SLAC-R-840

Measurements of Rates, Asymmetries, and Angular Distributions in B -> K l+ l- and B -> K* l+ l- Decays

Jonathan Hollar

Stanford Linear Accelerator Center
Stanford University
Stanford, CA 94309

SLAC-Report-840

Measurements of Rates, Asymmetries, and Angular Distributions in $B_d \to K\ell^+\ell^-$ and $B_d \to K^*\ell^+\ell^-$ Decays

by

Jonathan Hollar

A dissertation submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy
(Physics)

at the

University of Wisconsin – Madison

2006

MEASUREMENTS OF RATES, ASYMMETRIES, AND ANGULAR DISTRIBUTIONS IN $B_d \to K\ell^+\ell^-$ AND $B_d \to K^*\ell^+\ell^-$ DECAYS

Jonathan Hollar

Under the supervision of Professor Richard Prepost

At the University of Wisconsin — Madison

This dissertation describes studies of the rare decays $B_d \to K\ell^+\ell^-$ and $B_d \to K^*\ell^+\ell^-$, where $\ell^+\ell^-$ is either an e^+e^- or a $\mu^+\mu^-$ pair. These decays are highly suppressed in the Standard Model, and could be strongly affected by physics beyond the Standard Model. We measure the total branching fractions

$$\mathcal{B}(B_d \to K\ell^+\ell^-) = (0.34 \pm 0.07 \pm 0.03) \times 10^{-6},$$

$$\mathcal{B}(B_d \to K^*\ell^+\ell^-) = (0.78^{+0.19}_{-0.17} \pm 0.12) \times 10^{-6}.$$

In addition, we measure the partial branching fractions, relative abundance of muons to electrons, direct CP asymmetry, dilepton forward-backward asymmetry, and longitudinal polarization of the K^* in these modes. We also search for the lepton flavor-violating decays $B_d \to Ke^{\pm}\mu^{\mp}$ and $B_d \to K^*e^{\pm}\mu^{\mp}$. The measurements were performed at the SLAC PEP II storage ring running at the $\Upsilon(4S)$ resonance.

Contents

List of Figures

List of Tables

1 Introduction

The goal of particle physics is to understand the fundamental constituents of matter and their interactions. Our current best attempt at such a description is known as the Standard Model (SM) of particle physics. The SM has been highly successful in incorporating the known particles and forces, excepting gravity, into a common framework that can accurately predict particle interaction phenomena. The SM has survived over three decades of experimental tests of these predictions. Even though the SM has been enormously successful, it is widely believed that it can only be an approximation of a more fundamental theory. This section gives a brief overview of the SM, and of some of the reasons for believing that there exists physics beyond the SM.

1.1 The Standard Model

The Standard Model is a gauge theory based on the group $SU(3)_C \times SU(2)_L \times U(1)_Y$ of the strong and unified electromagnetic and weak ("electroweak") interactions (Table 1). The interactions of the SM have characteristic scale dependent coupling constants g_i, and are mediated by spin-1 gauge bosons. The electromagnetic (EM) force is mediated by massless photons, the weak force by massive W and Z bosons, and the strong force by massless gluons. The quarks and leptons that make up matter (Table 2) can be organized into three

Gauge boson	m (GeV/c^2)	Electromagnetic charge	Mediates
γ	$< 6 \times 10^{-17}$	0	electromagnetism
g	0	0	strong force
Z^0	91.1876 ± 0.0021	0	weak force
W^+	80.403 ± 0.029	$+1$	weak force

Table 1: Properties of the Standard Model gauge bosons. The columns from left are: particle identity, experimentally measured mass, electric charge, and the force mediated by the particle.

Leptons	Generation	m(GeV/c^2)	Electromagnetic charge	Interactions
e	1	0.000511	-1	weak, EM
ν_e	1	$< 3 \times 10^{-9}$	0	weak
μ	2	0.106	-1	weak, EM
ν_μ	2	$< 1.9 \times 10^{-4}$	0	weak
τ	3	$1.7770^{+0.00029}_{-0.00026}$	-1	weak, EM
ν_τ	3	< 0.018	0	weak
Quarks				
u	1	0.0015 to 0.003	$+2/3$	strong, weak, EM
d	1	0.003 to 0.007	$-1/3$	strong, weak, EM
c	2	1.25 ± 0.09	$+2/3$	strong, weak, EM
s	2	0.095 ± 0.025	$-1/3$	strong, weak, EM
t	3	174.2 ± 3.3	$+2/3$	strong, weak, EM
b	3	4.70 ± 0.07 (1S mass)	$-1/3$	strong, weak, EM

Table 2: Properties of the Standard Model quarks and leptons. The columns from left are: particle identity, generation, experimentally measured mass [1], electric charge, and the list of interactions the particle participates in.

generations, each containing a left-handed weak doublet and one or two right-handed weak singlets under the $SU(2)_L \times U(1)_Y$ symmetry group. The purely right-handed neutrinos interact only through the weak force, while the quarks and charged leptons also interact through the electromagnetic force. The quarks also carry one of three color charges of the strong interaction, and are triplets under the $SU(3)_C$ group. For each elementary particle of the SM there is an associated anti-particle carrying conjugated internal quantum numbers, such as electromagnetic charge.

Given the structure of the SM left-handed doublets of quarks (Q_i) and leptons (L_i), and the right-handed singlets of leptons (e_{Ri}) and up and down-type quarks (u_{Ri} and d_{Ri}):

$$L_1 = \begin{pmatrix} e^- \\ \nu_e \end{pmatrix}_L \qquad e_{R1} = e_R^- \qquad Q_1 = \begin{pmatrix} u \\ d \end{pmatrix}_L \qquad u_{R1} = u, d_{R1} = d \tag{1}$$

$$L_2 = \begin{pmatrix} \mu^- \\ \nu_\mu \end{pmatrix}_L \qquad e_{R2} = \mu_R^- \qquad Q_2 = \begin{pmatrix} c \\ s \end{pmatrix}_L \qquad u_{R2} = c, d_{R2} = s$$

$$L_3 = \begin{pmatrix} \tau^- \\ \nu_\tau \end{pmatrix}_L \qquad e_{R3} = \tau_R^- \qquad Q_3 = \begin{pmatrix} t \\ b \end{pmatrix}_L \qquad u_{R3} = t, d_{R3} = b,$$

the Lagrangian of the Standard Model interactions can be expressed as:

$$\mathcal{L}_I = -\frac{1}{4}G_{\mu\nu}^a G_a^{\mu\nu} - \frac{1}{4}W_{\mu\nu}^a W_a^{\mu\nu} - \frac{1}{4}B_{\mu\nu}^a B_a^{\mu\nu} + \tag{2}$$
$$\overline{L}_i i D_\mu \gamma^\mu L_i + \overline{e}_{Ri} i D_\mu \gamma^\mu e_{Ri} + \overline{Q}_i i D_\mu \gamma^\mu Q_i +$$
$$\overline{u}_{Ri} i D_\mu \gamma^\mu u_{Ri} + \overline{d}_{Ri} i D_\mu \gamma^\mu d_{Ri}.$$

Here D_μ indicates a covariant derivative that can be expressed for quarks in terms of the couplings g_i and the hypercharge Y as:

$$D_\mu \equiv \partial_\mu - ig_s - ig_2 \frac{1}{2}\sigma_a W_\mu^a - ig_1 \frac{Y_q}{2} B_\mu. \tag{3}$$

The covariant derivative acts on the fields G, W, and B, associated with the strong, weak, and electromagnetic interactions, respectively.

1.1.1   Electroweak symmetry breaking

Within the basic Standard Model framework, the fields of the Lagrangian $\mathcal{L}_I$ are massless. Attempting to introduce the experimentally observed masses from Tables 1 and 2 by the direct addition of mass-generating terms of the form $M_W W_\mu W^\mu$ will violate the local gauge invariance of the $SU(2)_L \times U(1)_Y$ group. The lack of mass-generating terms can be remedied by the introduction of electroweak symmetry breaking, in which the electroweak $SU(2)_L \times U(1)_Y$ symmetry is spontaneously broken to the electromagnetic $U(1)$ symmetry. Such spontaneous symmetry breaking is accomplished by adding to the Lagrangian an additional term of the form:

$$\mathcal{L}_S = (D^\mu \Phi)^\dagger (D_\mu \Phi) - \mu^2 \Phi^\dagger \Phi - \lambda (\Phi^\dagger \Phi)^2, \tag{4}$$

where Φ is a doublet of scalar fields:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}. \tag{5}$$

With $\mu^2 > 0$, the potential term $\mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2$ has a minimum at 0, as expected for a massless gauge boson. With $\mu^2 < 0$, the potential has minima at non-zero values of the vacuum expectation value v (Figure 1), where $v^2 = -\frac{\mu^2}{\lambda}$. After an appropriate gauge transformation, this spontaneous symmetry breaking allows the physical W and Z bosons to acquire masses of $M_W = \frac{1}{2} v g_2$ and $M_Z = \frac{1}{2} v \sqrt{g_2^2 + g_1^2}$, while the photon remains massless. Using precision data from μ decays, the value of v can be determined to be about 247 GeV [1]. The gauge boson masses are then be predicted to be [1] $M_W = (80.376 \pm 0.017)$ GeV/c^2 and $M_Z = (91.1874 \pm 0.0021)$ GeV/c^2 in the SM, in excellent agreement with the direct

5
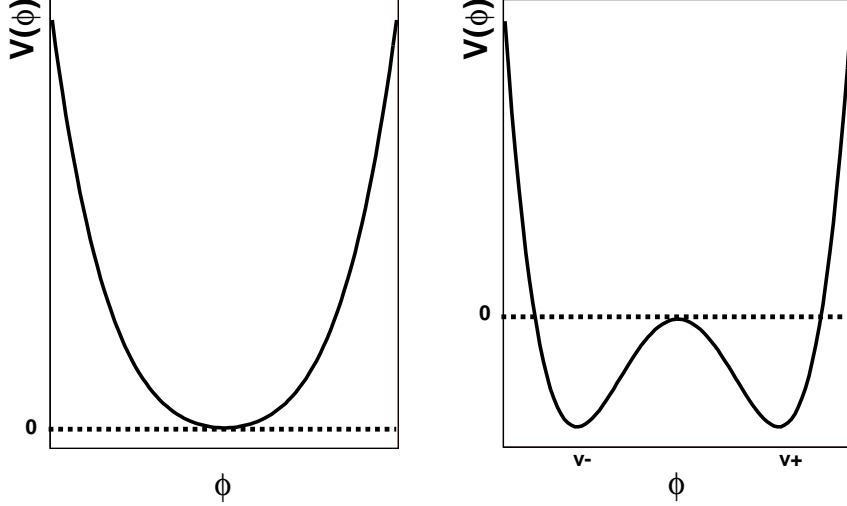
Figure 1: Higgs potential for $\mu^2 > 0$ (left) and $\mu^2 < 0$ (right).

measurements in Table 1. The masses of the SM quarks and leptons are similarly generated through coupling to this scalar Higgs field. Their masses, however, are proportional to unknown Yukawa couplings λ, and are thus not predicted in the SM.

After the symmetry is broken, one physical degree of freedom remains, corresponding to a neutral scalar Higgs boson of mass $M_H = \sqrt{-2\mu^2}$. Since the SM does not explicitly predict the mass of the Higgs boson it must be determined experimentally. The existing 95% CL lower limit on the Standard Model Higgs mass from direct searches is $M_H > 114.4$ GeV/c^2 [2]. Based on precision measurements of the W-boson and t quark masses as of 2005, the Higgs mass can be indirectly bounded from above at 95% CL to be $M_H < 186$ GeV/c^2 [1].

1.1.2 Neutral currents and the GIM mechanism

In weak neutral current interactions that occur through the exchange of a Z^0 boson, the quark flavor is conserved at the Z^0 vertex. However, effective flavor changing neutral currents
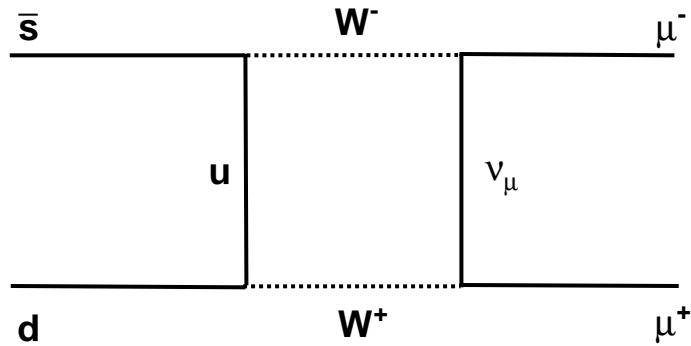
Figure 2: Example of an effective flavor changing neutral current kaon decay.

(FCNC's) are also possible through box diagrams. An example is the decay of a K^0 ($\bar{s}d$) into a $\mu^+\mu^-$ pair, as shown in Figure 2. In a model with only the three lightest quarks (u, d, s), a substantial rate would be predicted for this decay mode. The contribution from the same diagram containing a c quark rather than a u quark, however, nearly cancels this[1], leading to the suppression of effective FCNC's. The cancellation of the u and c quark contributions is the basis of the Glashow-Iliopoulos-Maiani (GIM) mechanism, which successfully explains the observed suppression of decays such as $K^0 \rightarrow \mu^+\mu^-$. The same argument can be extended to three complete generations, with the addition of the two heavy (b and t) quarks.

1.1.3 The CKM matrix and CP violation

In the Standard Model, mixing between quarks of different flavors in charged current interactions is described by a unitary 3×3 matrix, known as the Cabibbo-Kobayashi-Maskawa (CKM) matrix:

---

[1]The cancellation is not exact due to the mass difference between the u and c quarks

$$
\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}.
$$

The CKM matrix effectively transforms the mass eigenstates of the down-type quarks (d, s, b) to the weak eigenstates (d', s', b'). A 3×3 unitary matrix can be parameterized in terms of three angles and six phases, five of which can be removed by appropriate transformations. A single irreducible phase is then left, which allows for an asymmetry under the combined discrete operations of charge conjugation (C) and parity reversal (P). Within the Standard Model, this weak phase is the only source of such CP violating asymmetries in the quark sector. The hierarchy between elements of the CKM matrix is more apparent in the Wolfenstein parameterization:

$$
V_{CKM} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4),
$$

in which ρ, η is a point in the complex plane, and $\lambda \equiv |V_{us}|$ is used as an expansion parameter.

One of the unitarity conditions of the CKM matrix can be expressed in terms of the matrix elements as:

$$
V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0,
$$

which after rescaling describes a triangle in the complex plane with an apex at the point (ρ, η) (Figure 3). The experimental 90% CL bounds on the values of the CKM matrix elements as of 2004 are [1]:
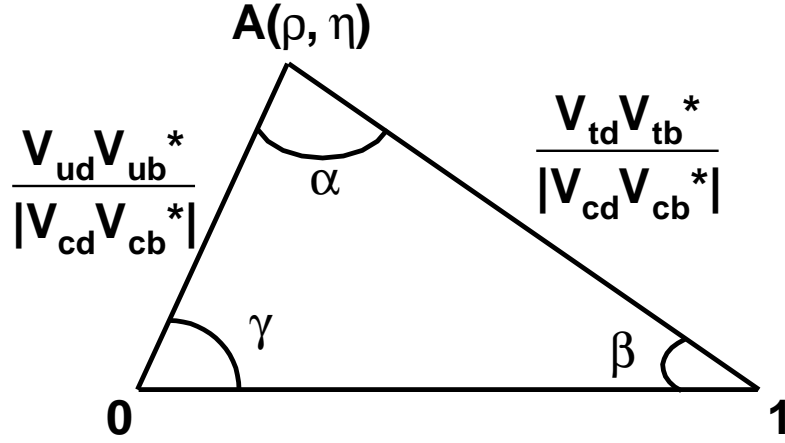
Figure 3: The unitarity triangle.

$$V_{CKM} = \begin{pmatrix} 0.97383^{+0.00024}_{-0.00023} & 0.2272^{+0.0010}_{-0.0010} & (3.96^{+0.09}_{-0.09}) \times 10^{-3} \\ 0.2271^{+0.0010}_{-0.0010} & 0.97296^{+0.00024}_{-0.00024} & (42.21^{+0.10}_{-0.80}) \times 10^{-3} \\ (8.14^{+0.32}_{-0.64}) \times 10^{-3} & (41.61^{+0.12}_{-0.78}) \times 10^{-3} & 0.999100^{+0.000034}_{-0.000004} \end{pmatrix}.$$

The experimental determination of all sides and angles of the Unitarity Triangle is a major ongoing activity of particle physics. To date all measurements are consistent with the hypothesis that the CKM mechanism is the only source of CP violation in the quark sector [3, 4].

1.2   New Physics

Despite the impressive agreement between experimental data and the Standard Model, there are compelling reasons to believe that there is new physics beyond the SM. Among these are:

- Higgs divergence. As discussed above, indirect constraints indicate the SM Higgs boson should have a mass of ~ 250 GeV/c^2 or less. The self-energy corrections (Figure 4)

to the Higgs mass, however, are quadratically divergent and can be many orders of magnitude larger than this. Therefore, a finely tuned cancellation between the loop contributions from spin-1/2 fermions and integer spin bosons is required to stabilize the Higgs mass.

- Unification of forces. The strengths of the strong, weak, and electromagnetic gauge couplings evolve as a function of the energy scale. In the Standard Model these couplings do not unify at any energy scale.

1.2.1 Supersymmetry

Among new physics theories, supersymmetric models are among the most well-motivated and thoroughly studied. Supersymmetry introduces a new set of superpartner bosons for each of the Standard Model fermions, and vice versa. The Standard Model quarks and leptons are paired with squarks ($\tilde{q}$) and sleptons ($\tilde{l}$); the gluons are paired with gluinos ($\tilde{g}$). In the minimal extension to the SM, the Higgs sector is expanded to include two Higgs doublets, whose ratio of vacuum expectation values is a free parameter $\tan(\beta)$. The extra degrees of freedom are expressed in additional physical Higgs bosons, both charged and neutral. The electroweak gauge bosons are paired with gauginos, which can mix with the Higgs sector to produce chargino ($\tilde{\chi}^+$) and neutralino ($\tilde{\chi}^0$) mass eigenstates.

Supersymmetry addresses several of the outstanding problems of the SM:
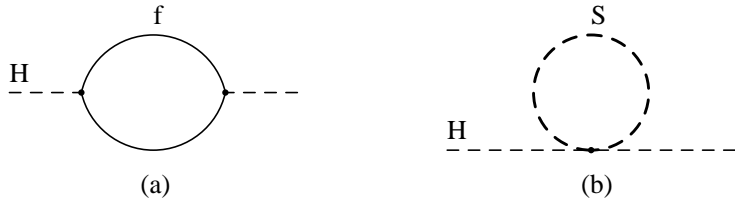
Figure 4: Examples of quadratically divergent corrections to the Higgs mass [5].
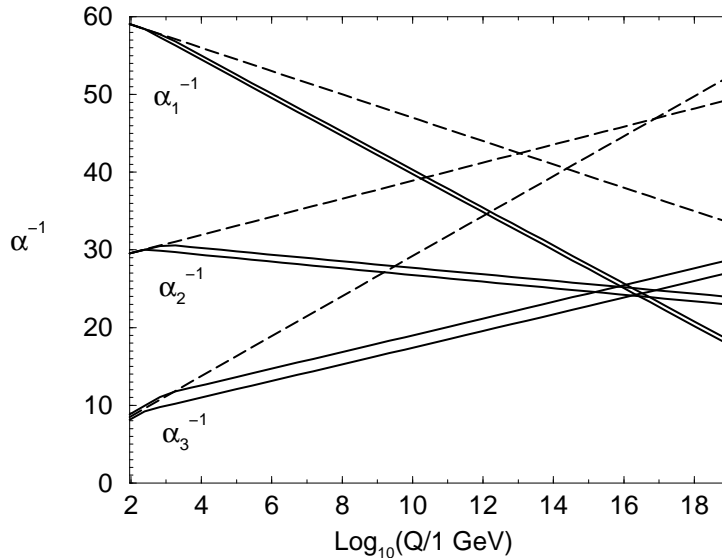
Figure 5: Gauge unification in supersymmetry. Evolution of inverse coupling constants as a function of the energy scale. The predictions of the SM (dashed lines) and a supersymmetric model (solid lines) are shown. The area between the solid lines represents the range allowed by varying α_3 evaluated at the Z mass from 0.113 to 0.123, and the superpartner mass threshold from 250 GeV to 1 TeV [5].

- The superpartners partially cancel the quadratic divergences in the Higgs self-energy, leaving much less severe logarithmic divergences.

- The inverse gauge couplings α_1, α_2, and α_3 of the electromagnetic, weak, and strong forces, respectively, can be unified at an energy scale of order $\sim 10^{16}$ GeV in some supersymmetric models (Figure 5).

To date, no superpartners have been observed experimentally. Therefore supersymmetry, if it exists, must be broken between the masses of particles and superparticles. A large number of supersymmetric models exist, with various mechanisms employed to break the symmetry. To have the desired effect of stabilizing the Higgs mass, it is generally argued that supersymmetry should be visible at energy scales of a few TeV or less. The current

limits from direct collider searches exclude superpartners up to masses of a few hundred GeV/c^2, depending on the model assumed [1].

2 The decay $B \to K^{(*)}\ell^+\ell^-$

The decays of B-mesons provide one opportunity to indirectly search for physics beyond the Standard Model. Rare decays in which the SM contribution is suppressed or forbidden are particularly promising in this regard. This section describes the physics of the rare decays $B \to K\ell^+\ell^-$ and $B \to K^*\ell^+\ell^-$, which are the topic of this thesis. Throughout the remainder of this document B is taken to refer to the B_d meson, unless otherwise indicated.

2.1 Theory

The decays $B \to K^{(*)}\ell^+\ell^-$, where $\ell^+\ell^-$ is an e^+e^- or $\mu^+\mu^-$ pair and $K^{(*)}$ is either a kaon or the $K^*(892)$ meson, are instances of the underlying quark-level process $b \to sl^+l^-$. As discussed in Section 1, flavor changing neutral currents are suppressed by the conservation of flavor at the Z vertex and by the GIM mechanism. They can, however, occur at greatly suppressed rates through loop and box diagrams. Three such SM diagrams contribute to this decay at leading order:
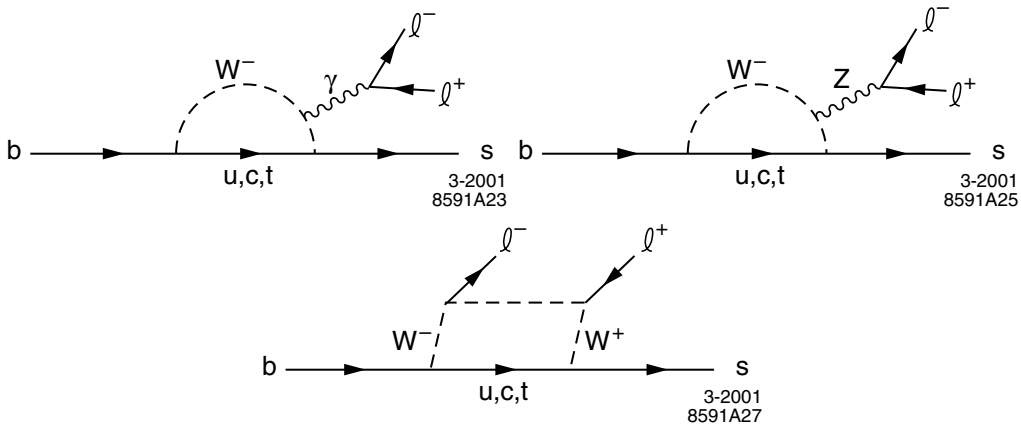
Figure 6: Standard Model Feynman diagrams contributing to $B \to K^{(*)}\ell^+\ell^-$ at leading order: photon penguin (above left), Z penguin (above right), W-box diagram (below center).
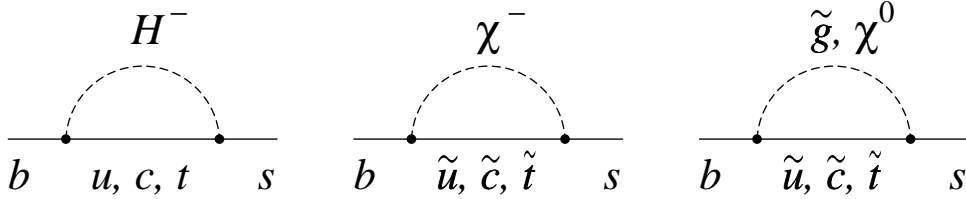
Figure 7: Examples of possible supersymmetric contributions to $B \to K^{(*)}\ell^+\ell^-$. From left: a charged Higgs diagram, a chargino/squark diagram, and a gluino(or neutralino)/squark diagram. These contributions can enter either the photon penguin or Z penguin of Figure 6.

1. A photon "penguin" diagram, proceeding via a virtual quark and a W-boson in the loop.

2. A Z penguin, also containing a virtual quark and a W-boson in the loop.

3. A W-box diagram, containing a virtual quark, two W-bosons, and a virtual neutrino.

In all three cases, the diagram with a virtual t quark dominates over those with light c and u quarks. In many theories beyond the Standard Model, new heavy particles can replace the t or W in the loop (Figure 7), leading to observable effects in the decay rates, asymmetries, and angular distributions for these modes. These effects have been studied in detail in supersymmetric theories [6, 7, 8, 9], as well as in a wide variety of alternative new physics models [10, 11, 12].

A second source of new physics effects is the potential existence of new penguin diagrams. In particular, the case in which the γ or Z boson is replaced with a neutral Higgs boson coupling to the lepton pair has been studied extensively [13, 14].

2.1.1 Wilson coefficients

The physics of heavy quark decays is more often described using the Operator Product Expansion (OPE) technique, which aims to separate the decay amplitude into a short-

Coefficient	Value	Operator structure
C_7^{eff}	-0.3094	$\frac{e}{g_s^2}m_b(\bar{s}_L\sigma^{\mu\nu}b_R)F_{\mu\nu}$
C_9^{eff}	4.2978	$\frac{e^2}{g_s^2}(\bar{s}_L\gamma_\mu b_L)\Sigma_l(\bar{l}\gamma^\mu l)$
C_{10}^{eff}	-4.4300	$\frac{e^2}{g_s^2}(\bar{s}_L\gamma_\mu b_L)\Sigma_l(\bar{l}\gamma^\mu\gamma_5 l)$

Table 3: Wilson coefficients in the SM. The first and second columns list the effective Wilson coefficient and it's numerical value evaluated at the scale $\mu = 4.6$ GeV [69], respectively. The third column shows the structure of the corresponding operator O_i, where Σ_l denotes a sum over lepton species and L and R denote the left and right-handed components of the fermion fields, respectively [7].

distance perturbative portion and a long-distance non-perturbative piece. In the case of the $b \rightarrow s\ell^+\ell^-$ transition this results in an effective Hamiltonian [15]:

$$\mathcal{H}_{eff} = -4\frac{G_F}{\sqrt{2}}V_{ts}^*V_{tb}\sum_{i=1}^{10}(|C_i(\mu)\mathcal{O}_i(\mu) + C_i'(\mu)\mathcal{O}_i'(\mu)|). \tag{6}$$

Here the terms $C_i(\mu)$ are the Wilson coefficients [16] describing the short distance physics above the energy scale μ; the terms $\mathcal{O}_i(\mu)$ are local operators describing the non-perturbative physics at scales below μ. The terms $C_i'(\mu)$ and $\mathcal{O}_i'(\mu)$ are the equivalent for right-handed currents, and are expected to be zero in the Standard Model. The G_F and $V_{ts}^*V_{tb}$ terms are the Fermi coupling constant [2] and the CKM matrix elements associated with the $b \rightarrow t \rightarrow s$ transition, respectively.

The Wilson coefficients are typically calculated at the scale of M_W, and then must be evolved down to the b mass scale. Physical observables calculated in the OPE are customarily rewritten in terms of "effective" Wilson coefficients C_i^{eff}, which are independent of the renormalization scheme. Only three of the ten Wilson coefficients are relevant to the $b \rightarrow$

---

[2]The Fermi coupling constant can be expressed in terms of the weak coupling and W mass:

$$G_F \equiv \frac{\sqrt{2}}{8}(\frac{g_2}{M_W})^2$$

$s\ell^+\ell^-$ decay: the electromagnetic operator C_7^{eff}, and the vector and axial vector terms C_9^{eff} and C_{10}^{eff}. The resulting dependence of the partial branching fraction on the Wilson coefficients is [7]:

$$\frac{d\Gamma(b \rightarrow X_s\ell^+\ell^-)}{d\hat{s}} \propto (1-\hat{s})^2((1+2\hat{s})(|C_9^{eff}|^2 + |C_{10}^{eff}|^2) + \tag{7}$$
$$4(1+\frac{2}{\hat{s}})|C_7^{eff}|^2 + 12Re(C_7^{eff}C_9^{eff*})),$$

where $\hat{s} \equiv q^2/m_b^2$ and $q^2 \equiv m_{\ell^+\ell^-}^2$. For very small values of $m_{\ell^+\ell^-}$, the rate is then dominated by the second term, proportional to the magnitude of C_7^{eff}. In the region of large $m_{\ell^+\ell^-}$, the rate is instead dominated by first term containing a C_9^{eff} and C_{10}^{eff}.

The numerical predictions for the SM Wilson coefficents, evaluated at $\mu = 4.6$ GeV [69], are given in Table 3. If new physics is present, one or more of the Wilson coefficients can be expected to deviate from these values, possibly including a change of sign.

2.1.2   Form factors

Experimentally, the exclusive channels $B \rightarrow K\ell^+\ell^-$ and $B \rightarrow K^*\ell^+\ell^-$ used in this analysis are more easily measured than the inclusive $b \rightarrow s\ell^+\ell^-$ decay. However, the use
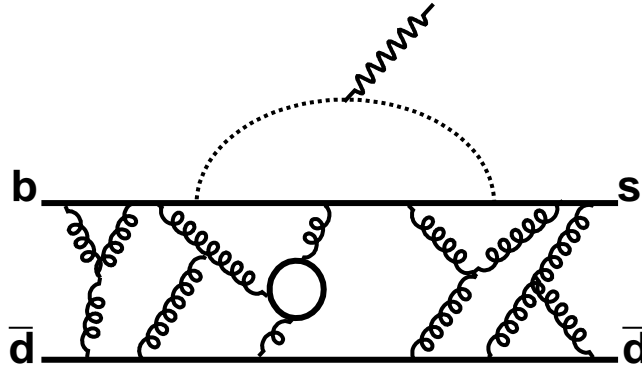
Figure 8: Example of strong interaction effects in exclusive decay modes
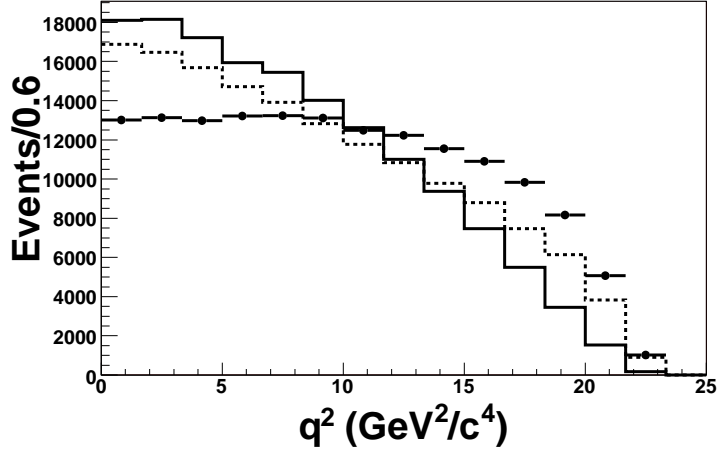
Figure 9: Standard Model q^2 spectrum in $B \to K\ell^+\ell^-$, as predicted in different form factor models. The models included are those of [19] (solid histogram), [20, 21] (dashed histogram), and [17] (points with error bars). The plots are all normalized to the same area.

of exclusive decay modes introduces complications with the theory predictions, due to the strong interaction effects involved in the $B \to K^{(*)}$ transition (Figure 8). The theoretical calculations therefore rely on form factor models to describe these hadronic effects in the $B \to K^{(*)}$ decays.

At present, there are several competing techniques used to calculate these form factors, including Light-Cone QCD Sum Rules (LCSR) [17, 18], the lattice-constrained constituent quark model [19], and three-point QCD Sum Rules [20, 21]. These techniques predict somewhat different results for the total branching fraction, and for the partial branching fraction as a function of q^2. The latter effect is illustrated in Figure 9, in which the Standard Model prediction for the q^2 spectrum in $B \to K\ell^+\ell^-$ is shown for each of these models. Branching fraction predictions based on a given form factor model also have associated theoretical uncertainties; typically these relative uncertainties are around $20 - 30\%$.
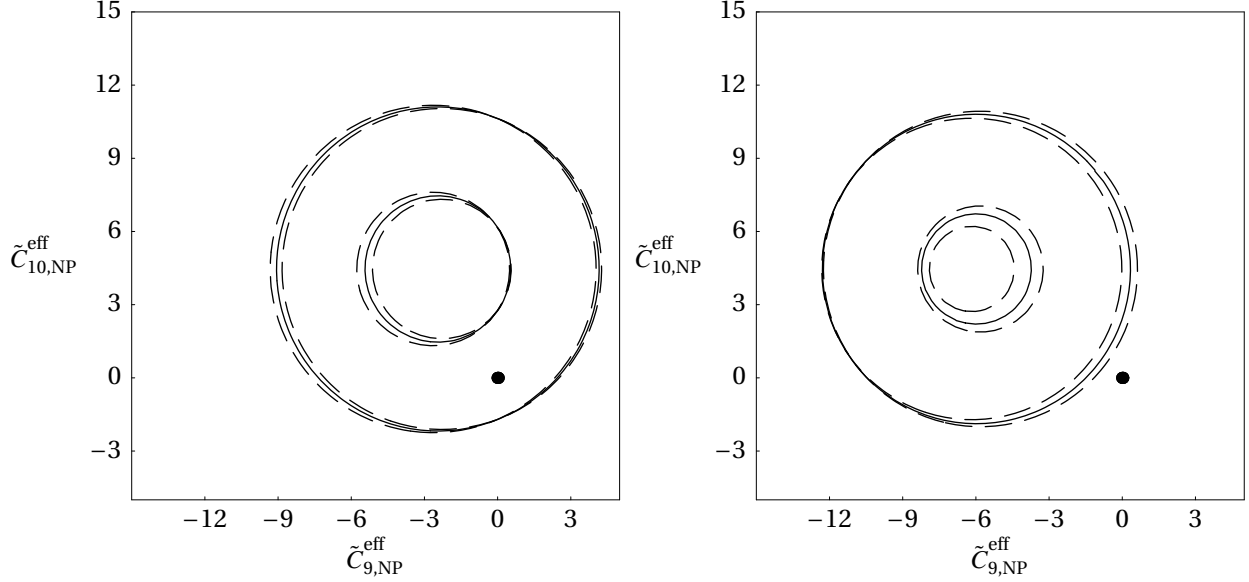
Figure 10: Existing constraints on new physics contributions to the Wilson coefficients. The 90% CL allowed region (annulus) in the C_9/C_{10} plane is shown for $C_7^{eff} = C_7^{eff}(SM)$ (left) and for $C_7^{eff} = -C_7^{eff}(SM)$ (right). The dashed lines indicate the uncertainty due to the $b \to s\gamma$ branching fraction. The plots are normalized so that the Standard Model values of C_9^{eff} and C_{10}^{eff} lie at the point (0,0) [22].

2.1.3 Correlations with other rare decays

The penguin amplitudes that contribute to $B \to K^{(*)}\ell^+\ell^-$ also appear in other rare decays of mesons. Correlations between rare decay modes can therefore be used to clarify the underlying physics. For example, the $b \to s\gamma$ branching fraction is proportional to the amplitude of the photon penguin, corresponding to the C_7^{eff} Wilson coefficient. The fact that the measured $b \to s\gamma$ branching fraction [23, 24, 25, 26] is in excellent agreement with the Standard Model prediction strongly constrains the magnitude of new physics contributions to C_7^{eff}. However the sign of C_7^{eff} has not been determined conclusively; several classes of new physics models allow these opposite sign solutions and cannot be ruled out by the $b \to s\gamma$ branching fraction alone.

Mode	World average ($\times 10^{-6}$)	Ali *et al.* ($\times 10^{-6}$)
$B \to K\ell^+\ell^-$	0.54 ± 0.08	0.35 ± 0.12
$B \to K^*\ell^+\ell^-$	1.05 ± 0.20	1.19 ± 0.39
$B \to Ke^+e^-$	$0.60^{+0.14}_{-0.12}$	0.35 ± 0.12
$B \to K\mu^+\mu^-$	$0.47^{+0.11}_{-0.10}$	0.35 ± 0.12
$B \to K^*e^+e^-$	$1.24^{+0.37}_{-0.32}$	1.58 ± 0.49
$B \to K^*\mu^+\mu^-$	$1.19^{+0.34}_{-0.29}$	1.19 ± 0.39

Table 4: World average $B \to K^{(*)}\ell^+\ell^-$ branching fractions, compared to a recent Standard Model based theoretical prediction [7].

Similarly, the Z penguin and W box amplitudes appear in decays such as $K \to \pi\nu\overline{\nu}$, $B \to K\nu\overline{\nu}$, and $B \to \ell^+\ell^-$ [27]. Unlike $b \to s\gamma$, the existing experimental limits on these branching fractions do not yet provide strong constraints on the C_9^{eff} and C_{10}^{eff} Wilson coefficients. The existing experimental constraints on new physics contributions to the Wilson coefficients are illustrated in Figure 10.

2.2   Observables

The $B \to K\ell^+\ell^-$ and $B \to K^*\ell^+\ell^-$ decays are unique among the rare B decays discussed in Section 2.1.3, in that they are three-body decays in which all particles in the final state are detectable. This section describes the observables studied in this analysis, and the relevance of each to new physics scenarios.

2.2.1   Branching fractions

Prior to the observation of $B \to K\ell^+\ell^-$ and $B \to K^*\ell^+\ell^-$ decays, searches were conducted by a number of experiments including CLEO [28, 29], UA1 [30], CDF [31, 32], and *BABAR* [33]. The $B \to K\ell^+\ell^-$ decay was finally observed by the Belle collaboration in 2002 in a sample of 31 million $B\overline{B}$ decays [34]. The first evidence for the $B \to K^*\ell^+\ell^-$ decay was then reported in 2003 by *BABAR* using a sample of 123 million $B\overline{B}$ pairs [35], and
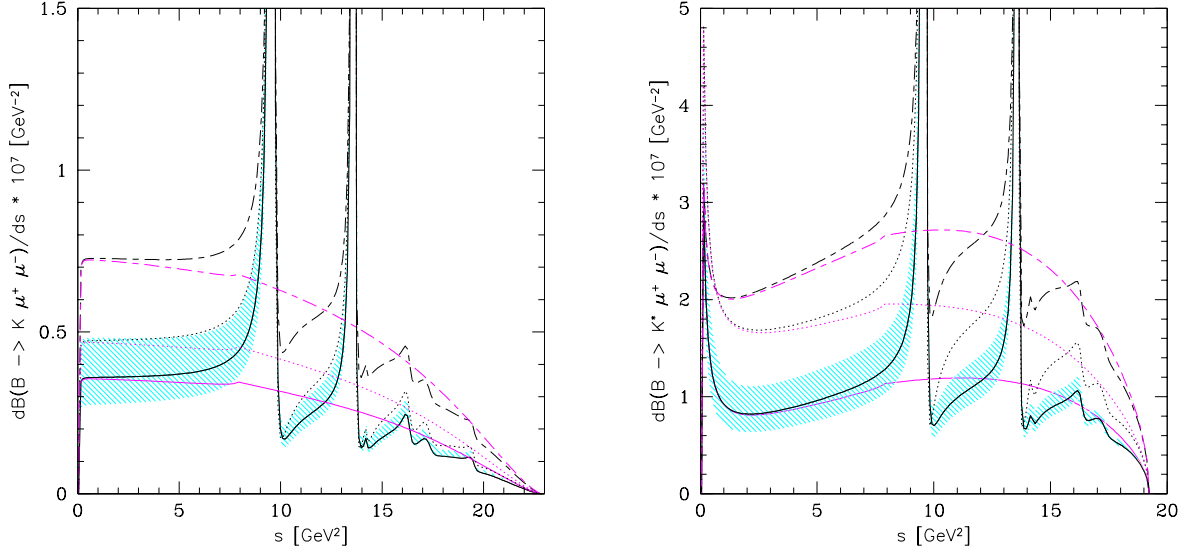
Figure 11: Partial branching fractions in $B \to K\ell^+\ell^-$ (left) and $B \to K^*\ell^+\ell^-$ (right), as a function of $s \equiv (m^2_{l^+l^-})$. The solid line corresponds to the Standard Model, with the shaded area representing the uncertainty due to the form factors. The dotted and long-short dashed lines represent allowed points in two supersymmetric models [6].

by Belle using a sample of 152 million $B\overline{B}$ pairs [36]. Table 4 lists the world average [1] of published experimental results prior to this one, together with a recent Standard Model based prediction of the total branching fractions. In addition, both Belle and *BABAR* have reported measurements of the semi-inclusive $B \to X_s\ell^+\ell^-$ rate, where X_s represents a final state with a kaon plus up to three pions [37, 38].

The existing measurements are generally consistent with the range of SM predictions; together with previous experimental limits [31], this already disfavors some possibilities, for example the "best enhancement scenario" considered in Ref. [6]. Currently, the experimental errors in these measurements are comparable to or smaller than the theoretical uncertainties due to the hadronic form factors. Therefore, in the absence of improvements to the form factor calculations, the total branching fractions in the exclusive modes will not be precision

tests of the Standard Model.

As discussed in section 2.1.1, the relationship between the branching fraction and the Wilson coefficients has a strong dependence on $m_{\ell^+\ell^-}$. Therefore, a measurement of the partial branching fraction as a function of $m_{\ell^+\ell^-}$ is sensitive to the relative contribution of the Wilson coefficients. While the partial rates also suffer from hadronic uncertainties, they are much less constrained by experimental data. The partial branching fractions are illustrated in Figure 11 for both the Standard Model and several new physics models in which the value of the C_7^{eff} Wilson coefficient is modified within the bounds allowed by the $b \to s\gamma$ measurement.

2.2.2 R_K and R_{K^*}

In contrast to the branching fractions, ratios and asymmetries in the rates can often be calculated with small theoretical uncertainties. In particular, for the $B \to K\ell^+\ell^-$ mode, the ratio:

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

has an extremely precise Standard Model prediction due to the almost complete cancellation of the form factors [14]:
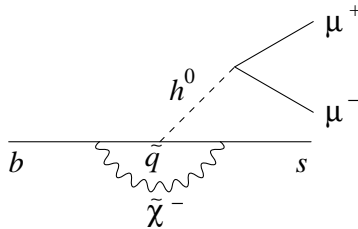
Figure 12: Example of a Feynman diagram for a Higgs penguin which can enhance R_K in supersymmetric models with large $\tan(\beta)$.
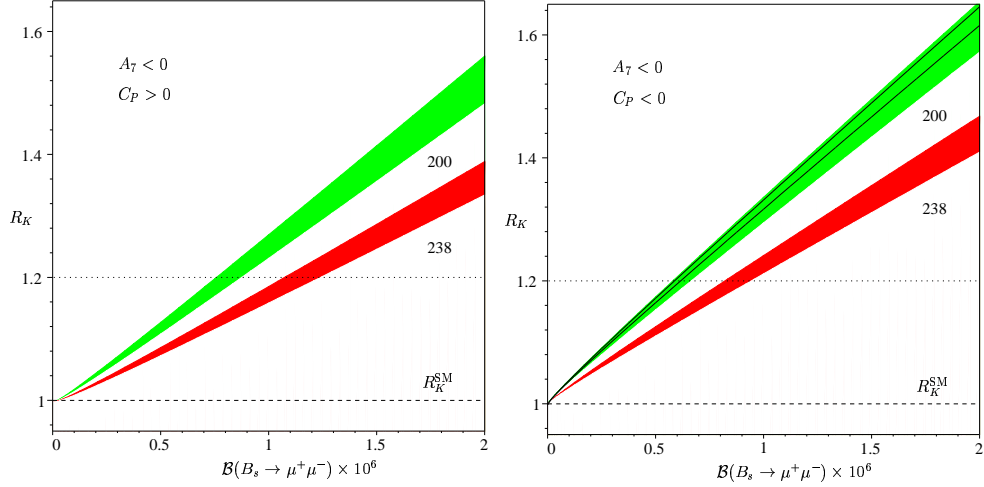
Figure 13: Correlation between R_K and $\mathcal{B}(B_s \to \mu^+\mu^-)$ for the cases $C_P > 0$ (left) and $C_P < 0$ (right). The two bands correspond to different values of the B_s meson decay constant [14].

$$R_K = 1.0000 \pm 0.0001.$$

The corresponding ratio in the $B \to K^*\ell^+\ell^-$ modes is complicated by the effect of near on-shell photons, which enhance the rate at low q^2. Unlike the $B \to K\ell^+\ell^-$ mode containing a pseudoscalar K meson, such diagrams do not violate angular momentum conservation when a vector K^* is involved. As the enhancement occurs below the threshold for producing a muon-pair, the effect is to decrease the muon to electron ratio. Thus the Standard Model prediction is $R_{K^*} \approx 0.75$, integrated over the full range of q^2 [7].

These ratios can be enhanced through diagrams with a neutral Higgs boson coupling to the lepton pair. For the Standard Model Higgs, this enhancement is negligible, as it is suppressed by the ratio $\frac{m_l m_b}{m_W{}^2}$. However, in supersymmetric models with large $\tan(\beta)$ this can occur at a measurable level [14, 39]. In the OPE formalism described in Section 2.1.1, these new contributions are described by the addition of new Wilson coefficient operators C_S

and C_P, where the S and P indicate scalar and pseudoscalar contributions, respectively [14]. The addition of new terms to the OPE contrasts with scenarios in which new physics only appears in the loop, where the three existing SM Wilson coefficients are modified.

Under the assumption that these scalar and pseudoscalar operators are left-handed, the ratio R_K is highly correlated with the branching fraction of the decay $B_s \to \mu^+\mu^-$, as shown in Figure 13. If right-handed operators C'_S and C'_P are allowed, this correlation breaks down. In this event R_K is sensitive to the sum of right and left handed operators, while $B_s \to \mu^+\mu^-$ is sensitive to their difference. Measurements of both R_K and $\mathcal{B}(B_s \to \mu^+\mu^-)$ are therefore needed to constrain contributions from these new operators [14]. While the $B_d \to \mu^+\mu^-$ decay is also sensitive to these effects, it is suppressed by the ratio of CKM elements $|V_{td}/V_{ts}|^2$, and is thus less likely than the B_s decay to be observed in this scenario.

2.2.3 CP asymmetry

The direct CP asymmetries, defined as:

$$A_{CP} \equiv \frac{\Gamma(\overline{B} \to \overline{K}^{(*)}\ell^+\ell^-) - \Gamma(B \to K^{(*)}\ell^+\ell^-)}{\Gamma(\overline{B} \to \overline{K}^{(*)}\ell^+\ell^-) + \Gamma(B \to K^{(*)}\ell^+\ell^-)}, \tag{8}$$

are expected to be vanishingly small in the Standard Model, of order 10^{-4} in the $B \to K^*\ell^+\ell^-$ mode [40]. The lack of a significant asymmetry is due to the dominance of the top quark contribution in the loop, leading to only one significant CKM contribution, $V_{ts}^*V_{tb}$, in the Hamiltonian. Observation of a large CP violating charge asymmetry in the branching fractions would therefore be strong evidence of a new source of non-SM CP violation [41].

2.2.4 Forward-backward asymmetry and K^* polarization

In the expressions for the total and partial rates, the Wilson coefficients usually enter quadratically[3], meaning that while their magnitude can be constrained from branching fraction measurements, their relative sign cannot be easily determined. Measurements of angular distributions, such as the dilepton forward-backward asymmetry A_{FB} and polarization of the K^*, are therefore required in order to resolve this ambiguity. The forward-backward asymmetry is defined as:

$$A_{FB}(\hat{s}) \equiv \frac{\int_0^1 \mathrm{d}\cos\theta^* \frac{\mathrm{d}^2\Gamma(B \to K^{(*)}\ell^+\ell^-)}{\mathrm{d}\cos\theta^* \, \mathrm{d}\hat{s}} - \int_{-1}^0 \mathrm{d}\cos\theta^* \frac{\mathrm{d}^2\Gamma(B \to K^{(*)}\ell^+\ell^-)}{\mathrm{d}\cos\theta^* \, \mathrm{d}\hat{s}}}{\mathrm{d}\Gamma(B \to K^{(*)}\ell^+\ell^-)/\mathrm{d}\hat{s}}, \qquad (9)$$

where $\hat{s} \equiv q^2/m_B^2$, and θ^* is the angle of the lepton with respect to the flight direction of the B meson. The angle θ^* is defined in the dilepton rest frame, with a sign determined by the CP state of the B meson. There are then two possible sign conventions, both of which have been used in the theoretical literature; this analysis follows the sign convention advocated in Ref. [42]. For a B^+ or B^0 meson, we define θ^* as the angle between the negatively charged lepton and the B. For a B^- or $\overline{B^0}$ meson, θ^* is the angle between the positively charged lepton and the B (Figure 14). Decays with $\cos(\theta^*) > 0$ are defined as "forward", while decays with $\cos(\theta^*) < 0$ are defined as "backward". As the $B \to K^0\ell^+\ell^-$ mode does not allow determination of its CP state directly from its decay products, it cannot be used to measure A_{FB}. In the $B \to K^*\ell^+\ell^-$ mode, the longitudinal polarization (F_L, following the notation of Ref. [43]) is defined analogously in terms of the angle θ_K, the angle between the kaon and the B meson calculated in the rest frame of the K^*.

The distribution of A_{FB} as a function of q^2 in $B \to K^*\ell^+\ell^-$ depends on the Wilson coefficients as [6]:

---

[3]The exception being the cross term proportional to $Re(C_7^{eff}C_9^{eff*})$ in Equation 7.

Figure 14: Schematic definition of $\cos\theta^*$.
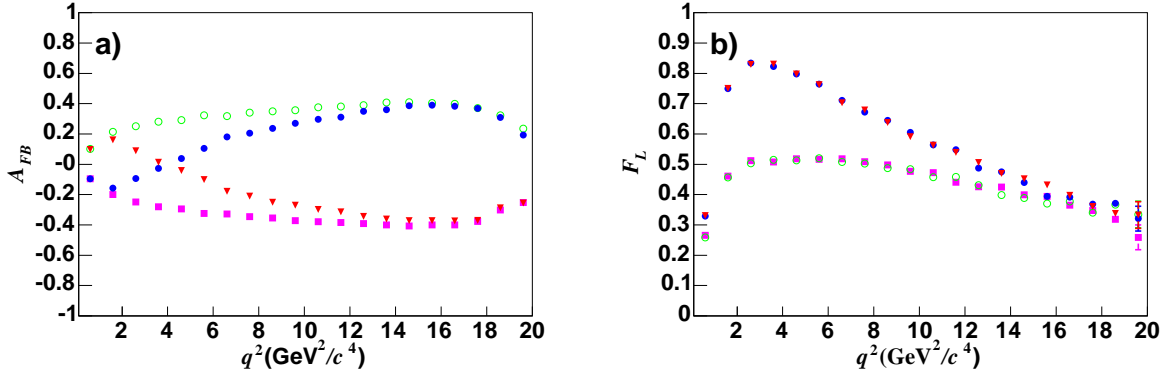
Figure 15: The q^2 dependence of a) A_{FB}, and b) F_L in $B \to K^* \ell^+ \ell^-$ decays. The points correspond to the cases of: the Standard Model (closed circles), $C_7^{eff} = -C_7^{eff}(SM)$ (open circles), $C_9^{eff} C_{10}^{eff} = -C_9^{eff} C_{10}^{eff}(SM)$ (triangles), $C_7^{eff}, C_9^{eff} C_{10}^{eff} = -C_7^{eff}(SM), -C_9^{eff} C_{10}^{eff}(SM)$ (squares).

$$\frac{dA_{FB}(B \to K^* l^+ l^-)}{d\hat{s}} \propto C_{10}^{eff}[Re(C_9^{eff}) + \frac{C_7^{eff}}{\hat{s}}]. \tag{10}$$

Models then fall into several general categories, defined by the relative sign of the Wilson coefficients. If C_9^{eff} has the sign expected in the SM, the four cases are:

1. $C_7^{eff} > 0$, $C_9^{eff} C_{10}^{eff} > 0$. The forward-backward asymmetry is positive at very low q^2 and negative at high q^2.

2. $C_7^{eff} > 0$, $C_9^{eff} C_{10}^{eff} < 0$. The forward-backward asymmetry is positive for all q^2.

3. $C_7^{eff} > 0$, $C_9^{eff} C_{10}^{eff} > 0$. The forward-backward asymmetry is negative for all q^2.

4. $C_7^{eff} < 0$, $C_9^{eff} C_{10}^{eff} < 0$. The forward-backward asymmetry is negative at very low q^2 and positive at high q^2. The Standard Model is included in this case.

Figure 15 shows the distribution of A_{FB} and the longitudinal K^* polarization F_L in $B \rightarrow K^* \ell^+ \ell^-$ decays for these four cases. The most dramatic deviations from the Standard Model can occur in cases 1 and 3 when the product of the C_9^{eff} and C_{10}^{eff} Wilson coefficients has the same magnitude but opposite relative sign as the Standard Model predicts, leading to an asymmetry which is large and negative at high q^2. A similar effect occurs at very low q^2, where C_7^{eff} is the dominant term, in case 2. However, since the magnitude of the asymmetry is smaller at low q^2, it is more difficult to distinguish these cases. In contrast to A_{FB}, the longitudinal K^* polarization is most sensitive to the sign of C_7^{eff}. While case 1 is almost indistinguishable from the Standard Model, cases 2 and 3 would lead to a reduced value of F_L at low q^2.

These cases have been considered in a number of specific new physics scenarios. In particular, the case in which C_7^{eff} has a similar magnitude but opposite sign as expected in the SM is a common feature of supersymmetric theories with a large $\tan(\beta)$ [6, 44]. Scenarios that can result in a large negative asymmetry at high q^2 have been investigated, for example, in Refs. [6, 11, 27, 42, 45]. Less dramatic effects, such as a small shift in the value of q^2 at which the asymmetry crosses zero, are also possible and have been widely studied.

In the $B \rightarrow K \ell^+ \ell^-$ decay mode, the forward-backward asymmetry is predicted to be identically zero for all regions of q^2 in all of these cases. The only exception to this comes if new scalar amplitudes of the type described in Section 2.2.2 are introduced [13, 46, 47]; however, even in that case any asymmetry is expected to be of order 0.01 or less for the

Mode	A_{FB} (low)	A_{FB} (high)	F_L (low)	F_L (high)
$B \to K^* \ell^+ \ell^-$ case 1	-0.03	-0.36	0.66	0.46
$B \to K^* \ell^+ \ell^-$ case 2	$+0.26$	$+0.39$	0.46	0.41
$B \to K^* \ell^+ \ell^-$ case 3	-0.26	-0.39	0.46	0.41
$B \to K^* \ell^+ \ell^-$ case 4 (SM)	$+0.03$	$+0.36$	0.67	0.48
$B \to K^{\pm} \ell^+ \ell^-$	0.00	0.00	N/A	N/A

Table 5: Predicted numerical values of A_{FB} and F_L in the Standard Model and new physics scenarios in which the magnitude of the Wilson coefficients is unchanged. The theory predictions are obtained from simulations based on the form factor model of [18].

muon channel [48]. Since no significant A_{FB} is expected even in the presence of new physics, the $B \to K\ell^+\ell^-$ mode acts primarily as an important cross-check of the analysis procedure.

A precise measurement of the shape of the A_{FB} and F_L distributions would require extremely large data samples. However even with the currently available data, the general features of these distributions can be determined. The *BABAR* data sample used in this analysis allows a measurement of A_{FB} and F_L integrated over two regions of dilepton mass:

- The "low" region above the photon pole but below the vetoed J/ψ resonance: $0.1 < q^2 < 8.41$ GeV$^2/c^4$.

- The "high" region above the J/ψ veto, excluding the region of the $\psi(2S)$ resonance: $q^2 > 10.24$ GeV$^2/c^4$.

The theoretical predictions for A_{FB} and F_L in these ranges of q^2 are listed in Table 5, for the cases where the Wilson coefficients have the magnitude expected in the Standard Model. In the low region, the regions of negative and positive A_{FB} roughly cancel in the SM and in case 1, leading to a small asymmetry. In cases 2 and 3, the asymmetry does not cross zero, resulting in a large positive or negative A_{FB}.

2.2.5 Lepton flavor-violation

The rates of the lepton flavor-violating decays $B \to Ke\mu$ and $B \to K^*e\mu$ are expected to be far below the current experimental sensitivity in the Standard Model with neutrino mixing. Any evidence for these decays would therefor be an indication of non-Standard Model contributions. There are few theoretical predictions for these decays, however the experimental limits can in principle be used to bound the contribution from leptoquark models [49].

3 PEP-II and the *BABAR* detector

The *BABAR* experiment is located at the Stanford Linear Accelerator Center (SLAC) at Stanford University. The primary goal of *BABAR* is the precision study of CP violation and rare decay processes in the B meson system. Complementary programs in charm and τ physics are also supported. This section describes the *BABAR* detector and the environment in which it operates.

3.1 PEP-II

The PEP-II facility is an asymmetric e^+e^- collider, in which the SLAC LINAC is used to inject 9.0 GeV electrons and 3.1 GeV positrons into high-energy (HER) and low-energy (LER) storage rings. The beams collide at a center of mass energy equal to the mass of the $\Upsilon(4S)$ particle (10.58 GeV/c^2), which has a branching fraction to B-meson pairs of nearly 100% [1]. With the asymmetric beam energies, the $\Upsilon(4S)$ system is Lorentz boosted by a factor $\beta\gamma = 0.56$. The boost allows the measurement of the B and $\overline{B}$ decay times critical for studying time-dependent CP violation. However, it is much less important for analyses of rare decays such as $B \to K^{(*)}l^+l^-$.

The machine has operated efficiently since 1999, delivering a total integrated luminosity of over 350 fb^{-1} (Figure 16). The data has been collected in five run periods, interrupted by

Run period	Beginning date	Ending date	Integrated luminosity (fb^{-1})
1	Oct. 1999	Nov. 2000	19.5
2	Feb. 2001	June 2002	60.3
3	Nov. 2002	June 2003	31.1
4	Sept. 2003	July 2004	99.8
5	April 2005	-	102.6

Table 6: Run periods and on-resonance luminosity available for analysis at *BABAR*.
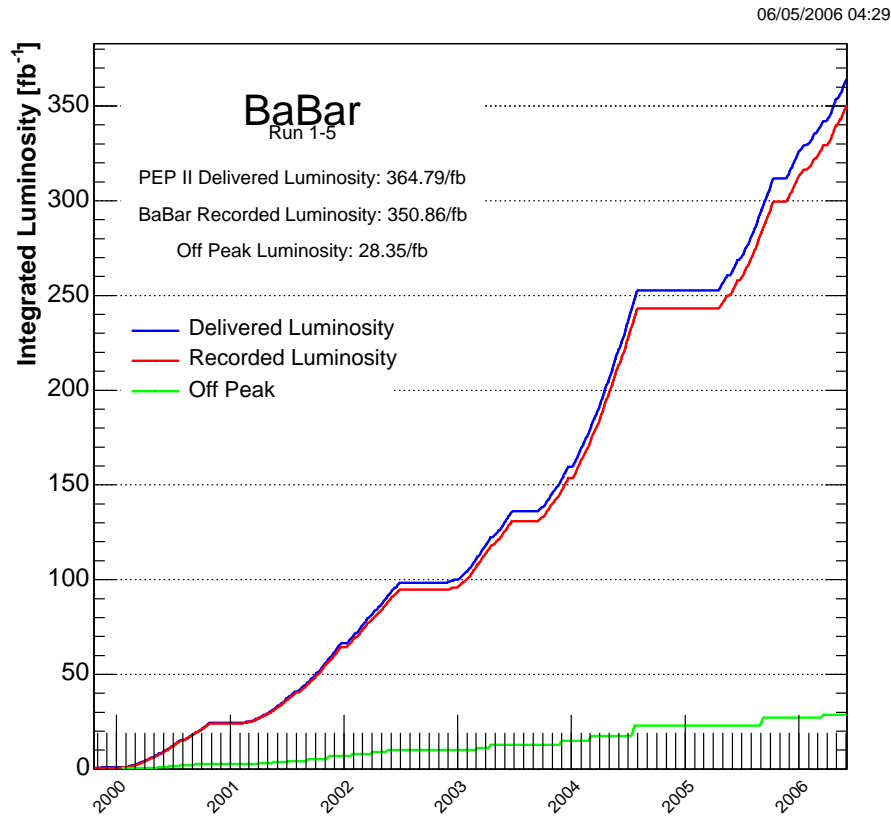
Figure 16: Integrated luminosity as a function of time.

shutdowns for maintenance and upgrades; the dates of the run periods are shown in Table 6. The analysis described in this document uses 208 fb^{-1} of data collected through the summer of 2004, corresponding to the end of run 4. Approximately $5-10\%$ of the data is collected at energies 40 MeV below the $\Upsilon(4S)$ resonance, to be used for study of backgrounds. Since early 2004, the machine has operated in "trickle" mode, allowing continuous injection of bunches to the storage ring. Prior to this, data-taking was typically halted every ~ 45 minutes to restore the electron and positron currents. To date, the best instantaneous luminosity achieved is 1.09×10^{34} cm^{-2} s^{-1}. The best 24-hour run period resulted in the collection of
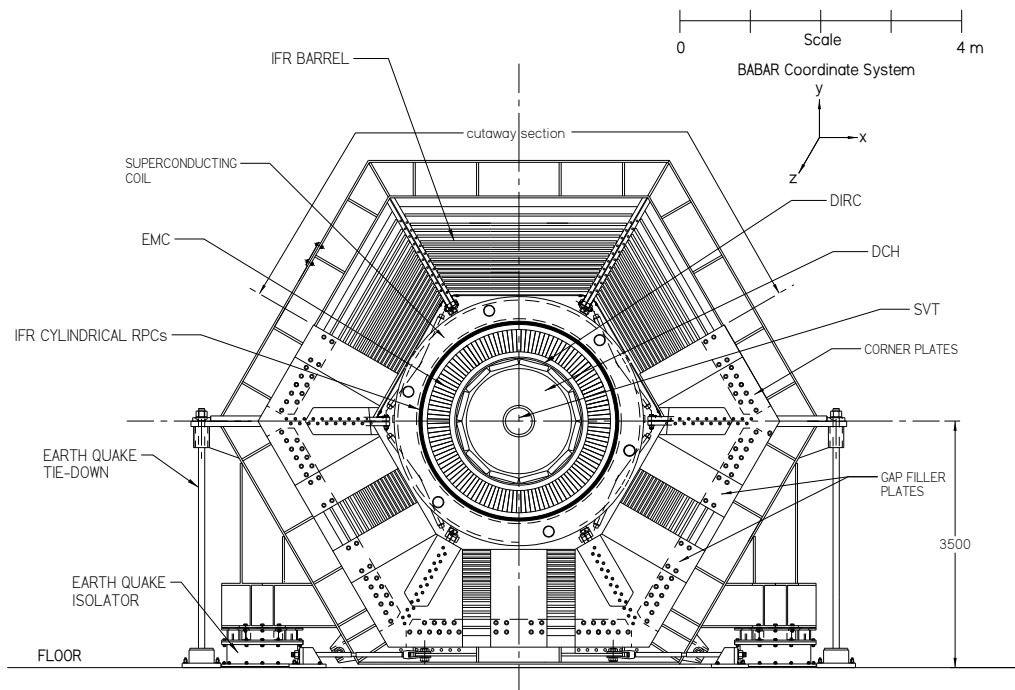
Figure 17: End view of the *BABAR* detector.

871 pb^{-1}.

3.2   The *BABAR* detector

BABAR [50] is a general purpose detector, designed to support a wide variety of analyses in flavor physics. The range of physics studied requires excellent vertexing, tracking of charged particles, and identification of both charged and neutral particles over a wide range of momentum and angular acceptance.

The layout of the detector is shown in Figure 17. The inner detector includes a silicon vertex tracker, drift chamber, ring-imaging Cherenkov detector, and electromagnetic calorimeter. Surrounding the inner detector is a superconducting solenoid, producing a $1.5T$ magnetic field. The steel flux return is instrumented for muon and neutral hadron identification.
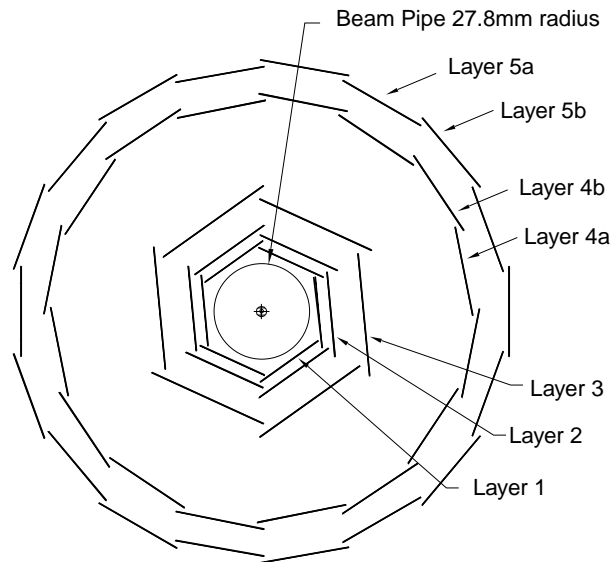
Figure 18: Layout of layers in the silicon vertex tracker.

3.2.1 Silicon Vertex Tracker

The silicon vertex tracker (SVT) consists of five layers of double-sided silicon microstrip detectors. The SVT provides vertexing information, tracking of low momentum pions, and contributes to hadron identification through measurement of the energy loss dE/dx.

The five layers, consisting of 6-18 modules each, are arranged at radial distances of 32-144 mm. The geometry of the SVT layers is shown in Figure 18. Each module is made up of 4-8 silicon strip detectors, oriented either parallel to the beam axis or transversely to the beam axis. While the modules of layers 1-3 are straight, the modules of layers 4 and 5 are arch-shaped with the forward and backward ends rotated toward the beam axis, as shown in Figure 19. The arch-shaped design was chosen in order to maximize the angular acceptance of the SVT, while reducing the angle of incidence with respect to the silicon for low polar angle tracks. The resulting polar angle coverage is $20° < \theta < 150°$. To ensure proper global alignment of the SVT with the outer detector, calibrations are performed on a run-by-run
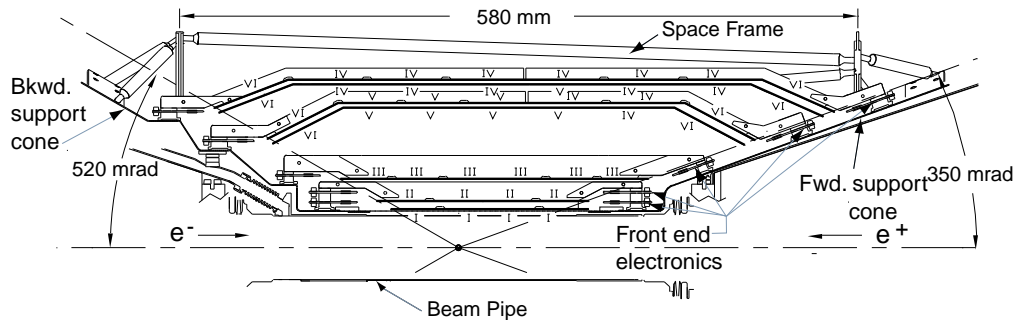
Figure 19: Side view of the SVT.

basis. Alignment of the individual silicon wafers within the SVT is performed on longer time scales, using large data control samples of Bhabha and dimuon events [51].

The SVT measures the track vertices in z with a resolution of $< 40\,\mu$m, compared to the $\sim 250\,\mu$m mean Δz separation between the decay vertices of the two B mesons produced in an event. The hit finding efficiency is above 97%, as determined from samples of Bhabha and dimuon events. The dE/dx information is useful for identification of hadrons with momenta below 500 MeV/c. In this momentum region the SVT provides approximately 2σ separation between kaons and pions, as defined by the difference in the mean measured dE/dx divided by the resolution [50].

3.2.2 Drift Chamber

The drift chamber (DCH) provides precision measurements of the momentum and angles of charged particles, as well as low momentum particle identification. The DCH has a cylindrical configuration, with a length of 2.8 m and inner and outer radii of 23.6 cm and 81 cm, respectively. The volume is filled with a gas mixture containing 80% helium, 20% isobutane, and trace amounts of water vapor. The interior contains 7,104 hexagonal cells
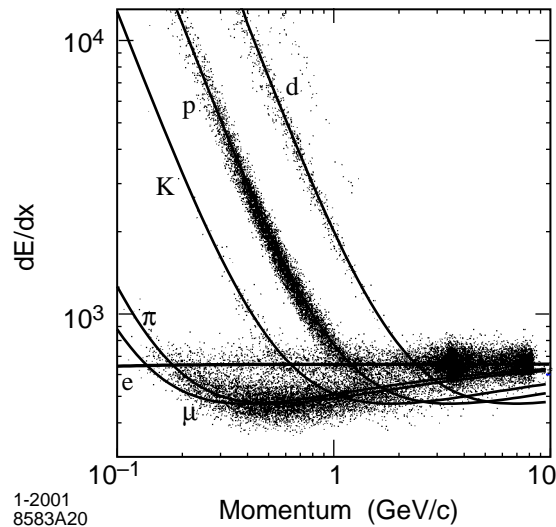
Figure 20: dE/dx measured in the DCH for different charged particle species. The points are derived from data control samples, and the solid lines show the predicted values for each particle species.

each consisting of a sense wire surrounded by ground wires. The cells are arranged in 40 layers, approximately half of which are oriented at angles with respect to the z-axis in order to give longitudinal position information. The first ~ 20 fb^{-1} of *BABAR* data were collected with the DCH wire layers operating at voltages of $1960V$ or $1900V$. The remainder was collected at an operating voltage of $1930V$.

The tracking algorithm uses a Kalman filter [52] approach to find helical tracks in the DCH. Secondary algorithms then attempt to associate unassigned hits to tracks, and to find tracks that do not originate from the interaction point. The tracks are then extrapolated and matched to any associated hits in the SVT [50].

The track-finding efficiency of the DCH, averaged over momentum and angle, is $\sim 96\%$, relative to the number of tracks found in the SVT. The resolution of the transverse momentum, measured with control samples of cosmic rays, is [53]:

$$\frac{\sigma(p_T)}{p_T} = 0.45\% + 0.13\% \times p_T \ \mathrm{GeV}/c.$$

Through the measured ionization energy loss (dE/dx), the DCH also provides efficient separation between charged kaons, pions, and protons. The dE/dx resolution is approximately 7% at low momentum; this allows $> 2\sigma$ separation of charged kaons and pions at momenta up to $\sim 0.7 \ \mathrm{GeV}/c$ (Figure 20).

3.2.3 Detector of Internally Reflected Cherenkov light

The detector of internally reflected Cherenkov light (DIRC) allows identification of long-lived charged hadrons with high momentum. Charged particles traveling faster than the speed of light in a medium of index of refraction n will emit Cherenkov light with a characteristic polar angle θ_C with respect to the track trajectory. The Cherenkov angle is related to the relativistic velocity $\beta = v/c$ of the particle by the expression

$$cos(\theta_C) = \frac{1}{\beta n}.$$

The measurement of the Cherenkov angle can then be combined with the DCH measurement of the track momentum and angle to determine the best hypothesis for the identity of the particle. In the DIRC, Cherenkov light from a charged particle is transmitted via total internal reflection within a set of 144 quartz bars. The light is delivered to a water filled standoff box at the backward end of the detector. The forward end is uninstrumented, thus a mirror is used to reflect light toward the instrumented, backward end. There an array of $\sim 11,000$ photomultiplier tubes (PMTs) covering the interior surface of the standoff box is used to reconstruct the opening angle of the ring of Cherenkov light emitted by the particle, correcting for the small difference between the index of refraction of quartz ($n = 1.473$) and the water ($n \approx 1.346$) [54]. The geometry of the DIRC is shown in Figure 21.
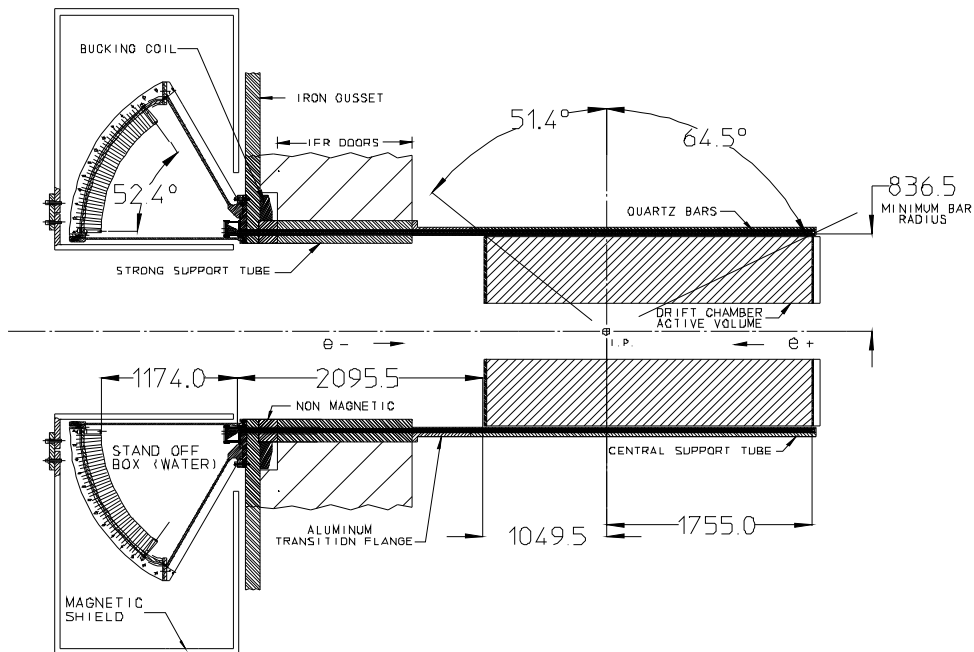
Figure 21: Geometry of the DIRC.

The DIRC reconstruction algorithm combines the tracking information provided by the DCH with the measured Cherenkov angle to discriminate between hadron species. In addition, the timing information provided by the PMTs is used to determine which track in the event produces the photons. Two algorithms are then used to find the best particle hypothesis for the track. A ring fitting algorithm evaluates the Cherenkov angle and difference between the number of expected and number of measured photons on a track-by-track basis. A second method evaluates a global likelihood, using all DIRC information in the event to find the best particle hypothesis for all tracks in the event [54].

The DIRC Cherenkov angle resolution is 2.5 mrad, as evaluated using samples of $e^+e^- \rightarrow \mu^+\mu^-$. The particle identification performance of the DIRC is evaluated using control samples of D^*, K_S^0, and Λ decaying to kaons, pions, and protons. The DIRC is particularly effective at separating charged kaons and pions with lab momenta above 0.7 GeV/c, as illustrated
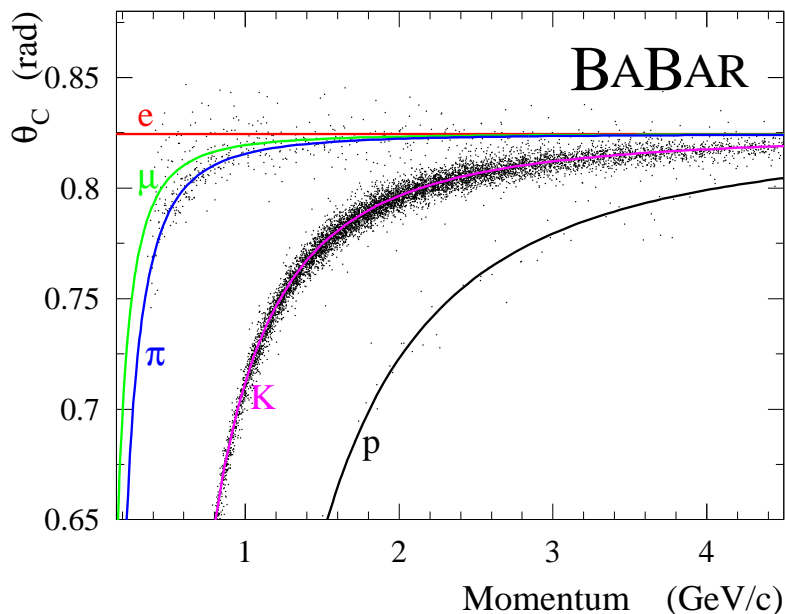
Figure 22: Cherenkov angle θ_C vs momentum for kaons (points). The solid lines correspond to the expected θ_C for other charged particle species.

in Figure 22. The kaon-pion separation power is quantified by the difference in the mean measured angle divided by the angular resolution. The separation power is approximately 4σ at a momentum of 3 GeV/c, and about 2.5σ at 4.1 GeV/c [54].

3.2.4 Electromagnetic Calorimeter

Measurement of neutral energies is accomplished via an electromagnetic calorimeter (EMC). The calorimeter allows identification of photons, and reconstruction of neutral π^0 and η hadrons. In addition, the ratio E/p (energy deposited in the calorimeter over momentum measured in the DCH) and the shower shape measured in the EMC provide the primary means of identifying electrons. The EMC consists of 6580 Thallium-doped cesium iodide crystals, split into barrel and endcap regions of the detector. The crystals in the barrel
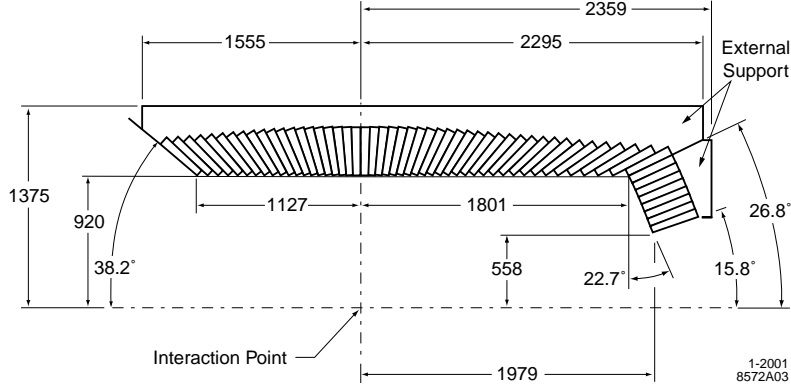
Figure 23: Side view of the crystal geometry in the electromagnetic calorimeter.

are arranged in 48 axially symmetric rings, while the endcap is a conic section consisting of 8 rings, with front and back surfaces tilted at 22.7° to the vertical. The range of polar angles covered by the EMC is 15.8° to 141.8°, or about 90% of the total solid angle in the center-of-mass system [50]. The geometry of the EMC is shown in Figure 23.

Control samples of electrons from Bhabha scattering and photons from π^0 and radiative χ_{c1} decays are used to determine the energy resolution of the calorimeter. The resolution can be expressed as a term proportional to the inverse fourth root of the energy added in quadrature ($\oplus$) with a constant term [50]:

$$\frac{\sigma_E}{E} = \frac{(2.32 \pm 0.30)\%}{\sqrt[4]{E(GeV)}} \oplus (1.85 \pm 0.12)\%.$$

The resolution measured in data control samples and in simulation as a function of photon energy is shown in Figure 24.

The shower shape provides discrimination between electrons and hadrons through both lateral and longitudinal shape variables. The lateral shape is described by the expression

$$LAT = \frac{\sum_{i=3}^{N} E_i r_i^2}{\sum_{i=3}^{N} E_i r_i^2 + E_1 r_0^2 + E_2 r_0^2},$$
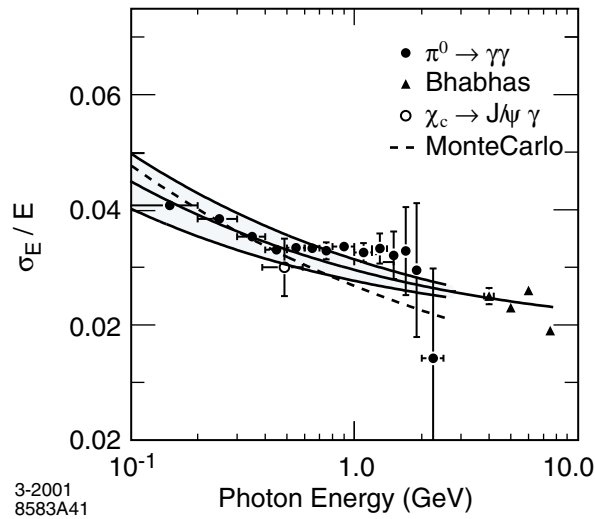
Figure 24: Electromagnetic calorimeter energy resolution vs photon energy.

where N is the number of crystals associated with a shower, E_i is the energy deposited in the i-th crystal, r_i is the lateral distance from the center of the shower and the i-th crystal, and r_0 is the average distance between two crystals. For electromagnetic showers most of
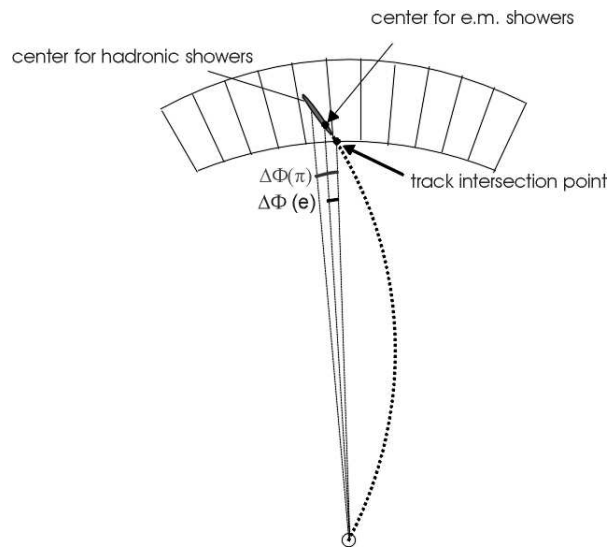
Figure 25: Schematic definition of the $\Delta\Phi$ variable in the calorimeter.
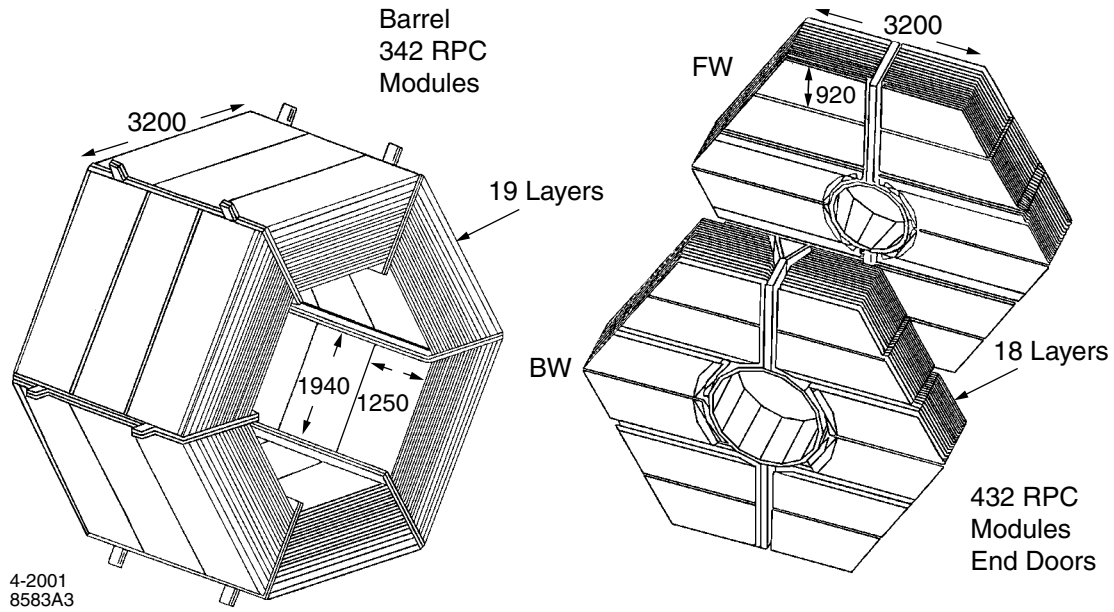
Figure 26: Geometry of the instrumented flux return.

the energy is deposited in a small number of crystals near the center of the shower, thus this variable is peaked near 0.3. In contrast hadronic showers deposit a larger fraction of their energy farther from the crystal center, resulting in a LAT distribution that peaks in the range 0.4-1.0. Information about the longitudinal shower shape can be extracted from $\Delta\Phi$, the difference between the polar angle where the track intersects the crystal face and the shower center (Figure 25. For a charged track curving in the XY plane, the center of the shower will be displaced in the angle Φ from the point of entry to the face of the calorimeter. In electromagnetic showers this displacement is generally small, leading to a strongly peaked $\Delta\Phi$ variable. In hadronic showers the maximum of the shower occurs farther from the point of entry, resulting in a broader distribution in $\Delta\Phi$.
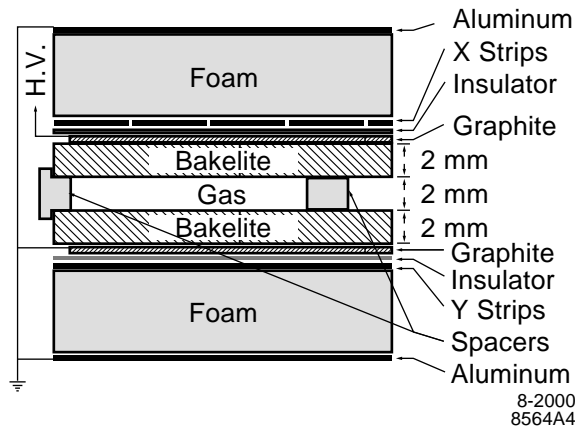
Figure 27: Cross-section of a Resistive Plate Chamber.

3.2.5 Instrumented Flux Return

The instrumented flux return (IFR) is designed to provide identification of muons and neutral hadrons. The IFR consists of approximately 2000 m^2 of Resistive Plate Chambers (RPC), interleaved with the steel of the magnet flux return. The barrel region was designed with 19 layers of RPCs and has a total iron thickness of 65 cm. The forward and backward endcaps were designed with 18 RPC layers and have a total iron thickness of 60 cm. For a particle traversing the detector at a polar angle of 31°, this corresponds to a total absorbing power of 5.4 λ, including the material in the EMC. Here λ is the nuclear interaction length, or mean distance traveled by a hadron between collisions with nuclei (about 17 cm for iron).

The RPCs are designed to detect streamers from ionizing particles passing through the IFR. Each RPC consists of a 2 mm gap filled with a gas mixture of isobutane, argon, and freon. Surrounding the gap are two layers of bakelite covered with linseed oil. Painted on the bakelite and connected to high voltage and ground are two conducting layers of graphite, with a nominal surface resistivity of 100 $k\Omega$/square. Streamers passing through the chamber induce signals in orthogonal rows of aluminum strips which cover the graphite and bakelite
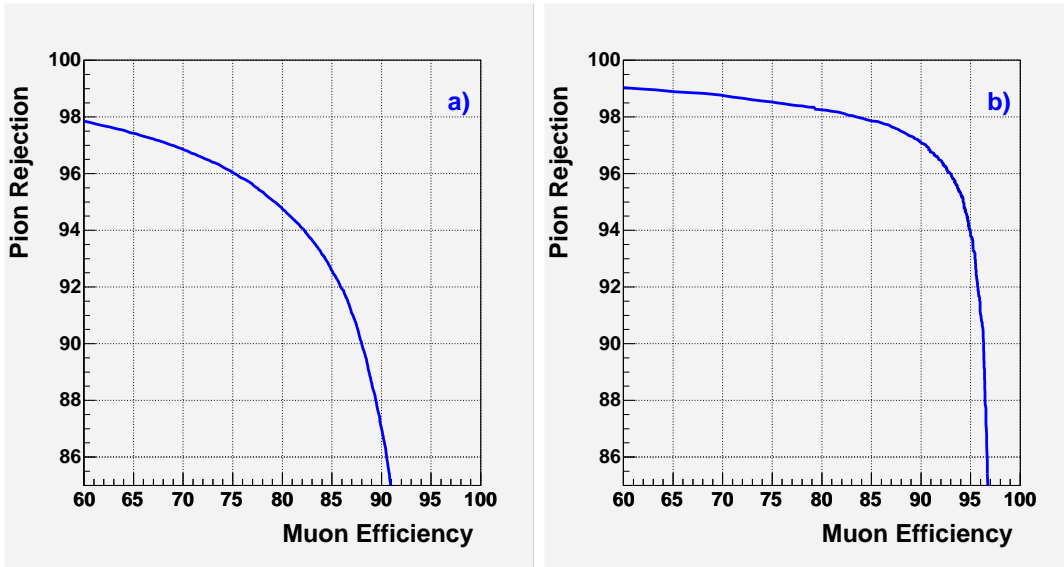
Figure 28: Muon efficiency vs. pion rejection for (a) the barrel region and (b) the forward endcap region of the IFR, for data collected in run 3, and muons and pions with momentum in the range $2.0 < p < 4.0$ GeV/c.

layers. The signals from the orthogonal strips provide a two-dimensional measurement of the streamer position.

The muon reconstruction uses a Kalman filter algorithm to find three-dimensional track clusters in the IFR, using the track extrapolated from the DCH as a starting point. As the Kalman steps through the layers of the IFR, two-dimensional clusters in the RPCs are associated with the track if they fall within a wide (17σ) window.

Many of the RPCs have suffered decreasing efficiencies over the lifetime of the BABAR experiment [55]. Multiple failure modes were identified, including improper curing of the linseed oil layer, increased bakelite resistance in the presence of dry gas, and degradation of the graphite conducting layer. Prior to Run 3, the chambers in the forward endcap were replaced with a new generation of RPCs, with additional brass and iron absorber added [56]. In the new configuration, the total absorbing power at an angle of 31° was increased to

$e^+e^- \rightarrow$	Cross-section (nb)
$b\overline{b}$	1.05
$c\overline{c}$	1.30
$s\overline{s}$	0.35
$u\overline{u}$	1.39
$d\overline{d}$	0.35
$\tau^+\tau^-$	0.94
$\mu^+\mu^-$	1.16
e^+e^-	~ 40

Table 7: Production cross-sections at $\sqrt{s} = 10.58$ GeV.

7.3 interaction lengths. In addition, a new double layer of belt RPCs was added to the region of overlap between the barrel and forward endcap. The new RPCs have generally performed well, with high and stable muon identification efficiencies [57]. The muon vs. pion discrimination in the barrel and forward endcap regions obtained in run 3 data is illustrated in Figure 28.

3.2.6 Trigger and environment

The production cross-sections at $\sqrt{s} = m(\Upsilon(4S))$ for $b\overline{b}$ and for light-quarks and leptons are listed in Table 7. The *BABAR* trigger is designed to maintain near 100% efficiency for $B\overline{B}$ events, with final event rates up to ~ 300 Hz for the data used in this analysis.

The trigger is implemented as a two tier system, with a Level 1 (L1) hardware trigger and a Level 3 (L3) software trigger. The L1 trigger is based on detection of charged tracks in the DCH, showers in the EMC, and tracks in the IFR. The resulting L1 trigger rate is approximately 1 kHz for a luminosity of 3×10^{33} cm^{-2} s^{-1} during normal operating conditions. The L3 software trigger further refines the event selection in order to reduce the rate of Bhabhas and beam background events. For data collected near the end of run 4 with luminosities near 10^{34} cm^{-2} s^{-1}, the final L3 trigger rate is about 300 Hz. Approximately

75% of the total trigger rate is allocated to physics events, and the remainder to calibration samples used to study the trigger performance [50]. The trigger efficiency is $> 99.9\%$ for high-multiplicity $B\overline{B}$ events, $\sim 95\%$ for continuum light quark pair events, and $90 - 95\%$ for $e^+e^- \to \tau^+\tau^-$ events [50].

4   Event Selection

The following decay modes of the B_d meson are studied in this analysis:

1. $B^{\pm} \rightarrow K^{\pm}l^+l^-$

2. $B^0 \rightarrow K^0_S l^+l^-$, where $K^0_S \rightarrow \pi^+\pi^-$

3. $B^0 \rightarrow K^{*0}l^+l^-$, where $K^{*0} \rightarrow K^{\pm}\pi^{\mp}$

4. $B^{\pm} \rightarrow K^{*\pm}l^+l^-$, where $K^{*\pm} \rightarrow K^0_S\pi^{\pm}$ and $K^0_S \rightarrow \pi^+\pi^-$

For the measurements of the branching fractions and asymmetries, the l^+l^- can be either an e^+e^- or a $\mu^+\mu^-$ pair, giving a total of eight signal modes. In the lepton flavor-violation search, the particle identification is modified to instead select candidate events with an $e^{\pm}\mu^{\mp}$ pair. This section describes the particle identification, kinematic selections, and background suppression strategy used to select signal events.

4.1   Particle identification

The analysis relies heavily on the charged particle identification abilities of the *BABAR* detector. Particle identification (PID) is performed by algorithms that combine information from the different sub-detector systems, usually with multivariate techniques such as neural networks or likelihood ratios. High purity control samples are then used to evaluate efficiencies and mis-ID rates. Tables binned in p, θ, ϕ, charge, and data taking time are used to correct for any discrepancies between data and Monte Carlo simulation in the particle selection.

In reconstructing $B \rightarrow K^{(*)}l^+l^-$ candidate events, we first require a well-measured track in the DCH for all particle species. The track quality requirement includes having a polar

angle within the detector acceptance, a distance of closest approach (DOCA) to the inter-action point of less than 1.5 cm in the XY plane, a DOCA of less than 10 cm along the z axis, and for the leptons a minimum of 12 hits in the DCH. We further require strict particle identification of both leptons, and kaon identification for charged kaon candidates. In the K^* modes, we require that the charged pion candidate from the K^* decay must fail the kaon identification algorithm.

4.1.1 Electron identification

Electron identification is based on the following quantities measured in the EMC, DIRC, and DCH:

- The ratio E/p of the shower energy deposited in the calorimeter to the track momentum measured in the DCH.

- The shower shape of the cluster in the EMC as described in Section 3.2.4.

- The difference between the dE/dX measured in the DCH and the expected dE/dX under the electron hypothesis.

- The Cerenkov angle θ_c measured in the DIRC.

The information from the individual subdetectors is combined into a global likelihood:

$$\mathcal{L}_{ipart} = \mathcal{L}_{ipart}^{EMC} \times \mathcal{L}_{ipart}^{DCH} \times \mathcal{L}_{ipart}^{DIRC}$$

where *ipart* is the particle species in question (e, π, K, or p). Events are selected or rejected based on the ratio of likelihoods, weighted by the *a priori* multiplicities of the particle species.

The efficiency of the electron selection is evaluated using samples of electrons kinemat-ically selected from radiative bhabha ($e^+e^- \rightarrow e^+e^-\gamma$) events. Pion misidentification is
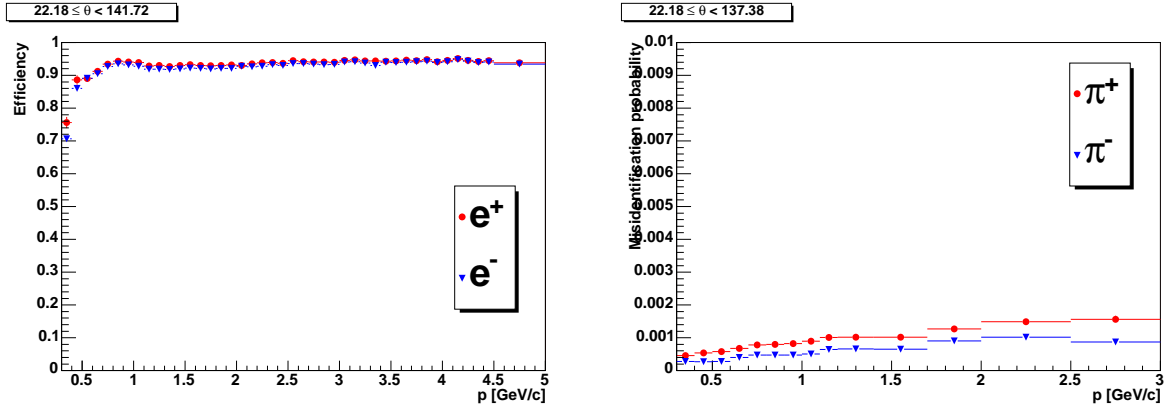
Figure 29: Electron ID efficiency (left) and pion misidentification rate (right) as a function of momentum for the selection used in this analysis. The performance is shown for positively (circles) and negatively (triangles) charged particles.

evaluated using control samples derived from decays of τ and K_s^0. From these samples, the efficiency is determined to be $\sim 94\%$ for electron momenta above 1 GeV/c, with a corresponding pion mis-ID probability of less than 0.1%. The efficiency and misidentification rates as a function of momentum are displayed in Figure 29.

4.1.2 Muon identification

Muon identification is accomplished by means of a neural network algorithm, using information from the IFR and EMC. The muon identification algorithm is described in more detail in Appendix A The detector quantities considered are:

- The number of measured interaction lengths of the muon candidate.

- The difference between the number of measured interaction lengths and the expected number of interaction lengths under the muon hypothesis.

- The continuity of the track in the IFR, defined as:
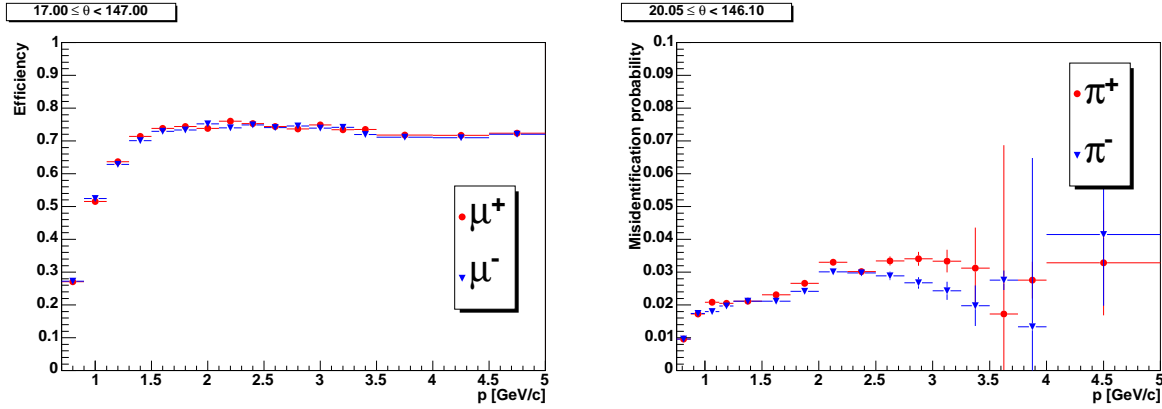
Figure 30: Muon ID efficiency (left) and pion misidentification rate (right) as a function of momentum for the selection used in this analysis. The performance is shown for positively (circles) and negatively (triangles) charged particles.

$$cont. = \frac{N_{layers}}{L_{last} - L_{first} + 1},$$

where L_{first} is the innermost layer hit, L_{last} the outermost layer hit, and N_{layers} is the total number of layers hit in a three-dimensional cluster.

- The average multiplicity of strips hit per layer.

- The standard deviation of the average strip multiplicity.

- The goodness of fit (χ^2/dof) of a third order polynomial fit to the hits in the three-dimensional cluster.

- The goodness of fit with respect to the track extrapolation from the DCH.

- The energy deposited in the EMC by the muon candidate.

The efficiency of the muon selection is evaluated using a control sample of $e^+e^- \rightarrow \mu^+\mu^-\gamma$ events. Pion misidentification is evaluated using control samples derived from D^* decays.

From these samples, the efficiency is determined to be $\sim 70\%$ for muon momenta above 1 GeV/c, with a corresponding pion mis-ID probability of $2-3\%$. The muon ID performance as a function of momentum is displayed in Figure 30. As a function of polar angle, the efficiency varies smoothly over the transition between the forward endcap and barrel regions. However, there is some loss in the region of overlap between the barrel and backward endcap of the IFR. The pion misidentification probability is typically a factor of two lower in the forward endcap region than in the barrel or backward endcap regions.

4.1.3 Kaon and pion identification

Hadron ID algorithms combine information from the SVT, DCH, and DIRC into a likelihood function for each particle hypothesis:

$$\mathcal{L}_{ihad} = \mathcal{L}_{ihad}^{DIRC} \times \mathcal{L}_{ihad}^{DCH} \times \mathcal{L}_{ihad}^{SVT}$$

where $ihad$ is the hadron species (π, K, or p). The detector quantities considered in the likelihood are:

- The difference between the dE/dX measured in the DCH and the expected dE/dX under the appropriate hadron hypothesis.

- The difference between the dE/dX measured in the SVT and the expected dE/dX under the appropriate hadron hypothesis.

- The Cherenkov angle θ_c measured in the DIRC.

- The number of observed photons in the DIRC.

- The quality of the track prior to reaching the DIRC.
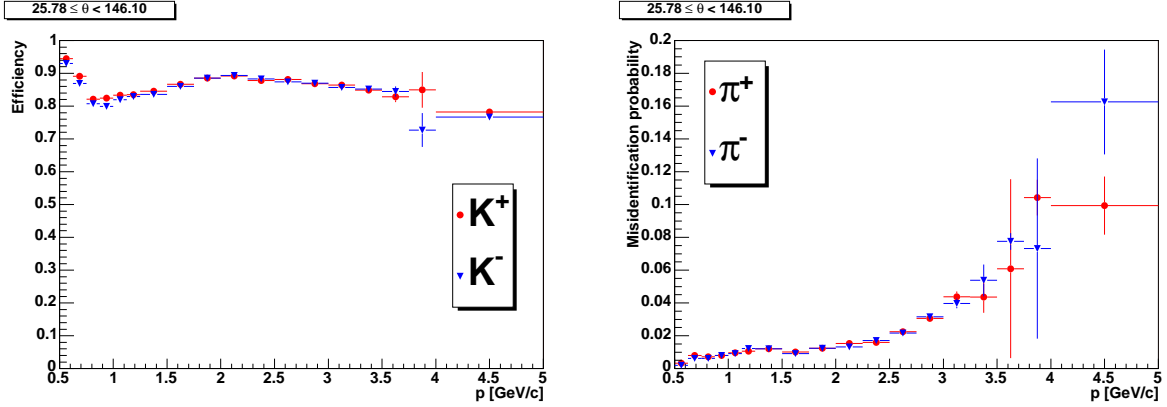
Figure 31: Kaon ID efficiency (left) and pion misidentification rate (right) as a function of momentum for the selection used in this analysis. The performance is shown for positively (circles) and negatively (triangles) charged particles.

In addition, kaon candidates are required to fail the electron ID described in Section 4.1.1.

The kaon efficiency is evaluated using a sample of kaons from the decay $D \to K\pi$, where the D is selected from the decay of a D^*. Pion misidentification is evaluated using pions from the same source. The hadron ID performance as a function of momentum is displayed in Figure 31. For kaons with momenta below ~ 1.0 GeV/c, where K/π discrimination is provided by the dE/dX measured in the DCH, the efficiency is $80 - 95\%$ with pion mis-ID rates below 1%. Above 1 GeV/c, K/π separation is primarily due to the Cherenkov angle measurement in the DIRC. The kaon efficiency in this range is stable at $80 - 90\%$, while the pion mis-ID increases with momentum to $10 - 15\%$ above 4 GeV/c. As noted previously, pions in the K^* modes are required to fail the kaon selection criteria. The pion selection efficiency is therefore $85 - 99\%$, while the kaon mis-ID probability is $5 - 20\%$.

4.2 Kinematic Selection

Events are selected if the decay products of the B meson satisfy the following kinematic requirements:

- Electrons must have momentum $p_{LAB} \geq 0.3$ GeV/c.

- Muons must have momentum $p_{LAB} \geq 0.7$ GeV/c.

- K_S^0 candidates must satisfy $0.4887 < m_{K_S^0} < 0.5073$ GeV/c^2, corresponding to a window of approximately 3σ in terms of the detector mass resolution.

- K^* candidates must satisfy $0.7 < m_{K^*} < 1.1$ GeV/c^2.

A significant fraction of electrons in the sample will radiate energy via bremsstrahlung, distorting the kinematics of the reconstructed B meson. Electrons are therefore paired with nearby photons which have $E_\gamma > 0.030$ GeV and lie within a 0.035 radian cone about the electron trajectory, in order to recover energy lost in such events. In approximately 30% of $B \to K^{(*)}e^+e^-$ signal events one or both of the electrons is brem-recovered in this way. In the $B \to Ke^+e^-$ modes, electrons are required to have an invariant mass $m_{e^+e^-} \geq 0.3$ GeV in order to remove background from photon conversions in the detector. In the $B \to K^*e^+e^-$ modes, there is a significant rate at low dilepton invariant mass, due to the pole in the photon penguin. Therefore the conversion veto is applied only if the radius of the vertex is greater than the 2 cm inner radius of the beam pipe, in order to preserve acceptance in these modes.

Signal B-meson candidates are selected based on kinematic variables derived from the momentum of the B in the lab frame $\mathbf{p_B}$, the energy of the B in the center-of-mass frame E_B^*, the energy and momentum of the $\Upsilon(4S)$ in the lab frame E_0 and $\mathbf{p_0}$, and the center-of-mass energy of the beams $\sqrt{s}$. These quantities are combined into two nearly uncorrelated variables m_{ES} and ΔE, defined as:

$$m_{ES} \equiv \sqrt{(s/2 + \mathbf{p_0} \cdot \mathbf{p_B})^2 / E_0^2 - p_B^2},$$
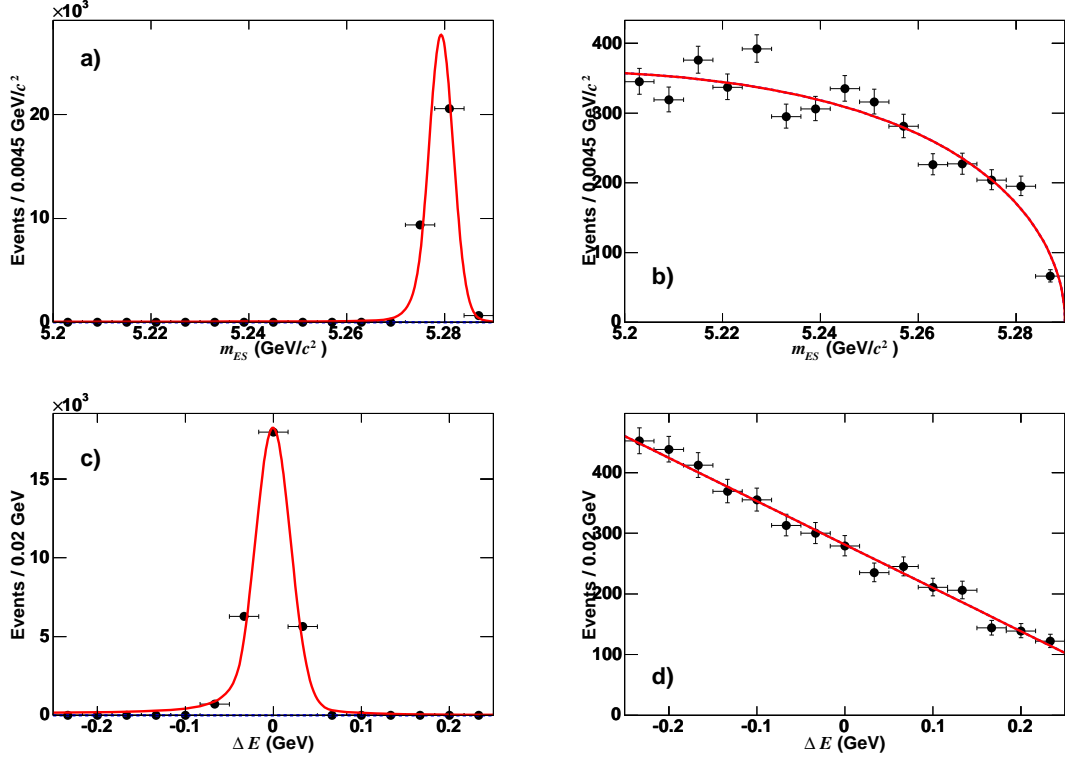
Figure 32: Plots of (a) m_{ES} in simulated signal, (b) m_{ES} in simulated background, (c) ΔE in simulated signal, and (d) ΔE in simulated background for the $B^{\pm} \to K^{\pm} \mu^+ \mu^-$ mode. The solid lines are fits with the parameterization described in Section 5. The normalization is arbitrary.

$$\Delta E \equiv E_B^* - \frac{\sqrt{s}}{2}.$$

Correctly reconstructed signal events will peak at the B mass in m_{ES}, and at zero in ΔE. The signal resolution for the final states considered in this analysis is approximately $\sigma_{m_{ES}} \approx 2.5$ MeV/c^2 in m_{ES} and $\sigma_{\Delta E} \approx 20$ MeV in ΔE. Combinatorial backgrounds follow a distribution which is linear in ΔE, and described by an empirically determined ARGUS threshold function [61] in the m_{ES} variable. The m_{ES} and ΔE distributions of simulated signal and combinatorial background events are shown in Figure 32.

We define three regions in the m_{ES}-ΔE plane that are used in the analysis:

- The fit region contains all the events which are used to extract the final signal yields. The fit region remains blinded until the event selection has been finalized. The boundaries of the fit region are: $5.2 < m_{ES} < 5.3$ GeV/c^2, $-0.25 < \Delta E < 0.25$ GeV, and $0.7 < m_{K^*} < 1.1$ GeV/c^2 for the modes with a K^*.

- The sideband region contains primarily combinatorial events, and is useful for studying such backgrounds. The sideband excludes the fit region but contains events which otherwise satisfy $m_{ES} > 5.0$ GeV/c^2 and $-0.5 < \Delta E < 0.5$ GeV.

- The signal region is a subset of the fit region that contains $85 - 95\%$ of the signal and has the boundaries: $5.2724 < m_{ES} < 5.2856$ GeV/c^2, $-0.07 < \Delta E < 0.05$ GeV for modes with muons, $-0.11 < \Delta E < 0.05$ GeV for modes with electrons, and $0.817 < m_{K^*} < 0.967$ GeV/c^2 for the modes with a K^*.

4.3 Background suppression

Backgrounds relevant for this analysis can be divided into two general categories:

- Combinatorial backgrounds are those which do not peak in the m_{ES} and ΔE variables.

- Peaking backgrounds are those which peak in one or more of the fit variables.

Combinatorial background is reduced by selecting on multivariate combinations of several variables. Peaking backgrounds are vetoed where possible, or measured using data control samples or Monte Carlo simulation when a veto is not possible.

4.3.1 Combinatorial background

Combinatorial background can arise from either continuum events, in which a light (u,d,s,c) quark pair is produced, or from $B\overline{B}$ events in which the decay products of two
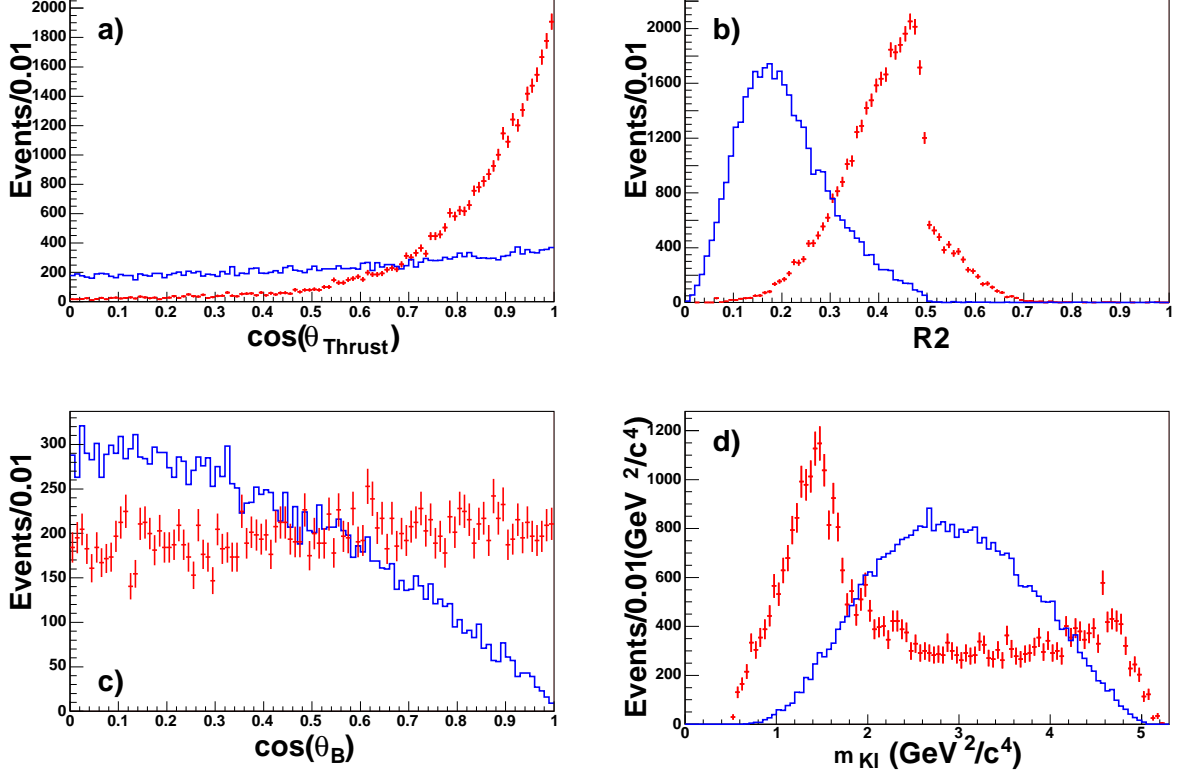
Figure 33: Plots of simulated signal (histograms) and continuum background (points with error bars) for variables used in the Fisher discriminant in the $B^{\pm} \to K^{\pm}e^{+}e^{-}$ mode: (a) $\cos(\theta_{thrust})$, (b) $R2$, (c) $\cos(\theta_B)$, and (d) m_{Kl}. For m_{Kl} only the $c\bar{c}$ component of the background is shown. The normalization is arbitrary.

B mesons are mis-reconstructed as a $B \to K^{(*)}l^{+}l^{-}$ candidate.

While true B decays tend to be spherical, continuum events are characteristically more jet-like. Therefore, these backgrounds can be discriminated against with the following event shape variables

- The angle between the event's thrust axis and the z axis in the CM frame $\cos(\theta_{thrust})$.

- The ratio $R2$ of the second to the zeroth Fox-Wolfram moment [58].

- The angle of the B candidate relative to the z axis in the CM frame $\cos(\theta_B)$.
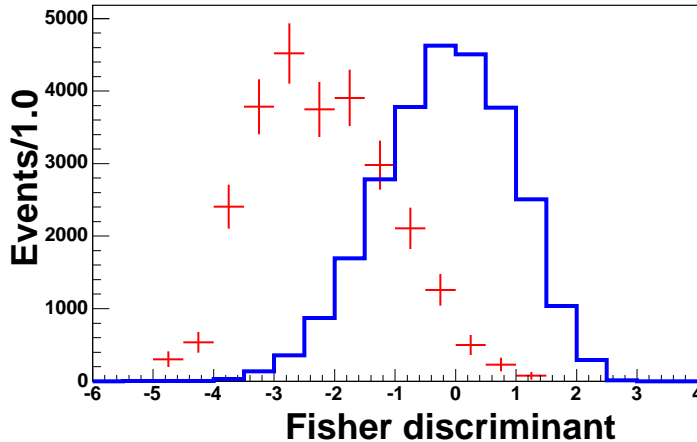
Figure 34: Fisher discriminant output for simulated signal (histograms) and continuum background (points with error bars) in the $B^{\pm} \rightarrow K^{\pm}e^+e^-$ mode. The normalization is arbitrary.

In addition, a significant fraction of the background from $c\bar{c}$ events is due to semileptonic charm decays such as $D \rightarrow Kl\nu$. In these decays, the invariant mass of the kaon and lepton of opposite charge must peak below the D meson mass, while signal events are broadly distributed at higher masses. Therefore the variable m_{Kl} is effective at rejecting $c\bar{c}$ events, though less so for uds events. The distributions of these variables for signal and background are shown in Figure 33.

These four variables are combined into a linear Fisher discriminant [59], which is trained using samples of off-resonance data and simulated signal. The training is performed separately for each of the eight signal modes. The output of the Fisher discriminant is shown in Figure 34, displaying a clear separation between signal and continuum background.

$B\overline{B}$ background overwhelmingly arises from events with two semileptonic decays of the form $B \rightarrow Xl\nu$, where X is most commonly a charm meson. These backgrounds are suppressed via a likelihood ratio composed of the following variables:

- The total missing energy in the event E_{miss}

- The vertex probability of the lepton pair $P_{vtx}(l^+l^-)$

- The vertex probability of the B candidate $P_{vtx}(B)$

- The angle of the B candidate relative to the z axis in the CM frame $\cos(\theta_B)$

The E_{miss} variable is particularly effective at rejecting events with two semileptonic decays, as these events will produce two undetected neutrinos. Loose requirements on the input variables are imposed to ensure that: $P_{vtx}(\ell^+\ell^-) > 10^{-10}$, $P_{vtx}(B) > 10^{-10}$, and $-4.0 < E_{miss} < 6.0$ GeV. The PDFs that comprise the likelihood are then derived by performing fits to these distributions for signal and generic $B\overline{B}$ Monte Carlo events as follows:

- The E_{miss} distributions are fit using a bivariate Gaussian.

- The log of the $P_{vtx}(\ell^+\ell^-)$ distribution is fit using a first-order polynomial plus an exponential.

- The log of the $P_{vtx}(B)$ distribution is fit using a first-order polynomial plus an exponential.

- The $\cos(\theta_B)$ distribution is fit using a second-order polynomial.

The distributions of these variables are shown in Figure 35. The likelihood is then calculated as:

$$\mathcal{L}_i = P(E_{miss})_i \times P(vtx(l^+l^-))_i \times P(vtx(B))_i \times P(\cos(\theta_B))_i$$

where i is the species (signal or background). The likelihoods are parameterized separately for each of the eight signal modes. The resulting output of the likelihood ratio is shown in
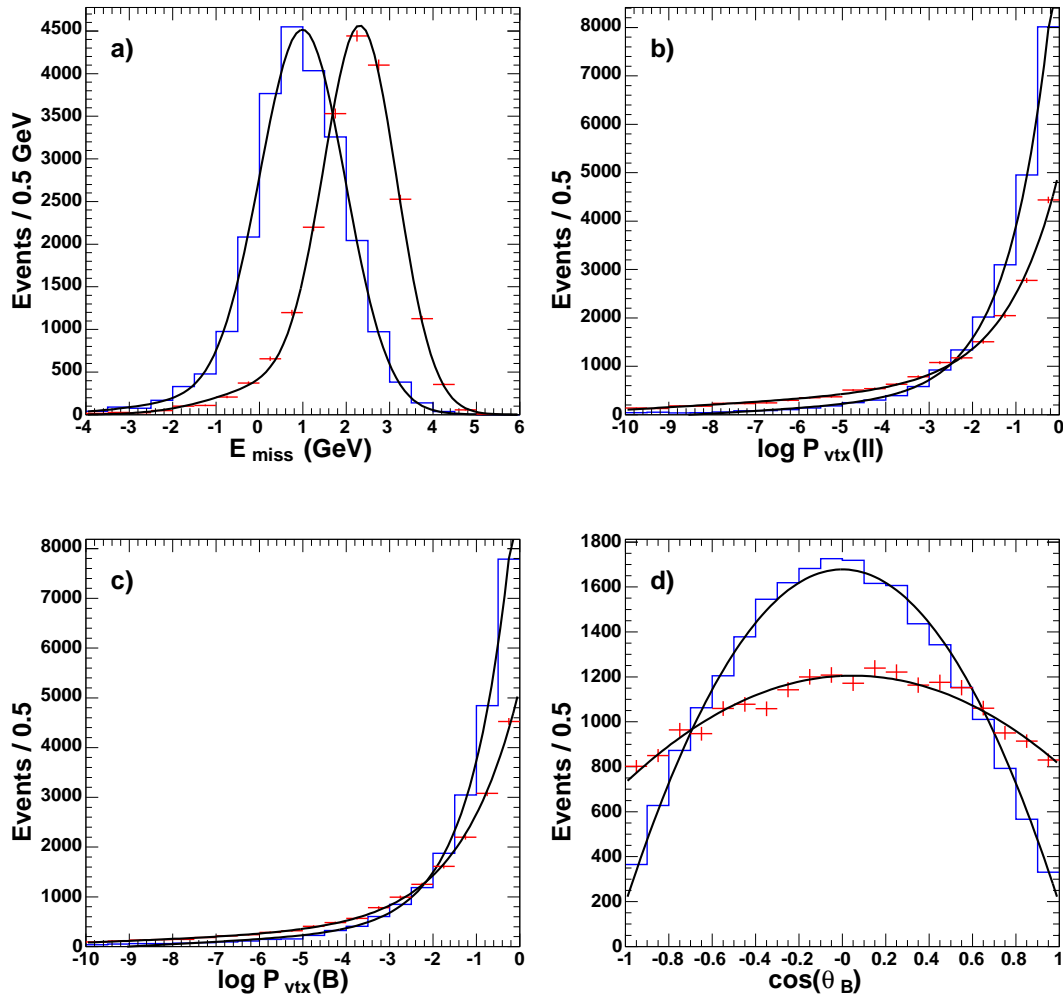
Figure 35: Plots of simulated signal (histograms) and $B\overline{B}$ background (points with error bars) for variables used in the $B\overline{B}$ likelihood in the $B^0 \rightarrow K^{*0}e^+e^-$ mode: (a) E_{miss}, (b) $\log(P_{vtx}(l^+l^-))$, (c) $\log(P_{vtx}(B))$, and (d) $\cos(\theta_B)$. The solid lines show the fits used to extract the PDFs. The normalization is arbitrary.

Figure 36; the combinatoric $B\overline{B}$ displays a clear peak at zero, while the signal peaks near a value of one.
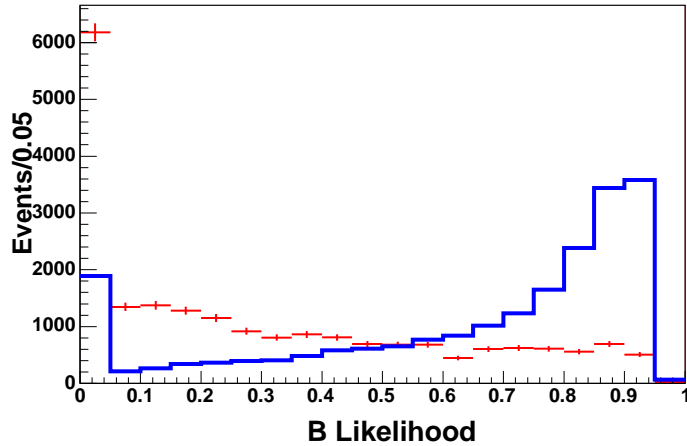
Figure 36: Likelihood output for simulated signal (histograms) and $B\overline{B}$ background (points with error bars) in the $B^0 \to K^{*0}e^+e^-$ mode. The normalization is arbitrary.

4.3.2 Optimization

We select events which pass minimal Fisher and likelihood selection criteria. We optimize this selection based on the figure of merit $S/\sqrt{S+B}$, where S is the expected signal and B the expected combinatorial background. In practice, the Fisher and likelihood are somewhat correlated. We account for this by performing a two-dimensional grid scan in these variables using simulated signal and background events, scaled to the number expected in $200 fb^{-1}$ of data. At each point we extract the expected number of signal and background events by performing a one-dimensional fit to the m_{ES} distribution. The optimization is done separately for each of the signal modes.

4.3.3 Peaking background

Backgrounds which peak in both m_{ES} and ΔE must be either vetoed or subtracted. These fall into three categories:

1. B decays to charmonium.
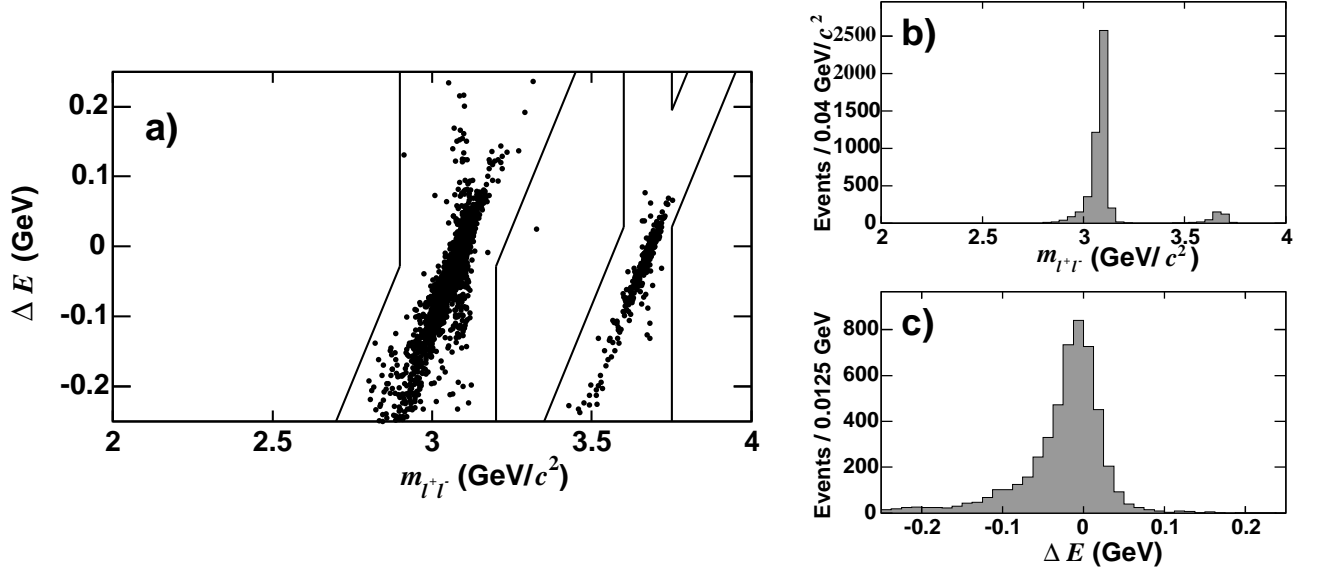
Figure 37: Charmonium veto regions in the $B \to K^+ e^+ e^-$ mode. The points are simulated $B \to J/\psi K$ and $B \to \psi(2S)K$ events with abundance equal to the number expected in $208 fb^{-1}$. The high density of points at the center of the veto region is illustrated by the projections onto $m_{l^+l^-}$ and ΔE, above and at left.

2. Hadronic B decays.

3. Rare B decays with $e^+ e^-$ pairs.

The largest source of peaking backgrounds are decays of the type $B \to J/\psi K^{(*)}$ and $B \to \psi(2S)K^{(*)}$, where the J/ψ or $\psi(2S)$ decays into a l^+l^- pair. These events are removed by means of a correlated veto in the $m_{\ell\ell}$-ΔE plane (Figure 37). Prior to the veto, charmonium decays contribute between 700 and 5000 events per decay mode. After the veto is applied, the residual charmonium background is estimated from Monte Carlo samples to be less than 1 event per decay mode.

The second largest source of peaking backgrounds are hadronic B decays to final states of the type $K^{(*)}h^+h^-$, where the hadron "h" can be either a charged kaon or pion. These will fake a signal candidate when both hadrons of the h^+h^- pair are misidentified as muons. In

practice, these are dominated by the $K\pi^+\pi^-$ final state, as events with a kaon misidentified as a muon will tend peak away from 0 in ΔE. These backgrounds are negligible in the electron channels, due to the much smaller probability of misidentifying hadrons as electrons.

The majority of these events originate from the decay $B \to D\pi$, with $D \to K^{(*)}\pi$. These can be vetoed by assuming the μ is a π, and removing events where the resulting $K^{(*)}\mu$ invariant mass lies between 1.84 and 1.90 GeV/c^2.

The remainder of the hadronic peaking background comes from three-body B decays such as $B \to K^{(*)}\pi^+\pi^-$, $B \to K^{(*)}K^+\pi^-$, and $B \to K^{(*)}K^+K^-$. As the branching fractions and kinematic distributions for these decays are not precisely measured, we estimate this background by constructing a data control sample of hadronic B decays. The control sample consists of events reconstructed as $B \to K^{(*)}\mu h$, where the h is either a K or a π. The μ is required to pass a loose muon identification selection in order to reduce the rate of this selection. The h is required to fail electron or muon identification, resulting in a sample which is composed of hadronic B decays. As the sample is reconstructed inclusively in data, we do not distinguish between purely non-resonant decays and decays that proceed through an intermediate resonace, such as $B \to K\rho^0$, with $\rho^0 \to \pi^+\pi^-$. The inclusive sample is then weighted by the probability to misidentify hadrons as muons. The peaking component of this weighted sample is then extracted from a fit to the m_{ES} distribution, resulting in an estimated hadronic peaking component of $0.35 - 1.60$ events per decay channel. The precision of this estimate is limited by the statistics of the control sample; the fractional uncertainties on the peaking background are $21\% - 68\%$ per decay mode.

Despite the photon conversion veto which removes events with low e^+e^- invariant mass, a small residual background remains from the rare decay $B \to K^*\gamma$, where the γ converts to an e^+e^- pair in the detector. In addition, the decays $B \to K^{(*)}\pi^0$ and $B \to K^{(*)}\eta$ will mimic the $B \to K^{(*)}e^+e^-$ signal when the π^0 or η undergoes a Dalitz decay to $e^+e^-\gamma$ and the

Mode	All q^2	$0.1 < q^2 < 8.41$ (GeV^2/c^4)	$q^2 > 10.24$ (GeV^2/c^4)
$B^\pm \to K^\pm e^+ e^-$	0.7 ± 0.2	0.6 ± 0.2	0.1 ± 0.1
$B^\pm \to K^\pm \mu^+ \mu^-$	2.3 ± 0.5	1.4 ± 0.4	0.9 ± 0.1
$B^0 \to K_s^0 e^+ e^-$	0.01 ± 0.01	0.01 ± 0.01	0.0
$B^0 \to K_s^0 \mu^+ \mu^-$	0.4 ± 0.1	0.3 ± 0.1	0.1 ± 0.04
$B^0 \to K^{*0} e^+ e^-$	3.0 ± 0.6	1.0 ± 0.5	0.6 ± 0.2
$B^0 \to K^{*0} \mu^+ \mu^-$	1.4 ± 0.8	0.5 ± 0.3	0.2 ± 0.1
$B^\pm \to K^{*\pm} e^+ e^-$	0.9 ± 0.2	0.2 ± 0.2	0.2 ± 0.1
$B^\pm \to K^{*\pm} \mu^+ \mu^-$	0.6 ± 0.3	0.2 ± 0.1	0.2 ± 0.1

Table 8: Mean expected peaking backgrounds in $208\,\text{fb}^{-1}$, for the individual $K^{(*)}\ell^+\ell^-$ decay modes after applying all selection requirements.

resulting γ is found by the brem-recovery algorithm. The backgrounds from these processes are estimated from Monte Carlo samples. As these decays do not produce $\mu^+\mu^-$ pairs, they are only relevant for the $B \to K^{(*)}e^+e^-$ modes.

Table 8 lists the total number of peaking background events expected from all sources in each decay mode. The muon backgrounds are dominated by the hadronic decays; the largest background in the $B \to K^* e^+ e^-$ is due to photon conversions from $K^*\gamma$, while the charmonium and Dalitz decays are dominant in $B \to K e^+ e^-$.

As described in Section 2.2.4, the partial branching fractions, K^* polarization, and A_{FB} are measured in the ranges $0.1 < q^2 < 8.41$ and $q^2 > 10.24$ (Figure 38). We therefore recompute the expected peaking backgrounds in each region for these measurements. As a function of q^2, all of the $K^*\gamma$ and $K^{(*)}\pi^0$ background is restricted to the region $0.0 < q^2 < 0.1$. The $K^{(*)}\eta$ and J/ψ events that escape the veto are limited to the region $0.1 < q^2 < 8.41$, while the non-vetoed $\psi(2S)$ background events fill the region $q^2 > 10.24$. The hadronic backgrounds occupy both the $0.1 < q^2 < 8.41$ and $q^2 > 10.24$ regions. The total peaking
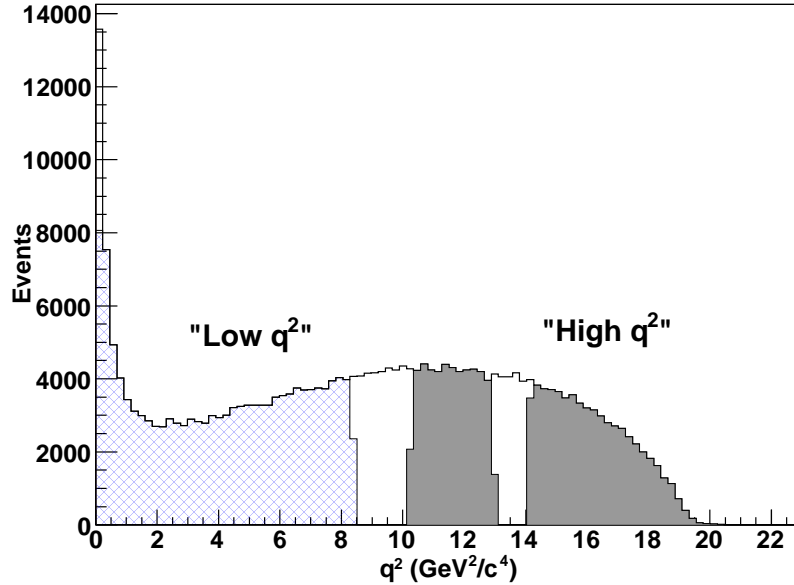
Figure 38: q^2 distribution in $B \to K^{*0}\mu^+\mu^-$ simulation. The "high" ($q^2 > 10.24$) and "low" ($0.1 < q^2 < 8.41$) regions are shown by the dark and light shaded areas, respectively.

backgrounds in each region are shown in Table 8.

4.4   Efficiencies

Once all selection criteria have been established, we evaluate the selection efficiency using simulated signal events. For the measurement of the partial branching fractions, we also evaluate efficiencies in each bin of q^2. The resulting efficiencies are shown in Table 9

There is a significant variation among the modes, which is due to a number of effects. The efficiencies in the electron channels are systematically higher due to the lower particle identification efficiency for muons. The $B \to K^*\ell^+\ell^-$ modes generally have lower efficiencies than the $B \to K\ell^+\ell^-$ modes; this is a result of the higher combinatorial background in the K^* modes, which leads to tighter optimal Fisher and likelihood selections. The efficiencies in $B^0 \to K_S^0\ell^+\ell^-$ are high at low q^2, but are degraded at high q^2 due to the low efficiency for

Mode	All q^2 (%)	$0.1 < q^2 < 8.41$ (%)	$q^2 > 10.24$ (%)
$B^\pm \to K^\pm e^+ e^-$	26.6	32.2	25.9
$B^\pm \to K^\pm \mu^+ \mu^-$	15.4	16.4	16.9
$B^0 \to K_S^0 e^+ e^-$	22.8	32.1	17.9
$B^0 \to K_S^0 \mu^+ \mu^-$	13.6	17.2	11.9
$B^0 \to K^{*0} e^+ e^-$	18.6	25.3	20.5
$B^0 \to K^{*0} \mu^+ \mu^-$	11.9	11.8	14.8
$B^\pm \to K^{*\pm} e^+ e^-$	15.7	22.7	17.2
$B^\pm \to K^{*\pm} \mu^+ \mu^-$	9.3	10.1	10.2

Table 9: Signal efficiency in % for individual $K^{(*)}\ell^+\ell^-$ decay modes after applying all selection requirements. The uncertainty due to simulation statistics is $\sim 0.1\%$.

finding very low momentum K_S^0 candidates. Here the "All q^2" efficiency is defined relative to the full q^2 range, including the photon pole and J/ψ resonance regions. The all q^2 efficiency can therefore be smaller than the efficiency in either of the two q^2 bins, as they do not include the pole or J/ψ regions.

5 Fit Procedure

To extract the signal yield, we perform an unbinned extended maximum likelihood fit. In the $B \to Kl^+l^-$ modes, we use a two-dimensional fit in the variables m_{ES} and ΔE. In the $B \to K^*l^+l^-$ modes, the mass of the K^* candidate is added as a third fit variable. The fit includes the following components:

- The signal is parameterized as a Crystal Ball function [60] in both m_{ES} and ΔE, and a relativistic Breit-Wigner line shape for the K^* mass. The Crystal Ball shape is an empirically determined function, which has a Gaussian core with a power law tail. The functional form is:

$$f(x) \propto \begin{cases} \exp(-\frac{(x-\overline{x})^2}{2\sigma^2}) & ; \quad (x-\overline{x})/\sigma > \alpha \\ A \times (B - \frac{x-\overline{x}}{\sigma})^{-n} & ; \quad (x-\overline{x})/\sigma \le \alpha \end{cases},$$

  where $A \equiv (\frac{n}{|\alpha|})^n \times \exp(-|\alpha|^2/2)$ and $B \equiv \frac{n}{|\alpha|} - |\alpha|$. The variables $\overline{x}$ and σ are the Gaussian peak and width, and α and n are the point at which the function transitions to the power function and the exponent of the power function, respectively. The tail parameters are fixed to the values obtained from fits to simulated signal events. The mean and width of m_{ES} and ΔE and the width of the K^* mass distribution are fixed to the values obtained from fits to the $J/\psi K^{(*)}$ data samples.

- The combinatorial background is modeled as an ARGUS threshold function [61] in m_{ES}, a linear term in ΔE, and a quadratic term in the K^* mass. The empirically derived form of the ARGUS function is $f(x) \propto x\sqrt{1-x^2} \exp\left[-\zeta(1-x^2)\right]$, where ζ is a fit parameter and $x = m_{ES}/E_b^*$.

- The feeddown background arises from events in which a non-resonant $B \to X_s l^+l^-$ decay is mis-reconstructed as signal, or a $B \to K^*l^+l^-$ decay is mis-reconstructed as

$B \to Kl^+l^-$. The feeddown background is parameterized as a Gaussian peak in m_{ES} and ΔE, with a floating mean and normalization. In general, the feeddown background will peak near the B mass in m_{ES}, but will be shifted into the lower ΔE sideband, due to the lost pion.

- The peaking background arising from the sources discussed in Section 4.3.3 uses the identical shape as the signal component, with the yield fixed in the fit.

In addition, the $B \to K^*l^+l^-$ modes include two additional components:

- The feedup background is the counterpart to the feeddown background, in which a $B \to Kl^+l^-$ decay is mis-reconstructed as $B \to K^*l^+l^-$. The feedup background is parameterized as a Gaussian peak in m_{ES} and ΔE, with a floating mean and normalization. The feedup background will peak near the B mass in m_{ES}, but will be shifted into the upper ΔE sideband, due to the added pion.

- A combinatorial K^* background is described by an Argus function in m_{ES}, a linear term in ΔE, and a relativistic Breit-Wigner shape in the K^* mass. The combinatorial K^* background comes from events which have a real K^* combined with two random leptons. Thus they will peak in the K^* mass, but not in m_{ES} or ΔE.

5.1   Angular fits

To extract F_L and A_{FB}, we add the angular distributions $\cos\theta_K$ or $\cos\theta^*$ as an additional fit dimension. As described in Section 2.2.4, $\cos\theta^*$ is defined as the angle between the ℓ^- (ℓ^+) and the B ($\overline{B}$), measured in the dilepton rest frame. The kaon decay angle $\cos\theta_K$ is similarly defined as the angle between the kaon and the B, measured in the K^* rest frame. The signal shape in $\cos\theta_K$ is described by an underlying differential distribution which depends on the longitudinal polarization F_L as:

$$\frac{1}{\Gamma}\frac{d\Gamma}{d\cos\theta_K}\bigg|_{K^*\ell^+\ell^-} = \frac{3}{2}F_L\cos^2\theta_K \quad +\tfrac{3}{4}(1-F_L)(1-\cos^2\theta_K). \tag{11}$$

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

$$\frac{1}{\Gamma}\frac{d\Gamma}{d\cos\theta^*}\bigg|_{K^*\ell^+\ell^-} = \frac{3}{4}F_L(1-\cos^2\theta^*) \quad +\tfrac{3}{8}(1-F_L)(1+\cos^2\theta^*) \quad +$$
$$A_{FB}\cos\theta^*. \tag{12}$$

In the $B \to K\ell^+\ell^-$ mode, the most general form for the angular distribution is:

$$\frac{1}{\Gamma}\frac{d\Gamma}{d\cos\theta^*}\bigg|_{K\ell^+\ell^-} = \frac{3}{4}(1-F_S)(1-\cos^2\theta^*) + \frac{1}{2}F_S + A_{FB}\cos\theta^*. \tag{13}$$

where F_S is the scalar contribution. As discussed in Section 2.2.4, the scalar component is expected to be small even in the presence of new physics. In the limit of zero scalar contribution the distribution reduces to:

$$\frac{1}{\Gamma}\frac{d\Gamma}{d\cos\theta^*}\bigg|_{K\ell^+\ell^-} = \frac{3}{4}(1-\cos^2\theta^*) + A_{FB}\cos\theta^*. \tag{14}$$

The results of fitting generated signal events to these functions are displayed in Figure 39.

The true angular distributions will be modified by detector acceptance and efficiency effects. To account for this, we define a signal shape PDFs which are the product of the true angular distribution with non-parametric histogram PDFs describing the efficiency as a function of $\cos\theta_K$ or $\cos\theta^*$. These are derived separately for each decay channel using signal Monte Carlo simulations. The effects of this efficiency correction are most pronounced in the
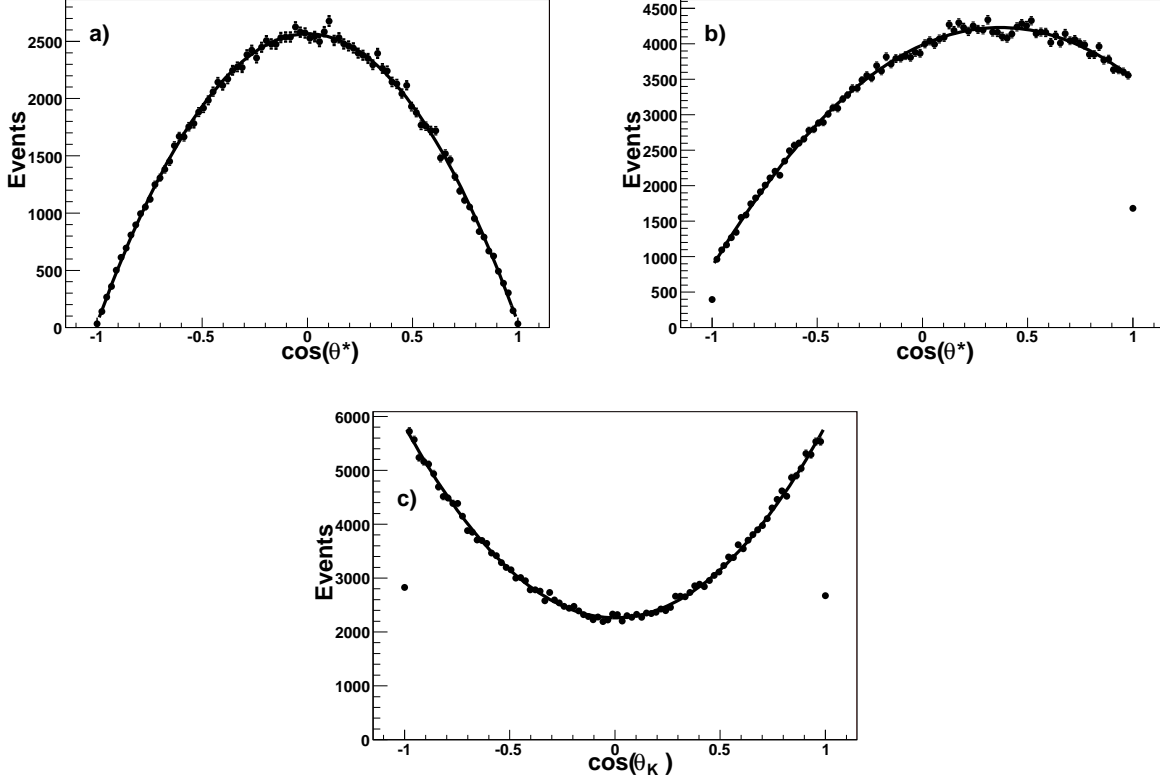
Figure 39: Angular fits to generated signal distributions. The distributions of (a) $\cos\theta^*$ in $B^\pm \to K^\pm \ell^+\ell^-$, (b) $\cos\theta^*$ in $B \to K^*\ell^+\ell^-$, and (c) $\cos\theta_K$ in $B \to K^*\ell^+\ell^-$ are shown. The points are generated according to the Standard Model Wilson coefficients and the form factor model of [17, 18]; the lines are fits to the generated events using the functional forms described in the text. The normalization is arbitrary.

low q^2 region near the extrema of the $\cos\theta^*$ distribution. In low q^2 events which are highly forward-backward asymmetric, there must be one very low momentum lepton. The minimum momentum requirement for identification of the leptons therefore reduces the acceptance for these events; the effect on the acceptance is more severe for the muon channels than the electron channels due to the higher momentum threshold for identifying muons. The variation of the efficiency with $\cos\theta^*$ is illustrated in Figure 40 for the $B^0 \to K^{*0}\ell^+\ell^-$ mode.

The combinatorial backgrounds are dominated by events with two semileptonic B decays,
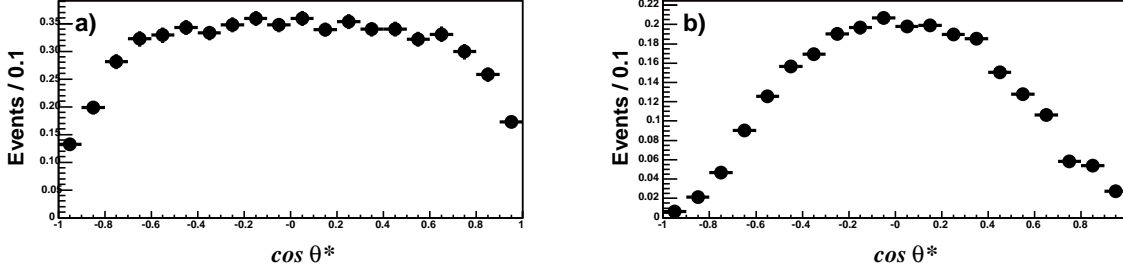
Figure 40: Signal efficiency as a function of $\cos\theta^*$ at low q^2 in the (a) $B^0 \to K^{*0}e^+e^-$ and (b) $B^0 \to K^{*0}\mu^+\mu^-$ channels.

which are in general asymmetric in $\cos\theta^*$. There are additional contributions from continuum backgrounds, "cascade" backgrounds such as $B \to D\ell\nu$ followed by $D \to K\ell\nu$, and events where one of the lepton candidates is a misidentified hadron. Each of these backgrounds has non-trivial angular distributions in $\cos\theta^*$, resulting in combinatoric background shapes which are highly forward-backward asymmetric(Figure 41). Rather than describing these backgrounds with an arbitrary continuous higher-order polynomial, we model them using a histogram PDF drawn from the $\cos\theta^*$ distribution in the m_{ES} and ΔE data sidebands.

As the $\cos\theta^*$ fit also has a weak dependence on F_L, we perform the angular fit in the $B \to K^*\ell^+\ell^-$ mode in two stages:

1. Perform a four-dimensional fit with $\cos\theta_K$ to extract the value of F_L.

2. Perform a four-dimensional fit with $\cos\theta^*$ to extract A_{FB}, where F_L is fixed to the measured value.

The measured A_{FB} has an additional systematic uncertainty due to the fixed value of F_L. In the $B^\pm \to K^\pm\ell^+\ell^-$ channel only a single four-dimensional fit with $\cos\theta^*$ is required to extract A_{FB}.
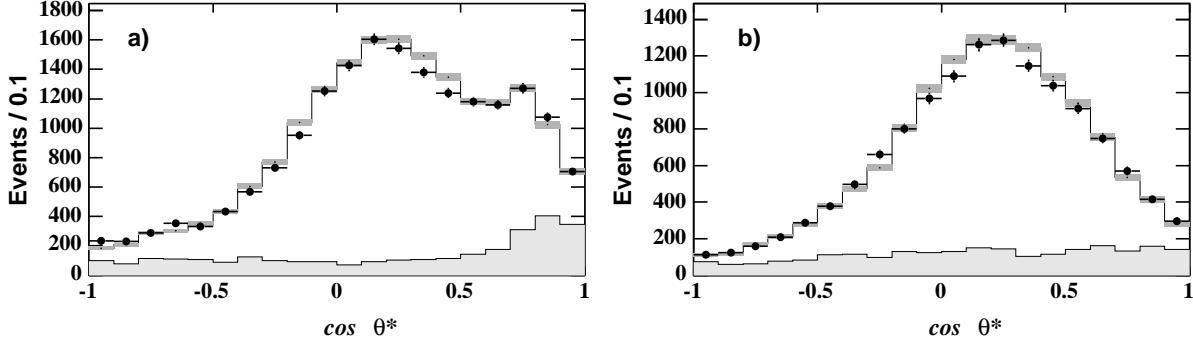
Figure 41: $\cos\theta^*$ combinatorial background shape in the (a) $B^0 \to K^{*0}e^+e^-$ and (b) $B^0 \to K^{*0}\mu^+\mu^-$ channels. The points are data from the m_{ES} and ΔE sidebands. The dark grey bands represent the total Monte Carlo prediction from all sources, where the width of the bands shows the uncertainty due to MC statistics. The light grey histogram represents the fraction of the MC background arising from continuum events.

5.2   Tests of fits in data control samples

Prior to unblinding and fitting the data, we test the performance of the fit technique using the high-statistics data control samples of $B \to J/\psi K^{(*)}$ and $B \to \psi(2S)K^{(*)}$ events.

5.2.1 J/ψ yield fits

We select $B \to J/\psi K^{(*)}$ events by reversing the charmonium veto described in Section 4.3.3. All other selection requirements are identical to those used for $B \to K^{(*)}\ell^+\ell^-$ signal candidates. We first fit each channel separately. We then perform fits with charged and neutral modes combined, with electrons and muons combine, and finally with all modes combined. When combining modes, we apply the world average measurements [1] as constraints on the ratio of charged to neutral B lifetimes, and on the ratio of muon to electron branching fractions.

The resulting projections of the branching fraction fits are shown for one decay mode in
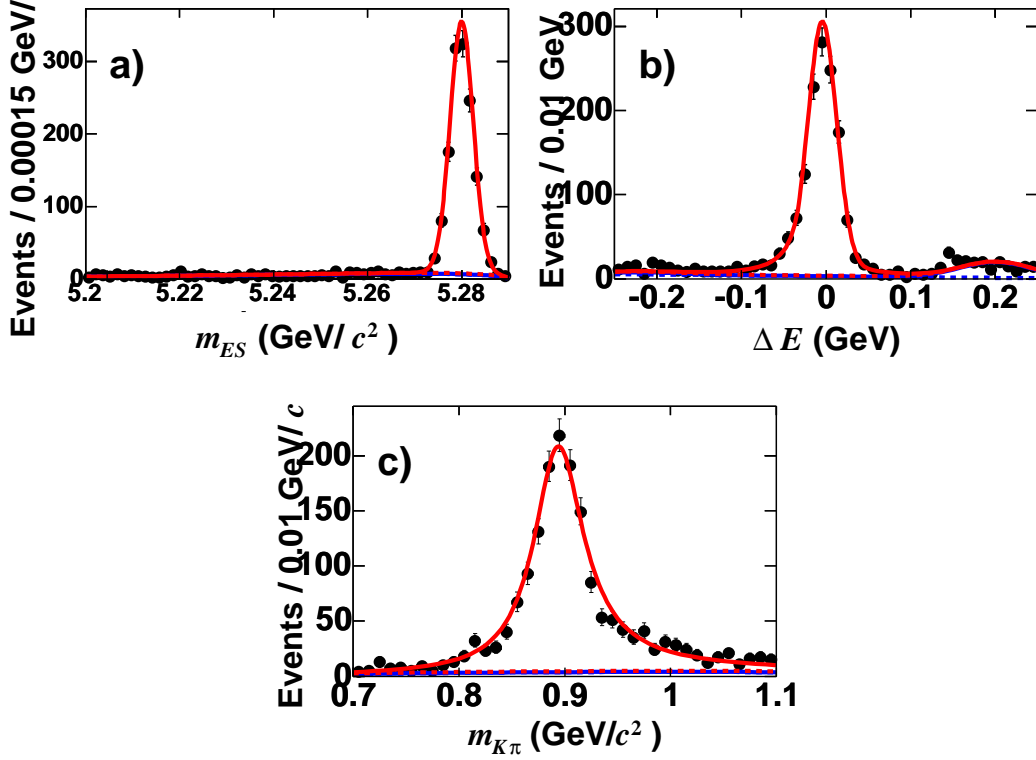
Figure 42: Fits to the $J/\psi(\to \mu^+\mu^-)K^{*0}$ data control sample. The fit projections onto (a) m_{ES}, (b) ΔE, and (c) $m_{K\pi}$ are shown for the background component (dashed lines) and the total fit (solid lines).

Figure 42. The results of all such fits to the full dataset are shown in Tables 10 and 11. We find good agreement with the world average branching fractions in all of the decay modes.

5.2.2 J/ψ Angular fits

We also test the angular fits to $\cos\theta^*$ and $\cos\theta_K$ using the charmonium samples. We derive background angular shapes from the m_{ES} and ΔE sidebands, and derive efficiency and acceptance corrections in the same manner as is done for signal. As in the branching fraction fits, we perform fits to extract A_{FB} and F_L in the individual modes, as well as in the combined modes. The fit projections for one decay mode are shown in figure 43. The
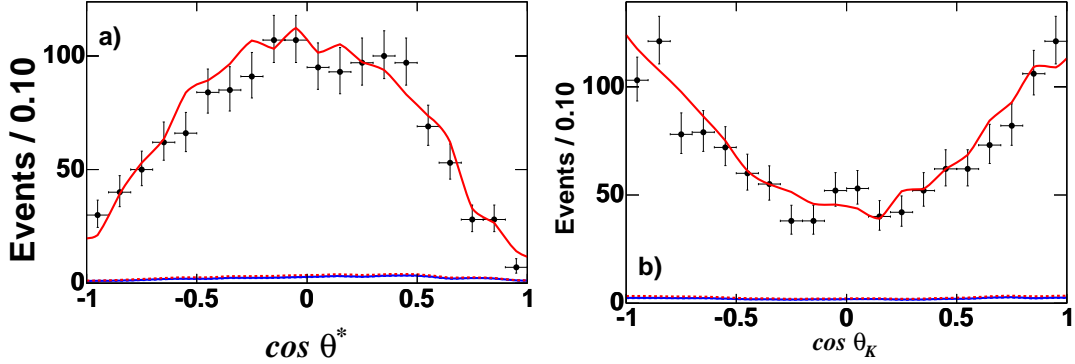
Figure 43: Fits to the angular distributions of the $J/\psi(\to \mu^+\mu^-)K^{*0}$ data control sample. The fit projections onto (a) $\cos\theta^*$ and (b) $\cos\theta_K$, are shown for the background component (dashed lines) and the total fit (solid lines). The non-smoothness results from the signal efficiency corrections and combinatoric background, which are modelled as histogram PDFs.

Mode	A_{CP} (%)	A_{FB} (%)	F_S (%)	Yield	Eff. (%)	$\mathcal{B}/10^{-6}$	PDG $\mathcal{B}/10^{-6}$
$B^\pm \to K^\pm e^+e^-$	-0.5 ± 1.5	-1.0 ± 1.0	3.9 ± 2.7	4746 ± 72	34.2	1019 ± 16	1022 ± 35
$B^\pm \to K^\pm \mu^+\mu^-$	0.4 ± 1.8	-0.3 ± 1.2	2.7 ± 2.9	3061 ± 57	21.8	1040 ± 19	1022 ± 35
$B^\pm \to K^\pm \ell^+\ell^-$	-0.1 ± 1.2	-0.7 ± 0.8	3.3 ± 2.0			1027 ± 12	1022 ± 35
$B^0 \to K_S^0 e^+e^-$	—	—	—	1286 ± 38	31.7	869 ± 26	872 ± 33
$B^0 \to K_S^0 \mu^+\mu^-$	—	—	—	846 ± 30	20.4	895 ± 32	872 ± 33
$B^0 \to K_S^0 \ell^+\ell^-$						879 ± 20	850 ± 50

Table 10: $J/\psi\ K$ yields, branching fractions, and asymmetries from fits to *BABAR* data.

resulting values of A_{FB} and F_L are listed in Tables 10 and 11. For the combined modes, we find values of A_{FB} which are consistent with zero to within 0.02. We find values of F_L which are consistent with previous measurements to within 0.05 [62]. If the scalar fraction F_S is allowed to float in the fits to $B^\pm \to J/\psi K^\pm$, a value consistent with zero to within 0.03 is obtained. These values are taken as a systematic uncertainty associated with modeling of detector efficiency and acceptance effects.

5.2.3 $\psi(2S)$ fits

The vetoed events in the $B \to \psi(2S)K^{(*)}$ sample provide an additional control sample of signal-like events. Although containing fewer events than the J/ψ sample, the $\psi(2S)$ sample

Mode	A_{CP} (%)	A_{FB} (%)	F_L (%)	Yield	Eff. (%)	$\mathcal{B}/10^{-6}$	PDG $\mathcal{B}/10^{-6}$
$B^0 \to K^{*0}e^+e^-$	2.0 ± 2.0	-0.1 ± 1.9	53.7 ± 1.7	3410 ± 71	28.1	1339 ± 28	1330 ± 60
$B^0 \to K^{*0}\mu^+\mu^-$	3.4 ± 2.5	-5.0 ± 2.8	51.9 ± 2.2	2095 ± 53	18.0	1294 ± 33	1330 ± 60
$B^0 \to K^{*0}\ell^+\ell^-$	2.5 ± 1.5	-1.6 ± 1.6	53.0 ± 1.3			1320 ± 21	1330 ± 60
$B^\pm \to K^{*\pm}e^+e^-$	1.5 ± 3.9	-0.8 ± 3.7	52.6 ± 3.3	1063 ± 44	24.8	1376 ± 57	1410 ± 80
$B^\pm \to K^{*\pm}\mu^+\mu^-$	-0.3 ± 4.7	-7.0 ± 5.3	47.3 ± 4.3	650 ± 32	13.7	1541 ± 76	1410 ± 80
$B^\pm \to K^{*\pm}\ell^+\ell^-$	0.7 ± 3.0	-2.9 ± 3.0	50.6 ± 2.6			1435 ± 46	1410 ± 80
$B \to K^*\ell^+\ell^-$	2.1 ± 1.4	-1.9 ± 1.4	52.5 ± 1.2				

Table 11: J/ψ K^* yields, branching fractions, and asymmetries from fits to *BABAR* data.

allows a cross-check of the fit performance in a higher region of q^2. We find the signal yield in all eight decay modes is consistent with previous measurements of the branching fraction [1].

5.3 Tests of fits in simulation

To estimate the expected precision for A_{FB}, and to test for large fit biases, we perform a series of embedded toy experiments. For each experiment, we generate background shapes from the PDF resulting from a fit to the fully simulated generic background samples. From the large sample of simulated signal, we embed events with the mean number of expected signal events in a sample of $200 fb^{-1}$ into the generated background samples. Between 200 and 500 experiments are performed for each bin of q^2.

Assuming the q^2 dependence of the form factor model of [17, 18], we find the expected precision for A_{FB} in each bin is $0.28 - 0.34$, with no significant bias due to the fit procedure (Figure 44). We also perform this study for several non-SM values of the asymmetry. We find that the fits that converge are unbiased, even for maximal values of the asymmetry. However, in cases where the true asymmetry is large, a significant fraction of the fits fail to converge to a minimum. In this case, only a one-sided limit on the asymmetry can be obtained.
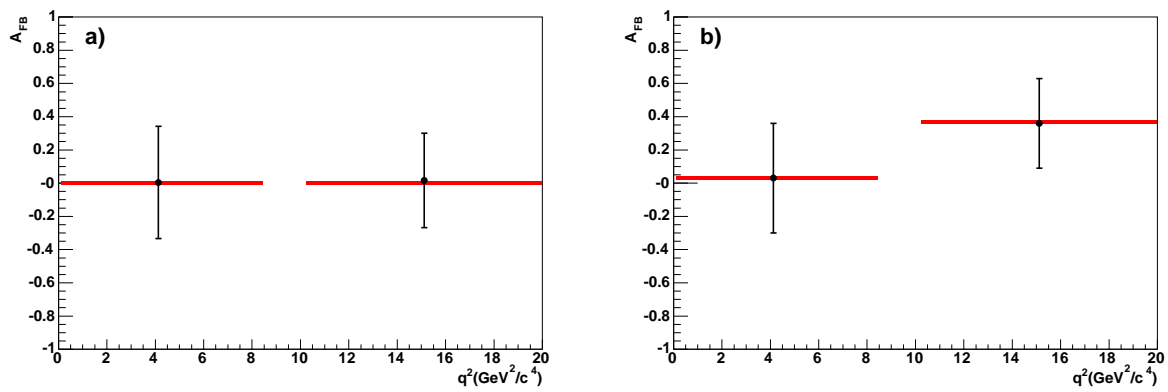
Figure 44: Expected precision on A_{FB} for (a) $B^{\pm} \to K^{\pm}\ell^+\ell^-$ and (b) $B \to K^*\ell^+\ell^-$. The points with error bars are the result of toy experiments. The lines are the SM predictions used to generate the events.

6 Systematics

This section describes the systematic uncertainties relevant to this analysis. Systematics related to the branching fractions, CP asymmetries, and angular distributions are each treated separately. In all cases, the evaluation of the systematic errors relies heavily on data control samples.

6.1 Branching fraction systematics

To determine systematic uncertainties on the total and partial branching fractions, we consider both signal efficiency uncertainties, and uncertainties related to the likelihood fit used to extract the signal yield.

6.1.1 Efficiency systematics

We use the agreement between data and Monte Carlo in the high statistics $B \to J/\psi K^{(*)}$ control sample to bound systematic uncertainties associated with the efficiency of the Fisher and likelihood background suppression selections. The level of agreement in the Fisher and likelihood distributions is illustrated in Figure 45. Systematics due to tracking efficiency and K_S reconstruction efficiency are based on standard *BABAR* studies of inclusive K_S and τ decay control samples. The uncertainty in the total number of $B\overline{B}$ events produced is taken from *BABAR* studies of the ratio of hadronic events to μ-pairs.

The model dependence uncertainties arise primarily from the variation of the signal efficiency as a function of q^2. As the various form-factor calculations produce somewhat different q^2 spectra, we generate event samples for each model to study the size of this effect. With the model of Refs. [17, 18] as a baseline, we take the largest variation produced by either of the alternate models [19, 20, 21] as the systematic error due to model dependence. In addition, for the K^* modes we consider possible effects due to the variation of efficiency

with the K^* polarization angle.

The list of all efficiency systematics considered in the individual branching fraction measurements for all q^2 are shown in Table 12. The largest sources of efficiency systematics are the model dependence and tracking efficiency. The latter is particularly significant for the high-multiplicity $B^\pm \to K^{*\pm}\ell^+\ell^-$ modes with five charged tracks in the final state.

6.1.2 Particle identification systematics

The data control samples described in Section 4.1 are used to correct for any discrepancies in the performance of the particle identification in data versus simulation. The statistics of these samples are sufficient to apply corrections as a function of momentum, polar angle, azimuthal angle, run period, and charge. PID systematics arise from possible differences between the PID performance measured in these control samples, and the PID performance on signal events. Specifically, the lepton identification is evaluated using low-multiplicity samples of $e^+e^- \to e^+e^-\gamma$ and $e^+e^- \to \mu^+\mu^-\gamma$ events. The PID could in principle be inferior in high-multiplicity B decay events, where there is greater overlap of tracks in the
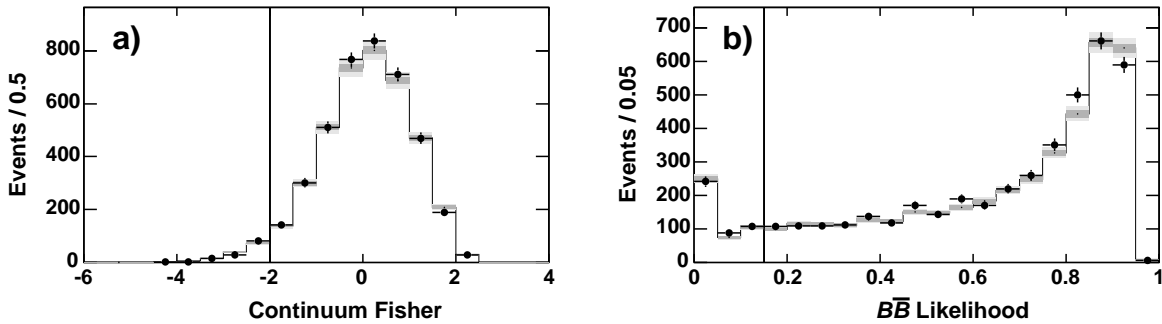
Figure 45: Fisher and likelihood distributions in the J/ψ control sample in the $B^\pm \to K^\pm e^+e^-$ channel. The points with error bars are data; the gray bands show the Monte Carlo simulation, with the width of the bands indicating the uncertainty in the MC prediction. Events to the right of the vertical line are selected.

Systematic	$K^{\pm}e^{+}e^{-}$	$K^{\pm}\mu^{+}\mu^{-}$	$K_S e^{+}e^{-}$	$K_S \mu^{+}\mu^{-}$	$K^{*0}e^{+}e^{-}$	$K^{*0}\mu^{+}\mu^{-}$	$K^{*\pm}e^{+}e^{-}$	$K^{*\pm}\mu^{+}\mu^{-}$
Trk eff. (e,μ)	±1.6	±1.6	±1.6	±1.6	±1.6	±1.6	±1.6	±1.6
Electron ID	±0.6	-	±0.6	-	±0.6	-	±0.6	-
Muon ID	-	±1.3	-	±1.3	-	±1.3	-	±1.3
Kaon ID	±0.6	±0.6	-	-	±0.6	±0.6	-	-
Pion ID	-	-	-	-	±0.2	±0.2	±0.2	±0.2
Trk eff. (K,π)	±1.4	±1.4	±2.8	±2.8	±2.8	±2.8	±4.2	±4.2
K_S eff.	-	-	±0.9	±0.9	-	-	±0.9	±0.9
$B\bar{B}$ counting	±1.1	±1.1	±1.1	±1.1	±1.1	±1.1	±1.1	±1.1
Fisher	±0.3	±0.5	±0.6	±1.1	±0.6	±1.0	±1.1	±2.2
$B\bar{B}$ likelihood	±0.6	±0.6	±0.9	±0.9	±0.9	±0.9	±1.7	±2.1
Model dep.	±1.5	±1.9	±8.3	±6.6	±1.1	±2.5	±1.3	±1.1
K^* polarization	-	-	-	-	±0.3	±1.8	±2.0	±1.6
MC statistics	±0.4	±0.5	±0.4	±0.5	±0.5	±0.6	±0.5	±0.7
Total	±3.7	±4.1	±9.6	±8.3	±4.9	±5.8	±6.8	±7.1

Table 12: The sources of efficiency related branching fraction systematic uncertainty (%) considered for the individual modes on a per-event basis.

EMC and IFR.

We evaluate these effects by reconstructing data samples of $B \to J/\psi K^{(*)}$ decays without requiring particle identification on the leptons. We then use this sample to measure the efficiency for identifying electrons and muons in data, compared to the efficiency measured in the radiative dilepton samples after appropriate weighting to match the kinematics of $J/\psi K^{(*)}$ decays. We find that the ratio of the PID efficiency for identifying two leptons in these two environments is $(98.6 \pm 0.6)\%$ for electrons, and $(96.4 \pm 1.3)\%$ for muons. Since these ratios are $\geq 2\sigma$ from 100%, we apply this difference as a correction of 1.4% per electron pair and 3.6% per muon pair. The uncertainty on this correction is used to obtain a PID systematic uncertainty of 0.6% per electron pair and 1.3% per muon pair. After applying the correction, we examine the momentum and theta spectrum in data versus simulation. We find good agreement in both electrons and muons (Figure 46).

The systematic uncertainties due to hadron identification are eevaluated in a similar manner. We reconstruct the samples of $B \to J/\psi K^{(*)}$ removing the particle identification
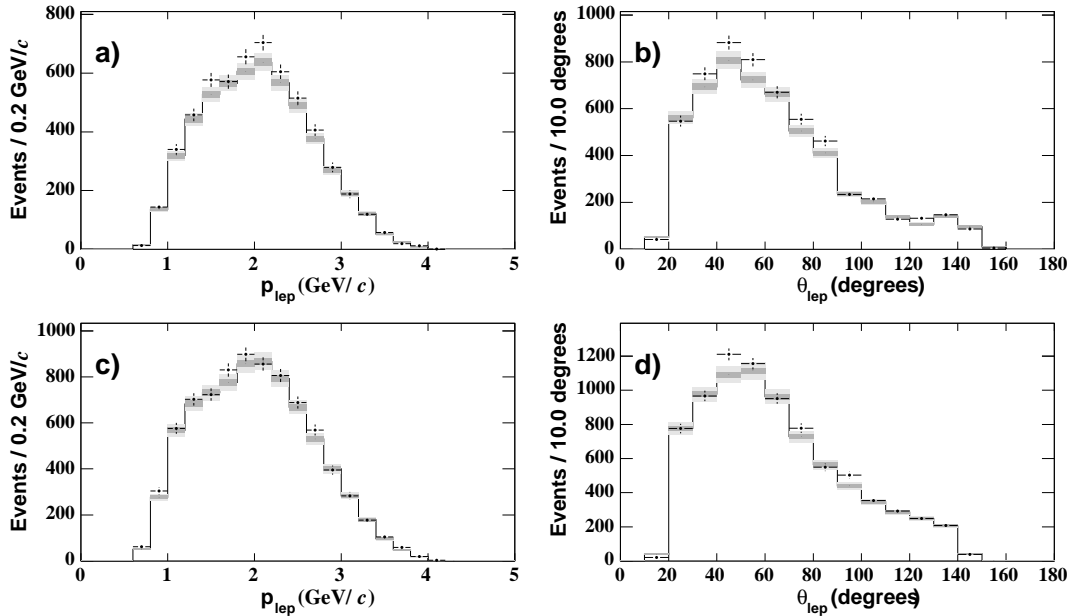
Figure 46: Kinematic distributions in leptons from the $B^{\pm} \to J/\psi K^{\pm}$ control sample: (a) muon lab momentum, (b) muon polar angle, (c) electron lab momentum, and (d) electron polar angle. The points with error bars are data; the gray bands show the Monte Carlo simulation after all corrections, with the width of the bands indicating the uncertainty in the MC prediction.

requirement first on the charged kaon and then on the charged pion. We then compare the efficiency measured in these samples to that obtained from the standard sample of D^* decays, weighted to match the kinematics of $J/\psi K^{(*)}$ decays. Using this method the ratio of the PID efficiency in these two environments is measured to be $98.8 \pm 0.6\%$ for kaons and $99.8 \pm 0.2\%$ for pions. Although the difference from 100% is less significant than in the case of the leptons, we apply a similar correction factor of 1.2% per charged kaon and 0.2% per pion. The systematic error is then taken as the uncertainty on this correction, giving 0.6% per charged kaon and 0.2% per charged pion. The systematic uncertainties due to particle identification for all species are included in the list of efficiency systematics in Table 12.

6.1.3 Fit systematics

The second class of systematic errors in the branching fraction measurements is associated with the fixed parameters in the unbinned maximum likelihood fit. We treat these by repeating the fit with the following variations:

- The m_{ES} and ΔE mean and width are varied by $\pm 1\sigma$, as determined from the charmonium control samples.

- The K^* width is varied by ± 4 MeV, determined by the data-MC difference in the charmonium control samples.

- The size of the radiative tail in the electron channels is varied by $\pm 15\%$.

- The amount of peaking background is varied by $\pm 1\sigma$.

- The fraction of K^*s in the combinatorial background is varied by $\pm 100\%$ of itself.

- The mean of the feedup and feeddown components in ΔE is varied by ± 20 MeV.

- The combinatorial background shapes in m_{ES} and ΔE are allowed to have a linear correlation.

The effects associated with signal shape, combinatorial background shape, and peaking background yield are then added in quadrature to obtain the total fit systematics. The resulting errors for the combined fits for all q^2 are $\pm 5.9\%$ for $B \rightarrow K\ell^+\ell^-$ and $\pm 12.8\%$ for $B \rightarrow K^*\ell^+\ell^-$.

6.2 CP asymmetry systematics

In determining the direct CP asymmetry, we consider systematics arising from detector efficiency effects, and from the CP asymmetry of the peaking background. We bound the

Systematic	low q^2	high q^2	all q^2
Signal shape	0.02	0.02	0.01
Combinatoric background shape	0.05	0.06	0.01
Peaking background shape	0.01	0.01	0.01
Efficiency	0.05	0.05	0.05
Total	0.07	0.08	0.05

Table 13: Systematic uncertainties on F_L measured in the $B \to K^* \ell^+ \ell^-$ mode.

former by comparing the CP asymmetry measured in the $B \to J/\psi K^{(*)}$ data control samples with the expected values. The CP asymmetry of the peaking background is determined in data control samples for the hadronic backgrounds, and in Monte Carlo simulation for the other background sources. The value of the peaking background CP asymmetry is fixed in the fit, and a systematic uncertainty is derived by varying it within its statistical uncertainty. The CP asymmetry of the combinatorial background is a free parameter in the fit, and does not contribute to the systematic uncertainty on the measured A_{CP}. The systematic uncertainties associated with variations of the background shapes are found to be negligible. The total systematic error on A_{CP} from all sources is 0.02 for the $B^{\pm} \to K^{\pm} \ell^+ \ell^-$ mode and 0.03 for the $B \to K^* \ell^+ \ell^-$; the efficiency and peaking background systematics contribute approximately equally to this total.

6.3 A_{FB} and F_L systematics

Systematic uncertainties considered in extracting the angular distributions include those due to the signal efficiency, those due to the angular shape of the backgrounds, and those due to the correlation between A_{FB} and F_L.

Systematics due to the angular dependence of the efficiency are bounded by the difference between the measured and expected values of A_{FB}, F_L, and F_S in the charmonium control samples listed in Tables 10 and 11.

The uncertainty due to the background shape is derived by repeating the fit with sev-

Systematic	low q^2	high q^2	all q^2
Signal shape	−	0.02	−
Combinatoric background shape	−	0.07	−
Peaking background shape	−	0.01	−
Efficiency	0.02	0.02	0.02
Total	0.02	0.08	0.02

Table 14: Systematic uncertainties on A_{FB} measured in the $B \to K^* \ell^+ \ell^-$ mode. The entries listed as "-" indicate that a minimum was not found for all variations of the fit; the treatment of these cases is described in Section 7.5.3

Systematic	low q^2	high q^2	all q^2
Signal shape	0.02	0.01	0.01
Combinatoric background shape	0.18	0.03	0.08
Peaking background shape	0.01	0.01	0.01
Efficiency	0.01	0.01	0.01
Total	0.18	0.03	0.08

Table 15: Systematic uncertainties on A_{FB} measured in the $B^\pm \to K^\pm \ell^+ \ell^-$ mode.

eral alternative shapes: (1) the shape of mis-reconstructed signal events is varied to match that of correctly reconstructed signal, (2) the angular distribution of the combinatorial background is drawn from a narrower region of m_{ES} and ΔE, (3) the angular distribution of the combinatorial background is drawn from the sample of events which fail the likelihood background suppression selection, and (4) the angular distributions of the peaking backgrounds are varied within their statistical uncertainties.

A further source of systematic uncertainty enters through the correlation between A_{FB} and F_L. The systematic due to this correlation is evaluated by varying F_L within its measured uncertainties in the fit to $\cos \theta^*$.

We categorize the effects as being due to the signal shape, combinatoric background shape, peaking background shape, or efficiency and acceptance; the systematics from these four categories are then added in quadrature to obtain the total systematic error. Tables 13-14 list the resulting systematic uncertainties on the measured A_{FB} and F_L in each q^2 bin

in the $B \to K^* \ell^+ \ell^-$ mode. In most cases the dominant systematic is due to the angular shape of the combinatoric background. Table 15 lists the systematic uncertainties in the $B^\pm \to K^\pm \ell^+ \ell^-$ mode; again the combinatoric background shape is dominant systematic. In the all q^2 case the statistics are sufficient to also float the scalar term F_S; the systematic uncertainty on this quantity is determined to be 0.46, also dominated by the combinatoric background shape.

7 Results

The selection criteria, peaking backgrounds, efficiency systematics, and fit procedure described in the previous sections all being established, we proceed to fit the data. This section presents the results of the analysis, and compares them to both theoretical predictions and to other experimental results.

7.1 Branching fractions

In order to report $B \to K\ell^+\ell^-$ and $B \to K^*\ell^+\ell^-$ branching fractions, we combine charged and neutral modes, and electron and muon modes, assuming the following constraints:

- In the $B \to K^*\ell^+\ell^-$ modes, the photon pole region is taken into account by fixing the ratio of $B \to K^*\mu^+\mu^-$ to $B \to K^*e^+e^-$ to 0.75 [7]. In the $B \to K\ell^+\ell^-$ modes, the ratio is fixed to 1.

- The world average ratio of lifetimes, $\frac{\tau(B^+)}{\tau(B^0)} = 1.071 \pm 0.009$ [1], is used to constrain the total width ratio $\frac{\Gamma(B^0)}{\Gamma(B^+)}$.

We further assume equal production rates for B^0 and B^+ in $\Upsilon(4S)$ decays, consistent with the *BABAR* measurement of this ratio [63]. The charge-averaged, lepton flavor-averaged results are then quoted in terms of an effective $B \to K^{(*)0}\mu^+\mu^-$ branching fraction. From the fits to the combined modes with these constraints, we find the branching fractions:

$$\mathcal{B}(B \to K\ell^+\ell^-) = (0.34^{+0.07}_{-0.07} \pm 0.03) \times 10^{-6},$$

$$\mathcal{B}(B \to K^*\ell^+\ell^-) = (0.78^{+0.19}_{-0.17} \pm 0.12) \times 10^{-6}.$$

The results of the final combined fit projections for the data integrated over all q^2 are shown in Figures 47- 48. The secondary peak in the lower ΔE sideband of the $B \to K\ell^+\ell^-$
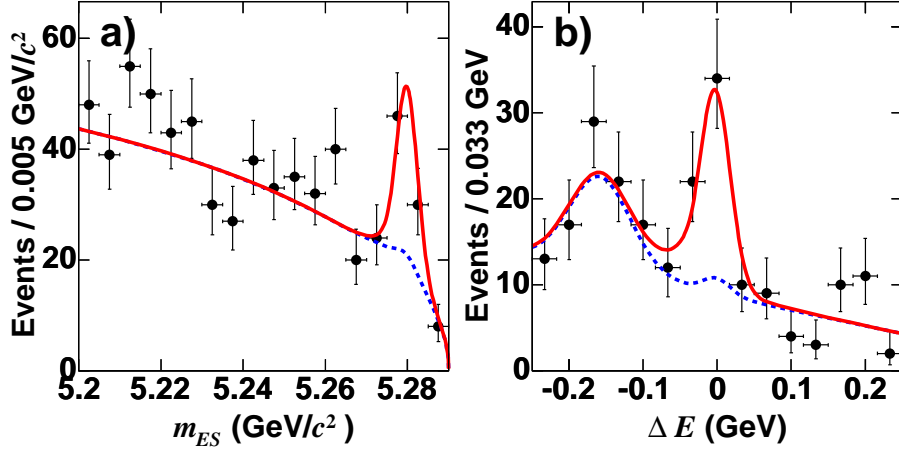
Figure 47: Combined data fit to $B \to K\ell^+\ell^-$. The projections onto (a) m_{ES} and (b) ΔE are shown for the background component of the fit (dashed line) and the total fit (solid line). For each variable plotted the signal region is selected for the other variable.

mode results from the feeddown component of the fit. We examine the events in this region, and find they are consistent with feeddown from $B \to K^*\ell^+\ell^-$ and non-resonant $b \to s\ell^+\ell^-$.

7.1.1   Signal significance

With a sufficiently large sample of events, the statistical significance of the signal can be simply estimated from the expression $\sqrt{-2\Delta \ln \mathcal{L}}$, where $\Delta \ln \mathcal{L}$ is the change in log likelihood between the nominal fit and the fit performed with the null signal hypothesis. To fully evaluate the significance of the measured signal, systematic uncertainties must be also incorporated. These are not necessarily Gaussian, and in the case of the background shape are bounded by repeating the fit with discrete choices of the background shape. We therefore incorporate systematics by simultaneously applying all variations which lower the signal yield and then re-evaluating the significance. With this prescription, the significance of the signal is 6.6σ for the $B \to K\ell^+\ell^-$ mode and 5.7σ for the $B \to K^*\ell^+\ell^-$ mode, including systematics. The significances without including systematics would be 6.9σ and 6.2σ, respectively.
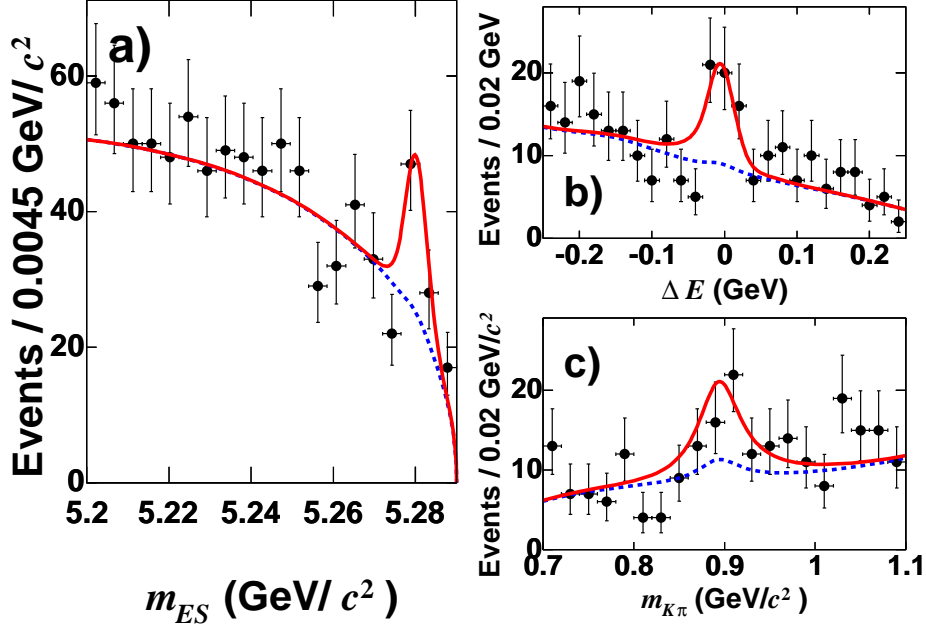
Figure 48: Combined data fit to $B \to K^* \ell^+ \ell^-$. The projections onto (a) m_{ES}, (b) ΔE, and (c) $m_{K\pi}$ are shown for the background component of the fit (dashed line) and the total fit (solid line). For each variable plotted the signal region is selected for the other variables.

7.1.2 Fits to subsamples

In addition to the combined result, we also apply the fit to the charged and neutral modes, and to the electron and muon modes separately, modifying the constraints appropriately. In the $B \to K^* \ell^+ \ell^-$ modes, we also perform the fit with the photon pole region excluded, modifying the constraint ratio of muon to electrons rates from 0.75 to 1. As expected, we find a substantial reduction in the $B \to K^* e^+ e^-$ yield by removing the pole region. The results of all such combined fits are summarized in Table 16. We find good agreement between all the subsamples.

As a further cross-check, we perform the branching fraction fit separately to each of the eight signal modes, with all constraints on the relationship between channels removed. The results are shown in Table 17. We find good agreement between the results obtained from

Mode	Effective signal yield	$\Delta\mathcal{B}_{\mathrm{eff}}$ (10^{-6})	$\Delta\mathcal{B}_{\mathrm{fit}}$ (10^{-6})	$\mathcal{B}$ (10^{-6})
$B \to Ke^+e^-$	$27.9^{+7.7}_{-6.9}$	±0.02	±0.01	$0.33^{+0.09}_{-0.08} \pm 0.02$
$B \to K\mu^+\mu^-$	$17.1^{+6.1}_{-5.3}$	±0.02	±0.03	$0.35^{+0.13}_{-0.11} \pm 0.03$
$B^\pm \to K^\pm\ell^+\ell^-$	$36.7^{+8.8}_{-8.0}$	±0.02	±0.02	$0.38^{+0.09}_{-0.08} \pm 0.03$
$B^0 \to K^0\ell^+\ell^-$	$8.2^{+4.4}_{-3.6}$	±0.02	±0.02	$0.29^{+0.16}_{-0.13} \pm 0.03$
$B \to K\ell^+\ell^-$	$45.0^{+9.7}_{-8.9}$	±0.02	±0.02	$0.34^{+0.07}_{-0.07} \pm 0.03$
$B \to K^*e^+e^-$	$36.1^{+11.2}_{-10.0}$	±0.06	±0.13	$0.97^{+0.30}_{-0.27} \pm 0.15$
$B \to K^*\mu^+\mu^-$	$20.7^{+8.1}_{-7.0}$	±0.08	±0.11	$0.90^{+0.35}_{-0.30} \pm 0.13$
$B^0 \to K^{*0}\ell^+\ell^-$	$45.2^{+11.6}_{-10.5}$	±0.06	±0.09	$0.81^{+0.21}_{-0.19} \pm 0.10$
$B^\pm \to K^{*\pm}\ell^+\ell^-$	$11.4^{+8.0}_{-6.7}$	±0.06	±0.21	$0.74^{+0.52}_{-0.43} \pm 0.22$
$B \to K^*\ell^+\ell^-$	$56.8^{+13.6}_{-12.4}$	±0.05	±0.10	$0.78^{+0.19}_{-0.17} \pm 0.12$
Pole excluded				
$B \to K^*e^+e^-$	$23.6^{+9.4}_{-8.3}$	±0.03	±0.11	$0.63^{+0.25}_{-0.22} \pm 0.11$
$B \to K^*\mu^+\mu^-$	$20.7^{+8.1}_{-7.0}$	±0.05	±0.11	$0.88^{+0.34}_{-0.30} \pm 0.12$
$B^0 \to K^{*0}\ell^+\ell^-$	$34.8^{+10.4}_{-9.3}$	±0.04	±0.10	$0.75^{+0.22}_{-0.20} \pm 0.10$
$B^\pm \to K^{*\pm}\ell^+\ell^-$	$9.5^{+7.0}_{-5.7}$	±0.05	±0.19	$0.73^{+0.53}_{-0.44} \pm 0.19$
$B \to K^*\ell^+\ell^-$	$44.3^{+12.2}_{-11.1}$	±0.04	±0.11	$0.73^{+0.20}_{-0.18} \pm 0.11$

Table 16: Branching fraction results for combined decay modes. From left the columns are: decay mode, effective fitted signal yield, efficiency related systematic error, fit related systematic error, and measured branching fraction.

the individual modes and the combined fit results.

7.1.3 Comparison of results

Figure 49 shows the results of the combined branching fraction measurement, along with the range of Standard Model predictions and the most recent published results from the Belle experiment [36]. The measured branching fractions are consistent with the range predicted by Ali *et al.* [7]; the $B \to K\ell^+\ell^-$ rate is significantly lower than the range predicted by Zhong *et al.* [64]. In both $B \to K\ell^+\ell^-$ and $B \to K^*\ell^+\ell^-$, the experimental uncertainties are substantially smaller than the theoretical uncertainties due to the form factors. We find good agreement with the results of the published Belle analysis [36]. However, a more recent preliminary Belle result reports a somewhat higher $B \to K^*\ell^+\ell^-$ branching fraction [65]
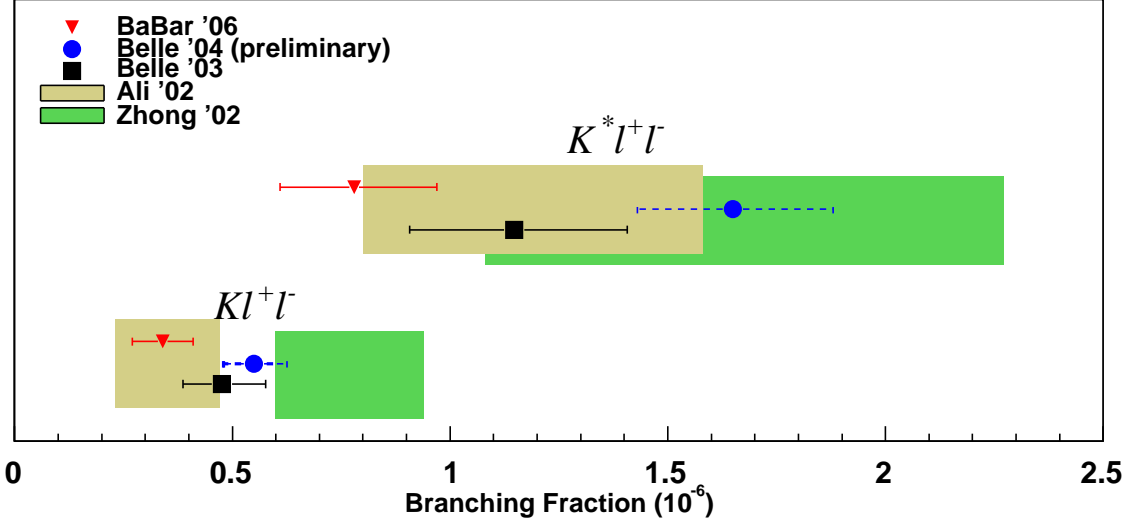
Figure 49: Total branching fractions measured in this analysis, compared to the predictions of Refs. [7] and [64], and to the recent results from the Belle collaboration [36, 65].

of $1.65^{+0.23}_{-0.22} \pm 0.09 \pm 0.04$, where the errors are due to statistics, systematics, and model dependence, respectively.

		Efficiency	$\Delta\mathcal{B}_{\text{eff}}$	$\Delta\mathcal{B}_{\text{fit}}$	$\mathcal{B}$
Mode	Signal yield	(%)	(10^{-6})	(10^{-6})	(10^{-6})
$B^{\pm} \to K^{\pm}e^{+}e^{-}$	$25.9^{+7.4}_{-6.5}$	26.4	± 0.02	± 0.02	$0.43^{+0.12}_{-0.11} \pm 0.03$
$B^{\pm} \to K^{\pm}\mu^{+}\mu^{-}$	$10.9^{+5.1}_{-4.3}$	15.2	± 0.02	± 0.04	$0.31^{+0.15}_{-0.12} \pm 0.04$
$B^{0} \to K^{0}e^{+}e^{-}$	$2.4^{+2.8}_{-2.0}$	22.6	± 0.01	± 0.01	$0.14^{+0.16}_{-0.11} \pm 0.02$
$B^{0} \to K^{0}\mu^{+}\mu^{-}$	$6.3^{+3.6}_{-2.8}$	13.3	± 0.04	± 0.03	$0.60^{+0.34}_{-0.27} \pm 0.05$
$B^{0} \to K^{*0}e^{+}e^{-}$	$29.4^{+9.5}_{-8.4}$	18.7	± 0.06	± 0.10	$1.03^{+0.33}_{-0.29} \pm 0.12$
$B^{0} \to K^{*0}\mu^{+}\mu^{-}$	$15.9^{+7.0}_{-5.9}$	11.7	± 0.08	± 0.11	$0.89^{+0.39}_{-0.33} \pm 0.14$
$B^{\pm} \to K^{*\pm}e^{+}e^{-}$	$6.2^{+7.0}_{-5.6}$	15.4	± 0.07	± 0.60	$0.77^{+0.87}_{-0.70} \pm 0.60$
$B^{\pm} \to K^{*\pm}\mu^{+}\mu^{-}$	$4.7^{+4.6}_{-3.4}$	9.0	± 0.10	± 0.13	$1.00^{+0.96}_{-0.71} \pm 0.16$

Table 17: Branching fraction results for individual decay modes. From left the columns are: decay mode, fitted signal yield, efficiency related systematic error, fit related systematic error, and measured branching fraction.

7.2 A_{CP}

From the fits to the data for all q^2, the direct CP asymmetries are determined to be

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

$$A_{CP}(B \to K^* l^+ l^-) = +0.03 \pm 0.23 \pm 0.03,$$

where the first error is statistical and the second is systematic. The measured values are consistent with the Standard Model expectation of a negligible A_{CP}. While there are no previous measurements of A_{CP} in the exclusive $B \to K^{(*)}\ell^+\ell^-$ modes, these results are consistent with the asymmetry of $A_{CP}(B \to X_s\ell^+\ell^-) = -0.22 \pm 0.26 \pm 0.02$ measured semi-inclusively by *BABAR* [37].

7.3 R_K and R_{K^*}

From the separate fits to the electron and muon samples shown in Table 16, the ratios of muon to electron branching fractions are determined to be

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

$$R_{K^*} = 0.91 \pm 0.45 \pm 0.06,$$

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

where the first error is statistical and the second systematic, respectively. In all three cases, the results are consistent with the predictions of the Standard Model (Figure 50). There are no previously published measurements of R_K or R_{K^*}; the results reported here are consistent with the preliminary measurements reported by the Belle collaboration [65].
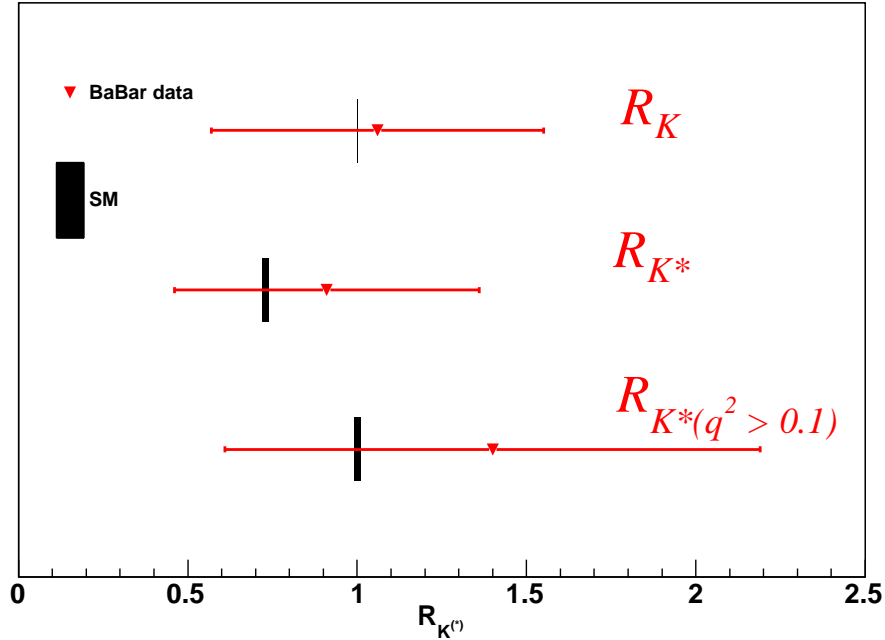
Figure 50: R_K and R_{K^*} measured in combined decay modes. The Standard Model predictions are shown by the vertical lines. The measurements of R_{K^*} with and without the pole region are highly correlated.

As discussed in Chapter 2.2, the measurement of R_K can be used along with the branching fraction $\mathcal{B}(B_s \to \mu^+\mu^-)$ to look for the contributions of neutral Higgs bosons in supersymmetric theories with large $\tan(\beta)$. The best published limits to date on the $\mathcal{B}(B_s \to \mu^+\mu^-)$ branching fraction come from the CDF [67] and D0 [68] collaborations; they report 90% CL upper limits of 1.5×10^{-7} and 5.0×10^{-7}, respectively. As shown in Figure 13, the lower of the two limits corresponds to a value of R_K of less than 1.1. If the new diagrams enter only through left-handed currents, the limits from R_K are therefore less stringent than those from $\mathcal{B}(B_s \to \mu^+\mu^-)$. If that assumption is removed, the correlation between R_K and $\mathcal{B}(B_s \to \mu^+\mu^-)$ breaks down. In that case the measured R_K gives new limits on the contribution of the neutral Higgs entering through right-handed currents.
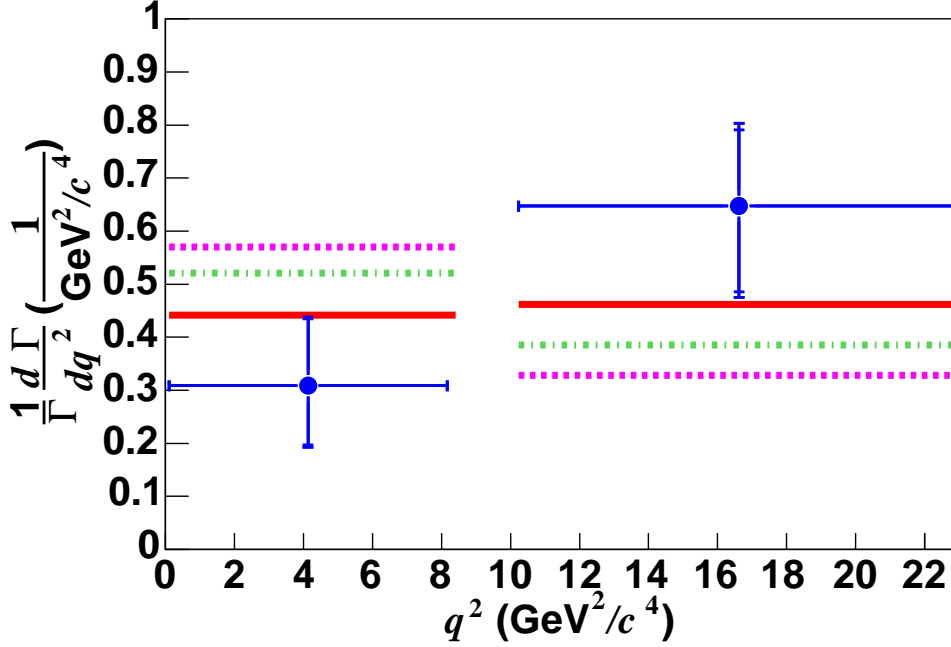
Figure 51: Partial branching fractions in the $B \to K\ell^+\ell^-$ mode. The points with error bars are data. The lines are the central values of the Standard Model predictions from the form factor models of Refs. [17, 18] (solid lines), [19] (dashed lines), and [20, 21] (dotted lines). Both data and theory are normalized to the total measured branching fraction.

7.4 Partial branching fractions

The partial branching fractions for $B \to K\ell^+\ell^-$ and $B \to K^*\ell^+\ell^-$ measured in the two regions of q^2 are shown in Table 18. The measured q^2 dependence of the rate is shown in Figures 51 and 52, compared to the q^2 dependence predicted by three Standard Model based form factor calculations. The q^2 distribution is generally in agreement with the range of predictions. While the measured central values are closest to the prediction of the Light Cone Sum Rules approach [17, 18], the statistical uncertainties are currently too large to rule out either of the alternative models. As with the total branching fraction, the high q^2 bin includes the estimated rate lost due to the $\psi(2S)$ veto.
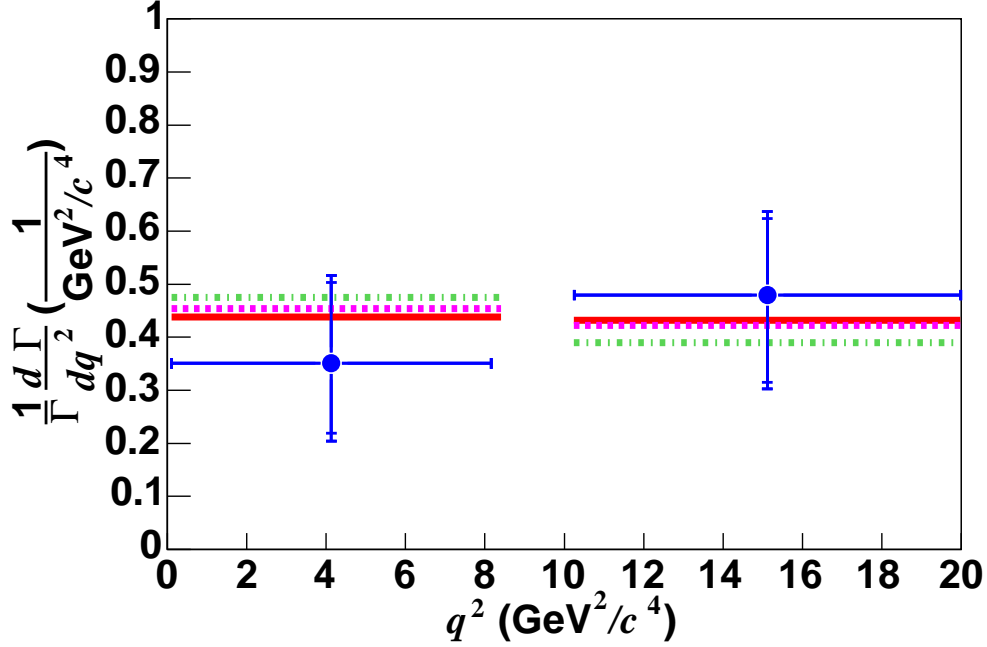
Figure 52: Partial branching fractions in the $B \to K^*\ell^+\ell^-$ mode. The points with error bars are data. The lines are the central values of the Standard Model predictions from the form factor models of Refs. [17, 18] (solid lines), [19] (dashed lines), and [20, 21] (dotted lines). Both data and theory are normalized to the total measured branching fraction.

$q^2(\text{GeV}^2/c^4)$	$\mathcal{B}(10^{-6})$	
	$B \to K\ell^+\ell^-$	$B \to K^*\ell^+\ell^-$
$0.1 < q^2 < 8.41$	$0.10^{+0.04}_{-0.04} \pm 0.01$	$0.27^{+0.12}_{-0.10} \pm 0.05$
$q^2 > 10.24$	$0.22^{+0.05}_{-0.05} \pm 0.02$	$0.37^{+0.13}_{-0.11} \pm 0.05$

Table 18: Partial branching fraction results in the combined $K^{(*)}\ell^+\ell^-$ decay modes in bins of q^2.

Numerically, the data can be compared to several recent predictions that have focused on the low q^2 region of $B \to K^*\ell^+\ell^-$, where the theoretical uncertainties are expected to be relatively small. The measured branching fractions are generally consistent with the range of predictions allowed by the form factor uncertainties, as shown in Table 19. However, the

theory predictions are done for somewhat narrower ranges of q^2. There are no previously published results for the partial branching fractions in the exclusive channels. The Belle collaboration has reported a preliminary result; as in the case of the total branching fraction, the preliminary Belle values are consistently higher than those reported here [65].

Source	q^2 range (GeV$^2/c^4$)	$\mathcal{B}$ (10^{-6})
BaBar data	0.10 - 8.41	$0.29^{+0.12}_{-0.10} \pm 0.05$
Beneke, *et al.* [69]	1.00 - 6.00	$0.33^{+0.04}_{-0.03} \times \left(\frac{A_0 \times 4GeV^2}{0.66}\right)^2$
Ali, *et al.* [70]	1.00 - 7.00	$0.29^{+0.06+0.03+0.02}_{-0.05-0.03-0.02}$
Ali, *et al.* [70] (minimal form factors)	1.00 - 7.00	$0.21^{+0.06+0.03+0.02}_{-0.05-0.03-0.02}$

Table 19: Measured partial rate in the low q^2 bin of $K^{(*)}\ell^+\ell^-$ compared to theoretical predictions. The $B^{\pm}$ rate is shown in order to compare directly with the theory predictions. The term A_0 in the prediction of Ref. [69] represents the residual dependence on one of the axial vector form factors.

7.5 Angular distributions

7.5.1 A_{FB} ($B \to K\ell^+\ell^-$)

The results for A_{FB} measured in the $B^{\pm} \to K^{\pm}\ell^+\ell^-$ decay mode are listed in Table 20. We first perform the fit under the assumption that the scalar contribution F_S is zero, as discussed in Section 5.1. With this constraint the $B^{\pm} \to K^{\pm}\ell^+\ell^-$ asymmetry acts primarily as a cross-check on the fitting procedure. In both q^2 regions, we obtain asymmetries that are consistent with zero; integrated over all $q^2 > 0.1$ GeV$^2/c^4$ we obtain

$$A_{FB}(B^{\pm} \to K^{\pm}\ell^+\ell^-) = 0.02^{+0.14}_{-0.19} \pm 0.05,$$

also consistent with zero.

In the case of the all q^2 fit, the statistics are sufficient to also allow for a non-zero scalar term in the fit. With this modification, we obtain the A_{FB} and scalar fraction results
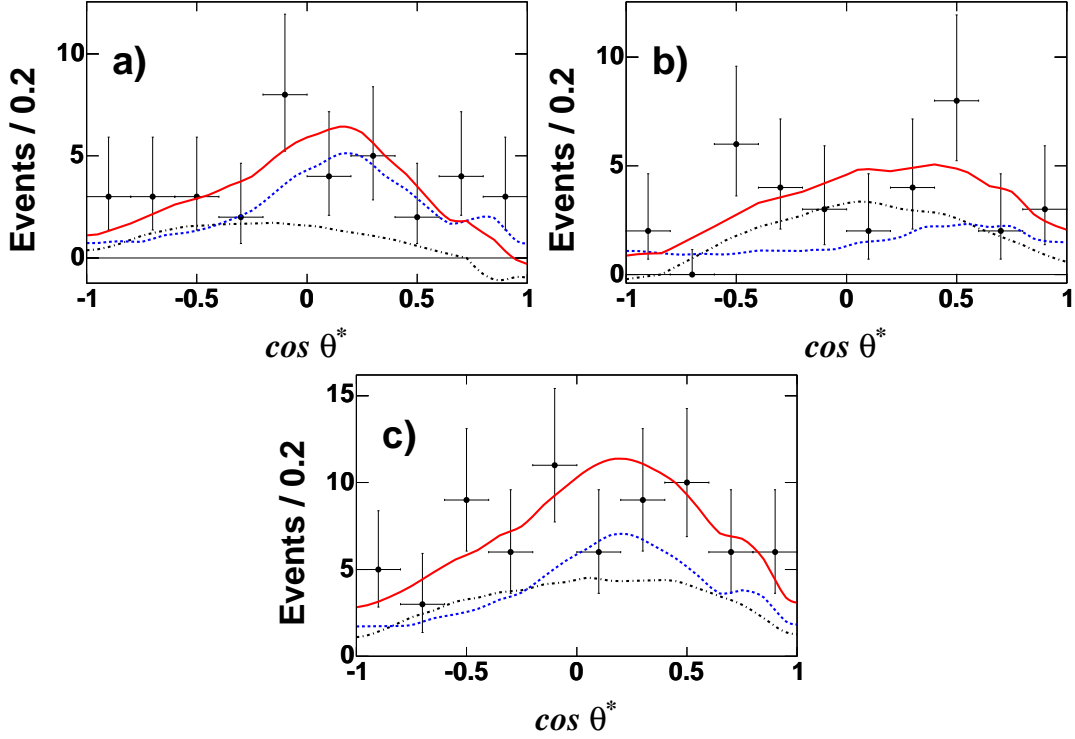
Figure 53: Distributions of the fit variable $\cos\theta^*$ in $B^\pm \to K^\pm \ell^+ \ell^-$ data (points), compared with projections of the combined fit (curves). The solid curve is the sum of all fit components, the dashed curve is the sum of all background components, and the dot-dashed curve is the signal component. The q^2 regions (a) $0.1 < q^2 < 8.41\,\mathrm{GeV}^2/c^4$, (b) $q^2 > 10.24\,\mathrm{GeV}^2/c^4$, and (c) $q^2 > 0.1\,\mathrm{GeV}^2/c^4$ are shown. The combined fits shown for (a) and (b) are performed by fixing F_S to zero, the fit shown in (c) is performed with F_S floating. The signal region is selected in m_{ES} and ΔE.

$$A_{FB}(B^\pm \to K^\pm \ell^+ \ell^-) = 0.15^{+0.21}_{-0.23} \pm 0.08,$$

$$F_S(B^\pm \to K^\pm \ell^+ \ell^-) = 0.81^{+0.58}_{-0.61} \pm 0.46,$$

where the errors are statistical and systematic, respectively. The correlation coefficient between A_{FB} and F_S is $+0.23$. With the scalar term included, we find no evidence for a non-zero asymmetry in $B^\pm \to K^\pm \ell^+ \ell^-$. The scalar fraction F_S is consistent with zero, however, the statistical errors and systematic errors due to uncertainties in the background shape

$q^2(\,\mathrm{GeV}^2/c^4)$	$B^\pm \to K^\pm \ell^+ \ell^-$ A_{FB}
$0.1 - 8.41$	$-0.49^{+0.51}_{-0.99} \pm 0.18$
> 10.24	$0.26^{+0.23}_{-0.24} \pm 0.03$
> 0.1	$0.02^{+0.14}_{-0.19} \pm 0.05$

Table 20: Results from fits to the combined $B^\pm \to K^\pm \ell^+ \ell^-$ decay mode in bins of q^2, with the scalar contribution fixed to zero. The columns from left to right are the fitted q^2 range, and the lepton forward-backward asymmetry A_{FB}. The first and second uncertainties are statistical and systematic, respectively.

are too large to provide a significant bound on this quantity. The data and 1-dimensional projections of the likelihood fit are shown for the $B^\pm \to K^\pm \ell^+ \ell^-$ modes in Figure 53.

The only other measurement of A_{FB} in $B^\pm \to K^\pm \ell^+ \ell^-$ comes from the Belle experiment. They measure a discrete asymmetry of $A_{FB}(B^\pm \to K^\pm \ell^+ \ell^-) = 0.10 \pm 0.14 \pm 0.01$, consistent with zero and with the results of this analysis. The Belle analysis does not attempt to extract the scalar contribution [66].

7.5.2 K^* polarization

The values of F_L measured in the two regions of q^2 and integrated over all q^2 are shown in Table 21. The longitudinal polarization of the K^* is consistent with the Standard Model prediction in both regions of q^2 (Figure 55); due the large statistical uncertainties the data cannot distinguish between the SM and models in which the C_7^{eff} Wilson coefficient has the opposite sign. There are no previous measurements of the K^* polarization in the $B \to K^* \ell^+ \ell^-$ decay mode. The data and 1-dimensional projections of the likelihood fit are shown for the $B^\pm \to K^\pm \ell^+ \ell^-$ modes in Figure 54.
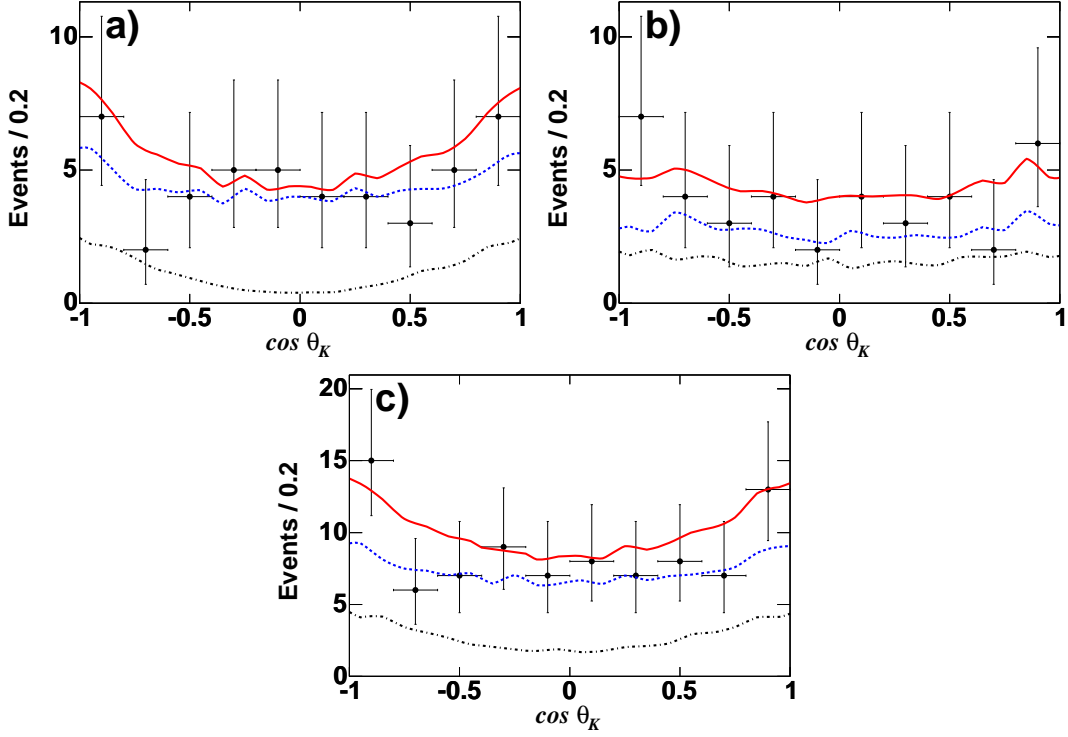
Figure 54: Distributions of the fit variable $\cos\theta_K$ in $B \to K^*\ell^+\ell^-$ data (points), compared with projections of the combined fit (curves). The solid curve is the sum of all fit components, the dashed curve is the sum of all background components, and the dot-dashed curve is the signal component. The q^2 regions (a) $0.1 < q^2 < 8.41\,\mathrm{GeV}^2/c^4$, (b) $q^2 > 10.24\,\mathrm{GeV}^2/c^4$, and (c) $q^2 > 0.1\,\mathrm{GeV}^2/c^4$ are shown. The signal region is selected in m_{ES}, ΔE, and m_{K^*}.

7.5.3 A_{FB} $(B \to K^*\ell^+\ell^-)$

The value of A_{FB} in the high q^2 region for $B \to K^*\ell^+\ell^-$ is large and positive, consistent with the prediction of the Standard Model. The likelihood scan of A_{FB} is shown in Figure 57. Numerically, we find that an asymmetry with the same magnitude but opposite sign as the SM is excluded at 3.0σ, including systematic errors and using the most conservative assumptions for the form factor model. We find that a negative asymmetry of any magnitude is disfavored at 2.2σ, including systematics. The positive asymmetry is consistent with the SM and disfavors models in which the sign of the product of C_9^{eff} and C_{10}^{eff} Wilson coefficients
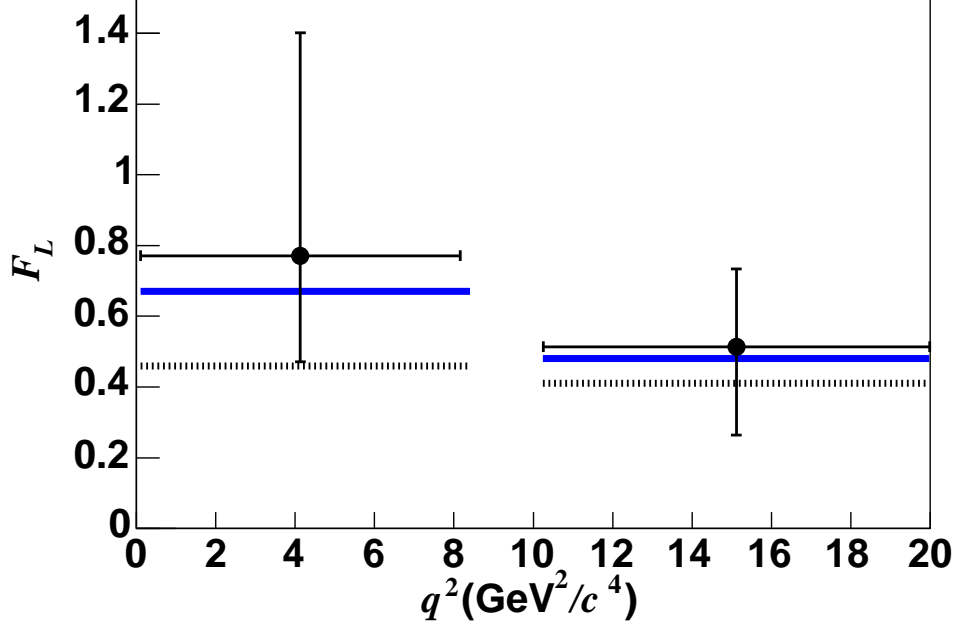
Figure 55: F_L measured in $B \to K^*\ell^+\ell^-$. The points with error bars are data. The lines are the predictions for the SM (solid lines) and wrong sign C_7^{eff} (dotted lines) scenario.

is opposite that of the SM, as these would lead to a negative A_{FB} at high q^2.

In the low q^2 range, the asymmetry is large and positive, but near the boundary at which

	$B \to K^*\ell^+\ell^-$	
$q^2(\mathrm{GeV}^2/c^4)$	F_L	A_{FB}
$0.1 - 8.41$	$0.77^{+0.63}_{-0.30} \pm 0.07$	> 0.19 (95%CL)
> 10.24	$0.51^{+0.22}_{-0.25} \pm 0.08$	$0.72^{+0.28}_{-0.26} \pm 0.08$
> 0.1	$0.63^{+0.18}_{-0.19} \pm 0.05$	> 0.55 (95%CL)

Table 21: Results from fits to the combined $K^*\ell^+\ell^-$ decay modes in bins of q^2. The columns from left to right are: fitted q^2 range, longitudinal K^* polarization F_L, and the lepton forward-backward asymmetry A_{FB}. The first and second uncertainties are statistical and systematic, respectively.
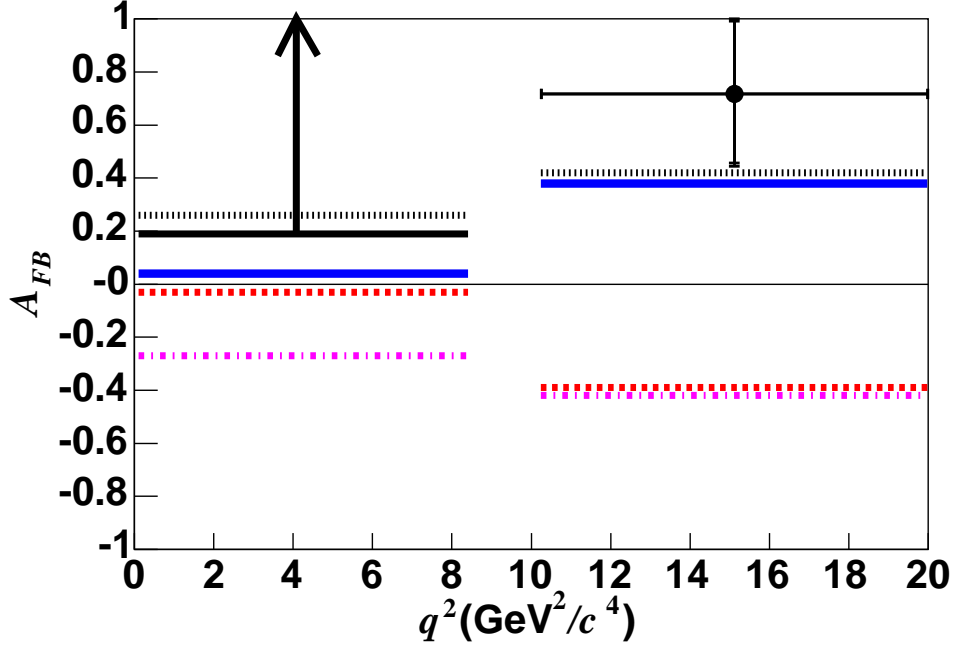
Figure 56: A_{FB} measured in $B \to K^* \ell^+ \ell^-$. The points with error bars are data; the arrow at low q^2 in A_{FB} represents the 95% CL lower limit. The lines are the predictions for the SM (solid lines), wrong sign C_7^{eff} (dotted lines), wrong sign $C_9^{eff} C_{10}^{eff}$ (dashed lines), and wrong sign C_7^{eff} and $C_9^{eff} C_{10}^{eff}$ (dot-dashed lines).

the likelihood function becomes negative and undefined. Thus, the fit fails to converge to a stable minimum. Therefore, we set a one-sided lower limit on the asymmetry. We generate toy experiments for a series of values of A_{FB}. We then define the 95% confidence level lower limit as the value of A_{FB} for which 5% of experiments give the behavior we observe. We incorporate systematics into this approach by smearing the parameters which are fixed in the fit when generating the toys. We repeat this study for each of the alternative background shapes discussed in Section 6, choosing the one which gives the weakest (lowest) lower limit. Using this procedure, we find at low q^2 the limit $A_{FB} > 0.19$ at 95% CL. The limit is slightly higher than the recent Standard Model predictions using the form factor models of
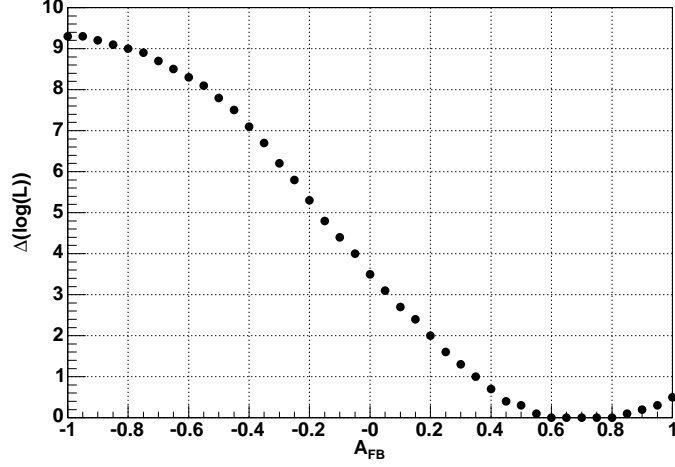
Figure 57: A_{FB} likelihood scan for the high q^2 region of $B \to K^* \ell^+ \ell^-$. The likelihood for each point is computed by fixing the value of A_{FB} and refitting the other parameters.

Refs. [18, 64], and the older calculation used in Ref. [7], all of which predict $A_{FB} = 0.03$ at low q^2. For an assumed $A_{FB} = 0.03$, the probability of obtaining our result is approximately 2%, equivalent to a 2.05 σ deviation for a Gaussian normal distribution. The form factor calculation of Ref. [20, 21] appears to allow a larger SM asymmetry at low q^2 which would be more consistent with the measured limit; the inconsistency between this approach and the other theoretical calculations has been discussed in Ref. [64]. In a similar manner we evaluate a one-sided bound on the value of A_{FB} integrated over all $q^2 > 0.1$ GeV^2/c^4. We find at 95% CL that $A_{FB} > 0.55$. The data and 1-dimensional projections of the likelihood fit are shown for the $B^\pm \to K^\pm \ell^+ \ell^-$ modes in Figure 58.

The Belle collaboration has reported a result using 357 fb^{-1} of data, in which they attempt to extract the Wilson coefficients directly from the q^2 and $\cos\theta^*$ distributions. They find that models in which the sign of the product of C_9^{eff} and C_{10}^{eff} is opposite the Standard Model are disfavored at 98.2% CL. The Belle result is therefore consistent with the result obtained here from the high q^2 data. They report a discrete forward-backward asymmetry
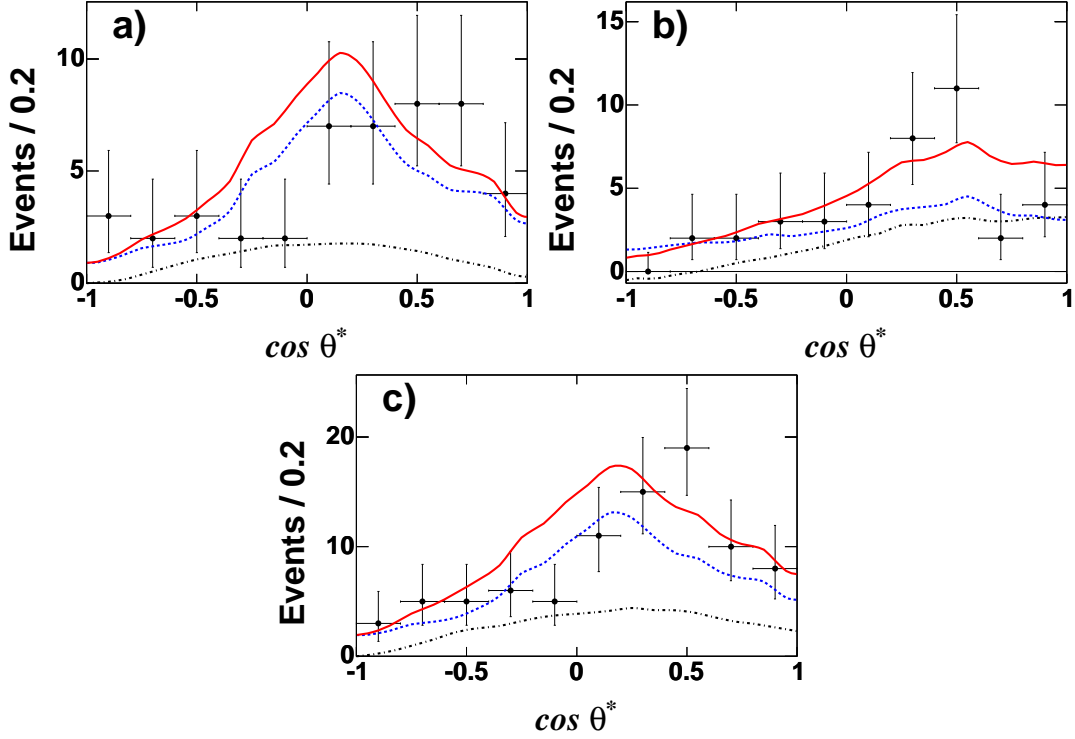
Figure 58: Distributions of the fit variable $\cos\theta^*$ in $B \to K^*\ell^+\ell^-$ data (points), compared with projections of the combined fit (curves). The solid curve is the sum of all fit components, the dashed curve is the sum of all background components, and the dot-dashed curve is the signal component. The q^2 regions (a) $0.1 < q^2 < 8.41\,\mathrm{GeV}^2/c^4$, (b) $q^2 > 10.24\,\mathrm{GeV}^2/c^4$, and (c) $q^2 > 0.1\,\mathrm{GeV}^2/c^4$ are shown. The combined fits shown for (a) and (c) are performed by fixing A_{FB} to its maximal physical value. The signal region is selected in m_{ES}, ΔE, and m_{K^*}.

integrated over all $q^2 > 0.1$ GeV^2/c^4 of $A_{FB}(B \to K^*\ell^+\ell^-) = 0.50 \pm 0.15 \pm 0.02$, where the errors are statistical and systematic [66].

7.6 Search for lepton flavor violation

We extract the signal in the lepton flavor-violating $B \to K^{(*)}e\mu$ modes in a manner identical to the other $B \to K^{(*)}l^+l^-$ modes. Since any processes which can generate these decays will not necessarily affect $e^+\mu^-$ and $e^-\mu^+$ equally, we first obtain the yield in each

Mode	Effective signal yield	90% CL limit ($\times 10^{-8}$)	Previous limit ($\times 10^{-8}$)
$B^{\pm} \to K^{\pm} e^{+} \mu^{-}$	$-3.5^{+2.1}_{-1.4} \pm 0.7$	9.1	$-$
$B^{\pm} \to K^{\pm} e^{-} \mu^{+}$	$-0.8^{+2.1}_{-1.3} \pm 0.5$	13	$-$
$B^{0} \to K^{*0} e^{+} \mu^{-}$	$1.1^{+3.6}_{-2.1} \pm 0.8$	53	$-$
$B^{0} \to K^{*0} e^{-} \mu^{+}$	$-1.1^{+3.5}_{-2.2} \pm 1.1$	34	$-$
$B^{\pm} \to K^{*\pm} e^{+} \mu^{-}$	$0.4^{+3.4}_{-2.3} \pm 1.0$	130	$-$
$B^{\pm} \to K^{*\pm} e^{-} \mu^{+}$	$-1.7^{+3.3}_{-2.0} \pm 0.8$	99	$-$
$B^{\pm} \to K^{\pm} e \mu$	$-3.2^{+1.7}_{-2.7} \pm 0.9$	9.1	80 [33]
$B^{0} \to K^{0} e \mu$	$-2.9^{+1.9}_{-1.3} \pm 1.5$	27	400 [33]
$B^{0} \to K^{*0} e \mu$	$0.9^{+4.6}_{-2.9} \pm 1.4$	58	340 [33]
$B^{\pm} \to K^{*\pm} e \mu$	$-0.2^{+4.2}_{-3.1} \pm 1.6$	140	790 [33]
$B \to K e \mu$	$-4.9^{+2.9}_{-1.9} \pm 1.2$	3.8	160 [71]
$B \to K^{*} e \mu$	$1.0^{+5.5}_{-3.7} \pm 2.3$	51	620 [71]

Table 22: Results for the lepton flavor-violating modes. The columns from left are: decay mode, effective signal yield, 90% CL upper limit, and the previous best experimental upper limit.

charge state separately, before obtaining combined yields assuming lepton-charge symmetry.

We find no evidence for a signal in any of these decays, and therefore set upper limits on the branching fractions. We determine the 90% confidence level by generating an ensemble of toy experiments and finding the point at which 10% of experiments will give a smaller yield.

The data and projections of the fit in the lepton-charge averaged, B charge-averaged modes are shown in Figures 59- 60. For the combined modes we find at 90% CL:

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

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

Prior to this, the most recent searches for these modes were conducted by *BABAR* using 20.7 fb^{-1} of data [33], and by the CLEO collaboration using a 9.2 fb^{-1} data sample [71]. The results of those searches compared to the results reported here are shown in Table 22. Most
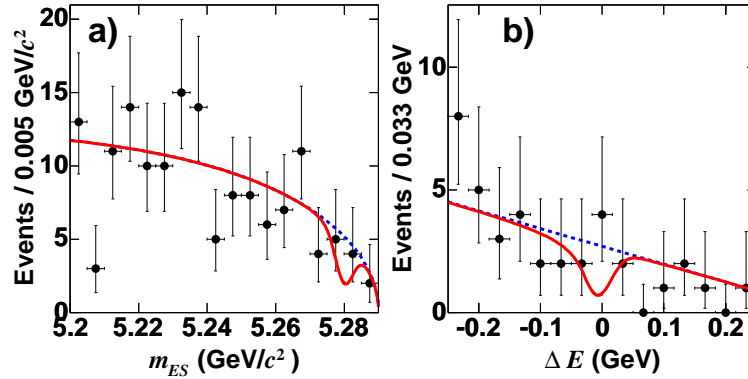
Figure 59: Combined data fit to $B \to Ke\mu$. The projections onto (a) m_{ES} and (b) ΔE are shown for the background component of the fit (dashed line) and the total fit (solid line). The signal region is selected in m_{ES} and ΔE.

Figure 60: Combined data fit to $B \to K^*e\mu$. The projections onto (a) m_{ES} and (b) ΔE are shown for the background component of the fit (dashed line) and the total fit (solid line). The signal region is selected in m_{ES}, ΔE, and m_{K^*}.
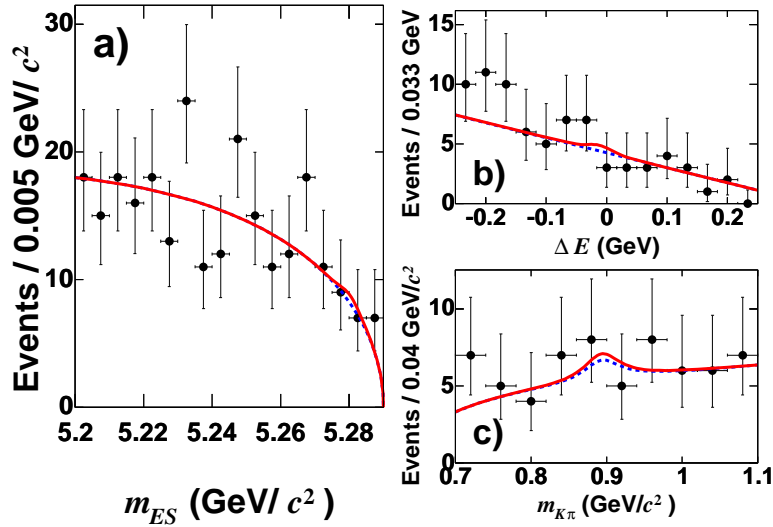
of the limits reported here are roughly an order of magnitude below the earlier *BABAR* and CLEO results, and are the most stringent limits on these processes to date.
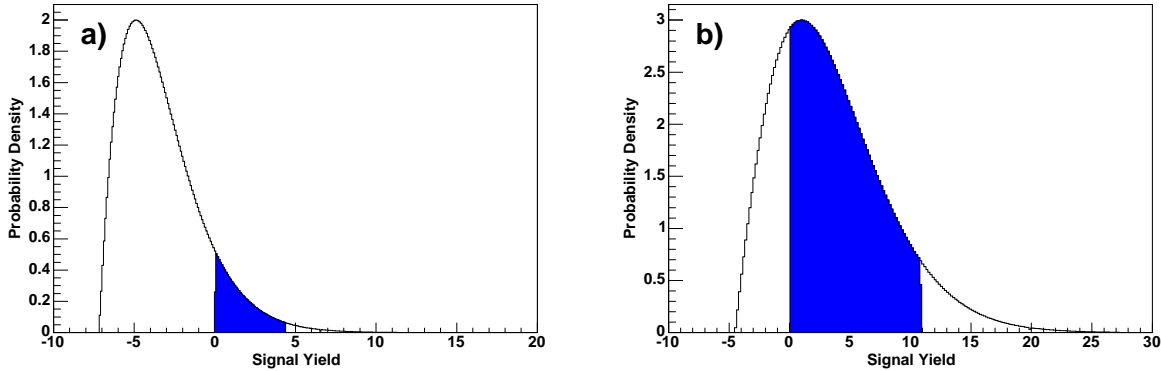
Figure 61: Example of a Bayesian limit-setting procedure in (a) $B \to Ke\mu$, and (b) $B \to K^*e\mu$. The open histogram shows the experimental likelihood, with the solid region showing the portion integrated to obtain the Bayesian 90% CL.

7.6.1 Discussion of $B \to K^{(*)}e\mu$ limits

The central values of the signal yield in several of the lepton flavor-violating modes are negative, as expected for a process whose true rate is close to zero in the presence of background. In the classical/frequentist approach used to derive the 90% CL, this necessarily results in upper limits that are somewhat lower than the analysis sensitivity, or limit that would have been obtained in a hypothetical experiment in which a signal yield of exactly zero was observed in data.

An alternative approach favored by some would be to perform a Bayesian analysis, incorporating prior beliefs into the limit setting procedure in order to make statements about the most probable true value of the branching fraction. One possible prior would consist of a step function at zero with a flat prior in the branching fraction for positive values. Such a prior would represent the belief that negative values are disallowed, and all positive values are equally likely outcomes. The experimental likelihood could then be integrated from zero to obtain an upper limit. The Bayesian procedure is illustrated in Figure 61 for

the combined modes. The resulting 90% credible upper limits from this procedure would be approximately $\mathcal{B}(B \to Ke\mu) < 11 \times 10^{-8}$ and $\mathcal{B}(B \to K^*e\mu) < 54 \times 10^{-8}$. Further discussion of the frequentist and Bayesian approaches can be found, for example, in Ref. [72].

As discussed in Ref. [49], lepton flavor-violating decays can be used to set limits on the ratio $\lambda_{ab}\lambda_{cd}/M_{LQ}^2$ in leptoquark models. Here M_{LQ} is the leptoquark mass, and λ is the leptoquark coupling; the subscripts $a - d$ run from one to three and denote the generation of the lepton and quark to which the leptoquark couples. A recent summary of leptoquark limits can be found in Ref. [73], showing the constraints obtained from rare meson decays, muon conversions in nuclear scattering, and direct searches. In the particular case of leptoquarks with $\lambda_{12}\lambda_{23}$ and $\lambda_{13}\lambda_{22}$ couplings, $B \to Ke\mu$ decays provide the strongest limits to date.

8   Summary and Conclusions

We have measured the branching fractions, partial branching fractions, direct CP asymmetry, ratio of muons to electrons, longitudinal K^* polarization, and forward-backward asymmetry in the rare decays $B \to K\ell^+\ell^-$ and $B \to K^*\ell^+\ell^-$.

The measured total branching fractions are:

$$\mathcal{B}(B \to K\ell^+\ell^-) = (0.34 \pm 0.07 \pm 0.03) \times 10^{-6},$$

$$\mathcal{B}(B \to K^*\ell^+\ell^-) = (0.78^{+0.19}_{-0.17} \pm 0.12) \times 10^{-6}.$$

Both the total and partial branching fractions as a function of q^2 are consistent with the range of Standard Model based predictions. The CP asymmetries in the total rate are consistent with the vanishingly small values predicted in the SM. The ratios of muons to electrons, R_K and R_{K^*}, are consistent with the SM predictions and begin to restrict the allowed contributions of non-SM Higgs penguin diagrams.

The longitudinal polarization of the K^* as a function of q^2 is consistent with the SM, however the present statistical precision of this measurement can not rule out models in which the sign of the C_7^{eff} Wilson coefficient is opposite that of the SM.

The value of A_{FB} measured in the $B^\pm \to K^\pm\ell^+\ell^-$ mode is consistent with zero in both q^2 ranges and integrated over all q^2. The value of A_{FB} at high q^2 in the $B \to K^*\ell^+\ell^-$ channel is found to be large and positive, consistent with the SM. The positive asymmetry disfavors models in which the sign of the product of C_9^{eff} and C_{10}^{eff} is opposite that of the SM. At low q^2, a positive asymmetry is also favored, with a limit that is consistent with most recent Standard Model predictions at the 2% level. i Finally, we have set new upper limits on the rates of the lepton-flavor violating decays $B \to K e\mu$ and $B \to K^* e\mu$ that are approximately one order of magnitude more stringent than previous searches.

8.1 Outlook

All of the measurements presented here are based on 208 fb^{-1} of data, and are limited by statistical uncertainties. The current goal of the *BABAR* Collaboration is to collect approximately five times the amount of data used in this analysis, for a total integrated luminosity of 1 ab^{-1}, before ending operations. At that point, the statistical errors on the total branching fractions of $B \to K\ell^+\ell^-$ and $B \to K^*\ell^+\ell^-$ would be comparable to the systematic uncertainties, assuming no further improvements to the systematic errors. The uncertainties on the asymmetries, ratios, and angular distributions will continue to be limited by statistics beyond that point.

References

[1] W. -M. Yao *et al.*, (Particle Data Group), J. Phys. G **33**, 1 (2006)

[2] G. Abbiendi, *et al.*, the ALEPH Collaboration, the DELPHI Collaboration, the L3 Collaboration, the OPAL Collaboration (The LEP Working Group for Higgs Boson Searches) Phys. Lett. B **565**, 61 (2003)

[3] M. Bona *et al.* (UTfit Collaboration), JHEP **028**, 0507 (2005).

[4] J. Charles *el al.* (CKMFitter Collaboration) Eur. Phys. J. C **41**, 1 (2005)

[5] S. P. Martin in *Perspectives on supersymmetry*, edited by G. L. Kane (World Scientific, Singapore, 1998).

[6] A. Ali *et al.*, Phys. Rev. D **61**, 074024 (2000).

[7] A. Ali *et al.*, Phys. Rev. D **66**, 034002 (2002).

[8] J. L. Hewett and J. D. Wells, Phys. Rev. D **55**, 5549 (1997).

[9] E. Lunghi *el al.*, Nucl. Phys. B **568**, 120 (2000).

[10] G. Burdman, Phys. Rev. D **52**, 6400 (1995).

[11] G. Eilam, J. L. Hewett, and T. G. Rizzo, Phys. Rev. D **34**, 2773 (1986); T. M. Aliev, A. Ozpineci, and M. Savci, Eur. Phys. J. C**29**, 265 (2003).

[12] T. M. Aliev, M. Savci, and A. Özpineci, Phys. Lett. B **410**, 216 (1997).

[13] Q. Yan, C. Huang, L. Wei, and S. Zhu, Phys. Rev. D **62**, 094023 (2000).

[14] G. Hiller and F. Krüger, Phys. Rev. D **69**, 074020 (2004).

[15] G. Buchalla, A. J. Buras and M. E. Lautenbacher, Rev. Mod. Phys. **68**, 1125 (1996).

[16] K. G. Wilson, Phys. Rev. **179**, 1499 (1969)

[17] P. Ball and R. Zwicky, Phys. Rev. D **71**, 014015 (2005).

[18] P. Ball and R. Zwicky, Phys. Rev. D **71**, 014029 (2005).

[19] D. Melikhov, N. Nikitin, and S. Simula, Phys. Rev. D **57**, 6814 (1998).

[20] P. Colangelo *et al.*, Phys. Rev. D **53**, 3672 (1996).

[21] P. Colangelo *et al.*, Phys. Rev. D **57**, 3186 (1998).

[22] P. Gambino, U. Haisch, M. Misiak, Phys. Rev. Lett. **94**, 061803 (2005).

[23] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. D. **72**, 052004 (2005).

[24] P. Koppenburg *et al.* (Belle Collaboration), Phys. Rev. Lett. **93**, 061803 (2004).

[25] K. Abe *et al.* (Belle Collaboration), Phys. Lett. B **511**, 151 (2001).

[26] S. Chen *et al.* (CLEO Collaboration), Phys. Rev. Lett. **87**, 251807 (2001).

[27] C. Bobeth *et al.*, Nucl. Phys. B **726**, 252 (2005).

[28] S. Anderson, *et al.* (CLEO Collaboration), Phys. Rev. Lett. **87** 181803 (2001).

[29] P. Avery, *et al.* (CLEO Collaboration), Phys. Lett. B **223** 470 (1989).

[30] C. Albajar, *et al.* (UA1 Collaboration), Phys. Lett. B **262** 163 (1991).

[31] T. Affolder, *et al.* (CDF Collaboration), Phys. Rev. Lett. **83** 3378 (1999).

[32] F. Abe, *et al.* (CDF Collaboration), Phys. Rev. Lett. **76** 4675 (1996).

[33] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. Lett. **88**, 241801 (2002).

[34] K. Abe *et al.* (Belle Collaboration), Phys. Rev. Lett. **88**, 021801 (2002).

[35] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. Lett. **91**, 221802 (2003).

[36] A. Ishikawa *et al.* (Belle Collaboration), Phys. Rev. Lett. **91**, 261601 (2003).

[37] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. Lett. **93** 261601 (2004).

[38] M. Iwasaki, *et al.* (Belle Collaboration), Phys. Rev. D **72**, 092005 (2005).

[39] Y. Wang and D. Atwood, Phys. Rev. D **68**, 094016 (2003).

[40] F. Krüger, L. M. Sehgal, N. Sinha, and R. Sinha, Phys. Rev. D **61**, 114028 (2000); Erratum-ibid. D **63**, 019901 (2001).

[41] F. Krüger and E. Lunghi, Phys. Rev. D **63**, 014013 (2001).

[42] G. Buchalla, G. Hiller, and G. Isidori, Phys. Rev. D **63**, 014015 (2001).

[43] F. Krüger and J. Matias, Phys. Rev. D **71**, 094009 (2005).

[44] T. Feldmann and J. Matias, JHEP **0301**, 074 (2003).

[45] A. J. Buras, R. Fleischer, S. Recksiegel, and F. Schwab, Nucl. Phys. B **697**, 133 (2004).

[46] C. H. Chen, C. Q. Geng, and A. K. Giri Phys. Lett. B **621**, 253 (2005).

[47] C. Bobeth, T. Ewerth, F. Krüger, and J. Urban Phys. Rev. D **64**, 074014 (2001).

[48] D. A. Demir and K. Olive, Phys. Rev. D **66**, 034015 (2002).

[49] S. Davidson, D. C. Bailey, and B. A. Campbell, Z. Phys. C **61**, 613 (1994).

[50] *BABAR* Collaboration, B. Aubert *et al.*, Nucl. Instrum. Methods A **479**, 1 (2001).

[51] C. Bozzi *et al.*, Nucl. Instrum. Methods A **461**, 162-167 (2001).

[52] P. Billoir, Nucl. Instrum. Meth A **225**, 352 (1984).

[53] Nucl. Instrum. Methods A **535**, 206-211 (2004).

[54] I. Adam *et al.*, Nucl. Instrum. Methods A **538**, 281-357 (2004).

[55] F. Anulli *et al.*, Nucl. Instrum. Methods A **494**, 455-463 (2002).

[56] F. Anulli *et al.*, Nucl. Instrum. Methods A **539**, 155-171 (2005).

[57] F. Anulli *et al.*, Nucl. Instrum. Methods A **552**, 276-291 (2005).

[58] G. C. Fox and S. Wolfram, Phys. Rev. Lett. **41**, 1581 (1978).

[59] R. A. Fisher, Ann. Eugenics **7**, 179 (1936).

[60] T. Skwarnicki (Crystal Ball Collaboration), Ph.D. thesis, Cracow Institute of Nuclear Physics (1986), DESY F31-86-02.

[61] H. Albrecht *et al.* (ARGUS Collaboration), Z. Phys. C **48**, 543 (1990).

[62] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. D **71**, 032005 (2005).

[63] B. Aubert, *et al.* (*BABAR* Collaboration), Phys. Rev. D **69**, 071101 (2004).

[64] M. Zhong , Y.L. Wu, and W.Y. Wang, Int. J. Mod. Phys. A**18**, 1959 (2003).

[65] K. Abe *et al.* (Belle Collaboration), BELLE-CONF-0415 [hep-ex/0410006].

[66] A. Ishikawa, *et al.* (Belle Collaboration), hep-ex/0603018, submitted to PRL

[67] D. Acosta, *et al.* (CDF Collaboration), Phys. Rev. Lett. **93**, 032001 (2004)

[68] V. M. Abazov, *et al.* (D0 Collaboration), Phys. Rev. Lett. **94**, 071802 (2005)

[69] M. Beneke, Th. Feldmann, and D. Seidel, Eur. Phys. J. C **41**, 173-188 (2005)

[70] A. Ali, G. Kramer, and G.Zhu, Eur. Phys. J. C , hep-ph/0601034

[71] K. W. Edwards *et al.* (CLEO Collaboration), Phys. Rev. D **65**, 111102 (2002).

[72] R. D. Cousins Am. J. Phys **63**, 398 (1995).

[73] S. Chekanov *et al.* (ZEUS Collaboration), Eur. Phys. J. C **44**, 463 (2005).

A   Muon identification algorithms in *BABAR*

In addition to the IFR hardware improvements described in Section 3.2.5, this analysis makes use of two significant improvements to the muon identification software that were unavailable to the previous *BABAR* analyses of $B \to K^{(*)}\mu^+\mu^-$:

1. Implementation of a Kalman filter for track reconstruction in the IFR.

2. Development of a neural network-based muon ID algorithm.

Training of the neural network is performed in two bins of momentum ($0.5 < p < 2.0$ GeV/c and $2.0 < p < 4.0$ GeV/c) and two bins of polar angle ($0.3 < \theta < 0.7$ and $0.7 < \theta < 2.7$), roughly separating the barrel and endcap regions of the IFR. The selection on the output is chosen to maintain a continuous transition in the muon ID performance between bins of p and θ. The neural network is also trained separately for different run periods, in order to account for the substantial variation in the performance of the RPCs over time. A typical example of the neural network output is shown in Figure 62.
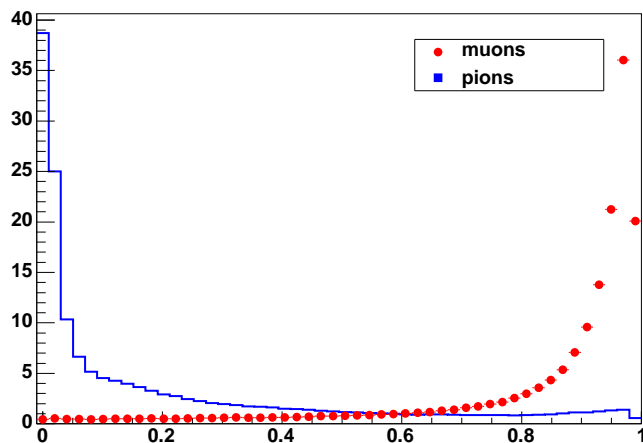

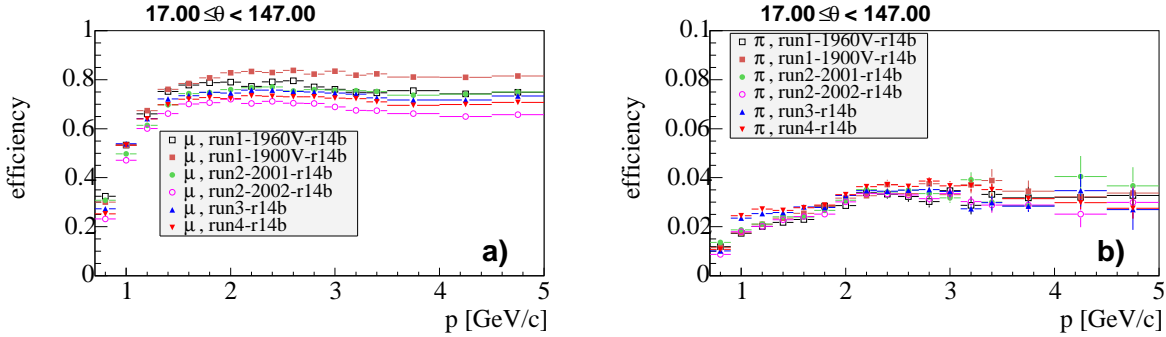Figure 62: Training output of the muon identification neural network.

Figure 63: Performance of the muon ID vs. momentum for different time periods for (a) muon efficiency, and (b) pion misidentification. The neural network "tight" selection used in this analysis is shown.

The network architecture comprises eight inputs (described in Section 4.1.2), one hidden layer with 16 nodes, and one output node. The training sample of muons is obtained from $e^+e^- \to \mu^+\mu^-\gamma$ data events, and the training sample of pions from three-prong τ decays. An example of the output of training is shown in Figure 62. Half of the available control sample events are used for training, and the remaining half are used for validation of the neural network performance. Eight different selection criteria are available; as the IFR performance varies over time these are tuned for either constant muon efficiency or constant pion misidentification. For the analysis of $B \to K^{(*)}\mu^+\mu^-$ we choose the "tight" criteria, which provides $\sim 70\%$ muon efficiency with a pion misidentification probability of $2-3\%$. The efficiency and misidentification probability versus momentum for different time periods is shown in Figure 63.

The effect on muon ID due to the Kalman filter and neural network is illustrated in Figure 64, for muons with momentum of ~ 2 GeV/c, near the peak of the spectrum for $B \to K^{(*)}\mu^+\mu^-$ events. The Kalman filter alone provides a noticeable improvement, particularly for the tighter selection criteria. The addition of the neural network is a further improvement
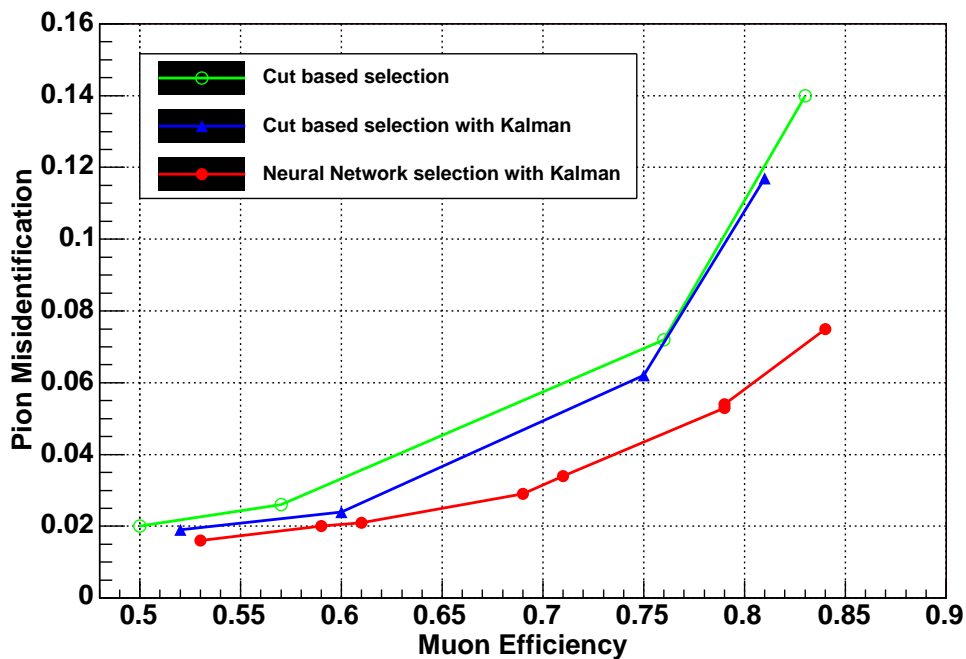
Figure 64: Comparison of different muon ID algorithms available in *BABAR*. All three algorithms are evaluated for 2 GeV/c muons selected from data collected during 2003. The points indicate the various selection criteria available to *BABAR* analyses.

for all selection criteria. Compared to the cut based selector used previously, the combination of the neural network and Kalman filter provides an improvement of up to $\sim 10\%$ in the absolute muon efficiency while retaining the same level of pion misidentification. For decays such as $B \to K^{(*)}\mu^+\mu^-$, which require identification of two muons, the increased efficiency significantly improves the sensitivity of the analysis.