Search for lepton-flavor-violating τ decays into a lepton and a vector meson using the full Belle data sample

The Belle collaboration
https://belle.kek.jp/

N. Tsuzuki ⓘ, K. Inami ⓘ, I. Adachi ⓘ, H. Aihara ⓘ, D.M. Asner ⓘ, H. Atmacan ⓘ,
T. Aushev ⓘ, R. Ayad ⓘ, V. Babu ⓘ, Sw. Banerjee ⓘ, P. Behera ⓘ, K. Belous ⓘ,
J. Bennett ⓘ, M. Bessner ⓘ, B. Bhuyan ⓘ, T. Bilka ⓘ, D. Biswas ⓘ, D. Bodrov ⓘ,
J. Borah ⓘ, A. Bozek ⓘ, M. Bračko ⓘ, P. Branchini ⓘ, T.E. Browder ⓘ, A. Budano ⓘ,
M. Campajola ⓘ, D. Červenkov ⓘ, M.-C. Chang ⓘ, B.G. Cheon ⓘ, K. Chilikin ⓘ,
H.E. Cho ⓘ, K. Cho ⓘ, S.-J. Cho ⓘ, S.-K. Choi ⓘ, Y. Choi ⓘ, S. Choudhury ⓘ,
D. Cinabro ⓘ, J. Cochran ⓘ, S. Das ⓘ, N. Dash ⓘ, G. De Nardo ⓘ, G. De Pietro ⓘ,
R. Dhamija ⓘ, F. Di Capua ⓘ, Z. Doležal ⓘ, T.V. Dong ⓘ, D. Dossett ⓘ, S. Dubey ⓘ,
D. Epifanov ⓘ, T. Ferber ⓘ, D. Ferlewicz ⓘ, B.G. Fulsom ⓘ, V. Gaur ⓘ, A. Giri ⓘ,
P. Goldenzweig ⓘ, Y. Guan ⓘ, K. Gudkova ⓘ, X. Han ⓘ, T. Hara ⓘ, K. Hayasaka ⓘ,
H. Hayashii ⓘ, M.T. Hedges ⓘ, D. Herrmann ⓘ, W.-S. Hou ⓘ, C.-L. Hsu ⓘ,
T. Iijima ⓘ, N. Ipsita ⓘ, A. Ishikawa ⓘ, R. Itoh ⓘ, M. Iwasaki ⓘ, W.W. Jacobs ⓘ,
E.-J. Jang ⓘ, S. Jia ⓘ, Y. Jin ⓘ, T. Kawasaki ⓘ, C. Kiesling ⓘ, C.H. Kim ⓘ,
D.Y. Kim ⓘ, K.-H. Kim ⓘ, Y.-K. Kim ⓘ, K. Kinoshita ⓘ, P. Kodyš ⓘ, T. Konno ⓘ,
A. Korobov ⓘ, S. Korpar ⓘ, E. Kovalenko ⓘ, P. Križan ⓘ, P. Krokovny ⓘ,
M. Kumar ⓘ, K. Kumara ⓘ, A. Kuzmin ⓘ, Y.-J. Kwon ⓘ, S.C. Lee ⓘ, J. Li ⓘ,
L.K. Li ⓘ, Y. Li ⓘ, J. Libby ⓘ, K. Lieret ⓘ, Y.-R. Lin ⓘ, D. Liventsev ⓘ, Y. Ma ⓘ,
A. Martini ⓘ, M. Masuda ⓘ, K. Matsuoka ⓘ, D. Matvienko ⓘ, S.K. Maurya ⓘ,
F. Meier ⓘ, M. Merola ⓘ, K. Miyabayashi ⓘ, R. Mizuk ⓘ, G.B. Mohanty ⓘ,
M. Nakao ⓘ, Z. Natkaniec ⓘ, A. Natochii ⓘ, L. Nayak ⓘ, M. Niiyama ⓘ,
N.K. Nisar ⓘ, S. Nishida ⓘ, S. Ogawa ⓘ, H. Ono ⓘ, P. Oskin ⓘ, G. Pakhlova ⓘ,
T. Pang ⓘ, S. Pardi ⓘ, H. Park ⓘ, J. Park ⓘ, S.-H. Park ⓘ, A. Passeri ⓘ, S. Paul ⓘ,
T.K. Pedlar ⓘ, R. Pestotnik ⓘ, L.E. Piilonen ⓘ, T. Podobnik ⓘ, E. Prencipe ⓘ,
M.T. Prim ⓘ, A. Rostomyan ⓘ, N. Rout ⓘ, G. Russo ⓘ, S. Sandilya ⓘ, A. Sangal ⓘ,

L. Santelj [iD], **V. Savinov** [iD], **G. Schnell** [iD], **C. Schwanda** [iD], **Y. Seino** [iD], **K. Senyo** [iD],
M.E. Sevior [iD], **W. Shan** [iD], **M. Shapkin** [iD], **C. Sharma** [iD], **J.-G. Shiu** [iD], **E. Solovieva** [iD],
M. Starič [iD], **M. Sumihama** [iD], **T. Sumiyoshi** [iD], **M. Takizawa** [iD], **U. Tamponi** [iD],
K. Tanida [iD], **F. Tenchini** [iD], **M. Uchida** [iD], **T. Uglov** [iD], **Y. Unno** [iD], **S. Uno** [iD],
P. Urquijo [iD], **Y. Ushiroda** [iD], **S.E. Vahsen** [iD], **G. Varner** [iD], **A. Vinokurova** [iD],
D. Wang [iD], **E. Wang** [iD], **M.-Z. Wang** [iD], **X.L. Wang** [iD], **S. Watanuki** [iD], **X. Xu** [iD],
B.D. Yabsley [iD], **W. Yan** [iD], **S.B. Yang** [iD], **J. Yelton** [iD], **Y. Yook** [iD], **L. Yuan** [iD],
Y. Zhai [iD], **V. Zhilich** [iD] and **V. Zhukova** [iD]

E-mail: noritsugu@hepl.phys.nagoya-u.ac.jp

ABSTRACT: Charged-lepton-flavor-violation is predicted in several new physics scenarios. We update the analysis of τ lepton decays into a light charged lepton ($\ell = e^{\pm}$ or $\mu^{\pm}$) and a vector meson ($V^0 = \rho^0$, ϕ, ω, K^{*0}, or $\overline{K}^{*0}$) using $980\,\mathrm{fb}^{-1}$ of data collected with the Belle detector at the KEKB collider. No significant excess of such signal events is observed, and thus 90% credibility level upper limits are set on the $\tau \to \ell V^0$ branching fractions in the range of $(1.7\text{--}4.3) \times 10^{-8}$. These limits are improved by 30% on average from the previous results.

Contents

1 Introduction

In the Standard Model, charged-lepton-flavor-violation (CLFV) is so strongly suppressed that it is undiscoverable by current experiments. Therefore, a discovery of a CLFV event indicates new physics (NP). Verifying various NP models requires many searches of various CLFV modes [1]. Whereas the CLFV constraints are much more stringent for μ-to-e than for τ through the precise measurements [2–4], we are interested in τ, the third-generation and heaviest lepton. So-called B-anomalies, which indicate NP effects in B semileptonic decays [5–16], also motivate the CLFV searches.

We focus on τ CLFV decays into a charged lepton ($\ell = e^{\pm}$ or $\mu^{\pm}$) and a neutral vector meson ($V^0 = \rho^0$, ϕ, ω, K^{*0}, or $\overline{K}^{*0}$). In refs. [17–22], the $\tau \to \mu\phi$ mode is a sensitive probe for leptoquark models that can explain the B-anomalies.[1] Some other NP models predict branching fractions of $\mathcal{O}(10^{-10})$–$\mathcal{O}(10^{-8})$ for $\tau \to \ell V^0$ [25–28].

We previously searched for $\tau \to \ell V^0$ events using $854\,\mathrm{fb}^{-1}$ of Belle data, and set 90% credibility level (C.L.) upper limits on the branching fractions in the range of (1.2–8.4) $\times 10^{-8}$ [29].[2] This paper reports an updated search for $\tau \to \ell V^0$ using the full $980\,\mathrm{fb}^{-1}$ Belle data set. The signal efficiency is improved through new event selection criteria, which allow more signal candidates first and then remove background events by a multivariate analysis.

---

[1]One of the B-anomalies which motivated the models described in those references is the $R(K^{(*)})$ anomaly reported by the LHCb experiment [23], but it disappeared in their updated analysis [24].

[2]In common high energy physics usage, this credibility level has been reported as "confidence level," which is a frequentist-statistics term.

2 Belle experiment

The Belle detector is a spectrometer that covers large solid angles of the e^+e^- collision events from the KEKB accelerator [30, 31]. The detector consists of a silicon vertex detector, a 50-layer central drift chamber (CDC), an array of aerogel threshold Cherenkov counters, time-of-flight scintillation counters, and an electromagnetic calorimeter composed of 8736 CsI(Tl) crystals (ECL). These devices are located inside a superconducting solenoid coil that provides a 1.5 T magnetic field. An iron flux return located outside of the coil is instrumented to detect K_L^0 mesons and identify muons. The Belle detector is described in detail elsewhere [32, 33].

Of the $980\,\text{fb}^{-1}$ data set, $703\,\text{fb}^{-1}$ was collected at the $\Upsilon(4S)$ resonance, $121\,\text{fb}^{-1}$ at the $\Upsilon(5S)$, $89\,\text{fb}^{-1}$ at an energy 60 MeV below the $\Upsilon(4S)$, $28\,\text{fb}^{-1}$ of energy-scans above the $\Upsilon(4S)$, and the remainder at and near the $\Upsilon(1\text{–}3S)$. Compared to the previous paper [29], the following data sets have been added: $78\,\text{fb}^{-1}$ at and near the $\Upsilon(5S)$, $38\,\text{fb}^{-1}$ at and near the $\Upsilon(1\text{–}3S)$, and $10\,\text{fb}^{-1}$ at an energy 60 MeV below the $\Upsilon(4S)$.

The e^+e^- collision events in the Belle detector are simulated by the Monte Carlo (MC) method. Signal MC events of $\tau \to \ell V^0$ are generated by a dedicated MC with KKMC and TAUOLA [34], where $\tau^+\tau^-$ pairs are initially produced and one of the τ's decays into ℓV^0 and the other decays generically. The numbers of generated signal MC events are 1.1×10^6 events at the $\Upsilon(4S)$ resonance, 0.4×10^6 events at the $\Upsilon(5S)$, 0.1×10^6 events at each of the $\Upsilon(1\text{–}3S)$, and 0.1×10^6 events at an energy 60 MeV below the $\Upsilon(4S)$. We assume a uniform CLFV decay angle in the τ rest frame. No specific NP model is assumed in the CLFV decay process, and the spin direction of V^0 is set randomly and independently of the spin of the mother τ. For background MC simulations, $e^+e^- \to q\bar{q}$ $(q = u, d, s, c)$, $e^+e^- \to \tau^+\tau^-$, Bhabha, and two-photon processes are generated by EvtGen [35], KKMC [34], BHLUMI [36], and AAFH [37], respectively. The detector responses are simulated by GEANT3 [38].

3 Reconstruction and event selection

To improve the signal selection efficiency, we loosen or change the event selection criteria in the previous paper [29] for this analysis. Some of the loose selection variables are used as inputs to the multivariate analysis, which is intended to reduce background events. All the event selection criteria have been optimized and are described in this section.

A signal τ is reconstructed from a lepton and a neutral vector meson. We separate the event into two hemispheres in the center-of-mass (c.m.) frame by a plane perpendicular to the thrust vector $(\vec{n}_T)$ [39, 40]. The thrust vector is obtained by maximizing the thrust $T = \Sigma_i |\vec{p}_i^{\,\text{c.m.}} \cdot \vec{n}_T| / \Sigma_i |\vec{p}_i^{\,\text{c.m.}}|$, where i runs over all tracks and photons, and $\vec{p}_i^{\,\text{c.m.}}$ is the momentum in the c.m. frame. Here, photons with energies above $0.1\,\text{GeV}$ are used. In the hemisphere that contains a τ CLFV decay (called signal side and τ_{sig}), V^0 is reconstructed as follows: ρ^0 from $\pi^+\pi^-$ within the reconstructed mass window of $0.445\text{–}1.08\,\text{GeV}/c^2$, ϕ from K^+K^- within $1.00\text{–}1.04\,\text{GeV}/c^2$, ω from $\pi^+\pi^-\pi^0$ within $0.7\text{–}0.9\,\text{GeV}/c^2$, K^{*0} from $K^+\pi^-$ within $0.7\text{–}1.1\,\text{GeV}/c^2$, and $\overline{K}^{*0}$ from $K^-\pi^+$ within $0.7\text{–}1.1\,\text{GeV}/c^2$. In the other hemisphere (called tag side), the other τ (τ_{tag}) is reconstructed from $\ell^\pm\nu\nu$, $\pi^\pm\nu$, $\pi^\pm\pi^0\nu$,

$\pi^{\pm}\pi^0\pi^0\nu$, or $\pi^{\pm}\pi^{\mp}\pi^{\pm}\nu$. The $\tau_{\text{tag}} \to \pi^{\pm}\pi^0\pi^0\nu$ and $\pi^{\pm}\pi^{\mp}\pi^{\pm}\nu$ decays were discarded in the previous paper [29]. This τ_{tag} information enables the suppression of background events that have no neutrinos in the tag side.

The signal $\tau \to \ell V^0$ events have a unique kinematical feature; the ℓV^0 invariant mass ($M_{\ell V^0}$) is close to the τ mass and the difference of the ℓV^0 energy from the beam energy in the c.m. frame (ΔE) is close to zero. The signal events within $1.65\,\text{GeV}/c^2 < M_{\ell V^0} < 1.90\,\text{GeV}/c^2$ and $|\Delta E| < 0.5\,\text{GeV}$ are reconstructed in this paper. We follow a blind analysis approach in this search by not looking at the signal candidates in the data set until finalizing the event selection and background estimation. The blind region is $1.75\,\text{GeV}/c^2 \leq M_{\ell V^0} < 1.81\,\text{GeV}/c^2$ and $|\Delta E| < 0.08\,\text{GeV}$ for the $\mu\rho^0$, $\mu\phi$ and $\mu K^{*0}(\overline{K}^{*0})$ modes, and $1.74\,\text{GeV}/c^2 \leq M_{\ell V^0} < 1.82\,\text{GeV}/c^2$ and $|\Delta E| < 0.1\,\text{GeV}$ for the other modes.

Charged tracks, photons, and π^0s should satisfy the following selection criteria. Each charged track or photon is within the fiducial volume defined by $-0.866 < \cos\theta < 0.956$, where θ is the polar angle with respect to the direction opposite to the e^+ beam in the laboratory frame. Charged tracks come from the interaction point; the distance of the closest point from the interaction point is less than 0.5 cm in the transverse direction and less than 3.0 cm in the longitudinal direction. Each π^0 is reconstructed from two photons inside the same hemisphere and the photon energy (E_γ) should be larger than 0.05 GeV. The π^0 mass window is $0.12\,\text{GeV}/c^2 < M_{\gamma\gamma} < 0.15\,\text{GeV}/c^2$, corresponding to $\pm 3\sigma$ in the π^0 mass resolution. A π^0 mass-constrained fit is performed to improve the energy resolution.

After reconstructing the signal and tag τ's, no extra charged tracks are allowed. We count the number of photons (n_γ) with E_γ larger than 0.1 GeV in the signal side, and require $n_\gamma \leq 3$ for the $\ell\omega$ mode, which includes a $\pi^0 \to \gamma\gamma$, and $n_\gamma \leq 1$ for the other modes.

Particle identification is effective in suppressing the main background events of three-hadron-track to the $\tau \to \ell V^0$ signal. We use likelihood ratios for electron identification ($\mathcal{P}(e)$) [41] and muon identification ($\mathcal{P}(\mu)$) [42]. The lepton identification criteria are $\mathcal{P}(e) > 0.9$ for electrons, and $\mathcal{P}(\mu) > 0.95$ and the momentum is larger than $0.6\,\text{GeV}/c$ for muons. The electron (muon) identification efficiency is 90% (75%), whereas the probability of misidentifying a pion as an electron (muon) is 0.1% (2%). These efficiencies and fake rates are different from the previous paper [29], because the track momentum requirements are loosened in this paper. The energy loss of an electron by bremsstrahlung is recovered by adding back the energy of every photon within 0.05 radians from the electron track direction into the electron momentum. To suppress low-multiplicity background events like Bhabha, $ee \to eeee$, or $ee \to ee\mu\mu$, an electron veto ($\mathcal{P}(e) < 0.9$) is applied to all hadron candidate tracks from V^0 and τ_{tag}.

For hadron identification, we use a binary likelihood ratio $\mathcal{P}(i|j) = \mathcal{L}_i/(\mathcal{L}_i + \mathcal{L}_j)$, where $\mathcal{L}_{i(j)}$ is the likelihood of particle i (j) [43] and i (j) is π, K, or p. The kaon identification criteria are $\mathcal{P}(K|\pi) > 0.6$ for both charged kaons from ϕ decay and $\mathcal{P}(K|\pi) > 0.8$ for the charged kaon from K^{*0} and $\overline{K}^{*0}$ decays. The kaon identification efficiency is 86% (77%), whereas the probability of misidentifying a charged pion as a kaon is 4% (2%) for the kaons from ϕ (K^{*0}, $\overline{K}^{*0}$). A kaon veto ($\mathcal{P}(K|\pi) < 0.6$) is applied to both charged pions from ρ^0 in the signal side, and 96% of pions are retained, whereas 14% of kaons are not vetoed. We do not apply this kaon veto for the charged pions from ω, K^{*0}, and $\overline{K}^{*0}$. To suppress

muons from kaons decaying inside the CDC ($K^\pm \to \mu^\pm \nu$), the kaon veto is also applied to the signal-side muon track for the hadronic tags ($\tau_{\text{tag}}^\pm \to \pi^\pm \nu$, $\pi^\pm \pi^0 \nu$, $\pi^\pm \pi^\mp \pi^\pm \nu$, or $\pi^\pm \pi^0 \pi^0 \nu$). For the μV^0 modes with the hadronic tags, a proton veto ($\mathcal{P}(p|K) < 0.6$ and $\mathcal{P}(p|\pi) < 0.6$) is applied for the tag-side tracks.

The signal events have one or two neutrinos from the τ_{tag} decay. We introduce some event selection criteria requiring one or more neutrinos in the tag side. The missing momentum due to the neutrino(s) is calculated by subtracting the vector sum of the momenta of all tracks and photons from the sum of the beam momenta. The missing energy is also calculated by subtracting the sum of the energy of all tracks and photons from the sum of the beam energy. Here, extra photons that are not used for the τ reconstruction are included. The transverse missing momentum is required to be larger than $0.5\,\text{GeV}/c$, and the missing energy in the c.m. frame ($E_{\text{miss}}^{\text{c.m.}}$) is required to be larger than $0\,\text{GeV}$. Events with missing particles other than neutrinos should be rejected as background events. These non-neutrino missing particles can arise in two ways: neutral particles pass through the gaps between the barrel and end-cap ECLs, and any particles go outside the CDC volume. Thus, the direction of the missing momentum is required not to point to such regions. The missing particles should be in the tag side and hence $\cos\theta_{\text{miss-tag}}^{\text{c.m.}} > 0$, where $\theta_{\text{miss-tag}}^{\text{c.m.}}$ is the angle between the missing momentum and the vector sum of the momenta of the tag-side tracks and photons in the c.m. frame. The neutrino angle with respect to the τ_{tag} momentum direction is restricted in a τ_{tag} two-body decay; thus $\cos\theta_{\text{miss-tag}}^{\text{c.m.}} < 0.85$ is also applied for the $\ell\rho^0$ modes with $\tau_{\text{tag}}^\pm \to \pi^\pm \nu$.

We require features of a generic τ decay in the tag side. The invariant mass of the particles including all photons in the tag hemisphere should be less than the τ mass ($1.777\,\text{GeV}/c^2$). For τ_{tag} decays into $\pi^\pm \pi^0 \nu$ ($\pi^\pm \pi^\mp \pi^\pm \nu$ or $\pi^\pm \pi^0 \pi^0 \nu$), the reconstructed mass of those pions is required to be $0.4\,\text{GeV}/c^2 < M_{\pi^\pm \pi^0} < 1.3\,\text{GeV}/c^2$ ($0.7\,\text{GeV}/c^2 < M_{3\pi} < 1.7\,\text{GeV}/c^2$), which corresponds to the mass of $\rho^\pm$ ($a_1^\pm$).

After the above event reconstruction, the background sources are the $q\bar{q}$ continuum ($q = u, d, s, c$), generic $\tau^+ \tau^-$, and low-multiplicity events. The low-multiplicity events especially contribute to the background events for eV^0 modes that have electron tracks. We suppress the low-multiplicity events first, and then use a maltivariate analysis tool to suppress the $q\bar{q}$ continuum and generic $\tau^+ \tau^-$ events.

The Bhabha events have tracks from photon conversion. To suppress these background events for the eV^0 modes, the invariant mass of the electron and one of the tracks from the V^0, assigned the electron-mass hypothesis, should be larger than $0.2\,\text{GeV}/c^2$. In addition, for the eK^{*0} and $e\overline{K}^{*0}$ modes, the invariant mass of the two tracks from the V^0, each assigned the electron-mass hypothesis, is required to be larger than $0.1\,\text{GeV}/c^2$. This event selection also suppresses some of the generic $\tau^+ \tau^-$ events, which have tracks from photon conversion.

The low-multiplicity background events are still not negligible for the events with electrons: $\tau \to eV^0$ or $\tau_{\text{tag}} \to e\nu\nu$. Because the missing particles of the low-multiplicity background events are the bremsstrahlung photons from the electron in the tag side, $\cos\theta_{\text{miss-tag}}^{\text{c.m.}}$ is close to one (figure 1). In addition, the missing energy is small for some low-multiplicity background events. For the $\mu\rho^0$ mode with $\tau_{\text{tag}}^\pm \to e^\pm \nu\nu$, $\cos\theta_{\text{miss-tag}}^{\text{c.m.}} < 0.99$ and $E_{\text{miss}}^{\text{c.m.}} > 0.4\,\text{GeV}$ selection criteria are applied. For the eV^0 modes with $\tau_{\text{tag}}^\pm \to e^\pm \nu\nu$
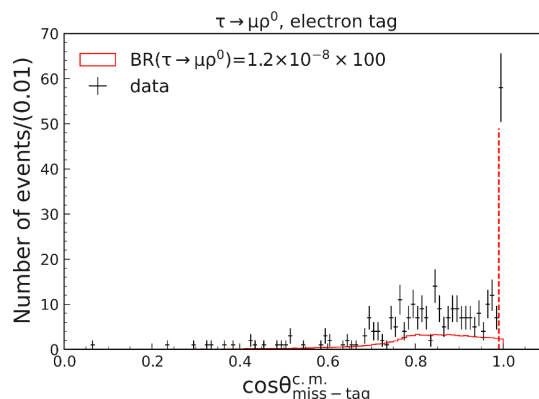
Figure 1. The $\cos\theta_{\mathrm{miss-tag}}^{\mathrm{c.m.}}$ distribution of the $\tau \to \mu\rho^0$ mode with a electron tag track after the reconstruction, particle identification, and photon conversion event suppression. Black points with error bars are the data outside the blind region. Red solid histogram is the signal MC. The signal MC is scaled to the number of events corresponding to 100 times as large branching fraction as the current upper limit. The red dashed line is the upper limit to remove the low-multiplicity events. The low-multiplicity events cluster around $\cos\theta_{\mathrm{miss-tag}}^{\mathrm{c.m.}} = 1$, whereas the other background events are linearly distributed in the region of $\cos\theta_{\mathrm{miss-tag}}^{\mathrm{c.m.}} > 0.8$.

or $\pi^\pm\nu$, $\cos\theta_{\mathrm{miss-tag}}^{\mathrm{c.m.}} < 0.97$ is applied. For the eV^0 modes with $\tau_{\mathrm{tag}}^\pm \to e^\pm\nu\nu$, $E_{\mathrm{miss}}^{\mathrm{c.m.}}$ should be larger than 0.4, 2.0, and 1.5 GeV for $e\phi$, $e\rho^0$, and the other eV^0 modes, respectively.

The remaining background events are mainly from the $q\bar{q}$ continuum ($q = u, d, s, c$) and generic $\tau^+\tau^-$ events, which have three charged pion tracks in the signal side. We use two-class Boosted Decision Tree (BDT) algorithms for signal and these background classification. The BDT library is LightGBM [44]. One BDT for each ℓV^0 mode is trained. These BDTs output signal probabilities using the following input variables:

- M_{V^0}, M_ν^2, $P_\nu^{\mathrm{c.m.}}$, T, P_ℓ^{sig}, $E_{\mathrm{tag}}^{\mathrm{hemi}}$, $\cos\theta_{\mathrm{miss-tag}}^{\mathrm{c.m.}}$

- (categorical variables) τ_{tag} decay mode, collision energy

- (additional for the $\ell\omega$ modes) $P_{\pi^0}^{\mathrm{sig}}$, E_γ^{low},

where M_{V^0} is the invariant mass of the vector meson, M_ν^2 is the missing mass squared, $P_\nu^{\mathrm{c.m.}}$ is the missing momentum in the c.m. frame, T is the magnitude of the thrust vector [39, 40], P_ℓ^{sig} is the momentum of the lepton in the signal side, $E_{\mathrm{tag}}^{\mathrm{hemi}}$ is the energy sum of the tracks and photons in the tag hemisphere, $P_{\pi^0}^{\mathrm{sig}}$ is the momentum of π^0 from ω and E_γ^{low} is the lower energy of the two photons from the π^0. The variables of neutrino kinematics (M_ν^2 and $P_\nu^{\mathrm{c.m.}}$) were not used for the event selection in the previous paper [29]. They are calculated from the momenta of the reconstructed τ_{sig} and τ_{tag}, where the energy of τ_{sig} is fixed to the half of the beam energy in the c.m. frame. The $q\bar{q}$ continuum background events can be effectively suppressed by a M_ν^2 selection in the hadronic tags, involving only one neutrino (figure 2).

The training, validation and evaluation of the BDTs are done with 40%, 10%, and 50% of the signal MC, respectively. Regarding the training and validation samples for
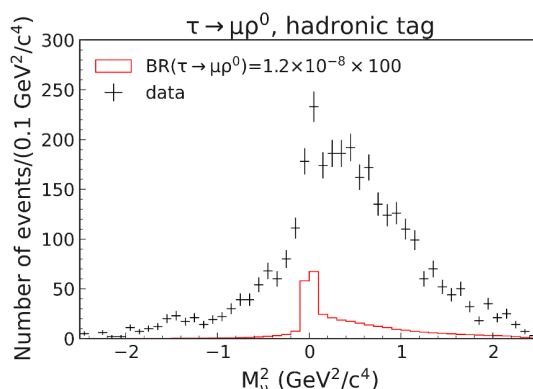
Figure 2. The M_ν^2 distribution of the $\tau \to \mu\rho^0$ mode with the hadronic tags after the event selection except for the requirement of the BDT output. Black points with error bars are the data outside the blind region. Red solid histogram is the signal MC. The signal MC is scaled to the number of events corresponding to 100 times as large a branching fraction as the current upper limit. The events constituting the upper tail of the signal distribution originate from wrong or missing π^0 in the tag side.

the background events, we utilize hadron background enhanced data that are obtained by removing the lepton identification for the signal-side leptons but with a lepton identification veto ($\mathcal{P}(e) \leq 0.9$ and $\mathcal{P}(\mu) \leq 0.95$) for all the signal-side tracks in the data. The hadron background enhanced data have a much larger number of events than the background data with the nominal selection criteria, whereas both data sets are composed mainly of three charged pions from τ decays or from continuum events. The training is done with 80% of the hadron background enhanced data and the validation is done with 20%. During BDT training, a weight is applied to each of the signal MC events such that the sum of the weights is equal to the number of background events. We monitor the area under curve (AUC) of the Receiver Operating Characteristic curve [45] using the validation samples after each training step, and choose the number of training steps with the best AUC score.

The event selection with the BDT output (BDT selection) is determined only by a target signal efficiency. The target signal efficiency is determined based on the signal efficiency with a cut-based event selection. In the cut-based event selection, the M_{V^0} windows correspond to $\pm 2\sigma$ of reconstructed mass distribution, and the M_ν^2 windows are set for each ℓV^0 mode and each τ_{tag} decay mode so that the expected number of background events inside the signal region (N_{BG}, see the next section) is approximately one or less. The target signal efficiency with the BDT selection is set as relatively 5% larger than that with the cut-based event selection, because we expect improvement in separating the signal events from the background events.

The finalized BDT selection shows similar N_{BG} to that of the cut-based event selection. The BDT selection is not applied to the $\ell\phi$ modes because N_{BG} in each of the two modes is small enough.
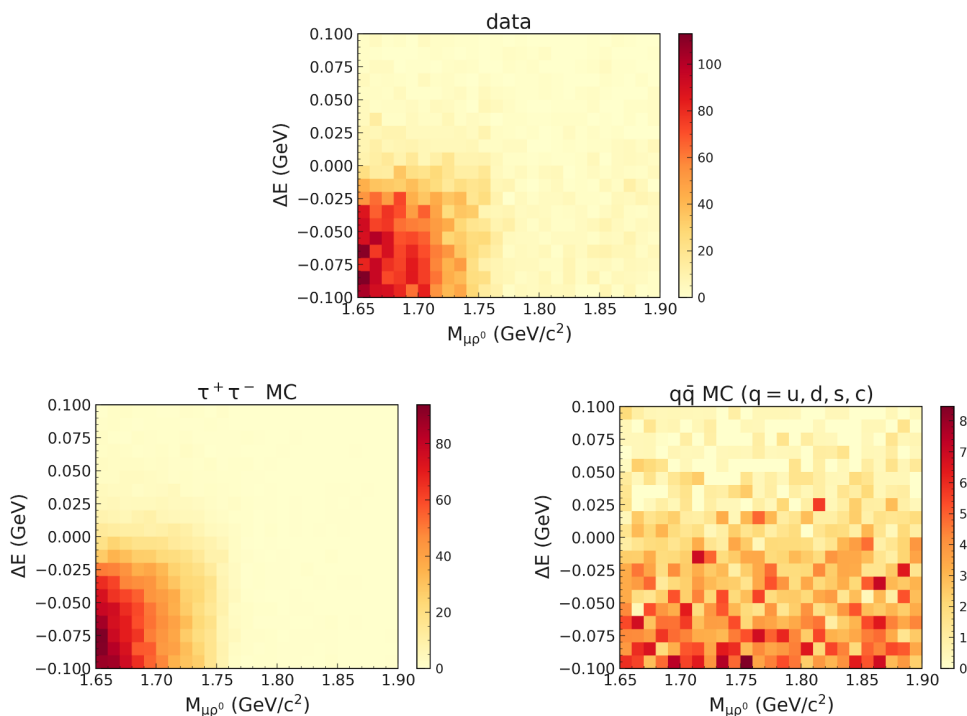
Figure 3. The $M_{\ell V^0}$ vs. ΔE distribution of the $\tau \to \mu\rho^0$ hadron background enhanced samples: the data (upper side), the generic $\tau^+\tau^-$ MC (lower left) and the $q\bar{q}$ continuum MC (lower right, $q = u, d, s, c$). The range of the ΔE axis is limited to the fitting region. The MC sets are scaled to the data. The low-multiplicity background events are negligible for the hadron background enhanced samples and are not shown in this figure.

4 Signal efficiency and background estimation

We define the signal region with an ellipse in the $M_{\ell V^0}-\Delta E$ plane inside the blind region. Most of the signal events cluster around $M_{\ell V^0} = 1.777\,\mathrm{GeV}/c^2$ and $\Delta E = 0\,\mathrm{GeV}$ with some correlation. The ellipse oblateness and the rotation angle are calculated from the covariance matrix of the signal MC distribution after the event selection. The center of the ellipse is the mean of the distribution. The ellipse size is determined to maximize the figure-of-merit (FOM) [46],

$$FOM = \frac{\varepsilon}{\frac{\alpha}{2} + \sqrt{N_{\mathrm{BG}}}}, \tag{4.1}$$

where ε is the signal efficiency inside the ellipse, α is the confidence coefficient ($\alpha = 1.64$ at 90% C.L.).

We estimate N_{BG} through interpolation from the sideband data. Here the sideband data is a set of data passing the event selection and inside the sideband region: $1.65\,\mathrm{GeV}/c^2 < M_{\ell V^0} < 1.9\,\mathrm{GeV}/c^2$ and $|\Delta E| < 0.1\,\mathrm{GeV}$ outside of the blind region. The interpolation is based on a function in the $M_{\ell V^0}-\Delta E$ plane. This function is obtained by fitting the distribution of the hadron background enhanced data within $|\Delta E| < 0.1\,\mathrm{GeV}$, and then it is scaled to the sideband data. Figure 3 shows the distributions of the hadron background

enhanced data and MC for the $\mu\rho^0$ mode. The function is:

$$F(M_{\ell V^0}, \Delta E) = f(M_{\ell V^0}) \times \frac{1}{1 + \exp[a_y(\Delta E - y_0)]} + c_0^{\text{flat}}, \tag{4.2}$$

$$f(x) = \begin{cases} \displaystyle\int_{x-5\sigma}^{x+5\sigma} g(x') \times \frac{1}{\sqrt{2\pi}\sigma} \exp\left[\frac{-(x-x')^2}{2\sigma^2}\right] \mathrm{d}x' & (V^0 = \rho^0, \omega) \\ c_1(x-x_0)^2 + c_0 & (V^0 = K^{*0}, \overline{K}^{*0}) \\ c_0 & (V^0 = \phi) \end{cases} \tag{4.3}$$

$$g(x) = \begin{cases} c_1[(x-x_0)^2 + k(x-x_0)] + c_0 & (x < x_0, V^0 = \rho^0) \\ c_1(x-x_0) + c_0 & (x < x_0, V^0 = \omega) \\ c_0 & (x \geq x_0) \end{cases}$$

where $f(x)$ represents the background distribution as a function of $M_{\ell V^0}$; c_1, c_0, x_0, and k are parameters that define the shape of the function; a_y represents sharpness of the sigmoid function along the ΔE axis; y_0 is the center of the sigmoid function; and c_0^{flat} is a term of flat background events in the $M_{\ell V^0}$–ΔE plane. We define $f(x)$ for each V^0 in eq. (4.3) and the functions for the $\ell\rho^0$ ($\ell\omega$) modes are smeared by a Gaussian with standard deviation (σ) of 6.6 (9.6) MeV/c^2. This σ corresponds to the mass resolution that affects the edge of the $M_{\ell V^0}$ distribution close to the τ mass for the $\tau^+\tau^-$ background. The edge is broad for the other modes owing to wrong mass assignment of fake kaons. The $\tau^+\tau^-$ background events for the $\ell\phi$ modes are included in c_0 because they are flat along the $M_{\ell V^0}$ axis in $1.65\,\text{GeV}/c^2 < M_{\ell V^0} < 1.9\,\text{GeV}/c^2$.

We obtain the optimal fit parameters by a likelihood fit using MINUIT [47]. The following region is excluded from the fitting to avoid $D^+ \to K^-\pi^+\pi^+$ and $D^+ \to \pi^+\phi$ background events, which cluster around the D meson mass: $1.83(1.82)\,\text{GeV}/c^2 \leq M_{\ell V^0} < 1.89\,\text{GeV}/c^2$ and $\Delta E < 0.04\,\text{GeV}$ for the μK^{*0} (eK^{*0}) and $\mu\phi$ ($e\phi$) modes.

The parameters of a_y, y_0, k, and x_0 are fixed at the fit results of the hadron background enhanced data within $|\Delta E| < 0.1\,\text{GeV}$. The fit uncertainties of these fixed parameters are included in the systematic uncertainty of N_{BG}. The other fit parameters correspond to the scale factors of each background component: generic $\tau^+\tau^-$ (c_1), and continuum and low-multiplicity background events (c_0 and c_0^{flat}). We fit the function floating these scale factors (c_1, c_0, and c_0^{flat}) to the sideband data. The same region around the D meson mass as for the fit to the hadron background enhanced data is excluded from the fitting for the $\ell\phi$ and ℓK^{*0} modes. The functions are integrated in the elliptical signal regions to deduce N_{BG}, which are in the range of 0.25–0.95.

Other systematic uncertainties on N_{BG} come from the difference of the $M_{\ell V^0}$–ΔE distributions between the sideband data and the hadron background enhanced data within $|\Delta E| < 0.1\,\text{GeV}$. The difference originates from the BDT selection and the lepton identification. We evaluate the amounts of changes of N_{BG} when the parameters—a_y, y_0, k, and x_0—are redetermined with another $M_{\ell V^0}$–ΔE distribution of the hadron background enhanced data changed by the BDT selection or weighted for the lepton identification as described below. The amounts of changes of N_{BG} are taken as the systematic uncertainties of N_{BG}.

Because the 80% of the hadron background enhanced data are used for the BDT training, we apply the BDT selection for the rest of data (the validation samples) and redetermine the parameters with those data. We estimate N_{BG} in the same way as previously described, and the amount of change from the nominal N_{BG} is obtained.

Each event of the sideband data has a pion misidentified as a lepton, which tends to have a lower momentum than the pions in the hadron background enhanced data. That is the reason of the difference in the $M_{\ell V^0}$–ΔE distributions because the fake rate, $R^{\mathrm{fake}}_{e(\mu)}(P, \theta)$, depends on the momentum P and θ of the track. We generate weighted hadron background enhanced data, where each event is weighted by the ratio of $R^{\mathrm{fake}}_{e(\mu)}(P, \theta)$ to $1 - R^{\mathrm{fake}}_e(P, \theta) - R^{\mathrm{fake}}_\mu(P, \theta)$ for the track, in order to conform the $M_{\ell V^0}$–ΔE distribution to the one of the sideband data. We redetermine the parameters with this weighted data and evaluate the change of N_{BG} again.

The statistical uncertainty of N_{BG} is calculated as follows: We generate 100 sets of pseudo-data for each mode in the $M_{\ell V^0}$–ΔE histogram. The content of each bin in the histogram is set randomly following a Poisson distribution, with the mean taken from the function fitted to the sideband data. We fit the function to each set of the pseudo-data to deduce N_{BG}, and the standard deviation of these N_{BG} is taken as the statistical uncertainty.

The major contribution to N_{BG} comes from the $M_{\ell V^0}$ flat term in eq. (4.2) (c_0 and c_0^{flat}), which corresponds to the continuum or low-multiplicity background events. The contribution of the generic $\tau^+\tau^-$ background events, which depends on $M_{\ell V^0}$, is about one-third as large as the other background contributions. We cannot distinguish the background components of the $\ell\phi$ modes through the fit to the data, because the generic $\tau^+\tau^-$ background events are distributed evenly along the $M_{\ell V^0}$ axis.

The systematic uncertainties of the expected number of signal events are listed in table 1. The dominant uncertainties are from the particle identification.

The track and photon energy resolutions in the MC are corrected such that the mass resolution of the $D^{(*)+}$ meson matches between the data and MC, where $D^{(*)+} \to K^-\pi^+\pi^+(\pi^0)$ is reconstructed with similar event selection criteria to the signal ones (e.g. $|\Delta E| < 0.5\,\mathrm{GeV}$). The uncertainty of the data mass resolution propagates to the uncertainties of the corrected energy resolutions. We generate two additional signal MC sets in which the track (photon) energy resolution is different by plus and minus its uncertainty, and take the half of the difference in the expected number of the signal events as the systematic uncertainty.

All the uncertainties in table 1 are summed in quadrature to yield the total systematic uncertainties shown in table 2.

5   Results

Figures 4 and 5 show the observed event distributions in the $M_{\ell V^0}$–ΔE plane. The observed number of events in the signal region (N_{obs}) has no excess over N_{BG}.

We set 90% C.L. upper limits on the branching fractions based on a Bayesian method with the use of Markov Chain Monte Carlo [49]. The probability density function of the branching fraction ($\mathcal{B}(\tau \to \ell V^0)$) is calculated assuming that N_{obs} follows a Poisson
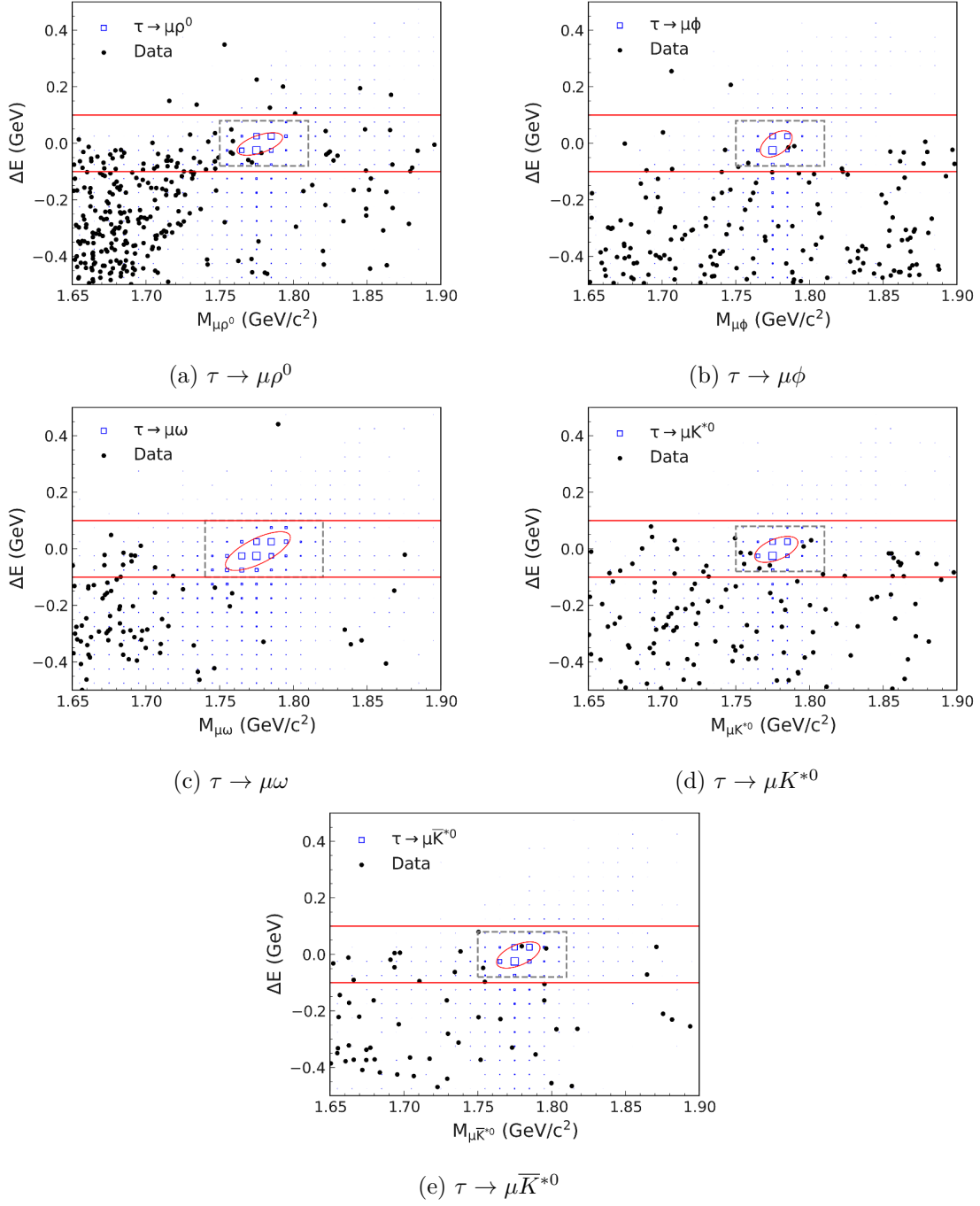
Figure 4. Observed event distributions of $M_{\ell V^0}$ vs. ΔE after the $\tau \to \mu V^0$ event selection. Black points are the data, blue squares show the signal MC distribution with an arbitrary normalization. The red elliptical lines are the signal regions. The estimations of the number of background events are done using the data between the red horizontal lines outside the blind regions (the gray dashed rectangles).
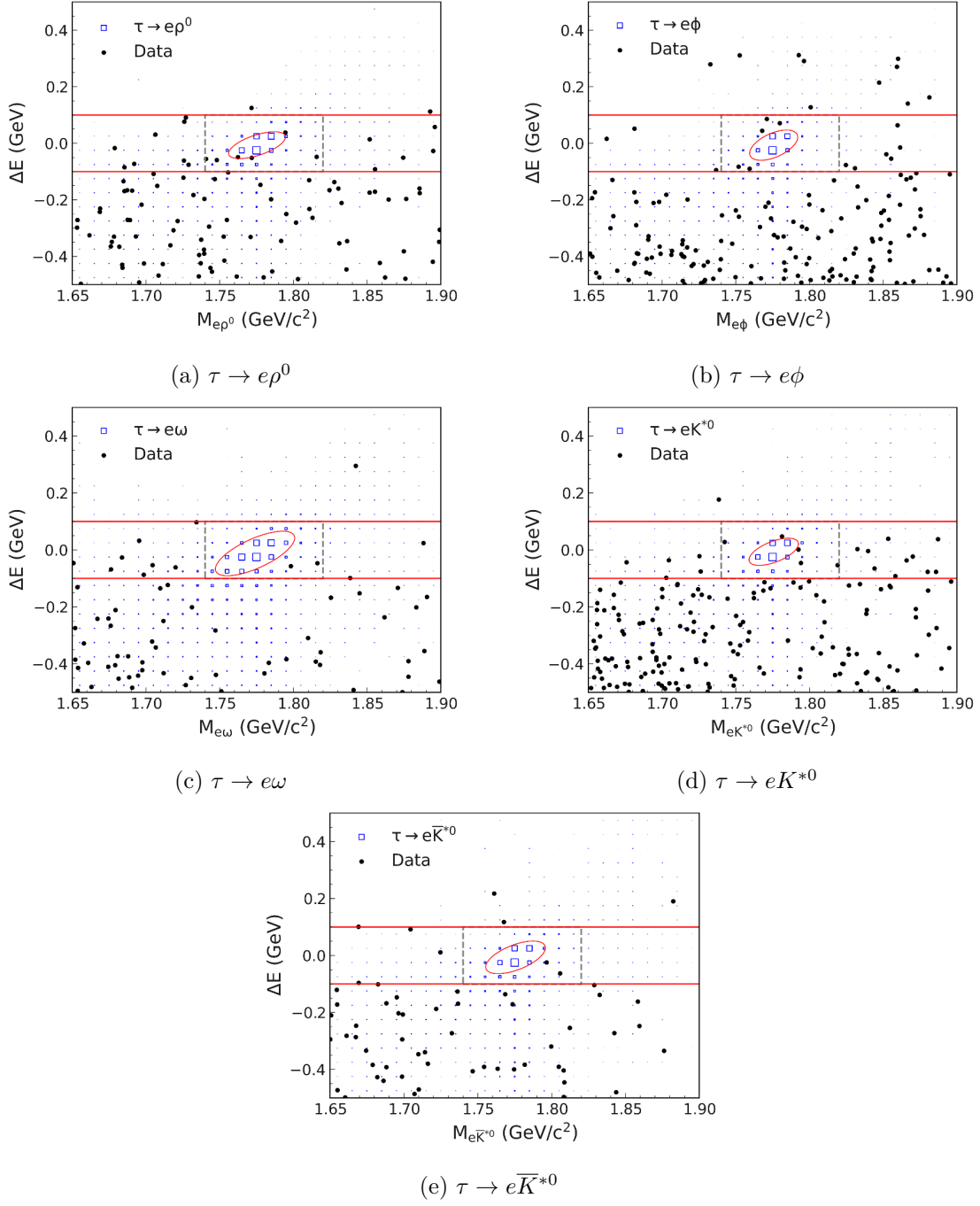
(a) $\tau \to e\rho^0$

(b) $\tau \to e\phi$

(c) $\tau \to e\omega$

(d) $\tau \to eK^{*0}$

(e) $\tau \to e\overline{K}^{*0}$

Figure 5. Observed event distributions of $M_{\ell V^0}$ vs. ΔE after the $\tau \to eV^0$ event selection. Black points are the data, blue squares show the signal MC distribution with an arbitrary normalization. The red elliptical lines are the signal regions. The estimations of the number of background events are done using the data between the red horizontal lines outside the blind regions (the gray dashed rectangles).

Source	σ_{syst} (%)
Integrated luminosity	1.4
$ee \to \tau\tau(\gamma)$ cross section [48]	0.3
$\mathcal{B}(\phi \to K^+K^-)$ and $\mathcal{B}(\omega \to \pi^+\pi^-\pi^0)$	1.2 and 0.7
Trigger efficiency	0.2–0.9
Tracking efficiency	$0.35 \times N_{\text{track}}$
Electron identification efficiency	$1.7 \times N_{\text{electron}}$
Muon identification efficiency	$1.8 \times N_{\text{muon}}$
$K^\pm$ and $\pi^\pm$ identification efficiency	1.6 (ρ^0), 1.8 (ϕ) and 1.1 $(K^{*0}$ and $\overline{K}^{*0})$
π^0 efficiency	$2.2 \times N_{\pi^0}$
Electron veto for hadrons	0.4–1.2
MC statistics	0.3–0.5
Track energy resolution	0.3–1.3
Photon energy resolution	0.0–0.4

Table 1. List of the systematic uncertainties of the expected number of signal events. The average number of tracks (particles) in the reconstructed $\tau^+\tau^-$ events for each signal mode is represented as $N_{\text{track(particle)}}$. When the uncertainty is different mode by mode, we show the range of the uncertainty.

distribution function whose mean value is the expected number of events (N_{exp}),

$$N_{\text{exp}} = L \times 2\sigma_{\tau\tau}\mathcal{B}(\tau \to \ell V^0) \times \varepsilon + N_{\text{BG}}, \tag{5.1}$$

where L is the integrated luminosity (980.4 ± 13.7 fb^{-1}), $\sigma_{\tau\tau}$ is the cross section of τ-pair production that is calculated with KKMC [48] (the weighted average of $\sigma_{\tau\tau}$ at all the beam energies is 0.916 ± 0.003 nb), and ε is the signal efficiency including the branching fraction of the V^0. We assume that these values ($L, \sigma_{\tau\tau}, \varepsilon,$ and N_{BG}) follow a Gaussian distribution with the width equal to the uncertainty of each value.

The upper limits on $\mathcal{B}(\tau \to \ell V^0)$ are listed in table 2. The average of the limits is better than that of the previous results using 854 fb^{-1} [29] by 30%. This is due to the additional 15% of integrated luminosity; the addition of $\pi^\pm\pi^\mp\pi^\pm\nu$ and $\pi^\pm\pi^0\pi^0\nu$ modes in τ_{tag} reconstruction, which increases the signal efficiency by 9.6%; and the event selection by multivariate analysis (BDT). The upper limit on $\mathcal{B}(\tau \to \mu\rho^0)$ is worse than that of the previous result, though the expected upper limit before unblinding is better. This is because we use the Bayesian limits instead of the Frequentist limits, which are negatively proportional to N_{BG} when N_{obs} is fixed.

6 Conclusion

To conclude, we searched for lepton-flavor-violating τ decays into one lepton and one vector meson using the full 980 fb^{-1} of Belle data. No statistically significant signal candidates are observed, and the 90% C.L. upper limits on the branching fraction are in the range of $(1.7–4.3) \times 10^{-8}$ for $\tau \to \mu V^0$ and $(1.7–2.4) \times 10^{-8}$ for $\tau \to e V^0$. The upper limits are

Mode	ε (%)	N_{BG}	σ_{syst} (%)	N_{obs}	$\mathcal{B}_{\mathrm{obs}}$ ($\times 10^{-8}$)
$\tau^{\pm} \to \mu^{\pm}\rho^0$	7.78	0.95±0.20(stat.) ±0.15(syst.)	4.6	0	< 1.7
$\tau^{\pm} \to e^{\pm}\rho^0$	8.49	0.80±0.27(stat.) ±0.04(syst.)	4.4	1	< 2.2
$\tau^{\pm} \to \mu^{\pm}\phi$	5.59	0.47±0.15(stat.) ±0.05(syst.)	4.8	0	< 2.3
$\tau^{\pm} \to e^{\pm}\phi$	6.45	0.38±0.21(stat.) ±0.00(syst.)	4.5	0	< 2.0
$\tau^{\pm} \to \mu^{\pm}\omega$	3.27	0.32±0.23(stat.) ±0.19(syst.)	4.8	0	< 3.9
$\tau^{\pm} \to e^{\pm}\omega$	5.41	0.74±0.43(stat.) ±0.06(syst.)	4.5	0	< 2.4
$\tau^{\pm} \to \mu^{\pm}K^{*0}$	4.52	0.84±0.25(stat.) ±0.31(syst.)	4.3	0	< 2.9
$\tau^{\pm} \to e^{\pm}K^{*0}$	6.94	0.54±0.21(stat.) ±0.16(syst.)	4.1	0	< 1.9
$\tau^{\pm} \to \mu^{\pm}\overline{K}^{*0}$	4.58	0.58±0.17(stat.) ±0.12(syst.)	4.3	1	< 4.3
$\tau^{\pm} \to e^{\pm}\overline{K}^{*0}$	7.45	0.25±0.11(stat.) ±0.02(syst.)	4.1	0	< 1.7

Table 2. The signal efficiency (ε), the expected number of background events (N_{BG}), total systematic uncertainty of the expected number of signal events (σ_{syst}), the number of observed events in the signal region (N_{obs}), and the observed 90% C.L. upper limits on the branching fraction ($\mathcal{B}_{\mathrm{obs}}$ (10^{-8})).

improved by 30% on average from the previous results. We achieve these improvements both with the reconsideration of the event selection criteria and with the $126\,\mathrm{fb}^{-1}$ of additional data set.

Acknowledgments

References

[1] A. Celis, V. Cirigliano and E. Passemar, *Model-discriminating power of lepton flavor violating τ decays*, *Phys. Rev. D* **89** (2014) 095014 [arXiv:1403.5781] [INSPIRE].

[2] MEG collaboration, *Search for the lepton flavour violating decay $\mu^+ \to e^+\gamma$ with the full dataset of the MEG experiment*, *Eur. Phys. J. C* **76** (2016) 434 [arXiv:1605.05081] [INSPIRE].

[3] SINDRUM collaboration, *Search for the Decay $\mu^+ \to e^+e^+e^-$*, *Nucl. Phys. B* **299** (1988) 1 [INSPIRE].

[4] SINDRUM II collaboration, *A Search for muon to electron conversion in muonic gold*, *Eur. Phys. J. C* **47** (2006) 337 [INSPIRE].

[5] BaBar collaboration, *Evidence for an excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays*, *Phys. Rev. Lett.* **109** (2012) 101802 [arXiv:1205.5442] [INSPIRE].

[6] BaBar collaboration, *Measurement of an Excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ Decays and Implications for Charged Higgs Bosons*, *Phys. Rev. D* **88** (2013) 072012 [arXiv:1303.0571] [INSPIRE].

[7] Belle collaboration, *Measurement of the branching ratio of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ relative to $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays with hadronic tagging at Belle*, *Phys. Rev. D* **92** (2015) 072014 [arXiv:1507.03233] [INSPIRE].

[8] BELLE collaboration, *Measurement of the τ lepton polarization and $R(D^*)$ in the decay $\bar{B} \to D^* \tau^- \bar{\nu}_\tau$*, *Phys. Rev. Lett.* **118** (2017) 211801 [arXiv:1612.00529] [INSPIRE].

[9] BELLE collaboration, *Measurement of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ with a semileptonic tagging method*, *Phys. Rev. Lett.* **124** (2020) 161803 [arXiv:1910.05864] [INSPIRE].

[10] LHCb collaboration, *Measurement of the ratio of branching fractions $\mathcal{B}(\bar{B}^0 \to D^{*+} \tau^- \bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \to D^{*+} \mu^- \bar{\nu}_\mu)$*, *Phys. Rev. Lett.* **115** (2015) 111803 [*Erratum ibid.* **115** (2015) 159901] [arXiv:1506.08614] [INSPIRE].

[11] LHCb collaboration, *Measurement of the ratio of the $B^0 \to D^{*-} \tau^+ \nu_\tau$ and $B^0 \to D^{*-} \mu^+ \nu_\mu$ branching fractions using three-prong τ-lepton decays*, *Phys. Rev. Lett.* **120** (2018) 171802 [arXiv:1708.08856] [INSPIRE].

[12] LHCb collaboration, *Test of Lepton Flavor Universality by the measurement of the $B^0 \to D^{*-} \tau^+ \nu_\tau$ branching fraction using three-prong τ decays*, *Phys. Rev. D* **97** (2018) 072013 [arXiv:1711.02505] [INSPIRE].

[13] BELLE collaboration, *Lepton-Flavor-Dependent Angular Analysis of $B \to K^* \ell^+ \ell^-$*, *Phys. Rev. Lett.* **118** (2017) 111801 [arXiv:1612.05014] [INSPIRE].

[14] LHCb collaboration, *Measurement of CP-Averaged Observables in the $B^0 \to K^{*0} \mu^+ \mu^-$ Decay*, *Phys. Rev. Lett.* **125** (2020) 011802 [arXiv:2003.04831] [INSPIRE].

[15] LHCb collaboration, *Angular Analysis of the $B^+ \to K^{*+} \mu^+ \mu^-$ Decay*, *Phys. Rev. Lett.* **126** (2021) 161802 [arXiv:2012.13241] [INSPIRE].

[16] LHCb collaboration, *Branching Fraction Measurements of the Rare $B_s^0 \to \phi \mu^+ \mu^-$ and $B_s^0 \to f_2'(1525) \mu^+ \mu^-$- Decays*, *Phys. Rev. Lett.* **127** (2021) 151801 [arXiv:2105.14007] [INSPIRE].

[17] C. Hati, J. Kriewald, J. Orloff and A.M. Teixeira, *The fate of V_1 vector leptoquarks: the impact of future flavour data*, *Eur. Phys. J. C* **81** (2021) 1066 [arXiv:2012.05883] [INSPIRE].

[18] L. Di Luzio et al., *Maximal Flavour Violation: a Cabibbo mechanism for leptoquarks*, *JHEP* **11** (2018) 081 [arXiv:1808.00942] [INSPIRE].

[19] J. Kumar, D. London and R. Watanabe, *Combined Explanations of the $b \to s \mu^+ \mu^-$ and $b \to c \tau^- \bar{\nu}$ Anomalies: a General Model Analysis*, *Phys. Rev. D* **99** (2019) 015007 [arXiv:1806.07403] [INSPIRE].

[20] A. Crivellin, C. Greub, D. Müller and F. Saturnino, *Importance of Loop Effects in Explaining the Accumulated Evidence for New Physics in B Decays with a Vector Leptoquark*, *Phys. Rev. Lett.* **122** (2019) 011805 [arXiv:1807.02068] [INSPIRE].

[21] A. Crivellin, D. Müller and F. Saturnino, *Flavor Phenomenology of the Leptoquark Singlet-Triplet Model*, *JHEP* **06** (2020) 020 [arXiv:1912.04224] [INSPIRE].

[22] P.S. Bhupal Dev, R. Mohanta, S. Patra and S. Sahoo, *Unified explanation of flavor anomalies, radiative neutrino masses, and ANITA anomalous events in a vector leptoquark model*, *Phys. Rev. D* **102** (2020) 095012 [arXiv:2004.09464] [INSPIRE].

[23] LHCb collaboration, *Test of lepton universality in beauty-quark decays*, *Nature Phys.* **18** (2022) 277 [arXiv:2103.11769] [INSPIRE].

[24] LHCb collaboration, *Test of lepton universality in $b \to s \ell^+ \ell^-$ decays*, arXiv:2212.09152 [INSPIRE].

[25] A. Ilakovac, *Lepton flavor violation in the standard model extended by heavy singlet Dirac neutrinos*, *Phys. Rev. D* **62** (2000) 036010 [hep-ph/9910213] [INSPIRE].

[26] Z.-H. Li, Y. Li and H.-X. Xu, *Unparticle-Induced Lepton Flavor Violating Decays* $\tau \to \ell(V^0, P^0)$, *Phys. Lett. B* **677** (2009) 150 [`arXiv:0901.3266`] [ɪɴSPIRE].

[27] A. Arhrib, R. Benbrik and C.-H. Chen, *Lepton flavor violating tau decays in type-III seesaw mechanism*, *Phys. Rev. D* **81** (2010) 113003 [`arXiv:0903.1553`] [ɪɴSPIRE].

[28] I. Pacheco and P. Roig, *Lepton flavour violation in hadron decays of the tau lepton within the littlest Higgs model with T-parity*, *JHEP* **09** (2022) 144 [`arXiv:2207.04085`] [ɪɴSPIRE].

[29] BELLE collaboration, *Search for Lepton-Flavor-Violating tau Decays into a Lepton and a Vector Meson*, *Phys. Lett. B* **699** (2011) 251 [`arXiv:1101.0755`] [ɪɴSPIRE].

[30] S. Kurokawa and E. Kikutani, *Overview of the KEKB accelerators*, *Nucl. Instrum. Meth. A* **499** (2003) 1 [ɪɴSPIRE].

[31] T. Abe et al., *Achievements of KEKB*, *PTEP* **2013** (2013) 03A001 [ɪɴSPIRE].

[32] BELLE collaboration, *The Belle Detector*, *Nucl. Instrum. Meth. A* **479** (2002) 117 [ɪɴSPIRE].

[33] BELLE collaboration, *Physics Achievements from the Belle Experiment*, *PTEP* **2012** (2012) 04D001 [`arXiv:1212.5342`] [ɪɴSPIRE].

[34] S. Jadach, B.F.L. Ward and Z. Was, *The Precision Monte Carlo event generator K K for two fermion final states in e^+e^- collisions*, *Comput. Phys. Commun.* **130** (2000) 260 [`hep-ph/9912214`] [ɪɴSPIRE].

[35] D.J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth. A* **462** (2001) 152 [ɪɴSPIRE].

[36] S. Jadach, E. Richter-Was, B.F.L. Ward and Z. Was, *Monte Carlo program BHLUMI-2.01 for Bhabha scattering at low angles with Yennie-Frautschi-Suura exponentiation*, *Comput. Phys. Commun.* **70** (1992) 305 [ɪɴSPIRE].

[37] F.A. Berends, P.H. Daverveldt and R. Kleiss, *Monte Carlo Simulation of Two Photon Processes. 2. Complete Lowest Order Calculations for Four Lepton Production Processes in electron Positron Collisions*, *Comput. Phys. Commun.* **40** (1986) 285 [ɪɴSPIRE].

[38] R. Brun et al., *GEANT Detector Description and Simulation Tool*, CERN-W5013 (1994) [`DOI:10.17181/CERN.MUHF.DMJ1`] [ɪɴSPIRE].

[39] S. Brandt, C. Peyrou, R. Sosnowski and A. Wroblewski, *The Principal axis of jets. An Attempt to analyze high-energy collisions as two-body processes*, *Phys. Lett.* **12** (1964) 57 [ɪɴSPIRE].

[40] E. Farhi, *A QCD Test for Jets*, *Phys. Rev. Lett.* **39** (1977) 1587 [ɪɴSPIRE].

[41] K. Hanagaki et al., *Electron identification in Belle*, *Nucl. Instrum. Meth. A* **485** (2002) 490 [`hep-ex/0108044`] [ɪɴSPIRE].

[42] A. Abashian et al., *Muon identification in the Belle experiment at KEKB*, *Nucl. Instrum. Meth. A* **491** (2002) 69 [ɪɴSPIRE].

[43] E. Nakano, *Belle PID*, *Nucl. Instrum. Meth. A* **494** (2002) 402 [ɪɴSPIRE].

[44] G. Ke et al., *LightGBM: A Highly Efficient Gradient Boosting Decision Tree*, in *Advances in Neural Information Processing Systems*, I. Guyon et al., eds., vol. 30, Curran Associates, Inc., (2017), https://proceedings.neurips.cc/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf.

[45] J.A. Hanley and B.J. McNeil, *The meaning and use of the area under a receiver operating characteristic (ROC) curve*, *Radiology* **143** (1982) 29.

[46] G. Punzi, *Sensitivity of searches for new signals and its optimization*, *eConf* **C030908** (2003) MODT002 [`physics/0308063`] [InSPIRE].

[47] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, *Comput. Phys. Commun.* **10** (1975) 343 [InSPIRE].

[48] S. Banerjee, B. Pietrzyk, J.M. Roney and Z. Was, *Tau and muon pair production cross-sections in electron-positron annihilations at $\sqrt{s} = 10.58\,GeV$*, *Phys. Rev. D* **77** (2008) 054012 [`arXiv:0706.3235`] [InSPIRE].

[49] A. Caldwell, D. Kollar and K. Kroninger, *BAT: The Bayesian Analysis Toolkit*, *Comput. Phys. Commun.* **180** (2009) 2197 [`arXiv:0808.2552`] [InSPIRE].