Search for $B^0 \rightarrow \tau^\pm \tau^\mp$ ($\ell = e, \mu$) with a hadronic tagging method at Belle

H. Atmacan,7 A. J. Schwartz,7 K. Kinoshita,7 I. Adachi,10,15 K. Adamczyk,64 H. Aihara,88 S. Al Said,81,40 D. M. Asner,3 V. Aulchenko,67 T. Ashe,20 R. Ayad,81 V. Babu,8 S. Bahinipati,24 M. Bauer,37 P. Behera,27 K. Belous,31 J. Bennett,55 F. Bernlochner,2 M. Bessner,18 V. Bhardwaj,23 B. Bhuyan,25 T. Bilka,5 J. Biswal,36 A. Bobrov,4,67 A. Bozek,64 M. Bračko,52,36 P. Branchini,33 T. E. Browder,18 A. Budano,33 M. Campajola,32,59 L. Cao,20 R. Ayad,81 V. Ba bu,8 S. Bahinipati,24 M. Bauer,37 P. Behera,27 K. Belous,31 J. Bennett,55 F. Bernlochner,2 M. Bessner,18 V. Bhardwaj,23 B. Bhuyan,25 T. Bilka,5 J. Biswal,36 A. Bobrov,4,67 A. Bozek,64 M. Bračko,52,36 P. Branchini,33 T. E. Browder,18 A. Budano,33 M. Campajola,32,59 L. Cao,20 R. Ayad,81 V. Ba bu,8 S. Bahinipati,24 M. Bauer,37 P. Behera,27 K. Belous,31 J. Bennett,55 F. Bernlochner,2 M. Bessner,18 V. Bhardwaj,23 B. Bhuyan,25 T. Bilka,5 J. Biswal,36 A. Bobrov,4,67 A. Bozek,64 M. Bračko,52,36 P. Branchini,33 T. E. Browder,18 A. Budano,33 M. Campajola,32,59 L. Cao,20 R. Ayad,81 V. Ba bu,8 S. Bahinipati,24 M. Bauer,37 P. Behera,27 K. Belous,31 J. Bennett,55 F. Bernlochner,2 M. Bessner,18 V. Bhardwaj,23 B. Bhuyan,25 T. Bilka,5 J. Biswal,36 A. Bobrov,4,67 A. Bozek,64 M. Bračko,52,36 P. Branchini,33 T. E. Browder,18 A. Budano,33 M. Campajola,32,59 L. Cao,20 R. Ayad,81 V. Ba bu,8 S. Bahinipati,24 M. Bauer,37 P. Behera,27 K. Belous,31 J. Bennett,55 F. Bernlochner,2 M. Bessner,18 V. Bhardwaj,23 B. Bhuyan,25 T. Bilka,5 J. Biswal,36 A. Bobrov,4,67 A. Bozek,64 M. Bračko,52,36 P. Branchini,33 T. E. Browder,18 A. Budano,33 M. Campajola,32,59 L. Cao,20 R. Ayad,81 V. Ba bu,8 S. Bahinipati,24 M. Bauer,37 P. Behera,27 K. Belous,31 J. Bennett,55 F. Bernlochner,2 M. Bessner,18 V. Bhardwaj,23 B. Bhuyan,25 T. Bilka,5 J. Biswal,36 A. Bobrov,4,67 A. Bozek,64 M. Bračko,52,36 P. Branchini,33 T. E. Browder,18 A. Budano,33 M. Campajola,32,59 L. Cao,20 R. Ayad,81 V. Ba bu,8 S. Bahinipati,24 M. Bauer,37 P. Behera,27 K. Belous,31 J. Bennett,55 F. Bernlochner,2 M. Bessner,18 V. Bhardwaj,23 B. Bhuyan,25 T. Bilka,5 J. Biswal,36 A. Bobrov,4,67 A. Bozek,64 M. Bračko,52,36 P. Branchini,33 T. E. Browder,18 A. Budano,33 M. Campajola,32,59 L. Cao,20 R. Ayad,81 V. Ba bu,8 S. Bahinipati,24 M. Bauer,37 P. Behera,27 K. Belous,31 J. Bennett,55 F. Bernlochner,2 M. Bessner,18 V. Bhardwaj,23 B. Bhuyan,25 T. Bilka,5 J. Biswal,36 A. Bobrov,4,67 A. Bozek,64 M. Bračko,52,36 P. Branchini,33 T. E. Browder,18 A. Budano,33 M. Campajola,32,59 L. Cao,20 R. Ayad,81 V. Ba bu,8 S. Bahinipati,24 M. Bauer,37 P. Behera,27 K. Belous,31 J. Bennett,55 F. Bernlochner,2 M. Bessner,18 V. Bhardwaj,23 B. Bhuyan,25 T. Bilka,5 J. Biswal,36 A. Bobrov,4,67 A. Bozek,64 M. Bračko,52,36 P. Branchini,33 T. E. Browder,18 A. Budano,33 M. Campajola,32,59 L. Cao,20 R. Ayad,81 V. Ba bu,8 S. Bahinipati,24 M. Bauer,37 P. Behera,27 K. Belous,31 J. Bennett,55 F. Bernlochner,2 M. Bessner,18 V. Bhardwaj,23 B. Bhuyan,25 T. Bilka,5 J. Biswal,36 A. Bobrov,4,67 A. Bozek,64 M. Bračko,52,36 P. Branchini,33 T. E. Browder,18 A. Budano,33 M. Campajola,32,59 L. Cao,20 R. Ayad,81 V. Ba bu,8 S. Bahinipati,24 M. Bauer,37 P. Behera,27 K. Belous,31 J. Bennett,55 F. Bernlochner,2 M. Bessner,18 V. Bhardwaj,23 B. Bhuyan,25 T. Bilka,5 J. Biswal,36 A. Bobrov,4,67 A. Bozek,64 M. Bračko,52,36 P. Branchini,33 T. E. Browder,18 A. Budano,33 M. Campajola,32,59 L. Cao,20 R. Ayad,81 V. Ba bu,8 S. Bahinipati,24 M. Bauer,37 P. Behera,27 K. Belous,31 J. Bennett,55 F. Bernlochner,2 M. Bessner,18 V. Bhardwaj,23 B. Bhuyan,25 T. Bilka,5 J. Biswal,36 A. Bobrov,4,67 A. Bozek,64 M. Bračko,52,36 P. Branchini,33 T. E. Browder,18 A. Budano,33 M. Campajola,32,59 L. Cao,20 R. Ayad,81 V. Ba bu,8 S. Bahinipati,24 M. Bauer,37 P. Behera,27 K. Belous,31 J. Bennett,55 F. Bernlochner,2 M. Bessner,18 V. Bhardwaj,23 B. Bhuyan,25 T. Bilka,5 J. Biswal,36 A. Bobrov,4,67 A. Bozek,64 M. Bračko,52,36 P. Branchini,33 T. E. Browder,18 A. Budano,33 M. Campajola,32,59 L. Cao,20 R. Ayad,81 V. Ba bu,8 S. Bahinipati,24 M. Bauer,37 P. Behera,27 K. Belous,31 J. Bennett,55 F. Bernlochner,2 M. Bessner,18 V. Bhardwaj,23 B. Bhuyan,25 T. Bilka,5 J. Biswal,36 A. Bobrov,4,67 A. Bozek,64 M. Bračko,52,36 P. Branchini,33 T. E. Browder,18 A. Budano,33 M. Campajola,32,59 L. Cao,20 R. Ayad,81 V. Ba bu,8 S. Bahinipati,24 M. Bauer,37 P. Behera,27 K. Belous,31 J. Bennet
| Page | Institution/Location |
|------|---------------------|
| 24   | Indian Institute of Technology Bhubaneswar, Satya Nagar 751007 |
| 25   | Indian Institute of Technology Guwahati, Assam 781039 |
| 26   | Indian Institute of Technology Hyderabad, Telangana 502285 |
| 27   | Indian Institute of Technology Madras, Chennai 600036 |
| 28   | Indiana University, Bloomington, Indiana 47408 |
| 29   | Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049 |
| 30   | Institute of High Energy Physics, Vienna 1050 |
| 31   | Institute for High Energy Physics, Protvino 142281 |
| 32   | INFN—Sezione di Napoli, I-80126 Napoli |
| 33   | INFN—Sezione di Roma Tre, I-00146 Rome |
| 34   | INFN—Sezione di Torino, I-10125 Torino |
| 35   | Advanced Science Research Center, Japan Atomic Energy Agency, Naka 319-1195 |
| 36   | J. Stefan Institute, 1000 Ljubljana |
| 37   | Institut für Experimentelle Teilchenphysik, Karlsruher Institut für Technologie, 76131 Karlsruhe |
| 38   | Kavli Institute for the Physics and Mathematics of the Universe (WPI), University of Tokyo, Kashiwa 277-8583 |
| 39   | Kennesaw State University, Kennesaw, Georgia 30144 |
| 40   | Department of Physics, Faculty of Science, King Abdulaziz University, Jeddah 21589 |
| 41   | Kitasato University, Sagamihara 252-0373 |
| 42   | Korea Institute of Science and Technology Information, Daejeon 34141 |
| 43   | Korea University, Seoul 02841 |
| 44   | Kyoto Sangyo University, Kyoto 603-8555 |
| 45   | Kyungpook National University, Daegu 41566 |
| 46   | Université Paris-Saclay, CNRS/IN2P3, IJCLab, 91405 Orsay |
| 47   | P.N. Lebedev Physical Institute of the Russian Academy of Sciences, Moscow 119991 |
| 48   | Faculty of Mathematics and Physics, University of Ljubljana, 1000 Ljubljana |
| 49   | Ludwig Maximilians University, 80539 Munich |
| 50   | Luther College, Decorah, Iowa 52101 |
| 51   | Malaviya National Institute of Technology Jaipur, Jaipur 302017 |
| 52   | Faculty of Chemistry and Chemical Engineering, University of Maribor, 2000 Maribor |
| 53   | Max-Planck-Institut für Physik, 80805 München |
| 54   | School of Physics, University of Melbourne, Victoria 3010 |
| 55   | University of Mississippi, University, Mississippi 38677 |
| 56   | University of Miyazaki, Miyazaki 889-2192 |
| 57   | Moscow Physical Engineering Institute, Moscow 115409 |
| 58   | Graduate School of Science, Nagoya University, Nagoya 464-8602 |
| 59   | Università di Napoli Federico II, I-80126 Napoli |
| 60   | Nara Women’s University, Nara 630-8506 |
| 61   | National Central University, Chung-li 32054 |
| 62   | National United University, Miao Li 36003 |
| 63   | Department of Physics, National Taiwan University, Taipei 10617 |
| 64   | H. Niewodniczanski Institute of Nuclear Physics, Krakow 31-342 |
| 65   | Nippon Dental University, Niigata 951-8580 |
| 66   | Niigata University, Niigata 950-2181 |
| 67   | Novosibirsk State University, Novosibirsk 630090 |
| 68   | Osaka City University, Osaka 558-8585 |
| 69   | Pacific Northwest National Laboratory, Richland, Washington 99352 |
| 70   | Panjab University, Chandigarh 160014 |
| 71   | Peking University, Beijing 100871 |
| 72   | University of Pittsburgh, Pittsburgh, Pennsylvania 15260 |
| 73   | Department of Physics, State Key Laboratory of Particle Detection and Electronics, University of Science and Technology of China, Hefei 230026 |
| 74   | Seoul National University, Seoul 08826 |
| 75   | Sogang University, Seoul 06978 |
| 76   | Université de Strasbourg, CNRS, IPHC, UMR 7178, 67037 Strasbourg |
| 77   | Sungkyunkwan University, Suwon 16419 |
| 78   | School of Physics, University of Sydney, New South Wales 2006 |
| 79   | Department of Physics, Faculty of Science, University of Tabuk, Tabuk 71451 |
The lepton-flavor-violating decays $B^0 \rightarrow \tau^\pm \ell^\mp$, where $\ell = (e, \mu)$, are promising modes in which to search for new physics. Recently, there have been indications of possible violation of lepton flavor universality (LFU) in $B^0 \rightarrow D^{(*)-}\tau^+\nu$ [2], $B^0 \rightarrow K^{(*)0}\ell^+\ell^-$ [3], and $B^0 \rightarrow K^{\pm}\ell^+\ell^-$ [4,5] decays. Other studies are less conclusive [6,7]. LFU violation is often accompanied by lepton flavor violation (LFV) in theoretical models [8]. The decay $B^0 \rightarrow \tau^\pm\ell^\mp$, like $B^0 \rightarrow D^{(*)-}\tau^+\nu$, connects a third-generation quark with a third-generation lepton. The decay can occur in principle via neutrino mixing [9]; however, the rate due to such mixing [10] is considerably below current or future experimental sensitivities. Thus, observing these decays would indicate new physics. Some new physics models give rise to branching fractions of $10^{-9}$ to $10^{-10}$. For example, Pati-Salam vector leptoquarks of mass 86 TeV/$c^2$ give branching fractions of 4.4 $\times$ 10$^{-9}$ for $B^0 \rightarrow \tau^\pm\ell^\mp$ and 1.6 $\times$ 10$^{-9}$ for $B^0 \rightarrow \tau^\pm\ell^\mp$ [11]. The general flavor-universal minimal supersymmetric Standard Model predicts branching fractions of up to about 2 $\times$ 10$^{-10}$ [12].

These decay modes have previously been studied by the CLEO [13], BABAR [14], and LHCb [15] experiments. No evidence for these decays has been found. The current most stringent upper limits are $B(B^0 \rightarrow \tau^\pm\mu^\mp) < 1.2 \times 10^{-5}$ [15] and $B(B^0 \rightarrow \tau^\pm\ell^\mp) < 2.8 \times 10^{-5}$ [14], both at 90% confidence level (CL). In this paper we report a search for $B^0 \rightarrow \tau^\pm\ell^\mp$ decays using the full Belle data sample of 711 fb$^{-1}$ recorded at the $\Upsilon(4S)$ resonance. This is the first such search from Belle.

II. DATASET AND DETECTOR DESCRIPTION

Our data sample consists of $(772 \pm 11) \times 10^6$ $B\bar{B}$ pairs produced in $e^+e^- \rightarrow \Upsilon(4S)$ events recorded by the Belle detector at the KEKB asymmetric-energy $e^+e^-$ collider [16]. The Belle detector is a large-solid-angle magnetic spectrometer that consists of a silicon vertex detector (SV), a 50-layer central drift chamber (CDC), an array of aerogel threshold Cherenkov counters (ACC), a barrel-like arrangement of time-of-flight (TOF) scintillation counters, and an electromagnetic calorimeter comprising CsI(Tl) crystals (ECL). All these detectors are located inside a superconducting solenoid coil that provides a 1.5 T magnetic field. An iron flux return yoke located outside the coil is instrumented with resistive-plate chambers (KLM) to detect $K^0_S$ mesons and to identify muons. Two inner detector configurations were used: for the first 152 $\times 10^6$ $B\bar{B}$ pairs, a 2.0 cm radius beam pipe and a three-layer SV were used; and for the remaining $620 \times 10^6$ $B\bar{B}$ pairs, a 1.5 cm radius beam pipe, a four-layer SV [17], and a small-cell inner drift chamber were used. A more detailed description of the detector is provided in Ref. [18].

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI. Funded by SCOAP3®.
We study properties of signal events, sources of background, and optimize selection criteria using Monte Carlo (MC) simulated events. These samples are generated using the software packages EVGEN [19] and PYTHIA [20], and final-state radiation is included via PHOTOS [21]. The detector response is simulated using GEANT3 [22]. We produce $B^0 \to \tau^\pm \ell^\mp$ MC events to calculate signal reconstruction efficiencies. To estimate backgrounds, we use MC samples that describe all $e^+e^- \to q\bar{q}$ processes. Events containing $e^+e^- \to BB$ with subsequent $b \to cW$ decay, and $e^+e^- \to q\bar{q}$ ($q = u, d, s, c$) continuum events, are both simulated with five times the integrated luminosity of Belle. Semileptonic $b \to u\ell\nu$ decays are simulated with 20 times the integrated luminosity. Rare $b \to s$ and $b \to u$ decays are simulated with 50 times the integrated luminosity.

### III. EVENT SELECTION

Our analysis uses a technique uniquely suited to $e^+e^-$ flavor factory experiments, in which the energy and momentum of the initial state are known. We first reconstruct a $B$ meson decaying hadronically; this is referred to as the “tag-side” $B$ meson ($B_{\text{tag}}$). We use the reconstructed $B_{\text{tag}}$ momentum and the $e^+e^-$ initial momentum to infer the momentum of the signal-side $B$ meson ($B_{\text{sig}}$). Because $B^0 \to \tau^\pm \ell^\mp$ are two-body decays, the momentum of the $\tau$ lepton can be inferred from the momentum of $B_{\text{sig}}$ and the momentum of $\ell^\mp$; thus, the $\tau^\pm$ does not need to be reconstructed. We define the “missing mass” as

$$M_{\text{miss}} = \sqrt{(E_{B_{\text{sig}}} - E_\ell)^2/c^4 - (p_{B_{\text{tag}}} - p_\ell)^2/c^2},$$

where $E_{B_{\text{sig}}}$ and $p_{B_{\text{sig}}}$ are the energy and momentum, respectively, of $B_{\text{sig}}$, and $E_\ell$ and $p_\ell$ are the corresponding quantities for $\ell^\mp$. The quantity $M_{\text{miss}}$ is the invariant mass of the unconstructed or missing particle and, for $B^0 \to \tau^\pm \ell^\mp$ decays, should peak at the mass of the $\tau$ lepton ($m_\tau = 1.776$ GeV/$c^2$ [23]). To improve the resolution in $M_{\text{miss}}$, we evaluate it in the $e^+e^-$ center-of-mass (c.m.) frame and substitute the beam energy $E_{\text{beam}}$ for $E_{B_{\text{tag}}}$. To avoid introducing bias in our analysis, we analyze the data in a “blind” manner, i.e., we finalize all selection criteria before viewing events in a region around $m_\tau$. This blinded region is $[1.65, 1.90]$ GeV/$c^2$, which corresponds to approximately 3.8σ in the resolution.

#### A. Tag-side selection

We first reconstruct $B_{\text{tag}}$ candidates in one of 1104 hadronic decay channels using a hierarchical algorithm based on the NeuroBayes neural network package [24]. The quality of $B_{\text{tag}}$ is represented by a single classifier output, $O_{\text{NN}}$, which ranges from 0 (backgroundlike) to 1 (signal-like). The output $O_{\text{NN}}$ is mainly determined by the $B_{\text{tag}}$ reconstruction. It includes event-shape information and significantly suppresses $e^+e^- \to q\bar{q}$ continuum events. In addition to $O_{\text{NN}}$, two other variables are used for selecting $B_{\text{tag}}$ candidates: the energy difference $\Delta E \equiv E_{B_{\text{tag}}} - E_{\text{beam}}$, and the beam-energy-constrained mass $M_{\text{bc}} \equiv \sqrt{E_{\text{beam}}^2/c^4 - |p_{B_{\text{tag}}}^\perp|^2/c^2}$, where $E_{B_{\text{tag}}}$ and $p_{B_{\text{tag}}}^\perp$ are the reconstructed energy and momentum, respectively, of $B_{\text{tag}}$. These quantities are evaluated in the $e^+e^-$ c.m. system. The $B_{\text{tag}}$ candidate is required to satisfy $|\Delta E| < 0.05$ GeV. For each signal mode, we choose selection criteria on $O_{\text{NN}}$ and $M_{\text{bc}}$ by optimizing a figure of merit (FOM). The FOM is defined as $\varepsilon_{\text{MC}}/\sqrt{N_B}$, where $\varepsilon_{\text{MC}}$ is the reconstruction efficiency of signal events as determined from MC simulation, and $N_B$ is the number of background events expected within the signal region $M_{\text{miss}} \in [1.65, 1.90]$ GeV/$c^2$. Based on FOM studies, we require $O_{\text{NN}} > 0.082$ for $B^0 \to \tau^\pm \mu^\mp$, $O_{\text{NN}} > 0.095$ for $B^0 \to \tau^\pm e^\mp$, and $M_{\text{bc}} > 5.272$ GeV/$c^2$ for both modes.

After all $B_{\text{tag}}$ selection criteria are applied, about 10% of $B^0 \to \tau^\pm \mu^\mp$ events and 8% of $B^0 \to \tau^\pm e^\mp$ events have multiple $B_{\text{tag}}$ candidates. For such events, we select a single $B_{\text{tag}}$ by choosing the candidate with the highest value of $O_{\text{NN}}$. This criterion selects the correct candidate 90% of the time, according to MC simulation.

#### B. Signal-side selection

To reconstruct the signal side, only tracks not associated with $B_{\text{tag}}$ are considered. Such tracks are required to originate from the interaction point (IP) and have an impact parameter $|dz| < 4.0$ cm along the $z$ axis, which points opposite the $e^+$ beam direction. We also require $dr < 2.0$ cm in the $x$--$y$ plane (transverse to the $e^+$ beam direction), where $dr = \sqrt{dx^2 + dy^2}$.

Charged tracks are identified by combining information from various subdetectors into a likelihood function $L_i$, where $i = e, \mu, \pi, K$, or $p$ [25]. Muon candidates are identified based on the response of the CDC and KLM [26]. A track with a likelihood ratio $R_\mu = L_\mu/(L_\mu + L_e + L_K) > 0.90$ is identified as a muon. The detection efficiency of this requirement is about 89%, and the pion misidentification rate is about 2%. Electron candidates are identified mainly using the ratio of the energy deposited in the ECL to the track momentum, the shower shape in the ECL, and the energy loss in the CDC. A track with a likelihood ratio $R_e = L_e/(L_e + L_{\text{hadrons}}) > 0.90$ is identified as an electron, where $L_{\text{hadrons}}$ is the product of probability density functions (PDFs) for hadrons [27]. The efficiency of this requirement is about 94%, and the pion misidentification rate is about 0.3%. We recover electron energy lost due to bremsstrahlung by searching for photons within a cone of radius 50 mrad centered around the electron momentum. If such a photon is found,
its four-momentum (assuming it originated at the IP) is added to that of the electron.

We require that $M_{\text{miss}}$ be in the range 1.40 to 2.20 GeV/c$^2$. Every muon or electron candidate satisfying this requirement is treated as a $B_{\text{sig}}$ candidate. After these selections, we find that less than 1% of $B^0 \rightarrow \tau^\pm \mu^\mp$ and $B^0 \rightarrow \tau^\pm e^\mp$ events have multiple $B_{\text{sig}}$ candidates. These fractions are consistent with those from MC simulations. For such events, in order to preserve efficiency, we retain all such candidates, i.e., we do not apply a best-candidate selection.

C. Background

After applying all selection criteria, a small amount of background remains. This background is studied using MC simulation and found to originate mainly from $b \rightarrow cW$ and $b \rightarrow u\ell\nu$ decays. These backgrounds are smoothly falling in the $M_{\text{miss}}$ distribution. However, for $B^0 \rightarrow \tau^\pm \mu^\mp$ candidates, two small peaks are observed: one at $M_{\text{miss}} \approx 1.869$ GeV/c$^2$ and the other at $M_{\text{miss}} \approx 2.010$ GeV/c$^2$. The former corresponds to $B^0 \rightarrow D^-\pi^+$ decays, while the latter corresponds to $B^0 \rightarrow D^+\pi^-$ decays, where in both cases the $\pi^+$ is misidentified as $\mu^+$. These $B^0 \rightarrow D^{(*)}\pi^+$ decays are taken into account when fitting the $M_{\text{miss}}$ distribution for the signal yield (described below).

D. Control samples

We use control samples of $B^0 \rightarrow D^{(*)}\pi^+$ decays to determine corrections to the shapes of the $B^0 \rightarrow \tau^\pm \ell^\mp$ PDFs used to fit for the signal yields (see Sec. IV). To identify $B^0 \rightarrow D^{(*)}\pi^+$ decays, we select pions on the signal side rather than leptons. Pion candidates are identified using $dE/dx$ measured in the CDC, time-of-flight information from the TOF, and the photon yield in the ACC. A track with a likelihood ratio $R_{\pi} = L_\pi/(L_\pi + L_K) > 0.90$ is identified as a pion [25]. All other selection criteria are the same as for the $B^0 \rightarrow \tau^\pm \ell^\mp$ search. In addition, we veto leptons by requiring that $R_\mu < 0.90$ and $R_e < 0.90$. With the above selection, the pion identification efficiency is about 95%, and the kaon misidentification rate is about 5%.

IV. MAXIMUM LIKELIHOOD FITS

We determine the $B^0 \rightarrow \tau^\pm \ell^\mp$ signal yields by performing an unbinned extended maximum-likelihood fit to the $M_{\text{miss}}$ distributions. The PDF used to model correctly reconstructed signal decays is a double Gaussian for $B^0 \rightarrow \tau^\pm \mu^\mp$ and the sum of three Gaussians for $B^0 \rightarrow \tau^\pm e^\mp$. These Gaussians are allowed to have different means. We also model misreconstructed signal decays in which the lepton selected is subject to final-state radiation or is not a direct daughter in the two-body $B^0 \rightarrow \tau^\pm \ell^\mp\bar{\nu}$ decay, i.e., it originates from $\tau^\pm \rightarrow \ell^\pm \nu\bar{\nu}$ or $\tau^\pm \rightarrow \pi^\pm \rightarrow \ell^\pm \nu\bar{\nu}$. This component is referred to as a “self-cross-feed” signal, and we model it with a double Gaussian and an exponential function. The fractions of self-cross-feed signal are fixed to the values obtained from MC simulation: (5.0 ± 0.2)% for $B^0 \rightarrow \tau^\pm \mu^\mp$ and (14.0 ± 0.3)% for $B^0 \rightarrow \tau^\pm e^\mp$. The self-cross-feed fraction is larger for the electron channel due to a larger contribution from $B^0 \rightarrow \tau^\pm e^\mp \gamma$ decays.

The shape parameters of the signal PDFs are obtained from MC simulations. We make corrections to these to account for small differences observed between the MC simulation and data. We obtain these correction factors by fitting the $M_{\text{miss}}$ distributions of the high-statistics $B^0 \rightarrow D^{(*)}\pi^+$ control samples. For the $B^0 \rightarrow D^{(*)}\pi^+$ samples, we fit both data and MC events and record small shifts observed in the means of the PDFs, and nominal differences in the widths. We apply these shifts for the means and scaling factors for the widths to the $B^0 \rightarrow \tau^\pm \ell^\mp$ signal PDFs. The uncertainties in these correction factors are accounted for when evaluating systematic uncertainties.

Background PDFs of all modes are modeled with exponential functions. The shape parameters for these background PDFs are all floated, along with the background and signal yields. The PDFs for misidentified $B^0 \rightarrow D^-\pi^+$ and $B^0 \rightarrow D^+\pi^-$ decays are taken to be a double Gaussian and the sum of three Gaussians, respectively.

We validate our fitting procedure and check for fit bias using MC simulations. We generate large ensembles of simulated experiments, in which the $M_{\text{miss}}$ distributions are generated from the PDFs used for fitting. We fit these ensembles and find that the fitted signal yields are consistent with the input values; the mean difference is

![Graph](image-url)
To assess the goodness of fit, we calculate these decays along with projections of the fit result are shown in Fig. 2. The \( \chi^2/n_{\text{dof}} \) values are 0.54 (\( n_{\text{dof}} = 44 \)) and 0.70 (\( n_{\text{dof}} = 44 \)) for \( B^0 \to \tau^+\mu^- \) and \( B^0 \to \tau^+e^- \), respectively. The fitted signal yields are \( N_{\text{sig}} = 1.8_{-0.6}^{+0.7} \) for \( B^0 \to \tau^+\mu^- \) and \( N_{\text{sig}} = 0.3_{-0.2}^{+0.8} \) for \( B^0 \to \tau^+e^- \). Both yields are consistent with zero. In the \( B^0 \to \tau^+\mu^- \) sample, we observe (17 ± 10) \( B^0 \to D^- \pi^+ \) events and (−2 ± 12) \( B^0 \to D^-\pi^+ \) events; these yields are consistent with expectations based on MC simulation.

### V. UPPER LIMIT CALCULATION

We calculate upper limits on \( N_{\text{sig}} \) and the branching fractions at 90% CL using a frequentist method. We first generate sets of MC-simulated events, with each set being equivalent to the Belle data sample. Both signal and background events are generated according to their respective PDFs. The number of background events generated is equal to that obtained from the data fit. We vary the number of input signal events, and for each value we generate an ensemble of 10,000 data sets. We fit these data sets and calculate the fraction \( f_{\text{sig}} \) that has a fitted signal yield less than that obtained from the Belle data (1.8 or 0.3 events). Our 90% CL upper limit on the number of signal events \( (N_{\text{UL}})_{\text{sig}} \) is the number of input signal events that has \( f_{\text{sig}} = 0.10 \). We convert \( N_{\text{UL}}^{\text{sig}} \) to an upper limit on the branching fraction \( (B^{\text{UL}}) \) via the formula

\[
B^{\text{UL}} = \frac{N_{\text{UL}}^{\text{sig}}}{2 \times N_{\text{BB}} \times f^{00} \times \epsilon}.
\]

In this expression, \( N_{\text{BB}} \) is the number of \( B\overline{B} \) pairs; \( f^{00} = 0.486 \pm 0.006 \) is the fraction that are \( B^0\overline{B}^0 \) [23]; and \( \epsilon \) is the signal efficiency including tag-side branching fractions and reconstruction efficiencies.

We include systematic uncertainties (discussed below) in \( B^{\text{UL}} \) as follows. We divide all systematic uncertainties into two types (see Table II): those arising from the numerator of Eq. (2) (“additive” uncertainties), and those arising from the denominator of Eq. (2) (“multiplicative” uncertainties). Additive uncertainties arise from fitting for the signal yield, while multiplicative uncertainties correspond to the number of \( B \) decays reconstructed. We account for the latter when generating MC data sets in our frequentist procedure. The number of signal events is varied randomly around the
TABLE I. Summary of the fit results for $N_{\text{sig}}$, and the resulting 90% CL upper limits $N_{\text{UL}}^{\text{UL}}$ and $B^{\text{UL}}$ (see text).

| Mode       | $\varepsilon$ ($\times 10^{-4}$) | $N_{\text{sig}}$ | $N_{\text{UL}}^{\text{UL}}$ | $B^{\text{UL}}$ ($\times 10^{-5}$) |
|------------|----------------------------------|-------------------|-----------------------------|-----------------------------------|
| $B^0 \rightarrow \tau^+\mu^-$  | 11.0                            | 1.8               | 12.4                        | 1.5                               |
| $B^0 \rightarrow \tau^+e^+$    | 9.8                             | 0.3               | 11.6                        | 1.6                               |

nominal input value by the total multiplicative uncertainty. Subsequently, after fitting an MC data set, we adjust the fitted value $N_{\text{sig}}$ by a value sampled from a Gaussian distribution with mean zero and a width equal to the total additive uncertainty. As a final step, to include possible fit bias, this value is shifted by an amount obtained by sampling a Gaussian distribution with a mean equal to the fit bias discussed earlier (the central value) and a width equal to the uncertainty in the bias. This final value is used when calculating $f_{\text{sig}}$. The resulting upper limits for $N_{\text{UL}}^{\text{UL}}$ and $B^{\text{UL}}$ are listed in Table I. These values are the same as the upper limits expected based on MC ($1.6 \times 10^{-5}$ for both modes), reflecting good agreement between the background levels observed in data and the MC.

VI. SYSTEMATIC UNCERTAINTIES

The systematic uncertainties in our measurement—aside from potential fit bias, which is treated separately when setting the upper limits—are listed in Table II. Uncertainties in the shapes of the PDFs used for the signal are evaluated by varying all fixed parameters by $\pm 1\sigma$; the resulting change in the signal yield is taken as the systematic uncertainty. The fixed parameters that are varied include the correction factors to the shapes as obtained from the $B^0 \rightarrow D^{(*)}\pi^+$ control samples. The fraction of the self-cross-feed signal is fixed to the MC value. We vary this fraction by $\pm 50\%$ and take the resulting change in the signal yield as the systematic uncertainty.

The reconstruction efficiency for $B_{\text{tag}}$ is evaluated via MC simulation. However, there is uncertainty arising from branching fractions for tagging modes that are not well measured, and from unknown decay dynamics of multi-body hadronic decays. To account for these effects, a correction factor to the reconstruction efficiency is applied. This correction is evaluated as done in Ref. [28], by comparing the number of events containing both a $B_{\text{tag}}$ and a semileptonic $B \rightarrow D^{(*)}\ell\nu$ decay in data and MC. As the branching fractions for $B \rightarrow D^{(*)}\ell\nu$ are precisely known, and their reconstruction efficiencies can be separately calculated, the difference between data and MC for $B_{\text{tag}}$ reconstruction can be extracted. The resulting correction factor is $0.64 \pm 0.03$. The uncertainty in this value is taken as a systematic uncertainty.

The systematic uncertainty due to charged track reconstruction is evaluated using $D^{(*)} \rightarrow D^0\pi^+$ decays, with $D^0 \rightarrow K_0^0\pi^+\pi^-$ and $K_+ \rightarrow \pi^+\pi^-$. The resulting uncertainty is $0.35\%$ per track. The uncertainty due to lepton identification is evaluated using $e^+e^- \rightarrow e^+e^-\gamma\gamma^\prime \rightarrow e^+e^-\ell^+\ell^-$ events. The resulting uncertainties are $1.6\%$ for muons and $1.8\%$ for electrons.

The systematic uncertainty in the signal reconstruction efficiency due to limited MC statistics is $< 0.1\%$ for both signal modes. The systematic uncertainty arising from the number of $BB$ pairs is $1.4\%$, and the known uncertainty on $f^{00}$ corresponds to a systematic uncertainty of $1.2\%$.

The total additive (in number of events) and multiplicative (in percent) systematic uncertainties are obtained by adding in quadrature all systematic uncertainties of that type.

TABLE II. Systematic uncertainties for the branching fraction measurement. Those listed in the upper section (“additive”) arise from fitting for the signal yield and are listed in number of events; those in the lower section (“multiplicative”) arise from the number of reconstructed $B$ decays and are listed in percent.

| Source              | $B^0 \rightarrow \tau^+\mu^-$ | $B^0 \rightarrow \tau^+e^+$ |
|---------------------|--------------------------------|-----------------------------|
| PDF shapes          | 0.7                            | 0.3                         |
| Self-cross-feed     | $< 0.1$                        | 0.1                         |
| Total (events)      | 0.7                            | 0.3                         |
| $B_{\text{tag}}$    | 4.5                            | 4.5                         |
| Track reconstruction| 0.3                            | 0.3                         |
| Lepton identification| 1.6                           | 1.8                         |
| MC statistics       | $< 0.1$                        | $< 0.1$                     |
| Number of $BB$ pairs| 1.4                            | 1.4                         |
| $f^{00}$ ($BB \rightarrow B^0B^0$ fraction) | 1.2 | 1.2 |
| Total (%)           | 5.1                            | 5.2                         |

VII. SUMMARY

We have searched for the lepton-flavor-violating decays $B^0 \rightarrow \tau^+\ell^-$ using the full Belle data set. We find no evidence for these decays and set the following upper limits on the branching fractions at 90% CL:

$$B(B^0 \rightarrow \tau^+\mu^-) < 1.5 \times 10^{-5},$$

$$B(B^0 \rightarrow \tau^+e^-) < 1.6 \times 10^{-5}.\quad (4)$$

Our result for $B^0 \rightarrow \tau^+\mu^-$ is very similar to a recent result from LHCb [15]. Our result for $B^0 \rightarrow \tau^+e^-$ is the most stringent limit to date, improving upon the previous limit by almost a factor of two. We find no indication of lepton flavor violation in these decays.
ACKNOWLEDGMENTS

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 for strong computing support; and the National Institute of Informatics, and Science Information Network 5 (SINET5) for valuable network support. We acknowledge support from 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. DP150103061, and No. FT130100303; Austrian Federal Ministry of Education, Science and Research (FWF) and FWF Austrian Science Fund No. P 31361-N36; the National Education, Science and Research (FWF) and FWF Higher Education of the Russian Federation, Agreement No. 113011 (2004); S. Fukuda et al. (KamLand Collaboration), Phys. Rev. Lett. 90, 021802 (2003); M. H. Ahn et al., Phys. Rev. Lett. 90, 041801 (2003).

[1] Throughout this paper, charge-conjugate modes are implicitly included unless noted otherwise.

[2] Y. Amhis et al. (Heavy Flavor Averaging Group), Eur. Phys. J. C 81, 226 (2021); and online update at https://hflav.web.cern.ch/.

[3] R. Aaij et al. (LHCb Collaboration), J. High Energy Phys. 08 (2017) 055.

[4] R. Aaij et al. (LHCb Collaboration), Phys. Rev. Lett. 122, 191801 (2019).

[5] R. Aaij et al. (LHCb Collaboration), arXiv:2103.11769.

[6] S. Wehle et al. (Belle Collaboration), Phys. Rev. Lett. 126, 161801 (2021).

[7] J. P. Lees et al. (BABAR Collaboration), Phys. Rev. D 86, 032012 (2012).

[8] S. L. Glashow, D. Guadagnoli, and K. Lane, Phys. Rev. Lett. 114, 091801 (2015).

[9] X.-G. He, G. Valencia, and Y. Wang, Phys. Rev. D 70, 113011 (2004); S. Fukuda et al. (Super-Kamiokande Collaboration), Phys. Rev. Lett. 85, 3999 (2000); Q. R. Ahmad et al. (SNO Collaboration), Phys. Rev. Lett. 89, 011301 (2002); K. Eguchi et al. (KamiLand Collaboration), Phys. Rev. Lett. 90, 021802 (2003); M. H. Ahn et al., Phys. Rev. Lett. 90, 041801 (2003).

[10] Specifically, a tiny neutrino mass in the eV range.

[11] A. D. Smirnov, Mod. Phys. Lett. A 33, 1850019 (2018).

[12] A. Dedes, J. Ellis, and M. Raidal, Phys. Lett. B 549, 159 (2002).

[13] A. Bornheim et al. (CLEO Collaboration), Phys. Rev. Lett. 93, 241802 (2004).

[14] B. Aubert et al. (BABAR Collaboration), Phys. Rev. D 77, 091104(R) (2008).

[15] R. Aaij et al. (LHCb Collaboration), Phys. Rev. Lett. 123, 211801 (2019).

[16] S. Kurokawa and E. Kikutani, Nucl. Instrum. Methods Phys. Res., Sect. A 499, 1 (2003); and other papers included.
in this volume; T. Abe et al., Prog. Theor. Exp. Phys. 2013, 03A001 (2013) and following articles up to 03A011.

[17] Z. Natkaniec et al. (Belle SVD2 Group), Nucl. Instrum. Methods Phys. Res., Sect. A 560, 1 (2006).

[18] A. Abashian et al. (Belle Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A 479, 117 (2002); also see detector section in J. Brodzicka et al., Prog. Theor. Exp. Phys. 2012, 04D001 (2012).

[19] D. J. Lange, Nucl. Instrum. Methods Phys. Res., Sect. A 462, 152 (2001).

[20] T. Sjöstrand, S. Mrenna, and P. Skands, J. High Energy Phys. 05 (2006) 026.

[21] E. Barberio, B. van Eijk, and Z. Was, Comput. Phys. Commun. 66, 115 (1991).

[22] R. Brun et al., CERN Report No. DD/EE/84-1, 1984.

[23] P. A. Zyla et al. (Particle Data Group), Prog. Theor. Exp. Phys. 2020, 083C01 (2020).

[24] M. Feindt, F. Keller, M. Kreps, T. Kuhr, S. Neubauer, D. Zander, and A. Zupanc, Nucl. Instrum. Methods Phys. Res., Sect. A 654, 432 (2011).

[25] E. Nakano, Nucl. Instrum. Methods Phys. Res., Sect. A 494, 402 (2002).

[26] A. Abashian et al., Nucl. Instrum. Methods Phys. Res., Sect. A 491, 69 (2002).

[27] K. Hanagaki, H. Kakuno, H. Ikeda, T. Iijima, and T. Tsukamoto, Nucl. Instrum. Methods Phys. Res., Sect. A 485, 490 (2002).

[28] A. Sibidanov et al. (Belle Collaboration), Phys. Rev. D 88, 032005 (2013).