Measurements of branching fractions, rate asymmetries, and angular distributions in the rare decays $B \to K\ell^+\ell^-$ and $B \to K^*\ell^+\ell^-$


(BABAR Collaboration)

1Laboratoire de Physique des Particules, F-74941 Annecy-le-Vieux, France
2Universitat de Barcelona, Facultat de Física Dept. ECM, E-08028 Barcelona, Spain
3Università di Bari, Dipartimento di Fisica and INFN, I-70126 Bari, Italy
4Institute of High Energy Physics, Beijing 100039, China
5University of Bergen, Institute of Physics, N-5007 Bergen, Norway

6Lawrence Berkeley National Laboratory and University of California, Berkeley, California 94720, USA
7University of Birmingham, Birmingham, B15 2TT, United Kingdom
8Ruhr Universität Bochum, Institut für Experimentalphysik 1, D-44780 Bochum, Germany
9University of Bristol, Bristol BS8 1TL, United Kingdom
10University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z1
11Brunel University, Uxbridge, Middlesex UB8 3PH, United Kingdom
12Budker Institute of Nuclear Physics, Novosibirsk 630090, Russia
13University of California at Irvine, Irvine, California 92697, USA
14University of California at Los Angeles, Los Angeles, California 90024, USA
15University of California at Riverside, Riverside, California 92521, USA
16University of California at San Diego, La Jolla, California 92039, USA
17University of California at Santa Barbara, Santa Barbara, California 93106, USA
18University of California at Santa Cruz, Institute for Particle Physics, Santa Cruz, California 95064, USA
19University of California Institute of Technology, Pasadena, California 91125, USA
20University of Cincinnati, Cincinnati, Ohio 45221, USA
21University of Colorado, Boulder, Colorado 80309, USA
22Colorado State University, Fort Collins, Colorado 80523, USA
23Universität Dortmund, Institut für Physik, D-44221 Dortmund, Germany
Technische Universität Dresden, Institut für Kern- und Teilchenphysik, D-01062 Dresden, Germany
Ecole Polytechnique, LLR, F-91128 Palaiseau, France
University of Edinburgh, Edinburgh EH9 3JZ, United Kingdom
Università di Ferrara, Dipartimento di Fisica e INFN, I-44100 Ferrara, Italy
Laboratori Nazionali di Frascati dell’INFN, I-00044 Frascati, Italy
Università di Genova, Dipartimento di Fisica e INFN, I-16146 Genova, Italy
Harvard University, Cambridge, Massachusetts 02138, USA
Universität Heidelberg, Physikalisches Institut, Philosophenweg 12, D-69120 Heidelberg, Germany
Imperial College London, London, SW7 2AZ, United Kingdom
University of Iowa, Iowa City, Iowa 52242, USA
Johns Hopkins University, Baltimore, Maryland 21218, USA
Universität Karlsruhe, Institut für Experimentelle Kernphysik, D-76021 Karlsruhe, Germany
Lawrence Livermore National Laboratory, Livermore, California 94550, USA
University of Liverpool, Liverpool L69 7ZE, United Kingdom
McGill University, Montréal, Québec, Canada H3A 2T8
Università di Milano, Dipartimento di Fisica e INFN, I-20133 Milano, Italy
University of Mississippi, University, Mississippi 38677, USA
Université de Montréal, Physique des Particules, Montréal, Québec, Canada H3C 3J7
Universidad de Napoli Federico II, I-80126, Napoli, Italy
NIKHEF, National Institute for Nuclear Physics and High Energy Physics, NL-1009 DB Amsterdam, The Netherlands
Université de Notre Dame, Notre Dame, Indiana 46556, USA
Ohio State University, Columbus, Ohio 43210, USA
University of Oregon, Eugene, Oregon 97403, USA
Università di Padova, Dipartimento di Fisica e INFN, I-35131 Padova, Italy
Universités Paris VI et VII, Laboratoire de Physique Nucléaire et de Hautes Energies, F-75252 Paris, France
University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA
Università di Perugia, Dipartimento di Fisica e INFN, I-06100 Perugia, Italy
Università di Pisa, Dipartimento di Fisica, Scuola Normale Superiore e INFN, I-56127 Pisa, Italy
Prairie View A&M University, Prairie View, Texas 77446, USA
Princeton University, Princeton, New Jersey 08544, USA
Università di Roma La Sapienza, Dipartimento di Fisica e INFN, I-00185 Roma, Italy
University of Rostock, D-18051 Rostock, Germany
Rutherford Appleton Laboratory, Chilton, Didcot, Oxon, OX11 0QX, United Kingdom
DSM/Dapnia, CEA/Saclay, F-91191 GIF-sur-Yvette, France
University of South Carolina, Columbia, South Carolina 29208, USA
Stanford Linear Accelerator Center, Stanford, California 94309, USA
Stanford University, Stanford, California 94305-4060, USA
State University of New York, Albany, New York 12222, USA
University of Tennessee, Knoxville, Tennessee 37996, USA
University of Texas at Austin, Austin, Texas 78712, USA
University of Texas at Dallas, Richardson, Texas 75083, USA
Università di Torino, Dipartimento di Fisica Sperimentale e INFN, I-10125 Torino, Italy
Università di Trieste, Dipartimento di Fisica e INFN, I-34127 Trieste, Italy
IFIC, Universitat de Valencia-CSIC, E-46071 Valencia, Spain
University of Victoria, Victoria, British Columbia, Canada V8W 3P6
Department of Physics, University of Warwick, Coventry CV4 7AL, United Kingdom
University of Wisconsin, Madison, Wisconsin 53706, USA
Yale University, New Haven, Connecticut 06511, USA

(Received 4 April 2006; published 5 May 2006)
We present measurements of the flavor-changing neutral current decays $B \to K\ell^+\ell^-$ and $B \to K^*\ell^+\ell^-$, where $\ell^+\ell^-$ is either an $e^+e^-$ or $\mu^+\mu^-$ pair and $K^{(*)}$ denotes either a kaon or the $K^{(*)}(892)$ meson, are manifestations of $b \to s\ell^+\ell^-$ flavor-changing neutral currents (FCNC). In the standard model (SM), these decays are forbidden at tree level and their predicted properties could be significantly modified by new physics at the electroweak scale. We measure the branching fractions $\mathcal{B}(B \to K\ell^+\ell^-) = (0.34 \pm 0.07 \pm 0.02) \times 10^{-6}$, $\mathcal{B}(B \to K^*\ell^+\ell^-) = (0.78^{+0.19}_{-0.17} \pm 0.11) \times 10^{-6}$, the direct $CP$ asymmetries of these decays, and the relative abundances of decays to electrons and muons. For two regions in $\ell^+\ell^-$ mass, above and below $m_{t\ell\phi}$, we measure partial branching fractions and the forward-backward angular asymmetry of the lepton pair. In these same regions we also measure the $K^*$ longitudinal polarization in $B \to K^*\ell^+\ell^-$ decays. Upper limits are obtained for the lepton-flavor-violating decays $B \to K e\mu$ and $B \to K^* e\mu$. All measurements are consistent with standard model expectations.

DOI: 10.1103/PhysRevD.73.092001

PACS numbers: 13.25.Hw, 13.20.He

I. INTRODUCTION

The decays $B \to K^{(*)}\ell^+\ell^-$, where $\ell^+\ell^-$ is either an $e^+e^-$ or $\mu^+\mu^-$ pair and $K^{(*)}$ denotes either a kaon or the $K^{(*)}(892)$ meson, are manifestations of $b \to s\ell^+\ell^-$ flavor-changing neutral currents (FCNC). In the standard model (SM), these decays are forbidden at tree level and can only occur at greatly suppressed rates through higher-order processes. At lowest order, three amplitudes contribute: (i) a photon penguin, (ii) a $Z$ penguin, and (iii) a $W^+W^-$ box diagram (Fig. 1). In all three, a virtual $t$ quark contribution dominates, with secondary contributions from virtual $c$ and $u$ quarks. Within the Operator Product Expansion (OPE) framework, these short-distance contributions are typically described in terms of the effective Wilson coefficients $C_7^{\gamma f}$, $C_9^{\gamma f}$, and $C_{10}^{\gamma f}$ [1]. Since these decays proceed via weakly-interacting particles with virtual energies near the electroweak scale, they provide a promising means to search for effects from new interactions entering with amplitudes comparable to those of the SM. Such effects are predicted in a wide variety of models [2–6].

In the SM the $B \to K\ell^+\ell^-$ branching fraction is predicted to be roughly $0.4 \times 10^{-6}$, while the $B \to K^*\ell^+\ell^-$ branching fraction is predicted to be about 3 times larger [4,7–12]. The $B \to K^*\ell^+\ell^-$ mode receives a significant contribution from a pole in the photon penguin amplitude at low values of $q^2 \equiv m_{\ell\ell}^2$, which is not present in $B \to K\ell^+\ell^-$ decays. Because of the lower mass threshold for producing an $e^+e^-$ pair, this enhances the $K^*e^+e^-$ final state relative to the $K^*\mu^+\mu^-$ state. Currently, theoretical predictions of the branching fractions have associated uncertainties of about 30% due to form factors that model the hadronic effects in the $B \to K$ or $B \to K^*$ transition. Previous experimental measurements of the branching fractions are consistent with the range of theoretical predictions, with experimental uncertainties comparable in size to the theoretical uncertainties [13,14].

With larger datasets, it becomes possible to measure ratios and asymmetries in the rates. These can typically be predicted more reliably than the total branching fractions. For example, the direct $CP$ asymmetry

$$A_{CP} = \frac{\Gamma(B \to K^{(*)}\ell^+\ell^-) - \Gamma(B \to K^{(*)}\ell^+\ell^-)}{\Gamma(B \to K^{(*)}\ell^+\ell^-) + \Gamma(B \to K^{(*)}\ell^+\ell^-)}$$

is expected to be vanishingly small in the SM, of order $10^{-4}$ in the $B \to K^*\ell^+\ell^-$ mode [15]. However it could be enhanced by new non-SM weak phases [16]. The ratio $R_K$, defined as

$$R_K = \frac{\Gamma(B \to K\mu^+\mu^-)}{\Gamma(B \to K e^+ e^-)},$$

also has a precise SM prediction of $R_K = 1.0000 \pm 0.0001$ [17]. In supersymmetric theories with a large ratio ($\tan\beta$) of vacuum expectation values of Higgs doublets, $R_K$ can be significantly enhanced. This occurs via penguin diagrams in which the $\gamma$ or $Z^0$ is replaced with a neutral Higgs boson that preferentially couples to the heavier muons [18]. In $B \to K^*\ell^+\ell^-$ this ratio is modified by the photon pole contribution, thus the SM prediction is $R_{K^*} = 0.75$ [4] with an estimated uncertainty of 0.01 [17] if the pole region is included, or $R_{K^*} \approx 1.0$ if it is excluded [17].

![Fig. 1. Examples of standard model Feynman diagrams for the decays $B \to K^{(*)}\ell^+\ell^-$. For the photon or $Z$ penguin diagrams on the left, boson emission can occur on any of the $b$, $t$, $c$, $u$, $s$, or $W$ lines.](image-url)
Additional sensitivity to non-SM physics arises from the fact that $B \to K^{(*)}\ell^+\ell^-$ transitions are three-body decays proceeding through three different electroweak penguin amplitudes, whose relative contributions vary as a function of $q^2$. Measurements of partial branching fractions and angular distributions as a function of the invariant momentum transfer $q^2$ are therefore of particular interest. The SM predicts a distinctive pattern in the forward-backward asymmetry

$$A_{FB}(s) \equiv \int_{-1}^{1} d\cos\theta \frac{d^2\Gamma(B\to K^{(*)}\ell^+\ell^-)\text{Sign}(\cos\theta)}{d^2\cos\theta} / dF(B\to K^{(*)}\ell^+\ell^-)/ds,$$

where $s \equiv q^2/m_R^2$ and $\theta$ is the angle of the lepton with respect to the flight direction of the $B$ meson, measured in the dilepton rest frame [19]. In the presence of non-SM physics, the sign and magnitude of this asymmetry can be altered dramatically [4,9,15]. In particular, at high $q^2$, the sign of $A_{FB}$ is sensitive to the sign of the product of the $C_{9}^{\text{eff}}$ and $C_{10}^{\text{eff}}$ Wilson coefficients. The value of $A_{FB}$ in $B \to K^{(*)}\ell^+\ell^-$ provides an important check on this measurement, as it is expected to result in zero asymmetry for all $q^2$ in the SM and many non-SM scenarios. This condition can be violated in models in which new operators such as a neutral Higgs penguin contribute significantly [18]. However even in this case the resulting asymmetry is expected to be of order 0.01 or less in the $B \to K^{(*)}\ell^+\ell^-$ mode for electron or muon final states [20]. In addition to $A_{FB}$, in $B \to K^{(*)}\ell^+\ell^-$ the fraction of longitudinal polarization $F_L$ of the $K^*$ can be measured from the angular distribution of its decay products. The value of $F_L$ measured at low $q^2$ is sensitive to effects from new left-handed currents with complex phases different from the SM, resulting in $C_{9}^{\text{eff}} = -C_7^{(\text{SM})}$, or effects from new right-handed currents in the photon penguin amplitude [21]. The predicted distributions of $A_{FB}(q^2)$ and $F_L(q^2)$ are shown for the SM and for several non-SM scenarios in Fig. 2. The non-SM scenarios correspond to those studied in Refs. [4,9,21].

Finally, the lepton-flavor-violating decays $B \to K^{(*)}\ell^+\mu^-$ can only occur at rates far below current experimental sensitivities in the context of the SM with neutrino mixing. Observation of these decays would therefore be an indication of contributions beyond the SM. For example, such decays are allowed in leptoquark models [6].

### II. DETECTOR AND DATASET

The results presented here are based on data collected with the $BABAR$ detector at the PEP-II asymmetric $e^+e^-$ collider located at the Stanford Linear Accelerator Center. The dataset comprises $229 \times 10^6 BB$ pairs, corresponding to an integrated luminosity of 208 fb$^{-1}$ collected on the Y(4S) resonance at a center-of-mass energy of $\sqrt{s} = 10.58$ GeV. An additional 12.1 fb$^{-1}$ of data collected at energies 40 MeV below the nominal on-peak energy is used to study continuum backgrounds arising from pair production of $u$, $d$, $s$, and $c$ quarks.

The $BABAR$ detector is described in detail in Ref. [22]. The measurements described in this paper rely primarily on the charged-particle tracking and identification properties of the detector. Tracking is provided by a five-layer silicon vertex tracker (SVT) and a 40-layer drift chamber (DCH) in a 1.5-T magnetic field produced by a superconducting magnet. Low momentum charged hadrons are identified by the ionization loss $(dE/dx)$ measured in the SVT and DCH, and higher momentum hadrons by a ring-imaging detector of internally reflected Cherenkov light (DIRC). A CsI(Tl) electromagnetic calorimeter (EMC) provides identification of electrons, and detection of photons. The steel in the instrumented flux return (IFR) of the superconducting coil is interleaved with resistive plate chambers, providing identification of muons and neutral hadrons.
III. EVENT SELECTION

We reconstruct signal candidates in eight final states: $B^+ \rightarrow K^+ \ell^+ \ell^-$, $B^0 \rightarrow K^0_S \ell^+ \ell^-$, $B^0 \rightarrow K^{*0} \ell^+ \ell^-$, $B^+ \rightarrow K^{*+} \ell^+ \ell^-$, where $K^0_S \rightarrow K^- \pi^+$, $K^{*0} \rightarrow K^0_S \pi^+$, $K^{*+} \rightarrow K^0_S \pi^+$, and $\ell$ is either an $e$ or $\mu$. Throughout this paper, charge-conjugate modes are implied.

Electrons are required to have momentum above 0.3 GeV/c and are identified using a likelihood ratio combining information from the EMC, DIRC, and DCH. Photons that lie in a small angular region around the electron direction and have $E > 30$ MeV are combined with electron candidates in order to recover bremsstrahlung energy. We suppress backgrounds due to photon conversions in the $B \rightarrow K e^+ e^-$ channels by removing $e^+ e^-$ pairs with invariant mass less than 0.03 GeV/c$^2$. As there is a significant contribution to the $B \rightarrow K e^+ e^-$ channels from the pole at low dielectron mass, we preserve acceptance by vetoing conversions in these channels only if the conversion radius is outside the inner radius of the beam pipe. Muons with momentum $p > 0.7$ GeV/c are identified with a neural network algorithm using information from the IFR and the EMC.

The performance of the lepton identification algorithms is evaluated using high-statistics data control samples. The electron efficiency is determined from samples of $p > 0.0025 \times 0.0135 \times 0.0135 \times 0.0255$ with electron candidates in order to recover bremsstrahlung energy. We suppress backgrounds due to photon conversions in the $B \rightarrow K e^+ e^-$ channels by removing $e^+ e^-$ pairs with invariant mass less than 0.03 GeV/c$^2$. As there is a significant contribution to the $B \rightarrow K e^+ e^-$ channels from the pole at low dielectron mass, we preserve acceptance by vetoing conversions in these channels only if the conversion radius is outside the inner radius of the beam pipe. Muons with momentum $p > 0.7$ GeV/c are identified with a neural network algorithm using information from the IFR and the EMC.

The performance of the lepton identification algorithms is evaluated using high-statistics data control samples. The electron efficiency is determined from samples of $p > 0.0025 \times 0.0135 \times 0.0135 \times 0.0255$ with electron candidates in order to recover bremsstrahlung energy. We suppress backgrounds due to photon conversions in the $B \rightarrow K e^+ e^-$ channels by removing $e^+ e^-$ pairs with invariant mass less than 0.03 GeV/c$^2$. As there is a significant contribution to the $B \rightarrow K e^+ e^-$ channels from the pole at low dielectron mass, we preserve acceptance by vetoing conversions in these channels only if the conversion radius is outside the inner radius of the beam pipe. Muons with momentum $p > 0.7$ GeV/c are identified with a neural network algorithm using information from the IFR and the EMC.

The performance of the lepton identification algorithms is evaluated using high-statistics data control samples. The electron efficiency is determined from samples of $p > 0.0025 \times 0.0135 \times 0.0135 \times 0.0255$ with electron candidates in order to recover bremsstrahlung energy. We suppress backgrounds due to photon conversions in the $B \rightarrow K e^+ e^-$ channels by removing $e^+ e^-$ pairs with invariant mass less than 0.03 GeV/c$^2$. As there is a significant contribution to the $B \rightarrow K e^+ e^-$ channels from the pole at low dielectron mass, we preserve acceptance by vetoing conversions in these channels only if the conversion radius is outside the inner radius of the beam pipe. Muons with momentum $p > 0.7$ GeV/c are identified with a neural network algorithm using information from the IFR and the EMC.

The performance of the lepton identification algorithms is evaluated using high-statistics data control samples. The electron efficiency is determined from samples of $p > 0.0025 \times 0.0135 \times 0.0135 \times 0.0255$ with electron candidates in order to recover bremsstrahlung energy. We suppress backgrounds due to photon conversions in the $B \rightarrow K e^+ e^-$ channels by removing $e^+ e^-$ pairs with invariant mass less than 0.03 GeV/c$^2$. As there is a significant contribution to the $B \rightarrow K e^+ e^-$ channels from the pole at low dielectron mass, we preserve acceptance by vetoing conversions in these channels only if the conversion radius is outside the inner radius of the beam pipe. Muons with momentum $p > 0.7$ GeV/c are identified with a neural network algorithm using information from the IFR and the EMC.

FIG. 3 (color online). Predicted distributions of (a) $A_{FB}(q^2)$ and (b) $F_L(q^2)$ in $B \rightarrow K^+ \ell^+ \ell^-$ for the two regions of $q^2$ considered. The lines represent the predictions of the SM (solid lines), $C_{eff}^{10} = -C_7^0(SM)$ (dotted lines), $C_9^{10} C_{eff}^{10} = -C_7^0 C_{10}(SM)$ (dashed lines), and $C_9^{10} C_{eff}^{10} = -C_7^0 C_{10}(SM)$ (dot-dashed lines) with the form factor model of Ref. [27]. In the case of $F_L$, the two solutions with $C_9^{10} C_{eff}^{10} = -C_7^0 C_{10}(SM)$ are not displayed; they are nearly identical to the two shown.
$-0.50 < \Delta E < 0.50 \text{ GeV}$. The signal is extracted by performing a multidimensional, unbinned maximum-likelihood fit in the region $5.20 < m_{ES} < 5.29 \text{ GeV}/c^2$ and $-0.25 < \Delta E < 0.25 \text{ GeV}$, which contains 100% of the signal candidates that pass all other selection requirements. This region remains blind to our inspection until all selection criteria are established. The events in the sideband with $5.00 < m_{ES} < 5.20 \text{ GeV}/c^2$, or $-0.50 < \Delta E < -0.25 \text{ GeV}$, or $0.25 < \Delta E < 0.50 \text{ GeV}$ are used to study the properties of the combinatorial background.

For the measurements of the partial branching fractions, $A_{FB}$, and $K^*$ polarization, we subdivide the sample into two regions of dilepton invariant mass. The first is the region above the pole and below the $J/\psi$ resonance, $0.1 < q^2 < 8.41 \text{ GeV}^2/c^4$, the second is the region $q^2 > 10.24 \text{ GeV}^2/c^4$, above the $J/\psi$ resonance. The $\psi(2S)$ resonance is explicitly excluded from this upper region as described in further detail in Sec. IV B. The lower bound of $0.1 \text{ GeV}^2/c^4$ in the first region is chosen to remove effects from the photon pole in the $B \to K^* e^+ e^-$ channel. The forward-backward asymmetry is extracted in each of these $q^2$ regions from the distribution of $\cos \theta^*$, which we define as the cosine of the angle between the $\ell^-$ ($\ell^+$) and the $B$ ($\bar{B}$) meson, measured in the dilepton rest frame. We do not measure $A_{FB}$ in the mode $B^0 \to K^0_S \ell^+ \ell^-$, in which the flavor of the $B$ meson cannot be directly inferred from the $K^0_S$. The $K^*$ polarization is similarly derived from the distribution of $\cos \theta_{K^*}$, defined as the cosine of the angle between the $K$ and the $B$ meson, measured in the $K^*$ rest frame. The predicted distributions of $A_{FB}$ and $F_L$ integrated over these two $q^2$ ranges are shown in Fig. 3 for both the SM and non-SM scenarios.

IV. BACKGROUND SOURCES

A. Combinatorial backgrounds

Combinatorial backgrounds arise either from the continuum, in which a $(u, d, s, o, c)$ quark pair is produced, or from $B\bar{B}$ events in which the decay products of the two $B$’s are misreconstructed as a signal candidate. We use the following variables computed in the CM frame to reject continuum backgrounds: (i) the ratio of second to zeroth Fox-Wolfram moments [23], (ii) the angle between the thrust axis of the $B$ and the remaining particles in the event, $\theta_{thrust}$, (iii) the production angle $\theta_p$ of the $B$ candidate with respect to the beam axis, and (iv) the invariant mass of the kaon-lepton pair with the charge combination expected from a semileptonic $D$ decay. The first three variables take advantage of the characteristic jetlike event shape of continuum backgrounds, versus the more spherical event shape of $B\bar{B}$ events. The fourth variable is useful for rejecting $c\bar{c}$ events. These frequently occur through decays such as $D \to K \ell \nu$, resulting in a kaon-lepton invariant mass which peaks below that of the $D$; for signal events the kaon-lepton mass is broadly distributed up to approximately the $B$ mass. These four variables are combined into a linear Fisher discriminant [24], which is optimized using samples of simulated signal events and off-resonance data. A separate Fisher discriminant is used for each of the decay modes considered in this analysis.

Combinatorial $B\bar{B}$ backgrounds are dominated by events with two semileptonic $B \to X \ell \nu$ decays. We discriminate against these events by constructing a likelihood ratio composed of (i) the vertex probability of the dilepton pair, (ii) the vertex probability of the $B$ candidate, (iii) the angle $\theta_p$ as in the Fisher discriminant, and (iv) the total missing energy in the event $E_{miss}$. Events with two semileptonic decays will contain at least two neutrinos; therefore the $E_{miss}$ variable is particularly effective at rejecting these backgrounds. The probability distribution functions (PDFs) used in the likelihood are derived by fitting simulated signal events and simulated $B\bar{B}$ events in which the signal decays are removed. We derive a separate likelihood parametrization for each decay mode.

We select those events that pass an optimal Fisher and $B\bar{B}$ likelihood requirement, based on the figure of merit $S/\sqrt{S+B}$ for the expected number of signal events $S$ and background events $B$. The selection is optimized simultaneously for the Fisher and likelihood, and is derived separately for each decay mode.

B. Peaking backgrounds

Backgrounds that peak in the $m_{ES}$ and $\Delta E$ variables in the same manner as the signal are either vetoed, or their rate is estimated from simulated data or control samples. The largest sources of peaking backgrounds are $B$ decays to charmonium: $B \to J/\psi K^{(*)}$ and $B \to \psi(2S) K^{(*)}$, where the $J/\psi$ or $\psi(2S)$ decays to a $\ell^+ \ell^-$ pair. We therefore remove events in which the dilepton invariant mass is consistent with a $J/\psi$ or $\psi(2S)$, either with or without bremsstrahlung recovery in the electron channels. In cases where the lepton momentum is mismeasured, or the bremsstrahlung recovery algorithm fails to find a radiated photon, the dilepton mass will be shifted from the charmonium mass. In addition, the measured $\Delta E$ will be shifted away from zero in a correlated manner. We account for this by constructing a two-dimensional veto region in the $m_{\ell^+\ell^-}$ vs. $\Delta E$ plane as shown in Fig. 4; the simulated points plotted demonstrate the expected background rejection. Within the veto region in data we find approximately 13700 $J/\psi$ events and 1000 $\psi(2S)$ events summed over all decay modes. These provide a high-statistics control sample useful for evaluating systematic uncertainties and selection efficiencies. The residual charmonium background after applying the veto is estimated from simulation to be between 0.0 and 1.6 events per decay mode.

Because of the 2–3% probability for misidentifying pions as muons, the $B \to K^{(*)} \mu^+ \mu^-$ channels also receive a significant peaking background contribution from hadronic $B$ decays. The largest of these are $B^+ \to \bar{D}^0 \pi^-$ where $D^0 \to K^- \pi^+$ or $D^0 \to K^+ \pi^-$, and $\bar{B}^0 \to D^+ \pi^-$.

092001-7
where \( D^+ \rightarrow \bar{K}^{*0}\pi^+ \). These are suppressed by removing events in which the \( K^{(*)}\mu \) invariant mass lies in the range \( 1.84 < m_{K^{(*)}\mu} < 1.90 \text{ GeV/}c^2 \). The remaining hadronic backgrounds come from charmless decays such as \( B \rightarrow K^{(*)}\pi^+\pi^- \), \( B \rightarrow K^{(*)}K^+\pi^- \), and \( B \rightarrow K^{(*)}K^+K^- \). We measure the peaking background from these processes using data control samples of \( B \rightarrow K^{(*)}\mu\mu \) events. These samples are selected with the same requirements as signal events, except hadron identification is required for the hadron candidate \( h \) in place of muon identification. This yields samples of predominantly hadronic \( B \) decays. We then weight each event by the muon misidentification rate for the hadron divided by its hadron identification efficiency. The hadronic peaking background is then extracted by a fit to the \( m_{ES} \) distribution of these weighted events. This results in a total hadronic peaking background measurement of 0.4–2.3 events per muon decay channel. These backgrounds are suppressed by a factor of approximately 400 in the \( B \rightarrow K^{(*)}\epsilon^+\epsilon^- \) channels due to the much lower probability of misidentifying pions as electrons.

There is an additional contribution to the peaking backgrounds in the electron channels from rare two-body decays. These include \( B \rightarrow K^*\gamma \) with the \( \gamma \) converting to an \( e^+e^- \) pair in the detector, and \( B \rightarrow K^{(*)}\pi^0 \) or \( B \rightarrow K^{(*)}\eta \), where the \( \pi^0 \) or \( \eta \) undergoes a Dalitz decay to \( e^+e^-\gamma \). These backgrounds are estimated from simulation to contribute 0.0–1.4 events per electron decay channel.

The sum of peaking backgrounds from all sources is summarized in Table I. As a function of \( q^2 \), all of the backgrounds from \( K^*\gamma \) and \( K^{(*)}\pi^0 \) are localized in the region \( 0.0 < q^2 < 0.1 \text{ GeV}^2/c^4 \). Backgrounds from \( J/\psi \) and \( K^{(*)}\eta \) populate the region \( 0.1 < q^2 < 8.41 \text{ GeV}^2/c^4 \), while the \( \psi(2S) \) backgrounds contribute only to the region \( q^2 > 10.24 \text{ GeV}^2/c^4 \). The hadronic backgrounds occupy both \( 0.1 < q^2 < 8.41 \text{ GeV}^2/c^4 \) and \( q^2 > 10.24 \text{ GeV}^2/c^4 \) regions.

### V. YIELD EXTRACTION PROCEDURE

We extract the signal yield and angular distributions using a multidimensional unbinned maximum-likelihood fit. For \( B \rightarrow K\epsilon^+\epsilon^- \), the total branching fraction is obtained from a two-dimensional fit to \( m_{ES} \) and \( \Delta E \). In the \( B \rightarrow K^*\ell^+\ell^- \) modes, we add the reconstructed \( K^* \) mass as a third fit variable. The signal shapes are parametrized in both \( m_{ES} \) and \( \Delta E \) by a Gaussian function plus a radiative tail described by an exponential power function. This takes the form

\[
\text{yield}(m_{ES}, \Delta E) = A \exp(-\frac{(m_{ES} - m_0)^2}{2\sigma_0^2}) + B \exp(-\frac{\Delta E^2}{2\sigma_{\Delta E}^2})
\]
The backgrounds that peak only in $m_{\pi}^{0.0025}$ and $m_{\pi}$ and the exponent of the power function, respectively. The $m_{ES}$ shape parameters $\tilde{x}$, $\sigma$, $\alpha$, and $n$ are assumed to have a $\Delta E$ dependence of the form $c_0 + c_2(\Delta E)^2$, determined empirically from simulation. The mean and width are fixed to the values derived by fitting the control sample of vetoed charmonium events. All other signal shape parameters are fixed to the values obtained from fits to simulated signal events. In the $B \rightarrow K^* \ell^+ \ell^-$ mode, the mass of the $K^*$ is parametrized with a relativistic Breit-Wigner line shape.

The background is modeled as a sum of terms describing (i) combinatorial background; (ii) peaking background; (iii) cross-feed backgrounds; and, (iv) in the $B \rightarrow K^* \ell^+ \ell^-$ modes, backgrounds that peak in $m_{K^*}$ at the $K^*$ mass but not in $m_{ES}$ and $\Delta E$. The combinatorial background is described by a product of an empirically derived threshold function in $m_{ES}$, a linear term in $\Delta E$, and the product of $\sqrt{m_{K^*} - m_K - m_\pi}$ and a quadratic function of $m_{K^*}$ for the $K^*$ modes. The form of the threshold function used to describe the background in $m_{ES}$ is $f(x) \propto x \sqrt{1 - x^2} \exp[-c(1 - x^2)]$, where $c$ is a fit parameter and $x = m_{ES}/E_{b}$. The peaking background component has the same shape as the signal, with normalization fixed to the estimates of the mean peaking backgrounds (Table I). The cross-feed component has a floating normalization to describe (a) background in $B \rightarrow K \ell^+ \ell^-$ ($B \rightarrow K^* \ell^+ \ell^-$) from $B \rightarrow K^* \ell^+ \ell^-$ ($B \rightarrow K^* \pi \ell^+ \ell^-$) events with a lost pion, and (b) background in $B \rightarrow K^* \ell^+ \ell^-$ from $B \rightarrow K^* \ell^+ \ell^-$ events with a randomly added pion. The backgrounds that peak only in $m_{K^*}$ are described by the signal shape in $m_{K^*}$ and the combinatorial background shape in $m_{ES}$ and $\Delta E$. The yield of this term is fixed to $((5 \pm 5)\%)$ of the total combinatorial background, as determined from simulation. As the shape parameters for term (i) and the normalizations for terms (ii) and (iii) are all free parameters of the fit, much of the background uncertainty propagates into the statistical uncertainty in the signal yield obtained from the fit.

The CP asymmetry is also extracted from the fit in the $B^+ \rightarrow K^* \ell^+ \ell^-$ and $B^0 \rightarrow K^* \ell^+ \ell^-$ channels, where the flavor of the $b$ quark can be inferred from the charge of the final state $K^{(*)}$ hadron. As this cannot be done in the case of $B^0 \rightarrow K^0 \ell^+ \ell^-$, we do not measure the CP asymmetry in that mode. The possibility of a nonzero CP asymmetry in the combinatorial background is accounted for by allowing its value to float in the fit. The CP asymmetry of the peaking background is fixed to the value expected from the relative composition of background sources.

The partial branching fractions are measured by repeating the fit with the sample partitioned into $q^2$ bins. The signal efficiencies and peaking backgrounds are recomputed for each region of $q^2$. To determine the forward-backward asymmetry and $K^*$ polarization in bins of $q^2$, we also utilize fits to the $\cos \theta^*$ and $\cos \theta_{K^*}$ angular distributions. We follow the treatment of Ref. [21] to parametrize the angular distributions for signal. The signal shape in $\cos \theta^*$ is described by an underlying differential distribution which depends on the fraction of longitudinal polarization $F_L$ as

$$\frac{1}{\Gamma} \int d\cos \theta_K = \frac{3}{4} F_L \cos^2 \theta_K + \frac{3}{4} (1 - F_L) (1 - \cos^2 \theta_K).$$

The underlying differential rate for signal in $\cos \theta^*$ is then described in terms of $F_S$ and the forward-backward asymmetry term $A_{FB}$ which enters linearly in $\cos \theta^*$:

$$\frac{1}{\Gamma} \int d\cos \theta^* = \frac{3}{4} F_L (1 - \cos^2 \theta^*) + \frac{3}{4} (1 - F_L) (1 + \cos^2 \theta^*) + A_{FB} \cos \theta^*.$$
VI. SYSTEMATIC UNCERTAINTIES

A. Branching fractions

In evaluating systematic uncertainties in the branching fractions, we consider both errors that affect the signal efficiency estimate, and errors arising from the maximum-likelihood fit. Sources of uncertainties that affect the efficiency are: charged-particle tracking (0.8% per lepton, 1.4% per charged hadron), charged-particle identification (0.5% per electron pair, 1.3% per muon pair, 0.2% per pion, 0.6% per kaon), the continuum background suppression selection (0.3%–2.2% depending on the mode), the \( B\bar{B} \) background suppression selection (0.6%–2.1%), \( K^0_S \) selection (0.9%), and signal simulation statistics (0.4%–0.7%). The estimated number of \( B\bar{B} \) events in our data sample has an uncertainty of 1.1%. We use the high-statistics sample of events that fail the charmonium veto to bound the systematic uncertainties associated with the continuum suppression Fisher discriminant, the \( B\bar{B} \) likelihood suppression selection, and charged-particle identification. The Fisher discriminant and \( B\bar{B} \) likelihood ratio for \( B^+ \to K^+ e^+ e^- \) are illustrated in Fig. 5 for data and simulation in the \( J/\psi \) control sample. An additional systematic uncertainty in the efficiency results from the choice of form factor model, which alters the \( q^2 \) distribution of the signal. We take this uncertainty to be the maximum efficiency variation obtained from a set of recent models [7,8,10,26,27]; the uncertainty is computed separately for each mode and varies in size from 1.1% to 8.3%.

Systematic uncertainties on the signal yields obtained from the maximum-likelihood fit arise from three sources: uncertainties in the parameters describing the signal shapes, uncertainties in the combinatorial background shape, and uncertainties in the peaking backgrounds. The uncertainties in the means and widths of the signal shapes are obtained by comparing data and simulated data in \( B \to J/\psi K^{(*)} \) control samples. For modes with electrons, we also vary the fraction of signal events in the tail of the \( \Delta E \) distribution by varying the exponent \( n \) in the exponential power function. Signal shape uncertainties are typically 2–4% of the signal yield. To evaluate the uncertainty due to the background shape, we reevaluate the fit yields with three different parametrizations: (i) an exponential shape for \( \Delta E \), (ii) a quadratic shape for \( \Delta E \), and (iii) an \( m_{ES} \) background shape parameter \( \xi \) which is linearly correlated with \( \Delta E \). In modes with a \( K^* \), we also vary the yield of the background component which peaks in \( m_{ES} \) but not in \( m_{ES} \) or \( \Delta E \) by 100% of itself. The induced uncertainty in the signal yield due to the background shape is 4–6% for \( B \to K^0 \ell^+ \ell^- \) modes and increases to 8–12% for \( B \to K^{(*)} \ell^+ \ell^- \) modes, where the backgrounds are generally larger. Uncertainties in the peaking background induce an uncertainty in the signal yields of 2–5%; this is obtained by varying the expected peaking background yield within its \( \pm 1\sigma \) uncertainties. The total systematic uncertainty in the fitted signal yield induces a systematic uncertainty \( \Delta B_{\text{fit}} \) in the measured branching fraction; this uncertainty is shown for each of the branching fraction fits in Tables II and III.

B. \( CP \) asymmetry

The systematic uncertainties in the measurement of \( A_{CP} \) include errors due both to detector efficiency effects and to the asymmetry in the peaking background component. The error associated with the detector efficiency is obtained by comparing the value of \( A_{CP} \) measured in the charmonium control samples with the expected value of zero; agreement with zero is obtained with a precision of 1.2% for \( B^+ \to K^+ \ell^+ \ell^- \) and 2.1% for \( B \to K^{(*)} \ell^+ \ell^- \). The uncertainty due to the peaking background is evaluated by varying the expected \( CP \) asymmetry of the peaking backgrounds within their uncertainties. The possible \( CP \) asymmetry in the charmonium and \( B \to K^{(*)} \gamma \) peaking backgrounds is highly constrained from previous measurements; any asymmetry in the Dalitz decays is suppressed by their relatively small contribution to the peaking background.
TABLE II. Results from fits to the individual $K^{(*)}\ell^+\ell^-$ decay modes for all $q^2$. The columns from left are: decay mode, fitted signal yield, signal efficiency, relative uncertainty on the branching fraction due to the systematic error on the efficiency estimate, systematic error from the fit, and the resulting branching fraction (with statistical and systematic errors).

<table>
<thead>
<tr>
<th>Mode</th>
<th>Yield</th>
<th>$\epsilon$ (%)</th>
<th>$\Delta B_{\text{eff}}$ (%)</th>
<th>$\Delta B_{\text{fit}}$ (10^{-6})</th>
<th>$B$ (10^{-6})</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K^+\mu^+\mu^-$</td>
<td>$10.9^{+5.4}_{-4.3}$</td>
<td>15.4</td>
<td>$\pm 4.1$</td>
<td>$\pm 0.03$</td>
<td>$0.31^{+0.15}_{-0.12}$</td>
</tr>
<tr>
<td>$K^0\ell^+\ell^-$</td>
<td>$2.4^{+3.8}_{-2.0}$</td>
<td>22.8</td>
<td>$\pm 9.6$</td>
<td>$\pm 0.01$</td>
<td>$0.13^{+0.10}_{-0.07}$</td>
</tr>
<tr>
<td>$K^0\mu^+\mu^-$</td>
<td>$6.3^{+3.6}_{-2.8}$</td>
<td>13.6</td>
<td>$\pm 8.3$</td>
<td>$\pm 0.04$</td>
<td>$0.59^{+0.33}_{-0.26}$</td>
</tr>
<tr>
<td>$K^{*0}\ell^+\ell^-$</td>
<td>$29.4^{+9.5}_{-8.4}$</td>
<td>18.6</td>
<td>$\pm 4.9$</td>
<td>$\pm 0.09$</td>
<td>$1.04^{+0.33}_{-0.29}$</td>
</tr>
<tr>
<td>$K^{*0}\mu^+\mu^-$</td>
<td>$15.9^{+7.0}_{-5.9}$</td>
<td>11.9</td>
<td>$\pm 5.8$</td>
<td>$\pm 0.11$</td>
<td>$0.87^{+0.38}_{-0.33}$</td>
</tr>
<tr>
<td>$K^{*+}\ell^+\ell^-$</td>
<td>$6.1^{+6.3}_{-5.3}$</td>
<td>15.7</td>
<td>$\pm 6.8$</td>
<td>$\pm 0.37$</td>
<td>$0.75^{+0.76}_{-0.58}$</td>
</tr>
<tr>
<td>$K^{*+}\mu^+\mu^-$</td>
<td>$4.7^{+6.8}_{-3.4}$</td>
<td>9.3</td>
<td>$\pm 7.1$</td>
<td>$\pm 0.13$</td>
<td>$0.97^{+0.94}_{-0.69}$</td>
</tr>
</tbody>
</table>

TABLE III. Results from fits to combined $K^{(*)}\ell^+\ell^-$ decay modes for all $q^2$. The columns from left are: decay mode combination, fitted signal yield, relative uncertainty on the branching fraction due to the systematic error on the efficiency estimate, systematic error on the branching fraction introduced by the systematic error on the fitted signal yield, and the resulting branching fraction (with statistical and systematic errors). The constraints for each combined fit are described in the text.

<table>
<thead>
<tr>
<th>Mode</th>
<th>Yield (events)</th>
<th>$\Delta B_{\text{eff}}$ (%)</th>
<th>$\Delta B_{\text{fit}}$ (10^{-6})</th>
<th>$B$ (10^{-6})</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K\ell^+\ell^-$</td>
<td>$36.7^{+8.8}_{-7.9}$</td>
<td>$\pm 3.7$</td>
<td>$\pm 0.02$</td>
<td>$0.35^{+0.13}_{-0.11}$</td>
</tr>
<tr>
<td>$K\mu^+\mu^-$</td>
<td>$8.2^{+4.5}_{-4.6}$</td>
<td>$\pm 9.0$</td>
<td>$\pm 0.02$</td>
<td>$0.29^{+0.16}_{-0.13}$</td>
</tr>
<tr>
<td>$K^{*}\ell^+\ell^-$</td>
<td>$45.5^{+8.9}_{-8.9}$</td>
<td>$\pm 4.6$</td>
<td>$\pm 0.02$</td>
<td>$0.34^{+0.07}_{-0.06}$</td>
</tr>
<tr>
<td>$K^{*}\mu^+\mu^-$</td>
<td>$36.2^{+11.2}_{-10.0}$</td>
<td>$\pm 5.2$</td>
<td>$\pm 0.13$</td>
<td>$0.97^{+0.30}_{-0.27}$</td>
</tr>
<tr>
<td>$K^{*}\ell^+\ell^-$</td>
<td>$45.3^{+10.6}_{-10.5}$</td>
<td>$\pm 5.0$</td>
<td>$\pm 0.08$</td>
<td>$0.81^{+0.21}_{-0.19}$</td>
</tr>
<tr>
<td>$K^{*+}\ell^+\ell^-$</td>
<td>$11.5^{+8.0}_{-8.6}$</td>
<td>$\pm 6.6$</td>
<td>$\pm 0.20$</td>
<td>$0.73^{+0.50}_{-0.42}$</td>
</tr>
<tr>
<td>$K^{*+}\ell^+\ell^-$</td>
<td>$57.1^{+13.7}_{-12.5}$</td>
<td>$\pm 5.3$</td>
<td>$\pm 0.10$</td>
<td>$0.78^{+0.19}_{-0.17}$</td>
</tr>
<tr>
<td>Pole excluded</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K^{*}\ell^+\ell^-$</td>
<td>$23.6^{+2.4}_{-2.3}$</td>
<td>$\pm 5.2$</td>
<td>$\pm 0.11$</td>
<td>$0.63^{+0.25}_{-0.22}$</td>
</tr>
<tr>
<td>$K^{*}\mu^+\mu^-$</td>
<td>$20.7^{+8.1}_{-7.0}$</td>
<td>$\pm 5.9$</td>
<td>$\pm 0.11$</td>
<td>$0.88^{+0.34}_{-0.30}$</td>
</tr>
<tr>
<td>$K^{*}\ell^+\ell^-$</td>
<td>$34.8^{+10.4}_{-9.3}$</td>
<td>$\pm 5.0$</td>
<td>$\pm 0.10$</td>
<td>$0.75^{+0.22}_{-0.20}$</td>
</tr>
<tr>
<td>$K^{*+}\ell^+\ell^-$</td>
<td>$9.5^{+7.0}_{-5.7}$</td>
<td>$\pm 6.6$</td>
<td>$\pm 0.19$</td>
<td>$0.73^{+0.53}_{-0.44}$</td>
</tr>
<tr>
<td>$K^{*+}\ell^+\ell^-$</td>
<td>$44.3^{+12.2}_{-11.1}$</td>
<td>$\pm 5.3$</td>
<td>$\pm 0.11$</td>
<td>$0.73^{+0.20}_{-0.18}$</td>
</tr>
</tbody>
</table>

In contrast, the hadronic peaking background in the muon modes could exhibit a significant $CP$ asymmetry; this is measured directly from the asymmetry of the hadronic control sample described in Section IV B with an uncertainty dominated by the statistics of the sample. This induces an uncertainty in the measured $A_{CP}$ of 1% for $B^+ \rightarrow K^+ \ell^+\ell^-$ and 2% for $B \rightarrow K^0 \ell^+\ell^-$. Other systematic uncertainties induced by the fitting procedure, as computed above for the branching fraction measurements, are found to be negligible.

C. Angular distributions

Systematic uncertainties related to the angular distributions of the efficiency are estimated by comparing the values of $A_{FB}$, $F_5$, and $F_L$ measured in the relevant charmonium control samples with their expected values. For $B \rightarrow J/\psi K^*$ and $B \rightarrow J/\psi K^{*0}$, we measure an $A_{FB}$ consistent with zero and with a precision of 0.01 and 0.02, respectively. For $B \rightarrow J/\psi K^{*+}$, we measure $F_L$ to be consistent with the previous BABAR measurement [28], with a preci-
sion of 0.05. For $B \to J/\psi K$ we measure $F_S$ consistent with zero and with a precision of 0.03.

Further systematic uncertainties are evaluated by repeating the fit with alternative shapes assumed for the background components: (i) the shape of misreconstructed signal events is fixed instead to the shape of correctly reconstructed signal, (ii) the combinatorial background shape is drawn from alternative ranges of $m_{ES}$ and $\Delta E$, and from the sample of events that fail the $BB$ likelihood selection, and (iii) the angular distributions of the peaking backgrounds are varied within their statistical uncertainties. Systematic uncertainties from backgrounds induce uncertainties in $F_L$ and $A_{FB}$ of 0.05–0.18, depending on the relative amount of background, and are the largest systematic uncertainty. $F_S$ is more sensitive to the background shape, with an induced systematic uncertainty of 0.45.

In the fit to $\cos \theta^*$ in the $B \to K^* \ell^+ \ell^-$ decay modes, the value of $F_L$ is fixed to the result obtained from the fit to the $\cos \theta_K$ distribution. This introduces an additional parametric uncertainty of 0.01 on the measured value of $A_{FB}$, which we evaluate by varying $F_L$ within the uncertainty of the measurement.

VII. RESULTS

A. Branching fractions

We first perform the fit separately for each of the eight decay modes to extract the branching fractions integrated over all $q^2$. In the branching fraction fits, the efficiency is defined such that the total branching fraction includes the estimated signal that is lost due to the charmonium vetos. The results for the individual decay modes are shown in Table II. We then perform a combined fit to the appropriate combinations of modes to extract the $B \to K^+ \ell^+ \ell^-$ and $B \to K^0 \ell^+ \ell^-$ branching fractions. We combine charged and neutral modes by constraining the total width ratio $\Gamma(B^0)/\Gamma(B^+)$ to the world average ratio of lifetimes $\tau(B^+)/\tau(B^0) = 1.071 \pm 0.009$ [29]. In the $B \to K^+ \ell^+ \ell^-$ mode, we add the additional constraint $\Gamma(B \to K^0 \mu^+ \mu^-)/\Gamma(B \to K^0 e^+ e^-) = 0.75$ to account for the enhancement due to the pole at low $q^2$ in the electron channel [4]. The final branching fractions are expressed in terms of the $B^0 \to K^{(*)0} \mu^+ \mu^-$ channels. With these constraints, we find the lepton-flavor averaged, $B$-charge-averaged branching fractions

$$\mathcal{B}(B \to K^+ \ell^+ \ell^-) = (0.34 \pm 0.07 \pm 0.02) \times 10^{-6},$$

$$\mathcal{B}(B \to K^0 \ell^+ \ell^-) = (0.78^{+0.19}_{-0.17} \pm 0.11) \times 10^{-6},$$

where the first error is statistical and the second systematic. The projections of the data overlayed with the combined fit results are shown in Figs. 6 and 7. The signal significance is computed as $\sqrt{2\Delta \ln(L)}$, where $\Delta \ln(L)$ is the difference between the likelihood of the best fit and that of the null signal hypothesis. Systematic uncertainties are incorporated in the significance estimate by simultaneously applying all variations that result in a lower signal yield before computing the change in likelihood. The significance of the signal including statistical and systematic uncertainties is 6.6 standard deviations for the $B \to K \ell^+ \ell^-$ mode and 5.7 standard deviations for the $B \to K^* \ell^+ \ell^-$ mode. The secondary peak in the $\Delta E$ sideband of $B \to K \ell^+ \ell^-$ results from the fit component describing events with a lost pion, either from $B \to K^* \ell^+ \ell^-$ or from events in which a $b \to s \ell^+ \ell^-$ decay results in a $K \pi \ell^+ \ell^-$ final state without

![Figure 6](092001-12)

**Figure 6** (color online). Distributions of the fit variables in $K \ell^+ \ell^-$ data (points), compared with projections of the combined fit (curves): (a) $m_{ES}$ distribution after requiring $-0.11 < \Delta E < 0.05$ GeV and (b) $\Delta E$ distribution after requiring $|m_{ES} - m_B| < 6.6$ MeV/$c^2$. The solid curve is the sum of all fit components, including signal; the dashed curve is the sum of all background components.

![Figure 7](092001-12)

**Figure 7** (color online). Distributions of the fit variables in $K^* \ell^+ \ell^-$ data (points), compared with projections of the combined fit (curves): (a) $m_{ES}$ after requiring $-0.11 < \Delta E < 0.05$ GeV and $0.817 < m_{K^*} < 0.967$ GeV/$c^2$, (b) $\Delta E$ after requiring $|m_{ES} - m_B| < 6.6$ MeV/$c^2$, $0.817 < m_{K^*} < 0.967$ GeV/$c^2$, and (c) $m_{K^*}$ after requiring $|m_{ES} - m_B| < 6.6$ MeV/$c^2$ and $-0.11 < \Delta E < 0.05$ GeV. The solid curve is the sum of all fit components, including signal; the dashed curve is the sum of all background components.
proceeding through an intermediate $K^*$ resonance. The normalization and mean $\Delta E$ of this component are free parameters in the fit. Examination of these events shows that the addition of a charged or neutral pion in results in $B \rightarrow K^+ \ell^+ \ell^-$ or $B \rightarrow K^0 \pi^+ \ell^+ \ell^-$ signal candidate. Using simulated signal decays, we find the effect of these events on the $B \rightarrow K^+ \ell^+ \ell^-$ signal yield is negligible.

We further perform a set of combined fits with the sample partitioned into final states containing muons and electrons, and into charged and neutral final states, modifying the constraints as appropriate. The results from all such fits are summarized in Table III.

If the pole region is removed by requiring $q^2 > 0.1 \text{ GeV}^2/c^4$, the constrained ratio between $B \rightarrow K^* \mu^+ \mu^-$ and $B \rightarrow K^* e^+ e^-$ in the combined fit is modified from 0.75 to 1. Repeating the combined fit with this modification, we obtain

$$B(B \rightarrow K^* \ell^+ \ell^-)_{(q^2>0.1 \text{ GeV}^2/c^4)} = (0.73^{+0.20}_{-0.18} \pm 0.11) \times 10^{-6}.$$ 

The results of the combined fits in the various subsamples with the pole region removed are shown in Table III. We observe good agreement in the branching fraction obtained in all of the subsamples, both with and without the pole region included. The measured total rates are consistent with the range of standard model rates predicted in Ref. [4]. The $B \rightarrow K^+ \ell^+ \ell^-$ rate is significantly lower than the range given by Ref. [12].

From the separate fits to the muon and electron channels integrated over all $q^2$, we obtain the ratios

$$R_K = 1.06 \pm 0.48 \pm 0.08,$$

consistent with the SM predictions of 1.00 and 0.75, respectively. If instead the pole region is excluded from the $B \rightarrow K^+ \ell^+ \ell^-$ channels, we find

$$R_{K^*,(q^2>0.1 \text{ GeV}^2/c^4)} = 1.40 \pm 0.78 \pm 0.10,$$

where this ratio is expected to be 1 in the SM.

### B. CP asymmetry

From the fit to the combined modes integrated over all $q^2$, we find the direct $CP$ asymmetries

$$A_{CP}(B^+ \rightarrow K^+ \ell^+ \ell^-) = -0.07 \pm 0.22 \pm 0.02,$$

$$A_{CP}(B \rightarrow K^+ \ell^+ \ell^-) = +0.03 \pm 0.23 \pm 0.03,$$

where the first error is statistical and the second systematic. The measured values in both channels are consistent with the SM expectation of a negligible direct $CP$ asymmetry.

### C. Partial branching fractions

The partial branching fractions obtained from the fits to $m_{FS}$, $\Delta E$, and $m_K \pi$ in two bins of $q^2$ are shown in Table IV. The results are generally consistent with the $q^2$ dependence predicted in recent standard model based form factor calculations (Fig. 8).

### D. $K^*$ polarization

The fit projections for the $\cos \theta_K$ distribution in bins of $q^2$ are shown in Fig. 12 of Appendix A. The resulting values for the fraction of longitudinal polarization $F_L$ are

<table>
<thead>
<tr>
<th>$q^2$ (GeV$^2$/c$^4$)</th>
<th>$B$ (10$^{-6}$)</th>
<th>$B \rightarrow K^+ \ell^+ \ell^-$</th>
<th>$F_L$</th>
<th>$A_{FB}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1 − 8.41</td>
<td>0.27$^{+0.12}_{-0.10}$ ± 0.05</td>
<td>0.77$^{+0.63}_{-0.30}$ ± 0.07</td>
<td>0.19 (95%CL)</td>
<td></td>
</tr>
<tr>
<td>&gt;10.24</td>
<td>0.37$^{+0.13}_{-0.11}$ ± 0.05</td>
<td>0.51$^{+0.22}_{-0.25}$ ± 0.08</td>
<td>0.72$^{+0.28}_{-0.26}$ ± 0.08</td>
<td></td>
</tr>
<tr>
<td>&gt;0.1</td>
<td>0.73$^{+0.20}_{-0.18}$ ± 0.11</td>
<td>0.63$^{+0.18}_{-0.19}$ ± 0.05</td>
<td>0.55 (95%CL)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$q^2$ (GeV$^2$/c$^4$)</th>
<th>$B$ (10$^{-6}$)</th>
<th>$B \rightarrow K^0 \ell^+ \ell^-$</th>
<th>$F_S$</th>
<th>$A_{FB}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1 − 8.41</td>
<td>0.10$^{+0.04}_{-0.04}$ ± 0.01</td>
<td>0</td>
<td>−0.49$^{+0.51}_{-0.99}$ ± 0.18</td>
<td></td>
</tr>
<tr>
<td>&gt;10.24</td>
<td>0.22$^{+0.05}_{-0.05}$ ± 0.02</td>
<td>0</td>
<td>0.26$^{+0.23}_{-0.24}$ ± 0.03</td>
<td></td>
</tr>
<tr>
<td>&gt;0.1</td>
<td>0.34$^{+0.07}_{-0.07}$ ± 0.02</td>
<td>0.81$^{+0.58}_{-0.61}$ ± 0.46</td>
<td>0.15$^{+0.21}_{-0.23}$ ± 0.08</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$q^2$ (GeV$^2$/c$^4$)</th>
<th>$B$ (10$^{-6}$)</th>
<th>$B \rightarrow K^+ \ell^+ \ell^-$</th>
<th>$F_L$</th>
<th>$A_{FB}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1 − 8.41</td>
<td>0.27$^{+0.12}_{-0.10}$ ± 0.05</td>
<td>0.77$^{+0.63}_{-0.30}$ ± 0.07</td>
<td>0.19 (95%CL)</td>
<td></td>
</tr>
<tr>
<td>&gt;10.24</td>
<td>0.37$^{+0.13}_{-0.11}$ ± 0.05</td>
<td>0.51$^{+0.22}_{-0.25}$ ± 0.08</td>
<td>0.72$^{+0.28}_{-0.26}$ ± 0.08</td>
<td></td>
</tr>
<tr>
<td>&gt;0.1</td>
<td>0.73$^{+0.20}_{-0.18}$ ± 0.11</td>
<td>0.63$^{+0.18}_{-0.19}$ ± 0.05</td>
<td>0.55 (95%CL)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$q^2$ (GeV$^2$/c$^4$)</th>
<th>$B$ (10$^{-6}$)</th>
<th>$B \rightarrow K^0 \ell^+ \ell^-$</th>
<th>$F_S$</th>
<th>$A_{FB}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1 − 8.41</td>
<td>0.10$^{+0.04}_{-0.04}$ ± 0.01</td>
<td>0</td>
<td>−0.49$^{+0.51}_{-0.99}$ ± 0.18</td>
<td></td>
</tr>
<tr>
<td>&gt;10.24</td>
<td>0.22$^{+0.05}_{-0.05}$ ± 0.02</td>
<td>0</td>
<td>0.26$^{+0.23}_{-0.24}$ ± 0.03</td>
<td></td>
</tr>
<tr>
<td>&gt;0.1</td>
<td>0.34$^{+0.07}_{-0.07}$ ± 0.02</td>
<td>0.81$^{+0.58}_{-0.61}$ ± 0.46</td>
<td>0.15$^{+0.21}_{-0.23}$ ± 0.08</td>
<td></td>
</tr>
</tbody>
</table>

092001-13
listed in Table IV. Combining all events with $q^2 > 0.1$ GeV$^2$/c$^4$, we find

$$F_L(B \rightarrow K^+ \ell^+ \ell^-)_{(q^2 > 0.1 \text{GeV}^2/c^4)} = 0.63^{+0.18}_{-0.19} \pm 0.05,$$

where the first error is statistical, and the second systematic.

The measured values of $F_L$ are consistent with the SM expectation in both $q^2$ ranges (Fig. 9) and integrated over all $q^2 > 0.1$ GeV$^2$/c$^4$. However, the large statistical uncertainties do not allow the determination of the sign of $C_7$ from this measurement at present.

### E. Lepton forward-backward asymmetry

The fit projections for the $\cos \theta^*$ distribution in the $B^+ \rightarrow K^+ \ell^+ \ell^-$ mode are shown in Fig. 13 of Appendix A. Combining all events with $q^2 > 0.1$ GeV$^2$/c$^4$, we find for the $B^+ \rightarrow K^+ \ell^+ \ell^-$ mode

$$A_{FB}(B^+ \rightarrow K^+ \ell^+ \ell^-)_{(q^2 > 0.1 \text{GeV}^2/c^4)} = 0.15^{+0.21}_{-0.23} \pm 0.08,$$

$$F_S(B^+ \rightarrow K^+ \ell^+ \ell^-)_{(q^2 > 0.1 \text{GeV}^2/c^4)} = 0.81^{+0.58}_{-0.61} \pm 0.46,$$

where the first errors are statistical, and the second systematic. The correlation coefficient between these two measurements is $+0.23$. Both $A_{FB}$ and $F_S$ are consistent with the SM prediction of zero. As a cross-check, we have also performed similar fits in the low and high $q^2$ regions for $A_{FB}$, where due to limited statistics $F_S$ must be fixed to zero; the resulting asymmetries are $-0.49^{+0.51}_{-0.99} \pm 0.18$ and $0.26^{+0.24}_{-0.23} \pm 0.03$, respectively, which again are both consistent with zero asymmetry.

The fit projections for the $\cos \theta^*$ distribution in the $B \rightarrow K^+ \ell^+ \ell^-$ mode are shown in Fig. 14 of Appendix A, and the resulting values of $A_{FB}$ listed in Table IV. We find a large positive asymmetry in the high $q^2$ region, consistent with the SM expectation. This disfavors new physics scenarios.
in which the product of the $C_{ij}^{\text{eff}}$ and $C_{ij}^{\text{eff}}$ Wilson coefficients have the same magnitude but opposite relative sign as in the SM, which would result in a large negative asymmetry at high $q^2$ (Fig. 9).

For the low $q^2$ region and the region integrated over all $q^2 > 0.1$ GeV$^2/c^4$, the $A_{FB}$ value corresponding to the maximum-likelihood experiments in the ensemble have a maximum likelihood asymmetry at high $q^2$ (Fig. 9).

The uncertainties in the other PDF parameters are accounted for by varying them randomly for each generated experiment in the ensemble according to normal distributions determined by the parameters’ measured central values and uncertainties. We account for systematic uncertainties that do not correspond to continuous PDF parameters, such as the choice of combinatorial background PDFs for $\cos\theta^*$, by generating ensembles for each PDF variation and choosing that which results in the lowest lower limit. With this method, we find $A_{FB} > 0.19$ at 95% CL for the low $q^2$ region. Combining all events with $q^2 > 0.1$ GeV$^2/c^4$, we find for the $B \rightarrow K^*\ell^+\ell^-$ mode at 95% CL

\[ A_{FB}(B \rightarrow K^*\ell^+\ell^-)(q^2 > 0.1 \text{GeV}^2/c^4) > 0.55. \]

The corresponding fit projections shown in Fig. 14 are produced by fixing the $A_{FB}$ of the signal component to its maximum physical value.

**F. Search for lepton-flavor-violation**

We extract the signal yield in the $B \rightarrow K\ell\mu$ and $B \rightarrow K^*\ell\mu$ final states in a similar manner as the $K^0(\ell^+\ell^-)$ decays, with the particle identification requirements modified to select $e^\pm\mu^\pm$ pairs. The signal efficiencies for these

![Image](https://example.com/image1.png)

**FIG. 11 (color online).** Distributions of the fit variables in $K\ell\mu$ data (points), compared with projections of the combined fit (curves): (a) $m_{ES}$ distribution after requiring $-0.11 < \Delta E < 0.05$ GeV and $0.817 < m_{ES} < 0.967$ GeV/c$^2$, (b) $\Delta E$ distribution after requiring $|m_{ES} - m_B| < 6.6$ MeV/c$^2$, and (c) $m_{K\pi}$ after requiring $|m_{ES} - m_B| < 6.6$ MeV/c$^2$ and $-0.11 < \Delta E < 0.05$ GeV. The solid curve is the sum of all fit components, including signal; the dashed curve is the sum of all background components.

![Image](https://example.com/image2.png)

**FIG. 10 (color online).** Distributions of the fit variables in $K\ell\mu$ data (points), compared with projections of the combined fit (curves): (a) $m_{ES}$ distribution after requiring $-0.11 < \Delta E < 0.05$ GeV and $0.817 < m_{ES} < 0.967$ GeV/c$^2$, (b) $\Delta E$ distribution after requiring $|m_{ES} - m_B| < 6.6$ MeV/c$^2$. The solid curve is the sum of all fit components, including signal; the dashed curve is the sum of all background components.

### TABLE V. Results from fits to lepton-flavor-violating decay modes

<table>
<thead>
<tr>
<th>Mode</th>
<th>Yield (%)</th>
<th>$\epsilon$ (%)</th>
<th>$B$ (10$^{-8}$)</th>
<th>$B$ UL (10$^{-8}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K^+ e^+\mu^-$</td>
<td>-3.5$^{+2.1}_{-1.4}$</td>
<td>12.6</td>
<td>$-12.1^{+7.4}_{-3.0}$</td>
<td>$\pm 2.3$</td>
</tr>
<tr>
<td>$K^+ e^-\mu^+$</td>
<td>-0.8$^{+1.6}_{-1.4}$</td>
<td>12.6</td>
<td>$-2.9^{+7.4}_{-4.4}$</td>
<td>$\pm 1.9$</td>
</tr>
<tr>
<td>$K^+ e^+\mu^-$</td>
<td>-3.2$^{+2.7}_{-1.7}$</td>
<td>12.6</td>
<td>$-11.9^{+1.3}_{-5.9}$</td>
<td>$\pm 3.2$</td>
</tr>
<tr>
<td>$K^0 e^\mu$</td>
<td>-2.9$^{+1.3}_{-1.4}$</td>
<td>12.5</td>
<td>$-30^{+19}_{-15}$</td>
<td>$\pm 15$</td>
</tr>
<tr>
<td>$K^0 e^-\mu^-$</td>
<td>1.1$^{+5.6}_{-2.1}$</td>
<td>10.4</td>
<td>$7^{+23}_{-13}$</td>
<td>$\pm 5$</td>
</tr>
<tr>
<td>$K^0 e^-\mu^+$</td>
<td>-1.1$^{+3.5}_{-2.2}$</td>
<td>10.4</td>
<td>$-7^{+22}_{-14}$</td>
<td>$\pm 7$</td>
</tr>
<tr>
<td>$K^0 e^\mu$</td>
<td>0.9$^{+4.6}_{-2.9}$</td>
<td>10.4</td>
<td>$6^{+59}_{-18}$</td>
<td>$\pm 9$</td>
</tr>
<tr>
<td>$K^{+*} e^+\mu^-$</td>
<td>0.4$^{+3.3}_{-2.3}$</td>
<td>10.0</td>
<td>$9^{+44}_{-23}$</td>
<td>$\pm 22$</td>
</tr>
<tr>
<td>$K^{+*} e^-\mu^+$</td>
<td>-1.7$^{+3.3}_{-2.0}$</td>
<td>10.0</td>
<td>$-32^{+63}_{-38}$</td>
<td>$\pm 15$</td>
</tr>
<tr>
<td>$K^{+*} e^+\mu^-$</td>
<td>-0.2$^{+4.2}_{-3.1}$</td>
<td>10.0</td>
<td>$-4^{+80}_{-59}$</td>
<td>$\pm 32$</td>
</tr>
<tr>
<td>$K\ell\mu$</td>
<td>-4.9$^{+2.9}_{-1.4}$</td>
<td>-</td>
<td>$-121^{+20}_{-16}$</td>
<td>$\pm 3.0$</td>
</tr>
<tr>
<td>$K^{+*} e\mu$</td>
<td>1.0$^{+5.5}_{-3.7}$</td>
<td>-</td>
<td>$48^{+26}_{-17}$</td>
<td>$\pm 11$</td>
</tr>
</tbody>
</table>
modes we find combined lepton-charge averaged, and therefore set upper limits on these processes. For the results of the combined fits are shown in Figs. 10 and 11. We find no evidence for a signal in any of these channels, and therefore set upper limits on these processes. For the combined lepton-charge averaged, $B$-charge-averaged modes we find

$$\mathcal{B}(B \to Ke\mu) < 3.8 \times 10^{-8},$$

$$\mathcal{B}(B \to K^e \mu e) < 51 \times 10^{-8},$$

at 90% CL. These limits are significantly more stringent than those of previous searches [30,31].

VIII. CONCLUSIONS

We have measured the branching fractions, partial branching fractions, direct CP asymmetries, ratio of muons to electrons, fraction of longitudinal $K^*$ polarization, and lepton forward-backward asymmetries in the rare FCNC decays $B \to K\ell^+\ell^-$ and $B \to K^*\ell^+\ell^-$. The branching fraction, $A_{CP}$, $R_K$, and $F_L$ results are all consistent with the standard model predictions for these decays. The values of $A_{FB}$ and the scalar contribution $F_S$ measured in the $B^+ \to K^+\ell^+\ell^-$ channel are consistent with the expected value of zero. In the $B \to K^*\ell^+\ell^-$ channel the large positive value of $A_{FB}$ at high $q^2$ is consistent with the SM and disfavors new physics scenarios in which the relative sign of the product of the $C_9$ and $C_{10}$ Wilson coefficients is opposite that of the SM. At low $q^2$ a positive value of $A_{FB}$ is also favored, with a 95% CL lower limit that is slightly above the SM prediction, as derived using the form factor models of Refs. [10,27].

In addition, we have obtained upper limits on the lepton-flavor-violating decays $B \to Ke\mu$ and $B \to K^e \mu e$ that are approximately one order of magnitude lower than those of previous searches.

We note that the Belle collaboration has recently reported [32] a measurement of the integrated forward-backward asymmetries, finding $\hat{A}_{FB}(B^+ \to K^+\ell^+\ell^-) = 0.10 \pm 0.14 \pm 0.01$ and $\hat{A}_{FB}(B \to K^*\ell^+\ell^-) = 0.50 \pm 0.15 \pm 0.02$. From a fit to the $\cos\theta^*$ and $q^2$ distributions, they conclude that scenarios in which the product of $C_9$ and $C_{10}$ has the opposite sign as expected in the SM are disfavored, consistent with the results reported here.

All of the measurements reported here are limited by statistical uncertainties, and can be improved with the addition of more data.

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), IHEP (China), CEA and CNRS-IN2P3 (France), BMBF and DFG (Germany), INFN (Italy), FOM (The Netherlands), NFR (Norway), MIST (Russia), and PPARC (United Kingdom). Individuals have received support from CONACyT (Mexico), Marie Curie EIF (European Union), the A.P. Sloan Foundation, the Research Corporation, and the Alexander von Humboldt Foundation.

APPENDIX: FITS TO ANGULAR DISTRIBUTIONS

In this Appendix we present plots of the $\cos\theta_K$ and $\cos\theta^*$ distributions in data, together with the projections of the combined fits used to extract $F_L$ and $A_{FB}$. Figure 12 shows the fitted $\cos\theta_K$ distributions for each of the $q^2$ bins considered in this analysis. Figs. 13 and 14 display the fitted

![Fig. 12](color online). Distributions of the fit variable $\cos\theta_K$ in $B \to K^*\ell^+\ell^-$ data (points), compared with projections of the combined fit (curves) after requiring $-0.11 < \Delta E < 0.05$ GeV, $|m_{ES} - m_{B}| < 6.6$ MeV/c$^2$, and $0.817 < m_{K\ell\ell} < 0.967$ GeV/c$^2$. The solid curve is the sum of all fit components, the dashed curve is the sum of all background components, and the dot-dashed curve is the signal component. The $q^2$ regions (a) $0.1 < q^2 < 8.41$ GeV$^2$/c$^4$, (b) $q^2 > 10.24$ GeV$^2$/c$^4$, and (c) $q^2 > 0.1$ GeV$^2$/c$^4$ are shown.
FIG. 13 (color online). Distributions of the fit variable $\cos \theta^*$ in $B^+ \rightarrow K^+ \ell^+ \ell^-$ data (points), compared with projections of the combined fit (curves) after requiring $-0.11 < \Delta E < 0.05$ GeV and $|m_{ES} - m_B| < 6.6$ MeV/$c^2$. The solid curve is the sum of all fit components, the dashed curve is the sum of all background components, and the dot-dashed curve is the signal component. The $q^2$ regions (a) $0.1 < q^2 < 8.41$ GeV$^2$/c$^4$, (b) $q^2 > 10.24$ GeV$^2$/c$^4$, and (c) $q^2 > 0.1$ GeV$^2$/c$^4$ are shown. The combined fits shown for (a) and (b) are performed by fixing $F_S$ to zero.

FIG. 14 (color online). Distributions of the fit variable $\cos \theta^*$ in $B \rightarrow K^* \ell^+ \ell^-$ data (points), compared with projections of the combined fit (curves) after requiring $-0.11 < \Delta E < 0.05$ GeV, $|m_{ES} - m_B| < 6.6$ MeV/$c^2$, and $0.817 < m_{K^*} < 0.967$ GeV/$c^2$. The solid curve is the sum of all fit components, the dashed curve is the sum of all background components, and the dot-dashed curve is the signal component. The $q^2$ regions (a) $0.1 < q^2 < 8.41$ GeV$^2$/c$^4$, (b) $q^2 > 10.24$ GeV$^2$/c$^4$, and (c) $q^2 > 0.1$ GeV$^2$/c$^4$ are shown. The combined fits shown for (a) and (c) are performed by fixing $A_{PB}$ to its maximal physical value.

cos$\theta^*$ distributions for each of the $q^2$ ranges for the $B^+ \rightarrow K^+ \ell^+ \ell^-$ and $B \rightarrow K^* \ell^+ \ell^-$ decay modes, respectively. For the fits to the cos$\theta^*$ distributions in the $B \rightarrow K^* \ell^+ \ell^-$ mode, the $K^*$ polarization $F_L$ is fixed to its measured value, as described in the text. The deviations from a smooth parabolic shape in the signal component are the result of the efficiency and acceptance corrections, which are described by nonparametric histogram PDFs.