Search for the Rare Decay $B^0_s \to \phi \mu^+\mu^-$ with the DØ Detector

We present a search for the flavor-changing neutral current decay $B_S^0 \rightarrow \phi \mu^+\mu^-$ using about 0.45 $fb^{-1}$ of data collected in $pp$ collisions at $\sqrt{s}=1.96$ TeV with the DØ detector at the Fermilab Tevatron Collider. We find an upper limit on the branching ratio of this decay normalized to
$B_s^0 \rightarrow J/\psi \phi$ of $\frac{B_s(B_s^0 \rightarrow J/\psi \phi)}{B_s(B_s^0 \rightarrow J/\psi \phi)} < 4.4 \times 10^{-3}$ at the 95% C.L. Using the central value of the world average branching fraction of $B_s^0 \rightarrow J/\psi \phi$, the limit corresponds to $B(B_s^0 \rightarrow \phi \mu^+ \mu^-) < 4.1 \times 10^{-6}$ at the 95% C.L., the most stringent upper bound to date.

PACS numbers: 13.20.He, 12.15.Mm, 14.40.Nd

The investigation of rare flavor changing neutral current (FCNC) $B$ meson decays has received special attention in the past since this opens 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 easier to calculate, exclusive decays with one hadron in the final state are experimentally easier to study. For instance, the exclusive decays $B_s^0 \rightarrow K^+ \ell^+ \ell^-$ and $B^\pm \rightarrow K^\pm \ell^+ \ell^-$ have already been measured at $B$-factories [1, 2] and were found to be consistent with the SM within the present experimental uncertainties. Related to the same quark-level transition of $b \rightarrow s \ell^+ \ell^-$ is the corresponding exclusive FCNC decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$ in the $B_s^0$ meson system. An observation of this decay or experimental upper limit on its rate will yield additional important information on the flavor dynamics of FCNC decays.

Within the SM, the decay rate for the $B_s^0 \rightarrow \phi \mu^+ \mu^-$ decay, neglecting the interference effects with the much stronger $B_s^0 \rightarrow J/\psi \phi$ and $B_s^0 \rightarrow \psi(2S) \phi$ resonance decays, is predicted to be of the order of $1.6 \times 10^{-6}$ [3] with about 30% uncertainty due to poorly known form factors. The interference effects with the $B_s^0$ resonance decay amplitudes are large, with their expected magnitude depending on the exact modeling of the charmonium states [4]. To separate experimentally the FCNC-mediated process $B_s^0 \rightarrow \phi \mu^+ \mu^-$, one has to restrict the invariant mass of the final state lepton pair to be outside the charmonium resonances. Presently, the only existing experimental upper bound on the $B_s^0 \rightarrow \phi \mu^+ \mu^-$ decay is given by CDF from the analysis of Run I data [5]. CDF sets an upper limit at the 95% C.L. of $B(B_s^0 \rightarrow \phi \mu^+ \mu^-) < 6.7 \times 10^{-5}$.

In this Letter, we report on a new experimental limit on the decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$, that is an order of magnitude more stringent than the existing limit. The $\phi$ mesons are reconstructed through their $K^+K^-$ decay mode and the invariant mass of the two muons in the final state is required to be outside the charmonium resonances. The events in our search are normalized to resonant decay $B_s^0 \rightarrow J/\psi \phi$ events. Using the $B_s^0 \rightarrow J/\psi \phi$ mode as the normalization channel has the advantage that the efficiencies to detect the $\phi \mu^+ \mu^-$ system in signal and normalization events are similar, and systematic effects tend to cancel.

The search uses a data set corresponding to approximately 0.45 fb$^{-1}$ of $pp$ collisions at $\sqrt{s} = 1.96$ TeV recorded by the DØ detector operating at the Fermilab Tevatron Collider. The DØ detector is described in detail elsewhere [6]. 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, which is located outside the calorimeter, consists of a layer of tracking detectors and scintillation trigger counters in front of 1.8 T toroidal magnets, followed by two more similar layers after the toroids, allowing for efficient muon detection out to pseudorapidity ($|\eta|$) of ±2.0.

Dimuon triggers were used in the data selection for this analysis. A trigger simulation was used to estimate the trigger efficiency for the signal and normalization samples. These efficiencies were also checked with data samples collected with single muon triggers. The event preselection starts with a loose selection of $B_s^0 \rightarrow \phi \mu^+ \mu^-$ candidates. These candidates are identified by requiring exactly two muons fulfilling quality cuts on the number of hits in the muon system and the two additional charged particle tracks to form a good vertex. The reconstructed invariant mass of the $B_s^0$ candidate should be within $4.4 < m_{\phi \mu^+ \mu^-} < 6.2$ GeV/c$^2$.

We then require the invariant mass of the two muons to be within $0.5 < m_{\mu^+ \mu^-} < 4.4$ GeV/c$^2$. In this mass region, the $J/\psi(\rightarrow \mu^+ \mu^-)$ and $\psi(2S)(\rightarrow \mu^+ \mu^-)$ resonances are excluded to discriminate against dominant resonant decays by rejecting the mass region $2.72 < m_{\mu^+ \mu^-} < 4.06$ GeV/c$^2$. The $J/\psi$ mass resolution in data is given by a Gaussian distribution with $\sigma = 75$ MeV/c$^2$. The rejected mass region then covers ±5$\sigma$ wide windows around the resonance masses.

The $\chi^2/d.o.f.$ of the two-muon vertex is required to be less than 10. The tracks that are matched to each muon are required to have at least three (four) measurements in the SMT (CFT) and the transverse momentum of each of the muons ($p_T$) is required to be greater than 2.5 GeV/c with $|\eta| < 2.0$ to be well inside the fiducial tracking and muon detector acceptances. In order to select well-measured secondary vertices, we define the two-dimensional decay length $L_{xy}$ in the plane transverse to the beamline, and require its uncertainty $\delta L_{xy}$ to be less than 0.15 mm. $L_{xy}$ is calculated as $L_{xy} = \frac{L_{xy}}{p_T^2}$, where $p_T^2$ is the transverse momentum of the candidate.
$B^0_s$ and $\vec{l}_{ctz}$ represents the vector pointing from the primary vertex to the secondary vertex. The uncertainty on the transverse decay length, $\delta L_{xy}$, is calculated by taking into account the uncertainties in both the primary and secondary vertex positions. The primary vertex itself is found for each event using a beam-spot constrained fit as described in Ref. [7].

Next, the number of $B^0_s \rightarrow \phi \mu^+\mu^-$ candidates is further reduced by requiring $p_T^B > 5$ GeV/c and asking the $B^0_s$ candidate vertex to have $\chi^2 < 36$ with 5 d.o.f. The two tracks forming the $\phi$ candidate are further required to have $p_T > 0.7$ GeV/c and their invariant mass within the range $1.008 < m_{\phi} < 1.032$ GeV/c$^2$. The successive cuts and the remaining candidates surviving each cut are shown in Table I. We apply the same selection for the resonant $B^0_s \rightarrow J/\psi \phi$ candidates except that the invariant mass of the muon pair is now required to be within $\pm 250$ MeV/c$^2$ of the $J/\psi$ mass.

For the final event selection, we require the candidate vertex to have $x > 36$ with 5 d.o.f. The two tracks forming the $\phi$ candidate are further required to have $p_T > 0.7$ GeV/c and their invariant mass within the range $1.008 < m_{\phi} < 1.032$ GeV/c$^2$. The successive cuts and the remaining candidates surviving each cut are shown in Table I. We apply the same selection for the resonant $B^0_s \rightarrow J/\psi \phi$ candidates except that the invariant mass of the muon pair is now required to be within $\pm 250$ MeV/c$^2$ of the $J/\psi$ mass.

For the final event selection, we require the invariant mass of the muon pair to be within $\pm 250$ MeV/c$^2$ of the $J/\psi$ mass. The analysis is carried out based on signal MC events in the $B^0_s$ mass region and on data events in regions outside the experimental signal window defined as $4.51 < m_{\phi} < 6.13$ GeV/c$^2$. A 44 MeV/c$^2$ mass shift in the mass region of interest is introduced to calibrate the D0 tracker.

In order to avoid biasing the analysis procedure, data candidates in the signal mass region are not examined until completion of the analysis, and events in the sideband regions around the $B^0_s$ mass are used instead. The expected mass resolution for $B^0_s \rightarrow \phi \mu^+\mu^-$ in the MC is 75 MeV/c$^2$. The start (end) of the upper (lower) sideband was chosen such that it is at least 270 MeV/c$^2$ away from the $B^0_s$ mass. The widths of the sidebands used for background estimation are chosen to be 540 MeV/c$^2$ each. The size of the blind signal region is $\pm 225$ MeV/c$^2$ which corresponds to a $\pm 3\sigma$ region around the $B^0_s$ mass. To determine the final limit on the branching fraction, we use a smaller mass region of $\pm 2.5\sigma$.

A random-grid search [11] was used to find simultaneously the optimal values of the discriminating variables: $L_{xy}/\delta L_{xy}$ as one of the discriminating variables, since it gives better discriminating power than the transverse decay length alone.

The fragmentation characteristics of the $b$ quark are such that most of its momentum is carried by the $B$ hadron. Thus the number of extra tracks near the $B^0_s$ candidate tends to be small. Therefore the second discriminant is an isolation variable, $I$, of the muon and kaon pairs, defined as:

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

Here, $\sum_{\text{track} \neq B} p_i$ is the scalar sum over all tracks excluding the muon and kaon pairs within a cone of $\Delta R < 1$ around the momentum vector $\vec{p}(\phi \mu^+\mu^-)$ of the $B^0_s$ candidate where $\Delta R = \sqrt{(\Delta \phi)^2 + (\Delta \eta)^2}$. The final discriminating variable used is the pointing angle $\alpha$, defined as the angle between the momentum vector $\vec{p}(\phi \mu^+\mu^-)$ of the $B^0_s$ candidate and the vector $\vec{l}_{ctz}$ between the primary and secondary vertices. This requirement ensures consistency between the direction of the decay vertex and the momentum vector of the $B^0_s$ candidate.

We generate signal Monte Carlo (MC) events for the decay $B^0_s \rightarrow \phi \mu^+\mu^-$ using a decay model which includes the NNLO improved Wilson coefficients [8] for the short-distance part. The form factors obtained from QCD light-cone sum rules are taken from Ref. [9]. These form factors were originally determined for $B \rightarrow K^\ast$ transitions and were compared with experimental measurements of the branching fraction $B(B^0_s \rightarrow K^\ast f \bar{f}^-)$ in Ref. [8]. Recently, new form factors for the $B^0_s \rightarrow \phi$ transition, obtained from the light cone QCD sum rules, were published [10]. The difference between the form factors in Ref. [8] and those in Ref. [10] reaches 20% for $m_{\mu^+\mu^-} < 1$ GeV/c$^2$, while elsewhere it remains well below 10%.

The analysis is carried out based on signal MC events in the $B^0_s$ mass region and on data events in regions outside the experimental signal window defined as $4.51 < m_{\phi} < 6.13$ GeV/c$^2$. A 44 MeV/c$^2$ mass shift in the mass region of interest is introduced to calibrate the D0 tracker.

The total signal efficiency relative to pre-selection of the three discriminating cuts is $(54 \pm 3)\%$ where the uncertainty is statistical only. After a linear interpolation of

<table>
<thead>
<tr>
<th>Cut</th>
<th>Value</th>
<th># candidates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good $B^0_s$ vertex</td>
<td>1555320</td>
<td></td>
</tr>
<tr>
<td>Mass region (GeV/c$^2$) excl. $J/\psi$</td>
<td>530892</td>
<td></td>
</tr>
<tr>
<td>Muon quality</td>
<td>276875</td>
<td></td>
</tr>
<tr>
<td>$\chi^2$/d.o.f. of vertex</td>
<td>127509</td>
<td></td>
</tr>
<tr>
<td>Muon $p_T$ (GeV/c)</td>
<td>73555</td>
<td></td>
</tr>
<tr>
<td>Muon $</td>
<td>\eta</td>
<td>$</td>
</tr>
<tr>
<td>Tracking hits</td>
<td>58012</td>
<td></td>
</tr>
<tr>
<td>$\delta L_{xy}$ (mm)</td>
<td>47522</td>
<td></td>
</tr>
<tr>
<td>$B^0_s$ candidate $p_T$ (GeV/c)</td>
<td>54399</td>
<td></td>
</tr>
<tr>
<td>$B^0_s$ $\chi^2$ vertex</td>
<td>53195</td>
<td></td>
</tr>
<tr>
<td>Kaon $p_T$ (GeV/c)</td>
<td>9639</td>
<td></td>
</tr>
<tr>
<td>$\phi$ mass (GeV/c$^2$)</td>
<td>2602</td>
<td></td>
</tr>
</tbody>
</table>

$\sum_{\text{track} \neq B} p_i$ is the scalar sum over all tracks excluding the muon and kaon pairs within a cone of $\Delta R < 1$ around the momentum vector $\vec{p}(\phi \mu^+\mu^-)$ of the $B^0_s$ candidate where $\Delta R = \sqrt{(\Delta \phi)^2 + (\Delta \eta)^2}$. The final discriminating variable used is the pointing angle $\alpha$, defined as the angle between the momentum vector $\vec{p}(\phi \mu^+\mu^-)$ of the $B^0_s$ candidate and the vector $\vec{l}_{ctz}$ between the primary and secondary vertices. This requirement ensures consistency between the direction of the decay vertex and the momentum vector of the $B^0_s$ candidate.
the sideband population for the whole data sample into the mass window signal region, we obtain an expected number of 1.6±0.4 background events with statistical uncertainty only.

Upon examining the data in the mass region, zero candidate events are observed in the signal region, consistent with the background events as estimated from sidebands. Figure 1 shows the remaining events populating the lower and upper sidebands. The Poisson probability of observing zero events for an expected background of 1.6 ± 0.4 is p = 0.22.

In the absence of an apparent signal, a limit on the branching fraction $B(B^0_s \rightarrow \phi \mu^+\mu^-)$ can be computed by normalizing the upper limit on the number of events in the $B^0_s$ signal region to the number of reconstructed $B^0_s \rightarrow J/\psi \phi$ events:

$$\frac{B(B^0_s \rightarrow \phi \mu^+\mu^-)}{B(B^0_s \rightarrow J/\psi \phi)} = \frac{N_{ul}}{N_{B^0_s}} \cdot \frac{\epsilon_{J/\psi \phi}}{\epsilon_{\phi \mu^+\mu^-}} \cdot B(J/\psi \rightarrow \mu^+\mu^-), \quad (2)$$

where $N_{ul}$ is the upper limit on the number of signal decays, estimated from the number of observed events and expected background events, and $N_{B^0_s}$ is the observed number of $B^0_s \rightarrow J/\psi \phi$ events. The measured branching fractions are $B(J/\psi \rightarrow \mu^+\mu^-) = (5.88 \pm 0.10) \times 10^{-2}$ and $B(B^0_s \rightarrow J/\psi \phi) = (9.3 \pm 3.3) \times 10^{-4}$ [13].

The different sources of relative uncertainty that enter into the limit calculation of $B(B^0_s \rightarrow \phi \mu^+\mu^-)$ are given in Table II. The largest uncertainty, 25%, is due to the background interpolation into the signal region and is based on the statistical uncertainty of the fit integral. The uncertainty on the number of observed $B^0_s \rightarrow J/\psi \phi$ events in the normalization channel is 14.8%.

The $p_T$ distribution of the $B^0_s$ in data is on average slightly higher than that from MC. Therefore, MC events for the signal and normalization channels have been reweighted accordingly and an additional uncertainty of 3.7% is applied. The CP-even signal MC events are generated with a $B^0_s$ lifetime of 1.44 ps [14]. To account for a possible efficiency difference related with the shorter lifetime of the CP-even $B^0_s$, the signal MC events are weighted according to the combined world average CP-even lifetime [15]. The efficiency difference is estimated to be 8% which is taken as an additional systematic uncertainty. The statistical uncertainty on the efficiency ratio $\epsilon_{J/\psi \phi}/\epsilon_{\phi \mu^+\mu^-}$ is found to be 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 do not include any theoretical uncertainty in our systematics uncertainty estimation.

The statistical and systematic uncertainties can be included in the limit calculation by integrating over probability functions that parameterize the uncertainties. We use a prescription [16] where we construct a frequentist confidence interval with the Feldman and Cousins [17] ordering scheme for the MC integration. The background
TABLE II: The relative uncertainties found for the upper limit on $\mathcal{B}$.

<table>
<thead>
<tr>
<th>Source</th>
<th>Relative Uncertainty [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td># of $B_s^0 \rightarrow J/\psi \phi$</td>
<td>14.8</td>
</tr>
<tr>
<td>$\epsilon_{J/\psi \phi}/\epsilon_{J/\psi \mu}$</td>
<td>7.5</td>
</tr>
<tr>
<td>MC weighting</td>
<td>3.7</td>
</tr>
<tr>
<td>CP-even lifetime</td>
<td>8.0</td>
</tr>
<tr>
<td>$\mathcal{B}(J/\psi \rightarrow \mu \mu)$</td>
<td>1.7</td>
</tr>
<tr>
<td>Total</td>
<td>18.9</td>
</tr>
<tr>
<td>Background uncertainty</td>
<td>25.0</td>
</tr>
</tbody>
</table>

is 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 Feldman and Cousins (FC) limit is

$$\frac{\mathcal{B}(B_s^0 \rightarrow \phi \mu^+ \mu^-)}{\mathcal{B}(B_s^0 \rightarrow J/\psi \phi)} < 4.4 \times 10^{-3}$$

at the 95% (90%) C.L. respectively. Taking a Bayesian approach [18] with a flat prior and the uncertainties treated as Gaussian distributions in the integration, we find an upper limit of $\mathcal{B}(B_s^0 \rightarrow \phi \mu^+ \mu^-)/\mathcal{B}(B_s^0 \rightarrow J/\psi \phi) < 7.4 (5.6) \times 10^{-3}$ at the 95% (90%) C.L., respectively.

Since we have fewer events observed than expected, we also quote the sensitivity of our search. 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 given by

$$\mathcal{B}(B_s^0 \rightarrow \phi \mu^+ \mu^-)/\mathcal{B}(B_s^0 \rightarrow J/\psi \phi) = 1.1 (1.2) \times 10^{-2}$$

at the 95% C.L. using the FC (Bayesian) approaches, respectively.

Using only the central value of the world average branching fraction [13] of $\mathcal{B}(B_s^0 \rightarrow J/\psi \phi) = (9.3 \pm 3.3) \times 10^{-4}$, the FC limit corresponds to $\mathcal{B}(B_s^0 \rightarrow \phi \mu^+ \mu^-) < 4.1 (3.2) \times 10^{-6}$ at the 95% (90%) C.L. respectively. This is presently the most stringent upper bound and can be compared with the SM calculation of $\mathcal{B}(B_s^0 \rightarrow \phi \mu^+ \mu^-) = 1.6 \times 10^{-6}$ of Ref. [3].

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 and KOSEF (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); The Swedish Research Council (Sweden); Research Corporation; Alexander von Humboldt Foundation; and the Marie Curie Program.

[10] Visitor from Helsinki Institute of Physics, Helsinki, Finland.