THE DIFFERENTIAL RATES OF EXCLUSIVE

CHARMLESS SEMILEPTONIC B MESON DECAY

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Matthew Randall Shepherd

August 2005

THE DIFFERENTIAL RATES OF EXCLUSIVE CHARMLESS
SEMILEPTONIC B MESON DECAY

Matthew Randall Shepherd, Ph.D.

Cornell University 2005

The rates of exclusive charmless semleptonic B meson decay provide experimental input necessary to extract the magnitude of Cabibbo-Kobayashi-Maskawa quark mixing matrix element V_{ub}, which can be used to search for physics beyond the Standard Model through precision tests of the parameters of the weak interaction. Using the CLEO detector at the Cornell Electron Storage Ring, we analyze 30.8 million B meson decays to measure the rates of decay for $B^0 \to \pi^- \ell^+ \nu$ and $B^0 \to \rho^- \ell^+ \nu$. The measurement is made in bins of the lepton decay angle in the W helicity frame and the four-momentum transfer to the virtual W, q^2, to minimize dependance on the theoretical form factors that govern the decay dynamics. The total rates are $\mathcal{B}(B^0 \to \pi^- \ell^+ \nu) = (1.32 \pm 0.15 \pm 0.11 \pm 0.02) \times 10^{-4}$ and $\mathcal{B}(B^0 \to \rho^- \ell^+ \nu) = (2.73 \pm 0.36 \pm 0.32 \pm 0.04) \times 10^{-4}$ where the errors are statistical, experimental systematic, and theoretical systematic respectively.

BIOGRAPHICAL SKETCH

Matthew was born on August 23, 1978 in Evansville, Indiana. He attended William Henry Harrison High School where his interest in physics was first kindled by Mike Kelly, his high school physics teacher. With visions of medical school in the future, he choose to pursue a degree in physics. In the fall of 1996 he headed just north of his hometown to begin his undergraduate studies at Indiana University in Bloomington.

The next four years at Indiana University forever changed his life in two ways. Most notably he met his future wife, Catherine Magee, then a ballet student at the university. In addition, through several years of work in the research group of Alex Dzierba, he realized his true love of physics and decided to pursue a Ph.D. in experimental elementary particle physics.

After receiving bachelors degrees in physics and mathematics from Indiana University in 2000, Matthew moved to Ithaca, New York to begin to embark on his Ph.D. at Cornell University. In early 2003 he began his dissertation research on exclusive charmless semileptonic B decays with Professor Lawrence Gibbons and, later that year, received an M.S. degree in physics.

In the summer of 2003 he and Catherine were married. They will be returning to Bloomington in the fall of 2005 where Matthew will begin an assistant professorship at his alma mater.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Motivation

One of the most outstanding objectives of modern particle physics is to understand the mechanism by which the matter-dominated universe that we live in was created. In 1967 Andrei Sakharov proposed three conditions which were necessary to allow baryogensis in the universe [1]. One of the three conditions can be elegantly accommodated for with the Standard Model of particle physics: the existence of CP violation. Is this natural source of CP violation *the unique source* of CP violation that Sakharov stated was necessary to produce the universe we live in? This is precisely the question that motivates this work!

1.1 Discovery of CP Violation

In 1957 Wu *et al.* experimentally verified that the weak interaction violates parity (P) by measuring the direction that electrons were emitted with respect to the nuclear spin in Cobalt 60 beta-decay [2]. The observation that the electrons were preferentially emitted in the direction opposite the spin of the nucleus meant that the process must be parity violating as the nuclear spin direction would remain unchanged under a parity transformation. In fact we know that the structure of the weak interaction maximally violates parity. Femi's traditional four-particle vertex of the weak interaction was replaced by the weak interaction Hamiltonian:

$$\mathcal{H} = \frac{G_F}{\sqrt{2}} \left[\bar{\psi}_1 \gamma_\mu \left(1 - \gamma_5\right) \psi_2 \right] \left[\bar{\psi}_3 \gamma^\mu \left(1 - \gamma_5\right) \psi_4 \right], \qquad (1.1)$$

where G_F is the Fermi coupling constant. Parity violation arises due to the $\gamma_\mu \left(1 - \gamma_5\right)$ or "vector$-$axial-vector" $(V - A)$ structure of the Hamiltonian. The

$(1 - \gamma_5)$ portion simply projects out the left-handed components of the ψs; therefore, by construction *only* left-handed particles or right-handed *anti*-particles interact weakly.

One might logically think that applying the parity operator (P), which changes "handedness", followed by the charge conjugation operator (C), which replaces particles with their anti-particles, would be a symmetry of the weak interaction. However, in 1964 Cronin and Fitch observed that this so-called CP symmetry was violated in the weak decays of neutral K mesons [3]. This groundbreaking observation sparked the search for natural ways to accommodate CP violation within the emerging Standard Model.

1.2 CP Violation in the Standard Model

To explain strangeness-violating weak decays of K mesons, Cabibbo, in 1963, proposed that quark eigenstates of the weak interaction were not the same as those of the strong interaction [4]. At the time, only u, d, and s quarks where known to exist, and Cabibbo's solution was that the s and d quarks of the weak interaction were different than those of the strong interaction. Mathematically:

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta_C & \sin\theta_C \\ -\sin\theta_C & \cos\theta_C \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}, \tag{1.2}$$

where the prime denotes the weak interaction states and θ_C, the Cabibbo angle, is the amount the that weak states are rotated from the strong states[1]. K mesons are therefore produced and bound through the strong force in states of d and s, and the subsequent hadrons produced in the decay would contain u and d quarks.

---

[1]Experimentally this angle is quite small – about $13°$

However, the intermediate weak decay proceeds through the s' and d' states. This allows $s \to d$ transitions and solves the problem of strangeness violating decays.[2]

Kobayashi and Maskawa realized that an extension of Cabibbo's idea could also elegantly account for CP violation in the Standard Model [5]. They predicted the existence of an additional two quarks, t and b, to add to the known u, d, and s, and the proposed c, quarks. Equation 1.2 could then be expanded to

$$
\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} . \tag{1.3}
$$

Now the matrix V, known as the CKM matrix, describes the rotation of the strong interaction quark eigenstates into the weak interaction eigenstates. The quark-level weak current in the Hamiltonian is then

$$
\mathcal{J}^\mu = \bar{\psi}_u \gamma^\mu \left(1 - \gamma_5\right) \psi_{d'} = \bar{\psi}_u \gamma^\mu \left(1 - \gamma_5\right) V \psi_d, \tag{1.4}
$$

where $\bar{\psi}_u$ denote the up-type (u, c, and t) quark and $\psi_{d'}$ and ψ_d denote the weak and strong eigenstates of the down-type quark. If we view the quarks as occurring in three-generations

$$
\begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix} , \tag{1.5}
$$

then the fact that V is not diagonal permits "generation-changing" decays such as $K \to \pi \ell \nu$, which is what initially motivated Cabibbo. However, extending the matrix further to incorporate a third generation of quarks as done by Kobayashi and Maskawa allows for CP violation.

---

[2]It also created a problem in that some decays predicted from Cabibbo's model were not observed. Glashow *et al.* solved this problem with the proprosal of the charm quark, whose presence in the "GIM mechanism" cancels out the unobserved decays. Of course the charm quark was later discovered.

If there exist strictly three generations of quarks then the CKM matrix must be a *unitary* complex rotation matrix. Any arbitrary matrix of such type can be specified by three angles and a single phase. These four parameters are fundamental parameters of the Standard Model. The common way [6] to parameterize V in terms of the three angles, θ_{12}, θ_{23}, and θ_{13} and the phase δ_{13} is

$$V = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix}, \quad (1.6)$$

where the c and s represent cosine and sine of the angles. The subscripts are chosen to indicate the generation labels. If we "remove" the effects of third generation by setting $\theta_{23} = \theta_{13} = 0$ we can recover Cabibbo's initial matrix (Equation 1.2) and identify θ_{12} as θ_C.

CP operating on a weak current of the form of Equation 1.4 transforms $V \to V^*$, thereby flipping the sign of the weak phase. If this phase is the only phase present in calculation of the matrix element, then the change in sign does not alter the magnitude of the amplitude. However if additional amplitudes contribute and they have relative phases which do not change sign under CP, the weak phase can produce CP violation.

Suppose that in the calculation of some amplitude A two indistinguishable processes contribute with magnitudes a_1 and a_2. These processes can be, for example, tree and penguin diagrams. In the case of some neutral mesons, a_1 could be direct decay to a final state and a_2 could be the amplitude to mix first then decay to the same final state. In any case a_1 and a_2, being complex amplitudes, have some relative weak phase δ_W and some other relative phase, like a strong
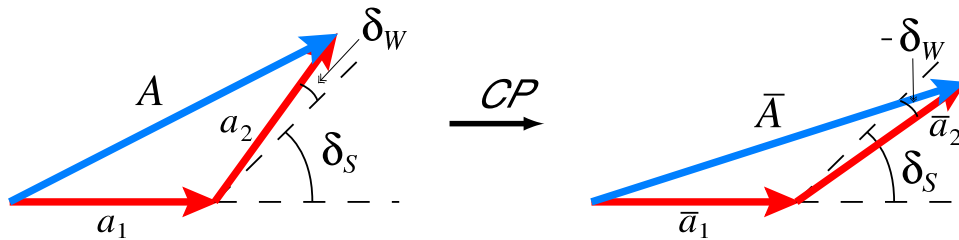
Figure 1.1: A graphical sketch of how CP is violated using the weak phase δ_W. Amplitudes a_1 and a_2 coherently contribute to the full amplitude A. Both a_1 and a_2 have relative weak phase δ_W and some other relative phase δ_S. Under CP the weak phase changes sign causing $|\bar{A}| > |A|$ and therefore CP violation.

phase, δ_S. We could then write

$$A = a_1 + a_2 e^{i(\delta_S + \delta_W)}. \tag{1.7}$$

Under CP, the sign of the relative weak phase is changed giving

$$\bar{A} = a_1 + a_2 e^{i(\delta_S - \delta_W)}. \tag{1.8}$$

Therefore as long as, $\delta_S \neq 0$ the amplitudes A and $\bar{A}$ will not be equal and the process is therefore CP violating as depicted graphically in Figure 1.1. Kobayashi and Maskawa proposed that this mechanism was responsible for the observed CP violation.

Realizing that the magic of the CKM matrix was in the complex phase, and noting that V is nearly diagonal, Wolfenstein [7] suggested the following parametrization

$$V \approx \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3 (\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3 (1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix}. \tag{1.9}$$

In this parametrization $\lambda \equiv \sin\theta_C$, and the expression above is accurate to $\mathcal{O}(\lambda^4)$. The magnitudes of the other off diagonal elements are set by the parameter A, and

CP violation is then carried by the $i\eta$ factor. Of the four Wolfenstein parameters, ρ and η are the least constrained by experimental measurements.

If we assume that all observable CP violation is due to the complex phase in the CKM matrix, then there should exist one unique value of (ρ, η) that satisfies all experimental results. Furthermore, the absence of additional quark generations or weak coupling of quarks to other particles would imply that V is unitary; therefore, the dot product of any two rows or columns must be equal to zero. To illustrate this condition pictorially, the third and first columns are dotted together, and the three components that must add to zero are visualized as a triangle in the complex plane as shown in Figure 1.2 with interior angles α, β, and γ. Each side represents one of the three products in the dot product. The unitarity condition is satisfied when the three complex numbers, vectors in the plane, form a closed triangle. It is common to normalize the base to unit length by dividing each side by the real number $V_{cd}V_{cb}^*$. When the variables ρ and η are recast as

$$\bar{\rho} \equiv \rho\left(1 - \frac{\lambda^2}{2}\right) \tag{1.10}$$

$$\bar{\eta} \equiv \eta\left(1 - \frac{\lambda^2}{2}\right), \tag{1.11}$$

we can plot the allowed position for the apex of the triangle as shown in Figure 1.3. The current best fit for the apex of the triangle is $(\bar{\rho}, \bar{\eta}) = (0.21, 0.34)$ [8].

One of the primary objectives of precision electro-weak physics is to over-constrain the apex of the triangle by measuring the the two sides and three angles independently. Any inconsistency in the allowed value of $(\bar{\rho}, \bar{\eta})$ would be a signal of CP violation outside of the Kobayashi-Maskawa mechanism and consequently a indication of new physics beyond the standard model.
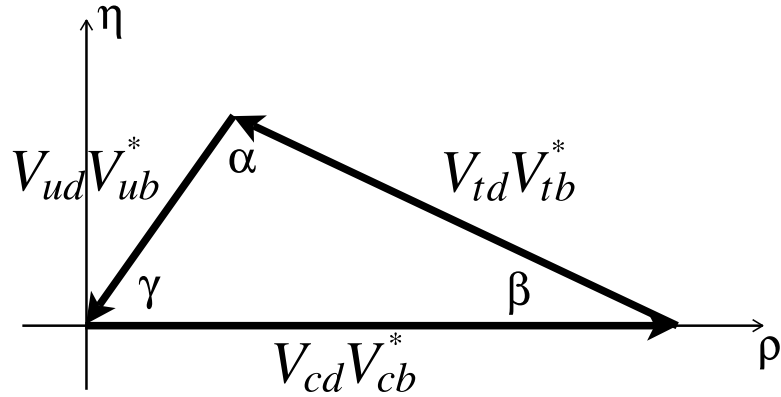
Figure 1.2: The unitarity triangle is constructed by dotting the first and third columns of the CKM matrix together. In the Wolfenstein parametrization $V_{cd}V_{cb}^*$ is purely real and is therefore the base of the triangle.
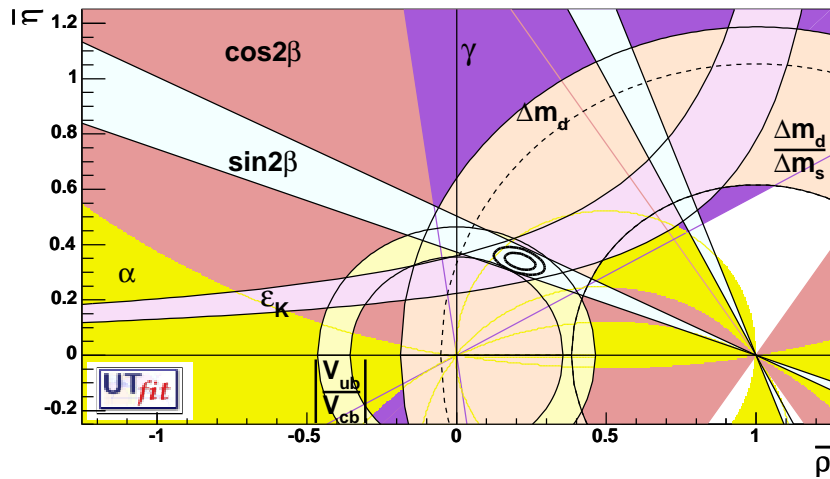
Figure 1.3: The current experimental constraints on the apex of the unitarity triangle in the $\bar{\rho}/\bar{\eta}$ plane are shown. The best fit is given by $(\bar{\rho}, \bar{\eta}) = (0.21, 0.34)$. The 99% and 95% confidence intervals are shown.

1.3 Connecting $B \to X_u \ell \nu$ with CP Violation

Decays of the B meson provide a wonderful laboratory for studying the unitarity triangle sketched in Figure 1.2. With the exception of the ϵ_K band, all of the experimental constraints pictured in the $\bar{\rho} - \bar{\eta}$ plane in Figure 1.3 come from the study of the B or B_s meson decay. Recall that CP violation is carried by the $i\eta$ portion of the CKM matrix and it is therefore the area of the triangle that sets the magnitude of the CP violation. Constraints on the angles α, β, and γ are provided by direct observation of CP violating processes. At least one angle must be measured to establish that the triangle is not degenerate with zero area and consequently no allowed CP violation. The constraints on the length of the sides, drawn as annuli in the $\rho - \eta$ plane, are derived from measurements of the magnitudes of the CKM matrix elements.

We would like to focus specifically on the determination of the magnitude of V_{ub}. The ratio $|V_{ub}|/|V_{cb}|$ provides the annular constraint about $(\bar{\rho}, \bar{\eta}) = (0, 0)$ pictured in Figure 1.3. Currently $|V_{cb}|$ is experimentally measured to roughly the 2% level; however, measurements of V_{ub} in exclusive decay channels with an uncertainty under 10% have yet to be made. Of all of the CKM matrix elements, the magnitude of V_{ub} is the least precisely determined. In order to further subject the Standard Model to precision testing and ultimately search for physics beyond the Standard Model, we must shrink the width of the $|V_{ub}|/|V_{cb}|$ annulus by providing a more precise value $|V_{ub}|$.

The magnitude of V_{ub} appears directly in the expressions for the decay rates of B hadrons to mesons containing u quarks. Specifically for semileptonic decay of

B meson into light-quark meson we can write the decay amplitude as

$$A(B \to X_u \ell \nu) = \frac{G_F}{\sqrt{2}} V_{ub} L^\mu H_\mu,\tag{1.12}$$

where L^μ and H_μ are the leptonic and hadronic currents. Semileptonic decay has the advantage that these currents are not coupled together through final state strong interactions. This comes at a price for the experimentalist, because the semileptonic channel is complicated by the undetectable final state neutrino. Additionally the extraction of V_{ub} is impeded by the theoretical uncertainty in the calculation of the hadronic current. Unfortunately measuring the bare $b \to u$ quark process is not an option – the measurement must be made with the quarks embedded inside of hadrons. The goal of this analysis is to provide a precision measurement of $B \to X_u \ell \nu$ that depends minimally on the evaluation of the hadronic current. By doing the measurement independent of the hadronic current, H_μ, the measurement will weather theoretical changes and can always be combined with state of the art calculations of H_μ to produce a precision value for $|V_{ub}|$.

Chapter 2

Theoretical Background

This chapter develops the expression for $\Gamma\left(B \to X_u \ell \nu\right)$ in terms of $|V_{ub}|$ and other kinematical and dynamical variables. I will first discuss the kinematic aspects of three-body decay in the context of the $V - A$ weak interaction. The chapter concludes with discussion of the decay form factors which are, theoretically, the least well understood aspects of the of the decay.

2.1 Decay Kinematics

We seek to write the decay rate, which we can measure experimentally, in terms of the amplitude that contains $|V_{ub}|$ presented in Equation 1.12. The following discussion parallels the analysis by Gilman and Singleton [9]. The differential decay rate can be written as

$$d\Gamma\left(B \to X_u \ell \nu\right) = \frac{1}{M_B}\left|A\left(B \to X_u \ell \nu\right)\right|^2 d\Pi_3, \tag{2.1}$$

where $d\Pi_3$ represents the allowed three-body phase space. The kinematics of the decay are sketched in Figure 2.1. Integrating over the decay angle of the meson in the rest frame of B the differential rate becomes

$$\frac{d\Gamma}{dq^2 d\Omega_l} = \frac{1}{2M_B^2}\frac{p_{X_u}}{(4\pi)^4}\left|A\left(B \to X_u \ell \nu\right)\right|^2, \tag{2.2}$$

where $d\Omega_l$ is the solid angle of the lepton in the virtual W rest frame. The amplitude squared is then

$$\left|A\left(B \to X_u \ell \nu\right)\right|^2 = \frac{G_F^2}{2}|V_{ub}|^2 L^{\mu\nu} H_\mu H_\nu^\dagger, \tag{2.3}$$

where the leptonic tensor, $L^{\mu\nu}$, has been constructed from the lepton current

$$L^\mu = \bar{u}_l \gamma^\mu (1 - \gamma_5) v_\nu \qquad (2.4)$$

and the hadronic current, H^μ, is given by

$$H^\mu = \langle p_{X_u} | J^\mu | p_B \rangle. \qquad (2.5)$$

Regardless of its form, H^μ can be expanded in the helicity basis of the virtual W. The spinless B meson causes the helicity of the W to be linked to the helicity of the hadronic system. The leptonic tensor is evaluated in the massless lepton limit and contracted with the hadronic current to give the final expression for the rate, integrated over the azimuthal angle of the lepton, in terms of the three W helicity amplitudes, H_+, H_-, and H_0, as

$$\frac{d\Gamma}{dq^2 d\cos\theta_{Wl}} = \frac{1}{128\pi^3} G_F^2 |V_{ub}|^2 p_{X_u} \frac{q^2}{M_B^2} \left[\frac{1}{2} (1 - \cos\theta_{Wl})^2 |H_+|^2 + \right.$$
$$\frac{1}{2} (1 + \cos\theta_{Wl})^2 |H_-|^2 +$$
$$\left. \sin^2\theta_{Wl} |H_0|^2 \right]. \qquad (2.6)$$

In order to develop an expression for the hadronic current we note that J^μ must carry the $V - A$ structure of the weak interaction. We first decompose J^μ into all possible vector and axial-vector combinations of the four-vectors in the decay. Each of these components is then scaled by a Lorentz invariant scale factor, which is the so-called form factor.

In the case of $B \to X_u^P \ell\nu$ where X_u^P is a psuedoscalar, the only four-vectors available are those of the initial and final meson, p_B^μ and $p_{X_u^P}^\mu$. Consequently we can only construct vectors $(p_B - p_{X_u^P})^\mu$ and $(p_B + p_{X_u^P})^\mu$. We can use these two vectors to write the vector portion of the hadronic current as

$$\langle p_{X_u^P} | V^\mu | p_B \rangle = f_+(q^2)(p_B + p_{X_u^P})^\mu + f_-(q^2)(p_B - p_{X_u^P})^\mu. \qquad (2.7)$$

The Lorentz invariant form factors f_+ and f_- scale the vector components. The pseudoscalar to pseudoscalar decay has no contributing axial vector component. Furthermore, conservation of angular momentum allows only the zero helicity state of the W. Substituting the expression above into Equation 2.3 we can identify for the pseudoscalar final state in the massless lepton limit

$$|H_\pm|^2 = 0 \tag{2.8}$$

$$|H_0|^2 = 4p_{X_u^P}^2 \frac{M_B^2}{q^2}\left|f_+(q^2)\right|^2 \tag{2.9}$$

and write the differential rate *to pseudscalar final states*, X_u^P, integrated over the lepton decay angle, θ_{Wl}, as

$$\frac{d\Gamma\left(B \to X_u^P \ell\nu\right)}{dq^2} = \frac{G_F^2 |V_{ub}|^2}{24\pi^3} p_{X_u^P}^3 \left|f_+(q^2)\right|^2. \tag{2.10}$$

For a vector meson, X_u^V, the polarization, ϵ, provides an additional four-vector from which we can construct the hadronic current. Analogously we can write the vector and axial-vector portions of the hadronic current as

$$\langle p_{X_u^V}, \epsilon|V_\mu|p_B\rangle = ig(q^2)\epsilon_{\mu\nu\rho\sigma}\epsilon^{*\nu}(p_B + p_{X_u^V})^\rho(p_B - p_{X_u^V})^\sigma \tag{2.11}$$

and

$$\langle p_{X_u^V}, \epsilon|A_\mu|p_B\rangle = f(q^2)\epsilon_\mu^* \quad + \quad a_+(q^2)(\epsilon^* \cdot p_B)(p_B + p_{X_u^V})_\mu$$
$$+ \quad a_-(q^2)(\epsilon^* \cdot p_B)(p_B - p_{X_u^V})_\mu, \tag{2.12}$$

where the form factors g, f, a_+, and a_- scale the individual components. Assembling the expressions above into the $V^\mu - A^\mu$ form and using Equations 2.3 we write the W helicity amplitudes for a final state vector particle, X_u^V, in the massless lepton limit as

$$|H_\pm|^2 = \left[f\left(q^2\right) \mp 2M_B p_{X_u^V} g\left(q^2\right)\right]^2 \tag{2.13}$$

$$|H_0|^2 = \frac{M_B^4}{4q^2 M_{X_u^V}^2}\left[\left(1 - \frac{1}{M_B^2}\left(M_{X_u^V}^2 + q^2\right)\right)f(q^2) + 4p_{X_u^V}^2 a_+(q^2)\right]. \tag{2.14}$$
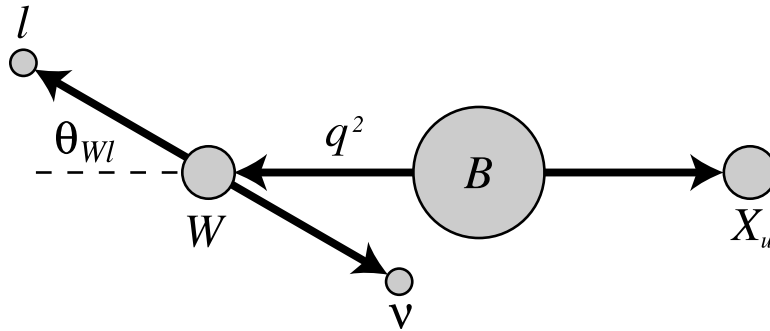
Figure 2.1: The relevant kinematic variables in $B \rightarrow X_u \ell \nu$ decay can are the four-momentum transfer to the lepton-neutrino system, q^2, and the decay angle of the lepton in the virtual W helicity frame, θ_{Wl}.

These expressions can be directly substituted into Equation 2.6 to express the differential decay rate for vector final states in terms of the lepton decay angle, θ_{Wl} and the three form factors, f, g, and a_+.

2.2   The Decay Form Factors

A key element of the rate calculation involves the theoretically challenging computation of the Lorentz invariant decay form factors. In the massless lepton limit of the pseudoscalar to pseudoscalar transition the only contributing form factor is $f_+(q^2)$. For the corresponding pseudoscalar to vector transition, three form factors, $f(q^2)$, $g(q^2)$, and $a_+(q^2)$, govern the decay. These form factors ultimately dictate the q^2 and, in the case of the vector final states, the θ_{Wl} dependance of the rate.

While much progress has been made in the theoretical community on techniques for calculating non-perturbative QCD interactions, at present the error on $|V_{ub}|$ is dominated by the uncertainty in the normalization of the form factor. From the

experimentalist's viewpoint, the shape of the form factors determine the overall signal reconstruction efficiency because the efficiency is typically not uniform in the kinematical variables. Therefore uncertainty in the shape produces a systematic error on the experimental measurement of the decay rate. In Section 5.1.2 I will discuss how this uncertainty is minimized.

2.2.1 Calculation Techniques

A variety of techniques have emerged for calculating form factors. The $b \rightarrow u$ transition is particularly difficult because the final state u quark is light and typically recoils with large momentum. The principles of Heavy Quark Symmetry (HQS) [10], which are useful in calculations of heavy-to-heavy $b \rightarrow c$ form factors, break down in the $b \rightarrow u$ case. Independent of HQS, there are several constituent quark model calculations available [11, 12, 13].

Lattice QCD is evolving as a method of directly computing the form factors to high precision. The action of the QCD Lagrangian is evaluated numerically on a lattice of discrete space-time points. In theory the lattice calculation provides a route to compute the form factors precisely because it can be done without approximation. Calculations were first done without the presence of light quarks and results were "chiraly extrapolated" to the actual light quark masses. The effects of quark loops were also ignored and the results were determined in the so-called "quenched" approximation. Recently progress has been made to overcome both of these hurdles. In particular we use the unquenched lattice calculations of Shigemitsu *et al.* for the $B \rightarrow \pi$ form factor in this analysis [16]. An additional limitation of the lattice calculations is that they are only valid at high q^2. Calculations are done in the rest frame of the B meson; therefore, at low q^2 the

high momentum of the recoil meson requires a prohibitively small lattice spacing to accurately compute the form factor.

The technique of Light Cone Sum Rules (LCSR) exploits the asymptotic freedom of QCD and provides complementary form factor data to that from lattice calculations. At low q^2 recoiling quarks are highly virtual, *i.e* on the "light cone," and QCD interactions are perturbative. Ball and Zwicky have used this technique to compute both $B \to \pi$ and $B \to \rho$ form factors [14, 15].

2.2.2   General Form Factor Behavior

At high q^2, the shape of the $B \to \pi$ form factor is dominated by the the presence of the B^* pole just beyond q^2_{max}. Figure 2.2 shows the unquenched lattice calculation of Shigemitsu *et al.* of the form factor as a function of q^2. The q^2 dependence of the rate is driven by the competition of f_+ with the p^3_π term in Equation 2.10.

The most striking features in the $B \to \rho$ form factors emerge from the relative populations of the three W helicity states. The left-handed nature of the weak interaction enhances the H_- component; therefore, the lepton decay angle spectrum favors the $(1 + \cos\theta_{Wl})^2$ shape (Equation 2.6). This ultimately results in a harder lepton momentum spectrum in $B \to \rho\ell\nu$ decays than in $B \to \pi\ell\nu$ decays. Calculations for the $B \to \rho$ form factor by Ball and Zwicky using LCSR are shown in Figure 2.3 [15]. The suppression of the H_+ W helicity state is clearly visible. In Section 5.1.2 I will discuss how we minimize our sensitivity to the lepton decay angle spectrum as derived from the relative strengths of the H_- and H_0 components which vary among different theoretical predictions.
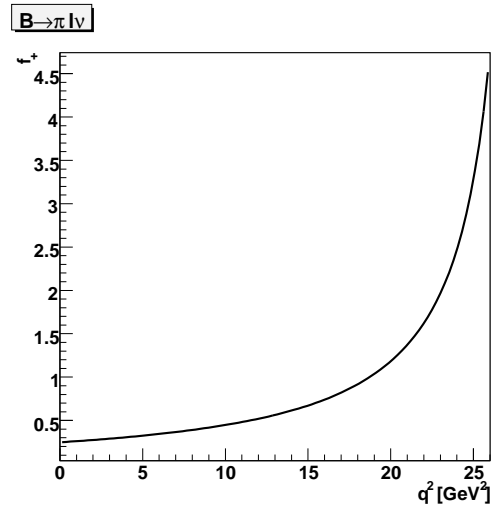
Figure 2.2: The $B \to \pi$ form factor, $f_+(q^2)$, as calculated by Shigemitsu *et al.* [16]. The presence of a pole at $M_{B^*}^2$ dominates the shape.
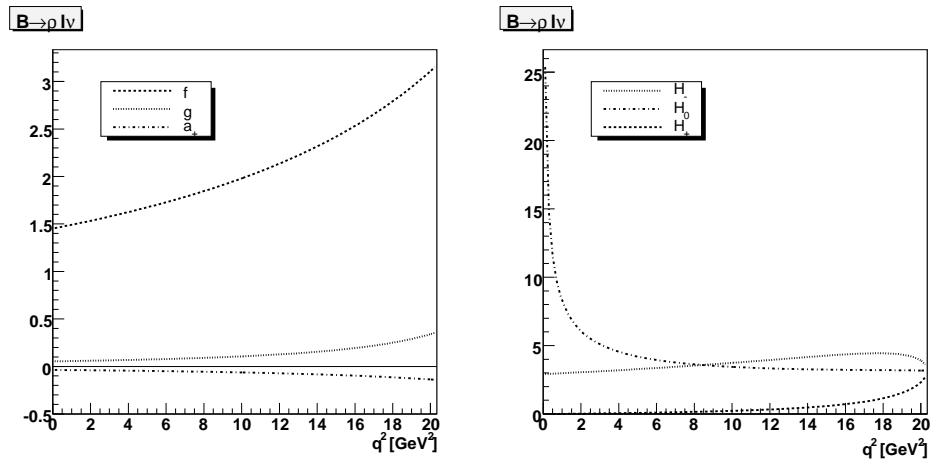
Figure 2.3: The $B \to \rho$ form factors plotted as a function of q^2 for f, g, and a_+ (left) and in the virtual W helicity basis (right).

In summary, we are now equipped with an understanding the importance of $|V_{ub}|$ and how we can access $|V_{ub}|$ through semileptonic decay of B mesons. The challenge that lies ahead is to measure the exclusive decay rate for $B \rightarrow \pi \ell \nu$ and $B \rightarrow \rho \ell \nu$, since this is the critical experimental input required for a precision measurement of $|V_{ub}|$. We strive to do this rate measurement in a way that is insensitive to the uncertainty in the decay form factors.

Chapter 3

Experimental Apparatus

With just the lightest quarks, u and d, and lightest lepton, the electron, all of the visible atoms of the universe can be constructed. The heavier quarks, like the b quark, and the phenomenology of CP violation discussed in the opening chapter influenced the evolution of the universe only at very early stages. Through the violent collisions of accelerated subatomic particles, we can recreate these conditions of the early universe in the laboratory and study the underlying physics.

3.1   The Cornell Electron Storage Ring

The Cornell Electron Storage Ring (CESR) is electron-positron storage ring with a circumference of 768 meters located on the Cornell University campus. A schematic drawing of the machine is shown in Figure 3.1. Electrons are produced and accelerated to roughly 200 MeV down the thirty-meter linac and injected into the synchrotron. Once in the synchrotron, the beam is accelerated to the full 5+ GeV and subsequently transferred to the storage ring. The process continues until the storage ring is full of electrons. Positrons are then produced as byproducts of the collision between the electron beam and a tungsten plate inserted into the linac. The positrons are collected and accelerated down the remainder of the linac before being injected into the synchrotron and the storage ring. The beams rotate in opposite directions within the same beam pipe following what is known as a "pretzel" orbit in order to avoid collisions away from the interaction region. The beams are steered into collision in the middle of the CLEO detector with a
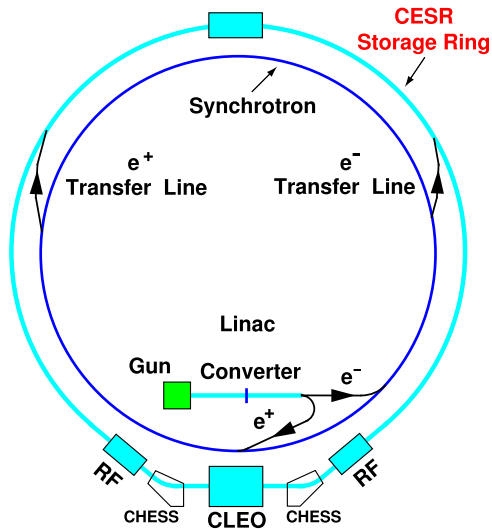
Figure 3.1: A schematic of the CESR machine.

total center of mass energy high enough to produce the $\Upsilon(4S)$ resonance, which immediately decays into a pair of B mesons.

3.2 The CLEO Detector

Data collected over roughly ten years with three different configurations of the CLEO detector is used for this analysis. While the individual detector components and performance have changed significantly over time, the fundamental principles and functionality of CLEO have remained constant. Like all particle physics detectors CLEO consists of host of specialized sub-detectors that work together to produce a complete picture of the products of an e^+e^- collision. Figures 3.2 and 3.3 provide a sketch of the major components of the CLEO II and III detectors. One can see that the shape and makeup of the detectors are similar. Let us tour of the CLEO detector from the interaction region radially outward, briefly discussing each of the detecting elements.
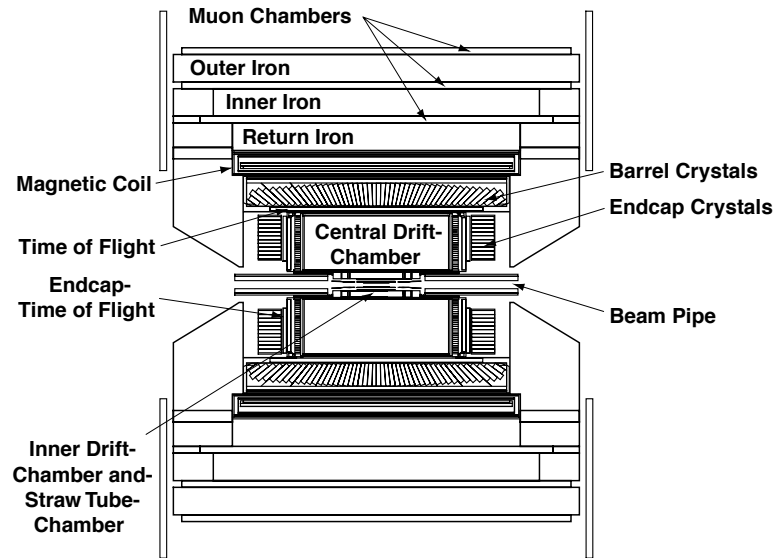
Figure 3.2: A side-view of the CLEO II detector.
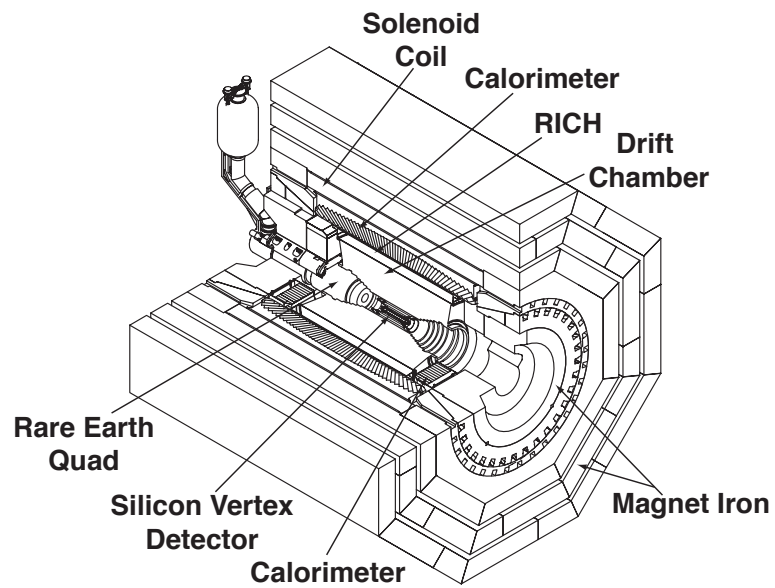
Figure 3.3: A 3D cutaway view of the CLEO III detector.

3.2.1 Charged Particle Tracking

Charged particle tracking devices occupy the first meter in radius from the beampipe. Charged particles leaving the interaction regions travel in helical trajectories due to a uniform magnetic field produced by a super-conducting solenoid positioned outside of the crystal calorimeter. A precision measurement of the trajectories of the decay products of a particle allows us provides the information necessary to reconstruct the kinematic variables of the parent.

All charged particle tracking devices in CLEO rely on ionization as their fundamental means of particle detection. In drift chambers, charged particles ionize the gas in the volume of the chamber. Electrons then subsequently drift to anode, or "sense", wires held at a couple thousand volts. As the electron nears the wire it is accelerated in the increasing electric field and it ionizes additional gas molecules, creating an electron avalanche of about 100,000 electrons at the wire. The electrical pulse then travels down the wire and is amplified and recorded by readout electronics. In silicon strip detectors this ionization produces electron/hole pairs in the bulk of a reverse-biased pn junction, and the resulting current is sensed by the strip providing the bias voltage.

In all three configurations of the CLEO detector (II, II.V, and III), a specialized tracking device was used to aid in the reconstruction of particles within tens of centimeters of the beam pipe. Ideally one would like an device capable of high resolution measurements that allow the separation of the two B vertices in the event. This desire must be balanced with the inherently noisy and intense radiation environment next to the beam pipe. CLEO II utilized a straw tube drift chamber with tubes running parallel to the beam pipe. Straw tubes consist of an anode wire placed axially in a cathode tube, a design suitable for the high rate environment

near the beam pipe. Both the CLEO II.V and III configurations utilized silicon strip detectors to accomplish high precision tracking near the interaction region [22, 23]. Wafers of *pn* doped silicon embedded with sensing strips were arranged in a multi-layer cylindrical pattern about the beam pipe. Strip spacing on wafers is at the 50-100 μm level allowing for precision position measurement on the order of tens of microns for tracks that pass through the wafer.

The majority of the tracking volume of all three configurations of the CLEO detector was occupied with an open-cell drift chamber [20, 21]. This design used cathode and anode wires strung parallel to the beam pipe to create an array of drift cells. Each cell is composed of a three by three grid of of wires with the central wire being the anode, or sense, wire, and the surrounding eight wires are cathode, or field, wires. When a charged particle passes through the cell, electrons drift to and avalanche at the sense wire leaving a pulse. Precision pulse time measurements record the total drift time and therefore allow determination of exactly where the particle passed through the cell. Position resolution at the 100 μm level, nearly a hundred times smaller than the overall cell size, can be achieve with this technique.

3.2.2   Particle Identification

Information about the identity of the particles can be gleaned from several detectors. Ionization per unit length, dE/dx, measured in the using the pulse heights on sense wires in the drift chamber provides a direct measurement of a particle's velocity. Charged particles loosing energy through ionization do so as a function of their mass, M, and velocity, β, according to the Bethe-Bloch equation [6]:

$$-\frac{dE}{dx} = \kappa z^2 \frac{Z}{A}\frac{1}{\beta^2}\left[\frac{1}{2}\ln\frac{2m_e c^2 \beta^2 \gamma^2 T_{\max}}{I^2} - \beta^2 - \frac{\delta}{2}\right]. \tag{3.1}$$

where T_{max} is the maximum kinetic energy transfer of the charged particle to an ionization electron in the gas volume of the drift chamber:

$$T_{\mathrm{max}} = \frac{2m_e c^2 \beta^2 \gamma^2}{1 + 2\gamma m_e/M + (m_e/M)^2}. \tag{3.2}$$

A momentum measurement from the drift chamber allows the the expression for dE/dx to be cast strictly as a function of the charged particle mass, and therefore a measurement of dE/dx can be used to determine the mass of the particle.

To supplement the particle identification information from the drift chamber, a time-of-flight (TOF) counter was utilized in the CLEO II and II.V detectors. Charged particles passing through this cylindrical arrangement of scintillator bars outside of the drift chamber produce light that is observed by a phototube. Precision measurement of the time of this light pulse with regard to the beam crossing time coupled with the path length measurement in the drift chamber provided a measurement of β which can be combined with momentum data to determine the identity of the particle.

The TOF counter was replaced by a Ring Imaging Cherenkov (RICH) detector in CLEO III [24]. Cherenkov radiation is emitted when charged particles traveling through a medium with an index of refraction n exceed the the speed of light in the medium, which is given by $\frac{1}{n}c$. The useful characteristic of the radiation that it is emitted along a cone about the velocity vector of the particle with an opening angle, θ_c, given by [6]

$$\cos\theta_c = \frac{1}{\beta n}. \tag{3.3}$$

The RICH detector consists of a LiF radiator in which the cone of Cherenkov radiation is produced. The cone subsequently expands in a short expansion volume, before the photons enter a chamber that contains a photosensitive gas. The electrons produced by the interaction of the Cherenkov photons with the gas drift to

anode wires and produce, just as in a drift chamber, an avalanche. Unlike the drift chamber however, the anode wires themselves are not read out. In the RICH detector pixelated cathode pads near the anodes are read out to give a two-dimensional image of the pattern of avalanches. From this image and careful knowledge of the geometry and track trajectory the Cherenkov cone opening angle and therefore particle velocity, β, can be determined.

3.2.3   Calorimetry

Photons and other neutral particles will escape the previously mentioned detectors because they are incapable of depositing energy through ionization. Photons are absorbed by a calorimeter of Thallium-doped Cesium-Iodide (CsI) scintillating crystals located out outside of the tracking volume and the supplementary particle identification detectors [25]. Photons entering a crystal produce a shower of electrons and photons through the repeating processes of pair-production and bremsstrahlung radiation. Due to the scintillating properties of the crystal, the intermediate electrons produce light that is registered by a photo-diode that is optically coupled to the crystal. The entire electromagnetic shower is contained within a small array of neighboring crystals that can be clustered together in order to find the precise location and energy of the incident photon.

Similar to photons, electrons shower electromagnetically in the calorimeter. We use this feature in conjunction with the presence of a drift chamber track pointing at the shower to identify electrons. Since the energy loss due to bremsstrahlung radiation is proportional to $1/m^2$, heavy muons pass through the calorimeter without showering and therefore only leave behind trace amounts of energy.

While hadrons interact with the calorimeter via the electromagnetic processes mentioned above, they also strongly interact with the nuclear matter in the crystals. These nuclear interactions result in the production of a variety of secondary hadrons. Some fraction of these secondary hadrons will be π^0s which immediately decay into two photons and therefore produce a photon-like shower. However, other secondary hadrons, such as charged pions, may even exit the crystal, traveling to neighboring crystals to produce an additional shower. For neutral hadrons such as K_L^0s and neutrons, hadronic showers provide the only kinematic information on the particle and consequently their energy is more poorly determined than photons or charged particles.

3.2.4 Muon Detection

In order to identify muons we search for tracks in proportional wire chambers outside of the calorimeter and sandwiched between layers of iron. Muons, which only minimally interact via the electromagnetic force, are capable of penetrating this massive amount of material and leaving signals in the muon chambers. A drift chamber track that points to one of these track stubs in the muon chamber is evidence that the charged particle was a muon.

3.3 Data Summary

The peak cross section of the process $e^+e^- \to \Upsilon(4S)$ is roughly one nb. Therefore for every fb^{-1} ($= 10^6$ nb^{-1}) of luminosity delivered by CESR approximately 2×10^6 B meson decays are observed by CLEO. Data taking is split between running with center of mass energy at the $\Upsilon(4S)$ mass and at 60 MeV/c^2 below the

Table 3.1: A summary of the three data sets used in this analysis. Due to variations in the beam energy the ratio of on-resonance luminosity to number of $B\bar{B}$ events is not constant.

Detector Configuration	$\mathcal{L}^{\text{on}}$ [fb^{-1}]	$\mathcal{L}^{\text{off}}$ [fb^{-1}]	$N_{B\bar{B}}$ [$\times 10^6$]
CLEO II	3.1	1.6	3.3
CLEO II.V	6.0	2.9	6.4
CLEO III	6.3	2.3	5.7
Total	**15.5**	**6.9**	**15.4**

$\Upsilon(4S)$, below $B\bar{B}$ threshold. The smaller, latter data set is used to understand the continuum processes that occur in addition to $\Upsilon(4S)$ production. Table 3.1 summarizes the integrated luminosity ($\mathcal{L}$) and number of $\Upsilon(4S) \rightarrow B\bar{B}$ events used in this analysis. Figure 3.4 shows the data collected per month throughout the life of the CLEO II, II.V, and III detectors. One notices an increasing trend in integrated luminosity as a function of accelerator development.
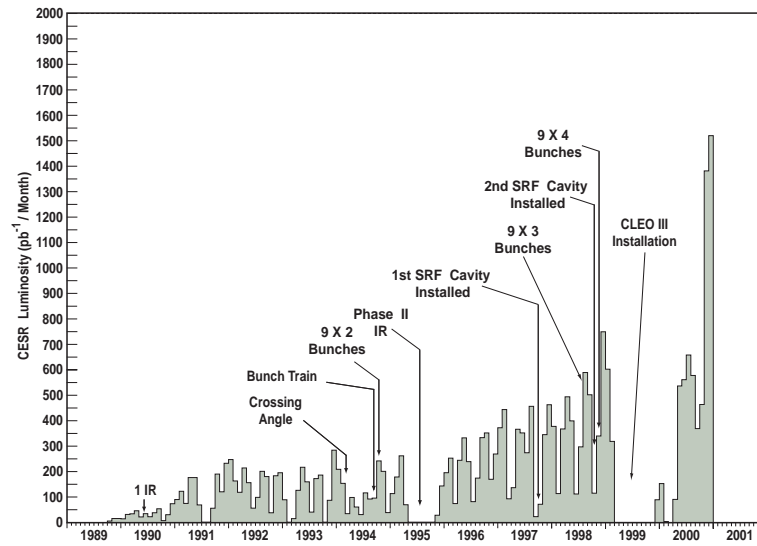
Figure 3.4: Data collected per month with the CLEO II (pre-1995), II.V (1995-1999), and III (post-1999) detectors. The plot highlights the how the integrated luminosity grows with accelerator developments.

Chapter 4

Experimental Technique

This chapter discusses the reconstruction and selection of the $B \to X_u \ell \nu$ candidates given the raw information from the detector. The idea is simple: create an algorithm that preserves as many true $B \to X_u \ell \nu$ decays as possible while rejecting fake candidates. The challenge arises in creating an implementing such an algorithm.

The only experimentally viable route to obtaining a value of $|V_{ub}|$ is through the semileptonic charmless decays of B mesons which are complicated by the undetectable neutrino in the final state. A key component of this analysis is therefore an algorithm which allows the neutrino to be "reconstructed" from the missing four-momentum in the event. Specifically

$$p_{miss} = p_{initial} - \sum p_{visible}, \tag{4.1}$$

where

$$p_{initial} = (2E_{beam}; -2\sin\theta_c E_{beam}, 0, 0) \tag{4.2}$$

is the initial four-momentum of the the two beams[1]. Ideally p_{miss} is strictly the four-momentum of the neutrino in the signal decay. In reality, however, p_{miss} may include multiple neutrinos, K_L^0's, or neutrons that go undetected along with other charged tracks and photons that either miss our detector or we have reconstructed improperly. To eliminate these events with flawed neutrino reconstruction, we must maximize the resolution of the reconstructed visible four-momentum in the event which, in turn, maximizes the neutrino resolution. As the neutrino resolution

---

[1]θ_c is the small (≈ 2 $mrad$) crossing angle of the beams

increases the kinematic requirements we place on the reconstructed $X_u \ell \nu$ become more effective at separating the signal events from the background events.

Candidate Reconstruction

In the following sections I will walk through the stages of candidate reconstruction with an eye towards optimizing the resolution on the four-momentum of visible particles produced in the $e^+ e^-$ collision. Initially we work to refine the raw information produced by the detector. From this refined set of visible objects we can then reconstruct the lepton, hadron, and ultimately neutrino daughters of the B decay.

4.1 The Fundamental Objects: Tracks and Showers

We can reduce every reconstructed particle in the detector down to a combination of two fundamental objects: tracks and showers. Ideally, we would like that each "track" corresponds to the trajectory of a single charged particle produced in an $e^+ e^-$ collision. We would like, similarly, a "shower" to correspond to the energy deposited in the calorimeter by a single neutral particle. The spatial location of the shower and the assumption that it came from the interaction region, we can deduce the trajectory of the neutral particle that produced it. In reality, such an ideal list of tracks and showers is not simple to produce.

4.1.1 Tracks

To enumerate the particles produced in the collision our goal is this: *count once and only once every charged particle coming from the interaction point.* We rely on the large acceptance of our tracking chamber to try to count *every particle.* Unfortunately there are many ways to double count particles, listed below are the leading contributors:

- Since the chamber is inside of a magnetic field, charged particles with low transverse momentum will curl inside of the detector. The pattern recognition software used to find tracks will frequently find multiple tracks from one curling track as showin in Figure 4.1. This is problematic since it leads to multiple counting of the same physical particle.

- Some particles, such as charged kaons, decay in flight. The decay produces a secondary charged particle with a different momentum than the parent and therefore the track appears to have a kink in it. The pattern recognition will identify both parts of the track as separate tracks.

- Occasionally, in the case of decays-in-flight or hard scattering where the kink is small, two tracks will be found. One contains the innermost and outermost hits while the other contains the hits around the kink. This "ghosting" effect produces two tracks with similar trajectories for the same physical particle.

Significant work has gone into the development of an algorithms packaged as *Trkman*, to recognize and remove these spurious tracks [26]. Tracks that are not flagged as curlers, ghosts, or decays-in-flight are said to be *TrkmanApproved.* From now on "tracks," unless explicitly stated, refer to *TrkmanApproved* tracks.
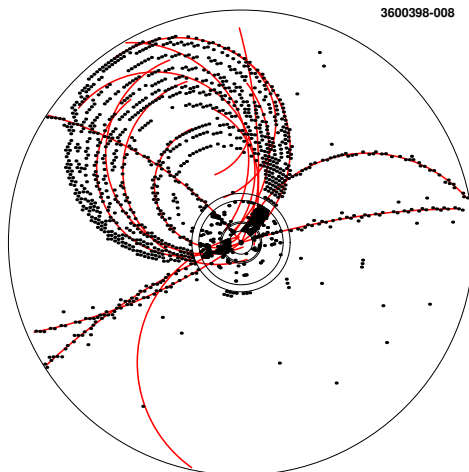
Figure 4.1: The pattern recognition code can find multiple tracks for a single curling particle.

4.1.2 Showers

Recall that the goal of the calorimeter is to measure every photon leaving the interaction region. As with tracks, there are methods of producing extra showers in the calorimeter that are not associated with photons coming from the interaction of the beams.

- All charged particles will deposit some energy in the calorimeter through ionization, hadronic interactions, or electromagnetic showers.

- Hadronic interactions within the calorimeter itself or the material just in front of the calorimeter can create particles that produce an additional separate secondary shower away from the primary shower. We call such showers *Splitoff* showers.

Showers produced by the first mechanism are eliminated by geometrically matching observed tracks and showers. Recall that in the case of decays-in-flight or

hadronic interactions in the tracking volume, one actually wants to match the secondary tracks, the ones produced by the decay or interaction, to showers in the calorimeter.

Eliminating the second source, the so-called splitoff showers, is more difficult. These showers are too far from any track projection to be removed by a simple proximity requirement without significant loss of real photon showers through accidental vetos. However, obtaining optimal missing energy resolution in neutrino reconstruction depends on their removal. Extensive documentation on the identification algorithm for Splitoff showers can be found elsewhere [27]. In summary the algorithm first classifies showers according to their location in the calorimeter, shower energy, and whether the shower is isolated. For each of these classes then information about the shower shape, location, and, if it exists, parent location is calculated and fed into a neural net for that classification. Showers that have a net output that is more splitoff-like are then removed from the event. We will later explore biases in the neural net algorithm as a potential systematic error.

4.1.3 K_S^0's and other "Vees"

The decay of neutral kaons into charged tracks deserves special consideration. On average there are roughly 0.6 K_S^0's per event. Because of the relative long K_S^0 lifetime, this decay frquently occurs after the K_S^0 has traveled a significant distance from the origin. This poses two problems for our reconstruction algorithm. First, because the charged tracks are displaced from the origin they are at a higher risk of being rejected by the *Trkman* track selection algorithm. Since the two charged tracks are all that one has to represent the K_S^0, losing one of them would mean an underestimate of the visible energy and momentum produced in the collision.

Secondly, even if the tracks survive the selection cuts the momentum used for each track is that of the track extrapolated back to the interaction point which is not a correct representation of the momentum of the particle.

To solve both of these problems we identify K_S^0's before filtering out the *TrkmanApproved* tracks. We then the kinematically refitted K_S^0 momentum and energy directly and eliminate the daughter tracks and any curler or ghost track associated with them.

To select K_S^0's we first locate oppositely charged tracks that appear to intersect at a point displaced from the origin. These tracks are then fit to a common origin in 3D space [28]. The χ^2 of such a fit describes the likelihood that the two tracks originated from the from the same point, $\vec{d}$, where $\vec{d}$ is measured from the beam spot. The fit also produces an error matrix, V, for $\vec{d}$. In addition the beam spot has some physical size that is approximately Gaussian in 3D and represented by an error matrix $\mathbf{B}$. With the "flight significance," FS, as

$$FS = \sqrt{\vec{d} \cdot (\mathbf{V} + \mathbf{B})^{-1} \cdot \vec{d}}. \tag{4.3}$$

we required $FS > 7.5$ to positively identify a K_S^0. It is worthwhile to note that those K_S^0's that travel farthest will be easiest to identify. Fortunately, those same K_S^0's are the most problematic for our neutrino reconstruction algorithm. For each selected K_S^0, we repeat the kinematic fit with the additional constraint that the momentum to point back at the beam spot, and all candidates that do not have a χ^2 for three degrees of freedom less than ten are dropped. The resulting $\pi^+\pi^-$ invariant mass spectrum, fitted to a double Gaussian plus a linear background, is shown in Figure 4.1.3. We keep candidates within 3σ of the K_S^0 mass, where σ is the width of the narrower, core Gaussian. These candidates are fit one last time with the invariant mass of the $\pi^+\pi^-$ constrained to that of the nominal K_S^0
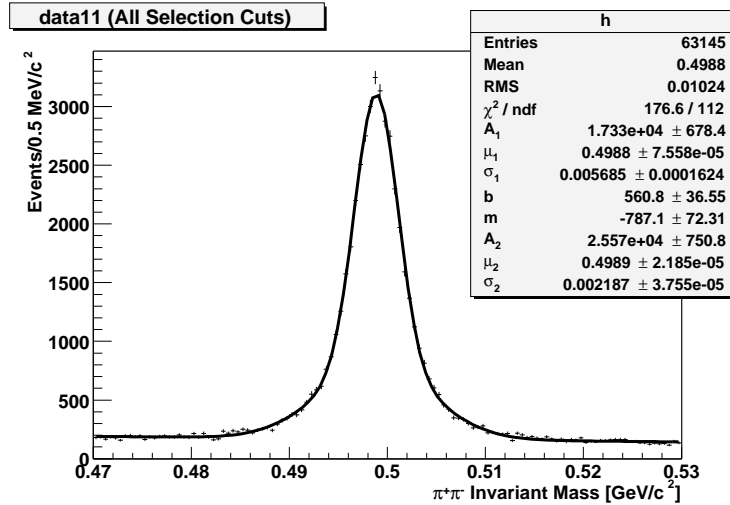
Figure 4.2: Typical $\pi^+\pi^-$ invariant mass distribution for K_S^0 candidates.

mass. The resulting four-momentum then replaces the daughters, along with all associated tracks and showers, in the subsequent reconstruction of the event.

One might ask whether such a sophisticated procedure should also be used for the decay $\Lambda \to \pi p$. The average number of Λ's per event is roughly thirty times less than the average number of K_S^0's so there is little to gain by fitting Λ's in a similar fashion. It should be noted that after all of the careful work the effect of "proper" K_S^0 reconstruction makes a nearly imperceptible difference in the reconstructed ν resolution. A more significant gain of K_S^0 reconstruction is the elimination of tracks from the event which in turn reduces combinatoric backgrounds in candidate reconstruction.

In summary, recall that initial goal of our procedure was a measurement of the momentum of every particle coming from the collision without any double counting. To accomplish this, we first identify K_S^0's in the event, removing all tracks and showers associated with the daughters of the K_S^0's. Next, we attempt to filter out the double counted tracks: ghosts, curlers, and decays-in-flight. Finally

we remove showers that are associated with particles already measured in the tracking system.

4.2 Particle Identification

We assume that remaining isolated showers are photons. Tracks, on the other hand, can be produced b a number of types of charged particles. Particle identification is necessary to select the semileptonic decay products calculate the observed four-momentum in the event that is used in neutrino reconstruction. This section summarizes how we distinguish among the many species of charged particles for each track.

4.2.1 Lepton Identification

One of the key signatures of the semileptonic $b \to u$ decay is the presence of a relatively high momentum lepton in the final state; therefore, we take care to identify leptons with high efficiency. Furthermore, additional leptons in the event usually indicate more than one neutrino was present in the event. This is problematic for neutrino reconstruction and we veto events with additional so-called "counting," as opposed to "signal," leptons. A high lepton fake rate will cause more events to be unnecessarily vetoed.

Electron Identificaiton

Electron identification has long been a recognized sport at CLEO. We use the accepted *Rochester Electron ID* algorithm. This algorithm was first developed for

CLEO II [29] and was recently revised and improved for use with CLEO II and CLEO III [30]. Here I review the core ideas of the algorithm and discuss how it is applied for this analysis.

The primary identifying mark of an electron is that it showers electromagnetically in the calorimeter. Therefore electrons tend to deposit all of their energy in the calorimeter in a photon-like shower. The hadronic interactions of pions, kaons, and protons with the calorimeter are much different because the characteristic nuclear interaction length of a hadronic shower is larger than that of an electromagnetic shower. Showers produced by hadronic interactions are distributed across a larger number of crystals with larger fluctuations in the energy deposition profile across the crystals. We can pick a set of variables then to discriminate between electrons and hadrons:

- E/p: The ratio of the energy of the shower to the momentum of the track as measured in the drift chamber. For reasons discussed above this is near unity for electrons and typically smaller for hadrons.

- $\chi_e^{dE/dx}$: Based on specific ionization data from the tracking chamber, this is the number of standard deviations that the dE/dx measurement for a particular track deviates from the nominal dE/dx of an electron with that track's momentum. The $\chi_e^{dE/dx}$ will generally be near zero for electrons and away from zero for hadrons.

- w: The RMS width of the shower – narrower for electrons than hadrons.

- $\Delta_\theta, \Delta_\phi$: The deviation of the depth-corrected shower center from the extrapolated track trajectory in the θ and ϕ directions. The large fluctuations in the energy profile of hadronic showers produce fluctuations in the reconstructed

locations of the shower and therefore Δ_θ and Δ_ϕ will be distributed more broadly about zero for hadrons than for electrons.

A pure sample of electrons and positrons can be obtained using radiative Bhabha events in order to compute the distributions in data for these key variables. Likewise, we can use the decay $D^* \to D\pi_s^\pm$, $D \to K^\mp\pi^\pm$, where the charge of the intermediate "slow" pion (π_s) tags the identity of the two daughter hadrons, to obtain pure samples of pions and kaons. These true distributions are then normalized to unit area to form probability distribution functions (pdf's). The pdf's are calculated in bins of track momentum in order to accommodate changes in shower shape as a function of momentum. A set of typical pdf's is shown in Figure 4.3

For a given track, we can combine the pdf's with the measured momentum (p) and key variable values (x_i) to obtain the ratio $\mathcal{L}$ of the probabilities that the track is an electron or a hadron. More formally, $\mathcal{L}$ the likelihood ratio, is given by

$$\mathcal{L} = \frac{\mathcal{P}(x,p)_e}{\mathcal{P}(x,p)_h}, \tag{4.4}$$

where $\mathcal{P}(x,p)_s$ is the probability that the track is an electron ($s = e$) or hadron ($s = h$). In turn,

$$\mathcal{P}(x,p)_s = \prod_i f_i^s(x_i;p), \tag{4.5}$$

where $f_i^s(x_i;p)$, is the pdf for a particle of type s in the ith key variable. In practice we compute $ln\ \mathcal{L}$ which turns the products into sums. Because this ratio is large for tracks that are electron like, we place a minimum requirement on $ln\ \mathcal{L}$ to identify electrons. We weigh the electron identification efficiency with the probability that a pion will fake an electron in determining the minimum acceptable value of $ln\ \mathcal{L}$. Specifically we find the minimum $ln\ \mathcal{L}$ such that the probability that a true pion will be misidentified as an electron is 0.2% (1.0%) in the barrel calorimeter (elsewhere).
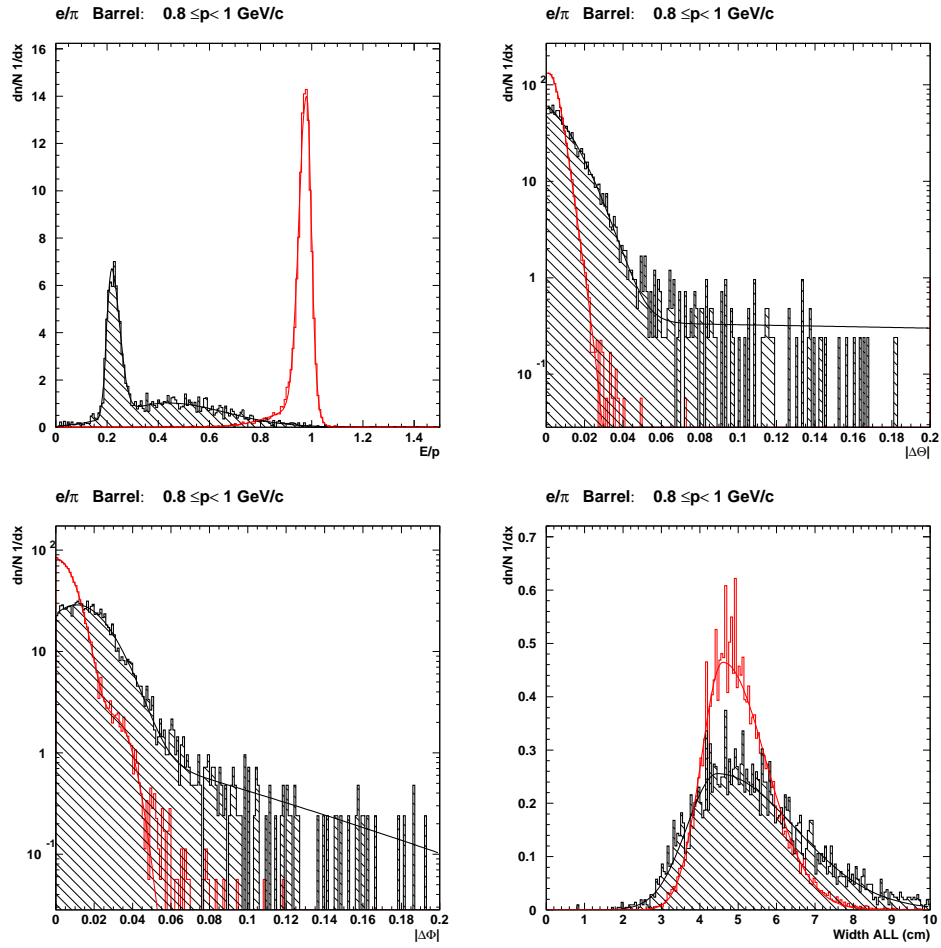
Figure 4.3: Typical pdf's for electron identification variables [30]: electron pdf shapes are shown by the open histogram, hadron shapes are shown hatched histogram.

Additionally we use time of flight (RICH) information in CLEO II (III) to veto tracks that may pass this minimum likelihood cut but otherwise look like hadrons. In CLEO II we can compute difference in χ^2 for the electron and kaon hypotheses as

$$\Delta\chi^2_{e/K} = \chi_e^{2^{TOF}} - \chi_K^{2^{TOF}}. \tag{4.6}$$

For true electrons this quantity tend negative, while for true kaons the $\chi_K^{2^{TOF}}$ will be at a minimum and the difference will tend to be positive. We require $\Delta\chi^2_{e/K} < 10$ to veto kaons that may fake electrons.

Using the RICH detector in CLEO III we can compute the likelihood that a particular track is an electron, pion, proton, or kaon. This allows us to reduce the probability that hadrons will fake electrons, especially in momentum regions where the dE/dx seems ambiguous. If the likelihood that particle of momentum p is of species, s, as determined by RICH information is $\mathcal{L}_s$ then we require:

$$\mathcal{L}_e > \mathcal{L}_\pi \quad \text{if} \quad p < 500 \ MeV/c, \tag{4.7}$$

$$\mathcal{L}_e > \mathcal{L}_K \quad \text{if} \quad 500 \ MeV/c < p < 800 \ MeV/c, \ \textbf{and} \tag{4.8}$$

$$\mathcal{L}_e > \mathcal{L}_p \quad \text{if} \quad 900 \ MeV/c < p < 1.7 \ GeV/c. \tag{4.9}$$

Tracks with momenta greater than 200 MeV/c that pass all of the criteria listed above are declared "electrons." The identification efficiency as a function of momentum is shown in Figure 4.4. From this set of electrons we select the "signal electrons" that will be used in reconstructing the exclusive $B \to X_u e\nu$ decays. To qualify as a signal electron the electron must also:

- have a momentum greater than 1.0 GeV/c

- have hits in at least 40% of the drift chamber layers the that track passed through
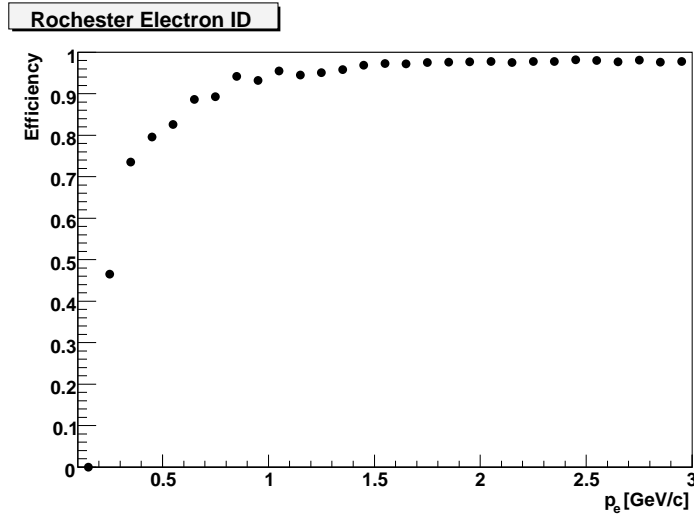
Figure 4.4: Electron identification efficiency as a function of momentum using the Rochester Electron ID algorithm. Data shown are for CLEO III – efficiencies in CLEO II and II.V are similar.

- have a distance of closest approach to the beam spot less than 2 mm in the plane transverse to the beam axis

- have a distance of closest approach to the beam spot less than 5 cm is the z direction.

These criteria yield a relatively pure sample of signal electrons with efficiency typically greater than 95% for electrons resulting from semileptonic $b \to u$ decay.

Identifying Muons

Much like electrons, muons have characteristic qualities that make them easy to identify. They are heavy particles that interact only via the electromagnetic and weak forces. At momenta relevant for CLEO, the primary mechanism of energy loss for muons traveling through material is ionization. As a result, muons can

penetrate through much more material than electrons or hadrons, whose ranges are limited by the much strong electromagnetic and additional hadronic interaction, respectively. We denote the depth that the muon penetrates by the number of hadronic interaction lengths, $\mathcal{X}_\mu$. Our muon chamber allows coverage over region $|\cos\theta| < 0.85$ (0.65) for CLEO II (III).[2] If a track within this fiducial volume that has not been identified as an electron satisfies

- $p > 1.2\ GeV/c$ **and** $\mathcal{X}_\mu \geq 5$ *or*

- $1.0\ GeV/c \leq p < 1.75\ GeV/c$ **and** $3 \leq \mathcal{X}_\mu < 5$,

we classify it as a "counting muon." Of these muons, we select a subset of "signal muons" for reconstructing the $B \to X_u \mu \nu$ decays that:

- have $\mathcal{X}_\mu \geq 5$,

- have hits in at least 40% of the drift chamber layers the that track passed through,

- have a distance of closest approach to the beam spot less than $2\ mm$ in the plane transverse to the beam axis,

- and have a distance of closest approach to the beam spot less than $5\ cm$ is the z direction.

The efficiencies for these "counting" and more-restrictive "signal" muons are shown in Figure 4.5. Because softer muons will lose all of their momentum through specific ionization before penetrating deeply into the detector, the $\mathcal{X}_\mu \geq 5$ requirement greatly reduces the identification efficiency for signal muons below $1.5\ GeV/c$.

---

[2]The reduced muon identification volume in CLEO III is due to excessive beam related noise in the endcap muon chambers that posed a problem with reconstruction.
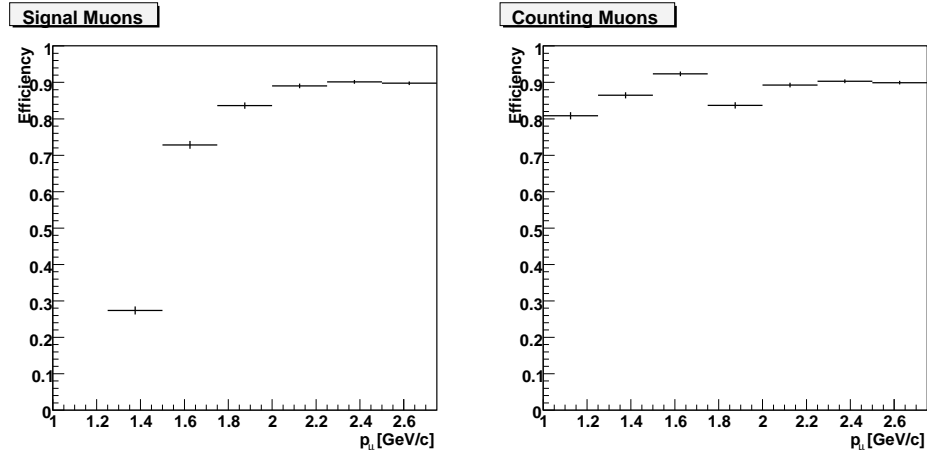
Figure 4.5: Muon identification efficiency as a function of momentum for signal muons (left) and counting muons (right). Data shown are for CLEO III – efficiencies in CLEO II and II.V are similar.

4.2.2 Hadron Identification

Once the electrons and muons have been identified, we then classify the remaining tracks in the event as pions, kaons, or protons using two pieces of information. First we use the probability that a particle is a particular species, which we obtain by combining the dE/dx in the drift chamber with the time of flight or RICH information. Let's call this the "detector" probability $\mathcal{P}_D^s$ that the particle is of type s. The second piece of information is the relative multiplicites of pion, kaons, and protons in generic B decay. That is, given a sample of charged particles at a particular momentum from a large number of B decays, what fraction will be pions, what fraction kaons, and what fraction protons? We will weight the probabilities based on the detector information by these production fractions, which we denote as $\mathcal{P}_P^s$. How do we determine $\mathcal{P}_D^s$ and $\mathcal{P}_P^s$?

For a set of true pions, kaons, or protons of a given momentum the time of flight and dE/dx will be Gaussian distributed variables. The raw dE/dx and time

of flight information in CLEO II can therefore be used to calculate an absolute probability that a particle is a particular species. If either the dE/dx or time of flight information is missing, for example, if the particle curls in the magnetic field and does not reach the time of flight counters, then we calculate $\mathcal{P}_D^s$ solely from the available measurement.

In CLEO III we use the dE/dx information along with information from the RICH detector. Unfortunately the calculation of pure detector probabilities for particular particle species is not as straightforward in CLEO III. For example in CLEO II we could determine the absolute probability that a particle was pion or a kaon, but in CLEO III we must ask whether a particle is more pion-like than kaon-like.

We can, however, mock up a number that behaves like $\mathcal{P}_D^s$. As an example I will discuss how we generate $\mathcal{P}_D^\pi$ and $\mathcal{P}_D^K$. The dE/dx reconstruction code provides $\chi_\pi^{dE/dx}$. The RICH system provides us the quantity $\chi_\pi^{2^{RICH}} = -2\ ln\ \mathcal{L}_\pi$, which is related to the likelihood that the track is a pion. However, $\chi_\pi^{2^{RICH}}$ only has meaning when compared with another likelihood, $\chi_K^{2^{RICH}}$, it has no meaning in an absolute sense. We can still assign relative pion and kaon probabilities using the construction

$$\tilde{\chi}^2 = \chi_K^{2^{dE/dx}} - \chi_\pi^{2^{dE/dx}} + \chi_K^{2^{RICH}} - \chi_\pi^{2^{RICH}}. \tag{4.10}$$

When $\tilde{\chi}^2$ is zero the track looks just as much like a pion as a kaon. For positive or negative numbers the track is more pion- or kaon-like respectively. We now map $\tilde{\chi}^2$ onto a function that looks like a probability by defining

$$\mathcal{P}_D^\pi = \begin{cases} 1 - \frac{1}{2}e^{-\tilde{\chi}^2} & \text{if } \tilde{\chi}^2 \geq 0 \\ \frac{1}{2}e^{\tilde{\chi}^2} & \text{if } \tilde{\chi}^2 < 0, \end{cases} \tag{4.11}$$

$$\mathcal{P}_D^K = 1 - \mathcal{P}_D^\pi. \tag{4.12}$$

In a similar fashion we define $\mathcal{P}_D^p$ with respect to $\mathcal{P}_D^K$. Using this technique for CLEO III we are able to generate the probabilities, $\mathcal{P}_D^s$, that can be used in the same way as the calculated CLEO II probabilities.

From our Monte Carlo model of generic B decays we obtain a parameterization of the production fractions $\mathcal{P}_P^s$ as a function of particle momentum p given by

$$
\mathcal{P}_P^K(p) = \begin{cases} 0.277p - 0.107p^2 & \text{if } p < 2.0 \; GeV/c \\ \\ 0.06 & \text{if } p \geq 2.0 \; GeV/c, \end{cases} \tag{4.13}
$$

$$
\mathcal{P}_P^p(p) = 0.02, \tag{4.14}
$$

$$
\mathcal{P}_P^\pi(p) = 1 - \mathcal{P}_P^K(p) - \mathcal{P}_P^p(p). \tag{4.15}
$$

We calculate a relative "probability" that the track is of species s, $\mathcal{P}^s$ by multiplying $\mathcal{P}_D^s \cdot \mathcal{P}_P^s(p)$. We then assign a particle an identification, s, based on which hypothesis, π, K, or p, gives the maximum $\mathcal{P}^s$. Note that pions are the dominant particle produced in B decay. Assuming detector information is inconclusive for a particular track we would still like to enforce what know about B decays to call that track a pion.

4.2.3 Monte Carlo Considerations

We need our analysis code to produce the same results, that is, efficiencies and fake rates, when run over Monte Carlo as when run over data. Perfect agreement is challenging in the area of lepton identification. Electron identification depends strongly on the shapes of calorimeter showers which can be difficult to model in Monte Carlo. The simulation also overestimates the rate for pions to fake muons. To remedy these problems we employ a special technique when identifying letpons in Monte Carlo.

In order to identify Monte Carlo electrons we consider only tracks that satisfy the quality requirements above and that are tagged to generated electrons. We measure the identification efficiency as a function of momentum in data using radiative Bhabha events embedded into hadronic events [30]. We throw a random number between zero and one and if the thrown number is less than the efficiency we identify the track as an electron. This ensures that we have the same identification efficiency on average in Monte Carlo that we do in data.

For muons we also only consider tracks which are tagged to generated muons. In CLEO II we have verified that the Monte Carlo correctly predicts the identification efficiency *for true MC muons.* With regard to CLEO III we apply an event weight that corrects the Monte Carlo identification efficiency for problems in simulating early features of the muon hardware readout . We will postpone a discussion of "event weights" in general until Section 5.2. Complete details of the CLEO III muon efficiency study can be found in Appendix A.

The procedure just outlined in the previous paragraph allows us reproduce the correct identification efficiency for true leptons. However, this procedure does not yield a background estimate for hadrons faking leptons. As will be discussed later, we require exactly one lepton in the event, and some fraction of the "extra leptons" arise from hadrons faking leptons. We therefore need measurements of the rates for hadrons to fake leptons both to estimate the fake signal lepton background contributions (see Section 5.3.3) and to simulate the efficiency loss when these fakes result in an extra lepton in the event.

We measure the probability for hadrons to fake leptons as a function of momentum in data, using clean samples of pions, kaons, and protons[3]. The probability

---

[3]Appendix A discusses this measurement for muons in CLEO III

that there will be no fake leptons in the event can be written as

$$\mathcal{P}_{!\text{fake}} = \prod_{hadrons} \left(1 - \mathcal{P}^s_{\mu\ \text{fake}}(p)\right) \left(1 - \mathcal{P}^s_{e\ \text{fake}}(p)\right), \qquad (4.16)$$

where $\mathcal{P}^s_{\mu\ \text{fake}}(p)$ or $\mathcal{P}^s_{e\ \text{fake}}(p)$ is the probability for a track of species s and momentum p to fake a muon or electron. A random number between zero and one will then be greater than $\mathcal{P}_{!\text{fake}}$ with a probability equal to that of having one or more fake leptons in the event. Events for which this is the case can therefore be vetoed in Monte Carlo at the same rate that they are in data.

It is only lepton identification efficiencies and fake rates in Monte Carlo that require sophisticated tuning to be certain that analyzing Monte Carlo and real data produce the same results. Studies have verified that the Monte Carlo does a sufficiently good job and modeling dE/dx, RICH, and time of flight information used in hadron identification within reasonable systematic uncertainties which we will explore in Chapter 6.

4.3   Signal Hadron Reconstruction

In our quest to reconstruct $B \rightarrow X_u \ell \nu$ decays we devote this section to identifying and reconstructing the X_u part of the decay. Given the work described above, the task of assembling unique pions and photons into signal hadrons is quite simple.

We would like to reconstruct X_u candidates of the following types: $\pi^{\pm}$, π^0, η, $\rho^{\pm}$, ρ^0, and ω. Table 4.1 summarizes the decay modes that we reconstruct for each of these hadrons. Note that all of the final states are a combination of charged pions and photons. We require that the pions be identified as pions by the particle

Table 4.1: A summary of the reconstructed hadron final states.

X_u	Final State
$\pi^\pm$	$\pi^\pm$
π^0	$\gamma\gamma$
η	$\gamma\gamma;\ \pi^+\pi^-\pi^0, \pi^0 \to \gamma\gamma$
$\rho^\pm$	$\pi^\pm\pi^0, \pi^0 \to \gamma\gamma$
ρ^0	$\pi^+\pi^-$
ω	$\pi^+\pi^-\pi^0, \pi^0 \to \gamma\gamma$

identification algorithm discussed above and that they not be the daughter of a reconstructed K_S^0.

In all instances where we reconstruct the the decay of η or π^0 to $\gamma\gamma$ we perform a kinematic fit that constrains the invariant mass of the photons to the nominal mass of the η or π^0. We require that this fitting procedure not "pull" the invariant mass more than two standard deviations from its raw value. This produces a resolution dependent cut on the invariant mass of the reconstructed photons. On average, this cut eliminates candidates more than 8 MeV/c^2 (26 MeV/c^2) from the nominal π^0 (η) mass.

In all other cases, kinematic fits are not performed. We accept η (ω) candidates decaying into $\pi^+\pi^-\pi^0$ that are within 10 (30) MeV/c^2 of the nominal η (ω) mass. In order to have sufficient efficiency when reconstructing the inherently wide ρ, we accept all $\pi\pi$ candidates within 285 MeV/c^2 of nominal ρ mass.

We will revisit the topic of signal hadron reconstruction briefly when we discuss $B \to X_u \ell\nu$ candidate selection later in Section 4.9. There we will mention additional requirements that can placed on the hadron candidates in order to increase the probability that we pick a real $B \to X_u \ell\nu$ decay over some other combination of random particles in the event.

4.4   Final Candidate Reconstruction

Let's begin this section by summarizing what we have discussed so far in this chapter. We have taken tracking and calorimetry information and attempted to produce a complete set of unique momentum measurements for each particle leaving the interaction region. From the tracks we first select the electrons and muons, and then we used remaining detector information and our knowledge of production fractions to classify the remaining hadrons as pions, kaons, or protons. Once the particle identification step is complete we now have a complete picture, *i.e.* the four-momenta, of all of the *visible* particles produced in the collision.

For our current purposes, let us assume for the moment that p_{miss}, as defined in Equation 4.1, is the measured four-momentum of the signal neutrino, p_ν. As can be seen in Figure 4.6, the resolution of the reconstructed neutrino is quite poor with regard to the expected resolution of the any of the visible particles in the event. We have an opportunity, given what we know about the rest of the $B \to X_u \ell \nu$ candidate, to refine the kinematic properties of this reconstructed neutrino.

Note that the missing energy component of the neutrino four-momentum is more poorly measured than the missing three-momentum. The two chief reasons for this are:

- Because momentum is a vector quantity, mistakes distributed uniformly in all directions tend to cancel out. With the scalar energy, similar mistakes such as a little extra noise spread out in the calorimeter have an additive effect.

- The momentum of track is derived directly from the curvature in the drift chamber. Determining its energy however requires knowing what its iden-
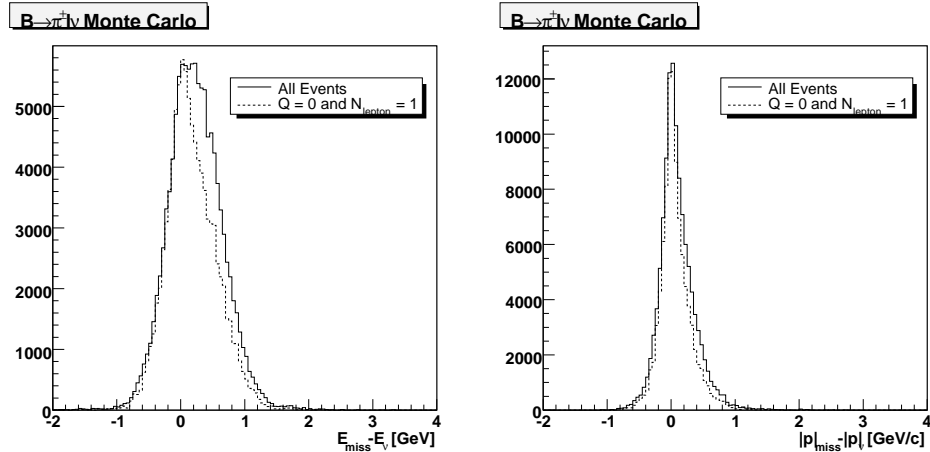
Figure 4.6: A plot of the difference between reconstructed and generated missing energy (left) and missing momentum (right). The resolutions can be improved by requiring the net charge of all tracks to be zero and the number of leptons in the event be just one, the signal lepton. Both of these criteria remove events with extra particles. (The dotted histogram has been scaled to allow an easy comparison.)

tity is. Therefore, the missing energy calculation is susceptible to particle identification mistakes.

Knowing that neutrinos are effectively massless we can use the magnitude of the of the missing momentum to set the energy of the neutrino. We refine the neutrino four-momentum by setting

$$p_\nu \to (|\vec{p}_\nu|, \vec{p}_\nu). \tag{4.17}$$

Using this modified neutrino four-vector we can now fold in the next piece of information: the energy of our reconstructed B decay should be equal to the energy of one of the beam particles. In the process $\Upsilon(4S) \to B\bar{B}$ each B carries away the same energy as one of the beam particles. We can therefore define the variable

ΔE as

$$\Delta E = E_{final} - E_{beam} \tag{4.18}$$

$$= (E_{X_u} + E_\ell + |\vec{p}_\nu|) - E_{beam}. \tag{4.19}$$

This variable is of great importance as the true signal candidates will peak at $\Delta E = 0$. Assuming that for a given signal decay the deviation from $\Delta E = 0$ arises predominantly from measurement errors in the magnitude of the missing momentum, we can scale the neutrino momentum $\vec{p}_\nu$ by a parameter α to force $\Delta E \to 0$. We can compute α from the expression above as

$$\alpha \equiv \frac{1}{|\vec{p}_\nu|} \left(E_{beam} - E_{X_u} - E_\ell \right). \tag{4.20}$$

We further refine our neutrino reconstruction by incorporating this energy adjustment and setting

$$p_\nu \to \left(\alpha \left| \vec{p}_\nu \right|, \alpha \vec{p}_\nu \right). \tag{4.21}$$

One final constraint remains: the mass of the reconstructed B meson should equal the nominal B mass. Since we know the energy of the B is equal to E_{beam}, a quantity known very well, we can define the reconstructed B mass, $M_{h\ell\nu}$ as:

$$M_{h\ell\nu} = \sqrt{E_{beam}^2 - |\vec{p_B}|^2} \tag{4.22}$$

$$= \sqrt{E_{beam}^2 - |\vec{p}_{X_u} + \vec{p}_\ell + \alpha \vec{p}_\nu|^2}. \tag{4.23}$$

True signal candidates will peak at $M_{h\ell\nu} = M_B$.

Finally, we assume that any deviation from the nominal value for a true signal candidate is due to an error in the reconstruction of the *direction* of the neutrino momentum. If we divide $\vec{p}_B$ into two components $\vec{p}_\nu$ and $\vec{p}_{h\ell}$, where $\vec{p}_{h\ell} \equiv \vec{p}_{X_u} + \vec{p}_\ell$, we can force $M_{h\ell\nu} \to M_B$ by rotating the neutrino momentum, $\vec{p}_\nu$, in the plane

defined by $\vec{p}_{h\ell} \times \vec{p}_\nu$. Specifically, we rotate the neutrino to enforce the relationship

$$\cos\theta_{\nu-h\ell} = \frac{E_{beam}^2 - M_B^2 - \alpha^2\vec{p}_\nu^2 - \vec{p}_{h\ell}^2}{2\,|\alpha\vec{p}_\nu|\,|\vec{p}_{h\ell}|}, \tag{4.24}$$

where $\theta_{\nu-h\ell}$ is the angle between $\vec{p}_{h\ell}$ and $\vec{p}_\nu$. We discard candidates for which the above expression for $\cos\theta_{\nu-h\ell}$ results in an unphysical angle. If we call $\mathbf{R}$ the rotation matrix that forces $M_{h\ell\nu} \to M_B$, we can reset the neutrino momentum, now incorporating all that we know about the candidate, to

$$p_\nu \to \left(\alpha\,|\vec{p}_\nu|\,, \alpha\mathbf{R}\vec{p}_\nu\right). \tag{4.25}$$

We use this highly refined value of the neutrino four-momentum to maximize our resolution in the momentum-transfer variable q^2:

$$q^2 \equiv |p_\nu + p_\ell|^2. \tag{4.26}$$

At this point I have outlined the algorithm for identifying and reconstruction $B \to X_u\ell\nu$ candidates within an event. In the coming sections I will discuss how to separate the real $B \to X_u\ell\nu$ from the enormous experimental backgrounds that are present. In the end we seek to represent the data in such a way that the signal appears as a peak on a smooth background for the purpose of extracting the signal with a fit. The variables ΔE and $M_{h\ell\nu}$ defined in this section permit the data to be represented in such a way.

Candidate Selection

Now that we have the ability to reconstruct $B \to X_u\ell\nu$ candidates we are faced with the daunting task of trying to separate the real signal candidates from the background. While I postpone a full discussion of fitting and extracting a signal

until the next chapter it is impossible to present the ideas of candidate selection without some general idea of how the data will be fit. When we attempt to fit the data we will have a collection events from the detector. Some fraction of these events will be signal and some fraction will be background. The purpose of the fit is to determine these relative fractions. Most importantly, we are interested in knowing the number of signal events. The ability of the fit to discriminate between signal and background depends on two things:

- the difference in shape between the signal and background

- and the statistical significance of the signal compared to the background.

In order to aid the fitter with the first of these we choose to bin the data two-dimensionally in the variables ΔE and $M_{h\ell\nu}$ as described in the previous section. In these variables the signal peaks and the background is relatively smooth allowing maximum shape discrimination. The second of these points will be the subject of the rest of this chapter. We must find a collection of analysis requirements which maximize the number of signal events with respect to background events.

It is important to note all of the selection studies were performed on Monte Carlo that is independent of the Monte Carlo used in the final fit. Furthermore, *no selection studies were performed using the actual data collected with the detector.* By using a Monte Carlo sample for tuning the analysis requirments that is independent of the sample used for fits to the data, we avoid biasing our final fit by accidentally tuning on a statistical fluctuation of the Monte Carlo.

4.5   Figures of Merit

To evaluate the impact of any particular selection criteria in the analysis it is necessary to define a "figure of merit" ($\mathcal{F}$), which in our case is directly related to the total significance of the extracted rates. The figure of merit becomes the yardstick for fine tuning the analysis. Because of its importance helping us achieve the maximum statistical significance of the result and its central role in later discussions, I undertake a rigorous development in this section. *Note:* in the discussion that follows I will use σ for statistical errors, s for systematic errors, and δ for combined systematic and statistical errors.

4.5.1   A Basic Figure of Merit

As a starting point, let's consider the figure of merit

$$\mathcal{F} = \frac{N_s}{\sqrt{N_s + N_b}}, \tag{4.27}$$

where N_s and N_b are the number of signal and background events in the signal bin. $\mathcal{F}$ is simply the significance of the signal because

$$\sigma_{N_s} = \sqrt{N_s}, \tag{4.28}$$

$$\sigma_{N_b} = \sqrt{N_b}, \tag{4.29}$$

and N_s by the quadrature sum yields the significance.

This simple figure of merit is inadequate for our use because it assumes that one can subdivide the signal and background contributions to the total yield with infinite precision. In both cases these distributions have finite statistics which should be considered when evaluating a figure of merit. We can expand the expression for

N_b as

$$N_b = N_{b \to c} + N_{b \to u} + N_{cont}. \tag{4.30}$$

The first and most significant component of the background comes from generic $b \to c$ decays and is modeled by a five-times luminosity set of Monte Carlo. Therefore the counting error on this distribution will be equal $\sqrt{5N_{b \to c}}$, but this counting error will enter with a factor of $\frac{1}{5}$. In addition we include an estimated 10% *systematic* error on $N_{b \to c}$ due to imperfect modeling of the Monte Carlo in the total uncertainty estimate of $N_{b \to c}$:

$$\sigma_{N_{b \to c}} = \sqrt{N_{b \to c} + \left(\frac{1}{5}\right)^2 5N_{b \to c}}, \tag{4.31}$$

$$s_{N_{b \to c}} = 0.1N_{b \to c}. \tag{4.32}$$

The background from the non-$B\bar{B}$ continuum, N_{cont}, will be derived from data taken below the $\Upsilon(4S)$ resonance. This data set is only one half of the size of the on-resonance data, so the distribution must be scaled up instead of down like the $b \to c$ distribution. Because the distribution does come from data there is no systematic component in the error. We can summarize the error on N_{cont} as

$$\sigma_{N_{cont}} = \sqrt{N_{cont} + 2^2 \frac{1}{2} N_{cont}}. \tag{4.33}$$

The final background component comes from other $b \to u$ decays that feed into the particular signal we are trying to extract. These decays are modeled with roughly twenty-times luminosity so we can neglect any counting error on the model distribution for the purposes of tuning our selection criteria. Simplifying we can rewrite the error on N_b:

$$\sigma_{N_b} = \sqrt{1.2N_{b \to c} + 3N_{cont} + N_{b \to u}}, \tag{4.34}$$

$$s_{N_b} = 0.1N_{b \to c}. \tag{4.35}$$

Using this more careful estimate of the error on the number of background events we can revise $\mathcal{F}$ to include the finite Monte Carlo statistics used in the measurement. For a single measured yield we therefore have

$$\mathcal{F} = \frac{N_s}{\sqrt{\sigma_{N_s}^2 + \sigma_{N_b}^2 + s_{N_{b \to c}}^2}}.$$

(4.36)

4.5.2 A Figure of Merit for a Rate Measurement

We are interested in producing a single rate measurement for the $\pi^- \ell^+ \nu$ and $\rho^- \ell^+ \nu$ modes in each q^2 bin and also for the η mode. In order to do this we must measure the efficiency corrected yield of these decays. We make this measurement from a combination of several independent rates. For example, in the $\pi^- \ell^+ \nu$ mode, we will have the $\pi^{\pm}\ell\nu$ and $\pi^0 \ell\nu$ samples, each of which will be divided into net charge ± 1 and net charge 0 samples. We average these independent measurements and enforce isospin symmetry to produce a final $\pi^- \ell^+ \nu$ measurement.

When averaging $\pi^{\pm}\ell\nu$ and $\pi^0 \ell\nu$ to determine a combined $B \to \pi l \nu$ event count cross feed exists and the two measurements are no longer independent. We will ignore correlations of this type in the final figure of merit calculation. However, there are correlated systematic errors, such as the 10% $b \to c$ modeling error, that cannot be ignored when averaging samples.

Assuming a set of cuts with signal efficiency ϵ_i produces a sample with N_{s_i} and N_{b_i} signal and background events one can calculate the efficiency corrected number of signal events, denoted $\mathcal{N}_i$, as

$$\mathcal{N}_i = \frac{1}{\epsilon_i} N_{s_i},$$

(4.37)

with statistical and correlated systematic errors:

$$\sigma_{\mathcal{N}_i} = \sqrt{\left(\frac{1}{\epsilon_i}\sigma_{N_{s_i}}\right)^2 + \left(\frac{\sigma_{\epsilon_i}}{\epsilon_i}\frac{1}{\epsilon_i}N_{s_i}\right)^2}, \tag{4.38}$$

$$s_{\mathcal{N}_i} = \frac{1}{\epsilon_i}s_{N_b}. \tag{4.39}$$

The error σ_{ϵ_i} is the uncertainty on the efficiency of the ith sample. For our tuning purposes we assume that this error is due strictly to the finite statistics of the N_{gen} signal Monte Carlo events used to evaluate the efficiency and does not contain any systematic component. The error, given by the binomial distribution, is therefore

$$\sigma_{\epsilon_i} = \sqrt{\frac{\epsilon_i\,(1-\epsilon_i)}{N_{gen}}}. \tag{4.40}$$

We can then make a weighted average of event measurements made with the n (assumed independent) samples. The average number of signal events $\mathcal{N}$ is given by

$$\mathcal{N} = \frac{\sum_{i=1}^{n} w_i \mathcal{N}_i}{\sum_{i=1}^{n} w_i}, \tag{4.41}$$

where we weight the samples by their overall significance using the weights

$$w_i = \frac{1}{\sigma_{\mathcal{N}_i}^2 + s_{\mathcal{N}_i}^2}. \tag{4.42}$$

Careful consideration is necessary when computing the error on the average given by Equation 4.41 if there are significant correlations in the sample. In the end we are interested in the significance of the average measurement, so an incorrect error calculation will easily lead us astray. Given that each independent measurement, $\mathcal{N}_i$, has an uncorrelated statistical error $\sigma_{\mathcal{N}_i}$ and a fully correlated systematic error $s_{\mathcal{N}_i}$, we can write the $n \times n$ covariance matrix as

$$C_{ij} = \begin{cases} \sigma_{\mathcal{N}_i}^2 + s_{\mathcal{N}_i}^2 & \text{if } i = j, \\ s_{\mathcal{N}_i} s_{\mathcal{N}_j} & \text{if } i \neq j. \end{cases} \tag{4.43}$$

Using this covariance matrix we can compute the variance of $\mathcal{N}$, $V_\mathcal{N}$,

$$V_\mathcal{N} \;=\; \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial \mathcal{N}}{\partial \mathcal{N}_i} C_{ij} \frac{\partial \mathcal{N}}{\partial \mathcal{N}_j} \tag{4.44}$$

$$=\; \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{w_i}{W} C_{ij} \frac{w_j}{W} \tag{4.45}$$

$$=\; \frac{1}{W^2}\left[\sum_{i=1}^{n} w_i^2 \left(\sigma_{\mathcal{N}_i}^2 + s_{\mathcal{N}_i}^2 \right) + 2\sum_{i=1}^{n}\sum_{j=i+1}^{n} w_i w_j s_{\mathcal{N}_i} s_{\mathcal{N}_j} \right] \tag{4.46}$$

$$=\; \frac{1}{W} + \frac{2}{W^2}\sum_{i=1}^{n}\sum_{j=i+1}^{n} w_i w_j s_{\mathcal{N}_i} s_{\mathcal{N}_j}, \tag{4.47}$$

where $W \equiv \sum_{i=1}^{n} w_i$. The first term in the last expression above is the contribution of the n independent uncorrelated samples, while the second term inflates the error appropriately to account for the correlations. Given the weighted average signal measurement and its variance we can trivially define a figure of merit related to the significance of the measurement as

$$\mathcal{F} = \frac{\mathcal{N}}{\delta \mathcal{N}} = \frac{\mathcal{N}}{\sqrt{V_\mathcal{N}}}. \tag{4.48}$$

This is the figure of merit that we use to tune the selection criteria for the analysis.

4.6   Continuum Suppression

The largest and fortunately easiest to suppress background comes from the production of $q\bar{q}$ pairs, where $q = u, d, s$, or c. Approximately 75% of the cross section for $e^+ e^- \rightarrow hadrons$ at the $\Upsilon(4S)$ is of this type. We are able to separate these events from those events in which a pair of B mesons was produced by examining the shape of the event. We do this using the technique of a Fisher Discriminant [31].

4.6.1 Event Shape Variables

If we have an event of the type $e^+e^- \to q\bar{q}$ where $q = u, d, s$, or c, the sum of the final state quark masses is much less than the total center of mass energy of the collision. To conserve four-momentum the quarks leave the interaction region back-to-back with large velocities. Because of this, they subsequently hadronize into two collinear jets of hadrons.

In our signal $B\bar{B}$ events, on the other hand, the particles tend to be distributed isotropically. In the case of $e^+e^- \to \Upsilon(4S) \to B\bar{B}$ the two B mesons are produced essentially at rest in the laboratory frame. They each subsequently explode into hadrons spraying particles uniformly in all directions. Furthermore the kinematics of a particular B decay is completely uncorrelated from the kinematics of the other B decay.

It is this shape difference, "jetty" versus isotropic, that allows us to discriminate between $B\bar{B}$ events and continuum $q\bar{q}$ events. In order utilize this shape difference we must define a suitable set of variables to quantitatively describe the shape of the events. One such variable is the "thrust axis," $\hat{t}_i$, for a set of i particles. We define the thrust axis as the unit vector that maximizes the thrust, t, where

$$t \equiv \sum_i \vec{p}_i \cdot \hat{t}_i. \tag{4.49}$$

It will be useful to consider the thrust axes of our candidate particles, the particles in the rest of the event, and the entire event. We can define these as $\hat{t}_{cand}$, $\hat{t}_{ROE}$, and $\hat{t}_{event}$ respectively.

To see the usefulness of this quantity, consider a continuum event where all of the particles are collimated into two jets. As a result, $\left| \hat{t}_{cand} \cdot \hat{t}_{ROE} \right|$ will peak near
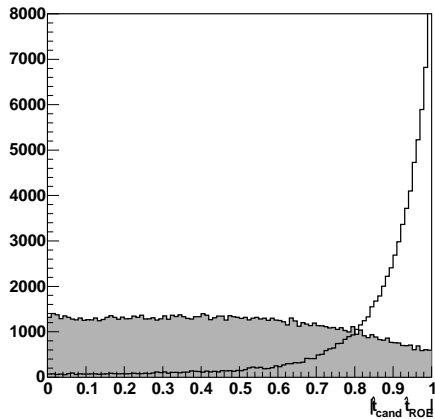
Figure 4.7: A comparison of $\left|\hat{t}_{cand} \cdot \hat{t}_{ROE}\right|$ for $B \rightarrow \pi\ell\nu$ signal Monte Carlo (shaded) and off-resonance continuum data (open).
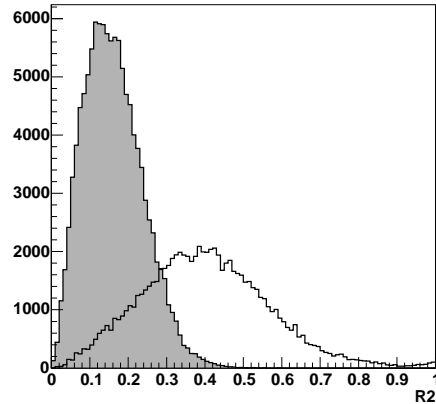
Figure 4.8: A comparison of $R2$ for $B \rightarrow \pi\ell\nu$ signal Monte Carlo (shaded) and off-resonance continuum data (open).

one (shown in Figure 4.7). That is to say, the two thrust axes will be aligned. We do not include the neutrino momentum in the calculation of $\hat{t}_{cand}$ to avoid the inherit correlation between it and the other particles in the event from which it is derived.

Additionally we use the ratio, $R2 \equiv H_2/H_0$, of the second to zeroth Fox-Wolfram moments [33]. The lth Fox-Wolfram moment for a set of i particles is defined as

$$H_l \equiv \left(\frac{4\pi}{2n+1}\right) \sum_{m=-l}^{l} \left| \sum_i Y_l^m(\Omega_i)\frac{|\vec{p}_i|}{\sqrt{s}} \right|^2 , \qquad (4.50)$$

where s is the total center of mass energy squared, and the Y_l^m's are the familiar spherical harmonics. The jet structure present in continuum events enhances the second moment and therefore the ratio $R2$ tends to one for jetty continuum events and to zero for isotropic $B\bar{B}$ events as shown in Figure 4.8.

The two variables defined above, $\left|\hat{t}_{cand} \cdot \hat{t}_{ROE}\right|$ and $R2$, provide very good dis-

crimination between $B\bar{B}$ and continuum events. Following previous work done on continuum supression [32] and in an attempt to gain further discriminating power we chose to define an additional set of ten variables. The first of these variables is the polar angle θ between thrust axis of the entire event, $\hat{t}_{event}$ and the beam axis. For $B\bar{B}$ events this variable will be randomly distributed, while for continuum events the $\hat{t}_{event}$ will align with the jet axis and be distributed with the angular dependence of the cross section for $e^+e^- \rightarrow q\bar{q}$ [34]

$$\frac{d\sigma}{d\Omega} \propto 1 + \cos^2\theta. \tag{4.51}$$

The remaining nine variables track the momentum flow of the event into nine double-cones about the thrust axis. Each double-cone spans $10°$ in polar angle from the thrust axis; therefore all nine, cover the entire solid angle. We write the fractional momentum flow of the j_i particles into the ith cone as

$$x_i = \frac{\sum_{j_i} |\vec{p}_{j_i}|}{\sum |\vec{p}|}, \tag{4.52}$$

where the sum in the denominator runs over all particles in the event. Certainly these variables are highly correlated with the previously defined event-shape variables. We use these twelve variables: $\left|\hat{t}_{cand} \cdot \hat{t}_{ROE}\right|$, $R2$, $\left|\hat{t}_{event} \cdot \hat{z}\right|$, and the nine x_i to as inputs to a Fisher discriminant that allows us to distinguish continuum events from $B\bar{B}$ events.

4.6.2 Constructing a Fisher Discriminant

Given an n dimensional parameter space the the Fisher discriminant effectively allows one to divide this space with an $n-1$ dimensional hyper-plane into two parts. We will call these the "signal" and "background" regions. For example, if we used

only two variables for continuum suppression, constructing a Fisher discriminant would essentially pick the line that best divided the two-dimensional space such that signal events most frequently fell on one side of the line while background events most frequently fell on the other side.

Operationally, we find a set of "Fisher weights" that determine the slope of the hyper-plane in each of the dimensions. These weights are picked by examining a set of pure simulated signal events and a set of simulated background events. Defining the twelve-dimensional covariance matrix of our continuum suppression variables as $\mathbf{C}$ and each variable has a mean μ_i, then the optimal Fisher weights, α_i are given by

$$\alpha_i \equiv \sum_{j=1}^{12} \left(\mathbf{C}^{sig} + \mathbf{C}^{bkg}\right)_{ij}^{-1} \times \left(\mu_i^{bkg} - \mu_i^{sig}\right). \tag{4.53}$$

We can then define the Fisher discriminant, $\mathcal{D}_\mathcal{F}$, for our set of twelve event-shape variables, λ_i, as

$$\mathcal{D}_\mathcal{F} \equiv \sum_{i=1}^{12} \alpha_i \lambda_i. \tag{4.54}$$

Examining the two equations above gives insight as to how the discriminant functions: we weight more heavily those variables with means that differ significantly in the two samples. The order of subtraction in the weight evaluation (Equation 4.53) pushes the discriminant more positive for background-like events and more negative for signal-like events. By cutting all events above some maximum $\mathcal{D}_\mathcal{F}$ we can remove a substantial fraction of the continuum-like events. From the multidimensional space perspective, the weights determine the orientation of the hyper-plane and the cut we choose translates the dividing plane in the remaining orthogonal dimension. We use the figure of merit to select the most optimal cut.
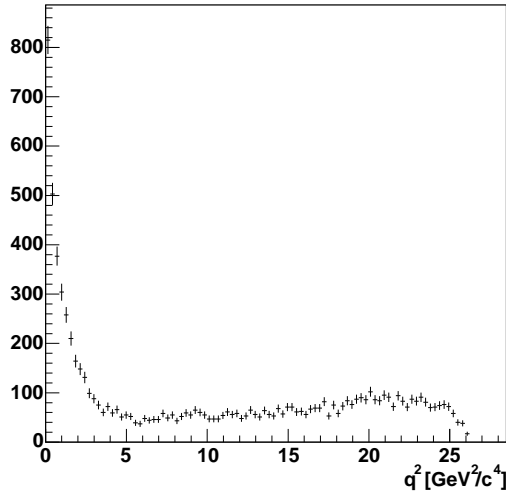
Figure 4.9: The reconstructed q^2 of $B \to \pi\ell\nu$ candidates in off-resonance data is shown. No continuum suppression cuts have been applied.

4.6.3   Cut Implementation and Optimization

As with most things, the implementation of a continuum suppression algorithm using a Fisher discriminant is not as simple as it first seems. Recall that the fundamental idea behind suppressing the the continuum background is the fact that the events are shaped differently. Within our signal modes there also exist a variety of different event shapes, $B \to \pi^-\ell\nu$ events look differently than $B \to \omega\ell\nu$ events because they have different final state multiplicities. More importantly, however, is the variation in shape across the q^2 spectrum for any particular mode. At low q^2 the signal leptons tend to come back-to-back against the fast recoiling hadrons mimicking the jet structure of a continuum event. For this reason, as shown in Figure 4.9, most of the continuum background leaks in at very low q^2.

We must take extra care to ensure that our continuum suppression algorithm does not introduce a q^2 bias into the reconstruction efficiency of our sample. Such
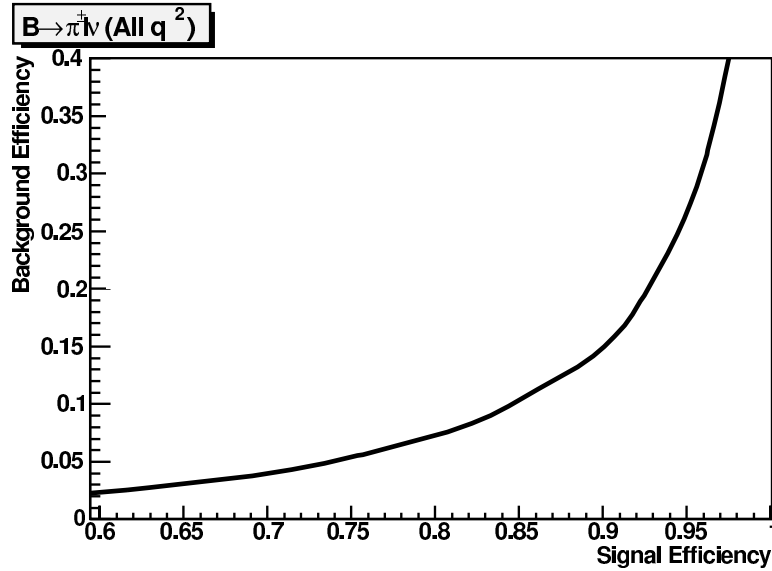
Figure 4.10: Plotting the background efficiency vs. signal efficiency for $B \to \pi^{\pm}\ell\nu$ candidates over all q^2 illustrates the typical performance of the continuum suppression algorithm.

a bias would prevent us from making a model independent measurement of the rate as we would have to rely on a theoretical prediction of the shape of the rate as a function of q^2 to determine our reconstruction efficiency.

We determine a set of weights, α_i, independently for each reconstructed mode and each reconstructed q^2 and $\cos\theta_{Wl}$ bin. We use simulated signal to make up the "signal" sample. We derive the "background" sample from an appropriately constructed mix of $q\bar{q}$ and τ-pair Monte Carlo. Once we have computed the weights, we pick the maximum allowed values for the $\mathcal{D}_{\mathcal{F}}$ that maximize our figure of merit. This technique performs quite well and on average allowing high background rejection at little cost in efficiency as can be seen in Figure 4.10.

4.7 $b \rightarrow c$ Background

After eliminating the large continuum background, we are left with the favored $b \rightarrow c$ decay of the B meson as our dominant background. As will be shown, some of this background can be reduced with lepton momentum and track multiplicity requirements. However, this background can not be eliminated as effectively as the continuum background. It remains as the dominant background that we must model and account for in the fit.

4.7.1 Lepton Momentum Requirement

Leptons arise from "charmed" B decay in one of two ways: "primary" leptons from $b \rightarrow c\ell\nu$ transitions and "secondary" leptons from $b \rightarrow c \rightarrow s\ell\nu$ decays. The characteristic spectra of these decays along with our signal $b \rightarrow u\ell\nu$ decays are shown in Figure 4.11. The lighter X_u final state permits the production of leptons with momenta higher than kinematically allowed by $b \rightarrow c\ell\nu$ states. This characteristic excess in the "lepton endpoint" provides a clear experimental indication of $b \rightarrow u\ell\nu$ decays.

One might think that the best way to make an exclusive measurement of $B \rightarrow X_u\ell\nu$ would be to require such a high momentum lepton. This however will also eliminate a significant fraction of the $B \rightarrow X_u\ell\nu$ rate and will depend on theoretical modeling of the decay to predict the efficiency of this cut.

In order to minimize the model dependence we require that the signal lepton have momentum greater than 1.0 GeV/c. This results in a very high efficiency for our signal modes while eliminating a significant fraction of the background
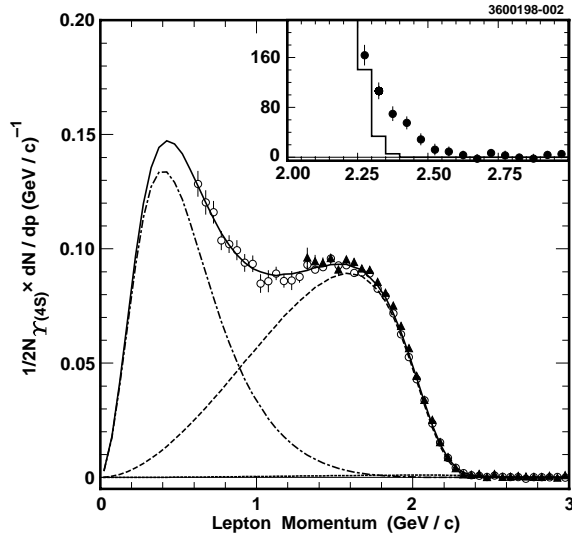
Figure 4.11: Model predictions for inclusive lepton spectra from primary (dash) and secondary (dot-dashed) are compared with data for electrons (hollow circles) and muons (solid triangles) [35]. The primary $b \to u\ell\nu$ component is shown as the finely dashed line that is roughly two orders of magnitude smaller than the primary $b \to c\ell\nu$ component. The inset shows the excess of data over the primary $b \to c\ell\nu$ indicating an observable $b \to u\ell\nu$ rate.

from $b \to c$ decays. Typical lepton momentum spectra, along with the momentum cutoff, are shown in Figure 4.12.

4.7.2 Track Multiplicity Criteria

One characteristic that separates our signal decays from typical B decays is the relatively low number of tracks in the final state. We follow the previous work done on reconstructing exclusive charmless decays by Boisvert [36] and apply the track multiplicity criteria summarized in Table 4.2. Figure 4.13 shows the track multiplicity for both generic B decay events and, as an example, signal $B \to \pi^{\pm}\ell\nu$ events. Note that in data the track multiplicity distribution tends to extend higher than shown in the simulation plots, which is most likely due to noisy data events
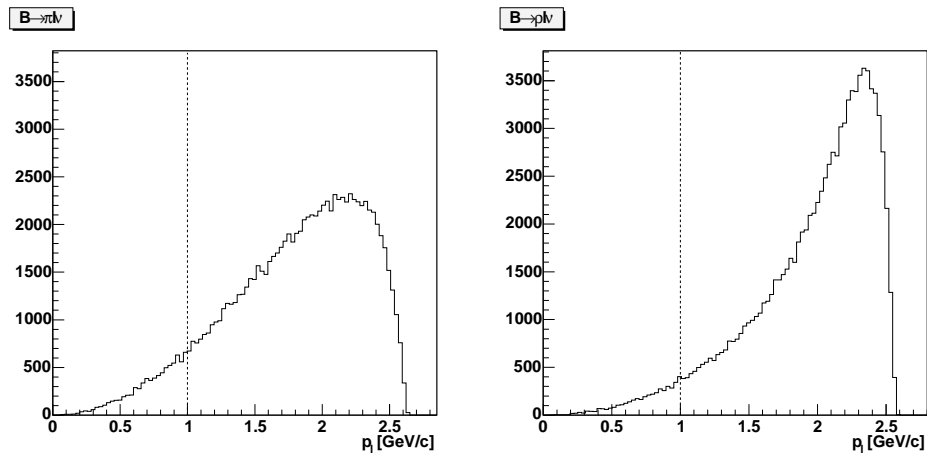
Figure 4.12: The lepton momentum spectrum in the B rest frame for $B \to \pi \ell \nu$ (left) and $B \to \rho \ell \nu$ (right) decays. The 1.0 GeV/c lepton momentum requirement is shown by the dotted line.

in which a large number of tracks are found. Placing a limit on the maximum number of tracks removes events of this type.

4.7.3 $B \to J/\psi K_L^0$

The decay of $B \to J/\psi K_L^0$ presents a unique peaking background that requires elimination. The background arises as follows. The K_L^0 escapes without detection

Table 4.2: A summary of the track multiplicity cuts used for each reconstructed mode.

Mode	Requirements on the Number of Tracks (N_{trk})
$\pi^{\pm} \ell \nu$	$4 \leq N_{trk} \leq 10$
$\pi^0 \ell \nu$	$4 \leq N_{trk} \leq 8$
$\eta \ell \nu, \eta \to \gamma\gamma$	$4 \leq N_{trk} \leq 8$
$\eta \ell \nu, \eta \to \pi^+ \pi^- \pi^0$	$4 \leq N_{trk} \leq 10$
$\rho^0 \ell \nu$	$6 \leq N_{trk} \leq 10$
$\rho^{\pm} \ell \nu$	$4 \leq N_{trk} \leq 10$
$\omega \ell \nu$	$4 \leq N_{trk} \leq 10$

Figure 4.13: The distribution of track multiplicity in signal $B \to \pi^{\pm}\ell\nu$ (left) and generic B decay (right) Monte Carlo.

and thereby fakes a neutrino. The J/ψ decays into a pair of leptons, one of which is identified as a signal lepton while the other is mistakenly identified as a charged pion. Should this happen we can easily create a $B \to \pi^{\pm}\ell\nu$ candidate with $M_{h\ell\nu} \approx M_B$ and $\Delta E \approx 0$. In order to eliminate this background we veto events in which the signal lepton, when paired with any oppositely charged track in the event, has an invariant mass, $M_{h\ell}$, that falls in the window: $3.060 < M_{h\ell} < 3.130 \; GeV/c^2$ or $3.675 < M_{h\ell} < 3.705 \; GeV/c^2$. The latter window removes decays of the type $B \to \psi(2S)K_L^0$.

With the exception of this peaking background, the majority of the generic B background tends to be smooth across the fit plane. Typically the background rises as one moves away from the signal bin towards lower $M_{h\ell\nu}$ and more negative ΔE. We use background shape as modeled in a range of bins to extrapolate this smooth background into the signal region. A more in depth discussion of binning will occur in the next chapter.

4.8 Neutrino Quality Cuts

There is no shortage of real leptons and hadrons in a typical $B\bar{B}$ decay event. What separates our signal from the background is the kinematic constraints realized in the ΔE and $M_{h\ell\nu}$ variables. However, this separation works only as good as we can measure the energy and momentum of our candidates, which is driven entirely by the resolution on the reconstructed neutrino. In order to select events with high quality neutrino reconstruction we apply the following cuts.

4.8.1 The V Cut

By far the most effective quality cut that can be made on the reconstructed neutrino is the V cut. We define

$$V \equiv \frac{M_{miss}^2}{2E_{miss}}, \tag{4.55}$$

and it is so-named because applying a cut on $|V|$ produces a vee shape in the missing mass squared/missing energy plane. To see why this variable is relevant consider the expression for the missing mass squared:

$$M_{miss}^2 = E_{miss}^2 - |\vec{p}_{miss}|^2. \tag{4.56}$$

Ideally we would like to require that $M_{miss}^2 \approx 0$ as should be true for real neutrinos. However, recall from Section 4.4 that the error on the raw neutrino four-momentum is dominated by the missing energy error. If we neglect the error on $\vec{p}_{miss}$ we can write write:

$$\delta M_{miss}^2 \propto 2E_{miss}\delta E_{miss} \tag{4.57}$$

In other words, as the missing energy increases, for example in the case of a high energy neutrino, the error on the M_{miss}^2 increases proportional to the missing

energy. In order to make constant fractional error cut on M_{miss}^2 we must account for this scaling and we do so by cutting on the ratio V.

One might question the effectiveness of this cut. After all, there are plenty of ordinary neutrinos in any typical B decay. The cut merely ensures that we have properly reconstructed a neutrino, not necessarily a signal neutrino. However, remember that requiring the $M_{h\ell\nu} \approx M_B$ and $\Delta E \approx 0$, as is done effectively in binning and fitting, places great kinematic constraints on the candidate. It is very difficult to take a real neutrino, lepton, and pion from a generic $b \rightarrow c$ decay and satisfy these requirements. The typical mechanism to produce the background comes when we add extra momentum to the neutrino by missing additional particles in the event. This increases the energy and momentum of the neutrino such that $M_{h\ell\nu}$ and ΔE comes close to that for signal candidates, but doing so also increases the effective mass of the "neutrino." This is where the V cut becomes effective at remove these background events. Figure 4.14 shows the how a V cut can be used to separate signal from background for candidates that are near the signal region in $M_{h\ell\nu}$ and ΔE.

In practice we tune the upper and lower edges of the V cut asymmetrically to optimize the figure of merit. We do this tuning independently for events with net charge, Q, 0 and $|Q| = 1$. Typically a $|Q| = 1$ event arises when a soft track is lost. If this happens in the other B decay of a signal event, for example in the soft pion of $D^* \rightarrow \pi D$, then we will only slightly perturb the reconstructed neutrino and the event is recoverable. However, we must expand the V window to make optimal use of these $|Q| = 1$ events. For $\pi\ell\nu$ candidates with $q^2 > 16 \; GeV^2$ we tighten the cut slightly avoid excess contamination of the high $b \rightarrow c$ background in this region. The V cut values are summarized in Table 4.3.
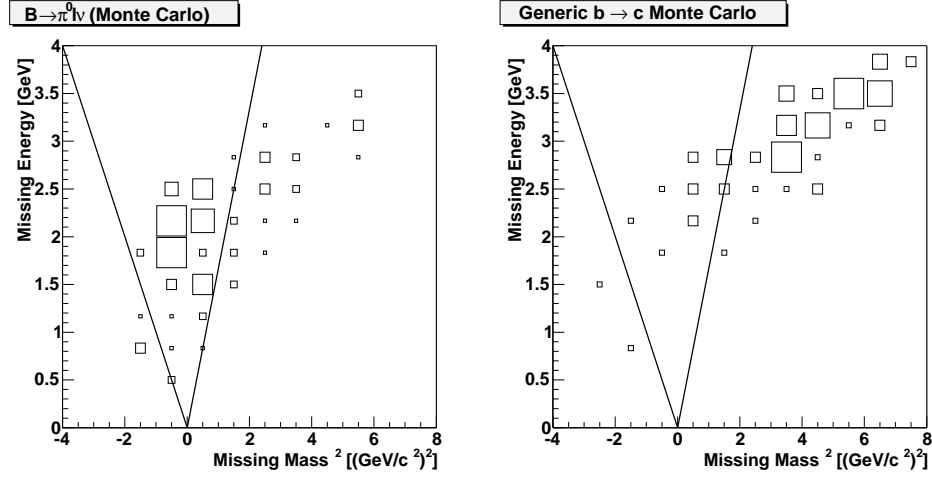
Figure 4.14: A possible V cut is overlaid on the E_{miss} vs. M^2_{miss} plane for $B \to \pi^0 \ell \nu$ (left) and generic $b \to c$ (right) Monte Carlo for candidates near the signal region in $M_{h\ell\nu}$ and ΔE. Real neutrinos end up inside of the vee.

Table 4.3: Values of the V ratio cut for various reconstructed modes. As discussed in Section 4.8.2 we only analyze the $Q = 0$ sample for the vector modes; therefore, no V cut for the $|Q| = 1$ sample is listed for these modes.

Mode	q^2	Q	V Requirement
$\pi^\pm \ell \nu, \pi^0 \ell \nu$	$q^2 < 16$	$Q = 0$	$-0.65 < V < 0.35$
$\pi^\pm \ell \nu, \pi^0 \ell \nu$	$q^2 < 16$	$|Q| = 1$	$-0.65 < V < 0.45$
$\pi^\pm \ell \nu, \pi^0 \ell \nu$	$q^2 > 16$	$Q = 0$	$-0.5 < V < 0.3$
$\pi^\pm \ell \nu, \pi^0 \ell \nu$	$q^2 > 16$	$|Q| = 1$	$-0.3 < V < 0.3$
$\eta \ell \nu$	all q^2	$Q = 0$	$-0.65 < V < 0.35$
$\eta \ell \nu$	all q^2	$|Q| = 1$	$-0.65 < V < 0.45$
$\rho^\pm \ell \nu, \rho^0 \ell \nu, \omega \ell \nu$	all q^2	$Q = 0$	$-0.5 < V < 0.3$

4.8.2   Net Charge

One of the clearest indicators that we have missed or improperly reconstructed charged tracks is when the total charge of all tracks does not add up to zero. By placing a cut on the net charge we improve the quality of the reconstructed neutrino and therefore increase the signal to background ratio in the $M_{h\ell\nu}/\Delta E$ plane. We find that in some cases such as the lost of a soft pion when the other B decays via $B \to D^*X, D^* \to \pi D$ we can still reconstruct the neutrino with some success. We therefore choose to require $|Q| \leq 1$ one for the $\pi\ell\nu$ and $\eta\ell\nu$ modes. Due to the larger backgrounds in the $\rho\ell\nu$ and $\omega\ell\nu$ modes allowing $|Q| = 1$ events provides little if any gain, we therefore require $Q = 0$ for both $\omega\ell\nu$ and $\rho\ell\nu$.

As will be discussed in the next chapter, we separate the sample of $\pi\ell\nu$ and $\eta\ell\nu$ modes based on $|Q|$. This avoids diluting the good $Q = 0$ signal with the $|Q| = 1$ background. Making a similar separation in the heavily background-contaminated $\rho\ell\nu$ modes also turns out systematic liability as requires very good modeling of the relative efficiency of the two $|Q|$ bins. This systematic error would most likely erase any potential statistical gain made by fitting the $|Q| = 1$ bins in the vector modes.

4.8.3   Additional Cuts

The V cut and the net charge requirement are key to increasing the significance of our reconstructed signal. In addition to these two we apply a few other cuts to promote quality neutrino reconstruction.

- We require that there be no additional leptons in the event. Additional

leptons are often produced with additional neutrinos – it is impossible to reconstruct more than one neutrino.

- We require that the polar angle of the reconstructed neutrino θ_ν satisfy $|\cos\theta_\nu| < 0.96$. This vetoes events with large missing momentum down the beam-pipe which can be due to lost tracks or two-photon events.[4]

- We require that all tracks in the event have z information. In a desperate effort to provide track fits the tracking code will sometimes return track information for just the transverse plane. Having no z fit will cause problems with neutrino reconstruction; therefore, we veto events with tracks of this type.

4.9   Selecting the Best Candidate

With the exception of the continuum suppression cut that considers the thrust axis of the candidate, up to this point we have mainly concerned ourselves with selecting a quality neutrino and lepton. These are in a sense "event-level" cuts. We now need to take the sample of reconstructed events and pick out the best $B \to X_u \ell \nu$ candidates. In general our strategy is to require that $M_{h\ell\nu}$ be greater than 5.175 GeV/c^2, and of the candidates that pass this requirement we select the one with ΔE closest to zero as the "best" candidate.

---

[4]These events produce hadrons in the detector from the collision of two photons radiated by the beam particles. Since the beam particles only radiate the photons they tend to escape detection along the beam line leaving large missing momentum.

4.9.1 Combinatoric Considerations

Selecting just one best candidate per event based on ΔE always comes with the risk of choosing the wrong candidate. Suppose we have an event with a true $B^0 \to \rho^- \ell^+ \nu$ candidate, but also many other candidates of different final states in the same event. The probability that one of these candidates happens to have $|\Delta E|$ less than our signal candidate is quite high. We therefore select up to one candidate for *each* of the final states: $\pi^0 \ell \nu$, $\pi^\pm \ell \nu$, $\eta \ell \nu$ with $\eta \to \gamma\gamma$, and $\eta \ell \nu$ with $\eta \to \pi^+ \pi^- \pi^0$.

In the $\rho^\pm \ell \nu$, $\rho^0 \ell \nu$, and $\omega \ell \nu$ modes selecting just one candidate each mode is still not sufficient enough to minimize the effects of the combinatoric background on the signal efficiency. This is especially true in the ρ modes where $\pi\pi$ candidates are accepted over a range of 570 MeV/c^2 in invariant mass. In the ρ (ω) modes we therefore further divide the accepted hadron candidate mass range into three bins of width 195 (20) MeV/c^2. The middle bin is centered on the nominal mass and we select up to one candidate *per bin*. This allows at most a total of six ρ (three charged and three neutral) and three ω candidates to be selected for each event. Unless otherwise noted, all plots, figures, and efficiencies shown in this work for the $\rho \ell \nu$ and $\omega \ell \nu$ modes are for the central mass bin only, although both of the side bins contribute significantly to the fit.

An additional source of combinatoric background arises in modes that have a π^0 in the final state. Typically there are many low energy showers in the calorimeter which give rise to multiple low energy π^0 candidates. Therefore if a final state contains a low energy π^0 that contains a low energy shower there are often several photons that can be substituted for this shower producing a set of valid candidates and therefore increasing the chance that the wrong candidate will be chosen. This
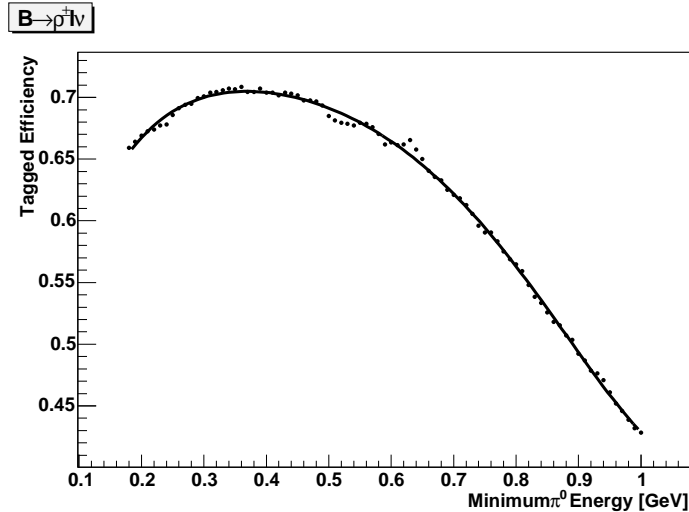
Figure 4.15: The efficiency for reconstructing the properly tagged $B \to \rho^{\pm}\ell\nu$ candidate as a function of the minimum π^0 energy cut. The initial increase in efficiency with increasing cut is due to the reduction of combinatoric background among the π^0 candidates in the final state.

effect is can be cleanly observed in $B \to \rho^{\pm}\ell\nu$ as shown in Figure 4.15. The π^0 in the charged ρ decay allows for multiple candidates at low energy. By increasing the energy cut we actually increase the efficiency of picking the true candidate. We choose to place minimum π^0 energy requirements of 250 and 300 MeV in the $\pi^0\ell\nu$ and $\rho^{\pm}\ell\nu$ final states respectively. Studies show that no significant gain is made by placing a cut on the π^0 energy in the η or ω decays.

Finally we can reduce combinatoric background in the $\omega\ell\nu$ final state by placing a cut on the Dalitz plot amplitude of the three-pion system. We required that the computed Dalitz plot amplitude for the candidate decay be at least twenty percent of the maximum Dalitz amplitude. This cut was chosen by examining the dependence of the figure of merit on the cut, and helps to remove candidates that are not likely to be real ω decays.

4.9.2 Signal Efficiencies

I conclude this section on candidate selection with a summary of the signal reconstruction efficiencies. The tables below show the luminosity weighted average efficiency over the CLEO II, II.V and III data sets. All efficiencies are determined by counting the number of signal events in the signal bin and dividing by the number of generated events. In computing the total efficiency for each final state over all q^2 and $\cos\theta_{Wl}$ we consider the production weighted average in each of the bins based on the form factor calculations used in the nominal fit. A full discussion of binning in the q^2 and $\cos\theta_{Wl}$ variables will follow in the next chapter. For completeness efficiencies in these bins are quoted now. In the $\eta\ell\nu$ and $\pi\ell\nu$ modes efficiencies are computed by summing over net charge bins. To produce a total $\eta\ell\nu$ efficiency the branching fractions for the two different reconstructed final states are folded in. In the $\rho\ell\nu$ and $\omega\ell\nu$ bins only the central mass bin is considered. By including the two side mass bins we typically find that the efficiency is increased by a factor of about 1.7.

Table 4.4: Reconstruction efficiencies in percent for $\pi\ell\nu$ modes. Bin numbering is summarized in Table 5.2.

Mode	q^2 Bin				Total
	1	2	3	4	
$\pi^\pm\ell\nu$	2.6	3.9	5.1	4.8	**4.4**
$\pi^0\ell\nu$	1.4	2.4	3.2	2.6	**2.6**

Throughout this chapter I have developed the technique for reconstructing our signal $B \to X_u \ell\nu$ decays from the basic low level detector information. A set of candidate and event cuts to enhance the ratio of signal to background has been discussed. With this optimized algorithm in place we forge ahead in the next chapter and confront how to represent the data and extract the signal yield.

Table 4.5: Reconstruction efficiencies in percent for $\rho\ell\nu$ and $\omega\ell\nu$ modes. Bin numbering is summarized in Table 5.3. The $\omega\ell\nu$ mode is reconstructed in just two bins, one covers the range of bins 1-4 while the other is bin 5. In all three final states the efficiency is quoted for the central mass bin.

Mode	$q^2/\cos\theta_{Wl}$ **Bin**					Total
	1	2	3	4	5	
$\rho^{\pm}\ell\nu$	0.5	0.9	1.1	1.1	0.8	**1.0**
$\rho^0\ell\nu$	1.1	1.5	2.1	1.9	1.8	**1.8**
$\omega\ell\nu$			0.8		0.5	**0.7**

Table 4.6: Reconstruction efficiencies for $\eta\ell\nu$ modes in percent. The total efficiency folds in the two decay mode branching fractions.

η Decay Mode	Efficiency
$\gamma\gamma$	2.8
$\pi^+\pi^-\pi^0$	1.4
Total	**1.4**

Chapter 5

Fitting the Data

After carefully reconstructing the data and applying all the the selection criteria to isolate events with candidates that look like signal $B \to X_u \ell \nu$ decays, we are left with a mixture of both signal and background events. It is impossible to examine any one event and know with absolute certainty that the event is a signal decay. We can estimate the signal fraction however be looking at distributions where signal and background are, on average, separated. By determining how signal-like or how background-like such a distribution looks we can estimate what fraction of the entire sample is signal.

In this chapter I will develop our fitting technique, which involves accounting for and modeling all backgrounds that may mimic our signal events. I also present a representation for the data that allows us maximum discriminating power between signal and background and therefore provides the most precise determination of the number of signal events. I will then discuss the binned maximum likelihood fit algorithm itself and exactly how the signal fraction is determined, and conclude by presenting branching fraction results for the signal $B \to X_u \ell \nu$ decay.

5.1 Binning

Binning is the art of dividing the data into collections of statistically independent[1] samples which can be simultaneously fit. The fit then attempts to determine

---

[1] Due to the possibility of selecting multiple candidate final states per event, *e.g.* $\pi^0 \ell \nu$ and $\rho^+ \ell \nu$, some of the binning presented here does not produce true statistically independent samples. However, because of the small number of multiple entries per

the relative fractions of signal and background by simply looking at the number of entries in each bin. In this analysis two underlying ideas drive the choice of the set of bins and the decision to do a binned fit:

- *The bin choices effectively separate the signal and background.* If we can choose variables where the signal events tend to end up in certain bins while the background events end up in others, this allows maximum discriminating power between signal and background in the fit.

- *A binned fit minimizes systematic uncertainties.* Neutrino reconstruction resolution drives shape of the signal in the key ΔE and $M_{h\ell\nu}$ variables. As we will see later, the leading experimental systematic errors are those associated with neutrino reconstruction. Performing an unbinned fit that is directly dependent on the modeled resolution would most likely increase the systematic errors on the final result. We can minimize our susceptibility to these resolution uncertainties by choosing a bin size comparable to the resolution in these variables.

5.1.1   The Fit Plane

We choose to bin the data in the variables $M_{h\ell\nu}$ and ΔE as defined in the previous chapter. Given these two variables we try to select a binning scheme in which much of the signal appears in one bin and we can extrapolate the background from the neighboring bins to estimate the amount of background in the signal bin. Figure 5.1 illustrates the choice of binning and how both the signal and the background populate the bins. We choose to number bins from right to left, top

event, the statistical correlations between bins are assumed to be negligible.
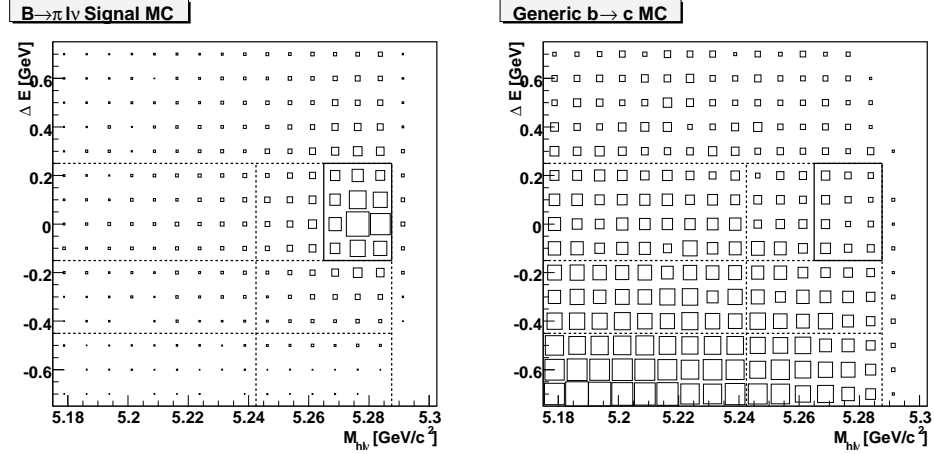
Figure 5.1: An illustration of the seven bins in ΔE and $M_{h\ell\nu}$ that are used to fit the data. Bins are numbered 1-7, right to left and top to bottom, starting with the "signal bin" shown in solid lines. Reconstructed $B \to \pi\ell\nu$ candidates are shown for signal Monte Carlo (left) and generic $b \to c$ Monte Carlo (right).

to bottom, where the first bin shown in the solid box is the "signal bin." The exact binning is summarized in Table 5.1. It is this seven-bin histogram that we fill for each "reconstructed mode." Throughout this work whenever an efficiency is quoted it will *always* be based on the contents of the *signal bin only*.

Table 5.1: The binning used in the $M_{h\ell\nu}$ and ΔE variables. The first bin in bold type is the "signal bin."

Bin #	$M_{h\ell\nu}$ Range $[GeV/c^2]$	ΔE Range $[GeV]$
1	$\mathbf{5.265 < M_{h\ell\nu} < 5.2875}$	$\mathbf{-0.15 < \Delta E < 0.25}$
2	$5.2425 < M_{h\ell\nu} < 5.265$	$-0.15 < \Delta E < 0.25$
3	$5.175 < M_{h\ell\nu} < 5.2425$	$-0.15 < \Delta E < 0.25$
4	$5.2425 < M_{h\ell\nu} < 5.2875$	$-0.45 < \Delta E < -0.15$
5	$5.175 < M_{h\ell\nu} < 5.2425$	$-0.45 < \Delta E < -0.15$
6	$5.2425 < M_{h\ell\nu} < 5.2875$	$-0.75 < \Delta E < -0.45$
7	$5.175 < M_{h\ell\nu} < 5.2425$	$-0.75 < \Delta E < -0.45$

5.1.2 q^2 and $\cos\theta_{W\ell}$ Binning

As discussed in the second chapter, the decay rate has a non-trivial theoretically predicted dependence on the variable q^2 and, additionally in the case of the vector final states, the lepton decay angle, θ_{Wl}. Experimentally this poses a significant problem as the reconstruction efficiency and cross-feed backgrounds are not flat in these two variables. For example, the 1.0 GeV/c lepton momentum cut results in a much higher efficiency for decays at high q^2. In order to remove the dependence on of the result on theoretical calculations of the form-factors we choose to bin the data coarsely in these variables and extract a rate independently for each bin. This accomplishes to things: first, we can make a relatively model-independent measurement of the total rate if our reconstruction efficiency is approximately flat *within any one bin.* Secondly, we can measure the differential rate directly in these variables and thus verify the shapes produced from theoretical calculations.

The $B \to \pi\ell\nu$ rate has its only non-trivial dependence in q^2. The $\theta_{W\ell}$ shape is forced to be $\propto \sin^2\theta_{W\ell}$ because of angular momentum conservation. For the most part, we maintain the same binning used in the previous publication [37]; however, we split the 0-8 GeV^2 bin into two parts: 0-2 GeV^2 and 2-8 GeV^2. As mentioned earlier, this is done to isolate the continuum background in the 0-2 GeV^2 bin. The binning for the $\pi\ell\nu$ modes is summarized in Table 5.2.

Table 5.2: q^2 binning for reconstructed $\pi\ell\nu$ modes

Bin #	q^2 Range $[GeV^2]$
1	$0 < q^2 < 2$
2	$2 < q^2 < 8$
3	$8 < q^2 < 16$
4	$16 < q^2$

In addition to q^2 the $B \to \rho\ell\nu$ rate depends on the lepton decay angle $\theta_{W\ell}$. This dependence arises from the relative rates for the three different W helicity states in the decay. The left-handed nature of the weak interactions prefers a W helicity that results in a typically harder lepton momentum spectrum than produced in $B \to \pi\ell\nu$. Past analyses have used this feature to place a to reduce the amount of $b \to c$ background in the fit by requiring the lepton momentum to be greater than 1.5 GeV/c and only considering candidates that have $\cos\theta_{Wl}$ greater than zero. However, doing so requires one to heavily rely on theoretical calculations to predict the efficiency of such a cut.

We choose a different method. We relax the lepton momentum cut to 1.0 GeV/c, the same as used for the $\pi\ell\nu$ decays and bin coarsely in the variables q^2 and $\cos\theta_{W\ell}$. By binning we can sample more of phase-space without diluting the signal strength in those regions where the signal to background ratio is high. Figure 5.2 shows the effect of a lepton momentum cut in the $\cos\theta_{W\ell}$ versus q^2 plane generator-level Monte Carlo that implements the 2004 Ball form-factor calculations [15]. Typical $b \to c\ell\nu$ backgrounds populate the region where $\cos\theta_{Wl}$ is less than zero[2]; therefore, a relatively high lepton momentum requirement will enhance the signal to background ratio. However, a high momentum requirement also means that an extrapolation over the remainder of the plane must be done based on a theoretical calculation. We choose a 1.0 GeV/c cut that includes a large fraction of the the rate and then divide the plane into the five bins shown in Figure 5.3. The 0-2 GeV^2 bin is used strictly for isolating the continuum background. The ranges are shown in Table 5.3.

[2]Typically $b \to c$ events enter the signal bin because extra missed particles boost the reconstructed neutrino momentum. This boost in the neutrino momentum causes the reconstructed $\cos\theta_{Wl}$ to be less than zero.
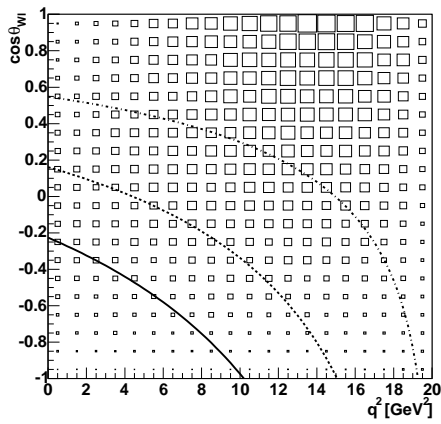
Figure 5.2: Lines indicating location of lepton momentum requirements of 1.0 (solid), 1.5 (dashed), and 2.0 (dot-dashed) GeV/c are overlaid on $B \rightarrow \rho\ell\nu$ generator-level Monte Carlo. The region below the line is excluded when the requirement is applied.
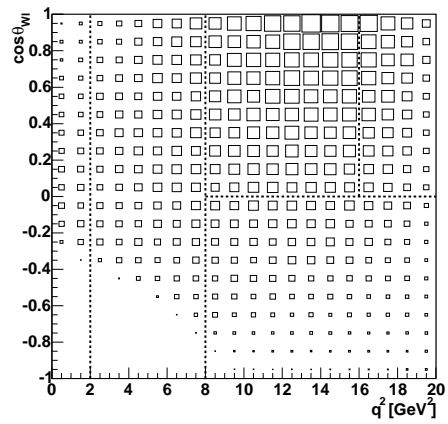
Figure 5.3: Binning in $\cos\theta_{W\ell}$ vs. q^2 is overlaid on toy signal $B \rightarrow \rho\ell\nu$ Monte Carlo. A 1.0 GeV/c lepton momentum cut has been applied while generating the plot.

Table 5.3: q^2 and $\cos\theta_{Wl}$ binning for $\rho\ell\nu$ modes

Bin #	q^2 Range $[GeV^2]$	$\cos\theta_{W\ell}$ Range
1	$0 < q^2 < 2$	$-1 < \cos\theta_{W\ell} < 1$
2	$2 < q^2 < 8$	$-1 < \cos\theta_{W\ell} < 1$
3	$8 < q^2 < 16$	$0 < \cos\theta_{W\ell} < 1$
4	$16 < q^2$	$0 < \cos\theta_{W\ell} < 1$
5	$8 < q^2$	$-1 < \cos\theta_{W\ell} < 0$

We fit directly for a rate in each of the four $\pi\ell\nu$ and five $\rho\ell\nu$ bins listed above. Due to low statistics we choose to sum over q^2 when reconstructing the $\eta\ell\nu$ mode. In the $\omega\ell\nu$ final state bins 1-4 of Table 5.3 are merged together, and bin 5 is reconstructed independently. This concentrates the low-statistics $\omega\ell\nu$ signal in one bin. As will be seen later, we use the $\omega\ell\nu$ mode to help constrain the $B \to \rho\ell\nu$ rate. The $B \to \eta\ell\nu$ rate is an independent parameter in the fit.

5.1.3 Decay Mode Binning

We bin the $\pi^\pm\ell\nu$, $\pi^0\ell\nu$, $\eta\ell\nu$, $\rho^\pm\ell\nu$, $\rho^0\ell\nu$, and $\omega\ell\nu$ modes independently of each other although their relative strengths will be constrained through isospin relations in the fit. In the $\eta\ell\nu$ mode, we further divide the reconstructed data into the two decay channels: $\eta \to \gamma\gamma$ and $\eta \to \pi^+\pi^-\pi^0$. These relative branching fractions for the η decay will be constrained in the fit.

As discussed in the previous chapter, combinatoric backgrounds are quite large in the vector final states due to the large intrinsic width of these mesons. Therefore we bin the ρ (ω) modes independently in three 190 (20) MeV/c^2 bins about the nominal ρ (ω) mass. When stating efficiencies and plotting data for these modes the central mass bin only is considered unless otherwise noted, but all three mass

bins are included in the fit.

5.1.4 Net Charge Binning

In order to promote quality neutrino reconstruction we require that our reconstructed event have net charge $|Q| \leq 1 (= 0)$ for pseudoscalar (vector) final states. The $|Q| = 1$ sample in the $\pi\ell\nu$ and $\eta\ell\nu$ channels has a considerably lower signal to background ratio, so we bin this sample separately for those final states. This prevents the high background in the $|Q| = 1$ sample from polluting the purer $Q = 0$ sample, but still allows the $|Q| = 1$ signal to contribute to the fit.

5.1.5 Bin Summary

The bin structure has been optimized in an attempt to both permit model-independent measurements to be made and to take advantage of regions where signal to background is high while allowing other more background prone regions to still contribute. Table 5.4 summarizes the 392 bins that are used in the fit.

Table 5.4: A summary of the bins used in the fit

Mode	$|Q|$	$q^2/\cos\theta_{W\ell}$	h_{decay}	M_h	$M_{h\ell\nu}/\Delta E$	Mode Total
$\pi^\pm\ell\nu$	2	4	1	1	7	56
$\pi^0\ell\nu$	2	4	1	1	7	56
$\eta\ell\nu$	2	1	2	1	7	28
$\rho^\pm\ell\nu$	1	5	1	3	7	105
$\rho^0\ell\nu$	1	5	1	3	7	105
$\omega\ell\nu$	1	2	1	3	7	42
Grand Total						**392**

5.2 Weights and Strengths

Before I further develop the components of the fit and fitting procedure a few lines should be devoted to defining the terms "weight" and "strength" as they will be used throughout the rest of this work.

As will be shown shortly appropriately selected weights can allow great flexibility and make up for multitude of simulation mistakes when dealing with the individual fit components. In general we have the ability to attach a weight to every candidate and it is these weights that we accumulate in the $M_{h\ell\nu}/\Delta E$ fit plane bins. For example if we realize that our simulation overestimates the π^0 efficiency by 4%, we can simply apply a weight of 0.96^n for candidates with n reconstructed π^0's and now the sum of the weights in the simulation will match what we expect in data. Typically weights vary from candidate to candidate within a bin. We can then write the weighted contents of a bin with N entries, $\tilde{N}$, as

$$\tilde{N} = \sum_{i=1}^{N} w_i = \bar{w}N, \tag{5.1}$$

where w_i are the individual weights. Re-weighting has implications when considering the statistical errors on the contents of a bin. For simplicity we make the assumption that the weights are roughly uniform within a bin and therefore the error on $\tilde{N}$ is the product of $\bar{w}$ and the statistical error on N. *We never apply weights to the reconstructed data that are being fit. We only re-weight the components with which we are fitting.*

When we refer to a "strength" we mean an overall scale factor for a fit component that spans all bins. Strengths are not event or even bin dependent adjustments. The strength of various components can be fixed or tied to freely floating parameters in the fit. If we were fitting a bin with D data entries to a single com-

ponent with bin contents $\tilde{N}$, we would vary the strength, s, such that $s\tilde{N} = D$.

Note that when we isolate any particular bin the only relevant numbers are the product of $s\bar{w}$ and the number of entries N. For any particular bin, the strength-adjusted sum of the weights is given by $s\bar{w}N$ and the appropriate Gaussian error is given by $s\bar{w}\sqrt{N}$.

5.3 Fit Components and Parameters

The goal is of the fit is to make a measurement of the number of signal events observed in the data sample. In order to do this we must understand not only the shape of the signal but also the shapes and normalization of the backgrounds in the fit.

5.3.1 Generic $b \to c$ Decays

Once our signal selection cuts have been applied, the largest remaining background is that from favored $b \to c$ decays of the two B mesons in the event. We use a Monte Carlo model based on the known inclusive and exclusive B branching fractions to charmed final states to model this background. We generate roughly a factor of five more simulated $B\bar{B}$ decays than we collect with the detector. We choose to float the strength of this component independently for each of the reconstructed modes, net charge bins, and hadronic submodes (for the η modes). This means that the simulation must properly model the rough shape of the background, but that the fit will correct for an overall scale in each of the independent modes. We find that this minimizes the effect of systematic uncertainties in our

$b \to c$ model. For our nominal fit this procedure introduces eleven free parameters.

We use the concept of event weighting to adjust this this sample of generic B decays, based on our present best knowledge of B decays. This allows us to fine tune and explore systematic dependences on the generated branching fractions and spectra of particles. For our nominal fit, we have applied the corrections listed below[3].

- By measuring the K_S^0 multiplicity in B decay and assuming that $N_{K_L^0} = N_{K_S^0}$ we can infer the K_L^0 multiplicity in B decay. We find that our simulation produces too few K_L^0, which ultimately results in better neutrino reconstruction on average in the simulation than in the data[38]. We therefore assign a weight of $(1.087)^{N_{K_L^0}}$ to each event which has the effect of increasing the total average number of K_L^0 per event by a factor of 1.072.

- The spectrum of secondary leptons, that is those from $b \to c \to s\ell\nu$ decay, affects both the neutrino reconstruction quality and the efficiency of our multiple lepton veto. We adjust this spectrum based on the convolution of the inclusive $B \to D^*X$ spectrum [39] and a measurement of the electron spectrum in inclusive semileptonic charm decay [40]. We divide this convolved spectrum by the generated spectrum and re-normalize based on the current world averages for inclusive $B \to DX$ and $D \to eX$ decay [6] to obtain the final set of weights.

- We correct the $B \to X_c\ell\nu$ branching fractions to match the latest CLEO result [41].

- Following the work done by Lipeles [42] we correct the $B \to D^*\ell\nu$ form

---

[3]These corrections and systematic uncertainties they introduce are discussed in greater detail in Chapter 6

factors to reflect more modern modeling and understanding of the shape of this decay.

- We eliminate decays of the type $B \to Y_c^{\text{baryon}} \ell \nu$, where Y_c^{baryon} is a charmed baryon. While these decays of this type are exist in our simulation, recent measurements [43] have set upper limits far below generated rates. Since these decays occupy a significant portion of the total inclusive semileptonic rate in our simulation we choose to boost up all other semileptonic decays so that the total inclusive rate remains unchanged.

5.3.2 Continuum Background

We use data taken approximately 60 MeV below the $\Upsilon(4S)$ resonance to model the residual $q\bar{q}$, two-photon, and τ-pair continuum background in the fit. Roughly $\frac{1}{2}$ $\left(\frac{1}{3}\right)$ of the total CLEO II (III) luminosity was taken at this off-resonance energy.

In order to model the shape of continuum background properly in the $M_{h\ell\nu}$ variable with the reduced off-resonance beam energy we must scale the value of $M_{h\ell\nu}$. We do so by setting

$$M_{h\ell\nu} \to \frac{E_{beam}^{\text{on}}}{E_{beam}^{\text{off}}} M_{h\ell\nu}, \tag{5.2}$$

where E_{beam}^{on} is the nominal single beam energy when running on the $\Upsilon(4S)$ and E_{beam}^{off} is the energy of the off-resonance point. In addition, recall that the last step of the neutrino reconstruction algorithm is to rotate the direction of the neutrino such that we force $M_{h\ell\nu} \to M_B$ in order to optimize q^2 resolution. For off-resonance data we rotate the neutrino to force $M_{h\ell\nu} \to M_B^{\text{eff}}$ where

$$M_B^{\text{eff}} \equiv \frac{E_{beam}^{\text{off}}}{E_{beam}^{\text{on}}} M_B. \tag{5.3}$$

Since we have at best $\frac{1}{2}$ of the total on-resonance luminosity taken at the off-resonance point and our continuum suppression algorithm is so effective at eliminating continuum events, the shape derived from this off-resonance data shows large statistical fluctuations from bin to bin. These fluctuations can be problematic when used directly as a component in the fit as they do not accurately represent the ideal smooth nature of the continuum background. We therefore adopt a continuum smoothing procedure.

In order to "smooth" the continuum data we reconstruct and bin the candidates without applying the continuum suppression cuts. This produces a relatively high-statistics distribution for the continuum component in the $M_{h\ell\nu}/\Delta E$ bins. We then re-weight the $M_{h\ell\nu}/\Delta E$ plane according a fitted bias function to reshape this high statistics distribution in the same fashion that our cuts would have shaped it. Finally we re-scale the distribution so that the total sum of the weights in $M_{h\ell\nu}/\Delta E$ fit plane is equivalent to the total number of events that passed our initial continuum suppression cuts. Figure 5.4 shows the effectiveness of this smoothing algorithm.

We study the bias that the suppression cuts introduce into the $M_{h\ell\nu}/\Delta E$ plane by removing the continuum suppression cuts and reconstructing the Monte Carlo in the $M_{h\ell\nu}/\Delta E$ plane[4]. We then apply the continuum suppression cuts to obtain a second set of $M_{h\ell\nu}/\Delta E$ histograms. Dividing this set by the first we obtain the bin-by-bin efficiency of our continuum suppression cuts across the plane. We fit a linear function of the form

$$\eta(M_{h\ell\nu}, \Delta E; \alpha, \beta) = \bar{\eta}\left(1 + \alpha\frac{\Delta E}{0.3~GeV} + \beta\frac{(M_{h\ell\nu} - 5.23125~GeV/c^2)}{0.0225~GeV/c^2}\right). \quad (5.4)$$

using the bin centers and contents of the efficiency plane to determine the param-

---

[4]To do this study we use a 5×5 array of bins in $M_{h\ell\nu}$ and ΔE. The bins range from $5.175 - 5.2875~GeV/c^2$ in $M_{h\ell\nu}$ and $-0.75 - 0.75~GeV$ in ΔE.
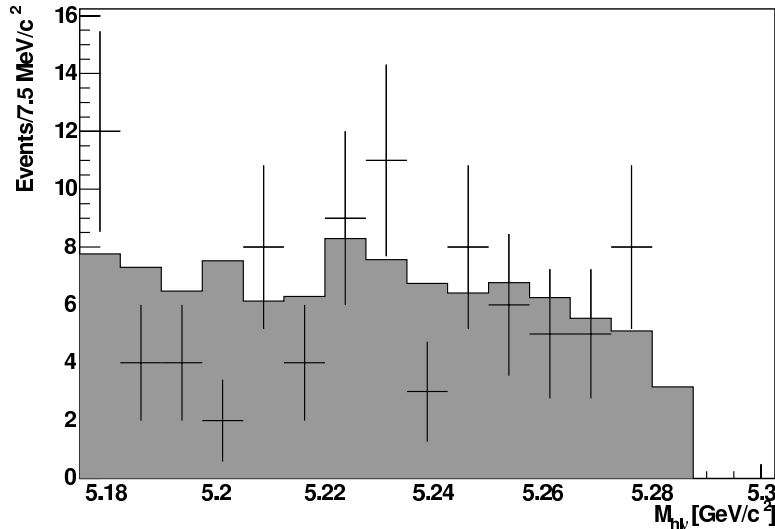
Figure 5.4: The $M_{h\ell\nu}$ distribution for $B \to \pi\ell\nu$ candidates from continuum background is shown. Events passing the continuum suppression cuts are displayed by points with errors. The smoothed representation of the continuum background is shown by the filled histogram.

eters α and β and their respective errors. The average efficiency across the plane is $\bar{\eta}$ and is fixed in the fit. This procedure is repeated in every bin for which we have a different set of Fisher discriminant weights.

A final correction must made before using the continuum component in the fit in order to be certain the statistical error bar on the smoothed sample is properly calculated. In the smoothing algorithm, statistical errors enter in two places:

- in the bin to bin fluctuations of the histograms that have no continuum suppression applied. These are the fluctuations present in the shaded histogram in Figure 5.4.

- in the total number of events the $M_{h\ell\nu}/\Delta E$ fit plane that is used to set the overall normalization of the smoothed distribution.

For each bin in the fit we maintain the same sum of weights. It can be shown that both of these statistical effects can be incorporated into a single number by using an "effective unweighted" number of events, N_i^{eff},

$$N_i^{\text{eff}} = \frac{N_i n}{N_i + n},$$

(5.5)

where N_i is the bin contents in the smoothed histogram and $n = \sum_i n_i$ is the sum of the entries in the $M_{h\ell\nu}/\Delta E$ plane for the sample with continuum suppression cuts applied. This sets the appropriate Gaussian error bar for the bin contents at $\bar{w}\sqrt{N_i^{\text{eff}}}$ where $\bar{w}$ is the average weight of all entries in the bin. This correction as stated is nearly insignificant; therefore, the fact that the statistical error on n is fully correlated for all seven $M_{h\ell\nu}/\Delta E$ bins in a particular histogram can be safely ignored.

We absolutely normalize this component in the fit by using the ratio of on- to off-resonance integrated luminosity $(\mathcal{L})$ and accounting for the energy dependence of the $e^+e^- \to q\bar{q}$ cross section. Specifically,

$$s = \left(\frac{E_{beam}^{\text{off}}}{E_{beam}^{\text{on}}} \right)^2 \frac{\mathcal{L}^{\text{on}}}{\mathcal{L}^{\text{off}}}$$

(5.6)

After smoothing, scaling, and properly accounting for all statistical fluctuations, the continuum component is fixed in the fit. From the point of view of the fit it is therefore treated as a straight subtraction with the appropriate error bars.

5.3.3 Fake Signal Leptons

In Section 4.2.3 we outlined a procedure for Monte Carlo lepton identification that forced the Monte Carlo efficiency and fake rates to agree with data. This relied on using only tracks tagged to a true lepton in Monte Carlo as a candidate

for an identified lepton. Therefore by construction our Monte Carlo simulation can never have real hadrons faking signal leptons. However, this can certainly happen in the real data, so this background must be accounted for.

We use Monte Carlo to model only the $B\bar{B}$ decays; therefore, we only need to determine the background component of fake signal leptons coming from $\Upsilon(4S)$ decays. This is done by first measuring the probability for pions, kaons, and protons to fake signal leptons using pure samples of these hadrons in data.[5] We then take a uniform subsample of the data (in this case, every 4th event) in which no lepton was identified. For each event we loop over the tracks in the event forcing our reconstruction code to identify each track first as an electron and then as a muon. With each track and each lepton species we then re-analyze the event and weight all candidates that come from that iteration of the analysis with the probability that the selected (hadron) track faked the forced lepton. The effect of this procedure is to produce a distribution of candidates from all potential signal leptons in the non-leptonic sample that is properly weighted by the probability that any one track will be identified as a lepton. We then normalize this distribution by the inverse of the sampling fraction.

As previously mentioned we only want to obtain this sample for $B\bar{B}$ decays. The continuum background distribution discussed in the section above is derived directly from data and therefore already has fake signal leptons in it. We therefore split the fake signal leptons distribution into off- and on-resonance components, both of which are fixed in the fit. The on-resonance component enters the fit with a strength of one, while, analogous to Equation 5.6 above, the off-resonance

---

[5]This exercise for muons in CLEO III is discussed in Appendix A.

component is given a strength, s, of

$$s = - \left(\frac{E_{beam}^{\text{off}}}{E_{beam}^{\text{on}}} \right)^2 \frac{\mathcal{L}^{\text{on}}}{\mathcal{L}^{\text{off}}}. \tag{5.7}$$

5.3.4 Signal $B \to X_u \ell \nu$ Decays

We generate signal Monte Carlo for each of the reconstructed modes. In each event we require that one of the B mesons decay to the desired signal mode while the other decays via our generic $b \to c$ decay model. We generate the signal decay using *phase space information only.* We can then use event weights to redistribute the sample according to the $V - A$ weak interaction coupling and any choice of form-factors. This allows us to explore the complete model dependence of our results by changing only the event weights used. For our nominal fit we use the recent unquenched LQCD results of Shigemitsu *et al.* [16] for the $\pi \ell \nu$ modes, LCSR results of Ball and Zwicky [15] for the $\rho(\omega) \ell \nu$ modes, and the ISGW2 model [11] for the $\eta \ell \nu$ mode.

We separate the signal Monte Carlo into the various q^2 and $\cos \theta_{Wl}$ bins based on the *true, generated* values of these variables. Note that this is different from the reconstructed values as the generated information gets smeared by our reconstruction resolution. We bin the fit by by the reconstructed values but vary the strengths of the signal sets binned by the generated kinematic variables. Through this process our experimental resolution and cross-feed is "unfolded" internally within the fit.

We seek to extract the rates in the bins listed in Tables 5.2 and 5.3. We therefore introduce one free parameter for each of the nine bins listed in these two tables. We fit specifically for the decay modes $B^0 \to \pi^- \ell^+ \nu$ and $B^0 \to \rho^- \ell^+ \nu$. The

relative rates of $B^0 \to \pi^- \ell^+ \nu$ and $B^+ \to \pi^0 \ell^+ \nu$ are constrained to be consistent with isospin symmetry assumptions, *i.e.* $\Gamma(B^0 \to \pi^- \ell^+ \nu) = 2\Gamma(B^+ \to \pi^0 \ell^+ \nu)$. Assuming that the total semileptonic width of charged and neutral B mesons is the same, we can write the strengths for the $\pi^0 \ell \nu$ components in terms of the freely floating $\pi^\pm \ell \nu$ strengths as

$$s_{\pi^0 \ell \nu} = \frac{1}{2} \frac{f_{+-}}{f_{00}} \frac{\tau_{B^+}}{\tau_{B^0}} s_{\pi^\pm \ell \nu}, \tag{5.8}$$

where f_{+-}/f_{00} is the ratio of charged to neutral B production at the $\Upsilon(4S)$ and τ_{B^+}/τ_{B^0} is the ratio of charged to neutral B lifetimes. For our nominal fit we use the current values of $f_{+-}/f_{00} = 1.026$ and $\tau_{B^+}/\tau_{B^0} = 1.078$ produced by the Heavy Flavor Averaging Group [44].

We constrain the relative strengths of the charged and neutral $\rho \ell \nu$ modes in a similar fashion. For the $\omega \ell \nu$ mode we argue the the quark content and mass of the omega is similar enough to the ρ^0 that we can assume $\Gamma(B^+ \to \omega \ell^+ \nu) \approx \Gamma(B^+ \to \rho^0 \ell^+ \nu)$. We therefore constrain $s_{\omega \ell \nu}$ to be the same as $s_{\rho^0 \ell \nu}$ which is related to $s_{\rho^\pm \ell \nu}$ through a relation similar to Equation 5.8 above. Note that even though we only *reconstruct* $\omega \ell \nu$ in two bins in the $q^2/\cos \theta_{Wl}$ plane, we divide the signal Monte Carlo into the same five *generated* bins as we do for the ρ modes as the strengths of the $\omega \ell \nu$ samples are tied directly to the $\rho \ell \nu$ parameters. In this way the fitted $\omega \ell \nu$ data act only to further constrain the $\rho \ell \nu$ fit.

In the $\eta \ell \nu$ modes we separate the generated signal Monte Carlo into three sub-samples according to η decay mode: $\gamma\gamma$, $\pi^+ \pi^- \pi^0$, and other η decays. Note that we only explicitly reconstruct the first two of these samples; however, the strengths of all three of the samples are constrained based on current branching fraction measurements [6].

It is impossible to overstate the importance of distinguishing between the re-

constructed modes that the data are fit in and the true generated Monte Carlo samples with strengths that are tied to parameters in the fit. For example in instances where large $\rho\ell\nu \to \pi\ell\nu$ cross-feed exists our fit will raise the cross-feed background in reconstructed $\pi\ell\nu$ bins simultaneously with raising the signal components in reconstructed $\rho\ell\nu$ bins. In the end the four $\pi\ell\nu$ parameters, five $\rho\ell\nu$ parameters, and one $\eta\ell\nu$ parameter add an additional ten free parameters to the fit. These are the only "handles" we provide the fitter tune the shape of the exclusive signal components and cross-feed backgrounds in the fit.

5.3.5 Other $B \to X_u\ell\nu$ Decays

A final background arises from other $B \to X_u\ell\nu$ decays that we are not exclusively reconstructing in the fit. In order to model this background we use a hybrid exclusive-inclusive Monte Carlo developed and documented by Meyer [45] that combines ISGW2 predictions of exclusive decays [11] with the inclusive lepton spectrum predicted using HQET by De Fazio and Neubert [18]. The parameters used in the heavy quark expansion are constrained by recent measurements of the $B \to X_s\gamma$ photon spectrum by CLEO [46]. The model attempts to generate resonant decays according to ISGW2 plus non-resonant decays such that the total generated spectrum matches theoretical predictions. We explicitly remove our exclusive decays from this simulation, and the remainder of the sample becomes the $B \to X_u^{\text{other}}\ell\nu$ component of the fit.

We constrain the strength of this component using the the recent measurement of the lepton endpoint branching fraction by the BaBar collaboration [47]. BaBar measures the branching fraction, $\mathcal{B}(B \to X_u e^+\nu)$, where the electron momentum is in the range of 2.2 to 2.6 GeV/c (the "endpoint region") to be $(2.35 \pm 0.22) \times 10^{-4}$.

Roughly 10% of the generated $B \rightarrow X_u^{\text{other}} \ell\nu$ spectrum has a lepton in this range. By summing over the strengths of the exclusively reconstructed modes and knowing the fraction that each mode contributes to the inclusive endpoint rate, we set strength of the $B \rightarrow X_u^{\text{other}} \ell\nu$ component in the fit force that the total $B \rightarrow X_u e^+ \nu$ rate to agree with the measured value.

5.3.6   Constraints Between Data Sets

Since we are doing a combined fit of the CLEO II, II.V, and III data we produce the fit components above for each data set and constrain their strengths to respect the integrated luminosity and total number of $B\bar{B}$ events in each of the data sets. This ensures that changes in shape or efficiency produced by various detector changes enter the final fit with appropriate strengths. Note that we fit the sum of all three data sets.

5.3.7   Parameter Summary

Table 5.5 provides a summary of all of the components in the fit and the free parameters that they introduce. We have listed the the components from top to bottom in the same order that they will appear in all plots in this work. Because of the coupled nature of the fit there is the potential for every component to exist in any plot. Armed with binned data and a selection of components to describe the data, we are ready to perform the fit and extract the signal.

Table 5.5: A comprehensive summary of the components and free parameters used in the fit. Components are listed top to bottom in the same order that they will be plotted throughout this work.

Name	# of Free Parameters	Notes	Plot Style
Signal $X_u\ell\nu$	4 $\pi\ell\nu$ + 5 $\rho\ell\nu$ + 1 $\eta\ell\nu$ = **10**	strengths are attached to bins of *generated* q^2 and $\cos\theta_{Wl}$; efficiency and cross-feed is unfolded in the fit; relative charged and neutral rates fixed by isospin relations	Signal π X-Feed ρ/ω X-Feed η X-Feed
Other $X_u\ell\nu$	0	from hybrid inclusive/exclusive model; strength is constrained by fixing the total inclusive $B \rightarrow X_u\ell\nu$ rate in the lepton endpoint	
Fake Signal Leptons ($B\bar{B}$ only)	0	constructed from non-leptonic data using measured fake rates; strength fixed by on-off subtraction	
$q\bar{q}$ Continuum	0	measured below the $\Upsilon(4S)$; strength fixed by considering $\mathcal{L}$ and cross section	
$b \rightarrow c$	11	from model of known inclusive and exclusive decays; floats freely in each reconstructed mode and each $|Q|$ bin	

5.4 Performing the Fit

We now have a set of data that is binned in the 392 bins listed in Table 5.4. The fit strategy is simply to vary the free parameters in the fit components until the sum of the fit components best match the data across all bins. To implement this strategy we use the technique of a binned likelihood fit.

5.4.1 A Binned Likelihood Fit

As an example let's consider one of the 392 fit bins. This bin, the ith bin, has data contents d_i. We can write the sum of the fit components for this bin as

$$f_i = \sum_{j=1}^{m} s_j \bar{w}_{ji} a_{ji},$$
(5.9)

where the index j runs over all of the fit components, s_j is the strength of the jth component and $\bar{w}_{ij}$ and a_{ji} are the average weight and number of entries in the ith bin of the jth fit component[6]. We then vary the s_j to minimize the difference between d_i and f_i over all bins simultaneously.

One such method of minimization is the method of maximum likelihood. We can write the Poisson distributed probability with an average of μ events to fluctuate to x events as:

$$\mathcal{P}(x; \mu) = e^{-\mu} \frac{\mu^x}{x!}.$$
(5.10)

We would then write the probability that the sum of the fit components, f_i, fluctuates to d_i events in the ith bin as $\mathcal{P}(d_i; f_i)$ which depends on the the strengths

---

[6]Throughout this discussion we use i to denote bins, where $n = 392$ is the total number of bins in the fit. We will use j to denote fit components, where m is the total number of components in the fit.

s_j through the expression for f_i above. We can then vary the s_j to maximize the likelihood

$$\mathcal{L} = \prod_{i=1}^{n} \mathcal{P}(d_i; f_i) \tag{5.11}$$

It is practically more useful; however, to maximize the likelihood ratio

$$\lambda = \prod_{i=1}^{n} \frac{\mathcal{P}(d_i; f_i)}{\mathcal{P}(d_i; d_i)}, \tag{5.12}$$

rather than the likelihood $\mathcal{L}$, because in the large statistics limit, $-2 \ln \lambda$ provides us with a χ^2 distributed variable. Note that in practice we simply minimize $-2 \ln \lambda$ itself and can therefore interpret the minimum value as a test of the goodness of fit. Note that the logarithm converts the product into a more algorithmically manageable sum and we can write

$$-2 \ln \lambda = -2 \sum_{i=1}^{n} d_i (\ln f_i - \ln d_i) + d_i - f_i. \tag{5.13}$$

5.4.2 Managing Finite Fit Component Statistics

Recall that our fit components are not smooth functional forms but rather produced by generating and analyzing simulation data as described in the previous section. Because we can only practically generate a finite amount of simulation data, it becomes necessary to account for the statistical fluctuations present in *the fit components themselves.* In order to do this we implement a method proposed by Barlow and Beeston[7] [48].

The core of the problem lies in the fact that some the jth fit component contributes a finite number of events, a_{ji}, to the ith bin where a_{ji} is actually Poisson

---

[7] What appears here is a condensed summary of the method presented in [48]. I omit technical details and special cases concerning the solution and merely attempt to give the reader a feel for the procedure. Where possible I have tried to maintain notation consistent with the reference. One exception being that Barlow and Beeston choose to use p to denote the strength of a component – I will continue to use s.

distributed about some mean number of events A_{ji}. It is the statistical fluctuations of a_{ji} from the mean A_{ji} that must be properly accounted for in the maximization procedure.. Therefore, instead of Equation 5.11, the likelihood we would really like to maximize is

$$\mathcal{L} = \prod_{i=1}^{n} \prod_{j=1}^{m} \mathcal{P}(d_i; f_i) \mathcal{P}(a_{ji}; A_{ji}), \tag{5.14}$$

or

$$\ln \mathcal{L} = \sum_{i=1}^{n} d_i \ln f_i - f_i + \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ji} \ln A_{ji} - A_{ji}. \tag{5.15}$$

Equation 5.9 becomes

$$f_i = \sum_{j=1}^{m} s_j \bar{w}_{ji} A_{ji}, \tag{5.16}$$

Again in practice we construct a ratio of likelihoods analogous to Equation 5.12 so that we may use the minimum as an indication of goodness of fit. Note that now we have accounted not only for the probability that the data contents d_i fluctuate to the predicted f_i, but also the probability that any *actual* contribution, a_{ji}, is a fluctuation of the A_{ji} expected in the infinite statistics limit.

Equation 5.15 includes $m \times n$ parameters A_{ji} and n parameters s_j that are unknown and must be determined in the process of maximizing the likelihood. We obtain a solution by differentiating Equation 5.15 with respect to these sets of variables and setting the derivatives to zero. This leaves n equations when differentiating with respect to the s_j, the jth of which looks like

$$\sum_{i=1}^{n} \frac{d_i A_{ji}}{f_i} - A_{ji} = 0. \tag{5.17}$$

Differentiating with respect to the A_{ji} leaves $n \times m$ equations. Indexed over i and j these equations have the form

$$\frac{d_i \bar{w}_{ji} s_j}{f_i} - \bar{w}_{ji} s_j + \frac{a_{ji}}{A_{ji}} - 1 = 0. \tag{5.18}$$

By defining a variable, t, where

$$t_i \equiv 1 - \frac{d_i}{f_i}, \tag{5.19}$$

we can recast Equation 5.18 as

$$A_{ji} = \frac{a_{ji}}{1 + \bar{w}_{ji} s_j t_i}. \tag{5.20}$$

Now for a given set of strengths, s_j, we have written the A_{ji} in terms of these s_j and n variables, t_i. In order to determine the n values t_i we must solve the set of n equations:

$$\frac{d_i}{1 - t_i} = f_i = \sum_{j=1}^{m} \bar{w}_{ji} s_j A_{ji} = \sum_{j=1}^{m} \frac{\bar{w}_{ji} s_j a_{ji}}{1 + \bar{w}_{ji} s_j t_i} \tag{5.21}$$

These are now n uncoupled equations which depend on the chosen set of s_j and the actual bin contents, a_{ji}, of the sources. Operationally these equations are solved using Halley's method to give values for the t_i. The t_i are then used to calculate the A_{ji}. Once the A_{ji} have been calculated then the likelihood can be determined for a given set of source strengths. The source strengths are then varied to minimize the likelihood.

5.5   Fit Results

Figures 5.5 - 5.10 show $M_{h\ell\nu}$ and ΔE projections for the nominal fit. These figures are generated by plotting $M_{h\ell\nu}$ for candidates with ΔE in the signal bin and vice versa. Note the the binning is chosen to show detailed peak structure and is somewhat different from the binning within these variables that fitter actually uses (see Table 5.1). We have summed over charged and neutral modes for the $\pi\ell\nu$ and $\rho\ell\nu$ modes. Within modes we maintain the same vertical scale between q^2 and

Table 5.6: A summary of the central values for the exclusive branching fractions for $B^0 \to \pi^- \ell^+ \nu$ obtained in the nominal fit. The errors shown are statistical only.

q^2 $[GeV^2]$	$\int \frac{d\Gamma}{dq^2} dq^2 / \Gamma_{Total} [10^{-4}]$
$q^2 < 2$	0.12 ± 0.07
$2 < q^2 < 8$	0.26 ± 0.07
$8 < q^2 < 16$	0.54 ± 0.08
$16 < q^2$	0.40 ± 0.07
Total	**1.32 ± 0.15**

$\cos \theta_{Wl}$ bins to demonstrate the distribution of both the signal and the background in these variables.

The nominal fit converges with $-2 \ln \lambda$ equal to 391 for 392 - 21 degrees of freedom. Recall that we fit explicity for the $\pi^{\pm} \ell \nu$ and $\rho^{\pm} \ell \nu$ yields. Given this yield, $Y_{\pi/\rho \ell \nu}$, we can obtain the branching fraction, averaged over lepton species, as

$$\mathcal{B}(B^0 \to \pi^-/\rho^- \ell^+ \nu) = \frac{Y_{\pi/\rho \ell \nu}}{4 f_{00} N_{\Upsilon(4S)}}, \qquad (5.22)$$

where f_{00} is the fraction of neutral B mesons produced at the $\Upsilon(4S)$. The factor four comes from a factor of two to convert $N_{\Upsilon(4S)}$ to N_B and another factor of two to average over electrons and muons. For the $\eta \ell \nu$ mode we obtain a similar expression:

$$\mathcal{B}(B^+ \to \eta \ell^+ \nu) = \frac{Y_{\eta \ell \nu}}{4 f_{+-} N_{\Upsilon(4S)}}, \qquad (5.23)$$

Tables 5.6 and 5.7 summarize the branching fractions and *statistical errors* obtained in the nominal fit for the $\pi \ell \nu$ and $\rho \ell \nu$ modes. In the $\eta \ell \nu$ mode we obtain the result: $\mathcal{B}(B^+ \to \eta \ell^+ \nu) = (0.45 \pm 0.25) \times 10^{-4}$.

In order to test that our fit properly models the shape of the various variables used in the candidate selection algorithm we plot these variables with the fit components superimposed. The values obtained in the nominal fit are used to

Table 5.7: A summary of the central values for the exclusive branching fractions for $B^0 \to \rho^- \ell^+ \nu$ obtained in the nominal fit. The errors shown are statistical only.

$q^2 \; [GeV^2]$	$\cos \theta_{Wl}$	$\int \frac{d\Gamma}{dq^2 \; d\cos\theta_{Wl}} dq^2 \; d\cos\theta_{Wl}/\Gamma_{Total}[10^{-4}]$
$q^2 < 2$	$-1 <\cos\theta_{Wl}< 1$	0.43 ± 0.19
$2 < q^2 < 8$	$-1 <\cos\theta_{Wl}< 1$	0.90 ± 0.19
$8 < q^2 < 16$	$0 <\cos\theta_{Wl}< 1$	0.71 ± 0.14
$16 < q^2$	$0 <\cos\theta_{Wl}< 1$	0.32 ± 0.06
$8 < q^2$	$-1 <\cos\theta_{Wl}< 0$	0.37 ± 0.17
Total		**2.73 ± 0.36**

Figure 5.5: $M_{h\ell\nu}$ projections of the nominal fit for $Q = 0$, $\pi^{\pm}\ell\nu$ and $\pi^0\ell\nu$ modes summed together. Fit components are as described in Table 5.5
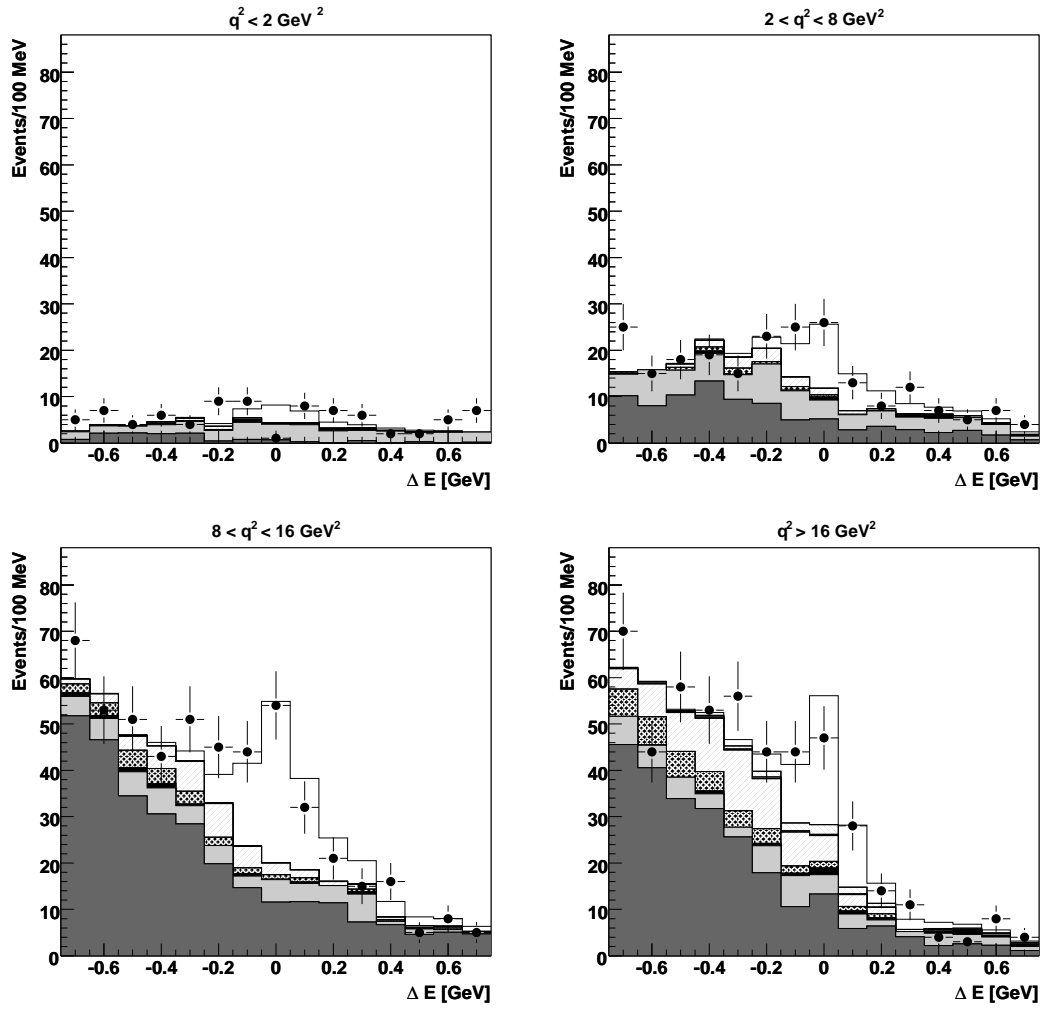
Figure 5.6: ΔE projections of the nominal fit for $Q = 0$, $\pi^\pm \ell \nu$ and $\pi^0 \ell \nu$ modes summed together. Fit components are as described in Table 5.5
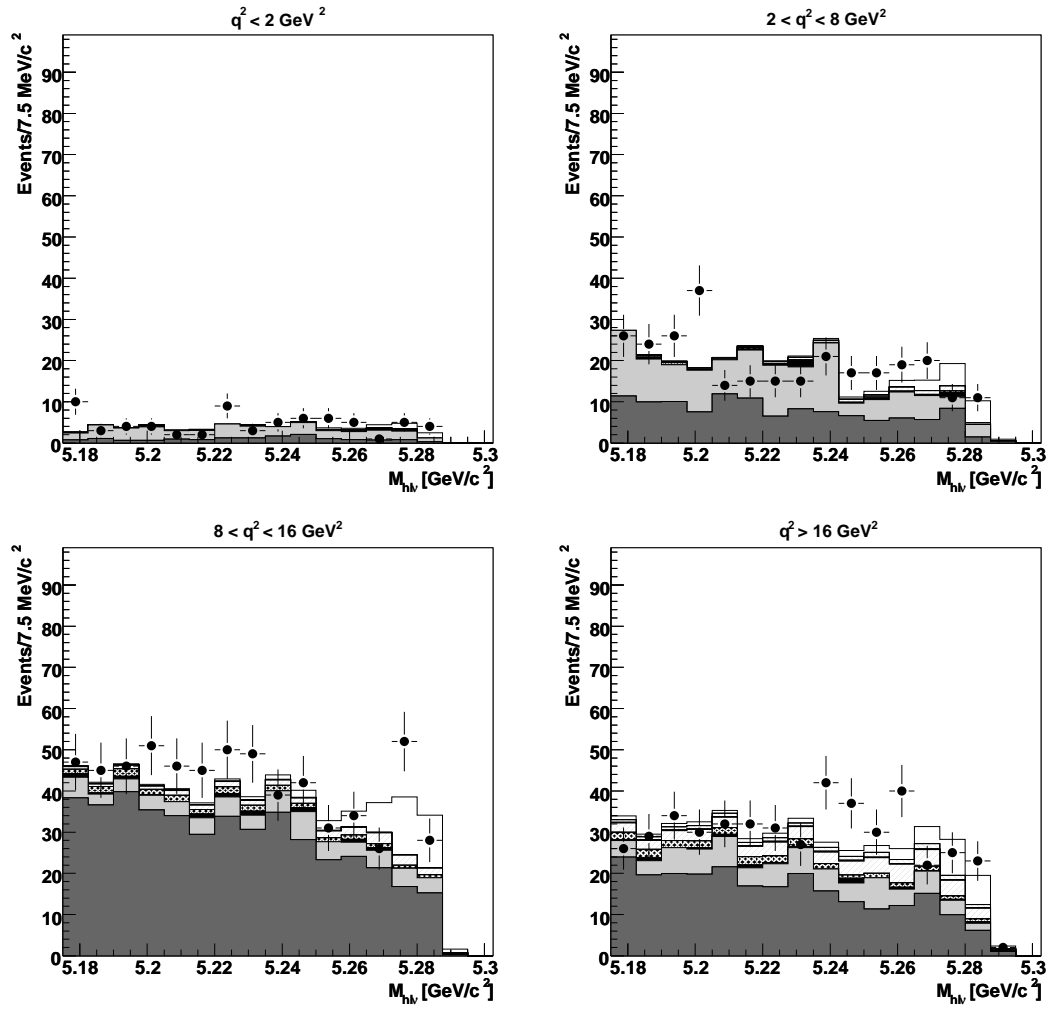
Figure 5.7: $M_{h\ell\nu}$ projections of the nominal fit for $|Q| = 1$, $\pi^{\pm}\ell\nu$ and $\pi^0\ell\nu$ modes summed together. Fit components are as described in Table 5.5
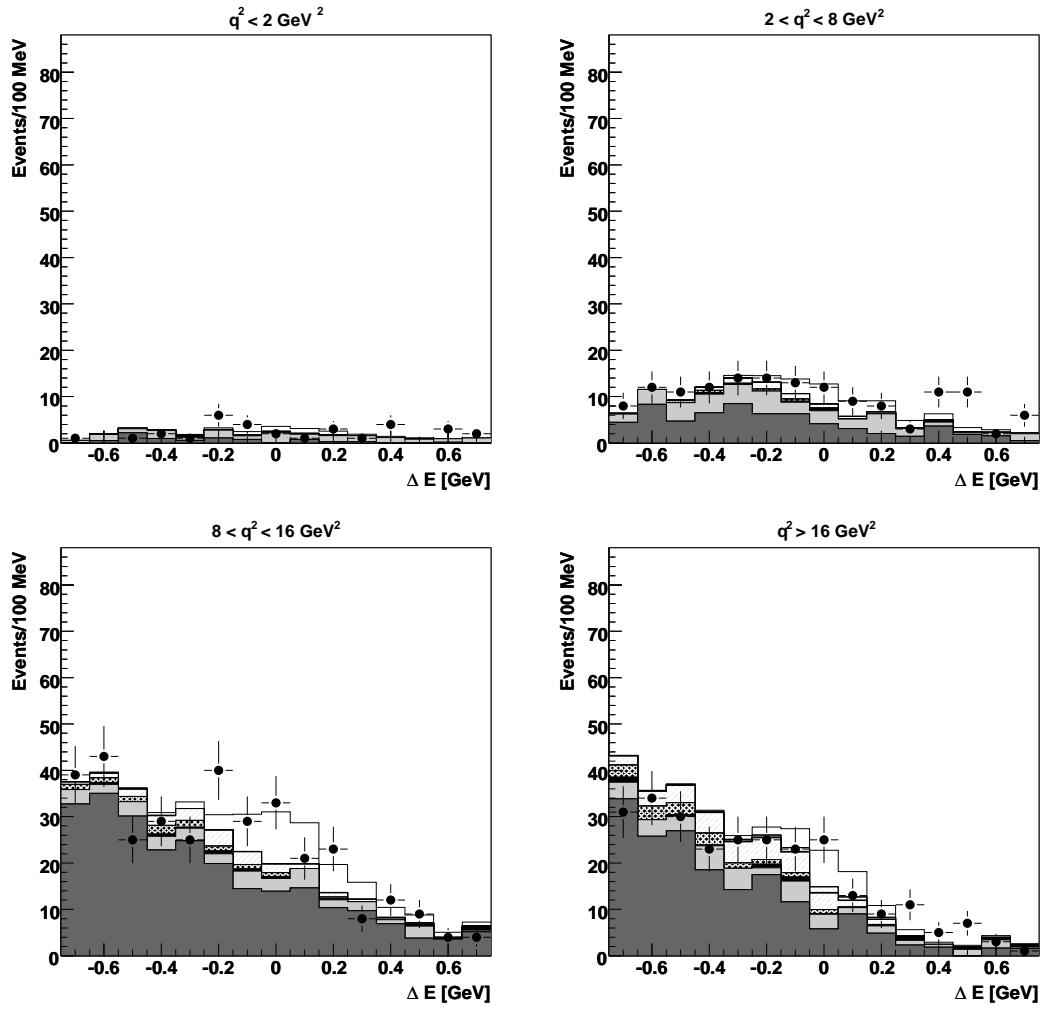
Figure 5.8: ΔE projections of the nominal fit for $|Q| = 1$, $\pi^{\pm}\ell\nu$ and $\pi^0\ell\nu$ modes summed together. Fit components are as described in Table 5.5

Figure 5.9: $M_{h\ell\nu}$ projections of the nominal fit for $Q = 0$, $\rho^{\pm}\ell\nu$ and $\rho^0\ell\nu$ modes summed together. Fit components are as described in Table 5.5
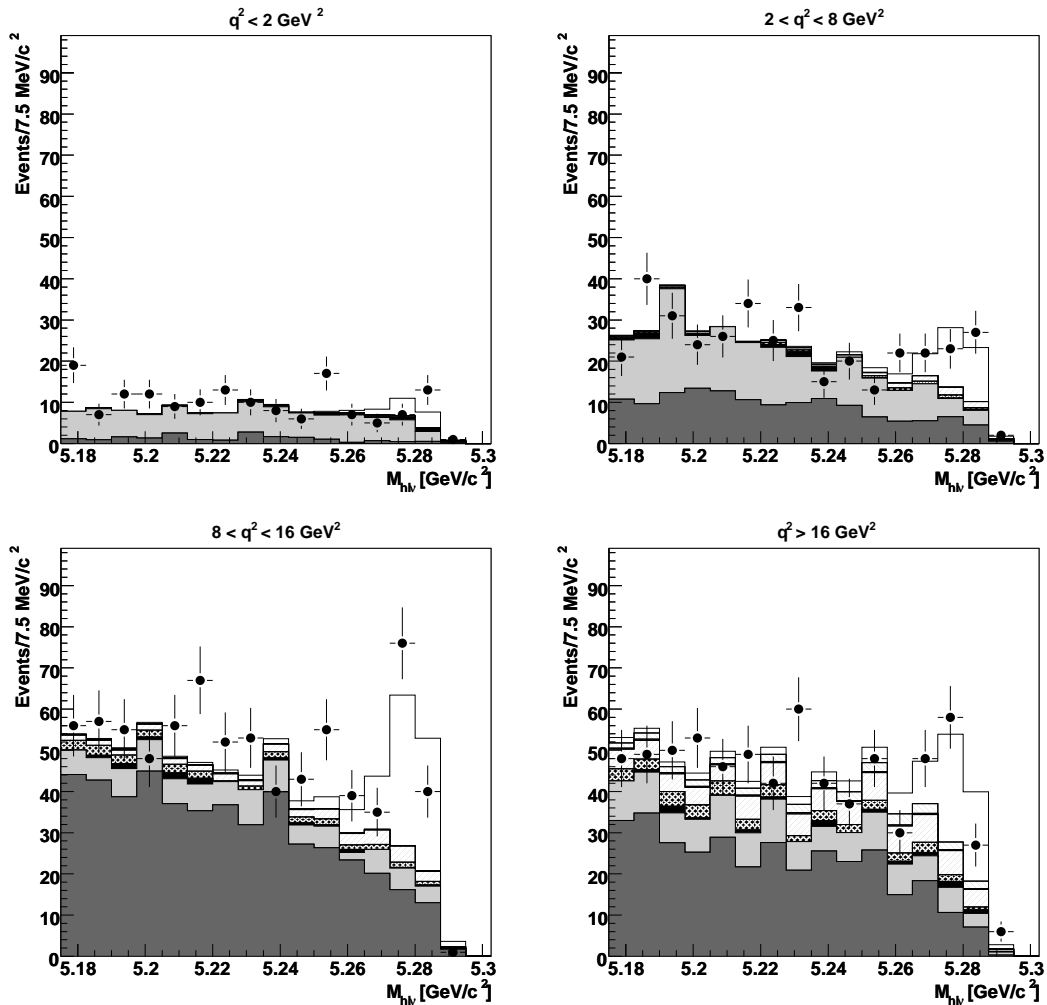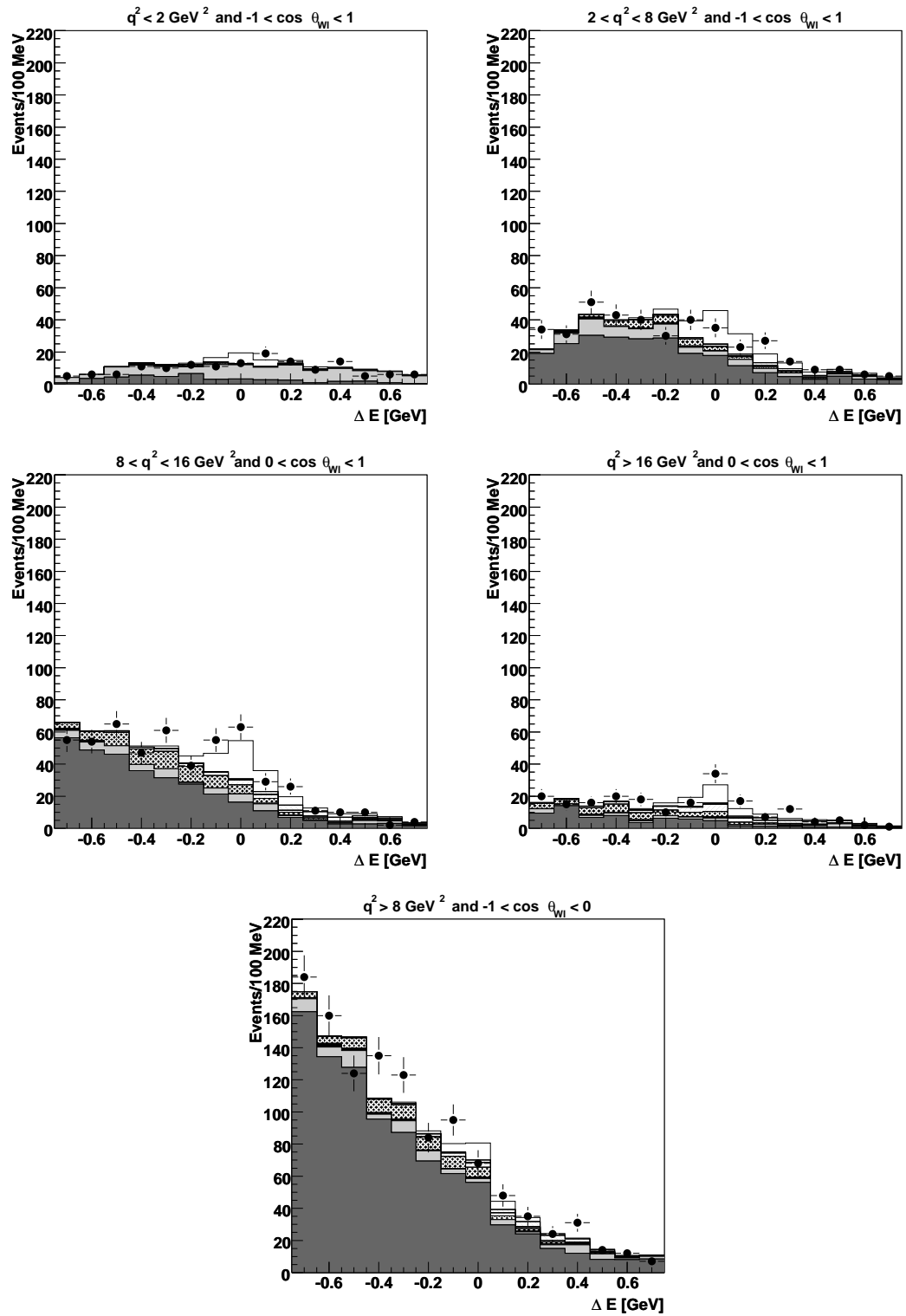
Figure 5.10: ΔE projections of the nominal fit for $Q = 0$, $\rho^{\pm}\ell\nu$ and $\rho^0\ell\nu$ modes summed together. Fit components are as described in Table 5.5
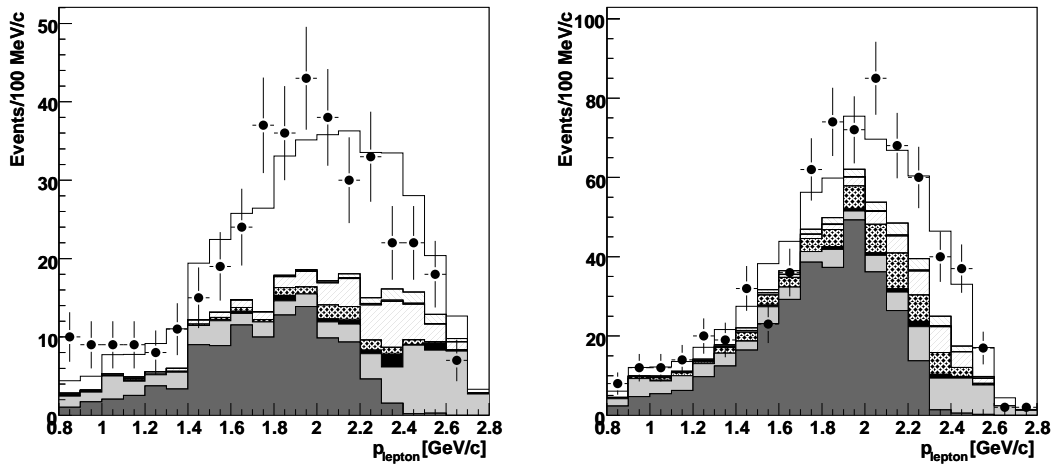
Figure 5.11: Lepton momentum projections of the nominal fit for $Q = 0$, $\pi^{\pm}\ell\nu$ and $\pi^0\ell\nu$ (left) and $\rho^{\pm}\ell\nu$ and $\rho^0\ell\nu$ (right). The data have been summed over q^2 and $\cos\theta_{Wl}$.

normalize the fit components.

Figure 5.11 shows the lepton momentum projections for both $\pi\ell\nu$ and $\rho\ell\nu$. As expected, the signal lepton spectrum for $\rho\ell\nu$ is noticeably harder than that for $\pi\ell\nu$. The agreement between the data and the sum of the fit components gives us confidence in the overall fit quality and our in ability to model the lepton momentum distribution of the fit components.

Figures 5.12 and 5.13 show the projections of $\cos\theta_{Wl}$ and the invariant mass of the two pions, $M_{\pi\pi}$, used to construct the ρ candidate for the $\rho\ell\nu$ modes. Again, the agreement between data and the sum of the fit components is encouraging. In the $\cos\theta_{Wl}$ variable the signal peaks near one and the $b \rightarrow c$ background is concentrated in the area where $\cos\theta_{Wl}$ is less than zero as expected. The $M_{\pi\pi}$ projection shows a nice peak in the signal component at the ρ mass. The peak in the highest bin is due to $B \rightarrow D^{(*)}\ell\nu$ backgrounds where the $D \rightarrow K\pi$ decay is identified as a $\rho \rightarrow \pi\pi$ decay. Recall that only the three bins about the ρ mass are
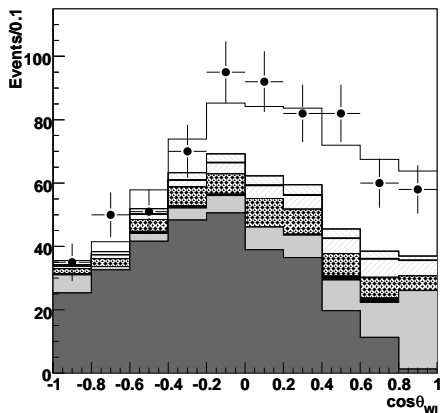
Figure 5.12: The $\cos\theta_{Wl}$ projection of the nominal fit, summed over q^2, for the $\rho^0\ell\nu$ and $\rho^\pm\ell\nu$ modes.
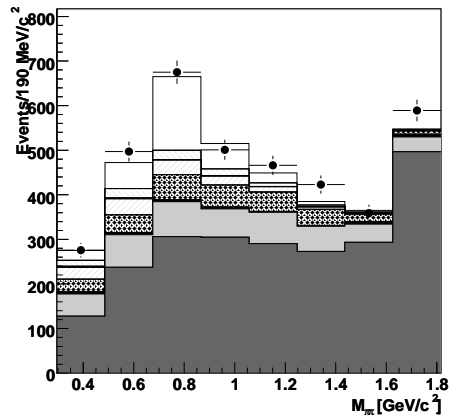
Figure 5.13: The projection of the two-pion invariant mass of the ρ candidates summed over q^2 and $\cos\theta_{Wl}$. The normalization of the fit components is determined from the nominal fit.

used in the fit.

The agreement shown above for various projections of the fit indicates that the simulation and fit components model the shapes observed in the data. The following chapter will evaluate the quality of this modeling in detail and assess the systematic errors on the results due to uncertainties present in the Monte Carlo and other inputs to the analysis algorithm.

Chapter 6

Systematic Uncertainties

The motivation behind evaluating a systematic error due to the uncertainty in a given aspect of the analysis is to ask oneself how the result would change if that aspect was different. These uncertainties can enter into the anlalysis in many ways, for example, in detector simulation, theoretical calculations, or even errors on other experimental results that are used the analysis. In this spirit we repeat the entire analysis multiple times changing each uncertain aspect individually to assess the impact of a given systematic uncertainty on the final result.

Wherever possible we have attempted to remove statistical fluctuations that may enter as a result of repeating the analysis. For example, in Monte Carlo electron identification we rely on random numbers coupled with data measured efficiencies to get the same identification efficiency in data as in Monte Carlo. When repeating the analysis for the sake of evaluation systematic errors we "freeze" this and other aspects of the analysis that depend on random numbers.

In all cases throughout this chapter I will quote errors in percentage deviation from the nominal results stated in the pervious chapter. We assess the error on the total rate for a particular exclusive channel independently of the error on the individual bins in q^2 or $\cos\theta_{Wl}$. Frequently the impact on the total measured rate will be smaller than the impact on the bins due to correlated changes among the bins.

6.1 Systematic Uncertainties in Neutrino Reconstruction

Neutrino reconstruction relies on our ability to reconstruct and identify every particle produced in a single collision. Since we are using Monte Carlo shapes in the $M_{h\ell\nu}/\Delta E$ plane to fit the data, it is necessary that all aspects of the Monte Calro, *e.g.* resolution, efficiency, etc., be the same as it is in data to avoid systematic bias.

Following the lead of the previous analysis [37] we are cautious not to underestimate our experimental systematic errors due to false cancellation. If we degrade the simulated resolution of the neutrino the simulated efficiency will drop. However, the fitted signal yield will also drop because the simulated signal bin now has a smaller fraction of signal events due to the degraded resolution on the neutrino. The lower raw yield combined with the lower efficiency can produce the same efficiency corrected yield and thereby negate the systematic error in neutrino resolution. Certainly this effect is "real" to some extent, but in cases where large cancellation occurs we do not trust that our simulation models this cancellation properly. In cases where the change in fit yield and efficiency have the same sign, *i.e.* there is some cancellation, we add an additional $\sqrt{2}/3$ of the magnitude of the cancellation, which is the minimum of the change in efficiency or yield. In other words, we assign an error on the canceling parts equal to one-third of their magnitude and add these errors in quadrature. The fraction one-third is purely an estimate, but is most likely conservative as there is a real cancellation effect.

Systematic errors due to neutrino reconstruction are summarized in Table 6.1 and are discussed in detail below. As expected, these systematic errors together tend to dominate the systematic error on the rate in most bins.

6.1.1 Track Efficiency and Resolution

In order to assess our susceptibility to systematic errors in the track finding efficiency of the Monte Carlo we repeat the analysis randomly dropping some fraction of the tracks. For CLEO II and II.V Berger notes [49] that systematic error on the Monte Carlo efficiency for track finding is 2.6% for tracks with momenta less than 250 MeV/c and 0.5% otherwise. The systematic error on low momentum tracking in CLEO III has not been extensively studied therefore, like Briere *et al.* [50], we use a conservative 5% error on low momentum tracks. The track finding error for high momentum tracks in CLEO III is taken to 0.5% in agreement with studies by Liu and Gao [51]. The analysis is repeated simultaneously dropping 2.6% (5%) of low momentum tracks for CLEO II/II.V (III) and 0.5% of high momentum tracks. This makes the conservative assumption that the tracking systematic error is fully correlated between CLEO II/II.V and CLEO III. The deviations from the nominal result are listed in Table 6.1 in the *Track Efficiency* row.

We explore potential biases due to errors in simulating track resolution by degrading the resolution in our Monte Carlo simulation. It is important to stress that errors in the result only arise when the resolution is *different* in Monte Carlo and reconstructed data. The absolute resolution is irrelevant. We choose to reset the momentum, $\vec{p}_{\mathrm{recon}}$, of each reconstruction track to

$$\vec{p}_{\mathrm{recon}} \to \vec{p}_{\mathrm{recon}} + \delta\vec{p}, \tag{6.1}$$

where

$$\delta\vec{p} \equiv \alpha \left(\vec{p}_{\mathrm{recon}} - \vec{p}_{\mathrm{true}} \right). \tag{6.2}$$

In Monte Carlo we determine the true momentum, $\vec{p}_{\mathrm{true}}$, by looking at the generator level information. This process increases the error on the reconstructed track by

some amount α. For CLEO II and II.V we choose α to be 10%. CLEO III studies of the D width in the decay of $D \to K\pi$ indicate that the Monte Carlo resolution should be smeared by 40% to broaden the width such that it agrees with data. The analysis is repeated smearing every Monte Carlo track by 10% (40%) in CLEO II/II.V (III), and the deviation from the nominal results are listed in the *Track Resolution* row of Table 6.1.

6.1.2 Shower Efficiency and Resolution

Shower reconstruction systematic errors are evaluated in a similar fashion to the tracking errors. Gritsan has limited the shower finding systematic error in CLEO II/II.V to 1.6% [52]. In CLEO III the shower finding systematic error is reduced to 1%. The analysis is again repeated while randomly dropping 1.6% (1%) of the showers in CLEO II/II.V (III) and the resulting systematic errors are noted in the *Shower Efficiency* row of Table 6.1.

Just as we did with tracks, the procedure outlined in Equations 6.1 and 6.2 is repeated with showers. We degrade all showers by 10%. Previous CLEO II/II.V studies and recent work by Muramatsu and Skwarnicki [53] indicate that actual resolution discrepancy is at or below the 10% level. The impact of this systematic error on the analysis is listed in the *Shower Resolution* row of Table 6.1.

6.1.3 Splitoff Simulation and Rejection

Potential systematic errors arise in the ability of the Monte Carlo to properly model so-called hadronic splitoff showers. In order to evaluate this error clean $\gamma\gamma \to K_S^0 K_S^0$ events with each K_S^0 decaying into $\pi^+\pi^-$ were compared in data and

Monte Carlo [54]. Based on the comparison of the multiplicity of hadronic showers in these events we add on average 0.03 showers per hadron to each event to explore systematic discrepancies between data and Monte Carlo. These showers are distributed uniformly throughout the detector and are drawn from a determined spectrum of splitoff showers which is concentrated at lower energies. Errors associated with this modeling mistake are shown in the *Splitoff Simulation* row of Table 6.1.

A additional systematic error in rejecting splitoff showers may arise in the variables used as inputs to the neural net that is used to identify the splitoff showers. In order to assess the impact of this systematic error we bias our neural net output in Monte Carlo. The bias procedure systematically pushes the neural net output for tagged photons in the "splitoff direction" and simultaneously pushes the output for showers not tagged to photons in the "photon direction" thus degrading the effectiveness of the algorithm. Meyer provides an extensive discussion on the implementation and effect of the bias [45]. Systematic changes in the result are listed in the *Splitoff Algorithm* row of Table 6.1.

6.1.4 Particle Identification

As mentioned earlier the ability to identify charged particles directly affects the missing energy resolution in neutrino reconstruction. We smear the time of flight (dE/dx) significance in CLEO II/II.V by $\frac{1}{2}$ ($\frac{1}{4}$) of the intrinsic resolution in the more pion-like direction. The degree of the smearing in each of these variables is thought to be a conservative estimate of the potential systematic error in detector information [55].

In CLEO III we find that degrading the proton and kaon probabilities based on detector information ($\mathcal{P}_D^s$ in Section 4.2.2) by 70% of themselves and turning off the production fraction weighting, *i.e.* setting $\mathcal{P}_D^s \to 1$, we produce systematic changes in the kaon and proton identification efficiencies at the 5% level. Studies using similar particle identification algorithms have limited systematic errors in CLEO III to less than 5% [56].

To determine the total impact due to systematic errors in particle identification we simultaneously degrade time of flight and dE/dx information in CLEO II/II.V and we degrade $\mathcal{P}_D^K$ and $\mathcal{P}_D^p$ by 70% of themselves in our CLEO III particle identification algorithm. The resulting error is shown in the *Particle Identification* row of Table 6.1.

6.1.5 K_L^0 Production and Energy Deposition

A substantial portion of the "missed particles" that inflate the resolution on the reconstructed neutrino are K_L^0's. We noted in Section 5.3.1 that by counting K_S^0's in B decay we infer that the K_L^0 multiplicity in the Monte Carlo should be increased by a factor of 1.072 [38]. The error on this determination is 1%, therefore we vary the weight up and down by 1% and average the absolute values of the deviation for a total error due to incorrect K_L^0 production in the Monte Carlo.

We also probe the effect of systematically increasing the energy deposited by K_L^0's in the calorimeter by increasing the energy of showers tagged to generator-level K_L^0's by 20%. This systematic error has a very minor impact on the result and is listed as K_L^0 *Energy Deposition* in Table 6.1.

6.1.6   Secondary Lepton Spectrum

As noted in Section 5.3.1 we correct the spectrum of secondary leptons based on the measurements of the inclusive $B \to D^*X$ spectrum [39] and the electron spectrum in inclusive semileptonic charm decay [40]. This spectrum is correlated with the spectrum of secondary neutrinos and therefore impacts neutrino resolution. We determine the allowable minimum and maximum weights as a function of momentum given the errors of these two measurements. We repeat the analysis using each of these sets of weights and then average the absolute values of the deviations to determine the potential systematic error due to uncertainties in the secondary lepton spectrum.

6.2   Additional Sources of Systematic Error

Aside from those systematic errors that directly impact the resolution of the reconstructed neutrino, there are a variety of other systematic errors to be considered. We attempt to isolate vulnerabilities in our analysis and or fit procedure due to dependancy on uncertain theoretical or experimental information. Again we proceed in the sprit of repeating the analysis and varying each unknown quantity independently to explore systematic effects.

6.2.1   Continuum Suppression

The continuum smoothing algorithm outlined in Section 5.3.2 relies on our ability to model the bias introduced in the $M_{h\ell\nu}/\Delta E$ plane by the continuum suppression cuts. Using an appropriate Monte Carlo sample the bias parameters,

Table 6.1: A summary of the systematic errors associated with neutrino reconstruction. The bin numbers are as listed in Tables 5.2 and 5.3 on pages 80 and 83. Errors from different sources are assumed to be uncorrelated. The bottom row shows the quadrature-sum of the errors.

| Systematic Error Source | $\delta\mathcal{B}\,(B^0 \to \pi^-\ell^+\nu)$ [%] | | | | | $\delta\mathcal{B}\,(B^0 \to \rho^-\ell^+\nu)$ [%] | | | | | |
	1	2	3	4	Total	1	2	3	4	5	Total
Track Efficiency	4.0	4.9	2.2	4.2	1.8	6.5	7.6	3.4	6.4	1.8	3.3
Track Resolution	1.5	1.7	1.6	1.3	1.4	9.2	13.8	2.4	4.1	40.5	1.9
Shower Efficiency	2.0	3.5	6.4	3.5	4.6	5.0	7.3	7.0	7.0	23.6	2.1
Shower Resolution	2.3	4.8	3.3	1.8	3.0	7.5	6.9	6.3	3.0	1.8	5.5
Splitoff Simulation	3.6	1.4	2.4	1.7	1.6	7.0	8.1	2.0	0.2	26.5	1.0
Splitoff Rejection	2.9	2.9	1.4	1.7	0.6	8.1	2.2	4.9	0.3	3.1	2.9
Particle Identification	0.6	2.2	0.7	4.3	1.1	7.3	12.6	1.1	3.6	18.0	2.3
K^0_L Production	0.2	0.1	0.3	0.2	0.2	0.6	0.9	0.1	0.3	4.0	0.2
K^0_L Energy Deposition	0.0	0.0	0.2	0.2	0.1	2.4	2.6	0.6	0.6	8.3	0.8
Secondary Lepton Spectrum	0.5	1.2	0.4	0.3	0.1	6.7	0.6	0.5	0.2	2.7	1.5
Total Neutrino Reconstruction Systematic Error	**7.0**	**8.9**	**8.3**	**7.7**	**6.3**	**20.7**	**24.2**	**11.6**	**11.4**	**57.7**	**8.2**

118

α and β, in Equation 5.4 are determined through a fit. We determine one pair of bias parameters and its error for each set of Fisher weights. This produces 26 pairs or 42 total parameters that are independent and uncorrelated. To assess the uncertainty due to the error on these bias parameters, we vary one of the parameters up and down one sigma as determined from the fit. We average the absolute value of the deviation from the nominal value and repeat the process for the remaining 45 parameters. The average deviations for each parameter are added in quadrature to produce the total systematic error due to our continuum suppression algorithm as listed in Table 6.2. Note that the dependance of the magnitude of this error on bin is as expected: low q^2 bins that contain large continuum backgrounds are affected most.

6.2.2 $\quad B \to X_c \ell \nu$

In the generic B decay simulation we have already applied weights to correct the generator level branching fractions of $B \to D\ell\nu$, $B \to D^*\ell\nu$, $B \to D^{**}\ell\nu$, and non-resonant $B \to X_c\ell\nu$ to agree with the previous CLEO result [41]. To evaluate sensitivity to this correction we vary each of these four rates up and down one sigma as reported in [41]. The magnitudes of the deviation are averaged for each variation to produce a systematic error for each of the four rates.

In addition we consider variations of the $B \to D^*\ell\nu$ form factor as it is the dominant $B \to X_c\ell\nu$ decay and a likely background. Following the example of Lipeles [42] we simultaneously vary the slope and curvature parameters, ρ^2 and c, of the Isgur-Wise function to span the set of allowed values. Specifically our nominal fit uses $(\rho^2, c^2) = (1.20, 0.7)$ and we vary these up to (1.51,1.39) and down to (0.88,0.0). The simultaneous motion accounts for correlations in the parameters.

We average the magnitude of the deviation and combine this in quadrature with the four variations of the branching fractions outlined in the previous paragraph. The full systematic uncertainty due to $B \to X_c \ell \nu$ decays is shown in Table 6.2. Again, as expected, maximum impact is seen at low q^2 and especially in the $\rho \ell \nu$ mode where $B \to X_c \ell \nu$ backgrounds are large.

6.2.3 Other $B \to X_u \ell \nu$

As noted in Section 5.3.5 we use a hybrid exclusive-inclusive Monte Carlo to model other $B \to X_u \ell \nu$ decays aside from the signal decay. We generate two additional sets of the Monte Carlo varying the heavy quark expansion parameters up and down as allowed by the CLEO $B \to X_s \gamma$ photon spectrum measurement [46]. We take the average magnitude of the deviation as the systematic error due to the modeling of the inclusive lepton spectrum in the decay.

Additionally the magnitude of the other $B \to X_u \ell \nu$ component in the fit is fixed by the BaBar endpoint measurement [47] as outlined in Section 5.3.5. We vary the endpoint rate within the errors allowed by the BaBar measurement and combine the average magnitude of the deviation in quadrature with the shape error described above to produce a total systematic error due to potential mis-modeling of the other $B \to X_u \ell \nu$ background.

6.2.4 Lepton Identification and Fake Leptons

Recall that electron identification in Monte Carlo is done by measuring electron efficiencies in data and then reproducing this efficiency in Monte Carlo using true generated electrons and random numbers. The statistical errors on the data mea-

surement of the efficiency dominate the systematic error on due to electron iden-
tification. Past studies have shown that muon identification efficiency in CLEO II
and II.V Monte Carlo agrees with data at the 2% level. The study discussed in
Appendix A shows that on average 2% errors or better can be achieved by applying
a weight to correct the efficiency which is applied in our nominal fit. All aspects of
our lepton identification algorithm are shown to be reproduced in Monte Carlo or
are artificially mocked up in Monte Carlo with accuracies at or below 2%. Since
we have one signal lepton per decay we simply take 2% as the systematic error on
the efficiency and therefore branching fraction due to lepton identification.

We rely on measurements of the rates for hadrons faking leptons in two ways in
our fit and therefore should consider potential systematic error due to statistical
errors in the measurement of the fake rates. First, we use the fake rates to produce
a fit contribution that is from $B\bar{B}$ events in which the only lepton in the event is
a fake signal lepton. As can be seen from the fit projections this contribution is
minute and is dominated by the statistical error on the on-resonance minus off-
resonance subtraction procedure. This statistical error is already accounted for
in the fit; therefore, we neglect potential systematic errors in the production of
this fit component. Secondly the data-measured fake rates are used to veto events
in Monte Carlo that have an additional lepton which arises from a hadron faking
a lepton. A systematic error in the fake rate would affect the efficiency of the
multiple-lepton veto and therefore affect the signal efficiency. We systematically
lower the fake rates by $\frac{1}{2}$ to probe the sensitivity of our result to the data-measured
fake rates. The deviations are listed in the *Fake Leptons* row of Table 6.2.

6.2.5 π^0 Identification

The CLEO III Monte Carlo is known to overestimate the π^0 efficiency. This is likely due to improper modeling of the pre-shower of the photon in the RICH detector or other material in front of the calorimeter. In data this pre-shower degrades the shower shape and produces an efficiency loss. Cassel *et al.* measure the discrepancy between data and Monte Carlo to be 4% [57]. Therefore in our nominal fit we apply a weight of 0.96 to CLEO III Monte Carlo signal decays that contain a π^0 in the final state. Note that π^0 finding is only used in the process of reconstructing signal decays. To explore the effect of systematic errors in CLEO III π^0 identification, we set the weight used in the fit to one, effectively removing the correction that exists in the nominal fit.

6.2.6   Number of $\Upsilon \to B\bar{B}$ Events

Based on luminosity, cross section, and event shape studies the error on the number of $B\bar{B}$ events in CLEO II + II.V is taken to be 2%. It is further assumed that the error is strongly correlated between the CLEO II and II.V data set such that the ratio the number $B\bar{B}$ events in each is well known. In order to determine the number of $B\bar{B}$ events in the CLEO III data set Alexander *et al.* measure the ratio of $B \to D\pi$ in CLEO II + II.V and in CLEO III [58] to an accuracy of 8%. Since we lock the relative ratios of CLEO II, II.V, and III in the fit and each of these detector configurations has a different reconstruction efficiency, we decide to probe what effect the uncertainty in the II + II.V : III ratio has on the result. The ratio is varied $\pm 8\%$ while keeping the total number of $B\bar{B}$ events constant and the resulting deviation is negligible. We combine the 2% error for the number of $B\bar{B}$

events in CLEO II and II.V with the 8% error on the ratio of number of $B\bar{B}$ events in II + II.V : III to produce a total error on the number of $B\bar{B}$ events of 3.6% that accounts for correlations between the two measurements. This translates directly to a 3.6% systematic error on the branching fractions from the uncertainty in the total number of $\Upsilon \to B\bar{B}$ events.

6.2.7 τ_{B^+}/τ_{B^0} and f_{+-}/f_{00}

The fixed relative strengths of the charged and neutral rates are sensitive to both the lifetime and production ratios of charged and neutral B mesons through Equation 5.8. We vary both the lifetime and production fraction ratios independently within the error bars produced by the Heavy Flavor Averaging Group [44]. Additionally the error on f_{+-}/f_{00} affects our calculation of the total number B^0 mesons produced through Equation 5.22. This effect combined with correlated effect on the charged/neutral constraint in the fit is listed in Table 6.2. The independent effect of the error on τ_{B^+}/τ_{B^0} is listed separately.

6.3 Dependance on Form Factors

Past attempts at measurement of the $B \to \pi \ell \nu$ and $B \to \rho \ell \nu$ rates have been plagued by the theoretical uncertainty regarding the form factors for these decays. As noted before, experimental efficiency is not uniform across the allowed kinematic space for these decays and the experimentalist must rely on theoretical calculations to extrapolate what is measured to the full rate. Our choice of binning in the q^2 and $\cos\theta_{Wl}$ variables allows a rate measurement with minimal dependance on theoretically calculated form factors.

Table 6.2: A summary of the experimental systematic errors. The bin numbers are as listed in Tables 5.2 and 5.3 on pages 80 and 83. Errors from different sources are assumed to be uncorrelated. Neutrino reconstruction systematic errors are copied directly from the totals summarized in Table 6.1.

| Systematic Error Source | $\delta\mathcal{B}\,(B^0 \to \pi^- \ell^+ \nu)$ [%] | | | | | $\delta\mathcal{B}\,(B^0 \to \rho^- \ell^+ \nu)$ [%] | | | | | |
	1	2	3	4	Total	1	2	3	4	5	Total
Neutrino Reconstruction	7.0	8.9	8.3	7.7	6.3	20.7	24.2	11.6	11.4	57.7	8.2
Continuum Suppression	9.8	2.9	0.9	1.0	1.1	9.3	1.8	1.4	1.3	3.4	1.5
$B \to X_c \ell \nu$	3.9	4.0	1.2	0.9	1.5	20.8	12.2	2.3	2.3	13.7	5.9
Other $B \to X_u \ell \nu$	2.0	1.4	2.0	5.0	2.3	5.6	2.5	7.6	4.9	16.6	3.1
Fake Leptons	4.7	1.2	2.3	0.5	1.8	3.1	2.6	1.6	1.0	5.4	1.2
Lepton Identification	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
π^0 Identification	0.4	0.2	0.2	0.2	0.1	1.0	0.9	1.4	0.9	3.3	1.4
Number of $\Upsilon \to B\bar{B}$	3.6	3.6	3.6	3.6	3.6	3.6	3.6	3.6	3.6	3.6	3.6
τ_{B^+}/τ_{B^0}	0.5	0.3	0.4	0.4	0.4	0.7	0.7	0.6	0.8	0.4	0.7
f_{+-}/f_{00}	0.5	0.8	0.7	0.7	0.7	0.1	0.0	0.2	0.2	0.7	0.1
Total Experimental Systematic Error	**14.3**	**11.1**	**9.9**	**10.2**	**8.3**	**31.7**	**27.7**	**14.9**	**13.4**	**62.2**	**11.6**

To evaluate the dependance on the form factors, we choose a set of calculations that span the extreme range of theoretical results. We then re-weight our signal Monte Carlo according to each of these form factor calculations to evaluate the systematic error of the form factor on the rate measurement. We vary the $B \to \pi\ell\nu$ and $B \to \rho\ell\nu$ independently as any change can affect all measured rates in the fit due to cross feed. We take the systematic error on rate due to the $B \to \pi\ell\nu$ form factor to be 1.7 times the RMS of the set of rates produced with the form factor variations and likewise for $B \to \rho\ell\nu$ rates. We inflate the RMS by a factor of 1.7 to account for potential variations of the internal parameters of any the calculated form factors. This inflation factor is purely an empirical result derived from exploring variations within individual models and was used in the previous CLEO analysis [37].

6.3.1   The $B \to \pi\ell\nu$ Form Factor

The unquenched lattice QCD result of Shigemitsu *et al.* [16] are used in the nominal fit. We probe the allowed space of form factors by also repeating the analysis using the calculations of Ball and Zwicky [14], Scora and Isgur (ISGW2) [11], and Feldmann and Kroll [13] for the $B \to \pi\ell\nu$ form factor. We use the nominal Ball and Zwicky result for the $B \to \rho\ell\nu$ form factor [15] throughout this study. The variation in q^2 dependance of the rate from these calculations is illustrated in Figure 6.1. We take the RMS of the results in each bin for the four calculations inflated by 70% of itself as the systematic error due to the $B \to \pi\ell\nu$ form factor uncertainty.
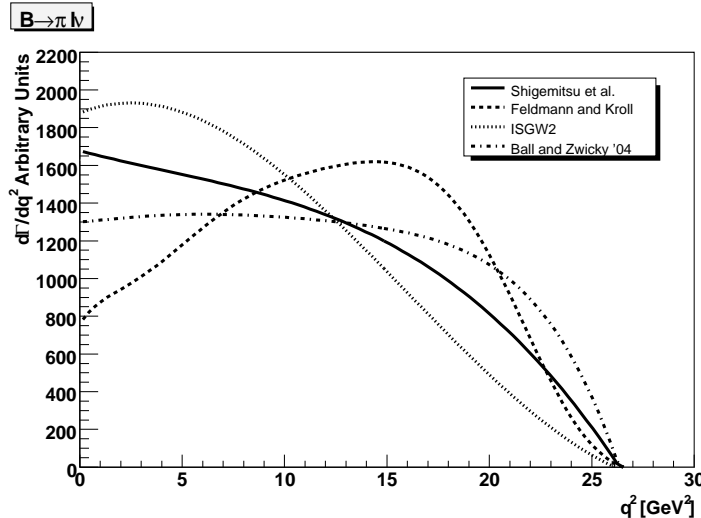
Figure 6.1: An illustration of the collection of theoretical predictions for the q^2 dependance of the $B \to \pi\ell\nu$ rate that is used to probe the systematic dependance of the result on the form factor calculation.

6.3.2 The $B \to \rho\ell\nu$ Form Factor

As previously mentioned, our nominal fit used the light-cone sum rules calculation of the $B \to \rho\ell\nu$ form factor of Ball and Zwicky [15]. The calculations of Melikhov and Stech [12], the UKQCD collaboration [17], and Scora and Isgur (ISGW2) [11] are used to test the dependance of the rate on a broad spectrum of possible form factors. The procedure is analogous to that used to evaluate the dependance on the $B \to \pi\ell\nu$ form factor. Throughout the study we use the nominal $B \to \pi\ell\nu$ form factor calculated by Shigemitsu *et al.* [16]. Figure 6.2 shows the the calculated rates as a function of q^2 and $\cos\theta_{Wl}$. Note that the variation between calculations for the rate as a function of $\cos\theta_{Wl}$ is significant; therefore, our attempt to coarsely bin in this variable is rewarded with reduced systematic error.

The systematic errors due to uncertainties in the signal decay form factors are
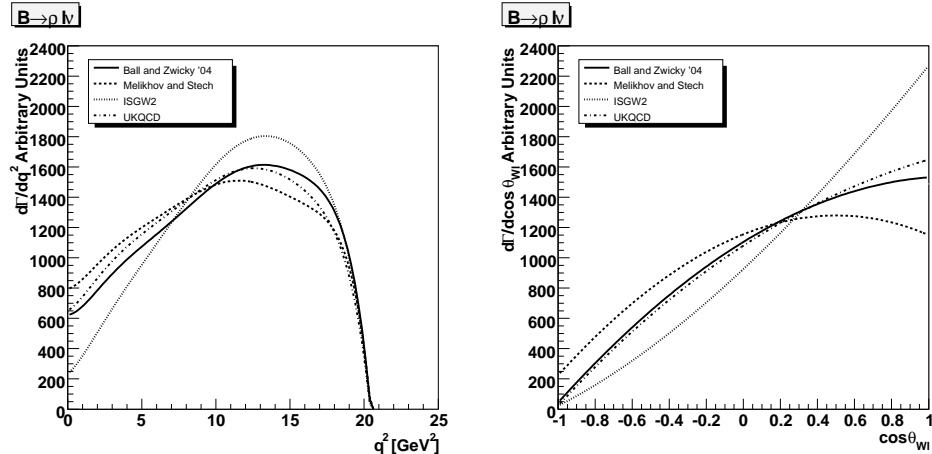
Figure 6.2: Shown are the predicted rate as a function of q^2 (left) and $\cos\theta_{Wl}$ (right) for the set of form factor calculations used to study the systematic error on the rate due to the uncertainty in the $B \to \rho\ell\nu$ form factor.

Table 6.3: The systematic errors associated with the individual $B \to \pi\ell\nu$ and $B \to \rho\ell\nu$ form factor uncertainties are listed below. The bin numbers are as listed in Tables 5.2 and 5.3 on pages 80 and 83.

Form Factor	$\delta\mathcal{B}\left(B^0 \to \pi^-\ell^+\nu\right)$ [%]					$\delta\mathcal{B}\left(B^0 \to \rho^-\ell^+\nu\right)$ [%]					
	1	2	3	4	Total	1	2	3	4	5	Total
$B \to \pi\ell\nu$	1.6	0.6	0.7	3.5	1.0	0.4	0.5	1.1	0.4	3.4	0.6
$B \to \rho\ell\nu$	0.5	0.8	1.5	1.2	1.0	2.8	5.7	1.4	2.0	9.6	1.4

summarized in Table 6.3. Dramatic improvements over the previous CLEO result have been made in minimizing the effect of the $B \to \rho\ell\nu$ form factor uncertainty. In fact the error bar on the total $B \to \rho\ell\nu$ rate due to form factor uncertainties has been reduced by an order of magnitude.

Chapter 7

Summary and Conclusions

I have presented a measurement of the of the $B^0 \to \pi^- \ell \nu$ and $B^0 \to \rho^- \ell \nu$ rates in bins of the momentum transfer variable, q^2, and the lepton decay angle, θ_{Wl}. The rates were measured using 15.4 million $\Upsilon(4S) \to B\bar{B}$ events recorded with the CLEO II, II.V, and III detectors. The analysis technique utilizes neutrino reconstruction in order to exclusively reconstruct the products of the semileptonic B decay. A binned likelihood fit is used to simultaneously extract the rates for the various decay modes and kinematic variable bins. Careful systematic studies were carried out to assess both the experimental and theoretical systematic errors.

7.1 Summary of the Rates

In order to produce a measurement with appropriate errors we combine the central values and statistical error bars determined in the fit in Chapter 5 with the fractional systematic errors that were studied in Chapter 6. Historically results have depended strongly on the choice of form factor used to analyze and interpret the data. In order to emphasize this, the systematic error is often separated show to independently show the portion due to form factor uncertainty. This dependance, studied in the latter portion of the previous chapter, has been greatly reduced in this analysis. To highlight this I will combine the uncertainties due to $B \to \rho \ell \nu$ and $B \to \pi \ell \nu$ form factors but still list this uncertainty separately from other systematic errors.

Tables 7.1 and 7.2 summarize the rate measurements and the relevant statistical

Table 7.1: A summary of the exclusive branching fractions for $B^0 \to \pi^- \ell^+ \nu$ obtained in this analysis. The errors shown are statistical, systematic, and form factor associated respectively.

$q^2 \; [GeV^2]$	$\int \frac{d\Gamma}{dq^2} dq^2 / \Gamma_{Total} [10^{-4}]$
$q^2 < 2$	$0.12 \pm 0.07 \pm 0.02 \pm 0.00$
$2 < q^2 < 8$	$0.26 \pm 0.07 \pm 0.03 \pm 0.00$
$8 < q^2 < 16$	$0.54 \pm 0.08 \pm 0.05 \pm 0.01$
$16 < q^2$	$0.40 \pm 0.07 \pm 0.04 \pm 0.01$
Total	$\mathbf{1.32 \pm 0.15 \pm 0.11 \pm 0.02}$

Table 7.2: A summary of the exclusive branching fractions for $B^0 \to \rho^- \ell^+ \nu$ obtained in this analysis. The errors shown are statistical, systematic, and form factor associated respectively.

$q^2 \; [GeV^2]$	$\cos \theta_{Wl}$	$\int \frac{d\Gamma}{dq^2 \; d\cos\theta_{Wl}} dq^2 \; d\cos\theta_{Wl} / \Gamma_{Total} [10^{-4}]$
$q^2 < 2$	$-1 < \cos\theta_{Wl} < 1$	$0.43 \pm 0.19 \pm 0.13 \pm 0.01$
$2 < q^2 < 8$	$-1 < \cos\theta_{Wl} < 1$	$0.90 \pm 0.19 \pm 0.25 \pm 0.05$
$8 < q^2 < 16$	$0 < \cos\theta_{Wl} < 1$	$0.71 \pm 0.14 \pm 0.11 \pm 0.01$
$16 < q^2$	$0 < \cos\theta_{Wl} < 1$	$0.32 \pm 0.06 \pm 0.05 \pm 0.01$
$8 < q^2$	$-1 < \cos\theta_{Wl} < 0$	$0.37 \pm 0.17 \pm 0.23 \pm 0.04$
Total		$\mathbf{2.73 \pm 0.36 \pm 0.32 \pm 0.04}$

and systematic errors. In most bins, the total error is dominated by the statistical component. With the lowered lepton momentum cut and fitting of data with $\cos\theta_{Wl} < 0$ in the $B \to \rho\ell\nu$ modes, we have managed to produce measurements for both modes that have a residual dependance on the decay form factors that is at or below the 2% level.

7.2   Comparison with Other Measurements

We can compare our results for the total $B^0 \to \pi^- \ell\nu$ and $B^0 \to \rho^- \ell\nu$ branching fractions with other independent results from the BaBar and Belle collaborations.
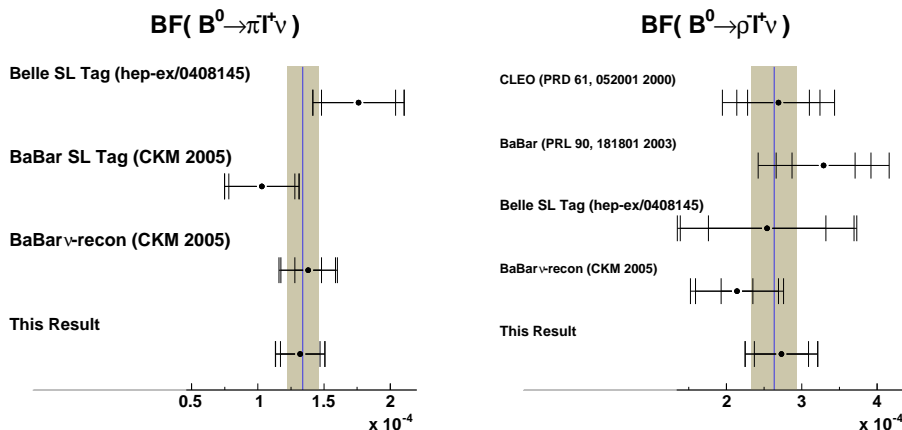
BF(B^0→π⁻l⁺ν)   **BF(B^0→ρl⁺ν)**

Belle SL Tag (hep-ex/0408145)   CLEO (PRD 61, 052001 2000)

BaBar SL Tag (CKM 2005)   BaBar (PRL 90, 181801 2003)

BaBar ν-recon (CKM 2005)   Belle SL Tag (hep-ex/0408145)

This Result   BaBar ν-recon (CKM 2005)

This Result

0.5   1   1.5   2   2   3   4
x 10^{-4} x 10^{-4}

Figure 7.1: The current world results for $B^0 \to \pi^- \ell \nu$ (left) and $B^0 \to \rho^- \ell \nu$ (right) branching fractions are shown. Error bars are statistical, experimental systematic, and form factor associated. The world average and one sigma error is shown. All of the errors are assumed to be uncorrelated in the computation of the average.

Figure 7.1 shows a comparison of the various independent measurements of the $B^0 \to \pi^- \ell \nu$ and $B^0 \to \rho^- \ell \nu$ branching fractions. The world average and one standard deviation error bar, assuming uncorrelated errors among the measurements, are shown. Not only do the results presented in this analysis agree remarkably with other measurements, they are also the most precise measurements to date of the total branching fractions. Furthermore as can be seen in Figure 7.1 for the first time a result for the total $B^0 \to \rho^- \ell \nu$ branching fraction has been produced with virtually no residual dependance on the $B^0 \to \rho^- \ell \nu$ form factor.

7.3 Final Thoughts

We have measured the $B^0 \to \pi^- \ell \nu$ and $B^0 \to \rho^- \ell \nu$ to better than the 15% and 18% level respectively, the best measurements available to date. Equally as important, we have successfully decoupled the experimental measurement from the

underlying theoretical uncertainty in the decay form factor. While the precision in the individual q^2 or $\cos\theta_{Wl}$ bins in this analysis is not high enough to effectively discriminate between the current form factor calculations, there are other routes to verification of form factors, such as high-statistics semileptonic D decay studies currently underway at CLEO. When the theoretical community produces verified, precise form factor calculations, either of the two branching fractions presented in this work can be used on its own to produce a sub-10% measurement for $|V_{ub}|$. We have therefore contributed an important experimental result that, when used in concert with other independent results and appropriate theoretical input, is capable pushing precision tests of the Standard Model to new extremes.

Appendix A

CLEO III Muon Identification Efficiency and Fake Rates

In general, we identify muons in two categories. The first category, so-called "counting muons," is subject to much looser requirements and used only when vetoing an event that has multiple leptons. The second category called "signal muons" is a smaller subset of these counting muons that is used for reconstructing signal $B \to X \mu \nu$ decays. The details of the identification algorithm are presented in Section 4.2.1. Here we examine the discrepancies between data and MC and make a measurement of the rate at which hadrons fakes muons.

A.1 Muon Efficiency

In order to check that our Monte Carlo simulation appropriately models the muon efficiency in data we measure the efficiency using radiative μ-pair events and compare this efficiency with what is observed for generated muons in the Monte Carlo.

A.1.1 True Muon Sample in Data

We need to select events of the type $e^+ e^- \to \mu^+ \mu^- \gamma$, where the γ is radiated by the μ that we wish to study. We do so by looking at events that have strictly two tracks and only one shower with energy > 1 GeV. We then require that
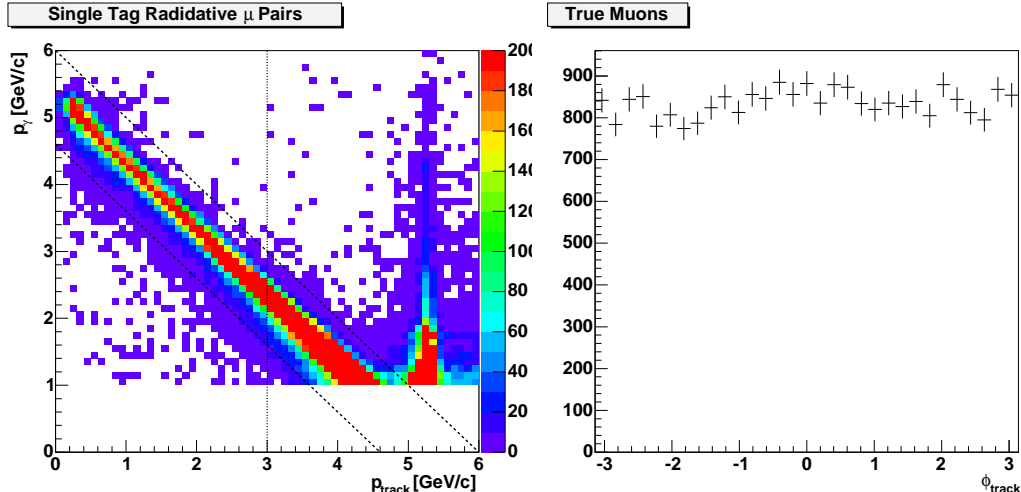
Figure A.1: (left) p_γ vs. p_{track} with selection criteria described in the text; (right) ϕ_{track}: no sign of cosmic ray contamination is visible

one of the tracks, call it μ_{tag}, satisfy the signal muon requirements above and also have $p_{tag} > 5.0\ GeV/c$. We can then plot p_γ vs. p_{track}, where p_{track} is the momentum of the *other* track in the event. This is shown in Figure A.1. The vertical band at $p_{track} \approx 5.2\ GeV/c$ is most likely due to pure $\mu - pair$ events where the interactions of one muon with the calorimeter produced the additional shower. For $p_{track} < 3.0\ GeV/c$ (the region we wish to study), we can isolate true muons by requiring $|p_\gamma + p_{track} - 5.3\ GeV/c| < 0.7\ GeV/c$. Finally we require that the track under study have $|\cos\theta| < 0.65$; therefore, we measure the efficiency within this fiducial volume. To verify that this sample is not contaminated by cosmic rays we plot ϕ_{track} also in Figure A.1 – no structure is observed.

A.1.2 True Muon Sample in Monte Carlo

Obtaining a sample of true muons in Monte Carlo is much simpler. We require that the generator-level particle be a muon and that it be matched to a

reconstructed track. In addition we require that this track be within the fiducial volume, $|\cos\theta| < 0.65$.

A.1.3   Divisions Among Run Ranges

Throughout this study we will define "CLEO III Data" as being data sets 6 through 14 inclusive. The corresponding "CLEO III Monte Carlo" is generated for all of those data sets and roughly in the appropriate proportions. We will use signal, $b \to ul\nu$ MC, as it provides a supply of leptons with momenta out to 2.8 GeV/c.

Midway through the CLEO III running a bug in the online muon sparsification code was discovered and repaired. Fixing this bug resulted in a noticeable change in efficiency; therefore, it is worthwhile to to compare efficiencies with and without the sparsification bug. All runs before run 116853 have the bug – from run 116853 onward the bug is fixed. The effect of the bug on the efficiency in data is show in Figure A.2.

A.1.4   Efficiency Comparison

Figure A.3 shows a comparison of the data and Monte Carlo efficiencies for muons as a function of momentum. From this plot one can conclude that the efficiency in Monte Carlo seems to be systematically too low by 1-2%. Figure A.4 shows a similar data/Monte Carlo comparison independently for runs before and after the sparsification fix. While the efficiency for runs taken with the bug is well modeled in the Monte Carlo, it seems that the Monte Carlo failed to reproduce the efficiency gain in data once the bug was fixed.
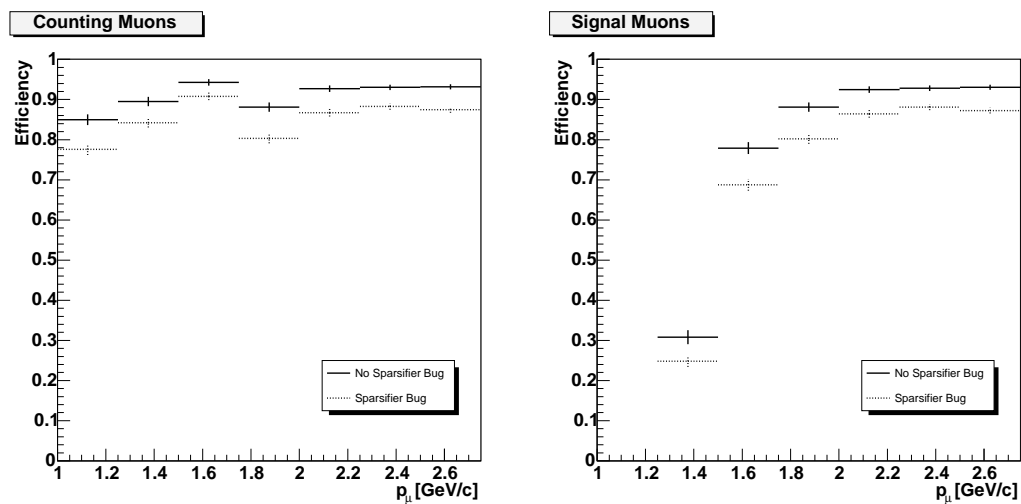
Figure A.2: Muon efficiencies in data for counting muons (left) and signal muons (right) for runs taken before and after the online sparsification bug was fixed.
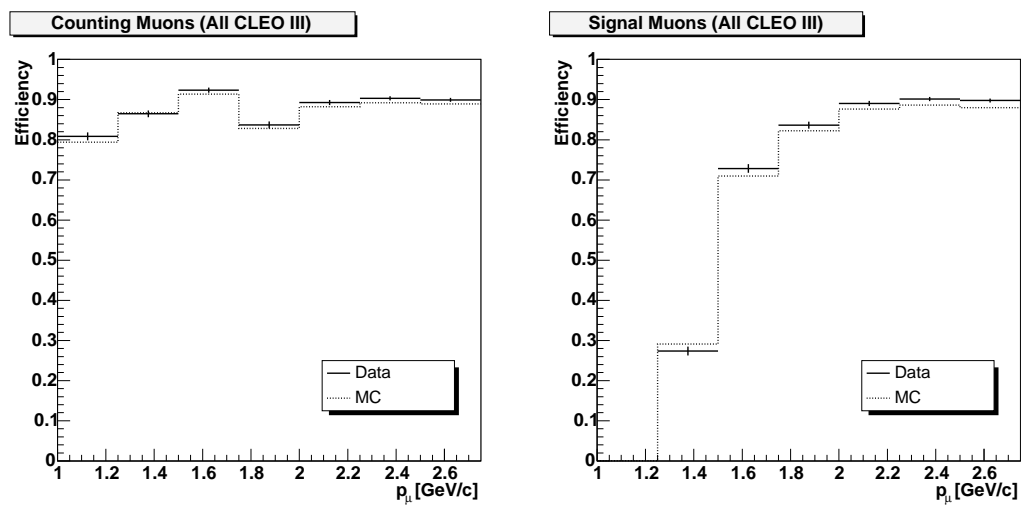

Figure A.3: Monte Carlo and data efficiencies as a function of momentum for counting muons (left) and signal muons (right) averaged over all CLEO III data.
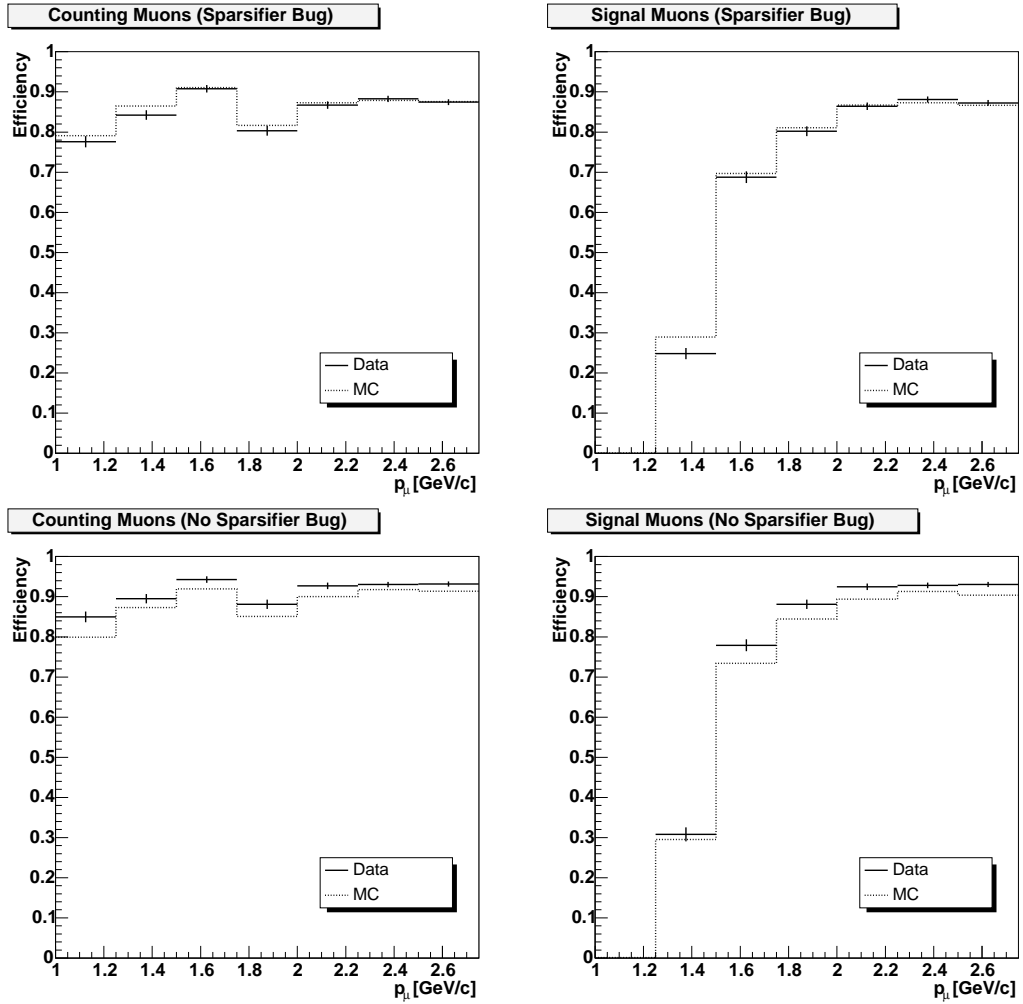
Figure A.4: Monte Carlo and data efficiencies for counting muons (left) and signal muons (right) for runs taken with the sparsification bug (top) and after the bug fix (bottom)

A.1.5 Efficiency Correction

In order to correct the data/Monte Carlo discrepancy we choose to apply a weight to each true muon in the Monte Carlo that is reconstructed as a signal or counting muon. Since the signal muons are strictly a subset of the counting muons we divide the sample into two parts: one for signal muons and one for counting but not signal muons. Note that this division only exists for muons with momenta less than 1.75 GeV/c. We further divide those samples in to two sets depending on whether the run was taken before or after the sparsification bug fix. The result is four sets of momentum dependent weights, one of which will be used for each Monte Carlo muon. These weights are given in the table below:

p_μ [GeV/c]	Run < 116853		Run $\geq$ 116853	
	Signal	Counting	Signal	Counting
1.00-1.25	–	0.981±0.018	–	1.065±0.017
1.25-1.50	0.858±0.049	1.033±0.028	1.044±0.060	1.016±0.033
1.50-1.75	0.987±0.021	1.032±0.062	1.061±0.020	0.881±0.071
1.75-2.00	0.989±0.015	–	1.044±0.014	–
2.00-2.25	0.997±0.011	–	1.035±0.009	–
2.25-2.50	1.009±0.009	–	1.016±0.008	–
2.50-2.75	1.007±0.010	–	1.029±0.009	–

A.2 Muon Fake Rates

We also need to measure the probability that a true pion, kaon, or proton will be reconstructed as a muon. Here we will make a data-based measurement and use this measurement directly in our analysis Monte Carlo instead of relying on
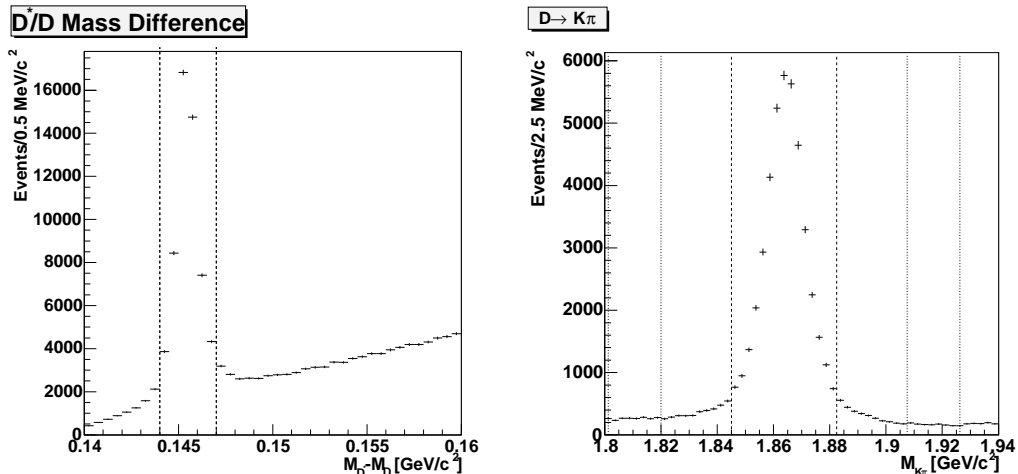
Figure A.5: (left) reconstructed $D^* - D$ mass, cut is shown by dashed line; (right) $K\pi$ mass for events that satisfy $D^* - D$ mass cut, signal region is indicated by the dashed line and the two sidebands are shown with dotted lines

the simulated fake rates. In order to measure the rate at which a true π, K, or p gets reconstructed as a muon we must have a pure source of these particles in data. We obtain a pure source of kaons and pions by reconstructing the decay $D^{*+} \rightarrow \pi^+ D^0$, $D^0 \rightarrow K^- \pi^+$ and its charge conjugate, where the charge of the intermediate π indicates which of the D daughters is the K or π. A pure source of protons can be obtained from reconstructing $\Lambda \rightarrow \pi p$.

A.2.1 π or K Faking μ

We run over the entire CLEO III data and collect a sample of D^* decays, where $D^{*+} \rightarrow \pi^+ D^0$, $D^0 \rightarrow K^- \pi^+$ (and the conjugate decay). We require only that all of the tracks in the decay be `TrkmanApproved`. The reconstructed $D^* - D$ mass is shown in Figure A.5. We require $|M_{D^*} - M_D - 145.5\ MeV/c^2| < 1.5\ MeV/c^2$. The D mass for events that satisfy this requirement is shown in also in Figure A.5. We choose the signal region to be between 1.845 and 1.8825 GeV/c^2 with two
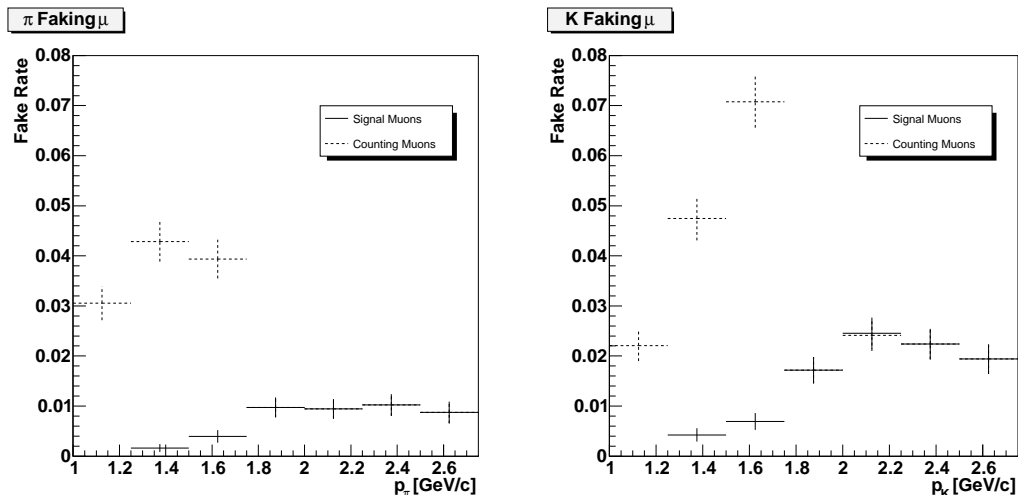
Figure A.6: The probability that a true pion (kaon) will fake a signal or counting muon is shown on the left (right).

sidebands: one between 1.80125 and 1.820 GeV/c^2 and one between 1.9075 and 1.92625 GeV/c^2.

Using the sign of the intermediate π we can obtain a pure sample of kaons and pions from the final D decay. We can then use the sample, with the appropriate sideband subtraction, to obtain the fake rate. We apply the $|\cos\theta| < 0.65$ cut to all tracks; therefore, we are determining the fake rate within this fiducial volume. Note that the signal muon ID requirements place additional constraints on DBCD and ZBCD. Because we are using the D^* decay chain we assume that these cuts have negligible impact on true kaons and pions from this decay, therefore they will not bias the signal fake rates. We also note that our reconstructed D^* decay is made up equally of both charge conjugations. Therefore, we present a fake rate that is averaged over charge. It has been noted in the past that the rate for K faking μ has some charge dependance [59]. We make no division in fake rates with regard to the hardware sparsification bug as the error bars on the resulting fake rates are equal to, if not larger than, the effect of the bug.

The fake rates for π and K faking μ are present in the table below and in Figure A.6. For tracks with momenta above 1.75 GeV/c the counting fake rate is *equivalent* to the signal fake rate.

p $[GeV/c]$	π faking $\mu[\%]$		K faking μ $[\%]$	
	Signal	Counting	Signal	Counting
1.00-1.25	0.0	3.05±0.34	0.0	2.20±0.31
1.25-1.50	0.16±0.08	4.29±0.40	0.42±0.13	4.75±0.44
1.50-1.75	0.39±0.12	3.93±0.38	0.69±0.16	7.07±0.52
1.75-2.00	0.97±0.20	–	1.72±0.26	–
2.00-2.25	0.94±0.20	–	2.45±0.31	–
2.25-2.50	1.02±0.22	–	2.24±0.31	–
2.50-2.75	0.87±0.22	–	1.94±0.30	–

A.2.2 p Faking μ

In order to obtain a pure sample of protons we reconstruct the decay $\Lambda \to \pi p$. We require that both of the tracks be `TrkmanApproved` and also we required that the vertex of the Λ be displaced from the beam spot by at least 10 cm. The reconstructed πp mass is shown in Figure A.7. We define the signal region from 1.1135 to 1.1185 GeV/c^2 and two sidebands: one from 1.105-1.1075 GeV/c^2 and one from 1.1245-1.127 GeV/c^2.

The displaced vertex requirement will bias tracks against our signal selection criteria. In addition, the efficiency for `TrkmanApproved` tracks drops as vertices become more displaced from the beam spot. Because of these systematic effects, and because low level of precision needed for the measured proton fake rate, we will assume fake rates for signal muons are the same as those for counting muons.
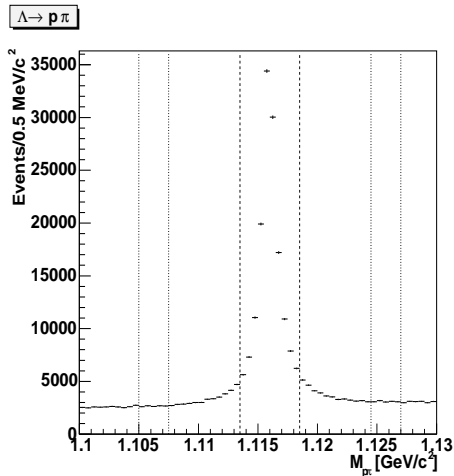
Figure A.7: reconstructed $p\pi$ invariant mass, the signal (sideband) region is indicated by dashed (dotted) lines
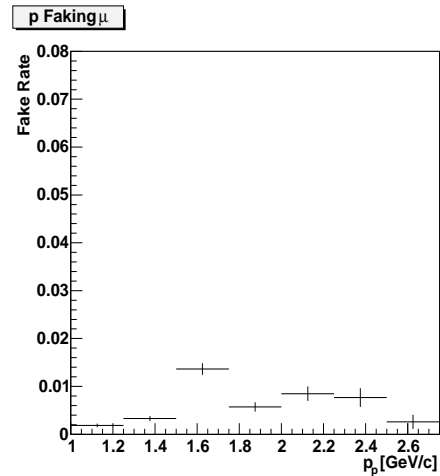
Figure A.8: probability that a proton will fake a counting muon – *error bars are statistical only*

As seen above, this is most likely a conservative over-estimate of the signal fake rate. In addition we will assign a conservative 100% error to the measured proton fake rates to account for systematic errors in the measurement. The fake rates are shown in the table below and in Figure A.8.

p [GeV/c]	p faking μ [%]
1.00-1.25	0.18
1.25-1.50	0.32
1.50-1.75	1.36
1.75-2.00	0.57
2.00-2.25	0.84
2.25-2.50	0.77
2.50-2.75	0.26

REFERENCES

[1] A. D. Sakharov, "Volation of CP invariance, C asymmetry, and baryon asymmetry of the universe," Pis. Zh. Eksp. Teor. Fiz **5**, 32-35 (1967).

[2] C. S. Wu, "Experimental Test of Parity Conservation in Beta Decay," Phys. Rev. **105**, 1413-1415 (1957).

[3] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, "Evidence for the 2π Decay of the $K(2)_0$ Meson," Phys. Rev. Lett. **13**, 138-140, (1964).

[4] N. Cabibbo, "Unitary Symmetry and Leptonic Decays," Phys. Rev. Lett. **10**, 531-533 (1963).

[5] M. Kobayashi and T. Maskawa, "CP-Violation in the Renormalizable Theory of Weak Interaction," Prog. Theo. Phys. **49**, 652-657 (1973).

[6] S. Eidelman *et al.*, "Review of Particle Physics", Phys. Let. B, **592**, 1 (2004).

[7] L. Wolfenstein, "Parametrization of the Kobayashi-Maskawa Matrix," Phys. Rev. Lett. **51**, 1945-1947 (1983).

[8] M. Bona *et al.*, UTFit: *http://utfit.roma1.infn.it*, current as of July 2005.

[9] F. J. Gilman and R. L. Singleton, Jr. , "Analysis of semileptonic decays of mesons containing heavy quarks," Phys. Rev. D **41** 142-150, (1990).

[10] M. Neubert, "Heavy Quark Symmetry," Phys. Rept. **245**: 259-396, (1994).

[11] D. Scora and N. Isgur, "Semileptonic Meson Decays in the Quark Model: An Update," Phys. Rev. D **52**, 2783 (1995).

[12] D. Melikhov and B. Stech, "Weak form-factors for heavy meson decays: an update," Phys. Rev. D **62**, 014006 (2000).

[13] T. Feldmann and P. Kroll, "Skewed parton distributions for $B \to \pi$ transitions," Eur. Phys. J. C **12**, 99 (2000).

[14] P. Ball and R. Zwicky, "New results on $B \to \pi, K, \eta$ decay form factors from light-cone sum rules," Phys. Rev. D **71**, 014015 (2005).

[15] P. Ball and R. Zwicky,"$B_{d,s} \to \rho, \omega, K^*, \phi$ decay form factors from light-cone sum rules revisited," Phys. Rev. D **71**, 014029 (2005).

[16] J. Shigemitsu *et al.* "Semileptonic B Decays with $N(F) = 2 + 1$ Dynamical Quarks," Lattice 2004, Batavia, Illinois (2004) [arXiv:hep-lat/0408019].

[17] L. Del Debbio *et al.* [UKQCD Collaboration], "Lattice constrained parametrizations of form-factors for semileptonic and rare radiative B decays," Phys. Lett. B **416**, 392 (1998).

[18] F. De Fazio and M. Neubert, "$B \to X_u \ell \bar{\nu}$ lepton decay distributions to order $\alpha(s)$," JHEP **9906**, 017 (1999).

[19] Y. Kubota *et al.*, "The CLEO II Detector," NIM A **320**, 66 (1992).

[20] D. G. Cassel *et al.*, "Design and Construction of the CLEO II Drift Chamber," NIM A **252**, 325 (1986).

[21] D. Peterson *et al.*, "The CLEO III Drift Chamber," NIM A **478**, 142 (2002).

[22] W. Ross, "The CLEO II.V Silicon Vertex Detector," NIM A **386**, 32 (1997).

[23] R. Kass *et al.*, "The CLEO III Silicon Vertex Detector," NIM A **501**, 32 (2003).

[24] R. Mountain *et al.*, "The CLEO III RICH Detector," NIM A **433**, 77 (1999).

[25] C. Bebek, "A Cesium Iodide Calorimeter with Photodiode Readout for CLEO II," NIM A **265** 258 (1988).

[26] S. Roberts, L. Gibbons, and E. Thorndike, "Trkman the Next Generation, The Quest for Track Quality," CLEO CBX 96-103.

[27] T. Meyer, *https://www.lns.cornell.edu/~cleo3/development/src/SplitoffProd/Doc.*

[28] S. Pappas, *https://www.lns.cornell.edu/~pappas/private/KinFit/FittingDoc.html.*

[29] C. H. Wang, "Electron Identification with CLEO II," CLEO CBX 91-52.

[30] C. S. Park, "CLEO III Rochester Electron Identification," CLEO CBX Draft v1.3, April 2003.

[31] R. A. Fisher, "The Use of Multiple Measurements in Taxonomic Problems," *Annals of Eugenics*, **7**:179, 1936.

[32] F. Würthwein and J. Alexander, "Details on Continuum Suppression as used in the Analysis of $B^0 \to K^+ \pi^-$ and $B^0 \to \pi^+ \pi^-$," CLEO CBX 93-41.

[33] G. Fox and S. Wolfram, "Observables for the Analysis of Event Shapes in $e^+ e^-$ Annihilation and Other Processes," Phys. Rev. Lett. **41**, 1581 (1978).

[34] See for example Chapter 1 of *An Introduction to Quantum Field Theory,* by M. Peskin and D. Schroeder.

[35] R. Wang, "Measurements of the Inclusive Semileptonic Branching Franction of B Mesons at the $\Upsilon(4S)$ resonance," Ph.D. Thesis, Univ. of Minnesota, (1994).

[36] V. Boisvert, "A Study of Exclusive Charmless Semileptonic B Decays with the CLEO Detector," Ph.D. Thesis, Cornell University, (2002).

[37] S. B. Athar et $al.$ [CLEO Collaboration], "Study of the q^2 dependence of $B \to \pi\ell\nu$ and $B \to \rho(\omega)\ell\nu$ decay and extraction of $|V_{ub}|$," Phys. Rev. D **68**, 072003 (2003).

[38] A. Warburton, private communication.

[39] L. Gibbons et $al.$ [CLEO Collaboration], "The inclusive decays $B \to DX$ and $B \to D^*X$," Phys. Rev. D **56**, 3783 (1997).

[40] C. S. Park, private communication.

[41] S. E. Csorna et $al.$ [CLEO Collaboration], "Moments of the B Meson Inclusive Semileptonic Decay Rate using Neutrino Reconstruction," Phys. Rev. D **70**, 032002 (2004).

[42] E. Lipeles, "A Study of the Fully Differential Inclusive Semileptonic B Meson Decay Rate," Ph.D. Thesis, California Institute of Technology, (2003).

[43] A. Lyon et $al.$ "Improvement of the Determination of V_{ub} using the $b \to s\gamma$ Photon Energy Spectrum," CLEO-CBX 01-59.

[44] Heavy Flavor Averaging Group, $http://www.slac.stanford.edu/xorg/hfag$, Dec. 2004.

[45] T. Meyer, "Limits on Weak Annihilation in Inclusive Charmless Semileptonic B Decays," Ph.D. Thesis, Cornell Univeristy, (2005).

[46] S. Chen et $al.$ [CLEO Collaboration], "Branching Fraction and Photon Energy Spectrum $b \to s\gamma$," Phys. Rev. Lett. **87**, 251807 (2001).

[47] B. Aubert et $al.$ [BaBar Collaboration], "Measurement of the Inclusive Electron Spectrum in Charmless Semileptonic B Decays Near the Kinematic Endpoint and Determination of $|V_{ub}|$, International Conference on High Energy Physics, Beijing, Chinia, (2004) [arXiv:hep-ex/0408075].

[48] R. Barlow and C. Beeston, "Fitting Using Finite Monte Carlo Samples," Comp.Phys.Comm. **77**, 219 (1993).

[49] B. Berger, "Tracking Efficiency Studies II: Recompress Results," CLEO CBX 00-32.

[50] R. Briere et $al.$, "Study of the Semileptonic Charm Decays: $D^0 \to \pi^-\ell^+\nu$ and $D^0 \to K^-\ell^+\nu$," CLEO CBX 04-12.

[51] F. Liu and Y. Gao, "Tracking Efficiency Systematics at CLEO III," CLEO CBX 02-13.

[52] A. Gritsan, "$\Gamma\left(\eta \to 3\pi\right)/\Gamma\left(\eta \to \gamma\gamma\right)$ and $\eta/\pi^0/\gamma$ Systematics," CLEO CBX 99-72.

[53] H. Muramatsu and T. Skwarnicki, "Inclusive Photon Analysis in $b\bar{b}$ System," CLEO CBX 04-24.

[54] A. Warburton, private communication.

[55] R. Briere, private communication.

[56] R. Briere *et al.*, "SSS Report on Hadron PID in CLEO III: Skims, Studies, and Systematics," CLEO CBX 03-37.

[57] D. Cassel *et al.*, "Determination of Tracking, K_S^0, π^0, and Particle Identification Efficiencies, CLEO CBX 05-07.

[58] J. Alexander *et al.*, "CLEO III Measurement of Rare Hadronic B Decays," CLEO CBX 02-37.

[59] A. Bellerive, "Fake Muons Study with DPTHMU > 3," CLEO CBX 97-59.