightarrow J/\psi K^{*0}\) analysis in Sec. IV. Section V then describes the acceptance calculation for both modes, followed by the results in Sec. VI.

II. CDF DETECTOR, DATA, AND MONTE CARLO SAMPLES

The data used in these analyses correspond to an integrated luminosity of 5.9 fb\(^{-1}\) and were collected by the CDF II detector from March 2002 to February 2010 using di-muon triggers. The CDF II detector is a general purpose, cylindrically symmetric detector. A more detailed description can be found elsewhere [7]. The subdetectors relevant for these analyses are briefly discussed here. Charged particle trajectories (tracks) are measured by a system comprising eight layers of silicon microstrip detector (SVX) and an open-cell wire drift chamber (COT), both immersed in a 1.4 T axial magnetic field. The silicon detector [8] extends from a radius of 1.5 to 22 cm and has a single-hit resolution of approximately 15 \(\mu\)m. The COT drift chamber [9] provides up to 96 measurements from radii of 40 to 137 cm and covers the range \(|\eta| \leq 1\) [10]. The combined COT + SVX charged particle momentum resolution is \(\sigma_{p_T}/(p_T)^2 = 0.07\%\) \([\text{GeV}/c]^2\), which leads to a mass resolution on the \(K^0_S \rightarrow \pi^+ \pi^-\) signal of 0.006 \text{GeV}/c\(^2\). Outside the calorimeters reside four layers of planar drift chambers [11] (CMU) that detect muons with transverse momentum \(p_T > 1.4\) \text{GeV}/c within \(|\eta| < 0.6\). Additional chambers and scintillators [12] (CMX) cover 0.6 < \(|\eta| < 1.0\) for muons with \(p_T > 2\) \text{GeV}/c.

The di-muon triggers collect a sample of \(J/\psi \rightarrow \mu^+ \mu^-\) candidates. At the first level of a three-level trigger system, an electronic track processor (XFT) [13] uses COT information to find tracks and extrapolate those with \(p_T > 1.5\) \((2.0)\) \text{GeV}/c to track segments in the CMU (CMX) muon-chambers. Events pass this first trigger level if two or more XFT tracks are matched to muon-chamber track segments. The second trigger level requires those tracks to have opposite charge and an appropriate opening angle in the plane transverse to the beam line. Finally, at level 3, full tracking information is used to reconstruct \(J/\psi \rightarrow \mu^+ \mu^-\) candidates. Events with a candidate in the mass range 2.7–4.0 \text{GeV}/c\(^2\) are accepted.

To identify \(B^0\) and \(B^0_s\) decay candidates, we pair \(J/\psi\) candidates with \(K^0_S \rightarrow \pi^+ \pi^-\) and \(K^{*0} \rightarrow K^+ \pi^-\) candidates. The reconstruction of \(K^0_S \rightarrow \pi^+ \pi^-\) and \(K^{*0} \rightarrow K^+ \pi^-\) candidates starts from pairs of oppositely charged tracks fit to a common interaction point (vertex). In the \(B^0 \rightarrow J/\psi K^0_S\) analysis, we reconstruct two tracks as pions and combine them to define a \(K^0_S\) candidate, where the invariant mass of the two pions is constrained to the known \(K^0_S\) mass [3]. In the \(B^0_s \rightarrow J/\psi K^{*0}\) analysis, we reconstruct the \(K^{*0}\) candidate from the combination of a \(\pi\) and a \(K\). If two \(K^{*0}\) candidates are reconstructed with the same tracks, with the only difference that the kaon and pion hypotheses are interchanged, we select the \(K^{*0}\) candidate whose mass is closer to the pole value of 896 \text{MeV}/c\(^2\). We perform a kinematic fit of each \(B\) candidate where the final-state tracks are constrained to come from a common decay point and the invariant mass of the muon pair is constrained to the known \(J/\psi\) mass [3]. These preliminary selection criteria for \(B^0\) and \(B^0_s\) candidates are listed in Table I. Additional selection criteria optimized for the individual channels are described in Secs. III and IV.

Simulated samples of \(B^0\) and \(B^0_s\) decays are used to optimize event selection, model signal distributions, and

<table>
<thead>
<tr>
<th>Variable (Units)</th>
<th>(B^0 \rightarrow J/\psi K^0_S)</th>
<th>(B^0_s \rightarrow J/\psi K^{*0})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(B^0/B^0_s) candidate four-track fit (\chi^2)</td>
<td>&lt;50</td>
<td>&gt;10(^{-5})</td>
</tr>
<tr>
<td>(B^0/B^0_s) candidate four-track fit probability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(B^0/B^0_s) candidate transverse momentum (p_T) (\text{GeV}/c)</td>
<td>&gt;6</td>
<td>&gt;4</td>
</tr>
<tr>
<td>(B^0/B^0_s) candidate impact parameter ((\mu m))</td>
<td>&lt;50</td>
<td></td>
</tr>
<tr>
<td>(B^0/B^0_s) candidate transverse decay length significance (L_{xy}/\sigma)</td>
<td>&gt;2</td>
<td></td>
</tr>
<tr>
<td>(J/\psi) candidate mass (\text{GeV}/c(^2))</td>
<td>&gt;3.05</td>
<td>&gt;2.8</td>
</tr>
<tr>
<td></td>
<td>&lt;3.15</td>
<td>&lt;3.3</td>
</tr>
<tr>
<td>(J/\psi) candidate 3-dimensional two-track fit (\chi^2)</td>
<td>&lt;30</td>
<td>&lt;30</td>
</tr>
<tr>
<td>(K) candidate mass (\text{GeV}/c(^2))</td>
<td>&gt;0.55</td>
<td>&gt;0.55</td>
</tr>
<tr>
<td></td>
<td>&lt;0.846</td>
<td>&lt;0.846</td>
</tr>
<tr>
<td>(K) candidate 3-dimensional two-track fit (\chi^2)</td>
<td>&lt;30</td>
<td>&lt;20</td>
</tr>
<tr>
<td>(K) candidate transverse decay length (L_{xy}) (cm)</td>
<td>&gt;0.5</td>
<td></td>
</tr>
<tr>
<td>(\mu) transverse momentum (p_T) (\text{GeV}/c)</td>
<td>&gt;1.5</td>
<td>&gt;1.5</td>
</tr>
<tr>
<td>(\Delta\phi) between the two muons (radians)</td>
<td>&lt;2.25</td>
<td>&lt;2.25</td>
</tr>
<tr>
<td>(\mu_1\text{charge} \times \mu_2\text{charge})</td>
<td>= -1</td>
<td>= -1</td>
</tr>
<tr>
<td>(\Delta z) in the beam line between the two (\mu) (cm)</td>
<td>&lt;5</td>
<td>&lt;5</td>
</tr>
<tr>
<td>(\pi) transverse momentum (p_T) (\text{GeV}/c)</td>
<td>&gt;0.5</td>
<td></td>
</tr>
</tbody>
</table>
assess systematic uncertainties. For our default Monte Carlo simulation (MC) samples, we generate single $b$ hadrons according to the predicted next-to-leading order QCD calculation [15]. For systematic studies, we also generate single $b$ hadrons according to momentum and rapidity spectra measured by CDF [7]. These hadrons are then decayed using the EVTGEN package [16] and fed into a GEANT simulation of the CDF detector [17]. The simulated data are then processed and reconstructed in the same manner as the detector data. In the case of $J/\psi K^{\pm 0}$ mode, it is necessary to specify the polarization parameters in the simulation. For both $B^0$ and $B^{0\prime}$, we use transversity basis [18] polarization amplitudes $|A_0|^2 = 0.6$ and $|A_1|^2 = 0.22$, which are similar to the PDG values of $|A_0|^2 = 0.571 \pm 0.008$ and $|A_1|^2 = 0.22 \pm 0.013$ [3]. For systematic acceptance studies, MC samples with other polarization values were generated.

In all of the MC samples generated, and throughout the analyses presented below, we assume that there is no CP violation in $B^{0\prime}$ mixing or decay. We also assume that equal numbers of $B^0$ and $B^{0\prime}$ mesons, as well as equal numbers of $B^0_s$ and $B^{0\prime}_s$ mesons, are produced in the $p\bar{p}$ collisions. In this untagged analysis that does not distinguish $B \to J/\psi K^0_S$ from $\bar{B} \to J/\psi K^0_S$, the observed yield is unaffected by CP violation provided that equal numbers of $B$ and $\bar{B}$ mesons are produced at $t = 0$.

III. $B^0 \to J/\psi K^{(892)}0$ ANALYSIS

We optimize the selection criteria to provide the highest likelihood for evidence of this mode. This is done by maximizing $S/(1.5 + \sqrt{B})$, where $S$ refers to the number of signal events and $B$ is the number of background events in the signal region. Reference [19] demonstrates that this quantity is well suited for discovery. For the signal sample, a $B^0 \to J/\psi K^{0\prime}$ MC sample is used. For the background sample, we use $J/\psi K^{0\prime}$ candidate events from data with the requirement that the reconstructed candidate mass $M_{B}$ falls in the range $5.6 \text{ GeV}/c^2 < M_{B} < 5.8 \text{ GeV}/c^2$. This “upper sideband” region contains events kinematically similar to the combinatorial background in the signal region and is not contaminated by residual signal events. We avoid using the sideband below the $B^0$ peak because it is contaminated with partially reconstructed $B$ decays such as $B^0 \to J/\psi K^{0\prime} \pi^0$. We optimize simultaneously over the transverse momenta $p_T(\pi^0)$ and $p_T(K^0\pi)$, the $B_{s}^{0}$ transverse decay length $L_{xy}(B_{s}^{0})$, and the $B_{s}^{0}$ decay kinematic-fit probability. The final cuts we use are $p_T(\pi^0) > 1.5 \text{ GeV}/c$, $p_T(K^0\pi) > 1.5 \text{ GeV}/c$, $L_{xy}(B_{s}^{0}) > 300 \mu$m, and fit probability greater than $10^{-5}$.

Particle identification using specific ionization $(dE/dx)$ in the COT was evaluated to further separate $K^{*0} \to K^+\pi^-$ from $\pi^+\pi^-$ and $K^+K^-$ backgrounds. Although further background reduction could be achieved, the corresponding reduction in signal efficiency rendered particle identification unprofitable, and we choose not to use it.

We determine the $B^0_s$ and $B^0$ yields using a binned likelihood fit in the candidate masses. We model the signal contributions with templates composed of three Gaussians obtained from fits to $B^0$ MC. The two dominant, narrow Gaussians model detector resolution effects and also account for cases where the identities of the $\pi$ and $K$ from the $K^{*0}$ decay are interchanged. As mentioned above, some events are identified where a single pair of tracks passes the selection requirements under both the $\pi-K$ and $K-\pi$ hypotheses. In those cases, we reconstruct the event using both sets of $\pi/K$ assignments and then choose a candidate whose particle assignment yields a reconstructed mass that is closer to the nominal $K^{*}(892)$ mass. This technique ensures that candidates are not used twice. Approximately 10% of $B \to J/\psi K^{*0}$ events are reconstructed with the incorrect $\pi-K$ assignment. These events peak at the $B$ masses, but have a significantly broader width. A wide Gaussian models misreconstructed signal events and other non-Gaussian resolution effects. The relative contributions, means, and widths of each Gaussian are fixed in the fit. The $B^{0}_s$ templates used in the fit are identical to $B^0$ templates, except for a shift of $86.8 \text{ MeV}/c^2$ in the mean value of the three Gaussians. This value corresponds to the known [3,20] mass difference between $B^0_s$ and $B^0$. The MC slightly underestimates the mass resolution, so the widths of the two narrow Gaussians are multiplied by a scale factor common to the $B^0$ and $B^{0}_s$ templates, which is allowed to float in the fit. The scale factor is not applied to the third Gaussian since the resolution effects are negligible compared to the other effects. Moreover, a common mass shift is added to the means of all Gaussian templates to account for a possible mass mismodeling in the MC. This mass shift is floating in the fit.

The $B^0 \to J/\psi K_{s}^{0}$ analysis has three primary background contributions: events with random track combinations (combinatorics), partially reconstructed $b$ hadrons, and $B^0 \to J/\psi \phi$ decays. Combinatorial background arises from sources such as a real $J/\psi$ plus two other tracks, where the $J/\psi$ could be either prompt or coming from a $B$ decay. Another source arises from false $J/\psi$ candidates reconstructed from misidentified hadrons. The combinatorial background is nonpeaking and accurately modeled in the fit with an exponential function.

Backgrounds from partially reconstructed $b$ hadrons come from multibody decays where a $\pi$, $K$, or $\gamma$ is not reconstructed, for example, the decay mode $B^0 \to J/\psi K^{*0} \pi^0$. We fit this background with two ARGUS functions [21], one for partially reconstructed $B^0$ and another for partially reconstructed $B^0_s$. The ARGUS function parametrization for $m < m_0$ is

$$f(m) = N_1 \times \sqrt{1 - \frac{m^2}{m_0^2}} \times e^{-C m^2/m_0^2},$$

where $m_0$ is the mass cutoff, $C$ the decay constant, and $N_1$ is the normalization. The function is set to zero for...
$m > m_0$. The ARGUS function for partially reconstructed $B^0$ has a fixed mass cutoff of $m(B^0) - m(\pi^0) = 5.140$ GeV/$c^2$, and the function for partially reconstructed $B^*_s$ has a fixed mass cutoff of $m(B_s^0) - m(\pi^0) = 5.220$ GeV/$c^2$. The decay constants of the two functions are constrained to be the same, and the normalizations are independent. Each ARGUS function is convoluted with a Gaussian having a width of 12 MeV/$c^2$ to account for detector resolution effects.

Since it is possible for $B^0 \rightarrow J/\psi \phi$ candidates to pass the $J/\psi K^{*0}$ reconstruction criteria, $B^0 \rightarrow J/\psi \phi$ must be considered as a background. We use a template consisting of two Gaussians, extracted from simulation, to model this background in the $J/\psi K^{*0}$ fit, where both Gaussians are primarily modeling detector resolution effects. We fix to the template values the widths, means, and relative contributions from each Gaussian in the final fit. We multiply the constant width of the narrower Gaussian by the same scale factor used in the signal templates. We constrain the $B^0 \rightarrow J/\psi \phi$ contribution in the $J/\psi K^{*0}$ fit by measuring the yield of $B^0 \rightarrow J/\psi \phi$ in the data using selection criteria efficient for reconstructing $B^0 \rightarrow J/\psi \phi$. We then use simulation to calculate the fraction of those $J/\psi \phi$ events that would satisfy the $J/\psi K^{*0}$ selection.

We perform a binned log likelihood fit to the $J/\psi K \pi$ invariant mass distribution using the templates for signals and the background functions described above. The mass distributions in data for $J/\psi K^{*0}$ candidates and the final fit appear in Fig. 1. The yields for $B^0 \rightarrow J/\psi K^{*0}$ and $B^*_s \rightarrow J/\psi K^{*0}$ signal are $9530 \pm 110$ and $151 \pm 25$, respectively. The ratio $N(B^0 \rightarrow J/\psi K^{*0})/N(B^0 \rightarrow J/\psi K^{*0})$ is $0.0159 \pm 0.0022$ (stat).

We determine the statistical significance of the $B^0 \rightarrow J/\psi K^{*0}$ signal by fitting the mass distribution without the $B^0$ contribution (background-only hypothesis). For likelihood $L$, we interpret $-2 \log L$ as a $\chi^2$ distribution. We use $\Delta \chi^2$ with 1 degree of freedom to determine that the probability of background fluctuations producing a comparable or greater signal is $8.9 \times 10^{-16}$ or $8.9 \sigma$. This is the first observation of the $B^0 \rightarrow J/\psi K^{*0}$ decay.

We consider several sources of systematic uncertainty in the measured ratio of $N(B^0 \rightarrow J/\psi K^{*0})/N(B^0 \rightarrow J/\psi K^{*0})$. The modeling of the $B^0$ and $B^*_s$ signal peaks can influence the ratio. To quantify the effect of the mis-modeling, we repeat the fit using two Gaussian templates instead of three for the signal. The fit value of $N(B^0 \rightarrow J/\psi K^{*0})/N(B^0)$ is shifted by $7 \times 10^{-4}$.

We vary the input mass difference between $B^0$ and $B^*_s$ in the templates within its uncertainty of 0.7 MeV/$c^2$. The difference in $N(B^0 \rightarrow J/\psi K^{*0})/N(B^0 \rightarrow J/\psi K^{*0})$ with the alternate templates is $2 \times 10^{-5}$. This is sufficiently small that we ascribe no systematic uncertainty due to the mass difference uncertainty.

The modeling of the combinatorial background is another source of systematic uncertainty. To explore the sensitivity to the choice parameterization, we use a power function ($f(m) = km^{\alpha}$, with $k$ and $\alpha$ free parameters) instead of an exponential. The overall fit quality with the power function is similar to that obtained using an exponential model. We assign the difference in relative yield as a background modeling systematic uncertainty of $2 \times 10^{-4}$ on the relative branching ratio.

In the likelihood fit, we allow the combinatorial background contribution to float. We performed a study to evaluate how the ratio of yields depends upon the specific, arbitrary choice of the fit range. We compare the main fit, which allows the combinatorial background to float over the entire fit range, to a control case where the combinatorial contribution is fitted in the upper sideband and extrapolated to the full mass range prior to the final fit. Because of the difference in the result from these two methods, we include a systematic uncertainty of 0.0050 on the $N(B_s^0 \rightarrow J/\psi K^{*0})/N(B^0 \rightarrow J/\psi K^{*0})$ ratio.

Several sources contribute to an uncertainty in the $B^0 \rightarrow J/\psi \phi$ contribution. While there is 30% uncertainty in the branching ratio, the dominant uncertainty arises from the uncertainty in the $B^0 \rightarrow J/\psi \phi$ template, given that we rely on MC to derive this. To perform a conservative assessment of this uncertainty, we repeated the fit while doubling the fraction of $B^0 \rightarrow J/\psi \phi$ candidates. The resulting shift

FIG. 1 (color online). (a) Invariant mass distribution in data for $J/\psi K^{*0}$ candidates and fit including the different contributions. (b) We enlarge the distribution in the signal region for more detail.
of $2 \times 10^{-4}$ is assigned as the uncertainty in the $B^0 \rightarrow J/\psi \phi$ contribution.

We add the different systematic uncertainty contributions, summarized in Table II, in quadrature, resulting in a final value of $N(B^0 \rightarrow J/\psi K^{*0})/N(B^0 \rightarrow J/\psi K^{*0})$ of $0.0159 \pm 0.0022$ (stat) $\pm 0.0050$ (syst).

### IV. $B^0 \rightarrow J/\psi K^0_S$ Analysis

The $B^0 \rightarrow J/\psi K^0_S$ decay has several differences compared to the $B^0 \rightarrow J/\psi K^{*0}$ decay. It contains a $K^0_S$, which has a relatively long lifetime of $c\tau = 2.68$ cm. We use the displacement between the reconstructed $K^0_S$ decay point and the reconstructed $B$ decay point in the event selection to reduce backgrounds such as $B^0 \rightarrow J/\psi K^{*0}$. Finally, as in the $B^0$ system, we expect the $B^0 \rightarrow J/\psi K^0_S$ signal to be smaller than that of the $B^0 \rightarrow J/\psi K^{*0}$ mode. Therefore, we use a neural network (NN) technique to take full advantage of all the kinematic variables and their correlations. We use the NEURALBAYES [22] NN package. The NN provides an output value close to $+1$ for signal-like events and near $-1$ for backgroundlike events.

We train the NN using simulated $B^0$ MC events as a signal sample. We use data from the upper sideband in the $J/\psi K^0_S$ candidate mass distribution, well separated from the signal region, as a background training sample. We use as inputs for the NN the quantities listed in Table III. These input quantities are chosen as good discriminating power which, alone or in combination, do not bias the mass spectrum. After the training, the NN achieves strong discrimination between signal and background, as shown in Fig. 2(a).

As in the $B^0 \rightarrow J/\psi K^{*0}$ analysis, we optimize the selection by maximizing $S/(1.5 + \sqrt{B})$. The signal $S$ is modeled using $B^0$ MC events in the reconstructed mass range $5.350$ GeV/$c^2 < M_B < 5.400$ GeV/$c^2$. The background $B$ is modeled using $J/\psi K^0_S$ candidates in data populating the mass range $5.430$ GeV/$c^2 < M_B < 5.480$ GeV/$c^2$. The figure of merit suggests a cut value in the NN response of 0.88, as shown in Fig. 2(b).

The fitting technique is similar to the $B^0 \rightarrow J/\psi K^{*0}$ analysis. We obtain the yields of $B^0 \rightarrow J/\psi K^0_S$ and $B^0 \rightarrow J/\psi K^0_S$ signals in a binned likelihood fit to the invariant mass distribution. We again model the $B^0$ and $B^0$ signal contributions with three Gaussian templates obtained from fitting $B^0 \rightarrow J/\psi K^0_S$ MC and use the mass difference between $B^0$ and $B^0$ for the formation of the $B^0 \rightarrow J/\psi K^0_S$ template. The two major sources of background in this analysis are combinatorial background and partially reconstructed $b$-hadron decays. We model these with the same functional forms used in the $B^0 \rightarrow J/\psi K^{*0}$ analysis. However, we include only one ARGUS function because the contribution of partially reconstructed $B^0$ is negligible. An additional background in this analysis is $\Lambda_{b,0} \rightarrow J/\psi \Lambda$ decays where the $p$ from the $\Lambda$ decay is assumed to be a $\pi$. In order to suppress the $\Lambda_{b,0}$ contribution, we apply a cut to the angular variable $\cos(\theta_{K^0_S,\pi_2})$, where $\theta_{K^0_S,\pi_2}$ is the angle between the $K^0_S$ candidate $p_T$ in the lab frame and the lower $p_T$ pion ($\pi_2$) in the $K^0_S$ center-of-mass frame. Cutting out events with $\cos(\theta_{K^0_S,\pi_2}) < -0.75$ removes 99.8% of the $\Lambda_{b,0}$ while retaining 86% of the $B^0$. The residual $\Lambda_{b,0}$ contamination is less than one event and is neglected. The invariant mass distribution for $J/\psi K^0_S$ and the fit result including the different contributions are shown in Fig. 3.

### TABLE II. Systematic uncertainties for the ratio of yields. All numbers in percent.

<table>
<thead>
<tr>
<th>Source</th>
<th>$\delta N(B^0 \rightarrow J/\psi K^{*0}) / N(B^0 \rightarrow J/\psi K^{*0})$ (%)</th>
<th>$\delta N(B^0 \rightarrow J/\psi K^{*0}) / N(B^0 \rightarrow J/\psi K^{*0})$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal modeling</td>
<td>4.4</td>
<td>4.6</td>
</tr>
<tr>
<td>Mass difference between $B^0$ and $B^*_0$</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Combinatorial background modeling</td>
<td>1.3</td>
<td>5.6</td>
</tr>
<tr>
<td>Combinatorial background contrib.</td>
<td>31.4</td>
<td>5.6</td>
</tr>
<tr>
<td>$B^*_0 \rightarrow J/\psi \phi$ contribution</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>31.8</td>
<td>9.2</td>
</tr>
</tbody>
</table>

### TABLE III. Variables used as input in the NN training.

<table>
<thead>
<tr>
<th>Input variables in the NN</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B^0/B^*_0$ candidate transverse momentump</td>
</tr>
<tr>
<td>$B^0/B^*_0$ candidate four-track decay point fit</td>
</tr>
<tr>
<td>$B^0/B^*_0$ candidate proper decay length</td>
</tr>
<tr>
<td>$B^0/B^*_0$ candidate impact parameter</td>
</tr>
<tr>
<td>$J/\psi$ candidate transverse momentump</td>
</tr>
<tr>
<td>$J/\psi$ candidate mass</td>
</tr>
<tr>
<td>$J/\psi$ candidate proper decay length</td>
</tr>
<tr>
<td>$J/\psi$ candidate impact parameter</td>
</tr>
<tr>
<td>$K^0_S$ candidate transverse momentump</td>
</tr>
<tr>
<td>$K^0_S$ candidate mass</td>
</tr>
<tr>
<td>$K^0_S$ candidate proper decay length</td>
</tr>
<tr>
<td>$K^0_S$ candidate impact parameter</td>
</tr>
<tr>
<td>$\pi$ transverse momentump</td>
</tr>
<tr>
<td>$\pi$ impact parameter</td>
</tr>
<tr>
<td>$\mu$ transverse momentump</td>
</tr>
<tr>
<td>$\mu$ impact parameter</td>
</tr>
<tr>
<td>$\mu$ cosine of the helicity angle in $J/\psi$ rest frame</td>
</tr>
</tbody>
</table>
We determine the yields of the $B^0 \rightarrow J/\psi K^0_s$ and $B^0 \rightarrow J/\psi K^0_S$ signals to be 5954 ± 79 and 64 ± 14, respectively. As with the $B^0 \rightarrow J/\psi K^0$ case, we determine the statistical significance of the $B^0 \rightarrow J/\psi K^0_S$ signal by fitting the mass distribution without the $B^0$ contribution (background-only hypothesis), a difference of 1 degree of freedom between the two hypotheses. For likelihood $\mathcal{L}$ we interpret $-2 \log \mathcal{L}$ as a $\chi^2$ and use the difference in that quantity to determine that the probability of background fluctuations producing a comparable or greater signal is $3.9 \times 10^{-13}$ or 7.2σ. The value of $N(B^0 \rightarrow J/\psi K^0_S)/N(B^0 \rightarrow J/\psi K^0)$ is 0.0108 ± 0.0019(stat).

The sources of systematic uncertainty are similar to the other analysis. In this case, the absolute uncertainties for the ratio are $6 \times 10^{-4}$ from the combinatorial background contribution, $6 \times 10^{-4}$ from the combinatorial background modeling, $5 \times 10^{-4}$ from the signal modeling, and $1.3 \times 10^{-5}$ from the mass difference between $B^0$ and $B^0$. The systematic uncertainties are summarized in Table II. We sum the contributions in quadrature resulting in a total systematic uncertainty of ±0.0010. The final value of $N(B^0 \rightarrow J/\psi K^0_S)/N(B^0 \rightarrow J/\psi K^0)$ is 0.0108 ± 0.0019(stat) ± 0.0010(syst).

V. ACCEPTANCE CALCULATION

To determine the ratio of branching ratios $\mathcal{B}(B^0 \rightarrow J/\psi K)/\mathcal{B}(B^0 \rightarrow J/\psi K)$, where $K$ represents $K^0_S$ or $K^{*0}$, the relative acceptances of $B^0 \rightarrow J/\psi K^0_S$ to $B^0 \rightarrow J/\psi K^0$ and $B^0 \rightarrow J/\psi K^{*0}$ to $B^0 \rightarrow J/\psi K^0$ need to be determined. We use MC samples to extract $A_{rel}$ as follows:

$$A_{rel} = \frac{N(B^0 \rightarrow J/\psi K_{\text{pass}})/N(B^0 \rightarrow J/\psi K_{\text{gen}})}{N(B^0 \rightarrow J/\psi K^0_{\text{pass}})/N(B^0 \rightarrow J/\psi K^0_{\text{gen}})},$$

where $N_{\text{gen}}$ is the number of MC generated signal events, $N_{\text{pass}}$ is the number of events passing all selection requirements, and $K$ represents $K^0_S$ or $K^{*0}$.

We determine the value for $A_{rel}$ to be 1.057 ± 0.010 for the $K^{*0}$ channel and 1.012 ± 0.010 for the $K^0_S$ channel. In both channels, the kinematics of the $B^0$ and $B^0$ final states are very similar to one another. The value for $A_{rel}$ in the $K^{*0}$ channel is larger than unity because the transverse decay length selection criteria removes more $B^0$ than $B^0$ because the $B^0$ lifetime is longer than the average $B^0$ lifetime. In the $K^0_S$ channel, $A_{rel}$ is close to unity because the $B^0$ lifetime is the long-lived $CP$ component and therefore close to the $B^0$ lifetime. We determine the statistical uncertainty on the
acceptances for $B^0$ and $B^0_s$, assuming binomial statistics. This MC statistical uncertainty is reported as a systematic uncertainty on $A_{\text{rel}}$.

The data sample utilized in this analysis was acquired using a number of variations on the $J/\psi \to \mu^+\mu^-$ trigger. We have verified that the acceptance calculation is robust and consistent across all kinematic variations of these triggers.

Several other effects contribute to the systematic uncertainty on $A_{\text{rel}}$. Uncertainty in $B^0$ and $B^0_s$ lifetimes introduces an uncertainty on the acceptance through the transverse decay length requirement. For $B^0_s \to J/\psi K^{(*)0}$, we generate different MC samples, varying the lifetimes by 1 standard deviation with respect to their measured values. We use the average measured value for $B^0$ and the evaluated $\tau_{B^0_s}$ value for $B^0_s$ [3]. The maximum deviation of $A_{\text{rel}}$ is 0.028, and we take this value as a systematic uncertainty.

For the $B^0_s \to J/\psi K^{(*)0}$ analysis, the procedure to evaluate the systematic uncertainty is slightly different. The $B^0_s \to J/\psi K^{(*)0}$ decay is an unknown admixture of CP-even and CP-odd states which have different lifetimes. The world-average currently gives $\Delta \Gamma_{B^0_s}/\Gamma_{B^0_s} = 0.092^{+0.051}_{-0.054}$ for $\Gamma_{B^0_s} = \frac{1}{2}(\Gamma_{B^0_s} + \Gamma_{B_s^{(*)}})$ [3], where $\Gamma_{B^0_s}$ and $\Gamma_{B_s^{(*)}}$ are the widths of the heavy and light mass eigenstates, respectively. If the $B^0_s$ were either all $B^0_s$ or all $B_s^{(*)}$, the maximum lifetime change would be 5%. To evaluate the effect on $A_{\text{rel}}$, we reweight the default $B^0_s \to J/\psi K^{(*)0}$ lifetime distribution. The reweighting is performed by normalizing the default lifetime distribution and comparing it to distributions with the lifetime increased or decreased by 5%. This leads to a maximum deviation on $A_{\text{rel}}$ of 0.046.

Another source of systematic uncertainty arises from the momentum spectra of the $B^0$ and $B^0_s$. Since we normalize our $B^0_s$ signal to the $B^0_s$ mode, we are sensitive only to mismodeling in the ratio of $p_T(B^0)$ versus $p_T(B^0_s)$, which should be quite small. We compare the default $B^0$ and $B^0_s$ samples which use a next-to-leading order QCD calculation [15] to the $p_T$ spectrum measured by CDF [7]. In the $B^0_s \to J/\psi K^{(*)0}$ analysis, the value of $A_{\text{rel}}$ varies by 0.029 when using these alternative production spectra, and we take this value as a systematic uncertainty. Likewise, for the $B^0_s \to J/\psi K^{(*)0}$ analysis, the change in $A_{\text{rel}}$ is 0.032.

Our relative acceptance is calculated assuming that the polarization in $B^0_s \to J/\psi K^{(*)0}$ is identical to the polarization in $B^0 \to J/\psi K^*$. Since we have no a priori knowledge of the actual polarization in the $B^0_s$ mode, we compute the systematic uncertainty by allowing all possible values for the polarization. We generated MC samples for $A_0 = 1$, $A_\parallel = 1$, and $A_\perp = 1$. The maximum variation from any of these polarizations leads to a systematic uncertainty on $A_{\text{rel}}$ of 0.261. Since the angular distributions arising from polarization are clearly the dominant systematic uncertainty, we have neglected the correlation between polarization and lifetime in assessing the uncertainties.

Table IV shows a summary of the systematic uncertainties on $A_{\text{rel}}$ for both measurements. Summing these contributions in quadrature, we find $A_{\text{rel}} = 1.057 \pm 0.010(\text{stat}) \pm 0.267(\text{syst})$ for the $K^{(*)0}$ analysis and $A_{\text{rel}} = 1.012 \pm 0.010(\text{stat}) \pm 0.042(\text{syst})$ for the $K^0_S$ analysis.

### Table IV. Systematic uncertainties for the relative acceptances. All numbers listed in percent.

<table>
<thead>
<tr>
<th>Source</th>
<th>$\delta A_{\text{rel}}(B^0_s \to J/\psi K^{(*)0})$ (%)</th>
<th>$\delta A_{\text{rel}}(B^0_s \to J/\psi K^{(*)0}_S)$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lifetime for $B^0$ and $B^0_s$</td>
<td>4.4</td>
<td>2.8</td>
</tr>
<tr>
<td>$B$ hadron $p_T$ spectrum</td>
<td>2.7</td>
<td>3.2</td>
</tr>
<tr>
<td>Polarization</td>
<td>24.7</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>25.3</td>
<td>4.2</td>
</tr>
</tbody>
</table>

VI. RESULTS

Using the values of $A_{\text{rel}}$ described above, we find

$$f_s/B(B^0_s \to J/\psi K^{(*)0})/f_d/B(B^0 \to J/\psi K^{(*)0}) = 0.0168 \pm 0.0024(\text{stat}) \pm 0.0068(\text{syst}) \quad (6)$$

and

$$f_s/B(B^0_s \to J/\psi K^{(*)0}_S)/f_d/B(B^0 \to J/\psi K^{(*)0}_S) = 0.0109 \pm 0.0019(\text{stat}) \pm 0.0011(\text{syst}). \quad (7)$$

To determine the ratio of branching ratios, we combine these results with the most recent CDF measurement [6] of $f_s/(f_s + f_d) \times B(D_s \to \phi \pi)$ and $f_s/f_d$ with the current world-average value [3] for $B(D_s \to \phi \pi)$ to yield $f_s/f_d = 0.269 \pm 0.033$. We quote the systematic uncertainty coming from the $f_s/f_d$ uncertainty as “frag”. The ratio of branching fractions to the reference $B^0$ decays are:

$$B(B^0_s \to J/\psi K^{(*)0})/B(B^0 \to J/\psi K^{(*)0}) = 0.062 \pm 0.009(\text{stat}) \pm 0.025(\text{syst}) \pm 0.008(\text{frag}) \quad (8)$$

and

$$B(B^0_s \to J/\psi K^{(*)0}_S)/B(B^0 \to J/\psi K^{(*)0}_S) = 0.041 \pm 0.007(\text{stat}) \pm 0.004(\text{syst}) \pm 0.005(\text{frag}). \quad (9)$$
The relative branching ratios observed for both modes are in good agreement with the expectation based upon the pure spectator model.

We use the world-average values for $\mathcal{B}(B^0 \to J/\psi K^{*0})$ and $\mathcal{B}(B^0 \to J/\psi K^0)$ [3] for normalization to calculate the absolute branching fractions:

$$\mathcal{B}(B^0_s \to J/\psi K^{*0}) = (8.3 \pm 1.2\text{(stat)} \pm 3.4\text{(syst)} \pm 1.0\text{(frag)}) \times 10^{-5}$$

(10)

and

$$\mathcal{B}(B^0_s \to J/\psi K^0) = (3.5 \pm 0.6\text{(stat)} \pm 0.4\text{(syst)} \pm 0.4\text{(frag)}) \times 10^{-5}.$$  

(11)

In conclusion, we present the first observation and branching ratio measurement of the Cabibbo-suppressed decays $B^0_s \to J/\psi K^{*0}$ and $B^0_s \to J/\psi K^0$. With larger data samples and additional analysis, these modes can be used to further explore the properties of the $B^0_s$ system.

[10] The CDF reference frame uses cylindrical coordinates, where $\theta$ and $\phi$ are the polar and azimuthal angles with respect to the proton beam. Pseudorapidity $\eta$ is defined as $-\ln(\tan(\theta/2))$. Transverse momentum $p_T$ is the charged particle momentum measured in the plane perpendicular to the beam line.