# A Study of the Fully Differential Inclusive Semileptonic $B$ Meson Decay Rate

Thesis by

Elliot Lipeles

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2004

(Submitted October 2003)

ii

# Acknowledgments

The completion of this thesis marks the end of my formal education. On the path to this point, I have benefited from the instruction, companionship, and example of many people. To start at my beginning, my parents have given me all the encouragement and opportunities I could need. It is certainly their love of learning that has given me my love of learning. Aaron, my brother, has been a companion in this experience and has lead the way into the world.

In my years spent in college, I benefited greatly from an early exposure to research under the tutelage of John Simpson. As a graduate student, my adviser Alan Weinstein showed me the beauty of careful and complete measurement. I have learned from him and the CLEO/Caltech post docs, Adi Bornheim, Steve Pappas, Jon Urheim, and Frank Würthwein, the techniques of high energy physics. The physics discussions I have had with them and with many other CLEO and Caltech scientists, too numerous to list, bring life to the science.

The research I have done for this thesis was possible only because of the contributions of many past and present CLEO collaboration and CESR staff members. I would especially like to thank my paper committee, Lawrence Gibbons, Ron Poling, Ed Thorndike, and honorary member Dan Cronin-Hennessy for their contributions to the analysis I am presenting here; Chris Bebek, Klaus Honscheid, Charlie Strohman, and Andreas Wolf who all taught me enormous amounts about electronics and programing; my companions in building the DAQ and commissioning the CLEO III detector, Valera Frolov, Tom Meyer, Todd Pedlar, and Greg and Jana Thayer; and my fellow CLEO/Caltech graduate students Anna Shapiro-Bridger and Werner Sun with whom I have passed hours doing problems sets and talking physics, politics, and philosophy.

Finally, I want to offer my gratitude to Maria Antonia Castro who has given me support and strength, and who has simply made my life larger.

# Abstract

We present a study of the fully differential inclusive semileptonic $B$ meson decay rate. Using a maximum likelihood fit, we extract the fractional contributions from the $B \to X_c l \nu$ processes with $X_c = D$, $D^*$, $D^{**}$, and nonresonant $X_c$, and the process $B \to X_u l \nu$. From the fit results, we extract moments of $B \to X_c l \overline{\nu}$ differential decay rate and the partial branching fraction of the $B \to X_u l \overline{\nu}$ decay in a restricted region of phase space. The region in which the $B \to X_u l \overline{\nu}$ partial branching fraction is measured is $M_X < 1.5$ GeV/$c^2$, $q^2 >11$ GeV$^2$/$c^4$. This measurements is used to extract CKM parameter $|V_{ub}| = (4.73 \pm 0.23 \pm 0.82 \pm 0.18 \pm 0.56 \pm 0.66) \times 10^{-3}$, where the uncertainties are due to statistics, detector systematics, $B \to X_c l \overline{\nu}$ model dependence, $B \to X_u l \overline{\nu}$ model dependence, and theoretical uncertainties. From the $\langle M_X^2 - \overline{M}_D^2 \rangle$ moment, the first moment of the photon energy spectrum in $B \to X_s \gamma$, and the semileptonic $B$ branching fraction, we extract the CKM parameter $|V_{cb}| = (4.12 \pm .10 \pm 0.09 \pm 0.16) \times 10^{-2}$, where the uncertainties are due to the measurement of the semileptonic $B$ decay rate, the moments measurements, and theoretical uncertainties. Both CKM parameter extractions use Heavy Quark Effective Theory (HQET) predictions for inclusive semileptonic $B$ decay. The measured moments are also used to test related predictions.

# Contents

# List of Figures

# List of Tables

# Chapter 1    Introduction

Particle physics is the study of the fundamental constituents of matter. Matter has been found to have several layers of structure: molecules, atoms, nuclei, and quarks and leptons. These constituents have been discovered through centuries of experimentation. The smallest constituents that have so far been observed are the quarks and leptons. In addition to these particles, forces have been discovered which govern the interactions between these particles: gravity, the electromagnetic force, the strong force, and the weak force. In this thesis, measurements of two of the fundamental properties of quarks are presented. These properties govern the interactions via the weak force of the bottom quark with the up and charmed quarks. These are studied in a process known as semileptonic $B$ meson decay.

The quarks and leptons and the forces that govern their known interactions, except for gravity, are described by the Standard Model of particle physics. This model has been enormously successful. It has been tested in a variety of ways over the past thirty years and is extraordinarily predictive and accurate. There are however experimental and conceptual reasons to expect the range of its applicability to be limited. Experimental particle physics is currently focused on testing the predictions of the Standard Model and searching for effects that it does not describe.

Two parameters of the Standard Model, known as $|V_{ub}|$ and $|V_{cb}|$, are measured in this thesis. Like the electromagnetic charge of the electron, these are fundamental properties of quarks. They are not predicted by the Standard Model and can only be determined by measurement. However, the Standard Model does predict a relationship between these two parameters and several other parameters. One of the purposes of the measurements presented is to contribute to the testing of this prediction. This test requires the combination of measurements like the ones presented and many others.

Also presented are measurements of the structure of $B$ mesons, which are particles composed of a bottom quark and an up or a down quark. The general name for particles composed of quarks is hadrons. The structure of hadrons is governed by interactions of the quarks via the strong force. Starting with nuclear-$\beta$ decay, semileptonic decays have been used to study the weak interactions of quarks and the structure of hadrons. The advantage of semileptonic decays is that the lepton part of the decay is not influenced by the strong force and therefore probes the internal structure of the hadron. Over the past century, semileptonic decay has been used to probe the structure of hadrons containing up, down, charmed, strange, and bottom quarks.

This chapter gives a brief overview of the Standard Model, focusing on the role played by the Cabibbo-Kobayashi-Maskawa matrix of which $|V_{ub}|$ and $|V_{cb}|$ are elements. Also in this chapter

is an introduction to $B$ meson decay and specifically semileptonic $B$ meson decay which is the physical process by which $V_{ub}$ and $V_{cb}$ are measured in this thesis. The effects of the strong force play an important role in $B$ meson decay, but these effects cannot be calculated reliably. This makes it difficult to relate the properties of $B$ meson decay measured in the lab to the Standard Model parameters $|V_{ub}|$ and $|V_{cb}|$. The particular choice of measurements in this thesis is designed to measure properties of the $B$ meson that can be related to these parameters using the recent theoretical progress in the calculation of inclusive heavy meson decay.

## 1.1   The Standard Model

The Standard Model describes the most basic constituents of matter. These constituents, elementary particles, are combined together into larger and larger objects to form the matter around us, which is referred to as bulk matter. As is widely known, bulk matter is composed of atoms, each of which is composed of electrons and a nucleus. This chain of substructure continues to the nucleus which contains protons and neutrons which are in turn made of quarks. The quarks and the electron are some of the elementary particles described by the Standard Model. At this time, these is no evidence to suggest that any of the particles of the Standard Model have substructure; they may be the last step in the chain. These particles are very far from everyday experience. In fact, there are many particles in the Standard Model whose existence only has an indirect effect on the properties of bulk matter and daily life. The particles and forces of the Standard Model are described below and are summarized in Table 1.1.

The Standard Model has a systematic structure, which is not understood. Parts of it are repeated three times, in what have been named *generations*. There are six quarks which can be grouped into the three generations. The quarks are called up (u) and down (d) in the first generation, strange (s) and charmed (c) in the second generation, and bottom (b) and top (t) in the third generation. The top and bottom quarks are sometimes also referred to as the truth and beauty quarks. The proton consists of two up quarks and a down quark bound together, although through the subtle effect of quantum mechanics all the generations play a role in its properties. Similarly the neutron is two down quarks and an up quark. The proton and the neutron are both members of a class of composite particles called baryons, all of which contain three quarks.

The quarks in the proton are held together by a force called the strong force, which is also responsible for holding together the protons and neutrons in atomic nuclei. The strong force is one of the three forces described by the Standard Model. The other two are the electromagnetic force and the weak force. The Standard Model does not describe gravity. One of the long term goals of particle physics is to learn how to combine the Standard Model with the theory of gravity called general relativity. The electromagnetic force plays an important role in the physics of bulk
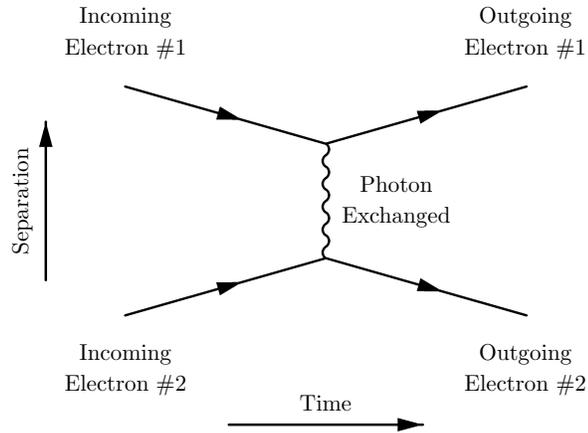
Figure 1.1: Diagram of the Photon as the Force Carrier of the Electromagnetic Force.

matter. It holds the electrons and nuclei of atoms together. It also holds the atoms in bulk matter together and also keeps one object from passing through another. The weak force does not play a large role in daily experience, but is important for the nuclear burning processes in the sun among other things. It is the only force which can convert one type of quark into another. The rate that the conversion takes place is governed by a set of numbers called the Cabibbo-Kobayashi-Maskawa (CKM) matrix. In the sun, protons are converted into neutrons, by changing one of the up quarks into a down quarks. The element of the CKM matrix important for that process is called $|V_{ud}|$. In this thesis two of these numbers are measured, $|V_{ub}|$ which governs the conversion of bottom quarks to up quarks, and $|V_{cb}|$ which governs the conversion of bottom quarks to charm quarks.

Another category of particles in the Standard Model is leptons. Unlike quarks, leptons are not affected by the strong force, they are only affected by the weak and the electromagnetic forces. There are also six leptons, which appear in three pairs, one for each generation. In each of these pairs is a charged lepton, which is affected by both the electromagnetic and weak forces, and a neutrino which is only affected by the weak force. Because the neutrinos are only affected by the weak force, they can pass through bulk matter, only interacting very rarely. There are neutrinos continually passing through the entire earth without being affected. The three pairs of leptons are the electron ($e$) and election neutrino ($\nu_e$), the muon ($\mu$) and muon neutrino ($\nu_\mu$), and the tau and ($\tau$) tau neutrino ($\nu_\tau$).

The forces can also be described in terms of particles which are said to be the mediators of that force. The photon is the mediator of the electromagnetic force. In this language when two electrons repel each other they exchange a photon. The photon carries energy and momentum from one of the of the electrons to the other giving it a kick (see Figure 1.1). Similarly the strong force is described by eight different particles called gluons and the weak force is described by three particles, the $W^+$, $W^-$, and $Z$ bosons. Each force also has an associated coupling, which indicates how often effects involving a the force occur. The strong force has the largest coupling, called $\alpha_s$. The coupling

Table 1.1: The Particles and Forces of the Standard Model.
The masses shown are from the Particle Data Group [1]. The uncertainties on the masses are only shown for the quarks where they are appreciable.

| | | EM Charge | Strong | Weak | Spin | Mass (GeV/$c^2$) |
|---|---|---|---|---|---|---|
| **Quarks:** | | | | | | |
| Up | $\begin{pmatrix} u \\ d \end{pmatrix}$ | $\frac{2}{3}$ | ● | ● | $\frac{1}{2}$ | 0.0015 to 0.0045 |
| Down | | $-\frac{1}{3}$ | ● | ● | $\frac{1}{2}$ | 0.005 to 0.0085 |
| Charm | $\begin{pmatrix} c \\ s \end{pmatrix}$ | $\frac{2}{3}$ | ● | ● | $\frac{1}{2}$ | 1.0 to 1.4 |
| Strange | | $-\frac{1}{3}$ | ● | ● | $\frac{1}{2}$ | 0.080 to 0.155 |
| Top | $\begin{pmatrix} t \\ b \end{pmatrix}$ | $\frac{2}{3}$ | ● | ● | $\frac{1}{2}$ | $174.3 \pm 5.1$ |
| Bottom | | $-\frac{1}{3}$ | ● | ● | $\frac{1}{2}$ | 4.0 to 4.5 |
| **Leptons:** | | | | | | |
| Electron | $\begin{pmatrix} e \\ \nu_e \end{pmatrix}$ | $-1$ | | ● | $\frac{1}{2}$ | 0.000510 |
| Electron Neutrino | | 0 | | ● | $\frac{1}{2}$ | $< 3 \times 10^{-9}$ |
| Muon | $\begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}$ | $-1$ | | ● | $\frac{1}{2}$ | 0.105 |
| Muon Neutrino | | 0 | | ● | $\frac{1}{2}$ | $< 3 \times 10^{-9}$ |
| Tau | $\begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}$ | $-1$ | | ● | $\frac{1}{2}$ | 1.777 |
| Tau Neutrino | | 0 | | ● | $\frac{1}{2}$ | $< 3 \times 10^{-9}$ |
| Higgs | H | 0 | | ● | 0 | $> 114.3$ |
| **Force Mediators:** | | | | | | |
| Electromagnetic | $\gamma$ | 0 | | | 1 | 0 |
| Strong | $g$ | 0 | ● | | 1 | 0 |
| Weak | $W^+$ | $+1$ | | ● | 1 | 80.423 |
| | $W^-$ | -1 | | ● | 1 | 80.423 |
| | $Z$ | 0 | | ● | 1 | 91.188 |
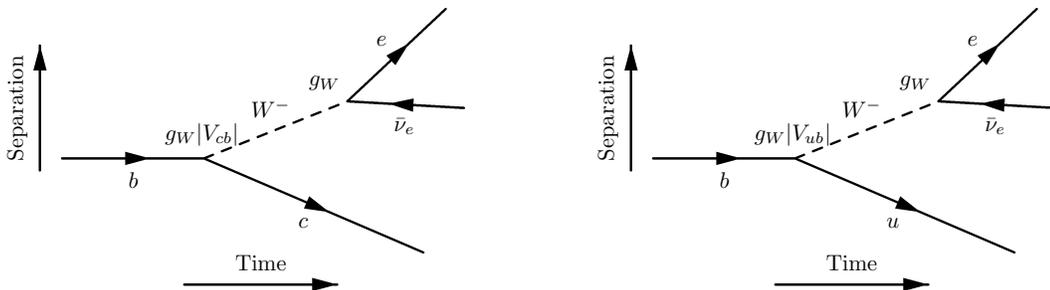
Figure 1.2: Semileptonic $b$ Quark Decay.
The strength of the coupling of the weak force $G_F$ discussed in the text is related to the $g_W$ shown in the diagrams by $G_F = \frac{\sqrt{2}g_W^2}{8m_W^2}$, where $m_W$ is the mass of the $W$ boson.

electromagnetic force, $\alpha_{EM}$, is a hundred times smaller, and the coupling of the weak force, $G_F$ is yet another thousand times smaller.

There is one more particle in the Standard Model, the Higgs. The Higgs field which is associated with the Higgs particle interacts with the quarks, leptons, and weak bosons to give them masses. The coupling of the the Higgs is the only thing that differentiates the three generations of quarks and leptons. The relationship between the Higgs and the CKM matrix is discussed in the next section. The Higgs is the only particle in the Standard Model which has not been observed experimentally.

Once particles have mass they can decay. Decay is a process in which the a parent particle ceases to exist and its mass is transfered either as mass or energy to two or more lighter daughter particles which are produced in the decay. Mass and energy are equivalent, as expressed by the famous equation $E = mc^2$. The mass of the original particle ends up in either the masses of the particles in the final state or in their relative motion (their kinetic energy). Particle decay is one of the most commonly studied processes in particle physics. Features of a decay can give information about the particles involved. The rate at which a particle decays is governed by the couplings involved. In particular, the rate at which the bottom quark decays into an up quark is governed by $|V_{ub}|$ and $G_F$. In this thesis, the rate that a bottom quark decays into an up quark, electron and an electron neutrino, which is written $b \rightarrow ue\nu_e$, is measured for the purpose of extracting $|V_{ub}|$. Similarly, the parameter $|V_{cb}|$ is measured in the decay $b \rightarrow ce\nu_e$. Both of these processes are displayed diagrammatically in Figure 1.2. Note that when the particle in the final state (on the right) is an up quark the coupling is $|V_{ub}|$ and when it is a charmed quark the coupling is $|V_{cb}|$.

The various particles of the Standard Model have properties("quantum numbers") which are conserved. The most familiar of such properties is the electromagnetic charge. The conservation of quantum numbers means that the sum of that number before and after any process is the same. There are some quantum numbers for which the product in the final state is equal to the product in the initial state. Because the weak force is the only force which can change one kind of quark

into another, each kind of quark has a quantum number associated with it which is conserved by electromagnetic and strong interactions, but not weak interactions. For instance for the strange quark the quantum number is strangeness and the strange quark has a strangeness of +1.

All the particles of the Standard Model have corresponding antiparticles or are their own antiparticles. Antiparticles have the same masses as the original particles and couple to the same forces as the original particle, but the signs of all the quantum numbers are reversed. For instance, the strange quark has strangeness +1 and electromagnetic charge $-1/3$, and the antistrange quark has strangness $-1$ and electromagnetic charge $+1/3$. Antiparticles are denoted with a line over them, so an antistrange quark is written $\bar{s}$.

The existence of antiquark states means that particle-antiparticle pairs can be produced by the electromagnetic and strong interactions. In the experiment discussed in this thesis, bottom quark pairs are produced via the electromagnetic interaction of an electron and its antiparticle the positron annihilating into a single photon which then produces a bottom quark and an antibottom quark (see Figure 1.3). In this process the initial state has a total electromagnetic charge of zero, because the electron and positron have opposite charges. Similarly the final state has zero electromagnetic charge, because the bottom quark and the antibottom quark have opposite charges. The energy of the collisions in this experiment is set so that the bottom quark pair produced cannot easily find each other to annihilate after the interaction and must therefore decay via the weak force which is the subject being studied. At low energies the quark and the antiquark can annihilate with each other via the strong or the electromagnetic forces.
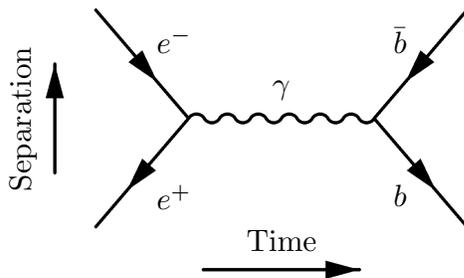


Figure 1.3: The Electromagnetic Production of $b\bar{b}$ Pairs.

The strong force is so strong that quarks can only be found in composite objects called hadrons. There are two classes for hadrons, baryons (previously mentioned) which contain three quarks and mesons which contain one quark and one antiquark. A meson is like an atom, but instead of being an electron and nucleus held together by the electromagnetic force, it is a quark and an antiquark held together by the strong force. If a quark in a hadron is hit with enough energy to knock it out of that hadron, there will be enough energy to create a new quark-antiquark pair, so that the struck quark will have a partner antiquark to make it a meson. The original hadron will be left with a
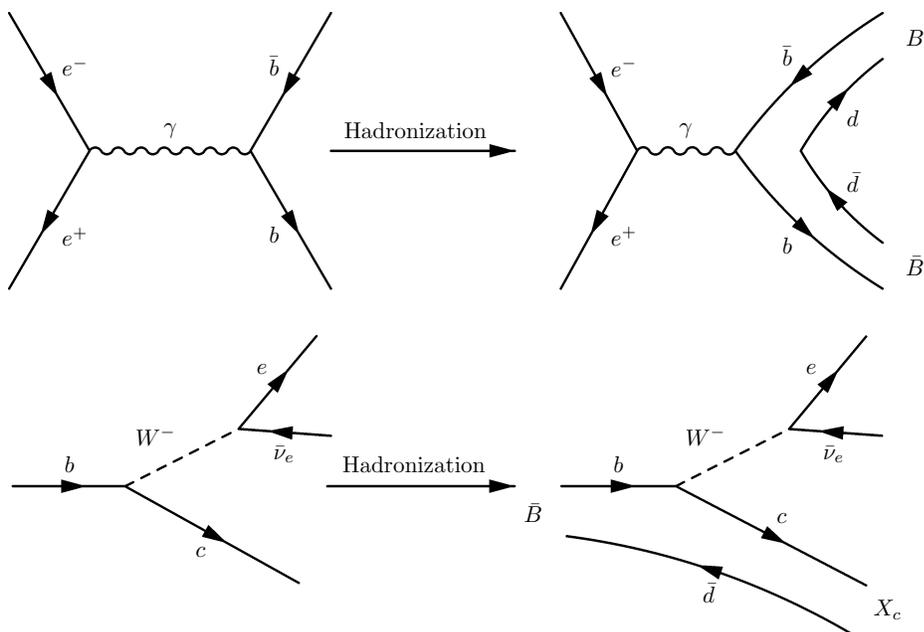
Figure 1.4: The Effects of Hadronization on the Processes $e^+e^- \to b\overline{b}$ and $b \to cl\nu$.

new quark to make it complete hadron again. This fact that quarks can only be found inside of hadrons is known as confinement and the process of creating new quarks or antiquarks to satisfy the constraints of confinement is called hadronization.

This makes it difficult, but not impossible, to relate the measured properties of hadrons containing quarks to the quark properties of interest. For this measurement, the bottom quarks are produced in the process $e^+e^- \to b\overline{b}$, but we observe $e^+e^- \to B\overline{B}$, where $B$ is the $B$ meson which contains a bottom quark and an up or down quark. The process of interest for measuring $|V_{cb}|$ is $b \to ce\overline{\nu}_e$, but what is observed is $B \to X_c e\overline{\nu}_e$, where $X_c$ is any set of hadrons containing the charmed quark. Similarly, $B \to X_u e\overline{\nu}_e$ is the result of the quark level process $b \to ue\nu_e$, where $X_u$ is any set of hadrons containing the up quark produced in the decay. The relationship between the quark level interaction of interest and the process involving hadrons which is observed is diagrammed in Figure 1.4. The relationship quark level and the hadron level is very difficult to calculate reliably. The focus of this thesis is to measure properties of the $B$ meson decay which can be related to quark level processes or can be used to refine the calculation of these processes.

The strength of the strong force, $\alpha_s$, varies with energy. At very large energies the strong force is only a little stronger than the weak and electromagnetic forces. In fact there is a hypothesis known as grand unified theory (GUT) that at energies $10^{13}$ times higher than the energies that have been probed in the laboratory, all three forces have equal strengths. As the energy of a process is reduced $\alpha_s$ increases and any calculation of the effects of the strong force becomes much more difficult. Below an energy called $\Lambda_{QCD} \approx 200$ MeV or approximately one quarter of the proton mass, the coupling $\alpha_s$

becomes large and calculations become impossible. The theory used to interpret the measurements presented in this analysis relies on the fact that the mass of the bottom and charmed quarks are considerably larger than $\Lambda_{QCD}$, which facilitates calculations of the properties of the decays.

## 1.2 The CKM Matrix

Without the Higgs, all the generations would be identical and there would be no transitions between the generations. The Higgs coupling not only gives the quarks and leptons mass, but the mass eigenstates are not equal to the weak eigenstates, allowing for transitions between generations. The mass basis has the advantage that the quark states can be identified experimentally by their masses. For instance, in this measurement the charm and up quark final states in $B$ mesons decay are separated by the energy released, which is related to the difference between the $B$ meson mass and the respective quark masses. Because the weak couplings of the three generations are all identical, linear combinations of the three weak eigenstates can be constructed so that the up type quarks ($u$, $c$ and $t$) are both weak and mass eigenstates. The weak eigenstates of the down type quarks are denoted $d'$, $s'$, and $b'$ and the mass eigenstates are denoted $d$, $s$, and $b$. The use of similar names for the weak and mass eigenstates is motivated by the experimental observation that the two bases are very similar. This does not have to be the case, and in fact in the lepton sector it is not the case.

These two bases are related by the Cabibbo-Kobayashi-Maskawa (CKM) matrix,

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.1}$$

This means that the weak eigenstates can be written in terms of the mass eigenstates as

$$|b'\rangle = V_{td}|d\rangle + V_{ts}|s\rangle + V_{tb}|b\rangle. \tag{1.2}$$

In the lepton sector there is an analogous matrix called Maki-Nakagawa-Sakata (MNS) matrix.

The weak interaction only produces transitions between the weak eigenstates of the same generation: ($u \leftrightarrow d'$), ($c \leftrightarrow s'$), and ($t \leftrightarrow b'$). Because the mass eigenstates are not the same as the weak eigenstates, transitions between the generations as defined by the masses does occur. Recall that the mass eigenstates are the states observed experimentally, because the quarks are identified by their masses. The probability for each transition is governed by the overlap of the weak eigenstate and the mass eigenstate. For instance, because $|V_{tb}|$ is much greater than $|V_{ts}|$ and $|V_{td}|$, the bottom quark mass eigenstate is mostly $b'$ and very little $s'$ and $d'$. The decays of the top quark to the bottom quark (emitting a $W^+$) are therefore much more common than decays of the top quark to

the strange and down quarks. So, the rate of the different quark decays is governed by the overlaps of the mass and weak eigenstates which is described by the CKM matrix.

Because the CKM matrix is a basis transformation, it must be unitary. This constraint reduces the number of free parameters in the CKM matrix. A completely general complex $3 \times 3$ matrix has 18 free parameters, the constraint that the matrix is unitary reduces this to nine free parameters. The phases of the quark fields in the Standard Model Lagrangian can be changed without changing the action and hence any observables. This can be exploited to remove another five free parameters from the CKM matrix (these are the five relative phases of the quark fields). This leaves a total of four free parameters. The Wolfenstein parameterization of the CKM matrix exploits the smallness of the off-diagonal elements to construct a representation in which the relationships between the elements are manifest:

$$
V_{\text{CKM}} = \begin{pmatrix} 1 - \frac{\lambda_c^2}{2} & \lambda_c & A\lambda_c^3(\rho - i\eta) \\ -\lambda_c & 1 - \frac{\lambda_c^2}{2} & A\lambda_c^2 \\ A\lambda_c^3(1 - \rho - i\eta) & -A\lambda_c^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda_c^4), \tag{1.3}
$$

where $\lambda_c \equiv |V_{us}| \approx .22$ is the Cabibbo angle. The use of $\lambda_c$ in the other off-diagonal terms is so that $A$, $\rho$, and $\eta$ are all order unity and has no physical motivation.

The unitarity constraint gives relationships between observables. In $B$ physics one commonly discussed relation is

$$
V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \tag{1.4}
$$

This term arises from multiplying the first column of the matrix and third row of the hermitian conjugate of the matrix. This relationship is of particular interest for testing flavor changing currents because the three numbers in the relation are all of similar scale, $\lambda_c^3$. When some of the terms are much larger than others, they must be measured with much higher fractional precision to be of use. A failure of such a relation would indicate either the presence of a fourth generation or that one of the measurements involved in the test has a contribution from a process that is not in the Standard Model. Equation 1.4 can be represented as a triangle in the complex plane. Each term of the equation is treated as a vector in the complex plane, the sum of which must be zero. Violations of unitarity would manifest themselves either as a failure of the triangle to close, failure of the sum of the angle to $180^o$, or a disagreement between the sides and the angles.

One of the main motivations measuring the parameters $|V_{ub}|$ and $|V_{cb}|$ is to contribute to testing the unitarity of the CKM matrix (Equation 1.4).

$$V_{ud}V_{ub}^* \qquad V_{td}V_{tb}^*$$

$$V_{cd}V_{cb}^*$$
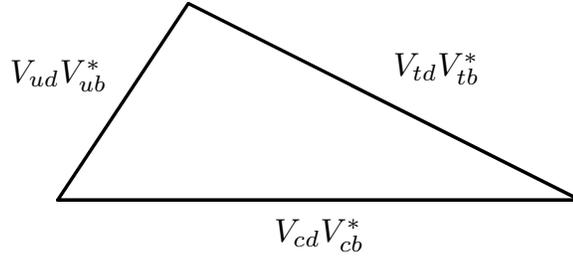
Figure 1.5: The CKM Unitarity Triangle for $B$ Physics.

## 1.2.1 Measuring CKM Parameters

Because of the nonperturbative nature of the strong force at low energy, it is often difficult to calculate the relationship between physical observables and the CKM parameters of interest; for example, relating the decay rate $\Gamma(B \to X_c l\overline{\nu})$ to the CKM parameter $|V_{cb}|$ requires an involved calculation and experimental input. To make the connection between quantities measured in the lab and the CKM parameters, a variety of symmetries and expansions are used. The symmetry calculations apply to exclusive decay modes of the hadrons, where the symmetry is used to relate the initial and final states. In some cases corrections due to violations of the symmetries can be calculated. The other class of calculations are of inclusive properties. These calculations rely on there being an energy in the process which is large compared to $\Lambda_{QCD}$. In this case the Operator Production Expansion (OPE) is used to separate the nonperturbative QCD physics at low energy from the perturbative QCD effects at high energy. Since the strength of the weak coupling is much smaller than the strengths of the strong and electromagnetic couplings, most of the measurements proceed by producing in a strong or electromagnetic process a pair of hadrons containing the quark of interest and the corresponding antiquark. The decay, which can only proceed through the weak interaction, is then studied. The exception to this is measurement of the parameters $|V_{cd}|$ and $|V_{cs}|$, where the calculations of the strong physics are limited. The variety of techniques used in the measurement of CKM parameters are diagrammed in Figure 1.6. Many of the measurement results are summarized in Table 1.2.

For transitions between the light quarks, the isospin and $SU(3)$ flavor symmetries are used. These symmetries arise from the fact that $m_u, m_d \ll \Lambda_{QCD}$ and $m_u, m_d, m_s \ll \Lambda_{QCD}$, respectively. In this limit, the effects of the strong force are much larger than the effects of the masses, so the properties of hadrons with different quark content can be related to each other. These symmetries allow the relationship between the structures of the incoming and outgoing hadrons to be calculated. The particular modes used are $n \to p^+ e^- \overline{\nu}_e$ for $|V_{ud}|$ and $K \to \pi e^- \overline{\nu}_e$ for $|V_{us}|$.

On the other end of the spectrum, the top quark is so heavy that the energy released its is much larger than $\Lambda_{QCD}$ and the strong effects can be treated perturbatively. In fact, it is a very good

Quark Decay:$|V_{ud}|,|V_{us}|,|V_{us}|,|V_{cb}|,|V_{ub}|,|V_{tb}|$          $W$ Decay:$|V_{cs}|$

Neutrino Scattering:$|V_{cd}|$          Mixing:$|V_{td}|,|V_{ts}|,\sin 2\beta$
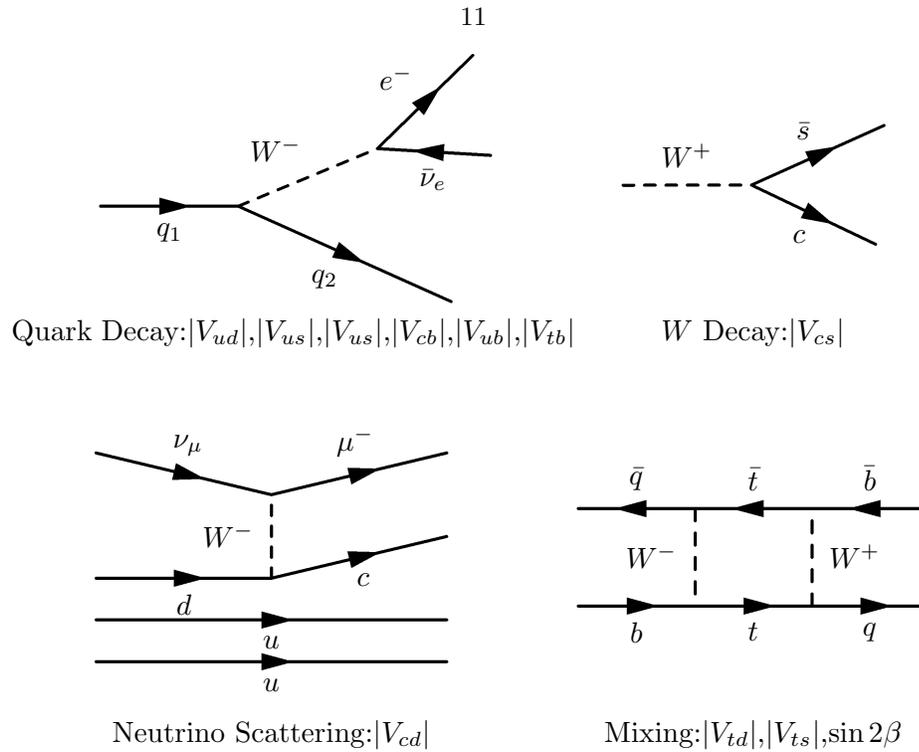
Figure 1.6: Processes Used to Measure CKM Parameters.

approximation to just neglect them. This technique allows the magnitude of $|V_{tb}|$ to be calculated, but is limited because of the difficulty in producing large numbers of top quarks. The other couplings to the top quarks $|V_{td}|$ and $|V_{ts}|$ are simply too small to be measured in top decay.

The masses of the $b$ and $c$ quarks are much larger than $\Lambda_{QCD}$, but not so much so that the nonperturbative effects of the strong force can be neglected. The large masses can be used in two different ways. For the case of $b$ to $c$ transitions, a symmetry between the $B$ and $D$ mesons occurs which allows $|V_{cb}|$ to be calculated. The second technique, which relies on the large energy release of the decays, is used in this analysis to extract both $|V_{cb}|$ and $|V_{ub}|$ from the $B \to X_c l \overline{\nu}$ and $B \to X_u l \overline{\nu}$ decay rates, respectively. The theoretical underpinnings of these measurements are described in the next chapter.

The CKM parameters $|V_{cd}|$ and $|V_{cs}|$ are particularly difficult to measure because the approximation that $m_c \gg \Lambda_{QCD}$ is not particularly good. The parameter $|V_{cd}|$ has been measured in the process $\nu_\mu p \to X_c \mu^+$, where a large energy transfer from the neutrino to the hadronic system allows the production process to be calculated. The parameter $|V_{cs}|$ has been measured in $W$ decay, where the $W$ mass provides a large energy.

$B$ physics can also be used to measure the parameters $|V_{td}|$ and $|V_{ts}|$ and their phases. These parameters appear in rate of $B^0 - \overline{B}^0$ and $B_s - \overline{B}_s$ mixing, respectively. Mixing is a process where a $B^0$ meson becomes a $\overline{B}^0$ meson or a $B_s$ meson becomes $\overline{B}_s$ meson through the exchange of two

$W$ bosons (see Figure 1.6). Unfortunately, there is no reliable way to calculate the hadronic physics involved in the overall rate of these processes. It is easier to extract $|V_{td}|/|V_{ts}|$ from measurements of both processes using the $SU(3)$ flavor symmetry discussed above.

Finally, there are several places in $B$ decay where the phase of a CKM matrix element can be extracted from the interference of two processes. In the case of mixing this can be particularly theoretically clean, only relying on the CP invariance of the strong force. In $B$ decays, calculation of these effects often uses isospin or $SU(3)$ flavor symmetry.

Table 1.2: Summary of CKM Parameter Measurements.
$\beta \equiv -\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}$. All of the measurements listed are from the Particle Data Group (PDG) [1]

| Element | PDG Value | Process | Symmetry/Expansion |
|---|---|---|---|
| $|V_{ud}|$ | $0.9734 \pm 0.0008$ | Nuclear Beta Decay | Isospin, $m_u, m_d \ll \Lambda_{QCD}$ |
| $|V_{us}|$ | $0.2196 \pm 0.0026$ | $K \to \pi e^- \overline{\nu}_e$ | $SU(3)$ flavor, $m_u, m_d, m_s \ll \Lambda_{QCD}$ |
| $|V_{cd}|$ | $0.224 \pm 0.0016$ | $\nu_\mu p \to X_c \mu^+$ | OPE, $(p_{\nu_\mu} - p_{\mu^-})^2 \gg \Lambda_{QCD}$ |
| $|V_{cs}|$ | $0.996 \pm 0.013$ | $W \to cs$ | OPE, $m_W \gg \Lambda_{QCD}$ |
| $|V_{cb}|$ | $0.0412 \pm 0.002$ | $B \to D^{(*)}l\overline{\nu}$ | Heavy Quark Spin-Flavor Symmetry, $m_b, m_c \gg \Lambda_{QCD}$ |
| | | $B \to X_c l\overline{\nu}$ | OPE, $m_b, m_c \gg \Lambda_{QCD}$ |
| $|V_{ub}|$ | $0.0036 \pm 0.0007$ | $B \to X_u l\overline{\nu}$ | OPE, $m_b \gg \Lambda_{QCD}$ |
| $|V_{tb}|$ | $0.94^{-0.31}_{-0.24}$ | $t \to bl\nu_l$ | OPE, $m_t \gg \Lambda_{QCD}$ |
| $|V_{td}|/|V_{td}|$ | - | $B^0$ Mixing/ $B_s$ Mixing | $SU(3)$ Flavor |
| $\sin 2\beta$ | $0.78 \pm 0.08$ | $B^0$ Mixing $+ B^0 \to \psi K_s$ | CP Symmetry of QCD |

## 1.3  $B$ Meson Decay

One of the primary goals of the study of $B$ mesons is to search for effects which are not described by the Standard Model, which are often referred to as *new physics*. The $B$ meson is a particularly good environment for this because the Standard Model contributions to bottom quark decay are small. Decays of the bottom quark to the top are kinematically not allowed, so all bottom quarks must either decay to the up and charmed quarks via the small couplings $|V_{ub}|$ and $|V_{cb}|$ or through virtual effect involving the top quark. New physics contributions may or may not be similarly suppressed.

In the Standard Model, decay of the $B$ meson can proceed via a very large variety of processes. These processes can be classified by their parton level interactions into semileptonic ($B \to X_c l\overline{\nu}$), hadronic with one charmed quark in the final state ($B \to X_c$), hadronic with two charmed quarks in the final state ($B \to X_{c\overline{c}}$), and rare processes. Semileptonic, hadronic, and some rare $B$ decay processes are diagrammed in Figure 1.7. Because $|V_{ub}|$ is ten times smaller than $|V_{cb}|$ and the decay rate is proportional to the CKM parameter squared, $b \to u$ processes only account for approximately 1% of $b$ quark decays and is considered rare. The other processes in the category of rare decays include the loop diagrams $b \to s\gamma$, $b \to sg$, $b \to sl^+l^-$, $b \to s\nu^+\overline{\nu}^-$, and these processes with the $s$

quark replaced by a $d$ quark.

The decay rate for the tree level processes is

$$\Gamma = \frac{G_F^2 m_b^5}{192\pi^3} \, |V_{\text{CKM}}|^2 \, N_c \, R \, \Phi, \qquad (1.5)$$

where

- $G_F$ is the weak coupling constant,

- $m_b$ is the $b$ quark mass,

- $|V_{\text{CKM}}|$ are the appropriate CKM matrix elements,

- $N_c$ is the number of colors for the final state (1 for semileptonic decays and 3 for the hadronic decays),

- $R$ is a correction due to QCD effects, and

- $\Phi$ is the available phase space for the decay.

For semileptonic decays, $|V_{\text{CKM}}|$ is $|V_{ub}|$ or $|V_{cb}|$. For hadronic decays of the form $b \rightarrow qq_1q_2$, the $|V_{\text{CKM}}|$ factor is $|V_{qb}||V_{q_1q_2}|$. In the limit that QCD corrections are negligible $R$ is one for the all the decays. The results of a recent calculation [2] of the values of $R$ are presented in Table 1.3 along with the other factors in the equation and the results of the complete calculation. The values of $m_b$ and $m_c$ can be calculated from the $B$ and $D$ meson masses using the formula presented in the next chapter (Equation 2.26). The formula above is accurate to order $(\Lambda_{QCD}/m_b)^2$. The moments results presented in this thesis can be used to improve both the calculation of the $b$ and $c$ quark masses and the $(\Lambda_{QCD}/m_b)^2$ corrections to the above formula.

The results for the total decay rate shown in Table 1.3 can be converted into a $B$ meson lifetime, $\tau_B = \hbar/\Gamma_B = 1.58$ ps. This compares reasonably to the $B^+$ and $B^0$ meson lifetimes of $1.653 \pm 0.028$ ps and $1.548 \pm 0.032$ ps, respectively [1]. The difference in the $B^+$ and $B^0$ lifetimes is due to terms of order $(\Lambda_{QCD}/m_b)^3$ which have been neglected in this calculation. Similarly the difference in the lifetimes between those mesons and the $B_s$ and $\Lambda_b$ hadrons occurs at order $(\Lambda_{QCD}/m_b)^2$.

From the entries in Table 1.3, the semileptonic branching fraction can also be calculated. It is simply the sum of the $b \rightarrow ce\nu_e$ and $b \rightarrow ue\nu_e$ branching fractions. The reason the semileptonic branching fraction is of particular interest is that electrons and muons have very distinctive signatures in the detector making it possible to measure the branching fraction with great precision. For the same reason the tau is often not included in the category of semileptonic decays although the tau is a lepton. For the remainder of this thesis, semileptonic will refer to only the electron and muon modes.
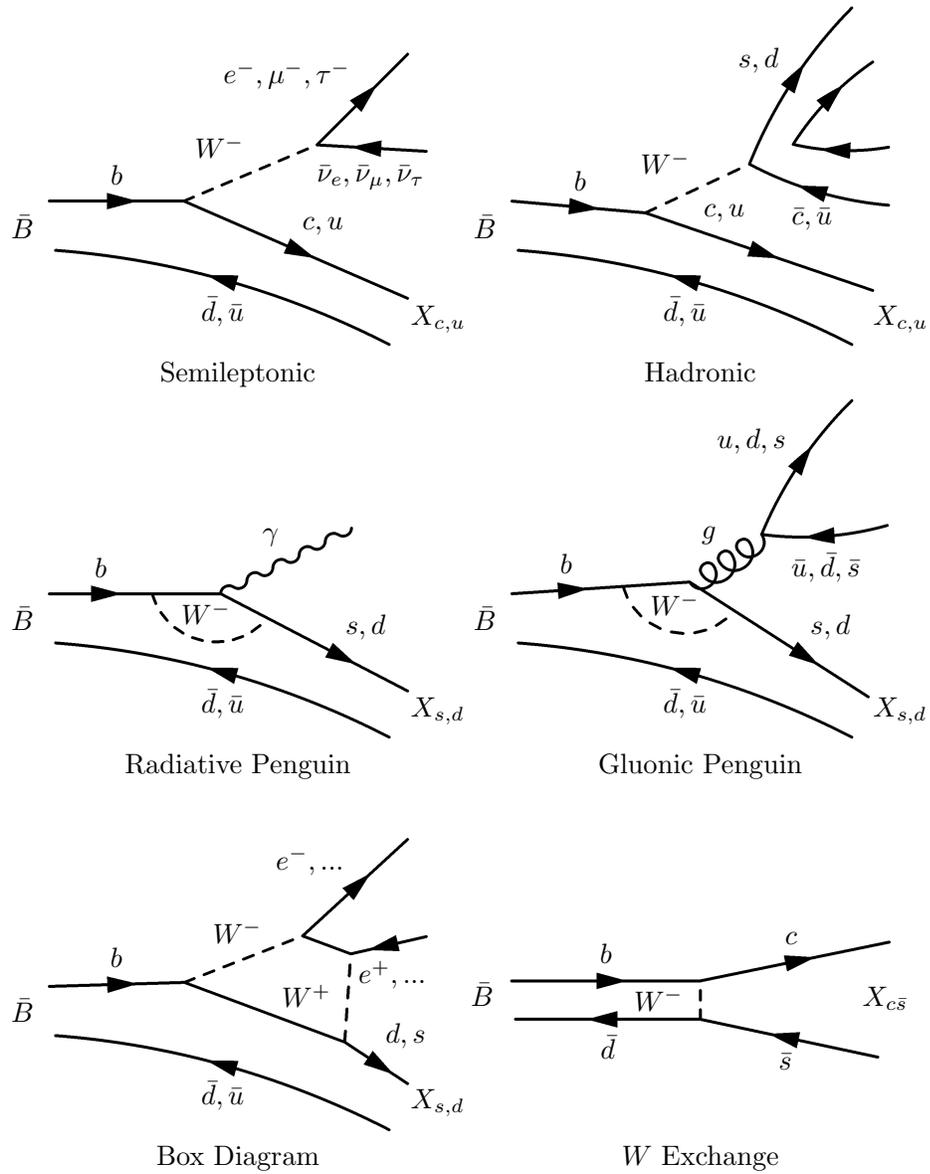
Figure 1.7:   Tree Level and Rare $B$ Decay Processes.

Table 1.3: Tree Level Contributions to $B$ Meson Decay.
For the CKM matrix elements the values $|V_{cb}|$ =0.041, $|V_{ub}|$ =0.0036, and $|V_{us}|$ =$|V_{cd}|$ =0.22 have been used. The radiative corrections are from reference [2].

| Mode | $N_c$ | $\Phi$ | $|V_{\mathrm{CKM}}|$ | $R$ | $\Gamma$ (GeV) | Branching Fraction |
|---|---|---|---|---|---|---|
| $c\overline{u}d$ | 3 | 0.52 | 0.041 | 1.17 | $1.79 \times 10^{-13}$ | 0.433 |
| $c\overline{u}s$ | 3 | 0.52 | 0.041×0.22 | 1.17 | $8.70 \times 10^{-15}$ | 0.021 |
| $c\overline{c}s$ | 3 | 0.25 | 0.041 | 1.44 | $1.06 \times 10^{-13}$ | 0.257 |
| $c\overline{c}d$ | 3 | 0.25 | 0.041×0.22 | 1.44 | $5.14 \times 10^{-15}$ | 0.012 |
| $ce\nu_e$ | 1 | 0.52 | 0.041 | 0.94 | $4.81 \times 10^{-14}$ | 0.116 |
| $c\mu\nu_\mu$ | 1 | 0.52 | 0.041 | 0.94 | $4.81 \times 10^{-14}$ | 0.116 |
| $c\tau\nu_\tau$ | 1 | 0.13 | 0.041 | 0.94 | $1.20 \times 10^{-14}$ | 0.029 |
| $u\overline{u}d$ | 3 | 1 | 0.0036 | 1.17 | $2.66 \times 10^{-15}$ | $6.43 \times 10^{-3}$ |
| $u\overline{u}s$ | 3 | 1 | 0.0036×0.22 | 1.17 | $1.29 \times 10^{-16}$ | $3.11 \times 10^{-4}$ |
| $u\overline{c}s$ | 3 | 0.52 | 0.0036 | 1.44 | $1.70 \times 10^{-15}$ | $4.12 \times 10^{-3}$ |
| $u\overline{c}d$ | 3 | 0.52 | 0.0036×0.22 | 1.44 | $8.25 \times 10^{-17}$ | $1.99 \times 10^{-4}$ |
| $ue\nu_e$ | 1 | 1 | 0.0036 | 0.94 | $7.13 \times 10^{-16}$ | $1.72 \times 10^{-3}$ |
| $u\mu\nu_\mu$ | 1 | 1 | 0.0036 | 0.94 | $7.13 \times 10^{-16}$ | $1.72 \times 10^{-3}$ |
| $u\tau\nu_\tau$ | 1 | 0.25 | 0.0036 | 0.94 | $1.78 \times 10^{-16}$ | $4.31 \times 10^{-4}$ |
| Sum of Tree Level Contributions | | | | | $4.14 \times 10^{-13}$ | $> 99\%$ |
| $2^{\mathrm{nd}}$ order process: $b \to s\gamma$, $b \to sg$, $b \to sl^+l^-$,... | | | | | | $\approx 10^{-3}$ |

In reference [2] the semileptonic branching fraction with the theoretical uncertainties assessed is

$$\mathcal{B}(B \to Xl\overline{\nu})^{\mathrm{theory}} = 12.0 \pm 1.6, \tag{1.6}$$

which is to be compared to the experimental value

$$\mathcal{B}(B \to Xl\overline{\nu})^{\mathrm{experiment}} = 10.87 \pm 0.18 \pm 0.30, \tag{1.7}$$

where the errors are statistical and systematic. The experimental value cited is from a recent measurement at BaBar [3]. It is similar to values measured at CLEO [4] and Belle [5], and a value measured in this analysis, all of which are listed in Table 6.5.

Another inclusive observable that can be derived from Table 1.3 is the number of charmed quarks expected per bottom quark decay,

$$
\begin{aligned}
n_c^{\mathrm{theory}} &= \frac{1}{\Gamma_B}\{\Gamma B \to X_c l\overline{\nu} + \Gamma(B \to X_c) + 2 \times \Gamma(B \to X_{c\overline{c}})\} & (1.8)\\
&= 1.24 \pm 0.05. & (1.9)
\end{aligned}
$$

The estimate of the theoretical uncertainty given above is also from reference [2]. This result is in moderate disagreement with the experimentally observed value [6] of

$$n_c^{\mathrm{experiment}} = 1.10 \pm 0.05. \tag{1.10}$$

The apparent disagreement is of marginal statistical significance. It is sometime referred to as the charm counting problem. The measurements of charm production in $B$ decay and the charm branching fractions needed to interpret them will likely improve the experimental measurement of $n_c$ in the near future. The uncertainties on the theoretical calculations can be improved by measurements of nonperturbative effects such as the moments measurements presented in this thesis. A true failure to agree would most likely indicate a misunderstanding of the theoretical uncertainties involved in the calculations. A more exciting possibility would be a new physics contribution causing the $B$ meson to decay to charmless final states more often than expected.

## 1.4   Semileptonic $B$ Meson Decay

Semileptonic $B$ decays are interesting because the matrix elements involved in the decay are more theoretically tractable than many other $B$ decay modes. This leads to two uses: the extraction of the CKM parameters $|V_{ub}|$ and $|V_{cb}|$, and the study of hadronic physics. Because the $B \rightarrow X_c l\overline{\nu}$ branching fraction is much larger than the $B \rightarrow X_u l\overline{\nu}$ branching fraction, the $B \rightarrow X_c l\overline{\nu}$ decays have been studied in much more detail. In this analysis, the shape of the inclusive semileptonic differential decay rate is analyzed by using a maximum likelihood fit to extract the contributions of the various exclusive modes. This is only possible because of the large theoretical and experimental knowledge of the dominant contributions.

In $B \rightarrow X_c l\overline{\nu}$, the exclusive contributions can be classified as $X_c=D$, $D^*$, $D^{**}$, and nonresonant $X_c$. Higher excited resonances have very small contributions and in the analysis presented they will be included in the nonresonant $X_c$ category, making the classification above complete. The $D$ and $D^*$ modes have been studied extensively in a variety of experiments. There are four $D^{**}$ mesons which can be classified as two narrow states, $D_1$ and $D_2^*$, and two broad states, $D_0^*$ and $D_1^*$ [7]. Only the narrow $D^{**}$ states have been observed experimentally. Furthermore, only one of the $D^{**}$ states, the $D_1$, has been observed in semileptonic $B$ decay. The measurement of these rates is made complicated by the small branching fractions and lack of clear signatures. This situation is only worse for the nonresonant $X_c$ final state which most likely consists of many different components each of which makes a very small contribution to the total decay rate. The nonresonant $X_c$ final state has not been observed unambiguously and it may never be possible to identify a large fraction of the exclusive contributions. The experimental knowledge about the exclusive modes and the final state mesons involved is summarized in Table 1.4.

The hadronic part of the final state in the $B \rightarrow X_u l\overline{\nu}$ process is also made up of a large number of exclusive final states including $X_u=\pi$, $\rho$, $\omega$, $\eta^{(\prime)}$, $a_1$, etc. The exclusive final states $\pi$, $\rho$, and $\eta$ have all been observed but together they account for less than one quarter of the total $B \rightarrow X_u l\overline{\nu}$ decay rate [8]. The partial decay rate of inclusive $B \rightarrow X_u l\overline{\nu}$ has been measured in the small endpoint

Table 1.4: Summary of the Knowledge of Exclusive Semileptonic $B$ Decays to Charm. The entries for the broad $D^{**}$ mesons are theoretical expectations as they have not yet been observed. The entries for the narrow $D^{**}$ mesons are for the neutral final state mesons. While the narrow charged $D^{**}$ have been observed with similar properties to the neutral mesons, semileptonic $B$ decay to the charged states has not been observed.

| | | Final State Meson Properties | | Semileptonic $B$ |
|---|---|---|---|---|
| | $X_c$ | Mass (MeV) | Width (MeV) | Branching Fraction |
| $D$ | | | | |
| | $D^+$ | $1869.3 \pm 0.5$ | $6.534 \pm 0.008 \times 10^{-10}$ | $2.11 \pm 0.17\%$ |
| | $D^0$ | $1864.1 \pm 0.5$ | $1.594 \pm 0.001 \times 10^{-9}$ | $2.15 \pm 0.22\%$ |
| $D^*$ | | | | |
| | $D^{*+}$ | $2006.7 \pm 0.5$ | $< 2.1$ | $4.6 \pm 21\%$ |
| | $D^{*0}$ | $2010.0 \pm 0.5$ | $0.096 \pm 0.022$ | $5.3 \pm 0.8\%$ |
| $D^{**}$ | | | | |
| | $D_0^*$ | $2460$ | $> 100$ | |
| | $D_1^*$ | $2460$ | $> 100$ | |
| | $D_1$ | $2422 \pm 2$ | $19 \pm 5$ | $5.6 \pm 1.6 \times 10^{-3}$ |
| | $D_2^*$ | $2459 \pm 2$ | $23 \pm 5$ | $< 8 \times 10^{-3}$ |
| Nonresonant $X_c$ | | No Measurements Available | | |

region where $B \to X_c l\overline{\nu}$ does not contribute [9]. This was the first evidence that $|V_{ub}|$ is not zero [10]. This region is also a very small part of the total rate.

In this analysis, we present measurements of the shape of the inclusive $B \to X_c l\overline{\nu}$ differential decay rate formulated as moments of the kinematical distributions, and a measurement of the partial decay rate of $B \to X_u l\overline{\nu}$ in a region of phase space where its contribution is not overwhelmed by the much larger $B \to X_c l\overline{\nu}$ contribution. The first of these measurements is used to extract nonperturbative properties of the $B$ meson, which are in turn used to improve the extraction of $|V_{cb}|$ from the measured semileptonic decay rate. These properties of the hadronic structure can also be used in calculations of fully hadronic $B$ meson decays, such as the ones presented in the previous section. The second measurement is used to extract $|V_{ub}|$.

The inclusive measurement technique presented here uses the approximate hermiticity of the detector and the well known initial state to infer the neutrino kinematics. All the available information from the charged lepton and the neutrino is used in a maximum likelihood fit to extract the composition of the data sample. Because the hadronic part of the final state is not reconstructed, a large number of complexities and limitations of exclusive reconstructions are avoided. The $X_c$ part of the results of the fit give a description of the composition that is by construction a good model of the shape of the total inclusive rate. While the exclusive branching fractions for $X_c = X_c = D$, $D^*$, $D^{**}$, and nonresonant $X_c$ are extracted with large systematic errors, the primary information provided is the shape of the inclusive spectrum summarized as moments of kinematical distributions. As previously mentioned, it would not be possible to measure the size and shape of the higher mass

contributions through exclusive reconstruction, because of their small size and the lack of clean experimental signatures with which to identify the modes.

All of the measurements presented can be compared directly to calculations of the properties of the inclusive processes $B \to X_u l \overline{\nu}$ and $B \to X_c l \overline{\nu}$. These calculations have advanced considerably in the past decade. Both the inclusive theory used to interpret the measurements and the exclusive theory needed to understand the individual modes are reviewed in the next chapter.

# Chapter 2    The Theory of Semileptonic $B$ Meson Decay

The coupling constants $|V_{ub}|$ and $|V_{cb}|$ appear in all tree level $B$ meson decays, which is the vast majority, but they remain difficult to measure because of the complexity of calculating the effects of the strong force. The strength of the QCD coupling, $\alpha_s$, makes the expansion in powers of the coupling constant, that is used in QED and weak calculations, impossible because the expansion does not converge.

At high energies perturbation theory is applicable, but $\alpha_s$ varies with the energy of the interaction. It becomes larger and larger as the energy is reduced, until it is infinite and perturbation theory is no longer applicable. To one loop order, $\alpha_s$ at an energy $\mu$ can be written as

$$\alpha_s(\mu) = \frac{1}{\beta_0 \ln(\mu^2/\Lambda_{QCD}^2)}, \tag{2.1}$$

where $\beta_0 = (33 - 2N_q)/12\pi$ is the one loop QCD $\beta$ function, $N_q$ is the number of light quarks, and $\Lambda_{QCD} \approx 200$ MeV is the point at which the one loop calculation of $\alpha_s$ becomes infinite. The parameter $\Lambda_{QCD}$ is often used as nominal scale of nonperturbative QCD effects. It is not actually the energy at which $\alpha_s$ becomes infinite, because this calculation becomes invalid long before that point.

The first step toward reducing the complexity of the calculation that is needed to relate measured quantities to $|V_{ub}|$ and $|V_{cb}|$ is to focus on semileptonic decays, where the calculation is simpler. Semileptonic decays have parton level processes $b \to ul\nu$ and $b \to cl\nu$. Because of confinement, what is actually observed are the processes $B \to X_u l\overline{\nu}$ and $B \to X_c l\overline{\nu}$, where $X_u$ and $X_c$ are any set of hadrons containing the $u$ and $c$ quarks from their respective parton level processes. Because the charged lepton and the neutrino in these processes do not interact strongly, the calculation of the matrix element can be simplified significantly. The matrix element for semileptonic decay can be factorized into two separate currents, one for the hadronic initial and final states and one for the leptonic system in the final state. This is written formally as

$$\begin{aligned} |\mathcal{M}|^2 &= |\langle X_q l\overline{\nu}|H_{\text{Weak}}|\overline{B}\rangle|^2 \\ &= |\langle X_q|J_\mu^{\text{had}}|\overline{B}\rangle\langle l\overline{\nu}|J_{\text{lep}}^\mu|0\rangle|^2 + \mathcal{O}(\alpha_{EM}, G_F), \end{aligned}$$

where $H_{\text{Weak}}$ is the weak Hamiltonian, and $J^{\text{had}}$ and $J_{\text{lep}}^\mu$ are hadronic and leptonic currents. These are describe in more detail in Section 2.2. As indicated in the equation, factorization is an approximation which neglects corrections of order the electromagnetic coupling $\alpha_{EM}$ and the weak coupling
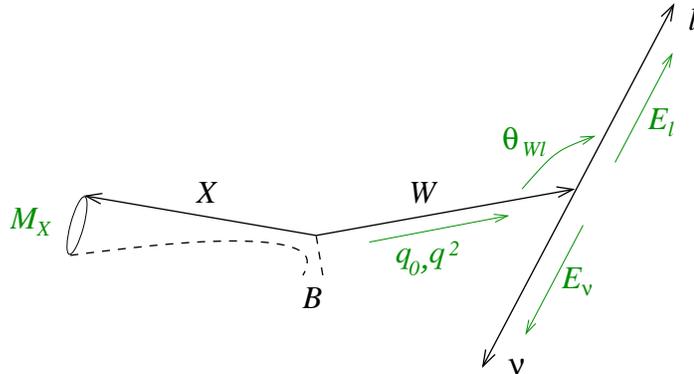
Figure 2.1: Graphical View the $B \rightarrow X l \overline{\nu}$ Kinematics.

$G_F$. Radiative decays such as $B \rightarrow X_s \gamma$ can also be factorized in a similar fashion. The leptonic current can be calculated very accurately, because strong interactions don't contribute, except by loop effects which are suppressed by two powers of $\alpha_{EM}$. The standard model predictions for the leptonic currents have been tested to very high precision in muon and tau decay [11]. The hadronic current, however, is very difficult to calculate.

In recent years, Heavy Quark Effective Theory (HQET) [7] has allowed a new level of precision in the calculation of the hadronic current. HQET is an effective theory of mesons which makes use of the fact that the bottom and charm quark masses are much larger than $\Lambda_{QCD}$. HQET predicts many properties of heavy meson spectroscopy, exclusive heavy to heavy decays rates, and inclusive heavy meson decay rates.

The measurements presented in this thesis make use of the HQET predictions of the $B$ meson decay rates to calculate $|V_{ub}|$ and $|V_{cb}|$. The exclusive theory is also used in the analysis because it is necessary to model the exclusive components of the inclusive decays in detail in order to understand the detector response and make the best use of the available data. This chapter will review heavy quark effective theory, the Lorentz structure of inclusive semileptonic decays, the HQET predictions for inclusive semileptonic decays, and the HQET predictions for exclusive decays. Finally, there is a section discussing radiative correction to these decays. In the following discussion, the heavy quarks will be referred to as $Q$, while quarks referred to as $q$ may be either heavy or light.

## 2.1    Kinematics of Semileptonic $B$ Decay

It is useful to treat the final state of the $B \rightarrow X l \overline{\nu}$ decay as a three-body final state where the hadronic system has a variable mass. In the $B$ meson rest frame, three independent variables are needed to fully describe the kinematics. There are several possible choices for those variables. Some of the possible variables are more relevant to detector effects and others are more directly related to

dynamics. The available variables, diagrammed in Figure 2.1, are

1. $E_\ell$, the energy of the charged lepton,

2. $E_\nu$, the energy of the neutrino,

3. $q^2$, the invariant mass of the charge lepton-neutrino pair, i.e., the mass of the virtual $W$,

4. $q_0$, the sum of the energies of the charged lepton and the neutrino, i.e., the energy of the virtual $W$,

5. $M_X$, the invariant mass of the recoiling hadronic system,

6. $\cos\theta_{W\ell}$, the helicity angle of the $W$ decay, defined as the angle between charged lepton in the $W$ rest frame and the $W$ in the $B$ frame.

Both $q^2$ and $M_X$ can be expressed simply in terms of the lepton kinematics in the $B$ rest frame. For the $\cos\theta_{W\ell}$ a longer calculation leads to

$$\cos\theta_{W\ell} \;=\; \frac{1}{\sqrt{q_0{}^2 - q^2}}\left[ E_\ell - \left(\frac{1 + \frac{m_\ell^2}{q^2}}{1 - \frac{m_\ell^2}{q^2}}\right) E_\nu \right] \approx \frac{E_\ell - E_\nu}{\sqrt{q_0{}^2 - q^2}}. \tag{2.2}$$

Equation 2.2 shows that the $\cos\theta_{W\ell}$ variable is related to the split of the energy between the charged lepton and the neutrino. For a moving charged lepton and a stopped neutrino, $\cos\theta_{W\ell}$ is 1, and for the converse $\cos\theta_{W\ell}$ is -1.

The $B$ mesons produced by the accelerator are very low momentum, but they are not at rest in the lab-frame. In the lab-frame, there are three additional kinematic variables corresponding to the flight direction and momentum of the $B$ meson. The reconstruction technique used in this analysis does not measure the variables related to the $B$ momentum direction. This leads to a degraded resolution on the reconstructed quantities, except for $q^2$ which is a Lorentz invariant that does not directly involve the parent direction. The resolution on the lepton energy and direction in the lab frame is much better than the resolution on the neutrino energy and direction. Because neutrino kinematics are measured with poor resolution, any variable derived from them will also have poor resolution (i.e., $q^2$, $M_X^2$, ...). Furthermore, only leptons with momentum in the lab frame of above 1 GeV/$c$ are used. A detailed Monte Carlo simulation is used to model the resolution and efficiency of the detector, as described in Section 3.4.

## 2.2 Lorentz Structure of Inclusive Semileptonic Decay

### 2.2.1 Phase Space

The differential decay rate for the $B \to Xl\overline{\nu}$ process is described in terms of a set of form factors which represent the most general hadronic current that is Lorentz invariant. In this section, the phase space for inclusive semileptonic $B$ decay is calculated and the relevant form-factors are identified. In both steps, the final state leptons are taken to be massless. This is a very good approximation, because the lepton energies are much larger than the mass. For the muon, heaviest lepton considered, corrections would be of order $(m_l/E_l)^2 \sim 0.01$. For decays involving the tau lepton, the mass should not be ignored, but the calculation would otherwise proceed in a similar manner.

The differential decay rate for a generic decay is

$$d\Gamma = \frac{(2\pi)^4}{2M}|\mathcal{M}|^2 d\Phi, \tag{2.3}$$

where $M$ is the mass of the parent, $\mathcal{M}$ is the matrix element of the decay, and $d\Phi$ is the phase space. The n-body phase space describes the density of possible final states for the n particles of the decay:

$$d\Phi(P; p_1, p_2, ...) = \delta^4(P - (\sum_i p_i)) \prod_j \frac{d^3 p_j}{(2\pi)^3 2E_j}. \tag{2.4}$$

In inclusive decays, final states with different numbers of particles are summed together. The number of particles in the final state is therefore not well defined. When calculating the inclusive $B \to Xl\overline{\nu}$ decay rate, the sum over the different hadronic final states and their phase space is performed as part of the calculation of the matrix element. The partial differential decay rate is

$$d\Gamma = \frac{d^3 p_l}{(2\pi)^3 2E_l} \frac{d^3 p_\nu}{(2\pi)^3 2E_\nu} \sum_X \sum_{\text{lepton spins}} \frac{(2\pi)^4}{2M_B}|\mathcal{M}|^2 \delta^4(p_B - p_l - p_{\nu_l} - p_X), \tag{2.5}$$

where $\sum_X$ indicates both the sum over the hadronic final states and the integral over the phase space of those states. Proceeding in the limit where the lepton masses are negligible, the differential elements can be written $d^3 p = E^2 dE \, dcos\theta \, d\phi$. Three of the four angles are just the orientation of the decay (two for the direction of the lepton and one rotation around the lepton axis). Because the $B$ meson is spinless, the orientation of the decay is isotropic in the $B$ rest frame and contains no information about the underlying physics. Integrating over the orientation gives a factor of $2(2\pi)^2$. The remaining angle is the angle between the lepton and neutrino directions, $cos\theta_{l\nu}$. After the integration, the partial differential decay rate is

$$d\Gamma = \sum_X \sum_{\text{lepton spins}} \frac{|\mathcal{M}|^2}{4M_B} E_\ell dE_\ell \, E_\nu dE_\nu \, dcos\theta_{l\nu} \, \delta^4(p_B - p_l - p_{\nu_l} - p_X). \tag{2.6}$$

Performing a final change of variables, using $q^2 = 2E_\ell E_\nu (1 - cos\theta_{l\nu})$, the partial differential decay rate in terms of the observables $E_\ell$, $E_\nu$, and $q^2$ is

$$\frac{d\Gamma}{dE_\ell dE_\nu dq^2} = \sum_X \sum_{\text{lepton spins}} \frac{|\mathcal{M}|^2}{8M_B} \delta^4(p_B - p_l - p_{\nu_l} - p_X). \tag{2.7}$$

## 2.2.2 Factorization

The matrix element for the $B \to Xl\bar{\nu}$ process can be written in terms of the low energy effective Hamiltonian for the weak process $b \to xl^-\bar{\nu}_l$ where $x = u$ or $c$:

$$H_W = \frac{G_F V_{xb}}{\sqrt{2}} J_{\text{lep}}^\mu J_\mu^{\text{had}}, \tag{2.8}$$

where $J_{\text{lep}}^\mu = \bar{l}\gamma^\mu(1 - \gamma^5)\nu_l$ and $J_\mu^{\text{had}} = \bar{x}\gamma_\mu(1 - \gamma^5)b$ are the leptonic and hadronic currents, respectively, and $V_{xb}$ is the CKM matrix element involved in the decay. Inserting the Hamiltonian and using the factorization approximation, the differential decay rate becomes

$$
\begin{aligned}
\frac{d\Gamma}{dE_\ell dE_\nu dq^2} &= \sum_X \sum_{\text{lepton spins}} \frac{|\langle Xl\bar{\nu}|H_W|\overline{B}\rangle|^2}{8M_B} \delta^4(p_B - p_l - p_{\nu_l} - p_X) \\
&= \sum_X \sum_{\text{lepton spins}} \frac{G_F^2 |V_{xb}|^2}{2} \frac{|\langle l\bar{\nu}|J_{\text{lep}}^\mu|0\rangle\langle X|J_\mu^{\text{had}}|\overline{B}\rangle|^2}{8M_B} \delta^4(p_B - p_l - p_{\nu_l} - p_X) \\
&= \frac{G_F^2 |V_{xb}|^2}{2\pi^3} L_{\alpha\beta} W^{\alpha\beta},
\end{aligned}
\tag{2.9}
$$

where $L_{\alpha\beta}$ and $W^{\alpha\beta}$ are called the leptonic and hadronic tensors, respectively.

In the Standard Model, the leptonic tensor is

$$
\begin{aligned}
L^{\alpha\beta} &= \sum_{\text{lepton spins}} \langle 0|J_{\text{lep}}^{\dagger\alpha}|l\bar{\nu}_l\rangle\langle l\bar{\nu}_l|J_{\text{lep}}^\beta|0\rangle \\
&= 2\left(p_l^\alpha p_{\nu_l}^\beta + p_l^\beta p_{\nu_l}^\alpha - g^{\alpha\beta}p_l \cdot p_{\nu_l} - i\epsilon^{\eta\beta\lambda\alpha}p_{l\eta}p_{\nu_l\lambda}\right),
\end{aligned}
\tag{2.10}
$$

where corrections occur at order $\alpha_{EM}$ and $G_F$ (the same order as the corrections to the factorization approximation) [12].

The hadronic tensor is defined as

$$W_{\alpha\beta} = \sum_X \frac{1}{2M_B}(2\pi)^3 \langle \overline{B}|J_\alpha^{\text{had}\dagger}|X\rangle\langle X|J_\beta^{\text{had}}|\bar{B}\rangle \, \delta^4(p_B - q - p_X), \tag{2.11}$$

where $q = p_l + p_{\nu_l}$. This current involves nonperturbative physics, so it cannot be calculated with an expansion in the coupling constant the same way the leptonic current was. Instead, HQET can be applied to expand this current in powers of $\frac{\Lambda_{QCD}}{M_B}$.

## 2.2.3   The Hadronic Tensor

The Lorentz structure of the hadronic current is constrained because it must be composed using only $v^\alpha$, the four-velocity of the parent $B$, and $p_X^\alpha$, the four-momentum of the hadronic system. Because of the conservation of energy and momentum expressed by the delta function in the $W^{\alpha\beta}$ definition, the momentum of the leptonic system $q$ can be used in place of $p_X^\alpha$. The most general tensor that can be composed of two available vectors is

$$
\begin{aligned}
W_{\alpha\beta} = {} & - g_{\alpha\beta}\, W_1(q^2, v\cdot q) + v_\alpha v_\beta\, W_2(q^2, v\cdot q) \\
& - i\epsilon_{\eta\beta\lambda\alpha} q^\eta v^\lambda\, W_3(q^2, v\cdot q) + q_\alpha q_\beta\, W_4(q^2, v\cdot q) \\
& + (v_\alpha q_\beta + q_\alpha v_\beta)\, W_5(q^2, v\cdot q) + (v_\alpha q_\beta - q_\alpha v_\beta)\, W_6(q^2, v\cdot q).
\end{aligned}
\tag{2.12}
$$

The form factors $W_i$ are functions of $q^2$ and $q_0 \equiv v\cdot q$ or equivalently $q^2$ and $M_X^2$.[1] In the limit of massless leptons, $q_\alpha L^{\alpha\beta}$ is zero, so the $W_4$, $W_5$, and $W_6$ terms do not contribute. This limit is used in everything that follows. The resulting inclusive semileptonic decay rate in terms of the unknown form factors $W_1$, $W_2$, and $W_3$ is

$$
\begin{aligned}
\frac{d\Gamma}{dE_\ell dE_\nu dq^2} = \frac{G_F^2 |V_{xb}|^2}{2\pi^3} \Big( & W_1(q^2, q_0)\, q^2 \\
& + W_2(q^2, q_0)\, (2E_l E_\nu - q^2/2) \\
& + W_3(q^2, q_0)\, q^2 (E_l - E_\nu) \Big).
\end{aligned}
\tag{2.13}
$$

Expressing the partial differential decay in terms of the variables $q^2$, $q_0$, and $\cos\theta_{W\ell}$ reveals the relationship between the $W_i$ functions and the helicity structure of the virtual $W$:

$$
\begin{aligned}
\frac{d\Gamma}{dq^2\, dq_0\, d\cos\theta_{W\ell}} = \frac{G_F^2 |V_{xb}|^2}{4\pi^3} \Big( & W_1(q^2, q_0)\, q^2 (q_0^2 - q^2)^{1/2} \\
& + W_2(q^2, q_0)\, (q_0^2 - q^2)^{3/2} \frac{1}{2}(1 - \cos^2\theta_{Wl}) \\
& + W_3(q^2, q_0)\, q^2 (q_0^2 - q^2)\, \cos\theta_{W\ell} \Big).
\end{aligned}
\tag{2.14}
$$

This derivation, which uses only the Lorentz structure of the decay and the approximations of factorization and negligible lepton mass, shows that the structure of the differential decay rate distribution in the $\cos\theta_{W\ell}$ variable is strongly constrained. In the context of HQET, the $W_i$ functions can be calculated using an operator product expansion. These calculations allow one to extract $|V_{ub}|$ and $|V_{cb}|$ from measurements of the partial decay rates of $B \to X_u l\bar\nu$ and $B \to X_c l\bar\nu$, respectively.

---

[1] In the literature, $W_6$ is often not considered because it violates time reversal invariance.

## 2.3 Heavy Quark Effective Theory

Strong interactions with heavy quarks behave very differently from strong interactions with light quarks. The difference arises from the fact that strong interactions are only strong at the energy scale near $\Lambda_{QCD}$ or below, which is much lower energy than the mass of the heavy quarks. In the case of charmed and bottom quarks, the masses are $\sim 1.5$ GeV/$c^2$ and $\sim 4.9$ GeV/$c^2$, respectively, and $\Lambda_{QCD}$ is $\sim 200$ MeV. While strong interactions can easily create a pair of quarks for which the mass of the quark is much less than the QCD scale, $m_q \ll \Lambda_{QCD}$, creating a pair of heavy quarks, $m_Q \gg \Lambda_{QCD}$, is highly suppressed. The effect of strong interactions on the momentum of a quark can only be of order $\Lambda_{QCD}$. Because the velocity of a particle is the momentum divided by the mass, the effect of strong interactions on a heavy quark's velocity is suppressed by $\Lambda_{QCD}/m_Q$. The velocity of a heavy quark in the infinite mass limit cannot be affected by the strong force and is therefore conserved. Because the velocity does not depend on the heavy quark mass, different flavor heavy quarks interact identically in the heavy quark limit. This is known as *heavy flavor symmetry*. The coupling of the strong force to the spin of the quark is via the chromomagnetic moment of the quark. Because the gyromagnetic ratio has the quark mass in the denominator, $\gamma = g_s \hbar / 2 m_Q$, the heavy quark spin also decouples from the strong interaction. The decoupling of the spin in the heavy quark limit leads to the *heavy quark spin symmetry*. These two symmetries have important consequences, especially for the exclusive decays of bottom hadrons to charmed hadrons, where the symmetries relate the initial and final states. These symmetries are only true in the heavy quark limit and are violated at order $\Lambda_{QCD}/m_Q$.

These observations can be formalized by writing the Standard Model Lagrangian as an expansion in $1/m_Q$. The derivation presented here follows the pedagogical text by Manohar and Wise [7]. The expansion is constructed by rewriting the Lagrangian in terms a of a set of fields $Q_v(x)$ which describe heavy quarks moving with velocity $v$ and a set of fields $\mathfrak{Q}_v(x)$ which describe the heavy antiquark excitations that we want to integrate out. These two sets of fields can be written in terms of the original QCD quark field $Q(x)$ as $Q_v(x) = e^{im_Q v \cdot x} \frac{1+\slashed{v}}{2} Q(x)$ and $\mathfrak{Q}_v(x) = e^{im_Q v \cdot x} \frac{1-\slashed{v}}{2} \mathfrak{Q}(x)$. Conversely, $Q(x) = e^{-im_Q v \cdot x}(Q_v(x) + \mathfrak{Q}_v(x))$. The $\frac{1\pm\slashed{v}}{2}$ terms project out the particle and antiparticle components respectively. The $Q_v(x)$ fields therefore have the property $\frac{1+\slashed{v}}{2}Q_v(x) = Q_v(x)$ and $\frac{1-\slashed{v}}{2}Q_v(x) = 0$. The reverse properties hold for the $\mathfrak{Q}_v(x)$ fields, which means that the $Q_v(x)$ and $\mathfrak{Q}_v(x)$ fields are orthogonal, $\bar{Q}_v(x)\mathfrak{Q}_v(x) = 0$.

The part of the QCD Lagrangian describing a quark is

$$\mathcal{L}_q = \bar{Q}(x)(i\slashed{D} - m_Q)Q(x), \tag{2.15}$$

where $D_\mu = \partial_\mu + ig A_\mu^A T^A$, $g$ is the QCD coupling constant, $A_\mu^A$ are the gluon fields, and $T^A$ are the $SU_{color}(3)$ generators. Specializing this to heavy quarks and writing it in terms of the fields, $Q_v(x)$,

we get

$$
\begin{aligned}
\mathcal{L}_Q &= (\bar{Q}_v + \bar{\mathfrak{Q}}_v)(i\slashed{D} + m_Q \slashed{v} - m_Q)(Q_v + \mathfrak{Q}_v) && (2.16) \\
&= (\bar{Q}_v + \bar{\mathfrak{Q}}_v)i\slashed{D}\left(\frac{1+\slashed{v}}{2}Q_v + \frac{1-\slashed{v}}{2}\mathfrak{Q}_v\right) + \bar{\mathfrak{Q}}_v 2m_Q \mathfrak{Q}_v && (2.17) \\
&= \bar{Q}_v(iv\cdot D)Q_v - \bar{\mathfrak{Q}}_v(iv\cdot D - 2m_Q)\mathfrak{Q}_v + \bar{Q}_v i\slashed{D}\mathfrak{Q}_v + \bar{\mathfrak{Q}}_v i\slashed{D}Q_v. && (2.18)
\end{aligned}
$$

To obtain the second line, the projection and orthogonality properties of the $Q_v(x)$ and $\mathfrak{Q}_v(x)$ fields are used. To get the third line the gamma matrix commutation relation $\{\gamma_\alpha, \gamma_\beta\} = 2g_{\alpha\beta}$ is also used. The structure of the Lagrangian shows that excitations of the antiquark fields $\mathfrak{Q}_v(x)$ are suppressed by $2m_Q$ with respect to $D$ operator which is of the scale of $\Lambda_{QCD}$.

The $\mathfrak{Q}_v(x)$ fields can be integrated out using the classical equation of motion:

$$
-(iv\cdot D - 2m_Q)\mathfrak{Q}_v + i\slashed{D}Q_v = 0 \;\longrightarrow\; \mathfrak{Q}_v = \frac{i\slashed{D}Q_v}{(2m_Q - iv\cdot D)}. \qquad (2.19)
$$

The Lagrangian can be then be expressed in terms of just the $Q_v(x)$ fields:

$$
\begin{aligned}
\mathcal{L}_Q &= \bar{Q}_v\left(iv\cdot D + i\slashed{D}\frac{i\slashed{D}}{(2m_Q - iv\cdot D)}\right)Q_v && (2.20) \\
&= \bar{Q}_v\left(iv\cdot D - \frac{\slashed{D}\slashed{D}}{2m_Q} + \mathcal{O}(1/m_Q^2)\right)Q_v && (2.21) \\
&= \bar{Q}_v iv\cdot DQ_v - \bar{Q}_v\frac{D^2}{2m_Q}Q_v - \bar{Q}_v\frac{\sigma^{\mu\nu}G_{\mu\nu}}{4m_Q}Q_v + \mathcal{O}(1/m_Q^2). && (2.22)
\end{aligned}
$$

In the last equation the $\mathcal{O}(1/m_Q)$ term has been regrouped into two terms. The first $\mathcal{O}(1/m_Q)$ term, $\frac{D^2}{2m_Q}$, corresponds to the kinetic energy of the heavy quark. The second term, $\frac{\sigma^{\mu\nu}G_{\mu\nu}}{4m_Q}$, depends on the heavy quark spin and corresponds to the chromomagnetic interaction of the heavy quark with the light degrees of freedom. Both terms violate the heavy quark flavor symmetry, but only the second term violates heavy quark spin symmetry.

## 2.3.1   Hadrons Containing Heavy Quarks

The Lagrangian derived in the previous section is the Lagrangian of an effective theory which is useful for describing the properties and interactions of hadrons containing a heavy quark. By switching to the $Q_v(x)$ fields, the mass and energy of the heavy quark have been subtracted out. The scale of the $D$ operator acting on the $Q_v(x)$ fields is therefore $\Lambda_{QCD}$, instead of $m_Q$ as it is for $D$ acting on the original $Q(x)$.

The expectation values of the operators in the Langragian appear in the calculations of many physical processes and are therefore of experimental interest. These expectation values measure properties of the heavy hadrons for which they are defined, but because they involve strong physics

at low energies, they cannot be calculated. They are conventionally defined as

$$\overline{\Lambda} \equiv \frac{1}{2}\langle H_Q | \bar{Q}_v (iv \cdot D) Q_v + \mathcal{H}_{\text{light}} | H_Q \rangle \tag{2.23}$$

$$\lambda_1 \equiv -\frac{1}{2}\langle H_Q | \bar{Q}_v D^2 Q_v | H_Q \rangle \tag{2.24}$$

$$\lambda_2 \equiv \frac{1}{16(J^2 - S_Q^2 - S_{\text{light}}^2)}\langle H_Q | \bar{Q}_v \sigma^{\mu\nu} G_{\mu\nu} Q_v | H_Q \rangle, \tag{2.25}$$

where $\mathcal{H}_{\text{light}}$ is the Hamiltonian describing the light degrees of freedom, $J$ is the spin of the hadron, $S_Q$ is the spin of the heavy quark, and $S_{\text{light}}$ is the spin of the light degrees of freedom. Alternative names for these expectation values used by some authors are $\mu_\pi^2 = -\lambda_1$ and $\mu_g^2 = 3\lambda_2$. In the heavy quark limit, $\overline{\Lambda}$ corresponds to the contribution to the meson mass of the light degrees of freedom. One of the goals of this thesis is to measure $\overline{\Lambda}$ and $\lambda_1$ as discussed in Section 2.4.3 ($\lambda_2$ can be determined from the meson masses, see below).

Because of the heavy quark flavor and spin symmetries, hadrons can be classified into categories for which the light degrees of freedom are the same and thus have the same $\overline{\Lambda}$, $\lambda_1$, and $\lambda_2$ parameters. For example, the parameters are the same for the $D$, $D^*$, $B$, and $B^*$ mesons. The parameters for states other than those mentioned are generally written with a subscript to define to which set of states it corresponds. E.g., $\overline{\Lambda}_s$, $\lambda_{1,s}$ and $\lambda_{2,s}$ for $D_s^{(*)}$ and $B_s^{(*)}$ mesons, and $\overline{\Lambda}_\Lambda$, $\lambda_{1,\Lambda}$ and $\lambda_{2,\Lambda}$ for the $\Lambda_c$ and $\Lambda_b$ baryons.

The masses of the heavy hadrons can be expressed in terms of these parameters as:

$$m_{H_Q} = m_Q + \overline{\Lambda} - \frac{\lambda_1}{2m_Q} + \frac{(J^2 - S_Q^2 - S_{\text{light}}^2)}{m_Q}\lambda_2 + \mathcal{O}(1/m_Q^2), \tag{2.26}$$

Using this relation, $\lambda_2$ can be calculated from the mass splitting of the pseudoscalar and vector states $H$ and $H^*$, respectively:

$$\lambda_2 = (m_{H^*} - m_H)\frac{m_Q}{2} = (m_{H^*} - m_H)\frac{\overline{m}_H}{2} \tag{2.27}$$

where $m_{\overline{H}} \equiv \frac{m_H + 3m_{H^*}}{4}$. The effect of $\overline{\Lambda}$ in this expression is suppressed by $\overline{\Lambda}/m_Q$. Using the PDG values [1], we get $\lambda_2 = 0.121 \pm .001\,\text{GeV}^2$ for the $B$ system and $0.139 \pm .001\,\text{GeV}^2$ for the $D$ system.

## 2.4 HQET, OPE, and Inclusive Semileptonic Heavy Meson Decay

One of the major applications of HQET is the calculation of inclusive semileptonic heavy meson decay properties. In Section 2.2, the differential decay rate was written in terms of a tensor

$$W^{\alpha\beta} = \sum_X \frac{1}{2M_B}(2\pi)^3 \langle \bar{B}|J_{\text{had}}^{\alpha}|X\rangle\langle X|J_{\text{had}}^{\beta}|\bar{B}\rangle\delta^4(p_B - q - p_X) \qquad (2.28)$$

which cannot be calculated with perturbative techniques. This current can be calculated using the operator product expansion (OPE) and HQET.

This calculation is complex and will be presented diagrammatically in order to make the basic features clear. The first step of the calculation is to use $\sum_X |X\rangle\langle X| = 1$ to remove the dependence on the structure of $X$ from the calculation:



$$(2.29)$$

The next step is to use the OPE, which is a short distance approximation in which the product of two operators that are separated by distance $x$ are expanded in terms of local operators and $x$ dependent coefficients. The expansion is written as

$$J_{\text{had}}^{\alpha}(0)J_{\text{had}}^{\beta}(x) = \sum_i C_i^{\alpha\beta}(x)\mathcal{O}_i^{QCD}(0), \qquad (2.30)$$

where $C_i^{\alpha\beta}(x)$ are calculable functions and $\mathcal{O}_i^{QCD}$ are operators of the full QCD. In diagram, the two currents are contracted into one:



$$(2.31)$$

The final step is to replace the QCD operators by the HQET operators:



$$(2.32)$$

The result is that the form factors of the hadronic current can be written in terms of expectations of the HQET operators

$$W^{\alpha\beta} = \sum_i C_i^{\alpha\beta}(q^2, q_0)\langle B|\mathcal{O}_i^{HQET}|B\rangle \qquad (2.33)$$

where the $C_i^{\alpha\beta}$ are calculable functions of the kinematical variables $q^2$ and $q_0$ and the $\langle B|\mathcal{O}_i^{HQET}|B\rangle$ are nonperturbative operator expectation values which are not specific to the semileptonic $B$ decay process. The $C_i^{\alpha\beta}$ also depend on the initial and final state quark masses, which must be converted to the meson masses using $\overline{\Lambda}$ and the heavy quark mass formula (Equation 2.26). The up quark mass is simply set to zero as it is much smaller than any of the other energies or masses involved. The $C_i^{\alpha\beta}$ have been calculated by several authors [14].

## 2.4.1  The Differential Decay Rate

From the hadronic current, the fully differential decay rate is calculated using Equation 2.9. Because the HQET operators are used in the expansion, the higher dimension terms in the OPE are suppressed by powers of $\Lambda_{QCD}/m_Q$. At zeroth order in $\Lambda_{QCD}/m_Q$, the differential decay rate only depends on the $m_Q$ parameter which is $m_H$. At first order in the expansion, there are no new terms in the OPE expansion, but the quark mass is no longer equal to the hadron mass and $\overline{\Lambda}$ is needed to translate between them. At order $(\Lambda_{QCD}/m_Q)^2$, terms proportional to $\lambda_1$ and $\lambda_2$ appear, as well as, terms proportional to $\overline{\Lambda}^2$. The highest order to which the semileptonic $B$ meson differential decay rate has been calculated is $(\Lambda_{QCD}/m_Q)^3$, where there are six nonperturbative parameters, which have been named $\rho_1$, $\rho_2$, $\tau_1$, $\tau_2$, $\tau_3$, and $\tau_4$. The nonperturbative parameters up to order $(\Lambda_{QCD}/m_Q)^2$ are defined in Equations 2.23 through 2.25 and are the same parameters as are used in the mass formula (Equation 2.26). The same nonperturbative parameters appear in the calculation of a large variety of $B$ meson decays including the $B \to X_c l\overline{\nu}$, $B \to X_u l\overline{\nu}$, $B \to X_s \gamma$, $B \to X_s l^+l^-$, the hadronic decays presented in Section 1.3, and other processes. The calculation of the $C_i^{\alpha\beta}(x)$ functions entails expansions in $\alpha_s(\mu)$ where $\mu$ is a hard scale, typically of order $m_Q$. A nonperturbative parameter extracted from the data with a calculation at a particular scale $\mu$ can only be used in other calculations using the same scale $\mu$.

The short distance approximation made in the operator product expansion means that the results for the differential decay rate cannot predict long-distance behavior such as which final state mesons are produced. Any comparison between the experiment and these calculations must sum over a number of final states. The short-distance approximation allows the calculation to be computed in terms of quarks instead of hadrons. Unfortunately, the theoretical uncertainty related to this assumption, known as quark-hadron duality, is not well understood. Whether a sufficiently large region of phase space and final state modes have been included in a calculation has to be evaluated on a case by case basis.

## 2.4.2 The Total Decay Rate

The largest region of phase space that can be included is all of it, which gives the total decay rate:

$$
\begin{aligned}
\Gamma(B \to X_c l \overline{\nu}) = \frac{G_F^2 |V_{cb}|^2}{192\pi^3} \overline{m}_B^5 \Big[ & 0.370 - 0.115\,\overline{\Lambda} - 0.012\,\overline{\Lambda}^2 + 0.\,\overline{\Lambda}^3 \\
& - 0.04\,\lambda_1 - 0.10\,\lambda_2 - 0.01\,\lambda_1\overline{\Lambda} + 0.02\,\lambda_2\overline{\Lambda} \\
& - 0.02\,\rho_1 + 0.02\,\rho_2 - 0.02\,\tau_1 + 0.\,\tau_2 \\
& - 0.03\,\tau_3 - 0.02\,\tau_4 - 0.040\,\epsilon - 0.022\,\epsilon_{BLM}^2 + 0.007\,\epsilon\overline{\Lambda} \Big].
\end{aligned}
\tag{2.34}
$$

This can be used to extract $|V_{cb}|$ from data. In order to do so, the values of nonperturbative parameters must be determined. The easiest parameter to determine is $\lambda_2$, because it is related to the mass splitting between the heavy pseudoscalar and vector mesons (see Equation 2.27). The other parameters can be determined by measuring moments of the differential decay rate.

## 2.4.3 Moments of the Differential Decay Rate

The OPE allows the full three-dimensional differential decay rate to be calculated, but because of the quark-hadron duality assumption, the prediction cannot be compared directly to the data. Instead moments of the kinematic distributions are measured and compared to predictions. To date the moments that have been measured are of the lepton energy [15] and hadronic invariant mass squared [16, 17] in $B \to X_c l \overline{\nu}$ events and of the photon energy in $B \to X_s \gamma$ events [13].

A moment,

$$
\langle M \rangle \quad = \quad \int d\vec{x}\; M(\vec{x})\; \frac{d\Gamma}{d\vec{x}},
\tag{2.35}
$$

can be expressed as,

$$
\begin{aligned}
\langle M \rangle = M_1 & + M_2\,\overline{\Lambda} + M_3\,\overline{\Lambda}^2 + M_4\,\overline{\Lambda}^3 \\
& + M_5\,\lambda_1 + M_6\,\lambda_2 + M_7\,\lambda_1\overline{\Lambda} + M_8\,\lambda_2\overline{\Lambda} \\
& + M_9\,\rho_1 + M_{10}\,\rho_2 + M_{11}\,\tau_1 + M_{12}\,\tau_2 \\
& + M_{13}\,\tau_3 + M_{14}\,\tau_4 + M_{15}\,\epsilon + M_{16}\,\epsilon_{BLM}^2 + M_{17}\,\epsilon\overline{\Lambda},
\end{aligned}
\tag{2.36}
$$

where $M_i$ are numerical constants which are calculated using the HQET and OPE technique described above and $\epsilon$ and $\epsilon_{BLM}^2$ express the size the $\alpha_s$ and $\alpha_s^2 \beta_0$ terms of the expansion. The default values of $\epsilon$ and $\epsilon_{BLM}^2$ are one. For all of the measurements mentioned above, the $M_i$ coefficients have been calculated [18, 19, 20, 21, 22]. In reference [22], Bauer and Trott have tabulated a large variety of moments. These calculations will be used to interpret the measurements presented here.
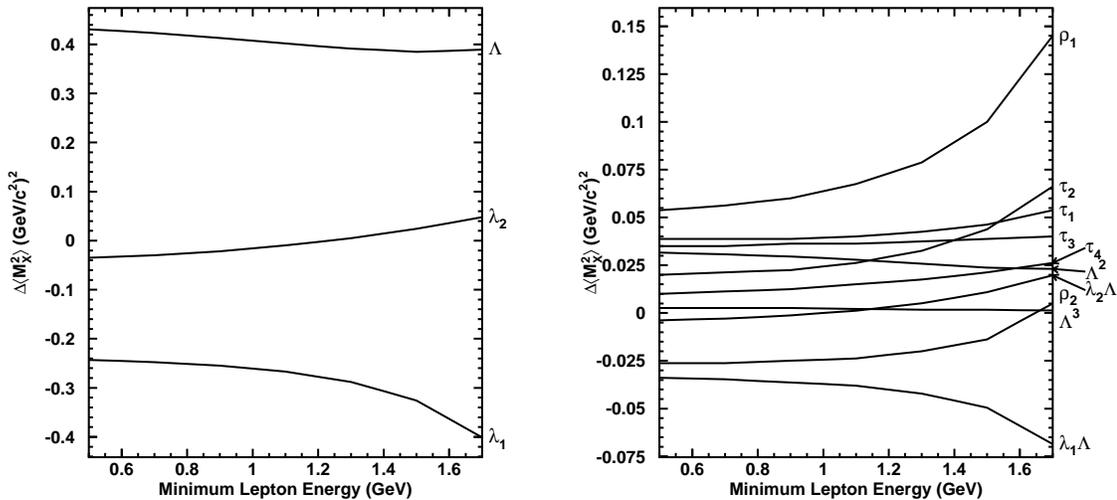
Figure 2.2: Variation of Corrections to $\langle M_X^2 - \overline{M}_D^2 \rangle$ as a Function of the Minimum Lepton Energy. The first and second order terms in the $\Lambda_{QCD}/m_Q$ operator product expansion are shown on the left and the third order terms are shown on the right.

Each of the nonperturbative parameters in the $\Lambda_{QCD}/m_Q$ expansion has a dimension corresponding to the order at which it occurs. The dimension of the first-order parameter $\overline{\Lambda}$ is GeV, the dimensions of the second-order parameters, $\lambda_1$ and $\lambda_2$, are $\text{GeV}^2$ and the dimensions of the third-order parameters, $\rho_1$, $\rho_2$, $\tau_1$, $\tau_2$, $\tau_3$, and $\tau_4$, are $\text{GeV}^3$. Because the nonperturbative parameters arise from strong interactions they are expected to be of order $\Lambda_{QCD}^n$. Because $\lambda_2$ can be calculated from the mass formula, a measured moment can be treated as a polynomial function of $\overline{\Lambda}$ and $\lambda_1$ where the third-order terms are varied by $\pm(0.5\text{GeV})^3$ as a measure of the theoretical uncertainty.

If a moment is sampling too little phase space for the quark-hadron duality to apply, the operator product expansion is expected to break down. One test of this is if the higher-order terms become large. Figure 2.2 shows the variation of the contribution of the nonperturbative parameters to the moment $\langle M_X^2 - \overline{M}_D^2 \rangle$ for $B \to X_c l\overline{\nu}$ process as a function of the minimum lepton energy cut [22]. As the phase space over which the moment is calculated is reduced, the higher-order terms get larger. Another way to test the quark-hadron duality assumption is to have a set of measurements which over constrains the nonperturbative parameters. This is one of the motivations for the moments measurement presented in this thesis.

## 2.4.4   Extracting $|V_{ub}|$ from Data

Another application of the HQET-OPE calculation is in the extraction of $|V_{ub}|$. Measuring $|V_{ub}|$ in inclusive semileptonic $B$ decay requires separating $B \to X_u l\overline{\nu}$ events from $B \to X_c l\overline{\nu}$ events. This

is usually done by measuring the partial branching fraction of the $B \to X_u l \overline{\nu}$ process in a region of phase space where $B \to X_c l \overline{\nu}$ does not contribute. This exploits the fact that the up quark is lighter than the charm quark, and the leptonic system will therefore have more energy. Theoretical input is then needed to relate the measured partial branching fraction to $|V_{ub}|$. Such calculations can be made with the HQET-OPE machinery.

The region of phase space most accessible to experiment is the lepton energy endpoint, $E_l > (m_B^2 - m_D^2)/2m_B$. This region was used in the first measurement of $B \to X_u l \overline{\nu}$ establishing a nonzero value of $|V_{ub}|$ [10]. The lepton energy endpoint region is very small and the standard operator product expansion cannot be used to relate the partial branching fraction in the endpoint region to $|V_{ub}|$. The problem can be understood by comparing the maximum lepton energy of a free bottom quark decay, $m_b/2 \approx 2.46$ GeV, to the maximum lepton energy of $B$ meson decay, $m_B/2 = 2.64$ GeV. The difference between the two end points $m_B/2 - m_b/2 = 180$ MeV is entirely due to nonperturbative effects and is of order $\Lambda_{QCD}$. The size of the region in which charm does not contribute is also of order $\Lambda_{QCD}$, $m_B/2 - (m_B^2 - m_D^2)/2m_B = 330$ MeV. This means that a large part of the region in which $B \to X_c l \overline{\nu}$ does not contribute can only be reached by nonperturbative effects. This can be addressed theoretically by resumming the condensate terms into a shape function [23], which can be measured in $B \to X_s \gamma$ decays. This has been done at CLEO [9, 13].

An alternative method is to move away from using the lepton energy to distinguish between $B \to X_u l \overline{\nu}$ and $B \to X_c l \overline{\nu}$. The kinematic variables that can be used to separate the two modes are $M_X^2$ and $q^2$. Using $M_X^2$ to separate $B \to X_u l \overline{\nu}$ from $B \to X_c l \overline{\nu}$ has similar problems as the lepton energy. The variable $q^2$ on the other hand does not [25]. Bauer and colleagues have shown that a combination of $q^2$ and $M_X^2$ can be used to define a region for which the partial differential decay rate can be predicted with good accuracy [26]. This technique is used in this analysis to extract $|V_{ub}|$.

## 2.5 HQET Predictions for Exclusive Semileptonic Heavy to Heavy Decays

In order to measure properties of the inclusive semileptonic $B$ meson decays, it is necessary to model the constituent exclusive modes in detail. The form factors of these modes can have important effects on the efficiency of the data selection and the resolutions of the measurements of kinematical variables. Projections of the differential decay rates used in this study are shown in Figure 2.3.

The exclusive spectrum is dominated ($\approx 80\%$) by $B \to D l \overline{\nu}$ and $B \to D^* l \overline{\nu}$. In addition to these modes, the other possible significant contributions are from four $D^{**}$ final states and from nonresonant multibody finals states such as $D^{(*)} n \pi$, $D_s^{(*)} K^{(*)}$, and $\Lambda_c p X$. The baryonic final states are constrained to be very small by the available phase space and direct searches [27].

The $q^2$ distributions of the $B \to D l \overline{\nu}$ and $B \to D^* l \overline{\nu}$ partial decay rates are very different. The
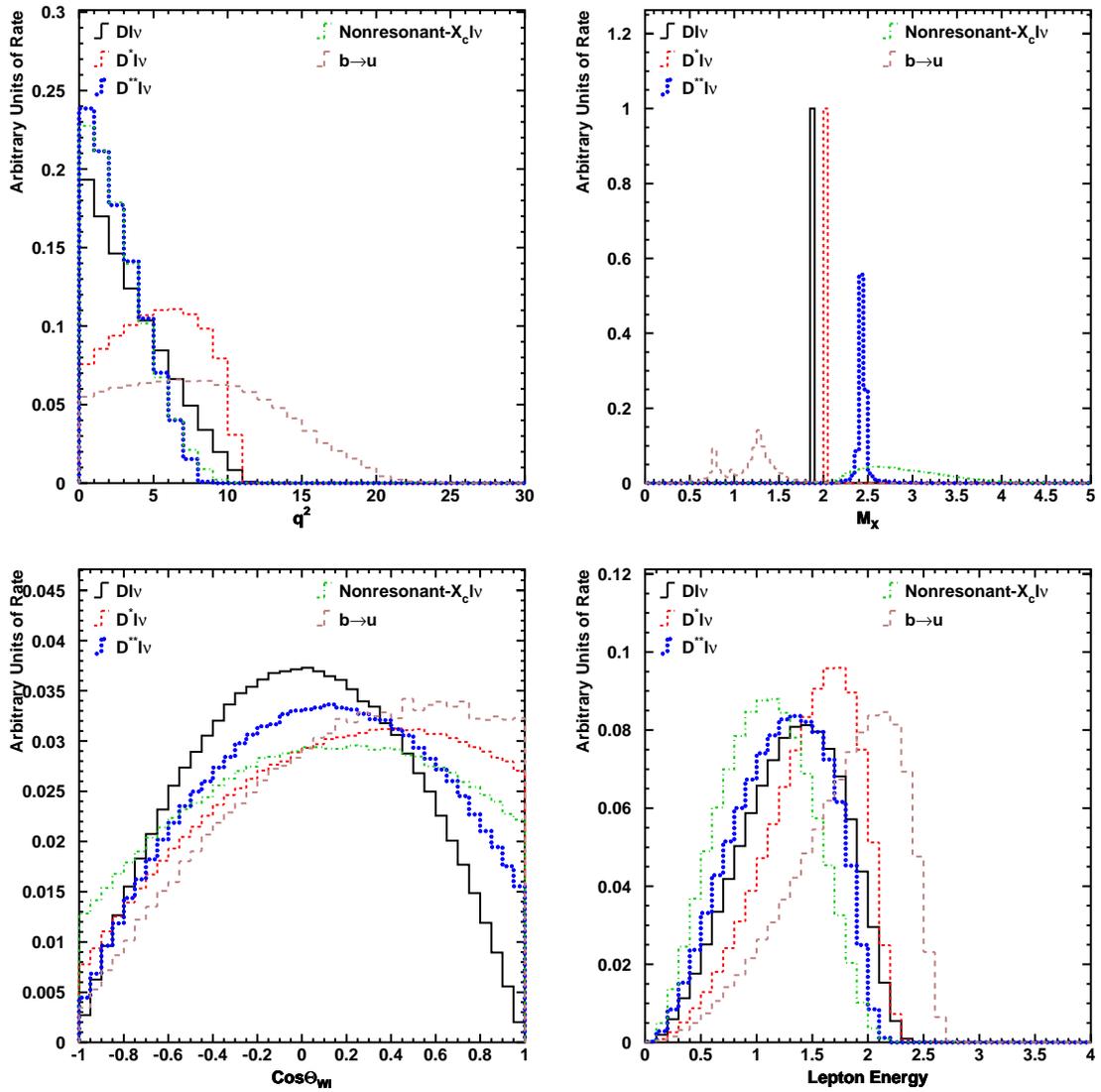
Figure 2.3: Projections of the Differential Decay Rates for the Individual Modes.
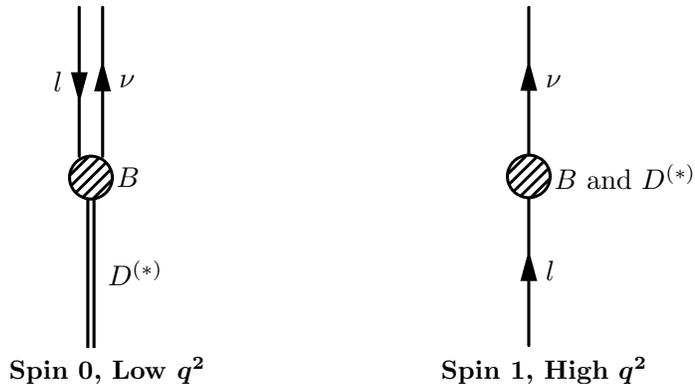
Figure 2.4:   Diagram of the Correlations between the Final State Spin and $q^2$.

difference is the direct result of the spin structure of the final state and the left-handedness of the weak current. The weak force only couples left-handed particles and right-handed antiparticles, so the lepton and antineutrino in the final state of a $\bar{B} \to D^+ l^- \bar{\nu}_l$ decay must be left and right-handed, respectively. Because the handedness describes the correlation of the momentum direction and the spin, a left-handed lepton and a right-handed anti-neutrino must be moving in the same direction to have a combined spin of zero (see Figure 2.4). Similarly, if they are moving in opposite directions they will have a spin of one. When a $B$ meson decays to a $D$ meson through a virtual $W$ boson, the spin of the $W$ must be zero to conserve spin, so the lepton and neutrino will prefer to move in the same direction which corresponds to low $q^2$. If the $B$ decays to a $D^*$ meson, the spin of the virtual $W$ must be one and the lepton and neutrino will prefer to move in opposite directions, which corresponds to high $q^2$. So $B \to D^* l \bar{\nu}$ decays will tend to have higher $q^2$ than $B \to D l \bar{\nu}$ decays.

Heavy quark spin and flavor symmetries provide strong constraints on the form factors of $B \to D l \bar{\nu}$ and $B \to D^* l \bar{\nu}$ decays. The Lorentz structure of these decays has two unknown form factors, $h_+(w)$ and $h_-(w)$, for $B \to D l \bar{\nu}$ and four more form factors, $h_V(w)$, $h_{A_1}(w)$, $h_{A_2}(w)$, and $h_{A_3}(w)$, for $B \to D^* l \bar{\nu}$, where $w = v \cdot v' = (m_B^2 + m_{D^{(*)}}^2 - q^2)/2m_B m_{D^{(*)}}$. In the heavy quark limit, two of the form factors are zero and those remaining are all equal:

$$h_+(w) = h_V(w) = h_{A_1}(w) = h_{A_3}(w) = \xi(w)$$
$$h_-(w) = h_{A_2}(w) = 0. \tag{2.37}$$

The universal form factor, $\xi(w)$, is known as the Isgur-Wise function [28].

At the maximum available $q^2$, the final state $D^{(*)}$ meson is at rest in the $B$ meson frame, $w = v \cdot v' = 1$. Because the two mesons are then not moving with respect to each other and are identical because in the approximation of heavy flavor symmetry, the value of the Isgur-Wise function at this point is one, $\xi(1) = 1$. Similarly, the form factors for the $D^{**}$ mesons and nonresonant decays are zero at $w = 1$ in the heavy quark limit. Both of these conclusions are modified by corrections

of order $\Lambda_{QCD}/m_Q$ and $\alpha_s(m_c)$. A complete review of heavy-to-heavy transitions can be found in reference [7].

### 2.5.1   $B \rightarrow Dl\overline{\nu}$

The differential decay rate for $B \rightarrow Dl\overline{\nu}$ can be expressed in terms of $\xi(w)$ as

$$\frac{d\Gamma}{dw} = \frac{G_F^2|V_{cb}|^2}{48\pi^3}M_B^5(1+r)^2r^3(w^2-1)^{3/2}\xi(w)^2, \tag{2.38}$$

where $r = M_D/M_B$. Because the normalization of the Isgur-Wise function is known at $w = 1$, it is usually expanded around that point. The first and second terms of the expansion are written as

$$\xi(w) = \xi(1)(1 - \rho_D^2(w-1) + c_D(w-1)^2). \tag{2.39}$$

Because the heavy quark limit is an approximation, the subscript $D$ is added to specify the final state mode. Dispersion relations can be used to relate the curvature $c_D$ to the slope $\rho_D^2$ [29],

$$c_D \approx 1.05\rho_D^2 - .15 \tag{2.40}$$

The PDG world average, $\rho_D^2 = 1.19 \pm 0.15 \pm 0.12$, is measured assuming this curvature relation [1].

### 2.5.2   $B \rightarrow D^*l\overline{\nu}$

The phenomenology of $B \rightarrow D^*l\overline{\nu}$ decays is similar to that of $B \rightarrow Dl\overline{\nu}$ decays, but with more form factors. Because the form factors of the $B \rightarrow D^*l\overline{\nu}$ decay are only equal in the heavy quark limit, the differential decay rate is written with three form factors instead of one. These are $h_{A_1}(w)$, $R_1(w) = h_V(w)/h_{A_1}(w)$, and $R_2(w) = (h_{A_3}(w) + r^*h_{A_2}(w))/h_{A_1}(w)$. In the heavy quark limit, $R_1(w)$ and $R_2(w)$ are one.

The full decay rate is differential in $q^2$, the $D^*$ decay angle $\cos\theta^*$, the $W$ decay angle $\cos\theta_{W\ell}$, and the angle between the $D^*$ decay plane and the $W$ decay plane $\chi$. Since in this analysis the $D^*$ is not reconstructed, neither $\cos\theta^*$ nor $\chi$ is measured. After integrating out these unmeasured angles the differential decay rate is

$$\frac{d\Gamma}{dw\,d\cos\theta_{W\ell}} = \frac{G_F^2|V_{cb}|^2M_B^5}{(4\pi)^3}r^{*3}(1-r^*)^2(w^2-1)(w+1)^2h_{A_1}(w)^2$$
$$\left[\sin^2\theta_{Wl}(1+\frac{w-1}{1-r^*}(1-R_2))^2 + \cos^2\theta_{W\ell}(1+\frac{w-1}{w+1}R_1^2) + 4R_1\cos\theta_{W\ell}\sqrt{\frac{w-1}{w+1}}\right], \tag{2.41}$$

where $r^* = M_{D^*}/M_B$. The form factor $h_{A_1}(w)$ is expanded analogously to $\xi(w)$, defining $\rho_{A_1}^2$ and $c_{A_1}$.

Like $B \to Dl\overline{\nu}$, dispersion relations can be used to constrain the shapes of the form factors. In reference [29], $h_{A_1}(w)$ is expanded in the variable $z \equiv \frac{\sqrt{w+1}-\sqrt{2}}{\sqrt{w+1}-\sqrt{2}} \approx \frac{1}{8}(w-1) - \frac{1}{32}(w-1)^2$ to get the relation:

$$
\begin{aligned}
h_{A_1}(z) &= h_{A_1}(z=0)(1 - 8\rho_{A_1}^2 z + (53\rho_{A_1}^2 - 15)z^2 + (231\rho_{A_1}^2 - 91)z^3) & (2.42)\\
h_{A_1}(w) &= h_{A_1}(w=1)(1 - \rho_{A_1}^2(w-1) + \frac{69\rho_{A_1}^2 - 15}{64}(w-1)^2 + \mathcal{O}((w-1)^3)) & (2.43)
\end{aligned}
$$

implying

$$
c_{A_1} = \frac{69\rho_{A_1}^2 - 15}{64}. \tag{2.44}
$$

In addition, the dispersion relations give $R_1(w)$ and $R_2(w)$ to be

$$
\begin{aligned}
R_1(w) &= R_1(1)(1 - .12(w-1) + .05(w-1)^2) & (2.45)\\
R_2(w) &= R_2(1)(1 + .11(w-1) - .06(w-1)^2) & (2.46)
\end{aligned}
$$

The PDG world average, $\rho_{A_1}^2 = 1.51 \pm 0.05 \pm 0.12$, is measured assuming these curvature relations [1].

### 2.5.3  $B \to D^{**}l\overline{\nu}$

There is very little known experimentally about the structure of the $B \to D^{**}l\overline{\nu}$ form factors or rate. Only one final state $B \to \bar{D}_1(2420)^0 l\nu_l$ has been observed directly [30]. In the context of HQET however some basic properties can be assessed. The main observation is that the form factors must vanish at $w = 1$. This leads to very steeply rising form factors as a function of $w$. The form factors of the various $B \to D^{**}l\overline{\nu}$ decays are calculated at order $\Lambda_{QCD}/m_Q$ in reference [31].

### 2.5.4  Nonresonant $B \to X_c l\overline{\nu}$

Even less is known about the $B \to X_c l\overline{\nu}$ nonresonant decay mode. Isgur suggests that the nonresonant component is composed of many different modes each of which contributes a very small amount to the total [32]. This case would make it nearly impossible to exclusively reconstruct individual modes in an experiment. For the $B \to D^{(*)}\pi l\nu$ subset of the nonresonant modes, Goity and Roberts have calculated the form factors using a combination of HQET and chiral perturbation theory [33].

## 2.6  Radiative Corrections

In any decay or scattering process, the charged particles involved emit photons. The process of photon emission is known as final state radiation and falls into a broader category of corrections to
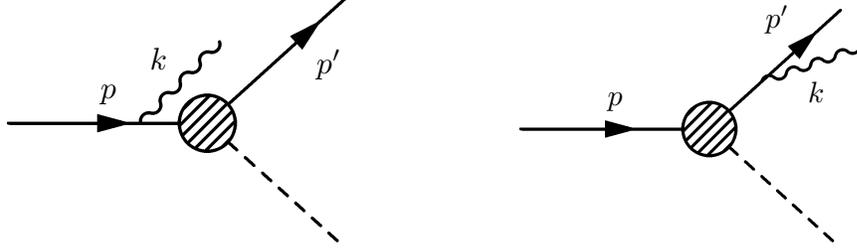
Figure 2.5: Contributions to Radiative Corrections.

tree-level processes known as radiative corrections. Final state radiation is generally a very small effect, but it becomes enhanced when the charged particle is an electron, due to the small mass.

The calculation of radiative corrections is not generally separable from the matrix element calculation of a given process, except in the soft photon limit [12] . The contributions to the radiative corrections in this limit are diagrammed in Figure 2.5.

The probability for the process with a radiated photons is

$$d\Gamma(X \to Y\gamma) = d\Gamma_B(X \to Y) \; \frac{\alpha}{4\pi^2} \int d^3k \, \frac{1}{k} \left( \frac{2p' \cdot p}{(p \cdot k)(p' \cdot k)} - \frac{p^2}{(p \cdot k)^2} - \frac{p'^2}{(p' \cdot k)^2} \right), \qquad (2.47)$$

where $d\Gamma_B(X \to Y)$ is the decay rate with no radiation, $p$, $p'$, and $k$, are the momentum four-vectors of the parent, the charged daughter, and the radiated photon.

The $p \cdot k$ and $p' \cdot k$ terms in the denominator come from the propagators of the parent or daughter between the hard event and the radiation of the photon. These terms can be written as $p^{(\prime)} \cdot k = kE^{(\prime)}(1 - \beta^{(\prime)} \cos\theta^{(\prime)})$, where $\beta^{(\prime)}$ and $E^{(\prime)}$ are the velocity and energy of the parent or daughter and $\theta^{(\prime)}$ is the angle between the parent or daughter and the photon direction. In this form it is clear that if $\beta$ or $\beta'$ is near one there will be an approximate pole in the $\cos\theta$ or $\cos\theta'$ distributions respectively. This pole causes the rate of photon emission from light particles to be large. It is convenient to work in the rest frame of the parent, where $\beta = 0$:

$$d\Gamma(X \to Y\gamma) = d\Gamma_B(X \to Y) \; \frac{\alpha}{2\pi} \int dk \, d\cos\theta \, \frac{1}{k} \left( \frac{2}{(1 - \beta' \cos\theta)} - 1 - \frac{m_{p'}^2}{E'^2(1 - \beta' \cos\theta)^2} \right), \; (2.48)$$

where $m_{p'}$ is the mass of the daughter particle. The photon emission probability diverges as $k$ gets small. This divergence is canceled by corrections to the photon-fermion vertex. There are whole chapters of quantum field theory textbooks devoted to this phenomenon [12].

While the effect of the divergence on the total rate is canceled, the number of photons emitted is infinite at very low photon energies. The total energy radiated by these photons is small, so the very soft part of the photon emission can be ignored. The effects of photon emission are then expressed as a *splitting* function $\rho(k)$ which is applied to a Born level differential decay rate to get the differential

decay rate in order to include radiated photons above an energy cut-off of $\epsilon$:

$$d\Gamma(X \to Y(\gamma)) = d\Gamma_B(X \to Y)\,\rho(k)dk. \qquad (2.49)$$

For simplicity, a one-dimensional splitting function is shown, although $\cos\theta$ could also be included. The splitting function divides the correction into a part which is negligible, $k < \epsilon$, where no correction is made and a part, $k > \epsilon$, where an appropriate photon is radiated:

$$\rho(k) = \left(1 - \frac{\alpha}{\pi}N\left(\ln\epsilon + \frac{3}{4}\right)\right)\delta(k) + \frac{\alpha}{\pi}N\left(\frac{1 + (1 - \frac{k}{k_{max}})^2}{2k}\right)\Theta(k - \epsilon), \qquad (2.50)$$

where $k_{max}$ is the maximum possible photon energy and $N = \ln(E'^2/m_{p'}^2) - 2$. The first term in the expression corresponds to not radiating a photon and the second term corresponds to radiating a photon of energy $k > \epsilon$. It should be noted that, in reality, there are always photons radiated, but $\epsilon$ is chosen as the scale at which they can be neglected. This scale is typically set by the experimental sensitivity. If photons below the energy $\epsilon$ cannot be detected, either directly or through effects on the kinematics of the other particles, they can be ignored. Notice that if $\epsilon$ is set very low, then the coefficient of the first term will be negative and the integral of the second term will be greater than zero. In that case, multiple photon emission would have to be considered. Section 5.2.3 discusses in detail how these corrections are applied in the analysis presented.

# Chapter 3  The CLEO II and II.V Detectors and Datasets

The data used for the measurements presented in this thesis were recorded by the CLEO detector which was constructed to analyze the $e^+e^-$ collisions produced by the Cornell Electron Storage Ring (CESR). An integrated luminosity of 9.4 fb$^{-1}$ was accumulated on the $\Upsilon(4S)$ resonance, $E_{cm} \approx 10.58$ GeV, and an additional 4.5 fb$^{-1}$ was taken 60 MeV below the $\Upsilon(4S)$ resonance where there is no $B\overline{B}$ production. The $\Upsilon(4S)$ is the first and largest resonance above $B\overline{B}$ production threshold, 10.56 GeV, and is the optimal operating point to produce $B$ mesons. The data taken below the $B\overline{B}$ production threshold is very useful for understanding the backgrounds from $e^+e^- \rightarrow q\overline{q}$ events, where $q = u, d, s,$ or $c$. The hadronic cross section as a function of center of mass energy in the region of the four lowest lying $\Upsilon$ resonances is shown in Figure 3.1. The flat offset from zero is due to $e^+e^- \rightarrow q\overline{q}$ events and the peaks are the $\Upsilon$ resonances as labeled.
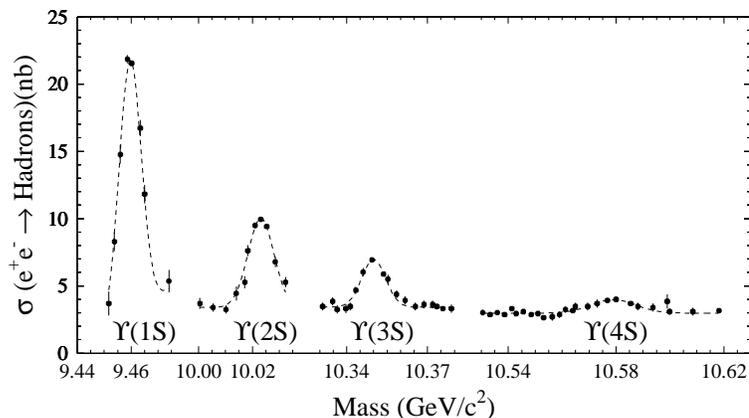


Figure 3.1: Hadronic Cross Section Near the $\Upsilon$ Resonances.

## 3.1  The Cornell Electron Storage Ring

CESR is the final stage of a series of machines which accelerate electrons and positrons to the energies necessary to produce $\Upsilon(4S)$ mesons. When CESR was first proposed, the $b$-quark had not yet been discovered, but by the time the construction was completed some of the $\Upsilon$ resonances had been observed at Fermilab [34]. Although the accelerator has a potential energy range of 9-16 GeV, CESR has been operated primarily around the upsilon resonances from 9 to 12 GeV. It has been tuned and upgraded over more than two decades to generate higher and higher luminosities. In addition to CLEO, a second detector CUSB was also operated at a second interaction region
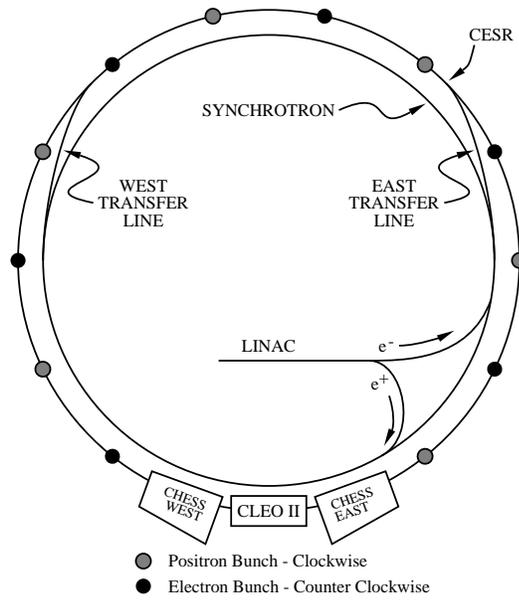
Figure 3.2: Diagram of the Accelerator Layout.

in CESR from 1979 to 1990. The high-energy, high-current beams stored in CESR are also used to produce synchrotron radiation for the Cornell High Energy Synchrotron Source (CHESS), which uses the X-rays for experiments in biology and solid state physics.

The electrons that form the beam originate in the electron gun, where they are extracted from a hot filament using a high voltage. When filling the storage ring, the electrons are accelerated in a 30-meter linac to 300 MeV before being injected into the synchrotron. The synchrotron then accelerates the electrons to the operating energy after which they are transfered to the storage ring where they are accumulated. When filling CESR with positrons, a tungsten target is placed in the path of the beam at the point in the linac where the electrons have an energy of 140 MeV. The positrons produced by the collision of the electron beam with the target are collected and accelerated to an energy of 200 MeV for injection into the synchrotron. The injection process is repeated at a rate of 60 Hz to accumulate beam currents of 200 to 300 mA. Figure 3.2 is a diagram which shows the layout of the components mentioned above.

After the injection of both beams is complete, they are brought into collision. The stored beams slowly lose particles through collisions with residual gas in the beam pipes or with the walls of the beam pipe. This causes a decrease in the rate of collisions leading to the need to refill the machine after a period of data taking. These periods, called runs, are typically 60 to 90 minutes and the refilling time is approximately 10 minutes. The length of the runs is chosen to optimize the total number of collision recorded.

The beams inside the storage ring are structured by the RF acceleration cavities as trains of bunches. A typical configuration consists of nine trains of three bunches each. Both the electron and positron beams are inside the same storage ring. When first operated, CESR had one bunch per beam and there were two detectors, CLEO in the south collision hall, and CUSB in the north collision hall. With one bunch per beam in the machine, the bunches collide twice per revolution, once in each of the collision halls. Over the years more bunches were added to increase luminosity and the CUSB detector was removed from the north area. This creates many more potential collision points than are desirable. To avoid collisions outside of the north and south interaction regions, the electron and positron bunches are perturbed from their nominal orbits. The bunches then oscillate around the nominal orbits missing each other as they cross paths. This configuration pioneered at CESR is known as pretzel orbits. To avoid interactions in the north area a high voltage electrostatic separator is used deflect the beams around each other.

The standard luminosity formula is

$$\mathcal{L} \;=\; \frac{f N_B n_{e^-} n_{e^+}}{4\pi \sigma_h \sigma_v}, \tag{3.1}$$

where f is the frequency at which a n individual bunch passes through the interaction region, $n_{e^-}$ and $n_{e^+}$ are the number of electrons and positrons in a bunch respectively, $N_B$ is the number of bunch pairs in the machine, and $\sigma_h$ and $\sigma_v$ are the horizontal and vertical sizes of the beam at the interaction point. The number of particles per bunch was in the range $2-3\times10^{11}$. The number of bunches was varied during the years of data collection; the peak luminosity occurred with a $9\times3=27$ bunch configuration. Because the particles are very near the speed of light, the frequency is defined by the circumference of the accelerator, $f = c\,/\,768\,\mathrm{m} \approx 3.9\times10^5\,\mathrm{Hz}$. The horizontal and vertical beams sizes are 300 $\mu$m and 6 $\mu$m respectively. The highest achieved luminosity in the CLEO II and CLEO II.V running period, $8.25\times10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, is much lower that would be predicted by Equation 3.1. This is because the equation only applies in the limit where the beam-beam interactions are negligible. There are many complicated effects which arise from the beam-beam interactions which degrade this predicted luminosity, causing it to saturate and scale with the bunch charge as opposed to the bunch charge squared.

## 3.2 The CLEO Detector

The data used in the measurements presented in this thesis were taken with two different configurations of the CLEO detector known as CLEO II (1990-1995) and CLEO II.V (1995-1999). CLEO is a general purpose barrel-shaped detector and covers 95% the $4\pi$ solid angle in both configurations. Tracking systems, an electromagnetic calorimeter and a time of flight system sit inside a 1.5 T super-

conducting solenoidal magnet. Outside of the magnet is a steel flux return which is constructed of three-layers of steel, between which muon chambers are located. Figures 3.3 and 3.4 show side and end views, respectively, of the layout of the CLEO II detector. For the CLEO II.V configuration, the innermost tracking system was replaced with a silicon vertex detector and the gas used in the main drift chamber was changed. The remaining systems are the same for the two configuration. This section will give an overview of the detector systems and the performance of the detector as a whole. A more detailed description of the CLEO II detector can be found in reference [35] and a description of the silicon vertex detector can be found in [36].

The coordinate system used to describe CLEO detector has the $x$ direction pointing toward the center of the ring, the $y$ direction pointing up, and the $z$ direction pointing east which is the direction that the electron beam travels. The detector is approximately symmetric about the beam axis. The angle with respect to the beam axis, $\cos\theta$, is often used in describing the coverage of the detector components. Also used in the description of the tracking system is the momentum of a particle transverse to the beam axis, $p_T$.

## 3.2.1   Tracking

The tracking systems occupy the volume inside a radius of 95 cm from the beam axis. In the CLEO II configuration, immediately outside a 3.5 cm gold-plated beryllium beam pipe is a six-layer straw tube chamber known as the Precision Tracking Layer (PTL). In the CLEO II.V configuration, the PTL and beam pipe are replaced by a three-layer silicon vertex detector and 2 cm gold-plated beryllium beam pipe. Vertex measurement is not important to the measurement presented in this thesis, so the difference does not play an important role. Outside the innermost tracker, is a ten-layer intermediate drift chamber known as the Vertex Detector (VD). The third and outermost tracking chamber is a 51-layer drift chamber known as the DR.

The PTL, the innermost tracking chamber, covers the radii from 4.5 cm to 7.7 cm. It is a six-layer straw tube detector, which provides $r - \phi$ measurement only. It is constructed of aluminized mylar tubes which are glued together to provide mechanical stability. There are 64 axial wires per layer, with the tube size increasing with radius so that the tubes maintain contact with their neighbors, as shown in Figure 3.5. The radius of the cells varies from approximately 2 mm for the inner layer to 4 mm for the outer layer. The hit resolution of the PTL is approximately 90 $\mu$m.

In the CLEO II.V configuration, the PTL is replaced by a three-layer silicon detector, the SVX (shown in Figure 3.6). There are 96 wafers in the SVX, all double side so that they give both $r - \phi$ and $r - z$ information. The strips on the $r - \phi$ side have a pitch of 28 $\mu$m, one fourth of which are read out, and on the $r - z$ side the pitch is 100 $\mu$m, all of which are read out. The hit resolution is approximately 19 $\mu$m on the $r - \phi$ side and varies from 25 to 55 $\mu$m on the $r - z$ side. This allows secondary vertices from $D$ and $\tau$ decays to be separated from the original collision point (the
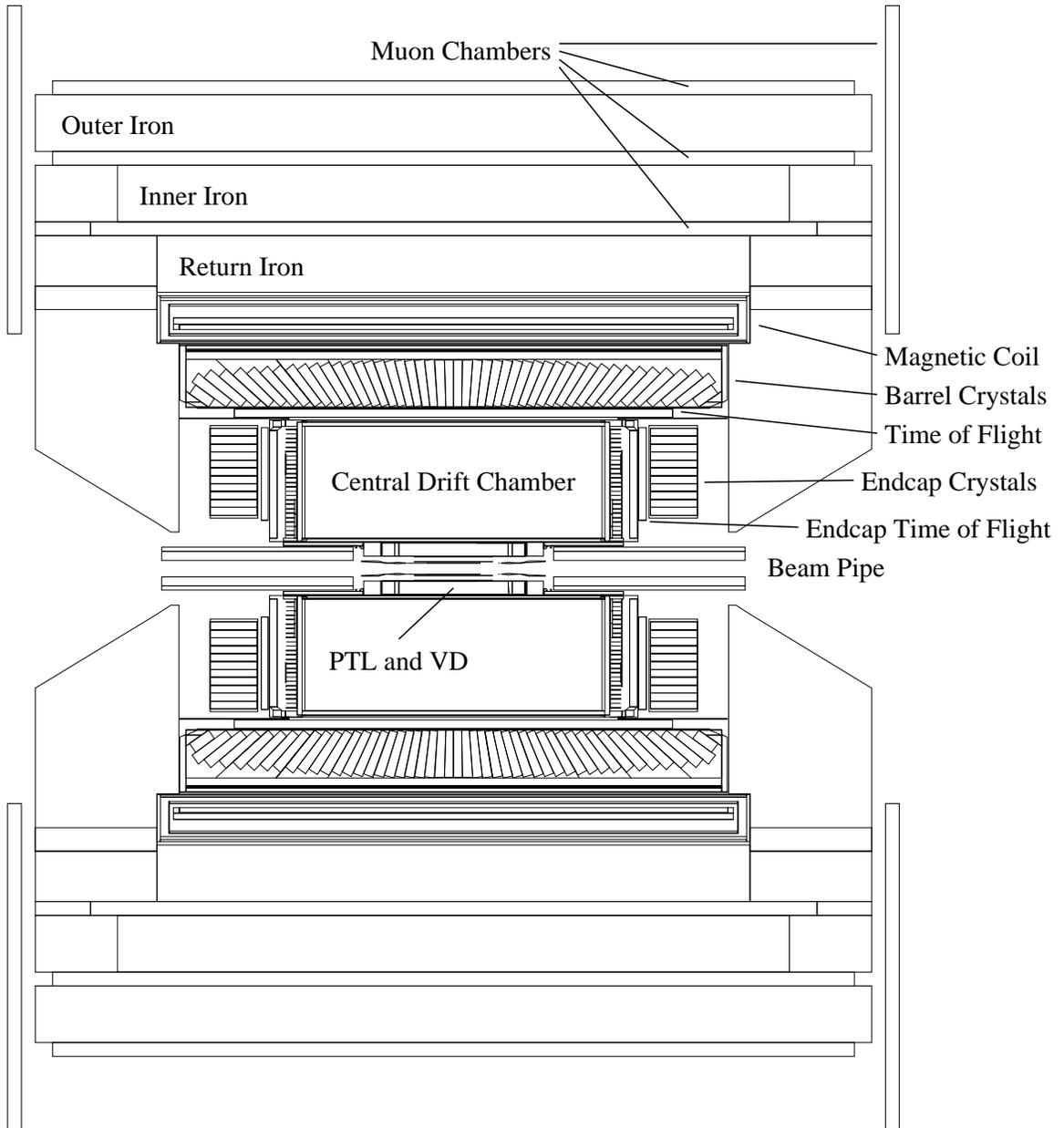
Figure 3.3: Side View of the CLEO II Detector.
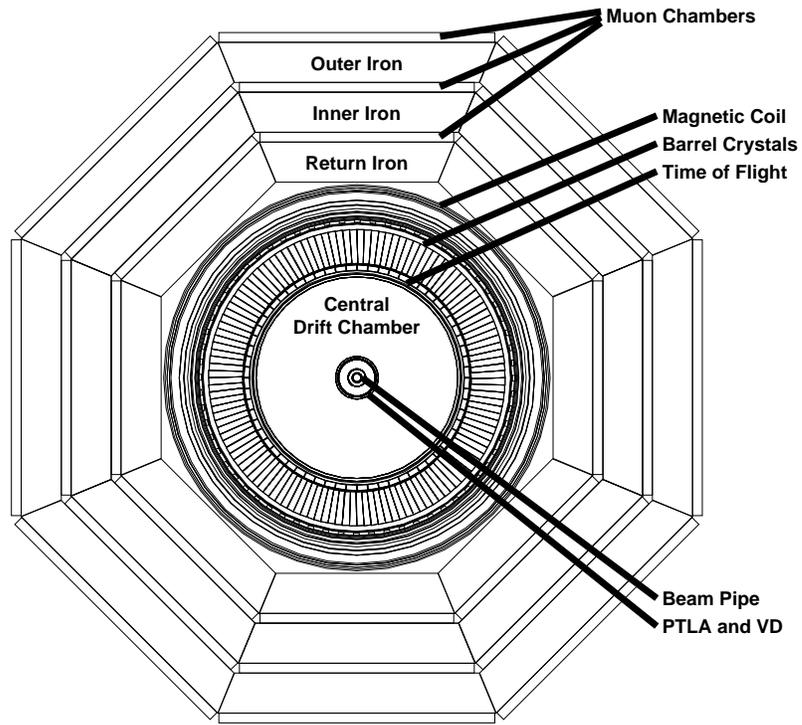
Figure 3.4: End View of the CLEO II Detector.
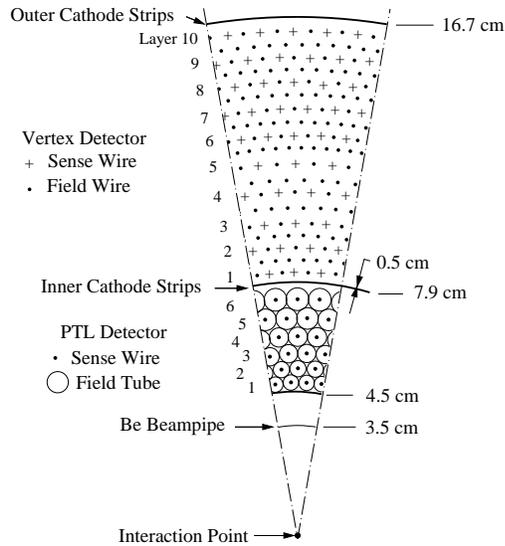


Figure 3.5: Diagram of the Inner (PTL) and Intermediate (VD) Tracking Chambers.
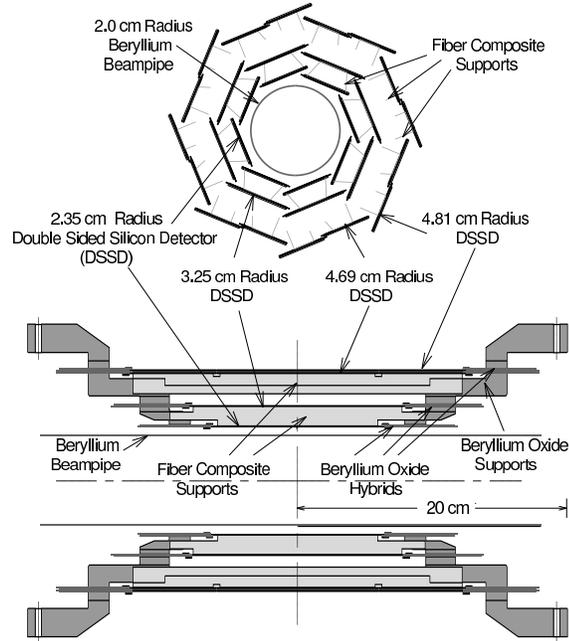
Figure 3.6: Diagram of Silicon Vertex Detector (SVX).

primary vertex). This capability does not play an important role in the measurement presented here.

The VD occupies the radii from 8.1 cm to 16.4 cm. It has ten axial layers with a cell radius of approximately 5 mm. The cell structure is shown in Figure 3.5. The sense wires in the VD are resistive and a $z$ measurement is made using the charge division between the two ends of the wire. Additional $z$ information is provided by instrumented cathode pads which define the field at the inner and outer edges of the detector. Tracks which have transverse momentum of 40 MeV/$c$ or greater and $|\cos\theta|$ of less than 0.9 will pass through all the layers of the VD.

The DR, the main tracking drift chamber, covers radii from 17.8 to 94.7 cm with 40 axial layers and 11 stereo layers. Stereo layers have wires which are at an angle to the beam axis, providing $z$ information when the hits are combined with the other layers. The axial layers are grouped in sets of three or five separated by a single stereo layer, as shown in Figure 3.7. The cells are all approximately rectangular and equal in size with a half-cell size of 7 mm. Inner and outer cathodes provide additional $z$ information. The hit resolutions varies from 100 $\mu$m in the central part of a cell to 200 $\mu$m for the outer part of the cell. The hit efficiency is approximately 85%. Tracks which have transverse momentum of 225 MeV/$c$ or greater and a $|\cos\theta|$ of less than 0.71 will pass through all the layers of the DR.

In the CLEO II configuration, all three chambers use a combination of 50% argon and 50% ethane as the drift gas. For the CLEO II.V configuration, the gas for the DR is changed to 60% helium

DR Outer Shell

DR Outer Cathode Strips

51
50
49
48
47
46
45
44
43
42
41
40

· Field Wires
∘ Stereo Sense Wires
• Axial Sense Wires

12
11
10
9
8
7
6
5
4
3
2
1

DR inner Cathodes
DR Inner Shell
VD Outer Shell
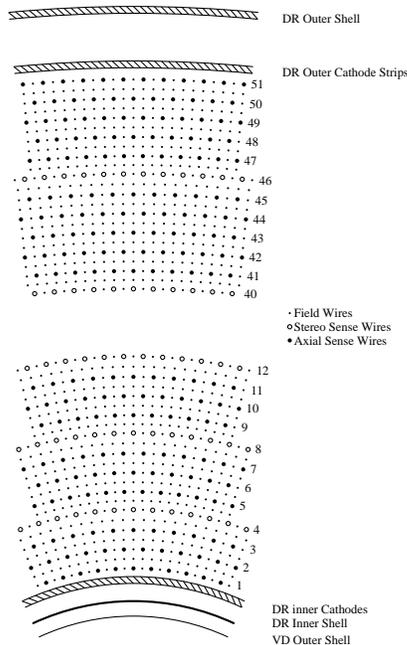
Figure 3.7: Diagram of the Main Drift Chamber (DR).

and 40% propane. The helium-propane mixture improves the hit efficiency and spatial resolution of the cells, improving the mean spatial resolution from 172 $\mu$m to 150 $\mu$m. The use of the helium-propane mixture also reduced the multiple scattering due to the gas particles, further improving the momentum resolution.

The information from the entire tracking system is processed by pattern recognition software which groups the hits into tracks. The three-momentum and the $r - \phi$ and $z$ impact parameters of each track are extracted from the sets of hits using a chi-squared minimization technique. The fitting software uses a *Kalman Filter* [37] algorithm to take into account the scattering due to the material through which the particles pass. For the CLEO II configuration, there is a total of 1.58% of a radiation length of material in the active region, the majority of which at the interface of the VD and DR. The overall transverse momentum resolution for the CLEO II configuration can be summarize as

$$\left(\frac{\sigma_{p_T}}{p_T}\right)^2 = (0.0017 p_T)^2 + (0.0050)^2. \tag{3.2}$$

The tracking resolution is improved by approximately 20% for the CLEO II.V configuration. The difference is not important for this measurement.

In addition to the momentum, the specific ionization of a track is measured using the pulse height information of the hits on the track. The rate of ionization energy loss, $dE/dx$, of a charged
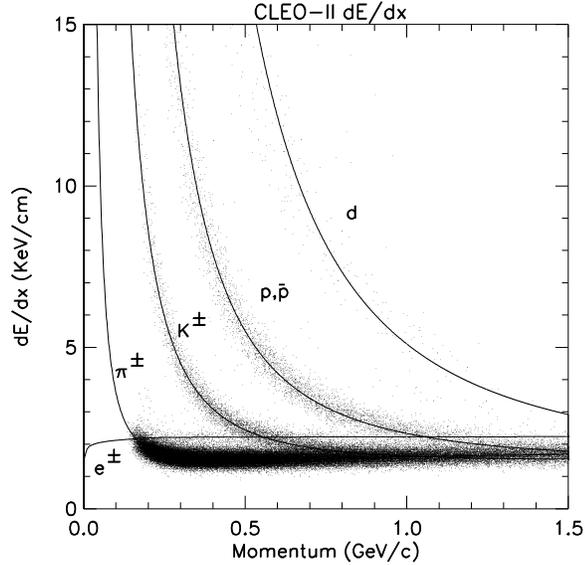
Figure 3.8: Measured Ionization versus Momentum for CLEO II.

particle is a function of the velocity, $\beta$, and not the mass of the particle. Combined with the track momentum, the velocity gives a measurement of the mass of the particle. Since the possible types of particles, $\pi$, $K$, $p$, $\mu$, and $e$ all have different masses, the $dE/dx$ measurement can be used to identify the particle type. The separation between the different possible particle hypotheses depends on the momentum of the particle. The energy loss versus momentum for the CLEO II configuration can be seen in Figure 3.8. The CLEO II.V performance is approximately 15% improved. To make the $dE/dx$ information easy to use, the variables

$$\sigma_{dE/dx}^h = \frac{(dE/dx)_{\mathrm{measured}} - (dE/dx)_{\mathrm{predicted}}^h}{\sigma((dE/dx)_{\mathrm{measured}})} \tag{3.3}$$

are constructed, where $(dE/dx)_{\mathrm{measured}}$ is the measured energy loss, $(dE/dx)_{\mathrm{predicted}}^h$ is the energy loss predicted for a particle of type $h = e$, $\mu$, $\pi$, $K$, or $p$ with the momentum of the track, and $\sigma((dE/dx)_{\mathrm{measured}})$ is the expected error in the energy loss.

**The TRKMNG Package**

The method of neutrino reconstruction described in Section 4.3 requires a best estimate of the set of charged particles produced in the event as well as their four-momenta. There are many pathologies of the pattern recognition which can cause more tracks to be identified than there were charged particles in the event. A package called TRKMNG has been developed to identify tracks which are the result of these pathologies and select the set of tracks most likely to represent the charged particles in the event.

The pathologies dealt with fall into the categories:

- **Curlers** occur when a particle has insufficient transverse momentum to reach the calorimeter, ($p_T < 225$ MeV). The particle will curl back to the beam axis and then back out again, until it stops from energy loss or exits the chamber in the z-direction. Each half loop may be identified as a separate track. Because the particle travels along a helix it will return close to the beam axis, but will be displaced along the axis from the interaction point. Furthermore, because of energy loss, the particle will also not return exactly to the beam axis. Kinematically, the tracks from a curler will either have similar $\phi$ and curvature, or opposite $\phi$ and curvature. Complex criteria based on these features are used to select the track most likely to be the initial outbound arc of the particle.

- **Ghosts** occur when the hits produced by one particle are reconstructed as two tracks. Because both tracks came from only one particle, there will be only one hit between the two tracks at each layer. The characteristics by which ghost pairs are identified are that tracks are parallel in the $r - \phi$ plane and there are large number of layers at which only one of the tracks has a hit and small number of layers at which both tracks have hits. Once a ghost pair is identified, the track with the larger number of stereo and cathode hits is kept.

- **Scattered particles and decays in flight** may also produce two tracks, one before the decay or scattering event and one after. The two tracks produced will meet at a point in the tracking chambers. The resulting track pair is identified by similar criteria to ghosts, but with the additional constraint that the hits from one track are inside the point at which the tracks meet and the hits from the other track are outside that point. In this case the inner track is used, except for track-shower matching where the outer track is used.

- **Low density** tracks usually occur when the density of hits in a region of the detector is large and a track is pieced together from hits left by multiple different particles or multiple passes of a curler. In this case, there are fewer hits on a track than would be expected from the 85% hit effiecieny of the DR. These tracks can be identified by the low number of hits relative to the number of layers crossed, and because they don't necessarily pass near the IP in the $r - \phi$ plane.

### 3.2.2   Time of Flight Detector

Outside of the main drift chamber is a time of flight system (TOF) for particle identification (see Figure 3.4). The barrel TOF consists of 64 strips of plastic scintillator which cover angles of $|\cos\theta| \leq 0.8$. The scintillator is 5 cm thick and has a short (2.1 ns) decay time. At both ends of the scintillator are light pipes which direct the photons to photomultiplier tubes that are outside the magnetic field.
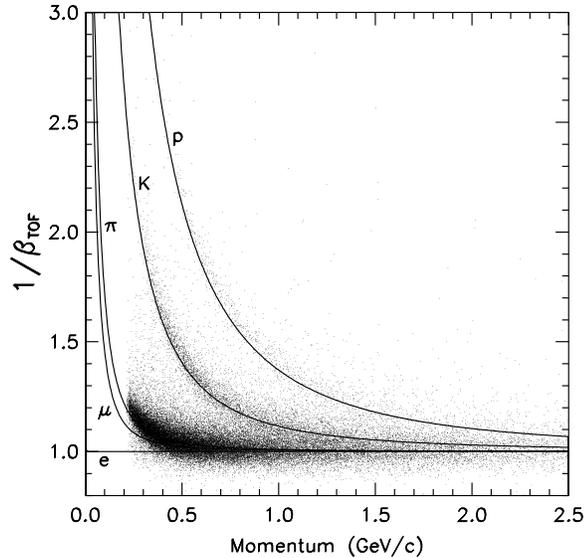
Figure 3.9: Measured Particle Velocity versus Momentum for CLEO II.

At each endcap the TOF has 28 wedges which cover the angles $0.8 < \cos\theta \le 0.95$. The wedges are the same material and thickness as the endcap, and are connected at the inner edge to photomultiplier tubes. The reference signal for the time measurement is derived from the bunch crossing time. The timing resolution of the TOF is 154 ps for pions. The measured $\beta$ is plotted as a function of the momentum in Figure 3.9. The separation between pions and kaons below momenta of 1 GeV/c is greater than two standard deviations. The TOF is also used in the first level of the trigger and for determining in which bunch the collision occurred. For the TOF information, variables $\sigma^h_{TOF}$ are defined analogously to those of the $dE/dx$ measurement.

### 3.2.3  Electromagnetic Calorimeter

The crystal calorimeter (CC) is constructed of 7800 cesium iodide crystals which are configured as a barrel with an inner radius of 1.02 m and two endcaps with a front plate located 1.25 m from the interaction point (see Figures 3.3 and 3.4). The crystals are thallium doped to increase the light output. At the back end of each crystal are four photodiodes which are summed in analog electronics before being digitized.

Cesium iodide is a very dense material (4.51 g/cm$^3$), with a short radiation length (1.83 cm), and a small Moliere radius (3.8cm). These properties allow for good angular resolution and a compact design. The crystals have a transverse size of 5cm, compatible with a Moliere radius, and are 30cm long, which is 16 radiation lengths. The barrel crystals are tapered, so that when packed together they form a projective geometry in which a photon produced at the interaction point will create a shower along the length of the crystal. This is done so that longitudinal fluctuations in the shower

will not affect the position measurement.

The energy and position resolutions depend on region of the detector because of the geometry and the amount of material in front of the calorimeter. The best resolution is in the central barrel, which covers the angles $|\cos\theta| < 0.71$, where the material is at a minimum. For the angles $|\cos\theta| > 0.71$, the photons must pass through the drift chamber endplate, cooling, and readout infrastructure. Geometry also plays a role in degrading the resolution outside of the central barrel. There is a region of overlap between the endcap and the barrel between $32°$ and $36°$ ($0.81 < |\cos\theta| < 0.85$), in which showers are split between the two. Also, the endcap is not projective.

The clustering of the light output measurements of each crystal is done in two stages. In the readout system, the crystal hits are saved if they either exceed a seed threshold or are neighbors of a crystal which exceeds the seed threshold. All measurements of more than 1 MeV in a crystal must be kept in order not to degrade the resolution. The reconstruction software applies gain and pedestal corrections and then sums the crystals in the cluster to get a single shower energy and position. The number of crystals included in the sum depends on the shower energy in order to minimize the effect of electronic noise on low energy showers. The resolution can be parameterized as

$$
\begin{aligned}
\text{Barrel}: \quad \frac{\sigma_E}{E}(\%) &= \frac{0.35}{E^{0.75}} + 1.9 - 0.1E \\
\sigma_\phi(\text{mr}) &= \frac{2.8}{\sqrt{E}} + 1.9 \\
\sigma_\theta(\text{mr}) &= 0.8\,\sigma_\phi\,\sin\theta \\
\text{Endcap}: \quad \frac{\sigma_E}{E}(\%) &= \frac{0.26}{E} + 2.5 \\
\sigma_\phi(\text{mr}) &= \frac{3.7}{\sqrt{E}} + 7.3 \\
\sigma_\theta(\text{mr}) &= \frac{1.4\sigma_\phi}{\sqrt{E}} + 5.6,
\end{aligned}
\tag{3.4}
$$

where E is the photon energy in GeV.

## The SPLITF Package

In addition to photons, showers are also produced by charged particles, $K_L^0$ mesons, and neutrons. In reconstructing a neutrino, it is important to remove the charged particle contributions, so as not to double count their energy (events with $K_L^0$ mesons and neutrons cannot be used in neutrino reconstruction effectively anyway, because their energy is poorly measured in the calorimeter). The first line of defense is to remove showers which have tracks pointing at them, but hadronic showers sometimes produce particles which travel sufficiently far before stopping that they are reconstructed as a separate shower. Figure 3.10 shows a conceptual representation of the effect. These secondary showers, known as splitoffs, can be identified with some efficiency by the shape of the shower. Hadronic showers tend to be broader than photon showers. In many CLEO analyses, this is quantified with the a variable $E_9/E_{25}$ which is the sum of the energy deposited in the central nine crystals of a shower divided by the energy deposited in the central 25 crystals. For photons this is in general
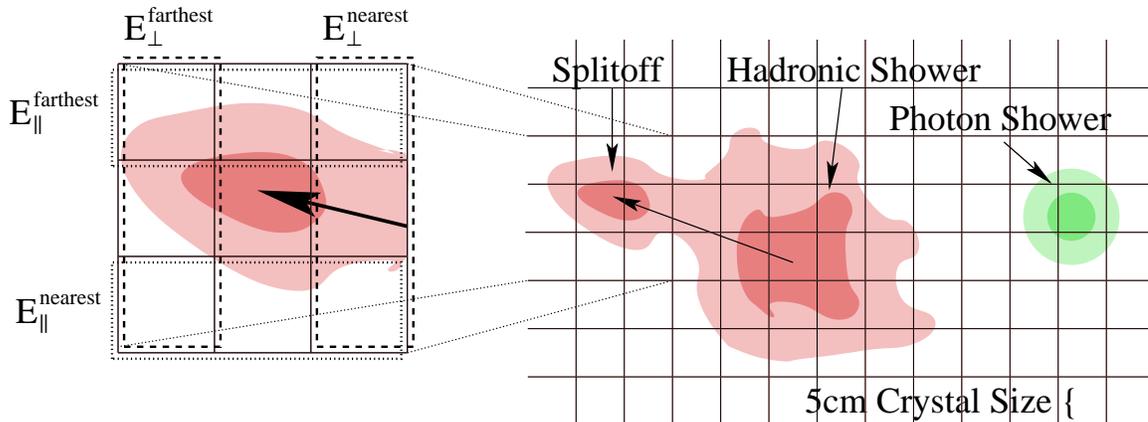
Figure 3.10: Conceptual Representation of a Splitoff Shower.

very close to one.

For neutrino reconstruction purposes, a more sophisticated method using a neural network has been developed. This package is known as SPLITF. For showers within a 25° cone of a track at the calorimeter, four scaled energy variables are constructed from the energy distribution in the three by three block of crystals centered on the most energetic crystal in the shower. The variables are constructed by summing the energies of the three crystals on each side of the block and dividing by the total energy of the entire block. These are then classified by their orientation. For the two sides closer to perpendicular to line of flight connecting the track impact point to the center of the block, the variables are called $E_\perp^{nearest}$ and $E_\perp^{farthest}$. Similarly, the variables for the remaining two sides, closer to parallel, are called $E_\parallel^{nearest}$ and $E_\parallel^{farthest}$. A diagram of the crystals used for each of these variables is shown in Figure 3.10. Figure 3.11 shows a Monte Carlo simulation of $E_\perp^{nearest}$ vs $E_\perp^{farthest}$ for photons and splitoffs. For a splitoff, the deposited energy tends to be along the line of flight and tends to be closer to the originating track. Photons have a more symmetric distribution of deposited energy. The neural network combines this information with the angle by which the shower is separated from the track, the angle of the line of flight relative to the grid of the crystals, and the $E_9/E_{25}$ of the shower which is matched to the track. A Monte Carlo simulation of semileptonic $B$ decays shows that on average the algorithm successfully rejects 185 MeV of energy deposited by splitoffs in an event while losing only 40 MeV of energy deposited by photons.

### 3.2.4 Superconducting Magnet

The magnetic field in the CLEO detector is provided by a NbTi superconducting coil. The coil sits in a helium cryostat outside the calorimeter (see Figure 3.3). It provides a highly uniform 1.5 Tesla magnetic field. The homogeneity of the field has been tested using an NMR probe and a Hall probe. The inhomogeneity was found to be less than 0.1% in the drift chamber volume. The accelerator's
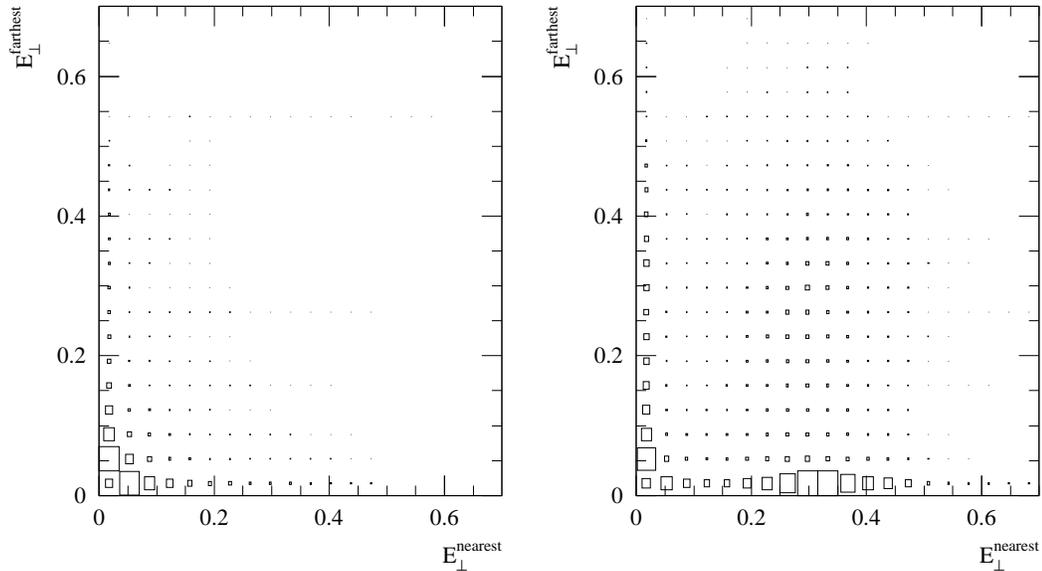
Figure 3.11: Monte Carlo Simulation of the Distribution of $E_\perp^{nearest}$ vs $E_\perp^{farthest}$. The distribution for photons is on the left and for splitoffs is on the right.

final focus quadruple magnets which sit inside the detector volume have small external fields which increase the inhomogeneity to 0.2% in the drift chamber volume. The stability of the magnetic field is monitored during data taking with a NMR probe located just outside the main drift chamber. The absolute value of the magnetic field used in the track reconstruction software is adjusted so that the measured $D^0$, $\phi$, and $K_S^0$ meson masses are their known values.

## 3.2.5   Muon Chambers

The barrel muon chambers are located in the steel flux return outside the solenoid at depths into the iron of 36, 72, and 108 cm from the solenoid. At normal incidence these correspond to 3, 5 and 7 interaction length between the interaction point and the respective chambers (the first interaction length is due to the rest of the detector). The barrel muon chambers cover the region $|\cos\theta| < 0.71$. There is a single set of muon chambers for each endcap which cover the ends of the detector extending the total muon identification coverage to 85% of the solid angle. Each chamber consists of three-layers of proportional counters which are 5 m long and 8.3 cm wide. The structure of the chambers is shown in Figure 3.12. Each counter has eight anode wires each of which are surrounded on three sides by a graphite coated plastic cathode. External copper pickup strips measure the coordinate parallel to the wires.

The chambers provide a resolution of 2.4 cm perpendicular to the wires and 5.3 cm along the wires. In order to match a track to a hit in the muon chambers, the track position is extrapolated out to the muon chamber position. This extrapolation has significant error because of the large

Copper Cathode Strip

Foam Separator

Anode Wire

PVC Shell

Graphite Cathode

Copper Ground
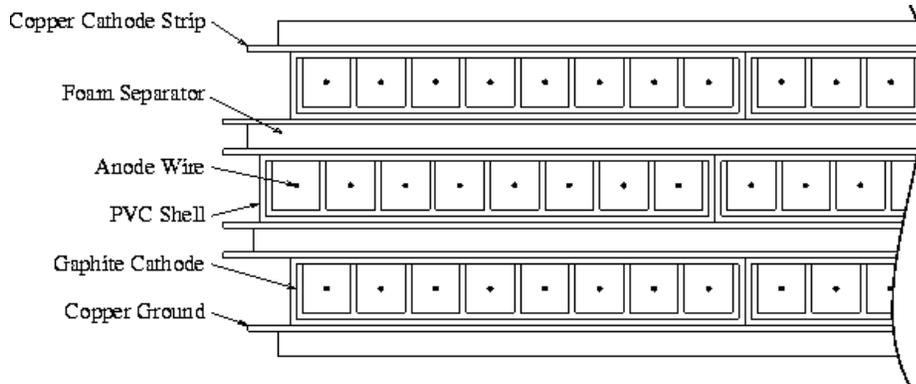
Figure 3.12: Diagram of the Structure of a Muon Chamber.

amount of material traversed. Perpendicular to the wires, the extrapolation resolutions for a 5.28 GeV muon are 3.7, 4.6 and 5.7 cm for the depths of 3, 5, and 7 interaction lengths respectively, and 7.2 cm for the endcaps. Parallel to the wires, the resolutions are 5.5, 7.0, 7.5, and 9.0 cm. A hit is considered to belong to a track if the $\chi^2$ between the extrapolated track position and the hit position is less than 16.

The efficiency of the muon chambers is measured using $e^+e^- \rightarrow \mu^+\mu^-\gamma$ events. These are identified by requiring one of the tracks to have a well-identified muon (momentum greater than 4.0 GeV and penetration to at least 7 interaction lengths). The other track in the event is then assumed to be a muon and the identification properties are examined. Figure 3.13 shows a Monte Carlo simulation of the distribution of penetration depth for all particles and for muons. Figure 3.14 shows the muon measured efficiencies.

### 3.2.6 Electron Identification

The most powerful variable for electron identification is the ratio of the energy deposited in the calorimeter over the momentum of the track, $E/p$. Electrons produce electromagnetic showers in the calorimeter depositing the majority of their energy, so $E/p$ is near one. Muons tend to pass through the calorimeter as minimum ionizing particles leaving behind very little energy, so $E/p$ is much less than one. Hadronic particles $\pi$, $K$, and $p$ shower hadronically either in the calorimeter or farther out in the detector. Hadronic showers have a larger spatial extent than electromagnetic showers and rarely deposit all their energy in the calorimeter, so the $E/p$ for hadrons tends to be smaller than one, but can fluctuate to near one. Finally, anti-protons can annihilate with a proton in the calorimeter causing a larger shower and a large $E/p$.

To combat the fluctuations of $E/p$ for hadrons, four additional variables are combined with $E/p$ in a likelihood. The shower shape is quantified by two variables $E_9/E_{25}$, mentioned earlier, and
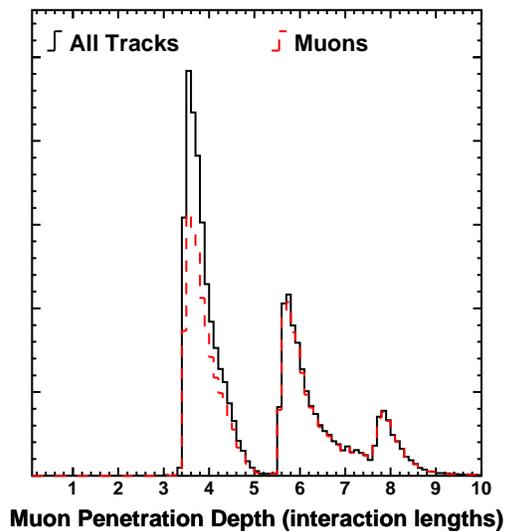
Figure 3.13: Monte Carlo Simulation of Muon Penetration Depth.
The bin at zero, which corresponds to observing no hits in the muon chambers, has been suppressed.
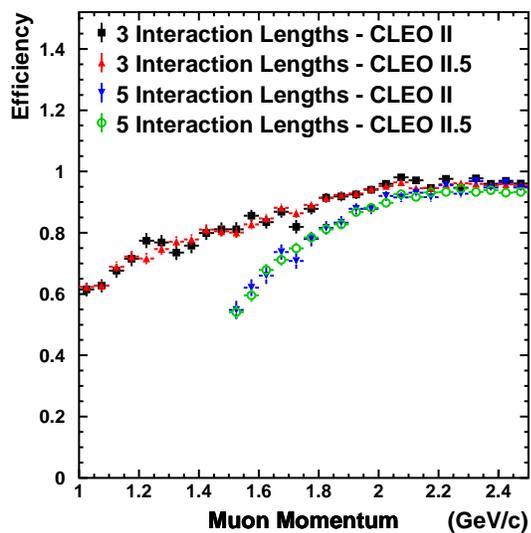


Figure 3.14: Muon Efficiencies Measured in Data.
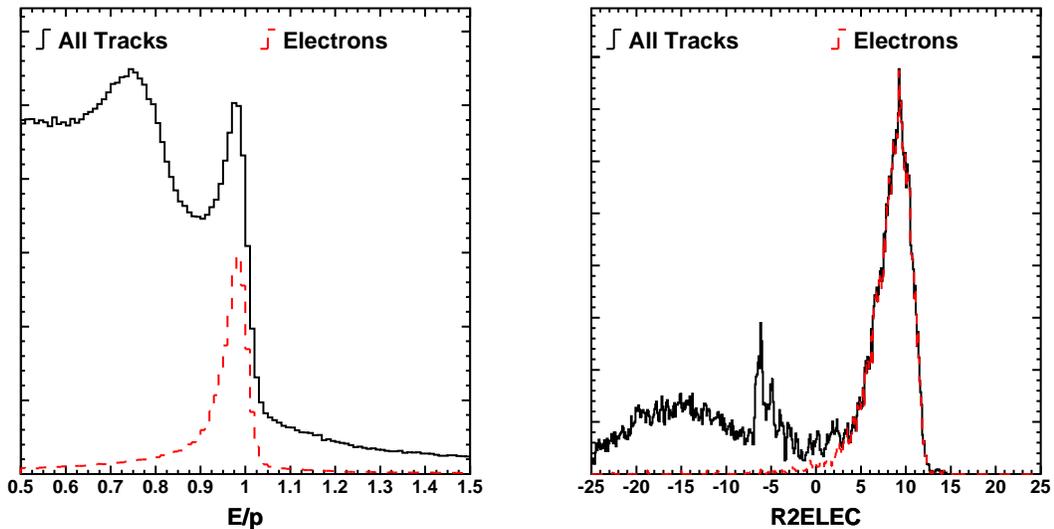
Figure 3.15: Monte Carlo Simulation of Electron ID.
The $E/p$ distribution with no lepton energy cut (left) and the $R2ELEC$ distribution with a 1.2 GeV lepton energy cut (right).

$W = \sqrt{\sum_j E_j R_j^2}/\sqrt{\sum_j E_j}$, where $E_j$ is the energy deposited in a crystal and $R_j$ is the radial distance from the shower center to the crystal. The quality of the track to shower match is also included, because hadronic shower fluctuations cause the shower center to be less correlated with the point of track incidence. Finally the $dE/dx$ variable $\sigma_{dE/dx}^e$ is included. The final variable used is

$$R2ELEC = \sum_{v}^{\#\ \text{Variables}} \ln(P_{e,v}/P_{\not e,v}), \tag{3.5}$$

where $P_{e,v}$ is the probability of measuring the observed value of the input variable, $v$, if the track is an electron, and $P_{\not e,v}$ is the analogous probability if the track is not an electron.

The efficiency of the electron identification is calculated using radiative Bhabha events. This is more complicated than the muon case, because electron identification signatures are in the inner detector, where the shower or the track might significantly overlap with the rest of the event. In order to account for the effect of the event environment, the data from isolated electrons in Bhabha events are added into hadronic events. The data are then reprocessed and the efficiency for finding the embedded electron is tabulated. The resulting measured electron efficiencies are shown in Figure 3.16.
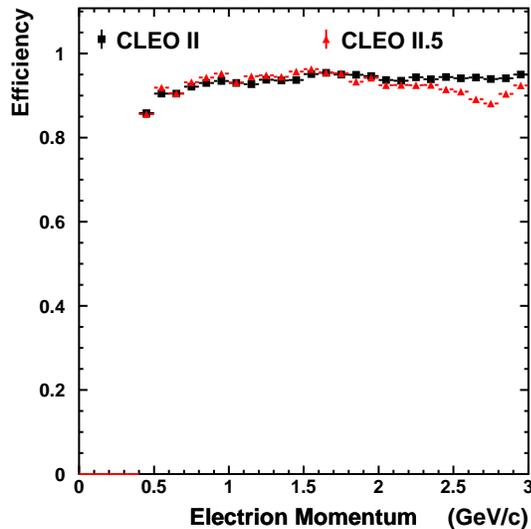
Figure 3.16: Electron Efficiencies Measured in Data.

## 3.2.7  Timing, Trigger, and Data Acquisition

To operate the detector components described above, a variety of electronics is necessary: the timing system, the trigger system, and the data acquisition system.

The timing system synchronizes all the time-related measurements. It sends signals to the wire chambers and time-of-flight system, which are used as the reference times from which the drift and flight times are measured. The clock which controls CESR operations is used to synchronize the signals to the beam crossing time. It is very important that the timing signals are very stable relative to the actual time of the collision. A 1 ns jitter, would make the 150 ps time-of-flight resolution useless, and would increase the wire chamber resolutions by $25-50$ $\mu$m. The stability achieved is 30 ps.

The trigger system is composed of four levels, each of which selects from its input a subset of events which are likely to be of interest to be saved. The first three levels, L0, L1, and L2, are implemented in hardware. The final trigger level, L3, is implemented in software. The L0 trigger gets signals from the intermediate tracking chamber(VD), time-of-flight (TOF), and calorimeter (CC). It selects from the 2.7 MHz beam crossing rate 20 KHz of events. The L1 trigger uses information from VD, TOF, CC, and the main drift chamber (DR) to further reduce the rate to approximately 25-50 Hz (the rate depends on the luminosity and specific trigger criteria which vary over the data taking period). Tracking information in the trigger is processed with "integer tracking" algorithms, which find tracks using look up tables and do not use hit timing information. These trackers can make a crude momentum measurement and determine the charge of the tracks. The L2 trigger makes further use of the VD and DR to eliminate beam-wall, beam-gas, and cosmic-ray events.

After an event is accepted by the L2 trigger, the data are digitized in parallel in 28 electronics crates which are part of the data acquisition system (DAQ). A controller in each crate is responsible for the digitization which takes approximately 3 $\mu$s per channel. The data are sparsified, to remove the channels which are consistent with there being no particles present. The crate controller then notifies a link board that the data are ready to be transfered. The link board stores the data in one of 32 buffers and waits for one of four VME based computers to collect the data (VME is a standard computer bus technology). Once the data are in the VME computers the trigger is reenabled. Another VME processor then uses the VME bus to retrieve this event data and the data from the trigger system and assembles them into a complete event record. The record is then transfered to a UNIX workstation.

The workstation runs the software trigger, L3, on the event, using vertex information to further suppress events that are not from $e^+e^-$ collisions. Bhabha events in which the final state electrons hit the endcaps are so numerous that some fraction of them are thrown out at this point. This is known as prescaling. The remaining events are written to magnetic tape for permanent storage. The trigger efficiency for $B\overline{B}$ events is over 99%. For a few low multiplicity processes, such as $\tau$-pair and two-photon events, the trigger efficiency is reduced. A simple reconstruction, known as PASS1, is also run on the data after the L3 trigger. This generates a set of histograms of detector and physics related quantities used for online data quality monitoring.

## 3.3  Data Reconstruction

After the data are collected, it is processed with a large reconstruction package, which translates the raw quantities measured by the detectors into tracks, showers, and particle identification quantities. This involves a large set of calibration constants which are tuned based on electronic and physics based calibrations. A pattern recognition package groups the drift chamber hits (and silicon hits in CLEO II.5) into tracks which are fit to extract their momenta and direction. The crystal hits are grouped into showers whose energies and positions are calculated using a calibrated algorithm. The tracks are matched to showers, hits in the muon chambers, and hits in the time of flight system. All of the particle identification quantities previously described are calculated in this software package. The event classification package is also run to classify the events as Bhabha, $\mu$ pair, $\tau$ pair, two photon, cosmic ray, junk, and hadrons. The results of the reconstruction software are stored for use by the collaboration. This stored data is the primary data set used for physics analysis.

## 3.4    Monte Carlo

The analysis presented in this thesis relies very heavily on a simulation of the data which uses the Monte Carlo technique. This simulation has two parts: a physics simulation and a detector simulation. The CLEO Monte Carlo simulation program is divided into two parts accordingly, QQ which simulates the physics of the collisions, and CLEOG which simulates the detector response.

Although the branching fractions and form factors of many exclusive $B$ decay modes have been measured, the sum of all measured modes only accounts for a very small fraction of the total inclusive rate. The sum of all measured branching fractions of a $B^0$ meson in the PDG [1] is less than 50%. It is therefore neccesary to use inclusive hadronization models such as JETSET [38] to "make up" the remaining contributions to be consistent with various published and unpublished inclusive measurements. One important measurement in this category is the inclusive charm production in $B$ mesons decays which is not consistent with theoretical predictions [6]. In addition to $B\overline{B}$ events, this analysis uses a JETSET based simulation to model backgrounds from $e^+e^- \rightarrow q\overline{q}$ events.

The detector simulation, CLEOG, is based on the GEANT package [39]. It receives the positions, three-momenta, and particles types of all the particles produced by the simulated collisions which have long enough lives to exit the beam pipe. CLEOG uses a detailed description of the detector material to add bremsstrahlung, multiple scattering, nuclear interactions and photon conversions. The detector noise properties are introduced to the simulation by merging into the simulated events data taken at beam crossings which are randomly selected independent of the trigger criteria described above. The simulated data is then reconstructed with the same software as is used for the real data.

Both the physics and the detector simulation have been tuned extensively using the 20 years of experience operating successive generations of the CLEO detector on the $\Upsilon(4S)$ resonance. Many studies of the inclusive spectra of both $B\overline{B}$ and $q\overline{q}$ have been used to correct the physics models. The detector simulation has also been tuned extensively to match the efficiency and resolution of each of the measurements. The result is a very reliable Monte Carlo, which often agrees with the data at the few percent level or better.

# Chapter 4   Event Selection and Reconstruction

In this analysis, we study the differential decay rate of inclusive semileptonic $B$ decays in the full three-dimensional space discussed in Section 2.1. In order to do this, events containing the signal mode are selected and the neutrino kinematics are reconstructed. In this section, these two aspects of the analysis are described.

## 4.1   Hadronic Event Selection

The accelerator produces $e^+e^-$ collisions at $\sqrt{s} = 10.580$ GeV. At this energy, there are several kinds of events produced, which can be classified as hadronic events and everything else. Hadronic events are $B\overline{B}$ events and the $e^+e^- \to q\overline{q}$ continuum events, where $q = u$, $d$, $s$, and $c$. The other kinds of events that can occur are Bhabha events, $\mu$-pair and $\tau$-pair production, two-photon events, $e^+e^- \to \gamma\gamma$ events, cosmic ray events, and collisions of the beam with residual gas in the beam pipe or with the pipe wall itself.

Hadronic events look very different from these other kinds of events and CLEO has a standard set of criteria which are used to identify them. The events are required to have at least three tracks (because we later require the sum of charges to be zero, this is effectively four). This strongly suppresses Bhabha, $\mu$-pair, $e^+e^- \to \gamma\gamma$ events, and cosmic ray events. If the number of tracks is three or four, the sum of the energy in the calorimeter is required to be less than 0.65 of the center of mass energy, $E_{cm}$. This uses the fact that electrons tend to deposit all of their energy in the calorimeter to further suppresses Bhabha events. The sum of the visible energy (track momenta and showers not matched to tracks) is required to be greater than 0.2 $E_{cm}$. This suppresses beam-wall, beam-gas, and two photon events, in which the beam particle (or particles) in the final state carry away much of the energy.

The only nonhadronic process that has a non-negligible contribution after this selection is $\tau$-pair production. These events can be moderately high multiplicity and deposit sufficient energy to pass the criteria. The selections used to improve the resolution on the neutrino kinematics significantly suppress this process, because $\tau$-pair events which have a lepton in the final state also have at least three neutrinos in the final state.

## 4.2   Lepton Identification

The available lepton identification information is described in Sections 3.2.5 and 3.2.6. We only consider signal leptons which have a momentum of at least 1.0 GeV/$c$. The minimum lepton energy considered is set by the rapid increase in the backgrounds below 1.0 GeV/$c$. The importance of accepting a large kinematic range in order to compare to the HQET predictions is discussed in Section 2.4.3.

The leptons are also required to fall within the barrel region of the detector ($|\cos\theta| < 0.71$). As this is a systematics limited analysis, the fiducial angular acceptance is determined by the availability of the measured lepton fake rates and efficiencies; there is no need to use regions that are less well modeled to increase the reconstruction efficiency. Electrons are required to have an $R2ELEC$ of greater than $3.0$[1]. For momenta between 1.0 and 1.5 GeV/$c$, muon candidates are required to penetrate at least 3 interaction lengths and above 1.5 GeV/$c$, candidates are required to penetrate at least 5 interaction lengths.

The candidate lepton is also required to pass some track quality constraints. The track is required to be consistent with the track coming from the IP: $r - \phi$ and $z$ impact parameters less than 2 mm and 5 cm respectively. The root mean square of the hit residual is required to be small, less than 0.5 mm.

## 4.3   Neutrino Reconstruction

Neutrinos are reconstructed using the approximate hermiticity of the CLEO detector. The sum of the four-vectors of all observed particles, $p^{\mu}_{visible}$, is subtracted from the well-known $e^+e^-$ initial state produced by the accelerator,

$$p^{\mu}_{\nu} = p^{\mu}_{e^+e^-} - p^{\mu}_{\text{visible}}, \qquad (4.1)$$

where $p^{\mu}_{e^+e^-} \approx (2E_{\text{beam}}, \vec{0})$. In order to achieve good resolution on the neutrino four-vector, an estimate of the set of visible particles and their properties must be made as accurately as possible. The packages described in Sections 3.2.1 and 3.2.3 are used to suppress fake tracks and showers respectively. In addition, only showers with energies above 25 MeV are used, because showers below this energy are predominantly not from photons (there are a lot of hadronic splitoff showers and very few photons below 25 MeV).

There are a few kinds of neutral particles, $K^0_S$ mesons and $\Lambda^0$ baryons, which travel several centimeters before decaying in the tracking chambers. The charged daughters of these particles meet at vertices which are measurably separated from the interaction point. In order to get the right

---

[1] see Section 3.2.6 for the definition of R2ELEC

contribution to the energy-momentum sum, the momenta of the daughter particles are evaluated at the separated vertex, not at the interaction point (where the momenta are usually evaluated). If this is not done, the curvature of the particles in the magnetic field will cause a shift in the direction of the corresponding track.

For all the tracks used in the neutrino four-vector calculation, a particle identification is needed in order to calculate the energy from the measured momentum. The particle identification is limited to $\pi$ or $K$. The proton hypothesis is ignored because protons are rare and correlated with either a anti-neutron or an anti-proton, so the chance of correctly reconstructing the event is very low. Similarly events with extra leptons, in addition to the signal lepton, are vetoed because of their correlation with extra neutrinos (see below). If they pass the veto they will be badly mismeasured anyway, and the particle misidentification will be a small effect.

The particle identification is done using both the $dE/dx$ and time-of-flight information. The $dE/dx$ is available on almost every track, but the time-of-flight efficiency is somewhat lower. Each variable is added, when available, to a total $\chi^2$ variable which is in turn used to calculate a probability, $P_x$, of the observed particle identification measurements for each hypothesis, $x = \pi$ or $K$. The particle type is then assigned to be a kaon if $R_{\pi K} P_K > P_\pi$ and otherwise it is assigned to be a pion. The value of $R_{\pi K}$ has been selected using a minimization of the RMS of the neutrino energy resolution using Monte Carlo simulated data. The resulting value, $R_{\pi K} = 0.25$, is larger than the value that would be expected from the ratio of kaons to pions in a typical $B\overline{B}$ events, $N_K/N_\pi \approx 0.1$. There are a few possible reasons for this. The momentum distributions of the two samples are not the same and mislabeling a low-momentum track creates a larger shift in the energy than mislabeling a high momentum track. The distributions of pions and kaons in an event are correlated. Most $B\overline{B}$ events have two kaons including both neutral and charged kaons. Finally, this optimization is made after the full reconstruction process, so events with $K_L^0$ mesons may be suppressed enhancing the $K^+/\pi^+$ ratio.

## 4.4  Selecting Events with Well-Reconstructed Neutrinos

Several effects can cause the inferred neutrino energy and momentum to be a poor estimator of the actual neutrino kinematics:

- Particles can be lost because the detector only covers 95% of $4\pi$ and it is not 100% efficient;

- Tracks and showers beyond those which correspond to the charged particles and photons in the event may be reconstructed;

- The long-lived neutral hadrons, $K_L^0$ mesons and neutrons, are not well reconstructed because they do not produce showers regularly, nor are these showers easily separated from the splitoff
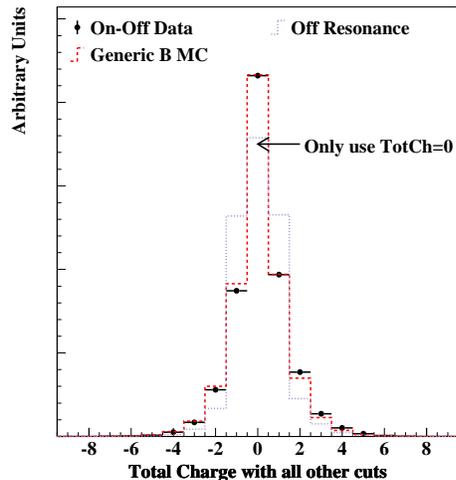
Figure 4.1: Distribution of the Total Charge.
The total charge distribution is shown for data where the off-resonance data has been used to subtract the non-$B\overline{B}$ contribution (points), the CLEO $B\overline{B}$ Monte Carlo simulation (long dash), and off-resonance data (short dash).

showers discussed in Section 3.2.3, nor do they reliably estimate the $K_L^0$ energy;

- There may be more neutrinos in the event that the one of interest;

- The particle identification, needed to get the energy part of the track four-vectors, is good but not completely effective;

- The energy and momentum measurements of the individual tracks and showers have finite resolutions, but this is a very small effect compared to the other reconstruction problems.

There are a few characteristics which can be used to reduce the effects of these reconstruction problems and correspondingly enhance the resolution on the neutrino four-momentum. The sum of the charges of the tracks in the event is required to be zero to suppress the effect of lost or fake tracks. Figure 4.1 shows the distribution of the total charge, before the cut. Because additional leptons tend to be correlated with additional neutrinos, events with additional identified leptons are excluded. Within the reconstruction resolution the measured neutrino should be massless.

The criteria for the "extra lepton veto" are looser than those for the signal lepton. If a track has a momentum of greater than 1.0 GeV and has any muon chamber hits associated with it, it is considered a possible muon and the event is vetoed. If a track has a momentum of greater than 0.4 GeV and a value of R2ELEC (see Section 3.2.6) greater than zero, it is considered a possible electron and the event is vetoed. The bulk of the extra leptons come from decays of the other $B$
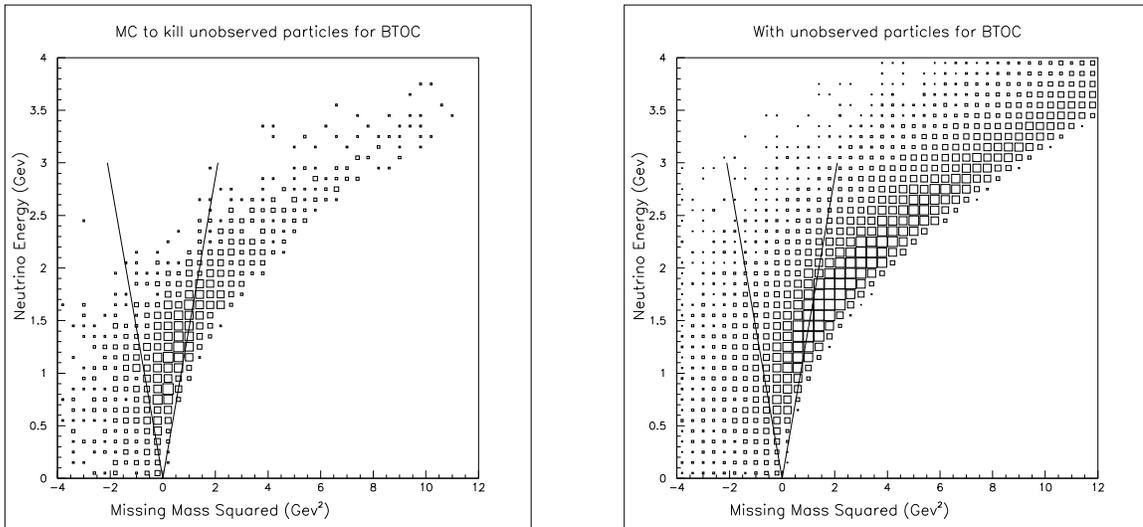
Figure 4.2: Reconstructed Neutrino Energy versus Missing Mass.
Both plots show reconstructed neutrino energy versus missing mass, as predicted by Monte Carlo simulation. On the left are events that have been selected using information about what was generated in the Monte Carlo event to have no missing particles ($K_L^0$s, extra $\nu$s, neutrons, and particles down the beam pipe). On the right are the events with missing particles, corresponding to poor resolution on the neutrino kinematics.

meson in the event or $D$ mesons. The only correlation the extra leptons might have with the inclusive semileptonic $B$ differential decay rate is through correlations with $D$ momenta in $B \to X_c l \overline{\nu}$ decays. This effect is small and should be correctly modeled by the Monte Carlo simulation.

Instead of using the missing mass squared, $M_{miss}^2 = p_\nu^2$, directly to implement the neutrino mass requirement, a variable, $V_{cut}$, is used which is proportional to the energy of a missing particle:

$$
\begin{aligned}
M_{miss}^2 &= (p_\nu^{true} + p_{lost})^2 \\
&= 2E_\nu^{true}(E_{lost} - |\overrightarrow{p}_{lost}|\cos\theta_{\nu-lost}) + M_{lost}^2 \\
V_{cut} &\equiv \frac{M_{miss}^2}{2E_\nu^{reco}} \propto |\overrightarrow{p}_{lost}|.
\end{aligned}
\tag{4.2}
$$

Figure 4.2 shows a Monte Carlo simulation of the distribution of reconstructed neutrino energy versus missing mass squared for events which have no missing particles and for events with missing particles. Missing particles are defined as $K_L^0$ mesons, neutrons, extra neutrinos, and particles not within the angular acceptance of the detector. The distribution for events with missing particles shows the correlation between the measured neutrino energy and the measured missing mass. In both histograms, the $V_{cut}$ requirement used is shown.

Because of data handling issues, two versions of $V_{cut}$ are used. One with all the particles treated as pions is called $V_{cut}^{NoPID}$, and one with the particle identification procedure defined above is called
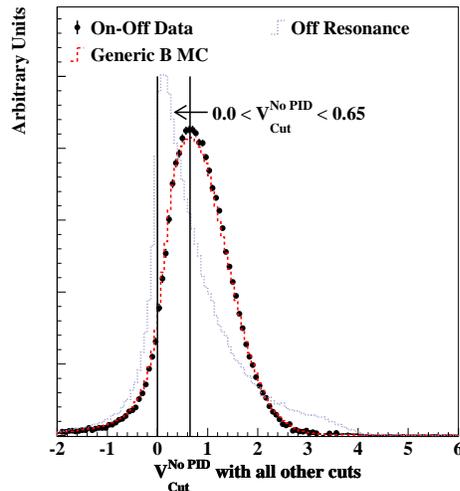
Figure 4.3: Distribution of $V_{cut}^{\mathrm{NoPID}}$.
The distribution of $V_{cut}^{\mathrm{NoPID}}$ is shown for data where the off-resonance data has been used to subtract the non-$B\overline{B}$ contribution (points), the CLEO $B\overline{B}$ Monte Carlo simulation (long dash), and off-resonance data (short dash).

$V_{cut}{}^2$. We select events for which $0 < V_{cut}^{\mathrm{NoPID}} < 0.65$ GeV and $|V_{cut}| < 0.35$ GeV. The requirement on $V_{cut}$ is much more restrictive and 99% of the events which pass the $V_{cut}$ requirement satisfy the $V_{cut}^{\mathrm{NoPID}}$ requirement. In Figure 4.3, the distribution of $V_{cut}^{\mathrm{NoPID}}$ in the data (with the non-$B\overline{B}$ contribution subtracted using off-resonance data) is compared to the Monte Carlo simulation for $B\overline{B}$ events. There is good agreement between the data and the Monte Carlo simulation for the $V_{cut}$ quantity, which is closely related to the expected neutrino momentum resolution.

Once the event selection is made, the four-vector of the neutrino is constrained to be massless by setting the energy of the reconstructed neutrino to be the magnitude of the reconstructed momentum vector. The momentum is considerably better measured than the energy, because it does not rely on particle identification. The resulting neutrino energy resolution is shown in Figure 4.4.

## 4.5 Continuum Suppression

Events from the $e^+e^- \rightarrow q\overline{q}$ continuum are a major source of background in this analysis. To suppress this contribution we make use of differences in the event topology between $B\overline{B}$ and continuum events. In continuum events, the particles in the final state tend to be aligned with the directions of the two

---

[2] This is necessary because the Monte Carl simulation uses random numbers to simulate the particle identification variables. In the analysis procedure, a copy of subset of the data (a skim) is made using loose event selection criteria. For the Monte Carlo data, this loose selection cannot use particle identification because the random numbers used in initial selection will not be correlated with those used when the data is processed a second time.
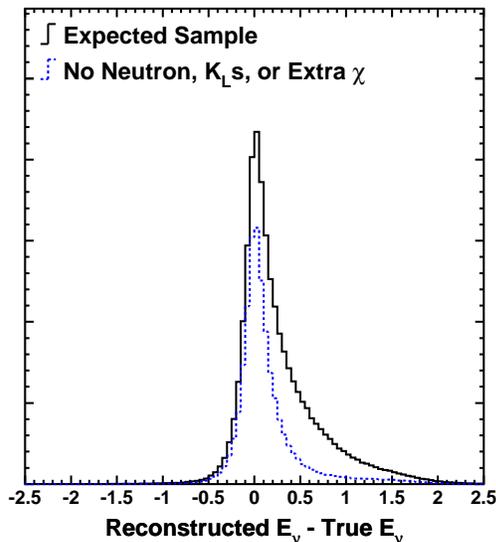
Figure 4.4: Distribution of Neutrino Energy Resolution.
The solid line is the Monte Carlo simulation of neutrino resolution and the dotted line is the same simulation with the events which have $K_L$ mesons, neutrons or extra neutrinos removed.

quarks produced forming two "jets." Because the quarks are produced with a $1 + \cos^2 \theta$ distribution with respect to the beam, the jets also tend be aligned with the beam. In $B\overline{B}$ events, the $B$ mesons are produced with a $\sin^2 \theta$ distribution, but this distribution is not evident because of the very small velocity of the $B$ mesons in the lab. Because $B$ mesons are pseudoscalars, they have no preferred direction and their decays are oriented randomly in the lab. The resulting features we exploit are that continuum events are back to back jets aligned with the beam, and $B\overline{B}$ events are spherical and randomly oriented in the detector.

The first selection we use is based on the normalized second Fox-Wolfram moment [40] $R_2$ which is defined as

$$R_2 = \frac{\sum_i \sum_j |\vec{p}_i| \, |\vec{p}_j| \, P_2(\cos \theta_{ij})}{\sum_i \sum_j |\vec{p}_i| \, |\vec{p}_j|}, \tag{4.3}$$

where the vectors $\vec{p}_i$ are the three-vectors of the observed tracks and unmatched showers, $\theta_{ij}$ is the angle between $\vec{p}_i$ and $\vec{p}_j$, and $P_2(x) = \frac{1}{2}(3x^2 - 1)$ is the second Legendre polynomial. Two particles in a continuum event are either in the same jet and hence aligned ($\cos \theta_{ij} \approx 1$) or in opposite jets with $\cos \theta_{ij} \approx -1$. For $B\overline{B}$ events the $\cos \theta_{ij}$ distribution is much flatter and $R_2$ tends to be lower than for continuum events, as shown in Figure 4.5. Requiring $R_2 < 0.4$ is more than 99% and 95% efficient for $B \to X_c l\overline{\nu}$ and $B \to X_u l\overline{\nu}$, respectively, while removing 60% of the continuum events.

Because continuum events are both jetty and the jets tend to point along the beam axis, continuum events tend to lose more particles down the beam pipe than do $B\overline{B}$ events. This can cause
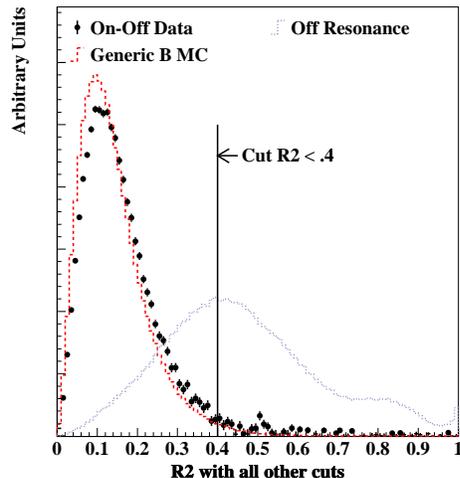
Figure 4.5: Distribution of $R_2$.
The distribution of $R_2$ is shown for data where the off-resonance data has been used to subtract the non-$B\overline{B}$ contribution (points), the CLEO $B\overline{B}$ Monte Carlo simulation (long dash), and off-resonance data (short dash).

the reconstructed neutrino momentum in continuum events to point along the beam axis, as shown in Figure 4.6. We require the cosine of the angle of the reconstructed neutrino momentum with the beam axis to be less than 0.95, $|\cos\theta_{\nu\cdot beam}| \leq 0.95$.

In analyses in which the $B$ mesons are fully reconstructed, the correlation between the orientation of the remaining particles and the candidate is often exploited. In a $B\overline{B}$ event, there will be no correlation, because the remaining particles make up the other $B$ meson in the event and the orientations of the two $B$ are not correlated, because the $B$ is a pseudoscalar. In a continuum event, the remaining particles will be elements of the same two jets from which the candidate was reconstructed and their orientation will be correlated with that of the candidate. The variable generally used is the angle between the thrust axis, $\vec{T}$, and some axis of the candidate. The thrust axis, $\vec{T}$, is the axis which maximizes

$$\text{Thrust Magnitude} = \sum_{i \in \{remaining\ particles\}} \vec{T} \cdot \vec{p_i}, \tag{4.4}$$

where $p_i$ are the three-momenta of the remaining particles.

In this analysis, we do not fully reconstruct the $B$ meson. This correlation can still be used because the $B$ mesons are almost at rest in the lab and the $B$ mass is well-known. The three vector of the unreconstructed hadronic part of the semileptonic $B$ candidate can be approximated by assuming a $B$ at rest and assigning the hadronic three vector to be $-\vec{q} = -(\vec{p_\ell} + \vec{p_\nu})$. The thrust
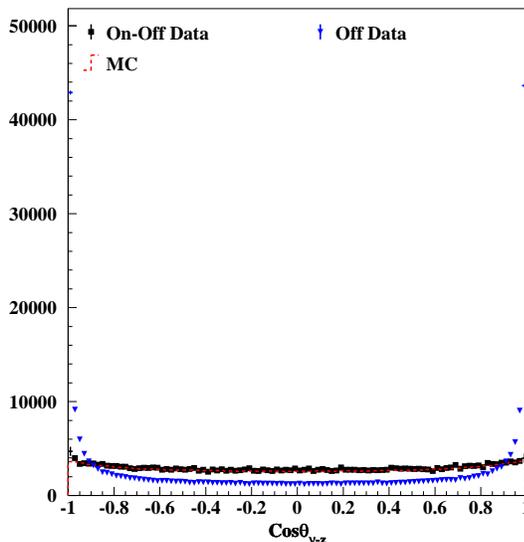
Figure 4.6: Distribution of $\cos\theta_{\nu\cdot beam}$.
The distribution of the cosine of the angle of the reconstructed neutrino momentum with the beam axis is shown for data where the off-resonance data has been used to subtract the non-$B\overline{B}$ contribution (squares), the CLEO $B\overline{B}$ Monte Carlo simulation (dashes), and off-resonance data (triangles).

of the other $B$ is then defined to be the axis, $\vec{T}$, which maximizes this quantity:

$$\text{Thrust Magnitude} \quad = \quad \left( \sum_{\substack{i\in\{showers+ \\ tracks\neq lepton\}}} \vec{T}\cdot\vec{p_i} \right) \quad - \quad \underbrace{\vec{T}\cdot(-\vec{q})}_{\substack{\text{approximate removal of the} \\ \text{hadronic part of the signal } B}} \qquad (4.5)$$

The second term should approximately cancel the contribution to the sum from the unreconstructed part of the signal $B$. This is exactly true for the case where the hadronic system does not decay (e.g., $B \to \pi l\overline{\nu}$). In the case where the hadronic system decays the hadronic daughters spread out perpendicular to the $\vec{q}$ axis. This biases the *Thrust Magnitude* to be maximized when $\vec{T}$ is perpendicular to the $\vec{q}$ axis. So the distribution of $\cos\theta_{T\cdot q}$ for $B$ decays will be flat with a slight rise for $\vec{T}$ perpendicular to $\vec{q}$ and continuum will be peaked for $\vec{T}$ parallel to $\vec{q}$, as shown in Figure 4.7.

This thrust variable is combined with other information into one variable using a neural network that is implemented with the JETNET package [41]. Each of the variables used in the neural network provides a small separation between $B\overline{B}$ and continuum events, but the information in the different variables is complementary. The neural network combines the variables into one variable. It is an effective method of optimizing a multidimensional cut to separate signal and background. The

variables used in the neural network are

- $R_2$,

- the angle between the $\vec{q}$ axis and the Z-axis, which exploits the correlation between the jet direction and the beam axis in continuum events,

- the angle between the $\vec{q}$ axis and the thrust axis, $\vec{T}$, and

- the fraction of the total energy lying in nine separate cones around the lepton direction.

The cones, sometimes called *virtual calorimeter* cones, are defined as

$$\frac{\sum_{i\in\{\text{particles in cone \#n}\}} |\vec{p_i}|}{\sum_{i\in\{\text{all particles}\}} |\vec{p_i}|}, \tag{4.6}$$

where the reconstructed neutrino is included. The cones contain information about the distribution of energy relative to the lepton direction. In continuum event, where the lepton is inside a jet, the energy will tend to be along the lepton direction, where as for a $B\overline{B}$ event the energy will be more evenly distributed. This information is similar to that used in the $R_2$ variable, but more finely segmented. The rescaling helps to reduce the sensitivity of the neural network to the kinematic variables used in the fit ($q^2$, $E_\ell$, and $E_\nu$) and thus reduces the possible bias. The neural net has 12 input nodes, 1 hidden layer with 12 nodes and an output node. It is trained with the off-resonance data, Monte Carlo simulated $B \to X_c l\overline{\nu}$ and $B \to X_u l\overline{\nu}$ events[3]. The full set of input variable are histogrammed in 4.7. The resulting separation is shown in Figure 4.8, along with the effect of the neural net cut on the efficiency as a function of various kinematic quantities. The neural net rejects 73% of the continuum, while rejecting only 6% of the $B \to X_u l\overline{\nu}$ signal and 8% of the $B \to X_c l\overline{\nu}$ signal.

## 4.6 $\psi$, Dalitz Decay, and $\gamma$ Conversion Veto Studies

It is common in analyses involving leptons to make special vetoes for $\psi$ decays, Dalitz decays, and $\gamma$ conversions, however this is not necessary in this analysis. This is largely because of the extra lepton veto (in all three cases the signal lepton is accompanied by a second lepton). Vetoes on these processes are usually made by removing events for which the invariant mass of the signal lepton and any other track in the event, which passes some loose lepton identification requirements, is in the range populated by the undesired process. Figure 4.9 shows the distribution of two track invariant mass between the signal track and any other opposite sign track before and after the extra lepton veto is applied (the plot is made with data). Most of the $\psi$ events are removed by the extra lepton veto.

---

[3] The off-resonance data is not used in the final fit so statistical correlations do not produce a bias, see Section 5.2.6.

69
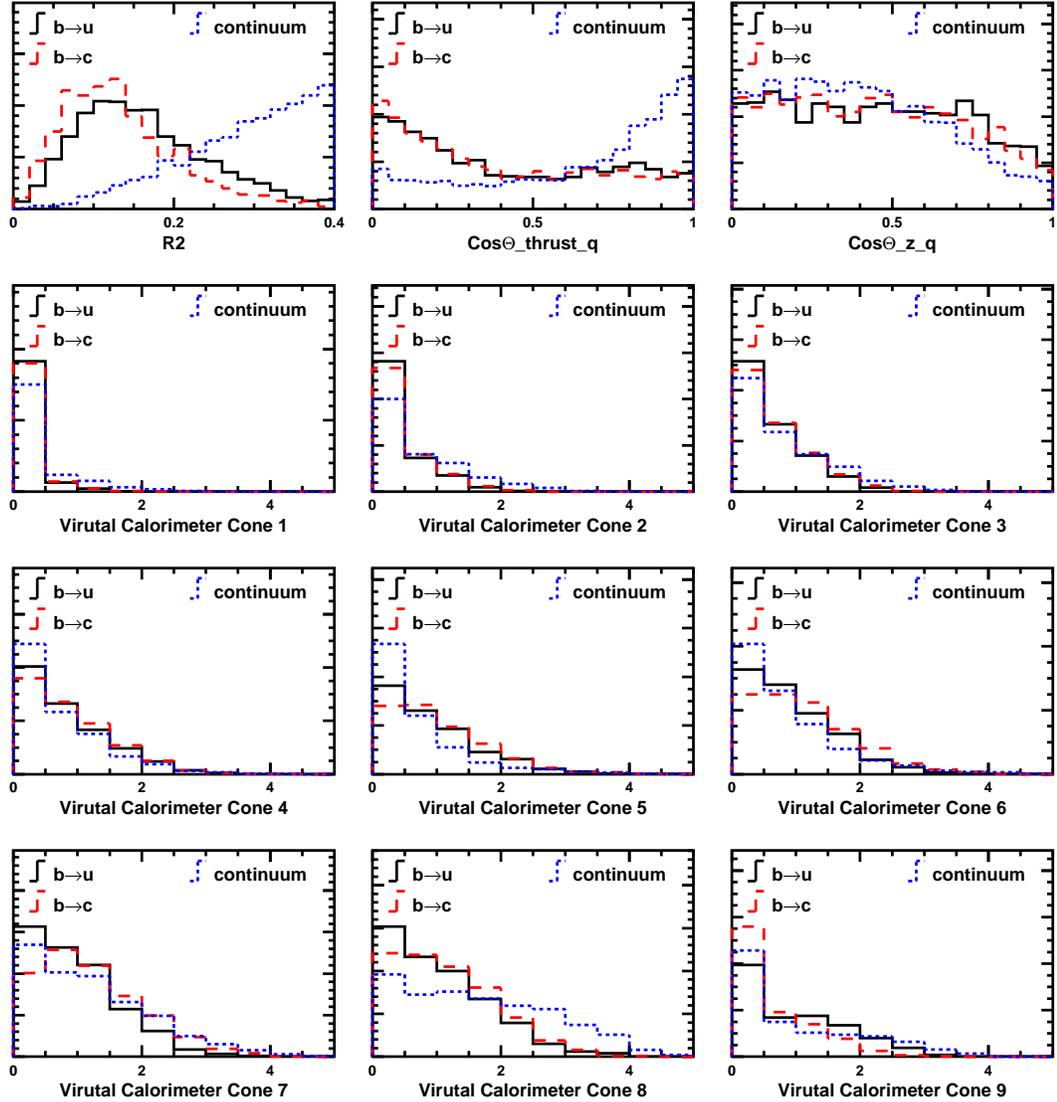


Figure 4.7: Neural Net Input Variables.
The variables are the $R_2$, the angle between the $\vec{q}$ axis and the thrust axis, the angle between the $\vec{q}$ axis and z axis, and the 9 *virtual calorimeter* cones. The $B \to X_u l\overline{\nu}$ (labeled $b \to u$) and $B \to X_c l\overline{\nu}$ (labeled $b \to c$) histograms are Monte Carlo simulated data and the continuum histogram is off-resonance data.
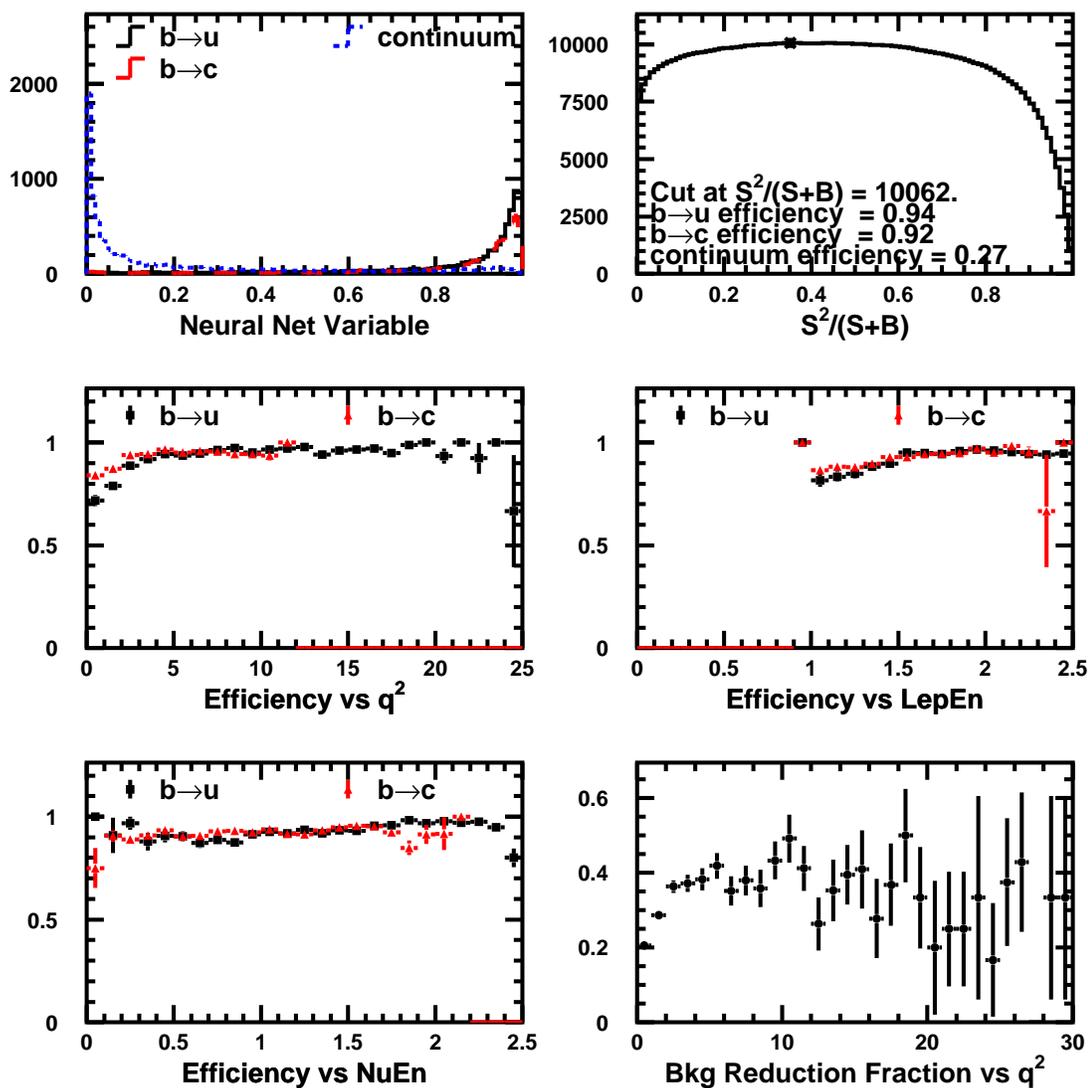
Figure 4.8: Neural Net Performance.
From the top left: the distribution of the neural network output, NN; the signal squared over signal+background optimization of the NN requirement; the efficiency as a function of a complimentary set of three kinematic variables; and the background reduction fraction as a function of $q^2$.
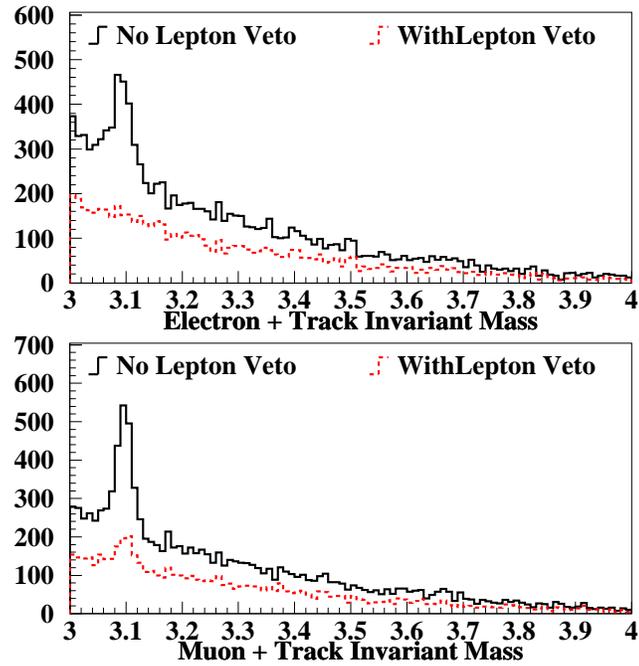
Figure 4.9: Track-Lepton Invariant Masses Distribution in $\psi$ Region.
Data histograms of the invariant mass of combinations of any track with the signal lepton for electrons (top) and from muons (bottom) both with and without the extra lepton veto.

Table 4.1 shows the Monte Carlo estimate of the composition of the secondary lepton background. It is dominated by charm with a small contribution from $\psi$ decays and very little Dalitz decays and gamma conversions. Based on this it is decided that vetoing these backgrounds is unnecessary since they are small and should be moderately well modeled by Monte Carlo simulation.

Table 4.1: Secondary Lepton Background Composition.
This table shows the fractional contribution of various processes to secondary lepton background as predicted by the $B\overline{B}$ Monte Carlo simulation. Each lepton type is separately normalized.

| Parent of Lepton | Electrons | Muons |
|---|---|---|
| $\gamma$ Conversion | 0.00 | 0.00 |
| Decay In Flight | 0.00 | 0.01 |
| $\tau$ | 0.05 | 0.04 |
| $D^0$ | 0.44 | 0.39 |
| $D^{\pm}$ | 0.41 | 0.38 |
| $D_s$ | 0.06 | 0.10 |
| $\Lambda_c$ | 0.00 | 0.00 |
| $\pi^0$ Dalitz | 0.01 | 0.00 |
| $\eta$ Dalitz | 0.00 | 0.00 |
| $\psi$ | 0.03 | 0.07 |
| $\psi'$ | 0.00 | 0.00 |
| total | 1.00 | 1.00 |
| non-open charm | 0.10 | 0.12 |

# Chapter 5   Extraction of Exclusive Branching Fractions from the Inclusive Distribution

We use a maximum likelihood fit to extract the contributions to the selected data set from the various exclusive semileptonic $B$ decay modes. These modes are $B \to Dl\overline{\nu}$, $B \to D^*l\overline{\nu}$, $B \to D^{**}l\overline{\nu}$, nonresonant $B \to X_c l\overline{\nu}$, and $B \to X_u l\overline{\nu}$. The data set also has significant backgrounds which must be modeled in the fit. The backgrounds are classified as secondary leptons, continuum leptons, and fake leptons. A secondary lepton is a real lepton in a $B\overline{B}$ event whose parent is not a $B$ meson. A continuum lepton is a real lepton in a $e^+e^- \to q\overline{q}$ continuum event. A fake lepton is a non-leptonic track from either a $B\overline{B}$ or a continuum event which is identified as a lepton. The only other possible sources of backgrounds are Bhabha events, $\mu$ pair events, $\gamma\gamma$ events, and $\tau$ pair events, none of which contribute significantly to the data sample (see Section 4.1). Table 5.1 shows a summary of the composition of the data sample and the models used for each component. The branching fractions for each of the semileptonic $B$ decay modes are determined from the fit results. After all cuts, 41411 events are observed in the CLEO II data set and 80440 events in the CLEO II.V data set. The overall efficiency varies from 1.5% for $B \to X_c l\overline{\nu}$ nonresonant to 4.2% for $B \to X_u l\overline{\nu}$.

Table 5.1: Composition of the Data Sample and Summary of the Models Used in the Fit.

| Mode | Fraction of Data Sample | Model |
|---|---|---|
| $B \to Dl\overline{\nu}$ | 0.118 | HQET |
| $B \to D^*l\overline{\nu}$ | 0.476 | HQET |
| $B \to D^{**}l\overline{\nu}$ | 0.084 | ISGW2 |
| Nonresonant $B \to X_c l\overline{\nu}$ | 0.033 | Goity and Roberts |
| $B \to X_u l\overline{\nu}$ | 0.016 | ISGW2 |
| Secondary Leptonss | 0.050 | CLEO $B$ decay model & measurements of semileptonic charm hadron decay |
| Fake Leptons | 0.132 | Data & measured lepton fake rates |
|     Electron | 0.002 | |
|     Depth Mu 3 | 0.100 | |
|     Depth Mu 5 | 0.030 | |
| Continuum Leptons | 0.089 | JETSET |

## 5.1   The Maximum Likelihood Technique

The fit used is a binned maximum likelihood fit in three-dimensions. This choice is motivated by the complicated structure of the efficiency and the reconstruction resolution. For this analysis, these

features can only be modeled by Monte Carlo simulation. An unbinned fit uses all the information available in the data, but requires a model of the probability at each data point which depends on the physics model, the efficiency, and the resolution (see Section 5.2). To use an unbinned fit for this analysis, a procedure would have to be defined to turn the Monte Carlo events into probabilities at each data point. The simplest solution is to group the events into bins and use the distribution as probability; this is exactly equivalent to a binned fit. More complicated procedures could be used to smooth the Monte Carlo simulated distributions, but this can add biases and is not necessary for this analysis. Binning involves throwing out the information about the values of variables within each bin, but with a good choice of binning, this is not a major loss. To this end, the variables $\cos\theta_{W\ell}$, $M_X^2$, and $q^2/q_0^2$ are chosen. The choice of $q^2/q_0^2$ has the same effect as making the size of the bins in $q^2$ become smaller as $q_0$ becomes smaller. This preserves information which would be lost if a coarse $q^2$ binning is used, and also ensures a reasonable population of Monte Carlo events in the large $q^2$ region, which would be a problem if a fine $q^2$ binning is used. A further advantage of both $q^2/q_0^2$ and $\cos\theta_{W\ell}$ is that the phase space boundaries in each of these dimensions do not depend on the other variables, so there will not be many partially populated bins at the these boundaries. For each dimension 15 bins are used, which is sufficient to resolve the important features of the distributions.

The background from fake leptons is much larger for muons than for electrons. The improved signal to background of the electron sample would be lost if the two samples where simply combined into one set of histograms when fitting. A fourth dimension is therefore added to the fit which has two bins, one for electrons and one for muons. This brings the total number of bins to $15 \times 15 \times 15 \times 2 = 6750$.

### 5.1.1 The Likelihood

The component histograms used in the fit are constructed with weighted Monte Carlo or data events. Because of the large number of bins, it is impossible to generate enough Monte Carlo events to be able perform the fit in the limit of infinite histogram statistics. Instead a method suggested by Barlow and Beeston [42] is used to include the statistics of the component histograms in the fit. This technique adds terms to the likelihood to include the probability of the component histograms' contents given unknown expectations. The unknown expectations are new parameters which are determined by the maximization of the likelihood. It is procedurally easier to work with the log-likelihood than with the likelihood. Maximizing the log-likelihood also maximizes the likelihood. The log-likelihood used is

$$\ln\mathcal{L} = \sum_{\text{bins}} \left( d_b \ln f_b - f_b \right) + \sum_{\text{modes}} \sum_{\text{bins}} \left( a_{mb} \ln A_{mb} - A_{mb} \right) \qquad (5.1)$$

where

$d_b$ is the number data events in bin $b$,

$a_{mb}$ is the number of events from mode $m$ in bin $b$,

$A_{mb}$ is the expected number of events from mode $m$ in bin $b$,

$f_b = \mathcal{N} \sum_m p_m w_{mb} A_{mb}$ is the expected number of data events in bin $b$,

$w_{mb}$ is the weight of events from mode $m$ in bin $b$ ,

$p_m$ is the fractional contribution of mode $m$ to the total yield, and

$\mathcal{N} = \sum_b d_b$ is the data yield.

The fit parameters, $A_{mb}$ and $p_m$, are determined by the maximization of the likelihood. The weights $w_{mb}$ allow the Monte Carlo simulated events used to model the individual modes to be modified to reflect corrections to simulation (see Section 5.2.1). The weights are also used to vary the simulation in the study of the systematic errors (see Chapter 6).

Both terms of the likelihood (Equation 5.1) are constructed from the Poisson probability for observing $\mu$ events when $\lambda$ events are expected,

$$\mathcal{P} = e^{-\lambda} \frac{\lambda^\mu}{\mu!}, \quad \longrightarrow \quad \ln \mathcal{P} = \mu \ln \lambda - \lambda - \ln \mu! \,. \tag{5.2}$$

The first term in the likelihood is the log of the Poisson probability of the observed number of data events in each bin, $d_b$, given the expectation, $f_b$, which depends on the fit parameters, $A_{mb}$ and $p_m$. The second term describes the Poisson probability of the data sets used to fill the component histograms, for which the observed is $a_{mb}$ and the expected is $A_{mb}$. The two terms are coupled by the appearance of $A_{mb}$ in $f_b$. The $\ln \mu!$ is dropped from the likelihood definition because it does not depend on the fit parameter, so it is just an additive constant and does not effect the location of the maximum of the likelihood.

There are two normalization conditions that need to be satisfied. The first is that $\sum_m p_m = 1$, so that the $p_m$ are fractional contributions of each mode to the total data sample. The other condition is that

$$\sum_b w_{mb} A_{mb} = 1. \tag{5.3}$$

This makes $w_{mb} A_{mb}$ the fraction of the events from mode $m$ in bin $b$. With these two conditions, $f_b$ is then a properly normalized prediction for the number of events in bin $b$ and $\sum_b f_b = \mathcal{N}$. Barlow and Beeston show that it is a property of the minimization that $\sum_b w_{mb} a_{mb} = \sum_b w_{mb} A_{mb}$, so if the inputs $w_{mb}$ and $a_{mb}$ satisfy $\sum_b w_{mb} a_{mb} = 1$, then the condition in Equation 5.3 will be satisfied.

To understand the structure of this likelihood, it is useful to consider the limit of a very large Monte Carlo sample and a significantly smaller data sample. The second term will then strongly constrain the expected number of Monte Carlo events, $A_{mb}$, to be close to the observed number of Monte Carlo events $a_{mb}$. Furthermore the $A_{mb}$ would not be significantly affected by the properties of the data. The $A_{mb}$ in the first term would then be equal to $a_{mb}$ and the first term would become exactly the commonly used binned likelihood which assumes infinite statistics.

## 5.1.2   Fit Implementation

The number of parameters which are determined by maximizing this likelihood is drastically larger than for a traditional binned likelihood. For a fit with $M$ modes and $B$ bins there are $M$ $p_m$'s and $M \times B$ $A_{mb}$'s, which makes $M \times (B + 1)$ parameters to be determined by maximization of the likelihood. For this analysis, the number of fit parameters is increased from $M = 10$ for a likelihood which assumes infinite Monte Carlo statistics to $10 \times (6750 + 1) = 67510$ for the likelihood described. Barlow and Beeston show that given a set of $p_m$'s, the minimization of the $A_{mb}$'s can be analytically simplified into the problem of finding the zero of one function for each bin [42]. In this analysis the remaining ten parameters, $p_m$ are minimized using the MINUIT package [43]. Except for cross-checks, the overall normalization is fixed so that $\mathcal{N}$ is the data yield. As a cross-check the $\mathcal{N}$ is determined by maximization of the likelihood. The shifts in the results were extremely small, as should be expected for the large number of events. The normalizations of the continuum background and the three categories of fake lepton background are also fixed in the fit, because they can be determined from the off-resonance data and the measured lepton fake rates (see Sections 5.2.6 and 5.2.7, respectively). This leaves only five parameters to be minimized by MINUIT, but the full $10 \times 6750$ $A_{mb}$ parameters must be minimized regardless.

The fitting package is tested with a simple Monte Carlo program. The program generates a new set of component histograms in which each bin is Poisson distributed assuming the original component histogram's contents as the mean. A similar procedure is used to generate a new data distribution using a sum of the original component histograms for the mean values. The result is a new set of data and mode histograms which have the same statistical properties as the true data and mode histograms. These are then fit twice, once using the Barlow and Beeston method and once assuming the infinite statistics limit for comparison. These resulting pull distributions for the six parameters determined by the fit are shown in Figure 5.1. They are consistent with Gaussians of mean zero and width one. The pull with the traditional likelihood (the first term of Equation 5.1 only), which assumes infinite statistics for the mode histograms, however, shows a significant bias.
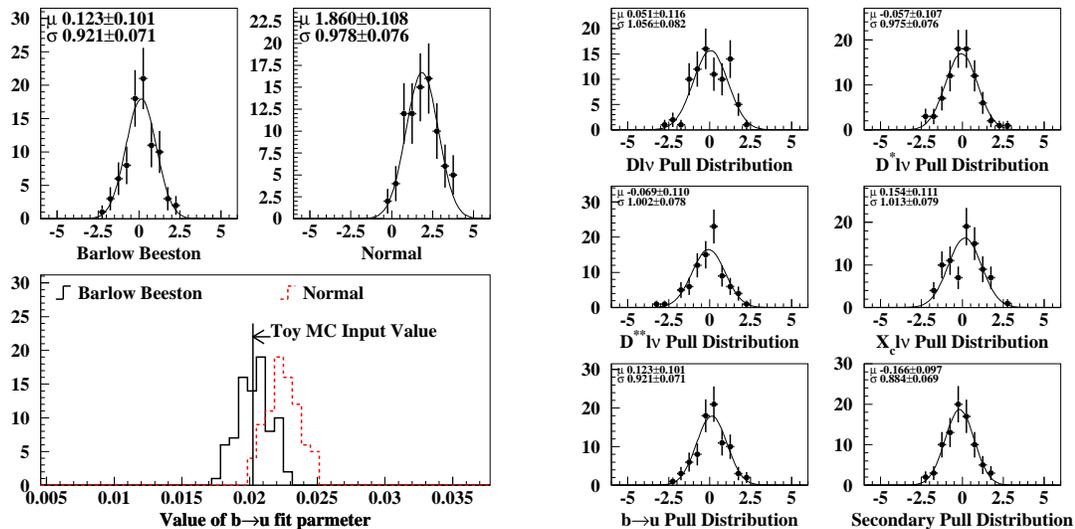
Figure 5.1:  Fitter Pull Distributions.
Comparison of the *Normal* likelihood function, which assumes infinite statistics for the data set used to generate the component histograms (i.e., the first term of Equation 5.1 only), and the Barlow and Beeston method, which accounts for possible statistical fluctuations in the component histograms (left). Pulls for all the fit parameters using the Barlow-Beeston method (right).

### 5.1.3   The Effect of the Variance of the Weights within a Bin

The likelihood above assumes that in each component histogram bin, the same weight is given to every event, but the weights used vary from event to event within a bin. For example, the form factors are input using weights that depend on the variables before the resolution is applied, but the component histograms are binned in the variables after the resolutions are applied. The variation of the weights within a bin causes the variance of the bin contents to have contributions from both fluctuations of the number of events and fluctuations of the average weight of the events. To account for the fluctuations in the average weight, an effective number of events, which is decreased from the true number of events, is used as the $a_{mb}$ in the likelihood.

The contents of the histograms, which are used to model the components in the fit, are

$$H_{mb} = \sum_{e \in \{\substack{\text{events in bin } b \\ \text{of model } m}\}} w_e, \qquad (5.4)$$

where $w_e$ is the weight for event $e$. If the Monte Carlo fluctuations are ignored, the expected number of data events in a bin $b$ is $\sum_m p_m H_{mb}$. The second term of the likelihood accounts for the variances

of the $H_{mb}$ values[1],

$$\text{Var}(H_{mb}) = \sum_{e \in \{\substack{\text{events in bin } b \\ \text{of model } m}\}} w_e^2. \tag{5.5}$$

To do this the values of $w_{mb}$ and $a_{mb}$ used in Equation 5.1 are

$$a_{mb} = \frac{H_{mb}^2}{\text{Var}(H_{mb})}$$

$$w_{mb} = \frac{1}{\sum_{b'} H_{mb'}} \frac{\text{Var}(H_{mb})}{H_{mb}}.$$

With these values, the likelihood will correctly take

$$w_{mb} a_{mb} = \frac{H_{mb}}{\sum_{b'} H_{mb'}} \tag{5.6}$$

as the fraction of events from model $m$ that falls in bin $b$, so that the definition of $f_b$ is in fact the expected number of events in bin $b$. It is easy to see that the values also satisfy the normalization requirement expressed in Equation 5.3, $\sum_b w_{mb} a_{mb} = 1$. The construction of the likelihood assumes that $a_{mb}$ is a Poisson distributed variable (i.e., $\text{Var}(a_{mb}) = a_{mb}$). The variance of $w_{mb} a_{mb}$ assumed by the fit is then

$$\text{Var}(w_{mb} a_{mb}) = w_{mb}^2 a_{mb} = \frac{\text{Var}(H_{mb})}{(\sum_{b'} H_{mb'})^2}, \tag{5.7}$$

which is the correct variance for the model's contribution in the bin.

To demonstrate how this works consider two cases, the trivial case of a constant weight $w$ for all $n$ events in the bin and the pathological case where one event in the bin gets a weight $w$ and all others get a weight of zero. In the first case,

$$a_{mb} = \frac{H_{mb}^2}{\text{Var}(H_{mb})} = \frac{(wn)^2}{w^2 n} = n \tag{5.8}$$

and

$$w_{mb} = \frac{1}{\sum_{b'} H_{mb'}} \frac{\text{Var}(H_{mb})}{H_{mb}} = \frac{1}{\sum_{b'} H_{mb'}} \frac{w^2 n}{wn} = \frac{w}{\sum_{b'} H_{mb'}}, \tag{5.9}$$

which is the standard case for which the likelihood was derived (note the $1/\sum_{b'} H_{mb'}$ term enforces the normalization condition). In the second case,

$$a_{mb} = \frac{H_{mb}^2}{\text{Var}(H_{mb})} = \frac{(w)^2}{w^2} = 1 \tag{5.10}$$

and

$$w_{mb} = \frac{1}{\sum_{b'} H_{mb'}} \frac{\text{Var}(H_{mb})}{H_{mb}} = \frac{1}{\sum_{b'} H_{mb'}} \frac{w^2}{w} = \frac{w}{\sum_{b'} H_{mb'}}, \tag{5.11}$$

---

[1] If the events were not weighted then $\text{Var}(H_{mb})$ would be equal to $H_{mb}$

which is Poisson fluctuations with one effective event in the bin, as it should be if all the other events have zero weight.

## 5.2   Component Modeling

The contributions to the data sample are summarized in Table 5.1. All of the signals and backgrounds except for the fake leptons are modeled with the Monte Carlo simulation introduced in Section 3.4. The accuracy of these models is subject to the simulation of both the underlying physics and the detector response. This section discusses the underlying physics models used in the fit and corrections applied to the default Monte Carlo simulation. The uncertainties on the accuracy of these models and the detector simulations are discussed in Chapter 6.

### 5.2.1   Corrections to the Underlying Event Description

The CLEO Monte Carlo includes a detailed description of the physics of $B$ decays and the decays of all the subsequent daughters. The description is based on a large variety of measurements and has been tuned to reproduce many inclusive spectra which have been measured in CLEO. Nonetheless, there are some spectra which are not well reproduced. Two cases, the number of $K_L^0$ mesons and the number of baryons produced in $B$ decays could affect the result presented here. Because $K_L^0$ mesons are not detected and baryonic events often have neutrons which are not detected, these must be modeled correctly in order to correctly model the neutrino momentum and direction resolution. The Monte Carlo simulate d events in which these processes are present are reweighted to reflect a better estimate of their rates than is included in the default CLEO Monte Carlo.

The average number of $K_L^0$ mesons produced in $B$ decays is equal to the average number of $K_S^0$ mesons (ignoring very small CP violating effects), which can be measured in the decay mode $K_S^0 \to \pi^+\pi^-$. A comparison of the number of $K_S^0$ mesons produced in the $B\overline{B}$ events with the CLEO Monte Carlo simulation shows that the simulation underestimates the number by 7.2%. The number of $K_L^0$ mesons is therefore corrected with the weight

$$w_e = c^{N_{K_L^0}}, \tag{5.12}$$

where $N_{K_L^0}$ is the number of $K_L^0$ mesons in event $e$, the correction factor $c$ is 1.072, and $w_e$ is the weight given to the event.

The number of neutron and protons produced in $B$ decays is significantly underestimated by the simulation. The rate $\mathcal{B}(B \to p/\overline{p}$ anything) in the simulation is 5.2%, but the PDG average is $8.0 \pm .4\%$ [1], 53% higher. The PDG average is a combination of measurements from ARGUS and CLEO which are consistent with each other. The number of $B \to$ baryon events is corrected by

giving events in which the other $B$ decayed to baryons a weight of 1.53.

## 5.2.2   Methods for the Application of Radiative Corrections

Experimental results in high-energy physics are often compared to calculations of the process being studied which do not include the radiative corrections. These calculations, without radiative corrections, are referred to as Born level calculations. To do this, the effect of radiative corrections on the results of interest are modeled and removed using the soft photon approximation, which ignores the changes of the decay matrix element due to radiation. The physics of radiative corrections is reviewed in Section 2.6. In this section the various methods available for the calculation of the corrections are discussed. In the following discussion, the emitting particle will be referred to as the lepton, but the equations are applicable to any charged particle.

There are two different approximations which are often used for applying radiative corrections: the splitting function method and the exponentiation method. The splitting function takes into account the full lepton energy distribution, but ignores corrections which are higher order in $\alpha$ (i.e., multiple photon emission). This method is used in the PHOTOS [44] program, which is a Monte Carlo simulation of radiative corrections discussed in detail in the next section. PHOTOS adds up to two photons per event raising the precision of the correction to order $\alpha^2$. The exponentiation method sums the soft corrections to all orders in $\alpha$, but ignores the hard corrections and the effect of the shape of the lepton energy spectrum. This is the basis of the Atwood and Marciano prescription [45].

In the splitting function method, the Born level differential decay rate, $d\Gamma_B/dE_\ell$, is convolved with a function, $\rho(k)$, to get the decay rate which is differential with respect to the lepton energy after radiation, $E_\ell'$,

$$\frac{d\Gamma}{dE_\ell'} = \int dE_\ell \int dk \, \frac{d\Gamma_B}{dE_\ell} \, \rho(k) \, \delta(E_\ell - E_\ell' - k), \tag{5.13}$$

where the lepton energy before radiation, $E_\ell$, and the emitted photon energy, $k$, are integrated over. This convolution is based on the equivalent photon approximation [12].

The exponentiation method is based on reweighting the spectrum to reflect the radiative corrections,

$$\frac{d\Gamma}{dE_\ell} = \frac{d\Gamma_B}{dE_\ell} \left[ \left( \frac{E_\ell^{\max} - E_\ell}{E_\ell^{\max} - \bar{E}_\ell} \right) \left( \frac{\bar{E}_\ell}{E_\ell} \right) \right]^{(2\alpha/\pi)\{(1/2\beta)\ln[(1+\beta)/(1-\beta)-1]\}}, \tag{5.14}$$

where $\bar{E}_\ell$ is the mean lepton energy, $E_\ell^{\max}$ is the maximum possible lepton energy, and $\beta = |P_\ell|/E_\ell$ is the velocity of the the lepton. The normalization of the effect is set by the use of $\bar{E}_\ell$, so that there is suppression of the decay rate at lepton energies higher than the mean and enhancement at lepton energies lower than the mean. This is an ad hoc choice, which forces the right global enhancement/suppression behavior[2]. At high observed lepton energy, this is a good approximation,

_____

[2] Atwood and Marciano also suggest fitting a constant to the data, which would appear as multiplicative factor in

because the dominant effect is the loss of high-energy leptons to lower energies. At low observed lepton energy, the dominant effect is the appearance of leptons that have "radiated down" from higher energies, which depends on what the Born level differential decay rate is at higher energies. Since, when calculating the correction to low lepton energies, the reweighting method only uses the Born level differential decay rate at low energies, it cannot correctly model the low energy region.

When more kinematic information than just the lepton energy is used (as in this analysis), the reweighting technique has another problem. Reweighting the Monte Carlo simulated events will properly modify the lepton energy distribution, but the distributions of the other variables will not be correct. For example the recoil mass squared, $(p_B^\mu - p_\ell^\mu - p_\nu^\mu)^2$, for the decay $B \to Dl\overline{\nu}(\gamma)$, will be $m_D^2$ regardless of the reweighting, but it should have a contribution of the form $(p_D^\mu + p_\gamma^\mu)^2$ from the events in which a photon is radiated. In the splitting function method, the energies of the final state particles in a each Monte Carlo simulated event are modified, so that all the distributions are correctly modeled.

Richter-Wąs [46] has compared both techniques with exact order $\alpha$ calculations of $\tau^+ \to e^+ \overline{\nu}_e \nu_\tau(\gamma)$ [47] and $B^+ \to D^0 e^+ \overline{\nu}_e(\gamma)$ [48][3] and a splitting function method which uses the collinear approximation. For $\tau^+ \to e^+ \overline{\nu}_e \nu_\tau(\gamma)$ at high electron energy, she finds that both methods agree well with the exact order $\alpha$ calculation, but at lower lepton energy the exponentiation method begins to fail, while the splitting function continues to perform well. The splitting function method agrees with the exact order $\alpha$ calculation at the $\approx 1\%$ level. This remarkable agreement is because there is no hadronic structure, so the soft photon approximation is very good. For $B^+ \to D^0 e^+ \overline{\nu}_e(\gamma)$ the agreement is significantly worse. The exponentiation method agrees with the exact $\alpha$ calculation at the $\approx 30\%$ level, while the splitting function method agrees at the $\approx 20\%$ level. There are no exact order $\alpha$ calculations for semileptonic $B$ decay modes other than $B \to Dl\overline{\nu}$, so measurements must rely on the soft photon approximation method and assign an appropriate uncertainty to the result. In this analysis the radiative corrections are assigned a 50% uncertainty.

### 5.2.3  Implementation Details of the Radiative Corrections

Radiative corrections are applied in the simulation of $B\overline{B}$ events using the PHOTOS package [44], which adds photons to the event description after all the other decay physics is simulated but before the detector response is simulated. A sample of Monte Carlo simulated events with the PHOTOS based corrections applied is not available for $e^+ e^- \to q\overline{q}$ continuum events, so a more approximate correction is applied after the detector simulation.

---

the term being exponentiated. This is not very practical in analyses in which the form-factors are being measured.

[3] The calculation of $B^+ \to D^0 e^+ \overline{\nu}_e(\gamma)$ used is actually rescaling of a calculation for $K^+ \to \pi^0 e^+ \overline{\nu}_e(\gamma)$.

**The Splitting Function**

Both methods use the same splitting function. Before presenting this function we introduce the variables,

$$\beta_0 \;=\; \sqrt{1 - \frac{4m_{ch}^2 m_P^2}{(m_{ch}^2 + m_P^2)^2}} \tag{5.15}$$

and

$$k_{max} \;=\; \frac{m_P^2 - (m_l + m_Y)^2}{2m_P}, \tag{5.16}$$

where

$m_{ch}$ is the mass of the radiating particle (e or $\mu$ here),

$m_P$ is the mass of the parent particle,

$m_Y$ is the mass of the system recoiling against the lepton (i.e., the hadronic system and the neutrino),

$k$ is the energy of the radiated photon, and

$k_{max}$ is the maximum possible radiated photon energy.

In terms of these variables the splitting function is

$$
\begin{aligned}
k f(k, cos\theta, \phi) = \frac{k}{4\pi m_P^2} \Bigg\{ & \delta(\frac{k}{m_P}) \frac{m_P}{k} \, (1 - \frac{\alpha}{\pi} N) \\
& + \Theta(k - \epsilon) \frac{\alpha}{\pi} \left( 1 + (1 - \frac{k}{k_{max}})^2 \right) \frac{m_P^2}{2k^2} \frac{2}{1 - \beta_0 cos\theta} \Bigg\},
\end{aligned}
\tag{5.17}
$$

where

$$N \;=\; \frac{1}{\beta_0} \ln\left(\frac{1 + \beta_0}{1 - \beta_0}\right) \left( \ln\frac{k_{max}}{\epsilon} - \frac{3}{4} + \frac{\epsilon}{k_{max}} - \frac{\epsilon^2}{4k_{max}^2} \right), \tag{5.18}$$

$\theta$ is the angle between the photon and the lepton direction and $\phi$ is the azimuthal angle of the photon around the lepton direction. The minimum photon energy considered, $\epsilon$, is 1 MeV, which is well below the detector sensitivity. This splitting function, $\rho(k)$, discussed in the previous section is a one-dimensional version of the function presented above and can be obtained by integrating over the other variables, $\rho(k) = \int\int d\cos\theta \, d\phi \, k f(k, cos\theta, \phi)$.

The first term of the splitting function above has an integral of $(1 - \alpha/\pi N)$ and corresponds to the probability of not radiating a photon. The second term, with integral $\alpha/\pi N$, corresponds to the

probability for radiating a photon of energy $k$ in the direction specified by $\phi$ and $\theta$. A correction to the splitting function is applied by rejecting some of the generated photons (i.e., moving probability from the second term to the first term). This correction is

$$W \quad = \quad \left(\frac{1 - \beta_0 cos\theta}{1 - \beta_1 cos\theta}\right) \left(1 - \frac{1 - \beta_1^2}{G_s(1 - \beta_1^2 cos^2\theta)}\right) \left(\frac{1 + \beta_1 cos\theta}{2}\right) \tag{5.19}$$

where

$$\beta_1 \quad = \quad \sqrt{1 - \frac{4m_{ch}^2 m_P^2(1 - \frac{2k}{m_P})}{(m_P^2(1 - \frac{2k}{m_P}) + m_{ch}^2 - m_Y^2)^2}} \tag{5.20}$$

$$G_s \quad = \quad \frac{1}{2}(1 - \frac{k}{k_{max}} + \frac{1}{1 - \frac{k}{k_{max}}}). \tag{5.21}$$

This correction accounts for changes in the phase space of the rest of the decay due to the emission of the photon and subleading terms which are neglected in Equation 5.17. When this splitting function is integrated over $\cos\theta$, it is the same function used by Richter-Wąs [46].

After the photon is generated, a correction is made to the lepton and $Y$ systems to reflect the transfer of energy and momentum to the photon[4]. The ad hoc proscription used in PHOTOS is to treat the decay as two sequential two-body decays. The first is the parent decaying into the photon and an off-shell parent. The second is of the off-shell parent decaying to the lepton and the $Y$ system (see Figure 5.2). The mass of the off-shell parent is $m_{p'}^2 = (p_p^\mu - k^\mu)^2 = m_p^2 - 2km_p$. In the off-shell parents frame, the momentum of the lepton, which is equal to the momentum of the Y system, is

$$q = \frac{[m_{p'}^2 - (m_{ch} + m_Y)^2][m_{p'}^2 - (m_{ch} - m_Y)^2]}{2m_{p'}}. \tag{5.22}$$

To change the momentum of the lepton and the $Y$ system from their values before radiation to $q$, Lorentz boosts along their flight directions are applied. Each of these systems is then boosted from the off-shell parent frame into the parent frame. Finally, a rotation is applied to both systems to force the $Y$ systems direction of flight to be unchanged.
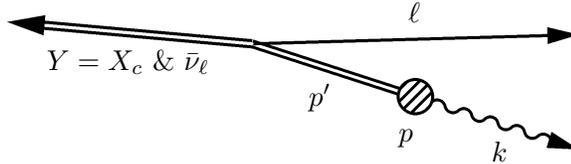


Figure 5.2: Sequential Two-Body Decays Used in PHOTOS Momentum Conservation Algorithm.

---

[4] The analytic methods used by Atwood and Marciano [45] and Richter-Wąs [46] do not conserve momentum.

**Application to Simulation**

When PHOTOS is used, the event description with the added photons is passed to the detector simulation. As mentioned above, a sample of Monte Carlo simulated $e^+e^- \to q\bar{q}$ continuum events with the PHOTOS correction applied before the detector simulation is not available. In this case, the corrections are applied using the same splitting function as PHOTOS, but after the detector simulation and the shift in the charged lepton and neutrino four-vectors due to the radiation is applied to the reconstructed variables. This is referred to as "myFSR". The two implementations of the PHOTOS splitting function have been checked against each other in detail to make sure they agree.

A complication arises when applying the correction after the detector simulation. Because the Monte Carlo events are generated without radiation, the simulation calculates the lepton identification efficiency for the lepton energy before radiation, which is higher than the lepton energy after radiation. The correct efficiency is obtained by reweighting the events with the ratio of the measured efficiency for the energy after radiation over the Monte Carlo simulation's efficiency for the energy before radiation,

$$\frac{\epsilon_{\text{Measured}}(E_\ell^{\text{after radiation}})}{\epsilon_{\text{Monte Carlo}}(E_\ell^{\text{before radiation}})}. \tag{5.23}$$

**The Effect of Radiative Corrections on the $M_X^2$ Reconstruction**

Radiative corrections play an important role in the measurement of the $M_X^2$ distribution. The reconstructed $M_{X,rec}^2$ is defined to be the mass squared of the system recoiling against the charged lepton and the neutrino. If a photon is radiated by the lepton in the event, it will be included in this definition of the recoil system,

$$M_{X,rec}^2 = (p_B^\mu - p_\ell^\mu - p_\nu^\mu)^2 = (p_X^\mu + p_\gamma^\mu)^2. \tag{5.24}$$

If radiation is neglected in the simulation, then the prediction of the $M_X^2$ distribution for a given mode will be biased low. Figure 5.3 shows the effect diagrammatically.

The application of the radiative corrections is complicated because approximately 72% of the photons are lost due to their low energy. Figure 5.4 shows the Monte Carlo simulation of the energy distribution of the radiated photons and the efficiency to find them. When a photon is lost in an event, it may cause the event to fail the $V_{cut}$ requirement[5]. If the event does pass the requirement, the reconstructed neutrino will be biased toward higher energy, pushing the reconstructed $M_X^2$ toward the true hadronic mass squared without the photon. If neglected, this would increase the measured $\langle M_X^2 - \overline{M}_D^2 \rangle$ moment with a 1.0 GeV lepton energy cut by 0.082 GeV$^2/c^4$, before detector effects are

---

[5] The $V_{cut}$ is described in Section 4.4.
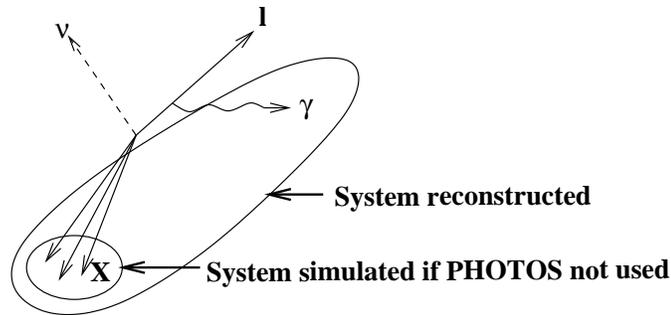
Figure 5.3: Diagram of the Effect of Radiative Corrections on the Reconstructed $M_X^2$.

included. This would be reduced to 0.037 GeV$^2/c^4$ after detector effects. For comparison, the total experimental error achieved on this quantity is 0.11 GeV$^2/c^4$. By using PHOTOS, the application of the radiative corrections correctly takes into account the detector effects.

### 5.2.4 $B \rightarrow Xl\overline{\nu}$ Modeling

The theoretical and experimental knowledge of heavy to heavy transitions is summarized in Section 2.5. In this section the specific models used in the nominal fit are discussed. In Section 6.2, we describe the variations of the model used to assess the dependence of the results on the models. Figure 5.5 shows the Monte Carlo simulation of the distributions of the reconstructed kinematical variables of the $B \rightarrow Xl\overline{\nu}$ modes. The fit exploits the differences in the distributions in the full three-dimensional kinematic space to determine the relative contributions of each mode.

#### $B \rightarrow Dl\overline{\nu}$ and $B \rightarrow D^*l\overline{\nu}$

Both $B \rightarrow Dl\overline{\nu}$ and $B \rightarrow D^*l\overline{\nu}$, have been measured extensively in variety of experiments [1]. Most of these measurements have been performed in the context of the HQET heavy to heavy predictions. In $B \rightarrow Dl\overline{\nu}$ decays, the only measured parameter is $\rho_D^2$, which is the slope of the Isgur-Wise function (see Section 2.5.1). Although this slope is the same for $B \rightarrow Dl\overline{\nu}$ and $B \rightarrow D^*l\overline{\nu}$ in the heavy quark limit, the parameters of the two processes are separated to allow experimental input on corrections to the heavy quark limit. In $B \rightarrow D^*l\overline{\nu}$ decays, there are three measured parameters, $\rho_{A_1}^2$, which is the slope of the function $h_{A_1}(w)$, and $R_1 = R_1(1)$ and $R_2 = R_2(1)$ which govern the helicity structure of the virtual $W$ (see Section 2.5.2). Measurements of these parameters and the values used in the fit are summarized in Table 5.2. For both of these modes, the theoretical predictions for the curvature are used. Because of a discrepancy in the $q^2$ distribution between data and the fit results (see Figure 5.14) the curvatures of the $B \rightarrow Dl\overline{\nu}$ and $B \rightarrow D^*l\overline{\nu}$ are altered to be half of their nominal values and large variations of the curvatures are used to determining the model dependence of the results.
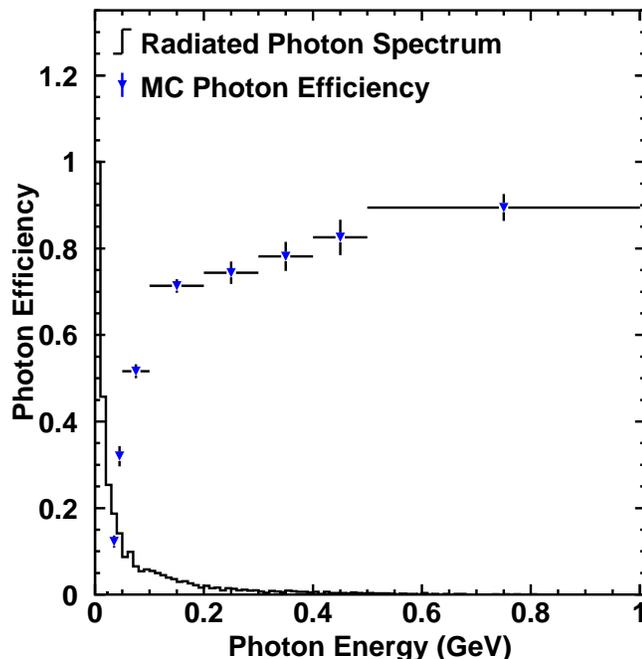
Figure 5.4: Radiated Photon Energy and Detection Efficiency Distributions.

The phase space for these decays vanishes at $w = 1$, so the exclusive measurements are not actually sensitive to the slope of the Isgur-Wise function at that point. Instead, experiments are sensitive to the slope in a range of $w$ defined by phase space and spin effects (see the factors of $w$ in Equations 2.38 and 2.41). Theory based curvature calculations connect the data to the slopes at $w = 1$. An observable better correlated to the actual quantity measured is the ratio of the form factor's value at either end of the region where the bulk of the phase space lies. For the $h_{A_1}(w)$ form factor of $B \to D^* l \overline{\nu}$ the ratio is

$$R_{w_1,w_2} \equiv \frac{h_{A_1}(w_1)}{h_{A_1}(w_2)} = \frac{1 - \rho_{A_1}^2 (w_1 - 1) + c_{A_1}(w_1 - 1)^2}{1 - \rho_{A_1}^2 (w_2 - 1) + c_{A_1}(w_2 - 1)^2}, \tag{5.25}$$

where $w_1$ and $w_2$ are the values of $w$ chosen to bound the region. A reasonable choice for $B \to D l \overline{\nu}$ is $w_1 = 1.15$ and $w_2 = 1.35$. To verify this, two CLEO measurements of $\rho_{A_1}^2$ can be compared, one assuming no curvature ($c_{A_1} = 0$), and one assuming the curvature given by Equation 2.44. The measurements are $\rho_{A_1}^2 = .92 \pm .12 \pm .06$ [49] and $\rho_{A_1}^2 = 1.61 \pm 0.09 \pm 0.21$ [50] respectively. The corresponding values of $R_{.15,.35}$ are $1.27 \pm .065 \pm .032$ and $1.28 \pm 0.062 \pm .145$, showing remarkable agreement. Equation 5.25 is used to convert between measurements which have assumed a curvature relation and other curvature assumptions. In the nominal fit, values of the curvatures, $c_D$ and $c_{A_1}$ are used that are different from the theoretical prediction used in measurements of $\rho_D$ and $\rho_{A_1}^2$ that contribute to the PDG world average (see Table 5.2). Equation 5.25 is therefore used to modify the
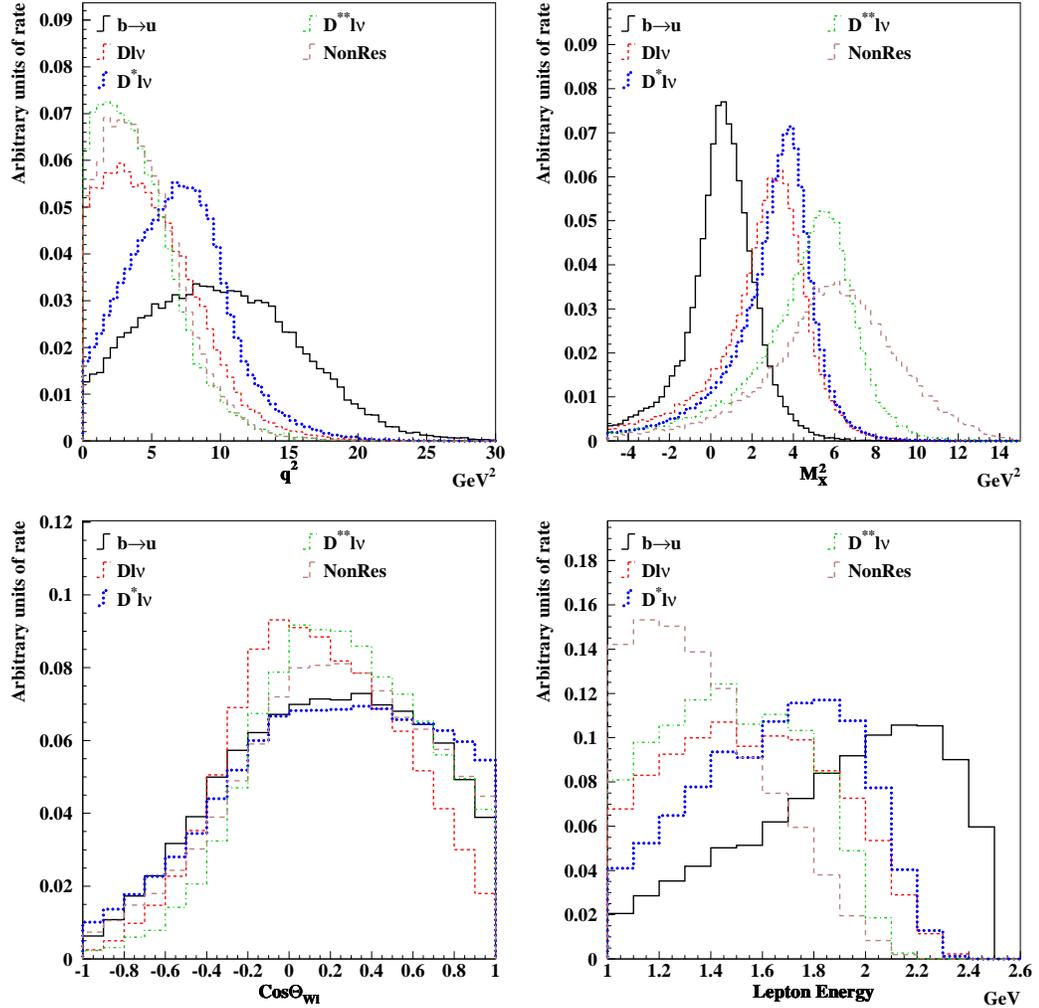
Figure 5.5:    Monte Carlo Simulation of the Decay Rate Distributions of Each of the $B \rightarrow Xl\overline{\nu}$ Modes.
The plots are of the distributions of the reconstructed variables $q^2$, $M_X^2$, $\cos\theta_{W\ell}$, and lepton energy. The modes shown are $B \rightarrow Dl\overline{\nu}$, $B \rightarrow D^*l\overline{\nu}$, $B \rightarrow D^{**}l\overline{\nu}$, nonresonant $B \rightarrow X_c l\overline{\nu}$, and $B \rightarrow X_u l\overline{\nu}$.

values of the $\rho$ parameters to maintain consistency with the measurements.

Table 5.2: Measurements of $B \to Dl\overline{\nu}$ to $B \to D^*l\overline{\nu}$ Form Factors.
The $c_D$ and $c_{A_1}$ parameters are derived from the $\rho_D^2$ and $\rho_{A_1}^2$ parameters with Equations 2.40 and 2.44. The parameter, $C_{R_1 R_2}$, is the correlation of the $R_1$ and $R_2$ measurements.

| Parameter | PDG | CLEO | Nominal for Fit |
|---|---|---|---|
| $\rho_D^2$ | $1.19 \pm 0.15 \pm 0.12$ | $1.27 \pm 0.25 \pm 0.14$ | 0.80 |
| $c_D$ | $1.10 \pm 0.20$ | $1.18 \pm 0.30$ | 0.56 |
| $\rho_{A_1}^2$ | $1.51 \pm 0.05 \pm 0.12$ | $1.61 \pm 0.09 \pm 0.21$ | 1.20 |
| $c_{A_1}$ | $1.39 \pm 0.14$ | $1.50 \pm 0.25$ | 0.70 |
| $R_1$ | - | $1.29 \pm .26$ | 1.29 |
| $R_2$ | - | $0.70 \pm .18$ | 0.70 |
| $C_{R_1 R_2}$ | - | 0.814 | - |

## $B \to D^{**}l\overline{\nu}$, Nonresonant $B \to X_c l\overline{\nu}$, and $B \to X_u l\overline{\nu}$

As discussed in Sections 2.5.3 and 2.5.4, there are very few measurements available for the $B \to D^{**}l\overline{\nu}$, nonresonant $B \to X_c l\overline{\nu}$ and $B \to X_u l\overline{\nu}$ modes, none of which are directly applicable to this analysis. It is therefore necessary to use theoretical models which are not based on a rigorous expansion for the simulation of these modes. For the $B \to D^{**}l\overline{\nu}$ and $B \to X_u l\overline{\nu}$ modes, the ISGW2 model [51] is used. This model is an updated version of the ISGW constituent quark model [52]. For the $B \to X_c l\overline{\nu}$ nonresonant mode, the Goity and Roberts model [33] is used, but with the resonant $D^*$ final state removed. The Goity and Roberts model is a calculation of $B \to D^{(*)}\pi\ell\nu$ in the context of HQET and chiral perturbation theory. These modes are not expected to saturate the nonresonant $B \to X_c l\overline{\nu}$ contribution, but because the hadronic final state is not reconstructed the detector response is expected to be similar enough to the other possible decay modes to be used. The lack of knowledge of mass distribution of the nonresonant mode is the dominant contribution to the model dependence of the hadronic recoil massed squared moment measurement (see Section 6.2.3).

## 5.2.5 Secondary Leptons

Secondary leptons are the background from leptons in $B\overline{B}$ events whose parent is not a $B$ meson. The largest contribution to this mode is from semileptonic decays of charmed hadrons, $B \to c \to \ell$. These are called cascades. The other contributions include $B \to \psi^{(\prime)}X$, $B \to X\tau\nu$, Dalitz decays of the $\pi^0$ and $\eta$, decays in flight of $\pi^\pm$ and $K^\pm$, and conversions of photons to $e^+e^-$ pairs in the detector. The Monte Carlo predictions for the contributions of each of these modes as a fraction of the total for each lepton type is shown in Table 4.1.

The trustworthiness of the simulations' prediction for the rate and normalization of the cascade contribution relies on the correct modeling of the basic physics processes which contribute. The

inclusive charm momentum spectra have been tuned to approximately agree with data (see Figure 5.6), and the semileptonic branching fractions of the charmed hadrons are adjusted to sum up to the inclusive measurements of charm semileptonic branching fractions. There is however a complication with the normalization of the production of charm in $B$ decays. The sum of measurements of the inclusive charm production in $B$ decays leaves an unexplained deficit compared to theoretical predictions (this is reviewed in the 2000 PDG [6]). This is the often discussed charm counting problem, and it is still awaiting improved measurements of the charm and $B$ branching fractions for a resolution. In the simulation, the charm production has been adjusted to make $b \rightarrow c$ transitions account for the bulk of $B$ decays and saturate the predicted level of charm production. Table 5.3 shows the PDG and CLEO simulation (QQ) values of the various input branching fractions. The result is a prediction for the overall cascade branching fraction, $\mathcal{B}(B \rightarrow c \rightarrow \ell)$, which is somewhat different for the PDG than for QQ.

Table 5.3: Prediction of $\mathcal{B}(B \rightarrow c \rightarrow \ell)$ Based on Summing Exclusive Modes.
This table shows inclusive $B$ meson branching fractions to various charmed hadrons and the inclusive semileptonic branching fractions for each of those hadrons. For each mode these are multiplied together to get the modes' contribution to the inclusive branching fraction $\mathcal{B}(B \rightarrow c \rightarrow \ell)$. These contributions are summed to get a prediction for the total branching fraction $\mathcal{B}(B \rightarrow c \rightarrow \ell)$. Values for both the PDG and the default CLEO $B\overline{B}$ Monte Carlo simulation (labeled QQ) are shown.

| Mode | $\mathcal{B}(B \rightarrow c)$ PDG | QQ | $\mathcal{B}(c \rightarrow \ell)$ PDG | QQ | $\mathcal{B}(B \rightarrow c \rightarrow \ell)$ PDG | QQ |
|---|---|---|---|---|---|---|
| $D^0$ | $63.5 \pm 2.9\%$ | $69.7\%$ | $6.75 \pm 0.29\%$ | $6.7\%$ | $4.39 \pm .27\%$ | $4.67\%$ |
| $D^\pm$ | $24.1 \pm 1.9\%$ | $29.6\%$ | $17.2 \pm 1.9\%$ | $16.7\%$ | $4.11 \pm .56\%$ | $4.94\%$ |
| $D_s^\pm$ | $10.0 \pm 2.5\%$ | $9.5\%$ | $8 \pm 5\%$ | $7.9\%$ | $.8 \pm .6\%$ | $.75\%$ |
| $\Lambda_c$ | $6.4 \pm 1.1\%$ | $7.0\%$ | $4.5 \pm 1.7\%$ | $4.48\%$ | $.29 \pm .12\%$ | $.31\%$ |
| Sum | | | | | $9.63 \pm .88\%$ | $10.67\%$ |

There are also a number of direct measurements of the branching fraction $\mathcal{B}(B \rightarrow c \rightarrow \ell)$. These come from a variety of methods including fits to the inclusive lepton spectrum in $B\overline{B}$ events from CLEO [53], a method using charge and angle correlations of events with two leptons from CLEO (referred to as "Double Tag") [54] and method using vertex information from LEP [1]. Table 5.4 summarizes the various measurements which show a sizable spread. The rate relevant to the size of the secondary lepton contribution to the data sample is that over 1.0 GeV. This means that the result labeled *Recent CLEO Double Tag* is the most directly applicable [55]. Because of the spread of the results, the contribution from the secondaries is allowed to vary in the fit.

## 5.2.6   Continuum Leptons

The off-resonance data provides a very good model of the contribution to the data sample from non-$B\overline{B}$ events, the majority of which are from $e^+e^- \rightarrow q\overline{q}$ continuum events. Unfortunately even
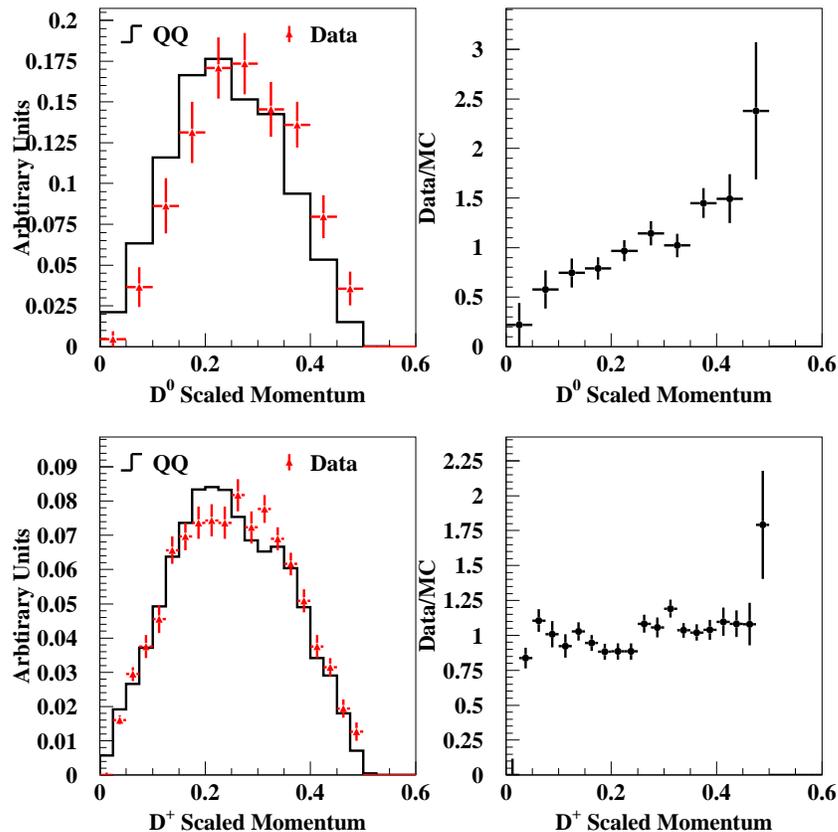
Figure 5.6: Comparison between CLEO $B$ Decay Model (QQ) and Data for the $D$ Momentum Spectra in $B\overline{B}$ Events.
$D^0$ momentum spectra (top) and $D^{\pm}$ momentum spectra (bottom) for the CLEO $B$ decay model (line) and measurement (dots). The right histograms show the data/MC ratio. The measurements are from reference [56].

Table 5.4: Predictions and Measurements of the Branching Fraction $\mathcal{B}(B \to c \to X\ell\nu)$.

| Source | $\mathcal{B}(B \to c \to X\ell\nu)$ |
|---|---|
| LEP $(B^0/B^{\pm}/B_s^0/b-baryon)$ | $8.0 \pm .30$ % |
| CLEO Lepton Spectrum Fit | $9.53 \pm .30$ % |
| CLEO Double Tag | $7.8\text{-}8.2 \pm .4$ % |
| Exclusive Sum (table 5.3) | $9.63 \pm .88$ % |
| CLEO $B\overline{B}$ MC | $10.7$ % |
| Recent CLEO Double Tag $(E_\ell \geq 1.0 GeV)$ | $\dfrac{\mathcal{B}(B \to c \to s\ell\nu)_{\text{data}}}{\mathcal{B}(B \to c \to s\ell\nu)_{\text{CLEO}B\overline{B}\text{ MC}}} \approx 1.07 \pm 0.03$ |

with the sophisticated handling of the histogram statistics, there is not enough off-resonance data to fill the three-dimensional histograms used in the fit. Instead, Monte Carlo simulated data is used to model the real leptons from continuum events, and a data based model is used for the fake leptons (see Section 5.2.7). The combination of the two models is tested with the off-resonance data (see Section 5.2.8).

The Monte Carlo simulation uses the JETSET [38] package in which the parameters have been tuned to reproduce the momentum spectra of the various charmed hadrons produced in $e^+e^- \to c\bar{c}$ events [56] This model is expected to be adequate because the bulk of the real leptons from continuum are the daughters of charmed hadrons and the decay of these hadrons into leptons is also tuned based on a variety of measurements. The decomposition of the continuum background based on the simulation is shown in Table 5.5.

Table 5.5: Sources of Real Leptons in Continuum.
This table shows the fractional composition of the real leptons in continuum Monte Carlo events after the full event selection.

| Parent of Lepton | Electrons | Muons |
|---|---|---|
| $\gamma$ Conversion | 0.00 | 0.00 |
| Decay In Flight | 0.00 | 0.00 |
| $\tau$ | 0.01 | 0.01 |
| $D^0$ | 0.47 | 0.47 |
| $D^\pm$ | 0.39 | 0.42 |
| $D_s$ | 0.07 | 0.09 |
| $\Lambda_c$ | 0.01 | 0.01 |
| $\pi^0$ Dalitz | 0.03 | 0.00 |
| $\eta$ Dalitz | 0.01 | 0.00 |
| $\psi$ | 0.00 | 0.00 |
| $\psi'$ | 0.00 | 0.00 |
| total | 1.00 | 1.00 |
| non-open charm | 0.05 | 0.03 |

### 5.2.7   Fake Leptons

The fake lepton background from both continuum and $B\overline{B}$ events is modeled by treating as leptons tracks in data events that are selected to be pion or kaons. The events in this sample are then weighted based on measured fake rates to simulate the fake lepton contribution. There is a technical complication with this approach: the particle identification of the selected track is itself in doubt. A track that is in the *pion* sample may in fact be a muon or a kaon (the electron and proton possibilities are small enough to be ignored). This is a large effect, the number of pions in the *kaon* sample is $\approx 30\%$ and muons make up $\approx 10\%$ of the *pion* and *kaon* samples. This means that before tracks from the *pion* sample are promoted to be fake leptons, the real muons and kaons must to removed and the number of pions lost to the *muon* and *kaon* samples must be taken into account. This is

done by constructing a matrix, $M_{\text{PID}}$, that carries the true distributions, $D_{\text{true}}$, to the observed distributions, $D_{\text{measured}}$, and then inverting the matrix,

$$
\begin{aligned}
D_{\text{measured}}\left(M_X^2, \frac{q^2}{q_0^2}, \cos\theta_{W\ell}\right) &= M_{\text{PID}}(E_{\text{track}}, \cos\theta)\, D_{\text{true}}\left(M_X^2, \frac{q^2}{q_0^2}, \cos\theta_{W\ell}\right) \\
\implies D_{\text{true}}\left(M_X^2, \frac{q^2}{q_0^2}, \cos\theta_{W\ell}\right) &= M_{\text{PID}}^{-1}(E_{\text{track}}, \cos\theta) D_{\text{measured}}\left(M_X^2, \frac{q^2}{q_0^2}, \cos\theta_{W\ell}\right).
\end{aligned}
\tag{5.26}
$$

As indicated, the matrix is a function of the track energy, $E_{\text{track}}$, and the cosine of the angle between the track and the beam axis, $\cos\theta$. The distribution can be in any variables of interest; the fit variables are indicated in the equations because they are of primary interest. The components of these three matrices, $M_{\text{PID}}$, $D_{\text{measured}}$, and $D_{\text{true}}$ are given by

$$
\begin{pmatrix} N_{\text{measured}}^{\pi} \\ N_{\text{measured}}^{K} \\ N_{\text{measured}}^{\mu} \end{pmatrix} = 
\begin{pmatrix}
\overbrace{f_{\pi|\pi} \quad f_{\pi|K}}^{\textbf{MC \& fake rates}} & \overbrace{f_{\pi|\mu}}^{\epsilon_\mu \text{ \& } \textbf{MC}} \\
f_{K|\pi} \quad f_{K|K} & f_{K|\mu} \\
\underbrace{f_{\mu|\pi} \quad f_{\mu|K}}_{\textbf{fake rates}} & \underbrace{f_{\mu|\mu}}_{\epsilon_\mu}
\end{pmatrix}
\begin{pmatrix} N_{\text{true}}^{\pi} \\ N_{\text{true}}^{K} \\ N_{\text{true}}^{\mu} \end{pmatrix}
\tag{5.27}
$$

where $f_{x|y}$ is the probability of identifying a particle of type $y$ as a particle of type $x$, $N_{\text{true}}^{x}$ is the true number of $x$ in a bin of the distribution and $N_{\text{observed}}^{x}$ is likewise the observed number of $x$. The sources for the various $f_{x|y}$ parameters are described below:

| Element | Source |
| --- | --- |
| $f_{\mu|\pi}$, $f_{\mu|K}$, $f_{e|\pi}$, $f_{e|K}$ | fake rates measured in data |
| $f_{\mu|\mu} = \epsilon_\mu$, $f_{e|e} = \epsilon_e$ | electron and muon efficiency measured in data |
| $f_{\pi|\pi}$, $f_{\pi|K}$, $f_{K|\pi}$, $f_{K|K}$ | MC $dE/dx$ simulation rescaled so that $1 - f_{\pi|\pi} - f_{K|\pi} = f_{\mu|\pi} + f_{e|\pi}$ |
| $f_{\pi|\mu}$, $f_{K|\mu}$ | MC $dE/dx$ simulation rescaled so that $1 - f_{\pi|\mu} - f_{K|\mu} = f_{\mu|\mu}$ |

Once the true distributions are inferred they can be multiplied by the fake rates to calculate the contributions from the pion, kaon, and muon samples. Procedurally a weight for each event in the samples can be calculated, avoiding the need to actually construct the intermediate distributions. Using the weight, the distribution of the fake lepton contribution in any kinematic variable is easy to construct. The weights are

$$
w_e = \left(f_{\ell|\pi} M_{\pi x}^{-1} + f_{\ell|K} M_{K x}^{-1}\right)
\tag{5.28}
$$

where $\ell =$ e or $\mu$ and x $= \pi$, k, or $\mu$. This takes into account both the imperfect separation between pions and kaon and the contamination of both of those with muons. Events in the muon
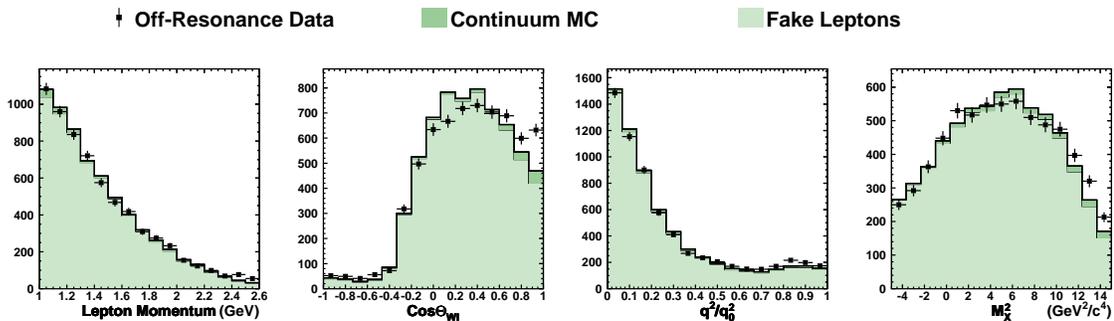
Figure 5.7: Comparison of Fake Lepton and Continuum Models with the Off-Resonance Data Electron Sample.
The sum of real electrons from the continuum Monte Carlo simulation (light) and fake electrons (dark) modeled as described in Section 5.2.7 compared to the off-resonance data for electrons from CLEO II.5. The variables are electron energy, $\cos\theta_{W\ell}$, $q^2/q_0^2$, and $M_X^2$ (left to right). The neural net cut has been removed to increase the sample size in the comparison.

sample have negative weights which subtract the excess of observed pion and kaons due to the muon contamination. This method predicts both the shapes and the total yields for the fake lepton contribution to the data sample.

## 5.2.8  Tests of the Fake Lepton and Continuum Lepton Model

The fake lepton and continuum model can be tested using the data taken off-resonance, where these are the only two contributions. Because the fake electron contribution is very small, the off-resonance electron sample is a good test of the continuum Monte Carlo simulation. Figure 5.7 compares the sum of the fake lepton and continuum models with the off-resonance data for the three variables used in the fit and for the lepton energy, which is most directly related to detector effects. The continuum Monte Carlo has been scaled so that the sum of the contributions of the two models is the data yield (the calculation described in Section 5.2.7 predicts the absolute fake lepton yield, so only the normalization of the continuum model is adjusted). The agreement is very good, but there is a small discrepancy in the $\cos\theta_{W\ell}$ distribution. To correct for the discrepancy a second order polynomial is fit to the (data - fake lepton)/continuum distribution. This polynomial is then applied as a weight to the continuum Monte Carlo events:

$$w_e = .905 - 0.095\cos\theta_{W\ell} + 0.487\cos^2\theta_{W\ell}. \tag{5.29}$$

Figure 5.8 shows the comparison after the correction has been applied. There is substantial improvement not only in the $\cos\theta_{W\ell}$ dimension, in which the correction is applied, but also in the other two fit variables, $q^2/q_0^2$ and $M_X^2$.
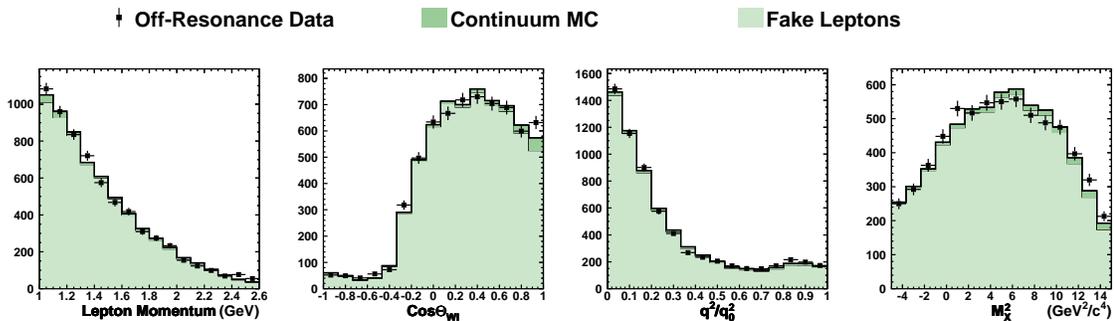
Figure 5.8:    Comparison of Fake Lepton and Continuum Models with the Off-Resonance Data Electron Sample with $\cos\theta_{W\ell}$ Correction Applied.
The set of histograms are the same as in Figure 5.7, but with the correction described in Equation 5.29 applied.
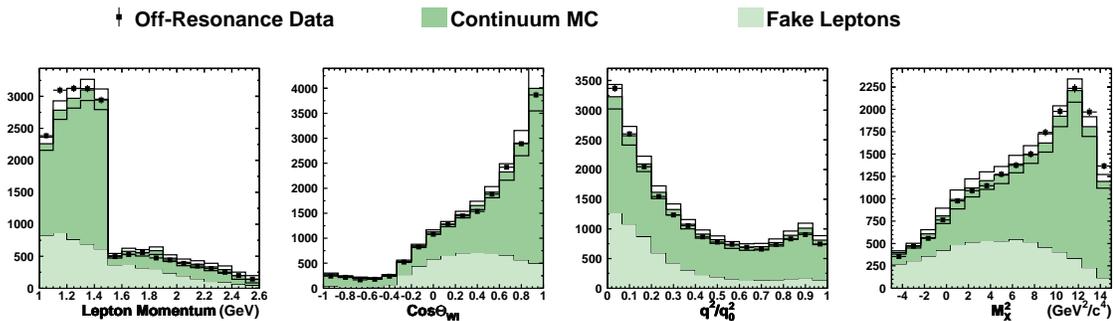


Figure 5.9: Comparison of Fake Lepton and Continuum Models with the Off-Resonance Data Electron Sample with $\cos\theta_{W\ell}$ Correction Applied.
Same as Figure 5.8, but for CLEO II5 muons. The solid lines above and below the top of the fake lepton contribution indicate the size of the uncertainty on the fake rates.

The muon sample has a much larger fake lepton contribution. In order to make a comparison between the off-resonance data and the sum of the continuum and fake lepton models, the scale of the continuum is set based on the off-resonance electrons (the same scale used in the electron sample comparisons). Figure 5.9 shows the comparison. For the low muon energy range, where a penetration of only three radiation lengths is required, the fake rate is much larger and the agreement slightly worse. To better quantify the quality of the agreement, the correction that would be necessary if the disagreement is assumed to be due to the measured fake rates is shown in Figure 5.10. With the neural net cut removed, the statistical precision of this test is better than the precision of the measured fake rates. A correction to fake rates based on the off-resonance data is therefore applied (and a variation is made to determine a the systematic uncertainty).

The comparison with the muon fake rate correction, the continuum $\cos\theta_{W\ell}$ correction, and
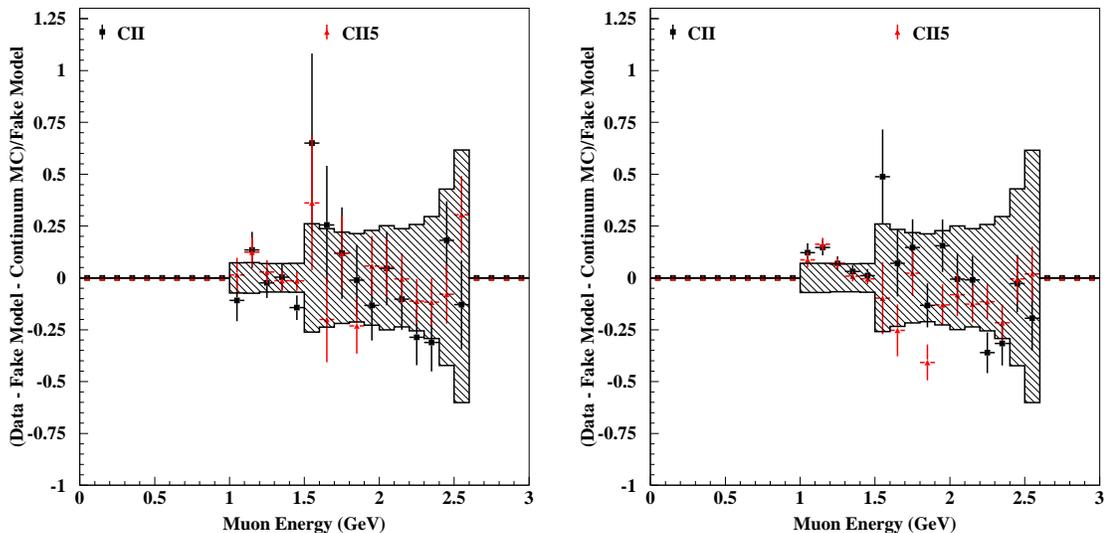
Figure 5.10:   Inferred Fake Rate Correction.
These plots show the correction that would be necessary if the full off-resonance discrepancy is assigned to the fake rates. The shaded region indicates the size of the statistical error on the fake rates. The left is with the standard analysis cuts. The right is with the neural network cut loosened to improve the statistical precision of the test.

the full selection used in the on-resonance fit is shown in Figures 5.11 and 5.12, for electrons and muons respectively. These comparisons validate the use of the Monte Carlo simulation to model the contribution of real leptons in continuum events.

The above comparisons used the electron component to fix the scale of the continuum leptons. This contribution could also be predicted from measurements of the inclusive hadron production cross-section at $\sqrt{s} = 10.52$ GeV which is 3.275 nb [58]. The two methods agree to better than 10%.

## 5.3   Goodness of Fit

For a maximum likelihood fit there is no value like the chi-squared that is a direct measure of whether the model is a good representation of the data. For this analysis, tests of the goodness of fit are further complicated by the fact that the results are systematics limited. In other words, the fit is not expected to be good as defined by a purely statistical test. Instead, the fit is expected to represent the data within the known systematic uncertainties.

One standard statistical test of the quality of the fit is the comparison of the distribution of the observed number of events per bin in the three-dimensional space with the expected number of events per bin (shown in Figure 5.13)[6]. This distribution shows that the data has small numbers of events in the bins which expect small numbers of events and large numbers of events in the

---

[6] This is equivalent to the likelihood per event distribution used in unbinned maximum likelihood fit.
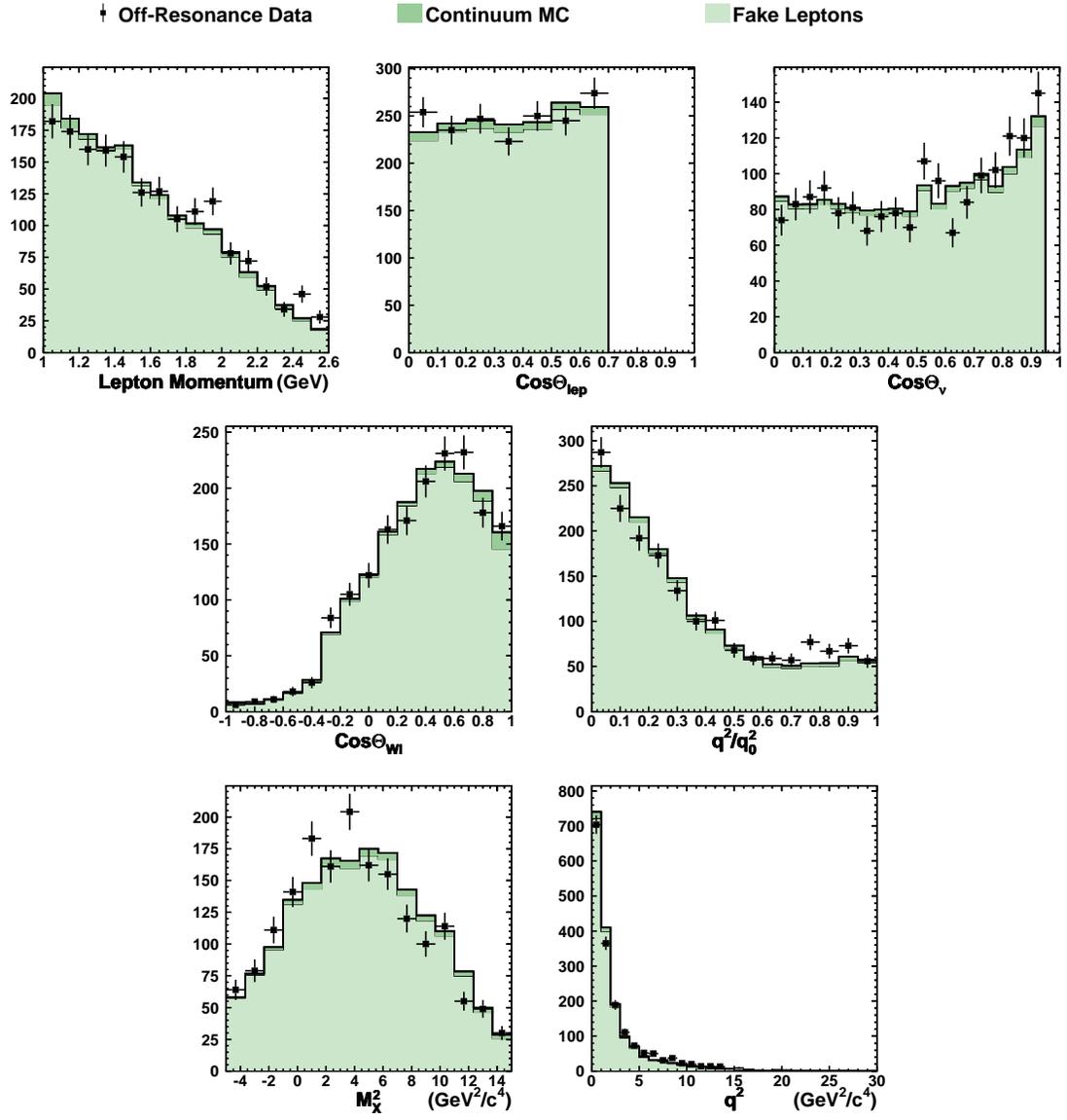
Figure 5.11: Comparison of Fake and Continuum Model with Off-Resonance Electron Data with All Corrections and All Cuts Including the Neural Net Cut.
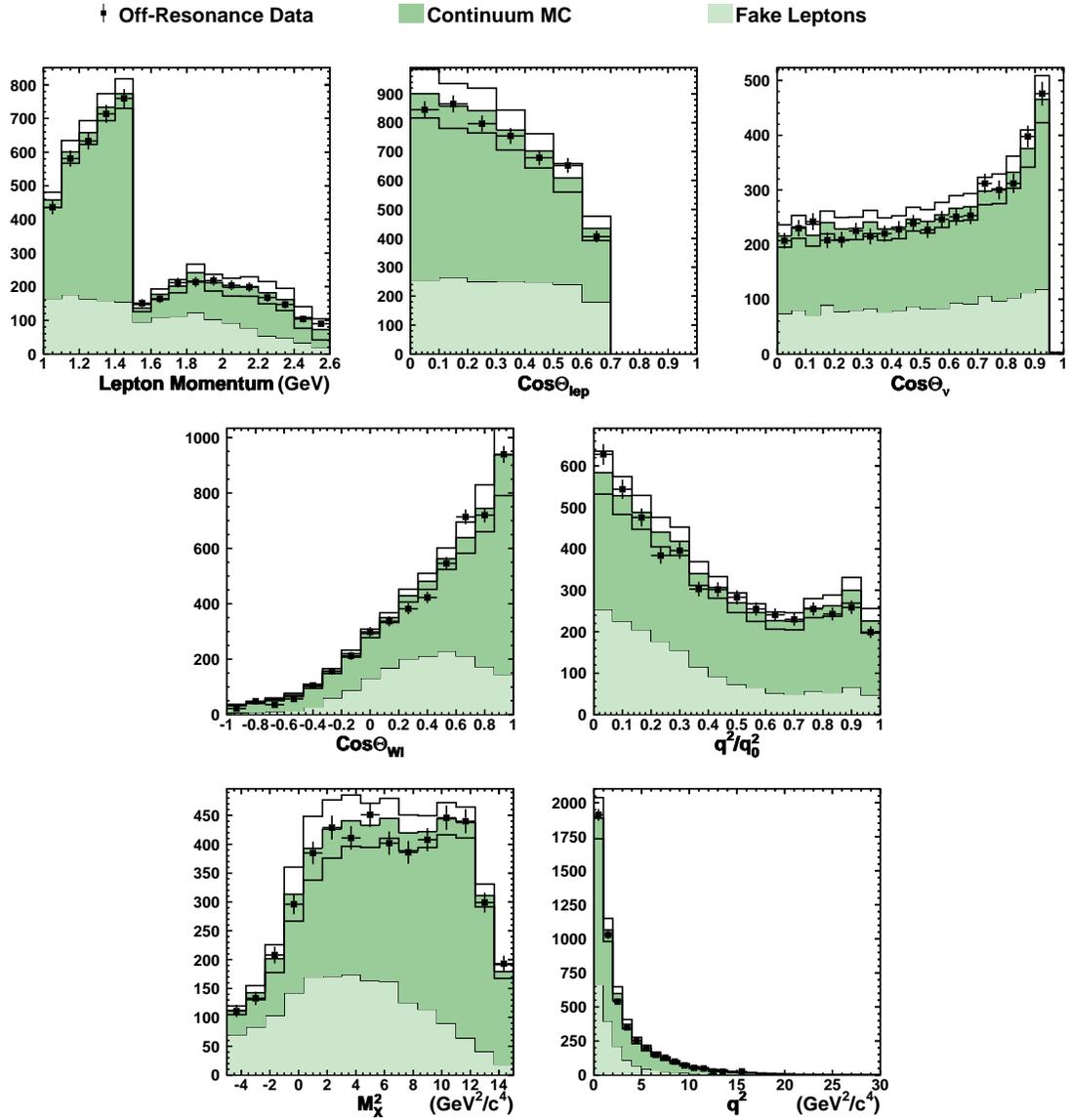
Figure 5.12: Comparison of Fake and Continuum Model with Off-Resonance Muon Data with All Corrections and All Cuts Including the Neural Net Cut.
The solid lines above and below the top of the fake lepton contribution indicate the size of the uncertainty on the fake rates.
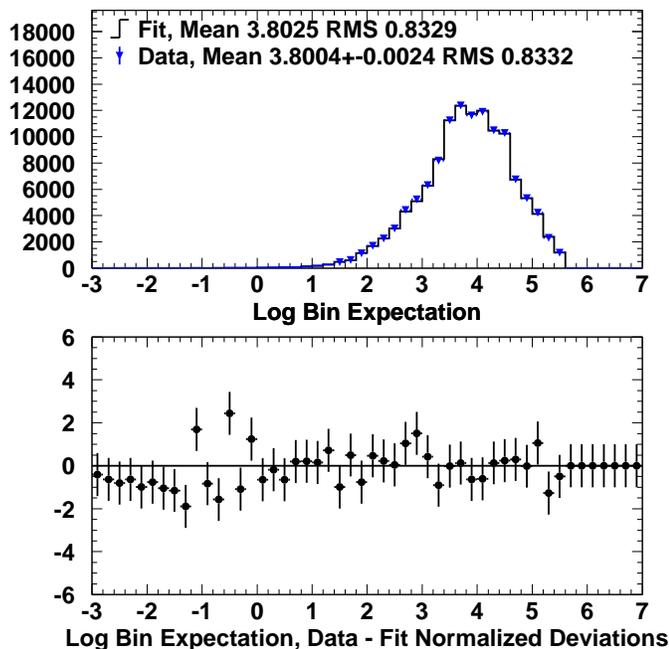
Figure 5.13: Log of the Expected Number of Events Per Bin Distribution.
The bottom plot shows normalized deviation of the histograms in the top plot. For each event an entry is made in the histogram at the log of the expected number of events, $\ln f_b$, for the bin in which the event falls. This is analogous to the likelihood per event distribution used in unbinned maximum likelihood fits.

bins which expect large numbers of events. This means that the tails of the distributions are well modeled. Also shown in Figure 5.13 is the normalized deviation, (data-fit)/statistical error, between the data and the prediction. The large fluctuations and apparent bias at the very low end of this distribution is due to the effect of Poisson statistics when fewer than one event is predicted per bin. The agreement of the mean and RMS of the distribution are indications that there is no bulk shift of the distributions.

Figures 5.14 and 5.15 show the projections of the fit results in the variables $q^2$, $M_X^2$, $\cos\theta_{W\ell}$, and $E_\ell$. The variable $q^2$ is shown in place of $q^2/q_0^2$, because $q^2$ is more easily interpreted physically (recall, the use $q^2/q_0^2$ is equivalent to varying the size of the $q^2$ bins as a function of $q_0$). The overall agreement is good, but there are a few regions which have systematic deviations. In the 4.0 to 7.0 GeV$^2$/$c^4$ region of the $q^2$ distribution there is an excess of data events above the fit; this is particularly visible for the electrons. Similarly in the high $\cos\theta_{W\ell}$ region there is an excess of data events. These are most likely due to imperfect modeling of the form factors of the $B \to Dl\overline{\nu}$ and $B \to D^*l\overline{\nu}$ decays. The $q^2$ disagreement is the motivation to use a curvature of the Isgur-Wise function which is half the predicted value (without this choice the discrepancy is clearly visible for muons as well). The $\cos\theta_{W\ell}$ deviation might be attributable to a $1\sigma$ shift in the $B \to D^*l\overline{\nu}$

form-factor parameter $R_1$.

Figure 5.16 shows the $M_X^2$ distribution in slices of $q^2$. This is useful because the signal shape and backgrounds change dramatically as a function of $q^2$. The backgrounds become very large for $q^2$ less than 2 GeV$^2/c^4$. For fake leptons, this is because there is no missing neutrino, so any deviation from zero is due entirely to resolution. For secondaries and continuum, it is because the maximum possible $q^2$ in $D$ meson decay is $q^2_{\max} = (M_D - M_K)^2 \approx 1.9$ GeV$^2/c^4$. In the $2 \leq q^2 \leq 11$ GeV$^2/c^4$ region, $B \rightarrow X_c l\overline{\nu}$ dominates but the composition varies considerably. Above 11 GeV$^2/c^4$ the $B \rightarrow X_u l\overline{\nu}$ component becomes visible (although still not large).

To assess whether the estimated systematic uncertainties cover the discrepancies between data and the fit, envelopes are constructed. The envelopes are the expected bound of the variations of the fit projections due to the systematic uncertainties. The various sources of systematic uncertainties are discussed in detail in the next chapter. For each projection there is an upper and lower envelope bound. These are calculated by summing in quadrature the deviations for each variation considered of the fit projections above and below nominal fit result. This procedure is analogous to the procedure used for calculating the systematic uncertainties on the physics results and is therefore an indicator of whether the estimated uncertainties in fact cover the true values of the physics parameters. The projections with envelopes are shown in Figures 5.17 and 5.18 for electrons and muons respectively.

## 5.4 Branching Fraction Results

The fit results give the fraction, $f_m$, of the data sample which is due to mode $m$. Using the fit results, the branching fractions are

$$\mathcal{B}_m = \frac{f_m \mathcal{N}}{\epsilon_m N_{B\overline{B}}},\tag{5.30}$$

where $\mathcal{N}$ is the number of events in the data sample, $\epsilon_m$ is the efficiency for mode $m$ to be in the sample, and $N_{B\overline{B}}$ is the number of $B\overline{B}$s in the data set. The total number of $B\overline{B}$s is calculated from the change in the hadronic cross-section between the off-resonance and on-resonance data sets[7].

The efficiency, $\epsilon_m$, is calculated from the Monte Carlo simulation. If there were no reweighting used, it would simply be the ratio of Monte Carlo simulated events which pass the event selection to the total number of simulated events generated. Because reweighting is used, this is somewhat more complicated and the approximation that the weights are not correlated is used:

$$\epsilon_m = \frac{\sum_s^{\#\text{selected}} w_1(s)w_2(s)...w_n(s)}{\sum_g^{\#\text{generated}} w_1(g)w_2(g)...w_n(g)} \approx \frac{\sum_s^{\#\text{selected}} w_1(s)w_2(s)...w_n(s)}{\langle w_1 \rangle_G \langle w_2 \rangle_G ... \langle w_n \rangle_G \sum_g^{\#\text{generated}} 1},\tag{5.31}$$

[7]It is assumed that the branching fraction of $\Upsilon(4S)$ to $B\overline{B}$ is 100%.
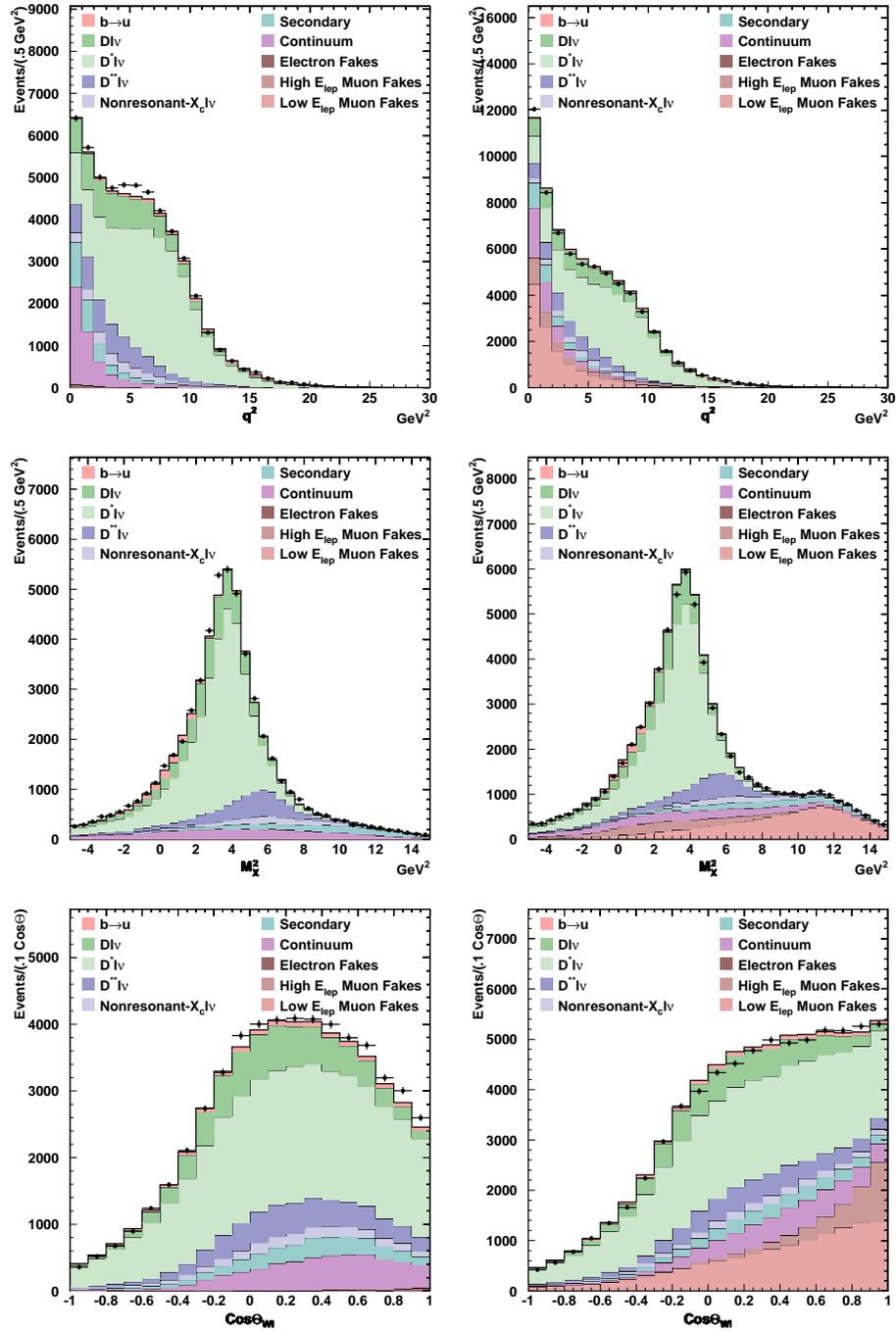
Figure 5.14: Fit projections for $q^2$, $M_X^2$, and $\cos\theta_{W\ell}$.
Projections of the fit result for $q^2$, $M_X^2$, and $\cos\theta_{W\ell}$ (top to bottom) for electrons(left) and muons(right) separately.
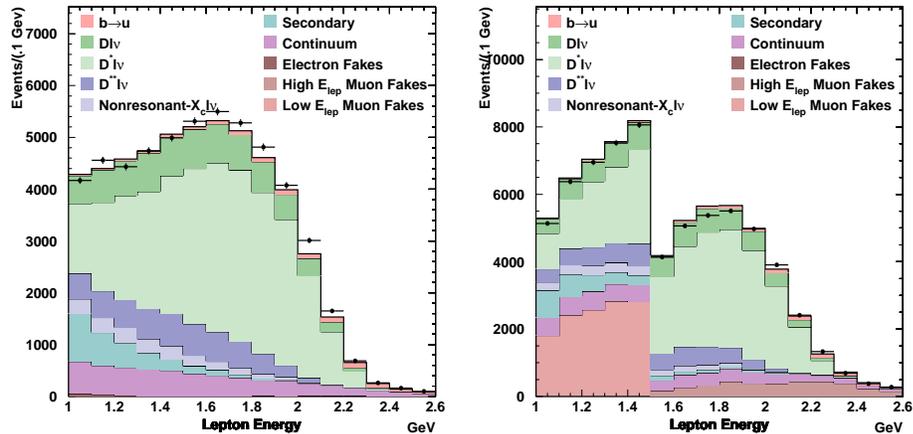
Figure 5.15: Fit projections for $E_\ell$.
Lepton energy projections for electrons (left) and muons (right).

where $w_i(x)$ are the weights on the event $x=s$ or $g$ and $\langle w_i \rangle_G$ are the averages of the weights over the set of generated events. All sums in the expression are only for the specific mode $m$ being considered. Between the first and second equations the approximation that the weights are uncorrelated is used. This should be a good approximation because the weights do not have large variances; most of them have the structure $1 + \delta$ where $\delta$ is much less than one.

The resulting branching fractions are shown in Table 5.6.

Table 5.6: Branching Fraction with Statistical Errors.

| Mode | Branching Fraction $(\times 10^{-2})$ |
|---|---|
| $B \to Dl\overline{\nu}$ | $1.919 \pm 0.076$ |
| $B \to D^*l\overline{\nu}$ | $6.374 \pm 0.060$ |
| $B \to D^{**}l\overline{\nu}$ | $1.509 \pm 0.073$ |
| Nonresonant $B \to X_c l\overline{\nu}$ | $0.695 \pm 0.070$ |
| $B \to X_u l\overline{\nu}$ | $0.115 \pm 0.008$ |
| Sum | $10.612 \pm 0.287$ |
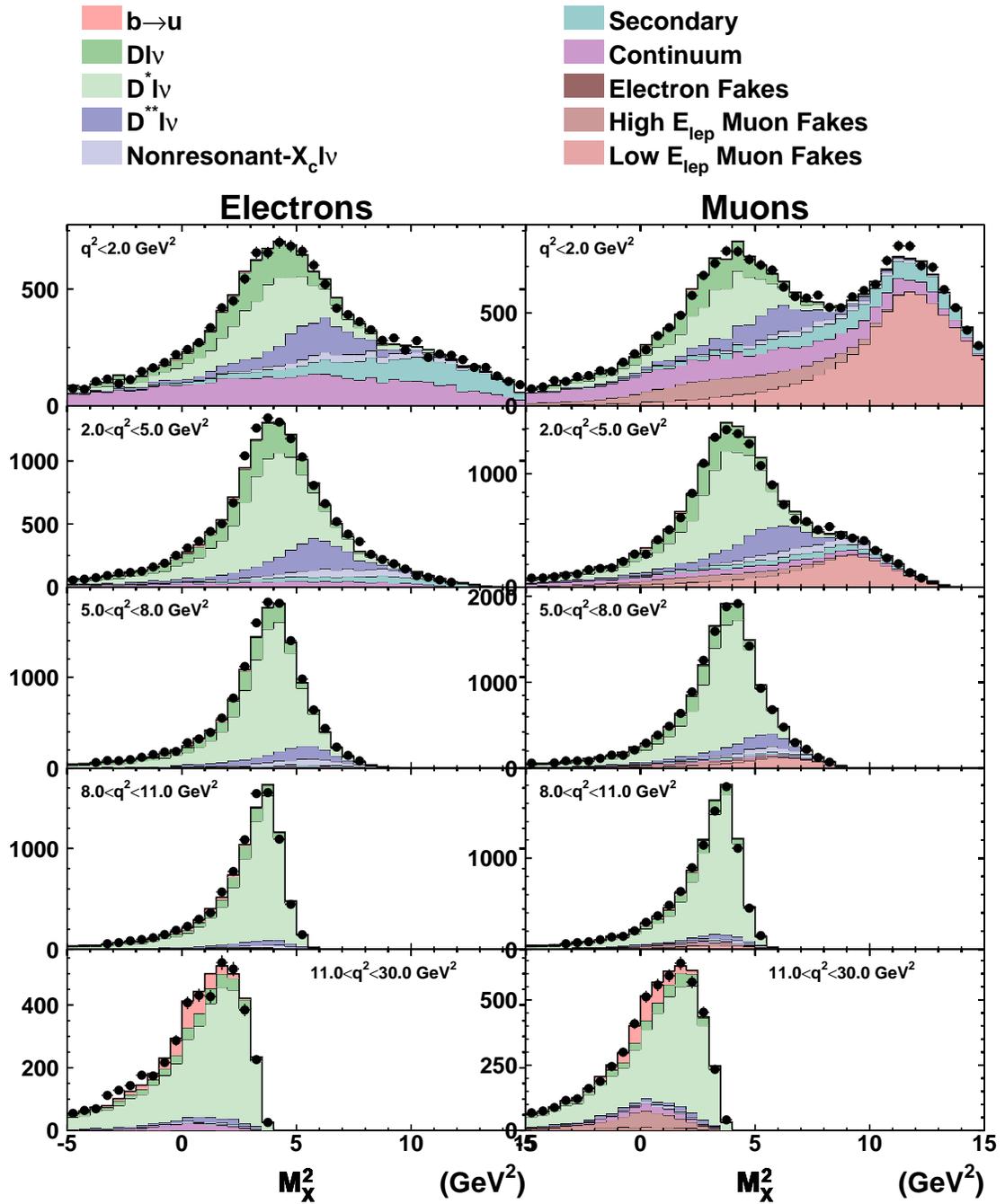
Figure 5.16: $M_X^2$ Distribution in Slices of $q^2$.
Electrons are on the left and muons are on the right. The slices from top to bottom are $q^2 < 2$, $2 < q^2 < 5$, $5 < q^2 < 8$, $8 < q^2 < 11$, and $11 < q^2 < 30$ GeV$^2/c^4$.
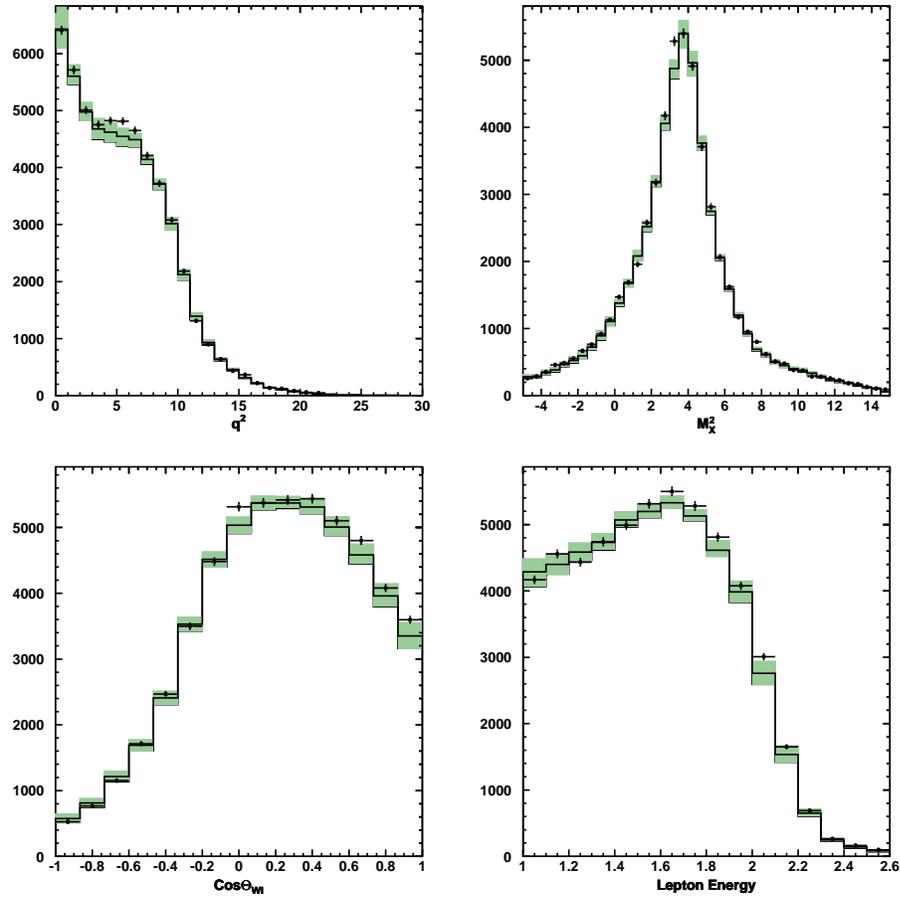
Figure 5.17: Systematics Envelope for Electrons.
Envelope defined by the deviations from the nominal fit of the systematic errors summed in quadrature. Plots are electrons only.
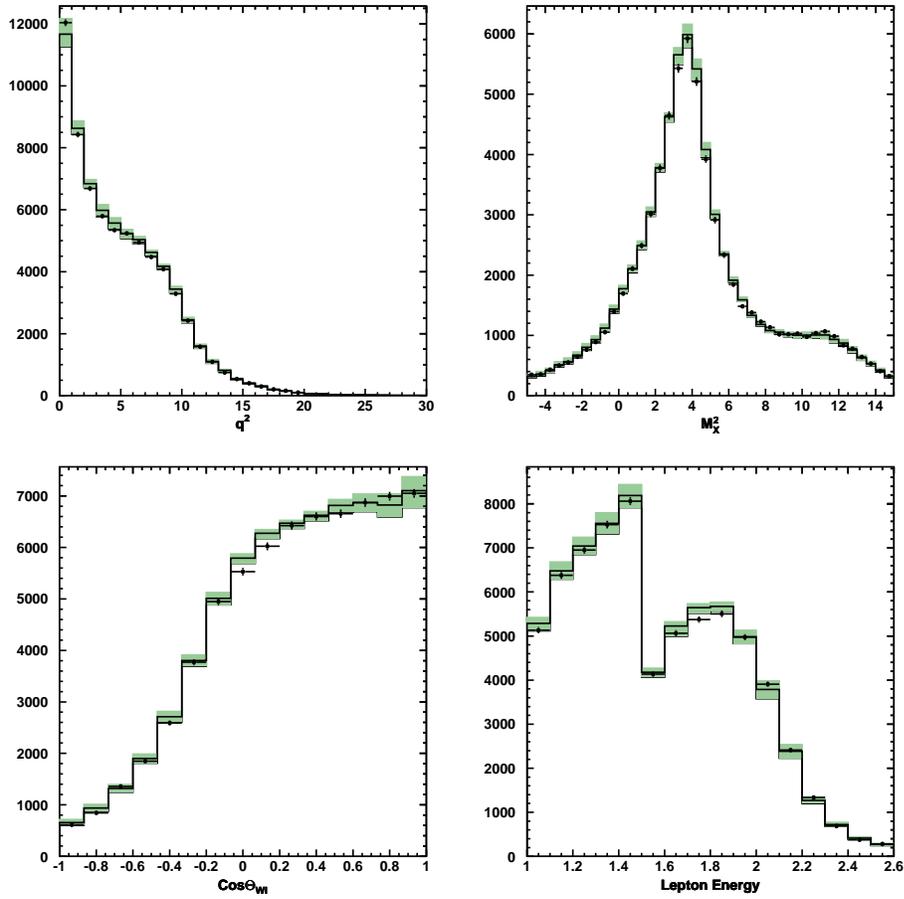
Figure 5.18: Systematics Envelope for Muons.
Envelope defined by the deviations from the nominal fit of the systematic errors summed in quadrature. Plots are muons only.

# Chapter 6  Systematics

The systematic uncertainties of the physics results in this analysis are much larger than the statistical errors. This is because of the uncertainties related to the neutrino reconstruction, the model dependence, and the very large data sample available. The neutrino reconstruction relies on correctly modeling of the physics of generic $B$ decay and the detector response to it. In general, the uncertainties are quantified by reweighting the simulated events to reflect a change in the input assumptions and then refitting. The shifts from the nominal results are combined into the two categories, detector systematics and model dependence. The detector systematics include the modeling the other $B$ decay in the event, the detector response, and radiative corrections. The model dependence is exclusively the dependence of the results on the $B \rightarrow Xl\overline{\nu}$ form factors.

## 6.1  Detector Systematics

### 6.1.1  Underlying Event

As discussed in Section 4.3, the detector is not sensitive to $K_L^0$ mesons, neutrons, and neutrinos. The data selection is designed to suppress events which contain these particles in the final state (except for the one neutrino being reconstructed), but they are still present in large numbers and affect the resolution on the neutrino kinematics significantly (see Figure 4.4). It is therefore neccesary to know how many of each of these particles are present in $B$ meson decays and to quantify the effects on the results of possible mismodeling of the $B$ decay physics. Because a higher multiplicity creates more chances to lose a particle, the charge particle and photon multiplicities must also be modeled.

**Number of $K_L^0$ Mesons**

The number of $K_L^0$ mesons should be equal to the number of $K_S^0$ mesons[1], which is straightforward to measure. A measurement of inclusive $K_S^0$ production in $B\overline{B}$ events (using off-resonance data to remove the continuum contribution) has been compared to the Monte Carlo simulations [59]. The result shows $(7.2 \pm 0.7)\%$ more of $K_S^0$ mesons in the data than in the simulation. Although the statistical uncertainty on this measurement is small, there are many systematics which have not been evaluated. The number of $K_L^0$ mesons per event is therefore varied by 10% of of the total number, a conservative estimation of the possible range of values. The weight corresponding to this correction is $w_e = c^{N_{K_L^0}}$ where $N_{K_L^0}$ is the number of $K_L^0$ mesons in an event and $c$ either 1.17 or

---

[1] CP violating effects could make the numbers of $K_S^0$ and $K_L^0$ mesons different, but it is a very small effect.

0.97, corresponding to raising and lowering the number of $K_L^0$ mesons in the event, respectively. In the nominal fit the value $c = 1.072$ is used (see Section 5.2.1).

**Number of Secondary Leptons**

The number of secondary leptons per $B$ decay affects this analysis in two ways. First is the secondary lepton can be used as the signal lepton. This case is only sensitive to the semileptonic decay rate for leptons with momenta above 1.0 GeV/$c$. This size of this contribution is determine by the maximization of the likelihood (see Section 5.2.5). The other way that the secondary rate affects this analysis is that secondary leptons are correlated with extra neutrinos which give a long high energy tail to the reconstructed neutrino energy. This effect depends on the total secondary lepton background, not just the part with lepton momenta above 1.0 GeV/$c$. The following discussion applies to the assessment of the systematic uncertainty due to the latter case (extra neutrinos).

The current knowledge of the number of secondary leptons is summarized in Table 5.4. The scatter in the predictions and measurements is fairly large, ranging from $\mathcal{B}(B \to c \to X\ell\nu) = 7.8\%$ to 10.7%. At the lower end of this range are the LEP measurments [1] and the older CLEO Double Tag analysis [54]. The meaurements from LEP are not directly applicable because they include the $\Lambda_b$ and $B_s$ parents in addition to the $B_\pm$ and $B_d$ which are of interest. The CLEO Double Tag has been superseded by a newer analysis, which is thought to be more reliable. The lowest end of the range it therefore excluded and the variations made are $\pm 10\%$ of the nominal 10.7% branching fraction. This is applied with a weight $w_e = c_{xl}^N$ where $N_{xl}$ is the number of lepton in addition to the signal lepton and $c$ is either 0.9 or 1.1.

**Number of Baryons**

As discussed in Section 5.2.1, the number of $B \to$ baryons in the physics simulations is significantly underestimated. The rate $\mathcal{B}(B \to p/\overline{p}$ anything) in the simulation is 5.2%, but the PDG average is $8.0 \pm .4\%$ [1], 53% higher. Although the error on this measurement is only 5%, the composition of the modes contributing to the difference is not understood (one of the reasons the simulation has not been corrected). This number of $B \to$ baryons in the final state is therefore varied by $\pm 20\%$ of the 8% rate of $\mathcal{B}(B \to p/\overline{p}$ anything), a significant fraction of the correction. Each event can have at most one baryon-antibaryon pair, so events with a baryon-antibaryon pair are given the weight $w_e = 1.23$, 1.53, or 1.85 for the low variation, the nominal and the high variation respectively.

**Charged and Neutral Multiplicities**

The data itself gives a good indicator of how well the charged particle and photon multiplicities are modeled. The observed track and shower multiplicities (shown in Figure 6.1) are directly related to these particle multiplicities, but have the added effects of efficiency and fakes signals. Assuming

there are not large cancelations in the mismodeling of these effects and the particle multiplicities, the track and shower multiplicities are a good measure of how well the particle multiplicities are modeled. To quantify the uncertainty, the simulation events are reweighted to reproduce the observed data multiplicities. This is acheived by weighting each event by the ratio of data to Monte Carlo simulation for the bin of multiplicity in which the event lies. The ratios are constructed with the Monte Carlo semileptonic $B$ branching fractions set to the fit results.

## 6.1.2   Detector Response

**Tracking Efficiency**

The absolute tracking efficiency has been studied in two ways in CLEO. The first uses $\tau$ pair events in which one $\tau$ has decayed to one charged particle and the other $\tau$ has decayed to three charged particles. Three of the four tracks are used to select this topology and then the efficiency to find the fourth track, which must be there because of charge conservation, is tabulated [60]. The other method used, is to embed the data from Monte Carlo simulated tracks in real data events and then process the data as usual [61]. The probability to find the track is then tabulated. The first method is more direct because it does not rely on simulation at all, but is also more limited because the event environment is not the same as the $B\overline{B}$ events of interest. Also the embedding gives more control for understanding the relative efficiency as a function of the transverse momentum. The result of the $\tau$ study is that at high momentum the efficiency is $\approx 95\%$ with an error of less than $0.5\%$. This makes the ineffiency $5.0 \pm 0.5\%$, where the uncertainty is comparable to the data versus Monte Carlo difference. Because in momentum regions of lower efficiency the uncertainty is higher, the error on the inefficiency is roughly constant across momenta. The uncertainty is determined by rewieghting events with lost tracks, whose number is proportional the ineffiency, by $\pm 10\%$. The weight used is $w_e = c^{N_{\mathrm{lost}}}$ where $N_{\mathrm{lost}}$ is the number of tracks lost in the event and $c$ is either 0.9 or 1.1, corresponding to raising and lowering the efficiency, respectively.

**Shower Efficiency**

There have been fewer studies of the photon effieciency of CLEO. The best study available embeds Monte Carlo simulated data from $\pi^0$ into $B\overline{B}$ events [62]. The result is an efficiency of $60\pm5\%$, where the uncertainty on the efficiency is comparable to the data versus Monte Carlo difference. There are extra complications related to $\pi^0$ reconstruction so this is not directly applicable. Furthermore, because reconstructing a $\pi^0$ requires finding two photons, the error on the efficiency for finding one photon should be approximately half that for finding a $\pi^0$. The above result translates to a $12.5\%$ uncertainy on the inefficiency to find a $\pi^0$. This then correponds to $6.25\%$ inefficiency to find a photon, which conservatively rounded to $10\%$. The weight used to implement these variations is
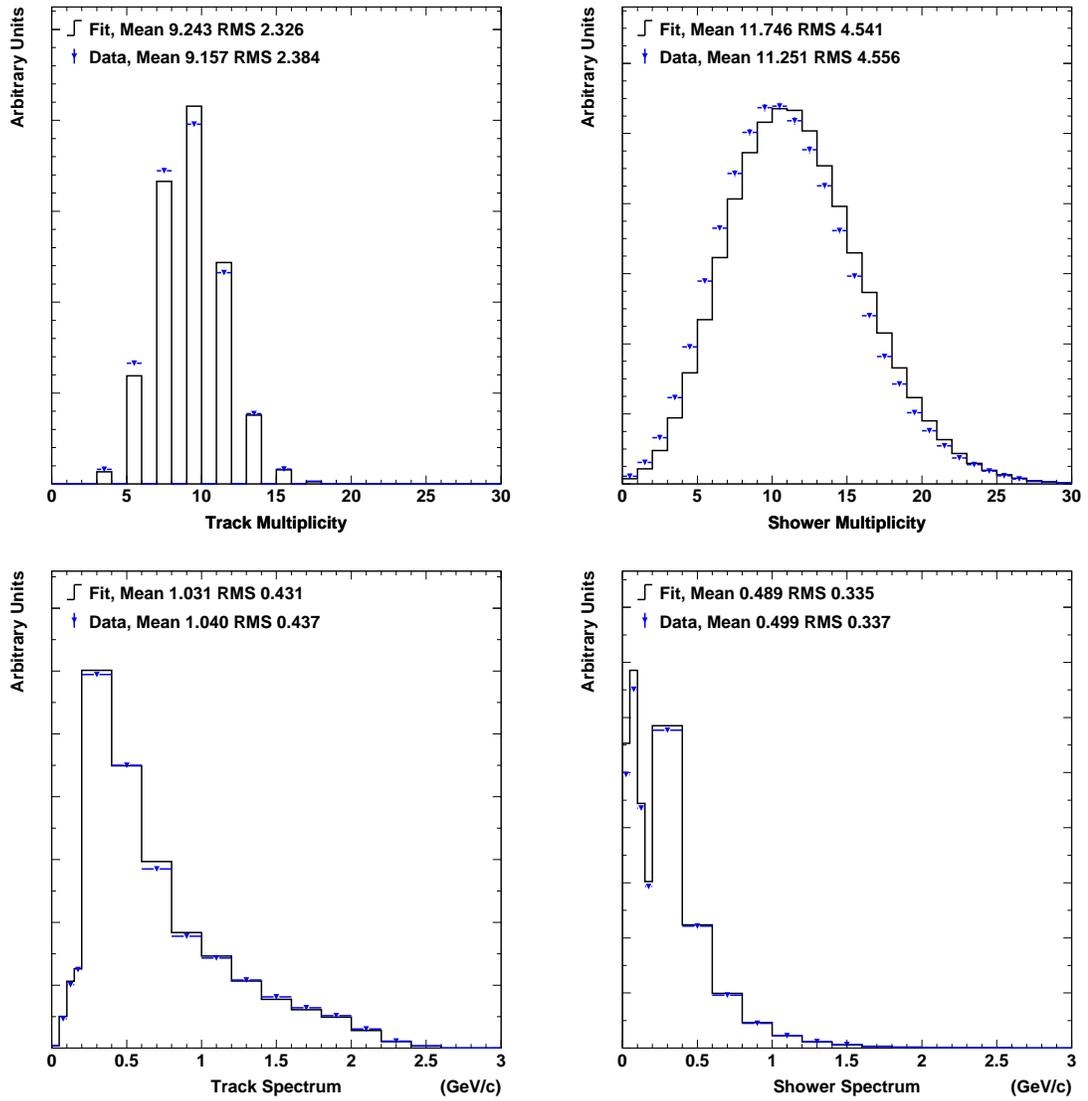
Figure 6.1: Track and Shower Spectra and Multiplicity Comparisions.
Comparision of the data and the simulation (Fit) of track (left) and shower (right) multiplicities are shown in the top row. Similar comparisons of the track (left) and shower (right) energy spectra are shown in the second row.

analogous to the track inefficiency case above.

### Fake Tracks

Fake tracks are tracks which are identified by the reconstruction software, but do not correspond to actual charged particles produced in the collisions. This includes such effects as curlers and ghosts, in which particles may be counted twice. Knowledge of the rate of fake tracks is limited. The only relevant measurement is that the number of $\tau$ to four track events observed is consistent with the Monte Carlo at the 10% level [60]. One basic test is to compare the distribution for the total charge of the event in data with the simulation. These agree at the $\approx 5\%$ level. The number of events with zero total charge is $\approx 45\%$. This indicates that at least 55% of event have least one error. This constrians the knowlegde of the error rate to be $\approx 5\%/55\% = 9\%$. This however includes both lost and fake tracks. Since fake tracks are less common than lost tracks, a large error in the fake track rate could be canceled by a smaller error in the track efficiency. A $\pm 10\%$ variation of the fake track rate is used for lack of a better constriant. The weight used for this variation is $w_e = c^{N_{\text{fake}}}$, where $N_{\text{fake}}$ is number of fake tracks in the event and $c$ is 1.1 or 0.9 for raising an lowering the number of fake tracks, respectively.

### Fake Showers

The rate of fake showers has been studied in two places. In $\tau^+\tau^- \to \ell^+ h^- \pi^0$ where $h^-$ is either a $\pi^-$ or a $K^-$, a comparison of the showers that are not matched to tracks and are not used in the reconstruction of the $\pi^0$ shows $10 - 20\%$ less showers in the data than in the simulation in the lowest energy bins (the agreement is better at higher energy) [63]. The other study of fake showers used $\gamma\gamma \to K_S^0 K_S^0$ where it is assumed that there are no real photons in the event. A 3% excess of showers compared to the simulation was found. A $\pm 10\%$ variation of the fake shower rate is used as a reasonable approximation of these results. The weights used to implement these variations is analogous to the fake track case above.

### Lepton Identification Efficiency

The lepton identification efficiency has been studied using Bhabha and $\mu$-pair events (see Sections 3.2.6 and 3.2.5, respectively). Unlike analyses which only use leptons, the effect of the lepton efficiency uncertainty is small compared to the other uncertainties. For the nominal fit the events are reweighted by the ratio of the measured efficiency to the Monte Carlo efficiency in order to apply the measurement. To assess the effect of the lepton efficiency uncertainty, this correction is removed. Figure 6.2 shows the difference between the efficiency used in the Monte Carlo simulation and the efficiency measured in the data. These are different by an amount somewhat larger than the statistical errors on the lepton efficiency measurements. The Monte Carlo simulation is therefore a
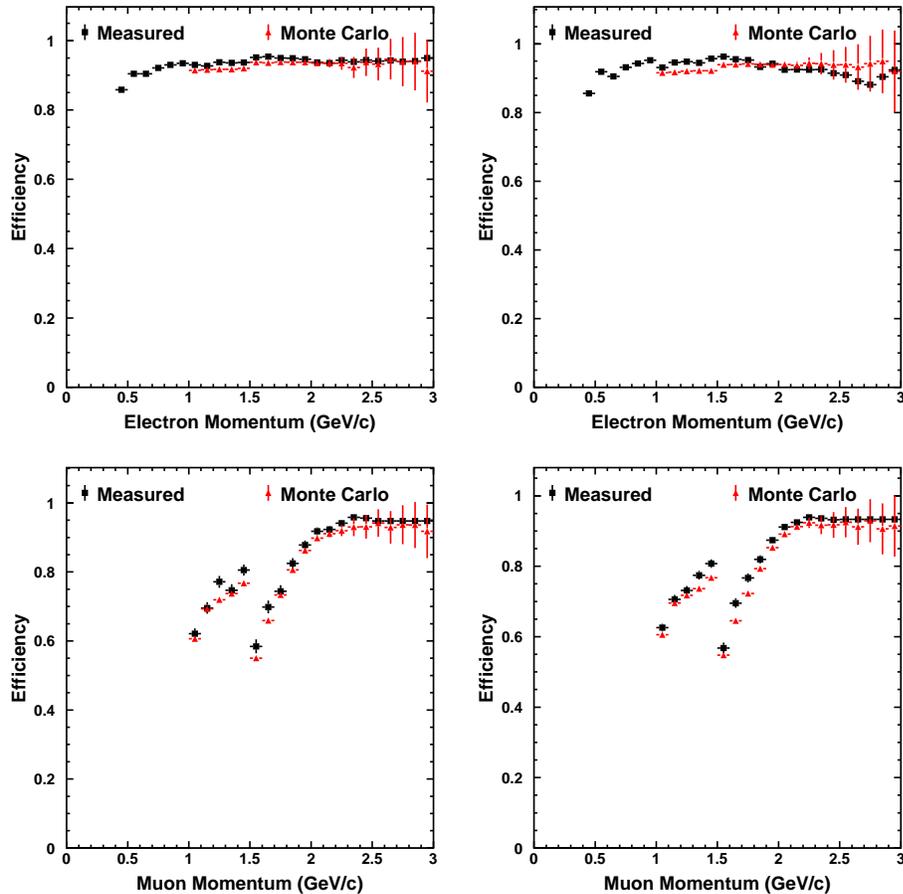
Figure 6.2: Comparison of the Monte Carlo Simulation and Measured Lepton Identification Efficiencies.
Electrons are shown on the top row and Muons on the bottom. CLEO II is shown on the left and CLEO II.V is shown on the right.

plausible alternative shape for the efficiency that should conservatively cover the uncertainty due to the lepton id efficiency.

### 6.1.3  Background Modeling

**Secondary Lepton Distribution**

Understanding the number and shape of the secondary leptons above 1 GeV is a very different problem than understanding the total number of secondary leptons discussed in Section 5.2.5. Only a very small tail of the secondary lepton distribution extends above 1 GeV, so the energy distribution of the leptons is very important. Because of the size of the uncertianty, the normalization of this component is not fixed, but is determined from the maximization of the likelihood.

To assess the effect of the uncertainty of the lepton energy distribution on the result, the model
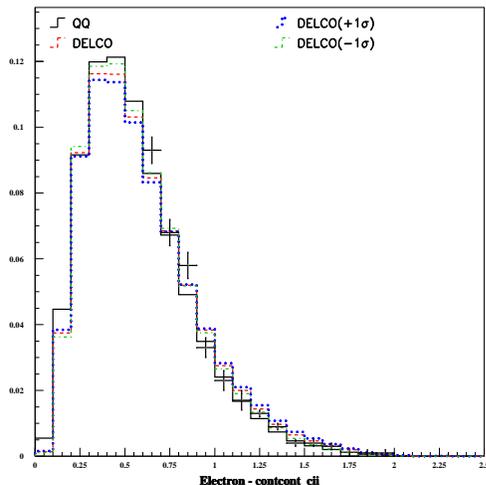
Figure 6.3: Models and Measurement of the Secondary Lepton Momentum Spectrum. Comparision of the CLEO measurement of the seconday lepton rate [54], the default simulation (QQ), and the CLEO/DELCO convolution model (see text) with $\pm 1\sigma$ variations.

of the distribution is reweighted to a model which is a convolution of the charm hadron spectra measured at CLEO and the lepton energy spectrum of $D$ semileptonic decays measured at the DELCO experiment[64]. Variations of this model based on the measurement uncertainties of the inputs are also used (they are labeled $\pm 1\sigma$ in the figures and tables). The uncertainty is due to the shape of the secondary lepton spectrum is taken to be the largest deviation from the nominal result of the three DELCO models. A comparison of the default simulation model (QQ), the three convolutions, and a CLEO measurement[54] is shown in Figure 6.3. The modified shapes are applied by weighting the seconadary lepton events by the ratios of the DELCO models to the QQ model as a function of the lepton energy.

**Continuum Normalization**

The normalization extracted from the off-resonance data agrees within 10% with the normalization calculated from the measured $e^+e^- \rightarrow q\overline{q}$ cross-section (see Section 5.2.8). The uncertainty of this normalization is therefore taken to be 10%. The shape of the model is validated using the off-resonance data, so no further uncertainty is assigned. As a cross-check of the shape, the $M_X^2$ distribution is reweighted to agree better with the off-resonance data. This shows a negligible variation in the fit results.

**Fake Leptons**

The uncertainty of measured lepton fake rates is the only important uncertainty related to the fake lepton background. The measurements include carefully evaluated statistical errors and systematic errors. These are propagated to the results by applying random Gaussian fluctutations to the rates used in the modeling of the fake lepton background. A set of twelve fluctuations is combined by taking the largest upward and downward variation of each result. To account for the possibility that the deviations of the measured fake rates from the true fake rates might be systematically high or low, the rates above and below 1.5 GeV are seperately varied by 10%.

### 6.1.4 Radiative Corrections

The knowledge of the radiative corrections is not complete. The PHOTOS splitting function described in Section 5.2.3 applies the same physics at $\mathcal{O}(\alpha)$ as the prescription of Atwood and Marciano [45]. The applicability of these calculation is not exact. They ignore the internal structure of the hadronic system. Richter-Wąs [46] has made a comparison of the PHOTOS and Atwood-Marciano proscriptions with an exact $\mathcal{O}(\alpha)$ calculation of the radiative corrections to the $B^- \to D^0 \ell^- \overline{\nu}$ differential decay rate and has found agreement at the 20% and 30% level, respectively. Because this result only applies to one of the modes, the uncertainty of the PHOTOS calculation is conservatively estimated to be $\pm 50\%$. This uncertainty is propagated to results by repeating the fit using a Monte Carlo sample without PHOTOS or the after detector simulation addition of radiative corrections. Half of the shift from the nominal is then used as the uncertainty.

### 6.1.5 Summary of the Detector Systematics

The shifts from nominal of the results for each of the variations are listed in Table 6.1. These are combined into a single detector systematic by summing in quadrature the largest variations in each category. The combined result is also shown in the table.

## 6.2 Model Dependence

### 6.2.1 $B \to Dl\overline{\nu}$ and $B \to D^* l\overline{\nu}$

The model dependence of the results due to the uncertainty of the $B \to Dl\overline{\nu}$ and $B \to D^* l\overline{\nu}$ form factors is assessed by varying the parameters of the form factors by their $(1\sigma)$ measurement errors. These measurements are summarized in Table 5.2. The uncertainty of the slope $B \to D^* l\overline{\nu}$ has been inflated, because of the large spread in the measurements [1]. Instead of the PDG value $\rho_{A_1}^2 = 1.51 \pm 0.13$, the value $\rho_{A_1}^2 = 1.51 \pm 0.20$ is used as the nominal measurement, before the modifications due to the curvature choice are applied (see Section 5.2.4).

Table 6.1: Detector Systematics of the Branching Fraction Results.
The top three rows are the branching fraction, its statistical error, and the detector related systematic error. The next two rows are the statistical and systematic errors as percentages of the branching fraction. The remaining rows are the individual contributions to the detector systematic as percentages of the branching fractions. Entries with /'s have the effect of raising the quantity on the left and lowering it on the right.

| | $B \to Dl\overline{\nu}$ | $B \to D^*l\overline{\nu}$ | $B \to D^{**}l\overline{\nu}$ | $X_c$ Nonres | $B \to X_ul\overline{\nu}$ | Sum |
|---|---|---|---|---|---|---|
| Branching Fraction $(\times 10^{-2})$ | 1.919 | 6.374 | 1.509 | 0.695 | 0.115 | 10.612 |
| Stastical Error | 0.076 | 0.060 | 0.073 | 0.070 | 0.008 | 0.287 |
| Detector Systematics | 0.193 | 0.650 | 0.302 | 0.245 | 0.029 | 1.089 |
| Fractional Statistical | 3.9 | 0.9 | 4.9 | 10.1 | 7.0 | 2.7 |
| Fractional Detector | 10.1 | 10.2 | 20.0 | 35.2 | 25.5 | 10.3 |
| Lepton Fake Rate | 0.8 / -0.6 | 0.4 / -0.3 | 2.7 / -1.8 | 6.9 / -2.8 | 12.6 / -6 .3 | 0.8 / -0.2 |
| DELCO $b \to c \to \ell$ shape | -0.6 | 0.1 | -2.8 | -1.7 | 0.3 | -0.5 |
| DELCO $+1\sigma$ $b \to c \to \ell$ shape | -0.8 | 0.1 | -4.3 | -2.0 | 0.1 | -0.8 |
| DELCO $-1\sigma$ $b \to c \to \ell$ shape | 0.2 | 0.1 | 0.6 | -0.7 | 1.7 | 0.1 |
| Continuum Norm $\pm 10\%$ | -3.6 / 2.6 | 0.4 / -0.5 | -2.9 / 6.7 | 1.7 / -6.3 | -4.1 / 4.6 | -0.8 / 0.7 |
| $\mu$ Fakes, $E_\ell < 1.5$ GeV, $\pm 10\%$ | -1.4 / 1.2 | -0.3 / 0.3 | 1.7 / -2.3 | -11.0 / 11.4 | -0.4 / 0.3 | -0.9 / 0.8 |
| $\mu$ Fakes, $E_\ell > 1.5$ GeV, $\pm 10\%$ | 0.2 / -0.2 | -0.2 / 0.2 | -1.4 / 1.4 | 1.4 / -1.5 | -4.9 / 5.1 | -0.3 / 0.3 |
| $\mathcal{B}(b \to c \to \ell)$ $\pm 10\%$ | 1.9 / -2.0 | 2.6 / -2.5 | 6.6 / -6.3 | -0.8 / 0.7 | -6.1 / 6.4 | 2.7 / -2.6 |
| $\mathcal{B}(b \to$ baryons$)$ $\pm 20\%$ | 1.7 / -0.9 | 3.9 / -3.0 | 4.9 / -3.9 | 1.7 / -1.0 | -0.1 / 0.9 | 3.4 / -2.6 |
| # $K_L^0$ | 2.4 / -2.4 | 2.6 / -2.5 | 7.0 / -6.7 | -1.3 / 1.2 | -7.2 / 7.7 | 2.8 / -2.7 |
| Track Efficiency | -4.8 / 4.9 | -5.8 / 6.0 | -9.6 / 10.8 | 0.9 / -1.7 | -1.7 / 1.7 | -5.6 / 5.9 |
| # Fake Tracks | 1.9 / -1.7 | 2.3 / -2.0 | 2.6 / -2.3 | -3.9 / 3.5 | 0.3 / -0.2 | 1.9 / -1.6 |
| Shower Efficiency | -2.2 / 2.5 | -1.4 / 1.7 | -3.9 / 4.8 | 1.8 / -1.4 | -3.8 / 3.7 | -1.7 / 2.1 |
| # Fake Showers | -3.1 / 2.7 | -0.5 / 0.2 | 4.5 / -3.9 | -28.2 / 27.5 | -14.7 / 15.3 | -2.2 / 2.0 |
| Force Trk Multiplicity | 2.5 | 2.0 | 3.1 | 2.2 | -3.0 | 2.2 |
| Force Shwr Multiplicity | 2.2 | 3.4 | 3.8 | 8.9 | -3.4 | 3.5 |
| Final State Radiation | -0.1 | -3.3 | -0.2 | 9.0 | -7.2 | -1.5 |
| Lepton Efficiency | 4.3 | 2.0 | 4.1 | 5.1 | -3.7 | 2.8 |

Because of the discrepancy between the data and the fit in the 4.0 to 7.0 GeV$^2$/$c^4$ region of the $q^2$ distribution, the curvature used in the nominal fit is $1/2 \pm 1/2$ of the theorectically predecited value. The correlation of the curvature and the slope measurements leads to a correction to the nominal values of $\rho_D^2$ and $\rho_{A_1}^2$ (see Section 5.2.4). The resulting four variations for the $B \to Dl\overline{\nu}$ Isgur-Wise function and four more variations for the $B \to D^*l\overline{\nu}$ form factor $\rho_{A_1}^2$ are shown in Table 6.2.

Table 6.2: Variations of the $B \to Dl\overline{\nu}$ and $B \to D^*l\overline{\nu}$ Form Factors.

|  | $B \to Dl\overline{\nu}$ | | $B \to D^*l\overline{\nu}$ | |
| --- | --- | --- | --- | --- |
| Variation | $\rho_D^2$ | $c_D$ | $\rho_{A_1}^2$ | $c_{A_1}$ |
| Nominal | 0.80 | 0.56 | 1.20 | 0.7 |
| Raise slope | 1.13 | 0.67 | 1.48 | 0.8 |
| Lower slope | 0.48 | .44 | 0.91 | 0.58 |
| Raise curvature | 1.22 | 1.12 | 1.51 | 1.39 |
| Lower curvature | 0.39 | 0.0 | 0.88 | 0.0 |

Because the measurements of $R_1$ and $R_2$ are highly correlated ($C_{R_1R_2} = 0.814$), the correlation must be taken into account in the determination of the model dependence. The covariance matrix of the two measurments is diagonalized and the eigen vectors and values are found. The eigen vectors are used to make correlated variation of $R_1$ and $R_2$, with the scale set by the square root of the eigen values of the matrix.

The variations of the $B \to Dl\overline{\nu}$ and $B \to D^*l\overline{\nu}$ form factors are illustrated in Figures 6.4 and 6.5 respectively.

## 6.2.2 $B \to D^{**}l\overline{\nu}$

The dependence of the results on the $B \to D^{**}l\overline{\nu}$ model is quantified by replacing the nominal $B \to D^{**}l\overline{\nu}$ model (ISGW2) with a calculation of the form factors to order $\Lambda_{QCD}/m_D$ using HQET [31]. The calculation has a significant number of unknown parameters, for which model dependent assumptions are used (as suggested by the authors) [31]. The main difference between this model and the ISGW2 model used in the nominal fit is in the ratios of the broad to narrow $B \to D^{**}l\overline{\nu}$ modes. Because ISGW2 obeys the constraints of the heavy quark symmetries, the two models are similar in their $\cos\theta_{W\ell}$ and $q^2$ distribution.

The $q^2$ slopes are also varied in an ad-hoc way, reweighting the $B \to D^{**}l\overline{\nu}$ events by $(w - 0.5)$ and $(-w + 2)$. The relatively small effect of this variation on the results gives confidence that the results are not very depenendent on the shape of the $B \to D^{**}l\overline{\nu}$ form factors which have not been measured experimentally.

The variations of $B \to D^{**}l\overline{\nu}$ form factors are illustrated in Figure 6.6.
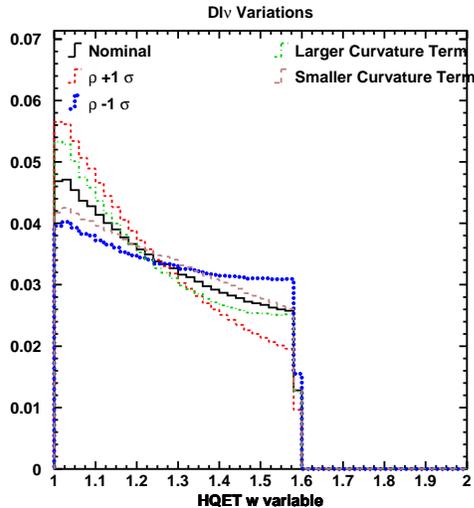
Figure 6.4: Variations of the $B \to Dl\overline{\nu}$ Form Factor, $\xi(w)_D$, Used in the Model Dependence Determination.

The distributions of the differential decay rate as a function of the HQET variable $w = v \cdot v' = (m_B^2 + m_D^2 - q^2)/2m_B m_D$, but have the leading spin and phase space factors divided out, so that they correspond to the Isgur-Wise function, $\xi(w)_D$.

### 6.2.3   Nonresonant $B \to X_c l\overline{\nu}$

The Goity and Roberts model is used in the nominal fit to describe the shape of the nonresonant $B \to X_c l\overline{\nu}$ component. Although this is only a model of the $B \to D^{(*)}\pi l\nu$ contribution to the full nonresonant component, it properly includes the helicity structure and $q^2$ distribution constraints implied by the structure of the weak currents and heavy quark symmetry. However, there is no reason to expect the Goity and Roberts model to accurately model the hadronic mass distribution. To assess corresponding model dependence, the hadronic mass distribution is reweighted into nine different Gaussians. The Gaussians have means of either 2.0, 2.75, or 3.5 GeV/$c^2$ and variances of either 0.25, 0.75, or 1.25 GeV$^2$/$c^4$ and are truncated at the low end so that $m_{X_c} \geq m_D + m_\pi$. The maximum shift of the results due to the Gaussians is used as the contribution to the total model dependence.

The $q^2$ slope of the nonresonant $B \to X_c l\overline{\nu}$ modes is also varied with the same weights that are used to vary the $q^2$ slope of the $B \to D^{**}l\overline{\nu}$ mode.

The variations of nonresonant $B \to X_c l\overline{\nu}$ form factors are illustrated in Figure 6.7.

### 6.2.4   $B \to X_u l\overline{\nu}$

The $B \to X_u l\overline{\nu}$ simulation is varied by using seperate sets of Monte Carlo events which are generated based on different models. The models used are:

- **ISGW2** contains only low mass resonances, it is described in [51].
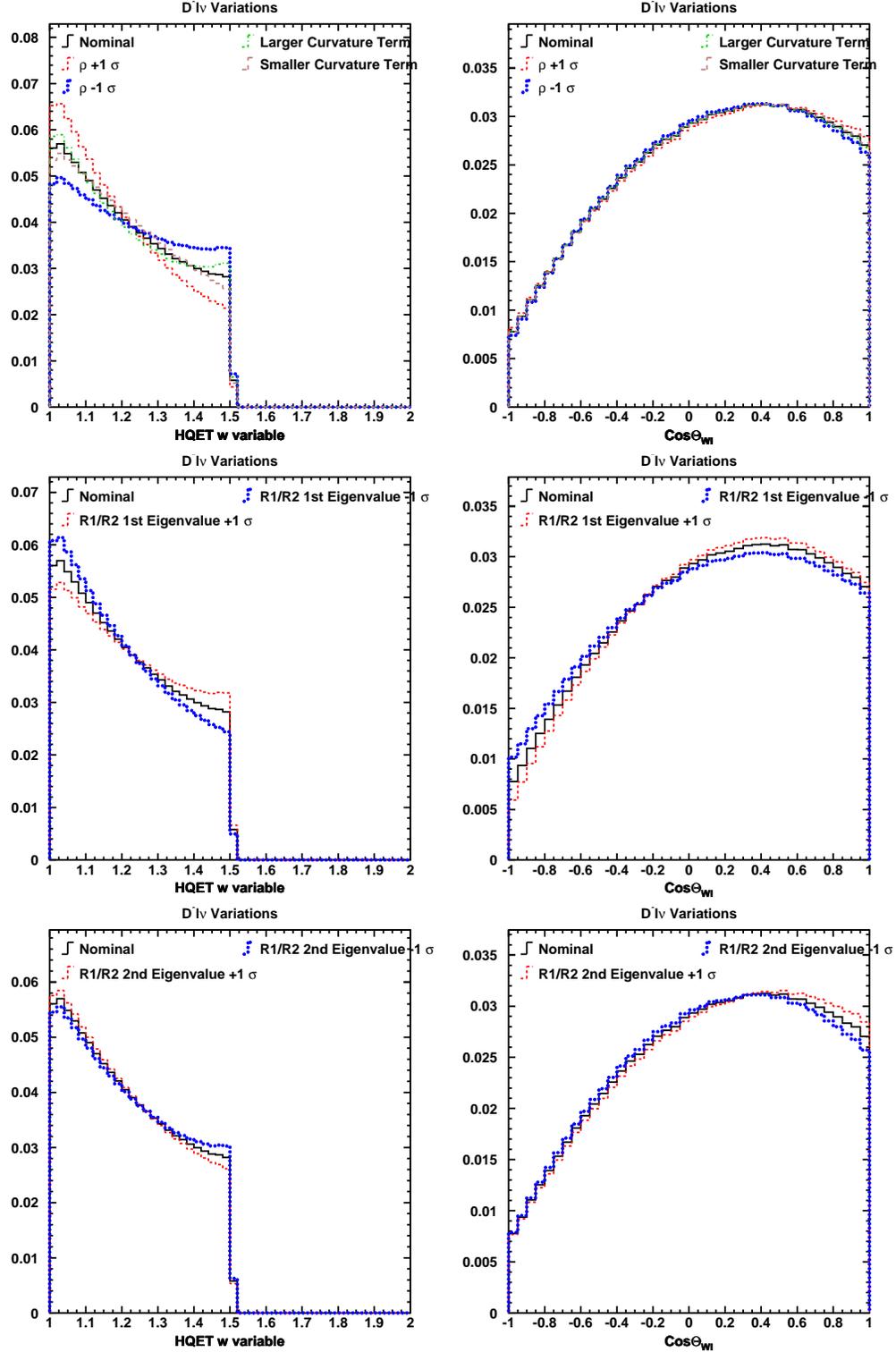
Figure 6.5: Variations of the $B \to D^* l \overline{\nu}$ Form Factors Used in the Model Dependence Determination. The distributions of the differential decay rate as a function of the HQET variable $w = v \cdot v' = (m_B^2 + m_{D^*}^2 - q^2)/2 m_B m_{D^*}$ (left) and $\cos \theta_{W\ell}$ (right). The $w$ distributions have the leading spin and phase space factors divided out, so that they correspond to the Isgur-Wise function, $\xi(w)_{D^*}$.
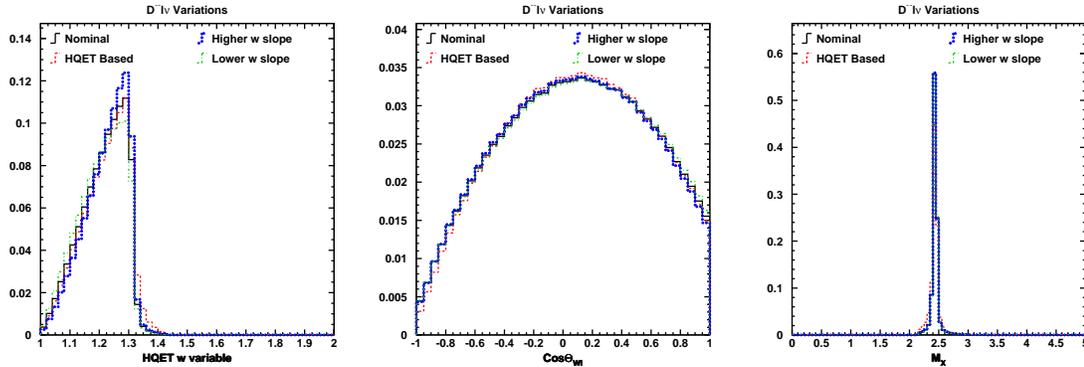
Figure 6.6: Variations of the $B \to D^{**}l\overline{\nu}$ Form Factors Used in the Model Dependence Determination.
The differential decay rate as a function of the HQET variable $w = v \cdot v' = (m_B^2 + m_{D^{**}}^2 - q^2)/2m_B m_{D^{**}}$ (left), $\cos\theta_{W\ell}$ (center), and $M_x$ (right).

- **InclGen** combines the ISGW2 model for the low mass resonances and a nonresonant contribution for the higher mass components. The mass distribution is designed to approximate the HQET inclusive prediction [65].

- **BSG** is similar to the **InclGen** model except that the mass distribution is based on the $B \to X_s\gamma$ spectral function [20, 13].

- **X0** is a purely non-resonant model which distributes the hadronic mass according to the $B \to X_s\gamma$ spectral function prediction.

- **Hi** is similar to **X0** but has the values of $\overline{\Lambda}$ and $\lambda_1$ increased, so that the hadronic systems are heavier and the distribution of masses is broader.

- **Lo** is similar to **X0** but has the values of $\overline{\Lambda}$ and $\lambda_1$ decreased, so that the hadronic systems are lighter and the distribution of masses is narrower.

To further study the model dependence the ISGW2 model is reweighted to have harder and softer $q^2$ spectra and higher and lower mean $\cos\theta_{W\ell}$.

The $M_X$, $E_\ell$, $q^2$, and $\cos\theta_{W\ell}$ projections of the various $B \to X_u l\overline{\nu}$ models are shown in Figure 6.8.

### 6.2.5 Summary of the Model Dependence

The shifts for each of the model variations are listed in Table 6.3. Like the dectector systematic variations, the shifts are summed in quadrature to get a single overall model dependence, which is also shown in the table. The $B \to X_u l\overline{\nu}$ model dependence of the $B \to X_u l\overline{\nu}$ branching fraction is discussed in detail in Chapter 8.
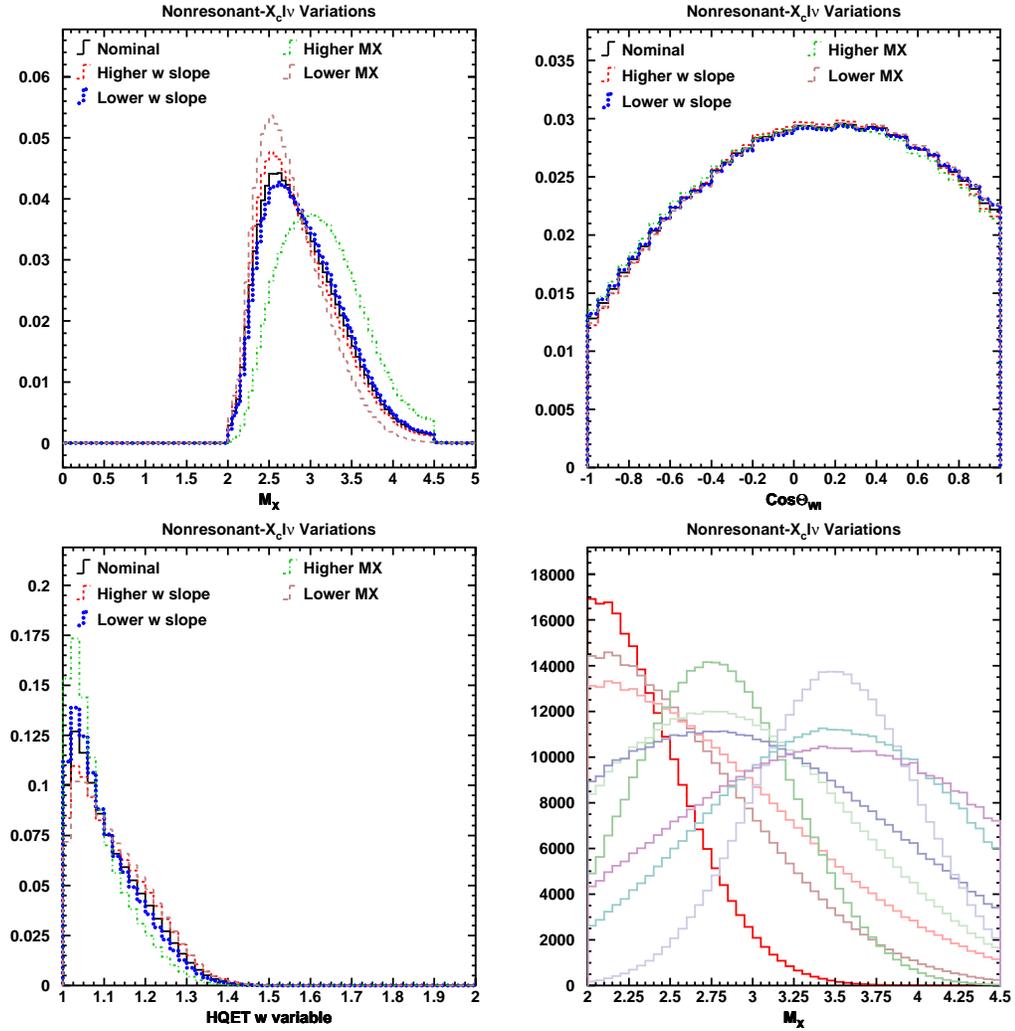
Figure 6.7: Variations of the Nonresonant $B \to X_c l\overline{\nu}$ Form Factors Used in the Model Dependence Determination.

The $M_X$ (top left), $\cos\theta_{W\ell}$ (top right), and HQET $w$ (bottom left) distributions of the differential decay rate for the nominal model, the variations of the w slope, and two polynomial reweightings of the $M_X$ distribution (which are used to show the correlations between the $M_X$ distribution and the other variables). The set of Gaussian $M_X$ distributions are shown in the bottom right plot. These are variations of the same distribution as is shown in the upper left plot.
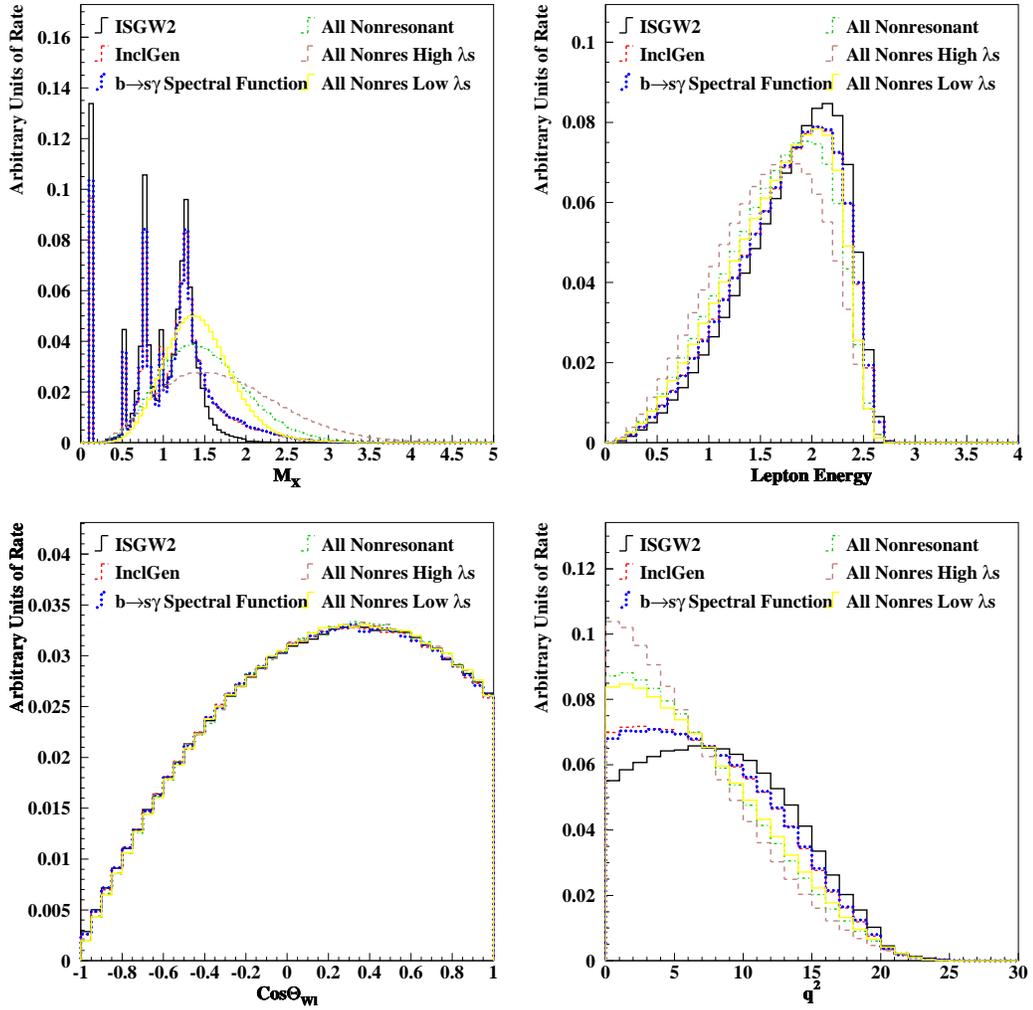
Figure 6.8:  $M_X$, $E_\ell$, $q^2$, and $\cos\theta_{W\ell}$ Projections of the Various $B \to X_u l\overline{\nu}$ Models.

Table 6.3: Model Dependence of the Branching Fraction Results.

The top three rows are the branching fraction, its statistical error, and its model dependence uncertainty. The next two rows are the statistical and systematic errors as percentages of the branching fraction. The remaining rows are the individual contributions to the detector systematic as percentages of the branching fraction. Entries with /'s have the effect of raising the quantity on the left and lowering it on the right.

| | $B \to Dl\overline{\nu}$ | $B \to D^*l\overline{\nu}$ | $B \to D^{**}l\overline{\nu}$ | $X_c$ Nonres | $B \to X_ul\overline{\nu}$ | Sum |
|---|---|---|---|---|---|---|
| Branching Fraction ($\times 10^{-2}$) | 1.919 | 6.374 | 1.509 | 0.695 | 0.115 | 10.612 |
| Stastical Error | 0.076 | 0.060 | 0.073 | 0.070 | 0.008 | 0.287 |
| Model Dependence | 0.743 | 0.859 | 0.516 | 0.616 | 0.231 | 0.746 |
| Fractional Statistical | 3.9 | 0.9 | 4.9 | 10.1 | 7.0 | 2.7 |
| Fractional Model Dep. | 38.7 | 13.5 | 34.2 | 88.7 | 201.5 | 7.0 |
| $B \to Dl\overline{\nu}$ $\rho$ param | 6.4 / -5.2 | -2.6 / 1.8 | 4.6 / -2.0 | -5.4 / 2.3 | 0.1 / 0.2 | -0.1 / 0.0 |
| $B \to Dl\overline{\nu}$ c param | 1.4 / -1.1 | -0.5 / 0.3 | 0.8 / -0.4 | -0.5 / 0.2 | 0.1 / -0.0 | 0.1 / -0.1 |
| $B \to D^*l\overline{\nu}$ $\rho$ param | 32.4 / -32.1 | -11.4 / 11.0 | 17.9 / -15.2 | -18.0 / 17.0 | 6.3 / -4.8 | 0.4 / -0.3 |
| $B \to D^*l\overline{\nu}$ $c_{A_1}$ param | -1.1 / 0.8 | 0.8 / -0.7 | -1.0 / 0.6 | 2.9 / -2.0 | 1.1 / -1.0 | 0.3 / -0.3 |
| $B \to D^*l\overline{\nu}$ R1/R2 1st eig-vec | -13.3 / 14.9 | 3.5 / -3.7 | -10.0 / 10.1 | 16.7 / -15.3 | -4.5 / 5.7 | -0.7 / 1.0 |
| $B \to D^*l\overline{\nu}$ R1/R2 2nd eig-vec | 9.9 / -10.4 | -4.0 / 4.2 | 3.2 / -3.4 | 0.9 / -0.7 | -3.3 / 3.5 | -0.1 / 0.2 |
| $B \to D^{**}l\overline{\nu}$ HQET model | -4.0 | 0.5 | 2.3 | 0.1 | 0.2 | -0.1 |
| $B \to D^{**}l\overline{\nu}$ w slope | -2.7 / 3.1 | 0.9 / -0.8 | -2.7 / 0.6 | 13.6 / -9.8 | -0.5 / 0.2 | 0.5 / -0.5 |
| $B \to X_cl\overline{\nu}$ NonRes w slope | -0.0 / -0.1 | 0.1 / -0.0 | -6.3 / 3.6 | 17.4 / -9.9 | -0.2 / 0.1 | 0.3 / -0.2 |
| $B \to X_cl\overline{\nu}$ Nonresonant Mass Dependence | | | | | | |
| Polynomial | -3.1 / 1.0 | 0.7 / -0.4 | 12.7 / -4.9 | -7.1 / -0.2 | 0.2 / 0.2 | 1.2 / -0.8 |
| Gauss $\mu = 2.0$ $\sigma^2 = .25$ | -1.7 | -2.9 | 1.6 | -13.5 | 2.5 | -2.7 |
| Gauss $\mu = 2.0$ $\sigma^2 = .75$ | 0.5 | -2.2 | -2.3 | 13.7 | 1.7 | -0.7 |
| Gauss $\mu = 2.0$ $\sigma^2 = 1.25$ | -0.0 | -1.7 | 0.8 | 24.8 | 1.4 | 0.7 |
| Gauss $\mu = 2.75$ $\sigma^2 = .25$ | -0.3 | -0.3 | 1.2 | -5.4 | 0.3 | -0.4 |
| Gauss $\mu = 2.75$ $\sigma^2 = .75$ | -0.9 | -0.7 | 4.9 | 22.0 | 0.8 | 1.6 |
| Gauss $\mu = 2.75$ $\sigma^2 = 1.25$ | -1.1 | -0.7 | 6.1 | 43.9 | 0.9 | 3.1 |
| Gauss $\mu = 3.5$ $\sigma^2 = .25$ | -6.2 | 1.0 | 25.6 | 4.3 | 0.5 | 3.4 |
| Gauss $\mu = 3.5$ $\sigma^2 = .75$ | -4.1 | 0.3 | 17.6 | 60.4 | 0.8 | 5.9 |
| Gauss $\mu = 3.5$ $\sigma^2 = 1.25$ | -3.5 | -0.0 | 15.2 | 81.8 | 1.0 | 6.9 |
| $B \to X_ul\overline{\nu}$ Model Dependence | | | | | | |
| $B \to X_s\gamma$ based | -0.9 | -0.1 | -0.2 | 0.1 | 54.9 | 0.3 |
| InclGen | -1.0 | -0.0 | -0.3 | 0.2 | 67.3 | 0.5 |
| All Nonres | -1.5 | -1.2 | -0.9 | 0.2 | 150.2 | 0.5 |
| All Nonres High Mass | -0.1 | -1.5 | -2.3 | -6.7 | 201.2 | 0.5 |
| All Nonres Low Mass | -2.7 | -1.0 | 0.6 | 0.9 | 120.2 | 0.4 |
| Hard $q^2$ | 3.0 | -0.1 | -1.0 | 0.6 | -19.4 | 0.2 |
| Soft $q^2$ | -3.0 | 0.2 | 0.8 | -0.6 | 17.2 | -0.2 |
| High $\cos\theta_{W\ell}$ | 0.3 | 0.1 | -0.2 | 0.2 | -4.4 | 0.0 |
| Low $\cos\theta_{W\ell}$ | -0.3 | -0.1 | 0.2 | -0.1 | 4.8 | -0.0 |
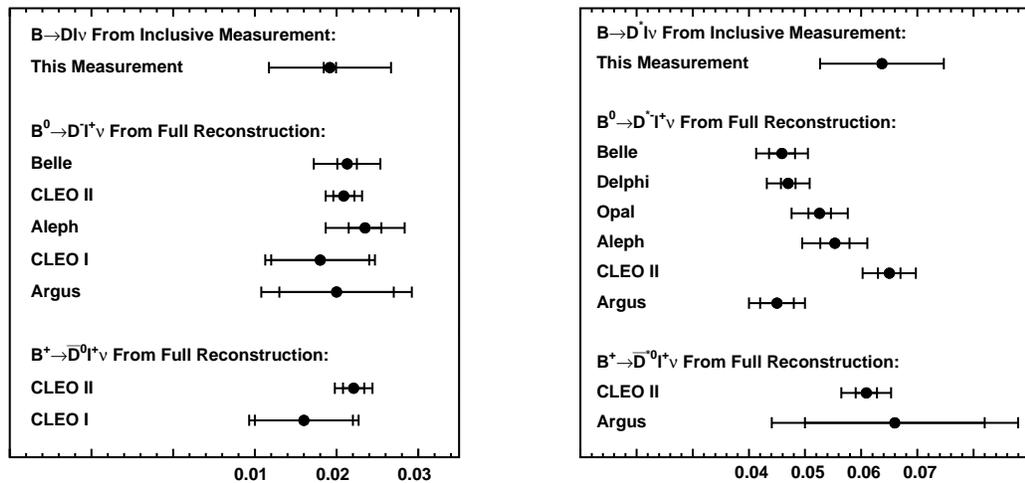
Figure 6.9: Comparison of the $\mathcal{B}(B \to Dl\overline{\nu})$ and $\mathcal{B}(B \to D^*l\overline{\nu})$ Measurements with Results that Use the Full Recontruction Method.

## 6.3 Branching Fractions with Systematics

The branching fractions with detector systematic uncertianties and model dependence are shown Table 6.4. The measurements of both $\mathcal{B}(B \to Dl\overline{\nu})$ and $\mathcal{B}(B \to D^*l\overline{\nu})$ are consistent with the variety of full reconstruction measurements available (see Figure 6.9). The inclusive rate measurement, $\mathcal{B}(B \to Xl\overline{\nu})$, is also consistent with the measurements that use the correlations between leptons to supress the secondary lepton background (see Table 6.5). Although the uncertainties of the individual branching fractions are large, the measurements are very correlated and the uncertainty on the moments, which is one of the two primary goals of this analysis not as large. The model that results from the fit is interesting because the exclusive modes sum up to the full inclusive rate by construction, unlike exclusive measurements which do not add up to the total rate.

Table 6.4: Branching Fraction with All Uncertainty Estimates.

| Mode | Branching Fraction $\pm$ Statistical $\pm$ Detector Sys. $\pm$ Model Dep. ($\times 10^{-2}$) |
|---|---|
| $B \to Dl\overline{\nu}$ | $1.919 \pm 0.076 \pm 0.193 \pm 0.743$ |
| $B \to D^*l\overline{\nu}$ | $6.374 \pm 0.060 \pm 0.650 \pm 0.859$ |
| $B \to D^{**}l\overline{\nu}$ | $1.509 \pm 0.073 \pm 0.302 \pm 0.516$ |
| Nonresonant $B \to X_c l\overline{\nu}$ | $0.695 \pm 0.070 \pm 0.245 \pm 0.616$ |
| $B \to X_u l\overline{\nu}$ | $0.115 \pm 0.008 \pm 0.029 \pm 0.231$ |
| Sum | $10.612 \pm 0.287 \pm 1.089 \pm 0.746$ |

Table 6.5: Comparison of Inclusive $\mathcal{B}(B \to Xl\overline{\nu})$ Rate with Previous Measurements.

| | $\mathcal{B}(B \to Xl\overline{\nu}) \pm$ Statistical $\pm$ Systematic $(\times 10^{-2})$ |
|---|---|
| This Measurement | $10.61 \pm 0.29 \pm 1.37$ |
| CLEO II [4] | $10.49 \pm 0.17 \pm 0.43$ |
| Babar [3] | $10.87 \pm 0.18 \pm 0.30$ |
| Belle [5] | $10.90 \pm 0.12 \pm 0.49$ |

## 6.3.1 Comparison of Electron and Muon Results

Because the electron and muon masses are much smaller than the lepton energies involved in $B$ decay, it is a good approximation to consider the leptons as massless. In this limit, the branching fractions measured using the electron-only sample should be the same as those measured in the muon-only sample. There are systematic differences between the two values: radiative corrections are a much larger effect for electrons and for muons, the efficiencies are lower, and the background from fake leptons is much larger. These difference can lead the results to be different by more than just the most directly related systematics uncertainties. This is because the fit intrinsically extrapolates from regions of higher sensitivity to regions of lower sensitivity. For instance the fake lepton background is large at low $q^2$, so for muons where this background is large, the contribution to any $B \to X_c l\overline{\nu}$ mode measured primarily at high $q^2$ and extrapolated down to low $q^2$. This extraplotion will depend on the model. For electrons, the low $q^2$ region has much less background, so the extrapolation uncertainty will be different. To quantify these effects, the comparison between electrons and muons is made by summing in quadrature the difference between the electron and muon results for each variation, the same way the branching fraction uncertainties are combined. In Table 6.6, the electron and muon results with their seperate systematic errors are shown along with the difference between the electron and muon results with the uncertainties calculated as described above.

Table 6.6: Comparison of the Branching Fraction Measured with the Electron and Muon Samples.

| Mode | Branching Fraction $\pm$ Statistical $\pm$ Detector Sys. $\pm$ Model Dep. $(\times 10^{-2})$ | | |
|---|---|---|---|
| | $e$ | $\mu$ | $\Delta(e-\mu)$ |
| $B \to Dl\overline{\nu}$ | $1.89 \pm 0.10 \pm 0.17 \pm 0.75$ | $1.92 \pm 0.11 \pm 0.22 \pm 0.73$ | $-0.03 \pm 0.15 \pm 0.16 \pm 0.09$ |
| $B \to D^*l\overline{\nu}$ | $6.47 \pm 0.08 \pm 0.64 \pm 0.87$ | $6.29 \pm 0.08 \pm 0.66 \pm 0.82$ | $0.17 \pm 0.12 \pm 0.22 \pm 0.09$ |
| $B \to D^{**}l\overline{\nu}$ | $1.63 \pm 0.09 \pm 0.46 \pm 0.55$ | $1.34 \pm 0.11 \pm 0.33 \pm 0.48$ | $0.29 \pm 0.14 \pm 0.26 \pm 0.18$ |
| Nonresonant | $0.66 \pm 0.08 \pm 0.23 \pm 0.79$ | $0.68 \pm 0.11 \pm 0.31 \pm 0.44$ | $-0.02 \pm 0.14 \pm 0.21 \pm 0.43$ |
| $B \to X_u l\overline{\nu}$ | $0.11 \pm 0.01 \pm 0.02 \pm 0.21$ | $0.11 \pm 0.01 \pm 0.03 \pm 0.24$ | $0.00 \pm 0.01 \pm 0.02 \pm 0.03$ |
| Sum | $10.77 \pm 0.38 \pm 1.07 \pm 0.93$ | $10.37 \pm 0.43 \pm 1.11 \pm 0.50$ | $0.40 \pm 0.58 \pm 0.36 \pm 0.44$ |

# Chapter 7  Moments and HQET Parameters

Moments of the measured differential decay rate can be directly compared to the HQET and OPE calculations. They are used to extract the HQET parameters $\overline{\Lambda}$ and $\lambda_1$, which are properties of the $B$ mesons and can be used to extract $|V_{cb}|$ from measurements of the inclusive decay rate $\mathcal{B}(B \rightarrow Xl\overline{\nu})$. The theoretical calculations are reviewed in Section 2.4.

In this chapter the following moments are calculated from the fit results:

- $\langle M_X^2 - \overline{M}_D^2 \rangle$ and $\langle (M_X^2 - \langle M_X^2 \rangle)^2 \rangle$, the first and second hadronic mass squared moments,

- $\langle q^2 \rangle$ and $\langle (q^2 - \langle q^2 \rangle)^2 \rangle$, the first and second lepton pair invariant mass squared moments,

- $\langle E_\ell \rangle$ and $\langle (E_\ell - \langle E_\ell \rangle)^2 \rangle$, the first and second lepton energy moments, and

- $R0 \equiv (\int_{E_\ell=1.7} d\Gamma/dE_\ell \, dE_\ell)/(\int_{E_\ell=1.5} d\Gamma/dE_\ell \, dE_\ell)$.

Although only two moments are needed to extract the parameters $\overline{\Lambda}$ and $\lambda_1$, measuring more moments can be used as a test of the theory. In this analysis the variation of the $\langle M_X^2 - \overline{M}_D^2 \rangle$ moment as a function of the lepton energy cut is compared to the prediction and is found to agree within the experimental and theoretical uncertainties. The lepton energy moments (including $R0$) have been measured with high precision in analyses which do not use neutrino reconstruction, but moments of $M_X^2$ and $q^2$ can only be measured using neutrino reconstruction or related techniques.

It is important to note that although the branching fraction measurements have large model dependence uncertainties, the fit is constructed so that they reproduce the inclusive differential decay distribution. Therefore, as the models of the exclusive components are varied, the description of the inclusive spectrum generated by summing up the exclusive components will only vary as a second order effect; the shape will be constrained by the data. The moments extracted from the sum of the models used in the fit will therefore have a smaller model dependence than the individual components. This also means that it is essential to calculate the moments for each model variation separately and then use the shifts from the nominal values to assess the model dependence (as opposed to propagating errors using no correlations). The fact that the inclusive shape of the data is used to extract the branching fraction used to calculate the moments is key to getting precise moments. If one were to attempt this calculation with a set of exclusive measurements, the result would be limited by the lack of measurements of the high recoil mass, high multiplicity modes and by the fact that the measurements would not be constrained to sum up to the inclusive branching fraction.

## 7.1   Calculating Moments

The moments are calculated from the branching fraction results, which combined with the physics models used in the fit form a description of the differential decay rate:

$$\frac{d\Gamma(B \to X_c l\overline{\nu})}{d\vec{x}} = \sum_{m \in \text{modes}} \frac{d\Gamma_m}{d\vec{x}} = \Gamma(B) \sum_{m \in \text{modes}} \mathcal{B}_m \frac{1}{\Gamma_m} \frac{d\Gamma_m}{d\vec{x}}, \tag{7.1}$$

where $\vec{x}$ is any complimentary set of three kinematic variables (e.g $E_\ell$, $E_\nu$, and $q^2$), and $\Gamma_m$ and $\mathcal{B}_m$ are the decay rates and branching fractions for the $b \to c$ semileptonic modes. The fit measures sizes of the contributions to the data sample of each mode, $\mathcal{B}_m$, in terms of a set of assumed shapes $1/\Gamma_m d\Gamma_m/d\vec{x}$. The shapes of the individual modes are not measured. The histograms used in the fit are the shapes $1/\Gamma_m d\Gamma_m/d\vec{x}$ with the detector efficiency and resolution applied. The moments are calculated using the branching fractions from the fit results and the shapes of the physics models used in the fit, but without the detector effects applied, (i.e., the true physics shapes as opposed to the reconstructed quantities used in the construction of the histograms used in the fit.)

The moments which are compared to the theory are calculated with a lepton energy cut. The theory can properly account for the cut, so extrapolating to the full range of lepton energies is unnecessary and would add additional model dependence. A small extrapolation/interpolation is made because the data used in the fit is selected to have a lab frame energy above a threshold (cut), but the moments are calculated with the cut applied using the lepton energy in the $B$ rest frame. The two frames are only different because of the small momentum of the $B$ meson in the lab frame ($|p_B| \approx 300 MeV$). The moments are calculated for a variety of lepton energy cuts. For each cut, the fit is repeated with the cut applied to the data in the lab frame, and the moment is calculated using the models in the $B$ rest frame. The cut applied in the lab frame is applied in terms of the lepton momentum, but the moment is calculated in terms of the energy; this is a very small interpolation. Similarly, the fit is conducted with the radiative corrections applied, but moments are calculated without the radiative corrections applied, so that they are comparable to the theoretical calculations which do not have the radiative corrections applied (see Section 7.1.1).

The moment of a variable, $M(\vec{x})$, with a cut, $C(\vec{x})$, is

$$\langle M \rangle_C \equiv \frac{\displaystyle\int M(\vec{x}) \frac{d\Gamma}{d\vec{x}} \Theta(C(\vec{x})) \, d\vec{x}}{\displaystyle\int \frac{d\Gamma}{d\vec{x}} \Theta(C(\vec{x})) \, d\vec{x}}, \tag{7.2}$$

where $\Theta$ is the step function ($\Theta(z) = 1$ for $z \geq 0$ and $\Theta(z) = 0$ for $z < 0$). The differential decay rate of each mode is calculated from the Monte Carlo simulations of the physics processes without

the simulation of the detector response or any selections:

$$\frac{d\Gamma_m}{d\vec{x}} = \frac{d\Gamma_m^0}{d\vec{x}} w(\vec{x}) = \sum_{g}^{\#\text{ generated}} w(\vec{x}_g), \tag{7.3}$$

where $d\Gamma_m^0/d\vec{x}$ is the differential decay rate used in the generation of the Monte Carlo sample and $w(\vec{x}_g)$ are the weights that are used to implement the measured form-factors and the variations discussed in Sections 5.2.4 and 6.2.

The calculation of the moments is implemented using two variables which are calculated from the models. The fraction of the decay rate for mode $m$ in the region defined by the the cut is

$$c_m \equiv \frac{\int \frac{d\Gamma_m}{d\vec{x}} \Theta(C(\vec{x})) d\vec{x}}{\int \frac{d\Gamma_m}{d\vec{x}} d\vec{x}} = \frac{\int \frac{d\Gamma_m^0}{d\vec{x}} w(\vec{x}) \Theta(C(\vec{x})) d\vec{x}}{\int \frac{d\Gamma_m^0}{d\vec{x}} w(\vec{x}) d\vec{x}} \tag{7.4}$$

$$= \frac{\sum_{g}^{\#\text{ generated}} w(\vec{x}_g) \Theta(C(\vec{x}_g))}{\sum_{g}^{\#\text{ generated}} w(\vec{x}_g)} \tag{7.5}$$

and the moment of the mode in that region is

$$m_m \equiv \frac{\int M(\vec{x}) \frac{d\Gamma_m}{d\vec{x}} \Theta(C(\vec{x})) d\vec{x}}{\int \frac{d\Gamma_m}{d\vec{x}} \Theta(C(\vec{x})) d\vec{x}} = \frac{\int M(\vec{x}) \frac{d\Gamma_m^0}{d\vec{x}} w(\vec{x}) \Theta(C(\vec{x})) d\vec{x}}{\int \frac{d\Gamma_m^0}{d\vec{x}} w(\vec{x}) \Theta(C(\vec{x})) d\vec{x}} \tag{7.6}$$

$$= \frac{\sum_{g}^{\#\text{ generated}} M(\vec{x}_g) w(\vec{x}_g) \Theta(C(\vec{x}_g))}{\sum_{g}^{\#\text{ generated}} w(\vec{x}_g) \Theta(C(\vec{x}_g))}. \tag{7.7}$$

Using the branching fractions from the fit results, $\mathcal{B}_m$, and $c_m$ and $m_m$ calculated from the models, the moment defined in Equation 7.2 is

$$\langle M \rangle_C = \frac{\sum_m m_m c_m \mathcal{B}_m}{\sum_m c_m \mathcal{B}_m}. \tag{7.8}$$

The quantities $m_m$ and $c_m$ depend only on the model. The measured branching fractions, $\mathcal{B}_m$, depend on the model, the detector simulation, and the data. Since the branching fractions are measured using the inclusive differential decay rate, when combined with the models use in the fit they should give a good description of the true differential decay rate. Mismodeling of a contribution may cause the branching fraction to be mismeasured, but the shape will still be well described. For instance, the main separation of the $D$ and $D^*$ modes is due to the $q^2$ distribution. If the $q^2$ slope of either of them is mismodeled, the relative rates of the modes will be affected, but the model of the $q^2$ distribution and its moments will only be weakly affected.

## 7.1.1 Radiative Corrections

Radiative corrections must be included in the physics models compared with the data, but radiative corrections are not included in the theoretical calculations of the moments. The moments results are therefore calculated with the effect of the radiative corrections removed. This is done using the fact the radiative corrections are not expected to change the branching fractions significantly, so the entire change to the differential decay rate is due to differences of the shape of the physics processes. The branching fractions are therefore extracted from the data using models which include the radiative corrections, as described in Section 5.2.3, and the moments are calculated with $c_m$ and $m_m$ variables extracted from models which do not have the radiative corrections included. The resulting moments are directly comparable with the theoretical calculations, in which radiative corrections have not been included.

## 7.1.2 Correlations and Systematic Uncertainties

### Statistical

The fit gives the fractional contribution of each mode to the data sample. These results are highly correlated. For instance if $B \to Dl\overline{\nu}$ is higher then $B \to D^*l\overline{\nu}$ must be lower because they have very similar $M_X^2$ distributions and the total rate in that range of $M_X^2$ is constrained. These correlations are propagated into the statistical error using the Gaussian approximation:

$$\mathrm{Var}[\langle M \rangle_C] \approx \sum_{m_1, m_2 \in \mathrm{modes}} \frac{d\langle M \rangle_C}{d\mathcal{B}_{m_1}} \frac{d\langle M \rangle_C}{d\mathcal{B}_{m_2}} \mathrm{Cov}[\mathcal{B}_{m_1}, \mathcal{B}_{m_2}], \tag{7.9}$$

where $\mathrm{Var}[\langle M \rangle_C]$ is the variance of the moment $\langle M \rangle_C$ and $\mathrm{Cov}[\mathcal{B}_{m_1}, \mathcal{B}_{m_2}]$ is the covariance matrix of the branching fractions derived from the likelihood maximization. Because all the moments results are derived from the same data set, they are correlated. The statistical covariance of two moments is

$$\mathrm{Cov}[\langle M_1 \rangle_C, \langle M_2 \rangle_C] \approx \sum_{m_1, m_2 \in \mathrm{modes}} \frac{d\langle M_1 \rangle_C}{d\mathcal{B}_{m_1}} \frac{d\langle M_2 \rangle_C}{d\mathcal{B}_{m_2}} \mathrm{Cov}[\mathcal{B}_{m_1}, \mathcal{B}_{m_2}]. \tag{7.10}$$

For two moments with different lepton energy cuts, the moment with the more restrictive cut is calculated from a fit to a subsample of the data used for the other moment. To calculate the resulting statistical correlation of the two moments, a Monte Carlo technique is used. Samples of events, which are the same size as the data sample, are generated using as the probability distribution histograms of the real data binned in the three fit dimensions and the lepton energy. Each sample is then fit multiple times, once for each cut of interest. The moments are then computed in the same way as for the data and the correlations between the results with different lepton energy cuts are then calculated.

**Detector Systematic and Model Dependence Uncertainties**

The detector systematic and model dependence uncertainties are propagated to the results using the same set of variation discussed in Chapter 6. For each variation the fit is repeated and a new set of branching fractions are determined. These branching fractions are then used to calculate a new set of moments. For the detector systematics, only the $\mathcal{B}_m$ values in Equation 7.8 differ between the variations, but for the model dependence the $c_m$ and $m_m$ are also different. The shifts of the moments from their nominal values due to each of the variations are combined into total detector systematic and model dependence uncertainties, by summing in quadrature the largest shift for each category of variation.

Determining the correlations of two moments due to the systematic uncertainties is more complicated. The correlation of two moments, $g$ and $h$, due to any individual variation is 100%. To see this consider the simplified example, in which there is only one systematic, the uncertainty on a parameter $\rho$. Then $g(\rho)$ and $h(\rho)$ would be functions of $\rho$ and the shifts in $g$ and $h$ due to a shift $\Delta\rho$ in $\rho$ would be $\Delta g = \frac{dg}{d\rho}\Delta\rho$ and $\Delta h = \frac{dh}{d\rho}\Delta\rho$. The shifts $\Delta g$ and $\Delta h$ are therefor 100% correlated. The covariance matrix which expresses the 100% correlation is

$$\begin{pmatrix} (\Delta g)^2 & (\Delta g)(\Delta h) \\ (\Delta g)(\Delta h) & (\Delta h)^2 \end{pmatrix},$$ (7.11)

where $\Delta g$ and $\Delta h$ are the shifts of each of the moments due to a particular variation. These covariances are summed over all the systematics as if they were statical uncertainties,

$$C[g,h]_{\text{systematics}} = \sum_{i \in \text{variations}} \begin{pmatrix} (\Delta g)_i^2 & (\Delta g)_i(\Delta h)_i \\ (\Delta g)_i(\Delta h)_i & (\Delta h)_i^2 \end{pmatrix}.$$ (7.12)

Notice that the diagonal elements of the resulting covariance matrix are just the uncertainties summed in quadrature and give the same total uncertainties on the moments as the previously defined method. Because each of the variations will have a different ratio of the shift $\Delta g$ to the shift $\Delta h$, the total covariance matrix is no longer 100% correlated.

For the systematic uncertainties on the individual moments, the largest deviation is used when multiple variations are used to determine the same uncertainty. For example when raising and lowering the tracking efficiency only the larger of the two shifts is used, using both would be double counting. For the correlations this simple prescription breaks down, since the question arises "the largest shift in which moment?" To address this, the ad-hoc choice is made to use the variation which produces the largest deviation of the 1.0 GeV lepton energy cut or the $\langle M_X^2 - \overline{M}_D^2 \rangle$ moment. Different choices of which variation to choose do not significantly affect the result.

## 7.2   Moments Results

Using the technique described above, we tabulate a variety of moments and their correlations. Table 7.1 shows first and second moments of $M_X^2$, $q^2$, and $E_\ell$, and the moment

$$R0 \equiv \frac{\int_{E_\ell = 1.7} \frac{d\Gamma}{dE_\ell} \, dE_\ell}{\int_{E_\ell = 1.5} \frac{d\Gamma}{dE_\ell} \, dE_\ell}. \tag{7.13}$$

The correlations between these moments with a 1.0 GeV lepton energy cut are shown in Table 7.2. The $\langle M_X^2 - \overline{M}_D^2 \rangle$ moment as a function of the lepton energy cut is shown in Table 7.3 and their correlations are shown in Table 7.4. Each of these moments is influenced in different ways by the various uncertainties. The shifts of the $\langle M_X^2 - \overline{M}_D^2 \rangle$ moment with lepton energy cuts of 1.0 GeV and 1.5 GeV for each of the detector systematics and the model dependence variations are shown in Tables 7.5 and 7.6, respectively. The full set of shifts for all the moments are in Appendix A. Because the $\langle M_X^2 - \overline{M}_D^2 \rangle$ moments at different lepton energies are very correlated, it is useful to consider as an alternative representation of the results the difference between the moments at 1.0 and 1.5 GeV,

$$\langle M_X^2 \rangle_{E_\ell \geq 1.0 \, \text{GeV}} - \langle M_X^2 \rangle_{E_\ell \geq 1.5 \, \text{GeV}} = (0.163 \pm 0.014 \pm 0.036 \pm 0.064) \, \text{GeV}^2/c^4,$$

which has a covariance with the $\langle M_X^2 - \overline{M}_D^2 \rangle_{E_\ell > 1.5 \, \text{GeV}}$ moment of $2.242 \times 10^{-3} \, \text{GeV}^4/c^8$ (which corresponds to a correlation coefficient of 0.486).

Table 7.1: Moments Results with $E_\ell > 1.0$ GeV and $E_\ell > 1.5$ GeV Lepton Energy Cuts. The errors on the entries in the table are the statistical, detector systematics, and model dependence, respectively. The definition of R0 with a 1.0 GeV cut is the definition of R0 in Equation 7.13 with the 1.5 GeV cut in the denoimator replaced by a 1.0 GeV cut.

| Moment | $E_\ell > 1.0$ GeV | $E_\ell > 1.5$ GeV |
|---|---|---|
| $\langle M_X^2 - \overline{M}_D^2 \rangle$ (GeV$^2$/$c^4$) | $0.456 \pm 0.014 \pm 0.045 \pm 0.109$ | $0.293 \pm 0.012 \pm 0.033 \pm 0.048$ |
| $\langle (M_X^2 - \langle M_X^2 \rangle)^2 \rangle$ (GeV$^4$/$c^8$) | $1.266 \pm 0.065 \pm 0.222 \pm 0.631$ | $0.629 \pm 0.031 \pm 0.088 \pm 0.113$ |
| $\langle E_\ell \rangle$ (GeV) | $1.551 \pm 0.001 \pm 0.006 \pm 0.008$ | $1.774 \pm 0.000 \pm 0.002 \pm 0.005$ |
| $\langle (E_\ell - \langle E_\ell \rangle)^2 \rangle$ (GeV$^2$) | $0.297 \pm 0.000 \pm 0.000 \pm 0.006$ | $0.176 \pm 0.000 \pm 0.000 \pm 0.006$ |
| R0 | $0.340 \pm 0.001 \pm 0.007 \pm 0.010$ | $0.608 \pm 0.000 \pm 0.004 \pm 0.008$ |
| $\langle q^2 \rangle$ (GeV$^2$/$c^4$) | $4.892 \pm 0.015 \pm 0.094 \pm 0.100$ | $5.287 \pm 0.020 \pm 0.073 \pm 0.095$ |
| $\langle (q^2 - \langle q^2 \rangle)^2 \rangle$ (GeV$^4$/$c^8$) | $2.852 \pm 0.002 \pm 0.003 \pm 0.047$ | $2.879 \pm 0.006 \pm 0.007 \pm 0.049$ |

### 7.2.1   Cross-Checks

The lepton energy moments, $\langle E_\ell \rangle$ and $R0$, for lepton energies above 1.5 GeV have been previously measured with smaller systematic errors in an analysis which only reconstructs a lepton [15]. These

Table 7.2: Statistical and Systematic Correlations of the Moments Results with a $E_\ell > 1.0$ GeV Lepton Energy Cut.

Starting with the top/left the rows/columns correspond to $\langle M_X^2 - \overline{M}_D^2 \rangle$, $\langle (M_X^2 - \langle M_X^2 \rangle)^2 \rangle$, $\langle E_\ell \rangle$, $\langle (E_\ell - \langle E_\ell \rangle)^2 \rangle$, R0, $\langle q^2 \rangle$, and $\langle (q^2 - \langle q^2 \rangle)^2 \rangle$. The units are not the same for all the entries; they are the product of the units for the row and the column given in Table 7.1.

Covariance Matrix:

$$
\begin{pmatrix}
1.42 \times 10^{-2} & 7.79 \times 10^{-2} & -5.10 \times 10^{-4} & 3.16 \times 10^{-4} & -4.20 \times 10^{-4} & -1.03 \times 10^{-2} & -2.13 \times 10^{-4} \\
 & 4.53 \times 10^{-1} & -3.26 \times 10^{-3} & 1.26 \times 10^{-3} & -2.96 \times 10^{-3} & -5.02 \times 10^{-2} & 1.60 \times 10^{-3} \\
 & & 7.06 \times 10^{-5} & 1.57 \times 10^{-5} & 9.03 \times 10^{-5} & 1.65 \times 10^{-5} & -2.06 \times 10^{-4} \\
 & & & 3.43 \times 10^{-5} & 3.11 \times 10^{-5} & -4.98 \times 10^{-4} & -1.22 \times 10^{-4} \\
 & & & & 1.21 \times 10^{-4} & -2.10 \times 10^{-4} & -3.14 \times 10^{-4} \\
 & & & & & 1.19 \times 10^{-2} & 1.56 \times 10^{-3} \\
 & & & & & & 2.22 \times 10^{-3}
\end{pmatrix}
$$

Correlation Matrix:

$$
\begin{pmatrix}
1.000 & 0.971 & -0.510 & 0.453 & -0.319 & -0.794 & -0.038 \\
 & 1.000 & -0.577 & 0.320 & -0.398 & -0.683 & 0.050 \\
 & & 1.000 & 0.319 & 0.974 & 0.018 & -0.522 \\
 & & & 1.000 & 0.481 & -0.779 & -0.444 \\
 & & & & 1.000 & -0.175 & -0.605 \\
 & & & & & 1.000 & 0.303 \\
 & & & & & & 1.000
\end{pmatrix}
$$

Table 7.3: $\langle M_X^2 - \overline{M}_D^2 \rangle$ versus the Lepton Energy Cut.

The errors on the entries in the table are the statistical, detector systematics, and model dependence, respectively.

| Cut (GeV) | $\langle M_X^2 - \overline{M}_D^2 \rangle$ (GeV$^2/c^4$) |
|---|---|
| $E_\ell > 1.0$ | $0.456 \pm 0.014 \pm 0.045 \pm 0.109$ |
| $E_\ell > 1.1$ | $0.422 \pm 0.014 \pm 0.031 \pm 0.084$ |
| $E_\ell > 1.2$ | $0.393 \pm 0.013 \pm 0.027 \pm 0.069$ |
| $E_\ell > 1.3$ | $0.364 \pm 0.013 \pm 0.030 \pm 0.054$ |
| $E_\ell > 1.4$ | $0.332 \pm 0.012 \pm 0.027 \pm 0.055$ |
| $E_\ell > 1.5$ | $0.293 \pm 0.012 \pm 0.033 \pm 0.048$ |

Table 7.4: Correlations of the $\langle M_X^2 - \overline{M}_D^2 \rangle$ Moments with Different Lepton Energy Cuts. The rows/columns correspond to progressively more restrictive lepton energy cut starting with 1.0 GeV and ending with 1.5 GeV.

Covariance Matrix

$$
\begin{pmatrix}
1.421 & 1.096 & 0.913 & 0.724 & 0.735 & 0.618 \\
 & 0.864 & 0.719 & 0.577 & 0.574 & 0.486 \\
 & & 0.612 & 0.495 & 0.499 & 0.422 \\
 & & & 0.413 & 0.418 & 0.352 \\
 & & & & 0.447 & 0.373 \\
 & & & & & 0.324
\end{pmatrix} \times 10^{-2} \ \mathrm{GeV}^4/c^8
$$

Correlation Matrix

$$
\begin{pmatrix}
1.000 & 0.989 & 0.979 & 0.944 & 0.922 & 0.910 \\
 & 1.000 & 0.990 & 0.965 & 0.924 & 0.918 \\
 & & 1.000 & 0.985 & 0.955 & 0.949 \\
 & & & 1.000 & 0.973 & 0.963 \\
 & & & & 1.000 & 0.981 \\
 & & & & & 1.000
\end{pmatrix}
$$

Table 7.5: Detector Systematics for $\langle M_X^2 - \overline{M}_D^2 \rangle$.
Nominal value and shifts of the $\langle M_X^2 - \overline{M}_D^2 \rangle$ moments due to the individual detector systematic uncertainties. All entries are in units of $\mathrm{GeV}^2/c^4$

| Variation | $E_\ell > 1.0$ GeV | $E_\ell > 1.5$ GeV |
|---|---|---|
| Nominal | 0.456 | 0.293 |
| Lepton Fake Rate | 0.014 / -0.006 | 0.017 / -0.030 |
| DELCO $b \to c \to \ell$ shape | -0.008 | 0.001 |
| DELCO $+1\sigma$ $b \to c \to \ell$ shape | -0.011 | 0.001 |
| DELCO $-1\sigma$ $b \to c \to \ell$ shape | -0.001 | 0.003 |
| Continuum Norm $\pm 10\%$ | 0.001 / 0.001 | 0.005 / -0.005 |
| $\mu$ Fakes, $E_\ell < 1.5$ GeV, $\pm 10\%$ | -0.011 / 0.011 | 0.000 / 0.000 |
| $\mu$ Fakes, $E_\ell > 1.5$ GeV, $\pm 10\%$ | -0.000 / 0.000 | 0.004 / -0.004 |
| $\mathcal{B}(b \to c \to \ell) \pm 10\%$ | 0.004 / -0.003 | 0.005 / -0.005 |
| $\mathcal{B}(b \to \mathrm{baryons}) \pm 20\%$ | 0.002 / -0.002 | 0.003 / -0.003 |
| # $K_L^0$ | 0.003 / -0.003 | 0.004 / -0.004 |
| Track Efficiency | 0.002 / -0.001 | -0.003 / 0.003 |
| # Fake Tracks | -0.009 / 0.008 | -0.003 / 0.003 |
| Shower Efficiency | 0.001 / 0.000 | 0.001 / -0.001 |
| # Fake Showers | -0.030 / 0.031 | -0.013 / 0.014 |
| Force Trk Multiplicity | 0.002 | 0.001 |
| Force Shwr Multiplicity | 0.011 | 0.009 |
| Final State Radiation | 0.021 | 0.011 |
| Lepton Efficiency | 0.006 | 0.004 |

Table 7.6: Model Dependence for $\langle M_X^2 - \overline{M}_D^2 \rangle$.

Nominal value and shifts of the $\langle M_X^2 - \overline{M}_D^2 \rangle$ moments due to the individual model variations. All entries are in units of GeV$^2/c^4$

| Variation | $E_\ell > 1.0$ GeV | $E_\ell > 1.5$ GeV |
|---|---|---|
| Nominal | 0.456 | 0.293 |
| $B \to Dl\overline{\nu}$ $\rho$ param | -0.006 / 0.005 | -0.006 / 0.005 |
| $B \to Dl\overline{\nu}$ $c_D$ param | -0.001 / 0.001 | -0.001 / 0.000 |
| $B \to D^*l\overline{\nu}$ $\rho$ param | -0.020 / 0.025 | -0.020 / 0.021 |
| $B \to D^*l\overline{\nu}$ $c_{A_1}$ param | 0.003 / -0.002 | 0.004 / -0.004 |
| $B \to D^*l\overline{\nu}$ R1/R2 1st eig-vec | 0.019 / -0.018 | 0.009 / -0.009 |
| $B \to D^*l\overline{\nu}$ R1/R2 2nd eig-vec | -0.000 / 0.000 | -0.002 / 0.002 |
| $B \to D^{**}l\overline{\nu}$ HQET model | 0.008 | 0.005 |
| $B \to D^{**}l\overline{\nu}$ $w$ slope | 0.014 / -0.014 | 0.008 / -0.008 |
| $B \to X_cl\overline{\nu}$ NonRes $w$ slope | 0.008 / -0.005 | 0.005 / -0.003 |
| $B \to X_cl\overline{\nu}$ Nonresonant Mass Dependence | | |
| Polynomial | 0.029 / -0.017 | 0.009 / -0.004 |
| Gauss $\mu = 2.0$ $\sigma^2 = .25$ | -0.078 | -0.016 |
| Gauss $\mu = 2.0$ $\sigma^2 = .75$ | -0.023 | -0.009 |
| Gauss $\mu = 2.0$ $\sigma^2 = 1.25$ | 0.003 | -0.002 |
| Gauss $\mu = 2.75$ $\sigma^2 = .25$ | -0.007 | 0.002 |
| Gauss $\mu = 2.75$ $\sigma^2 = .75$ | 0.024 | 0.006 |
| Gauss $\mu = 2.75$ $\sigma^2 = 1.25$ | 0.045 | 0.014 |
| Gauss $\mu = 3.5$ $\sigma^2 = .25$ | 0.059 | 0.019 |
| Gauss $\mu = 3.5$ $\sigma^2 = .75$ | 0.091 | 0.033 |
| Gauss $\mu = 3.5$ $\sigma^2 = 1.25$ | 0.102 | 0.040 |
| $B \to X_ul\overline{\nu}$ Model Dependence | | |
| $B \to X_s\gamma$ based | 0.001 | 0.002 |
| InclGen | 0.001 | 0.001 |
| All Non-Res | 0.003 | 0.003 |
| All Non-Res High Mass | -0.012 | -0.005 |
| All Non-Res Low Mass | 0.009 | 0.008 |
| Hard $q^2$ | -0.005 | -0.005 |
| Soft $q^2$ | 0.005 | 0.004 |
| High $\cos\theta_{W\ell}$ | -0.001 | -0.000 |
| Low $\cos\theta_{W\ell}$ | 0.001 | 0.000 |

are therefore good tests that the systematics assessments are adequate. In Table 7.7, the previously measured lepton energy moments are shown to agree within the systematic uncertainties with the results of this analysis. Also in the table are the previous hadronic mass moments results measured at CLEO [16], which are partially correlated with the results of this analysis, because they are based on a subset of the data used in the analysis and both use the same detector and physics simulations. The difference between the two results is largely due to different handling of the radiative corrections. The previous result did not simulate the detector response to the photon. Also the previous results did not use the $q^2$ and $\cos\theta_{W\ell}$ information, which is very useful for differentiating the $B \to Dl\overline{\nu}$ and $B \to D^*l\overline{\nu}$ modes.

Table 7.7: Comparison of the Moments Results with Previous Measurements.

R0

| | |
|---|---|
| this analysis | $0.608 \pm 0.001 \pm 0.004 \pm 0.009$ |
| Lepton Only [15] | $0.6187 \ \pm 0.0014 \ \pm 0.0016$ |

$\langle E_\ell \rangle$

| | |
|---|---|
| this analysis | $1.774 \pm 0.001 \pm 0.002 \pm 0.005$ GeV |
| Lepton Only [15] | $1.7810 \ \pm 0.0007 \ \pm 0.0009$ GeV |

$\langle M_X^2 - \overline{M}_D^2 \rangle$ with Lepton Energy $\geq 1.5$ GeV cut

| | |
|---|---|
| this analysis | $0.293 \pm 0.012 \pm 0.033 \pm 0.048$ GeV$^2/c^4$ |
| Previous CLEO [16] | $0.251 \pm 0.066$ GeV$^2/c^4$ |

$\langle (M_X^2 - \langle M_X^2 \rangle)^2 \rangle$ with Lepton Energy $\geq 1.5$GeV cut

| | |
|---|---|
| this analysis | $0.629 \pm 0.031 \pm 0.088 \pm 0.113$ GeV$^4/c^8$ |
| Previous CLEO [16] | $0.576 \pm 0.170$ GeV$^4/c^8$ |

To test for potential mismodeling, the moments of the reconstructed $M_X^2$ in data are compared to the simulation with the branching fractions set to the fit results (see Figure 7.1). This is a test of the quality of the fit, which is specific to how well the quantity being measured (i.e., $\langle M_X^2 - \overline{M}_D^2 \rangle$) is modeled by the data. The comparison is made in bins of lepton energy and shows good agreement across the range. The data has slightly higher means, which could be due to the shape of the signal or background composition, but are consistent within the assigned systematic errors.

In Table 7.8 the results using only the electron sample are compared to the results using only the muon sample. Like the analogous comparison for the branching fraction results (see Section 6.3.1), this is sensitive to mismodeling of the lepton efficiencies and fake rates, the radiative corrections, and the form factors. Again, the agreement for all of the results is good within the statistical errors. The comparison for the $\langle M_X^2 - \overline{M}_D^2 \rangle$ moment is particularly sensitive to the modeling of the radiative corrections, because the electrons have much larger corrections than the muons. As indicated in Table 7.5, the uncertainty due to the radiation corrections is assessed to be $\pm 0.021$GeV$^2/c^4$.
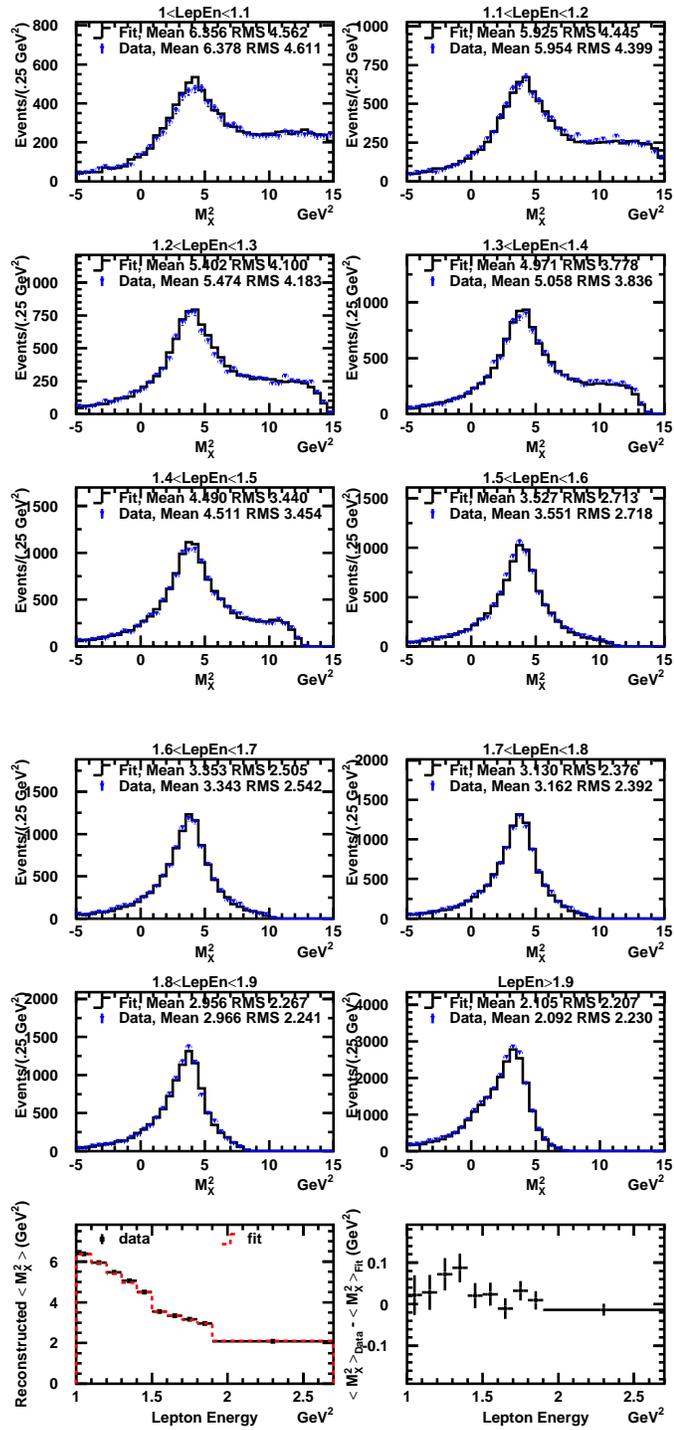
Figure 7.1: Comparision of the Reconstructed $M_X^2$ between Data and the Fit Results. Distribution of the reconstructed $M_X^2$ for data and the fit results in bins of lepton energy (first 10 plots). Summary of the moments of the reconstructed $M_X^2$ as a function of lepton energy (bottom row, first column) and the difference between these moments for data and fit (bottom row, second column).

Table 7.8: Comparison of Moments Results for Electrons and Muons with $E_\ell > 1.0$ GeV. The errors on the entries in the table are the statistical, detector systematics, and model dependence, respectively.

| Moment | Electron | Muon |
|---|---|---|
| $\langle M_X^2 - \overline{M}_D^2 \rangle$ (GeV$^2$/$c^4$) | $0.467 \pm 0.018 \pm 0.051 \pm 0.134$ | $0.434 \pm 0.023 \pm 0.047 \pm 0.100$ |
| $\langle (M_X^2 - \langle M_X^2 \rangle)^2 \rangle$ (GeV$^4$/$c^8$) | $1.215 \pm 0.078 \pm 0.206 \pm 0.802$ | $1.247 \pm 0.111 \pm 0.257 \pm 0.612$ |
| $\langle E_\ell \rangle$ (GeV) | $1.551 \pm 0.001 \pm 0.006 \pm 0.010$ | $1.552 \pm 0.001 \pm 0.005 \pm 0.007$ |
| $\langle (E_\ell - \langle E_\ell \rangle)^2 \rangle$ (GeV$^2$) | $0.296 \pm 0.000 \pm 0.000 \pm 0.003$ | $0.297 \pm 0.000 \pm 0.000 \pm 0.006$ |
| R0 | $0.339 \pm 0.001 \pm 0.007 \pm 0.013$ | $0.341 \pm 0.001 \pm 0.007 \pm 0.010$ |
| $\langle q^2 \rangle$ (GeV$^2$/$c^4$) | $4.880 \pm 0.020 \pm 0.099 \pm 0.094$ | $4.914 \pm 0.022 \pm 0.093 \pm 0.097$ |
| $\langle (q^2 - \langle q^2 \rangle)^2 \rangle$ (GeV$^4$/$c^8$) | $2.852 \pm 0.003 \pm 0.003 \pm 0.047$ | $2.853 \pm 0.004 \pm 0.003 \pm 0.047$ |

Understanding the final state radiation is particularly important in the measurement of the $\langle M_X^2 - \overline{M}_D^2 \rangle$ moment. In Table 7.9, the effect of the various methods of application of the radiative corrections are compared. "MyFSR" is the method of applying radiative corrections after the detector response discussed in Section 5.2.3. "MyFSR Improved" is the same as "MyFSR", but a correction is made to approximate the detector response to radiated photons. A random number is used to reject photons based on the photon efficiency extracted from the simulation (shown in Figure 5.4). If a photon is rejected, the energy of the photon is added to the reconstructed neutrino. The table shows that "MyFSR Improved" is closer to the PHOTOS based model than "MyFSR", which indicates that the discrepancy between the two methods is most likely a result of the different modeling of the detector response to the photon. The "MyFSR" method most closely resembles the method used in the previous CLEO analysis, and the result is also closer. The PHOTOS method is a better simulation of the detector response and the corresponding result, which is the central value of this analysis, is more reliable.

Table 7.9: Comparison of the Effect Different Final State Radiation Treatments on $\langle M_X^2 - \overline{M}_D^2 \rangle$. All entries are in units of GeV$^2$/$c^4$.

| FSR Method | $E_\ell > 1.0 GeV$ | $E_\ell > 1.5 GeV$ |
|---|---|---|
| PHOTOS | 0.456 | 0.293 |
| No FSR | 0.499 | 0.295 |
| "MyFSR" | 0.395 | 0.268 |
| "MyFSR Improved" | 0.438 | 0.303 |

The effect of including in the fit the $q^2$ and $\cos\theta_{W\ell}$ dimensions in addition to the $M_X^2$ dimension can be explored by removing these dimensions from the fit. Table 7.10 shows the variation of the $\langle M_X^2 - \overline{M}_D^2 \rangle$ result with a 1.0 GeV lepton energy cut for different sets of variables used in the fit. Adding the $q^2$ variable considerably reduces the detector systematics and model dependence uncertainties. This is because the backgrounds are concentrated at $q^2 \leq 2.0$ GeV, so adding the $q^2$

information allows the fit to better separate the signal from the background.

Table 7.10: $\langle M_X^2 - \overline{M}_D^2 \rangle_{E_\ell > 1.0}$ Changing Fit Dimensions Comparison.

| Dimensions in the Fit | $\langle M_X^2 - \overline{M}_D^2 \rangle$ |
|---|---|
| $M_X^2$ only | $0.436 \pm 0.031 \pm 0.063 \pm 0.177$ |
| $M_X^2$ & $q^2$ | $0.444 \pm 0.015 \pm 0.051 \pm 0.094$ |
| $M_X^2$ & $\cos\theta_{W\ell}$ | $0.461 \pm 0.020 \pm 0.080 \pm 0.170$ |
| $M_X^2$, $\cos\theta_{W\ell}$, & $q^2$ (full) | $0.456 \pm 0.014 \pm 0.045 \pm 0.109$ |

## 7.3   HQET and the Interpretation of the Moments

The measured moments can be used to determine the HQET parameters $\overline{\Lambda}$ and $\lambda_1$, which in turn can be used to refine the calculation of $|V_{cb}|$. It interesting to use the moments to over constrain the parameters and thereby test the HQET predictions. Each moment can be expressed as an expansion in powers of $\Lambda_{QCD}/m_b$. The terms at each order in the expansion are related to the HQET parameters of interest and have different effects on each of the various moments. The total decay rate for the process $B \to X_c l \overline{\nu}$, is just the zeroth moment and can be expanded in the same way. There is only one parameter at $\mathcal{O}(\Lambda_{QCD}/m_b)$, $\overline{\Lambda}$, which connects the b quark mass to the $B$ meson mass. At $\mathcal{O}((\Lambda_{QCD}/m_b)^2)$, the expansion has three terms two of which are related to the HQET parameters $\lambda_1$ and $\lambda_2$ and the last of which is related to $\overline{\Lambda}^2$. At $\mathcal{O}((\Lambda_{QCD}/m_b)^3)$, there are nine terms, six of which are related to the parameters $\rho_1$, $\rho_2$, $\tau_1$, $\tau_2$, $\tau_3$, and $\tau_4$. The remaining three are related to products of the lower order parameters: $\lambda_1 \overline{\Lambda}$, $\lambda_2 \overline{\Lambda}$, and $\overline{\Lambda}^3$.

The parameter $\lambda_2$ can be calculated using the measured $D^{(*)}$ and $B^{(*)}$ meson masses:

$$\lambda_2(m_b) = \frac{m_B^2 \, \Delta m_B - m_D^2 \, \Delta m_D}{2(m_B - \kappa(m_c)m_D)} = (0.1255 \pm 0.001), \tag{7.14}$$

where $\Delta m_D$ and $\Delta m_B$ are the $D^* - D$ and $B^* - B$ mass splittings, respectively, and $\kappa(m_c)$ is $(\alpha_s(m_c)/\alpha_s(m_b))^3/\beta_0 \approx 1.2$ [18]. The error shown reflects only the experimental uncertainty (dominated by the $B^* - B$ masssplitting). Theoretical errors are probably 10%, which still smaller than the other uncertainties in the extraction of $\overline{\Lambda}$ and $\lambda_1$. This differs from the mass formula presented in Section 2.3, because the calculation has been carried to $\mathcal{O}((\Lambda_{QCD}/m_b)^3)$. The mass splittings also constrain a combination of the third order parameters [18]:

$$\rho_2 - \tau_2 - \tau_4 = \frac{\kappa(m_c)m_B^2 \, \Delta m_B(m_D + \overline{\Lambda}) - m_D^2 \, \Delta m_D(m_B + \overline{\Lambda})}{m_B + \overline{\Lambda} - \kappa(m_c)(m_D + \overline{\Lambda})}. \tag{7.15}$$

Bauer et al. have calculated the terms of the expansion for moments $\langle M_X^2 - \overline{M}_D^2 \rangle$, $\langle (M_X^2 -$

$\langle M_X^2\rangle)^2\rangle$, $\langle E_\ell\rangle$ and $\langle (E_\ell - \langle E_\ell\rangle)^2\rangle$ as a function of the lepton energy cut [22]. They have also calculated the fraction of the differential decay rate above a lepton energy cut [21] and moments of the photon energy as a function of the minimum photon energy cut for the process $B \to X_s\gamma$. Their results are summarized as tables of coefficients $M_i$, which are used to calculate a moment, $\langle M\rangle$, with a cut using the expression:

$$
\begin{aligned}
\langle M\rangle = {} & M_1 + M_2\,\overline{\Lambda} + M_3\,\overline{\Lambda}^2 + M_4\,\overline{\Lambda}^3 \\
& + M_5\,\lambda_1 + M_6\,\lambda_2 + M_7\,\lambda_1\overline{\Lambda} + M_8\,\lambda_2\overline{\Lambda} \\
& + M_9\,\rho_1 + M_{10}\,\rho_2 + M_{11}\,\tau_1 + M_{12}\,\tau_2 \\
& + M_{13}\,\tau_3 + M_{14}\,\tau_4 + M_{15}\,\epsilon + M_{16}\,\epsilon_{BLM}^2 + M_{17}\,\epsilon\overline{\Lambda},
\end{aligned}
\tag{7.16}
$$

where $\epsilon$ and $\epsilon_{BLM}^2$ express the size of the $\alpha_s$ and $\alpha_s^2\beta_0$ terms of the expansion. The default values of $\epsilon$ and $\epsilon_{BLM}^2$ are one. The theory behind these calculations is outlined in Sections 2.3 and 2.4.

To explore the information provided by each moment, bands in the $\overline{\Lambda}$-$\lambda_1$ plane are constructed to show the constraints of the the measurements and the effects of the theoretical uncertainties. The theoretical uncertainties are due to the higher order terms in the $\Lambda_{QCD}/m_b$ expansion, the knowledge of at which scale to evaluate $\alpha_s$, and the higher order terms in the $\alpha_s$ expansion. The scale of the variation of the input moment is set by the combined statistical and systematics errors of the measurement. The $\tau_i$ third order HQET parameters are varied by $(0.5 \text{ GeV})^3$, while $\rho_2$ is constrained by Equation 7.15. The $\rho_1$ parameter is set to $(0.5 \pm 0.5)(0.5 \text{ GeV})^3$ to satisfy the bound that $\rho_1$ be positive, as indicated by the vacuum saturation approximation [18]. The $\epsilon$ denotes the terms proportional to $\alpha_s$ and is varied to reflect a change from the nominal $\alpha_s(m_b)$ by $\pm 25\%$ ($\epsilon = 1.0 \pm 0.25$), which corresponds to the range $\alpha_s(m_b/2)$ to $\alpha_s(2m_b)$. Similarly $\epsilon_{BLM}^2$, which corresponds to the corrections of order $\alpha_s^2\beta_0$, is varied by $\pm 100\%$ ($\epsilon_{BLM}^2 = 1.0 \pm 1.0$). The bands are calculated by scanning over values of $\overline{\Lambda}$. For each $\overline{\Lambda}$, many sets of Gaussian distributed random numbers are generated with variances appropriate for the parameters being varied. A value for $\lambda_1$ is then calculated using the each set of parameters, the $\overline{\Lambda}$, and Equation 7.16. The mean, $\mu$, and variance, $\sigma^2$ of the $\lambda_1$ values for each $\overline{\Lambda}$ are then calculated. The upper and lower bounds of the bands are then set to be $\mu \pm 1\sigma$ and should therefore be interpreted as one sigma errors.

The first plot in Figure 7.2 shows the resulting bands for the $\langle M_X^2 - \overline{M}_D^2\rangle_{E_\ell \geq 1.0 \text{ GeV}}$ and $\langle M_X^2 - \overline{M}_D^2\rangle_{E_\ell \geq 1.5 \text{ GeV}}$ moments and for the difference between the two. Also shown on in the first plot, is the bands from the first moment of the $B \to X_s\gamma$ photon energy spectrum [13] and the error ellipse from the previous lepton moments measurements [15]. All of these measurements are consistent within the uncertainties, which suggests that HQET is accurately predicting the relationships between them. As indicated in Table 7.4, the two $\langle M_X^2 - \overline{M}_D^2\rangle$ moments with different lepton energy cuts are highly correlated (correlation coefficient = .91), but as previously discussed
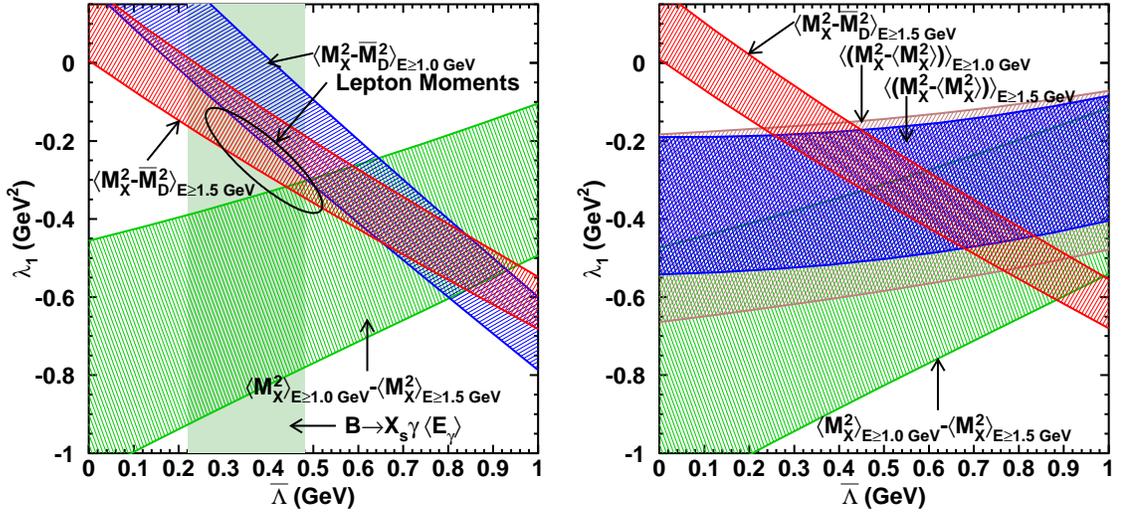
Figure 7.2: Constraints of the Moments in the $\overline{\Lambda}$-$\lambda_1$ Plane.

the difference is significantly less correlated. Also almost identical combinations of $\overline{\Lambda}$ and $\lambda_1$ appear in the theoretical calculations of the two $\langle M_X^2 - \overline{M}_D^2 \rangle$ moments causing the bands to be nearly parallel. The difference moment, $\langle M_X^2 \rangle_{E_\ell \geq 1.0\,\mathrm{GeV}} - \langle M_X^2 \rangle_{E_\ell \geq 1.5\,\mathrm{GeV}}$, is less correlated with the $\langle M_X^2 - \overline{M}_D^2 \rangle_{E_\ell \geq 1.5\,\mathrm{GeV}}$ moment and has a band nearly perpendicular to the $\langle M_X^2 - \overline{M}_D^2 \rangle_{E_\ell \geq 1.5\,\mathrm{GeV}}$ band, although much broader. In the second plot the $\langle (M_X^2 - \langle M_X^2 \rangle)^2 \rangle_{E_\ell \geq 1.0\,\mathrm{GeV}}$ and $\langle (M_X^2 - \langle M_X^2 \rangle)^2 \rangle_{E_\ell \geq 1.5\,\mathrm{GeV}}$ bands are shown. These are also nearly parallel, but they make a significant angle with the $\langle M_X^2 - \overline{M}_D^2 \rangle_{E_\ell \geq 1.5\,\mathrm{GeV}}$ band and therefore provide a useful constraint.

The variation of $\langle M_X^2 - \overline{M}_D^2 \rangle$ as function of lepton energy is shown in Figure 7.3. The HQET theory in the figure is constrained by the $\langle M_X^2 - \overline{M}_D^2 \rangle$ moment for a 1.5 GeV lepton energy cut and the the $B \to X_s \gamma$ first photon energy moment [13]. The photon energy moment constrains $\overline{\Lambda}$ to be $0.35 \pm 0.13$ GeV. The theory bands shown in the figure reflects the experimental errors on the two constraints in addition to the theoretical uncertainties which are assessed in a similar manner to the bands described above. Note that neither the band nor the data points are uncorrelated. The theory does however reasonably predict the variation of the moments as a function of the lepton energy.

Using the $\langle M_X^2 - \overline{M}_D^2 \rangle$ moment with a 1.5 GeV lepton energy cut and the $B \to X_s \gamma$ first photon energy moment, the calculated value of $\lambda_1$ is

$$\lambda_1 = (-0.21 \pm 0.10 \pm 0.06)\ \mathrm{GeV}^2, \tag{7.17}$$

where the errors are the combined experimental error from both measurements (statistical and systematic combined in quadrature) and the theoretical uncertainty calculated in the same way as
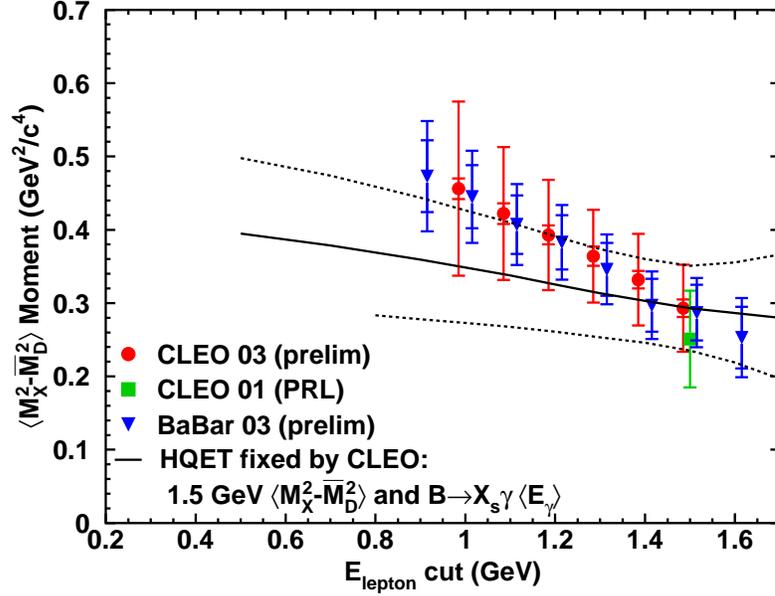
Figure 7.3: $\langle M_X^2 - \overline{M}_D^2 \rangle$ versus Lepton Energy Cut.
The CLEO03 data points are from the work presented here, the BaBar03 data points are from reference [17], and the CLEO01 data point is from [16].

the bands. Using the $\langle M_X^2 - \overline{M}_D^2 \rangle$ moment with a 1.0 GeV lepton energy cut in place of the 1.0 GeV moment, the value of $\lambda_1$ is $-0.13 \pm .145 \pm 0.13$ GeV$^2$.

Finally the value of $|V_{cb}|$ can be calculated using the equation

$$
\begin{aligned}
\Gamma(B \to X_c l\overline{\nu}) = \frac{G_F^2 |V_{cb}|^2}{192\pi^3} \overline{m}_B^5 \Big[ & 0.370 - 0.115\,\overline{\Lambda} - 0.012\,\overline{\Lambda}^2 + 0.\,\overline{\Lambda}^3 \\
& - 0.04\,\lambda_1 - 0.10\,\lambda_2 - 0.01\,\lambda_1\overline{\Lambda} + 0.02\,\lambda_2\overline{\Lambda} \\
& - 0.02\,\rho_1 + 0.02\,\rho_2 - 0.02\,\tau_1 + 0.\,\tau_2 \\
& - 0.03\,\tau_3 - 0.02\,\tau_4 - 0.040\,\epsilon - 0.022\,\epsilon_{BLM}^2 + 0.007\,\epsilon\overline{\Lambda} \Big],
\end{aligned}
\tag{7.18}
$$

which was introduced in Section 2.4.2. In order to apply this equation, the value of $\Gamma(B \to X_c l\overline{\nu})$ must be calculated from the measured branching fraction, $\mathcal{B}(B \to X_c l\overline{\nu})$. The decay rate $\Gamma(B \to X_c l\overline{\nu})$ is assumed to be the same for $B^+$ and $B^0$, but $\mathcal{B}(B \to X_c l\overline{\nu})$ will not be the same because their total decay rates are not equal, $\tau^+ = 1.542 \pm 0.016$ps $\neq \tau^0 = 1.674 \pm 0.018$ [1]. The branching fraction $\mathcal{B}(B \to X_c l\overline{\nu})$ has been measured to be $(10.7 \pm .4)\%$[1] at the $\Upsilon(4S)$ where the ratio of $B^+$ to $B^0$ produced is $f_{+-}/f_{00} = 1.04 \pm 0.08$ [1]. It should be noted that $\mathcal{B}(b \to X_c \ell\nu)$ where $b$ is a combination of $B^+$, $B^0$, $B_s$, and $b$ baryons has been measured at LEP, but these measurements are

---

[1] Measurements of the inclusive semileptonic branching fraction are $\mathcal{B}(B \to X l\overline{\nu})$ summarized in Table 6.5. The $B \to X_u l\overline{\nu}$ contribution $\approx 2 \times 10^{-3}$ must be subtracted from these to get the branching fraction for $B \to X_c l\overline{\nu}$.

not directly applicable because $\overline{\Lambda}$ and $\lambda_1$ are specific to the $B^+$ and $B^0$ (which should have very similar semileptonic decay properties because of the isospin symmetry). The measured branching fraction is

$$\mathcal{B}(B \to X_c l\overline{\nu}) = f_{+-}\frac{\Gamma(B \to X_c l\overline{\nu})}{\Gamma(B^+)} + f_{00}\frac{\Gamma(B \to X_c l\overline{\nu})}{\Gamma(B^0)}, \tag{7.19}$$

which implies

$$\Gamma(B \to X_c l\overline{\nu}) = \frac{\hbar}{c}\frac{\mathcal{B}(B \to X_c l\overline{\nu})}{(f_{+-}\tau^+ + f_{00}\tau^0)} = (4.35 \pm .21) \times 10^{-14}\text{GeV}. \tag{7.20}$$

Using the $\langle M_X^2 - \overline{M}_D^2 \rangle$ moment with a 1.5 GeV lepton energy cut, the $B \to X_s\gamma$ first photon energy moment, and the above value of $\Gamma(B \to X_c l\overline{\nu})$, $|V_{cb}|$ is

$$|V_{cb}| = (4.12 \pm .10 \pm 0.09 \pm 0.16) \times 10^{-2}, \tag{7.21}$$

where the errors are due to the measurement of $\Gamma(B \to X_c l\overline{\nu})$, the measurements of the two moments, and the theoretical uncertainties due to the third order terms and the radiative corrections.

# Chapter 8    $|V_{ub}|$ Extraction

In Section 6.2.5, we found that the branching fraction for $B \to X_u l\overline{\nu}$ determined from the fit results varies by a factor of three as the model of the $B \to X_u l\overline{\nu}$ differential decay rate is changed. Furthermore, the lepton energy information is included in the three kinematic variables that are used in the fit. The goal of using the neutrino reconstruction analysis is to not use the lepton energy to distinguish $B \to X_u l\overline{\nu}$ from $B \to X_c l\overline{\nu}$, because the theoretical calculations are less reliable (see Section 2.4.4). In this chapter, both of these problems are addressed and $|V_{ub}|$ is extracted from the fit results.

## 8.1    The Branching Fraction in the $B \to X_u l\overline{\nu}$ Sensitive Region

The fit is only sensitive to the $B \to X_u l\overline{\nu}$ rate in a small region of phase space where it is not completely overwhelmed by the $B \to X_c l\overline{\nu}$ contribution. The fit method extrapolates from this region to the region of phase space used in the fit. The efficiency correction used to calculate the branching fraction further extrapolates to the full kinematically allowed phase space. Each of these extrapolations is model dependent. The large model dependence of the measured branching fraction indicates that these models are not reliable enough to make a high precision extraction of $|V_{ub}|$. Instead, calculations based on heavy quark effect theory (HQET) and the operator product expansion (OPE) are used to extract $|V_{ub}|$ from the $B \to X_u l\overline{\nu}$ decay rate in a restricted region of phase space. This is a well controlled expansion and the theoretical uncertainties have been assessed by several authors [25, 26]. However, the fit's region of sensitivity does not coincide with these regions, and it is not possible to make cuts to isolate these regions, because of the poor resolution on the neutrino four-vector and hence $q^2$ and $M_X^2$. In order to calculate $|V_{ub}|$, a model is used to infer the partial branching fraction in a region of phase space where the HQET and OPE calculations are available. This prescription is designed to minimize the reliance on models and instead rely on the controlled expansion used in the HQET and OPE calculations. Figure 8.1 shows projections of kinematic variables with cuts selecting the region of $B \to X_u l\overline{\nu}$ sensitivity, $q^2 > 11$ GeV$^2$/$c^4$ and $M_X$ $<1.5$ GeV/$c^2$. Figure 8.2 shows the region of high $S^2/(S+B)$ ($B \to X_u l\overline{\nu}$ signal squared over signal plus background), relative to the region used for the nominal $|V_{ub}|$ result. The quantity $S^2/(S+B)$ is directly related to the information available in the fit.

Models of the $B \to X_u l\overline{\nu}$ differential decay rate and final state hadronic structure are used for two purposes in this analysis. The first is to simulate the shape of the distribution of the measured variables including the efficiency and resolution of the detector. This cannot be done with the HQET
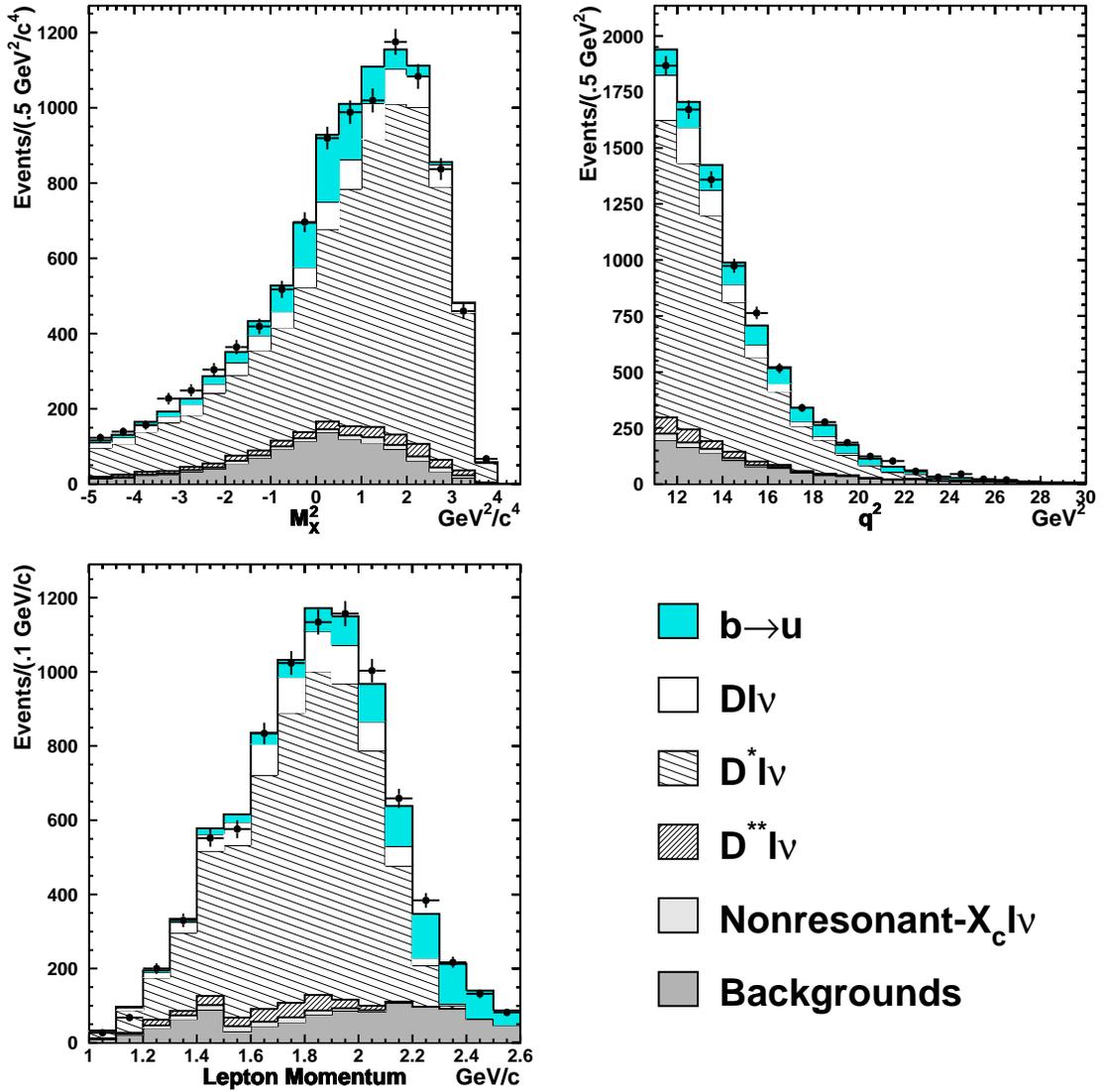
Figure 8.1: Projections Restricted to the Region of $B \rightarrow X_u l \overline{\nu}$ Sensitivity.
Projections of $M_X^2$ in the $q^2 > 11.0$ GeV$^2/c^4$ region (upper left), $q^2$ in the $M_X < 1.5$ GeV$/c^2$ region (upper right), and lepton energy in the $q^2 > 11.0$ GeV$^2/c^4$ and $M_X < 1.5$ GeV$/c^2$ region (lower left). The $q^2$ distribution for $M_X < 1.5$ GeV$/c^2$ extends to $q^2$ of zero, but the $B \rightarrow X_u l \overline{\nu}$ signal is concentrated in the region of $q^2$ above 11.0 GeV$^2/c^4$ that is shown.
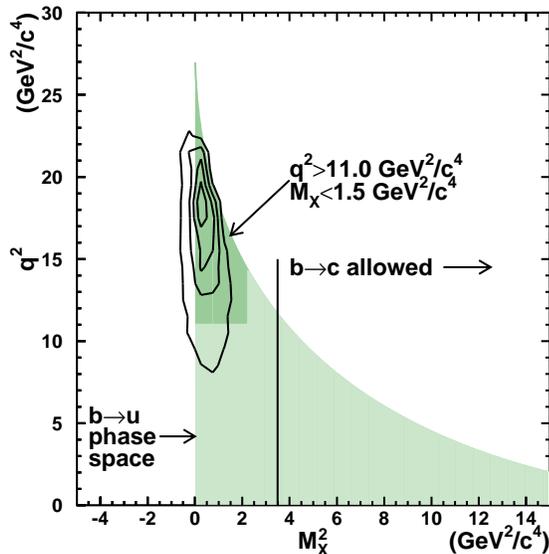
Figure 8.2: Region of High $S^2/(S+B)$ for the $B \to X_u l\overline{\nu}$ Signal.
The contours mark level curves of $S^2/(S+B)$ in the reconstructed variables. The phase space for $B \to X_u l\overline{\nu}$ and the region used for the nominal $|V_{ub}|$ result are in the true variables, but correlation is sufficient to make a qualitative comparison. Because the resolution is asymmetric, the region of high $S^2/(S+B)$ is shifted systematically to lower reconstructed $M_X^2$ than the true $M_X^2$ of the $B \to X_u l\overline{\nu}$ events.

and OPE calculations, because they do not predict specific hadronic final states. For example, the calculations predict the $B \to X_u l\overline{\nu}$ decay rate is above $q^2$ of 11 GeV$^2/c^4$, but not how much of that rate is $B \to \pi l\overline{\nu}$ and how much is $B \to \rho l\overline{\nu}$. The calculations are also limited in that they are not expected to predict features of the inclusive spectrum which correspond to energy scales that are small compared to $\Lambda_{QCD}$. The second use of models is to extrapolate and interpolate between the fit's region of sensitivity to $B \to X_u l\overline{\nu}$ and regions of interest for extracting $|V_{ub}|$. The fit is primarily sensitive to low $M_X^2$ and high $q^2$, but within this region the sensitivity is strongly biased toward high charged lepton momentum. Hence the models are relied on to extrapolate to the full range of lepton energies. The dependence of the sensitivity on the lepton energy can be seen in the lepton energy projection shown Figure 8.1. The background, and hence the $S^2/(S+B)$, varies dramatically as a function of the lepton energy. Issues related to the lepton energy sensitivity are discussed in more detail in the next section.

As shown in Figure 8.2, the region of sensitivity to $B \to X_u l\overline{\nu}$ does not have well-defined boundaries, nor is the sensitivity uniform within the region. In order to apply the theoretical calculations, a model is used to infer the partial branching fraction in a region, $\Delta\mathcal{B}_{\text{region}}$. This is a much more limited use of the model than is necessary to infer the full branching fraction, so a smaller model dependence is expected. For a particular model and region, the inferred partial branching

fraction is

$$\Delta\mathcal{B}_{\text{region}}^{\text{model}} = f_{\text{region}}^{\text{model}} \times \mathcal{B}(B \to X_u l\overline{\nu})^{\text{model}}, \tag{8.1}$$

where $\mathcal{B}(B \to X_u l\overline{\nu})^{\text{model}}$ is the previously presented branching fraction calculated from data assuming a model and $f_{\text{region}}^{\text{model}}$ is the fraction of the differential decay rate predicted by that model in the region being considered. As indicated by the subscripts and superscripts, both $\Delta\mathcal{B}_{\text{region}}^{\text{model}}$ and $f_{\text{region}}^{\text{model}}$ depend on the model and the region, but $\mathcal{B}(B \to X_u l\overline{\nu})^{\text{model}}$ only depends on the model and the data. Table 8.1 shows the partial branching fractions for a variety of regions and models (the values of $f_{\text{region}}^{\text{model}}$ for these regions is shown in Table 8.2). The models used are described in Section 6.2.4. They span a range of hadronic structures, from the ISGW2 model which includes only resonances to the all nonresonant models. The spread of the models is significantly less for the region bounded by $q^2 > 11$ GeV$^2/c^4$ and $M_X < 1.5$ GeV/$c^2$ than for the total rate. This is because, of the regions considered, that region most closely resembles the region of sensitivity of the fit (see Figure 8.2).

Table 8.1: Inferred $B \to X_u l\overline{\nu}$ Partial Branching Fractions, $\Delta\mathcal{B}_{\text{region}}^{\text{model}}$, for Various Regions and Models.
Each row is the results of a fit to the data with the specified model. The columns are the partial branching fractions in the specified region, which have be calculated from the fit result for the model and the fraction of the model in the specified region (see Table 8.2). The unit on the bounds are GeV/$c^2$ for the $M_X$ variable, GeV$^2/c^4$ for the $q^2$ variables, and GeV for the $E_\ell$ variable. All entries are in units of $10^{-3}$.

| Model Region | Total | $q^2 > 6$ $M_X < M_D$ | $q^2 > 8$ $M_X < 1.7$ | $q^2 > 11$ $M_X < 1.5$ | $q^2 > (M_B - M_{D^*})^2$ | $q^2 > (M_B - M_D)^2$ | $E_\ell > 2.2$ |
|---|---|---|---|---|---|---|---|
| Only Resonances | | | | | | | |
| ISGW2 | 1.15 | 0.73 | 0.57 | 0.35 | 0.33 | 0.38 | 0.26 |
| Hard $q^2$ | 0.93 | 0.67 | 0.56 | 0.37 | 0.34 | 0.39 | 0.24 |
| Soft $q^2$ | 1.35 | 0.76 | 0.57 | 0.33 | 0.30 | 0.36 | 0.28 |
| High $\cos\theta_{W\ell}$ | 1.10 | 0.70 | 0.55 | 0.34 | 0.31 | 0.37 | 0.27 |
| Low $\cos\theta_{W\ell}$ | 1.20 | 0.76 | 0.60 | 0.37 | 0.34 | 0.40 | 0.26 |
| Mix of Resonances and Nonresonant | | | | | | | |
| Bsg | 1.78 | 1.01 | 0.78 | 0.47 | 0.43 | 0.51 | 0.35 |
| InclGen | 1.92 | 1.08 | 0.83 | 0.50 | 0.46 | 0.54 | 0.38 |
| All Nonresonant | | | | | | | |
| Nominal | 2.87 | 1.33 | 0.97 | 0.56 | 0.52 | 0.63 | 0.40 |
| High Mass & Width | 3.46 | 1.30 | 0.93 | 0.52 | 0.51 | 0.62 | 0.38 |
| Low Mass & Width | 2.53 | 1.26 | 0.93 | 0.54 | 0.49 | 0.59 | 0.39 |
| Mininum | 0.93 | 0.67 | 0.55 | 0.33 | 0.30 | 0.36 | 0.24 |
| Maximum | 3.46 | 1.33 | 0.97 | 0.56 | 0.52 | 0.63 | 0.40 |
| Fractional Variation | 0.58 | 0.33 | 0.28 | 0.26 | 0.28 | 0.28 | 0.25 |

The central values for the partial branching fractions, $\Delta\mathcal{B}_{\text{region}}$, are taken to be the center of the range covered by the various $B \to X_u l\overline{\nu}$ models and the uncertainty is assigned to cover the full range of models. Table 8.3 shows the partial branching fractions for the various regions. The partial branching fraction of the region above 2.2 GeV has been previously measured at CLEO by reconstructing only a lepton and without using models as extensively. The result for that measurement is $(.228 \pm .015 \pm .035) \times 10^{-3}$, which is consistent with this measurement.

Table 8.2: Values of $f_{\text{region}}^{\text{model}}$ for Various Regions and Models.
The values for the regions with an additional 1.8 GeV lepton energy cut are used in the next section.

| Model Region | Total | $q^2 >6$ $M_X < M_D$ | $q^2 >8$ $M_X < 1.7$ | $q^2 >11$ $M_X < 1.5$ | $q^2 > (M_B - M_{D^*})^2$ | $q^2 > (M_B - M_D)^2$ | $E_\ell > 2.2$ |
|---|---|---|---|---|---|---|---|
| Only Resonances | | | | | | | |
| ISGW2 | 1.00 | 0.63 | 0.50 | 0.31 | 0.28 | 0.33 | 0.23 |
| Hard $q^2$ | 1.00 | 0.72 | 0.60 | 0.40 | 0.37 | 0.43 | 0.26 |
| Soft $q^2$ | 1.00 | 0.56 | 0.43 | 0.24 | 0.22 | 0.26 | 0.21 |
| High $\cos\theta_{W\ell}$ | 1.00 | 0.63 | 0.50 | 0.31 | 0.29 | 0.33 | 0.24 |
| Low $\cos\theta_{W\ell}$ | 1.00 | 0.63 | 0.50 | 0.31 | 0.28 | 0.33 | 0.22 |
| Mix of Resonances and Nonresonant | | | | | | | |
| Bsg | 1.00 | 0.57 | 0.44 | 0.26 | 0.24 | 0.28 | 0.20 |
| InclGen | 1.00 | 0.56 | 0.43 | 0.26 | 0.24 | 0.28 | 0.20 |
| All Nonresonant | | | | | | | |
| Nominal | 1.00 | 0.46 | 0.34 | 0.19 | 0.18 | 0.22 | 0.14 |
| High Mass & Width | 1.00 | 0.37 | 0.27 | 0.15 | 0.15 | 0.18 | 0.11 |
| Low Mass & Width | 1.00 | 0.50 | 0.37 | 0.21 | 0.19 | 0.23 | 0.15 |
| | | | | | | | |
| Regions defined with $E_\ell > 1.8$ GeV cut in addition to the indicated bounds | | | | | | | |
| Only Resonances | | | | | | | |
| ISGW2 | | 0.43 | 0.37 | 0.26 | 0.24 | 0.27 | |
| Hard $q^2$ | | 0.51 | 0.45 | 0.34 | 0.32 | 0.35 | |
| Soft $q^2$ | | 0.37 | 0.31 | 0.20 | 0.18 | 0.21 | |
| High $\cos\theta_{W\ell}$ | | 0.44 | 0.38 | 0.26 | 0.24 | 0.28 | |
| Low $\cos\theta_{W\ell}$ | | 0.42 | 0.36 | 0.25 | 0.23 | 0.26 | |
| Mix of Resonances and Nonresonant | | | | | | | |
| Bsg | | 0.38 | 0.32 | 0.22 | 0.20 | 0.23 | |
| InclGen | | 0.38 | 0.31 | 0.21 | 0.20 | 0.23 | |
| All Nonresonant | | | | | | | |
| Nominal | | 0.30 | 0.25 | 0.16 | 0.16 | 0.18 | |
| High Mass & Width | | 0.24 | 0.20 | 0.13 | 0.12 | 0.14 | |
| Low Mass & Width | | 0.33 | 0.27 | 0.18 | 0.17 | 0.19 | |

Table 8.3: Inferred $B \to X_u l\overline{\nu}$ Partial Branching Fractions, $\Delta\mathcal{B}_{\text{region}}$, for Various Regions with the Model Dependence Assessed.
The HQET calculations of the fraction of the $B \to X_u l\overline{\nu}$ decay rate in the region is also shown for reference [26].

| Region | $\Delta\mathcal{B}_{\text{region}} \pm$ Statistical $\pm$ Detector $\pm B \to X_c l\overline{\nu}$ Model Dependence $\pm B \to X_u l\overline{\nu}$ Model Dependence | $f_{\text{region}}^{\text{HQET}}$ |
|---|---|---|
| Total | $(2.19 \pm 0.15 \pm 0.56 \pm 0.21 \pm 1.27) \times 10^{-3}$ | 1.0 |
| $M_X < M_D$, $q^2 >6$ GeV$^2/c^4$ | $(1.00 \pm 0.07 \pm 0.25 \pm 0.10 \pm 0.33) \times 10^{-3}$ | $0.46 \pm 0.07$ |
| $M_X < 1.7$ GeV/$c^2$, $q^2 >8$ GeV$^2/c^4$ | $(0.76 \pm 0.05 \pm 0.19 \pm 0.07 \pm 0.21) \times 10^{-3}$ | $0.33 \pm 0.06$ |
| $M_X < 1.5$ GeV/$c^2$, $q^2 >11$ GeV$^2/c^4$ | $(0.44 \pm 0.03 \pm 0.11 \pm 0.04 \pm 0.11) \times 10^{-3}$ | $0.18 \pm 0.05$ |
| $q^2 > (M_B - M_{D^*})^2$ | $(0.41 \pm 0.03 \pm 0.10 \pm 0.04 \pm 0.11) \times 10^{-3}$ | $0.17 \pm 0.05$ |
| $q^2 > (M_B - M_D)^2$ | $(0.49 \pm 0.03 \pm 0.12 \pm 0.05 \pm 0.13) \times 10^{-3}$ | $0.20 \pm 0.05$ |
| $E_\ell > 2.2$ GeV | $(0.32 \pm 0.02 \pm 0.08 \pm 0.03 \pm 0.08) \times 10^{-3}$ | - |

In the context of HQET and the OPE, another fraction, $f_{\text{region}}^{\text{HQET}}$, relates the partial branching fraction to the total branching fraction. This fraction is a theoretical calculation and does not involve any models. It therefore has only theoretical uncertainties, as opposed to $f_{\text{region}}^{\text{model}}$ which are model dependent. The fractions, $f_{\text{region}}^{\text{HQET}}$, for the five regions of the $q^2$ and $M_X$ plane shown in Table 8.3 have been calculated with an evaluation of the theoretical uncertainty by Bauer *et al.* in reference [26]. Some of the regions have also been evaluated in references [25]. From the total rate, $|V_{ub}|$ can be extracted directly [66],

$$|V_{ub}| \quad = \quad \left[3.07 \pm 0.12 \times 10^{-3}\right] \left[ \frac{\Delta\mathcal{B}_{\text{region}}}{.001 \times f_{\text{region}}^{\text{HQET}}} \frac{1.6\text{ps}}{\tau_B} \right]^{1/2}.$$

The $|V_{ub}|$ results that correspond to the five regions and the total rate are shown in Table 8.4. The $|V_{ub}|$ results are calculated directly from the $\Delta\mathcal{B}_{\text{region}}$ and the uncertainties are simply halved because of the square root. A final uncertainty due to the calculation of the $f_{\text{region}}^{\text{HQET}}$ fractions is also included. The theoretical uncertainty is much more reliably determined than the model dependence uncertainty. Therefore, the region $M_X < 1.5$ GeV/$c^2$, $q^2 > 11$ GeV$^2$/$c^4$, which has the smallest $B \to X_u l\overline{\nu}$ model dependence, is taken as the central value. As previously mentioned, this is also the region that is most similar to the fit's region of sensitivity (see Figure 8.2) and is therefore the best summary of the data.

Table 8.4: $|V_{ub}|$ Results for the Various Regions.

| Region | $|V_{ub}| \pm$ Statistical $\pm$ Detector $\pm B \to X_c l\overline{\nu}$ Model Dependence $\pm B \to X_u l\overline{\nu}$ Model Dependence $\pm$ Theory Uncertainty |
|---|---|
| Total | $(4.55 \pm 0.16 \pm 0.58 \pm 0.22 \pm 1.31 \pm 0.18) \times 10^{-3}$ |
| $M_X < M_D$, $q^2 > 6$ GeV$^2$/$c^4$ | $(4.53 \pm 0.16 \pm 0.58 \pm 0.22 \pm 0.75 \pm 0.38) \times 10^{-3}$ |
| $M_X < 1.7$ GeV/$c^2$, $q^2 > 8$ GeV$^2$/$c^4$ | $(4.68 \pm 0.16 \pm 0.60 \pm 0.23 \pm 0.65 \pm 0.46) \times 10^{-3}$ |
| $M_X < 1.5$ GeV/$c^2$, $q^2 > 11$ GeV$^2$/$c^4$ | $(4.79 \pm 0.17 \pm 0.61 \pm 0.23 \pm 0.61 \pm 0.67) \times 10^{-3}$ |
| $q^2 > (M_B - M_{D^*})^2$ | $(4.78 \pm 0.17 \pm 0.61 \pm 0.23 \pm 0.66 \pm 0.74) \times 10^{-3}$ |
| $q^2 > (M_B - M_D)^2$ | $(4.74 \pm 0.17 \pm 0.60 \pm 0.23 \pm 0.65 \pm 0.64) \times 10^{-3}$ |

## 8.2 Reducing the Dependence on the Lepton Energy

The motivation for using the neutrino reconstruction to extract $|V_{ub}|$ is that there are regions of the $M_X$ and $q^2$ plane in which the partial decay rate of $B \to X_u l\overline{\nu}$ can be calculated more reliably than the partial decay rate in the lepton energy endpoint region. It is therefore troubling that the sensitivity to the $B \to X_u l\overline{\nu}$ contribution varies considerably with the lepton energy. In the previous section this is addressed by simply relying on the models to appropriately assess the model dependence of the result. But when using a set of models to assess the model dependence it is never
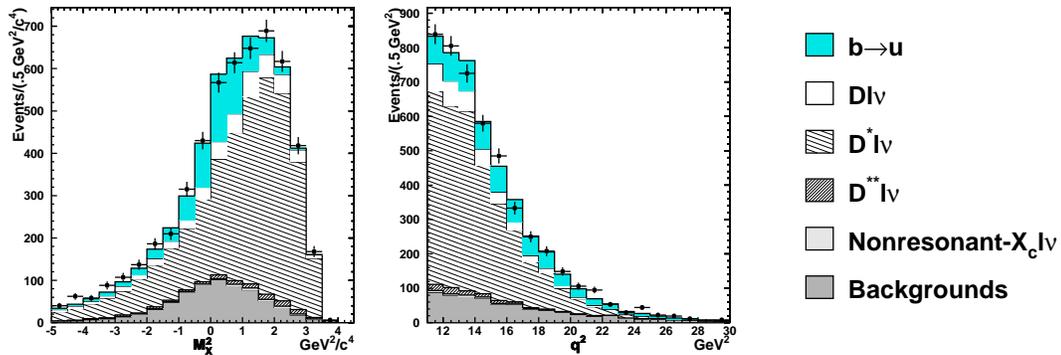
Figure 8.3: Projections Restricted to the Region of $B \to X_u l\overline{\nu}$ Sensitivity for $E_\ell > 1.8$ GeV. Projections of $M_X^2$ in the $q^2 > 11.0$ GeV$^2/c^4$ and $E_\ell > 1.8$ GeV region (left) and $q^2$ in the $M_X < 1.5$ GeV$/c^2$ and $E_\ell > 1.8$ GeV region (center). The $q^2$ distribution for $M_X < 1.5$ GeV$/c^2$ and $E_\ell > 1.8$ GeV extends to $q^2$ of zero, but the $B \to X_u l\overline{\nu}$ signal is concentrated in the region of $q^2$ above 11.0 GeV$^2/c^4$ that is shown.

clear when a sufficient variation has been made. The theoretical calculations on the other hand are expansions in which the higher order terms can be used to assign a more reliable uncertainty.

The simplest option for removing the unwanted lepton energy dependence is to remove the lepton energy information from the fit. The problem with this technique is that it removes too much information. The size of the $B \to X_u l\overline{\nu}$ contribution below a lepton energy of $\approx 1.8$ GeV is very small and the $B \to X_c l\overline{\nu}$ contribution is large. By placing a 1.8 GeV lepton energy requirement on the data sample and then using only the $M_X^2$ and $q^2$ variables in the fit, the amount of the lepton energy information used in the measurement is well defined. The $M_X^2$ and $q^2$ distributions in the $B \to X_u l\overline{\nu}$ sensitive region above 1.8 GeV are shown in Figure 8.3. Although calculations of $f_{\text{region}}^{\text{HQET}}$ in the $M_X^2$ and $q^2$ regions discussed in the previous section have not been calculated with a lepton energy cut, such a calculation is possible and would be more reliable than using a model dependent extrapolation.

Table 8.5 shows the inferred partial branching fractions for the fit to the data with a 1.8 GeV lepton energy cut using only the $M_X^2$ and $q^2$ variables in the fit. The results are shown for $M_X^2$ and $q^2$ regions previously discussed which involves an extrapolation to the full lepton energy range and those same regions with the 1.8 GeV lepton energy requirement in addition to the other boundaries which does not require an extrapolation of the lepton energy range. The apparent model dependence for the regions without the lepton energy requirement is slightly reduced from the three-dimensional fit. In both cases there is still an extrapolation from a region of high lepton energy to the full range. For the regions with the lepton energy requirement added, there is no extrapolation in the lepton energy dimension. For the largest region ($M_X < M_D$, $q^2 > 6$ GeV$^2/c^4$), there is a moderate difference between the model dependence assessment with and without the lepton energy cut, as expected. For

the smaller regions, which are more restrictive in energy the difference is not significant. One reason for the small change in the model dependence for the smaller regions, is that a very large fraction of the predicted rate falls above 1.8 GeV. In particular, for the nominal region, $M_X < 1.5$ GeV$/c^2$, $q^2 > 11$ GeV$^2/c^4$, $\approx 80\%$ of decay rate has a lepton energy above 1.8 GeV. The small change may also be an indication that the dominant difference between the models is their hadronic structure (i.e., resonant versus nonresonant) rather than the differential decay distribution. Although there is not a large change in the model dependence assessment, the results are intrinsically more reliable, because there is a smaller extrapolation.

Table 8.5: Inferred $B \to X_u l\overline{\nu}$ Partial Branching Fractions, $\Delta\mathcal{B}^{\text{model}}_{\text{region}}$, for Various Regions and Models Based on a Two-Dimensional Fit with $E_\ell > 1.8$ GeV Data.
The first set of numbers includes an extrapolation to the full range of lepton energies. In the second set of numbers, the regions are as specified, but with the additional restriction that $E_\ell > 1.8$ GeV. The unit on the bounds are GeV$/c^2$ for the $M_X$ variable, GeV$^2/c^4$ for the $q^2$ variables, and GeV for the $E_\ell$ variable. All entries are in units of $10^{-3}$.

| Model | Total | $q^2 > 6$ $M_X < M_D$ | $q^2 > 8$ $M_X < 1.7$ | $q^2 > 11$ $M_X < 1.5$ | $q^2 > (M_B - M_{D^*})^2$ | $q^2 > (M_B - M_D)^2$ | $E_\ell > 2.2$ |
|---|---|---|---|---|---|---|---|
| Only Resonances | | | | | | | |
| ISGW2 | 1.13 | 0.71 | 0.56 | 0.35 | 0.32 | 0.37 | 0.26 |
| Hard $q^2$ | 0.88 | 0.64 | 0.53 | 0.35 | 0.33 | 0.38 | 0.23 |
| Soft $q^2$ | 1.35 | 0.76 | 0.57 | 0.33 | 0.30 | 0.36 | 0.28 |
| High $\cos\theta_{W\ell}$ | 1.10 | 0.69 | 0.55 | 0.34 | 0.31 | 0.37 | 0.26 |
| Low $\cos\theta_{W\ell}$ | 1.16 | 0.73 | 0.57 | 0.36 | 0.33 | 0.38 | 0.25 |
| Mix of Resonances and Nonresonant | | | | | | | |
| Bsg | 1.77 | 1.00 | 0.77 | 0.46 | 0.43 | 0.50 | 0.35 |
| InclGen | 1.96 | 1.10 | 0.85 | 0.51 | 0.47 | 0.55 | 0.38 |
| All Nonresonant | | | | | | | |
| Nominal | 2.76 | 1.28 | 0.93 | 0.53 | 0.50 | 0.60 | 0.39 |
| High Mass & Width | 3.32 | 1.24 | 0.89 | 0.50 | 0.49 | 0.59 | 0.36 |
| Low Mass & Width | 2.30 | 1.15 | 0.85 | 0.49 | 0.45 | 0.53 | 0.35 |
| Minimum | 0.88 | 0.64 | 0.53 | 0.33 | 0.30 | 0.36 | 0.23 |
| Maximum | 3.32 | 1.28 | 0.93 | 0.53 | 0.50 | 0.60 | 0.39 |
| Fractional Variation | 0.58 | 0.33 | 0.27 | 0.24 | 0.26 | 0.26 | 0.25 |
| Regions defined with $E_\ell > 1.8$ GeV cut in addition to the indicated bounds | | | | | | | |
| Only Resonances | | | | | | | |
| ISGW2 | | 0.48 | 0.41 | 0.29 | 0.27 | 0.30 | |
| Hard $q^2$ | | 0.45 | 0.40 | 0.30 | 0.28 | 0.31 | |
| Soft $q^2$ | | 0.50 | 0.41 | 0.27 | 0.25 | 0.29 | |
| High $\cos\theta_{W\ell}$ | | 0.48 | 0.41 | 0.29 | 0.27 | 0.30 | |
| Low $\cos\theta_{W\ell}$ | | 0.48 | 0.41 | 0.29 | 0.27 | 0.31 | |
| Mix of Resonances and Nonresonant | | | | | | | |
| Bsg | | 0.67 | 0.56 | 0.38 | 0.36 | 0.41 | |
| InclGen | | 0.74 | 0.62 | 0.42 | 0.39 | 0.45 | |
| All Nonresonant | | | | | | | |
| Nominal | | 0.84 | 0.68 | 0.45 | 0.43 | 0.49 | |
| High Mass & Width | | 0.81 | 0.65 | 0.42 | 0.41 | 0.48 | |
| Low Mass & Width | | 0.76 | 0.62 | 0.41 | 0.38 | 0.44 | |
| Minimum | | 0.45 | 0.40 | 0.27 | 0.25 | 0.29 | |
| Maximum | | 0.84 | 0.68 | 0.45 | 0.43 | 0.49 | |
| Fractional Variation | | 0.30 | 0.26 | 0.25 | 0.27 | 0.27 | |

The partial branching fractions and $|V_{ub}|$ for fit to the data with the 1.8 GeV lepton energy cut using only the $M_X^2$ and $q^2$ dimensions are calculated in the same ways as for the full fit described in

the previous section. The resulting branching fractions are shown in Table 8.6 and the values of $|V_{ub}|$ are shown in Table 8.7. The central value is chosen to be the $|V_{ub}|$ calculated for the region $M_X < 1.5$ GeV/$c^2$, $q^2 > 11$ GeV$^2$/$c^4$, because this region most closely resembles the region of sensitivity of the fit and the lepton sensitivity is flat over more than 80% of available decay rate. The resulting central value for $|V_{ub}|$ is

$$|V_{ub}| \quad = \quad (4.73 \pm 0.23 \pm 0.82 \pm 0.18 \pm 0.56 \pm 0.66) \times 10^{-3}$$

where the errors are statistical, detector systematics, $B \rightarrow X_c l \overline{\nu}$ model dependence, $B \rightarrow X_u l \overline{\nu}$ model dependence, and theoretical uncertainty respectively The result is consistent with the CLEO measurement using the lepton energy endpoint [9], $E_\ell > 2.2$ GeV, $|V_{ub}| = (4.08 \pm 0.34 \pm 0.44 \pm 0.16 \pm 0.24) \times 10^{-3}$, where the errors are statistical, systematic, and two different theory errors. The first theory error is due to the calculation of the fraction of the total rate in the endpoint region, and the second theory error is due to the uncertainty of the relationship between the total rate and $|V_{ub}|$ (see Equation 8.2). Both of these sources of uncertainty are included in the single theory error quoted from the measurement presented here.

Table 8.6: Inferred $B \rightarrow X_u l \overline{\nu}$ Partial Branching Fractions, $\Delta \mathcal{B}_{\text{region}}$, for Various Regions Based on a Two-Dimensional Fit with $E_\ell > 1.8$ GeV Data.

| Region | $\Delta \mathcal{B}_{\text{region}} \pm$ Statistical $\pm$ Detector $\pm B \rightarrow X_c l \overline{\nu}$ Model Dependence $\pm B \rightarrow X_u l \overline{\nu}$ Model Dependence |
|---|---|
| Total | $(2.10 \pm 0.21 \pm 0.73 \pm 0.16 \pm 1.22) \times 10^{-3}$ |
| $M_X < M_D$, $q^2 > 6$ GeV$^2$/$c^4$ | $(0.96 \pm 0.10 \pm 0.33 \pm 0.07 \pm 0.32) \times 10^{-3}$ |
| $M_X < 1.7$ GeV/$c^2$, $q^2 > 8$ GeV$^2$/$c^4$ | $(0.73 \pm 0.07 \pm 0.25 \pm 0.06 \pm 0.20) \times 10^{-3}$ |
| $M_X < 1.5$ GeV/$c^2$, $q^2 > 11$ GeV$^2$/$c^4$ | $(0.43 \pm 0.04 \pm 0.15 \pm 0.03 \pm 0.10) \times 10^{-3}$ |
| $q^2 > (M_B - M_{D^*})^2$ | $(0.40 \pm 0.04 \pm 0.14 \pm 0.03 \pm 0.10) \times 10^{-3}$ |
| $q^2 > (M_B - M_D)^2$ | $(0.48 \pm 0.05 \pm 0.17 \pm 0.04 \pm 0.12) \times 10^{-3}$ |
| $E_\ell > 2.2$ GeV | $(0.31 \pm 0.03 \pm 0.11 \pm 0.02 \pm 0.08) \times 10^{-3}$ |
| | |
| $M_X < M_D$, $q^2 > 6$ GeV$^2$/$c^4$, $E_\ell > 1.8$ GeV | $(0.65 \pm 0.06 \pm 0.22 \pm 0.05 \pm 0.19) \times 10^{-3}$ |
| $M_X < 1.7$ GeV/$c^2$, $q^2 > 8$ GeV$^2$/$c^4$, $E_\ell > 1.8$ GeV | $(0.54 \pm 0.05 \pm 0.19 \pm 0.04 \pm 0.14) \times 10^{-3}$ |
| $M_X < 1.5$ GeV/$c^2$, $q^2 > 11$ GeV$^2$/$c^4$, $E_\ell > 1.8$ GeV | $(0.36 \pm 0.04 \pm 0.12 \pm 0.03 \pm 0.09) \times 10^{-3}$ |
| $q^2 > (M_B - M_{D^*})^2$, $E_\ell > 1.8$ GeV | $(0.34 \pm 0.03 \pm 0.12 \pm 0.03 \pm 0.09) \times 10^{-3}$ |
| $q^2 > (M_B - M_D)^2$, $E_\ell > 1.8$ GeV | $(0.39 \pm 0.04 \pm 0.14 \pm 0.03 \pm 0.10) \times 10^{-3}$ |

## 8.3  Comparison of This Method with the Lepton Energy Endpoint Method

The $|V_{ub}|$ measurement based on the $q^2$ and $M_X$ information has different theoretical assumptions than the lepton energy endpoint measurement. Both measurements rely on HQET and the OPE, but all orders of the OPE are needed for the lepton endpoint while for the $q^2$ and $M_X$ regions only

Table 8.7: $|V_{ub}|$ Results for the Various Regions Based on Two-Dimensional Fit.

| Region | $|V_{ub}| \pm$ Statistical $\pm$ Detector $\pm B \to X_c l\overline{\nu}$ Model Dependence $\pm B \to X_u l\overline{\nu}$ Model Dependence $\pm$ Theory Uncertainty |
|---|---|
| Total | $(4.45 \pm 0.22 \pm 0.77 \pm 0.17 \pm 1.29 \pm 0.17) \times 10^{-3}$ |
| $M_X < M_D$, $q^2 > 6$ GeV$^2/c^4$ | $(4.43 \pm 0.22 \pm 0.77 \pm 0.17 \pm 0.74 \pm 0.37) \times 10^{-3}$ |
| $M_X < 1.7$ GeV/$c^2$, $q^2 > 8$ GeV$^2/c^4$ | $(4.59 \pm 0.23 \pm 0.80 \pm 0.18 \pm 0.63 \pm 0.45) \times 10^{-3}$ |
| $M_X < 1.5$ GeV/$c^2$, $q^2 > 11$ GeV$^2/c^4$ | $(4.73 \pm 0.23 \pm 0.82 \pm 0.18 \pm 0.56 \pm 0.66) \times 10^{-3}$ |
| $q^2 > (M_B - M_{D^*})^2$ | $(4.72 \pm 0.23 \pm 0.82 \pm 0.18 \pm 0.60 \pm 0.73) \times 10^{-3}$ |
| $q^2 > (M_B - M_D)^2$ | $(4.68 \pm 0.23 \pm 0.81 \pm 0.18 \pm 0.60 \pm 0.64) \times 10^{-3}$ |

the first and second order terms are used. Of course it not possible to calculate the OPE to all orders, especially since at each order new nonperturbative parameters appear (e.g., $\overline{\Lambda}$, $\lambda_1$, $\lambda_2$, ...). Instead, the most singular terms at each order in the expansion are summed up into a shape function. This shape function contains all the nonperturbative effects at order $\Lambda_{QCD}/m_B$, but cannot be predicted from first principles. Like the $\overline{\Lambda}$, $\lambda_1$ and $\lambda_2$ parameters, the shape function is a property of the $B$ meson. It can therefore be measured using the photon energy spectrum of the $B \to X_s \gamma$ process, using a similar summation of the OPE terms, and applied to a calculation of the $B \to X_u l\overline{\nu}$ decay rate.

The first vulnerability of these measurements is due to the neglected terms. For the lepton endpoint measurement, this takes the form of an additional function which arises at order $\Lambda_{QCD}/m_B$. This function has a different coefficient in the $B \to X_s \gamma$ differential decay rate than it does in the $B \to X_u l\overline{\nu}$ differential decay rate [67]. For the $q^2$ and $M_X$ based measurement, the uncertainties are due to the higher order terms in the OPE that have been neglected, the uncertainty of which can be evaluated by varying the $(\Lambda_{QCD}/m_B)^3$ terms.

Another vulnerability is that the region used in the lepton endpoint measurement, $E_\ell > 2.2$, is very small and included in that region is the region $q^2 \approx q^2_{\max}$, where the daughter up quark has such low momentum that it can become part of the light degrees of freedom [68]. This is also an issue for the $q^2$ and $M_X$ based measurement, but the size of the region measured is larger ($\approx 18\%$ of the total rate compared to $\approx 13\%$), so the effect is diluted. If the size of the region could be further extended, this would become less important. For the largest region calculated by Bauer, et al. [26], the fraction of the total rate is $\approx 46\%$. The combined effect of these two neglected effects is that there is probably a higher theoretical uncertainty on the lepton endpoint measurement than the uncertainty on that measurement quoted above. This makes the two measurements more comparable in their precisions.

Finally, because the normalization of the shape function is extracted from $B \to X_s \gamma$, the assumption that $|V_{tb} V_{ts}^*| = |V_{cb}|$ and that there are no non-standard model contributions to $B \to X_s \gamma$ is implicit in the lepton endpoint measurement. These assumptions are acceptable, but should be kept in mind if the measurements of the CKM parameters are found to not obey the unitarity constraint.

# Chapter 9    Conclusion

We have studied the fully differential decay rate for inclusive semileptonic $B$ meson decay. The branching fractions of the exclusive final states $B \to Dl\overline{\nu}$, $B \to D^*l\overline{\nu}$, $B \to D^{**}l\overline{\nu}$, nonresonant $B \to X_cl\overline{\nu}$, and $B \to X_ul\overline{\nu}$ have been measured using a fit to the the differential decay rate. The measurements of the exclusive branching fraction form a complete model of composition of semileptonic $B$ decay. From this model, moments of the $B \to X_cl\overline{\nu}$ differential decay rate are extracted. The partial branching fraction of the $B \to X_ul\overline{\nu}$ decay in a restricted region of phase space is also extracted. These measurements are interpreted in the context of the HQET and OPE calculations in order to extract the CKM parameters $|V_{ub}|$ and $|V_{cb}|$. The value of $|V_{ub}|$ obtained is

$$|V_{ub}| = (4.73 \pm 0.23 \pm 0.82 \pm 0.18 \pm 0.56 \pm 0.66) \times 10^{-3}, \tag{9.1}$$

where the uncertainties are due to statistics, detector systematics, $B \to X_cl\overline{\nu}$ model dependence, $B \to X_ul\overline{\nu}$ model dependence, and theoretical uncertainties. For $|V_{cb}|$ the value obtained is

$$|V_{cb}| = (4.12 \pm .10 \pm 0.09 \pm 0.16) \times 10^{-2}, \tag{9.2}$$

where the uncertainties are due to the measurement of the semileptonic $B$ decay rate, the measurements required to constrain $\overline{\Lambda}$ and $\lambda_1$, and theoretical uncertainties due to the third-order terms of the $\Lambda_{QCD}/m_B$ expansion.

The moments measurements were also used to test the HQET and OPE predictions. In particular, good agreement is found in a comparison of the variation of the $\langle M_X^2 - \overline{M}_D^2 \rangle$ moment as a function of the minimum lepton energy requirement and the theoretical prediction (see Figure 7.3). Measurements of the $\langle (M_X^2 - \langle M_X^2 \rangle)^2 \rangle$ moments were also found to be consistent with the prediction (see Figure 7.2). The first and second moments of the $q^2$ distribution, $\langle q^2 \rangle$ and $\langle (q^2 - \langle q^2 \rangle)^2 \rangle$ have also been measured, although there are no calculations of the $q^2$ moments currently available with which to interpret the results.

The moments measurements have reached a point where improved constraints on the nonperturbative parameters, $\overline{\Lambda}$ and $\lambda_1$, are not needed. The uncertainty due to the third-order terms in the expansion dominate the errors in the measurement of $|V_{cb}|$. Higher-precision measurement could however be used to test the theory and to better understand the applicability of the quark-hadron duality assumption.

For the moments measurements both the detector systematic and model dependence uncertainties

151

are large. In future measurements, the detector systematics could be reduced by using control sample to better measure the detector response. This would be particularly useful for the quantifying the number of splitoff showers produced by hadronic tracks. This is difficult because a sample of events would be needed in which there are charged hadrons but there are no real photons. Most hadronic processes will produce both $\pi^\pm$ and $\pi^0$. Sometimes a hadronic resonance structure can be used to insure that there are no $\pi^0$ mesons and hence no real photons in the event. One example is $\Upsilon(3S) \to \Upsilon(1S)\pi^+\pi^-$, $\Upsilon(1S) \to l^+l^-$.

The largest contribution to the model dependence uncertainties of the moments measurements is the lack of knowledge of the nonresonant $B \to X_c l\overline{\nu}$ hadronic mass distribution. This could be improved by making more direct use of the data. The BaBar collaboration has used a technique that involves fully reconstructing the other $B$ meson in the event [17]. This substantially suppresses the background in the high hadronic mass region, at the expense of a very large decrease in the efficiency. They then use the distribution of high-mass events to directly measure the mass moment, without recourse to models. The final results have comparable uncertainties to the measurements presented here. This technique however throws away a lot of the information available from the models. In particular, the $D$ and $D^*$ masses are very well known and the uncertainty on the $D^{**}$ masses is small compared to the uncertainty on the nonresonant $X_c$ mass distribution. A hybrid of the technique presented here and the technique used by BaBar might provide an improved measurement.

Another area in which the HQET and OPE predictions can be tested is in charmed meson decays. The mass of the charmed quark, $\approx 1.5$ GeV, is substantially lower than that of the bottom quark $\approx 4.5$ GeV. The expansion parameter for the charmed system, $\Lambda_{QCD}/m_c \approx 1/7$, is dangerously close to unity, as opposed to the bottom system where it is $\Lambda_{QCD}/m_b \approx 1/22$. Although this is just a factor of three, it marks a considerable divide. It would not be surprising to find QCD effects at the 1 GeV energy scale. In fact the proton mass is just under a GeV and is generated primarily from nonperturbative QCD. It is considerably less likely to find a nonperturbative parameter nearing the bottom quark mass. It may still be of use to perform similar moments measurements in the charmed sector. Because the expansion parameter is smaller the effects of the higher order terms will be larger. If the expansion is found to work, it will give considerable confidence to the applications in the bottom system. Also, measurements using $D$ mesons produced at rest will likely be more precise than the measurements from $B$ meson decays.

For the $|V_{ub}|$ measurement, the technique presented is promising. The detector uncertainty could be improved by the same techniques suggested for the moments measurements. In addition, the $B \to X_u l\overline{\nu}$ measurement is sensitive to the $B \to K_L^0 X$ and $\mathcal{B}(B \to c \to \ell)$ rates which can also be measured directly. In order to improve the $B \to X_u l\overline{\nu}$ model dependence and the theoretical uncertainty, it is necessary to extend the region on sensitivity, which requires improving the signal to background in the $B \to X_u l\overline{\nu}$ region. There are several techniques involving reconstructing more

of the event that might be used to improve the resolution.

Both the lepton energy endpoint measurement and the $q^2$ and $M_X$ region measurement, presented here, relied heavily on the applicability of the HQET-OPE calculations in a very restricted region of phase space. A possible test of the shape function calculation used to interpret the lepton endpoint measurement is to calculate moments of the $q^2$ and $M_X$ distributions in the lepton endpoint region. Because a lepton energy requirement can be used to remove the vast majority of the $B \to X_c l\overline{\nu}$ background, the moments of the $B \to X_u l\overline{\nu}$ distributions could be measured cleanly. The moments can also be predicted with the same theory that is used to interpret the lepton endpoint partial branching fraction. A successful comparison would give significant confidence in the applicability of the calculations, and could be used to quantitatively limit the contributions of the higher order shape functions.

# Appendix A   Tables of the Systematic Studies for the Moments Results

The tables for the moment $\langle M_X^2 - \overline{M}_D^2 \rangle$ are presented in the text, see Tables 7.5 and 7.6.

Table A.1: Detector Systematics for $\langle (M_X^2 - \langle M_X^2 \rangle)^2 \rangle$.
Nominal value and shifts of the $\langle (M_X^2 - \langle M_X^2 \rangle)^2 \rangle$ moments due to the individual detector systematic uncertainties. All entries are in units of GeV$^4$/$c^8$

| Variation | $E_\ell > 1.0$ GeV | $E_\ell > 1.5$ GeV |
|---|---|---|
| Lepton Fake Rate | 0.047 / -0.019 | 0.034 / -0.071 |
| DELCO $b \to c \to \ell$ shape | -0.017 | 0.003 |
| DELCO $+1\sigma$ $b \to c \to \ell$ shape | -0.021 | 0.004 |
| DELCO $-1\sigma$ $b \to c \to \ell$ shape | -0.005 | 0.008 |
| Continuum Norm $\pm 10\%$ | 0.006 / -0.029 | 0.006 / -0.006 |
| $\mu$ Fakes, $E_\ell < 1.5$ GeV, $\pm 10\%$ | -0.068 / 0.070 | 0.000 / 0.000 |
| $\mu$ Fakes, $E_\ell > 1.5$ GeV, $\pm 10\%$ | 0.009 / -0.009 | 0.015 / -0.017 |
| $\mathcal{B}(b \to c \to \ell)$ $\pm 10\%$ | -0.015 / 0.015 | 0.002 / -0.002 |
| $\mathcal{B}(b \to$ baryons$)$ $\pm 20\%$ | -0.011 / 0.011 | -0.003 / 0.003 |
| # $K_L^0$ | -0.018 / 0.019 | -0.002 / 0.002 |
| Track Efficiency | 0.040 / -0.041 | 0.012 / -0.011 |
| # Fake Tracks | -0.041 / 0.038 | -0.012 / 0.011 |
| Shower Efficiency | 0.018 / -0.016 | 0.009 / -0.010 |
| # Fake Showers | -0.181 / 0.172 | -0.072 / 0.068 |
| Force Trk Multiplicity | 0.003 | 0.002 |
| Force Shwr Multiplicity | 0.037 | 0.019 |
| Final State Radiation | 0.088 | 0.025 |
| Lepton Efficiency | 0.022 | 0.017 |

Table A.2: Model Dependence for $\langle (M_X^2 - \langle M_X^2 \rangle)^2 \rangle$.
Nominal value and shifts of the $\langle (M_X^2 - \langle M_X^2 \rangle)^2 \rangle$ moments due to the individual model variations.
All entries are in units of GeV$^4/c^8$

| Variation | $E_\ell > 1.0$ GeV | $E_\ell > 1.5$ GeV |
|---|---|---|
| Nominal | 1.266 | 0.629 |
| $B \to Dl\overline{\nu}$ $\rho$ param | -0.016 / 0.004 | -0.005 / 0.000 |
| $B \to Dl\overline{\nu}$ $c_D$ param | -0.000 / -0.000 | 0.002 / -0.002 |
| $B \to D^*l\overline{\nu}$ $\rho$ param | -0.045 / 0.046 | 0.007 / 0.003 |
| $B \to D^*l\overline{\nu}$ $c_{A_1}$ param | 0.015 / -0.010 | 0.006 / -0.005 |
| $B \to D^*l\overline{\nu}$ R1/R2 1st eig-vec | 0.081 / -0.070 | 0.008 / -0.007 |
| $B \to D^*l\overline{\nu}$ R1/R2 2nd eig-vec | 0.028 / -0.028 | 0.015 / -0.016 |
| $B \to D^{**}l\overline{\nu}$ HQET model | 0.007 | 0.018 |
| $B \to D^{**}l\overline{\nu}$ $w$ slope | 0.077 / -0.058 | 0.019 / -0.016 |
| $B \to X_c l\overline{\nu}$ NonRes $w$ slope | 0.050 / -0.031 | 0.024 / -0.015 |
| $B \to X_c l\overline{\nu}$ Nonresonant Mass Dependence | | |
| Polynomial | 0.162 / -0.121 | 0.044 / -0.020 |
| Gauss $\mu = 2.0$ $\sigma^2 = .25$ | -0.543 | -0.156 |
| Gauss $\mu = 2.0$ $\sigma^2 = .75$ | -0.166 | -0.090 |
| Gauss $\mu = 2.0$ $\sigma^2 = 1.25$ | 0.004 | -0.059 |
| Gauss $\mu = 2.75$ $\sigma^2 = .25$ | -0.060 | 0.002 |
| Gauss $\mu = 2.75$ $\sigma^2 = .75$ | 0.150 | -0.003 |
| Gauss $\mu = 2.75$ $\sigma^2 = 1.25$ | 0.273 | 0.012 |
| Gauss $\mu = 3.5$ $\sigma^2 = .25$ | 0.397 | 0.071 |
| Gauss $\mu = 3.5$ $\sigma^2 = .75$ | 0.570 | 0.102 |
| Gauss $\mu = 3.5$ $\sigma^2 = 1.25$ | 0.615 | 0.104 |
| $B \to X_u l\overline{\nu}$ Model Dependence | | |
| $B \to X_s \gamma$ based | 0.002 | 0.002 |
| InclGen | 0.002 | 0.001 |
| All Non-Res | 0.010 | 0.006 |
| All Non-Res High Mass | -0.039 | -0.010 |
| All Non-Res Low Mass | 0.018 | 0.012 |
| Hard $q^2$ | 0.000 | -0.001 |
| Soft $q^2$ | -0.001 | -0.000 |
| High $\cos\theta_{W\ell}$ | -0.000 | -0.000 |
| Low $\cos\theta_{W\ell}$ | 0.000 | 0.000 |

Table A.3: Detector Systematics for $\langle E_\ell \rangle$.
Nominal value and shifts of the $\langle E_\ell \rangle$ moments due to the individual detector systematic uncertainties.
All entries are in units of MeV

| Variation | $E_\ell > 1.0$ GeV | $E_\ell > 1.5$ GeV |
|---|---|---|
| Nominal | 1550.944 | 1774.361 |
| Lepton Fake Rate | -0.894 / 0.428 | -0.697 / 1.302 |
| DELCO $b \to c \to \ell$ shape | 0.580 | -0.048 |
| DELCO $+1\sigma$ $b \to c \to \ell$ shape | 0.826 | -0.051 |
| DELCO $-1\sigma$ $b \to c \to \ell$ shape | 0.006 | -0.183 |
| Continuum Norm $\pm 10\%$ | 0.588 / -0.631 | -0.009 / 0.006 |
| $\mu$ Fakes, $E_\ell < 1.5$ GeV, $\pm 10\%$ | 0.743 / -0.692 | 0.000 / 0.000 |
| $\mu$ Fakes, $E_\ell > 1.5$ GeV, $\pm 10\%$ | -0.019 / 0.024 | -0.261 / 0.292 |
| $\mathcal{B}(b \to c \to \ell)$ $\pm 10\%$ | -0.182 / 0.184 | -0.144 / 0.140 |
| $\mathcal{B}(b \to$ baryons$)$ $\pm 20\%$ | 0.191 / -0.194 | 0.042 / -0.041 |
| # $K_L^0$ | -0.213 / 0.215 | -0.109 / 0.108 |
| Track Efficiency | -0.166 / 0.126 | -0.067 / 0.049 |
| # Fake Tracks | 0.536 / -0.482 | 0.172 / -0.160 |
| Shower Efficiency | 0.106 / -0.190 | -0.092 / 0.091 |
| # Fake Showers | 1.987 / -1.987 | 0.991 / -0.957 |
| Force Trk Multiplicity | -0.187 | -0.082 |
| Force Shwr Multiplicity | -0.438 | -0.311 |
| Final State Radiation | -1.730 | -0.455 |
| Lepton Efficiency | -0.681 | -0.294 |

Table A.4: Model Dependence for $\langle E_\ell \rangle$.

Nominal value and shifts of the $\langle E_\ell \rangle$ moments due to the individual model variations. All entries are in units of GeV

| Variation | $E_\ell > 1.0$ GeV | $E_\ell > 1.5$ GeV |
|---|---|---|
| Nominal | 1550.944 | 1774.361 |
| $B \to D l \overline{\nu}$ $\rho$ param | -1.024 / 0.824 | -0.771 / 0.496 |
| $B \to D l \overline{\nu}$ $c_D$ param | -0.106 / 0.031 | -0.262 / 0.068 |
| $B \to D^* l \overline{\nu}$ $\rho$ param | -4.443 / 3.067 | -3.128 / 3.313 |
| $B \to D^* l \overline{\nu}$ $c_{A_1}$ param | -0.580 / -0.082 | -0.082 / 0.799 |
| $B \to D^* l \overline{\nu}$ R1/R2 1st eig-vec | 4.422 / -5.763 | 2.540 / -2.441 |
| $B \to D^* l \overline{\nu}$ R1/R2 2nd eig-vec | 1.734 / -1.914 | 1.605 / -0.948 |
| $B \to D^{**} l \overline{\nu}$ HQET model | -0.357 | 0.005 |
| $B \to D^{**} l \overline{\nu}$ $w$ slope | -0.346 / 0.200 | -0.049 / 0.015 |
| $B \to X_c l \overline{\nu}$ NonRes $w$ slope | -0.231 / 0.114 | -0.100 / 0.054 |
| $B \to X_c l \overline{\nu}$ Nonresonant Mass Dependence | | |
| Polynomial | -1.229 / 0.788 | -0.099 / 0.095 |
| Gauss $\mu = 2.0$ $\sigma^2 = .25$ | 4.637 | 1.267 |
| Gauss $\mu = 2.0$ $\sigma^2 = .75$ | 1.332 | 0.436 |
| Gauss $\mu = 2.0$ $\sigma^2 = 1.25$ | 0.002 | 0.106 |
| Gauss $\mu = 2.75$ $\sigma^2 = .25$ | 0.429 | -0.095 |
| Gauss $\mu = 2.75$ $\sigma^2 = .75$ | -1.105 | -0.271 |
| Gauss $\mu = 2.75$ $\sigma^2 = 1.25$ | -2.142 | -0.624 |
| Gauss $\mu = 3.5$ $\sigma^2 = .25$ | -2.373 | -0.171 |
| Gauss $\mu = 3.5$ $\sigma^2 = .75$ | -4.361 | -1.217 |
| Gauss $\mu = 3.5$ $\sigma^2 = 1.25$ | -4.983 | -1.607 |
| $B \to X_u l \overline{\nu}$ Model Dependence | | |
| $B \to X_s \gamma$ based | 0.033 | -0.028 |
| InclGen | 0.072 | 0.007 |
| All Non-Res | -0.141 | -0.111 |
| All Non-Res High Mass | 0.492 | 0.135 |
| All Non-Res Low Mass | -0.273 | -0.235 |
| Hard $q^2$ | -0.143 | 0.016 |
| Soft $q^2$ | 0.188 | 0.014 |
| High $\cos \theta_{W\ell}$ | 0.020 | 0.012 |
| Low $\cos \theta_{W\ell}$ | -0.023 | -0.012 |

57

Table A.5: Detector Systematics for $\langle (E_\ell - \langle E_\ell \rangle)^2 \rangle$.
Nominal value and shifts of the $\langle (E_\ell - \langle E_\ell \rangle)^2 \rangle$ moments due to the individual detector systematic uncertainties. All entries are in units of $10^{-3} \times \text{GeV}^2$

| Variation | $E_\ell > 1.0$ GeV | $E_\ell > 1.5$ GeV |
|---|---|---|
| Nominal | 296.610 | 176.387 |
| Lepton Fake Rate | -0.008 / 0.011 | -0.187 / 0.272 |
| DELCO $b \to c \to \ell$ shape | 0.033 | -0.001 |
| DELCO $+1\sigma$ $b \to c \to \ell$ shape | 0.053 | 0.001 |
| DELCO $-1\sigma$ $b \to c \to \ell$ shape | -0.010 | -0.047 |
| Continuum Norm $\pm 10\%$ | 0.012 / -0.108 | 0.021 / -0.021 |
| $\mu$ Fakes, $E_\ell < 1.5$ GeV, $\pm 10\%$ | -0.094 / 0.099 | 0.000 / 0.000 |
| $\mu$ Fakes, $E_\ell > 1.5$ GeV, $\pm 10\%$ | 0.031 / -0.032 | -0.030 / 0.036 |
| $\mathcal{B}(b \to c \to \ell) \pm 10\%$ | -0.088 / 0.085 | -0.096 / 0.095 |
| $\mathcal{B}(b \to \text{baryons}) \pm 20\%$ | -0.052 / 0.050 | -0.026 / 0.025 |
| # $K_L^0$ | -0.091 / 0.091 | -0.108 / 0.107 |
| Track Efficiency | 0.110 / -0.120 | 0.096 / -0.108 |
| # Fake Tracks | -0.036 / 0.032 | 0.004 / -0.007 |
| Shower Efficiency | 0.046 / -0.054 | 0.032 / -0.041 |
| # Fake Showers | -0.244 / 0.205 | -0.044 / 0.020 |
| Force Trk Multiplicity | -0.011 | -0.036 |
| Force Shwr Multiplicity | 0.001 | -0.058 |
| Final State Radiation | 0.034 | -0.093 |
| Lepton Efficiency | 0.005 | -0.036 |

Table A.6: Model Dependence for $\langle (E_\ell - \langle E_\ell \rangle)^2 \rangle$.
Nominal value and shifts of the $\langle (E_\ell - \langle E_\ell \rangle)^2 \rangle$ moments due to the individual model variations. All entries are in units of $10^{-3} \times \text{GeV}^2$

| Variation | $E_\ell > 1.0$ GeV | $E_\ell > 1.5$ GeV |
|---|---|---|
| Nominal | 296.610 | 176.387 |
| $B \to Dl\overline{\nu}$ $\rho$ param | -0.697 / -0.006 | -0.326 / 0.028 |
| $B \to Dl\overline{\nu}$ $c_D$ param | -0.322 / 0.598 | 0.731 / 0.500 |
| $B \to D^*l\overline{\nu}$ $\rho$ param | -0.222 / 5.104 | -4.919 / 1.246 |
| $B \to D^*l\overline{\nu}$ $c_{A_1}$ param | 1.232 / 2.183 | -6.163 / -1.082 |
| $B \to D^*l\overline{\nu}$ R1/R2 1st eig-vec | 1.867 / 0.535 | -3.233 / -0.946 |
| $B \to D^*l\overline{\nu}$ R1/R2 2nd eig-vec | 0.479 / 0.620 | -5.409 / 0.336 |
| $B \to D^{**}l\overline{\nu}$ HQET model | 0.406 | -0.097 |
| $B \to D^{**}l\overline{\nu}$ $w$ slope | 0.520 / -0.037 | -0.138 / 0.381 |
| $B \to X_c l\overline{\nu}$ NonRes $w$ slope | 0.183 / 0.025 | 0.028 / 0.146 |
| $B \to X_c l\overline{\nu}$ Nonresonant Mass Dependence | | |
| Polynomial | 0.542 / -0.240 | -0.122 / -0.154 |
| Gauss $\mu = 2.0$ $\sigma^2 = .25$ | -1.377 | -0.353 |
| Gauss $\mu = 2.0$ $\sigma^2 = .75$ | -0.564 | -0.335 |
| Gauss $\mu = 2.0$ $\sigma^2 = 1.25$ | -0.205 | -0.394 |
| Gauss $\mu = 2.75$ $\sigma^2 = .25$ | -0.225 | 0.012 |
| Gauss $\mu = 2.75$ $\sigma^2 = .75$ | 0.372 | -0.449 |
| Gauss $\mu = 2.75$ $\sigma^2 = 1.25$ | 0.435 | -0.404 |
| Gauss $\mu = 3.5$ $\sigma^2 = .25$ | 1.180 | -0.097 |
| Gauss $\mu = 3.5$ $\sigma^2 = .75$ | 1.382 | -0.327 |
| Gauss $\mu = 3.5$ $\sigma^2 = 1.25$ | 1.233 | -0.407 |
| $B \to X_u l\overline{\nu}$ Model Dependence | | |
| $B \to X_s \gamma$ based | -0.007 | -0.010 |
| InclGen | -0.006 | 0.011 |
| All Non-Res | -0.004 | -0.031 |
| All Non-Res High Mass | 0.009 | 0.016 |
| All Non-Res Low Mass | -0.040 | -0.072 |
| Hard $q^2$ | 0.055 | 0.034 |
| Soft $q^2$ | -0.052 | -0.023 |
| High $\cos\theta_{W\ell}$ | 0.008 | 0.007 |
| Low $\cos\theta_{W\ell}$ | -0.009 | -0.007 |

Table A.7: Detector Systematics for $R0$.
Nominal value and shifts of the $R0$ moments due to the individual detector systematic uncertainties. All entries are in units of $10^{-3}$.

| Variation | $E_\ell > 1.0$ GeV | $E_\ell > 1.5$ GeV |
|---|---|---|
| Nominal | 339.842 | 608.008 |
| Lepton Fake Rate | -1.158 / 0.558 | -1.611 / 3.159 |
| DELCO $b \to c \to \ell$ shape | 0.765 | -0.137 |
| DELCO $+1\sigma$ $b \to c \to \ell$ shape | 1.093 | -0.151 |
| DELCO $-1\sigma$ $b \to c \to \ell$ shape | 0.003 | -0.443 |
| Continuum Norm $\pm10\%$ | 0.779 / -0.872 | -0.019 / 0.005 |
| $\mu$ Fakes, $E_\ell < 1.5$ GeV, $\pm10\%$ | 0.923 / -0.854 | 0.000 / 0.000 |
| $\mu$ Fakes, $E_\ell > 1.5$ GeV, $\pm10\%$ | -0.013 / 0.019 | -0.697 / 0.774 |
| $\mathcal{B}(b \to c \to \ell)$ $\pm10\%$ | -0.269 / 0.272 | -0.217 / 0.207 |
| $\mathcal{B}(b \to \text{baryons})$ $\pm20\%$ | 0.232 / -0.237 | 0.197 / -0.194 |
| # $K_L^0$ | -0.313 / 0.315 | -0.100 / 0.100 |
| Track Efficiency | -0.172 / 0.115 | -0.398 / 0.370 |
| # Fake Tracks | 0.678 / -0.610 | 0.485 / -0.446 |
| Shower Efficiency | 0.159 / -0.271 | -0.327 / 0.342 |
| # Fake Showers | 2.472 / -2.483 | 2.944 / -2.801 |
| Force Trk Multiplicity | -0.248 | -0.170 |
| Force Shwr Multiplicity | -0.560 | -0.754 |
| Final State Radiation | -2.228 | -1.105 |
| Lepton Efficiency | -0.884 | -0.788 |

Table A.8: Model Dependence for $R0$.

Nominal value and shifts of the $R0$ moments due to the individual model variations. All entries are in units of $10^{-3}$.

| Variation | $E_\ell > 1.0$ GeV | $E_\ell > 1.5$ GeV |
|---|---|---|
| Nominal | 339.842 | 608.008 |
| $B \to Dl\overline{\nu}$ $\rho$ param | -1.658 / 1.127 | -1.017 / 0.926 |
| $B \to Dl\overline{\nu}$ $c_D$ param | -0.193 / 0.092 | -0.212 / 0.340 |
| $B \to D^*l\overline{\nu}$ $\rho$ param | -6.227 / 5.436 | -5.446 / 4.626 |
| $B \to D^*l\overline{\nu}$ $c_{A_1}$ param | -1.201 / 0.538 | -1.609 / 0.902 |
| $B \to D^*l\overline{\nu}$ R1/R2 1st eig-vec | 6.373 / -8.316 | 4.359 / -4.935 |
| $B \to D^*l\overline{\nu}$ R1/R2 2nd eig-vec | 2.381 / -2.623 | 1.555 / -2.339 |
| $B \to D^{**}l\overline{\nu}$ HQET model | -0.372 | -0.631 |
| $B \to D^{**}l\overline{\nu}$ $w$ slope | 0.002 / 0.100 | -0.212 / 0.147 |
| $B \to X_c l\overline{\nu}$ NonRes $w$ slope | -0.326 / 0.159 | -0.363 / 0.246 |
| $B \to X_c l\overline{\nu}$ Nonresonant Mass Dependence | | |
| Polynomial | -0.742 / 0.476 | -0.332 / 0.271 |
| Gauss $\mu = 2.0$ $\sigma^2 = .25$ | 4.508 | 4.550 |
| Gauss $\mu = 2.0$ $\sigma^2 = .75$ | 1.042 | 1.774 |
| Gauss $\mu = 2.0$ $\sigma^2 = 1.25$ | -0.216 | 0.799 |
| Gauss $\mu = 2.75$ $\sigma^2 = .25$ | 0.337 | -0.191 |
| Gauss $\mu = 2.75$ $\sigma^2 = .75$ | -0.999 | -0.518 |
| Gauss $\mu = 2.75$ $\sigma^2 = 1.25$ | -2.101 | -1.333 |
| Gauss $\mu = 3.5$ $\sigma^2 = .25$ | -1.055 | -0.224 |
| Gauss $\mu = 3.5$ $\sigma^2 = .75$ | -3.731 | -3.104 |
| Gauss $\mu = 3.5$ $\sigma^2 = 1.25$ | -4.699 | -4.069 |
| $B \to X_u l\overline{\nu}$ Model Dependence | | |
| $B \to X_s \gamma$ based | 0.043 | -0.052 |
| InclGen | 0.093 | 0.009 |
| All Non-Res | -0.182 | -0.250 |
| All Non-Res High Mass | 0.633 | 0.336 |
| All Non-Res Low Mass | -0.363 | -0.507 |
| Hard $q^2$ | -0.173 | -0.061 |
| Soft $q^2$ | 0.233 | 0.125 |
| High $\cos\theta_{W\ell}$ | 0.029 | 0.018 |
| Low $\cos\theta_{W\ell}$ | -0.032 | -0.018 |

Table A.9: Detector Systematics for $\langle q^2 \rangle$.
Nominal value and shifts of the $\langle q^2 \rangle$ moments due to the individual detector systematic uncertainties. All entries are in units of GeV$^2$/$c^4$

| Variation | $E_\ell > 1.0$ GeV | $E_\ell > 1.5$ GeV |
|---|---|---|
| Nominal | 4.892 | 5.287 |
| Lepton Fake Rate | -0.011 / 0.006 | -0.019 / 0.033 |
| DELCO $b \to c \to \ell$ shape | 0.010 | -0.001 |
| DELCO $+1\sigma$ $b \to c \to \ell$ shape | 0.015 | -0.001 |
| DELCO $-1\sigma$ $b \to c \to \ell$ shape | -0.001 | -0.008 |
| Continuum Norm $\pm 10\%$ | 0.015 / -0.020 | 0.011 / -0.011 |
| $\mu$ Fakes, $E_\ell < 1.5$ GeV, $\pm 10\%$ | 0.005 / -0.003 | 0.000 / 0.000 |
| $\mu$ Fakes, $E_\ell > 1.5$ GeV, $\pm 10\%$ | 0.001 / -0.001 | -0.008 / 0.009 |
| $\mathcal{B}(b \to c \to \ell)$ $\pm 10\%$ | -0.006 / 0.006 | -0.007 / 0.007 |
| $\mathcal{B}(b \to \mathrm{baryons})$ $\pm 20\%$ | 0.003 / -0.003 | 0.004 / -0.004 |
| # $K_L^0$ | -0.008 / 0.008 | -0.008 / 0.009 |
| Track Efficiency | 0.003 / -0.004 | 0.002 / -0.002 |
| # Fake Tracks | 0.005 / -0.004 | 0.004 / -0.003 |
| Shower Efficiency | 0.006 / -0.007 | 0.001 / -0.001 |
| # Fake Showers | 0.013 / -0.014 | 0.018 / -0.017 |
| Force Trk Multiplicity | -0.004 | -0.005 |
| Force Shwr Multiplicity | -0.003 | -0.004 |
| Final State Radiation | -0.024 | -0.010 |
| Lepton Efficiency | -0.012 | -0.010 |

162

Table A.10: Model Dependence for $\langle q^2 \rangle$.
Nominal value and shifts of the $\langle q^2 \rangle$ moments due to the individual model variations. All entries are in units of GeV$^2/c^4$

| Variation | $E_\ell > 1.0$ GeV | $E_\ell > 1.5$ GeV |
|---|---|---|
| Nominal | 4.892 | 5.287 |
| $B \to Dl\overline{\nu}$ $\rho$ param | 0.022 / -0.010 | 0.029 / -0.014 |
| $B \to Dl\overline{\nu}$ $c_D$ param | 0.000 / 0.001 | 0.002 / -0.000 |
| $B \to D^*l\overline{\nu}$ $\rho$ param | 0.041 / -0.059 | 0.057 / -0.072 |
| $B \to D^*l\overline{\nu}$ $c_{A_1}$ param | -0.010 / 0.006 | -0.005 / 0.002 |
| $B \to D^*l\overline{\nu}$ R1/R2 1st eig-vec | -0.047 / 0.044 | -0.037 / 0.037 |
| $B \to D^*l\overline{\nu}$ R1/R2 2nd eig-vec | -0.006 / 0.005 | -0.004 / 0.004 |
| $B \to D^{**}l\overline{\nu}$ HQET model | -0.014 | -0.005 |
| $B \to D^{**}l\overline{\nu}$ $w$ slope | -0.013 / 0.015 | -0.012 / 0.012 |
| $B \to X_cl\overline{\nu}$ NonRes $w$ slope | -0.006 / 0.003 | -0.009 / 0.005 |
| $B \to X_cl\overline{\nu}$ Nonresonant Mass Dependence | | |
| Polynomial | -0.022 / 0.011 | -0.011 / 0.005 |
| Gauss $\mu = 2.0$ $\sigma^2 = .25$ | 0.029 | 0.008 |
| Gauss $\mu = 2.0$ $\sigma^2 = .75$ | 0.007 | -0.000 |
| Gauss $\mu = 2.0$ $\sigma^2 = 1.25$ | -0.006 | -0.006 |
| Gauss $\mu = 2.75$ $\sigma^2 = .25$ | 0.001 | -0.006 |
| Gauss $\mu = 2.75$ $\sigma^2 = .75$ | -0.017 | -0.012 |
| Gauss $\mu = 2.75$ $\sigma^2 = 1.25$ | -0.027 | -0.018 |
| Gauss $\mu = 3.5$ $\sigma^2 = .25$ | -0.044 | -0.025 |
| Gauss $\mu = 3.5$ $\sigma^2 = .75$ | -0.054 | -0.033 |
| Gauss $\mu = 3.5$ $\sigma^2 = 1.25$ | -0.058 | -0.037 |
| $B \to X_ul\overline{\nu}$ Model Dependence | | |
| $B \to X_s\gamma$ based | 0.001 | 0.001 |
| InclGen | 0.002 | 0.003 |
| All Non-Res | -0.001 | -0.002 |
| All Non-Res High Mass | 0.004 | -0.001 |
| All Non-Res Low Mass | -0.002 | -0.004 |
| Hard $q^2$ | -0.004 | -0.005 |
| Soft $q^2$ | 0.005 | 0.006 |
| High $\cos\theta_{W\ell}$ | 0.000 | 0.000 |
| Low $\cos\theta_{W\ell}$ | -0.000 | -0.000 |

Table A.11: Detector Systematics for $\langle (q^2 - \langle q^2 \rangle)^2 \rangle$.
Nominal value and shifts of the $\langle (q^2 - \langle q^2 \rangle)^2 \rangle$ moments due to the individual detector systematic uncertainties. All entries are in units of $10^{-3} \times \text{GeV}^4/c^8$

| Variation | $E_\ell > 1.0$ GeV | $E_\ell > 1.5$ GeV |
|---|---|---|
| Nominal | 2852.203 | 2879.355 |
| Lepton Fake Rate | -0.840 / 0.362 | 0.443 / -0.813 |
| DELCO $b \to c \to \ell$ shape | 0.421 | 0.032 |
| DELCO $+1\sigma$ $b \to c \to \ell$ shape | 0.574 | 0.006 |
| DELCO $-1\sigma$ $b \to c \to \ell$ shape | 0.036 | 0.522 |
| Continuum Norm $\pm 10\%$ | 0.149 / -0.198 | -1.549 / 1.581 |
| $\mu$ Fakes, $E_\ell < 1.5$ GeV, $\pm 10\%$ | 0.803 / -0.793 | 0.000 / 0.000 |
| $\mu$ Fakes, $E_\ell > 1.5$ GeV, $\pm 10\%$ | -0.044 / 0.049 | 0.303 / -0.380 |
| $\mathcal{B}(b \to c \to \ell) \pm 10\%$ | -0.068 / 0.048 | 0.465 / -0.514 |
| $\mathcal{B}(b \to \text{baryons}) \pm 20\%$ | 0.096 / -0.104 | -0.506 / 0.496 |
| # $K_L^0$ | -0.061 / 0.033 | 0.755 / -0.815 |
| Track Efficiency | -0.288 / 0.257 | -0.311 / 0.352 |
| # Fake Tracks | 0.548 / -0.506 | -0.015 / -0.001 |
| Shower Efficiency | -0.069 / -0.013 | -0.335 / 0.389 |
| # Fake Showers | 2.129 / -2.164 | 0.222 / -0.400 |
| Force Trk Multiplicity | -0.120 | 0.430 |
| Force Shwr Multiplicity | -0.534 | -0.363 |
| Final State Radiation | -1.567 | 0.019 |
| Lepton Efficiency | -0.499 | 0.525 |

Table A.12: Model Dependence for $\langle (q^2 - \langle q^2 \rangle)^2 \rangle$.
Nominal value and shifts of the $\langle (q^2 - \langle q^2 \rangle)^2 \rangle$ moments due to the individual model variations. All entries are in units of $10^{-3} \times \mathrm{GeV}^4/c^8$

| Variation | $E_\ell > 1.0$ GeV | $E_\ell > 1.5$ GeV |
|---|---|---|
| Nominal | 2852.203 | 2879.355 |
| $B \to Dl\overline{\nu}$ $\rho$ param | -7.112 / 5.057 | -7.769 / 5.075 |
| $B \to Dl\overline{\nu}$ $c_D$ param | 5.850 / -4.784 | 6.520 / -5.509 |
| $B \to D^*l\overline{\nu}$ $\rho$ param | 26.712 / -27.140 | 31.601 / -29.908 |
| $B \to D^*l\overline{\nu}$ $c_{A_1}$ param | 26.477 / -26.716 | 29.348 / -28.223 |
| $B \to D^*l\overline{\nu}$ R1/R2 1st eig-vec | -19.883 / 19.427 | -18.438 / 21.136 |
| $B \to D^*l\overline{\nu}$ R1/R2 2nd eig-vec | -1.072 / 0.859 | -0.992 / 1.583 |
| $B \to D^{**}l\overline{\nu}$ HQET model | 10.384 | 6.308 |
| $B \to D^{**}l\overline{\nu}$ $w$ slope | 12.122 / -11.871 | 8.247 / -7.658 |
| $B \to X_c l\overline{\nu}$ NonRes $w$ slope | 2.837 / -1.456 | 3.429 / -1.870 |
| $B \to X_c l\overline{\nu}$ Nonresonant Mass Dependence | | |
| Polynomial | 4.787 / -2.748 | 5.325 / -1.981 |
| Gauss $\mu = 2.0$ $\sigma^2 = .25$ | -4.284 | -1.719 |
| Gauss $\mu = 2.0$ $\sigma^2 = .75$ | -3.237 | -3.604 |
| Gauss $\mu = 2.0$ $\sigma^2 = 1.25$ | -1.232 | -3.332 |
| Gauss $\mu = 2.75$ $\sigma^2 = .25$ | -0.506 | 0.251 |
| Gauss $\mu = 2.75$ $\sigma^2 = .75$ | 1.706 | -1.168 |
| Gauss $\mu = 2.75$ $\sigma^2 = 1.25$ | 1.998 | -1.818 |
| Gauss $\mu = 3.5$ $\sigma^2 = .25$ | 11.531 | 9.983 |
| Gauss $\mu = 3.5$ $\sigma^2 = .75$ | 7.797 | 3.484 |
| Gauss $\mu = 3.5$ $\sigma^2 = 1.25$ | 5.928 | 0.658 |
| $B \to X_u l\overline{\nu}$ Model Dependence | | |
| $B \to X_s \gamma$ based | -0.022 | -0.177 |
| InclGen | -0.002 | -0.429 |
| All Non-Res | -0.161 | -0.080 |
| All Non-Res High Mass | 0.580 | 0.531 |
| All Non-Res Low Mass | -0.339 | -0.225 |
| Hard $q^2$ | 0.024 | 0.838 |
| Soft $q^2$ | 0.003 | -0.887 |
| High $\cos\theta_{W\ell}$ | 0.022 | -0.011 |
| Low $\cos\theta_{W\ell}$ | -0.025 | 0.013 |

# Bibliography

[1] K. Hagiwara *et al.* [Particle Data Group Collaboration], "Review of particle physics," Phys. Rev. D **66**, 010001 (2002).

[2] E. Bagan, P. Ball, V. M. Braun and P. Gosdzinsky, "Theoretical update of the semileptonic branching ratio of $B$ mesons," Phys. Lett. B **342**, 362 (1995) [Erratum-ibid. B **374**, 363 (1996)] [arXiv:hep-ph/9409440].

[3] B. Aubert *et al.* [BABAR Collaboration], "Measurement of the branching fraction for inclusive semileptonic $B$ meson decays," Phys. Rev. D **67**, 031101 (2003) [arXiv:hep-ex/0208018].

[4] B. Barish *et al.* [CLEO Collaboration], "Measurements of the $B$ semileptonic branching fraction with lepton tags," Phys. Rev. Lett. **76**, 1570 (1996).

[5] K. Abe *et al.* [Belle Collaboration], "Measurement of the inclusive semileptonic branching fraction of $B$ mesons and $|V_{cb}|$," Phys. Lett. B **547**, 181 (2002) [arXiv:hep-ex/0208033].

[6] C. Caso *et al.* [Particle Data Group Collaboration], "Review of particle physics," Eur. Phys. J. C **3**, 1 (1998).

[7] A. V. Manohar and M. B. Wise, *Heavy Quark Physics*, Cambridge Monogr. Part. Phys. Nucl. Phys. Cosmol. **10**, 1 (2000), provides a review and references to the orignal papers.

[8] S. B. Athar *et al.* [CLEO Collaboration], "Study of the $q^2$ dependence of $B \to \pi l \overline{\nu}$ and $B \to \rho(\omega) l \nu$ decay and extraction of $|V_{ub}|$," arXiv:hep-ex/0304019.

[9] A. Bornheim *et al.* [CLEO Collaboration], "Improved measurement of $|V_{ub}|$ with inclusive semileptonic $B$ decays," Phys. Rev. Lett. **88**, 231803 (2002) [arXiv:hep-ex/0202019].

[10] R. Fulton *et al.* [CLEO Collaboration], "Observation of $B$ meson semileptonic decays to non-charmed final states," Phys. Rev. Lett. **64**, 16 (1990).

[11] A. Stahl, *Physics with Tau Leptons*, Spinger, 2000.

[12] M. E. Peskin and D. V. Schroder, *An Introduction to Quantum Field Theory*, Addison-Wesley, 1995.

[13] S. Chen *et al.* [CLEO Collaboration], "Branching fraction and photon energy spectrum for $b \to s\gamma$," Phys. Rev. Lett. **87**, 251807 (2001) [arXiv:hep-ex/0108032].

[14] J. Chay, H. Georgi and B. Grinstein, "Lepton energy distributions in heavy meson decays from QCD," Phys. Lett. B **247**, 399 (1990); I. I. Bigi, M. A. Shifman, N. G. Uraltsev and A. I. Vainshtein, "QCD predictions for lepton spectra in inclusive heavy flavor decays," Phys. Rev. Lett. **71**, 496 (1993) [arXiv:hep-ph/9304225]; A. V. Manohar and M. B. Wise, "Inclusive semileptonic $B$ and polarized $\Lambda_b$ decays from QCD," Phys. Rev. D **49**, 1310 (1994).

[15] R. A. Briere *et al.* [CLEO Collaboration], "Measurement of the lepton energy in the decay $B \to Xl\overline{\nu}$ and determination of the heavy quark expansion parameters," [arXiv:hep-ex/0209024].

[16] D. Cronin-Hennessy *et al.* [CLEO Collaboration], "Hadronic mass moments in inclusive semileptonic $B$ meson decays," Phys. Rev. Lett. **87**, 251808 (2001) [arXiv:hep-ex/0108033].

[17] B. Aubert, *et al.* [BaBar Collaboration], "Measurement of the first hadronic spectral moment from semileptonic $B$ decays," [arXiv:hep-ex/0307046].

[18] M. Gremm, A. Kapustin, Z. Ligeti and M. B. Wise, "Implications of the $B \to X\ell\bar{\nu}_\ell$ lepton spectrum for heavy quark theory," Phys. Rev. Lett. **77**, 20 (1996) [arXiv:hep-ph/9603314]; M. Gremm and I. Stewart, "Order $\alpha_s^2\beta_0$ correction to the charged lepton spectrum in $b \to cl\bar{\nu}_l$ decays," Phys. Rev. D **55**, 1226 (1997) [arXiv:hep-ph/9609341]; M. Gremm and A. Kapustin, "Order $1/m_b^3$ corrections to inclusive semileptonic $B$ decay," Phys. Rev. D **55**, 6924 (1997) [arXiv:hep-ph/9603448].

[19] A. F. Falk and M. E. Luke, "Hadronic spectral moments in semileptonic $B$ decays with a lepton energy cut," Phys. Rev. D **57**, 424 (1998) [arXiv:hep-ph/9708327].

[20] Z. Ligeti, M. E. Luke, A. V. Manohar and M. B. Wise, "The $B \to X_s\gamma$ photon spectrum," Phys. Rev. D **60**, 034019 (1999) [arXiv:hep-ph/9903305]; C. W. Bauer, "Corrections to moments of the photon spectrum in the inclusive decay $B \to X_s\gamma$," Phys. Rev. D **57**, 5611 (1998) [Erratum-ibid. D **60**, 099907 (1999)] [arXiv:hep-ph/9710513].

[21] C. W. Bauer and M. Trott, "Reducing theoretical uncertainties in $m_b$ and $\lambda_1$," Phys. Rev. D **67**, 014021 (2003) [arXiv:hep-ph/0205039].

[22] C. W. Bauer, Z. Ligeti, M. Luke and A. V. Manohar, "$B$ decay shape variables and the precision determination of $|V_{cb}|$ and $m_b$," Phys. Rev. D **67**, 054012 (2003) [arXiv:hep-ph/0210027];

[23] M. Neubert, "Analysis of the photon spectrum in inclusive $B \to X_s\gamma$ decays," Phys. Rev. D **49**, 4623 (1994) [arXiv:hep-ph/9312311]; I. I. Bigi, M. A. Shifman, N. G. Uraltsev and A. I. Vainshtein, "On the motion of heavy quarks inside hadrons: Universal distributions and inclusive decays," Int. J. Mod. Phys. A **9**, 2467 (1994) [arXiv:hep-ph/9312359]; I. I. Bigi, M. A. Shifman, N. Uraltsev and A. I. Vainshtein, "Heavy quark distribution function in QCD

and the Ac$^2$m$^2$ model," Phys. Lett. B **328**, 431 (1994) [arXiv:hep-ph/9402225]; A. L. Kagan and M. Neubert, "QCD anatomy of $B \to X_s \gamma$ decays," Eur. Phys. J. C **7**, 5 (1999) [arXiv:hep-ph/9805303].

[24] A. F. Falk, Z. Ligeti and M. B. Wise, "$|V_{ub}|$ from the hadronic invariant mass spectrum in semileptonic $B$ decay," Phys. Lett. B **406**, 225 (1997) [arXiv:hep-ph/9705235].

[25] C. W. Bauer, Z. Ligeti and M. E. Luke, "A model independent determination of $|V_{ub}|$," Phys. Lett. B **479**, 395 (2000) [arXiv:hep-ph/0002161]; M. Neubert and T. Becher, "Improved determination of $|V_{ub}|$ from inclusive semileptonic $B$ meson decays," Phys. Lett. B **535**, 127 (2002) [arXiv:hep-ph/0105217].

[26] C. W. Bauer, Z. Ligeti and M. E. Luke, "Precision determination of $|V_{ub}|$ from inclusive decays," Phys. Rev. D **64**, 113004 (2001).

[27] G. Bonvicini *et al.* [CLEO Collaboration], "Study of semileptonic decays of $B$ mesons to charmed baryons," Phys. Rev. D **57**, 6604 (1998) [arXiv:hep-ex/9712008].

[28] N. Isgur and M. B. Wise, "Weak decays of heavy mesons in the static quark approximation," Phys. Lett. B **232**, 113 (1989); N. Isgur and M. B. Wise, "Weak transition form-factors between heavy mesons," Phys. Lett. B **237**, 527 (1990); N. Isgur and M. B. Wise, "Spectroscopy with heavy quark symmetry," Phys. Rev. Lett. **66**, 1130 (1991).

[29] C. G. Boyd, B. Grinstein and R. F. Lebed, "Model independent extraction of $|V_{cb}|$ using dispersion relations," Phys. Lett. B **353**, 306 (1995) [arXiv:hep-ph/9504235]; I. Caprini and M. Neubert, "Improved bounds for the slope and curvature of $\bar{B} \to D^{(*)} \ell \bar{\nu}$ form factors," Phys. Lett. B **380**, 376 (1996) [arXiv:hep-ph/9603414]; C. G. Boyd, B. Grinstein and R. F. Lebed, "Precision corrections to dispersive bounds on form factors," Phys. Rev. D **56**, 6895 (1997) [arXiv:hep-ph/9705252]; I. Caprini, L. Lellouch and M. Neubert, "Dispersive bounds on the shape of $\bar{B} \to D^{(*)} \ell \bar{\nu}$ form factors," Nucl. Phys. B **530**, 153 (1998) [arXiv:hep-ph/9712417].

[30] A. Anastassov *et al.* [CLEO Collaboration], "Investigation of semileptonic $B$ meson decay to P-wave charm mesons," Phys. Rev. Lett. **80**, 4127 (1998) [arXiv:hep-ex/9708035].

[31] A. K. Leibovich, Z. Ligeti, I. W. Stewart and M. B. Wise, "Semileptonic $B$ decays to excited charmed mesons," Phys. Rev. D **57**, 308 (1998) [arXiv:hep-ph/9705467].

[32] N. Isgur, "Nonresonant semileptonic heavy quark decay," Phys. Rev. D **60**, 074030 (1999) [arXiv:hep-ph/9904398].

[33] J. L. Goity and W. Roberts, "Soft pion emission in semileptonic $B$ meson decays," Phys. Rev. D **51**, 3459 (1995) [arXiv:hep-ph/9406236].

[34] S. W. Herb *et al.*, "Observation of a dimuon resonance at 9.5 GeV in 400 GeV proton - nucleus collisions," Phys. Rev. Lett. **39**, 252 (1977).

[35] Y. Kubota *et al.*, "The CLEO-II detector," Nucl. Instrum. Meth. A **320**, 66 (1992).

[36] T. S. Hill, "The CLEO-II silicon vertex detector," Nucl. Instrum. Meth. A **418**, 32 (1998).

[37] R. E. Kalman, J. Basic Eng. **82**, 34 (1961); R. E. Kalman and R. S. Bucy, J. Basic Eng. **83**, 95 (1961); P. Billoir, "Track fitting with multiple scattering: a new method ," Nucl. Instrum. Meth. A **225**, 352 (1984); P. Billoir, R. Frühwirth and M. Regler, "Track element merging strategy and vertex fitting in complex modular detectors," Nucl. Instrum. Meth. A **241**, 115 (1985); R. Frühwirth, Nucl. Instrum. Meth. A **262**, 444 (1987).

[38] T. Sjostrand, "The Lund Monte Carlo for jet fragmentation and $e^+e^-$ physics: jetset version 6.2," Comput. Phys. Commun. **39**, 347 (1986); T. Sjostrand and M. Bengtsson, "The Lund Monte Carlo for jet fragmentation and $e^+e^-$ physics: jetset version 6.3: an update," Comput. Phys. Commun. **43**, 367 (1987).

[39] R. Brun *et al.*, "GEANT 3.15: Detector description and simulation tool", CERN Report No. DD/EE/84-1 (1987).

[40] G. C. Fox and S. Wolfram, "Observables for the analysis of event shapes in $e^+e^-$ annihilation and other processes," Phys. Rev. Lett. **41**, 1581 (1978).

[41] C. Peterson, T. Rognvaldsson and L. Lonnblad, "JETNET 3.0: A versatile artificial neural network package," Comput. Phys. Commun. **81**, 185 (1994); (see also http://www.thep.lu.se/public_html/jetnet_30_manual/jetnet_30_manual.html).

[42] R. J. Barlow and C. Beeston, "Fitting using finite Monte Carlo samples," Comput. Phys. Commun. **77**, 219 (1993).

[43] F. James, "MINUIT 94.1: Function minimization and error analysis", CERN Program Library Long Writeup D506, (1994).

[44] E. Barberio and Z. Wąs, "PHOTOS: A universal Monte Carlo for QED radiative corrections. version 2.0," Comput. Phys. Commun. **79**, 291 (1994).

[45] D. Atwood and W. J. Marciano, "Radiative corrections and semileptonic $B$ decays," Phys. Rev. D **41**, 1736 (1990).

[46] E. Richter-Wąs, "QED bremsstrahlung in semileptonic $B$ and leptonic $\tau$ decays," Phys. Lett. B **303**, 163 (1993).

[47] T. Kinoshita and A. Sirlin, "Radiative corrections to fermi interactions," Phys. Rev. **113**, 1652 (1959).

[48] E. S. Ginsberg, "Radiative corrections to $K_{e3}^0$ decays and the Delta-I=1/2 rule. (erratum)," Phys. Rev. **171**, 1675 (1968) [Erratum-ibid. **174**, 2169 (1968)]; E. S. Ginsberg, "Radiative corrections to the $K_{l3}^{\pm}$ dalitz plot. (erratum)," Phys. Rev. **162**, 1570 (1967) [Erratum-ibid. **187**, 2280 (1969)]; E. S. Ginsberg, "Radiative corrections to $K_{\mu3}$ decays," Phys. Rev. D **1**, 229 (1970); E. S. Ginsberg, "Radiative corrections to the muon polarization in $K_{\mu3}^{\pm}$ decays," Phys. Rev. D **4**, 2893 (1971); E. S. Ginsberg and J. Smith, "Radiative corrections to the muon polarization in $K_{\mu3}^0$ decays," Phys. Rev. D **8**, 3887 (1973).

[49] B. Barish *et al.* [CLEO Collaboration], "Measurement of the $B \to D^* l\overline{\nu}$ branching fractions and $|V_{cb}|$," Phys. Rev. D **51**, 1014 (1995) [arXiv:hep-ex/9406005].

[50] R. A. Briere *et al.* [CLEO Collaboration], "Improved measurement of $|V_{cb}|$ using $B \to D^* l\overline{\nu}$ decays," Phys. Rev. Lett. **89**, 081803 (2002) [arXiv:hep-ex/0203032]; N. E. Adam *et al.* [CLEO collaboration], "Determination of the $B \to D^* l\overline{\nu}$ decay width and $|V_{cb}|$," Phys. Rev. D **67**, 032001 (2003) [arXiv:hep-ex/0210040].

[51] D. Scora and N. Isgur, "Semileptonic meson decays in the quark model: An update," Phys. Rev. D **52**, 2783 (1995) [arXiv:hep-ph/9503486].

[52] N. Isgur, D. Scora, B. Grinstein and M. B. Wise, "Semileptonic $B$ And $D$ Decays In The Quark Model," Phys. Rev. D **39**, 799 (1989).

[53] R. Wang, R. Poling, D. Perticone, "Inclusive measurement of $B$ meson semileptonic branching fractions," **CBX 93-71**.

[54] B. Barish *et al.* [CLEO Collaboration], "Measurements of the $B$ semileptonic branching fraction with lepton tags," Phys. Rev. Lett. **76**, 1570 (1996).

[55] Private communications with R. Poling and C. Stepaniak.

[56] Y. Kubota *et al.* [CLEO Collaboration], "Measurements of the inclusive semielectronic $D^0$ branching fraction," Phys. Rev. D **54**, 2994 (1996) [arXiv:hep-ex/9511014].

[57] D. Bortoletto *et al.* [CLEO Collaboration], Phys. Rev. D **37**, 1719 (1988) [Erratum-ibid. D **39**, 1471 (1989)].

[58] R. Ammar *et al.* [CLEO Collaboration], "Measurement of the total cross section for $e^+e^-$ →hadrons at $s^{1/2} = 10.52$ GeV," Phys. Rev. D **57**, 1350 (1998) [arXiv:hep-ex/9707018]; D. Besson and S. Kotov, "Updated note on R (update of cbx95-70)," **CBX 97-6**; Dave Besson and Ina Robertin, "Preliminary note on R," **CBX 95-70**.

[59] Private communications with A. Warburton.

[60] A. Weinstein, "Study of absolute tracking efficiency using 1 vs 3 pi events," **CBX 93-117**; M. Lohner, J. Smith, "Absolute tracking efficiency revisited," **CBX 95-31**.

[61] K. Bloom, R. Patterson and P. Drell, "Absolute track-finding efficiency measurements with embedded tracks," **CBX 96-101**; B. Berger, "Tracking efficiency studies II: Recompress results," **CBX 00-32**.

[62] B. Valant-Spaight, B. Berger, K. Ecklund, R. Patterson, "Slow $\pi^0$ efficiency in recompress data," **CBX 01-16**.

[63] J. Urheim, "Lepton-tagged analysis of $\tau^- \to h^- \pi^0 \nu_\tau$ decays," **CBX 93-127**.

[64] W. Bacino *et al.*, "The semileptonic decays of the d meson," Phys. Rev. Lett. **43**, 1073 (1979).

[65] F. De Fazio and M. Neubert, "$B \to X_u l \overline{\nu}$ decay distributions to order $\alpha_s$," JHEP **9906**, 017 (1999) [arXiv:hep-ph/9905351].

[66] A. H. Hoang, Z. Ligeti and A. V. Manohar, "$B$ decays in the upsilon expansion," Phys. Rev. D **59**, 074017 (1999) [arXiv:hep-ph/9811239]; N. Uraltsev, "Theoretical uncertainties in Gamma(sl)(b $\to$ u)," Int. J. Mod. Phys. A **14**, 4641 (1999) [arXiv:hep-ph/9905520].

[67] C. W. Bauer, M. E. Luke and T. Mannel, "Light-cone distribution functions for $B$ decays at subleading order in $1/m_b$," arXiv:hep-ph/0102089.

[68] A. K. Leibovich, Z. Ligeti and M. B. Wise, "Enhanced subleading structure functions in semileptonic $B$ decay," Phys. Lett. B **539**, 242 (2002) [arXiv:hep-ph/0205148].