Search for the Lepton Flavor Violating Decays $B^+ \to K^+\tau^\pm\ell^\mp$ ($\ell = e, \mu$) at Belle

S. Watanuki, G. de Marino, K. Trabelsi, I. Adachi, H. Aihara, D. M. Asner, H. Atmacan, V. Aulchenko, T. Aushev, R. Ayad, V. Babu, Sw. Banerjee, M. Bauer, P. Behera, K. Belous, M. Bessner, V. Bhardwaj, B. Bhuyan, D. Biswas, D. Bodrov, G. Bonvicini, J. Borah, A. Bozek, M. Bračko, P. Branchini, T. E. Browder, A. Budano, M. Campajola, L. Cao, D. Červenkov, M.-C. Chang, B. G. Cheon, K. Chilikin, K. Cho, S.-J. Cho, S.-K. Choi, Y. Choi, S. Choudhury, D. Cinabro, S. Das, G. De Nardo, G. De Pietro, R. Dhamija, F. Di Capua, T. V. Dong, D. Epifanov, T. Ferber, D. Ferlewicz, B. G. Fulsom, R. Garg, V. Gaur, A. Garmash, A. Giri, P. Goldenzweig, E. Graziani, T. Gu, Y. Guan, K. Gudkova, C. Hadjivasiliou, S. Halder, X. Han, T. Hara, K. Hayasaka, H. Hayashii, D. Herrmann, W.-S. Hou, C.-L. Hsu, K. Inami, G. Inguglia, N. Ipsita, A. Ishikawa, R. Itoh, M. Iwasaki, W. W. Jacobs, Q. P. Ji, S. Jia, Y. Jin, K. K. Joo, A. B. Kaliyar, H. Kichimi, C. H. Kim, D. Y. Kim, K.-H. Kim, Y.-K. Kim, K. Kinoshita, P. Kodyš, A. Korobov, S. Korpar, E. Kovalenko, P. Križan, P. Krokovny, T. Kuhr, M. Kumar, K. Kumara, A. Kuzmin, Y.-J. Kwon, J. S. Lange, M. Laurenza, S. C. Lee, P. Lewis, L. K. Li, Y. Li, L. Li Gioi, J. Libby, Y.-R. Lin, D. Liventsev, T. Matsuda, S. K. Maurya, F. Meier, M. Merola, F. Metzner, K. Miyabayashi, R. Mizuk, G. B. Mohanty, M. Nakao, L. Nayak, M. Nayak, N. K. Nisar, S. Nishida, H. Ono, P. Oskin, G. Pakhlova, S. Pardi, H. Park, J. Park, S.-H. Park, A. Passeri, T. K. Pedlar, R. Pestotnik, L. E. Piilonen, T. Podobnik, E. Prencipe, M. T. Prim, M. Röhrken, N. Rout, G. Russo, S. Sandilya, A. Sangal, L. Santelj, V. Savinov, G. Schnell, C. Schwanda, Y. Seino, K. Senyo, M. E. Sevior, W. Shan, M. Shapkin, J.-G. Shiu, B. Shwartz, F. Simon, E. Solovieva, M. Starič, M. Sumihama, T. Sumiyoshi, M. Takizawa, K. Tanida, F. Tenchini, M. Uchida, T. Uglov, Y. Unno, K. Uno, S. Uno, R. van Tonder, G. Varner, K. E. Varvell, D. Wang, E. Wang, M.-Z. Wang, E. Won, X. Xu, B. D. Yabsley, W. Yan, S. B. Yang, J. Yelton, Y. Yusa, Z. P. Zhang, V. Zhilich, and V. Zhukova

(Belle Collaboration)

(Received 13 December 2022; accepted 4 May 2023; published 28 June 2023)

We present a search for the lepton flavor violating decays $B^+ \to K^+\tau^\pm\ell^\mp$, with $\ell = (e, \mu)$, using the full data sample of 772×10^6 $B\bar{B}$ pairs recorded by the Belle detector at the KEKB asymmetric-energy e^+e^- collider. We use events in which one B meson is fully reconstructed in a hadronic decay mode. We find no evidence for $B^\pm \to K^\pm\tau\ell$ decays and set upper limits on their branching fractions at the 90% confidence level in the $(1 - 3) \times 10^{-5}$ range. The obtained limits are the world's best results.

DOI: 10.1103/PhysRevLett.130.261802

Recently, there has been a resurgence of interest in the study of leptoquark fields in light of discrepancies in semileptonic B decays [1], collectively known as B-physics anomalies, which challenge the assumed lepton flavor universality of fundamental interactions. These measurements have been obtained studying two different quark transitions: $b \to c\tau\nu$ [2] and $b \to s\ell\ell$, where $\ell = (e, \mu)$. If confirmed by further measurements, this would be clear evidence of new physics (NP) in which new heavy particles couple preferentially to second and third generation leptons. Many extensions of the standard model that include violation of lepton flavor universality predict lepton flavor violating processes in hadron decays with charged leptons in the final state [3]. In particular, the vector leptoquark U_1 has been identified as the only single-mediator solution [4,5]. In the minimal scenario, U_1 provides an interesting prediction, i.e., lower bounds on the lepton flavor violating $b \to s\tau^\mp\mu^\pm$ decay modes, for example $\mathcal{B}(B \to K\tau\mu) > 0.7 \times 10^{-7}$ [6]. The branching fractions for the two $\ell\tau$ charge combinations are not in general the same, as they depend on the details of the physics mechanism producing the decay.

Upper limits on the branching fractions for $B^+ \to K^+\tau^\pm\ell^\mp$ decays have been previously set at the 90% confidence level (C.L.) using hadronic B tagging by the

BABAR Collaboration between 1.5×10^{-5} and 4.5×10^{-5} [8]; the LHCb Collaboration has studied a single mode, using B^+ mesons from $B_{s2}^{*0} \to B^+ K^-$ decays, setting a limit $\mathcal{B}(B^+ \to K^+ \tau^+ \mu^-) < 3.9 \times 10^{-5}$ at the 90% C.L. [9].

In this Letter, we report a search for $B^+ \to K^+ \tau^\mp \ell^\pm$ decays using the full Belle data sample recorded at the $\Upsilon(4S)$ resonance. The inclusion of the charge-conjugate decay mode is implied. This is the first such search from Belle.

The analysis is based on the full data sample of 772×10^6 $B\bar{B}$ pairs collected with the Belle detector [10] at the KEKB asymmetric-energy e^+e^- collider [11]. The Belle detector is a large-solid-angle spectrometer, which includes a silicon vertex detector, a 50-layer central drift chamber (CDC), an array of aerogel threshold Cherenkov counters (ACC), time-of-flight scintillation counters, and an electromagnetic calorimeter (ECL) composed of CsI(Tl) crystals located inside a superconducting solenoid coil that provides a 1.5 T magnetic field. An iron flux return located outside the coil is instrumented to detect K_L^0 mesons and identify muons.

The analysis procedure is developed using Monte Carlo (MC) simulation based on events generated with EvtGen [12], which includes final-state radiation effects simulated by PHOTOS [13], and the detector response is simulated by GEANT3 [14]. The $B^+ \to K^+ \tau^\pm \ell^\mp$ decays are generated using a uniform three-body phase space model (PHSP); we also consider variations in the linear combinations of the relevant operators for the $b \to s\tau\ell$ transitions $\mathcal{O}_i$ and the relative Wilson coefficients $C_i^{\tau\ell}$, where $i = 9, 10, S, P$ [15].

In each event, we require a fully reconstructed hadronic $B^\pm$ decay, which we refer to as the tagged B-meson candidate or B_{tag}. This is done using the full event interpretation (FEI) algorithm [16], a machine-learning algorithm developed for B-tagged analyses at Belle and Belle II. It supports both hadronic and semileptonic tagging, reconstructing B mesons across more than 4000 individual decay chains. The training is performed in a hierarchical manner: final-state particles are first reconstructed from detector information, then unstable particles (e.g., D, D^*) are built up from these particles, and then reconstruction of B mesons is performed last. For each B_{tag} candidate reconstructed by the FEI, a value of the final multivariate classifier output, Σ_{FEI}, is assigned. Σ_{FEI} is distributed between zero and one, representing candidates identified as being backgroundlike and signallike, respectively.

For the hadronic FEI, the minimal number of tracks per event satisfying certain quality criteria is set to three, as the vast majority of B-meson chains include at least three charged particles, and such a criterion is useful for suppressing background from non-$B\bar{B}$ events. Requirements are placed on the impact parameters to ensure close proximity to the interaction point (IP), less than 0.5 cm in the transverse plane and less than 2.0 cm along the z axis (parallel to the e^+ beam). ECL clusters that are used for γ reconstruction are required to satisfy a region-dependent energy threshold criterion. All the intermediate states (π^0, J/ψ, K_S^0, and $D^{(*)}$ mesons) must pass loose cuts on the reconstructed invariant mass and only the best candidates in terms of Σ_{FEI} are kept. The FEI results in many B_{tag} candidates per event. The number of these candidates is reduced with selections on the beam-energy-constrained mass $M_{\text{bc}} = \sqrt{(E^*_{\text{beam}}/c^2)^2 - (p^*_{B_{\text{tag}}}/c)^2}$, and the energy difference $\Delta E = E^*_{B_{\text{tag}}} - E^*_{\text{beam}}$, where E^*_{beam} is the beam energy, and $E^*_{B_{\text{tag}}}$ and $p^*_{B_{\text{tag}}}$ are the energy and momentum of the B_{tag} candidate in the c.m. rest frame, respectively. The criteria applied are $M_{\text{bc}} > 5.27$ GeV/c^2 and $|\Delta E| < 0.1$ GeV. Finally, the candidate with the highest B_{tag} classifier output, Σ_{FEI}, is selected and a loose requirement $\Sigma_{\text{FEI}} > 0.001$, provides further background rejection with little signal loss.

We then search for the signal $B \to K\tau\ell$ decay in the rest of the event, which we refer to as the signal B-meson candidate or B_{sig}. The notation $B \to K\tau\ell$ refers to one of the following four final states that we consider, where in addition to the kaon of opposite charge to B_{tag} we associate the primary lepton, μ or e: $B^+ \to K^+\tau^+\mu^-$ and $B^+ \to K^+\tau^+e^-$ defined as $OS_{\mu,e}$ modes because the kaon and the primary lepton have opposite charge, and $B^+ \to K^+\tau^-\mu^+$ and $B^+ \to K^+\tau^-e^+$, defined as $SS_{\mu,e}$ modes. In all cases, we require that the τ decays to $\tau \to e\nu\bar{\nu}$, $\tau \to \mu\nu\bar{\nu}$, or $\tau \to \pi\nu$. The combined branching fraction for these decays is 46% [17]. The $\tau \to \rho\nu$ mode, despite not being explicitly reconstructed, significantly contributes to the $\tau \to \pi\nu$ candidates—by roughly one half—because of its large branching fraction (~25%).

We reconstruct $B^+ \to K\tau^\pm\ell^\mp$ decays by selecting three charged particles that originate from the vicinity of the IP and are not associated with the B_{tag}. We require impact parameters less than 1.0 cm in the transverse plane and less than 4.0 cm along the z axis. To reduce backgrounds from low-momentum particles, we require that tracks have a minimum transverse momentum of 100 MeV/c. From the list of selected tracks, we identify K^+ candidates using a likelihood ratio $\mathcal{R}_{K/\pi} = \mathcal{L}_K/(\mathcal{L}_K + \mathcal{L}_\pi)$, where $\mathcal{L}_K$ and $\mathcal{L}_\pi$ are the likelihoods for charged kaons and pions, respectively, calculated based on the number of photoelectrons in the ACC, the specific ionization in the CDC, and the time of flight as determined from the time-of-flight scintillation counters. We select kaons by requiring $\mathcal{R}_{K/\pi} > 0.6$, which has a kaon identification efficiency of 83% and a pion misidentification rate of 5%. Similarly, we select pions by requiring $\mathcal{R}_{\pi/K} > 0.6$, which has a pion identification efficiency of 84% and a kaon misidentification rate of 6%.

Muon candidates are identified based on information from the K_L^0 and muon subdetector (KLM). We require that candidates have a momentum greater than 0.8 GeV/c (enabling them to sufficiently penetrate the KLM), and a penetration depth and degree of transverse scattering consistent with those of a muon [18]. The latter information is used to calculate a normalized muon likelihood ratio $\mathcal{R}_\mu = \mathcal{L}_\mu/(\mathcal{L}_\mu + \mathcal{L}_K + \mathcal{L}_\pi)$, where $\mathcal{L}_\mu$ is the likelihood for muons, for which we require $\mathcal{R}_\mu > 0.9$. For this requirement, the average muon detection efficiency is 89%, with a pion misidentification rate of 1.5% [19].

Electron candidates are required to have a momentum greater than 0.5 GeV/c and are identified using the ratio of ECL cluster energy to the CDC track momentum, the shower shape in the ECL, the matching of the track with the ECL cluster, the specific ionization in the CDC, and the number of photoelectrons in the ACC. This information is used to calculate a normalized electron likelihood ratio $\mathcal{R}_e = \mathcal{L}_e/(\mathcal{L}_e + \mathcal{L}_{\mathrm{hadrons}})$, where $\mathcal{L}_{\mathrm{hadrons}}$ is a product of hadron likelihoods, for which we require $\mathcal{R}_e > 0.9$. This requirement has an efficiency of 92% and a pion misidentification rate below 1% [20].

After selecting one charged kaon, one prompt lepton (electron or muon), and the τ daughter (electron, muon, or pion) with the appropriate charge combination, we require that there are no other tracks than the ones associated to B_{tag} or B_{sig}. The charged kaon and the prompt lepton are uniquely determined to minimize χ^2 of the B_{sig} vertex fit for the prompt tracks. In case there are two possibilities in τ daughter particle identification, τ leptonic decay has priority. Unlike other B decays involving τs (e.g., $B \to \tau\nu$, $B \to D^*\tau\nu$), the $B \to K\tau\ell$ channel has the unique property of having the one or two neutrinos coming from the τ itself, allowing the signal yield to be extracted using the recoil mass, M_{recoil}, which should peak at the mass of the τ lepton.

This variable is easily obtained at B factories, because of the known initial kinematics and the full reconstruction of the other B in the event. In fact, if we consider the B_{sig}, the 4-momentum of the τ can be written as

$$p_\tau = p_{B_{\mathrm{sig}}} - p_K - p_\ell, \tag{1}$$

where $p_{B_{\mathrm{sig}}}$ is not known *a priori*. In the frame where the $\Upsilon(4S)$ resonance is at rest, the two B mesons are back to back, hence

$$\boldsymbol{p}^*_{B_{\mathrm{tag}}} = -\boldsymbol{p}^*_{B_{\mathrm{sig}}}. \tag{2}$$

Furthermore, the two Bs have the same energy, which is half the energy $\sqrt{s}$ of the $\Upsilon(4S)$:

$$E^*_{B_{\mathrm{tag}}} = E^*_{B_{\mathrm{sig}}} = \frac{\sqrt{s}}{2}. \tag{3}$$

In order to obtain the best resolution on the B variables, we replace $E^*_{B_{\mathrm{tag}}}$ with E^*_{beam}, but use the reconstructed $p^*_{B_{\mathrm{tag}}}$ rather than the average value $p^*_{\mathrm{beam}} = \sqrt{E^{*2}_{\mathrm{beam}}/c^2 - m_B^2 c^2}$. Using the Eq. (2) condition and the substitution of $E^*_B = E^*_{\mathrm{beam}}$ in Eq. (1), we obtain

$$\begin{cases} \boldsymbol{p}^*_\tau = -\boldsymbol{p}^*_{B_{\mathrm{tag}}} - \boldsymbol{p}^*_K - \boldsymbol{p}^*_\ell \\ E_\tau = E^*_{\mathrm{beam}} - E^*_K - E^*_\ell \end{cases} \Rightarrow M^2_{\mathrm{recoil}} = m^2_\tau = m^2_B + m^2_{K\ell} - 2(E^*_{\mathrm{beam}}E^*_{K\ell}/c^4 + p^*_{B_{\mathrm{tag}}}p^*_{K\ell}\cos\theta/c^2), \tag{4}$$

where θ is the angle between $\boldsymbol{p}^*_{B_{\mathrm{tag}}}$ and $\boldsymbol{p}^*_{K\ell}$ $(= \boldsymbol{p}^*_K + \boldsymbol{p}^*_\ell)$.

The main source of background consists of Cabibbo-favored transitions from B^+B^- events. For the *OS* configurations, where the primary lepton charge is opposite to the B_{sig} charge, the dominant background comes from semileptonic D decays: $B^+ \to \bar{D}^0(\to K^+\ell^-\bar{\nu}_\ell)X^+$. On the other hand, for the *SS* configurations the primary lepton and the B_{sig} have the same charge and the semileptonic B^+ decays like $B^+ \to \bar{D}^0(\to K^+X^-)X\ell^+\nu_\ell$ provide the three charged particles for the B_{sig} candidates. Events compatible with a $B^+_{\mathrm{sig}} \to \bar{D}^0(\to K^+\pi^-)X^+$ decay are rejected by vetoing candidates in the range $1.81\ \mathrm{GeV/c^2} < m_{K^+t^-} < 1.91\ \mathrm{GeV/c^2}$, where $m_{K^+t^-}$ is the invariant mass of the kaon K^+ and the oppositely charged particle t^-, that can be the prompt lepton or the τ daughter depending on the charge configuration. In the first case, only the $K\tau\mu$ modes show such a D^0 component because of the larger probability to identify a pion as a muon rather than an electron. In the *SS* case, the D^0 peak is much more prominent and relates to the $\tau \to \pi$ mode and is independent of the flavor of the primary lepton.

We further improve the signal selection using a boosted decision tree (BDT) classification. Two classifiers are trained for the background suppression. The first one is optimized to reduce the $B\bar{B}$ events and uses as inputs some kinematic information as well as the topology of the B_{sig} and information on the set of ECL clusters that are not used for the B_{sig} and B_{tag} reconstruction. In particular we use the following: the invariant mass $m_{K^+t^-}$, which helps in suppressing the combinatorial background from charm decays; the number of ECL clusters that are not associated with the reconstructed event and the sum of their energies; the extended Fox-Wolfram moments [21]; the distance from the IP of the signal vertex; and the distance of closest approach between the primary kaon and each of the other two signal tracks. For each mode only the ten most important variables are kept for the final training, the metrics being the information gain provided by each feature in all the decision trees used for the classifier. The threshold

t on the BDT response is optimized using a figure of merit [22], defined as

$$\mathcal{F}(t) = \frac{\epsilon(t)}{\frac{3}{2} + \sqrt{N_{\mathrm{bkg}}(t)}}, \tag{5}$$

where $\epsilon(t)$ is the efficiency for the cut t, N_{bkg} represents the number of background events surviving the cut t in the signal region defined as $1.68\ \mathrm{GeV}/c^2 < M_{\mathrm{recoil}} < 1.87\ \mathrm{GeV}/c^2$, which contains $\sim 80\%$ of the signal events. The MC sample used to estimate the background corresponds to a luminosity of twice that of data. After the cut on the first BDT output, a large fraction of the surviving background is coming from $q\bar{q}$ ($q = u, d, s, c$) events; for this reason a second BDT classifier is trained on these events. The input variables to suppress the continuum background are as follows: event-shape variables such as R_2 (ratio of the 2nd and the 0th Fox-Wolfram moments [23]) and the CLEO cones [24] and the angle θ_T between the thrust axes calculated from final-state particles for the B_{tag} and for the rest of the event in the c.m. frame.

We use control samples in order to evaluate systematic uncertainties related to data and MC discrepancies and to calibrate the signal shape probability density function (PDF) as it is fixed from MC simulation. The first control sample consists of $B^+ \to D^-\pi^+\pi^+$ events, generated in MC as the result of two-body $B^+ \to \bar{D}^{**0}(\to D^-\pi^+)\pi^+$ decays, where $\bar{D}^{**0}$ is either $\bar{D}_0^{*0}$ or $\bar{D}_2^{*0}$, with rates following Refs. [25,26]. This channel has similar topology to our signal as the D can be treated as the τ, allowing for a comparison of the performance of the first BDT classifier between data and MC [Fig. 1 (top) for $B^+ \to K^+\tau^+\mu^-$]. For the calibration of the efficiency of $q\bar{q}$ suppression a second control sample $B \to J/\psi K$ is used because of the similar final state while no usage of the B_{sig} topology is required [Fig. 1 (bottom) for $B^+ \to K^+\tau^+\mu^-$].

The signal yields for $B \to K\tau\ell$ decays are obtained by performing unbinned extended maximum-likelihood fits to the M_{recoil} distributions. The PDF used to model reconstructed signal decays consists of the sum of a reversed crystal ball function to model the main peak and the high-side power-law tail, and a broad Gaussian, with the same mean parameter, to describe the candidates with worse resolution due to imperfect B_{tag} reconstruction. The background events have a smooth shape in the M_{recoil} signal region, and are described by a 2nd-order Chebyshev polynomial. The yields are floated, as well as the background shape parameters while the parameters describing the signal PDF are fixed from the MC simulation. We apply corrections to these parameters to account for small differences between MC simulation and data. These correction factors are obtained from the $B^+ \to \bar{D}^{(*)0}\pi^+$ control samples where M_{recoil} is calculated from the pion from the B^+ decay and the B_{tag}. We validate our fitting procedure and check for fit bias using MC simulation. We generate large ensembles of simulated experiments in which the M_{recoil} distributions are generated from the PDFs used for fitting.

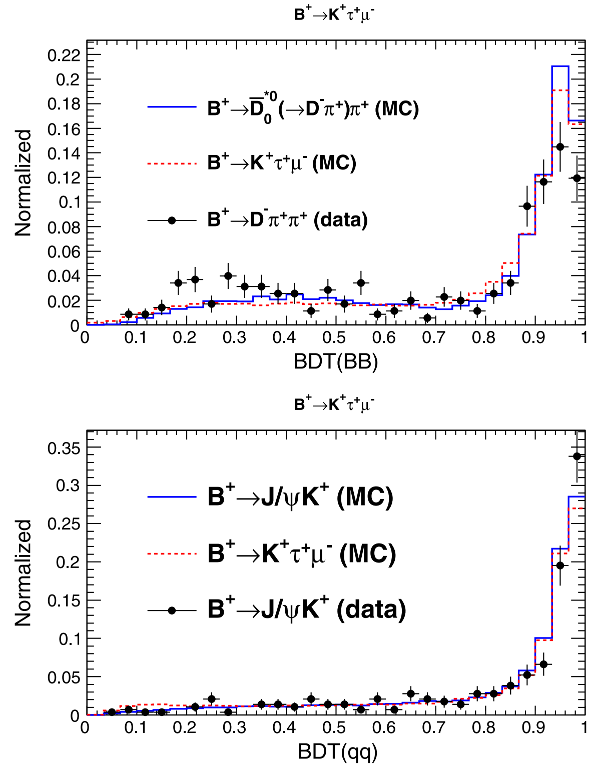

FIG. 1. The BDT response of the $B\bar{B}$ suppression for the $B^+ \to D^-\pi^+\pi^+$ mode (top), and the $q\bar{q}$ suppression for the $B^+ \to J/\psi K^+$ mode (bottom) for the $B^+ \to K^+\tau^+\mu^-$ case.

The M_{recoil} distributions for lepton flavor violating $B \to K\tau\ell$ decays along with projections of the fit result are shown in Fig. 2. The fitted signal yields listed in Table I are consistent with zero for all four modes.

We calculate the upper limit (UL) for these modes at the 90% C.L. using a frequentist method. In this method, for different numbers of signal events $N_{\mathrm{sig}}(\mathrm{gen})$, we generate 10 000 pseudo experiments with signal and background PDFs as obtained in the nominal data fit, with each set of events being statistically equivalent to our data sample of 772×10^6 $B\bar{B}$ pairs. We fit all these simulated datasets, and, for each value of $N_{\mathrm{sig}}(\mathrm{gen})$, we calculate the fraction of MC experiments that have $N_{\mathrm{sig}} \le N_{\mathrm{sig}}(\mathrm{data})$. The 90% C.L. upper limit is taken to be the value of $N_{\mathrm{sig}}(\mathrm{gen})$ (called here $N_{\mathrm{sig}}^{\mathrm{UL}}$) for which 10% of the experiments have $N_{\mathrm{sig}} \le N_{\mathrm{sig}}(\mathrm{data})$. The upper limit on the branching fraction is then derived using the formula

$$\mathcal{B}^{\mathrm{UL}} = \frac{N_{\mathrm{sig}}^{\mathrm{UL}}}{N_{B\bar{B}} \times 2 \times f^{+-} \times \varepsilon},$$

where $N_{B\bar{B}}$ is the number of $B\bar{B}$ pairs $= (772 \pm 11) \times 10^6$, f^{+-} is the branching fraction $\mathcal{B}(\Upsilon(4S) \to B^+B^-)$ for

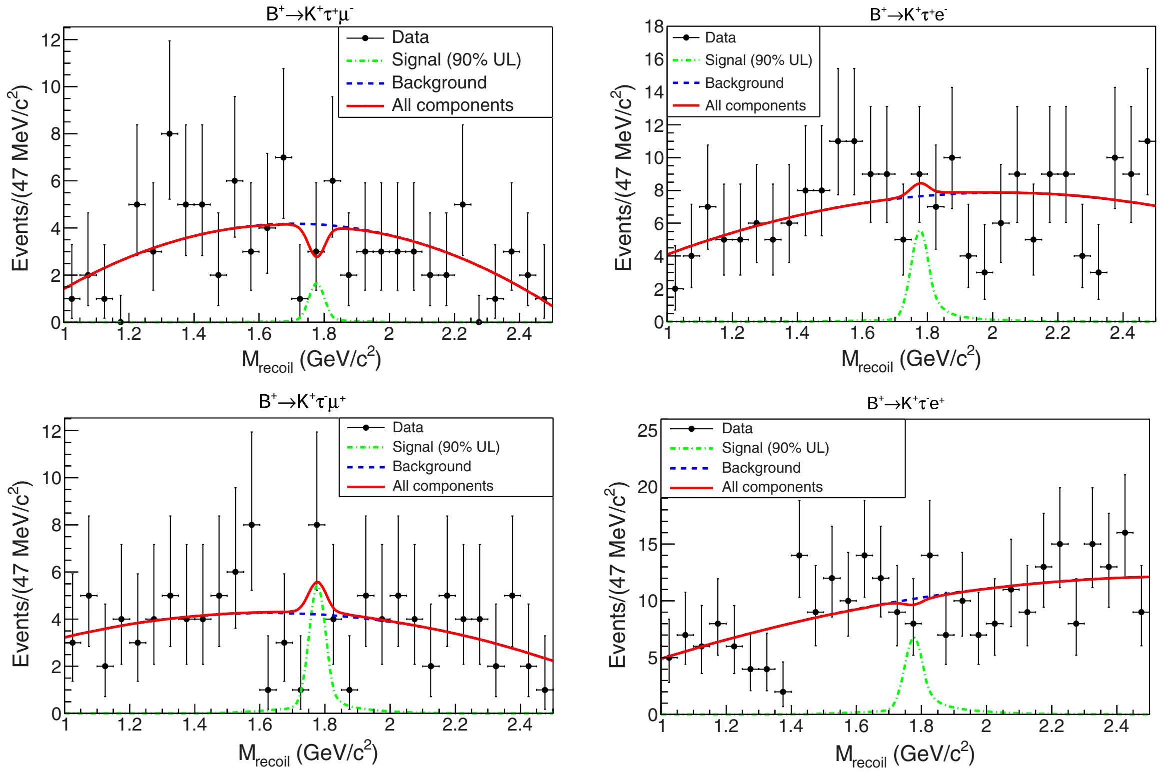

FIG. 2. Observed M_{recoil} distributions for the four $B \to K\tau\ell$ modes, along with projections of the fit result. The black dots show the data, the dashed blue curve shows the background component, and the solid red curve shows the overall fit result. The dash-dotted green curve shows the signal PDF, with a normalization corresponding to the 90% C.L. upper limit.

charged B decays (using 0.514 ± 0.006 [17]), and ε is the signal reconstruction efficiency. By default ε is obtained with signal PHSP MC samples [5], while we also consider a NP model with a combination of the effective operators $\mathcal{O}_{S,P}$ by reweighting the $q^2 = m^2_{\tau\ell}$ distribution which gives the smallest efficiency. The systematic uncertainty in $\mathcal{B}^{\text{UL}}$ is included by smearing the N_{sig} distribution obtained from the MC fits with the fractional systematic uncertainty. The results are listed in Table I.

The systematic uncertainties in our measurements are listed in Table II, where additive uncertainties arise from the signal yield, while multiplicative uncertainties are from the efficiency. Uncertainties in the shape of the PDFs used for the signal are evaluated by varying all fixed parameters by $\pm 1\sigma$, including the correction factors to the shapes obtained

TABLE I. Efficiencies, fit yields, and branching fraction expected or observed upper limits at the 90% C.L. for PHSP case. In parentheses the observed limits for the NP case.

Mode	$\varepsilon(\%)$	$\varepsilon^{\text{NP}}(\%)$	N_{sig}	$\mathcal{B}^{\text{UL}}(10^{-5})$
$B^+ \to K^+\tau^+\mu^-$	0.064	0.058	-2.1 ± 2.9	1.18/0.59 (0.65)
$B^+ \to K^+\tau^+e^-$	0.084	0.074	1.5 ± 5.5	1.34/1.51 (1.71)
$B^+ \to K^+\tau^-\mu^+$	0.046	0.038	2.3 ± 4.1	1.81/2.45 (2.97)
$B^+ \to K^+\tau^-e^+$	0.079	0.058	-1.1 ± 7.4	2.29/1.53 (2.08)

TABLE II. Contributions to the systematic uncertainties of the measured signal yields and branching fractions.

Source	$K^+\tau^+\mu^-$	$K^+\tau^+e^-$	$K^+\tau^-\mu^+$	$K^+\tau^-e^+$
Additive (events)				
PDF shape (mean)	0.09	0.01	0.08	0.08
PDF shape (width)	0.02	0.08	0.04	0.07
PDF shape (f_{sig})	0.28	0.16	0.11	0.16
Linearity	0.03	0.04	0.02	0.04
Total	0.30	0.18	0.14	0.20
Multiplicative (%)				
B_{tag} calibration	5.9	5.9	5.9	5.9
Track reconstruction	1.1	1.1	1.1	1.1
Kaon identification	1.3	1.4	1.3	1.3
Lepton identification	0.3	0.4	0.3	0.4
τ daughter identification	0.7	0.7	0.6	0.6
MC statistics	1.0	1.5	1.2	1.0
Number of $B\bar{B}$ pairs	1.4	1.4	1.4	1.4
BDT $B\bar{B}$ selection	10.6	10.0	12.7	12.6
BDT $q\bar{q}$ selection	8.8	8.6	9.2	6.6
f^{+-}	1.2	1.2	1.2	1.2
Total	15.3	14.8	17.0	15.7

from the $B^+ \to \bar{D}^{(*)0}\pi^+$ control samples, and varying the fraction of the Gaussian ($f_{\rm sig}$) by 10%. The resulting change in the signal yield is taken as the systematic uncertainty. The uncertainty related to the choice of the background PDF is evaluated and found to be negligible. Although no significant difference is observed between simulated and measured $N_{\rm sig}$ values, the uncertainties on the parameters of the linear regressions are used to evaluate the corresponding systematic uncertainties. The reconstruction efficiency for $B_{\rm tag}$ evaluated via MC simulation is corrected to account for differences between MC and data in the branching fractions and models used for hadronic B decays. This correction is evaluated by comparing the number of events containing both a $B_{\rm tag}$ and a semileptonic $B \to D\ell\nu$ [27]. The resulting correction factor is $85 \pm 5\%$ and the uncertainty in this value is taken as a systematic uncertainty.

The systematic uncertainty due to the charged track reconstruction is evaluated using $D^{*+} \to D^0\pi^+$ with $D^0 \to K_S^0\pi^+\pi^-$, resulting in an uncertainty of 0.35% per track. Uncertainties due to K^+ and π^+ (for $\tau \to \pi\nu$ mode) identification is 1.3%, as measured with a $D^{*+} \to D^0(K^-\pi^+)\pi^+$ sample. The uncertainty due to lepton identification is evaluated using $J/\psi \to \ell^+\ell^-$ events, resulting in an uncertainty of 0.3% for muons and 0.4% for electrons. The systematic uncertainty arising from the number of $B\bar{B}$ pairs is 1.4%. We compare the efficiency of the BDT selection between data and MC samples with the control channel $B^+ \to D^-\pi^+\pi^+$ for $B\bar{B}$ suppression and $B^+ \to J/\psi K^+$ for continuum suppression, the differences between data and MC simulation are assigned as a systematic uncertainty. We use a systematic uncertainty of 1.2% in the fraction f^{+-} [17].

We have searched for the lepton flavor violating decays $B^+ \to K^+\tau^\pm\ell^\mp$ using the full Belle dataset. We find no evidence for these decays and set the following upper limits on the branching fractions at the 90% C.L.:

$$\begin{aligned}
&\mathcal{B}(B^+ \to K^+\tau^+\mu^-) < 0.59 \times 10^{-5}\\
&\mathcal{B}(B^+ \to K^+\tau^+e^-) < 1.51 \times 10^{-5}\\
&\mathcal{B}(B^+ \to K^+\tau^-\mu^+) < 2.45 \times 10^{-5}\\
&\mathcal{B}(B^+ \to K^+\tau^-e^+) < 1.53 \times 10^{-5}. \qquad (6)
\end{aligned}$$

Our results are the most stringent limits to date.

This work, based on data collected using the Belle detector, which was operated until June 2010, was supported by the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan, the Japan Society for the Promotion of Science (JSPS), and the Tau-Lepton Physics Research Center of Nagoya University; the Australian Research Council including Grants No. DP180102629, No. DP170102389, No. DP170102204, No. DE220100462, No. DP150103061, No. FT130100303; Austrian Federal Ministry of Education, Science and Research (FWF) and FWF Austrian Science Fund No. P 31361-N36; the National Natural Science Foundation of China under Contracts No. 11675166, No. 11705209; No. 11975076; No. 12135005; No. 12175041; No. 12161141008; Key Research Program of Frontier Sciences, Chinese Academy of Sciences (CAS), Grant No. QYZDJ-SSW-SLH011; Project No. ZR2022JQ02 supported by Shandong Provincial Natural Science Foundation; the Ministry of Education, Youth and Sports of the Czech Republic under Contract No. LTT17020; the Czech Science Foundation Grant No. 22-18469S; Horizon 2020 ERC Advanced Grant No. 884719 and ERC Starting Grant No. 947006 "InterLeptons" (European Union); the Carl Zeiss Foundation, the Deutsche Forschungsgemeinschaft, the Excellence Cluster Universe, and the Volkswagen-Stiftung; the Department of Atomic Energy (Project Identification No. RTI 4002) and the Department of Science and Technology of India; the Istituto Nazionale di Fisica Nucleare of Italy; National Research Foundation (NRF) of Korea Grants No. 2016R1D1A1B02012900, No. 2018R1A2B3003643, No. 2018R1A6A1A06024970, No. RS202200197659, No. 2019R1I1A3A01058933, No. 2021R1A6A1A03043957, No. 2021R1F1A1060423, No. 2021R1F1A1064008, No. 2022R1A2C1003993; Radiation Science Research Institute, Foreign Large-size Research Facility Application Supporting project, the Global Science Experimental Data Hub Center of the Korea Institute of Science and Technology Information and KREONET/GLORIAD; the Polish Ministry of Science and Higher Education and the National Science Center; the Ministry of Science and Higher Education of the Russian Federation, Agreement 14.W03.31.0026, and the HSE University Basic Research Program, Moscow; University of Tabuk research Grants S-1440-0321, S-0256-1438, and S-0280-1439 (Saudi Arabia); the Slovenian Research Agency Grants No. J1-9124 and No. P1-0135; Ikerbasque, Basque Foundation for Science, Spain; the Swiss National Science Foundation; the Ministry of Education and the Ministry of Science and Technology of Taiwan; and the United States Department of Energy and the National Science Foundation. These acknowledgements are not to be interpreted as an endorsement of any statement made by any of our institutes, funding agencies, governments, or their representatives. We thank the KEKB group for the excellent operation of the accelerator; the KEK cryogenics group for the efficient operation of the solenoid; and the KEK computer group and the Pacific Northwest National Laboratory (PNNL) Environmental Molecular Sciences Laboratory (EMSL) computing group for strong computing support; and the National Institute of Informatics, and Science Information NETwork 6 (SINET6) for valuable network support.

[1] R. Aaij *et al.* (LHCb Collaboration), Nat. Phys. **18**, 277 (2022).
[2] Y. Amhis *et al.*, Phys. Rev. D **107**, 052008 (2023).
[3] S. L. Glashow, D. Guadagnoli, and K. Lane, Phys. Rev. Lett. **114**, 091801 (2015).
[4] A. Angelescu, D. Bečirević, D. A. Faroughy, F. Jaffredo, and O. Sumensari, Phys. Rev. D **104**, 055017 (2021).
[5] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.130.261802 for the diagram illustrating the leptoquark-mediated $B \to K\tau\ell$ decay and for the distributions of the selection efficiency as a function of the kinematic variables $M^2(\tau\ell)$ and $M^2(K\ell)$. These distributions can be used to recast the results for models having kinematics different from the uniform distribution in the phase space used for the signal simulation.
[6] The $R_{K^{(*)}}$ observables, which have long participated in the B anomalies, were recently found to be consistent with the standard model [7]. This may alter the mentioned lepton flavor violating predictions.
[7] LHCb Collaboration, arXiv:2212.09152.
[8] J. P. Lees *et al.* (*BABAR* Collaboration), Phys. Rev. D **86**, 012004 (2012).
[9] R. Aaij *et al.* (LHCb Collaboration), J. High Energy Phys. 06 (2020) 129.
[10] A. Abashian *et al.* (Belle Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **479**, 117 (2002).
[11] S. Kurokawa and E. Kikutani, Nucl. Instrum. Methods Phys. Res., Sect. A **499**, 1 (2003), and other papers in this volume; T. Abe *et al.*, Prog. Theor. Exp. Phys. **2013**, 03A001 (2013) and following articles up to 03A011.
[12] D. J. Lange, Nucl. Instrum. Methods Phys. Res., Sect. A **462**, 152 (2001).
[13] E. Barberio, B. van Eijk, and Z. Wąs, Comput. Phys. Commun. **66**, 115 (1991).
[14] R. Brun, F. Bruyant, M. Maire, A. C. McPherson, and P. Zanarini, *GEANT3: User's Guide GEANT 3.10, GEANT 3.11; rev. version* (CERN, Geneva, 1987).
[15] D. Bečirević, O. Sumensari, and R. Zukanovich Funchal, Eur. Phys. J. C **76**, 134 (2016).
[16] T. Keck *et al.*, Comput. Software Big Sci. **1**, 6 (2019); T. Kuhr, C. Pulvermacher, M. Ritter, T. Hauth, and N. Braun (Belle II Framework Software Group), Comput. Software Big Sci. **3**, 1 (2019); M. Gelb *et al.*, Comput. Software Big Sci. **2**, 9 (2018).
[17] R. L. Workman *et al.* (Particle Data Group), Prog. Theor. Exp. Phys. **2022**, 083C01 (2022).
[18] A. Abashian *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **491**, 69 (2002).
[19] E. Nakano, Nucl. Instrum. Methods Phys. Res., Sect. A **494**, 402 (2002).
[20] K. Hanagaki, H. Kakuno, H. Ikeda, T. Iijima, and T. Tsukamoto, Nucl. Instrum. Methods Phys. Res., Sect. A **485**, 490 (2002).
[21] S. H. Lee *et al.* (Belle Collaboration), Phys. Rev. Lett. **91**, 261801 (2003).
[22] G. Punzi, eConf **C030908**, MODT002 (2003).
[23] G. C. Fox and S. Wolfram, Phys. Rev. Lett. **41**, 1581 (1978).
[24] D. M. Asner *et al.* (CLEO Collaboration), Phys. Rev. D **53**, 1039 (1996).
[25] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. D **79**, 112004 (2009).
[26] K. Abe *et al.* (Belle Collaboration), Phys. Rev. D **69**, 112002 (2004).
[27] M. Gelb *et al.* (Belle Collaboration), Phys. Rev. D **98**, 112016 (2018).