SLAC-R-743

Study of Exclusive Charmless Semileptonic Decays of the B Meson

Amanda Jacqueline Weinstein

SLAC-Report-743
December 2004

Prepared for the Department of Energy
under contract number DE-AC02-76SF00515

STUDY OF EXCLUSIVE CHARMLESS SEMILEPTONIC DECAYS OF THE B MESON

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF PHYSICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Amanda Jacqueline Weinstein
December 2004

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Jonathan Dorfan
(Principal Adviser)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Vera Luth

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Patricia Burchat

Approved for the University Committee on Graduate Studies.

Foreword

The couplings of the quarks to the weak charged current are represented by the elements of the Cabibbo-Kobayashi-Maskawa (CKM) matrix. Precision tests of the relations between these elements provide a cross check of the Standard Model prediction for charged current interactions, and a potential probe for physics beyond the Standard Model. Among the smallest and least known of the CKM matrix elements is $|V_{ub}|$, which gives the probability of a weak-current transition from a bottom quark (b) to an up quark (u).

Charmless semileptonic decays of the B meson involve a $b \to u$ quark transition, along with the emission of a lepton and neutrino, and have decay rates directly proportional to $|V_{ub}|^2$. A measurement of $|V_{ub}|$ may be derived from the branching fraction for any semileptonic $b \to u$ transition to an exclusively reconstructed final state, such as $\pi^{0/-}\ell^+\nu$ or $\rho^{0/-}\ell^+\nu$. In this case, the dominant theoretical uncertainty in $|V_{ub}|$ comes from our limited knowledge of the process by which the b and u quarks are bound into the initial and final state mesons, described by a set of hadronic form factors for each decay mode.

In a sample of 83×10^6 $B\overline{B}$ events collected by the *BABAR* detector, we reconstruct 310 ± 31 $B^0 \to \pi^-\ell^+\nu$ decays and 117 ± 16 $B^+ \to \pi^0\ell^+\nu$ decays. We combine the information from both channels, using isospin symmetry and the measured ratio of B lifetimes, to obtain the branching fraction

$$\mathcal{B}(B^0 \to \pi^-\ell^+\nu) = 1.30 \pm 0.11(stat.)^{+0.26}_{-0.16}(syst.) \pm 0.07(FF_{\pi\ell\nu}) \times 10^{-4}. \tag{1}$$

With the large sample of $B\overline{B}$ events available, the measurement is currently limited by systematic uncertainties in the description of background.

The data are consistent with both quark model and light-cone sum rule predictions for the $B \to \pi\ell\nu$ form-factor. As a consequence, we consider both form-factor models in extracting $|V_{ub}|$ from the measured $\mathcal{B}(B^0 \to \pi^-\ell^+\nu)$ branching fraction. Based on the ISGW2 quark model predictions, we obtain

$$|V_{ub}|_{ISGW2} = (2.97^{+0.12}_{-0.13}(stat.)^{+0.28}_{-0.19}(syst.) \pm 0.08(FF_{\pi\ell\nu})) \times 10^{-3}, \quad (2)$$

and based on light-cone sum rule predictions, we obtain

$$|V_{ub}|_{Ball01} = (3.07^{+0.12}_{-0.13}(stat.)^{+0.29}_{-0.19}(syst.) \pm 0.08(FF_{\pi\ell\nu})^{+0.47}_{-0.43}(\Gamma_{thy})) \times 10^{-3}. \quad (3)$$

These results are consistent with, and of comparable precision to, existing measurements made using exclusively reconstructed decays, but have a substantially smaller statistical error. It is believed that both the experimental systematic and theoretical errors may be reduced with further study, and that this approach promises to provide one of the best measurements of $|V_{ub}|$.

Acknowledgments

There is no way I could possibly acknowledge all those who have contributed to my graduate experience. I wish only to highlight those to whom I owe a particular debt of gratitude. First and foremost, I thank my advisors, Jonathan Dorfan and Vera Luth, for their advice, their patience, and their support throughout my graduate career. Second, I would like to thank those who have worked with me on this analysis over the years, in particular Jochen Dingfelder, Eckhard Elsen, and Mike Kelsey. Mike and Jochen especially have been invaluable sounding boards and sources of moral support (not to mention the occasional lecture on C++ coding technique), and I am deeply grateful for their help.

I would like to thank the members of Group C, including my fellow students, for their friendship, their willingness to listen, and a great many useful physics discussions and tips, not to mention their tolerance of the fact that this measurement has occupied a rather appalling amount of their disk space. I would also like to thank Burt Richter for his very welcome support and equally welcome sage advice.

I would like to thank the members of my Reading and Examination Committees, Patricia Burchat, Aaron Roodman, and Helmut Wiedemann, for their time. I also owe a debt to the *BABAR* Collaboration as a whole, and to the PEP-II team, for making this research possible.

To Eleanor Mitchell, Mika Stratton, Anna Pacheco, and Maria Frank, all of whom have helped me navigate the treacherous waters of bureaucracy more than once in this past year, thank you.

Thank you, also, to all my friends, who saw me through the worst of things with frequent offers of food, friendship, and a wireless connection (not necessarily in that order or always in that combination), and who have stuck by me despite my obsessive focus, my frequent lapses in memory, and my completely crazy schedule.

And last but not least, to my family, especially my father, Marvin Weinstein, for everything.

Contents

List of Tables

List of Figures

Chapter 1

Theory of Semileptonic B Decays

1.1 Standard Model

The Standard Model of high-energy physics provides, to date, the most successful description we have of the fundamental interactions of matter. Formulated as a Lorentz-covariant quantum field theory invariant under transformations of the group $SU(3)_c \times SU(2)_L \times U(1)_Y$ [2], the Standard Model provides for the most part an elegant theoretical framework and a set of precise and well-tested predictions. Nonetheless, the model is generally felt to have both aesthetic and practical flaws, such as the fact that it possesses 19 free parameters whose values must be empirically determined, and it is generally believed that the Standard Model itself is only an effective form of some more fundamental high-energy theory.

Matter appears in the Standard Model in the form of spin-$\frac{1}{2}$ fermions, which couple to one another via spin-1 gauge bosons. The fermions divide into two parallel groupings: the six quarks, q^i, which transform non-trivially under all parts of the $SU(3)_c \times SU(2)_L \times U(1)_Y$ group, and the six leptons (l, ν), which transform as $SU(3)_c$ singlets. In other words, the quarks carry color-charge and participate in interactions governed by QCD, while the leptons carry no color-charge and are spectators in any purely strong interaction. The left-handed up and down-type quarks of the same generation transform as doublets under the $SU(2)_L$ gauge group, as do the left-handed lepton-neutrino pairs of each generation. Right-handed quarks and leptons all transform as $SU(2)_L$ singlets.

1.2 Electroweak Interactions

The Standard Model of electroweak interactions possesses an $SU(2)_L \times U(1)_Y$ gauge symmetry, with gauge coupling constant g and gauge bosons $W^i_\nu, i = 1, 2, 3$ associated with $SU(2)_L$ and gauge coupling constant $g^{'}$ and gauge boson B^μ associated with $U(1)_Y$. Rewriting $W^\pm = (W^1 \pm W^2)/\sqrt{2}$, and rotating the W^3 and B states by $\theta_W \equiv \tan^{-1}(g^{'}/g)$ (usually referred to as the weak mixing angle) produces a different set of fields, linear combinations of the first, which transform under a different U(1) and SU(2); the SU(2) is broken via spontaneous symmetry breaking and the Higgs mechanism.[1] The linear combination $A = B \cos\theta_W + W^3 \sin\theta_W$, which is the gauge boson of the unbroken U(1) symmetry, remains massless; this field is identified with the physical photon and the U(1) symmetry with electromagnetism. The fields $W^\pm$ and $Z = -B \sin\theta_W + W^3 \cos\theta_W$, no longer protected by the broken SU(2), acquire substantial masses, and may be identified as the charged and neutral massive vector bosons.

After spontaneous symmetry breaking, the electroweak Lagrangian for the fermion fields is [2]

$$\mathcal{L}_F = \sum_i^{n_f} \bar{\psi}_i \left(i\partial - m_i - \frac{g m_i H}{2M_W} \right) \psi_i - \frac{g}{2\sqrt{2}} \sum_i^{n_f} \bar{\psi}_i \gamma^\mu (1-\gamma^5) \left(T^+ W^+_\mu + T^- W^-_\mu \right) \psi_i$$
$$- e \sum_i^{n_f} q_i \bar{\psi}_i \gamma^\mu \psi_i A_\mu - \frac{g}{2\cos\theta_W} \sum_i \bar{\psi}_i \gamma^\mu (g^i_V - g^i_A \gamma^5) \psi Z_\mu \quad (1.1)$$

where the left-handed fermion fields $\psi_i = \begin{pmatrix} \nu_i \\ l^-_i \end{pmatrix}$ and $\psi_i = \begin{pmatrix} u_i \\ d^{'}_i \end{pmatrix}$ transform as doublets under SU(2). $T^+$and $T^-$are the weak SU(2) (isospin) raising and lowering operators,

[1]In the Higgs mechanism, a (Higgs) doublet of complex scalar fields is introduced, which couples to the gauge bosons W and B via the covariant derivative and to the fermions via a set of Yukawa couplings. By acquiring a vaccum expectation value, the Higgs doublet can "spontaneously" break the $SU(2) \times U(1)$ symmetry without requiring any explicit symmetry-breaking terms to be added to the original Lagrangian.

and the vector and axial weak couplings of fermions to the Z are

$$g_V^i \equiv t_{3,i} - 2q_i \sin^2\theta_W \tag{1.2}$$

$$g_A^i \equiv t_{3,i} \tag{1.3}$$

where $t_{3,i}$ is the weak isospin of fermion i ($+1/2$ for up-type quarks and neutrinos, $-1/2$ for down-type quarks and electrons, muons, and taus). $e \equiv g\sin\theta_W$ is the positron/electron charge, and the q_i are the fermion electric charges in units of e. The terms involving H are the fermion interactions with the physical neutral scalar field that is the residue of the original Higgs doublet after spontaneous symmetry breaking.

1.3 Strong Interactions

The Lagrangian for the strong interactions, which is symmetric under $SU(3)_c$, can be written

$$\mathcal{L}_{QCD} = -\frac{1}{4}F_{\mu\nu}{}^{(a)}F^{\mu\nu(a)} + i\sum_q \bar{\psi}_q^i\gamma^\mu(D_\mu)_{ij}\psi_q^j - \sum_q m_q\bar{\psi}_q^i\psi_{qi} \tag{1.4}$$

$$F_{\mu\nu}^{(a)} = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a - g_s f_{abc}A_\mu^b A_\nu^c \tag{1.5}$$

$$(D_\mu)_{ij} = \delta_{ij}\partial_\mu + ig_s\sum_a \frac{\lambda_{i,j}^a}{2}A_\mu^a \tag{1.6}$$

where g_s is the QCD coupling constant, the f^{abc} are the structure constants of the $SU(3)$ algebra, and the quark mass terms are the result of spontaneous symmetry breaking in the electroweak sector. The eight gluon fields A_μ^a, which transform as the adjoint representation of $SU(3)$, couple to each other as well as to the quark fields ψ, which carry color indices i, j in addition to the flavor index q.

QCD is *asymptotically free*: from an intuitive point of view, this implies that the QCD vacuum anti-screens color charge; i.e., the effective coupling between quarks vanishes at short distance scales, allowing quarks to behave as free particles, but increases a large distances scales. Strong coupling and asymptotic freedom are thought but have not been conclusively shown to imply another empirically determined feature of QCD, *confinement*, the fact that quarks do not appear free in nature, but are found as part of bound states known as hadrons.

Moreover, while QCD can be treated as perturbative where short-distance physics is concerned, the long-distance physics is completely non-perturbative, which poses an additional challenge in calculating matrix elements for hadron decay.

1.4 CKM Matrix and Current Experimental Constraints

In Equation 1.1, the fact that the neutrinos masses are ignored, and hence the neutrino are effectively degenerate in mass means that the weak eigenstates may be defined to coincide with the physical leptons (mass eigenstates). In the case of the quarks, sufficient freedom exists to do the same for the three up-type quarks (u, c, t) but the weak eigenstates d', s', b' referred to in Equation 1.1 are not the same as the quark mass eigenstates d, s, b.

The transformation between quark mass eigenstates and weak eigenstates is given by a unitary matrix, usually referred to as the Cabibbo-Kobayashi-Maskawa (CKM) matrix [3],

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} . \tag{1.7}$$

The fact that this matrix is not diagonal has a number of phenomenological implications, including the possibility of flavor-changing transitions between quarks of different generations.

The CKM matrix may be characterized in terms of three real rotation angles (θ_{12},

θ_{23}, θ_{13}), and a real phase δ_{13},

$$V = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13}. \end{pmatrix} \tag{1.8}$$

The single complex phase enters because of the three-generation nature of the CKM matrix: a mass-mixing matrix involving only two quark generations would be parameterizable solely in terms of real rotations. More quark generations would lead to a more complex parametrization involving multiple phases. The fact that δ_{13} is non-zero allows for the Standard Model mechanism of CP violation in weak decays. Since $|V_{ub}| = \sin\theta_{13}$ multiplies every term in the CKM matrix carrying that phase, the Standard Model mechanism for CP violation requires $|V_{ub}|$ to be non-zero.

1.4.1 The Wolfenstein Parametrization

It is often convenient to approximate the standard parametrization of the CKM matrix, as given in Equation 1.8, by what is effectively a Taylor series in $\lambda = \sin\theta_{12} \approx 0.22$ [4]:

$$V = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{1.9}$$

In this approximation, the hierarchy of the CKM matrix elements is made readily apparent; constraints on the CKM matrix are also often expressed in terms of constraints on the four parameters A, λ, ρ, and η.

1.4.2 CKM Tests and the Unitarity Triangle

As the CKM matrix represents an orthonormal change of basis, it is required to be unitary. Were there additional quark generations, the three-generation mass-mixing matrix would no longer be unitary by itself. If precise experimental measurements of

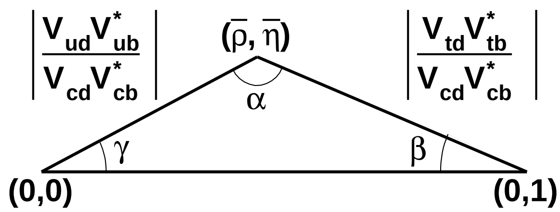


Figure 1.1: Illustration of the most common unitarity triangle, with the base normalized to unit length.

the CKM matrix elements were to show that it is not truly unitary, this could be a signature of additional quark generations or of a more fundamental problem with the Standard Model.

A convenient graphical representation may be derived from the fact that, if the CKM matrix is unitary, the scalar product of any two rows or columns must be zero. Any such equation may be graphically represented as a triangle in the complex plane; the triangle usually chosen is derived from the constraint applied to the first and third columns of the matrix, which gives

$$V_{ud}V^*_{ub} + V_{cd}V^*_{cb} + V_{td}V^*_{tb} = 0. \tag{1.10}$$

Dividing this expression by a factor of $V_{cd}V^*_{cb}$ gives the triangle shown in Figure 1.1; the base is of unit length, while the sides have lengths of $|V_{ud}V^*_{ub}/V_{cd}V^*_{cb}|$ and $|V_{td}V^*_{tb}/V_{cd}V^*_{cb}|$. In the Wolfenstein parametrization, $V_{ud}V^*_{ub}/V_{cd}V^*_{cb} \approx (1-\lambda^2/2)(-\rho - i\eta)$ and $V_{td}V^*_{tb}/V_{cd}V^*_{cb} \approx (1-\lambda^2/2)((\rho-1)+i\eta)$. In this parametrization, one can also identify the apex of the triangle with $(\bar{\rho}, \bar{\eta})$ where $\bar{\rho} = (1-\lambda^2/2)\rho$ and $\bar{\eta} = (1-\lambda^2/2)\eta$, and the angles β and $\gamma = \delta_{13}$ with the phases of the CKM matrix elements V_{td} and V_{ub}.

A selection of experimental results as quoted in Ref. [5] is summarized in Equation 1.11:

$$\begin{pmatrix} |V_{ud}| = 0.9737 \pm 0.0007 & |V_{us}| = 0.2210 \pm 0.0023 & |V_{ub}| = (4.26 \pm 0.13 \pm 0.50)10^{-3} \\ |V_{cd}| = 0.224 \pm 0.016 & |V_{cs}| = 0.995 \pm 0.014 & |V_{cb}| = 0.0419(1 \pm 0.025 \pm 0.17) \\ |V_{td}| = (9.2 \pm 1.4 \pm 0.5)10^{-3} & |V_{ts}| > 0.003 & |V_{tb}| \simeq 1 \end{pmatrix}, \tag{1.11}$$

where the $|V_{ud}|$ is an average of measurements of super-allowed nuclear β^+ decays, β^- decay of polarized neutrinos, and β^+ decay of π^+ mesons, $|V_{us}|$ is taken from a variety of measurements of hyperon and kaon decays, and is currently known at the 1 percent level, $|V_{cd}|$ is taken from measurements of dimuon production by neutrinos and antineutrinos, and $|V_{cs}|$ is determined using decays of real W bosons coupled with measurements of the other matrix elements. The value of $|V_{cb}|$ represents *BABAR*'s recent inclusive measurement, involving a fit to the hadronic mass moments, while the value of $|V_{ub}|$ is an average of several inclusive techniques from CLEO, BELLE, and *BABAR*. $|V_{td}|$ and $|V_{ts}|$ are determined using measurements of the mass splitting of B_d and B_s mesons, assuming $|V_{tb}| = 1.0$.

With $|V_{ud}|$ and $|V_{cd}|$ well known, knowledge of the side of the triangle opposite the angle β is limited by the experimental and theoretical uncertainties on $|V_{cb}|$ and $|V_{ub}|$, both of which are taken primarily from branching fraction measurements of semileptonic B decays. The experimental constraints on the unitarity triangle are summarized in Figure 1.4.2.

Current measurements of $|V_{ub}/V_{cb}|$, when combined with measurements of the mass splitting for B_d mesons (ΔM_d) and the corresponding limit for B_s mesons (ΔM_s), constrain the apex of the unitarity triangle to lie in the first or second quadrant of the complex plane. The overlap between these constraints, and those taken from measurements of either ϵ_K or sin2β further localize the apex $(\bar{\rho}, \bar{\eta})$ to the first quadrant. The precise measurements of sin2β from the *BABAR* and BELLE collaborations provide the tightest constraint.

Even assuming the most extreme possible values consistent with current experimental results, and assuming the precision of the sin2β measurement reaches 2% in the next few years, experimental knowledge of $|V_{ub}/V_{cb}|$ must be better than 10% before the triangle might be seen to fail to close.

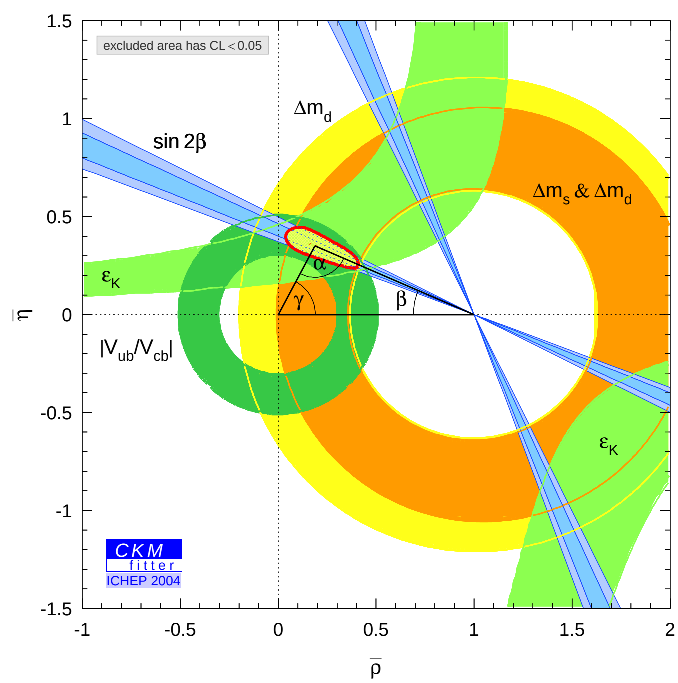


Figure 1.2: Current experimental constraints on the Unitarity Triangle, as given by the CKM Fitter Group for Summer 2004. The dark green ring shows the constraint from $|V_{ub}/V_{cb}|$.

1.5 Semileptonic B Decays

Flavor-changing quark transitions are accounted for by the charged-current weak interaction, which is the second term of the electroweak Lagrangian given in Equation 1.1 . The coupling of the W boson to free quarks may be taken from this term and written as

$$\sum_{i,j} \bar{u}_i \gamma^{\mu} \frac{1}{2}(1-\gamma_5) V_{ij} d_j \tag{1.12}$$

Likewise, the coupling of the W boson to an electron and neutrino is given by

$$-\frac{e}{2\sqrt{2}sin\theta_W}\left[W_{\mu}^{-}\bar{e}\gamma^{\mu}(1-\gamma^5)\nu + W_{\mu}^{+}\bar{\nu}\gamma^{\mu}(1-\gamma^5)e\right] \tag{1.13}$$

The factor $\gamma^{\mu}(1-\gamma_5)$ means the structure of the charged-current interaction is manifestly V-A for both quarks and leptons.

These two couplings allow for a process in which a quark may decay into a quark of different flavor, a lepton, and a neutrino. Particles containing heavy quarks (such as the B meson) can thus decay semileptonically in a manner analogous to nuclear beta decay.

The matrix element for a semileptonic B decay may be schematically represented by the Feynman diagram given in Figure 1.3. At tree-level this matrix element is merely the product of the leptonic and hadronic currents defined above, sandwiched between physical states. Higher order electroweak corrections are suppressed as powers of m_l/M_W and may be ignored; likewise, for processes where momenta are consistently small compared to M_W (i.e. for processes where the W is virtual) the W propagator may be approximated as $G_F/\sqrt{2} = g^2/8M_W^2$. In this approximation, the hadronic and lepton components of the matrix element factorize completely, and the semileptonic B decay amplitude may be written

$$A(\bar{B} \to m\ell^- \bar{\nu}) = \frac{G_F}{\sqrt{2}} V_{bq} L^{\mu} H_{\mu} \tag{1.14}$$

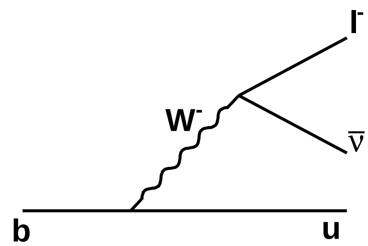


Figure 1.3: Feynman diagram for a semileptonic $b \to u\ell^-\bar{\nu}$ decay.

where the leptonic and hadronic currents are given by

$$L^\mu = \bar{u}_e \gamma^\mu (1-\gamma_5) v_\nu \tag{1.15}$$

$$H^\mu = < p_m | J^\mu_{\text{had}}(0) | P_B > \tag{1.16}$$

The hadronic current is not easily calculated, as the matrix element must be taken between the physical final states, mesons, rather than free quarks. This means that neither the higher-order perturbative corrections, nor the non-perturbative long-distance scale physics involved in hadronization, may be ignored. In general, rather than attempt to explicitly calculate this current, the difficult-to-calculate quantities for any given decay are bundled into a set of Lorentz-invariant form factors, and the current is rewritten in terms of these form factors and the four-vectors of the decay.

1.5.1 Form-factors and Helicity Amplitudes

The V-A structure of the hadronic current may be invoked, along with knowledge of the transformation properties of the final state meson, to identify the form factors of interest in semileptonic B decays. There are two categories. For decays with a pseudoscalar meson in the final state, the only available four-vectors are those of the

B meson and the pseudoscalar meson itself, and the current may be written as [6]

$$\langle P(p_P)|J_\mu(0)|B(p_B)\rangle = f_+(q^2)(p_B+p_P)_\mu + f_-(q^2)((p_B-p_P)_\mu \quad (1.17)$$

where p_P is the momentum of the final-state pseudoscalar meson P, p_B that of the parent B meson, and q^2 is the invariant mass of the virtual W. In the limit of a massless lepton the second term vanishes. Thus, for decays involving electrons and muons in the final state, the contribution from $f_-(q^2)$ may be neglected and the only form factor of interest is $f_+(q^2)$.

For decay with a vector meson V in the final state, with momentum P_V, there is an additional four-vector available, the polarization vector ϵ of the vector meson itself, and the matrix element of interest may be written [7]

$$\begin{aligned} &< V(p_V,\epsilon)|V_\mu(0)-A_\mu(0)|B(p_B)>= \\ &\quad -i\epsilon^*(m_B+m_V)A_1(q^2)+i(p_B+p)_\mu(\epsilon^* p_B)\frac{A_2(q^2)}{m_B+m_V} \\ &\quad +iq_\mu(\epsilon^* p_B)\frac{2m_V}{q^2}(A_3(q^2)-A_0(q^2))+\epsilon_{\mu\nu\rho\sigma}\epsilon^{*\nu}p_B^\rho p^\sigma\frac{2V(q^2)}{m_B+m_V} \end{aligned} \quad (1.18)$$

where

$$A_3(q^2) = \frac{m_B+m_V}{2m_V}A_1(q^2)-\frac{m_B-m_V}{2m_V}A_2(q^2) \quad (1.19)$$

$$A_0(0) = A_3(0) \quad (1.20)$$

$$< V|\partial_\mu A^\mu|B > = 2m_V(\epsilon^* p_B)A_0(q^2) \quad (1.21)$$

The differential decay rate for $B \to M\ell^-\bar{\nu}$ is proportional the square of the decay amplitude $\aleph$ [6]

$$\frac{d\Gamma}{dq^2 d\cos\theta_l d\Omega_m} = \frac{1}{2}\frac{p_m^*}{m_B^2(4\pi)^5}|\aleph|^2, \quad (1.22)$$

where p_m^* is the momentum of the final meson in the B rest frame, $\cos\theta_l$ the polar angle of the lepton in the W rest frame, and $d\Omega_m$ the differential solid angle for the direction of the final meson in the B rest frame.

It is convenient to expand the decay amplitude as a sum of terms, one for each of the three helicity states (+,-,0) of the W [6]:

$$\begin{aligned}|\aleph|^2 = \frac{G_F^2}{2}|V_{Qq}|^2 q^2 \Bigg[& \frac{1}{2}(1-\cos\theta_l)^2|H_+|^2 + \frac{1}{2}(1+\cos^2\theta_l)|H_-|^2 + \sin^2\theta_l|H_0|^2 \\ & + \frac{1}{2}\sin^2\theta_l(H_+H_-^* + H_+^*H_-) - \frac{1}{\sqrt{2}}(1-\cos\theta_l)(H_+H_0^* + H_+^*H_0) \\ & - \frac{1}{\sqrt{2}}(1+\cos\theta_l)(H_-H_0^* + H_-^*H_0)\Bigg] \end{aligned} \quad (1.23)$$

where the helicity amplitudes H_+, H_-, and H_0 are complex quantities depending on both q^2 and the direction of the final-state meson.

The helicity of the final-state meson is locked to that of the virtual W. Thus, for decays with a pseudoscalar final-state meson, such as the pion, $H_\pm(q^2) = 0$. Likewise, in the limit of massless leptons, only one form factor, $f_+(q^2)$, may appear in the decay amplitudes, and the remaining helicity amplitude, H_0, is trivially related to $f_+(q^2)$:

$$H_0 = -\frac{2m_B p_P^*}{q} f_+(q^2) \quad (1.24)$$

In the case of a vector final-state meson, the helicity amplitudes may be written in terms of the form factors as follows [8]

$$H_\pm(q^2) \propto (m_B + m_V)A_1(q^2) \mp \frac{2m_B|\vec{p_V}|}{m_B + m_V}V(q^2) \quad (1.25)$$

and

$$H_0 \propto \frac{1}{2m_V\sqrt{(q^2)}}\left[(m_B^2 - m_V^2 - q^2)(m_B + m_V)A_1(q^2) - \frac{4m_B^2|\vec{p_V}|^2}{m_B + m_V}A_2(q^2)\right] \quad (1.26)$$

The properties of the final-state meson (particularly whether or not it contains a charm quark) dictate which techniques may be used to calculate the form factors; this will be discussed in the following sections. Likewise, the transformation properties and available decay modes of the final-state meson dictate the angular dependence of the helicity amplitudes.

1.5.2 Form-Factors in Semileptonic Decays

Heavy Quark Effective Theory (HQET) is a powerful tool for handling certain classes of QCD calculations involving a heavy quark. In the case of a B meson, or indeed any meson containing one heavy ($m_Q >> \Lambda_{QCD}$) and one light ($m_q << \Lambda_{QCD}$) quark, the heavy quark may be treated, in the limit $m_Q \to \infty$, as a static color source, and the dynamics in this limit remain the same under exchange of heavy quark flavors. Since m_Q is finite, there are $1/m_Q$ corrections to this effective theory, which may be accommodated perturbatively.

In the case of $B \to X_c \ell\nu$ decays, where both the initial and final-state mesons contain a heavy quark, HQET may be applied to calculate form factors, particularly those for the decays $B \to D^* \ell\nu$. The standard parametrization of the HQET results for the $B \to D^* \ell\nu$ form factors are given in Appendix 2, while the derivation of the angular dependence of H_+, H_-, and H_0 for $B \to D^* \ell\nu$ for both $D^* \to D\pi$ and $D^* \to D\gamma$ is given in Appendix 1.

For charmless semileptonic B decays, the problem is less tractable; since the up and down quarks are light, HQET does not apply. Over the years, a number of different approaches—each with its own advantages and difficulties—have been used to calculate the form factors.

1.5.2.1 Quark Models (ISGW2)

Quark model calculations are used to calculate semileptonic decay form factors by postulating forms for the meson wave functions and using these estimated wave functions to calculate the matrix elements of the hadronic currents. Typically these matrix elements are calculated at one of the extreme values $q^2 = 0$ or $q^2 = q^2_{\mathrm{max}}$. The actual q^2 dependence of the form factor is then determined separately, usually by using some simple phenomenological ansatz.

The ISGW model, along with its update ISGW2 [9] is one such constituent quark model with relativistic corrections, and is the only such model considered in this analysis.[2] Calculations are normalized at $q^2 = q^2_{max}$, and extrapolated to low q^2 using

[2]This model is used as the default for the *BABAR* simulation of resonant hadronic states for all

an exponential governed by a "transition charge radius" ξ and an *ad hoc* relativistic correction factor[8]

$$F(q^2) = F(q^2_{max}) \left(1 + \frac{1}{6N}\xi^2(q^2_{max} - q^2)\right)^{-N}, \qquad (1.27)$$

where $N = 2$ for decays to pseudo-scalar mesons and $N = 3$ for decays to vector mesons. The theoretical uncertainties for these calculations are extremely difficult to quantify.

1.5.2.2 Lattice Calculations

Lattice QCD, which involves formulating an affective action on a discrete lattice, offers an alternative approach to calculating form factors for semileptonic decays. Present-day lattice results have a number of limitations, and carry a number of different sources of uncertainty. Quoted errors usually include the statistical error (which comes from the fact that path integrals are evaluated using Monte Carlo integration methods), discretization errors from the finite lattice spacing, and errors that come from the fact that lattice calculations are done on a finite-volume lattice and then extrapolated to the infinite-volume limit. In the case of $B \to \pi \ell \nu$, the chiral extrapolation to a physical final state pion is an additional source of error [10].

Since it is necessary, in current lattice calculations, to set the quark masses artificially high to avoid boundary effects, the result of a lattice calculation must also be extrapolated to the physical light quark masses; this is generally among the more significant uncertainties in any lattice calculation. Another significant uncertainty, whose size is particularly difficult to estimate, is the error due to the "quenching approximation," in which the effect of the quark vaccum polarization is omitted from the lattice simulation. Quenching errors have typically been taken to be on the order of 10%, but in general are difficult to estimate. Some lattice groups believe them to be as large as 30%[11]. Not all lattice results are quoted with errors that include an estimate of the quenching error. Recently, the first unquenched lattice results for $B \to \pi$ form factors have been released [12], [13], and comparison between these

charmless semileptonic decays, and also for most decays involving charm mesons, except for $D^* \ell \nu$.

results and the older quenched results suggest that the estimate of 10% quenching errors was in fact reasonbly accurate.

Present day lattices allow for a lattice spacing a of order $(2-3)\,\mathrm{GeV}^{-1}$; this is still substantially larger than the b quark wavelength, which is proportional to $1/m_b$ where $m_b \sim 5\,\mathrm{GeV}$. As a result, the b quark cannot be simulated directly on present-day lattices, and a number of possible approaches are used to circumvent this difficulty.

There are a plethora of lattice results available, calculated with a variety of different methods, some of which supersede the others. We will explicitly consider a handful of the most recent results. UKQCD99 [14] provides calculations of the $B \to \pi$ and $B \to \rho$ form factors calculated using relativistic quarks with full non-perturbative improvement; FNAL2001 [15] provides these same form factors using the hybrid approach, while JLQCD [16]uses NRQCD.

1.5.2.3 Parametrization of lattice results for $f_+(q^2)$

Becirevec and Kaidalov [17] offer a prescription for parameterizing $f_+(q^2)$ in terms of both the normal pole dominance of B^* and a second effective pole that simulates the effect of other higher resonance states:

$$f^+(q^2) = c_B \left(\frac{1}{1 - q^2/{m_B^*}^2} - \frac{\alpha}{1 - \frac{q^2/{m_B^*}^2}{\gamma}} \right) \tag{1.28}$$

where m_B^* is the B^* mass and $c_B {m_B^*}^2$ is the residue of the form factor at $q^2 = m_{B^*}$.

By fitting this parametrization to a set of lattice results at high q^2, the results can be effectively extrapolated to lower q^2. UKQCD99 has performed these fits to their own lattice data, and Becirevec and Kaidalov have applied the technique to the JLQCD results. A fit to the combined results of all three lattice collaborations for $f_+(q^2)$, using the Becirevic parametrization, is shown in Figure 1.4.

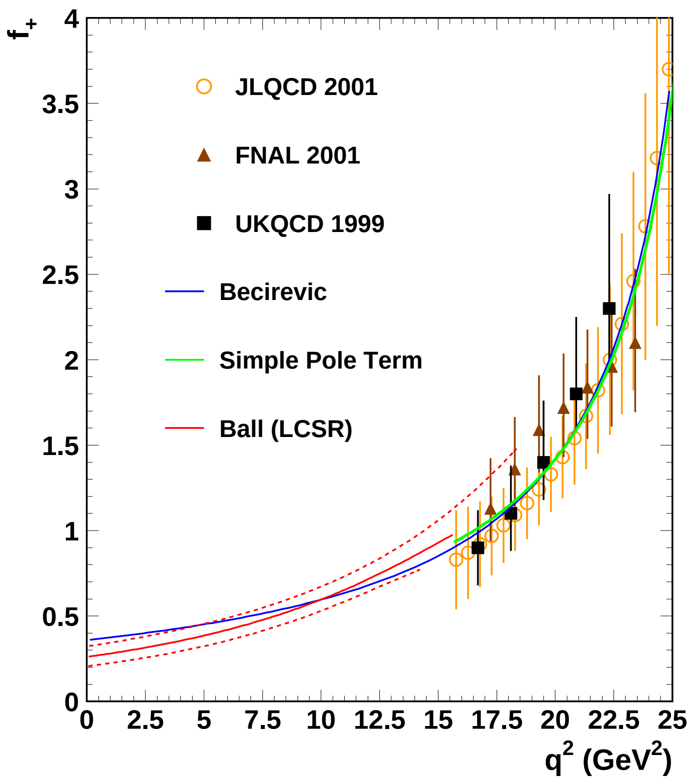


Figure 1.4: Lattice results for the $B \to \pi\ell\nu$ form factor $f_+(q^2)$ from the FNAL 2001 collaboration (yellow circles), the JLQCD2001 collaboration (red triangles), and the UKQCD99 collaboration (black squares). The parametrization of the Ball LCSR results is shown in red; the uncertainty in this parametrization is indicated by the dashed red lines. The simple-pole approximation at high q^2 is shown in green, and the result of a fit to the combined lattice results using the Becirevic parametrization is shown in blue. The error bars on the lattice points are those quoted by each collaboration, and in some cases are statistical and in others are statistical and systematic.

1.5.2.4 Light Cone Sum Rules: LCSR

Light-cone sum rules provide a non-perturbative approach to form factor calculations that is complementary to lattice formulations, albeit not quite as rigorous. It combines the concept of QCD sum-rules with the twist expansions characteristic of hard exclusive processes in QCD [7]. LCSR calculations are valid for low to intermediate values of q^2, i.e., they cover that range of q^2 not covered by lattice results.

Table 1.1: Parametrization of LCSR form factor calculations [1].

Decay	scalar	vector	vector	vector	vector
Form factor	f_+	A_1	A_2	A_0	V
F(0)	0.261	0.261	0.223	0.372	0.338
a	2.03	0.29	0.93	1.40	1.37
b	1.293	-0.415	-0.092	0.437	0.315
c	0.44				
q_0^2 [GeV^2]	15.7				

A universal parameterized form

$$F(q^2) = \frac{F(0)}{1 - a(q^2/m_B^2) + b(q^2/m_B^2)^2} \tag{1.29}$$

can be fitted to the explicit LCSR predictions for the form factor, and agrees with them to better than 2% for all form factors of interest (which includes the $B \to \pi\ell\nu$ form factor $f_+(q^2)$ and the $B \to \rho\ell\nu$ form factors $A_1(q^2)$, $A_2(q^2)$, $A_0(q^2)$, and $V(q^2)$ [1]). In the case of $f_+(q^2)$, this parametrization can be extended to high q^2 (i.e. $q_c^2 > 14 - 18\,\mathrm{GeV}^2$) by assuming vector meson pole dominance:

$$f_+(q^2) = \frac{c}{1 - (q^2/{m_B^*}^2)} \tag{1.30}$$

where $m_B^* = 5.32\,\mathrm{GeV}$ and c and q_c^2 are free parameters of the extended fit.

The values a, b, and $F(0)$ for each form factor are tabulated in Ref. [1] for a range of values of the standard input parameters, such as m_b; the nominal central values are reproduced in Table 1.1, along with the results of the extended parametrization for $f_+(q^2)$. Results of the LCSR extended fit parametrization for $f_+(q^2)$ are shown compared to both lattice points and the lattice fit parametrization in Figure 1.4. It is clear that for $f_+(q^2)$, the Becirevic parametrization of the lattice results and the LCSR parameterizations are both equivalent and consistent. The LCSR parameterizations of the $B \to \rho$ form factors $A_1(q^2)$, $A_2(q^2)$, $A_0(q^2)$, and $V(q^2)$ are shown in Figure 1.5 .

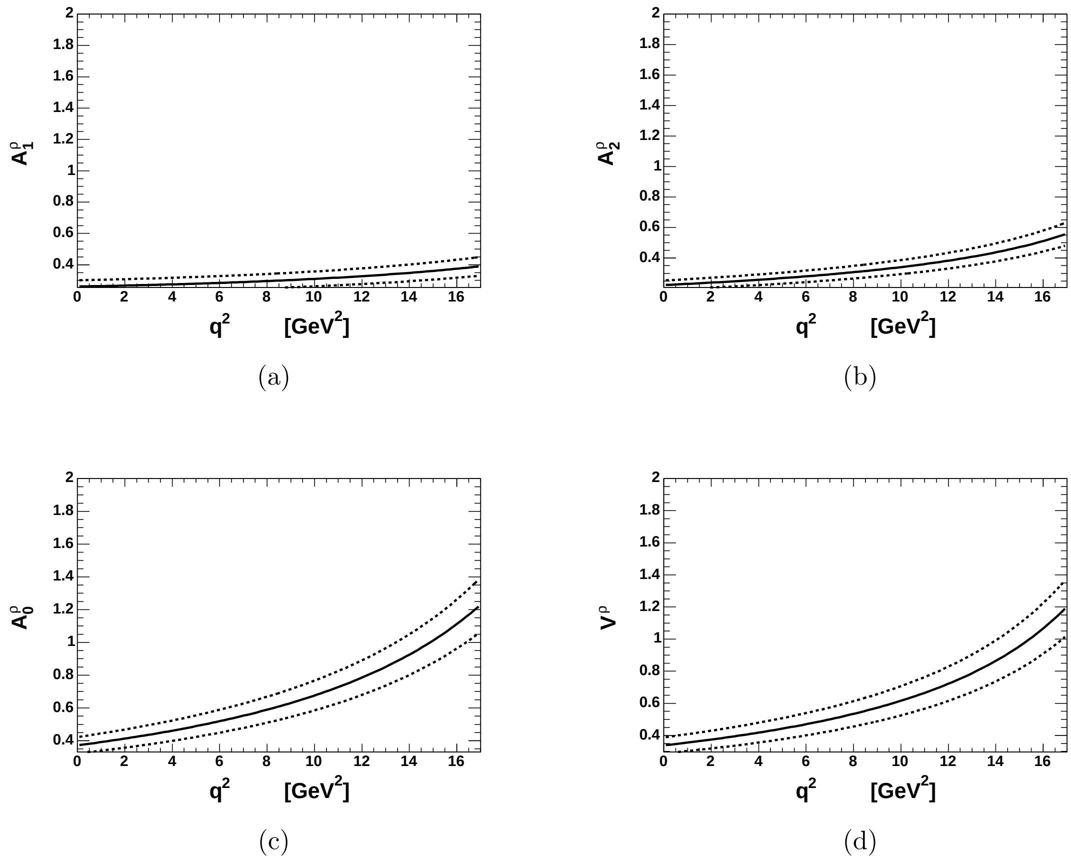


Figure 1.5: LCSR parameterizations of the $B \to \rho\ell\nu$ form factors (a) $A_1(q^2)$, (b) $A_2(q^2)$, (c) $A_0(q^2)$, and (d) $V(q^2)$. Nominal parameterizations are shown by the solid lines; the dashed lines indicate error ranges determined by varying the values of the inputs to the LCSR calculation.

1.5.3 $B \to \pi\ell\nu$

In the case of $B \to \pi\ell\nu$, equation 1.24 holds, and the formulas 1.22 and 1.23 for the differential decay rate, after the appropriate phase space integration, reduce to

$$\frac{d\Gamma}{dq^2 d\cos\theta_l} = \frac{G_F^2 {p_\pi^*}^3 |V_{ub}|^2}{32\pi^3} \sin^2\theta_l |f_+(q^2)|^2 \tag{1.31}$$

The differential decay rate for the nominal parameterized LCSR form factor is shown in Figure 1.6.

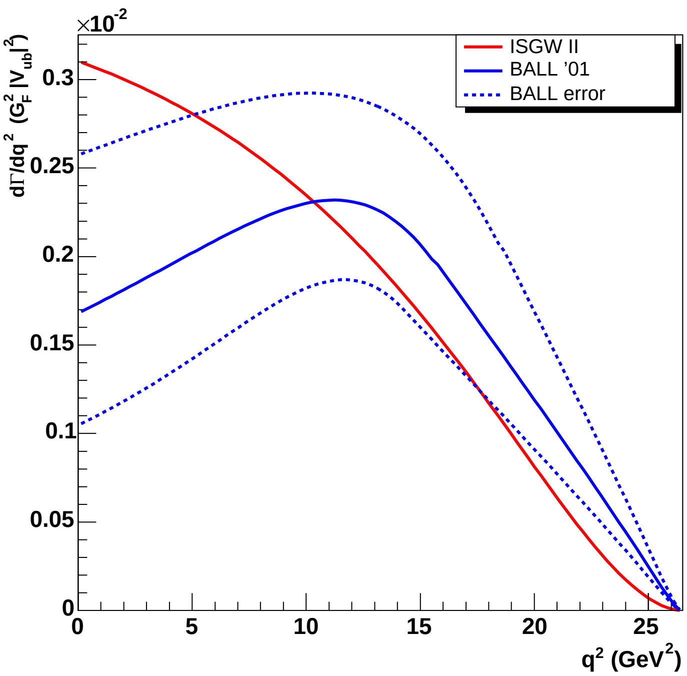


Figure 1.6: Differential decay rate for $B \to \pi\ell\nu$ (normalized to $1/G_F^2|V_{ub}|^2$), assuming ISGW2 form-factor model for $f_+(q^2)$(red), and the nominal extended LCSR parametrization for $f_+(q^2)$(blue). The dashed blue line indicates the error band on the LCSR-based differential decay rate resulting from the error band on $f_+(q^2)$.

1.5.4 $B \to \rho\ell\nu$

For $B \to \rho\ell\nu$ decays, the q^2-dependence of the helicity amplitudes H_0, H_+, and H_- may be expressed in terms of the form factors $A_0(q^2)$, $A_1(q^2)$, $A_2(q^2)$, and $V(q^2)$, as given by equations 1.25 and 1.26. The angular dependence of the helicity amplitudes is that of a vector meson decaying to two pseudoscalars. We will write each of the helicity amplitudes as a product of an angular component and a reduced amplitude ($\bar{H}_+$, $\bar{H}_-$, $\bar{H}_0$) which is a function of q^2 alone.

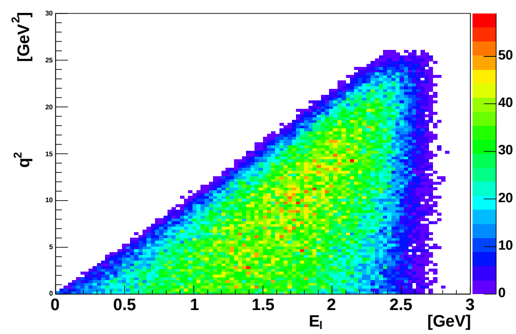
This case is exactly analogous to the case of $B \to D^*\ell\nu$ where the D^* decays to a D meson and pion, which is treated in detail in Appendix 1. The full form of the differential decay rate is

$$
\begin{aligned}
\frac{d\Gamma(B \to \rho l \bar{\nu})}{dq^2 d\cos\theta_l d\cos\theta d\cos\phi} &= \frac{3G_F^2|V_{ub}|^2 P_\rho q^2}{8(4\pi)^2 m_B^2}\mathcal{B}(\rho \to \pi\pi) \\
&\times \Big[(1-\cos\theta_l)^2\sin^2\theta_V|\bar{H}_+|^2 + (1+\cos\theta_l)^2\sin^2\theta_V|\bar{H}_-|^2 \\
&+ 4\sin^2\theta_l\cos^2\theta_V|\bar{H}_0|^2 + 2\sin^2\theta_l\sin^2\theta_V\cos(2\chi)\bar{H}_+\bar{H}_- \\
&+ 4\sin\theta_l(1-\cos\theta_l)\sin\theta_V\cos\theta_V\cos\chi\bar{H}_+\bar{H}_0 \\
&- 4\sin\theta_l(1+\cos\theta_l)\sin\theta_V\cos\theta_V\cos\chi\bar{H}_-\bar{H}_0\Big]
\end{aligned}
\tag{1.32}
$$

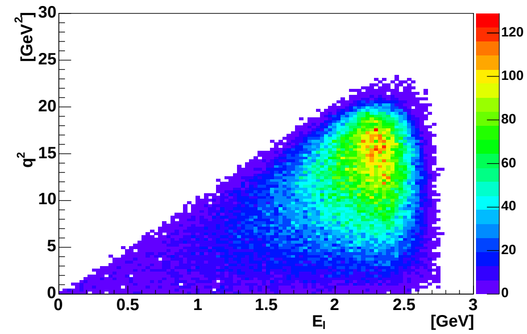
where θ_V is the polar angle of the one of the pions in the helicity frame of the ρ, and χ is the angle between the decay plane of the W and the ρ as defined in Appendix 1.

1.5.5 Properties of q^2 and Lepton Spectra

The exact q^2 and lepton energy spectra of a particular semileptonic B decay channel cannot be predicted without knowing the form factors for that decay, but a number of qualitative features of the lepton spectra may be deduced even in the absence of such information. The spin of the final-state meson affects the observed q^2 distribution; the decay rate for a pseudoscalar meson, such as $B \to \pi\ell\nu$ in equation 1.31, carries a factor of p_X^3 relative to the single factor of p_X in the decay rate for a final-state vector



(a)



(b)

Figure 1.7: Dalitz plots (q^2 vs. lepton energy) for the decays (a) $B \to \pi\ell\nu$ and (b) $B \to \rho\ell\nu$, assuming the ISGW2 form factor model.

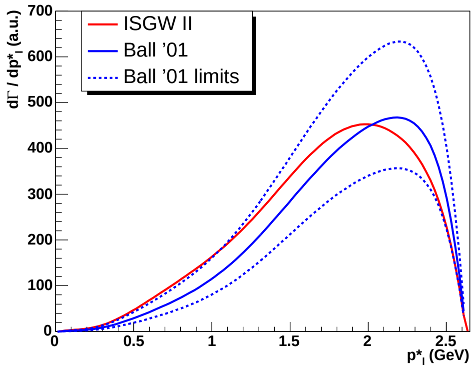
meson such as $B \to \rho\ell\nu$. The additional powers of p_X suppress the decay rate at high q^2 for $B \to \pi\ell\nu$ relative to $B \to \rho\ell\nu$ ("P-wave suppression"). Moreover, since the charged weak current is V-A in form, the processes $b \to c\ell^-\bar{\nu}$ and $b \to u\ell^-\bar{\nu}$ prefer a final state with a charged lepton of helicity $\lambda = -1/2$ and c and u quarks of helicity $\lambda = -1/2$; the configuration maximizing lepton energy—a collinear configuration in which the charged lepton recoils against the daughter quark and neutrino— is therefore allowed in semileptonic B decays, and the lepton-energy spectrum peaks at a higher energy than the neutrino spectrum. Thus, when q^2 is large, the lepton energy also tends to be large, and $B \to \rho\ell\nu$ tends to have both a harder lepton energy spectrum and a harder q^2 spectrum than $B \to \pi\ell\nu$ as shown in Figure 1.7.

Finally, both the lepton and q^2 spectra are influenced by the behavior of the form factors themselves, as shown in Figure 1.8 for the $B \to \pi\ell\nu$ decay channel.

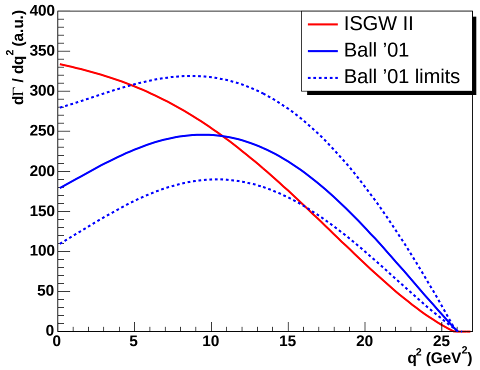
1.6 Extracting $|V_{ub}|$ from $B \to \pi\ell\nu$

In principle, the extraction of $|V_{ub}|$, given knowledge of the branching fraction $\mathcal{B}(B \to \pi\ell\nu)$, is straightforward. The total $B \to \pi\ell\nu$ decay rate

$$\Gamma(B \to \pi\ell\nu) = \int_0^{q^2_{\max}} \frac{G_F^2 p_\pi^{*3} |V_{ub}|^2}{24\pi^3} |f_+(q^2)|^2 dq^2 \qquad (1.33)$$



(a)



(b)

Figure 1.8: (a) Lepton energy spectrum for $B \to \pi\ell\nu$ decays for LCSR parametrization (solid blue), varied within errors (dashed blue) and for ISGW2 (red). (b) q^2 spectrum for $B \to \pi\ell\nu$ decays for LCSR parametrization (blue) varied within errors (dashed blue), and for ISGW2 (red).

is proportional to $|V_{ub}|^2$. In terms of the branching fraction $\mathcal{B}(B^0 \to \pi^-\ell^+\nu)$ times the total B decay width, we may write:

$$|V_{ub}| = \sqrt{\frac{\mathcal{B}}{\Gamma_{thy}\tau_{B^0}}}$$

where the exact value of $\Gamma_{thy} = \Gamma(B \to \pi\ell\nu)/|V_{ub}|^2$ depends on the form factor model used for the decay.

Chapter 2

The *BaBar* Detector

The *BABAR* detector was built to study the decays of B mesons produced by the PEP-II asymmetric-energy electron-positron collider. Both collider and detector are optimized for their primary purpose: the study of time-dependent CP asymmetries in neutral B meson decays. Nonetheless, the detector is sufficiently versatile to allow a full range of B physics measurements to be made. This section gives a brief description of the PEP-II collider and its performance, and details the various components of the *BABAR* detector [18][1].

2.1 The PEP-II Collider

The PEP-II B-factory is an asymmetric e^+e^- collider designed to operate at a luminosity of $3 \times 10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and above, at a center-of-mass energy of 10.58 GeV, which is the mass of the $\Upsilon(4S)$ resonance. As this resonance decays exclusively to $B^0\bar{B}^0$ and B^+B^- pairs, it provides an ideal laboratory for the study of B mesons.

The collider itself consists of a pair of storage rings which collide a 9.0 GeVelectron beam with a 3.1 GeV positron beam. The electron and positron beams are stored with respective currents of over 1.0 A and 1.5 A, in 1658 bunches approximately $120\mu\mathrm{m} \times 6\mu\mathrm{m} \times 9\,\mathrm{mm}$ in size. Dipole and quadrupole magnets are used to steer and focus the beams and bring them into collision; the centroid of the overlap of the beam

[1]Unless otherwise specified, all information in this chapter is derived from this reference.

profiles—the beam-beam interaction point, or IP—is monitored over time relative to the *BABAR* detector. The relative energies of the two beams are such that the Lorentz boost relative to the lab is $\beta\gamma = 0.56$.

2.2 The *BaBar* Detector

The study of B decays typically involves either partial or full reconstruction of the decay chain of the B meson down to the final-state particles: charged hadrons (π, K, p), charged leptons(e, μ), and photons. Intermediate states in the decay chain are reconstructed as composites of the final-state particles. Reconstruction of the decay chain and kinematics of the decay is rarely unambiguous, and optimal reconstruction requires good knowledge of the following:

- Momentum and charge of charged tracks,
- Particle identification of charged tracks,
- The energy and direction (momentum) of photons.

The *BABAR* detector consists of the five sub-detector components shown in Figure 2.1, each of which provides complementary information about the final-state products of the B decay. From the innermost to outermost, the sub-detectors, together with their primary tasks, are:

- Silicon Vertex Detector (SVT): Precise tracking of charged particles near the interaction region, and measurement of energy loss (dE/dX).
- Drift Chamber (DCH): Precise measurement of momentum and trajectory of charged particles, and measurement of energy loss (dE/dX).
- Ring-imaging Cerenkov Detector (DRC): Charged particle identification, particularly pion/kaon/proton discrimination.
- Electromagnetic Calorimeter (EMC): Position and energy measurement of photons and leptons. Hadron discrimination, and particularly electron identification.

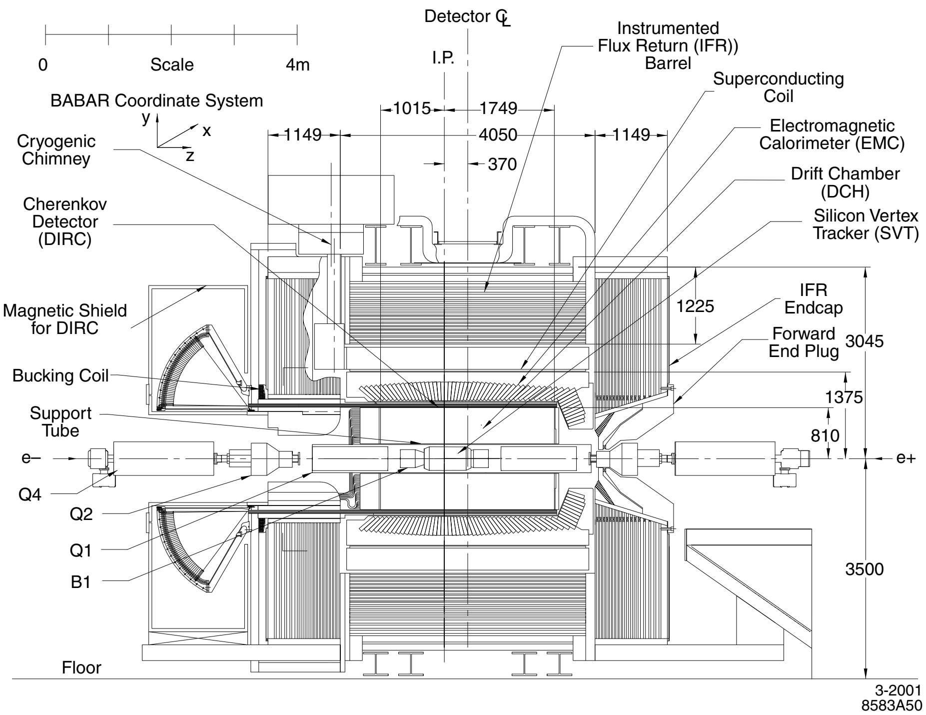


Figure 2.1: Longitudinal section of the *BABAR* detector.

- Instrumented Flux Return (Ifr): Neutral hadron and $\mu^{\pm}$ identification.

All detector components but the Ifr are embedded in the 1.5 T superconducting solenoid; the curvature of a charged track in the magnetic field allows determination of the momentum and charge of the particles.

All *BABAR* detector systems share a common electronics architecture. The front-end electronics (FEE) for any detector component are mounted directly on the detector system; the FEE chain digitizes the detector signals, buffers the digitized output, and forwards that information to the trigger system. Once triggered, the output of an FEE is transferred to storage via *readout modules*, or ROMs, which connect to the FEE circuits via 1.2 Gbits/s fiber optic cables and provide the standard interface between the detector-specific electronics of the FEE and the fast-control and timing system (FCTS) as well as the *event builder*. Subsystem-specific feature extraction, in which the revelant features of the raw data (e.g. integrated charge, shape, and timing of digitized waveforms) are extracted, is also done in the ROMs.

2.3 The Silicon Vertex Tracker

2.3.1 Layout and Electronics

The *BABAR* Svt has been designed to precisely reconstruct charged particle trajectories and decay vertices near the interaction region. It also provides a measurement of ionization loss (dE/dx) which is supplementary to that provided by the Dch.

The Svt layout is depicted in Figure 2.2. The detector consists of five layers of double-sided silicon strip sensors, organized into three sets of six modules for the inner three layers, and sixteen and eighteen modules for the outer two layers The silicon sensors are double-sided; on one side, the readout strips run parallel to the beam (ϕ strips), while on the other, they run transverse to to the beam axis (z strips). The readout pitch varies from 50 to 210 μm; in most cases *floating strips* (strips that are not read out) lie between two readout strips.

Modules in the inner three layers, which primarily provide position and angle information for measurement of the vertex position, are straight and positioned close to

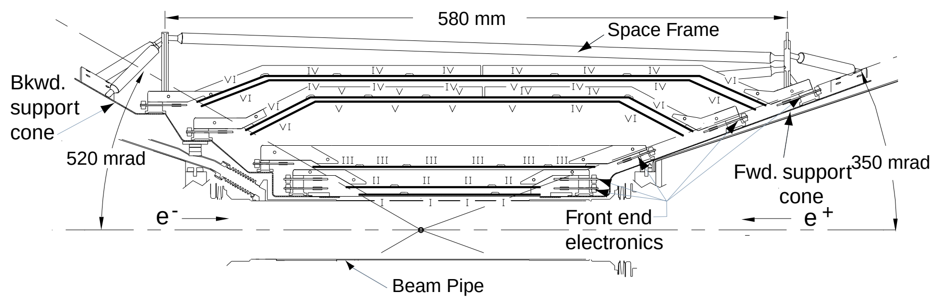


Figure 2.2: Longitudinal section of the SVT

the beam-pipe, in order to minimize the impact of multiple scattering on extrapolation to the vertex. Modules in the fourth and fifth layers are arch-shaped to increase solid angle coverage and the crossing angle for particles near the edges of the module acceptance. The forward acceptance of 350 mrad and the backward acceptance of 520 mrad, as well as the 32 mm radius of the innermost layer relative the interaction point, are determined by the radius of the beam pipe and the size and configuration of the magnets in the interaction region.

Data from the approximately 150,000 channels are delivered via fanout circuits to a custom integrated chip known as the ATOM (*A Time-Over-Threshold-Machine*). In the ATOM, the signal is processed by a charge-sensitive preamplifier and shaping circuit, and transformed by a programmable-threshold comparator into a pulse whose width is a quasi-logarithmic function of the collected charge. The comparator output is sampled at 15 MHz onto a 193 bin circular buffer. Upon receipt of a Level 1 (L1) trigger (see Section **??**), the time and time-over-threshold of the pulse are retrieved from the latency buffer, sparsified, and stored in a four-event buffer; if an L1 Accept is received, the time-over-threshold, the time stamp, and strip address are formatted, serialized, and delivered to the ROM.

2.3.2 Reconstruction and Performance

The reconstruction algorithm begins by discarding SVT hits with times more than 200 ns from the event time as determined by the DCH. The remaining in-time hits are passed on to the cluster finding algorithm, which derives the charge of the individual hits from their time-over-threshold values, and then groups hits from adjacent strips with consistent times into clusters. The cluster position is calculated from the positions of the individual strips, weighted by charge. The clusters are then passed on to both the SVT standalone and combined SVT-DCH tracking algorithms.

Accurate knowledge of both the SVT local and global alignments is critical if the SVT clusters are to be used in precise trajectory measurements. The SVT local alignment determines the relative positions of the individual SVT modules using primarily tracks from $e^+e^- \to \mu^+\mu^-$ events and cosmic rays; these track samples are supplemented with well isolated, high momentum tracks from hadronic events. The SVT global alignment determines the orientation of the SVT as a whole with respect to the DCH coordinate system using a sample of tracks with sufficient numbers of both SVT and DCH hits. The SVT and DCH components of these tracks are fit independently, and the differences between the respective track parameters (as a function of the six global alignment parameters) are minimized.

The local alignment is typically quite stable in time relative to the global SVT positioning. The derivation of the local alignments is complex; as a result, local alignments are performed relatively rarely, typically after magnet quenches or detector access. By contrast, the diurnal movement of the SVT with respect to the DCH requires that the global alignment procedure be performed approximately every 2-3 hours. The achieved spatial resolution for SVT hits, in both z and ϕ, varies between 20 and 40 μm depending on the angle of incidence of the track relative to the SVT module, while the mean dE/dx resolution for minimum-ionizing particles sampled over five layers is approximately 14%.

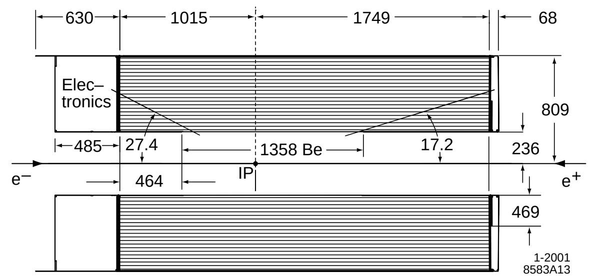


Figure 2.3: Longitudinal section of the DCH

2.4 The Drift Chamber

The BaBar drift chamber (DCH) is a tracking device which allows for the efficient detection of charged particles and precise measurement of their momenta, as well as the reconstruction of the decay vertices of long-lived particles such as the K_s^0, which may decay outside of the SVT. In addition, the drift chamber measures ionization loss (dE/dx), which provides particle identification information complementary to that provided by the other subsystems. This is particularly critical for low-momentum particles and those in the extreme forward and backward regions of the detector.

2.4.1 Design and Geometry

As shown in Figure 2.3, the DCH is enclosed by two concentric cylinders with radii of 236 mm and 809 mm, approximately 3m in length, and a pair of aluminum endplates. Gold-coated aluminum field wires form 7,104 densely packed hexagonal drift cells, each with a gold-coated tungsten-rhenium sense wires at the center. The cells are arranged in 40 cylindrical layers. Wires in 24 of the 40 layers are strung at small angles (between ±45mrad and ±76mrad) with respect to the z-axis, allowing the extraction of longitudinal as well as axial position information. The layers are grouped by four into ten superlayers; each layer of a superlayer has the same wire orientation (stereo angle) and equal numbers of cells. Sequential layers within a superlayer are staggered by a half a cell. The stereo angles of the superlayers alternate between the axial (A) and stereo (U and V) pairs, in the order AUVAUVAUVA, as shown in Figure 2.4.

The DCH coverage in azimuth is complete and uniform; the polar acceptance of

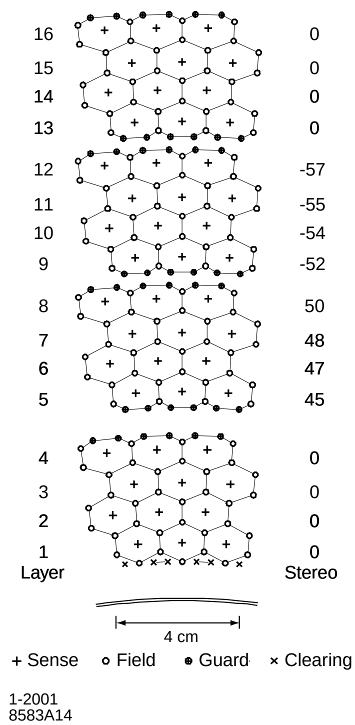


Figure 2.4: Schematic layout of the drift cells in the four inner DCH superlayers

the DCH as defined by the most extreme angle at which a particle from the origin crosses at least 20 layers, is 17.2° in the forward direction and 152.6° in the backward direction.

The need to minimize multiple scattering, which limits the track resolution below 1 GeV, dictates the choice of the physical materials used in the drift chamber construction, as well as the choice of a low-mass gas mixture (an 80:20 helium-isobutane mix). The inner cylindrical wall of the DCH is also kept thin to facilitate matching of SVT and DCH tracks and to minimize the background from photon conversions and interactions; the material in the outer wall and in the forward direction is minimized in order not to degrade the performance of the DRC and EMC.

2.4.2 Electronics and Readout

In order to keep the material in the forward direction to a minimum, the high-voltage (HV) distribution and all DCH readout electronics are mounted at the rear endplate of the chamber. The HV service boards provide the electrostatic potentials for sense, guard, and clearing wires, and pass both signals and ground to the front-end readout electronics. The front-end readout electronics consist of a set of wedge-shaped *Front-End Assemblies*, or FEAs, which plug into connectors on the back side of the HV service boards. Each of the wedge-shaped aluminum boxes of the FEAs contains from two to four amplifier/digitizer (ADB) boards.

The ADB boards themselves hold sets of two 4-channel amplifier integrated circuits(ICs) feeding a single 8-channel digitizer custom ASIC . The custom amplifier IC receives the input signal from the sense wire and produces both a discriminator output signal for the drift time measurement and a shaped analog signal for the dE/dx (integrated charge) measurement. The digitizer IC incorporates eight 4-bit Time-to-Digital Converters (TDCs) for time measurement and 6-bit 15MHz Flash-Analog-to-Digital Converters (FADCs) which sample 2.2 μs of the analog pulse.

Drift chamber-specific feature extraction algorithms convert the raw FADC and TDC information into an ordered set of drift times, the total charge, and a status word. The time and charge are both corrected channel-by-channel for time offsets,

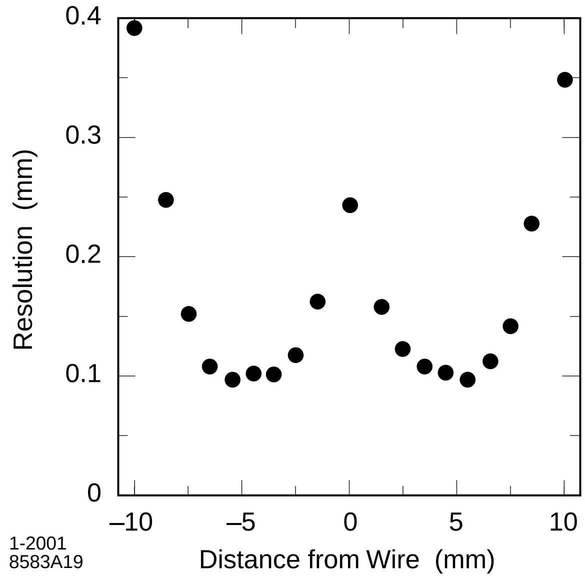


Figure 2.5: DCH single cell resolution.

pedestals, and gain constants, which are determined by a daily electronics calibration.

2.4.3 Calibration and Single-Cell Performance

Knowledge of both the drift time-to-distance relationship and the gas gain are required to determine the drift distance and ionization loss (dE/dx) from the recorded TDC times and accumulated charge. Calibrations for both the time-to-distance relation and dE/dx measurements were developed using cosmic ray data and then implemented for colliding beam data.

The relation between the measured drift time and drift distance is determined using tracks from e^+e^- scattering (Bhabha) and $\mu^+\mu^-$ production. A track trajectory is reconstructed using a set of "hits" (TDC times associated with particular drift cells); an estimated drift distance for a cell along the trajectory is determined by computing the distance of closest approach between the track and the signal wire. An average time-to-distance relation is determined for each layer, but separately for the right and left-hand sides of the sense wire, by fitting a sixth-order Chebychev polynomial to a set of estimated drift distances and measured drift times. Figure 2.5 shows the single-cell position resolution as a function of the drift distance for layer 18 of the DCH. The resolution is 100μm away from the boundaries of the cell, but worsens close to the sense wire and the outer cell boundary.

The specific energy loss ($\int_{cell}(dE/dx)dl$) for charged particles traversing the DCH is derived from the measurement of the total charge deposited in each drift cell, as computed by the feature extraction algorithm. The specific energy loss per track is computed as a truncated mean from the lowest 80 percent of the individual cell dE/dx measurements. Corrections are applied to compensate for changes in gas pressure and temperature, differences in cell geometry and charge collection, signal saturation due to space charge buildup, non-linearities in the most probable energy loss at large dip angles, and variation of charge collection as a function of entrance angle. The typical rms resolution, which is limited by the number of samples and Landau fluctuations, is about 7.5%.

2.4.4 Track Finding and Performance

Reconstruction of charged tracks relies on data from both the SVT and the DCH. Charged tracks are defined by five parameters: d_0 and z_0 (transverse distance and z coordinate at the point of closest approach of the track helix to the z axis), ϕ_0 (azimuthal angle), λ (dip angle with respect to the transverse plane), and $w = 1/p_t$ (track curvature). The track reconstruction builds on information from the Level 3 (L3) trigger and tracking algorithms (see Section 2.8.2), first refitting the trigger event time, t_0, and then performing helix fits to the hits found by the L3 tracking algorithm. A search for additional DCH hits that may belong to a track is performed, and additional track-finding algorithms employed to identify tracks which do not traverse the entire DCH or do not originate from the interaction point.

Tracks found by this algorithm are refit using a Kalman filter, which accounts for local variations in material and magnetic field along the fitting trajectory. They are extrapolated back into the SVT, where SVT track segments are added. Unassociated SVT hits are passed to a pair of standalone SVT track-finding algorithms.

By comparing the number of tracks found in the SVT that extrapolate into the

DCH acceptance to those actually found by the DCH the efficiency for DCH track-finding has been determined to be $98 \pm 1\%$. The tracking resolution in the four helix parameters and in tranverse momentum (p_t) are determined, in cosmic ray events, to be

$$\sigma_{d_0} = 23\,\mu\mathrm{m} \quad \sigma_{\phi_0} = 0.4\,\mathrm{mrad} \tag{2.1}$$

$$\sigma_{z_0} = 29\,\mu\mathrm{m} \quad \sigma_{\tan\lambda} = 0.53 \cdot 10^{-3} \tag{2.2}$$

and

$$\sigma_{p_t}/p_t = (0.13 \pm 0.01)\% \cdot p_t + (0.45 \pm 0.03)\% \tag{2.3}$$

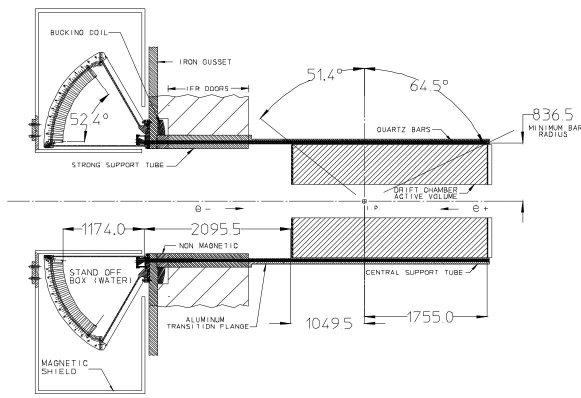
where p_t is measured in GeV.

2.5 The Ring-Imaging Cerenkov Detector (DRC)

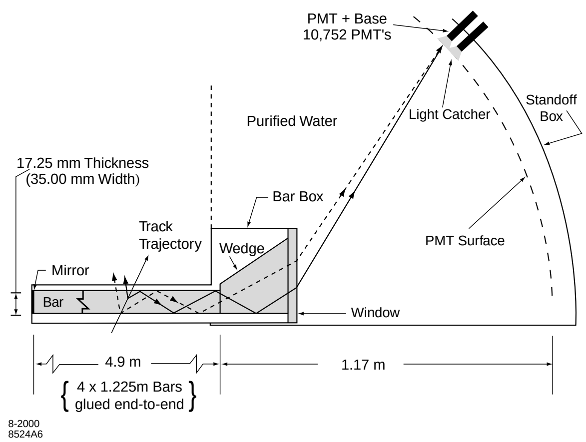
The DRC (Detector of Internally-Reflected Cherenkov light) is a novel ring-imaging Cherenkov radiation detector used for the identification of charged hadrons. The required momentum coverage of the DRC is dictated on the one hand by kaon tagging for time-dependent asymmetry measurements, where the typical momentum involved is less than 1 GeV, and on the other by K/π separation for the $B^0 \to \pi^+\pi^-/K^+\pi^-$ decays, where the relevant momenta lie between 1.7 and 4.2 GeV. The minimum transverse momentum required for a charged particle to traverse the DCH and reach the DRC is 280 MeV, which means there is no need for the DRC to have any sensitivity below this threshold.

2.5.1 Design and Geometry

The DRC consists of 144 synthetic fused-silica bars with a refractive index of 1.473, arranged in a 12-sided polygonal barrel around the DCH and an array of 10752 photomultiplier tubes mounted on the stand-off box (SOB) behind the rear IFR doors. The configuration is illustrated in Figure 2.6. Charged particles traversing the bars



(a)



(b)

Figure 2.6: (a) Elevated view of overall DRC geometry. (b) Bar/SOB transition region

emit Cherenkov radiation, which propagates by internal reflection to the photomultiplier array in the SOB, allowing reconstruction of the ring and determination of the Cherenkov angle.

The quartz bars are mounted in sets of 12 inside 12 aluminum bar boxes. These bars extend along the entire length of the DCH, covering polar angles down to 25.5° in the forward direction and 38.6° in the backward direction, and extend back through the IFR doors to the SOB. The water tank of the SOB flares out from the bars in a conical shape, with twelve sets of 896 29 mm diameter ETL 8125 photomultiplier tubes (PMTs) mounted on the back wall. At the end of each bar is a silica wedge prism, designed to recover photons at wide angles with respect to the bar axis via total reflection. The typical distance between the end of a bar and the photomultipliers is 1.17 m.

The index of refraction of water is 1.346, close enough to that of the bars to minimize reflection at the interface, and chromatic dispersion is minimized as well. The typical distance between a bar and the PMTs, along with the size of the bars and PMTs themselves, gives a geometric contribution to the single photon Cherenkov angle resolution of $\simeq 7\,\mathrm{mrad}$, a contribution somewhat larger than the approximately 5.4 mrad rms spread of the photon production and transmission dispersions. The overall single photon resolution is estimated to be about 10 mrad.

2.5.2 Electronics and Reconstruction

The DRC front-end electronics (FEE) are designed to measure the arrival time of each detected Cherenkov photon to an accuracy that is limited by the intrinsic 1.5 ns transit time spread of the PMTs, and to monitor pulse-height spectra in order to ensure that the PMTs are operating at a voltage which ensures a stable gain (HV plateau) and timing. The DRC FEE is mounted on the outside of the standoff box, and consists of a set of 168 DRC Front-end Boards (DFBs), each processing 64 PMT inputs. The four 16-channel custom-IC TDCs allow for an independent timing measurement for each TDC, while the single 8-bit flash ADC (FADC) multiplexes the pulse-height information for all 64 channels. The TDC has 0.5 ns binning, allowing the photon arrival time to be determined to better than the intrinsic 1.5 ns accuracy. The digitized information is shipped from the FEE to the ROMs via optical fibers.

Calibration of the DRC TDCs is achieved using 1 ns pulses from blue LED light pulsers; the LEDs, are pulsed at roughly 2kHz. Adjacent sectors are pulsed in a staggered fashion to prevent light crosstalk. Approximately 65,000 pulses per PMT are used in the calibration, to achieve a statistical accuracy of less than 0.1 ns. A complementary method compares observed and expected light arrival times associated with tracks in actual collision data; this method yields an improved resolution (about 15% better) and consistent results.

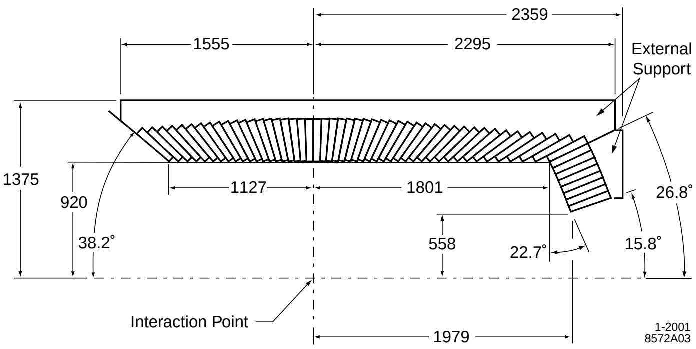
Reconstruction of the emission angle and the arrival time of the Cherenkov photons produced by a charged track in the DRC is done using observed space-time coordinates of the PMT signals, transformed into the Cherenkov coordinate system (θ_c and ϕ_c, the polar and azimuthal angles relative to the cone direction, and δt, the difference between the observed and expected arrival times). A set of three-dimensional vectors, from the end of a radiatior bar to the center of each coupled PMT, are extrapolated into the radiator bar using Snell's law and determined up to a 16-fold ambiguity. The uncertainties derive from the last reflection in the bar (forward/backward, left/right), whether the photon scattered off of the coupling wedge, and whether the photon initially propagated forward or backward. The timing resolution cannot provide competitive position information, but is used to suppress beam-induced background by about a factor of 40, to exclude other tracks in the same event, and to reduce the 16-fold ambiguity to a three-fold ambiguity. The reconstruction algorithm then maximizes the likelihood of the entire event, based on the individual track likelihoods for the electron, muon, pion, kaon, and proton hypotheses. When coupled with dE/dx information from the SVT and DCH, the DRC achieves better than 90% kaon identification efficiency, with a less than 3% pion misidentification rate, for tracks which intersect with the radiator bars and have momenta between 0.5 and 3 GeV.

2.6 The Electromagnetic Calorimeter

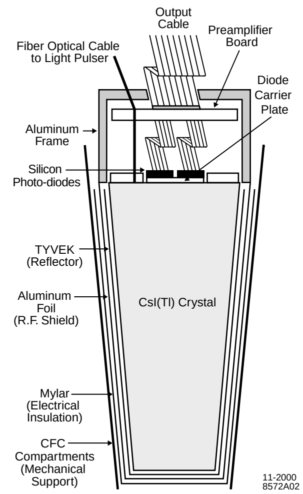
The electromagnetic calorimeter (EMC) is designed to measure electromagnetic showers with excellent efficiency, and provide energy and angular resolution over the energy range from 20 MeV to 9 GeV, allowing for the detection of photons from π^0 and η decays as well as from electromagnetic and radiative processes. Information about the shape of the electromagnetic showers detected by the EMC also makes the EMC the primary source of information for electron identification.

2.6.1 Geometry and Electronics

As sketched in Figure 2.7, the EMC consists of a cylindrical barrel, containing 48 rings of 120 Thalium-doped Cesium Iodide (CsI(Tl)) crystals, and a conical endcap,



(a)



(b)

Figure 2.7: (a) Longitudinal section of EMC. (b) Crystal housing with front-end electronics

containing rings of 120, 100, and 80 crystals each. It has full azimuthal coverage and a polar angle acceptance extending from 15.8° to 141.8°. The need for good energy and angular resolution dictated the choice of CsI(Tl), with its high light yield and small Molière radius; its short radiation length also allowed for a compact design. The fine segmentation provides the few mrad angular resolution needed to achieve good π^0 mass resolution above 2 GeV.

As shown in Figure 2.7, a pair of silicon photodiodes mounted at the end of each CsI(Tl) crystal registers the crystal light yield, and the signals are fed to a pair of low-noise preamplifiers. The amplified output is fed into the *custom auto-range encoding* (CARE) circuit; the total gain of the electronics chain is 256, 32, 4, or 1 for the four energy ranges 0-50 MeV, 50-400 MeV, 0.4-3.2 GeV, and 3.2-13.0 GeV. The two-fold redundancy of photodiodes and preamplifiers ensures reliability, since these components are inaccessible.

2.6.2 Reconstruction and Performance

A typical electromagnetic shower spreads over a number of adjacent crystals, forming a *cluster* of energy deposits. Pattern recognition algorithms have been developed to not only identify these clusters, but to differentiate clusters with a single energy maximum from those with multiple energy maxima, referred to as *bumps*. A cluster is required to contain at least one seed crystal with an energy above 10 MeV. Surrounding crystals are included in the cluster if their energy exceeds a 1 MeV threshold, or if they abut, in any direction, a crystal with at least 3 MeV of energy. Clusters are split into as many bumps as there are local maxima, and an iterative algorithm is used to determine the bump energies.

At low energies (around 6.13 MeV), a radioactive source calibration measures the fractional EMC energy resolution to be $5.0 \pm 0.8\%$; at higher energies (between 3 and 9 GeV). Bhabha scattering events are used to determine the resolution to be $1.9 \pm 0.07\%$. In the intermediate range (below 2 GeV), the energy resolution is inferred from the mass resolution of reconstructed $\pi^0 \to 2\gamma$ and $\eta \to 2\gamma$ decays, with the two photons of approximately equal energies.

The overall energy resolution may be parametrized by:

$$\frac{\sigma_E}{E} = \frac{(2.32 \pm 0.30)\%}{\sqrt[4]{E(\,\mathrm{GeV})}} \oplus (1.85 \pm 0.12)\%, \tag{2.4}$$

where the first term, dominant at low energies, encodes the statistical fluctuations in scintillation photon yield, beam background, and electronics noise, and the constant-energy term is associated with light leakage and absorption in front of and between crystals.

The angular resolution is determined solely from symmetric π^0 and η decays. A fit to an empirical parametrization of the energy dependence gives:

$$\sigma_\theta = \sigma_\phi = \left(\frac{3.87 \pm 0.07}{\sqrt{E(\,\mathrm{GeV})}} \oplus 0.00 \pm 0.04\right) \mathrm{mrad}, \tag{2.5}$$

which gives a resolution of about 12 mrad at low energies and 3 mrad at high energies. This slightly exceeds the performance predicted in simulation.

The reconstructed π^0 mass is measured to be 135.1 MeV and is stable to better than a percent over the full photon energy range, with a width of 6.9 MeV.

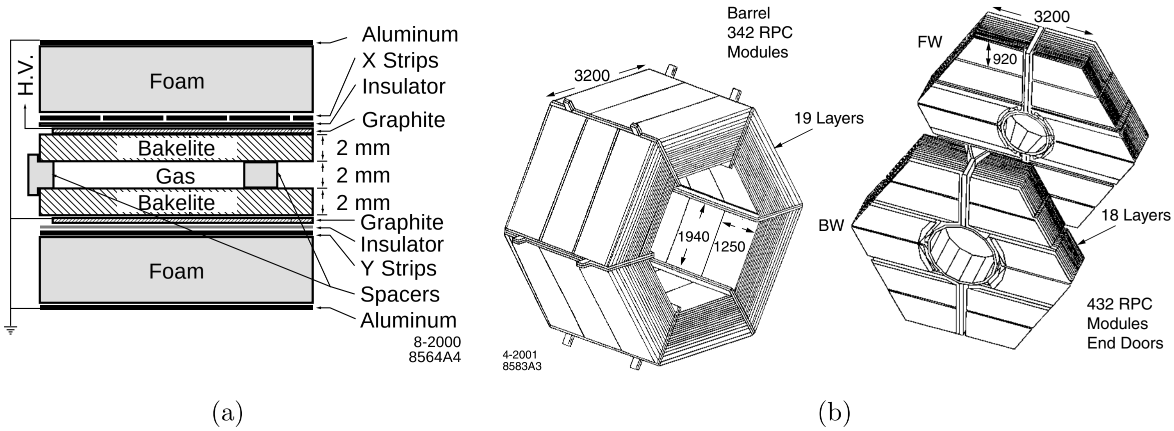


Figure 2.8: (a) Cross-section of a Resistive Plate Chamber (RPC). (b) Geometry of the IFRbarrel and endcaps

2.7 The Instrumented Flux Return

The Instrumented Flux Return (IFR) was designed to identify muons with high efficiency and good purity, and to detect neutral hadrons (primarily K_L^0 and neutrons) over a wide range of momenta and angles. It uses the steel flux return of the magnet as a hadron absorber. Single gap resistive plate chambers (RPCs) are installed the gaps of the finely segmented steel of the barrel and end doors of the flux return, as illustrated in Figure 2.8. There are 19 instrumented gaps (layers) arranged hexagonally around the barrel, and 18 in the end door steel.

2.7.1 The Resistive Plate Chambers

The RPCs detect streamers from ionizing particles via capacitive readout strips, and come in two geometric configurations. The planar RPCs consist of two bakelite sheets, 2 mm thick, separated by a 2 mm gap, and enclosed at the edge by a 7 mm wide frame. A uniform gap width is maintained by polycarbonate spacers, glued to the bakelite and spaced at distances of about 10 cm. The gap is filled with a mixture of 56.7% argon, 38.8% freon, and 4.5% isobutane at about 1500 torr of pressure. The two external surfaces of the bakelite sheets are coated in graphite; these graphite surfaces are connected to high voltage ($\simeq$ 8kV) and ground and protected by an insulating mylar film. The bakelite surfaces facing the gap are treated with linseed oil. The RPCs are operated in limited streamer mode and the signals read out capacitively, on both sides of the gap, by external electrodes made of aluminum strips on a mylar substrate.

The cylindrical RPCs are identical to the planar RPCs in terms of the gap thickness and spacer configuration, but are made of a set of resistive electrodes composed of conducting polymer and ABS plastic; these electrodes are laminated to fiberglass boards and foam to form a rigid cylindrical structure. Copper readout strips are attached to the fiberglass boards. Unlike the planar RPCs, no surface treatments of any kind have been applied.

The barrel RPCs are constructed in modules of $320 \times 130\,\mathrm{cm}^2$, with sets of three modules per gap and hexagonal face. A set of modules extending radially and covering the same region in z and ϕ define a *sector*. The readout strips in each barrel module are arranged with 32 strips perpendicular to the beam axis for z measurement, with another 96 strips orthogonal to the first set, and extending over the three-module set, to measure ϕ.

2.7.2 Reconstruction

Groups of adjacents hits in one of the two readout strip directions define a one-dimensional cluster; the IFR reconstruction joins clusters belonging to the same readout coordinate across layers to form two-dimensional clusters. Within each sector, these clusters are then grouped into a three-dimensional cluster, provided there are fewer than three IFR layers missing from one of the two coordinates.

Cosmic ray tests determined that during the early period of running, 75% of the active RPC modules had an efficiency of at least 90%. However, high-temperature conditions during early running, coupled with a problem in the curing of the linseed oil coating, caused a substantial fraction of the RPC modules to experience an ongoing degradation in efficiency.

The role and performance of the IFR in muon identification is discussed in section 2.9.2.

2.8 Trigger

PEP-II operating luminosities exceed the original design luminosity; at the design luminosity, beam-related background events occur at rates greater than 20 kHz, while useful events ($B\overline{B}$ $\tau^+\tau^-$, etc.) occur at rates of only a few Hz. The role of the trigger system is to select events of interest while rejecting the rest, thus reducing the total data-taking rate to level manageable for online reconstruction (less than 120 Hz). The trigger must operate with a high efficiency (at least 99% for $B\overline{B}$ events) and a low deadtime (no more than 1%).

The trigger is implemented as a two-tier system: Level 1 (L1), implemented in hardware, is designed to reduce the input rate to Level 3 to 1 kHz, while Level 3 (L3) is implemented in software. For historical reasons, there is no Level 2 trigger.

2.8.1 The Level 1 Trigger

The Level 1 (L1) trigger consists of three hardware components, each based on information from a specific subsystem: the DCT, based on DCH information, the EMT, based on EMC information, and the IFT, based on IFR information (IFT) and used primarily for diagnostic purposes.

2.8.1.1 The Drift Chamber Trigger

The DCT takes data from the DCH cells and implements a fast 24 module *Track Segment Finder* (TSF) using the ϕ coordinate and the drift times of the DCH hits. The segments are passed to the *Binary Link Tracker* (BLT), which bins the segments into *supercells* dividing the DCH into 32 ϕ bins and 10 superlayers. The BLT links segments in contiguous supercells, starting from the innermost drift-chamber layer. Eight transverse momentum discriminator modules (PTD) determine the number of tracks above a certain threshold. The output of the DCT is a set of 16 bit *trigger primitives* which categorize the BLT and PTD results into short tracks (traversing half the DCH), long tracks (traversing the entire DCH), and high P_t ($> 800\,\mathrm{MeV}$) tracks.

2.8.1.2 The Electromagnetic Calorimeter Trigger

For trigger purposes, the EMC is divided into 280 towers, 7 x 40 ($\theta \times \phi$). For each tower, all crystal energies above a 20 MeV threshold are summed and sent to the EMT every 269 ns. The towers are also grouped into the 40 *ϕ-sectors*. The patterns of energy deposition and arrival time in the ϕ-sectors are used to form a set of trigger primitives.

The DCT and EMT inputs are processed by a *global trigger* (GLT) to form specific L1 triggers. Processing times for the DCT and EMT both are about $5\,\mu$s with an additional $4\,\mu$s for the GLT to process and intitiate readout by the ROMs. The combined L1 triggers achieve nearly 100% efficiency for generic $B\overline{B}$ events.

2.8.2 The Level 3 Trigger

The Level 3 trigger (L3) involves a basic reconstruction of the event in both the DCH and EMC. As such it consists of a track finding algorithm for the DCH and a clustering algorithm for the EMC. The kinematics and topology of the reconstructed event, as given by the results of these two algorithms, allows the event to be categorized for acceptance or rejection.

The track-finding algorithm uses a Monte Carlo-derived lookup table of hit patterns in order to join track segments from the TSF to form a track. If a pattern of segment hits matches an entry in the lookup table, the reconstructed track is refined by an iterative fitting algorithm, which adds or drops hits based on their proximity to the fitted trajectory. The L3 clustering algorithm forms clusters from adjacent energy depositions that are within 1.3 μs of the event time, and have more than 20 MeV of energy. For clusters with at least 100 MeV of energy, the centroid, lateral energy profile, and average cluster time are calculated.

2.8.2.1 Performance

L3 information allows QED processes useful for calibration purposes, such as (radiative) Bhabha scattering or $e^+e^- \rightarrow 2\gamma$ events, to be identified and passed at reduced rates, while multiplicity criteria identify hadronic events from both $B\overline{B}$ decays and the continuum. The combined L1 and L3 trigger achieves better than 99.9% efficiency for $B\overline{B}$ events and better than 95% efficiency for continuum events.

2.9 Particle Identification

Particle identification requires a synthesis of the information available from multiple detector subsystems. *BABAR* does not identify particles via a unified likelihood or common algorithm; rather, identification of the different particle types is performed using a set of independently optimized selectors. Each of these individual selectors uses either a set of independent cuts on individual discriminating variables, or some type of likelihood function of these same discriminating variables. For a given type and technique of particle identification, a class of particle selectors exists, which ranges from a very loose criterion with high efficiency and low purity, to an extremely restrictive criterion with lower efficiency but the best achievable purity.

Particle identification selections are developed, and their performance determined, using a set of kinematically selected data control samples. Control samples exist for the following:

- Electrons: Pure samples of electron tracks are obtained from Bhabha and radiative Bhabha events, $e^+e^- \rightarrow e^+e^-e^+e^-$ events, and photon conversions.
- Muons: Samples of muons come from kinematically identified $\mu^+\mu^-(\gamma)$ events and two-photon production of $\mu^+\mu^-$ pairs, $e^+e^- \rightarrow e^+e^-\mu^+\mu^-$
- Pions: Charged pions from the process $K_s^0 \rightarrow \pi^+\pi^-$ are easily identified, as the K_S^0 lifetime $(0.89 \times 10^{-10}s)$ means its decay vertex is outside the interaction region. Various kinematic cuts, such as a cut on the $\pi\pi$ opening angle, and a cut on the $\pi\pi$ invariant mass, are used to suppress combinatoric background and backgrounds from photon conversions and Dalitz decays. A sample of pions from topologically selected 3-1 prong $\tau^+\tau^-$ events is used to provide an additional control sample for pions with momenta above 1.8 GeV.

- Kaons: A control sample of kaons is taken from the decay chain $D^* \to D^0\pi_{soft}$, $D^0 \to K^-\pi^+$; this control sample also provides an additional sample of pions.

- Protons: The proton control sample is extracted from decays $\Lambda \to p\pi$; this sample has a 98% purity overall.

2.9.1 Electron Identification

2.9.1.1 Criteria

Separation of electrons from muons and charged hadrons rests primarily on information obtained from the electron interaction with the electromagnetic calorimeter (Emc). All *BABAR* electron identification algorithms require a high-quality charged track (one with at least 12 drift-chamber hits) whose extrapolation to the front face of the Emc matches with an Emc cluster. This algorithm matches clusters within a combined distance in polar angle ($\Delta\Theta$) and azimuthal separation ($\Delta\Phi$), between the centroid of the Emc cluster and the impact point of the charged track on the Emc[19].

One of the primary discriminating variables for electron identification is E/p, the ratio of the energy of the Emc cluster to the measured momentum of the matched charged track. Electrons, which should deposit most or all of their energy in the electromagnetic calorimeter, typically have an E/p close to 1, while muons, which, interact with the Emc as minimum ionizing particles, have a small E/p, as do non-interacting charged hadrons.

Some fraction of charged hadrons, interacting hadronically with the Emc do deposit a substantially larger fraction of their energy, and may not easily be distinguished from electrons via E/p alone. Both the longitudinal and lateral development of an electromagnetic shower differ from that of a hadronic shower; while the *BABAR* Emc has no sensitivity to the longitudinal shower development, variables describing the lateral shower shape offer additional discriminating power. In this case, the shower shape variables used include: the number of calorimeter crystals in the reconstructed shower, the lateral moment of the shower (LAT), and a Zernike moment

(Z_{42}), which is sensitive to the azimuthal variation in lateral shower shape. Zernike moments are defined as:

$$A_{nm} = \sum_{r_i \leq R_0}^{n} \frac{E_i}{E} \cdot f_{nm}(\frac{r_i}{R_0}) \cdot e^{-im\phi_i}, \qquad R_0 = 15\,\mathrm{cm} \tag{2.6}$$

where

$$f_{nm}(\rho_i \equiv \frac{r_i}{R_0}) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)! \rho_i^{n-2s}}{s!((n+m)/2-s)!((n-m)/2-s)!} \tag{2.7}$$

.

At low energies, electron-hadron separation is also enhanced by taking into account both energy loss ($\mathrm{d}E/\mathrm{d}x$) in the drift chamber and the response of the DRC using both Cerenkov angle θ_C and the number of photons N_C.

2.9.1.2 Performance

Electron identification efficiency is measured using the samples of pure electrons discussed in section 2.9; the misidentification rates for pions, kaons, and protons are likewise extracted using their respective control samples. The average hadron fake rates per track are determined separately for positive and negative particles, taking into account the relative abundance from Monte Carlo simulation of $B\overline{B}$ events, with relative systematic uncertainties of 3.5%, 15% and 20% for pions, kaons, and protons, respectively. The resulting average fake rate per hadron track of $p_{lab} > 1.0\,\mathrm{GeV}/c$, is of the order of 0.1% for pions and 0.2% for kaons.

The efficiency and hadron misidentification rates for the most critical electron selection in this analysis are summarized in Figures 2.9, 2.10, 2.11, 2.12. For the electron selection, muon misidentification is essentially zero.

2.9.2 Muon Identification

2.9.2.1 Criteria

Charged tracks identified as muons are required to have a minimum momentum in the laboratory $p_{lab} > 1.0\,\mathrm{GeV}$; if the charged track has a matched cluster in the EMC

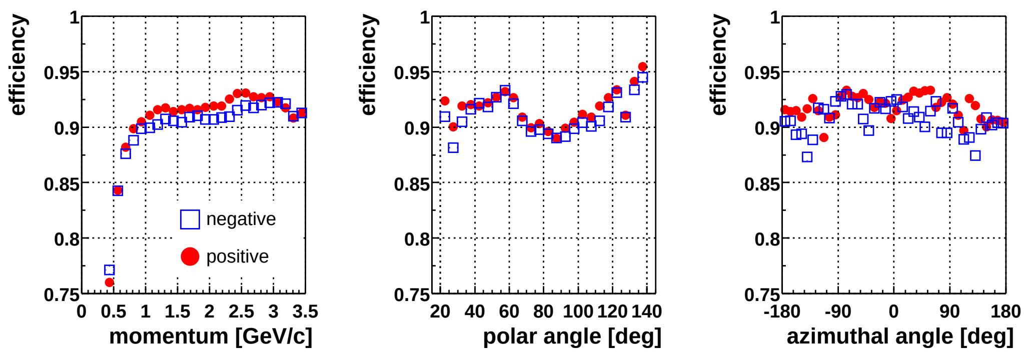


Figure 2.9: Efficiency of the electron selector as a function of laboratory momentum (left), polar angle (center), and azimuthal angle (right).

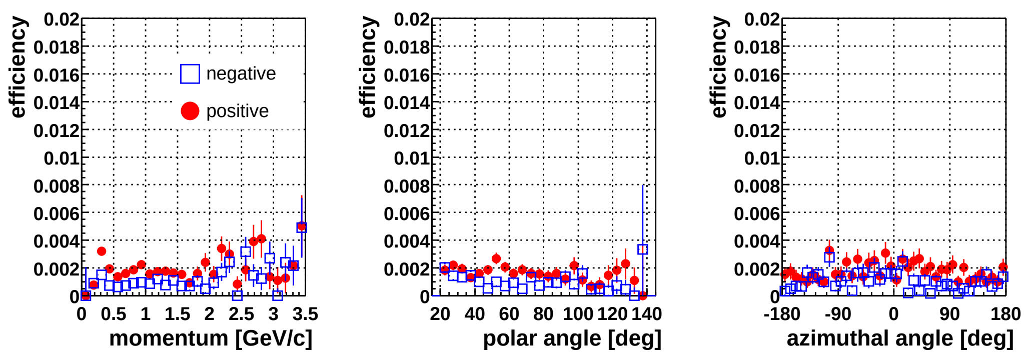


Figure 2.10: Pion misidentification probabilities of the electron selector as a function of laboratory momentum (left), polar angle (center), and azimuthal angle (right).

the energy deposited in the EMC is required to be consistent with a minimum ionizing particle, 50 MeV $< E_{cal} <$ 400 MeV. Muon identification, however, is primarily based on quantities characterizing the interaction of the muon with the IFR

- the number of IFR layers associated with the track has to be $N_L \geq 2$.
- the interaction lengths of material traversed by the track has to be $\lambda_{meas} > 2.2$.
- The number of interaction lengths expected for a muon of the measured momentum and angle to traverse is estimated by extrapolating the track up to the last

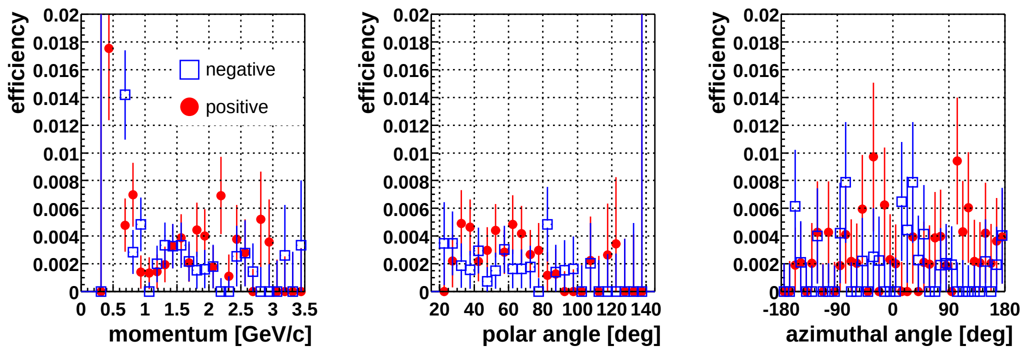


Figure 2.11: Kaon misidentification probabilities of the electron selector as a function of momentum (left), polar angle (center), and azimuthal angle (right).

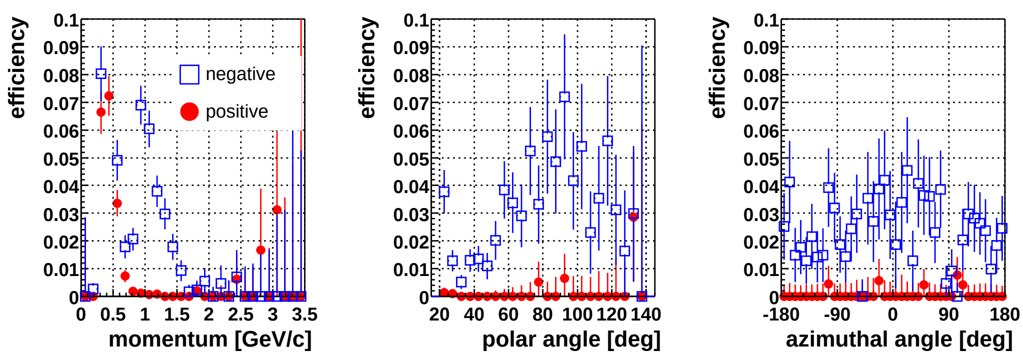


Figure 2.12: Proton misidentification probabilities of the electron selector as a function of momentum (left), polar angle (center), and azimuthal angle (right).

active layer of the IFR. This estimate takes into account the RPC efficiencies which are routinely measured and stored. For the difference $\Delta\lambda = \lambda_{exp} - \lambda_{meas}$ we require $\Delta\lambda < 1.0$, for tracks with momentum greater than 1.2 GeV/c. For track momenta between 0.5 GeV/c and 1.2 GeV/c, a variable limit is placed: $\Delta\lambda < [(p_{lab} - 0.5)/0.7]$.

- The *continuity* of the IFR cluster is defined as $T_c = N_L/(L - F + 1)$, where L and F are the last and first layers with a hit. T_c is expected to be 1.0 for muons penetrating an ideal detector whereas it is expected smaller for hadrons. We require $T_c > 0.3$ for tracks with $0.3 < \theta_{lab} < 1.0$ (i.e. in the Forward End Cap to remove beam background).

- The observed number of hit strips in each RPC layer is used to impose the conditions on the average number of hits, $\bar{m} < 8$, and the standard deviation, $\sigma_m < 4$.
- The strip clusters in the IFR layers are combined to form a track and fit to a third degree polynomial, with the quality of the fit selected by the condition $\chi^2_{fit}/DOF < 3$. In addition, the cluster centroids are compared to the extrapolated charged track, with the requirement $\chi^2_{trk}/DOF < 5$.

2.9.2.2 Performance

For the more restrictive muon selection, in the early running period, the average muon identification efficiency is approximately 76%, with a pion misidentification rate of $2 - 3\%$. While the misidentification rate is relative constant, the degradation of the RPCs results in an overall drop in the muon identification efficiency of $\simeq 1\%$ per month, leading overall to about a 20% drop over the period of interest.

2.10 Kaon Identification

2.10.1 Criterion and Performance

Charged kaon candidates are selected from a likelihood ratio formed from dE/dX information for low momentum tracks ($< 0.5 - 0.7$ GeV, depending on the selector) in both the SVT and DCH as well as DRC reconstruction information. There are two main variants of kaon identification currently in use, distinguished mainly by the way in which they use DRC information above the kaon Cherenkov threshold [20]. One uses a global likelihood function involving individual photon information [21], while the other uses a likelihood function which involves an intermediate fit to the common Cherenkov angle and time distribution [20]. For historical reasons, selectors of the second type are used to identify kaons in the selection of the $B \to D^* \ell \nu$ control sample 5.1 while the kaon veto used elsewhere in the analysis is based on the first. In all cases, rather than returning a single likelihood, the individual selectors use weighted cuts on the likelihood for the different particle hypotheses, and return a binary answer for a given track.

The average kaon identification efficiency, for a high-purity selection, is about 70% for momenta up to about 3.5 GeV, with a pion misidentification rate below 1%. The loosest selection has an efficiency above 90% with a pion misidentification rate of about 2%.

Chapter 3

Data Sample and Selection

This analysis is based on *BABAR* data recorded in the years 1999-2002. The following physics processes contribute to the data:

- $e^+e^- \to \Upsilon(4S) \to B\overline{B}$
- $e^+e^- \to q\bar{q}(\gamma)$ where $q = u, d, s, c$ (continuum)
- $e^+e^- \to e^+e^-(\gamma)$ (QED continuum). This category of events is produced by Bhabha scattering of the electron and positron beams; some percentage of these events are radiative. As the Bhabha cross-section is extremely high (approximately a factor of 40 larger than the $B\overline{B}$ cross-section), and as these events hold little physics interest aside from their use in calibration, events tagged as Bhabhas at the L3 level are not included in the trigger lines used in this analysis.
- $e^+e^- \to \mu^+\mu^-(\gamma)$, $\tau^+\tau^-$ (QED continuum). These are pure QED processes for muon and tau pair production.
- Other pure QED processes such as $e^+e^- \to 2\gamma$.

A total integrated luminosity of $75.5\,\mathrm{fb}^{-1}$ taken on the $\Upsilon(4S)$ resonance is used; this dataset includes events of all of the above types, including approximately 83 million $B\overline{B}$ pairs. An additional sample of $8.85\,\mathrm{fb}^{-1}$ was taken approximately

Table 3.1: Integrated luminosities, numbers of $B\bar{B}$ events, and $\sigma_{B\bar{B}}$ are obtained using standard *BABAR* tools. The first error in $N_{B\bar{B}}$ and $\sigma_{B\bar{B}}$ is statistical, the second is systematic.

Data sets	$\int L\,(\mathrm{pb}^{-1})$	$N_{B\bar{B}}$ (10^3)	$\sigma_{B\bar{B}}$ (nb)
2000-b1-on	8456	$9156 \pm 16 \pm 101$	$1.08 \pm 0.002 \pm 0.01$
2000-b1-off	1022	—	—
2000-b2-on	8794	$9650 \pm 16 \pm 106$	$1.10 \pm 0.002 \pm 0.01$
2000-b2-off	1384	—	—
2001-b1-on	35098	$39112 \pm 33 \pm 430$	$1.11 \pm 0.0009 \pm 0.01$
2001-b1-off	3292	—	—
2002-b1-on	23163	$25089 \pm 27 \pm 276$	$1.08 \pm 0.001 \pm 0.01$
2002-b1-off	3152	—	—
Σ*(on)*	75511	$83007 \pm 48 \pm 913$	$1.10 \pm 0.0006 \pm 0.01$
Σ*(off)*	8850	—	—

40 MeV below the $\Upsilon(4S)$ mass, below the threshold for $B\overline{B}$ production; thus, only the $q\bar{q}$ production and QED processes contribute.

Table 3.1 shows a summary of all data subsets used. The number of $B\overline{B}$ events was obtained by subtracting the luminosity scaled number of multihadron events in on- and off-resonance data samples, corrected for the $B\overline{B}$ efficiency. This measurement has an estimated systematic uncertainty of 1.6%. The luminosity of each data sample has been obtained by counting e^+e^- and $\mu^+\mu^-$ events in the central detector region. The absolute luminosity is determined with a systematic uncertainty of 1.5%. The relative luminosity for on- and off-resonance data is known to about 0.25%.

3.1 Preliminary Event Selection

Applying the following preliminary event selection criteria significantly reduces backgrounds from lepton pair production and $q\bar{q}$ events; it also reduces demands on CPU time and disk space by selecting for $B\overline{B}$ events likely to contain a primary semileptonic B decay.

The criteria are as follows:

- Basic trigger information: the preselection requires a primary trigger signal from either the DCH or the EMC.

- Event topology: Whereas B mesons are produced nearly at rest, and their decay products are uniformly distributed, $q\bar{q}$ events have a more jet-like event topology, with the angular distributions of the jets peaked along the beam axis. The second normalized Fox-Wolfram moment, R_2, is strongly sensitive to this difference in event shape; $R_2 = 0$ for a completely isotropic event, while $R_2 = 1$ for a completely collimated event. The preselection requires $R_2 < 0.6$. Figure 3.1 shows the $R2$ distribution for both off-resonance data and simulated $q\bar{q}$ and $B\overline{B}$ events.

- Multiplicity: lepton-pair events in particular have a much lower multiplicity of charged tracks than expected in $B\overline{B}$ events. The preselection requires that the number of charged tracks $N_{track} > 3$.

- Presence of a high-momentum lepton: For an event to contain a primarily semileptonic B decay that may be successfully reconstructed, it must also contain at least one lepton. Demanding that that lepton have a substantial momentum in the $\Upsilon(4S)$ rest frame suppresses background from photon pair-conversions and secondary decays. The preselection requires that the event contain at least one charged lepton with a momentum of $p^*_{tag} > 1.3\,\mathrm{GeV}/c$, measured in the rest frame of the $\Upsilon(4S)$ resonance. Figure 3.2 shows the momentum distribution for the highest-momentum lepton in the event for both off-resonance data and simulated $q\bar{q}$ and $B\overline{B}$ events.

Table 3.2 summarizes the percentage of Monte-Carlo $B\bar{B}$, $c\bar{c}$, and $u\bar{u}/d\bar{d}/s\bar{s}$ events surviving the preselection. The fact that the preselection efficiency for $c\bar{c}$ events is nearly twice as high as that for $u\bar{u}/d\bar{d}/s\bar{s}$ events is consistent with the fact that $u\bar{u}/d\bar{d}/s\bar{s}$ events rarely contain leptons, while $c\bar{c}$ events contain leptons from semileptonic decays of charm mesons. The 17.9% preselection efficiency for $B\bar{B}$ events is consistent with what we would expect if we kept mostly events containing a semileptonic B decay. The preselection efficiency for off-resonance data demonstrates that

Table 3.2: Table of efficiencies for the event-level preselection

Event type	Efficiency (%)	Nominal cross-section (nb)
generic $B\bar{B}$	17.9	1.09
off-resonance data	0.86	—
$c\bar{c}$	6.5	1.30
$u\bar{u}/dd/s\bar{s}$	3.5	2.03

the initial selection is more effective at suppressing pure QED continuum background as it is at suppressing $q\bar{q}$ events.

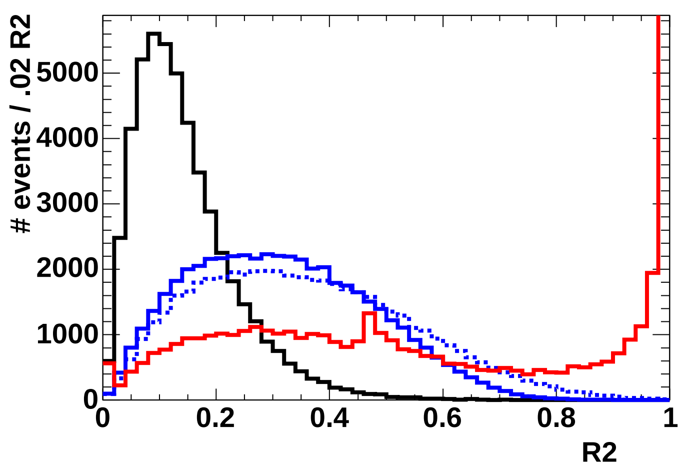


Figure 3.1: R2 distribution for simulated events: $B\bar{B}$ (black), $c\bar{c}$ (blue), and $u\bar{u}/d\bar{d}/s\bar{s}$ (blue, dashed), and for offpeak data (red).

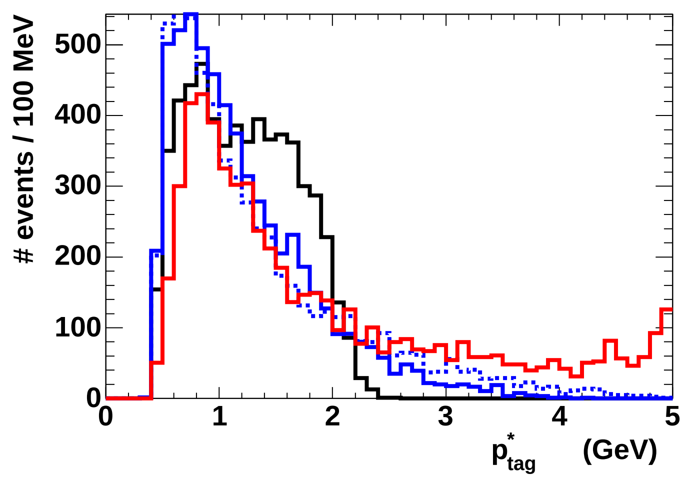


Figure 3.2: p^*_{tag} distribution for simulated events: $B\bar{B}$ (black), $c\bar{c}$ (blue), and $u\bar{u}/d\bar{d}/s\bar{s}$ (blue, dashed), and for offpeak data (red).

Chapter 4

Analysis: Candidate Selection

The goal of the analysis is to reconstruct the following exclusive charmless semileptonic B decays:

$$B^0 \to \pi^- \ell^+ \nu, \tag{4.1}$$
$$B^+ \to \pi^0 \ell^+ \nu, \tag{4.2}$$

and

$$B^0 \to \rho^- \ell^+ \nu, \tag{4.3}$$
$$B^+ \to \rho^0 \ell^+ \nu. \tag{4.4}$$

The candidate selection criteria can be divided into two main categories: those criteria designed to identify well-reconstructed signal decays (i.e., signal decays where all detectable elements of the decay are detected and correctly identified, and the neutrino is reconstructed with a reasonable resolution), and those designed to suppress background.

4.1 Signal Selection

From a physics standpoint, it is natural to partition a semileptonic decay into the daughters of the virtual W (lepton and neutrino), and the hadron. From the point of view of the reconstruction, however, it is more useful to partition the reconstructed semileptonic B decay into indirectly and directly reconstructed components: the neutrino, which is taken from the missing energy and momentum of the event, and the "Y" system, which consists of a charged lepton ($e^{\pm}$ or $\mu^{\pm}$), and the hadron, which is reconstructed from one or more charged and/or neutral pions.

4.1.1 Neutrino Reconstruction

"Neutrino reconstruction" is something of a misnomer, in the sense that the neutrino is not detected directly, but is instead taken to be the difference between the net four-momentum of the colliding beams and that of the tracks and showers captured in the detector. The four-momentum characterizing this different is

$$(\vec{p}_{miss}, E_{miss}) = (\vec{p}_{beams}, E_{beams}) - (\sum_i \vec{p}_i, \sum_i E_i), \tag{4.5}$$

where $\vec{p}_i$ and E_i are the three-momentum and energy, respectively, of the ith track or EMC shower measured in the laboratory system, and $\vec{p}_{beams}$, E_{beams} refer to the sum of the three-momenta and energies of the two colliding particles.

This technique involves several implicit assumptions: first, that the charged tracks and EMC showers used in the neutrino reconstruction come solely from the beam interaction, second, that the momentum and energy associated with a particle is counted only once, and third, that undetected momentum and energy in an event is associated only with the neutrino. None of these assumptions, of course, is entirely accurate. Beam background may produce additional reconstructed tracks and showers. Likewise, both physics processes and reconstruction errors may result in double-counting of the energy associated with a track or shower; for instance, if a shower resulting from the interaction of a charged particle in the EMC is not correctly associated with the corresponding track, it may be counted separately, adding energy to the event.

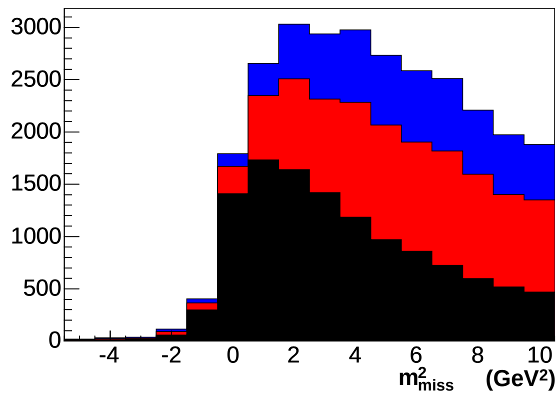


Figure 4.1: Monte Carlo simulation of m^2_{miss} for $B \to \pi\ell\nu$ events with no other cuts applied. Black: events with no K_L, neutrons or additional ν. Red indicates events with at least one K_L and/or neutron, blue those with an additional ν.

Finally, not only does the limited detector acceptance mean that a substantial fraction of tracks and showers go undetected, but either B decay may involve additional undetectable neutral particles, such as a K_L, neutron, additional primary neutrino, or secondary neutrino. Initial state radiation for $B\bar{B}$ events plays only a small role in contributing undetectable neutral energy, since the width of $\Upsilon(4S)$ resonance is small.

The degree to which these effects impact the neutrino measurement can only be assessed via simulation. Figure 4.1 shows the missing-mass squared of the event, defined as

$$m^2_{miss} = E^2_{miss} - |\vec{p}_{miss}|^2 \tag{4.6}$$

for a sample of simulated $B^0 \to \pi^-\ell^+\nu$ events. For a fully reconstructed event with a single semileptonic B decay, m^2_{miss} should be consistent with zero within measurement errors. Failure to detect one or more particles in the event, however, creates a highly asymmetric m^2_{miss} distribution that peaks above zero and has a substantial tail at large positive values. A substantial fraction of this tail comes from events where additional lost energy is unavoidable, such as those containing a K_L or additional neutrino. Nonetheless, the limited detector acceptance means that even for events with no additional undetectable neutrals, the missing-mass squared distribution still peaks above zero and has a significant positive tail. Double-counted and/or background tracks and EMC showers also have a visible impact, extending the missing-mass squared distribution to unphysical negative values.

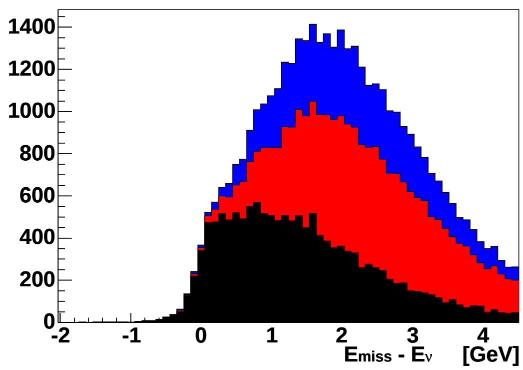
Figure 4.2 shows, for the same sample of simulated $B^0 \to \pi^- \ell^+ \nu$ events, the following quantities: the difference between the reconstructed missing energy and that of the true neutrino, $E_{miss} - E_\nu$, the difference between the magnitude of the reconstructed missing momentum and that of the true neutrino, $|\vec{p}_{miss}| - |\vec{p}_\nu|$, and the angle of the reconstructed neutrino with respect to the true neutrino direction, $\Delta\theta_{miss,\nu}$. As with the missing-mass squared distribution, all three of these distributions are highly asymmetric, with tails for large positive values (i.e. the measured value exceeds the true value). The reconstruction of the missing energy is substantially worse than that of the magnitude of the missing momentum; $E_{miss} - E_\nu$ is broader than $|\vec{p}_{miss}| - |\vec{p}_\nu|$, and unlike $|\vec{p}_{miss}| - |\vec{p}_\nu|$ peaks far above zero. This difference between the momentum and energy resolutions is not surprising. First, the missing momentum is a vector sum. Contributions from particle losses (or additional energy from extra tracks and EMC showers) do not add linearly as is the case for E_{miss}. Second, the energy calculation is more strongly influenced by the fact that all charged tracks are assumed to have the mass of a pion. By calculating the missing momentum and energy in the laboratory frame rather than in the rest from of the $\Upsilon(4S)$ resonance, these additional uncertainties are confined to the missing energy.

4.1.1.1 Track and Photon Selection

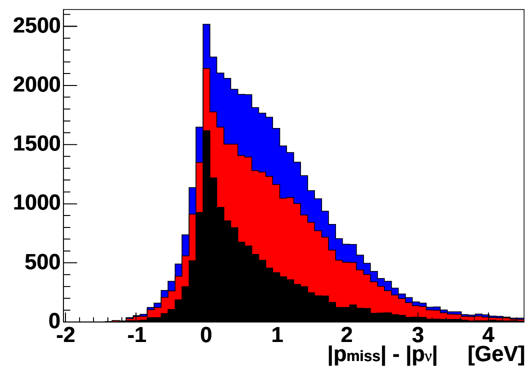
The track and shower selection used in neutrino reconstruction must balance the need to capture as much of the event as possible with the need to eliminate spurious momentum and energy from either beam background or reconstruction errors. To this end, an extremely loose track and shower selection is used, with additional cuts imposed to reject particular sources of badly measured or spurious track and showers.

Charged tracks are rejected if they meet any of the following criteria:

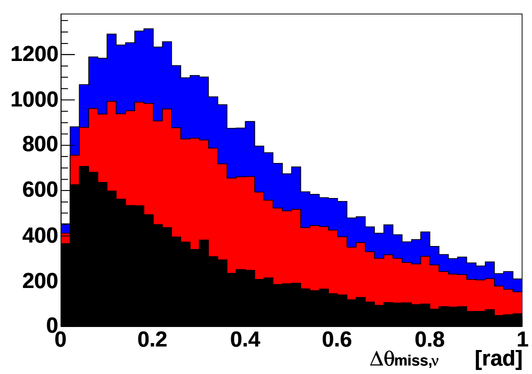
- the distance of closest approach to the beam spot in the x-y plane, $|d_{xy}| > 1.5\,\text{cm}$



(a)



(b)



(c)

Figure 4.2: Resolutions $E_{miss} - E_{\nu}$ (a), $|\vec{p}_{miss}| - |\vec{p}_{\nu}|$ (b), and $\Delta\theta_{miss}$ (c), before any cuts have been made. Events containing a K_L or additional ν are shown in red and blue, respectively. The remaining tails come from losses due to detector acceptance.

- the distance of closest approach to the beam spot along the z axis, $|d_z| \geq 10\,\mathrm{cm}$ These criteria reject fake tracks and background tracks not originating from the vicinity of the beam-beam interaction point.
- the laboratory momentum $p_{lab} > 10\,\mathrm{GeV}$. Such tracks are incompatible with the beam energies.

In addition, special criteria are used to identify and reject duplicate tracks and fake tracks in the SVT[22]:

- So-called "loopers" are tracks with a transverse momentum $p_\perp < 0.18\,\mathrm{GeV}$ that spiral inside the drift chamber. Looper candidates are identified as two tracks with a small difference in $p_\perp$, ϕ and θ. Of such a pair only the segment with the smallest distance $|d_z|$ to the beam IP is retained.
- Split tracks occur when the tracking algorithms assigns DCH hits generated by a single track to two different tracks. The resulting two tracks overlap in space. Only the track with the larger number of associated DCH hits, and thus the better resolution, is retained.
- Fake tracks in the SVT generally have no associated hits in the DCH. Thus tracks with transverse momentum $p_\perp > 0.2\,\mathrm{GeV}$ are rejected if they have no associated DCH hits.

These criteria remove roughly 13% of all low-momentum tracks. On average, the mean observed charged multiplicity per B meson is reduced by less than 1%.

A photon is selected as a local maximum of the energy depositions in the EMC which is not matched to a track. Any photon which has an energy less than 30 MeV, contains fewer than three crystals, or has a lateral moment of greater than 0.6, is rejected; these criteria remove spurious photons from electronic noise, low-energy photons from beam-related background, and help to suppress hadronic showers due to interactions of either K_L, neutrons or charged particles. Showers generated by the interaction of charged tracks in the calorimeter are not always successfully associated with the generating track; cutting on the angular separation between the centroid of

Table 4.1: Summary of Track Rejection Criteria.

Reject tracks with	Cut
distance in $x-y$ plane	$\lvert d_{xy}\rvert > 1.5\,\mathrm{cm}$
distance along z axis	$\lvert d_z\rvert \geq 10\,\mathrm{cm}$
maximum momentum	$p_{lab} > 10\,\mathrm{GeV}$
ghosts in SVT ($p_\perp > 200\,\mathrm{MeV}$)	$N_{Dch} = 0$
if	$\Delta p_\perp < 100\,\mathrm{MeV}$ w.r.t. other tracks AND
loopers ($p_\perp < 180\,\mathrm{MeV}$)	Same sign: $\lvert\Delta\phi\rvert < 220$ mrad &
	$\lvert\Delta\theta\rvert < 215$ mrad &
	Opposite sign: $\lvert\Delta\phi\rvert < 190$ mrad &
	$\lvert\Delta\theta\rvert < 300$ mrad
split tracks ($p_\perp < 350\,\mathrm{MeV}$)	$\lvert\Delta\phi\rvert < 220$ mrad &
	$\lvert\Delta\theta\rvert < 215$ mrad

a shower and the point of impact of the track in the EMC eliminates most of these showers (see Table 4.2).

Table 4.2: Summary of Photon Rejection Criteria.

Variable	Cut value
Shower energy	$E_{cal} > 30\,\mathrm{MeV}$
Shower shape	LAT> 0.6
	$N_{cry} \leq 2$
Nearest track impact	$\lvert d\theta\rvert < 30$ mrad
w.r.t. shower	$(-30 < Qd\phi < 70)$ mrad

4.1.1.2 Quality Cuts

The distributions shown in Figure 4.2 show an unacceptably broad missing-momentum resolution for the reconstructed neutrino, due primarily to undetected particles. Since it is impossible to either improve the detector acceptance or reduce the number of undetectable neutral particles in the average B decay, this resolution can only be improved by eliminating those events where lost particles have significantly compromised the neutrino reconstruction. This is accomplished by the following set of quality cuts:

- Q_{tot}: As the loss (or addition) of one or more charged particles impacts the measurement of the missing momentum and energy, it is logical to require that the event be charge balanced. On the other hand, the strict requirement $Q_{tot} = 0$ would lead to a significant loss of signal events, primarily due to the incomplete geometrical acceptance of the *BABAR* detector. A looser cut, $Q_{tot} \leq 1$, is made instead.

- $m^2_{miss}/2E_{miss}$: While the resolution of m^2_{miss} is broad, a cut on this variable is still a very powerful tool to reject events in which more than one particle is undetected or very poorly measured. The resolution in m^2_{miss} increases linearly with E_{miss}, a fact we take into consideration by cutting on $|m^2_{miss}/2E_{miss}|$ rather than on m^2_{miss}. The cut used, $|m^2_{miss}/2E_{miss}| < 0.4\,\mathrm{GeV}$, is illustrated in Figure 4.3.

- θ_{miss}: In cases where the signal neutrino lies outside the detector acceptance in either the forward or backward region, another lost particle, particularly a photon that is collinear with the neutrino, may add to the missing energy and momentum without adding appreciably to the missing-mass squared. Requiring $0.6 < \theta_{miss} < 2.9\,\mathrm{rad}$ suppresses these events even when the other quality cuts fail to do so.

Figure 4.4 shows the neutrino resolution for signal events selected by the neutrino quality cuts. The relative fraction of events containing a K_L or additional ν becomes quite small, while the resolutions for all quantities improve dramatically. The magnitude of the missing momentum, is now, on average, a fairly good measure of the true neutrino momentum. Since the missing energy, even after quality cuts, still gives a noticeably poorer approximation of the neutrino energy, it is preferable at this point to substitute $|p_{miss}|$ for E_{miss}.

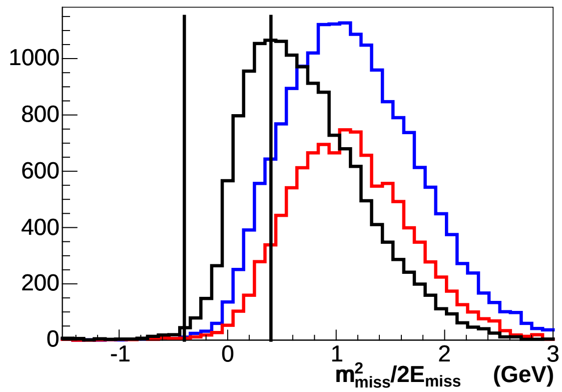


Figure 4.3: $m^2_{miss}/2E_{miss}$ for $B \to \pi\ell\nu$ events with no other cuts applied. Black: events with no K_L, neutrons, or additional ν. Red indicates events with at least one K_L/neutron, blue those with an additional ν. The vertical lines indicate the neutrino quality cut used in this analysis.

4.1.2 Y System Selection

While the track selection criteria used in neutrino reconstruction is kept as loose as possible in order to improve efficiency, it is preferable to impose stricter criteria when selecting the charged tracks for the signal (Y) lepton and hadron. The loss in efficiency is small, while the tighter criteria suppress background and ensure a more reliable momentum measurement. We require:

- $d_z < 3.0$ cm,
- the transverse momentum $p_\perp > 0.1\,\mathrm{GeV}/c$,
- and the number of associated drift chamber hits, $N_{Dch} \geq 12$.

In order to ensure a more reliable identification efficiency, electron and muon candidates are also required to have a laboratory polar angle in the range $0.409 < \theta_{lab} < 2.37$ rad (within the well-calibrated region of the EMC acceptance). Only lepton candidates with a momentum in the $\Upsilon(4S)$ frame $p_l^* > 1.3\,\mathrm{GeV}$ are considered, which restricts the analysis to a momentum range in which both electron and muon identification is reliable, and suppresses the background from secondary semileptonic decays and photon conversions.

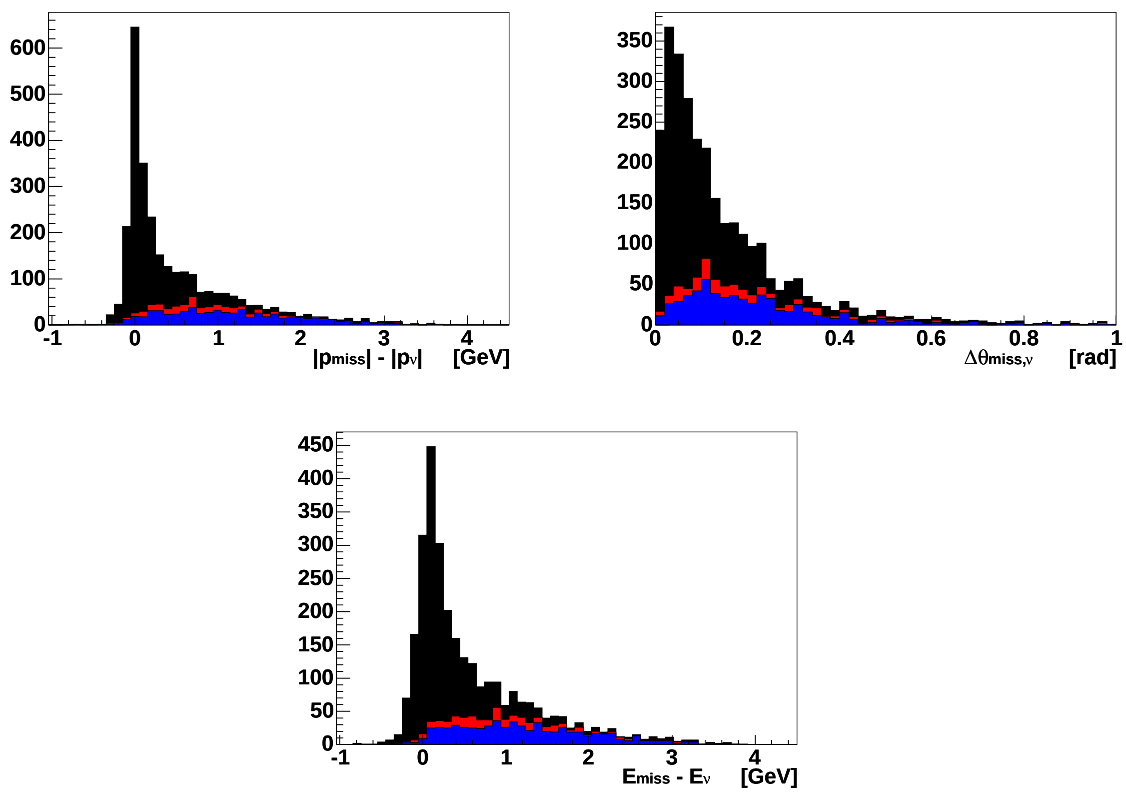


Figure 4.4: Resolutions (a) $|\vec{p}_{miss}|-|\vec{p}_\nu|$, (b) $\Delta\theta_{miss}$, and (c) $E_{miss}-E_\nu$, after requiring $m^2_{miss}/2E_{miss} < 0.4\,\mathrm{GeV}$. The component containing at least one K_L and/or neutron is shown in red; the component containing at least one additional ν in blue.

To search for the signal candidates, electron and muon candidates are combined with π^+, π^0, $\rho^+ \to \pi^\pm\pi^0$, and $\rho^0 \to \pi^+\pi^-$. For B^0 decays, the charge of the lepton is required to be opposite that of the hadron system, and all charged tracks in the Y system are included in the vertex fit, which must yield a minimum χ^2 probability of 0.1%. π or ρ candidates containing a track which is a signal-quality lepton are also vetoed, as are ρ candidates containing tracks which pass the tight kaon selection.

The π^0 candidates are reconstructed from pairs of photons. These photons must satisfy the same quality criteria used in the neutrino reconstruction. The effective mass is reconstructed under the assumption that both photons originate from the primary vertex, and only pairs with $|M(\gamma,\gamma) - M_{\pi^0}| < 17.5\,\mathrm{MeV}$ are retained. For ρ^+ and ρ^0 candidates, the effective mass of the two pions is restricted to $0.65 < M(\pi,\pi) < 0.85\,\mathrm{MeV}$.

4.1.2.1 Kinematic Consistency Requirements

The following criteria are designed to select those Y candidates for which the four-momentum $p_Y = p(\pi,\rho) + p_\ell$ is consistent with a $B \to Y\nu \to (\pi,\rho)\ell\nu$ decay.

In semileptonic B decays the four-momentum of the neutrino can be expressed as

$$p_\nu^2 = 0 = (p_B - p_Y)^2 = M_B^2 + M_Y^2 - 2(E_B E_Y - |\vec{p}_B||\vec{p}_Y|\cos\theta_{\mathrm{BY}}). \tag{4.7}$$

The momentum $|\vec{p}_B|$ and energy E_B are known; if we assume we have a perfectly reconstructed semileptonic decay, we can determine θ_{BY}, the angle between the B meson and the Y,

$$\cos\theta_{\mathrm{BY}} = \frac{2E_B E_Y - M_B^2 - M_Y^2}{2|\vec{p}_B||\vec{p}_Y|}. \tag{4.8}$$

Figure 4.5 shows the definition of this angle.

The distribution of this variable for $B^0 \to \pi^-\ell^+\nu$ signal candidates is shown in Figure 4.6; bremsstrahlung, which primarily affects electrons, creates the tail at unphysical negative values of $\cos\theta_{\mathrm{BY}}$. In order to reduce the number of events with multiple candidates per channel, the event-level preselection imposes a initial cut of $|\cos\theta_{\mathrm{BY}}| < 1.5$. The full selection requires $|\cos\theta_{\mathrm{BY}}| < 1.1$, retaining only candidates

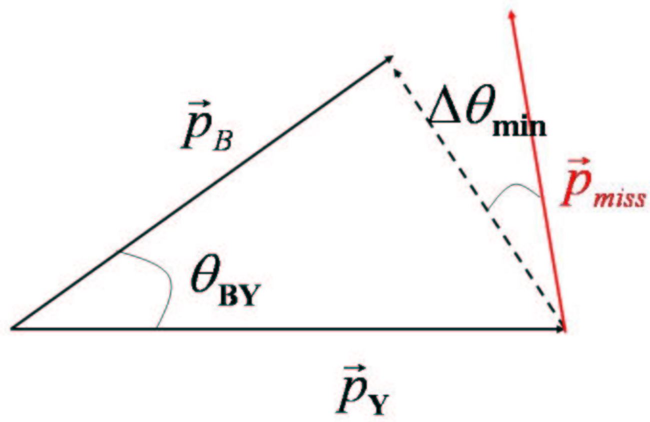


Figure 4.5: Illustration of the relationship between $\cos\theta_{\mathrm{BY}}$ and $\Delta\theta_{min}$

with a physical value of $\cos\theta_{\mathrm{BY}}$ within resolution.

The direction of the neutrino momentum, $\vec{p}_\nu$, may be inferred from the relation $\vec{p}_\nu = \vec{p}_B - \vec{p}_Y$; this direction may be compared with the direction of the missing momentum $\vec{p}_{miss}$, yielding an additional constraint.

Since only the magnitude and not the direction of $\vec{p}_B$ is known, $\vec{p}_\nu$ can only be determined up to a rotation around the Y direction. The angle $\Delta\theta$ between $\vec{p}_{miss}$ and $\vec{p}_\nu$ can be written as

$$\cos\Delta\theta = \frac{\vec{p}_\nu \cdot \vec{p}_{miss}}{|\vec{p}_\nu||\vec{p}_{miss}|} = \frac{\vec{p}_B\vec{p}_{miss} - \vec{p}_Y\vec{p}_{miss}}{|\vec{p}_\nu||\vec{p}_{miss}|} \tag{4.9}$$

with

$$\begin{aligned} \vec{p}_B\vec{p}_{miss} &= |\vec{p}_B||\vec{p}_{miss}|\cos\theta_{B,miss} \\ \vec{p}_Y\vec{p}_{miss} &= |\vec{p}_Y||\vec{p}_{miss}|\cos\theta_{Y,miss} \,. \end{aligned} \tag{4.10}$$

The smallest angular difference, denoted by $\Delta\theta_{min}$, occurs when the vectors lie in

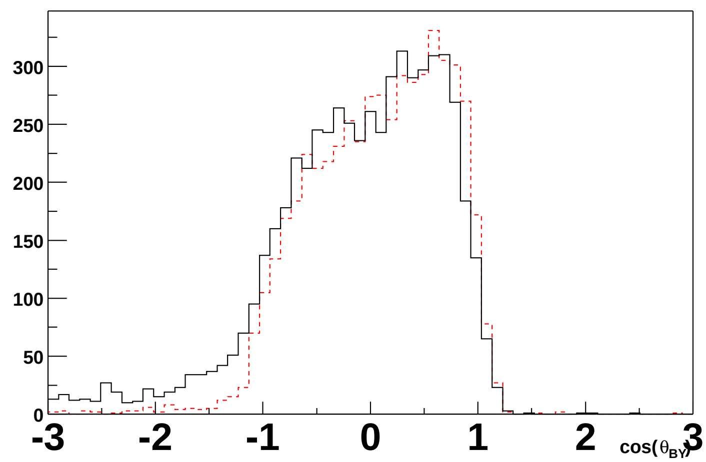


Figure 4.6: Monte Carlo simulation for $\cos\theta_{\rm BY}$ for $B \to \pi\ell\nu$ signal candidates; for electrons (black) and muons (red, dashed).

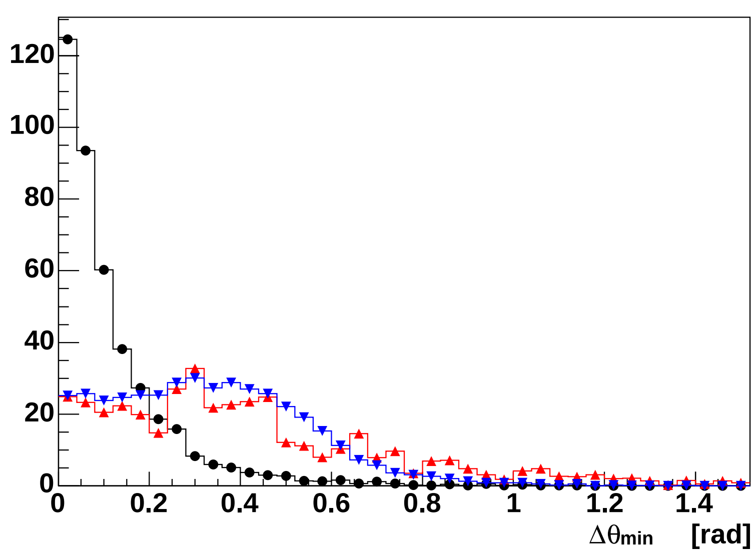


Figure 4.7: Monte Carlo simulation of $\Delta\theta_{min}$ distributions for reconstructed $B \to \pi\ell\nu$ candidates, after all other cuts have been applied. Shown are signal $B \to \pi\ell\nu$ decays (black, filled dots), background from generic B decays (blue, filled triangle), and background from $q\bar{q}$ events (red, open triangle). All distributions have been scaled to equal area.

a common plane, thus we have

$$\begin{aligned}\cos\theta_{B,miss} &= \cos\left(\theta_{B,Y}-\theta_{Y,miss}\right) \\ &= \cos\theta_{B,Y}\cos\theta_{Y,miss}+\sin\theta_{B,Y}\sin\theta_{Y,miss}\ .\end{aligned} \tag{4.11}$$

Taking into account the constraints $|\vec{p}_\nu| = E_B - E_Y$, $\Delta\theta_{min}$ can be calculated. The distribution of $\Delta\theta_{min}$ is shown in Figure 4.7. We require

$$|\Delta\theta_{min}| < 0.6. \tag{4.12}$$

4.2 Background Classification

In addition to candidates that represent successfully reconstructed decays in any signal channel, a variety of physics processes may, through errors in reconstruction, successfully mimic a signal decay. As will be discussed in section 4.4, there are several orders of magnitude more background candidates than successfully reconstructed signal decays for any channel. The different physics processes that produce background candidates occur at widely varying rates; they also have significantly different kinematics. As a result, it is inappropriate to treat this background in a monolithic fashion. Instead, we divide the simulated background into a set of *sources*, based on the origin of the candidate lepton.

The candidate lepton is the most natural criterion for a number of reasons. First, for the $B \to \rho\ell\nu$ signal channels, background candidates from $B\bar{B}$ events may contain reconstructed hadrons with tracks from both B decays. The reconstructed lepton, on the other hand, if its origin can be identified, is uniquely associated with one of the two B decays. Second, in general, the origin of the candidate lepton in simulation can be reliably identified. The lepton and q^2 spectra of the signal decays are of particular interest; to first order, we reconstruct both using only the candidate lepton and reconstructed neutrino. By distinguishing the candidate sources according to the origin of the lepton, we naturally divide the sources according their underlying physics and corresponding effect on these two spectra.

The various candidate sources and the critera used to determine them are shown in Figures 4.8 and 4.9. In general, an additional distinction is maintained between the sources coming from charged and neutral B decays, in order to be able to separate decay modes related by isospin and vary the relative B^+/B^0 production rates.

The background source where the lepton comes from the signal decay channel, but the hadron has been incorrectly reconstructed, is referred to as "combinatoric signal," and is something of a special case. It acts as signal in that the size of its contribution to the total yield is proportional to the branching fraction for the signal decay, but in the measurement of the signal q^2 spectrum, where the information from the reconstructed hadron is used to improve the resolution of the reconstructed q^2, it is a combinatoric background.

4.3 Simulation

In order to obtain the breakdown of background sources in data, we must rely on the predictions provided by simulation. This analysis utilizes a variety of samples with different simulated physics processes, most of them simulated by the `EvtGen` generator [23]. Regardless of the underlying physics involved, all simulated samples are then processed by the GEANT-based detector simulation and the output run through standard *BABAR* tracking and reconstruction algorithms [24], which treat them as if they were actual data.

Simulated samples of $B^0 \to \pi^-\ell^+\nu$, $B^+ \to \pi^0\ell^+\nu$, $B^0 \to \rho^-\ell^+\nu$, and $B^+ \to \rho^0\ell^+\nu$ decays used in this analysis were generated using the ISGW2 model; a secondary sample for each decay mode was generated using uniform phase space. The charmless semileptonic background is accounted for using exclusively produced samples of $B \to \omega\ell\nu$, $B \to \eta\ell\nu$ and $B \to \eta^{'}\ell\nu$, all of which are generated using ISGW2 form-factor models, and an inclusive sample of non-resonant charmless decays, which is generated with a smooth hadronic mass spectrum, using the triple differential decay rate $d\Gamma/dx/dz/dq^2$ defined by de Fazio and Neubert (FN) [25]. This generator cannot produce hadronic states with mass below $2m_\pi$.

The overall normalization of the non-resonant sample is adjusted to give the proper

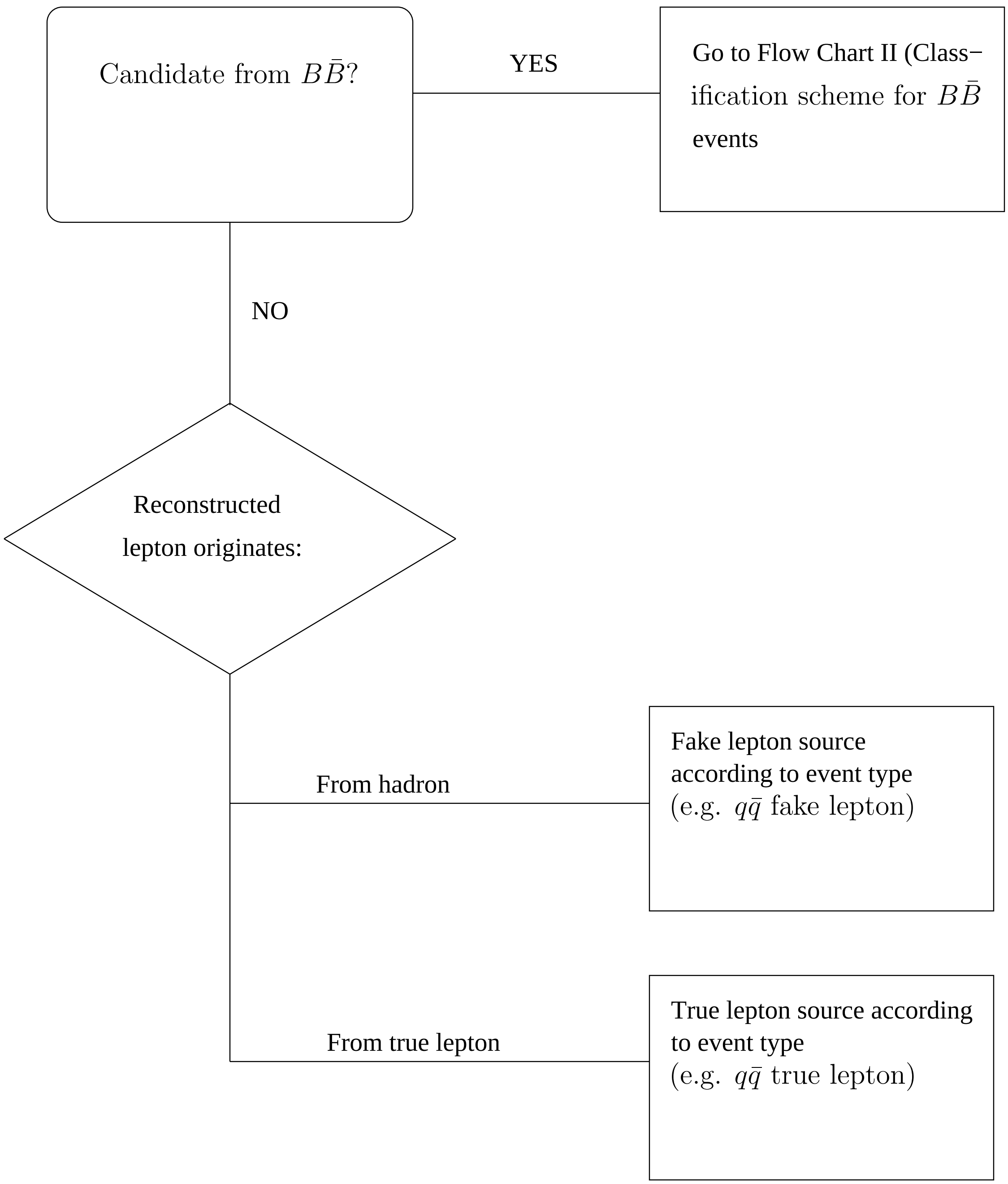


Figure 4.8: Flowchart illustrating classification procedure for candidates from simulated events.

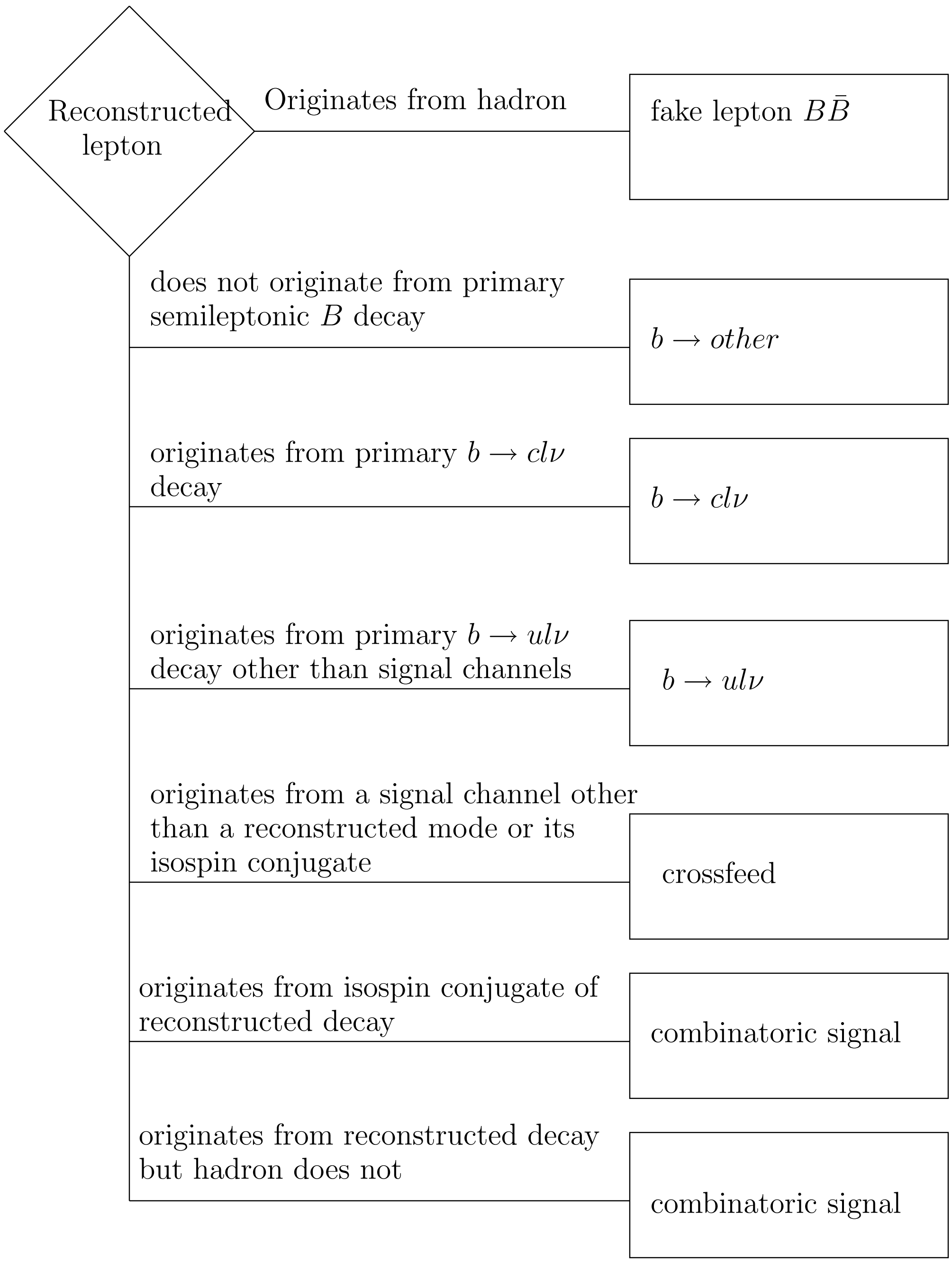


Figure 4.9: Classification procedure, in detail, for candidates from simulated $B\overline{B}$ events.

$b \to u\ell\bar{\nu}$ branching fraction, and the inclusive sample is weighted in bins of q^2, the lepton energy E_l, and the hadron mass m_X, such that the weighted combination of the non-resonant and resonant decay samples approximates the fraction of events per bin of the non-resonant sample alone [26].

Samples of simulated generic B^+B^- and $B^0\bar{B}^0$ events are used to extract the remaining background sources from B decays: the $b \to c\ell\nu$ source, where the candidate lepton is a primary lepton from a true $B \to X_c\ell\bar{\nu}$ decay, the $b \to$ *other* source, where the candidate lepton comes from a secondary decay, such as $B \to J/\psi \to l^+l^-$, and the $B\bar{B}$ fake lepton source, where the candidate lepton is a misidentified hadron from B decays.

The relative composition of the semileptonic B decays (both charmless and charm) is adjusted to yield the best agreement for the lepton spectra in all channels, before the full selection has been applied. While the branching fractions chosen take recent experimental results as a guide, they do not necessarily correspond to the current world averages for these branching fractions. The branching fractions used are given in Table 8.4.

The dominant semileptonic decays, $B \to D^*\ell\nu$, have been generated according to heavy quark effective theory (HQET), with the following values for the parameters—$R_1 = 1.4$, $R_2 = 0.8$, $\rho^2 = 0.85$—based on a preliminary *BABAR* measurement [27]. The decays to higher mass non-resonant states, $B \to D^{**}\pi\ell\nu$, are simulated according to Goity and Roberts [28]. The remaining decays are implemented according to ISGW2.

4.3.1 Continuum Background

Since the off-resonance data sample is very limited, we rely on simulation, both for the optimization of the event selection, and as an estimator of the normalization and shape of this background. Continuum background is simulated using the standard generator for $e^+e^- \to q\bar{q}$ with $q = u, d, s, c$ combined with fragmentation via JETSET [29]; in general, the pure QED processes such as lepton-pair production are not included in this simulation.

Table 4.3: Summary of Monte Carlo samples used in this analysis, specifying the equivalent luminosity for an assumed cross section σ or branching fraction.

Sample	Generator	# events [10^6]	L equiv. [fb^{-1}]	$\sigma[nb]$ or BR
B^+B^- generic		167	317.6	0.525
$B^0\bar{B}^0$ generic		164	312.9	0.525
$c\bar{c}$	JETSET	105	80.8	1.3
$u\bar{u}, d\bar{d}, s\bar{s}$	JETSET	166	81.9	2.03
$B^+ \to \eta\ell^+\nu$	ISGW2	0.04	432.9	$8.4 \cdot 10^{-5}$
$B^+ \to \eta^{'}\ell^+\nu$	ISGW2	0.08	865.8	$8.4 \cdot 10^{-5}$
$B \to H_u\ell\nu$ incl	FN	2	4035.9	$6.2 \cdot 10^{-4}$
$B^0 \to \pi^-\ell^+\nu$	ISGW2	0.3	4067.3	$1.33 \cdot 10^{-4}$
$B^+ \to \pi^0\ell^+\nu$	ISGW2	0.1	4124.6	$0.72 \cdot 10^{-4}$
$B^0 \to \rho^-\ell^+\nu$	ISGW2	0.3	2294.8	$2.69 \cdot 10^{-4}$
$B^+ \to \rho^0\ell^+\nu$	ISGW2	0.2	2495.9	$1.45 \cdot 10^{-4}$
$B^+ \to \omega^0\ell^+\nu$	ISGW2	0.2	2469.6	$1.45 \cdot 10^{-4}$

4.3.2 Sample Inventory

Table 4.3 shows a summary of the Monte Carlo simulated event samples and the equivalent luminosity. The $B\overline{B}$ simulation assumes equal production of B^+B^- and $B^0\bar{B}^0$, $i.e. f_{00} = f_{++} = 0.5$.

4.3.3 Corrections to Detector Simulation

A number of corrections are applied to the results of the detector simulation to compensate for known deficiencies. The primary corrections involve particle identification efficiencies and misidentification rates. For all selected decays with an identified electron or muon, candidates are assigned a weight based on the ratio of the identification efficiency as determined from control samples in data and Monte Carlo simulated events. Likewise, for a misidentified pion/kaon/proton, the candidate track is reweighted based on the ratio of the misidentification probabilities as determined in data and Monte Carlo. A similar weight is applied to the candidates to correct for the effect of lepton (and kaon) vetos applied to the charged hadron tracks.

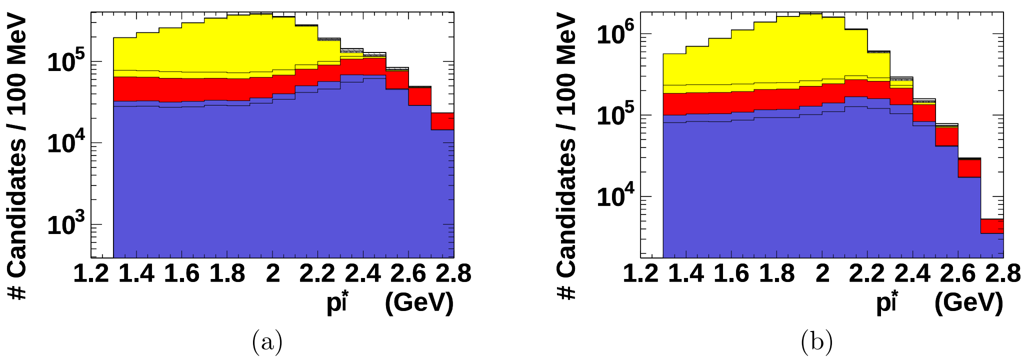


Figure 4.10: Lepton spectrum (on a log scale) for (a) $B^0 \to \pi^- \ell^+ \nu$ and (b) $B^0 \to \rho^- \ell^+ \nu$ candidates before any selection cuts have been applied. Candidates where the lepton comes from the signal channel are shown in white, candidates where the lepton comes from any other $b \to u\ell\nu$ decays in cross-hatch, candidates where the lepton comes from a $b \to c\ell\nu$ or non-semileptonic B decay in yellow, and candidates where the lepton comes from a $q\bar{q}$ decay (primarily $c\bar{c}$) in red. Candidates where the lepton is fake are shown in blue; most of these are fake muons from a $u\bar{u}/d\bar{d}/s\bar{s}$ event.

The standard *BABAR* π^0 efficiency correction is also applied [30]; any candidate π^0 where both photons are identified as "bumps" in the same EMC cluster [30] is assigned a weight according to the standard correction formula, which is a function of the π^0 energy.

4.4 Background Suppression

Figure 4.10 illustrates the composition of the lepton momentum (energy) spectrum for $B^0 \to \pi^- \ell^+ \nu$ and $B^0 \to \rho^- \ell^+ \nu$ candidates, before any of the selection cuts have been applied. In both cases, the signal is completely swamped by background candidates, particularly those candidates where the lepton comes from a $b \to c\ell\nu$ decay or a $q\bar{q}$ events. The initial size of these backgrounds is hardly surprising: the overall rate of $b \to c\ell\nu$ decays is roughly 50 times that of $b \to u\ell\nu$ decays, and about 500 times the rate of any of the exclusive decays reconstructed in this analysis. Likewise, the total cross-section for $q\bar{q}$ events is over three times that of $B\bar{B}$ events; while these events tend to contain fewer true leptons, this is offset by the high muon fake rate.

Sections 4.4.1 and 4.4.2 will evaluate those cuts used to actively select for signal candidates in terms of their effectiveness in suppressing the dominant backgrounds. While many of these cuts are quite effective, they are insufficient, motivating a further set of background suppression cuts.

4.4.1 Effect of Neutrino Reconstruction

The primary purpose of the neutrino reconstruction cuts is to select signal decays with a well-reconstructed neutrino over signal decays with a poorly-reconstructed neutrino. Nonetheless, the neutrino reconstruction cuts tend to provide a moderate suppression of one or both of the dominant backgrounds, for a variety of reasons.

For instance, $b \to c\ell\nu$ decays are subject to higher track losses due to their higher average multiplicity; decays of the charm meson can lead to an additional neutrino or K_L. These effects broaden the m^2_{miss} distribution and extend it to higher values. Thus the cut on $m^2_{\mathrm{miss}}/E_{\mathrm{miss}}$ suppresses background from $b \to c\ell\nu$ decays.

$q\bar{q}$ events can have a sizable missing energy (usually oriented along the beamline) and a small missing mass squared. Such events have a greater tendency to pass the other neutrino reconstruction cuts, but are more likely to have a small missing momentum, or a missing momentum which points at the holes in the forward or backward direction. By adding a cut on the magnitude of the missing momentum in the laboratory frame to the cut already made on the polar angle:

$$|p_{miss}| > 0.6 \text{ GeV}; \quad 0.6 < \theta_{miss} < 2.9\,\mathrm{rad} \tag{4.13}$$

we keep almost two-thirds of all signal events, while reducing the remaining backgrounds by a factor of 2-3 (see Tables 4.5 - 4.8).

Events containing a $b \to c\ell\nu$ also have a distribution in θ_{miss} which peaks more strongly in the forward direction than signal decays. This is due to the fact that the $b \to c\ell\nu$ neutrino spectrum is softer, and losses in the forward region due to acceptance have a larger impact on the final direction of the missing momentum. The extremely strict cut of $\theta_{miss} > 0.6$rad, which is in fact much wider than that

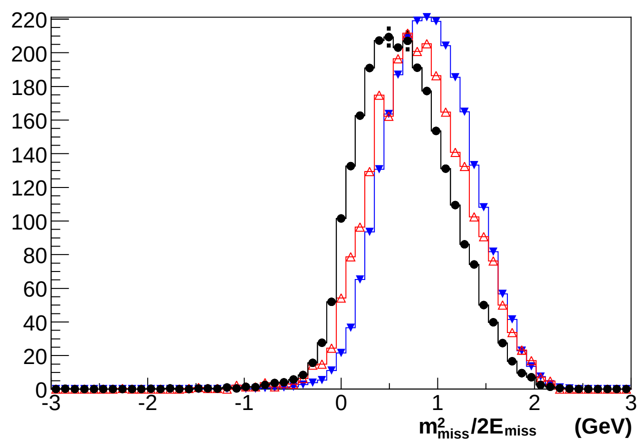
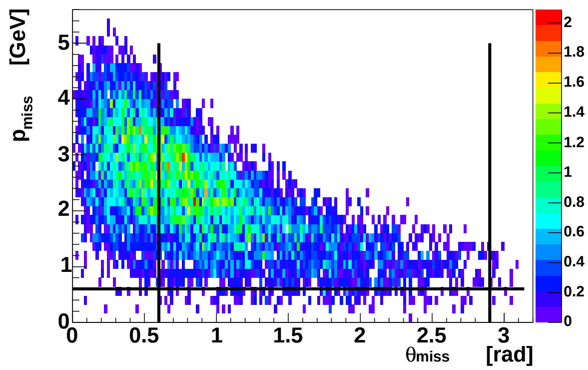


Figure 4.11: $m^2_{miss}/2E_{miss}$ for $B \to \pi\ell\nu$ events (black), continuum events (red), and events containing a $b \to c\ell\nu$ decay (blue), after all other analysis cuts have been applied.



(a)

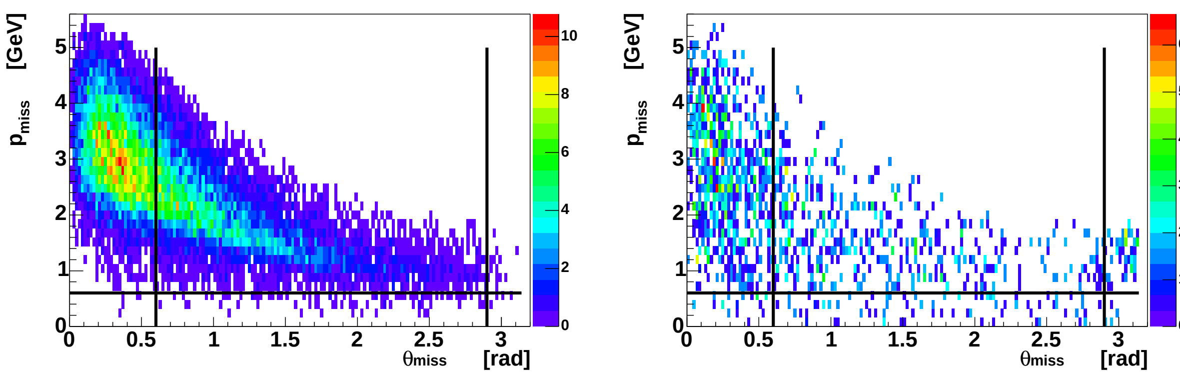


Figure 4.12: Monte Carlo simulation of the magnitude versus polar angle of the reconstructed neutrino momentum vector, for $B \to \pi\ell\nu$ selection after all other cuts are applied. Events above and between the black lines are kept. Shown are $B \to \pi\ell\nu$ signal (a), $b \to c\ell\nu$ (b) and $q\bar{q}$ (true and fake leptons) (c).

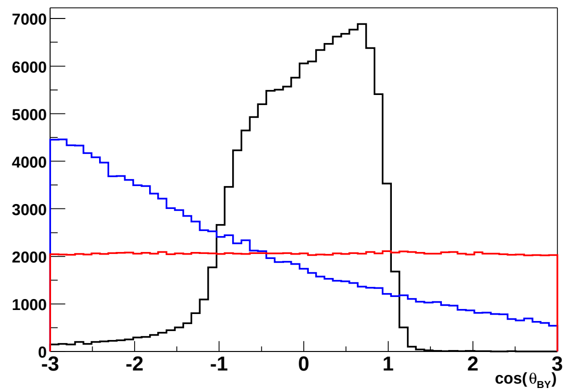


Figure 4.13: $\cos\theta_{\mathrm{BY}}$ distributions for the reconstructed $B \to \pi\ell\nu$ mode, after all other preselection cuts, but before any further selection has been applied. Shown are signal $B \to \pi\ell\nu$ decays (black), background from generic B decays (blue), and background from continuum (off-resonance data, red). All distributions have been scaled to equal area. 98% of the signal $B \to \pi\ell\nu$ distribution is captured by the histogram range, as opposed to 20% of the generic $B\bar{B}$ background and 32% of the continuum background.

required by the detector acceptance of $\theta > 0.3$rad, strongly suppresses candidates from $b \to c\ell\nu$ decays.

4.4.2 Effect of Kinematic Consistency Cuts

Figure 4.13 illustrates the effectiveness of the $\cos\theta_{\mathrm{BY}}$ variable at suppressing candidates where the lepton comes from one of the dominant background sources. As mentioned earlier, a loose cut of $|\cos\theta_{\mathrm{BY}}| < 1.5$ is used to reduce the number of background candidates stored for the analysis; this cut has a signal efficiency of about 90% and rejects about 90% of the $b \to c\ell\nu$ and $q\bar{q}$ background candidates. The further cut of $|\cos\theta_{\mathrm{BY}}| < 1.1$ has a signal efficiency of close to 97%, but rejects 25% of the remaining $b \to c\ell\nu$ decays and about 20-30% of the $q\bar{q}$ background (see Tables 4.5 - 4.8)) after all other cuts have been applied.

The cut $\Delta\theta_{min} < 0.6$rad provides some additional background suppression, as shown in Figure 4.7.

4.4.3 Kinematics of the Y System

The relative momentum distribution of the hadron and lepton, in the $\Upsilon(4S)$ frame, is dictated by the particular kinematics of a decay channel. In general, however, the correlations between p_h^* and p_l^*, in any channel, are substantially different for properly reconstructed signal candidates than for combinatoric background candidates, particularly those where the lepton originated either from a $b \to c\ell\nu$ decay or a $q\bar{q}$ event. We exploit this difference to further suppress both dominant backgrounds: for the two scalar decay modes, $B \to \pi\ell\nu$ and $B^+ \to \pi^0\ell^+\nu$, we require $p_\ell^* + p_\pi^* \geq\ 2.9\,\text{GeV}$ while for $B \to \rho\ell\nu$ and $B^+ \to \rho^0\ell^+\nu$ decays we require $0.25p_\ell^* + p_\rho^* \geq\ 1.375\,\text{GeV}$. The relevant distributions and cuts are illustrated in Figures 4.14 and 4.15.

4.4.4 Event Topology

Powerful background suppression variables may be constructed by separating the observed particles into two groups: each one corresponding to explicitly detected decay products of one of the two B mesons. The first group (the Y system) consists of the charged lepton and the signal hadron; the second group consists of all tracks and neutrals in the event that do not overlap with the Y system. This second group will be referred to as the "rest of the event," abbreviated *roe*. Variables constructed in this manner are sensitive both to the underlying topology of the event, and the apparent topology induced by reconstruction errors.

Two such variables are used in this analysis. The first variable, $|\cos\theta_{thrust}|$, is the absolute value of the cosine of the angle between the thrust axis of the Y system and the thrust axis of the rest of the event. The second variable, L_2, is derived from a second-order Legendre moment of all tracks in the rest of the event:

$$L_2 = \sum_i^{\text{roe}} p_i \times \frac{1}{2}(3\cos^2\theta_i - 1), \tag{4.14}$$

where all angles are measured with respect to the thrust axis of the signal Y system.

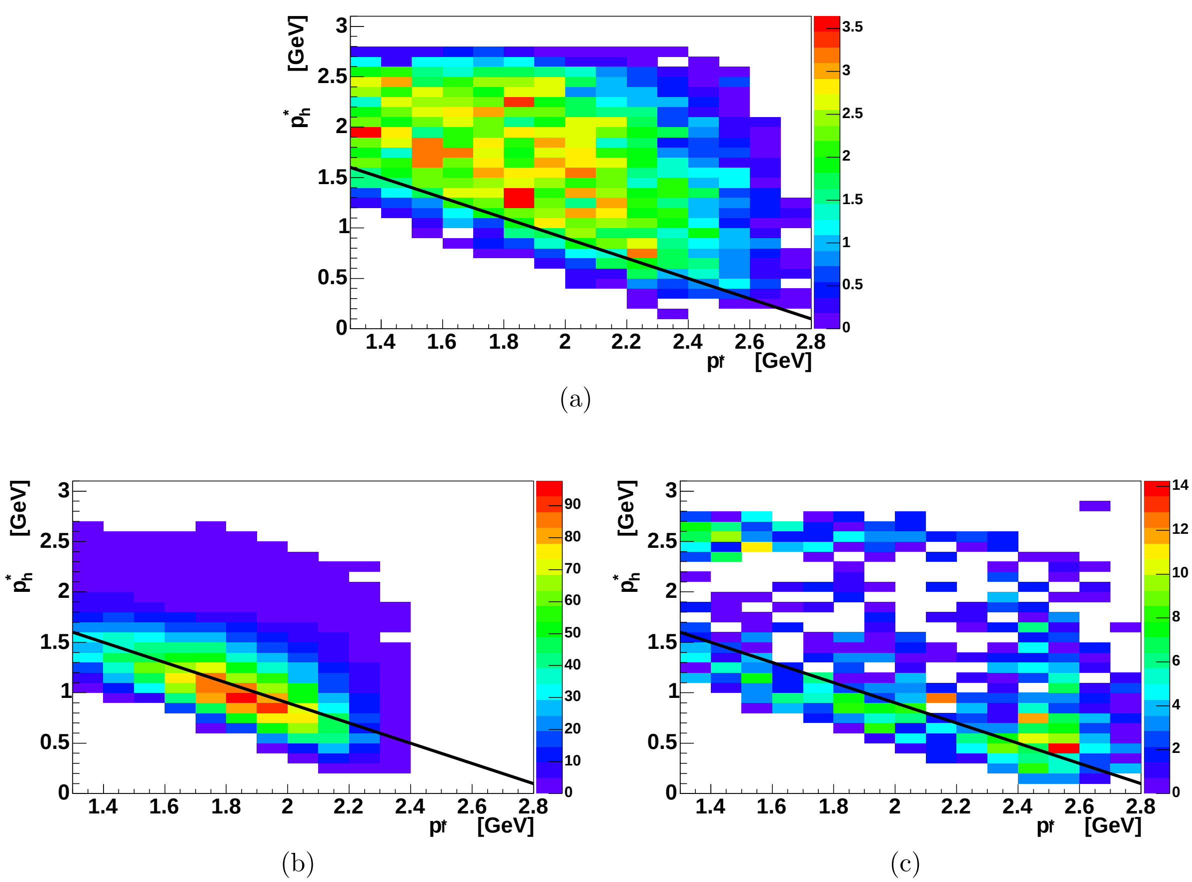


Figure 4.14: Monte Carlo simulation of the momentum (in the c.m.s) of the pion versus lepton candidates for $B \to \pi\ell\nu$ selection; all cuts but those on the lepton and hadron momentum have been applied. Candidates above the black line are kept. Shown are $B \to \pi\ell\nu$ signal (a), $b \to c\ell\nu$ background (b), and $q\bar{q}$ background (c).

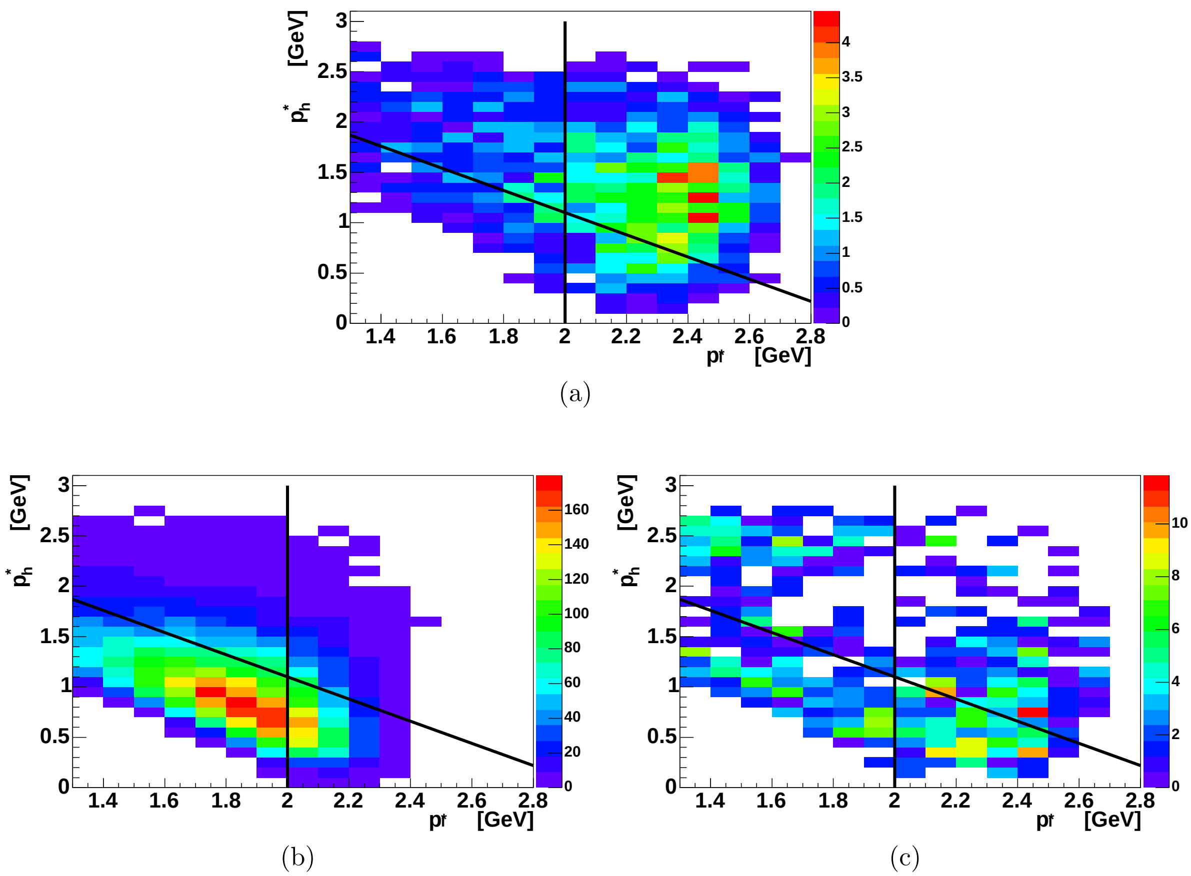
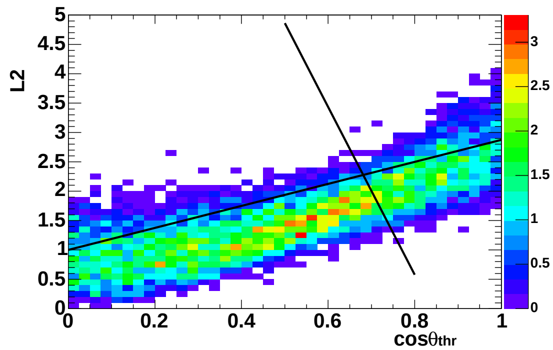
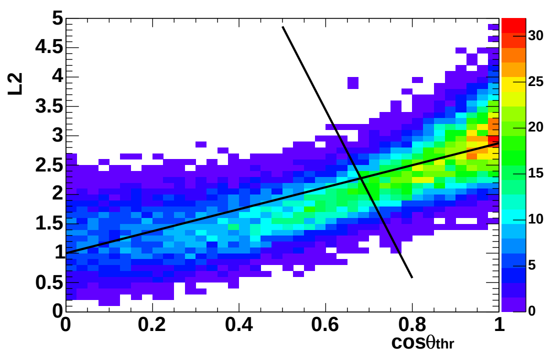


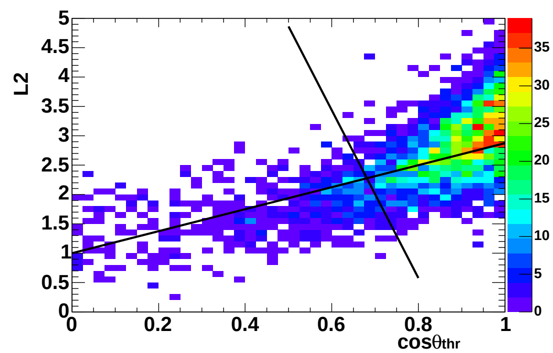
Figure 4.15: Monte Carlo simulation of the momentum (in the c.m.s) of the pion versus lepton candidates for $B \to \pi\ell\nu$ selection; all cuts but those on the lepton and hadron momentum have been applied. Candidates above the black line are kept. Shown are $B \to \rho\ell\nu$ signal (a), $b \to c\ell\nu$ background (b), and $q\bar{q}$ background (c).



(a)



(b)



(c)

Figure 4.16: Monte Carlo simulation of the second Legendre moment $L2$ vs. $\cos\theta_{thrust}$, for $B \rightarrow \pi\ell\nu$ candidates after all other cuts are applied. Candidates below the intersection of the two black lines are kept. Shown are MC candidates where the lepton is truth-matched to (a) a signal decay (b) a $b \rightarrow c\ell\nu$ decay (c) a real or fake lepton from a $q\bar{q}$ event.

This moment can be divided into a linear combination of two components:

$$L_0 = \sum_i^{\text{roe}} p_i \quad L_2 = \sum_i^{\text{roe}} p_i \cos^2\theta_i \tag{4.15}$$

L_0 is essentially the energy of the "other B" in the event and, as a consequence, is too strongly correlated with ΔE to be useful; L_2, a higher angular moment, is relatively uncorrelated with ΔE and may be used.

Both L2 and $\cos\theta_{thrust}$are particularly effective in suppressing $q\bar{q}$ background, since they are sensitive to the topological differences between $B\bar{B}$ events at the $\Upsilon(4S)$ resonance. The $\cos\theta_{thrust}$ distribution, for instance, tends to peak near a value of one for candidates from $q\bar{q}$ events, while being relatively flat for well-reconstructed signal candidates.

Nonetheless, this pair of variables is nearly as effective at suppressing combinatoric background from $b \to c\ell\nu$ decays. In general, the lepton dominates the thrust axis of the $b \to c\ell\nu$ decay; the thrust axis of the hadron daughter particles is naturally anti-correlated to the axis of the lepton, and thus to the thrust axis of the Y candidate. Since some of these daughter particles are dropped from the reconstructed candidate, they become included in the "other B", introducing a correlation between the two thrust axes, and $|\cos\theta_{thrust}|$ tends to peak at values closer to one.

Figure 4.16 shows the distributions of signal and background events as a function of L_2 and $\cos\theta_{thrust}$ for signal and background samples in the $B \to \pi\ell\nu$ mode. By choosing a two-dimensional cut

$$L_2 + 14.28\ \cos\theta_{thrust} \leq 14.28 \qquad L_2 - 1.875\ \cos\theta_{thrust} \leq 1.0 \tag{4.16}$$

we exploit the strong correlation between these two variables. We retain on the order of 75% - 78% of the signal events, while reducing the remaining $q\bar{q}$ events by about a factor of 10 and the remaining $b \to c\ell\nu$ events by about a factor of 3(see Tables 4.5 - 4.8).

4.4.5 Further Continuum Suppression

Backgrounds from lepton-pair production and Bhabha scattering are not included in the MC simulation; as a result, it is particularly important that we eliminate them from the final sample. Since the majority of these events have a low track and neutral multiplicity compared to actual B decays, a large portion of this background is easily suppressed by a cut requiring that any event contain at least four charged particles in addition to the candidate lepton, or three charged particles and at least two photons. While more stringent requirements might suppress these pure QED backgrounds to a greater degree, they would also impact the signal, since events containing a signal decay have, on average, a lower track and shower multiplicity than the average $B\bar{B}$ event. This is particularly true for $B \to \pi\ell\nu$ since the signal B contains only the lepton and either a single track or pair of photons from the signal pion.

A tighter cut on the Fox-Wolfram moment, $R2 < 0.4$, further suppresses pure QED backgrounds, and eliminates on the order of 20% of the remaining candidates from $q\bar{q}$ events while having a negligible impact on the signal efficiency (Tables 4.5 - 4.8).

Candidates from *showering Bhabhas* provide a background in the electron channel that is not included in our continuum simulation, and is difficult to eliminate with the usual Fox-Wolfram and track and shower multiplicity cuts. In these events, the positron in the backwards direction is captured as the signal lepton; the forward-going electron strikes the SVT or magnets and showers. The spray of particles splashes back across the forward part of the detector, resulting in a reconstructed event which no longer has either the characteristic topology or low multiplicity of an ordinary Bhabha. These candidates and events do, however, have a number of identifying characteristics: the candidate lepton is a positron pointing in the forward direction, and tends to have a fairly high momentum in the $\Upsilon(4S)$ frame. We veto any candidate with a positively-charged lepton where $\theta_l < 1.9\,\mathrm{rad}$ and $p_l^* > 2.2\,\mathrm{GeV}$; this has a negligible impact on the signal and all other backgrounds. While this cut does suppress a portion of the QED contribution to the continuum background, there are other QED sources that are not eliminated.

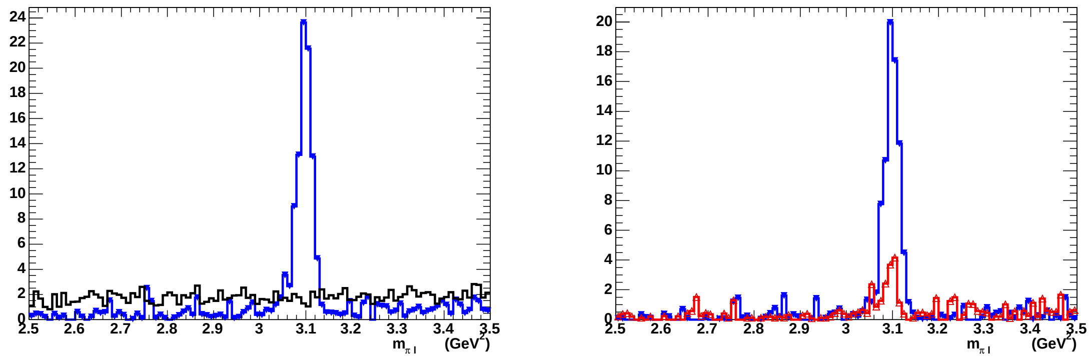
4.4.6 Suppressing Secondary Leptons

The majority of background candidates from $B\bar{B}$ events usually contain a primary lepton from a semileptonic B decay. A smaller fraction of background candidates contain a secondary lepton. Most of these sources (such a cascade semileptonic decay of a charm meson) produce leptons with a relatively soft momentum spectrum; the resulting backgrounds are suppressed by the requirement that the lepton have a momentum of at least $1.3\,\mathrm{GeV}(2.0\,\mathrm{GeV})$ in the $\Upsilon(4S)$ frame. Leptons from the decay $B \to J/\psi\, X; J/\psi \to \ell^+\ell^-$, on the other hand, have a hard spectrum; moreover, J/ψ will easily fake a $\pi - \ell$ candidate when the lepton veto fails for one of the two daughter tracks (about 10% of the time for electrons, 20% or more of the time for muons). If the J/ψ originates from $B \to J/\psi\, K_L$ decays or a B decay in which most of the remaining decay products are lost, the resulting missing momentum and energy may easily fake a neutrino, yielding a background candidate which legitimately has the energy and mass expected of a signal B. Fortunately, this background is easily identified by taking the effective mass of the lepton and the oppositely charged pion from the hadron candidate (assuming the lepton mass for both tracks). Figure 4.17 shows the sharply peaked background and the flat signal distribution for $B \to \pi\ell\nu$ candidates. Requiring $|M_{\ell\ell} - m_{J/\psi}| > 25\,\mathrm{MeV}$ removes most of this background.

4.5 Additional Variable Definitions

Kinematic consistency of the candidate decay with a B decay is checked using two variables, the beam energy substituted mass, m_{ES}, and the difference between the reconstructed and expected energy of the B candidate, ΔE. These variables will also serve as the basis of the final fit, discussed in Section 6. In the laboratory frame these variables are defined as [31]

$$\Delta E = \frac{P_B \cdot P_{beams} - s/2}{\sqrt{s}}, \tag{4.17}$$



(a) (b)

Figure 4.17: (a) Monte Carlo simulation of the $\pi\ell$ mass for $B \to \pi\ell\nu$ candidates, calculated using a lepton hypothesis for both tracks. Signal candidates are shown in black, candidates from non-semileptonic B decays are shown in blue. (b) The mass distribution for candidates from non-semileptonic B decays only, shown separately for electrons (red), and muons (blue).

where s refers total energy of the e^+e^- beam particles in the c.m. frame, P_B and P_{beams} denote the four-momenta of the B meson and the colliding beam particles.

$$m_{ES} = \sqrt{(s/2 + \vec{p}_B \cdot \vec{p}_{beams})^2/E_{beams} - \vec{p}_B^2} \tag{4.18}$$

The last quantity of interest is the q^2 of the underlying decay. Assuming a perfectly reconstructed signal candidate, there a number of possible ways to approximate this quantity. One approach uses only the available information from the candidate lepton and the reconstructed neutrino ($q^2_{\rm raw}$); the resolution of this variable may be improved by also using information from the candidate hadron ($q^2_{\rm corr}$). Another technique uses primarily information from the candidate hadron, along with assumptions about the momentum of the parent B, to calculate q^2 ($q^2_{\rm had}$).

Since q^2 for a semileptonic decay is simply the invariant mass of the virtual W, the simplest possible way of calculating q^2 is to take the invariant mass of the four-vector sum of the reconstructed lepton and neutrino:

$${q^2}_{\rm raw} = |(\vec{p^*}_l, E_l) + (\vec{p}_{\rm miss}, |\vec{p}_{\rm miss}|)|^2 \tag{4.19}$$

This raw q^2 has the virtue that it is the only method of q^2 calculation for which the q^2 is accurate regardless of whether the candidate hadron is perfectly reconstructed; however, the resolution of $q^2_{\mathrm{raw}}$is dominated by the resolution of the reconstructed neutrino, and is therefore severely effected by the large tails in the neutrino momentum resolutions.

The resolution of q^2_{raw} may be improved by using the information associated with the candidate hadron to compensate for the poor neutrino resolution. The magnitude of $p_{\vec{\mathrm{miss}}}$is allowed to vary by a factor of α, such that ΔE of the candidate is forced to zero.

It is also possible to infer the q^2 of the event by using

$$(p_l + p_\nu)^2 = q^2 = (p_b - p_{had})^2 \approx m_B^2 + m_{\mathrm{had}}^2 - 2m_B E_h^* = q^2_{\mathrm{had}} \tag{4.20}$$

where the B rest frame is approximated by the $\Upsilon(4S)$ frame.

4.6 Efficiencies

After applying all other selection cuts, we retain all candidates with $|\Delta E| < 0.9\,\mathrm{GeV}$ and $5.095 < m_{\mathrm{ES}} < 5.305\,\mathrm{GeV}$ ("fit region"). Unless otherwise specified, all quoted efficiencies are for this range in $\Delta E$and m_{ES}. Efficiencies are also quoted for the range $-0.15 < \Delta E < 0.25\,\mathrm{GeV}$ and $m_{\mathrm{ES}} > 5.255\,\mathrm{GeV}$ ("signal region"), which is the region of maximum signal-to-background ratio in all channels.

Table 4.4 summarizes the selection efficiencies for $B \to \pi/\rho\ell\nu$ decays; while tables 4.5 - 4.8 summarize the selection efficiencies for the individual decay modes compared to the principal background processes, as well as the "incremental efficiency" of a specific selection cut, that is, the efficiency of a cut after all other cuts have been applied.

The final signal selection efficiencies for all channels are relatively small (from around 0.5% to 1.5%). Background suppression is at the level of $10^{-5} - 10^{-6}$ for the dominant backgrounds and at the level of $10^{-3} - 10^{-4}$ for combinatoric signal and crossfeed (contributions from the pseudoscalar decay channels, $B \to \pi\ell\nu$, in the reconstructed vector decay modes, $B \to \rho\ell\nu$, and vice versa). This suppression is further enhanced by the differences in the $\Delta E$and $m_{\rm ES}$distributions for signal and background. In any decay channel, better than half the total efficiency for perfectly reconstructed signal decays lies in the region $-0.15 < \Delta E < 0.25\,\mathrm{GeV}$ and $m_{\rm ES} > 5.255\,\mathrm{GeV}$, while the efficiency for the dominant background components drops by at least an order of magnitude in this region.

It is clear from Tables 4.5 - 4.8 that the most effective background suppression cuts are the same in all reconstructed decay modes, with fairly consistent behavior across all channels. The kinematic consistency cuts provide moderate to significant background suppression with very little to moderate impact on the signal, while the cuts on $L2$ and $\cos\theta_{thrust}$ and on p_{miss} and θ_{miss} provide substantial background suppression (factors of 2-10 after all other cuts have been applied) at the price of a significant (about 40%) reduction in signal efficiency. It is the neutrino quality cut $m^2_{miss}/2E_{miss} < 0.4\,\mathrm{GeV}$, however, which, as an individual cut, reduces the signal efficiency the most—by approximately a factor of three after all other cuts have been applied.

There are some notable differences between the channels, however, particularly the fact that combinatoric signal candidates, while almost completely suppressed in the $B \to \pi\ell\nu$ channels, remain a significant contribution in the $B \to \rho\ell\nu$ channels. In addition, the cuts on p^*_l and p^*_h, while effective at suppressing background in all cases, have a more significant impact on the signal for the $B \to \rho\ell\nu$ channels.

Finally, the additional particle identification requirements for the hadron tracks are mainly effective in suppressing background from $B \to J/\psi X$ decays (as is the cut on the invariant mass of the Y candidate); in the $B^+ \to \rho^0\ell^+\nu$ channel, however, the kaon veto does provides an additional factor of two in suppressing $b \to c\ell\nu$ decays.

Table 4.4: Summary table of Monte Carlo-derived selection efficiencies, e and μ channels combined, for all reconstructed decay modes. Efficiencies in black are given for the entire sample used in the fit; efficiencies in blue are restricted to the signal region.

Mode	Absolute Efficiencies after all cuts(%)				
	signal	$B \to \pi\ell\nu$ comb.	$B \to \rho\ell\nu$ comb.	$B\overline{B}$ bkgd.	$q\bar{q}$
$B^0 \to \pi^- \ell^+ \nu$	1.426 ± 0.024	0.068 ± 0.004	0.649 ± 0.012	$(1.251 \pm 0.013)10^{-3}$	$(1.62 \pm 0.09)10^{-4}$
	0.801 ± 0.018	$(6.9 \pm 1.2)10^{-3}$	0.086 ± 0.004	$(6.90 \pm 0.33)10^{-5}$	$(1.81 \pm 0.30)10^{-5}$
$B^+ \to \pi^0 \ell^+ \nu$	0.959 ± 0.021	0.083 ± 0.004	0.375 ± 0.009	$(5.64 \pm 0.09)10^{-4}$	$(1.64 \pm 0.08)10^{-4}$
	0.586 ± 0.017	0.0165 ± 0.0017	0.0419 ± 0.0030	$(3.10 \pm 0.21)10^{-5}$	$(8.5 \pm 1.9)10^{-6}$
$B^0 \to \rho^- \ell^+ \nu$	0.382 ± 0.012	0.072 ± 0.004	0.312 ± 0.008	$(4.08 \pm 0.07)10^{-4}$	$(5.5 \pm 0.5)10^{-5}$
	0.212 ± 0.009	0.0119 ± 0.0018	0.089 ± 0.004	$(3.09 \pm 0.21)10^{-5}$	$(9.2 \pm 2.2)10^{-6}$
$B^+ \to \rho^0 \ell^+ \nu$	0.748 ± 0.019	0.0329 ± 0.0029	0.140 ± 0.005	$(2.41 \pm 0.05)10^{-4}$	$(4.9 \pm 0.5)10^{-5}$
	0.429 ± 0.015	$(2.6 \pm 0.7)10^{-3}$	0.0308 ± 0.0025	$(2.16 \pm 0.17)10^{-5}$	$(7.7 \pm 2.0)10^{-6}$

Table 4.5: Table of Monte Carlo-derived selection efficiencies, e and μ channels combined, for $B^0 \to \pi^- \ell^+ \nu$.

$B^0 \to \pi^- \ell^+ \nu$				
	Absolute Efficiencies after all cuts (%)			
	signal	$b \to c\ell\nu$	$b \to other$	$q\bar{q}$
Fit Region	1.426 ± 0.024	$(1.108 \pm 0.012)10^{-3}$	$(9.5 \pm 0.5)10^{-5}$	$(1.62 \pm 0.09)10^{-4}$
Signal Region	0.801 ± 0.018	$(5.34 \pm 0.27)10^{-5}$	$(9.5 \pm 1.5)10^{-6}$	$(1.81 \pm 0.30)10^{-5}$
	Efficiency of a single cut after all others are applied (%)			
	signal	$b \to c\ell\nu$	$b \to other$	$q\bar{q}$
Hadron: Lepton veto	100.0	88.00 ± 0.35	56.2 ± 2.0	97.9 ± 0.8
$\lvert m_{\pi\ell} - m_{\Psi}\rvert > 25\,\mathrm{MeV}$	98.18 ± 0.23	98.71 ± 0.13	66.1 ± 2.1	98.9 ± 0.6
Lepton fiducial cut: $0.4090 < \theta < 2.3720$	94.8 ± 0.4	94.38 ± 0.25	94.7 ± 1.2	85.5 ± 1.8
$\mathrm{Q_l} < 0 \lvert\lvert \theta_{\mathrm{l}} < 1.9\mathrm{rad} \rvert\rvert \mathrm{p_l^*} < 2.2\,\mathrm{GeV}$	99.33 ± 0.14	$99.95 \pm .03$	99.66 ± 0.32	97.7 ± 0.8
$\mathrm{p_l^*} + \mathrm{p_h^*} \geq 2.9$	92.3 ± 0.4	48.3 ± 0.4	86.3 ± 1.7	66.6 ± 2.2
$R2 < 0.4$	98.77 ± 0.19	99.50 ± 0.08	98.5 ± 0.7	94.3 ± 1.2
$\mathrm{nTrk} > 4 \lvert \mathrm{nTrk} = 4 \& \mathrm{nPhot} \geq 2$	99.85 ± 0.06	$99.92 \pm .03$	99.78 ± 0.26	100.0
$\mathrm{L2} + 14.28\cos(\theta_{\mathrm{thr}}) \leq 12.00;$ $\mathrm{L2} - 1.875\cos(\theta_{\mathrm{thr}}) \leq 1$	63.7 ± 0.6	33.03 ± 0.31	49.5 ± 1.9	11.4 ± 0.6
$0.7\,\mathrm{GeV} < \mathrm{p_{miss}}; 0.6 < \theta_{\mathrm{miss}} < 2.9$	61.1 ± 0.6	47.1 ± 0.4	40.0 ± 1.8	35.9 ± 1.6
$\lvert\Delta \mathrm{Q}\rvert \leq 1$	93.4 ± 0.4	89.01 ± 0.33	90.9 ± 1.5	91.6 ± 1.5
$\lvert\cos(\Theta_{\mathrm{BY}})\rvert \leq 1.1$	96.86 ± 0.29	73.6 ± 0.4	83.0 ± 1.9	76.6 ± 2.1
$\Delta(\theta_\nu) < 0.6\,\mathrm{rad}$	98.45 ± 0.21	89.91 ± 0.32	74.8 ± 2.1	74.0 ± 2.1
$\lvert \mathrm{m_{miss}^2}/(2 * \mathrm{E_{miss}})\rvert < 0.40$	27.1 ± 0.4	10.10 ± 0.11	25.2 ± 1.1	17.0 ± 0.9

Table 4.6: Table of Monte Carlo-derived selection efficiencies, e and μ channels combined, for $B^+ \to \pi^0 \ell^+ \nu$.

	$B^+ \to \pi^0 \ell^+ \nu$			
	Absolute Efficiencies after all cuts (%)			
	signal	$b \to c\ell\nu$	$b \to other$	$q\bar{q}$
Fit Region	0.959 ± 0.021	$(5.02 \pm 0.08)10^{-4}$	$(1.71 \pm 0.19)10^{-5}$	$(1.64 \pm 0.08)10^{-4}$
Signal Region	0.586 ± 0.017	$(2.67 \pm 0.19)10^{-5}$	$(7.6 \pm 2.5)10^{-7}$	$(8.5 \pm 1.9)10^{-6}$
	Efficiency of a single cut after all others are applied (%)			
	signal	$b \to c\ell\nu$	$b \to other$	$q\bar{q}$
Lepton fiducial cut: $0.4090 < \theta < 2.3720$	95.5 ± 0.5	94.2 ± 0.4	94.4 ± 2.5	79.9 ± 2.0
$\mathrm{Q_l} < 0 \| \|\theta_\mathrm{l} < 1.9\mathrm{rad}\| \|\mathrm{p_l^*} < 2.2\,\mathrm{GeV}$	99.55 ± 0.15	99.87 ± 0.06	99.7 ± 0.6	98.3 ± 0.7
$\mathrm{p_l^*} + \mathrm{p_h^*} \geq 2.9$	92.8 ± 0.6	38.1 ± 0.5	$49. \pm 4.$	55.1 ± 2.1
$R2 < 0.4$	98.14 ± 0.30	99.36 ± 0.13	98.9 ± 1.2	87.8 ± 1.7
$\mathrm{nTrk} > 4 \| \mathrm{nTrk} = 4 \& \mathrm{nPhot} \geq 2$	98.75 ± 0.25	99.49 ± 0.12	99.4 ± 0.9	99.0 ± 0.6
$\mathrm{L2} + 14.28\cos(\theta_\mathrm{thr}) \leq 12.00$; $\mathrm{L2} - 1.875\cos(\theta_\mathrm{thr}) \leq 1$	64.7 ± 0.9	29.1 ± 0.4	24.1 ± 2.6	11.6 ± 0.6
$0.7\,\mathrm{GeV} < \mathrm{p_{miss}}; 0.6 < \theta_\mathrm{miss} < 2.9$	60.5 ± 0.8	42.2 ± 0.5	39.0 ± 3.5	36.9 ± 1.6
$\|\Delta \mathrm{Q}\| \leq 1$	94.9 ± 0.5	86.4 ± 0.5	89.2 ± 3.3	85.8 ± 1.8
$\|\cos(\Theta_\mathrm{BY})\| \leq 1.1$	94.6 ± 0.5	71.6 ± 0.6	$78. \pm 4.$	75.4 ± 2.1
$\Delta(\theta_\nu) < 0.6\,\mathrm{rad}$	98.58 ± 0.26	91.3 ± 0.5	$81. \pm 4.$	85.6 ± 1.8
$\|\mathrm{m^2_{miss}}/(2 * \mathrm{E_{miss}})\| < 0.40$	30.0 ± 0.6	10.92 ± 0.17	15.3 ± 1.8	$20. \pm 1.0$

Table 4.7: Table of Monte Carlo-derived selection efficiencies, e and μ channels combined, for $B^0 \to \rho^- \ell^+ \nu$.

$B^0 \to \rho^- \ell^+ \nu$				
	Absolute Efficiencies after all cuts (%)			
	signal	$b \to c\ell\nu$	$b \to other$	$q\bar{q}$
Fit Region	0.382 ± 0.012	$(3.61 \pm 0.07)10^{-4}$	$(1.27 \pm 0.19)10^{-5}$	$(5.5 \pm 0.5)10^{-5}$
Signal Region	0.212 ± 0.009	$(2.32 \pm 0.17)10^{-5}$	$(1.6 \pm 0.3)10^{-6}$	$(9.2 \pm 2.2)10^{-6}$
	Efficiency of a single cut after all others are applied (%)			
	signal	$b \to c\ell\nu$	$b \to other$	$q\bar{q}$
Hadron: Kaon veto	99.49 ± 0.22	88.9 ± 0.6	$84. \pm 5.$	92.3 ± 2.5
$\mathrm{p}_\pi^{\mathrm{max}} > 0.4\,\mathrm{GeV}; \mathrm{p}_\pi^{\mathrm{min}} > 0.2\,\mathrm{GeV}$	97.7 ± 0.5	76.5 ± 0.7	$63. \pm 7.$	$75. \pm 4.$
Hadron: Lepton veto	99.16 ± 0.28	92.9 ± 0.5	$48. \pm 5.$	97.1 ± 1.6
$\lvert m_{\pi\ell} - m_\Psi \rvert > 25\,MeV$	98.0 ± 0.4	98.64 ± 0.23	$54. \pm 5.$	97.7 ± 1.4
Lepton fiducial cut: $0.4090 < \theta < 2.3720$	96.8 ± 0.5	96.1 ± 0.4	$91. \pm 4.$	$75. \pm 4.$
$\mathrm{Q_l} < 0 \lvert\lvert \theta_\mathrm{l} < 1.9\mathrm{rad} \rvert\rvert \mathrm{p}_\mathrm{l}^* < 2.2\,\mathrm{GeV}$	99.54 ± 0.21	99.83 ± 0.08	99.2 ± 1.4	92.1 ± 2.5
$\mathrm{p}_\mathrm{l}^* > 2.0\,\mathrm{GeV}$	76.2 ± 1.1	32.2 ± 0.5	19.0 ± 2.5	52.0 ± 3.5
$0.65\,\mathrm{GeV} < \mathrm{m}_\rho < 0.85\,\mathrm{GeV}$	73.9 ± 1.1	45.8 ± 0.7	$47. \pm 5.$	48.4 ± 3.4
$1.1\mathrm{p}_\mathrm{l}^* + \mathrm{p}_\mathrm{h}^* \geq 3.3$	82.3 ± 1.1	27.1 ± 0.5	$53. \pm 5.$	51.3 ± 3.5
$R2 < 0.4$	98.6 ± 0.4	99.60 ± 0.13	100.0	97.1 ± 1.6
$\mathrm{nTrk} > 4 \lvert \mathrm{nTrk} = 4 \& \mathrm{nPhot} \geq 2$	99.85 ± 0.12	99.88 ± 0.07	98.2 ± 2.1	98.6 ± 1.1
$\mathrm{L2} + 14.28\cos(\theta_{\mathrm{thr}}) \leq 12.00$; $\mathrm{L2} - 1.875\cos(\theta_{\mathrm{thr}}) \leq 1$	64.6 ± 1.2	36.0 ± 0.6	$38. \pm 4.$	9.7 ± 0.9
$0.7\,\mathrm{GeV} < \mathrm{p_{miss}}; 0.6 < \theta_{\mathrm{miss}} < 2.9$	61.7 ± 1.2	38.6 ± 0.6	$34. \pm 4.$	29.2 ± 2.4
$\lvert \Delta \mathrm{Q} \rvert \leq 1$	93.0 ± 0.7	87.0 ± 0.6	95.7 ± 3.1	91.3 ± 2.6
$\lvert \cos(\Theta_{\mathrm{BY}}) \rvert \leq 1.1$	96.2 ± 0.6	78.0 ± 0.7	$85. \pm 5.$	80.7 ± 3.4
$\Delta(\theta_\nu) < 0.6\,\mathrm{rad}$	97.8 ± 0.4	73.0 ± 0.8	$66. \pm 6.$	$62. \pm 4.$
$\lvert \mathrm{m}_{\mathrm{miss}}^2 / (2 * \mathrm{E_{miss}}) \rvert < 0.40$	28.0 ± 0.7	16.89 ± 0.31	26.1 ± 3.2	21.0 ± 1.8

Table 4.8: Table of Monte Carlo-derived selection efficiencies, e and μ channels combined, for $B^+ \to \rho^0 \ell^+ \nu$.

$B^+ \to \rho^0 \ell^+ \nu$				
	Absolute Efficiencies after all cuts (%)			
	signal	$b \to c\ell\nu$	$b \to other$	$q\bar{q}$
Fit Region	0.748 ± 0.019	$(2.12 \pm 0.05)10^{-4}$	$(7.3 \pm 1.3)10^{-6}$	$(4.9 \pm 0.5)10^{-5}$
Signal Region	0.429 ± 0.015	$(1.73 \pm 0.15)10^{-5}$	$(1.22 \pm 0.35)10^{-6}$	$(7.7 \pm 2.0)10^{-6}$
	Efficiency of a single cut after all others are applied (%)			
	signal	$b \to c\ell\nu$	$b \to other$	$q\bar{q}$
Hadron: Kaon veto	99.0 ± 0.3	$61. \pm 1.0$	$58. \pm 6.$	84.8 ± 3.5
$\mathrm{p}_\pi^{\max} > 0.4\,\mathrm{GeV}; \mathrm{p}_\pi^{\min} > 0.2\,\mathrm{GeV}$	96.9 ± 0.4	89.9 ± 0.7	$87. \pm 6.$	89.3 ± 3.1
Hadron: Lepton veto	98.1 ± 0.4	87.6 ± 0.8	$42. \pm 6.$	91.0 ± 2.9
$\lvert m_{\pi\ell} - m_\Psi \rvert > 25\,MeV$	98.0 ± 0.4	98.7 ± 0.3	$61. \pm 7.$	100.0
Lepton fiducial cut: $0.4090 < \theta < 2.3720$	96.5 ± 0.5	96.9 ± 0.4	$89. \pm 6.$	$81. \pm 4.$
$\mathrm{Q_l} < 0 \lvert\lvert \theta_\mathrm{l} < 1.9\mathrm{rad} \rvert\rvert \mathrm{p}_\mathrm{l}^* < 2.2\,\mathrm{GeV}$	99.6 ± 0.2	99.8 ± 0.1	$94. \pm 4.$	95.6 ± 2.1
$\mathrm{p}_\mathrm{l}^* > 2.0\,\mathrm{GeV}$	$76. \pm 1.0$	29.9 ± 0.6	$30. \pm 5.$	$60. \pm 4.$
$0.65\,\mathrm{GeV} < \mathrm{m}_\rho < 0.85\,\mathrm{GeV}$	$73. \pm 1.0$	45.9 ± 0.9	$52. \pm 7.$	$51. \pm 4.$
$1.1\mathrm{p}_\mathrm{l}^* + \mathrm{p}_\mathrm{h}^* \geq 3.3$	82.2 ± 0.9	29.6 ± 0.6	$62. \pm 7.$	$55. \pm 4.$
$R2 < 0.4$	98.2 ± 0.4	99.6 ± 0.2	100.0	93.7 ± 2.4
$\mathrm{nTrk} > 4 \lvert \mathrm{nTrk} = 4 \& \mathrm{nPhot} \geq 2$	100.0	100.0	100.0	100.0
$\mathrm{L2} + 14.28\cos(\theta_\mathrm{thr}) \leq 12.00$; $\mathrm{L2} - 1.875\cos(\theta_\mathrm{thr}) \leq 1$	63.9 ± 1.0	37.2 ± 0.8	$37. \pm 5.$	11.0 ± 1.1
$0.7\,\mathrm{GeV} < \mathrm{p_{miss}}; 0.6 < \theta_\mathrm{miss} < 2.9$	60.4 ± 1.0	42.9 ± 0.8	$37. \pm 6.$	35.0 ± 2.9
$\lvert\Delta \mathrm{Q}\rvert \leq 1$	94.7 ± 0.6	88.0 ± 0.8	$83. \pm 7.$	91.5 ± 2.8
$\lvert\cos(\Theta_\mathrm{BY})\rvert \leq 1.1$	96.5 ± 0.5	$77. \pm 1.0$	$86. \pm 6.$	$79. \pm 4.$
$\Delta(\theta_\nu) < 0.6\,\mathrm{rad}$	97.8 ± 0.4	$72. \pm 1.0$	$66. \pm 7.$	$66. \pm 4.$
$\lvert \mathrm{m_{miss}^2}/(2 * \mathrm{E_{miss}}) \rvert < 0.40$	30.7 ± 0.7	16.7 ± 0.4	$28. \pm 5.$	21.4 ± 1.9

Chapter 5

Monte Carlo Validation

As discussed in the previous chapter, this analysis relies on Monte Carlo simulation to identify critical backgrounds, "optimize" signal selection cuts, and determine the absolute efficiency of the signal selection. The previous chapter also illustrated that even after the full signal selection has been applied in any channel, a substantial irreducible background remains. In order to extract a branching fraction from the data, the data distribution in the variables ΔE and m_{ES} will be fit with the various signal and background contributions, and the Monte Carlo simulation must be relied upon to provide these shapes.

The Monte Carlo predictions in this regard cannot be trusted blindly; it is necessary to confront them with data. This chapter will assess Monte Carlo predictions for the distributions of variables with critical impact on the signal selection, the relative impact of the various selection cuts in data and Monte Carlo, and the shapes of the distributions to be used in the fit. In cases where the behavior of the backgrounds is of primary interest, this can be done directly, using the reconstructed decays in each of the signal decay channels; these samples, however, are too heavily contaminated by background at all stages of the selection to provide a stringent test of the modelling of the signal component alone. For this, control samples of reconstructed $B \to D^* \ell \nu$ decays are used.

5.1 Control Samples and Validation of Signal Distributions

In the case of the signal decay channels, the expected signal yield is small, and the background contamination high. A simple comparison of the distributions or efficiencies between data and Monte Carlo cannot determine, in the case of a discrepancy, whether the is is due to the modelling of a signal or background component. Likewise, there is always the danger of cancelling discrepancies, so that a mis-modelling of the background contributions might hide a problem with the modelling of the signal contribution.

In order to gain some independent understanding of the degree to which the Monte Carlo modelling of the signal contribution can be trusted, we turn to a control sample of exclusively reconstructed $B^0 \to D^{*-}\ell^+\nu$ decays with $D^{*-} \to \overline{D^0}\pi^-$. As these decays exceed the rate for $B^+ \to \rho^0\ell^+\nu$ by a factor of 30 (taking into account the $\overline{D^0}$ branching fractions, but not their detection efficiencies), they provide a high-purity sample of exclusive semileptonic decays that may then be subjected to the usual analysis selection. Since over 50% of the $b \to c\ell\nu$ background in this analysis comes from $B^0 \to D^{*-}\ell^+\nu$ decays, the control sample also provides a useful probe of the dominant background.

The $B \to D^*\ell\nu$ control sample is a useful if imperfect stand-in for the signal decay channels. On the one hand, as a set of semileptonic B decays, exclusively reconstructed with the same techniques used for the signal channels, it is the closest possible high-statistics analog. Moreover, the distributions of a number of the primary background background suppression variables: $L2$, $\cos\theta_{thrust}$, $m^2_{miss}/2E_{miss}$, all seem to have distributions that are relatively insensitive to the specific details of the decay channel in question. Even where the distribution of a variable varies widely for the background components, as it does for the variable $\cos\theta_{\mathrm{BY}}$, it is often fairly consistent for all well-reconstructed signal candidates, regardless of the actual channel. On the other, the underlying kinematics of the $B \to D^*\ell\nu$ control sample are somewhat different from those of any of the signal decay channels; for instance, the lepton spectrum is significantly softer. In addition, the presence of the low energy spectator pion impacts our ability to reconstruct D^* candidates and thus the overall detection efficiency for this sample. Fortunately, the overall selection efficiency of this analysis for $B^0 \to D^{*-}\ell^+\nu$ is of less interest than the degree to which the relative efficiency of a particular selection cut agrees for data and simulation. As a result, if the efficiencies of these selection cuts in data and Monte Carlo agree well for the control sample, they can be reasonably taken to agree well for the various signal channels.

The resolutions of the final fit variables ΔE and m_{ES} for correctly reconstructed signal candidates, are dominated by the resolution of the reconstructed neutrino. The shapes of these distributions, while sensitive to the particulars of the decay channel in the case of background, are relatively consistent across modes for the signal component— although, as illustrated in Figure 5.13, these distributions are in general somewhat broader for the control sample. As a result, the control sample can give us a reasonable but not exact measure of how well the Monte Carlo predicts resolutions for the neutrino and final fit variables.

5.1.0.1 Control Sample Selection

The following criteria are applied to select the $D^{*-}\ell^+\nu$ final state:

- A reconstructed $\overline{D^0}$ in one of the following channels: $\overline{D^0} \to K^+\pi^-$ or $K^+\pi^-\pi^0$. All charged tracks used in the $\overline{D^0}$ reconstruction are subject to the quality requirements of Sections 4.1.1.1 and 4.1.2.

- For the $K^+\pi^-\pi^0$ mode, the K^+ is required to pass tight kaon selection criteria; for $K^+\pi^-$ decays, the kaon is loosely required to be inconsistent with a pion.

- The effective mass of the $\overline{D^0}$ decay products should be within 17 MeV (34 MeV for $K^+\pi^-\pi^0$) of the nominal D^0 mass; the vertex fit probability must exceed 0.001;

- For $D^{*-} \to \bar{D}^0\pi^-$ decays, the total (p^π) and transverse (p_t^π) momentum of the slow π^- should fulfill the conditions $p_t^\pi < 0.05\,\mathrm{GeV}$ and $p^\pi < 0.450\,\mathrm{GeV}$.

- to reduce backgrounds from continuum and combinatorics, the momentum of the D^{*-} in the $\Upsilon(4S)$ rest frame is restricted to $0.5\,\mathrm{GeV} < p^*_{D^*} < 2.5\,\mathrm{GeV}$;

- the momentum of the D^{*-} and the charged lepton should point in opposite hemispheres, $\cos(D^*, \ell) < 0$, and D^{*-} and ℓ^+ should be consistent with a common vertex, with the χ^2 probability of the vertex fit > 0.1; and

- finally, the mass $D^* - D$ difference is restricted to $0.1432\,\mathrm{MeV} < \Delta m < 0.1478\,\mathrm{MeV}$.

The $B \to D^*\ell\nu$ candidates are reconstructed only for those events which pass the standard preselection.

5.2 Critical Cut Variable Distributions

In evaluating the agreement between data and Monte Carlo for various distributions, we will focus on those distributions involved in cuts with a significant impact on either signal or background (or both). Since the focus of the current measurement is $B \to \pi\ell\nu$, we will also illustrate the agreement here for the $B \to \pi\ell\nu$ signal decay channels only, with the understanding that while the general features in the $B \to \rho\ell\nu$ channels are quite similar, modulo some additional questions regarding the absolute reconstruction efficiencies and background contributions in these two modes.

Furthermore, the absolute efficiencies for the control sample channels are not a-priori well-modelled in the Monte Carlo, and there are also questions as to the value of the $B \to D^*\ell\nu$ branching fraction that should be used to provide an absolute normalization [32] [33]. Since the control sample is most useful in testing the shapes of distributions and the relative efficiencies of individual cuts, this is not currently of particular concern; we normalize the Monte Carlo to the data at the level of the initial candidate selection, rather than using the same absolute normalization as the signal decay channel comparisons.

Finally, it is important to note that none of the distributions shown in this chapter have been fitted; the relative size of the various contributions is determined by the default branching ratios put into the *BABAR* Monte Carlo.

5.2.1 Neutrino Reconstruction

Of particular interest is the level of agreement between data and Monte Carlo simulation for those variables involved in neutrino reconstruction. Figures 5.1 and 5.4 show the comparison of data to Monte Carlo the net charge of the event (ΔQ) and for the critical neutrino quality variable, $m^2_{miss}/2E_{miss}$.

While on the whole the net charge distributions for the control sample channels appear agree reasonably well between the Monte Carlo simulation and data, the signal channels show what appears to be a broadened charge distribution. That is, on average there appear to be fewer events with net charge 0 and more with a charge of at least ± 1 in the data than in the Monte Carlo. This would be consistent with the data reconstruction efficiency for charged tracks being slightly less than that in the Monte Carlo; Figures 5.2 and 5.3 show, however, that the agreement in charged and neutral multiplicity are actually quite good after most of the selection cuts have been applied, so the explanation is not that straightforward.

The distributions in $m^2_{miss}/2E_{miss}$ tend to look reasonably good (better than 10% agreement) over much of the range; the significant ($2-10$% deviations are begin near zero and extend to negative (unphysical) values of $m^2_{miss}/2E_{miss}$), with another significant deviation at large ($> 1.3\,\text{GeV}$) values of $m^2_{miss}/2E_{miss}$. While the discrepancy at large values of $m^2_{miss}/2E_{miss}$ is not critical, the difference at zero is enhanced by the neutrino quality cuts and will have a significant impact on the efficiency-related systematic uncertainties.

The control sample channels show good ($\simeq 10\%$) agreement for both $|p_{miss}|$ and θ_{miss}, while the signal decay channels show a consistent excess in the forward region, outside and at the edges of the detector acceptance ($\theta_{miss} < 0.5$), correlated with an excess for large values of $|p_{miss}|$. As this excess is not reflected in the control sample distributions, it seems likely to reflect a missing background component rather than a common reconstruction effect. Further comparisons, first of off-resonance subtracted data to $B\overline{B}$ Monte Carlo simulation alone, and second, of off-resonance data to $q\bar{q}$ Monte Carlo simulation, not only confirm this, but show that the missing component is in the continuum. One possibility is that there is a higher level of initial-state radiation for $q\bar{q}$ events in data; the radiated photon would be lost along the direction

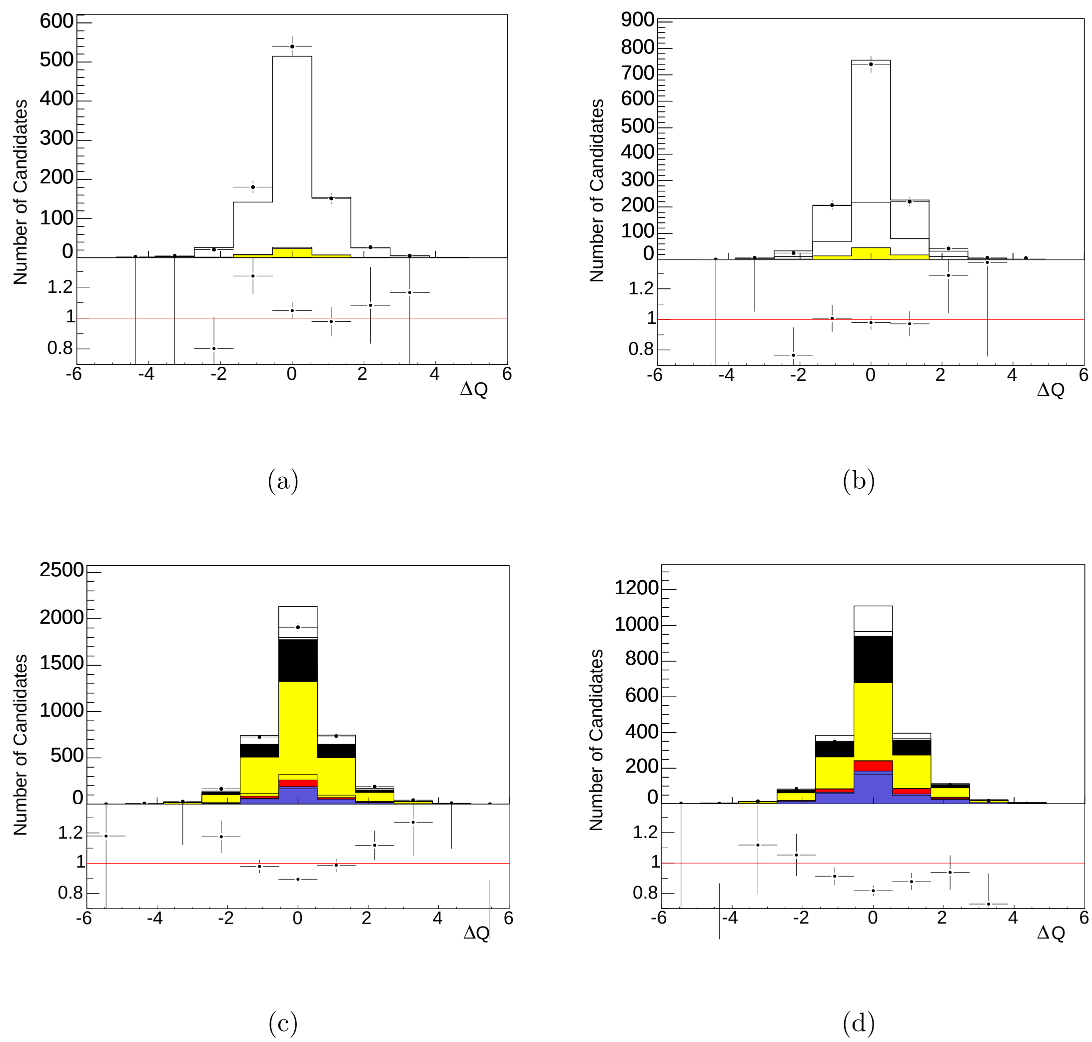
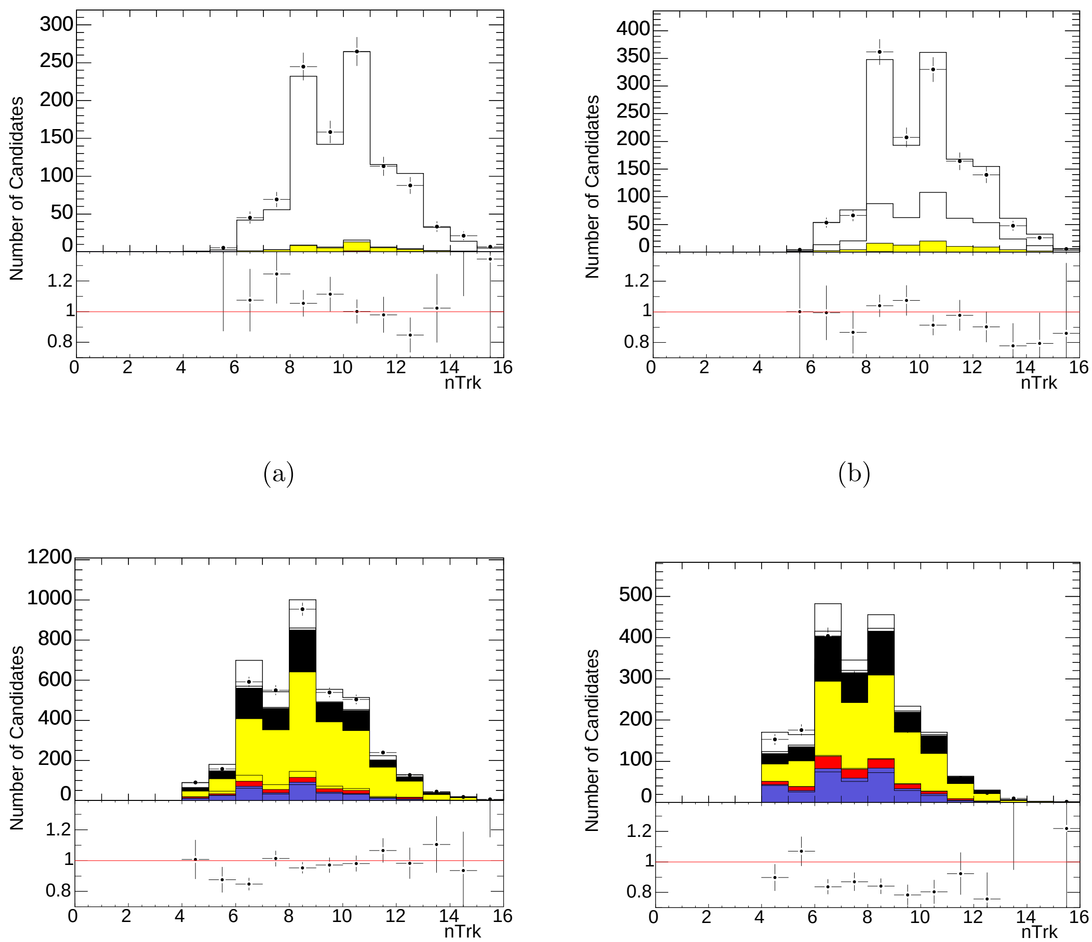


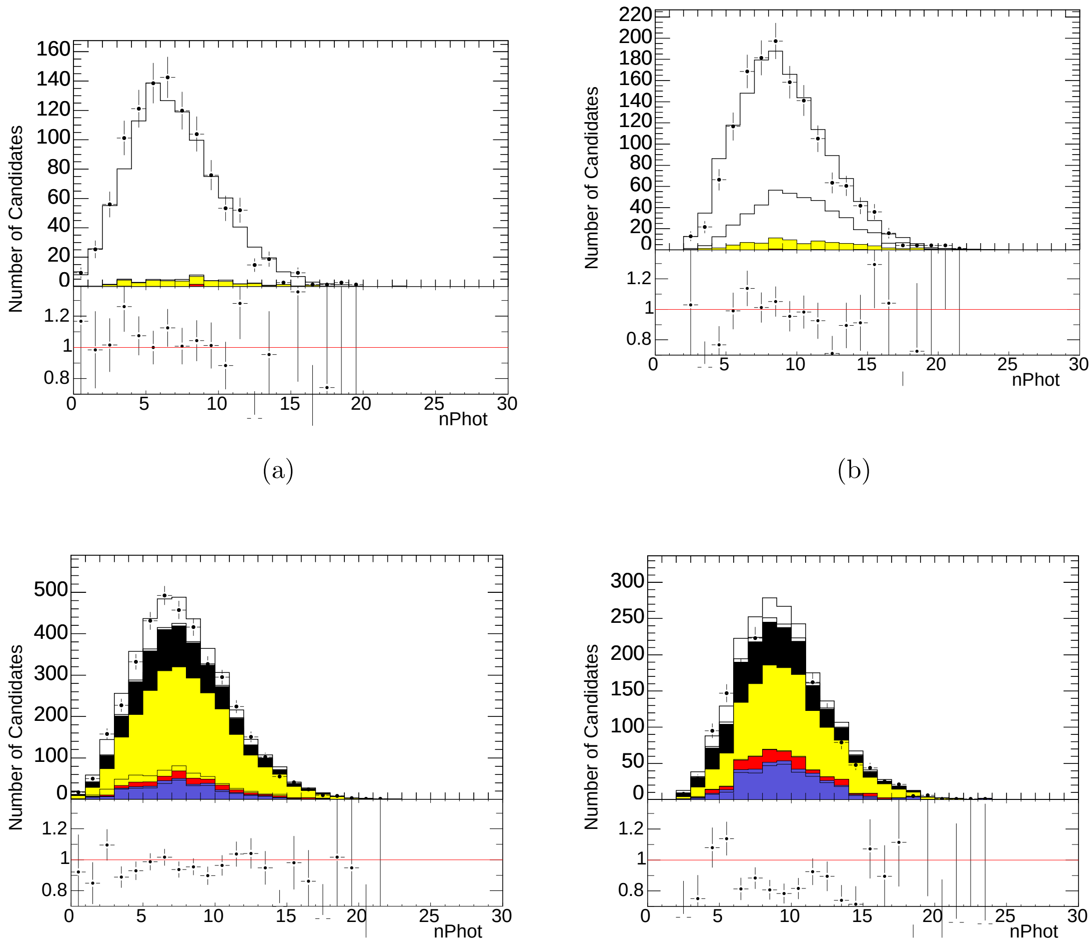
Figure 5.1: Data-Monte Carlo comparison for ΔQ, for (a) $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-$, (b) $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-\pi^0$, (c) $B^0 \to \pi^-\ell^+\nu$, (d) $B^+ \to \pi^0\ell^+\nu$ after all other selection cuts have been applied.



(a) (b)

(c) (d)

Figure 5.2: Data-Monte Carlo comparison for the number of charged tracks used in the neutrino reconstruction, for (a) $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-$, (b) $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-\pi^0$, (c) $B^0 \to \pi^-\ell^+\nu$, (d) $B^+ \to \pi^0\ell^+\nu$ after all cuts besides that on ΔQ have been applied.



(a) (b)

(c) (d)

Figure 5.3: Data-Monte Carlo comparison for the number of neutral showers used in the neutrino reconstruction, for (a) $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-$, (b) $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-\pi^0$, (c) $B^0 \to \pi^-\ell^+\nu$, (d) $B^+ \to \pi^0\ell^+\nu$ after cuts besides that on ΔQ have been applied.

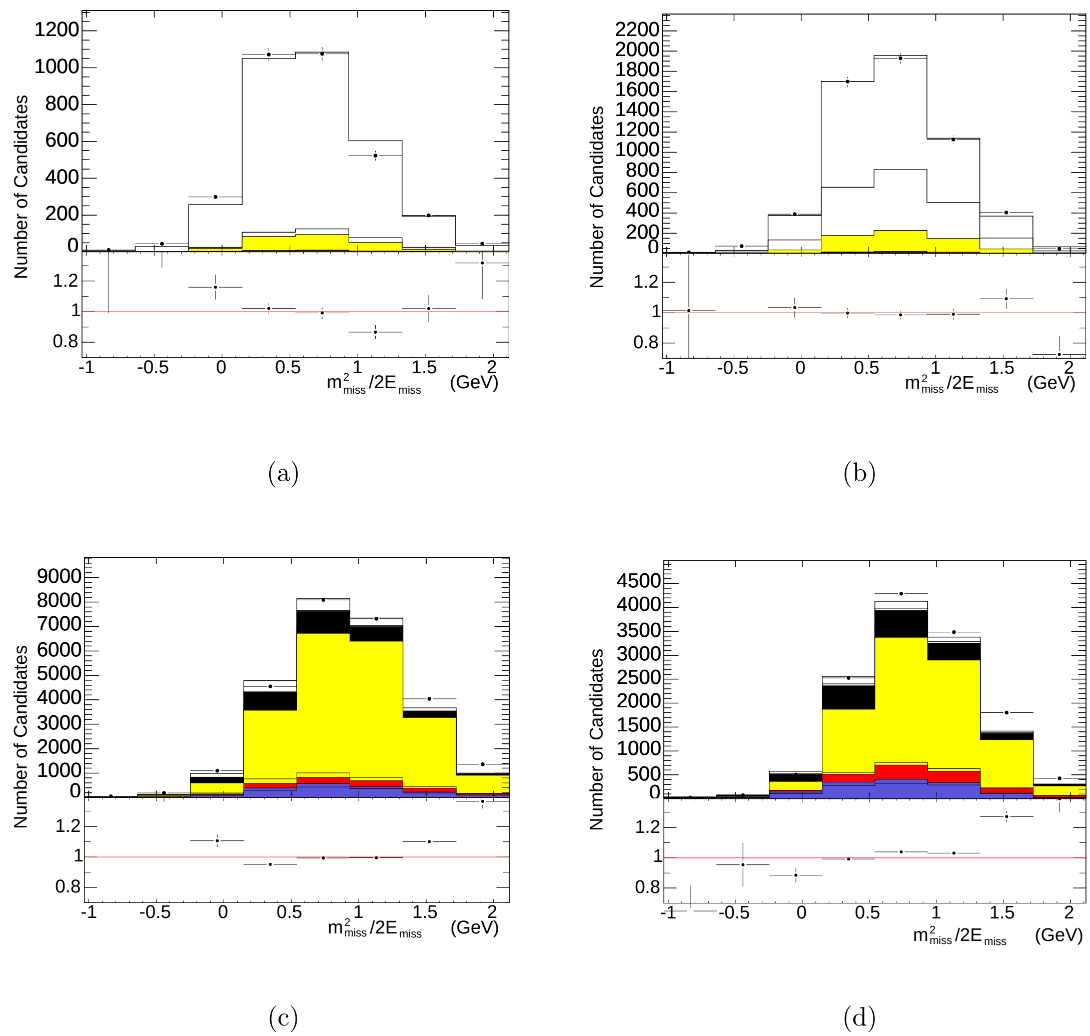


Figure 5.4: Data-Monte Carlo comparison for $m^2_{miss}/2E_{miss}$, for (a) $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-$, (b) $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-\pi^0$, (c) $B^0 \to \pi^-\ell^+\nu$, and (d) $B^+ \to \pi^0\ell^+\nu$ after all other selection cuts have been applied.

of the beam and would add to the apparent neutrino momentum.

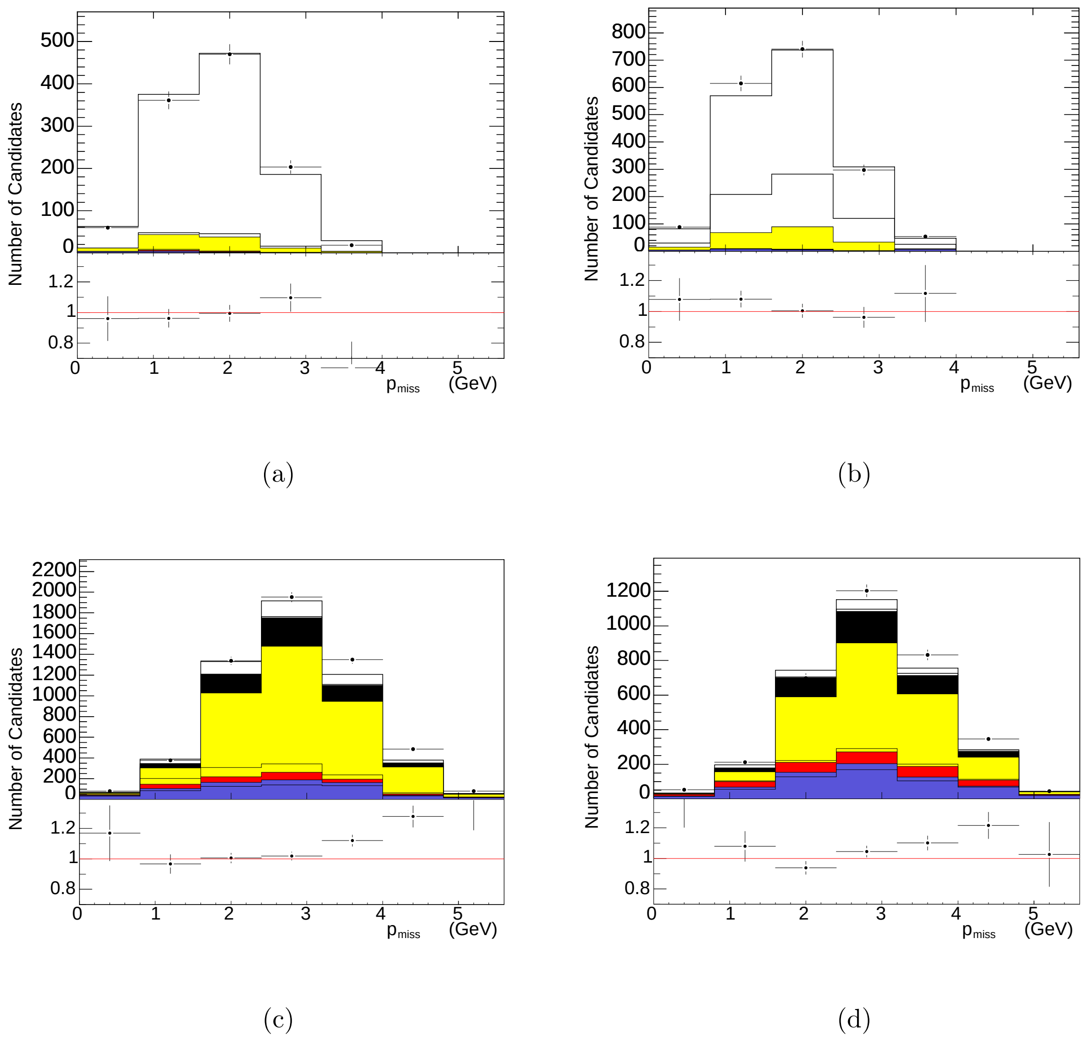
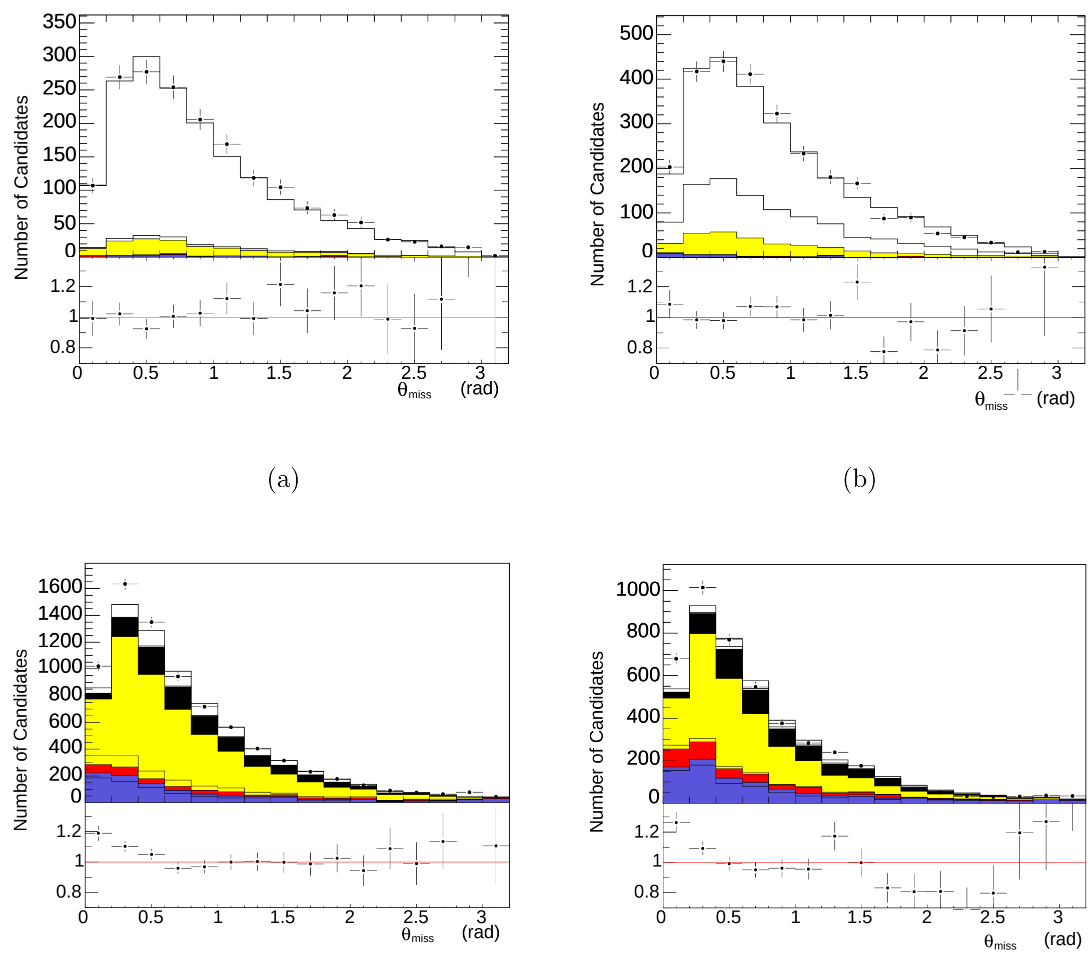


Figure 5.5: Data-Monte Carlo comparison for $|p_{miss}|$, for (a) $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-$, (b) $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-\pi^0$, (c) $B^0 \to \pi^-\ell^+\nu$, and (d) $B^+ \to \pi^0\ell^+\nu$ after all other selection cuts have been applied.

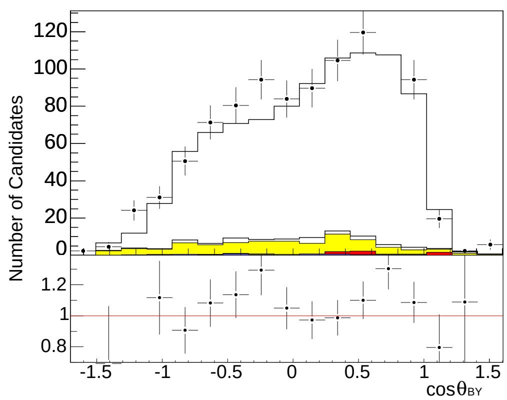
5.2.2 Kinematic Consistency

In general, the kinematic consistency variables $\cos\theta_{\rm BY}$ and $\Delta\theta_{min}$ appear reasonably well-described in the control samples, except for the expected resolution effect in

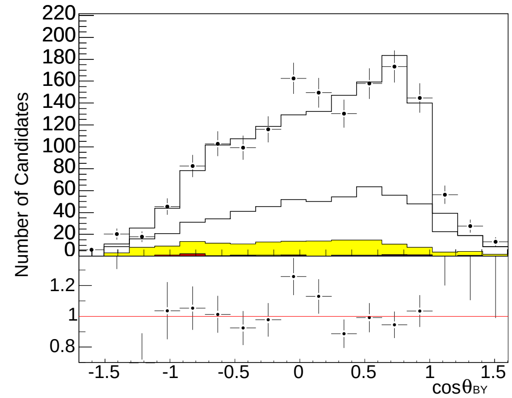


(a) (b) (c) (d)

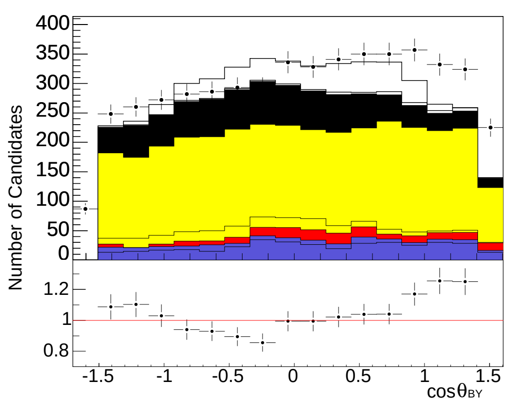
Figure 5.6: Data-Monte Carlo comparison for $|\theta_{miss}|$, for (a) $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-$, (b) $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-\pi^0$, (c) $B^0 \to \pi^-\ell^+\nu$, and (d) $B^+ \to \pi^0\ell^+\nu$ after all other selection cuts have been applied.



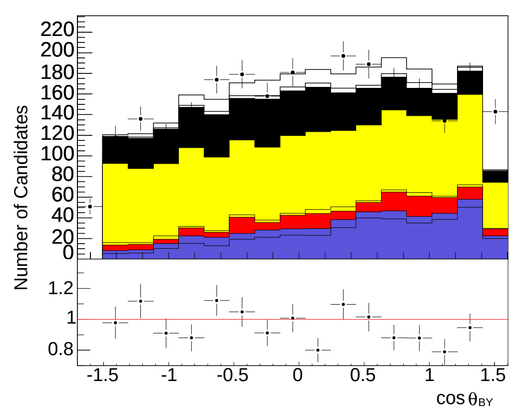
(a)



(b)



(c)



(d)

Figure 5.7: Data-Monte Carlo comparison for $\cos\theta_{\mathrm{BY}}$, for (a) $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-$, (b) $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-\pi^0$, (c) $B^0 \to \pi^-\ell^+\nu$ (d) $B^+ \to \pi^0\ell^+\nu$ after all other selection cuts have been applied.

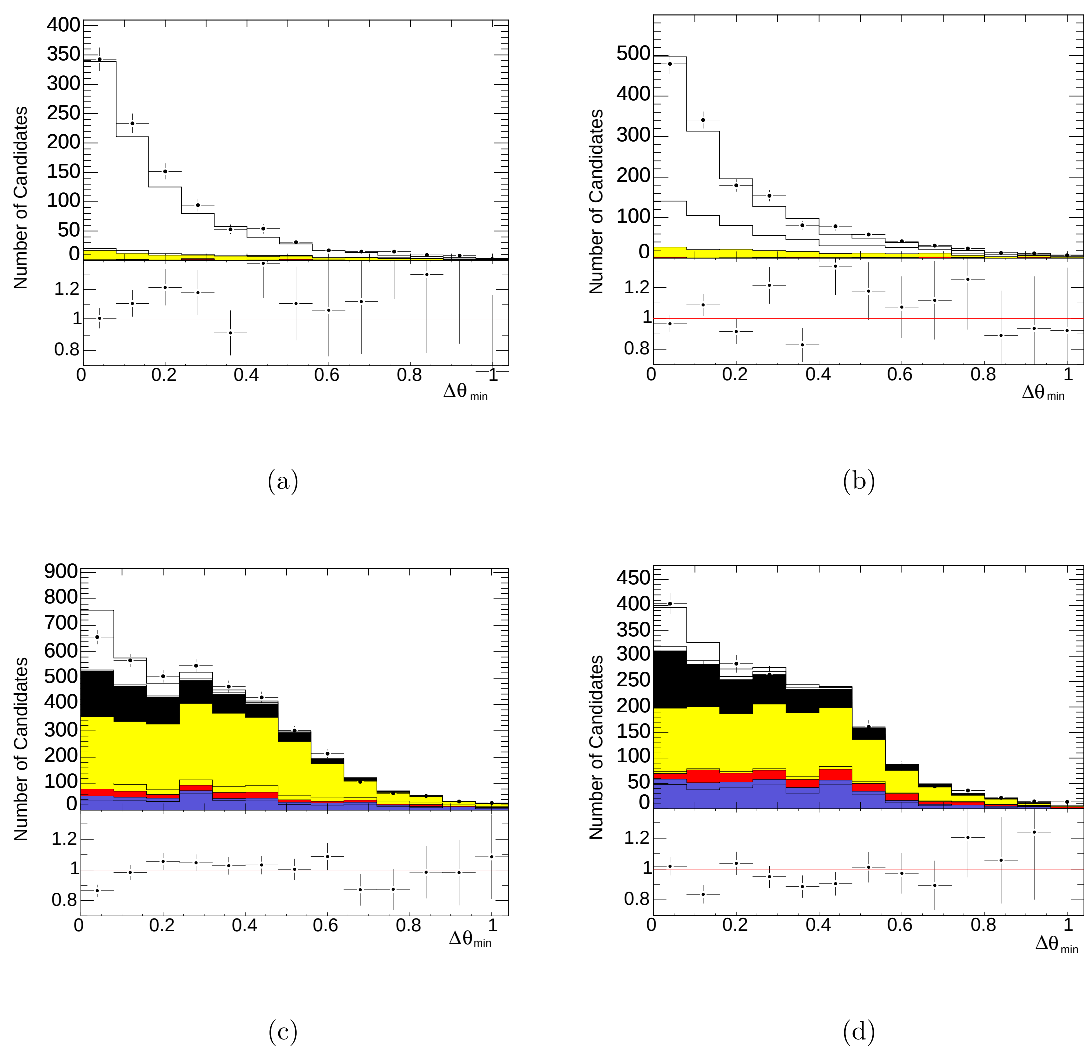


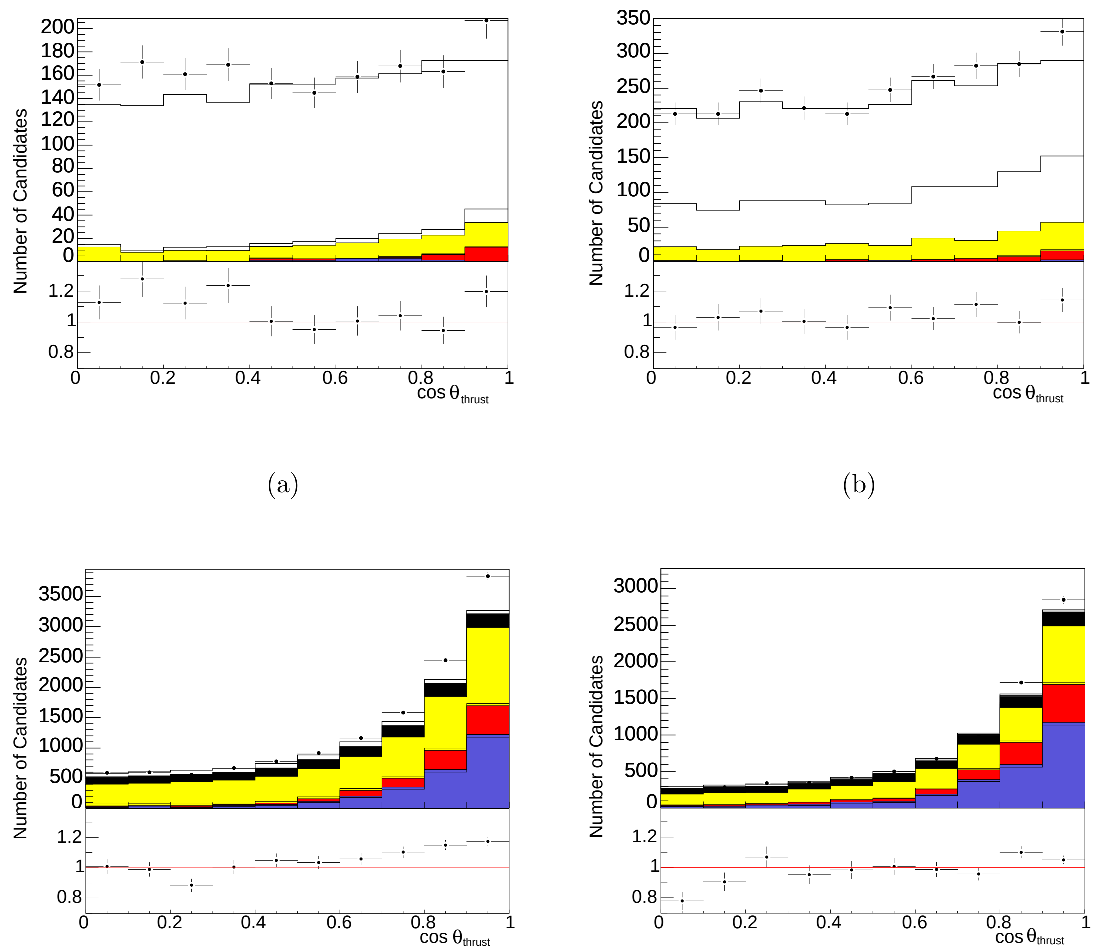
Figure 5.8: Data-Monte Carlo comparison for $\Delta\theta_{min}$, for (a) $B^0 \rightarrow D^{*-}\ell^+\nu, \bar{D}^0 \rightarrow K^+\pi^-$, (b) $B^0 \rightarrow D^{*-}\ell^+\nu, \bar{D}^0 \rightarrow K^+\pi^-\pi^0$, (c) $B^0 \rightarrow \pi^-\ell^+\nu$, (d) $B^+ \rightarrow \pi^0\ell^+\nu$ after all other selection cuts have been applied.

$\Delta\theta_{min}$. There are some significant anomalies in the $\cos\theta_{\rm BY}$ distribution for $B^0 \rightarrow \pi^-\ell^+\nu$, which may or may not be correlated with other discrepancies. Since a slightly different calculation of $\cos\theta_{\rm BY}$ was used for the preselection than in the final analysis, and since the two calculations perform identically on Monte Carlo but differently on data, edge effects in this distribution may also be an artifact of the $|\cos\theta_{\rm BY}| < 1.5$ preselection cut.

5.2.3 Background Suppression: Topology and Kinematics

Figures 5.9 and 5.10 show the comparison between data and Monte Carlo for the background suppression variables L2 and $\cos\theta_{thrust}$, for both the control sample channels and the $B^0 \rightarrow \pi^-\ell^+\nu$ and $B^+ \rightarrow \pi^0\ell^+\nu$ signal channels. In general, the shape agreement for both these variables in the control samples is good to about $10-15\%$, although there is a more significant anomaly at low values of $L2$ and $\cos\theta_{thrust}$ in the $B^0 \rightarrow D^{*-}\ell^+\nu, \bar{D}^0 \rightarrow K^+\pi^-$ channel that is not yet understood. On the other hand, the $B \rightarrow \pi\ell\nu$ decay channel in particular shows a significant ($\simeq 20\%$) excess at large values of both $L2$ and $\cos\theta_{thrust}$(a similar excess appears for the $B \rightarrow \rho\ell\nu$ decay channels which are not shown). Again, this discrepancy appears consistent with a missing component, and can be localized to the continuum. At this time, identifying the origin of this missing component has proved difficult. We know that QED background processes have not been included in the simulation, and at least one background component resulting from a QED process, the showering bhabhas, contributes at similar values of $L2$, $\cos\theta_{thrust}$, and $p_{\rm miss}$. However, whatever processes contribute to the remaining discrepancy, these candidates lack the clear topological, kinematic, and charge signature of the showering bhabhas, and they are thus difficult to either identify or suppress. At this time, it cannot be conclusively determined whether the problems come from a single QED process, a set of QED processes, or a problem in the $q\bar{q}$ simulation.

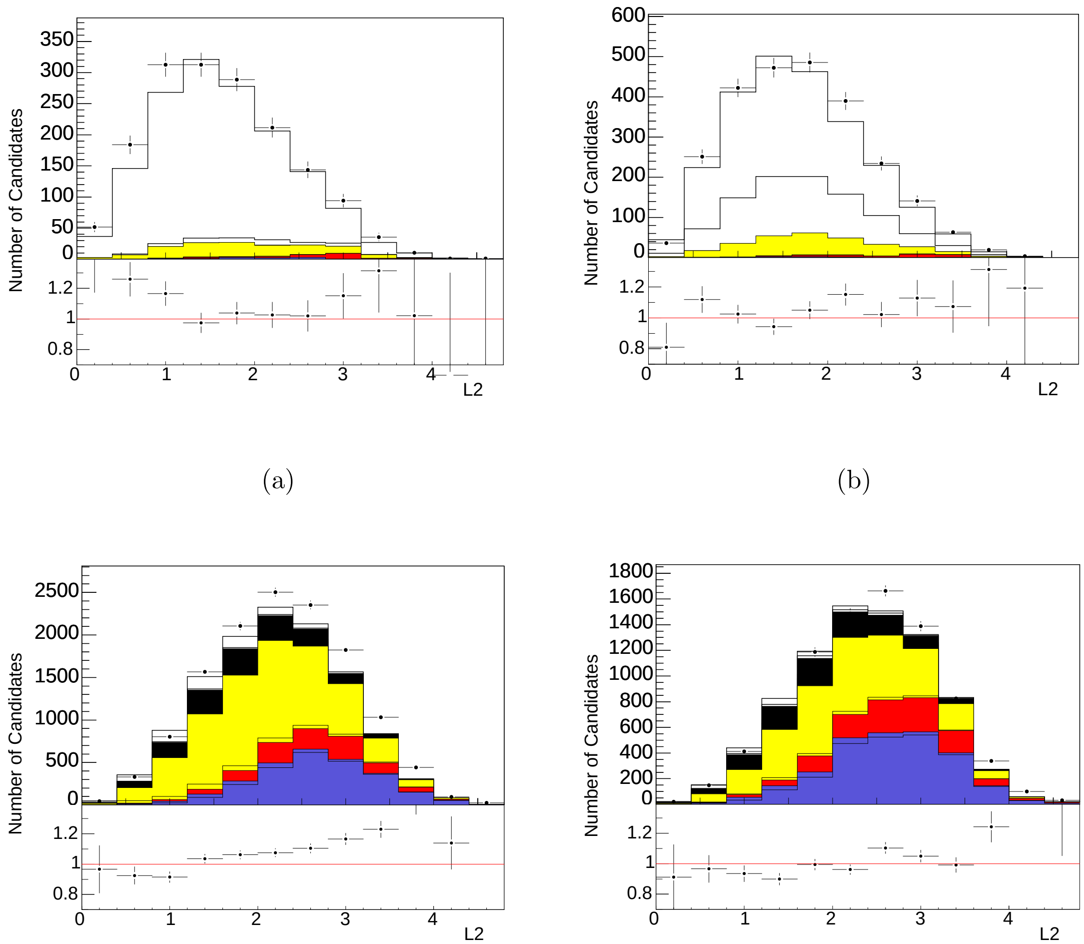
Figures 5.11 and 5.12 show the data/MC comparison for the lepton and hadron momentum in the $\Upsilon(4S)$frame. The general trend for the control sample spectra is to suggest that the data prefer a slightly harder lepton spectrum and a slightly softer



(a) (b)

(c) (d)

Figure 5.9: Data-Monte Carlo comparison for $\cos\theta_{thrust}$, for (a) $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-$, (b) $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-\pi^0$, (c) $B^0 \to \pi^-\ell^+\nu$ (d) $B^+ \to \pi^0\ell^+\nu$ after all other selection cuts but those on L2 and $\cos\theta_{thrust}$ have been applied.



(a) (b)

(c) (d)

Figure 5.10: Data-Monte Carlo comparison for L2, for (a) $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-$, (b) $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-\pi^0$, (c) $B^0 \to \pi^-\ell^+\nu$ (d) $B^+ \to \pi^0\ell^+\nu$ after all other selection cuts but those on L2 and $\cos\theta_{thrust}$have been applied.

hadron spectrum. In the case of the signal channels, the agreement for the lepton spectrum after all cuts have been applied looks quite good; the hadron momentum shows as excess at momenta below 0.7 GeV, which is more significant in the $B^0 \to \pi^- \ell^+ \nu$ channel than the $B^+ \to \pi^0 \ell^+ \nu$ channel (20% vs. 10%).

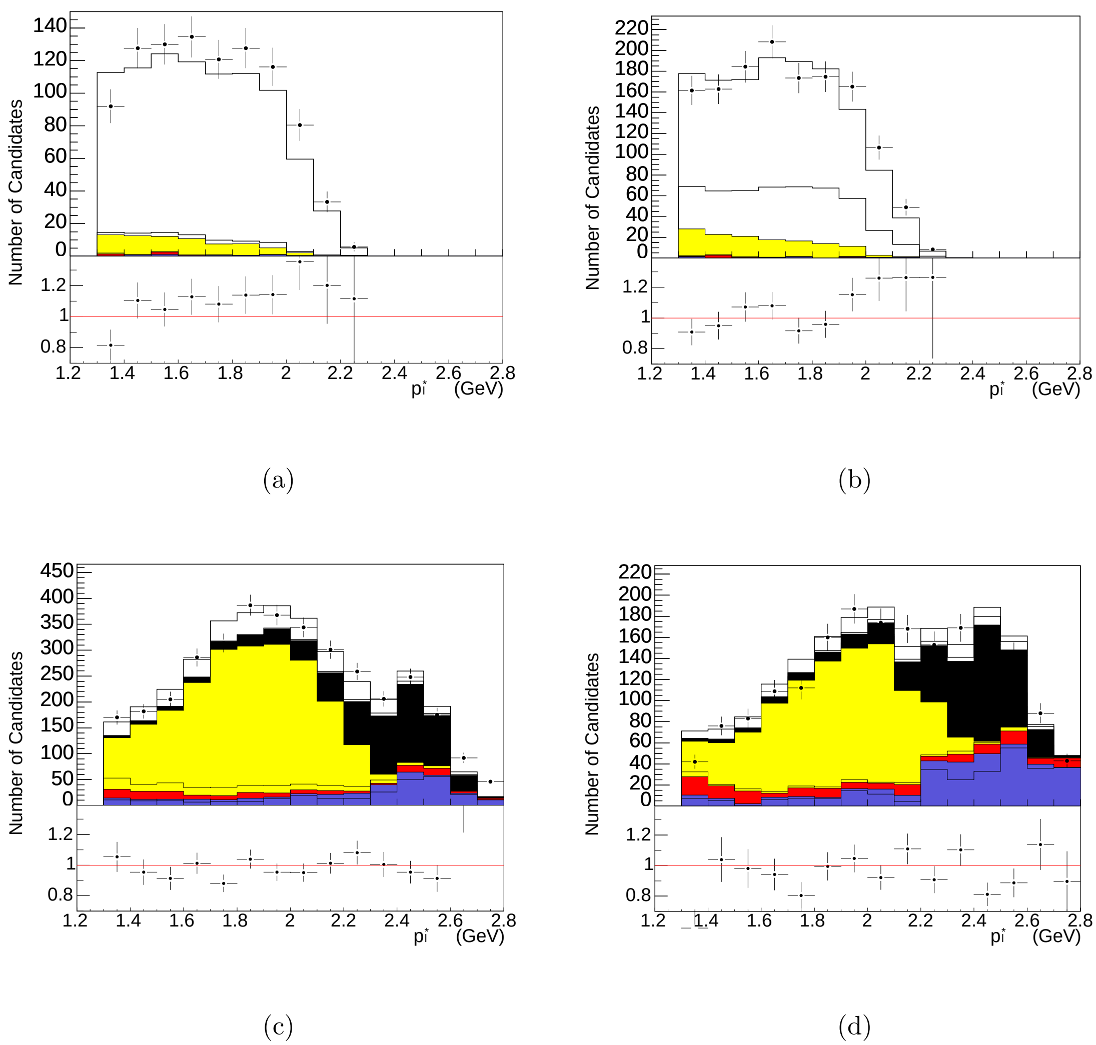


Figure 5.11: Data-Monte Carlo comparison for $|p_l^*|$, for (a) $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-$, (b) $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-\pi^0$, (c) $B^0 \to \pi^-\ell^+\nu$, (d) $B^+ \to \pi^0\ell^+\nu$ after all other selection cuts have been applied.

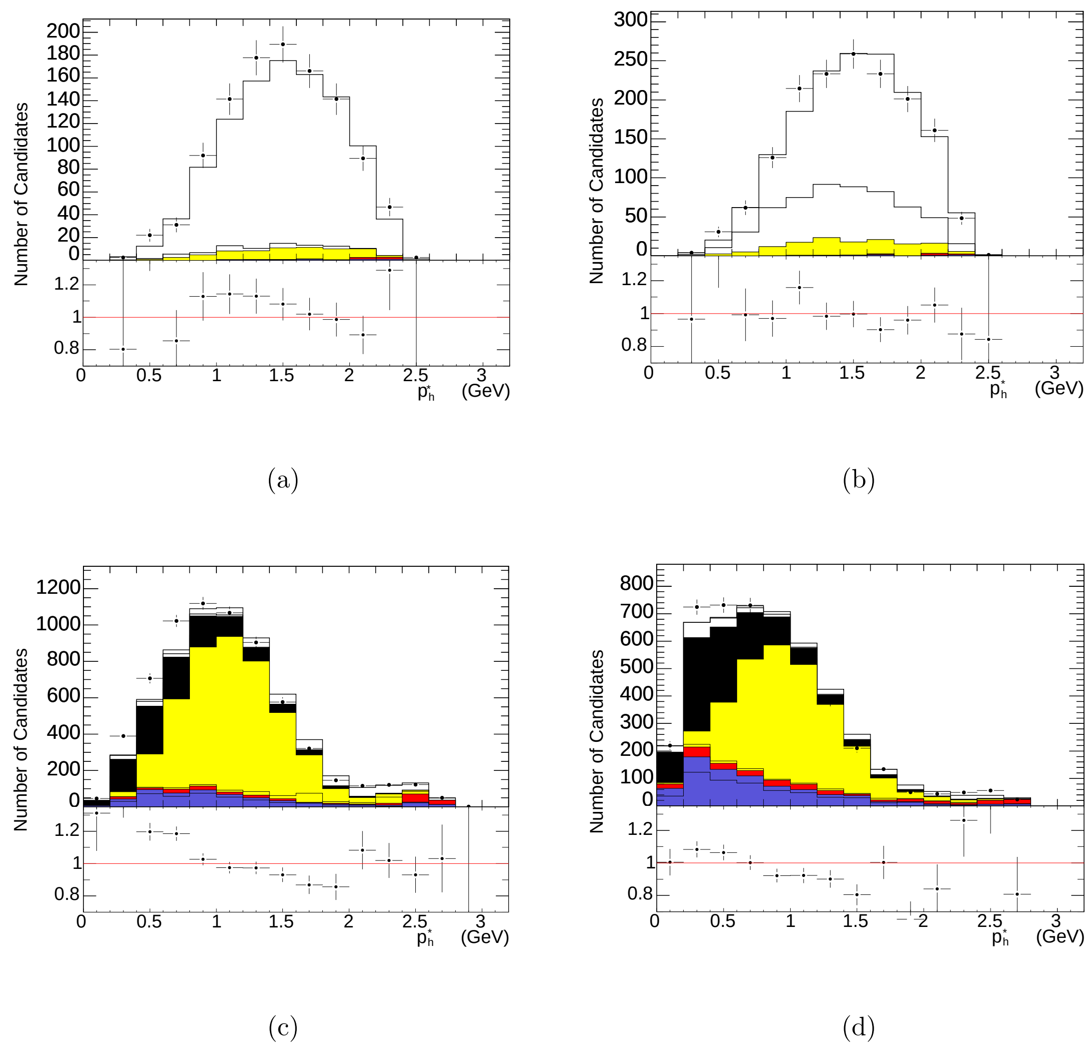


Figure 5.12: Data-Monte Carlo comparison for $|p_h^*|$, for (a) $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-$, (b) $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-\pi^0$, (c) $B^0 \to \pi^-\ell^+\nu$, (d) $B^+ \to \pi^0\ell^+\nu$ after all other selection cuts have been applied.

5.3 Efficiencies of Critical Cuts

The full selection criteria applied to signal $b \to u\ell\nu$ events are applied to the control samples, excepting only those cuts which exploit features specific to a given channel, such as hadron masses, or the correlation between the lepton and hadron momentum in the $\Upsilon(4S)$ frame.

Tables 5.1 and 5.2 show the evolution of the ratio of the yield in data over the yield in Monte Carlo, for the $\overline{D^0}$ decay channels $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-$ and $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-\pi^0$, as each of the more critical analysis cuts are applied. In general, those cuts which show the largest deviations between the data and Monte Carlo efficiencies are those involved in or dependent on the neutrino reconstruction: namely, the cuts on $m^2_{miss}/2E_{miss}$, $|p_{miss}|$, θ_{miss}, and $\Delta\theta_{min}$; all of these effects are at the level of $1-2\%$ in the incremental efficiency, except for the effect in $m^2_{miss}/2E_{miss}$, which is $8 \pm 4\%$ in the $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-$ channel and $2 \pm 3\%$ in the $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-\pi^0$ channel. This suggests that the Monte Carlo simulation underestimates the efficiency of the $m^2_{miss}/2E_{miss}$ cut by about $4-5\%$. $\cos\theta_{\mathrm{BY}}$, which was somewhat problematic in the previous section in terms of data/MC agreement, also shows an effect at the 1% level. The other critical cuts track to better than a percent between data and Monte Carlo.

In principle, the only selection efficiencies that are critical to know well are those for the signal selection; the relative fractions of the dominant backgrounds will be allowed to float in the fit. Nonetheless, it is somewhat instructive to repeat this exercise for the signal decay channels, as shown in Tables 5.3 and 5.4. In this case, it is not only the effect of a selection cut on an individual component that enters, but the effect of the selection cut on both signal and all background components. Thus any discrepancy in modelling the selection efficiency for a particular background component may affect the different decay channels, with their different background compositions, to a greater or lesser degree.

On average, all the deviations are somewhat larger, which is to be expected, since the unfitted distributions do not precisely adjust the relative sizes of the background components. Nonetheless, only a few cuts show discrepancies larger than 1.5%: the missing-mass squared cut, which shows a $6-10\%$ effect, the cut on $L2$ and $\cos\theta_{thrust}$, at a little over 3%, the cut on $\cos\theta_{\mathrm{BY}}$, at $4-5\%$, the cut on p_{miss} and θ_{miss}, which is a little over 3% for the $B^+ \to \pi^0\ell^+\nu$ channel, and the cut on $p_l^* + p_h^*$, which is about 5% for the $B^+ \to \pi^0\ell^+\nu$ channel. Again, the cut on the missing-mass squared shows the largest difference, although the sign of that effect is opposite to that seen in the control sample channels, likely because we are sensitive to the effect in the signal-channel background rather than the signal itself.

5.4 Resolution of Missing Momentum and Final Fit Variables

There is no way to test the accuracy of the Monte Carlo simulation directly in regards to the missing momentum resolution, since there is no available sample in data for which the neutrino momentum four-vector can be unambiguously determined. It can, however, be assessed indirectly, as it dominates the resolution of the variables ΔE m_{ES} and q^2; moreover, it is the resolution of these variables which directly impacts the final fit, and is thus of critical interest.

Figure 5.13 shows Monte Carlo derived resolutions for the magnitude and direction of the (measured) neutrino, as well as the Monte Carlo derived resolution for the q^2, for pure signal candidates for the signal channels $B \to \pi\ell\nu$ and $B^+ \to \pi^0\ell^+\nu$ and as well as the control sample channel $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-$ and $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-\pi^0$. The resolution for all three quantities shows only slight variations between the signal channels, but is noticeably broader in all cases for the control sample. This is not surprising: the softer neutrino spectrum for $B \to D^*\ell\nu$ decays means that any lost tracks or showers or undetected neutrals, such as K_L^0, will have a larger relative impact on the missing momentum measurement.

Figures 5.14 and 5.15 show the signal and sideband projections for ΔE and m_{ES} in the control samples. The general agreement is encouraging, although there is a clear broadening of both the ΔE and m_{ES} resolutions in data relative to the Monte Carlo. Figures 5.16 shows both q^2_{raw} and q^2_{corr} for the two control samples; the agreement in this case is also quite good, and the corrections applied to q^2_{corr} clearly improve the q^2 measurement as they should, eliminating most of the spillover beyond the kinematic limit.

5.5 Overview

In general, most of the critical variable distributions for this analysis are good to at the $10-20\%$ percent level. The $B^+ \to \pi^0 \ell^+ \nu$ channel shows some overall normalization effects, and there are clear signs, particularly in the $B^0 \to \pi^- \ell^+ \nu$ channel, that the insufficiency of off-resonance data and consequent use of continuum Monte Carlo leaves continuum background components unaccounted for; these components, however, are effectively suppressed by the full signal selection. The results from the control sample indicate that the majority of the cuts are well-modelled in terms of signal efficiency, with the significant exception of the cut on $m^2_{miss}/2E_{miss}$ and that the ΔE and $m_{\rm ES}$ resolutions are noticeably broader in data (in all likelihood, these two effects are correlated), although the reconstructed q^2 distributions. both raw and corrected, appear well-modelled. The discrepancies in critical cut variables, and in final fit variable resolutions, will be taken in account when systematic errors are assessed, the first by studying the stability of the fit results with respect to variations in cut values, and the second by studying stability of the fit results with respect to bin size.

Table 5.1: Initial yields and incremental efficiencies of critical selection cuts for both data and Monte Carlo for $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-$, e and μ channels combined. Cuts are applied sequentially. Where errors of 0.00 are shown, the actual error is less than .005.

	Incremental efficiency of a single cut (%)		
	Data	Monte Carlo	Data/MC
Candidate	17129 ± 129	19878 ± 70	
Lepton fiducial cut: $0.4090 < \theta < 2.3720$	96.42 ± 0.14	96.21 ± 0.07	1.00 ± 0.00
$\mathrm{Q_l} < 0 \| \|\theta_\mathrm{l} < 1.9\mathrm{rad}\|\|\mathrm{p_l^*} < 2.2\,\mathrm{GeV}$	99.96 ± 0.01	99.98 ± 0.01	1.00 ± 0.00
$\mathrm{p_l^*} > 1.3\,\mathrm{GeV}$	100.	100.	1.00 ± 0.00
$R2 < 0.4$	94.12 ± 0.18	94.18 ± 0.09	1.00 ± 0.00
$\mathrm{nTrk} > 4 \| \mathrm{nTrk} = 4 \& \mathrm{nPhot} \geq 2$	100.	100.	1.00 ± 0.00
$\mathrm{L2} + 14.28\cos(\theta_\mathrm{thr}) \leq 12.00$; $\mathrm{L2} - 1.875\cos(\theta_\mathrm{thr}) \leq 1$	62.2 ± 0.4	62.16 ± 0.18	1.00 ± 0.01
$0.7\,\mathrm{GeV} < \mathrm{p_{miss}}; 0.6 < \theta_\mathrm{miss} < 2.9$	46.8 ± 0.5	46.22 ± 0.22	1.01 ± 0.01
$\|\Delta\mathrm{Q}\| \leq 1$	88.0 ± 0.5	88.52 ± 0.21	0.99 ± 0.01
$\|\cos(\Theta_\mathrm{BY})\| \leq 1.1$	93.9 ± 0.4	94.81 ± 0.15	0.99 ± 0.00
$\Delta(\theta_\nu) < 0.6\,\mathrm{rad}$	77.7 ± 0.7	76.02 ± 0.31	1.02 ± 0.01
$\|\mathrm{m^2_{miss}}/(2 * \mathrm{E_{miss}})\| < 0.40$	29.4 ± 0.9	27.2 ± 0.4	1.08 ± 0.04

Table 5.2: Initial yields and incremental efficiencies of critical selection cuts for both data and Monte Carlo for $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-\pi^0$, e and μ channels combined. Cuts are applied sequentially. Where errors of 0.00 are shown, the actual error is less than .005.

	Incremental efficiency of a single cut (%)		
	Data	Monte Carlo	Data/MC
Candidate	33000 ± 179	39911 ± 102	
Lepton fiducial cut: $0.4090 < \theta < 2.3720$	96.7 ± 0.10	96.33 ± 0.05	1.00 ± 0.00
$\mathrm{Q_l} < 0 \| \|\theta_\mathrm{l} < 1.9\mathrm{rad}\| \|\mathrm{p_l^*} < 2.2\,\mathrm{GeV}$	99.97 ± 0.01	99.97 ± 0.00	1.00 ± 0.00
$\mathrm{p_l^*} > 1.3\,\mathrm{GeV}$	100.	100.	1.00 ± 0.00
$R2 < 0.4$	94.28 ± 0.13	94.43 ± 0.06	1.00 ± 0.00
$\mathrm{nTrk} > 4 \| \mathrm{nTrk} = 4 \& \mathrm{nPhot} \geq 2$	100.	100.	1.00 ± 0.00
$\mathrm{L2} + 14.28\cos(\theta_\mathrm{thr}) \leq 12.00$; $\mathrm{L2} - 1.875\cos(\theta_\mathrm{thr}) \leq 1$	59.48 ± 0.29	59.98 ± 0.13	0.99 ± 0.01
$0.7\,\mathrm{GeV} < \mathrm{p_{miss}}; 0.6 < \theta_\mathrm{miss} < 2.9$	45.5 ± 0.4	44.62 ± 0.16	1.02 ± 0.01
$\|\Delta\mathrm{Q}\| \leq 1$	87.7 ± 0.4	87.08 ± 0.16	1.01 ± 0.00
$\|\cos(\Theta_\mathrm{BY})\| \leq 1.1$	90.2 ± 0.4	91.51 ± 0.14	0.99 ± 0.00
$\Delta(\theta_\nu) < 0.6\,\mathrm{rad}$	74.9 ± 0.6	74.04 ± 0.24	1.01 ± 0.01
$\|\mathrm{m^2_{miss}}/(2 * \mathrm{E_{miss}})\| < 0.40$	24.5 ± 0.6	24.02 ± 0.27	1.02 ± 0.03

Table 5.3: Initial yields and incremental efficiencies of critical selection cuts for both data and Monte Carlo for $B^0 \to \pi^- \ell^+ \nu$, e and μ channels combined. Cuts are applied sequentially. Where errors of 0.00 are shown, the actual error is less than .005.

	Incremental efficiency of a single cut (%)		
	Data	Monte Carlo	Data/MC
Candidate	1236927 ± 1107	1061271 ± 828	
Hadron: Lepton veto	87.062 ± 0.030	87.487 ± 0.026	1.00 ± 0.00
$\lvert m_{\pi\ell} - m_{\Psi}\rvert > 25\,MeV$	99.06 ± 0.01	97.82 ± 0.01	1.01 ± 0.00
Lepton fiducial cut: $0.4090 < \theta < 2.3720$	92.36 ± 0.03	91.54 ± 0.02	1.01 ± 0.00
$\mathrm{Q_l} < 0 \Vert \theta_\mathrm{l} < 1.9\mathrm{rad} \Vert \mathrm{p}^*_\mathrm{l} < 2.2\,\mathrm{GeV}$	95.85 ± 0.02	96.61 ± 0.02	0.99 ± 0.00
$\mathrm{p}^*_\mathrm{l} > 1.3\,\mathrm{GeV}$	100.	100.	1.00 ± 0.00
$\mathrm{p}^*_\mathrm{l} + \mathrm{p}^*_\mathrm{h} \geq 2.9$	52.53 ± 0.05	51.74 ± 0.05	1.02 ± 0.00
$R2 < 0.4$	68.84 ± 0.07	69.68 ± 0.06	0.99 ± 0.00
$\mathrm{nTrk} > 4 \vert \mathrm{nTrk} = 4 \& \mathrm{nPhot} \geq 2$	98.60 ± 0.02	99.36 ± 0.01	0.99 ± 0.00
$\mathrm{L2} + 14.28\cos(\theta_\mathrm{thr}) \leq 12.00$; $\mathrm{L2} - 1.875\cos(\theta_\mathrm{thr}) \leq 1$	39.25 ± 0.08	40.48 ± 0.07	0.97 ± 0.00
$0.7\,\mathrm{GeV} < \mathrm{p_{miss}}; 0.6 < \theta_\mathrm{miss} < 2.9$	42.31 ± 0.14	42.73 ± 0.09	0.99 ± 0.00
$\lvert \Delta \mathrm{Q} \rvert \leq 1$	84.28 ± 0.15	85.22 ± 0.09	0.99 ± 0.00
$\lvert \cos(\Theta_\mathrm{BY}) \rvert \leq 1.1$	70.23 ± 0.21	74.03 ± 0.12	0.95 ± 0.00
$\Delta(\theta_\nu) < 0.6\,\mathrm{rad}$	84.20 ± 0.20	83.68 ± 0.12	1.01 ± 0.00
$\lvert \mathrm{m^2_{miss}} / (2 * \mathrm{E_{miss}}) \rvert < 0.40$	13.30 ± 0.20	14.07 ± 0.12	0.95 ± 0.02

Table 5.4: Initial yields and incremental efficiencies of critical selection cuts for both data and Monte Carlo for $B^+ \to \pi^0 \ell^+ \nu$, e and μ channels combined. Cuts are applied sequentially. Where errors of 0.00 are shown, the actual error is less than .005.

	Incremental efficiency of a single cut (%)		
	Data	Monte Carlo	Data/MC
Candidate	916495 ± 954	795110 ± 786	
Lepton fiducial cut: $0.4090 < \theta < 2.3720$	91.69 ± 0.03	90.72 ± 0.03	1.01 ± 0.00
$\mathrm{Q_l} < 0 \| \|\theta_\mathrm{l} < 1.9\mathrm{rad}\| \|\mathrm{p}_\mathrm{l}^* < 2.2\,\mathrm{GeV}$	92.14 ± 0.03	93.21 ± 0.03	0.99 ± 0.00
$\mathrm{p}_\mathrm{l}^* > 1.3\,\mathrm{GeV}$	100.	100.	1.00 ± 0.00
$\mathrm{p}_\mathrm{l}^* + \mathrm{p}_\mathrm{h}^* \geq 2.9$	42.41 ± 0.06	40.28 ± 0.05	1.05 ± 0.00
$R2 < 0.4$	64.98 ± 0.08	65.53 ± 0.08	0.99 ± 0.00
$\mathrm{nTrk} > 4 \| \mathrm{nTrk} = 4 \& \mathrm{nPhot} \geq 2$	96.28 ± 0.04	97.14 ± 0.03	0.99 ± 0.00
$\mathrm{L2} + 14.28\cos(\theta_\mathrm{thr}) \leq 12.00$; $\mathrm{L2} - 1.875\cos(\theta_\mathrm{thr}) \leq 1$	33.06 ± 0.10	34.21 ± 0.09	0.97 ± 0.00
$0.7\,\mathrm{GeV} < \mathrm{p_{miss}}; 0.6 < \theta_\mathrm{miss} < 2.9$	40.64 ± 0.19	42.11 ± 0.14	0.97 ± 0.01
$\|\Delta \mathrm{Q}\| \leq 1$	82.32 ± 0.23	83.24 ± 0.15	0.99 ± 0.00
$\|\cos(\Theta_\mathrm{BY})\| \leq 1.1$	70.33 ± 0.30	73.54 ± 0.20	0.96 ± 0.00
$\Delta(\theta_\nu) < 0.6\,\mathrm{rad}$	85.38 ± 0.28	85.75 ± 0.18	1.00 ± 0.00
$\|\mathrm{m}^2_\mathrm{miss}/(2 * \mathrm{E_{miss}})\| < 0.40$	13.99 ± 0.30	15.61 ± 0.20	0.90 ± 0.02

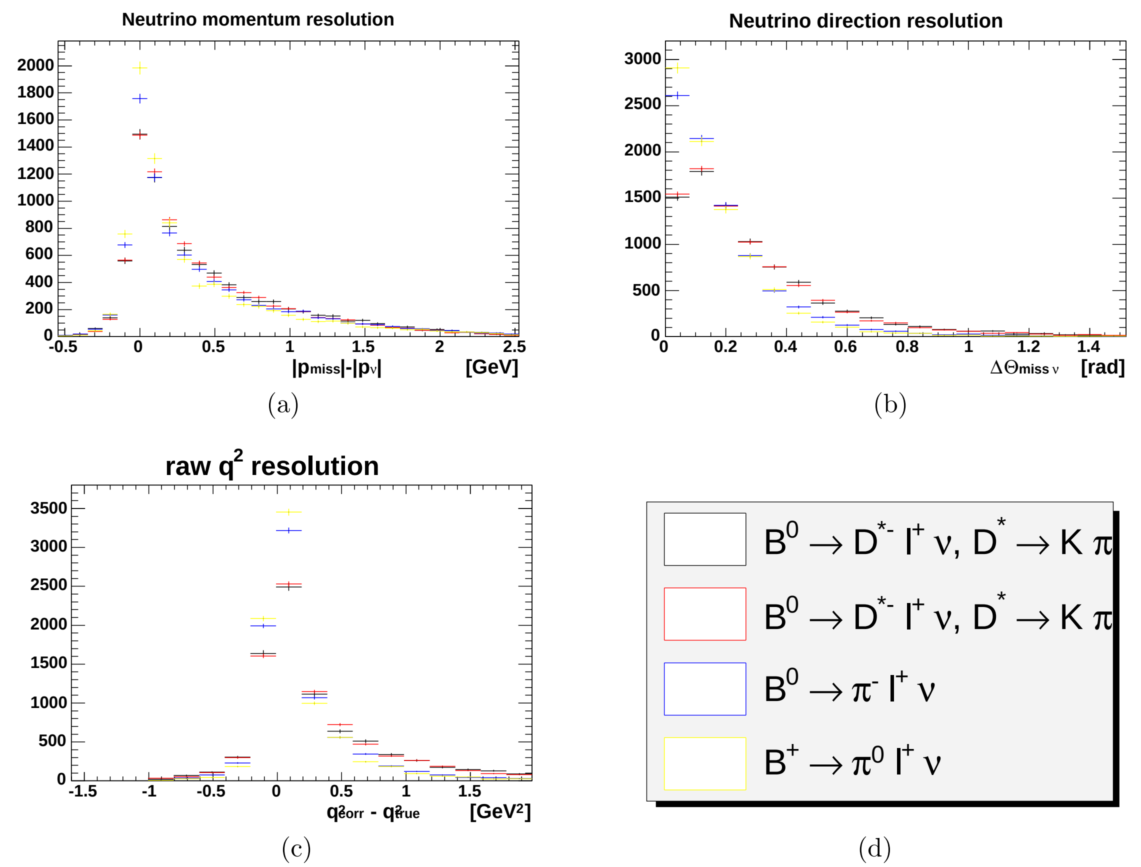


Figure 5.13: Neutrino and associated q^2 resolution for perfectly reconstructed Monte Carlo truth candidates for the control samples and both $B \to \pi\ell\nu$ decay channels, after the neutrino reconstruction cuts have been applied. (a) Difference between the magnitude of the missing momentum and that of the true neutrino (b) Angular difference ($\Delta\Theta_{miss\nu}$) between the direction of the missing momentum and the true neutrino (c) Raw q^2 resolution calculated using the lepton and missing momentum.

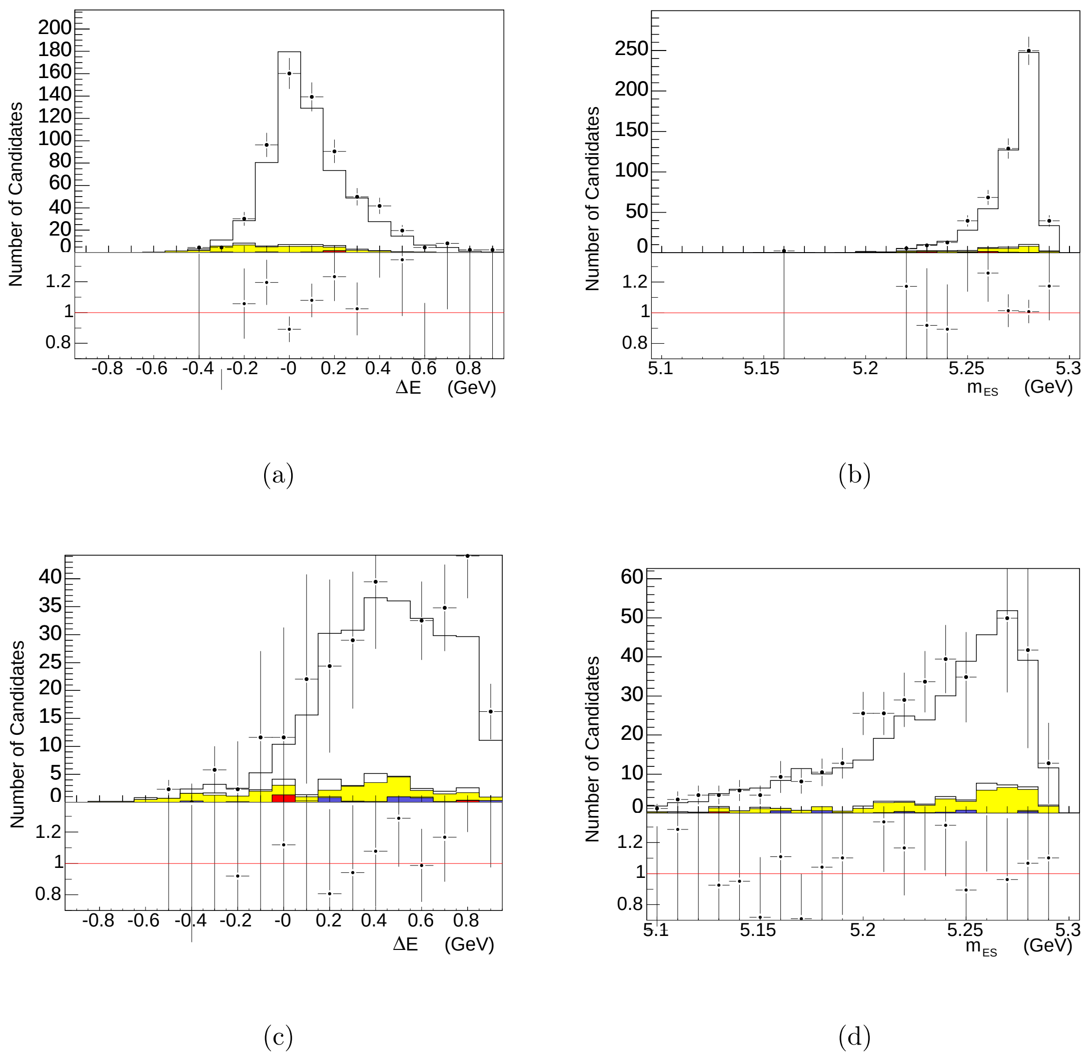


Figure 5.14: Data-Monte Carlo comparison for $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-$ after all selection cuts have been applied, for (a) ΔE projected in the m_{ES} signal band (b) m_{ES} projected in the ΔE signal band (c) ΔE projected in the m_{ES} sideband (d) m_{ES} projected in the ΔE sideband.

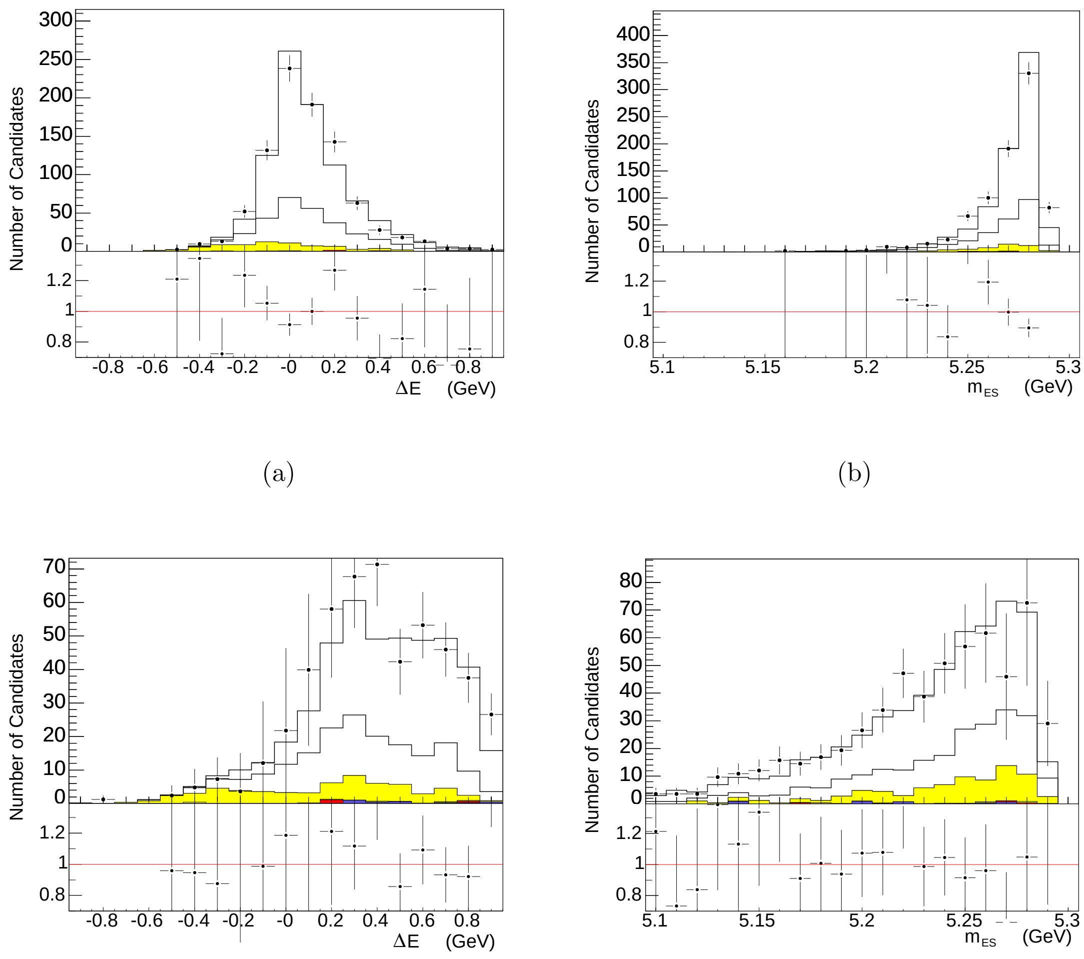


Figure 5.15: Data-Monte Carlo comparison for $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-\pi^0$ after all selection cuts have been applied, for (a) ΔE projected in the $m_{\rm ES}$ signal band (b) $m_{\rm ES}$ projected in the ΔE signal band (c) ΔE projected in the $m_{\rm ES}$ sideband (d) $m_{\rm ES}$ projected in the ΔE sideband.

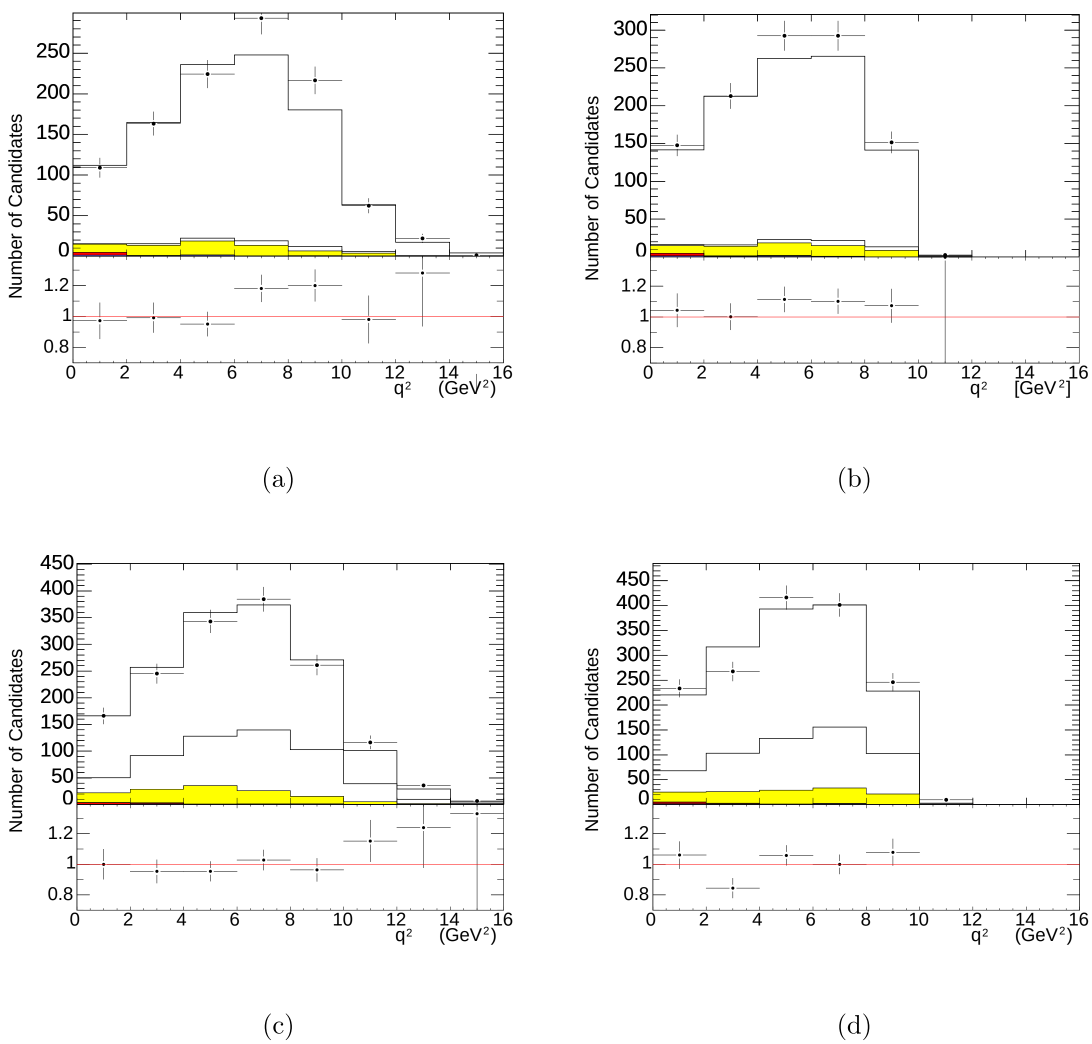


Figure 5.16: Data-Monte Carlo comparison for after all selection cuts have been applied, for (a) q^2_{raw}, $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-$, (b) q^2_{corr}, $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-$, (c) q^2_{raw}, $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-\pi^0$, (b) q^2_{corr}, $B^0 \to D^{*-}\ell^+\nu, \bar{D}^0 \to K^+\pi^-\pi^0$.

Chapter 6

Fit Technique and Definitions

To extract the signal yields for $B \to X_u \ell \nu$, we perform a fit to the selected samples for a set of exclusively-reconstructed decay modes. The scale factors for the signal and background contributions in the selected event samples are, in general, free parameters of the fit. All data distributions are fit simultaneously, and, where possible, with common fit parameters. A signal contribution in one reconstructed mode may come from the same source as background in another mode, and therefore these sources share a common fit parameter; likewise, where possible, all background contributions from a common physics source are taken to have a common scale factor.

The fit is performed using a set of two-dimensional distributions in the variables ΔE and m_{ES}. Since no analytic form for these distributions is available, a binned maximum-likelihood method is used, with the bin sizes chosen to optimize signal and background separation while retaining adequate statistics in most or all bins. The choice of a two-dimensional distribution in ΔE and m_{ES} is mandated by the fact that the two variables are significantly correlated for both signal and some background sources (Figure 6.1), making it impossible to factorize the likelihood function.

6.1 Fit Variables

All candidates that contribute inside the region

$$|\Delta E| < 0.9\,\text{GeV} \quad \text{and} \quad 5.095 < m_{ES} < 5.305\,\text{GeV} \tag{6.1}$$

are included in the fit; the standard partitioning of this two-dimensional space into bins is shown in Figure 6.1. The "signal region," which is defined for illustrative purposes only, is

$$-0.15 < \Delta E < 0.25\,\text{GeV} \quad \text{and} \quad m_{ES} > 5.255\,\text{GeV}. \tag{6.2}$$

and represents the part of the fit space where signal-to-background ratio is greatest for all modes. The size of the signal region is determined by the resolution of the signal ΔE and m_{ES} distributions. The substitution of $\vec{p}_{\text{miss}}$ for E_{miss} introduces the

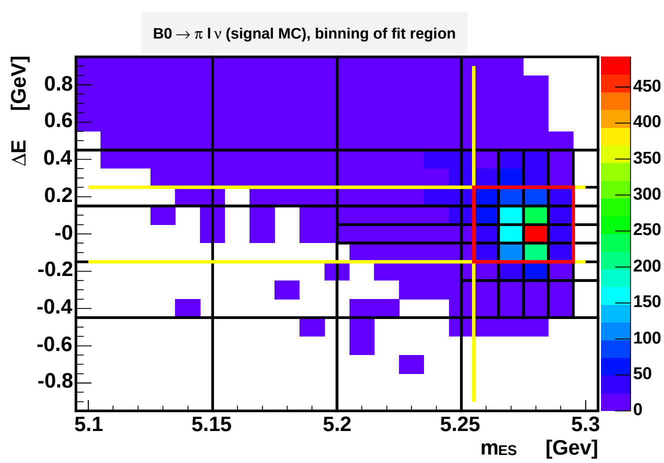


Figure 6.1: ΔE^* vs. M_{ES} distribution for $B \to \pi\ell\nu$ signal events, with the nominal binning of the fit superimposed in black. The "signal region" is shown in red, with the signal band boundaries indicated in yellow.

correlations between ΔE and m_{ES}. Moreover, the neutrino resolution dominates the resolution of both variables, such that the tails of the ΔE and m_{ES} distributions, where the correlation is most evident, are dominated by candidates from events where the neutrino has been poorly reconstructed due to lost tracks and undetected neutrals. Choice of a nominal ΔE vs. $m_{\rm ES}$ binning outside the signal region is dictated primarily by the shapes and statistical power of the background contributions and the size and shape of this correlated $\Delta E/m_{\rm ES}$ tail. Choice of a nominal ΔE vs. $m_{\rm ES}$ binning within the signal region is dictated by the resolution near the peak of the ΔE and $m_{\rm ES}$ distributions.

Figure 6.1 shows the two-dimensional distribution of events for $B \to \pi^- \ell^+ \nu$ with the boundaries of the bins marked.

6.2 Fit Structure

The data are divided into n bins in the two-dimensional distribution ΔE vs. $m_{\rm ES}$. If d_i is the number of selected candidates in bin i, and a_{ji} is the number of selected MC candidates from source j in this bin, then

$$N_D = \sum_{i=1}^{n} d_i, \qquad N_j = \sum_{i=1}^{n} a_{ji}, \tag{6.3}$$

where N_D is the total number of candidates in the data sample, and N_j is the total number in the MC sample for *source* j. We assume that there are m different source distributions that add up to describe the data. The predicted number of candidates in each bin, $f_i(P_j)$ can be written in terms of the strength of the individual contributions, $P_j \quad (j = 1, m)$, as

$$f_i = N_D \sum_{j=1}^{m} P_j w_{ji} a_{ji}/N_j = \sum_{j=1}^{m} p_j w_{ji} a_{ji}, \tag{6.4}$$

with $p_j = N_D P_j/N_j$; the weights w_{ji} account for the relative normalization of the samples and various other corrections.

In the case of multiple reconstructed decay channels, this generalizes to a set of parameters p_j^h, associated with the expected numbers of events $f_i^h(P_j^h)$, where h indexes the reconstructed channel.

A *source*, in this context, is completely defined by the physics which produces the candidate lepton, as discussed in Section 4.2 (index j), and the reconstructed signal decay channel to which the source candidates contribute (index h). From the perspective of the fit, a *fit source* consists of three pieces of information for each bin i: the numbers of selected candidates, a^h_{ji}, the associated bin-by-bin weights w'^h_{ji}, and an overall normalization W^h_j. [1] For example, the fit source representing the contribution of $B^0 \to \rho^- \ell^+ \nu$ decays ($l = e, \mu$ combined) to the reconstructed $B^0 \to \pi^- \ell^+ \nu$ channel would be encoded:

$$\text{Source}^h_{ji} = W^h_i \times w'^h_{ji} \times a^h_{ji} \tag{6.5}$$

where $h = B^0 \to \rho^- \ell^+ \nu$, $j = B^0 \to \pi^- \ell^+ \nu$, and the overall normalization

$$\begin{aligned} W^h_j &= N_{expected}/N_{generated} &(6.6)\\ &= \frac{\mathcal{B}(B^0 \to \rho^- \ell^+ \nu)_{MC} \times 2 \times N_{B\bar{B}}}{N_{generated}} &(6.7) \end{aligned}$$

with $N_{B\bar{B}}$ being the number of $B\overline{B}$ events in data.

Other Monte Carlo signal and $b \to u\ell\nu$ sources use the same procedure to determine their overall normalization. The remaining sources involving B decays ($b \to c\ell\nu$, $b \to other$) are normalized using the total number of $B\overline{B}$ events in the $B\overline{B}$ Monte Carlo, and the total number of $B\overline{B}$ events in the data. Continuum contributions are normalized to the same luminosity as the data, taking into account the production cross sections at the $\Upsilon(4S)$ resonance.

Rearranging this definition makes it clear that the fit sources include the candidate selection efficiencies. If we consider $\mathcal{P}^h_{ji}$ to represent the true, underlying probability density function for a given source h, j, such that $a^h_{ji} = N_{selected} \times \mathcal{P}^h_{ji}$ in the limit of

[1] This distinction is not as artificial as it might seem. The fit used in this analysis uses the raw statistics of each background source in order to correctly determine fit errors which take the limited Monte Carlo statistics into account. Moreover, we weight candidates individually in order to correct for particle identification efficiency, π^0 efficiency, etc. Each bin of a fit source in fact encodes two separate pieces of information: the raw number of selected candidates, a^h_{ji} and an average weight w^h_{ij}, defined such that the product of the average weight and the raw number of candidates is the weighted number of candidates in that bin.

large Monte Carlo statistics, then

$$\begin{aligned}
\text{Source}^h_{ji} &= W^h_j \times w'^h_{ji} \times a^h_{ji} && (6.8)\\
&= N_{expected} \times N_{selected}/N_{generated} \times \mathcal{P}^h_{ji} && (6.9)\\
&= N_{expected} \times \epsilon_{sel} \times \mathcal{P}^h_{ji} && (6.10)
\end{aligned}$$

The fit parameters associated with each source, p^h_j, represent a scale factor between the expected number of events from source h, j, as predicted by the Monte Carlo, and the number of events from source h, j, actually required by the data. It can be seen that for signal sources, this translates into a scale factor for the branching ratio assumed in the Monte Carlo. Taking our previous example:

$$p^h_j \times \text{Source}^h_{ji} = (p^h_j \mathcal{B}(B \to \rho\ell\nu)_{MC}) \times 2 \times N_{B\bar{B}} \epsilon_{sel} \mathcal{P}^h_{ji} \quad (6.11)$$

The full set of sources and associated parameters is summarized in Table 6.1; the numerical indices assigned to sources and channels will be used throughout.

Obviously, when defining the actual fit, the table above not only involves an excessive number of free parameters, but in fact represents four decoupled fits, one for each of the four reconstructed decay channels. We reduce the number of free parameters by fixing parameters associated with very small sources to a nominal value; we also share parameters associated with sources of a common physics process across signal decay channels. This allows us both to combine background sources (for instance, the $B^+ \to c\ell\nu$ and $B^0 \to c\ell\nu$ sources are generally allowed to share a single fit parameter, and thus act as a single background contribution in the fit), and to couple the individual channels into a single simultaneous fit.

One pair of sources per channel always shares the same fit parameters, regardless of which variation of the fit is used. These are the *signal* and *combinatoric signal* sources, first discussed in Chapter 4. The *signal* source consists of signal channel decays where both the lepton and hadron are correctly reconstructed, while the *combinatoric signal* source is populated by candidates where the lepton originates from the signal decay, but the hadron does not. The sources which originate from other signal decays (*i.e.*

Table 6.1: Most general possible set of sources and associated parameters for all reconstructed channels. $m = 1, 2$ corresponds to $\ell = e, \mu$, $n = 1, 2, 3$ corresponds to the three running period for years 2000, 2001, and 2002, and l indexes any other possible channel subdivision (e.g. lepton charge, reconstructed q^2, B flavor). Parameters associated with combinatoric signal sources are marked in red.

Fit Parameters by Source					
		Channel			
		$B^0 \to \pi^-\ell^+\nu$	$B^+ \to \pi^0\ell^+\nu$	$B^0 \to \rho^-\ell^+\nu$	$B^+ \to \rho^0\ell^+\nu$
Source Classification		1	2	3	4
signal	1	p_1^{1mnl}	p_1^{2mnl}	p_1^{3mnl}	p_1^{4mnl}
$B^0 \to \pi^-\ell^+\nu$	2	$p_2^{1mnl} = p_1^{1mnl}$	p_2^{2mnl}	p_2^{3mnl}	p_2^{4mnl}
$B^+ \to \pi^0\ell^+\nu$	3	p_3^{1mnl}	$p_3^{2mnl} = p_1^{2mnl}$	p_3^{3mnl}	p_3^{4mnl}
$B^0 \to \rho^-\ell^+\nu$	4	p_4^{1mnl}	p_4^{2mnl}	$p_4^{3mnl} = p_1^{3mnl}$	p_4^{4mnl}
$B^+ \to \rho^0\ell^+\nu$	5	p_5^{1mnl}	p_5^{2mnl}	p_5^{3mnl}	$p_5^{4mnl} = p_1^{4mnl}$
$B^0 \to u\ell\nu$	6	p_6^{1mnl}	p_6^{2mnl}	p_6^{3mnl}	p_6^{4mnl}
$B^+ \to u\ell\nu$	7	p_7^{1mnl}	p_7^{2mnl}	p_7^{3mnl}	p_7^{4mnl}
$B^0 \to c\ell\nu$	8	p_8^{1mnl}	p_8^{2mnl}	p_8^{3mnl}	p_8^{4mnl}
$B^+ \to c\ell\nu$	9	p_9^{1mnl}	p_9^{2mnl}	p_9^{3mnl}	p_9^{4mnl}
$B^0 \to other$	10	p_{10}^{1mnl}	p_{10}^{2mnl}	p_{10}^{3mnl}	p_{10}^{4mnl}
$B^+ \to other$	11	p_{11}^{1mnl}	p_{11}^{2mnl}	p_{11}^{3mnl}	p_{11}^{4mnl}
B^0B^0 fake l.	12	p_{12}^{1mnl}	p_{12}^{2mnl}	p_{12}^{3mnl}	p_{12}^{4mnl}
B^+B^- fake l.	13	p_{13}^{1mnl}	p_{13}^{2mnl}	p_{13}^{3mnl}	p_{13}^{4mnl}
$q\bar{q}$	14	p_{14}^{1mnl}	p_{14}^{2mnl}	p_{14}^{3mnl}	p_{14}^{4mnl}
$q\bar{q}$ fake l.	15	p_{15}^{1mnl}	p_{15}^{2mnl}	p_{15}^{3mnl}	p_{15}^{4mnl}

the source contribution of $B^0 \to \rho^-\ell^+\nu$ in the $B^0 \to \pi^-\ell^+\nu$ channel) are referred to as *crossfeed*.

While the combinatoric contribution of a signal source into its own channel (*i.e.* candidates in the $B^0 \to \pi^-\ell^+\nu$ channel where the lepton is from a $B^0 \to \pi^-\ell^+\nu$ decay but the hadron is not) is always considered combinatoric signal, additional sources may act as combinatoric signal or crossfeed, depending on the fit definition. For instance, in a variation of the fit in which we fit only for the combined branching fraction $B(B \to \pi\ell\nu)$ $(P_1^1 = P_2^2)$, the sources associated with P_1^2 and P_2^1 also act combinatoric signal, rather than crossfeed.

6.2.1 Two-Channel Fit

The two-channel fit uses data in only two of the four reconstructed decay channels: $B^0 \to \pi^- \ell^+ \nu$ and $B^+ \to \pi^0 \ell^+ \nu$. It is q^2-independent, and combines both lepton flavor channels and data from all run periods. The relative normalization parameter for sources of a common type is shared across decay channels; this reduces the possible eight free parameters associated with the $B^0 \to \pi^- \ell^+ \nu$ and $B^+ \to \pi^0 \ell^+ \nu$ sources in each channel to two. The nominal fit also uses the following isospin relations

$$\Gamma(B^0 \to \pi^- \ell^+ \nu) = 2\Gamma(B^+ \to \pi^0 \ell^+ \nu) \tag{6.12}$$

$$\Gamma(B^0 \to \rho^- \ell^+ \nu) = 2\Gamma(B^+ \to \rho^0 \ell^+ \nu) \tag{6.13}$$

to fix the relative ratio of the $B^0 \to \pi^- \ell^+ \nu$ to $B^+ \to \pi^0 \ell^+ \nu$ source and that of the $B^0 \to \rho^- \ell^+ \nu$ source to the $B^+ \to \rho^0 \ell^+ \nu$. This reduces the two free parameters associated with the $B \to \pi \ell \nu$ sources to a single parameter; the parameter associated with the $B^0 \to \rho^- \ell^+ \nu$ and $B^+ \to \rho^0 \ell^+ \nu$ sources is fixed, and the branching ratio used to calculate the source normalizations is $\mathcal{B}(B^0 \to \rho^- \ell^+ \nu) = 2.69 \times 10^{-4}$.

The background sources from B decays of different flavor are allowed to share a single fit parameter, combining, for instance, the $B^+ \to c\ell\nu$ and $B^0 \to c\ell\nu$ sources in any decay channel into a single effective $b \to c\ell\nu$ source. The $q\bar{q}$ true lepton and fake lepton sources are allowed to share a common fit parameters as well. All parameters associated with $b \to u\ell\nu$, $b \to$ *other*, and $B\overline{B}$ fake sources are fixed at 1.0, which fixes the levels of these sources to that predicted by the simulation.

The fit has very little independent sensitivity to either the $b \to$ *other* and $B\overline{B}$ fake contributions, both of which are quite small, and thus the parameters associated with these sources must remain fixed. With this sole exception, any of the constraints discussed in this section may be released for the purposes of systematic studies.

The remaining free parameters of the two-channel fit and the associated sources are summarized in Table 6.2. The combinatoric signal contributions are indicated in red, while crossfeed contributions (contributions from sources that cannot be considered

Table 6.2: Set of sources and associated parameters for the nominal two-channel fit. Combinatoric signal contributions are indicated in red.

Free Fit Parameters by Channel and Source			
		Channel	
		$B^0 \to \pi^- \ell^+ \nu$	$B^+ \to \pi^0 \ell^+ \nu$
Source		1	2
signal	1	p_1^{1mnl}	p_1^{1mnl}
$B^0 \to \pi^- \ell^+ \nu$	2	p_1^{1mnl}	p_1^{1mnl}
$B^+ \to \pi^0 \ell^+ \nu$	3	p_1^{1mnl}	p_1^{1mnl}
$b \to c \ell \nu$	8	p_8^{1mnl}	p_8^{1mnl}
$q\bar{q}$	14	p_{14}^{1mnl}	p_{14}^{1mnl}

signal or combinatoric signal for a particular decay channel, but would be signal in another reconstructed channel) are indicated in blue.

6.2.2 Four-Channel Fit

The four-channel fit is completely analogous to the two-channel fit, except that it uses data in all four of the reconstructed decay channels $B^0 \to \pi^- \ell^+ \nu$, $B^+ \to \pi^0 \ell^+ \nu$, $B^0 \to \rho^- \ell^+ \nu$, $B^+ \to \rho^0 \ell^+ \nu$, and allows the parameter associated with the $B^0 \to \rho^- \ell^+ \nu$ and $B^+ \to \rho^0 \ell^+ \nu$ sources to float. The free parameters of the four-channel fit and the associated sources are summarized in Table 6.3, with the combinatoric signal contributions indicated in red and crossfeed contributions in blue.

6.3 The Likelihood Fit

6.3.1 The Barlow-Beeston Method

The fit technique used in this analysis is a generalized binned likelihood method which takes into account the statistical fluctuations not only of the on- and off-resonance data samples but also of the various Monte Carlo samples. This method was introduced by R. Barlow and C. Beeston [34].

Table 6.3: Set of sources and associated parameters for the nominal four-channel fit. Combinatoric signal contributions are indicated in red; crossfeed contributions are marked in blue.

Free Fit Parameters by Channel and Source					
		Channel			
		$B^0 \to \pi^- \ell^+ \nu$	$B^+ \to \pi^0 \ell^+ \nu$	$B^0 \to \rho^- \ell^+ \nu$	$B^+ \to \rho^0 \ell^+ \nu$
Source		1	2	3	4
signal	1	p_1^{1mnl}	p_1^{1mnl}	p_1^{3mnl}	p_1^{3mnl}
$B^0 \to \pi^- \ell^+ \nu$	2	p_1^{1mnl}	p_1^{1mnl}	p_1^{3mnl}	p_1^{3mnl}
$B^+ \to \pi^0 \ell^+ \nu$	3	p_1^{1mnl}	p_1^{1mnl}	p_3^{3mnl}	p_3^{3mnl}
$B^0 \to \rho^- \ell^+ \nu$	4	p_4^{1mnl}	p_4^{1mnl}	p_1^{3mnl}	p_1^{3mnl}
$B^+ \to \rho^0 \ell^+ \nu$	5	p_5^{1mnl}	p_5^{1mnl}	p_1^{3mnl}	p_1^{3mnl}
$b \to c \ell \nu$	8	p_8^{1mnl}	p_8^{1mnl}	p_8^{1mnl}	p_8^{1mnl}
$q\bar{q}$	14	p_{14}^{1mnl}	p_{14}^{2mnl}	p_{14}^{3mnl}	p_{14}^{4mnl}

As discussed in Section 6.2, the data is divided into n bins in ΔE vs. m_{ES}. Let us first consider the likelihood for a single reconstructed channel h; we will drop the h index for convenience. The likelihood describes the data in terms of a set of Monte Carlo sources j, where the predicted number of event in each bin, $f_i(P_j)$ can be written in terms of the strength of the individual contributions,P_j:

$$f_i = N_D \sum_{j=1}^{m} P_j w_{ji} a_{ji} / N_j = \sum_{j=1}^{m} p_j w_{ji} a_{ji}, \tag{6.4}$$

with the fit parameters $p_j = N_D P_j / N_j$.

Since the Monte Carlo statistics are limited, the generated numbers of events a_{ji} have statistical fluctuations relative to the value A_{ji} expected for infinite statistics, and thus the more correct prediction for each bin is

$$f_i = \sum_{j=1}^{m} p_j w_{ji} A_{ji}. \tag{6.14}$$

If we assume Poisson statistics for both the data and MC samples, the total likelihood function $\mathcal{L}$ is the combined probability for the observed $\{d_i\}$ and $\{a_{ji}\}$ [34],

$$\ln \mathcal{L} = \sum_{i=1}^{n} d_i \ln f_i - f_i + \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ji}(\ln A_{ji} - A_{ji}). \tag{6.15}$$

The first term has the usual form associated with the uncertainty of the data, and the second term refers to the statistics of the Monte Carlo and is not dependent on data. There are $(n+1) \times m$ unknown parameters that we need to determine, the m relative normalization factors p_j which are of interest to the signal extraction and $n \times m$ values A_{ji}.

Barlow and Beeston [34] have shown that the problem can be significantly simplified. Namely, the $n \times m$ quantities A_{ji} can be determined by solving n simultaneous equations for A_{ji} of the form

$$\frac{d_i}{1 - t_i} = f_i = \sum_{j=1}^{m} p_j w_{ji} A_{ji} = \sum_{j=1}^{m} \frac{p_j w_{ji} a_{ji}}{1 + p_j w_{ji} t_i}, \tag{6.16}$$

with $A_{ji} = a_{ji}/(1 + p_j w_{ji} t_i)$ and the substitution $t_i = 1 - d_i/f_i$ (for $d_i = 0$ we define $t_i = 1$). At every step on the minimization of $-2\ln\mathcal{L}$ these n independent equations have to be solved. This procedure not only determines the parameters p_j, but also produces improved estimates for the various contributions A_{ji} in each bin.

The combined log-likelihood function is the sum of this expression for each decay channel

$$\ln \mathcal{L} = \sum_{h=1}^{n_{channel}} \ln \mathcal{L}_h = \sum_{h=1}^{n_{channel}} \left[\sum_{i=1}^{n} (d_i^h \ln f_i^h - f_i^h) + \sum_{i=1}^{n} \sum_{j=1}^{m} (a_{ji}^h \ln A_{ji}^h - A_{ji}^h) \right], \tag{6.17}$$

where the number of decay channels can be as small as two (in the case of the two-channel fit) or as large as twelve (in the case of a four-channel fit separated by year). For the two-channel fit, the parameters p_j^h are related as specified by Table 7.1.

6.3.2 Limited Monte Carlo Statistics and Zero Content Bins

A standard likelihood function treats sources as binned probability density functions. In this case, when a source contains a bin with no candidates, the fit interprets this to mean that there is zero probability this source may contribute to this bin in the data. In the limit of high statistics, this is perfectly correct. But if the statistics of the source itself are relevant, the possibility that this empty bin is the result of a downward statistical fluctuation must be taken into account.

Barlow and Beeston describe this in detail in [34]. In the case where, for bin i, the a_{ji} for one or more Monte Carlo sources j (and hence the A_{ji}) are zero, there are several possibilities. First, it may still be possible to accommodate the data in that bin without adjusting the values of the A_{ji}; if the fit can do so, it will make no further adjustment. If the data cannot be accommodated, then for the source j with the highest weight, the A_{ji} for that source is adjusted to make up the difference. In this case, the adjusted source is the one that would make the most significant additional contribution to the data in that bin.

The weight of a source j may be large because the associated fit parameter is large, or because the source comes from a sample with very few events, in which case the normalization of the source results in a large source weight. In the case where all fit parameters are of roughly equal strength, this results in the source with the weakest statistics being adjusted. This result is reassuring, since it is the source with the weakest statistics that is most likely to be subject to a statistical fluctuation of this type.

6.3.3 Extreme Low-Statistics Sources and Smoothing

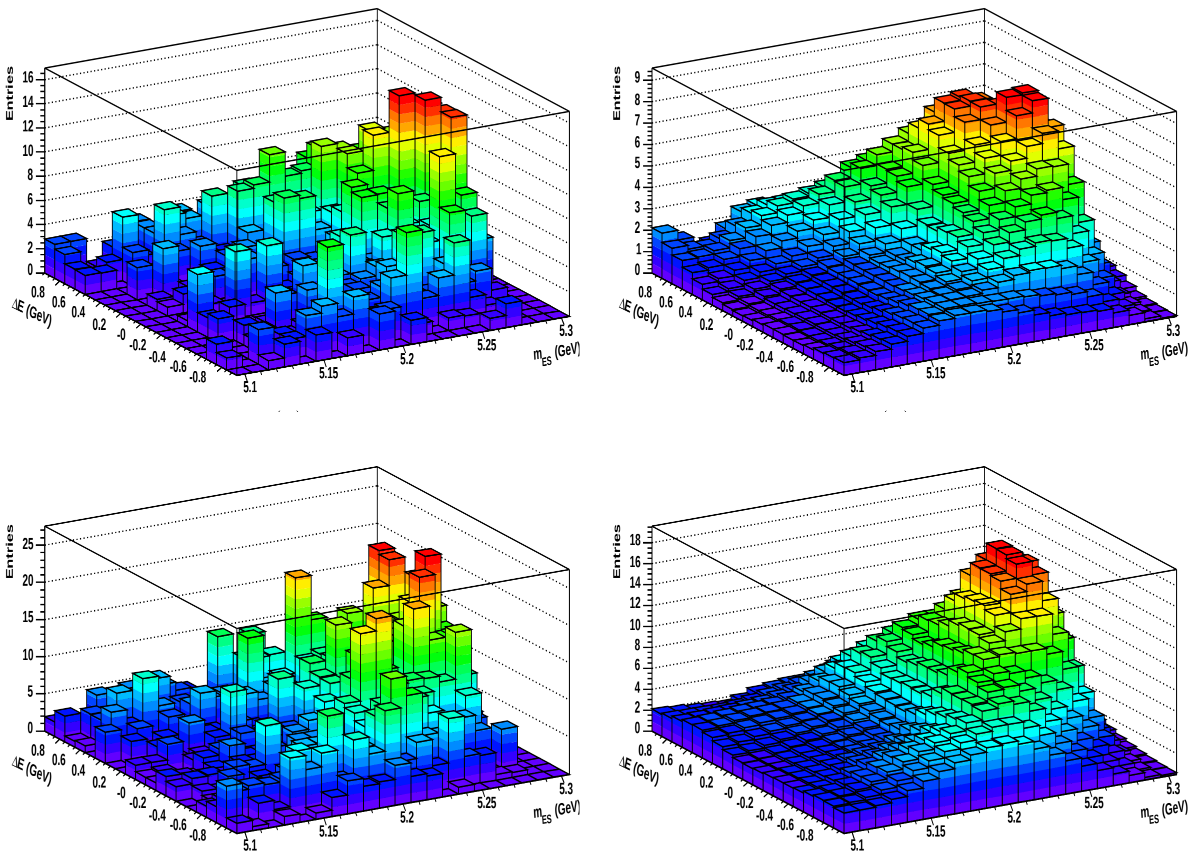
The prescription described above works quite well for sources of intermediate statistics—*i.e.* sources which are derived from a statistically limited sample, whose statistical errors must be accommodated in the fit, but which are nevertheless sufficiently well-populated that the source shape is well-determined. On the other hand, sources containing only a handful of candidates may pose a problem, particularly in the case where the source is statistically weak because the sample used to derive the source is statistically limited, but the source itself makes a significant contribution in the fit. In this situation, when there is no longer sufficient information for the source shape to be well-determined, the fit can only react to (and often amplify) statistically insignificant fluctuations. As a result, the fit becomes unstable.

This is the case, for instance, for the $q\bar{q}$ background source in the electron channel of every decay mode. Thus, for the $q\bar{q}$ true and fake lepton background in both the electron and muon channels, we derive smooth source shapes from the raw sources using a hexagonal diffusion smoothing algorithm [35]. Effectively infinite-statistics (~ -100000 candidates) sources are generated in each case using the smoothed distribution as a pdf, and then scaled to the integral of the original source pre-smoothing. Since the hexagonal diffusion algorithm, unlike the fit, deals with bin-to-bin correlations, it makes more effective use of the limited shape information in these weak sources. By incorporating the smoothed sources into the fit with apparently high statistics, the fit is forced to ignore the statistical contribution of that source.

In practice, in order to obtain sufficient statistics for the smoothing, we loosen the cut in L2 and $\cos\theta_{thrust}$ to $L2 + 14.28\cos\theta_{thrust} < 14$. Figures 6.2 and 6.3 show the dramatic effect the smoothing algorithm has on the $q\bar{q}$ source distributions.

6.3.4 Tests of Technique

To test the sensitivity of the fit to the number of events in a given background sample, a set of toy Monte Carlo studies were performed, with simplified functional forms used to mock up the signal Monte Carlo, $b \to c\ell\nu$, and $q\bar{q}$ distributions. In each case, functional forms were chosen that most closely approximated the shapes of the actual distributions: a modified Gaussian for the signal, and combinations of Landau and ARGUS functions for the backgrounds. The projected one-dimensional shapes of these distributions are shown in Figure 6.4. The functional forms were combined with a fixed set of weights and a ‘data" sample of a few thousand events produced from the sum via random number generator. Effective sources were also generated from these functional forms via random number generator, with the effective "$q\bar{q}$" source generated at varying statistical levels relative to the toy "data." For each level of



(c) (d)

Figure 6.2: Effect of smoothing on $q\bar{q}$ background sources in the $B^0 \to \pi^- \ell^+ \nu$ channel. The contribution from $q\bar{q}$ events from candidates containing real leptons is shown before smoothing (a) and after smoothing (b); the contribution from candidates containing fake leptons is also shown before smoothing (c) and after smoothing (d). The continuum suppression cut is set to $L2 + 14.28 \cos\theta_{thrust} < 14$

.

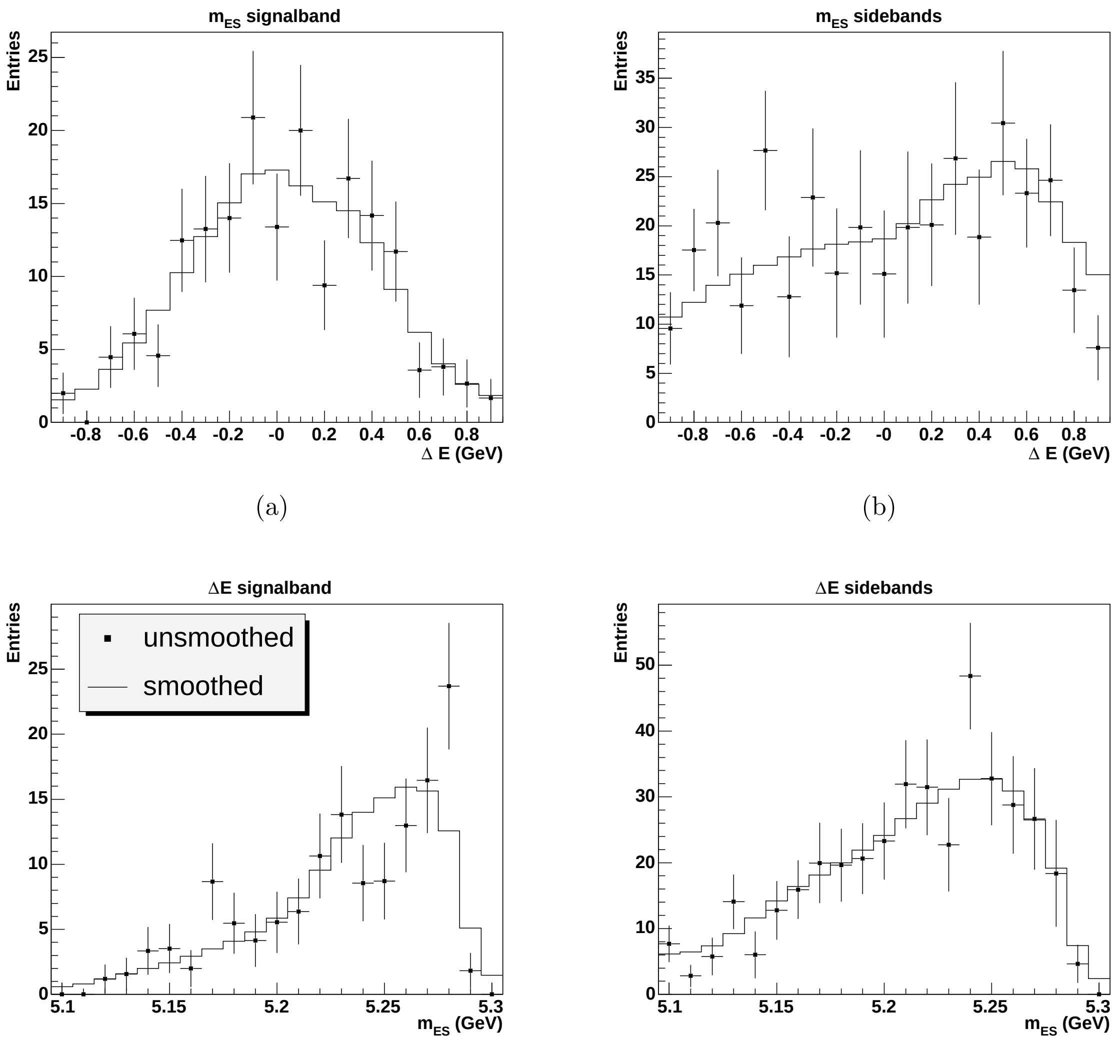
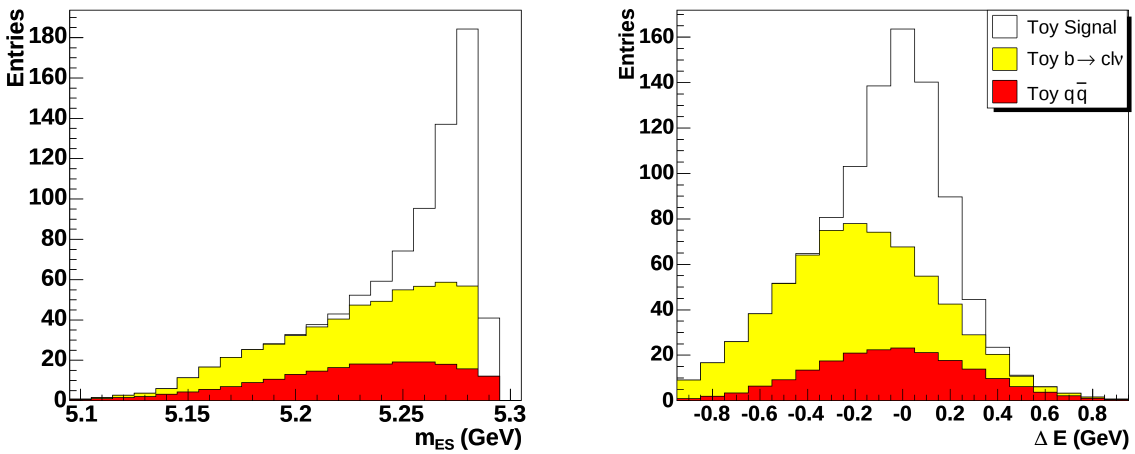


Figure 6.3: Signal (a,c) and sideband (b,d) projections of ΔE (a,b) and $m_{\rm ES}$ (c,d) distributions for candidates containing true leptons from $q\bar{q}$ MC, before smoothing (data points) and after (histograms). The continuum suppression cut is set to $L2 + 14.28\cos\theta_{thrust} < 14$



(a) (b)

Figure 6.4: One-dimensional projections of the toy source distributions used in tests of the fit technique.

“$q\bar{q}$” source statistics, a thousand datasets were randomly generated and fitted with randomly generated toy sources; the distribution of results for each of the three scale parameters is fitted with a single Gaussian. The results are summarized in Table 6.4.

Table 6.4: Toy model test of Barlow fit stability and bias, as a function of the relative size of the $q\bar{q}$background sample to the total number of events in the data (2800 candidates).

“$q\bar{q}$” stat./data	“signal”		“$b \to c\ell\nu$”		“$q\bar{q}$”	
	Mean	RMS	Mean	RMS	Mean	RMS
10	0.999	0.074	1.003	0.023	0.990	0.079
5	0.999	0.074	1.003	0.023	0.990	0.080
1	0.999	0.076	1.003	0.023	0.990	0.081
0.1	1.001	0.086	1.004	0.024	0.984	0.094
0.05	1.003	0.098	1.005	0.026	0.979	0.109

On average, the toy fits reproduce the input parameters to a few tenths of a percent for the “signal” and “$b \to c\ell\nu$” scale parameters, and within a percent for the “$q\bar{q}$” scale parameter, while the errors are $2-10\%$. We conclude that the fit technique has no significant instrinsic bias.

A second crosscheck was performed with the results of the actual two-channel fit to data, verifying that the weighted sums of the a_{ji} and of the adjusted A_{ji} are equal or nearly equal for each source. In general, this condition is met within small deviations, which are the consequence of a non-zero spread of weights in bins with low statistics.

	2000		2001		2002	
	$\sum w_i a_i$	$\sum w_i A_i$	$\sum w_i a_i$	$\sum w_i A_i$	$\sum w_i a_i$	$\sum w_i A_i$
signal	1296	1293	1552	1552	2130	2140
$B^0 \to \pi^- \ell^+ \nu$	51	51	71	71	91	91
$B^+ \to \pi^0 \ell^+ \nu$	62	62	98	98	63	63
$B^0 \to \rho^- \ell^+ \nu$	543	543	639	641	805	805
$B^+ \to \rho^0 \ell^+ \nu$	752	750	473	474	502	502
$B^+ \to \omega \ell^+ \nu$	334	334	217	217	225	225
$B^+ \to u \ell \nu$ (non-res.)	438	438	972	972	506	506
$B^0 \to u \ell \nu$ (non-res.)	462	462	828	828	462	462
$B^+ \to u \ell \nu$ (res.)	100	100	321	321	108	108
$B^0 \to u \ell \nu$ (res.)	114	114	332	332	95	95
$B^+ \to c \ell \nu$	752	751	1718	1721	1880	1895
$B^0 \to c \ell \nu$	552	552	1096	1098	1226	1228
$B^+ \to$ *other*	42	42	155	155	158	158
$B^0 \to$ *other*	49	48	90	90	131	131
$q\bar{q}$	100000	100000	100000	100001	99999	100003
$q\bar{q}$ fake l.	99999	100001	99999	100004	100000	100007

Table 6.5: Comparison of raw and adjusted source integrals for the two-channel fit. The abbreviations *res.* and *non-res.* indicate resonant and non-resonant $b \to u \ell \nu$ decays, respectively.

Chapter 7

Fit Results

7.1 Results of Fit to $B \to \pi\ell\nu$ Channels

The results of the nominal two-channel fit, along with a number of variations, are summarized in Table 7.1. The nominal fit result gives a 9% relative statistical error on the parameter $p_1^1 = \mathcal{B}(\mathcal{B} \to \pi\ell\nu)/\mathcal{B}(\mathcal{B} \to \pi\ell\nu)_{\mathcal{MC}}$, a 3% relative statistical error on the $b \to c\ell\nu$ background estimate (p_8^1), and a 10% relative statistical error on the $q\bar{q}$ background estimate (p_{14}^1). For this fit, the $\mathcal{B}(B^0 \to \rho^-\ell^+\nu)$ branching fraction is fixed at 2.69×10^{-4}. The fit also assumes a B^0 to B^+ production ratio, $f_{+-}/f_{00} = 1$, and adjusts the relative branching ratios for conjugate B^0 and B^+ decays to account for the difference in the measured B lifetimes, $\tau_{B^0}/\tau_{B^+} = 1.086 \pm 0.017$ [2].

The fit results are generally consistent within statistical errors for different running periods, as are the results from fitting the $B \to \pi e\nu$ and $B \to \pi e\mu$ channels separately. The electrons-only fit gives a smaller value for p_1^1 than the nominal fit, while the muons-only fit gives a higher result. There is a significant anti-correlated shift in the $q\bar{q}$ normalization parameter, which increases by a few sigma for the electron channel and decreases by approximately the same amount for the muon channel.

This effect suggests that it is driven, not by any systematic uncertainty in the electron and muon efficiencies, but by the uncertainty in the modelling of the $q\bar{q}$ background for the two channels. The lack of $q\bar{q}$ MC statistics required the combination of lepton channels and reconstructed decay modes for a loosened set of cuts

Table 7.1: Results of simultaneous fit to the decay channels $B^0 \to \pi^-\ell^+\nu$ and $B^+ \to \pi^0\ell^+\nu$. The fit marked "$B^0 \to \pi^-\ell^+\nu$ / $B^+ \to \pi^0\ell^+\nu$" is performed with the $B^0 \to \pi^-\ell^+\nu$ and $B^+ \to \pi^0\ell^+\nu$ scale parameters, and the $b \to c\ell\nu$ scale parameters for the $B^0 \to \pi^-\ell^+\nu$ and $B^+ \to \pi^0\ell^+\nu$ channels, floated separately. The simultaneous fit to both $B \to \pi\ell\nu$ and $B \to \rho\ell\nu$ decay channels also floats the $B \to \rho\ell\nu$ scale parameter, p_1^3, obtaining a value of $p_1^3 = 0.71 \pm 0.07$.

	Parameter			
Fit Type	p_1^1 ($B \to \pi\ell\nu$)	p_8^1 ($b \to c\ell\nu$)	p_{14}^1 ($q\bar{q}$)	χ^2/dof
Nominal	1.03 ± 0.09	0.99 ± 0.03	1.09 ± 0.10	387/273
e only	0.94 ± 0.12	0.93 ± 0.05	2.06 ± 0.38	318/257
μ only	1.12 ± 0.14	1.05 ± 0.05	0.79 ± 0.11	388/258
Year 2000 ($k = 1$)	0.98 ± 0.18	0.99 ± 0.06	0.89 ± 0.20	125/88
Year 2001 ($k = 2$)	0.90 ± 0.13	0.94 ± 0.05	1.32 ± 0.17	102/90
Year 2002 ($k = 3$)	1.22 ± 0.18	1.05 ± 0.07	1.02 ± 0.26	145/89
$B^0 \to \pi^-\ell^+\nu$ / $B^+ \to \pi^0\ell^+\nu$	1.03 ± 0.11 1.07 ± 0.15	0.94 ± 0.04 0.71 ± 0.07	1.49 ± 0.14	364/271
No Barlow-Beeston term	1.00 ± 0.09	0.96 ± 0.03	1.20 ± 0.09	389/273
$B \to \pi\ell\nu$ and $B \to \rho\ell\nu$ $p_1^3 = 0.71 \pm 0.07$	1.07 ± 0.09	1.05 ± 0.03	1.14 ± 0.10	686/532

in order to derive a well-defined $q\bar{q}$ background shape (see Section 6.3.3). It is reasonable to expect that the shape of the $q\bar{q}$ background for the two lepton channels is somewhat different, particularly in the case of the fake $q\bar{q}$ contribution. Thus the variation in the fit results in all likelihood reflects the error inherent in using the "averaged" $q\bar{q}$ background shape across all channels. This issue will be dealt with in greater detail in Section 8.5.4.

A rough measure of the goodness-of-fit is provided by

$$\chi^2/\mathrm{dof} = \sum_{i=1}^{n} \sum_{h=1}^{n_{channels}} \left(\frac{d_i^h - p_j^h w_{ji}^h a_{ji}^h}{d_i^h} \right)^2. \tag{7.1}$$

It is worth noting the the χ^2/dof given in these tables depends only on the statistical errors of the data and does not take the MC statistics into account, as do the pull distributions shown in Figures 7.1 and 7.2.

Data-MC comparisons for the signal and sideband projections in ΔE and m_{ES} after the nominal fit, are shown in Figures 7.1 and 7.2, along with the bin-by-bin pulls $((d_i^h - p_j^h w_{ji}^h a_{ji}^h)/\sqrt{d_i^h})$ in the ΔE-m_{ES} plane. The more substantial pulls for the $B^0 \to \pi^-\ell^+\nu$ channel reflect the fact that the $B^0 \to \pi^-\ell^+\nu$ data seems to favor a significantly broader ΔE and m_{ES} resolution, while the $B^+ \to \pi^0\ell^+\nu$ channel does not. This same effect drives the large pulls in the $b \to c\ell\nu$ and $q\bar{q}$ scale parameters for the channel-separated fits; the fit compensates for the difference in the ΔE and m_{ES} resolutions between the two channels by adjusting the relative ratio of the background components. Since the control sample channels show, in general, broader ΔE and m_{ES} resolutions for data than for Monte-Carlo, this is not surprising. Since there is also some channel-by-channel variation in the size of the effect for the control samples, it is also not surprising that one signal channel might be more heavily impacted. The exact origin of the discrepancy is still unknown.

The nominal fit takes the statistics of the Monte Carlo sources into account. We perform the fit dropping the second term of the likelihood (i.e. a normal maximum likelihood which ignores the Monte Carlo sources statistics), and find that for all parameters, the result is consistent with the results of the nominal fit.

The nominal fit fixes the $\mathcal{B}(B \to \rho\ell\nu)$ branching fraction. In order to test this assumption, we perform a simultaneous fit to all four reconstructed decay channels ($B^0 \to \pi^-\ell^+\nu$, $B^+ \to \pi^0\ell^+\nu$, $B^0 \to \rho^-\ell^+\nu$, $B^+ \to \rho^0\ell^+\nu$), allowing the $B \to \rho\ell\nu$ scale parameter, p_1^3, to float as well. This extended fit prefers a $B \to \rho\ell\nu$ branching ratio that is 30% smaller than that assumed elsewhere in this analysis, but the $B \to \pi\ell\nu$ scale parameter value, $p_1^1 = 1.07$, is completely consistent with the nominal fit result of $p_1^1 = 1.03 \pm 0.09$. Thus the fit for the $\mathcal{B}(B \to \pi\ell\nu)$ branching ratio does not appear to be particularly sensitive to the uncertainty in $\mathcal{B}(B \to \rho\ell\nu)$.

Figures 7.3 and 7.4 illustrate the data-Monte Carlo comparison, post-fit, for ΔE and m_{ES} projections in the $B^0 \to \rho^-\ell^+\nu$ and $B^+ \to \rho^0\ell^+\nu$ channels.

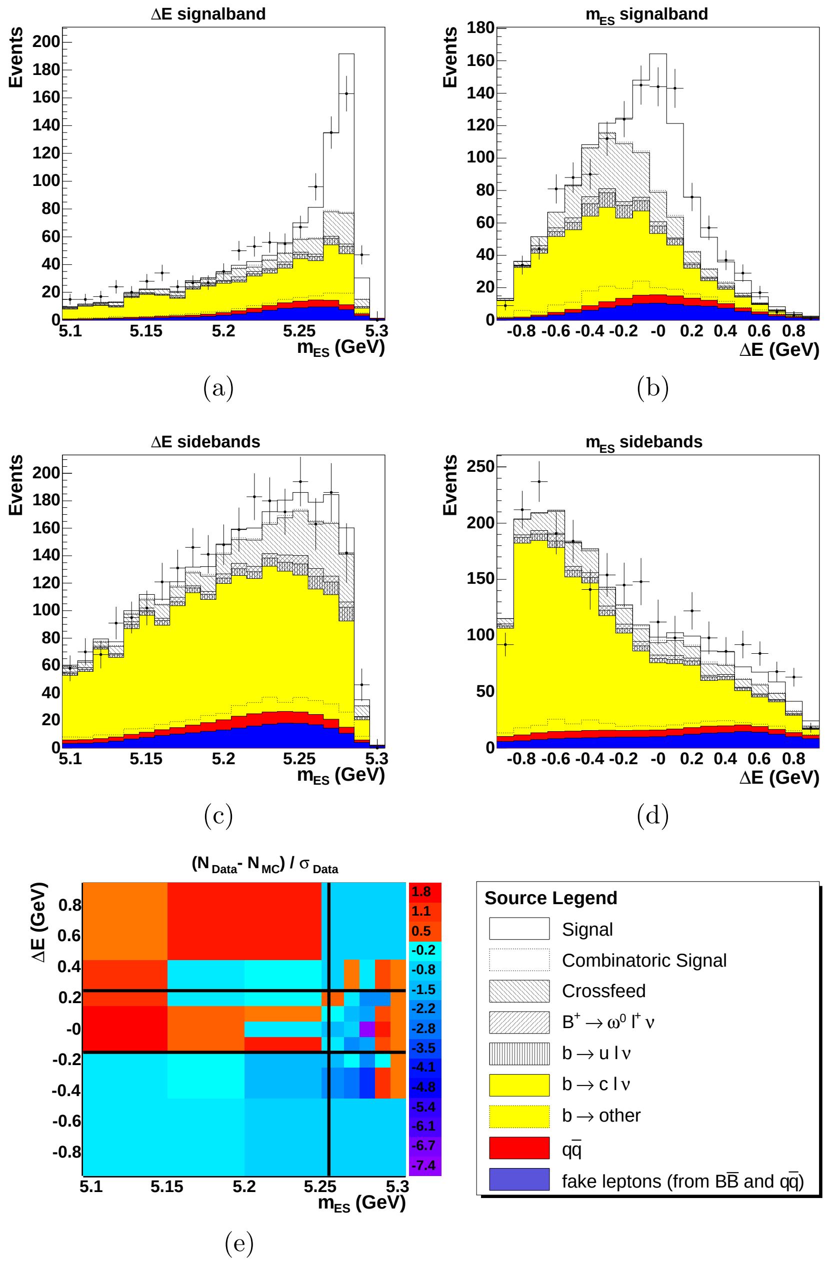


Figure 7.1: $B^0 \to \pi^- \ell^+ \nu$ channel, fitted signal band (a,b) and sideband (c,d) projections in ΔE and m_{ES} as well as the pulls in the ΔE-m_{ES} plane (e). For "combinatoric signal," the lepton originates from either a $B^0 \to \pi^- \ell^+ \nu$ or $B^+ \to \pi^0 \ell^+ \nu$ decay; for "crossfeed", it originates from either a $B^0 \to \rho^- \ell^+ \nu$ or $B^+ \to \rho^0 \ell^+ \nu$ decay.

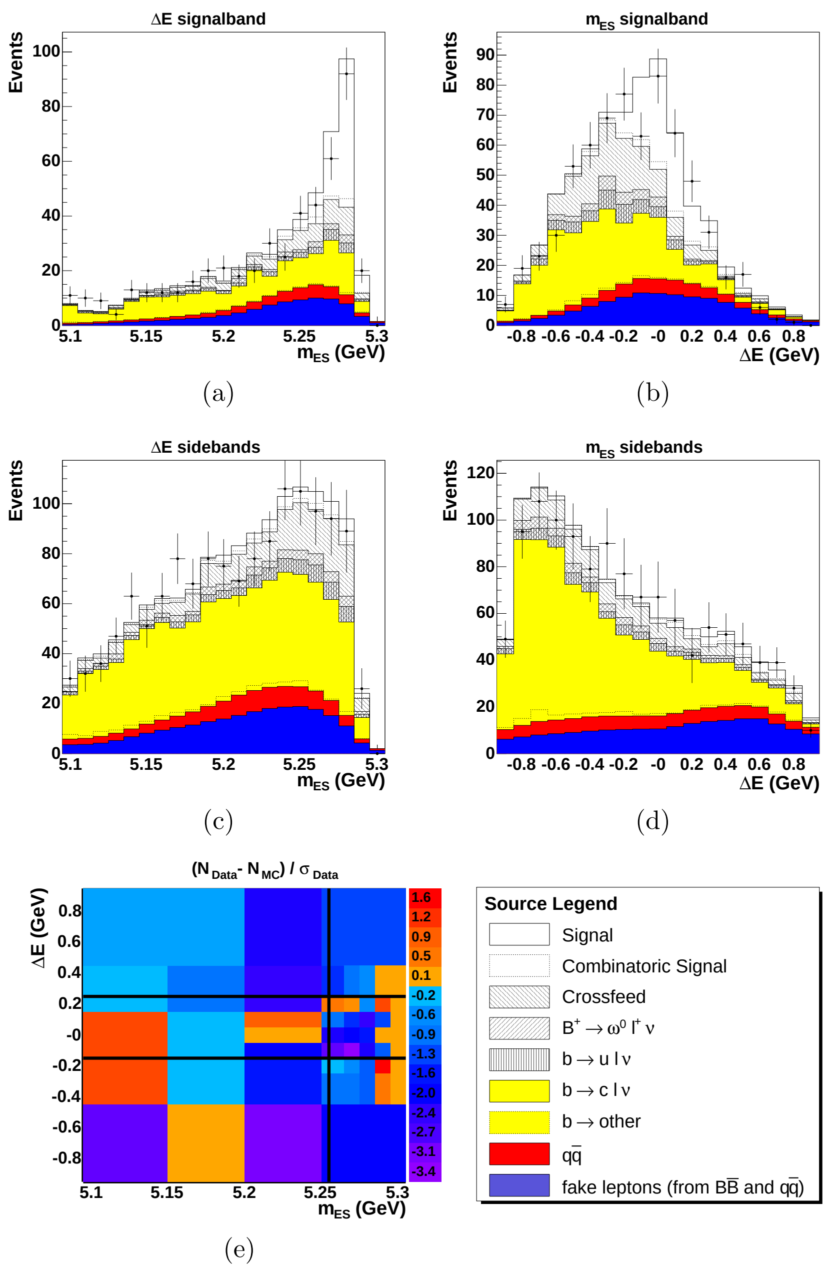


Figure 7.2: $B^+ \to \pi^0\ell^+\nu$ channel, fitted signal band (a,b) and sideband (c,d) projections in ΔE and m_{ES} as well as the pulls in the ΔE-m_{ES} plane (e). For "combinatoric signal," the lepton originates from either a $B^0 \to \pi^-\ell^+\nu$ or $B^+ \to \pi^0\ell^+\nu$ decay; for "crossfeed", it originates from either a $B^0 \to \rho^-\ell^+\nu$ or $B^+ \to \rho^0\ell^+\nu$ decay.

7.2 q^2 Dependence

While the current procedure fits only in ΔE and m_{ES} rather than ΔE, m_{ES}, and q^2, the post-fit q^2 spectra, projected for the signal region, contain interesting information, and are particularly useful when assessing the compatibility of the current results with different $B \to \pi$ form-factor models.

Section 4.5 discusses three different methods of computing the measured q^2 for a reconstructed $B \to \pi\ell\nu$ candidate. The q^2 resolutions for simulated $B \to \pi\ell\nu$ signal candidates, for each of these different methods, are shown in Figure 7.5. Core resolutions for q^2_{corr} and q^2_{had} are quite comparable, although the resolution of q^2_{had} is substantially more Gaussian. Since the practice in this measurement has been to sort the Monte Carlo derived distributions by the truth match of the lepton candidate, the preferred q^2 variable is q^2_{corr}, which will be used for all spectra shown here. We currently use five equal bins from $q^2 = 0$ to $q^2 = 25\,\mathrm{GeV}^2$ ($5\,\mathrm{GeV}^2$ per bin). For either the q^2_{corr} or q^2_{had} variables, the current binning results in a stability (the fraction of candidates generated in a given q^2 bin that reconstruct in that same q^2 bin) and purity (the fraction of candidates reconstructed in given q^2 bin that were also generated in that same q^2 bin) of better than 90%. Figure 7.6 shows the stability and purity as a function of q^2 for both of these variables, as well as for q^2_{raw}, which has a significantly poorer resolution and thus too low a stability and purity to provide a trustworthy measure of the q^2 spectrum with $5\,\mathrm{GeV}$ bins. As will become clear, the need to use information from the reconstructed hadron to improve the q^2 resolution has the unfortunate side effect of distorting the q^2 spectrum not only for semileptonic backgrounds, but for combinatoric signal. While the $B \to \pi\ell\nu$ channels have such small combinatoric signal contributions as to make this irrelevant, it as a significant impact on the $B \to \rho\ell\nu$ channels, and has implications for a future q^2 dependent measurement of $\mathcal{B}(B^0 \to \rho^-\ell^+\nu)$.

The q^2_{corr} spectra for both the $B^0 \to \pi^-\ell^+\nu$ and $B^+ \to \pi^0\ell^+\nu$ channels, after the two-channel fit, are shown in Figure 7.7. The background components contribute to different regions in q^2, with the true lepton $q\bar{q}$ component contributing only in the lowest q^2 bin, the $b \to c\ell\nu$ background contributing primarily to $10 < q^2 < 20\,\mathrm{GeV}^2$,

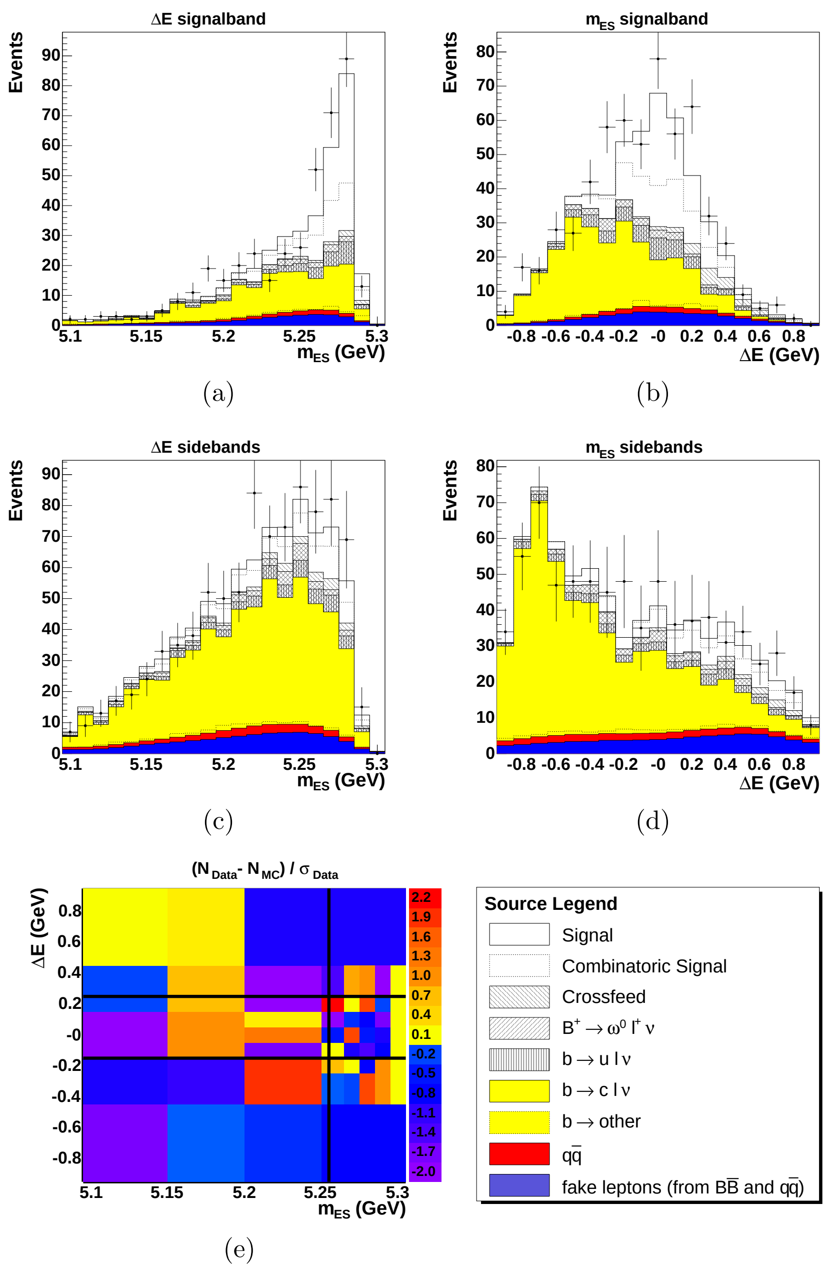


Figure 7.3: $B^0 \to \rho^- \ell^+ \nu$ channel, fitted signal band (a,b) and sideband (c,d) projections in ΔE and m_{ES} as well as the pulls in the ΔE-m_{ES} plane (e). For "combinatoric signal," the lepton originates from either a $B^0 \to \rho^- \ell^+ \nu$ or $B^+ \to \rho^0 \ell^+ \nu$ decay; for "crossfeed", it originates from either a $B^0 \to \pi^- \ell^+ \nu$ or $B^+ \to \pi^0 \ell^+ \nu$ decay.

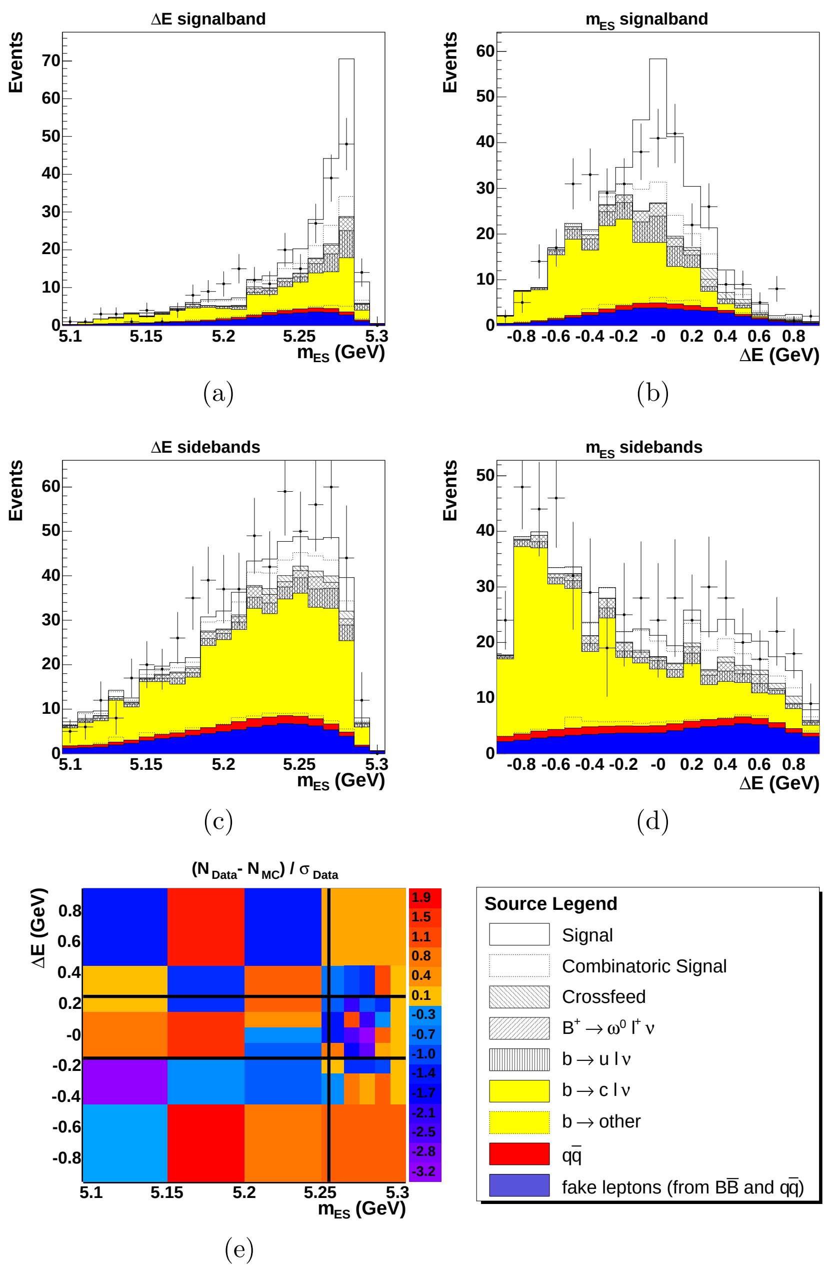
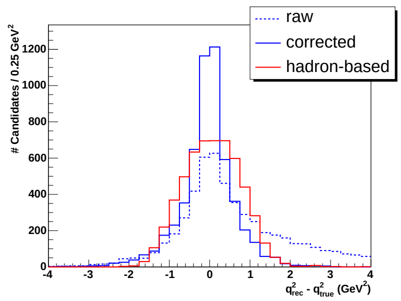
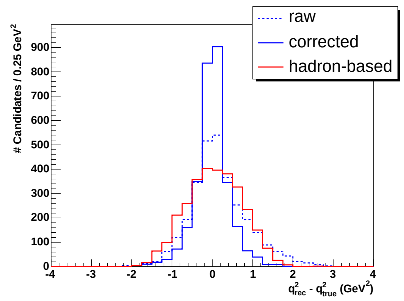


Figure 7.4: $B^+ \to \rho^0\ell^+\nu$ channel fitted signal band (a,b) and sideband (c,d) projections in ΔE and $m_{\rm ES}$ as well as the pulls in the ΔE-$m_{\rm ES}$ plane (e). For "combinatoric signal," the lepton originates from either a $B^0 \to \rho^-\ell^+\nu$ or $B^+ \to \rho^0\ell^+\nu$ decay; for "crossfeed", it originates from either a $B^0 \to \pi^-\ell^+\nu$ or $B^+ \to \pi^0\ell^+\nu$ decay.



(a)

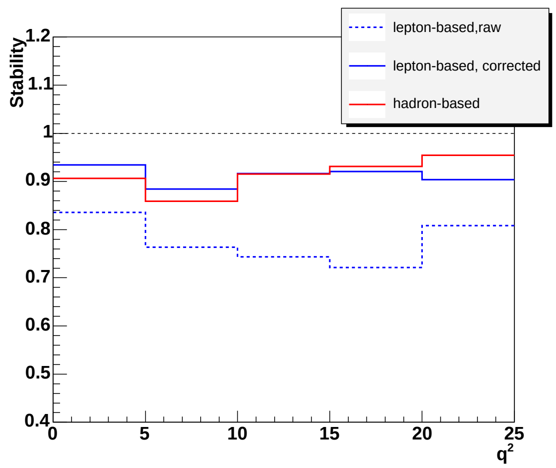


(b)

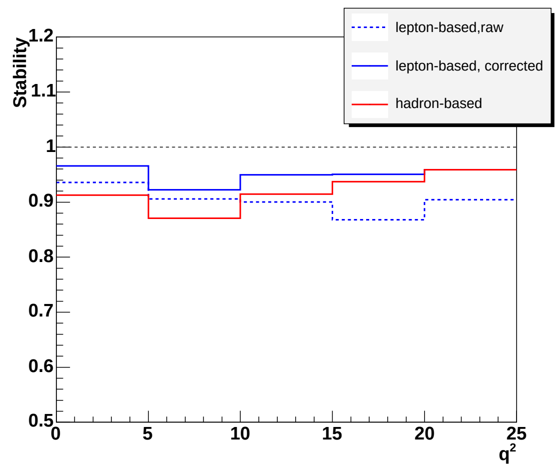
Figure 7.5: For pure signal $B^0 \to \pi^- \ell^+ \nu$ MC candidates, resolutions in (a) the fit region and (b) the signal region of the three q^2 calculations (q^2_{raw}(blue), q^2_{corr}(blue dashed), and q^2_{had})(red) with respect to the true q^2.

and the $B \to \rho\ell\nu$ crossfeed contributing primarily to $15 < q^2 < 25\,\mathrm{GeV}^2$. The fake lepton component, on the other hand, is relatively flat in q^2. In the case of semileptonic background candidates, the q^2 is consistently biased above their expected kinematic limit (i.e the $b \to c\ell\nu$ component contributes between 10 and $20\,\mathrm{GeV}^2$ when the true q^2 distribution cuts off at $10\,\mathrm{GeV}^2$). This is caused by two effects. On the one hand, in the signal region, the standard selection prefers candidates with poorly reconstructed neutrinos; these candidates tend to have a measured q^2 that is substantially (on the order of $5\,\mathrm{GeV}^2$) higher than the actual q^2. Outside the signal region, using hadron information to correct the q^2 derived from the lepton-neutrino pair also shifts the q^2 distribution. This is easily understood. For both $b \to c\ell\nu$ and $B \to \rho\ell\nu$ background candidates in the $B \to \pi\ell\nu$ channels, the lepton and one of daughter pions of the hadron are captured, giving a ΔE that peaks substantially below zero. Since the correction adjusts p_{miss} to bring ΔE to zero, this increases the measured q^2 for a background candidate well above its actual value. Nonetheless, the relative kinematics of the $B \to \rho\ell\nu$ and $b \to c\ell\nu$ contributions are preserved, with the $b \to c\ell\nu$ background having a softer measured q^2 spectrum than that from $B \to \rho\ell\nu$.

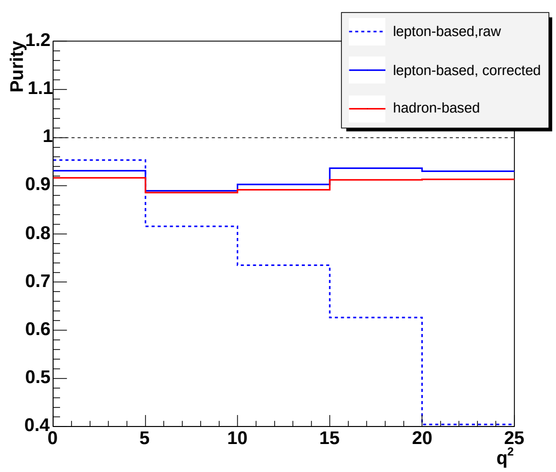
Figure 7.8 shows the q^2_{corr} distribution, after the four-channel fit, for both the $B^0 \to \rho^-\ell^+\nu$ and $B^+ \to \rho^0\ell^+\nu$ channels. The ISGW2 model is used for the signal Monte Carlo prediction. In the case of the $B^+ \to \rho^0\ell^+\nu$ channel, there appears to be a deficit of candidates in a single bin in q^2 ($5\,\mathrm{GeV}^2 < q^2_{corr} < 10\,\mathrm{GeV}^2$).



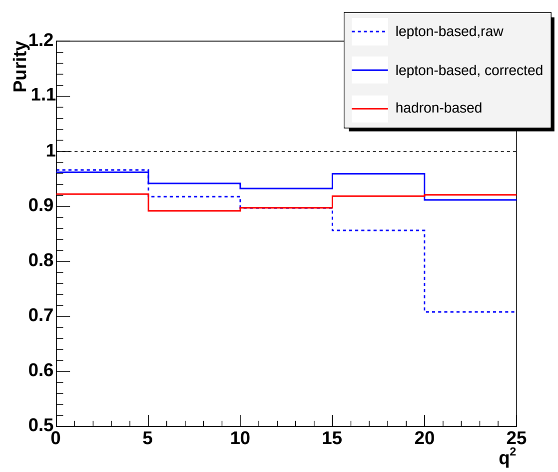
(a)



(b)

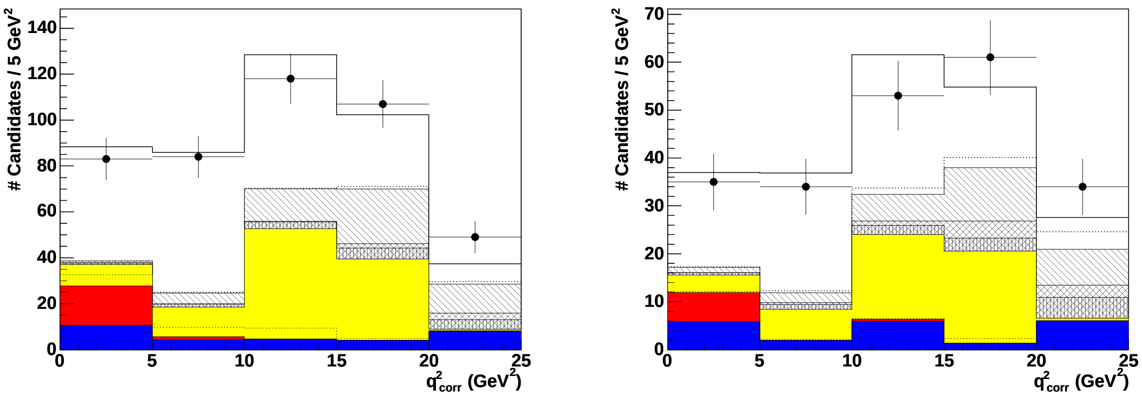


(c)



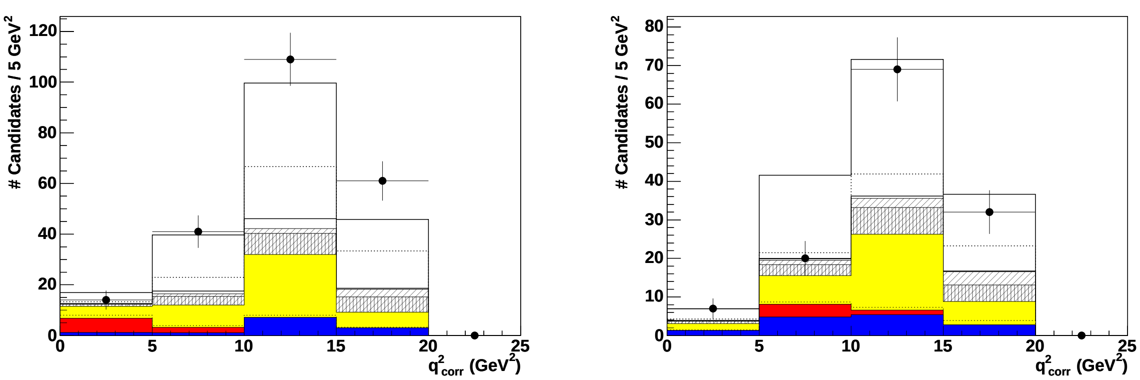
(d)

Figure 7.6: For pure signal $B^0 \to \pi^- \ell^+ \nu$ MC candidates, with the nominal q^2 binning, the stability in the fit plane and signal region (a,b) and purity in the fit plane and signal region (c,d) of the three q^2 calculations ($q^2_{\rm raw}$, $q^2_{\rm corr}$, and $q^2_{\rm had}$) with respect to the true q^2.



(a) (b)

Figure 7.7: q^2 spectra in the signal region for the $B^0 \to \pi^- \ell^+ \nu$ (a) and $B^+ \to \pi^0 \ell^+ \nu$ (b) channels, after the two-channel fit.



(a) (b)

Figure 7.8: q^2 spectra in the signal region for the $B^0 \to \rho^- \ell^+ \nu$ (a) and $B^+ \to \rho^0 \ell^+ \nu$ (b) channels, after the full four-channel fit.

Chapter 8

Systematic Errors

8.1 Stability with respect to critical cuts

As discussed in Chapter 4, the analysis depends on a set of background suppression cuts, and a subset of these cuts have a particularly strong impact on either or both signal or background. In order to test the stability of the fit results, each of the selection cuts is varied over a range of values, and the fit repeated for each case.

Figures 8.1 and 8.2 illustrate, and Table 8.1 summarizes, the impact of these variations of the critical selection variables on the fit scale parameter for $B \to \pi\ell\nu$. These figures in general show a wider scan range than that used to determine the corresponding spread. We do not consider a cut at $\Delta Q = 0$, for instance, since this requirement is known to be both too restrictive and too sensitive to the differences in tracking between data and Monte Carlo simulation. Likewise, it is inappropriate to consider a minimum cut on p_{miss} larger than 1.2 GeV, since we cut too deeply into the signal sample, and equally inappropriate to consider too loose a cut on θ_{miss}, since removing the cut in the forward direction includes a region of the EMC known to be poorly calibrated. We also cannot consider a cut on p_l^* of less than 1.3 GeV, since that is the minimum lepton momentum cut applied by the preselection. Finally, the requirement $|cosBY| \leq 1.1$ separates those candidates which are kinematically consistent with a $B \to \pi\ell\nu$ decay from those which are not by requiring $|\cos\theta_{\mathrm{BY}}|$ to be consistent, within experimental resolution, with a true cosine. Demanding that

Table 8.1: Studies of stability for the $B \to \pi\ell\nu$ scale parameter p_1^1 with respect to the variation of critical cuts.

Nominal Cut	Range of variation	Spread relative to nominal
$\|m_{miss}^2/2E_{miss}\| < 0.40\,\mathrm{GeV}$	$0.3-0.5\,\mathrm{GeV}$	+2.2% - 5.6%
$\|\Delta Q\| < 2$	$1-4$	+2.2%
$p_{miss} > 0.7\,\mathrm{GeV}$	$0.5-1.0\,\mathrm{GeV}$	+3.4%
$\theta_{miss} > 0.6\,\mathrm{rad}$	$0.4-0.6\,\mathrm{rad}$	+5.6%
$L2 + 14.28\cos\theta_{thrust} \leq 12.00$	$10-16$	+16%
$R2$	0.3-0.6 GeV	±1.1%
$\|\cos\theta_{\mathrm{BY}}\| < 1.1$	$1.1-1.3$	±0.55%
$p_l^* > 1.3\,\mathrm{GeV}$	$1.3-1.6$	$< 0.5\%$
$p_h^* + p_l^* > 2.9\,\mathrm{GeV}$	$2.75-3.1\,\mathrm{GeV}$	±5.6%

$|\cos\theta_{\mathrm{BY}}|$ be less than 1.0 moves from requiring that $|\cos\theta_{\mathrm{BY}}|$ be physical to relying on the detailed behavior of the $\cos\theta_{\mathrm{BY}}$ distributions for signal and background within this range; a tighter cut on $\cos\theta_{\mathrm{BY}}$ should therefore not be considered.

With these considerations in mind, we find that the largest instabilities are in the variations of $m_{miss}^2/2E_{miss}$, $L2+14.28\cos\theta_{thrust}$, and $p_l^*+p_h^*$. While the control sample distributions in $L2$, $\cos\theta_{thrust}$, p_l^* and p_h^* shown in Chapter 5 do show some deviations, they are small compared the corresponding discrepancies in the signal channels, where the background contributions dominate. The trend of the discrepancies in the signal channels match the trends in the fit; there are excesses in data at large $L2$, large $\cos\theta_{thrust}$, and small p_h^*, and the fit shows a trend towards larger $\mathcal{B}(B \to \pi\ell\nu)$ results when the nominal cuts are loosened to include these regions. As discussed in Chapter 5, this discrepancy is due to continuum background, and may either come from QED processes not included in the simulation, or from some problem in the $q\bar{q}$ simulation.

The fit results are sufficiently stable with respect to the cuts in ΔQ, p_{miss}, $R2$, and $\cos\theta_{\mathrm{BY}}$that we do not consider these as sources of systematic error. Since the cut on $p_l^* + p_h^*$ exploits the specific kinematics of the $B \to \pi\ell\nu$ decay, as determined by the $B \to \pi\ell\nu$ form factor, any systematic uncertainty we might assign would be completely correlated with the systematic error determined by varying the $B \to \pi\ell\nu$ form factor model, and thus is already taken into account. In fact, the $5-6\%$ variation in the effect of the $p_l^* + p_h^*$ cut will turn out to be comparable to the spread in results obtained by varying the $B \to \pi\ell\nu$ form factor model.

8.2 Uncertainties in Detector Performance and Reconstruction

8.2.1 Neutrino Reconstruction

The dependence of the fit results on the primary neutrino quality cut on $m_{miss}^2/2E_{miss}$ (see Section 8.1) is assumed to be the product of a number of different detector and reconstruction effects. Uncertainties in the charged and neutral particle reconstruction efficiencies, as well tracks and photons from beam background, fake tracks (ghosts), failures in matching neutral clusters to charged tracks, and showers from hadronic splitoffs all contribute, as will any uncertainties in the simulation of processes that affect the total rate of undetectable neutral particles (K_L^0, ν, neutrons). These effects not only contribute to the dependence of the fit results on the value of the primary neutrino quality cut, but also contribute to a bias due to uncertainties in the impact of the nominal $m_{miss}^2/2E_{miss}$ cut on signal efficiency.

The difference between the efficiency of the $m_{miss}^2/2E_{miss}$ cut in data and simulation, for the $B \to D^*\ell\nu$ control samples, is summarized in Table 8.2. As it is known that the simulation does not perfectly reflect the reconstruction efficiencies for both charged tracks and neutral showers, the following recipe, based on [36], is used to bound the impact of these differences:

- For MC tracks with $p_t > 0.2\,\text{GeV}$, eliminate tracks at random with a probability of 1%.
- For MC tracks with $p_t < 0.2\,\text{GeV}$, eliminate tracks at random with a probability of 1.6%.

Photons are also removed at random with a probability of 2.5%, independent of their energy.

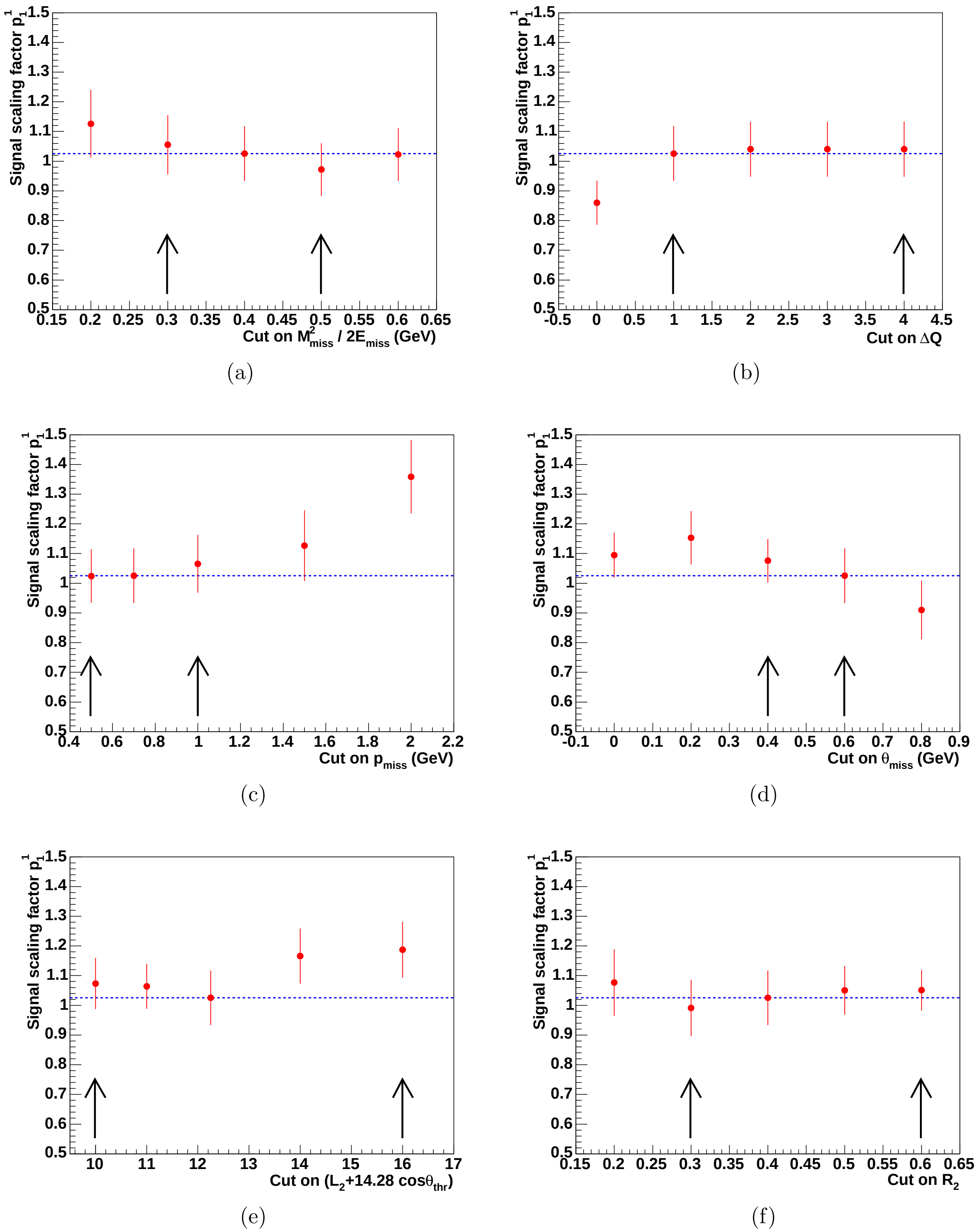


Figure 8.1: Fit results for the scale parameter for the $B \to \pi\ell\nu$ component, p_1^1, as a function of cut value for cuts on $m^2_{miss}/2E_{miss}$ (a) ΔQ (b) p_{miss} (c) θ_{miss} (d) $L2 + 14.28\cos\theta_{thrust}$ (e) $R2$ (f). The nominal fit value is always shown as a blue line, while the range of variation used in assigning systematics is indicated by the black arrows.

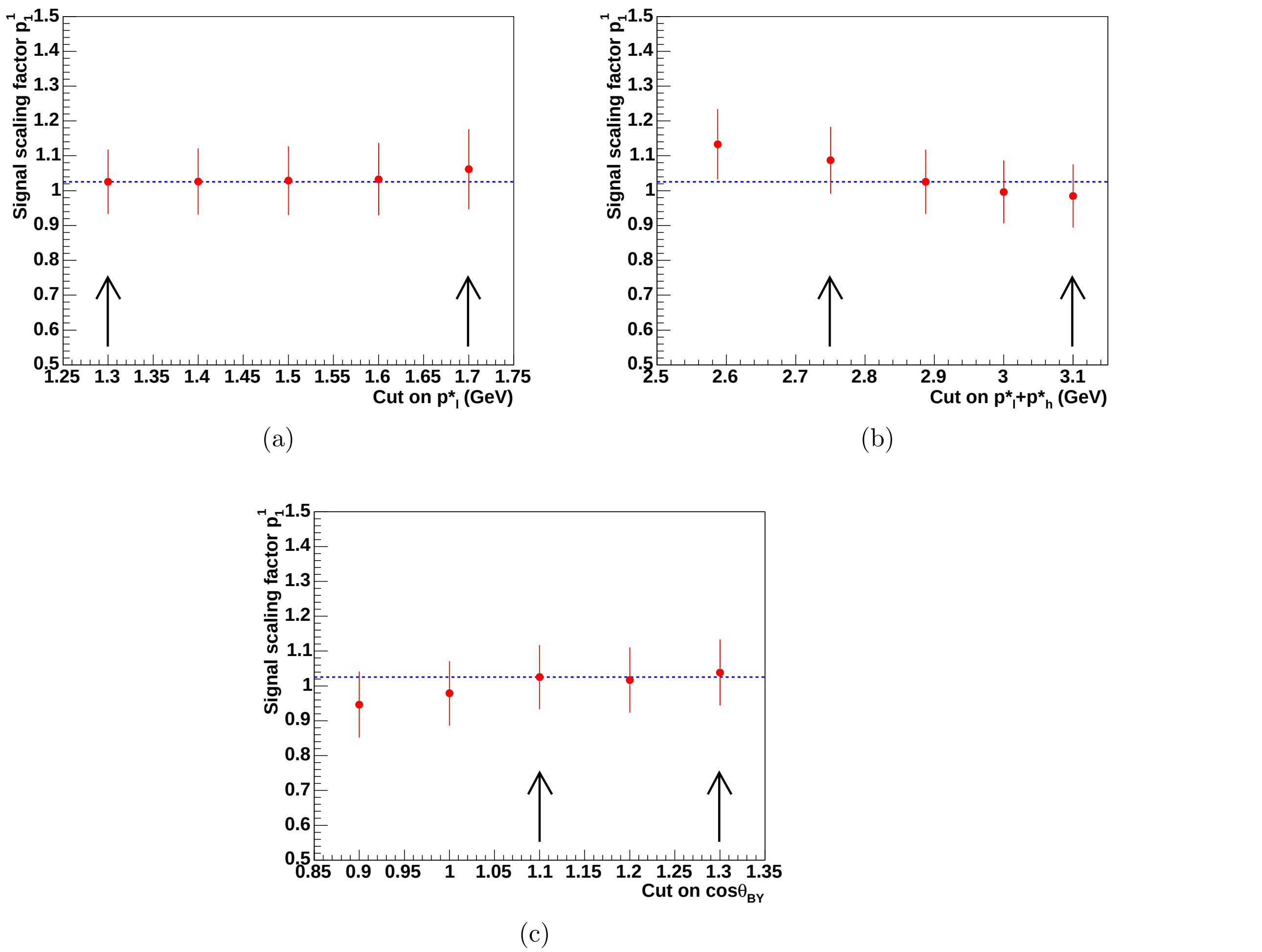


Figure 8.2: Fit results for the scale parameter for the $B \to \pi \ell \nu$ component, p_1^1, as a function of cut value for cuts on p_l^* (a) $p_l^* + p_h^*$ (b) $\cos\theta_{\rm BY}$(c). The nominal fit value is always shown as a blue line, while the range of variation used in assigning systematics is indicated by the black arrows.

While it is possible to remove a certain percentage of tracks or showers from the event to simulate a reduced reconstruction efficiency, it is not quite as easy to embed additional charged tracks or neutral showers in the simulated event to simulate additional tracks or neutrals from such effects as ghost tracks, hadronic splitoffs, or beam background. To gain some sense of what such an effect might entail, we repeat the studies discussed above, with the same killing procedure performed on the data rather than the simulation.

Table 8.2: Data/MC ratios for the control samples $B \to D^*\ell\nu$, with $D \to K\pi$ and $D \to K\pi\pi^0$, for the efficiency of the cut $m^2_{miss}/2E_{miss} < 0.4\,\mathrm{GeV}$.

	$B^0 \to D^{*-}\ell^+\nu$ $D^0 \to K^+\pi^-$	$B^0 \to D^{*-}\ell^+\nu$ $D^0 \to K^+\pi^-\pi^0$	Average
Nominal	1.08 ± 0.04	1.02 ± 0.03	1.05
reduced MC eff	1.15 ± 0.04	1.09 ± 0.03	1.12
reduced data eff.	1.03 ± 0.04	0.94 ± 0.03	0.99

We see that the efficiency of the m^2_{miss}/E_{miss} cut is consistently more efficient in data than in simulation; reducing the overall track and neutral reconstruction efficiency in simulation enhances this effect, while reducing the track and neutral reconstruction efficiency in data reduces it. We take these results to mean that there is an average 5% bias in our overall selection efficiency, with an uncertainty of $\pm 7\%$ on that bias given by the track and neutral killing exercise and additional uncertainty of $\pm 3\%$ given by the channel-to-channel variation in the control sample. We will correct for the efficiency bias in the final results, and take the residual associated uncertainty to be $\pm 7.6\%$.

The variations of the fit results with respect to cuts on both p_{miss} and θ_{miss} (discussed in Section 8.1) are also correlated with uncertainties in both the neutrino reconstruction and the continuum simulation. At this time, however, we cannot be certain that they are completely correlated with the other effects. We take this into account with an additional $\pm 5\%$ uncertainty in the neutrino reconstruction, yielding an overall systematic error due to the neutrino reconstruction of $\pm 9.1\%$.

8.2.2 Lepton Identification Efficiency

The particle identification efficiencies in Monte Carlo are adjusted to match those in data using a reweighting procedure described in Chapter 4. The particle identification efficiencies for electrons and muons are derived from a set of electron and muon control samples in data, and for a corresponding set of Monte-Carlo truth-based control samples for Monte Carlo simulation. The efficiencies are recorded in binned tables as a function of laboratory momentum and polar angle and lepton charge; each Monte Carlo candidate is weighted by the ratio of the appropriate data and Monte Carlo efficiency table entries.

The electron and muon control samples in data are primarily Bhabha scattering events and $\mu^+\mu^-$ production; this means that there is a bias in the efficiency tables resulting from the fact that the control sample events have a lower average track and photon multiplicity than the average multihadron event. We consider the scan of the cut in p_l^*, which shows no significant variation in p_1^1 for cuts of $p_l^* > 1.3\,\mathrm{GeV}$ to $p_l^* > 1.7\,\mathrm{GeV}$, as an indication that the effect is likely small for both the electron and muon channels. We also test the impact of this effect by applying a correction [37] for the overestimate of the efficiency for electrons and repeating the electrons-only fit. The resulting change in the fit result is about 1.2%. While we expect that the difference between muon efficiencies for control sample and multihadron events is of the same order, our knowledge of the performance of the muon system suggests that overall uncertainty in the muon identification efficiency is on the order of 4%. We take the average of these two uncertainties, and assign a systematic of $\pm 2.6\%$.

8.3 Sensitivity to binning

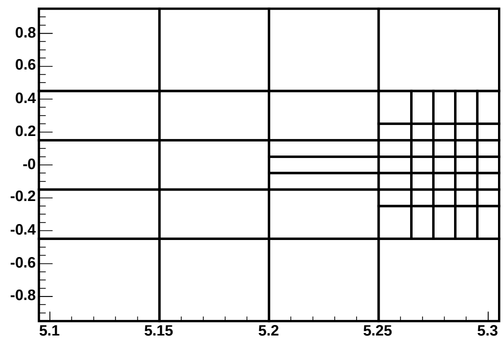
The standard fit described is a binned maximum likelihood fit, where the binning of the ΔE-m_{ES} fit region is chosen to be fine in the signal region and coarse in the region dominated by background. This maximizes sensitivity to the shape differences between signal and background components. In principle, for large statistics, the fit result should be insensitive to the choice of binning, but it is necessary to determine whether this is actually the case.

As the binning, or tiling, of the ΔE-$m_{\rm ES}$ region is irregular, we have chosen to use a set of four different tilings in the ΔE-$m_{\rm ES}$ plane. The known discrepancies between data and Monte Carlo simulation in terms of ΔE and $m_{\rm ES}$ provide some guidance as to nature of these tilings. In Chapter 5, we saw that for the control samples, the data demands a broader resolution in both ΔE and $m_{\rm ES}$ than that predicted by the simulation, an effect to which we are particularly sensitive near the endpoint of the $m_{\rm ES}$ distribution, and a larger population in the correlated ΔE-$m_{\rm ES}$ tail of the signal distribution relative to the core in the signal region. As a result, we have chosen the set of tilings illustrated in Figure 8.3, which combine coarse and fine binning variations for both signal and sideband.

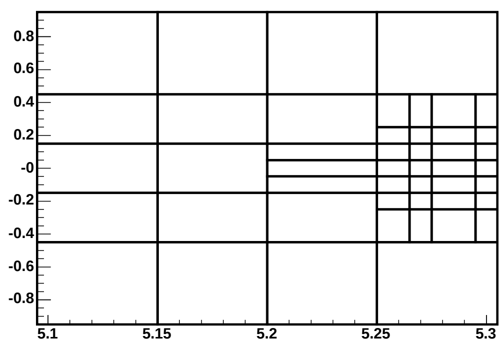
The two-channel fit is performed for each of these binnings, and the results summarized in the first section of Table 8.3.

The fits for tilings with fine signal bins produce values for all free parameters that are consistent within a few percent with the nominal fit results and with each other. The fits using tilings with extremely coarse signal bins produce parameter values consistent with each other within a few percent, but differ by about 6% from the nominal fit results. While all of these deviations are well within the 9% statistical error on the nominal fit value, the trend, and the relative size of the deviations, is consistent even for variations of the fit with more free parameters, such as that shown in the second section of Table 8.3 and discussed in Section 8.4. The fact that the fits with different tilings using the coarse signal bins are consistent with each other, despite their deviation from the other fit results, suggests that the fit is relatively insensitive to any variation in binning outside the signal region, and is most sensitive to significant changes in binning where it is most sensitive to the modelling of signal fit variable resolutions.

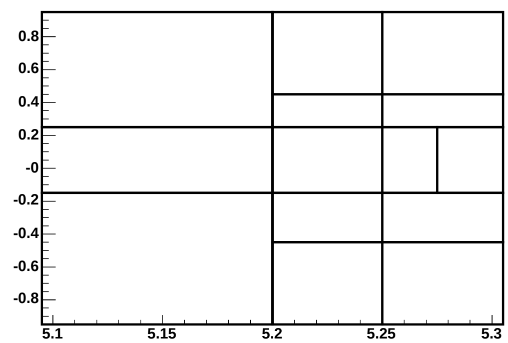
We assign a binning systematic of $\pm 3.5\%$.



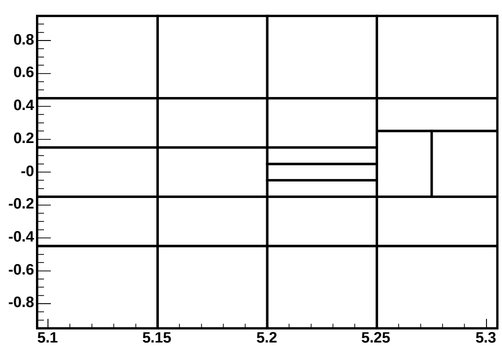
(a) Nominal



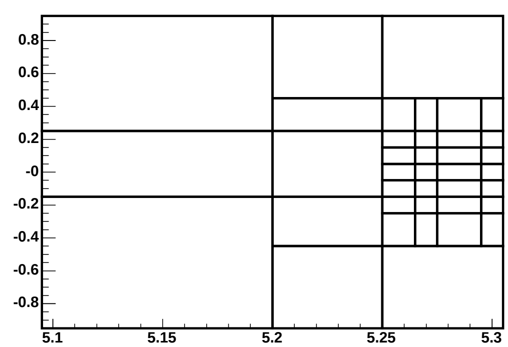
(b) Binning 1



(c) Binning 2



(d) Binning 3



(e) Binning 4

Figure 8.3: Binning of the ΔE-m_{ES} region used in the fit, illustrated for the nominal binning (a), a similar but slightly coarser binning in m_{ES} (b), a coarse binning (c), and coarsened binning confined to the signal region and sideband regions respectively (d,e).

8.4 Channel-by-channel variation of background parameters

One of the assumptions of the nominal fit is that the relative normalization of a common physics background can be shared across channels. We test the accuracy of this assumption, and the size of channel-by-channel efficiency effects, by floating the $b \to c\ell\nu$ and $q\bar{q}$ normalization parameters separately for the $B^0 \to \pi^-\ell^+\nu$ and $B^+ \to \pi^0\ell^+\nu$ channels. The results, given in Table 8.3, show a 10% percent downward shift in the $B \to \pi\ell\nu$ scale parameter (p_1^1) relative to the nominal fit value. This is matched by corresponding upward shift in the $q\bar{q}$ scale parameter for the $B^0 \to \pi^-\ell^+\nu$ channel, while the $q\bar{q}$ scale parameter for the $B^+ \to \pi^0\ell^+\nu$ channel, and the $b \to c\ell\nu$ scale parameters for both channels, are relatively stable. The upward shift in the $q\bar{q}$ scale parameter for the $B^0 \to \pi^-\ell^+\nu$ channel is significant (more than 2σ). We believe the fact that it is confined to the $B^0 \to \pi^-\ell^+\nu$ channel is symptomatic of a correlation with the observed resolution effect in that channel. Since the Monte Carlo simulation predicts a narrower resolution than that seen in the data, it also predicts a smaller population in the tail of the $B^0 \to \pi^-\ell^+\nu$ ΔE-m_{ES} distribution than required by the data. The fit uses the component which best approximates the shape of the $B \to \pi\ell\nu$ signal tail—the $q\bar{q}$ background—to compensate.

We do not consider this effect as an independent systematic uncertainty, since the impact of the broadened ΔE and m_{ES} resolution should be taken into account by the studies used to determine the binning systematic error.

8.5 Model Systematics

8.5.1 Isospin constraints, f_{+-}/f_{00}, and B lifetime

The nominal fit imposes the isospin relations given in Equations 6.12. The fit also assumes a B^0 to B^+ production ratio, $f_{+-}/f_{00} = 1$, and adjusts the relative branching ratios for conjugate B^0 and B^+ decays to account for the difference in the measured B lifetimes, $\tau_{B^0}/\tau_{B^+} = 1.086 \pm 0.017$ [2].

The isospin relations are not exact. Table 8.3 gives the result of a fit where sources are split based on the flavor of the B decay, and the associated parameters are floated separately. $p_1^1 = \frac{\mathcal{B}(B^0 \to \pi^- \ell^+ \nu)_{fit}}{\mathcal{B}(B^0 \to \pi^- \ell^+ \nu)_{MC}}$ and $p_2^2 = \frac{\mathcal{B}(B^+ \to \pi^0 \ell^+ \nu)_{fit}}{\mathcal{B}(B^+ \to \pi^0 \ell^+ \nu)_{MC}}$ are consistent within 1 sigma, as are the scale parameters for the $B^0 \to c\ell\nu$ and $B^+ \to c\ell\nu$ backgrounds. Thus the fit gives no experimental evidence for isospin breaking in the $B \to \pi\ell\nu$ channels, and no evidence for an f_{+-}/f_{00} significantly different from one.

Isospin breaking in the $B \to \pi\ell\nu$ channels is expected to be very small. On the other hand, the effect of isospin breaking in $B \to \rho\ell\nu$ is expected to be larger due to the effect of $\rho - \omega$ mixing. To estimate the size of this effect, we adjust the fixed scale parameters for the $B \to \rho\ell\nu$ sources to give $\frac{\mathcal{B}(B^0 \to \rho^- \ell^+ \nu)_{MC}}{\mathcal{B}(B^+ \to \rho^0 \ell^+ \nu)_{MC}} = 1.43$ (as opposed to the default value of 2.0) as suggested in [38] and [39]. The deviation between this value and the nominal fit result, given in Table 8.3, is less than 0.1%.

The most recent measurement gives $f_{+-}/f_{00} = 1.044 \pm 0.050$ [32] (which is still consistent with the assumptions in the nominal fit). We evaluate the related systematic by repeating the fit for the central value, $f_{+-}/f_{00} = 1.044$, and for the central value plus or minus the given error. Again, the deviation is less than 0.1%.

We assess the systematic due to the uncertainty in the B lifetime by varying the lifetime ratio within errors and adjusting the relative branching ratios for the B^0 and B^+ decays (for both signal and background components) to match. The deviation is also less than 0.1%.

8.5.2 $b \to u\ell\nu$ background composition and rate.

The $b \to u\ell\nu$ contribution is simulated by combining the $B \to \pi\ell\nu$, $B \to \rho\ell\nu$, $B^+ \to \omega\ell^+\nu$, $B^+ \to \eta\ell^+\nu$, and $B^+ \to \eta'\ell^+\nu$ resonances, and then adding non-resonant $b \to u\ell\nu$ decays in order to make up the total $b \to u\ell\nu$ branching fraction. The background from $B \to \rho\ell\nu$ decays has a separate scale parameter; the remainder of these decays, excluding $B \to \pi\ell\nu$, are considered part of a residual $b \to u\ell\nu$ background with a common scale parameter. Both of these scale parameters are fixed in the fit. The error on the $b \to u\ell\nu$ background composition, inclusive of $B \to \rho\ell\nu$, is dominated by the relative uncertainty in the resonant branching ratios, and the size of the non-resonant component.

We estimate the total error on the $b \to u\ell\nu$ background composition (including the uncertainty created by fixing $\mathcal{B}(B \to \rho\ell\nu)$) by repeating the fit for the following ranges of branching ratios:

- $\mathcal{B}(B^0 \to \rho^-\ell^+\nu) = 1.92 - 3.43 \times 10^{-4}$
- $\mathcal{B}(B^+ \to \eta\ell^+\nu) = 0.48 - 1.20 \times 10^{-4}$

where we vary $\mathcal{B}(B^+ \to \rho^0\ell^+\nu)$ and $\mathcal{B}(B^+ \to \omega\ell^+\nu)$ simultaneously with $\mathcal{B}(B^0 \to \rho^-\ell^+\nu)$, and $\mathcal{B}(B^+ \to \eta'\ell^+\nu)$ simultaneously with $\mathcal{B}(B^+ \to \eta\ell^+\nu)$, with their relative ratios set by isospin relations and τ_{B^0}/τ_{B^+}. The non-resonant component is always re-adjusted to reflect the branching ratios used. The spread in fit results is summarized in Table 8.6; we assign an associated systematic of $\pm 3.5\%$.

We also consider the additional error introduced by the uncertainty in the total $b \to u\ell\nu$ rate, $(2.24 \pm 0.27(stat.) \pm 0.26(syst.) \pm 0.39(theo)) \times 10^{-3}$ [40]; we vary the total $b \to u\ell\nu$ rate within these errors by increasing and decreasing the non-resonant $b \to u\ell\nu$ contribution, and then repeat the fit in each case. These results are also summarized in Table 8.6. We assign an additional systematic error of $\pm 2\%$.

8.5.3 $b \to c\ell\nu$ modelling and composition

We weight the $b \to c\ell\nu$ background candidates to reflect the $B \to D^*\ell\nu$ form-factor parameters measured by *BABAR* in [27], $R1 = 1.328 \pm 0.055 \pm 0.025 \pm 0.025$, $R2 = 0.920 \pm 0.044 \pm 0.020 \pm 0.013$, and $\rho^2 = 0.769 \pm 0.039 \pm 0.019$. We also adjust the average branching ratios for $B \to D\ell\nu$, $B \to D^*\ell\nu$, and $B \to D^{**}\ell\nu$ components to correct anomalies in the lepton momentum of the $B\overline{B}$ background for all reconstructed channels. We evaluate the effect of uncertainties in the form-factor parameters used in the simulation of the $B \to D^*\ell\nu$ component of the $b \to c\ell\nu$ background by repeating the fit with $\pm 1\sigma$ variation in each of the three $B \to D^*\ell\nu$ form-factor parameters; we also vary the composition of the $b \to c\ell\nu$ background by individually varying the branching ratios of the $B \to D\ell\nu$, $B \to D^*\ell\nu$, and $B \to D^{**}\ell\nu$ components within the ranges given in Table 8.4. In each case, the non-resonant component is set so that the total $b \to c\ell\nu$ branching ratio remains at the nominal value of 10.62%. The relative branching ratios for the B^0 and B^+ decay channels have been scaled from the average in the ratio of the B lifetimes.

Results are summarized in Table 8.5; we see about a 2% total spread, and we assign a systematic of 1%.

8.5.4 Modelling of continuum background

As discussed earlier, it was necessary to use an "average" smoothed shape for the continuum background in all channels; it is also clear that the error in this average shape has a significant impact on the fit. We estimate the associated systematic by deriving smoothed shape distributions for the $q\bar{q}$ contribution for three different samples:

- For the $B \to \pi\ell\nu$ channels only, e and μ channels combined.
- For the $B \to \pi\ell\nu$ and $B \to \rho\ell\nu$ channels combined, e only.
- For the $B \to \pi\ell\nu$ and $B \to \rho\ell\nu$ channels combined, μ only.

We repeat the fit using each of these distributions in place of the distribution usually used to estimate the $q\bar{q}$ component. In addition, since the smoothed distributions are derived for a looser values of the usual continuum suppression cut in $L2$ and $\cos\theta_{thrust}$, we take a set of smoothed distributions derived for different values of this continuum suppression cut and repeat the fit using these distributions in place of the nominal $q\bar{q}$ component. Finally, we repeat the nominal fit using the unsmoothed $q\bar{q}$ distribution.

The results are summarized in Table 8.5. The large variation in the fit result with respect to the cut in $L2 + 14.28\cos\theta_{thrust}$, discussed in Section 8.1, represents a systematic error due to our incomplete knowledge of the composition of the continuum background, and is taken into account with an error of +16%. This error can be reasonably expected to dominate the variation in fit results seen when smoothed backgrounds are used that are derived at different levels of this cut. We also, however, see a significant effect if we use separately derived smoothed shapes to describe the $q\bar{q}$ background for the $B \to \pi e\nu$ and $B \to \pi\mu\nu$ channels, rather than using a single shape for both. We take an additional systematic of $\pm 5\%$ to account for this variation.

8.5.5 Signal form-factor models

An additional error in measurement of the $B \to \pi\ell\nu$ branching ratio comes from the modelling of the $B \to \pi$ and $B \to \rho$ form-factors.

In the case of the $B \to \pi$ form-factor, it was established in Section 1.5.2.2 that parametrized fits to both the LCSR calculations performed by Patricia Ball and the lattice results for the JLQCD, FNAL01, and UKQCD99 collaborations give consistent results. Thus, we consider only two form-factor predictions—the ISGW2 model, which is used for the nominal fit, and the extended parametrization of the Ball LCSR calculation, which covers the entire q^2 range. We perform the fit for three variants of the Ball parametrization—the nominal values, and a pair of extreme values given by the errors quoted in Ref. [1]. The results are summarized in Table 8.6.

In all cases, the same sample of ISGW2-based $B \to \pi\ell\nu$ signal Monte-Carlo is used, and is reweighted candidate-by-candidate to reflect the new form-factor model. With the errors on the Ball parametrization taken into account, the fit gives results for the LCSR form-factor model that are consistent with those taken from ISGW2 (see Table 8.6). The spread in results between different points for the Ball parametrization is a couple of percent; the significant difference is the 5% downward shift in the fit result for the Ball parametrization relative to ISGW2. We assign an error on the $\mathcal{B}(B \to \pi\ell\nu)$ branching ratio of $\pm 5.2\%$.

Variation of the $B \to \rho\ell\nu$ form-factor model is expected to have a more modest impact, since it only contributes through the $B \to \rho\ell\nu$ background in the $B \to \pi\ell\nu$ channels. At this time, it is also difficult to do, since the current LCSR calculation does not extend above $q^2 > 15 - 17\,\mathrm{GeV}^2$. An updated calculation with an extended parametrization over the full range $0 < q^2 < 20\,\mathrm{GeV}^2$ is expected within the next few months [41], but for the moment, we choose not to evaluate this error.

8.6 B Counting

The number of $B\overline{B}$ events in data was obtained by subtracting the luminosity scaled number of multihadron events in on- and off-resonance data samples, corrected for the $B\overline{B}$ efficiency. This number has an uncertainty of $\pm 1.6\%$, and that uncertainty translates directly into an error of $\pm 1.6\%$ on the branching ratio.

8.7 Summary of Systematic Errors

Table 8.7 summarizes the various contributions to the total systematic uncertainty. We combine the detector, fit stability, and background-model systematics in quadrature, but keep separate the contribution from the $B \to \pi$ form factor. This gives us a total experimental systematic error of $^{+20}_{-12}\%$, and a theoretical error due to the $B \to \pi\ell\nu$ form factor variation of $\pm 5.2\%$.

	Parameter					
Fit Type (channel)	p_1^1 ($B \to \pi\ell\nu$)	p_8^1 ($b \to c\ell\nu$) ($B^0 \to \pi^-\ell^+\nu$)	p_8^2 ($b \to c\ell\nu$) ($B^+ \to \pi^0\ell^+\nu$)	p_{14}^1 ($q\bar{q}$) ($B^0 \to \pi^-\ell^+\nu$)	p_{14}^2 ($q\bar{q}$) ($B^+ \to \pi^0\ell^+\nu$)	χ^2/dof
Nominal	1.03 ± 0.09	0.99 ± 0.03		1.09 ± 0.10		387/273
Binning 1	1.06 ± 0.09	1.00 ± 0.03		1.05 ± 0.10		359/237
Binning 2	1.01 ± 0.10	1.01 ± 0.03		1.07 ± 0.11		124/81
Binning 3	1.01 ± 0.08	1.00 ± 0.03		1.09 ± 0.06		211/135
Binning 4	1.06 ± 0.09	1.01 ± 0.03		1.02 ± 0.12		273/183
Channel-sep.	0.96 ± 0.09	0.94 ± 0.04	0.98 ± 0.07	1.61 ± 0.20	0.94 ± 0.13	335/235

Isospin, f_{+-}/f_{00}, B lifetime						
	Parameter					
Fit Type	p_1^1 ($B^0 \to \pi^-\ell^+\nu$)	p_2^2 ($B^+ \to \pi^0\ell^+\nu$)	p_8^1 ($B^0 \to c\ell\nu$)	p_9^1 ($B^+ \to c\ell\nu$)	p_{14}^1 ($q\bar{q}$)	χ^2/dof
B^0/B^+	1.09 ± 0.10	0.85 ± 0.15	0.81 ± 0.42	1.10 ± 0.26	1.12 ± 0.12	383/271
$\tau_{B^0}/\tau_{B^+} = 1.103$	1.03 ± 0.09		0.96 ± 0.03		1.09 ± 0.12	388/273
$\tau_{B^0}/\tau_{B^+} = 1.069$	1.03 ± 0.09		0.98 ± 0.03		1.09 ± 0.11	387/273
$\rho^0/\rho^+ = 1.43$	1.02 ± 0.09		0.99 ± 0.02		1.10 ± 0.09	385/273
$f_{+-}/f_{00} = 1.044$	1.03 ± 0.09		1.00 ± 0.03		1.09 ± 0.11	387/273
$f_{+-}/f_{00} = 1.094$	1.03 ± 0.09		0.97 ± 0.03		1.08 ± 0.11	387/273
$f_{+-}/f_{00} = 0.994$	1.03 ± 0.09		0.99 ± 0.03		1.09 ± 0.11	387/273

Table 8.3: Studies of stability with respect to binning in ΔE and m_{ES} plane for the nominal two-channel fit, and for the two-channel fit with the relative normalizations of the dominant background components floated by channel. Result of fit with parameters floated separately by B flavor; effect of isospin breaking, uncertainty in f_{+-}/f_{00}, and uncertainty in the B lifetime.

Table 8.4: Branching ratio values and ranges of variation for $b \to c\ell\nu$ background components.

Mode	Average Branching Ratio (%)	Range
$B \to D\ell\nu$	2.10	2.08-2.39
$B \to D^*\ell\nu$	6.16	5.04-7.0
$B \to D^{**}\ell\nu$	1.5	0.27-2.31

Table 8.5: Summary of fit results for variations of the $B \to D^*\ell\nu$ form-factor and the $B \to D\ell\nu$, $B \to D^*\ell\nu$, and $B \to D^{**}\ell\nu$ branching ratios, and for variations of the shape ansatz for the $q\bar{q}$ background source(s).

Background Composition and Shape: Two-Channel Fit				
	Parameter			
FF parameter	p_1^1 $(B \to \pi\ell\nu)$	p_8^1 $(b \to c\ell\nu)$	p_{14}^1 $(q\bar{q})$	χ^2/dof
$R1 + 1\sigma$	1.03 ± 0.093	0.99 ± 0.03	1.09 ± 0.11	387/273
$R1 - 1\sigma$	1.03 ± 0.092	1.00 ± 0.03	1.09 ± 0.11	387/273
$R2 + 1\sigma$	1.03 ± 0.092	0.98 ± 0.03	1.09 ± 0.11	387/273
$R2 - 1\sigma$	1.02 ± 0.092	1.00 ± 0.03	1.09 ± 0.11	388/273
$\rho^2 + 1\sigma$	1.03 ± 0.092	0.98 ± 0.03	1.09 ± 0.11	387/273
$\rho^2 - 1\sigma$	1.03 ± 0.092	1.00 ± 0.03	1.09 ± 0.11	388/273
Branching ratio	p_1^1 $(B \to \pi\ell\nu)$	p_8^1 $(b \to c\ell\nu)$	p_{14}^1 $(q\bar{q})$	χ^2/dof
$B(B \to D\ell\nu) + 1\sigma$	1.02 ± 0.09	0.99 ± 0.03	1.09 ± 0.11	388/273
$B(B \to D\ell\nu) - 1\sigma$	1.03 ± 0.09	1.00 ± 0.03	1.09 ± 0.11	386/273
$B(B \to D^*\ell\nu) + 1\sigma$	1.03 ± 0.09	0.96 ± 0.03	1.10 ± 0.11	385/273
$B(B \to D^*\ell\nu) - 1\sigma$	1.01 ± 0.09	1.16 ± 0.04	1.08 ± 0.11	393/273
$B(B \to D^{**}\ell\nu) + 1\sigma$	1.02 ± 0.10	0.95 ± 0.03	1.11 ± 0.11	378/273
$B(B \to D^{**}\ell\nu) + 1\sigma$	1.03 ± 0.09	1.00 ± 0.03	1.07 ± 0.07	398/273
$q\bar{q}$ taken from combined $B \to \pi\ell\nu$ and $B \to \rho\ell\nu$ channels				
$cut(L2, \cos\theta_{thrust}) < 16$	1.02 ± 0.07	1.01 ± 0.03	1.03 ± 0.06	385/273
$cut(L2, \cos\theta_{thrust}) < 14$	1.03 ± 0.09	0.99 ± 0.03	1.09 ± 0.10	387/273
$cut(L2, \cos\theta_{thrust}) < 12$	1.08 ± 0.09	0.98 ± 0.03	1.11 ± 0.11	390/273
e only ($cut(L2, \cos\theta_{thrust}) < 12$)	1.04 ± 0.15	1.09 ± 0.04	0.67 ± 0.26	345/257
μ only ($cut(L2, \cos\theta_{thrust}) < 12$)	1.13 ± 0.14	1.09 ± 0.05	0.72 ± 0.09	384/258
unsmoothed	0.98 ± 0.10	0.92 ± 0.04	1.39 ± 0.16	458/273
$q\bar{q}$ taken from $B \to \pi\ell\nu$ alone				
$cut(L2, \cos\theta_{thrust}) < 16$	1.07 ± 0.07	0.97 ± 0.03	1.12 ± 0.10	380/273
$cut(L2, \cos\theta_{thrust}) < 14$	1.07 ± 0.09	0.96 ± 0.03	1.18 ± 0.11	381/273
$cut(L2, \cos\theta_{thrust}) < 12$	1.12 ± 0.09	0.94 ± 0.04	1.20 ± 0.12	387/273

Table 8.6: Summary of fit results for variations of the $B \to \pi$ ($B \to \pi\ell\nu$) and $B \to \rho$ ($B \to \rho\ell\nu$) form-factors, and for variations in the $b \to u\ell\nu$ background composition and size.

Form-Factor Variations				
	Parameter			
FF(pi, rho)	p_1^1 ($B \to \pi\ell\nu$)	p_8^1 ($b \to c\ell\nu$)	p_{14}^1 ($q\bar{q}$)	χ^2/dof
ISGW2,ISGW2	1.03 ± 0.09	0.99 ± 0.03	1.09 ± 0.10	387/273
Ball01,ISGW2	0.98 ± 0.09	0.99 ± 0.03	1.06 ± 0.11	389/273
Ball01hi,ISGW2	0.97 ± 0.09	0.99 ± 0.03	1.07 ± 0.11	390/273
Ball01lo,ISGW2	0.99 ± 0.09	0.99 ± 0.03	1.06 ± 0.11	388/273
$b \to u\ell\nu$ Composition and Rate: Two-Channel Fit				
Variation	Parameter			
$\mathcal{B}(B^0 \to \rho^-\ell^+\nu) + 1\sigma$	0.99 ± 0.09	0.95 ± 0.03	1.05 ± 0.10	403/273
$\mathcal{B}(B^0 \to \rho^-\ell^+\nu) - 1\sigma$	1.06 ± 0.09	1.04 ± 0.03	1.13 ± 0.11	377/273
$\mathcal{B}(B^+ \to \eta\ell^+\nu) + 1\sigma$	1.03 ± 0.09	0.99 ± 0.03	1.09 ± 0.10	387/273
$\mathcal{B}(B^+ \to \eta\ell^+\nu) + 1\sigma$	1.02 ± 0.09	0.99 ± 0.03	1.09 ± 0.10	388/273
$\mathcal{B}(b \to u\ell\nu) + 1\sigma$	1.01 ± 0.09	0.97 ± 0.03	1.08 ± 0.11	392/273
$\mathcal{B}(b \to u\ell\nu) - 1\sigma$	1.04 ± 0.09	1.01 ± 0.03	1.11 ± 0.11	383/273

Table 8.7: Summary of systematic errors

Error Contribution	$\delta\mathcal{B}_\pi/\mathcal{B}_\pi$ (%)
Neutrino reconstruction efficiency	± 9.1
Lepton identification efficiency	± 2.6
Binning	± 3.5
B Lifetime	< 0.5
Isospin Breaking	< 0.5
f_{+-}/f_{00}	< 0.5
$b \to u\ell\nu$ composition	± 3.5
$b \to u\ell\nu$ rate	± 2
$b \to c\ell\nu$ FF and composition	± 1
$q\bar{q}$ shape	± 5
$q\bar{q}$ composition	$+16$
B Counting	± 1.6
Total experimental syst.	$+20, -12$
$B \to \pi\ell\nu$ FF variation	± 5.2

Chapter 9

Summary and Outlook

9.1 Extraction of $\mathcal{B}(B^0 \to \pi^- \ell^+ \nu)$

The fit to the ΔE-m_{ES} distribution results in the scale factor

$$p_1^1 = \mathcal{B}(B^0 \to \pi^- \ell^+ \nu)/\mathcal{B}(B^0 \to \pi^- \ell^+ \nu)_{MC} = 1.03 \pm 0.09(stat) \tag{9.1}$$

where the statistical error takes in to account the statistical error of the data and of the various simulated signal and background components. The signal selection efficiency, as derived from Monte Carlo using ISGW2, is implicit in the fit, but we know from the control sample studies in Section 5.3 that the Monte Carlo selection efficiency is underestimated by 5% due to errors in simulating the efficiency of the neutrino reconstruction. Correcting the branching ratio to compensate, and using $\mathcal{B}(B^0 \to \pi^- \ell^+ \nu)_{MC} = 1.33 \times 10^{-4}$, we find

$$\mathcal{B}(B^0 \to \pi^- \ell^+ \nu) = (1.30 \pm 0.11(stat).) \times 10^{-4} \tag{9.2}$$

Taking into account the experimental systematic and form-factor errors of Section 8.7, we obtain

$$\mathcal{B}(B^0 \to \pi^- \ell^+ \nu) = (1.30 \pm 0.11(stat.)^{+0.26}_{-0.16}(syst.) \pm 0.07(FF_{\pi\ell\nu})) \times 10^{-4} \tag{9.3}$$

Given the Monte Carlo selection efficiencies in Tables 4.5 and 4.6, this corresponds to a sample of about 310 ± 31 correctly reconstructed $B^0 \rightarrow \pi^- \ell^+ \nu$ decays and 117 ± 16 correctly reconstructed $B^+ \rightarrow \pi^0 \ell^+ \nu$ decays. At this time, this measurement is clearly limited by systematic rather than statistical uncertainties, with the largest systematic uncertainties coming from detector simulation and the simulation of continuum backgrounds, rather than the uncertain knowledge of the $B \rightarrow \pi$ form-factor.

The fit assumes a specific q^2 dependence for the $B \rightarrow \pi \ell \nu$ signal decays, imposed by the assumed form factor; we have performed the fit for both the quark-model based ISGW2, and for variations of a parametrization based on both LCSR QCD and Lattice QCD calculations. Figure 9.1 compares the measured q^2 spectrum in data with the full Monte Carlo prediction for both the ISGW2 model and the nominal LCSR parametrization; Figure 9.2 shows the background-subtracted, efficiency-corrected spectra, normalized to the total $B \rightarrow \pi \ell \nu$ branching ratio. The measured q^2 distributions do not significantly depend on the form factor used to determine the selection efficiency. The efficiency-corrected distributions are consistent with both form factor calculations.

The measured branching ratio is consistent with the most recent CLEO result of $1.33 \pm 0.18(stat.) \pm 0.11(syst.) \pm 0.01(FF_{\pi\ell\nu}) \pm 0.07(FF_{\rho\ell\nu}) \times 10^{-4}$ [42]. As expected, we achieve a substantially smaller statistical error (an 8.5% relative error as opposed to CLEO's 13%); our experimental systematic error is at present substantially larger (a 23% relative error as opposed to CLEO's 8%).

9.2 Extraction of $|V_{ub}|$

We use the branching fraction derived in Section 9.1 and extract $|V_{ub}|$ according to two different form-factor model assumptions: the ISGW2 quark model and the extended parametrization of the 2001 Ball LCSR calculations.

The relation between the branching fraction $\mathcal{B}(B \rightarrow \pi \ell \nu)$ and $|V_{ub}|$ is given in Section 1.6:

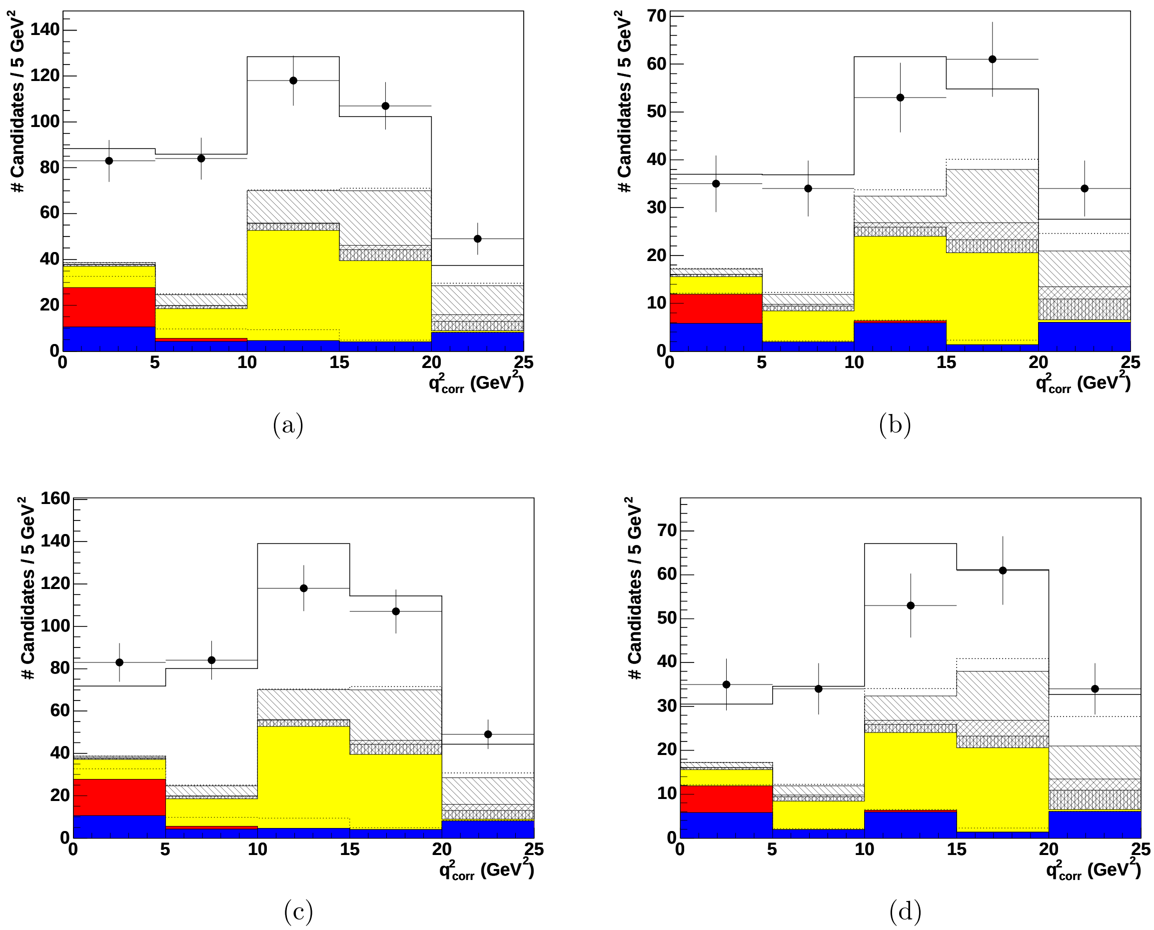
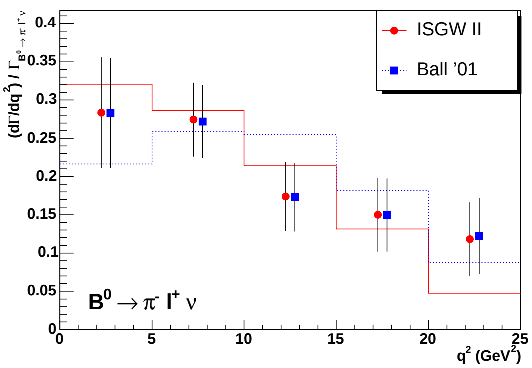
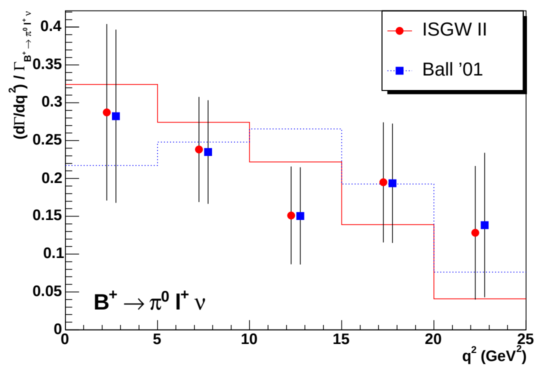


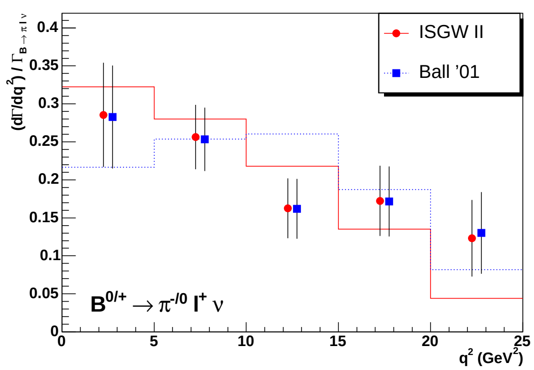
Figure 9.1: q^2 spectra in the signal region resulting from the two-channel fit. Shown are $B^0 \to \pi^- \ell^+ \nu$ with ISGW2 (a) $B^+ \to \pi^0 \ell^+ \nu$ with ISGW2 (b) $B^0 \to \pi^- \ell^+ \nu$ with Ball01 (c) and $B^0 \to \pi^- \ell^+ \nu$ with Ball01 (d). Errors shown are the statistical errors for data only.



(a)



(b)



(c)

Figure 9.2: Background-subtracted q^2 spectra in the signal region for $B^0 \to \pi^- \ell^+ \nu$ (a) and $B^+ \to \pi^0 \ell^+ \nu$ (b) resulting from the two-channel fit, normalized to the full $\mathcal{B}(B^0 \to \pi^- \ell^+ \nu)$ or $\mathcal{B}(B^+ \to \pi^0 \ell^+ \nu)$ branching ratio. Figure (c) shows the combined results for both channels. Data points in red are the results of the fit using the ISGW2 form factor model for signal, while data points in blue are the results of the fit using Ball01. The errors shown are statistical only for both data and Monte Carlo. The red and blue histograms show the ISGW2 and Ball01 predictions, respectively.

$$|V_{ub}| = \sqrt{\frac{\mathcal{B}}{\Gamma_{thy}\tau_{B^0}}}, \tag{9.4}$$

where Γ_{thy}, as defined in Section 1.6, is the integral of the differential decay rate for $B \to \pi\ell\nu$ over all of phase space, calculated for a particular form-factor model. In this case, we have taken the values

$$\Gamma_{thy} = 9.56 \times 10^{12} s^{-1} (\mathrm{ISGW2}) \tag{9.5}$$

$$\Gamma_{thy} = (8.93^{+3.16}_{-2.24}) \times 10^{12} s^{-1} (\mathrm{Ball01}) \tag{9.6}$$

where the quoted uncertainty on Γ_{thy} in the case of the Ball form-factor model derives from the uncertainty on $f_+(q^2)$ given in Ref. [43]. These uncertainties are derived by varying the standard inputs to the LCSR calculation within some nominal range; the inputs so varied include the normalization of the form-factor at $q^2 = 0$. In principle, a similar uncertainty should exist on the Γ_{thy} quoted for ISGW2, but the uncertainty in the ISGW2 form-factor normalization and shape cannot easily be quantified.

Using $\tau_{B^0} = 1.542 \times 10^{-12}$s, we find that

$$|V_{ub}|_{ISGW2} = (2.97^{+0.12}_{-0.13}(stat.)^{+0.28}_{-0.19}(syst.) \pm 0.08(FF_{\pi\ell\nu})) \times 10^{-3}, \tag{9.7}$$

and

$$|V_{ub}|_{Ball01} = (3.07^{+0.12}_{-0.13}(stat.)^{+0.29}_{-0.19}(syst.) \pm 0.08(FF_{\pi\ell\nu})^{+0.47}_{-0.43}(\Gamma_{thy})) \times 10^{-3}. \tag{9.8}$$

These results are consistent within errors. The $|V_{ub}|$ result extracted using the Ball form-factor model, which provides a more complete error estimate, has a 19% relative error on $|V_{ub}|$. The two dominant contributions are the large experimental systematic uncertainty in $\mathcal{B}(B \to \pi\ell\nu)$ (which translates into a $\simeq 10\%$ relative uncertainty in $|V_{ub}|$), and the uncertainty in Γ_{thy}, which derives primarily from the uncertainty in the form-factor normalization, $f_+(0)$.

We believe that the sources of the large experimental systematic uncertainties in both $\mathcal{B}(B \to \pi\ell\nu)$ and $|V_{ub}|$ can be effectively addressed in the next iteration of this analysis, ultimately leading to a much reduced systematic error. In the following sections, we discuss a program for reducing the experimental systematic uncertainties, and the prospects for measuring the $\mathcal{B}(B \to \pi\ell\nu)$ branching ratio as a function of q^2.

9.3 Improvements to Existing Analysis

Difficulty handling the continuum background constitutes one of the two dominant sources of systematic error, and contribute in two distinct ways. First, insufficient Monte Carlo statistics result in a large uncertainty in the shape of the smoothed $q\bar{q}$ distribution. Second, the continuum background is not well-modelled—possibly because QED backgrounds are not included in the continuum background simulation. This causes significant variations in the fit result when certain background suppression cuts are varied. Re-optimizing the signal selection cuts to reduce the overall fraction of background due to $q\bar{q}$ decays, while maintaining the current signal-to-background ratio in the signal region, may reduce the impact of this statistically-derived uncertainty in the $q\bar{q}$ shape on the current analysis. It should also be possible to address the continuum modelling uncertainty within the context of the current analysis, either by identifying and suppressing the additional sources of continuum background, or, if possible, by including the simulated source distributions in the fit. Addressing this problem should reduce to the total experimental systematic error to the level of $13\% za$.

The analysis will also extend the fit for $\mathcal{B}(B^0 \to \pi^- \ell^+ \nu)$ in five equal bins in q^2, although the weak continuum statistics may present a problem, and the systematic uncertainty due to the $\mathcal{B}(B^0 \to \rho^- \ell^+ \nu)$ form factor should also be properly determined as soon as the updates to the form factor model calculation [44] become available.

It is also preferable to control the error from the $\mathcal{B}(B \to \rho\ell\nu)$ branching ratio and form factor by performing a simultaneous fit to to the $B \to \pi\ell\nu$ and $B \to \rho\ell\nu$ modes. Assuming the discrepancies in the $B^+ \to \rho^0 \ell^+ \nu$ decay channel can be addressed, the analysis should be upgraded to the four-channel fit, and the branching ratio and associated systematics derived for both $\mathcal{B}(B^0 \to \pi^- \ell^+ \nu)$ and $\mathcal{B}(B^0 \to \rho^- \ell^+ \nu)$. Interpretation of the q^2 dependent fit for the vector modes will remain more challenging, due to the dilution of the q^2 distribution by combinatoric signal.

9.4 Future Analysis

A newer dataset covering four years worth of data ($\simeq 240\,\mathrm{fb}^{-1}$) is now available. This dataset uses a newer data format that provides access to more detailed information, for instance, information at the single-crystal level for EMC showers and at the hit level for tracks; it also incorporates several years of improvements in track and neutral reconstruction and in particle identification algorithms. Resources also exist to generate larger Monte Carlo samples.

In this section, we outline a future analysis which is based on the current analysis, but exploits the improved resources, statistics, and information available in the new dataset to significantly reduce all sources of systematic error.

9.4.1 Continuum

Using the resources of the larger dataset, we propose a $q\bar{q}$ Monte Carlo sample with an equivalent luminosity five to ten times that of the data, in order to reduce the systematic uncertainty associated with the $q\bar{q}$ background shape. In order to do this efficiently, we propose generating a set of $q\bar{q}$ samples with a generator-level cut requiring the presence of a lepton with $p_l^* > 1.1\,\mathrm{GeV}$. This would provide more than a factor of ten in effective luminosity. This Monte Carlo sample would be used only to derive the lepton component of the continuum background. The fake lepton component of the background would be taken from reweighted on- and off-resonance data, analyzed in a mode where the candidates use as "leptons" any track which is not an electron or muon, according to the loosest particle identification criteria available. These data would be reweighted according to measured particle fractions and misidentification rates.

9.4.2 Neutrino Reconstruction

The bias and associated uncertainty in our selection efficiency created by the neutrino quality cut on $m^2_{miss}/2E_{miss}$, and the discrepancies in ΔE and m_{ES} resolution between data and Monte Carlo, also constitute sources of significant and reducible systematic error in this analysis. The dataset used in this analysis was subject to a number of known problems in tracking and reconstruction that have since been corrected (including the fact that a slightly different track-cluster matching algorithm was used in data than in simulation). Second, recent studies [45] suggest that a dominant contribution to the negative missing-mass-squared tail are events with background photons, unmatched clusters, and hadronic splitoffs in the EMC. Hadronic splitoffs, in particular, are likely to significantly contribute to the current systematic error, since they are not distinguished from photons and it is known that the simulation poorly models hadronic interactions. With shower information at the single-crystal level, we can perform the detailed studies needed to develop an effective splitoff suppression algorithm.

Other potential sources of this discrepancy include errors in track-cluster matching, and additional tracks: fake tracks, ghosts (multiple reconstructed versions of a single particle track, with nearly identical parameters and sharing most or all hits), and tracks from particle decays-in-flight and interactions with material.

An additional feature, which should improve the overall neutrino reconstruction efficiency and resolution, if not the agreement between data and simulation, involves the particle hypothesis used to fit tracks used in the neutrino reconstruction. Currently, all tracks used in neutrino reconstruction are fitted with the pion hypothesis. The errors this introduces should track relatively well between data and MC simulation, and thus a change in this procedure is not likely to significantly impact the systematic uncertainty due to neutrino reconstruction. However, for kaon and proton tracks in particular, the error in the track energy is significant; this results in an error in the missing mass squared calculation and a correspondingly reduced efficiency for the missing mass squared cut. Refitting each track according to its most probable particle identification should reduce these errors and raise the overall selection efficiency, as should any improvements in the efficiency for charged track or neutral reconstruction.

9.4.3 Other future improvements

Recovery of bremsstrahlung photons would increase the selection efficiency in the electron channel by reducing the number of candidates in the bremsstrahlung-induced negative tail of $\cos\theta_{\mathrm{BY}}$; it would also improve the accuracy of the q^2 measurement. Improvements in muon identification available in the new data model should enhance the yield, although the muon detection efficiency is expected to drop significantly with time.

9.4.4 Prospects for q^2 dependent analysis

The q^2 dependent analysis will benefit from the newer dataset on both statistical and systematic level. With the improvement in efficiency, we can reasonably expect on the order of 1000 perfectly reconstructed $B^0 \to \pi^- \ell^+ \nu$ decays and 400 $B^+ \to \pi^0 \ell^+ \nu$ decays, translating to populations on the order of 30-200 events in each of the five bins of the q^2 distribution.

The systematic improvements discussed in Sections 9.4.1 and 9.4.2 should also significantly impact the systematic errors of the q^2 dependent fit. Figure 9.1 shows that, even when coarsely binned, the continuum contribution to the q^2 spectra is visibly affected by statistical fluctuations. Moreover, even if the two-dimensional smoothing algorithm is still employed to compensate for the weak statistics in ΔE and m_{ES} a separate smoothed distribution must be derived for each bin in q^2. Likewise, we must rely on Monte Carlo signal simulation in order to unfold the measured q^2 spectrum, and thus any improvements in tracking and neutral reconstruction will improve the agreement in neutrino—and thus q^2—resolution between data and Monte Carlo, and thus reduce the uncertainty in that unfolding procedure.

While the other dominant background, the $b \to c\ell\nu$ contribution, has only a modest impact on the full branching ratio systematic uncertainty, it may have a larger impact in the context of the q^2-dependent fit. While recent measurements have improved the knowledge and errors on the $B \to D^*\ell\nu$ form factor parameters

($B \to D^*\ell\nu$ decays contribute about 70% of the $b \to c\ell\nu$ background candidates), there are still significant uncertainties in the overall composition of the $b \to c\ell\nu$ background, including the size of the $B \to D^*\ell\nu$ branching fraction [32], the branching ratio and form factor for $B \to D\ell\nu$, and the relative branching ratios and widths of the various higher mass states collectively referred to as $B \to D^{**}\ell\nu$. These uncertainties have a significant impact on the background lepton spectrum, and can reasonably be expected to impact the q^2 distribution of the $b \to c\ell\nu$ background component as well. Independent studies are underway to address these uncertainties, both by resolving the $B \to D^*\ell\nu$ branching ratio anomaly and by measuring the higher-mass D^{**} states using a data sample tagged by exclusively reconstructed hadronic B decays. We could hope to benefit from these studies.

9.5 Outlook

In order for tests of the unitarity of the CKM matrix to be sensitive to physics beyond the Standard Model, we require a measurement of $|V_{ub}|$ with better than 10% precision. The current $|V_{ub}|$ measurement provided by this analysis has a precision of about 19%.

As $|V_{ub}|$ is proportional to the square root of $\mathcal{B}(B \to \pi\ell\nu)$, the relative statistical or experimental error on $|V_{ub}|$ is half the corresponding error on the branching ratio. Assuming the experimental systematic uncertainty in the branching fraction measurement in this analysis could be brought down to the 13% level, we could expect a 4.5% statistical and 7.5% experimental error on an extracted $|V_{ub}|$.

Form-factor shape uncertainties contribute twice to uncertainties in $|V_{ub}|$ once through the uncertainty in the branching ratio measurement (which is about 5%, and translates to about a 3% relative error on $|V_{ub}|$), and once as part of the overall uncertainty in Γ_{thy}. The dominant contribution to the Γ_{thy} uncertainty, however, and thus the dominant contribution to the uncertainty in $|V_{ub}|$, is not the uncertainty in the form-factor shape, but rather the uncertainty in the form-factor normalization. Current lattice and light-cone sum rule calculations quote errors in the form factor normalization at alone the 20% level, which would translate to an order 10% Γ_{thy}

error in $|V_{ub}|$; when combined with the current form-factor shape uncertainty, that error is in fact about 15%.

The future analysis program promises a factor of three in statistics, and the possibility of lowering the relative systematic uncertainty on the branching fraction to a level competitive with existing measurements (about 8%). Such a measurement would allow a $|V_{ub}|$ extraction with something like a 2.5% relative statistical error and a 4% relative systematic error. We could also reasonably expect, at that precision, to truly constrain the form-factor shape, so that it would make a smaller contribution to the total theoretical error on both the branching ratio and $|V_{ub}|$. At that point, the roughly 10% error on $|V_{ub}|$ due to uncertainties in the form factor normalization would become the limiting uncertainty.

Appendix A

Differential decay rate for $B \to D^*\ell\nu(D^* \to D\pi, D^* \to D\gamma)$

Equations 1.22 and 1.23 in Chapter 1 give, in full generality, the semileptonic decay rate for a B meson, where H_+, H_-, and H_0 contain not only the q^2 dependence of the decay rate but also the angular dependence of the decay of the final-state meson. For any particular decay (i.e. both the final-state meson and a particular decay mode of the final-state meson have been specified) it is preferable to rewrite the decay rate in terms of a set of reduced helicity amplitudes $\bar{H}_+$, $\bar{H}_-$, $\bar{H}_0$ which depend only on q^2, and to make the angular dependence explicit.

This appendix derives this angular dependence for $B \to D^*\ell\nu$, $B \to D\pi$, and $B \to D^*\ell\nu$, $B \to D\gamma$. The derivations, however, are applicable (with the appropriate substitutions) to any semileptonic B decay with a vector meson in the final state where the vector meson decays to two pseudoscalar mesons or to a pseudoscalar meson and a photon. Thus equation A.13 can also be used to give the decay rate for $B \to \rho\ell\nu$.

The angular dependence of H_+ H_- and H_0 may be determined by considering $D_m^{*\,J=1} \to PF$, where m indexes the z-component of the spin of the vector meson, and therefore the corresponding W helicity amplitude ($m = 1$ corresponds to H_+, etc.), and the final-state meson F has as yet unspecified transformation properties. The momentum vector p, θ, ϕ and helicity λ of F in the rest frame of the $D^*$is enough to fully specify the final state. The matrix element of the transition operator $\mathbf{O}$, between a $D^*$of definite angular momentum, $J_z = m$ and final state $|p, \theta, \phi, \lambda>$, is (reference Jacob and Wick)

$$< D_m^{*\,J=1}|\mathbf{O}|p, \theta, \phi, \lambda >=< D_m^{*\,J=1}|R^\dagger R\mathbf{O}R^\dagger R|p, \theta, \phi, \lambda > \tag{A.1}$$

where $R = R(\phi, \theta, -\phi)$ rotates $|p, \theta, \phi, \lambda >$ to $|p, 0, 0, \lambda >$. $\mathbf{O}$ is invariant under rotations, so $R\mathbf{O}R^\dagger = \mathbf{O}$, and $R^\dagger$ acts to the left on $|D_m^{*\,J=1}>$, yielding $\sum_{m'} < D^*_{m'}|D^J_{mm'}(\theta, \phi) = \sum_{m'} < D^*_{m'}|e^{i\phi m}e^{-i\phi m'}d^J_{mm'}(\theta)$.

Thus we have

$$< D_m^{*\,J=1}|\mathbf{O}|p, \theta, \phi, \lambda >= \sum_{m'} e^{i\phi m}e^{-i\phi m'}d^1_{mm'} < D^*_{m'}|\mathbf{O}|p, 0, 0, \lambda >=$$
$$e^{i\phi(m-\lambda)}d^1_{m\lambda} < D^*_\lambda|\mathbf{O}|p, 0, 0, \lambda > \tag{A.2}$$

Since $\mathbf{O}$ is a spin 0 operator, the only term that survives is the one proportional to the matrix element where $m' = \lambda$. The matrix element $< D^*_{m'}|\mathbf{O}|p, 0, 0, \lambda >$, which has no angular dependence, is a matrix element of a strong decay and is, again, difficult to calculate from first principles. It is easier to treat this matrix element as having amplitude one (any phases will become irrelevant in the decay rate), and restore the information by including the appropriate, empirically determined branching fraction $\mathcal{B}(\mathcal{D}^* \to \mathcal{PF})$ in the decay rate. In general, the decay rates for orthogonal final states (different values of λ) must be calculated separately and then summed.

In the case of vector meson decay to two pseudoscalars, such as $D^* \to D\pi$, λ must be zero, and there is only one final state to consider.

$$H_0 \propto d^1_{00} = \cos(\theta) \tag{A.3}$$
$$H_+ \propto e^{i\phi}d^1_{10} = e^{i\phi}\frac{-\sin\theta}{\sqrt{2}} \tag{A.4}$$
$$H_- \propto e^{-i\phi}d^1_{-10} = e^{-i\phi}\frac{\sin\theta}{\sqrt{2}} \tag{A.5}$$

The $d^j_{mm'}$ are tabulated in [2].

The decay rate for $B \to D^*\ell\nu$, $D^* \to D\pi$ then becomes:

$$\begin{aligned}
\frac{d\Gamma(B \to D^* l\bar{\nu}, D^* \to D\pi)}{dq^2 d\cos\theta_l d\cos\theta d\cos\phi} &= \frac{3G_F^2 |V_{cb}|^2 P_{D^*} q^2}{8(4\pi)^2 M_B^2} \mathcal{B}(D^* \to D\pi) \\
\times \Big[&(1-\eta\cos\theta_l)^2 \sin^2\theta |\bar{H}_+|^2 + (1+\eta\cos\theta_l)^2 \sin^2\theta |\bar{H}_-|^2 \\
&+ 4\sin^2\theta_l \cos^2\theta |\bar{H}_0|^2 + 2\sin^2\theta_l \sin^2\theta \cos(2\phi) \bar{H}_+ \bar{H}_- \\
&+ 4\eta \sin\theta_l (1-\eta\cos\theta_l) \sin\theta\cos\theta\cos\phi \bar{H}_+ \bar{H}_0 \\
&- 4\eta \sin\theta_l (1+\eta\cos\theta_l) \sin\theta\cos\theta\cos\phi \bar{H}_- \bar{H}_0 \Big]
\end{aligned} \tag{A.6}$$

In the case of a vector meson decaying to one pseudoscalar meson and a photon, such as $D^* \to D\gamma$, we have two possible final states, corresponding to the allowed transverse polarizations, $\lambda = \pm 1$, of the final-state photon. The decay rates for these two final states must be calculated separately and then summed.

For $\lambda = -1$,

$$H_0 \propto e^{i\phi} d^1_{0-1} = -e^{i\phi} \frac{\sin\theta}{\sqrt{2}} \tag{A.7}$$

$$H_+ \propto e^{2i\phi} d^1_{1-1} = e^{2i\phi} \frac{1-\cos\theta}{2} \tag{A.8}$$

$$H_- \propto d^1_{-1-1} = \frac{1-\cos\theta}{2} \tag{A.9}$$

For $\lambda = 1$,

$$H_0 \propto e^{-i\phi} d^1_{01} = e^{-i\phi} \frac{\sin\theta}{\sqrt{2}} \tag{A.10}$$

$$H_+ \propto d^1_{11} = \frac{1-\cos\theta}{2} \tag{A.11}$$

$$H_- \propto e^{-2i\phi} d^1_{-11} = e^{-2i\phi} \frac{1-\cos\theta}{2} \tag{A.12}$$

Using this and Equations 1.22 and 1.23 for the total decay rate, and averaging over the final-state photon polarization, we find that the total decay rate for $B \to D^* l\bar{\nu}, D^* \to D\gamma$ is:

$$\frac{d\Gamma(B \to D^* l\bar{\nu}, D^* \to D\gamma)}{dq^2 d\cos\theta_l d\cos\theta d\cos\phi} = \frac{3G_F^2|V_{cb}|^2 P_{D^*} q^2}{8(4\pi)^2 M_B^2} \mathcal{B}(D^* \to D\gamma)$$
$$\times \Big[(1-\eta\cos\theta_l)^2(1+\cos^2\theta)|\bar{H}_+|^2 + (1+\eta\cos\theta_l)^2(1+\cos^2\theta)|\bar{H}_-|^2$$
$$+ 4\sin^2\theta_l \sin^2\theta|\bar{H}_0|^2 - 2\sin^2\theta_l\sin^2\theta\cos(2\phi)\bar{H}_+\bar{H}_-$$
$$- 4\eta\sin\theta_l(1-\eta\cos\theta_l)\sin\theta\cos\theta\cos\phi\bar{H}_+\bar{H}_0$$
$$+ 4\eta\sin\theta_l(1+\eta\cos\theta_l)\sin\theta\cos\theta\cos\phi\bar{H}_-\bar{H}_0\Big] \quad \text{(A.13)}$$

In general, it is more useful to describe the decay rates in terms of angles that can be well-calculated for a measured decay: the polar angle of the D meson in the helicity frame of the D^*, and the dihedral angle χ between the decay plane of the leptons and the decay plane of the D^* in the rest frame of the B. An illustration of these angles and their definition is shown in Figure A.1. By considering the case where $\phi = 0$, one can see that the angle θ may be directly replaced by θ_D, since the in the D^* helicity frame the decays are back-to-back. Likewise, in the rest frame of the parent B, the $D^*$and virtual W are also back-to-back; at $\phi = 0$ for the pion or photon, the dihedral angle will be zero for an appropriate choice of sign convention. Thus the dihedral angle χ may be directly substituted for ϕ.

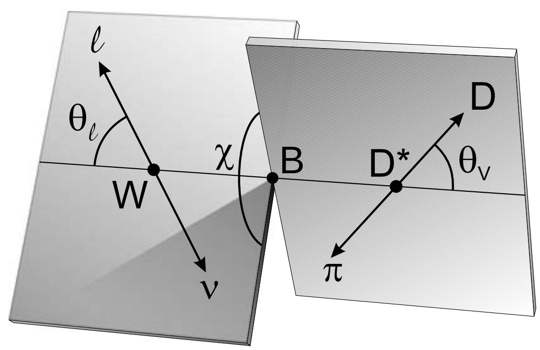


Figure A.1: Illustrated definition of the angles θ_l, θ_V, and χ.

Bibliography

[1] P. Ball, ECONF **C0304052**, WG101 (2003), hep-ph/0306251.

[2] K. Hagiwara *et al.*, Physical Review D **66**, 010001+ (2002).

[3] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).

[4] L. Wolfenstein, Presented at 4th Int. Symp. on Heavy Flavor Physics, Orsay, France, Jun 25-29, 1991.

[5] K. R. Schubert, Int. J. Mod. Phys. **A19**, 1004 (2004).

[6] F. J. Gilman and R. L. Singleton, Phys. Rev. **D41**, 142 (1990).

[7] P. Ball and V. M. Braun, Phys. Rev. **D58**, 094016 (1998), hep-ph/9805422.

[8] J. D. Richman and P. R. Burchat, Rev. Mod. Phys. **67**, 893 (1995), hep-ph/9508250.

[9] D. Scora and N. Isgur, Phys. Rev. **D52**, 2783 (1995), hep-ph/9503486.

[10] J. Shigemitsu *et al.*, Phys. Rev. **D66**, 074506 (2002), hep-lat/0207011.

[11] HPQCD, C. T. H. Davies *et al.*, Phys. Rev. Lett. **92**, 022001 (2004), hep-lat/0304004.

[12] J. Shigemitsu *et al.*, (2004), hep-lat/0408019.

[13] M. Okamoto *et al.*, (2004), hep-lat/0409116.

[14] UKQCD, K. C. Bowler *et al.*, Phys. Lett. **B486**, 111 (2000), hep-lat/9911011.

[15] A. X. El-Khadra, A. S. Kronfeld, P. B. Mackenzie, S. M. Ryan, and J. N. Simone, Phys. Rev. **D64**, 014502 (2001), hep-ph/0101023.

[16] JLQCD, S. Aoki *et al.*, Phys. Rev. **D64**, 114505 (2001), hep-lat/0106024.

[17] D. Becirevic and A. B. Kaidalov, Phys. Lett. **B478**, 417 (2000), hep-ph/9904490.

[18] BABAR, B. Aubert *et al.*, Nucl. Instrum. Meth. **A479**, 1 (2002), hep-ex/0105044.

[19] J. Back, http://www.slac.stanford.edu/bfroot/www/detector/calorimeter/-software/trkmatch/data/trkmatch.html, 2000, Internal *BABAR* webpage.

[20] G. Mancinelli *et al.*, Kaon selection at babar, 2001, BAD #116, *BABAR* Internal Analysis Document.

[21] G. Mancinelli *et al.*, An event likelihood algorithm for DRC based particle identification, 2003, BAD #165, *BABAR* Internal Analysis Document.

[22] A. Sarti *et al.*, Measurement of the inclusive charmless semileptonic branching ratio and the determination of V_{ub} based on the hadron mass spectrum in $b \to x_u \ell \nu$ decays, 2003, BAD #540, *BABAR* Internal Analysis Document.

[23] A. Ryd and D. Lange, Evtgen documentation, 2003, BAD #522, *BABAR* Internal Analysis Document.

[24] M. Bona *et al.*, Report of the tracking efficiency task force, 2002, BAD #324, *BABAR* Internal Analysis Document.

[25] F. De Fazio and M. Neubert, JHEP **06**, 017 (1999), hep-ph/9905351.

[26] D. Fortin, http://www.slac.stanford.edu/bfroot/www/physics/analysis/awg/-inclusivesl/common/dominiquehybrid.html/hybrid.html, 2004, Internal *BABAR* webpage.

[27] *BABAR* B. A. *et. al.*, Measurement of the $B \to D^*$ form factors in the semileptonic decay $B^0 \to D^{*-} \ell^+ \nu$ 2004.

[28] J. L. Goity and W. Roberts, Phys. Rev. **D51**, 3459 (1995), hep-ph/9406236.

[29] T. Sjostrand, Prepared for Particles and Fields 92: 7th Meeting of the Division of Particles Fields of the APS (DPF 92), Batavia, Illinois, 10-14 Nov 1992.

[30] F. di Lodovico, http://www.slac.stanford.edu/bfroot/www/physics/-analysis/awg/neutrals/validation/results/results.html, 2003, Internal *BABAR* webpage.

[31] W. T. Ford, Choice of kinematic variables in b meson reconstruction—take 3, 2000, BAD #53, *BABAR* Internal Analysis Document.

[32] S. Eidelman *et al.*, Physics Letters B **592**, 1+ (2004).

[33] M. Bomben *et al.*, Measurement of V_{cb} with the semileptonic $B^0 \to D^{*-}\ell^+\nu$ decays, 2004, BAD #776, *BABAR* Internal Analysis Document.

[34] R. J. Barlow and C. Beeston, Comput. Phys. Commun. **77**, 219 (1993).

[35] P. Bock, June 2004, Private communication via J. Dingfelder.

[36] M. B. *et al.*, Report on the tracking efficiency task force for 2001, 2001, BAD #324, *BABAR* Internal Analysis Document.

[37] T. Brandt, Measurement of the inclusive electron spectrum, moments and branching fraction for semileptonic b meson decays, 2003, BAD #636, *BABAR* Internal Analysis Document.

[38] J. L. Diaz-Cruz, G. Lopez Castro, and J. H. Munoz, Phys. Rev. **D54**, 2388 (1996), hep-ph/9605344.

[39] G. Lopez Castro, J. H. Munoz, and G. Toledo Sanchez, Phys. Rev. **D56**, 6025 (1997), hep-ph/9709266.

[40] BaBar, B. Aubert *et al.*, (2004), hep-ex/0408045.

[41] P. Ball, 2004, Private communication.

[42] CLEO, S. B. Athar *et al.*, Phys. Rev. **D68**, 072003 (2003), hep-ex/0304019.

[43] P. Ball and R. Zwicky, JHEP **10**, 019 (2001), hep-ph/0110115.

[44] P. Ball and V. M. Braun, Phys. Rev. **D58**, 094016 (1998), hep-ph/9805422.

[45] D. Brown, http://dnbmac3.lbl.gov/brownd/meetings/trackingnm2.pdf, 2004, Internal *BABAR* presentation.