We report on the measurement of ratios of branching ratio of decays $B \to K \mu \mu$ and $B \to Ke e$, dubbed as $R_K$, using a data sample of 711 fb$^{-1}$ at $\Upsilon(4S)$ resonance recorded by the Belle detector. In addition, we have measured the CP-averaged isospin asymmetry and differential branching ratio of these decays. These observables are calculated in the bins of di-lepton invariant mass squared, $q^2$. We have also searched for lepton-flavor-violating $B \to K \mu^\pm e^\mp$ decays. The upper limits on the branching ratio are $\mathcal{O}(10^{-8})$. We improve the existing upper limit on $B^0 \to K_S^0 \mu^\pm e^\mp$ by an order of magnitude, $\mathcal{B}(B^0 \to K_S^0 \mu^\pm e^\mp) < 3.8 \times 10^{-8}$ at 90% confidence level.
1. Introduction

Rare decays $B \to K^{(*)} \ell \ell$, $\ell = e$ or $\mu$, involve the quark transition $b \to s$, which are flavor changing neutral currents. These processes occur through penguin loop or box diagrams in the standard model (SM). The decays are highly suppressed because of loop-level contributions, having very small branching ratio (BR) of the order of $O(10^{-7})$. New physics can contribute by enhancing or suppressing the amplitude of the decay or modifying the angular distribution of the final state particles. The lepton-flavor-universality (LFU) ratio is defined as the ratio of BR’s of $B \to K \mu \mu$ and $B \to K ee$,

$$R_K = \frac{B \to K \mu \mu}{B \to K ee}$$

According to SM, this ratio should be consistent with unity. LHCb [1, 2] has reported a deviation of $2.5\sigma$ for a bin of $q^2 \in [1.1, 6.0]$ GeV$^2$/c$^4$, where $q^2$ is the invariant mass squared of the lepton pair. Belle [3] result on this observable for whole $q^2$ bin with 605 fb$^{-1}$ data sample was consistent with SM with large statistical uncertainty.

Another theoretically robust observable [4], where form factor related systematic cancels is CP-averaged isospin asymmetry,

$$A_I = \frac{\tau_{B^+}/\tau_{B^0}}{\tau_{B^+}/\tau_{B^0}} \frac{\mathcal{B}(B^0 \to K^0 \ell^+ \ell^-) - \mathcal{B}(B^+ \to K^+ \ell^+ \ell^-)}{\mathcal{B}(B^0 \to K^0 \ell^+ \ell^-) + \mathcal{B}(B^+ \to K^+ \ell^+ \ell^-)},$$

where $\tau_{B^+}/\tau_{B^0} = 1.076 \pm 0.004$ is the lifetime ratio of $B^+$ and $B^0$ [5]. The $A_I$ value is expected to be zero in SM. Belle [3] and BaBar [6] have previously reported this value to be significantly below zero, while LHCb [7] results are consistent with SM predictions.

In many theoretical models [8] lepton flavor violation (LFV) accompanies LFU violation. LHCb [9] has reported most stringent upper limit on LFV $B^+ \to K^+ \mu^+ e^-$ and $B^+ \to K^+ \mu^- e^+$ modes to be $6.4 \times 10^{-9}$ and $7.0 \times 10^{-9}$ at 90\% confidence level upper limit on the branching ratio. The $B^0 \to K_S^0 \mu^+ e^+$ upper limit was set by BaBar [10] to be $2.7 \times 10^{-7}$.

We have reported [11] the measurement of branching fraction in $B \to K \ell \ell$, $R_K$, and $A_I$ for $[0.1, 4.0]$, $[4.0, 8.12]$, $[1.0, 6.0]$, $[10.2, 12.8]$, $> 14.18$, and $> 0.1$ GeV$^2$/c$^4$ $q^2$ bins. We have also searched for LFV in $B \to K \mu^+ e^-$ decays using full data sample of 711 fb$^{-1}$ at $\Upsilon(4S)$ resonance recorded with the Belle detector.

2. Belle detector

The Belle detector [12] is a large-solid-angle magnetic spectrometer composed of a silicon vertex detector (SVD), a 50-layer central drift chamber (CDC), an array of aerogel threshold Cherenkov counters (ACC), a barrel-like arrangement of time-of-flight scintillation counters (TOF), and an electromagnetic calorimeter (ECL) comprising CsI(Tl) crystals. All these subdetectors are located inside a superconducting solenoid coil that provides a 1.5 T magnetic field. An iron flux-return yoke placed outside the coil is instrumented with resistive plate chambers (KLM) to detect $K_L^0$ mesons and muons. Two inner detector configurations were used: a 2.0 cm radius beam pipe and a three-layer SVD for the first sample of 140 fb$^{-1}$; and a 1.5 cm radius beam pipe, a four-layer SVD, and a small-cell inner CDC for the remaining 571 fb$^{-1}$ [13].
3. Selection Criteria

We reconstruct $B^+ \rightarrow K^+\ell\ell$ and $B^0 \rightarrow K^{0}\ell\ell$, where $\ell = e$ or $\mu$. The charged particles are selected near the interaction point, with impact parameters less than 1.0 cm in the transverse direction w.r.t the beam direction and less than 4.0 cm in the direction of the beam. The kaon candidates are selected with $R_{K/K} = L_K/(L_K + L_\pi) > 0.6$ where $L_K$ and $L_\pi$ are the likelihoods of kaon and pion, respectively. This selection has an efficiency of 92% with 7% pion misidentification rate. The $K^0_S$ are reconstructed from two oppositely charged tracks (assumed to be pions) and with the invariant mass within 3$\sigma$ around the $K^0_S$ nominal mass (sigma is the expected resolution on the invariant mass). The muon candidates are selected with $L_\mu > 0.9$, efficiency of 89% and the pion fake-rate of 1.5%. The electron selection efficiency is 92% with < 1% pion misidentification rate for a cut of $L_e > 0.9$. The $K^+$ or $K^0_S$ candidates are combined with oppositely charged pair of leptons to form a $B^+$ or $B^0$ meson. The kinematic variables which distinguish signal from background are beam energy constrained mass ($M_{bc}$) and energy difference ($\Delta E$):

$$M_{bc} = \frac{E_{beam}^2/c^4 - |p_B|^2/c^2}{\Delta E = E_B - E_{beam}}.$$  

Here $E_{beam}$ is the beam energy, and $E_B$ and $p_B$ are the energy and momentum of the $B$ candidate, respectively, in the center of mass frame. The candidates which satisfy $5.2 < M_{bc} < 5.3$ GeV/c$^2$ and $-0.1 < \Delta E < 0.25$ GeV are selected.

4. Background suppression and signal extraction

At this stage of analysis, we found significant contribution from continuum ($e^+e^- \rightarrow q\bar{q}$) and generic $B$ decays backgrounds. These background are suppressed by vertex quality requirements, event shape variables selection, and requirements on kinematic variables which distinguish signal from background. These variables are used to train a neural network (NN). The NN output obtained is translated to another variable using a log function,

$$O' = \log \frac{O - O_{\min}}{O_{\max} - O},$$

The $O_{\min}$ is the minimum NN cut applied to translate the variable and this cut reduced the background by $>75\%$ with $4 - 5\%$ signal efficiency loss.

The peaking background for each of the decay modes is studied separately. For $B^+ \rightarrow J/\psi(\ell\ell)K^+$ channel, the $[\pi^+J/\psi]$ background peaks at $M_{bc}$ signal region and shifted in $\Delta E$. This background component is kept fixed in the fit along with fraction of $B^+ \rightarrow J/\psi(\ell\ell)\pi^+$ to $B^+ \rightarrow J/\psi(\ell\ell)\pi^+$. For $B^+ \rightarrow K^+\mu\mu$, we found contribution from $B^+ \rightarrow D^0(\rightarrow K^+\pi^-)\pi^+$, where both muons are misidentified pions and are removed by invariant mass veto around the $D^0$ mass region, i.e., [1.85, 1.88] GeV/c$^2$. The peaking due to $B^+ \rightarrow J/\psi(\rightarrow \mu\mu)K^+$, misidentification of muon as a kaon, and kaon as a muon, is removed by $M_{K^+\mu^-}$ veto in [3.06 – 3.13] GeV/c$^2$. For LFV $B^+ \rightarrow K^{0}\mu^+\mu^-$ mode, the peaking backgrounds are coming from $B^+ \rightarrow J/\psi(ee)K^+$, with (a) the electron misidentified as a kaon and the kaon as a muon; this background is removed by vetoing events with $M_{K^{0}\mu^-} \in (2.95, 3.11)$ GeV/c$^2$; (b) the electron is misidentified as a muon; this
background is suppressed by vetoing events with \( M_{\mu^+\mu^-} \in (3.02, 3.12) \text{ GeV/c}^2 \). For \( B^+ \to K^+\mu^-e^+ \) channel, we found only the latter background and removed with \( M_{\mu^+\mu^-} \in (3.02, 3.12) \text{ GeV/c}^2 \) veto. For \( B^0 \to K_S^0\mu^0e^- \), the background from \( B^0 \to J/\psi(ee)K^0_S \) is removed with \( M_{\mu^+\mu^-} \in (3.04, 3.12) \text{ GeV/c}^2 \). A small contribution of \( B \to K\pi^+\pi^- \) is fixed in the fit by studying all the intermediate resonances.

The signal is extracted by performing an extended maximum likelihood fit in 3-dimensions, \( M_{bc}, \Delta E \) and \( O' \). The signal distributions in \( M_{bc}, \Delta E \) and \( O' \) are fitted with a Gaussian, Crystal Ball + Gaussian, and bifurcated Gaussian + Gaussian function, respectively. The continuum background distributions are fitted with ARGUS, Chebychev polynomial, and Gaussian function for \( M_{bc}, \Delta E \), and \( O' \), respectively. The generic \( B \) background distributions are fitted with ARGUS, exponential, and Gaussian for \( M_{bc}, \Delta E \), and \( O' \), respectively.

5. Results

We found 16736 \( \pm 130, 17010 \pm 130, 4961 \pm 71, \) and 4710 \( \pm 69 \) events for \( B^+ \to J/\psi(\mu\mu)K^+, B^+ \to J/\psi(\mu\mu)K^+, B^0 \to J/\psi(\mu\mu)K_S^0 \), and \( B^0 \to J/\psi(\mu\mu)K_S^0 \), respectively. The data fit results are shown in Figure 1. We find \( \mathcal{B}(B^+ \to J/\psi K^+) = 1.032 \pm 0.007 \pm 0.024 \) and \( \mathcal{B}(B^0 \to J/\psi K^0) = 0.902 \pm 0.010 \pm 0.026 \), which are the most precise measurements to date.

The \( B \to J/\psi(\to \ell\ell)K \) sample is used as a calibration of the signal PDF of \( B \to K\ell\ell \). The off-resonance sample, recorded 60 MeV below the \( \Upsilon(4S) \) resonance, is used to calibrate the continuum background. There are 137 \( \pm 14, 138 \pm 15, 27.3^{+6.6}_{-5.8}, \) and 21.8\( ^{+7.0}_{-6.1} \) signal events for \( B^+ \to K^+\mu\mu, \)
Figure 2: $R_{K^+}$ (left), $R_{K^0}$ (middle), and $R_K$ (right) results for $q^2 \in [0.1, 4.0]$ GeV$^2$/c$^4$ (blue), $q^2 \in [4.0, 8.12]$ GeV$^2$/c$^4$ (blue), $q^2 \in [1.0 - 6.0]$ GeV$^2$/c$^4$ (red), $q^2 \in [10.2 - 12.8]$ GeV$^2$/c$^4$ (blue), $q^2 > 14.18$ GeV$^2$/c$^4$ (blue), and whole $q^2$ interval (green).

Figure 3: $A_T(B \rightarrow K_{\mu\mu})$ (left), $A_T(B \rightarrow K_{ee})$ (middle), and $A_T(B \rightarrow K_{\ell\ell})$ (right) results for $q^2 \in [0.1, 4.0]$ GeV$^2$/c$^4$ (blue), $q^2 \in [4.0, 8.12]$ GeV$^2$/c$^4$ (blue), $q^2 \in [1.0 - 6.0]$ GeV$^2$/c$^4$ (red), $q^2 \in [10.2 - 12.8]$ GeV$^2$/c$^4$ (blue), $q^2 > 14.18$ GeV$^2$/c$^4$ (blue), and whole $q^2$ interval (green).

$B^+ \rightarrow K^*e\mu$, $B^0 \rightarrow K_{S0}\mu\mu$, and $B^0 \rightarrow K_{S0}ee$ modes, respectively. The branching ratio integrated over the full $q^2$ interval is $(5.99^{+0.45}_{-0.43} \pm 0.14) \times 10^{-7}$ and $(3.51^{+0.69}_{-0.60} \pm 0.10) \times 10^{-7}$ for $B^+ \rightarrow K^+\ell\ell$ and $B^0 \rightarrow K^0\ell\ell$, respectively. The $R_K$ is calculated using equation 1 and the results are shown in Figure 2. The observable is calculated for $q^2 \in [0.1, 4.0]$ GeV$^2$/c$^4$, $q^2 \in [4.0, 8.12]$ GeV$^2$/c$^4$, $q^2 \in [1.0 - 6.0]$ GeV$^2$/c$^4$, $q^2 \in [10.2 - 12.8]$ GeV$^2$/c$^4$, $q^2 > 14.18$ GeV$^2$/c$^4$, and whole $q^2$ interval, respectively. The $A_T$ is calculated in these bins using equation 2 and shown in Figure 3. The $d\mathcal{B}/dq^2$ is calculated for changed and neutral $B$ mesons, and the results are shown in Figure 4.

We have extracted the signal yield for LFV $B \rightarrow K_{\mu\mu}$ modes in the similar way as that of LFU $B \rightarrow K\ell\ell$ modes. We found $11.6^{+6.1}_{-5.5}, 1.7^{+3.6}_{-2.2}$, and $-3.3^{+4.0}_{-2.8}$ signal events for $B^+ \rightarrow K^+\mu^+\mu^-$, $B^+ \rightarrow K^+e^+\mu^-$, and $B^0 \rightarrow K_{S0}^{0}\mu^+e^-$ modes, respectively, and the signal enhanced projection ($|\Delta E| < 0.05$ and $O' > 1$) plots in $M_{bc}$ are shown in Figure 5. The upper limit on the signal yield is calculated using the frequentist approach. In this method, we have generated 1000 pseudo experiments using data fitting PDF with different number of signal events $N_{\text{sig}}(\text{gen})$. For each $N_{\text{sig}}(\text{gen})$ we calculated the fraction of experiments which yielded the fitted number of signal events $N_{\text{sig}} < N_{\text{sig}}(\text{observed})$; in turn this gives the 90% confidence level (CL) upper limit on the signal yield. The upper limit on the branching ratio at 90% CL is calculated using MC efficiency and signal
 Entries / (5.55556 MeV/c)
Entries / (5.55556 MeV/c)
Entries / (5.55556 MeV/c)
Entries / (5.55556 MeV/c)

\( \frac{d\mathcal{B}}{dq^2} \) measurement for \( B^+ \to K^\ast \mu \mu \) (upper left), \( B^+ \to K^\ast ee \) (upper right), \( B^0 \to K_S^0 \mu \mu \) (lower left), and \( B^0 \to K_S^0 ee \) (lower right) for \( q^2 \in [0.1, 4.0] \text{ GeV}^2/c^4 \) (blue), \( q^2 \in [4.0, 8.12] \text{ GeV}^2/c^4 \) (blue), \( q^2 \in [1.0 - 6.0] \text{ GeV}^2/c^4 \) (red), \( q^2 \in [10.2 - 12.8] \text{ GeV}^2/c^4 \) (blue), and \( q^2 > 14.18 \text{ GeV}^2/c^4 \) (blue). The band is the theoretical prediction from light-cone sum rule and lattice QCD [14, 15].

\[ \frac{d\mathcal{B}}{dq^2} \] plot of SM dimension fit for \( B^+ \to K^\ast e^-e^+ \) (left), \( B^+ \to K^\ast \mu^-\mu^+ \) (middle), and \( B^0 \to K_S^0 \mu^+\mu^- \) (right). The red, green, and cyan markers show signal, generic \( B \), and continuum events.

\[ M_{bc} \] yield and found to be \( 8.5 \times 10^{-8}, 3.0 \times 10^{-8}, \text{ and } 3.8 \times 10^{-8} \) for \( B^+ \to K^\ast e^-e^+, B^+ \to K^\ast \mu^-\mu^+, \text{ and } B^0 \to K_S^0 \mu^+\mu^- \), respectively.

6. Systematic Uncertainties

The lepton identification systematic uncertainty is calculated using \( J/\psi \to \ell \ell, \ell = e \text{ or } \mu \), and is 0.3% and 0.4% for each muon and electron. The \( K^+ \) and \( K_S^0 \) identification systematic uncertainties are 0.8% and 1.6% using \( D^{*+} \to D^0(K^-\pi^+)\pi^+_s \) and \( D^{*+} \to D^0(K_S^0\pi^0)\pi^+_s \) samples, respectively. The uncertainty due to limited MC statistics is 0.2%. The \( O > O_{min} \) selection
efficiency is calculated in $B \to J/\psi(\ell\ell)K$ and is $0.2 \sim 0.3\%$. The uncertainty due to PDF shape is obtained by varying fixed parameters in the fit within their uncertainties and is $0.1 \sim 0.6\%$. The $B \to K\pi\pi$ component fixed in the fit is varied within $\pm1\sigma$ to give a systematic uncertainty of $0.1 \sim 0.2\%$. The decay model systematic uncertainty is evaluated by comparing the signal MC efficiency from different models. A systematic uncertainty of $0.35\%$ is assigned per track due to track reconstruction efficiency. The systematic uncertainty on the number of $B\bar{B}$ pair is $1.4\%$. The $\mathcal{B}(Y(4S) \to B^{+}B^{-}) (\mathcal{B}(Y(4S) \to B^{0}\bar{B}^{0})$ systematic is $1.2\%$. The systematic uncertainties like hadron identification, track reconstruction, number of $B\bar{B}$ pairs and $\mathcal{B}(Y(4S) \to B^{+}B^{-}) (\mathcal{B}(Y(4S) \to B^{0}\bar{B}^{0})$ cancel out in the double ratio $R_{K}$. For $A_{f}$ the uncertainties due to lepton identification and number of $B\bar{B}$ pairs do not contribute.

7. Conclusion

The $\mathcal{B}(B^{\pm,0} \to J/\psi K^{\pm,0})$ results calculated with 711 fb$^{-1}$ data sample of Belle is the most precise measurement to date. The $R_{K}$ results are found to be consistent with the SM predictions for different $q^{2}$ bins. The $R_{K^{*}}$ value of Belle deviates from LHCb [1, 2] by $1.6\sigma$ for the bin of interest, $q^{2} \in [1.0, 6.0]$ GeV$^{2}$/c$^{4}$. The $A_{f}$ shows negative asymmetry for almost all $q^{2}$ bins and the maximum deviation of $2.6\sigma$ is found for $B \to K\mu\mu$ in $q^{2} \in [1.0, 6.0]$ GeV$^{2}$/c$^{4}$. The $d\mathcal{B}/dq^{2}$ values are consistent with the theoretical prediction for charged $B$ modes and below the expectations for neutral $B$. The upper limit for LFV $B \to K\ell\ell'$ modes are $O(10^{-8})$ at 90% CL. We improved the upper limit of $B^{0} \to K^{0}\mu^{+}\mu^{-}e^{+}e^{-}$ mode by an order of magnitude to $3.8 \times 10^{-8}$.

References


