



Search for the Rare Decay $B_s^0 \rightarrow \mu^+ \mu^- \phi$ with the DØ Detector*

The DØ Collaboration
URL <http://www-d0.fnal.gov>
(Dated: June 27, 2005)

We present in this note a search for the rare decay $B_s^0 \rightarrow \mu^+ \mu^- \phi$ using approximately 300 pb⁻¹ of Run II data collected with the DØ detector at the Fermilab Tevatron. Our search is based on a blind analysis hiding the signal region around the B_s^0 mass. The sideband regions below and above the blinded signal region are used to determine the shape and normalization of the background. Events are normalized to $B_s^0 \rightarrow J/\psi \phi$ to calculate a branching fraction or limit. The discrimination cuts have been optimized using a random grid search to enhance the signal sensitivity. The expected background for this optimization was 1.6 ± 0.4 events in the signal region. After unblinding the signal window, zero events have been observed.

Therefore, a frequentist upper limit on this branching fraction normalized to $B_s^0 \rightarrow J/\psi \phi$ is $\frac{\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^- \phi)}{\mathcal{B}(B_s^0 \rightarrow J/\psi \phi)} < 4.4 \times 10^{-3}$ at a 95% C.L. including statistical and systematic uncertainties on the background and signal efficiencies. Using the central value of the present world average value for $\mathcal{B}(B_s^0 \rightarrow J/\psi \phi) = 9.3 \pm 3.3 \times 10^{-4}$ without its uncertainty, the absolute limit is then $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^- \phi) < 4.1 \times 10^{-6}$ at a 95% C.L.

Preliminary Results for Summer 2005 Conferences

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

I. INTRODUCTION

The investigation of rare flavor changing neutral current (FCNC) B decay processes have received special attention in the past since they open up the possibility of precision tests of the flavor structure of the Standard Model (SM). In the SM FCNC decays are absent at tree level, but proceed at higher order through electroweak penguin and box diagrams. FCNC decays are sensitive to new physics, since decay amplitudes involving new particles interfere with SM amplitudes. Although inclusive FCNC decays like $B \rightarrow X_s \ell^+ \ell^-$ or $B \rightarrow X_s \gamma$ are theoretically cleaner to calculate, exclusive decays with one hadron in the final state are experimentally easier to access.

The decay $B_s^0 \rightarrow \mu^+ \mu^- \phi$ is an exclusive FCNC decay and is related to the transition of $b \rightarrow s \ell^+ \ell^-$ at quark level. Within the SM the decay rate $B_s^0 \rightarrow \mu^+ \mu^- \phi$ decay is predicted to be of the order of 1.6×10^{-6} [1] excluding long-distance effects from charmonium resonances. Including the long-distance contribution of charmonium states enhances the branching fraction of $B_s^0 \rightarrow \mu^+ \mu^- \phi$ by almost a factor of three, depending on the exact modeling of the charmonium states [2]. Due to poorly known form factors, SM calculations for exclusive decay rates may suffer from uncertainties of up to 30%.

In the two-Higgs doublet model, the branching fraction of the exclusive decay $B_s^0 \rightarrow \mu^+ \mu^- \phi$ might be enhanced [3], depending on the parameter values of $\tan \beta$ and the mass of the charged Higgs. However, in contrast to the purely leptonic decay $B_s^0 \rightarrow \mu^+ \mu^-$, enhancements due to new physics are expected to be less significant. Presently, the only existing experimental bound on $B_s^0 \rightarrow \mu^+ \mu^- \phi$ is given by CDF from a Run I search [4]. CDF set an upper limit at a 95% C.L. of $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^- \phi) < 6.7 \times 10^{-5}$.

In this note we present a search for the decay $B_s^0 \rightarrow \mu^+ \mu^- \phi$ decay. The ϕ mesons are reconstructed through their $K^+ K^-$ decay. Since we are interested in the non-resonant decay, which is mediated through a FCNC diagram, the invariant mass of the two muons is requested to be outside the charmonium resonances. The events in our search are normalized to resonant decay $B_s^0 \rightarrow J/\psi(\mu^+ \mu^-) \phi$ events. The event final state contains two muons and two kaon candidate tracks that form a ϕ candidate and is the same for signal channel and (resonant) normalization channel.

II. THE DØ DETECTOR

The search uses a data set of approximately 300 pb^{-1} of $p\bar{p}$ collisions at $\sqrt{s} = 1.96 \text{ TeV}$ 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 the 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. PRE-SELECTION

The pre-selection starts with a loose selection of B_s^0 candidates consisting of two muons and two oppositely charged tracks forming a good vertex. In this first selection step, the mass of the two kaon candidate tracks should be between $0.980 < m_\phi < 1.080 \text{ GeV}/c^2$ and the allowed mass of the loose B_s^0 candidate is required to be within $4.4 < m_B < 6.2 \text{ GeV}/c^2$.

We then require the invariant mass of the two muons between $0.5 < m_{\mu^+ \mu^-} < 4.4 \text{ GeV}/c^2$. In this mass region we exclude the $J/\psi(\rightarrow \mu^+ \mu^-)$ and $\psi'(\rightarrow \mu^+ \mu^-)$ resonances with cut-out regions that cover $\pm 5\sigma$ wide windows around the observed resonance masses. The J/ψ mass resolution in data is given by $\sigma = 75 \text{ MeV}$. We note in passing that the MC resolution is about 20% better than in data.

The two muons have to be of medium quality [6], requiring two of the three muon layers to be fired and a central track matched. The $\chi^2/d.o.f.$ of the two muon vertex is required to be $\chi^2/d.o.f. < 10$. The tracks that are matched to each muon leg need at least three hits in the SMT and four hits in the CFT. In addition, the transverse momentum of each of the muons is required to be greater than $2.5 \text{ GeV}/c$, and their pseudorapidity η has to be less than 2.0 in order to be well inside the fiducial tracking and muon region.

For surviving events, the two-dimensional decay length in the plane transverse to the beamline, L_{xy} , is calculated.

It is defined as the projection of the decay length vector \vec{l}_{Vtx} on the transverse momentum of the B_s^0 candidate:

$$L_{xy} = \frac{\vec{l}_{Vtx} \cdot \vec{p}_T^B}{p_T^B}. \quad (1)$$

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 required that $\delta L_{xy} < 150 \mu\text{m}$.

In the following, the number of B_s^0 candidates is further restricted by requiring $p_T^B > 5 \text{ GeV}/c$ and asking the B_s^0 candidate vertex to fulfill $\chi^2 < 36$ with 5 d.o.f. The two tracks that are combined with the two muons to the B_s^0 candidate should each have $p_T > 0.7 \text{ GeV}/c$. The two tracks forming the ϕ candidate are further restricted in their invariant mass to be within $1.008 < m_\phi < 1.032 \text{ GeV}/c^2$. The successive cuts and the remaining candidates surviving each cut are shown in Table I.

TABLE I: Number of candidate events surviving the cuts in data used in the pre-selection analysis.

Cut	Value	#candidates
Good vertex		1555320
Mass region (GeV/c^2)	$0.5 < m_{\mu^+\mu^-} < 4.4$ excl. $J/\psi, \psi(2S)$	530892
Muon Quality	two medium	276875
$\chi^2/d.o.f.$ of vertex	< 10	127509
Muon p_T (GeV/c)	> 2.5	73555
Muon $ \eta $	< 2.0	72350
Tracking hits	CFT > 3 , SMT > 2	58012
δL_{xy} (mm)	< 0.15	54752
B_s^0 Candidate p_T (GeV/c)	> 5.0	54399
B_s^0 χ^2 vertex	< 36	53195
Kaon p_T (GeV/c)	> 0.7	9639
ϕ mass (GeV/c^2)	$1.008 < m_\phi < 1.032$	2602

We apply the same selection for the resonant $B_s^0 \rightarrow J/\psi \phi$ candidates except that the invariant mass of the muon pair is now required to be within $\pm 250 \text{ MeV}/c^2$ of the J/ψ mass.

IV. DISCRIMINATING VARIABLES AND CUT OPTIMIZATION

We have generated 170k signal MC events for the decay $B_s^0 \rightarrow \mu^+\mu^-\phi$. The used decay model in EvtGen [7] includes the NNLO improved Wilson coefficients [8] for the short distance part with form factors obtained from QCD light cone sum rules taken from Reference [9]. These form factors are originally determined for $B \rightarrow K^*$ transitions and were confronted with experimental measurements on the branching fraction $\mathcal{B} \rightarrow K^*\ell^+\ell^-$ in Ref. [8]. Recently, new form factors on $B_s \rightarrow \phi$ obtained from light cone QCD sum rules were published [10] which differ in size but not in shape by 10-15% compared to the ones used in our MC.

The invariant mass spectra for the two muons from the signal MC is shown in Fig. 1.

For the final event selection we have used the same three discriminating variables that were already employed in the search for $B_s^0 \rightarrow \mu^+\mu^-$ [11]. The **isolation** variable I of the phi and muon pair is defined as:

$$I = \frac{|\vec{p}(\mu\mu\phi)|}{|\vec{p}(\mu\mu\phi)| + \sum_{\text{track } i \neq B} p_i(\Delta\mathcal{R} < 1)}.$$

$\sum_{\text{track } i \neq B} p_i$ 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 $\vec{p}(\mu^+\mu^-\phi)$ 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 data after pre-selection is shown

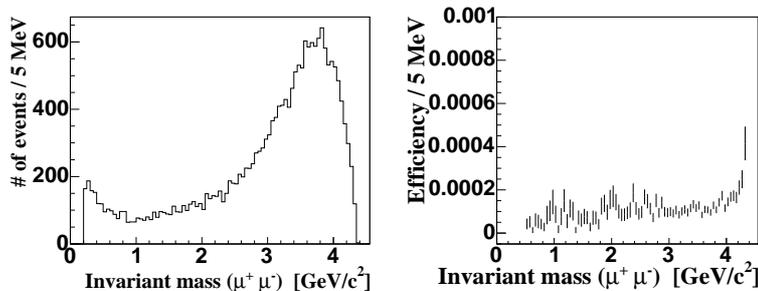


FIG. 1: Invariant mass spectra generated with the decay model [8] before selection cuts (left) and the efficiency as function of mass after all selection cuts (right).

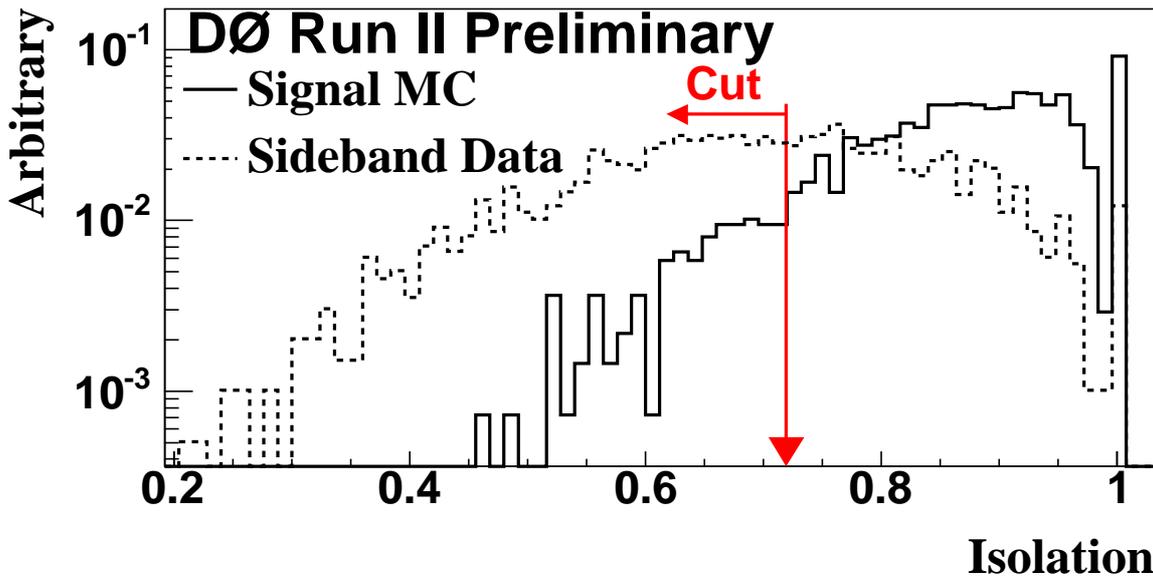


FIG. 2: Isolation variable after the pre-selection for data events from the sidebands and signal MC events.

in Fig. 2.

The **pointing angle** α is defined as the angle between the momentum vector $\vec{p}(\mu^+\mu^-\phi)$ of the B_s^0 candidate and the vector \vec{l}_{Vtx} pointing from the primary vertex to the secondary vertex. If the two muons and the ϕ are coming from the decay of a parent particle B_s^0 , the vector \vec{l}_{Vtx} should point into the same direction as $\vec{p}(\mu^+\mu^-\phi)$. 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. Figure 3 shows the distributions of the angle α for signal MC and (sideband) data after pre-selection.

To discriminate against short-lived background, we have used the **transverse decay length significance** $L_{xy}/\delta L_{xy}$ since it gives a better discriminating power than the transverse decay length alone. The error δL_{xy} is calculated by error propagation of the uncertainties on both the primary and secondary vertex positions. Figure 4 shows the distributions of the decay length significance for signal MC and data.

Before optimizing the cuts on these discriminating variables, we restrict ourselves to a mass region of interest of $4.51 < M_{\mu^+\mu^-\phi} < 6.13$ GeV/c^2 containing the signal region around the PDG world average value of the B_s^0 mass of $m_{B_s^0} = 5369.6 \pm 2.4$ MeV/c^2 [12]. The whole mass region of interest is shifted downward with respect to the world average B_s^0 mass by 44 MeV/c^2 in order to correct for the mass scale of the DØ tracker. The 44 MeV/c^2 mass shift was taken from the mean B_s^0 mass obtained from the fit to the $B_s^0 \rightarrow J/\psi\phi$ mass spectra without constraining the $\mu\mu$ -pair to the J/ψ mass.

The signal box is blinded during the whole analysis and is chosen to be sufficiently far away from the sidebands.

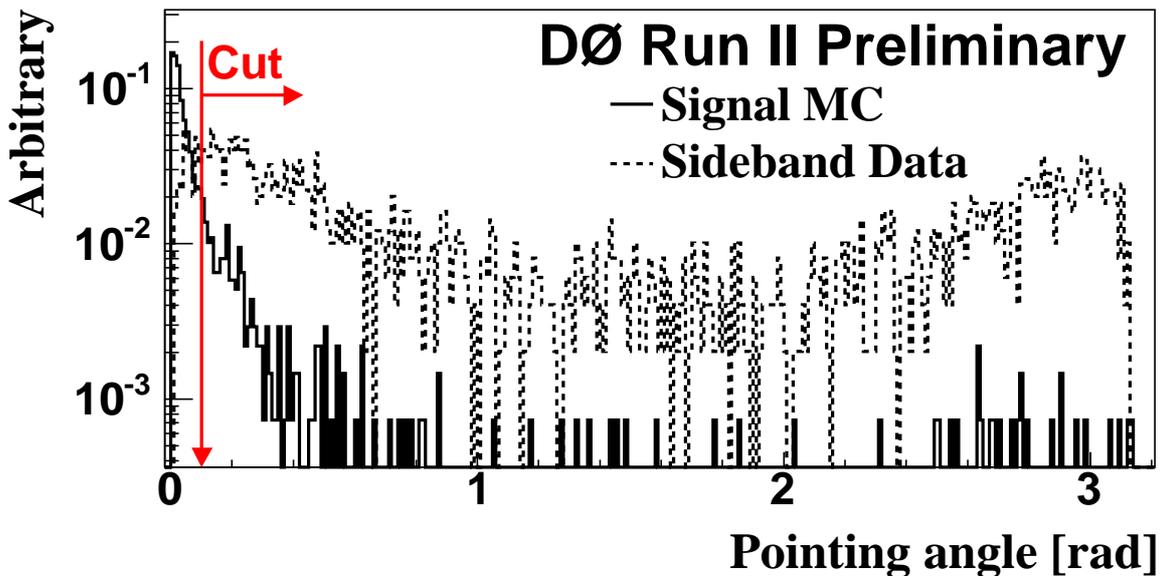


FIG. 3: The pointing angle α after the pre-selection for data events from the sidebands and signal MC events.

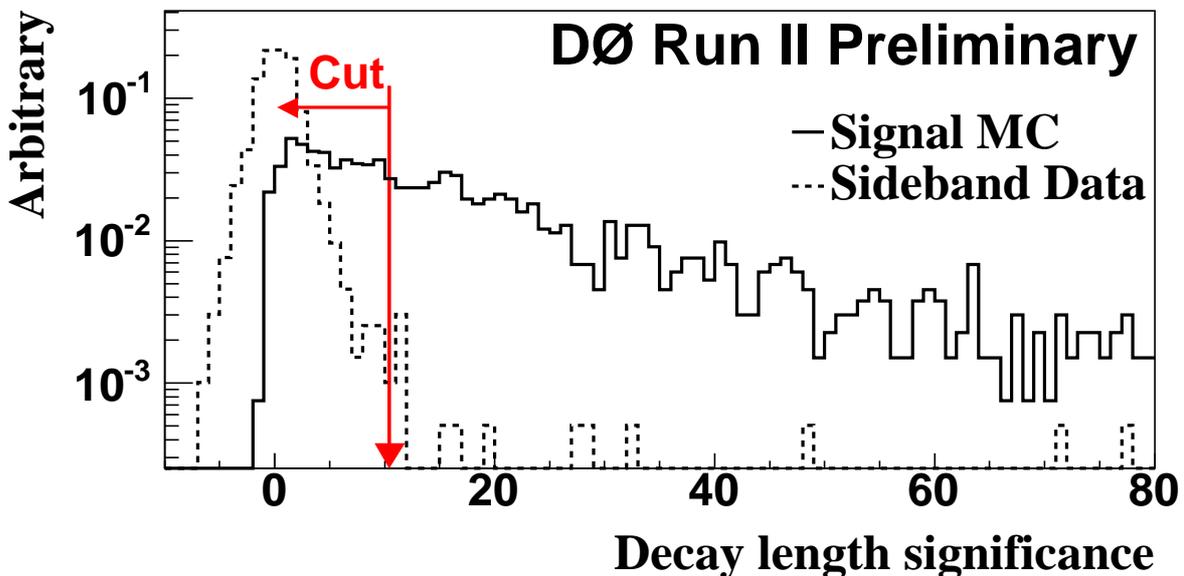


FIG. 4: The transverse decay length significance for data events from the sidebands and signal MC events.

Table II defines the regions for the sidebands and the blinded signal box that have been used. The signal region corresponds to a window of $\pm 270 \text{ MeV}/c^2$ around the (shifted) world average mass value of the B_s^0 . The expected mass resolution for $B_s^0 \rightarrow \mu^+ \mu^- \phi$ in the MC is $75 \text{ MeV}/c^2$. The $\pm 270 \text{ MeV}/c^2$ region corresponds therefore to $\pm 3.6\sigma$. After the cut optimization we shrink the blinded signal box to $\pm 2.5\sigma$ for calculating the limit.

To find the optimal set of cuts we use a Random Grid Search (RGS) [13] and an optimization criterion proposed by G. Punzi [14]. We maximize the ratio P defined as:

$$P = \frac{\epsilon_{\mu\mu\phi}}{\frac{a}{2} + \sqrt{N_{Back}}}. \quad (2)$$

$\epsilon_{\mu\mu\phi}$ 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

TABLE II: The invariant mass regions for signal and sidebands used for background estimation.

Region	min Mass (GeV/c ²)	max Mass (GeV/c ²)	width (MeV/c ²)
region of interest	4.5156	6.1356	1620
blinded signal region during optimization	5.0556	5.5956	540
final blinded signal region for limit	5.1381	5.5131	375
sideband I	4.5156	5.0556	540
sideband II	5.5956	6.1356	540

confidence level at which the signal hypothesis is tested. This number a should be defined before the statistical test and has been set to 2, corresponding to about 95% C.L. The optimization has been performed on the complete set of signal MC including the charmonium resonances to increase the number of cut combinations. This can be justified, since the discriminating variables do not depend on the invariant dimuon mass. The resulting cut values that were obtained from the maximized P are listed in table III

TABLE III: The optimized cuts and their relative MC signal efficiencies including their statistical uncertainties after maximizing P .

cut parameter	value (re-optimized)	MC eff. (%)
Opening angle (rad)	< 0.1	81±4
Decay length significance	> 10.3	73±4
Isolation	> 0.72	92±5

The total signal efficiency relative to pre-selection of the three discriminating cuts is $(54 \pm 3)\%$ with the uncertainty due to MC statistics. After a linear extrapolation of the sideband population for the whole data sample into the final signal region, we obtain an expected number of background events of 1.6 ± 0.4 . Figure 5 shows the remaining background events populating the lower and upper sidebands.

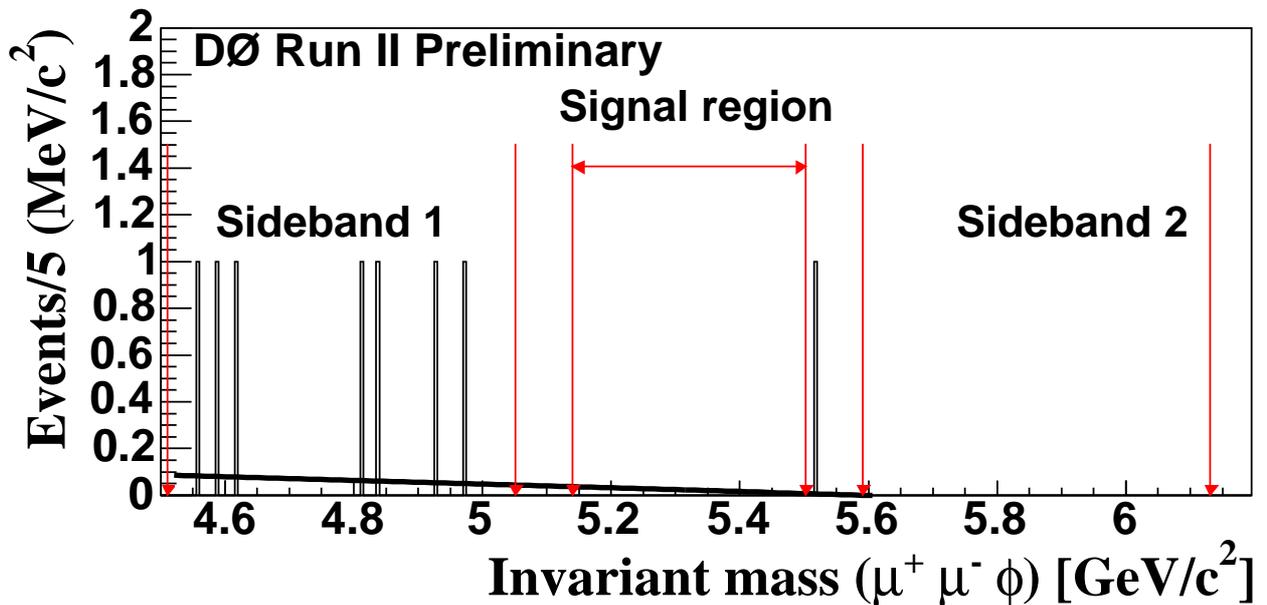


FIG. 5: The invariant mass distribution for the full data sample with our standard discriminating variables.

V. THE NORMALIZATION CHANNEL $B_s^0 \rightarrow J/\psi \phi$

In order to obtain a branching fraction limit for $B_s^0 \rightarrow \mu^+ \mu^- \phi$, we have used $B_s^0 \rightarrow J/\psi \phi$ events with $J/\psi \rightarrow \mu^+ \mu^-$ and $\phi \rightarrow K^+ K^-$ as normalization. As mentioned above, the same cuts were applied to the $B_s^0 \rightarrow J/\psi \phi$ candidates.

The contamination of muon pairs from the non-resonant $\mu^+\mu^-\phi$ decay in the resonant normalization region $J/\psi(\rightarrow \mu^+\mu^-)\phi$ is negligible. We therefore constrain the two muons to have an invariant mass equal to the J/ψ mass [12].

The mass spectrum of the reconstructed $B_s^0 \rightarrow J/\psi \phi$ for the full data sample is shown in Figure 6. A fit using a Gaussian function for the signal and a second order polynomial for the background yields 73 ± 10 B_s^0 events, where only the statistical uncertainty is given. When the $\mu^+\mu^-$ mass is constrained to the J/ψ mass, the mass resolution of the B_s^0 is found to be 27.3 MeV which compares very well to the resolution found in the MC simulation of 26.4 MeV.

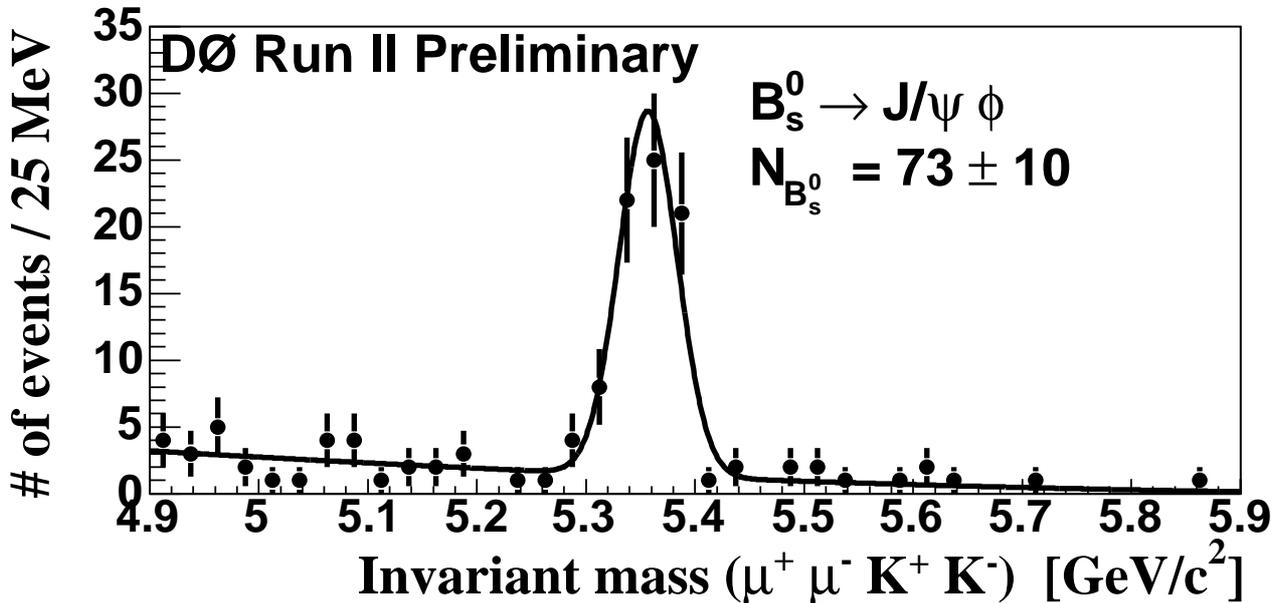


FIG. 6: The normalization channel $B_s^0 \rightarrow J/\psi \phi$ for the full data sample.

VI. CALCULATION OF THE LIMIT

In the absence of a signal in our search region we can calculate an upper limit on the ratio ($B_s^0 \rightarrow \mu^+\mu^-\phi/B_s^0 \rightarrow J/\psi \phi$) using

$$\frac{\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-\phi)}{\mathcal{B}(B_s^0 \rightarrow J/\psi \phi)} = \frac{\mu(n_{obs}, n_{back})}{N_{B_s^0}} \cdot \frac{\epsilon_{J/\psi \phi}}{\epsilon_{\mu\mu\phi}} \cdot \mathcal{B}(J/\psi \rightarrow \mu\mu), \quad (3)$$

where

- $\epsilon_{\mu\mu\phi}$ and $\epsilon_{J/\psi \phi}$ are the efficiencies of the signal and normalization channels, obtained from MC simulations (including trigger simulations), and
- $\mathcal{B}(J/\psi \rightarrow \mu\mu) = (5.88 \pm 0.1)\%$ [12].

The efficiencies $\epsilon_{\mu\mu\phi}$ and $\epsilon_{J/\psi \phi}$ are the global signal efficiencies for the search signal and normalization channel respectively including the pre-selection cuts and the acceptance. They are determined from MC yielding an efficiency ratio of $(\epsilon_{J/\psi \phi}/\epsilon_{J/\psi \mu\mu}) = 2.80 \pm 0.21$, where the uncertainties are due to MC statistics. Although the non-resonant decay $B_s^0 \rightarrow \mu^+\mu^-\phi$ has similar topology than the resonant normalization mode, its efficiency is smaller due to the rejection of dimuon events around the charmonium resonances.

In order to avoid large uncertainties associated with the poorly known branching fraction of $B_s^0 \rightarrow J/\psi \phi$, we normalize the limit on the non-resonant decay to $\mathcal{B}(B_s^0 \rightarrow J/\psi \phi)$ as it is given in Eq. 3. For the calculation of $\frac{\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-\phi)}{\mathcal{B}(B_s^0 \rightarrow J/\psi \phi)}$ we have used the fixed cuts of Table III that were obtained with our optimization method performed on the data sample. Figure 5 shows the invariant mass distribution after opening the box with no observed event remaining in the signal region. The Poisson probability of observing zero events for an expected background of 1.6 ± 0.4 is $p = 0.22$.

VII. UNCERTAINTIES

The different sources of relative uncertainties that go into the limit calculation of \mathcal{B} are given in Table IV. Basically, the branching fraction of $B_s^0 \rightarrow J/\psi \phi$ has the largest uncertainty, but its uncertainty cancels due to our normalization. The second largest uncertainty of 25% is due to the background interpolation into the signal region and is based on the statistical uncertainty of the fit integral. The statistical uncertainty on the number of observed $B_s^0 \rightarrow J/\psi \phi$ events as normalization channel is 14.0%. An additional uncertainty due to MC weighting is applied. Our MC for the normalization channel contained only CP-even states. The efficiency difference between the CP-even and CP-odd state for the normalization channel was estimated using MC to be 6% with an uncertainty on the ratio of 8%. Therefore another uncertainty concerning the CP odd-even eigenstates of the $B_s^0 \rightarrow J/\psi \phi$ events of 8% was assigned. The efficiency ratio $\epsilon_{J/\psi \phi}/\epsilon_{\mu\mu\phi}$ is found in this analysis as a single number from the MC and hence the correlations are taken into account. The statistical uncertainty on the ratio is 7.5%. The signal efficiency obtained from MC is based on the input for the NNLO Wilson coefficients and form factors of Ref. [8]. We did not include any theoretical uncertainty into our systematics.

TABLE IV: The relative uncertainties for calculating an upper limit of \mathcal{B} . Note that the first uncertainty is not taken into account since we normalize the limit to $\mathcal{B}(B_s^0 \rightarrow J/\psi \phi)$.

Source	Relative Uncertainty [%]
$\mathcal{B}(B_s^0 \rightarrow J/\psi \phi)$	35.5
$\mathcal{B}(J/\psi \rightarrow \mu\mu)$	1.7
$\epsilon_{J/\psi \phi}/\epsilon_{\mu\mu\phi}$	7.5
# of $B_s^0 \rightarrow J/\psi \phi$	13.7
MC weighting	1.1
pre-Geant weighting	3.5
CP odd-even lifetime differences	8.0
Total	18.0
background uncertainty	25.0

In case of significantly less observed events than expected, the unified frequentist approach [15] (aka Feldman-Cousins) and the Bayesian approach for calculating upper limits differ substantially. In particular, for zero observed events a Bayesian upper limit is independent from the number of expected background events, while in the unified frequentist approach the upper limit decreases when more background is expected. The question on which behavior is correct depends on how to link confidence intervals to probabilities of either the probability of a hypothesis given data (Bayes) or to the probability of data given a hypothesis (frequentist). Thus we quote our limit in both the unified and the Bayesian approaches.

The statistical and systematic uncertainties can be included in the limit calculation by integrating over probability functions that parameterize the uncertainties. We have used a prescription [16] where we construct a frequentist confidence interval with the Feldman and Cousins [15] ordering scheme for the MC integration. The background was modeled as a Gaussian distribution with its mean value equal to the expected number of background events and its standard deviation equal to the background uncertainty. Including the statistical and systematic uncertainties, the limit is $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^- \phi)/\mathcal{B}(B_s^0 \rightarrow J/\psi \phi) < 4.4 (3.5) \times 10^{-3}$ at 95% (90%) C.L. respectively. Using only the central value of the world average branching fraction [12] of $\mathcal{B}(B_s^0 \rightarrow J/\psi \phi) = 9.3 \pm 3.3 \cdot 10^{-4}$, this limit corresponds to $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^- \phi) = 4.1 (3.2) \cdot 10^{-6}$ at 95% (90%) C.L. respectively.

Taking a Bayesian approach [17] with flat prior and the uncertainties treated as gaussians in the integration, we find an upper limit of $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^- \phi)/\mathcal{B}(B_s^0 \rightarrow J/\psi \phi) < 7.4 (5.6) \times 10^{-3}$ at 95% (90%) C.L. respectively.

Since we have less events observed than expected, it is good statistical practice to also quote the sensitivity of our search, i.e., the ensemble average of all expected upper limits in the absence of a signal for a hypothetical repetition of the experiment. Assuming there is only background we calculate for each possible value of observation a 95% C.L. upper limit weighted by the Poisson probability of occurrence. Including the statistical and systematic uncertainties our sensitivity is then given by $\langle \mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^- \phi)/\mathcal{B}(B_s^0 \rightarrow J/\psi \phi) \rangle = 1.1 (1.2) \times 10^{-2}$ at 95% C.L. using the Feldman and Cousins (Bayesian) approach respectively.

VIII. CONCLUSIONS

We have presented a search for the rare decay $B_s^0 \rightarrow \mu^+ \mu^- \phi$ based on 300 pb^{-1} of data recorded with the $D\bar{O}$ detector. The search is carried out as a blind analysis hiding the signal region around the B_s^0 and normalizing the number of events to the number of reconstructed $B_s^0 \rightarrow J/\psi \phi$ candidates. The signal efficiency is determined using a MC simulation with improved NNLO Wilson coefficients for the short-distance part and form factors based on $B \rightarrow K^*$ transitions.

For the data set the expected background extrapolated from the sidebands is estimated to be 1.6 ± 0.4 events. Since no event after unblinding was found in the signal region, we calculate an upper limit at a 95% (90%) C.L. and obtain $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^- \phi) / \mathcal{B}(B_s^0 \rightarrow J/\psi \phi) < 4.4(3.5) \times 10^{-3}$ using the method proposed by Conrad et al. based on a Feldman and Cousins ordering scheme. Using the world average value for $\mathcal{B}(B_s^0 \rightarrow J/\psi \phi)$ this corresponds to an upper limit of $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^- \phi) = 4.1(3.2) \times 10^{-6}$ at 95% (90%) C.L. respectively.

We thank the staffs at Fermilab and collaborating institutions, and acknowledge support from the DOE and NSF (USA); CEA and CNRS/IN2P3 (France); FASI, Rosatom and RFBR (Russia); CAPES, CNPq, FAPERJ, FAPESP and FUNDUNESP (Brazil); DAE and DST (India); Colciencias (Colombia); CONACyT (Mexico); KRF (Korea); CONICET and UBACyT (Argentina); FOM (The Netherlands); PPARC (United Kingdom); MSMT (Czech Republic); CRC Program, CFI, NSERC and WestGrid Project (Canada); BMBF and DFG (Germany); SFI (Ireland); Research Corporation, Alexander von Humboldt Foundation, and the Marie Curie Program.

-
- [1] C. Q. Geng and C. C. Liu, *J. Phys. G* **29**, 1103 (2003), hep-ph/0303246.
 - [2] A. Deandrea and A. Polosa, *Phys. Rev. D* **64**, 074012 (2001), hep-ph/0105058.
 - [3] G. Erkol and G. Turan, *Eur. Phys. J. C* **25**, 575 (2002), hep-ph/0203038.
 - [4] T. Affolder *et al.* [CDF Collaboration], *Phys. Rev. D* **65**, 1111101.
 - [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 $D\bar{O}$ 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] <http://hep.ucsb.edu/people/lange/EvtGen>.
 - [8] A. Ali *et al.* *Phys. Rev. D* **66**, 034002 (2002), hep-ph/0112300.
 - [9] A. Ali *et al.* *Phys. Rev. D* **61**, 074024 (2000), hep-ph/9911273.
 - [10] P. Ball and R. Zwicky, *Phys. Rev. D* **71**, 014029 (2005).
 - [11] V. M. Abzov *et al.*, *Phys. Rev. Lett.* **94**, 071802 (2005), hep-ex/0410039.
 - [12] S. Eidelman *et al.* *Phys. Lett B* **592**, 1 (2004).
 - [13] H. B. Prosper, *Proceedings of the International Conference on Computing in High Energy Physics (CHEP'95)*, Rio de Janeiro, 18-22 September 1995.
 - [14] 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.
 - [15] G. J. Feldman and R. D. Cousins, *Phys. Rev. D* **57**, 3873 (1998).
 - [16] J. Conrad *et al.*, *Phys. Rev. D* **67**, 012002 (2003).
 - [17] I. Bertram *et al.*, A Recipe for the Construction of Confidence Limits, *D\bar{O} Note* 3476.