TOWARDS THE SEARCH FOR LEPTON FLAVOR VIOLATING $B \to K\tau\ell$ DECAYS AT BELLE AND BELLE II*

J.U. Rehman

on behalf of Belle/Belle II Collaboration

Institute of Nuclear Physics Polish Academy of Sciences
Radzikowskiego 152, 31-342 Kraków, Poland

Received 29 March 2026, accepted 20 July 2026,
published online 4 September 2026

The $B \to K\tau\ell$ decays (where $\ell = e, \mu$) are not allowed in the Standard Model (SM) at the tree level due to lepton flavor violation (LFV). In some New Physics (NP) models, their branching fractions are within the reach of the current experimental sensitivities. The Belle and Belle II experiments provide ideal conditions for their search. We used the basic kinematic constraints of the experiments to reconstruct them. A multivariate classifier score is used to model the signal and background. For the estimation of statistical upper limits (ULs), the Bayesian approach is used. To validate the analysis method, two control channel modes are used.

DOI:10.5506/APhysPolBSupp.19.5-A1

1. Introduction

The search for NP signatures includes finding the decays that are not allowed in the SM interactions and are sensitive to the NP effects. The decays $B \to K\tau\ell$ are allowed only at the loop level in the SM; however, the resulting branching fractions are extremely small due to tiny neutrino masses [1]. In the case of the leptoquark model, their predicted branching fractions are of the order of 10^{-7} [2], which are within the reach of the current experimental sensitivity. The presence of third-generation fermions makes them sensitive to NP effects.

The clean environments of the Belle and Belle II experiments, together with the known initial kinematic constraints, enable their search using the basic features of the experiments [3]. Their existing ULs at the confidence level (C.L.) of 90% are in the range of $(0.59\text{–}2.45) \times 10^{-5}$ and are set by the

* Presented at the 32nd Cracow Epiphany Conference on the Recent Results from Heavy Ion Physics, Kraków, Poland, 12–16 January, 2026.

Belle Collaboration, using the hadronic tagging method and mainly considering the leptonic τ-decay modes [4]. In this analysis, we use the semileptonic tagging method [5] to reconstruct the tagging side efficiently, and for the reconstruction of τ, we use decay modes $\tau \to (\pi/\rho)\nu_\tau$.

2. Experimental overview

The Belle II experiment [6] is the successor of the Belle experiment [7] located at SuperKEKB [8] and KEKB [9], respectively. The energies for e^- and e^+ are 8 (7) GeV and 3.5 (4) GeV for Belle (Belle II). The Belle experiment has collected around 1 ab^{-1} of data during its runtime (1999–2010), while the Belle II has started data collection from 2019 onward and so far has collected around 690 fb^{-1} of data. In both experiments, most of the data is collected in the resonance state $\Upsilon(4S)$, which decays to a pair of B mesons.

3. Analysis methodology

The produced B mesons are named B_{sig} and B_{tag}. The B_{sig} is reconstructed as $B^+ \to K^+\tau^\pm\ell^\mp$, where τ is reconstructed in the decay modes $\tau^\pm \to (\pi^\pm/\rho^\pm(\to \pi^\pm\pi^0))\nu_\tau$. Due to the presence of a neutrino on the B_{sig} side, we can reconstruct B_{sig} with a two-fold ambiguity as described in detail in [3]. To reconstruct B_{tag}, we use the inclusive semileptonic tagging method. In this type of tag, we combine a prompt lepton[1] with the remaining hadron system to form B_{tag}. An $\Upsilon(4S)$ candidate is formed by combining B_{sig} and B_{tag}. Some preliminary selections are made to discard the obvious background events. These include rejecting non-physical candidates and candidates from semileptonic B decays on the signal side that can mimic B_{sig}. After selections on the signal side, some loose selections on the B_{tag} side are also made to select only the semileptonic B decays on the tag side.

3.1. Boosted decision tree strategy

For effective separation of signal and background events, the boosted decision tree (BDT) approach is used, which is a machine learning technique that combines multiple decision trees using the different input features of signal and background events [10]. We used one million dedicated signal MC events $B^+ \to K^+\tau^\pm(\to \pi^\pm/\rho^\pm\nu_\tau)\ell^\mp$ to train (0.9 M) and test (0.1 M) the BDT. For background, we used the Belle generic $B\bar{B}$ MC samples, ten times the size of the Belle dataset, to train (nine times) and test (one time) the BDT. The variables used in BDT are $\Delta\cos\theta$ (minimum of the sum of

[1] A lepton produced directly from a B meson.

cosine angles between the signal-side momentum and tag-side momentum [3]), the mass of the hadronic system on the tag side, the total number of photons, the momentum of the tag-side lepton, the difference between the beam energy and the reconstructed B_{tag} energy, and the total number of leptons.

The trained BDT is validated on the 4.0 M dedicated generic signal MC samples $B^+ \to K^+\tau^\pm(\to \text{generic})\ell^\mp$ and the Belle generic $B\bar{B}$ and $q\bar{q}$ MC[2] one time the Belle dataset. The validated BDT score distributions (one such distribution for the $B^+ \to K^+\tau^+e^-$ mode is shown in Fig. 1) provide discrimination between signal and background events.

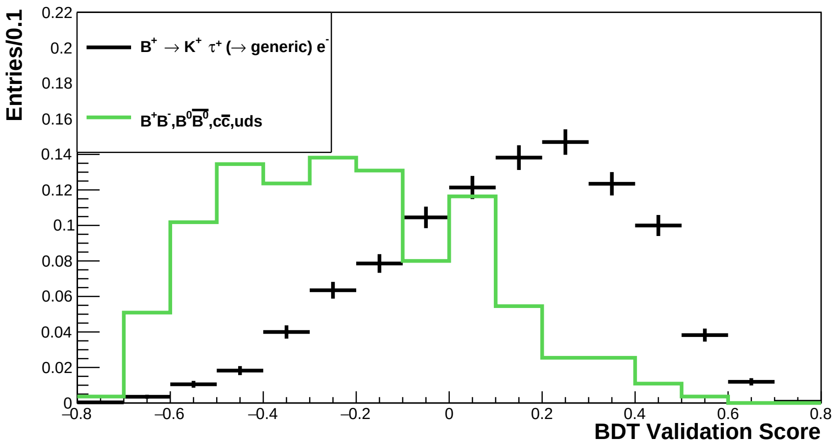

Fig. 1. BDT validation score distribution for $B^+ \to K^+\tau^+e^-$.

3.2. Fitting procedure

We perform a simultaneous fit on the $B^+ \to K^+\tau^\pm(\to \pi^\pm\nu_\tau)\ell^\mp$ and $B^+ \to K^+\tau^\pm(\to \rho^\pm\nu_\tau)\ell^\mp$ samples. The signal probability distribution function (pdf) is formed by combining the pdf for truly reconstructed events and the pdf for events that are misreconstructed as $\tau^\pm \to (\pi^\pm/\rho^\pm)\nu_\tau$ events. The background pdf is formed by adding the pdf for the $B\bar{B}$ background component and pdf for the $q\bar{q}$ background component. The branching fraction (BF) and the number of background events are kept as floating parameters in the fit.

The initial checks of the fit are performed on the MC data, and then the fit stability is checked on an ensemble of a thousand toy dataset. The pull distributions for the fit indicate an overall stable fit with small negative biases. The linearity test for the fit (taking the average value of BF in the

[2] $q\bar{q}$ is not the main background component in this analysis.

range of (0–5) $\times 10^{-5}$) yielded slopes (m) of around unity and y-intercepts (c) of approximately zero. The distribution of the linearity test for the $B^+ \to K^+\tau^-e^+$ mode is shown in Fig. 2.

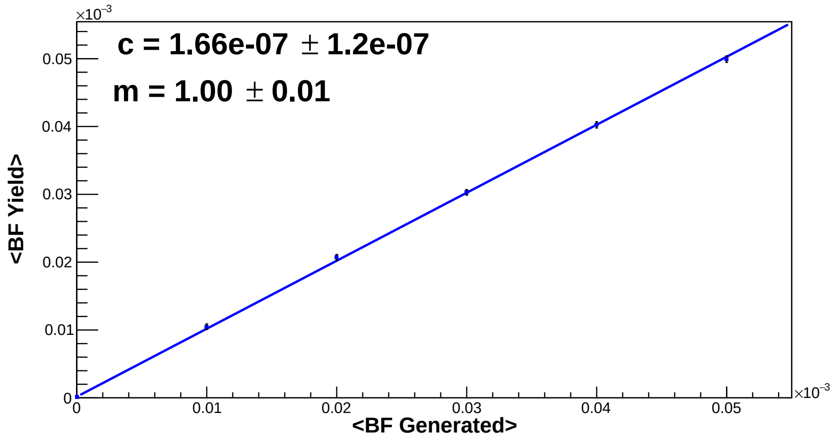

Fig. 2. The distributions of linearity test for the $B^+ \to K^+\tau^-e^+$ mode.

3.3. Statistical upper limits estimation

The Bayesian approach [11] is used to estimate the statistical ULs for the BF. We evaluated the negative log of the profile likelihood ratio (PLR), taking the luminosity five times that of the Belle luminosity[3]. Integration of the PLR over the physical range of BF to 90% of the PLR curve yields the ULs at 90% C.L. The statistical UL is estimated for one thousand toys dataset. A Landau fit is performed on the UL distribution, and the most probable value of the distribution is taken as the estimated statistical UL. The values of these estimated ULs are given in Table 1. It is worth mentioning that at this stage, the uncertainties are not included in the evaluation of these ULs.

Table 1. Estimated statistical ULs for all signal decay modes.

Signal decay mode	Estimated statistical ULs ($\times 10^{-5}$)
$B^+ \to K^+\tau^-e^+$	1.7
$B^+ \to K^+\tau^-\mu^+$	2.0
$B^+ \to K^+\tau^+e^-$	0.9
$B^+ \to K^+\tau^+\mu^-$	1.0

[3] Expected Belle + Belle II integrated luminosity in the near future.

4. Control modes studies

To validate the analysis approach on the actual data, the following two control channel modes are used:

$$B^+ \to J/\psi(\to \mu^+\mu^-)K^+;$$
$$B^+ \to \bar{D}^0(\to K^+\pi^-)\pi^+.$$

Both decays are kinematically similar to the signal decays and are reconstructed similarly to the signal decay modes. We used the complete Belle dataset and the Belle generic MC sample three times the Belle dataset (normalized to the Belle luminosity) for checking the consistency between the data and MC shapes. The agreement of the data and MC shapes is checked in various variable distributions. One such distribution of the variable $\Delta\cos\theta$ is shown in Fig. 3 for the $B^+ \to J/\psi(\to \mu^+\mu^-)K^+$ mode, where we can see a reasonable agreement of the shape between the data and MC. The data to MC ratio for $B^+ \to J/\psi(\to \mu^+\mu^-)K^+$ is 0.84 ± 0.06 and for $B^+ \to \bar{D}^0(\to K^+\pi^-)\pi^+$ is 0.98 ± 0.04.

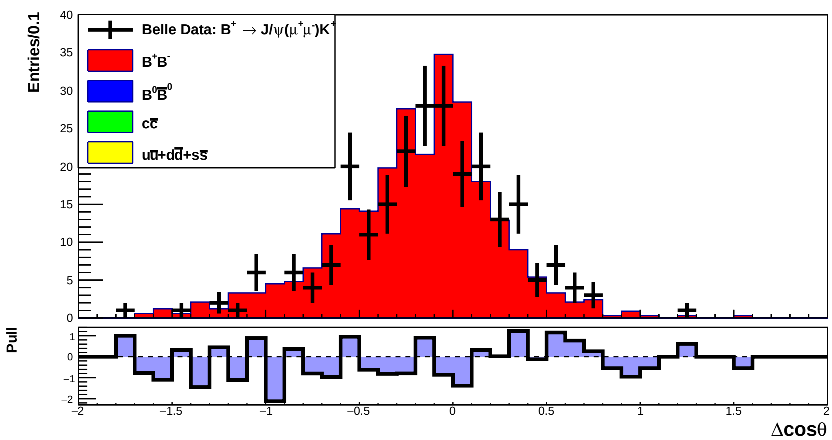

Fig. 3. $\Delta\cos\theta$ distribution for $B^+ \to J/\psi(\to \mu^+\mu^-)K^+$ mode.

5. Summary

The analysis framework for the search for four LFV signal decay modes $B^+ \to K^+\tau^\pm\ell^\mp$ has been developed. The statistical ULs are estimated using the Bayesian approach and are in a similar range as the existing ULs. The validation of the analysis approach is performed on the two control channel modes using the full Belle dataset and the Belle generic MC. The framework and tools developed in the analysis are fully transferable to be used in the same decay modes search in Belle II.

This work is partially supported by the Polish National Agency for Academic Exchange (NAWA) under the STER Programme — Internationalisation of Doctoral Schools, project No. BPI/STE/2023/1/00027/U/00001.

REFERENCES

[1] S. Sahoo, R. Mohanta, *Phys. Rev. D* **93**, 114001 (2016).

[2] A. Angelescu *et al.*, *Phys. Rev. D* **104**, 055017 (2021).

[3] J.U. Rehman, *Acta Phys. Pol. B Proc. Suppl.* **17**, 5-A38 (2024).

[4] S. Watanuki *et al.*, *Phys. Rev. Lett.* **130**, 261802 (2023).

[5] J. Stypula *et al.*, *Phys. Rev. D* **86**, 072007 (2012).

[6] J. Kahn, *CERN Proc.* **1**, 45 (2017).

[7] J. Brodzicka *et al.*, *Prog. Theor. Exp. Phys.* **2012**, 04D001 (2012).

[8] K. Akai, K. Furukawa, H. Koiso, *Nucl. Instrum. Methods Phys. Res. A* **907**, 188 (2018).

[9] S. Kurokawa, E. Kikutani, *Nucl. Instrum. Methods Phys. Res. A* **499**, 1 (2003).

[10] H. Voss, A. Höcker, J. Stelzer, F. Tegenfeldt, *PoS* **(ACAT)**, 040 (2007).

[11] L. Lista, «Statistical Methods for Data Analysis in Particle Physics», in: «Lecture Notes in Physics», Volume 909, *Springer*, Cham 2016.