Search for $B^+ \to \tau^+ \nu$
SEARCH FOR $B^+ \rightarrow \tau^+ \nu$

**Physical Review D 76, 052002 (2007)**

31 Imperial College London, London, SW7 2AZ, United Kingdom
32 University of Iowa, Iowa City, Iowa 52242, USA
33 Iowa State University, Ames, Iowa 50011-3160, USA
34 Johns Hopkins University, Baltimore, Maryland 21218, USA
35 Universität Karlsruhe, Institut für Experimentelle Kernphysik, D-76021 Karlsruhe, Germany
36 Laboratoire de l’Accélérateur Linéaire, IN2P3/CNRS et Université Paris-Sud 11, Centre Scientifique d’Orsay, B. P. 34, F-91898 ORSAY Cedex, France
37 Lawrence Livermore National Laboratory, Livermore, California 94550, USA
38 Queen Mary, University of London, E1 4NS, United Kingdom
39 University of Liverpool, Liverpool L69 7ZE, United Kingdom
40 University of London, Royal Holloway and Bedford New College, Egham, Surrey TW20 0EX, United Kingdom
41 University of Louisville, Louisville, Kentucky 40292, USA
42 University of Marylands, College Park, Maryland 20742, USA
43 University of Massachusetts, Amherst, Massachusetts 01003, USA
44 Massachusetts Institute of Technology, Laboratory for Nuclear Science, Cambridge, Massachusetts 02139, USA
45 McGill University, Montréal, Québec, Canada H3A 2T8
46 University of Miami, Dipartimento di Fisica and INFN, I-20133 Milano, Italy
47 University of Mississippi, University, Mississippi 38677, USA
48 Université de Montréal, Physique des Particules, Montréal, Québec, Canada H3C 3J7
49 Mount Holyoke College, South Hadley, Massachusetts 01075, USA
50 Università di Napoli Federico II, Dipartimento di Scienze Fisiche e INFN, I-80126, Napoli, Italy
51 NIKHEF, National Institute for Nuclear Physics and High Energy Physics, NL-1009 DB Amsterdam, The Netherlands
52 University of Notre Dame, Notre Dame, Indiana 46556, USA
53 Ohio State University, Columbus, Ohio 43210, USA
54 Pennsylvania State University, University Park, Pennsylvania 16802, USA
55 Princeton University, Princeton, New Jersey 08544, USA
56 Rutherford Appleton Laboratory, Chilton, Didcot, Oxon, OX11 0QX, United Kingdom
57 Université de Paris-Sud, Laboratoire de Physique de l’Acérrateur Linéaire, IN2P3/CNRS and Université Paris-Sud 11, Centre Scientifique d’Orsay, B. P. 34, F-91898 ORSAY Cedex, France
58 University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA
59 Università di Perugia, Dipartimento di Fisica e INFN, I-06100 Perugia, Italy
60 Università di Pisa, Dipartimento di Fisica, Scuola Normale Superiore e INFN, I-56127 Pisa, Italy
61 Prairie View A&M University, Prairie View, Texas 77446, USA
62 University of Rochester, Rochester, New York 14627, USA
63 Università di Roma La Sapienza, Dipartimento di Fisica and INFN, I-00185 Roma, Italy
64 Università di Rostock, D-18051 Rostock, Germany
65 Rutherford Appleton Laboratory, Chilton, Didcot, Oxon, OX11 0QX, United Kingdom
66 DSM/Dapnia, CEA/Saclay, F-91191 Gif-sur-Yvette, France
67 University of South Carolina, Columbia, South Carolina 29208, USA
68 Stanford Linear Accelerator Center, Stanford, California 94309, USA
69 Stanford University, Stanford, California 94305-4060, USA
70 State University of New York, Albany, New York 12222, USA
71 University of Tennessee, Knoxville, Tennessee 37996, USA
72 University of Texas at Austin, Austin, Texas 78712, USA
73 University of Texas at Dallas, Richardson, Texas 75083, USA
74 Università di Torino, Dipartimento di Fisica Sperimentale and INFN, I-10125 Torino, Italy
75 Università di Trieste, Dipartimento di Fisica e INFN, I-34127 Trieste, Italy
76 IFIC, Universitat de Valencia-CSIC, E-46071 Valencia, Spain
77 University of Victoria, Victoria, British Columbia, Canada V8W 3P6
78 Department of Physics, University of Warwick, Coventry CV4 7AL, United Kingdom
79 University of Wisconsin, Madison, Wisconsin 53706, USA
80 Yale University, New Haven, Connecticut 06511, USA
81 (Received 14 May 2007; published 18 September 2007)

---

*Deceased
†Also with Università di Perugia, Dipartimento di Fisica, Perugia, Italy
‡Also with Università della Basilicata, Potenza, Italy
§Also with Universitat de Barcelona, Facultat de Fisica, Departament ECM, E-08028 Barcelona, Spain
‖Also with IPPP, Physics Department, Durham University, Durham DH1 3LE, United Kingdom

052002-3
I. INTRODUCTION

In the standard model (SM), the purely leptonic decay $B^+ \to \tau^+ \nu$ [1] proceeds via quark annihilation into a $W^+$ boson. The branching fraction is given by

$$\mathcal{B}(B^+ \to \tau^+ \nu) = \frac{G_F^2 m_B m_\tau^2}{8\pi^2} \left[ 1 - \frac{m_\tau^2}{m_B^2} \right]^2 \tau_{B^+} f_{B^+} |V_{ub}|^2,$$  \hspace{1cm} (1)

where $V_{ub}$ is an element of the Cabibbo-Kobayashi-Maskawa quark-mixing matrix [2,3], $f_{B^+}$ is the $B$ meson decay constant, $G_F$ is the Fermi constant, $\tau_{B^+}$ is the $B^+$ lifetime, and $m_B$ and $m_\tau$ are the $B^+$ and $\tau$ masses. Physics beyond the SM, such as two-Higgs doublet models, could enhance or suppress $\mathcal{B}(B^+ \to \tau^+ \nu)$ through the introduction of a charged Higgs boson [4–6]. Using theoretical calculations of $f_{B^+}$ from lattice QCD and experimental measurements of $|V_{ub}|$ from semileptonic $B$ decays, this purely leptonic $B$ decay can be used to constrain the parameters of theories beyond the SM. Or, assuming that SM processes dominate and using the value of $|V_{ub}|$ determined from semileptonic $B$ decays, purely leptonic decays provide an experimental method of measuring $f_{B^+}$ with reduced theoretical error.

The branching fractions for $B^+ \to \mu^+ \nu$ and $B^+ \to e^+ \nu$ are suppressed by factors of $\sim 5 \times 10^{-3}$ and $\sim 10^{-7}$ with respect to $B^+ \to \tau^+ \nu$. However, a search for $B^+ \to \tau^+ \nu$ is experimentally challenging due to the large missing momentum from multiple neutrinos, which makes the signature less distinctive than in the other leptonic modes. The SM estimate of the branching fraction for $B^+ \to \tau^+ \nu$, using $|V_{ub}| = (4.31 \pm 0.30) \times 10^{-3}$ [7] and $f_{B^+} = 0.216 \pm 0.022$ GeV [8] in Eq. (1) is $(1.6 \pm 0.4) \times 10^{-4}$. In a previously published analysis using a sample of $223 \times 10^6$ $Y(4S)$ decays, the BABAR collaboration set an upper limit of $\mathcal{B}(B^+ \to \tau^+ \nu) < 2.6 \times 10^{-4}$ at the 90% confidence level (CL) [9]. The Belle Collaboration has reported evidence from a search for this channel using $449 \times 10^6$ $B$ meson pairs where the branching fraction was measured to be $\mathcal{B}(B^+ \to \tau^+ \nu) = (1.79^{+0.58}_{-0.45})(\text{stat.})^{+0.48}_{-0.38}(\text{syst.}) \times 10^{-4}$ [10].

II. THE BABAR DETECTOR AND DATA SET

The data used in this analysis were collected with the BABAR detector [11] at the PEP-II storage ring. The sample corresponds to an integrated luminosity of $346 \text{ fb}^{-1}$ at the $Y(4S)$ resonance (on-resonance) and $36.3 \text{ fb}^{-1}$ taken at 40 MeV below the $B \bar{B}$ production threshold (off-resonance) which is used to study background from $e^+ e^- \to f \bar{f}$ ($f = u, d, s, c, \tau$) continuum events. The on-resonance sample contains $(383 \pm 4) \times 10^6$ $Y(4S)$ decays. The detector components used in this analysis are the tracking system composed of a five-layer silicon vertex detector and a 40-layer drift chamber (DCH), the Cherenkov detector for charged $\pi^- K$ discrimination, a CsI calorimeter (EMC) for photon and electron identification, and an 18-layer flux return (IFR) located outside of the 1.5 T solenoidal coil and instrumented with resistive plate chambers for muon and neutral hadron identification. For the most recent $133 \text{ fb}^{-1}$ of data, a portion of the resistive plate chambers has been replaced with limited streamer tubes. We analyze the data from several data-taking periods separately to account for varying accelerator and detector conditions.

A GEANT4-based [12] Monte Carlo (MC) simulation is used to model signal efficiencies and physics backgrounds. The $\tau$ lepton decay is modeled using EvtGen [13]. Beam-related background and detector noise from data are overlaid on the simulated events. Simulation samples equivalent to approximately 3 times the accumulated data are used to model $B \bar{B}$ events, and samples equivalent to approximately 1.5 times the accumulated data are used to model underlying continuum events. We determine selection efficiencies for signal events using a MC simulation where one $B^+$ meson decays to $\tau^+ \nu$, while the other is allowed to decay into any final state.

III. ANALYSIS METHOD

Because of the presence of multiple neutrinos, the $B^+ \to \tau^+ \nu$ decay mode lacks the kinematic constraints which are usually exploited in $B$ decay searches in order to reject both continuum and $B \bar{B}$ backgrounds. The strategy adopted for this analysis is to reconstruct exclusively the decay of one of the $B$ mesons in the event, referred to as the
“tag” $B$. The remaining particle(s) in the event (the “recoil”) are assumed to come from the other $B$ and are compared with the signature expected for $B^+ \rightarrow \tau^+ \nu$. In order to avoid experimenter bias, the signal region in data is blinded until the final yield extraction is performed.

The tag $B$ is reconstructed in the set of semileptonic $B$ decay modes $B^- \rightarrow D^0 \ell^- \bar{\nu}_\ell X$, where $\ell$ denotes either electron or muon, and $X$ can be either nothing or a transition particle ($\pi^0$ or photon) from a higher mass charm state decay which we do not attempt to explicitly include in the tag $B$. However, we explicitly veto events where the best tag candidate is consistent with neutral $B$ decay, where the $X$ system is a single charged pion that can be combined with the $D^0$ to form a $D^{\pm}$ candidate.

The $B^+ \rightarrow \tau^+ \nu$ signal is searched for in both leptonic and hadronic $\tau$ decay modes constituting approximately 71% of the total $\tau$ decay width: $\tau^+ \rightarrow e^+ \nu \bar{\nu}$, $\tau^+ \rightarrow \mu^+ \nu \bar{\nu}$, $\tau^+ \rightarrow \pi^+ \nu \bar{\nu}$, and $\tau^+ \rightarrow \pi^0 \nu \bar{\nu}$. We do not consider the $\tau^+ \rightarrow \pi^+ \pi^- \pi^0 \bar{\nu}$ mode since we found it to be dominated by background events.

### A. Tag $B$ reconstruction

$D^0 \ell$ candidates are reconstructed by combining a $D^0$ with an identified electron or muon with momentum above 0.8 GeV/c in the $e^+ e^-$ center-of-mass (CM) frame (Fig. 1). The flight direction of the $D^0$ is required to intersect with the lepton track. Assuming that the massless neutrino is the only missing particle, we calculate the cosine of the angle between the $D^0 \ell$ candidate and the $B$ meson,

$$\cos \theta_{B-D^0\ell} = \frac{2E_B E_{D^0\ell} - m_B^2 - m_{D^0\ell}^2}{2|\vec{p}_B||\vec{p}_{D^0\ell}|},$$

where $(E_{D^0\ell}, \vec{p}_{D^0\ell})$ and $(E_B, \vec{p}_B)$ are the four-momenta of the $D^0 \ell$ and $B$ in the CM frame, and $m_{D^0\ell}$ and $m_B$ are the masses of the $D^0 \ell$ candidate and $B^+ \nu$ meson (the nominal mass [7] is used), respectively. $E_B$ and the magnitude of $|\vec{p}_B|$ are calculated from the beam energy: $E_B = E_{CM}/2$, where $E_{CM}$ is the CM energy of the beams, and $|\vec{p}_B|$ = \sqrt{E^2 - m_B^2}. Correctly reconstructed candidates populate $\cos \theta_{B-D^0\ell}$ in the range of $[-1, 1]$, whereas combinatorial backgrounds can take unphysical values outside this range. We retain events in the interval $-2.0 < \cos \theta_{B-D^0\ell} < 1.1$, where the upper bound takes into account the detector resolution and the less restrictive lower bound accepts those events where the $X$ is a soft transition particle from a higher mass charm state. Due to the semiexclusive nature of the tag $B$ reconstruction the $\cos \theta_{B-D^0\ell}$ distribution differs slightly from that measured in exclusive semileptonic $B$ decays.

![FIG. 1. $D^0$ invariant mass for tag $B$ candidates containing an (a) electron or (b) muon and the CM momentum of the tag $B$ lepton for tag $B$ candidates containing an (c) electron or (d) muon. On-resonance data (filled circles) are overlaid on the sum of $B \bar{B}$ MC (solid histogram) and nonresonant background MC (gray histogram), both of which have been normalized to the integrated data luminosity. Off-resonance data (open diamonds) are overlaid for comparison, and normalized to the on-resonance integrated luminosity.](image-url)
We reconstruct the $D^0$ candidates in four decay modes: $K^-\pi^+$, $K^-\pi^+\pi^-\pi^+$, $K^-\pi^+\pi^0\pi^0$, and $K^0_S\pi^+\pi^-$, only considering $K^0_S$ candidates decaying to charged pions. The charged tracks are required to meet particle identification criteria consistent with the particle hypothesis and are required to converge at a common vertex. The $\pi^0$ candidates are required to have invariant masses between 0.115 and 0.150 GeV/$c^2$ and the photon daughter candidates of the $\pi^0$ must have a minimum laboratory energy of 30 MeV and have shower shapes consistent with electromagnetic showers. The mass of the reconstructed $D^0$ candidates (Fig. 1) in the $K^-\pi^+, K^-\pi^+\pi^-\pi^+$, and $K^0_S\pi^+\pi^-$ modes is required to be within 20 MeV/$c^2$ of the nominal mass [7], while in the $K^-\pi^+\pi^0$ decay mode the mass is required to be within 35 MeV/$c^2$ of the nominal mass. These constraints are determined by fitting a single Gaussian function and a first-order polynomial to the mass distribution in signal MC and correspond to the 3σ positions on the Gaussian. Furthermore, the sum of the charges of all the particles in the event must be equal to zero. If more than one suitable $D^0\ell$ candidate can be reconstructed, the best candidate is taken to be the one with the largest probability of originating from a single vertex.

The tag reconstruction efficiency, including all $B$ and $D$ branching ratios, extracted from signal MC and averaged over all data-taking periods, is $(6.64 \pm 0.03) \times 10^{-3}$, where the error is due to the statistics of the signal MC sample. This corresponds to a tag $B$ yield of $(2.54 \pm 0.03) \times 10^6$ At this level of selection, we find that the MC models the data well in the electron channel of the semileptonic $B$ decay, but less so in the muon channel. The disagreement in the muon channel appears to derive largely from the continuum background and therefore should not affect the real semileptonic tags. The tag reconstruction efficiency is corrected for any data/MC disagreement using a control sample described in Sec. III D.

B. Selection of $B^+ \rightarrow \tau^+ \nu$ signal candidates

After the tag $B$ reconstruction, the recoil is studied for consistency with the signal modes. All selection criteria are optimized for each of the different signal $\tau$ decay modes. The optimization is performed by maximizing the signal significance, $s/\sqrt{s + b}$, for each channel using the signal $(s)$ and background $(b)$ MC and assuming a total branching fraction for $B^+ \rightarrow \tau^+ \nu$ of $1.0 \times 10^{-4}$, using the PRIM algorithm [14]. This algorithm simultaneously optimizes selection criteria over a number of variables by relaxing and tightening the constraints on all variables until a maximal significance is achieved, allowing only up to a fixed percentage of signal and background to be removed or restored with each iteration of the selection criteria.

All signal modes contain one charged particle that is identified as either an electron, muon, or pion using standard particle identification techniques. Both the $\tau^+ \rightarrow \pi^+ \bar{\nu}$ and the $\tau^+ \rightarrow \pi^+ \pi^0 \bar{\nu}$ modes contain a pion signal track. The signal track is required to have at least 12 hits in the DCH; its momentum transverse to the beam axis, $p_T$, is required to be greater than 0.1 GeV/$c$, and its point of closest approach to the interaction point must be less than 10.0 cm along the beam axis and less than 1.5 cm transverse to it. We demand the invariant mass of the signal $\pi^0$ be between 0.115 and 0.150 GeV/$c^2$. The daughter photon candidates must have a minimum energy of 50 MeV, and their shower shapes are required to be consistent with electromagnetic showers.

Background consists primarily of $B^+ B^-$ events in which the tag $B$ meson has been correctly reconstructed and the recoil contains one track and additional particles which are not reconstructed by the tracking detectors or calorimeter. These events typically contain one or more $K^0_L$ mesons, neutrinos, or particles that pass outside of the detector acceptance. $B^0\bar{B}^0$ and continuum events contribute background primarily to hadronic $\tau$ decay modes. In addition, some excess events in data, most likely from higher-order QED processes (such as two-photon fusion) that are not modeled in our MC simulation, are observed.

Backgrounds are suppressed relative to signal by imposing requirements on the kinematic and shape properties of the events. The missing mass is calculated as

$$M_{\text{miss}} = \sqrt{(E_{Y(4S)} - E_{\text{vis}})^2 - (\vec{p}_{Y(4S)} - \vec{p}_{\text{vis}})^2}. \quad (3)$$

Here $(E_{Y(4S)}, \vec{p}_{Y(4S)})$ is the four-momentum of the $Y(4S)$, known from the beam energies. The quantities $E_{\text{vis}}$ and $\vec{p}_{\text{vis}}$ are the total visible energy and momentum of the event, which are calculated by adding the energy and momenta, respectively, of all the reconstructed tracks and photons in the event. Continuum background is suppressed with two variables: the cosine of the angle between the signal candidate and the tag candidate thrust vectors (in the CM frame), $\cos\theta_T$, and the minimum invariant mass constructible from any three tracks in an event, $M_{\text{min}}^3$. For the background, the cosine of the thrust angle peaks at $\pm 1$, while the minimum invariant mass peaks strongly below 1.5 GeV/$c^2$, as can be seen in Fig. 2, where the signal and $\tau^+ \tau^-$ background MC are shown. We combine these two variables into a single quantity for use in the selection optimization algorithm. The projection uses the following empirically derived equation:

$$R_{\text{cont}} \equiv \sqrt{(3.7 - |\cos\theta_T|)^2 + (M_{\text{min}}^3/(\text{GeV}/c^2) - 0.75)^2}. \quad (4)$$

Applying selection criteria to $R_{\text{cont}}$ primarily removes background from $e^+ e^- \rightarrow \tau^+ \tau^-$, but also suppresses other continuum backgrounds. Since the $\tau^+ \rightarrow \pi^+ \pi^0 \bar{\nu}$ decay proceeds via an intermediate resonance ($\rho^+ \rightarrow \pi^+ \pi^0$), further background rejection can be achieved by applying requirements on the intermediate meson candidate. In events with more than one recoil $\pi^0$, the candidate with invariant mass closest to the nominal $\pi^0$ mass [7] is
chosen. The invariant mass of the reconstructed $\pi^+\pi^0$ signal candidates are required to lie between 0.64 and 0.86 GeV/$c^2$. A quantity analogous to $\cos\theta_{B-\tau^+\nu}$, as defined in Sec. III A, can be calculated for $\tau^+ \rightarrow \pi^+\pi^0\tilde{\nu}$ by replacing the $B$ with a $\tau$ and the $D^0\ell$ with $\pi^+\pi^0$ in Eq. (2). The analogous quantities of $|p_\ell|$ and $E_\ell$ are calculated assuming the $\tau$ is from the $B^+ \rightarrow \tau^+\nu$ decay and that the $B^+$ is almost at rest in the CM frame. We require $\cos\theta_{j-s}\sim0 < 0.87$.

We demand that there are no $K_L^0$ candidates reconstructed in the IFR. For the $\tau^+ \rightarrow \pi^+\tilde{\nu}$ channel, we demand that there are fewer than two candidate clusters in the EMC consistent with being deposited by a $K_L^0$. In the leptonic final states we demand that there are two or fewer $\pi^0$ candidates. For the $\tau^+ \rightarrow e^+\nu\tilde{\nu}$ mode, we reject events where a photon conversion creates the electron by requiring that the invariant mass of the signal and tag $B$ lepton pair be greater than 0.1 GeV/$c^2$. We impose mode-dependent selection criteria on the total momentum ($p_{\text{signal}}$) of the visible decay products of the $\tau$ candidate.

We further separate signal and background by exploiting the remaining energy ($E_{\text{extra}}$), calculated by summing the CM energy of the neutral clusters (with a minimum of 20 MeV in the laboratory frame) and tracks that are not associated with either the tag $B$ or the signal. Signal events tend to peak at low $E_{\text{extra}}$ values whereas background events, which tend to contain additional sources of neutral clusters, are distributed toward higher $E_{\text{extra}}$ values. The selection applied to $E_{\text{extra}}$ is optimized for the best signal significance, assuming the branching fraction is $1 \times 10^{-4}$ and was blinded for $E_{\text{extra}} < 0.5$ GeV in on-resonance data until the selection was finalized.

The signal selection efficiencies for the $\tau$ decay modes are determined from signal MC simulation and summarized in Table I. The signal efficiencies correspond to the fraction of events selected in a specific signal decay mode, given that a tag $B$ has been reconstructed. Signal selection efficiencies are further corrected by applying the factors provided in Table IV which are explained in later sections.

### C. Background estimation from $E_{\text{extra}}$ sidebands

We estimate our background from the data by studying events in a sideband region of $E_{\text{extra}}$. We define the sideband (sb) region as $E_{\text{extra}} > 0.5$ GeV, and also define signal regions (sig) in $E_{\text{extra}}$ using the appropriate signal mode-dependent selection. Any bias due to the signal tail extending into the sideband region is assumed to be negligible. After applying all other selection criteria, we compute from the background MC simulation the ratio of events in the sideband ($N_{\text{MC,sb}}$) and signal ($N_{\text{MC,sig}}$) regions,

$$R_{\text{MC}} = \frac{N_{\text{MC,sig}}}{N_{\text{MC,sb}}}.$$  (5)

Using the number of data events in the sideband ($N_{\text{data,sb}}$) and the ratio $R_{\text{MC}}$, the number of expected background events in the signal region in data ($N_{\text{exp,sig}}$) is estimated,

$$N_{\text{exp,sig}} = N_{\text{data,sb}} \cdot R_{\text{MC}}.$$  (6)

The sideband background projection (Table II) is taken as the number of expected background events.

The background estimate is validated by performing a similar test using sidebands in the $D^0$ mass distribution. We select events using $D^0$ mass sidebands between 4$\sigma$ and 9$\sigma$
above and below the nominal $D^0$ mass, with all other signal selection criteria applied. Candidates in these regions of the $D^0$ mass distribution are random combinations of kaons and pions, and represent a pure combinatoric background. We average the yields from the upper and lower sidebands and scale this using the ratio of the $D^0$ mass sideband and signal region. This yields a $D^0$ mass combinatoric background estimate in the $D^0$ mass signal region for both data ($N_{\text{comb}}^{\text{data}}$) and MC ($N_{\text{comb}}^{\text{MC}}$). The remaining component, in the MC, of the background which contains real $D^0$ mesons in the tag is then computed,

$$N_{\text{peak}}^{\text{MC}} = N_{\text{total}}^{\text{MC}} - N_{\text{comb}}^{\text{MC}},$$

and added to the combinatoric component (determined from data) to obtain an effective estimate of the total background,

$$N_{\text{total}}^{\text{predicted}} = N_{\text{peak}}^{\text{MC}} + N_{\text{comb}}^{\text{data}}.$$  \hfill (7)

This is done for each reconstructed signal decay channel. The method assumes that the background in the $E_{\text{extra}}$ signal region can be modeled by the combinatoric component of the $D^0$ mass distribution, taken from data, and the peaking component of the $D^0$ mass distribution, taken from MC simulations. Since it uses the $D^0$ mass sidebands, it is also statistically independent from the $E_{\text{extra}}$ sideband calculation.

We find very good agreement between the background prediction using the $D^0$ mass sidebands and that obtained from the projection of the $E_{\text{extra}}$ sideband. This agreement is demonstrated in Table II and validates our background estimation method.

| Signal mode | $e^+$ | $\mu^+$ | $\pi^+$ | $\pi^+ \pi^0$ |
|-------------|-------|---------|---------|-------------|
| $E_{\text{extra}}$ sideband | $44.3 \pm 5.2$ | $39.8 \pm 4.4$ | $120.3 \pm 10.2$ | $17.3 \pm 3.3$ |
| $D^0$ sideband | $44.2 \pm 6.4$ | $42.8 \pm 6.0$ | $113.4 \pm 11.6$ | $16.3 \pm 4.5$ |

D. Correction of tag $B$ yield and $E_{\text{extra}}$ simulation

The tag $B$ yield and $E_{\text{extra}}$ distribution in signal and background MC simulation are validated using control samples. These samples are further used to define corrections to efficiencies of selection criteria. “Double-tagged” events, for which both of the $B$ mesons are reconstructed in tagging modes, $B^- \rightarrow D^0 \ell^- \bar{\nu}_\ell X$ vs $B^+ \rightarrow D^0 \ell^+ \nu_\ell X$ are used as the primary control sample. “Single-tagged” events are also used where one $B$ decays via $B^- \rightarrow D^0 \ell^- \bar{\nu}_\ell X$ and the other $B$ decay is not constrained. The double-tagged sample is almost entirely free of continuum events.

We select double-tagged events by requiring that the two semileptonic $B$ candidates have opposite charge and do not share any particles. We also require that there are no additional tracks in the event. If there are more than two such independent tag $B$ candidates in the event then the two best candidates are selected as those with the largest probabilities of each originating from a common origin. The $D^0$ meson invariant mass is shown in Fig. 3 for the second tag in all double-tagged events.

We initially determine the tag efficiency using a signal MC where one of the two $B$ mesons always decays into a generic final state and the other always decays into a $\tau^+ \nu$ final state. We estimate the correction to the MC semileptonic tag efficiency by comparing the number of single- and double-tagged events in data and MC. We calculate the ratio of double-tagged to single-tagged events, and we use the ratio of this quantity from data and MC as a correction factor for the tag $B$ yield.

We determine the number of single-tagged events by subtracting the combinatoric component under the $D^0$ mass peak in events where one $B$ is tagged and the second is allowed to decay without constraint (Fig. 1). We determine this component by using $D^0$ mass sidebands between $4\sigma$ and $7\sigma$ above and below the nominal $D^0$ mass. A narrower sideband region is used for this study than in the background estimate validation due to the comparative flatness of the sidebands in this region and the large statistics available at this early stage of selection. We then average the yields from these combinatoric $D^0$ mass regions and scale by the ratio of the sideband and signal region widths. We perform this subtraction using events where the $D^0$ meson from one of the semileptonic tags is reconstructed as $D^0 \rightarrow K^- \pi^+$ and the second tag decays into any of our allowed final states. The resulting single-

![Graph of D^0 Candidate Mass](image)

FIG. 3. $D^0$ invariant mass from the recoil $B$ meson in double-tagged events. On-resonance data (black circles) are overlaid on the combined $B \bar{B}$ (solid histogram) and continuum (gray histogram) MC samples normalized to the data luminosity.
We additionally check the modeling of $E_{\text{extra}}$ by comparing samples of events where the signal and tag $B$ candidates are of the same sign. We find that for all signal modes, there is good agreement between the shape of the $E_{\text{extra}}$ from the background prediction and the data in this wrong-charge sample. In the pion channel, in particular, we find the data yield is higher than predicted from MC. This suggests that for a pure background sample, with a topology similar to that of signal, the $E_{\text{extra}}$ distribution is well modeled but the background estimate cannot be taken directly from the MC background simulation. This further validates our choice to take the background estimate from the $E_{\text{extra}}$ sideband in data and the signal-to-sideband ratio in MC simulation.

IV. STUDIES OF SYSTEMATIC UNCERTAINTIES

The main sources of uncertainty in the determination of the $B^+ \rightarrow \tau^+ \nu$ branching fraction are the tag reconstruction efficiency ($e_{\text{tag}}$), the efficiency of each signal mode ($e_i$), and the number of expected background events in the signal region for each signal mode.

An uncertainty of 1.1% enters the branching fraction calculation from the estimation of the number of $B^+ B^-$ events present in the data sample [15]. The tagging efficiency and yield in signal MC is corrected using the double-tagged and single-tagged samples. The tag $B$ yield systematic uncertainty is 3.6%, with a correction factor to the yield of 1.05. The systematic uncertainties on the signal efficiency depend on the $\tau$ decay mode and include effects such as the tracking of charged particles, particle identification, and the modeling of $\pi^0$ mesons.

The systematic uncertainty on the signal efficiency due to the mismodeling of the $E_{\text{extra}}$ variable is extracted using the double-tagged events. We extract the yield of candidates satisfying $E_{\text{extra}} < 0.5$ GeV. This yield is then compared to the number of candidates in the full sample. Comparing the ratio extracted from MC to that extracted from data yields a correction factor, the error of which is taken as the systematic uncertainty for $E_{\text{extra}}$. The systematic uncertainty for $E_{\text{extra}}$ is 3.4% with a correction of 0.99.

The systematic uncertainty on the modeling of $K^0$ candidates is extracted using the double-tagged events, similar to the method used for the $E_{\text{extra}}$ systematic evaluation. We quantify the data/MC comparison by extracting the yield with a cut demanding exactly zero (less than two) reconstructed IFR (EMC) measured $K^0$ candidates remaining, and extracting the yield with a sample where any number of $K^0$ candidates remain, and take the ratio of ratios from the MC and data. The systematic uncertainty for vetoing IFR (EMC) $K^0$ candidates is 3.3% (3.8%), with a correction factor on the efficiency of 0.99 (0.97).

A breakdown of the contributions to the systematic uncertainty for each signal mode is given in Table IV. We find that the most significant individual effects on the signal efficiency are from the modeling of the $E_{\text{extra}}$ and...
the $K^0_L$ vetoes. The uncertainties on each mode are combined by weighting them by the corrected efficiency for a given mode, using the efficiencies from Table I multiplied by the correction factors given in Table IV. The signal-mode-specific systematic uncertainties are summed in quadrature and then the sum is added linearly with the IFR $K^0_L$ and $E_{\text{extra}}$ uncertainties, which are correlated among the modes. The resulting overall systematic uncertainty on the signal efficiency is then added in quadrature with the uncertainties on the tag reconstruction and the number of $B\bar{B}$ pairs in the sample to give a total uncertainty of 6.6%.

### V. RESULTS

After finalizing the signal selection criteria, we measure the yield of events in each channel in the signal region of the on-resonance data. Table V lists the number of observed events in on-resonance data in the signal region, together with the expected number of background events in the signal region (taken from the $E_{\text{extra}}$ sideband prediction from Table II). Figure 5 shows the $E_{\text{extra}}$ distribution for all data and MC in the signal region, with signal MC shown for comparison. Figure 6 shows the $E_{\text{extra}}$ distribution separately for each of the signal modes.

We determine the $B^+ \rightarrow \tau^+ \nu$ branching fraction from the number of signal candidates $s_i$ in data for each $\tau$ decay mode, according to $s_i = N_{B\bar{B}} B(B^+ \rightarrow \tau^+ \nu) e_{\text{tag}} e_{\ell} \epsilon_i$, where $N_{B\bar{B}}$ is the total number of $B\bar{B}$ pairs in data. The results from each of our four signal decay channels ($n_{ch}$) are combined using the estimator $Q = \mathcal{L}(s + b)/\mathcal{L}(b)$, where $\mathcal{L}(s + b)$ and $\mathcal{L}(b)$ are the likelihood functions for signal plus background and background-only hypotheses, respectively:

$$\mathcal{L}(s + b) = \prod_{i=1}^{n_{ch}} \frac{e^{-(s_i + b_i)}(s_i + b_i)^{n_i}}{n_i!},$$

$$\mathcal{L}(b) = \prod_{i=1}^{n_{ch}} e^{-b_i} b_i^{n_i}/n_i!.$$  

We include the systematic uncertainties, including those of a statistical nature, on the expected background ($b_i$) in the likelihood definition by convolving it with a Gaussian function. The mean of the Gaussian is $b_i$, and the standard deviation ($\sigma_{b_i}$) of the Gaussian is the error on $b_i$ [16].

We calculate the branching fraction central value (including statistical uncertainty and uncertainty from the background) by scanning over signal branching fraction...
hypotheses between 0.0 and $3.0 \times 10^{-4}$ in steps of $0.025 \times 10^{-4}$ and computing the value of $\mathcal{L}(s + b)/\mathcal{L}(b)$ for each hypothesis [Fig. 7(a)]. The branching fraction is the hypothesis which minimizes $-2 \log(\mathcal{L}(s + b)/\mathcal{L}(b))$, and the statistical uncertainty is determined by finding the points on the likelihood scan that occur at one unit above the minimum. The systematic error is determined as detailed in Sec. IV and computed for the branching fraction as a fraction of the central value.

The upper limit at the 90% CL, including both statistical and systematic uncertainties, is determined by generating 5000 experiments for each of the aforementioned signal branching fraction hypotheses. Each generated experiment also includes the expected number of background events, and varies the generated number of background which will vary around the input hypotheses $s$ and $b$ according to the above procedures.

We determine the confidence level of a given signal hypothesis by finding the probability that the value of the estimator $Q$ in experiments generated according to a given composition (signal and background, $Q_{s+b}$, or only background, $Q_b$) is less than that observed in data ($Q_{\text{obs}}$). The 90% CL limit is determined by using the CLs method [17], in which we determine the signal hypothesis for which $P(Q_{s+b} < Q_{\text{obs}})/P(Q_b < Q_{\text{obs}}) = 1 - 0.90$, where $P(Q_{s+b} < Q_{\text{obs}}) (P(Q_b < Q_{\text{obs}}))$ is the probability that experiments generated assuming a given $s + b (b)$ hypothesis have a likelihood ratio lower than that observed in data [Fig. 7(b)].

We determine the branching fraction central value to be

$$B(B^+ \rightarrow \tau^+ \nu) = (0.9 \pm 0.6\text{(stat.)} \pm 0.1\text{(syst.)}) \times 10^{-4}$$

and set an upper limit at the 90% CL of

FIG. 6. $E_{\text{extra}}$ distribution after all selection criteria for (a) $\tau^+ \rightarrow e^+ \nu \bar{\nu}$, (b) $\tau^+ \rightarrow \mu^+ \nu \bar{\nu}$, (c) $\tau^+ \rightarrow \pi^+ \nu \bar{\nu}$, and (d) $\tau^+ \rightarrow \pi^+ \pi^0 \nu$. Background MC (solid histogram) has been normalized to the luminosity of the on-resonance data (black dots), and then additionally scaled according to the ratio of predicted background from data and MC as presented in Sec. III C. $B^+ \rightarrow \tau^+ \nu$ signal MC (dotted histogram) is normalized to a branching fraction of $10^{-3}$ and shown for comparison.
of 2 standard deviations. We interpret this result in the context of the standard model. Using the central value for \( \mathcal{B}(B^+ \rightarrow \tau^+ \nu) \) and taking the known values of \( G_F, m_B, m_{\tau} \), and \( \tau_B \) [7] we calculate, from Eq. (1), \( f_B \cdot |V_{ub}| = (7.5^{+2.0}_{-2.8} \text{ (stat.)} \pm 0.2(\text{syst.})) \times 10^{-4} \text{ GeV} \), where the uncertainties are non-Gaussian. Using the value of \( |V_{ub}| \) from [7] we extract \( f_B = 0.167^{+0.048}_{-0.066} \text{ GeV} \), where the uncertainty is dominated by the statistical uncertainty on the branching fraction central value.

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 U.S. Department of Energy and National Science Foundation, the Natural Sciences and Engineering Research Council (Canada), Institute of High Energy Physics (China), the Commissariat à l’Énergie 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 Science and Technology of the Russian Federation, Ministerio de Educación y Ciencia (Spain), and the Particle Physics and Astronomy Research Council (United Kingdom). Individuals have received support from the Marie-Curie IEF program (European Union) and the A. P. Sloan Foundation.

[1] Charge-conjugate modes are implied throughout the paper.
[2] N. Cabbibo, Phys. Rev. Lett. 10, 531 (1963).
[3] M. Kobayashi and T. Masakawa, Prog. Theor. Phys. 49, 1652 (1973).
[4] W. S. Hou, Phys. Rev. D 48, 2342 (1993).
[5] G. Isidori and P. Paradisi, Phys. Lett. B 639, 499 (2006).
[6] A. G. Akeroyd and C. H. Chen, Phys. Rev. D 75, 075004 (2007).
[7] W. M. Yao et al. (Particle Data Group), J. Phys. G 33, 1 (2006).
[8] A. Gray et al. (HPQCD Collaboration), Phys. Rev. Lett. 95, 212001 (2005).
[9] B. Aubert et al. (BABAR Collaboration), Phys. Rev. D 73, 057101 (2006).
[10] K. Ikado et al. (Belle Collaboration), Phys. Rev. Lett. 97, 251802 (2006).
[11] B. Aubert et al. (BABAR Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A 479, 1 (2002).
[12] S. Agostinelli et al. (GEANT4 Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A 506, 250 (2003).
[13] D. J. Lange, Nucl. Instrum. Methods Phys. Res., Sect. A 462, 152 (2001).
[14] J. Friedman and N. Fisher, Stat. Comput. 9, 123 (1999).
[15] B. Aubert et al. (BABAR Collaboration), Phys. Rev. D 67, 032002 (2003).
[16] L. Lista, Nucl. Instrum. Methods Phys. Res., Sect. A 517, 360 (2004).
[17] A. L. Read, J. Phys. G 28, 2693 (2002).