UNIVERSITY OF OKLAHOMA
GRADUATE COLLEGE

Semileptonic Decay of B Meson to Vector D Meson

A DISSERTATION
SUBMITTED TO THE GRADUATE FACULTY
in partial fulfillment of the requirements for the
degree of
Doctor of Philosophy

BY
Xu Fu
Norman, Oklahoma
1995

Semileptonic Decay of B Meson to Vector D Meson

A DISSERTATION APPROVED FOR
THE DEPARTMENT OF PHYSICS AND ASTRONOMY

BY

Patrick Skubic

George Kalbfleisch

Ronald Kantowski

John Cowan

Paul Goodey

To my parents

Acknowledgements

This thesis work would not have been possible without the entire CLEO collaboration and the dedicated CESR staff.

I would like to thank my adviser, Prof. Patrick Skubic, whose guidance and support have been of great help in my graduate career. I also want to thank the other members of my committee, Prof. George Kalbfleisch, Prof. Ronald Kantowski, Prof. John Cowan and Prof. Paul Goodey, who provided encouragement and direction in the course of my specialist exam and the preparation of this thesis.

I am grateful the $b \to c\ \ell\ \nu$ PTA group members, Prof. Persis Drell, Prof Bernie Gittleman, Maurice Sciveres, Chris O'Grady, Anders Ryd and others, their help to this work is invaluable. I am privileged to have Arne Freyberger supervising my analysis work. His insight, experience and consideration are central to my research in CLEO.

I am going to miss the "wings party". I had great time there with my officemates, Mark Lambrecht, Mark Wood and Bill Ross.

Most of all, I would especially like to thank my parents, without whom I would not have been possible.

Table of Contents

List of Tables

List of Figures

Abstract

Using data collected with the CLEO II detector at the Cornell Electron-positron Storage Ring (CESR) the exclusive semileptonic B meson decays $B^- \to D^{*0}\ \ell^-\ \bar{\nu}$ and $\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$ are studied. We present measurements of the branching fractions:

$$\begin{aligned}\mathcal{B}(\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}) &= (0.5/f_{00})[4.66 \pm 0.31(\text{stat.}) \pm 0.41(\text{sys.})]\% \\ \mathcal{B}(B^- \to D^{*0}\ \ell^-\ \bar{\nu}) &= (0.5/f_{+-})[5.20 \pm 0.50(\text{stat.}) \pm 0.66(\text{sys.})]\%,\end{aligned}$$

where f_{00} and f_{+-} are the neutral and charged B meson production fractions at the $\Upsilon(4S)$ resonance, respectively. By assuming $f_{+-} + f_{00} = 1$ we also determine the partial width:

$$\Gamma(\bar{B} \to D^*\ \ell\ \bar{\nu}) = [30.7 \pm 1.8(\text{stat.}) \pm 2.8(\text{sys.}) \pm 2.1(\text{lifetime})]\ ns^{-1}.$$

The product of the CKM matrix element $|V_{cb}|$ times the normalization of the decay form factor at the point of zero-recoil of the D^* meson, $\mathcal{F}(y = 1)$, is determined from a linear fit to the combined differential decay rate of the

exclusive $\bar{B} \to D^* \ell \bar{\nu}$ decays:

$$|V_{cb}|\mathcal{F}(1) = 0.0351 \pm 0.0019(\text{stat.}) \pm 0.0018(\text{sys.}) \pm 0.0008(\text{lifetime}).$$

The value for $|V_{cb}|$ is extracted using theoretical calculations of the form factor normalization. The bulk of this work has been published in 01 February 1995 issue of *Physical Review D*.

CHAPTER 1

Introduction

Particle physics concerns itself with the understanding of basic constituents and interactions of matter. High energy and small distance scales accessible only in the laboratory make this possible. Since the last half century more and more particles not observed in ordinary matter have been discovered. Our understanding of the interactions among these particles has advanced.

In this chapter, the motivation of this thesis work and an overview of the CLEO experiment are given. In chapter 2, we will discuss theoretical aspects of exclusive semileptonic B decays. In chapter 3 we will discuss in detail the measurement of exclusive semileptonic B meson branching fractions. In chapter 4 we will discuss in detail the extraction of the CKM matrix element $|V_{cb}|$.

Table 1
Summary of lepton properties.

Lepton	Name	Charge (e)	Mass (MeV)	Lifetime (sec)
e	Electron	-1	0.511	stable
ν_e	Neutrino	0	15×10^{-6}	stable
μ	Muon	-1	105.658	2.197×10^{-6}
ν_μ	Neutrino	0	< 0.27	stable
τ	Tau	-1	1,777	2.95×10^{-13}
ν_τ	Neutrino	0	< 31	stable

1.1 The Standard Model

The current set of fundamental particles and the quantized interactions is contained in a body of knowledge called the Standard Model.[1] It is the test of the validity of this model that comprises most particle physics research today. So far the Standard Model agrees with experiment well.

The standard model includes two types of fundamental particles: spin-1/2 fermions and spin-1 gauge bosons. In simple terms, matter is composed of fermions, and fermions interact with each other via the gauge bosons.

The fermions are divided into the three generations of leptons and three generations of quarks as shown in Tables 1 and 2.

The quarks are subject to electromagnetic, weak and strong interactions. The neutrinos interact through the weak interaction. The e, μ and τ are involved in both electromagnetic and weak interactions. The Gravitational interaction is not successfully accounted for in the Standard Model, and its effects are usually

Table 2
Summary of quark properties.

Quark	Name	Charge (e)	Mass (GeV)
d	down	$-1/3$	0.15
u	up	$+2/3$	0.15
s	strange	$-1/3$	0.50
c	charm	$+2/3$	1.50
b	beauty	$-1/3$	5.00
t	top	$+2/3$	174

Table 3
Properties of the four mediators of the electro-weak and strong interactions.

Mediator	Mass (GeV/c^2)	Spin	Interaction
Photon (γ)	0	1	Electromagnetic
$W^{\pm}$	80.22 ± 0.26	1	Weak
Z^0	91.187 ± 0.007	1	Weak
8 Gluons (g)	0	1	Strong

so small in particle physics that it can be ignored. At low energy and within the appropriate distance regime, the relative strengths of the strong, electromagnetic, weak, and gravitational forces are approximately $1 : 10^{-2} : 10^{-5} : 10^{-40}$, respectively.

The field quanta in the standard model are the gluon (g) for the strong interaction, the two intermediate vector bosons ($W^{\pm}$ and Z^{0}) for the weak interaction, and the massless photon (γ) for the electromagnetic interaction. The photon and the massive intermediate vector bosons have been directly observed, while the existence of the gluon has been confirmed indirectly.

Within the framework of the standard model, the weak and electromagnetic forces are unified into the electroweak force which have the same strength at higher energy (10^4 Gev). One of the primary goals of high energy physics is to further combine the fundamental forces, but this has not been fully accomplished.

The quark eigenstates of the weak interaction are linear combinations of the strong mass eigenstates, which can be described by the Cabibbo-Kobayashi-Maskawa[2] matrix:

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

The CKM matrix must be unitary in order to conserve probability. The unitary nature limits the numbers of the free parameters to only four, three real angles and one complex phase. The complex phase gives rise to CP violation.

One outstanding problem of the standard model is the large number of free parameters. The standard model includes at least eighteen free parameters: six quark and three lepton masses, four CKM parameters, the electron charge, the Higgs mass, the $W^{\pm}$ mass, the mixing angle θ_W, and the strong force coupling constant Λ_{QCD}. They are all determined by experiment.

CKM matrix element $|V_{cb}|$ is related to the coupling between b quark and c quark outside the b generation. This thesis work is an attempt to measure the decay rates of beauty flavored B meson decaying to charm flavored D^* mesons with the emission of leptons and to extract the CKM matrix element $|V_{cb}|$.

1.2 The Experiment

Two things have to be done before the physics analysis. First, large numbers of B mesons must be produced; second, the B meson decays must be recorded. The first task is achieved at the Cornell Electron Storage Ring (CESR) where an electron beam and a positron beam collide with each other at center-of mass energies around 10 GeV. The second task is achieved by the CLEO II detector where the momenta or the energies of B decay products are recorded, and the masses of the decay products are identified.

1.2.1 CESR

Cornell Electron Storage Ring[4] (CESR) was built in 1977-1979. Electrons are emitted from a activated metal cathode plate and accelerated to 150 MeV in a linear accelerator (linac). Positrons are produced in pair production ($e^- + nucleus \rightarrow e^- + nucleus + e^+ \; e^-$) by directing the 150 MeV electrons on a tungsten target. The electrons and positrons are accelerated to about 5GeV in the synchrotron before being deposited into the storage ring. Figure 1 shows the layout of CESR.

1.2.2 CLEO II

The CLEO II detector[5] is a solenoidal detector for detecting both charged and neutral particles with excellent resolution and efficiencies and has been in operation since 1989. The detector, as shown in Figure 2 and 3, consists of a charged particle tracking system surrounded by a time-of-flight scintillation system and an electromagnetic calorimeter, all immersed in a 1.5 Tesla field generated by a superconducting solenoid. The steel return yoke of the magnet is partially instrumented and thus is part of the muon identification system which is the outermost detector element. The detector components most relevant to this analysis are described in the following sections.

The Measurement of Charged Particle Momentum

The momentum of a charged particle is determined by measuring the curvature of its trajectory in a magnetic field.

Table 4
Achieved resolutions and material contributions in elements of the CLEO-II detector.

Detector Element	Resolution	Material in Front (% R.L.)
PTL	100 μm ($r-\phi$ plane) no longitudinal measurements	0.46
VD	150 μm ($r-\phi$ plane) 750 μm (z-direction)	1.38
DR	110 μm ($r-\phi$ plane - axial) 130 μm ($r-\phi$ plane - stereo)	2.51
All Drift Chambers	$(\delta p_t/p_t)^2 = (.0011 p_t)^2 + (.0067)^2$ $\sigma_\phi = 1 mrad$ $\sigma_\theta = 4 mrad$	
Ionization Loss	6.2% (5 GeV Bhabha electrons) 7.0% (minimum ionizing pions)	
Time-of-Flight	139 ps (5 GeV Bhabha electrons) 154 ps (minimum ionizing pions)	2.93
Barrel CsI Calorimeter	$\delta E/E[\%] = 0.35E^{-0.75} + 1.9 - 0.1E$ $\sigma_\phi = 2.8E^{-0.5} + 1.9$ $\sigma_\theta = 0.8\sigma_\phi \sin(\theta)$	18.0
Endcap CsI Calorimeter	$\delta E/E[\%] = 0.26E^{-1} + 2.5$ $\sigma_\phi = 3.7E^{-0.5} + 7.3$ $\sigma_\theta = 1.4E^{-0.5} + 5.6$	
Magnet	–	
Muon Chambers	2.4 cm (counters) 2.8 - 5 cm (strips)	Covers $\sim$ 7 nuclear absorption lengths

The tracking system is composed of three concentric drift chambers filled with 50:50 argon-ethane gas and it measures the trajectories of charged particles. The general principle is to measure the spatial distribution of the ionization left behind after a charged particle passes through the gas. For this purpose, the drift volume is divided up into "drift cells" consisting of cathode field wires and anode sense wires. At CESR energies, particles with heavy quarks (c or b) decay well within the beam pipe so that the tracking system primarily records their lighter long-lived decay products, such as pions, kaons, electrons, muons, and protons.

The measurement of Electromagnetic Shower Energy

High energy electrons and positrons undergo bremsstrahlung process when entering some materials. The emitted photons in turn carry a large fraction of the electron's or positron's initial energy and induce pair production in the materials. The total-absorption of these showers by the CLEO II CsI crystal calorimeter (CC) allows a measurement the energies and the spatial coordinates of the incident particles. The hadronic shower deposits less of its energy in the crystal and its shower shape is more spread out than the electromagnetic showers. These differences are utilized in electron identification.

Trigger and Data Acquisition

The CLEO II detector records the responses of a large number of electronic elements for further analysis. During the readout time, the experiment is unable to record further events. As a result, extreme selectiveness is required in the choice of the events recorded. A trigger system composed of fast digital electronics was established to process the primitive detector information available soon after beam crossing and to decide if precious live-time should be sacrificed to record an event. The primary consideration of the trigger design was the trade-off between recording as many diverse physical processes as possible, and reducing the dead-time. At the current CESR running condition of 7 bunches of electrons and positions per beam, the crossing rate is 2.7 MHz. The trigger system reduces the event recording rate to only about 25 Hz.

A VAX-3200 online computer with a data logging program provides the interface between the experimenter and the data control and acquisition system. The X-BUS, an eight-bit databus, handles the flow of control information to the detector elements, such as high voltage commands, trigger requirements and electronic calibration pulses, while the Y-BUS, a fast 16-bit databus transfers the data signals from the detector crate controllers to the online computer disk storage.

Data is stored on 4 mm tapes. In the off-line processing all events are skimmed for hadronic type decays, tau candidates, penguin type b-quark decays, *etc.* Events meeting criteria for any of these candidates are stripped from the raw data tape and converted to a quick access format denoted as *ROAR* data. This entire process is called PASS2*. Once the skimming procedure is completed, the results are again put onto disks accessed by the physicists carrying out data analysis.

*PASS1, which is an integral part of the data acquisition process, selects calibration type events such as Bhabha's and mu-pairs.

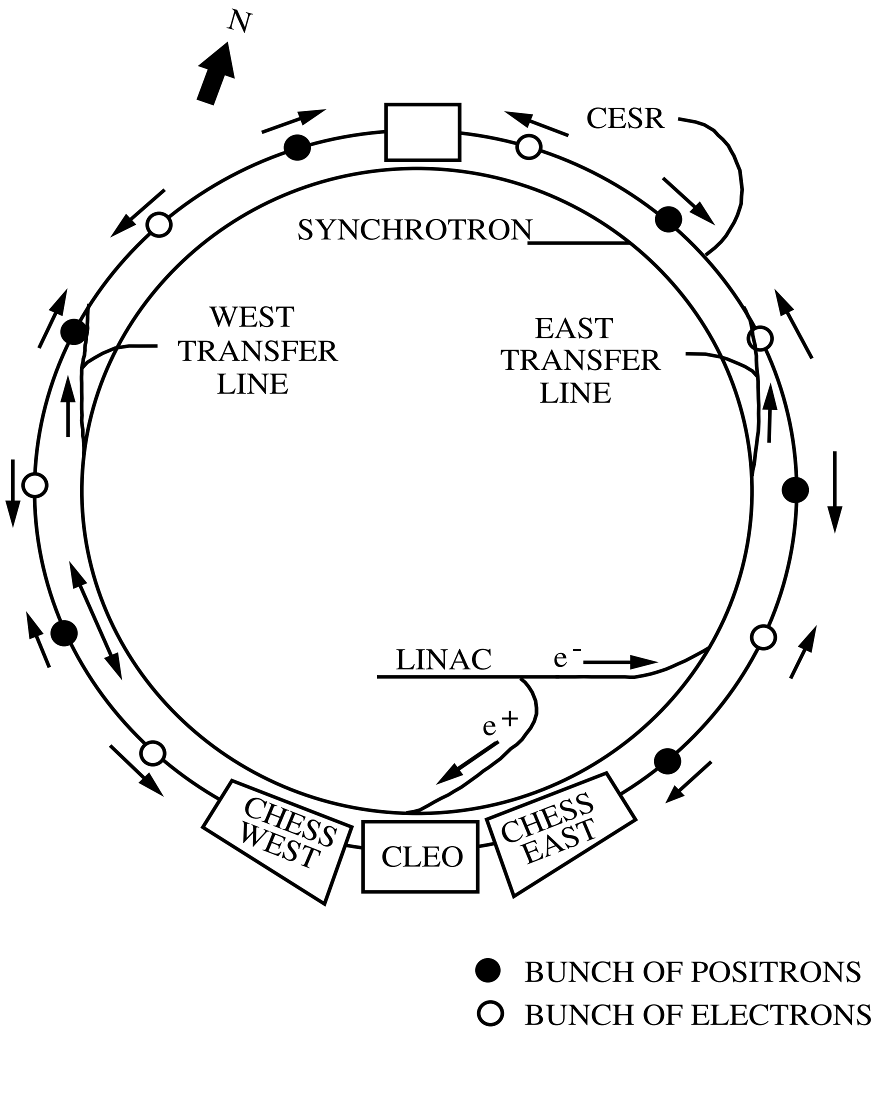

Figure 1
Schematic of the sychrotron and storage ring.

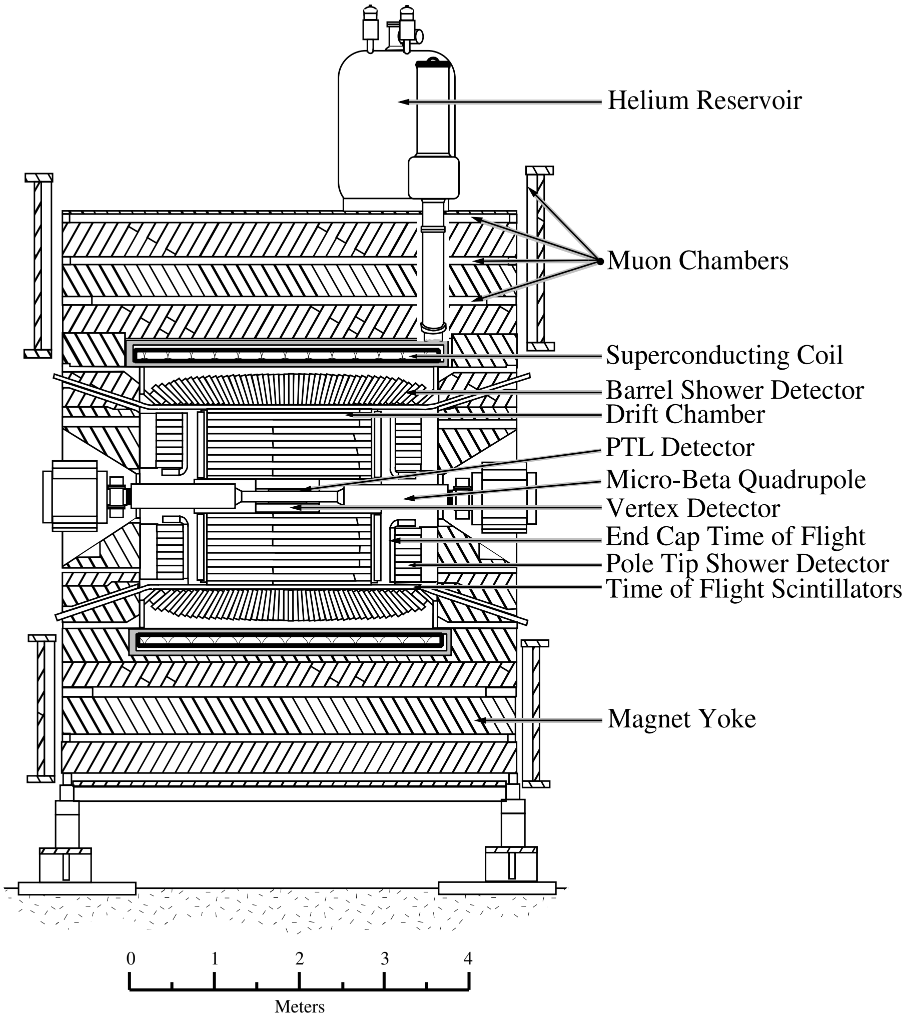

Figure 2
Side-view of the CLEO detector.

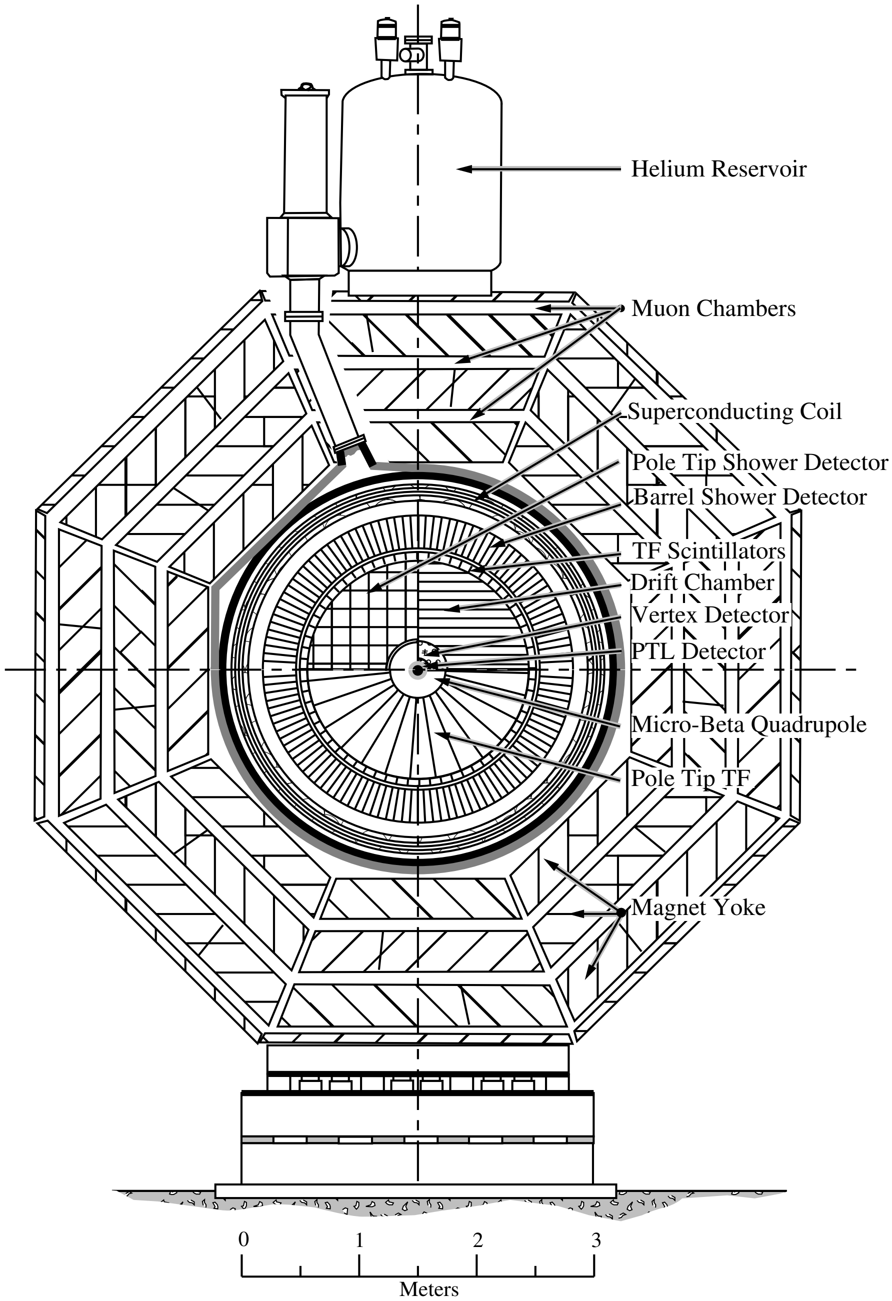

Figure 3
End-view of the CLEO detector.

CHAPTER 2

Models for exclusive semileptonic B meson decay

The semileptonic B meson decays are an important area to investigate because of their large branching fractions and simple theoretical interpretation.

We are going to concentrate on the exclusive decays* $\bar{B}^0 \rightarrow D^{*+}\ \ell^-\ \bar{\nu}$ and $B^- \rightarrow D^{*0}\ \ell^-\ \bar{\nu}$. First we will examine the spectator diagram of B decays. Then we will discuss different theoretical models.

2.1 Spectator Diagram

A subset of B meson decays such as $\bar{B} \rightarrow D^*\ \ell^-\ \bar{\nu}$ can be described by a spectator diagram, in which the light quark $d(u)$in the initial hadron state $B^0(B^-)$ still remains in the final hadron state $D^{*+}(D^{*0})$ while the b quark turns into a c quark by emitting a virtual W^-. The decay of W^- into leptons is separated

*Charge conjugation is implied throughout this study.

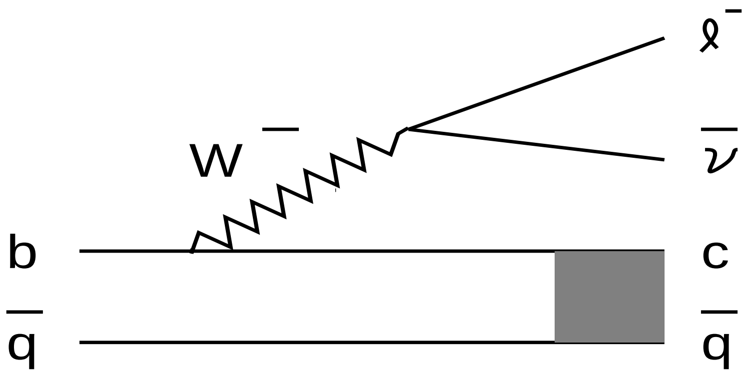

Figure 4
Spectator diagram for semileptonic B decay.

from the quark vertex. To lowest order, the light quark does not participate in the electroweak decay of the b quark and is called a "spectator". Figure 4 shows the Feynman diagram for the "spectator" decay.

In the B rest frame, the decay rate is given by[6]

$$d\Gamma(\bar{B} \to D^{*}\ \ell^{-}\ \bar{\nu}) = \frac{1}{2m_B} M^2 d\Pi^3 \tag{2.1}$$

where the $d\Pi^3$ is the three-body phase space factor and the matrix element M can be expressed as

$$M = \frac{G_F}{\sqrt{2}} |V_{cb}|^2 L^{\mu} H_{\mu}. \tag{2.2}$$

The leptonic and hadronic currents are given by

$$L^{\mu} = \bar{u}_{\ell}(p)\gamma^{\mu}(1-\gamma_5)v_{\nu}(p') \tag{2.3}$$

$$H^{\mu} = \langle k|V^{\mu}_{had} - A^{\mu}_{had}|K\rangle \tag{2.4}$$

where p and p' are the four-momentum of the final state ℓ and ν; and K and k are the four-momentum of the B meson and the D^* meson.

The V-A leptonic current at the W vertex is calculable. The matrix elements in the hadronic current between the two bound states B and D^* involve non-perturbative QCD † and are not calculable. The standard approach is to treat the bound states as ideal structureless particles and do an expansion. In the expansion the unknown bound states effects are absorbed into a set of Lorentz-invariant parameters called *form factors*.

Ignoring the lepton masses the hadronic matrix element of a pseudoscalar B meson decaying into a vector D^* final states are parameterized as[6]

$$\begin{aligned} \langle k, \epsilon | V_\mu | K \rangle &= \epsilon_{\mu\nu\rho\sigma} \epsilon^{*\nu} (K+k)^\rho (K-k)^\sigma \frac{V(q^2)}{m_B + m_{D^*}}, \\ \langle k, \epsilon | A_\mu | K \rangle &= (m_B + m_{D^*}) \epsilon^*_\mu A_1(q^2) - (\epsilon^* \cdot K)(K+k)_\mu \frac{A_2(q^2)}{m_B + m_{D^*}} \end{aligned} \tag{2.5}$$

Where ϵ is the D^* polarization, and $q^2 = (K-k)^2$. The four-momentum transfer q^2 is sometimes referred to as the mass of the virtual W^-. Thus there are three form factors, V, A_1 and A_2, for $\bar{B} \rightarrow D^* \ \ell^- \ \bar{\nu}$ decays . Once they are known, the exclusive decay rates and lepton spectra can be calculated in a straightforward manner.

---

†Quantum chromodynamics (QED) is the theory that describes strong forces.

2.2 The Isgur, Scora, Grinstein and Wise (ISGW) Model

In the ISGW [6] model, a non-relativistic calculation is performed to evaluate the hadron form factors. The non-relativistic approach is reliable since soft hadron recoils contribute significantly to semileptonic decay rates. It is important to know the form factors accurately at the zero-recoil region of the final state meson D^* where q^2 reaches its maximum value.

The ISGW model assumes the binding potential inside a meson is of the form

$$V(r) = -\frac{4\alpha_s}{3r} + c + br. \tag{2.6}$$

with α_s, b and c determined from experimental data. Simple harmonic oscillator wave function solutions are chosen for the Schrodinger equation. The Form factors are found to be

$$F(q^2) = Ce^{\left[-a\frac{q^2_{max}-q^2}{q^2_{max}}\right]}. \tag{2.7}$$

The explicit functional forms of form factors and the absolute decay rate in terms of $|V_{cb}|$ for the $\bar{B} \rightarrow D^* \ \ell^- \ \bar{\nu}$ decay are listed in Table 5. The Dalitz plot predicted by this model is shown in Figure 5.

The non-relativistic approach does not yield the appropriate q^2 dependence for the region of small q^2 . To correct for this a common normalization factor κ is chosen to modify the form factor slope to bring the the pion form factor into agreement with experiment data.

2.3 The Bauer, Stech and Wirbel (BSW) Model

The BSW [7] model calculates the form factors fully relativistically. Solutions to a relativistic scalar harmonic oscillator potential are used.

This model assumes pole-dominance for the form factors as follows

$$F(q^2) = \frac{F(q^2=0)}{1-q^2/M_p^2}. \tag{2.8}$$

The problem reduces to finding the values of the form factors at maximum-recoil($q^2 = 0$), which is the overlap between the initial state and the final state meson wave functions.

The explicit functional forms of the form factors and the absolute decay rates in terms of $|V_{cb}|$ for the $\bar{B} \to D^* \ \ell^- \ \bar{\nu}$ decay are listed in Table 5. The Dalitz plot predicted by this model is shows in Figure 5.

Although a relativistic approach is preferable, this model has its own limitation. There is a free parameter ω which relates to the average transverse momentum of the light quark, $< p_t^2 >= \omega^2$. ω is determined from the semileptonic D decay experimental data to be 0.4. Variation of ω in the range of ± 0.1 causes the $\bar{B} \to D^* \ \ell^- \ \bar{\nu}$ decay rates to change by 25%, so the uncertainty of ω limits the accuracy of the model prediction.

2.4 The Körner and Schuler (KS) Model

In the KS [8] model, quark wave functions are boosted from the B meson rest frame into the W^- rest frame. Equation 2.1 is re-written in terms of helicity amplitudes as

$$\frac{d\Gamma}{dq^2} = \frac{G_F^2}{96\pi^4 {m_B}^2}|V_{cb}|^2 P_{D^*} q^2 \times \left[|H_-|^2 + |H_+|^2 + |H_0|^2\right] \tag{2.9}$$

where $|H_\pm|$and$|H_0|$ are the three possible helicity states of the virtual W^-.

The relations between form factors and helicity amplitudes are

$$\begin{aligned} H_\pm(q^2) &= (m_B + m_{D^*})A_1(q^2) \mp \frac{2m_B\, P_{D^*}}{(m_B + m_{D^*})}V(q^2) \\ H_0(q^2) &= \frac{1}{2m_{D^*}q}\left[(m_B^2 - m_{D^*}^2 - q^2)(m_B + m_{D^*})A_1(q^2) - \frac{4m_B^2\, P_{D^*}^2}{(m_B + m_{D^*})}A_2(q^2)\right]. \end{aligned} \tag{2.10}$$

By comparing the helicity amplitudes with that of the free quark model at $q^2 = 0$, the form factors are found to be

$$F(q^2) = F(q^2 = 0)(\frac{1}{1 - q^2/M_p^2})^n. \tag{2.11}$$

The power n is determined from QCD power counting rules. The overlap $F(q^2 = 0)$ is taken from the BSW model.

The explicit functional forms of form factors and the absolute decay rate in terms of $|V_{cb}|$ for the $\bar{B} \to D^*\ \ell^-\ \bar{\nu}$ decay are listed in Table 5. The Dalitz plot predicted by this model is shown in Figure 5.

2.5 Heavy Quark Effective Theory (HQET)

Isgur and Wise [9] realized that new symmetries exist in the case that quark masses are much larger than the strong interaction scale $\Lambda_{QCD}(\Lambda_{QCD} \approx 200\text{MeV})$. For the hadron system built from a heavy quark and a light quark, the wave function of the light degrees of freedom of light quarks and gluons does not depend on the spin and mass of the heavy quark. The heavy quark is at rest in the hadron's rest frame. The heavy quark plays the role of the color source that binds the light degree of freedom.

HQET is not a phenomenological model. It is based on QCD calculations.

2.5.1 Universal Form Factor $\xi(y)$

In the heavy quark limit($m_Q \to \infty$), the hadron form factors contract into a single one, $\xi(q^2)$. $\xi(q^2)$ is normalized to one at zero-recoil. The conventional form factors can be expressed in terms of $\xi(q^2)$ as

$$V(q^2) = A_2(q^2) = A_1(q^2)\left[1 - \frac{q^2}{(m_B + m_{D^*})^2}\right]^{-1} = R^{*-1}\xi(q^2), \qquad (2.12)$$

where R^{*-1} is a constant determined by the meson masses .[10]

In the $\bar{B} \to D^* \ \ell^- \ \bar{\nu}$ decays, the masses of the two heavy quarks, the b in the initial state and the c in the final state are all much larger than Λ_{QCD}. The universal form factor $\xi(y)$ is calculable in the high q^2 region with corrections up to the order of $1/m_Q$. Corrections have been made to the symmetry breaking hadron form factors.

In order to compare to other form factor models, the convention of three form factors is still utilized. The form factors obtained by Neubert using QCD sum rules are listed in Table 5. The correponding Dalitz plot is shown in Figure 5. It is common to use a new variable

$$y \equiv v \cdot v' = \frac{m_B^2 + m_{D^*}^2 - q^2}{2 m_B m_{D^*}}, \tag{2.13}$$

where $v(v')$ is the 4-velocity of $B(D^*)$ in the B rest frame. $y = 1$ corresponds to q^2_{max}, the zero-recoil point.

2.5.2 Application: Determination of $|V_{cb}|$

Luke's theorem[11] shows that the non-perturbative correction to the normalization of $\xi(y = 1)$ in the finite quark mass is of the order $1/m_Q^2$. This leads to one of the earliest applications in HQET. The experimental study of $\bar{B} \to D^* \ \ell^- \ \bar{\nu}$ decays at D^* zero-recoil allows for a reliable determination of $|V_{cb}|$. This method is to a large extent free of hadron form factor uncertainties. In practice an extrapolation to wider region of y is done to increase the statistical power of the method.

Substituting Equations 2.10 and 2.12 into Equation 2.9, the differential decay rates of $\bar{B} \to D^* \ \ell^- \ \bar{\nu}$ can be rewritten as

$$\begin{aligned} \frac{d\Gamma}{dy} &= \frac{G_F^2}{48\pi^3} m_{D^*}^3 (m_B - m_D^*)^2 |V_{cb}|^2 \eta_A^2 \hat{\xi}^2(y) \times \\ & \sqrt{y^2 - 1}\Big[4y(y+1)\frac{1 - 2yr + r^2}{(1-r)^2} + (y+1)^2\Big], \end{aligned} \tag{2.14}$$

where η_A is the short-distance correction, $\hat{\xi}(y)$ incorporates the finite quark mass effects. Practically, it is convenient to use a compound function

$$\mathcal{F}(y) = \eta_A \hat{\xi}(y)$$

where $\mathcal{F}(1)$ is the product of the theoretically calculated η_A and $\hat{\xi}(1)$. Measurements of the differential decay rate $d\Gamma/dy$ will determine $|V_{cb}|\mathcal{F}(1)$ precisely and also provide information for the shape of the unknown form factor $\hat{\xi}(y)$.

Recent calculations by Neubert[12] give $\eta_A = 0.985 \pm 0.015$, $\hat{\xi}(1) = 0.945 \pm 0.025$, That leads to $\mathcal{F}(1) = 0.93 \pm 0.03$. Another calculation by Shifman *et al*[13] gives $\mathcal{F}(1) = 0.89 \pm 0.03$. Using these theoretical calculations, $|V_{cb}|$ can be extracted from an experimental measurement of $d\Gamma/dy$

2.6 Summary

Several form factor models have been examined briefly in this chapter. The similarity of the lepton momentum distribution and q^2 distribution between different form factor models are manifest in Figure 6 and Figure 7. This implies the dominance of the final hadron state spin, the V-A weak current and the heavy quark constraints. The detailed form factor functional forms will not change the shapes of these distributions. However, the different form factor predictions affect the amplitudes of the decay rate in the amount of 30%. The extraction of $|V_{cb}|$ later in this thesis will adopt the HQET method which is less sensitive to the hadron form factor uncertainties.

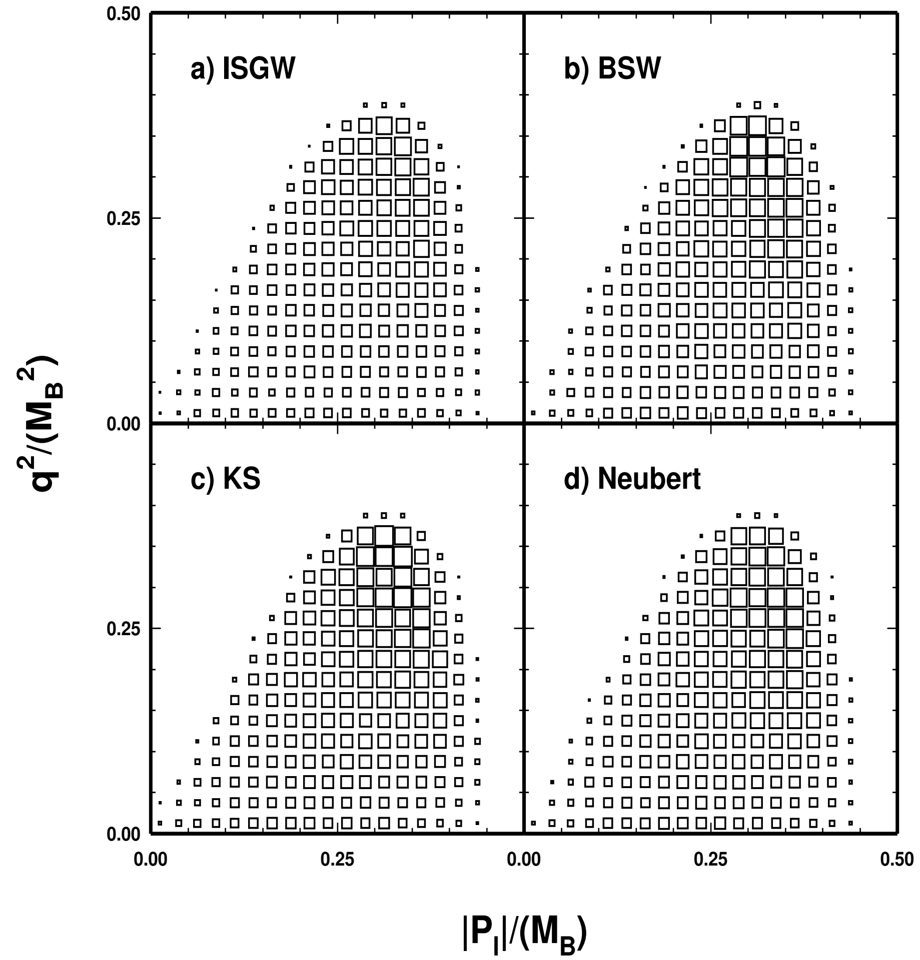

Figure 5

q^2 vs lepton momentum distributions predicted by various form factor models. q^2 is normalized to the mass squared of B meson; $|P_\ell|$ is normalized to the mass of B meson; $M_B = 5.28$ GeV. a) ISGW model; b) BSW model; c) KS model; d) Neubert model.

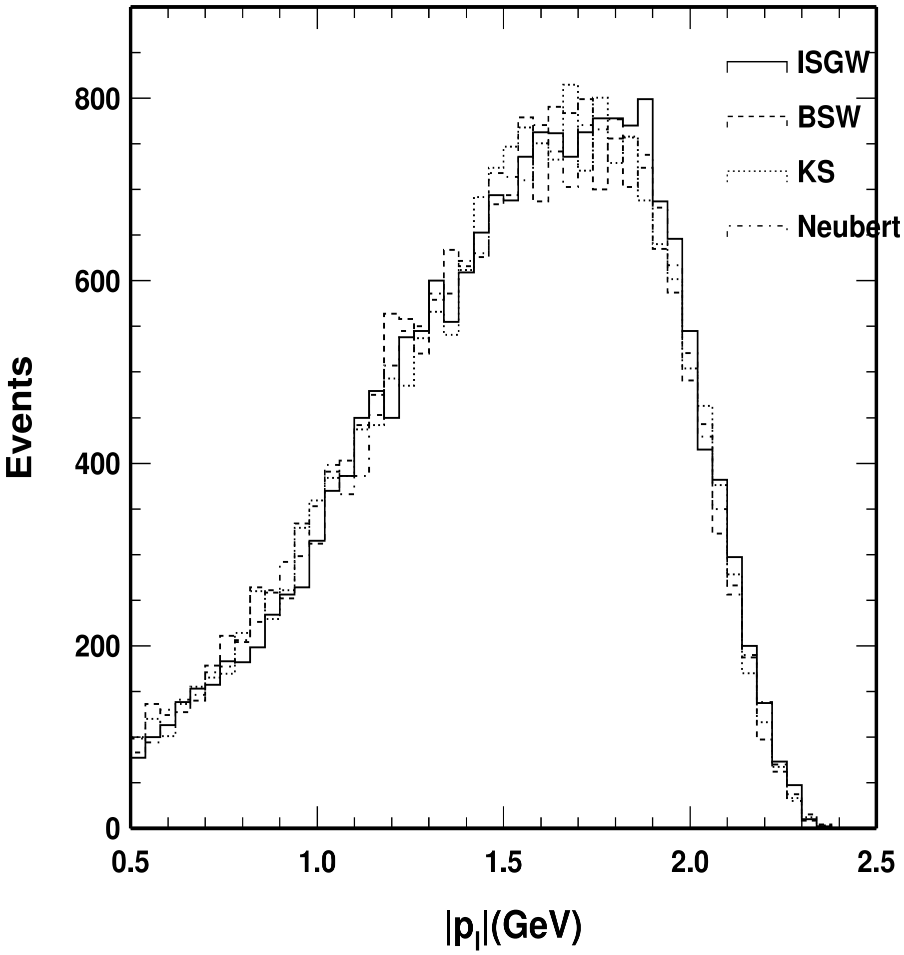

Figure 6

Lepton momentum distributions predicted by various form factor models. a) ISGW model; b) BSW model; c) KS model; d) Neubert model. The normalization is fixed to the same for each model.

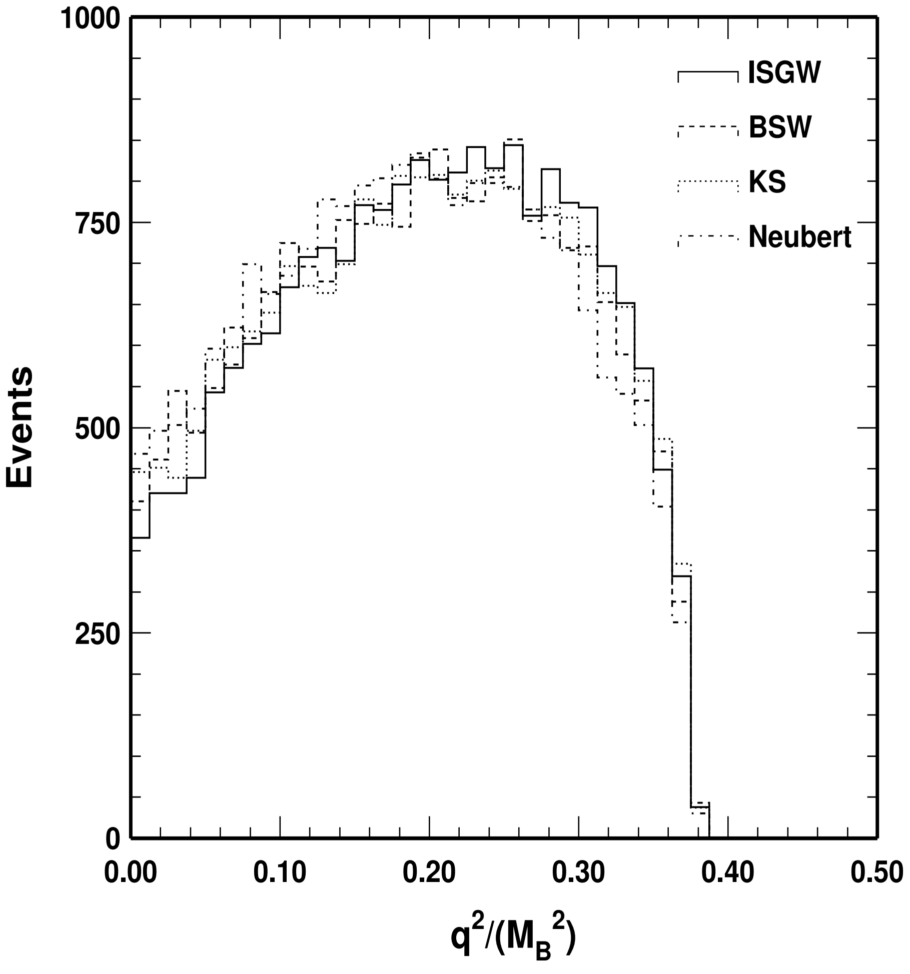

Figure 7

q^2 distributions predicted by various form factor models. q^2 is normalized to the B meson mass squared. a) ISGW model; b) BSW model; c) KS model; d) Neubert model. The normalization is fixed to the same for each model.

Table 5

Predictions for the form factor ratios and their q^2 or y dependence, and the predictions and q^2 dependence for the common form factor A_1. The pole forms have been abbreviated as: $P_1 = 1 - q^2/6.34^2$ and $P_2 = 1 - q^2/6.73^2$ where q is in GeV. $\mathcal{E}(q^2)$ stands for $\exp[-0.03(q^2_{max} - q^2)]$. For the form factor ratios of the Neubert model $N(y) = 2.5/(y+1)$.

	ISGW[6]	BSW[7]	KS[8]	Neubert[10]
$V(q^2)/A_1(q^2)$	1.27	$1.09P_2/P_1$	$1.00/P_1$	$N(y)[1.35 - 0.22(y-1) + 0.09(y-1)^2]$
$A_2(q^2)/A_1(q^2)$	1.14	1.06	$1.00/P_1$	$N(y)[0.79 + 0.15(y-1) - 0.04(y-1)^2]$
$A_1(q^2)$	$0.94\mathcal{E}(q^2)$	$0.65/P_2$	$0.70/P_1$	$0.86[2/(y+1)]^{0.6}$
$\frac{\Gamma(\bar{B}\to D^*\ \ell\ \bar{\nu})}{\lvert V_{cb}\rvert^2}$	24.6 ps^{-1}	21.9 ps^{-1}	25.8 ps^{-1}	29.0 ps^{-1}

CHAPTER 3

Branching Fractions of $\bar{B} \rightarrow D^* \ \ell^- \ \bar{\nu}$ Decays

3.1 Introduction

In high energy physics, a branching fraction is the probability of a parent particle decaying into a specified channel. A branching fraction is defined as the ratio of the number of decays into a specified channel, eg., $A \rightarrow B\ C$, to the total number of parents, A, in the sample; eg., $\mathcal{B}(A \rightarrow B\ C) = N(A \rightarrow B\ C)/N(A)$. Experimentally a branching fraction is measured by first imposing selection criteria on the data to obtain a clean sample of the reaction of interest. These selection criteria are called *cuts* and are imposed on the measured quantities (momentum, energy, angle ...) in the data to remove background events while retaining the reaction of interest. These cuts typically reduce the acceptance or efficiency of the detector to the specific reaction and the number of accepted or reconstructed reactions must be corrected for this loss. If $\epsilon(A \rightarrow B\ C)$ is the efficiency for reconstructing $A \to B\ C$, then $N(A \to B\ C) = n_{reconstructed}/\epsilon(A \to B\ C)$. In addition the number of the parent sample A must be determined.

In this chapter the selection criteria used to obtain a clean signal sample of $\bar{B} \to D^*\ \ell^-\ \bar{\nu}$ events is described in detail, along with the determination of the detection efficiency. We will first discuss the determination of the parent sample of B mesons.

3.2 Data Sample and Selection Criteria

3.2.1 Data Sample

$B\bar{B}$ pairs are produced on the $\Upsilon(4S)$ resonance via $e^+\ e^-$ collisions. Other events such as $e^+\ e^- \to c\bar{c}$, $e^+\ e^- \to u\bar{u}$, $e^+\ e^- \to d\bar{d}$, $e^+\ e^- \to e^+\ e^-$(Bhabha), $e^+\ e^- \to \gamma\ \gamma$ (gamma pairs), $e^+\ e^- \to \mu^+\ \mu^-$ (muon pairs) and $e^+\ e^- \to e^+\ e^-\ X$ (2-photon) also populate this energy region. They are called continuum background. CESR beams are tuned to collide at about 60 MeV below the $\Upsilon(4S)$ resonance for the purpose of estimating the the contribution from continuum. Figure 8 shows the cross section distribution in the Υ energy region.

A cross section is defined as the transition rate per incident flux per target particle. Figure 8 shows the cross section distribution in the Υ energy region. The cross sections for the above three QED* processes $e^+\ e^-$, $\gamma\ \gamma$ and $\mu^+\ \mu^-$

*Quantum Electrodynamics (QED) is the theory that describes electromagnetic forces.

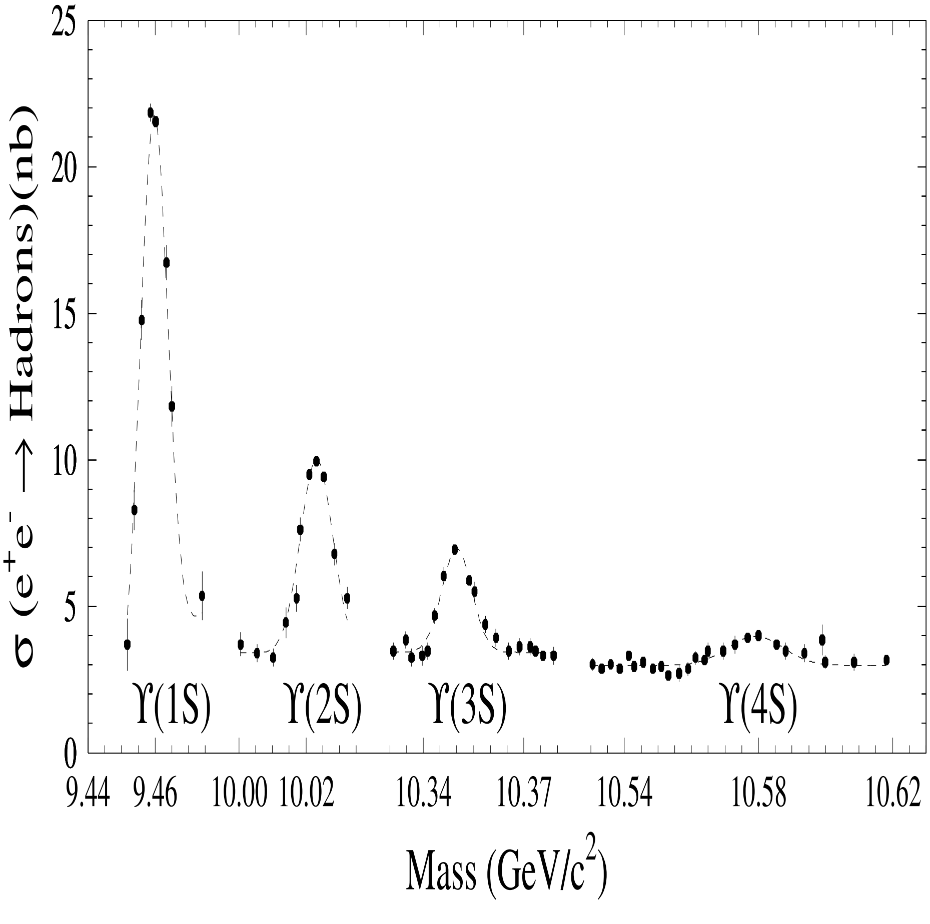

Figure 8
Observed Υ hadronic cross section distribution in the Υ energy region.

are precisely known from theory.[14] We use them to measure the total integrated luminosity in the CLEO II experiment .[15] The total integrated luminosity $\mathcal{L}$ is obtained by

$$\mathcal{L} = \frac{N}{\sigma} \tag{3.1}$$

where N is the number of specific transitions with cross section σ. The average systematic uncertainty on the total integrated luminosity, determined from the three processes, is 1.0%.[15]

This analysis is based on the data collected with the CLEO II detector at CESR during of 1990-1993 as shown in Table 6. The data sample consists of an total integrated luminosity of 1.55 fb^{-1} on the $\Upsilon(4S)$ resonance and 0.69 fb^{-1} at about 60 MeV below the $\Upsilon(4S)$ resonance.

We call those events that are consistent with originating from $\Upsilon(4S)$ decays or continuum charm production as *hadronic* events. The criteria that hadronic events must meet are as follows:

1. Each event must have at least three well-fitted charged tracks originating near the interaction point. This cut removes Bhabha events, gamma pairs or mu pairs, which fail the three-track requirement.

2. The visible energy (total energy of charged tracks plus that depositing in the calorimeter not matched to charged tracks) for each event must be greater than 15% of the center-of-mass energy. Furthermore, the visible energy must be between 20% to 90% of the energy deposited in the calorimeter. This cut rejects 2-photon events.

To determine the number of $B\bar{B}$ pairs $N_{B\bar{B}}$, one only needs to measure the cross section of $\Upsilon(4S)$ hadronic decays [16] once the total luminosity is known. Run conditions such as trigger efficiencies are different from one run block to another. The observed $\Upsilon(4S)$ cross section is measured in each run block separately in the following manner.

The run-by-run cross section is calculated by dividing the number of hadronic events by the integrated luminosity. The run-by-run cross section distribution in each run block is then fitted to Gaussian function to reduce effects from incompletely processed runs or incorrect luminosities. This run-by-run calculation is carried out twice for each run block, once for on-$\Upsilon(4S)$-resonance data, the other for continuum data. Subtraction of the two fitted means gives the observed continuum substracted $\Upsilon(4S)$ cross section $\sigma(\Upsilon(4S))$ for each run block.

The systematic uncertainty of this calculated cross section is dominated by the variation of cross section within each run block, which is about 1.5%. The total number of $B\bar{B}$ pairs is obtained from $\sigma(\Upsilon(4S)) \times \mathcal{L}$. The total uncertainty in the number of $B\bar{B}$ pairs is 1.8% ,[16] including statistical errors and systematic effects. We have already made the assumption that $N_{\Upsilon(4S)} = N_{B^0\bar{B}^0} + N_{B^+B^-}$. These results are summarized in Table 6.

3.2.2 Electron Identification

Electrons are identified[17] by utilizing many measured quantities. The most useful one is E/P, the ratio of energy deposited in the calorimeter to the momentum measured in the drift chamber. The other useful ones are specific ionization

Table 6

Dataset Summary. Each CLEO subset of data was recorded between major CESR down times. The integrated luminosity for individual subsets is in units of pb^{-1}. The error on the number of $B\bar{B}$ pairs includes statistical and systematical errors added in quadrature.

Subset	Date	Run	$\mathcal{L}$(ON/OFF)	$N_{B\bar{B}}$
4S2	11/90-06/91	36233-40095	462846./196990.	$501,300 \pm 9,700$
4S3	09/91-02/92	41386-45969	436095./208509.	$458,300 \pm 8,200$
4S4	03/92-05/92	46035-48057	213768./100878.	$231,140 \pm 4,400$
4S5	07/92-10/92	48939-51406	216243./105173.	$230,690 \pm 4,400$
4S6	11/92-01/93	51431-53701	232546./85004.	$248,400 \pm 4,500$
Total	11/90-01/93	36233-53701	$1.56fb^{-1}/0.70fb^{-1}$	1.67 ± 0.03 millions

losses (dE/dx), the distance between projection of the track and the calorimeter shower, and the shape of the shower in the calorimeter.

Distributions for each of these variables are made for electrons and non-electrons separately. The electron distributions come from embedding radiative Bhabha events into hadronic events where the event topology is closer to that for electrons from $B\bar{B}$ decays. The non-electron distributions come from electron-poor $\Upsilon(1S)$ hadronic events.

For each charged track the probabilities of being an electron(P_e) and non-electron ($P_{\not e}$) are calculated for all the variables. A log-likelihood is formed as follows:

$$R_e = \sum_{\text{variables}} \ln \frac{P_e}{P_{\not e}}, \tag{3.2}$$

The criterion for a good electron candidate is $R_e \geq 3$.

Electron detection efficiencies are determined from embedding Bhabha events into $\Upsilon(4S)$ events. Figure 9 shows the electron detection efficiency.

The probabilities of misidentifying a hadron as an electron (fake rates) are obtained from the lepton-poor $\Upsilon(1S)$ events. The typical fake rate is found to be 0.16%/track.

3.2.3 Muon Identification

Muon identification[18] utilizes the ability of muons to penetrate materials. It is done by matching the drift chamber track projection to hits registered in the muon counter. The muon acceptance decreases dramatically for muon momentum below 2.0 GeV because low-momentum μ's stop or undergo large scattering in the iron absorber. The amount of material or the equivalent number of nuclear interaction lengths traversed by the track is determined by proportional chambers embedded at depths of 36, 78, and 108 cm of the absorbing iron. In this work we require that a muon traverses at least five nuclear interaction lengths of iron.

The muon detection efficiencies are obtained from Monte Carlo simulations and the fake rates from the $\Upsilon(1S)$ events. Figure 9 shows the muon detection efficiency. The typical fake rate is 1.2%/track.

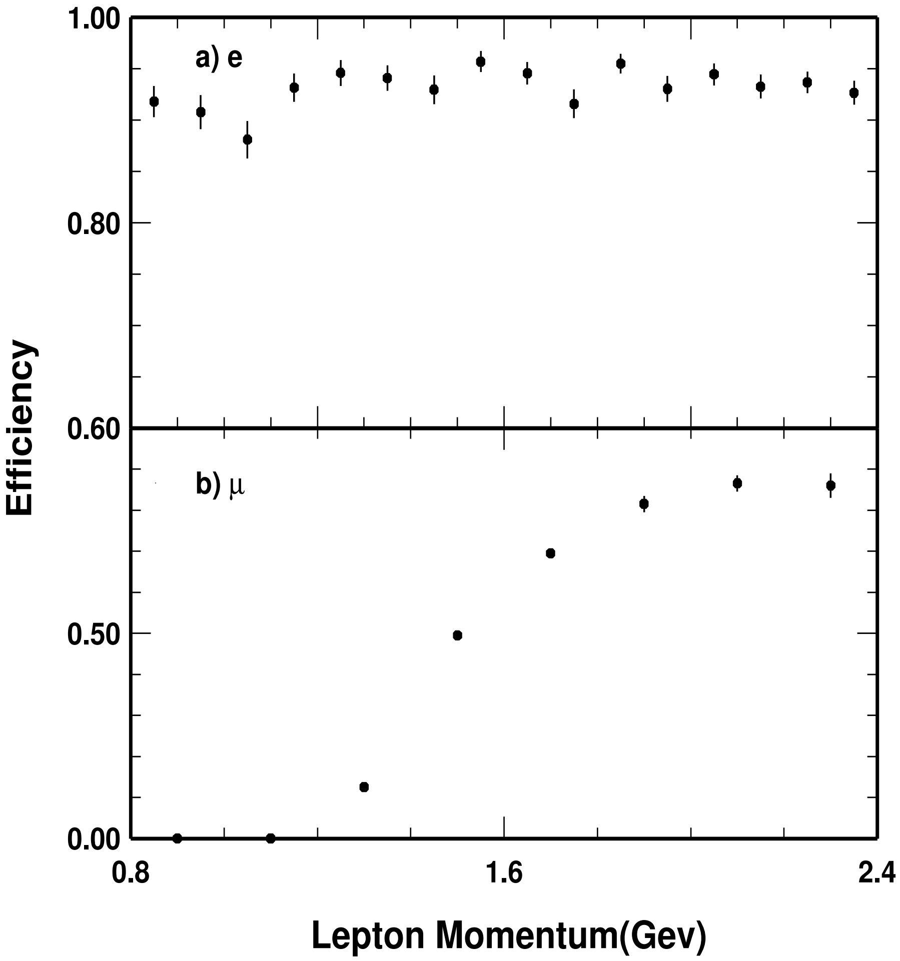

Figure 9
Lepton efficiency distribution for tracks in the polar angle region $|\cos\theta| < 0.7$. a) Electron candidates satisfying $R_e \geq 3$. b) Muon candidates satisfying $R_\mu \geq 5$.

3.2.4 Kaons and Pions

All the charged tracks (excluding electrons and muons) are required to pass the following criteria:

1. A track is not coming from a secondary vertex.
2. The x-y projection of the impact parameter is less than 5 mm.
3. The z (beam pipe direction) projection of the impact parameter is less than 5 cm.
4. All tracks pass "TRKMAN" algorithm.[20] This algorithm eliminates "ghost" tracks where track-fitting ambiguities exist. This algorithm also eliminates "curlers" that have both low transverse and low axial momentum components and make off-center spirals biased toward the beampipe.
5. All tracks have been energy loss corrected according to particle type hypothesis.

3.2.5 Slow Charged Pions

The slow charged pion π_s^+ from D^{*+} is treated differently since its momentum is only 40 MeV in the D^{*+} rest frame. The momenta of charged tracks are measured by the drift chamber (DR), and the geometry parameters of the DR and the magnetic field impose a lower limit on the charged track momentum that can be measured accurately.

A low momentum bound of 65 MeV and a polar angle limit of $45^o < \theta < 135^o$ are imposed on slow charged pions to assure reliable reconstruction efficiency. The second and third requirements in the section 3.2.4 on the impact parameters are excluded to increase the reconstruction efficiency.

3.2.6 Photons and Neutral Pions

Candidate π^0's are reconstructed from pairs of showers in the electromagnetic calorimeter which do not match the projection of any drift chamber track and have an invariant mass within 3σ of the measured π^0 mass ($\sigma = 5$ to 8 MeV, depending on shower energies and polar angles). Showers used in π^0 candidates must be in the polar angle region $|\cos\theta_\gamma| < 0.8$ and have an energy above 30 MeV. The π^0 momentum vector is reconstructed by constraining the shower position and energy measurements to produce the known π^0 mass.

3.3 Reconstruction of $\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$ and $B^- \to D^{*0}\ \ell^-\ \bar{\nu}$

Once all the ingredients are well understood, we combine them to search for the decay modes of interests:

$$\begin{array}{llll} \bar{B}^0 \to & D^{*+} & \ell^-\ \bar{\nu} & \\ & \hookrightarrow & D^0 {\pi^+}_s & \\ & & \hookrightarrow K^-\pi^+ & \end{array} \quad \text{and} \quad \begin{array}{llll} B^- \to & D^{*0} & \ell^-\ \bar{\nu} & \\ & \hookrightarrow & D^0 {\pi^0}_s & \\ & & \hookrightarrow K^-\pi^+ & \end{array}$$

In order to suppress the "jet-like" continuum background and favor the "spherical" resonance events, we require the ratio of the second to zeroth Fox-Wolfram[21] moments ($R2GL$) to be less than 0.4. Figure 10 shows the spectrum of the ratio $R2GL$.

Hadronic events that contain at least one lepton are selected. For each lepton in these events we search for D^0 candidates in the decay mode $D^0 \to K^- \pi^+$. The charges of the kaon and the lepton in each event are required to be the same. D^{*+} candidates are formed by adding π_s^+ to the D^0 candidates. D^{*0} candidates are formed by adding π_s^0 to theD^0 candidates.

The momentum of the D^* candidates must be less than 2.5 GeV to suppress non-$B\bar{B}$ decays as shown in Figure 11.

The cosine of opening angle between D^* and lepton from the same B must be preferentially less than zero as shown in Figure 12. A cut $|\cos(\theta_{D^*,\ell})| < 0$ is applied on data to assure that D^* and lepton are in opposite hemispheres.

Although we have no ability to detect the neutrinos emitted from $\bar{B} \to D^* \ell^- \bar{\nu}$ decay, a strong kinematic constraint can be obtained:

$$\begin{aligned} p_\nu^2 &= (p_B - p_{D^*\ell})^2 \\ &= (E_B - E_{D^*\ell})^2 - |\mathbf{p}_B|^2 - |\mathbf{p}_{D^*\ell}|^2 + 2|\mathbf{p}_B||\mathbf{p}_{D^*\ell}|\cos\Theta \\ &= MM^2 + 2|\mathbf{p}_B||\mathbf{p}_{D^*\ell}|\cos\Theta \end{aligned} \tag{3.3}$$

where $p's$ are the 4-momenta of particles in the subscript, and Θ is the angle between the 3-momenta $\mathbf{p}_B$ and $\mathbf{p}_{D^*\ell}$. MM^2 is traditionally called missing mass squared.[22] Figure 13 shows the $|\mathbf{p}_{D^*\ell}|$ vs MM^2 distribution for the $\bar{B} \to D^* \ell^- \bar{\nu}$ signal events. We define the lepton momentum range 1.4 GeV $< |\mathbf{p}_\ell| <$ 2.4 GeV and the kinematic boundary in Figure 13 as the Signal Region. As shown in Figure 14, the lower lepton momentum cut will suppress cascade decays and c to lepton decays. The upper lepton momentum cut is the kinematic upper limit of $b \to c\, \ell\nu$ decays. The event candidates that contain one D^* and one lepton and reside in the Signal Region are used in the invariance mass $M(K\pi)$ plot to fit for the $\bar{B} \to D^*\, \ell^-\, \bar{\nu}$ yields. The yields in the Signal Region are listed in Table 7 and Table 8. The determination of background is discussed in detail in the following section.

3.4 Background Estimate

As in all analyses, a major effort has been spent in understanding the background sources and correcting for them appropriately. There are five sources of backgrounds: combinatoric, correlated, uncorrelated, fake lepton and continuum background. They are discussed below in more detail.

3.4.1 Combinatoric Background

The combinatoric background arises from the process of forming the D^* candidate using the three particles K, π, and π_s. The flat and broad combinatoric background lies underneath the signal enhancement of interests as shown in Figures 15 and Figures 16. This background falls into two categories.

1) π_s not from a D^* plus any K, π combination. This category does not peak at $\delta m \equiv M(K\pi\pi_s) - M(K\pi)$ spectrum. It is shown in the Figure 17 and Figure 18 as the dashed lines. A cut around the nominal δm value eliminates most of

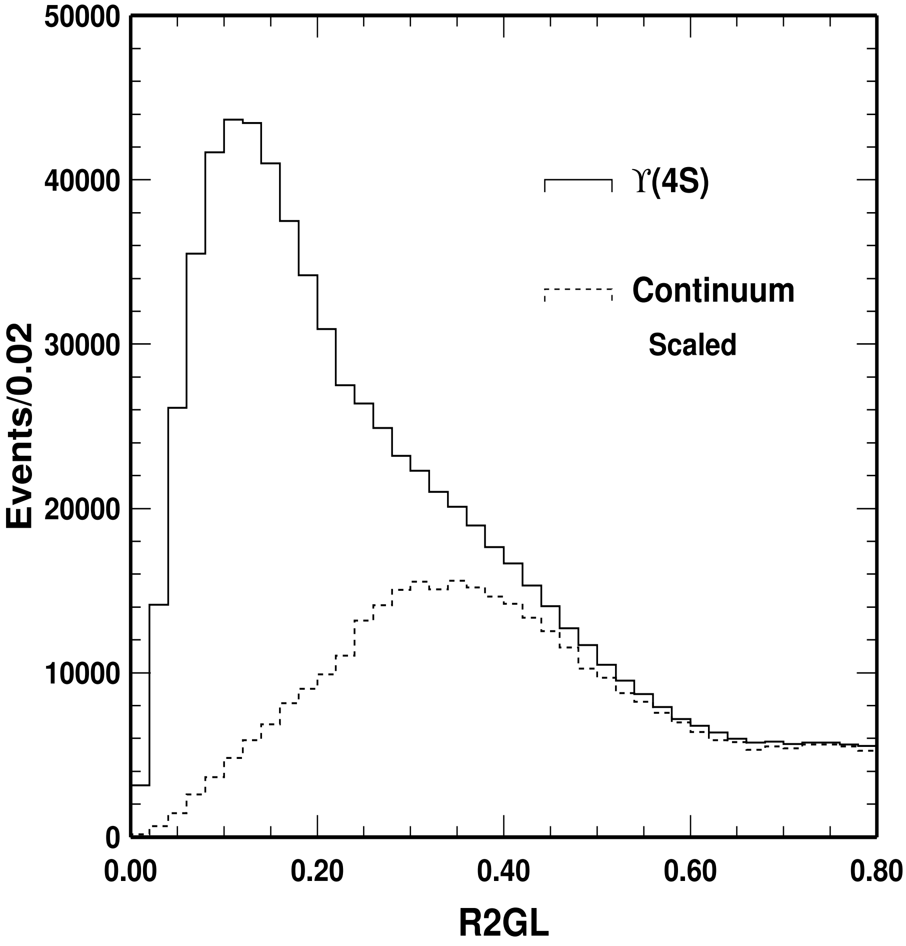

Figure 10

R2GL distribution for the dataset. The continuum is scaled properly according to the ratio of luminosities and ratio of run energies. The region of $R2GL > 0.4$ is saturated by continuum data.

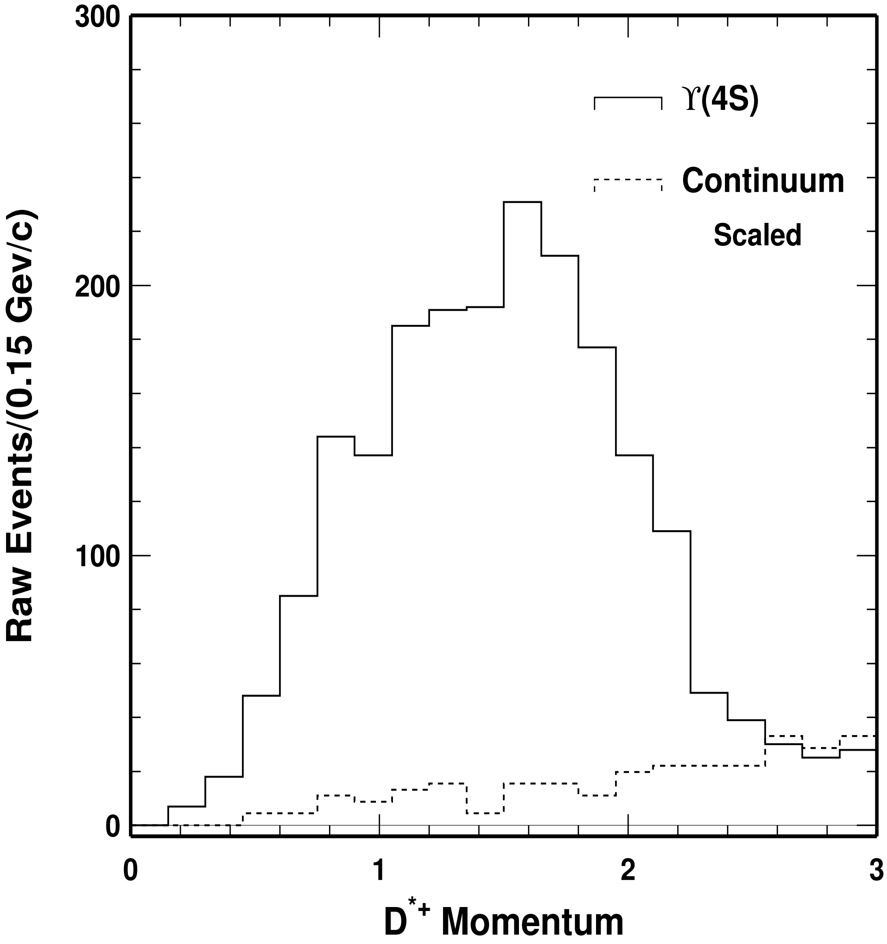

Figure 11
D^{*+} momentum distribution in data. The dashed histogram represents continuum data properly scaled according to the ratio of luminosities and ratio of run energies.

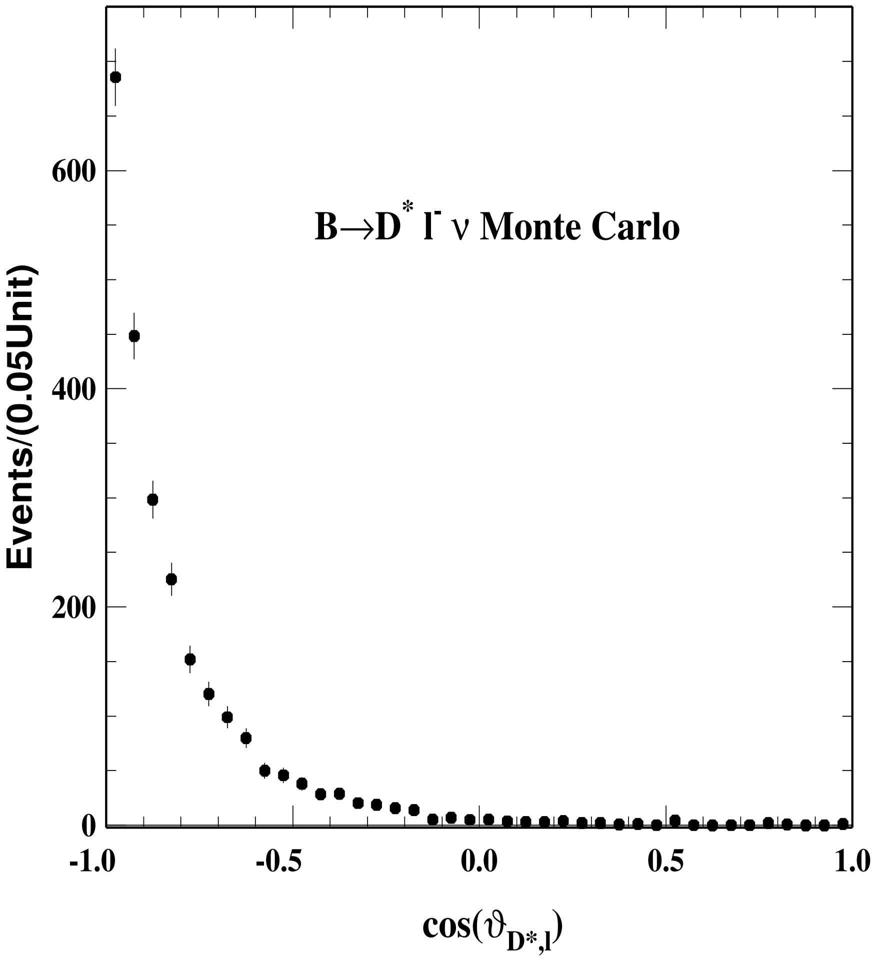

Figure 12
The distribution of cosine of the angle between D^* and lepton in the $\bar{B} \to D^* \ell^- \bar{\nu}$ Monte Carlo simulation.

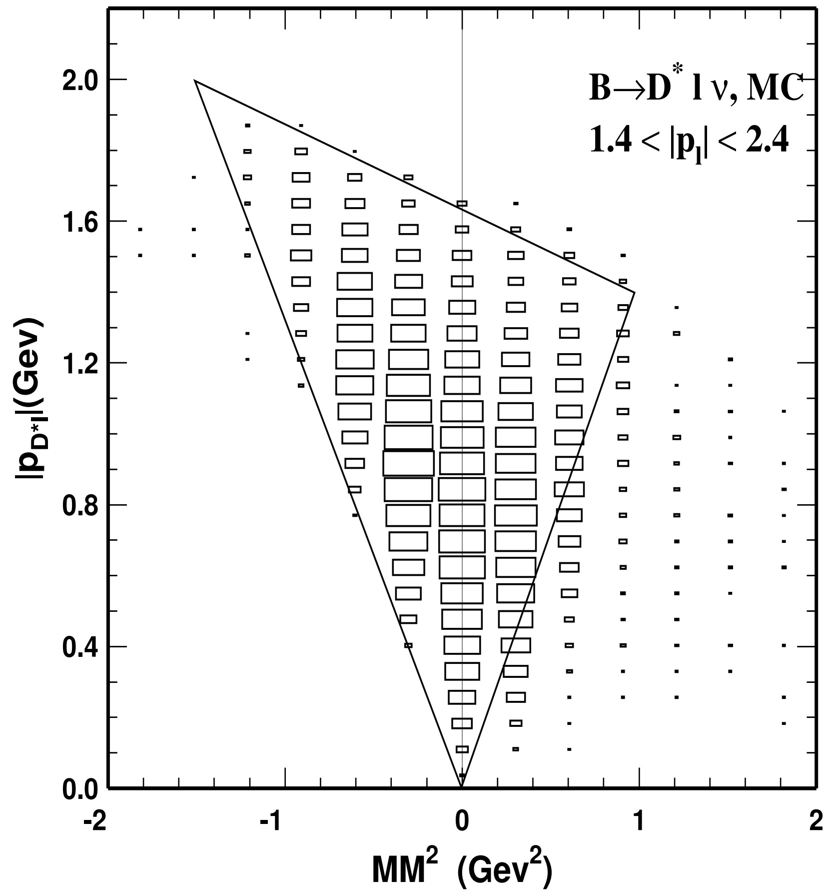

Figure 13

Kinematic boundary for $\bar{B} \to D^{*}\ \ell^{-}\ \bar{\nu}$ decays in $|\mathbf{p}_{D^{*}\ell}|$ vs MM^2 plane determined by $\bar{B} \to D^{*}\ \ell^{-}\ \bar{\nu}$ Monte Carlo simulation. The triangle ((0,0),(1.0,1.4),(-1.5,2.0)) defines the Signal Region. The reconstructed events outside the boundary are due to final state radiation and Bremsstrahlung.

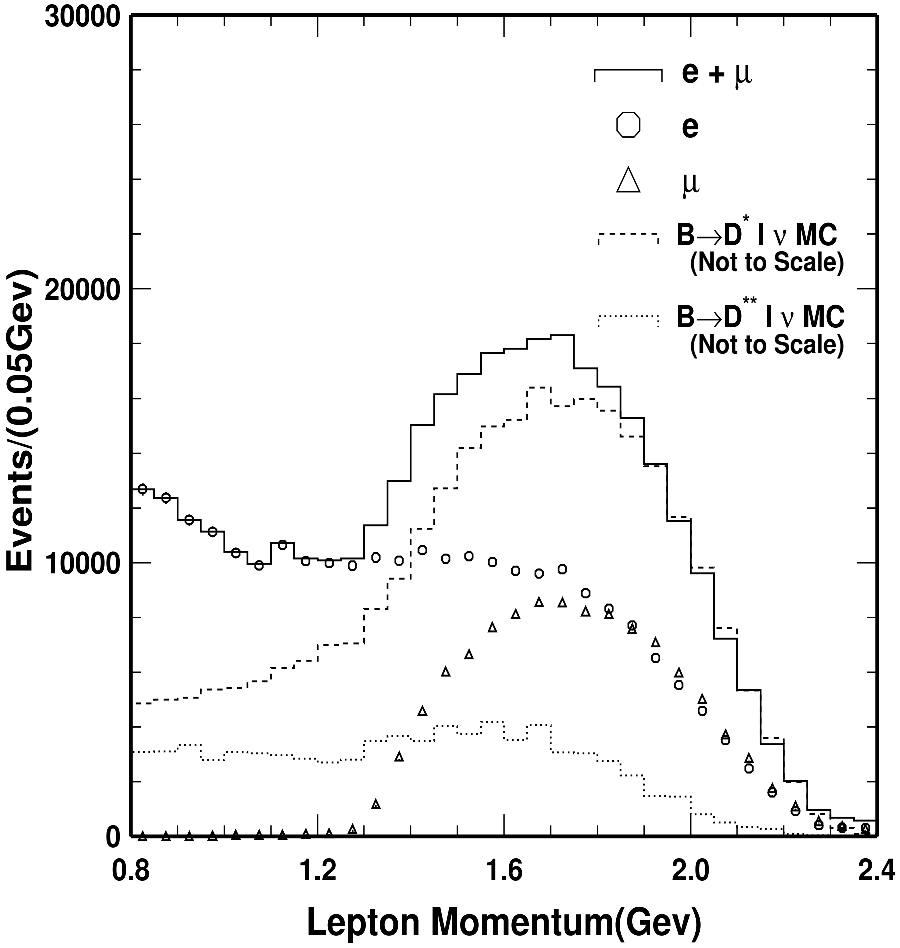

Figure 14

Lepton momentum distributions. The circles (triangles) are electrons (muons) in $\Upsilon(4S)$ data after continuum subtraction.

the background of this category. We cut at $|\delta m - 145.4| < 4.0$ MeV for D^{*-} and $|\delta m - 142.5| < 2.5$ MeV for D^{*0}. The residue of fake π_s in the δm signal region is modeled by the δm sideband. We use 10 MeV $< |\delta\text{m} - 145.4| <$ 50 MeV for D^{*-} and 5.5 MeV $< |\delta\text{m} - 145.4| <$ 17.5 MeV for D^{*0} as sideband regions. We then project the δm signal region and scaled sideband region onto $M(K\pi)$ to investigate the next category of background.

2) Random K, π combination plus any π_s. This category does not peak in the $M(K\pi)$ spectrum. It is removed when the $M(K\pi)$ spectra are fitted to the sum of two Gaussian functions and a first order polynomial. The sum of two Gaussian parts in Figure 19 and Figure 20 represent the $M(K\pi)$ signal. The first order polynomial parts in Figure 19 and Figure 20 represent this category of background.

To correct the residue of fake π_s in the δm signal, we subtract the fitting results of δm sideband projection from that δm signal projection. The results for the two modes are shown in Table 7 and Table 8.

3.4.2 Correlated D^*–lepton Background

The next major contribution to the background is from Correlated D^*–lepton pairs. "Correlated" means that the D^* and lepton are the decay product of the same B. For instance, $\bar{B} \to D^* \, \ell^- \, \bar{\nu}$ is reconstructed out of decay $\bar{B} \to D^* \, X \, \ell^- \, \bar{\nu}$. The characteristics of the background are studied with a Monte Carlo simulation of $\bar{B} \to D^{**} \, \ell^- \, \bar{\nu}$ decay based on the ISGW [6] model. The Monte Carlo sample

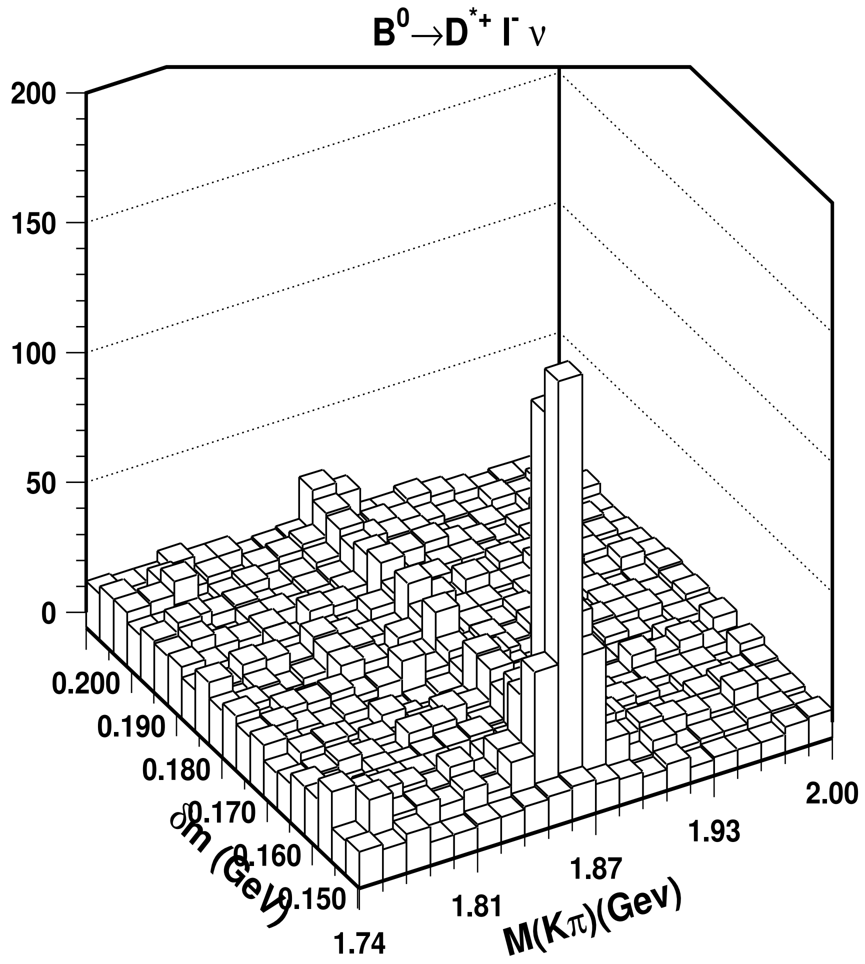

Figure 15
δm vs $M(K^-\pi^+)$ distribution from $\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$ decay in the Signal Region.

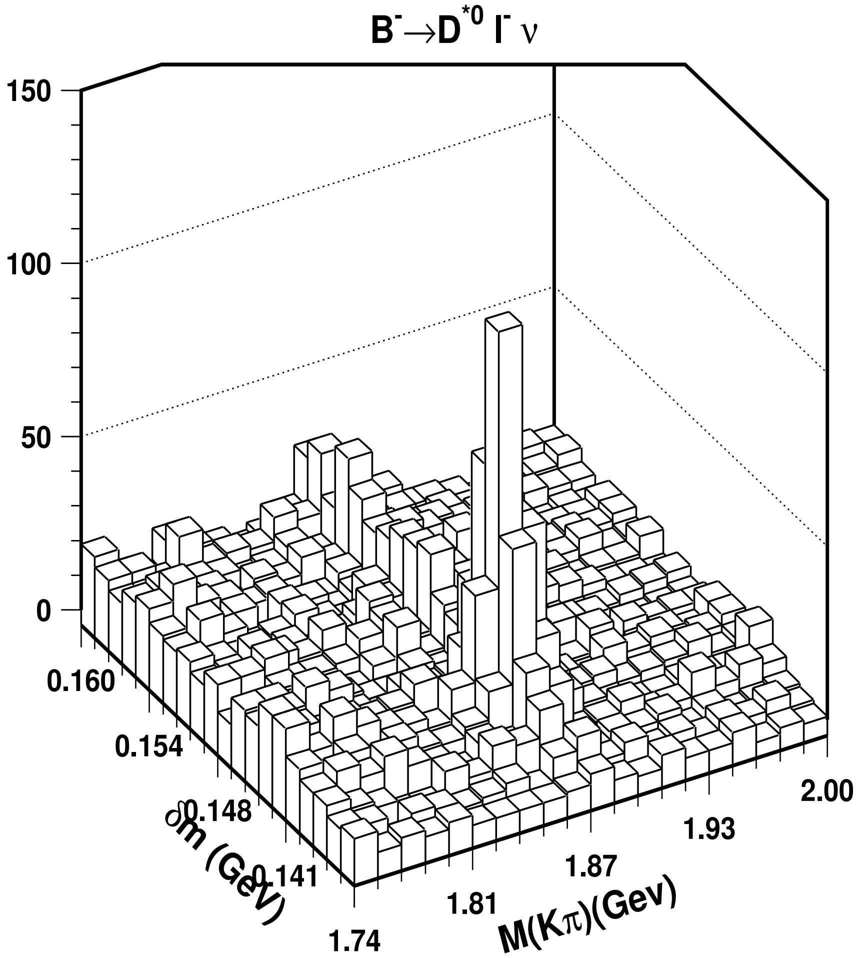

Figure 16
δm vs $M(K^-\pi^+)$ distribution from $B^- \to D^{*0}\ \ell^-\ \bar{\nu}$ decay in the Signal Region.

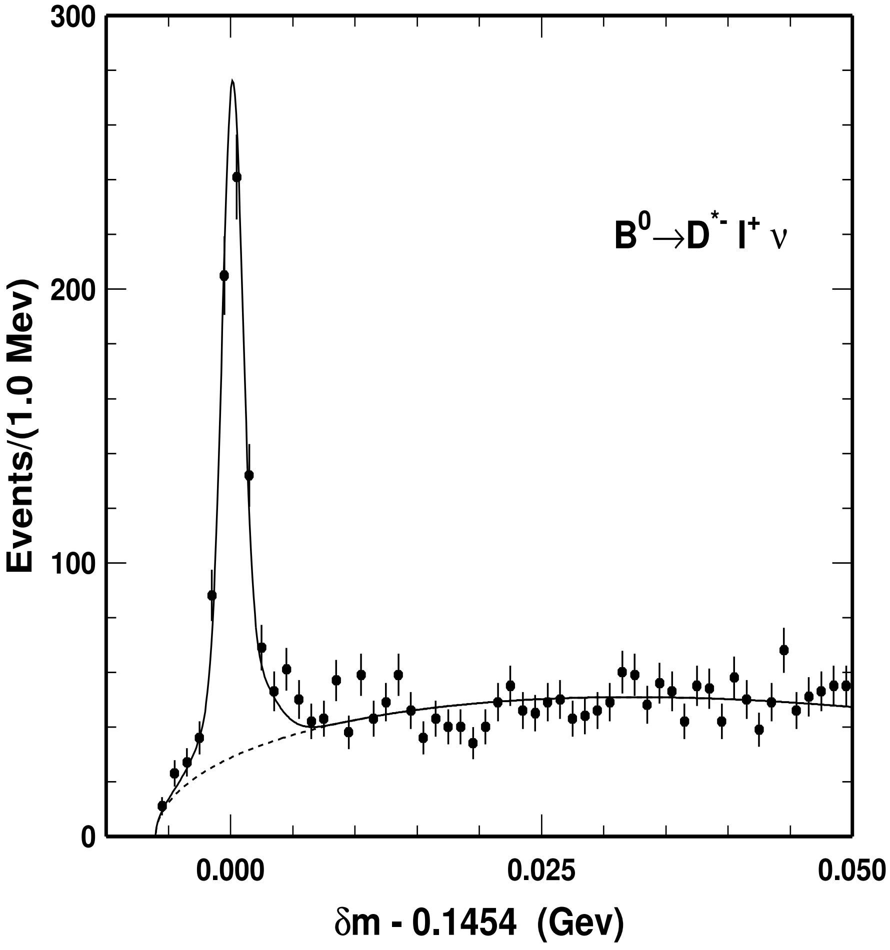

Figure 17

δm distribution for $\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$ decay in the Signal Region. The signal is fitted to the sum of two Gaussian functions. The background is fitted to the function $p_1(x+0.006)^{1/2} + p_2(x+0.006)^{3/2}$ with $x = \delta m - 0.1454$ and $p_i(i = 1, 2)$ are free parameters.

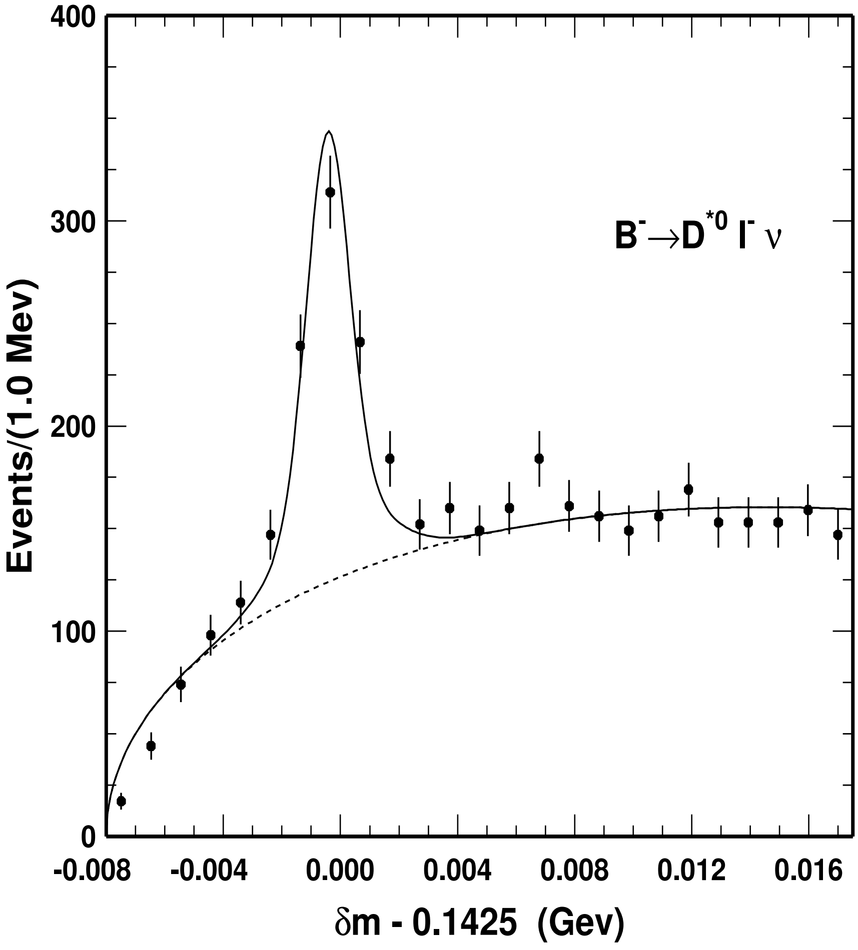

Figure 18

δm distribution for $B^- \rightarrow D^{*0}\ \ell^-\ \bar{\nu}$ decay in the Signal Region. The signal is fitted to the sum of two Gaussian functions. The background is fitted to the function $p_1(x+0.008)^{1/2}+p_2(x+0.008)^{3/2}$ with $x=\delta m-0.1425$ and $p_i(i=1,2)$ are free parameters.

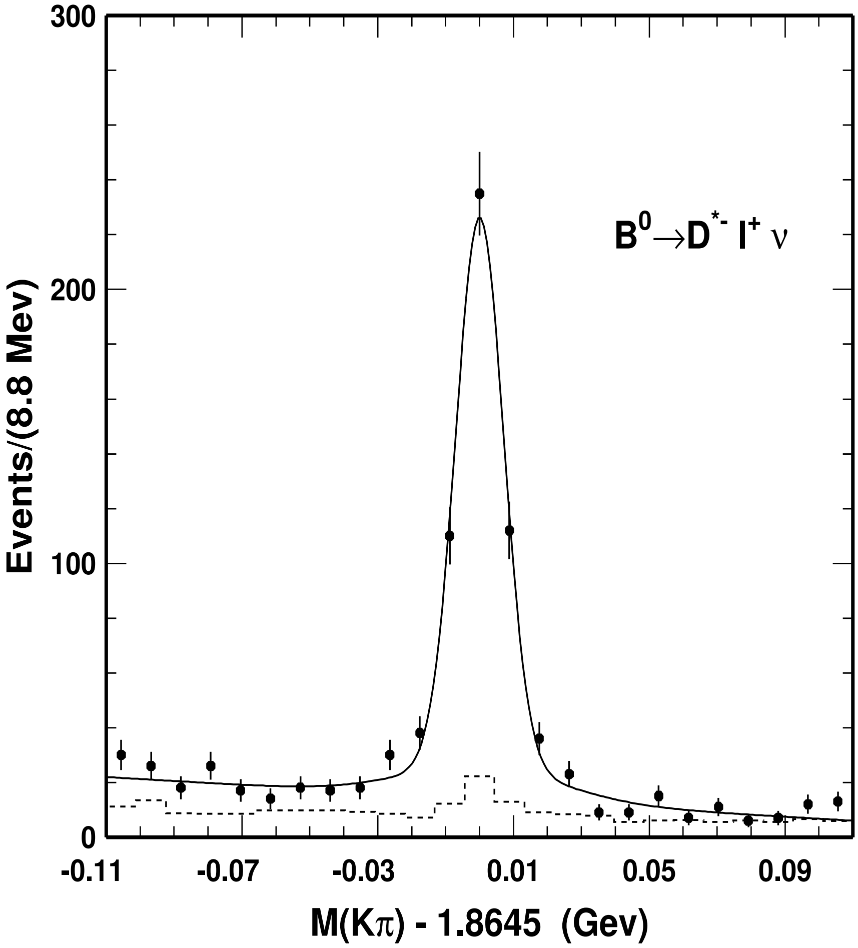

Figure 19

$M(K^-\pi^+)$ distribution for $\bar{B}^0 \rightarrow D^{*+}\ \ell^-\ \bar{\nu}$ decay in the Signal Region. The dots are the projection of the δm signal. The dashed histogram is the projection of the scaled δm sideband. The difference of the fitting results on these two histograms gives the $\bar{B}^0 \rightarrow D^{*+}\ \ell^-\ \bar{\nu}$ yields. The signal is fitted to the sum of two Gaussian functions. The background is fitted to a first order polynomial.

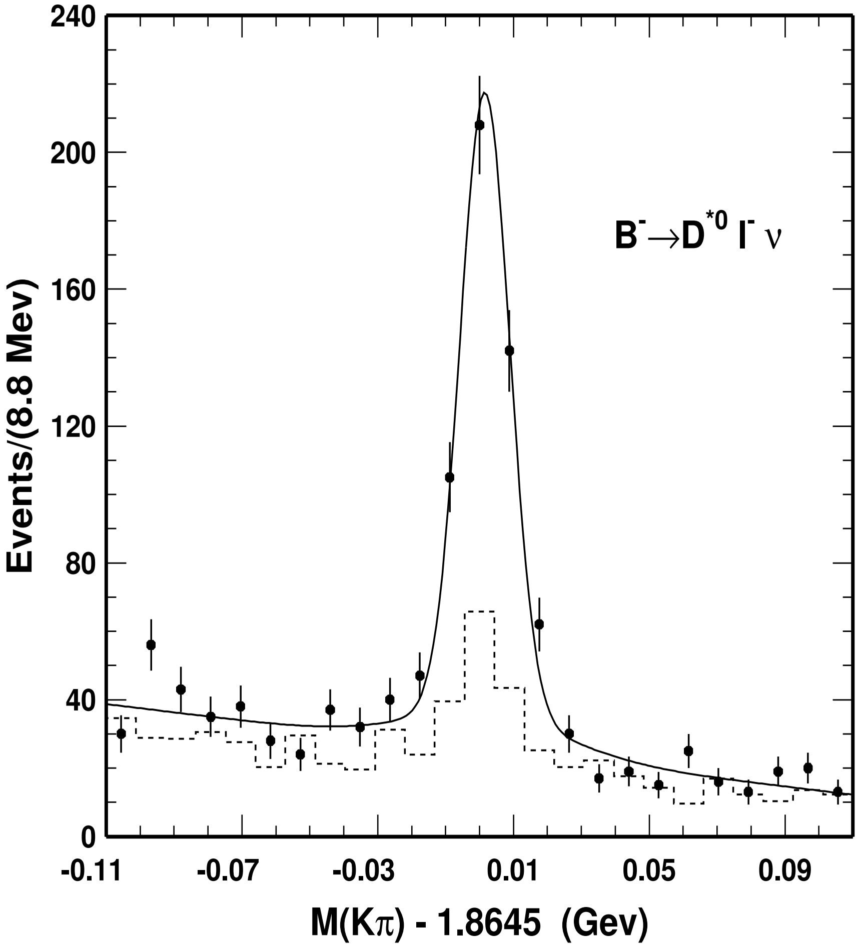

Figure 20

$M(K^-\pi^+)$ distribution for $B^- \rightarrow D^{*0}\ \ell^-\ \bar{\nu}$ decay in the Signal Region. The dots are the projection of the δm signal. The dashed histogram is the projection of the scaled δm sideband. The difference of the fitting results on these two histograms gives the $B^- \rightarrow D^{*0}\ \ell^-\ \bar{\nu}$ yields. The signal is fitted to the sum of two Gaussian functions. The background is fitted to a first order polynomial.

is generated with ${}^1P_1, {}^3P_1$ and ${}^3P_2$ D^{**} states with predicted abundances.[23] The broad resonant state ${}^3P_1$ is expected to approximate the non-resonant $(D^* \ \pi)$ contribution. Figure 21 shows the $|\mathbf{p}_{D^*\ell}|$ vs MM^2 distributions for this decay in two lepton momentum regions: 1.4 GeV $< |\mathbf{p}_\ell| <$ 2.4 GeV and 0.8 GeV $< |\mathbf{p}_\ell| <$ 1.4 GeV. The lower lepton momentum range provides a better handle on the $\bar{B} \to D^* \ X \ \ell^- \ \bar{\nu}$ yields since the lepton momenta tends to be soft in this case. Part (a) of Figure 21 shows a faction of $\bar{B} \to D^{**} \ \ell^- \ \bar{\nu}$ events leaking into the Signal Region. The dashed lines in part (b) of Figure 21 define a Correlated Background Region that can be used to constrain the amount of $\bar{B} \to D^{**} \ \ell^- \ \bar{\nu}$ events that leak into the Signal Region. The fact that the unreconstructed X always carries away energy causes the $|\mathbf{p}_{D^*\ell}|$ vs MM^2 distribution to shift to the positive side and $|\mathbf{p}_\ell|$ is softer.

We describe the Signal Region by using the first line in Equation (3.4). N_S is the total yield in the Signal Region after the correction for all other backgrounds discussed in the following sections. $N_s(\bar{B} \to D^* \ \ell^- \ \bar{\nu})$ is the number of $\bar{B} \to D^* \ \ell^- \ \bar{\nu}$ signal events. $N_c(\bar{B} \to D^* \ X \ \ell^- \ \bar{\nu})$ is the number of $\bar{B} \to D^* \ X \ \ell^- \ \bar{\nu}$ correlated background events that leak into the signal region.

$$\left.\begin{array}{rcrcr} N_S & = & N_s(\bar{B} \to D^* \ \ell^- \ \bar{\nu}) & + & N_c(\bar{B} \to D^* \ X \ \ell^- \ \bar{\nu}) \\ N_C & = & r_s N_s(\bar{B} \to D^* \ \ell^- \ \bar{\nu}) & + & r_c N_c(\bar{B} \to D^* \ X \ \ell^- \ \bar{\nu}) \end{array}\right\} \tag{3.4}$$

We describe the Correlated Background Region by using the second equation in Equation 3.4. N_C is the total yields in the Background Region after correcting for the backgrounds in the following sections. r_s $(= 0.033 \pm 0.006)$ is the ratio of the efficiency for $\bar{B} \to D^* \ \ell^- \ \bar{\nu}$ events in the Correlated Background Region to the efficiency in the Signal Region, and r_c $(= 1.5 \pm 0.7)$ is the ratio of the efficiency for $\bar{B} \to D^{**}\, \ell^-\, \bar{\nu}$ in the Correlated Background Region to the efficiency in the Signal Region. A conservative error of 50% is assigned to r_c since the Monte Carlo events are generated depending on theoretical models which describe the resonant structures only.

Solving the simultaneous equations in Equation 3.4 gives the results of N_S and N_C. Figure 22 and Figure 23 show the $M(K\pi)$ distributions of the correlated backround. The results are listed in Table 7 and Table 8.

Other sources of background such as $\bar{B} \to D^*\, \tau\, \nu$ and $\bar{B} \to D^*\, D^{(*)}_{(s)}$ have $|\mathbf{p}_{D^*\ell}|$ vs MM^2 distributions very similar to that of $\bar{B} \to D^{**}\, \ell^-\, \bar{\nu}$. Figure 24 shows the MM^2 distributions for these modes. The values of r_c are estimated by Monte Carlo simulation to be 100 ± 100 and 200 ± 200 for the two channels respectively. The contamination of these two channels to the Signal Region is negligible. Their contributions to the Background Region are estimated using various intermediate branching ratio[3,24] and efficiencies determined in Monte Carlo simulations. The results are listed in Table 7 and Table 8.

3.4.3 Uncorrelated Background

"Uncorrelated" background results from true $\Upsilon(4S)$ decays where the lepton comes from one B and the D^* comes from the opposite $\bar{B}$.

These physical processes include:

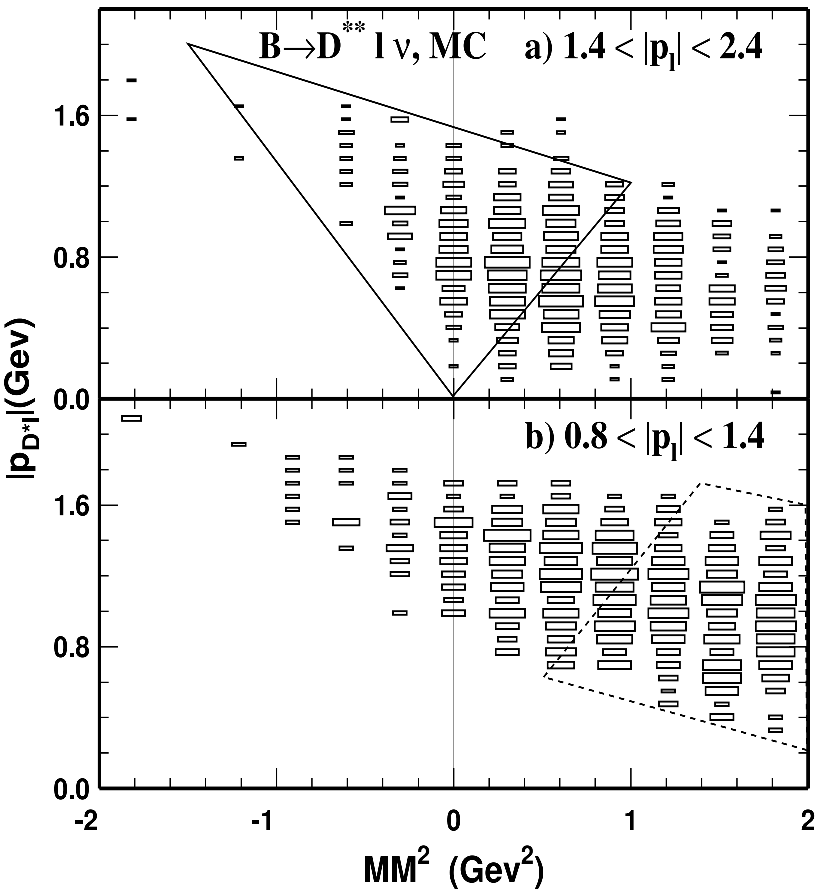

Figure 21

$\bar{B} \to D^{**}\ \ell^-\ \bar{\nu}$ Monte Carlo simulation. (a) The feed-down of $\bar{B} \to D^{**}\ \ell^-\ \bar{\nu}$ decays into the Signal Region(triangle). (b) The dashed lines define the kinematic boundary for Correlated Background Region.

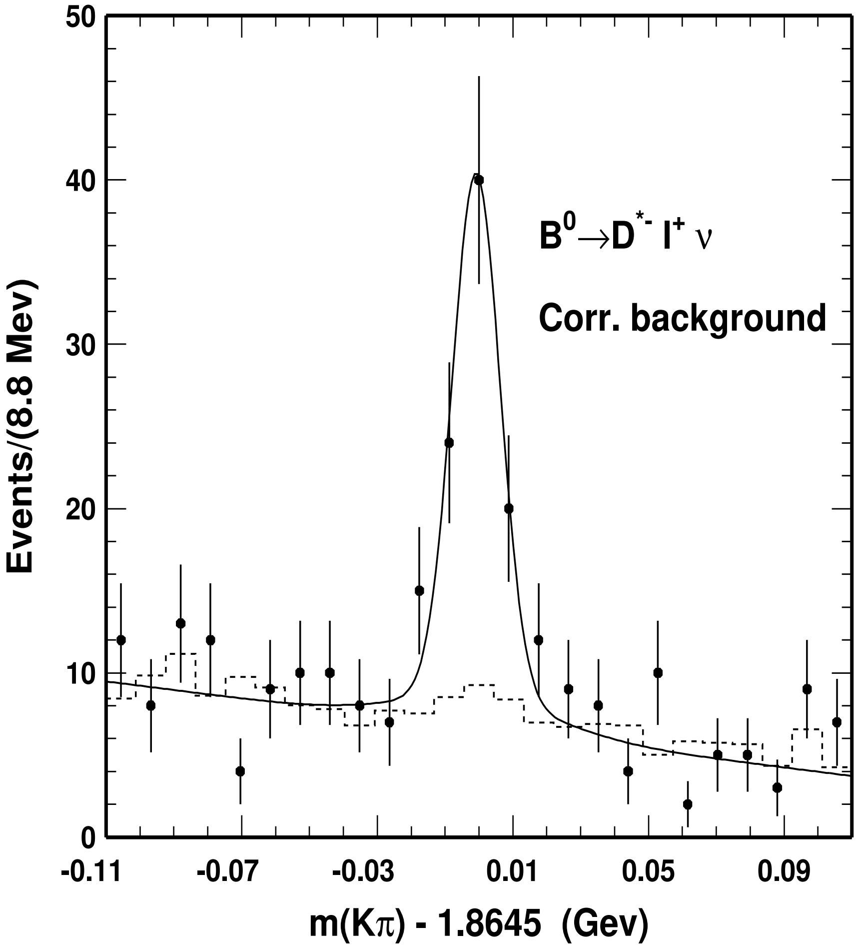

Figure 22

$M(K^-\pi^+)$ distribution for the Correlated Background Region of $\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$ decay. The dots are the projection of the δm signal. The dashed histogram is the projection of the scaled δm sideband. The signal is fitted to the sum of two Gaussian functions. The background is fitted to a first order polynomial.

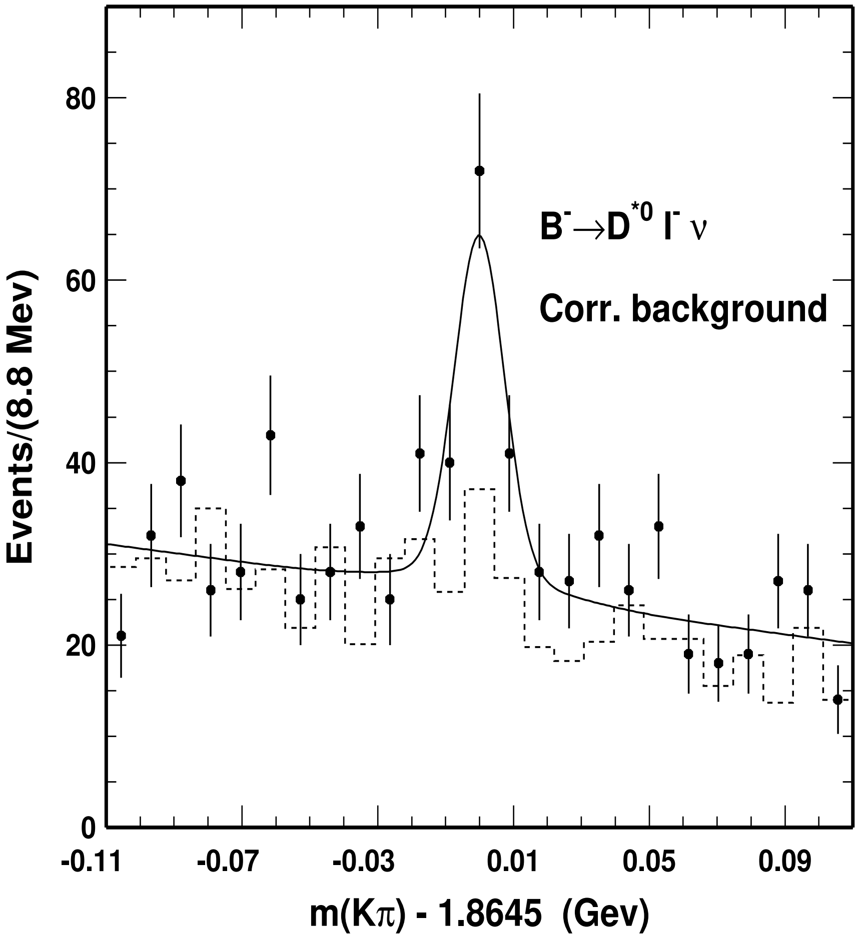

Figure 23

$M(K^-\pi^+)$ distribution for the Correlated Background Region of $B^- \to D^{*0}\ \ell^-\ \bar{\nu}$ decay. The dots are the projection of the δm signal. The dashed histogram is the projection of the scaled δm sideband. The signal is fitted to the sum of two Gaussian functions. The background is fitted to a first order polynomial.

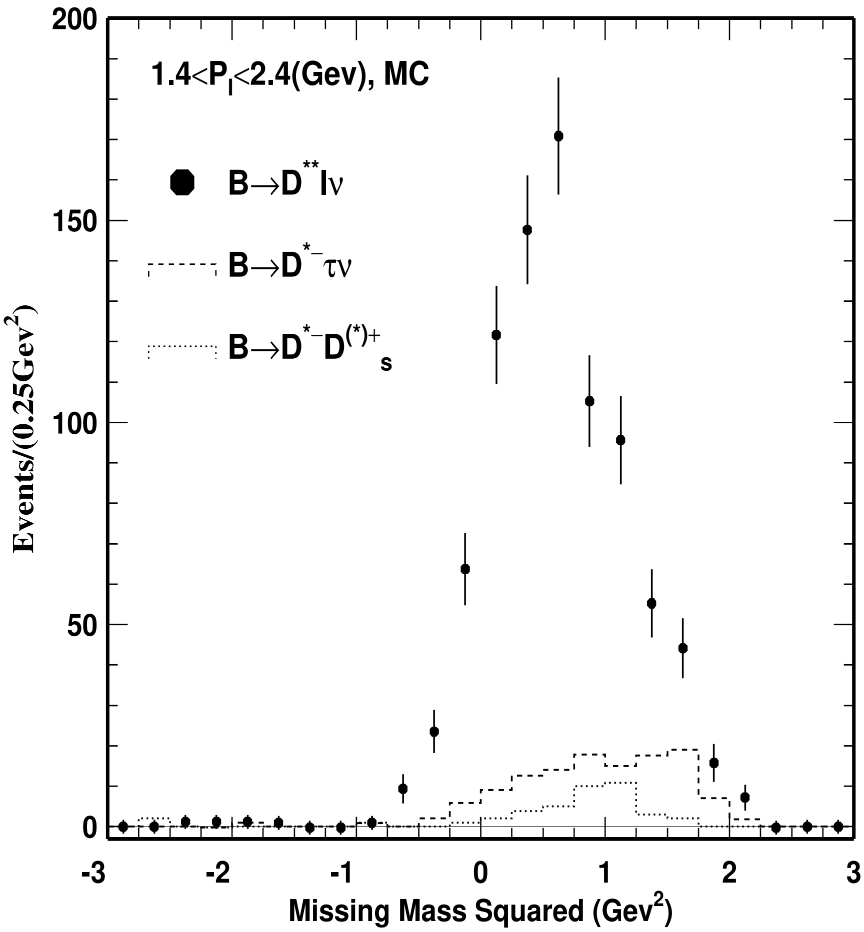

Figure 24

The MM^2 distribution for correlated background sources that pass all cuts. The solid points represent the $B \to D^{**} l\nu$ Monte Carlo events, the dashed histogram is for $\bar{B} \to D^{*}\tau^{-}\bar{\nu}_{\tau}$ Monte Carlo events and the dotted histogram is for $\bar{B} \to D^{*+}D_{s}^{(*)-}$ Monte Carlo events. The figures have been normalized so that the areas show the relative efficiencies between the three modes.

1. $(\bar{B}^0 \to D^{*+}X)(B^0 \to \bar{B}^0 \to l^- X)$

2. $(\bar{B} \to D^{*+}X)(B \to \bar{D}X)$ followed by $(\bar{D} \to l^- X)$

3. $(\bar{B} \to D^{*+}X)(B \to Xl\nu)$ where there is a double misidentification of the K and the π in the reconstruction of the D^0, or the D^0 decays to a Cabibbo suppressed channel.

4. $(\bar{B} \to D^{*+}X)(B \to \psi X)$ followed by $\psi \to l^+ l^-$

5. $(\bar{B} \to D^{*+}X)(B \to Xl_{fake})$

To subtract these backgrounds we utilize the fact that signal events are produced preferentially with the D^* and the lepton in opposite hemispheres as shown in Figure 12. For uncorrelated background the angle between the D^* and the lepton is isotropic since the B's are produced almost at rest at the $\Upsilon(4S)$. The "flipping" technique[25] developed in CLEO is used to estimate this uncorrelated background. We take all events in the sample with the D^* and the lepton in the same hemisphere and flip the lepton direction by 180 degrees. After subtracting the δm sideband we are left with the uncorrelated background to our sample. The results are listed in Table 7 and Table 8.

3.4.4 Fake lepton and Continuum Background

Fake leptons (misidentified hadrons) that combine with a D^* from the same B meson also contributes to correlated background. Contributions from fake leptons where the D^* and the lepton are uncorrelated is already removed using the "flipping" technique.

To obtain the number of fake leptons, we first perform the same analysis using tracks that are not identified as leptons and then scale down the yields by the appropriate momentum dependent misidentification probability per hadron.

We evaluate the contribution from continuum events using the 0.69 fb^{-1} of data taken at energies about 60 MeV below the $\Upsilon(4S)$. The OFF-$\Upsilon(4S)$ yield is scaled by a factor correcting for the differences in the total luminosity between the two datasets, and for the fact that the continuum QCD cross section various with the inverse square of the center of mass energy $(1/s)$:

$$f_{\mathcal{L}} = \frac{1.55 fb^{-1}}{0.69 fb^{-1}} \times \frac{(10.52\mathrm{GeV})^2}{(10.58\mathrm{GeV})^2} = 2.22 \tag{3.5}$$

We subtract the scaled OFF-$\Upsilon(4S)$ events from the total yield as shown in Table 7 and Table 8.

3.5 Reconstruction Efficiencies

We describe the procedure of determining the reconstruction efficiencies for $\bar{B} \to D^* \ \ell^- \ \bar{\nu}$ events and then discuss their systematic uncertainties in this section.

Table 7

A summary of the events in the $\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$ Signal Region and Correlated Background Region. In the top 2 rows the first error is statistical and the second is the systematic uncertainty due to the choice of functions used to fit the data. The bottom two rows have been computed using Equation (3.4) and the uncertainties in r_s and r_c have been propagated into the systematic error.

$D^{*+}l^-$	Signal Region	Correlated Background Region
δ_m Signal	$482 \pm 25 \pm 9$	$80 \pm 11 \pm 1$
Scaled δ_m Sideband	$27 \pm 3 \pm 1$	5 ± 2
Uncorrelated Background	9 ± 4	25 ± 5
Continuum	4 ± 4	0 ± 4
Correlated Lepton Fakes	4 ± 1	2 ± 1
$B \to D^*\tau\bar{\nu}$ and $B \to D^* D_s$	–	4 ± 3
Net Yield (N_S and N_C)	$438 \pm 26 \pm 9$	$44 \pm 13 \pm 1$
$N_s(\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu})$	$417 \pm 28 \pm 14$	–
$N_c(\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu})$	$21 \pm 10 \pm 10$	–

Table 8

A summary of the events in the $B^- \to D^{*0}\ \ell^-\ \bar{\nu}$ Signal Region and Correlated Background Region. In the top 2 rows the first error is statistical and the second is the systematic uncertainty due to the choice of functions used to fit the data. The bottom two rows have been computed using Equation (3.4) and the uncertainties in r_s and r_c have been propagated into the systematic error.

$D^{*0}l^-$	Signal Region	Correlated Background Region
δ_m Signal	$446 \pm 25 \pm 10$	$91 \pm 15 \pm 1$
Scaled δ_m Sideband	$103 \pm 8 \pm 3$	27 ± 6
Uncorrelated Background	12 ± 5	26 ± 8
Continuum	0 ± 5	4 ± 6
Correlated Lepton Fakes	3 ± 1	2 ± 1
$B \to D^*\tau\bar{\nu}$ and $B \to D^* D_s$	–	2 ± 2
Net Yield (N_S and N_C)	$328 \pm 27 \pm 10$	$30 \pm 19 \pm 1$
$N_s(B^- \to D^{*0}\ \ell^-\ \bar{\nu})$	$314 \pm 30 \pm 12$	–
$N_c(B^- \to D^{*0}\ \ell^-\ \bar{\nu})$	$14 \pm 12 \pm 6$	–

3.5.1 Procedures

The reconstruction efficiency for $\bar{B} \to D^* \ \ell \ \bar{\nu}$ events is estimated from a Monte Carlo simulation. The generator produces events obeying all kinematic constraints due to angular momentum conservation and the $V - A$ nature of the pseudoscalar to vector decay.[26] The form factor predictions of Neubert[10,27] are used. These events are passed through a full detector simulation[28] and then analyzed the same way as data.

The efficiencies for $\bar{B} \to D^* \ \ell \ \bar{\nu}$ events to be counted as signal are

$$\epsilon_s(\bar{B}^0 \to D^{*+} \ \ell^- \ \bar{\nu}) = [10.07 \pm 0.24 \pm 0.74]\% \tag{3.6}$$

$$\epsilon_s(B^- \to D^{*0} \ \ell^- \ \bar{\nu}) = [7.28 \pm 0.53 \pm 0.54]\%, \tag{3.7}$$

where the first error is uncorrelated between the two efficiencies and the second is correlated.

3.5.2 Systematic Uncertainties

Theoretical predictions for the unknown form factors of $\bar{B} \to D^* \ \ell^- \ \bar{\nu}$ decays results in uncertainties. The efficiencies with ISGW,[6] BSW,[7] and KS[8] models are compared to that with the Neubert model. The form factor ratios within the Neubert model are also varied to estimate the uncertainties. It is estimated that the model dependence contributes a 3% systematic uncertainty in the reconstruction efficiencies (see Table 10).

Another major uncertainty is from the modelling of the slow pion π_s. Figure 25

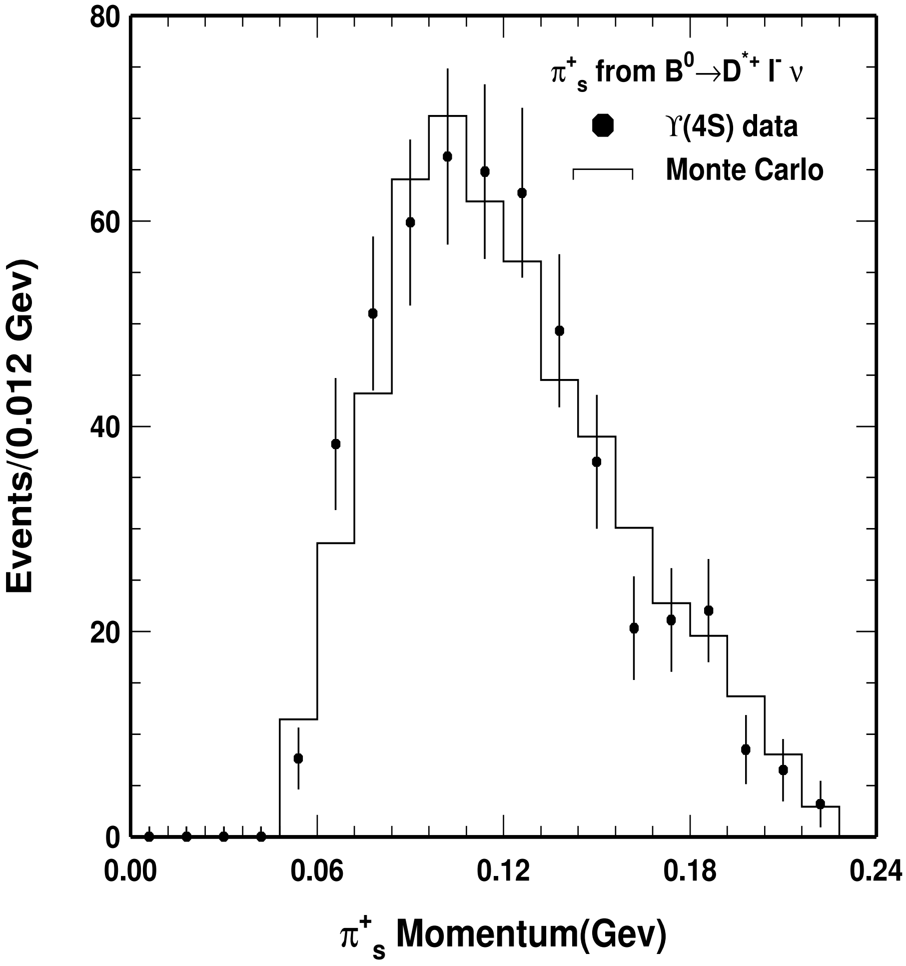

Figure 25
Charged slow pion momentum distribution for $\bar{B}^0 \to D^{*+}\,\ell^-\,\bar{\nu}$ decay. The Monte Carlo simulation is overlayed with arbitrary scale.

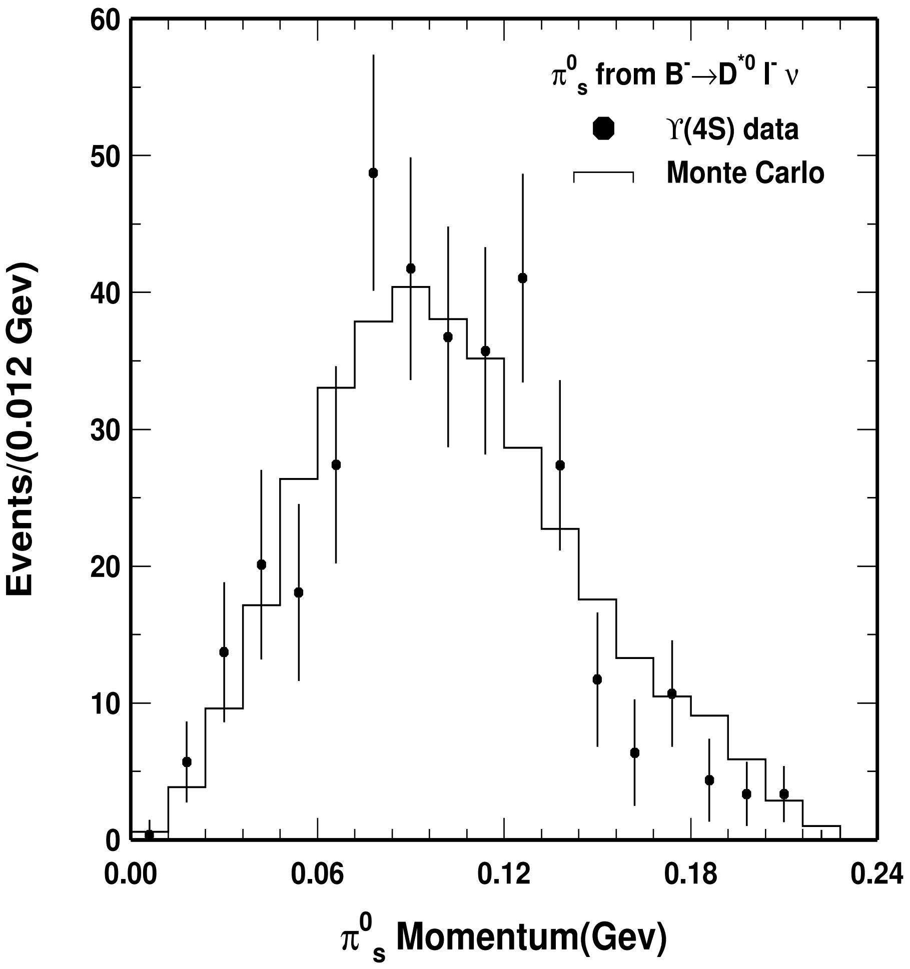

Figure 26
Neutral slow pion momentum distribution for $B^- \rightarrow D^{*0}\ \ell^-\ \bar{\nu}$ decay. The Monte Carlo simulation is overlayed with arbitrary scale.

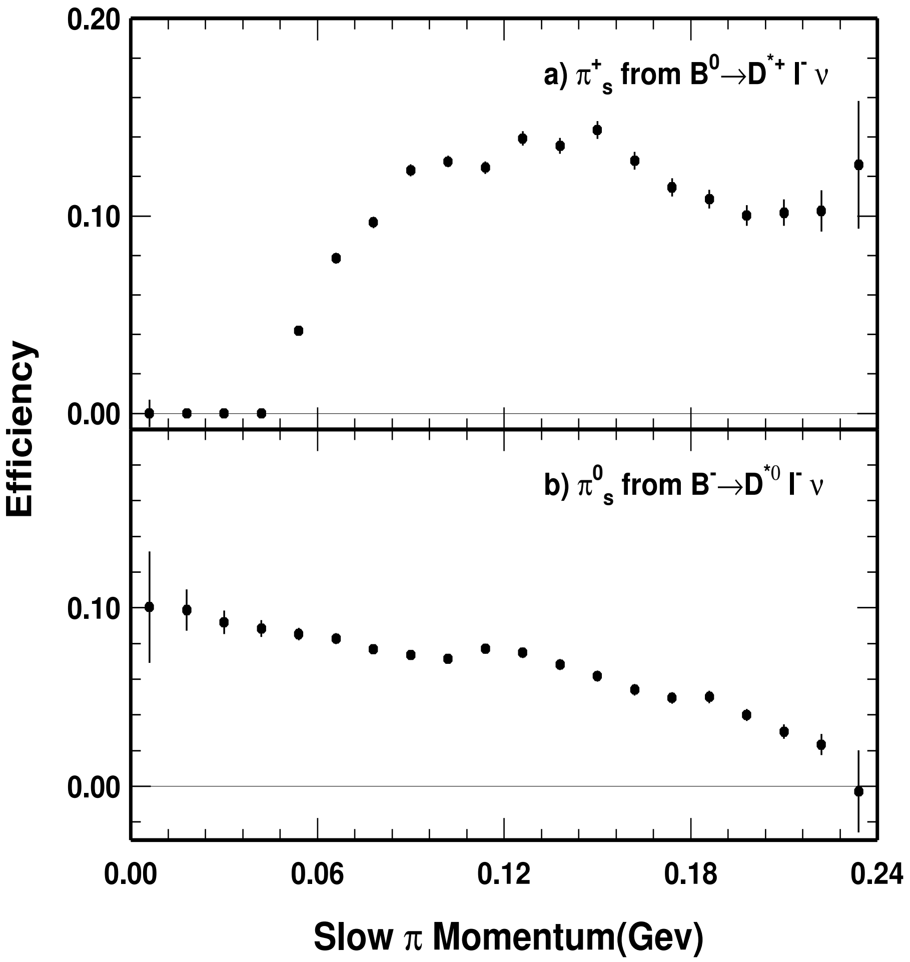

Figure 27
Slow pion momentum distribution of $\bar{B} \to D^* \ \ell^- \ \bar{\nu}$ efficiency.

and Figure 26 show the momentum distribution of charged and neutral slow pions. Part (a) of Figure 27 shows the momentum dependence of the efficiency of π_s^+ in $\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$ events. The spectrum is fairly flat above 100 Mev but decreases rapidly for $|p_{\pi_s^+}| < 100$ MeV. The detector has a poor detection capability for the π_s^+ with momentum below 100 MeV. Part (b) of Figure 27 shows the momentum dependence of the efficiency of π_s^0 in $B^- \to D^{*0}\ \ell^-\ \bar{\nu}$ events. Extensive study[29] has been performed to investigate the uncertainties on reconstruction efficiencies due to π_s. We quote the uncertainty of the efficiency of π_s^+ as 5% and the uncertainty of the efficiency ratio of neutral to charged slow pions as 7%. The errors on efficiencies also include: Monte Carlo statistics, model dependence, an uncertainty reflecting changes in efficiency corrected yields due to variation of cuts, the uncertainty in the normalization of the δ_m sidebands, the uncertainty in the lepton reconstruction and identification and 2% per track uncertainty in the reconstruction of K and π tracks with momenta above 250 MeV. The various contributions to the systematic uncertainty in the efficiency are listed in Table 9.

3.6 Results

With the numbers of reconstructed B mesons and efficiencies from Monte Carlo simulation, the $\bar{B} \to D^*\ \ell\ \bar{\nu}$ branching fractions can be calculated as follows,

$$\mathcal{B}(\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}) = \frac{N_0}{2 \times 2 \times N_{\Upsilon(4S)} f_{00} \mathcal{B}_{D^{*+}} \mathcal{B}_{D^0}} \tag{3.8}$$

$$\mathcal{B}(B^- \to D^{*0}\ \ell^-\ \bar{\nu}) = \frac{N_-}{2 \times 2 \times N_{\Upsilon(4S)} f_{+-} \mathcal{B}_{D^{*0}} \mathcal{B}_{D^0}} \tag{3.9}$$

Table 9

Contributions to the fractional errors of the detection efficiency for the $\bar{B}^0 \to D^{*+}\,\ell^-\,\bar{\nu}$ and $B^- \to D^{*0}\,\ell^-\,\bar{\nu}$ decays. Errors common to both modes are entered only once in the column labeled "Both".

Source	$\Delta\epsilon_s/\epsilon_s$ (%)		
	$B^0 \to D^{*+}\ell^-\nu$	Both	$B^- \to D^{*0}\ell^-\nu$
Form Factors		3.0	
δ_m Sideband Normalization	0.8		1.7
Variation of Cuts	2.3		1.0
Lepton Efficiency		2.2	
$D^0 \to K^-\pi^+$ Efficiency		4.0 *	
Slow π^+ Efficiency		5.0	
Slow π^0/π^+ Efficiency Ratio	–		7.0
Total Uncorrelated	2.4		7.3
Total Correlated		7.4	

$*$ This error is correlated with the systematic error of the CLEO II $D^0 \to K^-\pi^+$ branching fraction.

where

$$\begin{aligned} f_{00} &\equiv \mathcal{B}(\Upsilon(4S) \to B^0\bar{B}^0), \\ f_{+-} &\equiv \mathcal{B}(\Upsilon(4S) \to B^-B^+), \\ \mathcal{B}_{D^{*+}} &\equiv \mathcal{B}(D^{*+} \to D^0\pi^+), \\ \mathcal{B}_{D^{*0}} &\equiv \mathcal{B}(D^{*0} \to D^0\pi^0), \\ \mathcal{B}_{D^0} &\equiv \mathcal{B}(D^0 \to K^-\pi^+). \end{aligned}$$

$N_{\Upsilon(4S)}$ is the total number of $\Upsilon(4S)$ decays in the dataset. One factor of 2 in the denominator comes from the convention of ℓ representing e or μ. The other factor of 2 comes the assumption that each $\Upsilon(4S)$ decays to a $B\bar{B}$ pair.

3.6.1 Product of Branching Fractions

So far two independent efficiency-corrected yields for the $\bar{B} \to D^*\ \ell^-\ \bar{\nu}$ decays have been measured. They are†

$$N_0 = \frac{N_s(\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu})}{\epsilon_s(\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu})} = 4140 \pm 280 \pm 350 \tag{3.10}$$

$$N_- = \frac{N_s(B^- \to D^{*0}\ \ell^-\ \bar{\nu})}{\epsilon_s(B^- \to D^{*0}\ \ell^-\ \bar{\nu})} = 4310 \pm 410 \pm 470. \tag{3.11}$$

With the total number of $\Upsilon(4S)$ decays $N_{\Upsilon(4S)} = (1.65 \pm 0.01) \times 10^6$ events, the two branching fraction products are given directly as:

$$\mathcal{B}(\Upsilon(4S) \to B^0\bar{B}^0)\ \mathcal{B}(\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu})\ \mathcal{B}_{D^{*+}}\ \mathcal{B}_{D^0}$$

†Throughout this paper, the first error is statistical, the second systematic, the third due to B lifetime unless otherwise specified.

$$
\begin{aligned}
&= \frac{N_0}{2 \times 2 \times N_{\Upsilon(4S)}} \\
&= [6.20 \pm 0.42 \pm 0.53]\% && (3.12)
\end{aligned}
$$

$$
\begin{aligned}
&\mathcal{B}(\Upsilon(4S) \to B^+B^-)\, \mathcal{B}(B^- \to D^{*0}\, \ell^-\, \bar{\nu})\, \mathcal{B}_{D^{*0}}\, \mathcal{B}_{D^0} \\
&= \frac{N_-}{2 \times 2 \times N_{\Upsilon(4S)}} \\
&= [6.47 \pm 0.62 \pm 0.72]\% && (3.13)
\end{aligned}
$$

3.6.2 Branching Fractions of $\bar{B}^0 \to D^{*+}\, \ell^-\, \bar{\nu}$ and $B^- \to D^{*0}\, \ell^-\, \bar{\nu}$

To extract the B meson semileptonic branching fractions from Equations (3.8) and (3.9) the charm branching fractions in each product are required.

For the D^* branching fractions the CLEO II results[30] are used.

$$
\begin{aligned}
\mathcal{B}(D^{*+} \to D^0\pi^+) &= [68.1 \pm 1.0 \pm 1.3]\% \\
\mathcal{B}(D^{*0} \to D^0\pi^0) &= [63.6 \pm 2.3 \pm 3.3]\%
\end{aligned}
\qquad (3.14)
$$

For the D^0 branching fraction the CLEO II result[31] is used.

$$
\mathcal{B}(D^0 \to K^-\pi^+) = [3.91 \pm 0.08 \pm 0.07 \pm 0.16]\% \qquad (3.15)
$$

The value $\mathcal{B}(D^0 \to K^-\pi^+)$ being used is the one given in Reference [31] without final state radiation. The systematic error on the K and the π track reconstruction (third error) has been separated from other systematics (second error) because it is correlated with row 5 of Table 9. This correlation is an important advantage of using this CLEO II measurement of $\mathcal{B}(D^0 \to K^-\pi^+)$, as the error due to the K and the π track reconstruction will cancel in the determination of the $\bar{B} \to D^*\, \ell\, \bar{\nu}$ branching fractions. Thus the fractional systematic error on the results for $\mathcal{B}(\bar{B} \to D^* \ \ell \ \bar{\nu})$ will be reduced relative to the error on the product branching fractions.

The branching fractions of $\Upsilon(4S)$ are unmeasured in this work but they are needed to determine the $\mathcal{B}(\bar{B} \to D^* \ \ell^- \ \bar{\nu})$. If the assumption that $\Upsilon(4S)$ decays to $B\bar{B}$ pairs only with equal production of charged and neutral meson pairs is made, i.e., $f_{+-} = f_{00} = 0.5$, we obtain,

$$\mathcal{B}(\bar{B}^0 \to D^{*+} \ \ell^- \ \bar{\nu}) = [4.66 \pm 0.31 \pm 0.41]\%, \tag{3.16}$$

$$\mathcal{B}(B^- \to D^{*0} \ \ell^- \ \bar{\nu}) = [5.20 \pm 0.50 \pm 0.66]\%, \tag{3.17}$$

3.6.3 Partial decay rate of $\bar{B} \to D^* \ \ell^- \ \bar{\nu}$

Based on isospin invariance one can assume $\Gamma(\bar{B}^0 \to D^{*+} \ \ell^- \ \bar{\nu}) = \Gamma(B^- \to D^{*0} \ \ell^- \ \bar{\nu})$, and also assume $f_{+-} + f_{00} = 1$, Equation (3.8) and (3.9) can be combined to give

$$\Gamma(\bar{B} \to D^* \ \ell \ \bar{\nu}) = \frac{1}{4N_{\Upsilon(4S)}\mathcal{B}_{D^0}} \left[\frac{N_0}{\tau_{\bar{B}^0}\mathcal{B}_{D^{*+}}} + \frac{N_-}{\tau_{B^-}\mathcal{B}_{D^{*0}}} \right]. \tag{3.18}$$

Using the average of LEP and CDF measurements for the lifetimes, $\tau_{B^0} = [1.53 \pm 0.09]$ps and $\tau_{B^+} = [1.68 \pm 0.12]$ps,[32] yields,

$$\Gamma(\bar{B} \to D^* \ \ell \ \bar{\nu}) = [30.7 \pm 1.8 \pm 2.8 \pm 2.1] \ \mathrm{ns}^{-1}, \tag{3.19}$$

independent of the ratio f_{+-}/f_{00}.

The systematic errors on all the above results are dominated by the uncertainties in reconstructing the π_s^+ from $D^{*+}(5\%)$ and and the π_s^0 from $D^{*0}(8.6\%)$.

This partial rate will be used to extract the CKM matrix element $|V_{cb}|$ based on theoretical model predictions.

3.6.4 Consistency check on the assumption of $f_{+-}/f_{00} = 1$

The tradition assumption of $f_{+-}/f_{00} = 1$ is uncertain at a $5 - 10\%$ level.[33] With the independently measured $\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$ and $B^- \to D^{*0}\ \ell^-\ \bar{\nu}$ yields, one can perform a check on this assumption.

By taking the ratio of Eq. (3.9) over (3.8) and assuming only that $\Gamma(\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}) = \Gamma(B^- \to D^{*0}\ \ell^-\ \bar{\nu})$, one has

$$\frac{f_{+-}}{f_{00}}\frac{\tau_{B^-}}{\tau_{\bar{B}^0}} = \frac{N_-}{N_0}\frac{\mathcal{B}_{D^{*+}}}{\mathcal{B}_{D^{*0}}} = 1.03 \pm 0.13 \pm 0.13. \qquad (3.20)$$

where the first error is statistical and the second (systematic) error is dominated by the uncertainty on the slow pion efficiency ratio, $\epsilon_{\pi^0}/\epsilon_{\pi^+}$. Substituting the value for $\tau_{B^-}/\tau_{\bar{B}^0}$ from Reference [32],

$$\frac{f_{+-}}{f_{00}} = 0.94 \pm 0.12 \pm 0.12 \pm 0.09, \qquad (3.21)$$

where the third error is due to the input B lifetime ratio. The assumption of $f_{+-}/f_{00} = 1$ and our experimental result are in agreement within errors.

CHAPTER 4

CKM Matrix Element $|V_{cb}|$

4.1 Introduction

Two methods have been used to extract the CKM matrix element V_{cb} in this work. The first method uses the integrated decay rate $\Gamma(\bar{B} \rightarrow D^{*}\ \ell^{-}\ \bar{\nu})$ plus the theoretical model calculation of the hadron form factors. The second method is less model dependent where $|V_{cb}|$ is obtained by fitting the differential decay rate based on HQET.

4.2 Integrated Rate Method

There are several form factor models for exclusive semileptonic B decay that predict the magnitude and shape of the form factors in the decays $\bar{B} \rightarrow D^{*}\ \ell^{-}\ \bar{\nu}$.

Table 10
Model dependent predictions of the $\bar{B} \to D^* \, \ell \, \bar{\nu}$ partial width, detection efficiencies for each model, and $|V_{cb}|$ values derived from the measured partial width. An additional 3.5% systematic error due to B lifetime measurements is common to all the $|V_{cb}|$ values given.

	ISGW[6]	BSW[7]	KS[8]	Neubert[10]
$\Gamma(\bar{B} \to D^* \ell \bar{\nu})/\lvert V_{cb}\rvert^2$	24.6 ps^{-1}	21.9 ps^{-1}	25.8 ps^{-1}	29.0 ps^{-1}
$\epsilon_{\mathrm{model}}/\epsilon_{\mathrm{Neubert}}$	1.00	0.97	0.98	1
$\lvert V_{cb}\rvert \times 10^3$	$35.5 \pm 1.0 \pm 1.6$	$38.2 \pm 1.1 \pm 1.7$	$35.1 \pm 1.0 \pm 1.5$	$32.8 \pm 0.9 \pm 1.4$

Using these model predictions for the form factors and integrating over the whole q^2 space, one can then extract $|V_{cb}|$ using the measured integrated rate $\Gamma(\bar{B} \to D^* \, \ell^- \, \bar{\nu})$ in Equation 3.19. It is important to correct the $\Gamma(\bar{B} \to D^* \, \ell^- \, \bar{\nu})$ for the different efficiencies determined from Monte Carlo simulation of different models. The results for $|V_{cb}|$ with four different form factor models are given in Table 10.

4.3 Differential Rate Method

Another method for extracting $|V_{cb}|$ is to take advantage of HQET. We can expand the unknown Function, $\mathcal{F}(y)$, in Equation 2.14 to first order,

$$\mathcal{F}(y) = \mathcal{F}(1)[1 - \hat{\rho}^2(1 - y)] \tag{4.1}$$

where $\mathcal{F}(1)$ is a theoretically predicted quantity and y is defined in Equation 2.13.

We also define

$$|V'_{cb}| \equiv |V_{cb}|\mathcal{F}(1) \tag{4.2}$$

Then the HQET prediction can be written in a compact form

$$\frac{d\Gamma(\bar{B} \to D^* \ \ell^- \ \bar{\nu})}{dy} = \mathcal{G}(y, |V'_{cb}|, \hat{\rho}^2) \tag{4.3}$$

with y as physical variable and $|V'_{cb}|$ and $\hat{\rho}^2$ as parameters to be determined from the observed y distribution.

The approach adopted in this work is to use an unbinned maximum likelihood method to fit the smeared observed $\tilde{y}$ distribution directly[34] instead of fitting the efficiency corrected observable y distribution (See Appendix A). The equivalent of Equation (4.3) can be written as:

$$\frac{d\Gamma_{obs.}}{d\tilde{y}} = G(\tilde{y}, |V'_{cb}|, \hat{\rho}^2) \tag{4.4}$$

D^* and lepton pairs are selected in the Signal Region of the $|p_{D^*\ell}|$ vs. MM^2 plane as already described. However, since we are now performing an unbinned likelihood fit, we apply a $M_{K\pi}$ cut as well as a δ_m cut. The $M_{K\pi}$ cuts are $|M_{K\pi} - m_{D^0}| < 25$ MeV and $|M_{K\pi} - m_{D^0}| < 20$ MeV for $\bar{B}^0 \to D^{*+} \ \ell^- \ \bar{\nu}$ and $B^- \to D^{*0} \ \ell^- \ \bar{\nu}$ respectively, where m_{D^0} is the nominal D^0 mass. In order to have low background levels in the unbinned likelihood fit, the tighter $B^- \to D^{*0} \ \ell^- \ \bar{\nu}$ cut was chosen to optimize signal to background as much as possible, without introducing a systematic bias from reproduction of signal shapes by the Monte Carlo simulation.

4.3.1 Likelihood Function and Backgrounds

The full likelihood function is

$$
\begin{aligned}
&\mathcal{L}(|V'_{cb}|, \hat{\rho}^2) \\
= \quad & e^{-N(|V'_{cb}|,\hat{\rho}^2)-N_b} \prod_{i=1}^{n} \left(N(|V'_{cb}|, \hat{\rho}^2) \frac{G(\tilde{y}_i, |V'_{cb}|, \hat{\rho}^2)}{\int_1^{\tilde{y}_0} G(\tilde{y}_i, |V'_{cb}|, \hat{\rho}^2) d\tilde{y}} + \sum_b n_b P_b(\tilde{y}_i) \right),
\end{aligned}
\tag{4.5}
$$

with

$$N(|V_{cb}|, \hat{\rho}^2) = 4 N_{\Upsilon(4S)} f \mathcal{B}_{D^*} \mathcal{B}_{D^0} \tau_B \int_1^{y_0} \epsilon(y) \mathcal{G}(y, |V'_{cb}|, \hat{\rho}^2) dy \tag{4.6}$$

$$N_b = n_{\text{comb}} + n_{D^{**}l\bar{\nu}} + n_{\text{uncorr}} \tag{4.7}$$

$$\sum_b n_b P_b(\tilde{y}_i) = n_{\text{comb}} P_{\text{comb}}(\tilde{y}_i) + n_{D^{**}l\bar{\nu}} P_{D^{**}l\bar{\nu}}(\tilde{y}_i) + n_{\text{uncorr}} P_{\text{uncorr}}(\tilde{y}_i). \tag{4.8}$$

$N(|V'_{cb}|, \hat{\rho}^2)$ is the expected number of reconstructed signal events, where $\epsilon(y)$ is the efficiency for detecting $\bar{B} \to D^* \ \ell \ \bar{\nu}$ events determined from the Monte Carlo simulation as shown in Figure 28. τ_B is the B meson lifetime and f (meaning either f_{+-} or f_{00}), $N_{\Upsilon(4S)}$, $\mathcal{B}_{D^*}$ and $\mathcal{B}_{D^0}$ have been defined in Equations (3.8) and (3.9). The total number of background events passing our cuts is N_b. The exponential factor in Equation (4.5) comes from the Poisson probability discussed in the Appendix A.

The backgrounds are represented by Equation (4.7) and Equation (4.8). They have been divided into three classes: combinatorial, $\bar{B} \to D^* \ X \ \ell \ \bar{\nu}$, and uncorrelated background. The normalization and $\tilde{y}$ dependence of the backgrounds is represented by Equation (4.8). The combinatoric background yield and $\tilde{y}$ distribution,

$(n_{\text{comb}} P_{\text{comb}}(\tilde{y}))$, are obtained from sidebands in the δm distribution. The cor-

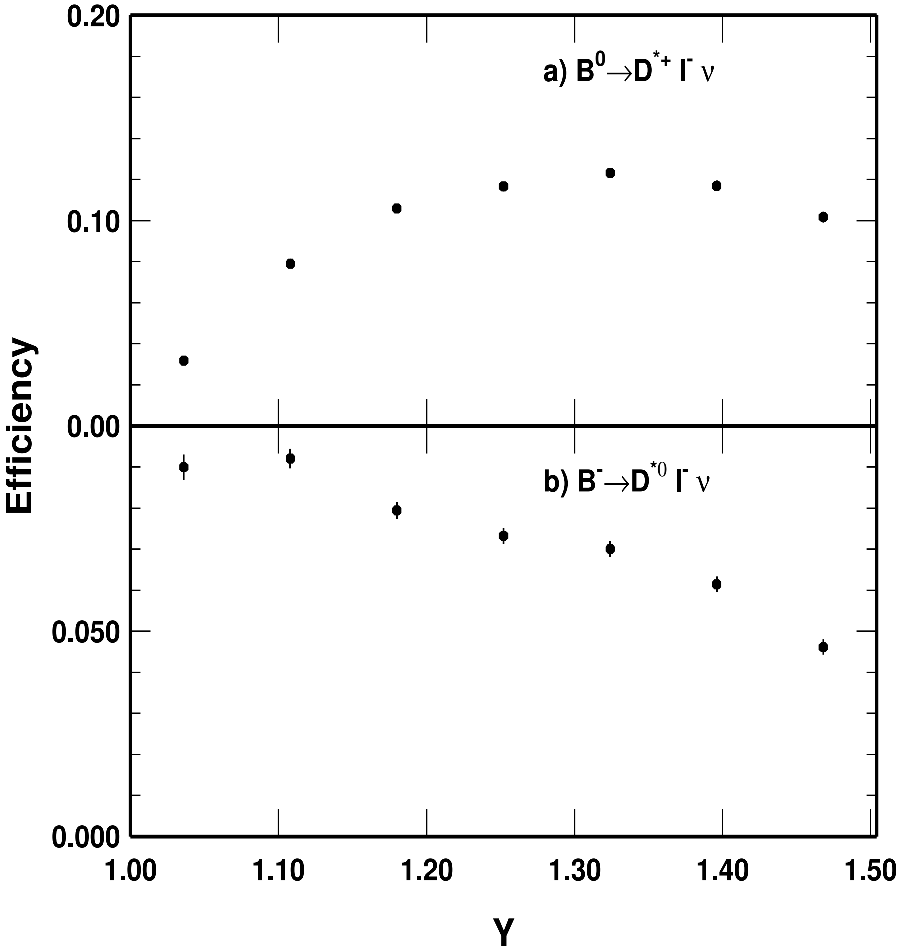

Figure 28
y distributions for $\bar{B} \to D^* \ \ell^- \ \bar{\nu}$ reconstruction efficiencies.

related background yield, $n_{D^{**}l\bar{\nu}}$, is obtained from Table 7 or 8 and its $\tilde{y}$ distribution, $P_{D^{**}l\bar{\nu}}(\tilde{y})$, is taken from $\bar{B} \to D^{**} \ \ell^- \ \bar{\nu}$ Monte Carlo simulation based on the ISGW model. The uncorrelated background yield and $\tilde{y}$ distribution, $n_{\mathrm{uncorr}} P_{\mathrm{uncorr}}(\tilde{y})$ is obtained from the "flipping" technique. All the $P's$ are normalized distribution densities. The three $\tilde{y}$ distributions of the backgrounds to $\bar{B}^0 \to D^{*+} \ \ell^- \ \bar{\nu}$ and $B^- \to D^{*0} \ \ell^- \ \bar{\nu}$ are shown in Figure 29.

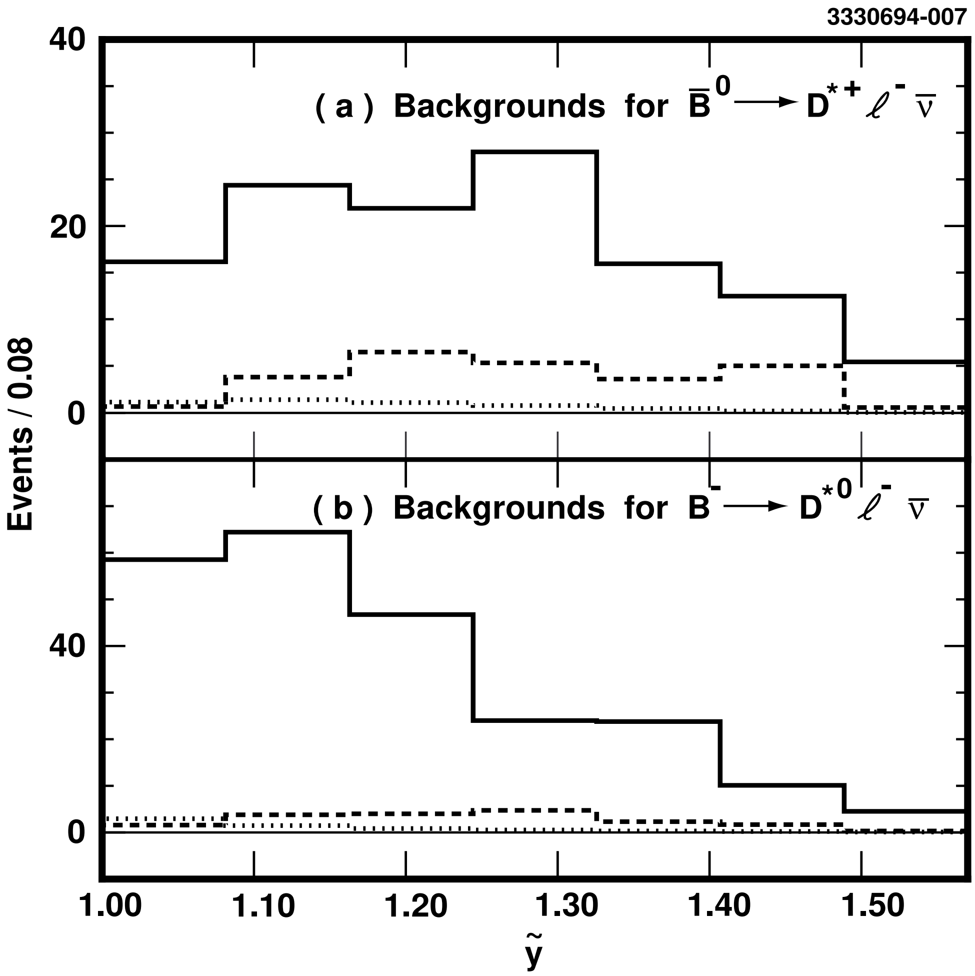

Figure 29

$\tilde{y}$ distributions for various backgrounds in $\bar{B} \to D^* \ \ell^- \ \bar{\nu}$ reconstruction. The Solid histogram represents the combinatoric background. The dashed histogram represents the correlated background. The dotted histogram represents the uncorrelated background. They are normalized to the background levels in the data.

4.3.2 Systematic Uncertainties

The contributions to the systematic error on $|V'_{cb}|$ and $\hat{\rho}^2$ are listed in Table 11. The errors on the absolute efficiencies are the same as for the branching fraction measurements. The errors that are detailed in Table 11 are specific to the $d\Gamma/dy$ fits and account for the uncertainty in the efficiency as a function of $\tilde{y}$ and uncertainties in the fitting procedure.

As a check of the fitting technique we have performed a binned fit to the efficiency corrected $d\Gamma/d\tilde{y}$ distribution, using an analytically smeared form,[29]

$$G'(\tilde{y}, |V'_{cb}|, \hat{\rho}^2) = \frac{1}{2\gamma\beta} \int_{y_{min}(\tilde{y})}^{y_{max}(\tilde{y})} \frac{\mathcal{G}(y, |V'_{cb}|, \hat{\rho}^2)}{\sqrt{y^2 - 1}} dy, \tag{4.9}$$

where $\gamma\beta$ is the known momentum to mass ratio of the decaying B meson and (y_{min},y_{max}) is the range of possible values of y given a value of $\tilde{y}$ (see Figure 30). The results agree well with the results obtained from the unbinned likelihood fit. The systematic uncertainties due to the fitting procedure were estimated by comparing the results of the unbinned and binned methods in 36 Monte Carlo samples, each with the same statistics as our data sample.

We have investigated the uncertainty in the shape of $\mathcal{F}(y)$ using several different functional forms proposed in the literature. The systematic error due to this source is less than 5% in $|V'_{cb}|$ as shown in Table 14.

Table 11

Estimates of the systematic error on V_{cb} and $\hat{\rho}^2$ determined from the fits to the $d\Gamma/dy$ distribution. The error on the absolute efficiency is taken from Table 9, excluding the 4% contribution from $K\pi$ reconstruction, which is also excluded from the error for $\mathcal{B}_{D^0}$.

source	$\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$		$B^- \to D^{*0}\ \ell^-\ \bar{\nu}$		$\bar{B} \to D^*\ \ell\ \bar{\nu}$	
	$\Delta V'_{cb}/V'_{cb}$	$\Delta\hat{\rho}^2/\hat{\rho}^2$	$\Delta V'_{cb}/V'_{cb}$	$\Delta\hat{\rho}^2/\hat{\rho}^2$	$\Delta V'_{cb}/V'_{cb}$	$\Delta\hat{\rho}^2/\hat{\rho}^2$
B lifetime	2.9%		3.6%		2.3%	
Absolute Efficiency	3.3%		4.8%		3.3%	
$\mathcal{B}_{D^*} \times \mathcal{B}_{D^0}$	1.9%		3.4%		2.1%	
Slow π efficiency shape	1.2%	3.5%			0.8%	2.4%
Fitting systematics	3%	9%	3%	9%	3%	9%
Total	5.8%	9.7%	7.5%	9.0%	5.5%	9.3%

4.3.3 Results and discussion

The data yields as a function of $\tilde{y}$ for $\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$ and $B^- \to D^{*0}\ \ell^-\ \bar{\nu}$ events are shown in Figure 31. The values for $|V'_{cb}|$ and $\hat{\rho}^2$ from the fit to these data are given in Table 12. The first error listed in Table 12 is statistical and the second error is an estimate of the systematic error, including the uncertainty in the B lifetimes. The statistical uncertainty in the backgrounds has been taken into account by varying the background levels in the fit by ± 1 standard deviation of their measured central values, and is included in the statistical error.

We have also fit the differential decay distribution of the two decay modes $\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$ and $B^- \to D^{*0}\ \ell^-\ \bar{\nu}$ simultaneously by calculating the likelihood of the two modes separately and maximizing their product. The result of the simultaneous fit is shown in Figure 32, along with the combined data and background level for the $\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$ and $B^- \to D^{*0}\ \ell^-\ \bar{\nu}$ modes. From a linear fit to these data we obtain the most precise measurement to date of $|V'_{cb}|$

and $\hat{\rho}^2$,

$$|V'_{cb}| = 0.0351 \pm 0.0019 \pm 0.0018 \pm 0.0008 \tag{4.10}$$

$$\hat{\rho}^2 = 0.84 \pm 0.12 \pm 0.08, \tag{4.11}$$

independent of f_{+-}/f_{00}; where the first error is statistical, the second is systematic, and the third is due to the input lifetimes.[36] The two parameters are highly correlated as shown in Figure 33. The correlation coefficient $\rho_{|V'_{cb}|,\hat{\rho}^2}$ equals 0.92.

To extract a value of $|V_{cb}|$ one requires a prediction for $\mathcal{F}(1)$. Table 13 gives the values of $|V_{cb}|$ extracted from the fit result using the most updated calculations of the normalization factor $\mathcal{F}(1)$. These results are consistent with previously published values of $|V_{cb}|$ (after scaling for differences in D and D^* branching fractions and B lifetimes) from both inclusive and exclusive analyses, although the unknown model dependence of the previous $|V_{cb}|$ extractions makes detailed comparisons difficult. Model dependent values of $|V_{cb}|$ in Table 10 are also consistent with this result.

The product $|V_{cb}|\mathcal{F}(1)$ as determined from our fits to the combined $\bar{B}^0 \to D^{*+}\,\ell^-\,\bar{\nu}$ and $B^- \to D^{*0}\,\ell^-\,\bar{\nu}$ data is shown in Fig. 34. The dotted lines show the contours for 1σ variations of the fit parameters. All the lines are functions of the variable y, which is equal to the D^* energy to mass ratio, E_{D^*}/m_{D^*}, evaluated in the B rest frame. The points are the data. As discussed before, we do not measure y but rather $\tilde{y}$, which is E_{D^*}/m_{D^*} in the laboratory frame. We have therefore binned the data in the average of the maximum and minimum possible values of y for each measured $\tilde{y}$. This average is denoted by $y_A(\tilde{y})$. Unfortunately $y_A(\tilde{y})$ is not an unbiased estimator of y, and although $y_A(\tilde{y})$ has the same kinematic limits as y, the smearing between these two variables is significant compared to the bin size. Therefore, the data shown this way do not exactly correspond to the product $|V_{cb}|\mathcal{F}(1)$. Note, however, that the lines shown are not the result of a fit to these data points, but were obtained from our unbinned fit in the $\tilde{y}$ space which correctly accounts for the boost between the B frame and the laboratory frame.

4.3.4 Test of Factorization

The parameters determined from the fit to the data can be used to obtain the functional form of $d\Gamma(\bar{B} \to D^* \ \ell \ \bar{\nu})/dq^2$; (where q^2 is the true q^2 of the decay, not the smeared variable measurable experimentally) which is useful to test hypotheses of factorization in B meson decay.[37] Fig. 35 shows $d\Gamma/dq^2$ as extracted from our data (upper solid line) along with the $d\Gamma/dq^2$ distributions derived from several popular models for comparison. The data points are also shown for illustration. As discussed above we do not measure q^2 experimentally and the lines shown are not a result of a fit to the data shown in the figure. Values of $d\Gamma(\bar{B} \to D^* \ \ell \ \bar{\nu})/dq^2$ as determined from the data are given in Table 15 for commonly used q^2 points.

Assuming factorization, it leads to [37]:

$$\frac{\Gamma(\bar{B}^0 \to D^{*+}h^-)}{d\Gamma(\bar{B} \to D^* \ \ell^- \ \bar{\nu})/dq^2|_{q^2=m_h^2}} = R \tag{4.12}$$

where R is calculated by theory.

Also given in Table 15 and Figure 35 are the values for the corresponding hadronic channels. Each of these values is obtained by dividing the experimental

Table 12

Values for $|V_{cb}|\mathcal{F}(1)$ and $\hat{\rho}^2$ determined by the fits to the $d\Gamma/d\tilde{y}$ distributions. The first error is statistical and the second systematic (including B lifetimes). The values for the individual modes, $\bar{B}^0 \rightarrow D^{*+}\ \ell^-\ \bar{\nu}$ and $B^- \rightarrow D^{*0}\ \ell^-\ \bar{\nu}$, assume $f_{00} = f_{+-} = 0.5$. The results of the combined fit (bottom two rows) are independent of this assumption.

Mode	$\lvert V_{cb}\rvert\mathcal{F}(1)\times 10^3$	$\hat{\rho}^2$
$\bar{B}^0 \rightarrow D^{*+}\ \ell^-\ \bar{\nu}$	$34.7 \pm 2.5 \pm 1.8$	$0.80 \pm 0.17 \pm 0.08$
$B^- \rightarrow D^{*0}\ \ell^-\ \bar{\nu}$	$35.7 \pm 2.8 \pm 2.4$	$0.94 \pm 0.20 \pm 0.08$
$\bar{B} \rightarrow D^*\ \ell\ \bar{\nu}$	$35.1 \pm 1.9 \pm 1.9$	$0.84 \pm 0.13 \pm 0.08$

Table 13

Values of $|V_{cb}|$ extracted from our results based on the predictions for the normalization of $\mathcal{F}(1)$ by various models.

Model	$\mathcal{F}(1)$	$\lvert V_{cb}\rvert \times 10^3$
Neubert[12]	0.93 ± 0.03	$37.7 \pm 2.0 \pm 2.1 \pm 1.2$
Shifman et al.[13]	0.89 ± 0.03	$39.4 \pm 2.1 \pm 2.2 \pm 1.3$

value $\Gamma(\bar{B}^0 \rightarrow D^{*+}h^-)$ by the theoretical value R. h^- stands for π, ρ, a_1, etc. The good agreement suggests that the factorization hypothesis works well for B decays within the q^2 range tested by the data.

Table 14

Results for fits to the $d\Gamma/dq^2$ distribution. The curve column indicates which curve in Figure 35 is represented by this function.

$\mathcal{F}(y)/\mathcal{F}(1)$	$\lvert V_{cb}\rvert\mathcal{F}(1)\times 10^3$	$\hat{\rho}^2$	Curve
$exp(-\hat{\rho}^2(y-1))$ ISGW	$36.4\pm 2.2\pm 2.0$	$1.10\pm 0.20\pm 0.10$	dashed
$\frac{2}{y+1}exp(-(2\hat{\rho}^2-1)\frac{y-1}{y+1})$ BSW	$37.2\pm 2.3\pm 2.0$	$1.33\pm 0.25\pm 0.13$	dotted
$(\frac{2}{y+1})^{2\hat{\rho}^2}$ POLE	$36.9\pm 2.3\pm 2.0$	$1.24\pm 0.22\pm 0.12$	dot-dashed
$1-\hat{\rho}^2(y-1)$ Linear	$35.1\pm 1.9\pm 1.9$	$0.84\pm 0.13\pm 0.08$	upper solid
$1-\hat{\rho}^2(y-1)+b(y-1)^2$ Quad.	$35.3\pm 3.2\pm 3.0$	$0.92\pm 0.64\pm 0.40$	lower solid
$b=0.15\pm 1.24\pm 0.90$			

Table 15

Values for the differential decay rate at selected q^2 points. These were calculated from Equation (2.14) with the results of the maximum likelihood fit given in Eqs. (4.10) and (4.11). The first error is statistical, the second is the systematic uncertainty due to $\mathcal{B}_{D^*}\times\mathcal{B}_{D^0}$, the third is the systematic uncertainty due to B lifetimes and the fourth is the rest of the systematic uncertainties as listed in Table 11. The factorization hypotheses values are also listed in column three.

$q^2(\mathrm{Gev}^2)$	$d\Gamma/dq^2$ $(\mathrm{ns}^{-1}\mathrm{GeV}^{-2})$	
	This measurement	Factorization
$m^2_{\pi^+}=0.020$	$1.65\pm 0.23\pm 0.07\pm 0.07\pm 0.26$	1.29 ± 0.33
$m^2_{\rho}=0.59$	$1.90\pm 0.23\pm 0.08\pm 0.09\pm 0.28$	1.33 ± 0.37
$m^2_{a_1}=1.59$	$2.33\pm 0.22\pm 0.10\pm 0.10\pm 0.31$	2.50 ± 0.73
$m^2_{D_s}=3.88$	$3.17\pm 0.16\pm 0.13\pm 0.14\pm 0.34$	
$m^2_{D_s^*}=4.45$	$3.33\pm 0.16\pm 0.14\pm 0.15\pm 0.34$	

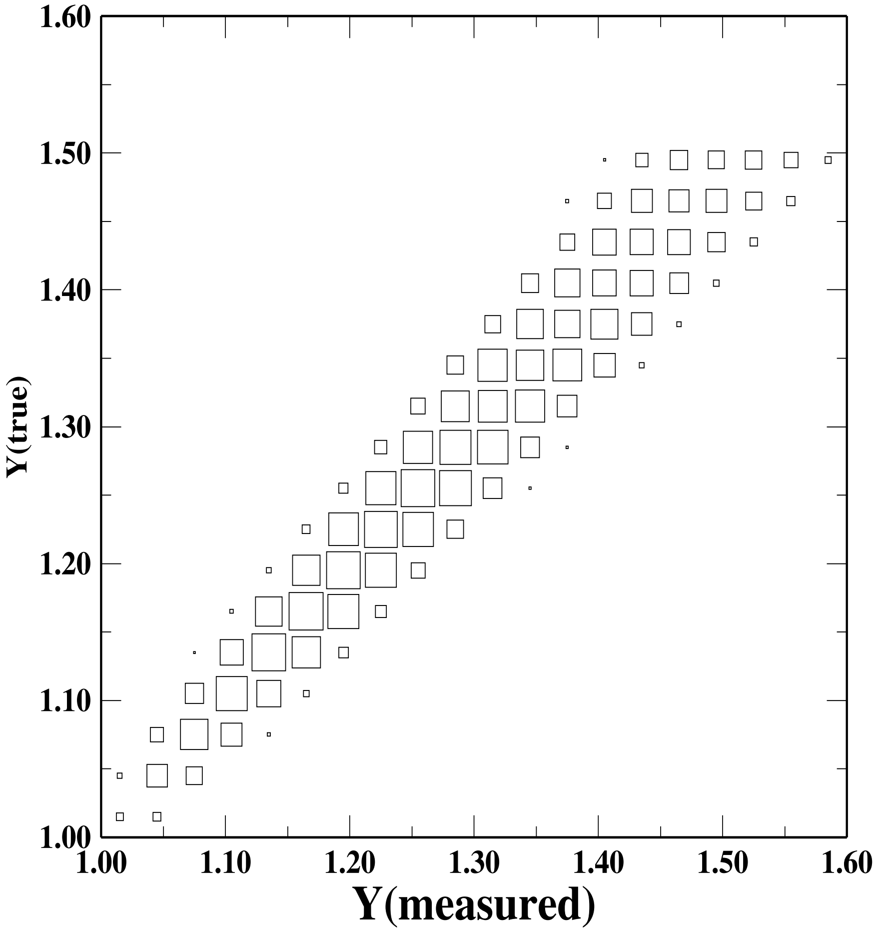

Figure 30
Unsmeared variable y versus smeared variable $\tilde{y}$. The averaged resolution of $\tilde{y}$ is 3%.

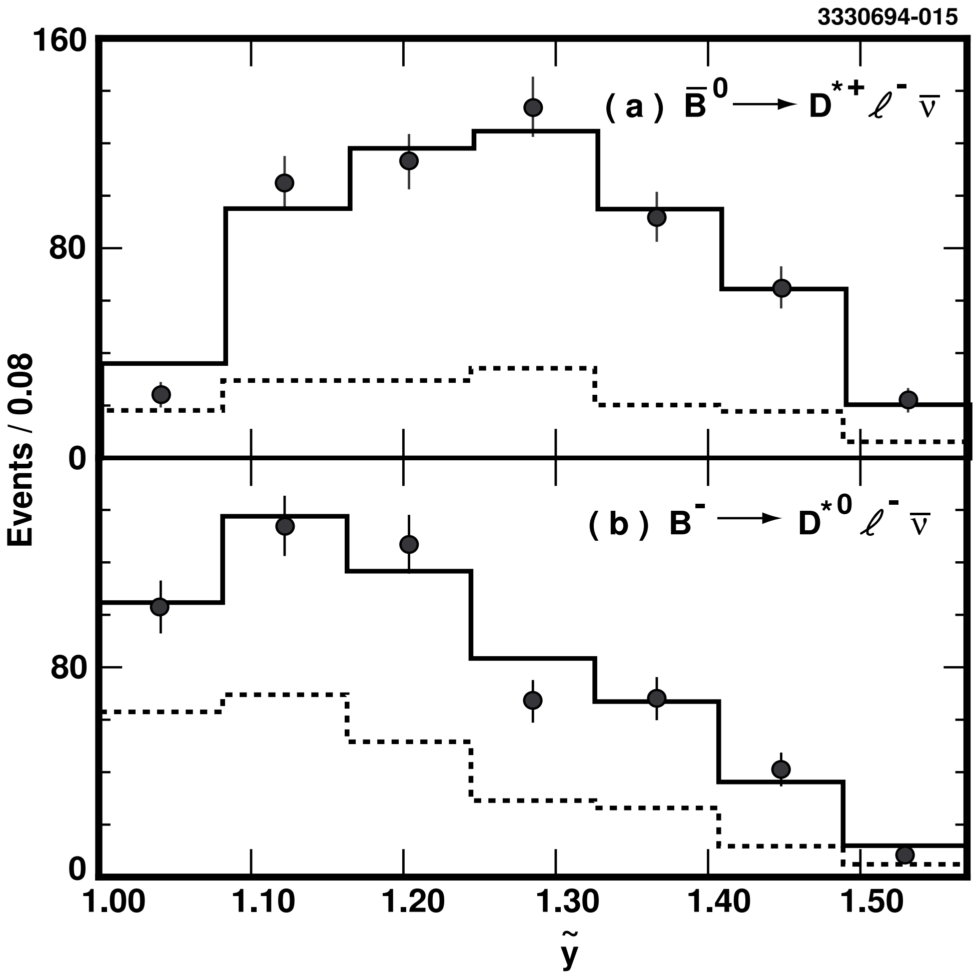

Figure 31

$\tilde{y}$ distributions for $\bar{B}^0 \rightarrow D^{*+}\ \ell^-\ \bar{\nu}$ and $B^- \rightarrow D^{*0}\ \ell^-\ \bar{\nu}$ decays. The dots are data. The dashed histograms are the sum the three background sources. The solid lines are the results of the unbinned maximum likelihood fits.

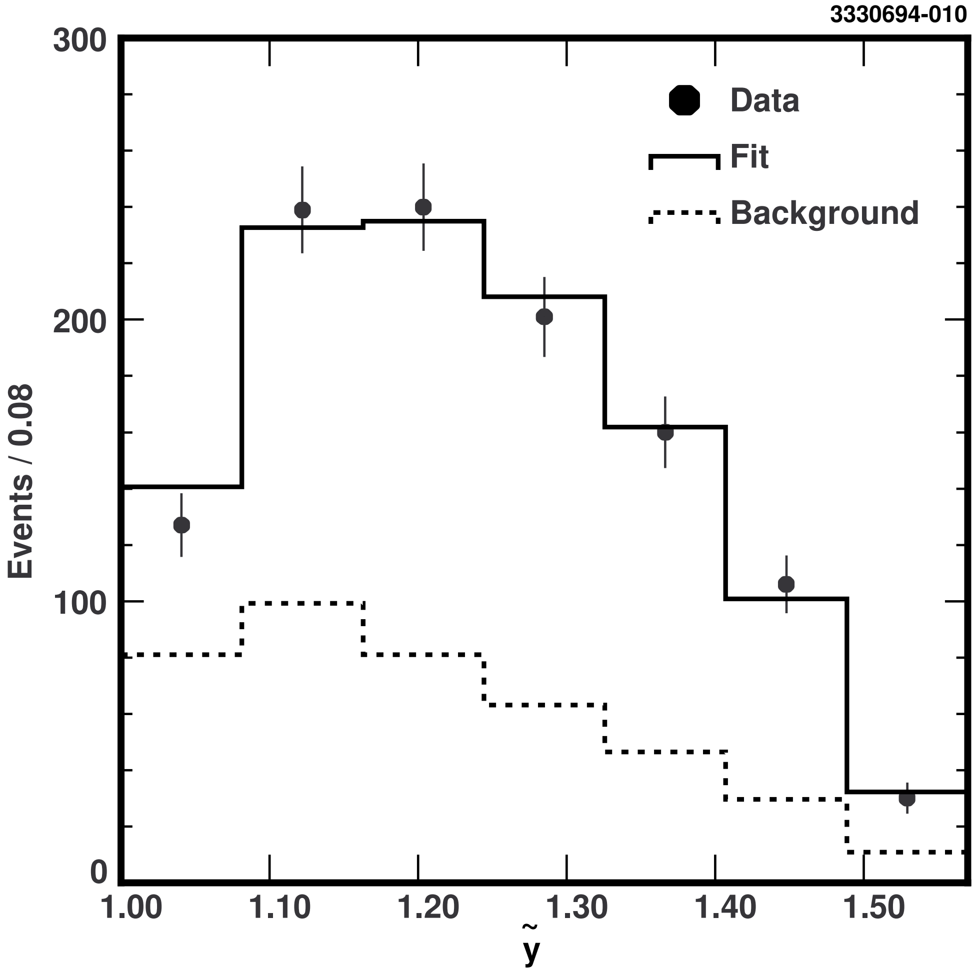

Figure 32
$\tilde{y}$ distributions for the combined $\bar{B}^0 \rightarrow D^{*+}\,\ell^-\,\bar{\nu}$ and $B^- \rightarrow D^{*0}\,\ell^-\,\bar{\nu}$ sample. The dots are data. The dashed histogram is the sum the three background sources. The solid line is the result of performing the unbinned maximum likelihood fit simultaneously on the combined sample.

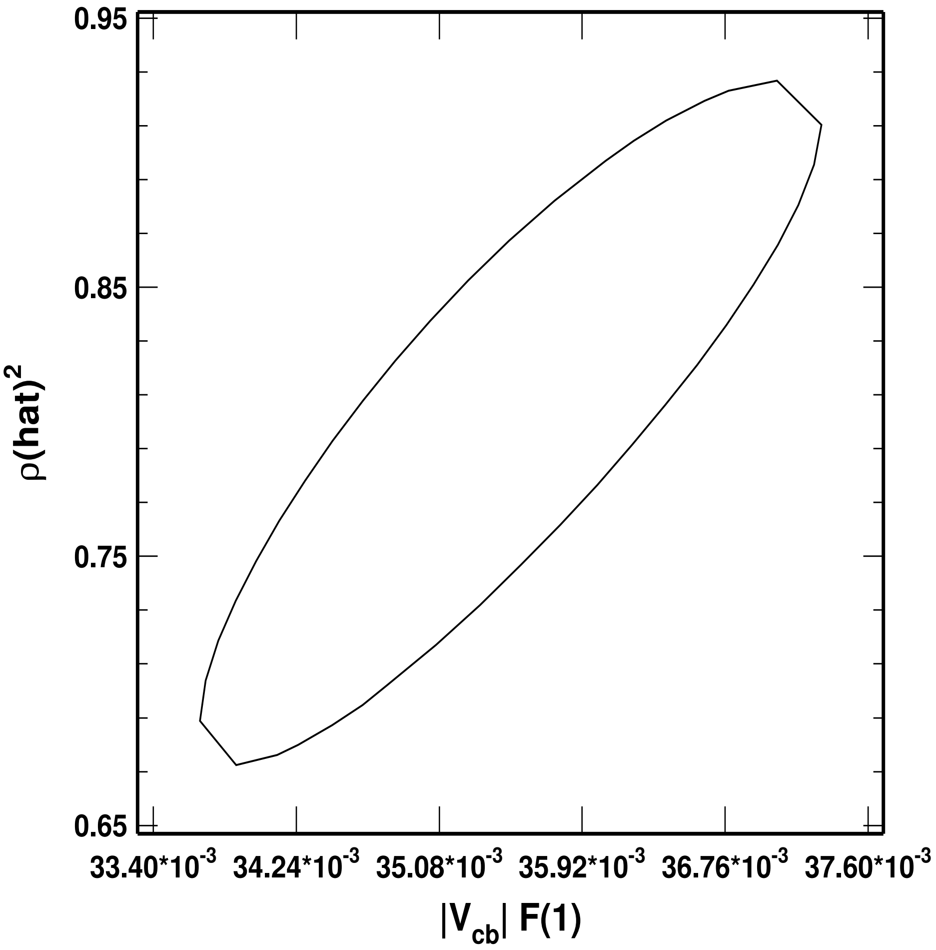

Figure 33

Contour plot of the two highly correlated variables $|V_{cb}|\mathcal{F}(1)$ and $\hat{\rho}^2$. The ellipse represents the points corresponding to likelihood function values being 0.5 away from the maximum value.

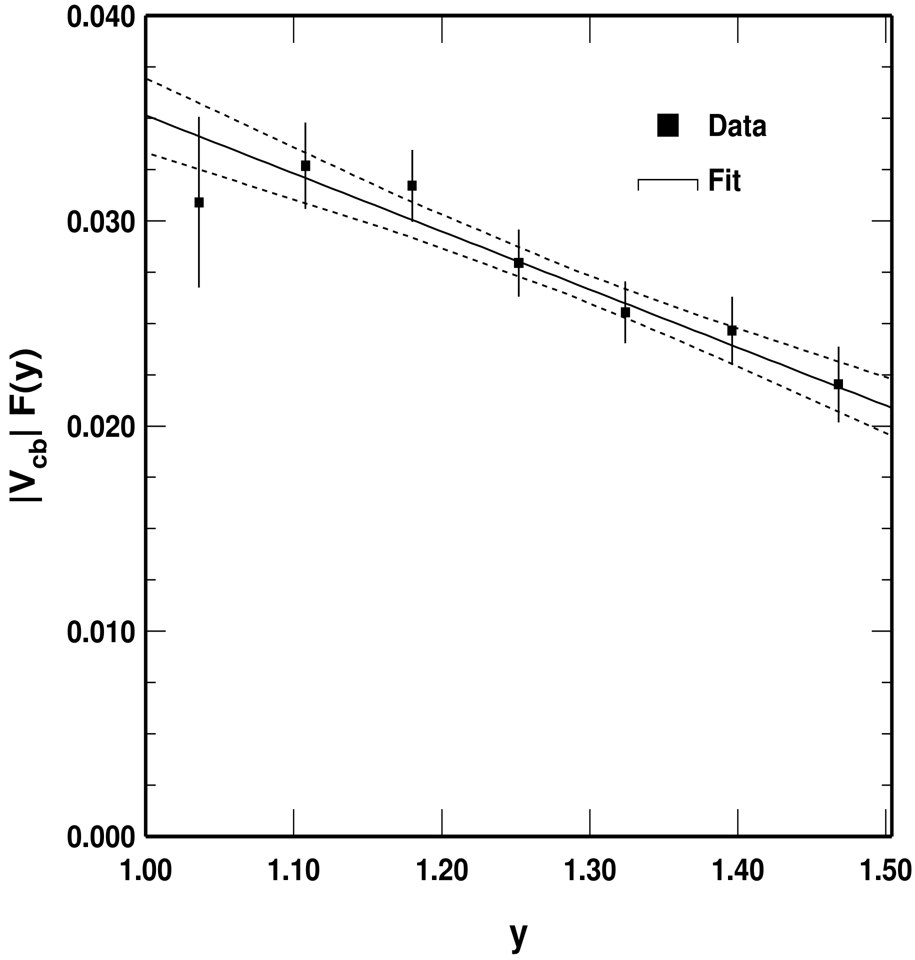

Figure 34
$|V_{cb}|\mathcal{F}(y)$ distribution after efficiency correction on the $\bar{B} \to D^* \ \ell^- \ \bar{\nu}$ data sample. The last data point at y_{max} has been corrected for smearing properly. The dashed lines represent the range of variation of the product $|V_{cb}|\mathcal{F}(y)$ corresponding the 1 σ variation of the fit parameters.

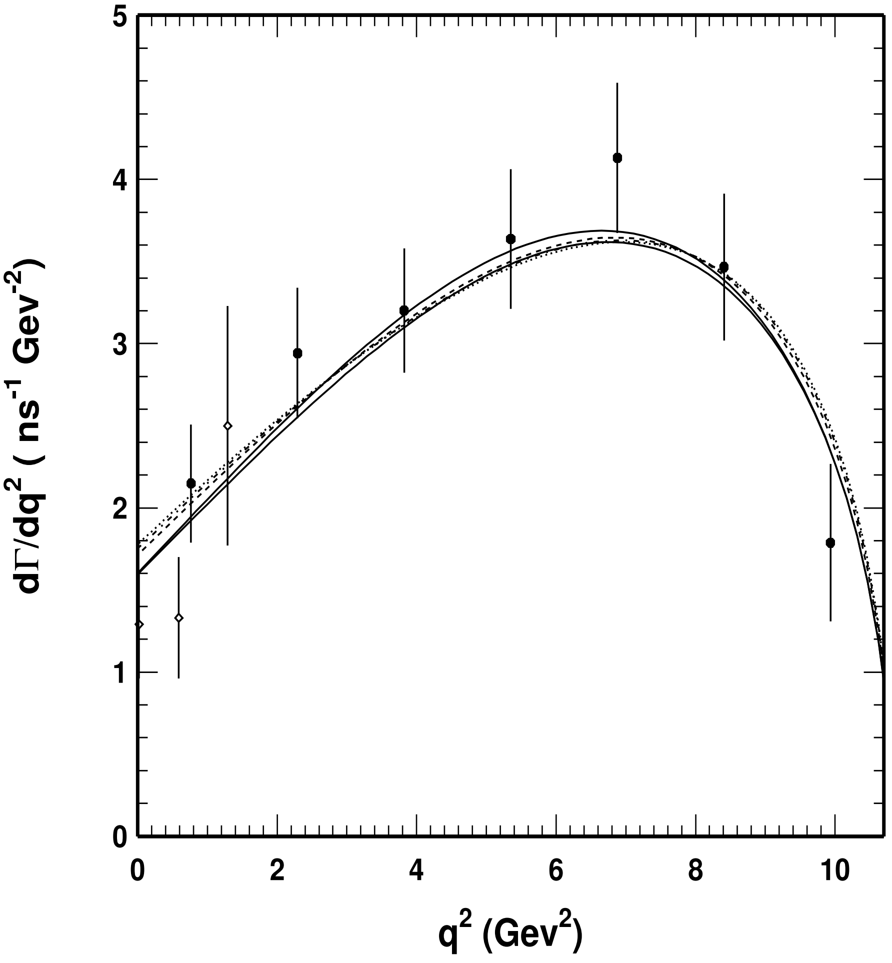

Figure 35

The $d\Gamma/dq^2$ distribution for $\bar{B} \to D^* \, \ell \, \bar{\nu}$ decays. The upper solid line corresponds to the linear fit, and the lower solid line corresponds to the quadratic fit to our data. The dashed line shows $d\Gamma/dq^2$ for the ISGW model, the dotted line is the BSW model, the dot-dashed line is the pole-type $\xi(y)$. The data are shown as solid points. As explained in the text the data are binned in a smeared variable which is not an unbiased estimator of q^2. Therefore the data points do not exactly correspond to the fits shown. The open diamonds are factorization hypotheses values obtained for hadronic channels at the mass squared of π, ρ, a_1, etc. The error bars on these points are the total errors on experimental value $\Gamma(\bar{B}^0 \to D^{*+} h^-)$ and total errors on theoretical value R quadratically combined.

CHAPTER 5

Results and Conclusions

5.1 Branching Fractions

In conclusion, using the CLEO II data sample we have reconstructed both $\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$ and $B^- \to D^{*0}\ \ell^-\ \bar{\nu}$ decay modes in the D^* decay chains: $D^{*+} \to D^0\pi^+$ and $D^{*0} \to D^0\pi^0$ with $D^0 \to K^-\pi^+$. We have measured the exclusive branching fractions,

$$\mathcal{B}(\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}) = (0.5/f_{00})[4.66 \pm 0.31(\text{stat.}) \pm 0.41(\text{sys.})]\% \quad (5.1)$$

$$\mathcal{B}(B^- \to D^{*0}\ \ell^-\ \bar{\nu}) = (0.5/f_{+-})[5.20 \pm 0.50(\text{stat.}) \pm 0.66(\text{sys.})]\%, \quad (5.2)$$

which depend on the relative production of charged and neutral B mesons at the $\Upsilon(4S)$ resonance. Comparison with previously published results in Table 16 shows this measurement provides higher precision and completeness (using both flavored B mesons).

Table 16

Comparison with previously published results for $\mathcal{B}(\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu})$ and $\mathcal{B}(B^- \to D^{*0}\ \ell^-\ \bar{\nu})$. All previous results have been rescaled to use the CLEO II D^0 and D^* branching ratios (except for the CLEO1.5 B^- number which depends nonlinearly on these branching fractions), and use $f_{+-} = f_{00} = 0.5$.

Experiment	$\mathcal{B}(\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu})$ (%)	$\mathcal{B}(B^- \to D^{*0}\ \ell^-\ \bar{\nu})$ (%)
CLEO II ($f_{+-} = f_{00} = 0.5$)	$4.66 \pm 0.31 \pm 0.41$	$5.20 \pm 0.50 \pm 0.66$
ARGUS[38,39]	$4.9 \pm 0.5 \pm 0.6$	$6.4 \pm 1.5 \pm 1.4$
ARGUS[40]	$4.5 \pm 0.3 \pm 0.4$	
CLEO 1.5[41,42]	$4.1 \pm 0.5 \pm 0.6$	$4.1 \pm 0.8 \pm 0.9$

We have determined the product

$$\frac{f_{+-}}{f_{00}} \frac{\tau_{B^+}}{\tau_{B^0}} = 1.03 \pm 0.13(\text{stat.}) \pm 0.13(\text{sys.}), \tag{5.3}$$

and by using existing B lifetime ratio measurements[32] and have solved for $f_{+-}/f_{00} = 0.94 \pm 0.12(\text{stat.}) \pm 0.12(\text{sys.}) \pm 0.09(\text{lifetime})$. This measurement verifies at the 10% level the assumption that $f_{+-} = f_{00}$, widely used to calculate B meson branching fractions with data collected at the $\Upsilon(4S)$ resonance. Also with the assumptions of equal partial widths and $f_{+-} + f_{00} = 1$, and using external B lifetime measurements, we have determined the partial width

$$\Gamma(\bar{B} \to D^*\ \ell\ \bar{\nu}) = [30.7 \pm 1.8(\text{stat.}) \pm 2.8(\text{sys.}) \pm 2.1(\text{lifetime})]\text{ns}^{-1}, \tag{5.4}$$

independent of f_{+-}/f_{00}.

Table 17
Comparison with other $|V_{cb}|\mathcal{F}(1)$ results. All other results were obtained using $\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$ decay and have been rescaled to use $\tau_{B^0} = 1.53 \pm 0.09\ ps$. However, the ARGUS result involves $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^+\pi^-$ modes and the old value of $\mathcal{B}(D^0 \to K^-\pi^+) = 3.65\%$ is still used.

Experiment	$\|V_{cb}\|\mathcal{F}(1) \times 10^3$
CLEO II	$35.1 \pm 1.9 \pm 2.0$
ARGUS[38]	$41.3 \pm 4.6 \pm 2.8$
ALEPH[43]	$39.2 \pm 4.4 \pm 3.5$

5.2 $|V_{cb}|$ Results

Taking advantage of theoretical constraints on the normalization and q^2 dependences of the form factors for the $\bar{B} \to D^*\ \ell\ \bar{\nu}$ decay provided by HQET, a combined fit to the differential decay distributions for $\bar{B}^0 \to D^{*+}\ \ell^-\ \bar{\nu}$ and $B^- \to D^{*0}\ \ell^-\ \bar{\nu}$ results in a determination of $|V_{cb}|\mathcal{F}(1)$,

$$|V_{cb}|\mathcal{F}(1) = 0.0351 \pm 0.0019(\text{stat.}) \pm 0.0018(\text{sys.}) \pm 0.0008(\text{lifetime}), \qquad (5.5)$$

which is also independent of f_{+-}/f_{00}. The explicit dependence on the B^+ and B^0 lifetimes is given in Reference [36]. The form factor normalization $\mathcal{F}(1)$ is believed to be predictable with small theoretical uncertainty, and this permits a precise determination of the CKM matrix element. Using the predictions for the normalization of $\mathcal{F}(1)$,[12,13] we obtain new values for $|V_{cb}|$. The results of $|V_{cb}|\mathcal{F}(1)$ obtained in other experiments are listed in Table 17 for comparison.

5.3 CKM Matrix re-visited

As discussed in Chapter 1, the unitary CKM matrix allows for only four independent parameters. In order to test the CKM hypothesis, it it necessary to make more than four measurements to be able to make an over-constrained fit. So far the elements in the first two rows of CKM matrix have been measured experimentally. Limits on the CKM matrix elements at 90% confidence level are[3]:

$$\begin{pmatrix} 0.9747 \ \mathit{to}\ 0.9759 & 0.218 \ \mathit{to}\ 0.224 & 0.002 \ \mathit{to}\ 0.005 \\ 0.218 \ \mathit{to}\ 0.224 & 0.9738 \ \mathit{to}\ 0.9752 & \mathbf{0.032 \ to\ 0.048} \\ 0.003 \ \mathit{to}\ 0.015 & 0.030 \ \mathit{to}\ 0.048 & 0.9988 \ \mathit{to}\ 0.9995 \end{pmatrix}. \tag{5.6}$$

with the last row derived from unitary. This work has measured $|V_{cb}|$ to be in the range from **0.035 to 0.041** according to the Neubert[12] model as shown in Table 13, which is with a precision higher than that of Particle Data Group value quoted above.

Applying unitary constraints to the first and the third columns yields:

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \tag{5.7}$$

This leads to the so-called "unitary-triangle" which summaries direct and indirect information on the CKM matrix. Using the fact that $V_{ud} \approx V_{tb} \approx 1$ simplifies Equation 5.7 to be: $V_{ub}^* + V_{cd}V_{cb}^* + V_{td} = 0$. Therefore, precision measurement of $|V_{cb}|$ plays an important role in the understanding of the CKM matrix.

Appendix A

Fitting Method

A.1 Introduction

Consider a physical distribution predicted by a theory

$$\mathcal{G}(y, \mu) \tag{A.1}$$

where y is the theory variable and μ refers the physics parameters to be determined. Normally one can utilize the Maximum Likelihood Method to determine the parameters from the distribution of the variable y in data. However, two problems arise for this work.

(1) Acceptance effects and smearing. The CLEO detector has limited acceptance and resolution. The assumption of $|\mathbf{p}_B| = 0$ results in unphysical values of y which correspond to $q^2 < 0$. In experiments one measures the smeared observation variable $\tilde{y}$ that distributes according to a new function $G(\tilde{y}, \mu)$. Both

acceptance effects and smearing effects are complicated.

(2) The standard Maximum Likelihood Method is used to determine the shape of the distribution only. In our case one of the parameters, $|V_{cb}|$, is directly related to the the total number of the events found. So MLM needs to be extended to determine the normalization also.

A.2 Acceptance Effect and Smearing Effect

The detector acceptance function is denoted by $A(\tilde{y})$. The smearing function is denoted by $S(\tilde{y}, y)$. $S(\tilde{y}, y)$ is a many-to-one mapping that relates many theory variables, $y's$, to one observation variable $\tilde{y}$. It is clear that

$$\int S(\tilde{y}, y)\mathcal{G}(y, \mu)dy = \mathcal{G}(y|\tilde{y}, \mu). \tag{A.2}$$

$y|\tilde{y}$ is the variable y that relates to $\tilde{y}$. The observation distribution $G(\tilde{y}, \mu)$ can be written as

$$G(\tilde{y}, \mu) = \int A(\tilde{y})S(\tilde{y}, y)\mathcal{G}(y, \mu)dy. \tag{A.3}$$

Monte Carlo simulation plays an important role. In Monte Carlo events, it's always possible to tag the theory variable y for a given observation variable $\tilde{y}_i$. We will demonstrate how to use Monte Carlo events to calculate the observable level distribution density $G(\tilde{y}, \mu)$. It is beneficial to first examine the definition of $G(\tilde{y}, \mu)$. Suppose that $\tilde{y}$ locates in the neighborhood V_i of width $\Delta\tilde{y}$,

$$G(\tilde{y}_i, \mu) \;\; = \;\; \lim_{V_i \to 0} \frac{N_i}{V_i}$$

$$\begin{aligned}
&= \frac{G(\tilde{y}_i,\mu)\Delta\tilde{y}_i}{\Delta\tilde{y}_i} \\
&= \frac{1}{\Delta\tilde{y}_i}\int A(\tilde{y}_i)S(\tilde{y}_i,y)\mathcal{G}(y,\mu)dy\Delta\tilde{y}_i \\
&= \frac{1}{\Delta\tilde{y}_i}[A(\tilde{y}_i)\mathcal{G}(y_i|\tilde{y}_i,\mu)\Delta\tilde{y}_i] \\
&= \frac{1}{\Delta\tilde{y}_i}[A(\tilde{y}_i)\Delta\tilde{y}_i\mathcal{G}(y_i|\tilde{y}_i,\mu)] \qquad \text{(A.4)}
\end{aligned}$$

The distribution of the reconstructed variable $\tilde{y}$ is just $A(\tilde{y}_i)$ if the generated y distribution in Monte Carlo sample is flat. $[A(\tilde{y}_i)\Delta\tilde{y}]$ means counting the reconstructed events inside volume V_i. $\mathcal{G}(y_i|\tilde{y},\mu)$ is also known. Equation (A.4) says $G(\tilde{y}_i,\mu)$ can be calculated by doing a weighted sum over the reconstructed Monte Carlo events that fall in the volume V_i. The weight on each reconstructed Monte Carlo event is the density $\mathcal{G}(y_i|\tilde{y},\mu)$ evaluated at the corresponding generator variable y. In reality, the Monte Carlo sample is always generated based on a model, say, $\mathcal{G}_{gen}(y,\mu)$. The model effect in $A(\tilde{y},\mu)$ can be removed by replacing $\mathcal{G}(y,\mu)$ with $\mathcal{G}(y,\mu)/\mathcal{G}_{gen}(y,\mu)$.

We can also write,

$$G(\tilde{y}_i,\mu) = \frac{1}{\Delta V_i}\sum_{\tilde{y}_j in V_i} W(\tilde{y}_j,\mu) \qquad \text{(A.5)}$$

where

$$W(\tilde{y}_j,\mu) = \frac{\mathcal{G}(y_j|\tilde{y}_j,\mu)}{\mathcal{G}_{gen}(y_j|\tilde{y}_j,\mu)}$$

Thus, the observation level distribution density $G(\tilde{y},\mu)$ is obtained via Monte Carlo if one has a very large Monte Carlo sample.

A.3 Extended Maximum Likelihood Method

The standard Maximum Likelihood method is well illustrated in some books .[44] Suppose there are n events with one dimensional variable $\tilde{y}$ corresponding to a distribution $G(\tilde{y}_i, \mu)$, $i = 1, 2, ..., n$. μ stands for the set of parameters to be determined. The Likelihood function is defined to be the joint probability:

$$\mathcal{L}(\mu) = \prod_{i=1}^{n} G(\tilde{y}_i, \mu) \tag{A.6}$$

where $G(\tilde{y}_i, \mu)$ is normalized. According to the principle of Maximum Likelihood method maximizing $\mathcal{L}(\mu)$ gives the most probable values for the parameters μ. Practically it is more convenient to minimize the minus Log-Likelihood function $-\log \mathcal{L}(\mu)$.

In this work $|V_{cb}|$ happens to be in the amplitude. The normalization process of $G(\tilde{y}_i, \mu)$ makes the parameter $|V_{cb}|$ drop out. The dependence of expected number of observed events on the parameter μ needs to be utilized. The expected numbers of events can be calculated as

$$N(\mu) = \int G(\tilde{y}, \mu) d\tilde{y} = \int \epsilon(y) \mathcal{G}(y, \mu) dy \tag{A.7}$$

A new joint probability "Extended Maximum Likelihood" $\mathcal{L}_{EML}(\mu)$ can be constructed as

$$\mathcal{L}_{EML}(\mu) = \frac{N(\mu)^n e^{-N(\mu)}}{n!} \prod_{i=1}^{n} G(\tilde{y}_i, \mu). \tag{A.8}$$

Maximizing the new $\mathcal{L}_{EML}(\mu)$ puts a constraint on the expected number of events and returns the amplitude $|V_{cb}|$ we are looking for.

An alternative form the $\mathcal{L}_{EML}(\mu)$ is simply

$$\mathcal{L}_{EML}(\mu) = e^{-N(\mu)} \prod_{i=1}^{n} N(\mu) G(\tilde{y}_i, \mu). \tag{A.9}$$

The factorial $n!$ can be discarded since n is a constant in the maximizing process.

A.4 Background Correction

Generally the experimental data sample contains signal events and background events of various sources. The latter leads to corrections.

For the case that both the number of background events, n_b, and the normalized background distribution density, $P(\tilde{y}_i)$, are known, the background effect can then be corporate into $\mathcal{L}_{EML}$ as

$$\mathcal{L}_{EML}(\mu) = e^{-N(\mu)-n_b} \prod_{i=1}^{n} \; N(\mu) G(\tilde{y}_i, \mu) + n_b P(\tilde{y}_i). \tag{A.10}$$

Note that the whole term after the product sign above is indeed normalized to $N(\mu) + n_b$.

For more than one source of background we have,

$$\begin{aligned} n_b &= \sum_j n_j \\ n_b P(\tilde{y}) &= \sum_j n_j P_j(\tilde{y}) \end{aligned}$$

Table 18
Results of Monte Carlo Tests on volume size.

ΔV_i	$\|V_{cb}\|$	$\hat{\rho}^2$
0.0016	0.0346 ± 0.0024	0.77 ± 0.17
0.004	0.0349 ± 0.0024	0.79 ± 0.17
0.01	0.0348 ± 0.0024	0.78 ± 0.17
0.02	0.0350 ± 0.0024	0.79 ± 0.17

A.5 Fitting Systematics

There are two intrinsic sources of systematic uncertainties for this fitting method.

(1)Non-linearity of $G(\tilde{y}, \mu)$ inside V_i. The approximation of $N_i = G(\tilde{y}_i, \mu)\Delta\tilde{y}$ requires function $G(\tilde{y}, \mu)$ to be linear in the cell. This can be achieved by reducing the volume size of V_i. It is important to insure that the fit results are stable and insensitive to the variation of the volume size.

Some Monte Carlo tests were done. A fixed number of reconstructed Monte Carlo events are used to do the re-weighting with difference volume sizes. Tests are performed on another independent Monte Carlo set used as "mock data".

Results for the tests are listed in Table 18.

The standard deviations of the fit values of the tests are $\sigma_{|V_{cb}|} = 0.0002$, $\sigma_{\hat{\rho}^2} = 0.01$. They are small compared to statistical errors.

(2) Monte Carlo statistics. The size of the Monte Carlo sample used to do the re-weighting is also important in getting accurate fitting results. The numbers of

Table 19
Results of tests on the Monte Carlo statistics.

subset	$\lvert V_{cb}\rvert$	$\hat{\rho}^2$
1	0.0346 ± 0.0024	0.77 ± 0.17
2	0.0347 ± 0.0024	0.77 ± 0.17
3	0.0344 ± 0.0024	0.76 ± 0.17
4	0.0352 ± 0.0024	0.82 ± 0.17

events in the sample should be large enough so that the fit values are stable.

Monte Carlo tests are performed. 8K reconstructed Monte Carlo events are divided in 4 subsets and each subset is used to do the re-weighting independently while the volume size is fixed. Test results are listed in Table 19. The standard deviations from the tests are $\sigma_{|V_{cb}|} = 0.0003$, $\sigma_{\hat{\rho}^2} = 0.03$. They are small compared to statistical errors.

A.6 Monte Carlo Verification Test

It's crucial to know the accuracy and the reliability of the fitter. A study is made to check whether the unbinned Maximum Likelihood method can reproduce the known input $\hat{\rho}^2$. Monte Carlo samples of different $\hat{\rho}^2$ are generated. Each sample after reconstruction is fed into the unbinned fitter to do the fit the same way it would be performed on data.

Table 20 shows the fitting results. The output parameters agree well with input parameters.

Table 20

Results of Monte Carlo verification Test. The input $\hat{\rho}^2 s$ are known generated values. The output $\hat{\rho}^2 s$ are fit values.

input $\hat{\rho}^2$	output $\hat{\rho}^2$
0.85	0.86 ± 0.05
0.95	0.88 ± 0.06
1.09	1.10 ± 0.06
1.22	1.20 ± 0.05
1.32	1.27 ± 0.05
1.43	1.40 ± 0.05
1.57	1.50 ± 0.06

List of References

1. S. Weinberg, A. Salam, and S. Glashow Rev. Mod. Phys. **52**, (1980) 515

2. N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963); M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 653 (1973).

3. Particle Data Group, K. Hikasa *et.al.*, Phys. Rev. D **50**, No. 3, Pt. I (1994).

4. B.D. McDaniel, IEEE Trans. on Nucl. Sci., **NS-28**, 1984 (1981).

5. CLEO collaboration, Y. Kubota *et.al.*, Nucl. Instr. and Meth. A **320**, pp 66-113 (1992).

6. N. Isgur, D. Scora, B. Grinstein, and M. B. Wise, Phys. Rev. D **39**, 799 (1989).

7. M. Bauer, B. Stech, M. Wirbel, Z. Phys. C **29**, 637 (1985).

8. J. G. Körner and G. A. Schuler, Z. Phys. C **38**, 511 (1989).

9. N. Isgur and M. Wise, Phys. Lett. B **232**, 113 (1989); N. Isgur and M. Wise, Phys. Lett. B **237**, 527 (1990); E. Eichten and B. Hill, Phys. Lett. B **234** 511 (1990); H. Georgi Phys. Lett. B **240** 447 (1990).

10. M. Neubert, SLAC Report No. SLAC-PUB-6263 (to appear in Physics Reports).

11. M. E. Luke, Phys. Lett. B **252**, 447 (1990); C. G. Boyd and D. E. Brahm, Phys. Lett. B **257**, 393 (1991).

12. M. Neubert, CERN Preprint CERN-TH.7395/94 (unpublished).

13. M. Shifman, N. Uraltsev and A. Vainshtein, University of Minnesota Report No. TPI-MINN-94/13-T (unpublished).

14. D.J. Borken and S.D. Drell, Relativistic Quantum Mechanics, Mcgraw-Hill, New York, 1984.

15. CLEO collaboration, G. Crawford *et.al.*, Nucl. Instr. and Meth. A **345**, pp 429-439 (1994).

16. D. Besson , CLEO Internal Memorandum *CBX* 92-75 (1992).

17. B. Gittelman *et. al.* , CLEO Internal Memorandum *CBX* 91-82 (1991).

18. D. Bortoletto *et. al.* , Syracuse University Internal Memorandum *HEPSY 5-91*, (1991).

19. H. Kroha, E. Thorndike, C. Wang, CLEO Internal Memorandum *CBX* 92-79 (1992).

20. S. Roberts , CLEO Internal Memorandum *CBX* 93-113 (1993).

21. G. Fox and S. Wolfram, Phys. Rev. Lett. **41**, 1581 (1978).

22. This technique was first successfully used by the ARGUS collaboration in H. Albrecht et al., ARGUS collaboration, Phys. Lett. **B197** (1987) 452.

23. The relative abundances used are $^1P_1$:$^3P_2$:$^3P_1$ = 63:21:16; Daryl J. Scora, Ph.D. Thesis, University of Toronto, 1993 p. 60 (unpublished).

24. P. Heiliger and L. M. Sehgal, Phys. Lett. B **229**, 409 (1989); ALEPH Collaboration, D. Buskulic *et.al.*, Phys. Lett. B **298** 479 (1992).

25. C. P. O'grady, Ph.D. Thesis, Cornell University 1993 (unpublished).

26. F. Gilman and R. Singleton, Jr., Phys. Rev. D **41**, 142 (1990).

27. By the "Neubert model" we mean the estimate of the shape of the form factor $\mathcal{F}(y)$ from Ref.[10]. This involves the use of a theoretically-inspired form for the Isgur-Wise function with one free parameter, $\rho^2 \approx 0.8$, which is estimated using QCD sum rules. This is not regarded as a precision estimate. Note that HQET makes no predictions about the shape of the Isgur-Wise function. The expressions for the A_1, A_2 and V form factors in this model have been provided in Table 5.

28. R. Brun *et.al.*, GEANT v. 3.14, CERN Report No. CERN DD/EE/84-1 (unpublished).

29. M. Garcia-Sciveres, Ph.D. thesis, Cornell University, 1994 (unpublished).

30. CLEO collaboration, F. Butler *et.al.*, Phys. Rev. Lett **69**, 2041 (1992).

31. CLEO collaboration, D. S. Akerib *et.al.*, Phys. Rev. Lett **71**, 3070 (1993). CLEO reports two values for $\mathcal{B}(D^0 \to K^-\pi^+)$, a value obtained using efficiencies from a Monte Carlo the does not include final state radiation and a value obtained using efficiencies from a Monte Carlo that does include final state radiation. We use the former, because our Monte Carlo simulation of D^0 decays did not include final state radiation.

32. W. Venus, in: Lepton and Photon Interactions, XVI International Symposium, eds. Persis Drell and David Rubin, AIP Press, New York (1994). The averages obtained from these measurements and used throughout this paper are:

$$\tau_{B^0} = 1.53 \pm 0.09 \text{ ps}$$
$$\tau_{B^+} = 1.68 \pm 0.12 \text{ ps}$$
$$\tau_{B^+}/\tau_{B^0} = 1.10 \pm 0.11.$$

33. G. P. Lepage, Phys. Rev. D **42**, 3251 (1990).

34. D. M. Schmidt, R. J. Morrison, M. S. Witherell, Nucl. Instr. and Meth. A **328**, 547 (1993).

35. The event selection for the unbinned likelihood fits use $M_{K\pi}$ cuts not present in background estimates listed in Tables 7 and 8. We therefore scale the numbers from these tables according to the change in δ_m sideband subtracted raw yields caused by the $M_{K\pi}$ cuts. This is an effect of order 1%.

36. The result for $|V_{cb}|\mathcal{F}(1)$ from the fit to the combined data scales as $1/\sqrt{\tau_{B^+} + \tau_{B^0}}$, as long as τ_{B^+}/τ_{B^0} remains constant. A change in the lifetime ratio has a non-trivial effect because it modifies the weights of the charged and neutral B components in the $d\Gamma/d\tilde{y}$ fit. By performing different fits with varying lifetime ratios within 2σ of the central value[32] (keeping $\tau_{B^+} + \tau_{B^0}$ fixed) we find, $\Delta|V_{cb}|\mathcal{F}(1) = 0.0012x - 0.003x^2$, where $x = \tau_{B^+}/\tau_{B^0} - 1.1$.

37. CLEO collaboration, M.S. Alam *et.al.*, Cornell preprint CLNS 94/1285(1994) (to be published on Phys. Rev. D).

38. ARGUS collaboration, H. Albrecht *et.al.*, Z. Phys. C **57**, 533 (1993).

39. ARGUS collaboration, H. Albrecht *et.al.*, Phys. Lett. B **275**, 195 (1992).

40. ARGUS collaboration, H. Albrecht *et.al.*, Phys. Lett. B **324**, 249 (1994).

41. CLEO collaboration, D. Bortoletto *et.al.*, Phys. Rev. Lett. **63**, 1667 (1989).

42. CLEO collaboration, R. Fulton *et.al.*, Phys. Rev. D **43**, 651 (1991).

43. I. Scot(ALEPH collaboration), to appear in: Preceeding of the 27th International Conference on High Energy Physics, Glasgow, Scotland, July 1994.

44. P. R. Bevington and D. K. Robinson, Data Reduction and Error Analysis for the Physical Sciences, McGraw-Hill,Inc (1992)

45. ARGUS collaboration, H. Albrecht *et.al.*, Phys. Lett. B **249**, 359 (1990).

46. CLEO collaboration, S. Henderson *et.al.*, Phys. Rev. D **45**, 2212 (1992).

47. CLEO collaboration, J. Bartelt *et.al.*, CLEO Report No. CLEO CONF 93-19 (unpublished).

48. ARGUS collaboration, H. Albrecht *et.al.*, Phys. Lett. B **197**, 452 (1987).

49. We use the symbol ℓ and the term lepton to denote either an electron or a muon.

50. M. Neubert, Phys. Lett. B **264**, 455 (1991).

51. A. Falk, M. Neubert and M. Luke, Nucl. Phys B **388**, 3363 (1992); A. F. Falk and M. Neubert, Phys. Rev. D **47**, 2965 (1993); M. Neubert, Phys. Rev. D **47**, 4063 (1993).

52. T. Mannel, CERN Report No. CERN-TH.7162/94 (unpublished).

53. Throughout this paper, charge conjugate decay modes are implied.

54. We investigated the feasibility of including other channels such as $D^0 \to K^-\pi^+\pi^+\pi^-$ and $D^0 \to K^-\pi^+\pi^0$, but the statistical significance of these modes is small due to the increased combinatoric backgrounds. Additionally, inclusion of these modes introduces new systematic errors, which offset the small statistical gain. Thus the overall significance of our result would not improve with more D^0 decay modes.

55. The calorimeter resolution quoted is for the region $|\cos\theta_\gamma| < 0.71$. The resolution is slightly worse for $0.71 < |\cos\theta_\gamma| < 0.8$

56. CLEO collaboration, D. Bortoletto *et.al.*, Phys. Rev. Lett. **69**, 2046 (1992).

57. CLEO Collaboration, S. Sanghera *et al.*, Phys. Rev. D **47** (1993) 791.

58. J. G. Layter *et al.*, Phys. Rev. D **7**, 2565 (1973).

59. D.Besson and T.Skwarnicki, Annu. Rev. Nucl. Part. Sci. **43** 333-378 (1993).

60. Some previous measurements used the $D^0 \to K^-\pi^-\pi^+\pi^+$ decay channel as well as $D^0 \to K^-\pi^+$, but we have scaled all results by the shift in the $D^0 \to K^-\pi^+$ branching fraction.

61. An additional form factor is required for non-zero lepton mass. The e and μ masses are neglected relative to the B and D^* masses.

62. Neubert writes the form factor $\mathcal{F}(y)$ as $\eta_A\hat{\xi}(y)$, where η_A is a calculable radiative correction for hard gluon processes ($\eta_A = 0.986 \pm 0.006^{10}$) and $\hat{\xi}(y)$ is the non-perturbative part of the form factor.

63. A. G. Frodesen, O. Skjeggestad and H. Tofte, Probability and Statistics in Particle Physics, Universitetsforlaget (1979)