Measurement of the Forward-Backward Asymmetry in the $B \to K(\pi)\mu^+\mu^-$ Decay and First Observation of the $B_s^0 \to \phi\mu^+\mu^-$ Decay

T. Aalto,22 B. Álvarez González,10 S. Amerio,42 D. Amidei,33 A. Anastassov,37 A. Anno,18 J. Antos,13 G. Apollinari,16 J.A. Appel,16 A. Apresyan,47 T. Arisawa,56 A. Artikov,14 J. Asaadi,52 W. Ashmanskas,16 B. Auerbach,59 A. Aurisano,52 F. Azfar,41 W. Badgett,16 A. Barbaro-Galtieri,27 V.E. Barnes,47 B.A. Barnett,24 P. Barria,45 P. Bartos,13 M. Bause,42 G. Bauer,31 F. Bedeschi,45 D. Beecher,29 S. Behari,24 G. Bellettini,45 J. Bellinger,58 D. Benjamín,15 A. Beretvas,16 A. Bhatti,49 M. Binkley,16 D. Bisello,45 I. Bizjak,29 K.R. Bland,5 C. Blocker,7 B. Blumenfeld,24 A. Bocci,15 A. Bodek,48 D. Bortoletto,47 J. Boudreau,46 A. Boveia,12 B. Brau,16 L. Brigliadori,66 A. Brisuda,13 C. Bromberg,34 E. Brucken,22 M. Bucciantonio,45 J. Budagov,14 H.S. Budd,38 S. Budd,23 K. Burket,16 G. Busetto,42 P. Bussey,20 A. Buzatu,32 S. Cabrera,15 C. Calancha,30 S. Camarda,4 M. Campanelli,34 M. Campbell,23 F. Canelli,16 A. Canepa,44 B. Carls,23 D. Carlsmith,58 R. Carosi,45 S. Carrillo,17 S. Carron,16 B. Casal,10 M. Casarosa,16 A. Castro,66 P. Catantini,16 D. Cau,53 V. Cavaliere,45 M. Cavalli-Sforza,4 A. Cerri,27 L. Cerri,29 Y.C. Chen,1 M. Chertok,8 G. Chiarelli,45 G. Chlachidze,16 F. Chlebana,16 K. Cho,26 D. Ciocio,45 A. Clark,12 D. Clark,7 G. Compostella,42 M.E. Convery,16 J. Conway,8 M.C. Corbo,43 M. Cordelli,18 C.A. Cox,5 D.J. Cox,8 F. Crescioli,45 C. Cuenca Almenar,59 J. Cuevas,10 R. Culbertson,16 D. Dagenhart,16 N. d’Ascanzo,14 M. Datta,16 P. de Barbaro,48 S. De Cecco,50 G. De Lorenzo,4 M. Dell’Orso,45 C. Deluca,4 L. Demortier,9 J. Deng,15 M. Deninno,6 F. Devoto,22 M. D’Errico,42 A. Di Canto,45 B. Di Ruza,45 J.R. Dittmann,5 M. D’Onofrio,28 S. Donati,45 P. Dong,16 T. Dorigo,42 K. Ebina,56 A. Elagin,52 A. Eppig,33 R. Erbacher,8 D. Errede,23 S. Errede,23 N. Ershadat,43 R. Eusebi,52 H.C. Fang,27 S. Farrington,41 M. Feindt,25 J.P. Fernandez,30 C. Ferrarizza,45 R. Field,17 G. Flanagan,35 R. Forrest,18 M.J. Frank,5 M. Franklin,21 J.C. Freeman,16 I. Furic,17 M. Gallinaro,49 J. Galyardt,13 J.E. Garcia,19 A.F. Garfinkel,47 P. Garosi,45 H. Gerberich,23 E. Gerchtin,16 S. Giagni,50 V. Giakounopoulos,3 P. Giannetti,45 K. Gibson,46 C.M. Ginsburg,16 N. Giokaris,3 P. Giromini,18 M. Giunta,45 G. Giurgiu,24 V. Glagolev,14 D. Glenzinski,16 M. Gold,36 D. Goldin,52 N. Goldschmidt,17 A. Golossanov,16 G. Gomez,10 G. Gomez-Caballos,31 M. Goncharov,31 O. González,30 I. Gorelov,36 A.T. Goshaw,15 K. Goulianos,49 A. Gresele,42 S. Grinstein,4 C. Grossi-Palcher,12 R.C. Group,16 J. Guimaraes da Costa,21 Z. Gunay-Unalan,34 C. Haber,27 S.R. Hahn,16 E. Halkiadakis,51 A. Hamaguchi,40 J.Y. Han,48 F. Happacher,18 K. Harai,54 D. Hare,51 M. Hare,55 R.F. Harr,57 K. Hatakeyama,5 C. Hays,41 M. Heck,25 J. Heinrich,44 M. Herndon,58 S. Hewamanelage,5 D. Hidas,51 A. Hocker,16 W. Hopkins,16 D. Horn,25 S. Hou,1 R.E. Hughes,38 M. Hurwit,12 U. Husemann,59 N. Hussain,32 M. Hussein,34 J. Huston,34 G. Introzzi,45 M. Iori,50 A. Ivanov,4 E. James,16 D. Jan,15 B. Jayatilaka,15 E.J. Jeon,26 M.K. Jha,6 S. Jindariani,16 W. Johnson,8 M. Jones,47 K.K. Joo,26 S.Y. Jun,11 T.R. Junk,16 T. Kumon,52 P.E. Karchin,37 Y. Kato,40 W. Ketchum,12 J. Keung,44 V. Khotilovich,52 B. Kilminster,16 D.H. Kim,26 H.S. Kim,26 H.W. Kim,26 J.E. Kim,26 M.J. Kim,18 S.B. Kim,26 S.H. Kim,54 Y.K. Kim,12 N. Kimura,56 S. Klimento,17 K. Kondo,56 D.J. Kong,26 J. Konigsberg,17 A. Korytov,17 A.V. Kotval,15 M. Kreps,25 J. Kroll,44 D. Krop,12 N. Krumnack,5 M. Kruse,15 V. Krutelyov,52 T. Kuhr,25 M. Kurata,54 S. Kwang,12 A.T. Laasanen,47 S. Lami,45 S. Lammel,16 M. Lancaster,29 R.L. Lander,8 K. Lannon,48 A. Lathi,51 G. Latino,45 I. Lazzizzera,42 T. LeCompte,2 E. Lee,52 H.S. Lee,12 J.S. Lee,26 S.W. Lee,52 S. Leocci,45 S. Leone,45 J.D. Lewis,16 C.-J. Lin,27 J. Linacre,41 M. Lindgren,16 E. Lipeles,44 A. Lister,19 D.O. Litvinvsev,16 C. Liu,46 Q. Liu,47 T. Liu,16 S. Lockwitz,59 N.S. Lockyer,44 A. Logino,59 D. Lucchesi,42 J. Lueck,25 P. Lujan,27 P. Lukens,16 G. Lungu,49 J. Lys,27 R. Lysak,13 R. Madak,16 K. Maeshima,16 K. Makhouli,31 P. Maksimovic,24 S. Malik,49 G. Manca,38 A. Manousakis-Katsikakis,3 F. Margaroli,47 C. Marino,25 M. Martínez,4 R. Martínez-Ballarín,30 M. Mastrodemos,50 M. Mathis,24 M.E. Mattson,57 P. Mazzanti,6 K.S. McFarland,48 P. McIntyre,52 R. McNulty,28 A. Mehta,52 P. Mehdi,22 A. Menzione,45 C. Mesropian,49 T. Miao,16 D. Mietlicki,33 A. Mitra,1 H. Miyake,54 S. Moed,21 N. Moggi,6 M.N. Mondragon,16 C.S. Moon,26 R. Moore,16 M.J. Moreno,16 J. Morlock,25 P. Movilla Fernandez,16 A. Mukherjee,16 Th. Muller,28 P. Murat,16 M. Mussini,69 J. Nachtman,59 Y. Nagai,54 J. Nagano,56 I. Nakano,39 A. Napiar,53 J. Nett,58 C. Neu,44 M.S. Neubauer,23 J. Nielsen,27 L. Nodulman,2 O. Norniella,23 E. Nurse,29 L. Oakas,41 S.H. Oh,15 Y.D. Oh,26 I. Oksuzian,17 T. Okusawa,40 R. Orava,22 L. Oortola,4 S. Pagan Griso,42 C. Pugliarone,53 E. Palencia,10 V. Papadimitriou,16 A.A. Paramonov,16 J. Patrick,16 G. Pauletta,53 M. Paulini,11 C. Paus,31 D.E. Pellett,8
We reconstruct the rare decays $B^+ \to K^+ \mu^+ \mu^-$, $B^0 \to K^*(892)^0 \mu^+ \mu^-$, and $B^0 \to (1020)\mu^+ \mu^-$ in a data sample corresponding to 4.4 fb$^{-1}$ collected in $pp$ collisions at $\sqrt{s} = 1.96$ TeV by the CDF II detector at the Tevatron Collider. Using $121 \pm 16 B^+ \to K^+ \mu^+ \mu^-$ and $101 \pm 12 B^0 \to K^{*0} \mu^+ \mu^-$ decays we report the branching ratios. In addition, we report the differential branching ratio and the muon forward-backward asymmetry in the $B^+$ and $B^0$ decay modes, and the $K^{*0}$ longitudinal polarization fraction in the $B^0$ decay mode with respect to the $\phi$ decay. These are consistent with the predictions, and most recent determinations from other experiments and of comparable accuracy. We also report the first observation of the $B^0 \to \phi \mu^+ \mu^-$ decay and measure its branching ratio $\text{BR}(B^0 \to \phi \mu^+ \mu^-) = [1.44 \pm 0.33 \pm 0.46] \times 10^{-6}$ using 27 $\pm$ 6 signal events. This is currently the most rare $B^0$ decay observed.

PACS numbers: 13.25 Hw, 13.20 He
The flavor-changing neutral current process $b \to s \ell \ell$ occurs in the standard model (SM) through higher order diagrams in which new physics contributions could arise. Accurate SM predictions make the $b \to s \ell \ell$ phenomenonology suited to uncover early indications of new physics [1–3], especially through observables like the lepton forward-backward asymmetry ($A_{FB}$) and the differential branching fraction (BR) as a function of dilepton mass $M_{\ell \ell}$. The $b \to s \ell \ell$ amplitudes can be described in terms of short distance operators and effective Wilson coefficients $C_{7,9,10}$. Some new physics models [4] allow the flipped sign of $C_7$. This results in the opposite sign of $A_{FB}$ in the small $q^2$ region ($q^2 \approx M_{\ell \ell}^2$). Recently, BaBar and Belle [4] measured an $A_{FB}$ in the $B^0 \to K^{*0} \ell^+ \ell^-$ decay larger than the SM expectation. The decay $B^0_d \to \phi(1020)\mu^+\mu^-$ has not been seen in previous searches by CDF [5] and D0 [6].

In this Letter we report an update of our previous analysis [5] of the rare decay modes $B^+ \to K^{*0}\mu^+\mu^-$, $B^0 \to K^{*0}\mu^+\mu^-$, and $B^0 \to \phi\mu^+\mu^-$ using an increased data sample of $p\bar{p}$ collisions at a center-of-mass energy of $\sqrt{s} = 1.96$ TeV corresponding to an integrated luminosity of 4.4 fb$^{-1}$, collected with the CDF II detector between March 2002 and January 2009. We update the BR measurements and also report the measurement of $A_{FB}$ in the $B^0 \to K^{*0}\mu^+\mu^-$ decay.

We reconstruct $B \to h \mu^+\mu^-$ candidates, where $B$ stands for $B^+, B^0$, or $B^0_d$, and $h$ stands for $K^+, K^{*0}$, or $\phi$, respectively. Charge conjugation is implied throughout the Letter. The $K^{*0}$ ($\phi$) meson is reconstructed in the decay $K^{*0} \to K^+\pi^-$ ($\phi \to K^+K^-$). We also reconstruct $B \to J/\psi h$ decays as normalization channels in BR measurements, because they have final states identical to those of the signals, resulting in a cancellation of many systematic uncertainties. The relative BR’s are:

$$BR(B \to h \mu^+\mu^-) = \frac{N_{h\mu^+\mu^-}}{N_{J/\psi h}} \frac{\varepsilon_{h\mu^+\mu^-}}{\varepsilon_{J/\psi h}} \times BR(J/\psi \to \mu^+\mu^-),$$

where $N_{h\mu^+\mu^-}$ ($N_{J/\psi h}$) is the $B \to h \mu^+\mu^-$ ($B \to J/\psi h$) yield, and $\varepsilon_{h\mu^+\mu^-}/\varepsilon_{J/\psi h}$ is the relative reconstruction efficiency determined from the simulation.

The CDF II detector is described in detail in Ref. [7] with the detector subsystems relevant for this analysis discussed in Ref. [8].

A sample of dimuon events is selected by the online trigger system. The trigger requires two opposite charged particles with transverse momentum $p_T \geq 1.5$ or 2.0 GeV/$c$ depending on the trigger condition, matched to the muon chambers. We use the muon chambers to detect muons within $|\eta| < 0.6$ and $0.6 < |\eta| < 1.0$ [3].

The trigger also requires $L_{xy} > 200$ $\mu$m, where $L_{xy}$ is the transverse displacement of their intersection from the beamline. The detail of the trigger system and selection requirements can be found in Ref. [3].

The offline loose event selection begins by looking for a common vertex of two trigger muons with one (two opposite-charge) reconstructed charged particle(s) to form a $B^+ \to K^+\mu^+\mu^-$ (a $B^0 \to K^{*0}\mu^+\mu^-$ or a $B^0 \to \phi\mu^+\mu^-$) candidate. The probability of the vertex fit $\chi^2$ is required to be greater than $10^{-3}$. All charged particle trajectories are required to be associated with hits in the silicon vertex detector and to have $p_T \geq 0.4$ GeV/$c$. In addition, we require $p_T(h) \geq 1.0$ GeV/$c$ and $p_T(B) \geq 4.0$ GeV/$c$. We require that the $B$ candidate’s decay is consistent with being displaced from the primary interaction point in the transverse plane by $L_{xy}(B)/\sigma(L_{xy}(B)) \geq 3$, and $\sigma(L_{xy}(B))$ is the estimated uncertainty of $L_{xy}(B)$. We also require that the $B$ candidate comes from the primary vertex by $|d_0(B)| \leq 120$ $\mu$m, where $d_0(B)$ is the distance of closest approach of the $B$ trajectory to the beamline.

For $B^0$ ($B^0_d$) candidates, the $K^+\pi^-$ ($K^+\pi^-$) mass must lie within 50 (10) MeV/$c^2$ of the world average $K^{*0}$ ($\phi$) mass [10]. The ambiguity of the mass assignment in the $K^{*0} \to K^+\pi^-$ decay is handled by choosing the combination with the $K^+\pi^-$ mass closer to the known $K^{*0}$ mass. This results in the correct mass assignments for about 92% of the decays as determined from the simulation. Particle identification is performed with the time-of-flight and the ionization energy loss ($dE/dx$) probabilities of the particle hypothesis. We require loose particle identification for both kaons and pions coming from the $K^{*0}$ meson or $\phi$ meson to reduce combinatorial background. This removes 15% of the $B$ mass sideband events while 99.5% of the signal is retained. We also require a muon likelihood [11] to suppress hadron tracks that produce false trigger muons.

Rare decay candidates with a dimuon mass near the $J/\psi (\psi')$ are rejected: 8.68 (12.86) $< q^2 < 10.09 (14.18)$ GeV$^2/c^2$. To eliminate the radiative charmonium decays that escaped rejection above, we remove candidates consistent with originating from a $B \to J/\psi (\psi')h$ decay followed by the decay of the $J/\psi (\psi')$ into two muons and a photon: $|\sqrt{(M(\mu\mu) - M_{PDG}^{J/\psi (\psi')})^2 - (M_{PDG}^{J/\psi (\psi')})^2}| < 100$ MeV/$c^2$, where the PDG superscript indicates known experimental averages [9] and $M(\mu\mu) < M_{PDG}^{J/\psi (\psi')}$. We also reject candidates if an opposite sign hadron-muon combination of the daughters, assigned the muon mass, satisfy $J/\psi$ or $\psi'$ mass within 40 MeV/$c^2$. This removes charmonium decays where one of the muons is misidentified as a hadron. We reject candidates in which two (three) track combinations are compatible within $\pm 25$ MeV/$c^2$ with $D^0 \to K^-\pi^+$ ($D^+ \to K^-\pi^+\pi^+$ or $D^0_s \to K^+K^-\pi^+$) decays for $B^+$, $B^0$, and $B^0_d$ decays, respectively. This removes $B \to D\pi$ ($D = D^0, D^+$, and $D^0_s$) decays where

Virginia, Charlottesville, VA 22906, aYarmouk University, Irbid 211-63, Jordan, aOn leave from J. Stefan Institute, Ljubljana, Slovenia,
two hadrons are misidentified as muons.

We train an Artificial Neural Network (NN) classifier on simulated signal and a sample of events representative of the background events in the signal region. To simulate the signal we use PyTHIA and EVTGEN based on the SM expectation. The background sample is obtained from the sidebands of the \(B\) invariant mass distribution. We take only the higher mass sideband for the \(B^+\) and \(B^0\) decays since the lower sideband is populated with physics background from partially reconstructed \(B\) meson decays. We use both sidebands for \(B_s^0\) decays. We use 7–10 observables based on \(B\) and daughter’s kinematics (e.g. \(p_T\) and mass), vertex qualities, and muon likelihoods. We optimize the NN threshold in order to maximize both the BR and the \(A_{FB}\) significance. For the \(B^+\) and \(B^0\) analysis we optimize the NN threshold by maximizing \(N_s/\sqrt{N_s + N_b}\), where \(N_s\) (\(N_b\)) is the expected number of signal (background) events. We determine \(N_s\) by Eq. (1) with the world average BR and NN cut efficiency of the simulated signal events, and determine \(N_b\) from the number of sideband events scaled to the signal region, which is defined as \(\pm 2\sigma\) from the world average \(B\) mass, and NN cut efficiency of the sideband events. For \(B_s^0\) decays, \(N_s\) is taken from a theoretical prediction. We maximize \(N_s/(5/2 + \sqrt{N_b})\).

The signal yield is obtained by an unbinned maximum log-likelihood fit to the \(B\) candidate invariant mass distribution. The likelihood is given by

\[
\mathcal{L} = \prod (f_{s\bar{s}} \mathcal{P}_{s\bar{s}} + (1 - f_{s\bar{s}}) \mathcal{P}_b),
\]

where \(f_{s\bar{s}}\) is the signal fraction, \(\mathcal{P}_{s\bar{s}}\) is the signal probability density functions (PDF) parametrized with two Gaussian distributions with different means, and \(\mathcal{P}_b\) is the background PDF modeled with a first- or second-order polynomial. The signal PDF’s are determined from the simulated signal and the \(B\) mass resolution is scaled by the ratio of the mass resolution in \(J/\psi h\) data and simulation, which ranges from 1.07 to 1.09. The background PDF’s are determined from sideband data. Fitted parameters are \(f_{s\bar{s}}\), the mean \(B\) mass, and the background shape. The fit range for \(B^+\) and \(B^0\) (\(B_s^0\)) decays is from 5.18 (5.00) to 5.70 GeV/c², to avoid the region dominated by the physics background.

While the contribution from charmless \(B\) decays is negligible due to the muon identification, we find a sizeable crosstalk between \(B^0 \rightarrow K^{*0}\mu^+\mu^-\) and \(B_s^0 \rightarrow \phi\mu^+\mu^-\) contributing approximately 1% of the signal, as estimated from simulation. These contributions, whose fractions are determined by simulation assuming the world average BR and the theoretical prediction, are subtracted from the fit results for the signal yields.

By optimized NN threshold we reject 99.5–99.8% of background events in the signal region. Figure 1 shows the \(B\) mass distributions. The statistical significance is

\[
s \equiv \sqrt{-2 \ln(\mathcal{L}_\text{null}/\mathcal{L}_\text{max})},
\]

where \(\mathcal{L}_\text{max}\) is obtained from a fit with the signal fraction free to float and the mean \(B\) meson mass fixed to the fitted value in the corresponding normalisation channel, and \(\mathcal{L}_\text{null}\) is the maximum likelihood obtained from a fit with \(f_{s\bar{s}} = 0\). Systematic uncertainty is not considered in the significance evaluation. We obtain \(s = 8.5\sigma, 9.7\sigma,\) and 6.3\(\sigma\) for \(B^+\), \(B^0\), and \(B_s^0\) decays, respectively. The observed yields are listed in Table I. This is the first observation of the \(B_s^0 \rightarrow \phi\mu^+\mu^-\) mode.

TABLE I: Summary of observed yields. The numbers in parentheses are the number of events in the signal region.

| Mode         | \(N_{B^+\mu^+\mu^-}\) | \(N_{J/\psi h}\) | \(\varepsilon_{B^+\mu^+\mu^-}/\varepsilon_{J/\psi h}\) |
|--------------|-------------------------|-----------------|--------------------------------------------------|
| \(B^+\)      | 121 ± 16 (218)          | 437.04 ± 245 (55296) | 0.434 ± 0.006                                   |
| \(B^0\)      | 101 ± 12 (140)          | 15815 ± 178 (22952)  | 0.477 ± 0.009                                   |
| \(B_s^0\)    | 27 ± 6 (40)             | 2930 ± 64 (3883)   | 0.498 ± 0.012                                   |

We do not apply a NN selection to \(J/\psi h\) channels, because these signals are of sufficient size and purity with the loose selection. To obtain the relative efficiency of Eq. (10), the NN cut efficiency of the loosely selected events is considered in addition to the relative efficiency of the loose selection.

The dominant source of systematic uncertainty for each BR measurement is the background PDF parameterization (3.9%) for \(B^+\), the discrepancy of the NN cut efficiency between data and simulation (4.8%) for \(B^0\), and particle identification (3.5%) for \(B_s^0\). For the absolute BR measurements we assign the uncertainties of the world average BR(\(B \rightarrow J/\psi h\))

Results of the relative BR (Eq. (1)) measurements are listed in Table II. The BR statistical uncertainties include the Poisson term from finite statistics of the sample. We also show the absolute BR which is obtained by replacing the normalization channel’s BR with the corresponding world average value.
TABLE II: Measured branching fractions of rare modes. First (second) uncertainty is statistical (systematic).

| Mode | Relative BR($10^{-8}$) | Absolute BR($10^{-8}$) |
|------|--------------------------|-------------------------|
| $B^+ \rightarrow K^+ \mu^+ \mu^-$ | 0.38 ± 0.05 ± 0.02 | 0.38 ± 0.05 ± 0.03 |
| $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ | 0.80 ± 0.10 ± 0.06 | 1.06 ± 0.14 ± 0.09 |
| $B_s^0 \rightarrow \phi \mu^+ \mu^-$ | 1.11 ± 0.25 ± 0.09 | 1.44 ± 0.33 ± 0.46 |

...respect to the dimuon mass. Events in the signal mass region are grouped into independent $q^2$ bins. Figure 2 (a, b) shows the differential BR for $B^+ \rightarrow K^+ \mu^+ \mu^-$ and $B^0 \rightarrow K^{*0} \mu^+ \mu^-$. The $A_{FB}$ and the $K^{*0}$ longitudinal polarization fraction ($F_L$) are extracted by an unbinned likelihood fit to the $\cos \theta_K$ and $\cos \theta_{\mu}$ distributions, respectively, where $\theta_{\mu}$ is the angle between the $\mu^+$ ($\mu^-$) direction and the direction opposite to the $B$ ($\bar{B}$) meson in the dimuon restframe, and $\theta_K$ is the angle between the kaon direction and the direction opposite to the $B$ meson in the $K^{*0}$ rest frame. The differential decay rates [2] are sensitive to $\cos \theta_K$ and $\cos \theta_{\mu}$ through the angular distributions given by $\frac{4}{3} F_L \cos^2 \theta_K + \frac{2}{3} (1 - F_L) (1 - \cos^2 \theta_K)$ for $\cos \theta_K$ and $\frac{1}{2} F_L (1 - \cos^2 \theta_{\mu}) + \frac{1}{2} (1 - F_L) (1 + \cos^2 \theta_{\mu}) + A_{FB} \cos \theta_{\mu}$ for $\cos \theta_{\mu}$. We measure $F_L$ and $A_{FB}$ for $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ and also $A_{FB}$ for $B^+ \rightarrow K^+ \mu^+ \mu^-$. Angular acceptances are obtained from simulated signal samples assuming unpolarized decays.

The contribution from decays with $K$-$\pi$ swapped $K^{*0}$ mesons distorts the signal distribution and swaps the sign of $\cos \theta_K$. This effect is considered by adding an additional signal-like term to the likelihood function. The contribution from decays with non-resonant $K$-$\pi$ is considered to be small [2] and neglected in the fit. For the $B^-$ decay, we set $F_L = 1$ and consider no scalar term [3].

The combinatorial background PDF shape is taken from the $B$ mass upper sideband that is used for the NN training. In the fit to $\cos \theta_K$ ($\cos \theta_{\mu}$) distribution, the only free parameter is $F_L$ ($A_{FB}$). For the $\cos \theta_{\mu}$ fit, the value of $F_L$ is fixed to the $\cos \theta_K$ fit result.

Most dominant source of systematic uncertainty for each angular fit is the fit bias near the physical boundary (0.02–0.07) for $F_L$ in $B^0$, the uncertainty of the $F_L$ fit (0.02–0.12) for $A_{FB}$ in $B^0$, and the angular background shape (0.01–0.07) for $A_{FB}$ in $B^+$. The total systematic uncertainties lie in the range 0.02–0.08 for $F_L$ in $B^0$, 0.05–0.25 for $A_{FB}$ in $B^0$, and 0.02–0.08 for $A_{FB}$ in $B^+$. The angular fit results are shown in Fig. 2 (c, d) and summarized in Table III. Results in the range $0 \leq q^2 < 4.3$ GeV$^2$/c$^2$ and $1 \leq q^2 < 6$ GeV$^2$/c$^2$ are also included.

In summary, we have updated our previous analysis of the flavor-changing neutral current decays $b \rightarrow s \mu \mu$ using data corresponding to an integrated luminosity of 4.4 fb$^{-1}$. We report the first observation of the $B_s^0 \rightarrow \phi \mu^+ \mu^-$, the most rare $B_s^0$ decay observed to date, and measure the total BR. We measure the total BR, differential BR, $A_{FB}$ of the $B^+ \rightarrow K^+ \mu^+ \mu^-$ and $B^0 \rightarrow K^{*0} \mu^+ \mu^-$, with respect to $q^2$. We also measure $F_L$ of $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ prior to $A_{FB}$. These are consistent and competitive with the other current best results. At present there is no evidence of discrepancy from the SM prediction.

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 World Class University Program, the National Research Foundation of Korea; the Science and Technology Facilities Council and the Royal Society, UK; the Institut National de Physique Nucleaire et Physique des Particules/CNRS; the Russian Foundation for Basic Research; the Minis-
TABLE III: Summary of $B^0 \rightarrow K^{*0}\mu^+\mu^-$ and $B^+ \rightarrow K^+\mu^+\mu^-$ fit results. Maximum $q^2$ is 19.30 (23.00) GeV$^2$/c$^2$ for $B^0$ ($B^+$).

| $q^2$ (GeV$^2$/c$^2$) | $\text{BR}(B^0)$ (10$^{-7}$) | $F_L(B^0)$ | $A_{FB}(B^0)$ | $\text{BR}(B^+)$ (10$^{-7}$) | $A_{FB}(B^+)$ |
|----------------------|-----------------------------|-------------|----------------|-----------------------------|----------------|
| (0.00, 2.00)         | 0.98 ± 0.40 ± 0.09          | 0.53 ± 0.34 ± 0.07 | 0.13 ± 0.27 ± 0.25 | 0.38 ± 0.16 ± 0.03 | −0.15 ± 0.39 ± 0.08 |
| (2.00, 4.30)         | 1.00 ± 0.38 ± 0.09          | 0.40 ± 0.33 ± 0.08 | 0.19 ± 0.40 ± 0.14 | 0.58 ± 0.19 ± 0.04 | 0.73 ± 0.32 ± 0.07 |
| (4.30, 8.68)         | 1.69 ± 0.57 ± 0.15          | 0.82 ± 0.19 ± 0.07 | −0.06 ± 0.30 ± 0.05 | 0.93 ± 0.25 ± 0.06 | −0.20 ± 0.17 ± 0.03 |
| (10.09, 12.86)       | 1.97 ± 0.47 ± 0.17          | 0.31 ± 0.18 ± 0.02 | 0.66 ± 0.23 ± 0.07 | 0.72 ± 0.17 ± 0.05 | −0.10 ± 0.17 ± 0.07 |
| (14.18, 16.00)       | 1.51 ± 0.36 ± 0.13          | 0.55 ± 0.18 ± 0.02 | 0.42 ± 0.16 ± 0.09 | 0.38 ± 0.12 ± 0.03 | 0.03 ± 0.16 ± 0.04 |
| (16.00, 19.30(23.00))| 1.35 ± 0.37 ± 0.12          | 0.09 ± 0.14 ± 0.03 | 0.70 ± 0.16 ± 0.10 | 0.35 ± 0.13 ± 0.02 | 0.07 ± 0.30 ± 0.03 |
| (0.00, 4.30)         | 1.98 ± 0.55 ± 0.18          | 0.47 ± 0.24 ± 0.03 | 0.21 ± 0.31 ± 0.05 | 0.96 ± 0.25 ± 0.06 | 0.36 ± 0.24 ± 0.06 |
| (1.00, 6.00)         | 1.60 ± 0.54 ± 0.14          | 0.50 ± 0.35 ± 0.03 | 0.43 ± 0.35 ± 0.06 | 1.01 ± 0.26 ± 0.07 | 0.08 ± 0.24 ± 0.07 |

[1] A. Ali, P. Ball, L. T. Handoko, and G. Hiller, Phys. Rev. D61, 074024 (2000).
[2] F. Kruger and J. Matias, Phys. Rev. D71, 094009 (2005).
[3] C. Bobeth, G. Hiller, and G. Piranishvili, J. High Energy Phys. 12 (2007) 040.
[4] B. Aubert et al. (BaBar Collaboration), Phys. Rev. D79, 031102 (2009); J. T. Wei et al. (Belle Collaboration), Phys. Rev. Lett. 103, 171801 (2009).
[5] T. Aaltonen et al. (CDF Collaboration), Phys. Rev. D79, 011104 (2009).
[6] V. M. Abazov et al. (D0 Collaboration), Phys. Rev. D74, 031107 (2006).
[7] D. Acosta et al. (CDF Collaboration), Phys. Rev. D71, 032001 (2005).
[8] T. Aaltonen et al. (CDF Collaboration), Phys. Rev. Lett. 102, 242002 (2009), and references therein.
[9] We use a cylindrical coordinate system in which $\theta$ is the polar angles with respect to the proton beamline and pseudorapidity $\eta = -\ln(\tan(\theta/2))$.
[10] K. Nakamura et al. (Particle Data Group), J. Phys. G37, 075021 (2010).
[11] T. Aaltonen et al. (CDF Collaboration), Phys. Rev. Lett. 100, 101802 (2008).
[12] T. Sjöstrand et al., Comput. Phys. Commun. 135, 238 (2001); D. J. Lange, Nucl. Instrum. Meth. A462, 152 (2001).
[13] C. Q. Geng and C. C. Liu, J. Phys. G29, 1103 (2003).
[14] G. Punzi, in Proc. of the Conference on Statistical Problems in Particle Physics, Astrophysics and Cosmology (Phystat 2003), (SLAC, Menlo Park, CA, 2003), p. 79.
[15] B. Aubert et al. (BaBar Collaboration), Phys. Rev. Lett. 102, 091803 (2009).