Search for $B^0_s \rightarrow \mu^+\mu^-$ and $B^0_d \rightarrow \mu^+\mu^-$ Decays in $p\bar{p}$ Collisions with CDF II

4 Baylor University, Waco, Texas 76798
5 Istituto Nazionale di Fisica Nucleare,
   University of Bologna, I-40127 Bologna, Italy
6 Brandeis University, Waltham, Massachusetts 02254
7 University of California, Davis, Davis, California 95616
8 University of California, Los Angeles, Los Angeles, California 90024
9 University of California, San Diego, La Jolla, California 92093
10 University of California, Santa Barbara, Santa Barbara, California 93106
11 Instituto de Fisica de Cantabria, CSIC-University of Cantabria, 39005 Santander, Spain
12 Carnegie Mellon University, Pittsburgh, PA 15213
13 Enrico Fermi Institute, University of Chicago, Chicago, Illinois 60637
14 Joint Institute for Nuclear Research, RU-141980 Dubna, Russia
15 Duke University, Durham, North Carolina 27708
16 Fermi National Accelerator Laboratory, Batavia, Illinois 60510
17 University of Florida, Gainesville, Florida 32611
18 Laboratori Nazionali di Frascati, Istituto Nazionale
di Fisica Nucleare, I-00044 Frascati, Italy
19 University of Geneva, CH-1211 Geneva 4, Switzerland
20 Glasgow University, Glasgow G12 8QQ, United Kingdom
21 Harvard University, Cambridge, Massachusetts 02138
22 Division of High Energy Physics, Department of Physics,
   University of Helsinki and Helsinki Institute of Physics, FIN-00014, Helsinki, Finland
23 University of Illinois, Urbana, Illinois 61801
24 The Johns Hopkins University, Baltimore, Maryland 21218
25 Institut für Experimentelle Kernphysik,
   Universität Karlsruhe, 76128 Karlsruhe, Germany
26 High Energy Accelerator Research Organization (KEK), Tsukuba, Ibaraki 305, Japan
27 Center for High Energy Physics: Kyungpook National University,
   Taegu 702-701; Seoul National University,
   Seoul 151-742; and SungKyungKwan University, Suwon 440-746; Korea
28 Ernest Orlando Lawrence Berkeley National Laboratory, Berkeley, California 94720
29 University of Liverpool, Liverpool L69 7ZE, United Kingdom
30 University College London, London WC1E 6BT, United Kingdom
31 Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
32 Institute of Particle Physics: McGill University, Montréal, Canada H3A 2T8; and University of Toronto, Toronto, Canada M5S 1A7
33 University of Michigan, Ann Arbor, Michigan 48109
34 Michigan State University, East Lansing, Michigan 48824
35 Institution for Theoretical and Experimental Physics, ITEP, Moscow 117259, Russia
36 University of New Mexico, Albuquerque, New Mexico 87131
37 Northwestern University, Evanston, Illinois 60208
38 The Ohio State University, Columbus, Ohio 43210
39 Okayama University, Okayama 700-8530, Japan
40 Osaka City University, Osaka 588, Japan
41 University of Oxford, Oxford OX1 3RH, United Kingdom
42 University of Padova, Istituto Nazionale di Fisica Nucleare, Sezione di Padova-Trento, I-35131 Padova, Italy
43 University of Pennsylvania, Philadelphia, Pennsylvania 19104
44 Istituto Nazionale di Fisica Nucleare Pisa, Universities of Pisa, Siena and Scuola Normale Superiore, I-56127 Pisa, Italy
45 University of Pittsburgh, Pittsburgh, Pennsylvania 15260
46 Purdue University, West Lafayette, Indiana 47907
47 University of Rochester, Rochester, New York 14627
48 The Rockefeller University, New York, New York 10021
49 Istituto Nazionale di Fisica Nucleare, Sezione di Roma I, University of Rome “La Sapienza,” I-00185 Roma, Italy
50 Rutgers University, Piscataway, New Jersey 08855
51 Texas A&I University, College Station, Texas 77843
52 Istituto Nazionale di Fisica Nucleare, University of Trieste/ Udine, Italy
53 University of Tsukuba, Tsukuba, Ibaraki 305, Japan
54 Tufts University, Medford, Massachusetts 02155
55 Waseda University, Tokyo 169, Japan
56 Wayne State University, Detroit, Michigan 48201
57 University of Wisconsin, Madison, Wisconsin 53706
Abstract

We report on a search for $B_s^0 \rightarrow \mu^+ \mu^-$ and $B_d^0 \rightarrow \mu^+ \mu^-$ decays in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV using 364 pb$^{-1}$ of data collected by the CDF II detector at the Fermilab Tevatron Collider. After applying all selection requirements, we observe no candidates inside the $B_s^0$ or $B_d^0$ mass windows, consistent with the background expectation. The resulting upper limits on the branching fractions are $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) < 1.5 \times 10^{-7}$ and $\mathcal{B}(B_d^0 \rightarrow \mu^+ \mu^-) < 3.9 \times 10^{-8}$ at 90% confidence level.

PACS numbers: 12.15.Mm, 13.20.He, 14.40.Nd
In the standard model (SM), Flavor Changing Neutral Current (FCNC) decays are highly suppressed and can only occur through higher order diagrams. The decay rate for the FCNC decay $B^0_s \rightarrow \mu^+\mu^-$ [1] is proportional to the CKM matrix element $|V_{ts}|^2$. The rate of $B^0_d \rightarrow \mu^+\mu^-$ decays is further suppressed by the ratio of $|V_{td}/V_{ts}|^2$. The SM expectations for these branching fractions are $\mathcal{B}(B^0_s \rightarrow \mu^+\mu^-) = (3.42 \pm 0.54) \times 10^{-9}$ and $\mathcal{B}(B^0_d \rightarrow \mu^+\mu^-) = (1.00 \pm 0.14) \times 10^{-10}$ [2], which are about two orders of magnitude smaller than the current experimental sensitivity. However, new physics contributions can significantly enhance these branching fractions. An observation of these decays at the Tevatron would be unambiguous evidence for physics beyond the SM. The best existing experimental bounds on these branching fractions are $\mathcal{B}(B^0_s \rightarrow \mu^+\mu^-) < 4.1 \times 10^{-7}$ [3] and $\mathcal{B}(B^0_d \rightarrow \mu^+\mu^-) < 8.3 \times 10^{-8}$ [4] at the 90\% confidence level (C.L.).

In minimal supersymmetric (SUSY) extensions of the SM (MSSM), additional diagrams involving SUSY particles also contribute to FCNC decay rates and the branching fraction $\mathcal{B}(B^0_{s,d} \rightarrow \mu^+\mu^-) \propto (\tan \beta)^6$, where $\tan \beta$ is the ratio of vacuum expectation values of the two neutral CP-even Higgs fields. Large values of $\tan \beta$ enhance the decay rate to a level observable by the Tevatron experiments [5]. For example, increases of one to three orders of magnitude are obtained in the minimal SO(10) models [6], which favor large values of $\tan \beta$. For the minimal flavor violating (MFV) models, $\mathcal{B}(B^0_{s,d} \rightarrow \mu^+\mu^-)$ remains suppressed relative to $B^0_s \rightarrow \mu^+\mu^-$ due to the CKM factor $|V_{td}/V_{ts}|^2$. This is not necessarily the case for non-MFV models such as R-parity violating SUSY [8], which can produce large enhancements, even for low values of $\tan \beta$, in either or both of the $B_s$ and $B_d$ FCNC decay rates. Thus, a simultaneous observation of $B^0_s \rightarrow \mu^+\mu^-$ and $B^0_d \rightarrow \mu^+\mu^-$ decays can be important in determining the flavor structure of the new physics. In the absence of an observation, any improvements to the limits can be used to constrain significantly many SUSY models [5–9].

In this Letter, we report on a search for $B^0_s \rightarrow \mu^+\mu^-$ and $B^0_d \rightarrow \mu^+\mu^-$ decays using 364 pb$^{-1}$ of data collected by the upgraded Collider Detector at Fermilab (CDF II). This data set includes 171 pb$^{-1}$ of data from our previous measurement [10]. We significantly improve the sensitivity of the search over our former analysis by doubling our data sample, extending the muon acceptance, and using a likelihood ratio technique for signal and background separation. The limits we present here are the most stringent to date and supersede our previous results.

The CDF II detector is described in detail in Ref. [11]. The inner tracking system is
composed of a silicon microstrip detector (SVX II) [12] surrounded by an open-cell wire drift chamber (COT) [13]. These tracking detectors are immersed in a 1.4 T magnetic field and measure $p_T$, charged particle momentum in the plane transverse to the beamline. Four layers of planar drift chambers (CMU) [14] detect muon candidates with $p_T > 1.4$ GeV/$c$ and provide coverage in the pseudorapidity range $|\eta| < 0.6$, where $\eta = -\ln(\tan \frac{\theta}{2})$ and $\theta$ is the angle of the track with respect to the beamline. The central muon extension (CMX) consists of conical sections of drift tubes and extends the coverage to $0.6 < |\eta| < 1.0$ for muon candidates with $p_T > 2.0$ GeV/$c$.

The data used in this analysis are selected by two classes of dimuon triggers: for the CMU-CMU (U-U) triggers both muon candidates are reconstructed in the CMU chambers, while for the CMU-CMX (U-X) triggers one of the muon candidates is reconstructed in the CMX chambers. The inclusion of the U-X trigger increases the signal acceptance by about 50%. The details of the trigger system and selection requirements can be found in Refs. [10, 11]. Since they have different sensitivities, we treat U-U and U-X channels separately, combining the results at the end.

The offline reconstruction begins by identifying two muon candidates of opposite charge which satisfy the online dimuon trigger requirements. To avoid regions of rapidly changing trigger efficiency, we omit CMU (CMX) muon candidates with $p_T < 2$ (2.2) GeV/$c$. The random combinatoric backgrounds are suppressed by requiring the vector sum of the muon transverse momenta to be $|\vec{p}_T^{\mu\mu}| > 4$ GeV/$c$. The remaining pairs of muon tracks are then refit under the constraint that they come from the same three-dimensional (3D) space point, and are required to satisfy vertex fit quality criteria. The 3D decay length is given by $L_{3D} = \vec{L} \cdot \vec{p}^{\mu\mu}/|\vec{p}^{\mu\mu}|$, where $\vec{L}$ is the displacement vector from the primary to the dimuon vertex. The primary vertex is determined using a constrained vertex fit of all tracks in the event, excluding the $\mu^+\mu^-$ pair and other secondary decay tracks. For each $B$-candidate we estimate the proper decay time $\tau = M_{\mu\mu} L_{3D}/|\vec{p}^{\mu\mu}|$, where $M_{\mu\mu}$ is the invariant mass and $\vec{p}^{\mu\mu}$ is the momentum vector of the dimuon system. Additional background is reduced by demanding $L_{3D} < 1.0$ cm, the uncertainty on $L_{3D}$ to be less than 150 $\mu$m, and $2\sigma_\lambda < \lambda < 0.3$ cm, where $\lambda = c \cdot \tau$ and $\sigma_\lambda$ is the total uncertainty on $\lambda$. There are 22459 (14305) dimuon candidates that fulfill all the above trigger and offline reconstruction requirements in the U-U (U-X) channel. At this stage, the data sample is dominated by random combinatoric background.

We model the signal $B^0_{s(\ell)} \to \mu^+\mu^-$ decays using the PYTHIA Monte Carlo (MC) pro-
FIG. 1: Distributions of the discriminating variables for data (dominated by background) shown in solid histogram and signal $B_{s(d)}^0 \rightarrow \mu^+\mu^-$ events shown in dashed histogram.

The PYTHIA events are passed through a full detector simulation and satisfy the same requirements as data. The $p_T$ spectrum of the $B$-hadron in the MC is weighted to match the measurement from Ref. [11]. The signal MC samples are used primarily for analysis optimization and for estimating the efficiency of selection requirements.

For the final event selection we use the following four variables which provide good discrimination of $B_{s(d)}^0 \rightarrow \mu^+\mu^-$ decays from background events: $M_{\mu\mu}$, $\lambda$, the 3D opening angle $\Delta\Theta$ between vectors $\vec{p}^{\mu\mu}$ and $\vec{L}$, and the $B$-candidate track isolation ($I$) [16]. These variables are also chosen because the correlations among them for background events are small. Figure 1 compares the distributions of these variables for data (which is background-dominated) to MC signal events. Based on the observed distributions, we apply two additional loose requirements, $I > 0.5$ and $\Delta\Theta < 0.7$ rad, to further reject background while maintaining 92% of signal efficiency. We collectively refer to all the selection requirements applied up to this point as the “baseline” requirements. In the data, 6242 (4908) events survive the baseline requirements in the U-U (U-X) sample.

To enhance signal and background separation we construct a multivariate likelihood ratio based on the input variables: $I$, $\Delta\Theta$, and $\lambda$ probability $P(\lambda) = e^{-\lambda/\tau_{BS(d)}}$, where $\tau_{BS(d)}$ is the world average $B_{s(d)}$ lifetime. We use the $P(\lambda)$ variable instead of $\lambda$ in constructing the
FIG. 2: The likelihood ratio distribution for signal (dashed) from PYTHIA MC and background (solid) from data sidebands in the U-U channel. Similar shapes are observed in the U-X channel. 

likelihood ratio because the $P(\lambda)$ distribution is nearly flat, and therefore better behaved in the likelihood. We define the likelihood ratio to be

$$L_R = \frac{\prod_i P_s(x_i)}{\prod_i P_s(x_i) + \prod_i P_b(x_i)},$$

(1)

where $x_1 = I$, $x_2 = \Delta \Theta$, $x_3 = P(\lambda)$, and $P_{s(b)}(x_i)$ is the probability that a signal (background) event has an observed $x_i$. The probability distributions for the signal events are obtained from the signal MC and the background distributions are taken from the data sidebands (defined below). The resulting $L_R$ distributions for the signal and background events are shown in Fig. 2.

The selection optimization is performed without using events in the search window in order to minimize biases in the optimization procedure. The search window is defined by $5.169 < M_{\mu\mu} < 5.469$ GeV/c$^2$, corresponding to approximately ±6 times the two-track invariant mass resolution, which is estimated to be $\sigma_M \approx 24$ MeV/c$^2$. The sideband regions $4.669 < M_{\mu\mu} < 5.169$ GeV/c$^2$ and $5.469 < M_{\mu\mu} < 5.969$ GeV/c$^2$ are used to estimate the background in the signal region and for extensive analysis cross-checks.

The $B_s^0 \rightarrow \mu^+\mu^-$ branching fraction is obtained by normalizing to the number of $B^+ \rightarrow J/\psi K^+ \rightarrow \mu^+\mu^- K^+$ decays collected from the same trigger sample. The $B^+ \rightarrow J/\psi K^+$ mode is reconstructed using similar baseline requirements as the signal mode but
including an additional $p_T > 1$ GeV/c requirement on the kaon candidate. The 90% C.L. upper limit on the branching fraction is given by

$$B(B_s^0 \rightarrow \mu^+ \mu^-)_{90\% C.L.} = \frac{N_{B_s^0}^{90\%}}{N_{B^+}^{90\%}} \cdot \frac{\alpha_{B^+}}{\alpha_{B_s^0}} \cdot \frac{\epsilon_{B_s^0}}{\epsilon_{B_s^0}^{\text{base}}} \cdot \frac{1}{\epsilon_{B_s^0}^{\text{LR}}} \cdot \frac{f_s}{f_s^u} \cdot B(B^+ \rightarrow J/\psi K^+),$$

(2)

where $N_{B_s^0}^{90\%}$ is the number of $B_s^0 \rightarrow \mu^+ \mu^-$ decays at the 90% C.L. for $N$ observed and $N_{bg}$ expected background events. The value of $N_{B_s^0}^{90\%}$ is estimated using the Bayesian approach of Ref. [17] assuming a flat prior for the $B(B_s^0 \rightarrow \mu^+ \mu^-)$ and incorporating Gaussian uncertainties into the limit. The number of reconstructed $B^+ \rightarrow J/\psi K^+$ candidates, $N_{B^+}$, determined from the data using sideband subtraction, is 1785 ± 60 for the U-U and 696 ± 39 for the U-X channel. The parameter $\alpha_{B_s^0}$ ($\alpha_{B^+}$) is the trigger acceptance and $\epsilon_{B_s^0}^{\text{base}}$ ($\epsilon_{B^+}^{\text{base}}$) is the efficiency of the baseline requirements for the signal (normalization) mode. We apply the likelihood ratio requirement only to the signal mode and therefore the efficiency of the likelihood ratio $\epsilon_{B_s^0}^{\text{LR}}$ appears only in the denominator of Eq. 2. We use the fragmentation ratio $f_s/f_s^u = 3.71 \pm 0.41$ [18]. The branching fraction $B(B^+ \rightarrow J/\psi K^+ \rightarrow \mu^+ \mu^- K^+)$ = $(5.88 \pm 0.26) \times 10^{-5}$ is obtained from Ref. [17]. The expression for $B(B_d^0 \rightarrow \mu^+ \mu^-)$ is derived by replacing $B_s^0$ with $B_d^0$ and the fragmentation ratio with $f_s/f_d = 1$.

The expected number of background events is estimated by extrapolating the number of sideband events passing the baseline requirements to the signal window and then scaling by the expected rejection factor $\kappa$ for a given $L_R$ requirement. The parameter $\kappa$ is determined from the background $L_R$ distribution, which we generate by randomly sampling the $\lambda$, $\Delta \Theta$, and $I$ distributions from the data sidebands to improve statistical precision on $\kappa$. The relative uncertainty on $N_{bg}$ is ±15% (±19%) for the U-U (U-X) channel. The dominant contribution comes from the limited statistics of the input distributions used to generate the $L_R$ distribution to determine $\kappa$.

We have cross-checked our background estimate procedure using control samples from the data: like sign $\mu^+ \mu^-$ events, $\mu^+ \mu^-$ events with $\lambda < 0$, and a fake-muon enhanced $\mu^+ \mu^-$ sample in which we demand at least one muon candidate fail the muon quality requirements. We compare our background predictions to the number of events observed in the search window for a wide range of $L_R$ requirements. No statistically significant discrepancies are
observed. For example, requiring $L_R > 0.99$ and summing over all control samples, we predict 3.8 (4.0) background events and observe 4 (3) in the U-U (U-X) channel. In addition we have considered background contributions from $B_{s(d)}^0 \rightarrow h^+ h^-$, where $h^\pm = \pi^\pm$ or $K^\pm$, in the signal window and determined that the contribution from those decays is negligible.

The acceptance ratio $\alpha_{B^+}/\alpha_{B^0_{s(d)}}$ obtained from the signal MC is $0.297 \pm 0.020$ (0.191 $\pm 0.013$) for the U-U (U-X) channel. The uncertainty includes contributions from systematic variations of the modeling of the $B$-hadron $p_T$ spectrum and longitudinal beam profile, and from the statistics of the MC sample.

The quantity $\epsilon^{base}$ includes the trigger and offline reconstruction efficiencies. The trigger efficiency is determined from inclusive data samples unbiased with respect to the triggers used here. The ratio of $B^+$ to $B^0_{s(d)}$ trigger efficiencies is measured to be less than 0.1% away from unity. We evaluate the single track COT, SVX II, and muon efficiencies using a data sample of inclusive $J/\psi \rightarrow \mu^+ \mu^-$ decays collected on single-muon triggers. The relevant double-track efficiencies are computed by convolution with $B^0_{s(d)} \rightarrow \mu^+ \mu^-$ and $B^+ \rightarrow J/\psi K^+$ MC events surviving the trigger requirements. The offline reconstruction efficiency between signal and normalization mode also largely cancels in the ratio with the exception of the kaon efficiency from the $B^+$ decay. Lastly, we obtain the efficiency of the remaining baseline requirements from the signal MC and cross-check the results by comparing $B^+$ data and MC. Combining all effects, we find $\epsilon^{base}_{B^+}/\epsilon^{base}_{B^0_{s(d)}} = 0.920 \pm 0.034$ (0.915 $\pm 0.034$) for the U-U (U-X) channel. The uncertainty is dominated by systematic uncertainties accounting for kinematic differences between $J/\psi \rightarrow \mu^+ \mu^-$ and $B^0_{s(d)} \rightarrow \mu^+ \mu^-$ decays.

The efficiency of the likelihood requirement $L^{L_R}_{B^0}$ is estimated from the signal MC. The efficiency varies from about 70% for $L_R > 0.90$ to 35% for $L_R > 0.99$. We assign to both channels a relative systematic uncertainty of $\pm 5\%$ based on comparisons of $B^+ \rightarrow J/\psi K^+$ MC and data samples.

We optimize the analysis based on the a priori expected 90% C.L. upper limit on $B(B^0_{s,d} \rightarrow \mu^+ \mu^-)$. The expected limit for a given set of optimization requirements is computed by summing the 90% C.L. limits over all possible observations $N$, weighted by the corresponding Poisson probability when expecting $N_{bg}$. We scan over a range of $L_R$ requirements and determine the optimal value to be $L_R > 0.99$. With the optimized selection requirements, the expected number of background events within the signal window of $\pm 60$ MeV/$c^2$ ($\pm 2.5 \sigma_M$) about the world average $B^0_s$ or $B^0_d$ mass [17] is $0.81 \pm 0.12$ (0.66 $\pm 0.13$), and the single-
event-sensitivity is \((1.0 \pm 0.2) \times 10^{-7}\) \([(1.5 \pm 0.3) \times 10^{-7}\) for the U-U [U-X] channel. Using these criteria we observe no events inside either the \(B^0\) or \(B^0_d\) signal box as shown in Fig. 3. The one candidate event that survived the previous analysis selection [10] has an \(L_R\) value of \(\approx 0.8\), and fail the selection requirements here. Using Eq. 2 and combining both channels taking into account the correlated uncertainties, we derive 90\% (95\%) C.L. limits of \(\mathcal{B}(B^0 \to \mu^+\mu^-) < 1.5 \times 10^{-7}\) (2.0 \times 10^{-7}) and \(\mathcal{B}(B^0_d \to \mu^+\mu^-) < 3.9 \times 10^{-8}\) (5.1 \times 10^{-8}). The new limits improve the previous limits [3, 4] by a factor of two and can be used to reduce significantly the allowed parameter space of a broad spectrum of SUSY models.

![Graph of mu+mu- invariant mass distribution versus likelihood ratio](image)

**FIG. 3:** The \(\mu^+\mu^-\) invariant mass distribution versus likelihood ratio for events satisfying baseline requirements for U-U (solid triangle) and U-X (open circle) channels. The \(B^0\) (solid box) and \(B^0_d\) (dashed box) signal regions are also shown. The one candidate event observed in the previous analysis is highlighted with a star symbol (at \(L_R \approx 0.8\)).

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the Bundesministerium für Bildung und Forschung, Germany; the Korean Science and Engineering Foundation and the Korean Research Foundation; the Particle Physics and Astronomy Re-
search Council and the Royal Society, UK; the Russian Foundation for Basic Research; the Comisión Interministerial de Ciencia y Tecnología, Spain; in part by the European Community’s Human Potential Programme under contract HPRN-CT-2002-00292; and the Academy of Finland.

[1] Throughout this paper inclusion of charge conjugate modes is implicit.


[16] The $B$-candidate isolation is defined as $I = |\vec{p}_T^{\mu\mu} |/(\sum_i |\vec{p}_T^i | + |\vec{p}_T^{\mu\mu} |)$, where the sum is over all tracks with $\sqrt{\Delta \eta^2 + \Delta \phi^2} \leq 1$; $\Delta \phi$ and $\Delta \eta$ are the azimuthal angle and pseudorapidity of track $i$ with respect to $\vec{p}_T^{\mu\mu}$.


[18] $f_x$ is the fraction of weakly decaying $B_x$ hadron in b quark fragmentation. We use values
from [17] and take into account the correlations between $f_s$ and $f_u$. 