DØnote 4869-CONF

Observation of the Decay $B_s^0 \to \psi(2S)\,\phi$ and a Measurement of $\mathcal{B}(B_s^0 \to \psi(2S)\,\phi)/\mathcal{B}(B_s^0 \to J/\psi\,\phi)$ with the DØ detector*

The DØ Collaboration
(Dated: July 15, 2005)

We report in this note the observation of the decay $B_s^0 \to \psi(2S)\,\phi$ using about 300 pb^{-1} of Run II data collected with the DØ detector at Tevatron. Using the decay $B_s^0 \to J/\psi\,\phi$ as normalization we report a measurement of the relative branching ratio:

$$\frac{\mathcal{B}(B_s^0 \to \psi(2S)\,\phi)}{\mathcal{B}(B_s^0 \to J/\psi\,\phi)} = 0.58 \pm 0.24 \text{ (stat)} \pm 0.06 \text{ (sys)} \pm 0.07(\mathcal{B}).$$

Preliminary Results for Summer 2005 Conferences

* In this note the charge conjugated states are included implicitly.

I. INTRODUCTION

Studies of the decays of B mesons to $\psi(2S)$ final states have contributed to knowledge of hadronic B-meson decays and provide insight into the interplay between weak and strong interactions. Historically the decay $B^{\pm} \rightarrow \psi(2S)\,K^{\pm}$ was first observed [1] at ARGUS, $B^0 \rightarrow \psi(2S)\,K^{*0}$ was observed [2] at CDF (Run I), and CLEO observed [3] $B^0 \rightarrow \psi(2S)\,K_S$ and $B^{\pm} \rightarrow \psi(2S)\,K^{\star\pm}$. Subsequently all these decay modes have been studied with more statistics by many experiments. The measurements show that the rates of $B^{\pm}$ and B^0 mesons decay to the $\psi(2S)$ final states are approximately 60% of the rates of the corresponding decay to the J/ψ final state. For the B_s^0 meson, up to now the decay $B_s^0 \rightarrow \psi(2S)\phi$ has only been observed by the ALEPH collaboration which reported [4] one candidate event in 1993 when they measured the B_s^0 mass. A branching ratio of $B_s^0 \rightarrow \psi(2S)\,\phi$ to $B_s^0 \rightarrow J/\psi\,\phi$ has not yet been measured. In this note we report on the observation of $B_s^0 \rightarrow \psi(2S)\,\phi$ and on a measurement of $\mathcal{B}(B_s^0 \rightarrow \psi(2S)\,\phi)$ relative to $\mathcal{B}(B_s^0 \rightarrow J/\psi\,\phi)$. The two-body decays of the pseudoscalar B_s^0 meson into two vector meson states $(J/\psi,\,\psi(2S))\,\phi$ are described as admixtures of CP eigenstates. Therefore, both CP even and odd eigenstates of the B_s^0 meson with lifetime differences contribute to the decay.

As control channel we have measured the branching ratio $\mathcal{B}(B^{\pm} \rightarrow \psi(2S)\,K^{\pm})$ relative to $\mathcal{B}(B^{\pm} \rightarrow J/\psi\,K^{\pm})$.

II. THE DØ DETECTOR

The search uses data set of approximately 300 pb^{-1} of $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV of Run II recorded by the DØ detector operating at the Fermilab Tevatron. The DØ detector is described elsewhere [5]. The main elements relevant for 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 within a 2 T superconducting solenoidal magnet. The muon detector located outside the calorimeter consists of a layer of tracking detectors and scintillation trigger counters in front of toroidal magnets (1.8 T), followed by two more similar layers after the toroids, allowing for efficient detection out to pseudorapidity (η) of about ± 2.0.

The data selected in this analysis were triggered by four versions of the scintillator based dimuon triggers. Trigger efficiencies for signal and normalization samples were estimated using a trigger simulation software package. These efficiencies were also checked with data samples collected with unbiased or single muon triggers.

III. EVENT SELECTION

The pre-selection starts with a loose selection of $B^{\pm}$ (B_s^0) candidates consisting of two identified muons and one (two oppositely) charged track(s) forming a good vertex. In the selection step for the B_s^0, the mass of the two kaon candidate tracks should be between $0.980 < m_\phi < 1.080$ GeV/c^2. The allowed mass of the loose B candidate is required to be within $4.4 < m_B < 6.2$ GeV/c^2. For $B^{\pm}(B_s^0) \rightarrow J/\psi\,K^{\pm}\,(\phi)$ and $B^{\pm}(B_s^0) \rightarrow \psi(2S)\,K^{\pm}\,(\phi)$ candidates, the invariant mass of the muon pair is required to be within 250 MeV/c^2 of the J/ψ and $\psi(2S)$ masses [7], respectively.

The $\chi^2/d.o.f.$ of the two muon vertex is required to be $\chi^2/d.o.f. < 16$. The transverse momentum of each of the muons is required to be greater than 2.0 GeV/c and their pseudorapidity has to be $|\eta| < 2.0$ to be well inside the fiducial tracking and muon region. Tracks that are matched to each muon leg need at least one hit in the SMT and one hit in the CFT.

For surviving events, the two-dimensional decay length of the B candidates in the plane transverse to the beamline L_{xy} is calculated. This length L_{xy} is defined as the projection of the decay length vector $\vec{l}_{Vtx}$ on the transverse momentum of the B-meson:

$$L_{xy} = \frac{\vec{l}_{Vtx} \cdot \vec{p}_T^{\,B}}{p_T^B} \tag{1}$$

The error on the transverse decay length δL_{xy} is calculated by taking into account the uncertainties on both the primary and secondary vertex positions. The primary vertex itself is found with a beam spot constrained fit. It is requested that δL_{xy} has to be smaller than 150 μm. The transverse momentum of the $\mu^+\mu^-$ system (either J/ψ or $\psi(2S)$) needs to be greater than 4 GeV/c

A. Reconstruction of $B^{\pm} \rightarrow (J/\psi, \psi(2S))K^{\pm}$ events

To reconstruct the decays $B^{\pm} \rightarrow (J/\psi, \psi(2S))K^{\pm}$ with $(J/\psi, \psi(2S)) \rightarrow \mu^+\mu^-$, the two muons are in addition requested to have at least the medium quality criterium fulfilled. The candidates are then constrained to have an invariant mass equal to the $J/\psi, \psi(2S)$ mass [7], respectively. The combined vertex fit of the J/ψ and the additional $K^\pm$ should not yield a χ^2 of more than 20 for 3 $d.o.f.$. The p_T of the $K^\pm$ should be larger than 0.9 GeV/c. Moreover, a collinearity between decay length vector of the $B^\pm$ and the combined momentum of J/ψ and $K^\pm$ in the transverse plane of at least 0.9 is requested. To remove prompt background the decay length significance $L_{xy}/\delta L_{xy}$ of the $B^\pm$ candidate has to be larger than 4. The obtained fit results, using a gaussian for the signal peak and a second order polynomial for the combinatorial background, are summarized in Table I for Monte Carlo and in Table II for the full data sample. The data yields are obtained by leaving all parameters floating. The mean for both channels decaying either in J/ψ or $\psi(2S)$ is lower than expected but the width are consistent with the MC expectations. The change in the yield if all values are fixed to the MC expectation will be discussed in section VI A. In Fig. 1 the resulting invariant mass distribution for the $B^\pm \to J/\psi K^\pm$ and in Fig. 2 the distribution for $B^\pm \to \psi(2S)K^\pm$ are shown.

Decay	Mean [MeV/c^2]	Width [MeV/c^2]	ε_{MC}
$B^\pm \to J/\psi K^\pm$	5278.8 ± 0.8	37.6 ± 0.8	$(1.14\pm0.02)\cdot10^{-3}$
$B^\pm \to \psi(2S)K^\pm$	5278.3 ± 1.0	29.0 ± 0.6	$(1.07\pm0.04)\cdot10^{-3}$

TABLE I: Summary of fitting results of $B^\pm \to (J/\psi, \psi(2S))K^\pm$ final states for MC.

Decay	Mean [MeV/c^2]	Width [MeV/c^2]	Yield
$B^\pm \to J/\psi K^\pm$	5273.7 ± 1.2	40.1 ± 1.2	1970±44
$B^\pm \to \psi(2S)K^\pm$	5269.9 ± 3.2	27.1 ± 2.9	149±18

TABLE II: Summary of fitting results of $B^\pm \to (J/\psi, \psi(2S))\, K^\pm$ final states for the data.

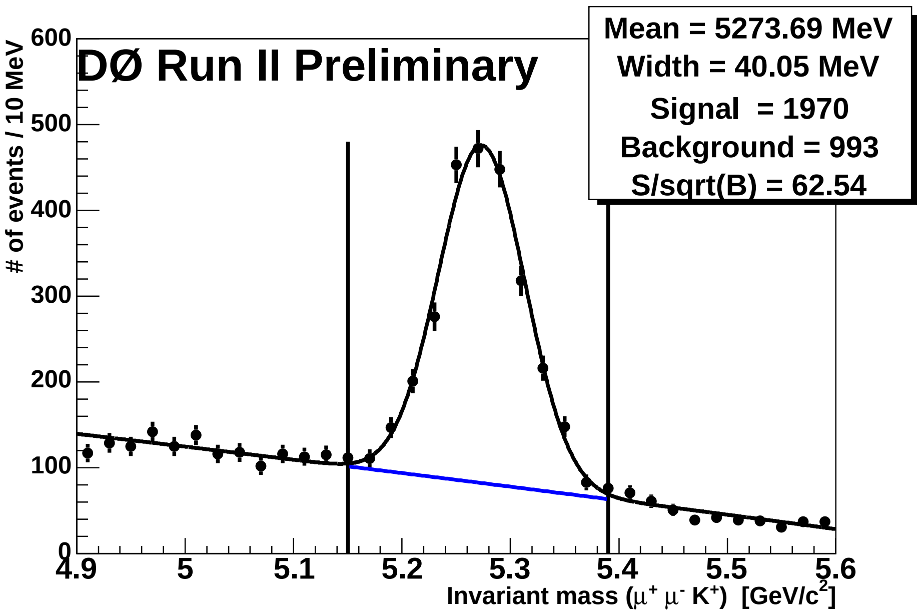

FIG. 1: Invariant mass distribution for the $B^\pm \to J/\psi\, K^\pm$ decay for the data sample.

B. Reconstruction of $B_s^0 \to J/\psi\, \phi$ and $B_s^0 \to \psi(2S)\, \phi$ events

To reconstruct the decays $B_s^0 \to (J/\psi, \psi(2S))\, \phi$ with $(J/\psi, \psi(2S)) \to \mu^+\mu^-$, the two muons are in addition requested to have at least a hit in the first layer of the muon system and are matched to a central track ($nseg > 0$ [6]).

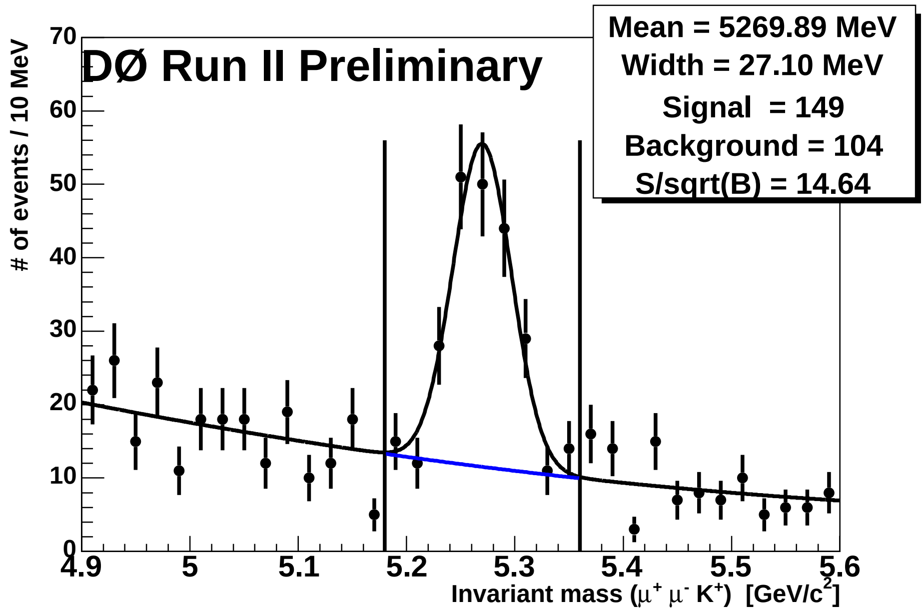

FIG. 2: Invariant mass distribution for the $B^{\pm} \to \psi(2S)\, K^{\pm}$ decay for the data sample.

They are constrained to have an invariant mass equal to the $J/\psi, \psi(2S)$ mass [7], respectively. The candidates are then combined with another pair of oppositely charged tracks (ϕ candidate), each with $p_T > 0.9$ GeV/c, to form a B_s^0 candidate vertex with a $\chi^2 < 36$ with 5 d.o.f.. Each of the kaon candidates needs at least one hit in the SMT and the ϕ candidate is required to have an invariant mass between 1.008 and 1.032 GeV/c^2. To remove prompt background the decay length significance $L_{xy}/\delta L_{xy}$ of the B_s^0 candidate has to be larger than 4. The obtained fit results, using a gaussian and a second order polynomial, are summarized in Table III for Monte Carlo and in Table IV for the full data sample. The yields for the $B_s^0 \to J/\psi\,\phi$ decay are obtained by leaving all parameters floating. For the $B_s^0 \to \psi(2S)\,\phi$ decay however, the values for the mean and width were fixed. The mean was set to the obtained value for the $B_s^0 \to J/\psi\,\phi$ and the width was obtained by scaling the expected width of the MC with the factor obtained from the MC/data difference [8] for the $B_s^0 \to J/\psi\,\phi$ decay. In Fig. 3 the resulting invariant mass distribution for the $B_s^0 \to J/\psi\,\phi$ and in Fig. 4 the distribution for $B_s^0 \to \psi(2S)\,\phi$ are shown.

Decay	Mean [MeV/c^2]	Width [MeV/c^2]	ε_{MC}
$B_s^0 \to J/\psi\,\phi$ (CP even)	5369.6 ± 1.3	26.6 ± 1.2	(1.94±0.08)·10^{-4}
$B_s^0 \to J/\psi\,\phi$ (CP odd)	5370.2 ± 1.3	28.0 ± 1.2	(2.05±0.09)·10^{-4}
$B_s^0 \to \psi(2S)\,\phi$ (CP even)	5370.1 ± 0.8	22.5 ± 0.7	(2.08±0.07)·10^{-4}

TABLE III: Summary of fitting results of $B_s^0 \to (J/\psi, \psi(2S))\,\phi$ final states for MC.

Decay	Mean [MeV/c^2]	Width [MeV/c^2]	Yield
$B_s^0 \to J/\psi\,\phi$	5356.7 ± 2.7	28.9 ± 2.3	200±18
$B_s^0 \to \psi(2S)\,\phi$	5356.7 (fixed)	24.4 (fixed)	13±8

TABLE IV: Summary of fitting results of $B_s^0 \to (J/\psi, \psi(2S))\,\phi$ final states for the data.

IV. DISCRIMINATING VARIABLES AND CUT OPTIMIZATION

To further enhance the signal to background ratio of the signal peak we have used the same three discriminating variables that were already employed in the search for $B_s^0 \to \mu^+\mu^-$ [9]. We restrict ourselves to a mass region of interest of $4.8 < M_{\psi(2S)\phi} < 6.0$ GeV/c^2 containing the signal region around the PDG [7] world average value of the

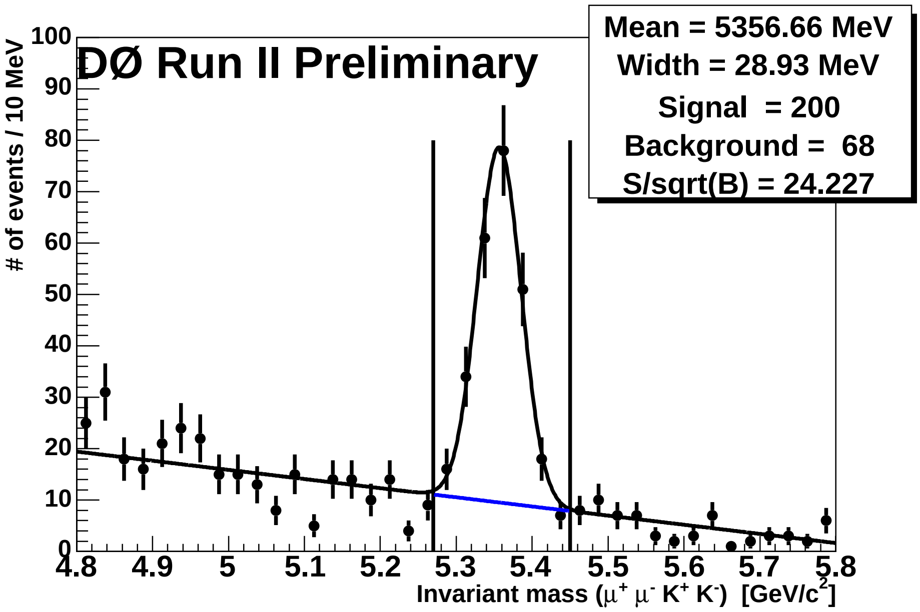

FIG. 3: Invariant mass distribution for the $B_s^0 \to J/\psi\,\phi$ decay for the data sample.

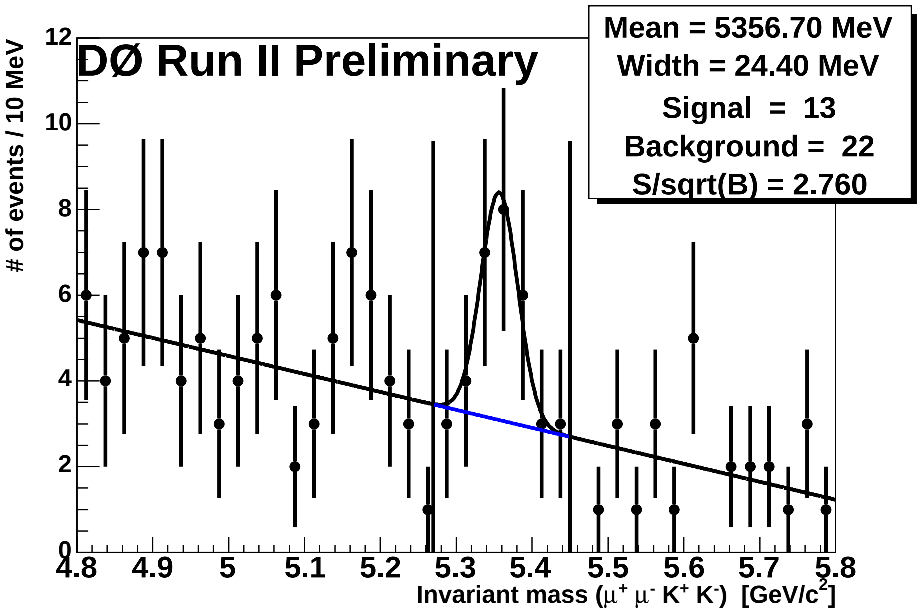

FIG. 4: Invariant mass distribution for the $B_s^0 \to \psi(2S)\,\phi$ decay for the data sample using a loose decay length significance cut of greater than four.

B_s^0 mass of $m_{B_s^0} = 5369.6 \pm 2.4$ MeV/c^2 and left the signal region hidden during optimization. Table IV defines the regions for the sidebands and the signal box that have been used. The given values translate the size of the signal region to a window of ± 150 MeV/c^2 around the B_s^0 mass. The expected mass resolution for $B_s^0 \to \psi(2S)\phi$ in the MC is ≈ 25 MeV/c^2, the chosen mass window is therefore sufficiently large to cover a $\pm 6\sigma$ window.

The cut optimization procedure is done on the pre-selected sample. We use as first discriminating variable the

isolation I of the ϕ meson and muon pair which is defined as:

$$I = \frac{|\vec{p}(\psi(2S)\phi)|}{|\vec{p}(\psi(2S)\phi)| + \sum\limits_{track\, i \neq B} p_i(\Delta\mathcal{R} < 1)}.$$

$\sum\limits_{track\, i \neq B} p_i$ is the scalar sum over all tracks excluding the muon and kaon pair within a cone of $\Delta\mathcal{R} < 1$ (where $\Delta\mathcal{R} = \sqrt{(\Delta\Phi)^2 + (\Delta\eta)^2}$) centered around the momentum vector $\vec{p}(\psi(2S)\phi)$ of the B_s^0 candidate.

All tracks that are counted in the isolation sum have the additional requirement that the z distance of the track to the z-vertex of the muon pair has to be smaller than 5 cm in order to avoid overlapping min-bias events from the same bunch crossing. The distribution of the isolation variable for signal MC and sideband data after pre-selection is shown in Fig. 5.

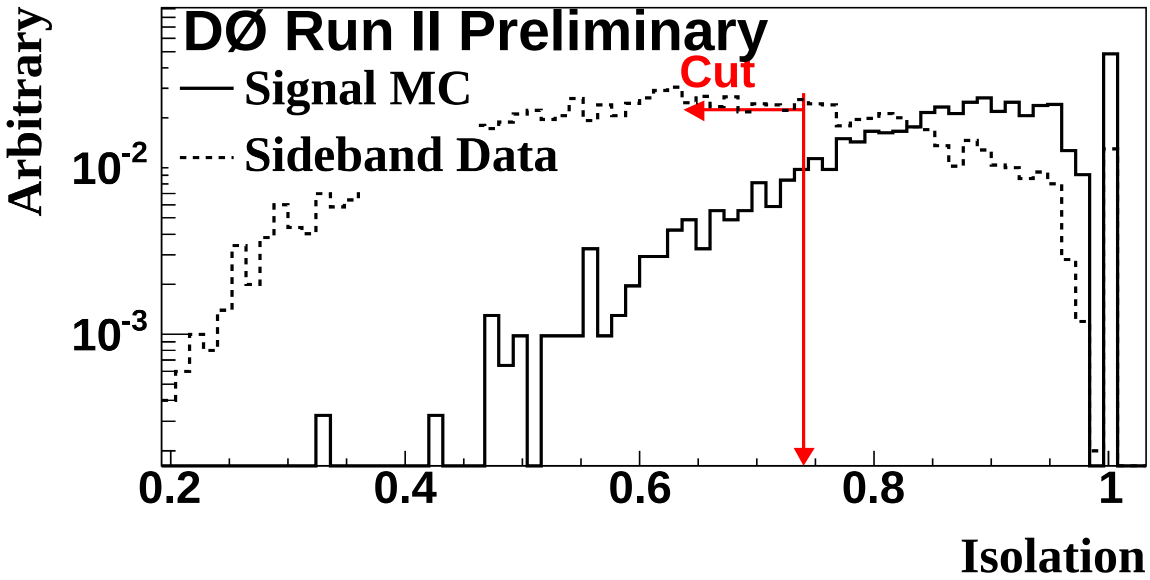

FIG. 5: Isolation variable after the pre-selection for sideband data and signal MC events.

The **pointing angle** α is used as second discriminating variable and is defined as the angle between the momentum vector $\vec{p}(\psi(2S)\,\phi)$ of the dimuon pair and the vector $\vec{l}_{Vtx}$ pointing from the primary vertex to the secondary vertex. If the muon pair is coming from the decay of a parent particle B_s^0, the vector $\vec{l}_{Vtx}$ should point into the same direction as $\vec{p}(\psi(2S)\phi)$. The angle α is well-defined and used as a consistency between the direction of the decay vertex and the flight direction of the B_s^0 candidate. Fig. 6 shows the distributions of the angle α for signal MC and data after pre-selection.

In order to discriminate against short-lived background, we have finally used the **transverse decay length significance** $L_{xy}/\delta_{L_{xy}}$ since it gives a better discriminating power then the transverse decay length alone. Figure 7 shows the distribution of the decay length significance for signal MC and data.

Region	min Mass (GeV/c^2)	max Mass (GeV/c^2)
region of interest	4.80	6.00
hidden signal region during optimization	5.22	5.52
sideband I	4.80	5.22
sideband II	5.52	6.00

TABLE V: The different four track invariant mass regions for signal and sidebands used for background estimation.

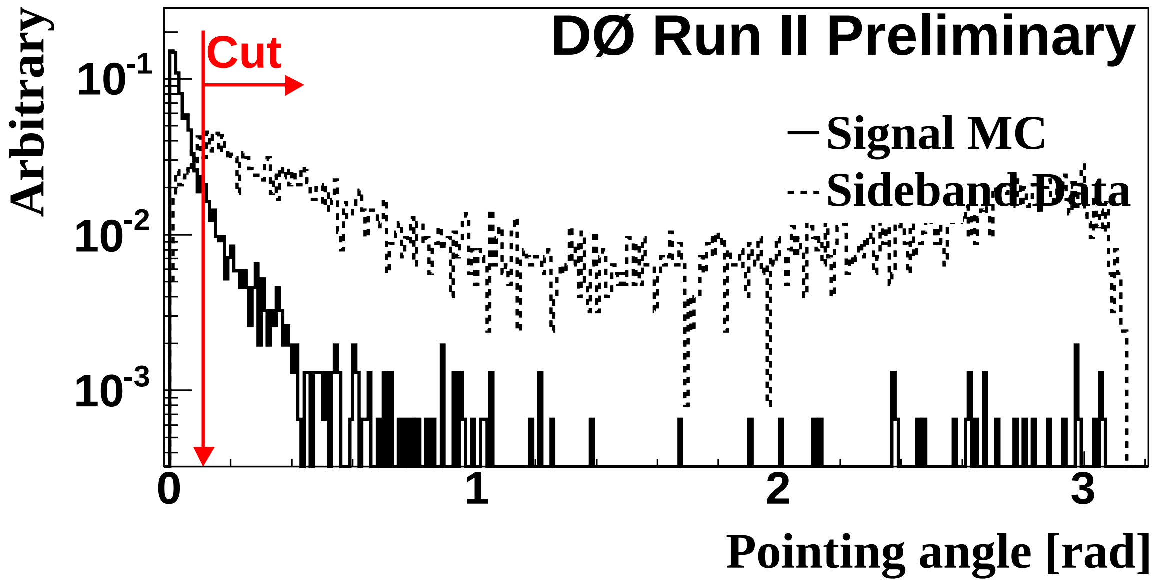

FIG. 6: The pointing angle α after the preselection for sideband data and signal MC events.

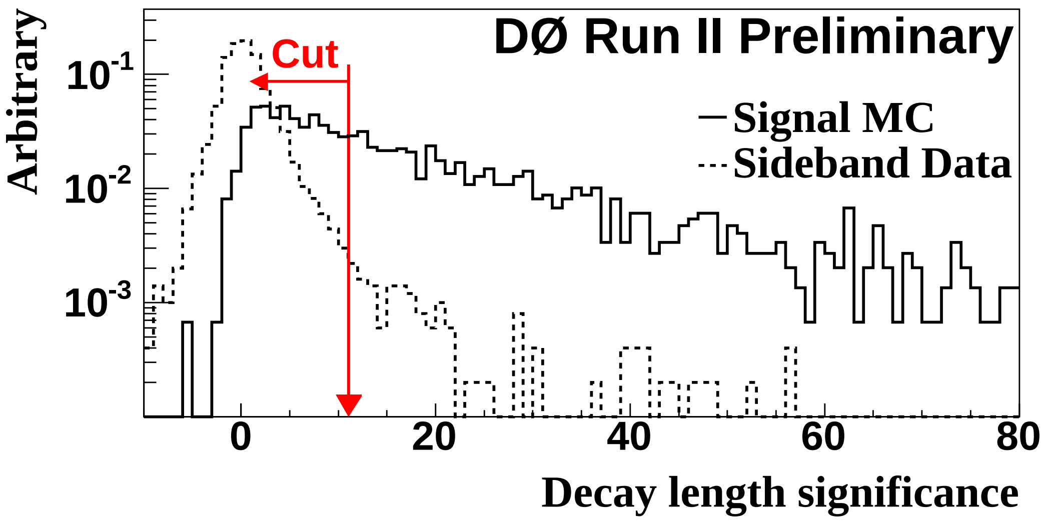

FIG. 7: The transverse decay length significance after the preselection for sideband data and signal MC events.

To find the optimal set of cuts we used a Random Grid Search (RGS) [10] and an optimization criterion proposed by G. Punzi [11]. The ratio P defined as

$$P = \frac{\epsilon_{\psi(2S)\phi}}{\frac{a}{2} + \sqrt{N_{Back}}} \tag{2}$$

was maximized. Here, $\epsilon_{\psi(2S)\phi}$ is the reconstruction efficiency of the signal MC after the pre-selection and N_{Back} is the expected number of background events extrapolated from the sidebands. The constant a is the number of sigmas corresponding to the confidence level at which the signal hypothesis is tested. This number a should be defined before the statistical test and has been set to 5, corresponding to an observation at a 5σ discovery. The resulting cut values that were obtained from the maximized P are listed in table VI

The total signal efficiency relative to pre-selection of the three discriminating cuts is $(50.4 \pm 2.5)\%$ with the uncertainty due to MC statistics. After a linear interpolation of the sideband population for the whole data sample into the final signal region we obtain an expected number of background events of $0.8{\pm}0.4$. Figure 8 shows the

cut parameter	cut value	MC efficiency (%)
Opening angle (rad)	< 0.11	79.6 ± 1.3
Decay length significance	> 11.1	69.9 ± 1.7
Isolation	> 0.74	90.9 ± 1.3

TABLE VI: The optimized cuts and their relative MC signal efficiencies after maximizing P.

remaining invariant mass distribution. In the signal region 11 candidate events are found. The probability that the expected background fluctuates into the observed 11 events or more is given by $p = 1.04 \cdot 10^{-9}$. The significance of this enhancement quoted in number of standard deviations corresponds to 5.99σ.

Assuming a gaussian-like signal peak added with a linear background we find in a log-likelihood fit 8.6±3.3 signal events with 1.8±1.3 background events. The probability of a background only fluctuation is $p = 1.1 \cdot 10^{-4}$. The significance S of the fitted signal peak can be evaluated as $S = \sqrt{-2\ln\frac{\mathcal{L}_0}{\mathcal{L}_{max}}} = 3.89$ with $\mathcal{L}_{max}$ being the Maximum-Likelihood of the best fit assuming both signal and background, while $\mathcal{L}_0$ being the best fit if the signal yield is set to zero. For the log-likelihood fit the mean and width of the gaussian have been fixed as described in section III B

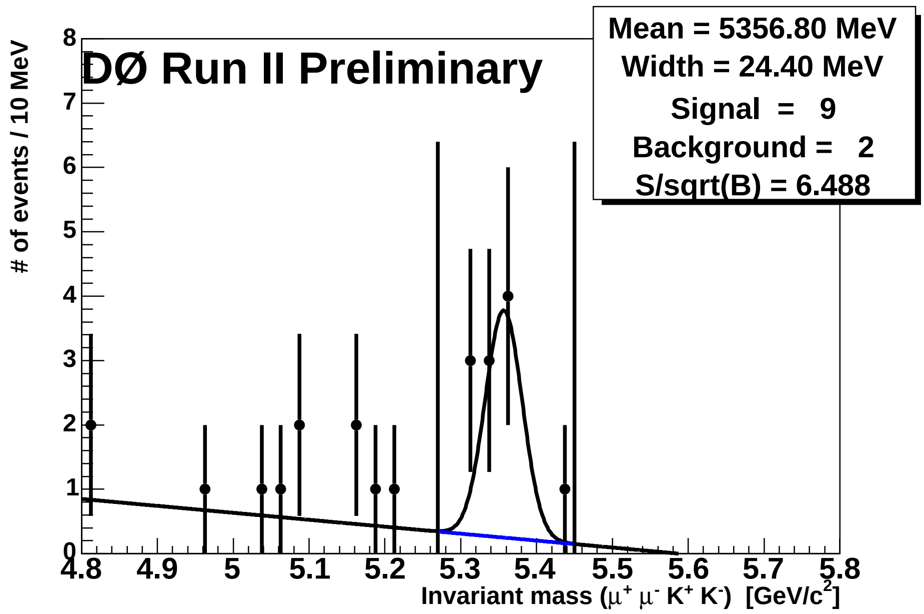

FIG. 8: Invariant mass distribution for the $B_s^0 \to \psi(2S)\phi$ decay for the data sample after cut optimization.

V. THE RELATIVE BRANCHING RATIO MEASUREMENTS

A. The measurement of $\mathcal{B}(B_s^0 \to \psi(2S)\,\phi)/\mathcal{B}(B_s^0 \to J/\psi\,\phi)$

Since we see a significant excess of signal events we use the $B_s^0 \to J/\psi\,\phi$ as normalization and calculate the ratio of the branching fractions as follows [12]:

$$\frac{\mathcal{B}(B_s^0 \to \psi(2S)\,\phi)}{\mathcal{B}(B_s^0 \to J/\psi\,\phi)} = \frac{N_{B_s^0 \to \psi(2s)\,\phi}}{N_{B_s^0 \to J/\psi\,\phi}} \cdot \frac{\epsilon_{J/\psi\,\phi}}{\epsilon_{\psi(2S)\,\phi}} \cdot \frac{\mathcal{B}(J/\psi \to \mu^+\mu^-)}{\mathcal{B}(\psi(2S) \to \mu^+\mu^-)} \tag{3}$$

where

- $\epsilon_{\psi(2S)\phi}$ and $\epsilon_{J/\psi\phi}$ are the efficiencies of the signal and normalization channels, obtained from MC simulations

- the measured branching fractions are $\mathcal{B}(J/\psi \to \mu^+\mu^-) = (5.88 \pm 0.1)\%$ and $\mathcal{B}(\psi(2S) \to \mu^+\mu^-) = (7.3 \pm 0.8) \cdot 10^{-3}$ [7].

The efficiencies $\epsilon_{\psi(2S)\phi}$ and $\epsilon_{J/\psi\phi}$ are the global signal efficiencies for the signal and normalization channel respectively including the pre-selection cuts and the acceptance. In Table VII the various MC efficiencies determined from MC needed to calculate the ratio of the branching ratio are given. The numbers refer to triggered MC events in the trigger simulation with uncertainties that are due to MC statistics.

	$\epsilon_{\psi(2S)\phi}$	$\epsilon_{J/\psi\phi}$	$\epsilon_{J\psi\phi}/\epsilon_{\psi(2S)\phi}$
RGS (CP even)	$(12.5 \pm 0.6) \cdot 10^{-5}$	$(11.5 \pm 0.7) \cdot 10^{-5}$	$(91.7 \pm 7.2)\%$

TABLE VII: Efficiencies for the two decay channels (both channels CP even only) after trigger and reconstruction with respect to the generated $b\bar{b}$-pairs.

The observed number of $B_s^0 \to J/\psi\,\phi$ events obtained from a gaussian fit after all discriminating cuts is 110 ± 11 events. This gives $\mathcal{B}(B_s^0 \to \psi(2S)\phi)/\mathcal{B}(B_s^0 \to J/\psi\phi) = 0.60 \pm 0.21$ (stat).

B. The measurement of $\mathcal{B}(B^\pm \to \psi(2S)\,K^\pm)/\mathcal{B}(B^\pm \to J/\psi\,K^\pm)$

For the measurement of the ratio $\mathcal{B}(B^\pm \to \psi(2S)\,K^\pm)$ to $\mathcal{B}(B^\pm \to J/\psi\,K^\pm)$ the values from Table I and II one can calculate the ratio $\mathcal{B}(B^\pm \to \psi(2S)K^\pm)/\mathcal{B}(B^\pm \to J/\psi K^\pm) = 0.57\pm$ 0.07 (stat). To calculate this ratio a similar formula as given in Eq. 3 has been used with simply exchanging the ϕ with the $K^\pm$ mesons.

A comparison between the obtained values in this analysis and published values is given in Table VIII.

Decay	This analysis	Reference
$\frac{\mathcal{B}(B_d^0 \to \psi(2S)\,K^0)}{\mathcal{B}(B_d^0 \to J/\psi K^0)}$		$0.82 \pm 0.13 \pm 0.12$ (PDG 04,[7])
$\frac{\mathcal{B}(B_d^0 \to \psi(2S) K^*)}{\mathcal{B}(B_d^0 \to J/\psi K^*)}$		$0.61 \pm 0.19 \pm 0.06$ (PDG 04,[7])
$\frac{\mathcal{B}(B^\pm \to \psi(2S) K^\pm)}{\mathcal{B}(B^\pm \to J/\psi K^\pm)}$	0.57 ± 0.07	$0.64 \pm 0.06 \pm 0.06$ (BaBar 02,[13])
$\frac{\mathcal{B}(B_s^0 \to \psi(2S)\phi)}{\mathcal{B}(B_s^0 \to J/\psi\phi)}$	0.58 ± 0.24	

TABLE VIII: The measured relative branching ratios from this analysis and published results.

VI. SYSTEMATIC STUDIES

For the measurement of the relative branching ratio different types of systematics are involved:

- Systematics due to the branching ratio $\mathcal{B}(J/\psi \to \mu^+\mu^-)\,/\,\mathcal{B}(\psi(2S) \to \mu^+\mu^-) \;=\; 8.05 \pm 0.89$. This uncertainty is given by the uncertainty on the single measured branching ratios assuming no correlations.
- Systematics due to the signal yield determination.
- Systematics due to the determination of the efficiencies $\epsilon_{\psi(2S)\phi}$ and $\epsilon_{J/\psi\phi}$. In the ratio one expects most effects to cancel out. This is due to the fact, that both decay modes have very similar topologies. For the B_s^0 decays the polarization however could be different. The generated signal MC was a pure CP even state, for the normalization channel however also a pure CP odd state was generated.

A. Measurement of $\mathcal{B}(B^\pm \to \psi(2S)\,K^\pm)/\mathcal{B}(B^\pm \to J/\psi\,K^\pm)$

All the relative uncertainties that go into the calculation of the relative branching ratio are given in table IX. The relative statistical uncertainties on $\epsilon_{\psi(2S)\,K}$ and $\epsilon_{J/\psi\,K}$ are 3.7% and 1.9% respectively. They are combined into one efficiency uncertainty number assuming no correlations. As mentioned in section III A the fit parameters (mean and width of the signal gaussian) for the $B^{\pm} \to \psi(2S)\,K^{\pm}$ candidates are smaller than expected, when all parameters are floating. To get an estimate of the signal yield variation due to that, we can fix these values to the ones obtained from MC and scaled by the data/MC difference for the $B^{\pm} \to J/\psi\,K^{\pm}$ events, as already done in section III B for the $B_s^0 \to \psi(2S)\,\phi$ events. We then obtain 157±18 events, which is an increase of 5%. We will take this as nominal value and use the difference to assign a systematic variation to the ratio. The relative branching ratio $\mathcal{B}(B^{\pm} \to \psi(2S)K^{\pm})/\mathcal{B}(B^{\pm} \to J/\psi K^{\pm})$ is then 0.60 ± 0.07 (stat) ± 0.04 (sys) ± 0.06 ($\mathcal{B}$).

Source	Relative Uncertainty [%]
$\mathcal{B}(J/\psi \to \mu\mu)$	1.7
$\mathcal{B}(\psi(2S) \to \mu\mu)$	11.0
Total $\mathcal{B}$	11.1
$\epsilon_{J/\psi K}/\epsilon_{\psi(2s)K}$	4.1
Signal yield	5
Total (sys)	4.1

TABLE IX: The relative uncertainties for the measurement of the relative branching ratio $\mathcal{B}(B^{\pm} \to \psi(2S)\,K^{\pm})/\mathcal{B}(B^{\pm} \to J/\psi\,K^{\pm})$

B. Measurement of $\mathcal{B}(B_s^0 \to \psi(2S)\,\phi)/\mathcal{B}(B_s^0 \to J/\psi\,\phi)$

All the relative uncertainties that go into the calculation of the relative branching ratio are given in table X. The relative statistical uncertainties on $\epsilon_{\psi(2S)\phi}$ and $\epsilon_{J/\psi\phi}$ are 4.5% and 5.6% respectively. They are combined into one efficiency uncertainty number assuming no correlations.

Source	Relative Uncertainty [%]
$\mathcal{B}(J/\psi \to \mu\mu)$	1.7
$\mathcal{B}(\psi(2S) \to \mu\mu)$	11.0
Total $\mathcal{B}$	11.1
$\epsilon_{J/\psi\phi}/\epsilon_{\psi(2s)\phi}$	7.2
CP odd-even ($J/\psi\phi$)	7.2
Total (sys)	10.2

TABLE X: The relative uncertainties for the measurement of the relative branching ratio $\mathcal{B}(B_s^0 \to \psi(2S)\phi)/\mathcal{B}(B_s^0 \to J/\psi\phi)$

The relative branching ratio $\mathcal{B}(B^{\pm} \to \psi(2S)\,\phi)/\mathcal{B}(B^{\pm} \to J/\psi\,\phi)$ is then 0.58 ± 0.24 (stat) ± 0.06 (sys) ± 0.07 ($\mathcal{B}$).

VII. CONCLUSIONS

We have presented the observation of the $B_s^0 \to \psi(2S)\phi$ decay with a significance of 5.99σ based on counting statistics. After a gaussian fit with linear background we find 8.6±3.3 signal events above 1.8±1.3 background events. We have measured the relative branching ratio to be $\mathcal{B}(B_s^0 \to \psi(2S)\,\phi)/\mathcal{B}(B_s^0 \to J/\psi\,\phi) = 0.58 \pm 0.24$ (stat) ± 0.06 (sys) ± 0.07 ($\mathcal{B}$) using about 300 pb^{-1}. As control check we have performed a measurement of the relative branching ratio $\mathcal{B}(B^{\pm} \to \psi(2S)\,K^{\pm})/\mathcal{B}(B^{\pm} \to J/\psi\,K^{\pm})$ to 0.60 ± 0.07 (stat) ± 0.04 (sys) ± 0.06 ($\mathcal{B}$).

[1] H. Albrecht *et al.* [ARGUS Collaboration], Phys. Lett. B **199**, 451 (1987).
[2] F. Abe *et al.* [CDF Collaboration], Phys. Rev. D **58**, 072001 (1998).
[3] S.J. Richichi *et al.* [CLEO2 Collaboration], Phys. Rev. D **63**, 031103R (2001).
[4] D. Buskulic *et al.* [ALEPH Collaboration], Phys. Lett. B **311**, 425 (1993).

[5] V. M. Abazov *et al.*, in preparation for submission to Nucl. Instrum. Methods Phys. Res. A, and T. LeCompte and H.T. Diehl, "The CDF and DØ Upgrades for Run II", Ann. Rev. Nucl. Part. Sci. **50**, 71 (2000).
[6] http://www-d0.fnal.gov/computing/algorithms/muon/p14muon-v02.ps.
[7] S. Eidelman *et al.* Phys. Lett B **592**, 1 (2004).
[8] $\sigma^{Data}_{B^0_s \to \psi(2S)\,\phi} = \frac{\sigma^{Data}_{B^0_s \to J/\psi\,\phi}}{\sigma^{MC}_{B^0_s \to J/\psi\,\phi}} \cdot \sigma^{MC}_{B^0_s \to \psi(2S)\,\phi}$ and using the CP even MC for $B^0_s \to J/\psi\,\phi$
[9] V. M. Abzov *et al.*, Phys. Rev. Lett. **94**, 071802 (2005), hep-ex/0410039.
[10] H. B. Prosper, Proceedings of the International Conference on Computing in High Energy Physics (CHEP'95), Rio de Janeiro, 18-22 September 1995
[11] G. Punzi, Talk given at the Conference on Statistical Problems in Particle Physics, Astrophysics and Cosmology (Phystat 2003), SLAC, Stanford, California, 8-11 September 2003, physics/0308063
[12] The same formula is also applied to calculate the branching ratio for the $B^{\pm}$ decay with ϕ replaced by $K^{\pm}$.
[13] B. Aubert *et al.* [BABAR Collaboration], Phys. Rev. D **65**, 032001 (2002), [hep-ex/0107025].