UNIVERSITY of CALIFORNIA
Santa Barbara

SEARCH FOR EXCLUSIVE ELECTROWEAK PENGUIN DECAYS AT *BABAR*

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

Doctor of Philosophy

in

Physics

by

Natalia Kuznetsova

Committee in charge:

Professor Jeffrey Richman, Chair
Professor Rollin Morrison
Professor Mark Srednicki

August 2001

The dissertation of Natalia Kuznetsova is approved:

Chair

July 2001

SEARCH FOR EXCLUSIVE ELECTROWEAK PENGUIN DECAYS AT *BABAR*

by

Natalia Kuznetsova

ACKNOWLEDGMENTS

I would like to thank Mike Witherell who was my thesis advisor from January 1997 till March 1999, when he became the director of Fermilab. I consider myself very fortunate that I have had an opportunity to work with Mike.

I am very grateful to Jeff Richman who took it over when Mike left. These two years of working with Jeff were very exciting and productive. I would particularly like to thank Jeff for his constant support, which I have always felt. In any situation, no matter how complicated or even seemingly hopeless, concerning physics or anything else, I knew I could count on Jeff's advice and kind encouragement, and this made it possible to keep going.

I would also like to thank Rolly Morrison and Claudio Campagnari, who have always been very helpful. Numerous enjoyable conversations with Claudio had a profound influence on my way of thinking about our field.

I would like to express my deepest gratitude to Anders Ryd who first introduced me to the world of the *BABAR* software, and was invariably providing great help and guidance despite his own extremely busy schedule. It was a privilege to work together with Anders on this analysis, which never would have reached the stage it has without Anders' remarkable insight.

Many friends and colleagues at *BABAR* have helped me over these years. Doug Roberts helped me greatly in many ways when he was a postdoc at UCSB and later. I would like to also thank David Lange, Ed Frank, Adil Hasan, Bill Lockman, Wouter Verkerke, Bob Jacobsen, Yury Kolomensky, Stephen Gowdy, and Dan Azzopardi (without whom I hardly would have kept sane during my first year at SLAC).

I am indebted to Gudrun Hiller, a SLAC theorist, who carefully read the theory chapter of this dissertation and made many excellent suggestions. Carsten Hast provided thoughtful comments on the detector chapter. David Lange read through the entire dissertation and helped improve both the content and the style and the grammar.

I would like to thank my Mother for offering unconditional support of all my educational endeavors. Yauncey and Juanita Newman, my dear Ohio friends, have been helping me throughout these years the way a loving family would.

Last but not least, I would like to thank Anton Eppich, my Santa Barbara officemate of two years, for many joyful hours of coffee drinking while discussing life in general. It was also a pleasure to share an office with Steve Levy at SLAC.

Curriculum Vitæ

Natalia Kuznetsova

Personal

Born	May 6, 1973 Minsk, Belarus

Education

1995 - 2001	Ph.D. Physics University of California, Santa Barbara
1993 - 1995	B.A. Physics *summa cum laude* with High Honors in Physics Kenyon College, Gambier, OH

Publications

2000	*Search for the Rare Decays* $B^+ \to K^+ \ell^+\ell^-$ *and* $B^0 \to K^{*0} \ell^+\ell^-$ *with the BABAR Detector*, contributed to the proceedings of DPF 2000, Columbus, Ohio, August 9 - 12, 2000; preprint hep-ex/0010071.

Abstract

SEARCH FOR EXCLUSIVE ELECTROWEAK PENGUIN DECAYS AT *BABAR*

by

Natalia Kuznetsova

This dissertation describes the search for the flavor-changing neutral current decays $B \to K\ell^+\ell^-$ and $B \to K^*(892)\ell^+\ell^-$, performed using a sample of $(22.7 \pm 0.4) \times 10^6$ $\Upsilon(4S) \to B\overline{B}$ decays collected with the *BABAR* detector at the PEP-II B Factory. The following final states have been reconstructed: $B^+ \to K^+\ell^+\ell^-$, $B^0 \to K^0\ell^+\ell^-$ ($K^0_s \to \pi^+\pi^-$), $B^+ \to K^{*+}\ell^+\ell^-$ ($K^{*+} \to K^0_s\pi^+$), and $B^0 \to K^{*0}\ell^+\ell^-$ ($K^{*0} \to K^+\pi^-$), where $\ell^+\ell^-$ is either an e^+e^- or $\mu^+\mu^-$ pair. The established 90% C.L. upper limits are:

$$\begin{aligned} \mathcal{B}(B \to K\ell^+\ell^-) &< 0.6 \times 10^{-6} \\ \mathcal{B}(B \to K^*\ell^+\ell^-) &< 2.5 \times 10^{-6}. \end{aligned}$$

These limits represent a significant improvement over previously published results and are close to Standard Model predictions.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The rare decays $B \to (K, K^*)\ell^+\ell^-$ are examples of flavor-changing neutral current (FCNC) processes that are highly suppressed in the Standard Model. They can only occur via higher-order processes, involving loops with flavor-changing $W^\pm$ vertices. In principle, new, heavy particles such as those predicted by the supersymmetry models, can enter the loops and lead to significant changes in both the decay rate and kinematic distributions. Therefore, these decay modes are very interesting from the point of view of searches for new physics.

The branching fractions for $B \to (K, K^*)\ell^+\ell^-$ are small in the Standard Model, of order 10^{-6}-10^{-7} for $B \to K\ell^+\ell^-$ and 10^{-6} for $B \to K^*\ \ell^+\ell^-$. Several experiments have searched for these modes. No evidence for signal has been observed. This document presents the results of search for $B \to (K, K^*)\ell^+\ell^-$ carried out with the *BABAR* detector, using the 1999-2000 data set consisting of 22.7×10^6 $B\overline{B}$ pairs. Due to the small statistics, a "blind" analysis has been performed to avoid any bias in the event selection. The modes considered include both all-charged track channels $B^+ \to K^+\ell^+\ell^-$ and $B^0 \to K^{*0}\ell^+\ell^-$ with $K^{*0} \to K^+\pi^-$, and three final states with a K^0_S or a π^0: $B^0 \to K^0_S\ell^+\ell^-$, $B^+ \to K^{*+}\ell^+\ell^-$ with $K^{*+} \to K^0_S\pi^+$, and $B^+ \to K^{*+}\ell^+\ell^-$ with $K^{*+} \to K^+\pi^0$. The $\ell^+\ell^-$ is either an e^+e^- or $\mu^+\mu^-$ pair, or an $e^\pm\mu^\mp$ combination.

1.1 Rare Decays of B Mesons

In order to put the $B \to (K, K^*)\ell^+\ell^-$ decays in prospective, it is useful to review other types of rare B decays. Examples, shown in Fig. 1.1, are:

- Hadronic decays involving a $b \to u$ transition. Such decays are suppressed with respect to $b \to c$ transitions by a factor of $|V_{ub}|/|V_{cb}|^2 \sim 0.01$ [1]. Figure 1.1a shows a $b \to u$ transition in the tree diagram for $\overline{B}^0 \to \pi^+\pi^-$.

- Gluonic penguins, involving an effective $b \to s$ or $b \to d$ transition which proceeds through a W exchange. Figure 1.1b shows the gluonic penguin diagram for $\overline{B}^0 \to K^-\pi^+$.

- Radiative penguin decays, which also involve an effective $b \to s$ transition, where one of the charged particles involved in the transition emits a photon. Figure 1.1c shows the diagram for a radiative penguin decay, $\overline{B}^0 \to K^*\gamma$.

The purely hadronic decays, $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$, have in fact both tree and penguin contributions. The tree diagram dominates the $B^0 \to \pi^+\pi^-$ rate, since the decay is $\sim \lambda^3$ (where $\lambda \approx 0.22$ is the sin of the Cabbibo angle), while the rate for the tree process in $B^0 \to K^+\pi^-$ is $\sim \lambda^4$. The gluonic penguin, on the other hand, dominates $B^0 \to K^+\pi^-$, since its contribution is $\sim \lambda^2$, whereas for $B^0 \to \pi^+\pi^-$ the penguin contribution is $\sim \lambda^3$. Both $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$ have been measured by CLEO [2], Belle [3], and *BABAR* [4]. The branching fractions from all three experiments are summarized in Table 1.1. There is a great deal of interest in these modes because of their relevance for studying direct CP-violation and measuring the CKM angle α [5].

Radiative penguin decays were first discovered by CLEO in 1993 ([10]; see [6] for the more recent results) with the measurement of the exclusive decay $B \to K^*\gamma$. This branching fraction has also been recently measured by both Belle [8] and *BABAR* [9]. In addition, CLEO [7], Belle [8], and ALEPH [11] have measured the inclusive rate $B \to X_s\gamma$. A large photon background from $\pi^0 \to \gamma\gamma$ decays

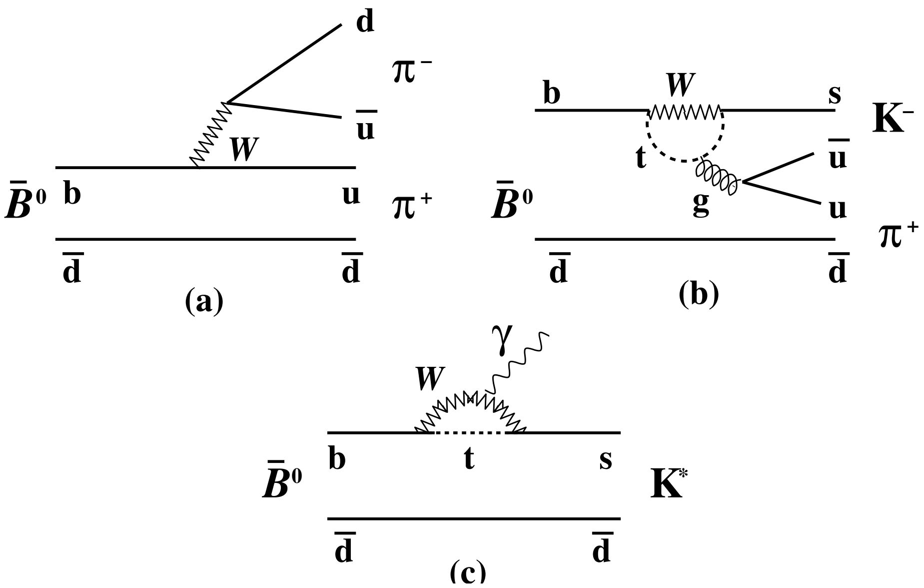


Figure 1.1: Feynman diagrams for various rare B decays. (a) The tree diagram for $\overline{B}^0 \rightarrow \pi^+\pi^-$. It is a doubly Cabbibo suppressed decay, involving, in particular, a $b \rightarrow u$ transition. (b) A gluonic penguin diagram for $\overline{B}^0 \rightarrow K^-\pi^+$. The W mediates the $b \rightarrow s$ transition. (c) The diagram for $\overline{B}^0 \rightarrow K^*\gamma$, which is an electromagnetic radiative penguin.

makes this a very difficult measurement. The measured branching fractions for both inclusive and exclusive $b \rightarrow s\gamma$ decays are listed in Table 1.1.

Summarizing, the existing experiments have reached the sensitivity to rare B decays with branching fractions $\mathcal{O}(10^{-4}\text{-}10^{-5})$, such as $B \rightarrow K\pi$, $B \rightarrow \pi\pi$, and $B \rightarrow K^*\gamma$. The next logical step is to explore processes with branching fractions of order 10^{-6}-10^{-7}, such as $B \rightarrow (K, K^*)\ell^+\ell^-$.

1.2 Plan of This Dissertation

The dissertation consists of the following sections:

- Chapter 2 is an overview of the mathematical formalism used to describe in-

Mode	$\mathcal{B}/10^{-5}$		
	CLEO	Belle	*BABAR*
$\pi^+\pi^-$	$(0.43^{+0.16}_{-0.14} \pm 0.05)$	$(0.59^{+0.24}_{-0.21} \pm 0.05$	$(0.41 \pm 0.10 \pm 0.07)$
$K^+\pi^-$	$(1.72^{+0.25}_{-0.24} \pm 0.12)$	$(1.87^{+0.33}_{-0.31} \pm 0.16)$	$(1.67 \pm 0.16^{+0.12}_{-0.17}$
$K^{*0}\gamma$	$(4.55^{+0.72}_{-0.68} \pm 0.34)$	$(4.96 \pm 0.67 \pm 0.45)$	$(4.39 \pm 0.41 \pm 0.27)$
$X_s\gamma$	$(31.5 \pm 3.5 \pm 2.9)$	$(33.7 \pm 5.3 \pm 4.2^{+5.0}_{-5.4})$	-

Table 1.1: The measured decay rates for some rare B decays from CLEO ([2], [6], [7]), Belle ([3], [8]) and *BABAR* ([4], [9]). The errors are statistical, systematic, and (where present) accounting for model dependence. K^{*0} denotes the $K^*(892)$ meson.

clusive and exclusive $B \to X_s \ell^+\ell^-$ decays. A general overview of the mathematical formalism commonly used to describe FCNC processes is given in Section 2.1. The formulas for the decay rates of $B \to (K, K^*)\ell^+\ell^-$ are derived in Section 2.2. The rates cannot be calculated exactly because of the presence of hadronic quark transitions, and are instead expressed in terms of form factors. There are a total of ten form factors relevant for the decays of interest: three in $B \to K\ell^+\ell^-$, and seven in $B \to K^*\ell^+\ell^-$. Not all of these form factors are independent due to the symmetry relations among them in the Large Energy Effective Limit, described in Section 2.3. Several models used for calculating the form factors, such as light cone QCD sum rules and quark models, are discussed in some detail in Section 2.5. The resulting predicted rates for the exclusive decays from some of these models are summarized in Section 2.5.5. Apart from the decay rates, the $B \to X_s\ell^+\ell^-$ decays present a number of other interesting quantities that in principle allow to discriminate between the Standard Model and certain new physics scenarios. One such variable is the forward-backward asymmetry in the $B \to K^*\ell^+\ell^-$ decays, described in Section 2.4. Possible modifications in the decay rates and the forward-backward asymmetry caused by new physics are discussed in Section 2.6.

- Chapter 3 summarizes the experimental status of the $B \to X_s \ell^+ \ell^-$ decays. The search for the inclusive decays carried out by CLEO is described in Section 3.1. The searches for the exclusive $B \to (K, K^*)\ell^+\ell^-$ decays are described in Section 3.2.1 (the CDF analysis), Section 3.2.2 (the CLEO analysis), and Section 3.2.3 (the Belle analysis). No signal has been observed yet by any of the experiments.

- Chapter 4 provides an overview of PEP-II and the *BABAR* detector. The Silicon Vertex Tracker (SVT) is described in Section 4.3 in detail. All other *BABAR* subdetectors are discussed briefly. Section 4.8 describes particle identification criteria for hadrons (Section 4.8.1) and leptons (Sections 4.8.2 and 4.8.3).

- A description of the data and Monte Carlo samples used in this analysis is given in Chapter 5.

- Chapter 6 describes the implementation of the model by Ali *et al.* [12] in the event generator for producing signal $B \to (K, K^*)\ell^+\ell^-$ Monte Carlo. The kinematic properties of the resulting signal events are discussed in detail in Section 6.2 using Dalitz plots and the distributions of various kinematic variables.

- Chapter 7 is a description of the event selection procedure. The motivations for performing this analysis as a blind search are summarized in Section 7.1. Since the selection of the event selection criteria is driven by what backgrounds must be suppressed, Section 7.2 gives an overview of the principal background categories. Section 7.3 is an overview of the three-stage event selection process. Section 7.11 addresses the issue of suppressing one of the most significant backgrounds for this analysis, the charmonium decays $B \to J/\psi(K, K^*)$ and $B \to \psi(2S)(K, K^*)$. Combinatorial backgrounds, arising from continuum *udsc* events and $B\overline{B}$ decays, are described in Sections 7.12 and 7.13, respectively. The procedure used to optimize the event selection criteria is described in Section 7.14. The effect of the event selection criteria on both signal and background is discussed in Section 7.16.

- Chapter 8 is devoted to the description of the various control samples. Because the search is performed as a blind analysis, the optimization of the event selection criteria must be performed using Monte Carlo simulated data; however, the efficiencies of the event selection criteria, as well as the consistency of various kinematic distributions, can be checked against control samples in data. Four samples are used in this analysis for the purpose: the three-body hadronic decay $B^+ \to \overline{D}^0\pi^+$ with $\overline{D}^0 \to K^+\pi^-$, described in Section 8.1; the charmonium sample $B \to J/\psi(K, K^*)$ and $B \to \psi(2S)(K, K^*)$, described in Section 8.2; $Ke\mu$ and $K^*e\mu$ combinations, described in Section 8.3; and the so-called "large sideband" – remote parts of the sideband, not used for either signal or background extraction, described in Section 8.4.

- Chapter 9 is a description of the extended unbinned maximum likelihood fit used to estimate the number of signal and background events in the signal region. The shapes used in the fit are described in Section 9.1. The validation of the fit yields is discussed in Section 9.3, and the results of fitting the $(K, K^*)e\mu$ sample in the data are discussed in Section 9.4.

- Chapter 10 is devoted to a discussion of the results of the analysis. Section 10.1 describes the procedure used to set 90% C.L. upper limits. Section 10.2 discusses the two-step procedure used to unblind the data and summarizes the 90% C.L. upper limits on the branching fractions of the individual modes. The limits on the combined modes are discussed in Section 10.3.

- Chapter 11 is a discussion of the systematic uncertainties.

- The future prospects for these decay modes are described in Chapter 12. There is little doubt that the signal will be observed by both the B factories and the Tevatron in the near future. On the other hand, a number of other interesting and important observables, such as the forward-backward asymmetry, are likely to be measured only after the advent of such experiments as LHCb [13] and BTeV [14].

Chapter 2

Theoretical predictions

There is a great deal of theoretical interest in the physics of effective flavor-changing neutral currents (FCNC) in B decays because they provide an excellent opportunity to test the Standard Model and thus potentially offer an insight into new physics. There are several variables of interest that can be measured in the $B \to X_s\, \ell^+\ell^-$ processes, all of which test the Standard Model:

- The decay rates of the exclusive channels $B \to (K, K^*)\ell^+\ell^-$ are likely to be the first experimentally measured quantities for these processes. The predicted rates for the exclusive modes suffer from large theoretical uncertainties in the description of the hadronic currents. Inclusive decays, on the other hand, are quite clean theoretically, but they are more difficult to measure experimentally.

- The Standard Model predicts that the forward-backward asymmetry A_{FB} in the $B \to K^*\ell^+\ell^-$ decays is sizable and has a well-defined - that is, largely independent of hadronic uncertainties - vanishing point. In some new physics scenarios, however, the zero is absent, and the behavior of the asymmetry itself (e.g., the overall sign) is quite different from the Standard Model predictions. This is discussed in more detail in Section 2.4.

- The forward-backward CP asymmetry FB_{CP} is defined as the asymmetry in A_{FB} for $B^0 \to K^*\ell^+\ell^-$ and its CP-conjugate channel. This quantity is expected to be very small in the Standard Model (below 10^{-3}), but it can be as large as 10% in some new physics models [15].

- The inclusive decay rates of the $B \to X_s\ell^+\ell^-$ and $B \to X_d\ell^+\ell^-$ processes can be used for extracting the ratio of the Cabibbo-Kabayashi-Maskawa (CKM) elements, $|V_{td}|/|V_{ts}|$. In the context of the Standard Model, the decay rates are dominated by loops with the heavy top quark. Therefore, in the effective $b \to s$ transitions, it is primarily three CKM elements – V_{tb}, V_{td}, and V_{ts} – that are involved in the process. Of these, only $|V_{tb}|$ has been measured directly at the Tevatron [16], yielding $|V_{tb}| = 0.99 \pm 0.15$. $|V_{td}|$ has so far been inferred from the measurements of Δm_d, the mass difference between B^0 and $\overline{B}^0$, although this method suffers from large theoretical uncertainties in the determination of the hadronic matrix element $f_{B_d}\sqrt{\mathcal{B}_{B_d}}$, where f_{B_d} is the leptonic decay constant and $\mathcal{B}_{B_d}$ is the so-called "B-parameter". Lattice QCD estimates for this quantity are 215 ± 35 MeV [17], leading to $0.007 < |V_{td}V_{tb}^*| < 0.010$. Measurements of the branching fractions for $b \to s\gamma$ processes have led to an inference of $|V_{ts}| = 0.033 \pm 0.007$ [18]. This exhausts the list of possible ways of measure the above mentioned CKM elements. FCNC processes can thus be of great help in over-constraining the CKM matrix. The possibility of measuring the ratio of the CKM elements $|V_{ts}|/|V_{td}|$ from the ratio of the inclusive rates, $B \to X_s\ell^+\ell^-$ and $B \to X_d\ell^+\ell^-$, has been theoretically studied in Refs. [19] and [20]. It was found that the ratio $\Delta\mathcal{R}$ of the branching fractions is stable, with a typical error of the order of a few percent.

The mathematical formalism most commonly used to describe FCNC processes is described in Section 2.1. It is applied to the exclusive decays $B \to K\ell^+\ell^-$ and $B \to K^*\ell^+\ell^-$ in Section 2.2. An overview of some of the form-factors models used to estimate hadronic matrix elements is given in Section 2.5. Possible alterations in the decay rates and in the behavior of the forward-backward asymmetry that could be caused by new physics are discussed in Section 2.6.

2.1 Methodology

Conventionally, FCNC processes are described using the operator product expansion (OPE) formalism. In this approach, the most general effective Hamiltonian for weak decays of hadrons is represented by [21]

$$\mathcal{H}_{\rm eff} = -4\frac{G_F}{\sqrt{2}}\sum_i V^i_{CKM}\, \mathcal{C}_i(\mu)\mathcal{O}_i, \tag{2.1.1}$$

where G_F is the Fermi constant, V^i_{CKM} is a combination of the relevant CKM matrix elements, $\mathcal{O}_i$ are the local operators relevant for the process in question, and $\mathcal{C}_i$ are the Wilson coefficients. The idea is to represent an interaction as a series expansion of local operators weighted by the Wilson coefficients, which can be thought of as effective "coupling constants" at a given scale μ, associated with "vertices" $\mathcal{O}_i$. The Wilson coefficients describe the features of the theory at scales higher than μ, and due to asymptotic freedom of QCD can be calculated perturbatively for a large enough μ. The operators, on the other hand, reflect physics contributions at scales lower than μ. The scale can therefore be viewed as separating short- and long-distance contributions, the former associated with the Wilson coefficients and the latter, with the operator matrix elements. μ is typically first chosen to be $\mathcal{O}(M_W)$, and subsequently evolved down to lower values (e.g., of the order of the b-quark mass for B decays) by renormalization group equations, integrating out heavy degrees of freedom. Since the amplitude cannot depend of the scale μ, the μ-dependence of the coefficients must eventually cancel that of the matrix elements. In a similar way, the dependence on the renormalization scheme must also cancel out.

For the FCNC processes in question, the operator basis consists of the following operators (assuming a massless s quark):

- **Current-current operators** (Fig. 2.1a)

$$\mathcal{O}_1 = (\overline{s}_\alpha\gamma^\mu b_{L\beta})(\overline{c}_\beta\gamma_\mu c_{L\alpha})$$

$$\mathcal{O}_2 = (\overline{s}_\alpha \gamma^\mu b_{L\alpha})(\overline{c}_\beta \gamma_\mu c_{L\beta}) \tag{2.1.2}$$

- **QCD penguin operators** (Fig. 2.1b)

$$\begin{aligned}
\mathcal{O}_3 &= (\overline{s}_\alpha \gamma^\mu b_{L\alpha}) \sum_{q=u,d,s,c,b} (\overline{q}_\beta \gamma_\mu q_{L\beta}) \\
\mathcal{O}_4 &= (\overline{s}_\alpha \gamma^\mu b_{L\beta}) \sum_{q=u,d,s,c,b} (\overline{q}_\beta \gamma_\mu q_{L\alpha}) \\
\mathcal{O}_5 &= (\overline{s}_\alpha \gamma^\mu b_{L\alpha}) \sum_{q=u,d,s,c,b} (\overline{q}_\beta \gamma_\mu q_{R\beta}) \\
\mathcal{O}_6 &= (\overline{s}_\alpha \gamma^\mu b_{L\beta}) \sum_{q=u,d,s,c,b} (\overline{q}_\beta \gamma_\mu q_{R\alpha})
\end{aligned} \tag{2.1.3}$$

- **Electroweak penguin operators** (Fig. 2.1c, Fig. 2.1d, Fig. 2.1e, Fig. 2.1f)

$$\begin{aligned}
\mathcal{O}_7 &= \frac{e}{16\pi^2} \overline{s}_\alpha \sigma_{\mu\nu} (m_b b_{R\alpha} F^{\mu\nu}) \\
\mathcal{O}_8 &= \frac{g_s}{16\pi^2} \overline{s}_\alpha T^a_{\alpha\beta} \sigma_{\mu\nu} (m_b b_{R\beta} G^{a\mu\nu}) \\
\mathcal{O}_9 &= \frac{e^2}{16\pi^2} (\overline{s}_\alpha \gamma^\mu b_{L\alpha})(\overline{\ell} \gamma_\mu \ell) \\
\mathcal{O}_{10} &= \frac{e^2}{16\pi^2} (\overline{s}_\alpha \gamma^\mu b_{L\alpha})(\overline{\ell} \gamma_\mu \gamma_5 \ell),
\end{aligned} \tag{2.1.4}$$

where α and β are color indices and summation over repeated indices is implied. $F_{\mu\nu}$ and $G^a_{\mu\nu}$ are the QED and QCD field strength tensors, respectively; e and g_s are electromagnetic and strong coupling constants, respectively; and $q_{L,R} \equiv [(1 \mp \gamma_5)/2]q$. Note that even though the diagrams in Fig. 2.1 show the W propagator, in the effective theory the W field is integrated out.

For an exclusive decay involving a transition from meson M to meson F, the Hamiltonian 2.1.1 leads to the following decay amplitude:

$$\mathcal{A}(M \to F) \equiv \langle F | \mathcal{H}_{\text{eff}} | M \rangle = -4 \frac{G_F}{\sqrt{2}} \sum_i V^i_{CKM}\ \mathcal{C}_i(\mu) \langle F | \mathcal{O}_i | M \rangle. \tag{2.1.5}$$

The evaluation of the amplitude 2.1.5 includes a perturbative calculation of the Wilson coefficients $\mathcal{C}_i(\mu)$ and a non-perturbative calculation of the hadronic matrix

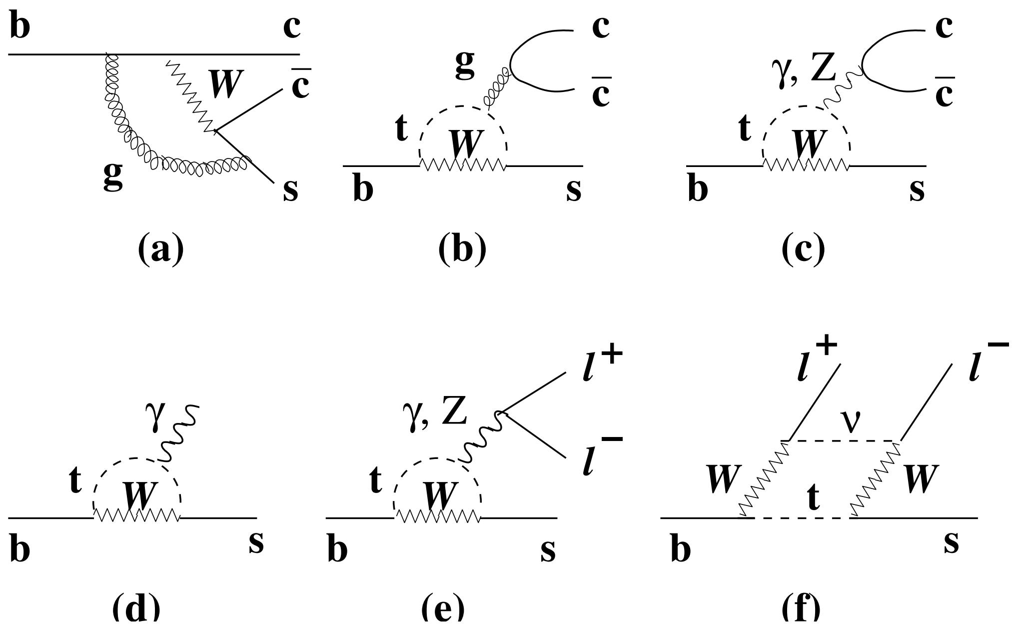


Figure 2.1: Penguin and box diagrams for the weak decay operators $\mathcal{O}_i$ relevant for the FCNC processes in question. Diagram (a) shows current-current interactions. Such interactions involve both short- and long- distance contributions, the latter originating from the gluon exchange shown schematically on the diagram as a single gluon. Diagram (b) shows a typical diagram for a four-fermion operator (a QCD penguin). Diagrams (c), (d), (e), and (f) are examples of electroweak penguins. Diagram (d) is sometimes referred to as a magnetic penguin (the γ is on-shell), and diagram (e), as a semileptonic penguin. Diagram (f) is a box diagram. Even though all of the diagrams show the W field, in the effective theory it is integrated out.

elements $\langle\mathcal{O}_i(\mu)\rangle$. The latter can be done using a number of different techniques, such as lattice calculations, QCD sum rules, chiral perturbation theory, and others. Because these techniques necessarily involve simplifying assumptions, the matrix elements $\langle\mathcal{O}_i(\mu)\rangle$ constitute the main source of theoretical uncertainty. However, inclusive decays of heavy mesons, such as the B meson, turn out to be easier to calculate. In this case, the amplitude is

$$\mathcal{A}(B \to X) = -4\frac{G_F}{\sqrt{2}} \sum_{f\,\mathrm{inX}} \sum_i V^i_{CKM}\ \mathcal{C}_i(\mu)\langle f|\mathcal{O}_i|B\rangle, \qquad (2.1.6)$$

where the sum is over all possible final states f. In the Heavy Quark Expansion formalism ([22], [23], [24]), the branching fraction can be expanded in inverse powers of the b quark mass, m_b:

$$\mathcal{B}(B \to X) = \mathcal{B}(b \to q) + \mathcal{O}\left(\frac{1}{m_b^2}\right). \qquad (2.1.7)$$

To first order, the decay is described by the decay of the b quark, which can be calculated perturbatively. The remaining terms in the expansion 2.1.7 turn out to affect the branching ratio by less than 10% [25].

2.2 Decay Rates for $B \to (K, K^*)\ \ell^+\ell^-$

At the quark level, the transition $b \to s\ell^+\ell^-$ is described by the Hamiltonian 2.1.1:

$$\mathcal{H}_{\mathrm{eff}} = -4\frac{G_F}{\sqrt{2}} V_{ts}^* V_{tb} \sum_{i=1}^{10} \mathcal{C}_i(\mu)\mathcal{O}_i, \qquad (2.2.1)$$

where the operators $\mathcal{O}_i$ were discussed in Section 2.1. This Hamiltonian leads to the following decay amplitude:

$$\begin{aligned}\mathcal{A}(b \to s\ell^+\ell^-) &= \frac{G_F\ \alpha}{\sqrt{2}\pi} V_{ts}^* V_{tb} \Big[\mathcal{C}_9^{\mathrm{eff}}(\hat{s})[\overline{s}\gamma_\mu b_L][\overline{\ell}\gamma^\mu \ell]\ +\ \mathcal{C}_{10}(\hat{s})[\overline{s}\gamma_\mu b_L][\overline{\ell}\gamma^\mu\gamma_5\ell] \\ &- 2\mathcal{C}_7^{\mathrm{eff}}(\hat{s})\hat{m}_b \left[\overline{s} i\sigma_{\mu\nu}\frac{\hat{q}^\nu}{\hat{s}} b_R\right][\overline{\ell}\gamma^\mu\ell]\Big], \qquad (2.2.2)\end{aligned}$$

where α is the fine structure constant; $s \equiv q^2 = (p_{l^+} + p_{l^-})^2$ is the squared sum of the leptons' four-momenta; $\hat{s} \equiv s/m_B^2$ is its normalization in terms of the B meson mass, m_B; and $\hat{m}_b \equiv m_b/m_B$ is the normalized mass of the b quark, evaluated at the scale μ. The mass of the strange quark is assumed to be zero, but the leptons are kept massive.

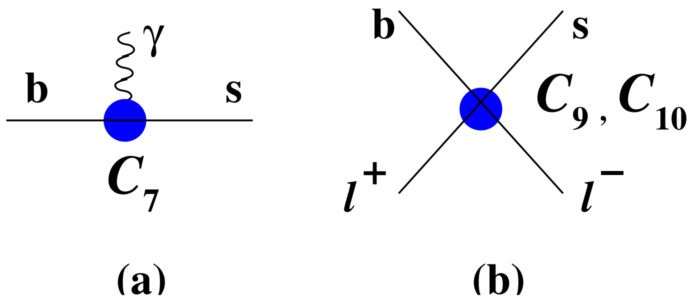


Figure 2.2: Effective vertex diagrams for (a) $\mathcal{C}_7$, and (b) $\mathcal{C}_9$, $\mathcal{C}_{10}$. $\mathcal{C}_7$ is associated with the γ penguin, and $\mathcal{C}_{10}$, with the Z penguin. Both γ and Z penguins contribute to $\mathcal{C}_9$. The contribution from the box diagram is also taken into account in $\mathcal{C}_9$ and $\mathcal{C}_{10}$.

Three Wilson coefficients, $\mathcal{C}_7$, $\mathcal{C}_9$, and $\mathcal{C}_{10}$, enter amplitude 2.2.2. Qualitatively, their contributions can be understood as follows. $\mathcal{C}_7$ is a "coupling constant" associated with the γ penguin shown in Fig. 2.1d, and $\mathcal{C}_9$ and $\mathcal{C}_{10}$ are associated with the semileptonic penguin depicted in Fig. 2.1e, as well as with the box diagram of Fig. 2.1f. This is schematically represented in Fig. 2.2, which shows effective vertex diagrams associated with the different Wilson coefficients. The coefficients $\mathcal{C}_9$ and $\mathcal{C}_{10}$ are chosen in such a way that $\mathcal{C}_{10}$ accounts mostly for the axial vector coupling. It can thus be referred to as the Z penguin coefficient. On the other hand, $\mathcal{C}_9$ accounts mainly for the vector coupling. The coefficient $\mathcal{C}_7$ in Eqn. 2.2.2 is labeled "effective", following the notation conventions of Ref. [26]. The other effective coefficient in Eqn. 2.2.2, $\mathcal{C}_9$, is defined in such a way as to absorb the matrix elements of the four-quark operators ($\langle \ell^+\ell^- s|\mathcal{O}_i|b\rangle$, $1 \le i \le 6$), examples of which are shown in Fig. 2.1a, Fig. 2.1b, and Fig. 2.1c.

$$\mathcal{C}_9^{\text{eff}}(\hat{s}) = \mathcal{C}_9 + Y(\hat{s}), \tag{2.2.3}$$

where $Y(\hat{s})$ incorporates the four-quark operator matrix elements. Its explicit form can be found in Ref. [27]. While light quarks do not contribute significantly to $Y(\hat{s})$, the charm loops – and in particular $c\overline{c}$ resonances – prevent perturbative calculations.

These resonances can be introduced into the theory by adding a phenomenological Breit-Wigner anzatz [28] to the perturbative part

$$Y(\hat{s}) = Y_{\text{pert}}(\hat{s}) + \frac{3\pi}{\alpha^2}\mathcal{C}^{(0)} \sum_{V_i = J/\psi, \psi', \cdots} k_i \frac{\Gamma(V_i \to \ell^+\ell^-) m_{V_i}}{m_{V_i}^2 - \hat{s} m_B^2 - i m_{V_i} \Gamma_{V_i}} \tag{2.2.4}$$

where $Y_{\text{pert}}(\hat{s})$ is the perturbative contribution, $C^{(0)}$ is a function of the four-quark Wilson coefficients $\mathcal{C}_i$ $(1 \leq i \leq 6)$, and k_i are phenomenological factors of order 2 [29]. Introducing such an anzatz naturally raises the question of the applicability of the parametrization away from the resonances. One approach, suggested by Ali, Mannel and Morozumi (AMM) [30], extrapolates the Breit-Wigner shape to these regions. Another approach, introduced by Krüger and Sehgal (KS) [31], uses the measured cross-section $\sigma(e^+e^- \to \text{hadrons})$ for the purpose, but perturbative contributions to this process and to $B \to X_s \ell^+\ell^-$ are not the same. There is at present no unique solution to the problem; the difference between these two approaches is included in the overall theoretical error.

The contributions of $\mathcal{C}_7^{\text{eff}}$, $\mathcal{C}_9^{\text{eff}}$, $\mathcal{C}_{10}$, and the interference between $\mathcal{C}_7^{\text{eff}}$ and $\mathcal{C}_9^{\text{eff}}$, to the inclusive decay rate R are shown in Fig. 2.3. It is apparent that the contribution from $\mathcal{C}_7^{\text{eff}}$, denoted R_7 in Fig. 2.3, is not by itself significant; however, the interference term $R_{7/9}$ (which is negative in the Standard Model) is quite important. The main contributions to the rate come from R_9 and R_{10}, and the two are roughly of the same size.

The matrix elements of the quark operators in Eqn. 2.2.2 cannot be calculated perturbatively, and must therefore be parametrized in terms of form factors. For the pseudoscalar-to-pseudoscalar meson transition $B \to K$, the three form factors $f_+(s)$, $f_0(s)$, and $f_T(s)$ are defined as follows: [12]

$$\begin{aligned}
\langle K(p)|\bar{s}\gamma_\mu b|B(p_B)\rangle &= f_+(s)\left\{(p_B+p)_\mu - \frac{m_B^2 - m_K^2}{s} q_\mu\right\} \\
&+ \frac{m_B^2 - m_K^2}{s} f_0(s) q_\mu \\
\langle K(p)|\bar{s}\sigma_{\mu\nu}q^\nu b|B(p_B)\rangle &= i\left\{(p_B+p)_\mu s - q_\mu(m_B^2 - m_K^2)\right\} \frac{f_T(s)}{m_B + m_K},
\end{aligned}$$

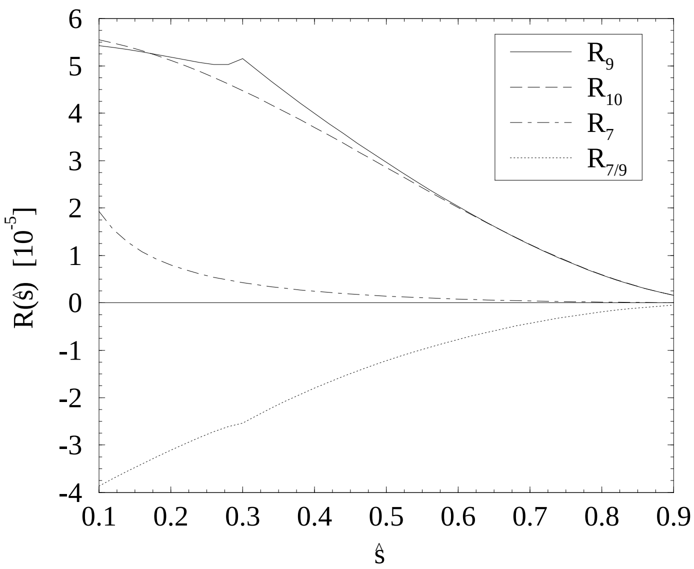


Figure 2.3: Top to bottom: the contributions to the inclusive $b \to s$ decay rate R from $\mathcal{C}_9^{\rm eff}$ (R_9), $\mathcal{C}_{10}$ (R_{10}), $\mathcal{C}_7^{\rm eff}$ (R_7), and the interference between $\mathcal{C}_7^{\rm eff}$ and $\mathcal{C}_9^{\rm eff}$ ($R_{7/9}$). Figure taken from [21].

(2.2.5)

where $q \equiv p_{l^+} + p_{l^-} = p_B - p_K$.

For the pseudoscalar-to-vector transition, $B \to K^*$, there are a total of seven form factors: four semileptonic and three penguin ones. The semileptonic form factors $V(s)$, $A_0(s)$, $A_1(s)$, and $A_2(s)$, are defined as:

$$\begin{aligned}\langle K^*(p)|(V-A)_\mu|B(p_B)\rangle &= i\epsilon^*_\mu(m_B+m_{K^*})A_1(s)\\ &+ i(p_B+p)_\mu(\epsilon^* p_B)\frac{A_2(s)}{m_B+m_{K^*}}\\ &+ iq_\mu(\epsilon^*_\mu p_B)\frac{2m_{K^*}}{s}(A_3(s)-A_0(s))\\ &+ \epsilon_{\mu\nu\rho\sigma}\epsilon^{*\nu}p_B^\rho p^\sigma\frac{2V(s)}{m_B+m_{K^*}},\end{aligned} \tag{2.2.6}$$

where

$$A_3(s) \equiv \frac{m_B + m_{K^*}}{2m_{K^*}} A_1(s) - \frac{m_B - m_{K^*}}{2m_{K^*}} A_2(s). \tag{2.2.7}$$

The penguin form factors $T_1(s)$, $T_2(s)$, and $T_3(s)$, are given by:

$$\begin{aligned} \langle K^*(p)|\overline{s}\sigma_{\mu\nu}q^\nu(1+\gamma_5)b|B(p_B)\rangle &= i\epsilon_{\mu\nu\rho\sigma}\epsilon^{*\nu}p_B^\rho p^\sigma 2T_1(s) \\ &+ T_2(s)\left\{\epsilon^*_\mu(m_B^2 - m_{K^*}^2) - (\epsilon^* p_B)(p_B+p)_\mu\right\} \\ &+ T_3(s)(\epsilon^* p_B)\left\{q_\mu - \frac{s}{m_B^2 - m_{K^*}^2}(p_B+p)_\mu\right\}, \end{aligned} \tag{2.2.8}$$

where ϵ^*_μ is the K^* polarization vector.

With these form factors, the amplitude 2.2.2 can be expressed as follows:

$$\mathcal{A} = \frac{G_F \alpha}{2\sqrt{2}\pi} V_{ts}^* V_{tb} m_B \left[\mathcal{T}^1_\mu \left(\overline{\ell}\,\gamma^\mu\,\ell\right) + \mathcal{T}^2_\mu \left(\overline{\ell}\,\gamma^\mu\,\gamma_5\,\ell\right)\right], \tag{2.2.9}$$

where

$$\begin{aligned} \mathcal{T}^1_\mu &= A'(\hat{s})\,\hat{p}_\mu + B'(\hat{s})\,\hat{q}_\mu \\ \mathcal{T}^2_\mu &= C'(\hat{s})\,\hat{p}_\mu + D'(\hat{s})\,\hat{q}_\mu, \end{aligned} \tag{2.2.10}$$

for $B \to K\ell^+\ell^-$; and for $B \to K^*\ell^+\ell^-$

$$\begin{aligned} \mathcal{T}^1_\mu &= A(\hat{s})\,\epsilon_{\mu\rho\alpha\beta}\epsilon^{*\rho}\hat{p}_B^\alpha\hat{p}_{K^*}^\beta - iB(\hat{s})\,\epsilon^*{}_\mu + iC(\hat{s})\,(\epsilon^*\cdot\hat{p}_B)\hat{p}_\mu + iD(\hat{s})\,(\epsilon^*\cdot\hat{p}_B)\hat{q}_\mu \\ \mathcal{T}^2_\mu &= E(\hat{s})\,\epsilon_{\mu\rho\alpha\beta}\epsilon^{*\rho}\hat{p}_B^\alpha\hat{p}_{K^*}^\beta - iF(\hat{s})\,\epsilon^*{}_\mu + iG(\hat{s})\,(\epsilon^*\cdot\hat{p}_B)\hat{p}_\mu + iH(\hat{s})\,(\epsilon^*\cdot\hat{p}_B)\hat{q}_\mu, \end{aligned} \tag{2.2.11}$$

where $p \equiv p_B + p_{(K,K^*)}$ and the hat over variables denotes the normalization in terms of m_B.

The functions used in Eqn. 2.2.10 and 2.2.11 are:

$$\begin{aligned} A'(\hat{s}) &= C_9^{\text{eff}}(\hat{s})\,f_+(\hat{s}) + \frac{2\hat{m}_b}{1+\hat{m}_K}C_7^{\text{eff}}\,f_T(\hat{s}) \\ B'(\hat{s}) &= C_9^{\text{eff}}(\hat{s})\,f_-(\hat{s}) - \frac{2\hat{m}_b}{\hat{s}}(1-\hat{m}_K)C_7^{\text{eff}}\,f_T(\hat{s}) \end{aligned}$$

$$
\begin{aligned}
C'(\hat{s}) &= C_{10}\, f_+(\hat{s}) \\
D'(\hat{s}) &= C_{10}\, f_-(\hat{s}) \\
A(\hat{s}) &= \frac{2}{1+\hat{m}_{K^*}} C_9^{\text{eff}}(\hat{s}) V(\hat{s}) + \frac{4\hat{m}_b}{\hat{s}} C_7^{\text{eff}} T_1(\hat{s}) \\
B(\hat{s}) &= (1+\hat{m}_{K^*}) \left[C_9^{\text{eff}}(\hat{s}) A_1(\hat{s}) + \frac{2\hat{m}_b}{\hat{s}} (1-\hat{m}_{K^*}) C_7^{\text{eff}} T_2(\hat{s}) \right] \\
C(\hat{s}) &= \frac{1}{1-\hat{m}_{K^*}^2} \left[(1-\hat{m}_{K^*}) C_9^{\text{eff}}(\hat{s}) A_2(\hat{s}) + 2\hat{m}_b C_7^{\text{eff}} \left(T_3(\hat{s}) + \frac{1-\hat{m}_{K^*}^2}{\hat{s}} T_2(\hat{s}) \right) \right] \\
D(\hat{s}) &= \frac{1}{\hat{s}} \Big[C_9^{\text{eff}}(\hat{s}) \left((1+\hat{m}_{K^*}) A_1(\hat{s}) - (1-\hat{m}_{K^*}) A_2(\hat{s}) - 2\hat{m}_{K^*} A_0(\hat{s}) \right) \\
&\qquad - 2\hat{m}_b C_7^{\text{eff}} T_3(\hat{s}) \Big] \\
E(\hat{s}) &= \frac{2}{1+\hat{m}_{K^*}} C_{10} V(\hat{s}) \\
F(\hat{s}) &= (1+\hat{m}_{K^*}) C_{10} A_1(\hat{s}) \\
G(\hat{s}) &= \frac{1}{1+\hat{m}_{K^*}} C_{10} A_2(\hat{s}) \\
H(\hat{s}) &= \frac{1}{\hat{s}} C_{10} \left[(1+\hat{m}_{K^*}) A_1(\hat{s}) - (1-\hat{m}_{K^*}) A_2(\hat{s}) - 2\hat{m}_{K^*} A_0(\hat{s}) \right] .
\end{aligned}
\tag{2.2.12}
$$

A number of techniques used to approximately calculate the form factors that appear in Eqn. 2.2.12 are briefly discussed below. Before describing them, however, it is worth pointing out that the different form factors are not all independent.

2.3 Symmetry Relations Among Form Factors

The success of the Heavy Quark Effective Theory (HQET) [32] in the description of heavy-to-heavy quark transition, such as $B \to (D, D^*)\ell\nu$, prompted many theorists to consider the application of the HQET ideas to heavy-to-light transitions. In its original form, the constraints imposed by HQET on such transitions are quite weak. However, in case of a large energy transfer E from the parent to the daughter hadron, one can make use of the Large Energy Effective Theory (LEET) [33]. Formally, the criteria for the applicability of LEET are:

$$(\Lambda_{\text{QCD}}, m') \;<<\; (M, E). \tag{2.3.1}$$

That is, the mass of the heavy parent hadron M should be large comparing to the characteristic QCD scale $\Lambda_{\rm QCD}$; and the energy E of the resulting daughter hadron in the rest frame of its parent should be large comparing to its mass m'. It is thus apparent that LEET will not be valid for large q^2, near the zero-recoil point.

The idea is that if the light daughter hadron receives a large energy transfer from the decaying parent, it may be assumed that the fast moving quark (the s quark for the $B \to K$ transition) carries most of the momentum of the resulting hadron. The other partons (the spectator quark and the gluons) are assumed to be soft, just as in the case of heavy-to-heavy transitions in HQET. These assumptions lead to a significant simplification of the formalism: it can be shown that to leading order in $1/M$, $1/E$, and α_s, the heavy-to-light matrix elements are functions of only three dimensionless form factors ξ, $\xi_{||}$, and $\xi_{\perp}$ [34], where the subscripts in the latter two form factors refer to the longitudinal and transverse polarizations of the resulting vector meson, respectively.

$$
\begin{aligned}
f_+(q^2) &= \xi(M, E_P) \\
f_0(q^2) &= \left(1 - \frac{q^2}{M^2 - m_P^2}\right) \xi(M, E_P) \\
f_T(q^2) &= \left(1 + \frac{m_P}{M}\right) \xi(M, E_P) \\
A_0(q^2) &= \left(1 - \frac{m_V^2}{M E_V}\right) \xi_{||}(M, E_V) + \frac{m_V}{M} \xi_{\perp}(M, E_V) \\
A_1(q^2) &= \frac{2E_V}{M + m_V} \xi_{\perp}(M, E_V) \\
A_2(q^2) &= \left(1 + \frac{m_V}{M}\right) \left[\xi_{\perp}(M, E_V) - \frac{m_V}{E_V} \xi_{||}(M, E_V)\right] \\
V(q^2) &= \left(1 + \frac{m_V}{M}\right) \xi_{\perp}(M, E_V) \\
T_1(q^2) &= \xi_{\perp}(M, E_V) \\
T_2(q^2) &= \left(1 - \frac{q^2}{M^2 - m_V^2}\right) \xi_{\perp}(M, E_V)
\end{aligned}
$$

$$T_3(q^2) \;=\; \xi_\perp(M, E_V) - \frac{m_V}{E}\left(1 - \frac{m_V^2}{M^2}\right)\xi_{||}(M, E_V), \tag{2.3.2}$$

where m_P, E_P, m_V, and E_V are the masses and energies of the pseudoscalar and vector daughter mesons, respectively. The leading corrections to the form factor ratios are of order $\mathcal{O}(\Lambda_{\rm QCD}/E_V)$, and are therefore small at low q^2, less than 10%. Symmetry-breaking effects from the exchange of hard gluons have also been shown to affect the symmetries by less than 10% [35]. Moreover, two form factor ratios do not receive any corrections of order α_s:

$$\begin{aligned} \frac{V(q^2)}{A_1(q^2)} &= \frac{(M+m_V)^2}{2E_V M} \\ \frac{T_2(q^2)}{T_1(q^2)} &= \left(1 - \frac{q^2}{M^2 - m_V^2}\right) \end{aligned} \tag{2.3.3}$$

This can be intuitively understood as follows [36]. The ratio of form factors is related to the helicity of the final state, just as it is for semileptonic decays ([37], [38]). When an infinitely heavy quark decays into a light quark, and the latter couples with the spectator quark to form a spin 1 bound state, there is a particular dominant helicity amplitude for the process, the "-" helicity. Hard gluons would have to be exchanged in order to form a "+" helicity state (which would also lead to α_s corrections), but the amplitude for that is suppressed by $\mathcal{O}(1/E_V)$, and is therefore small in LEET.

2.4 Forward-Backward Asymmetry

The symmetry relations between form factors discussed in Section 2.3 have important implications for some kinematic quantities that can be measured in the $B \to K^*\ell^+\ell^-$ decays, as we will now show. Because of the presence of non-perturbative form factors, the predictions for decay rates are model-dependent, with theoretical uncertainties of order 30%. However, for the $B \to K^*\ell^+\ell^-$ decays there exists a quantity called the forward-backward asymmetry A_{FB}, which can be computed with an error of $\sim 10\%$. It is defined in the following way:

$$A_{FB}(\hat{s}) = \frac{1}{d\Gamma(B \to K^*\ell^+\ell^-)/d\hat{s}} \int_{-1}^{1} d\cos\theta \frac{d^2\Gamma(B \to K^*\ell^+\ell^-)}{d\hat{s}d\cos\theta} \mathrm{sign}(\cos\theta), \quad (2.4.1)$$

where θ is the angle between the momentum of the B and the positively charged lepton in the dilepton center of mass frame (Fig. 2.4). The forward-backward asymmetry is zero for $B \to K\ell^+\ell^-$, but it is non-trivial for $B \to K^*\ell^+\ell^-$. Particularly interesting is the point $\hat{s}_0$ at which the asymmetry vanishes for this decay [12]

$$Re(\mathcal{C}_9^{\mathrm{eff}}(\hat{s}_0)) = -\frac{\hat{m}_b}{\hat{s}_0}\mathcal{C}_7^{\mathrm{eff}} \left\{ \frac{T_2(\hat{s}_0)}{A_1(\hat{s}_0)}(1 - \hat{m}_{K^*}) + \frac{T_1(\hat{s}_0)}{V(\hat{s}_0)}(1 + \hat{m}_{K^*}) \right\} \quad (2.4.2)$$

In order for Eqn. 2.4.2 to be satisfied, the following condition must be met:

$$\mathrm{sign}(\mathcal{C}_7^{\mathrm{eff}} Re(\mathcal{C}_9^{\mathrm{eff}})) = -1. \quad (2.4.3)$$

This condition is a signature of a Standard Model-like scenario. It is not satisfied in some new physics models. For example, in SUGRA models [39] with large $\tan\beta$, $\mathrm{sign}(\mathcal{C}_7^{\mathrm{eff}} Re(\mathcal{C}_9^{\mathrm{eff}})) = +1$, and there is no zero below the J/ψ resonance. Therefore, the vanishing point of the forward-backward asymmetry is a very good discriminating variable allowing to distinguish the Standard Model from new physics scenarios. Moreover, it is predicted to a very good precision. For instance, the value of s_0 in the model of Ref. [12] is

$$s_0 = 2.88^{+0.44}_{-0.28} \ \ (\mathrm{GeV}/c^2)^2, \quad (2.4.4)$$

known to about 10%. The relative model independence of this quantity can be established by examining the predictions from a number of different models, but it can also be put on a somewhat firmer ground by using the LEET limit, discussed in Section 2.3. Using Eqns. 2.3.3 and 2.4.2:

$$Re(\mathcal{C}_9^{\mathrm{eff}}(\hat{s}_0)) = -2\frac{\hat{m}_b}{\hat{s}_0}\mathcal{C}_7^{\mathrm{eff}} \frac{1 - \hat{s}_0}{1 + m_{K^*}^2 - \hat{s}_0} \quad (2.4.5)$$

Therefore, the precision with which this point is determined is limited only by the precision with which two quantities are known: $\hat{m}_b$ and the ratio of the effective

Wilson coefficients, $\mathcal{C}_7^{\text{eff}}$ and $\mathcal{C}_9^{\text{eff}}$. Equation 2.4.4 shows that the LEET limit is expected to work well because the zero of the asymmetry is located in the low q^2 region.

The forward-backward asymmetry, shown in Fig. 2.6, exhibits a q^2 dependence which can be qualitatively understood as follows. For low q^2, where the dominant interaction is mediated by photon exchange, the asymmetry is expected to be small. For large q^2, on the other hand, Z and W exchange dominates and the interaction has a $V-A$ parity-violating structure, leading to a large asymmetry. Incidentally, this qualitative picture can also help determine the sign of the asymmetry. In $B^0 \to K^*\ell^+\ell^-$ decay at large q^2, the recoiling $\overline{s}$ quark is predominantly helicity $\lambda = +1/2$. It can combine with the spectator quark to make a daughter meson with helicity $\lambda = 0$ or $+1$. For K^*, the vector meson must be predominantly helicity $\lambda = +1$. Therefore, the positively charged lepton must travel mostly in the direction opposite to that of the K^*, leading to a negative $\cos\theta_l$ and a negative asymmetry A_{FB}. Figure 2.4 illustrates the kinematic configuration. It is also interesting to note

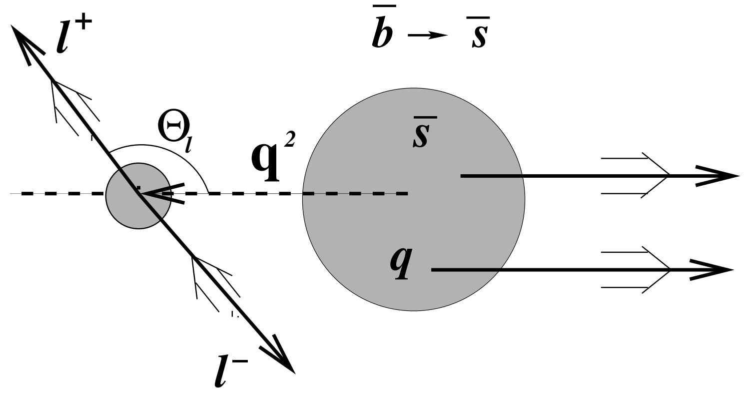


Figure 2.4: The kinematic configuration for the $B^0 \to K^*\ell^+\ell^-$ decay at large q^2.

that the configuration shown in Fig. 2.4 will lead to harder positrons than electrons. This means that the forward-backward asymmetry is in fact equivalent to the energy asymmetry:

$$A_{FB} = \frac{N(E_{\ell^+} > E_{\ell^-}) - N(E_{\ell^-} > E_{\ell^+})}{N(E_{\ell^+} > E_{\ell^-}) + N(E_{\ell^-} > E_{\ell^+})} \tag{2.4.6}$$

2.5 Overview of Form Factor Models

This section presents a brief review of some of the models used to calculate hadronic matrix elements defined in Eqn. 2.2.2. The models include lattice QCD, QCD sum rules and light-cone QCD sum rules, and quark models.

2.5.1 Lattice QCD

Lattice QCD is one of the most well-known methods for calculating form factors for hadronic transitions which involve non-perturbative QCD processes. The main limitation of this approach is the effective upper momentum cut-off arising from the finite lattice size. One is forced to extrapolate the predictions obtained at large q^2, where lattice predictions are reliable, to the entire physical range by using a properly chosen vacuum correlation function. An added difficulty is that the b quark cannot be simulated on the lattice directly because its mass is greater than the inverse lattice spacing ($\sim$ 2.73 GeV). One way of dealing with this problem is to perform calculations near the charm mass scale, and then use heavy quark symmetry [32] to extrapolate the results to the b-quark scale.

The APE and UKQCD Collaborations [40] have performed calculations of form factors for semileptonic decays of B and D mesons, as well as for $B \to K^*\gamma$ decay.

2.5.2 QCD Sum Rules

In the method of QCD sum rules ([41], [42]), hadrons are represented by their quark currents and interactions are introduced via a correlation function that is treated in the framework of the operator product expansion (OPE) formalism. The

OPE was briefly described in Section 2.1. The results of the perturbative calculation of the short-distance quark-gluon interactions are matched to a sum over hadronic states via a dispersion relation. The hadronic dispersion integrals are typically handled using the quark-hadron duality approximation, which is the assumption that the integrated physical hadronic and perturbative spectral densities produce the same results. One of the limitations of this method is that it can only be used for calculating the form factors in a well-defined region of q^2: very large values of q^2 lead to singularities in the correlation function. The method is complementary to lattice QCD in this respect. In practice, the form factors can be extrapolated to $q^2_{\rm max}$ using a proper fitting procedure.

An example of using QCD sum rules for calculating the branching fractions of $B \to (K, K^*)\ell^+\ell^-$ is given in Ref. [43].

2.5.3 Light-Cone QCD Sum Rules

The method of light-cone QCD sum rules (LCSR) combines the description of heavy-to-light transitions as hard, exclusive QCD processes with QCD sum rules for the description of the decaying hadron. As for all methods based on QCD sum rules, the method is restricted to low q^2 values and is thus complementary to lattice QCD. At sufficiently large $Q^2 \equiv -q^2$, the correlation function is dominated by the contributions from the region near the light cone, $x^2 = 0$. It can be shown [41] that an expansion in local operators around $x = 0$ is not applicable: the operators must be expanded near $x^2 = 0$ in components corresponding to different twists (a twist is defined as the difference between the dimension and the spin of a traceless, symmetric local operator). The calculations can thus be systematically improved by including higher-order twist contributions and radiative corrections to perturbative amplitudes.

The model described in Ref. [12] is based on the light-cone QCD sum rules approach. The vector form factors, computed in [44], include up to $4^{\rm th}$ order twist corrections [45, 46], as well as NLO radiative corrections which affect the resulting form factors by $\sim 10\%$. The $B \to K$ form factors also include twist four mass correction terms. The errors on the estimated branching fractions are dominated by the hadronic uncertainties and are of the order of 30%.

2.5.4 Quark Models

A number of quark models have been proposed in recent years for calculating form factors in $B \to (K, K^*)$ transitions: the light-cone quark model [47], the constituent quark model [48], the QCD relativistic potential model [49]. Reference [50] describes a relativistic constituent quark model that adopts a dispersion formalism. The approach is similar in many respects to the Light Cone QCD Sum Rules, but does not have the problem of a direct application at time-like values of q^2. For heavy-to-light transitions, all the form factors are required to satisfy the leading order Isgur-Wise relations [32]. This requirement is only fulfilled to $\sim 20\%$ for $B \to (K, K^*)$ transitions in this model, as the s quark is not heavy enough.

This model predicts form factors that are in general consistent with LCSR results. The differential decay rates are predicted with an estimated error of $\sim 50\%$. On the other hand, the forward-backward asymmetry is predicted to a $\sim 10\%$ accuracy.

2.5.5 Standard Model Summary

Table 2.1 summarizes the predictions of the branching fractions from various models within the Standard Model. All of the models predict the branching fractions for $B \to K\ell^+\ell^-$ of the order of 10^{-6}-10^{-7}, and of the order of 10^{-6} for $B \to K^*\ell^+\ell^-$. The theoretical uncertainties on the predictions are large enough to make it difficult to use these predictions for testing the Standard Model unless the presence of new physics leads to very dramatic changes. On the other hand, the predictions for the forward-backward asymmetry are largely model-independent, making the measure-

Model	$\mathcal{B}/10^{-6}$		
	$B \to K\ell^+\ell^-$	$B \to K^* e^+ e^-$	$B \to K^* \mu^+ \mu^-$
Lattice QCD [40]	-	1.15	1.15
QCD Sum Rules [43]	0.2	0.7	0.7
LCSR [12]	$0.57^{+0.16}_{-0.10}$	$2.3^{+0.7}_{-0.4}$	$1.9^{+0.6}_{-0.4}$
Quark Models [50]	0.62 ± 0.13	2.1 ± 0.7	1.5 ± 0.6

Table 2.1: Branching fraction predictions in various models within the SM: lattice QCD[40], QCD sum rules [43], quark models [50], and light-cone QCD sum rules (LCSR) [12].

ments of this quantity much more reliable in the searches for new physics. This point is discussed in more detail in the Section 2.6.

2.6 Physics Beyond the Standard Model

Numerous non-Standard Model scenarios have been considered for the $B \to (K, K^*)\ell^+\ell^-$ decays. Examples include multi-Higgs, technicolor, compositeness, left-right symmetric, and top-color models. Generally speaking, there are two mechanisms by which new physics can enter the weak Hamiltonian 2.1.1, both of which can happen in supersymmetry (SUSY): the values of the Wilson coefficients can be modified by introducing new virtual particles into the loops describing the $b \to s$ transitions; and new operators (e.g., operators with a different chirality quark current) can be introduced into the operator product expansion series.

In the framework of SUSY, it is customary to distinguish between two approaches: with and without minimal flavor violation. The models that do not introduce any additional flavor violation in addition to what already exists in the Standard Model belong to the latter, and those that do, to the former category. For example, supergravity (SUGRA) [39] models are effectively minimal flavor violation models, since the only extra flavor violation can come from the renormalization group, and it is not expected to be large. On the other hand, a generic minimal supersymmetric Standard Model (MSSM [51]) can be studied in the Mass Insertion Approximation (MIA [52]), which parametrizes the departure from the minimal flavor violation in terms of deltas or insertions, defined, e.g., in Ref. [53]. These insertions can result in large departures from the Standard Model. For example, the mass insertion $(\delta^u_{23})_{LR}$ - that is, squark mixing in the up sector, between the $2^{\rm nd}$ and the $3^{\rm rd}$ generations, and involving a helicity flip - can provide a major contribution to the bZs vertex [15]. It is curious to note that non-standard Z couplings can in principle be quite significant, for the following reasons. There are three broad categories of the possible local FCNC operators that can be generated at some new physics scale M_X: four-fermion operators (suppressed by $\mathcal{O}(1/M_X^2)$), magnetic operators (suppressed by $\mathcal{O}(1/M_X)$), and Z couplings (which do not need to be suppressed by any powers of $1/M_X$). Apart from the fact that the Z couplings are thus the least suppressed, they are also the least constrained by the experimental data: models with four-fermion operators are correlated with those entering $B-\overline{B}$ mixing, and magnetic penguins are constrained by $b\to s\gamma$. New physics introduced via non-standard bZs couplings can lead to branching fractions that would increase the branching fractions of the exclusive decays by factors of 5 to 10. Another striking characteristic of the model with a non-standard Z coupling is that the zero of the forward-backward asymmetry is essentially unchanged, but the asymmetry itself could change the sign.

To characterize quantitatively the departure of a given model from the Standard Model, it is convenient to define ratios of Wilson coefficients:

$$R_i(\mu) \equiv \frac{\mathcal{C}_i^{NP}+\mathcal{C}_i^{SM}}{\mathcal{C}_i^{SM}} = \frac{\mathcal{C}_i}{\mathcal{C}_i^{SM}}, \tag{2.6.1}$$

where $\mathcal{C}_i^{NP}$ and $\mathcal{C}_i^{SM}$ are the values of the $i^{\rm th}$ coefficient in a new physics (NP) scenario and the Standard Model, respectively. Of particular interest is the ratio R_7, which can be constrained using experimental data [7] on $B\to X\gamma$ in the following way [12]

$$0.80 \ < \ |\mathrm{R}_7(\mu = 4.8\ \mathrm{GeV/c^2})| \ < \ 1.20 \tag{2.6.2}$$

In the SUGRA models, no large deviations from the Standard Model are expected for low values of $\tan\beta$ (defined as the ratio of the Higgs vacuum expectation values v_2/v_1) ; however, for larger $\tan\beta$, leading to a positive $\mathcal{C}_7^{\rm eff}$, it is possible for the branching fractions to become significantly enhanced with respect to the Standard Model. Figure 2.5 illustrates this using

$$\tan\beta = 30, R_7 = -1.2, R_9 = 1.03, R_{10} = 1.0, \tag{2.6.3}$$

which results in a 30% enhancement of the $B \to K\ell^+\ell^-$ rate at low q^2, but the enhancement is difficult to distinguish from the non-perturbative uncertainties, which are themselves of that order, associated with the Standard Model calculation. The effect is much more pronounced in case of $B \to K^*\ell^+\ell^-$: the rate increases by about a factor of two, as shown in Fig. 2.5. The positive $\mathcal{C}_7^{\rm eff}$ also leads to a forward-backward asymmetry that is negative below the J/ψ resonance, so that there is no zero in the asymmetry in this region (Fig. 2.6).

In the MIA-SUSY model considered in Ref. [52], not only can $\mathcal{C}_7^{\rm eff}$ change, but the value of $\mathcal{C}_{10}$ can be significantly enhanced. Moreover, it can change its sign and become > 0. The branching fractions could in principle be enhanced by as much as factors of 5 or 6 if one of the so-called “best-enhancement scenarios” (one with $\mathcal{C}_7^{\rm eff} > 0$ and one with $\mathcal{C}_7^{\rm eff} < 0$) were realized. However, as was pointed out in Ref. [12], both of these scenarios are ruled out by the data. Nevertheless, using experimentally allowed values of R_7, R_9, and R_{10}, this model still predicts a significant enhancement of the rates comparing to the Standard Model. Figure 2.5 shows an example of such an enhancement calculated with the allowed values of the parameters:

$$R_7 = \pm 0.83, R_9 = 0.92, R_{10} = 1.61, \tag{2.6.4}$$

Because this model admits a solution with positive $\mathcal{C}_{10}$, the overall sign of the forward-backward asymmetry can change. Figure 2.6 illustrates this with

$$R_7 = \pm 0.83, R_9 = 0.79, R_{10} = -0.38, \tag{2.6.5}$$

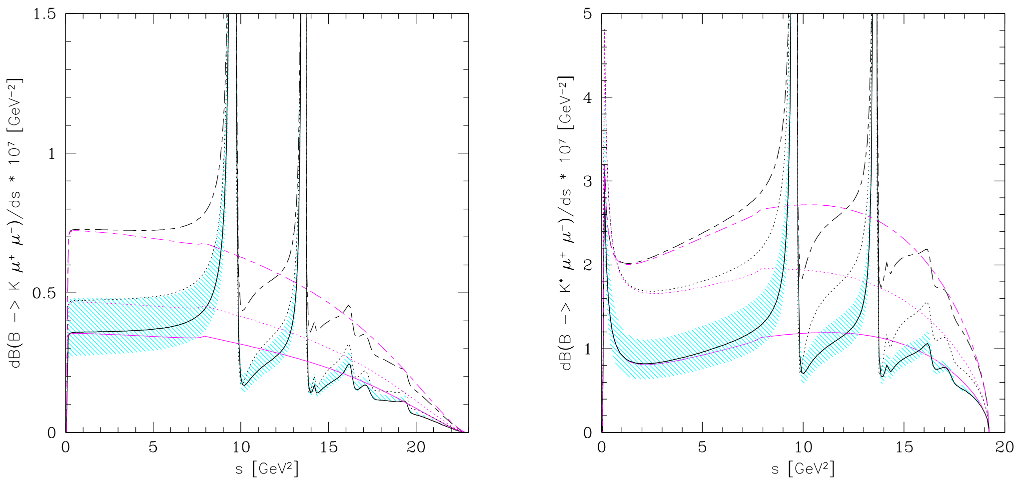


Figure 2.5: The dilepton invariant mass distribution in $B \to K\mu^+\mu^-$ (left plot) and $B \to K^*\mu^+\mu^-$ (right plot) decays, using the form factors from [12] as a function of $s \equiv q^2 = m_{\ell^+\ell^-}$. The $c\bar{c}$ resonant states are parametrized as in [31]. The solid line represents the SM and the shaded area depicts the uncertainties originating from the hadronic form factors. The dotted line corresponds to the SUGRA model with $R_7 = -1.2$, $R_9 = 1.03$ and $R_{10} = 1$. The long-short dashed lines correspond to the MIA-SUSY model with $R_7 = -0.83$, $R_9 = 0.92$ and $R_{10} = 1.61$. Figure taken from [12].

the so-called "best depression scenario" of [52].

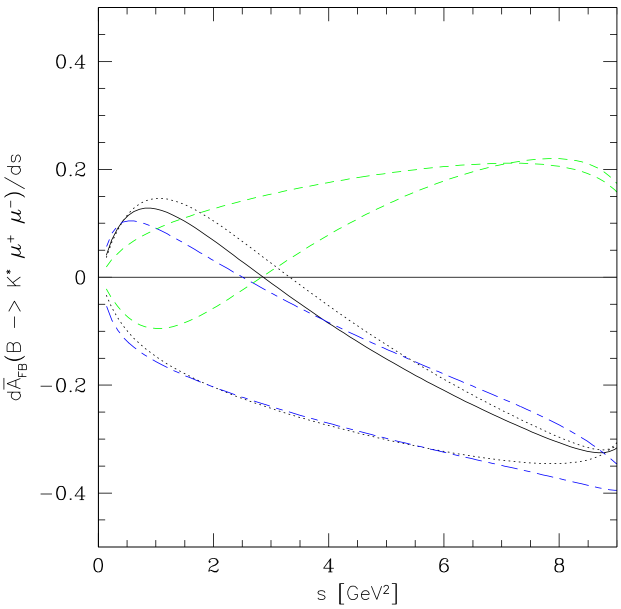


Figure 2.6: The normalized forward-backward asymmetry in $B \to K^*\mu^+\mu^-$ decay as a function of $s \equiv q^2 = m_{\ell^+\ell^-}$, using the form factors from [12]. The $c\bar{c}$ resonant states are parametrized as in [31]. The solid line denotes the SM prediction. The dotted (long-short dashed) lines correspond to the SUGRA (the MIA-SUSY) model, using the parameters given in Eqn. (2.6.3) (Eqn. (2.6.4)) with the upper and lower curves representing the $\mathcal{C}_7^{\rm eff} < 0$ and $\mathcal{C}_7^{\rm eff} > 0$ case, respectively. The dashed curves correspond to the MIA-SUSY models with the parameters given in Eqn. (2.6.5), the "best depression scenario" with $\mathcal{C}_{10} > 0$. Figure taken from [12].

Chapter 3

Experimental status

Searches for inclusive and exclusive $B \to X_s \ell^+ \ell^-$ decays have been carried out by a number of experiments other than *BABAR* : CLEO [54], CDF [55], Belle [3], and UA1 [56]. Results of these searches are briefly reviewed in this chapter.

3.1 Searches for Inclusive $b \to s\ell^+\ell^-$

CLEO has searched for inclusive $b \to s\ell^+\ell^-$ decays using data collected by the CLEO II detector at the Cornell Electron Storage Ring. The analysis is described in Ref. [54]. The data sample consists of $(3.30 \pm 0.06) \times 10^6$ $B\overline{B}$ pairs and 1.6 fb^{-1} of off-resonance data used for background studies. The data selection involves general hadronic event requirements based on the number of charged tracks, the visible energy in the detector, and the location of the event vertex. The two highest energy leptons of the opposite charge are selected and paired with a combination of hadronic particles X_s. The X_s is required to have a kaon candidate and from 0 to 4 prongs, and a mass below 1.8 GeV/c^2 to reduce the combinatorial background from $B\overline{B}$ events. The dilepton-X_s combination is required to be consistent with a B hypothesis. Continuum background is suppressed by requiring that the second order Fox-Wolfram moment R_2 [57] be less than 0.45, and the thrust axis of the event

Mode	Yield	$\mathcal{B}/10^{-5}$	SM $\mathcal{B}/10^{-5}$
$B \to X_s e^+ e^-$	10 ± 5	< 5.7	(0.8 ± 0.2)
$B \to X_s \mu^+ \mu^-$	12 ± 6	< 5.8	(0.6 ± 0.1)
$B \to X_s e^\pm \mu^\mp$	18 ± 8	< 2.2	0
$B \to X_s \ell^+ \ell^-$		< 4.2	

Table 3.1: Event yields and 90% C.L. upper limits on the branching fractions of the inclusive $B \to X_s \ell^+ \ell^-$ decays set by CLEO [54]. The 90% C.L. upper limit on the combined di-electron and di-muon mode $b \to s\ell^+\ell^-$ is also given. Expectations from the Standard Model (SM) [12] are listed for comparison.

$|\cos_{\rm thrust}| < 0.8$. To suppress the background from semileptonic B decays, the mass of the undetected particles in the event is required to be less than 1.5 GeV/c^2. Events with the dilepton mass $m_{\ell^+\ell^-}$ near the J/ψ and $\psi(2S)$ resonances are excluded.

After applying these selection criteria and subtracting the continuum background contribution, CLEO observes 10 ± 5 $X_s e^+ e^-$, 12 ± 6 $X_s \mu^+ \mu^-$, and 18 ± 8 $X_s e^\pm \mu^\mp$ candidate events, where the errors are statistical. The results are consistent with the expected numbers of $B\overline{B}$ background events obtained from the Monte Carlo simulation, which are 9 ± 1, 16 ± 2, and 39 ± 3, respectively (the errors are statistical only). The established 90% C.L. upper limits on the branching fractions, as well as the expectations from the Standard Model [12], are summarized in Table 3.1. The limits are consistent with the Standard Model predictions.

3.2 Searches for Exclusive $B \to (K, K^*)\ell^+\ell^-$ Modes

This section presents the searches for the exclusive decays carried out by CDF [55], CLEO [58], and Belle [3].

3.2.1 CDF Analysis

CDF has searched for the exclusive decays, $B^+ \to K^+\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$, using the 88 pb^{-1} data sample from the 1992-1993 and 1994-1995 runs,

Mode	Yield	Exp. bkg	$\mathcal{B}/10^{-6}$
$B^+ \to K^+\mu^+\mu^-$	4	0.4	< 5.2
$B^0 \to K^{*0}\mu^+\mu^-$	0	0.6	< 4.0

Table 3.2: Event yields, expected backgrounds, and 90% C.L. upper limits on the branching fractions for the exclusive $B \to (K, K^{*0})\mu^+\mu^-$ modes set by CDF [55].

combined [55]. CDF efficiently triggers on events containing two muon candidates. The di-lepton pair is combined with a K^+ or a K^{*0} candidate, where the latter is reconstructed in the $K^+\pi^-$ final state, to form a B^+ or a B^0 candidate, respectively. To further isolate the signal, the B candidate is required to have an isolation $I > 0.6$, where I is defined as the transverse momentum of the B candidate divided by the scalar sum of the transverse momenta of all the tracks in a narrow cone around the B momentum. The B candidate must also have a displaced vertex inconsistent with originating at the beam line. Resonant decays $B \to J/\psi(K, K^{*0})$ and $B \to \psi(2S)(K, K^{*0})$ are suppressed by excluding the regions with the dilepton mass $2.9 < m_{\mu^+\mu^-} < 3.3$ GeV/c^2 and $3.6 < m_{\mu^+\mu^-} < 3.8$ GeV/c^2. $J/\psi K$ events where the kaon and the muon are interchanged as each other are also removed. A signal region is selected by requiring that the mass of the B candidate m_B be $5.23 < m_B < 5.33$ GeV/c^2. The sideband regions are chosen in the range 5.38 - 5.88 GeV/c^2.

There are 4 events in the $K^+\mu^+\mu^-$ channel, and none in the $K^{*0}\mu^+\mu^-$ channel, passing all of the selection criteria. Table 3.2 summarizes the yields, estimated backgrounds, and 90% C.L. upper limits on the branching fractions set assuming all of the events in the signal box are in fact signal. The established limits on the branching fractions are normalized with respect to the tree-level processes $B^+ \to J/\psi K^+$ and $B^0 \to J/\psi K^{*0}$, where $J/\psi \to \mu^+\mu^-$. Figures 3.1 and 3.2 show the invariant mass of the B^+ and the B^0 candidates versus the dilepton mass for the events that passed all of the selection criteria.

Table 3.2 shows that CDF observed an excess of events in the $B \to K\mu^+\mu^-$ mode over the expected background. However, the backgrounds could not be estimated reliably enough: because the high-mass sideband region was used to optimize the event selection criteria, it was not possible to extrapolate the backgrounds from the sidebands into the signal region. This is the reason for the relative depletion of the number of events in the right-hand side sidebands shown in Figs. 3.1 and 3.2.

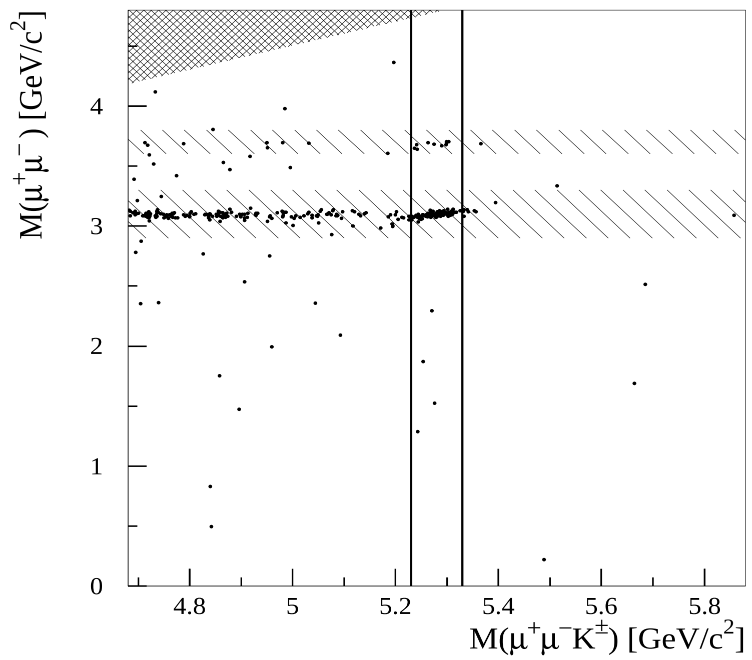


Figure 3.1: B candidate mass versus the invariant dilepton mass for the $B^+ \to K^+\mu^+\mu^-$ mode (CDF analysis). The two vertical lines indicate the signal region, and the hatched horizontal bands, the J/ψ and $\psi(2S)$ exclusion regions. The shaded area in the upper left corner is kinematically forbidden. Figure taken from [55].

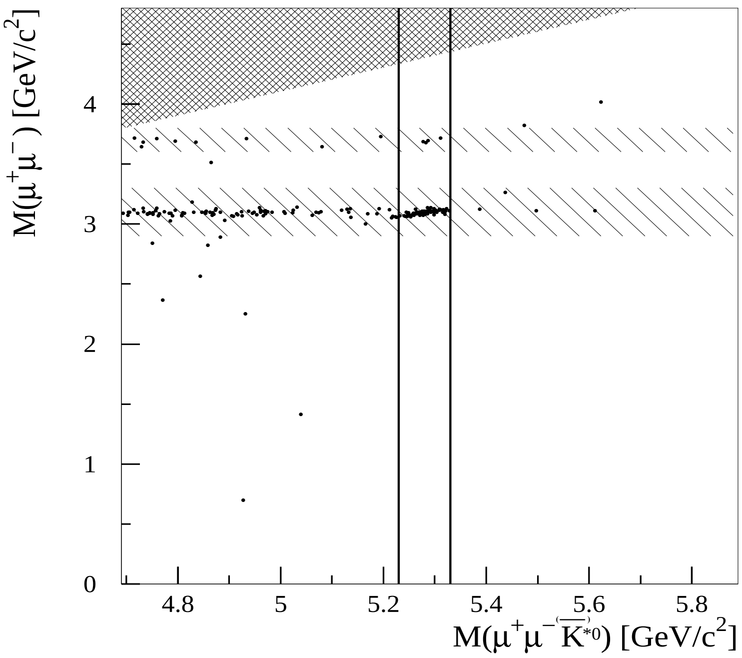


Figure 3.2: B candidate mass versus the invariant dilepton mass for the $B^0 \to K^{*0}\mu^+\mu^-$ mode (CDF analysis). The two vertical lines indicate the signal region, and the hatched horizontal bands, the J/ψ and $\psi(2S)$ exclusion regions. The shaded area in the upper left corner is kinematically forbidden. Figure taken from [55].

3.2.2 CLEO Analysis

CLEO has searched for the exclusive modes using a data set consisting of 9.6×10^6 produced $B\overline{B}$ events [58]. Two leptons, either electrons or muons, are combined with K or K^*, where charged kaons are identified using dE/dx and time-of-flight information. The so-formed B candidate is required to have a beam-constrained mass $m_{\text{cand}} \equiv \sqrt{E_{\text{beam}}^2 - P_{\text{cand}}^2}$ within 6.5 MeV$/c^2$ of the B mass; and

$\Delta E \equiv E_{\mathrm{cand}} - E_{\mathrm{beam}}$, within $\pm$ 60 MeV (for the muon channels) or within $\pm$ 70 MeV (for the electron channels), of zero. Additionally, for the electron modes there is a shift by -5 MeV in ΔE to account for radiation effects. The continuum background is suppressed using a Fisher discriminant [59], and the $B\overline{B}$ background is suppressed using the missing energy of the event. The charmonium background is vetoed by removing events with the dilepton mass in the following ranges:

$$\begin{aligned} 2.80 &< m_{ee} < 3.23 \ \ \mathrm{GeV}/c^2 \\ 2.90 &< m_{\mu\mu} < 3.20 \ \ \mathrm{GeV}/c^2 \\ 3.51 &< m_{ee} < 3.77 \ \ \mathrm{GeV}/c^2 \\ 3.55 &< m_{\mu\mu} < 3.74 \ \ \mathrm{GeV}/c^2. \end{aligned} \tag{3.2.1}$$

The resulting yields, together with estimated backgrounds, efficiencies, and the 90% C.L. upper limits established taking the backgrounds into account, are summarized in Table 3.3. The limits on the combined modes, $B \to K\ell^+\ell^-$ and $B \to K^*\ell^+\ell^-$ (the latter without the $m_{\ell^+\ell^-} < 0.5$ GeV/c^2 region), as well as on the weighted average over $K\ell^+\ell^-$ and $K^*\ell^+\ell^-$, are also given. The two modes are weighted by their relative efficiencies: $0.65\ \mathcal{B}(B \to K\ell^+\ell^-) + 0.35\ \mathcal{B}(B \to K^*\ell^+\ell^-)$. One can see from Table 3.3 that CLEO saw no indication of a signal.

3.2.3 Belle Analysis

Belle has searched for the exclusive decays using the 10.4 fb^{-1} 1999-2000 sample [3] in the following modes: $B^+ \to K^+\ell^+\ell^-$, $B^0 \to K^{*0}\ell^+\ell^-$ where $K^{*0} \to K^+\pi^-$, and $K^{*0} \to K^0_S\pi^0$, and $B^+ \to K^{*+}\ell^+\ell^-$, where $K^{*+} \to K^0_S\pi^+$ and $K^{*+} \to K^+\pi^0$. Both electron and muon channels are considered, with electrons required to have momenta greater than 0.5 GeV/c, and muons, greater than 1.0 GeV/c, in the laboratory frame. The event selection includes using likelihood ratios for background suppression. Continuum background is suppressed using a likelihood ratio containing $\cos\theta_B$ (cosine of the angle between the B candidate and the z axis in the CM frame),

Mode	Yield	Estimated bkg	ϵ (%)	$\mathcal{B}/10^{-6}$
$K^0 e^+ e^-$	1	0.10	5.3	< 7.6
$K^0 \mu^+ \mu^-$	0	0.21	4.1	< 7.8
$K^+ e^+ e^-$	1	0.95	16.5	< 2.3
$K^+ \mu^+ \mu^-$	1	0.74	11.1	< 3.4
$K \ell^+ \ell^-$	3	1.99 ± 0.35	3.7	< 1.49
$K^0 \pi^+ e^+ e^-$	0	0.35	1.9	< 12.8
$K^0 \pi^+ \mu^+ \mu^-$	0	0.27	1.5	< 15.6
$K^+ \pi^0 e^+ e^-$	3	0.27	1.5	< 46.0
$K^+ \pi^0 \mu^+ \mu^-$	0	0.49	0.8	< 29.3
$K^+ \pi^- e^+ e^-$	1	0.97	7.1	< 5.0
$K^+ \pi^- \mu^+ \mu^-$	0	1.24	5.2	< 4.6
$K^0 \pi^0 e^+ e^-$	0	0.11	0.7	< 35.8
$K^0 \pi^0 \mu^+ \mu^-$	0	0.10	0.2	< 117.3
$K^* \ell^+ \ell^-$	4	3.80 ± 0.57	18.8	< 2.94
Sum	7	5.79 ± 0.83	55.8	< 1.35

Table 3.3: The results from the CLEO search for exclusive $B \to (K, K^*)\ell^+\ell^-$ [58]: event yields, estimated backgrounds, signal efficiencies ϵ, and 90% C.L. upper limits on the branching fractions, computed after taking into account the systematic errors in the background estimate, but not in efficiency.

$\cos\theta_{\rm decay\ axis}$ (cosine of the angle between the signal thrust axis and the z axis in CM frame), and the Fisher discriminant, consisting of the virtual calorimeter variables and R_2. $B\overline{B}$ background is suppressed using a likelihood ratio containing the visible energy in the detector and $\cos\theta_B$. The background from π^0 Dalitz decays and γ conversions is suppressed by requiring that $m_{e^+e^-}$ be greater than 0.1 GeV/c^2 for the electron modes. Charmonium background is suppressed by using the following selection criteria:

$$\begin{aligned} -0.15 &< m_{e^+e^-} - m_{J/\psi[\psi(2S)]} < 0.07 \ \ \mathrm{GeV}/c^2 \\ -0.10 &< m_{\mu^+\mu^-} - m_{J/\psi[\psi(2S)]} < 0.05 \ \ \mathrm{GeV}/c^2 \end{aligned} \qquad (3.2.2)$$

Belle observes 4 events in the $B^0 \to K^{*0}\mu^+\mu^-$ channel, while expecting a

Mode	$\mathcal{B}/10^{-6}$
$B^+ \to K^+\mu^+\mu^-$	< 4.0
$B^0 \to K^0\mu^+\mu^-$	< 8.5
$B^0 \to K^{*+}\mu^+\mu^-$	< 8.7
$B^0 \to K^{*0}\mu^+\mu^-$	< 10.5
$B^+ \to K^+e^+e^-$	< 2.4
$B^0 \to K^0e^+e^-$	< 5.9
$B^0 \to K^{*+}e^+e^-$	< 20.8
$B^0 \to K^{*0}e^+e^-$	< 11.1

Table 3.4: The results from the Belle search for exclusive $B \to (K, K^*)\ell^+\ell^-$ [3]: 90% C.L. upper limits on the branching fractions, assuming zero background events

background of 1, and 4 events in the $B^+ \to K^+\mu^+\mu^-$ mode, while expecting 0.6. The statistical significance of the result is 2.7σ, and if the effect is real, the branching fractions for these modes would be two to three times larger than the Standard Model predictions. Table 3.4 summarizes the established 90% C.L. upper limits on the branching fractions, calculated assuming zero background events.

3.2.4 Summary: Established Upper Limits

None of the several experiments that have looked for inclusive or exclusive $B \to X_s\ell^+\ell^-$ processes observed any evidence for a signal. The established 90% C.L. upper limits on the branching fractions are typically two to ten times larger than the Standard Model predictions.

Chapter 4

The BABAR Detector

The primary motivation for constructing an asymmetric B-factory at the PEP-II e^+e^- storage ring at Stanford Linear Accelerator Center (SLAC) is to probe the Standard Model in the decays of neutral B mesons. The Standard Model predicts large CP-violating asymmetries in B decays to CP eigenstates. Because PEP-II operates at the Υ(4S) resonance, and the time integrated asymmetries at the Υ(4S) are equal to zero, the relative decay time of the B mesons must be measured. Due to the low B momentum in the $\Upsilon(4S)$ rest frame, this requires an asymmetric design of the machine and a vertexing device capable of precisely measuring the B decay vertex positions. While the asymmetries are expected to be large, the branching fractions of the B decay channels that can be used for CP violation studies are very small ($\sim 10^{-4}$). Therefore, the machine must provide a very high luminosity. PEP-II is designed to reach $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

4.1 PEP-II and the Interaction Region

PEP is an acronym for the Positron Electron Project. PEP-II is a two-ring e^+e^- storage system at SLAC. The high energy ring (HER) is filled with 9.0 GeV electrons, and the low energy ring (LER), with 3.1 GeV positrons. Typically, PEP-II operates at the center of mass energy of $E_{\mathrm{CM}} = 10.58\,\mathrm{GeV}$, which is the mass of the $\Upsilon(4\mathrm{S})$ resonance [1]. $\Upsilon(4\mathrm{S})$ decays into $B\overline{B}$ pairs almost 100% of the time. Some 12% of the data is taken about 40 MeV below the peak of the resonance to provide a sample of non-resonant background called *continuum* ($e^+e^- \to q\overline{q}$, where $q = u, d, s$ or c). The unequal beam energies give the $\Upsilon(4\mathrm{S})$ a boost in the laboratory frame with a $\beta\gamma$ of 0.56. As a consequence, the B mesons produced in the process $\Upsilon(4S) \to B\overline{B}$ are separated by $\sim$ 250 μm along the beam axis, defined as the z-axis.

A schematic representation of the acceleration and storage system is given in Fig. 4.1. An electron gun is used to create two electron bunches that are accelerated

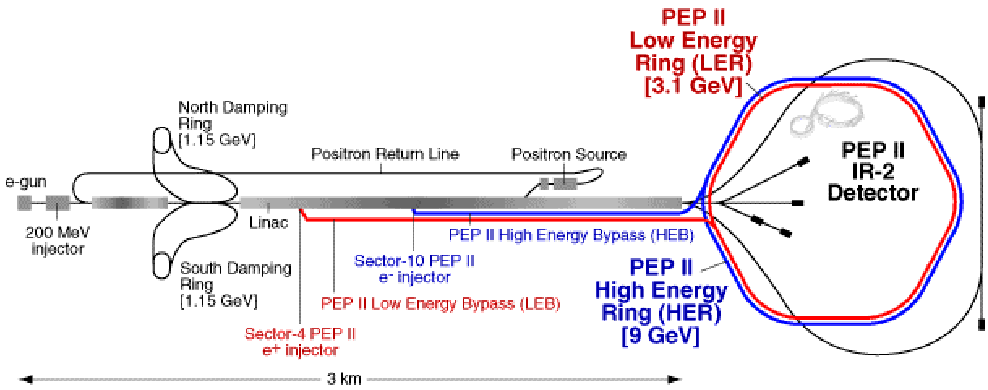


Figure 4.1: Schematic representation of the acceleration and storage system at PEP-II. Electrons and positrons are accelerated in the SLAC Linac and injected into the High and Low Energy Rings, respectively.

to approximately 1 GeV before entering one of the damping rings, whose purpose is to condense the beams. After that, the electrons are accelerated in the Linac (Linear Accelerator). One of the bunches is diverted to create a positron beam, which in turn passes through the damping ring and accelerates in the Linac. Upon reaching the design energies, the electron and positron beams are fed into the PEP-II storage rings, where they collide at the interaction region shown in Fig. 4.2. After the collision, the beams are separated magnetically by dipole magnets called B1,

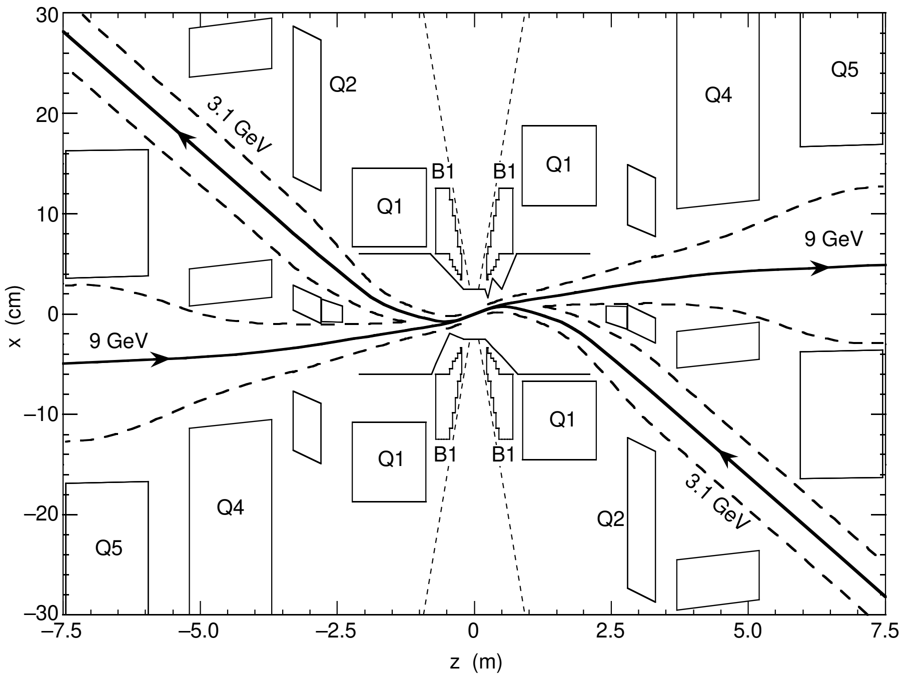


Figure 4.2: The interaction region layout. The beams collide head-on, and are separated magnetically before the next collision by the B1 dipole magnets. The focusing of the beams is achieved by using the quadrupole magnets, Q1, Q2, Q4, and Q5. The dashed lines show the 300 mrad detector acceptance cutoff.

located at ± 21 cm on either side of the interaction point (IP). The strong focusing of the beams is achieved by using quadrupole magnets. The Q1 quadrupole magnet is common to both high-energy and low-energy beams. Q2 is used to only focus the low-energy beam, whereas Q4 and Q5, only the high-energy beam. Both Q1 and B1 are permanent magnets, and Q2, Q4, and Q5 are standard iron magnets. The interaction point is surrounded by a water-cooled beryllium beam pipe with an outer radius of 2.8 cm, contributing about 1.08% of a radiation length at normal incidence. The Q1 magnet, two B1 magnets, the beam pipe, and the silicon vertex detector are assembled into a rigid support structure called the support tube. It has a diameter of roughly 43 cm, and contributes about 0.5% of a radiation length at normal incidence.

Parameter	Design	Typical
HER energy (GeV)	9.0	9.0
LER energy (GeV)	3.1	3.1
HER current (A)	0.75	0.7
LER current (A)	2.15	1.3
# bunches	1658	553-829
Bunch spacing (ns)	4.2	6.3-10.5
σ_x (μm)	110	120
σ_y (μm)	3.3	5.6
σ_z (mm)	9	9
Inst. luminosity (10^{33} cm^{-2}s^{-1})	3	2.5
Daily luminosity (pb^{-1}/d)	135	120

Table 4.1: PEP-II parameters at the design stage and typical after the first year of operation. HER and LER stand for high- and low- energy ring, respectively. σ_x σ_y, and σ_z are the rms size of the beam spot in the x, y, and z directions, respectively.

Some of the PEP-II parameters, at design stage and typical after the first year of operation, are summarized in Table 4.1. The high luminosity of the machine is achieved primarily by using high beam currents, a multi-bunch mode of operation with a large number of bunches, and strong focusing of the beams. Table 4.1 shows that the design luminosity has been achieved with fewer bunches than originally anticipated.

The high luminosity of PEP-II has important implications in terms of acceptable background levels for the proper functioning of the detector. Background sources include synchrotron radiation, interactions between the beam and the residual gas in the rings, and electromagnetic showers produced in beam-beam collisions. Beam-gas bremsstrahlung and Coulomb scattering of the beam particles off the residual gas in the rings dominate the Level 1 trigger rate, the instantaneous silicon vertex detector dose rates, and the total drift chamber current. Energy-degraded beam particles resulting from such interactions are bent by the separation dipole magnets horizontally into the beam pipe, resulting in occupancy peaks for almost all of the

BABAR subdetectors at $\phi = 0^\circ$ and 180°. The rate of this background is proportional to the product of the beam currents and the gas pressure in the rings. At higher luminosities the background from radiative Bhabha scattering is expected to be increasingly more important.

4.2 The Detector Overview

The *BABAR* detector, described in detail in Refs. [60]-[53], is depicted in Fig. 4.3. It was assembled at SLAC and detected the first collisions in May of 1999. The detector consists of five main components: a silicon vertex detector, a drift chamber, a dedicated particle identification system, an electromagnetic calorimeter, a superconducting solenoid providing a magnetic field of 1.5 T, and an instrumented flux return. The detector acceptance is $17^\circ < \theta_{\rm lab} < 150^\circ$ in the laboratory frame (-0.95 $< \theta_{\rm CM} <$ 0.87 in the center of mass frame).

The silicon vertex detector serves not only as a vertexing device but also as the only tracking system inside the 20 cm radius of the support tube, which is why it is often called the silicon vertex tracker (SVT). Outside the support tube, between the radii of about 22.5 cm and 80 cm, is the drift chamber (DCH). Apart from providing momentum measurements for tracks with transverse momenta $p_t > 100\,{\rm MeV/c}$, it is used as part of the trigger and also measures the ionization loss, dE/dx, for particle identification. A complementary particle identification system is used for tracks with momenta $0.5 < p < 4.5\,{\rm GeV/c}$: the detector of internally reflected Cherenkov light (DIRC). The cesium iodide (CsI) electromagnetic calorimeter (EMC) is used for detecting electromagnetic showers with energies from 20 MeV to 4 GeV. The instrumented flux return (IFR) serves as a muon and neutral hadron identification system.

An effort has been made to minimize the amount of material in the active detector volume in order to decrease the impact of multiple scattering. Figure 4.4 shows the distribution of material in the detector components.

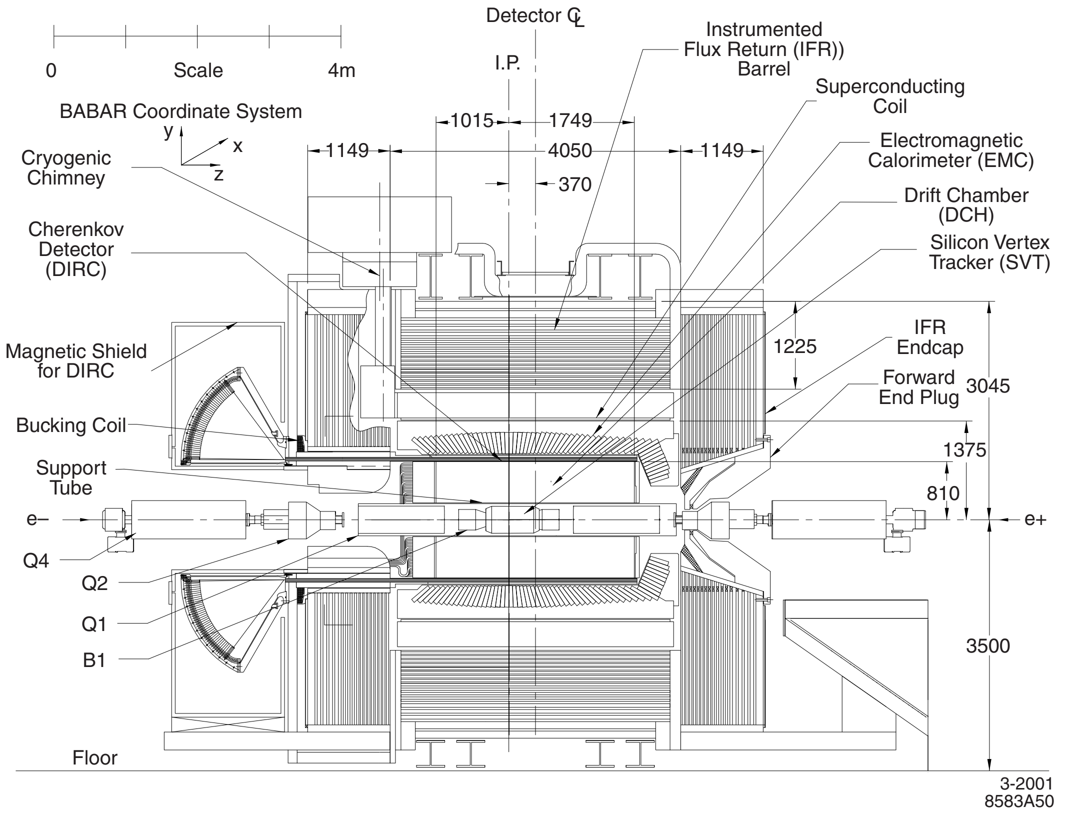


Figure 4.3: A cross-sectional view of the *BABAR* detector. Moving radially outward, the subdetectors are: (1) Silicon Vertex Tracker (SVT), (2) Drift Chamber (DCH), (3) Detector of Internally Reflected Cherenkov Light (DIRC), with its stand-off box (SOB) shown in the forward direction, (4) Electromagnetic Calorimeter (EMC), (5) magnet providing a 1.5 T field, (6) Instrumented Flux Return (IFR). All dimensions are in mm.

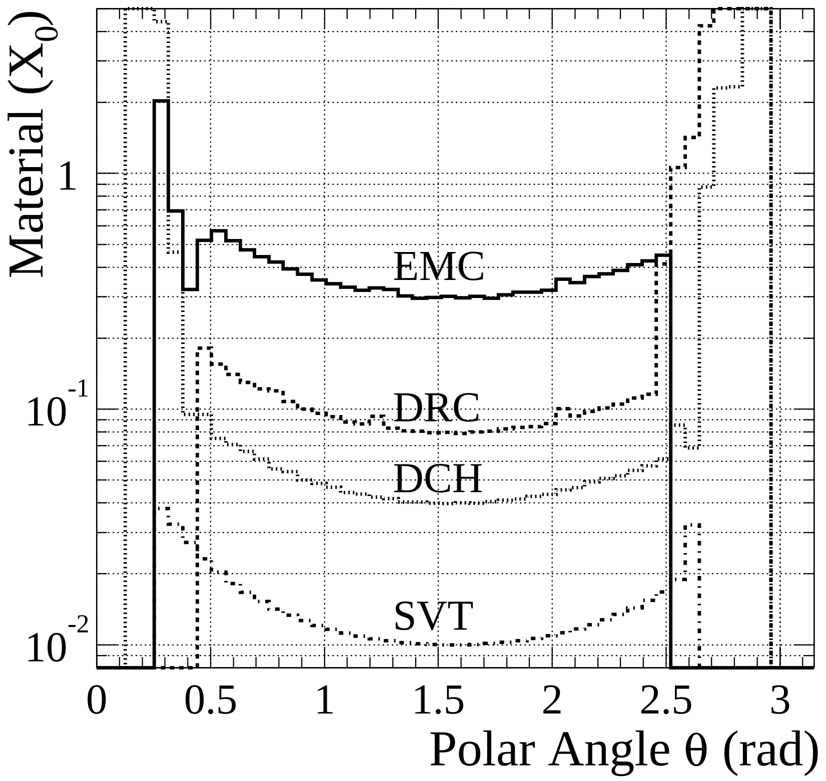


Figure 4.4: The distribution of material in the detector components as a function of the polar angle, in units of radiation lengths. At normal incidence, the beam pipe contributes about 1.08% X_0; the SVT, 4.0%; the support tube, 0.5%; the DCH, 1.08%; and the DIRC, 19%, making it a total of about 26.0% X_0 of material before the EMC.

4.3 Silicon Vertex Tracker

The design of the silicon vertex tracker (SVT) is dictated by its dual function as a vertexing and a tracking device. The main tracking device in *BABAR* is the drift chamber. However, because of the machine design, the inner radius of the drift chamber is relatively high (22.5 cm). In a 1.5 T magnetic field, this implies that the transverse momentum threshold for particles to reach the drift chamber is $p_t \sim 51$ MeV/c; and a p_t of at least 120 MeV/c is required for track reconstruction. Therefore, the SVT must be able to perform stand-alone pattern recognition and tracking for low momentum charged particles, which means that at least five layers are needed.

Good vertexing is crucial for CP-violation studies. Early studies of the detector showed that in order to measure the CP asymmetry in *BABAR*, where the mean separation between the B mesons is $\sim 250\,\mu$m, it is sufficient to determine the B decay vertex positions to a precision of about $80\,\mu$m. For a five-layer vertexing detector, this implies a $10 - 15\,\mu$m point resolution in the innermost layers. Resolutions of that order are routinely achieved by silicon microstrip devices.

The SVT is also used for measuring dE/dx for particle identification.

4.3.1 SVT Geometry

The SVT is a five layer double-sided silicon microstrip detector. Figure 4.5 shows the layout of the detector in the $x - y$ plane, and Fig. 4.6 shows an $r - z$ view of the upper half. The acceptance of the device is $17.2^\circ < \theta < 150^\circ$ in the laboratory frame, and is limited primarily by the location of the B1 dipole magnets (see Fig. 4.2). To reduce the effect of multiple scattering on the determination of the track impact parameters, it is important to minimize the amount of material between the IP and the first measuring plane. Therefore, the first layer of the SVT must be placed as close as possible to the beam pipe. The first three layers are arranged in a barrel structure, divided into sextants. Their primary goal is to provide precision

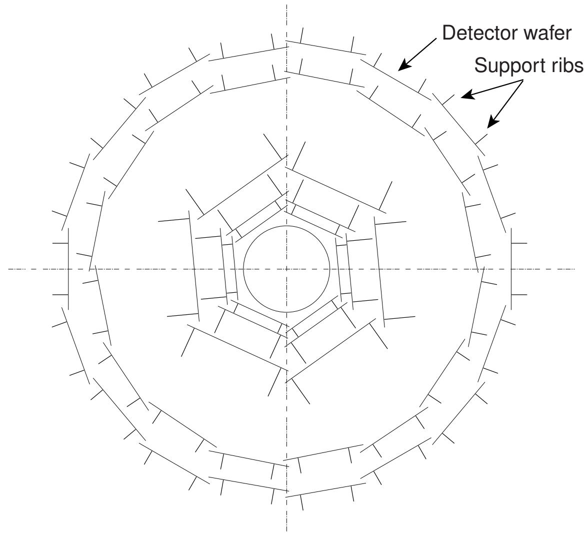


Figure 4.5: SVT layout in $x - y$ view. Only barrel wafers are shown. The modules in the inner three layers are rotated in ϕ by 5°, and have overlapping regions. The outer layers are divided into sublayers to provide a full ϕ coverage.

angle ($\phi - \theta$) and impact parameter measurements. The outer layers are needed for pattern recognition and stand-alone tracking. Their wedge end wafers minimize the amount of silicon needed for the solid angle coverage and reduce large incidence angles. Because of the arch geometry, it is not possible to tilt the wafers in ϕ. The overlap in ϕ is therefore provided by dividing the layers into sublayers (4a, 4b, and 5a and 5b, as shown in Fig. 4.5).

Within each layer, silicon wafers are combined into modules. There are 6 modules in the first three layers, arranged in a pinwheel geometry. Modules have 4 wafers each in the first two layers, and 6 wafers each in the third layer. Layers 4 and 5 have 16 and 18 modules with 7 and 8 wafers per module, respectively. The total number of wafers is 340. Different layers have, in general, different wafer types: there are six distinct wafer configurations. The differences are in the wafer geometry, pitch,

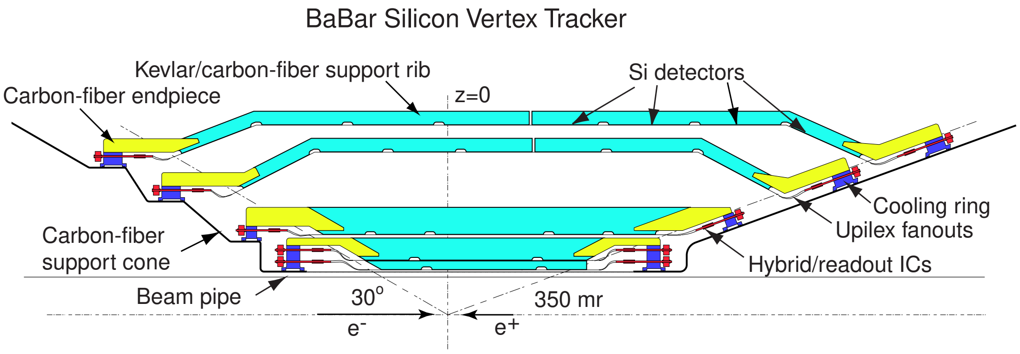


Figure 4.6: SVT layout: cross-section view of the upper half. The first layer radius is 3.3 cm, and the maximum fifth layer radius is 14.4 cm. The outer layers have an arch structure to minimize the amount of silicon needed for the solid angle coverage and to reduce large incidence angles.

and the number of strips. These parameters are summarized in Table 4.2. In the barrel wafers, the ϕ strips are oriented parallel to the beam pipe, and the z strips are perpendicular to it. The readout pitch in the inner layers is the smallest to provide the best possible resolution: it is 100 μm in z, and 50-100 μm in ϕ, depending on the region. The physical pitch is 50 μm in both views, but all of the z, and some ϕ, strips are "floating". The floating strips are capacitively coupled to the neighboring strips, but are not attached to a readout chip channel. The outer layers are not crucial for precision vertexing, and their strip pitch can therefore be larger.

Each module is divided into two halves, forward and backward. Within each half, from two to four silicon wafers are glued to a flexible fanout. The entire unit is called a Detector Fanout Assembly (DFA). The fanout connects the silicon strips with the front-end electronics. The fanouts are 50 μm thick and are made out of Upilex (a Kapton-like material). They are plated with a 150 nm chromium adhesion layer, followed by a 4.5 μm of electroplated copper conductive traces and another 150 nm chromium adhesion layer plus approximately 1 μm of gold. In the

Wafer type	I	II	III	IV	V	VI (wedge)
Dimensions (cm)						
z	4.24	4.54	4.39	6.76	5.39	6.80
ϕ	4.13	4.94	7.15	5.25	5.25	5.28→4.33
Strip Pitch (μm)						
z	50	50	50	105	105	105
ϕ	55	55	55	50	50	50→41
Floating Strip						
z	yes	yes	yes	yes	yes	yes
ϕ	partial	partial	yes	yes	yes	yes
Implant Type						
z	p	p	p	n	n	n
ϕ	n	n	n	p	p	p

Table 4.2: SVT wafer parameters for all six wafer types: ϕ and z dimensions, strip pitch, the presence of floating strips, and the implant type.

outer layers the number of available readout channels is less than the number of the z strips, and therefore the fanouts realize times 2 ganging (two z strips are connected in series by the fanout traces to the same readout channel).

Readout electronics are located entirely outside the active tracking volume, thus minimizing the amount of material in the acceptance region. The DFAs are wire-bonded to double-sided High-Density Interconnects (HDIs), which hold the readout chips. There are three different types of the HDIs, with 7 chips on each side for layers 1 and 2, 10 chips on each side for layer 3, and 9 chips total for layers 4 and 5. The HDI substrate is a 1.2 mm thick aluminum structure. Mechanically, the HDIs for each module are mounted on brass cooling rings, which are themselves mounted on the carbon-fiber support cones surrounding the B1 dipole magnets. These cones are connected to a carbon-fiber space frame, thus forming a rigid SVT support structure. Also part of the support elements are the module ribs and endpieces. The support ribs consist of a kevlar/epoxy piece surrounded by two thin carbon-fiber/epoxy pieces.

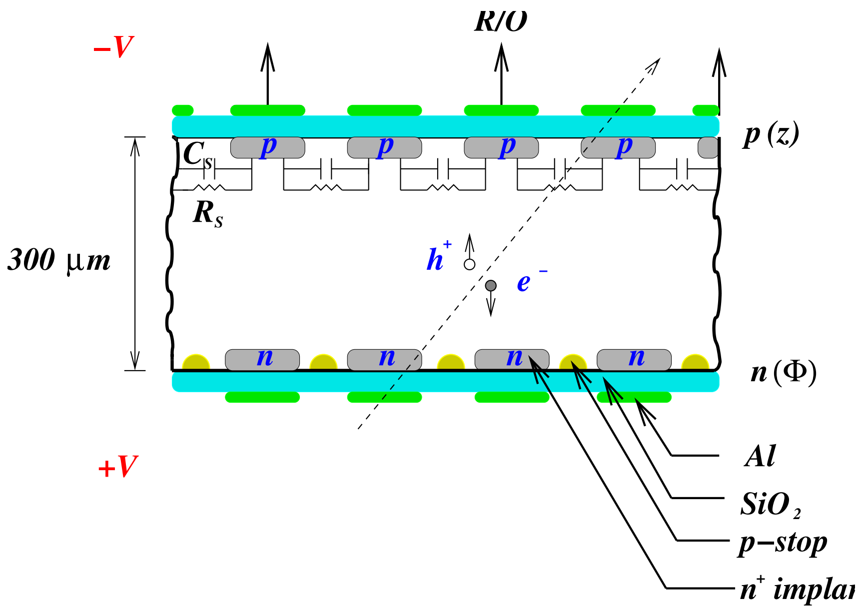


Figure 4.7: A simplified representation of an inner-layer SVT wafer. The implants are capacitively coupled to the aluminum readout strips. The p^+-stops on the $n(\phi)$ side serve to isolate the n^+ implants.

The wafers were manufactured by Micron Semiconductor[1]. A simplified diagram of a wafer is given in Fig. 4.7. Since $dE/dx \sim 3.8$ MeV/cm for a minimum ionizing particle (MIP), and the average energy required to create an electron-hole pair in silicon is ~ 3.6 eV, a 300 μm thick wafer is capable of producing large signals of $\sim$ 24,000 electron-hole pairs for a MIP. The wafers are n-type, with a resistivity of the order of a few kΩ-cm and a depletion voltage ranging from 15 V to 40 V. The n^+ and p^+ implants on two opposite sides are AC-coupled with the readout: there is a silicon-dioxide ($Si0_2$) layer between each implant and the aluminum readout strip. The implants are biased with polysilicon resistors, with resistances of 4-20 MΩ. The AC decoupling capacitance ranges from 20 pF/cm to 43 pF/cm, depending on the implant type and the strip pitch. The implant-to-back capacitance ranges from 0.17

[1]Micron Semiconductor Ltd., Lancing, UK

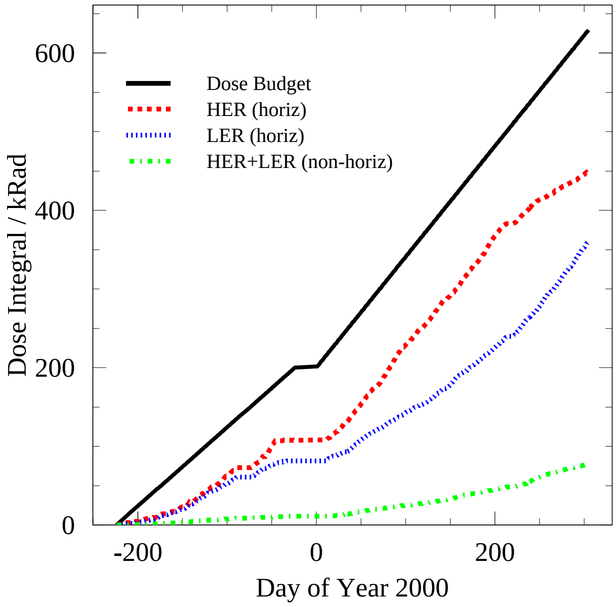


Figure 4.8: Accumulated radiation dose in the SVT, measured by three PIN diodes located in three different positions near the SVT. The top curve is the budget radiation dose in the SVT for the first year of running.

to 0.36 pF/cm, and the interstrip capacitance is 1 pF/cm. The wafer leakage current is typically about 50 nA/cm^2 on average at 10 V above the depletion voltage.

Because the SVT is the closest subdetector to the IP, the radiation doses are constantly monitored with 12 silicon photo-diodes, located at a radius of 3 cm from the beam pipe. Even though the instantaneous doses can be very high during beam injection, the lifetime of a detector subsystem is determined not by the instantaneous rates, but by the total integrated radiation dose. Figure 4.8 shows the total accumulated dose in the SVT. The dose is currently $\sim$ 30% below the allowed budget.

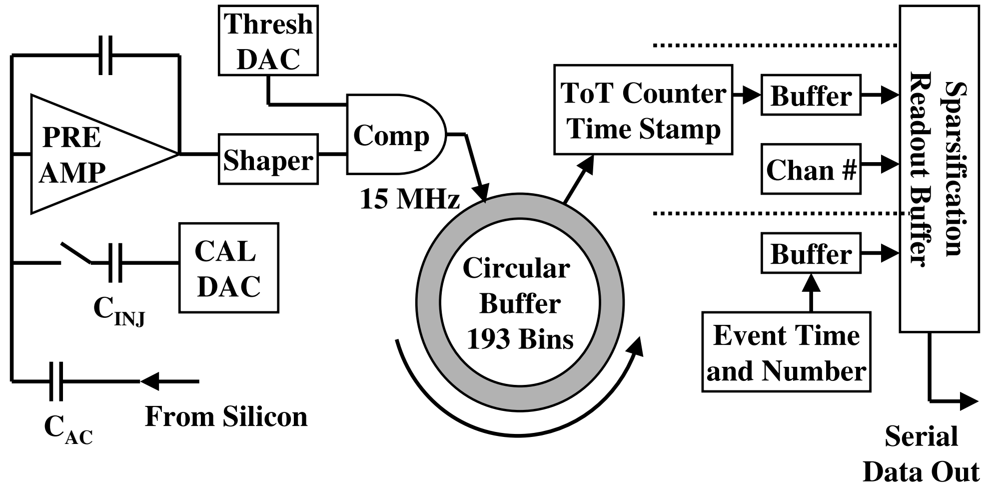


Figure 4.9: A schematic representation of the SVT readout chip.

4.3.2 SVT Front-End Electronics

Figure 4.9 shows a schematic representation of the readout AToM chip (AToM stands for *A Time Over* threshold *M*achine). The chip is manufactured by Honeywell[2] using a radiation-hard CMOS process. Because the bunch-crossing period at PEP-II is short ($\sim$ 4.2 ns), the chip must be able to acquire, amplify, digitize, buffer, and readout all input signals simultaneously. A buffering capability is provided to store signals for the duration of the Level-1 trigger latency, as well as for the digitized data until the readout command is received. The readout is based on the signal time-over-threshold (TOT).

There are 128 channels per chip, with a channel-to-channel pitch of 50 μm. The main components of each channel are an analog section, a comparator, a revolving buffer, a counter, and a back-end readout. The analog amplifier has two

[2]Honeywell Inc., Minneapolis, MN, USA

gain settings: 200 mV/fC (low) or 300 mV/fC (high). The amplifier is followed by a CR(RC)^2 shaper with variable peaking time (100, 200, 300, or 400 ns). At the comparator input, the shaped signal is compared with a pre-defined threshold, and if it exceeds the threshold, a TOT signal is generated. One TOT count corresponds to 67 ns. In principle, the incident charge can be extracted from the TOT value. The relationship between the charge and TOT, shown in Fig. 4.10, is close to logarithmic, giving finer segmentations for smaller charges. The output of the comparator is sampled with a frequency of 30 MHz, and is stored in the 193 bin revolving buffer. The revolving buffer digitizes the TOT information by synchronization with an external clock and stores it until the readout is initiated by the receipt of a Level 1 accept command from the data acquisition system. The data retrieval begins when a zero-to-one transition is detected in the buffer. As soon as this happens, a hit flag is raised, a hit time stamp is latched, and the TOT counter starts counting the number of ones in the buffer. This number is the digital TOT from which the initial signal charge can be calculated. Four bits for the digital TOT have been determined to produce sufficiently good position resolution.

4.4 The Drift Chamber

The main functionality provided by the *BABAR* Drift Chamber (DCH) is tracking for charged particles with momenta greater than approximately 100 MeV/c. In order to achieve very good momentum resolution, an effort has been made to minimize the effect of multiple scattering by using low mass materials in the construction of the chamber. Apart from tracking, the chamber is used for particle identification via the measurements of the ionization loss (dE/dx), and for Level 1 triggering.

4.4.1 Drift Chamber Overview

The drift chamber, shown in Fig. 4.11, is a 280 cm long cylinder with aluminum end-plates, strung with low-mass aluminum wires and filled with a 80:20

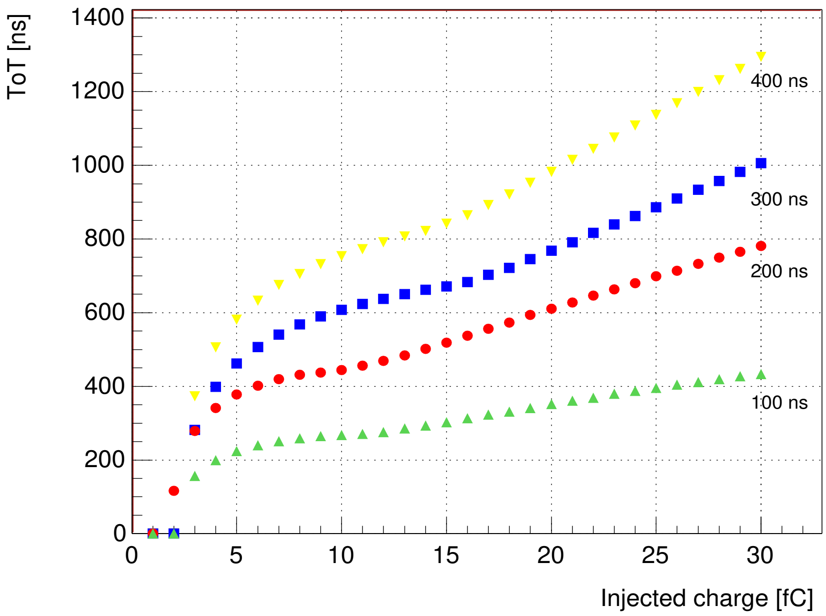


Figure 4.10: Time over threshold as a function of injected charge at four different peaking time settings. Figure courtesy Amedeo Perazzo.

mixture of Helium:Isobutane. Using a helium-based gas mixture allows to reduce multiple scattering and to achieve a dE/dx resolution of approximately 7.5%. The inner radius of the drift chamber is 23.6 cm and its outer radius is 80.9 cm. Because of the $\Upsilon(4S)$ boost along the beam (z) direction, the chamber forward-backward coverage is not the same: -1.11 $< z <$ 1.66 m. In order not to degrade the performance of the DIRC and the EMC in the forward region, there is less material in the forward part of the chamber than in the backward part. The forward end-plate is 12 mm thick while the rear end-plate is 24 mm thick; and all of the readout electronics are mounted on the rear end-plate. The inner cylinder is 1 mm beryllium, which corresponds to 0.28% radiation lengths. It was made as thin as possible in order to reduce the effect of photon conversions and facilitate track matching between the

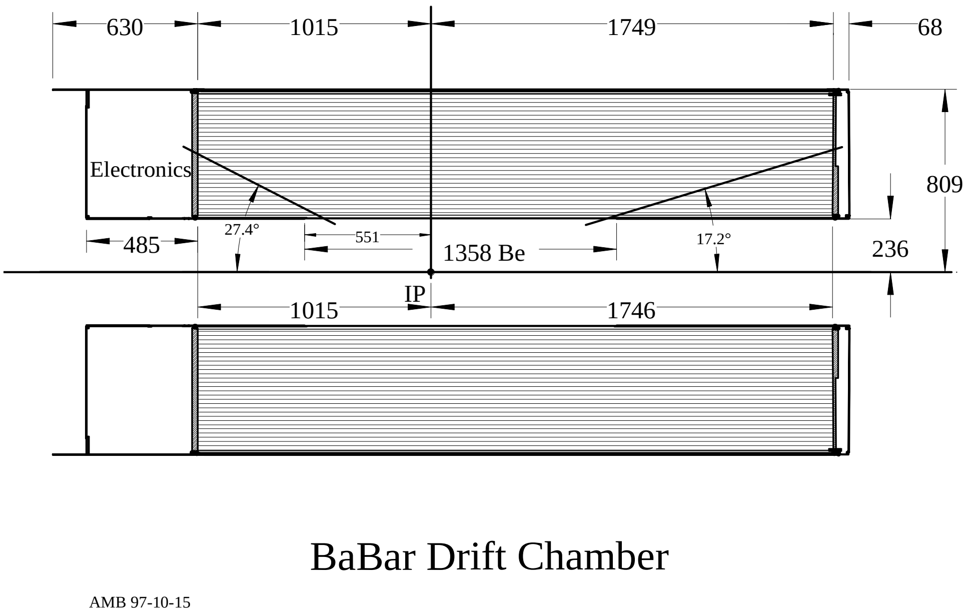


Figure 4.11: An $r - z$ view of the drift chamber (all dimensions are in mm). The asymmetric design of the chamber is a consequence of the $\Upsilon(4S)$ boost and serves to improve the forward tracking efficiency.

SVT and the DCH. The outer cylinder consists of two layers of carbon-fiber on a Nomex core, which corresponds to 1.5% radiation lengths. The total material inside the drift chamber contributes about 0.2% radiation lengths.

The wires are organized into 40 layers of small hexagonal cells, subdivided into 10 superlayers in an alternating axial (A) and stereo (U,V) pattern, as shown in Fig. 4.12. The stereo angle increases radially from $\pm$ 45 mrad to $\pm$ 76 mrad. The stereo measurements are used for extracting the longitudinal (z) position information. There are a total of 7,104 cells, which are hexagonal and small to achieve the needed pattern recognition efficiency. A typical cell size is 1.2×1.9 cm^2 along the radial and azimuthal directions, respectively. Figure 4.13 shows 100 ns contours of equal drift time in a typical cell. Each cell contains a 20 μm gold-plated tungsten-rhenium sense wire, surrounded by a grid of gold-plated aluminum field wires which are 120 μm

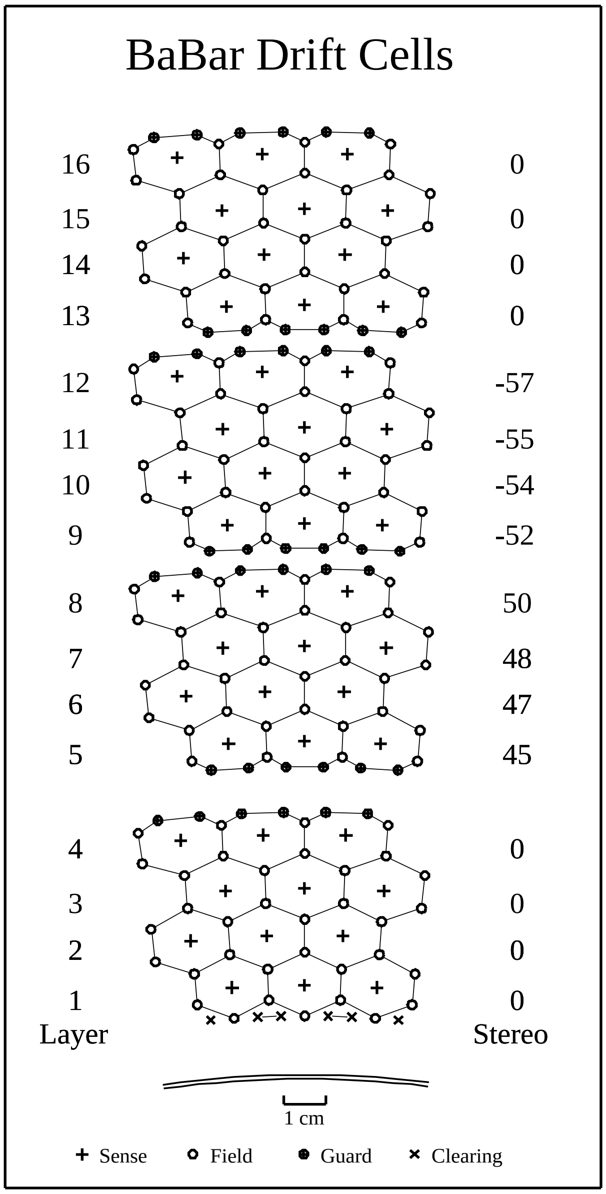


Figure 4.12: An illustration of the drift chamber layer arrangement. The layers are organized into 10 superlayers of 4 layers each, alternating in axial and stereo orientation.

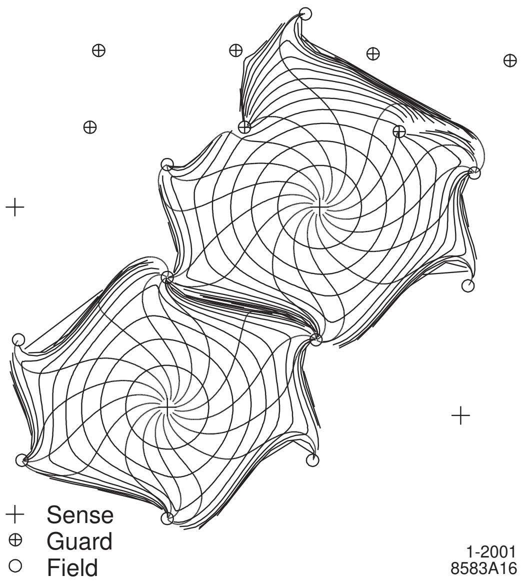


Figure 4.13: 100 ns isochrones in a typical drift chamber cell in a 1.5 T magnetic field. The isocrones are circular near the sense wires, but become distorted close to the field wires.

and 80 μm thick. Near the sense wire, isochrones are circular; however, they become distorted close to the field wires. A potential of 1900-1960 V is applied to the sense wires, while the field wires are grounded (except for the outer cells, where one of the cathode wires is held at 350 V). This potential difference gives an avalanche gain of about 5×10^4. The cells are designed to provide an average intrinsic spatial resolution better than 140 μm.

While the SVT dominates the impact parameter and angle measurements, the DCH dominates the transverse momentum p_t resolution. The data is well represented by the following formula:

$$\frac{\sigma p_t}{p_t} = (0.13 \pm 0.01)\%\ p_t + (0.45 \pm 0.03)\% \tag{4.4.1}$$

4.4.2 Tracking Efficiencies

BABAR has two tracking subsystems that can be considered independent of each other, the SVT and the DCH. This makes it possible to determine the absolute DCH tracking efficiency as the fraction of tracks found in the SVT that are also reconstructed in the DCH, provided that corrections for the fake tracks found in the SVT and the tracks that scattered in the support tube are applied. A small additional correction is applied to account for the fact that the two tracking subsystems are in fact linked together at one stage of the track finding algorithm. This prescription for determining track finding efficiencies can be applied to any data and/or Monte Carlo samples. Figure 4.14 shows the tracking efficiency as a function of transverse momentum and polar angle for a sample of multi-hadron events. The data from the 1999-2000 run were taken at two different DCH voltage settings, 1900 V and 1960 V, and the tracking efficiency was studied separately for both of these conditions. At 1960 V, the average efficiency is measured to be $(98 \pm 1)\%$ per track, for tracks with $p_t > 200\text{MeV}/c$ and polar angles $\theta > 500$ mrad. At 1900 V, the efficiency is lower by about 5% for tracks at normal incidence.

4.5 Detector of Internally Reflected Cherenkov Light (DIRC)

Good particle identification in *BABAR* is very important. For example, the decays used for B flavor tagging and rare B decays depend in a crucial way on a system capable of discriminating kaons from pions over a large range of solid angle and momentum. Good kaon/pion separation up to momenta of 0.7 GeV/c is obtained by measuring the ionization energy loss, dE/dx, in the silicon vertex detector and the drift chamber. A complimentary particle identification system is needed for higher momentum tracks. This system is the Detector of Internally Reflected Cherenkov Light (DIRC). This device is located entirely within the calorimeter volume. Thus,

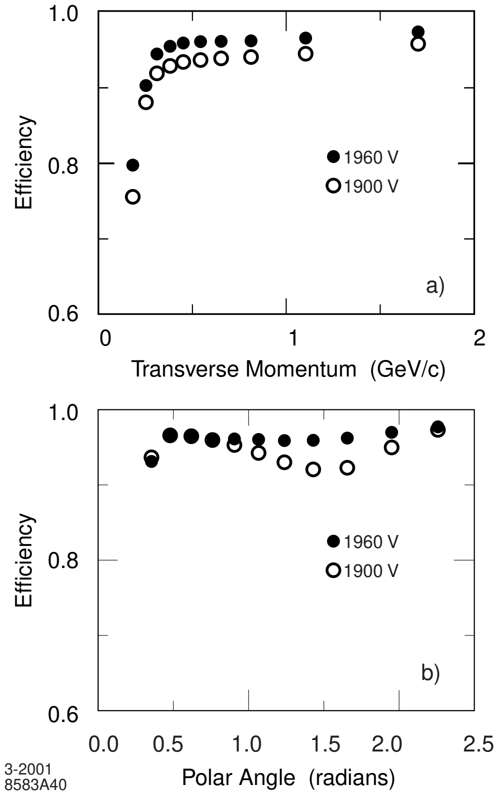


Figure 4.14: The track reconstruction efficiency in the DCH as a function of transverse momentum (a), and polar angle (b). The measurements at the DCH voltage of 1900 V (open circles) and 1960 V (filled circles) are shown.

it must be thin and uniform in radiation lengths.

4.5.1 DIRC Overview

The DIRC, shown in Fig. 4.15, consists of 12 bar boxes made of thin aluminum, each one containing 12 long quartz bars with 75 μm gaps between them. The bars, manufactured by Boeing[3], are 1.7 cm thick, 3.5 cm wide, and 4.9 m long. At the readout end, 9 cm long fused silica wedges are glued to the bars, providing

[3]The Boeing Company, Seattle, WA, USA

the interface to the SOB water. The standoff box (SOB) is made out of stainless steel and can hold about 6 m^3 of water. Its main components are a cone, a cylinder, and 12 sectors at the rear end at the distance of about 1.17 m from the end of the bars, containing photomultiplier tubes (PMTs). To reduce the amount of material in the forward direction, all of the readout electronics are mounted at the rear end of the bars. The angular coverage of the detector in the laboratory frame is -0.84 $< \cos\theta_{\rm lab} <$ 0.90.

The DIRC is based on the detection of the Cherenkov light produced by charged particles moving faster than the speed of light in quartz radiator bars. If the speed of a particle is v, and the medium index of refraction is n, then the particle will generate a cone of Cherenkov radiation with a half-opening angle $\theta_C = 1/\beta n$ with respect to the particle direction. If the index of refraction n is larger than that of the surrounding medium, then starting at a critical angle of incidence most of the light will be internally reflected multiple times, preserving the magnitude of the Cherenkov angle θ_C. In the DIRC quartz, $n_{\rm quartz} = 1.473$. Only one end of the bar is made instrumented: a mirror is placed on the other end, perpendicular to the bar axis. The light at that end is reflected toward the readout on the other end. Upon reaching the instrumented end of the bar, the light expands though the SOB water ($n \sim 1.346$) and is detected by an array of about 11,000 densely packed photomultiplier tubes, producing a Cherenkov ring. The Cherenkov angle is measured from the radius of the Cherenkov ring. Together with the particle momentum p measured by the drift chamber, it allows one to determine the particle mass m :

$$m = p\ \sqrt{n^2_{\rm quartz}\cos\theta^2_C\ -\ 1}. \tag{4.5.1}$$

The overall single photon resolution is about 10.2 mrad, dominated by how well the distance from the end of the quartz bar to the PMT is known and by the resolution contribution from the photon production. The resolution on a track scales as $\sigma_\gamma/\sqrt{N_\gamma}$, where σ_γ is the resolution on a single photon and N_γ is the number of

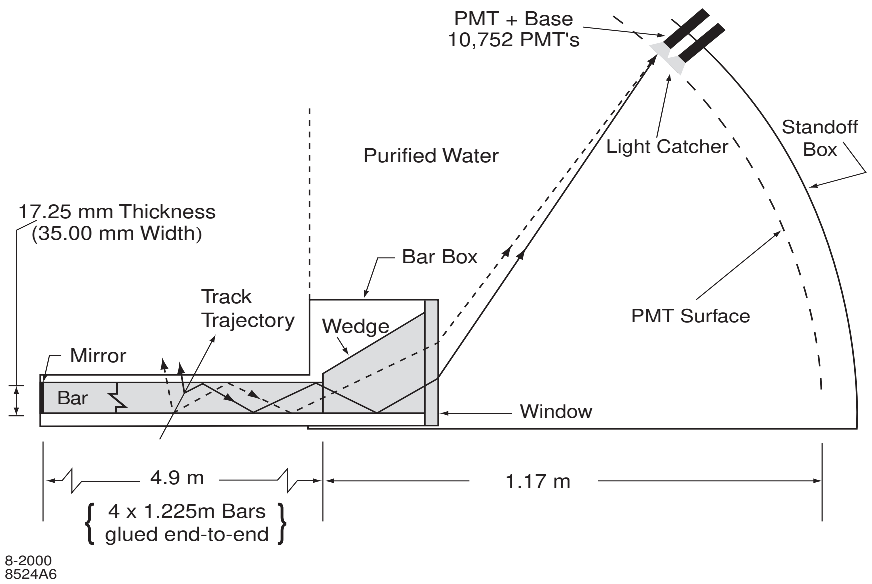


Figure 4.15: An illustration of the DIRC working principle. The Cherenkov light from a particle moving faster than the speed of light in quartz is internally reflected multiple times, until it reaches the instrumented end of the quartz bar. The light then expands through the standoff box filled with purified water and is detected by an array of photomultiplier tubes.

detected photons. N_γ is about 30 for normal incidence tracks. The average track Cherenkov angle resolution, measured in di-muon events, turns out to be 2.5 mrad, which leads to a $\sim 4\sigma$ separation between kaons and pions at 3 GeV/c.

4.6 The Electromagnetic Calorimeter

BABAR electromagnetic calorimeter (EMC), primarily used to detect photons and to separate electrons and positrons from charged hadrons, is made out of thalium dopped cesium iodide (CsI(Tl)) crystals. Because half of the photon energies from generic B decays are below 200 MeV, the calorimeter is designed to reconstruct low energy photons with an excellent efficiency and resolution.

4.6.1 EMC Layout and Technology

The EMC, depicted in Fig. 4.16, consists of a barrel section and a forward endcap. The barrel inner radius is 90 cm and its outer radius is 135.6 cm. It is forward-backward asymmetric to reflect the *BABAR* boost. The barrel angular coverage in the laboratory frame is $-0.80 < \cos\theta_{\rm lab} < 0.89$. There are a total of 5,760 CsI crystals in the barrel region, arranged in 48 θ rows of different crystal sizes, each with 120 identical crystals in ϕ. Each crystal is held in a 250 μm-thick carbon fiber composite compartment. The support structure of each such compartment allows the crystals to be grouped together into 280 modules of three crystals wide and seven crystals long. The forward endcap has a conical geometry, covering $0.89 < \cos\theta_{\rm lab} < 0.97$ in the laboratory frame. There are 820 crystals total in the endcap, grouped into 20 modules of 41 crystals each. They are tilted at 22.7° with respect to the vertical, thus minimizing the distance to the drift chamber endplates.

Crystals are trapezoidal in shape and range in length from 17.5 radiation lengths in the forward part of the barrel and the endcap to 16 radiation lengths in the backward part of the barrel. A typical crystal has a front face area of 4.7×4.7 cm^2 and a back face area of 6.0×6.0 cm^2. Scintillation light produced in the calorimeter crystals is detected by photodiodes attached to the crystal back faces. The signal is amplified and sent to a data acquisition board and to the Level 1 trigger system. The average light yield per crystal is about 7,300 photoelectrons/MeV, with a noise level of the order of 900 electrons.

The calorimeter is more than 96% efficient for photons with energies greater than 20 MeV. The 20 MeV energy limit is determined by beam background and the amount of material in front of EMC. The photon energy resolution, determined from direct measurements using a radioactive source at lower energies and Bhabha scattering at higher energies, is parametrized by

$$\frac{\sigma(E)}{E} = \frac{(2.32 \pm 0.30)\%}{E^{1/4}} \oplus (1.85 \pm 0.12)\%, \tag{4.6.1}$$

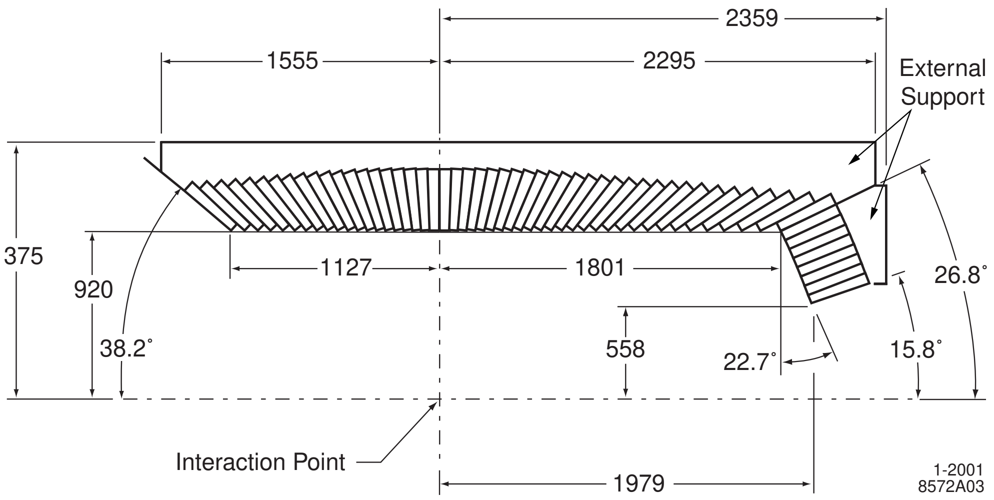


Figure 4.16: The $r-z$ view of the Electromagnetic Calorimeter. Only the top part is shown. It is apparent that the detector θ coverage is asymmetric. All dimensions are in mm.

where E is the photon energy in GeV. At higher photon energies, the resolution is dominated by losses in light collection, arising from leakage or absorption in the material between or in front of the crystals. At lower energies, the energy-dependent term dominates. It arises primarily from fluctuations in the number of photons.

The reconstructed π^0 mass has a width of 6.9 MeV/c^2. The mass resolution is dominated by the energy resolution at lower energies (below 2 GeV). At higher energies, the mass resolution is dominated by the angular resolution. The latter is determined primarily by the transverse crystal size. The angular resolution can be found from analyses of π^0 and η decays to two photons, and is well approximated by

$$\sigma_\theta = \sigma_\phi = \frac{(3.87 \pm 0.07)}{\sqrt{E}} \oplus (0.00 \pm 0.04) \text{ mrad.} \tag{4.6.2}$$

4.7 Instrumented Flux Return

The Instrumented Flux Return (IFR) has been designed to identify muons with momenta higher than ~ 1.0 GeV/c and to detect neutral hadrons, such as K_L^0 mesons.

4.7.1 IFR Layout

The IFR, shown in Fig. 4.17, consists of three parts: the barrel and the forward and backward endcaps. All of them are subdivided into sextants. The active detectors are 806 Resistive Plate Chambers (RPC), located in the gaps between the layers of steel. There are 19 RPC layers in the barrel, and 18 layers in the endcaps. Additionally, there are two layers of cylindrical RPCs between the EMC and the magnet. The thickness of the steel layers ranges from 2 cm in the inner 9 layers to 10 cm in the outermost layers.

RPCs are gas chambers enclosed between bakelite (which is a phenolic polymer) plates. In both the planar and cylindrical RPCs, the gap between the bakelite sheets is 2 mm, and the sheets themselves are 2 mm thick. One of the plates is kept at ≈ 8 kV, and the other is grounded, so that an ionizing particle crossing the gas gap will produce a quenched discharge. The gas used is a mixture of 56.7% Argon, 38.8% Freon-134a, and 4.5% Isobutane.

The efficiencies of the RPCs are routinely evaluated for both colliding beam events and cosmics. 75% of the RPCs are over 90% efficient. The average RPC efficiencies during the 1999-2000 data taking run were $\sim 78\%$ in the barrel and $\sim 87\%$ in the forward endcap region.

4.8 Particle Identification

Particle identification (PID) in the *BABAR* reconstruction code is addressed by developing so-called PID selectors. As a rule, several selectors are developed

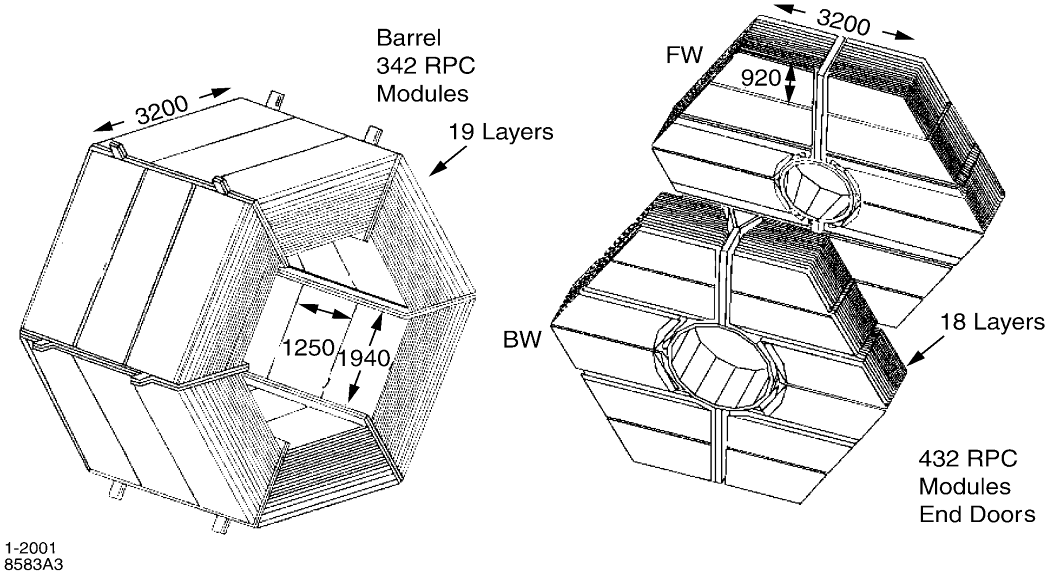


Figure 4.17: The Instrumented Flux Return consists of three sections: the barrel (left), containing 342 RPC modules between 19 layers of steel, and the forward and backward end-doors (right), with a total of 432 RPC modules in 18 layers. 32 more RPCs are placed in the two cylindrical layers (not shown here).

for a given particle species. The selectors differ in the tightness of the particle identification criteria, which range from very loose, characterized by high efficiency and high misidentification rate, to very tight, providing very pure samples at some expense of efficiency. This section describes the criteria used in several hadron and lepton selectors.

4.8.1 Kaon and Pion Identification

The kaon selector used in the analysis described in this dissertation is based on forming combined likelihood ratios using quantities from three *BABAR* subdetectors: SVT, DCH, and DIRC. The quantities used, as well as the momentum range in which a particular subdetector is applicable, are summarized in Table 4.3. For

System	Quantity	Momentum range	Requirements
SVT	dE/dx	$0.025 < p < 0.7$ & $p > 1.5$ GeV/c	> 3 dE/dx hits
DCH	dE/dx	$0.09 <$ p < 0.7 & $p > 1.5$ GeV/c	> 10 dE/dx hits
DIRC	N_γ, θ_C	$0.6 < p < 10.0$ GeV/c	exp. # γ for $e^- > 0$

Table 4.3: Detectors and quantities used in the kaon selector in the different momentum (p) ranges. A set of requirements must be satisfied for a particular detector to be used: at least 3 (10) dE/dx measurements for the SVT (the DCH), and a non-zero expected number of photons, calculated assuming an electron hypothesis, for the DIRC.

low momenta, the discriminating quantity is the ionization loss, dE/dx, from the SVT and the DCH. For higher momenta, the Cherenkov angle θ_C and the number of photons, N_γ, detected by the DIRC are added. The calculated probability for dE/dx from the SVT and DCH, and for the DIRC Cherenkov angle θ_C, is Gaussian. For the number of observed photons in the DIRC, N_γ, the probability includes a Poisson term. The total likelihood $\mathcal{L}$ for a given particle type is a product of the likelihoods from all three subdetectors. Details on the calculation of the likelihoods are given in Ref. [61]. The selection criteria used to separate kaons from pions and protons are:

$$\mathcal{L}_K > r_\pi \mathcal{L}_\pi \text{ \& } \mathcal{L}_K > r_p \mathcal{L}_p, \tag{4.8.1}$$

where $\mathcal{L}_K$ is the total likelihood for the kaon hypothesis; $\mathcal{L}_\pi$, for the pion hypothesis; $\mathcal{L}_p$, for the proton hypothesis; $r_p = 1$; and the value of r_π depends on the momentum range. The values of r_π and other parameters used in the "tight" and "very tight" kaon selector are summarized in Table 4.4.

The efficiency of the selector has been measured using inclusive D^* decays, with $D^{*+} \to D^0\pi^+$, $D^0 \to K^-\pi^+$. Figure 4.18 shows the kaon identification efficiency and the pion mis-identification rate for the very tight selector (in the tracking acceptance region). The average efficiency is $\sim$82%, and the average pion mis-identification rate is $\sim$2%. Pion identification is performed in the veto mode: pions are only required to not be identified by the tight kaon selector as kaons.

		Tight	Very tight
Subsystem		Momentum range (GeV/c)	
SVT		< 0.7	< 0.6
DCH		< 0.7	< 0.6
DIRC		> 0.7	> 0.6
Tight		Very tight	
Momentum (GeV/c)	r_π	Momentum (GeV/c)	r_π
< 2.7	1	$p < 2.5$	3
> 2.7	80	$p > 2.5$	200
$0.5 < p < 0.7$	15	$0.4 < p < 0.7$	20

Table 4.4: Tight and very tight kaon identification requirements. The information from three subsystems, SVT, DCH, and DIRC, is used for particles with momenta within the specified limits. The table also specifies the ratio r_π, defined in Eqn. 4.8.1, for different momentum ranges.

Parameter	Loose	Very tight
E/p	$0.65 < E/p < 5.0$	$0.88 < E/p < 1.3$
# EMC crystals	> 3	> 3
$< dE/dx$	$500.0 < dE/dx < 1000.0$	$500.0 < dE/dx < 850.0$
LAT	-	0.1$<$ LAT < 0.6
A_{42}	-	$0.0 < A_{42} < 0.1$
$\lvert\theta_{\mathrm{C}} - \theta_{\mathrm{C}}^{\mathrm{exp}}\rvert$	-	$< 3\sigma$

Table 4.5: Loose and very tight electron identification requirements. The meaning of the variables used is explained in the text.

4.8.2 Electron Identification

Table 4.5 summarizes the criteria used to select electrons in the two modes that are used in the analysis described in this dissertation: “loose” and “very tight”. The variables used for electron identification are [62]:

1. The ratio of the energy E deposited in the EMC to the track momentum p, E/p.

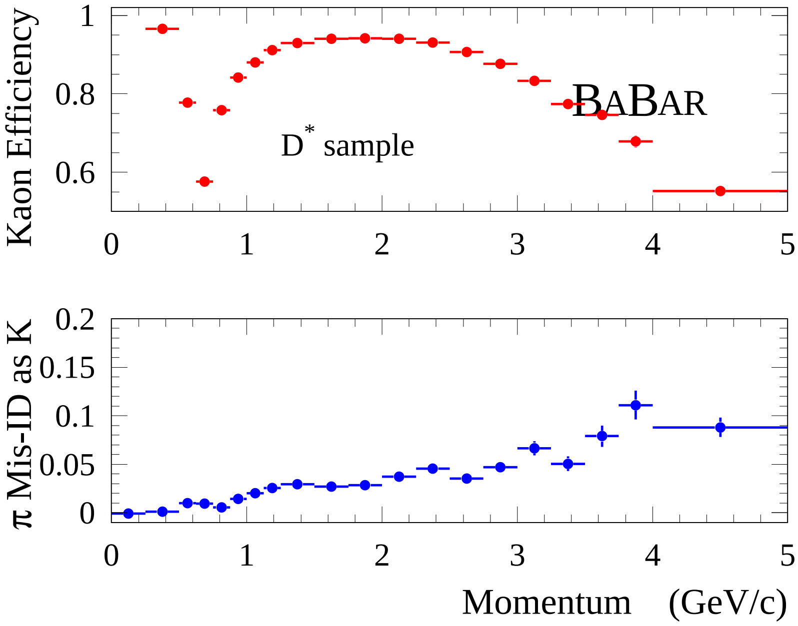


Figure 4.18: Kaon identification efficiency (upper plot) and pion mis-identification rate (lower plot) as a function of track momentum for the very tight kaon identification criteria.

2. The ionization loss in the silicon vertex detector and the drift chamber per unit length, dE/dx .

3. A lateral energy distribution variable, LAT, defined as

$$\mathrm{LAT} = \frac{\sum_{i=3}^{n} E_i r_i^2}{\sum_{i=3}^{n} E_i r_i^2 + E_1 r_0^2 + E_2 r_0^2}, \quad E_1 \geq E_2 \geq ... \geq E_n, \tag{4.8.2}$$

where n is the number of crystals in a shower, r_0 is the average distance between two crystal front faces ($\approx$ 5 cm), E_i is the energy deposited in the $i^{\rm th}$ crystal, and r_i is the distance between the $i^{\rm th}$ crystal and the shower center (the center of gravity). The discriminating power of this variable comes from the fact that electromagnetic showers are typically confined to one or two crystals, so that

LAT is smaller for such showers than for hadronic ones.

4. A Zernike moment, A_{42}, defined as

$$A_{42} = \sum_{r_i \leq R_0}^{4} \frac{E_i}{E} \cdot f_{42}\left(\frac{r_i}{R_0}\right) \cdot e^{-2i\phi_i}, \quad (4.8.3)$$

where $R_0 = 15$ cm, and

$$f_{42}\left(\rho_i \equiv \frac{r_i}{R_0}\right) = 4\rho_i^4 - 3\rho_i^2 \quad (4.8.4)$$

This variable is useful because the electromagnetic showers tend to be more regular than hadronic ones.

5. The measured DIRC Cherenkov angle, θ_{C}, of the e^+ or e^- candidate, which is compared to the expected Cherenkov angle, $\theta_{\mathrm{C}}^{\mathrm{exp}}$ calculated assuming an electron hypothesis.

The electron selector performance is shown in Fig. 4.19. The efficiency is measured using radiative Bhabha events, as well as four-prong $e^+e^- \rightarrow e^+e^-e^+e^-$ events. The pion fake rate is measured using pions from $K_S^0 \rightarrow \pi^+\pi^-$ decays and three-prong τ decays. The efficiency plateau for the very tight selector is 88.1% for tracks with momenta $0.5 < p < 2.0$ GeV/c, and the pion fake rate is 0.15%.

4.8.3 Muon Identification

Table 4.6 lists the criteria used to select muons in both "loose" and "very tight" modes. The variables used for muon identification are [63]:

1. The energy deposited in the EMC, E. The selection based on this variable is only applied to the tracks that have a matching cluster in the EMC.

2. The number of IFR layers in a cluster matched to the muon candidate, $\mathrm{N_L}$.

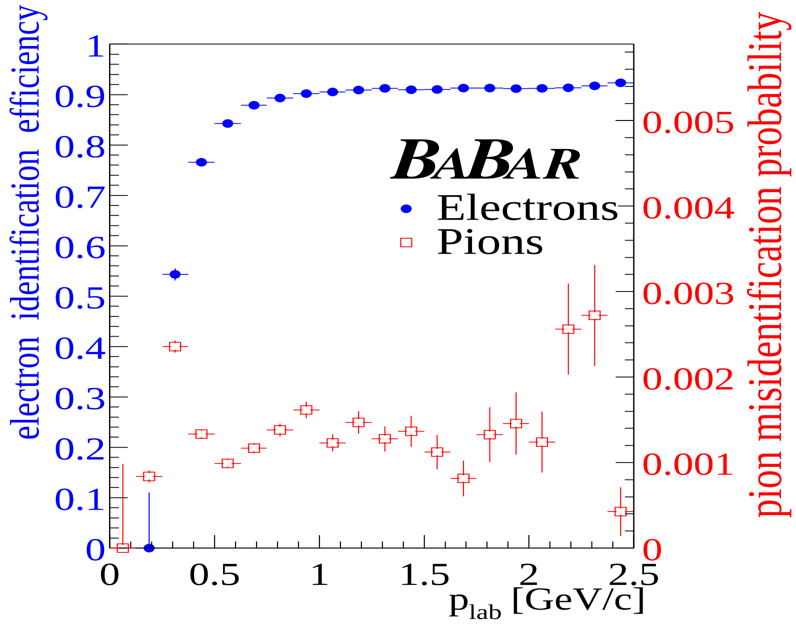


Figure 4.19: Electron identification efficiency and pion fake rate as a function of the track momentum in the laboratory frame for the very tight electron selector.

3. The measured and expected interaction lengths, λ and λ_{exp}, respectively, traversed by the track in the entire detector. $\Delta\lambda$ is the difference between the expected and measured interaction lengths. The expected interaction length is calculated assuming the muon hypothesis. To roughly estimate the contribution of the material in front of the IFR, one can consider a track passing through the EMC and the IFR at $\theta = 90°$. It will go through ~ 46 cm before reaching the IFR and 65 cm in the IFR. Since $\lambda_{\mathrm{CsI}} \approx 37$ cm, and $\lambda_{\mathrm{Fe}} \approx 17$ cm, the expected number of interaction lengths for such a track will be about 5. The material before the iron adds approximately one extra interaction length.

4. The variable χ^2_{trk}, which measures how closely the hit IFR strips in a cluster match the track extrapolation. It is computed using the coordinates of each

Parameter	Loose	Very tight
E	< 0.5 GeV	$0.05 < E < 0.4$ GeV
$\mathrm{N_L}$	≥ 2	≥ 2
$\Delta\lambda$	< 2	< 0.8
λ	> 2	> 2.2
χ^2_{trk}	< 7	< 5
χ^2_{fit}	< 4	< 3
$\mathrm{T_c}$	> 0.2	> 0.34
$\overline{\mathrm{m}}$	< 10	< 8
σ_{m}	< 6	< 4
K veto	-	tight

Table 4.6: Loose and very tight muon identification requirements. The meaning of the variables used is explained in the text.

hit in an IFR layer and the coordinates of the intersection point of the track extrapolated to the same layer. χ^2_{fit} is the χ^2 of a fit of all the hits in a given cluster to a 3$^{\mathrm{rd}}$ order polynomial. The discriminating power of both χ^2_{trk} and χ^2_{fit} comes from the fact that a pion or a kaon can interact in the material and produce a hadronic shower, leading to a more spread out distribution of hits, while a muon is expected to match the track extrapolation reasonably well.

5. The average multiplicity of hit strips per layer, $\overline{\mathrm{m}}$, and its standard deviation, σ_{m}.

6. The variable $\mathrm{T_c}$, defined as:

$$\mathrm{T_c} = \begin{cases} \frac{\mathrm{N_L}}{\mathrm{L_h}-\mathrm{F_h}+1} & \text{if cluster has a hit in the inner RPC} \\ \\ \frac{\mathrm{N_L}}{\mathrm{L_h}-\mathrm{F_h}} & \text{if cluster has no hits in the inner RPC} \end{cases}$$

where $\mathrm{F_h}$ and $\mathrm{L_h}$ are the first and last IFR hit layers in the cluster, respectively. $\mathrm{T_c}$ describes the continuity of the track in the IFR. This variable peaks at 1 for continuous tracks in an ideal detector (with a 100% detector efficiency and geometric acceptance). The selection based on this variable is applicable only to tracks in the angular range $0.3 < \theta < 1$.

7. Very tight muons are required not to be selected by the tight kaon selector, described in Section 4.8.1.

The performance of the muon selector for the very tight selection criteria is shown in Fig. 4.20. The plateau efficiency is ~70.2%, with the average pion (kaon) mis-identification rate ~2.3% (0.7%). The efficiency has been determined using data control samples with $\mu\mu ee$ and $\mu\mu\gamma$ final states, and the mis-identification rate has been determined using pions from three-prong τ decays and $K_S^0 \to \pi^+\pi^-$. Decays in flight constitute $\sim 2\%$ of the pion mis-identification rate.

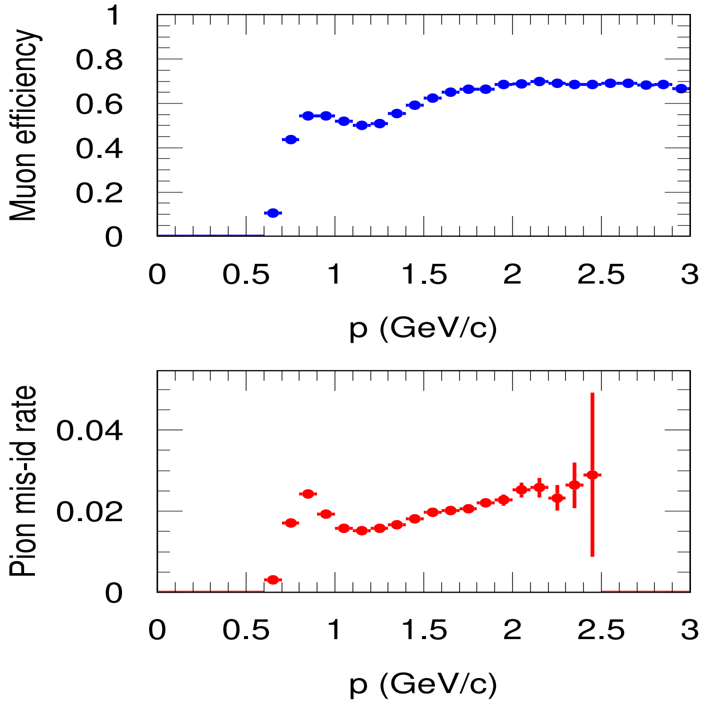


Figure 4.20: Muon efficiency (upper plot) and pion mis-identification rate (lower plot) as a function of the track momentum in the laboratory frame for the very tight selection criteria.

Chapter 5

Data and Monte Carlo samples

This chapter described the data and Monte Carlo simulation samples used in the analysis.

5.1 Data Sample

The data set used for the analysis was collected during the 1999-2000 data taking period, sometimes referred to as "Run 1". This period consisted of 3371 runs with run numbers 9931-17106. The total integrated luminosity of the sample taken on the $\Upsilon(4S)$ peak is 20.6 fb^{-1}, corresponding to $(22.7 \pm 0.4) \times 10^6$ $B\overline{B}$ pairs [64]. This data is divided into two blocks according to the drift chamber high voltage, which was raised from 1900 V to 1960 V approximately in the middle of the run. Table 5.1 lists the integrated luminosities per block. A fraction of the data was taken slightly below the peak of the $\Upsilon(4S)$ resonance, providing a sample of continuum data ($e^+e^- \rightarrow q\overline{q}$, where $q = u, d, s$ or c). The luminosity of this off-resonance sample is 2.6 fb^{-1}. Only good runs, as flagged by the *BABAR* data quality processing, are in the analysis. The main criteria for the goodness of run are the status and performance of all subdetectors, as well as the total luminosity of the run.

Block	DCH HV	$\mathcal{L}dt$ (fb^{-1})
1	1900 V	11.2
2	1960 V	9.4
TOTAL		20.6

Table 5.1: The structure of the 1999-2000 on-peak data sample. Roughly half of the data was taken with the high voltage (HV) of the drift chamber set at 1900 V, and the other half, with the voltage raised to 1960 V.

5.2 Monte Carlo Samples

The GEANT3-based [65] Monte Carlo simulation that is used in this analysis was produced in a simulation cycle called Simulation Production # 3 (SP3). It will thus be often referred to as "SP3 Monte Carlo simulation".

The generic SP3 Monte Carlo samples include $B\overline{B}$ events, with an approximately equal number of $B^0\overline{B}^0$ and B^+B^- events; continuum $c\overline{c}$ events; and continuum *uds* events. These samples are used extensively to perform data/simulation comparisons. For this purpose, the Monte Carlo samples need to be scaled to the data luminosity. The scaling factors F_{sc} are calculated in the following way. If N_{sam} is the number of Monte Carlo events in a given sample (e.g., the number of generic $B\overline{B}$ events), then:

$$\mathrm{F}_{\mathrm{sc}} = \frac{\sigma_{\mathrm{sam}}\,\mathcal{L}}{N_{\mathrm{sam}}}, \tag{5.2.1}$$

where $\mathcal{L}$ is the integrated data luminosity and σ_{sam} is the cross-section for the process in question. The number of events in each generic SP3 Monte Carlo sample, as well as the scaling factors F_{sc} and the cross-sections used in Eqn. 5.2.1, are listed in Table 5.2.

In order to estimate signal efficiencies for a given set of event selection criteria, a number of signal SP3 Monte Carlo samples have been generated. Table 5.3 summarizes the number of events in each of these samples; the branching fractions for $B \to (K, K^*)\ell^+\ell^-$ in the model of Ali, Ball, Handoki, and Hiller [12]; product

SP3 Sample	# Evts	σ	$\mathrm{F_{sc}}$
Generic $B\overline{B}$	8,436,700	1.11	2.72
Continuum $c\overline{c}$	12,086,400	1.35	2.54
Continuum uds	19,475,900	2.04	2.41

Table 5.2: The number of events in the generic $B\overline{B}$ and continuum $udsc$ SP3 Monte Carlo. Cross-sections σ and scaling factors $\mathrm{F_{sc}}$ used for data/Monte Carlo comparison are also given.

branching fractions; and the factors used to scale the samples to the data luminosity. In this case, the scaling factors are defined as:

$$\mathrm{F_{sc}} = \frac{\sigma_{\mathrm{sam}}\ \mathcal{L}\ \mathcal{B}_{\mathrm{prod}}}{N_{\mathrm{sam}}}, \tag{5.2.2}$$

where $\mathcal{B}_{\mathrm{prod}}$ is the product branching fraction for a given decay channel.

Using the branching fractions from the model of Ali *et al.* [12], it is easy to estimate how many signal events will be produced in the Run 1 data. Table 5.3 shows that in all of the modes, counting both electron and muon channels, there will be a total of about 150 produced events. With the signal efficiencies of the order of 10%, this translates to $\sim$ 15 reconstructed events.

Extensive use is also made of several specific final state Monte Carlo samples other than the signal. In particular, signal $B \to J/\psi(K, K^*)$ and $B \to \psi(2S)(K, K^*)$ SP3 Monte Carlo simulation is used for comparing the charmonium control sample in the data with the Monte Carlo. Table 5.4 summarizes the number of events in these Monte Carlo samples and the appropriate scaling factors calculated using Eqn. 5.2.2. The scaling factors refer to a particular J/ψ or $\psi(2S)$ final state, either di-electron or di-muon. The charmonium signal Monte Carlo samples were generated in such a way that in all of the J/ψ modes an equal fraction of the J/ψ mesons decayed into e^+e^- and $\mu^+\mu^-$. For the $\psi(2S)$ modes, the $\psi(2S)$ decayed into the di-lepton final state 30% of the time: 15% into e^+e^- and 15% into $\mu^+\mu^-$.

SP3 Sample	# Evts	$\mathcal{B}/10^{-6}$	$\mathcal{B}_{\text{prod}}/10^{-7}$	# Prod. Evts	$\text{F}_{\text{sc}}/10^{-3}$
$B^+ \to K^+e^+e^-$	18,826	0.57	5.7	12.9	0.7
$B^+ \to K^+\mu^+\mu^-$	20,000	0.57	5.7	12.9	0.7
$B^0 \to K^{*0}e^+e^-$, $K^{*0} \to K^+\pi^-$	18,022	2.3	15.3	34.1	1.9
$B^0 \to K^{*0}\mu^+\mu^-$, $K^{*0} \to K^+\pi^-$	14,000	1.9	12.7	29.5	2.1
$B^0 \to K^0_Se^+e^-$, $K^0_S \to \pi^+\pi^-$	18,000	0.57	2.0	4.4	0.2
$B^0 \to K^0_S\mu^+\mu^-$, $K^0_S \to \pi^+\pi^-$	20,000	0.57	2.0	4.4	0.2
$B^+ \to K^{*+}e^+e^-$, $K^{*+} \to K^0_S\pi^+$ $K^0_S \to \pi^+\pi^-$	19,283	2.3	5.3	11.4	0.6
$B^+ \to K^{*+}\mu^+\mu^-$, $K^{*+} \to K^0_S\pi^+$ $K^0_S \to \pi^+\pi^-$	16,154	1.9	4.3	9.1	0.6
$B^+ \to K^{*+}e^+e^-$, $K^{*+} \to K^+\pi^0$	20,000	2.3	7.6	17.3	0.9
$B^+ \to K^{*+}\mu^+\mu^-$, $K^{*+} \to K^+\pi^0$	18,000	1.9	6.3	14.3	0.8
TOTAL			66.9	150.3	

Table 5.3: The number of generated SP3 signal Monte Carlo events; predicted branching fractions $\mathcal{B}(B \to K^{(*)}\ell^+\ell^-)$ from the model of Ali, Ball, Handoki, and Hiller [12]; product branching fractions $\mathcal{B}_{\text{prod}}$; produced number of events in all of the Run 1 data, and scaling factors used to estimate signal yields for a given set of event selection criteria. In the model of Ref. [12], the branching fractions are calculated in the Standard Model using the Light Cone QCD Sum Rules. The theoretical uncertainty is estimated to be $\sim 35\%$.

SP3 Sample	# Evts	$\mathcal{B}/10^{-4}$	Ref	Prod. $\mathcal{B}/10^{-5}$	F_{sc}
$B^+ \to J/\psi K^+$	44,000	10.0	[66]	5.9	0.06
$B^+ \to J/\psi K^0$, $K^0 \to K^0_S$ $K^0_S \to \pi^+\pi^-$	33,000	8.3	[66]	1.7	0.02
$B^0 \to J/\psi K^{*0}$, $K^{*0} \to K^+\pi^-$	44,000	12.4	[66]	4.9	0.05
$B^+ \to J/\psi K^{*+}$, $K^{*+} \to K^0_S\pi^+$ $K^0_S \to \pi^+\pi^-$	44,000	13.7	[66]	1.8	0.02
$B^+ \to J/\psi K^{*+}$, $K^{*+} \to K^+\pi^0$	44,000	13.7	[66]	2.7	0.03

Table 5.4: The number of generated SP3 signal charmonium Monte Carlo events; branching fractions for $B \to J/\psi(K, K^*)$ with references; product branching fractions; and scaling factors used for data/Monte Carlo comparisons.

Because this analysis requires generic Monte Carlo samples that are several times larger than the data set for the optimization of the event selection criteria and background studies, and such samples of SP3 Monte Carlo were not available at the time the analysis was performed, a fast parametrized Monte Carlo simulation has been developed. This simulation, called *Pravda Monte Carlo*[1], is described in detail in Appendix A. A total of 200 fb^{-1} of generic $B\overline{B}$ and continuum *udsc* events have been generated using Pravda Monte Carlo.

5.3 Note on SP3 Monte Carlo Simulation

The SP3 Monte Carlo simulation does not adequately represent the data in all cases. The following steps have been taken to improve the simulation:

[1]Name invented by Anton Eppich

- The momentum resolution is worse in the data than in the Monte Carlo, likely because the amount of material in the simulation is underestimated. This problem is alleviated by introducing an artificial bias on the transverse momentum p_t and increasing the p_t resolution. The smearing and shifting parameters vary from mode to mode.

- The single photon efficiency is adjusted by randomly removing 2.5% of photons in the Monte Carlo. The photon energy deposition in the EMC is smeared by 1.5% and biased it by 7.5 MeV.

- Particle ID is applied via the so-called "PID killing tables". These tables are extracted from data control samples and summarize particle identification efficiencies and mis-identification rates for electrons, muons, and kaons, in bins of momentum, polar, and azimuthal angle. They work in the following way. First, Monte Carlo truth is used to identify the species of a given particle (e.g., an electron). Then, a random number is generated and compared to the efficiency entry for electrons of a given momentum, θ, and ϕ. If it is smaller than the efficiency, the candidate is considered to be an electron. A drawback of this method is that one can ask only one question for a given particle – e.g., is it an electron? If it was selected as an electron, it is impossible to determine whether it also passed, for example, kaon identification requirements.

  Because the control samples used to create the PID killing tables typically involve low multiplicity events, a small extra correction is needed to describe the PID performance in hadronic events. This correction is determined from the study of inclusive J/ψ production [67].

  The same PID killing tables have been used in Pravda Monte Carlo as in SP3 Monte Carlo. This ensures that the particle identification efficiencies and mis-identification rates are the same in the fast simulation as in the GEANT3-based simulation.

- The tracking efficiencies have also been determined using dedicated tables listing the tracking efficiencies in bins of momentum, θ, ϕ, and the event multiplicity. The method used for creating these tables is described in [68]. Two sets of tables have been created for the two Run 1 data blocks of Table 5.1.

Chapter 6

Kinematic Properties of Signal Events

The properties of $\Upsilon(4S) \rightarrow B\overline{B}$ events containing a signal $B \rightarrow K\ell^+\ell^-$ or $B \rightarrow K^*\ell^+\ell^-$ decay are modeled using the event generator EvtGen [69]. The generator is based on the model described in Ref. [12], which is a Standard Model calculation with the form factors computed using light cone QCD sum rules (LCSR). This chapter describes the details of the implementation of the model in EvtGen, as well as the resulting kinematic properties of the signal events.

6.1 Monte Carlo Event Generator for Signal Events

As was discussed in Section 2.2, the decay rate for the exclusive decays $B \rightarrow (K, K^*)\ell^+\ell^-$ cannot be calculated exactly because the hadronic matrix elements entering amplitude of Eqn. 2.2.2 cannot be evaluated perturbatively and are parametrized in terms of form factors given in Eqns. 2.2.5 and 2.2.6. In the model of Ref. [12], these form factors are parametrized in the following way:

$$F(\hat{s}) = F_0 \exp(c_1\hat{s} + c_2\hat{s}^2 + c_3\hat{s}^3), \qquad (6.1.1)$$

	f_+	f_0	f_T	A_1	A_2	A_0	V	T_1	T_2	T_3
F_0	0.319	0.319	0.355	0.337	0.287	0.471	0.457	0.379	0.379	0.260
c_1	1.465	0.633	1.478	0.602	1.172	1.505	1.482	1.519	0.517	1.129
c_2	0.372	-0.095	0.373	0.258	0.567	0.710	1.015	1.030	0.426	1.127
c_3	0.782	0.591	0.700	0	0	0	0	0	0	0

Table 6.1: Central values of the parameters used to parametrize form factors in the generator model. The quantities were calculated in Ref. [12] using light cone QCD sum rules.

$\mathcal{C}_1$	$\mathcal{C}_2$	$\mathcal{C}_3$	$\mathcal{C}_4$	$\mathcal{C}_5$	$\mathcal{C}_6$	$\mathcal{C}_7^{\text{eff}}$	$\mathcal{C}_9$	$\mathcal{C}_{10}$
-0.248	1.107	0.011	-0.026	0.007	-0.031	-0.313	4.344	-4.669

Table 6.2: The values of Wilson coefficients used in the generator model. The quantities were calculated in Ref. [12] using light cone QCD sum rules.

where $\hat{s} \equiv q^2/m_B^2$, and F_0, c_1, c_2, and c_3 are parameters evaluated in Ref. [12]. The central values of these parameters are summarized in Table 6.1, and Figs. 6.1 and 6.2 show the behavior of the form factors as a function of $\hat{s}$ for both $B \to K$ and $B \to K^*$ transitions.

The values of the Wilson coefficients entering Eqn. 2.2.2 that were used in the generator model are given in Table 6.2.

6.2 The Kinematics of Signal Events

The solid line in the plots of Fig. 2.5 shows the distribution of the dilepton invariant mass in $B \to (K, K^*)\mu^+\mu^-$ decays in the model of Ali *et al.* [12]. The plots show both short- and long- distance contributions, the latter arising from the $c\bar{c}$ resonances. One can see that in the regions of $m_{\ell^+\ell^-}$ near the J/ψ and $\psi(2S)$ resonances interference occurs between the penguin amplitude and the amplitude for charmonium production. This interference enhances the rate on the low-mass side of the resonance and suppresses it on the high side. These effects are modeled in

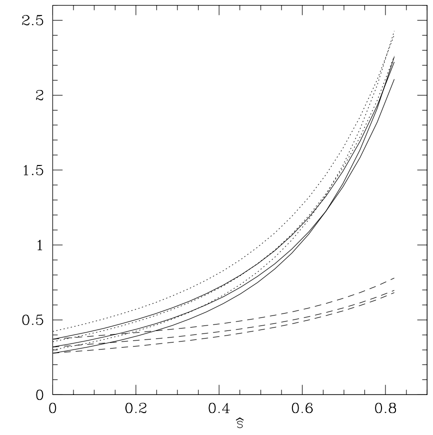


Figure 6.1: The form factors for the $B \to K$ transition as a function of $\hat{s} \equiv q^2/m_B^2$. The solid curve corresponds to f_+; the dotted curve, to f_T; and the dashed curve, to f_0. The form factors were evaluated in Ref. [12] using the light cone QCD sum rule approach. Both central values and theoretical uncertainties are shown. Figure taken from Ref. [12].

(a) (b)

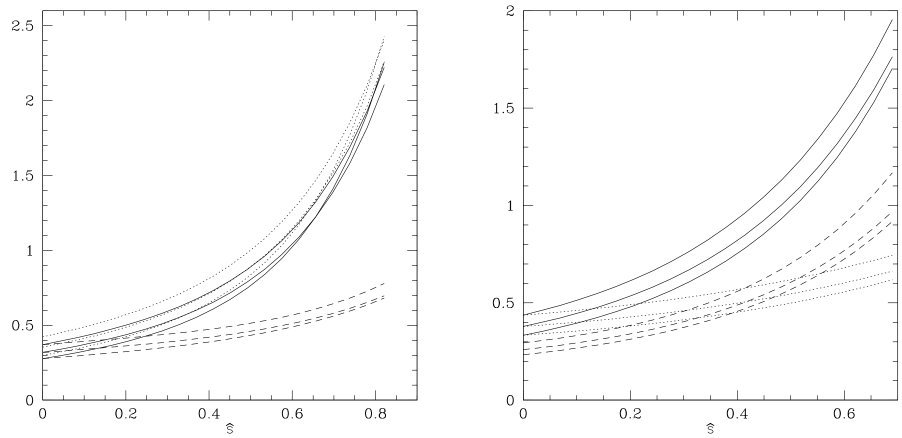


Figure 6.2: The form factors for the $B \to K^*$ transition as a function of $\hat{s} \equiv q^2/m_B^2$. (a) The solid, dotted, dashed and short long dashed curves correspond to V, A_0, A_1, A_2, respectively. (b) The solid, dotted and dashed curves correspond to T_1, T_2, T_3, respectively. The form factors were evaluated in Ref. [12] using the light cone QCD sum rule approach. Both central values and theoretical uncertainties are shown. Figure taken from Ref. [12].

Ref. [12], but not in the Monte Carlo generator. However, since $B \to J/\psi(K, K^*)$ and $B \to \psi(2S)(K, K^*)$ are major backgrounds in the analysis, most of these mass regions are eliminated, as described in Section 7.11. Small residual interference effects are expected for masses close to the exclusion regions, but these are ignored in the present study. Further enhancements in the rate, along with related interference effects, are expected from charmonium resonances with masses above the $\psi(2S)$. These effects are relatively small compared to the expected electroweak penguin contribution, and they are also ignored in this analysis.

The Dalitz plot and q^2 distributions resulting from the matrix elements computed in this model are shown in Fig. 6.3 for $B \to (K, K^*)\mu^+\mu^-$, and in Fig. 6.4 for $B \to (K, K^*)e^+e^-$. The Dalitz plot variables are q^2, the invariant dilepton mass squared, and E_{lep}, the lepton energy in the B meson rest frame, where the positive lepton is selected for the case of $B^+ \to K^+\ell^+\ell^-$ and $B^0 \to K^{*0}\ell^+\ell^-$ (the negative lepton is selected for the charge-conjugate states). The kinematics of the $B \to K\ell^+\ell^-$ decay are relatively simple and resemble the semileptonic decay $B^+ \to \overline{D}^0\, \ell^+\nu_\ell$. The decay rate is largest at $q^2 \equiv m_{\ell^+\ell^-} = 0$, and it decreases with increasing q^2, as shown in Figs. 6.3c and 6.4c. The distribution of the lepton decay angle θ, defined as a polar angle in the rest frame of the $\ell^+\ell^-$ system measured with respect to the direction of motion of the dilepton system in the B rest frame, and the lepton energy spectrum in the B rest frame, are shown in Figs. 6.5a and 6.5b, respectively, for $B \to K\mu^+\mu^-$. The leptons exhibit a $\sin^2\theta$ angular distribution. It is evident that for $B \to K\ell^+\ell^-$ there is no forward-backward asymmetry in the lepton angular distribution.

For $B \to K^*\ell^+\ell^-$, the q^2 distribution peaks at large values, as shown in Figs. 6.3d and 6.4d for $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$, respectively. The large values of q^2 correspond to a slow K^* recoil in the B rest frame. There is also a pole in the amplitude at $q^2 = 0$, where the photon becomes real. Comparing Figs. 6.3d and 6.4d, one can see that the pole is about a factor of three larger for the $B \to K^*e^+e^-$ mode than for the $B \to K^*\mu^+\mu^-$ mode. The decay $B^+ \to K^+\ell^+\ell^-$ cannot have a pole at $q^2 = 0$ because $B^+ \to K^+\gamma$ is forbidden by

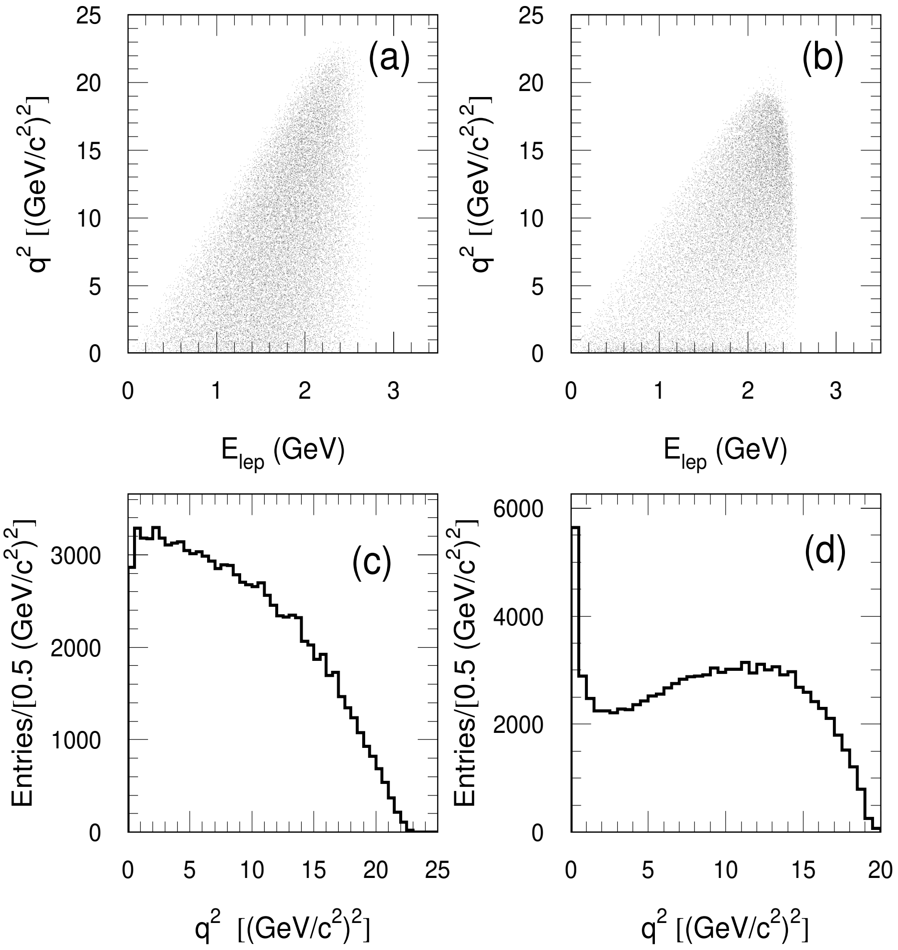


Figure 6.3: Generator-level Monte Carlo distributions for $B^+ \to K^+\mu^+\mu^-$. (a) Dalitz plot for $B^+ \to K^+\mu^+\mu^-$. (b) Dalitz plot for $B^0 \to K^{*0}\mu^+\mu^-$. The Dalitz plot variables are E_{lep}, the energy of the lepton with the same charge as the kaon in the B rest frame, and the invariant dilepton mass squared $q^2 \equiv m^2_{\mu^+\mu^-}$. (c) q^2 distribution for $B^+ \to K^+\mu^+\mu^-$. (d) q^2 distribution for $B^0 \to K^{*0}\mu^+\mu^-$. Note the pole at $q^2 = 0$.

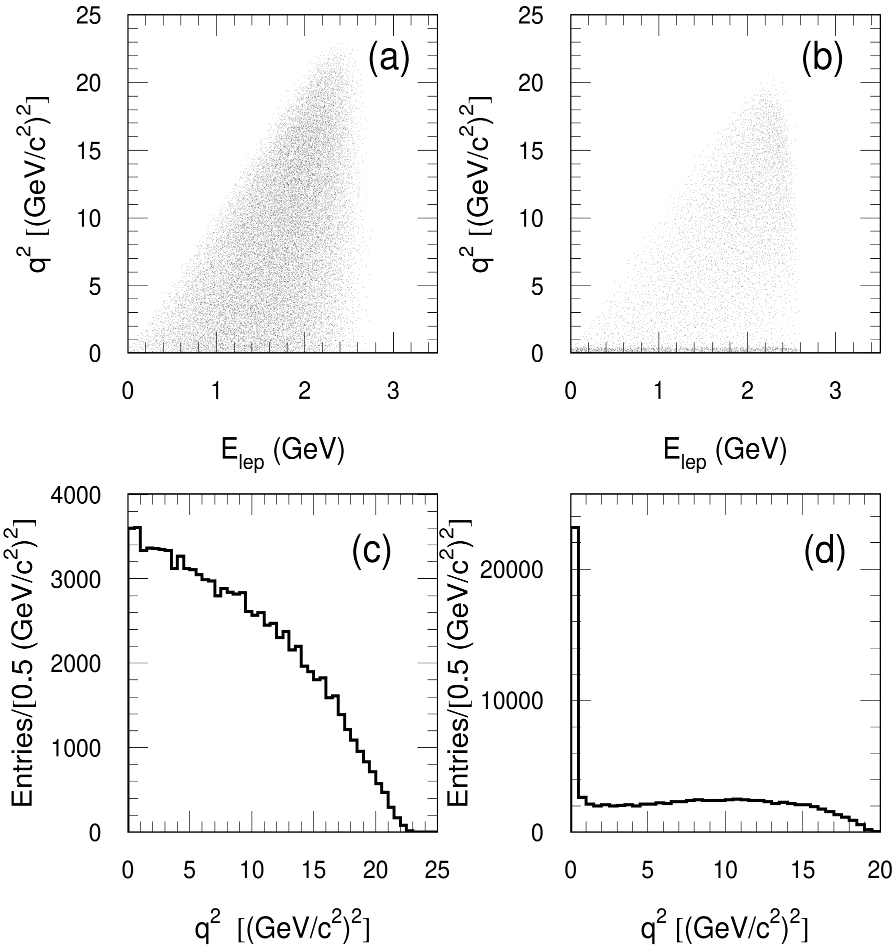


Figure 6.4: Generator-level Monte Carlo distributions for $B^+ \to K^+e^+e^-$. (a) Dalitz plot for $B^+ \to K^+e^+e^-$. (b) Dalitz plot for $B^0 \to K^{*0}e^+e^-$. The Dalitz plot variables are $E_{\rm lep}$, the energy of the lepton with the same charge as the kaon in the B rest frame, and the invariant dilepton mass squared $q^2 \equiv m^2_{e^+e^-}$. (c) q^2 distribution for $B^+ \to K^+e^+e^-$. (d) q^2 distribution for $B^0 \to K^{*0}e^+e^-$. Note the pole at $q^2 = 0$.

angular momentum conservation. Also unlike $B^+ \to K^+\ell^+\ell^-$, this channel has a non-trivial lepton forward-backward asymmetry (Fig. 6.5c). For relatively large q^2, where $V-A$ coupling dominates, the kinematics of the decay start to resemble those for $B^0 \to D^{*-}\ell^+\nu_\ell$, discussed extensively in Refs. [37] and [38]: for $B^0 \to K^{*0}\,\ell^+\ell^-$, the positron tends to be harder than the electron, and for the charge-conjugate mode it is the electron that tends to have higher momentum. The lepton energy spectrum for $B \to K^*\mu^+\mu^-$ in the B rest frame is also shown in Fig. 6.5d.

For reference, the Dalitz plots after applying the analysis selection criteria are shown in Fig. 7.16 for the charged track modes, and in Fig. 7.17 for the modes with a K^0_S or a π^0.

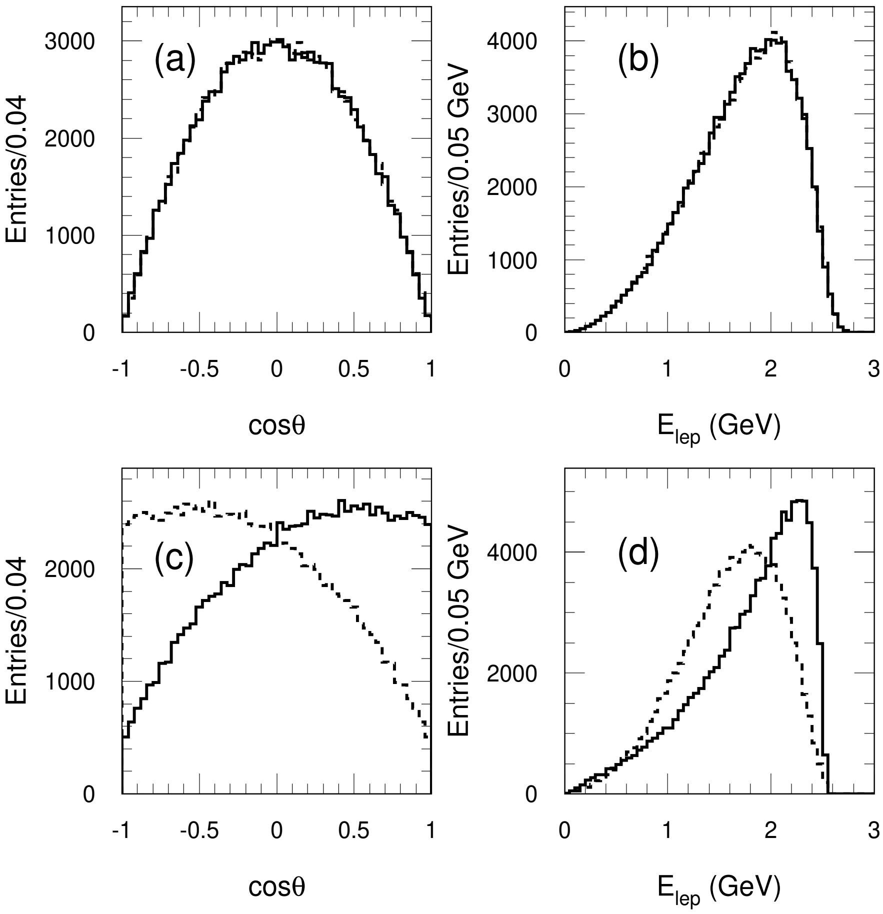


Figure 6.5: Generator-level Monte Carlo simulated distributions for signal processes. (a) Decay angle for μ^+ (solid line) and μ^- (dashed line) in the $B^+ \to K^+\mu^+\mu^-$ decay. The distributions are essentially identical, indicating that there is no lepton forward-backward asymmetry. (b) μ^+ (solid line) and μ^- (dashed line) energy in the B rest frame in the $B^+ \to K^+\mu^+\mu^-$ decay. (c) Decay angle for μ^+ (solid line) and μ^- (dashed line) in the $B^0 \to K^{*0}\mu^+\mu^-$ decay. Unlike $B^+ \to K^+\mu^+\mu^-$, the forward-backward asymmetry is non-zero. (d) μ^+ (solid line) and μ^- (dashed line) energy in the B rest frame in the $B^0 \to K^{*0}\mu^+\mu^-$ decay.

Chapter 7

Event Selection

This chapter describes the criteria used to reconstruct the following B decay channels: $B^+ \to K^+\ell^+\ell^-$, $B^0 \to K^{*0}\ell^+\ell^-$ ($K^{*0} \to K^+\pi^-$), $B^0 \to K^0_S\ell^+\ell^-$, and $B^+ \to K^{*+}\ell^+\ell^-$, where $K^{*+} \to K^0_S\pi^+$ and $K^{*+} \to K^+\pi^0$. The l is an electron or a muon.

7.1 Note on Blind Analyses

There are a number of potential problems associated with low statistics analyses.

- Optimizing the event selection criteria by looking at the signal region in data may lead to enhancing a statistical fluctuation and therefore the observation of a "signal" that is not reproducible in other data samples.

- Using the sideband region in data for optimizing the event selection criteria can also be dangerous. This procedure may lead to underestimating the background if the same sideband is intended to be used for estimating the number of background events in the signal region, One method of avoiding any bias in the event selection is the technique of a blind analysis [70]. In this approach, the signal region in data is kept hidden until all of the selection criteria have been finalized and fixed. Keeping the data sideband regions intended to be used for background extrapolation hidden as well ensures that there will be no bias in the determination of the background.

In a blind analysis, all of the selection criteria must be tuned either on simulated data or on a data sample excluded from the final measurement. It is of course possible that the event selection criteria optimized on such samples will not be the most optimal, but they are guaranteed not to be biased. In order to verify that simulated data adequately reproduce the real data, it is useful to compare the efficiencies of various selection criteria and the shapes of kinematic distributions in the simulated data and in the so-called control samples in data. The control samples are data sub-samples that are excluded from the final measurement.

The search described in this dissertation is performed as a blind analysis.

7.2 Main Background Categories

$B \to (K, K^*)\ell^+\ell^-$ decays have a very clean experimental signature, with two leptons, a kaon or a K^*, and a fully reconstructed B. However, even though the signature for the signal is clear, backgrounds are still a significant problem because the signal is very rare. The main challenge of the analysis lies in understanding and suppressing the backgrounds. This section is devoted to the discussion of backgrounds and various ways of suppressing them.

- The single most significant background for this analysis is the charmonium decays $B \to J/\psi(K, K^*)$ and $B \to \psi(2S)(K, K^*)$, with J/ψ or $\psi(2S) \to \ell^+\ell^-$. Such events must be rejected with a very high efficiency, since their topology is the same as that of the signal events.

- Apart from $B \to J/\psi(K, K^*)$ and $B \to \psi(2S)(K, K^*)$ events, there are less significant backgrounds that can peak in the signal region. An example is $B^+ \to \overline{D}^0\pi^+$, with $\overline{D}^0 \to K^+\pi^-$. If both pions are misidentified as leptons (most often muons), such events will have all the characteristics of the signal.

- Combinatorial background, arising from combining random leptons and kaons in $B\overline{B}$ and continuum decays, does not peak in the signal region. After applying most of the analysis selection criteria, lepton candidates in the continuum sample arise from semileptonic D decays or from hadrons faking a lepton signature. In the $B\overline{B}$ sample, leptons arise from semileptonic B decays and from hadronic fakes, although there is also a small residual contribution from charmonium decays. In both continuum and $B\overline{B}$ events, electrons and positrons coming from γ conversions in material and π^0 Dalitz decays also contribute at a low level.

The background categories, as well as the effect of the analysis selection criteria on them, are discussed further in Section 7.16.

7.3 Event Selection Overview

There are several levels on which events are selected for this analysis. In the order of increasing tightness, the event selection criteria include skim, code, and ntuple selection requirements. Skim and code selections are described in detail in this section, and ntuple cuts are discussed here only briefly. They are described in more detail in Section 7.15.

- **Skim event selection criteria.** The loosest event selection, summarized in Table 7.1, occurs at the skim level. The first four requirements in Table 7.1 are used to select events consistent with coming from $B\overline{B}$ decays as opposed to continuum. The variables used are defined as follows. The BGFMultiHadron tag requires events to have at least 3 charged tracks and a normalized second-order Fox-Wolfram moment [57] below 0.98. The Fox-Wolfram moment is an

Variable	Value
BGFMultiHadron tag	True
R2All	≤ 0.5
Number of charged tracks	≥ 4
L3DchOut = True or L3EmcOut = True	
Number of loose leptons	≥ 2

Table 7.1: Skim event selection criteria. The first four requirements are used to select events consistent with $B\overline{B}$ decays. The events are also required to have at least two loosely identified leptons, which can be two electrons, two muons, or an electron-muon pair.

event topology variable, allowing to discriminate between the more spherical $B\overline{B}$ events and the more jet-like continuum. The Fox-Wolfram moment entering the BGFMultiHadron tag is calculated using only charged tracks. We also require the Fox-Wolfram moment of the event calculated using both charged tracks and unmatched neutral clusters, R2All, to be less than 0.5. L3DchOut and L3EmcOut are Level 3 trigger tags, set to true if the event was selected by the trigger on the basis of the DCH or EMC response, respectively.

Events are also required to have at least two loosely identified oppositely charged leptons. The loose lepton identification criteria were given in Table 4.5 for electrons and Table 4.6 for muons. The leptons can be either two electrons, two muons, or an electron-muon pair. Electrons are required to have momenta greater than 0.5 GeV/c and muons, greater than 1.0 GeV/c, in the laboratory frame.

- **Code event selection criteria.** Events that passed the skim criteria are considered further in the code. Only loosely identified electron (muon) candidates with momenta greater than 0.5 GeV/c (1.0 GeV/c) in the laboratory frame are used at this level. No particle identification criteria are used for either kaon or pion candidates. Leptons are combined with a K or a K^* candidate to form

Variable	Value
m_{ES}	> 5.0 GeV/c^2
$\lvert\Delta E\rvert$	< 0.8 GeV
Electrons and muons satisfy loose selection criteria.	

Table 7.2: Code event selection criteria: a grand sideband region is selected in the m_{ES} vs. ΔE plane.

a B candidate. Two kinematic variables are calculated for this candidate: the beam-energy substituted mass m_{ES} and the energy difference ΔE:

$$\begin{aligned} m_{\mathrm{ES}} &\equiv \sqrt{\frac{\left(\frac{1}{2}s + \mathbf{p_0}\,\mathbf{p_1}\right)^2}{E_0^2} - p_1^2}, \\ \Delta E &\equiv \frac{2q_0\,q_1 - s}{2\sqrt{s}}, \end{aligned} \tag{7.3.1}$$

where $q_0 \equiv \{E_0; \mathbf{p_0}\}$ is the four-momentum of the initial state for the process $e^+e^- \to B\overline{B}$, $q_1 \equiv \{E_1; \mathbf{p_1}\}$ and $q_2 \equiv \{E_2; \mathbf{p_2}\}$ are the four-momenta of the B mesons, and $s \equiv q_0^2 = (q_1 + q_2)^2$ is the center of mass energy squared. These variables are discussed in detail in Ref. [71].

If m_{ES} and ΔE satisfy the criteria listed in Table 7.2, the event is stored in the ntuple for further analysis.

For continuum events the center of mass energy is lower, so m_{ES} constructed by the prescription given above is always below $m_0^{\mathrm{cont}} \sim 5.27$ GeV/c^2. Therefore, an explicit m_{ES} correction is made for events with beam energy E_{beam} less than 5.28 GeV, bringing the m_{ES} endpoint to $m_0^{B\overline{B}} = 5.29$ GeV/c^2:

$$m_{\mathrm{ES}} \;\to\; m_{\mathrm{ES}} + (5.29 - E_{\mathrm{beam}}) \tag{7.3.2}$$

The shift of Eqn. 7.3.2 produces the same effect as explicit re-scaling by $(m_0^{B\overline{B}} - m_0^{\mathrm{cont}})/m_B \sim 20$ MeV/5.29 GeV, which is of the order of 10^{-3}.

- **Ntuple event selection criteria.** The event selection criteria applied at the ntuple level are described in detail in Section 7.15. This section presents only a brief overview of the main ideas.

  Both leptons and hadrons are required to pass a tighter set of particle identification requirements comparing to what was required at the skim and code selection level. Electrons are also required to pass the γ conversion veto requirement described in Section 7.10.

  The K^{*0} in the $B^0 \to K^{*0}\ell^+\ell^-$ channel is reconstructed in the $K^+\pi^-$ final state. The K^{*+} in the $B^+ \to K^{*+}\ell^+\ell^-$ decays is reconstructed in both $K^0_S\pi^+$ and $K^+\pi^0$ final states. The mass of the K^* candidate is required to be within 75 MeV/c^2 of the mean K^* mass. The charged kaon daughters of the K^* are required to be identified as kaons, and the pion candidates must satisfy the kaon veto requirement, which means that they must not be identified as kaons. The selection of K^0_S and π^0 candidates is described in Sections 7.6 and 7.7, respectively. The K^0_S is reconstructed in the charged track mode only: $K^0_S \to \pi^+\pi^-$.

  Charmonium background, $B \to J/\psi(K, K^*)$ and $\psi(2S)(K, K^*)$, is suppressed by applying a veto in the ΔE vs. $m_{\ell^+\ell^-}$ plane, as described in Section 7.11. Charmonium events can pass this veto if one of the leptons and the kaon are mis-identified as each other ("swapped"). There is also a feed-up from the $J/\psi K$ and $\psi(2S)K$ modes to the $B \to K^*\ell^+\ell^-$ channels that is not suppressed by this veto. Both the swapped candidates and the feed-up contributions are explicitly forbidden.

  The combinatorial background from continuum events is suppressed using a four-variable Fisher discriminant described in Section 7.12, and the $B\overline{B}$ combinatorial background is suppressed using a likelihood ratio, which combines several vertexing variables and the missing energy of the event in the center of mass frame. This is described in Section 7.13.

Finally, a signal box is chosen in the "grand sideband" (the ΔE vs. m_{ES} plane): $5.272 < m_{\mathrm{ES}} < 5.286$ GeV (a 3σ selection, obtained from signal Monte Carlo) and -0.11 (-0.07) $< \Delta E < 0.05$ GeV for the electron (muon) channels. The cut in ΔE in the electron channels is asymmetric in order to account for the effects of radiation. For the muon channels, where the effects of bremsstrahlung are not expected to be large, the slight asymmetry in the ΔE cut is introduced to match the data. The nature of the asymmetry in the data is not yet understood. It must be emphasized that this definition of the signal box is used only for the purpose of tuning the event selection criteria. The actual extraction of signal and background yields is performed using a two-dimensional fit in the ΔE vs. m_{ES} plane, described in Chapter 9.

Figure 7.1 shows the distribution of Monte Carlo truth tagged signal events in the ΔE vs. m_{ES} plane for the four charged track modes, and Fig. 7.2, for the modes with a K^0_S or a π^0. Three distinct regions are selected in this plane: region A, which is the signal region; region B, the sideband that can be used for extracting the number of background events in the signal region; and region C, which is called the "large sideband". Both region A and region B are blind in the data, whereas region C is not: it is one of the control samples.

7.4 Particle Identification

At the initial stage of the analysis, the leptons are required to pass the loose identification requirements, and there are no particle identification requirements on either charged kaons or pions. At the ntuple level, the lepton identification requirements are made as tight as possible. The very tight lepton identification criteria are described in Section 4.8.2 (electrons) and Section 4.8.3 (muons). At this level, particle identification requirements are also imposed on charged kaons and pions from the K^*. Kaons must pass the very tight particle identification criteria described in Section 4.8.1. For pions, particle identification is performed in the veto mode: they

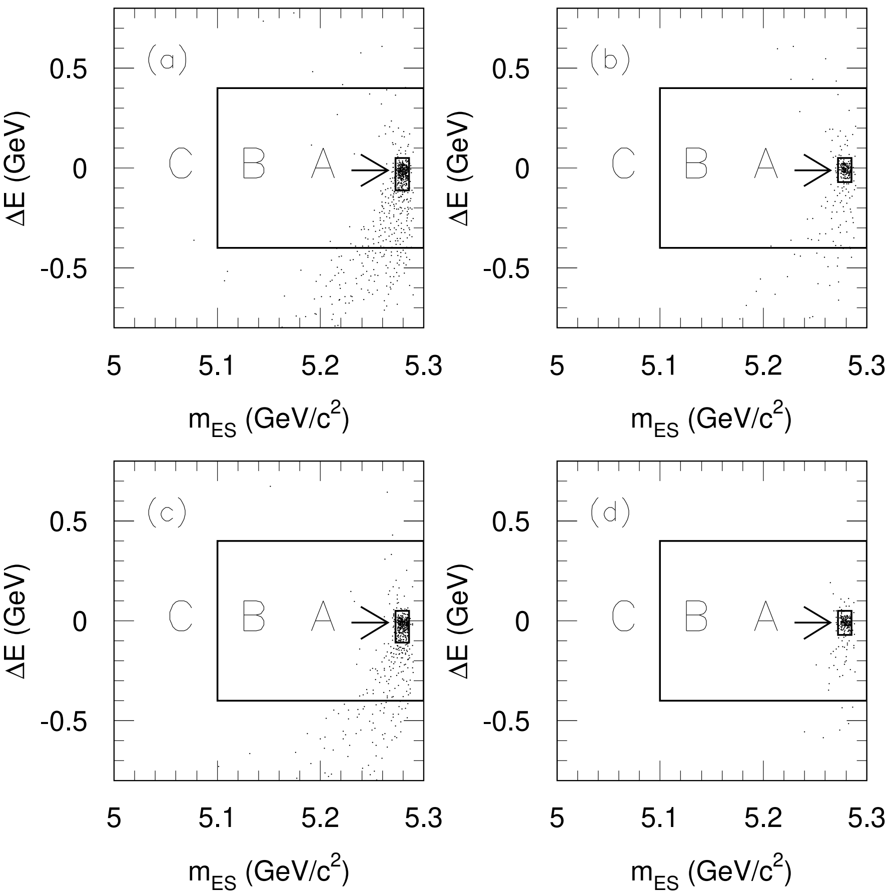


Figure 7.1: ΔE vs. m_{ES} distributions for truth tagged signal SP3 Monte Carlo. The four charged track modes are shown: (a) $B^+ \to K e^+ e^-$, (b) $B^+ \to K \mu^+ \mu^-$, (c) $B^0 \to K^{*0} e^+ e^-$, and (d) $B^0 \to K^{*0} \mu^+ \mu^-$. Region A is the signal region. Region B can be used for extrapolating the background into the signal region. Only a part of it, with $|\Delta E| < 0.25$ GeV and $m_{\mathrm{ES}} > 5.2$ GeV/c^2, is in fact used in the fitter described in Chapter 9. Both A and B are blinded in the data. Region C is the "large sideband", and is not blinded in data.

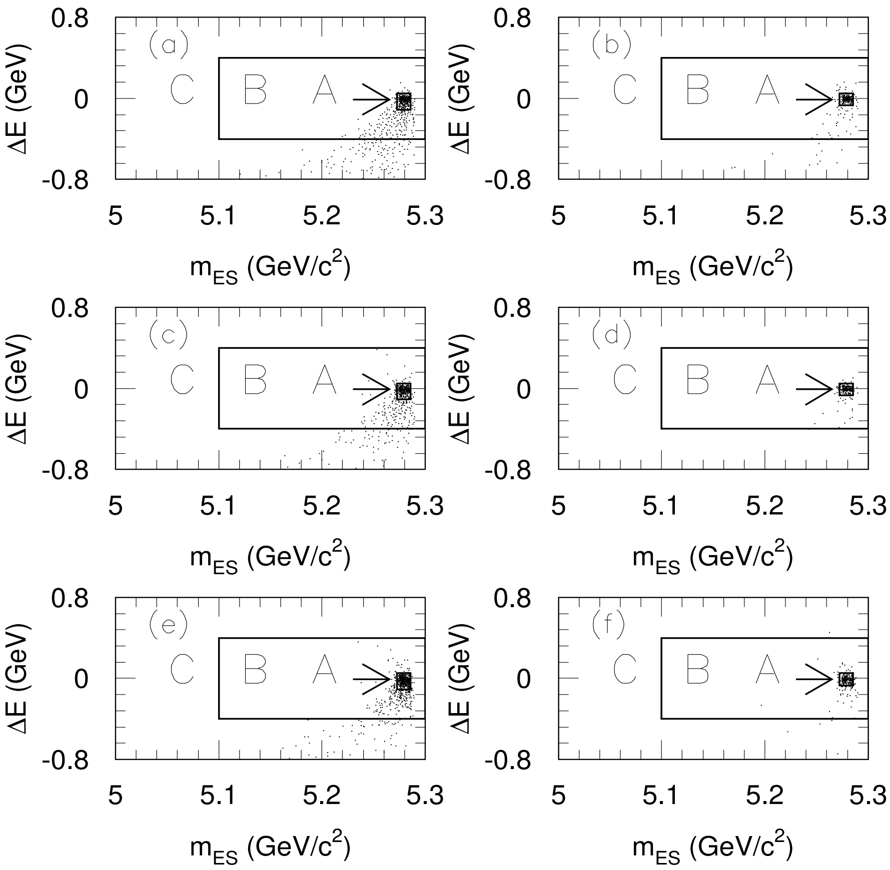


Figure 7.2: ΔE vs. m_{ES} distributions for truth tagged signal SP3 Monte Carlo. The six modes with a K_S^0 or a π^0 are shown: (a) $B^0 \to K_S^0 e^+e^-$, (b) $B^0 \to K_S^0\mu^+\mu^-$, (c) $B^+ \to K^{*+}e^+e^-$ with $K^{*+} \to K_S^0\pi^+$, (d) $B^+ \to K^{*+}\mu^+\mu^-$ with $K^{*+} \to K_S^0\pi^+$, (e) $B^+ \to K^{*+}e^+e^-$ with $K^{*+} \to K^+\pi^0$, and (f) $B^+ \to K^{*+}\mu^+\mu^-$ with $K^{*+} \to K^+\pi^0$. Region A is the signal region. Region B can be used for extrapolating the background into the signal region. Only a part of it, with $|\Delta E| < 0.25$ GeV and $m_{\mathrm{ES}} > 5.2$ GeV$/c^2$, is in fact used in the fitter described in Chapter 9. Both A and B are blinded in the data. Region C is the "large sideband", and is not blinded in data.

Variable	GoodTracksVeryLoose	GoodTracksLoose
p_t	> 0.0 GeV/c	> 0.1 GeV/c
# DCH hits	> 0	> 12
$\lvert x_0 \rvert$	< 1.5 cm	< 1.5 cm
$\lvert z_0 \rvert$	< 10.0 cm	< 10.0 cm

Table 7.3: The track selection criteria for GoodTracksVeryLoose and GoodTracksLoose. The meaning of the variables used is explained in the text.

are required not to be identified as tight kaons. There is no particle identification for the pions from K_S^0 mesons.

The main reason for the tightest possible lepton and kaon particle identification criteria is the need to keep the background from hadrons faking a lepton signature and pions faking kaons low.

7.5 Track Selection

In this analysis, three types of tracking criteria are used for charged particles: ChargedTracks, GoodTracksVeryLoose, and GoodTracksLoose. ChargedTracks are all tracks found by *BABAR* tracking, with no additional requirements. The criteria used to select GoodTracksVeryLoose and GoodTracksLoose are summarized in Table 7.3. The variables used in the selection are:

- The transverse momentum of the track, p_t.
- The number of drift chamber hits on the track, # DCH hits.
- The absolute value of the closest distance between the track and the beam spot in the $x - y$ plane, $|x_0|$.
- The absolute value of the closest distance between the track and the beam spot in the z direction, $|z_0|$.

The leptons are required to be GoodTracksLoose, and the hadrons, GoodTracksVeryLoose, except for the charged pions from the K_S^0 mesons, which remain ChargedTracks in order to avoid the loss of efficiency associated with the impact parameter requirements.

7.6 Selection of K_S^0 Mesons

K_S^0 mesons are used in two modes: $B^0 \to K_S^0 \ell^+ \ell^-$, and $B^+ \to K^{*+} \ell^+ \ell^-$ with $K^{*+} \to K_S^0 \pi^+$. The K_S^0 mesons are reconstructed in the charged pion channel only. The reconstruction algorithm proceeds as follows. Two oppositely charged pions are combined and form a K_S^0 candidate, first by attempting to find a common vertex, and if this fails, by four-momentum addition. All K_S^0 candidates with masses $0.46 < m_{K_S^0} < 0.54$ GeV/c^2 are stored in the ntuple. A number of other requirements, summarized in Table 7.4, are applied at the ntuple level. The quantities used in the selection are shown in Fig. 7.3. They include the mass of the K_S^0 candidate $m_{K_S^0}$; the flight distance, defined as the distance between the K_S^0 vertex and the reconstructed candidate B vertex; $\cos K_S$, defined as the cosine of the angle between the flight distance vector and the K_S^0 momentum; and the probability of the K_S^0 vertex. The selection requirement on the flight distance uses the fact that K_S^0 mesons are long-lived. $\cos K_S$ is roughly 1 for signal K_S^0, while it can be -1 for background. This variable is so sharply peaked at $\pm$ 1 because the K_S^0 itself is used to form the B vertex. If such a vertex could be formed, the K_S^0 must be either approximately aligned or anti-aligned with the flight distance vector.

7.7 Selection of π^0 Candidates

There is only one mode with π^0 mesons in this analysis: $B^+ \to K^{*+} \ell^+ \ell^-$ with $K^{*+} \to K^+ \pi^0$. The π^0 candidates are formed as described in Ref. [72], using all neutral bumps with $E_\gamma > 30$ MeV, and LAT < 0.8 (LAT is defined in Section 4.8.2).

Variable	Value
$m_{K^0_S}$	$\|m_{K^0_S} - 0.498)\| < 0.0093$ GeV/c^2
flight	> 0.1 cm
$cosK_S$	> 0
K^0_S vtx prob	> 0.001

Table 7.4: K^0_S selection criteria. The meaning of the variables used is explained in the text.

The candidates are required to have a raw mass $0.10 < m_{\pi^0} < 0.16$ GeV/c^2. Later on, at the ntuple level, this requirement is made tighter: $|m_{\pi^0} - 0.133| < 0.0175$ GeV/c^2, as shown in Fig. 7.4. Only resolved π^0 candidates are used, which means that the constituent photons form two distinct bumps.

7.8 Selection of K^* Candidates

K^* candidates are reconstructed in three different modes: $K^{*0} \to K^+\pi^-$, $K^{*+} \to K^0_S\pi^+$, and $K^{*+} \to K^+\pi^0$. The reconstruction proceeds as follows. No particle identification criteria are applied to either charged kaon or pion candidates, but the four-momenta of the charged tracks are re-scaled assuming a kaon or a pion mass, respectively. All oppositely charged kaons and pions are combined and form K^{*0} candidates. For the K^{*+}, K^0_S and π^0 candidates are combined with charged pions and kaons, respectively. The selection of K^0_S mesons is described in Section 7.6, and the selection of π^0 candidates, in Section 7.7. In all modes, the particles are combined using four-momentum addition. The resulting K^* candidate is selected if its mass $0.7 < m_{K^*} < 1.1$ GeV/c^2. Later on, at the ntuple level, the mass requirement is made tighter: m_{K^*} must be within 75 MeV/c^2 of the mean K^* mass (0.892 GeV/c^2).

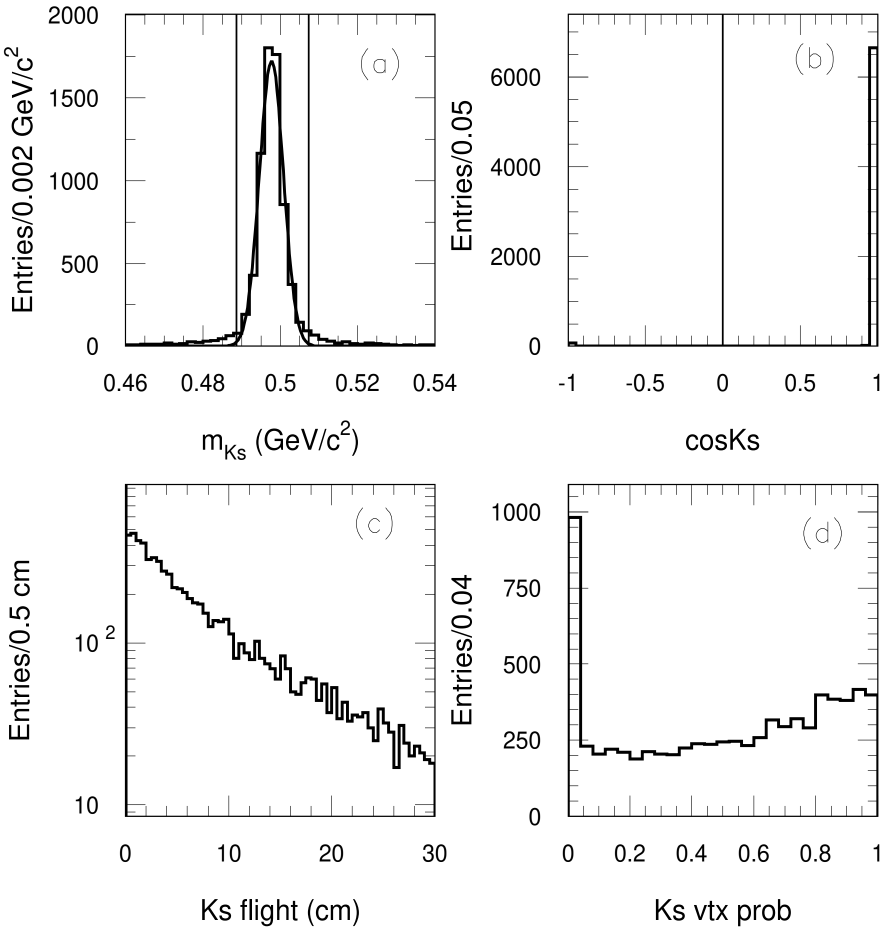


Figure 7.3: Quantities used for the K_S^0 selection, shown here for truth tagged signal $B^0 \to K_S^0 e^+ e^-$ SP3 Monte Carlo. Only skim and code selection criteria have been applied. (a) K_S^0 mass candidate, fitted to a Gaussian with $\mu = 0.498$ GeV/c^2, and $\sigma = 3.1$ MeV/c^2. The lines indicate the values of a 3σ selection on the mass. (b) $\cos K_S$, defined in the text. This variable is required to be positive. (c) K_S^0 flight distance, defined in the text. (d) K_S^0 vertex probability.

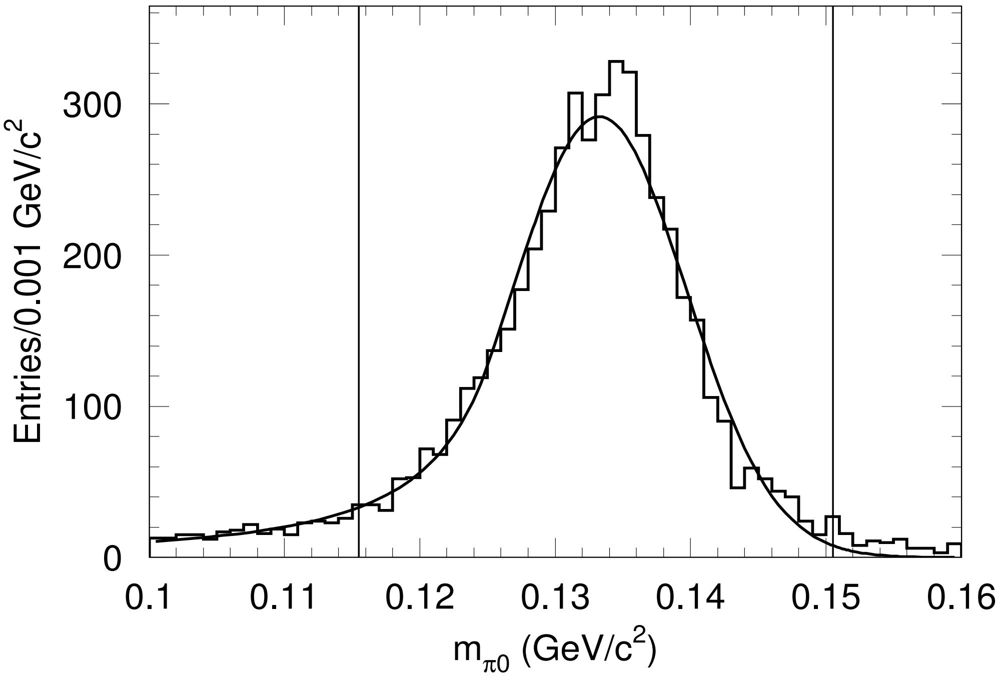


Figure 7.4: The mass of the selected π^0 candidates shown for truth tagged signal $B^+ \to K^{*+}e^+e^-$ ($K^{*+} \to K^+\pi^0$) SP3 Monte Carlo. The distribution is fitted to the Crystal Ball shape [73]. The central value of the mass obtained from the fit is 0.133 GeV/c^2, and the resolution, 6.4 MeV/c^2. The lines indicate the values of the ntuple selection criteria on the π^0 mass.

7.9 Selection of Unique B Candidates per Event

All B candidates with m_{ES} and ΔE satisfying the code event selection criteria described in Section 7.3 are kept in the ntuple. There are no restrictions on the number of B candidates per event.

A unique B candidate is selected at the ntuple level. In the modes involving a slow pion, such as those with $K^{*0} \to K^+\pi^-$ and $K^{*+} \to K^0_S\pi^+$, many multiple combinations are caused by the so-called "ghost track" pions, which are tracks found twice in the SVT using slightly different combinations of hits. Therefore, in these modes the selection of a unique B candidate is based on choosing the combination that has a pion with the largest number of SVT hits. For $B^+ \to K^+\ell^+\ell^-$ and $B^+ \to K^{*+}\ell^+\ell^-$ ($K^{*+} \to K^+\pi^0$), the combination with the maximal number of DCH hits on the kaon track is selected.

If the number of SVT (DCH) hits on the pion (kaon) track is the same for different B candidates, the first combination that occurred is selected. This method is also used for the $B^0 \to K^0_S\ell^+\ell^-$ mode.

7.10 Suppression of e^+ and e^- Originating From γ Conversions

Electrons originating from γ conversions in material are a background for the signal modes with electron-positron pairs. In order to suppress this background, the standard *BABAR* conversions suppression algorithm is used. The algorithm pairs up a loosely identified electron or positron candidate with all other oppositely charged tracks in the event, whose four momenta have been re-scaled assuming an electron mass hypothesis. The resulting dilepton candidate is tested to check whether it satisfies the requirements listed in Table 7.5. The quantities used in the test are the invariant mass $m_{e^+e^-}$ of the dilepton pair and the distance d_{xy} between the two tracks in the $x-y$ plane and in the z direction (d_z). If the criteria of Table 7.5 are satisfied, the electron or positron candidate is assumed to have come from a γ conversion.

Figure 7.5 shows a number of quantities for positrons that were identified by the above described algorithm as originating from γ conversions in the off-resonance data sample. Only skim and code event selection criteria have been applied. The peaks in Fig. 7.5a and the structure in Fig. 7.5c are easily identifiable. The broad peak below $R = 10$ cm corresponds to the beam pipe and the first three layers of the

Variable	Value
$\lvert d_{xy}\rvert$	< 0.5 cm
d_z	< 1.0 cm
$m_{e^+e^-}$	< 30 MeV/c^2

Table 7.5: Parameters used to test whether an e^+e^- pair could come from a γ conversion. The meaning of the variables used is explained in the text. e^+e^- pairs satisfying these criteria are considered to be conversions candidates.

SVT. The two peaks at $10 < R < 15$ cm correspond to the fourth and fifth layers of the SVT. The peak at $R \sim 22$ cm corresponds to the support tube, and the peak at $R \sim 24$ cm, to the beryllium inner cylinder of the DCH. A significant fraction of conversions leptons originate at the area corresponding to the support cones and cooling rings of the SVT, where the mounted electronics and cooling equipment provide a high density of material. This material is not adequately simulated in the *BABAR* GEANT3-based Monte Carlo, as shown in Fig. 7.6.

For the $B \to K^*\ell^+\ell^-$ modes, where there is a pole at $q^2 = 0$, the criteria listed in Table 7.5 are used to reject conversion candidates only if their radial vertex position $R > 2$ cm, since the radius of the beam pipe is ~ 2 cm. In practice, some real conversions can pass this requirement because the resolution on the vertex position of two nearly parallel tracks is not perfect.

It is important to point out that the background from conversions does not peak in the signal region.

7.11 Suppression of the Charmonium Background

In order to suppress $B \to J/\psi(K, K^*)$ and $B \to \psi(2S)(K, K^*)$ events, it is not sufficient to simply remove events with dilepton masses consistent with those of the J/ψ and $\psi(2S)$. The reason is that bremsstrahlung or track mismeasurement can result in a large departure of the dilepton mass from the J/ψ or $\psi(2S)$ masses.

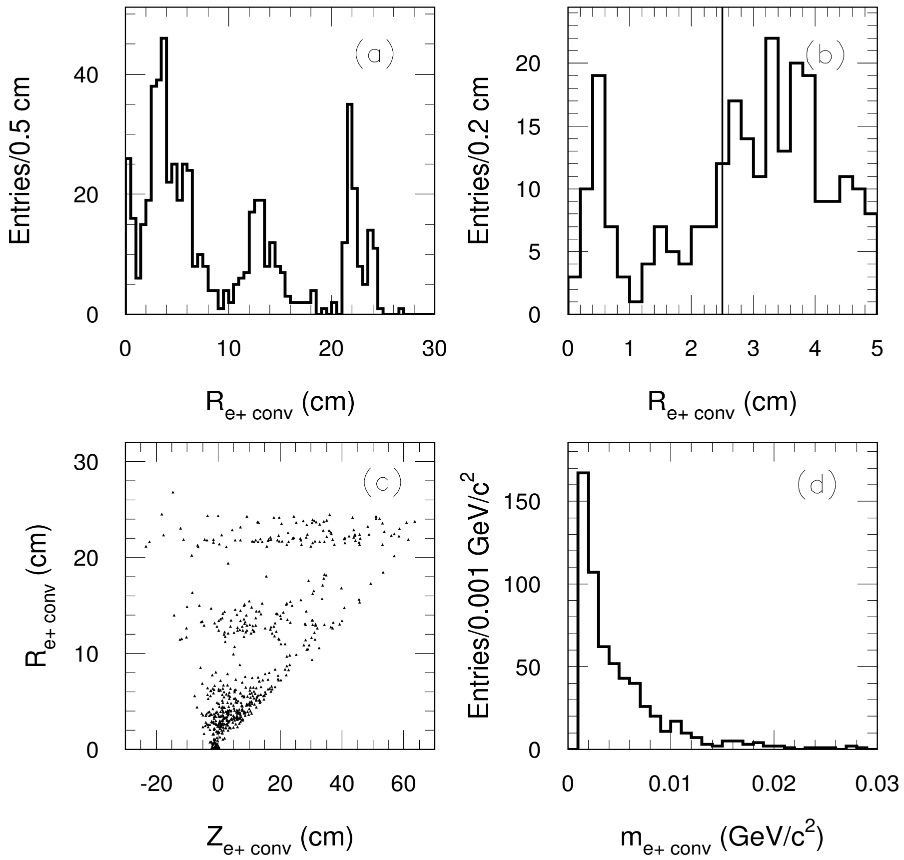


Figure 7.5: Positrons identified by the conversions finder as originating from γ conversions in the off-resonance data sample. Only skim and code cuts have been applied. (a) The $x - y$ positions of the conversion vertices. (b) A close-up on the region from 0 to 5 cm. The line corresponds to the position of the beryllium beam pipe. The algorithm finds a significant number of conversions in front of the beam pipe, which are e^+e^- pairs from π^0 Dalitz decays and γ conversions with poorly resolved vertices. (c) The distribution of the $x - y$ positions of the conversion vertices vs. their z positions. (d) The distribution of the conversion candidate invariant mass.

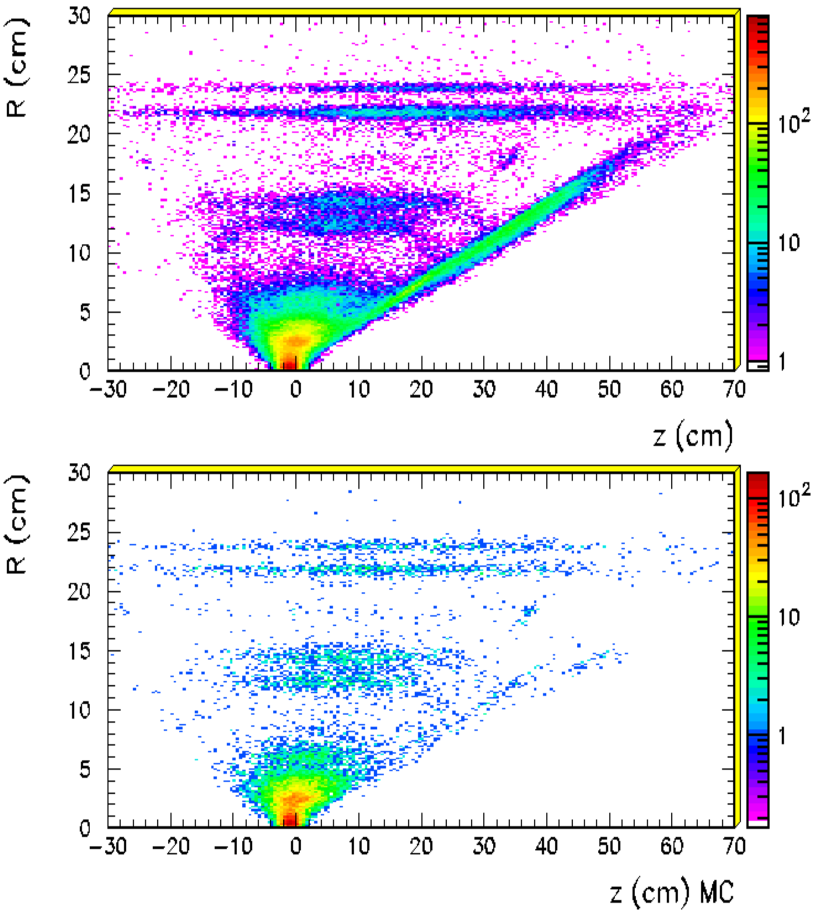


Figure 7.6: Position of the conversion vertices in the $R - z$ plane in data (upper plot), and Monte Carlo (lower plot) in hadronic events. The conversions originating at the support cone of the SVT are not adequately simulated in the GEANT3-based Monte Carlo: there is a clear excess of events in the forward region of the detector in the data plot. Figure courtesy Patrick Robbe.

In principle, removing events with dilepton masses close to $m_{J/\psi}$ and $m_{\psi(2S)}$ may be sufficient to eliminate these events from the *signal* region, as shown in Fig. 7.8a and Fig. 7.8b, where signal charmonium Monte Carlo is shown in the signal ΔE region. However, it is desirable to remove these events not only from the signal region, but

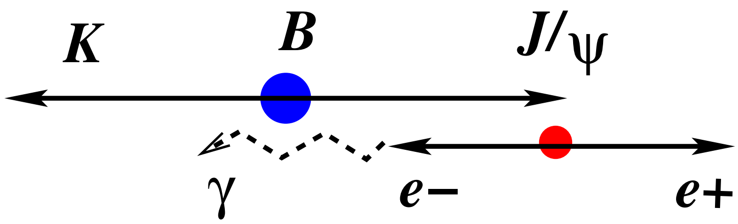


Figure 7.7: The kinematic configuration corresponding to the minimal change in ΔE for a given change in $m_{\ell^+\ell^-}$ in radiative J/ψ events. The photon emitted by the electron from the decay $J/\psi \to e^+e^-$ is moving in the direction opposite to that of the J/ψ in the B rest frame.

also from the background sideband region, since it simplifies the understanding of the backgrounds and makes it easier to perform a two-dimensional fit in the ΔE vs. $m_{\rm ES}$ plane. To remove charmonium background from the sideband region, a correlated selection on ΔE and $m_{\ell^+\ell^-}$ is applied. The correlation comes about as follows. Consider a radiative $J/\psi \to e^+e^-$ decay. In the J/ψ rest frame, $m_{e^+e^-} = E^*_{e^+} + E^*_{e^-}$, where $E^*_{e^+}$ and $E^*_{e^-}$ are the energies of the positron and the electron, respectively. Now, suppose one of the leptons emitted a 200 MeV photon, reducing $m_{e^+e^-}$ from the mean J/ψ mass, 3.1 GeV/c^2, to 2.9 GeV/c^2. This will also result in a shift in ΔE. Let us consider the worst case scenario, where a given shift in $m_{e^+e^-}$ results in the smallest possible change in ΔE. This happens when the photon is emitted in the direction opposite to the direction of the J/ψ in the B rest frame, as depicted in Fig. 7.7. Since in the B rest frame $\beta_{J/\psi} = 0.48$ and $\gamma_{J/\psi} = 1.14$, the change in the energy of the dilepton system in the laboratory frame is:

$$\Delta E_{\rm min} = \gamma_{J/\psi}\ (E^*_\gamma - \beta E^*_\gamma) = 1.14\ (1 - 0.48)\ 200 = 120\ (\text{MeV}). \tag{7.11.1}$$

Therefore, by carefully choosing a correlated $m_{\ell^+\ell^-}$ vs. ΔE cut, charmonium events can be removed from the sideband. Figure 7.8c and 7.8d show the two-dimensional charmonium veto. The regions between the close pairs of vertical lines correspond to the nominal J/ψ and $\psi(2S)$ resonance regions, and the triangular regions veto the events with bremsstrahlung and track mismeasurement. In the electron channels, the effects of bremsstrahlung are much more substantial than in the muon channels, and the veto regions are subsequently much larger. The numerical criteria for the vetoes are summarized in Table 7.6.

Mode	Veto
$J/\psi \to e^+e^-$	if 3.17 $< m_{l^+l^-} <$3.4 GeV/c^2, then $\Delta E >$ (0.68 $m_{l^+l^-}$-1.99) or $\Delta E <$ (1.43 $m_{l^+l^-}$-4.57) if $m_{l^+l^-} <$ 3.02 GeV/c^2, then $\Delta E >$ (0.68 $m_{l^+l^-}$-1.99)
$\psi(2S) \to e^+e^-$	if 3.76 $< m_{l^+l^-} <$ 4.0 GeV/c^2, then $\Delta E >$ (0.75 $m_{l^+l^-}$-2.68) or $\Delta E <$ (1.3 $m_{l^+l^-}$-4.9) if $m_{l^+l^-} <$ 3.61 GeV/c^2, then $\Delta E >$ (0.75 $m_{l^+l^-}$-2.68)
$J/\psi \to \mu^+\mu^-$	$\Delta E >$ 1.11 $m_{l^+l^-}$-3.33 or $\Delta E <$ 1.11 $m_{l^+l^-}$-3.60 and $m_{l^+l^-} <$ 3.05 GeV/c^2 or $m_{l^+l^-} >$ 3.15GeV/c^2
$\psi(2S) \to \mu^+\mu^-$	$\Delta E < m_{l^+l^-}$-3.8 or $\Delta E > m_{l^+l^-}$-3.55 and $m_{l^+l^-} <$ 3.64 GeV/c^2 or $m_{l^+l^-} >$ 3.74 GeV/c^2

Table 7.6: Charmonium veto numerical definitions. Events are required to satisfy these criteria to *pass* the vetoes.

There are several ways for the charmonium events to pass this veto. For example, $B \to J/\psi K$ and $B \to \psi(2S)K$ events can pass the veto if the kaon is misidentified as a lepton (most often a muon). To be more specific, consider the case of a $B^+ \to J/\psi K^+$ decay, where $J/\psi \to \mu^+\mu^-$. If the K^+ and the μ^+ are interchanged ("swapped"), the resulting "dilepton" mass may be such that the event will escape the charmonium veto and peak in ΔE and $m_{\rm ES}$. To suppress this background, the four-momentum of the kaon candidate is re-calculated assuming a muon hypothesis, and the invariant mass m_{swap} of this kaon and the μ^- is calculated. Removing events with m_{swap} close to the J/ψ or $\psi(2S)$ mass is an efficient way to suppress the background from such "swapped" charmonium events.

Another way in which $B \to J/\psi K$ and $B \to \psi(2S)K$ events can pass the charmonium veto is the feed-up to the $K^*\ell^+\ell^-$ modes. If a J/ψ or a $\psi(2S)$ decayed radiatively, then ΔE can be brought back into the signal region – without

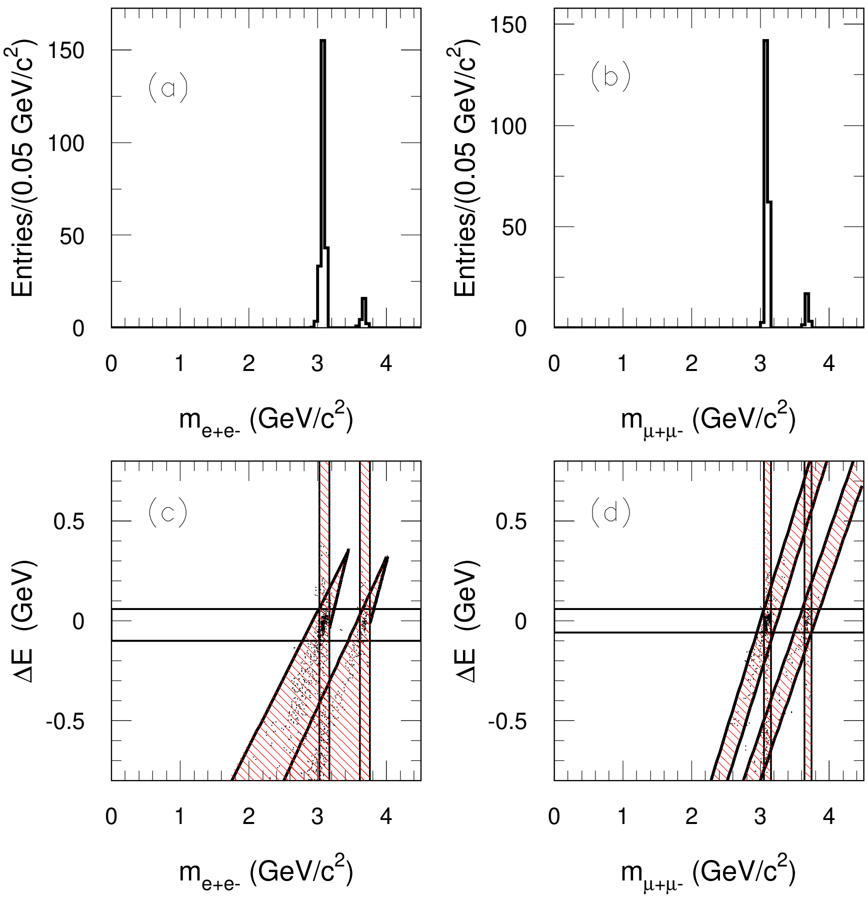


Figure 7.8: (a) and (b) Di-electron and di-muon mass, respectively, in the signal $B \to J/\psi K$ and $B \to \psi(2S)K$ SP3 Monte Carlo after applying all of the $K\ell^+\ell^-$ analysis event selection criteria except those on $m_{\rm ES}$ in the signal ΔE region. (c) and (d) The hatched areas show the regions in the ΔE vs. $m_{\ell^+\ell^-}$ plane that are used to exclude $J/\psi(K, K^*)$ and $\psi(2S)(K, K^*)$ events for electron and muon channels, respectively. The sample used is the signal $B \to J/\psi K$ and $B \to \psi(2S)K$ SP3 Monte Carlo. All of the $K\ell^+\ell^-$ analysis selection criteria have been applied except those on ΔE. The horizontal lines show the signal cut in ΔE.

altering the reduced dilepton mass – by combining the kaon with a pion from the other B. Such an event can escape the charmonium veto and mimic the $B \to K^*\ell^+\ell^-$ signal. Figure 7.9a and 7.9b illustrate the problem. The feed-up background can be suppressed by checking the consistency of the event with the hypothesis of $B \to J/\psi K$ and $B \to \psi(2S)K$, where J/ψ or $\psi(2S) \to \ell^+\ell^-\gamma$. In such an event, the energy of the photon in the center-of-mass frame is:

$$E^*_\gamma = E^*_{\text{beam}} - E^*_{e^+} - E^*_{e^-} - E^*_K, \tag{7.11.2}$$

where E^*_{beam} is the beam energy, and $E^*_{e^+}$, $E^*_{e^-}$, and E^*_K are the positron, electron, and kaon energies, respectively, in the center of mass frame. Supposing that this photon was emitted in the direction of the e^+ or e^-, it is easy to compute the mass of the assumed J/ψ or $\psi(2S)$ candidate. For example, supposing that the photon was aligned with the positron, and setting the masses of the leptons to zero:

$$m_{\gamma e^+e^-} = 2\ \left[E^*_{e^-}\ (E^*_{e^+} + E^*_\gamma)\ (1 - \cos\theta^*)\right], \tag{7.11.3}$$

where θ^* is the angle between the positron and the electron in the CM frame. If this mass is close to the J/ψ or $\psi(2S)$ mass, the next step is to check whether the dilepton-γ pair makes a B with the kaon candidate. This can be done by calculating the B momentum in the CM frame and checking whether it is close to ~ 330 MeV/c:

$$\mathbf{p}^*_{\gamma e^+e^-K} \equiv \mathbf{p}^*_B = \mathbf{p}^*_K + \mathbf{p}^*_{e^-} + \mathbf{p}^*_{e^+}\ \frac{E^*_{e^+} + E^*_\gamma}{E^*_{e^+}} \tag{7.11.4}$$

Table 7.7 summarizes the feed-up veto criteria, and Fig. 7.9b and 7.9c show the effect of applying these vetoes in the $K^{*0}e^+e^-$ and $K^{*0}\mu^+\mu^-$ modes, respectively, using the signal $B \to J/\psi K$ and $B \to \psi(2S)K$ SP3 Monte Carlo. The amount of the feed-up is significantly reduced after applying the veto. Note that the veto is somewhat tighter for the modes with a π^0, where the mass and momentum parameters for the electron channel are 0.1 GeV and 0.2 GeV/c^2, respectively, and for the muon channel, they are 0.07 GeV and 0.15 GeV/c^2, respectively.

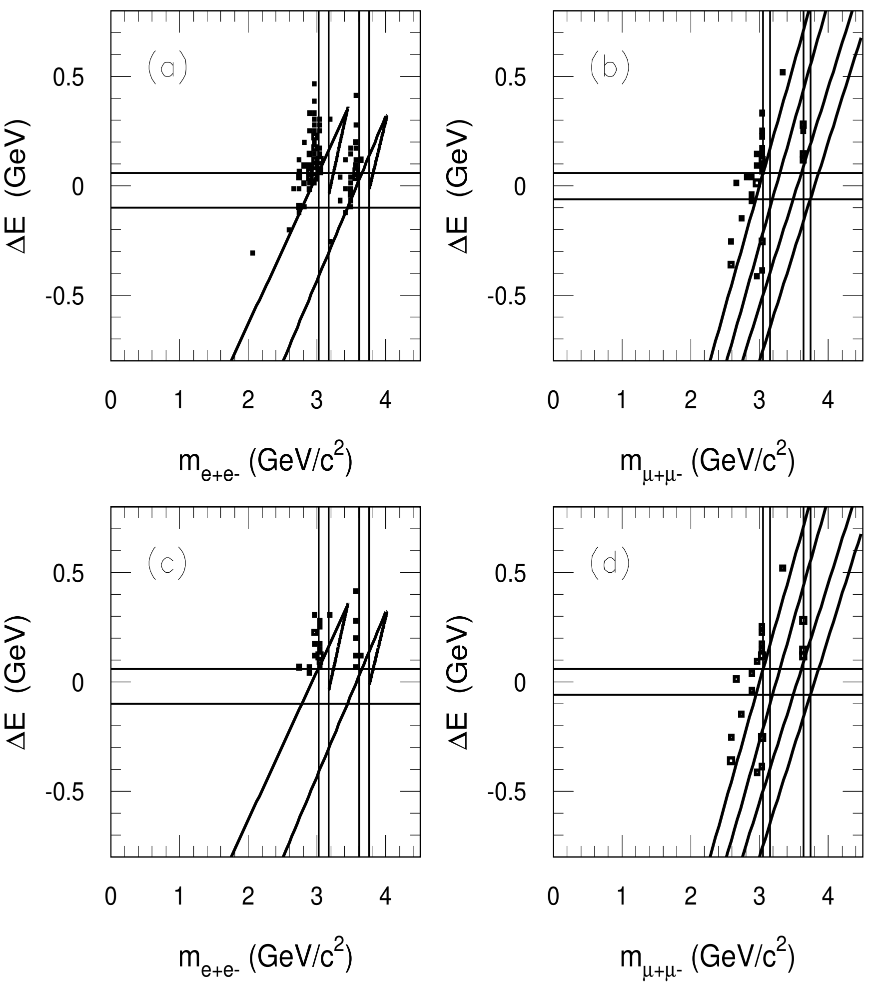


Figure 7.9: (a) and (b) $B \to J/\psi K$ and $B \to \psi(2S)K$ events feeding up to $K^{*0}e^+e^-$ and $K^{*0}\mu^+\mu^-$, respectively. The sample used is signal $B \to J/\psi K$ and $B \to \psi(2S)K$ SP3 Monte Carlo. All event selection criteria have been applied except the signal selection requirements on ΔE. (c) and (d) The same after applying the feed-up veto. The number of events in the signal region has been significantly reduced.

Feeding-up mode	Veto
$J/\psi(\rightarrow e^+e^-)K$	$\lvert m_{\gamma e^+e^-}$ - 3.097 $\rvert > 0.07$ GeV/c^2 and $\lvert p^*_{\gamma e^+e^-K}$ - 0.33 $\rvert > 0.18$ GeV/c
$\psi(2S)(\rightarrow e^+e^-)K$	$\lvert m_{\gamma e^+e^-}$ - 3.69 $\rvert > 0.07$ GeV/c^2 and $\lvert p^*_{\gamma e^+e^-K}$ - 0.33 $\rvert > 0.18$ GeV/c
$J/\psi(\rightarrow \mu^+\mu^-)K$	$\lvert m_{\gamma\mu^+\mu^-}$ - 3.097 $\rvert > 0.05$ GeV/c^2 and $\lvert p^*_{\gamma\mu^+\mu^-K}$ - 0.33 $\rvert > 0.12$ GeV/c
$\psi(2S)(\rightarrow \mu^+\mu^-)K$	$\lvert m_{\gamma\mu^+\mu^-}$ - 3.69 $\rvert > 0.05$ GeV/c^2 and $\lvert p^*_{\gamma\mu^+\mu^-K}$ - 0.33 $\rvert > 0.12$ GeV/c

Table 7.7: Feed-up veto definitions. Events are required to satisfy these criteria to *pass* the vetoes.

It is also possible to veto the feed-up by applying the charmonium veto to the dilepton-kaon combination. However, this method results in a much lower signal efficiency because ignoring the pion lowers ΔE. For example, such a veto would be only 77% efficient for signal in the $B^0 \rightarrow K^{*0}e^+e^-$ channel, whereas the feed-up veto described above is 98.5% efficient after all other event selection criteria have been applied.

7.12 Continuum Suppression

The background from continuum events is combinatorial, arises from combining random leptons and kaons. The leptons can be real, coming from semileptonic D decays or (for the electron channels) from π^0 Dalitz decays and γ conversions in material. They can also be hadrons faking a lepton signature. There are a number of variables that can be used for suppressing this background. Table 7.8 defines four of them: R_2, $|\cos\theta_B|$, $|\cos \mathrm{Thrust}|$, and $m_{K\ell}$. These variables are shown in Fig. 7.10 for $B^+ \rightarrow K^+e^+e^-$, comparing SP3 signal Monte Carlo and off-resonance data. The Fox-Wolfram moment, R_2, tends to have lower values for the more spherical $B\overline{B}$ events than for the jet-like continuum. $\cos\theta_B$ is distributed as $\sin^2\theta$ for signal, and

Variable	Definition
R_2	normalized 2nd order Fox-Wolfram moment
$\cos\theta_{\mathrm{Thrust}}$	cosine of the angle between the signal thrust axis and the z axis in CM frame
$\cos\theta_B$	cosine of the angle between the B candidate and the z axis in CM frame
$m_{K,\ell}$	invariant mass of the kaon and the oppositely charged lepton

Table 7.8: Continuum suppression variable definitions.

is roughly flat for the continuum background (strictly speaking, the quark jets follow the $1+\cos^2\theta$ distribution, but this angular dependence is not reflected in the randomly formed B candidates). cos Thrust is flat for signal, and peaks sharply at $\pm$ 1 for continuum. Finally, $m_{K\ell}$ helps discriminate against the D semileptonic decays, since it will tend to peak below the D mass for real leptons and kaons coming from processes such as $\overline{D}^0 \to K^+\ell^-\overline{\nu}$. The lepton and kaon candidate particles used in the calculation of this variable are different for different modes. For $B^+ \to K^+\ell^+\ell^-$, they are K^+ and ℓ^- (charge-conjugation is always implied). For $B^0 \to K^{*0}\ell^+\ell^-$ ($K^{*0} \to K^+\pi^-$), K^+ and ℓ^- are used as well. For the $B^0 \to K^0_S\ell^+\ell^-$ mode, there is no information about the charge correlations, so it is necessary to consider both $m_{K^0_S\ell^+}$ and $m_{K^0_S\ell^-}$. For $B^+ \to K^{*+}\ell^+\ell^-$ ($K^{*+} \to K^0_S\pi^+$), the charged pion tags the flavor of the assumed D. If the pion is positively charged, K^0_S and ℓ^+ are used to calculate $m_{K\ell}$. For $B^+ \to K^{*+}\ell^+\ell^-$ ($K^{*+} \to K^+\pi^0$), K^+ and ℓ^- are again used to calculate $m_{K\ell}$.

In order to enhance the discriminating power of these variables, they are combined into a Fisher discriminant. The Fisher discriminant is a linear combination of four input variables: $\mathcal{F} = \sum_{i=1}^{4} \alpha_i x_i$, where the coefficients (or the weights) α_j are defined as [59]

$$\alpha_j \equiv \sum_{i=1}^{4} (U^{\mathrm{sig}}_{ij} + U^{\mathrm{bkg}}_{ij})^{-1} \times (\mu^{\mathrm{bkg}}_i - \mu^{\mathrm{sig}}_i), \tag{7.12.1}$$

where U_{ij}^{sig} and U_{ij}^{bkg} are the covariance matrices for the input variables for signal and background, respectively, and μ_i^{sig} and μ_i^{bkg} are the mean values of the x_i's. The coefficients of the Fisher discriminants are determined mode-by-mode using SP3 signal Monte Carlo and off-resonance data as the background sample. The discriminants are shown in Figs. 7.11 and 7.12 for the charged track modes and for the modes with a K_S^0 or a π^0, respectively. The discriminants are often re-scaled so as to be between 0 and 1; however, since they are defined as polynomials, there is no physical reason for them to be constrained to [0,1]. Figures 7.11 and 7.12 show that the Fisher discriminants are not the same for different modes. This is explained by the fact that they reflect not only the properties of the signal, but also those of the background, which are not the same for different channels. For example, lepton fakes constitute a more significant background for the muon modes than for the electron ones.

7.13 $B\overline{B}$ Suppression

The combinatorial background from $B\overline{B}$ events can be suppressed in a way similar to the suppression of the continuum background described in Section 7.12. The variables used for suppressing this background are defined in Table 7.9, and shown in Fig. 7.13 for the $B^+ \to K^+e^+e^-$ mode. Both B and di-lepton vertex probabilities are roughly flat for signal, and peak sharply at 0 for the background. The significance of the di-lepton separation in z is broader for the background than for signal. The missing energy E_{miss} is expected to be larger for the background because of the presence of semileptonic decays with missing neutrinos. $\cos\theta_B$ is used again, this time for discriminating against the $B\overline{B}$ background where it is expected to be flat.

In principle, it is possible to combine these variables into a Fisher discriminant. However, Fisher discriminants are not particularly effective in combining non-Gaussian variables. For example, it is apparent from Fig. 7.13a that only the

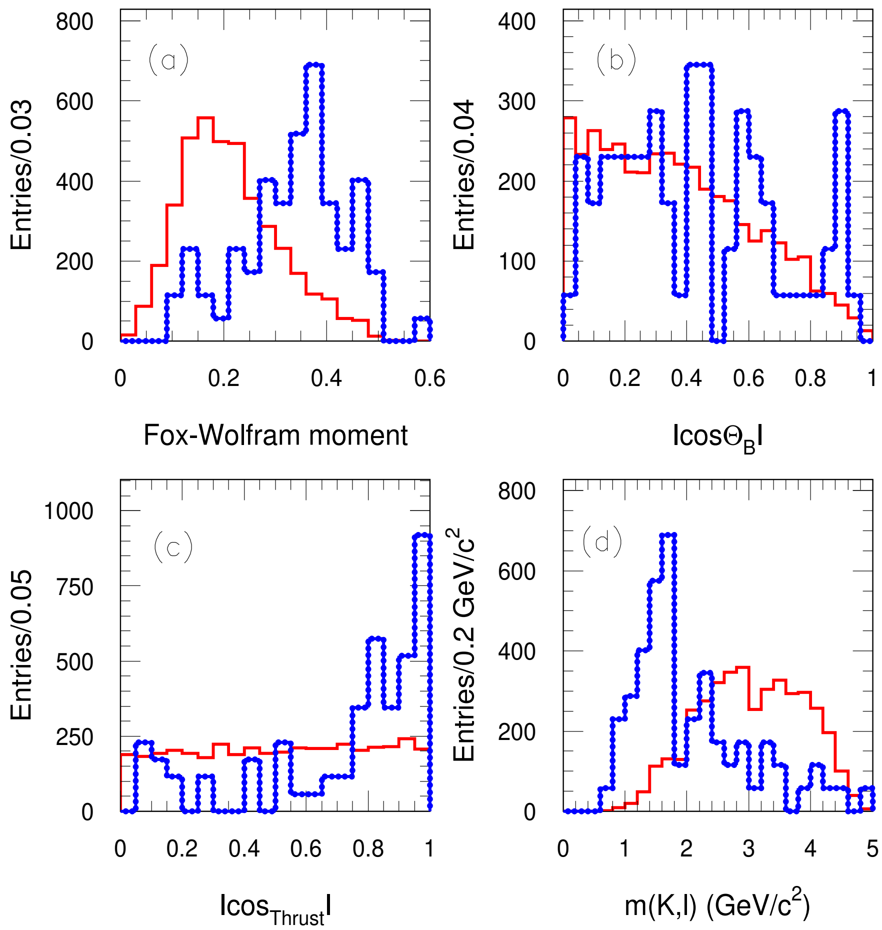


Figure 7.10: The constituent variables used in a Fisher discriminant, shown for the $B^+ \to K^+e^+e^-$ mode after applying all of the event selection criteria except those on the Fisher discriminant itself, ΔE, and m_{ES}. Solid line: truth tagged signal SP3 Monte Carlo, dotted line: off-resonance data. (a) the Fox-Wolfram moment R_2; (b) $\cos\theta_B$; (c) $\cos\theta_{\mathrm{Thrust}}$; (d) $m_{K,l}$: note the peak below the D mass for continuum. The signal and off-resonance plots are normalized to the same total area to compare the shapes.

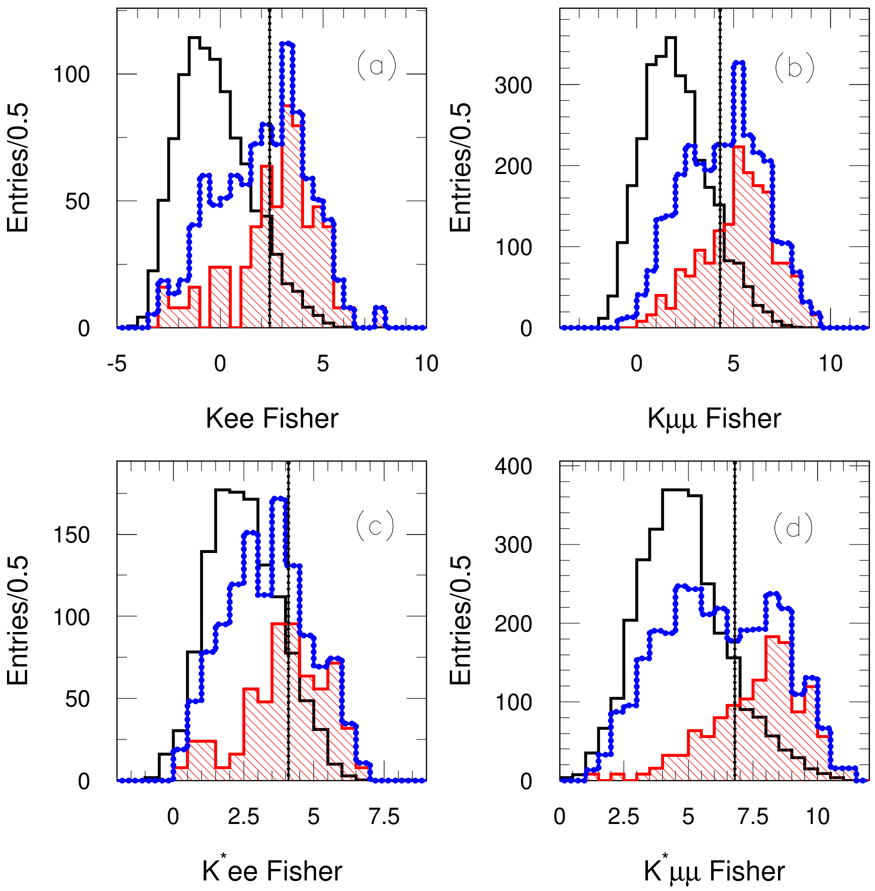


Figure 7.11: Fisher discriminants for the four charged track modes. (a) $B^+ \to K^+e^+e^-$, (b) $B^+ \to K^+\mu^+\mu^-$, (c) $B^0 \to K^{*0}e^+e^-$, and (d) $B^0 \to K^{*0}\mu^+\mu^-$. Solid line: truth tagged signal Monte Carlo; hatched histogram: off-resonance data; dotted line: off-resonance data + $B\overline{B}$ Monte Carlo. The total background and signal histograms are normalized to the same total area. All event selection criteria have been applied except those on the Fisher discriminant itself, ΔE, and m_{ES}. The lines indicate the values of the selection requirements.

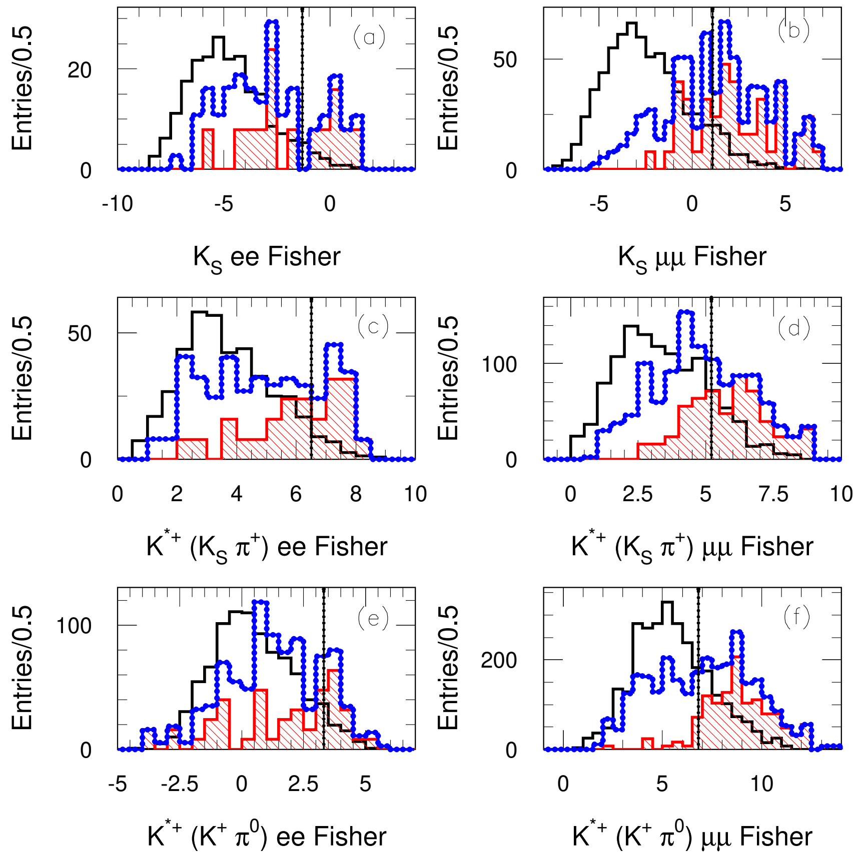


Figure 7.12: Fisher discriminants for the modes with a K^0_S or a π^0. (a) $B^0 \to K^0_S e^+e^-$, (b) $B^0 \to K^0_S \mu^+\mu^-$, (c) $B^+ \to K^{*+}e^+e^-$ with $K^{*+} \to K^0_S\pi^+$, (d) $B^+ \to K^{*+}\mu^+\mu^-$ with $K^{*+} \to K^0_S\pi^+$, (e) $B^+ \to K^{*+}e^+e^-$ with $K^{*+} \to K^+\pi^0$, and (f) $B^+ \to K^{*+}\mu^+\mu^-$ with $K^{*+} \to K^+\pi^0$. Solid line: truth tagged signal Monte Carlo; hatched histogram: off-resonance data; dotted line: off-resonance data + $B\overline{B}$ Monte Carlo. The total background and signal histograms are normalized to the same total area. All event selection criteria have been applied except those on the Fisher discriminant itself, ΔE, and $m_{\rm ES}$. The lines indicate the values of the cuts.

Variable	Definition
$P_{\rm vtx}^{B}$	B vertex probability
$P_{\rm vtx}^{\ell^+\ell^-}$	di-lepton vertex probability
$\Delta z_{\ell^+\ell^-}/\sigma\Delta z_{\ell^+\ell^-}$	significance of the lepton separation in z
$\cos\theta_B$	cosine of the angle between the B candidate and the z axis in CM frame
$E_{\rm miss}$	missing energy of the event calculated in the CM frame.

Table 7.9: $B\overline{B}$ suppression variable definitions.

first few bins of the probability distribution are useful for discriminating between signal and background. An alternative approach is therefore adapted. The variables are combined into a likelihood ratio, $R_{B\overline{B}}$, defined as the ratio of the probability for an event to be signal divided by the probability for an event to be $B\overline{B}$ background:

$$\mathcal{R}_{B\overline{B}} \equiv \frac{P_{\rm sig}}{P_{\rm bkg}}, \tag{7.13.1}$$

where $P_{\rm sig}$ and $P_{\rm bkg}$ are products of the probability distributions of the variables listed in Table 7.9. The individual probability distributions are obtained by parametrizing the constituent variables and fitting them to find the values of the parameters. The B and di-lepton vertex probabilities, the significance of the di-lepton separation in z, and the missing energy are all parametrized using a double Gaussian shape. $\cos\theta_B$ is parametrized as $\sin^2\theta$ for signal, and as a constant for background. The variables combined into $P_{\rm sig}$ and $P_{\rm bkg}$ are assumed to be independent. In fact, some of the variables are somewhat correlated; however, even so the likelihood ratio is superior to simple square selection requirements. Figures 7.14 and 7.15 shows the likelihood ratios for the charged track modes and for the modes with a K_S^0 or a π^0, respectively.

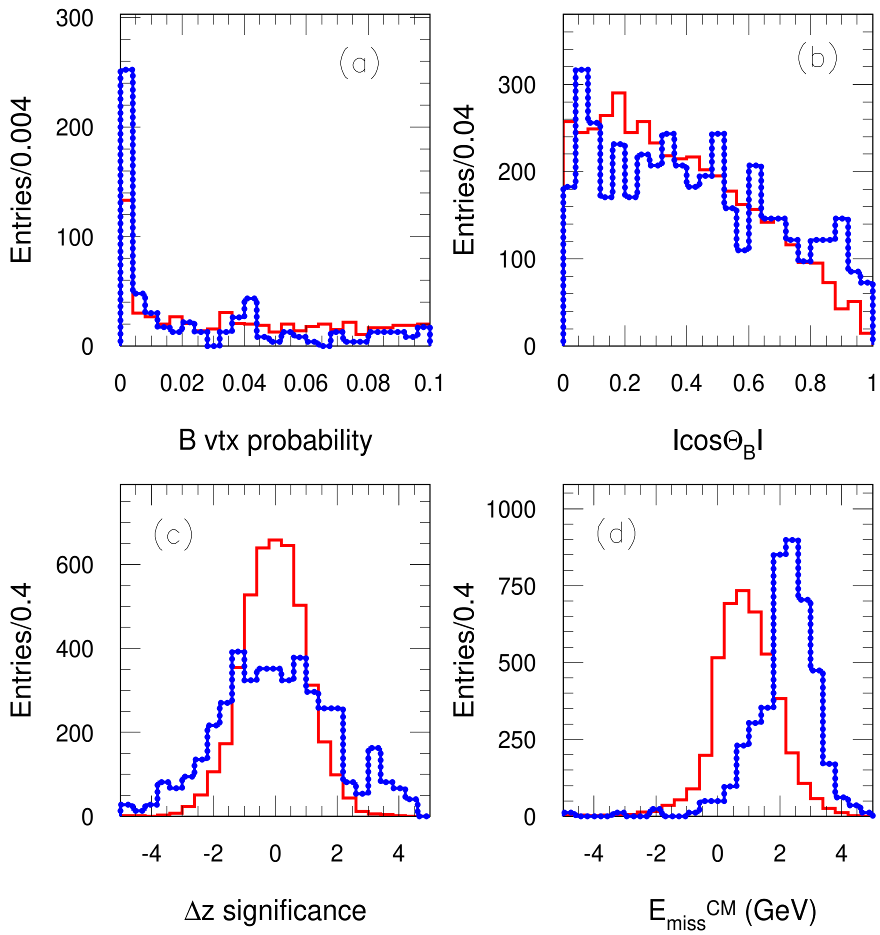


Figure 7.13: The variables used in the $B\overline{B}$ suppression variable, shown for the $B^+ \to K^+e^+e^-$ mode after applying all of the event selection criteria except those on the variable itself, ΔE, and m_{ES}. Solid line: truth tagged signal Monte Carlo, dotted line: generic $B\overline{B}$ SP3 Monte Carlo. (a) the B vertex probability; (b) $|\cos\theta_B|$; (c) the dilepton z separation significance; (d) the missing energy of the event in the CM frame. Signal and background are normalized to the same total area.

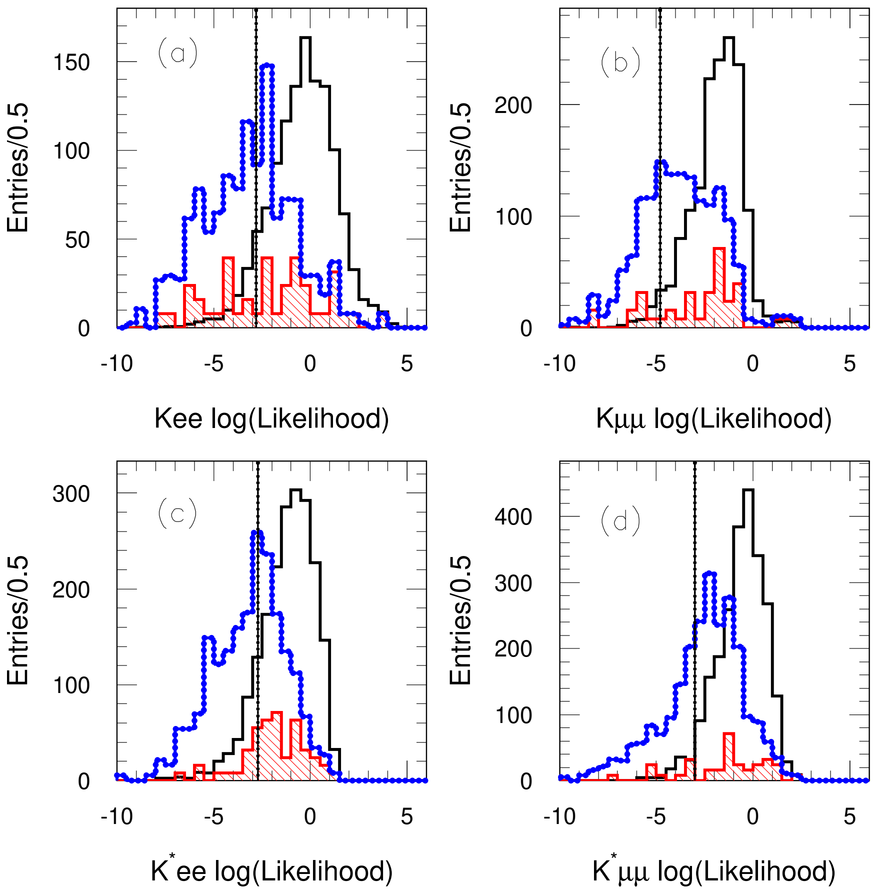


Figure 7.14: $B\overline{B}$ suppression variables for the charged track modes. (a) $B^+ \to K^+e^+e^-$, (b) $B^+ \to K^+\mu^+\mu^-$, (c) $B^0 \to K^{*0}e^+e^-$, and (d) $B^0 \to K^{*0}\mu^+\mu^-$. Solid line: truth tagged signal Monte Carlo; hatched histogram: off-resonance data; dotted line: off-resonance data + $B\overline{B}$ Monte Carlo. The total background and signal histograms are normalized to the same total area. All event selection criteria have been applied except those on the $B\overline{B}$ suppression variables themselves, ΔE, and $m_{\rm ES}$. The lines indicate the values of the cuts.

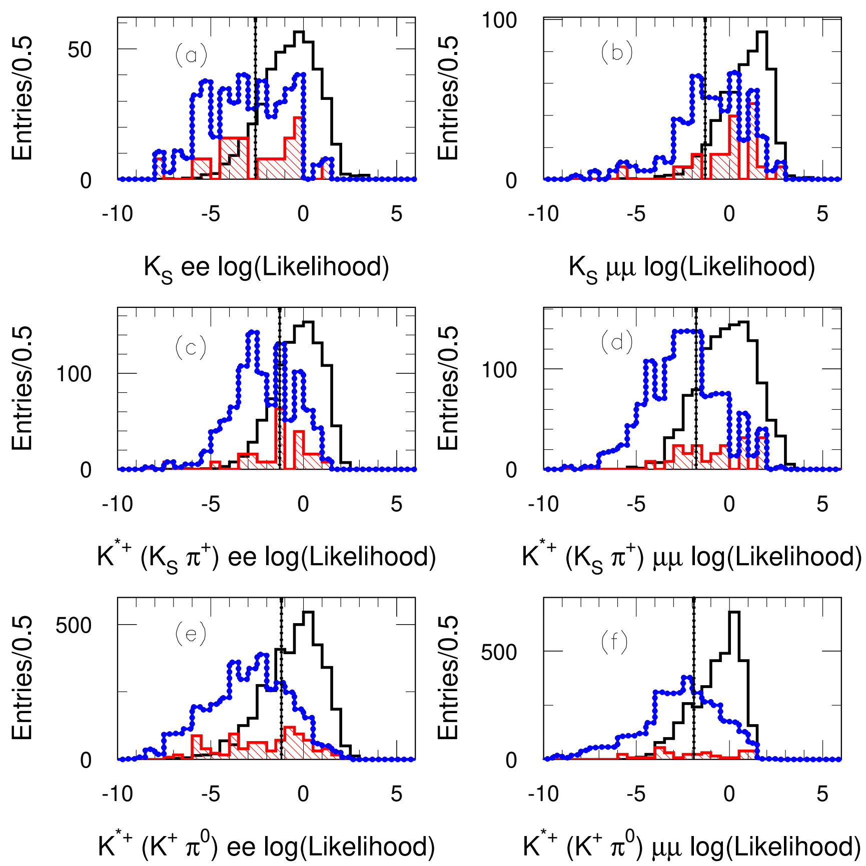


Figure 7.15: $B\overline{B}$ suppression variables for the modes with a K^0_S or a π^0. (a) $B^0 \to K^0_S e^+e^-$, (b) $B^0 \to K^0_S \mu^+\mu^-$, (c) $B^+ \to K^{*+}e^+e^-$ with $K^{*+} \to K^0_S\pi^+$, (d) $B^+ \to K^{*+}\mu^+\mu^-$ with $K^{*+} \to K^0_S\pi^+$, (e) $B^+ \to K^{*+}e^+e^-$ with $K^{*+} \to K^+\pi^0$, and (f) $B^+ \to K^{*+}\mu^+\mu^-$ with $K^{*+} \to K^+\pi^0$. Solid line: truth tagged signal Monte Carlo; hatched histogram: off-resonance data; dotted line: off-resonance data + $B\overline{B}$ Monte Carlo. The total background and signal histograms are normalized to the same total area. All event selection criteria have been applied except those on the $B\overline{B}$ suppression variables themselves, ΔE, and $m_{\rm ES}$. The lines indicate the values of the cuts.

7.14 The Optimization of Event Selection Criteria

Because this is a blind analysis, the optimization of the event selection criteria is performed using Monte Carlo and data samples that will not be used for extracting the number of signal and background events. The off-resonance data set is the ideal sample for studying the effects of a given set of event selection criteria on the continuum background. For the $B\overline{B}$ background, the role of this sample is played by the generic $B\overline{B}$ SP3 Monte Carlo sample. However, the statistics in both the off-resonance data sample and the SP3 $B\overline{B}$ Monte Carlo are too low for this type of analysis. It is therefore necessary to rely on a very large sample of fast parametrized Monte Carlo simulation called Pravda Monte Carlo, described in Appendix A. However, fast Monte Carlo simulation is only used for extracting the *shapes* of various distributions; the off-resonance data and SP3 $B\overline{B}$ Monte Carlo determine the normalization.

The following procedure is used for the optimization of the event selection criteria. First, the number of events in regions $A+B+C$, defined in Fig. 7.1, is counted in the SP3 $B\overline{B}$ Monte Carlo or in the off-resonance data. Let us call this number N_{ABC}^{SP3}. The number of events in $A+B+C$ in Pravda Monte Carlo is also counted. Let us call this number N_{ABC}^{PMC}. Using Pravda Monte Carlo, one determines the fraction of events in the signal region A relative to the number of events in regions $A+B$. This is done by performing a two-dimensional fit described in Chapter 9 in regions $A+B$ and integrating the fit shapes over A. Let us call this fraction f_A^{PMC}. Then, if a simple count of events in regions $A+B$ in Pravda Monte Carlo gives N_{AB}^{PMC}, the number of background events Bkg in the signal region is given by

$$\mathrm{Bkg} = f_A^{\mathrm{PMC}}\; N_{AB}^{\mathrm{PMC}}\; \frac{N_{ABC}^{\mathrm{SP3}}}{N_{ABC}^{\mathrm{PMC}}}. \tag{7.14.1}$$

Essentially, the idea is to scale the number of events from regions $A+B+C$ to the signal region A, from Pravda Monte Carlo to SP3 $B\overline{B}$ Monte Carlo or off-resonance data. The reason for doing so is that the statistics are the largest in

$A + B + C$. One of the most crucial assumptions in the procedure is that the fast Monte Carlo simulation models the shape of the ΔE and m_{ES} distributions used in the fitter reasonably well. The validation of this assumption is described in Chapters 8 and 9.

To determine the number of signal events Sig for a given set of event selection criteria, one can simply count the number of events in region A using signal SP3 Monte Carlo. Having obtained both the number of signal events Sig and the number of background events Bkg, two variables are computed for the optimization of the cuts. One is $\mathrm{Sig}^2/(\mathrm{Sig} + \mathrm{Bkg})$, which is maximal for the most optimal criteria. The other is the probability that the background will fluctuate to the number of observed events, described in Appendix B, which is minimal for the best criteria. The two variables typically give the same values for the most optimal cuts; in cases where they do not, the preference is given to the more optimal cut from the point of view of the probability variable.

Only two event selection criteria in this analysis need to be optimized following the procedure described above: the cut involving the Fisher discriminant and the cut involving the $B\overline{B}$ suppression variable. They are optimized concurrently. Most other event selection requirements are dictated by the properties of the particular background they are designed to suppress. For example, the charmonium veto has been chosen so as to allow the smallest possible number of signal charmonium SP3 events to pass it. Similarly, very tight lepton and kaon identification criteria were chosen in order to maximally suppress the backgrounds from fakes.

7.15 Ntuple Event Selection Criteria

This section lists all of the ntuple event selection criteria with their efficiencies for both signal and background. Note that in this section, the tables listing the cut efficiencies for background have been produced without applying the single B candidate per event requirement.

- Table 7.10 summarizes the event selection criteria for the $B^+ \to K^+e^+e^-$ mode, and Table 7.11 lists the efficiencies for these event selection criteria for signal SP3 and Pravda Monte Carlo $B^+ \to K^+e^+e^-$, as well as for $B\overline{B}$ and continuum background. All of these criteria have been discussed in the previous chapters. The conversion veto for the electron channels was described in detail in Section 7.10. The J/ψ and $\psi(2S)$ direct and swap vetoes were discussed in Section 7.11. The origin of the D^0 veto was discussed in Section 7.2: the background from the $B^+ \to \overline{D}^0\pi^+$ ($\overline{D}^0 \to K^+\pi^-$) mode, where the hadrons are mis-identified as leptons, is vetoed by re-assigning the hadron mass hypotheses to the corresponding particles and excluding the resulting mass combinations around the D^0 mass. The Fisher discriminant and the $B\overline{B}$ suppression variables were described in Sections 7.12 and 7.13, respectively. Table 7.11 indicates that the least efficient event selection criteria for this mode are the kaon identification requirements and the ΔE cut. The requirement that the kaon be within the DIRC acceptance affects the former, and bremsstrahlung lowers the efficiency of the latter. The final signal efficiency is 14.4% for SP3 and 15.1% for Pravda Monte Carlo.

- Tables 7.12 and 7.13 list the event selection criteria and their efficiencies for the $B^+ \to K^+\mu^+\mu^-$ mode. For this mode, the two most inefficient cuts are the kaon and muon identification requirements. Muon identification efficiency is lower in *BABAR* than in CLEO or Belle, for reasons described in [74]. Besides, the requirement that the muons have momenta greater than 1.0 GeV/c in the laboratory frame lowers the efficiency of the skim and code event selection requirements. The final signal efficiency is 9.6% for SP3 and 11.3% for Pravda Monte Carlo.

- Tables 7.14 and 7.15 summarize the $B^0 \to K^{*0}e^+e^-$ event selection definitions and efficiencies, respectively. In the definition of the conversion veto, R is the $x-y$ position of the conversion candidate vertex. The requirement that R be greater than 2 cm for the K^* modes is needed to keep the peak at $q^2 = 0$, and is so chosen because the radius of the beam pipe is ~ 2.5 cm. The suppression of the $J/\psi K$ and $\psi(2S)K$ feed-up background was described in Section 7.11. Table 7.15 shows that, similarly to $B^+ \to K^+e^+e^-$, the most inefficient event selection criteria are kaon ID and ΔE. However, the requirements on the Fisher discriminant and especially on the $B\overline{B}$ suppression variable are less efficient for this mode than for $B^+ \to K^+e^+e^-$. This is because the $K^*\ell^+\ell^-$ modes, unlike the $K\ell^+\ell^-$ ones, are dominated by the $B\overline{B}$ background, since the presence of a K^* increases the combinatorial background contribution. The final signal efficiency is 7.8% for SP3 and 9.0% for Pravda Monte Carlo.

- Tables 7.16 and 7.17 describe the event selection criteria and their efficiencies for the $B^0 \to K^{*0}\mu^+\mu^-$ mode. Similarly to $B^+ \to K^+\mu^+\mu^-$, the most inefficient criteria are those involving muon and kaon identification. The final signal efficiency is 6.4% for SP3, and 7.5% for Pravda Monte Carlo.

- The event selection criteria and their efficiencies for $B^0 \to K^0_S e^+e^-$ are summarized in Tables 7.18 and 7.19. The selection of K^0_S mesons was discussed in Section 7.6. The final signal efficiency is 13.1% for SP3 and 19.2% for Pravda Monte Carlo.

- Tables 7.20 and 7.21 describe the event selection criteria and their efficiencies for $B^0 \to K^0_S\mu^+\mu^-$. Muon identification requirements are primarily responsible for lowering the efficiency. The final signal efficiency is 9.1% for SP3, and 14.4% for Pravda Monte Carlo.

- The event selection criteria and efficiencies for the $B^+ \to K^{*+}e^+e^-$ ($K^{*+} \to K^0_S\pi^+$) mode are summarized in Tables 7.22 and 7.23. The most inefficient selection criteria are the $B\overline{B}$ suppression cuts and the ΔE cut. The final signal efficiency is 6.3% for SP3 and 9.6% for Pravda Monte Carlo.

- Tables 7.24 and 7.25 summarize the event selection criteria and their efficiencies for $B^+ \to K^{*+}\mu^+\mu^-$ ($K^{*+} \to K^0_S\pi^+$). Muon identification requirements are the most inefficient cut for this mode. The final signal efficiency is 4.2% for SP3, and 7.3% for Pravda Monte Carlo.

- Tables 7.26 and 7.27 summarize the event selection criteria and their efficiencies for $B^+ \to K^{*+}e^+e^-$ ($K^{*+} \to K^+\pi^0$). The feature of lower $B\overline{B}$ suppression and ΔE cut efficiencies, observable in other $K^*e^+e^-$ modes, is also present here. The final signal efficiency is 3.6% for SP3 and 4.8% for Pravda Monte Carlo.

- The event selection criteria definitions and efficiencies for $B^+ \to K^{*+}\mu^+\mu^-$ ($K^{*+} \to K^+\pi^0$) are given in Tables 7.28 and 7.29. The muon identification cut is only $\sim$ 62% efficient for signal. The final signal efficiency is 2.1% for SP3 and 3.1% for Pravda Monte Carlo.

Event selection	Value
Tracking	e^+, e^- GoodTracksLoose, kaon GoodTrackVeryLoose
Conversions	e^+ and e^- must *fail* criteria of Table 7.5 of Section 7.10
$(J/\psi, \psi(2S))K$ veto	see Table 7.6 of Section 7.11
$(J/\psi, \psi(2S))K$ swap	$m_{swap} > 3.13$ GeV/c^2 or $m_{swap} < 3.06$ GeV/c^2 and $m_{swap} < 3.68$ GeV/c^2 or $m_{swap} > 3.70$ GeV/c^2
D^0 swap veto	$m_{D^0} < 1.84$ GeV/c^2 or $m_{D^0} > 1.9$ GeV/c^2
Fisher $\mathcal{F}$	$\mathcal{F} < 2.4$
Lepton ID	very tight electrons
Kaon ID	very tight kaon
$B\overline{B}$ suppression $\mathcal{R}_{B\overline{B}}$	$\log(\mathcal{R}_{B\overline{B}}) >$ -2.8
ΔE	-0.11 $< \Delta E <$ 0.05 GeV
$m_{\rm ES}$ mass	5.272 $< m_{\rm ES} <$ 5.286 GeV/c^2

Table 7.10: Ntuple cut definitions for the $B^+ \to K^+e^+e^-$ mode.

Selection	$e_{\rm sig}$	$e_{\rm fast}$	$e_{B\overline{B}}$	$e_{\rm offres}$	$\frac{e^2_{\rm sig}}{e_{B\overline{B}}}$	$\frac{e^2_{\rm sig}}{e_{\rm offres}}$	$\frac{N_{B\overline{B}}}{N_{\rm offres}}$
Skim & code selection	0.539	0.588	0.002	0.002	131.222	120.588	0.296
Conversions	0.990	0.992	0.987	0.755	0.992	1.296	0.386
Tracking	0.963	1.000	0.736	0.846	1.260	1.096	0.336
$J/\psi K$ veto	0.846	0.840	0.558	0.821	1.283	0.873	0.229
$\psi(2S)K$ veto	0.873	0.865	0.554	0.825	1.376	0.924	0.153
$(J/\psi, \psi(2S))K$ swap	0.961	0.961	0.951	0.980	0.971	0.943	0.149
D^0 swap veto	0.985	0.987	0.976	0.981	0.995	0.990	0.148
$\mathcal{F}$	0.914	0.873	0.764	0.448	1.093	1.864	0.253
$\mathcal{R}_{B\overline{B}}$	0.919	0.921	0.549	0.755	1.538	1.119	0.184
Electron ID	0.849	0.839	0.287	0.095	2.508	7.613	0.558
Kaon ID	0.750	0.755	0.268	0.090	2.097	6.228	1.658
ΔE	0.763	0.755	0.119	0.115	4.892	5.047	1.710
$m_{\rm ES}$	0.979	0.978	0.000	0.000	-	-	-
Total	0.144	0.151	0.000	0.000	-	-	-

Table 7.11: Ntuple cut efficiencies for the $B^+ \to K^+e^+e^-$ mode. Event selection criteria are applied one after another, with the efficiency defined as the fraction of events passing the previous set of event selection criteria plus a new requirement. $e_{\rm sig}$ and $e_{\rm fast}$ are the requirement efficiency for the truth tagged SP3 and Pravda signal Monte Carlo, respectively; $e_{B\overline{B}}$, for SP3 generic $B\overline{B}$; and $e_{\rm offres}$, for the off-resonance data sample.

Event selection	Value
Tracking	muons GoodTracksLoose, kaon GoodTrackVeryLoose
$(J/\psi, \psi(2S))K$ veto	see Table 7.6 of Section 7.11
$(J/\psi, \psi(2S))K$ swap	$m_{swap} > 3.17$ GeV/c^2 or $m_{swap} < 3.03$ GeV/c^2 and $m_{swap} > 3.70$ GeV/c^2 or $m_{swap} < 3.68$ GeV/c^2
D^0 swap veto	$m_{D^0} < 1.84$ GeV/c^2 or $m_{D^0} > 1.9$ GeV/c^2
Fisher $\mathcal{F}$	$\mathcal{F} < 4.3$
Lepton ID	very tight muons
Kaon ID	very tight kaon
$B\overline{B}$ suppression $\mathcal{R}_{B\overline{B}}$	$\log(\mathcal{R}_{B\overline{B}}) >$ -4.8
ΔE	-0.07 $< \Delta E <$ 0.05 GeV
$m_{\rm ES}$ mass	$5.272 < m_{\rm ES} < 5.286$ GeV/c^2

Table 7.12: Ntuple cut definitions for the $B^+ \to K^+\mu^+\mu^-$ mode.

Selection	$e_{\rm sig}$	$e_{\rm fast}$	$e_{B\overline{B}}$	$e_{\rm offres}$	$\frac{e^2_{\rm sig}}{e_{B\overline{B}}}$	$\frac{e^2_{\rm sig}}{e_{\rm offres}}$	$\frac{N_{B\overline{B}}}{N_{\rm offres}}$
Skim & code selection	0.392	0.441	0.002	0.002	76.828	80.940	0.339
Tracking	0.967	1.000	0.855	0.776	1.092	1.204	0.374
$J/\psi K$ veto	0.906	0.911	0.754	0.939	1.089	0.874	0.300
$\psi(2S)K$ veto	0.876	0.871	0.776	0.885	0.989	0.867	0.263
$(J/\psi, \psi(2S))K$ swap	0.925	0.930	0.930	0.970	0.920	0.882	0.252
D^0 swap veto	0.983	0.985	0.965	0.974	1.002	0.993	0.250
$\mathcal{F}$	0.881	0.831	0.668	0.243	1.161	3.193	0.687
$\mathcal{R}_{B\overline{B}}$	0.968	0.977	0.736	0.921	1.275	1.019	0.549
Muon ID	0.594	0.603	0.333	0.121	1.061	2.912	1.507
Kaon ID	0.762	0.758	0.232	0.114	2.507	5.103	3.068
ΔE	0.916	0.956	0.052	0.062	16.071	13.439	2.566
$m_{\rm ES}$	0.992	0.991	0.133	0.000	7.388	-	-
Total	0.096	0.113	0.000	0.000	-	-	-

Table 7.13: Ntuple cut efficiencies for the $B^+ \to K^+\mu^+\mu^-$ mode. Event selection criteria are applied one after another, with the efficiency defined as the fraction of events passing the previous set of event selection criteria plus a new requirement. $e_{\rm sig}$ and $e_{\rm fast}$ are the requirement efficiency for the truth tagged SP3 and Pravda signal Monte Carlo, respectively; $e_{B\overline{B}}$, for SP3 generic $B\overline{B}$; and $e_{\rm offres}$, for the off-resonance data sample.

Event selection	Value
Tracking	e^+, e^- GoodTracksLoose K, π from K^{*0} GoodTracksVeryLoose
Conversions	if $R > 2$ cm, e^+, e^- must fail criteria of Table 7.5 of Section 7.10
$(J/\psi, \psi(2S))K^{*0}$ veto	see Table 7.6 of Section 7.11
$(J/\psi, \psi(2S))K^+$ feed-up	see Table 7.7 of Section 7.11
K^{*0} mass	$\|m_{K^{*0}} - 0.892\| < 0.075$ GeV/c^2
Fisher $\mathcal{F}$	$\mathcal{F} < 4.1$
Lepton ID	very tight electrons
Kaon ID	very tight kaon
Pion ID	NotAKaon (veto)
$B\overline{B}$ suppression $\mathcal{R}_{B\overline{B}}$	$\log(\mathcal{R}_{B\overline{B}}) >$ -2.7
ΔE	-0.11 $< \Delta E <$ 0.05 GeV
m_{ES} mass	$5.272 < m_{\mathrm{ES}} < 5.286$ GeV/c^2

Table 7.14: Ntuple cut definitions for the $B^0 \to K^{*0}e^+e^-$ mode.

Selection	e_{sig}	e_{fast}	$e_{B\overline{B}}$	e_{offres}	$\frac{e^2_{\mathrm{sig}}}{e_{B\overline{B}}}$	$\frac{e^2_{\mathrm{sig}}}{e_{\mathrm{offres}}}$	$\frac{N_{B\overline{B}}}{N_{\mathrm{offres}}}$
Skim & code selection	0.423	0.489	0.018	0.010	10.100	18.346	0.584
Conversions	0.980	0.979	0.987	0.828	0.972	1.159	0.697
Tracking	0.927	1.000	0.628	0.681	1.367	1.261	0.643
$J/\psi K^*$ veto	0.858	0.852	0.483	0.758	1.526	0.972	0.409
$\psi(2S)K^*$ veto	0.864	0.852	0.568	0.804	1.314	0.928	0.289
$J/\psi K$ feed-up	0.978	0.979	0.996	0.999	0.959	0.957	0.288
K^* mass	0.857	0.870	0.374	0.391	1.962	1.879	0.276
$\mathcal{F}$	0.838	0.775	0.769	0.345	0.912	2.034	0.616
$\mathcal{R}_{B\overline{B}}$	0.862	0.867	0.489	0.782	1.520	0.950	0.385
Electron ID	0.832	0.849	0.311	0.082	2.225	8.418	1.457
Kaon ID	0.734	0.724	0.234	0.172	2.306	3.127	1.976
Pion ID	0.986	0.982	0.854	0.911	1.140	1.068	1.852
ΔE	0.773	0.766	0.117	0.098	5.096	6.117	2.224
m_{ES}	0.972	0.978	0.038	0.000	24.550	-	-
Total	0.078	0.090	0.000	0.000	-	-	-

Table 7.15: Ntuple cut efficiencies for the $B^0 \to K^{*0}e^+e^-$ mode. Event selection criteria are applied one after another, with the efficiency defined as the fraction of events passing the previous set of event selection criteria plus a new requirement. e_{sig} and e_{fast} are the requirement efficiency for the truth tagged SP3 and Pravda signal Monte Carlo, respectively; $e_{B\overline{B}}$, for SP3 generic $B\overline{B}$; and e_{offres}, for the off-resonance data sample.

Event selection	Value
Tracking	μ^+, μ^- GoodTracksLoose K, π from K^{*0} GoodTracksVeryLoose
$(J/\psi, \psi(2S))K^{*0}$ veto	see Table 7.6 of Section 7.11
$(J/\psi, \psi(2S))K^+$ feed-up	see Table 7.7 of Section 7.11
K^{*0} mass	$\lvert m_{K^{*0}} - 0.892\rvert < 0.075$ GeV/c^2
Fisher $\mathcal{F}$	$\mathcal{F} < 6.8$
Lepton ID	very tight muons
Kaon ID	very tight kaon
Pion ID	NotAKaon (veto)
$B\overline{B}$ suppression $\mathcal{R}_{B\overline{B}}$	$\log(\mathcal{R}_{B\overline{B}}) >$ -3.0
ΔE	-0.07 $< \Delta E <$ 0.05 GeV
m_{ES} mass	5.272 $< m_{\mathrm{ES}} <$ 5.286 GeV/c^2

Table 7.16: Ntuple cut definitions for the $B^0 \to K^{*0}\mu^+\mu^-$ mode.

Selection	e_{sig}	e_{fast}	$e_{B\overline{B}}$	e_{offres}	$\frac{e^2_{\mathrm{sig}}}{e_{B\overline{B}}}$	$\frac{e^2_{\mathrm{sig}}}{e_{\mathrm{offres}}}$	$\frac{N_{B\overline{B}}}{N_{\mathrm{offres}}}$
Skim & code selection	0.331	0.381	0.015	0.007	7.477	14.664	0.631
Tracking	0.931	1.000	0.720	0.619	1.203	1.401	0.735
$J/\psi K^*$ veto	0.902	0.898	0.670	0.901	1.213	0.903	0.547
$\psi(2S)K^*$ veto	0.848	0.842	0.775	0.820	0.928	0.876	0.517
$J/\psi K$ feed-up	0.974	0.979	0.999	0.999	0.950	0.950	0.517
K^* mass	0.852	0.872	0.396	0.398	1.830	1.822	0.545
$\mathcal{F}$	0.842	0.802	0.768	0.284	0.924	2.502	1.392
$\mathcal{R}_{B\overline{B}}$	0.951	0.975	0.669	0.849	1.352	1.065	1.097
Muon ID	0.620	0.587	0.355	0.131	1.085	2.945	2.976
Kaon ID	0.733	0.715	0.171	0.094	3.147	5.727	5.416
Pion ID	0.987	0.986	0.867	0.806	1.125	1.210	5.827
ΔE	0.916	0.961	0.049	0.069	17.266	12.163	4.105
m_{ES}	0.987	0.984	0.000	0.000	-	-	-
Total	0.064	0.075	0.000	0.000	-	-	-

Table 7.17: Ntuple cut efficiencies for the $B^0 \to K^{*0}\mu^+\mu^-$ mode. Event selection criteria are applied one after another, with the efficiency defined as the fraction of events passing the previous set of event selection criteria plus a new requirement. e_{sig} and e_{fast} are the requirement efficiency for the truth tagged SP3 and Pravda signal Monte Carlo, respectively; $e_{B\overline{B}}$, for SP3 generic $B\overline{B}$; and e_{offres}, for the off-resonance data sample.

Event selection	Value
Tracking	e^+, e^- GoodTracksLoose
Conversions	e^+, e^- must fail criteria of Table 7.5 of Section 7.10
$(J/\psi, \psi(2S))K^0_S$ veto	see Table 7.6 of Section 7.11
Fisher $\mathcal{F}$	$\mathcal{F}$ < -1.3
K^0_S selection	See Table 7.4 of Section 7.6
Lepton ID	very tight electrons
$B\overline{B}$ suppression $\mathcal{R}_{B\overline{B}}$	$\log(\mathcal{R}_{B\overline{B}})$ > -2.6
ΔE	-0.11 < ΔE < 0.05 GeV
m_{ES} mass	5.272 < m_{ES} < 5.286 GeV/c^2

Table 7.18: Ntuple cut definitions for the $B^0 \rightarrow K^0_S e^+ e^-$ mode.

Selection	e_{sig}	e_{fast}	$e_{B\overline{B}}$	e_{offres}	$\frac{e^2_{\mathrm{sig}}}{e_{B\overline{B}}}$	$\frac{e^2_{\mathrm{sig}}}{e_{\mathrm{offres}}}$	$\frac{N_{B\overline{B}}}{N_{\mathrm{offres}}}$
Skim & code selection	0.395	0.540	0.001	0.001	160.302	209.134	0.420
Conversions	0.989	0.992	0.984	0.814	0.994	1.202	0.508
Tracking	0.965	1.000	0.773	0.922	1.206	1.011	0.426
$J/\psi K$ veto	0.846	0.847	0.535	0.789	1.337	0.907	0.289
$\psi(2S)K$ veto	0.884	0.870	0.596	0.818	1.312	0.955	0.210
K^0_S selection	0.867	0.958	0.132	0.152	5.671	4.932	0.183
$\mathcal{F}$	0.948	0.931	0.871	0.694	1.031	1.293	0.229
$\mathcal{R}_{B\overline{B}}$	0.878	0.863	0.439	0.743	1.755	1.037	0.135
Electron ID	0.840	0.848	0.396	0.035	1.782	19.993	1.520
ΔE	0.786	0.761	0.075	0.000	8.238	-	-
m_{ES}	0.977	0.977	0.000	-	-	-	-
Total	0.131	0.192	0.000	0.000	-	-	-

Table 7.19: Ntuple cut efficiencies for the $B^0 \rightarrow K^0_S e^+ e^-$ mode. Event selection criteria are applied one after another, with the efficiency defined as the fraction of events passing the previous set of event selection criteria plus a new requirement. e_{sig} and e_{fast} are the requirement efficiency for the truth tagged SP3 and Pravda signal Monte Carlo, respectively; $e_{B\overline{B}}$, for SP3 generic $B\overline{B}$; and e_{offres}, for the off-resonance data sample.

Event selection	Value
Tracking	μ^+, μ^- = GoodTracksLoose
$(J/\psi, \psi(2S))K^0_S$ veto	see Table 7.6 of Section 7.11
Fisher $\mathcal{F}$	$\mathcal{F} < 1.1$
K^0_S selection	See Table 7.4 of Section 7.6
Lepton ID	very tight muons
$B\overline{B}$ suppression $\mathcal{R}_{B\overline{B}}$)	$\log(\mathcal{R}_{B\overline{B}}) >$ -1.3
ΔE	-0.07 $< \Delta E <$ 0.05 GeV
$m_{\rm ES}$ mass	5.272 $< m_{\rm ES} <$ 5.286 GeV/c^2

Table 7.20: Ntuple cut definitions for the $B^0 \to K^0_S\mu^+\mu^-$ mode.

Selection	$e_{\rm sig}$	$e_{\rm fast}$	$e_{B\overline{B}}$	$e_{\rm offres}$	$\frac{e^2_{\rm sig}}{e_{B\overline{B}}}$	$\frac{e^2_{\rm sig}}{e_{\rm offres}}$	$\frac{N_{B\overline{B}}}{N_{\rm offres}}$
Skim & code selection	0.288	0.398	0.001	0.001	100.003	136.081	0.438
Tracking	0.973	1.000	0.911	0.898	1.040	1.056	0.444
$J/\psi K$ veto	0.895	0.904	0.733	0.935	1.091	0.855	0.348
$\psi(2S)K$ veto	0.899	0.903	1.029	0.943	0.785	0.857	0.313
K^0_S selection	0.858	0.951	0.147	0.114	5.002	6.441	0.402
$\mathcal{F}$	0.924	0.882	0.683	0.340	1.248	2.509	0.809
$\mathcal{R}_{B\overline{B}}$	0.929	0.947	0.558	0.840	1.546	1.028	0.538
Muon ID	0.596	0.597	0.330	0.130	1.075	2.734	1.368
ΔE	0.947	0.963	0.015	0.059	60.956	15.239	0.342
$m_{\rm ES}$	0.989	0.988	0.000	0.000	-	-	-
Total	0.091	0.144	0.000	0.000	-	-	-

Table 7.21: Ntuple cut efficiencies for the $B^0 \to K^0_S\mu^+\mu^-$ mode. Event selection criteria are applied one after another, with the efficiency defined as the fraction of events passing the previous set of event selection criteria plus a new requirement. $e_{\rm sig}$ and $e_{\rm fast}$ are the requirement efficiency for the truth tagged SP3 and Pravda signal Monte Carlo, respectively; $e_{B\overline{B}}$, for SP3 generic $B\overline{B}$; and $e_{\rm offres}$, for the off-resonance data sample.

Event selection	Value
Tracking	e^+, e^- GoodTracksLoose π GoodTrackVeryLoose
Conversions	if $R > 2$ cm, e^+, e^- must fail criteria of Table 7.5 of Section 7.10
$(J/\psi, \psi(2S))K^{*+}$ veto	see Table 7.6 of Section 7.11
$(J/\psi, \psi(2S))K^0_S$ feed-up	see Table 7.7 of Section 7.11
K^{*+} mass	$\lvert m_{K^{*+}} - 0.892\rvert < 0.075$ GeV/c^2
Fisher $\mathcal{F}$	$\mathcal{F} < 6.5$
K^0_S selection	See Table 7.4 of Section 7.6
Lepton ID	very tight electrons
Pion ID	NotAKaon (veto)
$B\overline{B}$ suppression $\mathcal{R}_{B\overline{B}}$	$\log(\mathcal{R}_{B\overline{B}}) >$ -1.3
ΔE	-0.11 $< \Delta E <$ 0.05 GeV
$m_{\rm ES}$ mass	$5.272 < m_{\rm ES} < 5.286$ GeV/c^2

Table 7.22: Ntuple cut definitions for the $B^+ \to K^{*+}e^+e^-$ $(K^{*+} \to K^0_S\pi^+)$ mode.

Selection	$e_{\rm sig}$	$e_{\rm fast}$	$e_{B\overline{B}}$	$e_{\rm offres}$	$\frac{e^2_{\rm sig}}{e_{B\overline{B}}}$	$\frac{e^2_{\rm sig}}{e_{\rm offres}}$	$\frac{N_{B\overline{B}}}{N_{\rm offres}}$
Skim & code selection	0.301	0.435	0.012	0.006	7.666	15.637	0.656
Conversions	0.972	0.972	0.985	0.830	0.960	1.139	0.779
Tracking	0.950	1.000	0.631	0.744	1.432	1.214	0.660
$J/\psi K^*$ veto	0.850	0.848	0.491	0.745	1.471	0.970	0.435
$\psi(2S)K^*$ veto	0.867	0.854	0.647	0.818	1.163	0.919	0.344
$J/\psi K^0_S$ feed-up	0.973	0.978	0.998	0.999	0.949	0.949	0.344
K^{*+} mass	0.861	0.866	0.373	0.391	1.987	1.892	0.328
$\mathcal{F}$	0.919	0.869	0.901	0.554	0.936	1.523	0.533
$\mathcal{R}_{B\overline{B}}$	0.779	0.747	0.382	0.755	1.589	0.804	0.270
K^0_S selection	0.808	0.933	0.123	0.111	5.311	5.903	0.300
Electron ID	0.835	0.835	0.283	0.048	2.463	14.555	1.771
Pion ID	0.981	0.982	0.898	0.824	1.071	1.167	1.930
ΔE	0.783	0.761	0.177	0.000	3.458	-	-
$m_{\rm ES}$	0.985	0.980	0.000	-	-	-	-
Total	0.063	0.096	0.000	0.000	-	-	-

Table 7.23: Ntuple cut efficiencies for the $B^+ \to K^{*+}e^+e^-$ $(K^{*+} \to K^0_S\pi^+)$ mode. Event selection criteria are applied one after another, with the efficiency defined as the fraction of events passing the previous set of event selection criteria plus a new requirement. $e_{\rm sig}$ and $e_{\rm fast}$ are the requirement efficiency for the truth tagged SP3 and Pravda signal Monte Carlo, respectively; $e_{B\overline{B}}$, for SP3 generic $B\overline{B}$; and $e_{\rm offres}$, for the off-resonance data sample.

Event selection	Value
Tracking	μ^+, μ^- GoodTracksLoose π GoodTrackVeryLoose
$(J/\psi, \psi(2S))K^{*+}$ veto	see Table 7.6 of Section 7.11
$(J/\psi, \psi(2S))K^0_S$ feed-up	see Table 7.7 of Section 7.11
K^{*+} mass	$\lvert m_{K^{*+}} - 0.892\rvert < 0.075$ GeV/c^2
Fisher $\mathcal{F}$	$\mathcal{F} < 5.2$
K^0_S selection	See Table 7.4 of Section 7.6
Lepton ID	very tight muons
Pion ID	NotAKaon (veto)
$B\overline{B}$ suppression $\mathcal{R}_{B\overline{B}}$	$\log(\mathcal{R}_{B\overline{B}}) >$ -1.8
ΔE	-0.07 $< \Delta E <$ 0.05 GeV
m_{ES} mass	5.272 $< m_{\mathrm{ES}} <$ 5.286 GeV/c^2

Table 7.24: Ntuple cut definitions for the $B^+ \to K^{*+}\mu^+\mu^-$ ($K^{*+} \to K^0_S\pi^+$) mode.

Selection	e_{sig}	e_{fast}	$e_{B\overline{B}}$	e_{offres}	$\frac{e^2_{\mathrm{sig}}}{e_{B\overline{B}}}$	$\frac{e^2_{\mathrm{sig}}}{e_{\mathrm{offres}}}$	$\frac{N_{B\overline{B}}}{N_{\mathrm{offres}}}$
Skim & code selection	0.225	0.331	0.009	0.004	5.716	12.462	0.702
Tracking	0.953	1.000	0.753	0.699	1.207	1.300	0.756
$J/\psi K^*$ veto	0.890	0.892	0.667	0.900	1.187	0.880	0.560
$\psi(2S)K^*$ veto	0.850	0.844	0.797	0.823	0.908	0.878	0.542
$J/\psi K^0_S$ feed-up	0.976	0.978	0.999	0.999	0.952	0.952	0.542
K^* mass	0.864	0.866	0.388	0.396	1.924	1.886	0.531
$\mathcal{F}$	0.843	0.809	0.760	0.278	0.936	2.554	1.451
$\mathcal{R}_{B\overline{B}}$	0.830	0.828	0.475	0.782	1.449	0.880	0.881
K^0_S selection	0.789	0.917	0.101	0.090	6.134	6.947	0.998
Muon ID	0.602	0.603	0.291	0.131	1.247	2.767	2.214
Pion ID	0.981	0.983	0.902	0.842	1.066	1.143	2.373
ΔE	0.950	0.960	0.036	0.188	25.061	4.817	0.456
m_{ES}	0.986	0.988	0.000	0.000	-	-	-
Total	0.042	0.073	0.000	0.000	-	-	-

Table 7.25: Ntuple cut efficiencies for the $B^+ \to K^{*+}\mu^+\mu^-$ ($K^{*+} \to K^0_S\pi^+$) mode. Event selection criteria are applied one after another, with the efficiency defined as the fraction of events passing the previous set of event selection criteria plus a new requirement. e_{sig} and e_{fast} are the requirement efficiency for the truth tagged SP3 and Pravda signal Monte Carlo, respectively; $e_{B\overline{B}}$, for SP3 generic $B\overline{B}$; and e_{offres}, for the off-resonance data sample.

Event selection	Value
Tracking	e^+, e^- GoodTracksLoose, K GoodTrackVeryLoose
Conversions	if $R > 2$ cm, e^+, e^- must fail criteria of Table 7.5 of Section 7.10
$(J/\psi, \psi(2S))K^{*+}$ veto	see Table 7.6 of Section 7.11
$(J/\psi, \psi(2S))K^+$ feed-up	see Table 7.7 of Section 7.11
K^{*+} mass	$\lvert m_{K^{*+}} - 0.892\rvert < 0.075$ GeV/c^2
Fisher $\mathcal{F}$	$\mathcal{F} < 3.3$
Lepton ID	very tight electrons
Kaon ID	very tight kaon
π^0 selection	see Section 7.7
$B\overline{B}$ suppression $\mathcal{R}_{B\overline{B}}$	$\log(\mathcal{R}_{B\overline{B}}) >$ -1.2
ΔE	-0.11 $< \Delta E <$ 0.05 GeV
$m_{\rm ES}$ mass	$5.272 < m_{\rm ES} < 5.286$ GeV/c^2

Table 7.26: Ntuple cut definitions for the $B^+ \to K^{*+}e^+e^-$ ($K^{*+} \to K^+\pi^0$) mode.

Selection	$e_{\rm sig}$	$e_{\rm fast}$	$e_{B\overline{B}}$	$e_{\rm offres}$	$\frac{e^2_{\rm sig}}{e_{B\overline{B}}}$	$\frac{e^2_{\rm sig}}{e_{\rm offres}}$	$\frac{N_{B\overline{B}}}{N_{\rm offres}}$
Skim & code selection	0.247	0.377	0.025	0.012	2.446	4.914	0.646
Conversions	0.977	0.975	0.988	0.854	0.966	1.119	0.748
Tracking	0.945	1.000	0.683	0.818	1.310	1.092	0.624
$J/\psi K^*$ veto	0.842	0.824	0.508	0.754	1.396	0.940	0.420
$\psi(2S)K^*$ veto	0.856	0.844	0.465	0.734	1.578	0.999	0.266
$J/\psi K$ feed-up	0.971	0.973	0.992	0.998	0.950	0.945	0.265
π^0 mass	0.930	0.901	0.692	0.676	1.250	1.280	0.271
K^* mass	0.860	0.867	0.388	0.375	1.909	1.970	0.280
$\mathcal{F}$	0.927	0.879	0.930	0.695	0.924	1.237	0.374
$\mathcal{R}_{B\overline{B}}$	0.680	0.660	0.303	0.555	1.524	0.834	0.205
Electron ID	0.844	0.846	0.196	0.060	3.640	11.880	0.668
Kaon ID	0.746	0.741	0.227	0.119	2.456	4.696	1.277
ΔE	0.724	0.709	0.107	0.100	4.893	5.242	1.368
$m_{\rm ES}$	0.965	0.963	0.000	0.000	-	-	-
Total	0.036	0.048	0.000	0.000	-	-	-

Table 7.27: Ntuple cut efficiencies for the $B^+ \to K^{*+}e^+e^-$ ($K^{*+} \to K^+\pi^0$) mode. Event selection criteria are applied one after another, with the efficiency defined as the fraction of events passing the previous set of event selection criteria plus a new requirement. $e_{\rm sig}$ and $e_{\rm fast}$ are the requirement efficiency for the truth tagged SP3 and Pravda signal Monte Carlo, respectively; $e_{B\overline{B}}$, for SP3 generic $B\overline{B}$; and $e_{\rm offres}$, for the off-resonance data sample.

Event selection	Value
Tracking	μ^+, μ^- GoodTracksLoose, K GoodTrackVeryLoose
$(J/\psi, \psi(2S))K^{*+}$ veto	see Table 7.6 of Section 7.11
$(J/\psi, \psi(2S))K^+$ feed-up	see Table 7.7 of Section 7.11
K^{*+} mass	$\lvert m_{K^{*+}} - 0.892\rvert < 0.075$ GeV/c^2
Fisher $\mathcal{F}$	$\mathcal{F} < 6.8$
Lepton ID	very tight muons
Kaon ID	very tight kaon
π^0 selection	see Section 7.7
$B\overline{B}$ suppression $\mathcal{R}_{B\overline{B}}$	$\log(\mathcal{R}_{B\overline{B}}) >$ -1.9
ΔE	-0.07 $< \Delta E <$ 0.05 GeV
$m_{\rm ES}$ mass	5.272 $< m_{\rm ES} <$ 5.286 GeV/c^2

Table 7.28: Ntuple cut definitions for the $B^+ \to K^{*+}\mu^+\mu^-$ ($K^{*+} \to K^+\pi^0$) mode.

Selection	$e_{\rm sig}$	$e_{\rm fast}$	$e_{B\overline{B}}$	$e_{\rm offres}$	$\frac{e^2_{\rm sig}}{e_{B\overline{B}}}$	$\frac{e^2_{\rm sig}}{e_{\rm offres}}$	$\frac{N_{B\overline{B}}}{N_{\rm offres}}$
Skim & code selection	0.188	0.292	0.021	0.011	1.681	3.333	0.638
Tracking	0.947	1.000	0.814	0.747	1.102	1.201	0.695
$J/\psi K^*$ veto	0.897	0.896	0.690	0.914	1.165	0.880	0.525
$\psi(2S)K^*$ veto	0.842	0.843	0.759	0.816	0.935	0.869	0.488
$J/\psi K$ feed-up	0.973	0.973	0.998	0.999	0.948	0.948	0.488
π^0 mass	0.930	0.898	0.682	0.679	1.270	1.275	0.490
K^* mass	0.853	0.850	0.390	0.381	1.866	1.910	0.502
$\mathcal{F}$	0.711	0.666	0.602	0.165	0.841	3.069	1.833
$\mathcal{R}_{B\overline{B}}$	0.799	0.813	0.444	0.730	1.439	0.875	1.115
Muon ID	0.604	0.588	0.292	0.115	1.250	3.178	2.834
Kaon ID	0.698	0.701	0.180	0.058	2.703	8.364	8.770
ΔE	0.861	0.853	0.035	0.091	20.922	8.161	3.421
$m_{\rm ES}$	0.976	0.984	0.200	0.000	4.766	-	-
Total	0.021	0.031	0.000	0.000	-	-	-

Table 7.29: Ntuple cut efficiencies for the $B^+ \to K^{*+}\mu^+\mu^-$ ($K^{*+} \to K^+\pi^0$) mode. Event selection criteria are applied one after another, with the efficiency defined as the fraction of events passing the previous set of event selection criteria plus a new cut. $e_{\rm sig}$ and $e_{\rm fast}$ are the cut efficiency for the truth tagged SP3 and Pravda signal Monte Carlo, respectively; $e_{B\overline{B}}$, for SP3 generic $B\overline{B}$; and $e_{\rm offres}$, for the off-resonance data sample.

7.16 Event Selection Summary

This section summarizes the effect of the analysis selection criteria on both signal and background and gives an estimate of the expected signal and background yields in the entire Run 1 data set after applying all of the cuts. It also discusses the categories of background events that tend to pass various subsets of the analysis selection requirements.

7.16.1 Dalitz Plots

Dalitz plots offer one of the best ways to observe the effect of the analysis selection criteria on both signal and background. The Dalitz plot variables used in this section are the same as those described in Section 6.2: q^2, the invariant dilepton mass squared, and $E_{\rm lep}$, the lepton energy in the B meson rest frame. For the latter, the positive lepton is selected for the modes $B^+ \to K^+\ell^+\ell^-$, $B^0 \to K^{*0}\ell^+\ell^-$, and $B^+ \to K^{*+}\ell^+\ell^-$, and the negative lepton is selected for the charge-conjugate channels. For the $B^0 \to K^0_S\ell^+\ell^-$ mode, the positive lepton is used.

Figures 7.16 and 7.17 show the Dalitz plots for truth tagged signal SP3 Monte Carlo, for the charged track modes and the modes with a K^0_S or a π^0, respectively. All of the event selection criteria have been applied except those on ΔE and $m_{\rm ES}$. Several features of the plots are easily recognizable. The two large horizontal empty bands correspond to the charmonium vetoes. The "chopped" edges appear because of the requirement that the electron momenta be greater than 0.5 GeV/c, and the muon momenta, greater than 1.0 GeV/c, in the laboratory frame. The vertical band at $E_{\rm lep} \sim 1.7$ GeV, visible particularly well in the $B^+ \to K^+\mu^+\mu^-$ plot in Fig. 7.16b, corresponds to the charmonium swap veto: for $B^+ \to K^+\mu^+\mu^-$, the swapped mass is directly proportional to $E^2_{\rm lep} \equiv E^2_{\mu^+}$, since $m_{\rm swap} \equiv m^2_{K^+\mu^-} = (m^2_B - m^2_{\mu^+}) - 2m_B E^2_{\mu^+}$. There is a hint of this band on the $K^+e^+e^-$ plot as well, but because the density of data points on the plot is larger, it is not as pronounced. Figures 7.18 and 7.19 show the one-dimensional projections of

q^2 for the charged track modes and the modes with a K^0_S or a π^0, respectively. One can see that after all the event selection criteria the pole at $q^2 = 0$ is preserved for the $K^* e^+ e^-$ modes, but not for the $K^* \mu^+ \mu^-$ ones.

In order to observe the effect of the analysis selection criteria on the $B\overline{B}$ background, the entire available fast Monte Carlo $B\overline{B}$ sample is used to increase statistics. The Dalitz plots for this background are shown in Figs. 7.20 and 7.21 for the charged track modes and the modes with a K^0_S or a π^0, respectively. These plots indicate that $B\overline{B}$ events tend to concentrate in the middle of the Dalitz plot. The q^2 projections are shown in Figs. 7.22 and 7.23 for the charged track modes and the modes with a K^0_S or a π^0, respectively.

Figures 7.24 and 7.25 show where continuum $c\overline{c}$ and uds events land in the Dalitz plots, for the charged track modes and the modes with a K^0_S or a π^0, respectively. Fast continuum $udsc$ Monte Carlo is used to increase statistics. Continuum events are distributed mostly around the boundary of the Dalitz plots. This is a reflection of the fact that the leptons typically come from two opposite jets, leading to $\cos\theta = \pm\, 1$, where θ is the lepton decay angle. They can also come from the same jet, which would give a q^2 close to zero. Figures 7.26 and 7.27 show the one-dimensional q^2 projections for the charged track modes and the modes with a K^0_S or a π^0, respectively. There are very large peaks at low q^2. The events with low q^2 in the electron channels typically involve a di-lepton pair with one real and one fake lepton. An example of an event resulting in a low q^2 in the $B^+ \to K^+ e^+ e^-$ channel is the semileptonic D decay, $D^0 \to K^- e^+ \nu$, where e^+ is real and K^- is faking an electron. Other categories of events that are likely to result in low q^2 are photon conversions, π^0 Dalitz decays, and η and ω dilepton decays. For the muon channels, in addition to events with one real and one fake muon, the low q^2 region includes a significant fraction of events where both muons are fake.

It is interesting to note that the distributions of background events in the Dalitz plots are almost complementary to each other: the $B\overline{B}$ background tends to concentrate in the middle of the Dalitz plot, while continuum prefers the edges.

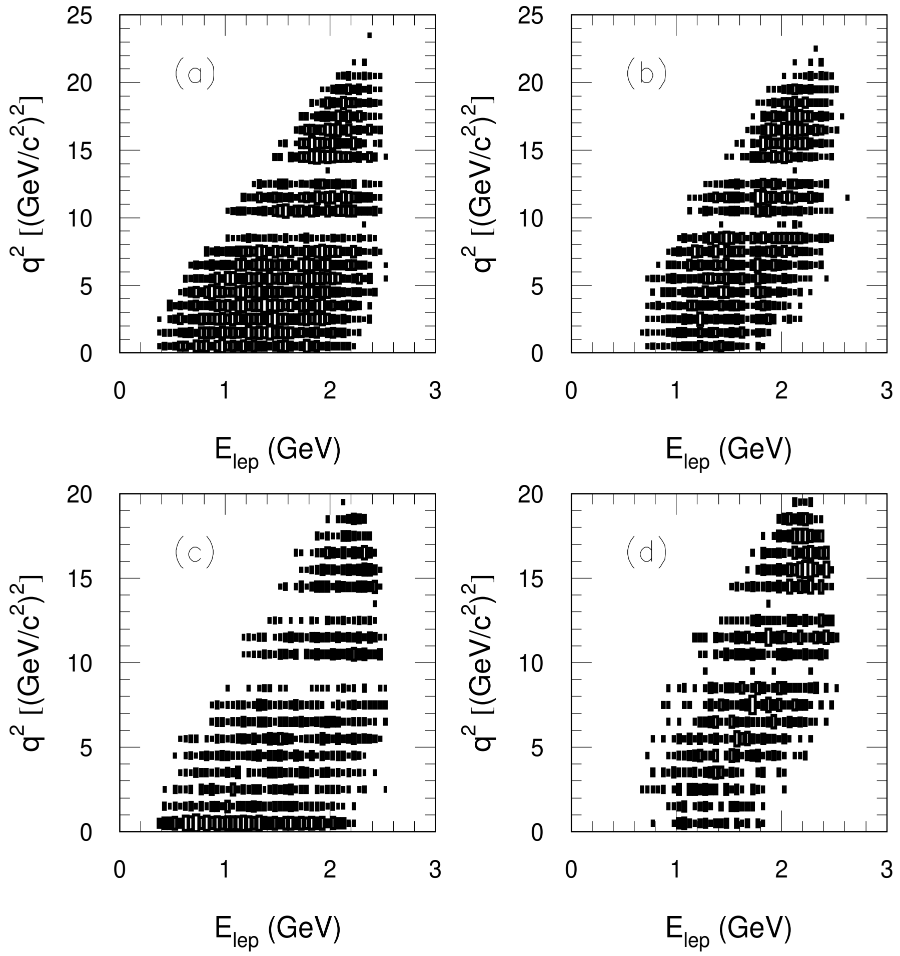


Figure 7.16: Dalitz plots for truth tagged SP3 signal Monte Carlo for the charged track modes. (a) $B^+ \to K^+ e^+ e^-$, (b) $B^+ \to K^+ \mu^+ \mu^-$, (c) $B^0 \to K^{*0} e^+ e^-$, and (d) $B^0 \to K^{*0} \mu^+ \mu^-$. All of the event selection criteria have been applied except those on ΔE and m_{ES}.

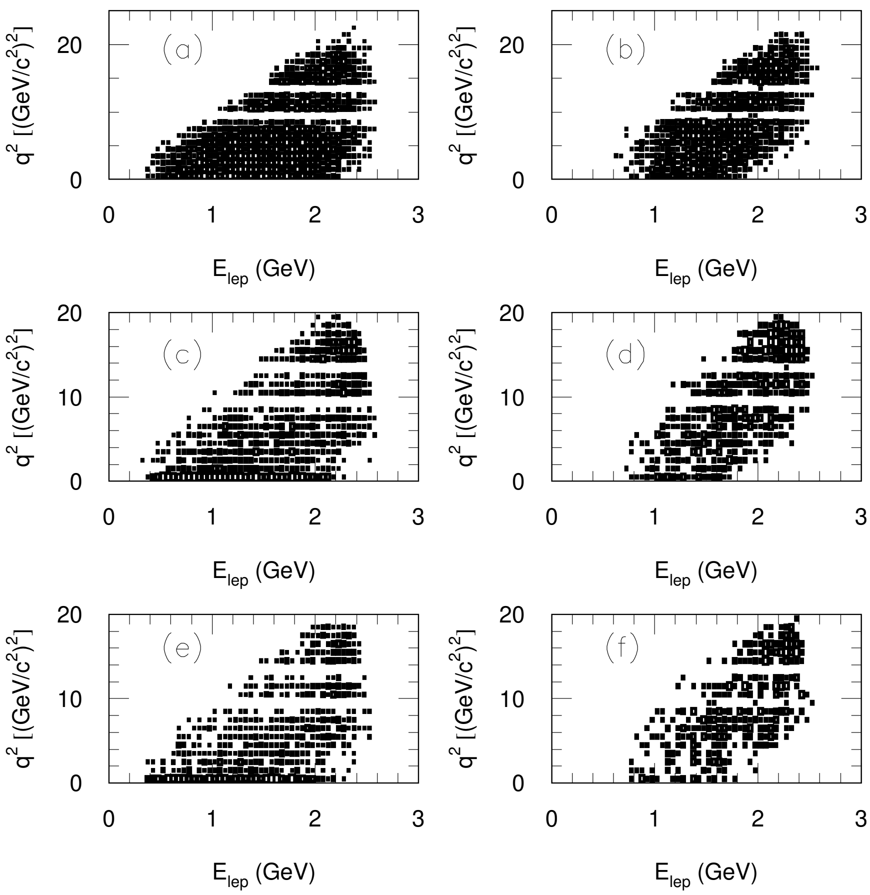


Figure 7.17: Dalitz plots for truth tagged SP3 signal Monte Carlo for the modes with a K_S^0 or a π^0. (a) $B^0 \to K_S^0 e^+e^-$, (b) $B^0 \to K_S^0 \mu^+\mu^-$, (c) $B^+ \to K^{*+} e^+e^-$ with $K^{*+} \to K_S^0\pi^+$, (d) $B^+ \to K^{*+}\mu^+\mu^-$ with $K^{*+} \to K_S^0\pi^+$, (e) $B^+ \to K^{*+}e^+e^-$ with $K^{*+} \to K^+\pi^0$, and (f) $B^+ \to K^{*+}\mu^+\mu^-$ with $K^{*+} \to K^+\pi^0$. All of the event selection criteria have been applied except those on ΔE and $m_{\rm ES}$.

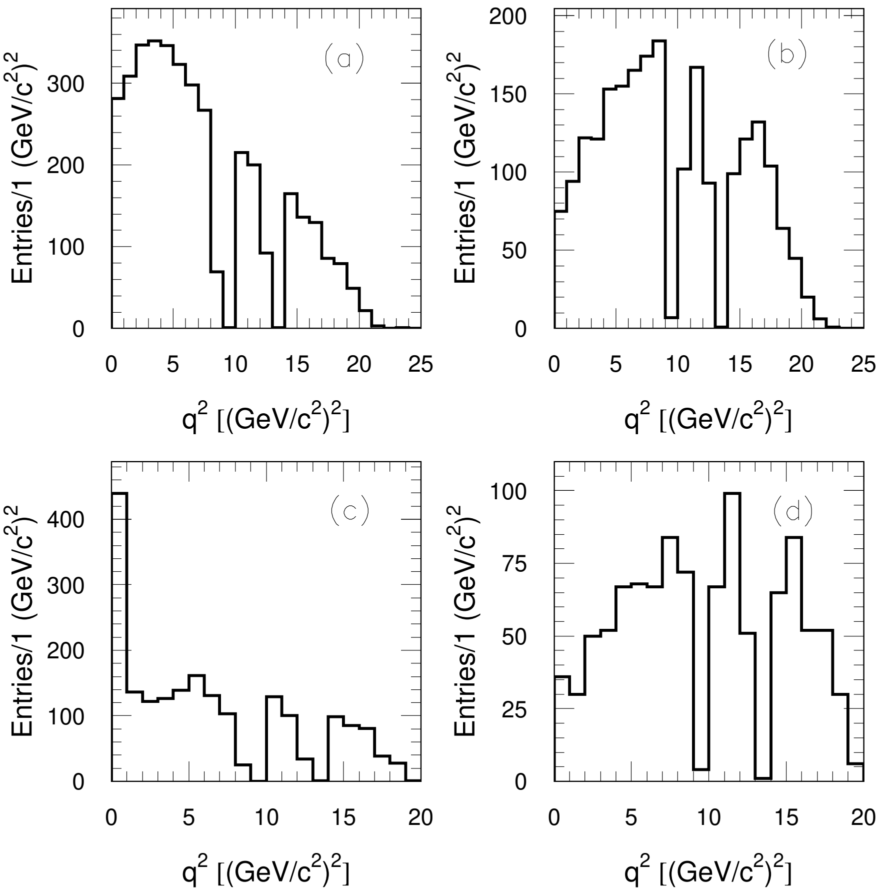


Figure 7.18: q^2 distributions for truth tagged SP3 signal Monte Carlo for the charged track modes. (a) $B^+ \to K^+e^+e^-$, (b) $B^+ \to K^+\mu^+\mu^-$, (c) $B^0 \to K^{*0}e^+e^-$, and (d) $B^0 \to K^{*0}\mu^+\mu^-$. All of the event selection criteria have been applied except those on ΔE and $m_{\rm ES}$.

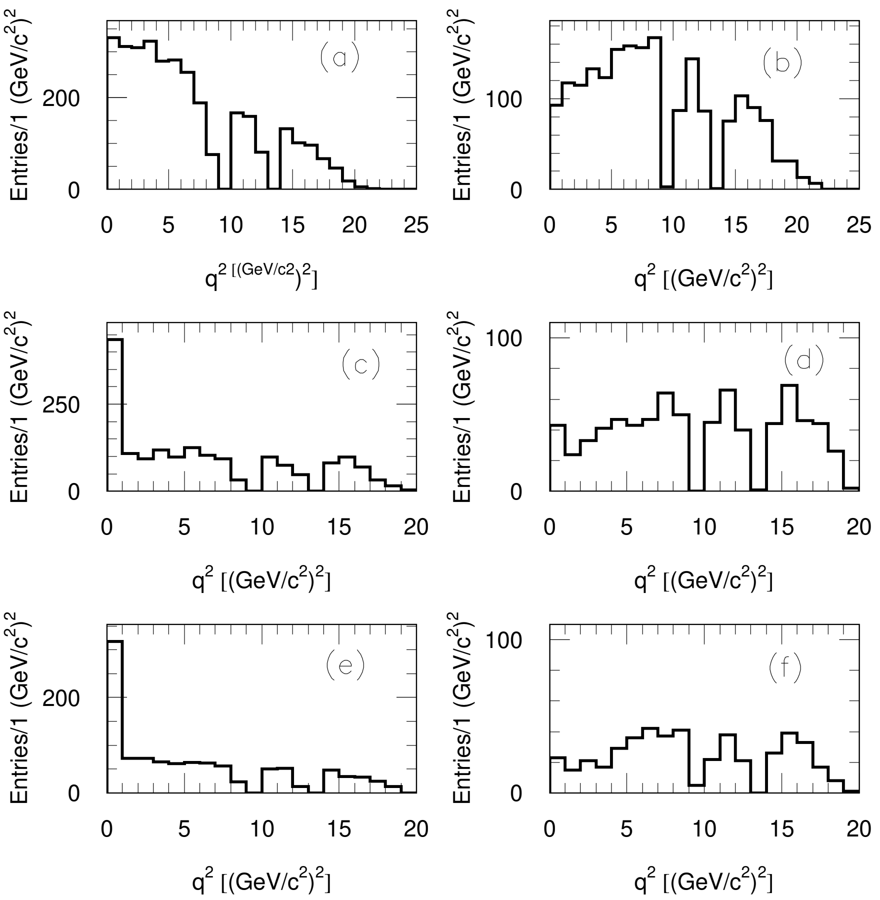


Figure 7.19: q^2 distributions for truth tagged signal Monte Carlo for the modes with a K^0_S or a π^0. (a) $B^0 \to K^0_S e^+ e^-$, (b) $B^0 \to K^0_S \mu^+ \mu^-$, (c) $B^+ \to K^{*+} e^+ e^-$ with $K^{*+} \to K^0_S \pi^+$, (d) $B^+ \to K^{*+} \mu^+ \mu^-$ with $K^{*+} \to K^0_S \pi^+$, (e) $B^+ \to K^{*+} e^+ e^-$ with $K^{*+} \to K^+ \pi^0$, and (f) $B^+ \to K^{*+} \mu^+ \mu^-$ with $K^{*+} \to K^+ \pi^0$. All of the event selection criteria have been applied except those on ΔE and m_{ES}.

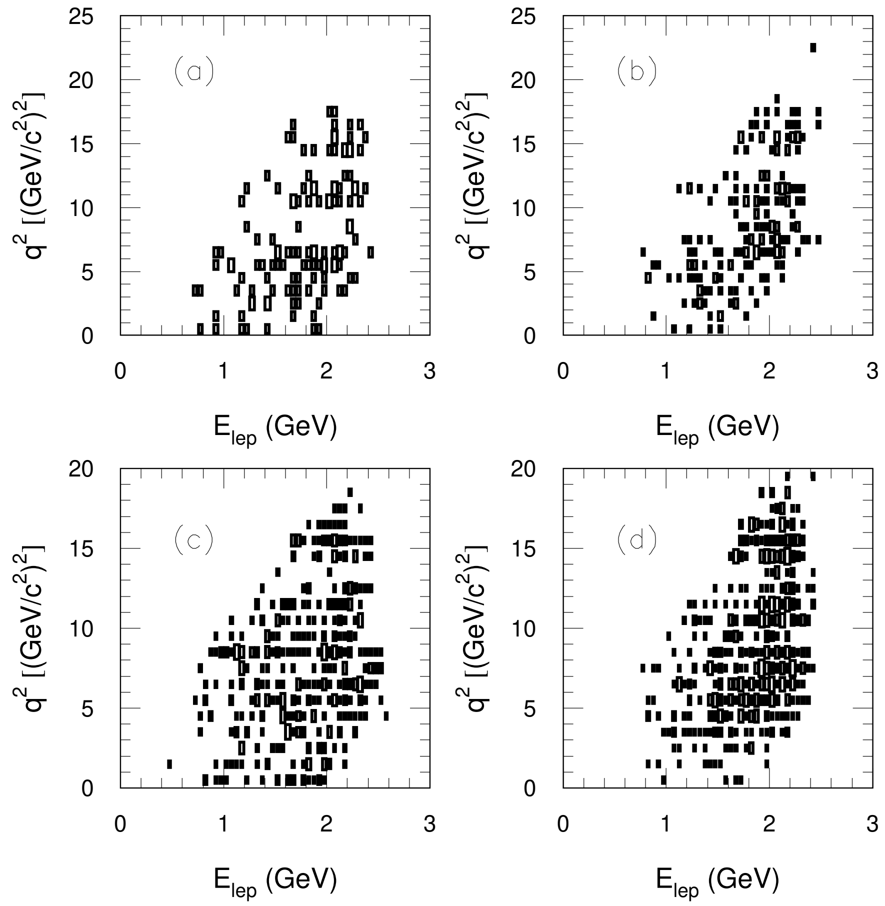


Figure 7.20: Dalitz plots for Pravda Monte Carlo generic $B\overline{B}$ Monte Carlo sample for the charged track modes. (a) $B^+ \to K^+e^+e^-$, (b) $B^+ \to K^+\mu^+\mu^-$, (c) $B^0 \to K^{*0}e^+e^-$, and (d) $B^0 \to K^{*0}\mu^+\mu^-$. All of the event selection criteria have been applied with loose cuts on ΔE and $m_{\rm ES}$: $|\Delta E| < 0.25$ GeV and $m_{\rm ES} > 5.2$ GeV/c^2.

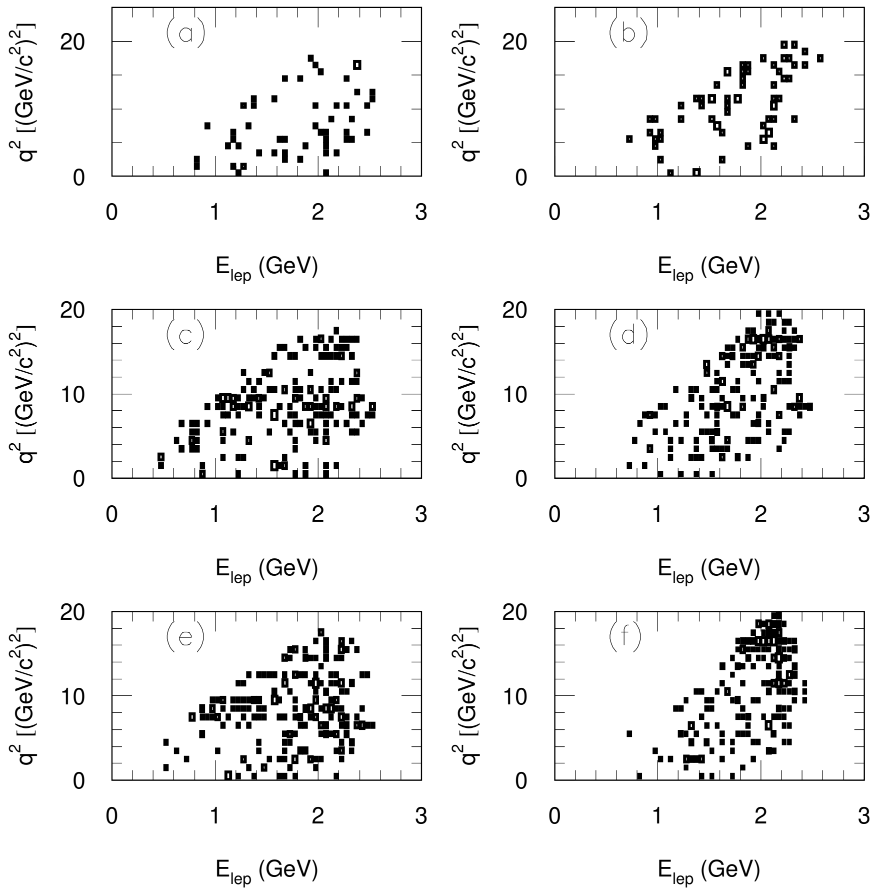


Figure 7.21: Dalitz plots for the generic $B\overline{B}$ Pravda Monte Carlo sample for the modes with a K_S^0 or a π^0. (a) $B^0 \to K_S^0 e^+e^-$, (b) $B^0 \to K_S^0\mu^+\mu^-$, (c) $B^+ \to K^{*+}e^+e^-$ with $K^{*+} \to K_S^0\pi^+$, (d) $B^+ \to K^{*+}\mu^+\mu^-$ with $K^{*+} \to K_S^0\pi^+$, (e) $B^+ \to K^{*+}e^+e^-$ with $K^{*+} \to K^+\pi^0$, and (f) $B^+ \to K^{*+}\mu^+\mu^-$ with $K^{*+} \to K^+\pi^0$. All of the event selection criteria have been applied with loose cuts on ΔE and m_{ES}: $|\Delta E| < 0.25$ GeV and $m_{\mathrm{ES}} > 5.2$ GeV$/c^2$.

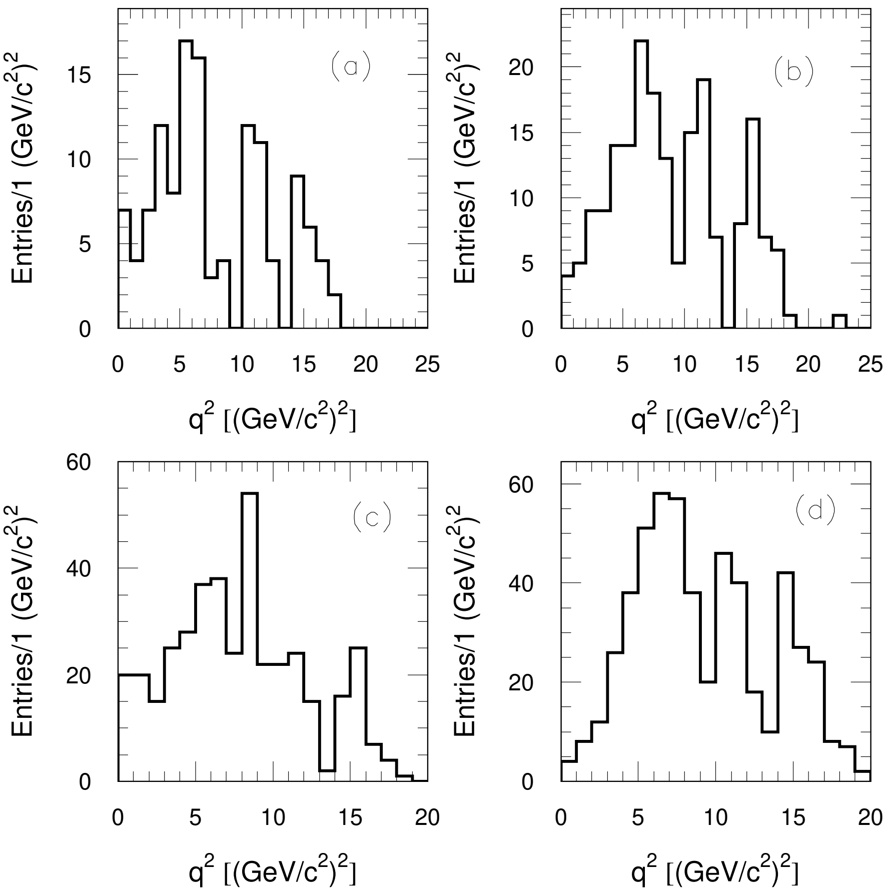


Figure 7.22: q^2 distributions for the generic $B\overline{B}$ Pravda Monte Carlo sample for the charged track modes. (a) $B^+ \to K^+e^+e^-$, (b) $B^+ \to K^+\mu^+\mu^-$, (c) $B^0 \to K^{*0}e^+e^-$, and (d) $B^0 \to K^{*0}\mu^+\mu^-$. All of the event selection criteria have been applied with loose cuts on ΔE and m_{ES}: $|\Delta E| < 0.25$ GeV and $m_{\text{ES}} > 5.2$ GeV/c^2.

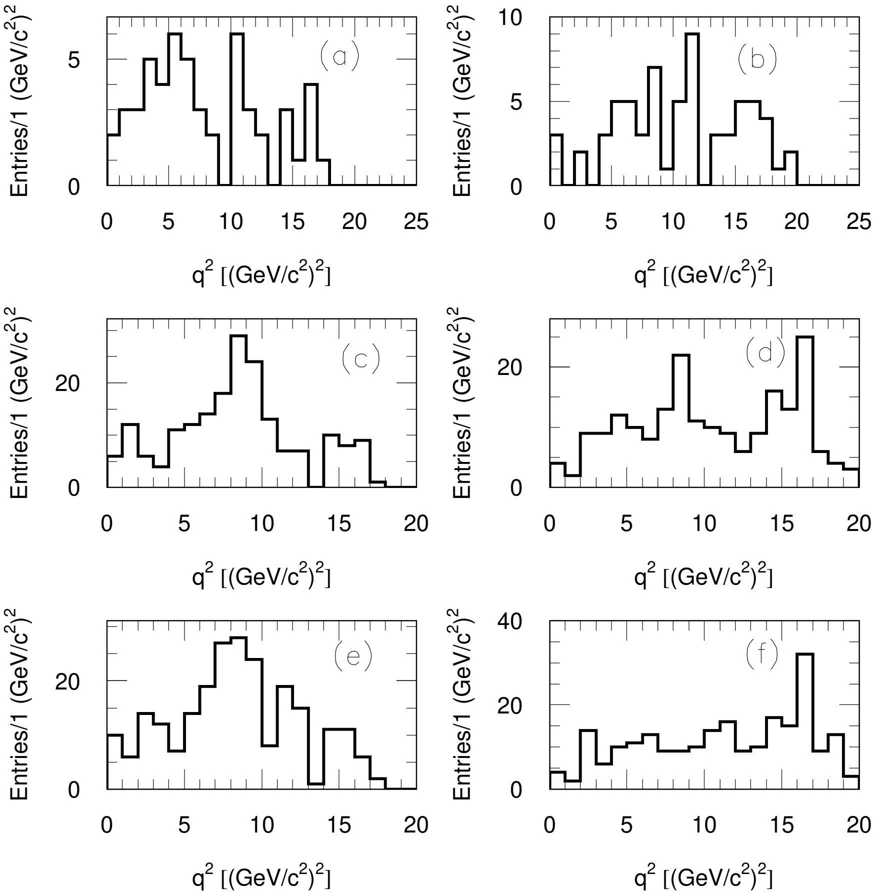


Figure 7.23: q^2 distributions for the generic $B\overline{B}$ Pravda Monte Carlo sample for the modes with a K^0_S or a π^0. (a) $B^0 \to K^0_S e^+e^-$, (b) $B^0 \to K^0_S \mu^+\mu^-$, (c) $B^+ \to K^{*+}e^+e^-$ with $K^{*+} \to K^0_S\pi^+$, (d) $B^+ \to K^{*+}\mu^+\mu^-$ with $K^{*+} \to K^0_S\pi^+$, (e) $B^+ \to K^{*+}e^+e^-$ with $K^{*+} \to K^+\pi^0$, and (f) $B^+ \to K^{*+}\mu^+\mu^-$ with $K^{*+} \to K^+\pi^0$. All of the event selection criteria have been applied with loose cuts on ΔE and $m_{\rm ES}$: $|\Delta E| < 0.25$ GeV and $m_{\rm ES} > 5.2$ GeV/c^2.

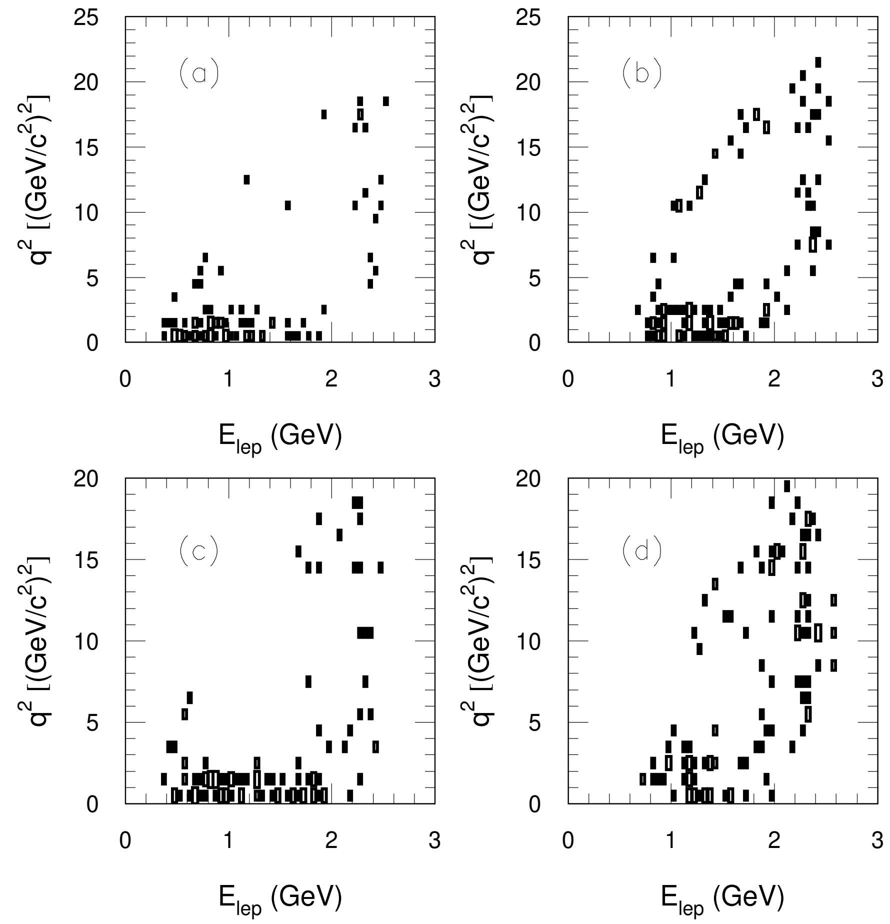


Figure 7.24: Dalitz plots for the continuum ($c\bar{c}$ + uds) Pravda Monte Carlo sample for the charged track modes. (a) $B^+ \to K^+e^+e^-$, (b) $B^+ \to K^+\mu^+\mu^-$, (c) $B^0 \to K^{*0}e^+e^-$, and (d) $B^0 \to K^{*0}\mu^+\mu^-$. All of the event selection criteria have been applied, with loose cuts on ΔE and $m_{\rm ES}$: $|\Delta E| < 0.25$ GeV and $m_{\rm ES} > 5.2$ GeV$/c^2$.

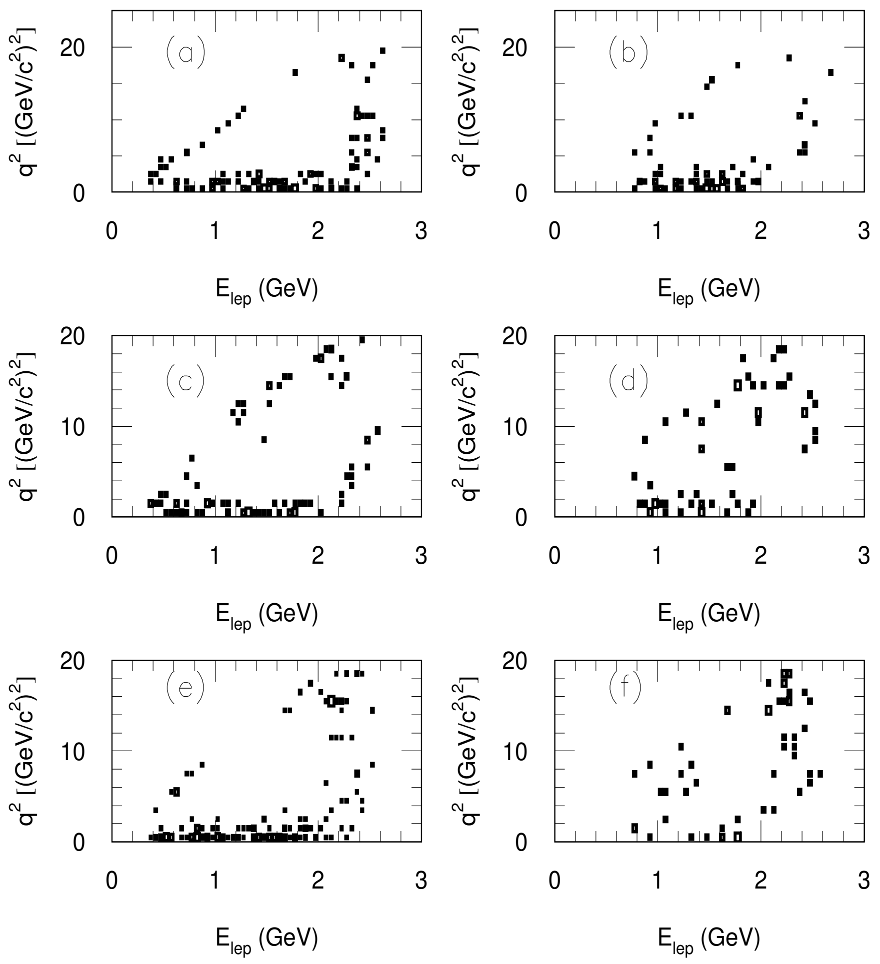


Figure 7.25: Dalitz plots for the continuum ($c\bar{c}$ + uds) Pravda Monte Carlo sample for the modes with a K_S^0 or a π^0. (a) $B^0 \to K_S^0 e^+ e^-$, (b) $B^0 \to K_S^0 \mu^+ \mu^-$, (c) $B^+ \to K^{*+} e^+ e^-$ with $K^{*+} \to K_S^0 \pi^+$, (d) $B^+ \to K^{*+} \mu^+ \mu^-$ with $K^{*+} \to K_S^0 \pi^+$, (e) $B^+ \to K^{*+} e^+ e^-$ with $K^{*+} \to K^+ \pi^0$, and (f) $B^+ \to K^{*+} \mu^+ \mu^-$ with $K^{*+} \to K^+ \pi^0$. All of the event selection criteria have been applied with loose cuts on ΔE and $m_{\rm ES}$: $|\Delta E| < 0.25$ GeV and $m_{\rm ES} > 5.2$ GeV$/c^2$.

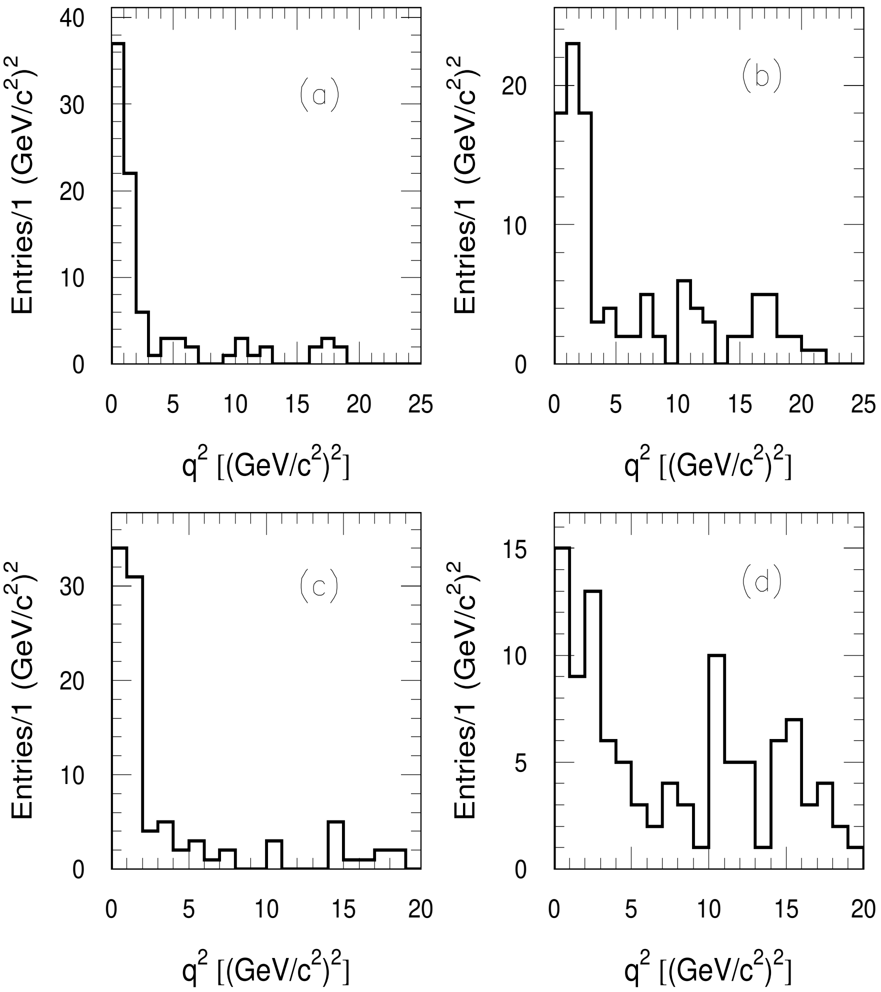


Figure 7.26: q^2 distributions for the continuum ($c\overline{c}$ + uds) Pravda Monte Carlo sample for the charged track modes. (a) $B^+ \to K^+ e^+ e^-$, (b) $B^+ \to K^+ \mu^+ \mu^-$, (c) $B^0 \to K^{*0} e^+ e^-$, and (d) $B^0 \to K^{*0} \mu^+ \mu^-$. All of the event selection criteria have been applied, with loose cuts on $\varDelta E$ and $m_{\rm ES}$: $|\varDelta E| < 0.25$ GeV and $m_{\rm ES} > 5.2$ GeV$/c^2$.

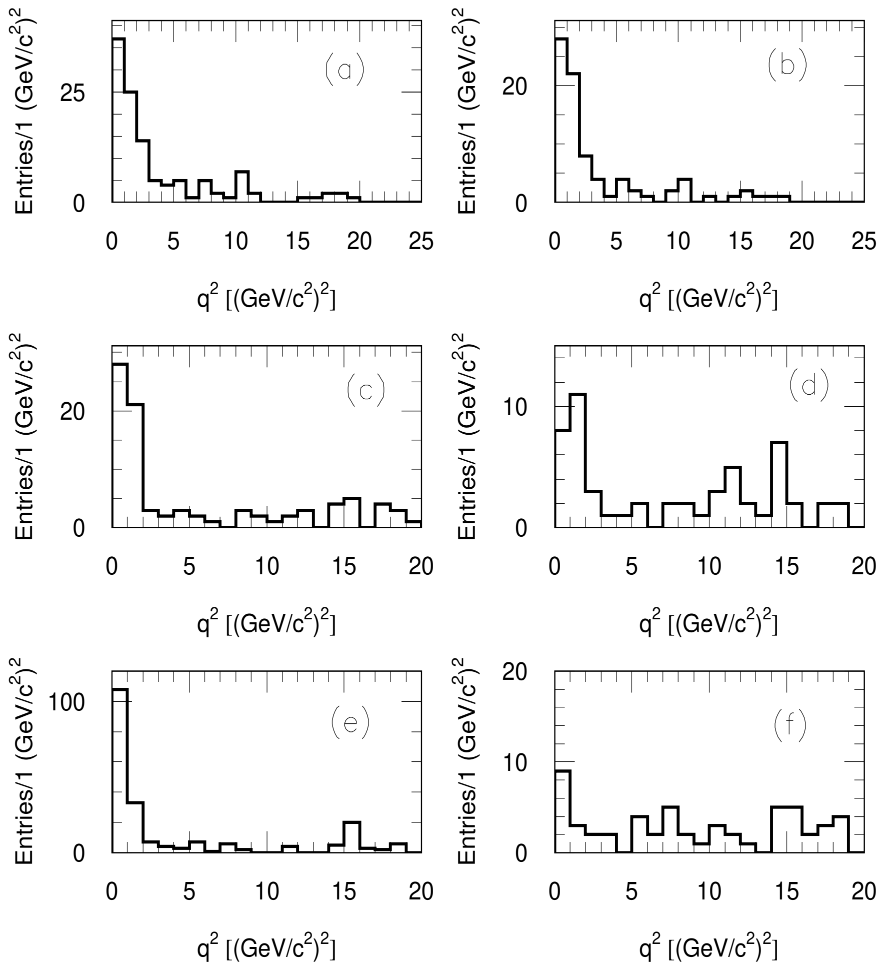


Figure 7.27: q^2 distributions for the continuum ($c\bar{c}$ + uds) Pravda Monte Carlo sample for the modes with a K^0_S or a π^0. (a) $B^0 \to K^0_S e^+e^-$, (b) $B^0 \to K^0_S\mu^+\mu^-$, (c) $B^+ \to K^{*+}e^+e^-$ with $K^{*+} \to K^0_S\pi^+$, (d) $B^+ \to K^{*+}\mu^+\mu^-$ with $K^{*+} \to K^0_S\pi^+$, (e) $B^+ \to K^{*+}e^+e^-$ with $K^{*+} \to K^+\pi^0$, and (f) $B^+ \to K^{*+}\mu^+\mu^-$ with $K^{*+} \to K^+\pi^0$. All of the event selection criteria have been applied with loose cuts on ΔE and $m_{\rm ES}$: $|\Delta E| < 0.25$ GeV and $m_{\rm ES} > 5.2$ GeV/c^2.

Mode	ϵ (%)	# sig	# $B\overline{B}$	# Off-res	S/B
$B^+ \to K^+ e^+ e^-$	14.3 ± 0.4	1.9	0.3	0.7	1.8
$B^0 \to K^0_S e^+ e^-$	13.2 ± 0.4	0.6	0.2	0.4	1.2
$B^0 \to K^{*0} e^+ e^-$	7.9 ± 0.3	2.7	1.9	0.8	1.0
$B^+ \to K^{*+}_{K^0_S} e^+ e^-$	6.6 ± 0.3	0.7	0.9	0.3	0.6
$B^+ \to K^{*+}_{\pi^0} e^+ e^-$	4.3 ± 0.2	0.7	1.4	1.2	0.3
$B^+ \to K^+ \mu^+ \mu^-$	9.6 ± 0.3	1.3	0.3	0.6	1.3
$B^0 \to K^0_S \mu^+ \mu^-$	9.1 ± 0.3	0.4	0.1	0.2	1.3
$B^0 \to K^{*0} \mu^+ \mu^-$	6.5 ± 0.3	1.9	1.0	0.4	1.4
$B^+ \to K^{*+}_{K^0_S} \mu^+ \mu^-$	4.4 ± 0.2	0.4	0.5	0.5	0.4
$B^+ \to K^{*+}_{\pi^0} \mu^+ \mu^-$	2.6 ± 0.2	0.4	0.5	0.3	0.5
TOTAL		11.0	12.5		0.9
TOTAL w/o the π^0 modes		9.9	9.0		1.1

Table 7.30: Summary of signal efficiencies ϵ and the number of expected signal and background events in the Run 1 data set. A ratio of signal (S) to background (B) is also given. The signal used is non-truth tagged SP3 signal Monte Carlo. $K^{*+}_{K^0_S}$ denotes the channel with $K^{*+} \to K^0_S \pi^+$, and $K^{*+}_{\pi^0}$, with $K^{*+} \to K^+ \pi^0$. The efficiencies do not include the branching fractions for the constituent decays.

7.16.2 Summary of Signal Efficiencies and Yields

Table 7.30 summarizes the signal efficiencies and the number of expected signal and background events after applying all of the the analysis selection criteria. The signal efficiencies listed in Table 7.30 are defined as the fraction of non-truth tagged SP3 signal Monte Carlo events in the signal box defined in Section 7.3. They do not include the branching fractions for the constituent decays. The number of expected signal events, S, is determined by counting the number of events in the signal region after applying all of the analysis event selection criteria in SP3 signal Monte Carlo and scaling the number to the Run 1 luminosity using the branching fractions estimates of Ali *et al.* [12]. The number of background events, B, is estimating using the procedure described in Section 7.14, which involves scaling the number of events from regions $A+B+C$ to region A using both SP3 $B\overline{B}$ Monte Carlo and off-resonance data and Pravda Monte Carlo. The resulting expectations are approximately 10 signal events and 9 background events total in all of the modes except the π^0 ones, which are not expected to be averaged into the final result.

7.16.3 Background Categories

It is useful to not only estimate the level of the expected background, but also to understand its composition to check whether there are any particular background event categories that are likely to pass most of the analysis event selection criteria. Of particular concern are the backgrounds that can potentially peak in the signal region. There are two main categories of such backgrounds: the contributions from charmonium events and the hadronic backgrounds that may peak in the signal region if the hadrons in the final state are misidentified as leptons.

Tables 7.31 and 7.32 summarize the number of expected J/ψ and $\psi(2S)$ events, respectively, in the signal region after applying all of the event selection criteria. These estimates were obtained by using SP3 signal $B \to J/\psi(K, K^*)$ and $B \to \psi(2S)(K, K^*)$ Monte Carlo described in Section 5.2. For example, the first entry in Tables 7.31 corresponds to the number of events in the $B^+ \to J/\psi K^+$ sample, where $J/\psi \to e^+e^-$ and $\mu^+\mu^-$, passing all of the $B^+ \to K^+e^+e^-$ selection requirements, and scaled to the data luminosity. The error listed is statistical only. If no events passed all of the event selection criteria, the error corresponds to one observed event. It is evident that no statistically significant charmonium contribution is expected in any of the modes.

In order to estimate the contribution from hadronic backgrounds to the residual background composition, 250,000 events were generated in several specific hadronic final states. Table 7.33 summarizes the expected number of hadronic events in the signal region after applying all of the analysis event selection criteria, scaled to

Mode	$J/\psi K$	$J/\psi K^0_S$	$J/\psi K^{*0}$	$J/\psi K^{*+}_{K^0_S}$	$J/\psi K^{*+}_{\pi^0}$
Ke^+e^-	0.06±0.06	0.00±0.02	0.05±0.05	0.00±0.02	0.00±0.03
$K\mu^+\mu^-$	0.00±0.06	0.00±0.02	0.00±0.05	0.00±0.02	0.00±0.03
$K^{*0}e^+e^-$	0.12±0.09	0.02±0.02	0.05±0.05	0.00±0.02	0.00±0.03
$K^{*0}\mu^+\mu^-$	0.00±0.06	0.00±0.02	0.00±0.05	0.00±0.02	0.03±0.03
$K^0_Se^+e^-$	0.06±0.06	0.00±0.02	0.00±0.05	0.02±0.02	0.00±0.03
$K^0_S\mu^+\mu^-$	0.00±0.06	0.00±0.02	0.00±0.05	0.02±0.02	0.00±0.03
$K^{*+}_{K^0_S}e^+e^-$	0.00±0.06	0.09±0.05	0.00±0.05	0.04±0.03	0.00±0.03
$K^{*+}_{K^0_S}\mu^+\mu^-$	0.00±0.06	0.05±0.03	0.00±0.05	0.00±0.02	0.00±0.03
$K^{*+}_{\pi^0}e^+e^-$	0.00±0.06	0.00±0.02	0.05±0.05	0.00±0.02	0.08±0.05
$K^{*+}_{\pi^0}\mu^+\mu^-$	0.06±0.06	0.00±0.02	0.00±0.05	0.00±0.02	0.00±0.03

Table 7.31: The expected number of $J/\psi(K, K^*)$ events in the signal region after applying all of the analysis event selection criteria in Run 1. $K^{*+}_{K^0_S}$ denotes the channel with $K^{*+} \to K^0_S\pi^+$, and $K^{*+}_{\pi^0}$, with $K^{*+} \to K^+\pi^0$. The charmonium samples used are signal $J/\psi(K, K^*)$ SP3 Monte Carlo. The errors are statistical only. If no events were observed, the error corresponds to one observed event.

the Run 1 luminosity. In no channel does the expected hadronic contribution exceed 0.2 events.

The decays $B^0 \to K^*\gamma$ can also be a source of peaking background if the photon converts into e^+e^-. Having generated 22,000 $B^0 \to K^{*0}\gamma$ events, it is established that the expected number of these events passing all of the $B^0 \to K^{*0}e^+e^-$ event selection criteria in the Run 1 data sample does not surpass 0.1.

Generic $B\overline{B}$ and continuum background is not expected to peak in the signal region. In order to facilitate the description of the effects of the analysis event selection criteria on this background, several specific categories of background have been selected. They are summarized in Table 7.34 for the $B\overline{B}$ background, and in Table 7.35 for the continuum. Figures 7.28 and 7.29 show the $B\overline{B}$ background categories after applying all of the analysis event selection criteria except ΔE and $m_{\rm ES}$,

Mode	$\psi(2S)K$	$\psi(2S)K_S^0$	$\psi(2S)K^{*0}$	$\psi(2S)K^{*+}_{K_S^0}$	$\psi(2S)K^{*+}_{\pi^0}$
Ke^+e^-	0.04±0.04	0.00±0.01	0.00±0.04	0.00±0.01	0.00±0.02
$K\mu^+\mu^-$	0.00±0.04	0.00±0.01	0.00±0.04	0.00±0.02	0.00±0.02
$K^{*0}e^+e^-$	0.00±0.04	0.00±0.01	0.08±0.05	0.00±0.01	0.02±0.02
$K^{*0}\mu^+\mu^-$	0.00±0.04	0.00±0.01	0.00±0.04	0.00±0.02	0.00±0.02
$K_S^0e^+e^-$	0.00±0.04	0.00±0.01	0.00±0.04	0.00±0.01	0.00±0.02
$K_S^0\mu^+\mu^-$	0.00±0.04	0.00±0.01	0.00±0.04	0.00±0.02	0.00±0.02
$K^{*+}_{K_S^0}e^+e^-$	0.00±0.04	0.01±0.01	0.04±0.04	0.00±0.01	0.00±0.02
$K^{*+}_{K_S^0}\mu^+\mu^-$	0.00±0.04	0.00±0.01	0.00±0.04	0.00±0.02	0.00±0.02
$K^{*+}_{\pi^0}e^+e^-$	0.04±0.04	0.00±0.01	0.04±0.04	0.00±0.01	0.00±0.02
$K^{*+}_{\pi^0}\mu^+\mu^-$	0.00±0.04	0.00±0.01	0.00±0.04	0.00±0.02	0.00±0.02

Table 7.32: The expected number of $\psi(2S)(K,K^*)$ events in the signal region after applying all of the analysis event selection criteria in Run 1. $K^{*+}_{K_S^0}$ denotes the channel with $K^{*+} \rightarrow K_S^0\pi^+$, and $K^{*+}_{\pi^0}$, with $K^{*+} \rightarrow K^+\pi^0$. The charmonium samples used are signal $\psi(2S)(K,K^*)$ SP3 Monte Carlo. The errors are statistical only. If no events were observed, the error corresponds to one observed event.

which are required to satisfy the loose cuts $|\Delta E| < 0.25$ GeV and $m_{\rm ES} > 5.2$ GeV/c^2, for the charged track modes and modes with a K_S^0 or a π^0, respectively. The main background sources differ somewhat for different modes, although there are similar trends in all of the electron channels and in all of the muon channels. In the electron channels, the bulk of the background comes from semileptonic B decays and charmonium events. In the muon channels, semileptonic decays and fake muons are the major background contributions. The fact that fake leptons remain a significant background for the muon channels is a consequence of the fact that the fake rate is higher for muons than it is for electrons. Figures 7.30 and 7.31 show the residual background components in the continuum sample for the charged track modes and modes with a K_S^0 or a π^0, respectively. They show that for the electron channels, the leptons are often real, whereas for the muon modes, fakes still dominate the

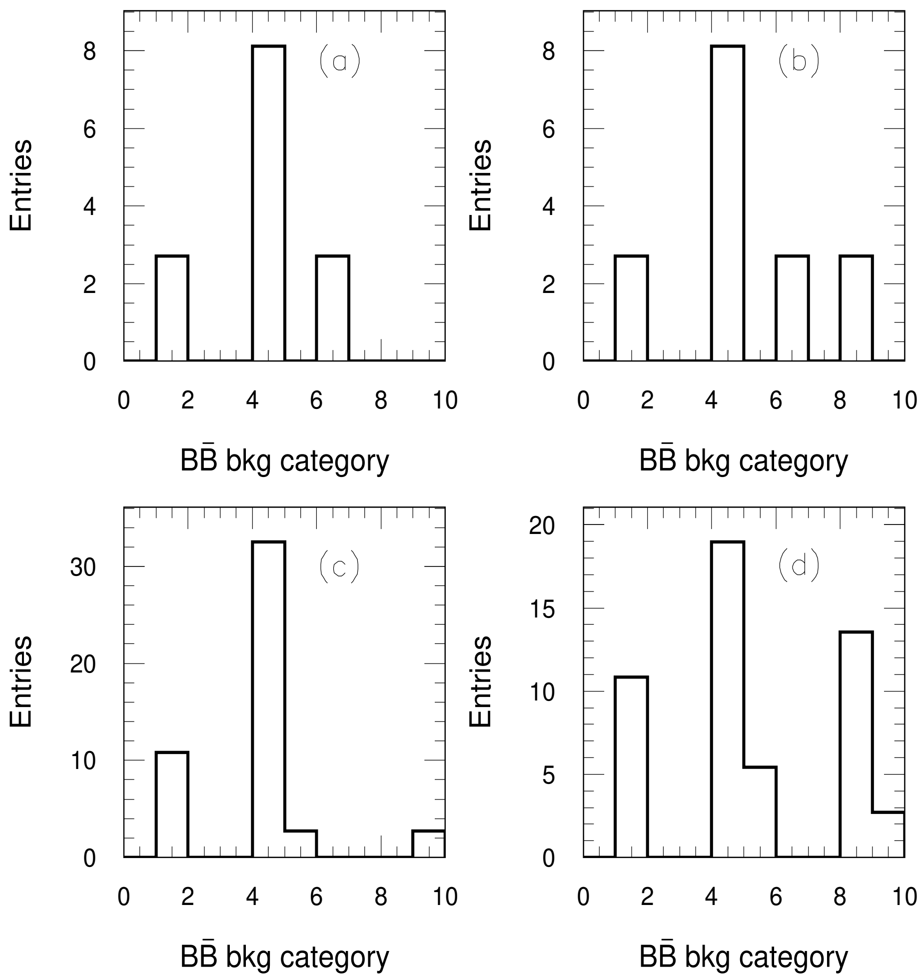


Figure 7.28: $B\overline{B}$ background categories in SP3 generic $B\overline{B}$ sample for the charged track modes. (a) $B^+ \to K^+e^+e^-$, (b) $B^+ \to K^+\mu^+\mu^-$, (c) $B^0 \to K^{*0}e^+e^-$, and (d) $B^0 \to K^{*0}\mu^+\mu^-$. All of the analysis event selection criteria, including loose ΔE and $m_{\rm ES}$ cuts, have been applied ($|\Delta E| < 0.25$ GeV and $m_{\rm ES} > 5.2$ GeV/c^2).

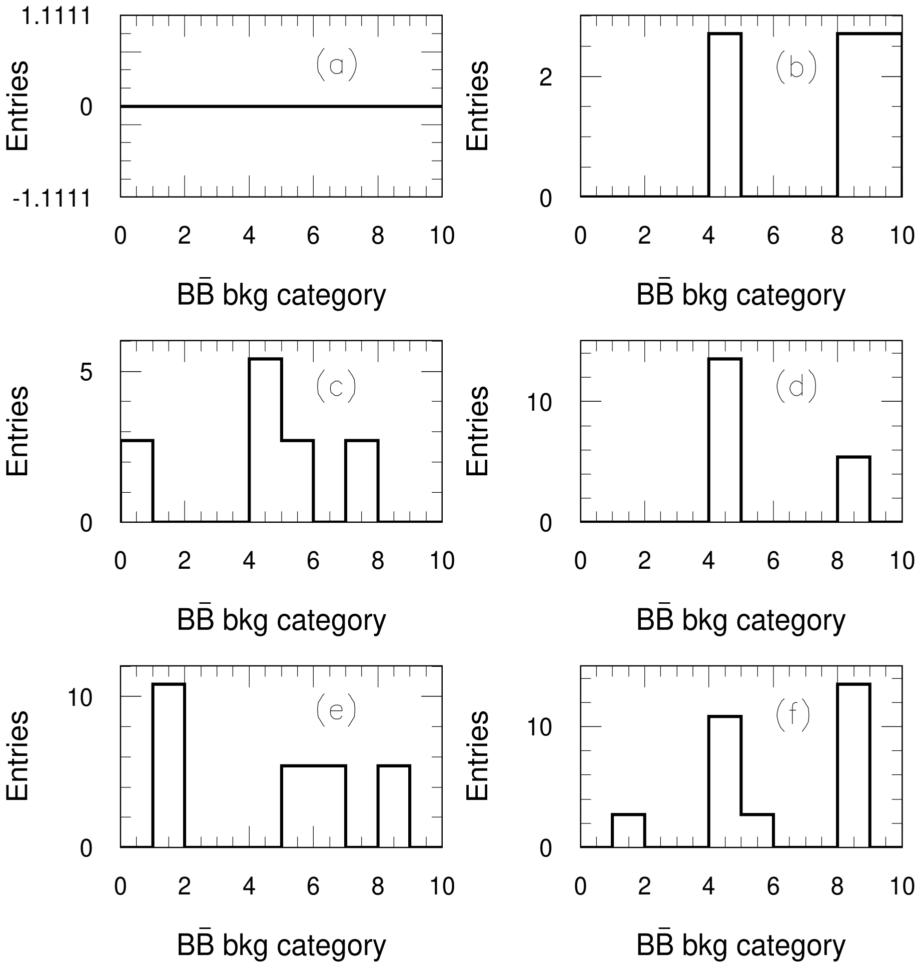


Figure 7.29: $B\overline{B}$ background categories in SP3 generic $B\overline{B}$ sample for the modes with a K_S^0 or a π^0. (a) $B^0 \to K_S^0 e^+e^-$, (b) $B^0 \to K_S^0 \mu^+\mu^-$, (c) $B^+ \to K^{*+}e^+e^-$ with $K^{*+} \to K_S^0\pi^+$, (d) $B^+ \to K^{*+}\mu^+\mu^-$ with $K^{*+} \to K_S^0\pi^+$, (e) $B^+ \to K^{*+}e^+e^-$ with $K^{*+} \to K^+\pi^0$, and (f) $B^+ \to K^{*+}\mu^+\mu^-$ with $K^{*+} \to K^+\pi^0$. All of the analysis event selection criteria, including loose ΔE and m_{ES} cuts, have been applied ($|\Delta E| < 0.25$ GeV and $m_{\text{ES}} > 5.2$ GeV/c^2).

continuum background. It is worth noticing that none of the residual background categories are expected to peak in the signal region.

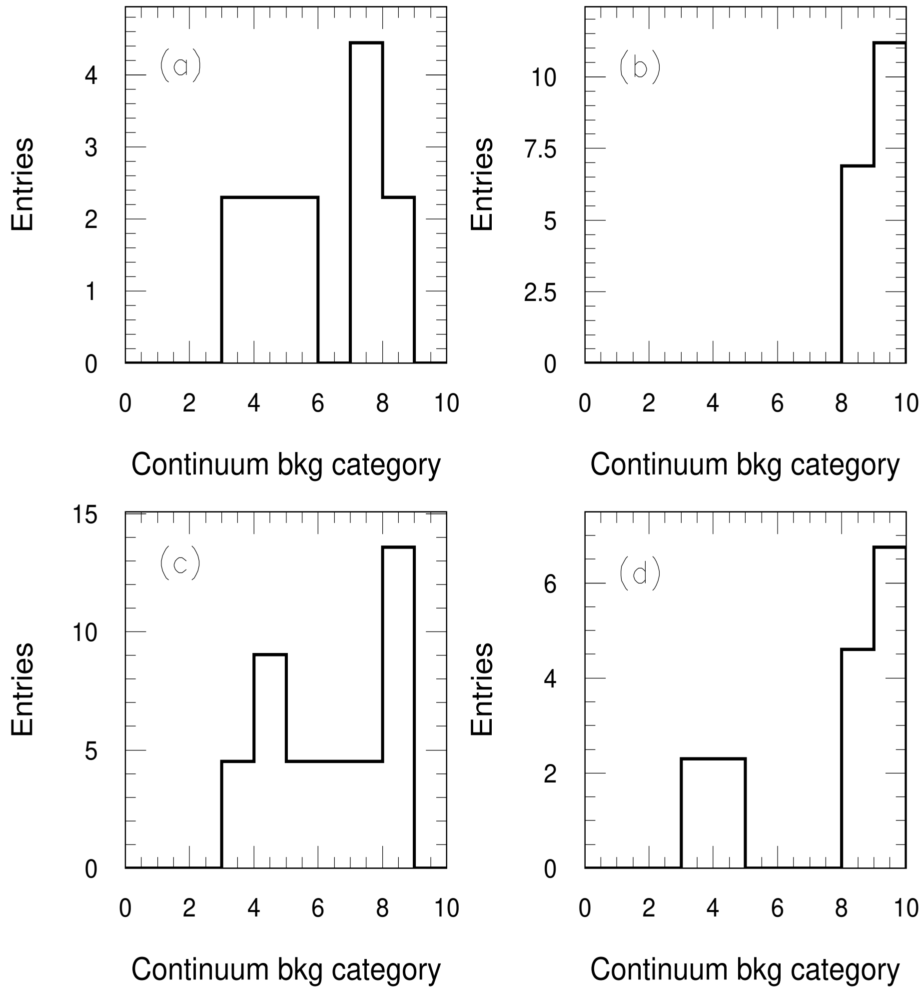


Figure 7.30: Continuum background categories in SP3 $c\bar{c}$ + uds sample for the charged track modes. (a) $B^+ \to K^+e^+e^-$, (b) $B^+ \to K^+\mu^+\mu^-$, (c) $B^0 \to K^{*0}e^+e^-$, and (d) $B^0 \to K^{*0}\mu^+\mu^-$. All of the analysis event selection criteria, including loose ΔE and m_{ES} cuts, have been applied ($|\Delta E| < 0.25$ GeV and $m_{\mathrm{ES}} > 5.2$ GeV/c^2).

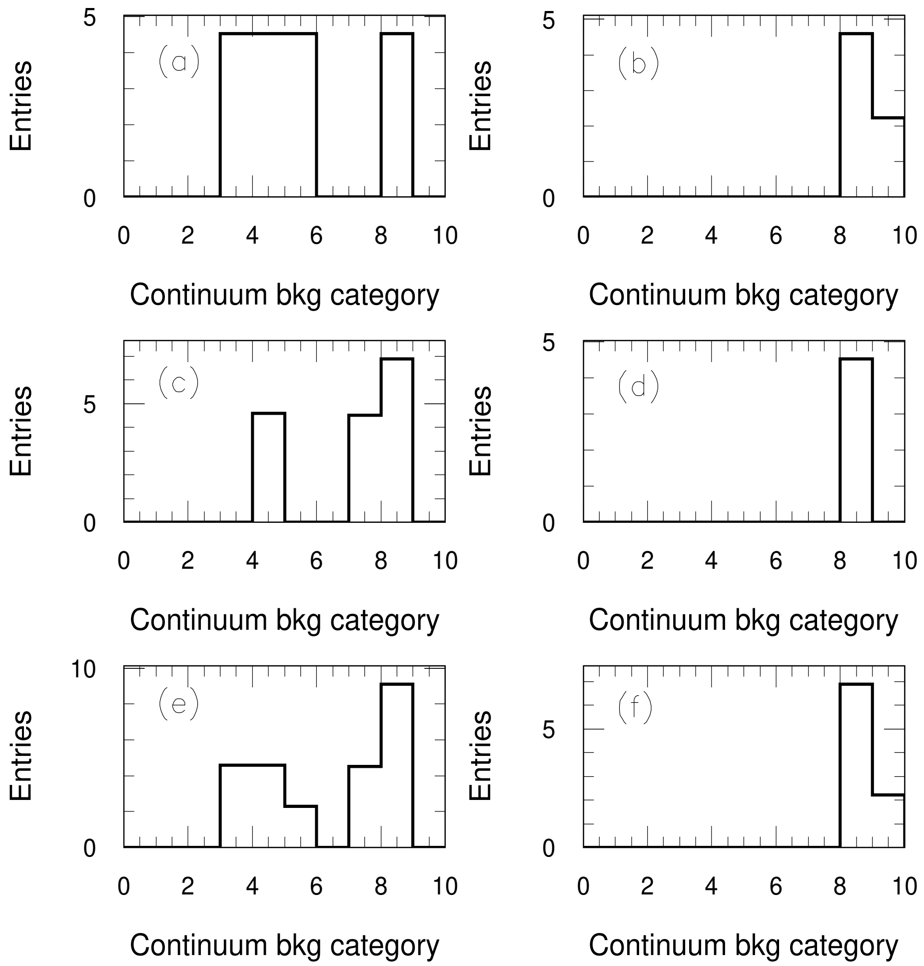


Figure 7.31: Continuum background categories in SP3 $c\bar{c} + uds$ sample for the modes with a K^0_S or a π^0. (a) $B^0 \to K^0_S e^+ e^-$, (b) $B^0 \to K^0_S \mu^+ \mu^-$, (c) $B^+ \to K^{*+} e^+ e^-$ with $K^{*+} \to K^0_S \pi^+$, (d) $B^+ \to K^{*+} \mu^+ \mu^-$ with $K^{*+} \to K^0_S \pi^+$, (e) $B^+ \to K^{*+} e^+ e^-$ with $K^{*+} \to K^+ \pi^0$, and (f) $B^+ \to K^{*+} \mu^+ \mu^-$ with $K^{*+} \to K^+ \pi^0$. All of the analysis event selection criteria, including loose ΔE and $m_{\rm ES}$ cuts, have been applied ($|\Delta E| < 0.25$ GeV and $m_{\rm ES} > 5.2$ GeV$/c^2$).

Mode	Product $\mathcal{B}$	# Prod. evts	# Exp. evts	
			Ke^+e^-	$K\mu^+\mu^-$
$K^+\pi^+\pi^-$ NR	$< 2.8 \times 10^{-5}$	< 640	< 0.005	< 0.06
$\overline{D}^0(\to K^+\pi^-)\pi^+$	2.0×10^{-4}	4600	0.00 ± 0.02	0.05 ± 0.03
			$K^{*0}e^+e^-$	$K^{*0}\mu^+\mu^-$
$K^+\pi^-\pi^+\pi^-$ NR	$< 2.3 \times 10^{-4}$	< 5300	$< 0.00 \pm 0.02$	< 0.09
$D^+(\to K^-\pi^+\pi^-)\pi^-$	2.7×10^{-4}	6200	0.000 ± 0.002	0.20 ± 0.06
$\overline{D}^0(\to K^+\pi^-)\pi^+\pi^-$ NR	$< 6.1 \times 10^{-5}$	< 1400	$< 0.000 \pm 0.006$	$< 0.000 \pm 0.006$
$\overline{D}^0(\to K^+\pi^-)\rho^0$	$< 1.5 \times 10^{-5}$	< 350	$< 0.000 \pm 0.001$	$< 0.006 \pm 0.003$
$K^{*0}(\to K^+\pi^-)\phi$	$< 7 \times 10^{-6}$	< 160	$< 0.000 \pm 0.001$	< 0.001
			$K^0_S e^+e^-$	$K^0_S\mu^+\mu^-$
$K^{*+}(\to K^0_S\pi^+)\pi^+$	$< 1.6 \times 10^{-5}$	< 370	$< 0.000 \pm 0.002$	< 0.02
$D^-(\to K^0_S\pi^-)\pi^+$	5.1×10^{-5}	1200	0.00 ± 0.01	0.10 ± 0.02
$K^0_S\phi(\to K^+K^-)$	$< 1 \times 10^{-5}$	< 230	$< 0.000 \pm 0.001$	$< 0.000 \pm 0.001$
			$K^{*+}_{K^0_S}e^+e^-$	$K^{*+}_{K^0_S}\mu^+\mu^-$
$K^{*+}_{K^0_S}\rho^0$	$< 2 \times 10^{-4}$	< 4600	< 0.02	< 0.2
$\overline{D}^0(\to K^0_S\pi^+\pi^-)\pi^+$	8.8×10^{-5}	2000	0.0	0.01
			$K^{*+}_{\pi^0}e^+e^-$	$K^{*+}_{\pi^0}\mu^+\mu^-$
$K^{*+}_{\pi^0}\pi^+\pi^-$ NR	$< 3.7 \times 10^{-4}$	< 8500	$< 0.00 \pm 0.03$	< 0.3
$\overline{D}^0(\to K^+\pi^-\pi^0)\pi^+$	7.4×10^{-4}	1.7×10^4	0.0 ± 0.1	0.07 ± 0.07
$\overline{D}^{*0}(\to D^0(\to K^+\pi^-)\pi^0)\pi^+$	1.1×10^{-4}	< 2500	0.04	0.07 ± 0.03
$K^{*+}_{\pi^0}\phi$	$< 6.8 \times 10^{-6}$	< 160	$< 0.000 \pm 0.001$	$< 0.000 \pm 0.001$

Table 7.33: The expected number of hadronic events in the signal region in the Run 1 data after applying all of the analysis event selection criteria. $K^{*+}_{K^0_S}$ denotes the channel with $K^{*+} \to K^0_S\pi^+$, and $K^{*+}_{\pi^0}$, with $K^{*+} \to K^+\pi^0$. The product branching fractions $\mathcal{B}$ for the modes not yet observed are listed as limits. The errors are statistical only. If no events were observed, the error corresponds to one observed event.

Category #	Description
0	unknown
1	both leptons are from J/ψ
2	both leptons are from $\psi(2S)$
3	both leptons are from γ or π^0
4	both leptons are from primary semileptonic B decays
5	one lepton is from primary, other from secondary semileptonic decay of the same B
6	one lepton is from primary B semileptonic decay , other from secondary semileptonic decay of the other B
7	both leptons are from secondary semileptonic B decays
8	one lepton is fake
9	both leptons are fake

Table 7.34: The $B\overline{B}$ background categories used to study the effect of various event selection criteria on background suppression.

Category #	Description
0	unknown
1	both leptons are from primary D semileptonic decays
2	one lepton is from primary, other from secondary semileptonic decay of the same D
3	one lepton is from primary D semileptonic decay , other from secondary semileptonic decay of a different D
4	both leptons are from secondary D decays
5	one lepton comes from a γ conversion
6	one lepton comes from a π^0 Dalitz decay
7	both leptons come from a γ or a π^0 decay
8	one lepton is fake
9	both leptons are fake

Table 7.35: The continuum background categories used to study the effect of various event selection criteria on background suppression.

Chapter 8

Control Samples

This analysis is performed as a blind search, and thus the signal and the nearest sideband regions are hidden in the data. Only Monte Carlo and off-resonance data are used for optimizing the event selection criteria and estimating the expected number of signal and background events. In order to verify the efficiencies of the selection requirements and the shapes of various kinematic distributions, several control samples in the on-resonance data sample can be used without any risk of biasing the analysis. A control sample is a data subset excluded from the final measurement, used for comparing data and Monte Carlo for quality control and sanity checks. The samples used in this analysis include three-body hadronic decays $B^- \to D^0\pi^-$ with $D^0 \to K^-\pi^+$; charmonium decays $B \to J/\psi(K, K^*)$ and $B \to \psi(2S)(K, K^*)$; the $(K, K^*)e\mu$ combinatorial sample; and the "large sideband".

8.1 $B^- \to D^0\pi^-$, $D^0 \to K^-\pi^+$ Decays

The decays $B^- \to D^0\pi^-$ with $D^0 \to K^-\pi^+$ (the charged-conjugate channel is always implied) are 3-body decays with no leptons and a kaon. They are thus useful for checking the efficiency of the kaon identification requirements. They are also useful for measuring the resolution of ΔE and $m_{\rm ES}$, which can be compared with Monte Carlo predictions.

The event selection criteria used to extract a clean sample of $B^- \to D^0\pi^-$, $D^0 \to K^-\pi^+$ events are summarized in Table 8.1. The first five selection criteria in Table 8.1 are applied at the skim level. The BGFMultiHadron tag and R2All are defined in Chapter 7. The BchToD0Pi tag selects candidate events using the following selection criteria:

- A D^0 candidate is reconstructed in one of three modes: $D^0 \to K^+\pi^-$, $D^0 \to K^+\pi^-\pi^0$, $D^0 \to K^+\pi^-\pi^+\pi^-$. All charged tracks are required to be GoodTracksLoose, but there are no particle identification requirements applied to either kaon or pion candidates. The invariant mass of the D^0 candidate, formed using four-momentum addition, is required to be within $\pm$ 40 MeV/c^2 of the mean D^0 mass for the two charged track modes, and within $\pm$ 70 MeV/c^2 of the mean D^0 mass for the mode with a π^0.

- The momentum of the candidate D^0 is required to be greater than 1.3 MeV/c in the center-of-mass frame.

- The invariant mass of the three tracks, computed using four-momentum addition, is required to be within 5.0 and 5.5 GeV/c^2

- The reconstructed energy of the D^0 candidate should be within $\pm$ 300 MeV of one-half the center of mass energy.

The remaining selection criteria in Table 8.1 are applied at the ntuple level. The requirement that m_{D^0} be within $\pm$ 0.0218 MeV/c^2 of $m_{D^0}^{\mathrm{PDG}}$ is a 3.5σ cut derived from signal Monte Carlo. The last two requirements are 3.5σ ΔE and m_{ES} cuts, also derived from signal Monte Carlo.

Figures 8.1a and 8.1b show the comparison of the ΔE and m_{ES} distributions, respectively, in this control sample in the on-resonance data vs. SP3 $B\overline{B}$ Monte Carlo plus off-resonance data. It is evident that the ΔE distribution in the on-resonance data is broader and shifted to the left. The fact that the data does

Variable	Value
BGFMultiHadron tag	True
R2All	≤ 0.9
Number of charged tracks	≥ 3
L3DchOut = True or L3EmcOut = True	
BchToD0Pi tag	True
All three tracks = GoodTracksLoose	
$\lvert m_{D^0} - m_{D^0}^{\mathrm{PDG}}\rvert$	< 0.0218 MeV/c^2
$\lvert \Delta E\rvert$	< 0.063 MeV
$\lvert m_{\mathrm{ES}} - m_{\mathrm{ES}}^0\rvert$	< 0.00696 MeV/c^2

Table 8.1: Event selection criteria used to select $B^- \to D^0\pi^-$, $D^0 \to K^-\pi^+$. The first five requirements are skim requirements; the rest are ntuple requirements.

not agree with the Monte Carlo is an illustration of some of the problems with SP3 Monte Carlo that were mentioned in Section 5.3. The broadening is a consequence of the fact that the momentum resolution is worse in the data than in the Monte Carlo, probably because of the presence of material not simulated in SP3. The shift to the left is not understood, but its possible causes include the uncertainty on the boost, the imperfection of alignment, the presence of extra material not simulated in the Monte Carlo, and **B**-field effects.

In order to better match the data, the Monte Carlo was adjusted using a so-called smearing technique. This involves increasing the transverse momentum resolution p_t for all charged tracks by 48%, and introducing a bias of 0.16% on the central value of p_t. The smearing parameters are *ad hoc* corrections, derived from other data samples. The effect of this smearing on ΔE and m_{ES} is shown in Fig. 8.1c and 8.1d, respectively. The agreement clearly improves. The comparison of the ΔE and m_{ES} shapes in Fig. 8.1 also shows that m_{ES} is much less sensitive to the smearing effect than ΔE. This is explained by the fact that the resolution of ΔE is dominated by the momentum resolution, whereas the m_{ES} resolution is determined by the beam energy spread [71].

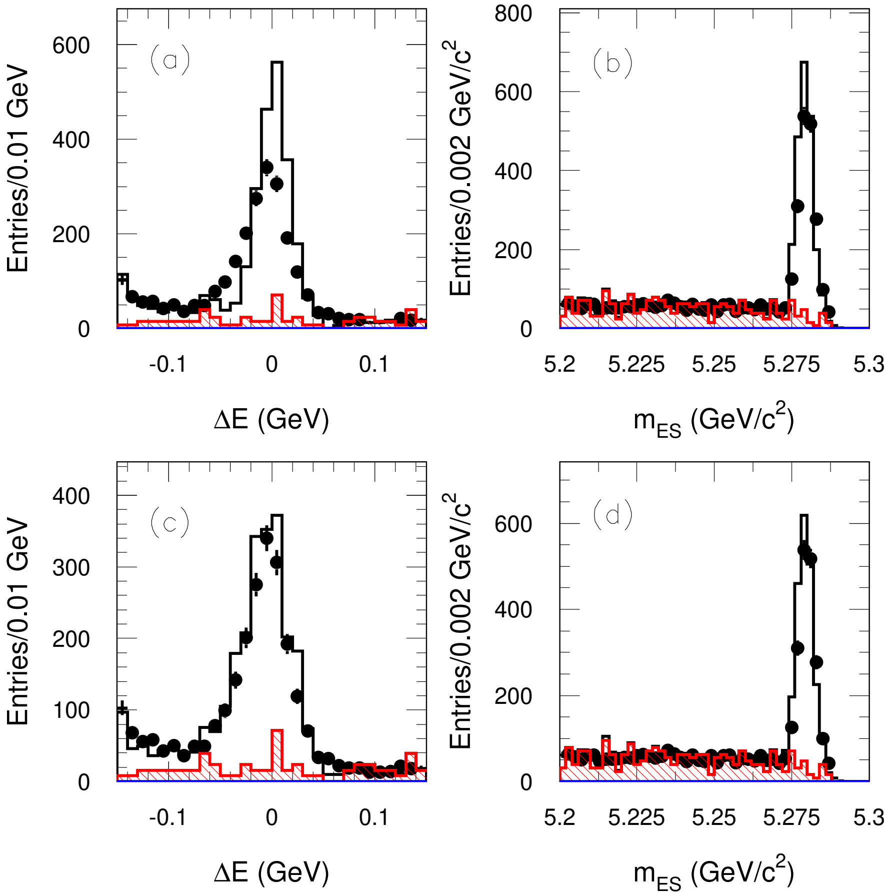


Figure 8.1: Comparison of ΔE and m_{ES} shapes in the $B^- \to D^0\pi^-$, $D^0 \to K^-\pi^+$ control sample. The dots are on-resonance data, the hatched histogram is off-resonance data, and the solid line is the sum of off-resonance data and $B\overline{B}$ Monte Carlo. (a) and (b) ΔE and m_{ES} distributions, respectively, after applying all of the event selection criteria in Table 8.1 plus a very tight kaon identification requirement. No smearing of the Monte Carlo has been applied. (c) and (d) The same after the Monte Carlo has been smeared and shifted, as described in the text. The normalization is absolute.

8.2 Charmonium Control Sample

Charmonium decays, $B \to J/\psi(K, K^*)$ and $B \to \psi(2S)(K, K^*)$, are essentially signal modes for a particular value of $q^2 \equiv m_{\ell^+\ell^-} = m_{J/\psi}$ and $m_{\psi(2S)}$. A sample of charmonium events can therefore be used to verify the efficiencies of almost all of the analysis selection criteria. It is also useful for checking that the distributions of various quantities of interest are adequately reproduced in the Monte Carlo. This section discusses the comparison between the charmonium control sample in the data and SP3 signal $B \to J/\psi(K, K^*)$ and $B \to \psi(2S)(K, K^*)$ Monte Carlo. The charmonium Monte Carlo samples used in the comparison are described in Section 5.2.

Several steps have been taken in order to account for the known differences between data and Monte Carlo:

- For all charged tracks, the transverse momentum resolution and its central value are adjusted in the Monte Carlo using the smearing technique described in Section 8.1. The central value is biased by 0.16% in all of the modes. The factors by which the resolution on p_t is increased are listed in Table 8.2. These factors are exactly the same as those that *BABAR* used in its measurement of the charmonium branching fractions [66], since in the comparison between data and Monte Carlo the *BABAR* results for the charmonium decays are used wherever available. Table 8.2 shows that the p_t resolution smearing is not the same for electron and muon channels. It is larger for the electron modes to account for the effects of radiation. For the muon modes, the smearing applied is also different for the Monte Carlo simulating the 1900 V data and the 1960 V data.

- Lepton and kaon particle identification "killing tables" are used to match particle identification efficiencies in data and Monte Carlo. The tables are described in Section 5.3. Different tables are used for the Monte Carlo simulating the

	1900 V		1960 V	
Mode	p_t smearing	PID corr	p_t smearing	PID corr
Electrons	1.48	0.988	1.48	0.920
Muons	1.32	0.996	1.14	0.950

Table 8.2: Smearing and particle identification (PID) correction factors used in the SP3 charmonium signal Monte Carlo for the two different DCH high voltage settings.

two blocks of data with different drift chamber voltages. A small residual particle identification correction is applied to account for the differences in the low multiplicity environment of most particle identification control samples and the relatively high multiplicity environment of charmonium $B\overline{B}$ decays. This correction is computed in Ref. [67], and is listed in Table 8.2.

- B candidates are weighted by the product of the data/Monte Carlo relative tracking efficiencies of their constituent GoodTracksLoose. The relative efficiencies are obtained using tracking efficiency tables described in Section 5.3. Different tables are used for the 1900 V and 1960 V data.

- Modes with K^0_S mesons have been corrected for the difference in the K^0_S reconstruction efficiency in data and in the Monte Carlo, as described in Ref. [66]. In particular, the efficiency for finding K^0_S mesons is 5% (6%) higher in the data for the electron (muon) modes.

The Monte Carlo, with these adjustments, is compared with the data. Figures 8.2 and 8.3 compare the ΔE shapes for the charged track modes and the modes with a K^0_S or a π^0, respectively. The event selection criteria applied are the reversed charmonium vetoes and all of the other analysis selection requirements including the signal cut in ΔE and $m_{\rm ES}$. The Monte Carlo appears to be modeling the shapes reasonably well. In fact, it may be over-correcting the shift to the left in ΔE for the $B^+ \to K^+\ell^+\ell^-$ modes, although for the K^* modes the shift may need to be slightly increased to better match the data. Figures 8.4 and 8.5 compare the $m_{\rm ES}$ shapes for the charged track modes and the modes with a K_S^0 or a π^0, respectively. These plots show that the data distributions are slightly but systematically shifted to the right.

The yield of events in the data and in the Monte Carlo Table 8.3. They agree within statistics in all of the modes.

Mode	Data Yield	MC Yield
$J/\psi(\to e^+e^-)K^+$	236 ± 15	228.0 ± 2.5
$J/\psi(\to e^+e^-)K_S^0$	57 ± 8	68.9 ± 0.9
$J/\psi(\to e^+e^-)K^{*0}$	112 ± 11	105.5 ± 1.6
$J/\psi(\to e^+e^-)K^{*+}$ $(K^{*+} \to K_S^0\pi^+)$	31 ± 6	40.1 ± 0.6
$J/\psi(\to e^+e^-)K^{*+}$ $(K^{*+} \to K^+\pi^0)$	35 ± 6	36.8 ± 0.7
$J/\psi(\to \mu^+\mu^-)K^+$	189 ± 14	187.2 ± 2.3
$J/\psi(\to \mu^+\mu^-)K_S^0$	56 ± 7	58.8 ± 0.9
$J/\psi(\to \mu^+\mu^-)K^{*0}$	116 ± 11	108.1 ± 1.7
$J/\psi(\to \mu^+\mu^-)K^{*+}$ $(K^{*+} \to K_S^0\pi^+)$	32 ± 6	30.2 ± 0.6
$J/\psi(\to \mu^+\mu^-)K^{*+}$ $(K^{*+} \to K^+\pi^0)$	25 ± 5	27.4 ± 0.6

Table 8.3: The yields of the charmonium events in data and signal $B \to J/\psi(K, K^*)$ SP3 Monte Carlo after applying all of the analysis event selection criteria. The errors are statistical only. The normalization is absolute.

The charmonium control sample can also be used to perform checks of other variables of interest that are used in the analysis. For example, the Fisher discriminants are compared in Figs. 8.6 and 8.7 for the charged track modes and the modes with a K_S^0 or a π^0, respectively. The comparison of the $B\overline{B}$ suppression variables is shown on Figs. 8.8 and 8.9 for the charged track modes and the modes with a K_S^0 or a π^0, respectively. None of these plots indicates any serious discrepancies between data and Monte Carlo.

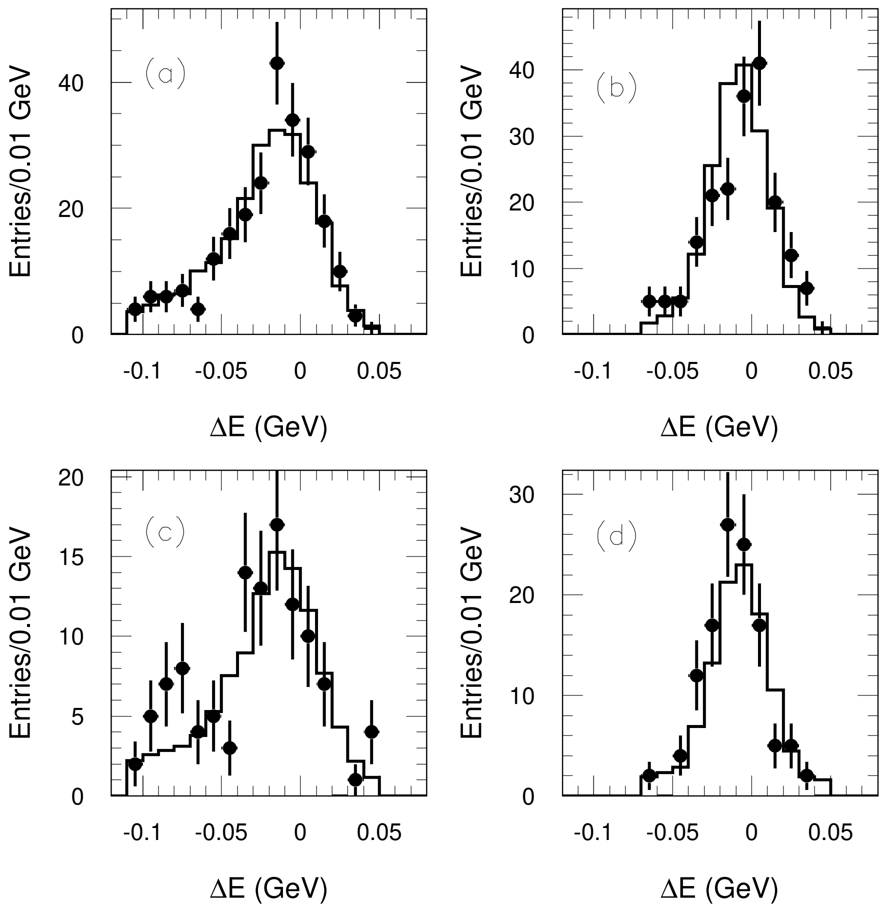


Figure 8.2: Comparison of ΔE shapes in the charmonium control sample for the charged track modes. (a) $B^+ \to J/\psi(\to e^+e^-)K^+$; (b) $B^+ \to J/\psi(\to \mu^+\mu^-)K^+$; (c) $B^0 \to J/\psi(\to e^+e^-)K^{*0}$; (d) $B^0 \to J/\psi(\to \mu^+\mu^-)K^{*0}$. The dots are the on-resonance data, and the solid histogram is the signal charmonium SP3 Monte Carlo. All of the event selection criteria have been applied (with the reversed charmonium veto), including the signal ΔE and m_{ES} cut. The normalization is absolute.

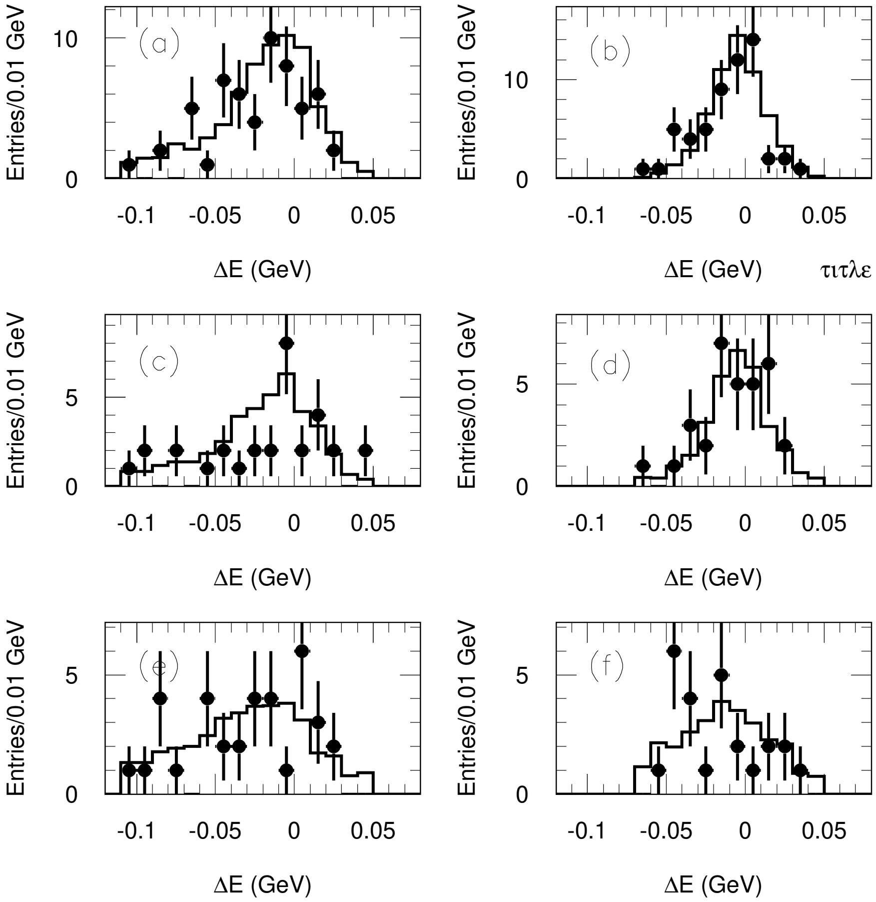


Figure 8.3: Comparison of ΔE shapes in the charmonium control sample for the modes with a K^0_S or a π^0. (a) $B^0 \to J/\psi(\to e^+e^-)K^0_S$; (b) $B^0 \to J/\psi(\to \mu^+\mu^-)K^0_S$; (c) $B^+ \to J/\psi(\to e^+e^-)K^{*+}$ with $K^{*+} \to K^0_S\pi^+$; (d) $B^+ \to J/\psi(\to \mu^+\mu^-)K^{*+}$ with $K^{*+} \to K^0_S\pi^+$; (e) $B^+ \to J/\psi(\to e^+e^-)K^{*+}$ with $K^{*+} \to K^+\pi^0$; (f) $B^+ \to J/\psi(\to \mu^+\mu^-)K^{*+}$ with $K^{*+} \to K^+\pi^0$. The dots are the on-resonance data, and the solid histogram is the signal charmonium SP3 Monte Carlo. All of the event selection criteria have been applied (with the reversed charmonium veto), including the signal ΔE and $m_{\rm ES}$ cut. The normalization is absolute.

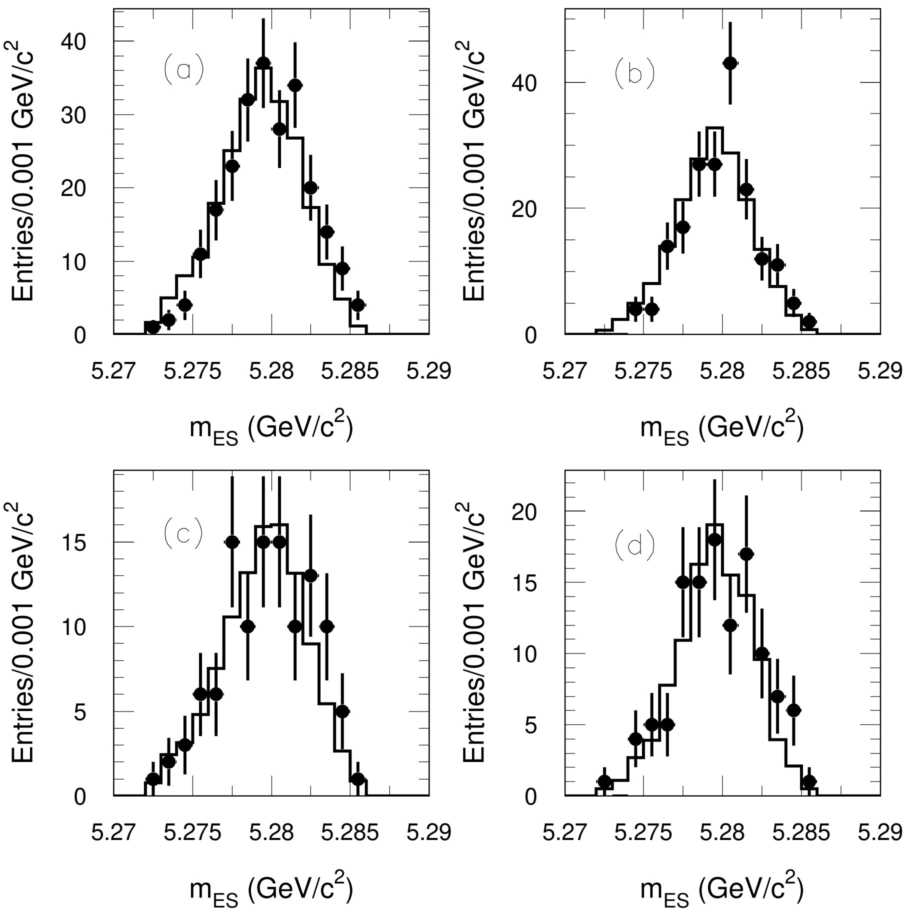


Figure 8.4: Comparison of m_{ES} shapes in the charmonium control sample for the charged track modes. (a) $B^+ \to J/\psi(\to e^+e^-)K^+$; (b) $B^+ \to J/\psi(\to \mu^+\mu^-)K^+$; (c) $B^0 \to J/\psi(\to e^+e^-)K^{*0}$; (d) $B^0 \to J/\psi(\to \mu^+\mu^-)K^{*0}$. The dots are the on-resonance data, and the solid histogram is the signal charmonium SP3 Monte Carlo. All of the event selection criteria have been applied (with the reversed charmonium veto), including the signal ΔE and m_{ES} cut. The normalization is absolute.

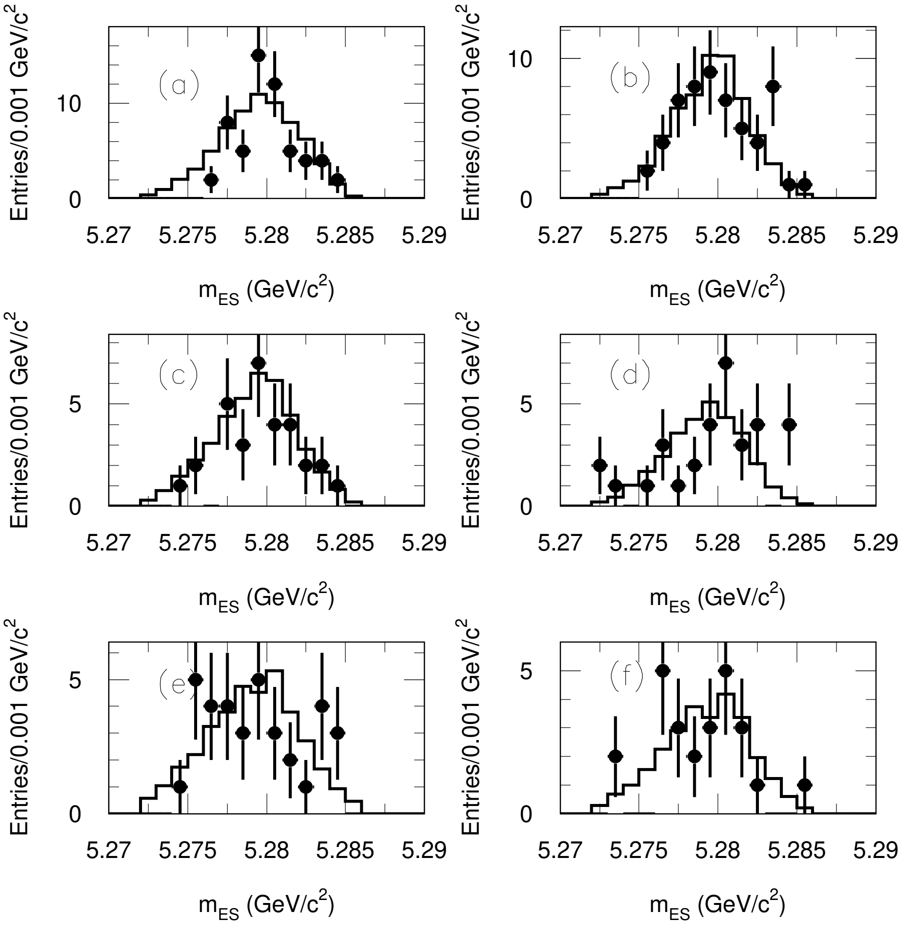


Figure 8.5: Comparison of m_{ES} shapes in the charmonium control sample for the modes with a K^0_S or a π^0. (a) $B^0 \to J/\psi(\to e^+e^-)K^0_S$; (b) $B^0 \to J/\psi(\to \mu^+\mu^-)K^0_S$; (c) $B^+ \to J/\psi(\to e^+e^-)K^{*+}$ with $K^{*+} \to K^0_S\pi^+$; (d) $B^+ \to J/\psi(\to \mu^+\mu^-)K^{*+}$ with $K^{*+} \to K^0_S\pi^+$; (e) $B^+ \to J/\psi(\to e^+e^-)K^{*+}$ with $K^{*+} \to K^+\pi^0$; (f) $B^+ \to J/\psi(\to \mu^+\mu^-)K^{*+}$ with $K^{*+} \to K^+\pi^0$. The dots are the on-resonance data, and the solid histogram is the signal charmonium SP3 Monte Carlo. All of the event selection criteria have been applied (with the reversed charmonium veto), including the signal ΔE and m_{ES} cut. The normalization is absolute.

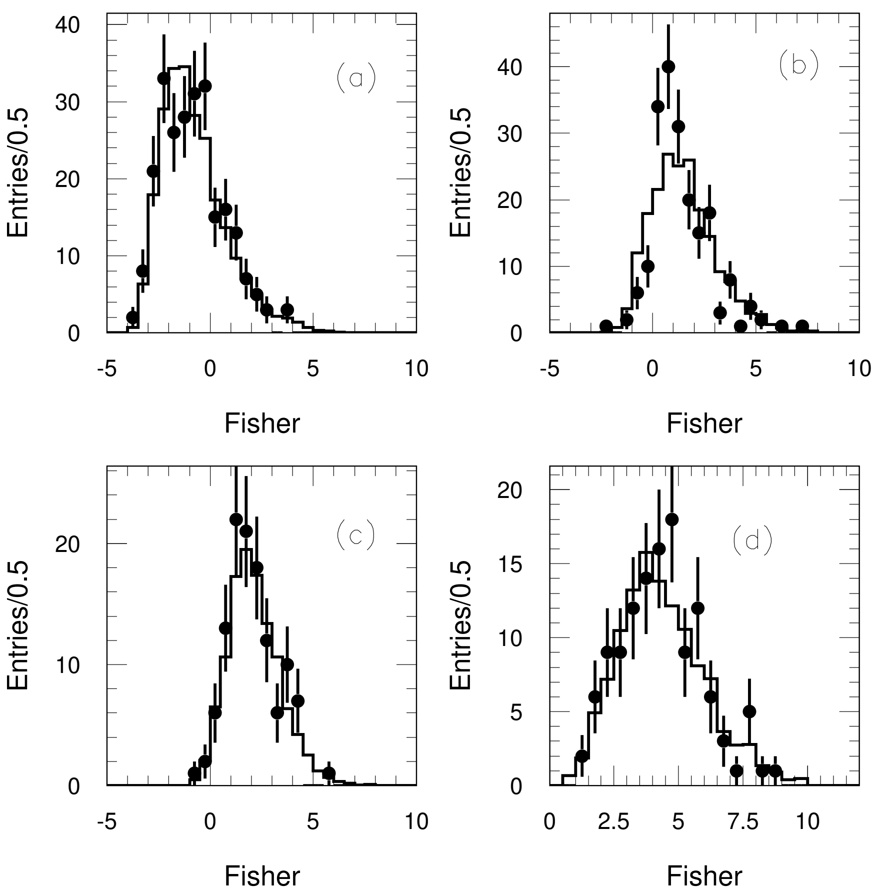


Figure 8.6: Comparison of Fisher discriminant shapes in the charmonium control sample for the charged track modes. (a) $B^+ \to J/\psi(\to e^+e^-)K^+$; (b) $B^+ \to J/\psi(\to \mu^+\mu^-)K^+$; (c) $B^0 \to J/\psi(\to e^+e^-)K^{*0}$; (d) $B^0 \to J/\psi(\to \mu^+\mu^-)K^{*0}$. The dots are the on-resonance data, and the solid histogram is the signal charmonium SP3 Monte Carlo. All of the event selection criteria have been applied (with the reversed charmonium veto), including the signal ΔE and $m_{\rm ES}$ cut, except those on the Fisher discriminants themselves. The normalization is absolute.

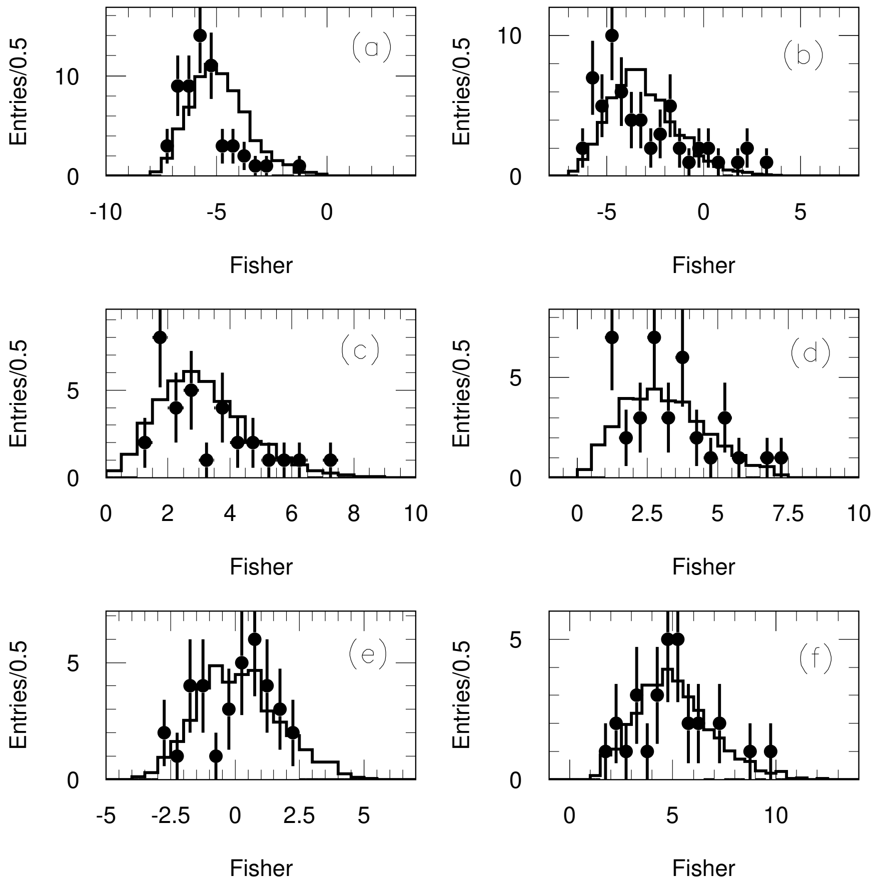


Figure 8.7: Comparison of Fisher discriminant shapes in the charmonium control sample for the modes with a K_S^0 or a π^0. (a) $B^0 \to J/\psi(\to e^+e^-)K_S^0$; (b) $B^0 \to J/\psi(\to \mu^+\mu^-)K_S^0$; (c) $B^+ \to J/\psi(\to e^+e^-)K^{*+}$ with $K^{*+} \to K_S^0\pi^+$; (d) $B^+ \to J/\psi(\to \mu^+\mu^-)K^{*+}$ with $K^{*+} \to K_S^0\pi^+$; (e) $B^+ \to J/\psi(\to e^+e^-)K^{*+}$ with $K^{*+} \to K^+\pi^0$; (f) $B^+ \to J/\psi(\to \mu^+\mu^-)K^{*+}$ with $K^{*+} \to K^+\pi^0$. The dots are the on-resonance data, and the solid histogram is the signal charmonium SP3 Monte Carlo. All of the event selection criteria have been applied (with the reversed charmonium veto), including the signal ΔE and $m_{\rm ES}$ cut, except those on the Fisher discriminants themselves. The normalization is absolute.

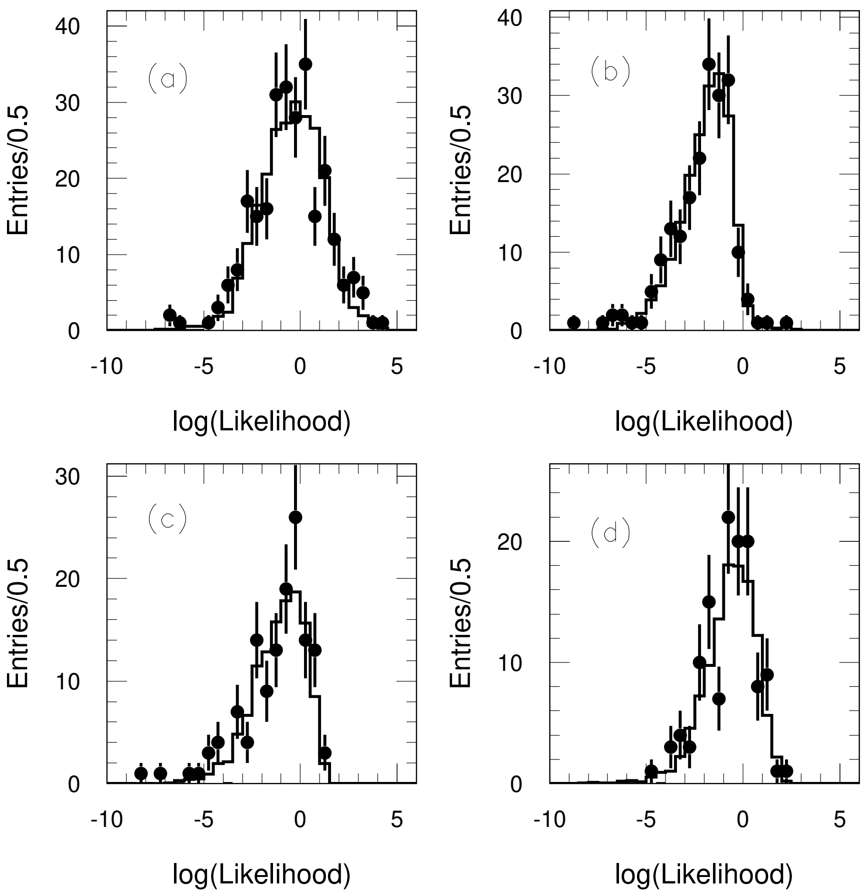


Figure 8.8: Comparison of the $B\overline{B}$ suppression variable shapes in the charmonium control sample for the charged track modes. Plotted are the logs of the likelihoods. (a) $B^+ \to J/\psi(\to e^+e^-)K^+$; (b) $B^+ \to J/\psi(\to \mu^+\mu^-)K^+$; (c) $B^0 \to J/\psi(\to e^+e^-)K^{*0}$; (d) $B^0 \to J/\psi(\to \mu^+\mu^-)K^{*0}$. The dots are the on-resonance data, and the solid histogram is the signal charmonium SP3 Monte Carlo. All of the event selection criteria have been applied (with the reversed charmonium veto), including the signal ΔE and $m_{\rm ES}$ cut, except those on the $B\overline{B}$ suppression variables. The normalization is absolute.

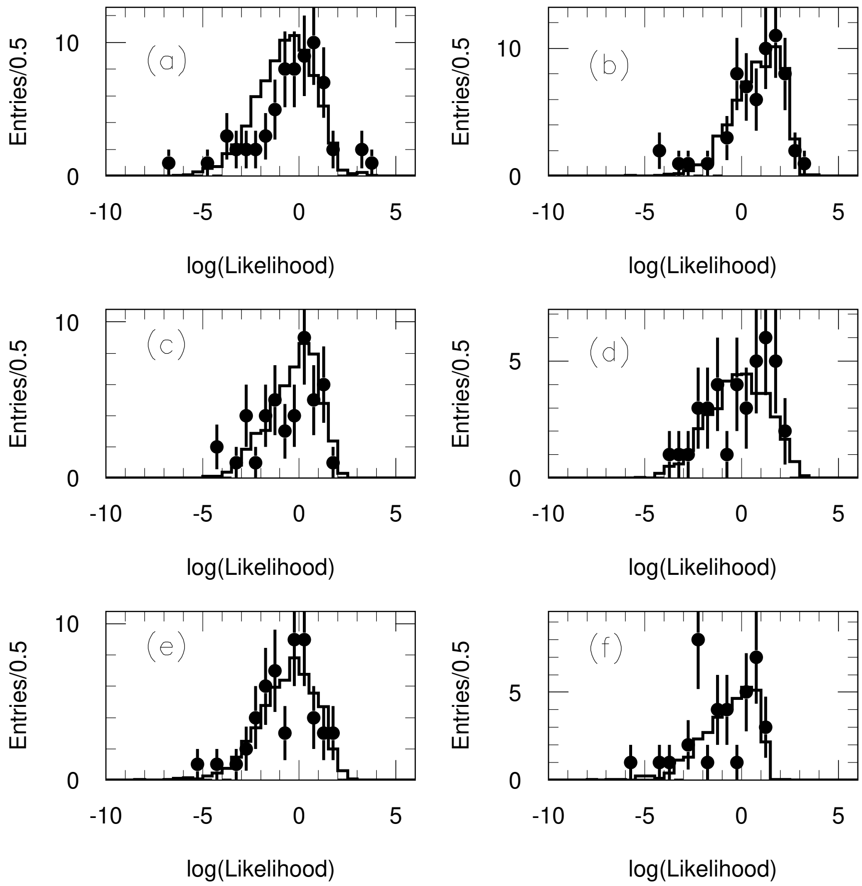


Figure 8.9: Comparison of the $B\overline{B}$ suppression variable shapes in the charmonium control samples for the modes with a K^0_S or a π^0. Plotted are the logs of the likelihoods. (a) $B^0 \to J/\psi(\to e^+e^-)K^0_S$; (b) $B^0 \to J/\psi(\to \mu^+\mu^-)K^0_S$; (c) $B^+ \to J/\psi(\to e^+e^-)K^{*+}$ with $K^{*+} \to K^0_S\pi^+$; (d) $B^+ \to J/\psi(\to \mu^+\mu^-)K^{*+}$ with $K^{*+} \to K^0_S\pi^+$; (e) $B^+ \to J/\psi(\to e^+e^-)K^{*+}$ with $K^{*+} \to K^+\pi^0$; (f) $B^+ \to J/\psi(\to \mu^+\mu^-)K^{*+}$ with $K^{*+} \to K^+\pi^0$. The dots are the on-resonance data, and the solid histogram is the signal charmonium SP3 Monte Carlo. All of the event selection criteria have been applied (with the reversed charmonium veto), including the signal ΔE and m_{ES} cut, except those on the $B\overline{B}$ suppression variables. The normalization is absolute.

8.3 $(K, K^*)e\mu$ Combinatorial Control Sample

A significant fraction of the total background is formed by random leptons that are not produced in a single B decay. This background can be studied using $Ke\mu$ and $K^*e\mu$ events in the data. In this control sample, the comparison between data and Monte Carlo is performed after applying tracking, particle identification, the Fisher discriminant, and the $B\overline{B}$ suppression event selection criteria. The Fisher and $B\overline{B}$ suppression variables used are the same as those for the corresponding electron channels.

The yields of events in the data and in the Monte Carlo are listed in Table 8.4. They are consistent within statistical errors. The worst level of agreement in the event yield between the on-resonance data sample and the SP3 $B\overline{B}$ Monte Carlo plus off-resonance data sample is at a 1.9σ level in the $K^{*+}e\mu$ ($K^{*+} \to K^0_S\pi^+$) mode. The comparison of the on-resonance data and Pravda Monte Carlo event yields gives the worst results for the $Ke\mu$ mode, where it is at a 3.3σ level.

Figures 8.10 and 8.11 show the comparison of ΔE and $m_{\rm ES}$ shapes for the charged track modes and the modes with a K^0_S or a π^0, respectively. Because there is not enough statistics to show these distributions after applying the signal

Mode	Data	SP3 + off-res	Pravda MC
$Ke\mu$	112.0 ± 10.6	165.2 ± 31.4	75.0 ± 2.8
$K^{*0}e\mu$	150.0 ± 12.2	176.8 ± 29.2	141.3 ± 3.8
$K^0_Se\mu$	36.0 ± 6.0	66.2 ± 21.6	41.4 ± 2.1
$K^{*+}e\mu$ ($K^{*+} \to K^0_S\pi^+$)	74.0 ± 8.6	43.3 ± 13.4	84.6 ± 3.0
$K^{*+}e\mu$ ($K^{*+} \to K^+\pi^0$)	121.0 ± 11.0	112.3 ± 24.3	96.1 ± 3.2

Table 8.4: $(K, K^*)e\mu$ yields in data and Monte Carlo. The Monte Carlo has been normalized to the data luminosity. The event selection criteria applied include tracking, particle ID, the Fisher and $B\overline{B}$ suppression cuts with the same values as those used for the corresponding di-electron channels, and $|\Delta E| < 0.25$ GeV, $m_{\rm ES} > 5.2$ GeV/c^2. The errors are statistical only. If no events were observed in the Monte Carlo samples, an error corresponding to one observed event is assigned.

cuts on ΔE and m_{ES}, the variables are selected in the range $|\Delta E| < 0.25$ GeV and $m_{\mathrm{ES}} > 5.2$ GeV/c^2 (the fit region, discussed further in Chapter 9). It is apparent that the statistics is still not sufficient to make a good comparison, but there appear to be no major discrepancies in the shapes of the distributions. Pravda Monte Carlo provides a larger statistics simulation sample. The comparisons of ΔE and m_{ES} shapes in data and Pravda Monte Carlo are shown in Figs. 8.12 and 8.13 for the charged track modes and the modes with a K^0_S or a π^0, respectively. The agreement is good.

8.4 Large Sideband

A region of the grand sideband referred to as "large sideband" is selected by requiring $|\Delta E| > 0.5$ GeV or $m_{\mathrm{ES}} < 5.1$ GeV/c^2. It is shown as region C in Fig. 7.1. While regions A (signal) and B (nearest background) are hidden in the data, region C is a control sample that is used for checking the understanding of the combinatorial background.

Figures 8.14 and 8.15 show the comparison of the ΔE shapes in data and SP3 Monte Carlo for the charged track modes and the modes with a K^0_S or a π^0, respectively. All of the analysis selection criteria have been applied, with ΔE and m_{ES} selected as specified above. Figures 8.16 and 8.17 compare the ΔE shapes in data and Pravda Monte Carlo for the charged track modes and the modes with a K^0_S or a π^0, respectively. Figures 8.18 and 8.19 compare the m_{ES} distributions in data and SP3 Monte Carlo for the charged track modes and the modes with a K^0_S or a π^0, respectively. Similarly, Figs. 8.20 and 8.21 show the comparison of the m_{ES} shapes in data and Pravda Monte Carlo for the charged track modes and the modes with a K^0_S or a π^0, respectively.

Table 8.5 summarizes the event yields. In general, the agreement between data and SP3 Monte Carlo is good. The worst disagreement in the event yield between the on-resonance data and the $B\overline{B}$ SP3 Monte Carlo plus off-resonance

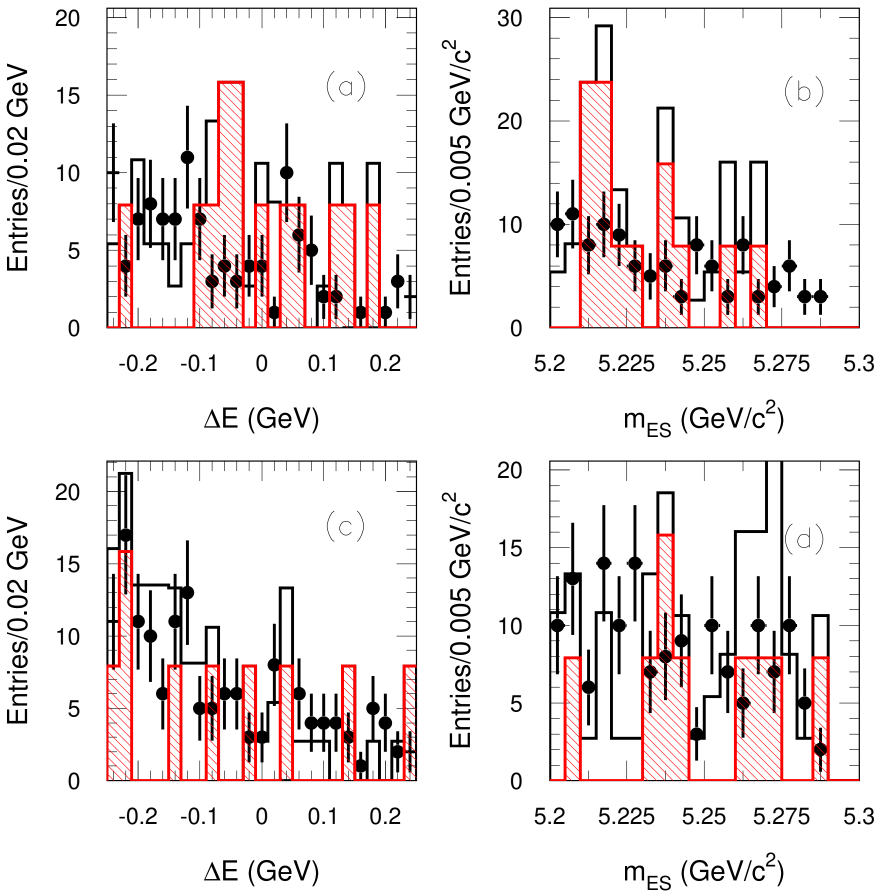


Figure 8.10: Comparison of ΔE and m_{ES} shapes in the $Ke\mu$ and $K^{*0}e\mu$ control sample. (a) ΔE for $Ke\mu$; (b) m_{ES} for $Ke\mu$; (c) ΔE for $K^{*0}e\mu$; (d) m_{ES} for $K^{*0}e\mu$. The dots are on-resonance data, the hashed histogram is off-resonance data, and the solid histogram is the sum of off-resonance data and generic SP3 $B\overline{B}$ Monte Carlo. The event selection criteria applied include tracking, particle ID, the Fisher and $B\overline{B}$ suppression cuts with the same values as those used for the corresponding di-electron channels, and $|\Delta E| < 0.25$ GeV, $m_{\mathrm{ES}} > 5.2$ GeV$/c^2$. The normalization is absolute.

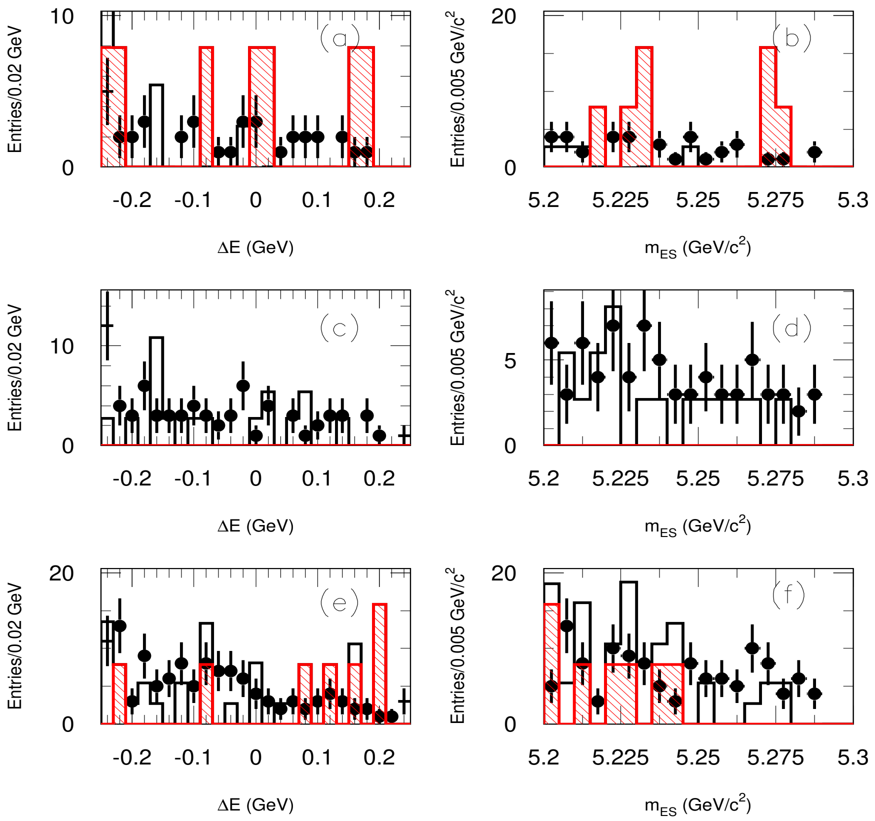


Figure 8.11: Comparison of ΔE and m_{ES} shapes in the $e\mu$ control sample for the modes with a K^0_S or a π^0. (a) ΔE for $K^0_S e\mu$; (b) m_{ES} for $K^0_S e\mu$; (c) ΔE for $K^{*+}e\mu$ ($K^{*+} \to K^0_S\pi^+$); (d) m_{ES} for $K^{*+}e\mu$ ($K^{*+} \to K^0_S\pi^+$); (e) ΔE for $K^{*+}e\mu$ ($K^{*+} \to K^+\pi^0$); (f) m_{ES} for $K^{*+}e\mu$ ($K^{*+} \to K^+\pi^0$). The dots are on-resonance data, the hashed histogram is off-resonance data, and the solid histogram is the sum of off-resonance data and generic SP3 $B\overline{B}$ Monte Carlo. The event selection criteria applied include tracking, particle ID, the Fisher and $B\overline{B}$ suppression cuts with the same values as those used for the corresponding di-electron channels, and $|\Delta E| < 0.25$ GeV, $m_{\mathrm{ES}} > 5.2$ GeV/c^2. The normalization is absolute.

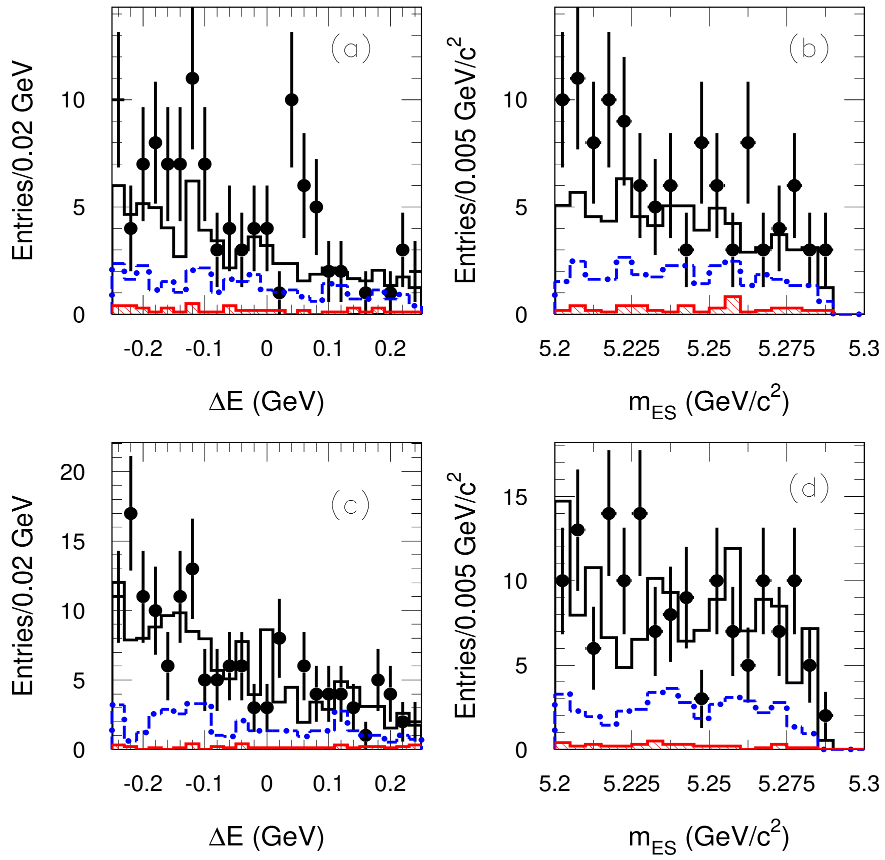


Figure 8.12: Comparison of ΔE and m_{ES} shapes in the $Ke\mu$ and $K^{*0}e\mu$ control sample. (a) ΔE for $Ke\mu$; (b) m_{ES} for $Ke\mu$; (c) ΔE for $K^{*0}e\mu$; (d) m_{ES} for $K^{*0}e\mu$. The dots are on-resonance data, the hashed histogram is Pravda *uds* continuum, the dashed histogram is the sum of Pravda Monte Carlo *uds* and $c\overline{c}$ continuum, and the solid histogram is the sum of continuum and $B\overline{B}$ Pravda Monte Carlo. The event selection criteria applied include tracking, particle ID, the Fisher and $B\overline{B}$ suppression cuts with the same values as those used for the corresponding di-electron channels, and $|\Delta E| < 0.25$ GeV, $m_{\mathrm{ES}} > 5.2$ GeV$/c^2$. The normalization is absolute.

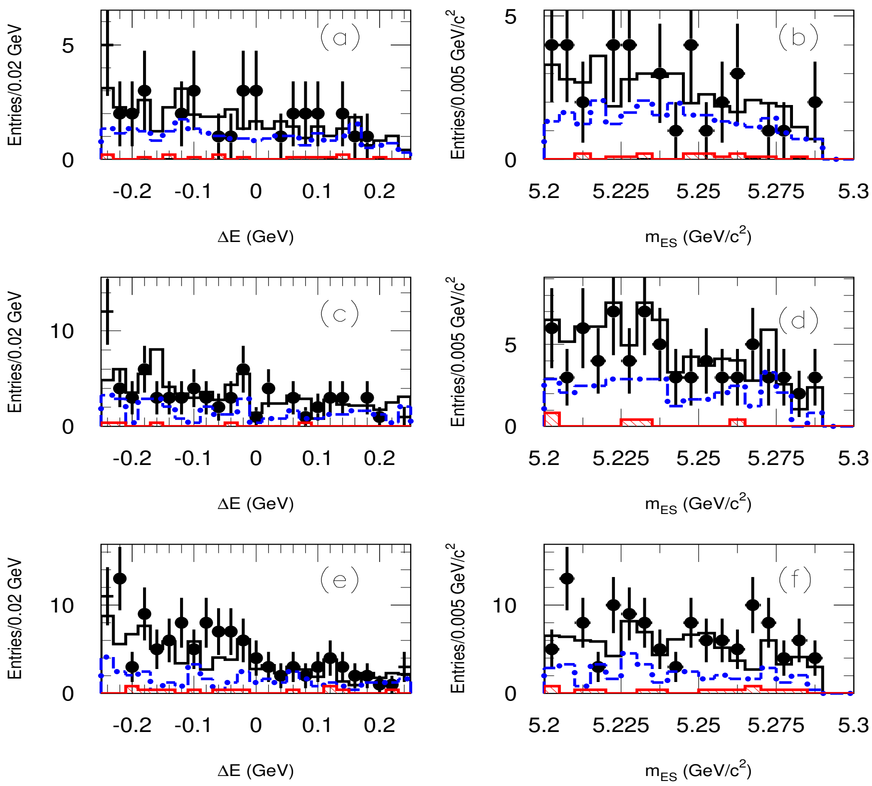


Figure 8.13: Comparison of ΔE and $m_{\rm ES}$ shapes in the $e\mu$ control sample for the modes with a K^0_S or a π^0. (a) ΔE for $K^0_S e\mu$; (b) $m_{\rm ES}$ for $K^0_S e\mu$; (c) ΔE for $K^{*+}e\mu$ ($K^{*+} \to K^0_S\pi^+$); (d) $m_{\rm ES}$ for $K^{*+}e\mu$ ($K^{*+} \to K^0_S\pi^+$); (e) ΔE for $K^{*+}e\mu$ ($K^{*+} \to K^+\pi^0$); (f) $m_{\rm ES}$ for $K^{*+}e\mu$ ($K^{*+} \to K^+\pi^0$). The dots are on-resonance data, the hashed histogram is Pravda uds continuum, the dashed histogram is the sum of Pravda Monte Carlo uds and $c\overline{c}$ continuum, and the solid histogram is the sum of continuum and $B\overline{B}$ Pravda Monte Carlo. The event selection criteria applied include tracking, particle ID, the Fisher and $B\overline{B}$ suppression cuts with the same values as those used for the corresponding di-electron channels, and $|\Delta E| < 0.25$ GeV, $m_{\rm ES} > 5.2$ GeV$/c^2$. The normalization is absolute.

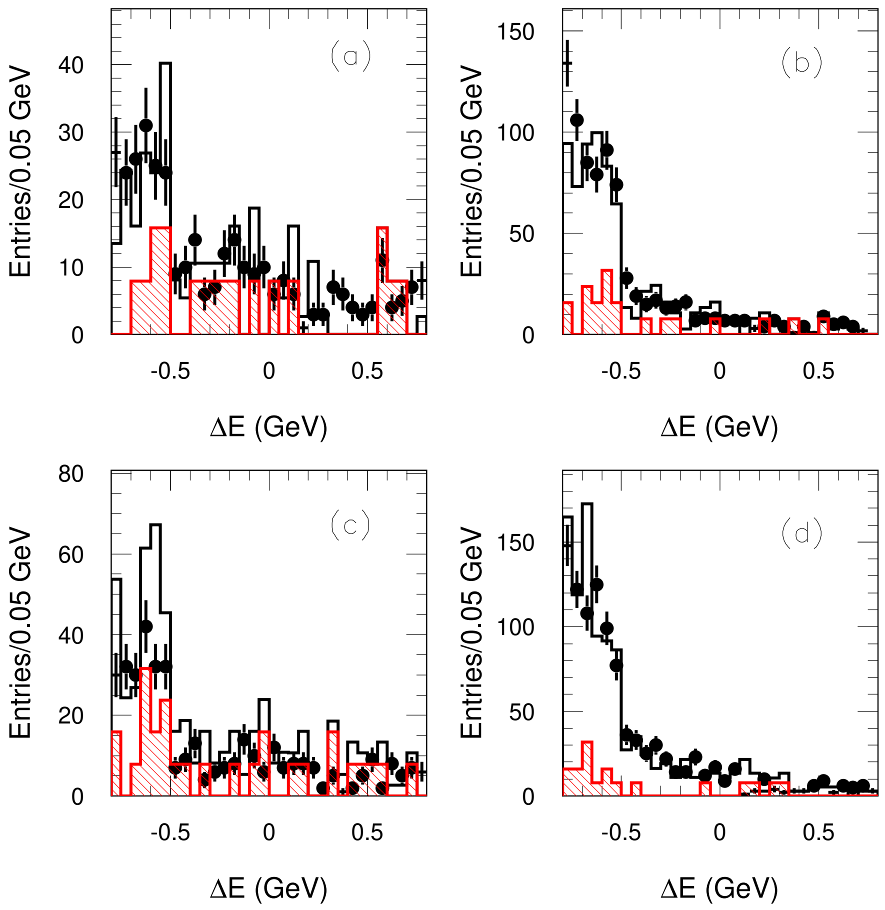


Figure 8.14: Comparison of the ΔE shapes in the large sideband for the charged track modes after all of the analysis selection criteria (with the large sideband selection requirements on ΔE and $m_{\rm ES}$). (a) $B^+ \to K^+ e^+ e^-$; (b) $B^+ \to K^+ \mu^+ \mu^-$; (c) $B^0 \to K^{*0} e^+ e^-$; (d) $B^0 \to K^{*0} \mu^+ \mu^-$. The dots are on-resonance data, the hashed histogram is off-resonance data, and the solid histogram is the sum of off-resonance data and generic SP3 $B\overline{B}$ Monte Carlo. The normalization is absolute.

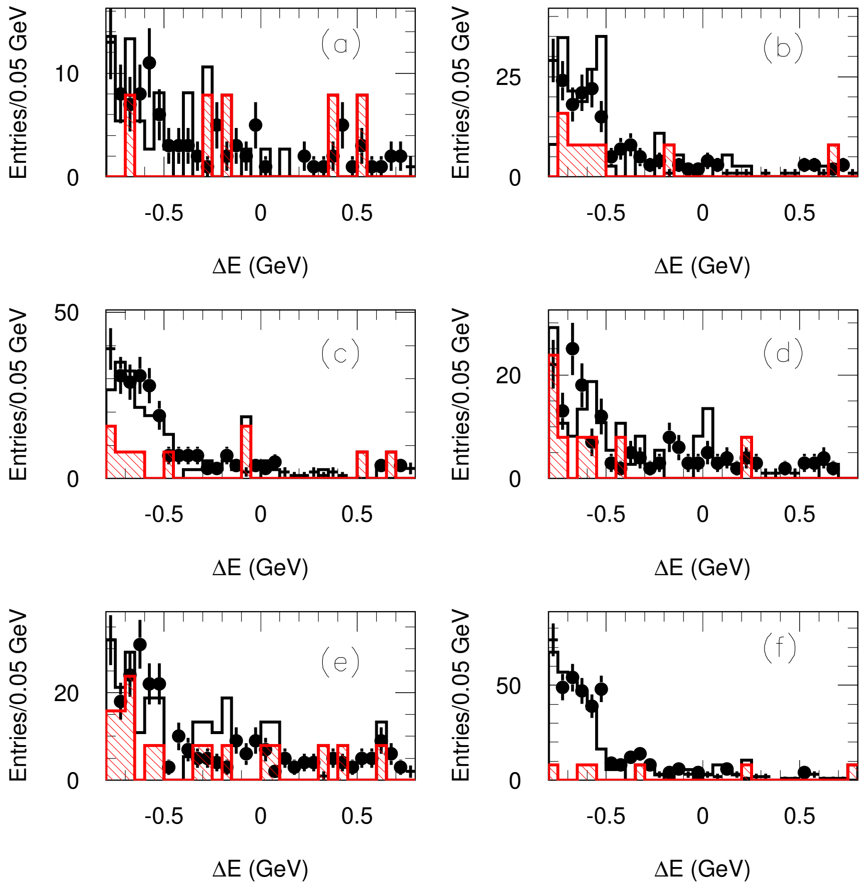


Figure 8.15: Comparison of the ΔE shapes in the large sideband for the modes with a K^0_S or a π^0 after all of the analysis selection criteria (with the large sideband selection requirements on ΔE and m_{ES}). (a) $B^0 \to K^0_S e^+e^-$; (b) $B^0 \to K^0_S \mu^+\mu^-$; (c) $B^+ \to K^{*+}e^+e^-$ $(K^{*+} \to K^0_S\pi^+)$; (d) $B^+ \to K^{*+}\mu^+\mu^-$ $(K^{*+} \to K^0_S\pi^+)$; (e) $B^+ \to K^{*+}e^+e^-$ $(K^{*+} \to K^+\pi^0)$; (f) $B^+ \to K^{*+}\mu^+\mu^-$ $(K^{*+} \to K^+\pi^0)$. The dots are on-resonance data, the hashed histogram is off-resonance data, and the solid histogram is the sum of off-resonance data and generic SP3 $B\overline{B}$ Monte Carlo. The normalization is absolute.

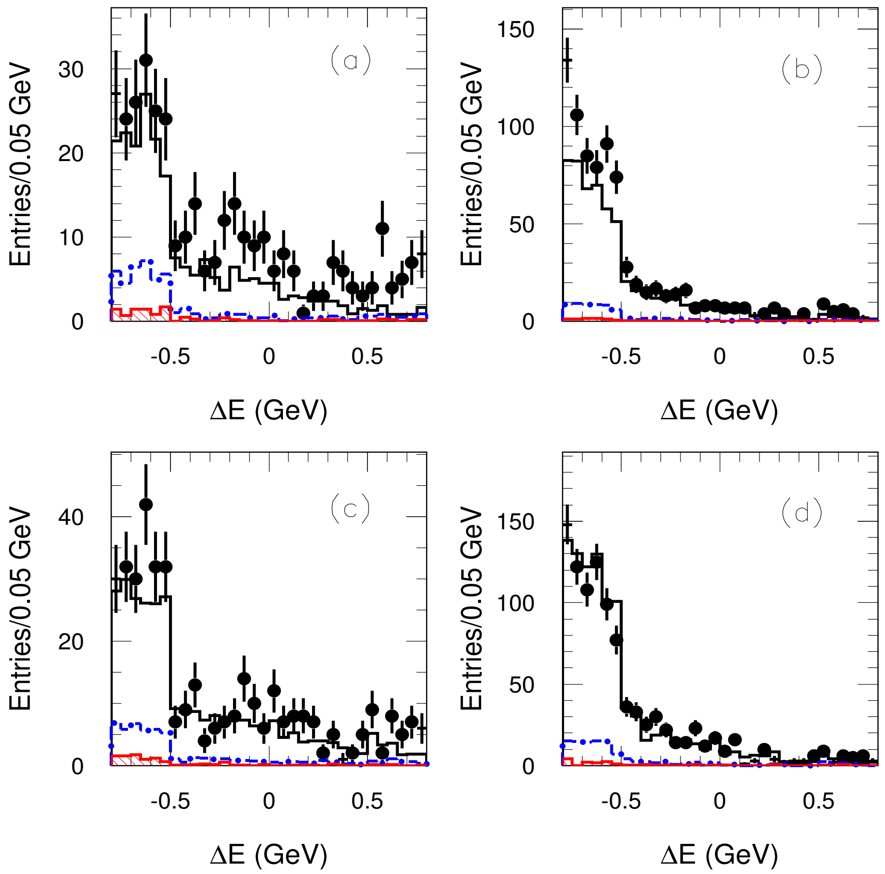


Figure 8.16: Comparison of the ΔE shapes in the large sideband for the charged track modes after all of the analysis selection criteria (with the large sideband selection requirements on ΔE and m_{ES}). (a) $B^+ \to K^+ e^+ e^-$; (b) $B^+ \to K^+ \mu^+ \mu^-$; (c) $B^0 \to K^{*0} e^+ e^-$; (d) $B^0 \to K^{*0} \mu^+ \mu^-$. The dots are on-resonance data, the hashed histogram is Pravda *uds* continuum, the dashed histogram is the sum of Pravda Monte Carlo *uds* and $c\bar{c}$ continuum, and the solid histogram is the sum of continuum and $B\overline{B}$ Pravda Monte Carlo. The normalization is absolute.

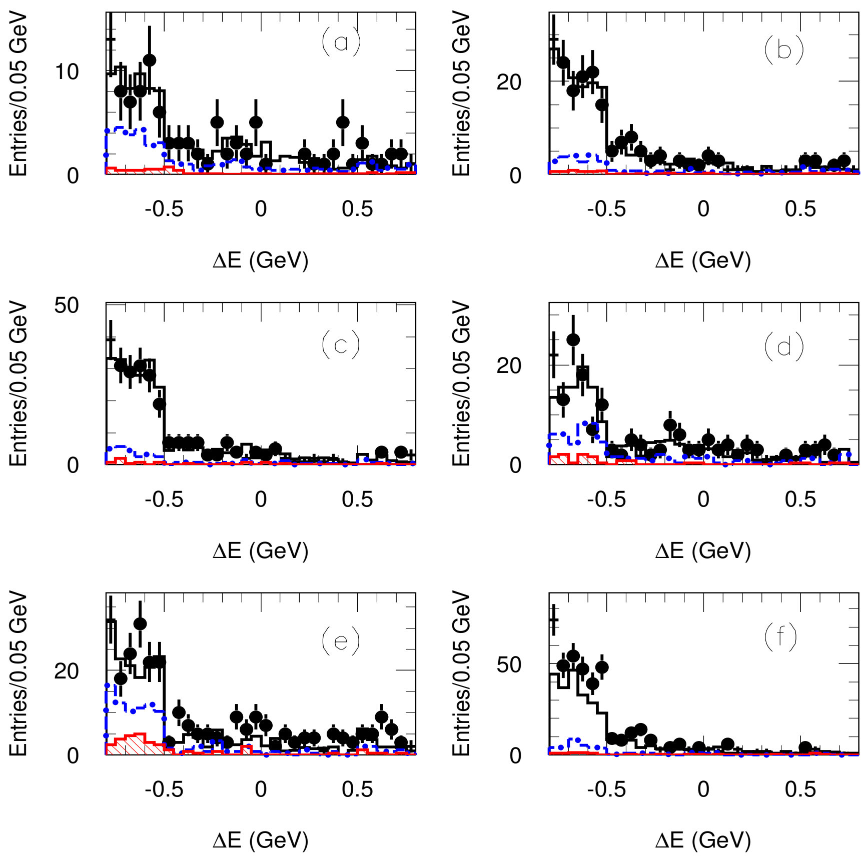


Figure 8.17: Comparison of the ΔE shapes in the large sideband for the modes with a K_S^0 or a π^0 after all of the analysis selection criteria (with the large sideband selection requirements on ΔE and $m_{\rm ES}$). (a) $B^0 \to K_S^0 e^+ e^-$; (b) $B^0 \to K_S^0 \mu^+\mu^-$; (c) $B^+ \to K^{*+} e^+ e^-$ $(K^{*+} \to K_S^0 \pi^+)$; (d) $B^+ \to K^{*+}\mu^+\mu^-$ $(K^{*+} \to K_S^0\pi^+)$; (e) $B^+ \to K^{*+} e^+ e^-$ $(K^{*+} \to K^+\pi^0)$; (f) $B^+ \to K^{*+}\mu^+\mu^-$ $(K^{*+} \to K^+\pi^0)$. The dots are on-resonance data, the hashed histogram is Pravda *uds* continuum, the dashed histogram is the sum of Pravda Monte Carlo *uds* and $c\overline{c}$ continuum, and the solid histogram is the sum of continuum and $B\overline{B}$ Pravda Monte Carlo. The normalization is absolute.

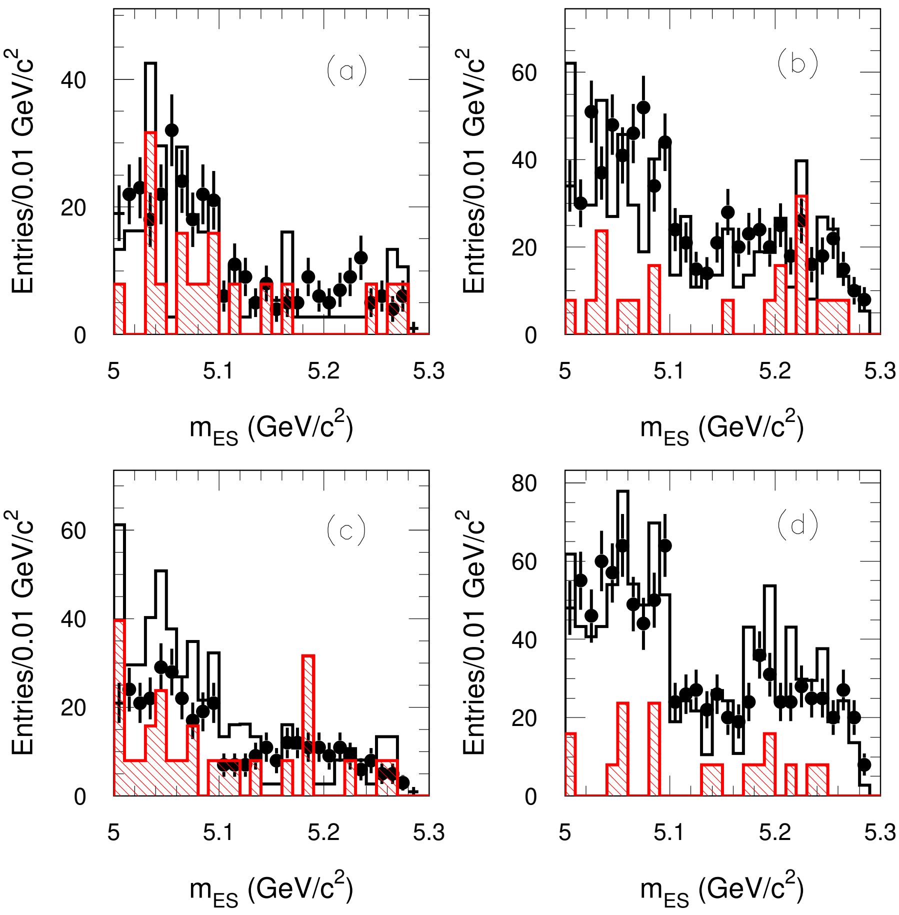


Figure 8.18: Comparison of the $m_{\rm ES}$ shapes in the large sideband for the charged track modes after all of the analysis selection criteria (with the large sideband selection requirements on ΔE and $m_{\rm ES}$). (a) $B^+ \to K^+e^+e^-$; (b) $B^+ \to K^+\mu^+\mu^-$; (c) $B^0 \to K^{*0}e^+e^-$; (d) $B^0 \to K^{*0}\mu^+\mu^-$. The dots are on-resonance data, the hashed histogram is off-resonance data, and the solid histogram is the sum of off-resonance data and generic SP3 $B\overline{B}$ Monte Carlo. The normalization is absolute.

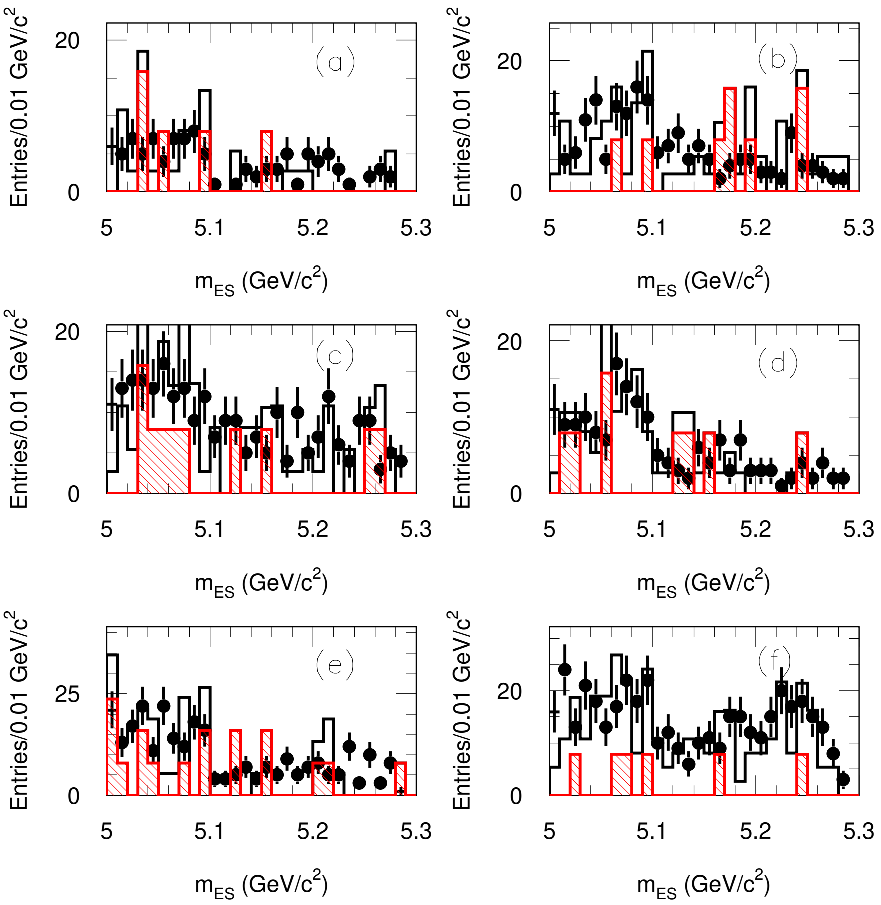


Figure 8.19: Comparison of the $m_{\rm ES}$ shapes in the large sideband for the modes with a K^0_S or a π^0 after all of the analysis selection criteria (with the large sideband selection requirements on ΔE and $m_{\rm ES}$). (a) $B^0 \to K^0_S e^+e^-$; (b) $B^0 \to K^0_S\mu^+\mu^-$; (c) $B^+ \to K^{*+}e^+e^-$ $(K^{*+} \to K^0_S\pi^+)$; (d) $B^+ \to K^{*+}\mu^+\mu^-$ $(K^{*+} \to K^0_S\pi^+)$; (e) $B^+ \to K^{*+}e^+e^-$ $(K^{*+} \to K^+\pi^0)$; (f) $B^+ \to K^{*+}\mu^+\mu^-$ $(K^{*+} \to K^+\pi^0)$. The dots are on-resonance data, the hashed histogram is off-resonance data, and the solid histogram is the sum of off-resonance data and generic SP3 $B\overline{B}$ Monte Carlo. The normalization is absolute.

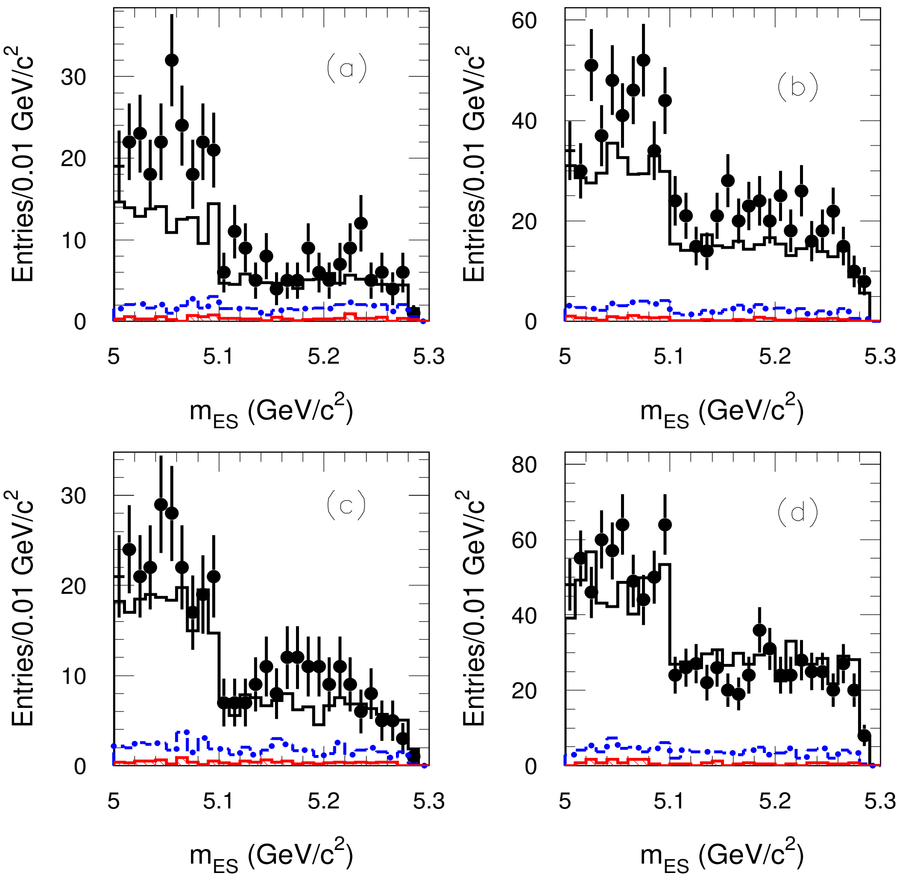


Figure 8.20: Comparison of the m_{ES} shapes in the large sideband for the charged track modes after all of the analysis selection criteria (with the large sideband selection requirements on ΔE and m_{ES}). (a) $B^+ \to K^+e^+e^-$; (b) $B^+ \to K^+\mu^+\mu^-$; (c) $B^0 \to K^{*0}e^+e^-$; (d) $B^0 \to K^{*0}\mu^+\mu^-$. The dots are on-resonance data, the hashed histogram is Pravda *uds* continuum, the dashed histogram is the sum of Pravda Monte Carlo *uds* and $c\bar{c}$ continuum, and the solid histogram is the sum of continuum and $B\overline{B}$ Pravda Monte Carlo. The normalization is absolute.

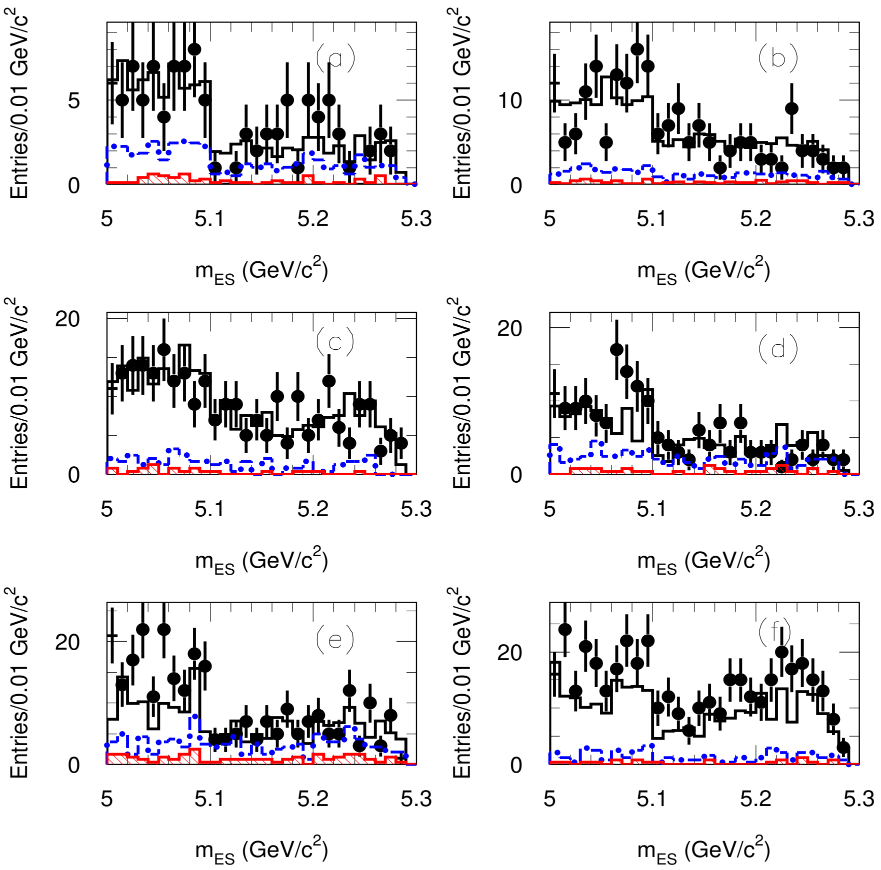


Figure 8.21: Comparison of the m_{ES} shapes in the large sideband for the modes with a K_S^0 or a π^0 after all of the analysis selection criteria (with the large sideband selection requirements on ΔE and m_{ES}). (a) $B^0 \to K_S^0 e^+e^-$; (b) $B^0 \to K_S^0\mu^+\mu^-$; (c) $B^+ \to K^{*+}e^+e^-$ $(K^{*+} \to K_S^0\pi^+)$; (d) $B^+ \to K^{*+}\mu^+\mu^-$ $(K^{*+} \to K_S^0\pi^+)$; (e) $B^+ \to K^{*+}e^+e^-$ $(K^{*+} \to K^+\pi^0)$; (f) $B^+ \to K^{*+}\mu^+\mu^-$ $(K^{*+} \to K^+\pi^0)$. The dots are on-resonance data, the hashed histogram is Pravda *uds* continuum, the dashed histogram is the sum of Pravda Monte Carlo *uds* and $c\overline{c}$ continuum, and the solid histogram is the sum of continuum and $B\overline{B}$ Pravda Monte Carlo. The normalization is absolute.

Mode	Data	SP3+off-res	Pravda MC
$K^+e^+e^-$	344.0 ± 18.5	315.8 ± 40.0	219.0 ± 4.8
$K^+\mu^+\mu^-$	785.0 ± 28.0	710.7 ± 52.4	574.2 ± 7.7
$K^{*0}\ e^+e^-$	376.0 ± 19.4	549.3 ± 51.4	298.0 ± 5.6
$K^{*0}\ \mu^+\mu^-$	993.0 ± 31.5	1038.7 ± 59.6	981.3 ± 10.1
$K^0_Se^+e^-$	105.0 ± 10.2	101.9 ± 22.0	102.2 ± 3.3
$K^0_S\mu^+\mu^-$	195.0 ± 14.0	206.8 ± 29.8	189.6 ± 4.4
$K^{*+}e^+e^-\ (K^{*+}\to K^0_S\pi^+)$	174.0 ± 13.2	168.9 ± 28.1	150.6 ± 3.9
$K^{*+}\mu^+\mu^-\ (K^{*+}\to K^0_S\pi^+)$	257.0 ± 16.0	236.2 ± 32.4	258.5 ± 5.2
$K^{*+}e^+e^-\ (K^{*+}\to K^+\pi^0)$	278.0 ± 16.7	248.3 ± 37.1	205.0 ± 4.6
$K^{*+}\mu^+\mu^-\ (K^{*+}\to K^+\pi^0)$	413.0 ± 20.3	358.9 ± 34.9	284.3 ± 5.4

Table 8.5: Large sideband yields in data and Monte Carlo, normalized to the luminosity of the on-resonance sample. The errors are statistical only. If no events were observed in the Monte Carlo samples, an error corresponding to one observed event is assigned.

data is at a 3σ level in the $K^{*0}e^+e^-$ mode. The comparison of the on-resonance data with Pravda Monte Carlo shows that the fast Monte Carlo is not modeling the number of events in the large sideband quite well. However, the comparison of the shapes of ΔE and m_{ES}, shown in Figs. 8.16, 8.17, 8.20, and 8.21, shows that there are no drastic disagreements. Pravda Monte Carlo cannot be expected to model the remote parts of the sideband very well, since some categories of events that are found there are either not represented in Pravda Monte Carlo at all (e.g., photon conversions) or are not adequately simulated (e.g., mismeasured tracks).

8.5 Control Sample Summary

Summarizing the results of the comparison between data and Monte Carlo in the four different control samples considered above, generic SP3 $B\overline{B}$ Monte Carlo plus off-resonance data appear to agree reasonably well with the on-resonance data. The comparison of the event yields and kinematic distributions in the charmonium control sample, described in Section 8.2, confirms that the charmonium background is understood and that the shapes of various kinematic distributions for signal are well-modeled in the Monte Carlo. The comparison of the background samples, such as the $(K, K^{*})e\mu$ combinations described in Section 8.3, and the large sideband described in Section 8.4, shows that the combinatorial background is largely understood. The disagreement in event yields between the on-resonance data and SP3 $B\overline{B}$ Monte Carlo plus off-resonance data is at a 3σ level in the worst comparison. The comparison of various distributions in the on-resonance data and in Pravda Monte Carlo shows that while the normalizations do not always agree, the shapes appear to be modeled reasonably well.

Chapter 9

Unbinned Maximum Likelihood Fit

In order to extract the number of signal and background events, an extended unbinned maximum likelihood fit is performed. The likelihood function is given by:

$$\mathcal{L} = \frac{e^{-(n_{\rm sig}+n_{B\overline{B}}+n_{\rm cont})}}{N!} \prod_{i=1}^{N} (n_{\rm sig} P_i^{\rm sig} + n_{B\overline{B}} P_i^{B\overline{B}} + n_{\rm cont} P_i^{\rm cont}), \tag{9.0.1}$$

where $n_{\rm sig}$, $n_{B\overline{B}}$, and $n_{\rm cont}$ are the number of signal, $B\overline{B}$ and continuum background events, respectively, which are extracted from the fit; $P_{\rm sig}$, $P_{B\overline{B}}$, and $P_{\rm cont}$ are the probability density functions (PDFs) for the corresponding hypotheses; N is the total number of input events; and i is the index over the input events. The advantages of using such a fit as opposed to a simple cut-and-count method are the increased sensitivity to signal achieved by including the signal shapes, a better determination of the background, and the fact that the fit makes it more straightforward to evaluate the significance of the signal in the presence of background.

The fitter has been written using a *BABAR* developed toolkit called RooFitTools [75]. Even though the fitter is flexible enough to allow all of the parameters to float, the approach used in the analysis described in this dissertation is to fix the shapes of the PDFs using Monte Carlo samples, and only let the number of events in each of the three fit categories float. Both the number of continuum events and the number of $B\overline{B}$ events are constrained to be non-negative, but no such constraint is imposed on the number of signal events. This issue is discussed in more detail in Section 9.2. The fitter performs a two-dimensional fit in the ΔE vs. m_{ES} plane. In principle, the sensitivity of the fit may be improved by including more variables that separate signal from background, such as the Fisher discriminants, the $B\overline{B}$ suppression variables, and the K^* mass.

9.1 Fitter Shapes

This section describes the shapes used to parametrize signal and background probability density functions.

- **Signal shapes:** $P^{sig}_{m_{\mathrm{ES}}}$ and $P^{sig}_{\Delta E}$, the signal m_{ES} and ΔE probability density functions, respectively, are parametrized using the Crystal Ball shape [73]. Both the standard form of the Crystal Ball shape, with a single Gaussian, and a modified form, with a double Gaussian, are used.

$$P^{\mathrm{sig}}_{m_{\mathrm{ES}}} = \frac{1}{N} \begin{cases} e^{-\frac{(x-\mu)^2}{2\sigma^2}} & \text{if } x > \mu - \alpha\,\sigma \\ \\ \frac{(n/\alpha)^n\; e^{-\frac{\alpha^2}{2}}}{[(\mu - x)/\sigma\, +\, n/\alpha\, -\, \alpha]^n} & \text{if } x \le \mu - \alpha\,\sigma, \end{cases} \tag{9.1.1}$$

$$P^{\mathrm{sig}}_{\Delta E} = \frac{1}{N} \begin{cases} \beta\, e^{-\frac{(x-\mu)^2}{2\sigma^2}} + (1-\beta)\, e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}} & \text{if } x > \mu - \alpha\,\sigma \\ \\ (n/\alpha)^n \frac{\beta\; e^{-\frac{\alpha^2}{2}} + (1-\beta)\; e^{-\frac{(\mu-\mu_2-\alpha\;\sigma)^2}{2\sigma_2^2}}}{[(\mu - x)/\sigma\, +\, n/\alpha\, -\, \alpha]^n} & \text{if } x \le \mu - \alpha\,\sigma, \end{cases} \tag{9.1.2}$$

  where N is the normalization factor, μ and σ are the mean and the width, respectively, of the narrow Gaussian, and μ_2 and σ_2 are the mean and the width, respectively, of the broader Gaussian; β characterizes the fraction of events in the two Gaussian components. The parameters α and n describe the radiative tail: α is a cross-over point from the Gaussian core to the radiative tail in units of the Gaussian peak with σ. If $\alpha > 0$, the tail is on the left of the peak. n is the tail power parameter: the lower it is, the longer the tail.

The two PDFs are not assumed to be independent: all four parameters of the m_{ES} PDF depend on ΔE. The dependence is quadratic: e.g., the μ parameter in the m_{ES} parameterization in Eqn. 9.1.1 is assumed to be of the form $\mu = \mu_0 + \mu_1\ \Delta E^2$, where μ_0 is the value μ would have in the absence of m_{ES}-ΔE correlations, and μ_1 is a constant parameter. The total signal PDF is thus:

$$P_{\mathrm{sig}}\left(m_{\mathrm{ES}}, \Delta E\right) = P^{\mathrm{sig}}_{\Delta E}\left(\Delta E\right)\ P^{\mathrm{sig}}_{m_{\mathrm{ES}}}\left(m_{\mathrm{ES}}; \Delta E\right), \tag{9.1.3}$$

where it is explicitly indicated that the PDF for m_{ES} is also a function of ΔE.

The fits to the signal SP3 Monte Carlo are shown in Figs. 9.1-9.10. The event selection criteria applied include all of the analysis selection requirements and the fit region cut on ΔE and m_{ES}: $|\Delta E| < 0.25$ GeV and $m_{\mathrm{ES}} > 5.2$ GeV/c^2. Tables 9.1 and 9.2 give examples of the fit parameters for signal SP3 $B^0 \to K^{*0}e^+e^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ samples, respectively.

Table 9.3 shows the efficiencies for signal events to lie in the fit region. Since there is no requirement that the events be within the signal box in ΔE and m_{ES}, the efficiencies are higher than those listed in Table 7.30, particularly for the electron channels. Given the efficiencies, one can calculate the expected number of signal events in the Run 1 data. It is about 14 events in all the modes, and 12 events without counting the π^0 modes.

- **Continuum background:** The ΔE probability density function, $P^{\mathrm{cont}}_{m_{\mathrm{ES}}}$, for the continuum background is parametrized as a straight line, and the m_{ES} PDF, $P^{\mathrm{cont}}_{m_{\mathrm{ES}}}$, as an Argus function [76]:

$$P^{\mathrm{cont}}_{m_{\mathrm{ES}}} = \frac{1}{N} m_{\mathrm{ES}}\ \sqrt{E_b^2 - m_{\mathrm{ES}}^2}\ e^{-\xi\left(E_b^2 - m_{\mathrm{ES}}^2\right)}, \tag{9.1.4}$$

ΔE parameter	Value
μ	(-10.4 $\pm$ 1.2) MeV
σ	(18.5 $\pm$ 1.0) MeV
α	0.70 $\pm$ 0.09
n	1.32 $\pm$ 0.20
μ_2	(50.0 $\pm$ 0.4) MeV
σ_2	(50.0 $\pm$ 11.6) MeV
β	0.94 $\pm$ 0.01
m_{ES} parameter	Value
μ_0	(5.2797 $\pm$ 0.0001) GeV/c^2
σ_0	(2.25 $\pm$ 0.07) MeV/c^2
α_0	1.58 $\pm$ 0.10
n_0	1.70 $\pm$ 0.19
μ_1	(-8.82 $\pm$ 2.05) $\times 10^{-2}$
σ_1	(2.18 $\pm$ 0.21) $\times 10^{-1}$
α_1	-5.00 $\pm$ 0.55
n_1	-3.00 $\pm$ 2.10

Table 9.1: An example of the fit parameters from the fit to the signal SP3 $B^0 \to K^{*0} e^+ e^-$ sample.

where N is the normalization factor, E_b is the beam energy, and ξ is the shape parameter. For this background, the ΔE and m_{ES} PDFs are assumed to be independent, so that the total PDF, P_{cont}, is given by:

$$P_{\mathrm{cont}}\left(m_{\mathrm{ES}}, \Delta E\right) = P^{\mathrm{cont}}_{\Delta E}\left(\Delta E\right)\ P^{\mathrm{cont}}_{m_{\mathrm{ES}}}\left(m_{\mathrm{ES}}\right). \tag{9.1.5}$$

An example of this parametrization is given in Fig. 9.11, which shows the ΔE and m_{ES} fits to the Pravda Monte Carlo continuum sample in the $B^+ \to K^+ \mu^+ \mu^-$ mode.

- **Generic $B\overline{B}$ background:** Parametrizing the $B\overline{B}$ background is somewhat complicated because ΔE and m_{ES} are not independent. The correlation is taken into account using the same method as for the signal parametrization.

ΔE parameter	Value
μ	(-8.28 $\pm$ 0.78) MeV
σ	(17.8 $\pm$ 0.7) MeV
α	1.79 $\pm$ 0.05
n	0.50 $\pm$ 0.08
μ_2	(50.0 $\pm$ 0.6) MeV
σ_2	(50.0 $\pm$ 0.2) MeV
β	0.94 $\pm$ 0.01
m_{ES} parameter	Value
μ_0	(5.2800 $\pm$ 0.0002) GeV/c^2
σ_0	(2.07 $\pm$ 0.14) MeV/c^2
α_0	1.29 $\pm$ 0.20
n_0	2.08 $\pm$ 0.42
μ_1	(-1.55 $\pm$ 0.72) $\times 10^{-1}$
σ_1	(5.00 $\pm$ 0.14) $\times 10^{-1}$
α_1	5.00 $\pm$ 0.68
n_1	-3.00 $\pm$ 2.57

Table 9.2: An example of the fit parameters from the fit to the signal SP3 $B^0 \to K^{*0}\mu^+\mu^-$ sample.

The PDF for ΔE, $P^{B\overline{B}}_{\Delta E}$, is parametrized as a quadratic exponential:

$$P^{B\overline{B}}_{\Delta E} = \frac{1}{N}\exp\left(s\Delta E + t\Delta E^2\right) \tag{9.1.6}$$

where s and t are constant coefficients of the polynomial, and N is the normalization constant.

The PDF for m_{ES}, $P^{B\overline{B}}_{m_{\mathrm{ES}}}$, is parametrized as an Argus function, whose functional form is given in Eqn. 9.1.4, with a shape parameter ξ that is a function of ΔE:

$$\xi = c_0 + c_1\ \Delta E + c_2\ (\Delta E)^2 \tag{9.1.7}$$

where c_0, c_1, and c_2 are constant parameters.

The total PDF for $B\overline{B}$ background is thus:

$$P_{B\overline{B}}\left(m_{\mathrm{ES}}, \Delta E\right) = P^{B\overline{B}}_{\Delta E}\left(\Delta E\right)\ P^{B\overline{B}}_{m_{\mathrm{ES}}}\left(m_{\mathrm{ES}}; \Delta E\right), \tag{9.1.8}$$

Mode	Efficiency (%)	Expected yield
$B^+ \to K^+e^+e^-$	17.5 ± 0.4	2.3
$B^0 \to K^0_S\, e^+e^-$	$15.7 \pm 0.$	0.7
$B^0 \to K^{*0}e^+e^-$	10.2 ± 0.3	3.5
$B^+ \to K^{*+}e^+e^-\ (K^{*+} \to K^0_S\pi^+)$	8.5 ± 0.3	1.0
$B^+ \to K^{*+}e^+e^-\ (K^{*+} \to K^+\pi^0)$	6.6 ± 0.3	1.1
$B^+ \to K^+\mu^+\mu^-$	10.5 ± 0.3	1.4
$B^0 \to K^0_S\, \mu^+\mu^-$	9.6 ± 0.3	0.4
$B^0 \to K^{*0}\mu^+\mu^-$	8.0 ± 0.3	2.4
$B^+ \to K^{*+}\mu^+\mu^-\ (K^{*+} \to K^0_S\pi^+)$	5.5 ± 0.3	0.5
$B^+ \to K^{*+}\mu^+\mu^-\ (K^{*+} \to K^+\pi^0)$	4.1 ± 0.2	0.6
TOTAL		13.8
TOTAL w/o π^0 modes		12.1

Table 9.3: The efficiency for SP3 signal Monte Carlo events to be in the fit region in ΔE and m_{ES} after passing all of the other analysis selection criteria. The expected number of signal events in the Run 1 data is also given.

where $P^{B\overline{B}}_{m_{\mathrm{ES}}}$ is not only a function of m_{ES} but also of ΔE.

The $B\overline{B}$ background parametrization is shown in Fig. 9.12 for the Pravda Monte Carlo $B\overline{B}$ sample in the $B^0 \to K^{*0}\mu^+\mu^-$ mode.

9.2 A Note on Low Statistics Likelihood Fits

In cases of low statistics, maximum likelihood fits may converge in such a way that the signal component becomes very negative whereas the background components appear to be higher than the data would suggest[1]. To understand how this may occur, let us consider the simple case of a likelihood fit with only two components, signal and background, in a sample with a fixed number of events N.

[1]The discussion in this section is indebted to an informal note entitled *Low Statistics Likelihood Fits Are Tricky* written by Frank Porter (Caltech)

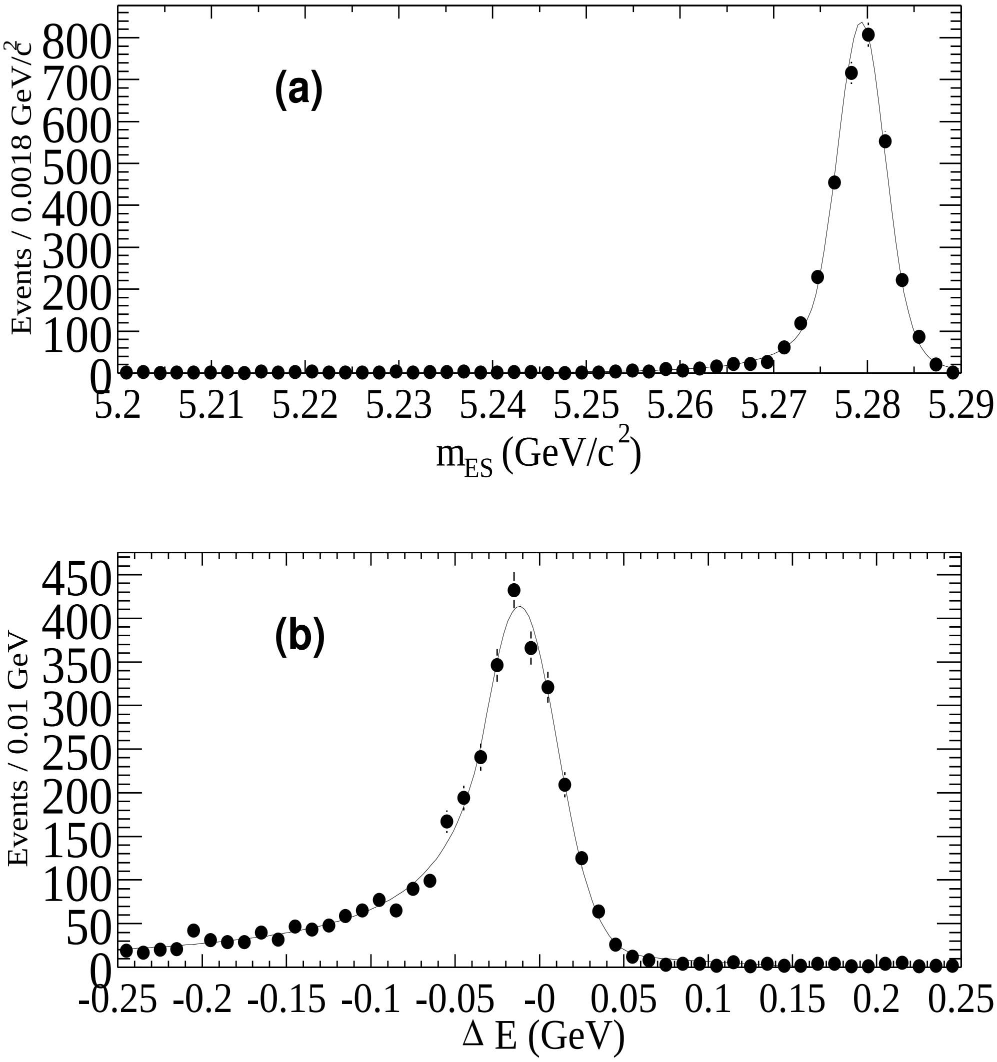
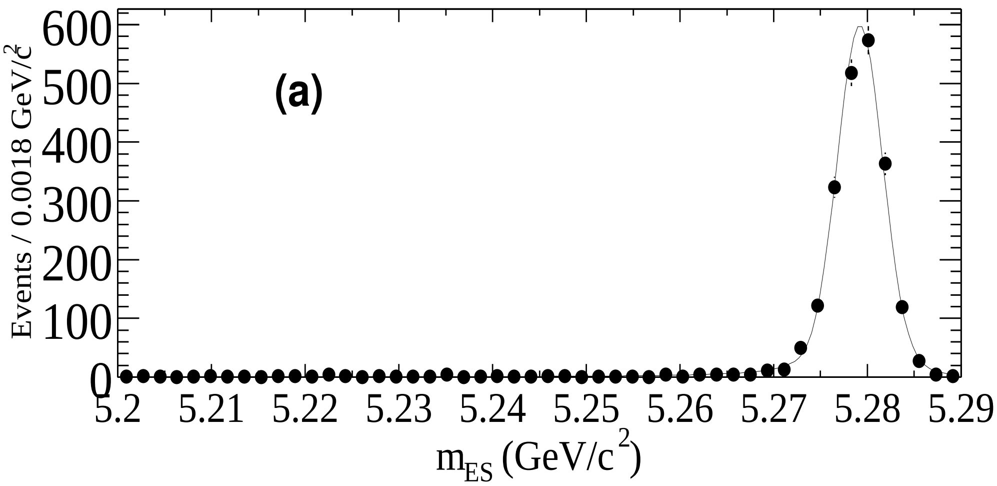


Figure 9.1: (a) m_{ES} and (b) ΔE distributions in the $B^+ \to K^+e^+e^-$ SP3 signal Monte Carlo, fitted to the shapes given in Eqn. 9.1.1. The points with error bars are the Monte Carlo, and the lines are the fitter shapes. All of the event selection criteria have been applied, with loose ΔE ($|\Delta E| < 0.25$ GeV) and m_{ES} ($m_{\mathrm{ES}} > 5.2$ GeV/c^2) cuts.



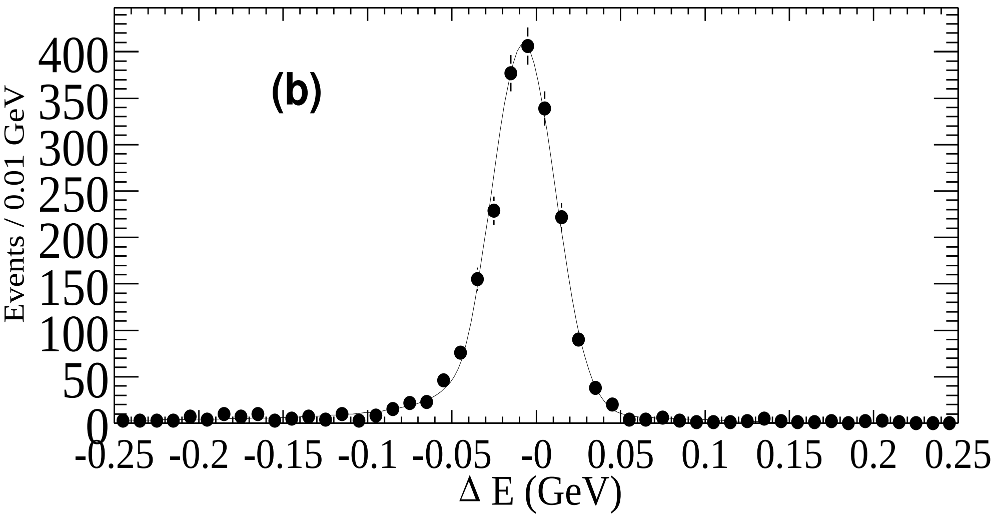
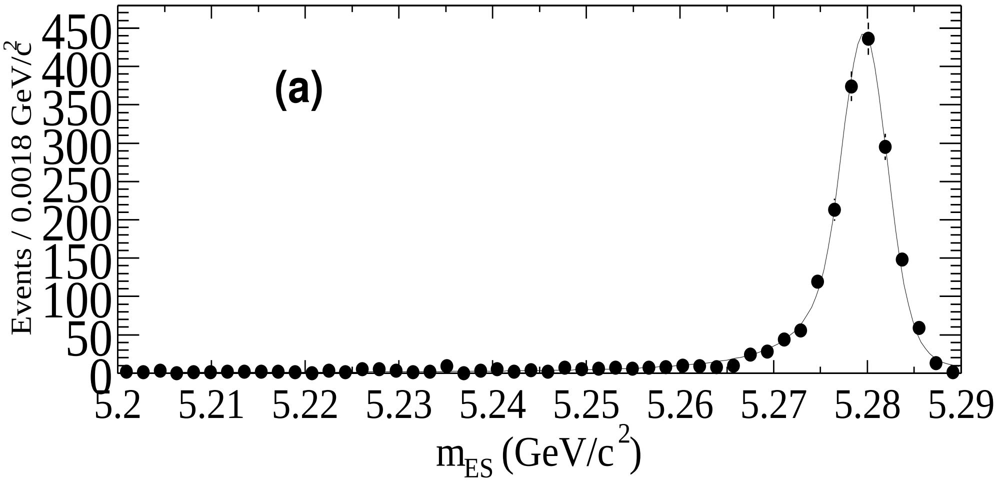


Figure 9.2: (a) m_{ES} and (b) ΔE distributions in the $B^+ \to K^+\mu^+\mu^-$ SP3 signal Monte Carlo, fitted to the shapes given in Eqn. 9.1.1. The points with error bars are the Monte Carlo, and the lines are the fitter shapes. All of the event selection criteria have been applied, with loose ΔE ($|\Delta E| < 0.25$ GeV) and m_{ES} ($m_{\mathrm{ES}} > 5.2$ GeV/c^2) cuts.



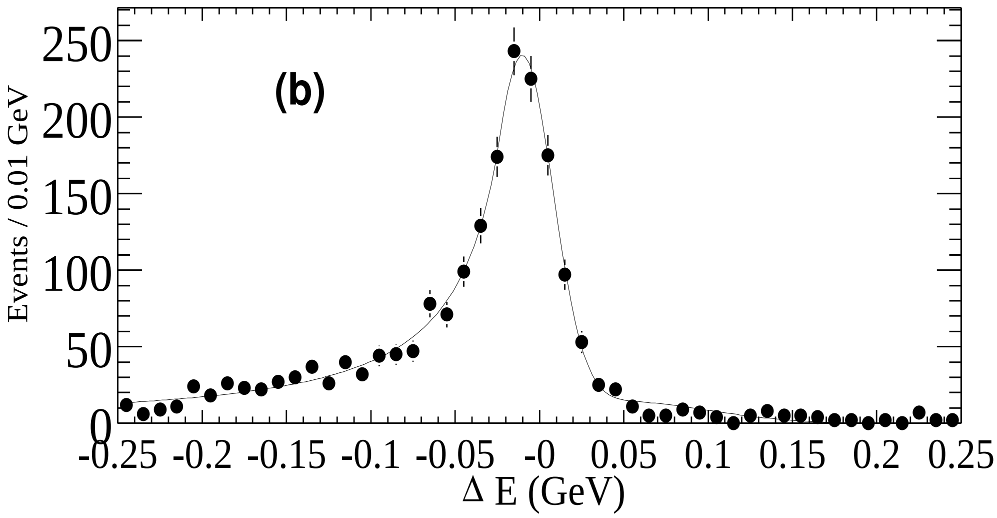


Figure 9.3: (a) m_{ES} and (b) $\mathit{\Delta E}$ distributions in the $B^0 \to K^{*0}e^+e^-$ SP3 signal Monte Carlo, fitted to the shapes given in Eqn. 9.1.1. The points with error bars are the Monte Carlo, and the lines are the fitter shapes. All of the event selection criteria have been applied, with loose $\mathit{\Delta E}$ ($|\mathit{\Delta E}| < 0.25$ GeV) and m_{ES} ($m_{\mathrm{ES}} > 5.2$ GeV/c^2) cuts.

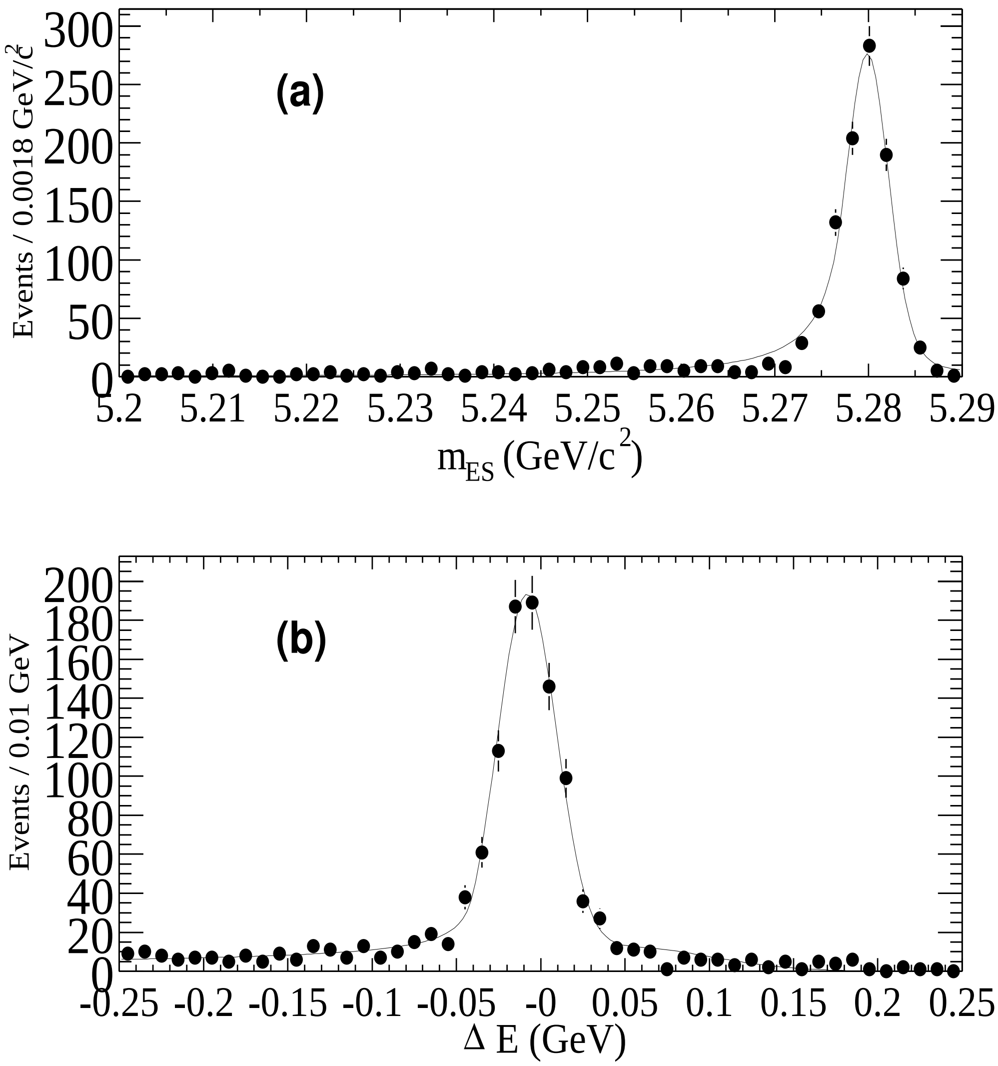


Figure 9.4: (a) $m_{\rm ES}$ and (b) ΔE distributions in the $B^0 \to K^{*0}\mu^+\mu^-$ SP3 signal Monte Carlo, fitted to the shapes given in Eqn. 9.1.1. The points with error bars are the Monte Carlo, and the lines are the fitter shapes. All of the event selection criteria have been applied, with loose ΔE ($|\Delta E| < 0.25$ GeV) and $m_{\rm ES}$ ($m_{\rm ES} > 5.2$ GeV/c^2) cuts.

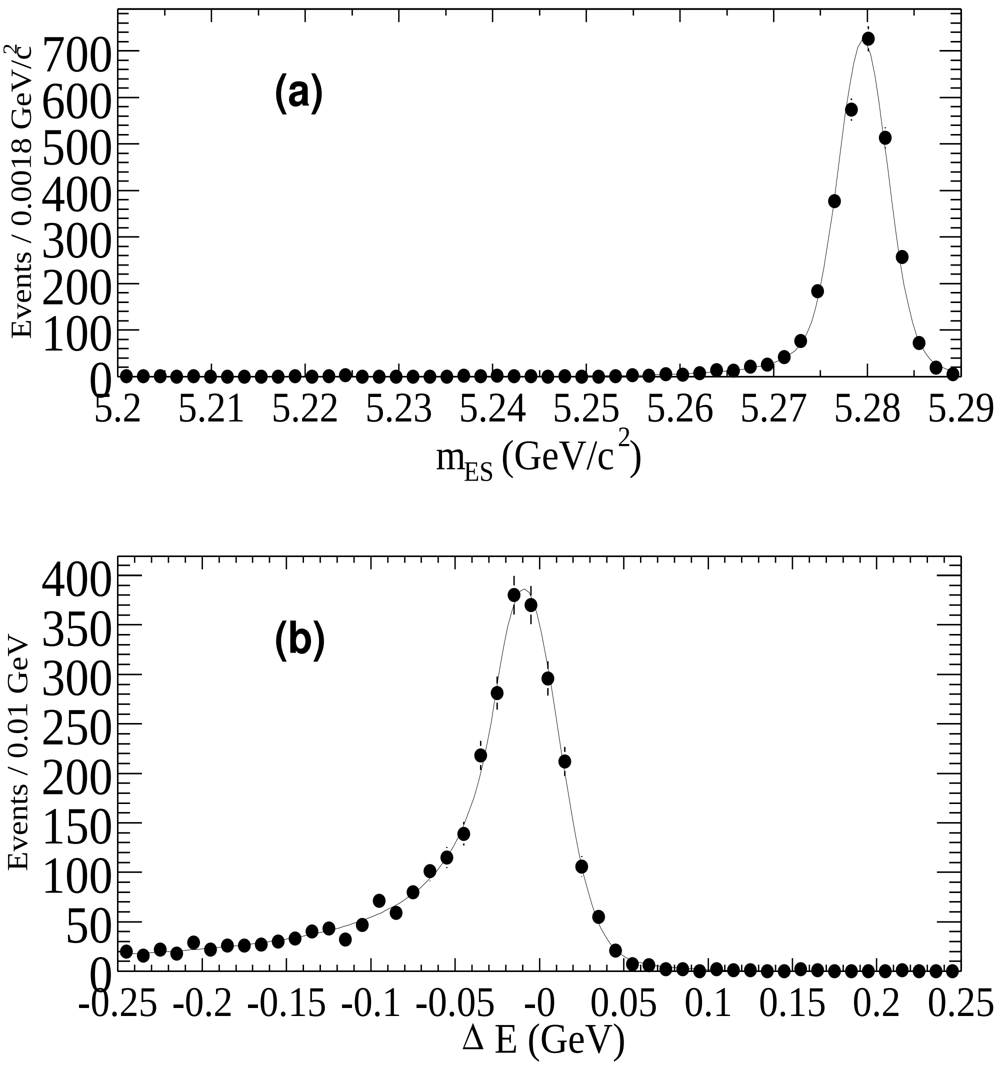


Figure 9.5: (a) m_{ES} and (b) ΔE distributions in the $B^0 \rightarrow K_S^0 e^+ e^-$ SP3 signal Monte Carlo, fitted to the shapes given in Eqn. 9.1.1. The points with error bars are the Monte Carlo, and the lines are the fitter shapes. All of the event selection criteria have been applied, with loose ΔE ($|\Delta E| < 0.25$ GeV) and m_{ES} ($m_{\mathrm{ES}} > 5.2$ GeV/c^2) cuts.

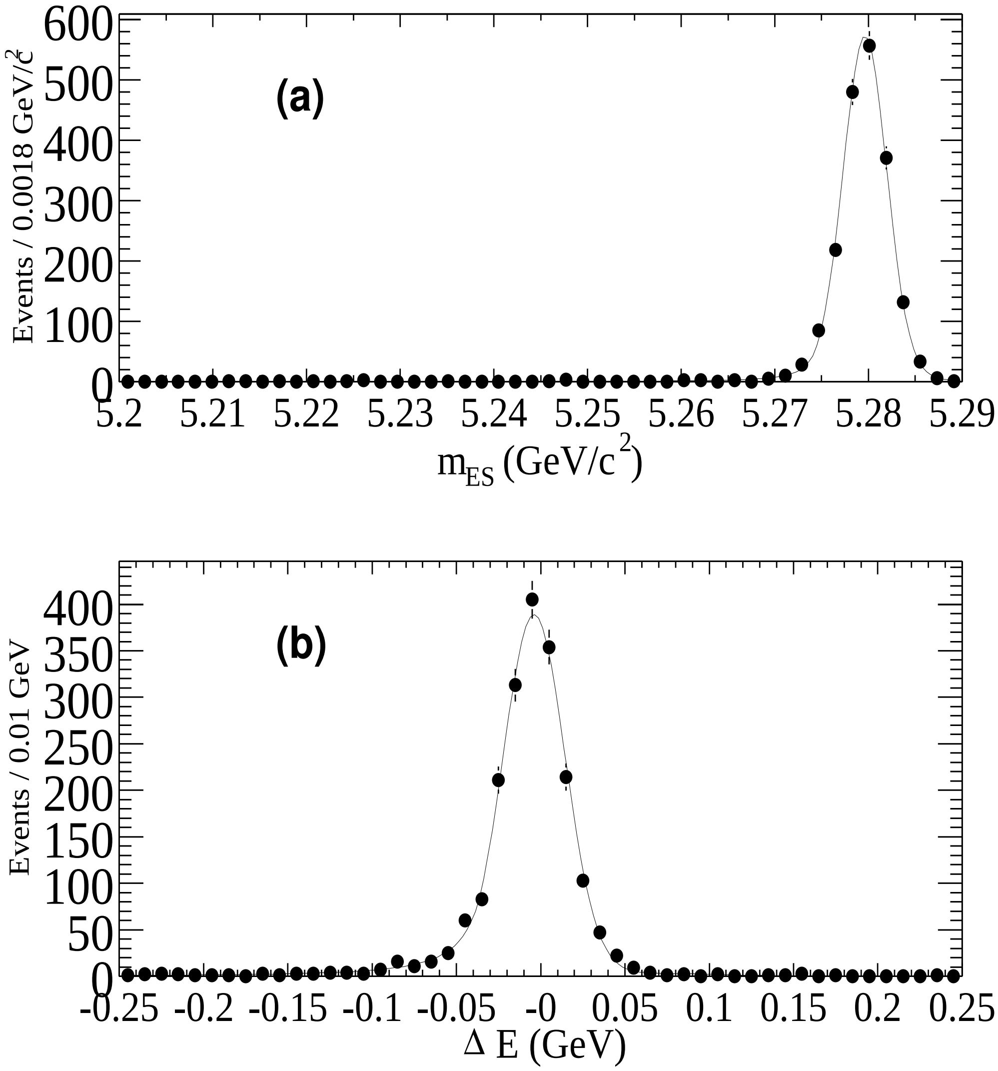


Figure 9.6: (a) m_{ES} and (b) ΔE distributions in the $B^0 \to K^0_S\mu^+\mu^-$ SP3 signal Monte Carlo, fitted to the shapes given in Eqn. 9.1.1. The points with error bars are the Monte Carlo, and the lines are the fitter shapes. All of the event selection criteria have been applied, with loose ΔE ($|\Delta E| < 0.25$ GeV) and m_{ES} ($m_{\mathrm{ES}} > 5.2$ GeV/c^2) cuts.

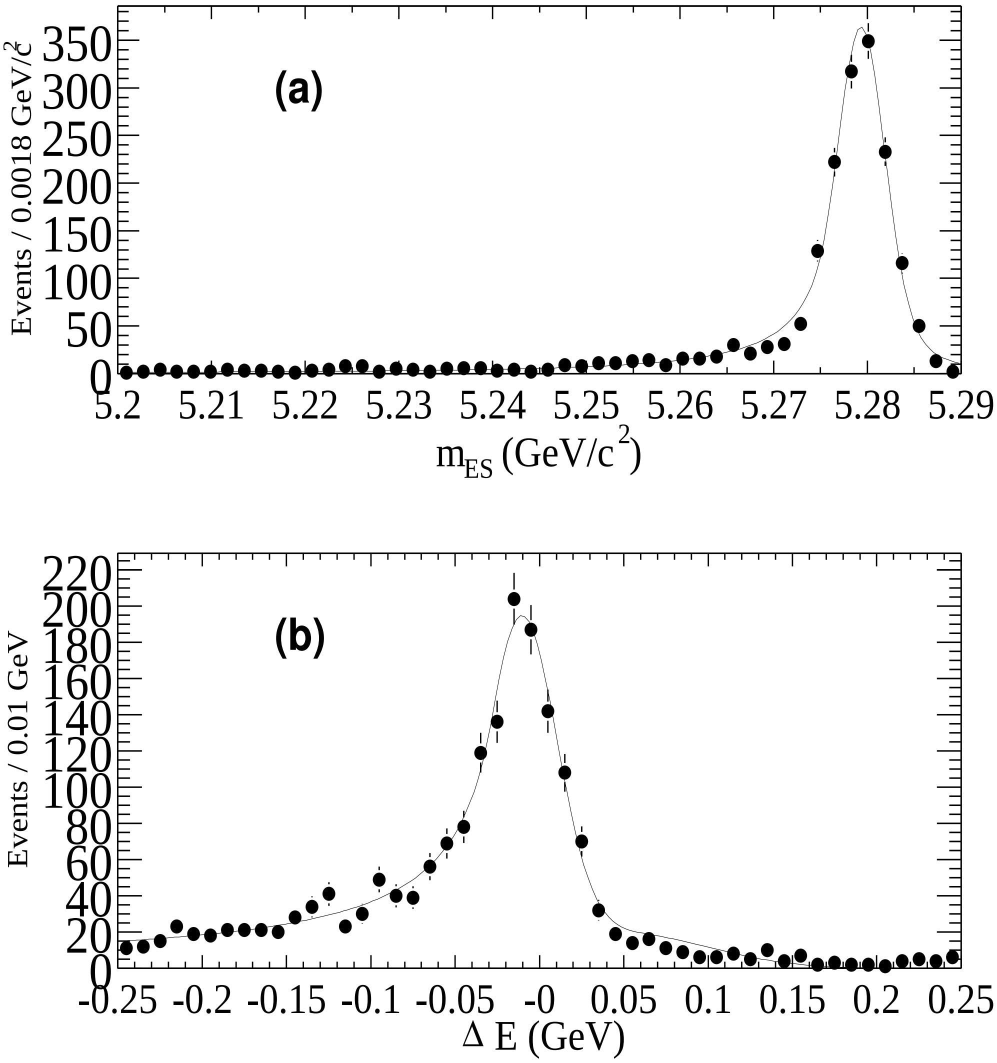


Figure 9.7: (a) m_{ES} and (b) ΔE distributions in the $B^+ \to K^{*+} e^+ e^-$, $K^{*+} \to K^0_S \pi^+$, SP3 signal Monte Carlo, fitted to the shapes given in Eqn. 9.1.1. The points with error bars are the Monte Carlo, and the lines are the fitter shapes. All of the event selection criteria have been applied, with loose ΔE ($|\Delta E| < 0.25$ GeV) and m_{ES} ($m_{\mathrm{ES}} > 5.2$ GeV/c^2) cuts.

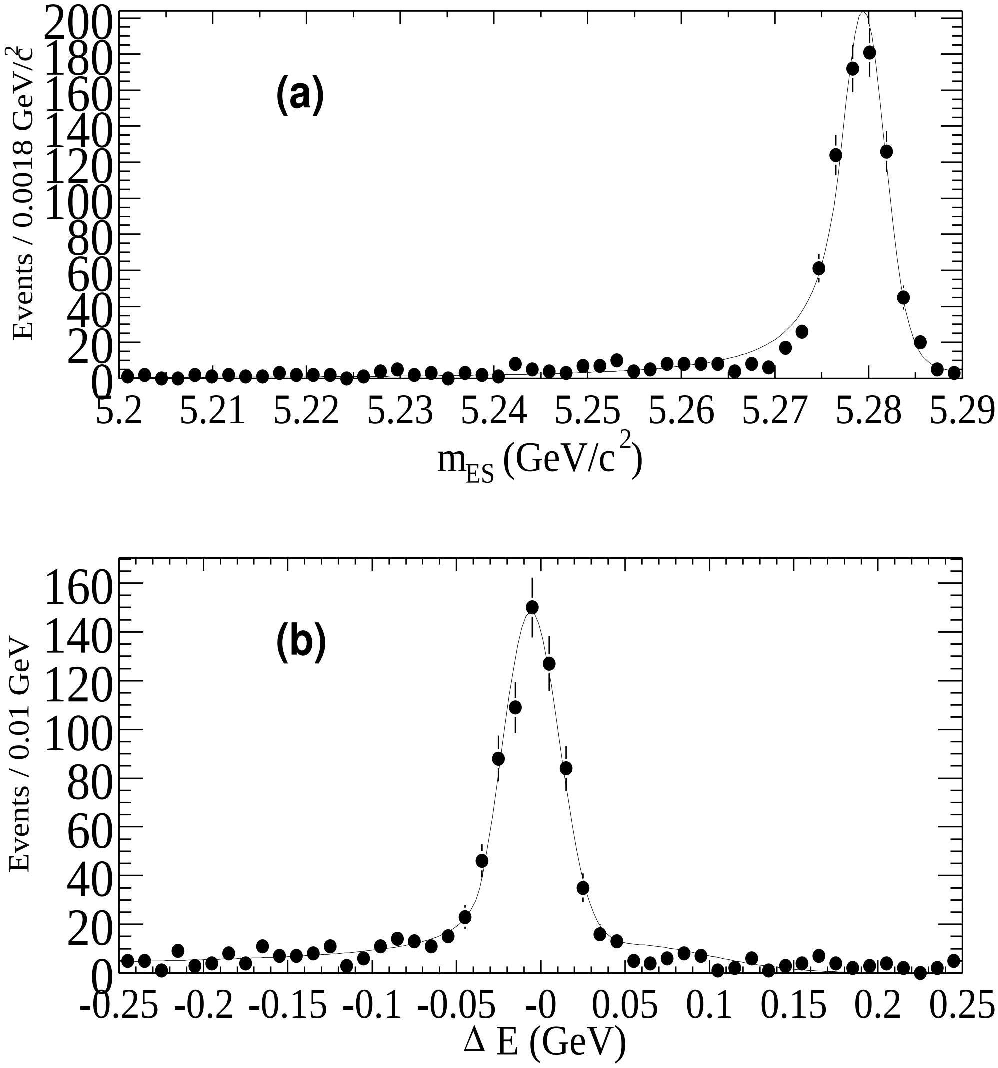


Figure 9.8: (a) $m_{\rm ES}$ and (b) $\varDelta E$ distributions in the $B^+ \rightarrow K^{*+}\mu^+\mu^-$, $K^{*+} \rightarrow K^0_S\pi^+$, SP3 signal Monte Carlo, fitted to the shapes given in Eqn. 9.1.1. The points with error bars are the Monte Carlo, and the lines are the fitter shapes. All of the event selection criteria have been applied, with loose $\varDelta E$ ($|\varDelta E| < 0.25$ GeV) and $m_{\rm ES}$ ($m_{\rm ES} > 5.2$ GeV/c^2) cuts.

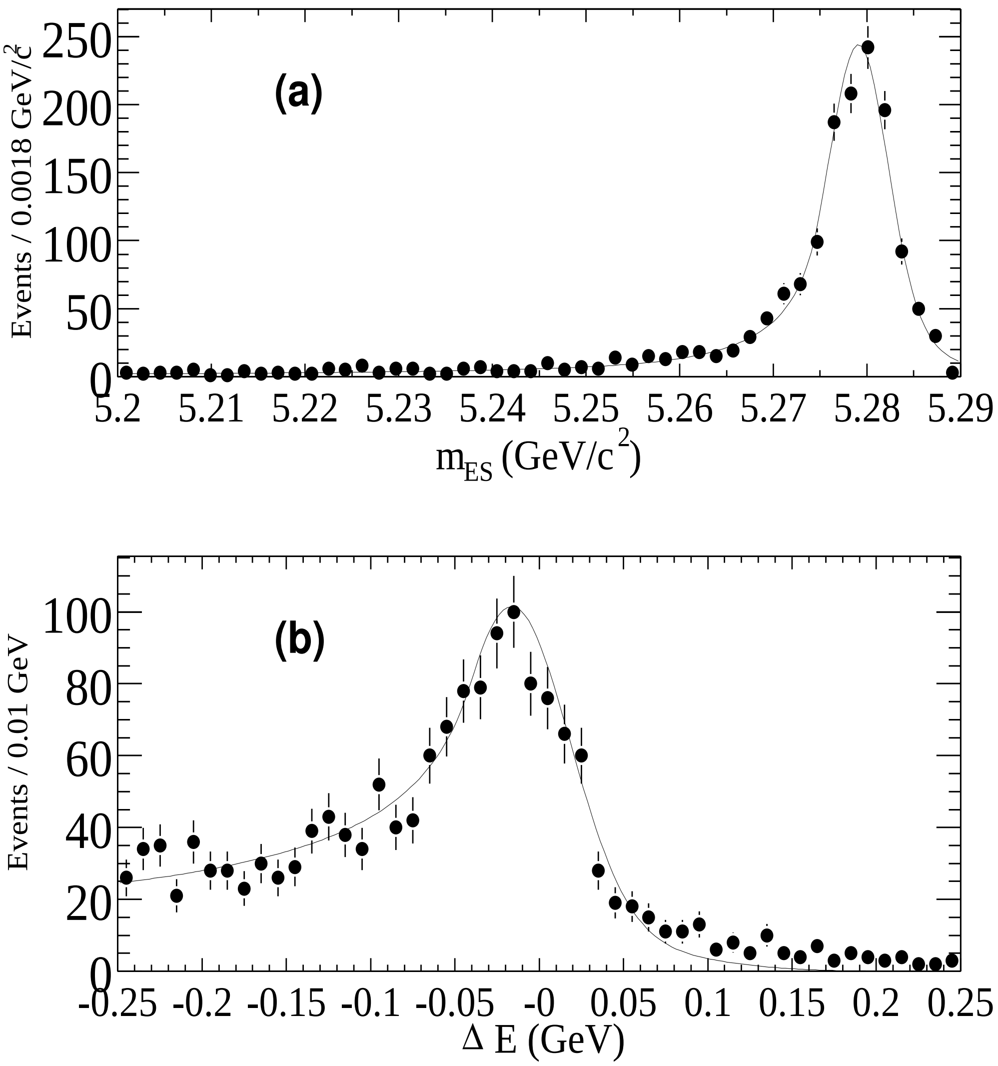


Figure 9.9: (a) $m_{\rm ES}$ and (b) ΔE distributions in the $B^+ \rightarrow K^{*+}e^+e^-$, $K^{*+} \rightarrow K + \pi^0$, SP3 signal Monte Carlo, fitted to the shapes given in Eqn. 9.1.1. The points with error bars are the Monte Carlo, and the lines are the fitter shapes. All of the event selection criteria have been applied, with loose ΔE ($|\Delta E| < 0.25$ GeV) and $m_{\rm ES}$ ($m_{\rm ES} > 5.2$ GeV/c^2) cuts.

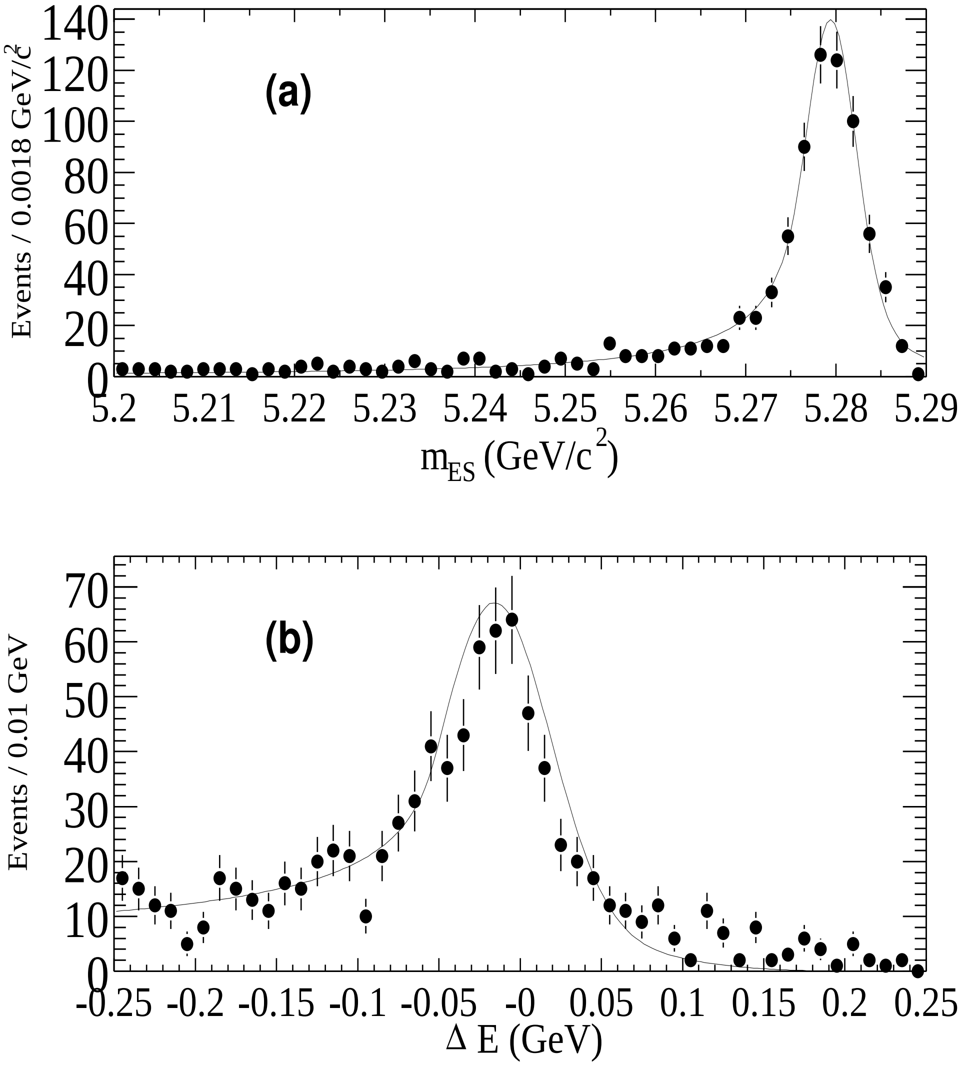


Figure 9.10: (a) m_{ES} and (b) ΔE distributions in the $B^+ \rightarrow K^{*+}\mu^+\mu^-$, $K^{*+} \rightarrow K^+\pi^0$, SP3 signal Monte Carlo, fitted to the shapes given in Eqn. 9.1.1. The points with error bars are the Monte Carlo, and the lines are the fitter shapes. All of the event selection criteria have been applied, with loose ΔE ($|\Delta E| < 0.25$ GeV) and m_{ES} ($m_{\mathrm{ES}} > 5.2$ GeV/c^2) cuts.

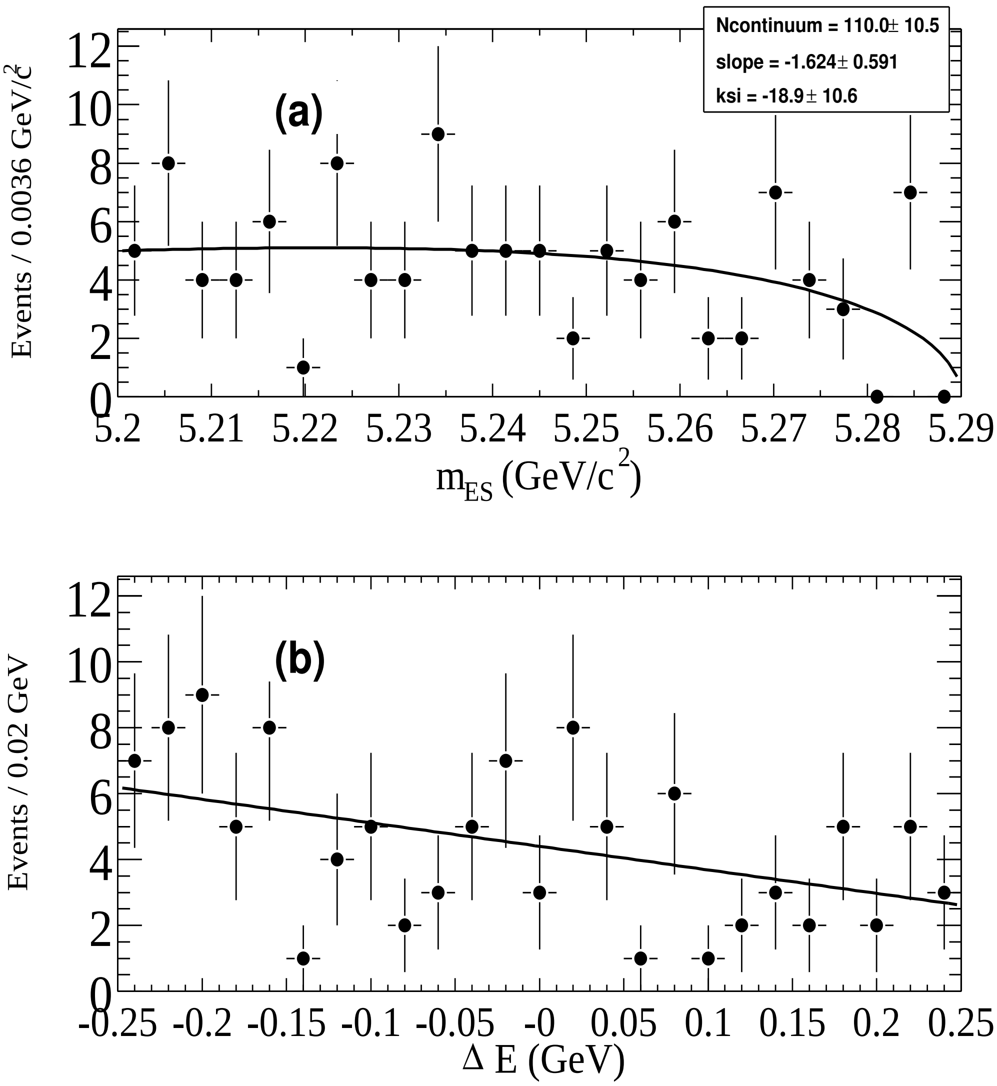
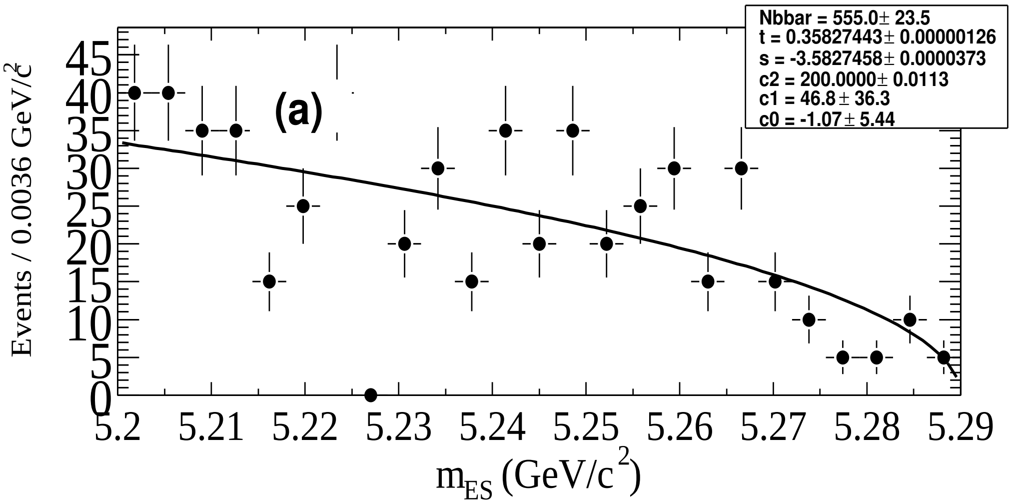


Figure 9.11: The shapes used to parametrize ΔE and $m_{\rm ES}$ for the continuum background, shown for the Pravda continuum *udsc* sample in the $B^+ \rightarrow K^+\mu^+\mu^-$ mode. (a) $m_{\rm ES}$ is fit to the Argus function. (b) ΔE is fit to a straight line. All of the analysis selection criteria have been applied, with a loose selection on ΔE and $m_{\rm ES}$: $|\Delta E| < 0.25$ GeV and $m_{\rm ES} > 5.2$ GeV$/c^2$.



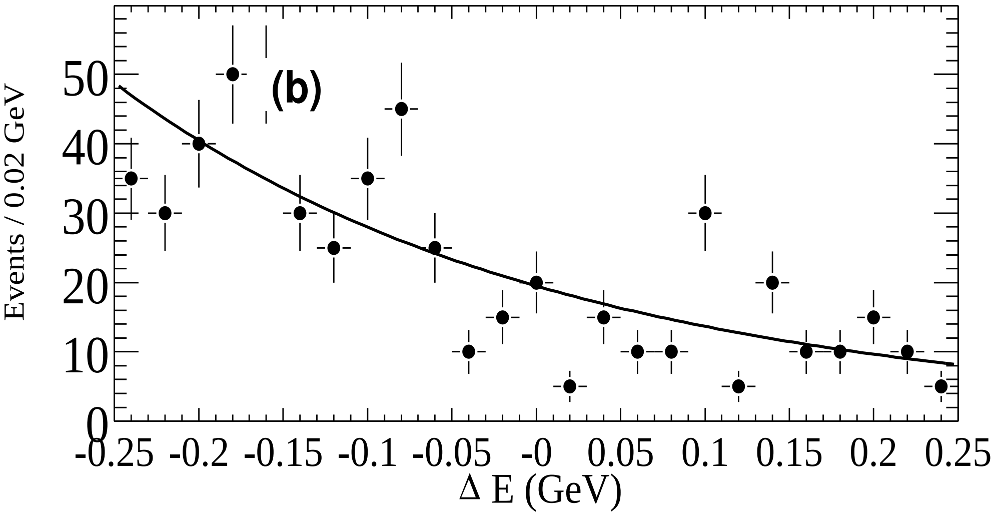


Figure 9.12: The shapes used to parametrize ΔE and m_{ES} for the $B\overline{B}$ background, shown for the Pravda Monte Carlo $B\overline{B}$ sample in the $B^0 \to K^{*0}\mu^+\mu^-$ mode. (a) m_{ES} is fit to the Argus function. (b) ΔE is fit to a quadratic exponential. All of the analysis selection criteria have been applied, with a loose selection on ΔE and m_{ES}: $|\Delta E| < 0.25$ GeV and $m_{\mathrm{ES}} > 5.2$ GeV/c^2.

Then the likelihood function is given by:

$$\mathcal{L} = \prod_{i=1}^{N} \left[\hat{n}_{\rm sig} P_{\rm sig} + \hat{n}_{\rm bkg} P_{\rm bkg}\right], \tag{9.2.1}$$

where $\hat{n}_{\rm sig}$ and $\hat{n}_{\rm bkg}$ are the normalized number of signal and background events, respectively, and $P_{\rm sig}$ and $P_{\rm bkg}$ are the total probability density functions for signal and background, respectively. A maximum likelihood fit will attempt to find a solution of the following equation:

$$-\sum_{i=1}^{N} \frac{\partial \ln \mathcal{L}}{\partial \hat{n}_{\rm sig}} = -\sum_{i=1}^{N} \frac{\partial \ln\left[\hat{n}_{\rm sig} P_{\rm sig} + (1-\hat{n}_{\rm sig}) P_{\rm bkg}\right]}{\partial \hat{n}_{\rm sig}} = \sum_{i=1}^{N} \frac{P_{\rm bkg} - P_{\rm sig}}{\hat{n}_{\rm sig} P_{\rm sig} + (1-\hat{n}_{\rm sig}) P_{\rm bkg}} = 0. \tag{9.2.2}$$

It is evident that if $P_{\rm sig} \to 0$, which may happen if there are very few or no events in the region where the signal is expected to peak, the solution will correspond to $\hat{n}_{\rm sig} \to -\infty$. An example is shown in Figs. 9.13 and 9.14, which show fits to one of the 20 fb^{-1} Pravda Monte Carlo samples described in Section 9.3 in the $B^+ \to K^+\mu^+\mu^-$ mode.

One way to solve this problem is to constrain the number of signal events to be always non-negative. However, this approach would ignore negative fluctuations in the data, so that meaningful averaging of results over a large number of experiments performed under the same conditions would not be possible. The solution adapted in the present analysis is to constrain the likelihood function to be always non-negative. This allows negative signal yields while avoiding the problem of very large negative fluctuations. However, this is only the first step towards solving the problem, since the artificial constraint on the likelihood function makes averaging the results still somewhat dubious.

9.3 Validation of the Fitter Yields

The fitter performance is checked by verifying that event yields obtained from the fits to samples with known numbers of signal and background events are

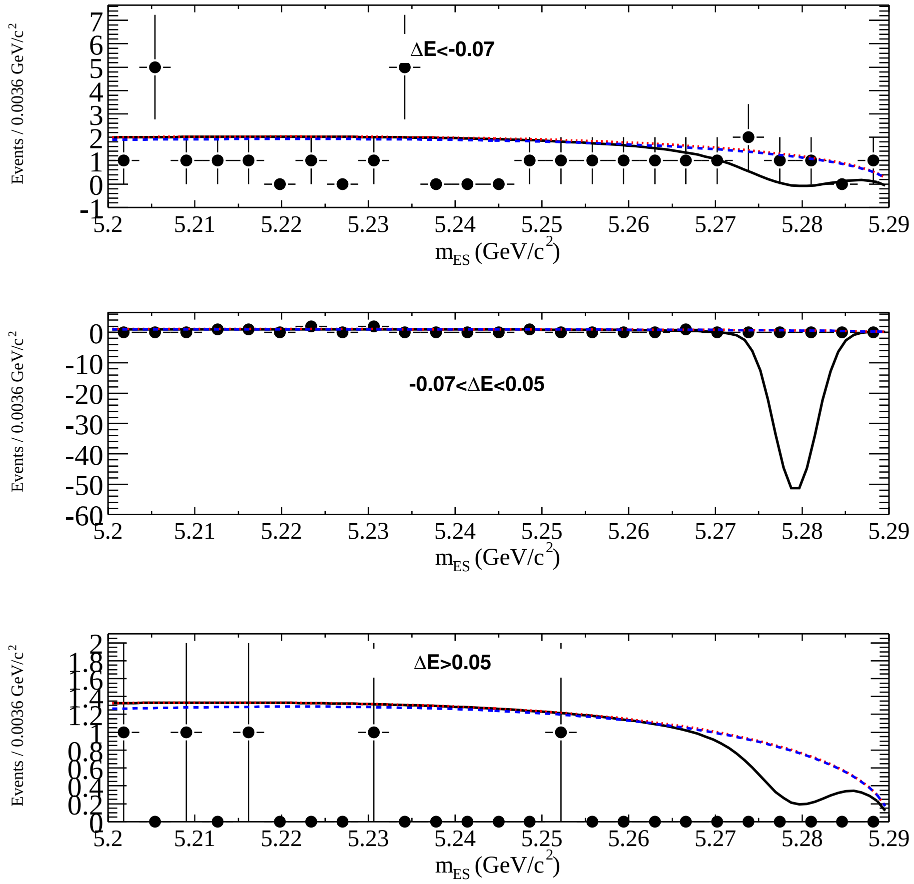


Figure 9.13: An example of a fit with a very large negative signal yield. The plots show the $m_{\rm ES}$ distributions for fits to a 20 fb^{-1} Pravda Monte Carlo background plus signal sample in the $B^+ \to K^+\mu^+\mu^-$ channel. The fit is displayed in 3 slices of ΔE: the upper plot is for $\Delta E <$ -0.07 GeV, the middle plot is for -0.07 $< \Delta E <$ 0.05 GeV, and bottom plot is for $\Delta E >$ 0.05 GeV. The middle plot thus corresponds to the signal region in ΔE. The points with error bars are the Monte Carlo. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied, with loose ΔE ($|\Delta E| <$ 0.25 GeV) and $m_{\rm ES}$ ($m_{\rm ES} >$ 5.2 GeV/c^2) cuts.

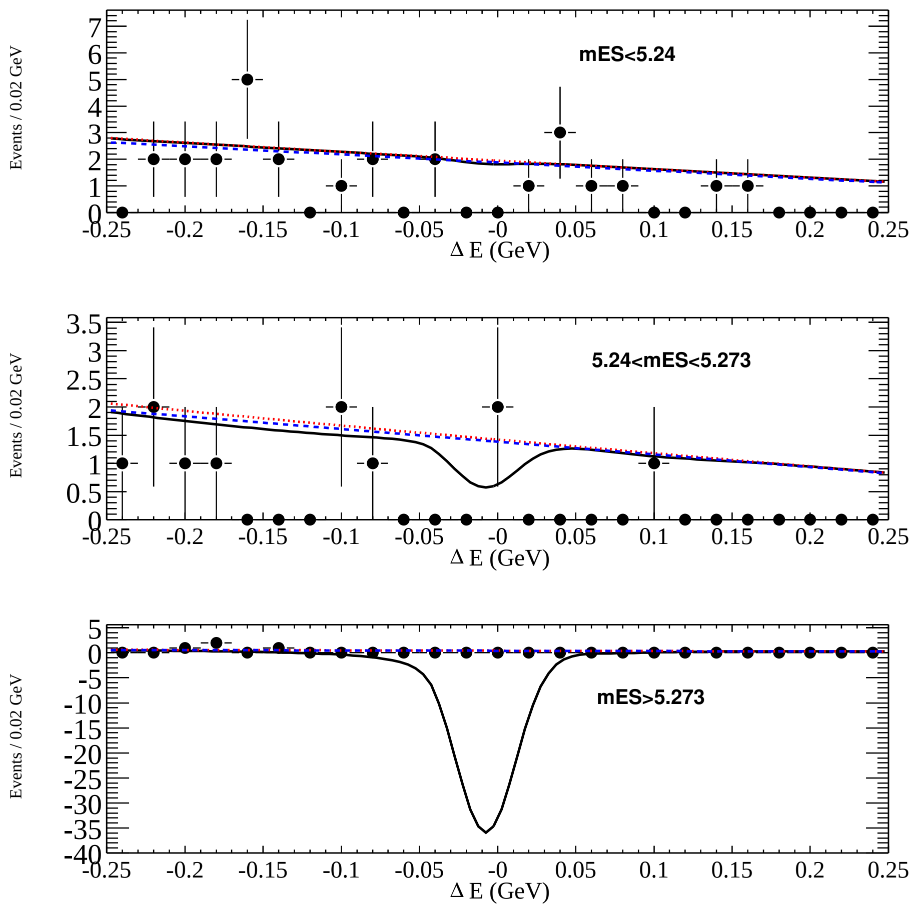


Figure 9.14: An example of a fit with a very large negative signal yield. The plots show the ΔE distributions for fits to a 20 fb^{-1} Pravda Monte Carlo background plus signal sample in the $B^+ \to K^+\mu^+\mu^-$ channel. The fit is displayed in 3 slices of $m_{\rm ES}$: the top plot is for $m_{\rm ES}$ < 5.24 GeV/c^2, the middle plot is for 5.24 < $m_{\rm ES}$ < 5.27 GeV/c^2, and bottom plot is for $m_{\rm ES}$ > 5.27 GeV/c^2. The bottom plot thus corresponds to the signal region in $m_{\rm ES}$. The points with error bars are the Monte Carlo. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied, with loose ΔE ($|\Delta E|$ < 0.25 GeV) and $m_{\rm ES}$ ($m_{\rm ES}$ > 5.2 GeV/c^2) cuts.

Mode	Signal yield
$B^+ \to K^+e^+e^-$	-1.5 $\pm$ 1.8
$B^0 \to K^0_S\ e^+e^-$	3.2 $\pm$ 2.5
$B^0 \to K^{*0}e^+e^-$	-1.2 $\pm$ 3.9
$B^+ \to K^{*+}e^+e^-$ $(K^{*+} \to K^0_S\pi^+)$	-3.4 $\pm$ 15.3
$B^+ \to K^{*+}e^+e^-$ $(K^{*+} \to K^+\pi^0)$	-13.4 $\pm$ 8.2
$B^+ \to K^+\mu^+\mu^-$	1.9 $\pm$ 2.6
$B^0 \to K^0_S\ \mu^+\mu^-$	-0.2 $\pm$ 1.2
$B^0 \to K^{*0}\mu^+\mu^-$	1.2 $\pm$ 3.7
$B^+ \to K^{*+}\mu^+\mu^-$ $(K^{*+} \to K^0_S\pi^+)$	0.3 $\pm$ 2.5
$B^+ \to K^{*+}\mu^+\mu^-$ $(K^{*+} \to K^+\pi^0)$	-3.1 $\pm$ 4.5

Table 9.4: The yields of signal events obtained from fits to the 200 fb^{-1} Pravda Monte Carlo sample consisting of only $B\overline{B}$ and continuum backgrounds. The error on the yield is statistical.

consistent with the expectations. The entire 200 fb^{-1} sample of Pravda Monte Carlo is used for this purpose.

If a sample of Pravda Monte Carlo includes only background events, then the fit is expected to yield the number of signal events consistent with zero. Table 9.4 summarizes the results of fits to the 200 fb^{-1} of Pravda Monte Carlo consisting only of background events. No statistically significant excess of signal events is observed.

It is also necessary to verify that the fitter yields the expected number of events in samples with the signal present. Such samples are created using SP3 signal Monte Carlo and 10 samples of background Pravda Monte Carlo events of 20 fb^{-1} each. Both signal and background events are selected by sampling a Poisson distribution with a mean μ equal to the desired number of events. The algorithm for generating random numbers distributed according to a Poisson distribution is described in Ref. [1]. The results from these fits are shown in Table 9.5 for the

	$K^+e^+e^-$	$K^+\mu^+\mu^-$	$K^{*0}e^+e^-$	$K^{*0}\mu^+\mu^-$
Fit #	$\mu = 2.4$	$\mu = 1.5$	$\mu = 4.1$	$\mu = 2.8$
1	1.7 ± 2.2	2.8 ± 2.4	0.2 ± 2.9	6.1 ± 3.2
2	4.2 ± 2.3	0.9 ± 1.2	4.7 ± 2.9	3.2 ± 2.4
3	-0.1 ± 0.7	1.8 ± 1.8	4.9 ± 3.0	4.2 ± 2.4
4	3.9 ± 2.3	1.2 ± 1.3	3.3 ± 2.5	7.2 ± 3.2
5	1.4 ± 1.5	-0.1 ± 0.6	1.3 ± 2.2	2.9 ± 2.2
6	0.8 ± 1.1	3.2 ± 2.0	6.9 ± 3.5	3.5 ± 2.4
7	-0.1 ± 0.6	-0.1 ± 1.3	-0.5 ± 1.2	3.0 ± 2.1
8	2.7 ± 1.8	-0.1 ± 0.6	2.2 ± 2.2	2.8 ± 2.1
9	-0.1 ± 0.6	2.9 ± 2.1	3.5 ± 2.6	-0.3 ± 0.9
10	5.0 ± 2.5	2.0 ± 1.9	8.5 ± 3.3	0.8 ± 1.7
Sum	19.4 ± 5.4	14.5 ± 5.2	35.0 ± 8.5	33.4 ± 7.4
# Sig.	23	13	41	26

Table 9.5: Signal yields from fits to ten 20 fb^{-1} Pravda Monte Carlo samples with a given number of signal and background events for the charged track modes. The errors are statistical only. μ is the mean number of signal events used in a Poisson distribution to select the number of signal events in each of the ten samples. The sum is the total yield obtained by adding the results of all 10 fits. The last row shows the total number of signal events drawn in each mode.

charged track modes and in Table 9.6 for the modes with a K^0_S or a π^0. In general, the signal yields obtained from the fitter are consistent with the expectations.

9.4 Fitting the $(K, K^*)e\mu$ Control Sample

Fitting the $(K, K^*)e\mu$ events in the data provides another check of the fitter performance. No signal is expected in this sample, and therefore the fitter should yield results consistent with zero.

20,000 signal $B \to (K, K^*)e\mu$ SP3 Monte Carlo events have been generated using the phase space model. This Monte Carlo is used to establish the signal shapes for $(K, K^*)e\mu$ fits. The yields of signal events obtained from these fits are summarized in Table 9.7, and the fits themselves are shown in Figs. 9.15-9.19. The yields in all

Fit #	$K^0_S e^+e^-$	$K^0_S \mu^+\mu^-$	$K^{*+}_{K^0_S} e^+e^-$	$K^{*+}_{K^0_S} \mu^+\mu^-$	$K^{*+}_{\pi^0} e^+e^-$	$K^{*+}_{\pi^0} \mu^+\mu^-$
	$\mu = 0.8$	$\mu = 0.4$	$\mu = 1.2$	$\mu = 0.6$	$\mu = 1.6$	$\mu = 0.8$
1	2.0 ± 1.4	2.5 ± 1.8	3.1 ± 2.6	2.7 ± 2.3	-2.4 ± 3.4	0.8 ± 2.1
2	1.0 ± 1.0	0.7 ± 1.0	2.0 ± 2.2	-0.3 ± 0.9	1.6 ± 4.2	1.4 ± 2.4
3	0.8 ± 1.1	-0.1 ± 0.6	5.2 ± 3.1	4.0 ± 2.7	-3.8 ± 1.1	2.3 ± 2.6
4	1.2 ± 1.4	-0.1 ± 0.6	0.7 ± 2.0	-0.3 ± 0.8	-2.7 ± 8.9	-0.1 ± 1.8
5	1.2 ± 1.2	1.0 ± 1.1	1.0 ± 2.2	1.6 ± 1.8	2.7 ± 4.2	-0.8 ± 2.6
6	0.0 ± 0.5	-0.1 ± 0.5	2.7 ± 2.6	-0.4 ± 0.9	5.5 ± 5.5	-0.1 ± 2.4
7	0.0 ± 0.5	1.1 ± 1.2	0.8 ± 2.1	-0.3 ± 0.7	3.6 ± 4.5	-0.8 ± 3.7
8	0.0 ± 0.5	-0.1 ± 0.6	1.3 ± 1.9	-0.6 ± 1.6	2.1 ± 5.1	-0.9 ± 1.5
9	0.0 ± 0.7	-0.1 ± 0.5	4.1 ± 2.6	-0.2 ± 1.0	-3.0 ± 3.4	-0.7 ± 2.2
10	-1.0 ± 0.5	0.9 ± 1.1	0.5 ± 1.3	-0.3 ± 0.7	3.8 ± 4.0	3.4 ± 3.2
Sum	6.0 ± 3.0	5.7 ± 3.1	21.4 ± 7.3	5.9 ± 4.8	7.4 ± 15.2	4.5 ± 8.0
# Sig.	7	4	17	10	12	13

Table 9.6: Signal yields from fits to ten 20 fb^{-1} Pravda Monte Carlo samples with a given number of signal and background events for the modes with a K^0_S or a π^0. $K^{*+}_{K^0_S}$ denotes the mode with $K^{*+} \to K^0_S\pi^+$, and $K^{*+}_{\pi^0}$, the mode with $K^{*+} \to K^+\pi^0$. The errors are statistical only. μ is the mean number of signal events used in a Poisson distribution to select the number of signal events in each of the ten samples. The sum is the total yield obtained by adding the results of all 10 fits. The last row shows the total number of signal events drawn in each mode.

Mode	Signal yield
$K^+e^\pm\mu^\mp$	-0.6 ± 1.0
$K^0_S e^\pm\mu^\mp$	0.6 ± 2.8
$K^{*0}e^\pm\mu^\mp$	0.8 ± 1.7
$K^{*+}e^\pm\mu^\mp$ ($K^{*+} \to K^0_S\pi^+$)	-0.4 ± 0.8
$K^{*+}e^\pm\mu^\mp$ ($K^{*+} \to K^+\pi^0$)	-2.6 ± 8.3

Table 9.7: Signal yields from the fits to the $(K, K^*)e\mu$ control samples in data. The errors are statistical only.

the channels are consistent with zero.

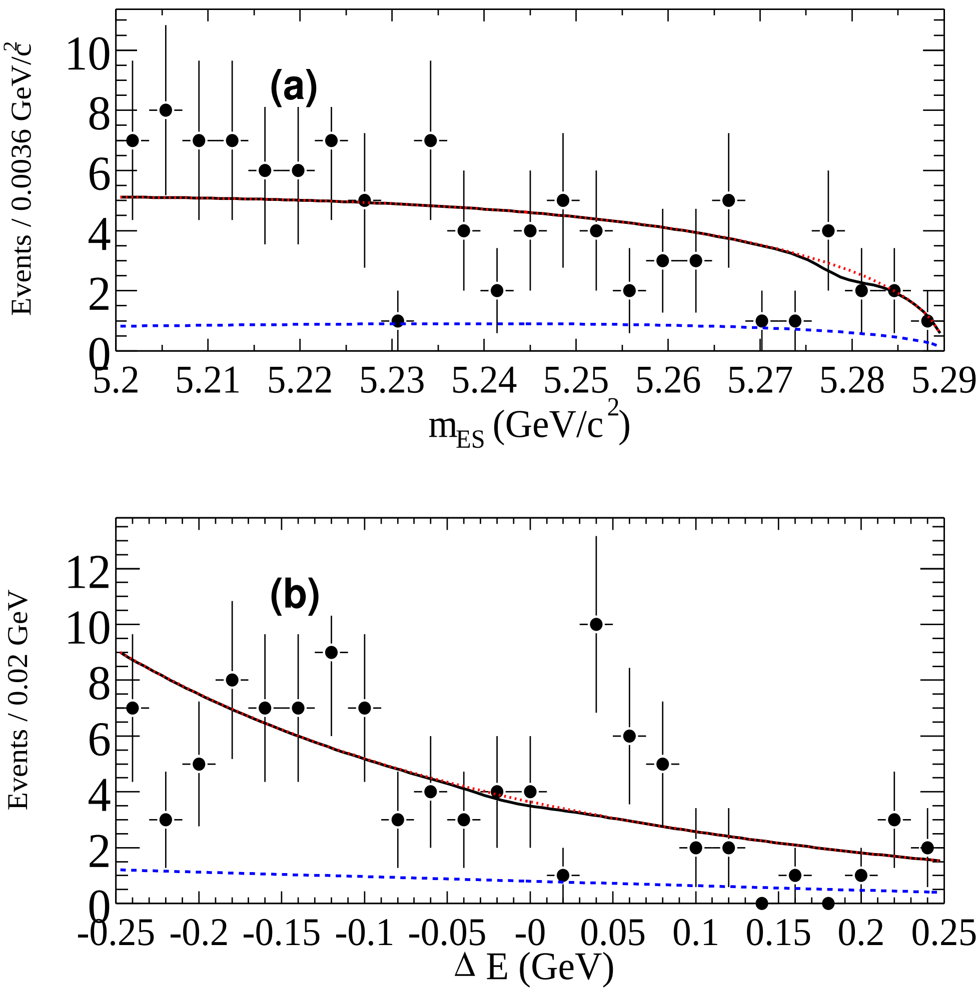


Figure 9.15: (a) $m_{\rm ES}$ and (b) $\mathit{\Delta}E$ distributions for fits to the data in the $Ke\mu$ sample. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied, with loose $\mathit{\Delta}E$ ($|\mathit{\Delta}E| < 0.25$ GeV) and $m_{\rm ES}$ ($m_{\rm ES} > 5.2$ GeV/c^2) cuts.

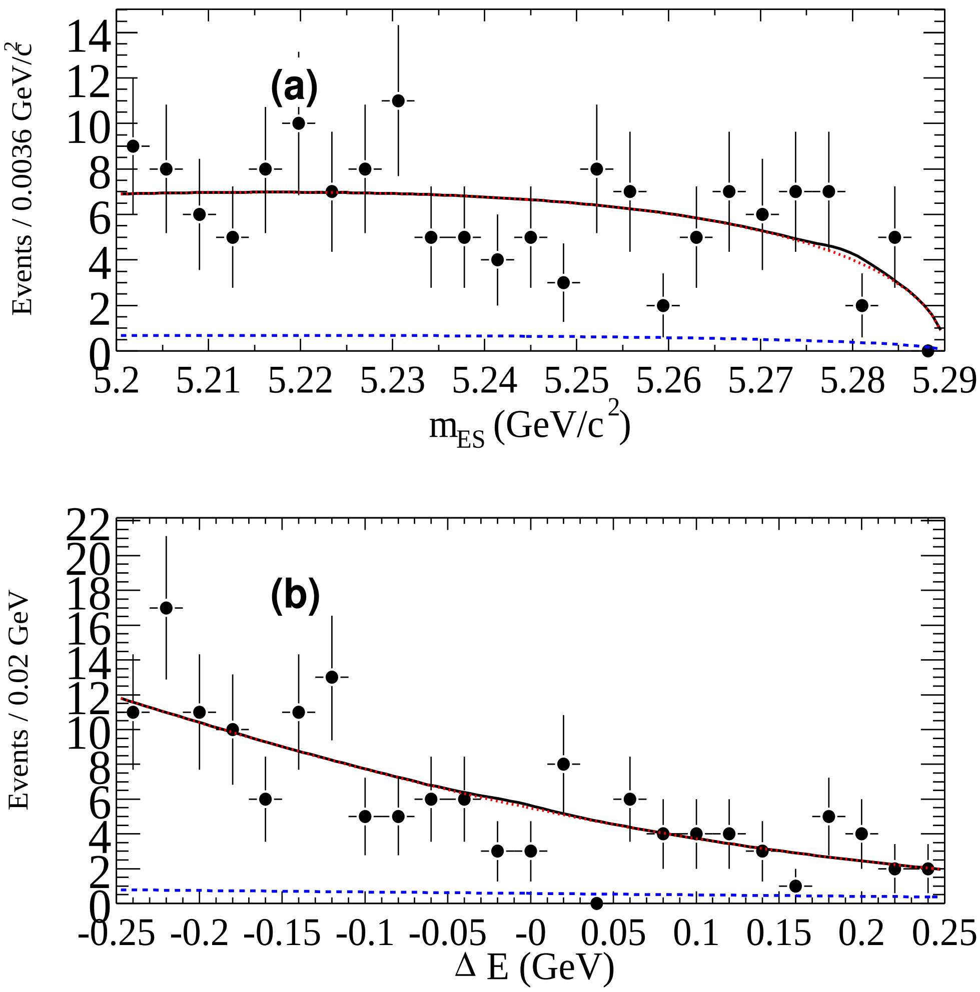


Figure 9.16: (a) $m_{\rm ES}$ and (b) ΔE distributions for fits to the data in the $K^{*0}e\mu$ sample. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied, with loose ΔE ($|\Delta E| < 0.25$ GeV) and $m_{\rm ES}$ ($m_{\rm ES} > 5.2$ GeV/c^2) cuts.

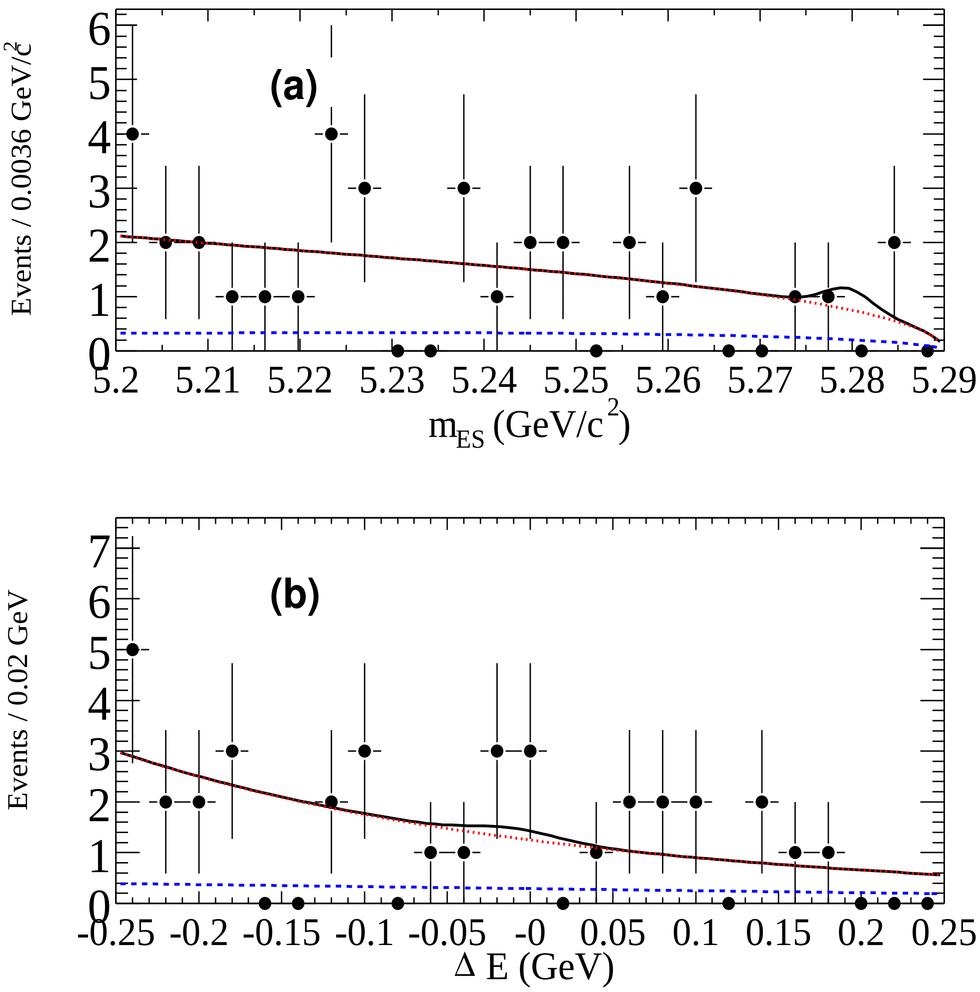


Figure 9.17: (a) $m_{\rm ES}$ and (b) ΔE distributions for fits to the data in the $K^0_S e\mu$ sample. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied, with loose ΔE ($|\Delta E| < 0.25$ GeV) and $m_{\rm ES}$ ($m_{\rm ES} > 5.2$ GeV/c^2) cuts.

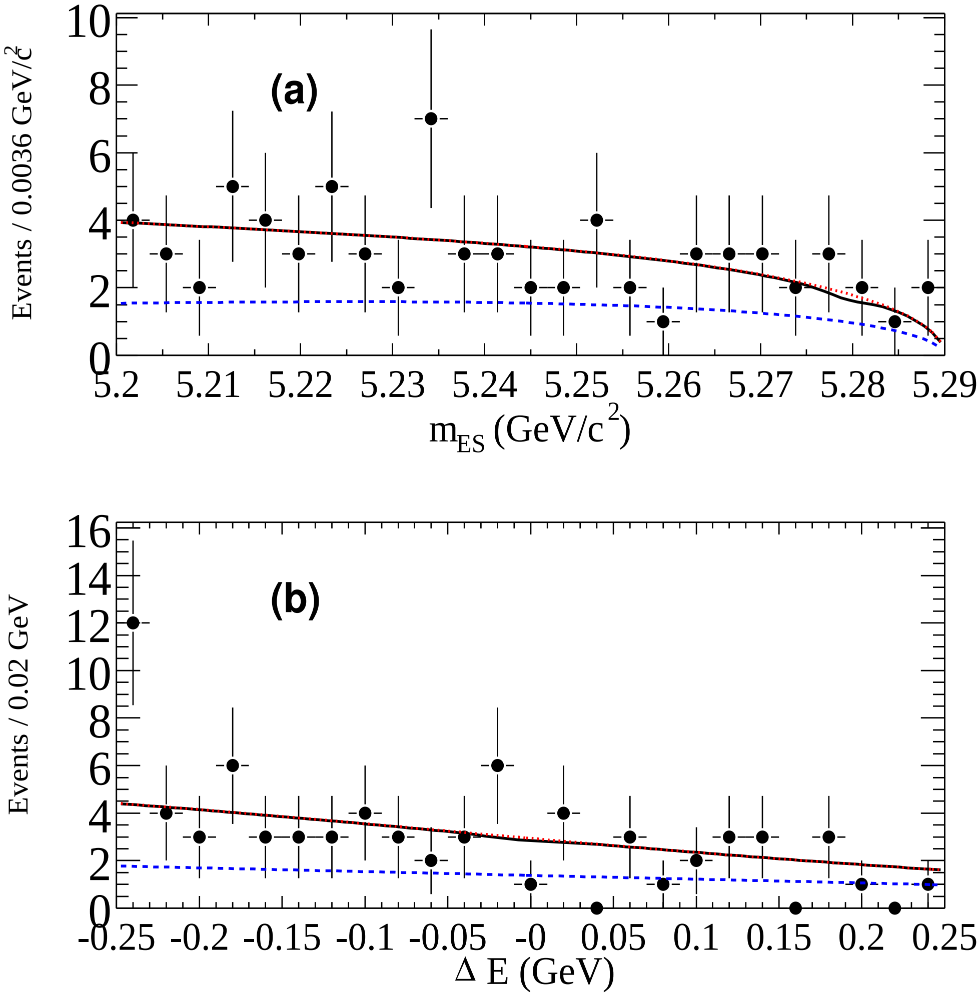


Figure 9.18: (a) m_{ES} and (b) ΔE distributions for fits to the data in the $K^{*+}e\mu$ ($K^{*+} \to K^0_S\pi^+$) sample. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied, with loose ΔE ($|\Delta E| < 0.25$ GeV) and m_{ES} ($m_{\mathrm{ES}} > 5.2$ GeV/c^2) cuts.

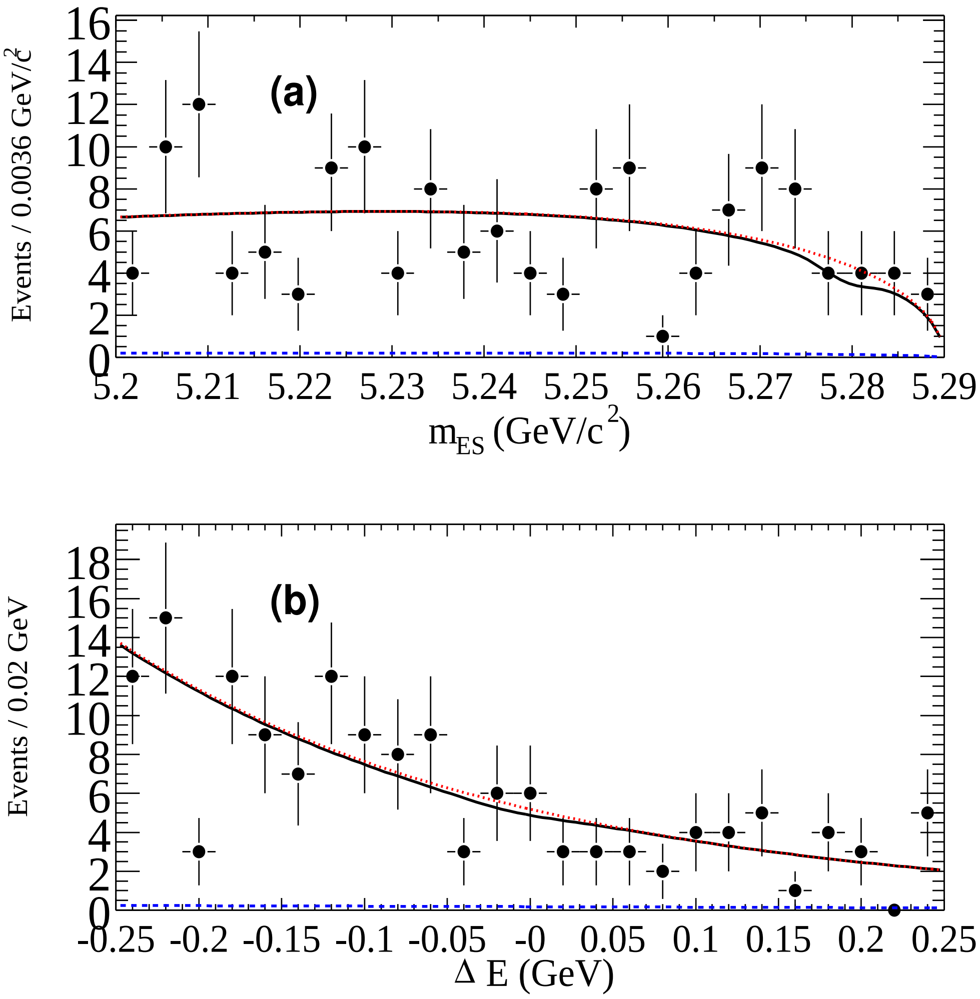


Figure 9.19: (a) m_{ES} and (b) ΔE distributions for fits to the data in the $K^{*+}e\mu$ ($K^{*+} \rightarrow K^+\pi^0$) sample. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied, with loose ΔE ($|\Delta E| < 0.25$ GeV) and m_{ES} ($m_{\mathrm{ES}} > 5.2$ GeV$/c^2$) cuts.

Chapter 10

Results

10.1 The Procedure for Setting Upper Limits

In this analysis, the following procedure is used to establish 90% C.L. upper limits. Let us suppose that the number of observed signal events, given by the fit described in Chapter 9, is n_0. Generating a very large number of toy Monte Carlo samples according to the fit PDF (Eqn. 9.0.1), with a varying mean number of signal events, one can find the mean number of input signal events n_1 such that 90% of the cases yield a signal greater than n_0. The yield n_1 is called the *90% C.L. signal event yield* and is used to set a 90% C.L. upper limit on the branching fraction in the following way:

$$\frac{n_1}{N_{B\overline{B}}\ \mathcal{B}\ \epsilon}, \tag{10.1.1}$$

where $N_{B\overline{B}}$ is the number of produced $B\overline{B}$ events in the data sample used, $\mathcal{B}$ is the product branching fraction of the secondary decays (e.g., that of $K^{*0} \to K^+\pi^-$ for the $B^0 \to K^{*0}e^+e^-$ mode), and ϵ is the signal efficiency.

The number of input signal events in the toy Monte Carlo fits is drawn from a Poisson distribution with the mean equal to the number of events observed in the data. If a negative number was observed, the input number of signal events

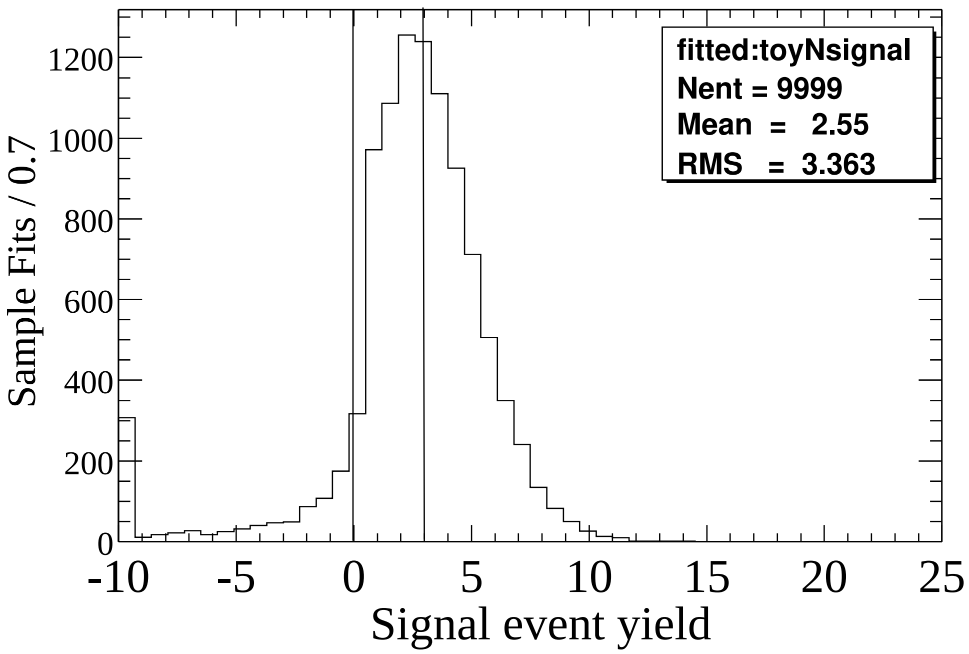


Figure 10.1: An illustration of the procedure used to find the 90% C.L. event yield for establishing the 90% C.L. upper limit for the $B^+ \to K^+e^+e^-$ data sample. The fit yielded a negative signal yield, but $n_0 = 0$ is used as the number of observed events. The plot shows the results of 10,000 toy Monte Carlo experiments with the input mean number of signal events $n_1 = 3.0$. For this n_1, roughly 90% of the cases yield a signal greater than zero, establishing 3.0 as the 90% C.L. event yield.

is taken to be zero. This corresponds to setting a Bayesian upper limit with a flat prior [77]. The procedure is illustrated in Fig. 10.1, which shows the results of 10,000 toy Monte Carlo experiments generating the $B^+ \to K^+e^+e^-$ fit PDF. As will be discussed in Section 10.2, a negative number of signal events was obtained for the $B^+ \to K^+e^+e^-$ channel after unblinding the data; therefore, in accordance with the procedure described above, zero events is used as the observed signal yield.

Mode	MC bkg	Data bkg	$B\overline{B}$ bkg	Cont bkg
$B^+ \to K^+e^+e^-$	1.0	1.0	0.8	1.2
$B^0 \to K^0_S e^+e^-$	0.6	0.3	0.2	0.3
$B^0 \to K^{*0}e^+e^-$	2.7	2.0	2.0	2.6
$B^+ \to K^{*+}e^+e^-$ $(K^{*+} \to K^0_S\pi^+)$	1.2	1.5	1.6	1.3
$B^+ \to K^+\mu^+\mu^-$	0.9	1.1	0.9	1.2
$B^0 \to K^0_S\mu^+\mu^-$	0.3	0.2	0.2	0.3
$B^0 \to K^{*0}\mu^+\mu^-$	1.4	1.5	1.3	1.8
$B^+ \to K^{*+}\mu^+\mu^-$ $(K^{*+} \to K^0_S\pi^+)$	1.0	0.8	0.8	0.9
TOTAL	9.1	8.4	7.8	9.6

Table 10.1: Comparison of expected and observed background levels in the signal region. The MC background is the background level predicted from the Monte Carlo (Table 7.30). The data background is the number of background events obtained by fitting region B in data and integrating the total fit shape over the signal region. $B\overline{B}$ and continuum backgrounds are estimated using only the $B\overline{B}$ or continuum shape in the fit to the data.

10.2 Results of Unblinding

Unblinding the data is a two-step procedure. As a first step, the nearest sideband (region B) is unblinded for the four charged track modes and the modes with a K^0_S. The modes with a π^0 remain blinded since it is felt that the event selection criteria for these modes may be improved further. Extrapolating the background shapes from region B into the signal region, one can estimate the background level in the signal region and compare it to the predictions from the Monte Carlo (which are given in Table 7.30). The results of this comparison are summarized in Table 10.1.

The purpose of this comparison is to make sure that the backgrounds in data are not significantly larger from the Monte Carlo predictions. If they were, this would indicate the presence of a background not considered in the simulation, and it would have to be understood before unblinding the signal region. Since the agreement between predicted and observed background levels is quite good, one can proceed to the second step of the unblinding procedure and unblind the signal region (region A).

Figures 10.2 and 10.4 show ΔE vs. m_{ES} distributions for the four charged track modes and the modes with a K^0_S, respectively, after unblinding both signal and near sideband regions. No events are observed in the signal box in the $B^+ \to K^+e^+e^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ modes. There is a single event observed in $B^+ \to K^+\mu^+\mu^-$; however, it is located so close to the boundary of the signal box as to make it likely that it is in fact a background event. Six events passed all of the analysis selection criteria in the $B^0 \to K^{*0}e^+e^-$ mode. However, a significant fraction of these events resides in a low ΔE region. Figure 10.3 shows the dilepton mass distributions for these events. It is evident that they are not all concentrated very close to the charmonium resonance or at zero (which would suggest they come from γ conversions or π^0 Dalitz decays). There are two events in the signal box in the $B^0 \to K^0_S e^+e^-$ channel; zero in the $B^0 \to K^0_S\mu^+\mu^-$ channel; two in the $B^+ \to K^{*+}e^+e^-$ ($K^{*+} \to K^0_S\pi^+$) mode, both of which have rather low ΔE and are very close to the boundary of the signal box; and one event in the $B^+ \to K^{*+}\mu^+\mu^-$ ($K^{*+} \to K^0_S\pi^+$) mode.

There are thus a total of 12 events in the signal boxes in all of the modes, and Table 10.1 shows that the number of background events in the signal region, obtained by using the fit to the data, is 8. Table 7.30 shows that the expected number of signal events in the signal box is 10. Thus, the expected number of events is 18, and the observed number of events (12 ± 3.5) is within approximately 2σ of the expectations. However, in this analysis the number of signal events is determined not by counting the number of events in the signal box, but by the unbinned maximum likelihood fit described in Chapter 9. Figures 10.5–10.28 show the results of the fit to the data. This signal yield is given in Table 10.2, which also summarizes the 90% C.L. signal yields obtained using the procedure described in Section 10.1; the efficiencies for signal events to lie in the fit region; the relative systematic error

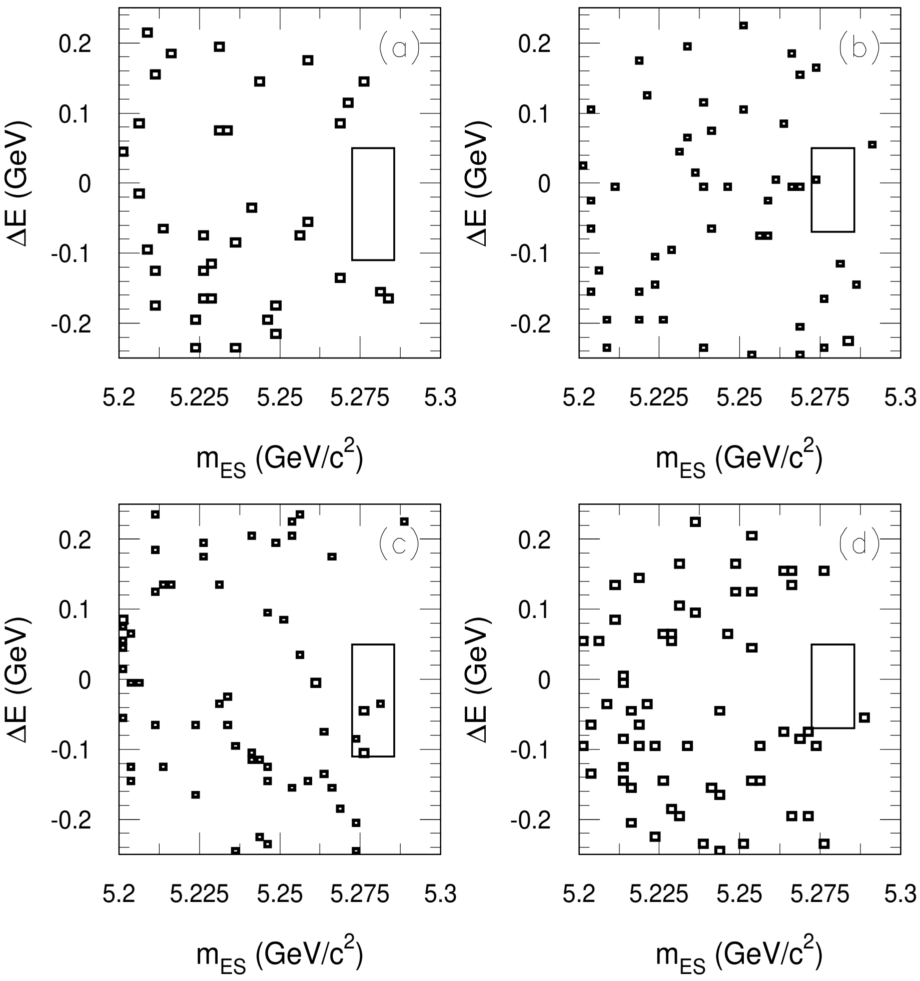


Figure 10.2: The results of unblinding the signal and near sideband regions for the four charged track modes. (a) $B^+ \to K^+e^+e^-$, (b) $B^+ \to K^+\mu^+\mu^-$, (c) $B^0 \to K^{*0}e^+e^-$, and (d) $B^0 \to K^{*0}\mu^+\mu^-$. All of the event selection criteria have been applied. The smaller boxes shown in the plots correspond to the signal region (region A).

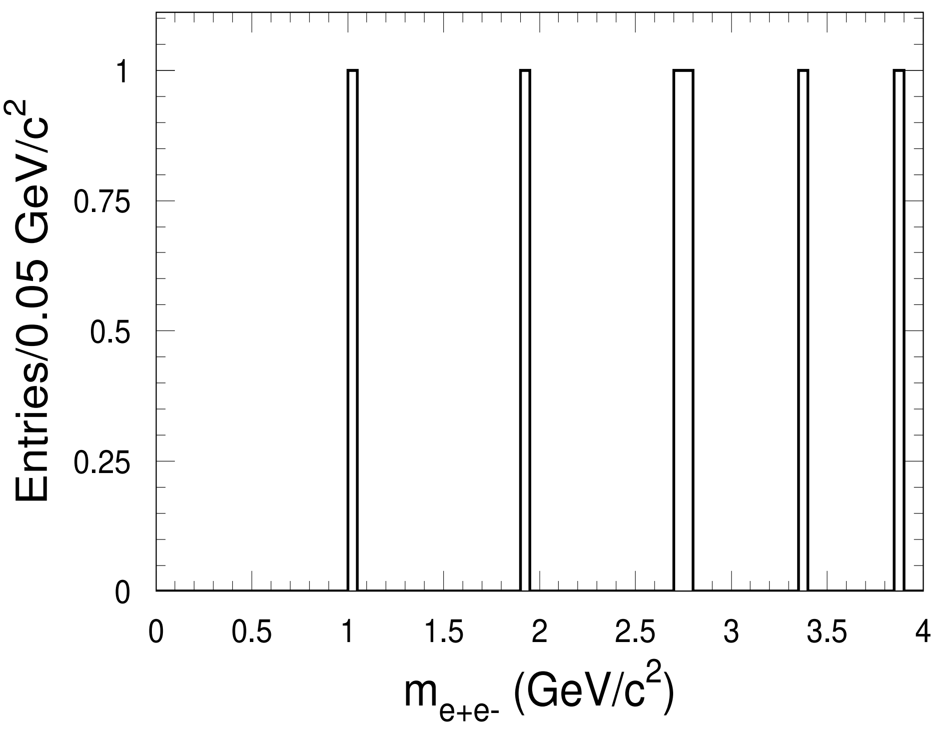


Figure 10.3: The dilepton mass distribution for the six data events passing all of the $B^0 \to K^{*0}e^+e^-$ analysis selection criteria.

on the signal efficiency and on the signal yield; and the 90% C.L. upper limits on the branching fractions. The systematic errors are discussed in detail in Chapter 11. Note that since the $(K, K^*)e\mu$ sample in the data has only been used for cross-checks, not for cut tuning, we can set 90% C.L. upper limits on the branching fractions for $B \to (K, K^*)e^{\pm}\mu^{\mp}$. These are also summarized in Table 10.2.

Figures 10.29-10.34 show event displays for the six events in the $B^0 \to K^{*0}e^+e^-$ signal box. Figures 10.35 and 10.36 show the event display for the two $B^0 \to K^0_S e^+e^-$ candidate events in the signal box. Figures 10.37 and 10.38 show the event displays for the two events in the $B^+ \to K^{*+}e^+e^-$ ($K^{*+} \to K^0_S\pi^+$) signal box. Finally, Fig. 10.39 shows the event display for the single $B^+ \to K^{*+}\mu^+\mu^-$ ($K^{*+} \to K^0_S\pi^+$) event in the signal box. All of the event displays look like ordinary $B\overline{B}$ events. The event shown in Fig. 10.33 may suggest a jet-like continuum event, however.

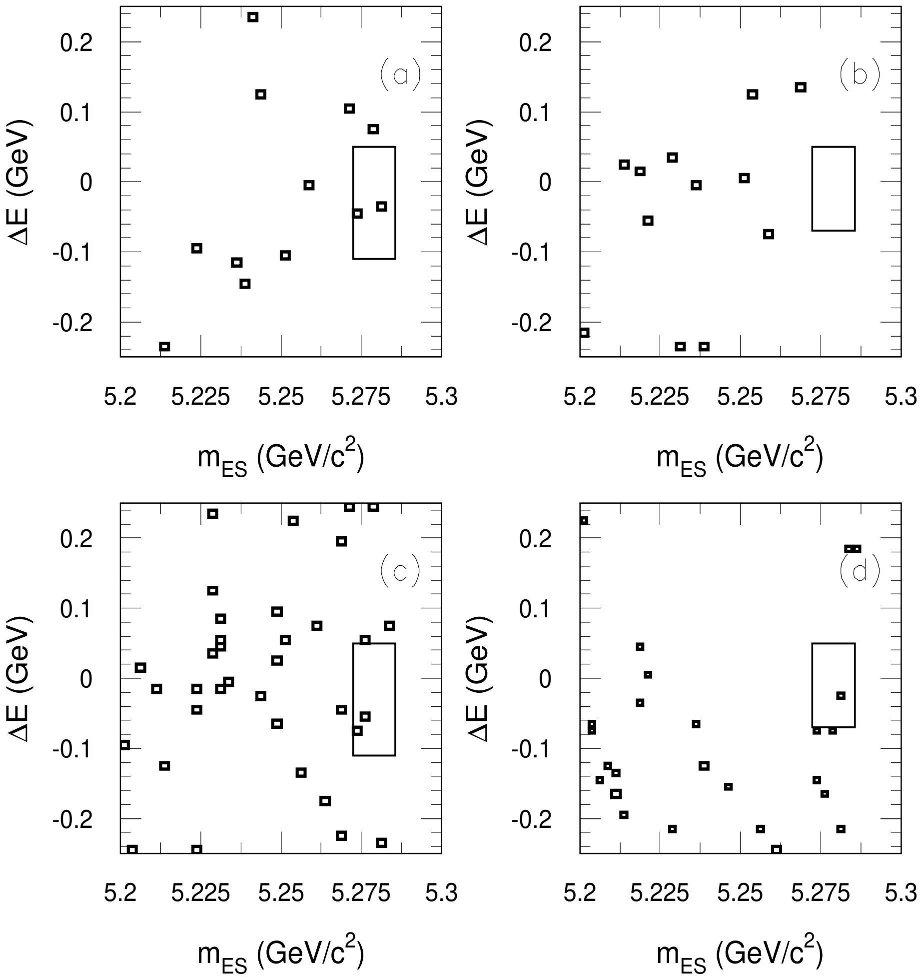


Figure 10.4: The results of unblinding the signal and near sideband regions for the modes with a K^0_S. (a) $B^0 \to K^0_S e^+e^-$, (b) $B^0 \to K^0_S \mu^+\mu^-$, (c) $B^+ \to K^{*+}e^+e^-$ ($K^{*+} \to K^0_S\pi^+$), and (d) $B^+ \to K^{*+}\mu^+\mu^-$ ($K^{*+} \to K^0_S\pi^+$). All of the event selection criteria have been applied. The smaller boxes shown in the plots correspond to the signal region (region A); however, the number of events in the signal boxes is not the signal yield.

Mode	Signal yield	90% C.L. yield	ϵ (%)	$(\Delta\mathcal{B}/\mathcal{B})_{\epsilon}$ (%)	$(\Delta\mathcal{B}/\mathcal{B})_{\rm yield}$ (%)	$\mathcal{B}/10^{-6}$	$\mathcal{B}/10^{-6}$ (90% C.L.)
$B^+ \to K^+e^+e^-$	-0.2	3.0	17.5	$\pm$8.6	$\pm$10.6	-0.1	0.9
$B^+ \to K^+\mu^+\mu^-$	-0.2	2.8	10.5	$\pm$8.6	$\pm$10.6	-0.1	1.3
$B^0 \to K^{*0}e^+e^-$	2.5	6.7	10.2	$\pm$10.5	$\pm$10.6	1.6	5.0
$B^0 \to K^{*0}\mu^+\mu^-$	-0.3	3.6	8.0	$\pm$10.8	$\pm$10.6	-0.2	3.6
$B^0 \to K^0e^+e^-$	1.3	5.0	15.7	$\pm$9.3	$\pm$10.6	1.1	4.7
$B^0 \to K^0\mu^+\mu^-$	0.0	2.9	9.6	$\pm$11.4	$\pm$10.6	0.0	4.5
$B^+ \to K^{*+}e^+e^-$	0.1	3.8	8.5	$\pm$11.4	$\pm$10.6	0.1	10.0
$B^+ \to K^{*+}\mu^+\mu^-$	1.0	4.3	5.8	$\pm$12.0	$\pm$10.6	3.3	17.5
$B^+ \to K^+e^\pm\mu^\mp$	-0.6	3.3	16.8	$\pm$8.6	$\pm$10.6	-0.2	1.0
$B^0 \to K^{*0}e^\pm\mu^\mp$	0.6	4.5	11.9	$\pm$10.6	$\pm$10.6	0.3	2.7
$B^0 \to K^0e^\pm\mu^\mp$	0.8	3.3	14.6	$\pm$10.3	$\pm$10.6	0.7	3.3
$B^+ \to K^{*+}e^\pm\mu^\mp$	-0.4	3.7	9.3	$\pm$11.7	$\pm$10.6	-0.8	8.7

Table 10.2: The results of the fits to the unblinded data. The yield refers to the number of signal events given by the fit. ϵ is the signal efficiency, defined as the efficiency for the signal events to be in the fit region. It does not include the branching fractions of the secondary decays. $(\Delta\mathcal{B}/\mathcal{B})_{\epsilon}$ is the relative systematic uncertainty on the signal efficiency, and $(\Delta\mathcal{B}/\mathcal{B})_{\rm yield}$, on the signal yield. The last column lists 90% C.L. upper limits on the individual branching fractions.

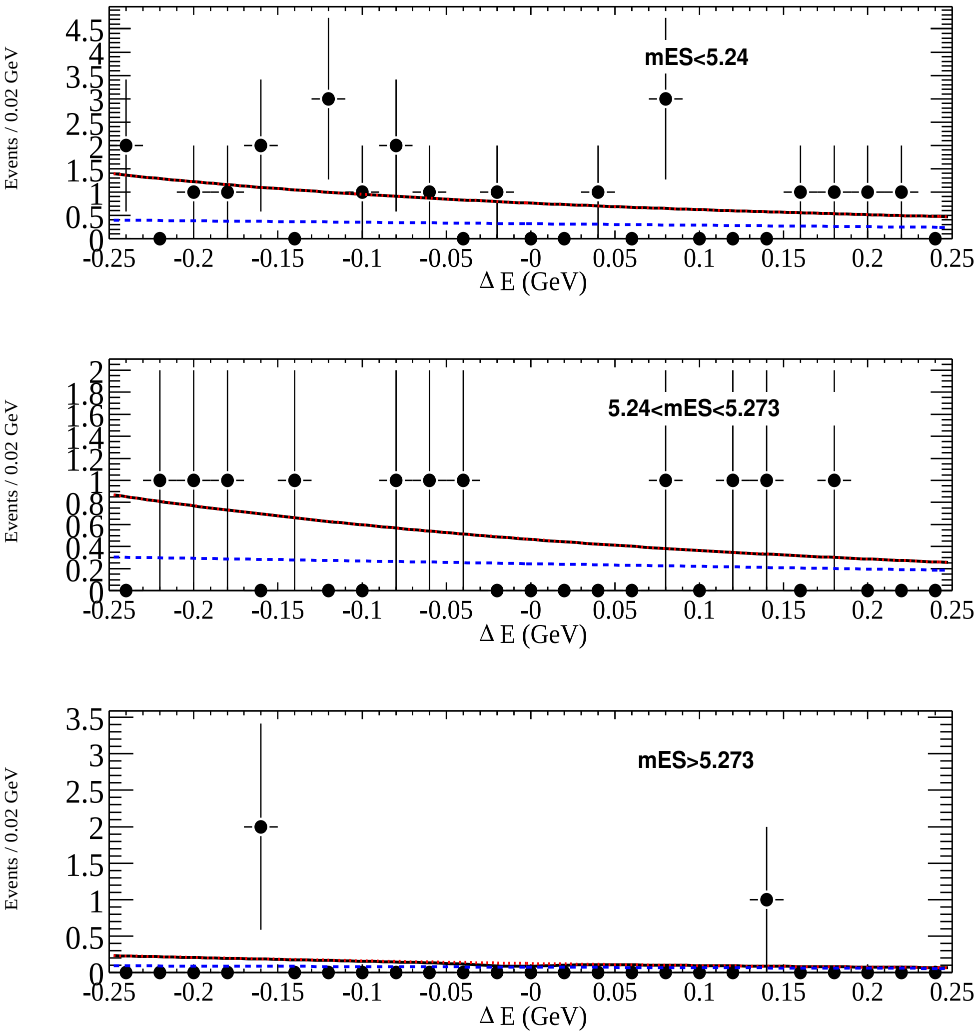


Figure 10.5: ΔE distributions for fits to all of the data in the $B^+ \to K^+e^+e^-$ mode. The fit is displayed in 3 slices of m_{ES}: the upper plot is for $m_{\mathrm{ES}} < 5.24$ GeV/c^2, the middle plot is for $5.24 < m_{\mathrm{ES}} < 5.27$ GeV/c^2, and the bottom plot is for $m_{\mathrm{ES}} > 5.27$ GeV/c^2. The bottom plot thus corresponds to the signal region in m_{ES}. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied.

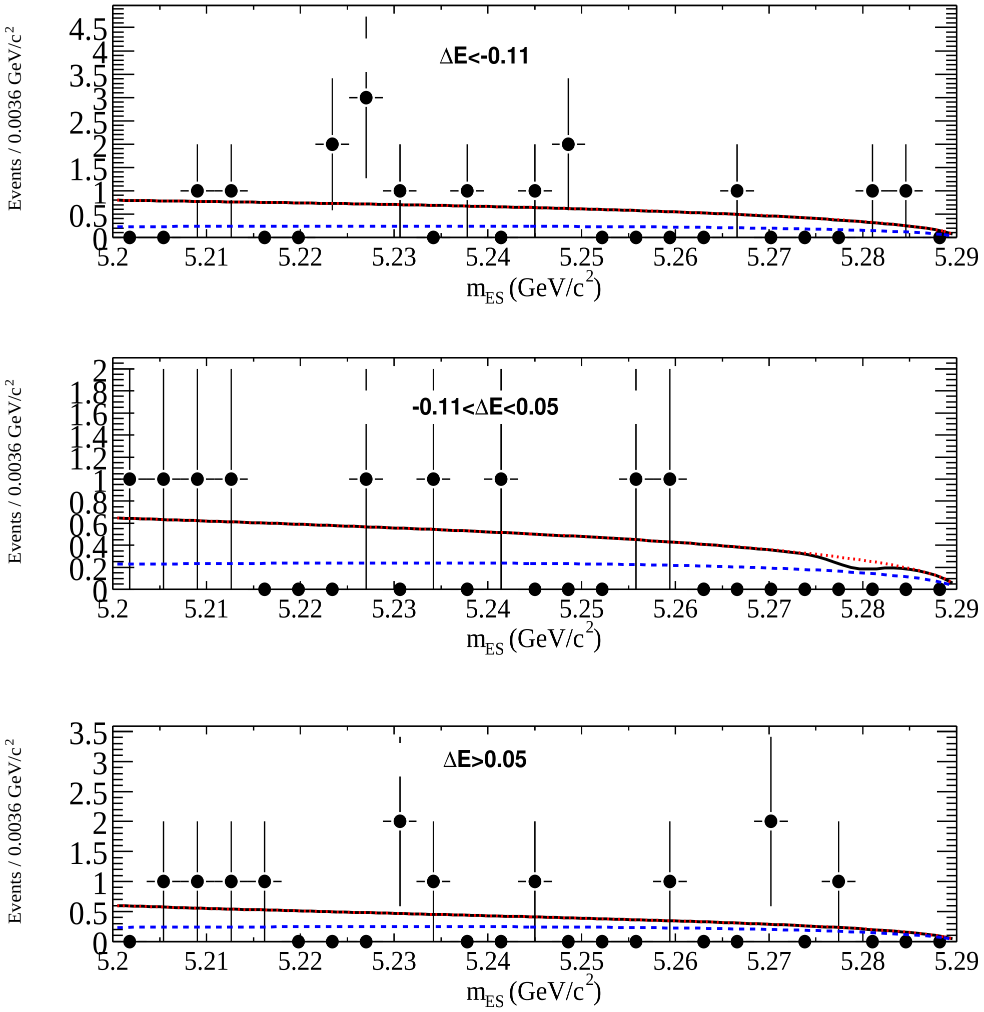
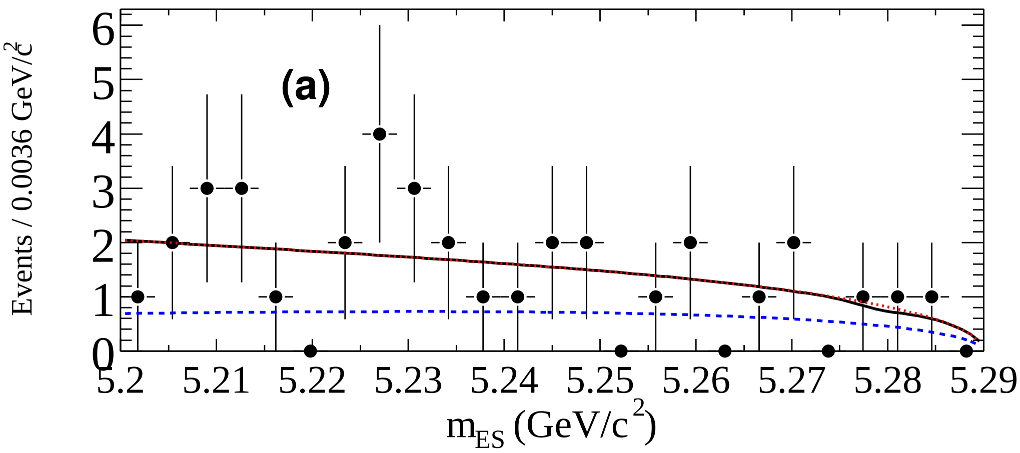


Figure 10.6: m_{ES} distributions for fits to all of the data in the $B^+ \to K^+e^+e^-$ mode. The fit is displayed in 3 slices of ΔE: the upper plot is for $\Delta E <$ -0.11 GeV, the middle plot is for -0.11 $< \Delta E <$ 0.05 GeV, and the bottom plot is for $\Delta E >$ 0.05 GeV. The middle plot thus corresponds to the signal region in ΔE. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied.



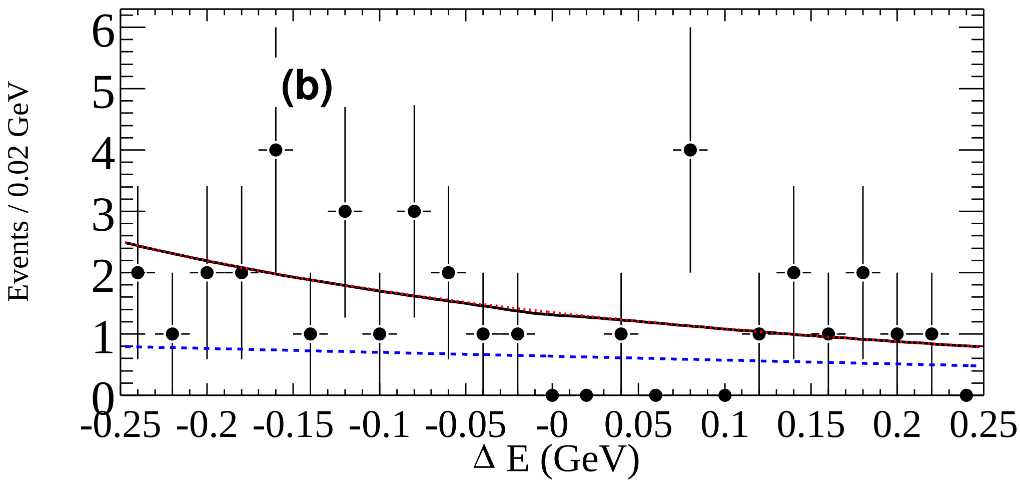


Figure 10.7: (a): m_{ES} and (b): ΔE distributions for fits to all of the data in the $B^+ \to K^+e^+e^-$ mode. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied.

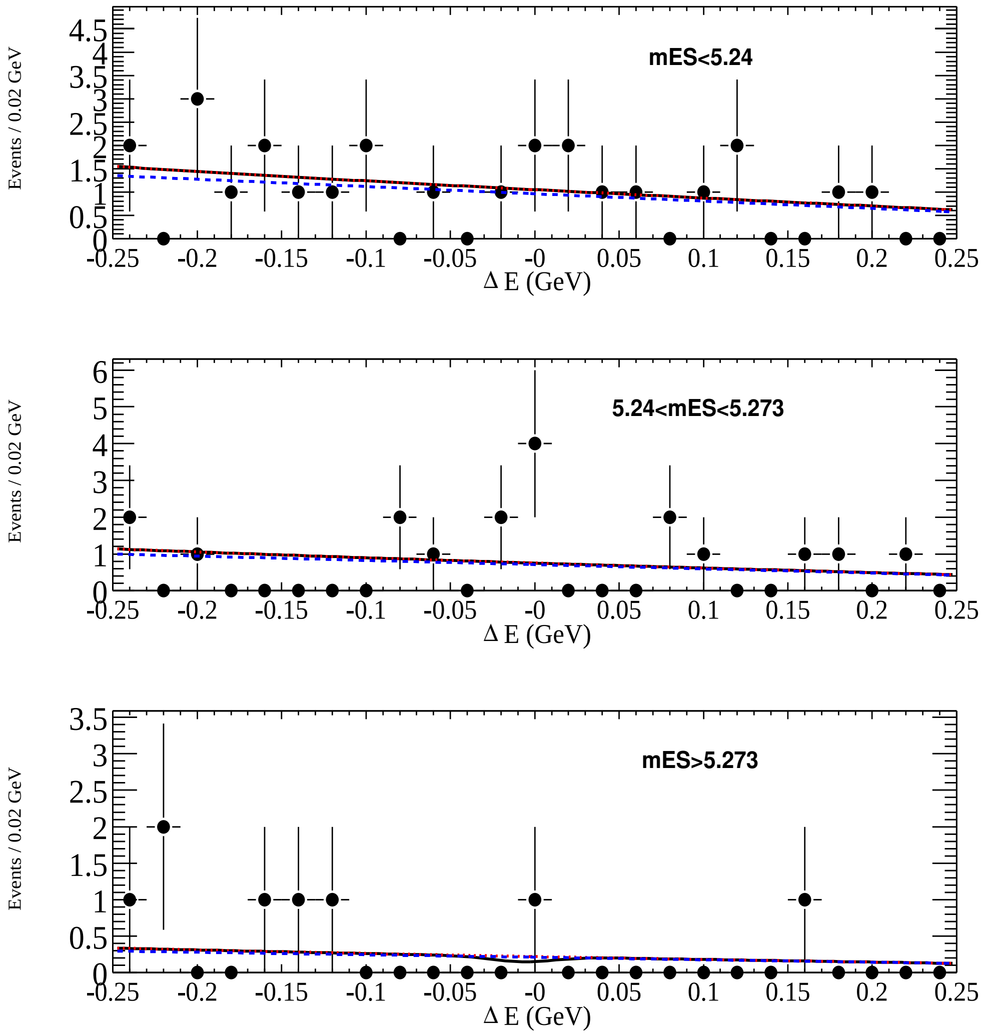


Figure 10.8: ΔE distributions for fits to all of the data in the $B^+ \to K^+\mu^+\mu^-$ mode. The fit is displayed in 3 slices of $m_{\rm ES}$: the upper plot is for $m_{\rm ES} < 5.24$ GeV/c^2, the middle plot is for $5.24 < m_{\rm ES} < 5.27$ GeV/c^2, and the bottom plot is for $m_{\rm ES} > 5.27$ GeV/c^2. The bottom plot thus corresponds to the signal region in $m_{\rm ES}$. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied.

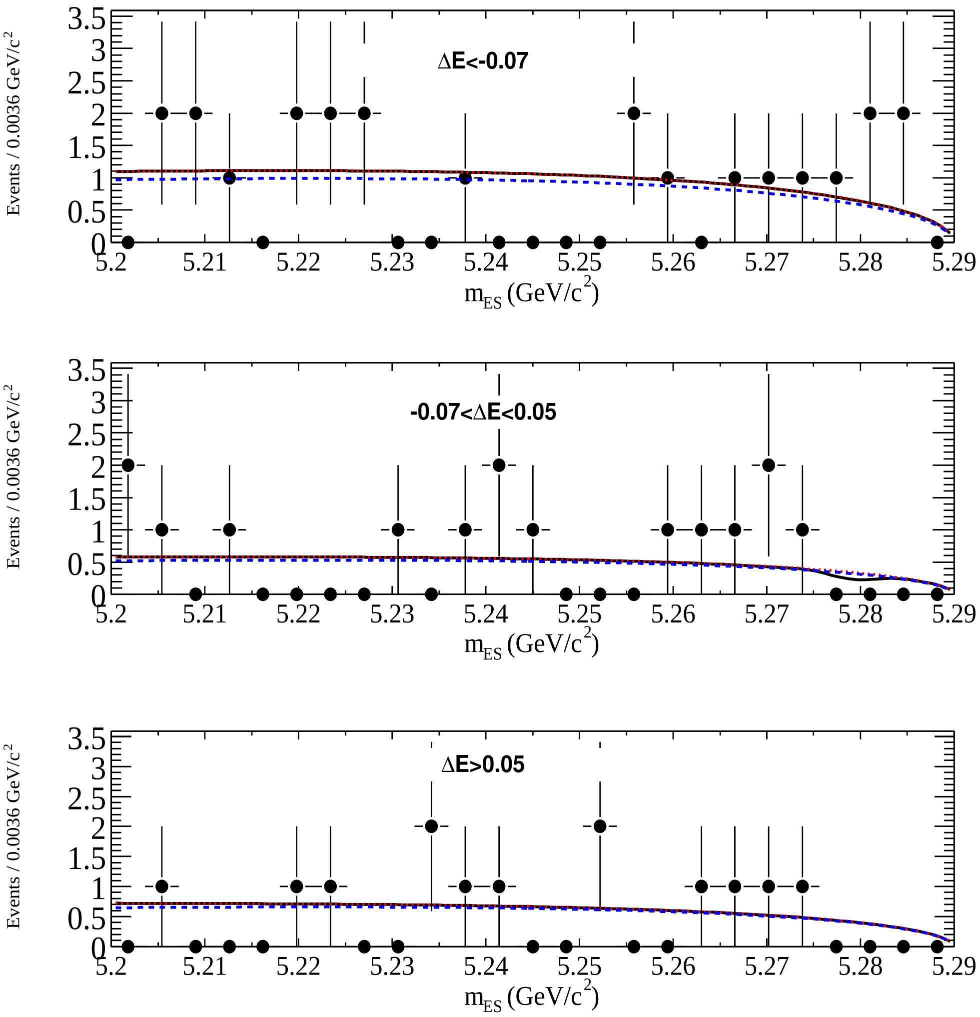


Figure 10.9: m_{ES} distributions for fits to all of the data in the $B^+ \to K^+\mu^+\mu^-$ mode. The fit is displayed in 3 slices of ΔE: the upper plot is for $\Delta E <$ -0.07 GeV, the middle plot is for -0.07 $< \Delta E <$ 0.05 GeV, and the bottom plot is for $\Delta E >$ 0.05 GeV. The middle plot thus corresponds to the signal region in ΔE. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied.

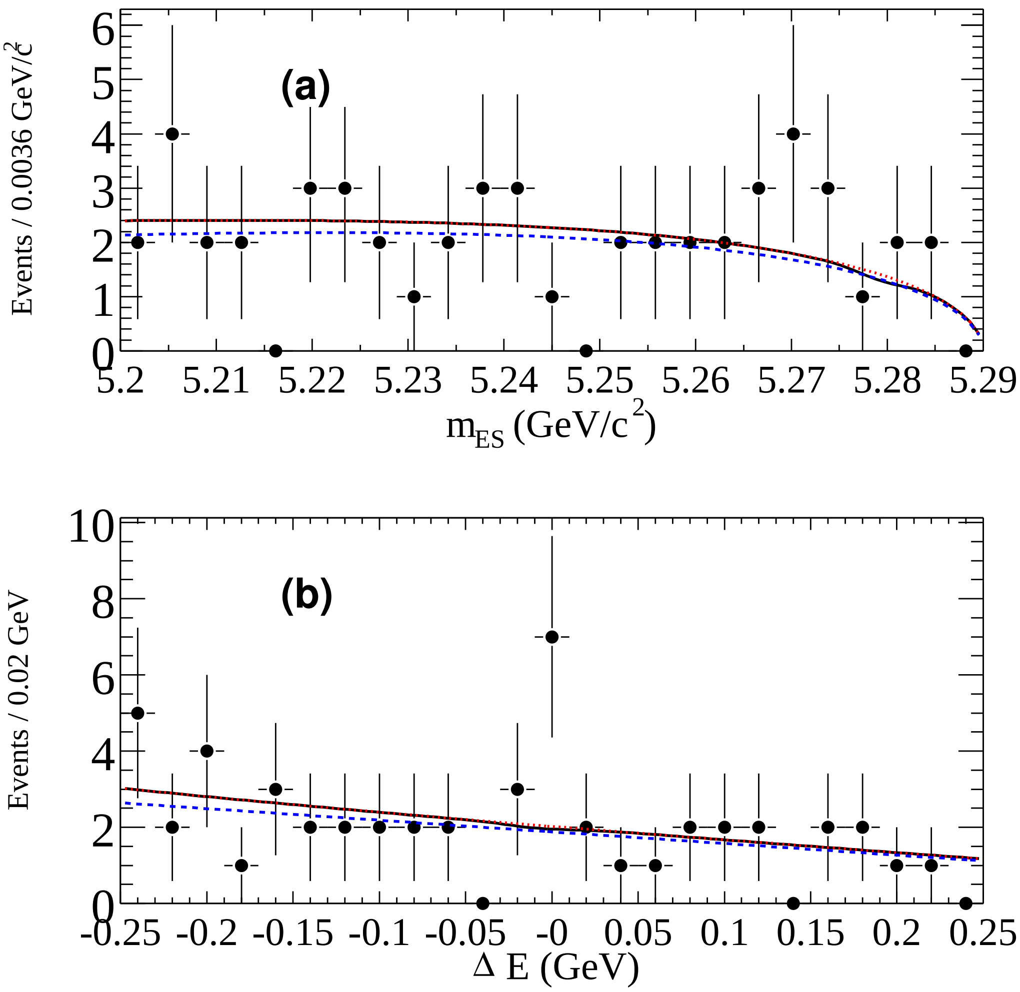


Figure 10.10: (a): m_{ES} and (b): ΔE distributions for fits to all of the data in the $B^+ \to K^+\mu^+\mu^-$ mode. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied.

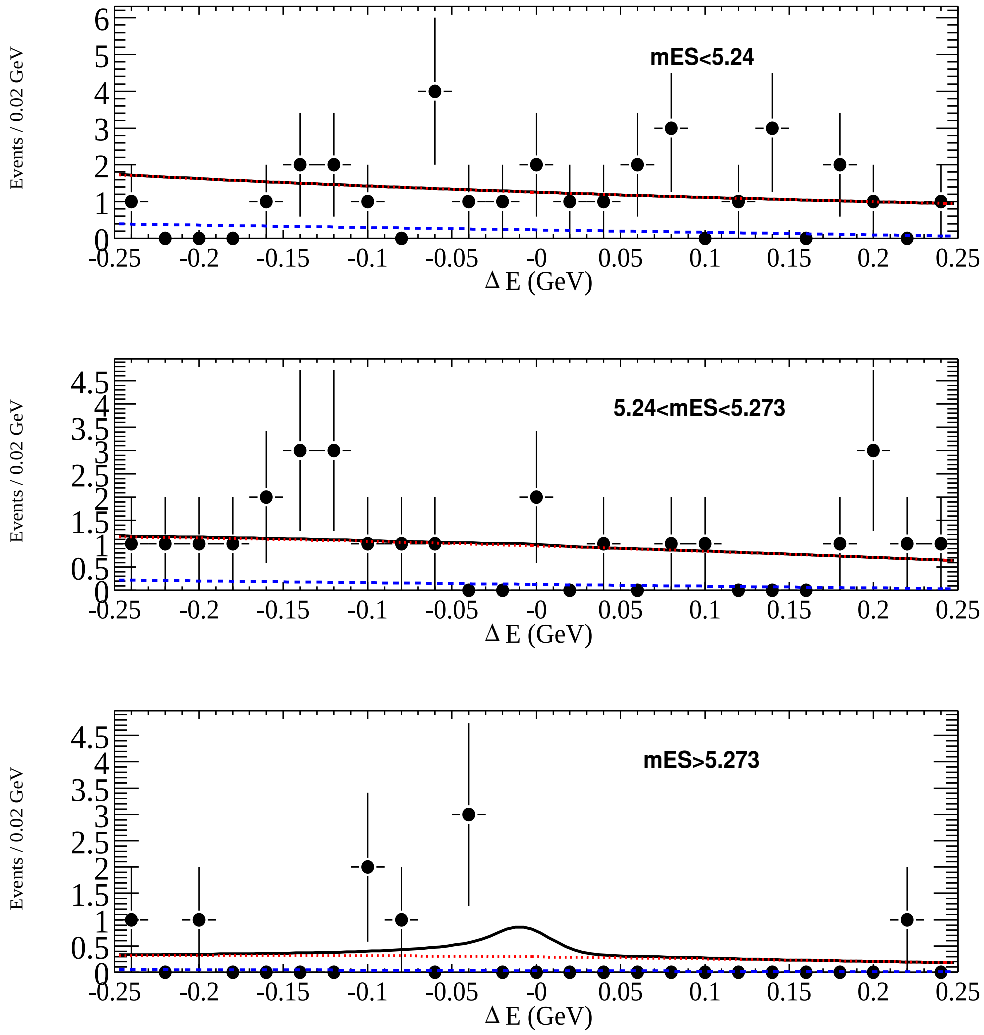


Figure 10.11: ΔE distributions for fits to all of the data in the $B^0 \to K^{*0}e^+e^-$ mode. The fit is displayed in 3 slices of m_{ES}: the upper plot is for $m_{\mathrm{ES}} < 5.24$ GeV/c^2, the middle plot is for $5.24 < m_{\mathrm{ES}} < 5.27$ GeV/c^2, and the bottom plot is for $m_{\mathrm{ES}} > 5.27$ GeV/c^2. The bottom plot thus corresponds to the signal region in m_{ES}. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied.

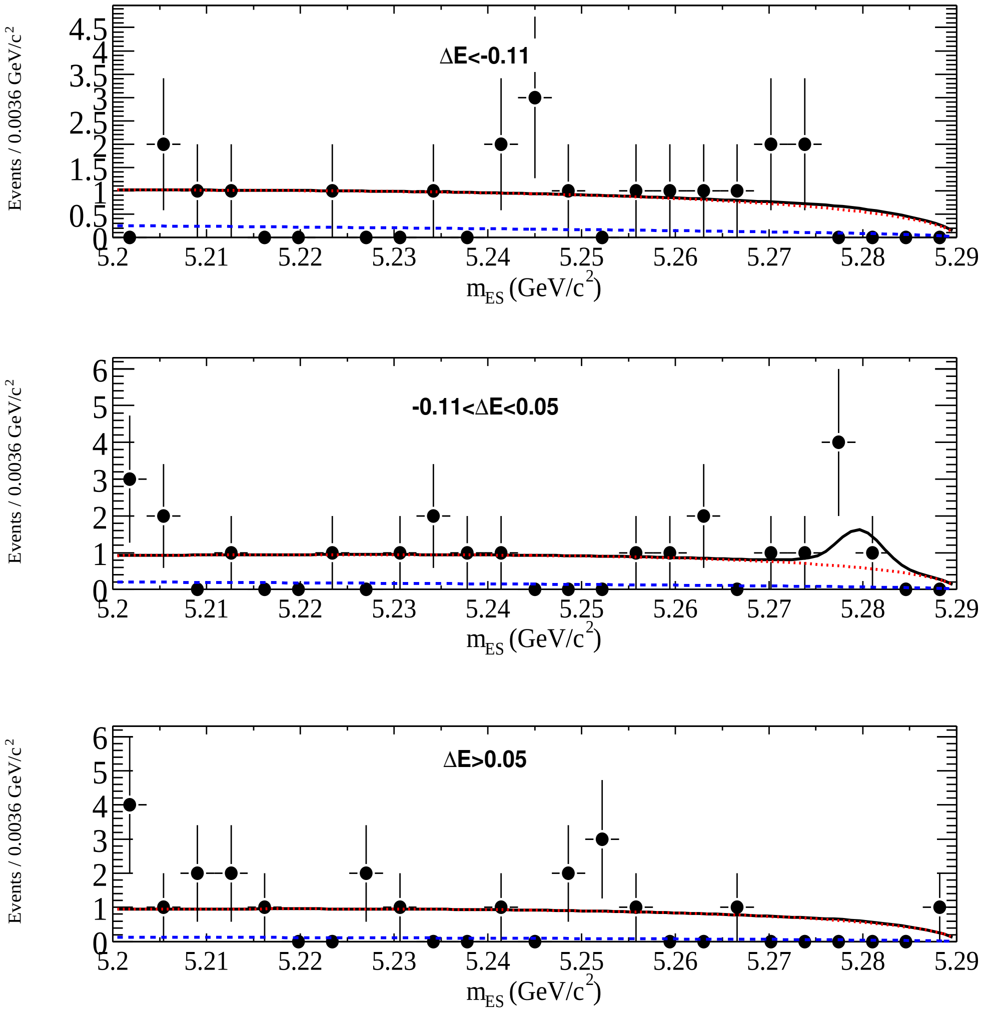


Figure 10.12: m_{ES} distributions for fits to all of the data in the $B^0 \to K^{*0}e^+e^-$ mode. The fit is displayed in 3 slices of ΔE: the upper plot is for $\Delta E <$ -0.11 GeV, the middle plot is for -0.11 $< \Delta E <$ 0.05 GeV, and the bottom plot is for $\Delta E >$ 0.05 GeV. The middle plot thus corresponds to the signal region in ΔE. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied.

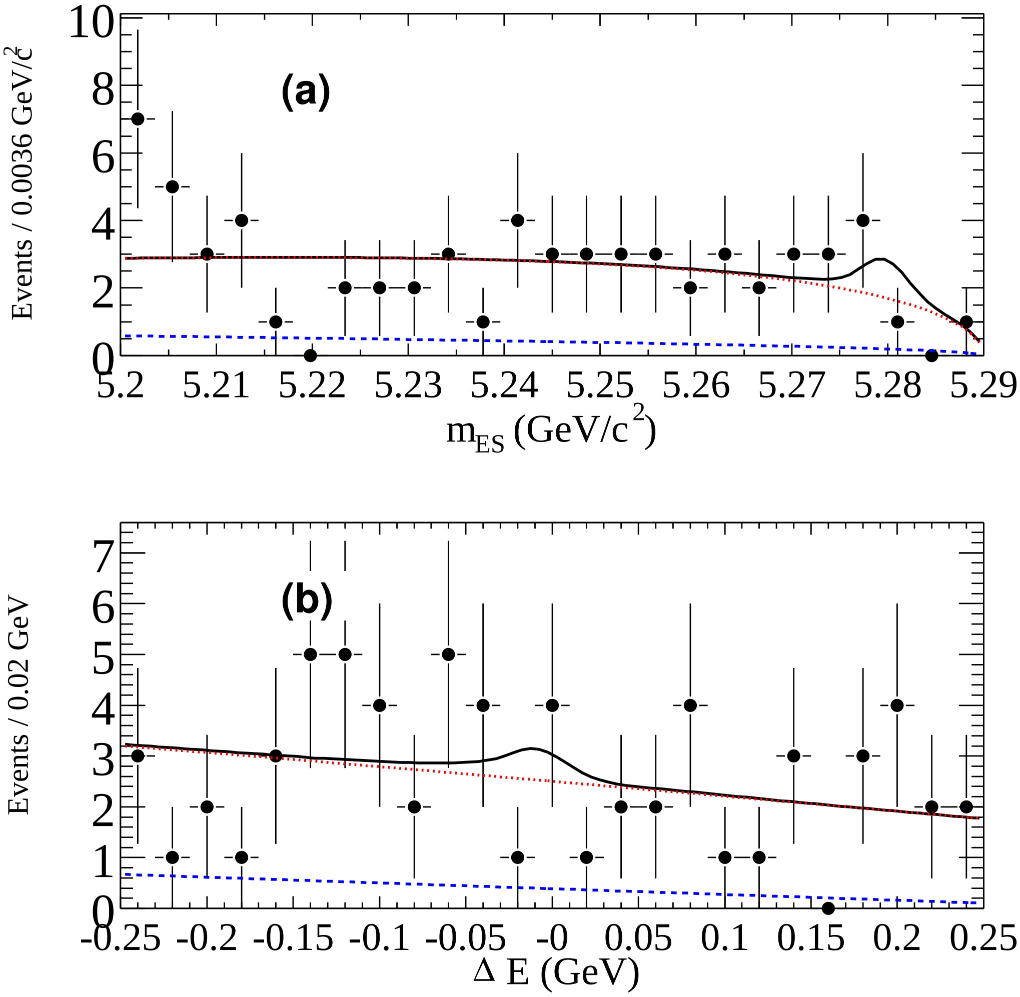


Figure 10.13: (a): m_{ES} and (b): ΔE distributions for fits to all of the data in the $B^0 \to K^{*0}e^+e^-$ mode. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied.

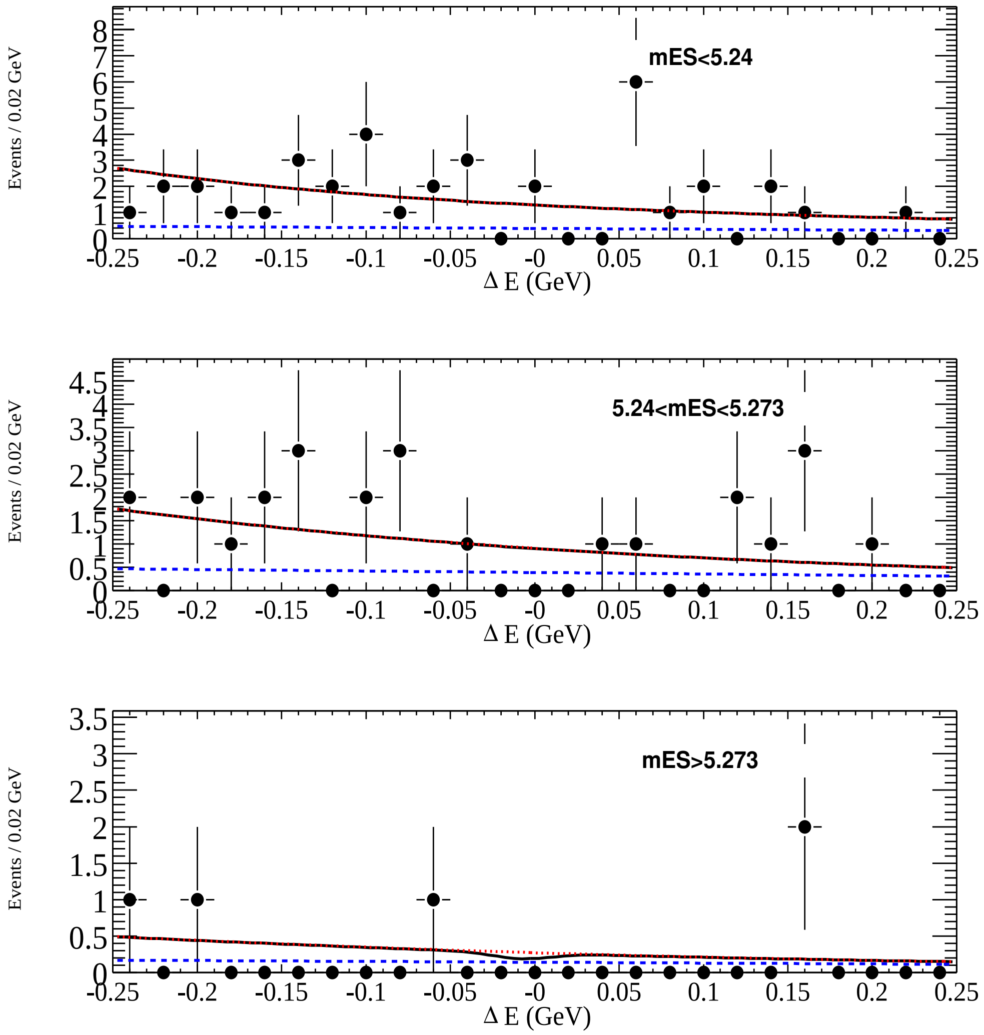


Figure 10.14: ΔE distributions for fits to all of the data in the $B^0 \to K^{*0}\mu^+\mu^-$ mode. The fit is displayed in 3 slices of m_{ES}: the upper plot is for $m_{\text{ES}} < 5.24\ \text{GeV}/c^2$, the middle plot is for $5.24 < m_{\text{ES}} < 5.27\ \text{GeV}/c^2$, and the bottom plot is for $m_{\text{ES}} > 5.27\ \text{GeV}/c^2$. The bottom plot thus corresponds to the signal region in m_{ES}. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied.

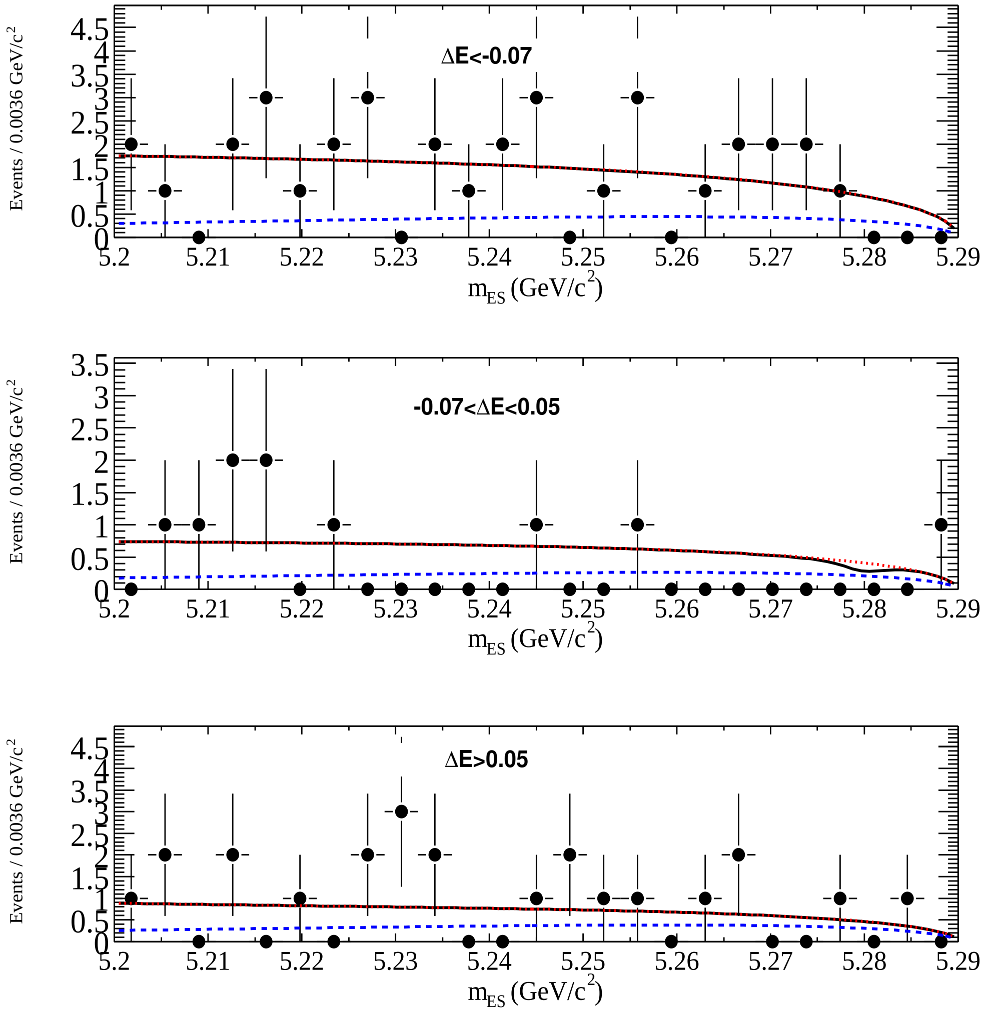


Figure 10.15: m_{ES} distributions for fits to all of the data in the $B^0 \to K^{*0}\mu^+\mu^-$ mode. The fit is displayed in 3 slices of ΔE: the upper plot is for $\Delta E <$ -0.07 GeV, the middle plot is for -0.07 $< \Delta E <$ 0.05 GeV, and the bottom plot is for $\Delta E >$ 0.05 GeV. The middle plot thus corresponds to the signal region in ΔE. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied.

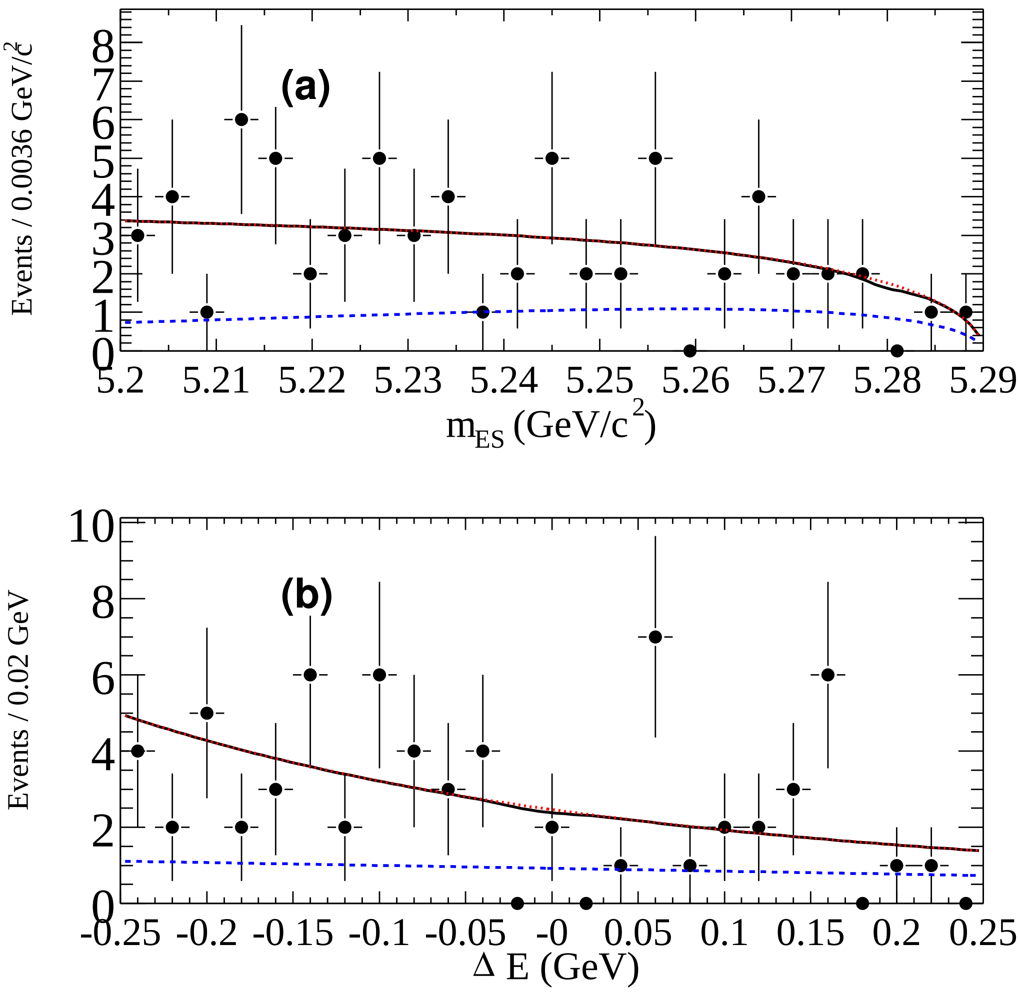


Figure 10.16: (a): m_{ES} and (b): ΔE distributions for fits to all of the data in the $B^0 \to K^{*0}\mu^+\mu^-$ mode. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied.

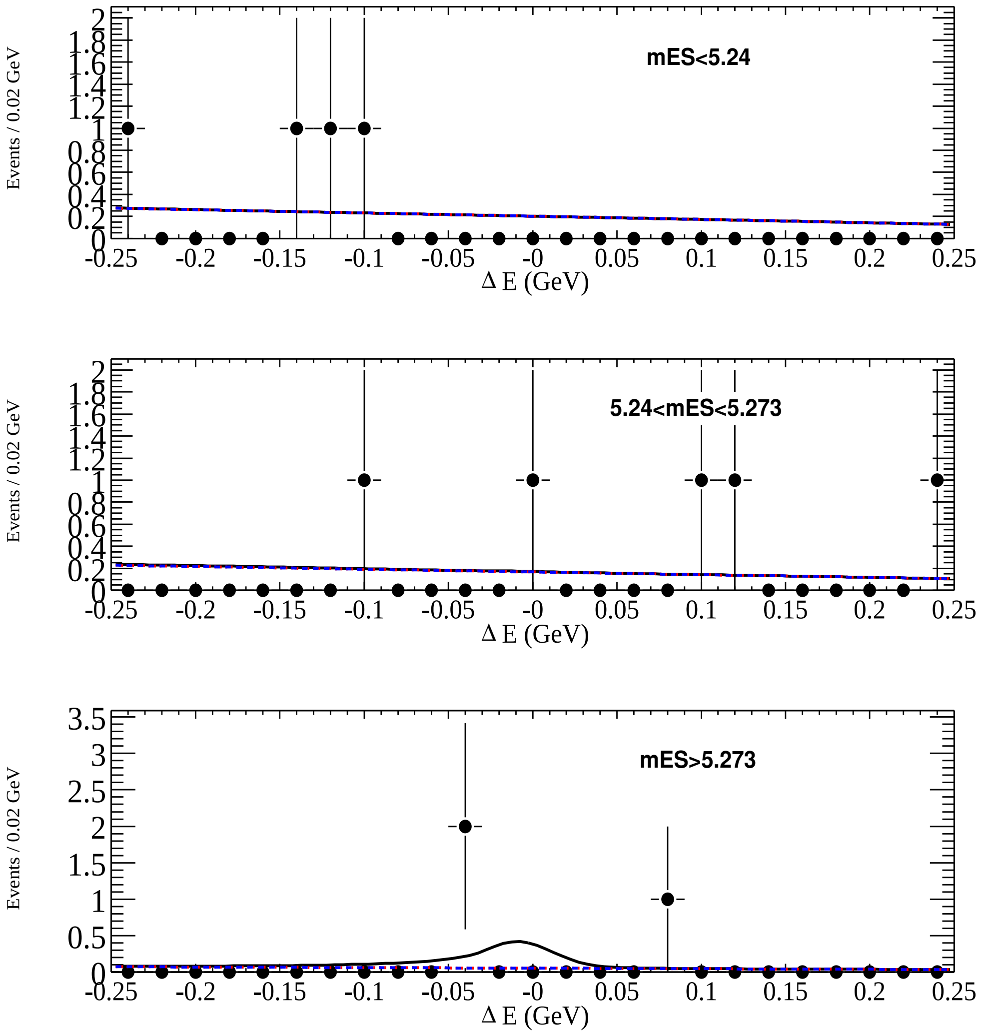


Figure 10.17: ΔE distributions for fits to all of the data in the $B^0 \to K^0_S e^+ e^-$ mode. The fit is displayed in 3 slices of m_{ES}: the upper plot is for $m_{\mathrm{ES}} < 5.24$ GeV/c^2, the middle plot is for $5.24 < m_{\mathrm{ES}} < 5.27$ GeV/c^2, and the bottom plot is for $m_{\mathrm{ES}} > 5.27$ GeV/c^2. The bottom plot thus corresponds to the signal region in m_{ES}. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied.

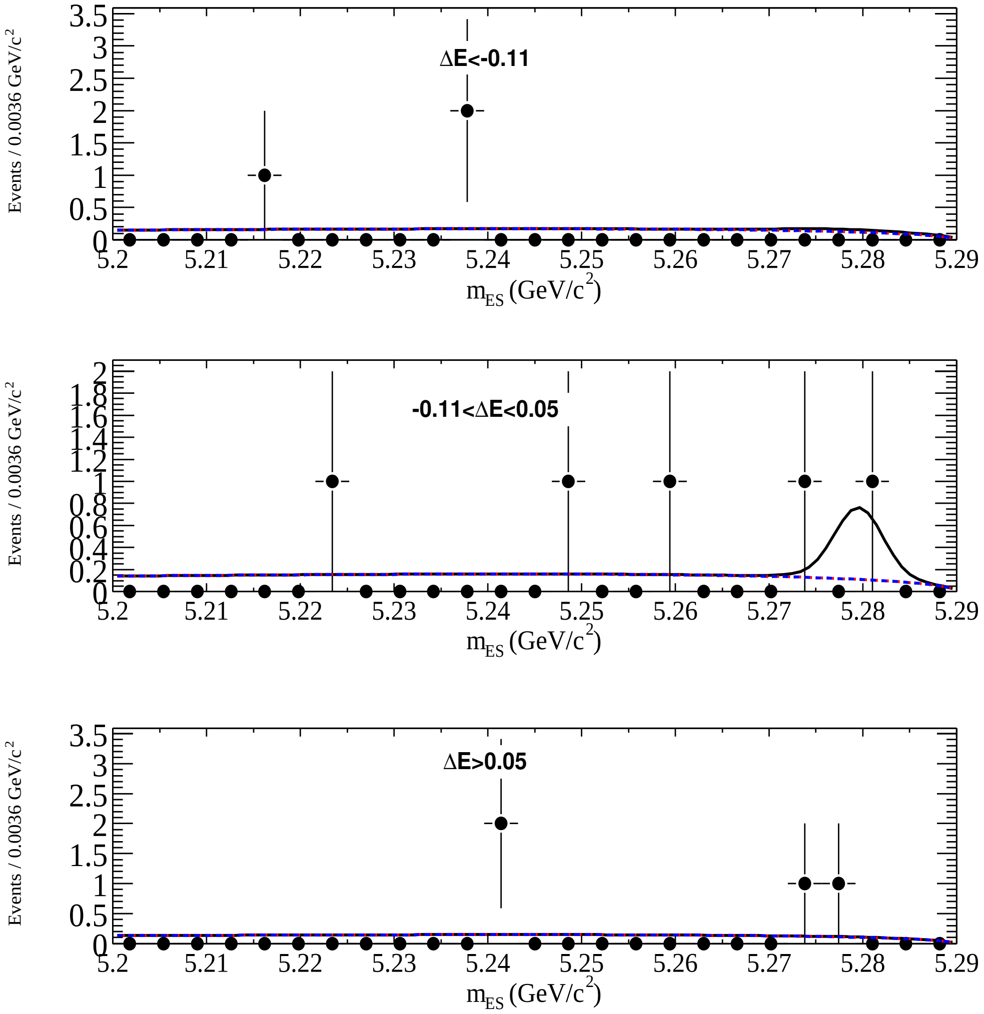


Figure 10.18: m_{ES} distributions for fits to all of the data in the $B^0 \rightarrow K^0_S e^+e^-$ mode. The fit is displayed in 3 slices of ΔE: the upper plot is for $\Delta E <$ -0.11 GeV, the middle plot is for -0.11 $< \Delta E <$ 0.05 GeV, and the bottom plot is for $\Delta E >$ 0.05 GeV. The middle plot thus corresponds to the signal region in ΔE. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied.

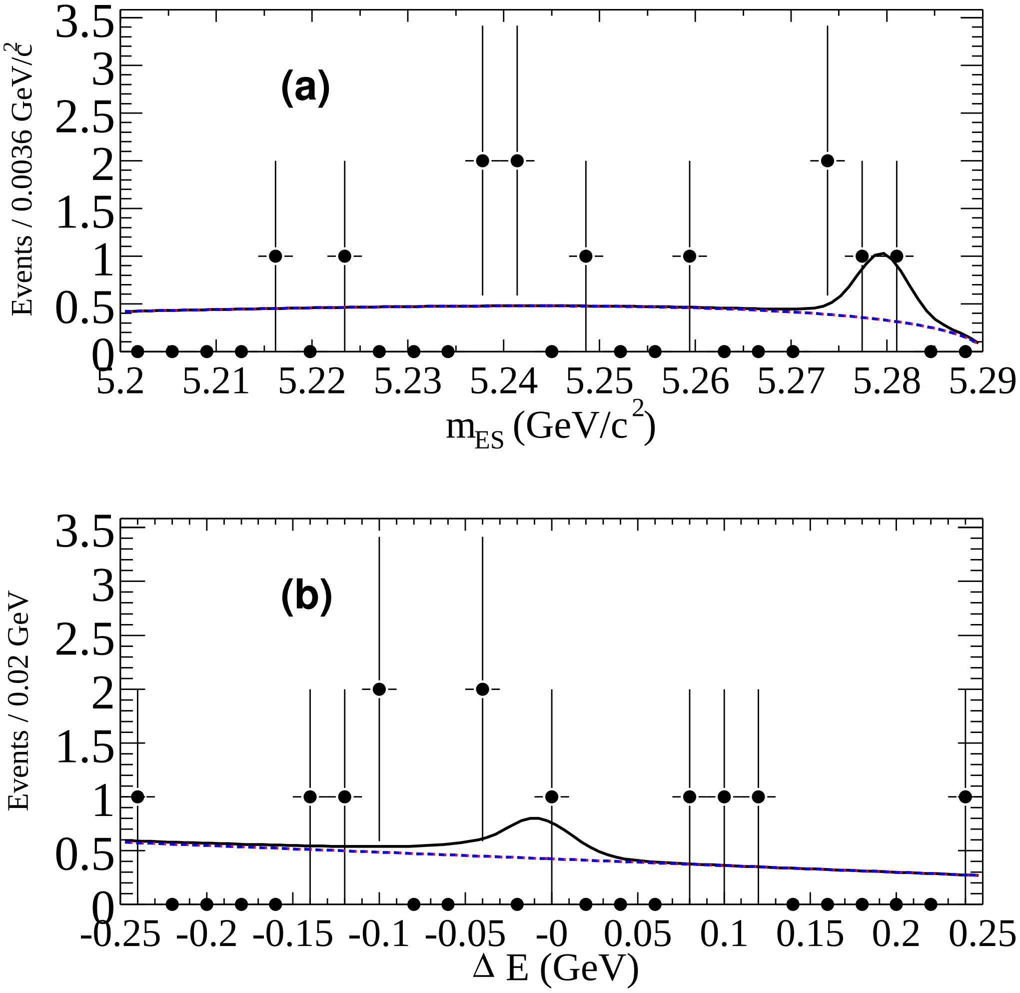


Figure 10.19: (a): m_{ES} and (b): ΔE distributions for fits to all of the data in the $B^0 \to K_S^0 e^+ e^-$ mode. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied.

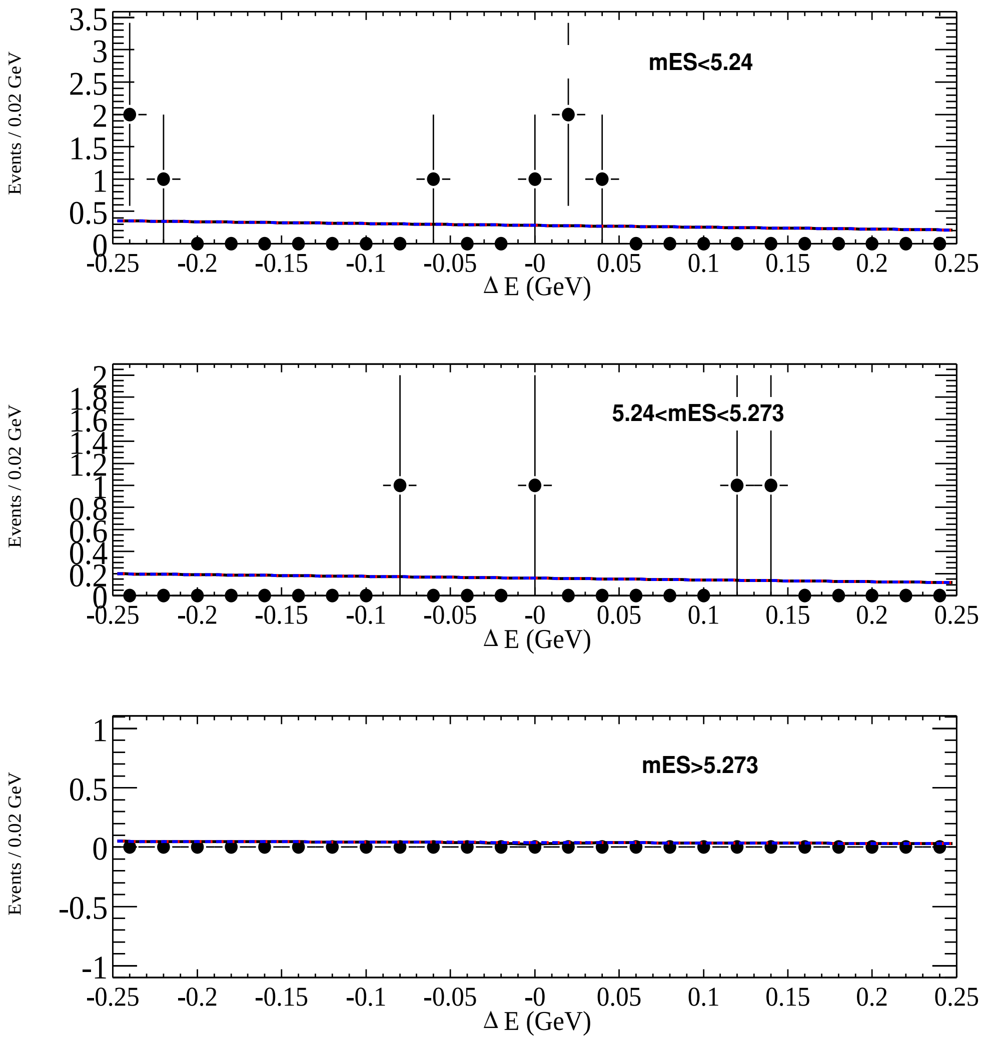


Figure 10.20: ΔE distributions for fits to all of the data in the $B^0 \to K_S^0\mu^+\mu^-$ mode. The fit is displayed in 3 slices of $m_{\rm ES}$: the upper plot is for $m_{\rm ES} < 5.24$ GeV/c^2, the middle plot is for $5.24 < m_{\rm ES} < 5.27$ GeV/c^2, and the bottom plot is for $m_{\rm ES} > 5.27$ GeV/c^2. The bottom plot thus corresponds to the signal region in $m_{\rm ES}$. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied.

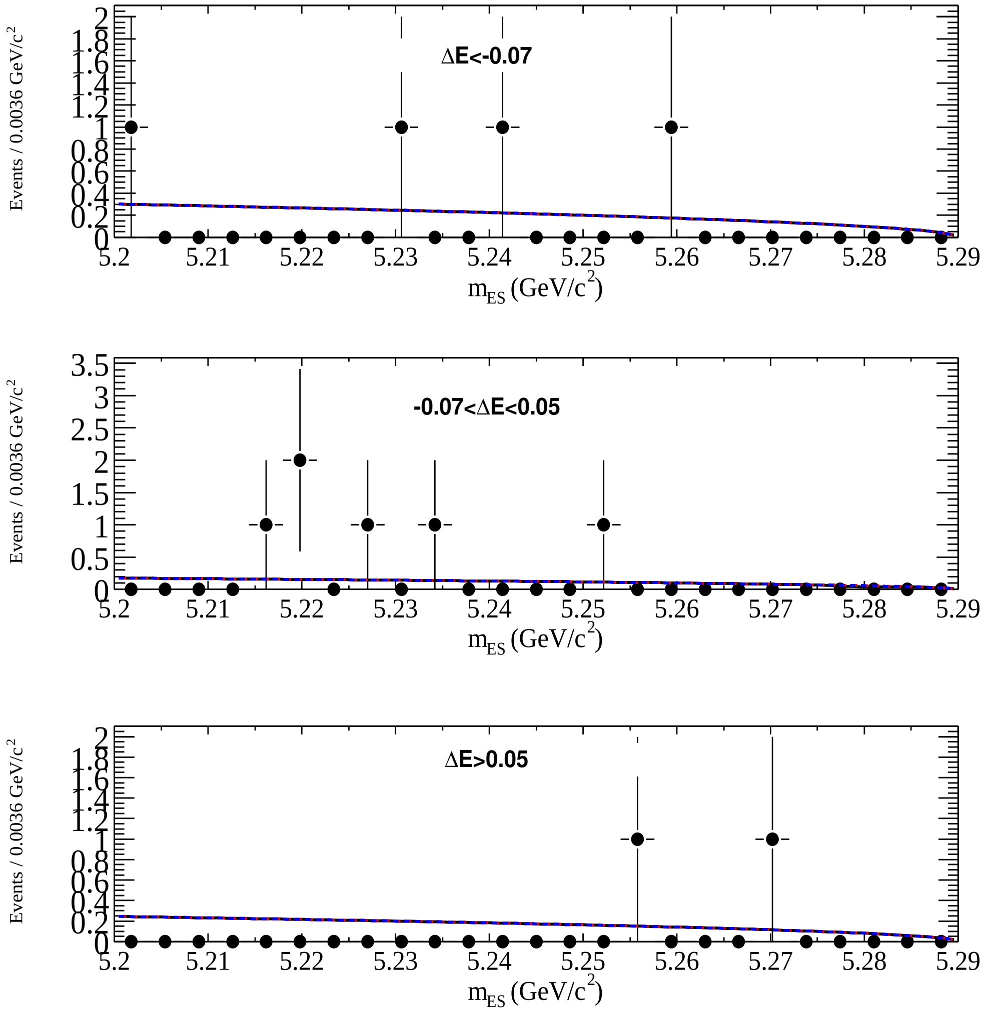


Figure 10.21: m_{ES} distributions for fits to all of the data in the $B^0 \rightarrow K^0_S\mu^+\mu^-$ mode. The fit is displayed in 3 slices of ΔE: the upper plot is for $\Delta E <$ -0.07 GeV, the middle plot is for -0.07 $< \Delta E <$ 0.05 GeV, and the bottom plot is for $\Delta E >$ 0.05 GeV. The middle plot thus corresponds to the signal region in ΔE. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied.

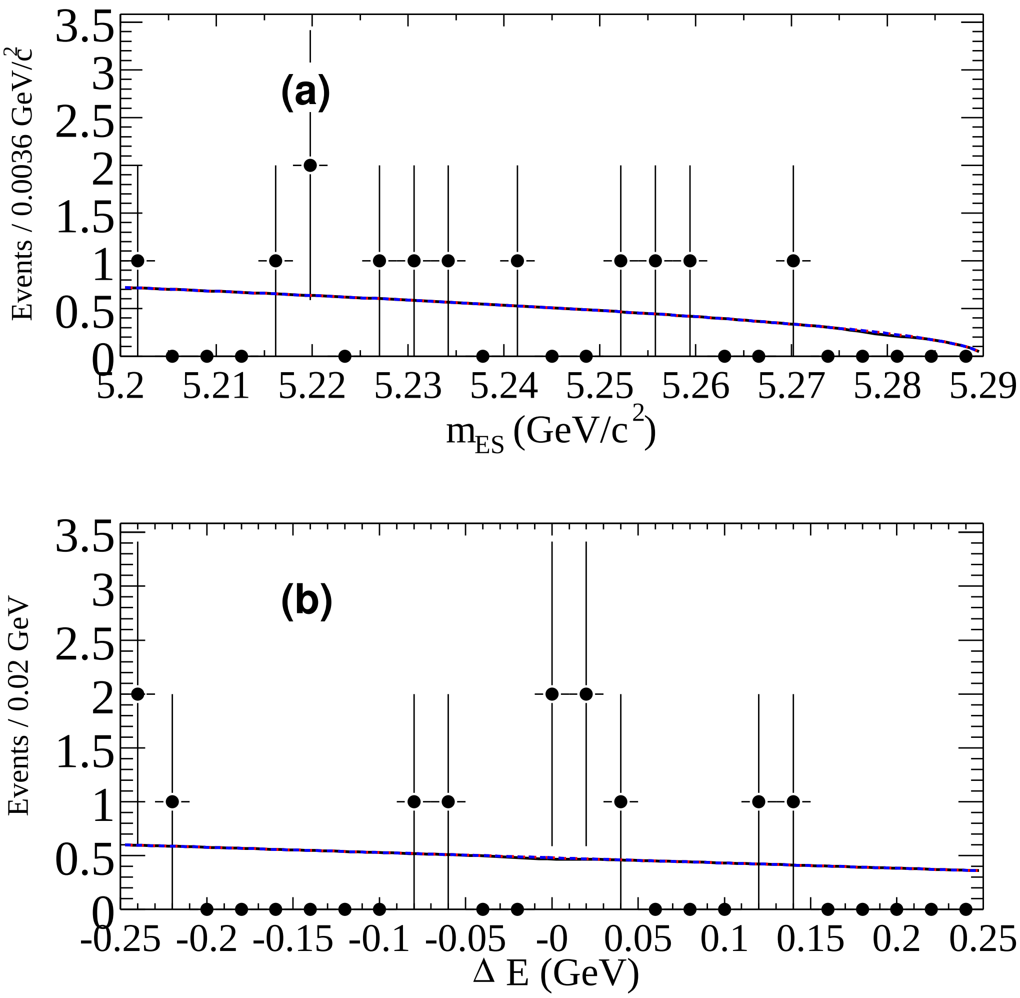


Figure 10.22: (a): m_{ES} and (b): ΔE distributions for fits to all of the data in the $B^0 \to K^0_S\mu^+\mu^-$ mode. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied.

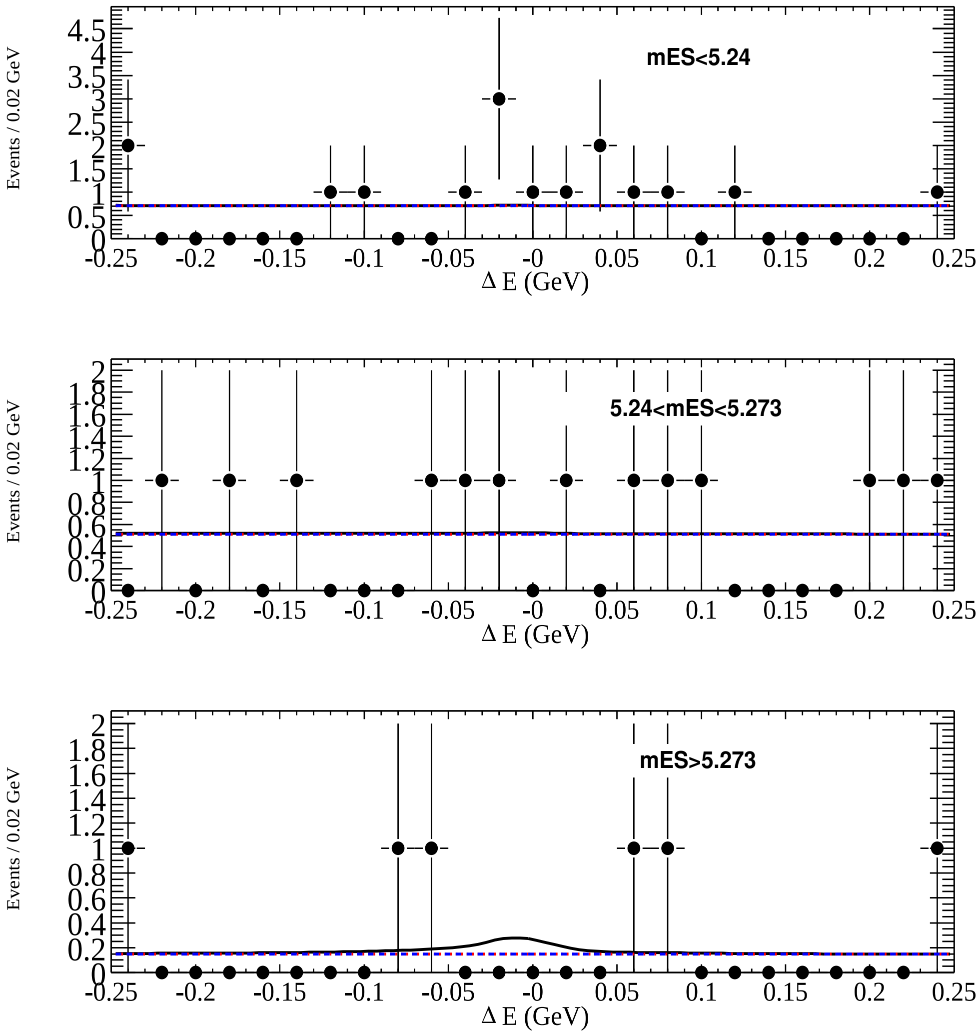


Figure 10.23: ΔE distributions for fits to all of the data in the $B^+ \to K^{*+} e^+ e^-$ ($K^{*+} \to K^0_S \pi^+$) mode. The fit is displayed in 3 slices of m_{ES}: the upper plot is for $m_{\mathrm{ES}} < 5.24$ GeV/c^2, the middle plot is for $5.24 < m_{\mathrm{ES}} < 5.27$ GeV/c^2, and the bottom plot is for $m_{\mathrm{ES}} > 5.27$ GeV/c^2. The bottom plot thus corresponds to the signal region in m_{ES}. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied.

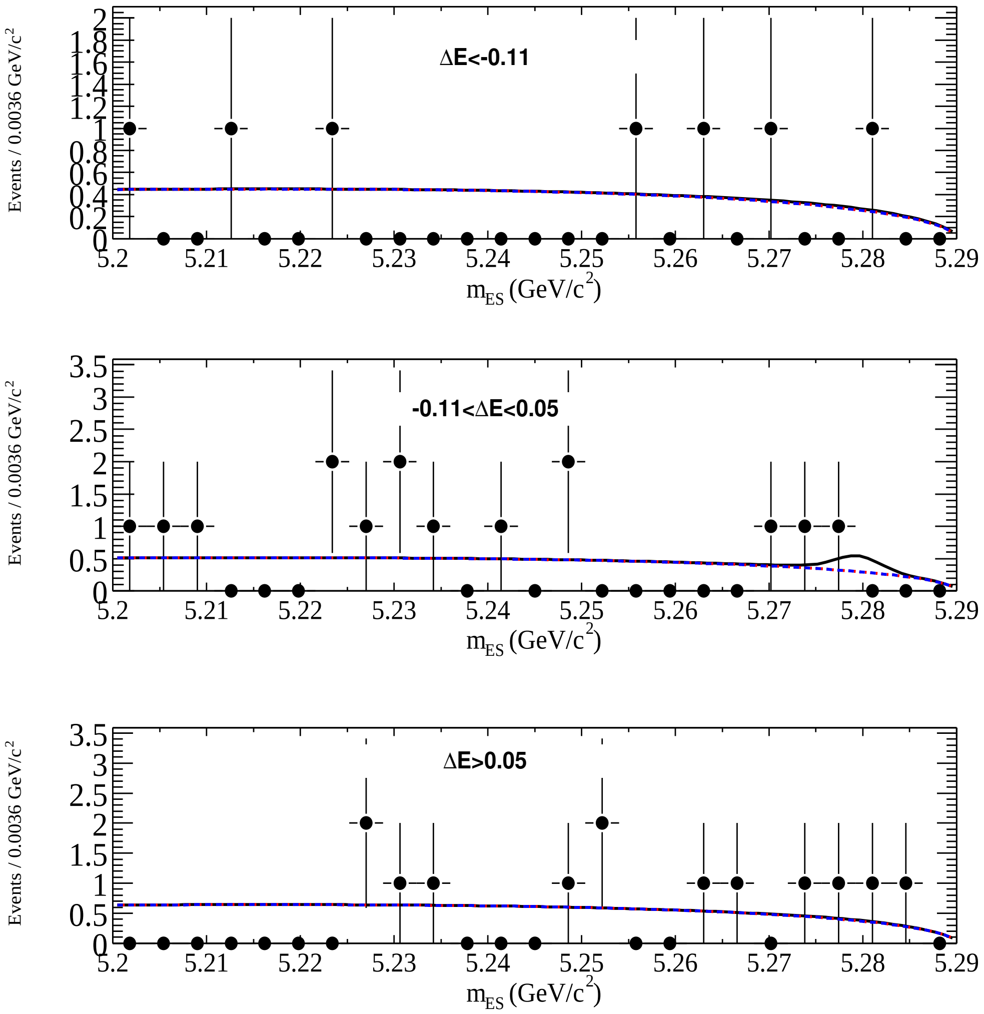


Figure 10.24: m_{ES} distributions for fits to all of the data in the $B^+ \rightarrow K^{*+}e^+e^-$ ($K^{*+} \rightarrow K^0_S\pi^+$) mode. The fit is displayed in 3 slices of ΔE: the upper plot is for $\Delta E <$ -0.11 GeV, the middle plot is for -0.11 $< \Delta E <$ 0.05 GeV, and the bottom plot is for $\Delta E >$ 0.05 GeV. The middle plot thus corresponds to the signal region in ΔE. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied.

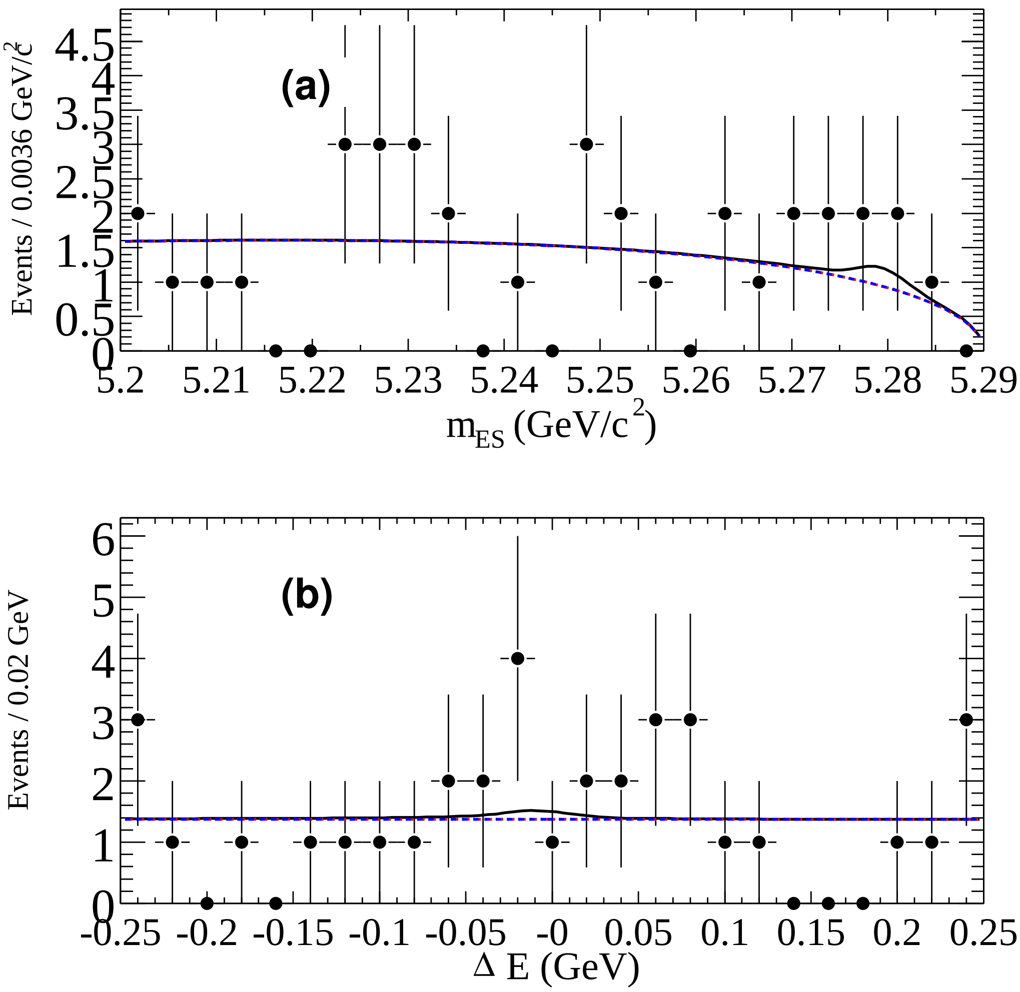


Figure 10.25: (a): $m_{\rm ES}$ and (b): ΔE distributions for fits to all of the data in the $B^+ \to K^{*+}e^+e^-$ ($K^{*+} \to K^0_S\pi^+$) mode. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied.

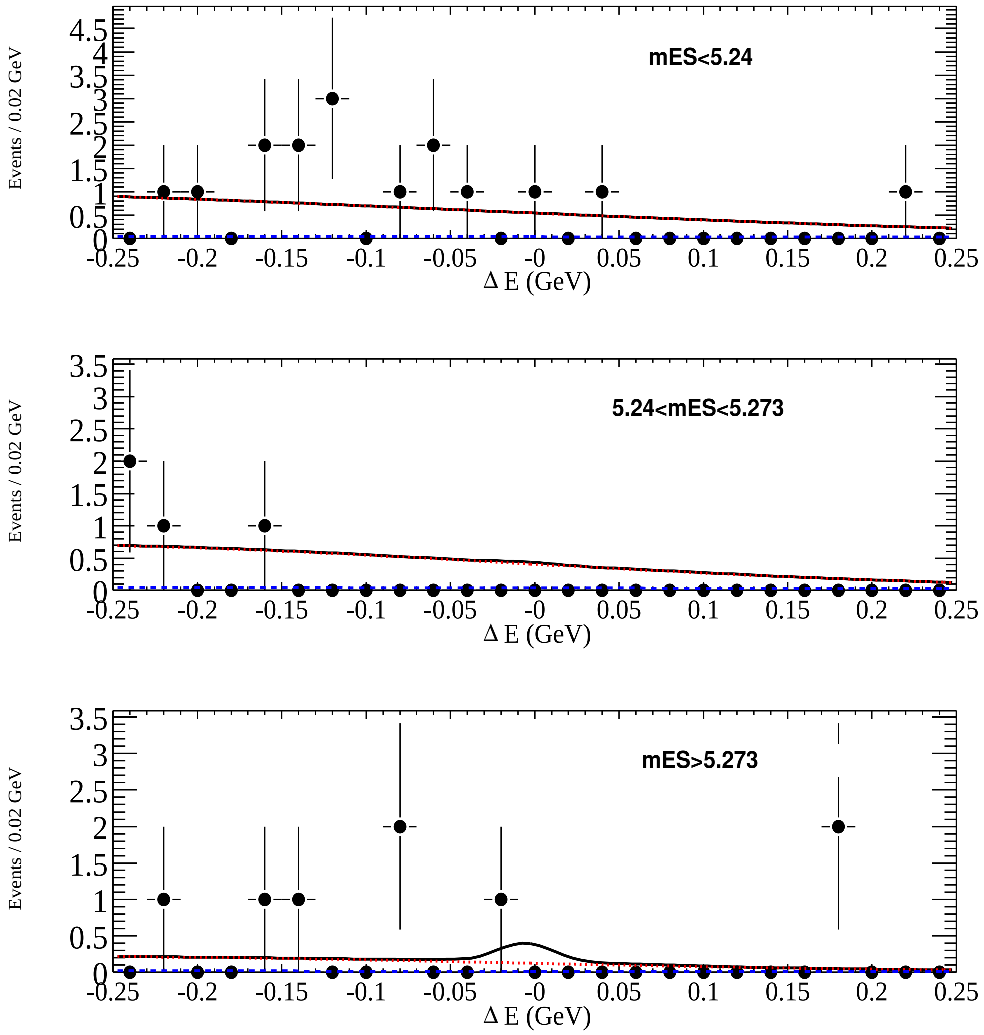


Figure 10.26: ΔE distributions for fits to all of the data in the $B^+ \to K^{*+}\mu^+\mu^-$ ($K^{*+} \to K^0_S\pi^+$) mode. The fit is displayed in 3 slices of $m_{\rm ES}$: the upper plot is for $m_{\rm ES} < 5.24$ GeV/c^2, the middle plot is for $5.24 < m_{\rm ES} < 5.27$ GeV/c^2, and the bottom plot is for $m_{\rm ES} > 5.27$ GeV/c^2. The bottom plot thus corresponds to the signal region in $m_{\rm ES}$. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied.

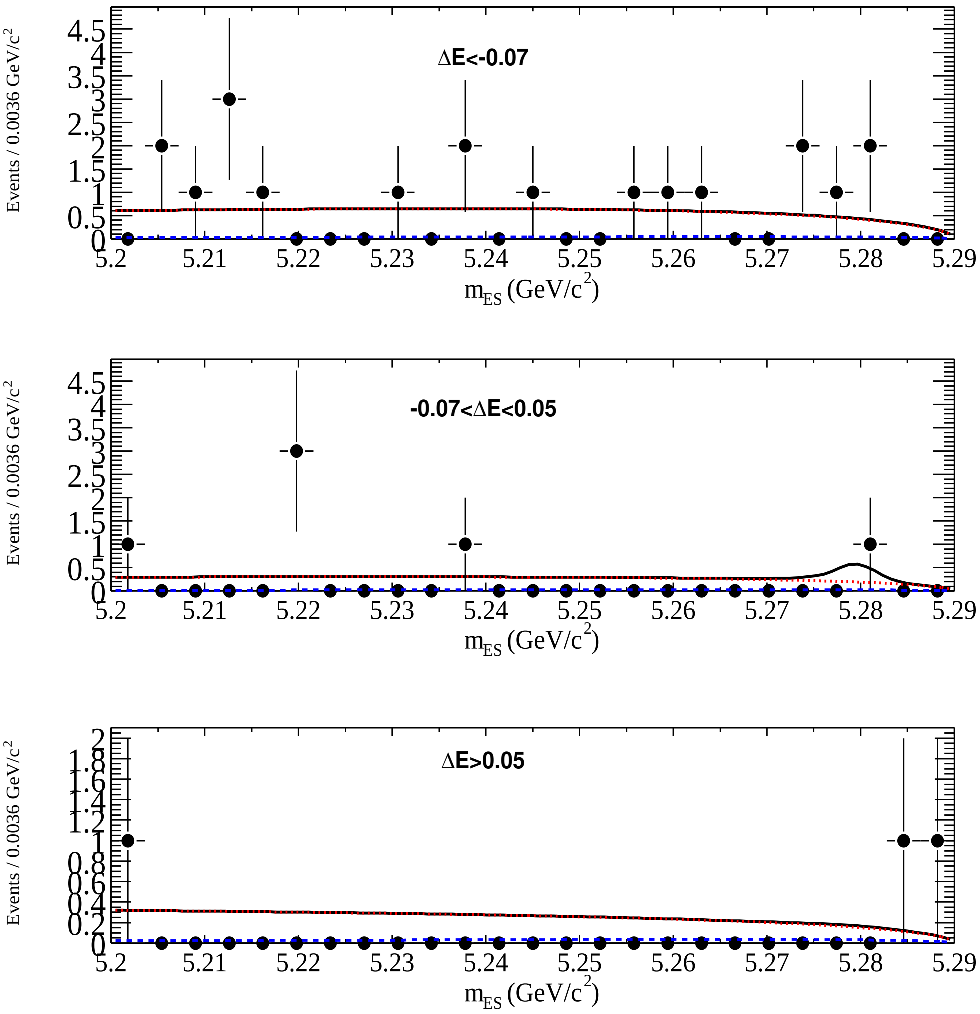


Figure 10.27: m_{ES} distributions for fits to all of the data in the $B^+ \to K^{*+}\mu^+\mu^-$ ($K^{*+} \to K^0_S\pi^+$) mode. The fit is displayed in 3 slices of ΔE: the upper plot is for $\Delta E <$ -0.07 GeV, the middle plot is for -0.07 $< \Delta E <$ 0.05 GeV, and the bottom plot is for $\Delta E >$ 0.05 GeV. The middle plot thus corresponds to the signal region in ΔE. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied.

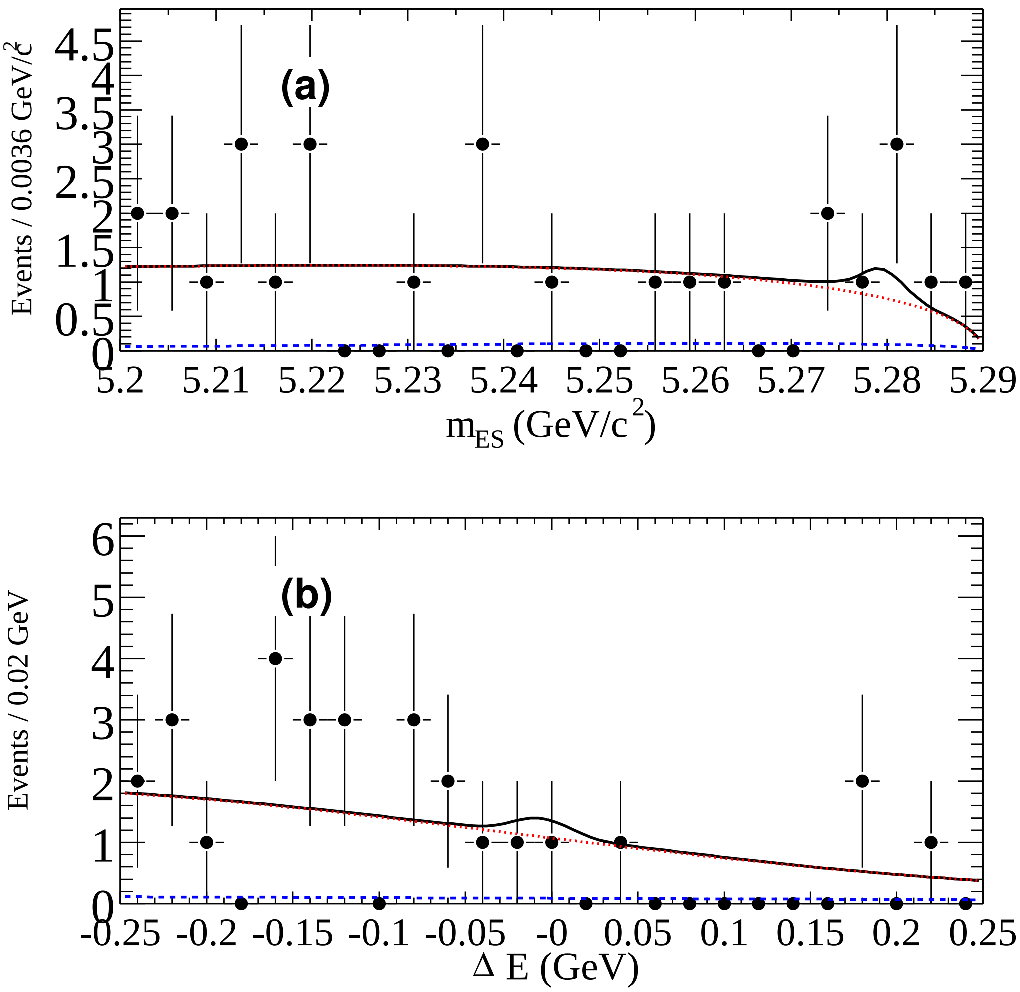


Figure 10.28: (a): $m_{\rm ES}$ and (b): ΔE distributions for fits to all of the data in the $B^+ \to K^{*+}\mu^+\mu^-$ ($K^{*+} \to K^0_S\pi^+$) mode. The points with error bars are the data. Dashed line: continuum background; dotted line: the sum of $B\overline{B}$ and continuum backgrounds; solid line: the sum of total background and signal contributions. All of the event selection criteria have been applied.

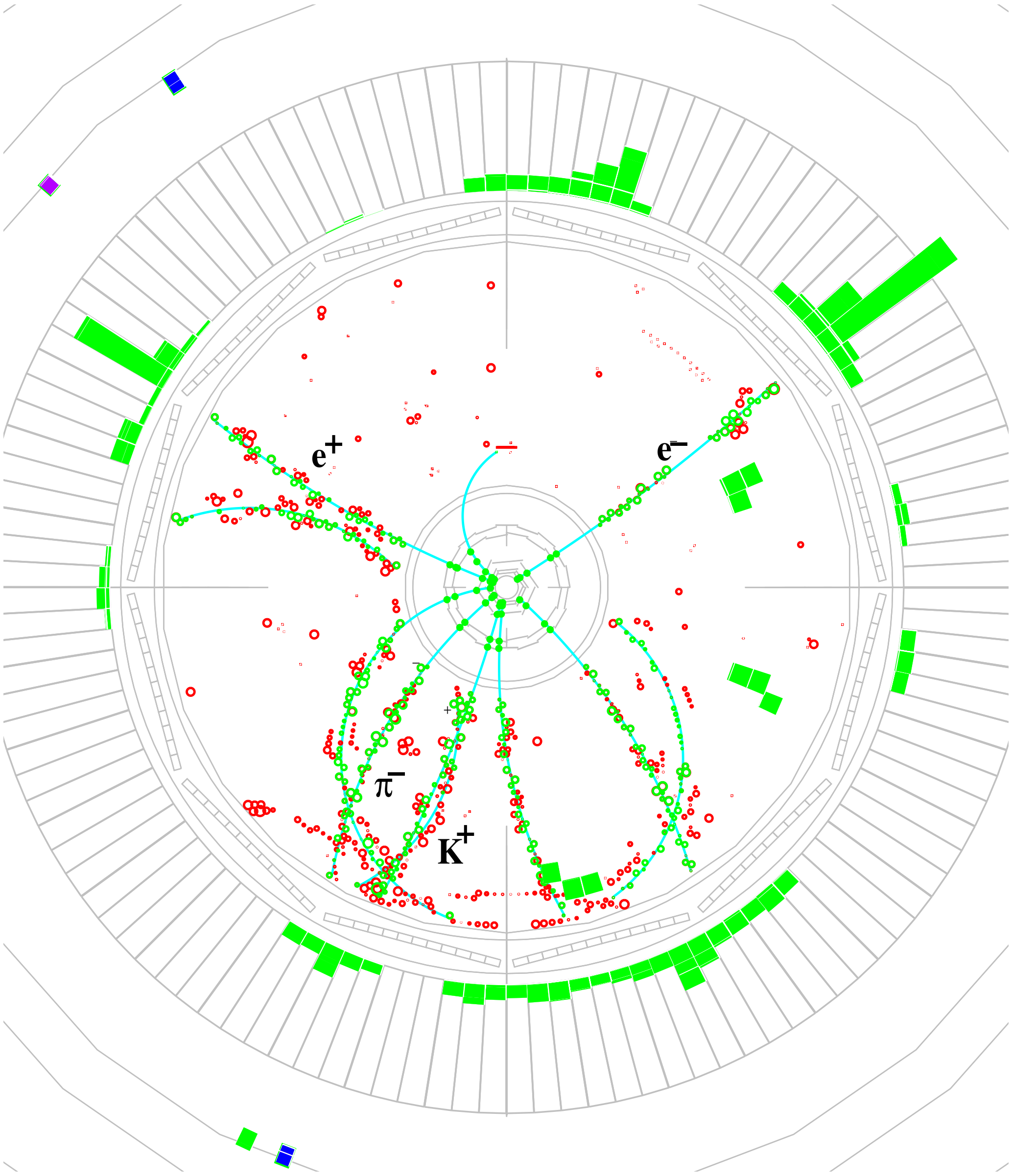


Figure 10.29: Event display for one of the $B^0 \to K^{*0}e^+e^-$ candidate events in the signal box: run 12437.

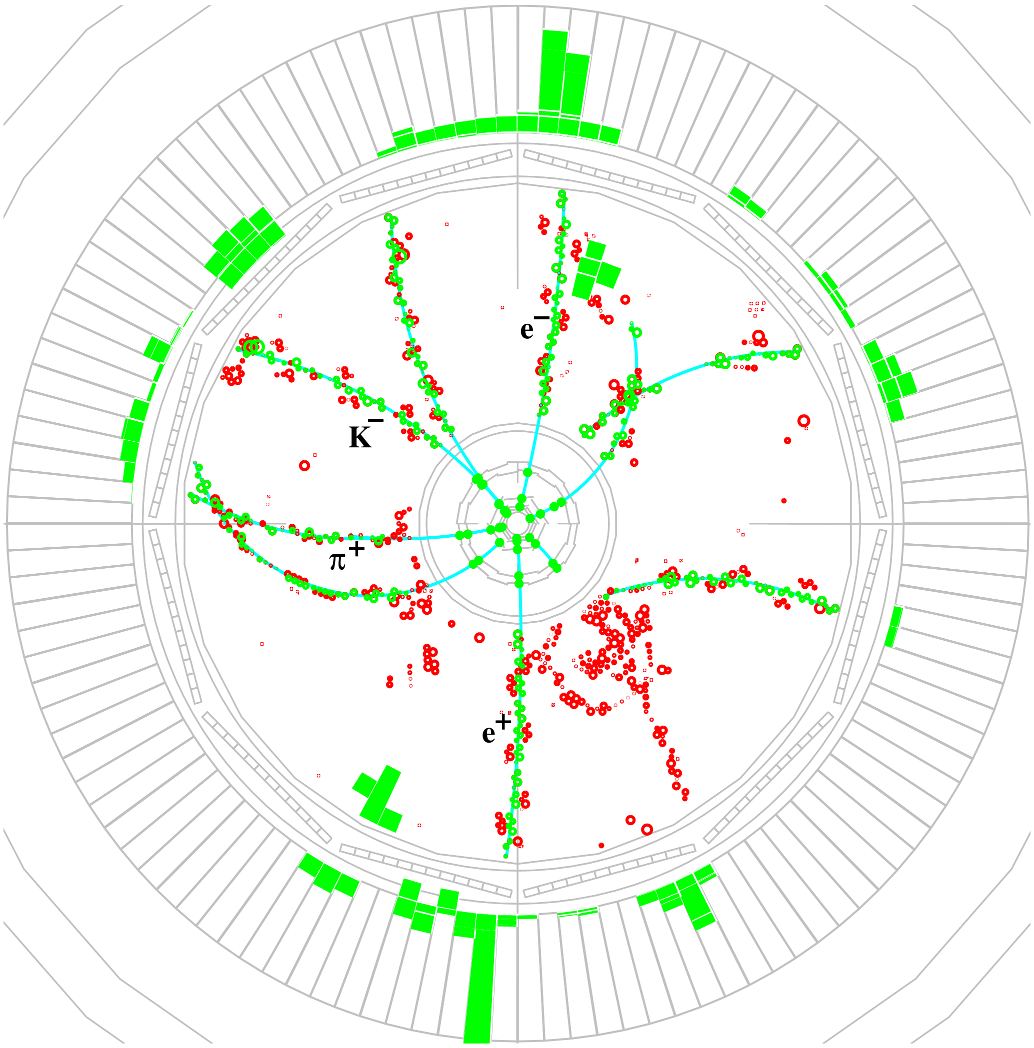


Figure 10.30: Event display for one of the $B^0 \to K^{*0}e^+e^-$ candidate events in the signal box: run 14184.

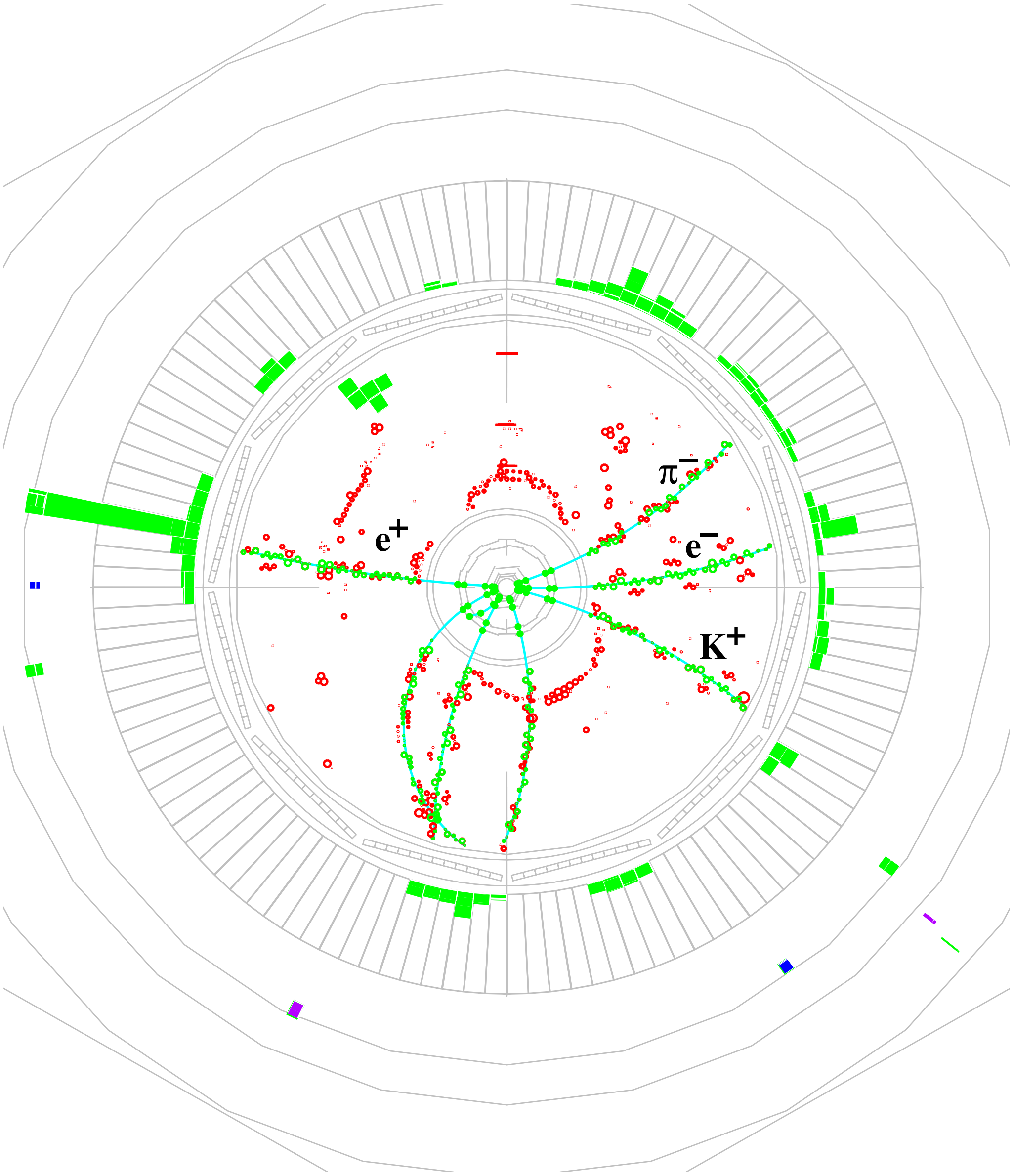


Figure 10.31: Event display for one of the $B^0 \to K^{*0}e^+e^-$ candidate events in the signal box: run 14284.

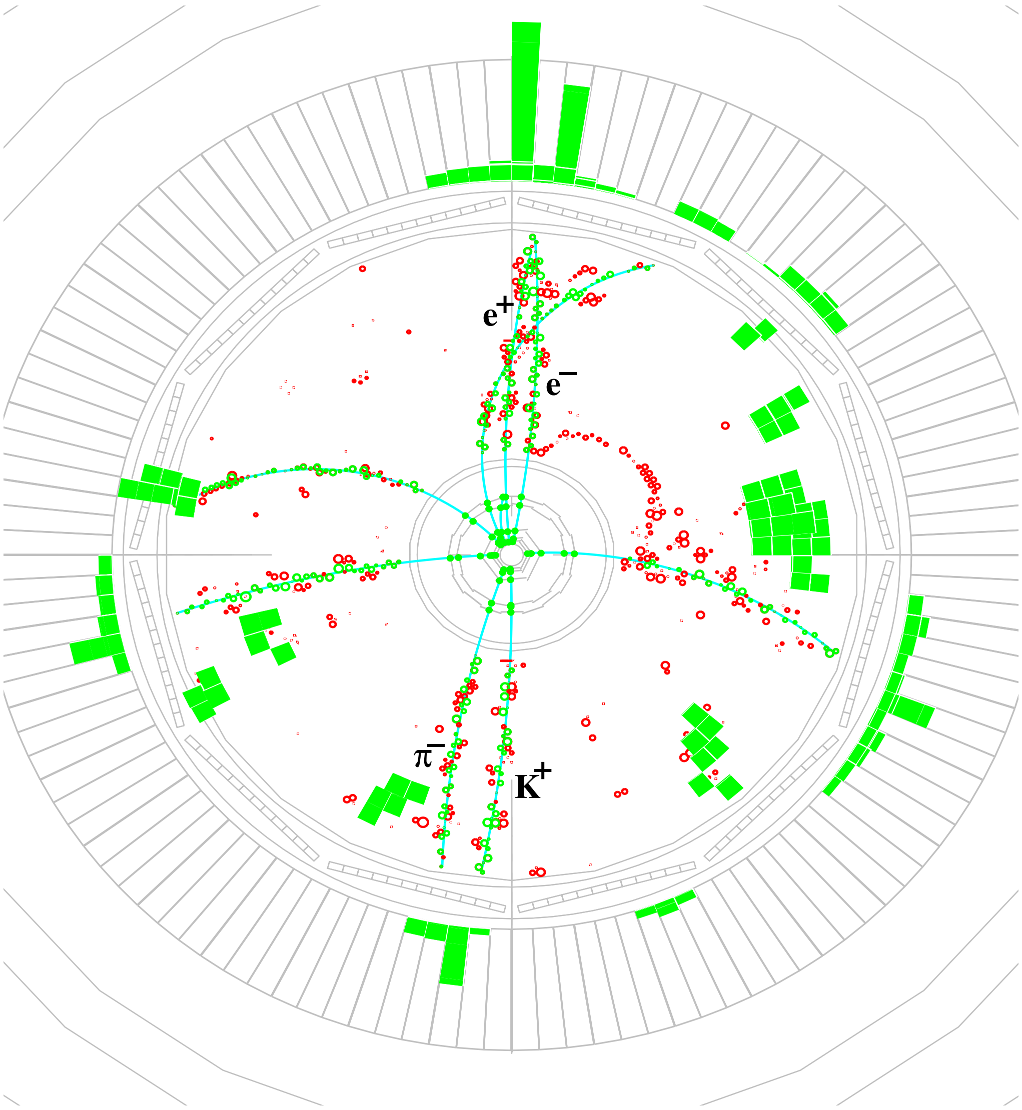


Figure 10.32: Event display for one of the $B^0 \to K^{*0}e^+e^-$ candidate events in the signal box: run 14438.

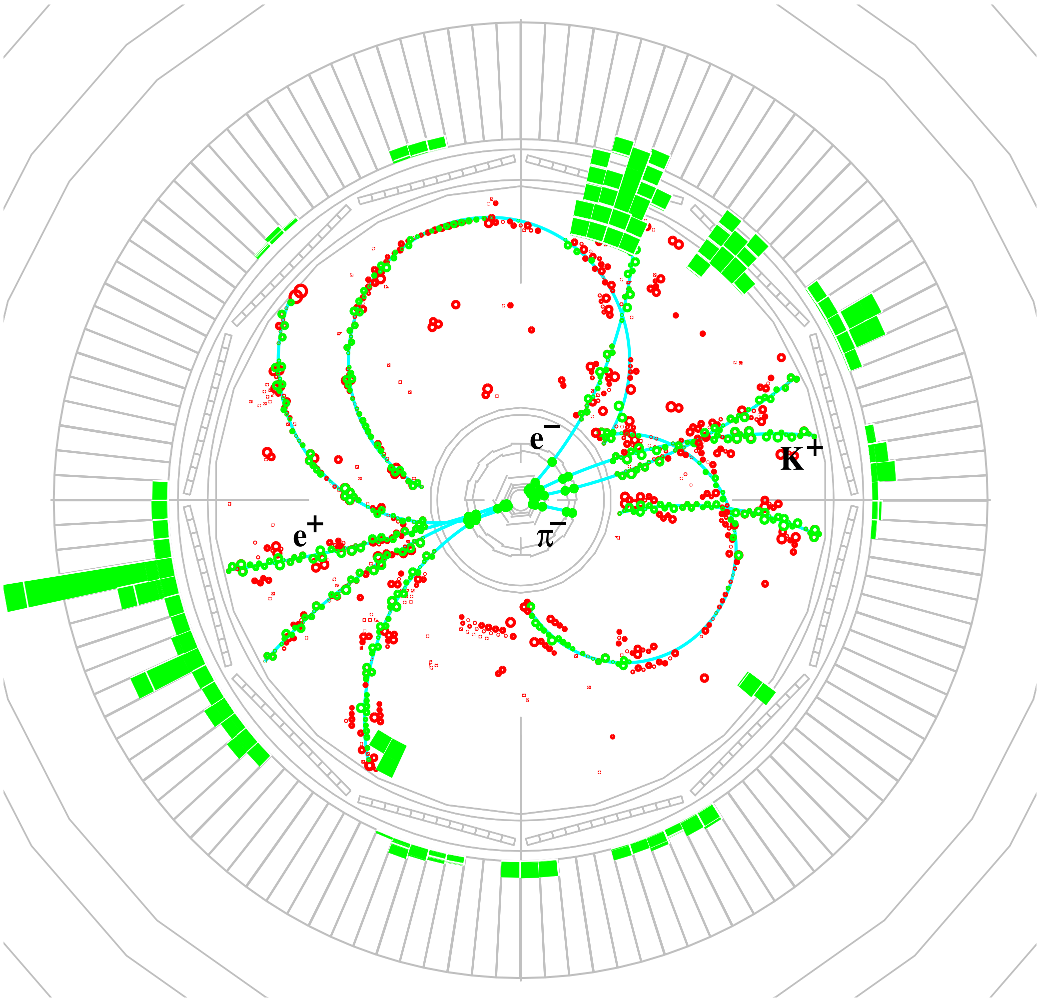


Figure 10.33: Event display for one of the $B^0 \to K^{*0} e^+ e^-$ candidate events in the signal box: run 14819.

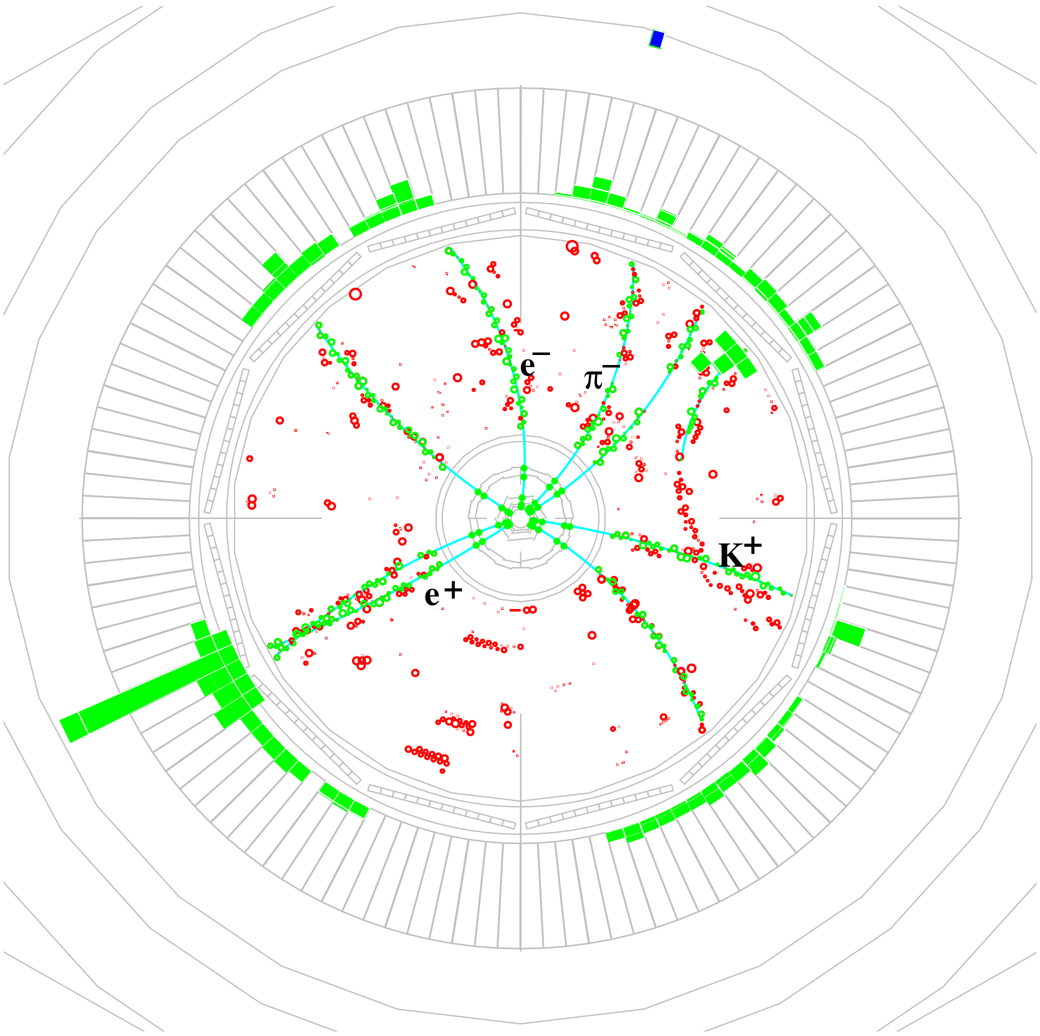


Figure 10.34: Event display for one of the $B^0 \to K^{*0}e^+e^-$ candidate events in the signal box: run 16244.

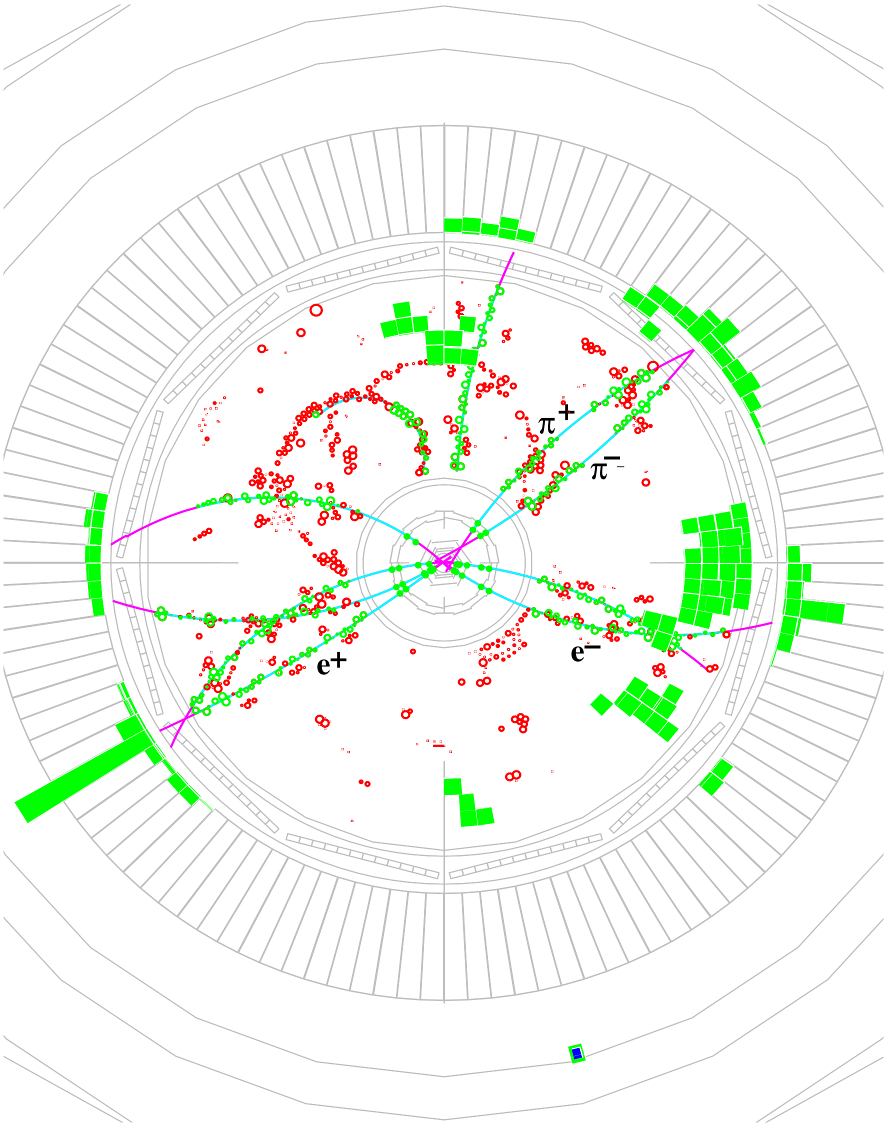


Figure 10.35: Event display for one of the $B^0 \to K_S^0 e^+ e^-$ candidate events in the signal box: run 16326. Note the displaced vertex of the K_S^0 on the right.

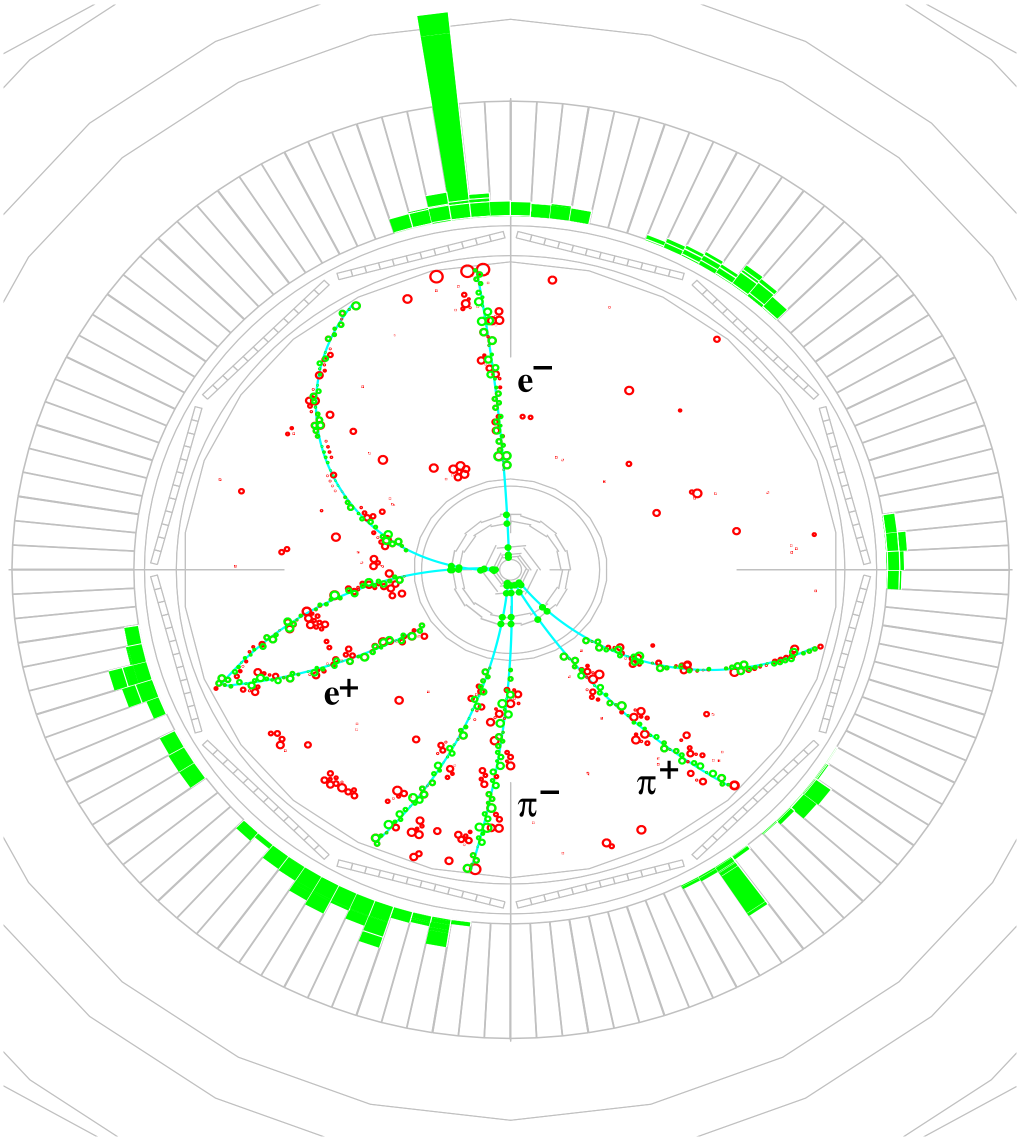


Figure 10.36: Event display for one of the $B^0 \to K^0_S e^+ e^-$ candidate events in the signal box: run 16690.

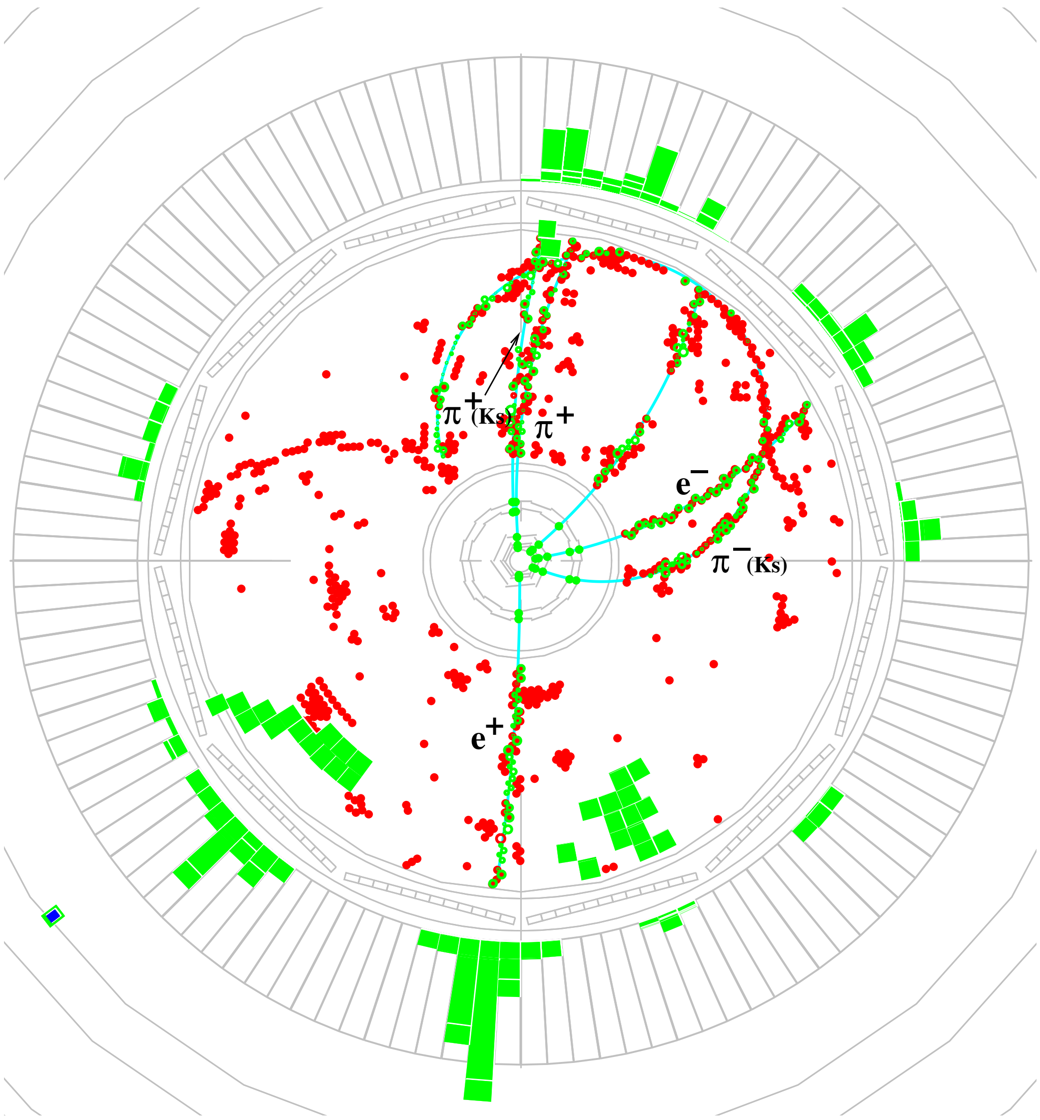


Figure 10.37: Event display for one of the $B^+ \to K^{*+}e^+e^-$ ($K^{*+} \to K^0_S\pi^+$) candidate events in the signal box: run 16703. The pions with Ks in parentheses come from the K^0_S.

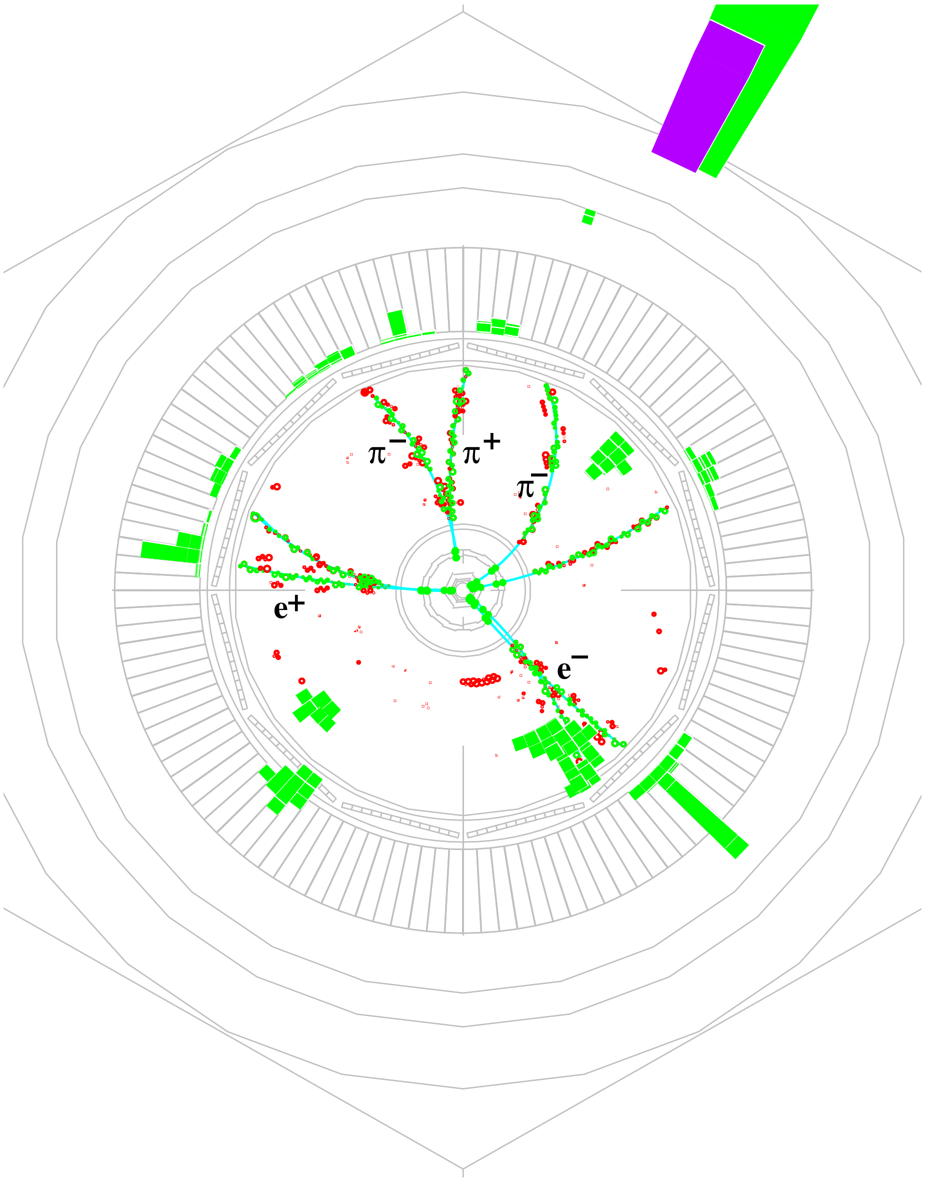


Figure 10.38: Event display for one of the $B^+ \to K^{*+}e^+e^-$ ($K^{*+} \to K^0_S\pi^+$) candidate events in the signal box: run 12784. The $\pi^+\pi^-$ pair comes from the K^0_S (note the displaced vertex).

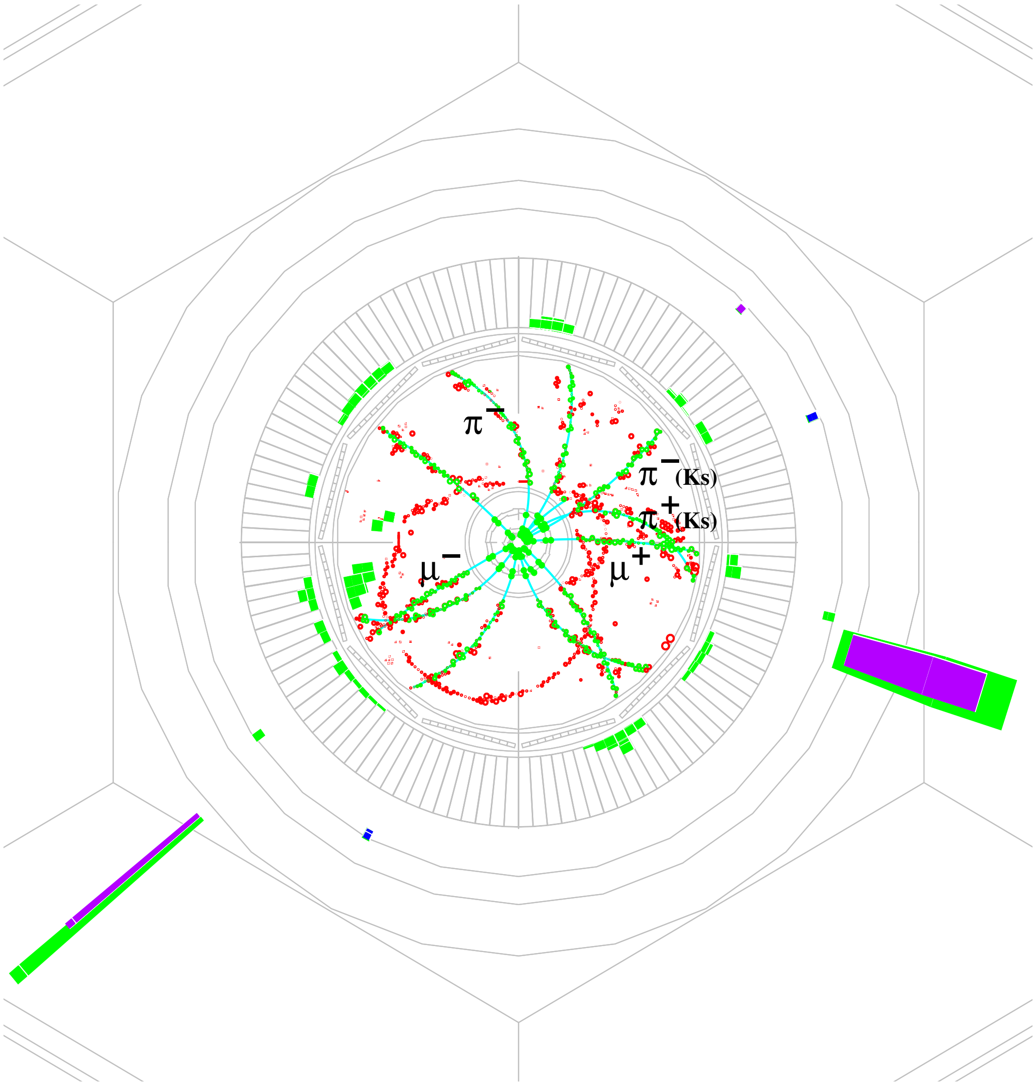


Figure 10.39: Event display for the $B^+ \to K^{*+}\mu^+\mu^-$ ($K^{*+} \to K^0_S\pi^+$) candidate event in the signal box: run 16024. The pions with Ks in parentheses come from the K^0_S.

10.3 Upper Limits on Combined Modes

The data samples containing $B^+ \to K^+e^+e^-$, $B^+ \to K^+\mu^+\mu^-$, $B^0 \to K^0_Se^+e^-$, and $B^0 \to K^0_S\mu^+\mu^-$ candidate events are statistically independent, and can therefore be simply averaged to reach a better sensitivity to the $B \to K\ell^+\ell^-$ mode. In order to combine the different modes, a combined likelihood function is formed as a product of the individual likelihoods, whose fit parameter $n_{\rm sig}$ is corrected for the differences in the mode-by-mode efficiency and secondary branching fractions. The fit thus gives the efficiency-corrected average signal event yield. For example, the total likelihood $\mathcal{L}^{\rm tot}$ for combining $B^+ \to K^+e^+e^-$ and $B^+ \to K^+\mu^+\mu^-$ is

$$\begin{aligned}\log\mathcal{L}^{\rm tot} \;=\;& \log\mathcal{L}^{Ke^+e^-}\left(\epsilon_{Ke^+e^-}n_{\rm sig}^{K\ell^+\ell^-}, n_{\rm cont}^{Ke^+e^-}, n_{B\overline{B}}^{Ke^+e^-}\right) + \\ & \log\mathcal{L}^{K\mu^+\mu^-}\left(\epsilon_{K\mu^+\mu^-}n_{\rm sig}^{K\ell^+\ell^-}, n_{\rm cont}^{K\mu^+\mu^-}, n_{B\overline{B}}^{K\mu^+\mu^-}\right), \qquad (10.3.1)\end{aligned}$$

where $n_{\rm sig}^{K\ell^+\ell^-}$ is the efficiency-corrected average signal yield, $\epsilon_{Ke^+e^-}$ is the signal efficiency for $B^+ \to K^+e^+e^-$, and $\epsilon_{K\mu^+\mu^-}$ is the signal efficiency for $B^+ \to K^+\mu^+\mu^-$. After determining $n_{\rm sig}^{K\ell^+\ell^-}$, the 90% C.L. upper limit on the combined branching fraction is placed in the way described in Section 10.1.

In principle, the K^* modes, $B^0 \to K^{*0}e^-e^+$, $B^0 \to K^{*0}\mu^-\mu^+$, $B^+ \to K^{*+}e^-e^+$, and $B^+ \to K^{*+}\mu^-\mu^+$, can also be combined in the same way, but in this case the procedure is less straightforward because the predicted branching fractions for the electron and muon channels are not the same due to the presence of a pole at $q^2 = 0$ (compare, for example, Figs. 6.3d and 6.4d). A possible way of solving this problem, used, in particular, in the analysis described in Ref. [58], is to combine the electron and muon modes only in the $m_{\ell^+\ell^-} > 0.5$ GeV$/c^2$ region, thus removing the pole. The approach adapted in this analysis is to combine these modes using a constraint on the relative predicted branching fractions, $\mathcal{B}(B \to K^*e^+e^-)/\mathcal{B}(B \to K^*\mu^+\mu^-) = 2.3/1.9 = 1.2$, where the branching fractions for $B \to K^*e^+e^-$ and for $B \to K^*\mu^+\mu^-$ are taken from the model of Ali *et al.* [12] and are summarized in Table 5.3. This conservative approach in effect establishes the 90% C.L. upper limit on the $B \to K^* e^+ e^-$ branching fraction.

The combined fits described above yield the following central values for the branching fractions:

$$\begin{aligned} \mathcal{B}(B \to K\ell^+\ell^-) &= (0.0 \pm 0.3) \times 10^{-6} \\ \mathcal{B}(B \to K^*\ell^+\ell^-) &= (0.7 \pm 1.1) \times 10^{-6}, \end{aligned} \tag{10.3.2}$$

where the error is statistical only. Since no evidence for a signal has been observed, 90% C.L. upper limits are placed using the procedure described in Section 10.1:

$$\begin{aligned} \mathcal{B}(B \to K\ell^+\ell^-) &< 0.6 \times 10^{-6} \\ \mathcal{B}(B \to K^*\ell^+\ell^-) &< 2.5 \times 10^{-6}. \end{aligned} \tag{10.3.3}$$

There upper limits have been calculated taking into account the systematic uncertainties discussed in detail in Chapter 11.

Chapter 11

Systematic Uncertainties

As was discussed in Chapter 10, no evidence for signal has been seen in any of channels, and therefore only upper limits can be set. There are two categories of systematic effects that affect the limits: the uncertainties on the the extracted number of signal events, and the uncertainties on the signal efficiency. An additional source of systematic uncertainty is the counting of the produced B mesons in the data sample. This uncertainty is known to be 1.6% [64].

The sources of systematic error affecting the signal efficiency include:

- The systematic uncertainty on the tracking efficiency. It is 1.2% for GoodTracksLoose and 1.3% for GoodTracksVeryLoose [68]. A conservative systematic uncertainty of 2.0% is assigned for the tracking efficiency of the relatively slow charged pions from the K^* decays ($K^{*0} \to K^+\pi^-$ and $K^{*+} \to K^0_S\pi^+$). For K^0_S mesons, the systematic error is also 1.3% per daughter, which implies that it is 2.6% per K^0_S.

- The systematic uncertainties associated with electron and muon identification. These uncertainties are summarized in [67]. For the very tight lepton identification criteria, they are 2.7% for the electron channels, and 2.0% for the muon channels.

- The very tight kaon identification criteria, the systematic uncertainty on which is 2% [78]. The same uncertainty is assumed for pion identification, since it is in fact a kaon veto.
- The uncertainty on the signal efficiency arising from the Monte Carlo statistics
- The uncertainty on the modeling of the Fisher discriminant and $B\overline{B}$ suppression cuts in the Monte Carlo. The efficiencies of these cuts are compared in the J/ψ control sample in the data and in the signal SP3 $J/\psi(K, K^*)$ Monte Carlo, and the difference is taken as a measure of the systematic uncertainty.
- The uncertainty on the modeling of the K^0_S selection efficiency for the K^0_S modes, which is determined in exactly the same way as the efficiency of the Fisher and $B\overline{B}$ suppression selection criteria.
- The uncertainty on the signal efficiency arising from model dependence. This uncertainty can be estimated by using several form factor models. Two models other than that of Ref. [12] have been considered: a QCD sum rules model of Ref. [49], and a quark model of Ref. [50]. For each model, 100,000 Pravda Monte Carlo signal events were generated and the corresponding signal efficiencies were compared. Figure 11.1 shows the distribution of $m^2_{\ell^+\ell^-}$ in generator-level Monte Carlo for all three models. It is evident that the differences between the models are at the level of a few percent. However, since the models considered may not span the entire range of theoretical possibilities, the uncertainty arising from model dependence is conservatively taken to be 5%.

The sources of systematic uncertainty affecting the extraction of the signal yield are:

- There is an uncertainty associated with the parametrization of the background shape. It is evaluated by taking the largest difference between the nominal 90% C.L. signal yield and the signal yield obtained assuming the background is purely $B\overline{B}$ or continuum. In addition, there is a statistical uncertainty associated with the procedure used to extract the 90% C.L. event yields. Overall,

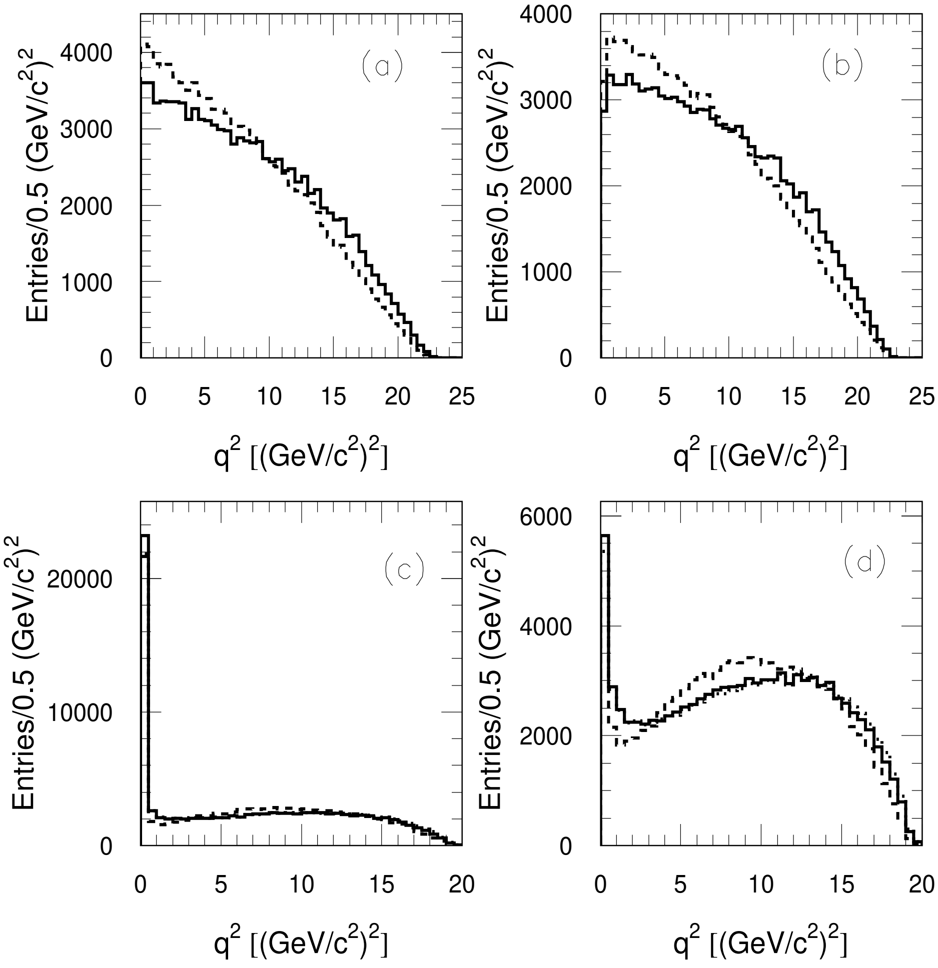


Figure 11.1: The dilepton mass squared distributions for three different theoretical models: the Light Cone QCD Sum Rule model of Ref. [12] (solid line), the QCD Sum Rules model of Ref. [49] (dashed line), and the quark model of Ref. [50] (dotted line). Generator-level Monte Carlo for (a) $B^+ \to K^+e^+e^-$, (b) $B^+ \to K^+\mu^+\mu^-$, (c) $B^0 \to K^{*0}e^+e^-$, and (d) $B^0 \to K^{*0}\mu^+\mu^-$.

the uncertainty on the background shapes is taken to be 7.0% for all of the modes.

- The uncertainty associated with the parametrization of the signal shapes in the fit is estimated by using the parameters from fits to the J/ψ control sample in data.

  Let us describe the procedure in more detail. First, a clean sample of J/ψ events is selected by applying all of the analysis event selection criteria including *either* ΔE *or* m_{ES} cut. Suppose a sample has been chosen where the m_{ES} cut has been applied. Then the parameters describing the ΔE shape can be found from the fit. Repeating the procedure using the sample where the ΔE cut has been applied, one can determine the parameters describing the m_{ES} shape in data. The measure of the systematic uncertainty on the signal shape is the largest difference between the nominal 90% C.L. signal yield and the signal yield obtained using both the central values of the new signal shape parameters and their variations within $\pm\sigma$ of the statistical error. This uncertainty is found to be 8.0% for all of the modes.

Table 11.1 is a summary of all systematic uncertainties for the four charged track modes; Table 11.2, for the modes with a K^0_S; and Table 11.3, for the combined modes. For establishing the systematic uncertainty on the combined modes, the errors on the lepton identification and Monte Carlo statistics are assumed to be uncorrelated for the electron and muon channels.

The uncertainties affecting the signal efficiency are added in quadrature. The total relative systematic uncertainty $(\Delta\mathcal{B}/\mathcal{B})_\epsilon$ on the signal efficiency also includes the 1.6% uncertainty on the B counting. The uncertainties on the extraction of the signal yield are also added in quadrature to obtain $(\Delta\mathcal{B}/\mathcal{B})_{\text{yield}}$. In setting the 90% C.L. upper limits the systematic errors on the signal efficiency and on the signal yield are added in quadrature, and the limit is inflated by the corresponding factor.

Error	$(\Delta\mathcal{B}/\mathcal{B})$ (%)			
	$K^+e^+e^-$	$K^+\mu^+\mu^-$	$K^{*0}e^+e^-$	$K^{*0}\mu^+\mu^-$
B counting	1.6	1.6	1.6	1.6
Tracking	3.7	3.7	5.6	5.6
e, μ ID	2.7	2.0	2.7	2.0
K, π ID	2.0	2.0	4.0	4.0
Fisher cut	2.0	2.0	2.0	2.0
$B\overline{B}$ cut	3.0	3.0	3.0	3.0
MC signal statistics	2.8	3.3	3.9	4.9
Model dependence	5.0	5.0	5.0	5.0
TOTAL for ϵ	8.6	6.5	10.5	10.8
	Uncertainties on yield for 90% C.L. limit			
Bkg shapes	7.0	7.0	7.0	7.0
Signal shapes	8.0	8.0	8.0	8.0
TOTAL for yield	10.6	10.6	10.6	10.6

Table 11.1: Systematic uncertainties for the charged track modes. The uncertainties are of two types: those affecting the signal efficiency ϵ, and those affecting the 90% C.L. signal yield.

Error	$(\Delta\mathcal{B}/\mathcal{B})$ (%)			
	$K^0_S e^+e^-$	$K^0_S \mu^+\mu^-$	$K^{*+}_{K^0_S} e^+e^-$	$K^{*+}_{K^0_S} \mu^+\mu^-$
B counting	1.6	1.6	1.6	1.6
Tracking	5.0	5.0	6.8	6.8
e, μ ID	2.7	2.0	2.7	2.0
K, π ID	0.0	0.0	2.0	2.0
Fisher cut	2.0	2.0	2.0	2.0
$B\overline{B}$ cut	3.0	3.0	3.0	3.0
K^0_S selection	4.0	4.0	4.0	4.0
MC signal statistics	3.0	3.5	4.1	5.8
Model dependence	5.0	5.0	5.0	5.0
TOTAL for ϵ	9.3	11.4	11.4	12.0
	Uncertainties on yield for 90% C.L. limit			
Bkg shapes	7.0	7.0	7.0	7.0
Signal shapes	8.0	8.0	8.0	8.0
TOTAL for yield	10.6	10.6	10.6	10.6

Table 11.2: Systematic uncertainties for the modes with a K^0_S. The uncertainties are of two types: those affecting the signal efficiency ϵ, and those affecting the 90% C.L. signal yield.

Error	$(\Delta\mathcal{B}/\mathcal{B})$ (%)	
	$B \to K\ell^+\ell^-$	$B \to K^*\ell^+\ell^-$
B counting	1.6	1.6
Tracking	4.0	5.9
e, μ ID	1.7	1.7
K, π ID	1.5	3.5
Fisher cut	2.0	2.0
$B\overline{B}$ cut	2.5	4.0
K^0_S selection	2.0	2.1
MC signal statistics	2.0	3.0
Model dependence	5.0	4.0
TOTAL for ϵ	8.5	10.6
	Uncertainties on yield for 90% C.L. limit	
Bkg shapes	7.0	7.0
Signal shapes	8.0	8.0
TOTAL for yield	10.6	10.6

Table 11.3: Systematic uncertainties for the combined modes. The uncertainties are of two types: those affecting the signal efficiency ϵ, and those affecting the 90% C.L. signal yield.

Chapter 12

Future Prospects

A number of interesting and important discoveries will be made in the $B \to X_s \ell^+ \ell^-$ decays with larger data samples. This chapter discusses some of the future prospects for these modes at *BABAR* and, for comparison, at the Tevatron.

How large a data set is required to perform the measurement of the exclusive branching fractions? Figure 12.1 shows the probability for background to fluctuate to the number of observed events as a function of the number of signal events μ_{sig}. This probability is described in more detail in Appendix B. It is apparent that for $\mu_{\text{bkg}} = \mu_{\text{sig}}$, which is roughly the case for the analysis described in this dissertation, the probability reaches $\mathcal{O}(10^{-4})$ for about 40 signal events. Assuming these 40 events are the sum of the four charged track modes, 100 fb^{-1} of data are needed to make the discovery. Obtaining this much data is no longer a remote possibility: *BABAR* is planning to obtain a data set of this size in a very near future. Table 12.1 summarizes the expectations for *BABAR* and CDF for the next several years. The prediction for CDF is a scaling from [55]. The reasoning goes as follows. It was pointed out in [55] that CDF expected 0.5 signal $K^{*0}\mu^+\mu^-$ events in the approximately 100 pb^{-1} of their Run 1 data set. Assuming that with an updated CDF detector and a better trigger CDF could expect about twice this amount (that is, 1 event in 100 pb^{-1}), one can obtain the 20 $K^{*0}\mu^+\mu^-$ events in 2 fb^{-1} that are listed in Table 12.1.

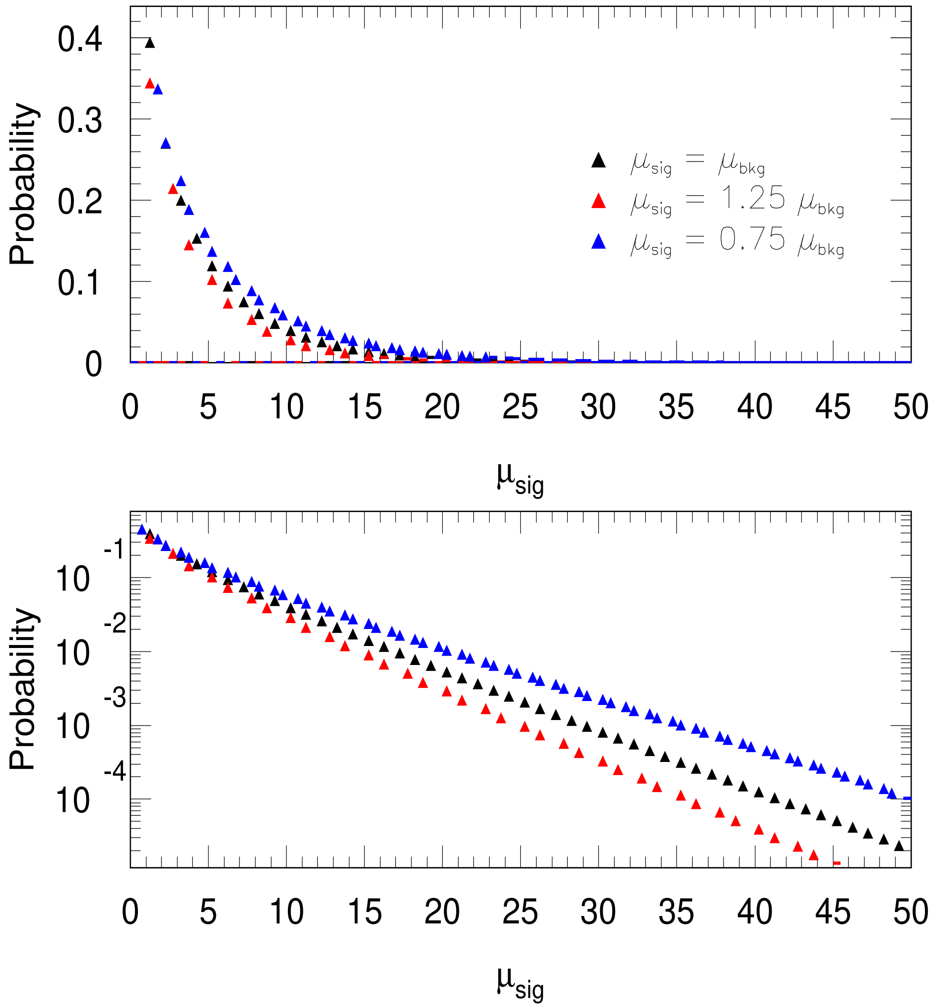


Figure 12.1: Probability for background $\mu_{\rm bkg}$ to fluctuate to the number of observed events as a function of the number of signal events $\mu_{\rm sig}$ on a linear (upper plot) and logarithmic scale (lower plot). Three different scenarios are considered: $\mu_{\rm bkg} = \mu_{\rm sig}$ (middle curve), $\mu_{\rm bkg} = 1.25\ \mu_{\rm sig}$ (lower curve), and $\mu_{\rm bkg} = 0.75\ \mu_{\rm sig}$ (upper curve).

Year	*BABAR* $\int \mathcal{L}$ dt (fb^{-1})	*BABAR* # $K^{*0}\ell^+\ell^-$ evts	CDF $\int \mathcal{L}$ dt (fb^{-1})	CDF # $K^{*0}\mu^+\mu^-$ evts
2003	230	60	2	20
2007	1300	340	15	150

Table 12.1: Predicted yields of $K^{*0}\ell^+\ell^-$ events for the next several years for *BABAR* and CDF.

As was mentioned in Chapter 2, apart from the branching fractions, the $B \to (K, K^*)\ell^+\ell^-$ decays offer a number of other interesting quantities that can be used for testing the Standard Model. Consider the forward-backward asymmetry A_{FB} in the $B \to K^*\ell^+\ell^-$ decays. In Section 2.2, it was pointed out that, unlike the decay rates, the zero of the asymmetry is very clean theoretically: it is known to about 10%. Figure 12.2 shows the forward-backward asymmetry as a function of q^2 for 50,000 generated $B^0 \to K^{*0}\mu^+\mu^-$ signal events, which are required to be truth tagged. The distributions are fitted to a straight line to determine the zero of the asymmetry to first approximation. The zeros of the asymmetries determined from the fits are summarized in Table 12.2. Therefore, with the amount of data

Mode	q_0^2 (GeV/c^2)^2
$B^0 \to K^{*0}\mu^+\mu^-$	2.81 ± 0.68
$\overline{B}^0 \to \overline{K}^{*0}\mu^+\mu^-$	3.75 ± 0.58

Table 12.2: Zeros of the forward-backward asymmetry for Monte Carlo-truth tagged signal Pravda Monte Carlo $B^0 \to K^{*0}\mu^+\mu^-$. The errors are statistical only.

corresponding to about 33 ab^{-1}, it is possible to determine the asymmetry to $\sim$ 15-25% accuracy. It is unlikely, however, that either the B factories or the Tevatron will obtain enough data to make this measurement; but the LHCb ([13], [79]) and BTeV [14] should be able to measure it.

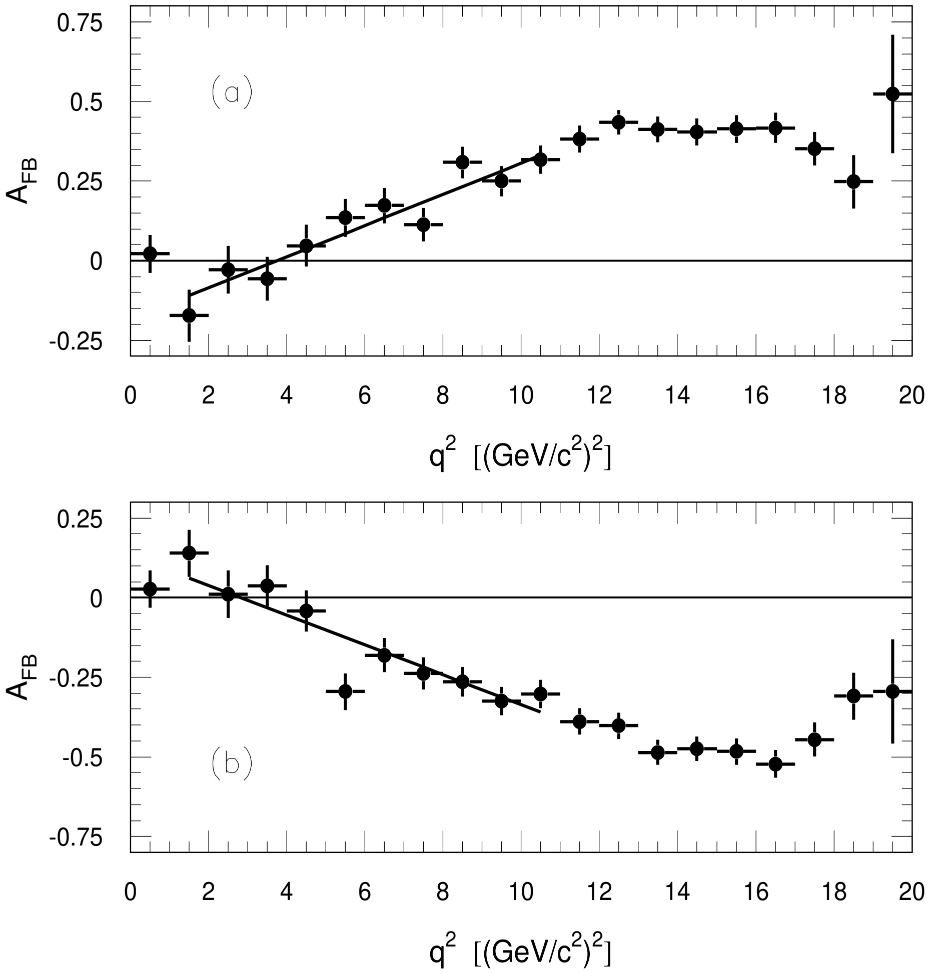


Figure 12.2: Fit to the forward-backward asymmetry for Monte Carlo-truth tagged Pravda Monte Carlo signal (a) $\overline{B}^0 \rightarrow \overline{K}^{*0}\mu^+\mu^-$ and (b) $B^0 \rightarrow K^{*0}\mu^+\mu^-$. The asymmetry was determined using 50,000 of Pravda Monte Carlo signal events and fitting it to a straight line. The errors shown are statistical.

Chapter 13

Conclusion

The search for the flavor-changing neutral current processes $B \to K\ell^+\ell^-$ and $B \to K^*\ell^+\ell^-$, described in this dissertation, did not yield any evidence for a signal. The established 90% C.L. upper limits on the combined modes, $B \to K\ell^+\ell^-$ and $B \to K^*\ell^+\ell^-$, are significantly better than previously published results ([58], [3]), and are at the level of the theoretical predictions from the model of Ali *et al.* [12], summarized in Table 5.3. Figure 13.1 compares the established 90% C.L. upper limits with those set by CLEO [58], CDF [55], and Belle [3], as well as with the theoretical predictions from the model of Ali *et al.* [12]. The theoretical points include a 35% uncertainty.

The analysis reached the expected sensitivity for the presently considered data sample of 22.7×10^6 $B\overline{B}$ events. With larger data samples, which are expected to be available in the next few years, experiments will be able to make a conclusive statement on the branching fractions of $B \to (K, K^*)\ell^+\ell^-$ and their level of agreement with the Standard Model. However, the exploration of these processes will not stop with their discovery. A number of other interesting observables allowing one to probe the Standard Model will be measured in the next decade or so, with the advent of experiments such as LHCb [13] and BTeV [14].

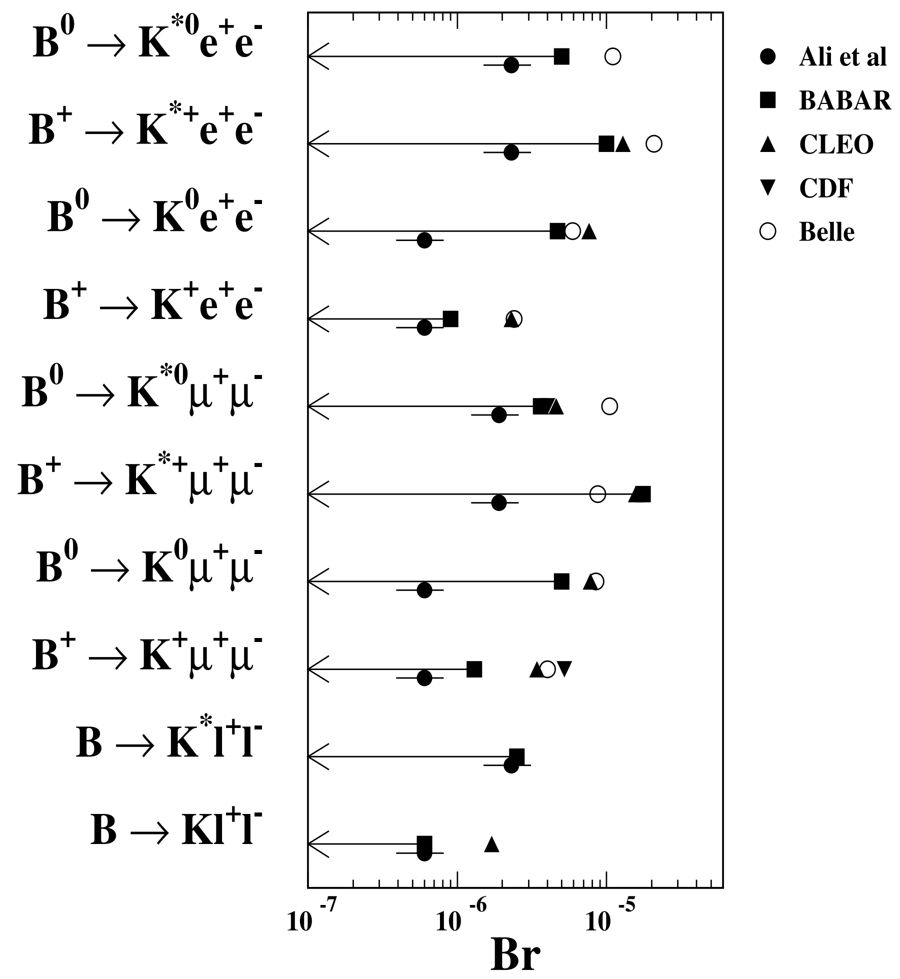


Figure 13.1: Summary of the theoretical predictions from the model of Ali *et al.* [12] and 90% C.L. upper limits on the branching fractions for the individual and combined exclusive $B \to (K, K^*)\ell^+\ell^-$ modes, established by *BABAR* (squares), CLEO [58] (upward pointing triangles), CDF [55] (downward pointing triangles), and Belle [3] (unfilled circles). The theoretical predictions (filled circles) are shown with error bars to emphasize that the predictions are known to 35%.

Appendix A

Pravda Monte Carlo

Pravda Monte Carlo is not a standard *BABAR* tool, and must therefore be described in some detail here. It is a fast parametrized simulation that takes the true parameters of a given track or neutral cluster and smears them, in a rough but plausible way parametrizing the response of the detector.

A.1 Charged Tracks

The track finding probability is parametrized in the following way. All tracks are required to be within -0.93 $< \cos\theta <$ 0.975, where $\cos\theta = p_z/p$ is the cosine of the polar angle of the track in the laboratory frame. If the track is within these angular limits, the probability is given by:

$$P = \left[1 - \exp\left(\frac{\cos\theta - \cos\theta_{\max}}{0.005}\right)\right]\left[1 - \exp\left(\frac{-(\cos\theta - \cos\theta_{\min})}{0.005}\right)\right] \times \left(1. - \frac{0.005}{p_t}\right)\left[1 - \exp\left(-\frac{(p_t - p_t^{\mathrm{th}})}{0.015}\right)\right], \quad \text{(A.1.1)}$$

where $\cos\theta_{\max}$ = 0.975, $\cos\theta_{\min}$ = -0.93, p_t is the transverse momentum, and p_t^{th} is the transverse momentum threshold, tabulated in Table A.1. The first two terms simply enforce the (somewhat smoothed out) acceptance requirement, and the last

Particle Type	p_t^{th} (GeV/c)
π	0.05
K	0.15
p	0.25
μ	0.05
e	0.04

Table A.1: Transverse momentum thresholds for different particle species used for parametrizing tracking in Pravda Monte Carlo.

term describes the turn-on of the tracking efficiency at $p_t = p_t^{\mathrm{th}}$, as shown in Fig. A.1a.

The position error matrix is diagonal, with all of the elements set equal to

$$U_{ii} = \left(0.0015 \ + \ \frac{0.005}{p}\right)^2, \ i = 1, 2, 3 \tag{A.1.2}$$

The position of the charged track vertex is smeared with a Gaussian of $\sigma = U_{ii}$. The momentum error matrix is also diagonal, with all of the elements set equal to

$$U_{jj} = \frac{1}{2}\sigma_{p_t}^2, \ j = 4, 5, 6 \tag{A.1.3}$$

where σ_{p_t} is given by [60]:

$$\frac{\sigma_{p_t}}{p_t} = (0.27 + 0.17\ p_t)\%, \text{ if } p_t > 0.7 \text{ GeV}/c. \tag{A.1.4}$$

For lower momenta, σ_{p_t} is multiplied by an extra term:

$$[\ 1. \ + \ 2.\ (0.7 - p_t)\ ], \text{ if } p_t < 0.7 \text{ GeV}/c, \tag{A.1.5}$$

and for very low momenta, by yet another term:

$$\frac{0.18}{p_t}\ [\ 1 \ + \ (0.18 - p_t)\ 80.0\ ], \text{ if } p_t < 0.18 \text{ GeV}/c. \tag{A.1.6}$$

The plot of σ_{p_t}/p_t is given in Fig. A.1b. The momentum of the track is smeared with a Gaussian of $\sigma = U_{jj}/\sqrt{2}$.

For electrons, in order to simulate the effects of radiation, the momentum p is shifted by an exponential sampling term of magnitude 0.8 p.

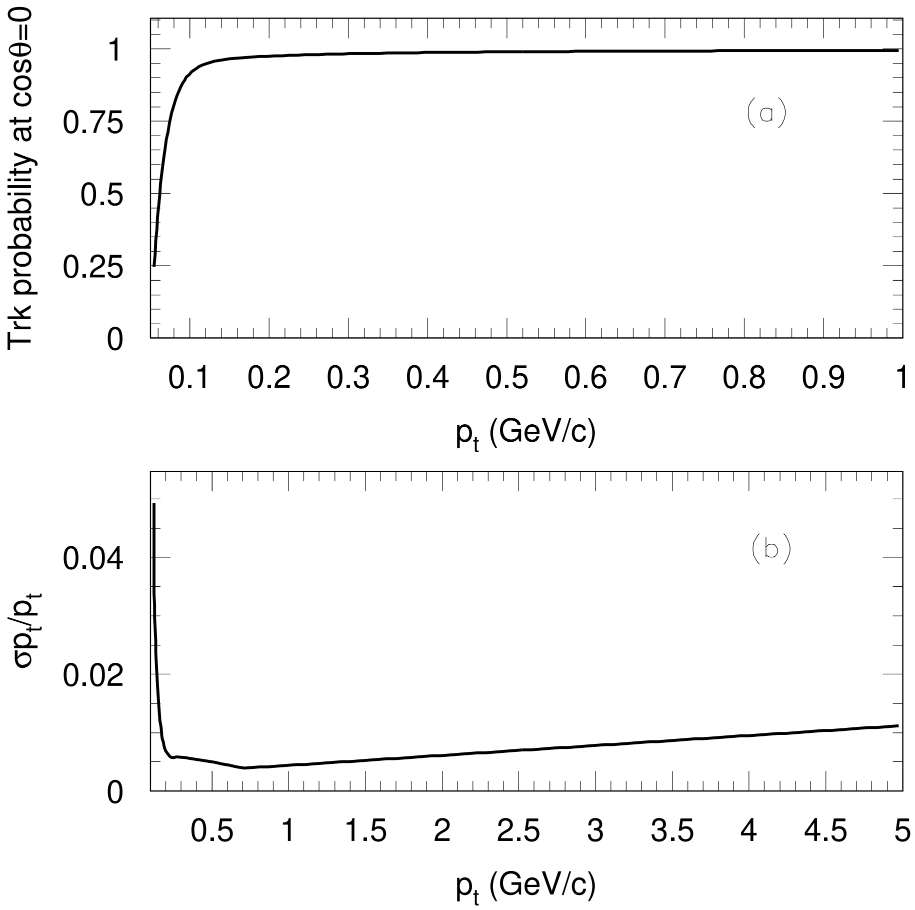


Figure A.1: (a) The parametrization of the tracking efficiency as a function of transverse momentum for pions at $\cos\theta = 0$. Note the turn-on of the efficiency at $p_t^{\rm th} = 0.05$ GeV/c. (b) The parmetrization of the momentum resolution.

A.2 Neutral Clusters

The parametrization of neutral clusters is very similar to charged tracks. The neutral finding probability is assumed to be perfect within the EMC acceptance

(-0.8 $< \cos\theta <$ 0.97).

The momentum magnitude is smeared with a Gaussian of sigma [74]:

$$\frac{\sigma_E}{E} = \frac{2.32\%}{\sqrt[4]{E}} \oplus 1.85\%, \tag{A.2.1}$$

where E is the photon energy. The central value is shifted by an exponential sampling term of amplitude 0.4 E. The momentum components, p_x, p_y, and p_z, are re-scaled to the new magnitude of the momentum. They are subsequently smeared by the angular resolution:

$$\begin{aligned} \Delta p_x &= p\,(-\sin\theta\sin\phi\ \sigma_\phi + \cos\theta\cos\phi\ \sigma_\theta) \\ \Delta p_y &= p\,(\sin\theta\cos\phi\ \sigma_\phi + \cos\theta\sin\phi\ \sigma_\theta) \\ \Delta p_z &= p\,(-\sin\theta\ \sigma_\theta), \end{aligned} \tag{A.2.2}$$

where [74]:

$$\sigma_\phi = \sigma_\theta = \frac{3.87}{\sqrt{E}}\ \text{mrad} \tag{A.2.3}$$

All three components of the momentum error matrix are smeared by a Gaussian with $\sigma = \sigma_E$.

The position vector of the photon is in fact set equal to the momentum vector, since what matters here is that it points in the right direction for Monte Carlo truth matching. The position error matrix is smeared by:

$$\begin{aligned} U(1,1) &= \sin\phi^2\sin\theta^2\ \sigma_\phi^2 + \cos\phi^2\cos\theta^2\ \sigma_\theta^2 \\ U(2,2) &= \sin\theta^2\cos\phi^2\ \sigma_\phi^2 + \sin\phi^2\cos\theta^2\ \sigma_\theta^2 \\ U(3,3) &= \sin\theta^2\ \sigma_\theta^2 \end{aligned} \tag{A.2.4}$$

An attempt has been made to simulate the energy deposition by K_L^0's. The deposited energy is set to 0.125 times the original K_L^0 energy $E_{K_L^0}$, and smeared with a Gaussian with $\sigma = 0.25\ E_{K_L^0}$ as long as the resulting energy is greater than 0.05 GeV and less than $E_{K_L^0}$. Everything else is a carbon copy of the γ case, except that the ϕ resolution is made worse by a factor of 3.

In order to make Pravda Monte Carlo more realistic, 2% of the tracks are added to the list of tracks several times as "ghost tracks". Additionally, in 2% of the events for the tracks with energies greater than 0.2 GeV a split-off is simulated as an energy deposition distributed as a Gaussian with $\sigma = 0.025$ GeV. The split-offs are subsequently treated just like photons.

An early version of Pravda attempted to simulate all of the quantities used for particle identification, such as the number of interaction lengths traversed by a particle in the detector, the Cherenkov angle in the DIRC, etc.. With the increased complexity of the criteria used for particle ID, it was decided to use the PID killing tables described in Section 5.3 to generate lists of electrons, muons, and kaons.

Figure A.2 shows the comparison of several distributions of the kaon candidate in the $B^+ \rightarrow K^+e^+e^-$ mode in a generic $B\overline{B}$ sample of Pravda Monte Carlo and SP3 Monte Carlo of the same size. The agreement in both the shapes and the normalization is very good.

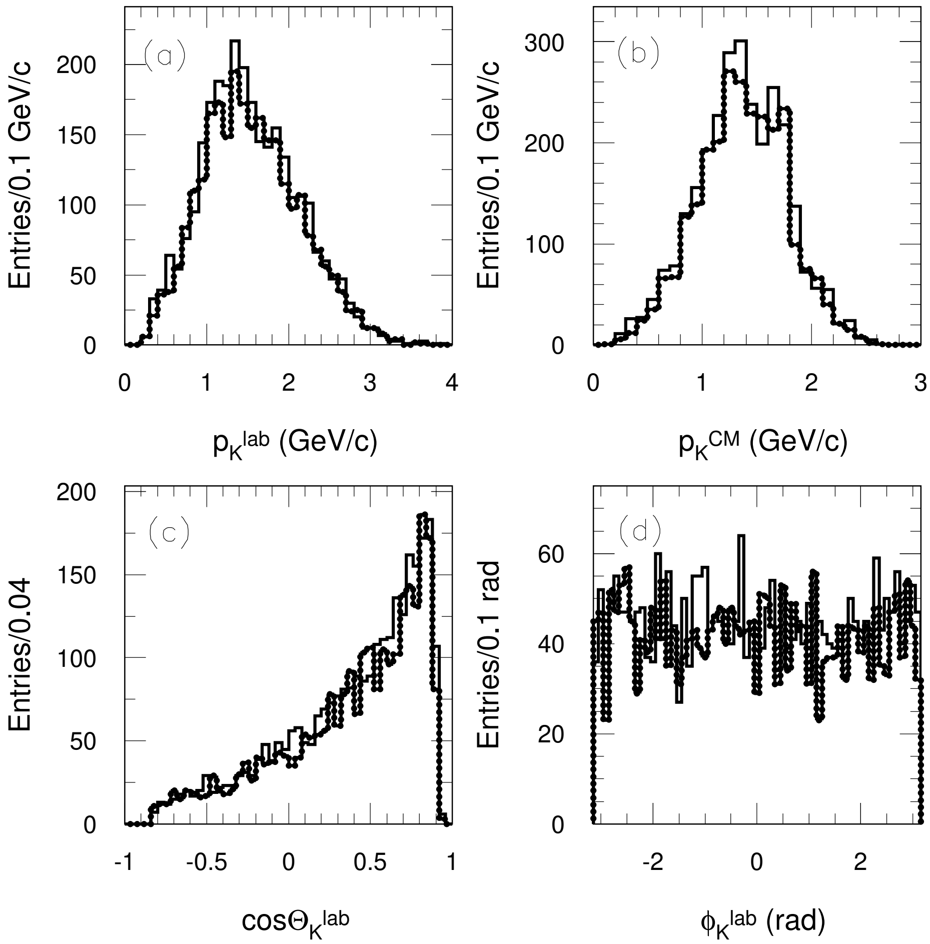


Figure A.2: The comparison of Pravda Monte Carlo with SP3 simulation. The plots show several quantities for the kaon in the $B^+ \to K^+e^+e^-$ mode in generic $B\overline{B}$ Monte Carlo samples of Pravda (dashed lines) and SP3 (solid lines) events. (a) and (b) show kaon momentum spectra in the laboratory and CM frames, respectively. (c) and (d) show polar and azimuthal angles, respectively. The normalization is absolute. The event selection criteria applied are tight electron ID on both e^+ and e^-, very tight kaon ID on the kaon, and $R_2 < 0.5$.

Appendix B

Event Selection Optimization Variable

In order to properly optimize the event selection criteria so as to achieve the best sensitivity to signal, the question that needs to be answered is this: what is the probability that, for a given number of signal and background events, the number of background events is greater than or equal to the number of observed events? One can show that this probability is given by the following equation:

$$P(\mu_{\rm sig}, \mu_{\rm bkg}) = \sum_{N_{\rm sig}, N_{\rm bkg}} \sum_{N_{\rm obs}} \sum_{N \geq N_{\rm obs}} \frac{\mu_{\rm bkg}^{N}}{N!} e^{-\mu_{\rm bkg}} \delta_{N_{\rm obs}, N_{\rm sig}+N_{\rm bkg}} \times$$

$$\frac{\mu_{\rm bkg}^{N_{\rm bkg}}}{N_{\rm bkg}!} e^{-\mu_{\rm bkg}} \frac{\mu_{\rm sig}^{N_{\rm sig}}}{N_{\rm sig}!} e^{-\mu_{\rm sig}} =$$

$$\sum_{N_{\rm obs}} \sum_{N \geq N_{\rm obs}} e^{-2\mu_{\rm bkg}} e^{-\mu_{\rm sig}} \sum_{N_{\rm sig}=0}^{N_{\rm obs}} \frac{\mu_{\rm bkg}^{N} \, \mu_{\rm bkg}^{N_{\rm obs}-N{\rm sig}} \mu_{\rm sig}^{N_{\rm sig}}}{N! N_{\rm sig}! (N_{\rm obs} - N_{\rm sig})!} =$$

$$\sum_{N_{\rm obs}} \sum_{N \geq N_{\rm obs}} e^{-2\mu_{\rm bkg}} e^{-\mu_{\rm sig}} \frac{\mu_{\rm bkg}^{N+N_{\rm obs}}}{N! N_{\rm obs}!} \left[\sum_{N_{\rm sig}=0}^{N_{\rm obs}} \left(\frac{\mu_{\rm sig}}{\mu_{\rm bkg}} \right)^{N_{\rm sig}} \frac{N_{\rm obs}!}{N_{\rm sig}! (N_{\rm obs} - N_{\rm sig})!} \right] =$$

$$\sum_{N_{\rm obs}} \sum_{N \geq N_{\rm obs}} e^{-2\mu_{\rm bkg}} e^{-\mu_{\rm sig}} \frac{\mu_{\rm bkg}^{N+N_{\rm obs}}}{N! N_{\rm obs}!} \left[\frac{(\mu_{\rm bkg} + \mu_{\rm sig})^{N_{\rm obs}}}{\mu_{bacl}^{N_{\rm obs}}} \right] =$$

$$e^{-2\mu_{\text{bkg}}} e^{-\mu_{\text{sig}}} \sum_{N=0}^{\infty} \sum_{N_{\text{obs}}=0}^{N} \frac{(\mu_{\text{bkg}} + \mu_{\text{sig}})^{N_{\text{obs}}} \mu_{\text{bkg}}^{N}}{N_{\text{obs}}! N!} =$$

$$e^{-2\mu_{\text{bkg}}} e^{-\mu_{\text{sig}}} \sum_{N=0}^{\infty} \frac{\mu_{\text{bkg}}^{N}}{N!} \left[\frac{e^{\mu_{\text{bkg}}+\mu_{\text{sig}}}}{N!} \, \Gamma(N+1, \mu_{\text{sig}} + \mu_{\text{bkg}}) \right] =$$

$$e^{-\mu_{\text{bkg}}} \sum_{N=0}^{\infty} \frac{\mu_{\text{bkg}}^{N} \, \Gamma(N+1, \mu_{\text{sig}} + \mu_{\text{bkg}})}{\Gamma(N+1)^2},$$

where μ_{sig} and μ_{bkg} are the mean number of signal and background events, respectively; Γ is the gamma function; and N is the index over the number of events.

In practice, it is sufficient to calculate only the first several terms of the sum ($N = 15$ is used in the analysis described in this dissertation), since afterwards adding more terms does not significantly affect the result.

It has been observed that in most cases minimizing this probability gives the same results as maximizing signal squared over signal plus background.

Bibliography

[1] Particle Data Group, D. E. Groom *et al.*, Eur. Phys. Jour. **C15**, 1 (2000).

[2] D. Cronin-Hennessy *et al.* (CLEO Collaboration), CLNS 99/1650, CLEO 99-18; preprint hep-ex/0001010.

[3] T. Iijima (Belle Collaboration) *B meson rare decays*, presented at the 4th International Conference On B Physics And CP Violation (BCP4), Ise-Shima, Japan, 19-23 Feb, 2001; preprint hep-ex/0105005.

[4] A. Hoecker (*BABAR* Collaboration) *Hadronic B decays to charmless final states and to* $J/\psi K^*$, presented at the 4th International Conference On B Physics And CP Violation (BCP4), Ise-Shima, Japan, 19-23 Feb, 2001; preprint hep-ex/0104028.

[5] R. Fleischer, Acta Phys. Polon. **B31** 2633 (2000);
M. Gronau, Nucl. Phys. Proc. Suppl. **65** 245 (1998).

[6] T. E. Coan *et al.* (CLEO Collaboration), Phys. Rev. Lett. **84**, 5283 (2000).

[7] S. Ahmed *et al.* (CLEO Collaboration), CLEO CONF 99-10; preprint hep-ex/9908022.

[8] Y. Ushiroda (Belle Collaboration) *Radiative B meson Decay*, presented at the 4th International Conference On B Physics And CP Violation (BCP4), Ise-Shima, Japan, 19-23 Feb, 2001; preprint hep-ex/0104045.

[9] M. Convery (*BABAR* Collaboration) *Charmless B Decays from BABAR ($B^0 \to K^{*0}\gamma$ and $B^0 \to \pi^+\pi^-, K^+\pi^-, K^+K^-$)*, presented at the 15th Les Rencontres De Physique De La Vallee D'Aoste: Results And Perspective In Particle Physics, La Thuile, Valle d'Aosta, Italy, 4-10 Mar 2001; preprint hep-ex/0106043.

[10] R. Ammar *et al.* (CLEO Collaboration), Phys. Rev. Lett. **71** 674 (1993).

[11] R. Barate *et al.* (ALEPH Collaboration), Phys. Lett. **B429**, 169 (1998).

[12] A. Ali, P. Ball, L. T. Handoko, G. Hiller, Phys. Rev. **D61**, 074024 (2000).

[13] S. Amato *et al.* (LHCb Collaboration), *LHCb Technical Proposal*, CERN-LHCC-98-4, CERN-LHCC-P-4 (1998).

[14] A. Kulyavtsev *et al.* (BTeV Collaboration), *BTeV - an Experiment to Measure Mixing, CP Violation, and Rare Decays of Beauty and Charm at the Fermilab Collider*; preprint hep-ex/0006037.

[15] G. Isidori, *Z penguins and rare B decays*, presented at the 30th International Conference On High-Energy Physics (ICHEP 2000), Osaka, Japan , Jul 27 - Aug 2, 2000; preprint hep-ph/0009024;
G. Buchalla, G. Hiller, G. Isidori, Phys. Rev. **D63** 014015 (2001).

[16] M. Narain, In: C. Campagnari (ed.), *Proceedings of the 7th International Symposium on Heavy Flavors, Santa Barbara, California, 1997* (World Scientific, Singapore, 1999) p. 379.

[17] T. Draper, Nucl.Phys.Proc.Suppl. **73** 43 (1999).

[18] A. Ali, In: C. Campagnari (ed.), *Proceedings of the 7th International Symposium on Heavy Flavors, Santa Barbara, California, 1997* (World Scientific, Singapore, 1999) p. 196.

[19] A. Ali and G. Hiller, Eur. Phys. J. **C8**, 619 (1999).

[20] C. S. Kim, T. Morozumi, A. I. Sanda, Phys. Rev. **D56**, 7240 (1997).

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

[22] J. Chay, H. Georgi and B. Crinstein, Phys. Lett. **B247** 399 (1990).

[23] I. I. Bigi, N. G. Uraltsev and A. I. Vainshtein, Phys. Lett. **B293** 430 (1992) [Erratum-*ibid.* **B297** 477 (1993)];
I. I. Bigi, M. A. Shifman, N. G. Uraltsev and A. I. Vainshtein, Phys. Rev. Lett. **71** 496 (1993);
I. Blok, L. Koyrakh, M. A. Shifman and A. I. Vainshtein, Phys. Rev. **D49** 3356 (1994) [Erratum-*ibid.* **D50** 3572 (1994)].

[24] A. V. Manohar and M. B. Wise, Phys. Rev. **D49** 1310 (1994).

[25] M. Neubert, Phys. Rep. **245** 259 (1994).

[26] A. Buras, *et al.*, Nucl. Phys. **B424**, 374 (1994).

[27] A. Buras and M. Münz, Phys. Rev. **D52** 186 (1995).

[28] G. Deshpande, J. Trampetic and K. Panose, Phys. Rev. **D39** 1461 (1989); C. S. Lim, T. Morozumi and A. I. Sanda, Phys. Lett. **B218** 343 (1989).

[29] M. Bauer, B. Stech and M. Wirbel, Z. Phys. **C34** 103 (1987).

[30] A. Ali, T. Mannel and T. Morozumi, Phys. Lett. **B273** 505 (1991).

[31] F. Krüger and L. M. Sehgal, Phys. Lett. **B380**, 199 (1996).

[32] N. Isgur and M. B. Wise, Phys. Rev. **D42**, 2388 (1990); Phys. Lett. **B237**, 527 (1990); Phys. Lett. **B232**, 113 (1989).

[33] M. J. Dugan and B. Grinstein, Phys. Lett. **B255**, 583 (1991).

[34] J. Charles *et al.*, Phys. Rev. **D60** 014001 (1999).

[35] M. Beneke and Th. Feldmann, Nucl. Phys. **B592**, 3 (2001).

[36] G. Burdman and G. Hiller, Phys. Rev. **D63**, 113008 (2001).

[37] J. D. Richman and P. R. Burchat, Rev. Mod. Phys. **67**, 893 (1995).

[38] J. D. Richman, *Les Houches Lectures on Heavy-Quark Physics and CP Violation* (Elsevier Science, 1997).

[39] T. Goto *et al.*, Phys. Rev. **D55**, 4273 (1997);
T. Goto, Y. Okada and Y. Shimizu, Phys. Rev. **D58**, 094006 (1998);
T. Goto, Y. Okada and Y. Shimizu, KEK-TH-611; preprint hep-ph/9908499.

[40] C. R. Allton *et al.* (APE Collaboration), Phys. Lett. **B345**, 513 (1995);
K. C. Bowler *et al.* (UKQCD Collaboration), Phys. Rev. **D51**, 4955 (1995);
L. Del Debbio *et al.* (UKQCD Collaboration), Phys. Lett. **B416**, 392 (1998).

[41] P. Colangelo and A. Khodjamirian, In: M. Shifman (ed.), the Boris Ioffe Festschrift *At the Frontier of Particle Physics / Handbook of QCD*, (World Scientific, Singapore, 2001), to be published.

[42] M. A. Shifman, A. I. Vainshtein and V. .I. Zakharov, Nucl. Phys. **B147** 385 (1979).

[43] P. Colangelo *et al.*, Phys. Rev. **D53** 3672 (1996) [Erratum-*ibid.* **D57** 3186 (1998)]; Phys. Lett. **B395**, 339 (1997).

[44] P. Ball and V. M. Braun, Phys. Rev. **D58**, 094016 (1998).

[45] P. Ball *et al.*, Nucl. Phys.**B529**, 323 (1998).

[46] P. Ball and V. M. Braun, Nucl. Phys.**B543**, 201 (1999).

[47] W. Jaus and D. Wyler, Phys. Rev. **D41**, 3405 (1990);
C. Q. Geng and C. P. Kao, Phys. Rev. **D54**, 5636 (1996).

[48] B. Stech, Phys. Lett. **B354**, 447 (1995).

[49] P. Colangelo *et al.*, Eur. Phys. Jour. **C8** 81 (1999).

[50] D. Melikhov, N. Nikitin and S. Simula, Phys. Lett. **B410**, 290 (1997).

[51] P. Cho, M. Misiak and D. Wyler, Phys. Rev. **D54** 3329 (1996).

[52] E. Lunghi, Nucl. Phys. **B568** 120 (2000).

[53] P. F. Harrison and H. R. Quinn (eds.) *The BABAR Physics Book: Physics at an Asymmetric B Factory*, SLAC-R-0504 (1998).

[54] S. Glenn *et al.* (CLEO Collaboration), Phys. Rev. Lett. **80** 2289 (1998).

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

[56] C. Albajar *et al.* (UA1 Collaboration), Phys. Lett. **B262**, 163 (1991).

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

[58] S. Anderson *et al.* (CLEO Collaboration), CLNS 01/1739, CLEO 01-11; preprint hep-ex/0106060.

[59] *Details on continuum suppression as used in the analysis of* $B^0 \to K^+\pi^-$ *and* $B^0 \to \pi^+\pi^-$, CLEO CBX 93-41 (1993).

[60] *BABAR Technical Design Report*, SLAC-R-95-457 (1995).

[61] *Kaon Selection at the BABAR Experiment*, *BABAR* Analysis Document #116 (2000).

[62] *Cut-Based Electron Identification*, *BABAR* Analysis Document #90 (2001).

[63] *Muon Identification in the BaBar Experiment*, *BABAR* Analysis Document #60 (2000).

[64] *Measurement of the Number of Upsilon(4S) Mesons Produced in Run 1 (B Counting)*, *BABAR* Analysis Document #134 (2001).

[65] GEANT Detector Description and Simulation Tool, Version 3.21, CERN Program Library W5103, CERN (unpublished) (1995).

[66] *Measurement of exclusive charmonium branching ratios using data from BABAR's first run*, *BABAR* Analysis Document #113 (2001).

[67] *Lepton Identification Efficiencies and Systematic Errors*, *BABAR* Analysis Document # 126 (2001).

[68] *Measurement of the GoodTrackLoose efficiency using SVT tracks BABAR* Analysis Document # 157 (2001).

[69] D. Lange, Nucl. Instr. and Methods A **462**, 152 (2001).

[70] *Blind Analyses in BABAR* , *BABAR* Analysis Document # 91 (2000).

[71] *Choice of Kinematic Variables in B Meson Resonstruction – Take 3*, *BABAR* Analysis Document #53 (2000).

[72] *Reconstruction of* π^0 *in BABAR* , *BABAR* Analysis Document #20 (2000).

[73] T. Skwarnicki, *A Study of the Radiative Cascade Transitions Between the Upsilon-Prime and Upsilon Resonances*, DESY F31-86-02 (thesis, unpublished) (1986).

[74] B. Aubert *et al.* (*BABAR* Collaboration) *The BABAR Detector*, SLAC-PUB-8569; preprint hep-ex/0105044.

[75] *A User's Guide to the RooFitTools Package for Unbinned Maximum Likelihood Fitting*, *BABAR* Analysis Document #18 (2001).

[76] H. Albrecht *et al.*, (ARGUS Collaboration), Phys. Lett. **B185**, 218 (1987).

[77] G. J. Feldman and R. D. Cousins, Phys. Rev. **D57**, 3873 (1998).

[78] S. Spanier, private communication.

[79] P. Ball *et al.*, *B decays at the LHC*; preprint hep-ph/0003238.