Search for $B^+ \rightarrow K^+ \tau^+\tau^-$ at the BaBar experiment


(The BaBar Collaboration)

1Laboratoire d’Annecy-le-Vieux de Physique des Particules (LAPP), Université de Savoie, CNRS/IN2P3, F-74941 Annecy-le-Vieux, France
2Universitat de Barcelona, Facultat de Fisica, Departament ECM, E-08028 Barcelona, Spain
3INFN Sezione di Bari e Dipartimento di Fisica, Università di Bari, I-70126 Bari, Italy
4University of Bergen, Institute of Physics, N-5007 Bergen, Norway
5Lawrence Berkeley National Laboratory and University of California, Berkeley, California 94720, USA
6Ruhr Universität Bochum, Institut für Experimentalphysik 1, D-44780 Bochum, Germany
7University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z1
8Budker Institute of Nuclear Physics SB RAS, Novosibirsk 630090, Novosibirsk State University, Novosibirsk 630090, Novosibirsk State Technical University, Novosibirsk 630092, Russia
9University of California at Irvine, Irvine, California 92697, USA
10University of California at Riverside, Riverside, California 92521, USA
11University of California at Santa Cruz, Institute for Particle Physics, Santa Cruz, California 95064, USA
12California Institute of Technology, Pasadena, California 91125, USA
We search for the rare flavor-changing neutral current process $B^+ \rightarrow K^+ \tau^+ \tau^-$ using data from the BABAR experiment. The data sample, collected at the center-of-mass energy of the $\Upsilon(4S)$ resonance, corresponds to a total integrated luminosity of 424 fb$^{-1}$ and to 471 million $B\bar{B}$ pairs. We reconstruct one $B$ meson, produced in the $\Upsilon(4S) \rightarrow B\bar{B}$ decay, in one of many hadronic decay
The flavor-changing neutral current process \( B^+ \rightarrow K^+ \tau^+ \tau^- \) [1] is highly suppressed in the standard model (SM), with a predicted branching fraction in the range \( 1 - 2 \times 10^{-7} \) [2, 3]. This decay is forbidden at tree level and only occurs, at lowest order, via one-loop diagrams. The SM contributions, shown in Fig. 1, include the electromagnetic penguin, the \( Z \) penguin, and the \( W^+W^- \) box diagrams. Rare semi-leptonic \( B \) decays such as \( B^+ \rightarrow K^+ \tau^+ \tau^- \) can provide a stringent test of the SM and a fertile ground for new physics searches. Virtual particles can enter in the loop and thus allow to probe, at relatively low energies, new physics at large mass scales. Measurements of the related decays, \( B^+ \rightarrow K^+ \ell^+ \ell^- \) where \( \ell = e \) or \( \mu \), have been previously published by \( \bar{B}\bar{A}\bar{B} \) [4] and other experiments [5]-[8], and exhibit some discrepancy with the SM expectation [9].

The decay \( B^+ \rightarrow K^+ \tau^+ \tau^- \) is the third family equivalent of \( B^+ \rightarrow K^+ \ell^+ \ell^- \) and hence may provide additional sensitivity to new physics due to third-generation couplings and the large mass of the \( \tau \) lepton [10]. An important potential contribution to this decay is from neutral Higgs boson couplings, where the lepton-lepton-Higgs vertices are proportional to the mass squared of the lepton [11]. Thus, in the case of the \( \tau \), such contributions can be significant and could alter the total decay rate. Additional sources of new physics and their effect on the \( B^+ \rightarrow K^+ \tau^+ \tau^- \) branching fraction and the kinematic distributions of the \( \tau^+ \tau^- \) pair are also discussed in Refs. [12]-[24]. These new physics scenarios do not necessarily have the same impact on the \( B^+ \rightarrow K^+ \psi(2S), \psi(2S) \rightarrow \tau^+ \tau^- \) decay, and thus the latter will only be considered if a visible signal is present.

We report herein a search for \( B^+ \rightarrow K^+ \tau^+ \tau^- \) with data recorded by the \( \bar{B}\bar{A}\bar{B} \) detector [25] at the \( e^+e^- \) PEP-II collider at the SLAC National Accelerator Laboratory. This search is based on 424 fb\(^{-1}\) of data [26] collected at the center-of-mass (CM) energy of the \( \Upsilon(4S) \) resonance, where \( \Upsilon(4S) \) decays into a \( B\bar{B} \) pair. We use hadronic \( B \) meson tagging techniques, where one of the two \( B \) mesons, referred to as the \( B_{\text{tag}} \), is reconstructed exclusively via its decay into one of several hadronic decay modes. The remaining tracks, clusters, and missing energy in the event are attributed to the signal \( B \), denoted as \( B_{\text{sig}} \), on which the search for \( B^+ \rightarrow K^+ \tau^+ \tau^- \) is performed. We consider only leptonic decays of the \( \tau \): \( \tau^+ \rightarrow e^+\nu_\tau \tau^- \) and \( \tau^+ \rightarrow \mu^+\nu_\tau \overline{\nu}_\tau \), which results in three signal decay topologies with a charged \( K \), multiple missing neutrinos, and either \( e^+\nu_e \), \( \mu^+\nu_\mu \) or \( e^+\mu^- \) in the final state. The neutrinos are accounted for as missing energy in any signal event where a charged kaon and lepton pair are identified and extra neutral activity, including \( \pi^0 \) candidates, is excluded.

Simulated Monte Carlo (MC) signal and background events, generated with EvtGen [27], are used to develop signal selection criteria and to study potential backgrounds. The detector response is simulated using GEANT4 [28]. Signal MC events are generated as \( \Upsilon(4S) \rightarrow B^+B^- \), where one \( B \) decays according to its measured SM branching fractions [29] and the other \( B \) decays via \( B^+ \rightarrow K^+ \tau^+ \tau^- \) according to the model described in Ref. [30]. Within this model, referred to as LCSR, a light-cone sum rule approach is used to determine the form factors that enter into the parameterization of the matrix elements describing this decay. Signal events are also reweighted to a model based on the unquenched lattice QCD calculations of the \( B \rightarrow K\ell\ell \) form factors [2] for the determination of the signal efficiency, and the two theoretical approaches are then compared to evaluate the model-dependence of our measurement. Because of the low efficiency of the hadronic \( B_{\text{tag}} \) reconstruction, “dedicated” signal MC samples are also generated for this analysis, where one \( B \) decays exclusively through \( B^{\pm} \rightarrow D^0\pi^\pm, D^0 \rightarrow K^-\pi^+ \) while the other \( B \) meson decays via the signal channel. This ensures that more events pass the hadronic \( B_{\text{tag}} \) reconstruction and allows for increased statistics in the distributions of discriminating variables in the signal sample. Only variables that are independent of the \( B_{\text{tag}} \) decay mode are considered with the dedicated signal MC sample. To avoid potential bias, this dedicated sample is not used to evaluate the final signal selection efficiency. Back-
ground MC samples consist of $B^+B^-$ and $B^0\bar{B}^0$ decays and continuum events, $e^+e^-\to f\bar{f}$, where $f$ is a lepton or a quark. The $B\bar{B}$ and $e^+e^-\to e\bar{e}$ MC-simulated samples are produced with an integrated luminosity ten times that of data, whereas the remaining continuum samples have an integrated luminosity that is four times larger.

The signal selection of $B^+\to K^+\tau^+\tau^-$ events is preceded by the full hadronic reconstruction of the $B_{\text{tag}}$ meson, via $B\to SX$ [31]. Here, $S$ is a seed meson, $D^{*(+)0}$, $D^{*(+)±}$, $D_s^{±}$ or $J/\psi$, and $X$ is a combination of at most five charged or neutral kaons and pions with at most two neutral $\pi^0$ or $K_S^0$ candidates. The $D$ seeds are reconstructed in the decay modes $D^+\to K_S^0\pi^+$, $K^0_S\pi^+\pi^0$, $K^0_S\pi^+\pi^-\pi^0$, $K^+\pi^+$, $K^+K^-\pi^+\pi^0$; $D^0\to K^0\pi^+$, $K^-\pi^+$, $K^-\pi^+\pi^0$, $K^-\pi^+\pi^\pm\pi^0$; $D_s^+\to \phi\pi^+$, $K_S^0K^+$; and $J/\psi\to e^+e^-$, $\mu^+\mu^-$, respectively. $K^0_S$ and $\phi$ candidates are reconstructed via their decay to $\pi^+\pi^−$ and $K^+K^−$, respectively.

We select $B_{\text{tag}}$ candidates using two kinematic variables: $m_{\text{ES}} = \sqrt{(E_{\text{CM}}/2)^2 - \vec{p}^2_{\text{ES}}}$ and $\Delta E = E^\text{CM}_{\text{ES}} - E_{\text{B}_{\text{tag}}}$, where $E_{\text{B}_{\text{tag}}}$ and $\vec{p}_{\text{ES}}$ are the CM energy and three-momentum vector of the $B_{\text{tag}}$, respectively, and $E^\text{CM}_{\text{ES}}$ is the CM beam energy. A properly reconstructed $B_{\text{tag}}$ has $m_{\text{ES}}$ consistent with the mass of a $B$ meson and $\Delta E$ consistent with 0 GeV. We require $5.20 < m_{\text{ES}} < 5.30 \text{ GeV}/c^2$ and $-0.12 < \Delta E < 0.12 \text{ GeV}$, where the $m_{\text{ES}}$ range includes a sideward region for background studies. On average, about two $B_{\text{tag}}$ candidates per event satisfy these requirements, where the multiplicity is usually related to whether or not a soft $\pi^0$ is included in the exclusive reconstruction. If there are multiple $B_{\text{tag}}$ candidates per event, the $B_{\text{tag}}$ candidate in the highest purity mode is chosen. The purity of a $B_{\text{tag}}$ decay mode is determined from MC studies and is defined as the fraction, ranging from zero to one, of $B_{\text{tag}}$ candidates with $m_{\text{ES}} > 5.27 \text{ GeV}/c^2$ that are properly reconstructed within the given mode. If more than one $B_{\text{tag}}$ candidate with the same purity exists, the one with the smallest $|\Delta E|$ is chosen.

The hadronic $B_{\text{tag}}$ reconstruction results in both charged and neutral $B$ mesons. Since the $B_{\text{tag}}$ is fully reconstructed, its four-vector is fully determined and thus that of the $B_{\text{sig}}$ can be calculated. The latter is obtained using $|\vec{p}_{\text{B}_{\text{sig}}}| = \sqrt{(E_{\text{CM}}/2)^2 - m_B^2}$, where $\vec{p}_{\text{B}_{\text{sig}}}$ is the three momentum vector of $B_{\text{sig}}$ in the CM frame and $m_B$ is the mass of the $B$ meson, with the direction of $\vec{p}_{\text{B}_{\text{sig}}}$ opposite to that of $\vec{p}_{\text{B}_{\text{tag}}}$. The missing momentum four-vector, $p_{\text{miss}}$, is determined by subtracting the CM four-momentum of all “signal-side” tracks and clusters from that of the $B_{\text{sig}}$.

$B^+\to K^+\tau^+\tau^-$ signal events are required to have a charged $B_{\text{tag}}$ candidate with $m_{\text{ES}} > 5.27 \text{ GeV}/c^2$ and missing energy, $E_{\text{miss}}$, given by the energy component of $p_{\text{miss}}$, greater than zero. Furthermore, to reduce contamination from mis-reconstructed events with high-multiplicity $B_{\text{tag}}$ decay modes, the purity of $B_{\text{tag}}$ candidates is recalculated at this point after also requiring that there remain only three charged tracks in the event not used in the $B_{\text{tag}}$ reconstruction (corresponding to the track multiplicity in signal events). This purity is more relevant to the signal selection, since only charged $B_{\text{tag}}$ decay modes reconstructed with low multiplicity $B_{\text{sig}}$ events are considered. Signal events with a purity greater than 40% are retained.

Continuum events are further suppressed using a multivariate likelihood selector, which consists of six event-shape variables. These include the magnitude of the $B_{\text{tag}}$ thrust, defined as the axis which maximizes the sum of the longitudinal momenta of an event’s decay products, and its component along the beam axis and the ratio of the second-to-zeroth Fox-Wolfram moment [32]. The remaining variables are the angle of the missing momentum vector, $\vec{p}_{\text{miss}}$, with the beam axis, the angle between $\vec{p}_{\text{B}_{\text{tag}}}$ and the beam axis, and the angle between the thrust axis of the $B_{\text{tag}}$ and that of the $B_{\text{sig}}$ in the CM frame. The six event-shape variables discriminate between $B\bar{B}$ events, where the spin-zero $B$ mesons are produced almost at rest and the decay daughters consequently produce an isotropic distribution, and continuum events. In the latter, fermions are initially produced with higher momentum, resulting in a more collinear distribution of the final decay products. We require the likelihood ratio

$$L = \frac{\prod_i P_B(x_i)}{\prod_i P_B(x_i) + \prod_i P_Q(x_i)} > 0.50,$$

where $P(x_i)$ are probability density functions, determined from MC samples, that describe the six event-shape variables for $B\bar{B}$, $P_B(x_i)$, and continuum, $P_Q(x_i)$, events. This requirement removes more than 75% of the continuum events while retaining more than 80% of (signal and background) $B\bar{B}$ MC events.

A signal selection is then applied on the charged tracks and neutral clusters that are not used in the $B_{\text{tag}}$ reconstruction. $B^+\to K^+\tau^+\tau^-$ candidates are required to possess exactly three charged tracks satisfying particle identification (PID) requirements consistent with one charged $K$ and an $e^+e^−$, $\mu^+\mu^−$, or $e^+e^−$ pair. The PID selection algorithms for charged tracks are based on multivariate analysis techniques that use information from the BABAR detector subsystems [25]. The $K^{±}$ is required to have a charge opposite to that of $B_{\text{tag}}$. Furthermore, events with $3.00 < m_{e^+e^-} < 3.19 \text{ GeV}/c^2$ are discarded to remove backgrounds with a $J/\psi$ resonance. The invariant mass of the combination of the $K$ with the oppositely charged lepton must also lie outside the
region of the $D^0$ mass, i.e. $m_{K^-\ell^+} < 1.80\text{GeV}/c^2$ or $m_{K^+\ell^-} > 1.90\text{GeV}/c^2$, to remove events where a pion coming from the $D^0$ decay is misidentified as a muon. Moreover, events with $\gamma \rightarrow e^+e^-$ are removed by requiring the invariant mass of each electron with any other oppositely charged track in the event to be greater than 50 MeV/$c^2$. Background events with $\pi^0$ candidates, reconstructed from a pair of photons with individual energies greater than 50 MeV, a total CM energy greater than 100 MeV, and an invariant mass ranging between 100 and 160 MeV/$c^2$, are rejected. Additional calorimeter clusters not explicitly associated with $B_{\text{tag}}$ daughter particles may originate from other low-energy particles in background events. We therefore define $E_{\text{extra}}^*$ to be the energy sum of all neutral clusters with individual energy greater than 50 MeV that are not used in the $B_{\text{tag}}$ reconstruction.

The normalized squared mass of the $\tau^+\tau^-$ pair is given by $s_B = (p_{B_{\text{tag}}}-p_K)^2/m_B^2$, where $p_{B_{\text{tag}}}$ and $p_K$ are the four-momentum vectors of $B_{\text{tag}}$ and of the kaon, respectively, in the laboratory frame. The large mass of the $\tau$ leptons in signal events kinematically limits the $s_B$ distribution to large values. A requirement of $s_B > 0.45$ is applied. A peaking distribution about the $\psi(2S)$ $s_B$ value is not observed, and thus the contribution of this background is considered negligible.

At this point in the selection, remaining backgrounds are primarily $B\bar{B}$ events in which a properly reconstructed $B_{\text{tag}}$ is accompanied by $B_{\text{sig}} \rightarrow D^*(\pm)\tau\bar{\tau}$, with $D^*(\pm) \rightarrow K\ell\bar{\ell}$, and thus have the same detected final-state particles as signal events. A multi-layer perceptron (MLP) neural network [33], with eight input variables and one hidden layer, is employed to suppress this background. The input variables are: (i) the angle between the kaon and the oppositely charged lepton, (ii) the angle between the two leptons, and (iii) the momentum of the lepton with charge opposite to the $K$, all in the $\tau^+\tau^-$ rest frame, which is calculated as $p_{B_{\text{tag}}} - p_K$; (iv) the angle between the $B_{\text{tag}}$ and the oppositely charged lepton, (v) the angle between the $K$ and the low-momentum lepton, and (vi) the invariant mass of the $K^+\ell^-\bar{\ell}$ pair, all in the CM frame. Furthermore, the final input variables to the neural network are (vii) $E_{\text{extra}}^*$ and (viii) the residual energy, $E_{\text{res}}$, which here is effectively the missing energy associated with the $\tau^+\tau^-$ pair and is calculated as the energy component of $p_{\text{residual}}^\tau = p_{B_{\text{tag}}}^\tau - p_K^\tau - p_{\ell^+\ell^-}^\tau$, where $p_{B_{\text{tag}}}^\tau$, $p_K^\tau$ and $p_{\ell^+\ell^-}^\tau$ are the four-momenta vectors in the $\tau^+\tau^-$ rest frame of the $B_{\text{tag}}$, $K$, and lepton pair in the event, respectively. $E_{\text{res}}$ has, in general, higher values for signal events than generic $B\bar{B}$ and continuum events due to the higher neutrino multiplicity. A neural network is trained and tested using randomly split dedicated signal MC and $B^+\bar{B}^-$ background events, for each of the three channels: $e^+e^-$, $\mu^+\mu^-$, and $e^+\mu^-\bar{\mu}$. The results are shown in Fig. 2 for the three modes combined. The last step in the signal selection is to require that the output of the neural network is $> 0.70$ for the $e^+e^-$ and $\mu^+\mu^-$ channels and $> 0.75$ for the $e^+\mu^-\bar{\mu}$ channel. This requirement is optimized to yield the most stringent upper limit in the absence of a signal.

The branching fraction for each of the signal modes, $i$, is calculated as:

$$B_i = \frac{N_i^{\text{obs}} - N_i^{\text{bkg}}}{\epsilon_i^{\text{sig}} N_{B\bar{B}}},$$

where $N_{B\bar{B}} = 471 \times 10^6$ is the total number of $B\bar{B}$ pairs in the data sample, assuming equal production of $B^+B^-$ and $B^0\bar{B}^0$ pairs in $\Upsilon(4S)$ decays, and $N_i^{\text{obs}}$ is the number of data events passing the signal selection. The signal efficiency, $\epsilon_i^{\text{sig}}$, and the background estimate, $N_i^{\text{bkg}}$, are determined for each mode from the signal and background MC yields after all selection requirements.

For each mode, $N_i^{\text{bkg}}$ consists of two components: background events that have a properly reconstructed $B_{\text{tag}}$ and thus produce a distribution in $m_{\text{ES}}$ which peaks at the $B$ mass, and combinatorial background events composed of continuum and $B\bar{B}$ events with misreconstructed $B_{\text{tag}}$ candidates which do not produce a peaking structure in the $m_{\text{ES}}$ signal region. After the MLP output requirement, peaking background events comprise $84\%$ of the total $N_i^{\text{bkg}}$ for all three modes. To reduce the dependence on MC simulation, the combinatorial background is extrapolated directly from the yield of data events in the $m_{\text{ES}}$ “sideband” region ($5.20 < m_{\text{ES}} < 5.26\text{GeV}/c^2$), after the full signal selection. The yield of sideband data events is scaled by the ratio, determined from MC, of combinatorial background in the $m_{\text{ES}}$ signal region to the $m_{\text{ES}}$ sideband region, and used to estimate the combinatorial background component of data in the signal region.

The peaking background is determined using $B^+B^-$
background MC, while data in the final signal region is kept blinded to avoid experimentalist bias. Because of the large uncertainties on the branching fractions of many of the $B_{\text{tag}}$ decay modes as well as their associated reconstruction effects, there is a discrepancy in the $B_{\text{tag}}$ yield of approximately 10% between MC and data, independent of the signal selection. A $B_{\text{tag}}$ yield correction is therefore determined by calculating the ratio of data to $B^+B^-$ MC events before the final MLP requirement. The data sample after this requirement contains a sufficiently large background contribution after the $s_B$ requirement, which consists mainly of $B^+B^-$ events (>96%) according to MC simulation, to allow for a data-driven correction without unblinding the final signal-region. This correction factor is determined to be 0.913 ± 0.020, where the uncertainty is statistical only, and is applied to the MC reconstruction efficiency for both signal and background events.

The $B_{\text{tag}}$ yield is also cross-checked using a $B^+ \to D^0 \ell^+ \nu$, $D^0 \to K^- \pi^+$ control sample, which is selected using the same signal selection discussed above, but with requiring one track to satisfy pion instead of lepton PID and reversing the $D^0$ veto, such that $1.80 < m_{K^- \pi^+} < 1.90$ GeV/$c^2$. These criteria are also applied to the full background MC and the resulting sample is found to consist mainly of peaking $B^+B^-$ events, which the MLP neural network is trained to classify as background. Before the MLP requirement, good agreement between data and MC is found in all the distributions of the input variables of the $B^+ \to D^0 \ell^+ \nu$, $D^0 \to K^- \pi^+$ samples, as shown in Fig. 3 for the $m_{K^- \pi^+}$ distribution. These samples are then run through the MLP neural network and a detailed comparison of the MLP output and the input variables, after the full signal selection, is performed.

The results for each signal channel are then combined to determine $B(B^+ \to K^+ \tau^+ \tau^-)$. This is done using a frequentist approach by finding the value of $B$ that maximizes the product of the Poisson likelihoods of observing $N_{\text{obs}}^i$ in each of the signal channels. Branching fraction uncertainties and limits are determined using the method described in Ref. [34], taking into account the statistical and systematic uncertainties on $N_{\text{bkg}}$ and $\epsilon_{\text{sig}}$.

Systematic uncertainties associated with the level of data-MC agreement are determined for most of the variables used in the signal selection. The determination of the $B_{\text{tag}}$ yield correction is anti-correlated with the extrapolation of the combinatorial background from the $m_{\text{ES}}$ sideband, as both use the combinatorial background shape from MC. Therefore, only one systematic uncertainty on the $B_{\text{tag}}$ yield and combinatorial background estimate is evaluated, using a simulated MC sample composed of background events with the same luminosity as the data sample. Accounting for the anti-correlation, the effect of varying the value of the $B_{\text{tag}}$ yield correction on the final signal efficiency and background estimate is determined to be 1.2% and 1.6%, respectively. The uncertainty associated with the theoretical model is evaluated by reweighting the $s_B$ distribution of the dedicated signal MC sample to the LCSR [30] theoretical model and to that of Ref. [35] and determining the difference in signal efficiency, which is calculated to be 3.0%. The resonant $B \to K^+ \psi(2S), \psi(2S) \to \tau^+ \tau^-$ decay has a negligible background contribution and thus only non-resonant models are used to estimate the theoretical uncertainty, especially since the kinematics of any new physics sources are not well-known. Additional uncertainties on $\epsilon_{\text{sig}}$ and $N_{\text{bkg}}$ arise due to the modeling of PID selectors (4.8% for $e^+e^-$, 7.0% for $\mu^+\mu^-$, and 5.0% for $e^+\mu^-$) and the $\pi^0$ veto (3.0%). The level of agreement between data and MC is evaluated using the $B^+ \to D^0 \ell^+ \nu$, $D^0 \to K^- \pi^+$ control sample before and after the MLP requirement. Comparison of both the overall yields as well as the distributions of the input and output variable results in a systematic uncertainty of 2.6%. Other potential sources of systematic uncertainties have been investigated, including those associated with the assumption that charged and neutral $B$ candidates are produced at equal rates, the continuum likelihood suppression, $B_{\text{tag}}$ purity, track multiplicity, $E_{\text{miss}}$ and $s_B$ selection criteria, and are all implicitly accounted for in the $B_{\text{tag}}$ yield correction uncertainty. Correlations between the signal efficiency and the background estimate due to common systematic errors are included, but are found to have a negligible effect on the final branching fraction results.

The final signal efficiencies, background estimates and observed yields of each signal mode are shown in Table I, with the associated branching fraction significance. The yields in the $e^+e^-$ and $\mu^+\mu^-$ channels show consistency with the expected background estimate. The signal yield in the $e^+\mu^-$ channel is approximately equal to the sum of

---

**FIG. 3:** (color online) Invariant-mass distribution of the $K^-\pi^+$ pair in the $B^+ \to D^0 \ell^+ \nu$, $D^0 \to K^- \pi^+$ samples after all signal selection criteria are applied, except for the final requirement on the MLP output. The data (points) are overlaid on the expected combinatorial (hatched) plus $m_{\text{ES}}$-peaking (solid line) background contributions.

---
ACKNOWLEDGMENTS

We are grateful for the excellent luminosity and machine conditions provided by our PEP-II colleagues, and for the substantial dedicated effort from the computing organizations that support BABAR. The collaborating institutions wish to thank SLAC for its support and kind hospitality. This work is supported by DOE and NSF (USA), NSERC (Canada), CEA and CNRS-IN2P3 (France), BMBF and DFG (Germany), INFN (Italy), FOM (The Netherlands), NFR (Norway), MES (Russia), MINECO (Spain), STFC (United Kingdom). Individuals have received support from the Marie Curie EIF (European Union) and the A. P. Sloan Foundation (USA).

* Now at: Wuhan University, Wuhan 43072, China
† Now at: Università di Bologna and INFN Sezione di Bologna, I-47921 Rimini, Italy
‡ Now at: University of Huddersfield, Huddersfield HD1 3DH, UK
§ Now at: University of South Alabama, Mobile, Alabama 36688, USA
¶ Also at: Università di Sassari, I-07100 Sassari, Italy
[1] Charge conjugation is implied throughout the entire paper.
R. Aaij et al [LHCb Collaboration], JHEP 07, 133 (2012);
R. Aaij et al. [LHCb Collaboration], JHEP 02, 105 (2013);