BABAR-PUB-11/023

SLAC-PUB-14951

A search for the decay modes $B^{\pm} \to h^{\pm}\tau\ell$

J. P. Lees, V. Poireau, and V. Tisserand
Laboratoire d'Annecy-le-Vieux de Physique des Particules (LAPP), Université de Savoie, CNRS/IN2P3, F-74941 Annecy-Le-Vieux, France

J. Garra Tico and E. Grauges
Universitat de Barcelona, Facultat de Fisica, Departament ECM, E-08028 Barcelona, Spain

D. A. Milanes[a], A. Palano[ab], and M. Pappagallo[ab]
INFN Sezione di Bari[a]; Dipartimento di Fisica, Università di Bari[b], I-70126 Bari, Italy

G. Eigen and B. Stugu
University of Bergen, Institute of Physics, N-5007 Bergen, Norway

D. N. Brown, L. T. Kerth, Yu. G. Kolomensky, and G. Lynch
Lawrence Berkeley National Laboratory and University of California, Berkeley, California 94720, USA

H. Koch and T. Schroeder
Ruhr Universität Bochum, Institut für Experimentalphysik 1, D-44780 Bochum, Germany

D. J. Asgeirsson, C. Hearty, T. S. Mattison, and J. A. McKenna
University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z1

A. Khan
Brunel University, Uxbridge, Middlesex UB8 3PH, United Kingdom

V. E. Blinov, A. R. Buzykaev, V. P. Druzhinin, V. B. Golubev, E. A. Kravchenko, A. P. Onuchin, S. I. Serednyakov, Yu. I. Skovpen, E. P. Solodov, K. Yu. Todyshev, and A. N. Yushkov
Budker Institute of Nuclear Physics, Novosibirsk 630090, Russia

M. Bondioli, D. Kirkby, A. J. Lankford, M. Mandelkern, and D. P. Stoker
University of California at Irvine, Irvine, California 92697, USA

H. Atmacan, J. W. Gary, F. Liu, O. Long, and G. M. Vitug
University of California at Riverside, Riverside, California 92521, USA

C. Campagnari, T. M. Hong, D. Kovalskyi, J. D. Richman, and C. A. West
University of California at Santa Barbara, Santa Barbara, California 93106, USA

A. M. Eisner, J. Kroseberg, W. S. Lockman, A. J. Martinez, T. Schalk, B. A. Schumm, and A. Seiden
University of California at Santa Cruz, Institute for Particle Physics, Santa Cruz, California 95064, USA

C. H. Cheng, D. A. Doll, B. Echenard, K. T. Flood, D. G. Hitlin, P. Ongmongkolkul, F. C. Porter, and A. Y. Rakitin
California Institute of Technology, Pasadena, California 91125, USA

R. Andreassen, Z. Huard, B. T. Meadows, M. D. Sokoloff, and L. Sun
University of Cincinnati, Cincinnati, Ohio 45221, USA

P. C. Bloom, W. T. Ford, A. Gaz, M. Nagel, U. Nauenberg, J. G. Smith, and S. R. Wagner
University of Colorado, Boulder, Colorado 80309, USA

R. Ayad* and W. H. Toki
Colorado State University, Fort Collins, Colorado 80523, USA

Submitted to Physical Review D

Work supported by US Department of Energy contract DE-AC02-76SF00515.

SLAC National Accelerator Laboratory, Menlo Park, CA 94025

B. Spaan
Technische Universität Dortmund, Fakultät Physik, D-44221 Dortmund, Germany

M. J. Kobel, K. R. Schubert, and R. Schwierz
Technische Universität Dresden, Institut für Kern- und Teilchenphysik, D-01062 Dresden, Germany

D. Bernard and M. Verderi
Laboratoire Leprince-Ringuet, Ecole Polytechnique, CNRS/IN2P3, F-91128 Palaiseau, France

P. J. Clark and S. Playfer
University of Edinburgh, Edinburgh EH9 3JZ, United Kingdom

D. Bettoni[a], C. Bozzi[a], R. Calabrese[ab], G. Cibinetto[ab], E. Fioravanti[ab], I. Garzia[ab], E. Luppi[ab], M. Munerato[ab], M. Negrini[ab], L. Piemontese[a], and V. Santoro
INFN Sezione di Ferrara[a]; Dipartimento di Fisica, Università di Ferrara[b], I-44100 Ferrara, Italy

R. Baldini-Ferroli, A. Calcaterra, R. de Sangro, G. Finocchiaro, P. Patteri, I. M. Peruzzi,[†] M. Piccolo, M. Rama, and A. Zallo
INFN Laboratori Nazionali di Frascati, I-00044 Frascati, Italy

R. Contri[ab], E. Guido[ab], M. Lo Vetere[ab], M. R. Monge[ab], S. Passaggio[a], C. Patrignani[ab], and E. Robutti[a]
INFN Sezione di Genova[a]; Dipartimento di Fisica, Università di Genova[b], I-16146 Genova, Italy

B. Bhuyan and V. Prasad
Indian Institute of Technology Guwahati, Guwahati, Assam, 781 039, India

C. L. Lee and M. Morii
Harvard University, Cambridge, Massachusetts 02138, USA

A. J. Edwards
Harvey Mudd College, Claremont, California 91711

A. Adametz, J. Marks, and U. Uwer
Universität Heidelberg, Physikalisches Institut, Philosophenweg 12, D-69120 Heidelberg, Germany

H. M. Lacker and T. Lueck
Humboldt-Universität zu Berlin, Institut für Physik, Newtonstr. 15, D-12489 Berlin, Germany

P. D. Dauncey
Imperial College London, London, SW7 2AZ, United Kingdom

P. K. Behera and U. Mallik
University of Iowa, Iowa City, Iowa 52242, USA

C. Chen, J. Cochran, W. T. Meyer, S. Prell, and A. E. Rubin
Iowa State University, Ames, Iowa 50011-3160, USA

A. V. Gritsan and Z. J. Guo
Johns Hopkins University, Baltimore, Maryland 21218, USA

N. Arnaud, M. Davier, D. Derkach, G. Grosdidier, F. Le Diberder, A. M. Lutz, B. Malaescu, P. Roudeau, M. H. Schune, A. Stocchi, and G. Wormser
Laboratoire de l'Accélérateur Linéaire, IN2P3/CNRS et Université Paris-Sud 11, Centre Scientifique d'Orsay, B. P. 34, F-91898 Orsay Cedex, France

D. J. Lange and D. M. Wright
Lawrence Livermore National Laboratory, Livermore, California 94550, USA

I. Bingham, C. A. Chavez, J. P. Coleman, J. R. Fry, E. Gabathuler, D. E. Hutchcroft, D. J. Payne, and C. Touramanis
University of Liverpool, Liverpool L69 7ZE, United Kingdom

A. J. Bevan, F. Di Lodovico, R. Sacco, and M. Sigamani
Queen Mary, University of London, London, E1 4NS, United Kingdom

G. Cowan
University of London, Royal Holloway and Bedford New College, Egham, Surrey TW20 0EX, United Kingdom

D. N. Brown and C. L. Davis
University of Louisville, Louisville, Kentucky 40292, USA

A. G. Denig, M. Fritsch, W. Gradl, A. Hafner, and E. Prencipe
Johannes Gutenberg-Universität Mainz, Institut für Kernphysik, D-55099 Mainz, Germany

K. E. Alwyn, D. Bailey, R. J. Barlow,[‡] G. Jackson, and G. D. Lafferty
University of Manchester, Manchester M13 9PL, United Kingdom

E. Behn, R. Cenci, B. Hamilton, A. Jawahery, D. A. Roberts, and G. Simi
University of Maryland, College Park, Maryland 20742, USA

C. Dallapiccola
University of Massachusetts, Amherst, Massachusetts 01003, USA

R. Cowan, D. Dujmic, and G. Sciolla
Massachusetts Institute of Technology, Laboratory for Nuclear Science, Cambridge, Massachusetts 02139, USA

D. Lindemann, P. M. Patel, S. H. Robertson, and M. Schram
McGill University, Montréal, Québec, Canada H3A 2T8

P. Biassoni[ab], N. Neri[a], F. Palombo[ab], and S. Stracka[ab]
INFN Sezione di Milano[a]; Dipartimento di Fisica, Università di Milano[b], I-20133 Milano, Italy

L. Cremaldi, R. Godang,[§] R. Kroeger, P. Sonnek, and D. J. Summers
University of Mississippi, University, Mississippi 38677, USA

X. Nguyen, M. Simard, and P. Taras
Université de Montréal, Physique des Particules, Montréal, Québec, Canada H3C 3J7

G. De Nardo[ab], D. Monorchio[ab], G. Onorato[ab], and C. Sciacca[ab]
INFN Sezione di Napoli[a]; Dipartimento di Scienze Fisiche, Università di Napoli Federico II[b], I-80126 Napoli, Italy

M. Martinelli and G. Raven
NIKHEF, National Institute for Nuclear Physics and High Energy Physics, NL-1009 DB Amsterdam, The Netherlands

C. P. Jessop, K. J. Knoepfel, J. M. LoSecco, and W. F. Wang
University of Notre Dame, Notre Dame, Indiana 46556, USA

K. Honscheid and R. Kass
Ohio State University, Columbus, Ohio 43210, USA

J. Brau, R. Frey, N. B. Sinev, D. Strom, and E. Torrence
University of Oregon, Eugene, Oregon 97403, USA

E. Feltresi[ab], N. Gagliardi[ab], M. Margoni[ab], M. Morandin[a],
M. Posocco[a], M. Rotondo[a], F. Simonetto[ab], and R. Stroili[ab]
INFN Sezione di Padova[a]; Dipartimento di Fisica, Università di Padova[b], I-35131 Padova, Italy

S. Akar, E. Ben-Haim, M. Bomben, G. R. Bonneaud, H. Briand, G. Calderini,
J. Chauveau, O. Hamon, Ph. Leruste, G. Marchiori, J. Ocariz, and S. Sitt
Laboratoire de Physique Nucléaire et de Hautes Energies,
IN2P3/CNRS, Université Pierre et Marie Curie-Paris6,
Université Denis Diderot-Paris7, F-75252 Paris, France

M. Biasini[ab], E. Manoni[ab], S. Pacetti[ab], and A. Rossi[ab]
INFN Sezione di Perugia[a]; Dipartimento di Fisica, Università di Perugia[b], I-06100 Perugia, Italy

C. Angelini[ab], G. Batignani[ab], S. Bettarini[ab], M. Carpinelli[ab],¶ G. Casarosa[ab], A. Cervelli[ab], F. Forti[ab],
M. A. Giorgi[ab], A. Lusiani[ac], B. Oberhof[ab], E. Paoloni[ab], A. Perez[a], G. Rizzo[ab], and J. J. Walsh[a]
INFN Sezione di Pisa[a]; Dipartimento di Fisica, Università di Pisa[b]; Scuola Normale Superiore di Pisa[c], I-56127 Pisa, Italy

D. Lopes Pegna, C. Lu, J. Olsen, A. J. S. Smith, and A. V. Telnov
Princeton University, Princeton, New Jersey 08544, USA

F. Anulli[a], G. Cavoto[a], R. Faccini[ab], F. Ferrarotto[a], F. Ferroni[ab],
M. Gaspero[ab], L. Li Gioi[a], M. A. Mazzoni[a], and G. Piredda[a]
INFN Sezione di Roma[a]; Dipartimento di Fisica,
Università di Roma La Sapienza[b], I-00185 Roma, Italy

C. Bünger, O. Grünberg, T. Hartmann, T. Leddig, H. Schröder, C. Voss, and R. Waldi
Universität Rostock, D-18051 Rostock, Germany

T. Adye, E. O. Olaiya, and F. F. Wilson
Rutherford Appleton Laboratory, Chilton, Didcot, Oxon, OX11 0QX, United Kingdom

S. Emery, G. Hamel de Monchenault, G. Vasseur, and Ch. Yèche
CEA, Irfu, SPP, Centre de Saclay, F-91191 Gif-sur-Yvette, France

D. Aston, D. J. Bard, R. Bartoldus, C. Cartaro, M. R. Convery, J. Dorfan, G. P. Dubois-Felsmann,
W. Dunwoodie, M. Ebert, R. C. Field, M. Franco Sevilla, B. G. Fulsom, A. M. Gabareen, M. T. Graham,
P. Grenier, C. Hast, W. R. Innes, M. H. Kelsey, P. Kim, M. L. Kocian, D. W. G. S. Leith, P. Lewis, B. Lindquist,
S. Luitz, V. Luth, H. L. Lynch, D. B. MacFarlane, D. R. Muller, H. Neal, S. Nelson, M. Perl, T. Pulliam,
B. N. Ratcliff, A. Roodman, A. A. Salnikov, R. H. Schindler, A. Snyder, D. Su, M. K. Sullivan, J. Va'vra,
A. P. Wagner, M. Weaver, W. J. Wisniewski, M. Wittgen, D. H. Wright, H. W. Wulsin, C. C. Young, and V. Ziegler
SLAC National Accelerator Laboratory, Stanford, California 94309 USA

W. Park, M. V. Purohit, R. M. White, and J. R. Wilson
University of South Carolina, Columbia, South Carolina 29208, USA

A. Randle-Conde and S. J. Sekula
Southern Methodist University, Dallas, Texas 75275, USA

M. Bellis, J. F. Benitez, P. R. Burchat, and T. S. Miyashita
Stanford University, Stanford, California 94305-4060, USA

M. S. Alam and J. A. Ernst
State University of New York, Albany, New York 12222, USA

R. Gorodeisky, N. Guttman, D. R. Peimer, and A. Soffer
Tel Aviv University, School of Physics and Astronomy, Tel Aviv, 69978, Israel

P. Lund and S. M. Spanier
University of Tennessee, Knoxville, Tennessee 37996, USA

R. Eckmann, J. L. Ritchie, A. M. Ruland, C. J. Schilling, R. F. Schwitters, and B. C. Wray
University of Texas at Austin, Austin, Texas 78712, USA

J. M. Izen and X. C. Lou
University of Texas at Dallas, Richardson, Texas 75083, USA

F. Bianchi[ab] and D. Gamba[ab]
INFN Sezione di Torino[a]; Dipartimento di Fisica Sperimentale, Università di Torino[b], I-10125 Torino, Italy

L. Lanceri[ab] and L. Vitale[ab]
INFN Sezione di Trieste[a]; Dipartimento di Fisica, Università di Trieste[b], I-34127 Trieste, Italy

F. Martinez-Vidal and A. Oyanguren
IFIC, Universitat de Valencia-CSIC, E-46071 Valencia, Spain

H. Ahmed, J. Albert, Sw. Banerjee, F. U. Bernlochner, H. H. F. Choi, G. J. King, R. Kowalewski, M. J. Lewczuk, I. M. Nugent, J. M. Roney, R. J. Sobie, and N. Tasneem
University of Victoria, Victoria, British Columbia, Canada V8W 3P6

T. J. Gershon, P. F. Harrison, T. E. Latham, and E. M. T. Puccio
Department of Physics, University of Warwick, Coventry CV4 7AL, United Kingdom

H. R. Band, S. Dasu, Y. Pan, R. Prepost, and S. L. Wu
University of Wisconsin, Madison, Wisconsin 53706, USA

We present a search for the lepton flavor violating decay modes $B^{\pm} \to h^{\pm}\tau\ell$ ($h = K, \pi$; $\ell = e, \mu$) using the *BABAR* data sample, which corresponds to 472 million $B\overline{B}$ pairs. The search uses events where one B meson is fully reconstructed in one of several hadronic final states. Using the momenta of the reconstructed B, h, and ℓ candidates, we are able to fully determine the τ four-momentum. The resulting τ candidate mass is our main discriminant against combinatorial background. We see no evidence for $B^{\pm} \to h^{\pm}\tau\ell$ decays and set a 90% confidence level upper limit on each branching fraction at the level of a few times 10^{-5}.

PACS numbers: 13.25.Hw, 14.40.Nd

I. INTRODUCTION

The standard model (SM) of electroweak interactions does not allow charged lepton flavor violation or flavor changing neutral currents (FCNC) in tree-level interactions [1]. Lepton flavor violating decays of B mesons can occur at the one-loop level through processes that involve neutrino mixing, but these are highly suppressed by powers of m_ν^2/m_W^2 [2] and have predicted branching fractions many orders of magnitude below the current experimental sensitivity. However, in many extensions of the SM, B decays involving lepton flavor violation and/or FCNC interactions are greatly enhanced [2–5]. In some cases, decays involving the second and third generations of quarks and leptons are particularly sensitive to physics beyond the SM [3].

Until recent years, experimental information on B decays to final states containing τ leptons has been weak or absent. The presence of at least one neutrino from the τ decay prevents direct reconstruction of the τ, making it difficult to distinguish $B \to X\tau$ decays from the abundant semileptonic $B \to X\ell\nu$; $\ell = e, \mu$ decays. The high-luminosity B factory experiments have developed the technique of using a fully-reconstructed hadronic B decay (the "tag" B) to determine the three-momentum of the other B (the "signal" B) in $\Upsilon(4S) \to B\overline{B}$ events, which enables the τ to be indirectly reconstructed. This technique assigns all detected tracks and neutral objects to either the tag B or the signal B. Recent applications of this technique by *BABAR* are the searches for $B^+ \to K^+\tau\mu$ [6], $B^0 \to \ell^{\pm}\tau^{\mp}$ [7] and $B^+ \to \tau^+\nu$ [8]. We present an update of our search for $B^+ \to K^+\tau\mu$ [6] and the first search for the decays $B^+ \to K^+\tau e$, $B^+ \to \pi^+\tau\mu$, and $B^+ \to \pi^+\tau e$ [9].

The signal branching fraction is determined by using the ratio of the number of $B \to h\tau\ell$ ($h = K^{\pm}, \pi^{\pm}$) signal candidates to the yield of control samples of $B^+ \to \overline{D}^{(*)0}\ell^+\nu$; $\overline{D}^0 \to K^+\pi^-$ events from a fully reconstructed hadronic $B^{\pm}$ decay sample. Continuum background is suppressed for each decay channel using a likelihood ratio based on event shape information, unassociated calorimeter clusters, and the quality of muon identification for channels that have a muon in the final state. Final signal candidates are selected requiring the indirectly recon-

*Now at the University of Tabuk, Tabuk 71491, Saudi Arabia
†Also with Università di Perugia, Dipartimento di Fisica, Perugia, Italy
‡Now at the University of Huddersfield, Huddersfield HD1 3DH, UK
§Now at University of South Alabama, Mobile, Alabama 36688, USA
¶Also with Università di Sassari, Sassari, Italy

structed τ mass to fall in a narrow window around the known τ mass. The yield and estimated background in the τ mass signal window are used to estimate and set upper limits on the signal branching fractions. We followed the principle of a blind analysis, to avoid experimenter's bias, by not revealing the number of events in the signal window until after all analysis procedures were decided.

II. DATA SAMPLE AND DETECTOR DESCRIPTION

We use a data sample of 472 million $B\overline{B}$ pairs in 429 fb^{-1} of integrated luminosity, delivered by the PEP-II asymmetric-energy e^+e^- collider and recorded by the *BABAR* experiment at the SLAC National Accelerator Laboratory. This corresponds to the entire $\Upsilon(4S)$ data sample.

The *BABAR* experiment is described in detail elsewhere [10]. Trajectories of charged particles are reconstructed by a double-sided, five-layer silicon vertex tracker (SVT) and a 40-layer drift chamber (DCH). The SVT provides precision measurements for vertex reconstruction and stand-alone tracking for very low momentum tracks, with transverse momentum less than 120 MeV/c. The tracking system is inside a 1.5 T superconducting solenoid. Both the SVT and the DCH provide specific ionization (dE/dx) measurements that are used in particle identification (PID). Just beyond the radius of the DCH lies an array of fused silica bars which are part of the detector of internally reflected Cherenkov radiation (DIRC). The DIRC provides excellent charged-hadron PID. A CsI(Tl) crystal electromagnetic calorimeter (EMC) is used to reconstruct photons and identify electrons. The minimum EMC cluster energy used in this analysis is 30 MeV. The iron of the flux return for the solenoid is instrumented (IFR) with resistive plate chambers and limited streamer tubes, which are used in the identification of muons.

Monte Carlo (MC) simulated samples for our $B \to h\tau\ell$ signals and for all relevant SM processes are generated with EvtGen [11]. We model the *BABAR* detector response using GEANT4 [12]. The $B \to h\tau\ell$ decays are generated using a uniform three-body phase space model and the background MC sample combines SM processes: $e^+e^- \to \Upsilon(4S) \to B\overline{B}$, $e^+e^- \to q\bar{q}$ ($q = u, d, s, c$), and $e^+e^- \to \tau^+\tau^-$. The number of simulated Monte Carlo events corresponds to integrated luminosities equivalent to three times the data for $B\overline{B}$ events and two times the data for the continuum processes. Each Monte Carlo sample is reweighted to correspond to an integrated luminosity equivalent to the data.

The data and MC samples in this analysis are processed and generated with consistent database conditions determined from the detector response and analyzed using *BABAR* analysis software release tools.

III. EVENT RECONSTRUCTION

In each event, we require a fully reconstructed hadronic $B^\pm$ decay, which we refer to as the "tag" B meson candidate or $B_{\rm tag}$. We then search for the signal $B \to h\tau\ell$ decay in the rest of the event, which we refer to as the "signal" B meson candidate or $B_{\rm sig}$. The notation $B \to h\tau\ell$ refers to one of the following eight final states that we consider, where the primary hadron h is a K or π and the primary lepton ℓ is a μ or e: $B^+ \to K^+\tau^-\mu^+$, $B^+ \to K^+\tau^+\mu^-$, $B^+ \to K^+\tau^-e^+$, $B^+ \to K^+\tau^+e^-$, $B^+ \to \pi^+\tau^-\mu^+$, $B^+ \to \pi^+\tau^+\mu^-$, $B^+ \to \pi^+\tau^-e^+$, and $B^+ \to \pi^+\tau^+e^-$. In all cases, we require that the τ decays to a "one-prong" final state ($\tau \to e\nu\overline{\nu}$, $\tau \to \mu\nu\overline{\nu}$, and $\tau \to (n\pi^0)\pi\nu$ with $n \geq 0$). The branching fraction for τ decays to a one-prong final state is 85%.

The $\Upsilon(4S) \to B^+B^-$ decay requires the $B_{\rm sig}$ three-momentum to be opposite from that of the $B_{\rm tag}$ ($-\vec{p}_{\rm tag}$) and the $B_{\rm sig}$ energy to be equal to the beam energy ($E_{\rm beam}$) in the e^+e^- center-of-mass (CM) reference frame [13]. These constraints allow us to reconstruct the τ indirectly using

$$\begin{aligned} \vec{p}_\tau &= -\vec{p}_{\rm tag} - \vec{p}_h - \vec{p}_\ell, \\ E_\tau &= E_{\rm beam} - E_h - E_\ell, \\ m_\tau &= \sqrt{E_\tau^2 - |\vec{p}_\tau|^2}, \end{aligned}$$

where (E_τ, $\vec{p}_\tau$), (E_h, $\vec{p}_h$), and (E_ℓ, $\vec{p}_\ell$) are the corresponding four-momenta of the reconstructed signal objects. The indirectly reconstructed τ mass (m_τ) peaks sharply at the true τ mass in $B \to h\tau\ell$ signal events and has a very broad distribution for combinatorial background events. To avoid experimental bias, we did not look at events in the data with m_τ within ± 175 MeV/c^2 of the nominal τ mass until all analysis procedures were established.

A. Tag B reconstruction

The $B_{\rm tag}$ is fully reconstructed in one of many final states [14] of the form $B^- \to D^{(*)0}X^-$. The notation $D^{(*)0}$ refers to either a D^0 or a D^{*0} which decays to either $D^0\gamma$ or $D^0\pi^0$. The D^0 is reconstructed in the $K^-\pi^+$, $K^-\pi^+\pi^-\pi^+$, $K^-\pi^+\pi^0$, and $K^0_S\pi^+\pi^-$ channels, with $K^0_S \to \pi^+\pi^-$ and $\pi^0 \to \gamma\gamma$. The X^- represents a system of charged and neutral hadrons composed of $n_1\pi^\pm$, $n_2K^\pm$, $n_3K^0_S$, and $n_4\pi^0$; subject to the constraints $n_1 + n_2 \leq 5$, $n_3 \leq 2$, $n_4 \leq 2$, and total charge -1.

Each distinct $B_{\rm tag}$ decay mode has an associated *a priori* purity, defined as the number of peaking events divided by the number of peaking plus combinatorial events, where peaking and combinatorial yields are obtained from fits to $m_{\rm ES} \equiv \sqrt{E^2_{\rm beam} - |\vec{p}_{\rm tag}|^2}$ distributions for each distinct $B_{\rm tag}$ decay mode. We only consider $B_{\rm tag}$ decay modes with a purity greater than 10% and choose the $B_{\rm tag}$ candidate with the highest purity in the event. If there is more than one $B_{\rm tag}$ candidate with the same

purity, we choose the one with reconstructed energy closest to the beam energy. The $B_{\rm tag}$ candidate must have $m_{\rm ES} > 5.27\,{\rm GeV}/c^2$ and $E_{\rm tag}$ within three standard deviations of $E_{\rm beam}$. A charged $B_{\rm tag}$ candidate is properly reconstructed in approximately 0.25% of all $B\overline{B}$ events.

B. Particle identification

PID algorithms are used to identify kaons, pions, protons, muons, and electrons. We use an error correcting output code (ECOC) algorithm [15] with 36 input variables to identify electrons, pions, and protons. The ECOC combines multiple bootstrap aggregated decision tree binary classifiers trained to separate e, π, K, and p. The most important inputs for electron identification are the EMC energy divided by the track momentum, several EMC shower shape variables, and the deviation from the expected values divided by the measurement uncertainties of the Cherenkov angle and of the dE/dx for the e, π, K, and p hypotheses. Neutral clusters in the EMC that are consistent with bremsstrahlung radiation are used to correct the momentum and energy of electron candidates. A γ candidate from an $e^{\pm}$ track is consistent with bremsstrahlung radiation if the corresponding three-momenta are within $|\Delta\theta| < 35\,{\rm mrad}$ and $|\Delta\phi| < 50\,{\rm mrad}$, with respect to the polar and azimuthal angles of the beam axis.

Muons and kaons are identified using a bagging decision trees [16] algorithm with 30 (36) input variables for the muon (kaon) selection. For muons, the most important input variables are the number and position of the hits in the IFR, the difference between the expected and measured DCH dE/dx for the muon hypothesis, and the energy deposited in the EMC. For kaons, the most important variables are the kaon and pion likelihoods based on the measured Cherenkov angle in the DIRC and the difference between the expected and measured dE/dx for the kaon hypothesis.

We define several quality levels of particle identification for use in the analysis. The "loose" levels have higher efficiency but also higher misidentification probabilities. The "tight" levels have lower misidentification probabilities and efficiencies. Table I summarizes the selection efficiency and misidentification probabilities of the PID selection algorithms used. A "very loose" (VL) K-PID algorithm is used for identifying the primary K in $B \to K\tau\ell$, while a "very tight" (VT) K-PID algorithm, with lower efficiency but much smaller misidentification probability, is used to reject $B_{\rm sig}$ candidates where a non-kaon track passes the the VT K-PID criteria. Four quality levels of μ-PID are used. In order of decreasing efficiency and misidentification probability, they are Very Loose (VL), Loose (L), Tight (T), and Very Tight (VT).

TABLE I: PID efficiencies and misidentification probabilities for the algorithms used in the analysis. The values are approximate and representative only for the laboratory frame momentum ($p_{\rm lab}$) specified (when given). More than one algorithm is used for kaons and muons. The abbreviations VL, L, T, and VT stand for selection quality levels Very Loose, Loose, Tight, and Very Tight.

Type	Efficiency	misidentification probability
K-VL	$> 95\%$	$< 6\%$ for π and μ with $p_{\rm lab} < 3.5\,{\rm GeV}/c$
K-VT	$> 85\%$	$\approx 1\%$ for π and μ with $p_{\rm lab} < 3.5\,{\rm GeV}/c$
π	$> 98\%$	$< 20\%$ for K
p	$\approx 80\%$	$< 0.5\%$ for K, π, μ, e
μ-VL	$\approx 90\%$	$< 15\%$ for π with $p_{\rm lab} < 1.25\,{\rm GeV}/c$, $< 4\%$ for π with $p_{\rm lab} > 1.25\,{\rm GeV}/c$
μ-L	$\approx 80\%$	$< 5\%$ for π with $p_{\rm lab} < 1.25\,{\rm GeV}/c$, $< 2\%$ for π with $p_{\rm lab} > 1.25\,{\rm GeV}/c$
μ-T	$\approx 75\%$	$< 3\%$ for π with $p_{\rm lab} < 1.25\,{\rm GeV}/c$, $\approx 1\%$ for π with $p_{\rm lab} > 1.25\,{\rm GeV}/c$
μ-VT	$\approx 70\%$	$< 2\%$ for π with $p_{\rm lab} < 1.25\,{\rm GeV}/c$, $< 1\%$ for π with $p_{\rm lab} > 1.25\,{\rm GeV}/c$
e	95%	$< 0.2\%$ for π, K, p

C. Signal B reconstruction

The eight $B \to h\tau\ell$ decay modes are independently analyzed. Tracks for the signal B reconstruction must satisfy the following criteria: the distance of closest approach (DOCA) to the beam axis in the transverse plane must be less than 1.5 cm; the z position of the DOCA point must be less than 2.5 cm from the primary vertex of the event; the transverse momentum must be > 50 MeV/c; and the momentum must be < 10 GeV/c. After selecting the best $B_{\rm tag}$ candidate, we require exactly three tracks satisfying the above criteria remain in the event (excluding the $B_{\rm tag}$ daughters) and that the sum of the charges of these tracks be the opposite of the $B_{\rm tag}$ candidate charge. We refer to these three tracks as the $B_{\rm sig}$ daughters.

We require the primary hadron, which is the h in $B \to h\tau\ell$, to be one of the two $B_{\rm sig}$ daughters with the same charge as the $B_{\rm sig}$ candidate. The primary hadron must pass the K-VL-PID criteria for the $B \to K\tau\ell$ modes and the π-PID criteria for the $B \to \pi\tau\ell$ modes. For the $B \to K\tau\ell$ modes, if both of the $B_{\rm sig}$ daughters with the same charge meet the minimal K-PID criteria, the one with the highest K-PID quality level is selected as the primary $K^{\pm}$. If they have the same K-PID quality level, we choose the one with the lower momentum as the primary $K^{\pm}$. For the $B \to \pi\tau\ell$ modes, if both $B_{\rm sig}$ daughters with the same charge meet the π-PID criteria, we choose the one that gives m_τ closest to the true τ mass. This algorithm does not produce an artificial peak

in the signal window of the background m_τ distribution. Once the primary hadron candidate has been assigned, the τ daughter and primary lepton are uniquely defined for a given $B \to h\tau\ell$ mode from the remaining two $B_{\rm sig}$ daughters based on their electric charge.

The primary lepton, which is the ℓ in $B \to h\tau\ell$, must pass either the e-PID or the loosest μ-PID criteria (μ-VL). We remove events where any of the three $B_{\rm sig}$ daughters passes the p-PID criteria, or where any of the three $B_{\rm sig}$ daughters passes the K-VT-PID criteria, with the exception of the $K^\pm$ in $B \to K\tau\ell$.

By requiring exactly three $B_{\rm sig}$ daughters, we are restricting the selection to one-prong τ decays. For each of the eight $B \to h\tau\ell$ modes, we divide the selection into three τ decay channels: electron, muon, and pion. From now on, we use "modes" to refer to types of $B \to h\tau\ell$ decays and "channels" to refer to types of τ decays. The three τ decay channels are analyzed in parallel, with different background rejection criteria applied. If the τ daughter satisfies the e-PID criteria, the event is assigned to the electron channel. If the τ daughter does not satisfy the e-PID, but does satisfy the μ-VL-PID criteria, the event is assigned to the muon channel. If the τ daughter passes neither the e-PID or the μ-VL-PID, the event is assigned to the pion channel. This ensures that an event does not get double counted and categorized into another τ decay channel for a given $B \to h\tau\ell$ mode.

Background events with a $B \to h(c\bar{c})$; $(c\bar{c}) \to \ell^+\ell^-$ decay can pass our signal selection criteria. We remove events in the electron (muon) and pion τ decay channels of the $B \to h\tau e$ ($B \to h\tau\mu$) modes if the invariant mass of the primary lepton and τ daughter, $m_{\ell\ell}$, is consistent with a dilepton charmonium decay: $3.03 < m_{\ell\ell} < 3.14\,{\rm GeV}/c^2$ for the J/ψ or $3.60 < m_{\ell\ell} < 3.75\,{\rm GeV}/c^2$ for the $\psi(2S)$. The core dilepton invariant mass resolution for these charmonium decays is on the order of 12 MeV/c^2. These charmonium vetos effectively remove the charmonium background at a minimal cost in signal efficiency. We also require $m_{\ell\ell} > 0.1\,{\rm GeV}/c^2$ for $B \to h\tau e$ candidates in the electron and pion channels to remove candidates where the primary electron and the τ daughter are consistent with originating from a photon conversion.

D. $B\overline{B}$ background and the $m(K\pi)$ invariant mass requirement

After the selection described above, the dominant background is due to $B\overline{B}$ events, where the $B_{\rm tag}$ is properly reconstructed. However, the largest background source differs depending on the charge of the primary lepton relative to the charge of the $B_{\rm sig}$ candidate.

When the primary lepton charge is the same as the $B_{\rm sig}$ charge, such as a $B^+ \to K^+\tau^-\ell^+$ candidate, the dominant background comes from semileptonic B decays, such as $B^+ \to \overline{D}^{(*)0}\ell^+\nu$; $\overline{D}^0 \to K^+X^-$, where X^- contains a π^-, e^-, or μ^- and perhaps other charged and/or neutral daughters that are not reconstructed. For example, the final state tracks $K^+\pi^-\ell^+$ are identical for this background with $\overline{D}^0 \to K^+\pi^-$ and the $B^+ \to K^+\tau^-\ell^+$ signal decay with $\tau^- \to \pi^-\nu_\tau$. On the other hand, when the primary lepton charge is opposite to the $B_{\rm sig}$ charge, such as for a $B^+ \to K^+\tau^+\ell^-$ candidate, the dominant background comes from semileptonic D decays, such as $B^+ \to \overline{D}^{(*)0}X^+$; $\overline{D}^0 \to K^+\ell^-\bar{\nu}_\ell$.

To reduce these backgrounds, we reject $B_{\rm sig}$ candidates where two of the $B_{\rm sig}$ daughters are kinematically compatible with originating from a charm decay, as described below. For the four $B \to K\tau\ell$ modes, we define the variable $m(K\pi)$ as the invariant mass of the primary K and the $B_{\rm sig}$ daughter that has opposite charge to this K. In computing $m(K\pi)$, the non-K track is assumed to be a pion. Distributions of $m(K\pi)$ for the background and signal MC are shown in Fig. 1 for $B \to K\tau\mu$. For the four $B \to \pi\tau\ell$ modes, we define $m(K\pi)$ by combining two $B_{\rm sig}$ daughters that have opposite charge. Of the two $B_{\rm sig}$ daughters with the same charge as the $B_{\rm sig}$ candidate, we choose the one with the highest K-PID quality level. We assume that the kaon is one of the $B_{\rm sig}$ daughters with the same charge as the $B_{\rm sig}$ candidate and the pion is the $B_{\rm sig}$ daughter with the opposite charge as the $B_{\rm sig}$ candidate. If the two $B_{\rm sig}$ daughters with the same charge as the $B_{\rm sig}$ candidate have the same K-PID quality level, we use the daughter with higher momentum as the kaon in the $m(K\pi)$ calculation.

We require $m(K\pi) > 1.95\,{\rm GeV}/c^2$. This rejects between 97% and 99% of the background while retaining between 32% and 37% of the signal for the $B^+ \to h^+\tau^-\ell^+$ modes. For the $B^+ \to \pi^+\tau^+\ell^-$ modes, the $m(K\pi)$ requirement rejects 85% and 89% of the $\pi^+\tau^+\mu^-$ and $\pi^+\tau^+e^-$ background while retaining 72% and 65% of the signal, respectively. For the $B^+ \to K^+\tau^+\ell^-$ modes, the $m(K\pi)$ requirement rejects 92% and 96% of the $K^+\tau^+\mu^-$ and $K^+\tau^+e^-$ background while retaining 63% and 62% of the signal, respectively.

IV. $B \to D^{(*)0}\ell\nu$ CONTROL SAMPLE

We select a control sample of semileptonic B decays of the form $B^+ \to \overline{D}^{(*)0}\ell^+\nu$; $\overline{D}^0 \to K^+\pi^-$ by requiring $m(K\pi)$ to be near the D^0 mass, $1.845 < m(K\pi) < 1.885\,{\rm GeV}/c^2$. The $D^{(*)0}\ell\nu$ control sample has a negligible amount of combinatorial background. In our search for $B \to h\tau\ell$, we normalize the $B \to h\tau\ell$ branching fraction by using the measured $D^{(*)0}\ell\nu$ yield taken from the control sample. We determine the relative amounts of B mesons that decay to $\overline{D}^0$, $\overline{D}^{*0}$, and higher resonances ($\overline{D}^{**0}$) using the reconstructed CM energy difference

$$\begin{aligned} E_\nu &= p_\nu = |-\vec{p}_{\rm tag} - \vec{p}_K - \vec{p}_\pi - \vec{p}_\ell\,|\,, \\ \Delta E_{D\ell\nu} &= E_K + E_\pi + E_\ell + E_\nu - E_{\rm beam}. \end{aligned}$$

For $B^+ \to \overline{D}^0\ell^+\nu$ decays, $\Delta E_{D\ell\nu}$ is centered at zero. The missing neutral particles from $\overline{D}^{*0}$ and $\overline{D}^{**0}$ decays shift $\Delta E_{D\ell\nu}$ in the negative direction.

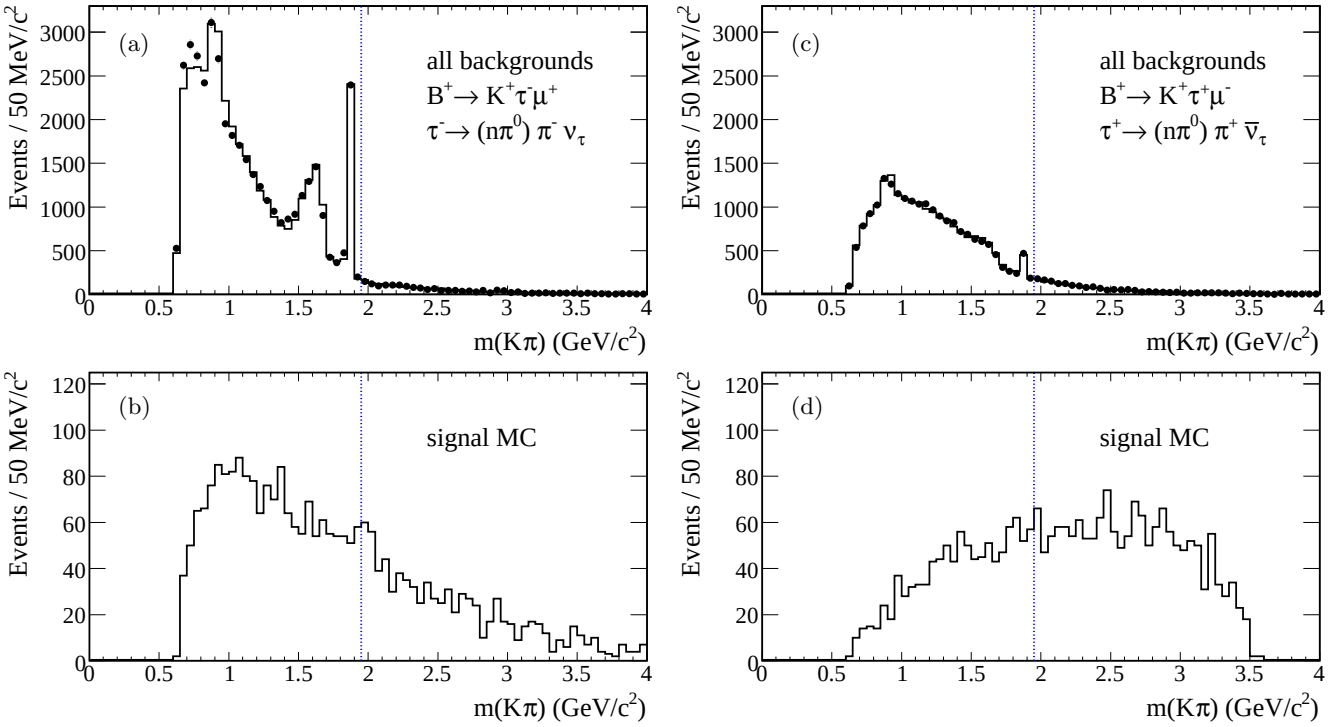

FIG. 1: Distributions of $m(K\pi)$ for the (a, b) $B^+ \to K^+\tau^-\mu^+$; $\tau^- \to (n\pi^0)\pi^-\nu_\tau$ and (c, d) $B^+ \to K^+\tau^+\mu^-$; $\tau^+ \to (n\pi^0)\pi^+\bar{\nu}_\tau$ channels. The top row shows the data (points) compared with the background MC (solid line). The area of the background MC distribution has been normalized to the area of the data distribution. The bottom row shows the (b) $B^+ \to K^+\tau^-\mu^+$ and (d) $B^+ \to K^+\tau^+\mu^-$ signal MC. The normalization of the bottom row is arbitrary. The dotted vertical line is at $1.95\,\mathrm{GeV}/c^2$, which is the minimum allowed value of $m(K\pi)$ for the signal selection. The peak in the top row just below $1.95\,\mathrm{GeV}/c^2$ is from $\overline{D}^0 \to K^+\pi^-$ decays.

The expected observed yields of $D\ell\nu$ and $h\tau\ell$ as functions of their branching fractions are given by

$$N_{D\ell\nu} = N_0\,\mathcal{B}_{D\ell\nu}\,\epsilon_{\rm tag}^{D\ell\nu}\,\epsilon_{D\ell\nu}, \quad (1)$$

$$N_{h\tau\ell} = N_0\,\mathcal{B}_{h\tau\ell}\,\epsilon_{\rm tag}^{h\tau\ell}\,\epsilon_{h\tau\ell}, \quad (2)$$

where N_0 is the number of $B\overline{B}$ events, $\mathcal{B}_{D\ell\nu}$ ($\mathcal{B}_{h\tau\ell}$) is the branching fraction for $B \to D\ell\nu$ ($B \to h\tau\ell$), $\epsilon_{\rm tag}^{D\ell\nu}$ ($\epsilon_{\rm tag}^{h\tau\ell}$) is the $B_{\rm tag}$ reconstruction efficiency in $B\overline{B}$ events that contain a $D\ell\nu$ ($h\tau\ell$) decay on the signal side, $\epsilon_{D\ell\nu}$ ($\epsilon_{h\tau\ell}$) is the signal-side reconstruction efficiency for $D\ell\nu$ ($h\tau\ell$), and the symbol $D\ell\nu$ represents either $B^+ \to \overline{D}^{*0}\ell^+\nu$ or $B^+ \to \overline{D}^0\ell^+\nu$. Solving for the expected $h\tau\ell$ event yield gives

$$N_{h\tau\ell} = \mathcal{B}_{h\tau\ell}\,\epsilon_{h\tau\ell}\,S_0, \quad (3)$$

where we have defined a common factor

$$S_0 = \frac{N_{D\ell\nu}}{\mathcal{B}_{D\ell\nu}\,\epsilon_{D\ell\nu}}\left(\frac{\epsilon_{\rm tag}^{h\tau\ell}}{\epsilon_{\rm tag}^{D\ell\nu}}\right). \quad (4)$$

Table II gives the tag-side efficiency ratios determined from MC samples. We find the ratios to be close to one, indicating that the signal-side decay does not strongly influence the tag-side reconstruction efficiency, and does not depend on the primary lepton or hadron flavor.

TABLE II: Tag-side reconstruction efficiency ratios determined from MC samples. The uncertainty includes both statistical and systematic sources.

Efficiency Ratio	μ modes	e modes
$\epsilon_{\rm tag}^{K\tau\ell}/\epsilon_{\rm tag}^{D\ell\nu}$	0.96 ± 0.05	0.98 ± 0.07
$\epsilon_{\rm tag}^{\pi\tau\ell}/\epsilon_{\rm tag}^{D\ell\nu}$	0.95 ± 0.04	0.97 ± 0.06

Figure 2 shows the results of unbinned maximum likelihood fits of the $\Delta E_{D\ell\nu}$ distributions for the $B^+ \to \overline{D}^{(*)0}\mu^+\nu$ and $B^+ \to \overline{D}^{(*)0}e^+\nu$ control samples. The fits have independent $\overline{D}^0$, $\overline{D}^{*0}$, and $\overline{D}^{**0}$ components. Any residual combinatorial background is included in the $\overline{D}^{**0}$ component. The $\overline{D}^0$ and $\overline{D}^{*0}$ component probability density functions (PDFs) are each modeled with the sum of a Gaussian and a Crystal Ball function [17]. The $\overline{D}^{**0}$ component PDF is the sum of a Gaussian and a bifurcated Gaussian, which has different width parameters above and below the mean. The overall normalization of each component, the core Gaussian mean and width of the $\overline{D}^0$ component, and the relative fraction of the

Crystal Ball function within the $\overline{D}^{*0}$ component are all parameters of the likelihood that are varied in its maximization.

The results of the $\Delta E_{D\ell\nu}$ maximum likelihood fits and S_0 calculations are given in Table III. We use the following branching fractions [18] in the calculation of S_0: $\mathcal{B}(B^- \to D^0\ell^-\bar{\nu}) = (2.23 \pm 0.11)\%$, $\mathcal{B}(B^- \to D^{*0}\ell^-\bar{\nu}) = (5.68 \pm 0.19)\%$, and $\mathcal{B}(D^0 \to K^-\pi^+) = (3.87 \pm 0.05)\%$. The four determinations of S_0 are all consistent with each other, as expected.

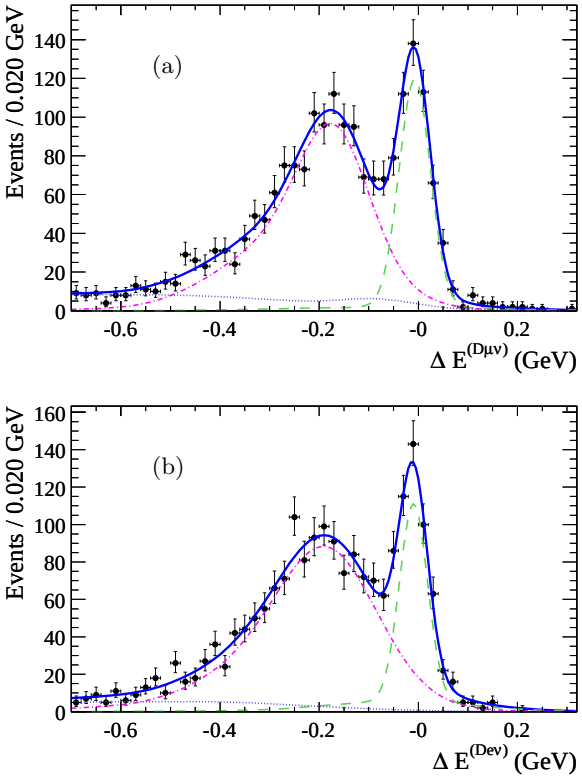

FIG. 2: (color online) Distributions of the three-component $\Delta E_{D\ell\nu}$ unbinned maximum likelihood fits of the data for the (a) $B \to D^{(*)0}\mu\nu$ and (b) $B \to D^{(*)0}e\nu$ control samples. In each plot, the points represent the data, the solid blue curve is the sum of all PDFs, the long-dashed green curve is the D^0 component, the dot-dashed purple curve is the D^{*0} component, and the dotted blue curve is the D^{**0} component, which also includes any residual combinatorial background.

V. CONTINUUM BACKGROUND REJECTION

After the $m(K\pi) > 1.95\,\mathrm{GeV}/c^2$ requirement, the $B\overline{B}$ background is highly suppressed. The remaining background is dominated by continuum quark-pair production ($e^+e^- \to q\bar{q}$; $q = u, d, s, c$). We combine the variables described in this section in a likelihood ratio

$$L_R = \frac{\prod_i P_s(x_i)}{\prod_i P_s(x_i) + \prod_i P_b(x_i)} \qquad (5)$$

TABLE III: Results of the $\Delta E_{D\ell\nu}$ maximum likelihood fits and S_0 calculations. The uncertainties on $N_{D\ell\nu}$ and $\epsilon_{D\ell\nu}$ are statistical. The efficiency $\epsilon_{D\ell\nu}$ is determined from a Monte Carlo sample. The uncertainty on S_0 includes the uncertainties on the B and D branching fractions.

$D\ell\nu$ mode	$N_{D\ell\nu}$	$\epsilon_{D\ell\nu}$	S_0
$D^0\mu\nu$	513 ± 38	$(47.8 \pm 0.9)\%$	$(12.0 \pm 1.2) \times 10^5$
$D^{*0}\mu\nu$	1234 ± 49	$(50.8 \pm 0.5)\%$	$(10.7 \pm 0.8) \times 10^5$
$D^0e\nu$	484 ± 46	$(48.2 \pm 0.9)\%$	$(11.4 \pm 1.5) \times 10^5$
$D^{*0}e\nu$	1368 ± 58	$(52.2 \pm 0.5)\%$	$(11.7 \pm 1.1) \times 10^5$

where x_i is one of a set of variables that discriminate against background, and $P_s(x_i)$ ($P_b(x_i)$) is the PDF for variable x_i in signal (background) events.

The variables used in the L_R calculation are:

- $\boldsymbol{|\cos\theta_{\rm thr}|}$ the absolute value of the cosine of the angle $\theta_{\rm thr}$ between the $B_{\rm tag}$ thrust axis and the thrust axis of the remainder of the event ($\equiv B_{\rm sig}$); the thrust axis is defined as the direction $\hat{a}$ which maximizes $\sum_j \hat{a}\cdot\vec{p}_j$, where j represents all particles assigned to a particular B candidate,
- $\boldsymbol{\sum E_{\rm cal}}$ the scalar sum of all EMC neutral cluster energy that is not associated with the $B_{\rm tag}$ candidate or bremsstrahlung radiation from any e candidates, where the threshold cluster energy is 100 MeV (50 MeV) in the forward (barrel) region of the detector,
- **primary μ-PID quality level**, where, for the $B \to h\tau\mu$ modes, we include the highest quality level (VL, L, T, VT) of the primary μ candidate, and
- **secondary μ-PID quality level**, where we include the highest quality level (VL, L, T, VT) of the τ-daughter μ candidate, if applicable.

We fit histograms of the $|\cos\theta_{\rm thr}|$ and $\sum E_{\rm cal}$ signal and background MC distributions using polynomials of up to order eight to define the PDFs for those variables. The PDFs for the muon-PID quality level are normalized histograms, with one bin for each muon-PID quality level.

For each of the eight signal B decay modes, we construct a distinct L_R for each of the three τ channels (e, μ, and π). This corresponds to 24 different likelihood ratios. In the final selection, described in section VIII, we impose a minimum L_R requirement for each τ channel in each of the eight $B \to h\tau\ell$ modes.

Figure 3 shows the background and signal $|\cos\theta_{\rm thr}|$ distributions for the π channel of the $B^+ \to K^+\tau^-\mu^+$ mode. The continuum background peaks sharply near $|\cos\theta_{\rm thr}| = 1$ because the events have a back-to-back jet-like topology. The signal $|\cos\theta_{\rm thr}|$ distribution is

roughly uniform because the detected decay products in $B\overline{B}$ events are more isotropically distributed.

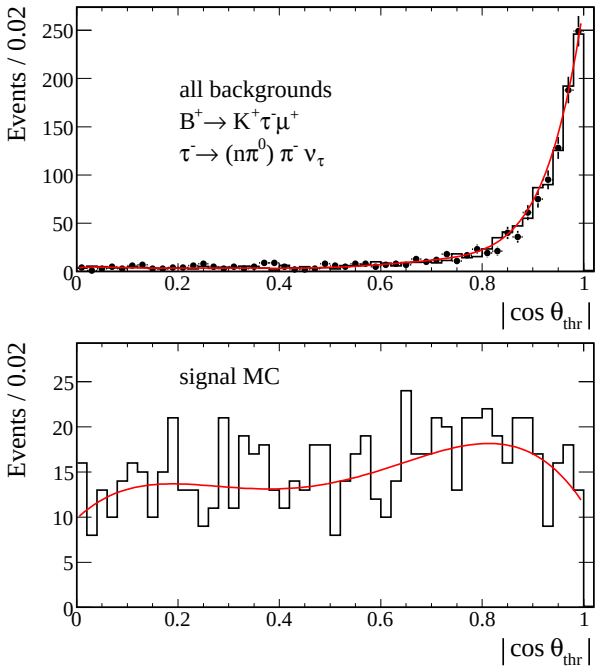

FIG. 3: Distributions of $|\cos\theta_{\rm thr}|$ for background (top) and signal MC (bottom), for the $B^+ \to K^+\tau^-\mu^+$; $\tau^- \to (n\pi^0)\pi^-\nu_\tau$ channel. The points (solid line) in the top figure are the data (background MC). The background MC has been normalized to match the area of the data distribution. The normalization of the signal MC is arbitrary. The solid red curve is the result of the polynomial fit of the MC distribution.

Figure 4 shows $\sum E_{\rm cal}$ distributions for the three τ channels of the $B^+ \to K^+\tau^-\mu^+$ mode. The events where $\sum E_{\rm cal} = 0$, due to the absence of unassociated neutral clusters above the minimum energy threshold, are not included in the polynomial fit and treated separately. The $\sum E_{\rm cal} = 0$ events are plotted below zero in Figure 4 for clarity. The signal MC $\sum E_{\rm cal}$ distributions peak at zero, as expected, while the background rarely has $\sum E_{\rm cal} = 0$ but rather has a distribution that peaks between 1 and 2 GeV. The signal MC $\sum E_{\rm cal}$ distributions for the π channel extend to higher values, compared to the e and μ channels, due to hadronic τ decays that produce a single $\pi^\pm$ with one or more neutral pions.

Figure 5 shows background and signal MC L_R distributions for the $B^+ \to K^+\tau^-\mu^+$; $\tau^- \to (n\pi^0)\pi^-\nu_\tau$ channel. The background peaks sharply near zero and the signal peaks sharply near one. The value of the L_R selection for each τ channel in each of the eight signal modes is chosen by determining the lowest upper limit on the branching fractions under the null hypothesis with MC pseudo-experiments. We vary the minimum L_R requirement in intervals of 0.05.

VI. SIGNAL AND BACKGROUND ESTIMATION

In our signal selection, we require the indirectly reconstructed τ mass m_τ to be within $\pm 60\,{\rm MeV}/c^2$ of the world average τ mass $1.777\,{\rm GeV}/c^2$ [18]. The relative signal efficiency after the m_τ signal window requirement is around 84% (78%) for the $B \to h\tau\mu$ ($B \to h\tau e$) modes. We optimized the m_τ signal windows, considering windows in the range of ± 50 MeV/c^2 to ± 175 MeV/c^2. Our optimization metric was the average expected signal branching fraction 90% confidence level upper limit from a set of toy experiments simulating background-only datasets. In each toy experiment, we generate a value for the observed number of events in the signal window using a random number that we take from a Poisson distribution with the mean value set to the expected number of background events. We find that a m_τ signal window of ± 60 MeV/c^2 gives the lowest expected branching fraction upper limits for all τ decay channels.

The background distribution in m_τ is very wide and slowly varying. We use a broad m_τ sideband from 0 to $3.5\,{\rm GeV}/c^2$, excluding the signal window, to estimate the background in the m_τ signal window with

$$b \;=\; R_b\, N_{sb}, \qquad (6)$$

where b is the number of background events in the signal window, N_{sb} is the number of background events in the m_τ sideband, and R_b is the expected signal-to-sideband ratio (b/N_{sb}). The ratio R_b is determined from the ratio of selected background events in the m_τ signal window (b) and the m_τ sideband (N_{sb}) in the background Monte Carlo.

Figures 6 and 7 show the observed, signal MC, and background MC m_τ distributions for the $B \to K\tau\ell$ and $B \to \pi\tau\ell$ modes, respectively. Table IV gives the results for the observed numbers of sideband events $N_{sb,i}$, signal-to-sideband ratios $R_{b,i}$, expected numbers of background events b_i, numbers of observed events n_i, and signal efficiencies $\epsilon_{h\tau\ell,i}$ for each τ channel i. All of the observed numbers of events n_i in the m_τ signal window are statistically consistent with the expected backgrounds b_i, thus there is no evidence for any $B \to h\tau\ell$ decay.

VII. SYSTEMATIC UNCERTAINTIES

Since we normalize our $B \to h\tau\ell$ signals using the $B \to D^{(*)0}\ell\nu$ control sample, many systematic uncertainties cancel, such as the ones coming from the absolute $B_{\rm tag}$ efficiency uncertainty and the tracking efficiency uncertainty. We evaluate systematic uncertainties on the efficiency of the minimum L_R requirement by varying the signal and background PDFs for each L_R. We use the $B \to D^{(*)0}\ell\nu$ control sample in place of the signal Monte Carlo as a variation of the signal $\sum E_{\rm cal}$ PDF. A uniform distribution is used in place of the nominal polynomial

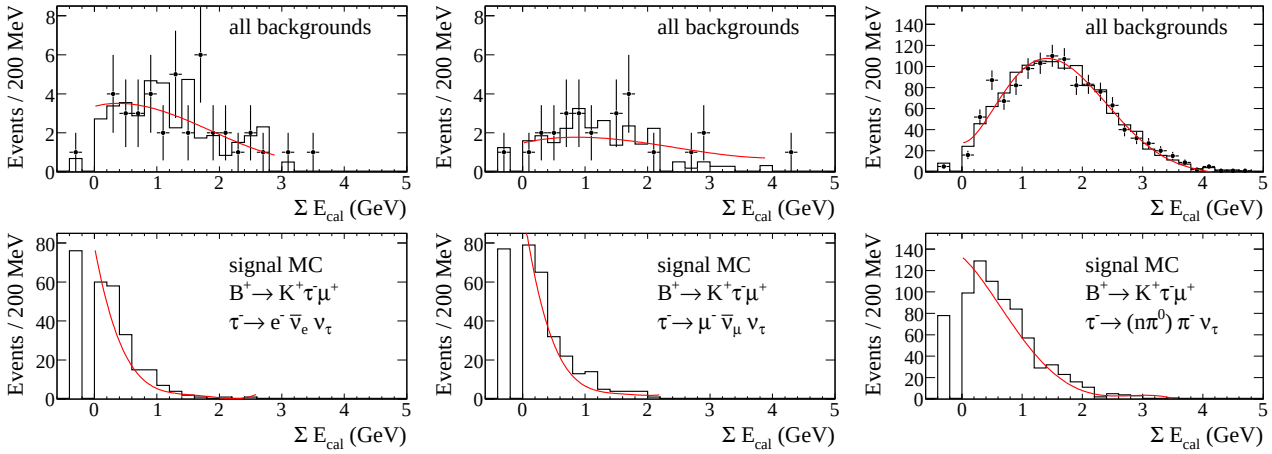

FIG. 4: Distributions of $\sum E_{\rm cal}$ for background (top) and signal MC (bottom), for the $B^+ \to K^+\tau^-\mu^+$ mode; $\tau^- \to e^-\bar{\nu}_e\nu_\tau$ (left), $\tau^- \to \mu^-\bar{\nu}_\mu\nu_\tau$ (middle), and $\tau^- \to (n\pi^0)\pi^-\nu_\tau$ (right). The events where $\sum E_{\rm cal} = 0$ have been separated from the main distribution and plotted in a bin below zero for clarity. The points (solid line) in the top figure are the data (background MC). The background MC has been normalized to match the area of the data distribution. The normalization of the signal MC is arbitrary. The solid red curve is the result of the polynomial fit of the MC distribution.

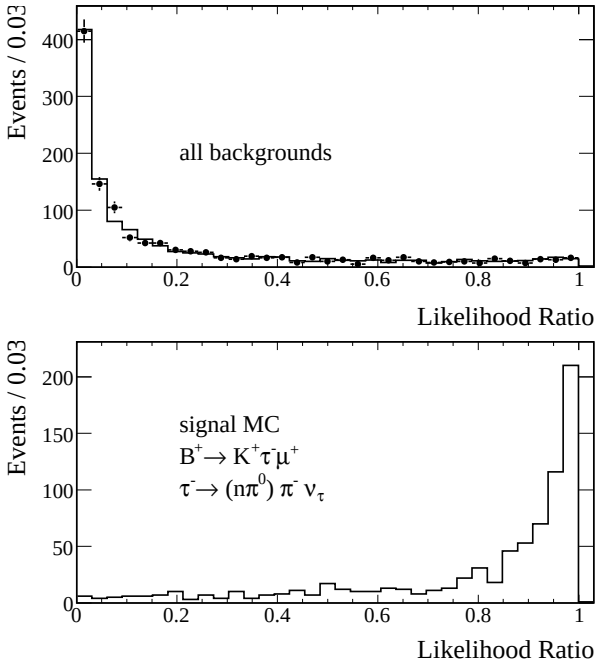

FIG. 5: Likelihood ratio (L_R) output distributions of background (top) and signal MC (bottom), for the $B^+ \to K^+\tau^-\mu^+$; $\tau^- \to (n\pi^0)\pi^-\nu_\tau$ channel. The points (solid line) in the top figure are the data (background MC). The background MC has been normalized to match the area of the data distribution. The normalization of the signal MC is arbitrary.

fit as the variation of the signal $|\cos\theta_{\rm thr}|$ PDF. The efficiency for each lepton PID level is varied by ±2.5% for the VL,L, and T levels and ±3.2% for the VT level. The data m_τ sideband is used in place of the Monte Carlo as a variation of the background PDFs.

Our largest sources of systematic uncertainty come from variations in modeling the data distributions of the $\sum E_{\rm cal}$ and $|\cos\theta_{\rm thr}|$ L_R inputs when compared to the nominal background MC PDFs. The changes in $\epsilon_{h\tau\ell,i}$ from the variations are added in quadrature. We determine systematic uncertainties as high as 1.1%, with the largest ones coming from the $B^+ \to K^+\tau^+e^-$; $\tau^+ \to e^+\nu_e\bar{\nu_\tau}$ and $\tau^+ \to (n\pi^0)\pi^+\bar{\nu_\tau}$ channels.

The $B \to \pi\tau\ell$ modes require π-PID, while the $B \to D^{(*)0}\ell\nu$ control sample requires K-PID. We evaluate a systematic uncertainty on $\epsilon_{\pi\tau\ell}/\epsilon_{D\ell\nu}$ by measuring the π-PID and K-PID efficiencies using the $B \to D^{(*)0}\ell\nu$ control sample with and without the K-PID or π-PID requirements. The measured efficiencies in data are consistent with the MC simulation. Based on the results from the $B \to D^{(*)0}\mu\nu$ and $B \to D^{(*)0}e\nu$ samples, we assign systematic uncertainties of 1.8% and 1.0% to $\epsilon_{\pi\tau\mu}/\epsilon_{D\mu\nu}$ and $\epsilon_{\pi\tau e}/\epsilon_{De\nu}$, respectively.

The uncertainty on the signal-to-sideband ratio R_b is the statistical uncertainty from the Monte Carlo sample used to determine its value. Figures 6 and 7 show good agreement between the Monte Carlo and observed data distributions in the sidebands. No additional systematic uncertainty is included in R_b.

The tag efficiency ratio $\epsilon_{\rm tag}^{h\tau\ell}/\epsilon_{\rm tag}^{D\ell\nu}$ is evaluated using two independent Monte Carlo samples: one where the tag-side B meson decays to all possible final states and another where the tag-side B meson is forced to decay to most of the modes that comprise the tag-side reconstruction. The value of the ratio is taken from the first sample. The systematic error on the ratio is the difference in the ratio between the two samples. The overall uncertainty on the ratio, given in Table II, is the sum in quadrature of the statistical and systematic uncertainties

on the ratio.

VIII. BRANCHING FRACTION RESULTS

We determine the branching fraction for each of the eight $B \to h\tau\ell$ modes using a likelihood function which is the product of three Poisson PDFs, one for each of the three τ channels. The expected number of events in a particular τ channel is given by

$$n_i = N_{h\tau\ell,i} + b_i \quad (7)$$
$$n_i = \mathcal{B}_{h\tau\ell}\ \epsilon_{h\tau\ell,i}\ S_0 + b_i, \quad (8)$$

where $N_{h\tau\ell,i}$ (b_i) is the expected number of signal (background) events in channel i. Total uncertainties on the signal efficiency $\epsilon_{h\tau\ell,i}$, common factor S_0, and expected background b_i are included by convolving the likelihood with Gaussian distributions in $\epsilon_{h\tau\ell,i}$, S_0, and b_i.

We set 90% confidence intervals on the branching fractions of the eight $B \to h\tau\ell$ modes assuming uniform three-body phase space decays using the likelihood ratio ordering principle of Feldman and Cousins [19] to construct the confidence belts.

The 90% C.L. upper limits on the $B \to h\tau\ell$ branching fractions are between 1.5×10^{-5} and 7.4×10^{-5}. Table IV includes the final results for the $B \to h\tau\ell$ branching fraction and 90% C.L. upper limits. In Table V, we give combined results for $\mathcal{B}(B^+ \to h^+\tau\ell) \equiv \mathcal{B}(B^+ \to h^+\tau^-\ell^+) + \mathcal{B}(B^+ \to h^+\tau^+\ell^-)$ with the assumption $\mathcal{B}(B^+ \to h^+\tau^-\ell^+) = \mathcal{B}(B^+ \to h^+\tau^+\ell^-)$.

In the analysis of Black, Han, He, and Sher [4], the $B \to K\tau\mu$ and $B \to \pi\tau\mu$ branching fractions are proportional to $\Lambda_{\bar{b}s}^{-4}$ and $\Lambda_{\bar{b}d}^{-4}$, which are the new physics energy scales for the corresponding fermionic effective operators for these decays. Using the limits $\mathcal{B}(B^+ \to \pi^+\tau\mu) < 7.2 \times 10^{-5}$ and $\mathcal{B}(B^+ \to K^+\tau\mu) < 4.8 \times 10^{-5}$, we improve the model-independent bounds on the energy scale of new physics in flavor-changing operators reported in [4] from $\Lambda_{\bar{b}d} > 2.2$ TeV and $\Lambda_{\bar{b}s} > 2.6$ TeV to $\Lambda_{\bar{b}d} > 11$ TeV and $\Lambda_{\bar{b}s} > 15$ TeV.

IX. SUMMARY AND CONCLUSIONS

We have searched for the lepton flavor violating decays $B \to h\tau\ell$. We find no evidence for these decays and set 90% C.L. upper limits on the branching fractions of a few times 10^{-5}. The results for the $B \to K\tau\mu$ mode supersede our previous result [6]. The results for $B \to K\tau e$, $B \to \pi\tau\mu$, $B \to \pi\tau e$ modes are the first experimental limits for these decays. We use our results to improve model-independent limits on the energy scale of new physics in flavor-changing operators [4] to $\Lambda_{\bar{b}d} > 11$ TeV and $\Lambda_{\bar{b}s} > 15$ TeV.

Acknowledgments

We are grateful for the extraordinary contributions of our PEP-II colleagues in achieving the excellent luminosity and machine conditions that have made this work possible. The success of this project also relies critically on the expertise and dedication of the computing organizations that support *BABAR*. The collaborating institutions wish to thank SLAC for its support and the kind hospitality extended to them. This work is supported by the US Department of Energy and National Science Foundation, the Natural Sciences and Engineering Research Council (Canada), the Commissariat à l'Energie Atomique and Institut National de Physique Nucléaire et de Physique des Particules (France), the Bundesministerium für Bildung und Forschung and Deutsche Forschungsgemeinschaft (Germany), the Istituto Nazionale di Fisica Nucleare (Italy), the Foundation for Fundamental Research on Matter (The Netherlands), the Research Council of Norway, the Ministry of Education and Science of the Russian Federation, Ministerio de Ciencia e Innovación (Spain), and the Science and Technology Facilities Council (United Kingdom). Individuals have received support from the Marie-Curie IEF program (European Union) and the A. P. Sloan Foundation (USA).

[1] S. Weinberg, Phys. Rev. Lett., **19**, 1264 (1967); S. L. Glashow, J. Iliopoulos, and L. Maiani, Phys. Rev. D **2**, 1285 (1970); see also J. Erler and P. Langacker in [18] and B. Kayser in [18].
[2] X.-G. He, G. Valencia, and Y. Wang, Phys. Rev. D **70**, 113011 (2004).
[3] M. Sher and Y. Yuan, Phys. Rev. D **44**, 1461 (1991).
[4] D. Black, T. Han, H.-J. He, and M. Sher, Phys. Rev. D **66**, 053002 (2002).
[5] T. Fujihara *et al.*, Phys. Rev. D **73**, 074011 (2006).
[6] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. Lett. **99**, 201801 (2007).
[7] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. D **77**, 091104(R) (2008).
[8] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. D **77**, 011107(R) (2008).
[9] Charge-conjugate decays are implied throughout the paper.
[10] B. Aubert *et al.* (*BABAR* Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A**479**, 1 (2002).
[11] D. Lange, Nucl. Instrum. Methods Phys. Res., Sect. A**462**, 152 (2001).
[12] S. Agostinelli *et al.*, (GEANT4 Collaboration),Nucl. Instrum. Methods Phys. Res., Sect. A**506**, 250 (2003).
[13] All energies and momenta are in the e^+e^- center-of-mass reference frame unless explicitly stated otherwise.
[14] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. D **74**, 031103 (2006).

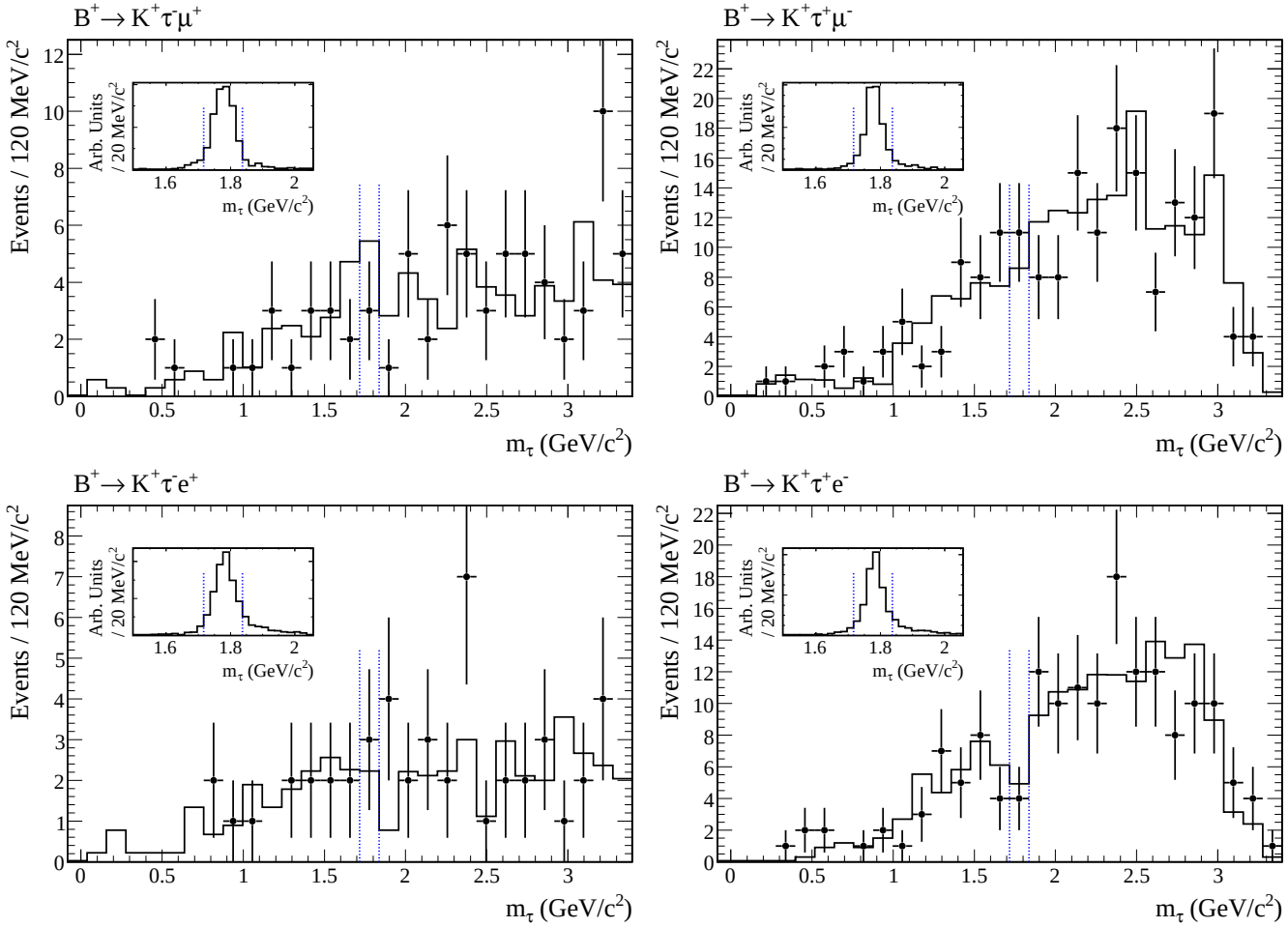

FIG. 6: Observed distributions of the τ invariant mass for the $B \to K\tau\ell$ modes. The distributions show the sum of the three τ channels (e, μ, π). The points with error bars are the data. The solid line is the background MC which has been normalized to the area of the data distribution. The dashed vertical lines indicate the m_τ signal window range. The inset shows the m_τ distribution for signal MC.

[15] T. G. Dietterich and G. Bakiri, Journal of Artificial Intelligence Research, **2**, 263 (1995).

[16] I. Narsky, contribution to PHYSTAT 2005 proceedings, ed. L. Lyons and M. K. Unel, Imperial Coll. Press, 2006; arXiv:physics/0507157v1.

[17] M. J. Oreglia, Ph. D Thesis, SLAC-236 (1980), Appendix D; J. E. Gaiser, Ph.D Thesis, SLAC-255 (1982), Appendix F; T. Skwarnicki, Ph.D Thesis, DESY F31-86-02 (1986), Appendix E.

[18] K. Nakamura *et al.* (Particle Data Group), JP G **37**, 075021 (2010) and 2011 partial update for the 2012 edition (URL: http://pdg.lbl.gov).

[19] G. J. Feldman and R. D. Cousins, Phys. Rev. D **57**, 3873 (1998).

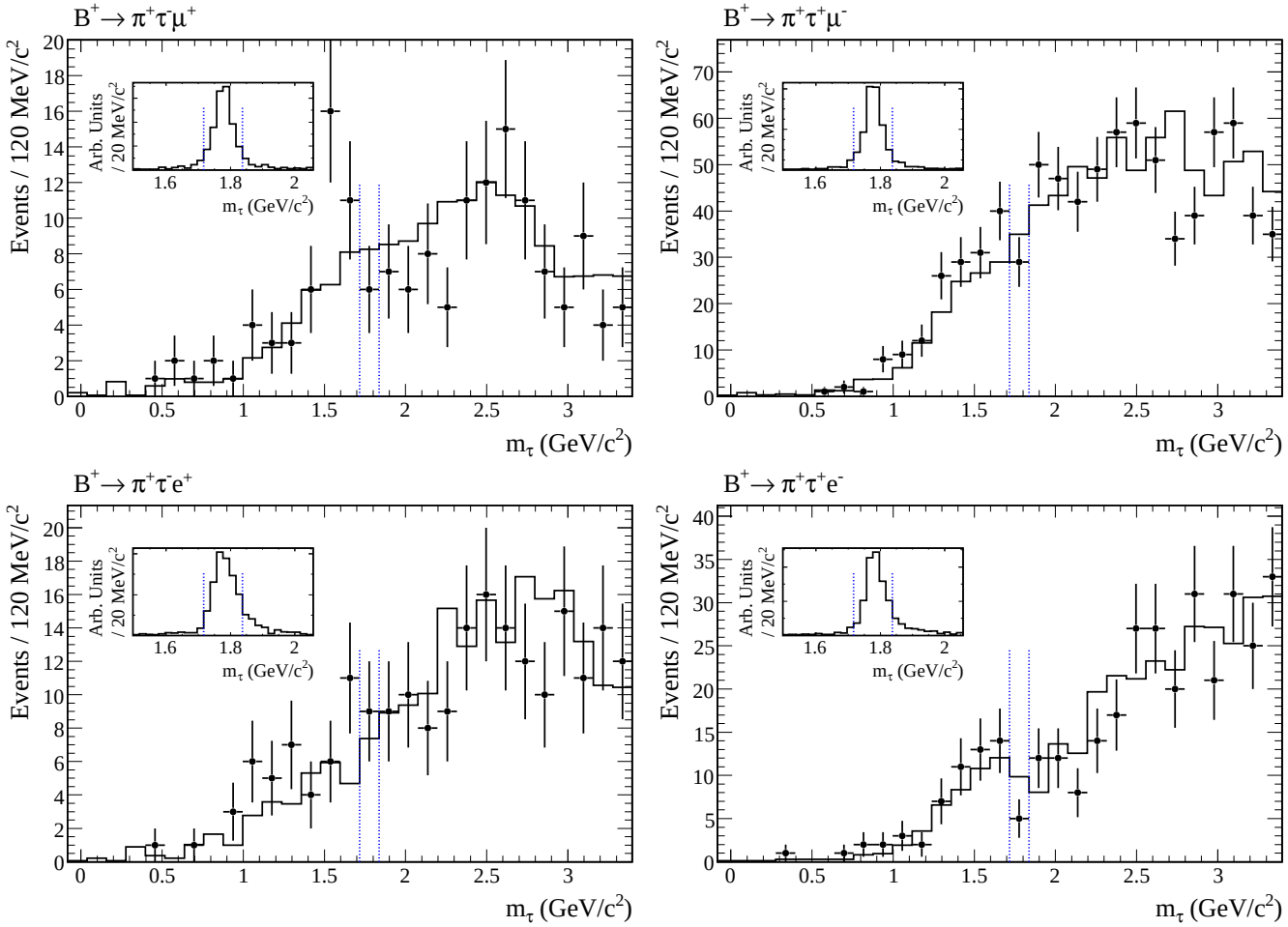

FIG. 7: Observed distributions of the τ invariant mass for the $B \to \pi\tau\ell$ modes. The distributions show the sum of the three τ channels (e, μ, π). The points with error bars are the data. The solid line is the background MC which has been normalized to the area of the data distribution. The dashed vertical lines indicate the m_τ signal window range. The inset shows the m_τ distribution for signal MC.

TABLE IV: Results for the observed sideband events $N_{sb,i}$, signal-to-sideband ratio $R_{b,i}$, expected background events b_i, number of observed events n_i, signal efficiency $\epsilon_{h\tau\ell,i}$ (assuming uniform three-body phase space decays) for each τ channel i and $B \to h\tau\ell$ [9] branching fraction central value and 90% C.L. upper limits (UL). All uncertainties include statistical and systematic sources.

							$\mathcal{B}(B \to h\tau\ell)$ $(\times 10^{-5})$	
Mode	τ channel	$N_{sb,i}$	$R_{b,i}$	b_i	n_i	$\epsilon_{h\tau\ell,i}$	central value	90% C.L. UL
$B^+ \to K^+\tau^-\mu^+$	e	22	0.02 ± 0.01	0.4 ± 0.2	2	$(2.6 \pm 0.2)\%$	$0.8\ ^{+1.9}_{-1.4}$	< 4.5
	μ	4	0.08 ± 0.05	0.3 ± 0.2	0	$(3.2 \pm 0.4)\%$		
	π	39	0.045 ± 0.020	1.8 ± 0.8	1	$(4.1 \pm 0.4)\%$		
$B^+ \to K^+\tau^+\mu^-$	e	5	0.03 ± 0.01	0.2 ± 0.1	0	$(3.7 \pm 0.3)\%$	$-0.4\ ^{+1.4}_{-0.9}$	< 2.8
	μ	3	0.06 ± 0.03	0.2 ± 0.1	0	$(3.6 \pm 0.7)\%$		
	π	153	0.045 ± 0.010	6.9 ± 1.5	11	$(9.1 \pm 0.5)\%$		
$B^+ \to K^+\tau^-e^+$	e	6	0.095 ± 0.020	0.6 ± 0.1	2	$(2.2 \pm 0.2)\%$	$0.2\ ^{+2.1}_{-1.0}$	< 4.3
	μ	4	0.025 ± 0.010	0.1 ± 0.1	0	$(2.7 \pm 0.6)\%$		
	π	33	0.045 ± 0.015	1.5 ± 0.5	1	$(4.8 \pm 0.6)\%$		
$B^+ \to K^+\tau^+e^-$	e	8	0.10 ± 0.06	0.8 ± 0.5	0	$(2.8 \pm 1.1)\%$	$-1.3\ ^{+1.5}_{-1.8}$	< 1.5
	μ	3	0.045 ± 0.020	0.1 ± 0.1	0	$(3.2 \pm 0.7)\%$		
	π	132	0.035 ± 0.010	4.6 ± 1.3	4	$(8.7 \pm 1.2)\%$		
$B^+ \to \pi^+\tau^-\mu^+$	e	55	0.017 ± 0.010	0.9 ± 0.6	0	$(2.3 \pm 0.2)\%$	$0.4\ ^{+3.1}_{-2.2}$	< 6.2
	μ	10	0.11 ± 0.04	1.1 ± 0.4	2	$(2.9 \pm 0.4)\%$		
	π	93	0.035 ± 0.010	3.3 ± 0.9	4	$(2.8 \pm 0.2)\%$		
$B^+ \to \pi^+\tau^+\mu^-$	e	171	0.012 ± 0.003	2.1 ± 0.5	2	$(3.8 \pm 0.3)\%$	$0.0\ ^{+2.6}_{-2.0}$	< 4.5
	μ	89	0.04 ± 0.01	3.6 ± 0.9	4	$(4.8 \pm 0.3)\%$		
	π	512	0.050 ± 0.005	25 ± 3	23	$(9.1 \pm 0.6)\%$		
$B^+ \to \pi^+\tau^-e^+$	e	1	0.050 ± 0.025	0.1 ± 0.1	1	$(2.0 \pm 0.8)\%$	$2.8\ ^{+2.4}_{-1.9}$	< 7.4
	μ	16	0.025 ± 0.010	0.4 ± 0.2	1	$(2.8 \pm 0.3)\%$		
	π	172	0.035 ± 0.008	6.0 ± 1.4	7	$(5.8 \pm 0.3)\%$		
$B^+ \to \pi^+\tau^+e^-$	e	31	0.033 ± 0.013	1.0 ± 0.4	0	$(2.9 \pm 0.3)\%$	$-3.1\ ^{+2.4}_{-2.1}$	< 2.0
	μ	247	0.012 ± 0.005	3.0 ± 1.2	2	$(4.6 \pm 0.4)\%$		
	π	82	0.07 ± 0.03	5.7 ± 2.5	3	$(3.7 \pm 1.0)\%$		

TABLE V: Branching fraction central values and 90% C.L. upper limits (UL) for the combination $\mathcal{B}(B^+ \to h^+\tau\ell) \equiv \mathcal{B}(B^+ \to h^+\tau^-\ell^+) + \mathcal{B}(B^+ \to h^+\tau^+\ell^-)$ with the assumption $\mathcal{B}(B^+ \to h^+\tau^-\ell^+) = \mathcal{B}(B^+ \to h^+\tau^+\ell^-)$. The uncertainties include statistical and systematic sources.

	$\mathcal{B}(B \to h\tau\ell)$ $(\times 10^{-5})$	
Mode	central value	90% C.L. UL
$B^+ \to K^+\tau\mu$	$0.0\ ^{+2.7}_{-1.4}$	< 4.8
$B^+ \to K^+\tau e$	$-0.6\ ^{+1.7}_{-1.4}$	< 3.0
$B^+ \to \pi^+\tau\mu$	$0.5\ ^{+3.8}_{-3.2}$	< 7.2
$B^+ \to \pi^+\tau e$	$2.3\ ^{+2.8}_{-1.7}$	< 7.5