CDF note 10047

Measurement of Forward-Backward Asymmetry in $B \to K^{(*)}\mu^+\mu^-$ and First Observation of $B_s^0 \to \phi\mu^+\mu^-$

The CDF Collaboration
URL http://www-cdf.fnal.gov
(Dated: January 25, 2010)

We study $B^+ \to K^+\mu^+\mu^-$, $B^0 \to K^{*0}\mu^+\mu^-$ and $B_s^0 \to \phi\mu^+\mu^-$ decays in 4.4 fb^{-1} of data collected by the CDF Run II detector with the dimuon trigger. We fully reconstruct one B meson for each decay. Selection criteria are optimized to maximize statistical significance using an artificial neural network technique. In addition to $B^+ \to K^+\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$, we obtain 27 ± 6 $B_s^0 \to \phi\mu^+\mu^-$ decays, corresponding to a 6.3σ statistical significance. This is the first observation of this decay and the rarest B_s^0 decay observed so far. Each branching ratio (BR) is computed using $B^+ \to J/\Psi K^+$, $B^0 \to J/\Psi K^{*0}$, and $B_s^0 \to J/\Psi\phi$ as a reference, respectively.

We then measure differential BR and muon forward-backward asymmetry ($\mathrm{A_{FB}}$) with respect to square of dimuon mass in $B^+ \to K^+\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$. For these analyses we divide the dimuon spectrum into five or six bins. $\mathrm{A_{FB}}$ is measured from a muon angular distribution in the dimuon restframe. In case of $B^0 \to K^{*0}\mu^+\mu^-$, the K^{*0} longitudinal polarization $\mathrm{F_L}$ is measured from the kaon angle in K^{*0} restframe. These differential BR, $\mathrm{A_{FB}}$ and $\mathrm{F_L}$ are measured for the first time at a hadron collider and show comparable resolutions as B-factory measurements.

Preliminary Results for Autumn 2009 Conferences

I. INTRODUCTION

A flavor changing neutral current (FCNC) process is promising probe to test the standard model (SM) precisely and to explore physics beyond SM (BSM), since it is forbidden at the tree level in the SM and occurs via higher order diagrams. Some of the BSM models enhance the decay amplitude via the loop diagrams.

Although the branching ratio (BR) of each $b \to s\ell\ell$ decay, one of the FCNC process, is quite small as $\mathcal{O}(10^{-6})$, the decay is experimentally clean due to opposite sign leptons. Among many $b \to s\ell\ell$ decays, the exclusive channels $B^+ \to K^+\ell^+\ell^-$ and $B^0 \to K^{*0}\ell^+\ell^-$ have been observed at Belle [1, 2] and BaBar [3]. However, the analogous decay $B_s^0 \to \phi\mu^+\mu^-$ has not been observed despite searches by CDF [4] and D0 [5].

Typically B decays amplitudes are effectively calculated using the Wilson coefficients [6] through Operator Product Expansion. The $b \to s\ell\ell$ decay is effectively described by three dominant coefficients: $C_7^{\rm eff}$, $C_9^{\rm eff}$, and $C_{10}^{\rm eff}$. In the SM framework, $C_7^{\rm eff}$ comes from the photon penguins and $C_9^{\rm eff}$ ($C_{10}^{\rm eff}$) comes from the vector (axial vector) component of weak diagrams. They appear in γ, Z penguin diagrams, and W box diagrams at the lowest order. The partial decay rate is proportional to the magnitude of $C_7^{\rm eff}$ ($C_9^{\rm eff}$ and $C_{10}^{\rm eff}$) for small q^2 (large q^2), where $q^2 \equiv M_{\ell\ell}^2 c^2$. Forward-backward asymmetry ($\mathrm{A_{FB}}$) arises from the interference between $C_{10}^{\rm eff}$ term and the others. The BR and $\mathrm{A_{FB}}$ are then sensitive to Wilson coefficients, which can indicate whether the underlying dynamics is governed by the SM or BSM physics like supersymmetry (SUSY) [7], technicolor [8] or fourth generation [9].

Recently BaBar [10] and Belle [11] updated their $\mathrm{A_{FB}}$ measurements in the $B^0 \to K^{*0}\ell^+\ell^-$. It is interesting that both experiments showed larger $\mathrm{A_{FB}}$ than the SM expectation including small q^2 region.

In year 2007 CDF analyzed $B^+ \to K^+\mu^+\mu^-$, $B^0 \to K^{*0}\mu^+\mu^-$ and $B_s^0 \to \phi\mu^+\mu^-$ with integrated luminosity 924 pb^{-1} [4], and reported measurements of the BR of $B^+ \to K^+\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$, which are consistent with the B-factory measurements. For $B_s^0 \to \phi\mu^+\mu^-$, we set the upper limit since the statistics was not sufficient.

This note documents an update of the analysis of the rare decay modes $B^+ \to K^+\mu^+\mu^-$, $B^0 \to K^{*0}\mu^+\mu^-$, and $B_s^0 \to \phi\mu^+\mu^-$ at CDF RunII using approximately 4.4 fb^{-1} of data, collected with the CDFII detector between March 2002 and January 2009. We update the BR measurements and also report the first result of $\mathrm{A_{FB}}$ measurement with $B^0 \to K^{*0}\mu^+\mu^-$ at a hadron collider.

II. THE CDF II DETECTOR

CDF II is a general-purpose particle detector, built with cylindrical symmetry around one of the two collision points of the Tevatron $p\bar{p}$ collider operating at a center-of-mass energy of $\sqrt{s} = 1.96$ TeV, and is described in detail elsewhere [12]. Charged particles are detected with the tracking system, immersed in a 1.4 T solenoidal magnetic field. A 7 layer silicon tracking system, ranging in radius from 1.5 to 28 cm measures precisely the point of origin of charged particle trajectories (tracks), a large drift chamber provides 96 measurements per tracks between 40 and 137 cm radii, allowing an accurate determination of the charged particle's momentum. The drift chamber also provides charged particle identification through the measurement of specific ionization energy loss (dE/dx). A time-of-flight (ToF) detector provides additional particle identification. Drift chambers referred to as CMU and CMX are located at the outermost radial extent of the detector, to detect muons within $|\eta| < 0.6$ and $0.6 < |\eta| < 1.0$, respectively, where $\eta = -\ln(\tan\theta/2)$ and θ is the angle of the track with respect to the beamline. The muon candidates are required to have a momentum transverse to the beamline, p_T, greater than 1.5 or 2.0 GeV/c, depending on the trigger selection.

III. DATA SAMPLE & EVENT SELECTION

We reconstruct $B \to h\mu^+\mu^-$ event as signal candidates, where B stands for B^+, B^0, or B_s^0, and h stands for K^+, K^{*0}, or ϕ respectively. The K^{*0} is reconstructed in the mode $K^{*0} \to K^+\pi^-$, and the ϕ is reconstructed as $\phi \to K^+K^-$. For BR measurements, we also reconstruct $B \to J/\Psi h$ decays as normalization channels, which have final states identical to those of the rare decay modes, resulting in a cancellation of many systematic uncertainties in the ratio of BR.

Signal candidates are selected online with a three-level trigger system. The first trigger level requires the presence of two charged particle trajectories with $p_T \geq 1.5$ GeV/c or 2.0 GeV/c, matched to activity in the muon chambers. At the second level, the trigger rate is reduced by requiring that the muons have opposite charge and that the opening angle in the projection transverse to the beamline is less than 120°. At the third trigger level, the event is fully reconstructed and we select events where the muons are reconstructed in the silicon detector and their intersection has a transverse displacement from the beamline of at least 200 μm.

Candidates for $B \to h\mu^+\mu^-$ and $B \to J/\Psi h$ modes are selected by constructing a vertex of two muons that satisfy the trigger requirements with one reconstructed charged particle to form a $B^+ \to K^+\mu^+\mu^-$ candidate, or with two

reconstructed particles of opposite charge to form $B^0 \to K^{*0}\mu^+\mu^-$ or $B_s^0 \to \phi\mu^+\mu^-$ candidate. The χ^2 probability of the vertex fit is required to be 10^{-3} or higher.

All charged particle trajectories are required to be observed in at least three layers of the silicon vertex detector and have $p_T \geq 0.4\,\mathrm{GeV}/c$. In addition, we require $p_T(h) \geq 1.0\,\mathrm{GeV}/c$ and $p_T(B) \geq 4.0\,\mathrm{GeV}/c$. A lower threshold on the lifetime significance, $ct/\sigma_{ct} \geq 3$ and an upper threshold on the impact parameter of the B candidate, $|d_0(B)| \leq 120\mu\mathrm{m}$, help to reduce background with little loss of signal. For B^0 candidates the $K^+\pi^-$ mass must lie within $50\,\mathrm{MeV}/c^2$ of the K^{*0} mass, and for B_s^0 candidates the K^+K^- mass must lie within $10\,\mathrm{MeV}/c^2$ of the ϕ mass. The ambiguity of the mass assignment in $K^{*0} \to K^+\pi^-$ decay is handled by choosing the combination whose $K^+\pi^-$ mass is closest to the PDG value of the K^{*0} mass. The results in the right mass assignments for about 92% of the decays.

Particle identification is obtained as the combined log likelihood (CLL) that is the logarithm of ratio among ToF and dE/dx probability of each particle hypothesis. We require CLL ≥ -4 for kaon and pion coming from K^{*0} or ϕ to reduce combinatorial background, if CLL is available. We also require muon likelihood $\mathrm{L}(\mu) \geq 0.1$ to suppress hadron track that produced fake trigger muons.

To select the normalization samples, the dimuon invariant mass is required to be within $50\,\mathrm{MeV}/c^2$ of the J/Ψ mass. For the selection of the $B \to h\mu^+\mu^-$ signal, four different vetoes are applied to reduce peaking background,

- Candidates with a dimuon mass near the J/Ψ and ψ' are rejected: $8.68 < M^2(\mu^+\mu^-) < 10.09\,\mathrm{GeV}/c^2$ and $12.86 < M^2(\mu^+\mu^-) < 14.18\,\mathrm{GeV}/c^2$.
- We reject candidates consistent within originating from a $B \to J/\Psi(')h$ decay followed by the decay of the $J/\Psi(')h$ into two muons and a photon: $|(M(\mu\mu\mathrm{h}) - M_B^{\mathrm{PDG}}) - (M(\mu\mu) - M_{\mathrm{J}/\Psi(')}^{\mathrm{PDG}})| > 100\mathrm{MeV}/c^2$, where $M(\mu\mu) < \mathrm{M}_{\mathrm{J}/\Psi(')}^{\mathrm{PDG}}$.
- Candidates with an opposite-sign hadron-muon combination (with the muon mass assigned to both particles) within $40\,\mathrm{MeV}/c^2$ of the J/Ψ or Ψ' mass are rejected. This removes $B \to J/\Psi h$ and $B \to \Psi' h$ decays where one of the hadrons is misidentified as a muon.
- We reject candidates in which two-track combinations are compatible within $\pm 25\,\mathrm{MeV}/c^2$ with $D^0 \to K^-\pi^+$ decays, or three-track combinations are compatible within $\pm 25\,\mathrm{MeV}/c^2$ with $D^+ \to K^-\pi^+\pi^+$ or $D_s^+ \to K^+K^-\pi^+$ decays. This removes $B \to D\pi$ decays where two hadrons are misidentified as muons.

After loose selections above, rare decay channels are tightly selected with using multivariate analysis technique. We adopt the MLP model of Artificial Neural Network (NN) implemented in the TMVA package [13]. To describe the true signal, we use PYTHIA signal Monte-Carlo (MC) simulation. The p_T distribution of signal MC is reweighted by scale factors to match distribution measured in corresponding control channel with its MC. The background events are sampled from the B invariant mass sideband. Because $B^+ \to K^+\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ have significant physics backgrounds in the lower B mass sideband region, we took only the higher side (5σ to 15σ higher than B PDG mass, where $\sigma = 20\mathrm{MeV}/c^2$). $B_s^0 \to \phi\mu^+\mu^-$ use both sideband.

We choose optimized NN cut value that maximizes both BR and $\mathrm{A_{FB}}$ significance. For $B^+ \to K^+\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ we minimize $S/\sqrt{S+B}$, where S is the estimate of the expected yield, and B is the expected background. S is determined from the world averaged BR of rare channels, the observed yield of control channels, and the relative efficiencies described later. B is determined from number of loosely selected events that is dominated by combinatorial background. For $B_s^0 \to \phi\mu^+\mu^-$, S is taken from a theoretical expectation [14] and we minimize $S/(5/2 + \sqrt{B})$ [15], since $B_s^0 \to \phi\mu^+\mu^-$ is not yet observed and we aim at a statistical significance of 5σ.

IV. YIELD FIT AND BR MEASUREMENT

The signal yield is obtained by an unbinned maximum log-likelihood fit in the B invariant mass distribution. The likelihood is composed by the event-by-event signal probability density function (PDF) and background PDF:

$$\mathcal{L} = \prod (f_{\mathrm{sig}}\mathcal{P}_{\mathrm{sig}} + (1 - f_{\mathrm{sig}})\mathcal{P}_{\mathrm{bg}}), \tag{1}$$

where f_{sig} is the signal fraction, $\mathcal{P}_{\mathrm{sig}}$ is the signal PDF parametrized with two Gaussian having the different means, and $\mathcal{P}_{\mathrm{bg}}$ is the background PDF that is described by first or second order polynomial. The signal shape is determined from the signal MC but the B mass resolution is scaled by the resolution ratio of $J/\Psi h$ data with MC.

Possible sources of peaking background are considered for charmless B decays and cross-talk among rare decays. Although contribution from charmless B decays are negligible due to good muon identification, we find sizable cross-talk between $B^0 \to K^{*0}\mu^+\mu^-$ and $B_s^0 \to \phi\mu^+\mu^-$ as $\sim 1\%$ contribution to the observed signal MC yields. These contributions, whose fractions are determined by simulation, are subtracted from the fit results for the signal yields.

For the yield fit, in addition to the statistical uncertainty derived from $f_{\rm sig}$, we take into account further uncertainty coming from fluctuation of total number of event, $\sqrt{N_{\rm tot}}$, where $N_{\rm tot}$ is total number of event in the fit region.

From the B mass fit, we obtained 120 ± 16, 101 ± 12, and 27 ± 6 signals for $B^+ \to K^+\mu^+\mu^-$, $B^0 \to K^{*0}\mu^+\mu^-$, and $B_s^0 \to \phi\mu^+\mu^-$, respectively. Fig. 1 shows B mass plot for each rare decay. Statistical significance s is defined as

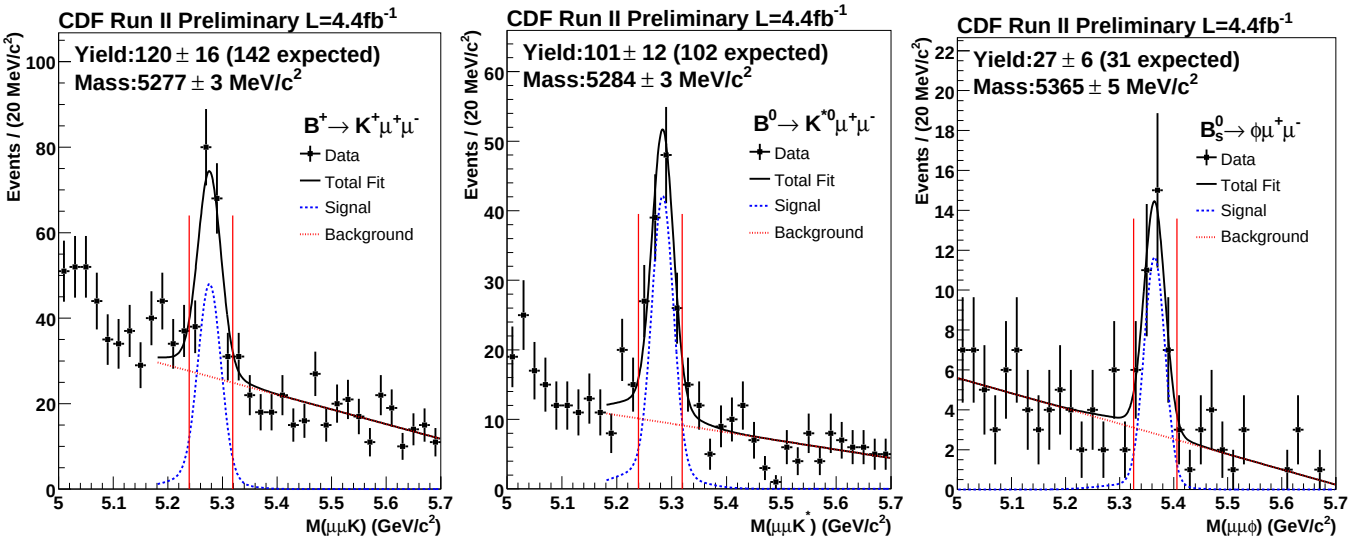

FIG. 1: The B invariant mass of $B^+ \to K^+\mu^+\mu^-$, $B^0 \to K^{*0}\mu^+\mu^-$, and $B_s^0 \to \phi\mu^+\mu^-$ for 4.4 fb^{-1}, respectively. Histogram shows the data. Solid, dashed, and dotted curve shows total fit, signal PDF and background PDF, respectively. Fit range for $B^+ \to K^+\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ ($B_s^0 \to \phi\mu^+\mu^-$) is from 5.18 GeV/c^2 (5.00 GeV/c^2) to 5.70 GeV/c^2. Vertical lines show the signal region.

a difference of log likelihood from null hypothesis: $s \equiv \sqrt{-2\ln(\mathcal{L}_{\rm null}/\mathcal{L}_{\rm max})}$, where $\mathcal{L}_{\rm null}$ is maximum log likelihood obtained from a fit with $f_{\rm sig} = 0$ condition. We fix the B mean mass to the fitted value in the corresponding control channel. We obtain $s = 8.5\sigma$, 9.7σ, and 6.3σ for $B^+ \to K^+\mu^+\mu^-$, $B^0 \to K^{*0}\mu^+\mu^-$, and $B_s^0 \to \phi\mu^+\mu^-$, respectively. This is the first observation of the $B_s^0 \to \phi\mu^+\mu^-$ mode. Obtained yields are consistent with world average and theoretical expectation.

The relative BR to the corresponding control channel is then measured by following formula:

$$\frac{\mathcal{B}(B \to h\mu^+\mu^-)}{\mathcal{B}(B \to J/\Psi h)} = \frac{N^{\rm NN}_{h\mu^+\mu^-}}{N^{\rm loose}_{J/\Psi h}} \frac{\epsilon^{\rm loose}_{J/\Psi h}}{\epsilon^{\rm loose}_{h\mu^+\mu^-}} \frac{1}{\epsilon^{\rm NN}_{h\mu^+\mu^-}} \times \mathcal{B}(J/\Psi \to \mu^+\mu^-), \tag{2}$$

where $N^{\rm NN}_{h\mu^+\mu^-}$ ($N^{\rm loose}_{J/\Psi h}$) is $B \to h\mu^+\mu^-$ ($B \to J/\Psi h$) yields after the optimal NN cut (at the loose selection), $\varepsilon^{\rm loose}_{h\mu^+\mu^-}/\varepsilon^{\rm loose}_{J/\Psi h}$ is the relative efficiency at the loose selection, $\varepsilon^{\rm NN}_{h\mu^+\mu^-}$ is the NN cut efficiency to the loosely selected event. These efficiencies are estimated by signal MC. We do not apply a NN selection to $J/\Psi h$ channels, because the channels have sufficient statistics with the loose selection. The relative efficiency and NN cut efficiency are obtained from MC. We find the relative (NN cut) efficiencies of 0.76 ± 0.01 (0.569 ± 0.002), 0.77 ± 0.01 (0.624 ± 0.003), 0.76 ± 0.02 (0.654 ± 0.004) for $B^+ \to K^+\mu^+\mu^-$, $B^0 \to K^{*0}\mu^+\mu^-$, and $B_s^0 \to \phi\mu^+\mu^-$, respectively.

A. Systematic Uncertainty for the BR measurement

The following systematic errors are evaluated and are summarized in Table I and II:

- Branching Ratios
  We take the branching ratios of a B meson decaying to corresponding normalization mode and $J/\Psi \to \mu^+\mu^-$ from PDG [16] and assign the systematics from these uncertainties.

- Decay Model
  We evaluate the difference in relative efficiency between the default decay model for the rare modes [7] and other models [17–19].

- $p_T(B)$ Spectrum
  Since signal MC is reweighted with using $p_T(B)$ spectrum, the uncertainty related to the $p_T(B)$ spectrum uncertainty is included by switching it with the spectrum of different B meson.

- Trigger Turn-on
  The muon trigger efficiency close to the p_T threshold is not well known. We assign the uncertainty comparing nominal result with analyses using different p_T thresholds from the trigger requirement: 1.6 (2.1), GeV/c for CMU (CMX).

- Particle ID
  To consider different behavior of particle identification between data and MC, the uncertainty is evaluated by repeating the analysis several times employing different cuts of particle identification for both kaon and pion independently.

- Low Momentum Hadrons
  The signal modes have approximately 10% more low momentum tracks than normalization modes, and the simulation models the track reconstruction efficiency at small momentum to an accuracy of approximately 2%. We therefore assign a systematic error on the relative efficiency of $2\% \times 10\% = 0.2\%$.

- B_s^0 Lifetime Difference
  We find an uncertainty of 1.0% on the relative efficiency of $B_s^0 \to \phi\mu^+\mu^-$ decays due to the unknown fraction of short-lived CP-even state in the rare mode. In addition, the uncertainty on $\Delta\Gamma/\Gamma$ contributes another 0.8%, resulting in a total uncertainty of 1.3%.

- Polarization
  We evaluate the effect of the unknown fraction of J/Ψ produced with a longitudinal polarization by varying its fraction by 1σ from the PDG value.

- Control Mode Statistics
  The yields in the normalization modes have associated statistical errors that are included as a systematic error.

- $B^+ \to J/\Psi\pi^+$ Contribution
  Cabibbo-suppressed $B^+ \to J/\Psi\pi^+$ decays contribution are neglected in the nominal fit. Its contribution is estimated with additional signal PDF that has same width as $B^+ \to J/\Psi K^+$ but different mean. We find a 0.5% contribution from the fitted yield.

- MC Statistics
  The relative efficiencies between the signal and normalization modes are obtained from MC samples of finite statistics. We evaluate the uncertainty from the corresponding statistical errors.

- NN Cut
  The uncertainty of the NN cut efficiency due to different NN discriminant behavior between data and MC is studied varying the NN cut from 0.7 to 1.0 at four points.

- Signal PDF
  The signal PDF shape is determined from both data and signal MC. The uncertainties is determined by varying the shape parameters up to $\pm 1\sigma$, where σ is the statistical uncertainty of each quantity.

- Background PDF
  The uncertainty of different background PDF assumption is evaluated by switching a linear shape assumption with second order polynomial.

- Peaking Background
  We find a cross-talk among rare decays up to 1.3%. Number of signal events is corrected from the fitted yield and the difference are taken as the systematic uncertainty.

Source	$B^+ \to K^+\mu^+\mu^-$	$B^0 \to K^{*0}\mu^+\mu^-$	$B_s^0 \to \phi\mu^+\mu^-$
Theory model	0.7	2.0	3.4
MC reweight	2.6	1.4	1.4
Trigger turn-on	< 0.1	3.7	2.5
Particle ID	-	0.3	3.5
Low momentum hadrons	0.2	0.2	0.2
B_s lifetime difference	-	-	1.3
Polarization	-	0.3	0.7
Control mode statistics	0.6	1.2	2.4
B$\to J/\psi\pi$ Contribution	0.5	-	-
MC statistics	1.4	1.8	2.3
NN cut	2.7	4.8	1.6
Efficiency total	4.2	6.9	6.9

TABLE I: Efficiency systematic errors as percentage contributions.

Source	$B^+ \to K^+\mu^+\mu^-$	$B^0 \to K^{*0}\mu^+\mu^-$	$B_s^0 \to \phi\mu^+\mu^-$
Efficiency	4.2	6.9	6.9
$\mathcal{B}(J/\Psi \to \mu^+\mu^-)$	1.0	1.0	1.0
Signal PDF	0.4	0.5	0.2
Background PDF	3.9	2.8	3.1
Peaking BG	0.1	1.3	0.7
$\mathcal{B}(B \to J/\Psi h)$	3.5	4.5	30.8
Total	6.8	8.9	31.7

TABLE II: Total systematic errors as percentage contributions.

B. Relative and absolute branching ratio

We measure the branching fractions of rare decays relative to the corresponding reference channels as follows:

$$\mathcal{B}(B^+ \to K^+\mu^+\mu^-)/\mathcal{B}(B^+ \to J/\Psi K^+) = [0.38 \pm 0.05(\text{stat}) \pm 0.02(\text{syst})] \times 10^{-3},$$
$$\mathcal{B}(B^0 \to K^{*0}\mu^+\mu^-)/\mathcal{B}(B^0 \to J/\Psi K^{*0}) = [0.80 \pm 0.10(\text{stat}) \pm 0.06(\text{syst})] \times 10^{-3},$$
$$\mathcal{B}(B_s^0 \to \phi\mu^+\mu^-)/\mathcal{B}(B_s^0 \to J/\Psi\phi) = [1.11 \pm 0.25(\text{stat}) \pm 0.09(\text{syst})] \times 10^{-3}.$$

Here we note that the dimuon mass spectrum is cut off due to charmonium veto $8.68 < M^2(\mu^+\mu^-) < 10.09\,\text{GeV}/c^2$ and $12.86 < M^2(\mu^+\mu^-) < 14.18\,\text{GeV}/c^2$, though both acceptance losses are corrected suitably.

The absolute branching ratio is then obtained by replacing the control channel's branching ratio with the corresponding PDG [16] value:

$$\mathcal{B}(B^+ \to K^+\mu^+\mu^-) = [0.38 \pm 0.05(\text{stat}) \pm 0.03(\text{syst})] \times 10^{-6},$$
$$\mathcal{B}(B^0 \to K^{*0}\mu^+\mu^-) = [1.06 \pm 0.14(\text{stat}) \pm 0.09(\text{syst})] \times 10^{-6},$$
$$\mathcal{B}(B_s^0 \to \phi\mu^+\mu^-) = [1.44 \pm 0.33(\text{stat}) \pm 0.46(\text{syst})] \times 10^{-6}.$$

This is the first measurement of the $B_s^0 \to \phi\mu^+\mu^-$ branching fraction, the rarest B_s^0 decay observed so far. These numbers are consistent with our previous results [4] and other B-factory measurements [11, 20].

V. DIFFERENTIAL BRANCHING RATIO

We measure the differential BR with respect to the dimuon mass. The signal region is divided into five or six q^2 bins, where $q^2 \equiv M_{\mu\mu}c^2$. Since each q^2 bin has different number of signal and background, we obtain them by the same procedure used in the global yield fit. During the fit, we fix the mean of the B mass and BG slope to the number obtained from the global fit, therefore f_{sig} is only floated. Fig. 2 and Tables III and IV show the differential branching fraction for $B^0 \to K^{*0}\mu^+\mu^-$ and $B^+ \to K^+\mu^+\mu^-$.

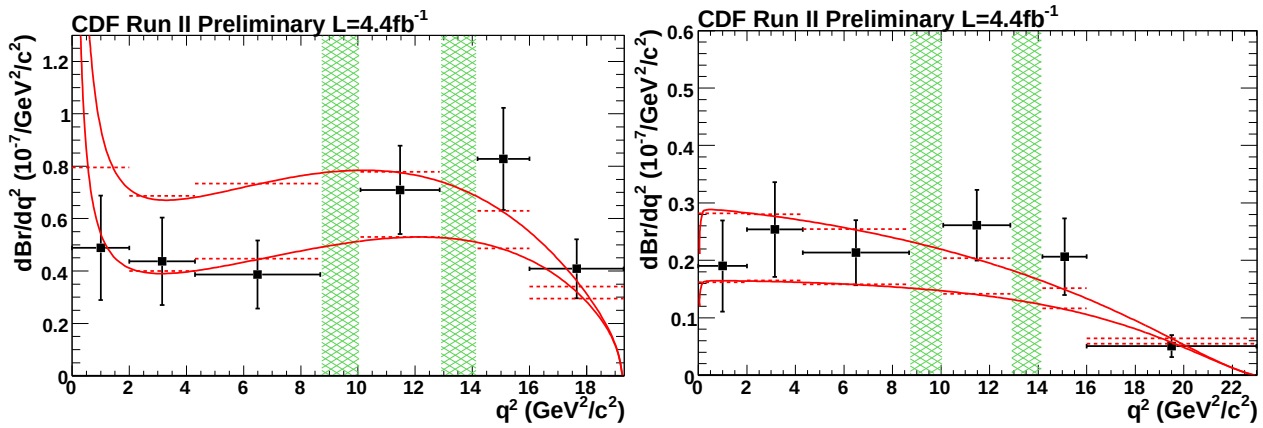

FIG. 2: Differential BR of $B^0 \to K^{*0}\mu^+\mu^-$ (left) and $B^+ \to K^+\mu^+\mu^-$ (right). Hatched regions are charmonium veto regions. Solid lines are the SM expectation [7], which use maximum- and minimum- allowed form factor.

q^2 (GeV$^2/c^2$)	$d\mathcal{B}/dq^2(10^{-8}/(\mathrm{GeV}^2/c^2))$	$\mathcal{B}(10^{-7})$
0.00-2.00	$4.88 \pm 2.00(\mathrm{stat}) \pm 0.43(\mathrm{syst})$	$0.98 \pm 0.40(\mathrm{stat}) \pm 0.09(\mathrm{syst})$
2.00-4.30	$4.37 \pm 1.67(\mathrm{stat}) \pm 0.39(\mathrm{syst})$	$1.00 \pm 0.38(\mathrm{stat}) \pm 0.09(\mathrm{syst})$
4.30-8.68	$3.87 \pm 1.30(\mathrm{stat}) \pm 0.34(\mathrm{syst})$	$1.69 \pm 0.57(\mathrm{stat}) \pm 0.15(\mathrm{syst})$
10.09-12.86	$7.10 \pm 1.69(\mathrm{stat}) \pm 0.63(\mathrm{syst})$	$1.97 \pm 0.47(\mathrm{stat}) \pm 0.17(\mathrm{syst})$
14.18-16.00	$8.28 \pm 1.95(\mathrm{stat}) \pm 0.73(\mathrm{syst})$	$1.51 \pm 0.36(\mathrm{stat}) \pm 0.13(\mathrm{syst})$
16.00-19.30	$4.08 \pm 1.13(\mathrm{stat}) \pm 0.36(\mathrm{syst})$	$1.35 \pm 0.37(\mathrm{stat}) \pm 0.12(\mathrm{syst})$

TABLE III: Differential BR of $B^0 \to K^{*0}\mu^+\mu^-$.

q^2 (GeV$^2/c^2$)	$d\mathcal{B}/dq^2(10^{-8}/(\mathrm{GeV}^2/c^2))$	$\mathcal{B}(10^{-7})$
0.00-2.00	$1.90 \pm 0.79(\mathrm{stat}) \pm 0.13(\mathrm{syst})$	$0.38 \pm 0.16(\mathrm{stat}) \pm 0.03(\mathrm{syst})$
2.00-4.30	$2.54 \pm 0.82(\mathrm{stat}) \pm 0.17(\mathrm{syst})$	$0.58 \pm 0.19(\mathrm{stat}) \pm 0.04(\mathrm{syst})$
4.30-8.68	$2.13 \pm 0.57(\mathrm{stat}) \pm 0.14(\mathrm{syst})$	$0.93 \pm 0.25(\mathrm{stat}) \pm 0.06(\mathrm{syst})$
10.09-12.86	$2.61 \pm 0.62(\mathrm{stat}) \pm 0.18(\mathrm{syst})$	$0.72 \pm 0.17(\mathrm{stat}) \pm 0.05(\mathrm{syst})$
14.18-16.00	$2.06 \pm 0.67(\mathrm{stat}) \pm 0.14(\mathrm{syst})$	$0.38 \pm 0.12(\mathrm{stat}) \pm 0.03(\mathrm{syst})$
16.00-23.00	$0.51 \pm 0.19(\mathrm{stat}) \pm 0.03(\mathrm{syst})$	$0.35 \pm 0.13(\mathrm{stat}) \pm 0.02(\mathrm{syst})$

TABLE IV: Differential BR of $B^+ \to K^+\mu^+\mu^-$.

VI. FORWARD-BACKWARD ASYMMETRY MEASUREMENT

The forward-backward asymmetry ($\mathrm{A_{FB}}$) and K^{*0} longitudinal polarization ($\mathrm{F_L}$) are extracted from $\cos\theta_\mu$ and $\cos\theta_K$ distributions, respectively, where θ_μ is the helicity angle between μ^+ (μ^-) direction and the opposite of the B ($\overline{B}$) direction in the dimuon restframe, and θ_K is the angle between the kaon direction and the direction opposite to the B meson in the K^{*0} rest frame. Fig.3 shows the schematic view of the $B^0 \to K^{*0}\mu^+\mu^-$ angular distribution.

We perform an unbinned maximum likelihood method to extract $\mathrm{A_{FB}}$ and $\mathrm{F_L}$:

$$\mathcal{L} = \prod (f_{\mathrm{sig}}\mathcal{P}_{\mathrm{sig}}(M_B)\mathcal{F}_{\mathrm{sig}}(\cos\theta) + (1 - f_{\mathrm{sig}})\mathcal{P}_{\mathrm{bg}}(M_B)\mathcal{F}_{\mathrm{bg}}(\cos\theta)), \tag{3}$$

where f_{sig} is the signal fraction, $\mathcal{P}_{\mathrm{sig}}$ ($\mathcal{P}_{\mathrm{bg}}$) is the signal (background) PDF of the B mass shape and $\mathcal{F}_{\mathrm{sig}}$ ($\mathcal{F}_{\mathrm{bg}}$) is the signal (background) PDF of the angular shape. In contrast to BR measurement, we consider only the statistical uncertainty derived from f_{sig} for $\mathrm{A_{FB}}$ and $\mathrm{F_L}$ measurement, since we only need the composition of our dataset.

As mentioned above, the differential decay rate in $\cos\theta_K$ is sensitive to $\mathrm{F_L}$ [21]:

$$\frac{1}{\Gamma}\frac{d\Gamma(B^0 \to K^{*0}\mu^+\mu^-)}{d\cos\theta_K} = \frac{3}{2}\mathrm{F_L}\cos^2\theta_\mathrm{K} + \frac{3}{4}(1 - \mathrm{F_L})(1 - \cos^2\theta_\mathrm{K}), \tag{4}$$

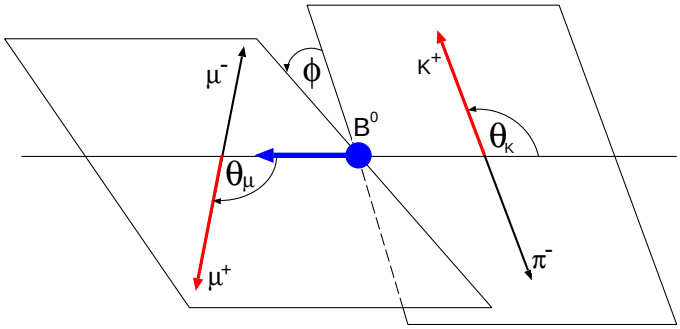

FIG. 3: Schematic view of the $B^0 \to K^{*0}\mu^+\mu^-$ angular distribution.

and the differential decay rate in $\cos\theta_\mu$ is sensitive to $\mathrm{F_L}$ and $\mathrm{A_{FB}}$:

$$\frac{1}{\Gamma}\frac{d\Gamma(B^0 \to K^{*0}\mu^+\mu^-)}{d\cos\theta_\mu} = \frac{3}{4}\mathrm{F_L}(1-\cos^2\theta_\mu) + \frac{3}{8}(1-\mathrm{F_L})(1+\cos^2\theta_\mu) + \mathrm{A_{FB}}\cos\theta_\mu. \tag{5}$$

Angular acceptances of $\cos\theta_K$ and $\cos\theta_\mu$ are considered as a 25-bin histograms which are obtained from phase space signal MC. Sizable contribution from the signal which consists of K-π swapped K^{*0} is seen as a fit bias. Especially, since K-π swapped event swaps the sign of $\cos\theta_\mu$, the fraction works as a dilution factor of asymmetry. To model this effect in the fit we add additional signal-like term that considers opposite $\cos\theta_\mu$ sign. Because we found that the $\cos\theta_K$ acceptance was also affected, a similar PDF is added to the $\cos\theta_K$ function.

Finally, signal PDF are described as:

$$\begin{aligned}
\mathcal{F}^K_{\mathrm{sig}}(\cos\theta_K) \equiv\ & (1-f_{\mathrm{swap}})[\frac{3}{2}\mathrm{F_L}\cos^2\theta_\mathrm{K} + \frac{3}{4}(1-\mathrm{F_L})(1-\cos^2\theta_\mathrm{K})] \times \varepsilon^\mathrm{K}(\cos\theta_\mathrm{K})] \\
& + f_{\mathrm{swap}}[\frac{3}{2}\mathrm{F_L}\cos^2\theta_\mathrm{K} + \frac{3}{4}(1-\mathrm{F_L})(1-\cos^2\theta_\mathrm{K})] \times \varepsilon^\mathrm{K}_{\mathrm{swap}}(\cos\theta_\mathrm{K})], \qquad (6)\\
\mathcal{F}^\mu_{\mathrm{sig}}(\cos\theta_\mu) \equiv\ & (1-f_{\mathrm{swap}})[\frac{3}{4}\mathrm{F_L}(1-\cos^2\theta_\mu) + \frac{3}{8}(1-\mathrm{F_L})(1+\cos^2\theta_\mu) \\
& + \mathrm{A_{FB}}\cos\theta_\mu] \times \varepsilon^\mu(\cos\theta_\mu)] \\
& + f_{\mathrm{swap}}[\frac{3}{4}\mathrm{F_L}(1-\cos^2\theta_\mu) + \frac{3}{8}(1-\mathrm{F_L})(1+\cos^2\theta_\mu) \\
& - \mathrm{A_{FB}}\cos\theta_\mu] \times \varepsilon^\mu_{\mathrm{swap}}(\cos\theta_\mu)], \qquad (7)
\end{aligned}$$

where f_{swap} is the swap fraction, ε^K, $\varepsilon^K_{\mathrm{swap}}$, ε^μ and $\varepsilon^\mu_{\mathrm{swap}}$ are the angular acceptance of true $\cos\theta_K$, swapped $\cos\theta_K$, true $\cos\theta_\mu$ and swapped $\cos\theta_\mu$ events, respectively. For $B^+ \to K^+\mu^+\mu^-$, we set $\mathrm{F_L} = 1$.

The combinatorial background PDF shape is taken from higher B mass sideband, $3\sigma < M_B \leq 15\sigma$, where $\sigma = 20\mathrm{MeV}/c^2$. They are divided into five or six q^2 bins same as the signal and described as a 25-bin histogram.

1. *Control sample*

As control samples for $\mathrm{A_{FB}}$ and $\mathrm{F_L}$ measurements, we fit $\mathrm{F_L}$ and $\mathrm{A_{FB}}$ ($\mathrm{A_{FB}}$) in $B^0 \to J/\Psi K^{*0}$ ($B^+ \to J/\Psi K^+$) decays and compare them with expectation. Acceptance function and background PDF are taken from their own signal MC and B mass sideband data with same procedure as rare channels. We apply NN weight trained for corresponding rare channels to the control samples and cut by same optimal point for the rare channels.

Table V shows the fit result. The measured $\mathrm{F_L}$ is consistent with other measurements, $0.556 \pm 0.009(\mathrm{stat}) \pm 0.010(\mathrm{syst})$ [22] and $0.574 \pm 0.012(\mathrm{stat}) \pm 0.009(\mathrm{stat})$ [23] within the error. However we note that we do not consider S-wave interference effect in our fit. Both measured $\mathrm{A_{FB}}$ are consistent with zero as expectation.

A. Systematics Uncertainties for the angular analysis

The following systematic errors for $\mathrm{F_L}$ and $\mathrm{A_{FB}}$ are evaluated in each q^2 bin independently:

Mode	$N_{\rm sig}$	$\mathrm{F_L}$	$\mathrm{A_{FB}}$
$B^0 \to J/\Psi K^{*0}$	9433.7 ± 101.4	0.572 ± 0.008	0.011 ± 0.009
$B^+ \to J/\Psi K^+$	27439.4 ± 66.8	-	$0.002^{+0.002}_{-0.003}$

TABLE V: Results of the angular analysis on the control sample. Statistical error is only shown.

- Signal Fraction and B Mass Shape
  The uncertainties originated from B mass fit are evaluated by varying the signal fraction and shape parameters up to $\pm 1\sigma$, where σ is statistical uncertainty of each quantity that is obtained from data or signal MC.
- Angular Acceptance
  The angular acceptance is described as a 25-bin histogram that is taken from phase-space signal MC. We evaluate the uncertainty by different binning (20 and 30).
- Angular Background
  For the central fit, we take the background shape from the B mass sideband region $3\sigma < M_B \leq 15\sigma$, where $\sigma = 20\mathrm{MeV}/c^2$. Systematics due to uniformity of the background is investigated by an angular fit using different sideband regions: $3\sigma < M_B \leq 9\sigma$ and $9\sigma < M_B \leq 15\sigma$.
- K-π Swap
  The K-π swap fraction is obtained from signal MC as $7.1 \pm 0.1\%$ in the fit region. To consider discrepancy between data and MC, we switch the fraction to 5% and 10%, and repeat the fit.
- Fit Bias
  Using simulation, we found that if the true value of $\mathrm{F_L}$ or $\mathrm{A_{FB}}$ was close to the physical boundary, the fitted value was affected. Possible fit bias is considered using various combinations of Wilson coefficient. Numerical calculation of each Wilson coefficient is based on Ref. [7] and scaled from -2 to $+2$.
- Trigger Bias
  The trigger dependence on angular acceptances is studied by applying the $p_T > 1.6$ (2.1) GeV/c cut for CMU (CMX) muon to the acceptance function.
- $\mathrm{F_L}$ Fit
  The $\mathrm{F_L}$ statistical error is the largest source of systematic uncertainty for the $\mathrm{A_{FB}}$ fit. The uncertainty is evaluated by varying $\mathrm{F_L}$ up to both positive and negative 1σ.

B. Fit Results

We measure $\mathrm{F_L}$ and $\mathrm{A_{FB}}$ for $B^0 \to K^{*0}\mu^+\mu^-$ and also $\mathrm{A_{FB}}$ for $B^+ \to K^+\mu^+\mu^-$. Fit results with six q^2 bins are shown in Fig. 4. Fig. 5 (6) shows $\cos\theta_K$ ($\cos\theta_\mu$) distribution for $K^{*0}\mu^+\mu^-$. Our choice of bins was driven by allowing easier combination with Belle [11] (most precise result currently).

Fig. 7 shows the fit results with five q^2 bin analysis which combines together the first and the second q^2 bin of the six bins analysis.

Table IX (X) summarize the fit results for $B^0 \to K^{*0}\mu^+\mu^-$ ($B^+ \to K^+\mu^+\mu^-$).

VII. CONCLUSION

- We updated the analysis of FCNC decays $b \to s\mu\mu$ to the 4.4 fb^{-1} sample.
- We report the first observation of the $B_s^0 \to \phi\mu^+\mu^-$, the rarest B_s^0 ever observed, with 6.3σ and measure its BR to be $\mathcal{B}(B_s^0 \to \phi\mu^+\mu^-) = [1.44 \pm 0.33(\mathrm{stat}) \pm 0.46(\mathrm{syst})] \times 10^{-6}$.
- We obtain measurements of BR of the other modes consistent and competitive with current best results: $\mathcal{B}(B^+ \to K^+\mu^+\mu^-) = [0.38 \pm 0.05(\mathrm{stat}) \pm 0.03(\mathrm{syst})] \times 10^{-6}$ and $\mathcal{B}(B^0 \to K^{*0}\mu^+\mu^-) = [1.06 \pm 0.14(\mathrm{stat}) \pm 0.09(\mathrm{syst})] \times 10^{-6}$.
- We report the first measurement of $\mathrm{A_{FB}}$ in hadron collisions, consistent and competitive with current best results.

	Bin#1	Bin#2	Bin#3	Bin#4	Bin#5	Bin#6	Bin#1+Bin#2
signal fraction	0.001	0.020	0.062	0.003	0.009	0.011	0.007
B mass shape	0.001	0.004	0.014	0.010	0.004	0.008	0.001
angular eff.	0.048	0.008	0.008	0.003	0.004	0.003	0.019
angular BG	0.025	0.005	0.015	0.002	0.002	0.005	0.013
K-π swap	0.014	0.011	0.004	0.004	0.006	< 0.001	0.011
fit bias	0.036	0.073	0.024	0.015	0.019	0.021	0.016
trigger bias	0.006	0.003	0.003	0.006	< 0.001	0.003	0.004
total	0.067	0.077	0.071	0.020	0.022	0.026	0.031

TABLE VI: $\mathrm{F_L}$ systematic uncertainty for $B^0 \to K^{*0}\mu^+\mu^-$.

	Bin#1	Bin#2	Bin#3	Bin#4	Bin#5	Bin#6	Bin#1+Bin#2
signal fraction	0.071	0.009	0.007	0.007	< 0.001	0.009	0.009
B mass shape	0.034	0.018	0.006	0.005	0.001	0.003	0.022
angular eff.	0.189	0.036	0.017	0.006	0.021	0.010	0.012
angular BG	0.078	0.031	0.023	0.003	0.013	0.002	0.010
K-π swap	0.013	0.023	0.009	0.009	0.008	0.002	0.018
fit bias	0.036	0.032	0.037	0.046	0.053	0.031	0.024
trigger bias	0.008	0.012	0.010	0.003	< 0.001	0.001	0.017
$\mathrm{F_L}$ fit	0.114	0.119	0.015	0.051	0.073	0.090	0.024
total	0.250	0.136	0.052	0.070	0.094	0.096	0.051

TABLE VII: $\mathrm{A_{FB}}$ systematic uncertainty for $B^0 \to K^{*0}\mu^+\mu^-$.

	Bin#1	Bin#2	Bin#3	Bin#4	Bin#5	Bin#6	Bin#1+Bin#2
signal fraction	0.017	0.054	0.021	0.004	0.007	0.018	0.034
B mass shape	0.028	0.012	0.002	0.002	0.010	0.010	0.012
angular eff.	0.019	0.031	0.005	0.003	0.024	0.007	0.006
angular BG	0.059	0.012	0.018	0.066	0.027	0.008	0.006
fit bias	0.020	0.024	0.011	0.008	0.009	0.011	0.044
trigger bias	0.009	0.036	0.012	0.004	0.007	< 0.001	0.009
total	0.074	0.078	0.032	0.067	0.040	0.026	0.058

TABLE VIII: $\mathrm{A_{FB}}$ systematic uncertainty for $B^+ \to K^+\mu^+\mu^-$.

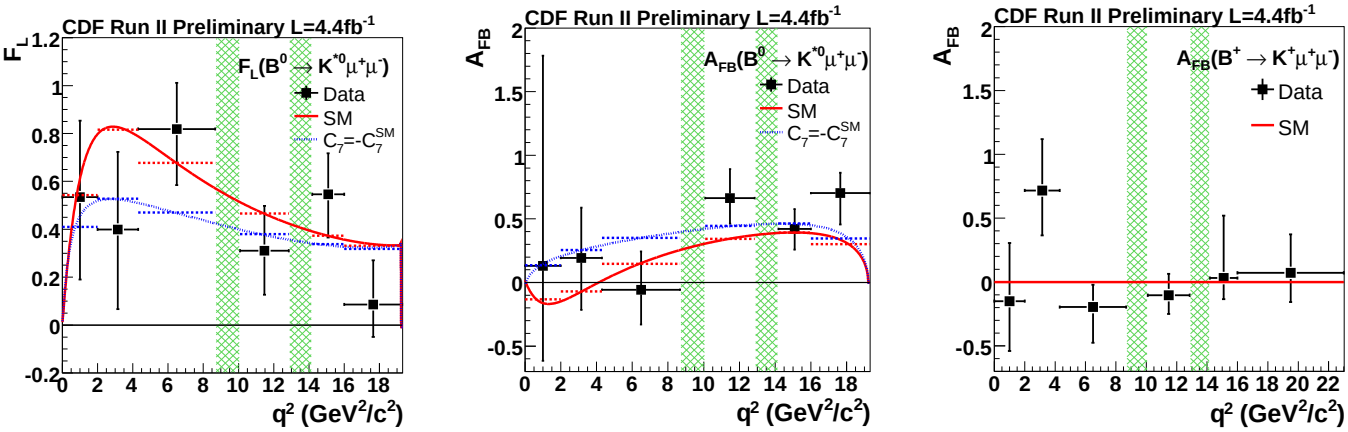

FIG. 4: $\mathrm{F_L}$ and $\mathrm{A_{FB}}$ fit results with respect to six q^2 bins for $B^0 \to K^{*0}\mu^+\mu^-$ and $B^+ \to K^+\mu^+\mu^-$. From left to right, $\mathrm{F_L}$ and $\mathrm{A_{FB}}$ for $B^0 \to K^{*0}\mu^+\mu^-$ and $\mathrm{A_{FB}}$ for $B^+ \to K^+\mu^+\mu^-$ are shown. Histogram is the fit result, solid (dotted) curve is the SM ($C_7 = -C_{7\mathrm{SM}}^{\mathrm{eff}}$) expectation [7], dashed line is the averaged expectation in each q^2 bin, and hatched regions mean charmonium veto.

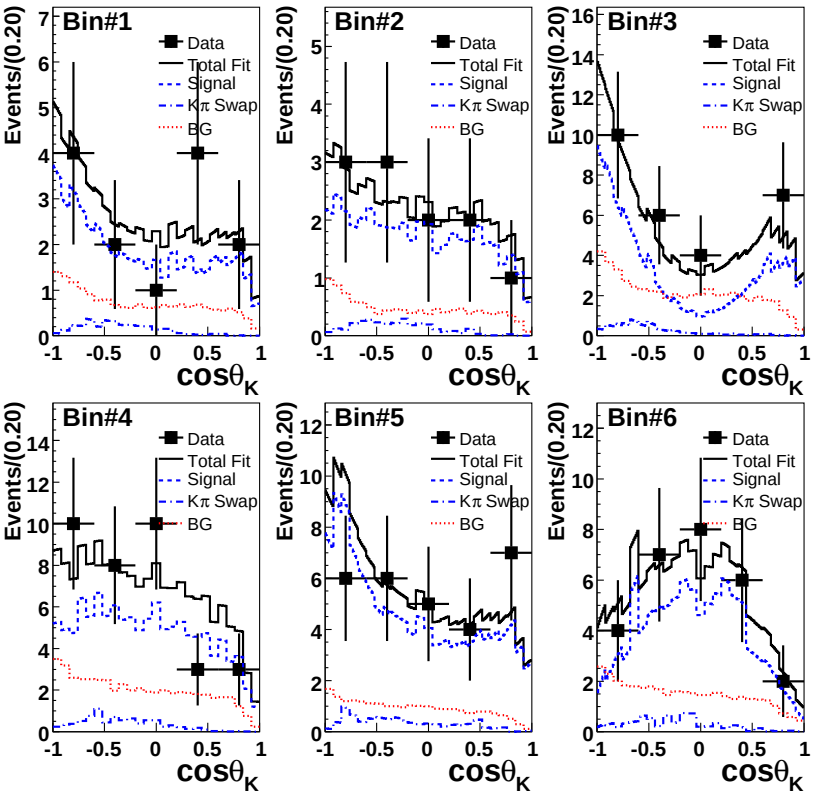

FIG. 5: $\cos\theta_K$ distributions for $B^0 \to K^{*0}\mu^+\mu^-$.

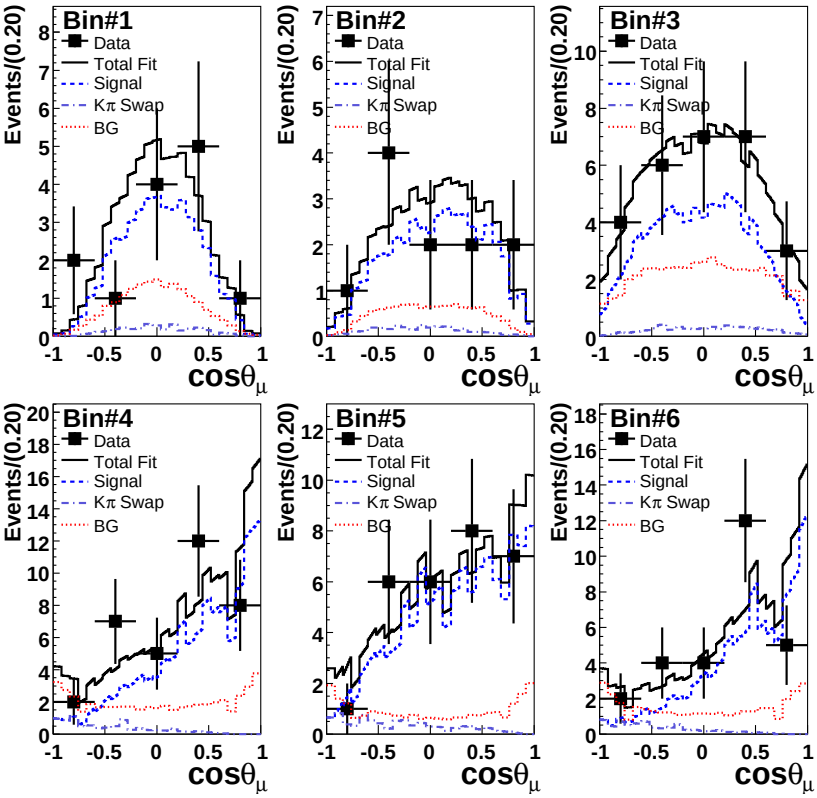

FIG. 6: $\cos\theta_\mu$ distributions for $B^0 \to K^{*0}\mu^+\mu^-$.

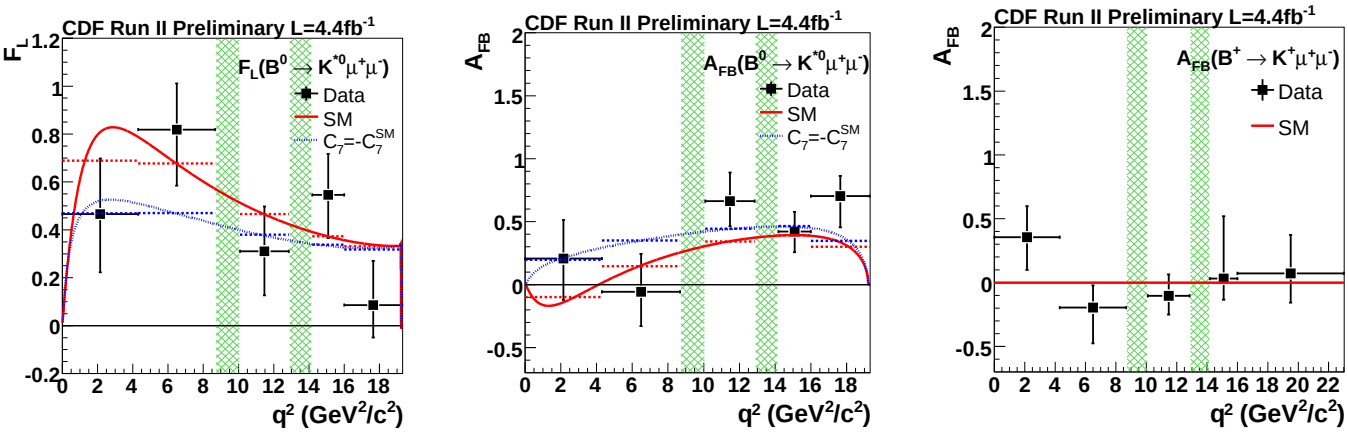

FIG. 7: F_L and A_{FB} fit results with respect to five q^2 bin for $B^0 \to K^{*0}\mu^+\mu^-$ and $B^+ \to K^+\mu^+\mu^-$. From left to right, F_L and A_{FB} for $B^0 \to K^{*0}\mu^+\mu^-$ and A_{FB} for $B^+ \to K^+\mu^+\mu^-$ are shown. Histogram is the fit result, solid (dotted) curve is the SM ($C_7 = -C_{7\,\mathrm{SM}}^{\mathrm{eff}}$) expectation [7], dashed line is the averaged expectation in each q^2 bin, and hatched regions mean charmonium veto.

q^2 (GeV$^2/c^2$)	N_{sig}	$\mathcal{B}(10^{-7})$	F_L	A_{FB}
0.00-2.00	8.52 ± 3.05	$0.98 \pm 0.40 \pm 0.09$	$0.53^{+0.32}_{-0.34} \pm 0.07$	$+0.13^{+1.65}_{-0.75} \pm 0.25$
2.00-4.30	8.91 ± 2.79	$1.00 \pm 0.38 \pm 0.09$	$0.40^{+0.32}_{-0.33} \pm 0.08$	$+0.19^{+0.40}_{-0.41} \pm 0.14$
4.30-8.68	16.86 ± 5.31	$1.69 \pm 0.57 \pm 0.15$	$0.82^{+0.19}_{-0.23} \pm 0.07$	$-0.06^{+0.30}_{-0.28} \pm 0.05$
10.09-12.86	25.71 ± 5.38	$1.97 \pm 0.47 \pm 0.17$	$0.31^{+0.19}_{-0.18} \pm 0.02$	$+0.66^{+0.23}_{-0.20} \pm 0.07$
14.18-16.00	21.91 ± 3.95	$1.51 \pm 0.36 \pm 0.13$	$0.55^{+0.17}_{-0.18} \pm 0.02$	$+0.42^{+0.16}_{-0.16} \pm 0.09$
16.00-19.30	19.78 ± 4.78	$1.35 \pm 0.37 \pm 0.12$	$0.09^{+0.18}_{-0.14} \pm 0.03$	$+0.70^{+0.16}_{-0.25} \pm 0.10$
0.00-4.30	17.43 ± 4.13	$1.98 \pm 0.55 \pm 0.18$	$0.47^{+0.23}_{-0.24} \pm 0.03$	$+0.21^{+0.31}_{-0.33} \pm 0.05$

TABLE IX: Summary of $B^0 \to K^{*0}\mu^+\mu^-$ fit results. First (second) error is statistical (systematic). Last row shows the result of combined first and second bin fit.

q^2 (GeV$^2/c^2$)	N_{sig}	$\mathcal{B}(10^{-7})$	F_L	A_{FB}
0.00-2.00	11.58 ± 4.60	$0.38 \pm 0.16 \pm 0.03$	-	$-0.15^{+0.46}_{-0.39} \pm 0.08$
2.00-4.30	18.02 ± 5.48	$0.58 \pm 0.19 \pm 0.04$	-	$+0.72^{+0.40}_{-0.35} \pm 0.07$
4.30-8.68	34.53 ± 8.87	$0.93 \pm 0.25 \pm 0.06$	-	$-0.20^{+0.17}_{-0.28} \pm 0.03$
10.09-12.86	29.15 ± 6.24	$0.72 \pm 0.17 \pm 0.05$	-	$-0.10^{+0.17}_{-0.15} \pm 0.07$
14.18-16.00	15.98 ± 4.64	$0.38 \pm 0.12 \pm 0.03$	-	$+0.03^{+0.49}_{-0.16} \pm 0.04$
16.00-23.00	13.94 ± 5.00	$0.35 \pm 0.13 \pm 0.02$	-	$+0.07^{+0.30}_{-0.23} \pm 0.02$
0.00-4.30	29.37 ± 7.15	$0.96 \pm 0.25 \pm 0.06$	-	$+0.36^{+0.24}_{-0.26} \pm 0.06$

TABLE X: Summary of $B^+ \to K^+\mu^+\mu^-$ fit results. First (second) error is statistical (systematic). Last row shows the result of combined first and second bin fit.

Acknowledgments

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 Ministerio de Ciencia e Innovación, and Programa

Consolider-Ingenio 2010, Spain; the Slovak R&D Agency; and the Academy of Finland.

[1] K. Abe *et al.* (BELLE), Phys. Rev. Lett. **88**, 021801 (2002), arXiv:hep-ex/0109026.
[2] A. Ishikawa *et al.* (BELLE), Phys. Rev. Lett. **91**, 261601 (2003), arXiv:hep-ex/0308044.
[3] B. Aubert *et al.* (BABAR), Phys. Rev. Lett. **91**, 221802 (2003), arXiv:hep-ex/0308042.
[4] T. Aaltonen *et al.* (CDF), Phys. Rev. **D79**, 011104 (2009), arXiv:0804.3908.
[5] V. M. Abazov *et al.* (D0), Phys. Rev. **D74**, 031107 (2006), arXiv:hep-ex/0604015.
[6] K. G. Wilson, Phys. Rev. **179**, 1499 (1969).
[7] A. Ali, P. Ball, L. T. Handoko, and G. Hiller, Phys. Rev. **D61**, 074024 (2000), arXiv:hep-ph/9910221.
[8] Y. Su, Phys. Rev. **D56**, 335 (1997), arXiv:hep-ph/9612485.
[9] T. M. Aliev, A. Ozpineci, and M. Savci, Eur. Phys. J. **C29**, 265 (2003), arXiv:hep-ph/0301078.
[10] B. Aubert *et al.* (BABAR), Phys. Rev. **D79**, 031102 (2009), arXiv:0804.4412.
[11] J. T. Wei *et al.* (BELLE), Phys. Rev. Lett. **103**, 171801 (2009), arXiv:0904.0770.
[12] F. Abe, *et al.*, Nucl. Instrum. Methods Phys. Res. A **271**, 387 (1988); D. Amidei, *et al.*, Nucl. Instum. Methods Phys. Res. A **350**, 73 (1994); F. Abe, *et al.*, Phys. Rev. D **52**, 4784 (1995); P. Azzi, *et al.*, Nucl. Instrum. Methods Phys. Res. A **360**, 137 (1995); The CDFII Detector Technical Design Report, Fermilab-Pub-96/390-E.
[13] A. Hocker *et al.*, PoS **ACAT**, 040 (2007), arXiv:physics/0703039.
[14] C. Q. Geng and C. C. Liu, J. Phys. **G29**, 1103 (2003), arXiv:hep-ph/0303246.
[15] G. Punzi, Comments on likelihood fits with variable resolution, 2004, arXiv:physics/0401045.
[16] C. Amsler *et al.* (Particle Data Group), Phys. Lett. **B667**, 1 (2008).
[17] D. Melikhov and B. Stech, Phys. Rev. **D62**, 014006 (2000), arXiv:hep-ph/0001113.
[18] P. Colangelo, F. De Fazio, and P. Santorelli, Phys. Rev. **D51**, 2237 (1995), arXiv:hep-ph/9409438.
[19] P. Ball and R. Zwicky, Phys. Rev. **D71**, 014029 (2005), arXiv:hep-ph/0412079.
[20] B. Aubert *et al.* (BABAR), Phys. Rev. Lett. **102**, 091803 (2009), arXiv:0807.4119.
[21] F. Kruger and J. Matias, Phys. Rev. **D71**, 094009 (2005), arXiv:hep-ph/0502060.
[22] B. Aubert *et al.* (BABAR), Phys. Rev. **D76**, 031102 (2007), arXiv:0704.0522.
[23] R. Itoh *et al.* (Belle), Phys. Rev. Lett. **95**, 091601 (2005), arXiv:hep-ex/0504030.