The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/128890

Please be advised that this information was generated on 2017-09-02 and may be subject to change.
Evidence for the Rare Decay $B \to K^\mp \ell^+ \ell^-$ and Measurement of the $B \to K^{*0} \ell^+ \ell^-$ Branching Fraction

We present evidence for the flavor-changing neutral current decay $B \to K^+ \ell^+ \ell^-$ and a measurement of the branching fraction for the related process $B \to K \ell^+ \ell^-$, where $\ell^+ \ell^-$ is either an $e^+ e^-$ or a $\mu^+ \mu^-$ pair. These decays are highly suppressed in the standard model, and they are sensitive to contributions from new particles in the intermediate state. The data sample comprises $123 \times 10^6$ $\Upsilon(4S) \to B\bar{B}$ decays collected with the BABAR detector at the SLAC PEP-II $e^+ e^-$ storage ring. Averaging over $K^{(*)}$ isospin and lepton flavor, we obtain the branching fractions $\mathcal{B}(B \to K^+ \ell^+ \ell^-) = (0.65^{+0.11}_{-0.10} \pm 0.04) \times 10^{-6}$ and $\mathcal{B}(B \to K^0 \ell^+ \ell^-) = (0.88^{+0.33}_{-0.29} \pm 0.10) \times 10^{-6}$, where the uncertainties are statistical and systematic, respectively. The significance of the $B \to K^+ \ell^+ \ell^-$ signal is over $8\sigma$, while for $B \to K^+ \ell^+ \ell^-$ it is $3.3\sigma$. 

DOI: 10.1103/PhysRevLett.91.221802  PACS numbers: 13.20.He
Rare decays of $B$ mesons that involve loop diagrams in the standard model (SM) provide a promising means to search for effects beyond the SM [1]. The decays $B \to K^+ \ell^+ \ell^-$ and $B \to K^* \ell^+ \ell^-$, where $\ell^\pm$ are charged leptons and $K^*$ is the $K^*(892)$ meson, result from one-loop processes that transform the $b$ quark in the initial-state $B$ meson into an $s$ quark in the final-state $K^{(*)}$ meson. In the SM, three amplitudes contribute at leading order: an electromagnetic (EM) penguin, a $Z$ penguin, and a $W^+W^-$ box diagram. The penguin diagrams involve the emission and absorption of a $W$ boson. The presence of three SM electroweak amplitudes makes $B \to K^{(*)} \ell^+ \ell^-$ more complex than $B \to K^* \gamma$, which proceeds solely through an EM penguin.

Because of their loop structure, these decays are highly suppressed, with SM branching fractions expected to be roughly $0.5 \times 10^{-6}$ for $B \to K^+ \ell^+ \ell^-$ and about 3 times that for the $B \to K^* \ell^+ \ell^-$ modes [1–3]. Because of the complexity of strong interaction effects, however, theoretical uncertainties on the rates are currently at least 35% (Ali et al. [1]). Both $B \to K^+ e^+ e^-$ and $B \to K^* \mu^+ \mu^-$ receive a contribution from the pole in the EM penguin amplitude at $q^2 = m_{\gamma\gamma}^2 = 0$, but the enhancement in the electron mode is significantly larger due to a lower $q^2$ threshold. This pole does not contribute to $B \to K e^+ e^-$ or to $B \to K \mu^+ \mu^-$; thus, the difference in those partial widths is expected to be negligible. An important consequence of the loop structure of these decays is that their branching fractions and kinematic distributions can be significantly affected by the presence of new particles or couplings, such as those predicted in models based on supersymmetry [1].

Recently, substantial progress has been made in experimental studies of these decays. The Belle Collaboration has observed $B \to K^+ \ell^+ \ell^-$, as well as the inclusive $B \to X_s \ell^+ \ell^-$ decay [4]. Limits on these and similar modes have been set by BABAR [5], Belle [4], CLEO [6], and CDF [7]. Our new measurements are based on a data sample 6 times larger than that used for our previously published results. We study eight final states: $B^+ \to K^+ \ell^+ \ell^-$, $B^0 \to K_S^0 \ell^+ \ell^-$, $B^{+} \to K^{*+} \ell^+ \ell^-$, and $B^0 \to K^{0*} \ell^+ \ell^-$, where $K^{*0} \to K^+ \pi^-$, $K^{*+} \to K_S^0 \pi^+$, $K_S^0 \to \pi^+ \pi^-$, and $\ell$ is either an $e$ or a $\mu$. Throughout this Letter, charge-conjugate modes are implied.

We analyze data collected with the BABAR detector [8] at the PEP-II storage ring at the Stanford Linear Accelerator Center. The data sample comprises 113.1 fb$^{-1}$ recorded on the $Y(4S)$ resonance, yielding $(122.9 \pm 1.4) \times 10^6 B\bar{B}$ decays, and an off-resonance sample of 12.0 fb$^{-1}$ used to study continuum background.

We select events that include two oppositely charged leptons ($e^+e^-$, $\mu^+\mu^-$), a kaon (either $K^\pm$ or $K_S^0$), and, for the $B \to K^+ \ell^+ \ell^-$ modes, a $\pi^\pm$ that combines with a kaon to form a $K^*$ candidate. Electrons are identified primarily in the CsI(Tl) electromagnetic calorimeter, while muons are identified by their penetration through iron plates of the magnet flux return. Electron (muon) candidates are required to satisfy $p > 0.5$ (1.0) GeV/c. Bremsstrahlung photons from electrons are recovered by combining an electron candidate with up to one photon with $E > 30$ MeV in a small angular region around the initial electron direction. Photon conversions and $\pi^0$ Dalitz decays are removed by vetoing all low-mass $e^+e^-$ pairs, except in $B \to K^{*0} e^+ e^-$ modes, where we preserve acceptance at low mass by retaining pairs that intersect inside the beam pipe.

$K^\pm$ candidates are tracks with $dE/dx$ and Cherenkov angle consistent with a kaon. $\pi^\pm$ candidates are tracks that do not satisfy the $K^\pm$ selection. $K_{S}^{0}$ candidates are reconstructed from two oppositely charged tracks with an invariant mass consistent with the $K_{S}^{0}$ mass and a common vertex displaced from the primary vertex by at least 1 mm.

True $B$ signal decays produce narrow peaks in the distributions of two kinematic variables, which can be fitted to extract the signal and background yields. For a candidate system of $B$ daughter particles with masses $m_i$ and three-momenta $\mathbf{p}_i$ in the $Y(4S)$ center-of-mass (CM) frame, we define $m_{ES} = \sqrt{E_b^2 - \sum |\mathbf{p}_i|^2}$ and $\Delta E = \sum \sqrt{m_i^2 + |\mathbf{p}_i|^2} - E_b$, where $E_b$ is the beam energy in the CM frame. For signal events, the $m_{ES}$ distribution peaks at the $B$ meson mass with resolution $\sigma = 2.5$ MeV/$c^2$, and the $\Delta E$ distribution peaks near zero, with a typical width $\sigma = 20$ MeV. In $B \to K^+ \ell^+ \ell^-$ channels, we perform a two-dimensional unbinned maximum-likelihood fit to the distribution of $m_{ES}$ and $\Delta E$ in the region $m_{ES} > 5.2$ GeV/$c^2$ and $|\Delta E| < 0.25$ GeV. In $B \to K^{*0} \ell^+ \ell^-$ decays, we perform a three-dimensional fit to $m_{ES}$, $\Delta E$, and the kaon-pion invariant mass in the region $0.7 < m_{K\pi} < 1.1$ GeV/$c^2$.

Backgrounds arise from three main sources: random combinations of particles from $q\bar{q}$ events produced in the continuum, random combinations of particles from $Y(4S) \to B\bar{B}$ decays, and $B$ decays to topologies similar to the signal modes. The first two ("combinatorial") backgrounds typically arise from pairs of semileptonic decays and produce broad distributions in $m_{ES}$ and $\Delta E$ compared to the signal. The third source arises from modes such as $B \to J/\psi K^{(*)}$ (with $J/\psi \to \ell^+ \ell^-$) or $B \to K^{(*)} \pi \pi$ (with pions misidentified as muons), which have shapes similar to the signal. All selection criteria are optimized with GEANTS [9] simulated data or with data samples outside the full fit region.

We suppress combinatorial background from continuum processes using a Fisher discriminant [10], which is a linear combination of variables with coefficients optimized to distinguish between signal and background. The variables (defined in the CM frame) are (1) the ratio of second- to zeroth-order Fox-Wolfram moments [11] for the event, computed using all charged tracks and neutral energy clusters; (2) the angle between the thrust axis of the $B$ candidate and that of the remaining particles in the event; (3) the production angle $\theta_B$ of the $B$ candidate with...
respect to the beam axis; and (4) the masses of $K\ell$ pairs with charge correlation consistent with $D$ decay.

We exclude dilepton pairs consistent with the $K^0/\bar{K}^0$ from the charmonium mass, while $3_{221802-5}^{221802-5}$ computed without bremsstrahlung photon recovery. When a $D^0$ is vetoed events where the $D^0$ by vetoing events where the $D$ or $K^0$ by simulation to be $3_{221802-5}^{221802-5}$ $m_{ES}$ and $3_{221802-5}^{221802-5}$ $K^0$ events in $75\text{ GeV}/c^2$. Missing energy provides the strongest suppression of combinatorial backgrounds from charmonium data control samples.

The background is modeled as the sum of three terms: (1) a combinatorial background shape with floating normalization, written as the product of an ARGUS function [12] in $m_{ES}$, an exponential in $\Delta E$, and the product of $\Delta m_{K^*}-m_K-m_\pi$ and a quadratic function of $m_{ES}$ for the $K^*$ modes; (2) a peaking background contribution, with the same shape as the signal, but with normalization fixed to measured peaking backgrounds; and (3) terms with floating normalization to describe (a) background in $B\to K^+\ell^+\ell^-$ ($B\to K^*\ell^+\ell^-$), (b) background in $B\to K^+\ell^+\ell^-$ ($B\to K^*\ell^+\ell^-$) events with a lost pion, and (b) background in $B\to K^+\ell^+\ell^-$ ($B\to K^+\ell^+\ell^-$) events with a randomly added pion. In the $K^*$ modes, we allow an additional background (4) that uses our combinatorial shape in $m_{ES}$ and $\Delta E$, but peaks in $m_{K^*}$ at the $K^*$ mass. Because the normalizations for terms (1), (3), and (4) are floating, as are the combinatorial background shape parameters, much of the uncertainty in the background is propagated into the statistical uncertainty on the signal yield obtained from the fit.

Table I lists signal yields and branching fractions for each mode. The relative systematic uncertainties on the efficiency, $\Delta B_{B^0}/B$, arise from charged-particle tracking (1.0% per lepton, 1.7% per charged hadron), particle identification (1.1% per electron, 1.6% per muon, 0.9% per pion, 0.9% per kaon), the continuum suppression cut [(0.8–2.8)%], the $B\bar{B}$ suppression cut [(1.4–5.0)%], $K^0_S$ selection (3.8%), signal simulation statistics [(0.7–1.4)%], theoretical model dependence of the efficiency [(4–7)%], depending on the mode), and the number of $B\bar{B}$ events (1.1%). Uncertainties on efficiencies due to model dependence of form factors are taken to be the full range of variation from a set of models [2].

The systematic uncertainties on the fit yields, $\Delta B_{fit}$, arise from three sources: uncertainties in the parameters describing the signal shapes, possible correlation between $m_{ES}$ and $\Delta E$ 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 simulation for the charmonium control samples. For modes with electrons, we also vary the fraction of signal events in the tail of the $\Delta E$ distribution. To evaluate sensitivity to the background parametrization, we allow additional parameters and a correlation between $m_{ES}$ and $\Delta E$.

Table I also lists results from simultaneous fits to combinations of $B\to K^+\ell^+\ell^-$ modes and combinations of $B\to K^-\ell^+\ell^-$ modes, where the relative branching fractions for the contributing modes are constrained. $B^0$ and $B^+$ production rates are constrained to be equal, and the ratio of their total widths is constrained to be...
TABLE I. Results from the fits to $B \to K^{(*)} \ell^+ \ell^-$ modes. The columns are, from left to right, fitted signal yield; the signal efficiency, $\epsilon$ (not including the branching fractions for $K'$ and $K^0$ decays); the fractional systematic error on the selection efficiency, $\Delta B_\epsilon / B$; the systematic error from the fit, $\Delta B_{\text{fit}}$; and the branching fraction central value ($B$) with its statistical and total systematic uncertainties. For the branching fractions averaged over different channels (lower part of table), simultaneous, constrained fits are performed to extract an efficiency-corrected signal yield that averages over the included channels. The modes with significance $>3\sigma$ are $K^+ e^+ e^-$ ($8.4\sigma$), $K^0 \mu^+ \mu^-$ ($4.1\sigma$), $K^- e^+ e^-$ ($7.8\sigma$), $K^{*+} \ell^- (8.4\sigma)$, and $K^{*+} \ell^- (3.3\sigma)$.

<table>
<thead>
<tr>
<th>Mode</th>
<th>Signal yield (%)</th>
<th>$\epsilon$</th>
<th>$\Delta B_\epsilon / B$ (%)</th>
<th>$\Delta B_{\text{fit}}$ (%)</th>
<th>$B$ ($10^{-6}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K^+ e^+ e^-$</td>
<td>24.7 $^{+5.0}_{-4.8}$</td>
<td>19.2</td>
<td>$\pm 6.3$ $0.02$</td>
<td>1.05 $^{+0.25}_{-0.22}$ $0.07$</td>
<td></td>
</tr>
<tr>
<td>$K^+ \mu^+ \mu^-$</td>
<td>0.7 $^{+2.0}_{-1.2}$</td>
<td>8.5</td>
<td>$\pm 7.6$ $0.02$</td>
<td>0.07 $^{+0.11}_{-0.12}$ $0.02$</td>
<td></td>
</tr>
<tr>
<td>$K^0 e^+ e^-$</td>
<td>$-1.8^{+2.2}_{-1.5}$</td>
<td>20.1</td>
<td>$\pm 8.4$ $0.08$</td>
<td>$-0.21^{+0.15}_{-0.16}$ $0.08$</td>
<td></td>
</tr>
<tr>
<td>$K^0 \mu^+ \mu^-$</td>
<td>5.9 $^{+2.0}_{-1.3}$</td>
<td>8.6</td>
<td>$\pm 8.8$ $0.02$</td>
<td>1.63 $^{+0.35}_{-0.33}$ $0.14$</td>
<td></td>
</tr>
<tr>
<td>$K^{*0} e^+ e^-$</td>
<td>12.4 $^{+6.3}_{-5.2}$</td>
<td>13.6</td>
<td>$\pm 7.6$ $0.08$</td>
<td>1.11 $^{+0.47}_{-0.46}$ $0.11$</td>
<td></td>
</tr>
<tr>
<td>$K^{*0} \mu^+ \mu^-$</td>
<td>4.5 $^{+1.5}_{-1.0}$</td>
<td>6.4</td>
<td>$\pm 10.1$ $0.07$</td>
<td>0.86 $^{+0.38}_{-0.38}$ $0.11$</td>
<td></td>
</tr>
<tr>
<td>$K^{*+} e^+ e^-$</td>
<td>$0.6^{+1.3}_{-1.0}$</td>
<td>10.2</td>
<td>$\pm 10.7$ $0.28$</td>
<td>$20.1^{+0.34}_{-0.35}$ $0.28$</td>
<td></td>
</tr>
<tr>
<td>$K^{*+} \mu^+ \mu^-$</td>
<td>4.2 $^{+1.3}_{-0.8}$</td>
<td>4.8</td>
<td>$\pm 12.7$ $0.15$</td>
<td>3.07 $^{+0.28}_{-0.27}$ $0.42$</td>
<td></td>
</tr>
<tr>
<td>$K e^+ e^-$</td>
<td>91 $^{+20}_{-19}$</td>
<td>$\pm 6.5$</td>
<td>$0.02$</td>
<td>0.74 $^{+0.18}_{-0.16}$ $0.05$</td>
<td></td>
</tr>
<tr>
<td>$K \mu^+ \mu^-$</td>
<td>55 $^{+20}_{-15}$</td>
<td>$\pm 7.4$</td>
<td>$0.01$</td>
<td>0.45 $^{+0.14}_{-0.18}$ $0.03$</td>
<td></td>
</tr>
<tr>
<td>$K^{*0} e^+ e^-$</td>
<td>80 $^{+15}_{-14}$</td>
<td>$\pm 6.4$</td>
<td>$0.01$</td>
<td>0.65 $^{+0.13}_{-0.14}$ $0.04$</td>
<td></td>
</tr>
<tr>
<td>$K^{*0} \mu^+ \mu^-$</td>
<td>121 $^{+34}_{-33}$</td>
<td>$\pm 7.8$</td>
<td>$0.08$</td>
<td>0.98 $^{+0.20}_{-0.18}$ $0.11$</td>
<td></td>
</tr>
<tr>
<td>$K^{*+} e^+ e^-$</td>
<td>156 $^{+41}_{-35}$</td>
<td>$\pm 10.1$</td>
<td>$0.09$</td>
<td>1.27 $^{+0.61}_{-0.56}$ $0.16$</td>
<td></td>
</tr>
<tr>
<td>$K^{*+} \mu^+ \mu^-$</td>
<td>108 $^{+36}_{-35}$</td>
<td>$\pm 8.1$</td>
<td>$0.07$</td>
<td>0.88 $^{+0.33}_{-0.36}$ $0.10$</td>
<td></td>
</tr>
</tbody>
</table>

1.085 $\pm 0.017$ [13]. All branching fractions from simultaneous fits are expressed relative to the $B^0$ total width. The projections of the fit on $m_{\text{ES}}$ and $\Delta E$ are shown in Fig. 1 for the simultaneous fit to the $B \to K^{*+} \ell^- \ell^-$ channels. We assume that all four $B \to K^{*0} \ell^- \ell^-$ modes have equal partial widths. A signal is evident at the $B$ mass in $m_{\text{ES}}$ and at $\Delta E = 0$. Figure 2 shows projections of the simultaneous fit to all $B \to K^{*0} \ell^- \ell^-$ modes. Here, the partial width ratio of electron and muon modes is constrained to be $\Gamma(B \to K^{*0} e^+ e^-) / \Gamma(B \to K^{*0} \mu^+ \mu^-) = 1.33$ from the model of Ali et al. [1]. Our simultaneous fit result is expressed as a $B^0 \to K^{*0} \mu^+ \mu^-$ branching fraction.

The significance of the $B \to K^{*+} \ell^- \ell^-$ signal from the simultaneous fit is $\sim 8\sigma$, computed as $\sqrt{\Delta \log L}$, where $\Delta \log L$ is the likelihood difference between the best fit and the null-signal hypothesis. We account for systematic uncertainties in the significance by simultaneously including all effects that individually lower the fit yields prior to computing the change in likelihood. The significance of the $B \to K^{*+} \ell^- \ell^-$ signal, including all systematic uncertainties, is $3.3\sigma$ ($3.8\sigma$ not including them).

In summary, we have observed signals for $B \to K^{*0} \ell^- \ell^-$, averaged over lepton type ($e^+ e^-$ and $\mu^+ \mu^-$) and $B$ charge, and we have obtained the first evidence for $B \to K^{*+} \ell^- \ell^-$, similarly averaged. We obtain

$$B(B \to K^{*+} \ell^- \ell^-) = (0.65^{+0.14}_{-0.13} \pm 0.04) \times 10^{-6},$$

$$B(B \to K^{*0} \ell^- \ell^-) = (0.88^{+0.33}_{-0.29} \pm 0.10) \times 10^{-6},$$

where the first error is statistical and the second is systematic. Our branching fraction for $B \to K^{*+} \ell^- \ell^-$ is slightly higher than our previous limit $0.51 \times 10^{-6}$ (90% confidence level) [5] and is in agreement with the Belle result $(0.75^{+0.25}_{-0.21} \pm 0.09) \times 10^{-6}$ [4]. Our $B \to K^{*0} \ell^- \ell^-$ branching fraction is consistent with previous 90% confidence level limits from BABAR ($<3.1 \times 10^{-6}$ for $K^{*0} \ell^- \ell^-$) [5] and Belle ($<3.1 \times 10^{-6}$ for $K^{*+} \ell^- \ell^-$) [4]. These results are consistent with the range of predictions based on the standard model [1–3].

FIG. 1 (color online). Distributions of the fit variables in $K^{*+} \ell^+ \ell^-$ data (points), compared with projections of the simultaneous fit (curves): (a) $m_{\text{ES}}$ after requiring $-0.11 < \Delta E < 0.05$ GeV and (b) $\Delta E$ distribution after requiring $m_{\text{ES}} - m_{\text{BL}} < 6.6$ MeV/$c^2$ (2.6σ). The solid curve is the sum of all fit components, including signal; the dashed curve is the sum of all background components.

FIG. 2 (color online). Distributions of the fit variables in $K^{*0} \ell^+ \ell^-$ data (points), compared with projections of the simultaneous fit (curves): (a) $m_{\text{ES}}$ after requiring $-0.11 < \Delta E < 0.05$ GeV and $0.817 < m_{\text{ES}} < 0.967$ GeV/$c^2$, (b) $\Delta E$ after requiring $m_{\text{ES}} - m_{\text{BL}} < 6.6$ MeV/$c^2$ (2.6σ), $0.817 < m_{\text{ES}} < 0.967$ GeV/$c^2$, and (c) $m_{\text{ES}}$ after requiring $m_{\text{ES}} - m_{\text{BL}} < 6.6$ 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.
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 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 the A. P. Sloan Foundation, Research Corporation, and Alexander von Humboldt Foundation.

*Also with Università di Perugia, Perugia, Italy.
†Also with Università della Basilicata, Potenza, Italy.
‡Also with IFIC, Instituto de Fisica Corpuscular, CSIC-Universidad de Valencia, Valencia, Spain.
§Deceased.

[12] The function is \( f(x) \approx \frac{x}{\sqrt{1-x^2}} \exp[-\zeta(1-x^2)]\), where \( \zeta \) is a fit parameter and \( x = \frac{m_0}{E}\), ARGUS Collaboration, H. Albrecht et al., Z. Phys. C 48, 543 (1990).