Search for the Lepton-Flavor-Violating Decay $\tau^- \rightarrow \mu^- \eta$ at Belle


(Belle Collaboration)

1Budker Institute of Nuclear Physics, Novosibirsk
2Chiba University, Chiba
3University of Cincinnati, Cincinnati, Ohio 45221
4Gyeongsang National University, Chinju
5University of Hawaii, Honolulu, Hawaii 96822
6High Energy Accelerator Research Organization (KEK), Tsukuba
7Institute of High Energy Physics, Chinese Academy of Sciences, Beijing
8Institute of High Energy Physics, Vienna
9Institute for Theoretical and Experimental Physics, Moscow
10J. Stefan Institute, Ljubljana
11Kyungpook National University, Taegu
12Korea University, Seoul
13Institut de Physique des Hautes Énergies, Université de Lausanne, Lausanne
14University of Ljubljana, Ljubljana
15University of Maribor, Maribor
16University of Melbourne, Victoria
17Nagoya University, Nagoya
18Nara Women's University, Nara
19National Lien-Ho Institute of Technology, Miaoli
20National Tsing-Hua University, Taipei
21H. Niewodniczanski Institute of Nuclear Physics, Krakow
22Niigata University, Niigata
23Osaka City University, Osaka
24Osaka University, Osaka
25Panjab University, Chandigarh
26Peking University, Beijing
27Princeton University, Princeton, New Jersey 08545
28University of Science and Technology of China, Hefei
29Seoul National University, Seoul
30Sungkyunkwan University, Suwon
31University of Sydney, Sydney, New South Wales
32Tata Institute of Fundamental Research, Bombay
33Toho University, Funabashi

© 2004 The American Physical Society
Among the possible lepton flavor violating (LFV) decays of the $\tau$ lepton, $\tau^{-} \rightarrow \mu^{-}\eta$ is the process that provides the most stringent bound on Higgs-mediated LFV. Sher [1] has pointed out that a flavor nondiagonal lepton-lepton-Higgs Yukawa coupling could be induced if slepton mixing is large. The $\mu$-$\tau$-Higgs vertex is particularly promising since mixing between left-handed smuons and staus is large in many supersymmetric models [2]. This mechanism initially led various authors [3] to study the enhancement of the LFV decay $\tau^{-} \rightarrow \mu^{-}\eta$. The $\tau^{-} \rightarrow \mu^{-}\eta$ branching fraction is given by

$$B(\tau^{-} \rightarrow \mu^{-}\eta) = 0.84 \times 10^{-6} \times \left( \frac{\tan \beta}{100 \text{ GeV}} \right)^6 \left( \frac{m_{\Lambda}}{60} \right)^4,$$

where $m_{\Lambda}$ is the pseudoscalar Higgs mass and $\tan \beta$ is the ratio of the vacuum expectation values ($\langle H_u \rangle / \langle H_d \rangle$). In such models, $\tau^{-} \rightarrow \mu^{-}\eta$ and $\tau^{-} \rightarrow 3\mu$ are particularly sensitive to LFV at large tan$\beta$.

Previous experimental studies of $\tau^{-} \rightarrow \mu^{-}\eta$ by ARGUS [4] and CLEO [5] set 90% confidence level upper limits on the branching fraction of $7.3 \times 10^{-5}$ from 0.387 fb$^{-1}$ of data, and $9.6 \times 10^{-6}$ from 4.68 fb$^{-1}$ of data, respectively. We present here a new search based on a data sample of 84.3 fb$^{-1}$, equivalent to 76.9M $\tau^+\tau^-$ pairs, collected at the Y(4S) resonance with the Belle detector at the KEKB asymmetric $e^+e^-$ collider [6]. A description of the detector can be found in Ref. [7].

For Monte Carlo (MC) studies, the following programs were used to generate background (BG) events: KORALB/TAULA [8] for $\tau^+\tau^-$ processes, QQ [9] for $B\bar{B}$ and continuum, BHLUMI [10] for Bhabha, KKMC [11] for $\mu^+\mu^-$, and AAFH [12] for two-photon processes. The $\tau^{-} \rightarrow \mu^{-}\eta$ decay is initially assumed to have a uniform angular distribution in the $\tau$'s rest frame. The Belle detector response is simulated by a GEANT3 [13] based program. Kinematical variables are evaluated in the laboratory frame, unless denoted by the superscript 'cm' in which case they are evaluated in the center-of-mass frame. Two $\eta$ decay modes are considered in this analysis: $\eta \rightarrow \gamma \gamma$ ($B = 39.43 \pm 0.26\%$) and $\eta \rightarrow \pi^+ \pi^- \pi^0$ ($B = 22.6 \pm 0.4\%$) [14].

For $\eta \rightarrow \gamma \gamma$, we search for events containing exactly two oppositely charged tracks and two or more photons, two of which form an $\eta$. The events should be consistent with a $\tau^+\tau^-$ event, in which one $\tau$ decays to $\mu\eta$ and the other $\tau$ decays to a charged particle other than a muon with any number of $\gamma$’s and neutrinos.

To select candidate events we require the momentum of each track, $p$, and the energy of each photon, $E_\gamma$, to satisfy $p > 0.1 \text{ GeV/c}$ and $E_\gamma > 0.1 \text{ GeV}$. The tracks and photons are required to be detected in the barrel or end cap of Belle: $-0.866 < \cos \theta < 0.956$. To exclude Bhabha, $\mu^+\mu^-$, and two-photon events, we require the total energy to be between 5 and 10 GeV in the cm frame, as shown in Fig. 1(a).

In the cm frame the events are subdivided into two hemispheres by a plane perpendicular to the thrust axis. The signal side should contain a muon and two photons. Muons are identified using a likelihood ratio $L_\mu/(L_\mu + L_K + L_\pi)$, where the $L$ are likelihoods for the various track identification hypotheses [15]; we require $P_{\mu} > 0.9$, accepting $87.5 \pm 0.3\%$ of muons but only $1.4 \pm 0.1\%$ of pions in the sample under study. (A similar ratio $P_e$ [16] is used to identify electrons.) An $\eta$ meson produced in a two-body $\tau$ decay has on average a higher momentum than $\eta$ mesons from other sources. Therefore, the photons used in $\eta$ candidates are required to have a rather high energy $E_\gamma > 0.22 \text{ GeV}$. To reduce background, events are rejected when two $\gamma$’s, one from the...
signal side and the other from the tagging side (γ′), have a resolution-normalized π⁰ mass in the range −5 < Sπ⁰ < 5, where \( S_{\pi^0} = (m_{\gamma\gamma} - 0.135 \text{ GeV/c}^2) / \sigma_{\pi^0} \) and \( \sigma_{\pi^0} \) is in the range 5–8 MeV/c². This π⁰ veto rejects 86% of BG events while retaining 75% of the signal. To further reduce BG, the cosine of the opening angle between the \( \mu \) and \( \gamma \gamma \) on the signal side is required to satisfy \( 0.5 < \cos \theta_{\mu - \gamma \gamma} < 0.95 \), as shown in Fig. 1(b).

In the tagging side hemisphere, the charged track should not be a muon (\( P_\mu < 0.6 \) is imposed), but may be either an electron, i.e., have an electron likelihood ratio \( P_e > 0.9 \) [16] or a hadron (\( P_\mu < 0.6 \) and \( P_e < 0.9 \)). If an electron is found, the number of photons and the electron momentum are constrained by \( n_\gamma \leq 2 \) and \( p_e > 0.7 \text{ GeV/c} \). If a hadron is found, the constraints are \( n_\gamma \geq 2 \) and \( p_{\text{had}} > 0.1 \text{ GeV/c} \).

The following two criteria are imposed on the missing momentum and energy in the event. To ensure that the missing particles are neutrinos rather than γ’s or charged particles that are outside of the detector acceptance, we require that the direction of the missing momentum should satisfy \( -0.866 < \cos \theta_{\text{miss}} < 0.956 \). Since neutrinos are emitted only on the tagging side, the direction of the missing momentum should be contained on the tagging side: \( \cos \theta_{\text{thrust-miss}} < -0.55 \). The correlation between the missing momentum, \( p_{\text{miss}} \) and the missing mass squared, \( m_{\text{miss}}^2 \), shown in Figs. 1(c) and 1(d) for signal and generic \( \pi^+ \pi^- \) MC, is utilized for additional BG rejection.

The \( \eta \) candidate is selected based on the signal side \( \gamma \gamma \) invariant mass in terms of the resolution-normalized \( \eta \) mass, \( -5 < S_\eta < 3 \), where \( S_\eta = (m_{\gamma\gamma} - 0.547 \text{ GeV/c}^2) / \sigma_{\eta} \) and \( \sigma_{\eta} \) is 12 MeV/c². The resulting \( S_\eta \) distributions for signal and generic \( \pi^+ \pi^- \) MC and data are shown in Fig. 2(a).

The application of these selection criteria to the data set results in a sample of 18 events. The detection efficiency is determined from MC studies to be \( e(2\gamma) = 9.3\% \). In MC, small backgrounds from the three following processes survive: 8.6 ± 2.2 events from generic \( \tau^+ \tau^- \), 2.5 ± 1.8 events from \( \mu^+ \mu^- \), and 5.8 ± 2.2 events from the continuum.

For the \( \eta \rightarrow \pi^+ \pi^- \pi^0 \) mode, we search for events containing four charged tracks (net charge = 0) and two or more photons. Because of the higher multiplicity compared to the \( \eta \rightarrow \gamma \gamma \) mode, the detection efficiency is smaller; however, the extra reconstruction constraint in the \( \eta \) decay chain improves the background rejection power. The selection criteria are similar to those in the \( \eta \rightarrow \gamma \gamma \) case with the differences listed below.

The minimum photon energy requirement is reduced from 0.1 to 0.05 GeV, since the photons from this decay mode have a softer energy distribution compared to those in \( \eta \rightarrow \gamma \gamma \). The signal side hemisphere should have three tracks and two or more photons. One track must be a muon (\( P_\mu > 0.9 \)), but particle identification is not performed on the other two tracks—they are treated as pions. We also require that one \( \pi^0 \) be reconstructed from the photons in the signal hemisphere, such that \( -5 < S_{\pi^0} < 5 \). Figure 2(b) shows the reconstructed mass of \( \eta \) candidates.

**FIG. 1** (color online). Some kinematical distributions from MC and data: (a) total energy, \( E_{\text{total}} \), (b) cosine of the opening angle between the \( \mu \) and \( \gamma \gamma \) on the signal side, (c) \( m_{\text{miss}} \) vs \( p_{\text{miss}} \) for signal MC events, (d) \( m_{\text{miss}}^2 \) vs \( p_{\text{miss}} \) for generic decays of \( \tau^+ \tau^- \) MC. The shaded histograms represent the signal \( \tau^+ \tau^- \) MC, the hatched histograms represent the combined \( \mu^+ \mu^- \) and Bhabha MC, the open histograms represent the combined \( BB \), continuum, and two-photon MC, and the triangles represent the data. The selected regions in (a) and (b) are indicated by the arrows. The selected regions in (c) and (d) are the areas between the two lines: \( p_{\text{miss}} > -2.615 \times m_{\text{miss}}^2 - 0.191 \) and \( p_{\text{miss}} > 1.238 \times m_{\text{miss}}^2 - 0.869 \).

**FIG. 2** (color online). (a) Invariant mass of \( \gamma \gamma \) in terms of the resolution-normalized \( \eta \) mass, \( S_\eta^\gamma \), in the \( \eta \rightarrow \gamma \gamma \) case, and (b) \( \eta \) mass from \( \eta \rightarrow \pi^+ \pi^- \pi^0 \) reconstruction. Signal and generic \( \tau^+ \tau^- \) MC distributions are indicated by the shaded and open histograms, and the triangles represent the data. The selection region is indicated by the arrows.
After the application of these requirements, 67 events remain in the data, while the generic $\tau^+\tau^-$ MC predicts a contribution of $38.0 \pm 4.6$ events, and the continuum MC predicts $15.6 \pm 3.5$ events. The detection efficiency is $\epsilon(3\pi) = 5.6\%$.

The final evaluation of the number of signal candidates is performed by defining a signal region in the $M_{\mu\eta}\Delta E$ plane, where the candidate $\mu\eta$ system should have an invariant mass ($M_{\mu\eta}$) close to the $\tau$-lepton mass and an energy close to the beam energy in the cm frame, i.e., $\Delta E = E_{\mu\eta}^\text{cm} - E_{\text{beam}}^\text{cm} \approx 0$. Figures 3(a) and 3(b) show scatterplots of the signal MC in the $M_{\mu\eta}\Delta E$ plane for the $\eta \rightarrow \gamma\gamma$ and $\eta \rightarrow \pi^+\pi^-\pi^0$ modes, respectively. The signal exhibits a long low-energy tail due to initial-state radiation and calorimeter energy leakage for photons. Parametrizing the $M_{\mu\eta}$ and $\Delta E$ signal distributions with asymmetric Gaussian functions, the $M_{\mu\eta}$ and $\Delta E$ resolutions are found to be $\sigma_{M_{\mu\eta}}^{\text{low/high}} = 25.8 \pm 0.7/15.3 \pm 0.4$ MeV/$c^2$ and $\sigma_{\Delta E}^{\text{low/high}} = 69.7 \pm 3.0/34.7 \pm 1.2$ MeV for the $\eta \rightarrow \gamma\gamma$ mode, and $\sigma_{M_{\mu\eta}}^{\text{low/high}} = 13.8 \pm 0.4/9.0 \pm 0.4$ MeV/$c^2$ and $\sigma_{\Delta E}^{\text{low/high}} = 44.4 \pm 2.3/22.6 \pm 1.3$ MeV for the $\eta \rightarrow \pi^+\pi^-\pi^0$ mode, where the “low/high” superscript indicates the lower/higher energy side of the peak. To optimize the sensitivity, we take an elliptically shaped signal region in the $M_{\mu\eta}\Delta E$ plane with a signal acceptance, $\Omega$, of 90% as shown in Fig. 3.

Figure 3 shows the final data distributions for a $\pm 10\sigma_{M_{\mu\eta}/\Delta E}$ region in the $M_{\mu\eta}\Delta E$ plane. In the signal region, there are no events in either the data or background MC. Outside the signal region, seven events for the $\eta \rightarrow \gamma\gamma$ mode and two events for the $\eta \rightarrow \pi^+\pi^-\pi^0$ mode are observed in data, while background MC predicts $3.7 \pm 2.4$ and $0.0_{-0.0}^{+4.0}$ events, respectively. The observed data yields are consistent with MC.The BG yield in the signal region, estimated from the sidebands, is found to be $0.5 \pm 0.2$ for $\eta \rightarrow \gamma\gamma$ and $0.0_{-0.0}^{+4.0}$ events for $\eta \rightarrow \pi^+\pi^-\pi^0$.

As no events are observed, an upper limit on the branching fraction, at 90% C.L., is given by

$$B(\tau^- \rightarrow \mu^-\eta) < \frac{2.3}{2(\epsilon\Omega \times B_\eta) \times N_{\tau^-\tau^+}}.$$

where $B_\eta$ is the branching fraction of $\eta$ decay to either $\gamma\gamma$ or $\pi^+\pi^-\pi^0$. The calculated upper limits, at 90% C.L., are thus found to be $4.6 \times 10^{-7}$ for the $\eta \rightarrow \gamma\gamma$ mode, and $13.1 \times 10^{-7}$ for the $\eta \rightarrow \pi^+\pi^-\pi^0$ mode. Combining the two decay modes, we obtain $\epsilon\Omega \times B_\eta = 4.4\%$ and $B(\tau^- \rightarrow \mu^-\eta) < 3.4 \times 10^{-7}$ at 90% C.L.

The systematic uncertainties on the detection sensitivity, $2(\epsilon\Omega \times B_\eta) \times N_{\tau^-\tau^+}$, arise from the track reconstruction efficiency of the tag and the muon (2.0% for both $\eta \rightarrow \gamma\gamma$ and $\eta \rightarrow \pi^+\pi^-\pi^0$ modes), $\eta$ reconstruction efficiency (2.0% and 4.2%, the latter value includes the uncertainties in the tracking efficiency for charged pions and the $\pi^0$ reconstruction efficiency), $\pi^0$ veto (5.5% and none), muon identification efficiency (4.0% and 4.0%), trigger efficiency (1.4% and 1.4%), beam background (2.3% and 2.1%), luminosity (1.4% and 1.4%), $B_\eta$ (0.7% and 1.8%), and MC statistics (1.3% and 2.1%). Adding all of these components in quadrature, the total uncertainty is evaluated to be 8.1% for $\eta \rightarrow \gamma\gamma$ and 7.3% for $\eta \rightarrow \pi^+\pi^-\pi^0$. For the combination of the two decay modes the systematic uncertainty is $\pm 7.9\%$.

This systematic uncertainty is included in the upper limit following Ref. [17], where the detection sensitivity, $2(\epsilon\Omega \times B_\eta) \times N_{\tau^-\tau^+}$, is modeled by a Gaussian distribution having a width given by the systematic error quoted above. There is no appreciable effect on the branching fraction, $B$.

The angular distribution of the $\tau^- \rightarrow \mu^-\eta$ decay has a strong dependence on the LFV interaction structure [18],

![Figure 3](image-url)
and spin correlations between the $\tau$'s at the signal and tagged sides must be considered. To evaluate the maximum possible variation, $V - A$ and $V + A$ interactions are assumed; no statistically significant difference in the $M_{\mu\eta}\Delta E$ distribution or in the efficiency is found compared to the case of the uniform distribution.

As a result, we obtain an upper limit on the branching fraction for the LFV $\tau^- \to \mu^- \eta$ decay of

$$B(\tau^- \to \mu^- \eta) < 3.4 \times 10^{-7},$$

at 90% C.L. This result improves the previous upper limit, $B(\tau^- \to \mu^- \eta) < 9.6 \times 10^{-6}$ [5], by a factor of 30.

Using Eq. (1), which was derived in a seesaw MSSM with a specific neutrino mass texture, our upper limit restricts the allowed parameter space for $m_A$ and $\tan \beta$, as indicated in Fig. 4, where our boundary is indicated in the cases of 90% and 95% C.L. Figure 4 also shows the 95% C.L. constraints from high energy collider experiments at LEP [19] and CDF [20]. Our result has a sensitivity close to that of the CDF experiment, achieved by searching for $pp \to A/\phi b\bar{b} \to b\bar{b}b\bar{b}$, where $\phi$ is a $CP$-even neutral Higgs state and $A$ is a $CP$-odd state in the MSSM.

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 NII for valuable computing and SuperSINET network support. We are grateful to Y. Okada for fruitful discussions on theoretical aspects of $\tau$ spin correlations. We also thank A. Dedes and J. Hisano for important and useful comments about model dependence of limits on SUSY parameters. We acknowledge support from MEXT and JSPS (Japan); ARC and DEST (Australia); NSFC (Contract No. 10175071, China); DST (India); the BK21 program of MOEHRD and the CHEP SRC program of KOSEF (Korea); KBN (Contract No. 2P03B 01324, Poland); MIST (Russia); MESS (Slovenia); NSC and MOE (Taiwan); and DOE (USA).

*On leave from Nova Gorica Polytechnic, Nova Gorica.