Search for the Lepton-Flavor-Violating Leptonic $B$ Decays

$B^0 \rightarrow \mu^\pm \tau^\mp$ and $B^0 \rightarrow e^\pm \tau^\mp$

A. Bornheim,1 E. Lipeles,1 S. P. Pappas,1 A. J. Weinstein,1 R. A. Briere,2 G. P. Chen,2 T. Ferguson,2 G. Tatishvili,2 H. Vogel,2 M. E. Watkins,2 N. E. Adam,3 J. P. Alexander,3 K. Berkelman,3 D. G. Cassel,3 J. E. Duboscq,3 K. M. Ecklund,3 R. Ehrlich,3 L. Fields,3 R. S. Galik,3 L. Gibbons,3 B. Gittelman,3 R. Gray,3 S. W. Gray,3 D. L. Hartill,3 B. K. Heltsley,3 D. Hertz,3 L. Hsu,3 C. D. Jones,3 J. Kandaswamy,3 D. L. Kreinick,3 V. E. Kuznetsov,3 H. Mahlke-Krüger,3 T. O. Meyer,3 P. U. E. Onyisi,3 J. R. Patterson,3 D. Peterson,3 J. Pivarski,3 D. Riley,3 J. L. Rosner,3 A. Ryd,3 A. J. Sadoff,3 H. Schwarthoff,3 M. R. Shepherd,3 W. M. Sun,3 J. G. Thayer,3 D. Urner,3 T. Wilksen,3 M. Weinberger,3 S. B. Athar,4 P. Avery,4 L. Breva-Newell,4 R. Patel,4 V. Potlia,4 H. Stoeck,4 J. Yelton,4 P. Rubin,5 B. I. Eisenstein,6 G. D. Gollin,6 I. Karliner,6 D. Kim,6 N. Lowrey,6 P. Naik,6 C. Sedlack,6 M. Selen,6 J. J. Thaler,6 J. Williams,6 J. Wiss,6 K. W. Edwards,7 D. Besson,8 K. Y. Gao,9 D. T. Gong,9 Y. Kubota,9 S. Z. Li,9 R. Poling,9 A. W. Scott,9 A. Smith,9 C. J. Stepaniak,9 J. Urheim,9 Z. Metreveli,10 K. K. Seth,10 A. Tomaradze,10 P. Zwebner,10 J. Ernst,11 K. Arts,12 K. K. Gan,12 H. Severini,13 P. Skubic,13 D. M. Asner,14 S. A. Dytman,14 S. Mehrabian,14 J. A. Mueller,14 V. Savinov,14 Z. Li,15 A. Lopez,15 H. Mendez,15 J. Ramirez,15 G. S. Huang,16 D. H. Miller,16 V. Pavlunin,16 B. Sanghi,16 E. I. Shibata,16 I. P. J. Shipsey,16 G. S. Adams,17 M. Chasse,17 J. P. Cummings,17 I. Danko,17 J. Napolitano,17 D. Cronin-Hennessy,18 C. S. Park,18 W. Park,18 J. B. Thayer,18 E. H. Thorndike,18 T. E. Coan,19 Y. S. Gao,19 F. Liu,19 R. Stroynowski,19 M. Artuso,20 C. Boulahouache,20 S. Blusk,20 J. Butt,20 E. Dambasuren,20 O. Dorjkhaidav,20 N. Menaa,20 R. Mountain,20 H. Muramatsu,20 R. Nandakumar,20 R. Redjimi,20 R. Sia,20 T. Skwarnicki,20 S. Stone,20 J. C. Wang,20 K. Zhang,20 A. H. Mahmood,21 S. E. Csorna,22 G. Bonvicini,23 D. Cinabro,23 and M. Dubrovin23

(CLEO Collaboration)

1California Institute of Technology, Pasadena, California 91125
2Carnegie Mellon University, Pittsburgh, Pennsylvania 15213
3Cornell University, Ithaca, New York 14853
4University of Florida, Gainesville, Florida 32611
5George Mason University, Fairfax, Virginia 22030
6University of Illinois, Urbana-Champaign, Illinois 61801
7Carleton University, Ottawa, Ontario, Canada K1S 5B6 and the Institute of Particle Physics, Canada
8University of Kansas, Lawrence, Kansas 66045
9University of Minnesota, Minneapolis, Minnesota 55455
10Northwestern University, Evanston, Illinois 60208
11State University of New York at Albany, Albany, New York 12222
12Ohio State University, Columbus, Ohio 43210
13University of Oklahoma, Norman, Oklahoma 73019
14University of Pittsburgh, Pittsburgh, Pennsylvania 15260
15University of Puerto Rico, Mayaguez, Puerto Rico 00681
16Purdue University, West Lafayette, Indiana 47907
Abstract

We have searched a sample of 9.6 million $B\bar{B}$ events for the lepton-flavor-violating leptonic $B$ decays, $B^0 \rightarrow \mu^{\pm}\tau^{\mp}$ and $B^0 \rightarrow e^{\pm}\tau^{\mp}$. The $\tau$-lepton was detected through the decay modes $\tau \rightarrow \ell \nu \bar{\nu}$, where $\ell = e, \mu$. There is no indication of a signal, and we obtain the 90% confidence level upper limits $\mathcal{B}(B^0 \rightarrow \mu^{\pm}\tau^{\mp}) < 3.8 \times 10^{-5}$ and $\mathcal{B}(B^0 \rightarrow e^{\pm}\tau^{\mp}) < 1.1 \times 10^{-4}$.

*On leave of absence from University of Chicago.
We report results of a search for two lepton-flavor-violating leptonic decays of $B$ mesons: $B^0 \to \mu^\pm \tau^\mp$ and $B^0 \to e^\pm \tau^\mp$. These modes are forbidden in the conventional Standard Model by the lepton-flavor conservation law. However, they are predicted to occur in many theories “beyond the standard model,” for example multi-Higgs-boson extensions, theories with leptoquarks, supersymmetric models without R-parity, and Higgs-mediated decay in supersymmetric seesaw models. The recent discovery of neutrino oscillation, while not leading to predictions of observable rates for lepton-flavor-violating decays, nonetheless heightens interest in them. The decays we searched for involve both third generation quarks and third generation leptons. Decays of this variety have been less extensively searched for than those involving only first or second generation quarks or leptons. Discovery of such decays at levels of our sensitivity would be clear evidence of physics beyond the Standard Model. Currently the best limits on the branching fractions are at levels of our sensitivity would be clear evidence of physics beyond the Standard Model.

In this analysis, we search for the decays $B^0 \to \mu^\pm \tau^\mp$ and $B^0 \to e^\pm \tau^\mp$, which are identified by shower energy to momentum ratio ($E/P$) and shower shape. We require that the secondary lepton, from the signal $\nu\bar{\nu}$ pair, be greater than 0.6 (1.0) GeV/c for $\mu$ ($e$). We “measure” the 4-momentum of the neutrino pair as the missing visible 4-momentum in the event: $E_{\nu\bar{\nu}} = 2E_{\text{beam}} - \sum E_i$, $\vec{P}_{\nu\bar{\nu}} = -\Sigma \vec{P}_i$, where sums are over all observed (charged and neutral) particles.

We define two neural net variables. $NN_{BB\nu}$ is a neural net variable used to suppress backgrounds from $BB\nu$ decays. We calculate three inputs: beam-constrained mass $\sqrt{E_{\text{beam}}^2 - \vec{P}_{\text{cont}}^2}$, $\Delta E$ (the energy of the $B$ candidate, and $\cos \theta_{tB}$ (the cosine of the angle between the momenta of primary lepton and $B$ candidate). We feed them into a neural net and train it with signal and $BB\nu$ Monte Carlo simulations for each mode. $NN_{cont}$ is a neural net variable to suppress backgrounds from continuum. We calculate five inputs: $R_0$ (the ratio of second and zeroth Fox-Wolfram moments of the event), $S$ (the sphericity), thrust of the event, $\cos \theta_{tB}$ (the cosine of the angle between the $\vec{p}_L - \vec{p}_\nu$ and the thrust axis of the rest of the event), and $\cos \theta_{\nu\bar{\nu}, \vec{p}_{L'} + \vec{p}_{\nu\bar{\nu}}}$ (the cosine of the angle between the momenta of neutrino pair and lepton pair), then feed
them into a neural net and train it with signal and continuum Monte Carlo for each mode. The nominal neural net range is from 0.0 to 1.0. We cut in the 2D space defined by $NN_{BB}$ and $NN_{cont}$, requiring $NN_{BB} > NN_{BB}^{cut}$, $NN_{cont} > NN_{cont}^{cut}$ and also

$$\frac{(NN_{BB} - NN_{BB}^{cut})}{(1 - NN_{BB}^{cut})} + \frac{(NN_{cont} - 1)}{(1 - NN_{cont})} > 0.$$ 

We define two $\tau$-mass variables. The first is the conventionally defined invariant mass of the reconstructed $\tau$, $M_{\ell\bar{\nu}} \equiv \sqrt{(E_\ell + E_{\bar{\nu}})^2 - (\vec{p}_\ell + \vec{p}_{\bar{\nu}})^2}$. The second $\tau$-mass variable makes use of the fact that, with perfect measurements of all quantities, $\Delta E = 0$, and hence we can use $E_{beam} - E_\ell - E_{\bar{\nu}}$ for $E_{\ell\bar{\nu}}$, yielding $M_{\ell\bar{\nu},\Delta E=0} \equiv \sqrt{(E_{beam} - E_\ell)^2 - (\vec{p}_\ell + \vec{p}_{\bar{\nu}})^2}$. We further define $\Delta M_\tau \equiv M_{\ell\bar{\nu}} - M_\tau$ and $\Delta M_{\ell\bar{\nu},\Delta E=0} \equiv M_{\ell\bar{\nu},\Delta E=0} - M_\tau$, where $M_\tau$ is the nominal $\tau$ mass, 1777 MeV.

By examining the angular distribution of electrons, positrons, and missing momentum, in off-resonance data in a $|\vec{p}_\ell|$ sideband region ($2.0 < |\vec{p}_\ell| < 2.2$ GeV/c and $2.5 < |\vec{p}_\ell| < 2.7$ GeV/c), we see clear evidence of the two-photon-fusion process. There are sharp peaks in the forward directions for electrons and positrons (“forward” being the direction of the beam particle of the same charge). Also, the missing momentum peaks sharply in the opposite direction from a detected $e^+$ or $e^-$, indicating an $e^-$ or $e^+$ lost down the beam pipe. By eliminating events with $|\cos\theta_{miss}| > 0.90$ (0.95 for ($\mu$, $\mu$)), we considerably reduce this background.

We compare Monte Carlo samples with data using the $|\vec{p}_\ell|$ sideband region defined above. In Fig. 1 we show distributions in $NN_{cont}$, $NN_{BB}$, $\Delta M_\tau$, and $\Delta M_{\ell\bar{\nu},\Delta E=0}$, for off-resonance-subtracted on-resonance data and absolutely normalized $BB$ Monte Carlo. Agreement is good. In Fig. 2 we show distributions for the same variables for off-resonance data and absolutely normalized continuum ($e^+e^- \rightarrow q\bar{q}$, $q = u,d,s,c$) Monte Carlo. We have not included a Monte Carlo for the inclusive multi-hadronic two-photon-fusion process, lacking a trustworthy simulation of this process. For ($\mu$, $e$), shown in Fig. 2 agreement is good, indicating that the remaining contribution from two-photon fusion is small. For ($e$, $e$), not shown, data exceeds continuum Monte Carlo, indicating a sizeable remaining contribution from two-photon fusion.

We measure the ratio of data to Monte Carlo yields in the $|\vec{p}_\ell|$ sideband region, denoting by $R_{BB}$ the ratio of off-resonance-subtracted on-resonance data to $BB$ Monte Carlo, and by $R_{cont}$ the ratio of off-resonance data to continuum Monte Carlo. $R$’s are measured with loose selection criteria applied: $NN_{cont} > 0.5$, $NN_{BB} > 0.5$, $|\cos\theta_{miss}| < 0.9$ (0.95 for ($\mu$, $\mu$)), and $|\Delta M_\tau| < 2.0$ GeV. Values so obtained are given in Table I. One sees that $R_{BB}$ differs little from 1.0, while $R_{cont}$ is less well behaved, particularly for the ($e$, $e$) mode, indicating that the two-photon-fusion background is present. To estimate backgrounds in the signal $|\vec{p}_\ell|$ region correctly, we scale the signal-region yields from Monte Carlo by $R$. Because we perform a direct subtraction of off-resonance data, the accuracy of the continuum background prediction is not critical for our results.

We optimized our selection criteria on $NN_{BB}$, $NN_{cont}$, $\Delta M_\tau$, and $\Delta M_{\ell\bar{\nu},\Delta E=0}$ to obtain the best upper limit when the true branching fraction is zero. This optimization procedure made use of signal and background Monte Carlo samples, and scaled the background samples by $R_{BB}$ or $R_{cont}$ as described above. The optimized selection criteria, found separately for each mode, are shown in Table I.

The number of events satisfying all selection criteria is shown, for each mode, in Table I along with the background estimate. We find 15 ($\mu$, $e$) candidates, with 23.7 expected from
background; we find 4 ($\mu$, $\mu$) candidates, with 9.0 expected from background; we find 14 ($e$, $e$) candidates, with 11.6 expected from background; we find 6 ($e$, $\mu$) candidates, with 5.1 expected from background. Thus there are a total of 39 events with 49.4 expected from background. The probability that a true mean of 49.4 will give rise to a yield of 39 or more events is 93%. With no indication of signal, we obtain the branching fraction upper limits.

We calculate upper limits at 90% confidence level. There is some probability of observing the off-resonance-subtracted on-resonance yield that we do observe, or less, if the branching fractions for $B^0 \to \mu^+\tau^-$ and $B^0 \to e^+\tau^-$ are zero. We take the 90% confidence level.
TABLE I: The rows of $NN_{BB}^{cut}$, $NN_{cont}^{cut}$, $\Delta M_\tau$ (GeV), and $\Delta M_{\tau,\Delta E=0}$ (GeV) indicate the optimized selection criteria for each mode. $R_{BB}$ and $R_{cont}$ are measured ratios between data yields and Monte Carlo with loose selection criteria in the $|\vec{p}_\ell|$ sideband region. $N_{ON}(N_{OFF})$ is the number of observed events satisfying the optimized selection criteria in the signal region of $|\vec{p}_\ell|$ from on(off)-resonance data samples. $N_{obs}$ is the number of observed events from off-resonance-subtracted on-resonance data, $N_{ON}=1.99N_{OFF}$. $N_{BB}$ is the $B\bar{B}$ background estimate from Monte Carlo, scaled by $R_{BB}$; in the absence of signal it should be comparable with $N_{obs}$. $0.5N_{cont}$ is $(1/1.99)$ times the continuum background estimate from Monte Carlo, scaled by $R_{cont}$, which should be comparable with $N_{OFF}$. $\epsilon$ is the signal detection efficiency including $\tau$ decay branching fraction. $BR UL$ is the branching ratio upper limit at 90% confidence level with systematic error considered.

| $(\ell, \ell')$ | $(\mu, e)$ | $(\mu, \mu)$ | $(e, e)$ | $(e, \mu)$ |
|----------------|-----------|-------------|---------|-----------|
| $NN_{BB}^{cut}$ | 0.725 | 0.875 | 0.675 | 0.825 |
| $NN_{cont}^{cut}$ | 0.700 | 0.775 | 0.700 | 0.475 |
| $-2.0<\Delta M_\tau$ | $<2.00$ | $<1.40$ | $<1.50$ | $<1.40$ |
| $-2.0<\Delta M_{\tau,\Delta E=0}$ | $<0.25$ | $<0.25$ | $<0.30$ | $<0.25$ |

$R_{BB}$ | 1.21±0.06 | 1.06±0.07 | 1.04±0.07 | 0.94±0.07 |
$R_{cont}$ | 1.03±0.27 | 1.52±0.32 | 4.30±0.95 | 0.64±0.38 |

$N_{ON}$ | 19 | 10 | 28 | 6 |
$N_{OFF}$ | 2 | 3 | 7 | 0 |
$N_{obs}$ | 15.0±5.2 | 4.0±4.7 | 14.0±7.5 | 6.0±2.4 |

$N_{BB}$ | 23.7±2.7 | 9.0±1.4 | 11.6±1.4 | 5.1±0.8 |
$0.5N_{cont}$ | 1.8±0.6 | 0.4±0.2 | 3.1±1.0 | 0.5±0.3 |
$\epsilon$ (%) | 1.57 | 0.63 | 0.96 | 0.58 |
$BR UL(10^{-4})$ | 0.55 | 0.87 | 1.64 | 1.46 |

upper limit to be that value of the branching fraction which reduces the above-mentioned probability by a factor of 10. The ingredients needed for the calculation are: (1) the observed off-resonance-subtracted on-resonance yield; (2) the true mean for the background contribution from $B\bar{B}$ processes, and (3) less critically, the true mean of the background contribution from non-resonance processes. To allow for the uncertainty in the background estimates, we changed $R_{BB}$ and $R_{cont}$ in the unfavorable directions by $1\sigma$, i.e. $-1\sigma$ for $R_{BB}$ and $+1\sigma$ for $R_{cont}$.

We use Monte Carlo simulation to determine the efficiency for detecting the signal modes. The decays $B \rightarrow \mu(e)\pm\tau^\mp$ are generated with the $\tau$ lepton unpolarized. For a $\tau$ lepton polarization as given by $V - A$, the secondary lepton is boosted (has its average lab energy increased), which in turn increases the efficiency. For the opposite polarization, as given by $V + A$, the secondary lepton is deboosted, and the efficiency is lowered. The fractional
changes in efficiency, averaged over the four modes, are +11% for $V - A$, $-8\%$ for $V + A$. Our upper limits are quoted for unpolarized $\tau$’s.

Systematic errors are of two varieties – those on the estimate of signal detection efficiencies, and those on the estimate of backgrounds. The dominant contributors to the former are lepton identification efficiency uncertainties (contributing $\pm 3.5\%$ per lepton, relative, in the efficiency), and missing-four-vector-simulation uncertainties ($\pm 5.4\%$), giving a relative uncertainty in the overall efficiency of $\pm 7.4\%$ for $(e, \mu)$ and $(\mu, e)$, and $\pm 8.9\%$ for $(e, e)$ and $(\mu, \mu)$. The background uncertainties are handled by varying the $R_{BB}$ and $R_{cont}$ as mentioned above. The errors shown on the backgrounds in Table I include statistical and systematic errors.

There is no universally agreed-upon procedure for including systematic errors in upper-limit estimates. We conservatively vary the background by 1.0 standard deviations, and decrease the efficiency by 1.0 standard deviations for each mode and the results are as shown in Table II.

To combine the results from two leptonic modes, $\tau \to e\nu\bar{\nu}$ and $\tau \to \mu\nu\bar{\nu}$, we simply add the yields, add the backgrounds, and add the efficiencies.

In this way we obtain our final results:

$$B(B \to \mu^\pm \tau^\mp) < 3.8 \times 10^{-5},$$
$$B(B \to e^\pm \tau^\mp) < 1.1 \times 10^{-4},$$

both at 90% confidence level. These results are significant improvements over previously published limits [3].

In summary, we have searched for the decays $B \to \mu^\pm \tau^\mp$ and $B \to e^\pm \tau^\mp$. We find no indication of a signal, and obtain upper limits on the branching fractions.

We gratefully acknowledge the effort of the CESR staff in providing us with excellent luminosity and running conditions. This work was supported by the National Science Foundation, the U.S. Department of Energy, the Research Corporation, and the Texas Advanced Research Program.

[1] M. Sher and Y. Yuan, Phys. Rev. D 44, 1461 (1991); J. Pati and A. Salam, Phys. Rev. D 10, 275 (1974); S. Davidson, D. Bailey, and B. A. Campbell, Z. Phys. C 61, 613 (1994); D. Guetta, J. M. Mira, and E. Nardi, Phys. Rev. D 59, 034019 (1999); A. Dedes, J. Ellis and M. Raidal, Phys. Lett. B549 159 (2002).
[2] Y. Fukuda et al. (Super-Kamiokande Collaboration), Phys. Rev. Lett. 85 3999 (2000); Q.R. Ahmed et al. (SNO Collaboration), Phys. Rev. Lett. 89 011301 (2002); K. Eguchi et al. (KamLAND Collaboration), Phys. Rev. Lett. 90 041801 (2003).
[3] R. Ammar et al. (CLEO Collaboration), Phys. Rev. D 49, 5701 (1994).
[4] Y. Kubota et al. (CLEO Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A 320, 66 (1992); T. Hill, Nucl. Instrum. Methods Phys. Res., Sect. A 418, 32 (1998).
[5] G. Fox and S. Wolfram, Phys. Rev. Lett. 41, 1581 (1978).