



Precise measurements of exclusive $b \rightarrow s\mu^+\mu^-$ decay amplitudes using the full CDF data set

The CDF Collaboration

URL <http://www-cdf.fnal.gov>

(Dated: July 12, 2012)

We study the flavor-changing neutral current (FCNC) decays $B^+ \rightarrow K^+\mu^+\mu^-$, $B^0 \rightarrow K^*(892)^0\mu^+\mu^-$, $B_s^0 \rightarrow \phi(1020)\mu^+\mu^-$, $B^0 \rightarrow K_S^0\mu^+\mu^-$, $B^+ \rightarrow K^*(892)^+\mu^+\mu^-$, and $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ in 9.6 fb^{-1} of data collected by the CDF detector with a dimuon trigger. Besides the total branching fraction and the differential branching fraction measurements, we measure the isospin asymmetry A_I between neutral and charged B mesons. We also measure the muon forward-backward asymmetry A_{FB} , K^* longitudinal polarization F_L , the transverse polarization asymmetry $A_T^{(2)}$, and time-reversal-odd charge-and-parity asymmetry A_{im} from the angular distributions in $B \rightarrow K^{(*)}\mu^+\mu^-$ decays, as a function of $q^2 = M_{\mu\mu}^2 c^2$, where $M_{\mu\mu}$ is the dimuon mass. No deviations from the standard model predictions are observed.

I. INTRODUCTION

Rare decays of b hadrons mediated by the flavor-changing neutral current (FCNC) process $b \rightarrow s\mu^+\mu^-$ occur in the standard model (SM) with $\mathcal{O}(10^{-6})$ rates through higher order amplitudes [1, 2]. A variety of beyond-the-standard-model (BSM) theories, on the other hand, favor enhanced rates for these FCNC decays [2–4].

Experimental measurements on branching ratios and their dependence on the di-lepton mass distributions, therefore probe SM predictions. Exclusive decays of $B \rightarrow K^{(*)}\mu^+\mu^-$ have been observed by BaBar [5], Belle [6], CDF [7], and LHCb [8]. CDF also reported the first observations of $B_s^0 \rightarrow \phi(1020)\mu^+\mu^-$ [7] and $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ [9]. Although there is no significant departure from the SM has been found thus far, recently LHCb reported the significance of 4.4 standard deviation from zero in the isospin asymmetry between the branching ratio of $B^+ \rightarrow K^+\mu^+\mu^-$ and $B^0 \rightarrow K^0\mu^+\mu^-$ [10].

On the other hand, one can extract more detailed information from the angular distributions of the decay products [2, 11, 12]. The full differential decay distribution for the decay $B \rightarrow K^*\mu^+\mu^- \rightarrow K\pi\mu^+\mu^-$ are described by four independent kinematic variables; the di-muon invariant mass squared $q^2 \equiv M_{\mu\mu}^2 c^2$, the angle θ_μ between the μ^+ (μ^-) direction and the direction opposite to the B (\bar{B}) meson in the di-muon rest frame, the angle θ_K between the kaon direction and the direction opposite to the B meson in the K^* rest frame, and the angle ϕ between the two planes formed by the di-muon and the K - π systems. The distributions of θ_μ , θ_K , and ϕ are projected from the full differential decay distribution and can be parametrized with four angular observables, A_{FB} , F_L , $A_T^{(2)}$ and A_{im} [11, 13, 14]:

$$\begin{aligned} \frac{1}{\Gamma} \frac{d\Gamma}{d\cos\theta_K} &= \frac{3}{2} F_L \cos^2\theta_K + \frac{3}{4} (1 - F_L)(1 - \cos^2\theta_K); \\ \frac{1}{\Gamma} \frac{d\Gamma}{d\cos\theta_\mu} &= \frac{3}{4} F_L (1 - \cos^2\theta_\mu) + \frac{3}{8} (1 - F_L)(1 + \cos^2\theta_\mu) + A_{FB} \cos\theta_\mu; \\ \frac{1}{\Gamma} \frac{d\Gamma}{d\phi} &= \frac{1}{2\pi} \left[1 + \frac{1}{2} (1 - F_L) A_T^{(2)} \cos 2\phi + A_{im} \sin 2\phi \right]; \end{aligned} \quad (1)$$

where $\Gamma \equiv \Gamma(B \rightarrow K^*\mu^+\mu^-)$, A_{FB} is the muon forward-backward asymmetry, F_L is the K^* longitudinal polarization fraction, $A_T^{(2)}$ is the transverse polarization asymmetry [11], and A_{im} is the triple-product asymmetry of the transverse polarizations [13].

In the SM framework the theory calculations of the angular observables at different q^2 ranges are described in terms of the short-distance Wilson coefficients $C_{7,9,10}^{\text{eff}}$. Since various BSM models have different K^* spin components from the SM, the measurement of F_L could constrain BSM parameter regions [11]. The quantity A_{FB} is expected to be small at low q^2 and to have a large positive value at high q^2 in the SM. Contributions from BSM can change the magnitude and the sign of A_{FB} . The quantity $A_T^{(2)}$ is expected to have a small value at low q^2 and a large negative value at high q^2 in the SM [15]. However, the existence of right-handed current could enhance its value up to $\pm\mathcal{O}(1)$ [11, 14, 16]. The quantity A_{im} is expected to be very close to zero for all accessible q^2 , and thus very sensitive to the possible existence of right-handed current in BSM physics.

BaBar [17], Belle [6], CDF [18], and LHCb [8] have reported measurements of A_{FB} and F_L in the $B \rightarrow K^*\ell^+\ell^-$ decay modes. CDF [18] and LHCb also measured $A_T^{(2)}$ and A_{im} . All measurements are consistent with the SM so far.

In this note we report the updated measurements of exclusive $b \rightarrow s\mu^+\mu^-$ decays using the data corresponds to an integrated luminosity of 9.6 fb^{-1} $p\bar{p}$ collisions at a center-of-mass energy of $\sqrt{s} = 1.96 \text{ TeV}$ collected with the CDF II detector during whole Run II period. Besides the measurements of total and differential branching ratios in $B^+ \rightarrow K^+\mu^+\mu^-$, $B^0 \rightarrow K^*(892)^0\mu^+\mu^-$, $B_s^0 \rightarrow \phi\mu^+\mu^-$, $B^0 \rightarrow K^0\mu^+\mu^-$, $B^+ \rightarrow K^*(892)^+\mu^+\mu^-$, and $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$, we measure the combined branching ratio assuming isospin symmetry and the isospin asymmetry between neutral and charged B mesons.

To cancel dominant systematic uncertainties, decay rates for each rare channel $H_b \rightarrow h\mu^+\mu^-$ are measured relative to the corresponding resonant channel $H_b \rightarrow J/\psi h$, used as a normalization and cross-check of the whole analysis, where H_b stands for B^+ , B^0 , B_s^0 , and Λ_b^0 , and h stands for K^+ , K^{*0} , ϕ , K_S^0 , K^{*+} , and Λ .

We also report measurements of the angular distributions in the decay $B \rightarrow K^{(*)}\mu^+\mu^-$. We report the updated measurements of A_{FB} , F_L , $A_T^{(2)}$, and A_{im} with higher precision.

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 $\sqrt{s} = 1.96 \text{ TeV}$ and is described in detail

elsewhere [19]. Charged particles are detected with the tracking system, immersed in a 1.4 T solenoidal magnetic field. A seven-layer silicon tracking system, ranging in radius from 1.5 to 22 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. A second set of muon drift chambers (CMP), located behind the CMU and an additional steel absorber, covers $|\eta| < 0.6$. The muon candidates are required to have a momentum p_T transverse to the beamline greater than 1.5, 2.0, or 3.0 GeV/c, depending on the trigger selection.

III. DATA SAMPLE & EVENT SELECTION

We reconstruct $H_b \rightarrow h\mu^+\mu^-$ event as signal candidates and $H_b \rightarrow J/\psi h$ event as normalization channels. For K^{*0} , K^{*+} , ϕ , and Λ reconstruction, we collect $K^+\pi^-$, $K_S^0\pi^+$, K^+K^- , and $p^+\pi^-$ combinations, respectively. The K_S^0 meson is reconstructed in the mode $K_S^0 \rightarrow \pi^+\pi^-$.

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 100 μm . The standard selection which we call the LXY trigger. To increase the signal acceptance we use another third level trigger, the CMUP trigger, which requires a muon track that has hits on both CMU and CMP chambers and $p_T \geq 3.0$ GeV/c. We do not require the intersection displacement for the CMUP trigger. Since we accept the CMUP trigger only when the LXY trigger fails, both triggers are exclusive. For the final update we introduce a new category of dimuon trigger that requires both muon candidates match with displaced tracks reconstructed by the silicon vertex trigger (SVT) [20]. We also employ additional CMX sectors which were unstable in early CDF operation but worked well later. These new trigger categories increase the acceptance by 8%.

Candidates for $H_b \rightarrow h\mu^+\mu^-$ and $H_b \rightarrow J/\psi h$ modes are selected by constructing a vertex of two muons that satisfy the trigger requirements with one reconstructed charged particle to form a $B^+ \rightarrow K^+\mu^+\mu^-$ candidate, or with two reconstructed particles of opposite charge to form $B^0 \rightarrow K^{*0}\mu^+\mu^-$, $B_s^0 \rightarrow \phi\mu^+\mu^-$, $B^0 \rightarrow K_S^0\mu^+\mu^-$, or $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ candidate. The $B^+ \rightarrow K^{*+}\mu^+\mu^-$ candidate is selected by the two trigger muons with a K_S^0 candidate and a charged particle. The χ^2 probability of the vertex fit is required to be 10^{-3} or higher.

All charged particle trajectories except K_S^0 and Λ daughters are required to be observed in at least three layers of the silicon vertex detector. The charged tracks have $p_T \geq 0.4$ GeV/c, while K_S^0 and K^{*+} daughters have $p_T \geq 0.3$ GeV/c. We require $p_T(h) \geq 1.0$ GeV/c and $p_T(H_b) \geq 4.0$ GeV/c. A lower threshold on the lifetime significance, $ct(H_b)/\sigma_{ct}(H_b) \geq 5$ and an upper threshold on the impact parameter of the H_b candidate, $|d_0(H_b)| \leq 120\mu\text{m}$, help to reduce background with little loss of signal. We require $ct(V^0) \geq 0.5$ cm and $ct(V^0)/\sigma_{ct}(V^0) \geq 3$, where V^0 stands for K_S^0 and Λ . The $\pi^+\pi^-$ mass must lie within $20\text{ MeV}/c^2$ of the K_S^0 (PDG) mass, where PDG refers to the world average mass [21]. The K^+K^- and $p\pi^-$ combinations must have invariant mass within $10\text{ MeV}/c^2$ of the ϕ and Λ mass. We require the $K^+\pi^-$ mass must be in the range $-50 < M(K^+\pi^-) - M(K^{*0}) < 100\text{ MeV}/c^2$ where $M(K^{*0})$ is the K^{*0} pole mass, and the $K_S^0\pi^+$ mass must lie within $\pm 100\text{ MeV}/c^2$ of the K^{*+} pole mass. The ambiguity of the mass assignment in $K^{*0} \rightarrow K^+\pi^-$ decay is handled by choosing the combination whose $K^+\pi^-$ mass is closest to $M(K^{*0})$. This results in the correct mass assignments for about 92% of the decays.

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

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

- Candidates with a dimuon mass near the J/ψ and ψ' are rejected: $8.68 < M^2(\mu^+\mu^-) < 10.09\text{ GeV}^2/c^4$ and $12.86 < M^2(\mu^+\mu^-) < 14.18\text{ GeV}^2/c^4$.
- We reject candidates consistent within originating from a $H_b \rightarrow J/\psi(')h$ decay followed by the decay of the $J/\psi(')h$ into two muons and a photon: $|(M(\mu\mu h) - M_{H_b}^{\text{PDG}}) - (M(\mu\mu) - M_{J/\psi(')}^{\text{PDG}})| > 100\text{ MeV}/c^2$, where $M(\mu\mu) <$

$M_{J/\psi(\prime)}^{\text{PDG}}$.

- Candidates with an opposite-sign hadron-muon combination (with the muon mass assigned to both particles) within $40 \text{ MeV}/c^2$ of the J/ψ or ψ' mass are rejected. This removes $H_b \rightarrow J/\psi h$ and $H_b \rightarrow \psi' h$ decays where one of the hadrons is misidentified as a muon.

In addition to the common peaking backgrounds above, we find sizable cross-feed between $B^0 \rightarrow K_S^0 \mu^+ \mu^-$ and $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$. To eliminate the cross-feed, we introduce Armenteros-Podolanski parameterization [22]. We utilize the longitudinal momentum asymmetry $\alpha \equiv (q_L^+ - q_L^-)/(q_L^+ + q_L^-)$, where $q_L^{+(-)}$ is the longitudinal momentum of the positive (negative) daughter tracks relative to the direction of the V^0 . We reject candidates satisfy $-0.26 < -1.9M(\pi\pi) + |\alpha| < -0.15$ for $B^0 \rightarrow K_S^0 \mu^+ \mu^-$ and $4.73 < 3.6M(p\pi) + |\alpha| < 4.78$ for $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ where the masses are measured in GeV/c^2 . With the cut we remove 90% (76%) of the cross-feed while the signal loss is 7% (11%) for $B^0 \rightarrow K_S^0 \mu^+ \mu^-$ ($\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$).

After the loose selections described above, rare decay channels are tightly selected with a multivariate analysis technique using the NeuroBayes [23] artificial Neural Network (NN) package. To describe the true signal, we use a PYTHIA Monte Carlo (MC) simulation. Some distributions of the simulated signal, $p_T(B)$, isolation [24], and the energy depositions of muon candidates at the electromagnetic and hadron calorimeters, are weighted by a series of scale factors to ensure that the distributions from the corresponding control channel match those of the simulation. The background events are sampled from the H_b invariant mass sideband. Because $B \rightarrow K^{(*)} \mu^+ \mu^-$ candidates have significant physics backgrounds in the lower H_b mass sideband region, we take only the higher side (100 to $300 \text{ MeV}/c^2$ higher than H_b PDG mass). $B_s^0 \rightarrow \phi \mu^+ \mu^-$ and $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ upper both sidebands of the same size.

We choose an optimized NN cut value that maximizes both branching ratio and A_{FB} significance. We minimize $N_s/\sqrt{N_s + N_b}$, where N_s is the estimate of the expected yield and N_b is the expected background. The expected N_s is determined from the world averaged branching ratio of rare channels, the observed yield of control channels, and the relative efficiencies described later. The N_b is determined from the number of loosely selected events that is dominated by combinatorial background.

IV. YIELD FIT AND BRANCHING RATIO MEASUREMENT

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

$$\mathcal{L} = \prod (f_{\text{sig}} \mathcal{P}_{\text{sig}} + (1 - f_{\text{sig}}) \mathcal{P}_{\text{bg}}), \quad (2)$$

where f_{sig} is the signal fraction, \mathcal{P}_{sig} is the signal PDF parametrized with two Gaussians having the different means, and \mathcal{P}_{bg} is the background PDF that is described by a first or second order polynomial. The signal shape is determined from the signal MC but the H_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 H_b decays and cross-talk among rare decays. Although contributions from charmless H_b decays are negligible due to muon identification, we find cross-talk between $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ and $B_s^0 \rightarrow \phi \mu^+ \mu^-$ to contribute $\sim 1\%$ to the observed signal MC yields. These contributions, whose fractions are determined by simulation, are subtracted from the fit results for the signal yields.

The fit regions are from $5.18 \text{ GeV}/c^2$ to $5.70 \text{ GeV}/c^2$ for $B \rightarrow K^{(*)} \mu^+ \mu^-$, from $5.00 \text{ GeV}/c^2$ to $5.70 \text{ GeV}/c^2$ for $B_s^0 \rightarrow \phi \mu^+ \mu^-$, and from $5.30 \text{ GeV}/c^2$ to $6.00 \text{ GeV}/c^2$ for $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$. During the yield fit, floating parameters are f_{sig} , H_b mean mass, and the background shape. We fix the H_b mass width for the rare decays while it is floated for the control samples. The mass distribution for the rare mode is determined from the control mode. Besides the statistical uncertainty derived from f_{sig} , we take into account further uncertainty $\sqrt{N_{\text{tot}}}$ coming from fluctuation of total number of events where N_{tot} is total number of events in the fit region. Fig. 1 shows H_b mass plot for each rare decay, and the yields are listed in Tab. IV.

Statistical significance s is defined as a difference of log likelihood from null hypothesis: $s \equiv \sqrt{-2 \ln(\mathcal{L}_{\text{null}}/\mathcal{L}_{\text{max}})}$, where $\mathcal{L}_{\text{null}}$ is the maximum log likelihood obtained from a fit with $f_{\text{sig}} = 0$ condition. We fix the H_b mean mass to the fitted value in the corresponding control channel. The observed yields and s are summarized in Table IV.

The relative branching ratio to the corresponding control channel is:

$$\frac{\mathcal{B}(H_b \rightarrow h \mu^+ \mu^-)}{\mathcal{B}(H_b \rightarrow J/\psi h)} = \frac{N_{h\mu^+\mu^-}^{\text{NN}}}{N_{J/\psi h}^{\text{loose}}} \frac{\epsilon_{J/\psi h}^{\text{loose}}}{\epsilon_{h\mu^+\mu^-}^{\text{loose}}} \frac{1}{\epsilon_{h\mu^+\mu^-}^{\text{NN}}} \times \mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-), \quad (3)$$

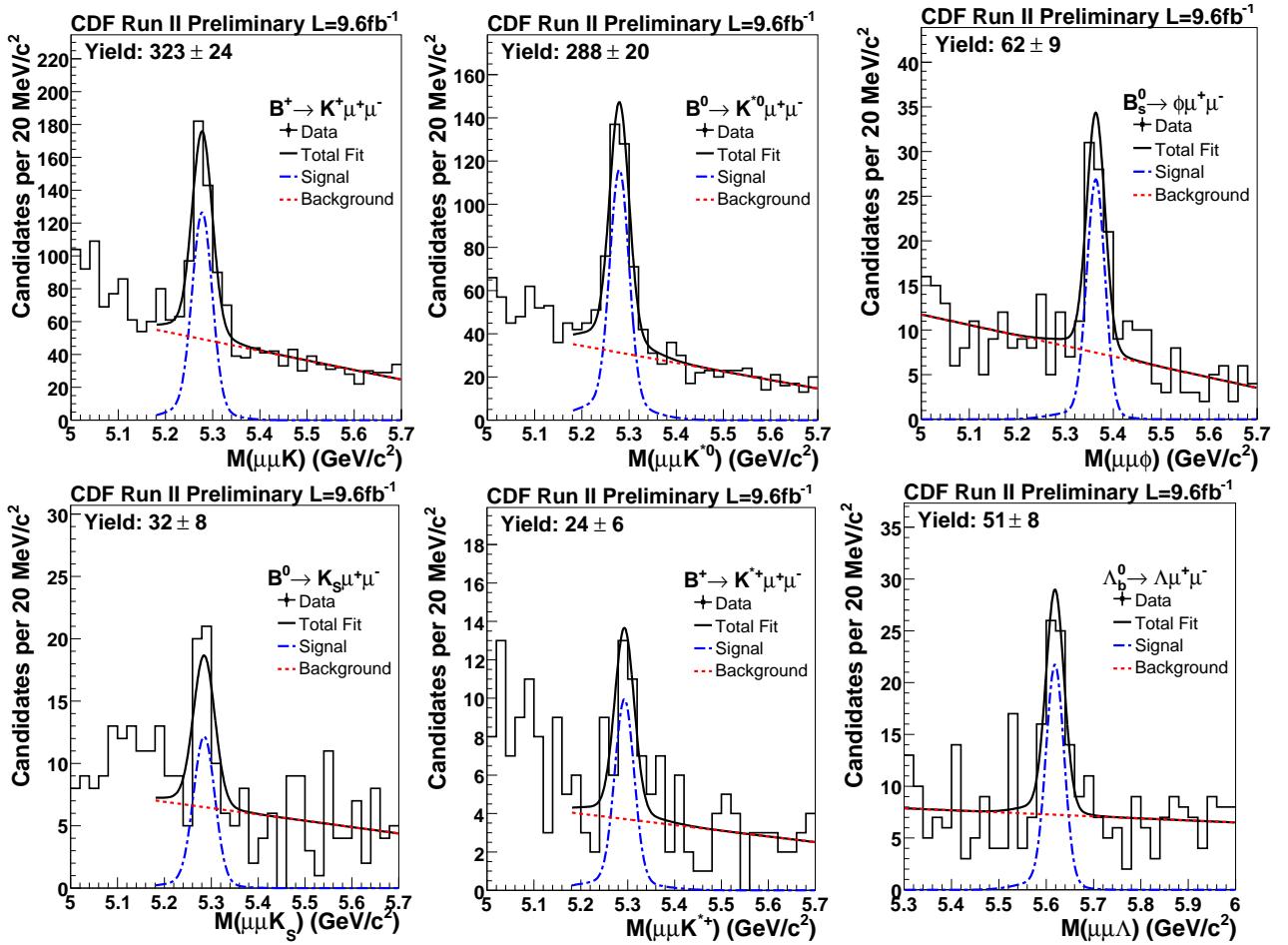


FIG. 1: (From top-left to bottom-right) The H_b invariant mass of $B^+ \rightarrow K^+ \mu^+ \mu^-$, $B^0 \rightarrow K^{*0} \mu^+ \mu^-$, $B_s^0 \rightarrow \phi \mu^+ \mu^-$, $B^0 \rightarrow K_S^0 \mu^+ \mu^-$, $B^+ \rightarrow K^{*+} \mu^+ \mu^-$, and $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$, for 9.6 fb^{-1} , respectively.

Mode	$N_{h\mu^+\mu^-}$	$s (\sigma)$	$N_{J/\psi h}$	$\epsilon_{J/\psi h}^{\text{loose}} / \epsilon_{h\mu^+\mu^-}^{\text{loose}}$	$\epsilon_{h\mu^+\mu^-}^{\text{NN}}$
$B^+ \rightarrow K^+ \mu^+ \mu^-$	319 ± 23 (490)	15.6	101900 ± 300 (118200)	0.81 ± 0.01	0.522 ± 0.001
$B^0 \rightarrow K^{*0} \mu^+ \mu^-$	288 ± 20 (234)	16.8	39900 ± 300 (56700)	0.77 ± 0.02	0.655 ± 0.002
$B_s^0 \rightarrow \phi \mu^+ \mu^-$	62 ± 9 (66)	8.9	6570 ± 80 (8000)	0.80 ± 0.01	0.770 ± 0.002
$B^0 \rightarrow K_S^0 \mu^+ \mu^-$	32 ± 8 (63)	4.6	13200 ± 100 (15000)	0.83 ± 0.01	0.454 ± 0.001
$B^+ \rightarrow K^{*+} \mu^+ \mu^-$	24 ± 6 (36)	4.2	6330 ± 90 (8000)	0.74 ± 0.02	0.483 ± 0.002
$\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$	51 ± 7 (34)	7.6	2450 ± 60 (3640)	0.84 ± 0.01	0.533 ± 0.002

TABLE I: Summary of observed yields and efficiencies. The numbers in parentheses are the number of events in the signal region. s is the statistical significance.

where $N_{h\mu^+\mu^-}^{\text{NN}}$ ($N_{J/\psi h}^{\text{loose}}$) is $H_b \rightarrow h\mu^+\mu^-$ ($H_b \rightarrow J/\psi h$) yields after the optimal NN cut (at the loose selection), $\epsilon_{h\mu^+\mu^-}^{\text{loose}} / \epsilon_{J/\psi h}^{\text{loose}}$ is the relative efficiency at the loose selection, $\epsilon_{h\mu^+\mu^-}^{\text{NN}}$ 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 these channels have sufficient signal to background ratios with the loose selection. The relative efficiency and NN cut efficiency are obtained from MC and are listed in Tab. IV. We study the NN cut efficiency with the relevant control sample. The rare channel's NN cut efficiency is corrected by the efficiency ratio between data and MC.

A. Systematic Uncertainty for the branching ratio measurement

The following systematic errors are evaluated and are summarized in Tables II and III:

- Branching Ratios

We take the branching ratios of a H_b meson decaying to corresponding normalization mode and $J/\psi \rightarrow \mu^+ \mu^-$ from the PDG [21] and assign the systematics from these uncertainties. For $\Lambda_b^0 \rightarrow J/\psi \Lambda$, we use the latest D0's measurement [25].

- Decay Model

We evaluate the difference in relative efficiency between the default decay model for the rare modes [26] and an alternative model [2]. In case of $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$, we compare our default decay model [27] with Ref. [28].

- MC reweighting

The uncertainty of MC reweighting is studied by comparison between MC with and without reweighting.

- 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, and 3.1, GeV/c for CMU, CMX, CMUP triggered muons, respectively.

- 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 each kaon track 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 $2\% \times 10\% = 0.2\%$ [29].

- B_s^0 Lifetime Difference

We find an uncertainty of 0.2% on the relative efficiency of $B_s^0 \rightarrow \phi \mu^+ \mu^-$ decays due to the unknown fraction of short-lived CP -even state in the rare mode. Furthermore, the uncertainty on $\Delta\Gamma/\Gamma$ contributes another 0.1%.

- Polarization

For $B \rightarrow J/\psi K^*$ and $B_s^0 \rightarrow J/\psi \phi$, 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. For $\Lambda_b^0 \rightarrow J/\psi \Lambda$, polarization of Λ_b^0 is unknown. We study the variation of the acceptance by varying the J/ψ helicity from -1 to 1. We also compare the polarized MC and default phase space MC.

- Control Mode Statistics

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

- $B^+ \rightarrow J/\psi \pi^+$ Contribution

Cabibbo-suppressed $B^+ \rightarrow J/\psi \pi^+$ decay contributions are neglected in the nominal fit. Its contribution is estimated with additional signal PDF taken from simulated $B^+ \rightarrow J/\psi \pi^+$ events. We find a 0.6% contribution from the fitted yield.

- MC Statistics

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

- NN Cut

The uncertainty of the NN correction is taken from uncertainty of correction function which is determined at the optimal cut point.

- Signal PDF

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

Source	$K^+\mu^+\mu^-$	$K^{*0}\mu^+\mu^-$	$\phi\mu^+\mu^-$	$K_S^0\mu^+\mu^-$	$K^{*+}\mu^+\mu^-$	$\Lambda\mu^+\mu^-$
Theory model	0.9	0.7	1.7	0.5	0.7	3.4
MC reweight	2.6	0.1	2.5	2.5	1.8	2.9
Trigger turn-on	1.3	0.8	4.0	5.5	8.6	4.8
Particle ID	-	1.5	4.4	-	-	-
Low p_T hadrons	0.2	0.2	0.2	0.2	0.2	0.2
B_s lifetime	-	-	0.2	-	-	-
Polarization	-	0.5	0.1	-	0.5	6.6
Control mode stat.	0.4	0.7	1.4	1.0	1.7	2.6
MC stat.	0.1	0.2	0.1	0.1	0.2	0.1
NN cut	0.2	0.2	0.3	0.2	0.6	0.5
Efficiency total	3.0	2.1	6.8	6.1	9.0	9.7

TABLE II: Efficiency systematic errors as percentage contributions.

Source	$K^+\mu^+\mu^-$	$K^{*0}\mu^+\mu^-$	$\phi\mu^+\mu^-$	$K_S^0\mu^+\mu^-$	$K^{*+}\mu^+\mu^-$	$\Lambda\mu^+\mu^-$
Efficiency	3.0	2.1	6.8	6.1	9.0	9.7
$\mathcal{B}(J/\psi \rightarrow \mu^+\mu^-)$	1.0	1.0	1.0	1.0	1.0	1.0
Signal PDF	0.2	0.3	0.4	0.1	0.2	0.1
Background PDF	2.5	1.3	3.0	2.6	0.8	0.1
Peaking BG	0.5	1.3	1.1	0.3	0.6	0.3
$\mathcal{B}(H_b \rightarrow J/\psi h)$	3.4	4.5	30.8	3.7	5.6	29.7
Total	5.3	5.4	31.7	7.7	10.7	31.3

TABLE III: Total systematic errors as percentage contributions.

- Background PDF

The uncertainty of different background PDF assumptions are evaluated by switching a linear shape assumption with an exponential function.

- Peaking Background

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

B. Relative and absolute branching ratio

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

$$\begin{aligned}
 \mathcal{B}(B^+ \rightarrow K^+\mu^+\mu^-)/\mathcal{B}(B^+ \rightarrow J/\psi K^+) &= [0.44 \pm 0.03(\text{stat}) \pm 0.02(\text{syst})] \times 10^{-3}, \\
 \mathcal{B}(B^0 \rightarrow K^{*0}\mu^+\mu^-)/\mathcal{B}(B^0 \rightarrow J/\psi K^{*0}) &= [0.85 \pm 0.07(\text{stat}) \pm 0.03(\text{syst})] \times 10^{-3}, \\
 \mathcal{B}(B_s^0 \rightarrow \phi\mu^+\mu^-)/\mathcal{B}(B_s^0 \rightarrow J/\psi \phi) &= [0.90 \pm 0.14(\text{stat}) \pm 0.07(\text{syst})] \times 10^{-3}, \\
 \mathcal{B}(B^0 \rightarrow K^0\mu^+\mu^-)/\mathcal{B}(B^0 \rightarrow J/\psi K^0) &= [0.38 \pm 0.10(\text{stat}) \pm 0.03(\text{syst})] \times 10^{-3}, \\
 \mathcal{B}(B^+ \rightarrow K^{*+}\mu^+\mu^-)/\mathcal{B}(B^+ \rightarrow J/\psi K^{*+}) &= [0.62 \pm 0.18(\text{stat}) \pm 0.06(\text{syst})] \times 10^{-3}, \\
 \mathcal{B}(\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-)/\mathcal{B}(\Lambda_b^0 \rightarrow J/\psi \Lambda) &= [2.75 \pm 0.48(\text{stat}) \pm 0.27(\text{syst})] \times 10^{-3}.
 \end{aligned}$$

The absolute branching ratio is obtained by replacing the control channel's branching ratio with the corresponding

PDG [21] value or the latest measurement [25]:

$$\begin{aligned}
\mathcal{B}(B^+ \rightarrow K^+ \mu^+ \mu^-) &= [0.45 \pm 0.03(\text{stat}) \pm 0.02(\text{syst})] \times 10^{-6}, \\
\mathcal{B}(B^0 \rightarrow K^{*0} \mu^+ \mu^-) &= [1.14 \pm 0.09(\text{stat}) \pm 0.06(\text{syst})] \times 10^{-6}, \\
\mathcal{B}(B_s^0 \rightarrow \phi \mu^+ \mu^-) &= [1.17 \pm 0.18(\text{stat}) \pm 0.37(\text{syst})] \times 10^{-6}, \\
\mathcal{B}(B^0 \rightarrow K^0 \mu^+ \mu^-) &= [0.33 \pm 0.08(\text{stat}) \pm 0.03(\text{syst})] \times 10^{-6}, \\
\mathcal{B}(B^+ \rightarrow K^{*+} \mu^+ \mu^-) &= [0.89 \pm 0.25(\text{stat}) \pm 0.09(\text{syst})] \times 10^{-6}, \\
\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-) &= [1.95 \pm 0.34(\text{stat}) \pm 0.61(\text{syst})] \times 10^{-6}.
\end{aligned}$$

We also calculate the combined branching ratio assuming isospin symmetry:

$$\begin{aligned}
\mathcal{B}(B \rightarrow K \mu^+ \mu^-) &= [0.40 \pm 0.03(\text{stat}) \pm 0.02(\text{syst})] \times 10^{-6}, \\
\mathcal{B}(B \rightarrow K^* \mu^+ \mu^-) &= [1.10 \pm 0.08(\text{stat}) \pm 0.06(\text{syst})] \times 10^{-6},
\end{aligned}$$

We take the lifetime ratio of B^+ meson to B^0 meson, $\tau_{B^+}/\tau_{B^0} = 1.071 \pm 0.009$ from the PDG [21] and assume equal production of B^0 and B^+ .

These numbers are consistent with our previous results [7] and B-factory measurements [6, 17].

V. DIFFERENTIAL BRANCHING RATIO

We measure the differential branching ratio with respect to the dimuon mass. The signal region is divided into six exclusive and two additional q^2 bins, where $q^2 \equiv M_{\mu\mu}^2 c^2$. Since each q^2 bin has a different amount of signal and background, we fit for the signal in each q^2 bin with the same procedure used for the global fits. During the fit, we fix the mean of the H_b mass and BG slope to the value obtained from the global fit, therefore only f_{sig} is floated. Fig. 2 and Tables IV-VIII show the differential branching ratios for $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$, $B_s^0 \rightarrow \phi \mu^+ \mu^-$, $B^+ \rightarrow K^+ \mu^+ \mu^-$, $B^0 \rightarrow K^0 \mu^+ \mu^-$, combined $B \rightarrow K \mu^+ \mu^-$, $B^0 \rightarrow K^{*0} \mu^+ \mu^-$, $B^+ \rightarrow K^{*+} \mu^+ \mu^-$, and combined $B \rightarrow K^* \mu^+ \mu^-$.

We also calculate the isospin asymmetry A_I between neutral and charged B mesons:

$$A_I^{(*)} \equiv \frac{\mathcal{B}(B^0 \rightarrow K^{(*)0} \mu^+ \mu^-) - \mathcal{B}(B^+ \rightarrow K^{(*)+} \mu^+ \mu^-)}{\mathcal{B}(B^0 \rightarrow K^{(*)0} \mu^+ \mu^-) + \mathcal{B}(B^+ \rightarrow K^{(*)+} \mu^+ \mu^-)}, \quad (4)$$

with respect to q^2 . Tables X and XI show the combined $K \mu^+ \mu^-$ and $K^* \mu^+ \mu^-$ differential branching ratios, and Table XII and Fig. 3 show the isospin asymmetry between $B^+ \rightarrow K^+ \mu^+ \mu^-$ ($B^0 \rightarrow K^{*0} \mu^+ \mu^-$) and $B^0 \rightarrow K^0 \mu^+ \mu^-$ ($B^+ \rightarrow K^{*+} \mu^+ \mu^-$).

q^2 range	N_s	$d\mathcal{B}/dq^2 [10^{-8}(\text{GeV}^2/c^2)^{-1}]$	$\mathcal{B}(10^{-7})$
[0.00, 2.00)	22.4 ± 6.3	$1.80 \pm 0.53 \pm 0.12$	$0.36 \pm 0.11 \pm 0.02$
[2.00, 4.30)	53.0 ± 8.1	$3.16 \pm 0.54 \pm 0.18$	$0.73 \pm 0.12 \pm 0.04$
[4.30, 8.68)	109.2 ± 12.6	$3.05 \pm 0.38 \pm 0.18$	$1.34 \pm 0.17 \pm 0.08$
[10.09, 12.86)	47.8 ± 8.7	$1.86 \pm 0.36 \pm 0.11$	$0.51 \pm 0.10 \pm 0.03$
[14.18, 16.00)	42.7 ± 6.7	$2.33 \pm 0.42 \pm 0.13$	$0.42 \pm 0.08 \pm 0.02$
[16.00, 23.00)	45.5 ± 8.0	$0.72 \pm 0.14 \pm 0.04$	$0.50 \pm 0.10 \pm 0.03$
[0.00, 4.30)	74.7 ± 10.3	$2.55 \pm 0.38 \pm 0.15$	$1.10 \pm 0.16 \pm 0.06$
[1.00, 6.00)	97.9 ± 11.7	$2.69 \pm 0.35 \pm 0.16$	$1.35 \pm 0.18 \pm 0.08$

TABLE IV: Differential branching ratio of $B^+ \rightarrow K^+ \mu^+ \mu^-$. First (second) error is statistical (systematic).

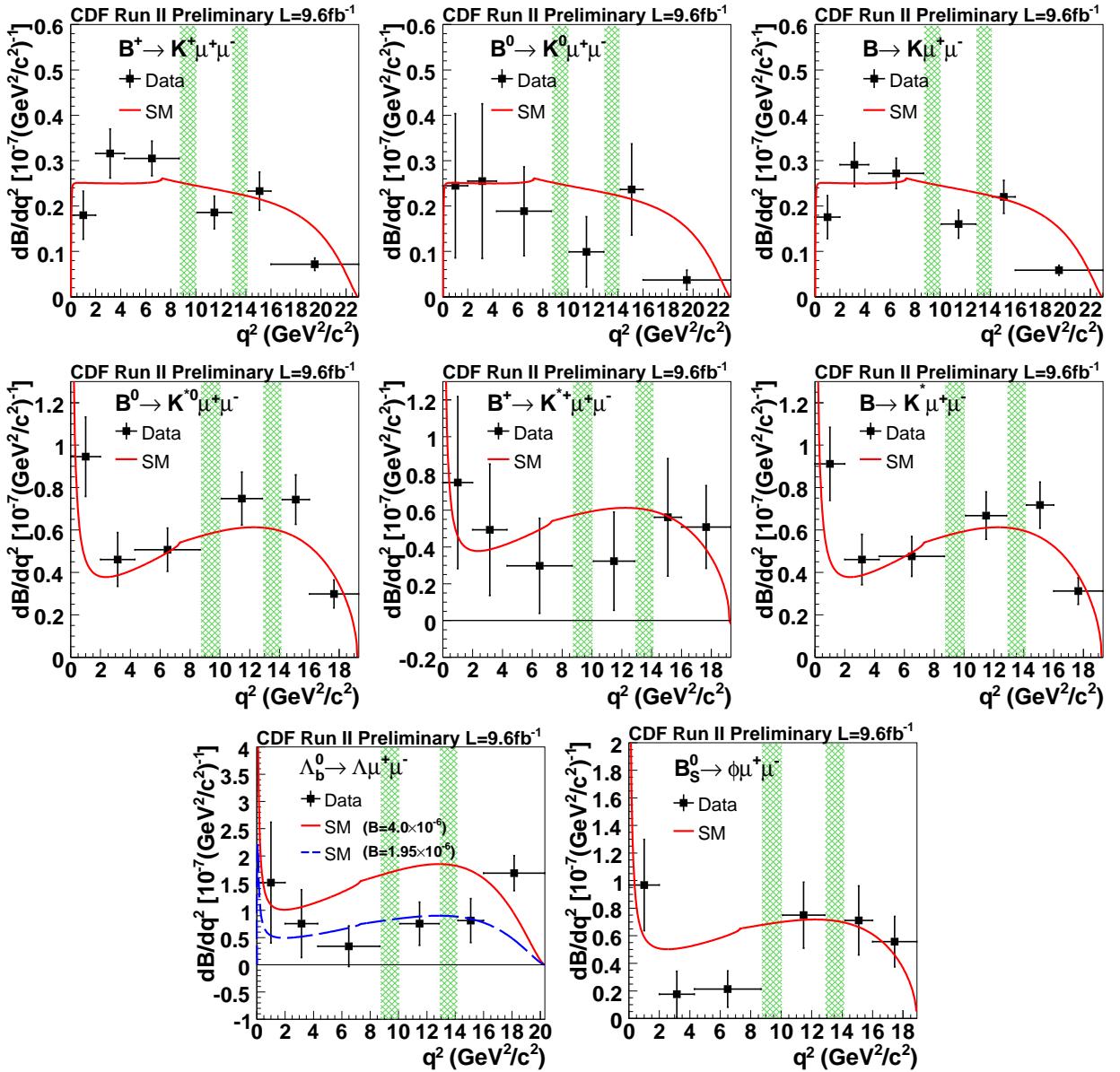


FIG. 2: Differential branching ratio of $B^+ \rightarrow K^+ \mu^+ \mu^-$, $B^0 \rightarrow K^0 \mu^+ \mu^-$, $B \rightarrow K^* \mu^+ \mu^-$, $B^0 \rightarrow K^{*0} \mu^+ \mu^-$, $B^+ \rightarrow K^{*+} \mu^+ \mu^-$, and $B \rightarrow K \mu^+ \mu^-$, $B_s^0 \rightarrow \phi \mu^+ \mu^-$, and $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$, respectively (top-left to bottom-right). Statistical error only is shown. Hatched regions are charmonium veto regions. The solid line is the SM expectation [4, 26, 27]. The long dashed line in $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ plot is the SM expectation scaled to our total branching ratio measurement.

q^2 range	N_s	$d\mathcal{B}/dq^2 [10^{-8} (\text{GeV}^2/\text{c}^2)^{-1}]$	$\mathcal{B}(10^{-7})$
[0.00, 2.00)	39.4 ± 6.8	$9.46 \pm 1.88 \pm 0.55$	$1.89 \pm 0.38 \pm 0.11$
[2.00, 4.30)	24.7 ± 6.3	$4.61 \pm 1.27 \pm 0.30$	$1.06 \pm 0.29 \pm 0.07$
[4.30, 8.68)	61.7 ± 11.9	$5.07 \pm 1.02 \pm 0.37$	$2.22 \pm 0.45 \pm 0.16$
[10.09, 12.86)	73.1 ± 11.2	$7.48 \pm 1.25 \pm 0.46$	$2.07 \pm 0.35 \pm 0.13$
[14.18, 16.00)	54.2 ± 7.1	$7.43 \pm 1.17 \pm 0.42$	$1.35 \pm 0.21 \pm 0.08$
[16.00, 19.30)	36.9 ± 7.4	$2.98 \pm 0.66 \pm 0.18$	$0.98 \pm 0.22 \pm 0.06$
[0.00, 4.30)	64.1 ± 9.3	$6.87 \pm 1.11 \pm 0.41$	$2.95 \pm 0.48 \pm 0.18$
[1.00, 6.00)	45.6 ± 10.9	$3.94 \pm 0.97 \pm 0.28$	$1.97 \pm 0.49 \pm 0.14$

TABLE V: Differential branching ratio of $B^0 \rightarrow K^{*0} \mu^+ \mu^-$. First (second) error is statistical (systematic).

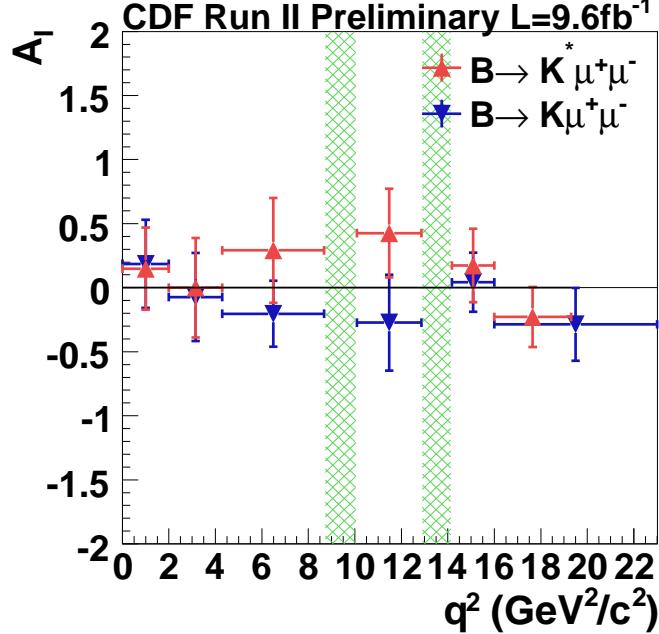


FIG. 3: Isospin assymetry of $B \rightarrow K\mu^+\mu^-$ and $B \rightarrow K^*\mu^+\mu^-$. Statistical error is only shown. Hatched regions are charmonium veto regions and upper (lower) triangle shows $B \rightarrow K\mu^+\mu^-$ ($B \rightarrow K^*\mu^+\mu^-$).

q^2 range	N_s	$d\mathcal{B}/dq^2 [10^{-8}(\text{GeV}^2/\text{c}^2)^{-1}]$	$\mathcal{B}(10^{-7})$
[0.00, 2.00)	13.1 ± 3.1	$15.80 \pm 4.59 \pm 5.01$	$3.16 \pm 0.92 \pm 1.00$
[2.00, 4.30)	1.4 ± 2.1	$1.18 \pm 1.79 \pm 0.37$	$0.27 \pm 0.41 \pm 0.09$
[4.30, 8.68)	3.6 ± 3.8	$1.47 \pm 1.55 \pm 0.46$	$0.64 \pm 0.68 \pm 0.20$
[10.09, 12.86)	16.1 ± 4.6	$8.12 \pm 2.47 \pm 2.57$	$2.25 \pm 0.69 \pm 0.71$
[14.18, 16.00)	9.0 ± 3.1	$6.12 \pm 2.33 \pm 1.94$	$1.11 \pm 0.42 \pm 0.35$
[16.00, 18.90)	17.4 ± 3.5	$7.95 \pm 2.03 \pm 2.52$	$2.31 \pm 0.59 \pm 0.73$
[0.00, 4.30)	14.9 ± 4.0	$7.62 \pm 2.27 \pm 2.42$	$3.28 \pm 0.98 \pm 1.04$
[1.00, 6.00)	5.2 ± 3.5	$2.06 \pm 1.40 \pm 0.65$	$1.03 \pm 0.70 \pm 0.33$

TABLE VI: Differential branching ratio of $B_s^0 \rightarrow \phi\mu^+\mu^-$. First (second) error is statistical (systematic).

q^2 range	N_s	$d\mathcal{B}/dq^2 [10^{-8}(\text{GeV}^2/\text{c}^2)^{-1}]$	$\mathcal{B}(10^{-7})$
[0.00, 2.00)	2.9 ± 1.6	$2.45 \pm 1.59 \pm 0.21$	$0.49 \pm 0.32 \pm 0.04$
[2.00, 4.30)	4.8 ± 3.1	$2.55 \pm 1.70 \pm 0.35$	$0.59 \pm 0.39 \pm 0.08$
[4.30, 8.68)	8.4 ± 4.2	$1.89 \pm 0.98 \pm 0.28$	$0.83 \pm 0.43 \pm 0.12$
[10.09, 12.86)	3.8 ± 2.9	$0.99 \pm 0.78 \pm 0.15$	$0.28 \pm 0.22 \pm 0.04$
[14.18, 16.00)	6.7 ± 2.4	$2.36 \pm 1.01 \pm 0.19$	$0.43 \pm 0.18 \pm 0.04$
[16.00, 23.00)	5.2 ± 2.9	$0.37 \pm 0.22 \pm 0.04$	$0.26 \pm 0.15 \pm 0.03$
[0.00, 4.30)	8.0 ± 3.5	$2.60 \pm 1.21 \pm 0.28$	$1.12 \pm 0.52 \pm 0.12$
[1.00, 6.00)	9.1 ± 4.1	$2.22 \pm 1.04 \pm 0.28$	$1.11 \pm 0.52 \pm 0.14$

TABLE VII: Differential branching ratio of $B^0 \rightarrow K^0\mu^+\mu^-$. First (second) error is statistical (systematic).

q^2 range	N_s	$d\mathcal{B}/dq^2[10^{-8}(\text{GeV}^2/c^2)^{-1}]$	$\mathcal{B}(10^{-7})$
[0.00, 2.00)	3.1 ± 1.5	$7.50 \pm 4.68 \pm 0.88$	$1.50 \pm 0.94 \pm 0.18$
[2.00, 4.30)	2.7 ± 1.8	$4.94 \pm 3.58 \pm 0.63$	$1.14 \pm 0.82 \pm 0.15$
[4.30, 8.68)	3.9 ± 3.3	$2.98 \pm 2.59 \pm 0.84$	$1.30 \pm 1.13 \pm 0.37$
[10.09, 12.86)	3.6 ± 2.9	$3.23 \pm 2.67 \pm 0.73$	$0.89 \pm 0.74 \pm 0.20$
[14.18, 16.00)	4.3 ± 2.1	$5.61 \pm 3.20 \pm 0.71$	$1.02 \pm 0.58 \pm 0.13$
[16.00, 19.30)	6.3 ± 2.2	$5.09 \pm 2.25 \pm 0.57$	$1.68 \pm 0.74 \pm 0.19$
[0.00, 4.30)	5.7 ± 2.3	$6.06 \pm 2.91 \pm 0.74$	$2.61 \pm 1.25 \pm 0.32$
[1.00, 6.00)	8.4 ± 2.8	$7.12 \pm 2.76 \pm 0.85$	$3.56 \pm 1.38 \pm 0.43$

TABLE VIII: Differential branching ratio of $B^+ \rightarrow K^{*+}\mu^+\mu^-$. First (second) error is statistical (systematic).

q^2 range	N_s	$d\mathcal{B}/dq^2[10^{-8}(\text{GeV}^2/c^2)^{-1}]$	$\mathcal{B}(10^{-7})$
[0.00, 2.00)	3.2 ± 2.3	$15.10 \pm 11.10 \pm 4.73$	$3.02 \pm 2.22 \pm 0.95$
[2.00, 4.30)	2.7 ± 2.2	$7.56 \pm 6.23 \pm 2.37$	$1.74 \pm 1.43 \pm 0.54$
[4.30, 8.68)	3.2 ± 3.5	$3.42 \pm 3.75 \pm 1.08$	$1.50 \pm 1.64 \pm 0.47$
[10.09, 12.86)	7.2 ± 3.7	$7.55 \pm 3.96 \pm 2.36$	$2.09 \pm 1.10 \pm 0.65$
[14.18, 16.00)	6.2 ± 2.9	$8.13 \pm 4.04 \pm 2.54$	$1.48 \pm 0.74 \pm 0.46$
[16.00, 20.30)	29.2 ± 4.0	$16.83 \pm 3.22 \pm 5.27$	$7.24 \pm 1.39 \pm 2.26$
[0.00, 4.30)	6.0 ± 3.2	$10.45 \pm 5.76 \pm 3.27$	$4.50 \pm 2.48 \pm 1.41$
[1.00, 6.00)	5.2 ± 3.3	$6.34 \pm 4.14 \pm 1.99$	$3.17 \pm 2.07 \pm 0.99$

TABLE IX: Differential branching ratio of $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$. First (second) error is statistical (systematic).

q^2 range	$d\mathcal{B}/dq^2[10^{-8}(\text{GeV}^2/c^2)^{-1}]$	$\mathcal{B}(10^{-7})$
[0.00, 2.00)	$1.76 \pm 0.47 \pm 0.10$	$0.35 \pm 0.09 \pm 0.02$
[2.00, 4.30)	$2.91 \pm 0.48 \pm 0.16$	$0.67 \pm 0.11 \pm 0.04$
[4.30, 8.68)	$2.72 \pm 0.34 \pm 0.15$	$1.19 \pm 0.15 \pm 0.07$
[10.09, 12.86)	$1.60 \pm 0.31 \pm 0.09$	$0.44 \pm 0.09 \pm 0.03$
[14.18, 16.00)	$2.20 \pm 0.37 \pm 0.11$	$0.40 \pm 0.07 \pm 0.02$
[16.00, 23.00)	$0.59 \pm 0.11 \pm 0.03$	$0.41 \pm 0.08 \pm 0.02$
[0.00, 4.30)	$2.40 \pm 0.34 \pm 0.13$	$1.03 \pm 0.15 \pm 0.06$
[1.00, 6.00)	$2.48 \pm 0.31 \pm 0.14$	$1.24 \pm 0.16 \pm 0.07$

TABLE X: Combined differential branching ratio of $B \rightarrow K\mu^+\mu^-$. First (second) error is statistical (systematic).

q^2 range	$d\mathcal{B}/dq^2[10^{-8}(\text{GeV}^2/c^2)^{-1}]$	$\mathcal{B}(10^{-7})$
[0.00, 2.00)	$9.12 \pm 1.73 \pm 0.49$	$1.82 \pm 0.35 \pm 0.10$
[2.00, 4.30)	$4.61 \pm 1.19 \pm 0.27$	$1.06 \pm 0.27 \pm 0.06$
[4.30, 8.68)	$4.76 \pm 0.94 \pm 0.33$	$2.08 \pm 0.41 \pm 0.15$
[10.09, 12.86)	$6.68 \pm 1.12 \pm 0.40$	$1.85 \pm 0.31 \pm 0.11$
[14.18, 16.00)	$7.18 \pm 1.09 \pm 0.38$	$1.31 \pm 0.20 \pm 0.07$
[16.00, 19.30)	$3.12 \pm 0.63 \pm 0.17$	$1.03 \pm 0.21 \pm 0.06$
[0.00, 4.30)	$6.72 \pm 1.03 \pm 0.37$	$2.89 \pm 0.44 \pm 0.16$
[1.00, 6.00)	$4.24 \pm 0.91 \pm 0.27$	$2.12 \pm 0.46 \pm 0.13$

TABLE XI: Combined differential branching ratio of $B \rightarrow K^*\mu^+\mu^-$. First (second) error is statistical (systematic).

q^2 (GeV $^2/c^2$)	$A_I(B \rightarrow K\mu^+\mu^-)$	$A_I(B \rightarrow K^*\mu^+\mu^-)$
0.00-2.00	$0.19 \pm 0.34(\text{stat}) \pm 0.05(\text{syst})$	$0.15 \pm 0.32(\text{stat}) \pm 0.06(\text{syst})$
2.00-4.30	$-0.07 \pm 0.34(\text{stat}) \pm 0.07(\text{syst})$	$0.00 \pm 0.39(\text{stat}) \pm 0.07(\text{syst})$
4.30-8.68	$-0.20 \pm 0.26(\text{stat}) \pm 0.08(\text{syst})$	$0.29 \pm 0.41(\text{stat}) \pm 0.13(\text{syst})$
10.09-12.86	$-0.27 \pm 0.37(\text{stat}) \pm 0.08(\text{syst})$	$0.43 \pm 0.35(\text{stat}) \pm 0.10(\text{syst})$
14.18-16.00	$0.04 \pm 0.23(\text{stat}) \pm 0.05(\text{syst})$	$0.17 \pm 0.29(\text{stat}) \pm 0.07(\text{syst})$
16.00-23.00	$-0.29 \pm 0.28(\text{stat}) \pm 0.06(\text{syst})$	$-0.23 \pm 0.23(\text{stat}) \pm 0.06(\text{syst})$
0.00-4.30	$0.04 \pm 0.24(\text{stat}) \pm 0.06(\text{syst})$	$0.10 \pm 0.25(\text{stat}) \pm 0.07(\text{syst})$
1.00-6.00	$-0.06 \pm 0.24(\text{stat}) \pm 0.07(\text{syst})$	$-0.26 \pm 0.21(\text{stat}) \pm 0.07(\text{syst})$

TABLE XII: Isospin assymetry of $B \rightarrow K\mu^+\mu^-$ and $B \rightarrow K^*\mu^+\mu^-$

VI. ANGULAR ANALYSIS ON $B \rightarrow K^{(*)}\mu^+\mu^-$ DECAYS

The muon forward-backward asymmetry (A_{FB}) and K^* longitudinal polarization (F_L) are extracted from $\cos\theta_\mu$ and $\cos\theta_K$, respectively, where θ_μ is the helicity angle between μ^+ (μ^-) direction and the opposite of the B (\bar{B}) direction in the dimuon rest frame, and θ_K is the angle between the kaon direction and the direction opposite to the B meson in the K^* rest frame. The transverse polarization asymmetry ($A_T^{(2)}$) and the T -odd CP asymmetry (A_{im}) are extracted from ϕ , where the angle ϕ is the angle between the two decay planes of the dimuon pair and $k\pi$ decay pair. Fig. 4 shows the schematic view of the $B^0 \rightarrow K^{*0}\mu^+\mu^-$ angular distribution.

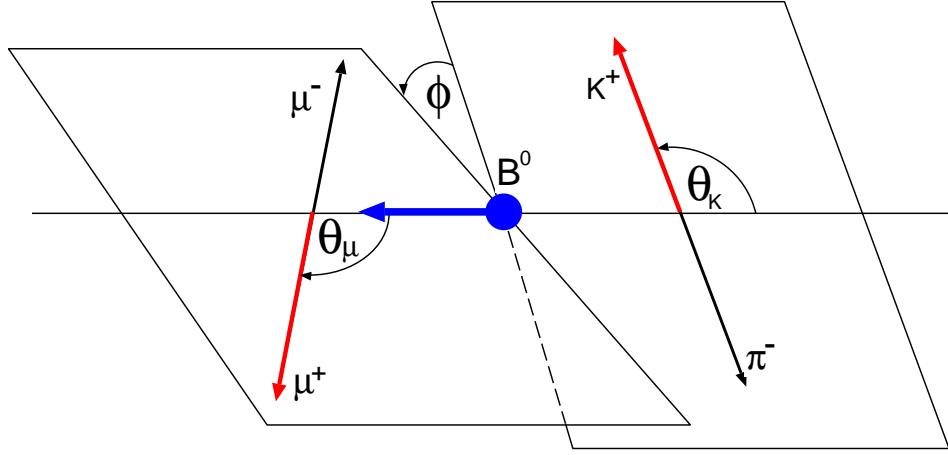


FIG. 4: Schematic view of the $B^0 \rightarrow K^{*0}\mu^+\mu^-$ angular distribution.

For $B \rightarrow K^*\mu^+\mu^-$, these angular observables can be expressed in terms of the polarization amplitudes:

$$A_{FB} \equiv \frac{3}{2} \frac{\operatorname{Re}(A_{\perp L} A_{\parallel L}^*) - \operatorname{Re}(A_{\perp R} A_{\parallel R}^*)}{|A_0|^2 + |A_\parallel|^2 + |A_\perp|^2}, \quad (5)$$

$$F_L \equiv \frac{|A_0|^2}{|A_0|^2 + |A_\parallel|^2 + |A_\perp|^2}, \quad (6)$$

$$A_T^{(2)} \equiv \frac{|A_\perp|^2 - |A_\parallel|^2}{|A_\perp|^2 + |A_\parallel|^2}, \quad (7)$$

$$A_{im} \equiv \frac{\operatorname{Im}(A_{\perp L} A_{\parallel L}^*) + \operatorname{Im}(A_{\perp R} A_{\parallel R}^*)}{|A_0|^2 + |A_\parallel|^2 + |A_\perp|^2}, \quad (8)$$

where A_0 , A_\parallel , and A_\perp are K^* 's transversity basis polarization.

We perform an unbinned maximum likelihood fit to extract the angular observables:

$$\mathcal{L} = \prod (f_{\text{sig}} \mathcal{P}_{\text{sig}}(M_B) \mathcal{F}_{\text{sig}} + (1 - f_{\text{sig}}) \mathcal{P}_{\text{bg}}(M_B) \mathcal{F}_{\text{bg}}), \quad (9)$$

where f_{sig} is the signal fraction, \mathcal{P}_{sig} (\mathcal{P}_{bg}) is the signal (background) PDF of the B mass shape and \mathcal{F}_{sig} (\mathcal{F}_{bg}) is the signal (background) PDF of the angular shape. In contrast to branching ratio measurement, we consider only the statistical uncertainty derived from f_{sig} for A_{FB} and 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 F_L [11]:

$$\frac{1}{\Gamma} \frac{d\Gamma(B \rightarrow K^*\mu^+\mu^-)}{d\cos\theta_K} = \frac{3}{2} F_L \cos^2\theta_K + \frac{3}{4} (1 - F_L)(1 - \cos^2\theta_K), \quad (10)$$

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

$$\frac{1}{\Gamma} \frac{d\Gamma(B \rightarrow K^*\mu^+\mu^-)}{d\cos\theta_\mu} = \frac{3}{4} F_L (1 - \cos^2\theta_\mu) + \frac{3}{8} (1 - F_L)(1 + \cos^2\theta_\mu) + A_{FB} \cos\theta_\mu. \quad (11)$$

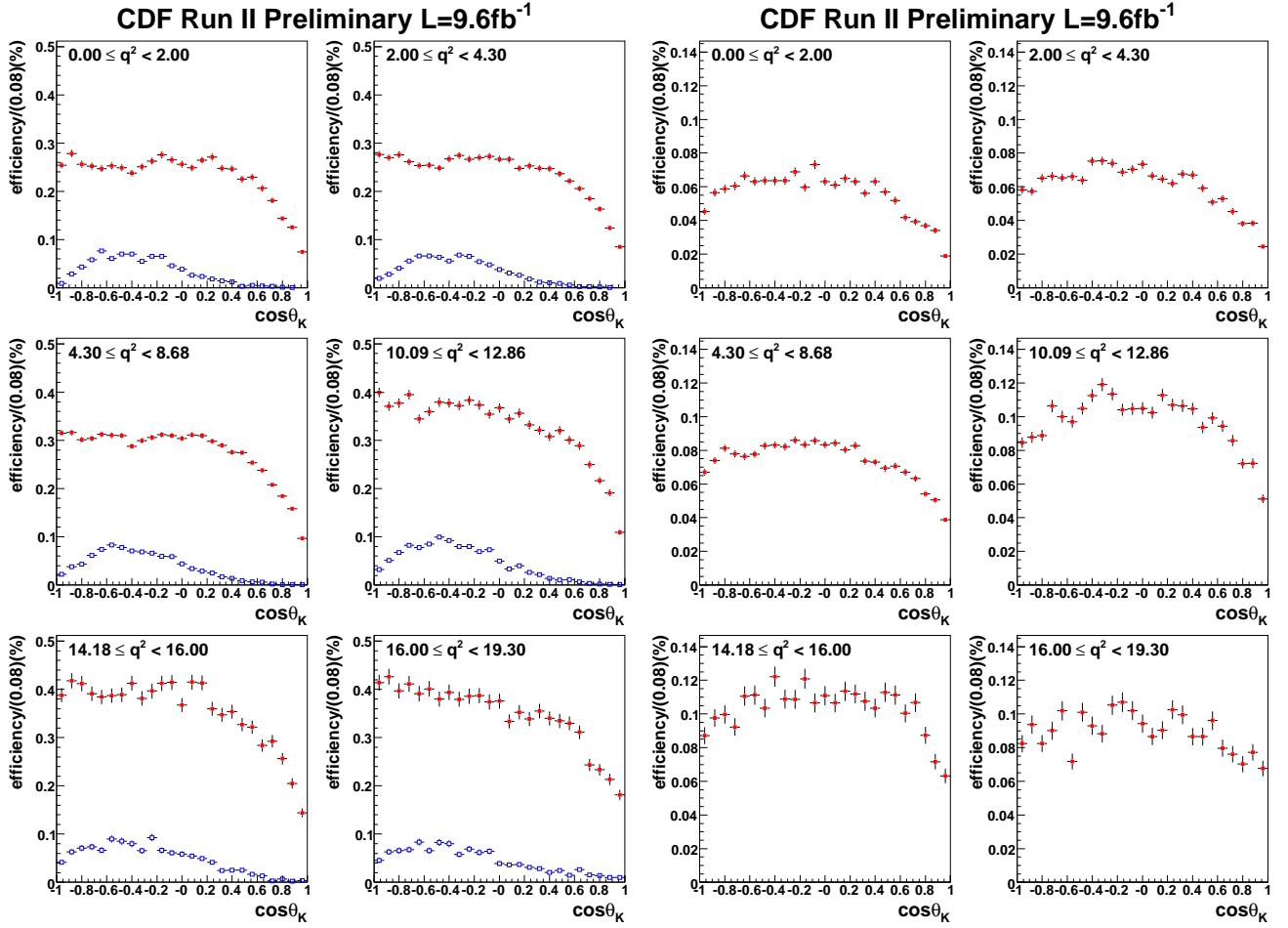


FIG. 5: $\cos \theta_K$ acceptance of $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ (left two column) and $B^+ \rightarrow K^{*+} \mu^+ \mu^-$ (right two column). Filled circle is true signal and open circle shows $K\pi$ swapped signal.

The differential decay rate in ϕ depends on $A_T^{(2)}$ and A_{im} [14]:

$$\frac{1}{\Gamma} \frac{d\Gamma(B \rightarrow K^* \mu^+ \mu^-)}{d\phi} = \frac{1}{2\pi} \left[1 + \frac{1}{2}(1 - F_L) A_T^{(2)} \cos 2\phi + A_{im} \sin 2\phi \right]. \quad (12)$$

Angular acceptances of $\cos \theta_K$, $\cos \theta_\mu$, and ϕ are considered as a 25-bin histograms which are obtained from phase space signal MC. A sizable contribution to the signal which consists of $K\pi$ swapped K^{*0} is handled as a fit bias. Since a $K\pi$ swap inverts the sign of $\cos \theta_\mu$, the fraction works as a dilution to the asymmetry. To model this effect in the fit we add an additional signal-like term that considers opposite $\cos \theta_\mu$ sign. Because we find that the other angular acceptances are also affected, similar PDF's are added to the acceptance functions. Figs. 5-7 show the angular acceptance for each q^2 bin.

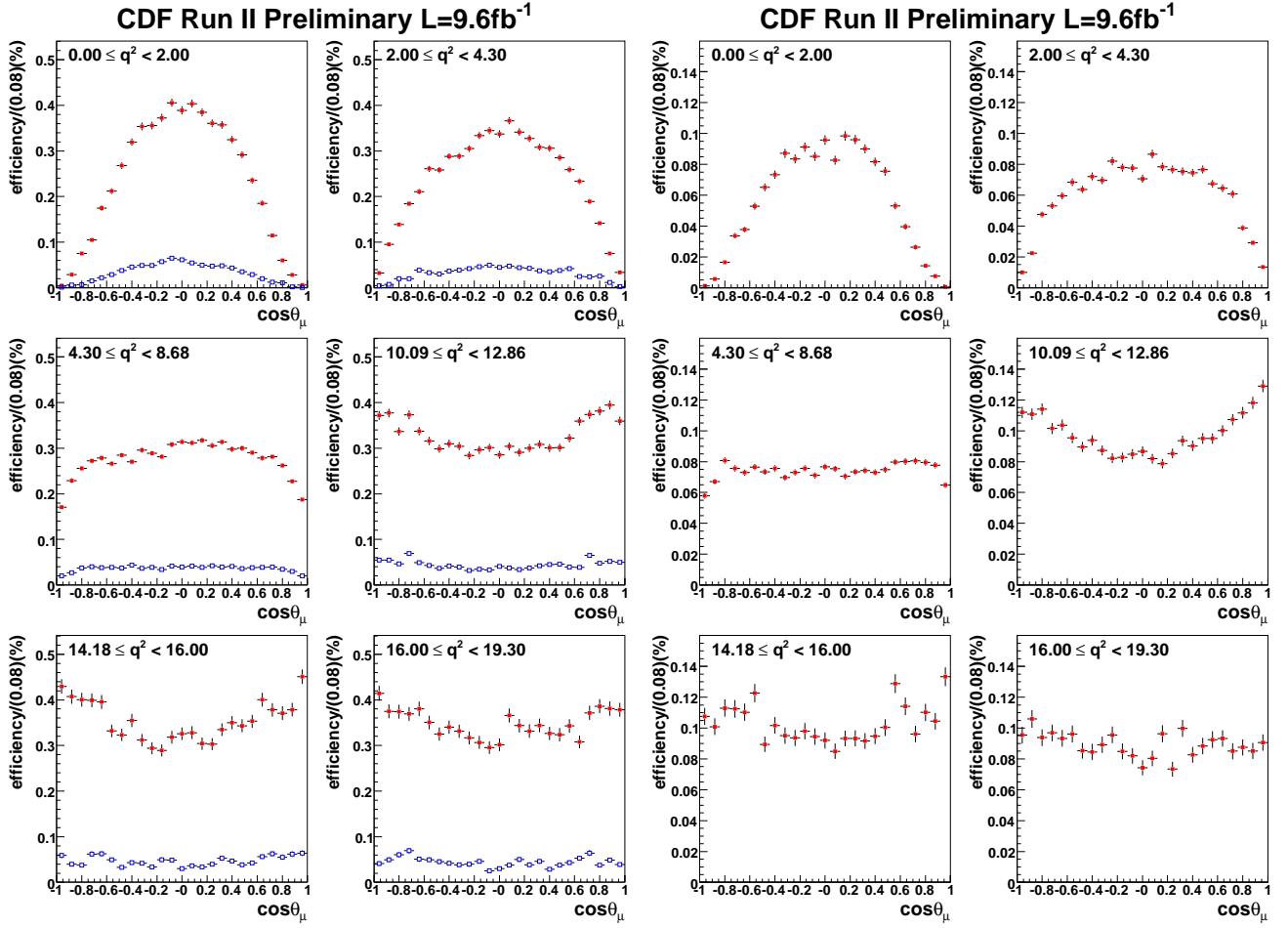


FIG. 6: $\cos \theta_\mu$ acceptance of $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ (left two column) and $B^+ \rightarrow K^{*+} \mu^+ \mu^-$ (right two column). Filled circle is true signal and open circle shows $K\pi$ swapped signal.

The signal PDF are described as:

$$\begin{aligned} \mathcal{F}_{\text{sig}}^K(\cos \theta_K) &\equiv (1 - f_{\text{swap}}) \left[\frac{3}{2} F_L \cos^2 \theta_K + \frac{3}{4} (1 - F_L)(1 - \cos^2 \theta_K) \right] \times \varepsilon^K(\cos \theta_K) \\ &\quad + f_{\text{swap}} \left[\frac{3}{2} F_L \cos^2 \theta_K + \frac{3}{4} (1 - F_L)(1 - \cos^2 \theta_K) \right] \times \varepsilon_{\text{swap}}^K(\cos \theta_K), \end{aligned} \quad (13)$$

$$\begin{aligned} \mathcal{F}_{\text{sig}}^\mu(\cos \theta_\mu) &\equiv (1 - f_{\text{swap}}) \left[\frac{3}{4} F_L (1 - \cos^2 \theta_\mu) + \frac{3}{8} (1 - F_L)(1 + \cos^2 \theta_\mu) \right. \\ &\quad \left. + A_{FB} \cos \theta_\mu \right] \times \varepsilon^\mu(\cos \theta_\mu) \\ &\quad + f_{\text{swap}} \left[\frac{3}{4} F_L (1 - \cos^2 \theta_\mu) + \frac{3}{8} (1 - F_L)(1 + \cos^2 \theta_\mu) \right. \\ &\quad \left. - A_{FB} \cos \theta_\mu \right] \times \varepsilon_{\text{swap}}^\mu(-\cos \theta_\mu), \end{aligned} \quad (14)$$

$$\begin{aligned} \mathcal{F}_{\text{sig}}^\phi(\phi) &\equiv (1 - f_{\text{swap}}) [1 + \frac{1}{2} (1 - F_L) A_T^{(2)} \cos 2\phi + A_{im} \sin 2\phi] \times \varepsilon^\phi(\phi) \\ &\quad + f_{\text{swap}} [1 + \frac{1}{2} (1 - F_L) A_T^{(2)} \cos 2\phi + A_{im} \sin 2\phi] \times \varepsilon_{\text{swap}}^\phi(\phi + \pi), \end{aligned} \quad (15)$$

where f_{swap} is the swap fraction, ε^K , $\varepsilon_{\text{swap}}^K$, ε^μ , $\varepsilon_{\text{swap}}^\mu$, ε^ϕ , and $\varepsilon_{\text{swap}}^\phi$, are angular acceptance of true $\cos \theta_K$, swapped $\cos \theta_K$, true $\cos \theta_\mu$, swapped $\cos \theta_\mu$, true ϕ , and swapped ϕ events, respectively. For $B^+ \rightarrow K^+ \mu^+ \mu^-$, we expect the forward-backward asymmetry to be very small [30]. The PDF is described by Eq. (14) with no swap term, and we set $F_L = 1$ since K^+ is a pseudoscalar meson. We consider no scalar term [12].

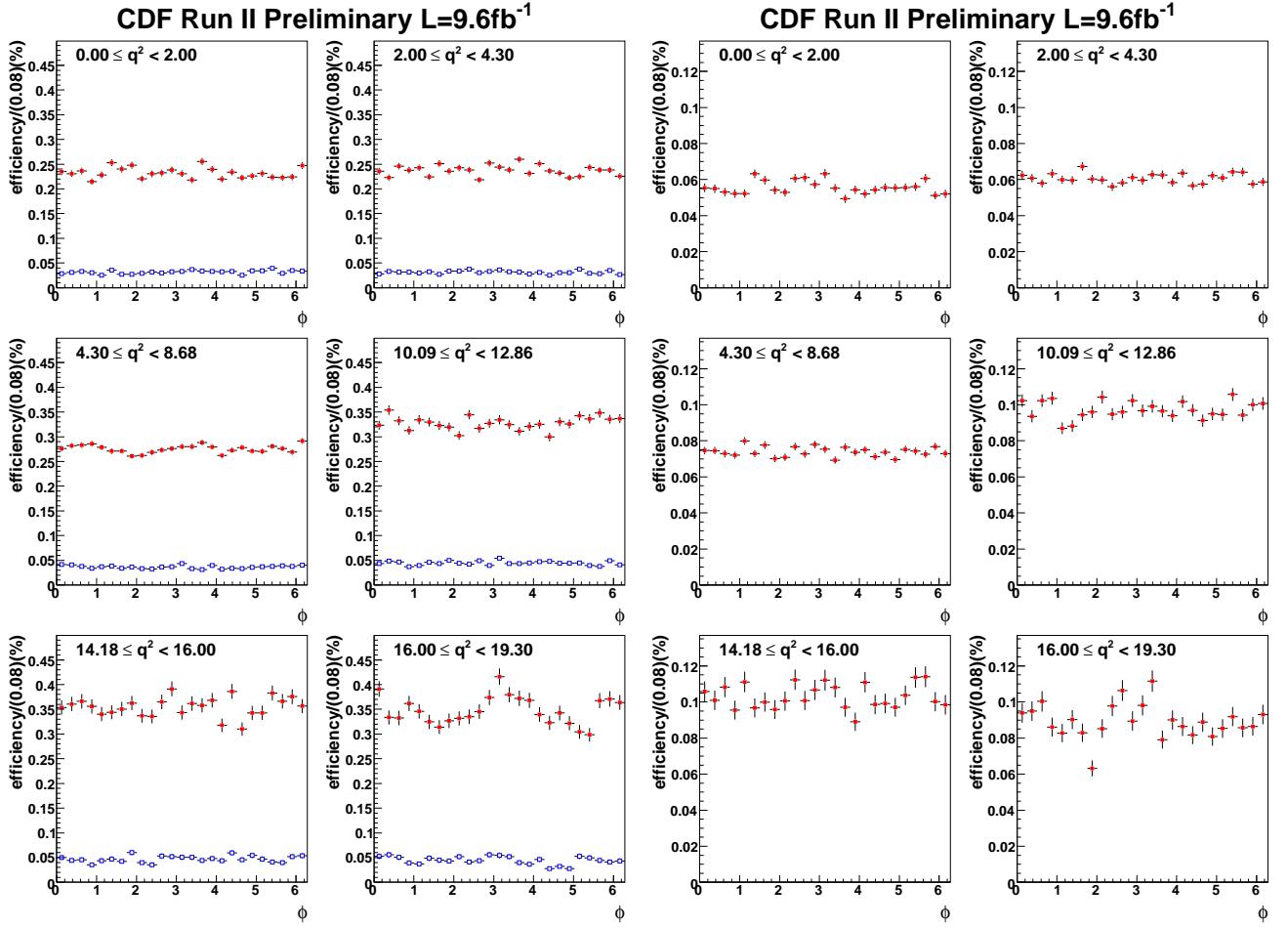


FIG. 7: ϕ acceptance of $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ (left two column) and $B^+ \rightarrow K^{*+} \mu^+ \mu^-$ (right two column). Filled circle is true signal and open circle shows $K\pi$ swapped signal.

The combinatorial background PDF shape is taken from the high-side B -mass sideband, $60 < M - M_B \leq 300 \text{ MeV}/c^2$. The sideband sample is divided into the same q^2 bins as the signal and is described by a polynomial function. We employ third-order (forth-order) polynomials for $\cos \theta_K$ ($\cos \theta_\mu$), and linear functions for the ϕ distributions. f_{swap} is evaluated from the signal MC ($7.5 \pm 0.1\%$) and fixed during the fit. Table XIII summarizes the parameters used for the angular analysis. Possible uncertainties due to the fixed parameters are assessed as the systematic errors.

	$\cos \theta_K$	$\cos \theta_\mu$	ϕ
f_{sig}	fixed	fixed	fixed
B mass shape	fixed	fixed	fixed
f_{swap}	fixed	fixed	fixed
BG slope	fixed	fixed	fixed
F_L	float	fixed	fixed
A_{FB}	-	float	-
$A_T^{(2)}$	-	-	float
A_{im}	-	-	float

TABLE XIII: Parameters used for each angular fit

Besides the single $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ fit, we perform the simultaneous fit with $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ and $B^+ \rightarrow K^{*+} \mu^+ \mu^-$ to increase the sensitivity. In the fit, we use different signal fractions, B mass and background shapes, angular

acceptances, and background functions for the charged and neutral samples. The angular observables are the only common parameters between two data sets.

A. Systematics Uncertainties for the angular analysis

The following systematic errors for F_L , A_{FB} , $A_T^{(2)}$, and A_{im} are evaluated in each q^2 bin independently:

- Signal Fraction and B Mass Shape

The uncertainties originating from the B mass fit are evaluated by varying the signal fraction and shape parameters by $\pm 1\sigma$, where σ is the statistical uncertainty of each quantity that is obtained from data or signal MC. This is the dominant uncertainty for the all angular observables.

- 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 using alternate binning (20 and 30 bins).

- Angular Background

For the central fit, we take the background shape from the B mass sideband region and fit the shape. The modeling uncertainty is evaluated by varying each shape parameters by $\pm 1\sigma$. For each evaluation the other shape parameters are evaluated by the shape fitting. Additional modeling uncertainty is also evaluated by increasing the order of the polynomial in the fit by one.

- $K\pi$ Swap

The $K\pi$ swap fraction is obtained from signal MC as $7.5 \pm 0.1\%$ in the fit region. This estimate was justified by the consistency of measured F_L in $B \rightarrow J/\psi K^*$ between data and MC. To allow for the discrepancy between data and MC in the rare decay, we switch the fraction to 5% and 10%, which are conservative variations, and repeat the fit.

- Peaking BG

The largest source of peaking BG for $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ is $B_s^0 \rightarrow \phi \mu^+ \mu^-$ feed-down that is reconstructed as $B^0 \rightarrow K^{*0} \mu^+ \mu^-$. The effect is estimated with special ensemble test which adds misreconstructed $B_s^0 \rightarrow \phi \mu^+ \mu^-$ MC angular distribution.

- Fit Bias

Using simulation, we find that if the true value of the angular observable is close to the physical boundary, the fitted value is affected. Possible fit bias is considered using the ensemble test whose input value is set to observed numbers.

- Trigger Bias

The trigger dependence on angular acceptances is studied by applying $p_T > 1.6\text{GeV}/c$, $2.1\text{GeV}/c$, and $3.1\text{GeV}/c$ cut for each CMU, CMX, CMP triggered muon to the acceptance function.

- F_L Fit

The uncertainty is evaluated by varying F_L up to both positive and negative 1σ . The uncertainty in A_{im} measurement is negligibly small.

B. Fit Results

We measure F_L , A_{FB} , $A_T^{(2)}$, and A_{im} for $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ and $B \rightarrow K^* \mu^+ \mu^-$. For $B^+ \rightarrow K^+ \mu^+ \mu^-$ we measure only A_{FB} . Fit results for $B^0 \rightarrow K^{*0} \mu^+ \mu^-$, $B \rightarrow K^* \mu^+ \mu^-$, and $B^+ \rightarrow K^+ \mu^+ \mu^-$ are shown in Figs. 8-10. Figs. 11, 12, and 13 show $\cos \theta_K$, $\cos \theta_\mu$, and ϕ distribution for $K^{*0} \mu^+ \mu^-$, respectively. Our choice of bins is driven by allowing easier combination with Belle [6] and LHCb [8]. Tables XXIII, XXIV, and XXV summarize the fit results for $B^0 \rightarrow K^{*0} \mu^+ \mu^-$, $B \rightarrow K^* \mu^+ \mu^-$, and $B^+ \rightarrow K^+ \mu^+ \mu^-$, respectively.

q^2 (GeV^2/c^2)	0.00- 2.00	2.00- 4.30	4.30- 8.68	10.09- 12.86	14.18- 16.00	16.00- 19.30	0.00- 4.30	1.00- 6.00
Signal fraction	0.014	0.063	0.035	0.016	0.021	0.073	0.003	0.059
B mass shape	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.001
Angular eff.	0.006	0.000	0.001	0.004	0.005	0.002	0.003	0.002
Angular BG	0.037	0.059	0.045	0.031	0.033	0.031	0.030	0.049
K - π swap	0.004	0.004	0.003	0.004	0.003	0.001	0.005	0.005
Fit bias	0.007	0.002	0.003	0.003	0.001	0.013	0.002	0.002
Trigger bias	0.002	0.001	0.002	0.004	0.004	0.003	0.002	0.002
Peaking BG	0.006	0.009	0.009	0.004	0.001	0.005	0.005	0.007
Total	0.042	0.087	0.058	0.036	0.040	0.081	0.031	0.078

TABLE XIV: F_L systematic uncertainty for $B^0 \rightarrow K^{*0} \mu^+ \mu^-$.

q^2 (GeV^2/c^2)	0.00- 2.00	2.00- 4.30	4.30- 8.68	10.09- 12.86	14.18- 16.00	16.00- 19.30	0.00- 4.30	1.00- 6.00
Signal fraction	0.070	0.093	0.013	0.051	0.010	0.070	0.015	0.020
B mass shape	0.000	0.001	0.000	0.000	0.000	0.000	0.001	0.002
Angular eff.	0.008	0.106	0.005	0.005	0.001	0.009	0.047	0.014
Angular BG	0.084	0.080	0.023	0.016	0.017	0.042	0.044	0.039
K - π swap	0.002	0.006	0.007	0.025	0.009	0.026	0.000	0.012
Fit bias	0.005	0.014	0.012	0.026	0.054	0.011	0.003	0.039
Trigger bias	0.008	0.016	0.001	0.001	0.002	0.001	0.008	0.001
Peaking BG	0.001	0.011	0.003	0.011	0.005	0.007	0.006	0.002
F_L fit	0.023	0.007	0.023	0.010	0.046	0.007	0.004	0.018
Total	0.112	0.164	0.038	0.066	0.074	0.087	0.067	0.064

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

q^2 (GeV^2/c^2)	0.00- 2.00	2.00- 4.30	4.30- 8.68	10.09- 12.86	14.18- 16.00	16.00- 19.30	0.00- 4.30	1.00- 6.00
Signal fraction	0.156	0.195	0.040	0.039	0.091	0.039	0.101	0.386
B mass shape	0.002	0.006	0.011	0.002	0.002	0.002	0.001	0.009
Angular eff.	0.010	0.029	0.012	0.016	0.018	0.012	0.008	0.032
Angular BG	0.066	0.229	0.317	0.041	0.100	0.028	0.046	0.185
K - π swap	0.008	0.007	0.006	0.001	0.002	0.002	0.006	0.017
Fit bias	0.054	0.022	0.035	0.026	0.017	0.033	0.040	0.044
Trigger bias	0.011	0.036	0.017	0.006	0.013	0.006	0.016	0.020
Peaking BG	0.009	0.043	0.027	0.027	0.001	0.002	0.000	0.012
F_L fit	0.230	0.456	0.058	0.045	0.040	0.116	0.143	0.622
Total	0.291	0.551	0.329	0.083	0.143	0.131	0.187	0.758

TABLE XVI: $A_T^{(2)}$ systematic uncertainty for $B^0 \rightarrow K^{*0} \mu^+ \mu^-$.

q^2 (GeV^2/c^2)	0.00- 2.00	2.00- 4.30	4.30- 8.68	10.09- 12.86	14.18- 16.00	16.00- 19.30	0.00- 4.30	1.00- 6.00
Signal fraction	0.072	0.014	0.065	0.041	0.023	0.010	0.039	0.124
B mass shape	0.001	0.000	0.001	0.000	0.000	0.000	0.000	0.001
Angular eff.	0.001	0.007	0.001	0.003	0.003	0.002	0.002	0.004
Angular BG	0.018	0.032	0.040	0.024	0.009	0.030	0.016	0.041
K - π swap	0.000	0.002	0.000	0.000	0.000	0.002	0.001	0.000
Fit bias	0.030	0.004	0.038	0.043	0.011	0.016	0.034	0.077
Trigger bias	0.004	0.006	0.000	0.003	0.003	0.005	0.001	0.001
Peaking BG	0.002	0.008	0.004	0.003	0.001	0.001	0.005	0.001
F_L fit	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Total	0.080	0.037	0.085	0.064	0.027	0.036	0.055	0.152

TABLE XVII: A_{im} systematic uncertainty for $B^0 \rightarrow K^{*0} \mu^+ \mu^-$.

q^2 (GeV^2/c^2)	0.00- 2.00	2.00- 4.30	4.30- 8.68	10.09- 12.86	14.18- 16.00	16.00- 19.30	0.00- 4.30	1.00- 6.00
Signal fraction	0.013	0.037	0.032	0.024	0.025	0.067	0.009	0.053
B mass shape	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.001
Angular eff.	0.006	0.001	0.001	0.004	0.004	0.002	0.003	0.002
Angular BG	0.038	0.055	0.041	0.029	0.031	0.025	0.030	0.043
$K\pi$ swap	0.004	0.004	0.003	0.004	0.003	0.001	0.004	0.005
Fit bias	0.009	0.016	0.002	0.003	0.004	0.007	0.003	0.002
Trigger bias	0.002	0.001	0.002	0.004	0.003	0.003	0.002	0.002
Peaking BG	0.006	0.009	0.009	0.004	0.001	0.005	0.005	0.007
Total	0.042	0.069	0.053	0.038	0.040	0.072	0.032	0.069

TABLE XVIII: F_L systematic uncertainty for $B \rightarrow K^* \mu^+ \mu^-$.

q^2 (GeV^2/c^2)	0.00- 2.00	2.00- 4.30	4.30- 8.68	10.09- 12.86	14.18- 16.00	16.00- 19.30	0.00- 4.30	1.00- 6.00
Signal fraction	0.057	0.090	0.022	0.056	0.021	0.048	0.027	0.018
B mass shape	0.001	0.001	0.000	0.000	0.000	0.000	0.001	0.001
Angular eff.	0.007	0.105	0.005	0.005	0.002	0.007	0.045	0.008
Angular BG	0.079	0.080	0.021	0.014	0.016	0.025	0.042	0.032
$K\pi$ swap	0.000	0.005	0.003	0.039	0.009	0.011	0.000	0.002
Fit bias	0.006	0.035	0.007	0.023	0.046	0.008	0.005	0.023
Trigger bias	0.007	0.015	0.000	0.001	0.001	0.001	0.008	0.006
Peaking BG	0.001	0.011	0.003	0.011	0.005	0.007	0.006	0.002
F_L fit	0.027	0.003	0.018	0.012	0.045	0.015	0.005	0.016
Total	0.101	0.165	0.037	0.076	0.071	0.059	0.069	0.048

TABLE XIX: A_{FB} systematic uncertainty for $B \rightarrow K^* \mu^+ \mu^-$.

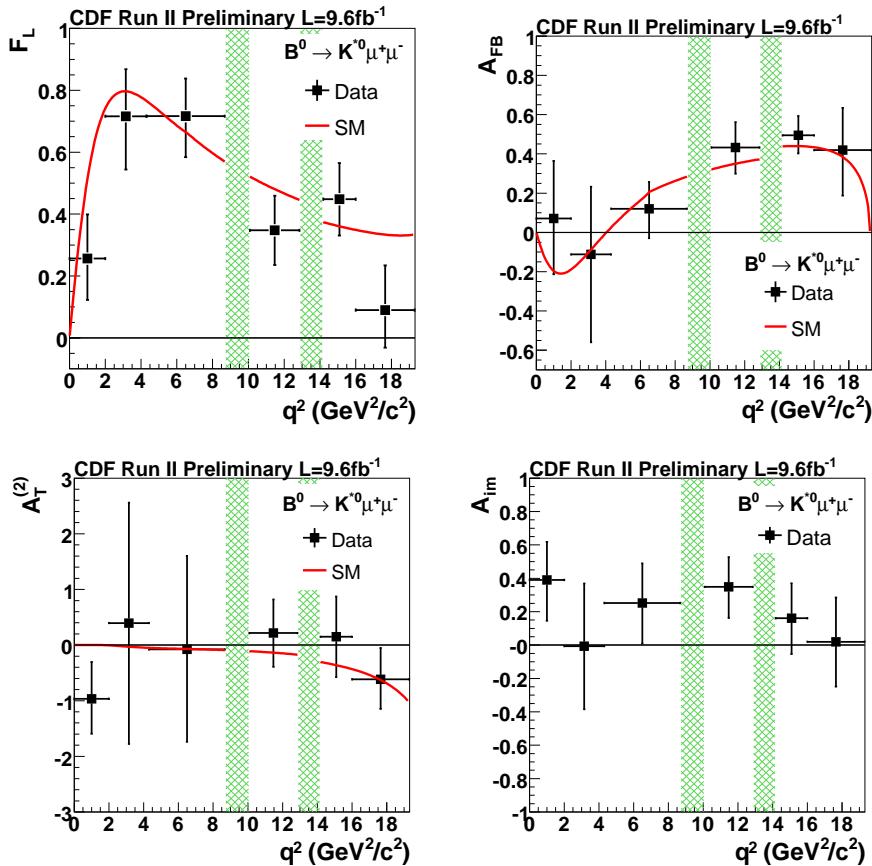
q^2 (GeV^2/c^2)	0.00- 2.00	2.00- 4.30	4.30- 8.68	10.09- 12.86	14.18- 16.00	16.00- 19.30	0.00- 4.30	1.00- 6.00
Signal fraction	0.195	0.204	0.479	0.056	0.063	0.031	0.119	0.137
B mass shape	0.002	0.008	0.010	0.001	0.001	0.002	0.001	0.006
Angular eff.	0.009	0.023	0.013	0.016	0.015	0.011	0.006	0.028
Angular BG	0.073	0.194	0.294	0.042	0.085	0.031	0.049	0.102
$K\pi$ swap	0.009	0.002	0.001	0.001	0.001	0.002	0.007	0.007
Fit bias	0.049	0.069	0.053	0.009	0.006	0.034	0.033	0.027
Trigger bias	0.012	0.034	0.017	0.006	0.011	0.007	0.015	0.015
Peaking BG	0.009	0.043	0.027	0.027	0.001	0.002	0.000	0.012
F_L fit	0.268	0.020	0.302	0.033	0.025	0.103	0.179	0.076
Total	0.343	0.296	0.641	0.084	0.111	0.118	0.224	0.192

TABLE XX: $A_T^{(2)}$ systematic uncertainty for $B \rightarrow K^* \mu^+ \mu^-$.

q^2 (GeV^2/c^2)	0.00- 2.00	2.00- 4.30	4.30- 8.68	10.09- 12.86	14.18- 16.00	16.00- 19.30	0.00- 4.30	1.00- 6.00
Signal fraction	0.060	0.063	0.056	0.045	0.017	0.022	0.051	0.111
B mass shape	0.001	0.001	0.001	0.000	0.000	0.001	0.001	0.001
Angular eff.	0.001	0.007	0.002	0.003	0.003	0.002	0.002	0.003
Angular BG	0.021	0.030	0.034	0.024	0.009	0.026	0.017	0.030
$K\pi$ swap	0.000	0.002	0.000	0.000	0.000	0.001	0.001	0.000
Fit bias	0.026	0.011	0.043	0.035	0.012	0.007	0.020	0.062
Trigger bias	0.004	0.005	0.001	0.003	0.003	0.005	0.001	0.001
Peaking BG	0.002	0.008	0.004	0.003	0.001	0.001	0.005	0.001
F_L fit	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Total	0.069	0.072	0.079	0.062	0.024	0.035	0.058	0.131

TABLE XXI: A_{im} systematic uncertainty for $B \rightarrow K^* \mu^+ \mu^-$.

q^2 (GeV^2/c^2)	0.00- 2.00	2.00- 4.30	4.30- 8.68	10.09- 12.86	14.18- 16.00	16.00- 23.00	0.00- 4.30	1.00- 6.00
Signal fraction	0.027	0.088	0.000	0.003	0.004	0.042	0.043	0.016
B mass shape	0.001	0.001	0.000	0.000	0.000	0.001	0.001	0.001
Angular eff.	0.030	0.008	0.002	0.001	0.001	0.002	0.022	0.012
Angular BG	0.069	0.042	0.010	0.014	0.009	0.018	0.013	0.010
Fit bias	0.035	0.004	0.004	0.025	0.002	0.002	0.011	0.005
Trigger bias	0.008	0.007	0.001	0.000	0.000	0.001	0.002	0.000
Total	0.088	0.098	0.011	0.029	0.010	0.045	0.052	0.023

TABLE XXII: A_{FB} systematic uncertainty for $B^+ \rightarrow K^+ \mu^+ \mu^-$.FIG. 8: Angular analysis results with respect to six q^2 bins for $B^0 \rightarrow K^0 \mu^+ \mu^-$. From top-left to bottom-right, F_L , A_{FB} , $A_T^{(2)}$, and A_{im} plot are shown. Histogram is the fit result, solid curve is the SM expectation [31], and hatched regions mean charmonium veto.

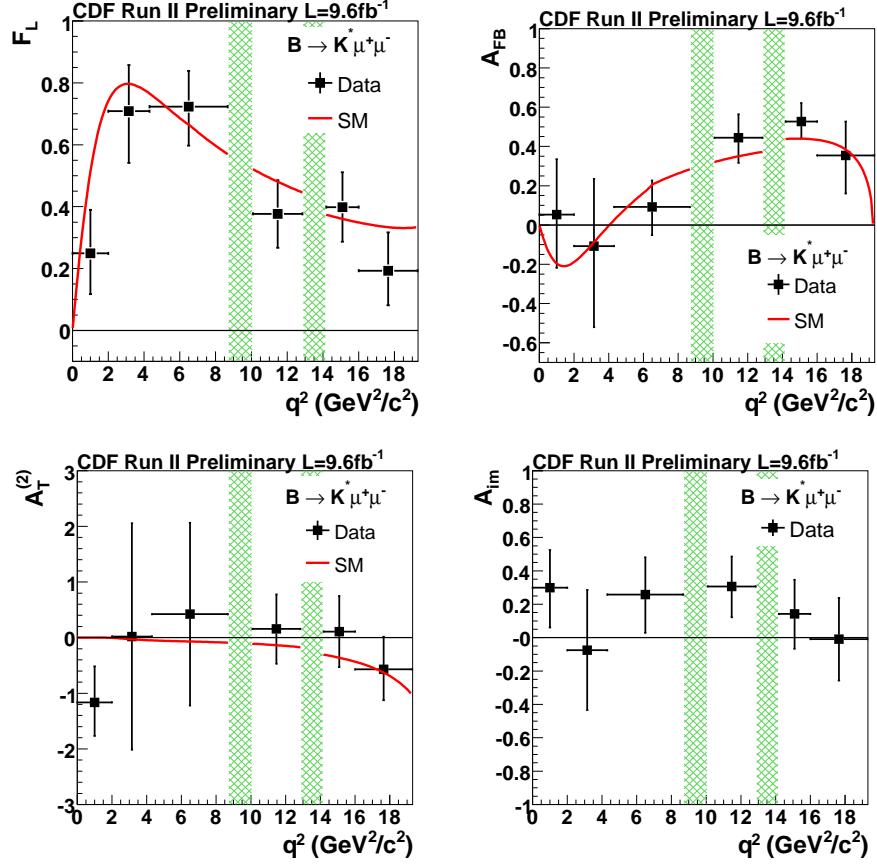


FIG. 9: Angular analysis results with respect to six q^2 bins for $B \rightarrow K^* \mu^+ \mu^-$. From top-left to bottom-right, F_L , A_{FB} , $A_T^{(2)}$, and A_{im} plot are shown. Histogram is the fit result, solid curve is the SM expectation [31], and hatched regions mean charmonium veto.

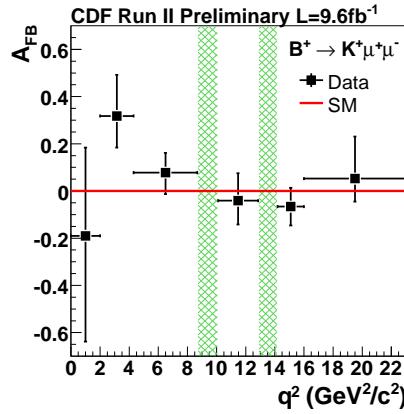
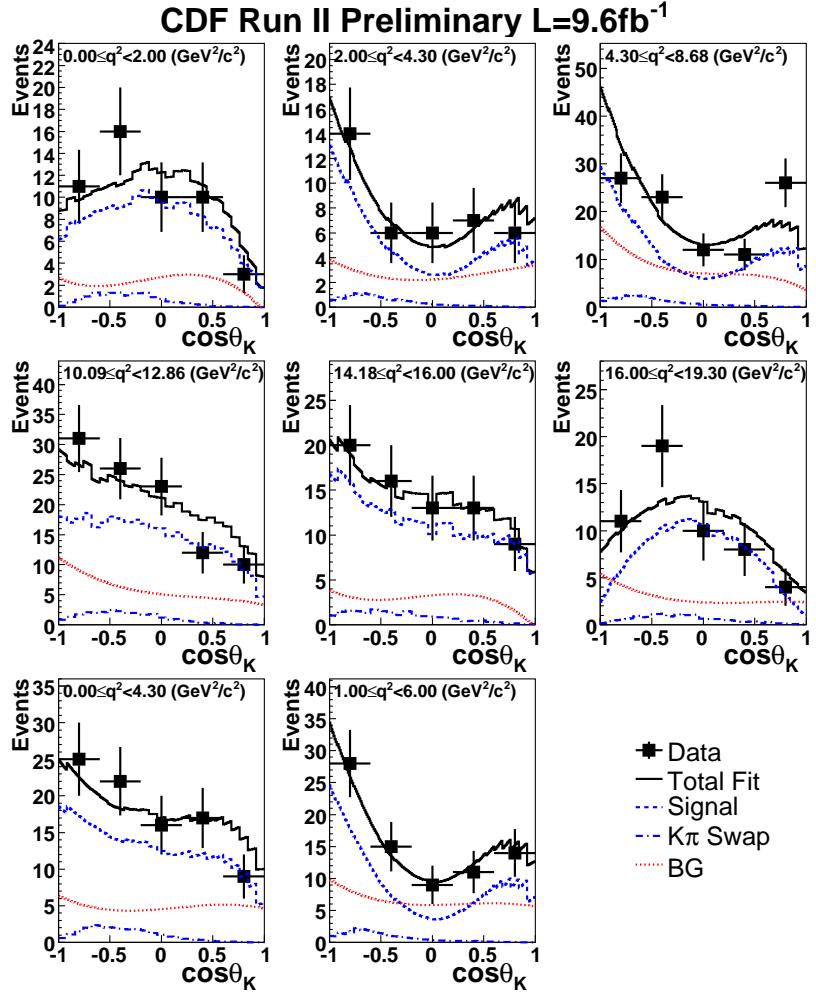
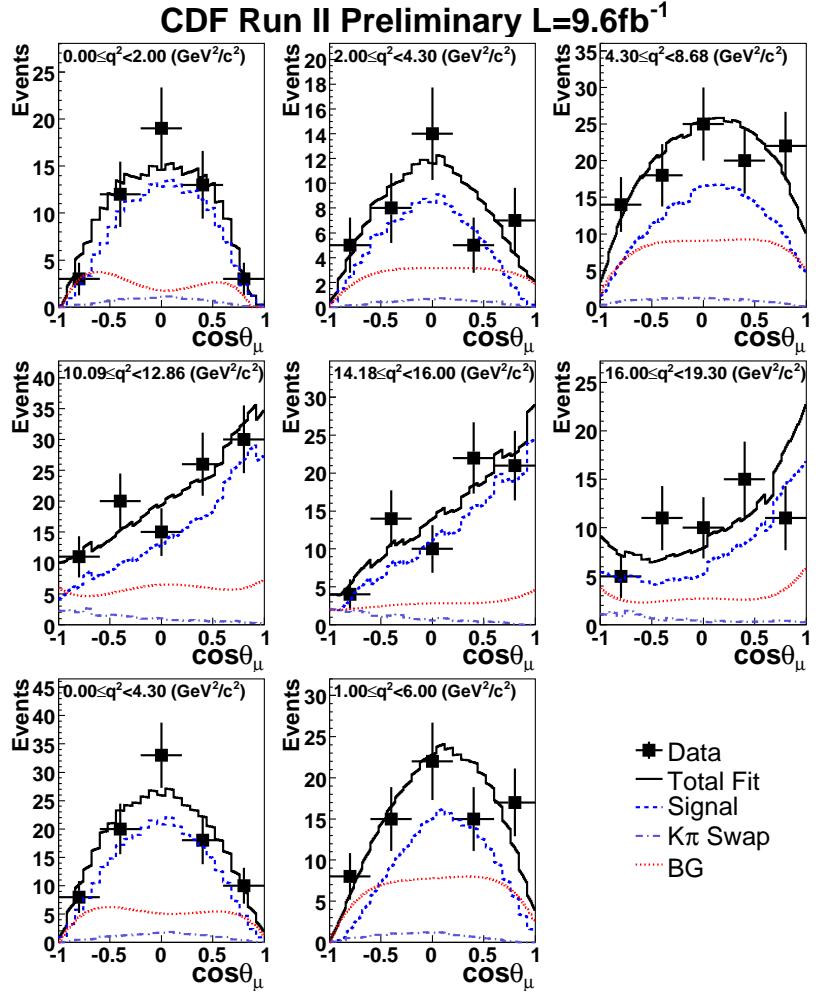


FIG. 10: A_{FB} result as a function of q^2 for $B^+ \rightarrow K^+ \mu^+ \mu^-$. Histogram is the fit result, solid curve is the SM expectation, and hatched regions mean charmonium veto.

FIG. 11: $\cos \theta_K$ distributions for $B^0 \rightarrow K^{*0} \mu^+ \mu^-$.

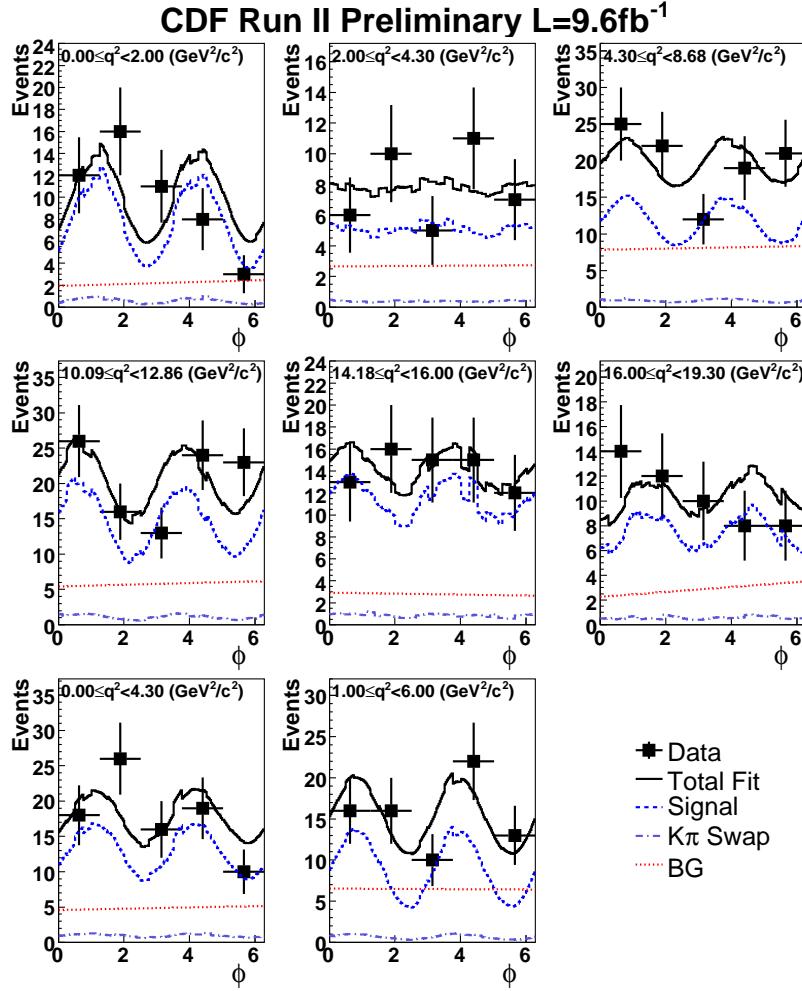
q^2 range	F_L	A_{FB}	$A_T^{(2)}$	A_{im}
[0.00, 2.00)	$0.26^{+0.14}_{-0.13} \pm 0.04$	$0.07^{+0.29}_{-0.28} \pm 0.11$	$-0.97^{+0.66}_{-0.63} \pm 0.29$	$0.39^{+0.23}_{-0.24} \pm 0.08$
[2.00, 4.30)	$0.72^{+0.15}_{-0.17} \pm 0.09$	$-0.11^{+0.34}_{-0.45} \pm 0.16$	$0.40^{+2.16}_{-2.17} \pm 0.55$	$-0.01^{+0.38}_{-0.38} \pm 0.04$
[4.30, 8.68)	$0.72^{+0.12}_{-0.13} \pm 0.06$	$0.12^{+0.14}_{-0.15} \pm 0.04$	$-0.08^{+1.68}_{-1.66} \pm 0.33$	$0.25^{+0.24}_{-0.24} \pm 0.09$
[10.09, 12.86)	$0.35^{+0.11}_{-0.11} \pm 0.04$	$0.43^{+0.13}_{-0.13} \pm 0.07$	$0.22^{+0.60}_{-0.61} \pm 0.08$	$0.35^{+0.18}_{-0.19} \pm 0.06$
[14.18, 16.00)	$0.45^{+0.12}_{-0.12} \pm 0.04$	$0.49^{+0.10}_{-0.09} \pm 0.07$	$0.15^{+0.72}_{-0.72} \pm 0.14$	$0.16^{+0.21}_{-0.22} \pm 0.03$
[16.00, 19.30)	$0.09^{+0.14}_{-0.12} \pm 0.08$	$0.42^{+0.22}_{-0.23} \pm 0.09$	$-0.62^{+0.56}_{-0.53} \pm 0.13$	$0.02^{+0.26}_{-0.27} \pm 0.04$
[0.00, 4.30)	$0.44^{+0.11}_{-0.11} \pm 0.03$	$-0.04^{+0.23}_{-0.23} \pm 0.07$	$-0.59^{+0.68}_{-0.67} \pm 0.19$	$0.25^{+0.21}_{-0.21} \pm 0.05$
[1.00, 6.00)	$0.78^{+0.13}_{-0.15} \pm 0.08$	$0.29^{+0.25}_{-0.21} \pm 0.06$	$-0.45^{+2.24}_{-2.22} \pm 0.76$	$0.51^{+0.28}_{-0.29} \pm 0.15$

TABLE XXIII: Summary of $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ fit results. First (second) error is statistical (systematic).

FIG. 12: $\cos \theta_\mu$ distributions for $B^0 \rightarrow K^{*0} \mu^+ \mu^-$.

q^2 range	F_L	A_{FB}	$A_T^{(2)}$	A_{im}
[0.00, 2.00)	$0.25^{+0.14}_{-0.13} \pm 0.04$	$0.05^{+0.28}_{-0.27} \pm 0.10$	$-1.16^{+0.65}_{-0.60} \pm 0.34$	$0.30^{+0.23}_{-0.24} \pm 0.07$
[2.00, 4.30)	$0.71^{+0.15}_{-0.17} \pm 0.07$	$-0.11^{+0.34}_{-0.41} \pm 0.16$	$0.02^{+2.04}_{-2.04} \pm 0.30$	$-0.08^{+0.36}_{-0.36} \pm 0.07$
[4.30, 8.68)	$0.72^{+0.12}_{-0.13} \pm 0.05$	$0.09^{+0.14}_{-0.14} \pm 0.04$	$0.42^{+1.64}_{-1.64} \pm 0.64$	$0.26^{+0.22}_{-0.23} \pm 0.08$
[10.09, 12.86)	$0.38^{+0.11}_{-0.11} \pm 0.04$	$0.44^{+0.12}_{-0.13} \pm 0.08$	$0.16^{+0.62}_{-0.62} \pm 0.08$	$0.31^{+0.18}_{-0.18} \pm 0.06$
[14.18, 16.00)	$0.40^{+0.11}_{-0.11} \pm 0.04$	$0.53^{+0.09}_{-0.09} \pm 0.07$	$0.11^{+0.64}_{-0.64} \pm 0.11$	$0.14^{+0.20}_{-0.21} \pm 0.02$
[16.00, 19.30)	$0.19^{+0.12}_{-0.11} \pm 0.07$	$0.35^{+0.17}_{-0.19} \pm 0.06$	$-0.57^{+0.59}_{-0.56} \pm 0.12$	$-0.01^{+0.25}_{-0.25} \pm 0.04$
[0.00, 4.30)	$0.43^{+0.11}_{-0.11} \pm 0.03$	$-0.04^{+0.23}_{-0.22} \pm 0.07$	$-0.76^{+0.66}_{-0.65} \pm 0.22$	$0.15^{+0.20}_{-0.20} \pm 0.06$
[1.00, 6.00)	$0.76^{+0.12}_{-0.14} \pm 0.07$	$0.19^{+0.17}_{-0.21} \pm 0.05$	$-0.07^{+1.90}_{-1.90} \pm 0.19$	$0.42^{+0.24}_{-0.26} \pm 0.13$

TABLE XXIV: Summary of $B \rightarrow K^* \mu^+ \mu^-$ fit results. First (second) error is statistical (systematic).

FIG. 13: ϕ distributions for $B^0 \rightarrow K^{*0} \mu^+ \mu^-$.

q^2 range	F_L	A_{FB}	$A_T^{(2)}$	A_{im}
[0.00, 2.00)	-	$-0.19^{+0.37}_{-0.45} \pm 0.09$	-	-
[2.00, 4.30)	-	$0.32^{+0.17}_{-0.13} \pm 0.10$	-	-
[4.30, 8.68)	-	$0.08^{+0.08}_{-0.09} \pm 0.01$	-	-
[10.09, 12.86)	-	$-0.04^{+0.12}_{-0.10} \pm 0.03$	-	-
[14.18, 16.00)	-	$-0.07^{+0.08}_{-0.06} \pm 0.01$	-	-
[16.00, 23.00)	-	$0.05^{+0.18}_{-0.10} \pm 0.05$	-	-
[0.00, 4.30)	-	$0.23^{+0.22}_{-0.17} \pm 0.05$	-	-
[4.30, 6.00)	-	$0.13^{+0.11}_{-0.10} \pm 0.02$	-	-

TABLE XXV: Summary of $B^+ \rightarrow K^+ \mu^+ \mu^-$ fit results. First (second) error is statistical (systematic).

VII. CONCLUSION

- We update the analysis of FCNC decays $b \rightarrow s\mu^+\mu^-$ to the 9.6 fb^{-1} sample that is whole CDF data sample.
- We obtain measurements of branching ratio of the exclusive $b \rightarrow s\mu^+\mu^-$ final states:

$$\begin{aligned}\mathcal{B}(B^+ \rightarrow K^+\mu^+\mu^-) &= [0.45 \pm 0.03(\text{stat}) \pm 0.02(\text{syst})] \times 10^{-6}, \\ \mathcal{B}(B^0 \rightarrow K^{*0}\mu^+\mu^-) &= [1.14 \pm 0.09(\text{stat}) \pm 0.06(\text{syst})] \times 10^{-6}, \\ \mathcal{B}(B_s^0 \rightarrow \phi\mu^+\mu^-) &= [1.17 \pm 0.18(\text{stat}) \pm 0.37(\text{syst})] \times 10^{-6}, \\ \mathcal{B}(B^0 \rightarrow K^0\mu^+\mu^-) &= [0.33 \pm 0.08(\text{stat}) \pm 0.03(\text{syst})] \times 10^{-6}, \\ \mathcal{B}(B^+ \rightarrow K^{*+}\mu^+\mu^-) &= [0.89 \pm 0.25(\text{stat}) \pm 0.09(\text{syst})] \times 10^{-6}, \\ \mathcal{B}(\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-) &= [1.95 \pm 0.34(\text{stat}) \pm 0.61(\text{syst})] \times 10^{-6}.\end{aligned}$$

- We update differential branching ratios in the channels all above and measure isospin asymmetry.
- We update A_{FB} in $B^+ \rightarrow K^+\mu^+\mu^-$ and A_{FB} , F_L , $A_T^{(2)}$, and A_{im} , in $B^0 \rightarrow K^{*0}\mu^+\mu^-$ and combined $B \rightarrow K^*\mu^+\mu^-$.
- All results are consistent with other experiments and the SM.

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 Korean 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 Nucléaire 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; the Academy of Finland; and the Australian Research Council (ARC).

-
- [1] D. Melikhov, N. Nikitin, and S. Simula, Phys. Rev. D **57**, 6814 (1998).
[2] A. Ali *et al.*, Phys. Rev. D **61**, 074024 (2000).
[3] Q. Chan and Y.-H. Gao, Nucl. Phys. B **845**, 179 (2011).
[4] Y.-M. Wang, M. J. Aslam, and C.-D. Lu, Eur. Phys. J. C **59**, 847 (2009).
[5] B. Aubert *et al.* (BaBar Collaboration), Phys. Rev. Lett. **102**, 091803 (2009), arXiv:0807.4119.
[6] J. T. Wei *et al.* (Belle Collaboration), Phys. Rev. Lett. **103**, 171801 (2009), arXiv:0904.0770.
[7] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **106**, 161801 (2011), arXiv:1101.1028.
[8] R. Aaij *et al.* (LHCb Collaboration), Phys. Rev. Lett. **108**, 181806 (2012), arXiv:1112.3515.
[9] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **107**, 201802 (2011), arXiv:1107.3753.
[10] R. Aaij *et al.* (LHCb Collaboration), (2012), arXiv:1205.3422.
[11] F. Kruger and J. Matias, Phys. Rev. D **71**, 094009 (2005).
[12] C. Bobeth, G. Hiller, and G. Piranishvili, J. High Energy Phys. 12 (2007) 040.
[13] U. Egede, T. Hurth, J. Matias, M. Ramon, and W. Reece, JHEP **11**, 032 (2008), arXiv:0807.2589.
[14] S. Descotes-Genon, D. Ghosh, J. Matias, and M. Ramon, JHEP **06**, 099 (2011), arXiv:1104.3342.
[15] C. Bobeth, G. Hiller, and D. van Dyk, JHEP **07**, 098 (2010), arXiv:1006.5013.
[16] A. K. Alok *et al.*, (2010), arXiv:1008.2367.
[17] B. Aubert *et al.* (BaBar Collaboration), Phys. Rev. D **79**, 031102 (2009).
[18] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **108**, 081807 (2012), arXiv:1108.0695.
[19] F. Abe, *et al.*, Nucl. Instrum. Methods Phys. Res. A **271**, 387 (1988); D. Amidei, *et al.*, Nucl. Instrum. 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.
[20] W. Ashmanskas *et al.*, Nucl. Instrum. Methods A **518**, 532 (2004).

- [21] K. Nakamura *et al.* (Particle Data Group), J. Phys. **G37**, 075021 (2010).
- [22] J. Podolanski and R. Armenteros, Phil. Mag. **45**, 13 (1954).
- [23] M. Feindt, (2004), arXiv:physics/0402093.
- [24] Isolation I is defined as the $p_T(B)$ divided by the scalar sum of the p_T values of all tracks in a cone of $\Delta R \equiv \sqrt{\Delta\eta^2 + \Delta\phi^2} = 1.0$, where $\Delta\eta$ and $\Delta\phi$ are the difference in pseudorapidity and azimuthal angle of the B candidate and each track.
- [25] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. **D84**, 031102 (2011), arXiv:1105.0690.
- [26] P. Ball and R. Zwicky, Phys. Rev. D **71**, 014015 (2005); P. Ball and R. Zwicky, Phys. Rev. D **71**, 014029 (2005).
- [27] T. M. Aliev, K. Azizi, and M. Savci, Phys. Rev. D **81**, 056006 (2010).
- [28] T. M. Aliev and M. Savci, Eur. Phys. J. **C50**, 91 (2007), arXiv:hep-ph/0606225.
- [29] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. D **79**, 011104 (2009).
- [30] C. Bobeth, T. Ewerth, F. Kruger, and J. Urban, Phys. Rev. **D64**, 074014 (2001), arXiv:hep-ph/0104284.
- [31] We draw theory curves using the EOS code by D. van Dyk *et al.* (<http://project.het.physik.tu-dortmund.de/eos/>). However, for illustration purposes, we extend these curves into kinematical regions ($q^2 < 1$, $7 < q^2 < 8.68$, $10.09 < q^2 < 12.86$) where their reliability is known to be approximate.