INVESTIGATION OF SEMILEPTONIC B MESON DECAYS TO P-WAVE CHARM MESONS

INVESTIGATION OF SEMILEPTONIC B MESON DECAYS TO P-WAVE CHARM MESONS

By

ALAIN BELLERIVE, Department of Physics

A Thesis
Submitted to the Faculty of Graduate Studies and Research
in Partial Fulfillment of the Requirements
of the Degree of
Doctor of Philosophy

McGill University, Montréal

DOCTOR OF PHILOSOPHY (1997) MCGILL UNIVERSITY
High Energy Physics Montréal, Québec

TITLE: **Investigation of Semileptonic *B* Meson Decays to P-Wave Charm Mesons**

AUTHOR: Alain Bellerive
Department of Physics

SUPERVISOR: Prof. David B. MacFarlane

NUMBER OF PAGES: xvii, 256

Abstract

This thesis presents an investigation of semileptonic B meson decays with a narrow P-wave charm meson in the final state. The data sample consists of 3.29×10^6 $B\bar{B}$ events collected with the CLEO II detector at the Cornell Electron-positron Storage Ring. The P-wave charm mesons are reconstructed in the chain of decays: $D_J^0 \to D^{*+}\pi^-$, $D^{*+} \to D^0\pi^+$, $D^0 \to K^-\pi^+$ or $D^0 \to K^-\pi^+\pi^0$. Study of the decay $B^- \to D^{*+}\pi^-\ell^-\bar{\nu}_\ell$ reveals useful information about the deficit observed in inclusive charm semileptonic B decays and the effective couplings of the W boson to heavy quark mesons. The results obtained for the exclusive semileptonic product branching fractions are $\mathcal{B}(B^- \to D_1^0\ell^-\bar{\nu}_\ell)\mathcal{B}(D_1^0 \to D^{*+}\pi^-) = (0.373 \pm 0.085 \pm 0.052 \pm 0.024)\%$ and $\mathcal{B}(B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell)\mathcal{B}(D_2^{*0} \to D^{*+}\pi^-) < 0.16\%$ (90% C.L.). The assumption $\mathcal{B}(D_1^0 \to D^{*+}\pi^-) = 67\%$ and $\mathcal{B}(D_2^{*0} \to D^{*+}\pi^-) = 20\%$ implies $\mathcal{B}(B^- \to D_1^0\ell^-\bar{\nu}_\ell) = (0.56 \pm 0.13 \pm 0.08 \pm 0.04)\%$ and $\mathcal{B}(B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell) < 0.8\%$ (90% C.L.). These results indicate that at least 18% of the total B semileptonic rate is still unaccounted for by the observed exclusive decays, $\bar{B} \to D^0\ell\bar{\nu}_\ell$, $\bar{B} \to D^*\ell\bar{\nu}_\ell$, $\bar{B} \to D_1\ell\bar{\nu}_\ell$, and $\bar{B} \to D_2^*\ell\bar{\nu}_\ell$. Furthermore, the first measurement of the q^2 spectrum for $B^- \to D_1^0\ell^-\bar{\nu}_\ell$ is presented. The present analysis also suggests that the $\Lambda_{\mathrm{QCD}}/m_Q$ corrections beyond the HQS prescriptions might be significant in the theoretical treatment of the dynamics of B semileptonic decays to excited charm mesons.

Résumé

Cette thèse présente une étude de la désintégration semileptonique du méson B produisant un méson charmé radialement excité dans l'état final. L'échantillon de données contient 3.29×10^6 événements $B\bar{B}$ recueillis par le détecteur CLEO II à l'anneau de collisions CESR à Cornell. Les mésons radialement excités sont reconstruits dans la réaction en chaîne: $D_J^0 \to D^{*+}\pi^-$, $D^{*+} \to D^0\pi^+$, $D^0 \to K^-\pi^+$ ou $D^0 \to K^-\pi^+\pi^0$. L'étude de la désintégration $B^- \to D^{*+}\pi^-\ell^-\bar{\nu}_\ell$ révèle de l'information pertinente au sujet du déficit observé dans les désintégrations inclusives du méson B ainsi que sur le couplage effectif entre le boson W et les mésons formés de quarks lourds. Les résultats obtenus pour les produits de rapports d'embranchement exclusifs sont $\mathcal{B}(B^- \to D_1^0\ell^-\bar{\nu}_\ell)\mathcal{B}(D_1^0 \to D^{*+}\pi^-) = (0.373 \pm 0.085 \pm 0.052 \pm 0.024)\%$ et $\mathcal{B}(B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell)\mathcal{B}(D_2^{*0} \to D^{*+}\pi^-) < 0.16\%$ (90% C.L.). La supposition $\mathcal{B}(D_1^0 \to D^{*+}\pi^-) = 67\%$ et $\mathcal{B}(D_2^{*0} \to D^{*+}\pi^-) = 20\%$ implique $\mathcal{B}(B^- \to D_1^0\ell^-\bar{\nu}_\ell) = (0.56 \pm 0.13 \pm 0.08 \pm 0.04)\%$ et $\mathcal{B}(B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell) < 0.8\%$ (90% C.L.). Ces résultats indiquent qu'au moins 18% du taux de désintégration semileptonique du méson B provient d'autres sources que les désintégrations exclusives antérieurement observées, $\bar{B} \to D^0\ell\bar{\nu}_\ell$, $\bar{B} \to D^*\ell\bar{\nu}_\ell$, $\bar{B} \to D_1\ell\bar{\nu}_\ell$, et $\bar{B} \to D_2^*\ell\bar{\nu}_\ell$. De plus, la première mesure du spectre q^2 pour $B^- \to D_1^0\ell^-\bar{\nu}_\ell$ est présentée. Cette analyse suggère que les corrections Λ_{QCD}/m_Q au delà des prescriptions de HQS peuvent être significatives pour un traitement théorique adéquat de la dynamique des désintégrations du méson B en mésons charmés excités.

Acknowledgments

This dissertation builds upon the careful and detailed work of many people. First and foremost, I wish to thank David Cassel and Persis Drell of Cornell University for their constant guidance and motivation. They were always available for long discussions and keen insights concerning my work. I could simply say that it was (and still is) a pleasure to interact with David and Persis. I will always be grateful for their great expertise in my thesis analysis. I am also grateful to my thesis supervisor David MacFarlane for his advice and support. He helped make my graduate career as enjoyable and rewarding as possible.

With great pleasure, I acknowledge the encouragement provided by early mentors: Claude Leroy (Université de Montréal) and Jean-Claude Baril (CEGEP de Trois-Rivières). I also express my gratitude to (in alphabetical order): Roy Briere, Dave Cinabro, Lawrence Gibbons, René Janicek, David Kreinick, Rob Kutschke, David London, Ken McLean, Scott Menary, Popat Patel, Ritchie Patterson, Daniel Peterson, Chris O'Grady, Jeffrey Richman, Anders Ryd, and Ed Thorndike for many useful discussions. Special thanks to my paper committee for their patience and wisdom: Alan Weinstein (chair), John Bartelt, Karl Berkelman, and Anders Ryd.

It has been a pleasure to work with my fellow officemates René Janicek and Ken McLean. Their help and warm friendship deserve recognition. Special cheers for their review of this thesis.

Anders Ryd, David Lange, and Véronique Boisvert are to be credited for their effective work in writing and testing the Monte Carlo generator EvT used in this analysis. Thanks to Suzanne Jones and Mauricio Garcia-Sciveres for letting me use part of the tracking CBX for the calibration chapter. Special thanks to Suzanne for her great help in both my tracking duties and my physics analysis.

I wish to thank the CESR staff and all the members of the CLEO collaboration. Through their combined efforts and cooperation they have made the CLEO experiment a very active and internationaly well-recognized experiment. In particular, I gratefully acknowledged the help of the tracking group and the B-to-lepton PTA working group during my stay at Cornell. As a guest for several years at Cornell University, I thank the staff of Wilson and Newman Laboratories for their hospitality.

I would also like to express my gratitude to the Department of Physics at McGill University, le Fonds Québécois pour la Formation de Chercheurs et l'Aide à la Recherche, and the Natural Sciences and Engineering Research Council of Canada for financial support.

Overall, graduate school was a wonderful experience. It would have been impossible, though, without the support of my relatives and all my friends. I say a very special *merci* to my mother, father, brother, and sister who, together with Jean and Denise, have been an inexhaustible source of motivation over the years.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The origin of our material world has always been of fundamental interest to mankind. Throughout the millennia, philosophers and scientists have searched for a better understanding of the visible transformations of nature. In the extraordinarily diverse phenomena of nature, the questions about what matter is made of and how it is bound together have occupied the thoughts of those who are now called physicists.

The description and interpretation of experimental observations of physical systems constitute much of a physicist's task. To probe the physical phenomena of our universe, we must deal with concepts such as motion, particles, forces, fields, and symmetries. Theoretical models use such concepts in order to correlate, understand, and explain experimental observations and quantitative measurements. There would be no physical science without theories based on laws fundamental enough to predict the behavior of physical systems. Whenever a discrepancy arises between theory and experiment, new theories and concepts must be worked out to explain the discrepancy. Development of new experiments and formulation of new theories are at the forefront of the scientific enterprise.

The bulk of this dissertation will be concerned with the study of modern physics; more precisely, the description of the dynamics of elementary particles and the fundamental forces that govern them. The remainder of this chapter provides a brief overview of the development of ideas and concepts in particle physics. The Standard Model is introduced, followed by a more detailed outline of this thesis.

1.1 Fundamental Forces of Nature

In the modern study of the fundamental constituents of matter and their interactions, the Standard Model (SM) attempts to explain all the phenomena of particle physics in terms of three distinct types of elementary particles. The first two are the leptons and the quarks, both spin-$\frac{1}{2}$ fermions. The other is the gauge bosons, carriers of four distinct types of fundamental forces:

Gravity Gravity governs the attraction between two massive objects. Its range is infinite. The mediator of the gravitational interaction is, in theory, a spin-2 boson: the graviton. The effects of the gravitational force are well understood in macroscopic systems such as the solar system. However, due to the small mass of elementary particles, the gravitational interaction between them is negligible compared to the other three forces of nature. For the description of interactions between elementary particles, we shall therefore focus on the electromagnetic, strong, and weak forces.

Electromagnetic Force Like the gravitational force, the electromagnetic (EM) force was first observed as an extranuclear phenomenon. Most of us are familiar with electric and magnetic phenomena. The massless photon is the mediator of the EM interaction which governs the attraction and repulsion between charged objects. Consequently, the EM force is responsible for the bound state of the electrons in the atoms and for the arrangement of atoms in molecules. The range of the EM force is infinite.

Strong Force In the Standard Model, hadrons (like neutrons and protons) are considered to be made of quarks bound together by the strong force. The carriers of the strong interaction are called gluons and are massless spin-1 bosons like the photon. The gluons have no electric charge but couple to their own color charge and that of the quarks. The strong interactions between quarks have in principle infinite range because the gluons have zero mass. The residuals of the strong interactions between quarks give rise to the nuclear force which holds together the nucleons. For this reason, the effect of the strong force between hadrons is of short range and vanishes when the colorless hadrons are far apart.

Weak Force The weak interaction is more subtle because of its very short range (10^{-18} meters). It is responsible for the instability of some nuclei via β-decay. In fact, the weak interaction leads to the disintegration of a neutron to a proton, an electron, and a massless neutral lepton named the neutrino. Weak decays are slow processes when compared to typical EM and strong decays. The weak interaction is associated with the exchange of massive bosons called the W and Z bosons. The short range nature of the weak force arises because its mediators are massive. In the Standard Model of elementary particles, the electromagnetic and the weak interactions are unified in the so-called electroweak interaction.

The goal of elementary particle physics is to unravel the properties of matter at the deepest level [1]. Investigations at the subnuclear scales (below 10^{-15} m) requires high energy particles and, consequently, the development of particle accelerators for producing such particles. Modern particle accelerators create, for very short times, an energy density thought to have prevailed in an early stage of the universe, when the fundamental particles and the forces that govern them were beginning to form. Table 1.1 lists the fundamental interactions of relevance in particle physics. Particle experiments enable physicists to study the fundamental building blocks of our universe in the hope of gaining a deeper understanding of the ultimate constituents of matter.

Interaction	Particle	Range (m)	Coupling
Electromagnetic	γ	∞	10^{-2}
Strong	g	10^{-15}	1
Weak	W^+, W^-, Z^0	10^{-18}	10^{-6}

Table 1.1: The fundamental interactions of relevance in particle physics, the particles that mediate them, their range, and their relative coupling strength.

In the next sections, a history of the discovery of the known elementary particles and of the development of the Standard Model is given. The reader can find a full chronological list of papers concerning the major discoveries in the field of particle physics in Reference [2].

1.2 History of Particle Physics in the 20^{th} Century

The ancient Greek philosophers were the first to talk about the concept of elementary building blocks of matter. The earliest of these philosophers were concerned with the natural world and its processes. Democritus, one of the last great natural philosophers, agreed with his predecessors that transformations in nature could not be due to the fact that anything is actually changing. He assumed that everything was built up of tiny invisible blocks which were immutable and eternal. He called them "a-toms", which meant "un-cuttable". The Greek theory of matter remained a philosophical detail until the first modern atomic theory was developed in the early nineteenth century by J. Dalton. After concentrating on the idea of gases as consisting of particles, Dalton was led to the assumption that the free particles (atoms, as he called them) are all alike. To Dalton, as for Democritus, the atom was indivisible and was therefore an "elementary particle", and every substance in the universe was made up of a different combination of a few different kinds of atoms. Consequently, the theoretical paradigm of the late 19^{th} century was not adequate to prepare chemists and physicists for the discoveries that took place between 1890 and 1920.

In 1887, J. J. Thomson gave proof of the independent existence of a negatively charged particle. Thomson applied rudimentary electromagnetic laws, developed in the mid-1800's, to observe what was then called the electron. Following this discovery, it was clear that a revision of the atomic theory was required. The atom could no longer be regarded as the ultimate unit of matter. The next attack on the structure of the atom came from a rapid succession of experimental discoveries: Röntgen discovered X rays (1895) and Becquerel discovered radioactivity (1896). In 1902, Lord Kelvin and J. J. Thomson proposed the spherical atom which is essentially a blob of positively and negatively charged matter. At the same time, Planck introduced the novel idea of quanta. This was followed in 1905 by the photon theory of Einstein (accepted slowly even by Planck) to describe the photoelectric effect. In 1911, Ernest Rutherford had worked out most of the details of his theory of scattering of α particles by an atom. He presented the atom as a small, massive, dense, positively charged nucleus surrounded by a cloud of light electrons. The modern atomic model was born.

Applying the idea of the quantum to the Rutherford model, Bohr (in 1913) explained the hydrogen spectrum with admirable precision. This discovery was the starting point for tumultuous developments in atomic physics that culminated in the late 1920's with the establishment of quantum mechanics by De Broglie, Heisenberg, Schrödinger, Pauli, and Dirac.

Before 1930, the only elementary particles known were the electron and the proton (hydrogen nucleus). In 1932, J. Chadwick discovered the neutron in studying $\alpha\, Be \rightarrow$ nucleus n. Soon afterward, C. D. Anderson found the positron in cosmic ray photographs from cloud chambers. As measured by Chadwick, the mass of the neutron turned out to be somewhat greater than the combined mass of the proton and the electron. This is responsible for the most striking property of the neutron: its instability. The study of the nuclear β decay, and thus the instability of the neutron, led later to the theory of the weak interaction. In fact, conservation of energy rules out the simple scheme $n \rightarrow p + e$. One of the consequences is the need to introduce an extra fundamental particle, the neutrino. Such a particle was postulated by Pauli to explain spin balance and energy conservation of nuclei involved in nuclear β decay.

The same year, Fermi also developed a universal quantitative theory of the emission of β rays by neutrons and some nuclei. Fermi proposed a contact interaction (or a current-current invariant amplitude) between the four fermions present in the β decay. The following year, Yukawa proposed the idea of heavy quanta to mediate forces between elementary particles. An attempt to build a universal $V - A$ (Vector-Axial vector) weak interaction was finally made by Feynman and Gell-Mann in 1958.

In the 1940's, Tomonaga, Feynman, and Schwinger developed a quantum field theory of electrodynamics (QED), which is a cornerstone in the analysis of elementary particle interactions. In the process, Feynman developed a pictorial technique to describe interactions of elementary particles: Feynman diagrams.

Other elementary particles discovered in the period between the late nineteen thirties and the early nineteen fifties include the muon (1936), the π meson (1947), the kaon (1951), and several baryons (Λ, Σ and Ξ). In 1955, the antiproton was produced as predicted by the Dirac theory of antimatter for spin-$\frac{1}{2}$ particles.

In principle, the electron neutrino can be detected by observing the inverse β processes such as $\nu_e + n \to e^- + p$ and $\bar{\nu_e} + p \to e^+ + n$. Reines and Cowan used an intense antineutrino flux from a nuclear reactor and claimed the discovery of the invisible particle in 1959. In 1962, a group at the Brookhaven National Laboratory (BNL) proton synchrotron found evidence for more than one kind of neutrinos. With a neutrino beam ($p\,Be \to \nu_\mu\,X$) they discovered the muon neutrino ($\nu_\mu\,Al \to \mu^\pm\,X$) and demonstrated the doublet structure of the leptons.

Back in the 1950's, Lee and Yang proposed that parity might not be conserved in weak interactions, and almost immediately experimental tests uncovered violations of parity in weak decay processes. In 1964, CP violation in kaon decays was also discovered. CP violation is the lack of invariance of the amplitude of a physical process under the combined operation of charge conjugation (C) and parity (P).

By the mid nineteen sixties many hadrons (mesons and baryons) had been discovered. The Cabibbo angle had already been introduced to predict the semileptonic decay rate of the hyperon. In 1964, Gell-Mann and Zweig independently came up with a unifying theoretical framework that explained the multitude of hadronic states: observed hadrons could be interpreted as bound states of just three fundamental spin-$\frac{1}{2}$ particles, together with their antiparticles. These particles were called quarks and were required to have fractional electric charge Q/e of $+\frac{2}{3}$, $-\frac{1}{3}$, and $-\frac{1}{3}$. They were called the up, down and strange quark respectively. Using group theory, one can build mesons ($q\bar{q}$) and baryons (qqq), and predict the mass splitting between different members of a given supermultiplet. The mass of the sss state (the Ω^- baryon) was estimated to be about 1.68 GeV/c^2 even before it was discovered at Brookhaven. During the same time period, Greenberg resolved the apparent contradiction between the quark model and the Pauli principle (for spin-$\frac{1}{2}$ fermions) by adding a quantum number called color for the quarks. To find experimental evidence for the fundamental quarks within hadrons, Friedman, Kendall, Taylor, and collaborators investigated deep inelastic scattering of high-energy electrons by nucleons at the Stanford Linear Accelerator Center (SLAC) in the late 1960's. Their analysis indeed showed the dynamical effects of point-like constituents within hadrons. Hence, the quark model highlighted a new simplicity in particle physics and established the quark as an elementary constituent of matter.

Then Glashow, Salam, and Weinberg developed the electroweak theory, which is often compared to the unification of the electric and magnetic interaction by Faraday and Maxwell. This new theory made remarkable predictions, including the existence of a neutral and a charged massive mediator for the weak force: the Z^0 and the $W^{\pm}$ bosons. Until 1973, all observed weak processes were consistent with the hypothesis that they were associated with the exchange of the $W^{\pm}$ bosons. At the Centre Européen de Recherche Nucléaire (CERN), in 1973, neutral current reactions were observed in heavy-liquid bubble chambers. The electroweak theory has evolved into what is now known as the Standard Model (SM) of the electroweak interaction.

By the mid-1970's, all hadrons discovered could be accounted for by the three quarks u, d and s proposed by Gell-Mann and Zweig. In 1974, two completely independent experiments, one at BNL and the other at SLAC, found a new particle. This particle is now called J/ψ and represents the first discovered bound state of a new type of quark known as charm. Clearly the discovery of the J/ψ forced a major overhaul of the quark model. Previously, theorists had suggested a fourth quark based on symmetry between hadrons and leptons. In 1970, Glashow, Iliopoulos, and Maiani (GIM) had already postulated the introduction of a new quark to remove the flavored-changing neutral current contribution in kaon decays. Later in 1974, a second resonance, the ψ', was found and thus re-enforced the presence of charmonium states and a whole new family of charm particles. In 1976, at SLAC, the D^0 meson ($c\bar{u}$ state) was discovered in the $D^0 \to K^-\pi^+$ mode. Since then, other charm mesons (D^* and D_s) and charm baryons (Λ_c for example) have been found.

In the mean time, Perl and his group produced evidence for the existence of a new heavy lepton. The discovery was based on the observation of $e\mu$ events in e^+e^- collisions attributed to $e^+e^- \to \tau^+\tau^-$, where $\tau^+ \to e^+\nu_e\bar{\nu}_\tau$ and $\tau^- \to \mu^-\bar{\nu}_\mu\nu_\tau$. It was

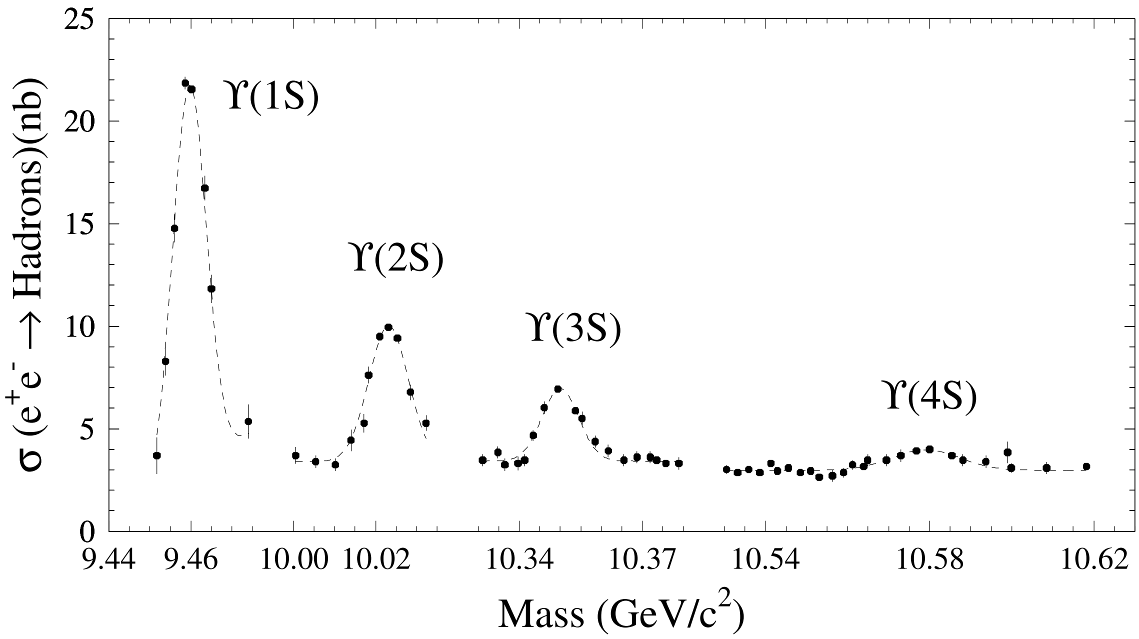


Figure 1.1: The hadronic cross-section for e^+e^- annihilations around $\sqrt{s} = 10$ GeV$/c^2$ measured by the CUBS collaboration.

suggested that the τ lepton is part of a third lepton family (τ^-, ν_τ) and that the three leptons (e^-, μ^-, τ^-), and their associated neutrinos $(\nu_e, \nu_\mu, \nu_\tau)$, occur in pairs. The tau neutrino has never been detected directly, but its existence is essential for lepton number conservation.

Evidence for the fifth quark, the bottom quark b, associated with the quantum number beauty, came from the discovery in 1977 of the lightest bottonium state, the $\Upsilon(1S)$ meson. The experiment was performed at the Fermi National Laboratory (FNAL) under the direction of Leon Lederman. It was similar in design to the one at the BNL that led to the discovery of charm. A beam of protons was allowed to strike a nuclear target surrounded by a spectrometer set up to look for di-lepton events. The $\Upsilon(1S)$ and $\Upsilon(2S)$ mesons were observed as narrow states a year later in e^+e^- annihilation at the DORIS ring in Hamburg, and in 1979 at the Cornell Electron-positron Storage Ring (CESR) at Cornell University (see Figure 1.1). Later at CESR, the $\Upsilon(3S)$ and $\Upsilon(4S)$ states were identified. Subsequently, the B mesons ($B^- = b\bar{u}$,

$\bar{B}^0 = b\bar{d}$, $B^+ = \bar{b}u$, and $B^0 = \bar{b}d$) were also discovered.

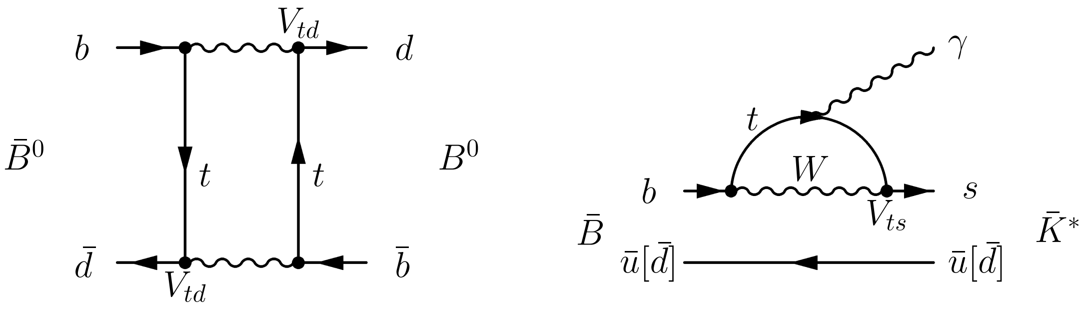
In 1979, first direct confirmation of gluon jets in $e^+e^- \to 3$ jets was found at the the Deutsches Elektronen Synchrotron (DESY). In the early 1970, Nambu, Fritzsch, and Gell-Mann had proposed a theory of the strong interaction in which the massless gluon carries the strong force. In this theory, neither the quarks nor the gluons are free color-neutral particles.

It was not until 1983 before the W and Z bosons were detected. They were first produced at CERN in $p\bar{p}$ collisions in the reactions: $p\bar{p} \to W^{\pm}X^{\mp}$ or $p\bar{p} \to Z^0X^0$, where $X^{\mp}$ and X^0 are any allowed hadronic states. The heavy bosons were detected in their leptonic decay modes: $W^- \to \ell^-\bar{\nu}_\ell$ (or $W^+ \to \ell^+\nu_\ell$) and $Z^0 \to \ell^+\ell^-$, where $\ell = e$ or μ.

The Z and W bosons can also be produced in e^+e^- collision. Hence, in the mid-1980, two e^+e^- colliders tuned to the Z^0 mass ($\sqrt{s} = 91$ GeV) were commissioned. The Stanford Linear Collider (SLC) was the first to produce results. The Large Electron Positron (LEP) collider at CERN started, soon after, producing Z^0s at a much higher rate (10^3 per day). The SLC and LEP experiments concentrated on exploring the weak interaction properties of the Z^0 boson and provided an intensive test of the Standard Model. By measuring the total decay rate of the Z^0 boson, they were able to show that the number of light SM neutrino types (or the number of lepton families) is $N_\nu = 3$.

During the mid and late-1980's, two multipurpose detectors known as ARGUS and CLEO, located respectively at DESY and Cornell University, studied the B meson extensively. Both experiments made some major contributions in understanding the properties of the b quark. In 1987, ARGUS found an unequivocal signature for $B^0\bar{B}^0$ oscillation. Evidence for the transition of a b quark into a first generation u quark was made by CLEO and ARGUS in 1989. In 1993, CLEO found evidence of the penguin decay $b \to s\gamma$. Figure 1.2 shows the Feynman diagrams for oscillation and penguin decays of the $\bar{B}$ meson.

In the 1990's, Isgur and Wise made the remarkable observation that heavy quark transitions can be expressed in terms of a single universal function. This step in understanding the physical properties of hadrons containing a single heavy quark



(a) (b)

Figure 1.2: Feynman diagrams for (a) $B\bar{B}$ oscillation (or mixing), and (b) $b \to s\,\gamma$. The evidence for oscillation and penguin decays opened up new prospects for the study of the b quark properties.

led to the development of the Heavy Quark Effective Theory (HQET). The basis of HQET followed Shuryak's idea that a hadron containing a heavy quark resembles the hydrogen atom with a fixed nucleus at its center [3]. The discovery of HQET, which by now is a well established part of the theoretical framework in particle physics, initiated a better description of heavy quark systems in a well-defined kinematic regime for nonperturbative strong interaction physics. The heavy quark symmetries provided a clear picture of an exclusive semileptonic decay of the B meson. This opened new doors for model-independent determination of certain weak mixing parameters of the Cabibbo-Kobayashi-Maskawa (CKM) matrix. Figure 1.3 illustrates the weak coupling of the b quark to a c quark in a semileptonic decay of the $\bar{B}$ meson.

Following the predictions of the Standard Model and the LEP measurements, it had been anticipated that a sixth quark (the top quark) of charge $+\frac{2}{3}$ should exist as a partner of the bottom quark. The search for the top quark was one of the most outstanding challenges in high energy physics since the discovery of the Upsilon mesons. Finally, in 1995, the CDF and the D0 experiments at FNAL observed the top quark in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV. They searched for events consistent with the production and decay of $t\bar{t}$ pairs in the mode $t\bar{t} \to WWb\bar{b}$. The discovery of the

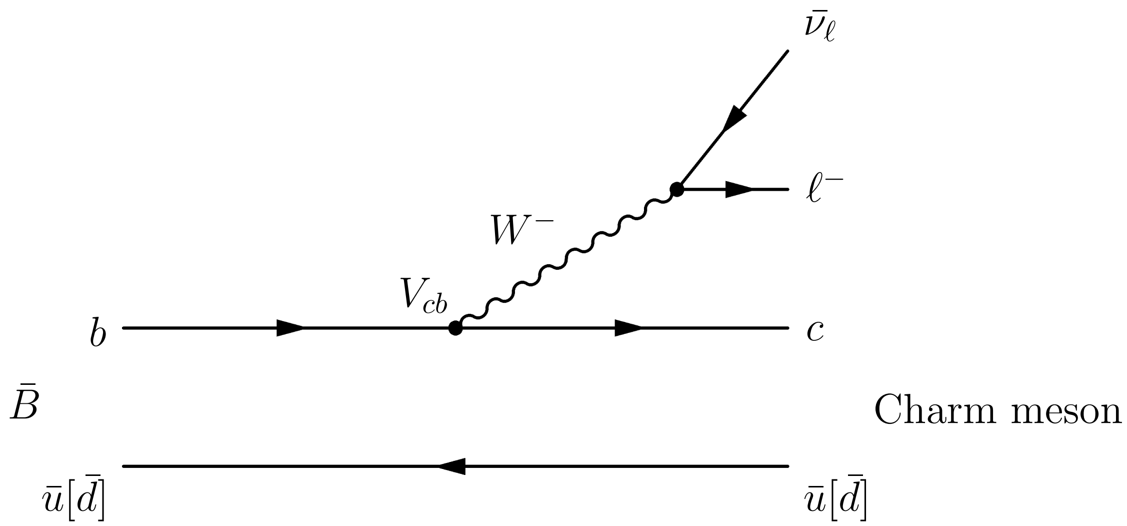


Figure 1.3: Feynman diagram for the semileptonic decay of the $\bar{B}$ meson. This figure shows explicitly the weak coupling of the b quark to a c quark. Exclusive semileptonic decays of heavy-light mesons played a fundamental role in the development of HQET.

top quark confirmed the triplet structure of the quarks. Like the leptons, six types (or flavors) of quarks occur in three pairs (or generations).

In the theoretical framework of the Standard Model, there are presently two fundamental questions at the forefront of high energy physics. The first concerns the origin of mass generation in the electroweak sector via the Higgs mechanism. The other deals with the origin of CP violation.

The Higgs boson is a neutral scalar boson predicted by the SM, but which has not yet been observed. The Higgs field is required for a renormalizable theory of electroweak interactions incorporating the massive W and Z bosons. By searching directly for the SM Higgs particle at LEP1, the LEP experiments have set an unambiguous lower limit of 65 GeV/c^2 on its mass. At LEP2 ($\sqrt{s}$ up to 192 GeV), the SM Higgs will be discovered if its mass is less than about 95 GeV/c^2. The ultimate effort in the search of the Higgs boson is the construction of the Large Hadron Collider (LHC) at CERN. The LHC is a high luminosity pp collider with $\sqrt{s} = 14$ TeV. The LHC experiments will extend the search for the SM Higgs boson up to a mass of 1000 GeV/c^2.

In the Standard Model, the structure of the CKM matrix has significant implications. For three quark generations, the CKM matrix allows a phase which would lead to a possible violation of the CP symmetry. Besides its importance in understanding the structure of the weak interaction, CP asymmetry is necessary to explain the predominance of matter over antimatter in the universe. However, the CP violation in the SM might not be sufficient and CP violation beyond the SM may be required to produce a large enough effect. Experimentally, CP violation has been observed to be very small in K meson decays, but there is still no proof that the CKM matrix is the true source of CP violation in the weak sector. CP asymmetries are predicted by the SM in B meson decays.

Today, the CLEO II detector has collected about five million $B\bar{B}$ pairs in e^+e^- collisions at the $\Upsilon(4S)$ resonance ($\sqrt{s} = 10.58$ GeV). The CLEO collaboration measured with high precision several hadronic, leptonic, and semileptonic branching fractions of the B meson. Consequently, CLEO has opened up new and exciting prospects for exploring and elucidating the weak properties of the bottom quark in high luminosity facilities. Currently, major laboratories around the world (Cornell, CERN, DESY, FNAL, KEK, and SLAC) have been defining a scientific program to search for CP violation in B meson decays.

This brief overview of the evolution of ideas and concepts in particle physics provides a natural introduction to a more detailed description of the minimal Standard Model.

1.3 The Standard Model

The Standard Model (SM) provides a general description of the physics currently accessible with modern particle accelerators. However, many predictions of the SM still have to be verified. We are still searching for the Higgs boson, the particle responsible for the breakdown of the electroweak symmetry, and we are still seeking a better understanding of the electroweak properties of the neutral and charged mediators of the weak interaction, the Z and W bosons. The SM postulates that matter is composed of fundamental spin-$\frac{1}{2}$ quarks and spin-$\frac{1}{2}$ leptons interacting via spin-1 gauge bosons [4]. It incorporates three generations of quarks and leptons grouped into doublets, and gives a unified description of the transitions among them [5]. The SM particles are shown below [6].

Leptons

$$\begin{pmatrix} e \\ \nu_e \end{pmatrix} \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix} \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix} \tag{1.1}$$

Quarks

$$\begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix} \tag{1.2}$$

The quarks are subject to electromagnetic, weak, and strong interactions. The electron, the muon, and the tau leptons are involved in both the electromagnetic and weak interactions, while the massless neutrinos interact only through the weak interaction. As previously mentioned, the gravitational interaction in particle physics is generally neglected. The dynamics of interacting particles in the SM are described by the interaction terms in the Lagrangian. Since the unified electromagnetic and weak interactions are invariant under weak isospin $\mathrm{SU}(2)_L$ and weak hypercharge $\mathrm{U}(1)_Y$, the electroweak Lagrangian contains a $\mathrm{SU}(2) \times \mathrm{U}(1)$ symmetry. The electroweak Lagrangian contains three terms: one for the weak charge current, one for the weak neutral current, and one for the electromagnetic neutral current. Explicitly:

$$\begin{aligned} \mathcal{L} &= \mathcal{L}(\text{Weak CC}) + \mathcal{L}(\text{Weak NC}) + \mathcal{L}(\text{em NC}) \\ &= \frac{g}{\sqrt{2}}\,(J_\mu^- \, W_\mu^+ + J_\mu^+ \, W_\mu^-) + \frac{g}{\cos\theta_W}\,(J_\mu^0 - \sin^2\theta_W \, J_\mu^{\text{em}})\, Z_\mu + e\, J_\mu^{\text{em}} A_\mu, \end{aligned} \tag{1.3}$$

where $W_\mu^\pm$, Z_μ, and A_μ represent the field operators for the physical gauge bosons $W^\pm$, Z^0, and γ, respectively. The coupling constants for the weak and electromagnetic interactions are related by the weak mixing angle, θ_W:

$$e = g\,\sin\theta_W\,. \tag{1.4}$$

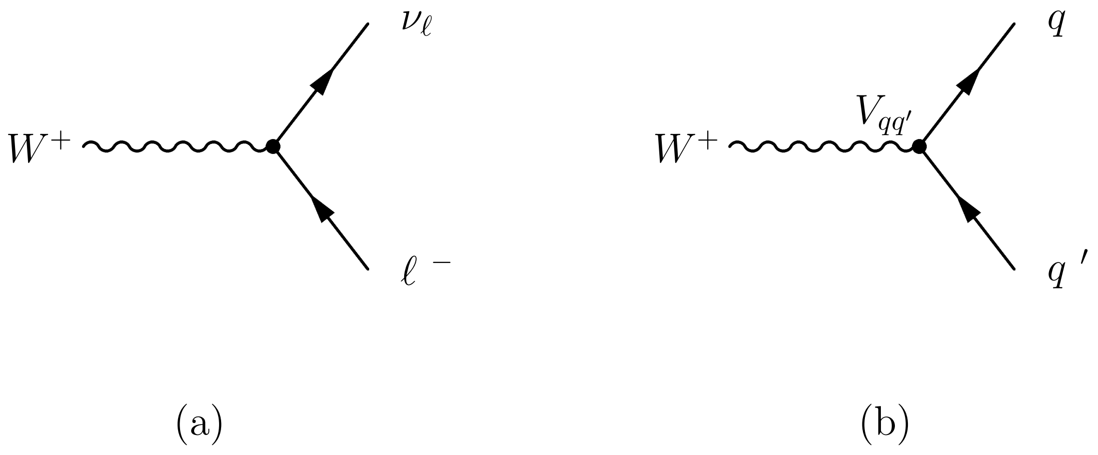


Figure 1.4: Basic weak charged current involving a W boson. In (a) the W couples to the lepton doublets, and in (b) to the quark doublets.

As one can see, the SM provides a clear picture of the weak interaction. Quarks and leptons decay via the weak charged current mediated by the W boson, as shown in Figure 1.4. The weak charged current which mediates the decay process of Figure 1.4 can be written for leptons as:

$$J_\mu^+ = (\bar{\nu}_e, \bar{\nu}_\mu, \bar{\nu}_\tau)\,\frac{1}{2}\,\gamma_\mu\,(1-\gamma^5)\begin{pmatrix} e \\ \mu \\ \tau \end{pmatrix}, \tag{1.5}$$

and for quarks as:

$$J_\mu^+ = (\bar{u}, \bar{c}, \bar{t})\,\frac{1}{2}\,\gamma_\mu\,(1-\gamma^5)\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}. \tag{1.6}$$

Quark transitions $q \to Wq'$ have strengths that depend on the flavor of the quarks involved. The coupling at the W vertex is proportional to $|V_{qq'}|$, where $V_{qq'}$ is a complex number. In the SM, generation-changing transitions between quarks are allowed via the V-A charged current. The coupling in $\ell \to W\nu_\ell$ is also governed by a V-A charged current, but the leptonic number is conserved in the SM when neutrinos are massless [7]. Therefore, generation-changing transitions between leptons are not allowed in the minimal SM.

For three generations of quarks with SU(2) $\times$ U(1) as the gauge group, the relation between the quark mass eigenstates and the weak eigenstates (denoted with a prime in Equation (1.6)) in weak transitions is governed by a 3×3 matrix called the Cabibbo-Kobayashi-Maskawa (CKM) matrix [8]. The CKM matrix is a generalization of the Cabibbo hypothesis known since 1963 [9]. The quark mixing matrix V can be expressed as:

$$V = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}, \tag{1.7}$$

such that

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.8}$$

Since the elements of the CKM matrix can be complex, a total of eighteen numbers are needed to describe all the terms of the matrix. By imposing unitarity, and by redefining the quark fields to remove unphysical phases, the numbers of parameters can be reduced from eighteen to four. These four parameters can be chosen as three angles $(\theta_{12}, \theta_{13}, \theta_{23})$ and one phase (δ), and the CKM matrix can be written as the product of three separate matrices:

$$V = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{1.9}$$

where $c_{ij} = \cos\theta_{ij}$, $s_{ij} = \sin\theta_{ij}$, and i, j denote the quark generations. The middle matrix in Equation (1.9) has been chosen to incorporate the phase δ because it describes a rotation between quarks that are two generations apart. Multiplying these matrices, we obtain:

$$V = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}. \tag{1.10}$$

Based on the empirical observation that the mixing angles have a hierarchical structure such that we can expand in powers of the Cabibbo angle $\lambda = s_{12} = 0.22$, with $s_{23} = A\lambda^2$ and $s_{13}e^{-i\delta} = A\lambda^3(\rho - i\eta)$. The CKM matrix takes the form [10]:

$$V \simeq \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix}. \tag{1.11}$$

In order to have a more complete description of the fundamental interactions in particle physics, one must include the gluonic fields in the SM framework. The theory which describes the strong interaction in the Standard Model is called Quantum Chromodynamics (QCD). To describe QCD, the electroweak Lagrangian is extended to include an SU(3) color symmetry. The mediator of the strong force, the gluon, couples to the color charge of the quark and therefore belongs to an octet representation of SU(3). Although QCD is not tested to the same extent as QED, it is nevertheless in impressive agreement with a large body of experimental data. The favored form of the strong interaction potential for short interquark distances ($r \lesssim R_{\text{hadron}} \simeq 1/\Lambda_{\text{QCD}} \simeq 1$ fm) is:

$$V_{\text{QCD}} \simeq -\frac{4\,\alpha_s}{3\,r}, \tag{1.12}$$

where α_s is the strong coupling constants between quarks and gluons. At large distances ($r > 1$ fm), a confining term must be added to the Coulomb type potential to confine quarks inside hadrons.

Although the SM has great predictive power, it contains many free parameters. The gauge coupling constants ($\alpha_{\rm em}$, G_F, α_s), the parameters of the Higgs field (m_Z, $m_{\rm Higgs}$), the fermions (quarks and leptons) masses, and the CKM matrix elements all have to be determined experimentally.

1.4 Thesis Outline

Having summarized the basic concepts underlying the study of elementary particle physics, we are now better able to focus on the main goals of this thesis.

In the next chapter, the principal heavy-light mesons, along with the production and decay dynamics of the B meson, are discussed; this is followed by the analysis motivation and objectives. In Chapter 3, the theory of B semileptonic decays is summarized. The experimental apparatus used to make the measurements is introduced in Chapter 4. A review of the tracking system calibration procedure is presented in Chapter 5. In Chapter 6, the event selection are described; while Chapter 7 is devoted to the experimental results. The final chapter contains interpretations and conclusion. A list of the CLEO collaborators and an overview of the CLEO terminology are given in the appendices, along with details of specific analysis studies. Enjoy!

Chapter 2

B Meson Physics

In recent years, our understanding of heavy flavor physics has advanced significantly. The study of heavy quark systems provides a rich source of information about particle physics. The primary subject of this dissertation is new experimental results on the weak decays of the B meson. As noted in the introduction chapter, a B meson is a bound state of a $\bar{b}$ quark and a lighter u or d quark. B mesons can be produced copiously in colliding beam machines. Here, we are mainly interested in the properties of the B meson in a physical process in which the bottom quark decays semileptonically to excited charm states.

In the next sections, a review of non-strange charm and bottom mesons is presented. This is followed by a general description of the weak decay properties of the B meson.

2.1 Heavy-light Mesons

In the quark model, a meson is a bound state of a quark and an antiquark. The bottom and charm quarks are both heavy compared to the QCD scale. The principal non-strange heavy-light mesons containing one c or one b quark, and a light antiquark, are listed in Table 2.1, along with their quark composition, masses, and quantum numbers. In the beauty sector, the B mesons are spin-0 mesons. In the charm sector, the D mesons are spin-0 meson; while for the D^*, quark-quark spin coupling leads to spin-1. The D_J mesons have a relative angular momentum $L = 1$ between the charm and the light quarks and hence are called P-wave charm mesons. The quark spins of the D_J can sum to $S = 0$ or $S = 1$, leading to four spin-parity states $J^P = 1^+$ or $0^+, 1^+$, and 2^+. The D_J form two doublets: the D_1 and D_2^*, which are narrow states, and the D_0^* and D_1^*, which are believed to be broad states (more details on the P-wave charm mesons can be found in Section 3.3.1). The D' and the $D^{*\prime}$ mesons are the radially excited states (n=2) $J^P = 0^-$ and $J^P = 1^-$ respectively. They are predicted to be broad states. The orbitally and radially excited charm mesons are sometimes denoted D^{**}. In this dissertation, the notation D_J is used to refer to the P-wave charm mesons, and D^{**} is employed to described all orbitally and radially excited charm mesons. In the present analysis we focus on the decay of the B^- meson to a D_1^0 or a D_2^{*0} meson, and a lepton-neutrino pair.

2.2 *B* Meson Production

When an electron and a positron collide and annihilate, the resulting energy leads to the creation of new matter in the form of lepton and quark pairs. At relatively low center-of-mass energies (*i.e.*, $q^2 \ll m_Z^2$), quark and lepton pair production from e^+e^- annihilation proceeds via a virtual photon as shown in Figure 2.1.

Quarks can never appear as free particles in a final state. In e^+e^- annihilation, the quarks produced are initially free, but as they separate to distance $\mathcal{O}(1 \text{ fm})$ the increasing strength of the strong interaction converts their kinetic energy into additional quarks which combine to form mesons. This process is called fragmentation. Near $\sqrt{s} \simeq 10$ GeV, the energy is sufficient for the production of $b\bar{b}$ pairs. The hadronic cross-section for $e^+e^- \to q\bar{q}$ around $\sqrt{s} = 10$ GeV is shown in Figure 1.1. The $\Upsilon(1S)$, $\Upsilon(2S)$, and $\Upsilon(3S)$ are the lowest $b\bar{b}$ states. These states are below the threshold for open beauty production and therefore decay relatively slowly to non-bottom hadrons via triple gluon exchange, which severely limit their hadronic widths. At the fourth resonance, there is enough energy in the excited state to create a light quark pair ($u\bar{u}$ or $d\bar{d}$) and to produce a pair of B mesons. The $\Upsilon(4S)$ indeed decays predominantly into pairs of B mesons by the OZI favored channel shown in

Charm Mesons			
Hadron Symbol	Quark Content	Mass (GeV/c^2)	$n\ ^{2S+1}L_J$
D^0	$c\bar{u}$	1.865	$1\ ^1S_0$
D^{*+}	$c\bar{d}$	2.010	$1\ ^3S_1$
D_1^0	$c\bar{u}$	2.422	$1\ ^1P_1$
D_0^{*0}	$c\bar{u}$	$\sim$2.360	$1\ ^3P_0$
D_1^{*0}	$c\bar{u}$	$\sim$2.420	$1\ ^3P_1$
D_2^{*0}	$c\bar{u}$	2.459	$1\ ^3P_2$
D'	$c\bar{u}$	$\sim$2.580	$2\ ^1S_0$
$D^{*\prime}$	$c\bar{u}$	$\sim$2.640	$2\ ^3S_1$

Bottom Mesons			
Hadron Symbol	Quark Content	Mass (GeV/c^2)	$n\ ^{2S+1}L_J$
B^-	$b\bar{u}$	5.279	$1\ ^1S_0$
$\bar{B}^0$	$b\bar{d}$	5.279	$1\ ^1S_0$

Table 2.1: In this table, the principal non-strange charm and bottom mesons are listed, along with their quark composition, masses, and quantum numbers. Each meson listed has its antiparticle with the opposite quark content. The mass values are taken from the Particle Data Group compilation [6]. Only the mesons of interest in this thesis are listed. The broad states (D_0^{*0}, D_1^{*0}) and $(D', D^{*\prime})$ have not yet been observed directly and the masses given are theoretical predictions based on heavy-light spectroscopy. Note that we follow the convention that the $\bar{B}$ contains the b quark and the B the $\bar{b}$ quark.

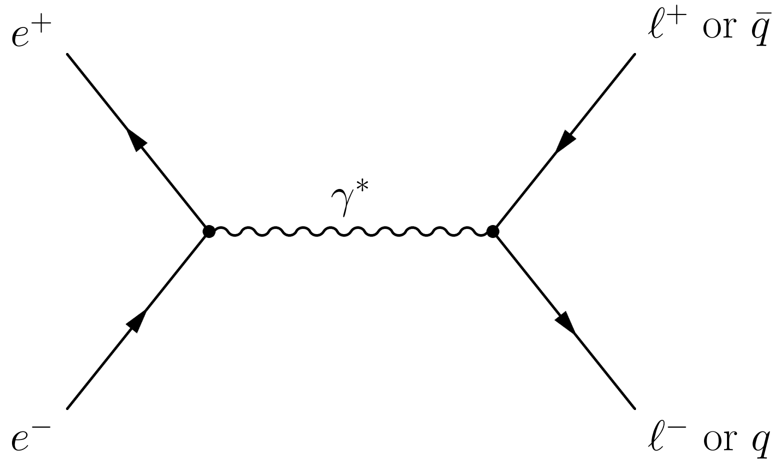


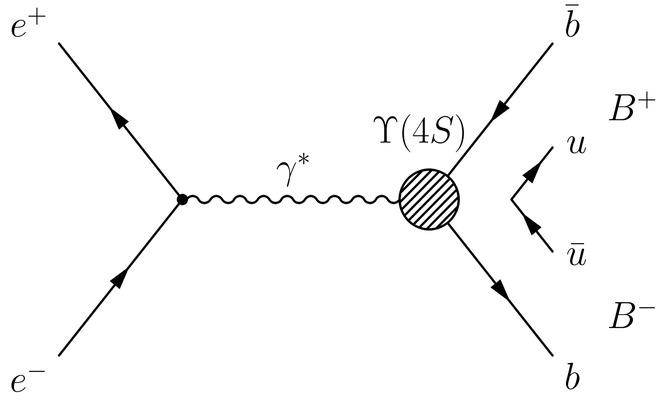
Figure 2.1: First order QED contribution to $e^+e^- \to \ell^+\ell^-$ or $q\bar{q}$. At $q^2 \ll m_Z^2$, all leptons and all quarks, with the exception of the top quark, can be produced.

Figure 2.2. This implies that the 4S state has a much greater width than the 1S, 2S, and 3S states.

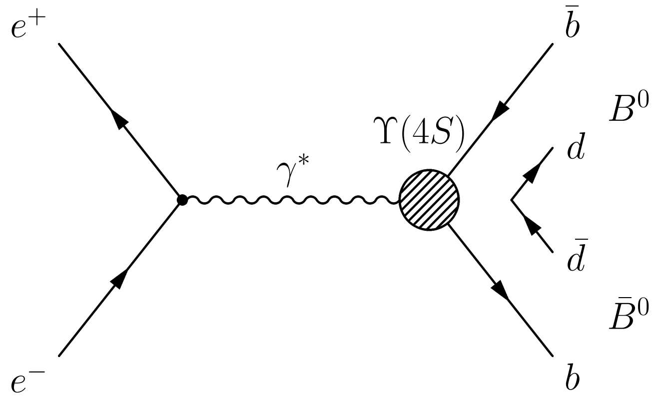
At the $\Upsilon(4S)$ resonance, a $B\bar{B}$ event is an e^+e^- interaction which results in $\Upsilon(4S) \to B\bar{B}$ and a continuum event is an e^+e^- interaction which results in $q\bar{q}$ hadronization rather than producing an $\Upsilon(4S)$ meson. As one can see in Figure 1.1, the relative cross-section $\sigma(e^+e^- \to \Upsilon(4S))/\sigma(e^+e^- \to q\bar{q}) \simeq 1/3$ at $\sqrt{s} = 10.58$ GeV/c^2.

2.3 *B* Meson Decay

Once produced, the B mesons decay weakly with a lifetime of about 10^{-12} sec [11]. The simplest model for B meson decay is called the spectator model. In this model, the bottom quark decays via the weak charged current mediated by the W boson and the light quark acts as a mere spectator. The spectator model decay is shown in Figure 2.3(a). In the CKM scheme, the bottom quark can decay into a c quark or the lighter u quark, with amplitude proportional to $|V_{cb}|$ or $|V_{ub}|$ respectively. At the



(a)



(b)

Figure 2.2: $B\bar{B}$ production mechanism in e^+e^- collisions at the $\Upsilon(4S)$ resonance. In (a) $\Upsilon(4S) \to B^+B^-$ and (b) $\Upsilon(4S) \to B^0\bar{B}^0$.

$\Upsilon(4S)$ resonance, one can therefore produce B mesons and study the fundamental couplings to the W boson for the transition of b to c quarks and the much rarer process of changing a b quark into a first generation u quark. In the Wolfenstein parameterization of the CKM matrix in Equation (1.11), the ratio $|V_{ub}/V_{cb}|$ is given by:

$$|V_{ub}/V_{cb}| = \lambda\,(\rho^2 + \eta^2)^{1/2}. \tag{2.1}$$

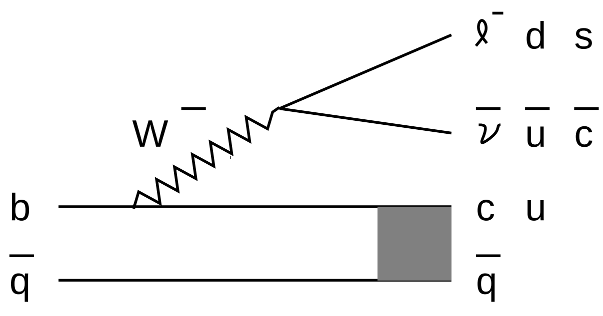
The CKM constraints are normally summarized in the literature by the unitarity triangle in the $\rho - \eta$ complex plane [12, 13]. The unitarity triangle is a simple geometrical representation of the unitarity of the CKM matrix:

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0, \tag{2.2}$$

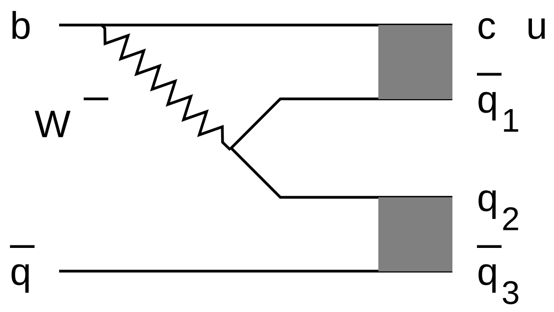
as represented in Figure 2.4. Already, many direct and indirect constraints can be extracted from the existing data from K and B decays. The measurement of $|\epsilon_K|$, the CP violating parameter in K decays, $x_d = \Delta M/\Gamma$, the mixing parameter in $B^0\bar{B}^0$ mixing, and the current measurements of $|V_{cb}|$ and $|V_{ub}|$ limit the allowed region in the $\rho - \eta$ space. These constraints have been discussed extensively in many places [14], they are summarized in Figure 2.5.

More complicated B meson decays include the color suppressed, annihilation, W exchange, and penguin decays, as shown in Figure 2.3. Penguin diagrams in particular are sensitive to several CKM matrix elements and the CKM phase. The goal for the next generation of high luminosity B facilities is to restrict (ρ, η) space, and therefore measure the CKM phase and the CP asymmetry in the b quark sector. The major aim of CP violation studies in B decays is to make enough independent measurements of the CKM parameters (ρ, η) so as to allow a check of the validity of the SM. High precision measurements may reveal some inconsistency and lead to new physics beyond the SM. Indeed, the large baryon excess in the universe suggests that CP violation in the weak interaction is not sufficient.

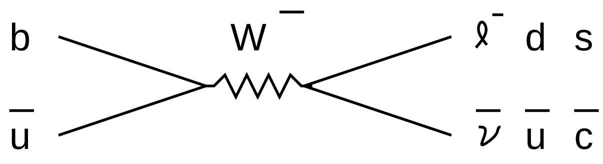
Figure 2.6 summarizes the methods available for measurements of all the CKM matrix elements. One should note the particular importance of B meson decays in the extraction of many of these elements.



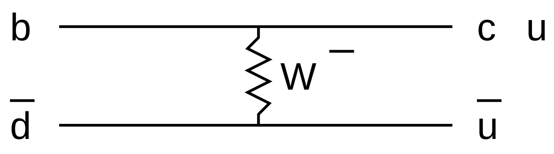
(a) Spectator



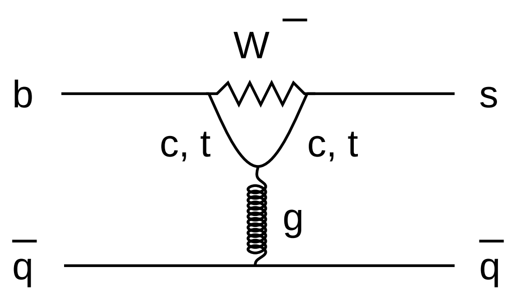
(b) Color Mixed



(c) Annihilation



(d) Exchange



(e) Penguin

Figure 2.3: Quark-level Feynman diagram for B decays:(a) spectator, (b) color mixed, (c) annihilation, (d) exchange, and (e) penguin.

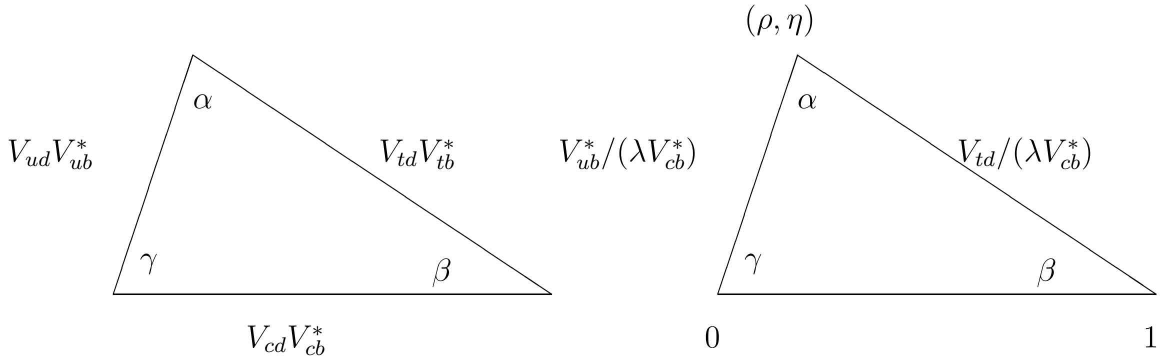


Figure 2.4: The unitarity triangle.

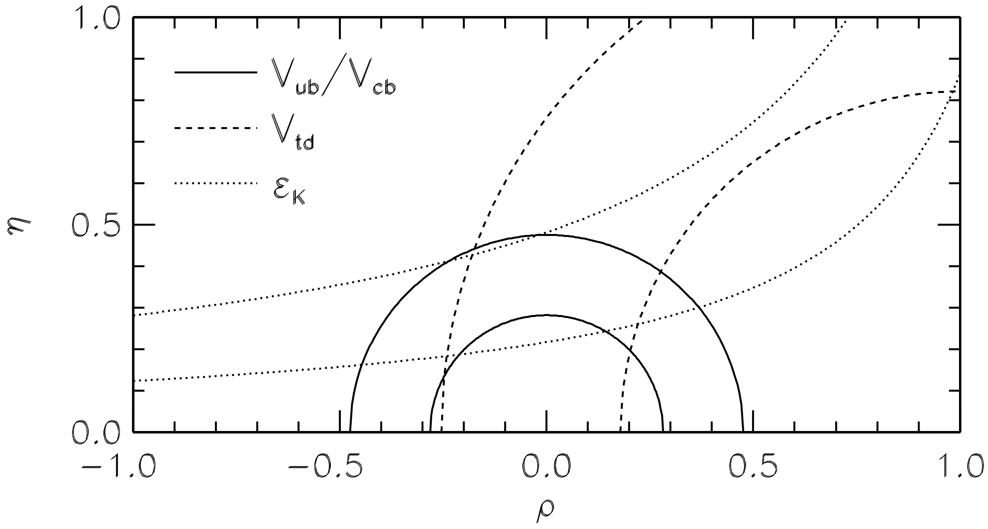


Figure 2.5: The allowed region in the $\rho - \eta$ plane is the intersection of the three region defined by experimental measurements in K and B decays [15]. We used $|V_{cb}| = (39.6 \pm 1.7) \times 10^{-3}$, $|V_{ub}| = (3.3 \pm 0.83) \times 10^{-3}$, $|\epsilon_K| = 2.26 \times 10^{-3}$, and measurements of mixing parameter x_d lead to $|V_{td}| = (8.7^{+2.2}_{-1.5}) \times 10^{-3}$. Theoretical value of $\sqrt{B_B}\, f_B = 200 \pm 40$ is used. More details on the constraints on (ρ, η) can be found in References [12, 13, 14].

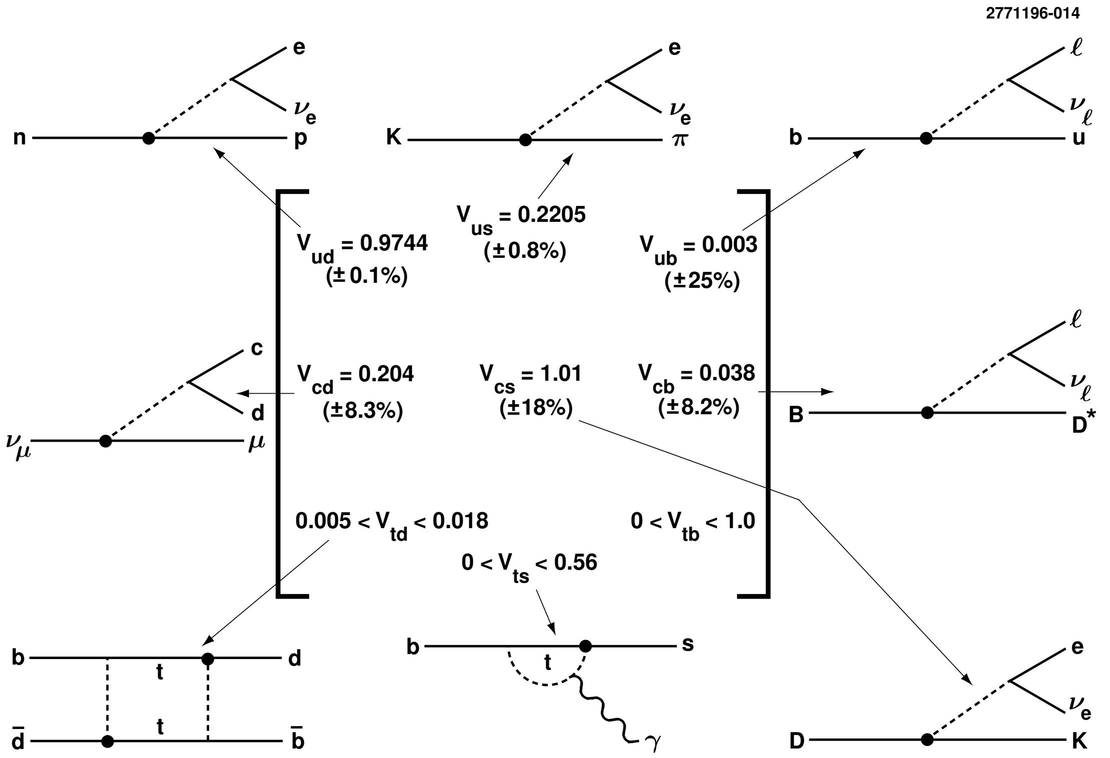


Figure 2.6: The measured values of CKM matrix elements and a schematic diagram indicating the processes used to measured them [16].

2.4 *B* Semileptonic Decays to Charm Mesons

The B meson offers a great variety of experimental probes for weak processes [17]. One of these is through the study of the fundamental couplings to the W boson in the transition of b to c quarks. The virtual W^* boson, produced in the weak process $b \to W^*c$, can decay to a lepton-neutrino pair because the W boson not only couples to quarks but also to leptons. Hence, the B meson can decay semileptonically to a charm meson, a lepton, and a neutrino. The presence of a single charged lepton in the final state of a semileptonic decay of the B meson provides a clear experimental signature for a weak process mediated by a W boson. Because semileptonic decays are both relatively simple and experimentally accessible, they are presently the primary tool for investigating the effective couplings of the W boson to the b quark.

Today, one of the major tasks in heavy flavor physics is to test the predictions of Heavy Quark Effective Theory (HQET) in heavy quark decays such as B semileptonic decays. The simplicity of the spectator decay leads to straightforward and reliable theoretical predictions for B decays. In a semileptonic B decay, the decay amplitude can be written as the product of a leptonic current and a hadronic current. The leptonic current is simple from a theoretical point of view, but the hadronic current contains some non-trivial strong interaction effects. Nevertheless, semileptonic decays provide a relatively simple environment for studying QCD.

The amplitude for a semileptonic $\bar{B}$ meson decay to a charm meson $X \equiv X_{c\bar{q}}$, where q is a light quark (q=u or d), takes the form:

$$\mathcal{M}(\bar{B} \to X\,\ell^- \bar{\nu}_\ell) = -i\,\frac{G_F}{\sqrt{2}}\,V_{cb}\,L^\mu\,H_\mu, \tag{2.3}$$

where the leptonic current, L^μ, can be written in terms of Dirac spinors $\bar{\ell}$ and ν_ℓ:

$$L^\mu = \bar{\ell}\,\gamma^\mu\,(1-\gamma_5)\,\nu_\ell\,, \tag{2.4}$$

and the hadronic current, H_μ, can be expressed in terms of the quark current sandwiched between the meson states:

$$H_\mu = \langle X|\bar{c}\,\gamma^\mu\,(1-\gamma_5)\,b|\bar{B}\rangle\,. \tag{2.5}$$

The hadronic current contains information about the structures of the mesons $\bar{B}$ and X. The decay rate for $\bar{B} \to X\,\ell^- \bar{\nu}_\ell$ is related to the transition amplitude by

$$d\,\Gamma = \frac{(2\pi)^4}{2E_{\bar{B}}}|\mathcal{M}|^2\,\frac{d^3\mathbf{p}_X}{2(2\pi)^3E_X}\,\frac{d^3\mathbf{p}_\ell}{2(2\pi)^3E_\ell}\,\frac{d^3\mathbf{p}_{\bar{\nu}_\ell}}{2(2\pi)^3E_{\bar{\nu}_\ell}}\,\delta^4(P_{\bar{B}}-P_X-P_\ell-P_{\bar{\nu}_\ell}), \tag{2.6}$$

where $P_i = (E_i, \mathbf{p}_i)$ is the four-momentum vector for a given particle ($i = \bar{B}, X, \ell$ or $\bar{\nu}_\ell$). The semileptonic decay rate for $\bar{B} \to X\,\ell^- \bar{\nu}_\ell$ depends on the momentum transfer q^2, the mass of the virtual W^*. In the $\bar{B}$ rest frame,

$$q^2 = m_{W^*}^2 = (P_\ell + P_{\bar{\nu}_\ell})^2 = (P_B - P_X)^2 = m_B^2 + m_X^2 - 2\,m_B\,E_X\,, \tag{2.7}$$

where the mass of the $\bar{B}$ meson is m_B and the mass of the charm meson is m_X. At the $\Upsilon(4S)$ resonance, B mesons are produced almost at rest in the laboratory frame.

In the rest frame of the $\bar{B}$, the four-velocity transfer $w = (v_B \cdot v_X)$ is related to q^2 by:

$$w = \frac{m_B^2 + m_X^2 - q^2}{2\, m_B\, m_X} \tag{2.8}$$

This quantity is in fact the relativistic $\gamma_X = 1/\sqrt{1-\beta_X^2}$ of the meson X in the $\bar{B}$ rest frame. Heavy Quark Symmetry (HQS) often refers to the momentum transfer of the light constituent, which has the typical scale $\Lambda_{\rm QCD}\,(v_B \cdot v_X - 1)$.

A high value of q^2 (small recoil) corresponds to the meson X being created at rest and the lepton and the neutrino being produced nearly back-to-back. The zero-recoil configuration is when $q^2 = q^2_{\rm max} = (m_M - m_X)^2$. At small q^2 (large recoil), the charm quark recoils against the virtual W^* and initially moves rapidly away from the spectator quark. In order to form a bound state X, gluons must be exchanged between the charm quark and the light degrees of freedom. QCD predictions near $q^2 = q^2_{\rm min}$ are non-trivial because the hadronic system is highly disturbed. Consequently, the q^2 distribution is affected by the dynamics of the formation of the hadronic system. At maximum recoil, $q^2 \simeq m_\ell^2$, which is nearly zero for $\ell = e$ or μ.

The fact that the electron and the muon are nearly massless implies a definite spin structure for $\bar{B} \to X\,\ell^- \bar{\nu}_\ell$. For light leptons, the $\ell^-\,\bar{\nu}_\ell$ system has a helicity $\lambda = -1$ because the term $\frac{1}{2}(1-\gamma_5)$ in Equation (1.5) automatically selects a left-handed electron and a right-handed antineutrino. The $V - A$ structure of the weak current leads directly to a characteristic dependence of the lepton energy on the spin and the q^2 of the virtual W^* boson in semileptonic decays. Therefore, since q^2 and E_ℓ are affected by the spin structure of the decay $\bar{B} \to X\,W^*$, study of these quantities provides information on the coupling to the W boson and on the hadronic transition matrix.

In semileptonic decays of a spinless B meson, the total angular momentum of the daughter charm meson X must cancel the total angular momentum of the lepton-neutrino system. In a $b \to c$ semileptonic decay, the charm quark helicity is predominantly $\lambda = -1/2$ and manifests itself as the probability for the meson X to have helicity $\lambda = -1$ or 0 rather than $\lambda = +1$. Several useful observables relate the helicity dynamics to the polarization, Γ_L/Γ_T, of a particular B semileptonic decay [17]: the forward-backward asymmetry, $\bar{A}_{\rm FB}$, of the lepton in the W rest frame, and the longitudinal to transverse polarization, $\bar{A}_{\rm POL}$, of the meson X. These are defined as:

$$\bar{A}_{\rm FB} = \frac{3}{4}\frac{(\Gamma_- - \Gamma_+)}{(\Gamma_L + \Gamma_T)}\,, \tag{2.9}$$

and

$$\bar{A}_{\rm POL} = 2\frac{\Gamma_L}{\Gamma_T} - 1\,. \tag{2.10}$$

where $\Gamma_L = \Gamma_0$ and $\Gamma_T = \Gamma_+ + \Gamma_-$. The decay rate for helicity state $\lambda = 0$ is Γ_0 and the decay rate for helicity states $\lambda = \pm 1$ is $\Gamma_\pm$.

This thesis presents an experimental investigation of P-wave charm meson ($X = D_J$) production in semileptonic B meson decays using data collected with the CLEO II detector. The exclusive decays $B^- \to D_1^0 \ell^- \bar{\nu}_\ell$ and $B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell$ [18] are studied by reconstructing the decay channel $D_J^0 \to D^{*+}\pi^-$ using the decay chain $D^{*+} \to D^0\pi^+$, and $D^0 \to K^-\pi^+$ or $D^0 \to K^-\pi^+\pi^0$ [19].

In the semileptonic decay of a B meson to a D_J meson, q^2 has some predictable features. At large q^2 (small recoil), the D_J is moving slowly and is nearly unpolarized (*i.e.*, $\lambda = -1, 0, +1$ are present in approximately equal amounts). As q^2 decreases the $\lambda = +1$ component of the D_J is suppressed. At $q^2 \to 0$, the lepton and the neutrino are collinear, forcing a pure $\lambda = 0$ state.

Other kinematic variables, such as E_{D_J} and the angles θ_J, θ_ℓ, χ and α (defined in Figures 2.7 and 2.8) also describe the dynamics of the decay $B^- \to D_J^0 \ell^- \bar{\nu}_\ell$, with $D_J^0 \to D^{*+}\pi^-$ followed by $D^{*+} \to D^0\pi_s^+$. The quantity E_{D_J} is the D_J meson energy in the B^- rest frame. In the rest frame of the D_J meson, θ_J is the decay angle of the D^*. Similarly, in the rest frame of the virtual W^* boson, θ_ℓ is the decay angle of the lepton. The angle χ is the angle between the decay plane of the W^* and the decay plane of the D_J, measured in the B rest frame. Finally, the angle α is the D^* helicity angle. The helicity angle α is defined as the angle between the D_J and D momenta, both measured in the D^* rest frame.

In principal, the full dynamics of $B^- \to D_J^0 \ell^- \bar{\nu}_\ell$ can be studied by examining all the available kinematic variables describing the decay. In practice, statistical limitations force us, at the moment, to study the q^2 distribution of the virtual W^* boson alone.

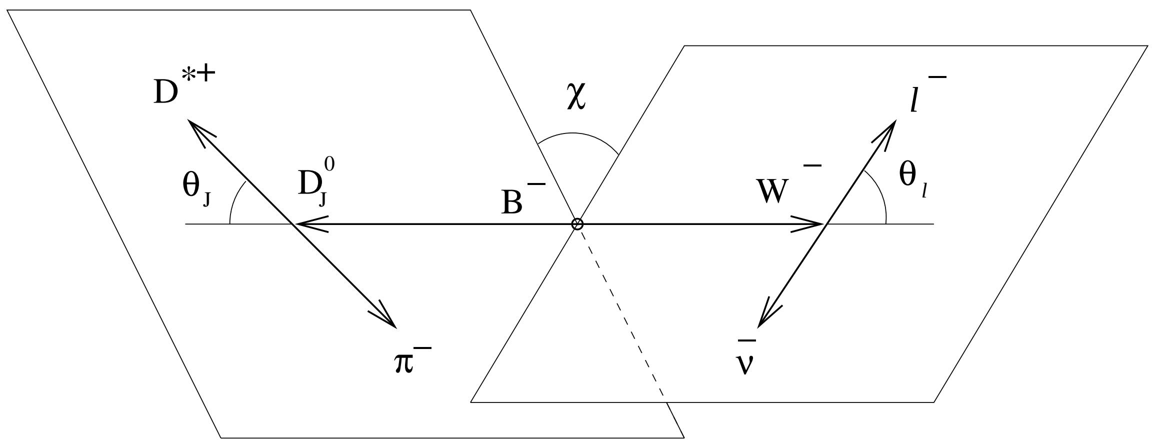


Figure 2.7: Kinematic variables used to describe the decay $B^- \rightarrow D_J^0 \ell^- \bar{\nu}_\ell$, with $D_J^0 \rightarrow D^{*+}\pi^-$. The decay of the B^- meson, W^- boson and D_J^0 meson are shown in their respective rest frames.

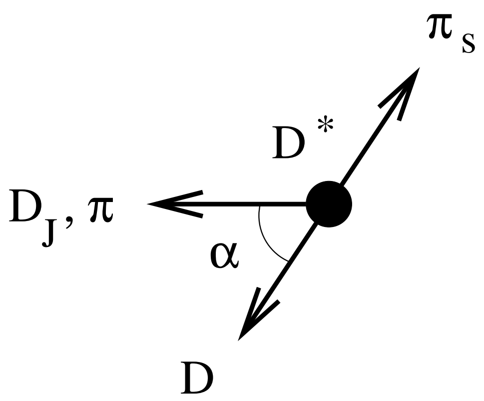


Figure 2.8: Definition of the D^* helicity angle.

2.5 Motivation and Analysis Objectives

The B meson decays to many different states and gives rise to a large number of secondary decay products. The principal decay modes of the B meson can be subdivided into three categories according to the final particles produced. These are: leptonic, semileptonic, and hadronic decays. The primary objective of this thesis is to measure two exclusive B semileptonic branching fractions: $\mathcal{B}(B^- \to D_1^0\ell^-\bar{\nu}_\ell)$ and $\mathcal{B}(B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell)$. A branching fraction is simply the probability of a parent particle to decay via a specific channel. The next sections contain a summary of the experimental and theoretical motivation and objectives of this analysis.

2.5.1 Deficit in Inclusive B Semileptonic Decays

In B semileptonic decays, one expects the sum of all the exclusive modes to saturate the inclusive rate. At the $\Upsilon(4S)$, where the b hadrons are a mixture of $\bar{B}$ mesons ($b\bar{u}$ or $b\bar{d}$ states), the inclusive B semileptonic branching fraction ($\mathcal{B}_{\rm SL}$) is:

$$\begin{aligned}
\mathcal{B}_{\rm SL} &= \sum_{i=u,c} \mathcal{B}(b \to q_i\ell\bar{\nu}_\ell) = \mathcal{B}(b \to u\ell\bar{\nu}_\ell) + \mathcal{B}(b \to c\ell\bar{\nu}_\ell) \\
&= \sum_{i=\text{Hadrons}} \mathcal{B}(\bar{B} \to H_i\ell\bar{\nu}_\ell), \qquad (2.11)
\end{aligned}$$

where H_i is any allowed hadronic final state.

In the inclusive approach, the sum over all possible final states is considered, ignoring the detailed breakdown among the individual decay modes. Experimentally, the inclusive B semileptonic decay branching fraction is obtained by counting the number of leptons from b quarks. At the $\Upsilon(4S)$, the total inclusive B semileptonic branching fraction has been measured many different ways by the ARGUS and the CLEO experiments (see Figure 2.9). One method relies on the measurement of the single-lepton spectrum. This is called the spectral fitting method because the observed inclusive lepton spectrum is composed of leptons from the b hadrons (primary leptons) and leptons from charm decays (secondary leptons). The second method considers events with two leptons. This technique uses the charge and angular correlations in dilepton events to extract the primary lepton spectrum. Results from both methods are given below and summarized in Figure 2.9.

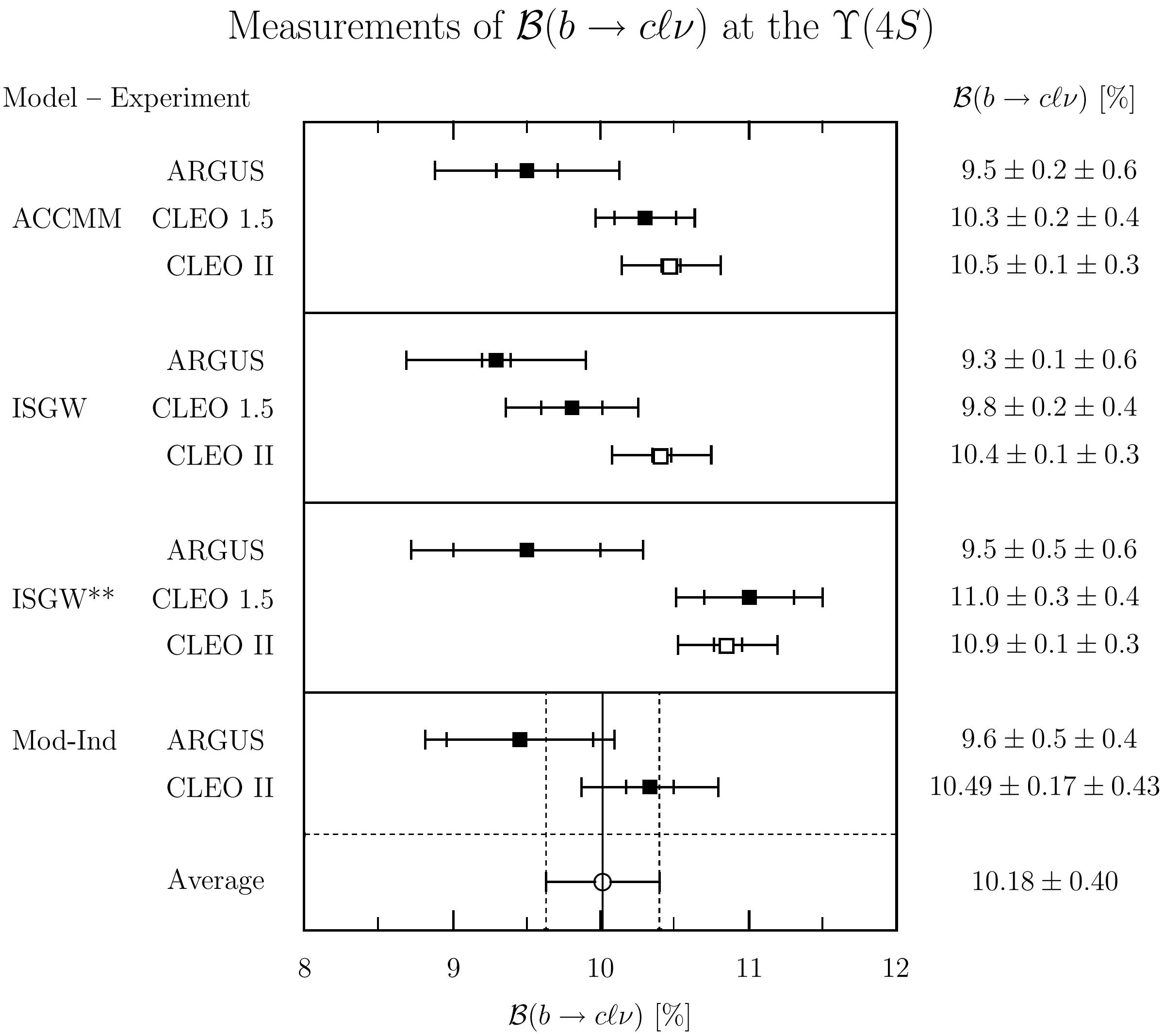


Figure 2.9: Measurements of the inclusive B semileptonic branching fraction using the spectral method and the dilepton method. This is an example of a summary of many experimental results from ARGUS and CLEO. Details can be found in References [15, 17, 20]. The results from the spectral analysis are given for each models used: ACCMM [21], ISGW [22], and ISGW** [23]. The dilepton results are the model-independent measurements. Source [15, 20].

In the spectral fitting method, the ARGUS and CLEO collaborations used theoretical models to describe the primary lepton spectrum. Depending on the model used they measured [15, 20]:

$$\mathcal{B}_{\mathrm{SL}} = (9.3 - 11.0)\%. \tag{2.12}$$

The ARGUS collaboration introduced the dilepton method which reduces the model dependence because of its ability to separate the primary and the secondary lepton contributions. Theoretical models are therefore only used to extrapolate to momenta below the detector acceptance. The average of the two results from CLEO [24] and ARGUS [25] gives [15, 20]:

$$\mathcal{B}_{\mathrm{SL}} = (10.18 \pm 0.40)\%. \tag{2.13}$$

The exclusive approach to semileptonic decays is to measure individually the branching fractions of all the possible modes. The decays $\bar{B} \to X_{c\bar{q}}\ell\bar{\nu}_\ell$, where $X_{c\bar{q}}$ is a charm hadron, account for the majority ($\sim 98.5\%$) of the total B semileptonic rate. Early phenomenological descriptions of exclusive B semileptonic decays expected $\bar{B} \to D\ell\bar{\nu}_\ell$ and $\bar{B} \to D^*\ell\bar{\nu}_\ell$ to saturate the total rate. Presently, there is general agreement among a number of measurements of the exclusive semileptonic $\bar{B}$ meson decays, $\bar{B} \to D\ell\bar{\nu}_\ell$ and $\bar{B} \to D^*\ell\bar{\nu}_\ell$. Together they account for approximately 70% of the inclusive $\mathcal{B}_{\mathrm{SL}}$ branching fraction because

$$\mathcal{B}(\bar{B} \to D^{(*)}\ell\bar{\nu}_\ell) \equiv \mathcal{B}(\bar{B} \to D\ell\bar{\nu}_\ell) + \mathcal{B}(\bar{B} \to D^*\ell\bar{\nu}_\ell) = (6.99 \pm 0.36)\%. \tag{2.14}$$

Figure 2.10 summarizes recent results for $\bar{B} \to D^{(*)}\ell\bar{\nu}_\ell$. These results contrast with the situation in D semileptonic decays, where $D \to \bar{K}\bar{\ell}\nu_\ell$ and $D \to \bar{K}^*\bar{\ell}\nu_\ell$ saturate the total rate.

Table 2.2 shows the contributions from $\bar{B} \to D\ell\bar{\nu}_\ell$ and $\bar{B} \to D^*\ell\bar{\nu}_\ell$ to the inclusive B semileptonic rate. These results indicate that a substantial fraction ($\gtrsim$30%) of the inclusive $\bar{B}$ semileptonic rate is from modes other than $D\ell\bar{\nu}_\ell$ and $D^*\ell\bar{\nu}_\ell$. Since the branching fraction for $b \to u\ell\bar{\nu}_\ell$ is known to be small, the missing exclusive rate in B semileptonic decays must be sought among $b \to c\ell\bar{\nu}_\ell$ decays to higher mass D_J states or nonresonant hadronic states with a D or D^* and other hadrons. Hence, the

Decay Mode	Branching Fraction
$\bar{B} \to D\ell\bar{\nu}_\ell$	$(1.94 \pm 0.26)\%$
$\bar{B} \to D^*\ell\bar{\nu}_\ell$	$(5.05 \pm 0.25)\%$
$\mathcal{B}_{\rm SL}$	$(10.18 \pm 0.40)\%$
Inclusive - Exclusive	$(3.19 \pm 0.54)\%$

Table 2.2: Contribution to the B meson inclusive semileptonic branching fraction. This clearly shows the need to include higher resonance contributions to saturate the exclusive rate. The quoted value for $B_{\rm SL}$ is based on the dilepton method because this has very little model dependence. The values in this table are taken from Figures 2.9 and 2.10.

study of $B^- \to D_J^0\ell^-\bar{\nu}_\ell$ provides very useful input in resolving the difference between the known exclusive channels and the inclusive B semileptonic decay rate.

2.5.2 Precise Measurements in *B* Semileptonic Decays.

As illustrated in Figure 2.6, the weak decays of the B meson provide direct input into the determination of many of the CKM matrix elements. Inclusive and exclusive semileptonic decays of the B meson are especially important for the determination of the element $|V_{cb}|$.

In an inclusive analysis, the measurement of $B_{\rm SL}$ is sensitive to the shape of the lepton spectrum predicted by the theoretical models, while the extraction of $|V_{cb}|$ from $B_{\rm SL}$ is sensitive to the overall normalization. In both cases, the contribution from higher mass states, such as P-wave charm mesons, is important.

In the exclusive approach, the golden modes for the extraction of $|V_{cb}|$ are $\bar{B} \to D^{(*)}\ell\bar{\nu}$. Results for $\mathcal{B}(B^- \to D_J^0\ell^-\bar{\nu}_\ell)$ are crucial for background estimation in making precise measurements of $|V_{cb}|$ and $\mathcal{B}(\bar{B} \to D^{(*)}\ell\bar{\nu})$. Therefore, to provide a complete understanding of $\bar{B} \to D^{(*)}\ell\bar{\nu}$, it is essential to study the semileptonic decays of the B meson to higher resonant states such as the narrow D_J mesons.

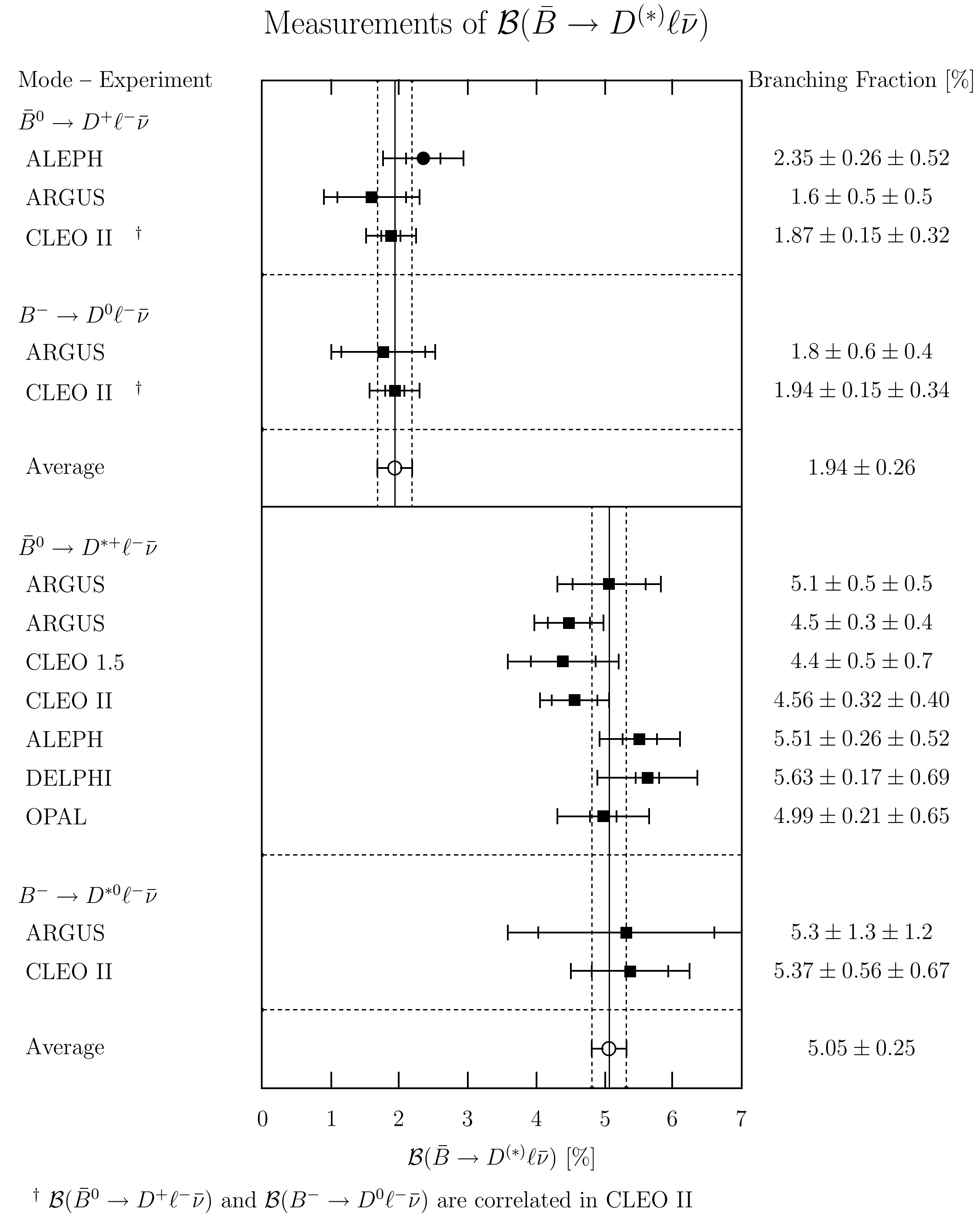


Figure 2.10: Summary of the measurements of exclusive decays $\bar{B} \to D^{(*)}\ell\bar{\nu}_\ell$. Source [15, 20].

2.5.3 Dynamics of Heavy Quark Decays

As one can see, semileptonic B decays have a special standing in our understanding of the phenomenology of the weak interaction. B meson decays not only provide means to study the weak couplings, but also a way to study how the strong interactions affect the weak processes. The decays $B^- \to D_J^0 \ell^- \bar{\nu}_\ell$ can then reveal information about the dynamics of heavy quark decays and the level of heavy quark symmetry breaking. Semileptonic decays of the B meson to D, D^*, and narrow D_J mesons are probably the only modes in the beauty sector that experimentalists and theorists can use to understand the breakdown of heavy quark symmetries, and to study the effects of the strong interaction in the limit of nonperturbative QCD.

At higher luminosity e^+e^- facilities, the experimental errors on certain CKM elements will become negligible. It is therefore important to understand the breakdown of the heavy quark symmetries so that the theoretical uncertainties on the CKM matrix elements $|V_{ub}|$ and $|V_{cb}|$ can be reduced to a few percents. Precise measurements of the CKM matrix elements complement a measurement of the CP asymmetry in the SM, which is presently one of the main goals in particle physics [26].

2.5.4 Analysis Strategy

In summary, the aim of the present analysis is to investigate the production of orbitally excited charm mesons in B semileptonic decays. We provide measurements for $\mathcal{B}(B^- \to D_1^0 \ell^- \bar{\nu}_\ell)$ and $\mathcal{B}(B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell)$, when $D_J^0 \to D^{*+}\pi^-$ followed by $D^{*+} \to D^0\pi^+$ and $D^0 \to K^-\pi^+$ or $D^0 \to K^-\pi^+\pi^0$. Although we are statistically limited, a q^2 distribution for $B^- \to D_1^0 \ell^- \bar{\nu}_\ell$ is extracted.

Evidence for the D^{**} states in semileptonic decays have been reported previously by ARGUS [27], CLEO [28], ALEPH [29], OPAL [30], and DELPHI [31]. All these analyses confirmed the presence of higher mass states contribution in B semileptonic decays. Exclusive measurements of $B^- \to D_1^0 \ell^- \bar{\nu}_\ell$ and $B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell$ have been presented at conferences by CLEO [32, 33]. This thesis reports updated measurements of these two decay modes [34].

The procedure for measuring a branching fraction is straightforward. The first step is to reconstruct the decay products of the B meson in a specific channel. In our case $B^- \to D_J^0 \ell^- \bar{\nu}_\ell$, where $D_J^0 = D_1^0$ or D_2^{*0}. The second step is to determine the reconstruction efficiency by a Monte Carlo simulation. This simulation relies on theoretical predictions for the decay under study and also on a detailed description of the CLEO II detector. With the number of reconstructed B mesons and the reconstruction efficiency, a product branching fraction can be calculated as follows:

$$\begin{aligned} \mathcal{P}(D_J) &= \mathcal{B}(B^- \to D_J^0 \ell^- \bar{\nu}_\ell)\, \mathcal{B}(D_J^0 \to D^{*+}\pi^-) \\ &= \frac{n_{D_J}/\varepsilon_{D_J}}{2 \times 2\, N_{\Upsilon(4S)}\, f_{+-}\, \mathcal{B}(D^{*+} \to D^0\pi^+)\, \mathcal{B}(D^0 \to K^-\pi^+(\pi^0))}, \end{aligned} \tag{2.15}$$

where

$\boldsymbol{\mathcal{B}(B^- \to D_J^0 \ell^- \bar{\nu}_\ell)}$**:** The branching fraction for $B^- \to D_J^0 \ell^- \bar{\nu}_\ell$.

$\boldsymbol{n_{D_J}}$**:** The number of reconstructed $B^- \to D_J^0 \ell^- \bar{\nu}_\ell$ events.

$\boldsymbol{\varepsilon_{D_J}}$**:** The reconstruction efficiency for $B^- \to D_J^0 \ell^- \bar{\nu}_\ell$.

$\boldsymbol{N_{\Upsilon(4S)}}$**:** The number of $\Upsilon(4S)$ is our data sample.

$\boldsymbol{f_{+-}}$**:** The branching fraction for $\Upsilon(4S) \to B^+B^-$.

$\boldsymbol{\mathcal{B}(D_J^0 \to D^{*+}\pi^-)}$**:** The branching fraction for $D_J^0 \to D^{*+}\pi^-$.

$\boldsymbol{\mathcal{B}(D^{*+} \to D^0\pi^+)}$**:** The branching fraction for $D^{*+} \to D^0\pi^+$.

$\boldsymbol{\mathcal{B}(D^0 \to K^-\pi^+(\pi^0))}$**:** The branching fraction for $D^0 \to K^-\pi^+(\pi^0)$.

The factor of (2×2) in the denominator comes from the fact that ℓ represents e and μ and that each $\Upsilon(4S)$ decays to a $B\bar{B}$ pair.

The q^2 distribution for $B^- \to D_J^0 \ell^- \bar{\nu}_\ell$ is simply the differential decay rate $d\Gamma(B^- \to D_J^0 \ell^- \bar{\nu}_\ell)/dq^2$. The decay rate is related to the branching fraction by $\Gamma_J = \Gamma(B^- \to D_J^0 \ell^- \bar{\nu}_\ell) = \mathcal{B}(B^- \to D_J^0 \ell^- \bar{\nu}_\ell)/\tau_{B^-}$. The B^- lifetime is τ_{B^-}. Then,

$$\frac{d\Gamma_J}{dq^2} = \frac{n_{D_J}(q^2)/\varepsilon_{D_J}(q^2)}{4\, \tau_{B^-}\, N_{\Upsilon(4S)}\, f_{+-}\, \mathcal{B}(D_J^0 \to D^{*+}\pi^-)\, \mathcal{B}(D^{*+} \to D^0\pi^+)\, \mathcal{B}(D^0 \to K^-\pi^+(\pi^0))}, \tag{2.16}$$

where $n_J(q^2)$ and $\varepsilon_J(q^2)$ are, respectively, the numbers of reconstructed B^- mesons and the reconstructed efficiency as a function of q^2.

Chapter 3

Theoretical Models for Semileptonic B decays

At the $\Upsilon(4S)$ resonance, inclusive branching fraction measurements involve an admixture of B mesons. Limits obtained for non-$B\bar{B}$ decay of the $\Upsilon(4S)$ are consistent with $\mathcal{B}(\Upsilon(4S) \to B\bar{B}) = 100\%$ [24]. Furthermore, we assume that the branching fractions of $\Upsilon(4S)$ to charged and neutral $B\bar{B}$ pairs are each 50%. An experimental result from CLEO II [28] agrees with $f_+/f_0 \equiv \mathcal{B}(\Upsilon(4S) \to B^+B^-)/\mathcal{B}(\Upsilon(4S) \to B^0\bar{B}^0) = 1$. The inclusive semileptonic branching fraction at the $\Upsilon(4S)$ resonance is then related to the total decay width (Γ_{TOT}) and the semileptonic decay width (Γ_{SL}) of the B meson by:

$$\mathcal{B}_{\mathrm{SL}} = \frac{\Gamma_{\mathrm{SL}}}{\Gamma_{\mathrm{TOT}}} = \tau_B\, \Gamma_{\mathrm{SL}}\,, \tag{3.1}$$

since the τ_{B^+}/τ_{B^0} is consistent with unity [6]. The branching fraction for an exclusive semileptonic decay of a $\bar{B}$ meson is given by:

$$\mathcal{B}(\bar{B} \to H\,\ell\bar{\nu}_\ell) = \frac{\Gamma(\bar{B} \to H\,\ell\bar{\nu}_\ell)}{\Gamma_{\mathrm{TOT}}}\,, \tag{3.2}$$

where $\Gamma(\bar{B} \to H\,\ell\bar{\nu}_\ell)$ is the partial width for $\bar{B} \to H\,\ell\bar{\nu}_\ell$. The state H denotes a particular hadronic final state kinematically allowed in semileptonic $\bar{B}$ decays. Theoretical predictions exist for inclusive and exclusive semileptonic B decays [17]. Models which describe the phenomenology and dynamics of semileptonic decays are discussed in this chapter.

3.1 Free Quark Model

The simplest description of the B meson decay treats the spectator quark as a free particle. The free quark model was developed in the scheme of inclusive decays and therefore leads to prediction for the inclusive lepton energy spectrum.

The partial width for the inclusive semileptonic decay of a free quark Q can be written as:

$$\Gamma(Q \to q\,\ell^- \bar{\nu}_\ell) = \frac{G_F^5 m_Q^5}{192\pi^3} |V_{qQ}|^2\, I(x)\,, \tag{3.3}$$

where the phase factor $I(x)$ for QED radiative corrections is given by:

$$I(x) = 1 - 8x^2 + 8x^6 - x^8 - 24x^4 \ln x\,, \tag{3.4}$$

for $x = m_q/m_Q$. The factor $I(x)$ is close to one for $b \to u\,\ell\bar{\nu}_\ell$ and approximately 0.5 for $b \to c\,\ell\bar{\nu}_\ell$. Here, the analogy with muon decay is obvious

$$\Gamma(\mu^- \to e^- \nu_\mu \bar{\nu}_\ell) = \frac{G_F^5 m_\mu^5}{192\pi^3}\,. \tag{3.5}$$

In the free quark model, the exchange of gluons between quarks in a semileptonic decay modifies the rate. Two diagrams contributing to first order gluon radiation are illustrated in Figure 3.1. These corrections, tabulated in the function $g(x)$, modify Equation (3.3), which becomes:

$$\Gamma(Q \to q\,\ell^- \bar{\nu}_\ell) = \frac{G_F^5 m_Q^5}{192\pi^3} |V_{qQ}|^2\, I(x) \left[1 - \frac{2}{3\pi}\alpha_s g(x)\right]. \tag{3.6}$$

The ACCMM model [21] was one of the first models to incorporate bound state effects to the free quark model. These effects can significantly modify the lepton energy spectrum. In the ACCMM model, the momentum of the light quark within the decaying meson is modeled by a Gaussian distribution $\phi(p)$ which has the form:

$$\phi(p) = \frac{4}{\sqrt{\pi} p_F^3} \exp\left(-\frac{p^2}{p_F^2}\right). \tag{3.7}$$

The parameter p_F is the Fermi momentum (p_F = 150 MeV/c to 300 MeV/c). The free quark spectator model gives a prediction for the lepton energy spectrum for

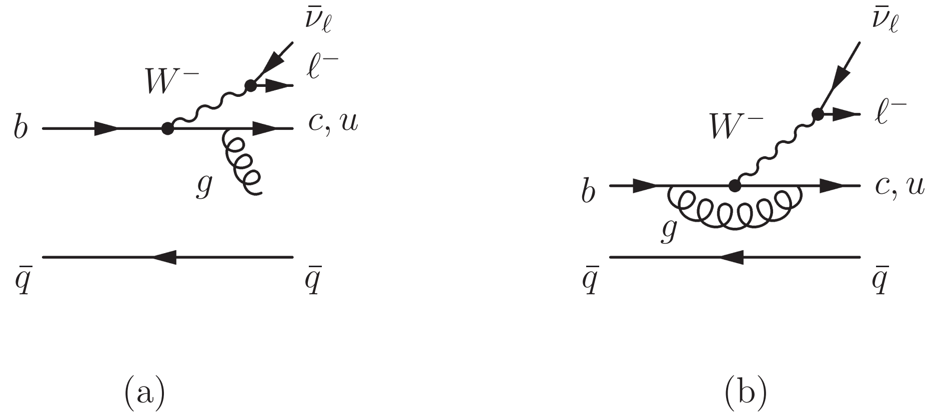


Figure 3.1: First order gluon radiation in B semileptonic decays. These diagrams illustrate real (a) and virtual (b) gluon bremsstrahlung. In (a) the radiated gluon can be emitted from the initial quark as well. In (b) the gluon can be emitted and reabsorbed entirely on either the initial and final quark propagator lines as well.

semileptonic decays of the B meson to charm mesons. In the b quark rest-frame, the partial decay width is:

$$\frac{d\Gamma(b \to c\,\ell\bar{\nu}_\ell)}{dy} = \frac{G_F^5 m_b^5}{192\pi^3}|V_{cb}|^2\,\Phi(x,y)\,G(x,y), \tag{3.8}$$

where $x = m_c/m_b$ and $y = 2E_\ell/m_b$. The phase space factor is $\Phi(x,y)$ and $G(x,y)$ incorporates the effects of gluon radiation [21]. To compute the lepton energy spectrum, the decay distribution in the b quark rest frame is boosted to the B meson frame. The spectator quark in this model is assumed to have a definite mass m_{sp}, but the b quark is a virtual particle of variable mass

$$m_b^2 = m_B^2 + m_{\mathrm{sp}}^2 - 2m_B\sqrt{m_{\mathrm{sp}}^2 + p^2}. \tag{3.9}$$

Thus, the ACCMM model has three free parameters: the Fermi momentum, the effective mass for the light degrees of freedom m_{sp}, and the mass of the daughter quark $m_q = m_c$. The lepton energy spectrum of the ACCMM model for $b \to c\,\ell\bar{\nu}_\ell$ is shown in Figure 3.2. The inclusive ACCMM spectrum of lepton energy from $b \to c \to y\,\ell\bar{\nu}_\ell$ decays is also shown in Figure 3.2.

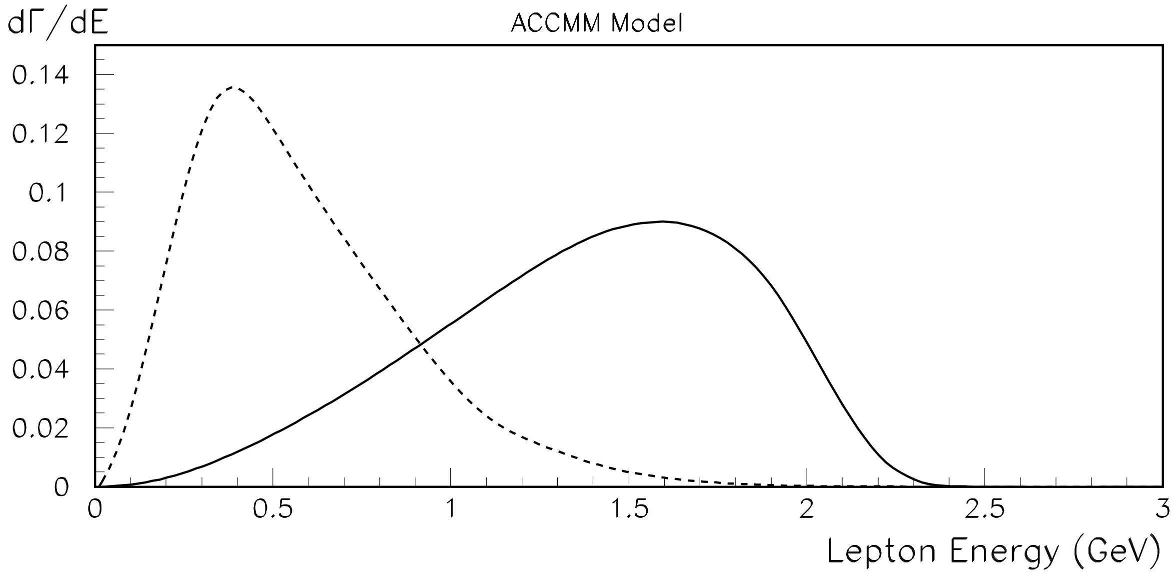


Figure 3.2: The predicted $b \to c\,\ell\bar{\nu}_\ell$ (solid) and $b \to c \to y\,\ell\bar{\nu}_\ell$ (dashed) lepton energy spectra for the ACCMM model. These spectra are based on the fit to the data described in Reference [35]. They have been corrected for detector acceptance and efficiencies. The spectator quark mass is taken to be $m_{\rm sp} = 150$ MeV$/c^2$. The Fermi momentum and the c quark mass are determined from the fit: $p_F = 265 \pm 25$ MeV$/c$ and $m_c = 1.670 \pm 0.025$ GeV$/c^2$.

3.2 Form Factor Models

In a semileptonic decay, the hadronic current can be constructed from the available four-vectors, which are the velocity and spin-polarization vectors, and from Lorentz-invariant coefficients called form factors. The form factors describe the wave functions overlap of the initial and final state hadrons. In this approach, all the QCD effects are swept into the form factors. Consequently, the form factor models take advantage of the fact that the strong interactions can be isolated in the hadronic current of the semileptonic decay amplitude.

In the case of $B^- \to D_1^0\ell^-\bar{\nu}_\ell$ and $B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell$ (see Figure 3.3), the hadronic matrix element of the vector and axial currents ($V^\mu = \bar{c}\,\gamma^\mu\, b$ and $A^\mu = \bar{c}\,\gamma^\mu\gamma_5\, b$) can

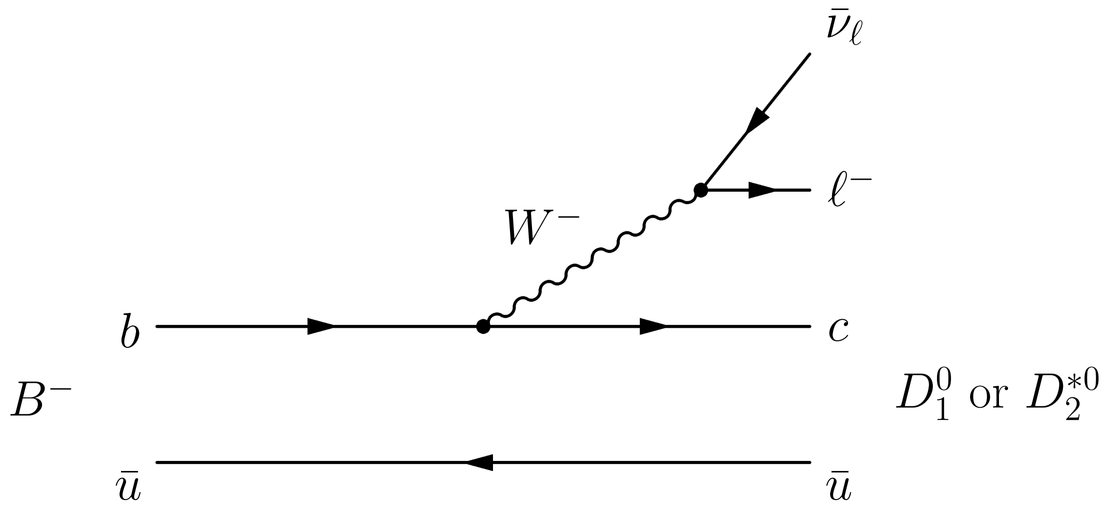


Figure 3.3: Feynman diagram for $B^- \to D_1^0 \ell^- \bar{\nu}_\ell$ and $B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell$.

be parameterized as [36]:

$$
\begin{aligned}
\langle D_1(v',\epsilon)|V^\mu|B(v)\rangle &= \sqrt{m_{D_1} m_B}\,[f_{V_1}\epsilon^{*\mu} + (f_{V_2}v^\mu + f_{V_3}v'^\mu)(\epsilon^* \cdot v)]\,, \\
\langle D_1(v',\epsilon)|A^\mu|B(v)\rangle &= \sqrt{m_{D_1} m_B}\, i\, f_A\, \varepsilon^{\mu\alpha\beta\gamma}\epsilon^*_\alpha v_\beta v'_\gamma\,, \\
\langle D_2^*(v',\epsilon)|A^\mu|B(v)\rangle &= \sqrt{m_{D_2^*} m_B}\,[k_{A_1}\epsilon^{*\mu\alpha}v_\alpha + (k_{A_2}v^\mu + k_{A_3}v'^\mu)\epsilon^*_{\alpha\beta}v^\alpha v^\beta]\,, \\
\langle D_2^*(v',\epsilon)|V^\mu|B(v)\rangle &= \sqrt{m_{D_2^*} m_B}\, i\, k_V\, \varepsilon^{\mu\alpha\beta\gamma}\epsilon^*_{\alpha\sigma} v^\sigma v_\beta v'_\gamma\,, \qquad (3.10)
\end{aligned}
$$

where the form factors f_i and k_i are dimensionless functions of w (or q^2). The four velocity of the B^- (D_J^0) is v (v') and the polarization of the D_J^0 is ϵ^*. The differential decay rates for $B^- \to D_1^0 \ell^- \bar{\nu}_\ell$ and $B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell$ can be written in terms of the form factors f_i and k_i. With $r_1 = m_{D_1}/m_B$ and $r_2 = m_{D_2^*}/m_B$, one can write:

$$
\begin{aligned}
\frac{d\Gamma_1}{dw} &= \frac{d\Gamma(\bar{B} \to D_1 \ell \bar{\nu}_\ell)}{dw} \\
&= \frac{G_F^2 |V_{cb}|^2 m_B^5 r_1^3}{48\pi^3}\sqrt{w^2-1}\left\{2(1-2wr_1+r_1^2)[f_{V_1}^2 + (w^2-1)f_A^2]\right. \\
&\quad \left. +[(w-r_1)f_{V_1} + (w^2-1)(f_{V_3}+r_1 f_{V_2})]^2\right\}\,, \qquad (3.11)
\end{aligned}
$$

and

$$
\frac{d\Gamma_2}{dw} = \frac{d\Gamma(\bar{B} \to D_2^* \ell \bar{\nu}_\ell)}{dw}
$$

$$= \frac{G_F^2|V_{cb}|^2 m_B^5 r_2^3}{144\pi^3}(w^2-1)^{3/2}\Big\{3(1-2wr_2+r_2^2)[k_{A_1}^2+(w^2-1)k_V^2] \\ +2[(w-r_2)k_{A_1}+(w^2-1)(k_{A_3}+r_2k_{A_2})]^2\Big\} . \quad (3.12)$$

Based on the equations for the differential decay rate, one can see that the dynamics of the semileptonic decays fully depend on the q^2 dependence of the form factors. In theoretical calculations, the form factors are normally computed for some particular q^2 and extrapolated to other values of q^2. The variation of the form factors with q^2 described the nonperturbative QCD physics. To understand the information contained in the form factors, one must therefore understand the decay dynamics of mesons containing a single heavy quark.

3.3 Heavy Quark Effective Theory

In the limit of an infinite mass quark, the Heavy Quark Effective Theory (HQET) makes definite predictions about the properties and decays of mesons containing one heavy quark [3, 37, 38]. By heavy quark, we mean a quark with a mass $m_Q \gg \Lambda_{\rm QCD}$ and a Compton wavelength $\lambda_Q \sim 1/m_Q \ll 1/\Lambda_{\rm QCD}$. The heavy quark in a heavy-light meson rest frame moves nonrelativistically with momentum of the order of $\Lambda_{\rm QCD}$. The scale of the typical momenta exchange between the heavy and light constituents is set by the size of a typical hadron $R_{\rm hadron} \simeq 1/\Lambda_{\rm QCD}$. Then, the soft gluons, which keep the mesons in a bound state, are only able to resolve distances much larger than λ_Q. This means that, in the limit of $m_Q \to \infty$, the soft gluons which couple to the light degrees of freedom are not able to probe the quantum numbers of the heavy quark. In other words, the light degrees of freedom of a heavy-light meson are blind to the flavor and spin orientation of the heavy quark.

With the approximate spin-flavor symmetry highlighted by the Heavy Quark Symmetry (HQS), useful descriptions of heavy quark systems can be made. By the HQS criteria, the top, bottom, and charm quarks are heavy; and the strange, down, and up quarks are light. Hadronic systems such as the B, $D^{(*)}$, and D_J mesons can therefore be studied in the limit of HQS. As $m_Q \to \infty$, the heavy quark and the meson have the same velocity causing the shape and normalization of the wave function of the light degrees of freedom to be independent of the mass and the spin of the heavy quark. However, the HQS is broken by effects of the order $\Lambda_{\rm QCD}/m_Q$ because the mass of the heavy quarks are not truly infinite. HQET leads to an operator product expansion of the Lagrangian as a series of local higher dimension operators multiplied by powers of $\Lambda_{\rm QCD}/m_Q$. Consequently, the effective QCD Lagrangian in HQET is a systematic expansion and it is possible to treat the $\Lambda_{\rm QCD}/m_Q$ terms as corrections to the prediction based on the infinite mass limit.

3.3.1 P-wave Charm Mesons

As discussed in Section 2.1, there are four states of D_J mesons. Parity and angular momentum conservation restrict the decays available to the four states. The 2^+ state can decay to either $D\pi$ or $D^*\pi$ through D-wave decays and both 1^+ states can decay only to $D^*\pi$ through S or D-wave decay, while the 0^+ state can decay only to $D\pi$ and must go through S-wave decay.

In a heavy-light meson such as a D_J meson, the typical velocity of the heavy quark is $|\mathbf{v}_Q| = |\mathbf{p}_Q|/m_Q \sim \Lambda_{\rm QCD}/m_Q$. In the limit of infinite heavy quark mass, the quark Q is essentially a stationary source of color field, and the spin of the heavy quark S_Q and the spin of the light quark S_q decouple. In this configuration, the D_J mesons are described by the total angular momentum of the light constituents $j = S_q + L$ and the total angular momentum of the meson $J = j + S_Q$. The D_J mesons then make up two doublets, $j = 1/2$ and $j = 3/2$.

According to the approximate spin-flavor symmetry in HQET, the total angular momentum of the light constituents is a conserved quantity, and the members of the $j = 3/2$ doublet are predicted to decay only in a D-wave and to be relatively narrow. The $j = 1/2$ mesons are predicted to decay only in an S-wave and to be relatively broad. In this analysis we study the semileptonic decays of the B meson to final states containing the narrow (j=3/2) excited charm mesons: the ${}^jL_J = {}^{3/2}P_2$ and ${}^{3/2}P_1$, labeled by the Particle Data Group and referred to here as the D_2^* and D_1, respectively.

When the D_J meson decays to $D^*\pi$, the helicity of the D^* is used to describe

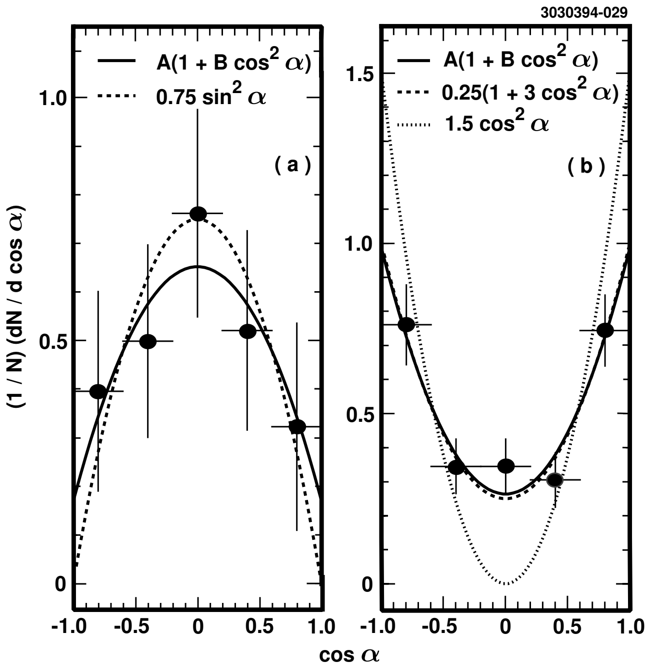


Figure 3.4: Measured helicity distribution $\cos\alpha$ for the decays (a) $D_2^{*0} \to D^{*+}\pi^-$ and (b) $D_1^0 \to D^{*+}\pi^-$. Data taken from [39].

the dynamics of the decay. Regardless of the initial polarization of the D_J, the following helicity angular distribution are predicted:

$$\frac{dN}{d\cos\alpha} \propto \begin{cases} 1 + 3\cos^2\alpha & (D_1 \text{ state with pure D-wave}) \\ \sin^2\alpha & (D_2^* \text{ state}) \end{cases} \tag{3.13}$$

Experimental results from CLEO II confirmed the predictions of HQET [39]. The $\cos\alpha$ distributions of the D_1 and D_2^* states agree with the expected $1 + 3\cos^2\alpha$ and $\sin^2\alpha$ distributions, as shown in Figure 3.4.

As noted above, the ${}^{3/2}P_2$ state can decay to both $D\pi$ and $D^*\pi$. HQET relates the decay rates of $D_2^* \to D\pi$ and $D_2^* \to D^*\pi$. When phase space and barrier-penetration corrections are included in the calculation (since $D\pi$ and $D^*\pi$ have different contribution depending if the invariant mass is close to threshold or not), models [40, 41] predict

$$R = \frac{\mathcal{B}(D_2^* \to D\pi)}{\mathcal{B}(D_2^* \to D^*\pi)} = 1.5 \;\; \text{to} \;\; 3.0\,. \tag{3.14}$$

CLEO II measured the ratio R [39]:

$$R = 2.2 \pm 0.7 \pm \pm 0.6\,. \tag{3.15}$$

The experimental result agrees well with the HQET prediction.

3.3.2 *B* Semileptonic Decays

Theoretical predictions of the transition matrix for semileptonic decays are complicated by the strong interaction. Some QCD effects are not calculable perturbatively and are parameterized in terms of a set of form factors which describe how the strong interaction modifies the weak process underlying the semileptonic decay. The spin-flavor symmetry in HQET simplifies the description of the heavy quark decay by reducing the number of independent form factors. HQET provides relations between the form factors and gives estimation for the size of the HQS breaking effects.

In a heavy quark to heavy quark decay $Q \to Q'$, both quarks have the same static color field properties. In such a decay, the configuration of the light degrees of freedom does not change if the quark $Q(v_Q, s_Q)$ with velocity v_Q and spin s_Q is replaced by the quark $Q'(v_Q, s_{Q'})$ with different flavor or spin, but with the same velocity. Even though the initial and final quarks have different masses (but are both sufficiently heavy), important simplifications occur. The form factors can be related to a *universal* form factor [38] called the Isgur-Wise function. In HQET, each doublet of charm mesons has its form factors described by its own Isgur-Wise function.

3.3.3 $\bar{B} \to D^{(*)}\ell\bar{\nu}_\ell$ Decays

In HQET, B semileptonic decays to D and D^* are related to a single universal Isgur-Wise function $\xi(w)$. The differential decay rate is given by [42]:

$$\frac{d\Gamma(\bar{B} \to D\ell\bar{\nu}_\ell)}{dw} = \frac{G_F^2 |V_{cb}|^2 m_B^5 r^3}{48\pi^3}(1+r)^2(w^2-1)^{3/2}\mathcal{F}_D^2(w)\,, \tag{3.16}$$

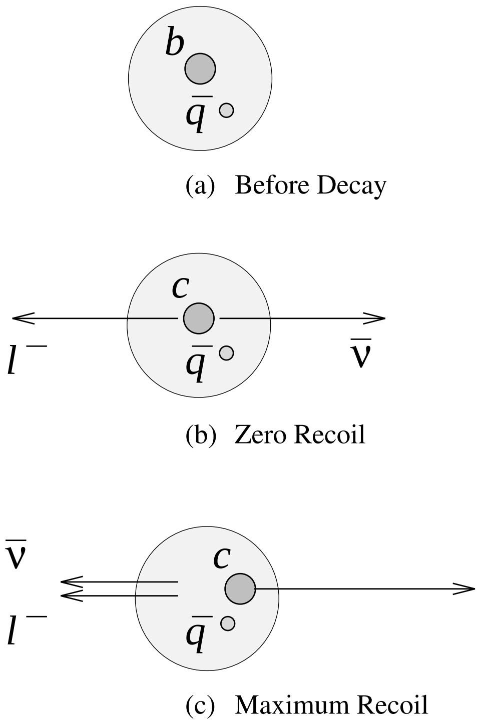


Figure 3.5: Kinematic configurations for the semileptonic decay of a $\bar{B}$ meson from Reference [17]: (a) before the $\bar{B}$ decay, (b) decay configuration at zero recoil, and (c) at maximum recoil.

and

$$\frac{d\Gamma(\bar{B} \to D^* \ell \bar{\nu}_\ell)}{dw} = \frac{G_F^2 |V_{cb}|^2 m_B^5 r^3}{48\pi^3}(1-r)^2\sqrt{w^2-1}\mathcal{F}_{D^*}^2(w) \times \left[4w(w+1)\frac{1-2wr+r^2}{(1-r)^2} + (w+1)^2\right], \quad (3.17)$$

where $r = m_{D^{(*)}}/m_B$. In the limit of infinitely heavy quarks, $\mathcal{F}_{D^{(*)}}(w)$ becomes $\eta\,\xi(w)$, and HQET predicts that $\xi(1) = 1$. The parameter η is a perturbatively calculable QCD correction. The normalization to unity of the Isgur-Wise function at zero recoil arises because there is a full overlap of the wave-function of the light degrees of freedom at $w = (v \cdot v') = 1$. At the point of equal velocity $v = v'$, the light constituents have identical configurations before and after the weak decay of the heavy quark causing no gluon exchange between the light and the heavy quark. Figure 3.5 shows the kinematic configurations for the semileptonic decay of a B meson.

For finite heavy quark masses, $\mathcal{F}_{D^{(*)}}(1)$ can be estimated in the framework of HQET [43]. In the case of $B \to D^* \ell \bar{\nu}_\ell$, it has been shown [44] that corrections of order $\Lambda_{\rm QCD}/m_Q$ are identically zero at $w = 1$, and that the leading corrections arise at order $(\Lambda_{\rm QCD}/m_Q)^2$. The second-power corrections to $\mathcal{F}_{D^*}(1)$ are then parametrically suppressed and can be estimated with an accuracy of better than 4% [3, 43]. In the case of $B \to D \ell \bar{\nu}_\ell$, the form factors are not protected against $\Lambda_{\rm QCD}/m_Q$ corrections, but such corrections turn out to be small and calculable [45].

Theoretically, $\bar{B} \to D^{(*)} \ell \bar{\nu}_\ell$ is perfectly suited for the determination of the CKM element $|V_{cb}|$. The normalization of the hadronic form factors at zero recoil allows a model independent measurement of $\mathcal{F}_{D^{(*)}}(1)\,|V_{cb}|$ [28, 46, 47]. But as mentioned before, the q^2 (or w) dependence of the Isgur-Wise function must be determined separately using nonperturbative techniques. Even if HQET does not predict the form of $\xi(w)$, it relates all the form factors to them. In the framework of HQET, the differential rates and the kinematic variables of $B \to D^{(*)} \ell \bar{\nu}_\ell$ provide information on the shape of $\xi(w)$, which then provide information on nonperturbative QCD effects [47, 48]. Measurements of $|V_{cb}|$ and form factor studies at CLEO II and elsewhere are reviewed in great detail in Reference [17].

3.3.4 $\bar{B} \to D_J \ell \bar{\nu}_\ell$ Decays

The use of HQET resulted in a dramatic improvement in our understanding of exclusive decays $B \to D^{(*)} \ell \bar{\nu}_\ell$. In the infinite heavy quark limit, even the semileptonic decay form factors of B meson into either D_J doublet are given by just one Isgur-Wise function. The leading Isgur-Wise functions for the orbitally excited charm mesons $j = 1/2$ and $j = 3/2$ are denoted by $\zeta(w)$ and $\tau(w)$ respectively. The form factors f_i and k_i of Equations (3.11) and (3.12) are functions of the leading Isgur-Wise function $\tau(w)$. We used the term *leading* Isgur-Wise function only to distinguish $\tau(w)$ from the functions that appear at order $\Lambda_{\rm QCD}/m_Q$, which are sometimes called *subleading* Isgur-Wise functions. The differential decay rate for $B^- \to D_1^0 \ell^- \bar{\nu}_\ell$ and $B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell$ can be written as [42]:

$$\begin{aligned}\frac{d\Gamma_1}{dw} &= \frac{G_F^2|V_{cb}|^2 m_B^5 r_1^3}{48\pi^3}(w^2-1)^{3/2}\,\mathcal{F}_{D_1}^2(w)\times \\ &\frac{2}{3}(w+1)\left[(w-1)(1+r_1)^2+w(1-2wr_1+r_1^2)\right]\,, \qquad (3.18)\end{aligned}$$

and

$$\begin{aligned}\frac{d\Gamma_2}{dw} &= \frac{G_F^2|V_{cb}|^2 m_B^5 r_2^3}{48\pi^3}(w^2-1)^{3/2}\,\mathcal{F}_{D_2^*}^2(w)\times \\ &\frac{2}{3}(w+1)\left[(w+1)(1-r_2)^2+3w(1-2wr_2+r_2^2)\right]\,, \qquad (3.19)\end{aligned}$$

where the functions $\mathcal{F}_{D_1}(w)$ and $\mathcal{F}_{D_2^*}(w)$ can be related to the leading Isgur-Wise function $\tau(w)$ and the subleading Isgur-Wise functions.

Models are needed to extrapolate the form factors for all kinematically allowed q^2. Several models try to calculate the q^2 dependence of the form factors within the HQS prescriptions. In the finite mass limit, there are calculable $\Lambda_{\rm QCD}/m_Q$ corrections to the QCD effective Lagrangian in HQET. At zero recoil, the finite mass $\Lambda_{\rm QCD}/m_Q$ corrections to the transition matrix elements can be written in terms of the leading Isgur-Wise function and the meson mass splittings. Away from zero recoil, there are perturbative α_s and nonperturbative $\Lambda_{\rm QCD}/m_Q$ corrections to the prediction of the $m_Q \to \infty$ limit. Theoretical predictions exist for both the infinite and finite heavy quark limits. They are discussed in the next sections.

Measurements of the $B^- \to D_1^0\ell^-\bar{\nu}_\ell$ and $B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell$ branching fractions allow a test of the various theoretical predictions. With data from semileptonic decays, one can then compare:

1. the different form factors of a given mode and check the HQET relations among them;

2. the different B semileptonic modes to each other.

A useful variable to look at is:

$$\mathcal{R} = \frac{\mathcal{B}(B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell)}{\mathcal{B}(B^- \to D_1^0\ell^-\bar{\nu}_\ell)}. \qquad (3.20)$$

Precise measurements of the differential decay rate $d\Gamma_{1,2}/dw$ would determine (in principle) the shape of $\mathcal{F}_{D_1}(w)$ and $\mathcal{F}_{D_2^*}(w)$, which provide insights on the nonperturbative QCD dynamics of the decay. Unfortunately, precise measurements require large and clean samples of $B^- \to D_1^0 \ell^- \bar{\nu}_\ell$ and $B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell$ decays. Nevertheless, even a crude measurement of $\mathcal{R}$ would give us some handle on the breaking of HQS.

Finally, the ultimate theoretical goal is to relate the inclusive decays to the sum of the exclusive channels. In the infinite mass limit, the inclusive sum of the probabilities for semileptonic decays into hadronic states is equal to the probability for the free quark transition. The Bjorken sum rule [49, 50] is obtained by expanding this probability in powers of $(w-1)$ and keeping terms of first order only

$$\rho^2 = \frac{1}{4} + \sum_m \frac{|\zeta^{(m)}(1)|^2}{4} + 2\sum_n \frac{|\tau^{(n)}(1)|^2}{3}, \tag{3.21}$$

where ρ^2 is the slope of the universal Isgur-Wise function for $\bar{B} \to D^{(*)}\ell\bar{\nu}_\ell$: $\xi(w) = 1 - \rho^2(w-1) + \mathcal{O}[(w-1)^2]$. The subscript m and n are the radially excited indices; $m = n = 0$ correspond to the orbitally excited states (D_1, D_2^*) and (D_0^*, D_1^*). The experimental observations of semileptonic B decays into excited D_J mesons suggest $\rho^2 > 1/4$.

3.3.5 Dynamics of $\bar{B} \to D_J \ell \bar{\nu}_\ell$ Revisited

Now that we know a little more about semileptonic decays and HQET, let's go back and try to understand the dynamics of the semileptonic decay of a B meson into an orbitally excited meson in more detail.

First we note that when the B meson decays to $D_J W^*$, the light quark gains one unit of angular momentum. This implies that the hadronic matrix element $\langle D_J(v', \epsilon)|(V^\mu - A^\mu)|B(v)\rangle$ vanishes at zero recoil for any $\tau(1)$ since the B meson and the (D_1, D_2^*) mesons are in different heavy quark spin symmetry multiplets. Put another way, at zero recoil, the wave functions of the light degrees of freedom before and after the action of the weak current are totally orthogonal. Consequently, we expect the D_J wave function to have a larger overlap with a state in which the daughter c quark has a large velocity relative to the spectator quark. In such circumstances, the typical value of q^2 must be pushed significantly lower before a D_J can be produced. Remember that the charm quark gets a large relative velocity at low q^2. Given that low q^2 production should be enhanced by this effect, it is clear that we should get a lot of helicity zero D_J, since at $q^2 = 0$ the lepton and neutrino are collinear with a net spin of zero along the decay axis. This forces the daughter meson to have helicity zero, regardless of its spin, since the B meson is spinless. As a consequence, in HQET, the helicity zero contribution is expected to make up the majority of the total decay rate $\Gamma_{1,2}$.

By looking at Equation (3.10), we should notice that at zero recoil ($v = v'$ configuration) only the form factors f_{V_1} and k_{A_1} can contribute to the rates. This is easily explained by the fact that v' dotted into the polarization $\epsilon^{*\mu}$ or $\epsilon^{*\mu\alpha}$ vanishes, which implies that near zero recoil one form factor dominates each decay rate.

The complete details of the dynamics of the decay depends on the q^2 dependence of the form factors, and thus depends on some theoretical assumptions about the form of the heavy-light meson wave functions and of the Isgur-Wise functions. The importance of the corrections from the infinite heavy quark mass limit is more model dependent for $B \to D_J \ell \bar{\nu}_\ell$ than for $B \to D^{(*)} \ell \bar{\nu}_\ell$ because at no kinematic point can the leading Isgur-Wise function $\tau(w)$ be normalized. The leading $\Lambda_{\mathrm{QCD}}/m_Q$ corrections can nevertheless be divided into two classes: corrections to the current itself and corrections to the states. In HQET, a nonlocal effective *action* is derived [3] which allows an expansion in powers of $\Lambda_{\mathrm{QCD}}/m_Q$ for the external current and the QCD Lagrangian (containing the heavy quark spinor field). This effective theory therefore models the long-distance physics of QCD accurately. However, it cannot describe the short-range physics. The heavy quark participates in strong interactions through its coupling to gluons with momenta in the range $\Lambda_{\mathrm{QCD}} < \mu < m_Q$. The effective theory of HQET provides an appropriate description at long distance scales $\mu \ll m_Q$; but at large scales, α_s is small and thus perturbative QCD can be used to compute the short distance effects. However, perturbative calculations for each $\mathcal{O}(\alpha_s)$ diverge logarithmically as $m_Q \to \infty$. So to rearrange (renormalize) the perturbative contributions to be finite at each order of $\Lambda_{\mathrm{QCD}}/m_Q$, these must be matched to the calculations of

Parameter	ISGW2
b	0.18 GeV^2
c	-0.81 GeV
α_s	$0.60 \rightarrow 0.30$
$m_u = m_d$	0.33 GeV
m_s	0.55 GeV
m_c	1.82 GeV
m_b	5.20 GeV

Table 3.1: Parameters of the constituent quark potential model ISGW2.

the effective theory at some intermediate mass scale $\Lambda_{\text{QCD}} < \mu < m_Q$. The calculation of such matching corrections is in principle straightforward, but rather tedious in practice. The influence of perturbative QCD corrections in B to charm meson semileptonic decays is believed to be small [3, 51, 52]. In practice, each model has its own way to deal with the details beyond the HQET limit.

3.4 The ISGW2 Model

The ISGW2 [51] is a form factor model based on HQET. It is an improved version of the original ISGW model of Isgur, Scora, Grinteins, and Wise [22]. ISGW2's calculation for semileptonic meson decays is based on a nonrelativistic constituent quark potential model, with an assumed Coulomb plus linear potential:

$$V(r) = -\frac{4\alpha_s}{3r} + c + br\,. \tag{3.22}$$

The quark model parameters used by ISGW2 are summarized in Table 3.1. ISGW2 incorporates the HQS constraints between the form factors and on the slopes of the form factors near zero recoil. Matching requirements of HQET are also included in the ISGW2 calculations. ISGW2 is consistent with the restriction of HQS breaking at the order of Λ_{QCD}/m_Q. It includes the two leading order breaking effects of HQS:

the heavy quark kinematic energy which breaks the flavor symmetry, and the color magnetic moment interaction of the heavy quark with the color field (or chromomagnetic effects) which breaks both the spin and the flavor symmetry. Such a calculation is expected to be reliable near zero recoil where the mesons, and thus the individual quarks, have small momenta. Ad hoc relativistic corrections are included to better describe the dynamics of semileptonic decays at larger recoil.

ISGW2 used Schrödinger harmonic oscillator (SHO) wave functions to describe the quark position wave functions [51]. The wave function variational parameters are denoted by β. The parameter β_B describes the B meson and β_X the meson in the final state. The form factors for the (D_1, D_2^*) doublet are related to the transition matrix by:

$$\begin{aligned}
\langle D_1(p_X,\epsilon)|V^\mu|B(v)\rangle &= r\epsilon^{*\mu} + [s_+(p_B+p_X)^\mu + s_-(p_B-p_X)^\mu](\epsilon^*\cdot p_B)\,, \qquad (3.23)\\
\langle D_1(p_X,\epsilon)|A^\mu|B(v)\rangle &= i\,v\,\varepsilon^{\mu\alpha\beta\gamma}\epsilon^*_\alpha(p_B+p_X)_\beta(p_B+p_X)_\gamma\,,\\
\langle D_2^*(p_X,\epsilon)|A^\mu|B(v)\rangle &= k\,\epsilon^{*\mu\nu}p_{B\,\nu} + [b_+(p_B+p_X)^\mu + b_-(p_B-p_X)^\mu]\epsilon^*_{\alpha\beta}p_B^\alpha p_B^\beta\,,\\
\langle D_2^*(p_X,\epsilon)|V^\mu|B(v)\rangle &= i\,h\,\varepsilon^{\mu\alpha\beta\gamma}\epsilon^*_{\alpha\sigma}p_B^\sigma(p_B+p_X)_\beta(p_B-p_X)_\gamma\,,
\end{aligned}$$

where the momentum of the mesons are $p_X = m_X v'$ and $p_B = m_B v$. The form factors for the D_1 are given by:

$$\begin{aligned}
v &= \left[\frac{\tilde{m}_B\beta_B}{4\sqrt{2}m_b m_c \tilde{m}_X} + \frac{(\tilde{w}-1)}{6\sqrt{2}}\frac{m_d}{\tilde{m}_X\beta_B}\right]F_5^{(v)}, \qquad (3.24)\\
r &= \frac{\tilde{m}_B\beta_B}{\sqrt{2}}\left[\frac{1}{\mu_+} + \frac{m_d\tilde{m}_X}{3m_c\beta_B^2}(\tilde{w}-1)^2\right]F_5^{(r)},\\
s_+ + s_- &= \frac{m_d}{\sqrt{2}\tilde{m}_B\beta_B}\left[1 - \frac{m_d}{m_c} + \frac{m_d}{2\mu_+}\frac{\beta_B^2}{\beta_{BX}^2}\right]F_5^{(s_+ + s_-)},\\
s_+ - s_- &= \frac{m_d}{\sqrt{2}m_c\beta_B}\left[\frac{(4-\tilde{w})}{3} - \frac{m_d m_c}{2\tilde{m}_X\mu_+}\frac{\beta_B^2}{\beta_{BX}^2}\right]F_5^{(s_+ - s_-)},
\end{aligned}$$

and for the D_2^* by:

$$\begin{aligned}
h &= \frac{m_d}{2\sqrt{2}\tilde{m}_B\beta_B}\left[\frac{1}{m_c} - \frac{m_d\beta_B^2}{2\mu_-\tilde{m}_X\beta_{BX}^2}\right]F_5^{(h)}, \qquad (3.25)\\
k &= \frac{m_d}{\sqrt{2}\beta_B}(\tilde{w}+1)F_5^{(k)},
\end{aligned}$$

$$b_+ + b_- = \frac{m_d^2}{4\sqrt{2} m_c m_b \tilde{m}_B \beta_B} \frac{\beta_X^2}{\beta_{BX}^2} \left[1 - \frac{m_d}{2\tilde{m}_B} \frac{\beta_X^2}{\beta_{BX}^2} \right] F_5^{(b_+ + b_-)},$$
$$b_+ - b_- = \frac{-m_d}{\sqrt{2} m_b \tilde{m}_X \beta_B} \left[1 - \frac{m_d m_b}{2\mu_+ \tilde{m}_B} \frac{\beta_X^2}{\beta_{BX}^2} + \frac{m_d}{4m_c} \frac{\beta_X^2}{\beta_{BX}^2} \left(1 - \frac{m_d}{2\tilde{m}_B} \frac{\beta_X^2}{\beta_{BX}^2} \right) \right] F_5^{(b_+ - b_-)}.$$

The tilde parameters are $\tilde{m}_B = m_b + m_d$, $\tilde{m}_X = m_c + m_d$ and $(\tilde{w} - 1) = (q^2_{\max} - q^2)/(2\bar{m}_B \bar{m}_X)$. The physical spin-weighted mass averages are $\bar{m}_B = \frac{1}{4} m_B + \frac{3}{4} m_{B^*}$ and $\bar{m}_X = \frac{3}{8} m_{D_1} + \frac{5}{8} m_{D_2^*}$. The factors $\mu_\pm = (1/m_c \pm 1/m_b)^{-1}$ and $\beta_{BX}^2 = \frac{1}{2}(\beta_B^2 + \beta_X^2)$ The ISGW2 form factors in the equations above are related to the f_i and k_i by simple transformations [53].

As one can see, all the form factors for the doublet (D_1, D_2^*) are proportional to a universal function $F_5^{(\alpha)}$

$$F_5^{(\alpha)} = \left(\frac{\bar{m}_B}{\tilde{m}_B} \right)^{n_B(\alpha)} \left(\frac{\bar{m}_X}{\tilde{m}_X} \right)^{n_X(\alpha)} \left(\frac{\tilde{m}_X}{\tilde{m}_B} \right)^{\frac{1}{2}} \left(\frac{\beta_B \beta_X}{\beta_{BX}^2} \right)^{\frac{5}{2}} \left[1 + \frac{r_{BX}^2}{12} (q^2_{\max} - q^2) \right]^{-2}, \quad (3.26)$$

where r_{BX}^2 is called the charge radius of the decay; which is typically $r_{BX}^2 \simeq 0.2$ GeV$^{-2}c^4$ when X is a charm meson. The parameters $n_B(\alpha)$ and $n_X(\alpha)$ are half-integers; their values depend on the form factor α (see Table XIII in [51] for details). The q^2 and lepton momentum distributions obtained with the ISGW2 model for $B^- \to D_1^0 \ell^- \bar{\nu}_\ell$ and $B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell$ are shown in Figure 3.6. The areas of the distributions are proportional to the decay rates predicted. Monte Carlo simulations at CLEO rely mainly on the ISGW2 model to describe the dynamics of exclusive B semileptonic decays [54].

3.5 More Models

3.5.1 SISM and VO Models

Both the model of Suzuki, Ito, Sawada, and Matsuda (SISM) [55], and the model of Veseli and Olsson (VO) [42] explicitly incorporate the principles of HQET in their calculations. They do not calculate higher order Λ_{QCD}/m_Q corrections beyond the HQS prescription. They assume the bottom and charm quarks heavy enough compared to the QCD scale parameter Λ_{QCD} so that there are very little QCD interactions in the

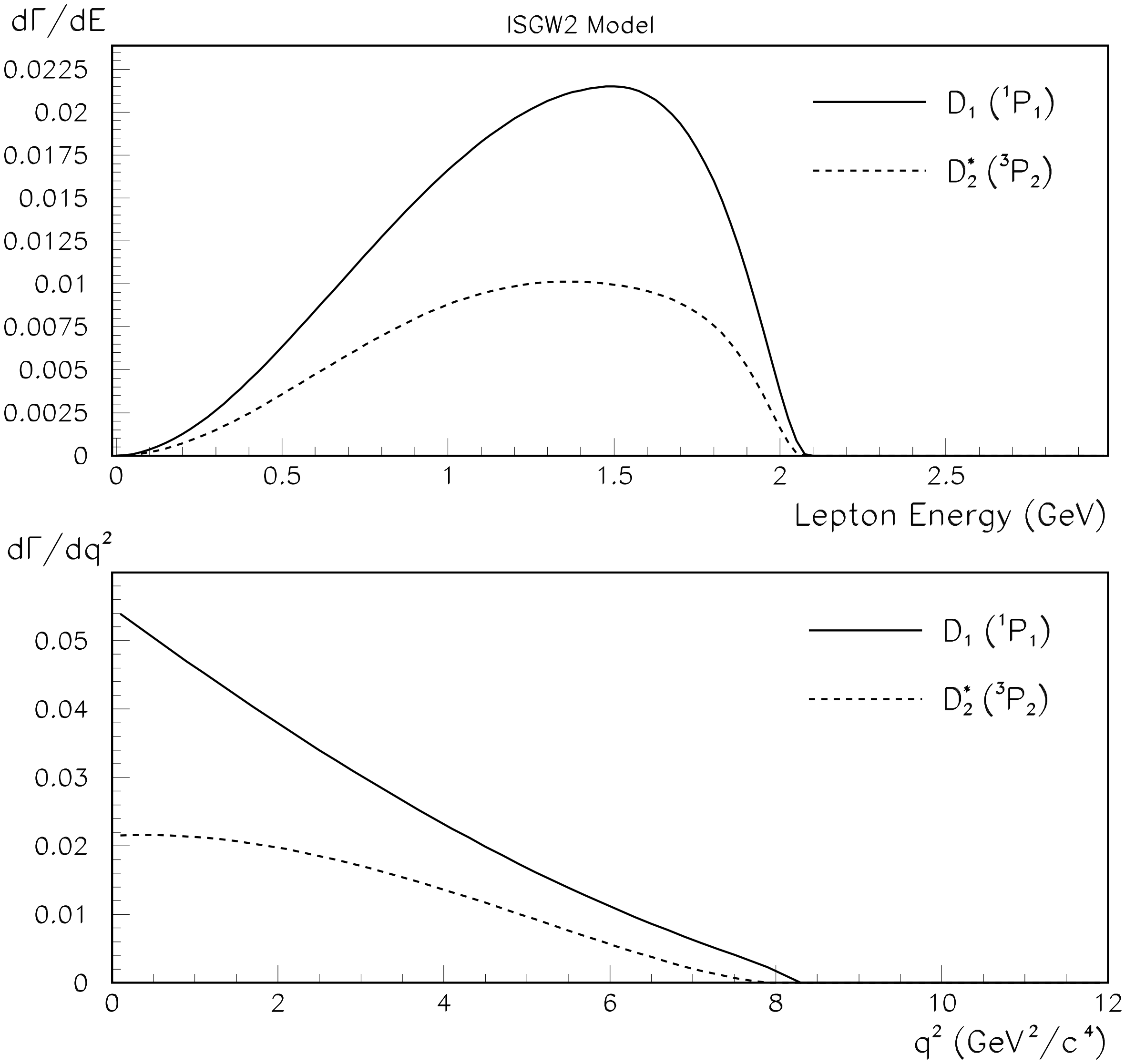


Figure 3.6: The predicted $B^- \to D_1^0 \ell^- \bar{\nu}_\ell$ and $B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell$ lepton energy and q^2 spectra for the ISGW2 model. The areas of the distributions are proportional to the decay rates; hence, $\Gamma_1 \simeq 2\,\Gamma_2$ in ISGW2.

heavy flavored meson. In the $m_Q \to \infty$ limit, the decay rates are simply given by Equations (3.18) and (3.19) with $\mathcal{F}_{D_1}(w) = \mathcal{F}_{D_2^*}(w) = \tau(w)$.

The authors of the SISM model use a similar method to the one developed by Isgur, Scora, Grinstien, and Wise [22]. They assume that the constituent quarks are bounded by the potential of Equation (3.22) with $b = 0.1568$ GeV^2, $c = -0.563$ GeV, and a fix value of $\alpha_s = 0.5$. The matching conditions of HQET are not part of their calculation. They approximate the relativistic recoil effects with a compensation factor κ as it was prescribed in the original ISGW model. Their leading Isgur-Wise function has the form:

$$\tau(w) = \frac{1}{\sqrt{2}} \left(\frac{2\beta_{PD}\beta_{SB}}{\beta_{PD}^2 + \beta_{SB}^2} \right)^{5/2} \exp\left[\frac{-\Lambda^2/\kappa^2}{2(\beta_{PD}^2 + \beta_{SB}^2)} (w^2 - 1) \right] \frac{\Lambda}{\beta_{SB}} \sqrt{w^2 - 1}\,, \qquad (3.27)$$

where $\Lambda = m_X m_d/(m_c + m_d)$. The βs are their SHO wave function parameters.

Veseli and Olsson employ three qualitatively different hadronic models to describe the light degrees of freedom: the Dirac equation with scalar confinement (DESC), the Salpeter equation with vector confinement (SEVC), and the relativistic flux tube confinement (RFTC) [42]. All three models involve a short range Coulomb potential with a fixed α_s. The various parameters are chosen to best fit the heavy-light data. The DESC, SEVC, and RFTC methods are all consistent with each other.

3.5.2 CNP Model

Colangelo, Nardulli, and Paver (CNP) use QCD sum rules to evaluate the B semileptonic transition matrices to excited charm mesons [56]. The QCD sum rules provide an independent approach to the evaluation of these matrices. In the CNP approach, the infinite heavy quark mass limit is taken. The procedure to obtain the sum rules is standard [57], but we do not intend to discuss it in detail. First, they take m_Q to be finite. Then, they estimate the sum rules and perform the limit $m_Q \to \infty$. They compute the sum rules (and thus the form factors) at $q^2 = 0$ where it is easier to include perturbative $\mathcal{O}(\alpha_s)$ corrections. They point out that $\Lambda_{\mathrm{QCD}}/m_Q$ corrections may turn out to be important, but, nevertheless, use the heavy quark limit in their estimation of the decay rates.

3.5.3 SHJL Model

While ISGW2, SISM, and VO all used nonrelativistic quark models, Sutherland, Holdom, Jaimungal, and Lewis (SHJL) [58] employ a relativistic quark model. The only model parameters are some momentum dependent quark mass functions. Their approach is based on a relativistic field theory with the flavor and spin symmetries of QCD. The interaction vertices between the heavy quark and the light degrees of freedom are defined as a product of γ-matrix structure and a damping factor. The structure of the interaction vertices is chosen to be identically equal to the form determined by HQS. The vertex form factors take the place of the Coulomb plus linear potential. Their model does not rely on any expansion in $\Lambda_{\mathrm{QCD}}/m_Q$, but it somehow satisfies the requirements of HQET. In earlier publications, they explicitly checked that the model satisfies HQET constraints [59].

The interaction Lagrangian of SHJL includes the light-quark triplet (u, d, s) and triplets of heavy-light meson fields with the heavy quark Q $(Q = c, b)$. Perturbative α_s corrections to the currents could be added, but this has not been done. On the other hand, the model makes predictions for all possible terms which break HQS, including the heavy quark kinematic energy and chromomagnetic effects. SHJL believe that nonresonant contributions in B semileptonic decays are not negligible. They give predictions for the rates in both the infinite and the finite heavy quark mass limit. Their results suggest nonperturbative departures from the heavy quark limit.

3.5.4 LLSW Model

As mentioned before, HQS implies that in the $m_Q \to \infty$ limit matrix elements of the weak currents between the B meson and an excited charm meson vanish at zero recoil. However, it is possible that at order $\Lambda_{\mathrm{QCD}}/m_Q$ these matrix elements are not zero. Leibovich, Ligeti, Stewart, and Wise (LLSW) investigated exclusive semileptonic decays into D_1 and D_2^* mesons using HQET [36]. Their calculation includes order of $\Lambda_{\mathrm{QCD}}/m_Q$ corrections. At zero recoil, the $\Lambda_{\mathrm{QCD}}/m_Q$ corrections to the transition matrix elements are written in terms of the leading Isgur-Wise function $\tau(w)$

and the known meson mass splittings. Their conclusion is that, since most of the phase space for B semileptonic decays to P-wave charm mesons is near zero recoil, $\Lambda_{\rm QCD}/m_Q$ corrections can be very important.

In HQET, the effective heavy quark field $h_\nu(x)$ is related to the original field $Q(x)$ by

$$Q(x) = \exp(-im_Q\, v_Q \cdot x)\, h_\nu(x). \tag{3.28}$$

In the finite mass limit however, the field $Q(x)$ is

$$Q(x) = \exp(-im_Q\, v_Q \cdot x) \left[1 + \frac{i\,\not{D}}{2m_Q} + \mathcal{O}\left(\frac{\Lambda_{\rm QCD}}{m_Q}\right)^2\right] h_\nu(x), \tag{3.29}$$

where D is the covariant derivative in QCD. By putting the field $Q(x)$ in the QCD Lagrangian $\mathcal{L} = \bar{Q}(i\,\not{D} - m_Q)Q$, one finds:

$$\mathcal{L} = \mathcal{L}_{\rm HQET} + \delta\mathcal{L}. \tag{3.30}$$

The LLSW calculation includes two extra terms in $\delta\mathcal{L}$. The first term deals with the heavy quark kinematic energy and the second with chromomagnetic effects. In the infinite heavy quark mass limit, all the form factors are given by the leading Isgur-Wise function $\tau(w)$. At order $\Lambda_{\rm QCD}/m_Q$ unknown subleading Isgur-Wise functions must be introduced to describe the $b \to c$ current. In the LLSW model, $\Lambda_{\rm QCD}/m_Q$ corrections to the current itself are computed and the HQS constraints on the relations between the form factors near zero recoil are respected. The corrections originating from the matching of the $b \to c$ flavor changing current onto the effective theory and the $\Lambda_{\rm QCD}/m_Q$ corrections are taken care of in a systematic way.

In a more recent publication [52], the authors separate the contributions to the rate into the different helicity decompositions of the D_1 and D_2^* mesons, since the $\Lambda_{\rm QCD}/m_Q$ corrections affect these differently. They present two different approximations to the decay rate (approximations A and B). In approximation A, they treat $(w-1)$ as order $\Lambda_{\rm QCD}/m_Q$ and expand the decay rate in these parameters. In approximation B, the known order $\Lambda_{\rm QCD}/m_Q$ contributions to the form factor is kept, as well as the full w dependence of the decay rate. The results of both approximations are comparable.

In summary, the LLSW model predicts the differential decay rates. It includes $\Lambda_{\mathrm{QCD}}/m_Q$ corrections with some model dependence on the form of the subleading Isgur-Wise functions. The model dependence is larger away from zero recoil. The HQET matching conditions are respected at large recoil with α_s corrections. They give predictions on the ratio of branching fractions $\mathcal{R}$, in which the order $\Lambda_{\mathrm{QCD}}/m_Q$ corrections turn out to be important.

3.5.5 G&R Model

The model of Goity and Roberts (G&R) [60] treats soft pion emission in B semileptonic decays in the framework of the heavy quark limit. G&R then provide a description of exclusive B semileptonic decays to nonresonant and resonant hadronic states such as $\bar{B} \to D\pi\ell\bar{\nu}_\ell$ and $\bar{B} \to D^*\pi\ell\bar{\nu}_\ell$. The various effective coupling constants and the form factors are obtained using a chiral quark model and HQET. The quark contributions to the QCD Lagrangian separates naturally into two pieces: the first contribution comes from the light quarks (u, d, s) whereas the second is due to heavy quarks (c, b, t). The light-quark sector has an approximate flavor chiral symmetry because the current quark masses are all very small on the typical hadron energy scale [61]. On the other hand, the dynamics of the heavy quark depend only on its velocity and are independent of its mass and spin. Hence, their model includes both the chiral symmetry of the light quarks and the heavy quark symmetry for low-energy meson interactions with the pion (called Goldstone boson in the the $\mathrm{SU}(3)_L \times \mathrm{SU}(3)_R$ flavor chiral symmetry).

The Feynman diagrams describing the process $\bar{B} \to D^{(*)}\pi\ell\bar{\nu}_\ell$ appear in Figure 3.7. In the $G\&R$ model, the intermediate mesons $\tilde{B}$ and $\tilde{D}$ are either the ground state meson D, D^*, B, and B^* or the excited states D^{**} and B^{**}. The lowest chiral Lagrangian is expanded to the lowest order $\mathcal{O}(p_\pi)$. The expansion places restrictions on the momentum quantum numbers of the D^{**} and B^{**}. It turns out that the well established D_1 and D_2^* states are not included in their analysis because the doublet (D_1, D_2^*) contribution appears only at higher powers of the slow pion momentum.

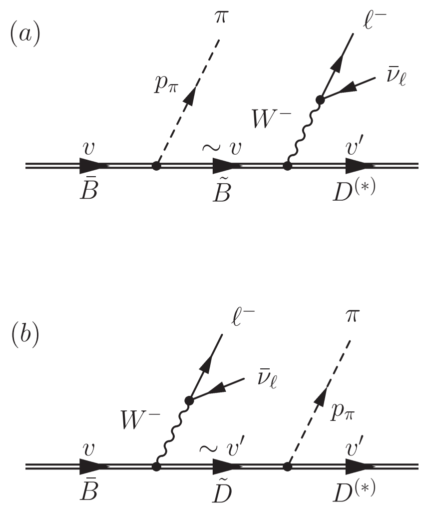


Figure 3.7: Feynman diagrams for $\bar{B} \to D^{(*)}\pi\ell\bar{\nu}_\ell$. The dashed line represents the soft pion. The mesons $\tilde{B}$ and $\tilde{D}$ are either ground states or excited states mesons.

After the expansion in p_π, the remaining states are described by a set of independent form factors. Their respective leading Isgur-Wise functions have an exponential form with no $\Lambda_{\rm QCD}/m_Q$ and no α_s corrections. Harmonic oscillator wave function solutions of the Coulomb plus linear potential are used to calculate the form factors. Similar work has been performed by other theorists; namely Cheng *et al.* [62] and Lee *et al.* [63]. The Goity and Roberts analysis is an improvement over previous work since it includes some of the radially excited states.

As mentioned earlier, the B semileptonic rate is far from being saturated by the resonant decays $\bar{B} \to D\ell\bar{\nu}_\ell$ and $\bar{B} \to D^*\ell\bar{\nu}_\ell$. Therefore, it is reasonable to assume that nonresonant decays $\bar{B} \to D^{(*)}\pi\ell\bar{\nu}_\ell$ may contribute to the inclusive rate. At CLEO, we employ a hybrid version of the standard G&R model to describe the nonresonant decays $\bar{B} \to D^{(*)}\pi\ell\bar{\nu}_\ell$ in generic B semileptonic Monte Carlo [54]. We do not use the G&R model of resonant decays because it does not include the exclusive semileptonic decays of the B meson to the D_1 or the D_2^* meson. The hybrid model only considers the diagram with a $\tilde{B}$ (see Figure 3.7(a)) since we only care about nonresonant pion emission. Doing so removes some possibly important interference terms between the two diagrams in Figure 3.7. Our approach is nevertheless believed to be adequate [64]. Appendix C presents the CLEO modification to the standard G&R model.

3.6 Summary

In the previous sections we discussed many theoretical models which attempt to describe the semileptonic decays of the B meson to excited charm mesons. Most models evaluate the decay rate and the absolute branching fraction for the exclusive process $B^- \to D_J^0 \ell^- \bar{\nu}_\ell$ on the basis of HQET. Some models also provide predictions for the ratio of branching fractions $\mathcal{R}$. It is important to compare the experimental results for $B^- \to D_1^0 \ell^- \bar{\nu}_\ell$ and $B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell$ with the predictions given by these models. Much can be learned by investigating possible discrepancies between the experimental results and the theoretical predictions.

Some theoretical models discussed earlier do not take into account perturbative QCD corrections, spectator effects, and deviations from exact HQS. According to several calculations [3, 51, 52], the influence of $\mathcal{O}(\alpha_s)$ matching corrections in B semileptonic decays to charm mesons is small. It remains to be seen whether the neglect of the $\Lambda_{\mathrm{QCD}}/m_Q$ corrections is significant in the theoretical treatment of $B^- \to D_1^0 \ell^- \bar{\nu}_\ell$ and $B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell$.

In the last chapter, a discussion on the agreement between the experimental results presented in this thesis and the theoretical calculation is given. It will be interesting to see if we can or cannot discriminate between models and at the same time learn about the level of the heavy quark symmetry breaking in HQET.

Chapter 4

Experimental Apparatus

Particle experiments study the products of collisions of high energy particles, which are generally produced at accelerators. For instance, hadrons containing bottom quarks can be produced in a wide variety of modern experiments: e^+e^- colliders, $p\bar{p}$ colliders, and fixed target experiments. In e^+e^- machines, large multi-layered detectors surround the collision point. Each layer of the detector serves a separate function in tracking and identifying each of the many particles that may be produced in a single collision. The data used in this thesis were taken with the CLEO II detector located at the Cornell Electron-positron Storage Ring (CESR). The CLEO II detector is operated by a collaboration of over 100 physicists from many institutions. A list of the CLEO collaborators and institutions is given in Appendix A. In this chapter, we will review the accelerator CESR and the detector CLEO II.

4.1 CESR

The Cornell Electron-positron Storage Ring (CESR) was constructed between 1977 and 1979, and started operating in 1979. CESR accelerates electrons and positrons to energies of 4.5 GeV to 6.0 GeV in a ring of 768 meters in circumference. The two beams are concentrated along the circular path by powerful magnets and brought together to collide in the center of the CLEO II detector. The electrons and the positrons circulate in opposite directions in the same ring of bending magnets. The beams are composed of more than 10^{12} electrons and positrons ($\sim$ 100 mA/beam), separated into bunches of a few centimeters length, traveling at almost the speed of light. Figure 4.1 shows a schematic diagram of the CESR facility.

The electrons and positrons are accelerated in a three-step process. The electrons are produced by heating a cathode, and then injected in the linear accelerator (linac) where they are accelerated to about 300 MeV. The positrons are produced by irradiating a target with electrons. When the incident 140 MeV electrons hit the target located down the linac, showers of low energy electrons and positrons are produced. The positrons are collimated by a magnetic field and accelerated to 200 MeV in the remaining length of the linac. The beams from the linac are injected into the synchrotron and accelerated to 5 GeV before being transferred to the main storage ring (see Figure 4.1). Each electron or positron spends less than one-hundredth of a second in the synchrotron, but once it is transferred to the storage ring it must coast there for several hours. This puts very strict demands on the precision of the ring magnets and on the quality of the vacuum. At the moment, CESR operates with nine trains of two bunches each. Typically, at the interaction region, a bunch is about 20 mm long, 0.3 mm wide, and 0.008 mm high. This leads to an average luminosity of roughly 10^{32} cm^{-2}sec^{-1} and a best integrated luminosity of 22 pb^{-1}/day.

4.2 An Overview of the CLEO II Detector

At CESR, the products of e^+e^- collisions are studied with a complex detection apparatus known as the CLEO II detector. The CLEO II detector, illustrated in Figures 4.2, 4.3, and 4.4, is a multipurpose high energy physics detector incorporating excellent charged and neutral particle detection and measurement. It was installed in the CESR south interaction region in 1988-89. Data taking began in October 1989. CLEO II operates near the threshold for open beauty production and it has logged to date the largest sample of B mesons in the world using data collected at the $\Upsilon(4S)$ resonance.

CLEO II consists of several independent detectors, whose combined role is to measure with high precision the energy, direction, charge, and type of particles

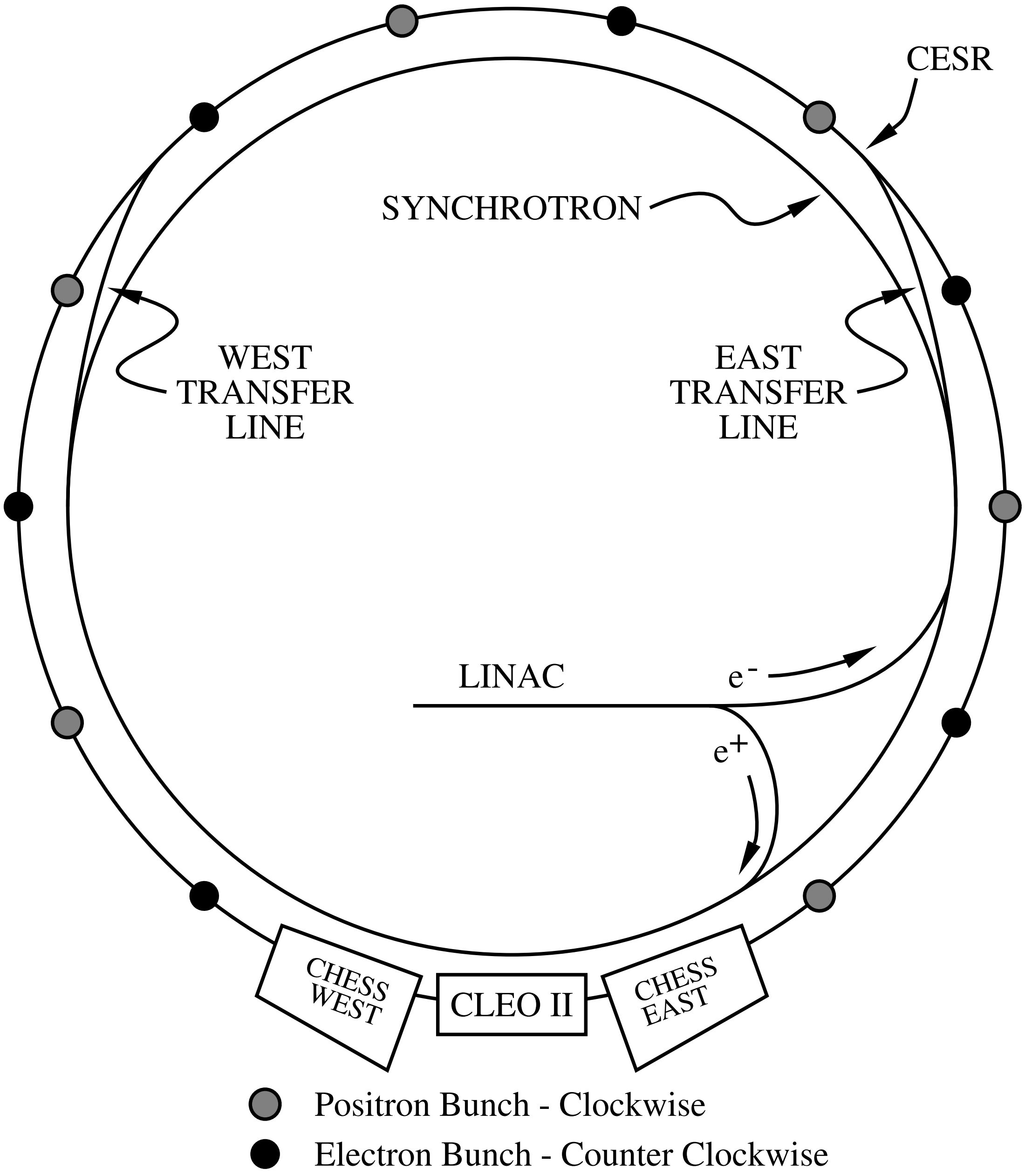


Figure 4.1: The CESR facility at the Wilson Laboratory on the campus of Cornell University, Ithaca, NY, USA.

produced inside the beam pipe in every e^+e^- annihilation. Apart from the elusive neutrino, which traverses matter with a very small interaction probability, no other particle should be able to escape the CLEO fiducial volume without leaving some sign of its passage. Particles such as K_L and neutrons do interact with the detector, but usually in a region not well instrumented making their reconstruction rather difficult. All other neutral and charged particles leave distinct signal in the CLEO II detector.

A superconducting magnet deflects charged particles so that their charge and momentum can be derived from the curvature of the measured trajectories. Inside the magnet coil are the CsI crystal calorimeter, where the energies of the electrons, positrons, and photons are measured; the time-of-flight counters, used for particle identification and trigger purposes; and various tracking chambers. The closest sub-detector to the interaction region is the precision tracking chamber, followed by the vertex chamber and the main drift chamber. Charged particles passing through the gas in these detectors knock electrons out of the gas atoms; the electrons generated in the ionization process are then attracted to positively charged wires and generate a pulse signal. Penetrating muons are expected to pass through the iron of the magnet and make its way to the layers of muons chambers where it can be detected in a second set of gaseous counters. The data from all these sub-detectors are collected and digitized with high speed electronics, and then combined into a data record for each event before being written on a permanent storage medium. Afterward, computers allow detector monitoring and physics analysis.

A full technical report on the CLEO II detector can be found in Reference [65]. The individual components of the CLEO II detector are described in the next sections, moving radially outward from the interaction point. This is followed by a brief overview of the trigger and the data acquisition systems, and the CLEO II Monte Carlo simulation. Appendix B describes the CLEO terminology.

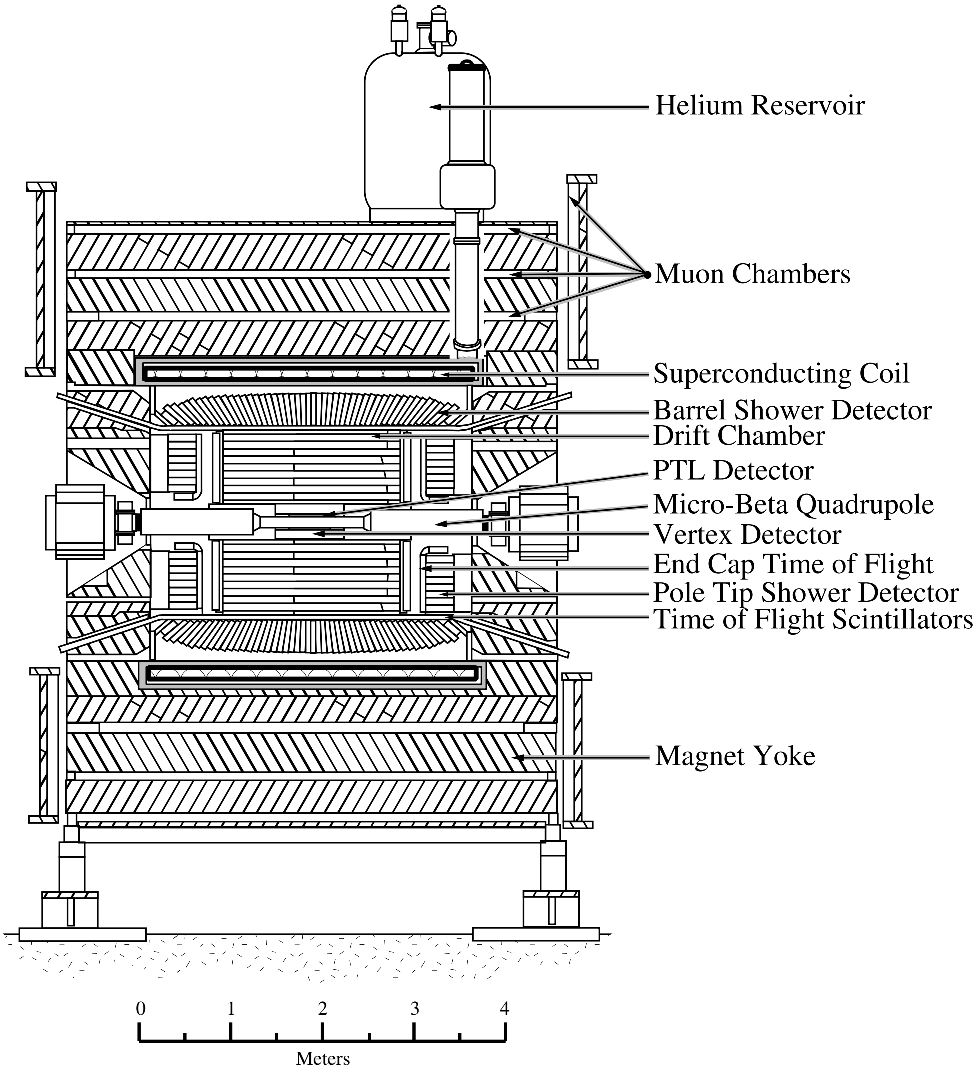


Figure 4.2: Side-view of the CLEO II detector.

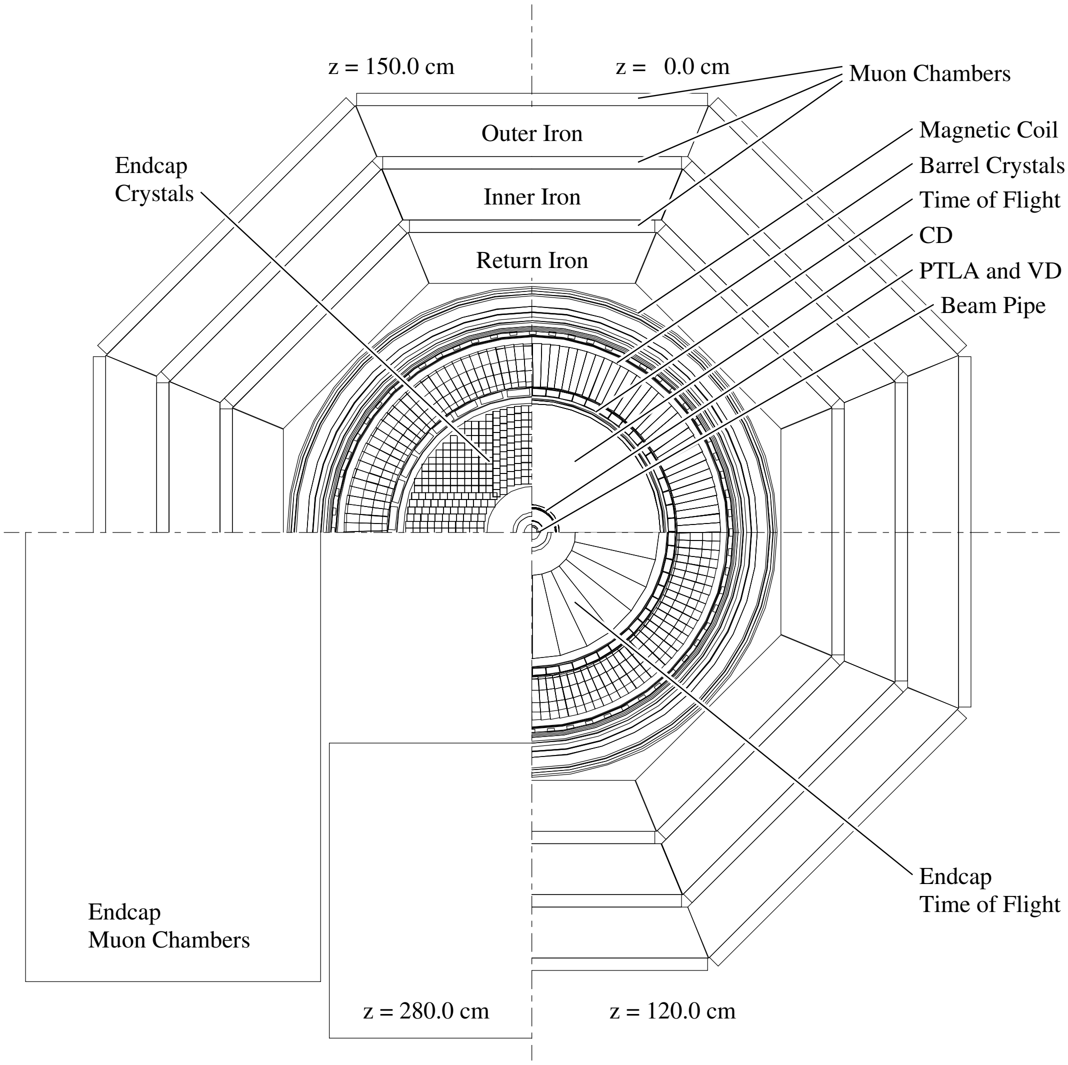


Figure 4.3: End-view of the CLEO II detector.

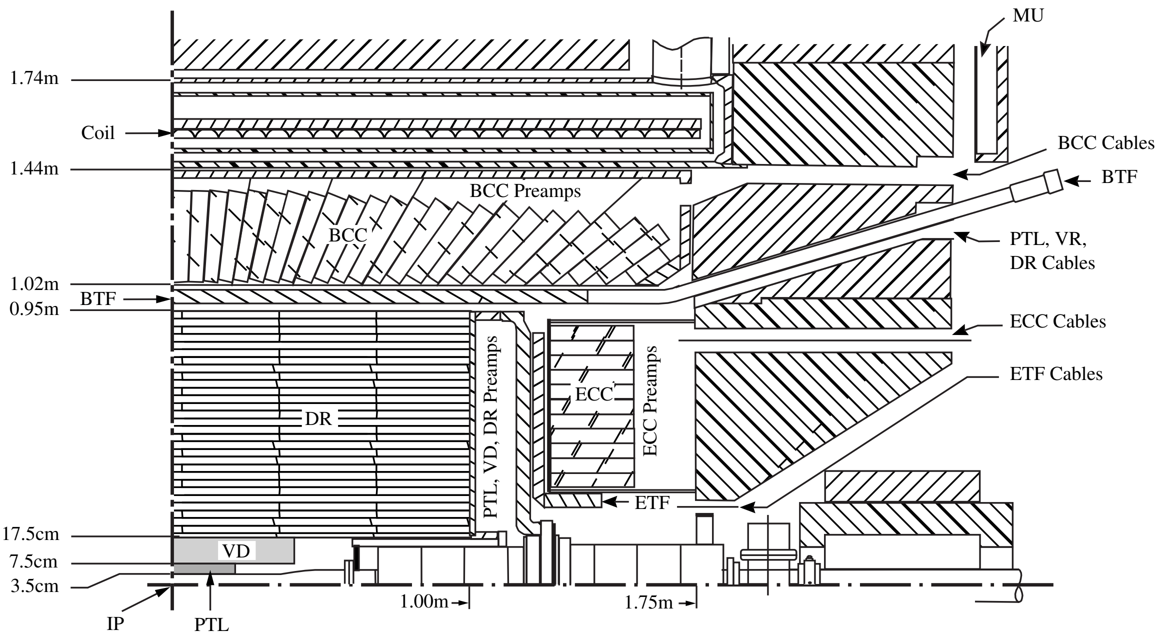


Figure 4.4: One quadrant of the CLEO II detector [66].

4.3 Central Tracking Chambers

The CLEO II central tracking system (CD) is a set of cylindrical wire chambers and of cathode strips. While the cathode pads were built to allow z coordinate measurements, the wire chambers were designed to facilitate 3-dimensional track recognition and to measure the momentum and specific ionization (dE/dx) of individual tracks. The wire chambers are the Precision Tracking Layers (PTL), the Vertex Detector (VD), and the outer Drift Chamber (DR). A more detailed description of the tracking system is given in Chapter 5.

The PTL is the innermost of the three tracking chambers; it occupies the region between 4.7 cm and 7.2 cm from the beamline. Its purpose is to precisely measure the origin of the tracks. The PTL is a six-layer straw tube drift chamber and has the smallest drift cell size of any of the chambers. The VD has 10 axial wire layers with radii from 8.4 cm to 16.0 cm and 70 cm in length. The structure of the inner chambers (PTL and VD) is shown in Figure 4.5. The inner and outer walls of the VD and DR are covered with cathodes segmented in ϕ and z. The DR

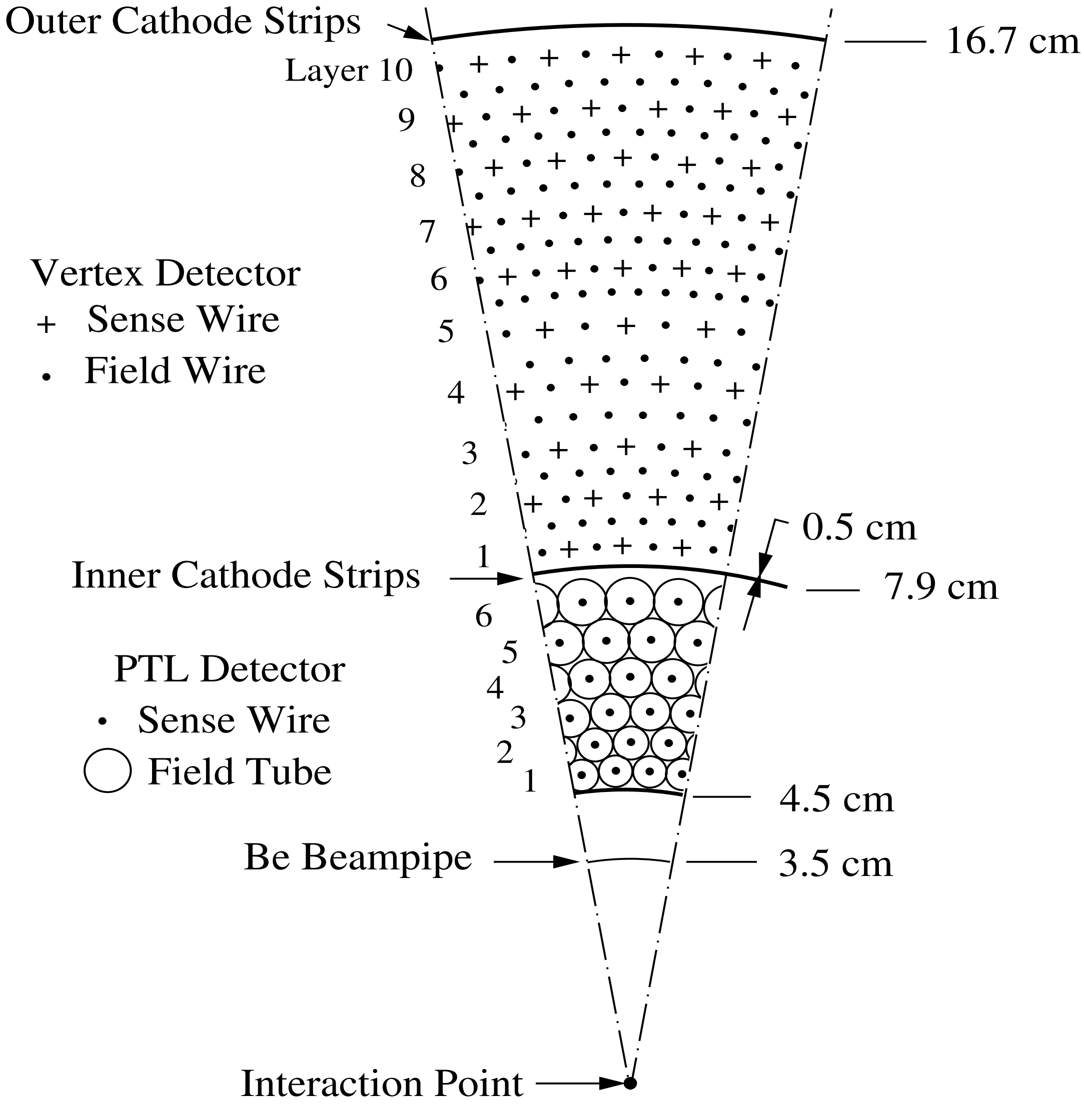


Figure 4.5: End-view of the inner tracking chambers (PTL and VD) showing the pattern of the drift cells.

(Figure 4.6) is the most important of the tracking chambers. It has 51 layers: 40 layers of axial wires and 11 layers of stereo wires. The DR provides most of the lever arm needed for the determination of charged particle momenta. It occupies radii from 17.5 cm to 95 cm and is two meters long.

The tracking chambers subtend 92% of 4π of solid angle with a transverse momentum resolution of

$$(\frac{\delta p_\perp}{p_\perp})^2 = (0.0015\, p_\perp)^2 + (0.0050)^2\,, \tag{4.1}$$

where $p_\perp$ is in GeV/c. The angular resolutions of the CD are:

$$\sigma_\phi = 1 \text{ mrad} \quad \text{and} \quad \sigma_\theta = 4 \text{ mrad}. \tag{4.2}$$

A description of the CD offline calibration is given in Chapter 5. The complete geometry and overview of the CLEO II drift chambers is described in References [67].

4.4 Time-of-Flight System

The main purpose of the time-of-flight (TOF) system of the CLEO II detector is to determine the velocities of charged particles by measuring their flight time from the interaction region. Using the momentum parameters measured in the drift chambers, one can then identify the particles by their mass

$$\frac{1}{\beta} = \frac{c \times \text{TOF}}{\text{Arc length}} = \sqrt{1 + \left(\frac{mc^2}{|\mathbf{p}|c}\right)^2}\,. \tag{4.3}$$

The TOF system also serves as the primary component of the fast trigger for data acquisition (see Section 4.8).

The TOF is made of plastic scintillation counters monitored by photomultiplier tubes. It is situated directly outside the DR and is divided into two sections: the barrel and the endcaps. The barrel section covers 86% of 4π and consists of 64 scintillation counters. The light from the counters is carried to phototubes by ultraviolet transparent lucite light guides. The long light guides allow the phototubes to be mounted outside the iron flux return of the magnet. Each barrel counter is viewed

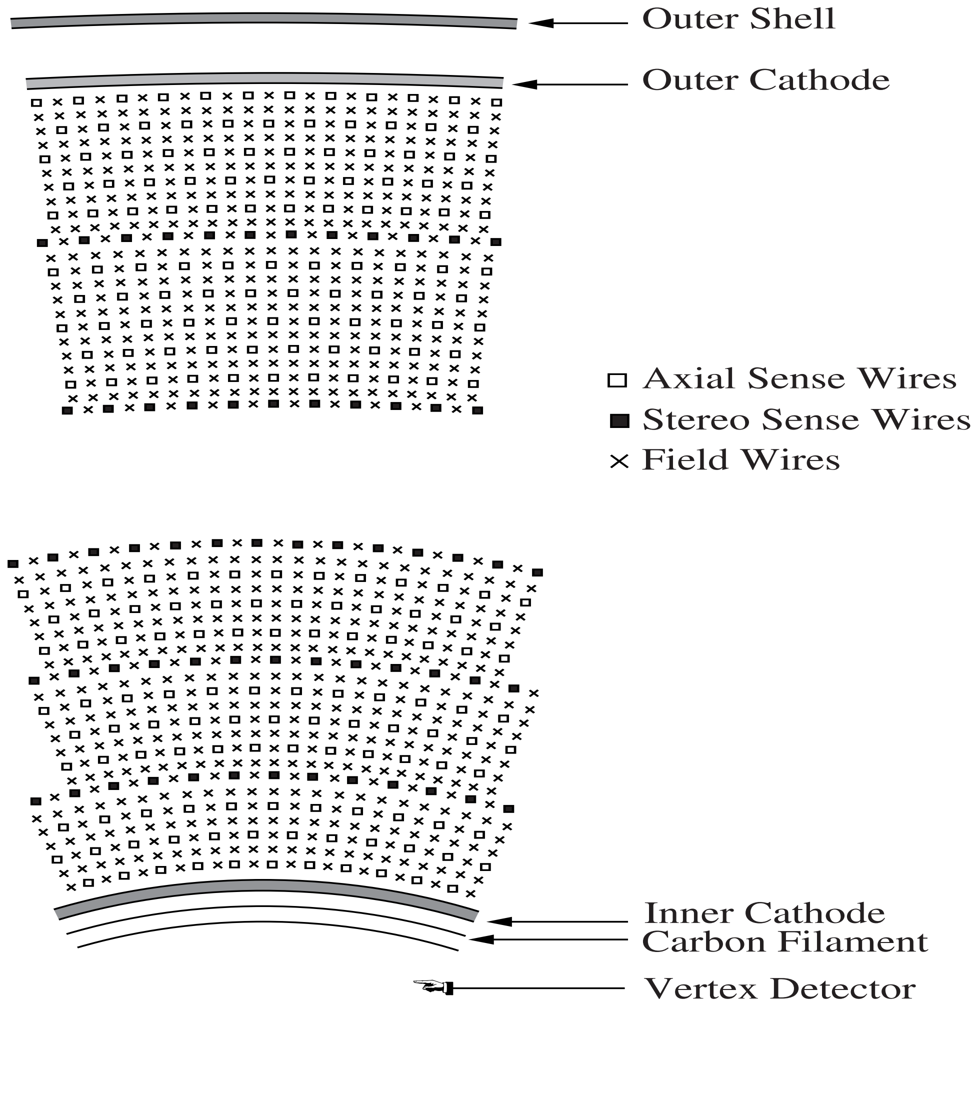


Figure 4.6: DR drift cell structure. The radial dimension of each drift cell size is about 14 mm. More details on the main tracking chamber layer topology can be found in Chapter 5 and References [65, 67].

by two phototubes. Each endcap section contains 28 scintillation counters arranged like pie slices around the beampipe. They extend the coverage to 96% of 4π. The phototubes for the endcap are glued directly to the square face of the scintillator prism and operate in the 1.5 Tesla solenoidal magnetic field. The resolution of the barrel (endcap) TOF is about 150 ps (300 ps) per tube for Bhabha events.

4.5 CsI Calorimeter

A calorimeter is a device where a particle's energy and position are measured with the total (or partial) absorption method. In such devices, the measured energy of the incident particle is contained in what is called a shower. The CLEO II electromagnetic calorimeter fulfills a fourfold purpose:

1. It measures the total energy and position of electrons, as well as the total energy, position, and direction of photons.
2. Analysis of the energy and shape of the shower permit separation of hadrons and muons from electrons and photons.
3. It is used to construct an energy trigger. In addition, the total measured energy is used for event classification.
4. It provides online and offline luminosity monitoring.

In the electromagnetic (EM) calorimeter, electrons and photons transfer all their energy to the material in a succession of bremsstrahlung and e^+e^- pair production, which in turn produce secondary electrons and photons in a chain reaction know as an EM shower. An electron is identified by matching the energy of the shower in the calorimeter to the associated track momentum measured in the central tracking detector. A photon is associated with individual showers when no charged track points back to the EM cluster. A full description of electron and photon identification is given in Section 6.6.1 and Section 6.8, respectively.

The CLEO II electromagnetic calorimeter (CC) is composed of 7800 thallium-doped cesium iodide (CsI) crystals of dimension $\sim$ 5 cm (2.7 r.l.) $\times$ 5 cm (2.7 r.l.)

$\times$ 30 cm (16 r.l.), with photodiode readout inside the 1.5 Tesla solenoid magnet. The dimension and segmentation of the crystals of the calorimeter were chosen to minimize leakage and maximize light transmission. The CsI calorimeter is located just outside the time-of-flight counters, but more importantly inside the magnet coil (see Figure 4.4). This configuration greatly reduces the amount of material which the particles must traverse before they reach the calorimeter. It also reduces the volume, and consequently the cost of the calorimeter. The calorimeter consists of a barrel and two endcaps, which together cover 95% of the solid angle. The barrel coverage starts at 32°, overlapping with the endcap (between 32° and 36°). The endcap coverage goes down to 18°. The barrel contains 6144 tapered blocks arranged in a nearly vertex-pointing geometry of 48 z-rows with 128 azimuthal segments in each. Each endcap holds 828 rectangular crystals stacked inside a cylindrical holder. Four silicon photodiodes mounted on a lucite window on the rear face of each crystal convert the scintillation light from the CsI into electrical signals. For redundancy, each of the photodiodes is closely connected to an independent preamplifier. The cable from each preamplifier travels outside the CLEO II detector for summing and digitization. The calibration of the electronics, relative crystal-to-crystal gains, and absolute energy scale have successfully maintained excellent resolution and stability over time.

The energy and angular resolutions of the CsI calorimeter are:

$$\frac{\sigma_E}{E}[\%] = \frac{0.35}{E^{0.75}} + 1.9 - 0.1E\,, \tag{4.4}$$

$$\sigma_\phi[\text{mrad}] = \frac{2.8}{\sqrt{E}} + 1.9 \quad \text{and} \quad \sigma_\theta[\text{mrad}] = 0.8\,\sigma_\phi \sin\theta \tag{4.5}$$

for the barrel section, and

$$\frac{\sigma_E}{E}[\%] = \frac{0.26}{E} + 2.5\,, \tag{4.6}$$

$$\sigma_\phi[\text{mrad}] = \frac{3.7}{\sqrt{E}} + 7.3 \quad \text{and} \quad \sigma_\theta[\text{mrad}] = \frac{1.4}{\sqrt{E}} + 5.6 \tag{4.7}$$

for the endcap section, with E being the photon energy in GeV.

4.6 Magnet Coil

The CLEO II magnet was designed to produce a uniform solenoidal magnetic field along the beam line so that charged particles in the tracking chambers would follow helical paths. Its superconducting solenoid coil provides an axial 1.5 Tesla magnetic field, uniform to within 0.2% over 95% of the tracking system volume. It is 3.1 m in diameter, 3.5 m long, and operates with a current of 3300 Amperes. The magnet is cooled by a liquid helium system. A 700 litre Dewar located above the detector (see Figure 4.2) delivers the cold liquid helium to a manifold at the bottom of the coil. This manifold supplies riser pipes fastened to the outside of the coil shell where the helium absorbs heat and decreases in density. The buoyancy of the lower density liquid helium and gas bubbles rising in the pipes draws the liquid around the cooling system.

4.7 Muon Chambers

Muons are very penetrating and long-lived particles. They are identified by placing a large piece of iron absorber in the path of the particles produced in the collision. The absorber stops essentially all particles, except the penetrating muons. Muon counters are placed outside the absorber. Muon identification relies on track matching between the hits in the muon chambers and the hits in the central tracking chambers.

At CLEO, the muon detectors are planar drift chambers outside the CsI calorimeter and the magnet coil. They consist of a barrel portion and two endcap portions. Each portion was designed to maximize the solid angle coverage and detection efficiency, and to minimize the number of hadrons which are misidentified as muons. The barrel muon counters are embedded in the iron flux return of the magnet. The barrel section consists of eight octants that are symmetrically positioned parallel to the beam line. Each octant is composed of three layers of 36 cm thick iron absorber which amounts to a minimum of 2.2 nuclear absorption lengths per layer. The muon counters are in the 9 cm gaps between the iron absorber at depths of 36, 72, and 108 cm, which corresponds to roughly 3, 5, and 7 absorption lengths ($\lambda = 16.7$ cm in

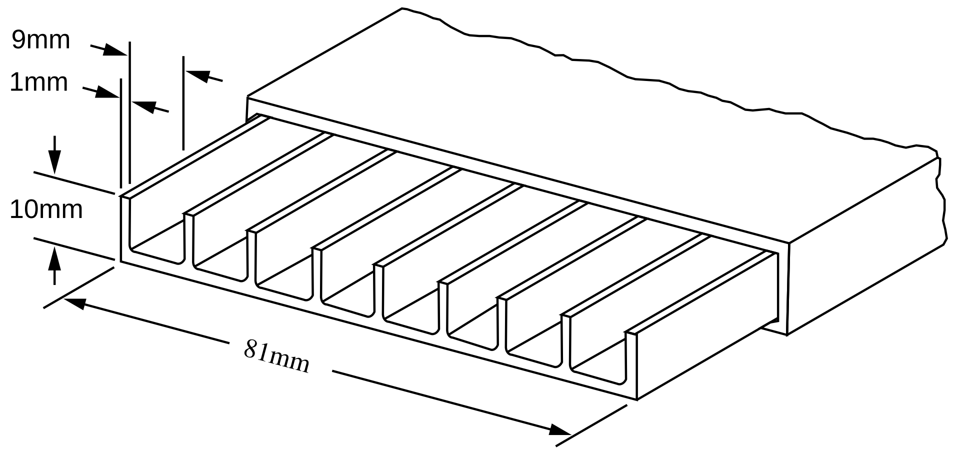


Figure 4.7: Cross-section of a plastic muon proportional counter.

iron). The endcap section is composed of additional counters which cover the forward and backward region. The positions of the muon chambers are shown in Figures 4.2 and 4.3.

Every muon chamber is composed of a set of plastic streamer counters (Iarocci counters) operating in proportional mode (see Figure 4.7). A set of counters in an iron gap is called a superlayer. Each superlayer is composed of three layers of counters. The cross-section of a superlayer is shown in Figure 4.8. The superlayers are about 5 m long and 8.3 cm wide. They are constructed from 8 rectangular plastic tubes with a bore of about 9 mm by 9 mm. Three inner sides of the tube are coated with graphite to form a cathode and electrically isolate the anode. The side of the tubes without graphite has 8 cm wide copper strips mounted perpendicular to the wires providing z measurements. The anode wire is positioned in the center of the tube and operates at about 2400 Volts with a 50:50 Argon-Ethane gas mixture. The counters spatial resolution is about 2.4 cm; and the copper strips give a spatial resolution in z of 2.8 to 5.5 cm.

The identification efficiencies and solid angle coverage for each of the three barrel superlayers are shown in Table 4.1. The solid angle coverage is limited by

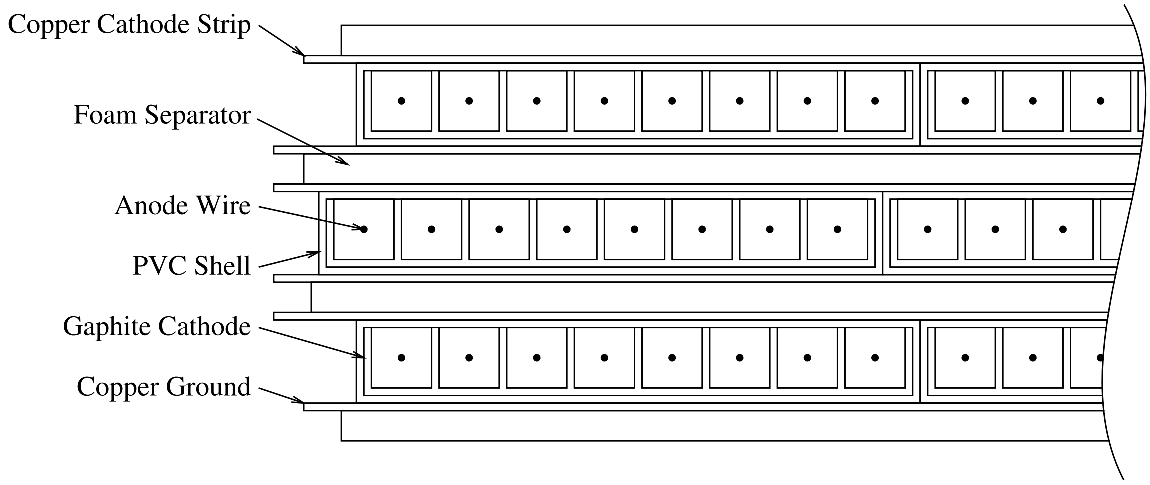


Figure 4.8: Partial cross-section of a muon detector superlayer, showing the slightly staggered three layers of proportional counters.

the light guides and phototubes of the barrel TOF to about $0.85 \times 4\pi$ of the total solid angle. The efficiencies were determined with muons from $e^+e^- \to \mu^+\mu^-$. A full description of the construction and performance of the CLEO II muon detectors can be found in Reference [68].

Depth	Solid Angle	Efficiency (%)
$> 3\lambda$	$0.85 \times 4\pi$	98.6 ± 1.6
$> 5\lambda$	$0.82 \times 4\pi$	97.5 ± 1.6
$> 7\lambda$	$0.79 \times 4\pi$	89.5 ± 1.5

Table 4.1: The solid angle coverage of the various barrel muon chambers and their efficiencies for detecting muons from $e^+e^- \to \mu^+\mu^-$.

4.8 Trigger and Data Acquisition

The massive amount of data produced by the CLEO II detector is digitized, collected, analyzed, and reduced to physics results with the help of computers. At every interesting e^+e^- interaction, the data acquisition (DAQ) system processes the electronic signals from the detector elements in a temporary storage medium (called a buffer), reduces the data rate to a manageable level, records the events of interest on a permanent storage medium, and controls and monitors the detector performance. While the information for the event is stored in buffers, trigger processors perform rapid but crude pattern recognition algorithms to select events suited for calibration and physics analysis.

Electrons and positrons in CESR cross each other at a rate of 3.6 MHz, which is far too rapid to be accommodated by the data storage and data analysis components of CLEO. Fortunately, most of the interactions are physically uninteresting and the actual rate of interesting annihilations is only a few Hz. CLEO II uses a hierarchical three-level trigger system; the three stages are called Level 0 (L0), Level 1 (L1), and Level 2 (L2). An additional software filter, called Level 3 (L3), is applied before data storage.

The L0 trigger is the first link of the DAQ system. It is designed to make fast and efficient decisions about whether or not charged and neutral particles have been produced in CLEO II. Because the L0 trigger system is confronted with the highest data rates, it uses information from a fraction of the detector channels. The L0 trigger receives input from the TOF scintillators, the VD tracking chamber, and the CC calorimeter. The TOF is the fastest device in CLEO II; the signals from the phototubes are ready in about 55 ns. The L0 criteria reduce the crossing frequency to a rate on the order of 10 kHz. Whenever any of the L0 requirements are met (see Reference [69] for more details), all gates to the detectors are disabled and the L1 trigger is initiated.

The L1 trigger takes more information from the detector and uses it to make better informed decisions about the event. It uses information from the TOF, VD, DR, and the CC. Typically, L0 and L1 require a few microseconds to eliminate uninteresting events. Overall the L0 and L1 requirements reduce the trigger rate to about 50 Hz.

Higher level triggers face much lower rates and perform more sophisticated event rejection algorithms. The L2 trigger uses more detailed tracking information and reduces the overall read-out rate by another factor of two. An accept flag at the Level 2 forces the detector signals to be sent to the L3 software filter. The L3 filter reduces the rate by 30% to 40%, depending on beam conditions. Events that pass the L3 requirement are then stored permanently on magnetic tapes for data reconstruction. The overall CLEO II trigger efficiency for $B\bar{B}$ events is 99.8%.

An accept signal from the L2 trigger allows CLEO to be read-out. CLEO is read-out in a common stop mode. The closing gate is set by CESR after each beam crossing. The actual data acquisition system can currently read events at 50 Hz with a 10% deadtime [70]. This means that the readout of the front-end electronics for each detector component is completed within 2 ms. To reduce the amount of data read out after a trigger, each of the electronic signals has to pass certain cuts (this process is called data sparsification) before they are sent to a buffer. The digitization of the entire CLEO II detector takes about 2.2 ms, and the digitization and sparsification take about 13.5 ms. The event size of a typical hadronic event is about 8 kbytes, which, given a 25 Hz triggering rate, requires a bandwidth of 200 kbytes/sec.

Online, a set of control and monitoring computers provides a user interface for the detector supervisors and ensures that the detector is performing correctly. Offline, diagnostic programs are used to monitor and calibrate the CLEO II sub-detectors. Bhabha and muon pair events are recorded online for calibration purposes. The L2 trigger has the capability to prescale these events by accepting only a predetermined fraction of two-track triggers. After calibration, the data stored on magnetic tapes are processed with the reconstruction program PASS2. The task of PASS2 is to transform the raw information (hits and clusters) into quantities required for physics analysis. The data processed and compressed with the program PASS2 are stored permanently on disk for later physics analysis.

4.9 Monte Carlo Simulation

A detailed set of Monte Carlo programs are used to simulate events in the CLEO II detector. This task is divided into two sections: the event generator and the detector simulation.

In this analysis, the event generator used is called QQ. We used two versions of the QQ program: the first one is the default program QQ and the second one is an extension to QQ called EvT [54]. The default QQ is used to describe continuum decays and non-semileptonic B decays. The EvT program was developed to describe B semileptonic decays. Unlike the default QQ, EvT is able to handle the full angular correlation among the decay products of the B meson.

In the simulation of a $B\bar{B}$ event, the event generator initially produces a virtual photon from an e^+e^- annihilation. The virtual photon decays to a pair of $B\bar{B}$ mesons. The $B\bar{B}$ pair is decayed according to the QQ decay table that includes the masses of all known particles, their measured branching fractions and lifetimes. Theoretical predictions for several expected, but not yet observed, modes and states are also used.

The continuum events are generated using the JETSET7.3 [71] software package from CERN. This set of routines produces non-resonant $q\bar{q}$ pairs based on the parton shower model according to the LUND formulation. Any unstable particles produced in the hadronization process are decayed using the default QQ decay table.

The Monte Carlo simulation of the detector is handled by a program called CLEOG. CLEOG is based on the GEANT software package from CERN [72]. CLEOG contains parameterizations of hadronic interactions between particles produced by QQ and the nuclear matter of the CLEO II detector. Its routines deal with electromagnetic shower development, decay in flight, multiple scattering, energy losses, Compton scattering, pair production, annihilation, ionization, delta-ray production, and bremsstrahlung. The output of CLEOG is then compressed with the PASS2 processor, which is the same program that also compresses the real data.

An adequate simulation of the CLEO II detector is essential for understanding the efficiencies and backgrounds for a given physical process under study. The optimization of our event-selection procedure, the estimation of various backgrounds, and the determination of the efficiencies for the signal events all rely on Monte Carlo simulation.

In order to ensure that the Monte Carlo simulation is reasonably correct, it is important to compare results of the simulation to measured data. Whenever CLEOG was used, the detector efficiency was tuned to match the values obtained from data. In this analysis great care is placed on the detector's ability for precise modeling of charged particle tracking, lepton identification, and hadron reconstruction. This topic is covered in more detail during the discussion of systematic uncertainties presented in Chapter 7.

Chapter 5

Calibration of the Tracking Chambers

A drift chamber is a single gas volume detector consisting of a large number of drift cells arranged in some pattern, each cell being a simple detector. The initial momentum of a charged particles governs its path in the detector and therefore the measurement of its position. Since the CLEO II drift chambers are placed in a 1.5 Tesla solenoidal magnetic field, charged particles move on helical trajectories. Ideally, such a helix is described by five parameters.

The purpose of a drift chamber is to sample the trajectory of charged particles at several points. From these samples, along with assumptions about the functional form of the trajectory, the momentum can be reconstructed. This is achieved by a program which interprets the measurements of each drift cell and finds the five parameters associated with a single particle trajectory. CLEO uses the programs DUET [73] and TRIO [74] for this purpose. The inputs required by a track reconstruction program are a list of the cells which were traversed by at least one charged particle, the coordinates of the sense wire in each of those cells, and a drift distance value for each cell that recorded a hit. The first two inputs are sufficient for coarse track reconstruction. The last input permits precise track reconstruction suitable for physics analysis.

The chambers consist of many drift cells, each of which consists of one sense wire connected to readout electronics. When a charged particle passes through a drift cell it ionizes some of the gas molecules in the cell volume. Then, an electric field directed away from the sense wire causes the released electrons to drift towards the sense wire, away from the ionized molecules. Very close to the sense wire, the electric field is so strong that the drifting electrons trigger an avalanche, which produces a macroscopic electric pulse on the sense wire. This pulse then travels to the readout electronics, where, if it satisfies a discriminator threshold, it closes a circuit consisting of a capacitor at a reference voltage grounded through a resistor [65]. This capacitor then discharges through the resistor until a gate from the trigger system breaks the circuit or until the circuit is reset. If the circuit is closed by a trigger [69], the voltage left on the capacitor is, for the purposes of track reconstruction, the only reading of that drift cell. When a Level 2 trigger causes the entire detector to be read out, the output of the drift chambers consists of a list of drift cells which recorded a hit (*i.e.*, a list of drift cells in which there was a large enough pulse created on the sense wire to satisfy the discriminator) as well as voltage reading for each of these cells. The voltage reading is an integer number between 0 and 4096, corresponding to an analog to digital conversion of the voltage left on the capacitor, and is known as the TDC. The goals of the calibration of the drift chambers for position measurements are:

1. To determine, from each drift cell output (TDC), a drift distance value with the smallest possible error.

2. To measure the relative positions of all sense wires as precisely as possible.

The final performance of the track reconstruction depends on how well these goals can be met. It is therefore desirable to have a procedure that will quickly and reliably lead to the achievement of these goals for each new data set, with minimal human effort. The tracking chambers calibration procedure is described in this chapter [75]. More details can be found in [76, 77, 78].

The main drift chambers of the CLEO II detector not only measure the trajectories of charged particles, but also their specific ionizations dE/dx. The whole volume of the main tracking chamber is filled with drift cells in order to maximize both the ionization collection for dE/dx measurement and the acceptance for low momentum and small angle tracks. The charged particle identification procedure is described later in Section 6.7.

5.1 Definitions

It is useful to clearly define the concepts involved in the calibration procedure. This section is devoted to the definition of the calibration and tracking terminology. It makes the following sections shorter and easier to understand.

Track Our best estimate of the trajectory of a stable charged particle (π, K, p, e, or μ) is called a track. In a solenoidal magnetic field, trajectories are assumed to be helical. They are fully specified by five independent track parameters {DA,CU,FI,CT,Z0}.

- **DA** Distance of closest approach of the track to the coordinate origin in the $r - \phi$ plane.
- **CU** Signed curvature of the track. CU $= q/2\rho$, where ρ is the radius of curvature and q is the electric charge ($q = \pm 1$).
- **FI** Azimuthal angle of a tangent to the track at the point of closest approach.
- **CT** Cotangent of the polar angle (θ) or the tangent of the dip angle.
- **Z0** The value of the z coordinate of the track at the point of closest approach.

DCA This is the Distance of Closest Approach of a track to a given sense wire, and it is a function of the track parameters. A DCA value is calculated for each drift cell that records a hit. It is often called the *calculated* or the *projected* distance of a track to the sense wire (see Figure 5.1).

Entrance Angle (α) This is the angle between the tangent to a track at the point of closest approach to the sense wire and a radial line that connects the point of closest approach to the coordinate origin. Again, α is a function of the track parameters, and a value of α is calculated for every cell that a track passes through. In the approximation that any track is locally a straight line in every cell, α and DCA uniquely specify a particle trajectory through a drift cell.

TDC As explained in the introduction, this is the output of a drift cell. It is a number between 0 and 4096 corresponding to the digitized voltage left on the timing capacitor when the trigger occured.

Drift Time (TM) This is the time elapsed between the ionization of the gas in a drift cell and the arrival of the drift electrons to the sense wire. TM stands for *measured* time.

Raw Time The TDC is a measure of the time elapsed between the firing of a drift cell electronics and the stop from the trigger. It is converted to time units relative to the trigger stop by the *electronic* or *online* calibration [79, 80]. The result is called a raw time.

Time Zero The time zero is the time between the trigger stop and the beam crossing. Thus, to derive a drift time from a raw time, one needs to know the time lapse between the ionization of the gas in the drift cell and the trigger stop. This lapse is the time zero minus the transit time of the charged particle from the interaction point to the given drift cell. The time zero does not depend on track parameters and, since the trigger stop occurs at a fixed time after a beam crossing [65], the time zero is roughly a constant.

In summary, the drift time is given by the time zero, minus the raw time, minus the transit time (Arc length/$\beta\, c$), minus the propagation time (see Figure 5.2). The propagation time is the time between the arrival of the drift electrons to the sense wire and the firing of the drift cell electronics. The transit and propagation time corrections do depend on the track parameters.

The time zeros are calculated via an iterative process that minimizes the time residuals. The time zero is one of two essential offline calibration constants.

Drift Distance (DM) When the firing of a drift cell is caused by ionization from a charged particle (rather than noise), the drift time contains information about the DCA of that charged particle. The measurement of the drift distance is our best estimate of the DCA given the drift time. DM stands for *measured* drift distance (see Figure 5.1).

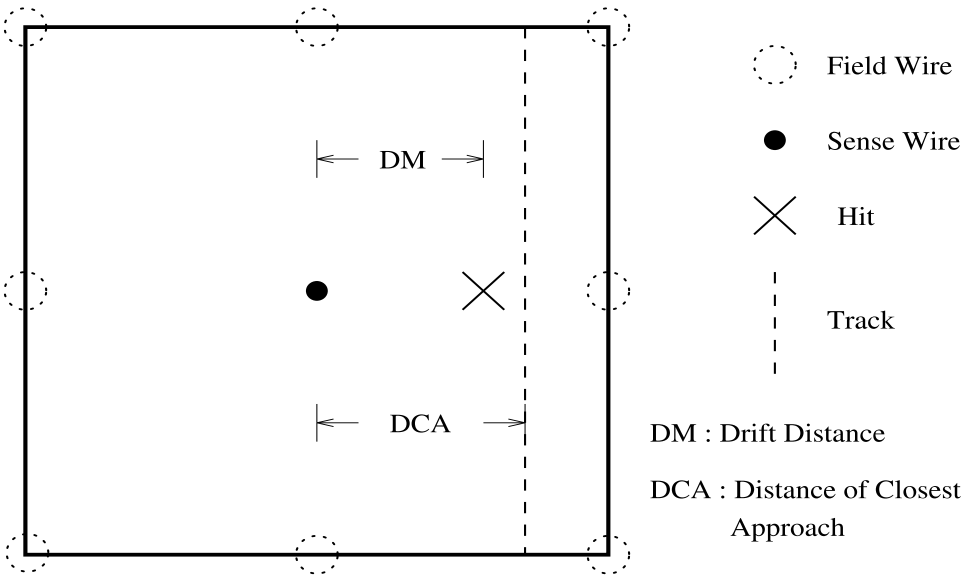
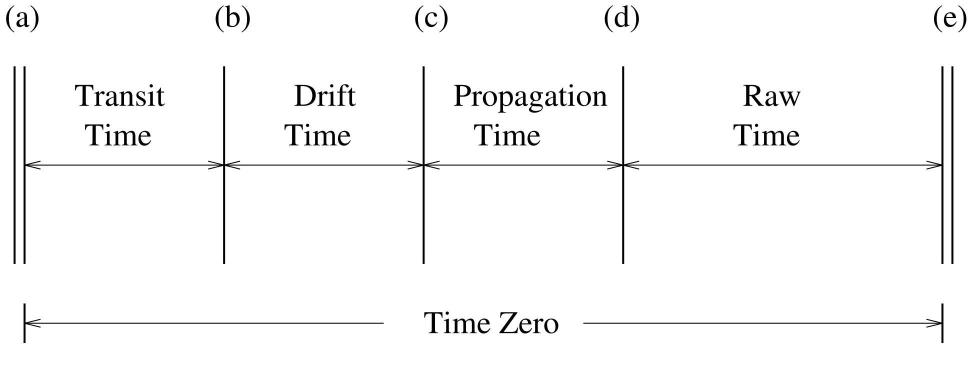


Figure 5.1: Representation of a typical DR II drift cell.



Drift Time = Time Zero - Raw Time - Transit Time - Propagation Time

(a) Beam Crossing
(b) Ionization of the gas in the drift cell
(c) The drifting electrons arrive at the sense wire
(d) The signal pulse crosses the discriminator threshold
(e) Trigger stop

Figure 5.2: Timing structure of the CLEO II drift chambers.

D-T Function The Drift-Time function is a one-to-one mapping from drift time to approximate drift distance. Ideally, the drift distance is exactly a function of drift time; however, as will be discussed in the next section, the drift distance in general also depends on other variables, such as entrance angle. Fortunately, the dependence on other variables is typically small and can be treated as a correction to the D-T function. In the CLEO II offline constants, there is one D-T function for every layer of drift cells. Each of these is parameterized by a variable number of connected straight line segments of equal width in drift time, where zero drift time always maps onto zero drift distance. The D-T function is the second essential offline calibration constant.

Ambiguity (A) The DCA is not enough to tell us whether the track passed to the left or to the right of the sense wire. Any track crossing a drift cell always has two possible paths characterized by the same DCA. Which side of the wire the track is on is referred to as the ambiguity, A, of the hit assignment. In CLEO coordinates, if ϕ at the point of closest approach is less than ϕ of the sense wire, then $A = -1$; otherwise $A = +1$. The ambiguity of each hit is resolved by the track fitting algorithm.

Calculated Time (TC) It is simply DCA converted into a time. In other words, TC is Time(DCA). When TC is computed, all the corrections to the D-T function are considered in the distance to time inversion process.

Spatial Residual A spatial residual (RES) is a measure of how close the drift distance of a hit is to the DCA of a fitted track in a cell, in units of distance. Recall that the drift distance is an estimate of the DCA based on the drift cell measurement. Whereas the DCA is the projected distance: the actual distance from the sense wire to the reconstructed particle trajectory. Residuals are defined with a particular sign convention: RES = A(DM - DCA). With this sign convention, the average residual of many tracks in any given cell is sensitive to offsets in the sense wire position. For example, if the track reconstruction algorithm thinks that ϕ of the sense wire is less than it actually is, the DCA will be overestimated for $A = 1$ and underestimated for $A = -1$, while the drift distance will not be similarly biased. This will lead to negative shifted residuals for both ambiguities.

Time Residual A time residual (TRES) is a measure of how close the drift distance is to the DCA, in units of time. It is given by TRES = TM - TC. Note that since the D-T function maps drift time to the approximate drift distance, Time(DCA) is the inverse of not just the D-T function, but also of all subsequent corrections. Thus, the calculation of Time(DCA) is not trivial. An alternative way to calculate time residuals is: TRES = (DM - DCA) $\div$ dD/dT, where dD/dT is the derivative of the D-T function at the given drift time. This method is exact in the limit of small TRES, even if corrections to the D-T function are ignored in calculating dD/dT, but it is inaccurate for large TRES because the derivative of the D-T function is not constant.

5.2 Constraints From Hardware Design

The physical characteristics of the drift chambers largely determine the form of the calibration constants. It is crucial to learn in advance what limits on the performance of track reconstruction are imposed by the hardware. Doing so focuses the calibration effort on areas where gains can be made and prevents wasted effort in areas where improvement is not possible.

A drift cell detects charged particles through their interactions with gas molecules. Unfortunately, each interaction also affects the motion of the charged particle that the drift chamber is trying to measure. In general, any two particles on exactly the same trajectory (but at different times) will produce the same TDC value in a given drift cell only if they ionize gas continuously along their path. It is therefore desirable that the charged particle mean free path, $\bar{x}$, be small compared to a typical DCA. For the argon-ethane gas mixture used in CLEO II, $\bar{x} \approx 200$ μm, whereas 0 mm $<$ DCA $<$ 7 mm. Note that very close to the sense wire (DCA $\approx \bar{x}$) the drift distance resolution is inevitably limited by $\bar{x}$. However, the main limitation related to $\bar{x}$ is that information about the initial conditions of a particle's motion is diluted in every collision with the wires and the gas (multiple scattering). Overall, the multiple scattering induced on the particles by the wires and the gas have comparable contributions. The effect of multiple scattering depends on the thickness of obstructing material in the chambers (in radiation lengths), the strength of the magnetic field, and the length over which position measurements are made. For the CLEO II hardware configuration, the expected transverse momentum resolution is [65, 67, 81]

$$\left(\frac{\delta p_\perp}{p_\perp}\right)^2 = (0.0015\, p_\perp)^2 + (0.0050)^2, \tag{5.1}$$

where $p_\perp$ is in GeV/c. The first term depends on the drift distance resolution and hence on calibration, whereas the second term accounts for multiple scattering. This equation makes very clear that for low momentum tracks ($p_\perp < 1$ GeV/c) it is not useful to push the calibration near the above theoretical limit of 0.0015. It further suggests that it is appropriate to calibrate the tracking system with Bhabha events, even though most particles in physics analyses are of much lower momentum, because it is at high momentum that calibration dominates the momentum resolution.

The CLEO II tracking system is embedded in a 1.5 T magnetic field (along z, the beam axis) which is another important consideration for calibration. Due to the magnetic field, the drifting electrons do not, in general, follow a straight path to the sense wire. For a fixed DCA, the exact trajectory, and hence the drift time, depend on the details of both the electric field and the magnetic field (since the electromagnetic force is given by $\mathbf{F} = q\mathbf{E} + q\mathbf{v} \times \mathbf{B}$) and on the type of gas in the chamber [82]. Therefore, the D-T function potentially varies with entrance angle and ambiguity.

However, if the electric field is invariant under rotations about the sense wire, then the D-T function is constrained by symmetry to be the same for all entrance angles and ambiguities. This is clearly the case in the PTL drift cells because of their circular geometry, but it is also true near any sense wire. The surface of any wire is an equipotential and therefore, the field lines of the electric field are radial close to the wire. Whatever the particular case may be, for small DCA, the D-T function will neither depend on entrance angle nor the ambiguity, regardless of the electrostatic boundary conditions far away from the sense wire. This is an important physical constraint in converting drift times to drift distances.

In general, the exact trajectory of the electrons (or drift lines) deviates from the direction of the electric field **E**, by an angle named the Lorentz angle η. In the particular case of perpendicular **E** and **B** fields, the Lorentz angle is predicted to follow:

$$\tan(\eta) = w\tau \simeq \kappa(E)\upsilon B/E, \tag{5.2}$$

where E is the magnitude of the electric field, B is the magnitude of the magnetic field, ω and τ are respectively the cyclotron frequency (eB/m) and the effective time between collisions, υ is the drift velocity ($\upsilon \sim 50\mu$m/ns), and the factor $\kappa(E)$ depends on the electric field and the gas in the chamber.

Figure 5.3 shows a computer simulation of the isochrones in a typical DR II cell. An isochrone is a line connecting all points of closest approach that lead to the same drift time. Near the sense wire the isochrones are circular, indicating no dependence of the D-T function on entrance angle or ambiguity. Therefore, the D-T functions in all DR II layers are constrained to be identical close to the sense wire, and there is only one D-T function per layer, with separate entrance angle dependent corrections for each layer added for large DCA.

5.3 Method

From the calibration point of view, the CLEO II tracking system allows one to handle all cells of a particular drift chamber (PTL, VD, or DR) with the same basic calibration constants. Remaining fine tuning of the D-T functions and time zeros is treated by small corrections. Because of deviation caused by fluctuation in operating conditions, a set of calibration "constants" are normally valid for a given segment of data.

There are several ways to find the D-T functions, time zeros, and all other calibration constants of a drift chamber. The goal is to find the constants that lead to the smallest χ^2 between the drift distances of all data hits and the DCA values predicted by track reconstruction. We do not intend to perform a simultaneous minimum χ^2 fit to all CLEO II data with every calibration constant as a free parameter.

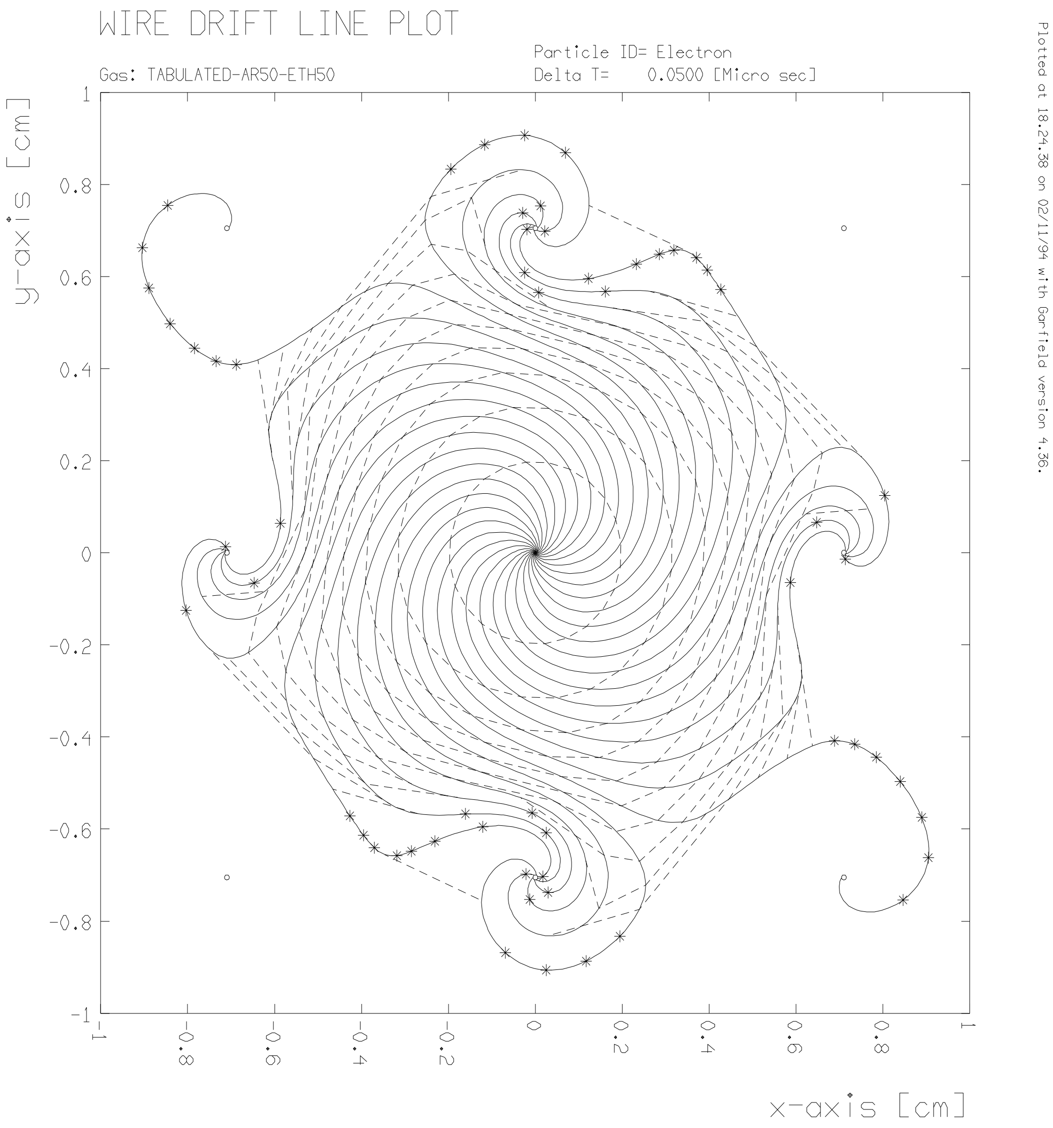


Figure 5.3: Simulation of the DR II cell of an axial layer bounded by axial layers. Drift lines (solid line in 15° steps) and isochrones (dashed line in 50 ns steps) for Argon-Ethane with B=1.5 T and HV=2000 Volts. The typical radius of a DR II cell is about 7 mm.

This ideal situation is not practical, because the number of free parameters would exceed $10,000$. Instead, for CLEO II, we determine these constants empirically using Bhabha events. We find the solution by successive iterations on the same data sample, where we determine some fraction of the free parameters in each iteration by starting with the D-T functions. The method is based on an accurate determination of the D-T functions, followed by calculation of all other constants from spatial and time residuals. The advantage of this method is that it should naturally converge on the desired goal, because the combined χ^2 of all the data is nothing but the sum of the squares of all residuals (appropriately normalized). Because of this, and because we use only Bhabha data rather than all data, we need to understand and impose physical constraints that the constants must respect, in order to avoid falling into local minima of this iterative χ^2 minimization. The first step is to reconstruct non-radiative Bhabha events with approximate D-T functions and time zeros. Improved D-T functions are then obtained from track parameters generated by an iterative least square fit which minimizes:

$$\chi^2 = \sum_{i=1}^{N} (\mathrm{RES}_i^2/\sigma_i^2), \tag{5.3}$$

where N is the number of hits on the track and σ_i is the spatial resolution.

After the D-T functions are found, other constants, such as time zeros, are extracted. The iterative χ^2 minimization is repeated with the new D-T functions, the new time zeros and other updated constants until it converges. In the following sections, a detailed description of how the D-T functions are found and parameterized is given.

5.3.1 The Track Fits

For reasons discussed earlier, we choose Bhabha events for calibration. Bhabhas have the disadvantage, however, that they are not produced isotropically, but exhibit a strong $\cos^2\theta$ dependence which peaks along the beam axis. Therefore, in order to assure that our constants are not biased by over-representation of small polar angle tracks, we filter Bhabhas so as to flatten their distribution in $\cos\theta$. This is accomplished through a routine named `levcut` which monitors the $\cos\theta$ distribution of events selected for constants-finding and rejects events through a simple negative-feedback check. As soon as the large $\cos\theta$ bins pass some threshold, small polar angle events are rejected until the large polar angle bins catch up. See Figure 5.4.

After an appropriate event selection procedure is established, one must decide which hits belong to each track. This is done by applying a crude track-fitting algorithm (in our case, TRIO) which provides us with track parameters for the two tracks in the event. These track parameters are used to identify which hits on the hit map of the event fall within a road about the track. Hits meeting the road requirement are stored in a hit list. The road is sufficiently wide so that misalignments are insignificant. Neglecting to do so can result in cases in which perfectly good hits are discarded, resulting in low statistics and biased constants.

Once we have a list of hits we begin fitting the Bhabha events. Various fits using subsets of the hit list will be carried out for different purposes, but for now we will concentrate on the most general fit to the event data: a fit which includes all hits from all three chambers and cathodes. The five track parameters defined earlier are determined by this fit. They will henceforth be referred to as the *base* track parameters. They will serve as the default track parameters unless a parameter is specifically requested to be re-determined in a more specialized fit.

Although there are actually two physical tracks in the event, the electron and the positron, we treat the two as if they were one particle entering the detector from one side, scattering elastically at the DR-VD interface, scattering elastically at the the origin, scattering again at the DR-VD interface, and exiting on the other side. This entity, the two tracks treated as one, will be called a *dualtrack*. The scattering at the origin is necessary to account for initial and/or final state radiation. We do not use radiative Bhabha events, but all Bhabha events radiate to some extent. We describe the dualtrack with one set of five track parameters, two angles (one in r-ϕ and one in r-z) describing a kink at the origin, and an angle for each DR-VD interface crossing, to account for the scattering.

If an event is determined, based on all the criteria mentioned so far, to be suitable for calibration, it may then be fitted for the base track parameters. These

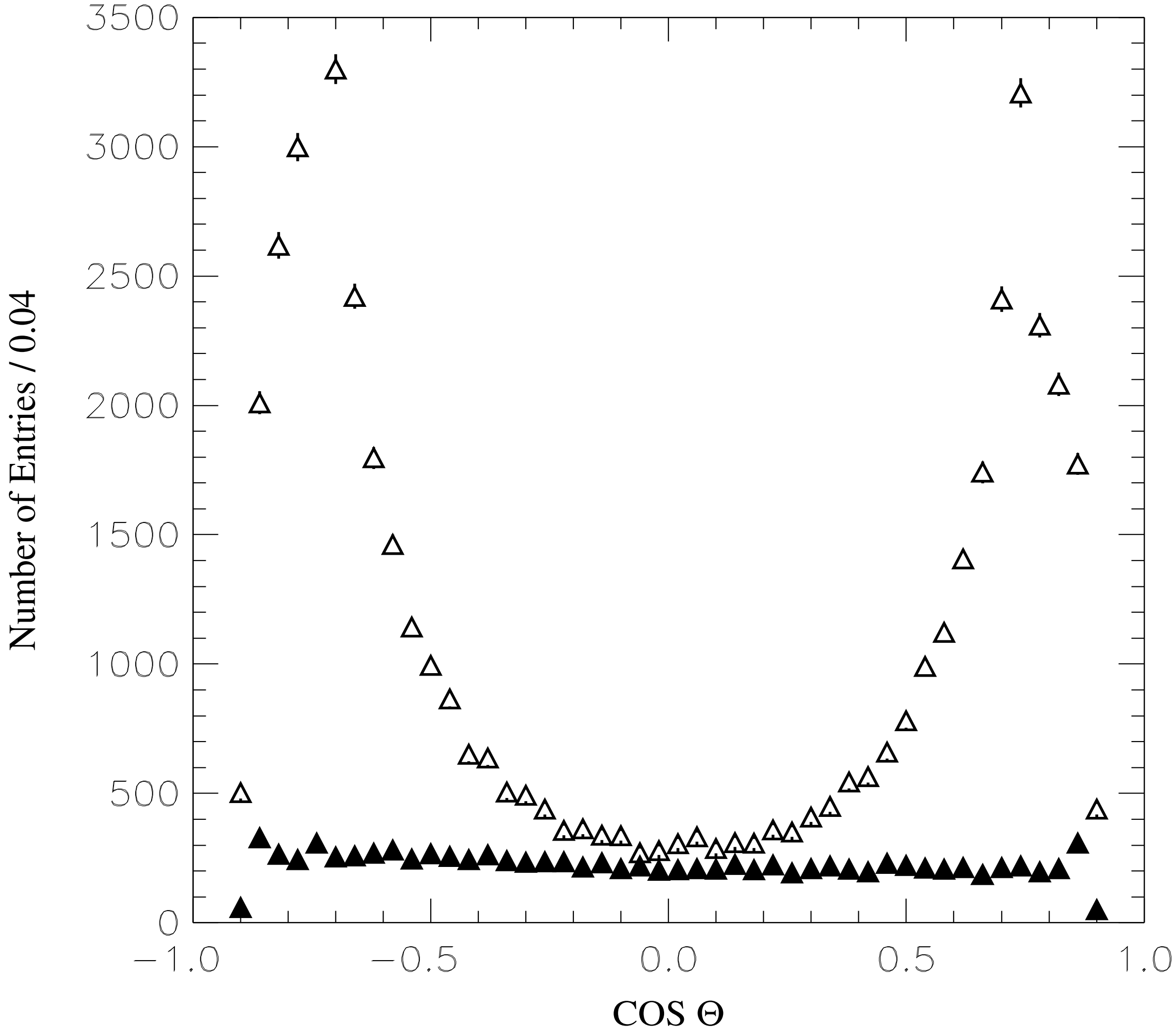


Figure 5.4: The distribution in $\cos\theta$ for Bhabha events. Open triangles: without `levcut`. Filled triangles: with `levcut`.

track parameters will be very important for calibration, so we want to take steps to ensure that obviously spurious hits will be ignored. To do so we fit the dualtrack once, look at the residual of each hit, and reject hits with residuals greater than 4 standard deviations (σ) [83].

Rejected hits are flagged and will never be used in any future fit. After all the hits have been checked and the bad ones flagged, we fit the dualtrack again and throw out hits with residuals greater than 2σ. Now we carry out the fit for the base track parameters and save them. We also save for every hit (including those that were rejected for use in the fit) an address, layer number, wire number, TDC, measured drift distance, pulse height, arc length to the hit, residual, and DCA. Now that we have the base track parameters, we can carry out the specialized fits designed specifically for obtaining various calibration constants.

Currently there are three specialized fits, also called *internal* fits, one for each chamber. The specialized fit information obtained will be used to calibrate each device: the DR fit for the DR axial and stereo layers, the VD fit for the VD axial layers, and the PTL fit for the PTL axial layers. In an internal fit, hits from one device *only* are used as input so that the trajectory of a particle within that device can be determined independent of relative alignment between devices. (At this stage, misalignments may not have been corrected and hence relative rotations, offsets, and tilts may exist between one device and another; if we fit hits from each device separately, this systematic bias is eliminated.) We also demand that the measured drift distances of the hits used in an internal fit fall within a window corresponding to a region of the cell with good resolution. A minimum number of hits per dualtrack passing this cut is also required. Keep in mind that for the purposes of calibration, unlike physics analyses, we want to favor events in which we can determine with confidence the track parameters. We are concerned not with efficiency, but with accurate track parameters. The above requirements are aimed at selecting the best hits and hence, favoring well-reconstructed tracks which will make it possible to converge on the true constants. A description of each internal fit follows.

DR Fit: The DR fit refits for CU, FI, DA, CT, and Z0 using only DR axial hits and all cathodes. Notice that all the track parameters except kinks are redetermined for this fit. By doing so we are assuming that all five of them can be reliably determined without the help of PTL or VD information. This is clearly true for CT and Z0, since the PTL and VD anodes contribute no z information. It is also true of CU, since the DR, with its large span in radius, affords us the greatest lever-arm with which to determine the curvature. Because the DR provides information only far away from the origin, it makes for poor determination of FI and DA in single track fits. The dualtrack fit imposes physical constraints on these parameters, so we believe that they are well determined with the DR alone. To favor events with a maximum number of reliable hits, we require that hits used in the fit fall between 10% and 70% of the cell radius. Such requirement are mainly based on the large uncertainty on the Lorentz angle in the outer cell regions where the electric field is no longer cylindrical.

VD Fit: The VD fit refits for FI and DA using only hits from the VD anodes. It is appropriate to refit for FI and DA, and not CU, Z0 or CT, when using the VD alone, since the VD provides good information about what happened close to the interaction point, but offers little lever-arm for measuring CU and obviously no new z information. This fit excludes hits in the outer 10% of the cell. This requirement, in conjunction with the fact that the fit is unbiased by geometry misalignments, is expected to generate better values for FI and DA than those of the base track parameters.

PTL Fit: The PTL fit refits for FI and DA using only hits from the PTL. The reasoning here is identical to that applied to the VD. Here hits within the entire cell are allowed into the fit because with a maximum of 12 possible hits on the dualtrack, statistics become a limitation. Nevertheless, this window could potentially be optimized.

For each internal fit, as for the general fit mentioned first, a set of information is stored: the new track parameters and an arc length, residual, and DCA for every hit. From this information a variety of quantities (*e.g.*, TRES, entrance angle, z position of hits) will be calculated, plotted and used to determine the constants. The internal fits are labeled by the parameters IFIT. The DR fit has IFIT = 1, the VD fit has IFIT = 2, and the PTL fit has IFIT = 3.

5.3.2 The Drift Functions

Overview

In principle, the drift function depends only on the cell geometry, the properties of the gas, the magnetic field strength, and the voltage of the sense wire. Since all cells in a layer are held at the same voltage and share the rest of the determinants as well, we make one drift function per layer with the exception of a few badly behaved layers.

The gravity sag ($\sim 60\ \mu$m) and all geometrical distortions of the field wires are corrected for in the calculation of the drift distance within DUET. The sense wires are theoretically in an unstable equilibrium, being attracted to all surrounding field wires. The tension on each wire is thought to be enough to overcome electrostatic distortions [84] and it is not corrected for in the track-fitting procedure.

The extraction of the drift functions is an iterative process. Before one begins, one must make an estimate of the drift function; this estimate could be the most naive drift function imaginable, given by:

$$d(t) = t \times R_{\text{cell}}/(t_{\max} - t_{\min}) \tag{5.4}$$

where R_{cell} is the cell radius and $t_{\max}$ and $t_{\min}$ are the maximum and minimum possible drift times. A better estimate, if it is available, is a drift function from a previous data set. This is generally what is used. Use of an approximate drift function means that there will be a large uncertainty in the measured drift distances. But because a large number of hits is used in the track fit, the error in the track parameters, and hence DCA, will be small [81]. Thus, a plot of DCA versus drift time yields a better drift function than the original form used to determine drift distances. Samples of DCA versus drift time, the scatter plot and the result of a fit to the corresponding profile histogram, can be found in Figures 5.5 and 5.6. The profile histogram [85] of DCA averages versus time has been fit using a piecewise linear scheme (described in Appendix D). The result of the fit is saved and will go on to become the starting point for the next iteration. Eventually the error in the drift distances is dominated by the intrinsic error in the measured raw times and further iterations cannot improve the drift function.

There are complications to this simple scheme. Far from the wire DCA is nearly a function of time; there the distribution of DCA for a given drift time bin is sharply peaked. Close to the wire, however, these distributions broaden due to ionization statistics and misassigned ambiguities creating a background underneath them. The closer the track passes to the wire, the more difficult it is for the fitter to resolve the ambiguity, creating a broad background peaked at zero.

The greatest complication to determining the drift function arises in the case where the isochrones are not rotationally invariant. One can easily see in Figure 5.3 that this effect is most dramatic at the edge of the cell. The result of this effect is that for every possible drift time there are many possible drift distances. This leads to a drift function which cannot be described by one set of offsets and slopes per layer, but must be parameterized in such a way that the effect of the distorted isochrones is modeled. Several techniques have been developed to address these issues. The specific parameterization of the drift function will be discussed for each device.

VD and DR Axial Layer D-T Functions

The D-T functions for these two devices are made in the same way and so will be discussed together. As stated earlier, the drift function is found by plotting the DCA versus drift time and fitting the function to a piecewise linear function with typically ten to twenty segments, each of which is described by a slope and a y-intercept. Thus, the D-T function is parameterized by:

$$D(t) = A_i + B_i t \quad \text{where } (i-1)W_t < t < iW_t \ \cdot \tag{5.5}$$

Here i denotes the segment index and W_t the width of a segment in picoseconds. The DCA versus drift time data can be broken into two regions: the region sufficiently far from the wire that ionization statistics have not significantly smeared the drift distance and the region close to the wire where we want to minimize the effect of this smearing.

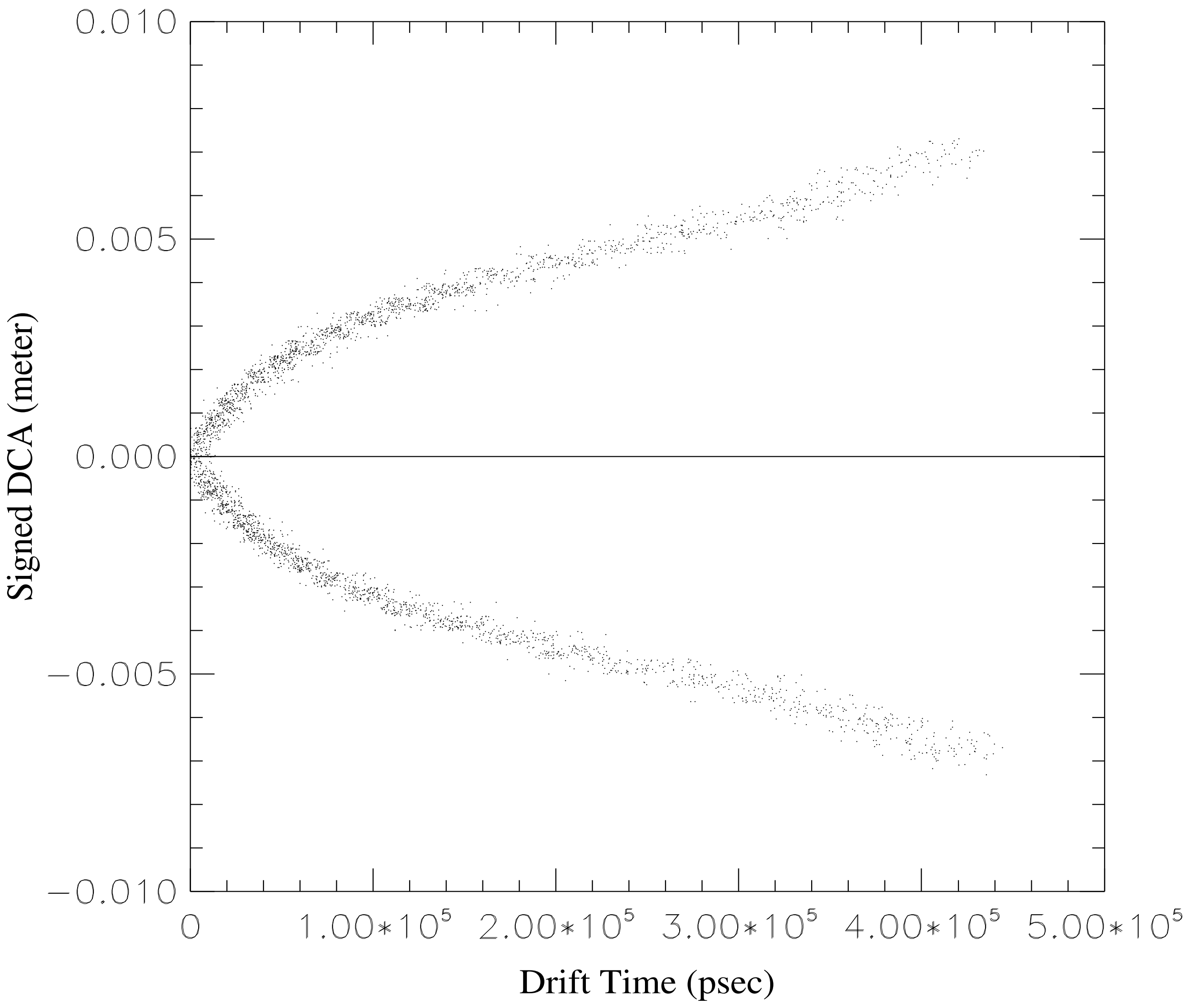


Figure 5.5: Scatter plot of $A \times$ DCA (in meters) versus the drift time (in ps) for CD Layer 26 (DR Layer 10) [86].

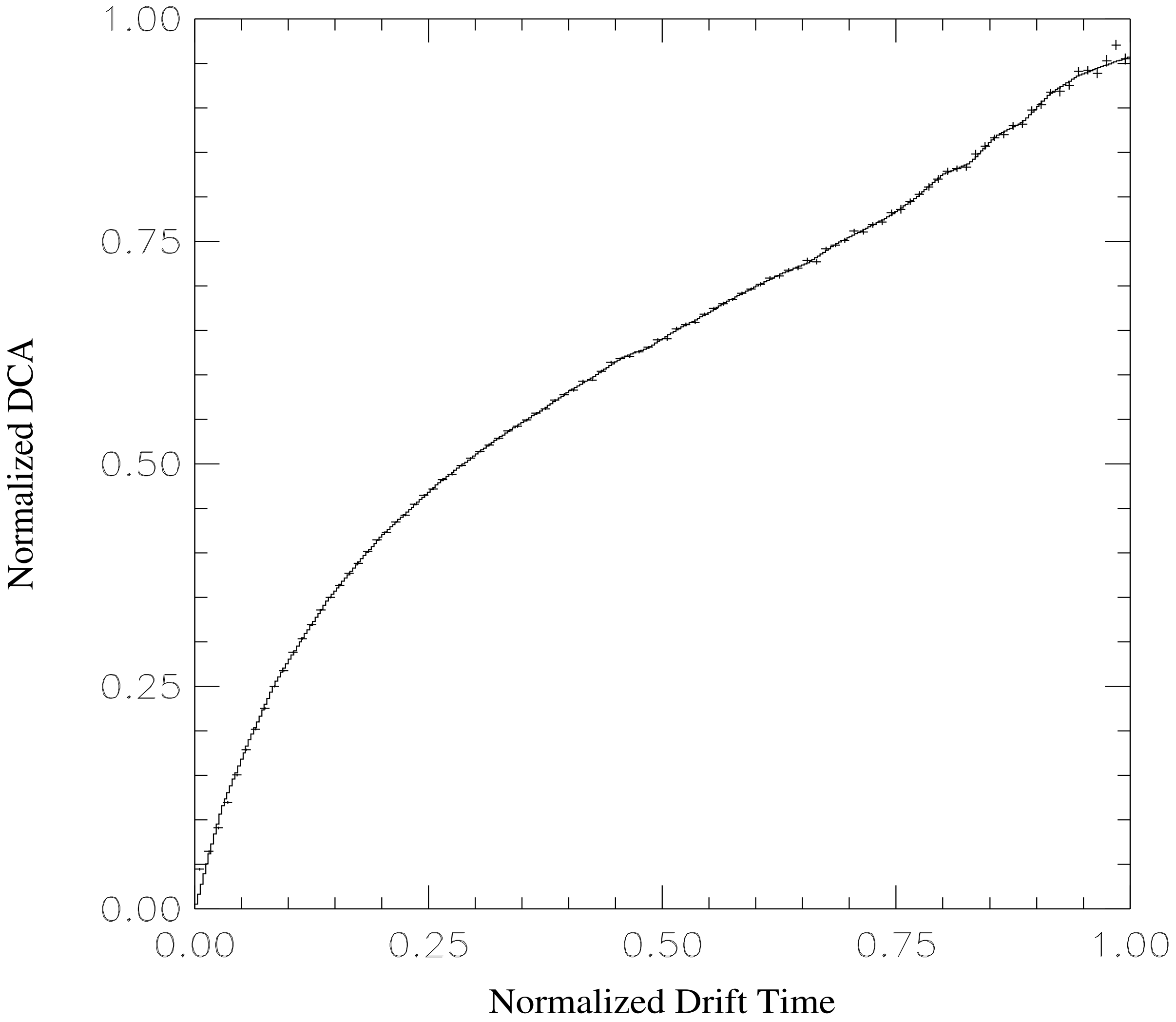


Figure 5.6: D-T function for CD Layer 26. The profile histogram of the normalized drift distance as function of the normalized drift time is represented by the crosses. The line is the result of the piecewise fit.

In the outer region it is sufficient to take the simple mean of DCA for each time bin. This is accomplished by taking a slice in time, folding the positive DCA distribution on top of the negative DCA distribution (Figure 5.5), and plotting the mean versus TM as a profile histogram (Figure 5.6). The error on the mean is plotted as the error for each entry. This treatment works far from the wire where the DCA distribution is sharply peaked and the background is low. Close to the wire, however, ionization statistics smear the distribution and ambiguity mistakes add a background. Both effects pull the mean away from the peak, rendering the profile histogram inadequate in this region (see Figure 5.7). Thus, for these small-time points we project DCA for each time slice and fit the distribution to a double Gaussian over a background function. The means of the two Gaussians are then averaged to obtain one DCA value for each time slice

To gain statistics in the small-time plots, we average small-time data from many layers together. This also imposes the desired constraint that all drift function are the same close to the wire, since one set of small time points is used for all layers in the average. Note that to be absolutely correct we should only combine layers with the same sense wire voltage. This is because even for the case of a saturated gas, where the drift speed is independent of voltage, the Lorenz angle and hence the drift function do depend on the voltage. Since all the DR layers are held at 2000 volts, we combine all the DR anode small-time data and use one set of small-time points for all DR layers. Even though the drift lines are bent in this case, they are bent by the same amount close to the wire, so the drift function is the same in this region for all the DR layers. We also average small-time data for the VD anodes. However, the VD voltages range from about 1900 volts to 2400 volts, so we assume that the effect of the electric field is not too different in different layers to significantly change the shape of the D-T function close to the wire.

Note that this practice of sharing small-time data between different layers will only work if the time zeros are sufficiently well known, since a time zero offset will shift the small-time data along the time axis. Otherwise, the small-time data will be systematically shifted relative to the outer cell data points for layers whose time residual does not happen to be close to the time residual averaged over all the

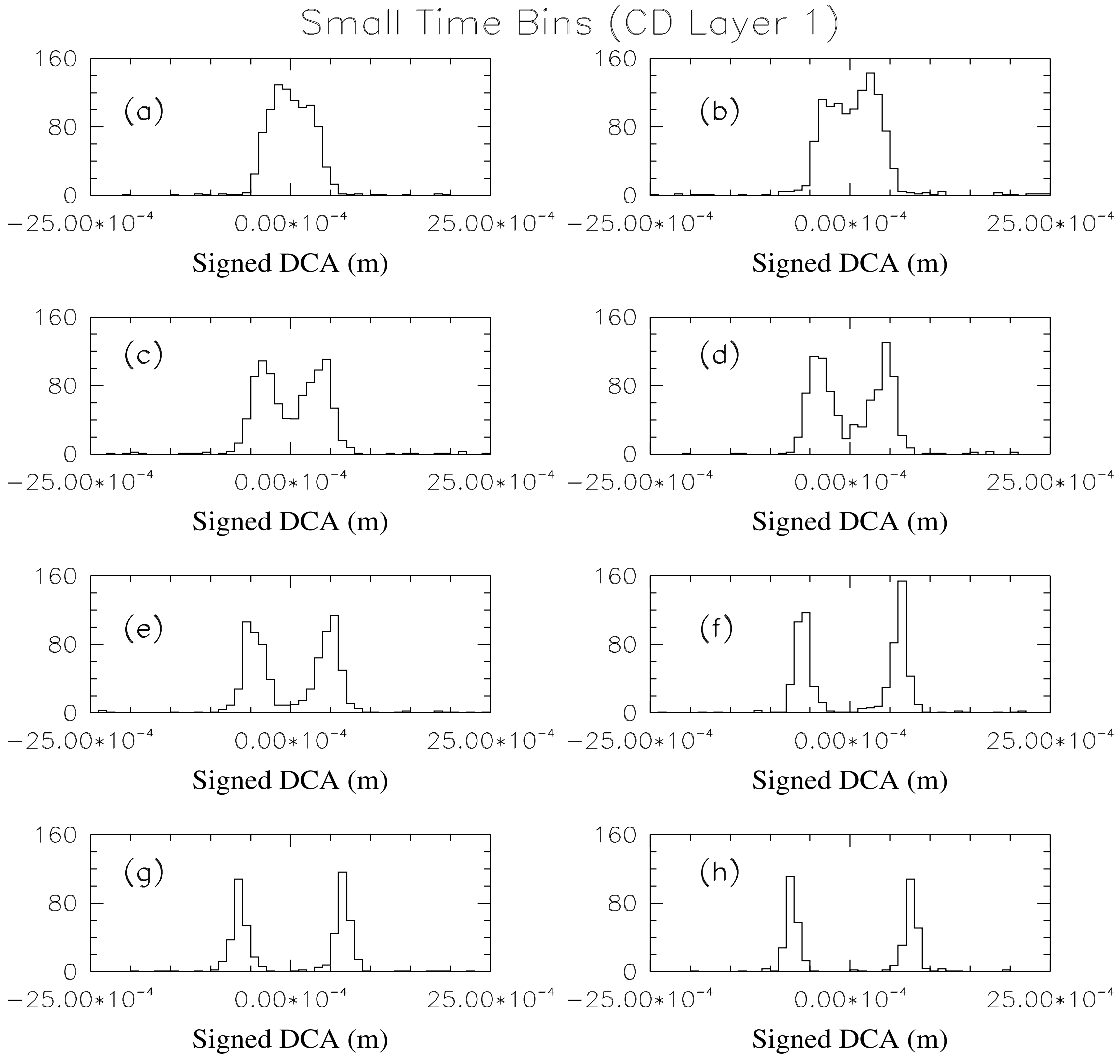


Figure 5.7: Drift distance for different small-time bins in the first layer of the PTL (*i.e.*, CD Layer 1 [86]): (a) -2 ns $\leq$ TM $<$ 0 ns, (b) 0 ns $\leq$ TM $<$ 2 ns, (c) 2 ns $\leq$ TM $<$ 4 ns, (d) 4 ns $\leq$ TM $<$ 6 ns, (e) 6 ns $\leq$ TM $<$ 8 ns, (f) 10 ns $\leq$ TM $<$ 12 ns, (g) 12 ns $\leq$ TM $<$ 14 ns, (h) 16 ns $\leq$ TM $<$ 18 ns.

layers. For the VD, small-time data consists of drift times 27 ns and smaller and for the DR, 39 ns and smaller.

We have seen how the drift time data are obtained. They are then fit according to the parameterization described by Equation (5.5) above. This parameterization is not sufficient, however, due to distorted isochrone effects far from the wire that make the DCA a function of entrance angle as well as measured time. To model this effect we divide DCA versus drift time into bins given by the sign of the ambiguity and sign of the entrance angle. Since each of these can take on two values, there are four possible permutations of the two. Thus the parameterization now becomes:

$$D(t) = A_i + B_i t + C_{jk} t^2, \quad \text{where } (i-1)W_t < t < iW_t, \tag{5.6}$$

and where j and k each take on two values given by the signs of the entrance angle and ambiguity. The quadratic corrections for CD layer 26 is shown in Figure 5.8. This is definitely not the best we could do. A better parameterization would be one which introduced a more finely grained entrance angle dependence in the drift function. An entrance angle correction based on the shape of the drift line in Figure 5.3 was developed for the recompress data samples.

This quadratic parameterization is inappropriate for a few badly behaved layers. A badly behaved layer is one adjacent to a cathode layer. In this case the field is highly distorted. For such layers, the drift time data diverges so much and the maximum drift time is so different for positive and negative ambiguities that the quadratic parameterization fails. Here we use the simple linear parameterization (Equation (5.5)), but make two D-T functions for each layer, one for each ambiguity, as shown in Figure 5.9. The earlier arguments used to justify averaging small-time data from many layers still apply; the "bad behavior" occurs in the outer part of the cell. Thus, the same small-time data is used for these layers as for the other layers.

PTL D-T Funtions

The PTL drift functions are generated in much the same way as the VD and DR D-T functions, with a few differences. As before, the DCA is plotted versus drift time for each layer and fitted to a piecewise linear function. The data are comprised of

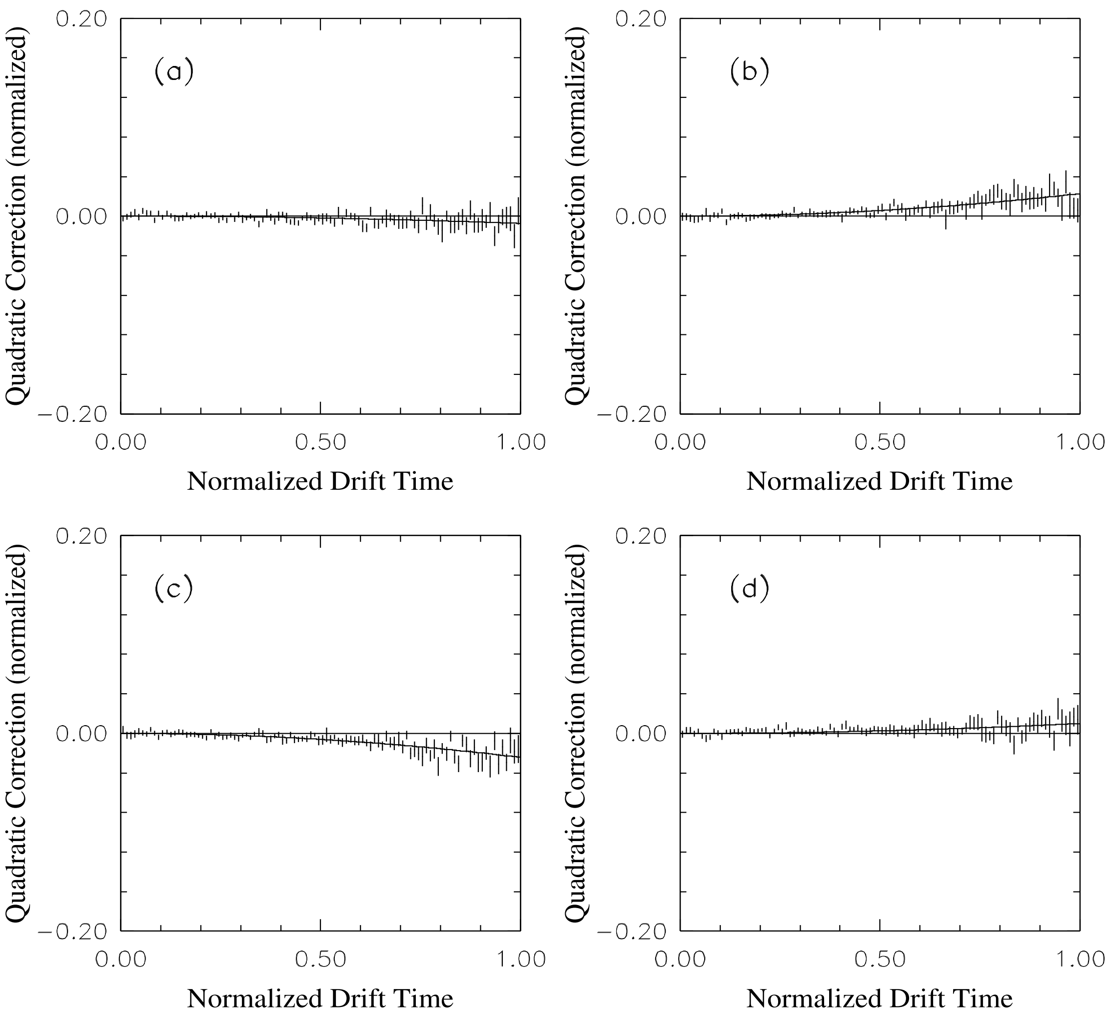


Figure 5.8: Quadratic Corrections for CD Layer 26. (a) $A = -1$ and negative entrance angles, (b) $A = +1$ and negative entrance angles, (c) $A = -1$ and positive entrance angles, and (d) $A = +1$ and positive entrance angles.

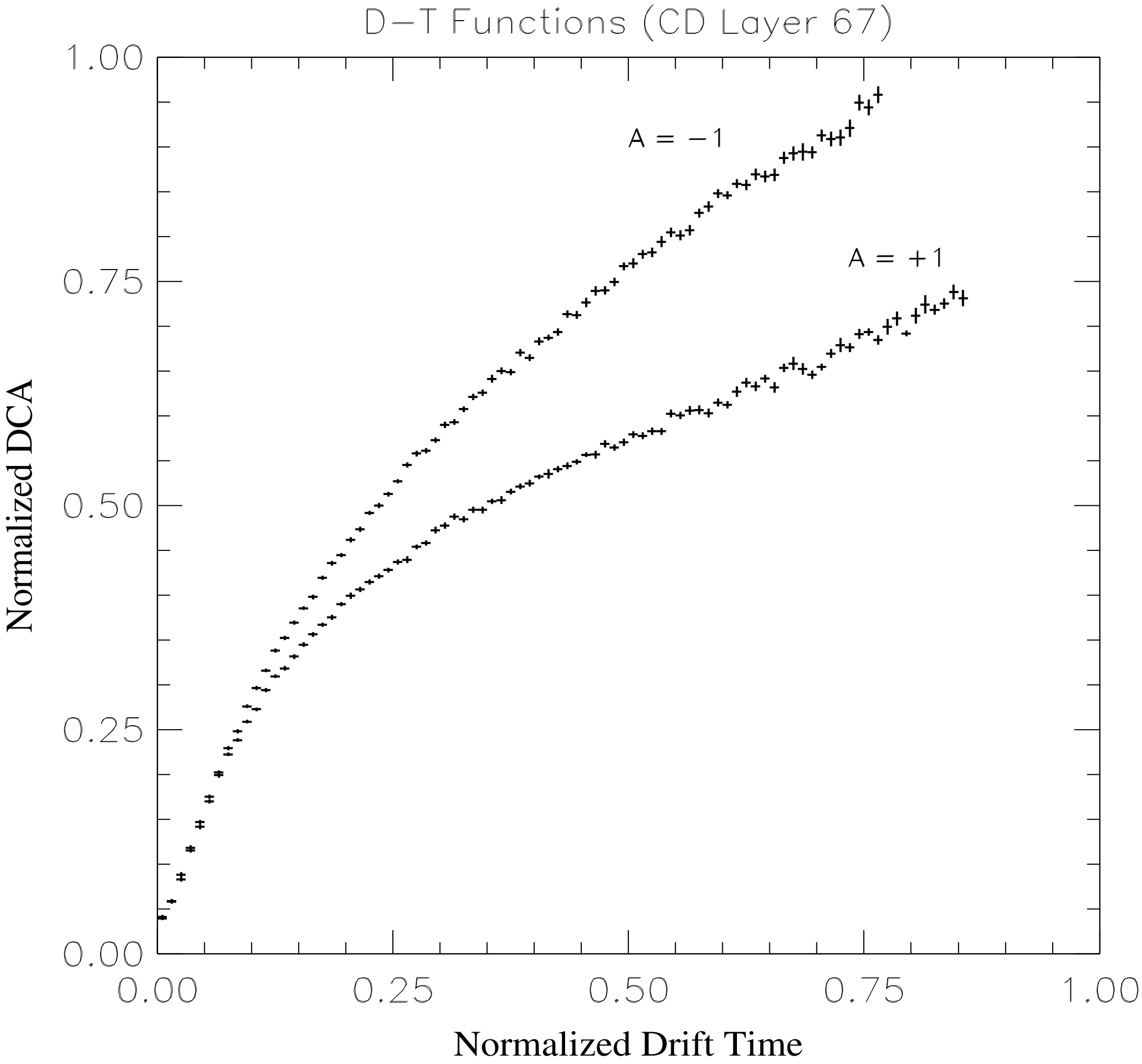


Figure 5.9: Two-sided D-T function for CD Layer 67.

small-time data and outer cell data in the same way. However, since DME (DiMethyl Ether) is not a saturated gas, the drift speed depends on the voltage of the sense wire and because the PTL layers are held at different voltages, small-time data may not be shared among different layers. In the PTL, small-time data consist of drift times 14 ns and smaller.

The drift function is fitted in the same manner as before, but to many more segments, typically one-hundred. The large number of segments is necessary to model the rapidly changing drift speed of DME; the speed can range from 5 μm/ns to 100 μm/ns over a cell. The large number of segments can lead to a drift function that is bumpy, though, and so a constraint is applied to the fit to keep the drift function smooth; the first derivative is required to decrease monotonically. This is a physical trait of the PTL D-T functions which we observe directly and, therefore, a valid constraint.

The PTL D-T functions are simpler than the VD and DR in that the circular symmetry of the straw tubes provides rotationally invariant isochrones and therefore no dependence on ambiguity or entrance angle. The simple linear parameterization of Equation (5.5) is sufficient.

DR stereo D-T Funtions

The stereo layer D-T functions are quite difficult to calibrate effectively. Since the sense wire is not parallel to the z-axis, its position within the cell changes with z. As a result, the D-T function which we are trying to measure is z-dependent. We could consider making several drift functions for each stereo layer in bins of z, but obtaining sufficient statistics would be impractical. Instead we make one drift function integrated over z for each stereo layer and then apply a z-dependent correction. The corrections are made by plotting, in profile format, space residual versus z in two bins of ambiguity, five bins of drift distance, and two bins of stereo angle[1]. This translates to twenty plots. The plots are fit to a piecewise linear scheme (Figure 5.10) and the

[1] A stereo layer is like an axial layer which has been grasped at each end and rotated in opposite directions in ϕ, or in other words, twisted about the z axis. Obviously there are two choices of direction to twist. In the DR alternate stereo layers have alternating senses of this twist, or alternating sign of the stereo angle.

result becomes the correction to the simple first order drift function:

$$\begin{aligned} &D(t,z) = A_i + B_i t + C_{klmn} + D_{klmn} z \\ &\text{where } (i-1)W_t < t < iW_t \text{ and } (m-1)W_z < z < mW_z, \end{aligned} \tag{5.7}$$

where *klmn* denote the range over ambiguity, drift distance, z bin, and stereo angle. We openly admit that this fix does not reflect a deep understanding of the actual behavior of these layers [87]. It is merely a means of compensating for the error in the drift function introduced by averaging over z. Nevertheless, this correction results in a 5% improvement in resolution for stereo hits with $|z| > 0.5$ m.

5.3.3 The Time Zeros

The timing is critical in the drift chambers because every nanosecond of jitter adds 25-50 μm to the tracking resolution. As discussed in the introduction, a charged particle traversing a drift cell ionizes the gas in the tracking chamber and produces an electronic pulse on the sense wire. The wire pulse then travels along the wire and, if it crosses the discriminator threshold, starts the discharge of the timing capacitor. Using the z coordinate convention of CLEO II, the propagation time of the wire signal can be described as:

$$T_p = \begin{cases} L/2v + z/v & \text{East Readout} \\ L/2v - z/v & \text{West Readout} \end{cases} \tag{5.8}$$

where the term $L/2v$ is the time the pulse takes to travel half the length of the sense wire plus all the electronic delays that may be incorporated before the pulse starts the timing circuit. Consequently,

$$L/2v = l/2v + \text{delay}_{\text{wire}}, \tag{5.9}$$

where l is the physical length of the sense wire and $\text{delay}_{\text{wire}}$ is the delay for a given wire pulse to travel down the electronics. On the preamplifier board, the route of the signal from individual wires is different. In other words, $\text{delay}_{\text{wire}}$ is a function of the channel associated with a given sense wire. The relative start time for a wire input

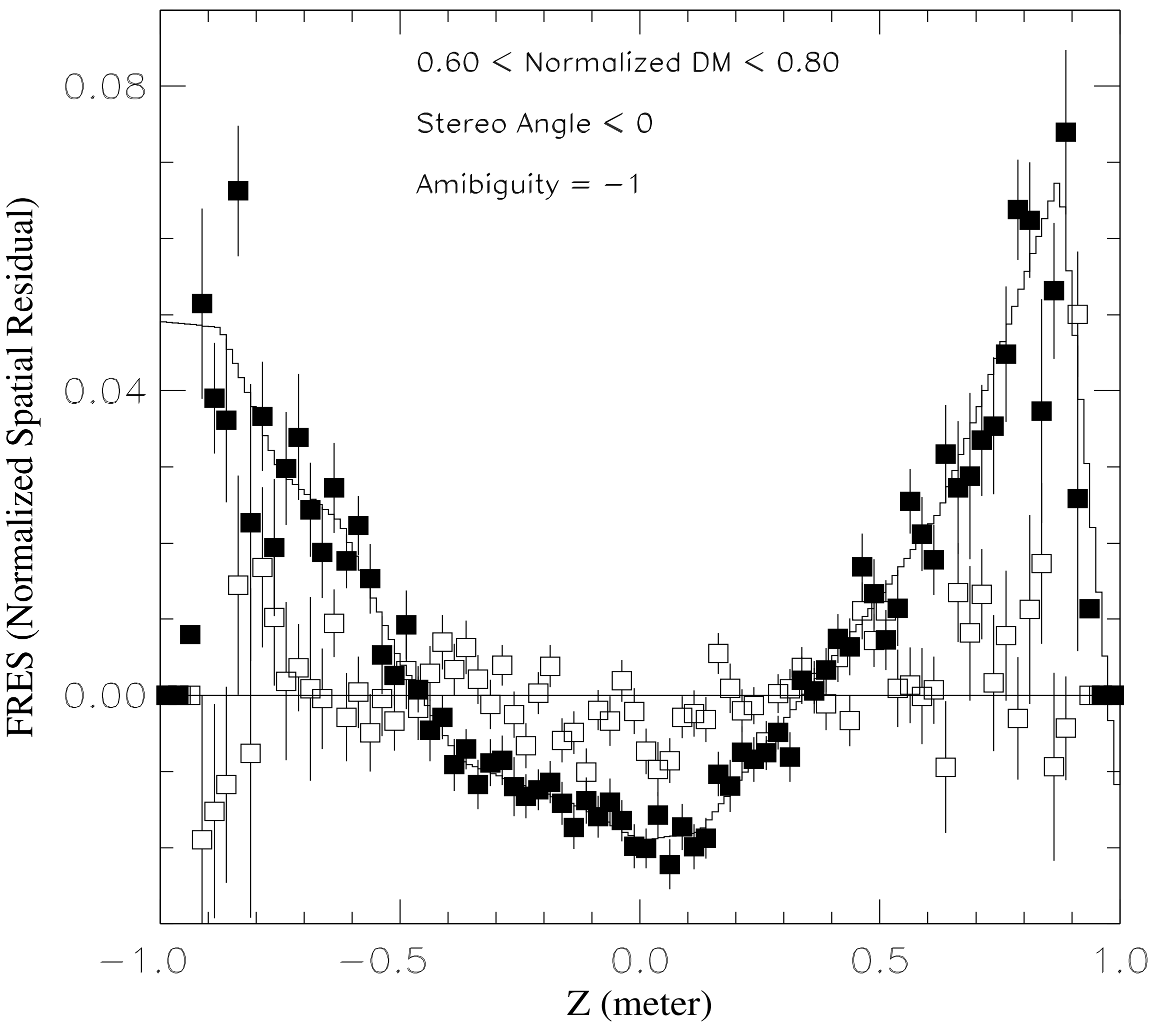


Figure 5.10: The normalized spatial residual (FRES) versus Z for the DR stereo layers with $A = -1$, stereo angle < 0 and $0.6 <$ DM < 0.8. The filled squares are the normalized spatial residuals with no correction. The lines segments represent the fit. The open squares are the normalized spatial residuals after the correction has been applied.

on a DR preamplifier board is shown in Figure 5.11(b). To remove this non-linear wire-to-wire dependence, each channel (n) has its own TDC-to-time relation. The TDC-to-time relation is obtained with a calibration pulse [79]:

$$\mathrm{TDC}(n) = s_0 + s_1 t + s_2 t^2 + \cdots, \tag{5.10}$$

where the term s_0 removes most of the wire-to-wire differences. As one can see in Figure 5.12, the route of the calibration pulse is slightly different than the one taken by the wire signal. The pulser circuit is attached directly to the side of each DR preamplifier board. The pulse signal travels from the front of the board to each preamplifier. The wire-to-wire variation left over has a linear dependence due to the layout of the preamplifier boards. The relative start time for a calibration input pulse is shown in Figure 5.11(a). Figure 5.11(c) shows the remaining delay (*i.e.*, $\Delta T = \mathrm{delay}_{\mathrm{pulse}} - \mathrm{delay}_{\mathrm{wire}}$) and its linear dependence on the channel (or wire) number. Therefore, the total propagation delay on the preamplifier board can be parameterized as $\Delta T = \mathrm{SLOPE} \times n + \mathrm{OFFSET}$, where n represents the channel number. Thus, channel-to-channel corrections can be performed based on the dependence shown in Figure 5.11(c). By design, there is no significant propagation delay on the PTL and VD preamplifier boards. To reduce the delays and the electronic noise, the preamplifier boards are mounted as close as possible to the end of the sense wires. In the PTL and the VD, coaxial cables (∼95 cm long) run from the sense wires, attached to the endplate, to the preamplifiers boards. In the PTL, the calibration pulse does not travel through these coaxial cables because it is sent directly to the preamplifier boards. Consequently, length variation in the coaxial cables between the sense wires and the endplate introduces arbitrary delays in the PTL channels. Without knowing the exact length of those cables, it is impossible to find the functional dependence of the delays versus the channel number; and, one cannot constrain this dependence to be linear. In the VD, the calibration pulse is sent through the coaxial cables to the power supply board and then to the preamplifier boards. Therefore, in the PTL and the VD, the total propagation delay on the preamplifier boards has SLOPE = 0. In the DR, the preamplifier boards are directly mounted on the endplate. Since the propagation delay in the DR preamplifier board is significant (see Figure 5.11(c)), the

SLOPE is not set to zero and has to be extracted.

In the drift chambers, when the CLEO II detector is read out, the timing capacitors are stopped by the lowest trigger, Level 0 [65]. Since each chamber operates in common stop mode, its global time zero (T_0) is defined by its own trigger stop relative to the known interaction time. Depending on the device, some fix delays may be added between the Level 0 gate and the actual trigger stop of a given channel. These delays are due to the fact that the route of the stop signal to the circuit which closes the timing capacitor is different for each preamplifier board and for each channel on a given preamplifier board. Again, all the stop signal delays are removed channel-by-channel during the *electronic* calibration because each channel has its own offset (s_0) in the TDC-to-time relation (Equation (5.10)).

Before any precise track reconstruction, the tracking algorithm needs a drift distance for each sense wire that recorded a hit. The time zero is then essential to map time to distance correctly. The time zero is divided between a crate time zero, a preamplifier board time zero, and a wire time zero:

$$T_0 = T_{\rm crate} + T_{\rm preamp} + T_{\rm wire}. \tag{5.11}$$

This structure was implemented to allow for compensation of offset variations on different time scales. The crate component was intended for run by run variations, but since these turn out to be negligible the crate time zeros are currently not varied. The values currently in use[2] have been fixed from an online calibration [79].

The preamplifier board time zero is determined by the *offline* calibration. The time zero is computed on a preamplifier board basis to incorporate the propagation delay in the preamplifier board. The form of the preamplifier board time zero for a given channel n is described as:

$$T_{\rm preamp}(n) = {\rm SLOPE} \times n + {\rm OFFSET}. \tag{5.12}$$

The wire time zero is a time shift for each individual wire in the drift chambers. Since each channel or wire has its own TDC-to-time relation, the major time shift between

[2]The CESR timing implies the $T_{\rm crate}$ for the PTL, VD and the DR to be 447 ns for seven bunch running and 340 ns for nine bunch running.

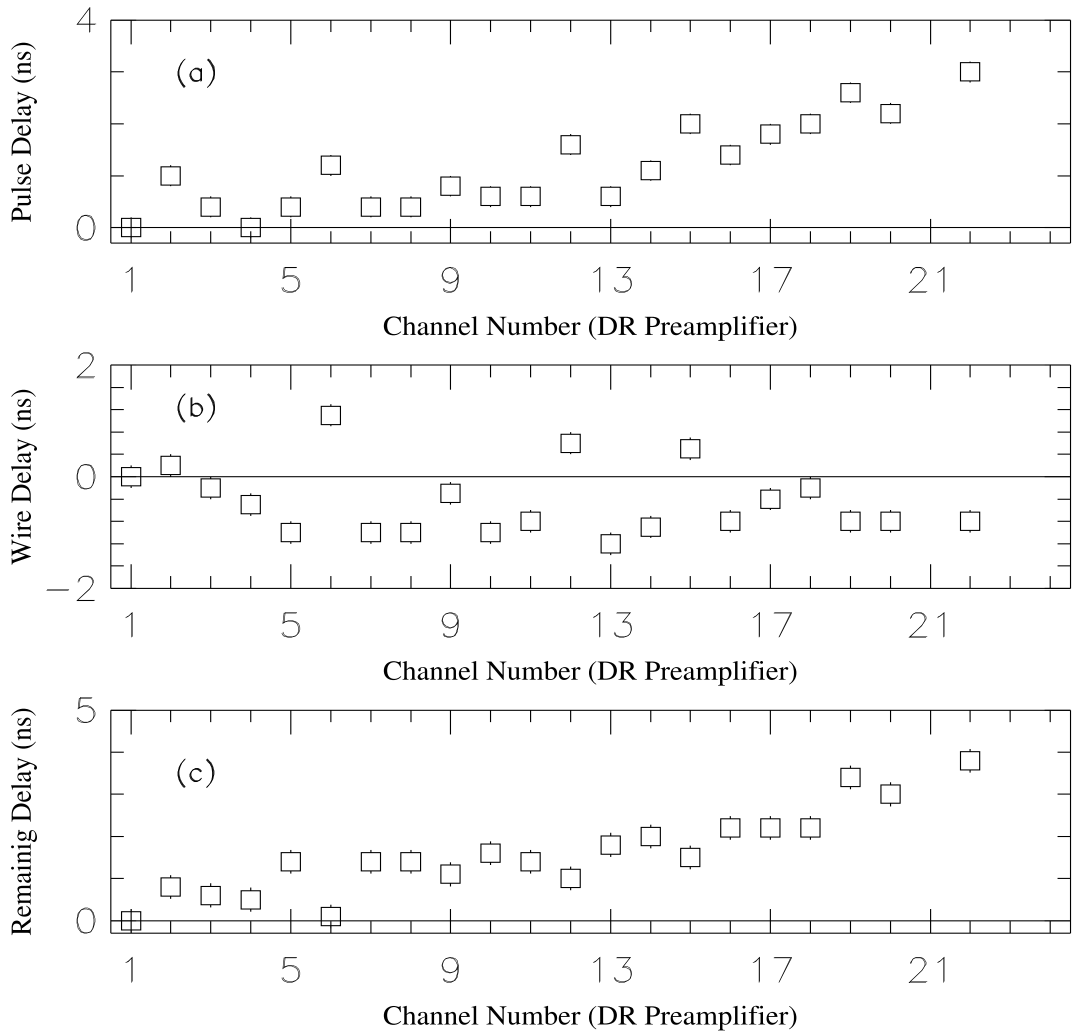


Figure 5.11: Propagation delay in the DR preamplifier board: (a) relative start time for a calibration input pulse, (b) relative start time for a wire input pulse, and (c) the difference between the relative start time for the calibration and the wire pulses. It is crudely a straight line. It represents the time taken by the calibration pulse to propagate across the preamplifier board.

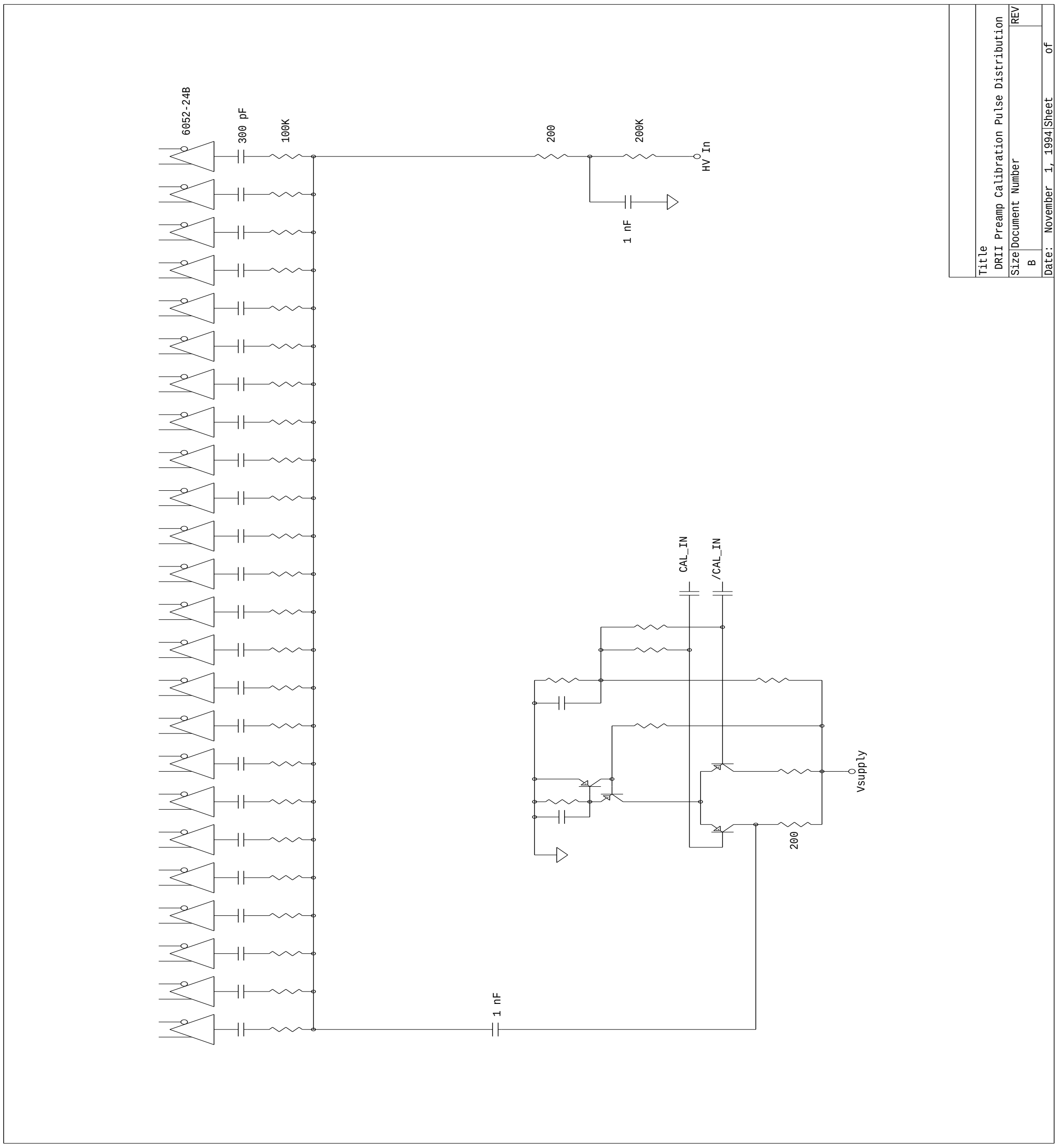


Figure 5.12: DR Preamplifier board pulser circuit. The difference in path between the calibration and signal pulses add a linear dependence versus channel number for the propagation delay.

wires has already been taken out. It is obvious that T_{wire} absorbs all other time zero fine tuning. In summary, one can see the time zero as a first order approximation ($T_{\rm crate}$), plus a correction for propagation delays in the electronics ($T_{\rm preamp}$), and plus a wire correction ($T_{\rm wire}$).

As mentioned earlier, the time zeros are determined via a minimization of the time residuals. An accurate determination of time residuals is then needed. The time residuals are obtained from the internal fit to each dualtrack. For a given channel, a value of TRES greater than zero implies that the drift time is greater that the calculated time and vice versa.

$$\mathrm{TRES} = \mathrm{TM} - \mathrm{TC} > 0 \Longleftrightarrow \mathrm{TM} > \mathrm{TC}. \tag{5.13}$$

Therefore, a positive (negative) value of TRES implies an overestimate (underestimate) of the time zero for this particular channel.

The extraction of the time residuals for the PTL and for the VD is based on the determination of the time offset for each preamplifier board with no channel-to-channel dependence. In the DR, time residuals are found for the group of wires connected to a given preamplifier board. The same sample of Bhabha events used for the determination of the D-T functions is used for determining the time zeros. The preamplifier board time residuals are determined and subtracted from the preamplifier board time zeros and the time residuals are re-measured with the new time zeros until the time residuals converge to zero for all preamplifier boards. In general, the convergence of TRES is reliable (see Figure 5.13). New time zeros are updated before and after the extraction of a new D-T function.

The common belief is that steady running conditions modulate the timing of the chambers by less than a few nanoseconds and do not actually change the shape of the D-T functions. The adjustment of the time zeros as a function of time is important in order to maintain a good spatial resolution in each cell. The goal of the calibration procedure is to make |TRES| less than one nanosecond.

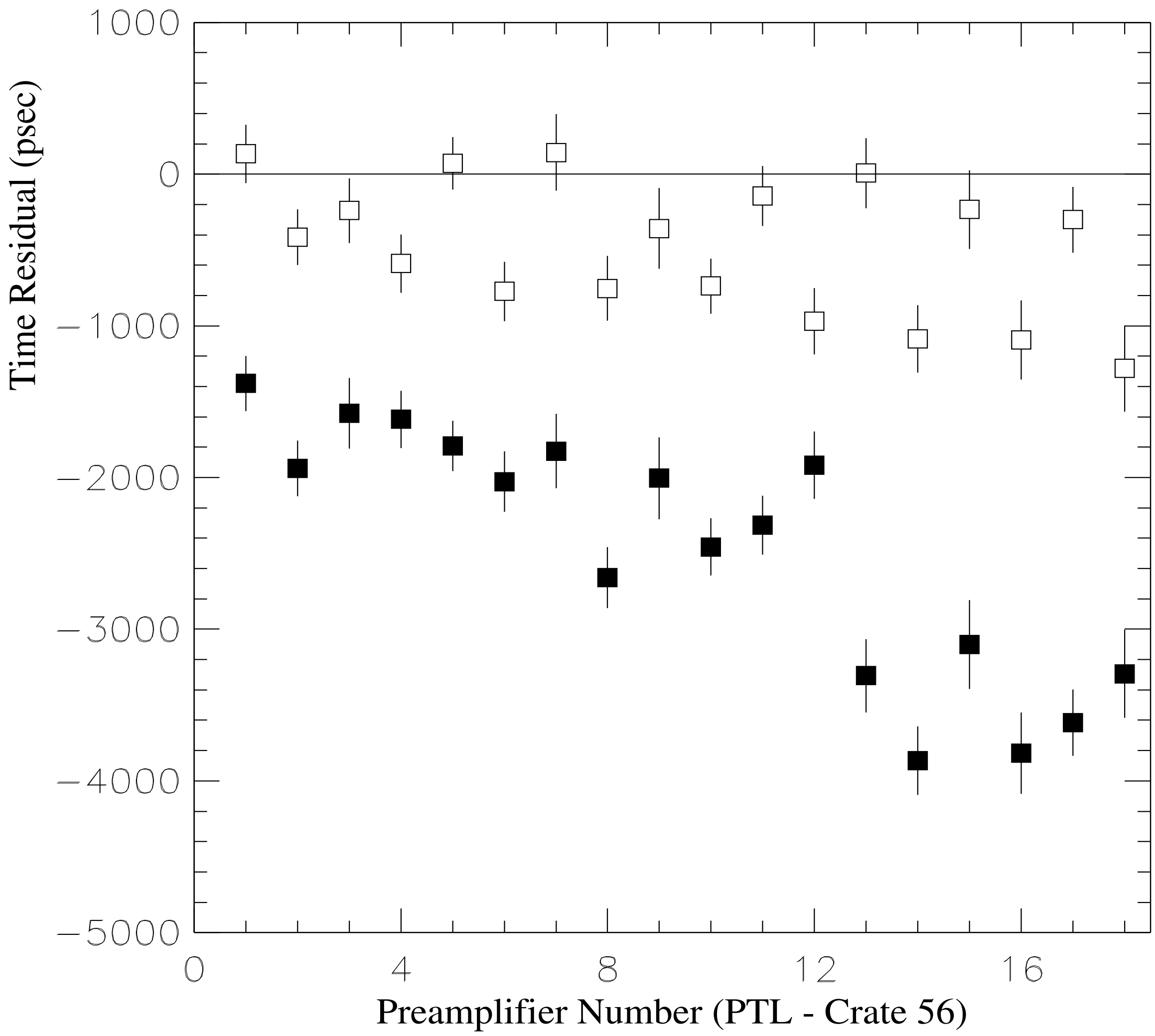


Figure 5.13: Time Residual versus preamplifier board number for the PTL (Crate 56). The filed squares show TRES before any re-adjustment. The open squares represents TRES after one iteration.

Signal Propagation Corrections

The propagation time (T_p) in the chambers is the lapse of time between the arrival of the drift electrons to the sense wire and the firing of the drift cell electronics. By using Equation (5.8) and the convention:

$$\mathcal{Z} = \begin{cases} -z & \text{East Readout} \\ +z & \text{West Readout} \end{cases}$$

the propagation time is simply given by:

$$T_p = L/2v - \mathcal{Z}/v. \tag{5.14}$$

The first term is just the time that the pulse takes to travel the length of the wire for hits at z=0, plus all the electronics delays that may be incorporated before the pulse crosses the discriminator threshold. Consequently, the constants term $L/2v$ is already included in the preamplifier board time zero. Thus, $T_p = -\mathcal{Z}/v$. The value of the (mean) speed of the pulse along the wire is assumed to be the speed of light ($v = c$) for all devices.

The VD, which was designed to perform charge division, has highly resistive sense wires and has electronic readouts at each end. These two features make another signal propagation correction necessary. An empirical correction takes into account non-linear effects such as signal reflection at the readout electronics and transmission line effects that cause the propagation speed to depend on the frequency composition of the pulse.

The nonlinearity of the signal propagation along the wire can be seen in the distribution of the time residual as a function of the $\mathcal{Z}$ coordinate (see Figure 5.14). The signal propagation correction is parameterized by a piecewise fit of seven connected straight line segments of equal width in $\mathcal{Z}$ (see Appendix D). The width of each segment is 10 cm. The number of segments and their widths are optimized to fit the data points with the smallest χ^2. Then, for the VD,

$$T_p = -\mathcal{Z}/v + \text{TSCOR}(\mathcal{Z}) \quad \text{where } \text{TSCOR}(\mathcal{Z}) = a_m\mathcal{Z} + b_m \ \cdot \tag{5.15}$$

The coefficients a_m and b_m are simply the slopes and the intercepts of the line segments in the piecewise fit of the time residual distribution versus $\mathcal{Z}$. Such a correction gives a 12 μm improvement in the width of the residual distribution.

5.3.4 Summary: Drift Time and Drift Distance

The goal of the drift chambers is to sample the trajectory of a charged particle with the information of each individual drift cell. Along with the calibration *constants*, a drift time and a drift distance are measured for all the cells that recorded a hit. In this section, the equations for TM and DM are summarized. The measured time or drift time is given by:

$$\mathrm{TM} = T_0 - T(\mathrm{TDC}) - \mathrm{Arclength}/\beta c - T_p, \tag{5.16}$$

where

$$\begin{aligned} \text{Raw time} &= T(\mathrm{TDC}) \\ \text{Transit time} &= \mathrm{Arclength}/\beta c \\ T_p &= \begin{cases} -\mathcal{Z}/v & \text{In the PTL and DR} \\ -\mathcal{Z}/v + \mathrm{TSCOR}(\mathcal{Z}) & \text{In the VD} \end{cases} \end{aligned}$$

with

$$\mathrm{TSCOR}(\mathcal{Z}) = a_m \mathcal{Z} + b_m \quad \text{where } (m-1)W_{\mathcal{Z}} < \mathcal{Z} < mW_{\mathcal{Z}} \ \cdot \tag{5.17}$$

Further, the measured drift distance is given by:
In the PTL,

$$D(t) = A_i + B_i t \quad \text{where } (i-1)W_t < t < iW_t \ , \tag{5.18}$$

while in CD layers 7, 16, 17 and 67 (two-sided drift functions),

$$D(t) = A_{ij} + B_{ij} t \quad \text{where } (i-1)W_t < t < iW_t \ , \tag{5.19}$$

and in the stereo layers,

$$\begin{aligned} &D(t,z) = A_i + B_i t + C_{klmn} + D_{klmn} z\,, \\ &\text{where } (i-1)W_t < t < iW_t \text{ and } (m-1)W_z < z < mW_z, \end{aligned} \tag{5.20}$$

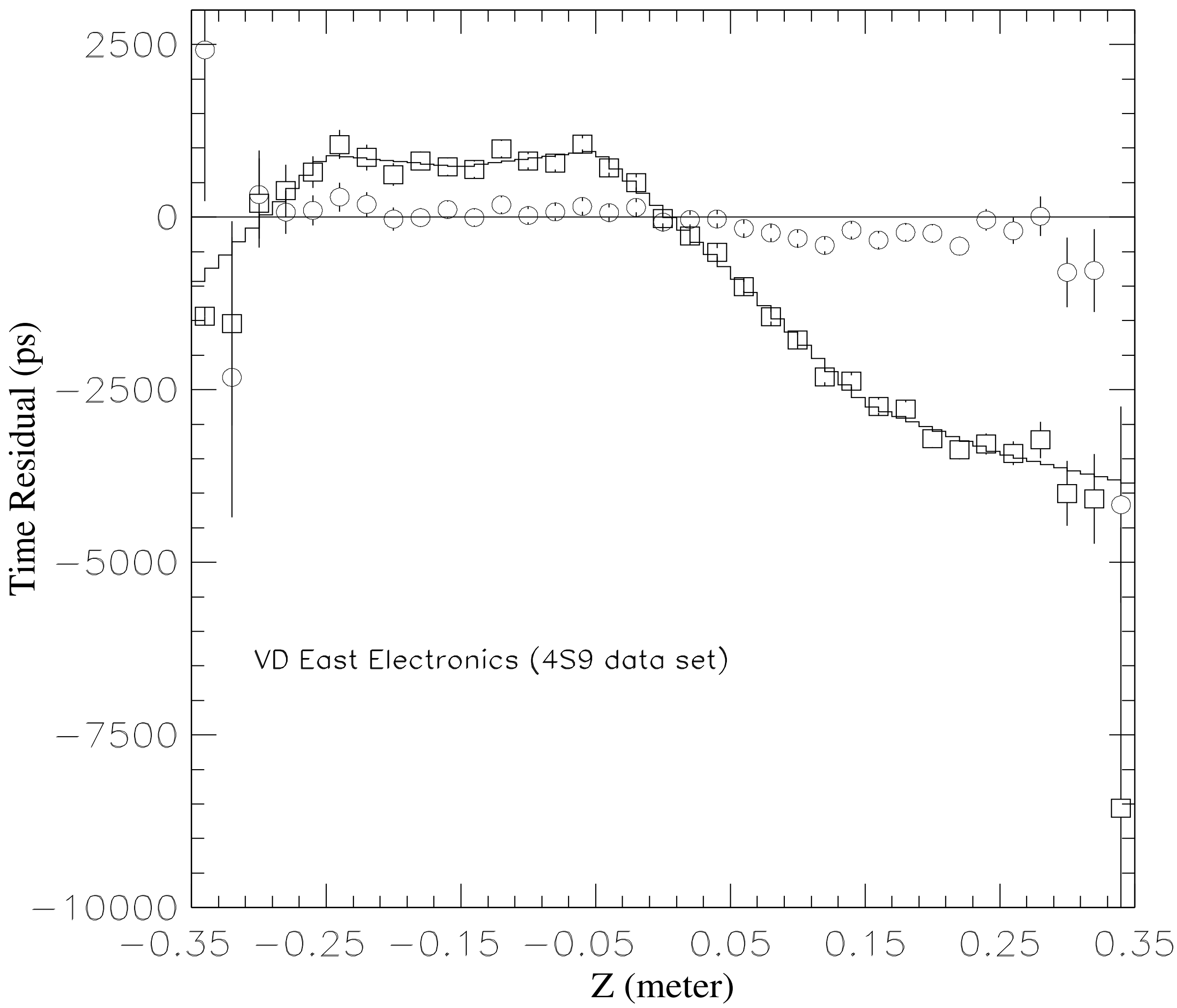


Figure 5.14: Time residuals (for all values of pulse height) as a function of $\mathcal{Z}$ for the east readout of the VD anodes. A fit to this raw distribution determined a_m and b_m. The squares are the average time residual without correction and the line is the result of the fit. The circles are the time residuals after the correction has been applied.

In all the other layers of the VD and DR,

$$D(t) = A_i t + B_i t + C_{jk} t^2 \quad \text{where } (i-1)W_t < t < iW_t \ , \tag{5.21}$$

where

- t : The drift time TM.
- W_t: The width of a time segment in picoseconds.
- W_z: The width of a z segment in meters.
- A_i and B_i: Intercept and Slope of the piecewise fit for the i^{th} segment.
- The indices j and k refer to the sign of the entrance angle and the sign of the ambiguity, respectively.
- A_{ij} and B_{ij}: Intercept and Slope of the piecewise fit for the i^{th} segment and a given ambiguity.
- C_{jk}: Quadratic correction coefficient.
- The indices l, m and n refer to drift distance bin, z bin and the sign of the stereo angle, respectively.
- C_{klmn} and D_{klmn}: The correction coefficients to the stereo layer D-T function.

5.3.5 Geometry Alignment

Between each data run, the CLEO II pole tips are opened for periodic detector maintenance. The displacement of the pole tips can change the relative positions of the chambers. Then, before the start of each data compression, one has to look at a set of diagnostic histograms in order to detect any alignment anomalies. The geometry alignment procedure is described in this section.

The alignment of the tracking chambers relies on a moment analysis of the residuals. Using the fact that the tracking chambers are uniformly illuminated, one can project a track from the main drift chambers into the inner chambers and compute the track residuals. The inner chambers (PTL and VD) are aligned relative to the DR because the main drift chamber defines the CLEO coordinate system. The geometry alignment constants are therefore *weighted* track shifts. A track shift ($\Re_{\mathrm{IFIT},i}$) for a given internal fit is defined as the DCA of a hit (label i) calculated during the DR internal fit minus the DCA of the same hit calculated during the internal fit IFIT. Then, a VD track shift is $\Re_{2,i} = A \times (\mathrm{DCA}_{1,i} - \mathrm{DCA}_{2,i})$, and a PTL track shift is $\Re_{3,i} = A \times (\mathrm{DCA}_{1,i} - \mathrm{DCA}_{3,i})$. Obviously, $\Re_{1,i} = 0$ by construction. The weights in the definition of the geometry alignment variables are a function of the radius of the layer (R), the azimuthal angle, and/or the polar angle (ϕ_{IFIT} and/or θ_{IFIT}) of the dualtrack found by the internal fit IFIT. The geometry constants are:

$$
\begin{aligned}
\mathrm{GTRYY} &= \Re_{\mathrm{IFIT},i} \times \cos\phi_{\mathrm{IFIT}} \\
\mathrm{GTRXX} &= \Re_{\mathrm{IFIT},i} \times \sin\phi_{\mathrm{IFIT}} \\
\mathrm{GRTXY} &= \Re_{\mathrm{IFIT},i}/R_i \\
\mathrm{GTLYZ} &= \Re_{\mathrm{IFIT},i} \times \cos\phi_{\mathrm{IFIT}} \times \cos\theta_{\mathrm{IFIT}}/R_i \\
\mathrm{GTLXZ} &= \Re_{\mathrm{IFIT},i} \times \sin\phi_{\mathrm{IFIT}} \times \cos\theta_{\mathrm{IFIT}}/R_i
\end{aligned}
\tag{5.22}
$$

The relative position between two Cartesian coordinate systems can be described by:

$$
\begin{aligned}
\triangle y \quad &= \text{Translation along the y-axis} \\
\triangle x \quad &= \text{Translation along the x-axis} \\
\omega \quad &= \text{Rotation about the z-axis} \\
\delta_y \quad &= \text{Tilt about the y-axis} \\
\delta_x \quad &= \text{Tilt about the x-axis}
\end{aligned}
\tag{5.23}
$$

Consider two small translations $x \xrightarrow{\mathcal{A}} x + \triangle x = x(1+\epsilon_x)$ and $y \xrightarrow{\mathcal{A}} y + \triangle y = y(1+\epsilon_y)$ represented by the transformation matrix

$$\mathcal{A} = \begin{pmatrix} 1+\epsilon_x & 0 & 0 \\ 0 & 1+\epsilon_y & 0 \\ 0 & & 1 \end{pmatrix}, \tag{5.24}$$

and consider three small rotations represented by the transformation matrices

$$\mathcal{B} = \begin{pmatrix} \cos\omega & \sin\omega & 0 \\ -\sin\omega & \cos\omega & 0 \\ 0 & 0 & 1 \end{pmatrix} \simeq \begin{pmatrix} 1 & \omega & 0 \\ -\omega & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{5.25}$$

$$\mathcal{C} = \begin{pmatrix} \cos\delta_y & 0 & \sin\delta_y \\ 0 & 1 & 0 \\ -\sin\delta_y & 0 & \cos\delta_y \end{pmatrix} \simeq \begin{pmatrix} 1 & 0 & \delta_y \\ 0 & 1 & 0 \\ -\delta_y & 0 & 1 \end{pmatrix}, \tag{5.26}$$

$$\mathcal{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\delta_x & \sin\delta_x \\ 0 & -\sin\delta_x & \cos\delta_x \end{pmatrix} \simeq \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \delta_x \\ 0 & -\delta_x & 1 \end{pmatrix}. \tag{5.27}$$

One can carry out the transformation from a given Cartesian coordinate system to another by means of successive translations and rotations performed in an unspecified sequence since we are considering rather small transformations ($\mathcal{G}$) of the form $\mathbf{r} \xrightarrow{\mathcal{G}} \mathbf{r} + \delta\mathbf{r}$ with $|\delta\mathbf{r}|/|\mathbf{r}| \ll 1$. In CLEO, $\delta\mathbf{r}$ is the relative position of the VD or PTL center to the DR center. Hence the matrix of the complete alignment transformation can be written as:

$$\mathcal{G} = \mathcal{ABCD} \simeq \begin{pmatrix} 1+\epsilon_x & \omega & \delta_y \\ -\omega & 1+\epsilon_y & \delta_x \\ -\delta_y & -\delta_x & 1 \end{pmatrix}, \tag{5.28}$$

where all the second order terms had been dropped. Symbolically, the alignment transformation can be written as followed:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \mathcal{G} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + \triangle x + \omega y + \delta_y z \\ -\omega x + y + \triangle y + \delta_x z \\ -\delta_y x - \delta_x y + z \end{pmatrix}. \tag{5.29}$$

Since the drift distance are projections in the $r - \phi$ plane, the track shifts ($\Re_{\mathrm{IFIT},i}$) are also projections in the the $r - \phi$ plane. In our discussion, $dz = (z' - z)$ is then irrelevant and only $dx = (x' - x)$ and $dy = (y' - y)$ matter. Equation (5.29) becomes

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x + \triangle x + \omega y + \delta_y z \\ y + \triangle y - \omega x + \delta_x z \end{pmatrix}. \tag{5.30}$$

It leads to the difference $\delta \mathbf{r} = dx\ \hat{x} + dy\ \hat{y} = (x' - x)\hat{x} + (y' - y)\hat{y}$ between the Cartesian coordinate system of reference and another Cartesian coordinate system (in our case the DR is the reference and the VD or the PTL is the other system). In the DR Cartesian coordinate system, the coordinate of a projected hit is given by P' and the wire position is given by (x', y'); and in the other Cartesian coordinate system, the coordinate of the same hit is given by P and the wire position is given by (x, y). See Figure 5.15. Using the definition that $R \equiv$ radius of the layer, $D \equiv A \times \mathrm{DCA}_{\mathrm{IFIT}}$ and $D' \equiv A \times \mathrm{DCA}_1$ for a given hit, we have[3]

$$\begin{aligned} x &= R\cos\phi + D\sin\phi \\ y &= R\sin\phi - D\cos\phi \\ z &= R\cot\theta \end{aligned} \tag{5.31}$$

and

$$\begin{aligned} x' &= R\cos\phi' + D'\sin\phi' \\ y' &= R\sin\phi' - D'\cos\phi' \end{aligned} \tag{5.32}$$

[3] In Equations (5.31) and (5.32), the assumption of infinite radius of curvature (i.e. $\mathrm{CU} = q/2\rho = 0$) is made for the calculation of $(x, y)_{wire}$ in the PTL and VD; which is a fair approximation for a Bhabha.

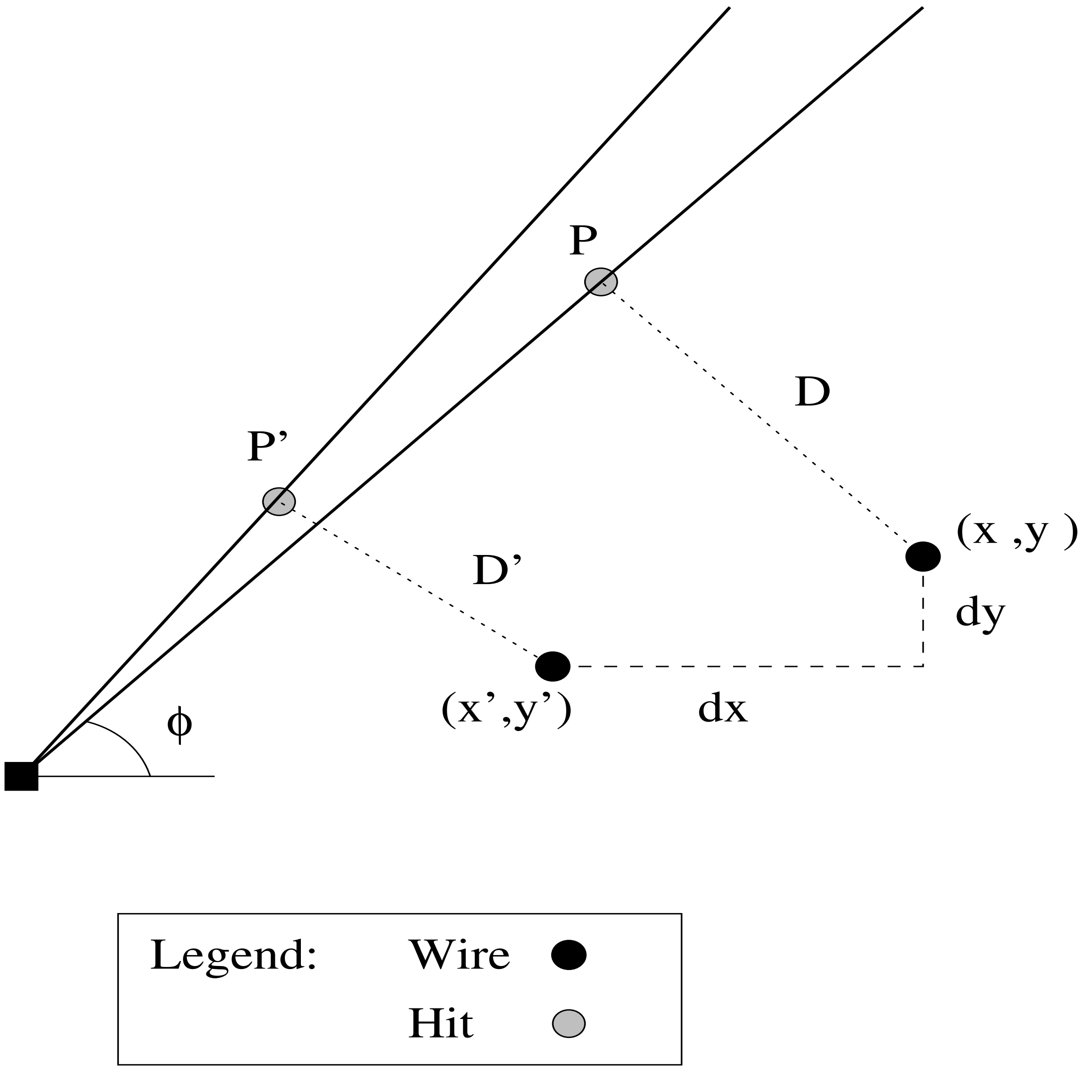


Figure 5.15: Geometry alignment parameters dx and dy are shown. The prime Cartesian coordinate system is the DR (or CLEO II) coordinate system. Then, D' represent signed $\mathrm{DCA}_{1,i}$ and D represent signed $\mathrm{DCA}_{\mathrm{IFIT},i}$.

Then, we have

$$\begin{pmatrix} R\cos\phi' + D'\sin\phi' \\ R\sin\phi' - D'\cos\phi' \end{pmatrix} = \begin{pmatrix} \triangle x + R\cos\phi + D\sin\phi + \omega(R\sin\phi - D\cos\phi) + \delta_y R\cot\theta \\ \triangle y + R\sin\phi - D\cos\phi - \omega(R\cos\phi + D\sin\phi) + \delta_x R\cot\theta \end{pmatrix}. \quad (5.33)$$

Taking

$$\begin{aligned} R\cos\phi' &= R\cos(\phi+\delta\phi) \simeq R(\cos\phi - \delta\phi\sin\phi) \simeq R\cos\phi \\ R\sin\phi' &= R\sin(\phi+\delta\phi) \simeq R(\sin\phi + \delta\phi\cos\phi) \simeq R\sin\phi \end{aligned} \quad (5.34)$$

and dropping the second order terms (ωD), Equation (5.33) becomes[4]

$$\begin{pmatrix} D'\sin\phi' \\ -D'\cos\phi' \end{pmatrix} = \begin{pmatrix} \triangle x + D\sin\phi + \omega R\sin\phi + \delta_y R\cot\theta \\ \triangle y - D\cos\phi - \omega R\cos\phi + \delta_x R\cot\theta \end{pmatrix}. \quad (5.35)$$

Then,

$$D'^2 = (\triangle x + (D+\omega R)\sin\phi + \delta_y R\cot\theta)^2 + (\triangle y - (D+\omega R)\cos\phi + \delta_x R\cot\theta)^2. \quad (5.36)$$

In expanded form, with the second order terms dropped, Equation (5.36) can be written as

$$D'^2 = (D+\omega R)^2\{1 + \frac{2\sin\phi}{(D+\omega R)}(\triangle x + \delta_y R\cot\theta) - \frac{2\cos\phi}{(D+\omega R)}(\triangle y + \delta_x R\cot\theta)\}. \quad (5.37)$$

Since $\sqrt{1+2x} \simeq (1+x)$ for $x \ll 1$, the projected drift distance in the Cartesian coordinate system of reference becomes

$$D' = D + \omega R + (\triangle x\sin\phi - \triangle y\cos\phi) + (\delta_y\sin\phi - \delta_x\cos\phi)R\cot\theta\,. \quad (5.38)$$

Therefore, one can write down the track shift $\Re_{\mathrm{IFIT}} = (D' - D) \equiv A \times (\mathrm{DCA}_1 - \mathrm{DCA}_{\mathrm{IFIT}})$ as a function of the relative position between the two Cartesian coordinate

[4]One must keep in mind that our framework implies a relatively small misalignment.

systems considered[5]

$$\Re_{\mathrm{IFIT}} = \omega R + (\triangle x \sin\phi - \triangle y \cos\phi) + (\delta_y \sin\phi - \delta_x \cos\phi) R \cot\theta. \tag{5.39}$$

The ambiguity ($A = \pm 1$) and the size of D and D' depend on the relative sign and size of $\triangle x$, $\triangle y$, ω, δ_x and δ_y. Thus, the sign of $\Re_{\mathrm{IFIT}}$ depends on the relative sign and size of the elements of the transformation matrix $\mathcal{G}$. The absolute sign relation between $\Re_{\mathrm{IFIT}}$ and the elements of $\mathcal{G}$ given by Equation (5.39) was verified with MC simulation. A given translation or rotation was incorporated and the sign of the geometry alignment parameters were determined.

Assuming that the Bhabha events used for the CD calibration are isotropic[6] in ϕ and θ, it follows that

$$\begin{aligned}
&\sum \cos\phi = N < \cos\phi >= 0\\
&\sum \sin\phi = N < \sin\phi >= 0\\
&\sum \sin\phi\cos\phi = \tfrac{1}{2}\sum \sin 2\phi = \tfrac{N}{2} < \sin 2\phi >= 0\\
&\sum \cos^2\phi \cot\theta = N < \cos^2\phi\cot\theta >= 0\\
&\sum \sin^2\phi \cot\theta = N < \sin^2\phi\cot\theta >= 0\\
&\sum \sin\phi\cos\phi\cot\theta = N < \sin\phi\cos\phi\cot\theta >= 0\\
&\sum \cos^2\phi = N < \cos^2\phi >= \tfrac{N}{2}\\
&\sum \sin^2\phi = N < \sin^2\phi >= \tfrac{N}{2}
\end{aligned} \tag{5.40}$$

with N being the total number of hits. The means in Equation (5.40) are calculated with

$$< f(\phi,\theta) >= \textstyle\int_0^\pi \int_0^{2\pi} f(\phi,\theta)\sin\theta d\theta d\phi \ / \ \int_0^\pi \int_0^{2\pi} \sin\theta d\theta d\phi. \tag{5.41}$$

On taking $\Re_{\mathrm{IFIT}}$ from Equation (5.39) and substituting it into Equation (5.22), and then summing over all hits (i.e. the sum of all hits for a layer followed by the sum on

[5]All the approximations lead to a simple formula that can be related to $\Re_{\mathrm{IFIT}}$. Keeping all the terms in Equation (5.33) would be a more complete method. A full χ^2 minimization of the relative position between the inner chambers and the main drift chamber would be more precise [87]. Equation (5.39) is a first order approximation of the full story.

[6]A Bhabha event is isotropic in ϕ and since we use the `levcut` selection criteria, the Bhabha events selected are also isotropic in θ.

all layers), one finds

$$\begin{aligned}
\sum \mathrm{GTRYY} &= -\triangle y \sum \cos^2\phi = -\tfrac{N}{2}\triangle y \\
\sum \mathrm{GTRXX} &= \quad \triangle x \sum \sin^2\phi = \quad \tfrac{N}{2}\triangle x \\
\sum \mathrm{GRTXY} &= \quad N\omega \\
\sum \mathrm{GTLYZ} &= -\delta_x \sum \cos^2\phi \cos^2\theta / \sin\theta = -\tfrac{N}{4}\delta_x \\
\sum \mathrm{GTLXZ} &= \quad \delta_y \sum \sin^2\phi \cos^2\theta / \sin\theta = \quad \tfrac{N}{4}\delta_y
\end{aligned} \tag{5.42}$$

Therefore, from Equations (5.42), one can fully determine (to first order) the relative position between the main drift chamber and the inner chambers by simply taking the mean of the geometry constants,

$$\begin{aligned}
\overline{\mathrm{GTRYY}} &= -\tfrac{1}{2}\triangle y \\
\overline{\mathrm{GTRXX}} &= \quad \tfrac{1}{2}\triangle x \\
\overline{\mathrm{GRTXY}} &= \quad \omega \\
\overline{\mathrm{GTLYZ}} &= -\tfrac{1}{4}\delta_x \\
\overline{\mathrm{GTLXZ}} &= \quad \tfrac{1}{4}\delta_y
\end{aligned} \tag{5.43}$$

Finally, the geometry alignment parameters for the PTL and the VD are given by:

$$\begin{aligned}
\triangle y &= -2 \times \overline{\mathrm{GTRYY}} \\
\triangle x &= \quad 2 \times \overline{\mathrm{GTRXX}} \\
\omega &= \quad \overline{\mathrm{GRTXY}} \\
\delta_x &= -4 \times \overline{\mathrm{GTLYZ}} \\
\delta_y &= \quad 4 \times \overline{\mathrm{GTLXZ}}
\end{aligned} \tag{5.44}$$

5.4 Constants Extraction

The task of making CLEO II tracking constants consists of three steps:

1. Making the histograms from which constants are extracted.

2. Calculating the constants from these histograms and installing them in libraries.
3. Checking the quality of the constants by monitoring the residuals and by running diagnostics on samples of muon pairs.

A detailed description of the calibration software (for pre-recompress data) is given in Reference [77].

5.5 Constants Quality and Monitoring

The accuracy of individual position measurement in the tracking chambers is defined to be the intrinsic spatial resolution. The position measurement resolution is shown in Figure 5.16 for all the drift chambers. The inverse square of the intrinsic spatial resolution is used as the weight [83] in the fit (see Equation (5.3)). To verify that the D-T functions and time zeros has been determined correctly, a constants monitoring packages called KALI was developed. For more details consult [78].

Muon Pairs Diagnostic

To check the quality of the constants, we run DUET on muon pairs ($e^+e^- \to \mu^+\mu^-$). Testing the constants with muons allows us to test the single track fit quality in a data sample that was not used for calibration. Momentum resolution, miss distance[7], acollinearity[8], and Z0 match[9] for muons are monitored with various diagnostic histograms. The momentum distribution for a sample of muon pairs is shown in Figure 5.17. The measured momentum resolution is around 54 Mev/c depending on the data set and the running conditions. This is slightly higher than the expected $\delta p_\perp =$ 49 Mev/c at $p_\perp = 5.280$ GeV/c from Equation (5.1), which is really only a simplified representation of performance expected for a complex tracking system and does not account for initial state radiation. The average miss distance for muon pairs events is

[7] The miss distance is the separation near the interaction point, in the $r-\phi$ plane, between the positive and negative tracks of a two tracks event.

[8] The acollinearity is the kink at the origin (in radians) between the two tracks in the event. It is defined as $MOD(\phi_1 - \phi_2 + 2\pi, 2\pi) - \pi$.

[9] The Z0 match is the missing Z0 at the point of closest: $\mathrm{Z0}_1 - \mathrm{Z0}_2$.

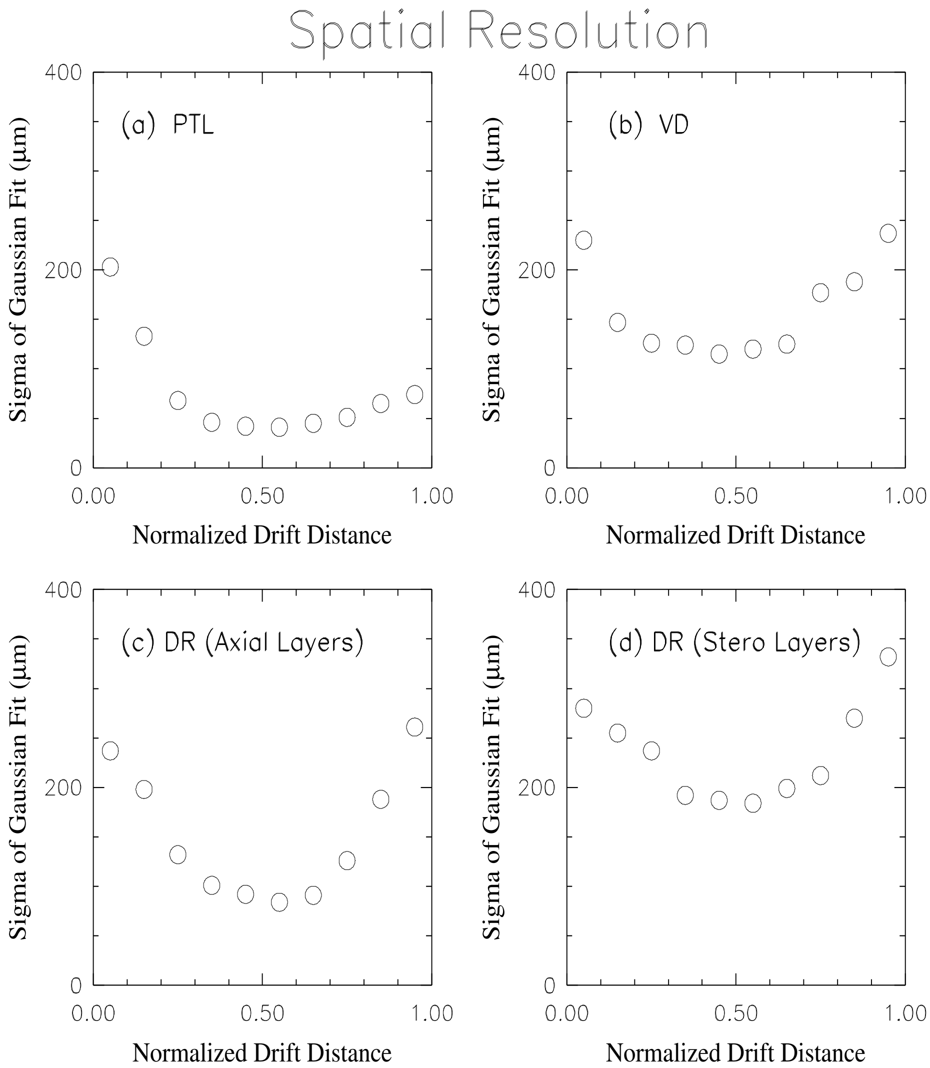


Figure 5.16: Position measurement resolution versus the normalized position within the cell: (a) PTL layers, (b) VD layers, (c) DR axial layers, and (d) DR stereo layers.

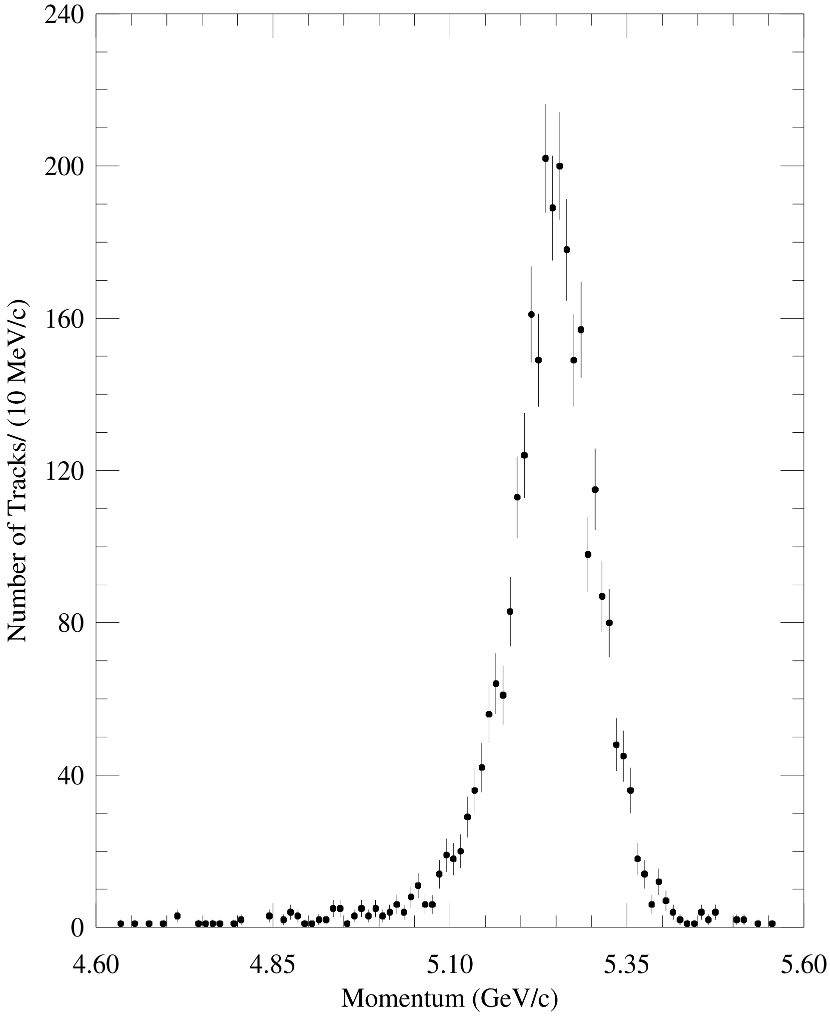


Figure 5.17: Observed momentum of muons in the process $e^+e^- \to \mu^+\mu^-$ at $E_{\text{beam}} =$ 5.280 GeV.

about 95 μm, the resolution on the acollinearity is around 1.0 mrad and the average Z0 match is about 1.6 mm.

Tracking Simulation and MC Tuning

The simulation of the drift chambers in CLEOG is fairly complex and requires a detailed understanding of the physics involved in the ionization process and in the signal pulse readout. The MC parameters are tuned to reproduce the track-finding efficiency and the resolution within the drift cells of a given layer. They are also adjusted to incorporate a good description of the ADC, TDC, pulse height, and time and spatial residual distributions. Currently, the drift chambers MC simulation is in fairly good agreement with the data. It reproduces the behavior of the PTL, VD, and DR in term of hit efficiencies, occupancy, and resolutions for individual layers within $\mathcal{O}(1\%)$ for the anode layers and $\mathcal{O}(3\%)$ for the cathode layers. The CLEOG hit simulation in the PTL, VD, and DR is summarized in great details in Reference [88].

Chapter 6

Event Selection

The decay studied in this thesis is $B^- \to D^{*+}\pi^-\ell^-\bar{\nu}_\ell$, where $D^{*+} \to D^0\pi^+$, and $D^0 \to K^-\pi^+$ or $D^0 \to K^-\pi^+\pi^0$ [19]. We are principally sensitive to the narrow D_J^0 resonances which decay to $D^{*+}\pi^-$. In what follows, we refer to the pion from the D^{*+} as the "slow pion".

Reconstruction of this state requires a knowledge of the 4-momenta of the observable decay products: the pions, a kaon, and a lepton. We make use of photon candidates for the reconstruction of the single $\pi^0 \to \gamma\gamma$ required. The 4-momentum of a neutral pion is determined with the CsI calorimeter position and energy information. Since the π^0 is a short lived meson ($c\tau = 25.1$ nm), one can assume that the massless photons originate from the interaction region and then measure the 3-momentum and the energy of the $\gamma\gamma$ pair. For a charged particle, the 3-momentum is obtained by measuring the curvature of its path in the magnetic field. Further measurements are needed to determine the energy of a charged particle, and therefore fix a value for its mass. Since the mass of a charged particle uniquely defines its identity, such measurements are generally referred to as "particle identification". This chapter focuses on the aspects of particle detection, identification, and reconstruction relevant to the channel under study.

6.1 Data Sample

B mesons are produced copiously in electron-positron collisions at CESR. The data used in this thesis were collected between November 1990 and April 1995 by the CLEO II detector. The data consists of a sample with an integrated luminosity of 3.11 fb^{-1} on the $\Upsilon(4S)$ resonance (ON Resonance), corresponding to 3.29×10^6 $B\bar{B}$ events, and a further 1.61 fb^{-1} at a center-of-mass energy ~ 55 MeV below the $\Upsilon(4S)$ resonance (OFF Resonance). These represent almost a tenfold increase in luminosity over those used in similar studies of B meson decays. The data are divided into fifteen different subsets labeled from 4S2 through 4SG, as summarized in Table 6.1. All changes in detector or running condition from one data set to another have been taken into account in the Monte Carlo simulation of the CLEO II detector used in the analysis. Overall, the operation and the calibration of the CLEO II detector have maintained good stability over time [75].

6.2 Monte Carlo Samples

Monte Carlo (MC) simulations are needed to study the optimization of the statistical significance of the signal relative to the backgrounds and to compute the reconstruction efficiencies of our analysis technique. To model the signal events and the various backgrounds, several MC samples were generated. The first sample consists of 21.6×10^6 generic $B\bar{B}$ events and the second is a sample of 5.2×10^6 continuum events. Many signal MC data sets were also generated to investigate each decay mode.

The generic $B\bar{B}$ MC sample corresponds to roughly 6.4 times the size of the actual sample of $B\bar{B}$ events in data. The generic decay of the B meson is handled by a decay table which contains all the measured and expected branching fractions of all the exclusive hadronic, leptonic, and semileptonic decay modes of the B meson [54]. The simulation of the semileptonic decays of the B meson relies on the ISGW2 [51] and the G&R hybrid [89] models, as described in Chapter 3. The MC generator EvT takes into account the angular correlation among the decay products, which provides an accurate description of the decay dynamics of the semileptonic decay of

Data Set	Date of Data Collection	Luminosity (pb^{-1}) ON $\Upsilon(4S)$	Luminosity (pb^{-1}) OFF $\Upsilon(4S)$
4S2	Nov. 90 - Jun. 91	462	197
4S3	Sep. 91 - Feb. 92	436	209
4S4	Apr. 92 - May. 92	214	101
4S5	Jul. 92 - Oct. 92	216	105
4S6	Nov. 92 - Jan. 93	232	85
4S7	Mar. 93 - Jul. 93	285	177
4S8	Aug. 93 - Sep. 93	188	94
4S9	Nov. 93 - Jan. 94	230	117
4SA	Jan. 94 - Feb. 94	138	54
4SB	Mar. 94 - May. 94	85	64
4SC	Jun. 94 - Aug. 94	115	36
4SD	Sep. 94 - Oct. 94	53	50
4SE	Oct. 94 - Nov. 94	71	62
4SF	Nov. 94 - Nov. 94	89	66
4SG	Jan. 95 - Apr. 95	293	192
Total Luminosity		3107	1609

Table 6.1: Data sets summary.

a B meson. The B semileptonic branching fractions used in the generic $B\bar{B}$ MC are listed in Table 6.2. The branching fractions for $\bar{B} \to D^{(*)}\ell\bar{\nu}_\ell$ come from the average of experimental measurements by CLEO II (see Figure 2.10). The branching fraction for $\bar{B} \to D_1\ell\bar{\nu}_\ell$ is the value measured by ALEPH [29]. The contribution of $\bar{B} \to D_2^*\ell\bar{\nu}_\ell$ is chosen to be consistent with the ALEPH measurement [90]. Based on measurement of inclusive production of leptons in B semileptonic decays, we expect the presence of other exclusive decay modes. Pioneering measurements of $\bar{B} \to D^{**}\ell\bar{\nu}_\ell$ by ARGUS [27] and CLEO [28] indicate the possible presence of resonant and nonresonant contributions from $D\pi\ell\bar{\nu}_\ell$ and $D^*\pi\ell\bar{\nu}_\ell$. More recent measurements from the LEP experiments [29, 30, 31] confirm the presence of $D\pi$ and $D^*\pi$ states in B

State	Decay Mode	Assumed $\mathcal{B}$ (%)
$1\,^1S_0$	$\bar{B} \to D\ell\bar{\nu}_\ell$	1.80
$1\,^3S_1$	$\bar{B} \to D^*\ell\bar{\nu}_\ell$	4.90
$1\,^1P_1$	$\bar{B} \to D_1\ell\bar{\nu}_\ell$	0.74
$1\,^3P_2$	$\bar{B} \to D_2^*\ell\bar{\nu}_\ell$	0.43
$1\,^3P_1$	$\bar{B} \to D_0^*\ell\bar{\nu}_\ell$	0.26
$1\,^3P_0$	$\bar{B} \to D_1^*\ell\bar{\nu}_\ell$	0.26
$2\,^1S_0$	$\bar{B} \to D'\ell\bar{\nu}_\ell$	0.26
$2\,^3S_1$	$\bar{B} \to D^{*\prime}\ell\bar{\nu}_\ell$	0.53
NR	$\bar{B} \to D\pi\ell\bar{\nu}_\ell$	0.76
NR	$\bar{B} \to D^*\pi\ell\bar{\nu}_\ell$	0.24

Table 6.2: Assumed branching fractions for the exclusive semileptonic decays of the B meson in generic $B\bar{B}$ MC. The nonresonant contribution states are label by NR.

semileptonic decays. The contributions from (D_0^*, D_1^*), $(D', D^{*\prime})$, and the nonresonant (NR) states in our generic $B\bar{B}$ MC were set to saturate the inclusive rate so that $\mathcal{B}_{\rm SL} = 10.18\%$. The relative amount of each state was constrained to be consistent with the existing CLEO measurement of the inclusive lepton energy spectrum for $b \to c\ell\bar{\nu}_\ell$ [24]. The branching fractions for the cascade decays $D_J \to D^{(*)}\pi$, $D^{(*)\prime} \to D^{(*)}\pi$, and $D^{*\prime} \to D_1^*\pi$ were determined from isospin symmetry and estimations of the decay matrix elements [91]. They are summarized in Table 6.3. The D^* and D branching fractions are taken from the Particle Data Group compilation [6].

The second MC sample was generated to check the importance of non-$B\bar{B}$ background. It turns out that the contamination of our signal sample by continuum events is fairly small. A reliable way to deal with the continuum background is to perform a subtraction, as described in Section 7.3.2.

Our MC signal samples contain 70×10^3 $B^- \to D_1^0\ell^-\bar{\nu}_\ell$ events and 70×10^3 $B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell$ events for the $D^0 \to K^-\pi^+$ mode, and 230×10^3 $B^- \to D_1^0\ell^-\bar{\nu}_\ell$ events and 230×10^3 $B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell$ events for the $D^0 \to K^-\pi^+\pi^0$ mode. These independent samples are used to compute the signal detection efficiencies.

Decay Mode	Assumed $\mathcal{B}$ (%)	Decay Mode	Assumed $\mathcal{B}$ (%)
$D_1 \to D^*\pi^\pm$	67.0	$D' \to D^*\pi^\pm$	65.9
$D_1 \to D^*\pi^0$	33.0	$D' \to D^*\pi^0$	33.7
$D_2^* \to D^*\pi^\pm$	20.9	$D' \to D^*\gamma$	0.4
$D_2^* \to D^*\pi^0$	10.3	$D^{*\prime} \to D^*\pi^\pm$	24.3
$D_2^* \to D\pi^\pm$	45.9	$D^{*\prime} \to D^*\pi^0$	12.2
$D_2^* \to D\pi^0$	22.9	$D^{*\prime} \to D\pi^\pm$	42.1
$D_0^0 \to D\pi^\pm$	67.0	$D^{*\prime} \to D\pi^0$	20.2
$D_0^0 \to D\pi^0$	33.0	$D^{*\prime} \to D^*\gamma$	0.1
$D_1^* \to D^*\pi^\pm$	67.0	$D^{*\prime} \to D\gamma$	0.2
$D_1^* \to D^*\pi^0$	33.0	$D^{*\prime} \to D_1^*\pi$	0.9

Table 6.3: Branching fractions for the various D^{**} decay modes used in our generic $B\bar{B}$ MC.

6.3 Selection Criteria Optimization

In the investigation of $D^*\pi$ production in B semileptonic decays, one wants to choose selection criteria that are as efficient as possible for our signal, while retaining good rejection power for the various backgrounds. In this analysis, we mainly use our $B\bar{B}$ and continuum MC samples to optimize the statistical significance of the signal observation. Some of the cuts used are based solely on expectations for the physics of a semileptonic B decay, and others rely on the calculation of a figure of merit (F) which maximizes signal over background. We define

$$F = \frac{\mathtt{S}^2}{\mathtt{S}+\mathtt{B}}\,, \tag{6.1}$$

where $\mathtt{S}$ is the number of reconstructed signal events and $\mathtt{B}$ is the number of non-signal events which pass the selection cut(s) under study.

When a specific background process is not modeled in the MC, wrong-sign and sideband samples turn out to be reliable tools for modeling combinatorial backgrounds and optimizing their rejection. No optimization was ever performed on data that could contain real signal events.

After a careful study of the $B\bar{B}$ and continuum MC events, the background in this analysis can be divided into several well-defined components:

1. Our main background, both in the $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^0$ channels, is the combination of a real D^{*+} and a random pion. Since $\mathcal{B}(\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell)$ is expected to be an order of magnitude larger than $\mathcal{B}(B^- \to D_1^0\ell^-\bar{\nu}_\ell)$ and $\mathcal{B}(B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell)$, we need to pay special attention to this background.

2. The second background that we must consider in our analysis is the continuum background (or non-$B\bar{B}$ background). This background is modeled by measuring the signal yield using OFF Resonance data.

3. Another background arises from fake leptons. The fake lepton background is the contribution in which a $D^{*+}\pi^-$ is paired with a hadron misidentified as a lepton. This contribution is estimated by performing an analysis where non-leptons are treated as leptons and the result re-normalized using known estimates of the fake rates.

4. Uncorrelated background (background from events in which the $D^{*+}\pi^-$ comes from the $\bar{B}$ and the lepton from the B) can also contaminate our signal. In such cases, the lepton comes from a cascade decay $\bar{b} \to \bar{c} \to \ell^-$ from the second B meson in the event.

5. Correlated background (background from events in which $\bar{B} \to D^{*+}\pi^-X\ell^-\bar{\nu}_\ell$ and the other B decays generically) produces a real $D^{*+}\pi^-\,\ell^-$ combination and can therefore mimic our signal.

6. Real lepton background from $\tau^- \to \ell^-\bar{\nu}_\ell\nu_\tau$ decays ($\ell = e$ or μ). We found no contribution to our yields from this background.

6.4 Global Event Shape Criteria

There are two global event shape criteria used in our analysis. They are both intended to select $B\bar{B}$ events. The first is an event class cut (called KLASGL) and the second is an event shape cut (called R_2 or R2GL).

An event is classified as a possible hadronic final state if the following requirements are met.

- The event must contain a minimum of three charged tracks.
- The total visible energy in the event must be greater than 15% of the total center-of-mass energy.
- The energy observed in the calorimeter has to be between 15% and 90% of the total center-of-mass energy.
- The location of the primary vertex for the event must be within $\pm$ 2 cm and $\pm$ 5 cm of the beam spot in the $r-\phi$ plane and z-direction respectively.

To further reduce non-$B\bar{B}$ background, each event is required to satisfy the ratio of Fox-Wolfram [92] moments $R_2 < 0.4$. R_2 is a measure of the isotropy of the momentum distribution. The smaller the value of R_2, the more isotropic the event. The R_2 parameter is then very useful for distinguishing $B\bar{B}$ events, which tend to be isotropic, from continuum events, which tend to be more jet-like. The distributions of R_2 for $B\bar{B}$ and continuum events are shown in Figure 6.1.

6.5 Track Selection

Charged particle detection is crucial in the present analysis. In Appendix B, we give a description of the CLEO variables used for track selection. All charged tracks (with the exception of the slow pion from the D^*) must meet the following criteria.

- The track must be in the fiducial volume of the drift chambers: $|\cos\theta| < 0.92$. The angle θ is the angle of the track with respect to the beam line.
- The track must originate from the vicinity of the e^+e^- interaction point. We require: DBCD $<$ 5 mm and Z0CD $<$ 5 cm and KINCD $=$ 0. The impact parameters DBCD and Z0CD are measured in the $r-\phi$ plane and along the z-direction respectively. The vertex flag KINCD $=$ 0 selects tracks from the primary vertex.

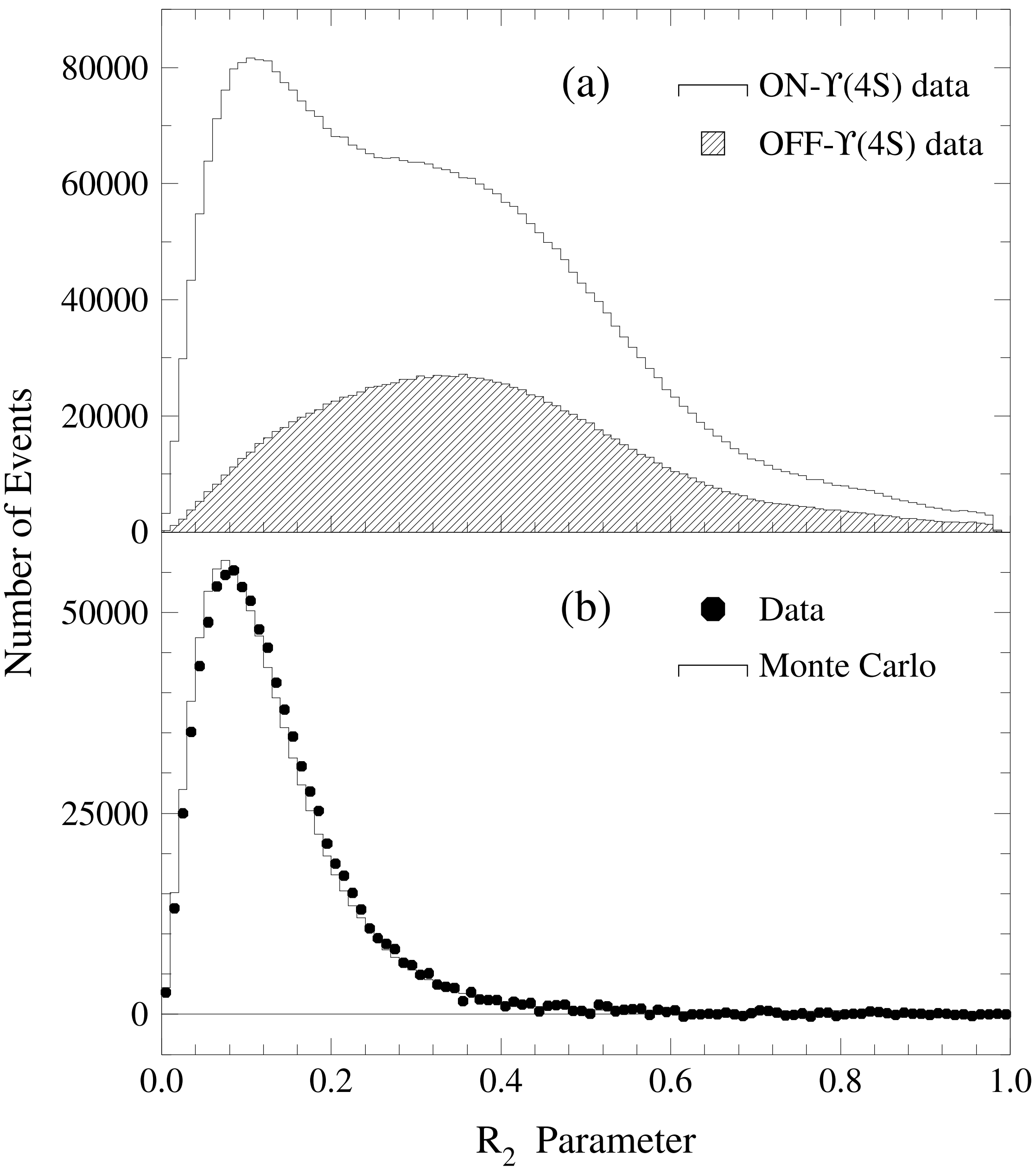


Figure 6.1: The R_2 distribution in data and MC simulation: (a) R_2 distributions derived from ON Resonance data (unshaded) and OFF Resonance data (shaded). (b) R_2 distribution of $\Upsilon(4S) \to B\bar{B}$ decays derived after scaled continuum subtraction (data points). The superimposed histogram shows the same distribution derived from generic $B\bar{B}$ MC simulation. We require $R_2 < 0.4$. Source [93].

- The track must pass the TRKMAN requirements [94]. The software package TRKMAN eliminates spurious ghost pairs, curlers, backsplash, and scattered tracks[1].

- The track must have good dE/dx information.

The slow charged pion from the D^{*+} (labeled π^+_{slow}) is treated differently since its momentum is about 40 MeV/c in the D^{*+} rest frame. A low momentum bound of 65 MeV/c and a polar angle limit of $|\cos\theta| < 0.71$ are imposed on π^+_{slow} to assure a reliable reconstruction efficiency [95]. Kinematic constraints for our signal also put an upper bound of 250 MeV/c on the momentum of the slow pion. We therefore limit our slow pion candidates to the momentum range of 65 MeV/c to 250 MeV/c. No dE/dx information is required. We further require the slow pion to pass the TRKMAN requirements.

Another basic requirement for the slow pion candidates is that they must originate from the primary vertex, and thus have KINCD = 0. In the remainder of this section, we describe a special vertex cut which was developed to increase the reconstruction efficiency of the slow pion.

Based on work presented in Reference [93], we find that the DBCD distribution is shifted from zero for tracks with momenta below 250 MeV/c. The mean of the DBCD distribution as a function of the particle momentum is displayed in Figure 6.2 for a sample of charged slow pions from signal MC events for $\bar{B} \to D_1 \ell \bar{\nu}_\ell$, where $D_1 \to D^*\pi$ followed by $D^* \to D\pi^\pm_{\text{slow}}$. It is important to note that the shift of DBCD is of opposite sign for positively and negatively charged tracks. The observed shifts are due to an underestimation of the scattering material at the PTL-VD and the VD-DR walls, which causes a mismeasurement of the track parameter ϕ for low

[1] A ghost pair is two tracks fitted to the same set of hits. A track with insufficient momentum to reach the outer edge of the main drift chamber may spiral many times in the tracking chambers. Multiple tracks formed from the spirals are called curlers. A track with enough momentum can exit the main drift chamber, enter the calorimeter, lose energy and reenter the drift chambers. Such tracks are called backsplash. Occasionally, a particle will scatter in the material of the detector and kink. Sometimes it will interact with the material and might create many other charged particles. Some other times it may simply decay in flight. In each of these three later cases, one or several tracks may intersect the point of scatter or decay: such tracks are classified as scattered tracks. The role of TRKMAN is to map a set of hits to only one track.

momentum tracks.

The observed values in MC of the mean of DBCD ($\overline{\mathrm{DBCD}}$) and the width of DBCD (σ_{DBCD}), averaged over positively and negatively charged tracks, are fit to the functional form:

$$\overline{\mathrm{DBCD}}(p_\pi) = Q \times a(1 + bp_\pi + cp_\pi^2)e^{-dp_\pi}, \tag{6.2}$$

$$\sigma_{\mathrm{DBCD}}(p_\pi) = a(1 + bp_\pi + cp_\pi^2)e^{-dp_\pi}, \tag{6.3}$$

with $p_\pi \equiv |\mathbf{p}_\pi|$ being the track momentum measured in GeV/c and Q being the charge of the track. The parameterization of the mean of DBCD and the error on DBCD versus the momentum of the slow pion is shown in Figure 6.2(c). The fit parameters are listed in Table 6.4. The functional form of $\overline{\mathrm{DBCD}}(p_\pi)$ and $\sigma_{\mathrm{DBCD}}(p_\pi)$ shown in Figure 6.2 for charged slow pions from $D^* \to D\pi^\pm_{\mathrm{slow}}$ in MC is consistent with the functional form obtained with charged slow pions from $D^* \to D\pi^\pm_{\mathrm{slow}}$ in data [93]. We define:

$$\chi^2_{\mathrm{DBCD}}(p_\pi) = \frac{[\mathrm{DBCD} - \overline{\mathrm{DBCD}}(p_\pi)]^2}{\sigma^2_{\mathrm{DBCD}}(p_\pi)}, \tag{6.4}$$

and select only slow pion candidates with $\chi^2_{\mathrm{DBCD}} < 9$ (*i.e.*, a 3σ cut). Similarly, we define:

$$\chi^2_{\mathrm{Z0CD}} = \frac{[\mathrm{Z0CD} - \overline{\mathrm{Z0CD}}]^2}{\sigma^2_{\mathrm{Z0CD}}}, \tag{6.5}$$

In contrast to $\overline{\mathrm{DBCD}}$ and σ_{DBCD}, the mean and width of Z0CD do not depend on the momentum or the charge of the slow pion, and are fixed to $\overline{\mathrm{Z0CD}} = 1.2$ mm and $\sigma_{\mathrm{Z0CD}} = 14.2$ mm. Again, those values were determined by MC simulation. We accept candidates with $\chi^2_{\mathrm{Z0CD}} < 9$. The introduction of the vertex cut described above improves $\mathtt{S}^2/(\mathtt{S}+\mathtt{B})$ by 17% when we run on generic $B\bar{B}$ MC events, mainly by removing combinations of real D mesons and spurious slow pions.

6.6 Lepton Identification

The identification of leptons is essential to our analysis. At the $\Upsilon(4S)$, the detection of a fast lepton strongly suggests the presence of a B semileptonic decay. Electrons

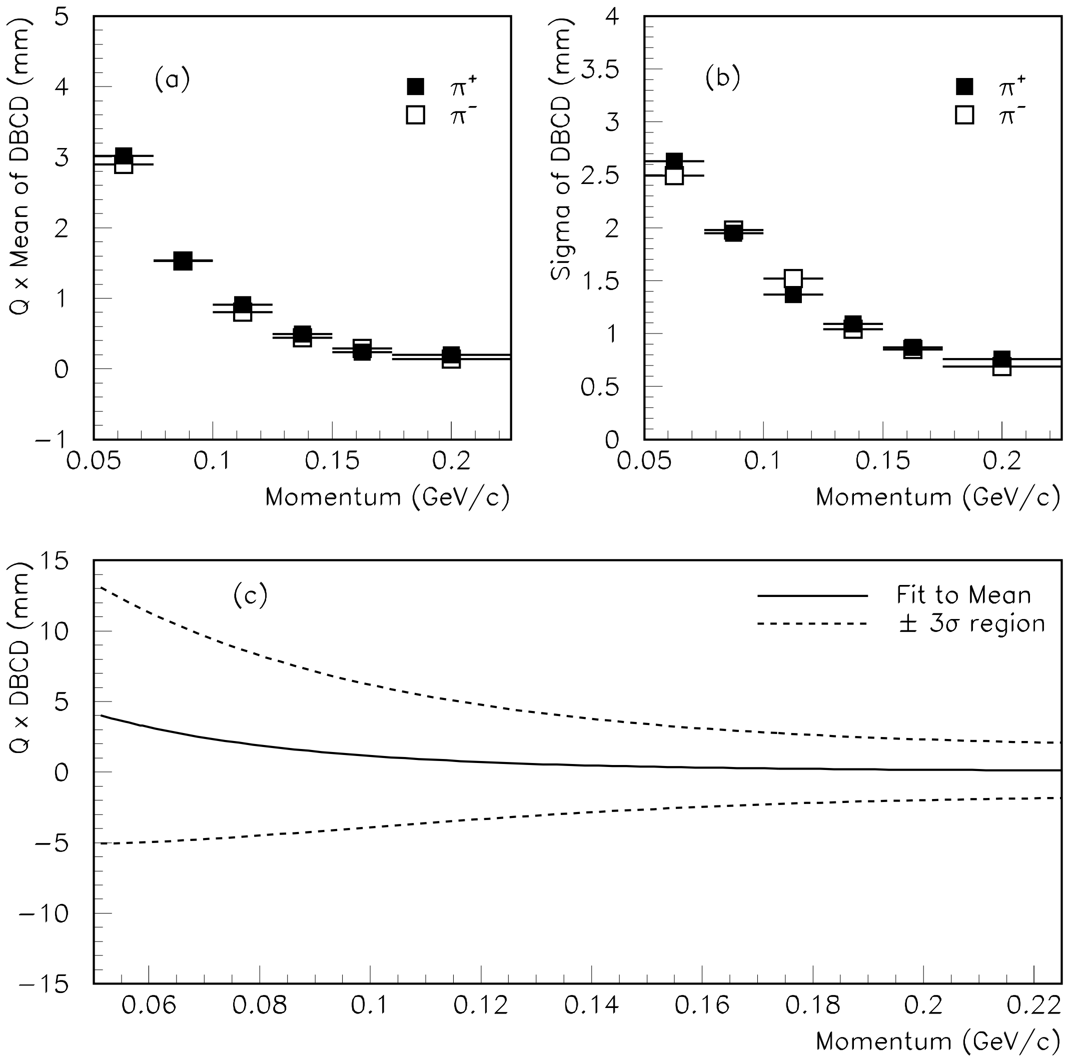


Figure 6.2: (a) The mean of DBCD times the charge of the track and (b) standard deviation of DBCD as a function of the momentum of the slow pion. The sample of slow pions comes from $\bar{B} \to D_1 \ell \bar{\nu}_\ell$ MC events when $D_1 \to D^* \pi$ and $D^* \to D\pi_{\text{slow}}$. In (c), the solid line traces the mean of DBCD times the charge of the track (averaged over π^+ and π^-) while the dashed lines indicate the $\pm\ 3\sigma$ boundary.

Parameter	Mean of DBCD	Width of DBCD
a	15.99	5.716
b	-7.51	-6.35
c	25.34	17.17
d	19.51	6.008

Table 6.4: Fit parameters for the mean and width of DBCD distributions as a function of track momentum as derived by a MC simulation [93].

and muons produce very distinctive signatures in the CLEO II detector through their characteristic interactions with matter. The electrons and the muons leave tracks in the drift chambers and their charges and momenta are calculated from the curvature of these tracks. The electrons deposit essentially all their energy in the CsI calorimeter while the muons leave trails in the muon chambers.

6.6.1 Electron Identification

Electron identification relies primarily upon several independent measured quantities [96]:

E*/*p The most sensitive variable for identifying electrons is the ratio of the energy (E) deposited in the calorimeter to the momentum ($p \equiv |\mathbf{p}|$) of the track pointing to the cluster. The quantity E/p is close to one for electrons, and smaller for all other charged particles. The discrimination of electrons from hadrons or muons from the ratio E/p is illustrated in Figure 6.3.

dE*/*dx The specific ionization (dE/dx) measured in the drift chambers is also a powerful piece of information for identifying electrons (see Figure 6.5). The difference between the measured and the predicted ionization loss for an electron peaks at zero, whereas the hadron response is shifted lower by about two standard deviations.

Track match Another quantity useful in electron identification is the distance between the projection of the track and the calorimeter shower. A matching requirement between the track and the shower provides good discrimination between electrons and other neutral and charged particles.

Cluster shape The last quantity used for electron detection is the shape of the shower. Electromagnetic showers tend to deposit all their energy in a few crystals very close to the center of the cluster. We use variables which measure the lateral development of the shower to distinguish electrons from hadrons.

For studying efficiencies and rejections rates, distributions for each of these variables are made for electrons and non-electrons separately. The electron sample comes from embedding radiative Bhabha events into hadronic events where the event topology is close to that for an electron from B semileptonic decays. The non-electron sample comes from $\Upsilon(1S)$ hadronic events which are known to have very few leptons in them.

For each charged track the probabilities of being an electron (P_e) and non-electron ($P_{\not e}$) are calculated for all variables. We then combine this information by computing a log-likelihood ratio defined as:

$$\mathcal{L}_e = \sum_{\text{variables}} \ln\left(\frac{P_e}{P_{\not e}}\right). \tag{6.6}$$

For a track to be identified as an electron we require $\mathcal{L}_e$ to be greater than 3.0 and $|\cos\theta| < 0.92$. We demand the electron momentum to be between 0.8 GeV/c and 2.0 GeV/c. The lower bound is set to minimize the contribution from secondary leptons from charm decays (*c.f.,* Figure 3.2). The upper bound is just below the kinematic limit for $B^- \to D_1^0\ell^-\bar{\nu}_\ell$ and $B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell$ (see Figure 3.6).

The electron detection efficiency is about 94% in the momentum range of 0.8 GeV/$c \le |\mathbf{p}| < 2.0$ GeV/c [97]. Electron efficiencies are obtained from embedding Bhabha events in non-leptonic $\Upsilon(4S)$ events. The probabilities of misidentifying a hadron as an electron (called fake rates) are obtained from $\Upsilon(1S)$ hadronic events. The typical fake rate is found to be 0.1% to 0.2% per track [98].

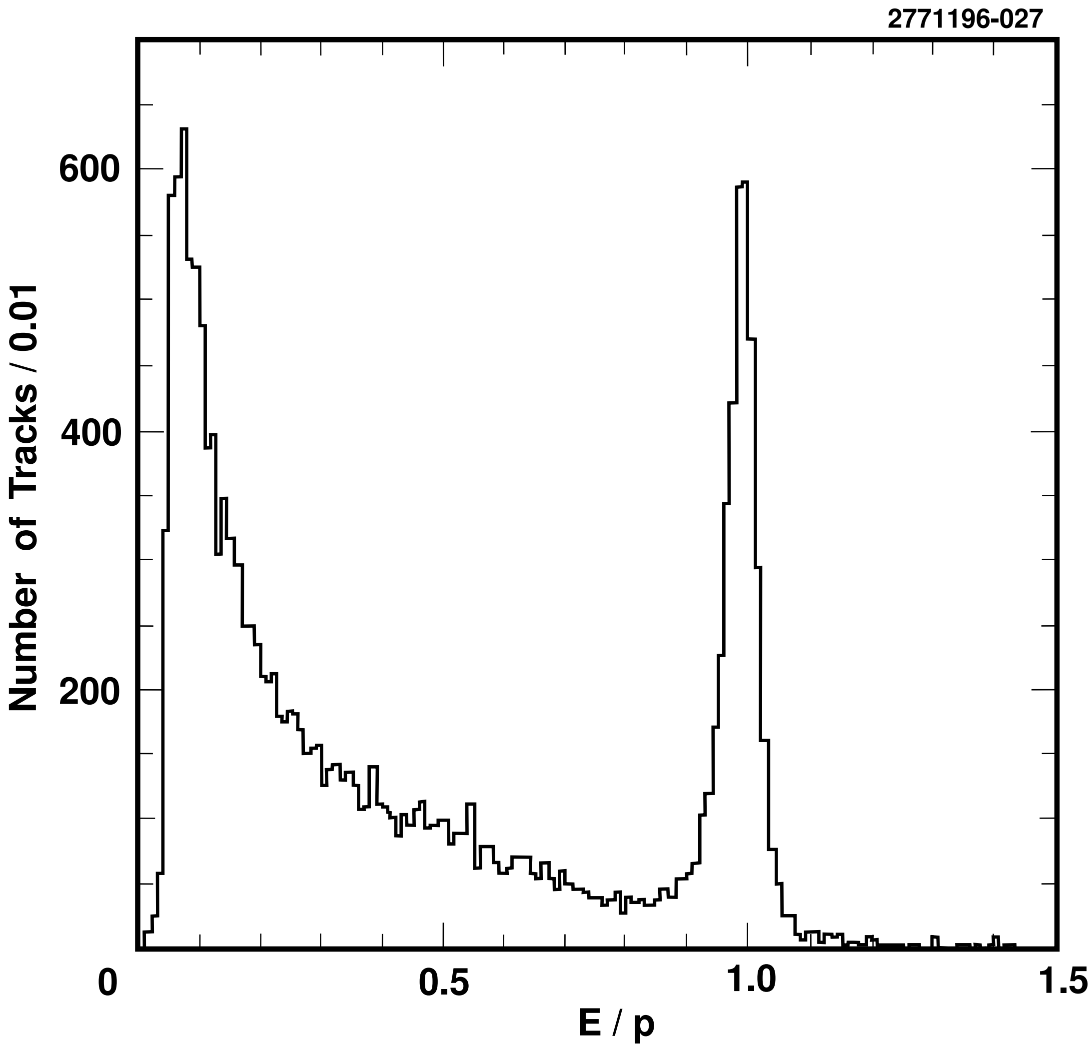


Figure 6.3: The ratio of the EM cluster energy to the momentum of the track pointing to the cluster. The peak at $E/p = 1$ is due to electron and the tail for $E/p < 1$ is due to hadrons and muons.

6.6.2 Muon Identification

Muon identification relies upon penetration through the layers of iron absorber to the various levels of the muon chambers. Each superlayer of the muon chambers is preceded by approximately two nuclear absorption lengths of iron. This arrangement leads to a muon-penetration threshold of 1.0 GeV/c to the first superlayer, 1.5 GeV/c to the second superlayer, and 2.0 GeV/c to the third superlayer. We used the MUTR package for the basic muon identification. MUTR provides us with DPTHMU and MUQUAL. DPTHMU is the depth that the muon traveled, *i.e.*, the number of nuclear absorption lengths (λ) that the muon traveled. MUQUAL is a track quality flag which correlates hit patterns in the muon detector with the projected trajectories of the particles found in the central tracking detector. In the track matching algorithm, multiple scattering in the calorimeter and the iron, and the deflection caused by the magnetic field in the flux return are taken into account.

In this analysis, muons are required to have a good match (MUQUAL=0) between hits registered in the muon chambers and the extrapolated drift chamber track. Furthermore, muon candidates must satisfy the following acceptance cuts:

- For 1.0 GeV/$c \leq |\mathbf{p}| < 1.5$ GeV/c: $|\cos\theta| < 0.85$ and $3 \leq$ DPTHMU < 5
- For 1.5 GeV/$c \leq |\mathbf{p}| \leq 2.0$ GeV/c: $|\cos\theta| < 0.82$ and DPTHMU ≥ 5

where θ is the angle of the muon candidate with respect to the beam axis. The angular and momentum coverage are constrained by the acceptance of the muon counters [68]. The high energy cutoff at 2.0 GeV/c is again set by the kinematics of the decay.

The muon detection efficiency for the different levels of the muon chambers, as a function of the muon momentum, is shown in Figure 6.4. The probabilities of misidentifying a hadron as a muon are much higher than the probabilities of misidentifying a hadron as an electron. The individual fake probabilities are determined by running the muon identification package on hadronic tracks. A detailed study of muon fake rates is presented in Appendix E. The hadron fake probabilities are shown in Figures E.3, E.4, and E.5.

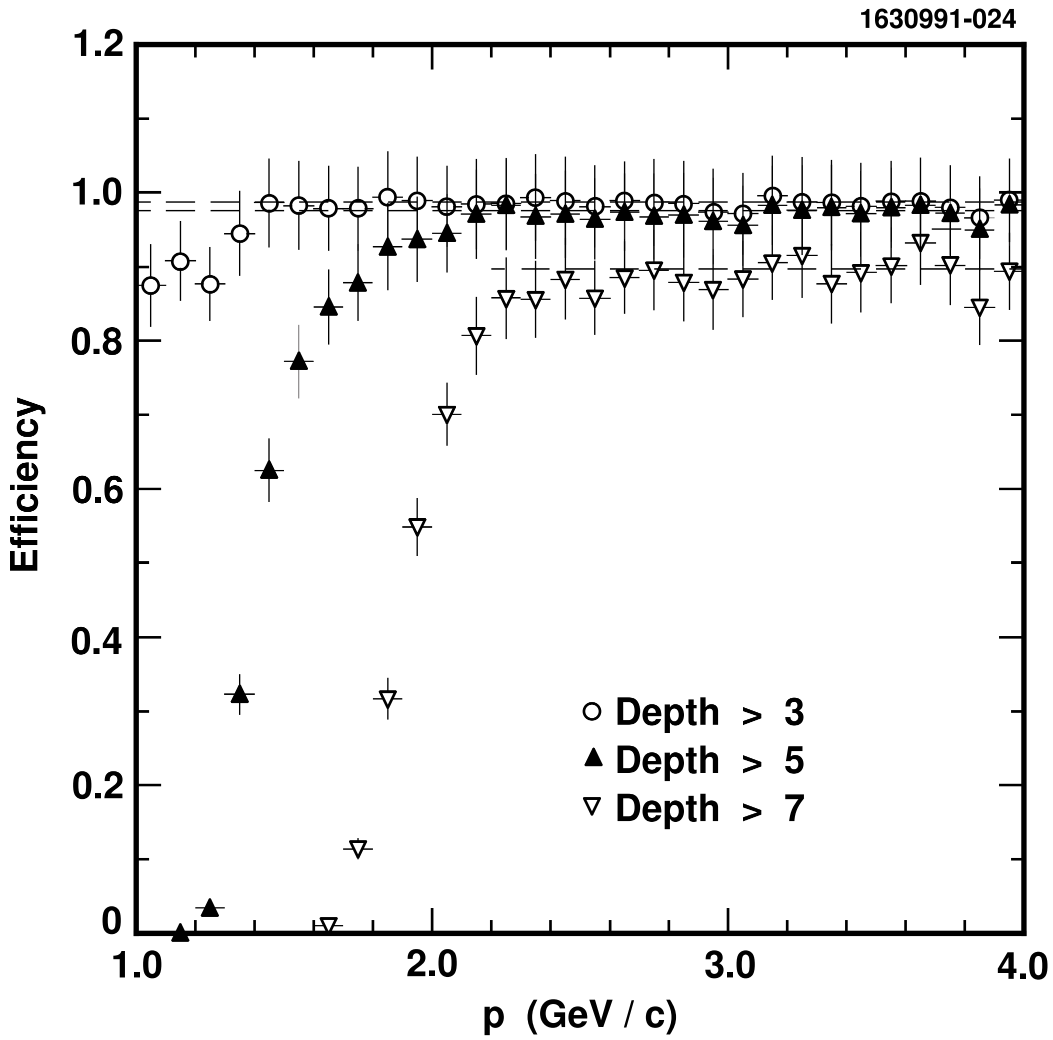


Figure 6.4: Muon identification efficiency versus momentum for three different depth requirements [65]: DPTHMU > 3, DPTHMU > 5, and DPTHMU > 7.

6.7 Charged Hadron Identification

At CLEO II, hadron identification has traditionally been based only on dE/dx measurements in the main drift chamber. The TOF counters provide particle identification information as well, but the agreement between data and MC simulation of the TOF counter for hadrons has never been entirely reliable [99]. The present analysis relies heavily on the MC reconstruction efficiency for our signal. The use of the TOF would introduce a larger systematic uncertainty on the hadron identification and it would only provide a marginal reduction of the combinatoric backgrounds. Therefore, we do not use the TOF identification capability in the hadron selection for $B^- \to D^{*+}\pi^-\ell^-\bar{\nu}_\ell$.

A relativistic particle passing through argon-ethane (50:50) at atmospheric pressure makes a collision about every 200 μm and transfers energy to the gas via ionization. In practice, we measure the amount of charge (the pulse height) collected on every wire and normalize it to the estimated track-length in the cell. Then, the specific ionization of a hadron candidate is determined by computing the truncated mean of the normalized pulse height over the ensemble of cells associated with the track. Using the truncated mean eliminates the long tail to high depositions created by the Landau distribution and leads to a more Gaussian behavior for the mean dE/dx result. For each measurement of the energy-loss, we require the hadron candidate to pass through at least four drift cells. We demand dE/dx information for all the hadrons with the exception of the slow pion.

The specific ionization depends on the speed or $\beta\gamma$ of a relativistic particle [100]. Since the momentum of the particle is related to its mass by $|\mathbf{p}| = \beta\gamma\, m$, one can parameterize the energy-loss of the particle versus its momentum and then determine its mass. Figure 6.5 shows dE/dx as a function of momentum for different types of particles at CLEO II. As one can see, the dE/dx measurements yield good separation of kaons from pions up to momenta of roughly 700 MeV/c. These particle identification capabilities are useful for the reduction of the $\pi - K$ combinatoric backgrounds.

Finally, for the purpose of identifying charged pions and kaons, we use the

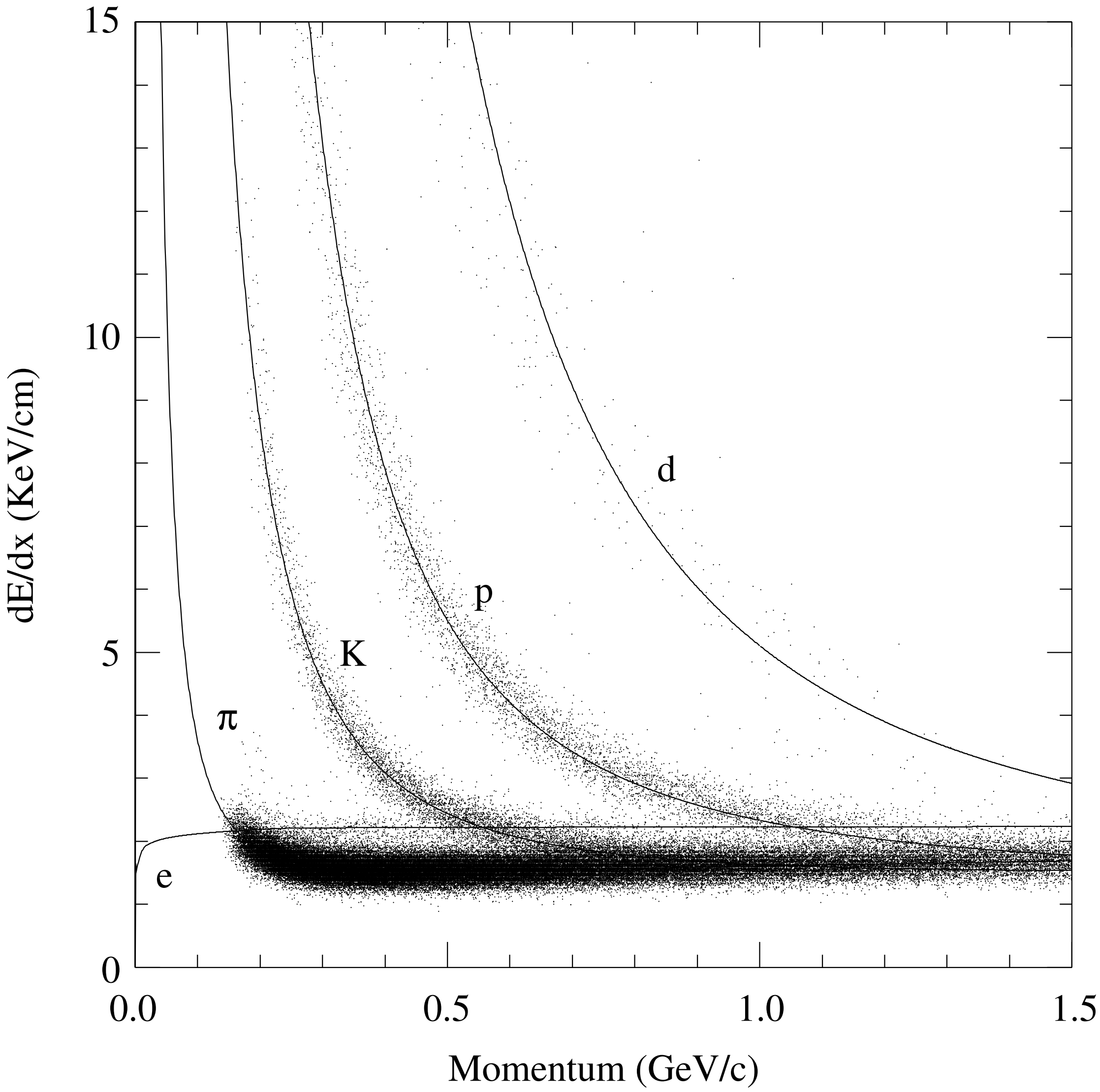


Figure 6.5: Specific ionization curves versus momentum for various species of hadron. One can identify bands corresponding to electrons, pions, kaons, protons and deuterons. The latter are produced predominantly through beam-wall interactions.

mean dE/dx as follows. We compute, for $i = \pi$ or K,

$$(\chi_i^{dE/dx})^2 = \frac{\left[\left(\frac{dE}{dx}\right)_{\text{Measured}} - \left(\frac{dE}{dx}\right)_{\text{Expected}_i}\right]^2}{\sigma^2_{\text{Expected}_i}}. \tag{6.7}$$

Among primary tracks, charged pions are the most abundant particle species produced in $B\bar{B}$ events. One therefore needs to apply more stringent particle identification criteria when selecting kaon candidates to further reduce $\pi - K$ misidentification. Hence, charged pion and kaon candidates, with the exception of the slow pion from the decay of the D^{*+}, are required to have ionization losses in the drift chamber within 3.0 and 2.5 standard deviations (σ), respectively, of those expected for the hypothesis under consideration. This corresponds to $(\chi_\pi^{dE/dx})^2 < 9.0$ and $(\chi_K^{dE/dx})^2 < 6.25$.

6.8 Neutral Pion Reconstruction

The neutral pion reconstruction at CLEO II relies on the hermiticity and the excellent photon detection capability of the CsI calorimeter. Not only is the energy resolution of the calorimeter exceptional, but its fine granularity permits very good position resolution as well. This last feature is critically important in reconstructing $\pi^0 \to \gamma\gamma$. In the π^0 reconstruction algorithm, photon candidates must not be matched to charged tracks projected from the drift chambers and are required to have a cluster shape consistent with that expected for photons. The π^0s used for the reconstruction of $D^0 \to K^-\pi^+\pi^0$ are required to have two individual showers, corresponding to photons with energies of at least 50 MeV. At least one of the two photons is required to be in the good barrel region, where θ_γ, the polar angle of the photon with respect to the beam line, satisfies $|\cos\theta_\gamma| < 0.71$.

All π^0 candidates must have an unconstrained invariant mass $M(\gamma\gamma)$ [101] within 2 standard deviations of the nominal π^0 mass ($\sigma_{\gamma\gamma} = 5$ MeV/c^2 to 8 MeV/c^2, depending on the shower energies and polar angles).

To obtain optimal momentum resolution for the π^0 candidate, we perform a kinematic fit to constrain the measured invariant mass $M(\gamma\gamma)$ to the known π^0 mass using the method of the Lagrange Multipliers. Each photon ($i = 1, 2$) has a measured energy E_i, an azimuthal angle ϕ_i, and a polar angle θ_i, with the associated errors σ_{E_i}, σ_{ϕ_i} and σ_{θ_i} given by Equations (4.4)-(4.7). The mass of the pion candidate is given by $M(\gamma\gamma) = 2\,E_1 E_2(1-\cos\psi)$, where $\psi = \psi(\phi_1,\theta_1,\phi_2,\theta_2)$ is the opening angle between the two photons. In the Lagrange Multipliers minimization [102], the unconstrained observables of the neutral pion are represented by the vector $\vec{y} = (E_1,\phi_1,\theta_1,E_2,\phi_2,\theta_2)$. The constrained or adjustable parameters, represented by the vector $\vec{\eta} = (\hat{E}_1,\hat{\phi}_1,\hat{\theta}_1,\hat{E}_2,\hat{\phi}_2,\hat{\theta}_2)$, are allowed to float within their respective errors to minimize

$$\chi^2(\vec{\eta},\lambda) = \sum_{j=1}^{6}\left(\frac{y_j-\eta_j}{\sigma_j}\right)^2 + 2\,\lambda[M(\gamma\gamma) - m_{\pi^0}]. \tag{6.8}$$

The constrained mass is obtained with the adjusted value $\vec{\eta} = (\hat{E}_1,\hat{\phi}_1,\hat{\theta}_1,\hat{E}_2,\hat{\phi}_2,\hat{\theta}_2)$ which satisfy the conditions $\partial\chi^2(\vec{\eta},\lambda)/\partial\eta_i = 0$ and $\partial\chi^2(\vec{\eta},\lambda)/\partial\lambda = 0$. Subsequently, we use $\hat{E}_i$, $\hat{\phi}_i$, and $\hat{\theta}_i$ in the calculation of the D^0 invariant mass $M(K^-\pi^+\pi^0)$.

6.9 D^0 Reconstruction

Once the lepton and all the hadrons have been reconstructed, we combine them to search for the decay mode of interest:

$$\begin{aligned} B^- \longrightarrow\ & \mathrm{D}^{*+}\pi^-\ell^-\bar{\nu}_\ell \\ & \hookrightarrow D^0\pi^+ \\ & \qquad \hookrightarrow D^0 \to K^-\pi^+ \text{ or } D^0 \to K^-\pi^+\pi^0. \end{aligned}$$

In the reconstruction of our signal, we take advantage of the charge correlation between the lepton and the kaon in the D^0 decay by requiring the lepton (ℓ^-) and the kaon (K^-) to have the same sign. The D^0 candidates are required to have a scaled momentum $x_D = |\mathbf{p}_D|/\,[E^2_{\text{beam}} - M^2(D)]^{\frac{1}{2}} < 0.475$, which is the kinematically allowed range for a D^0 meson from the signal. The scaled momentum requirement suppresses fast D^0s from continuum events. The reconstruction of the D^0 meson in the $K^-\pi^+$ and $K^-\pi^+\pi^0$ modes is outlined below.

6.9.1 The $D^0 \to K^-\pi^+$ Mode

$D^0 \to K^-\pi^+$ candidates are formed from two tracks identified as a K^- and a π^+. The invariant mass $M(K^-\pi^+)$ for candidates which pass all the selection criteria described above is shown in Figure 6.6. The $K^-\pi^+$ combinations are required to have an invariant mass within 16 MeV/c^2 ($\sim 2\sigma$) of the nominal D^0 mass [6].

6.9.2 The $D^0 \to K^-\pi^+\pi^0$ Mode

$D^0 \to K^-\pi^+\pi^0$ candidates are formed from a combination of a neutral pion and two tracks identified as a K^- and a π^+. False particle combinations produce a background for $D^0 \to K^-\pi^+\pi^0$ that is inherently worse than for $D^0 \to K^-\pi^+$. To reduce these random combinations, we enforce a minimum momentum of 800 MeV/c for the D^0 candidates.

In addition, we select regions of the $D^0 \to K^-\pi^+\pi^0$ Dalitz plot to take advantage of the resonant substructure of the decay. A Dalitz plot shows the fundamental kinematic variables of a three-body final state as a scatter plot. Departure from a uniform distribution occurs due to angular momentum and parity conservation and/or resonant substructures and their interferences. The resonant substructures of the $D^0 \to K^-\pi^+\pi^0$ decay have been measured accurately by the E691 collaboration [103]. The resonant decays $D^0 \to K^-\rho^+$, $D^0 \to K^*(892)^-\pi^+$, and $D^0 \to \bar{K}^*(892)^0\pi^0$ all contribute to the rate for $D^0 \to K^-\pi^+\pi^0$. The Dalitz plot for these decays is shown in Figure 6.7. One notes the strong departure from a uniform density due to the $K^*(892)$ and the ρ resonances. The Dalitz probability is defined as a function of the decay amplitude ($\mathcal{M}$) for $D^0 \to K^-\pi^+\pi^0$:

$$W_{\text{Dalitz}} = \frac{\mathcal{M}^2[M(K^-\pi^+), M(K^-\pi^0)]}{\mathcal{M}^2_{\text{max}}}, \tag{6.9}$$

where we use E691 results [103] to compute the decay amplitude for each $K^-\pi^+\pi^0$ combination. To reduce the amount of fake D^0 in the $K^-\pi^+\pi^0$ mode, we cut on the Dalitz probability and on the π^0 energy. Figure 6.8 shows E_{π^0} versus W_{Dalitz} for signal MC and for background, which demonstrates that a two-dimensional cut improves $\mathtt{S}^2/(\mathtt{S}+\mathtt{B})$. A MC signal sample ($\mathtt{S}$) was used together with a background sample

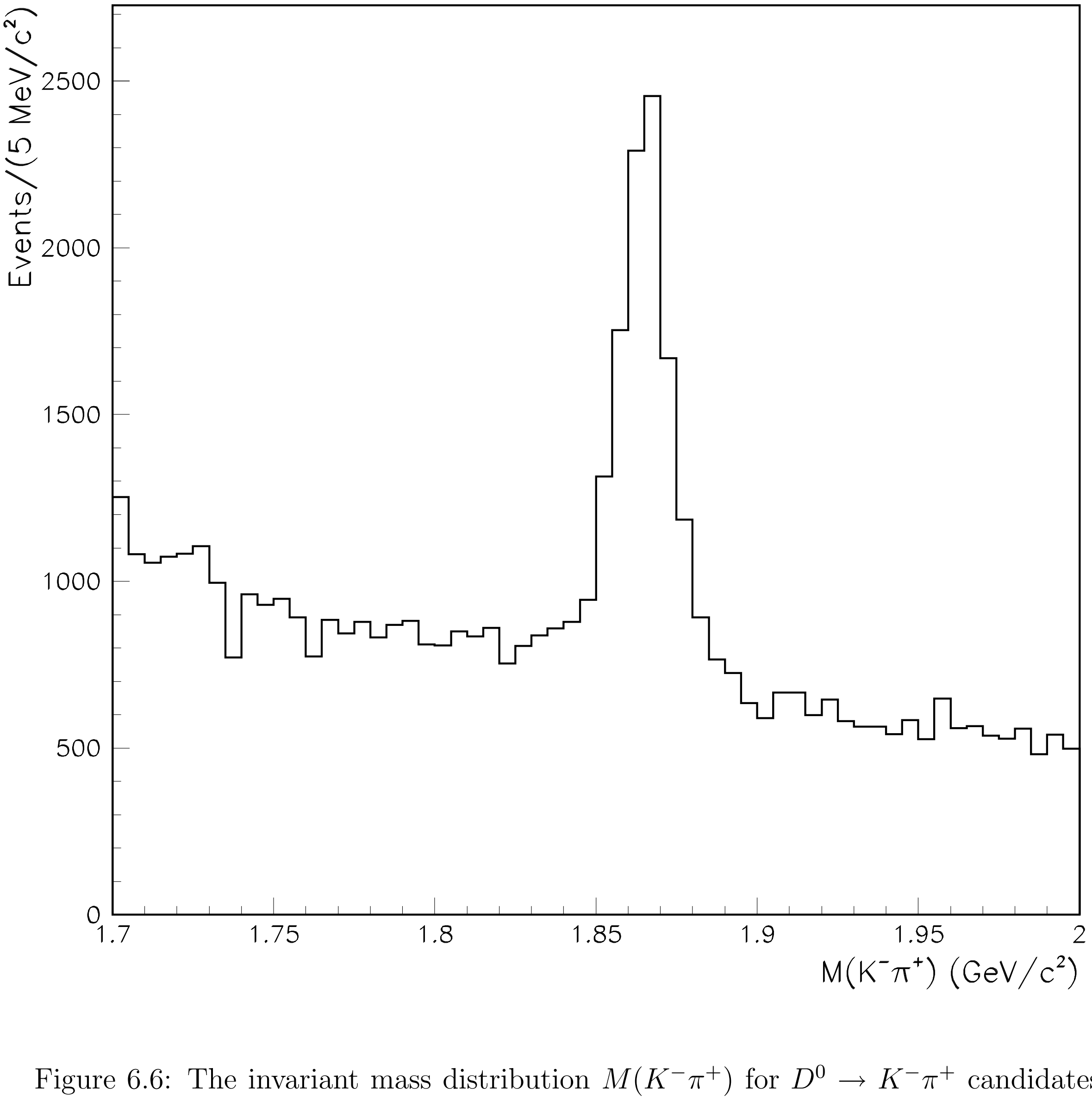


Figure 6.6: The invariant mass distribution $M(K^-\pi^+)$ for $D^0 \to K^-\pi^+$ candidates in data.

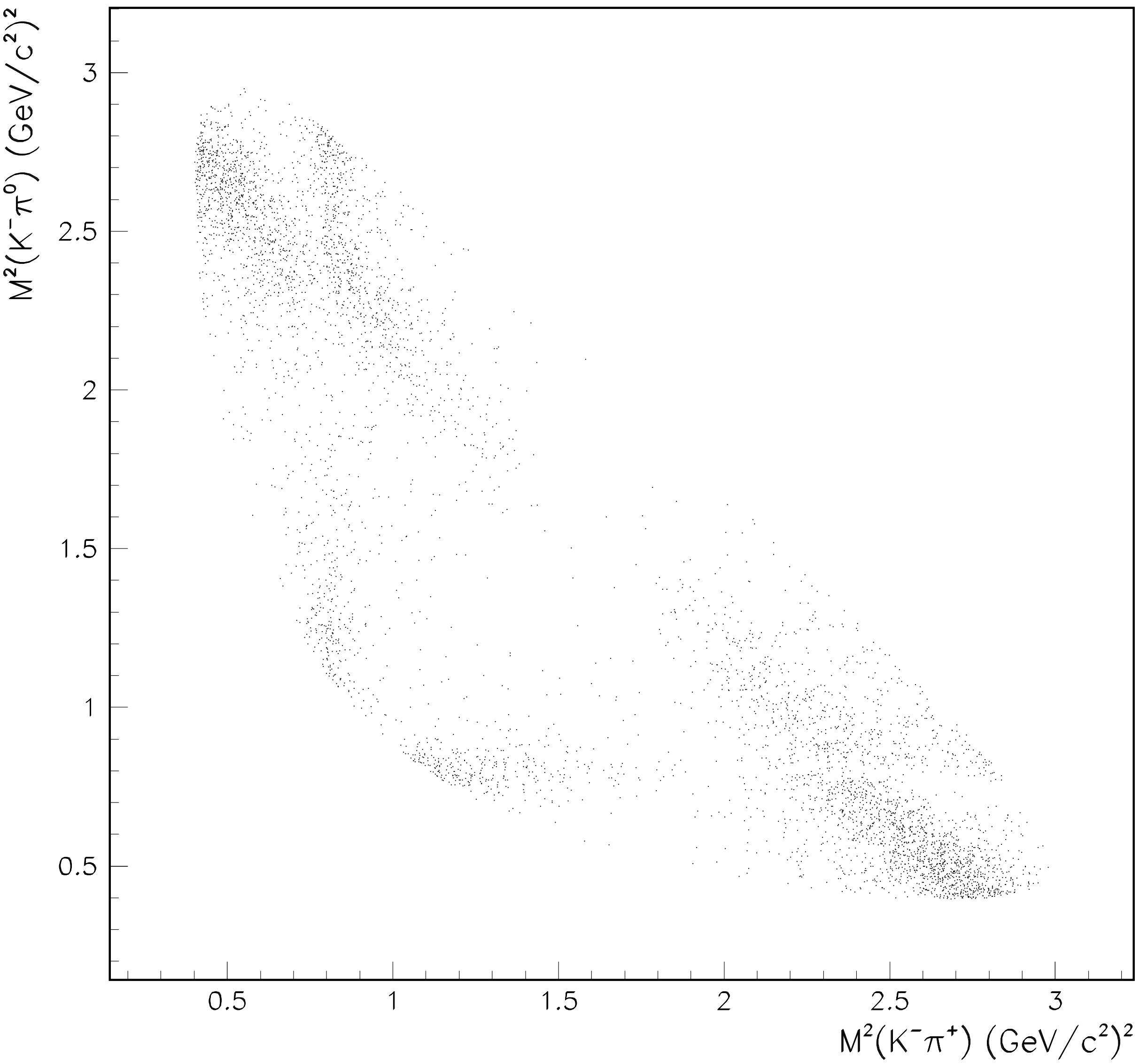


Figure 6.7: The Dalitz plot for the decay $D^0 \rightarrow K^-\pi^+\pi^0$ from MC signal events based on the measurement by the E691 collaboration. The horizontal and vertical bands are due to $K^*(892)^-$ and $\bar{K}^*(892)^0$ resonances respectively, while the diagonal band is due to the $\rho^+(770)$ resonance.

(B) taken from the D^0 sideband in real $\bar{B}^0 \rightarrow D^{*+}\ell^-\bar{\nu}_\ell$ events[2]. The solid line in Figure 6.8 shows the cut used to select $D^0 \rightarrow K^-\pi^+\pi^0$. We keep all candidates with $E_{\pi^0} \geq \frac{2}{W_{\mathrm{dalitz}}-15} + \frac{1}{4}$. The introduction of a two-dimensional cut in E_{π^0} vs W_{Dalitz} space improves $\mathtt{S}^2/(\mathtt{S}+\mathtt{B})$ by 15% over two one-dimensional cuts of $W_{\mathrm{Dalitz}} > 35$ and $E_{\pi^0} > 350$ MeV.

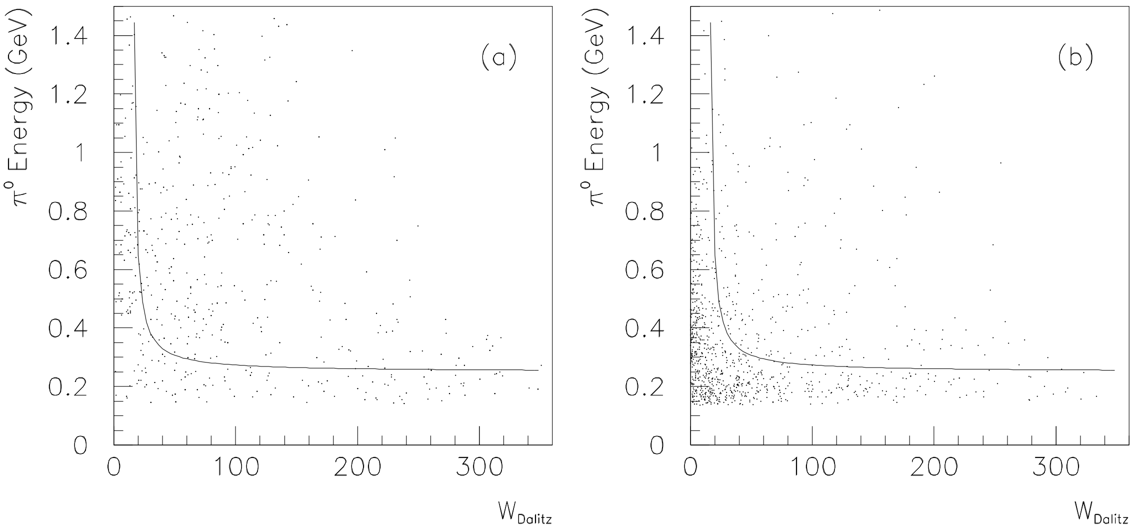


Figure 6.8: The energy of the neutral pion E_{π^0} versus the Dalitz probability of the decay $D^0 \rightarrow K^-\pi^+\pi^0$. E_{π^0} versus W_{Dalitz} is shown for signal MC in (a) and for background in (b). The two-dimensional cut is the solid line. We accept candidates with $E_{\pi^0} \geq \frac{2}{W_{\mathrm{dalitz}}-15} + \frac{1}{4}$.

The invariant mass of the D^0 candidates which meet all the $K^-\pi^+\pi^0$ selection criteria is shown in Figure 6.9. Only candidates with an invariant mass within 25 MeV/c^2 ($\sim 2\sigma$) of the nominal D^0 mass are accepted.

[2]We do not optimize $\mathtt{S}^2/(\mathtt{S}+\mathtt{B})$ using a D^0 sideband from our D^0_J sample in data because statistical limitations may lead to bias. The D^0 sideband from our D^0_J candidates is used as a consistency check. The fake D^0 mesons from exclusive $\bar{B}^0$ semileptonic decays to D^{*+} mesons mimic the dynamics of the fake D^0s in $B^- \rightarrow D^0_J\ell^-\bar{\nu}_\ell$ quite well. We could not use generic $B\bar{B}$ MC because the Dalitz decay $D^0 \rightarrow K^-\pi^+\pi^0$ is not modeled properly in the generator used for the generic $B\bar{B}$ MC.

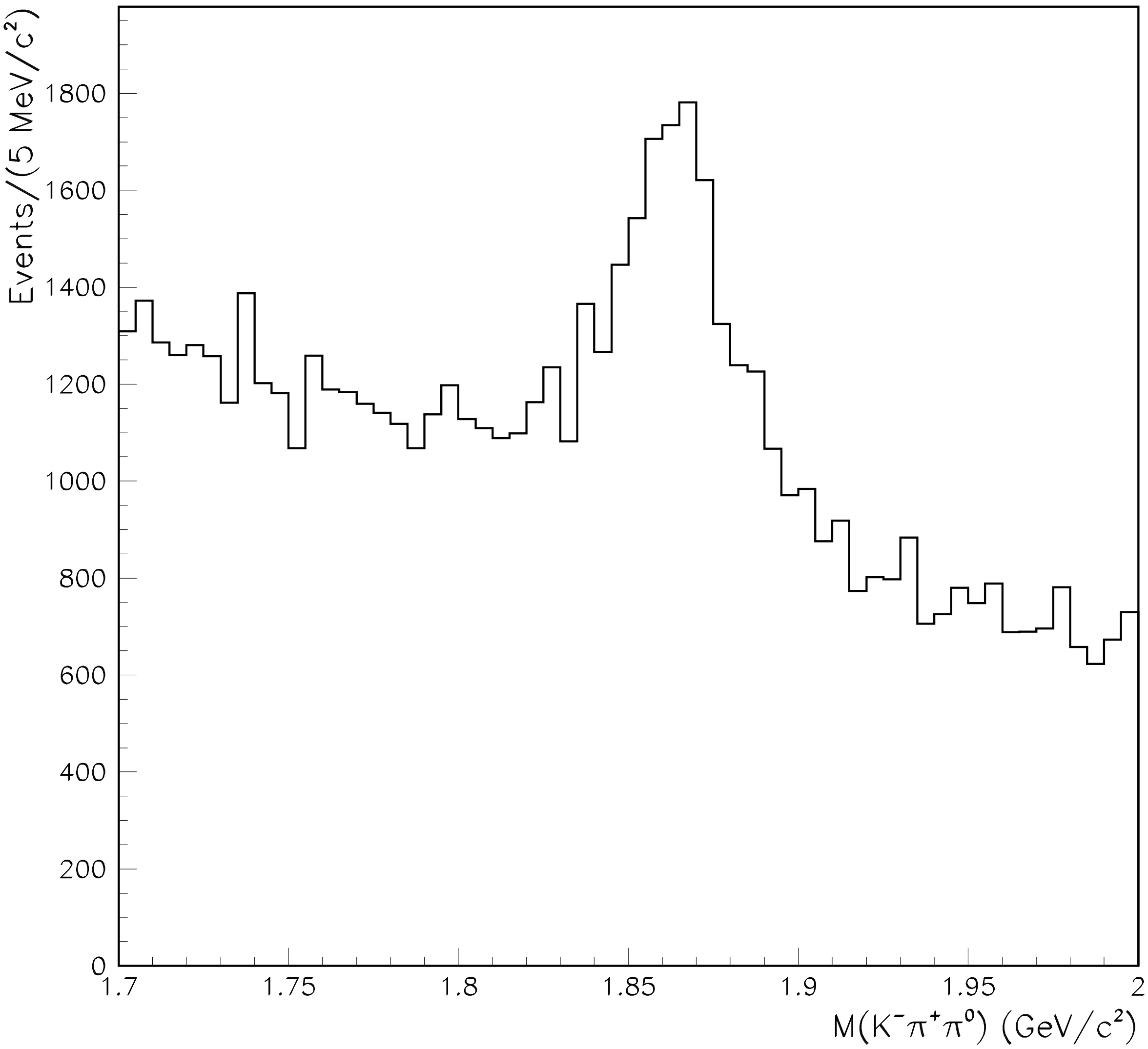


Figure 6.9: The invariant mass distribution $M(K^-\pi^+\pi^0)$ for $D^0 \rightarrow K^-\pi^+\pi^0$ candidates in data. Notice the additional combinatoric background compared to the $K^-\pi^+$ mode (Figure 6.6) due to soft photons forming false neutral pion candidates.

6.10 $D^{*+} \to D^0\pi^+$ Reconstruction

D^{*+} candidates are reconstructed in the channel $D^{*+} \to D^0\pi^+$. The small phase space available for this decay has significant consequences. The first is that the intrinsic width of the D^{*+} resonance is small so that the observed width is dominated by detector resolution. The second is that the detector resolution on the mass difference $\delta m = M(D^0\pi^+) - M(D^0)$ is much better than that for $M(D^0\pi^+)$ because the tracking errors on the $D^0\pi^+$ and D^0 masses are highly correlated and mainly cancel in the difference. The resolution on δm can be written as

$$\begin{aligned} \sigma[\delta m] &= \sigma[M(D\pi) - M(D)] = \sigma\left[\frac{M^2(D\pi) - M^2(D)}{M(D\pi) + M(D)}\right] \qquad (6.10) \\ &= \frac{\sigma[M^2(D\pi) - M^2(D)]}{M(D\pi) + M(D)} - \frac{M(D\pi) - M(D)}{M(D\pi) + M(D)}\sigma[M(D\pi) + M(D)]. \end{aligned}$$

The next step is to use the approximation $M(D^*) - M(D) \ll M(D^*) + M(D)$ and the fact that $M^2(D\pi) - M^2(D) = m_\pi^2 + 2\,\mathbf{p}_D \cdot \mathbf{p}_\pi$. This gives

$$\sigma[\delta m] \simeq \frac{2\,\sigma[\mathbf{p}_D \cdot \mathbf{p}_\pi]}{M(D\pi) + M(D)}. \qquad (6.11)$$

As one can see, most of the uncertainties from the D have canceled. For typical D^* candidates, the opening angle between the D and the slow π is small. Since the momenta of the charged particles are well measured in the drift chambers, the error on the mass difference is dominated by the error on the $D - \pi$ opening angle. In the CLEO II detector, the resolution on the mass difference δm is about 1 Mev/c^2. The small width of the δm peak allows very little combinatoric background under the signal.

The slow pion used to form the D^{*+} must satisfy the selection criteria presented in Section 6.5. We further require $x_{D^*} = |\mathbf{p}_{D^*}|/\,[E_{\text{beam}}^2 - M^2(D^*)]^{\frac{1}{2}} < 0.495$ and the reconstructed mass difference $\delta m = M(D^0\pi^+) - M(D^0)$ to be within 2 MeV/c^2 of the known $D^{*+} - D^0$ mass difference [6]. In Figure 6.10, the δm distributions are shown in the $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^0$ channels after all the selection criteria, except the δm cut, have been applied.

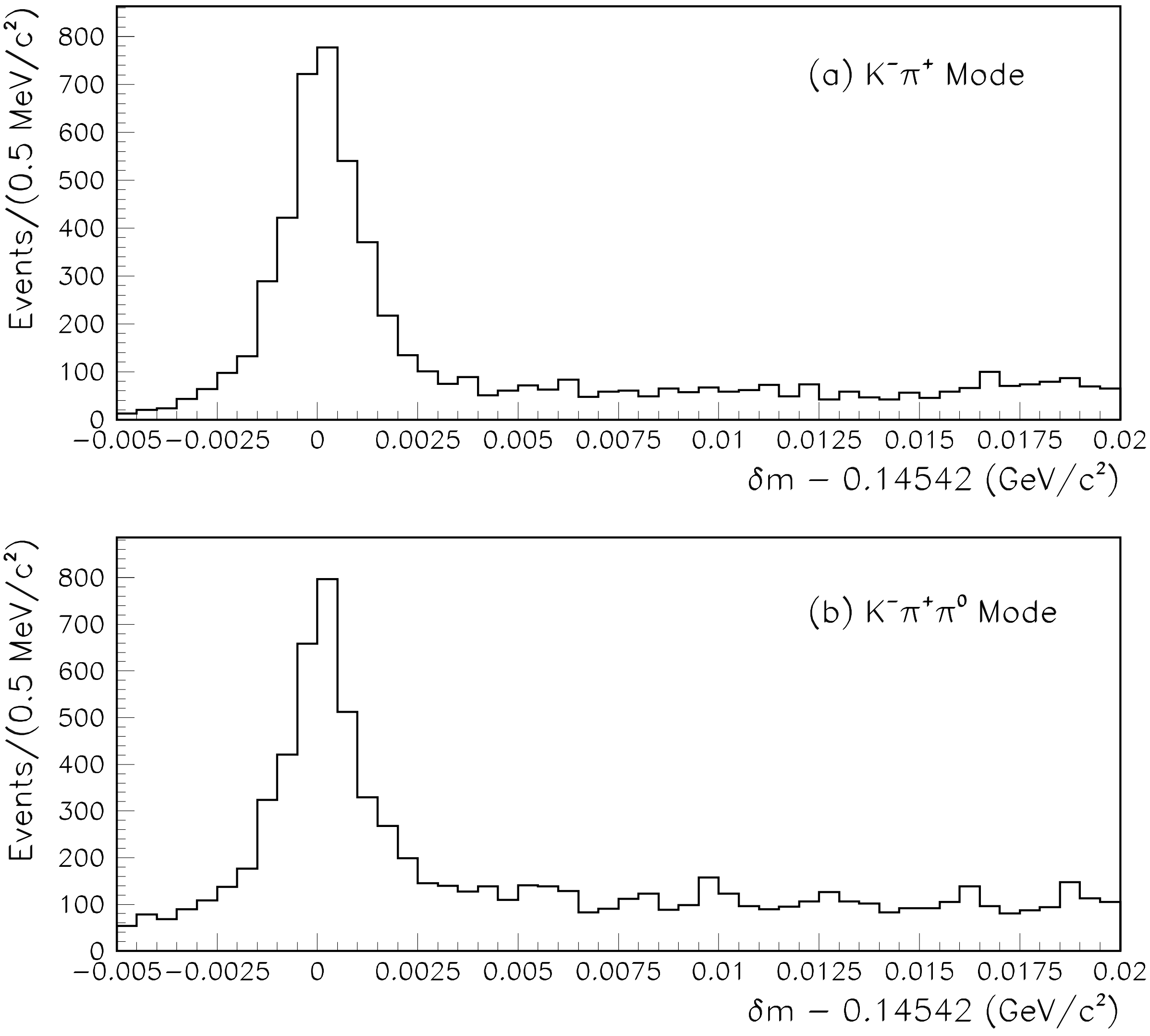


Figure 6.10: The mass difference $\delta m = M(D^0\pi^+) - M(D^0) - 0.14542$ in data for the (a) $D^0 \to K^-\pi^+$ and (b) $D^0 \to K^-\pi^+\pi^0$ modes.

6.11 $D_J^0\,\ell^-$ Candidates

The D^{*+} candidate is then combined with an additional π^- in the event to form a D_J^0 candidate. The D_J^0 candidates must have a scaled momentum $x_{D_J} < 0.5$, the kinematic limit from B decays. These D_J^0 candidates are then paired with right-sign leptons to form $D_J^0\,\ell^-$ candidates for $B^- \to D_J^0 \ell^- \bar{\nu}_\ell$ decays. By right-sign lepton we mean that a D_J^0 must be paired with a negatively charged lepton, while a $\bar{D}_J^0$ requires a positively charged lepton. The D_J^0 and $\bar{D}_J^0$ are distinguished by the sign of the kaon form the D^0 or $\bar{D}^0$ decay.

At CESR, the energy of the B meson (E_B) must be equal to the beam energy (E_{beam}), which is precisely known by machine optics. Hence, one can determine the magnitude of the momentum of the B meson from E_{beam} and the known B meson mass [6]:

$$|\mathbf{p}_B| = \sqrt{E_{\text{beam}}^2 - m_B^2} \tag{6.12}$$

In our case, $E_{\text{beam}} = E_B = E_{D_J\ell} + E_{\bar{\nu}_\ell}$ and $\mathbf{p}_B = \mathbf{p}_{D_J\ell} + \mathbf{p}_{\bar{\nu}_\ell}$. Even with no ability to detect the neutrino emitted from $B^- \to D_J^0 \ell^- \bar{\nu}_\ell$, a kinematic constraint on the magnitude of the neutrino momentum can be obtained:

$$|\mathbf{p}_{\bar{\nu}_\ell}|^2 = |\mathbf{p}_{D_J\ell}|^2 + |\mathbf{p}_B|^2 - 2\,|\mathbf{p}_B||\mathbf{p}_{D_J\ell}|\,\cos\theta_{B-D_J\ell} \tag{6.13}$$

Here, $\theta_{B-D_J\ell}$ is the angle between the B momentum and the $D_J\,\ell$ momentum, as shown in Figure 6.11. The B meson four-momentum is $(E_B, \mathbf{p}_B)$ and the sum of the D_J and the lepton four-momentum is $(E_{D_J\ell}, \mathbf{p}_{D_J\ell})$.

There is significant background in this analysis from real D^{*+}s combined with pions that are not from D_J^0 mesons. To suppress this background we select $D_J^0\,\ell^-$ pairs that are consistent with $B^- \to D_J^0 \ell^- \bar{\nu}_\ell$ decays and reject $D^{*+}\,\ell^-$ pairs that are consistent with $\bar{B}^0 \to D^{*+} \ell^- \bar{\nu}_\ell$. Although we do not know the direction of the B meson or the $\bar{\nu}_\ell$, we use conservation of momentum to construct the angle $\theta_{B-D_J\ell}$ between the momenta of the B and the $D_J^0\,\ell^-$:

$$\begin{aligned} \cos\theta_{B-D_J\ell} &= \frac{|\mathbf{p}_{D_J\ell}|^2 + |\mathbf{p}_B|^2 - |\mathbf{p}_{\bar{\nu}_\ell}|^2}{2|\mathbf{p}_B||\mathbf{p}_{D_J\ell}|} \\ &= -\frac{m_B^2 + M^2(D_J\ell) - 2E_B E(D_J\ell)}{2|\mathbf{p}_B||\mathbf{p}_{D_J\ell}|}. \end{aligned} \tag{6.14}$$

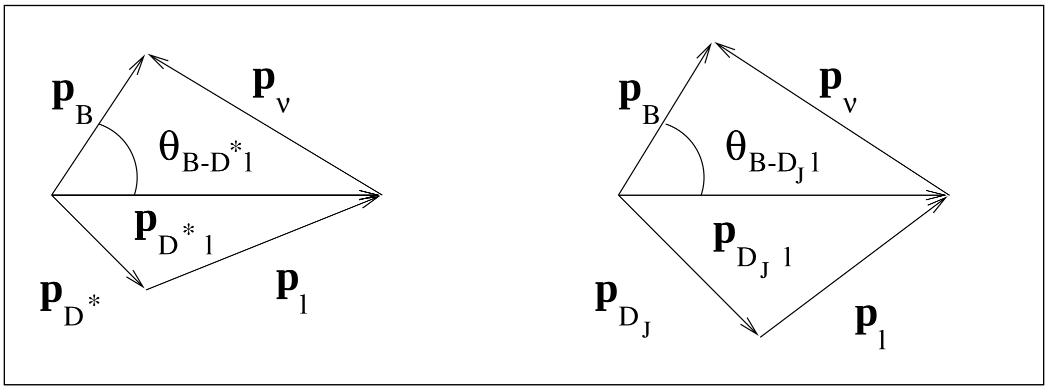


Figure 6.11: Definitions of the angles $\theta_{B-D^*\ell}$ and $\theta_{B-D_J\ell}$

For true $B^- \to D_J^0\ell^-\bar{\nu}_\ell$ decays, $\cos\theta_{B-D_J\ell}$ will have physical values, $|\cos\theta_{B-D_J\ell}| \leq 1$. For background decays, in which particles from the B decay chain are missing (*e.g.*, $\bar{B} \to D_J^0 X\ell^-\bar{\nu}_\ell$), or in which an extra random pion has been added to a true semileptonic B decay with a D^{*+} in the final state, $\cos\theta_{B-D_J\ell}$ is shifted toward non-physical values. Therefore, we require $D_J^0\,\ell^-$ candidates to have $|\cos\theta_{B-D_J\ell}| \leq 1$. When the requirement $|\cos\theta_{B-D_J\ell}| \leq 1$ is applied, 93% of the $B^- \to D_J^0\ell^-\bar{\nu}_\ell$ decays are retained and 60% of the background is rejected (see Figure 6.12).

Similarly, we construct the angle $\cos\theta_{B-D^*\ell}$, between the B and the $D^{*+}\,\ell^-$ momentum:

$$\begin{aligned}\cos\theta_{B-D^*\ell} &= \frac{|\mathbf{p}_{D^*\ell}|^2 + |\mathbf{p}_B|^2 - |\mathbf{p}_{\bar{\nu}_\ell}|^2}{2|\mathbf{p}_B||\mathbf{p}_{D^*\ell}|} \\ &= -\frac{m_B^2 + M^2(D^*\ell) - 2E_B E(D^*\ell)}{2|\mathbf{p}_B||\mathbf{p}_{D^*\ell}|}.\end{aligned} \tag{6.15}$$

For true $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$ decays, $|\cos\theta_{B-D^*\ell}| \leq 1$. For $B^- \to D_J^0\ell^-\bar{\nu}_\ell$ decays, in which the correct D^{*+} and lepton have been used (and the D_J^0 daughter π^- has been ignored) in the computation of $|\mathbf{p}_{\bar{\nu}_\ell}|$, $\cos\theta_{B-D^*\ell}$ is shifted toward negative values. We therefore require $\cos\theta_{B-D^*\ell} < -1$. This cut substantially reduces the dominant background to this analysis which comes from true $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$ decays, where the D^{*+} is combined with a random pion in the event to make a D_J^0 candidate. When the requirement $\cos\theta_{B-D^*\ell} < -1$ is applied, 65% of the $B^- \to D_J^0\ell^-\bar{\nu}_\ell$ decays are retained and 78% of the background is rejected. The main contribution to the background is $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$, as illustrated in Figure 6.13. The $\cos\theta_{B-D^*\ell}$ cut removes 92% of events from $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$.

The definitions of $\cos\theta_{B-D_J\ell}$ and $\cos\theta_{B-D^*\ell}$ are summarized in Figure 6.11. When the requirements $|\cos\theta_{B-D_J\ell}| \leq 1$ and $\cos\theta_{B-D^*\ell} < -1$ are applied together, they retain 60% of the $B^- \to D_J^0\ell^-\bar{\nu}_\ell$ decays and reject 89% of the background remaining after all other cuts. Figures 6.12 and 6.13 show the $\cos\theta_{B-D_J\ell}$ and $\cos\theta_{B-D^*\ell}$ distributions for MC signal events and for background events from generic $B\bar{B}$ MC. The background from $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$ is the hatched region. The $\cos\theta_{B-D_J\ell}$ and $\cos\theta_{B-D^*\ell}$ cuts remove almost all ($\sim 96\%$) of the $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$ background.

Another useful cut to reject uncorrelated background is to require the D_J^0 and the lepton to be in opposite hemispheres: $\cos\theta_{D_J\ell} < 0$. The angle $\theta_{D_J\ell}$ is the angle between the D_J^0 and the lepton in the laboratory frame. The signal is strongly peaked near $\cos\theta_{D_J\ell} = -1$ because of the $V - A$ structure of the weak coupling in $B^- \to D_J^0\ell^-\bar{\nu}_\ell$. Figure 6.14 shows the $\cos\theta_{D_J\ell}$ distributions for signal and background. One should notice that the background also tends to peak near $\cos\theta_{D_J\ell} = -1$. The reason is that the remainder of the $\bar{B} \to D^{(*)}\ell\bar{\nu}_\ell$ background tends to be back-to-back in the lab frame. The cut on $\cos\theta_{D_J\ell}$ is conservative in terms of introducing model dependence on the acceptance. Uncorrelated background (real or fake D_J^0 from the $\bar{B}$ combined with a real or a fake lepton from the B) is uniform in $\cos\theta_{D_J\ell}$. After all the selection criteria have been applied to the generic $B\bar{B}$ MC events, the bulk of the remaining uncorrelated background is negligible

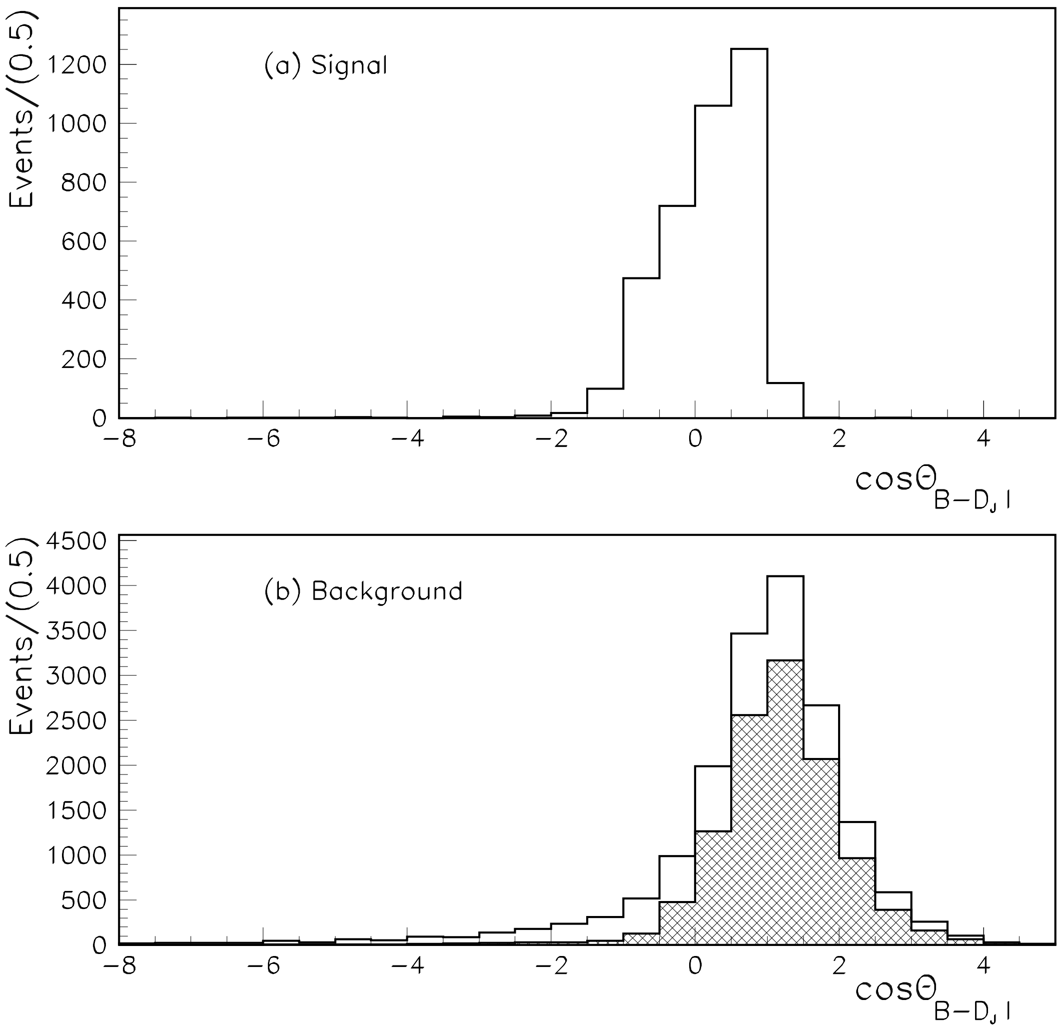


Figure 6.12: The $\cos\theta_{B-D_J\ell}$ distributions for (a) signal MC events and (b) generic $B\bar{B}$ background MC events. The cut $|\cos\theta_{B-D_J\ell}| \leq 1$ retains 93% of the $B^- \to D_J^0\ell^-\bar{\nu}_\ell$ decays and rejects 60% of the background. The hatched area is the contribution from $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$ events. The $\cos\theta_{B-D_J\ell}$ cut removes 61% of events from $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$.

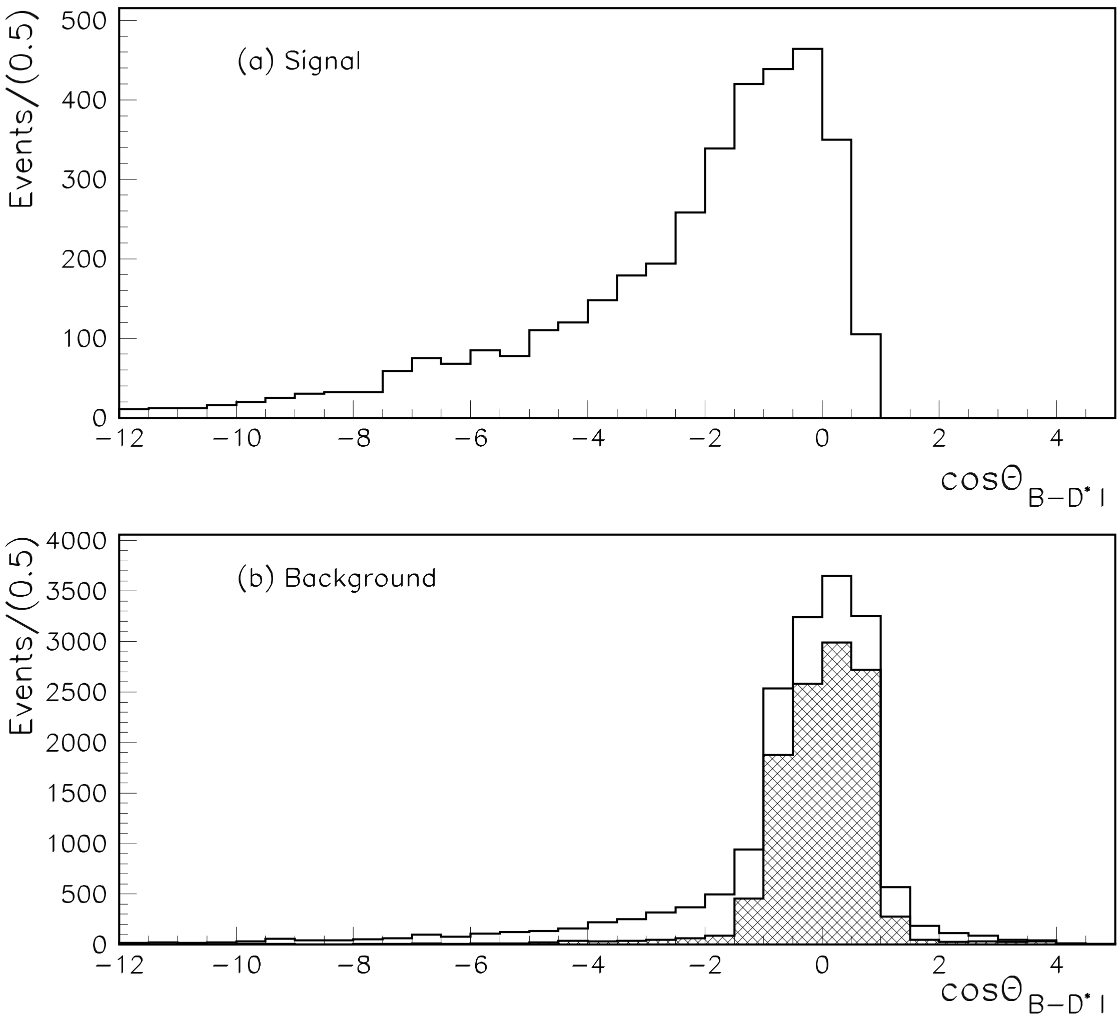


Figure 6.13: The $\cos\theta_{B-D^*\ell}$ distributions for (a) signal MC events and (b) generic $B\bar{B}$ background MC events. The cut $\cos\theta_{B-D^*\ell} < -1$ retains 65% of the $B^- \to D_J^0 \ell^- \bar{\nu}_\ell$ decays and rejects 78% of the background. The hatched area is the contribution from $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$ events. The $\cos\theta_{B-D^*\ell}$ cut removes 92% of events from $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$.

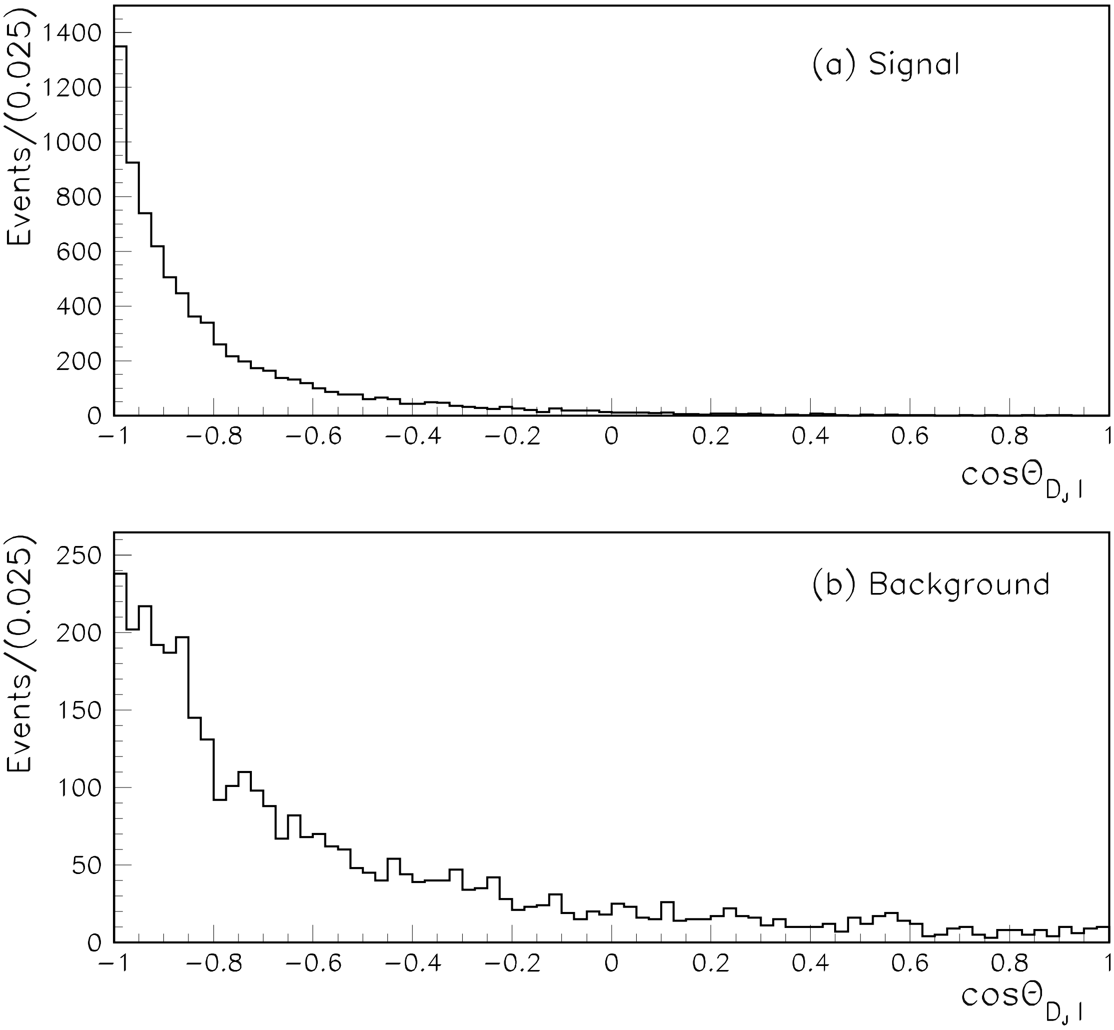


Figure 6.14: The $\cos\theta_{D_J\ell}$ distributions after all cuts except the $\cos\theta_{D_J\ell} < 0$. Signal MC in (a) and background from generic $B\bar{B}$ MC in (b).

Chapter 7

Experimental Results

In the previous chapter, we introduced the event selection criteria and reconstruction methods needed to extract a data sample that should contain a significant contribution from the signal under study: narrow $D_J^0 \to D^{*+}\pi^-$ contributions to semileptonic B decays. In this chapter, we turn our attention to the quantitative aspects of extracting the branching fractions for $B^- \to D_1^0\ell^-\bar{\nu}_\ell$ and $B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell$ from this data sample. We also measure the q^2 distribution for $B^- \to D_1^0\ell^-\bar{\nu}_\ell$. In addition, we look at the sensitivity of this analysis to nonresonant $D^{*+}\pi^-$ production in semileptonic B decays.

After applying all the selection requirements, the mass difference $\delta M_J \equiv M(D^{*+}\pi^-) - M(D^{*+})$ is calculated for each $D_J^0\ \ell^-$ candidate. Evidence for $B^- \to D_1^0\ell^-\bar{\nu}_\ell$ and $B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell$ would be seen as enhancements at the known mass difference in this distribution. We use δM_J rather than $M(D_J^0)$ for the same reason that we used δm for the D^{*+} reconstruction. In the computation of δM_J most of the D^{*+} contributions to the $M(D^{*+}\pi^-)$ and $M(D^{*+})$ errors cancel. The data and the fit described below are shown in Figures 7.1 and 7.2 for both decay modes of the D^0 meson. While there is a statistically significant enhancement in the region of the narrow D_1^0 resonance at $\delta M_J \simeq 0.412$ GeV/c^2, the evidence for a signal for the other narrow resonance, D_2^{*0}, at $\delta M_J \simeq 0.449$ GeV/c^2 is not compelling.

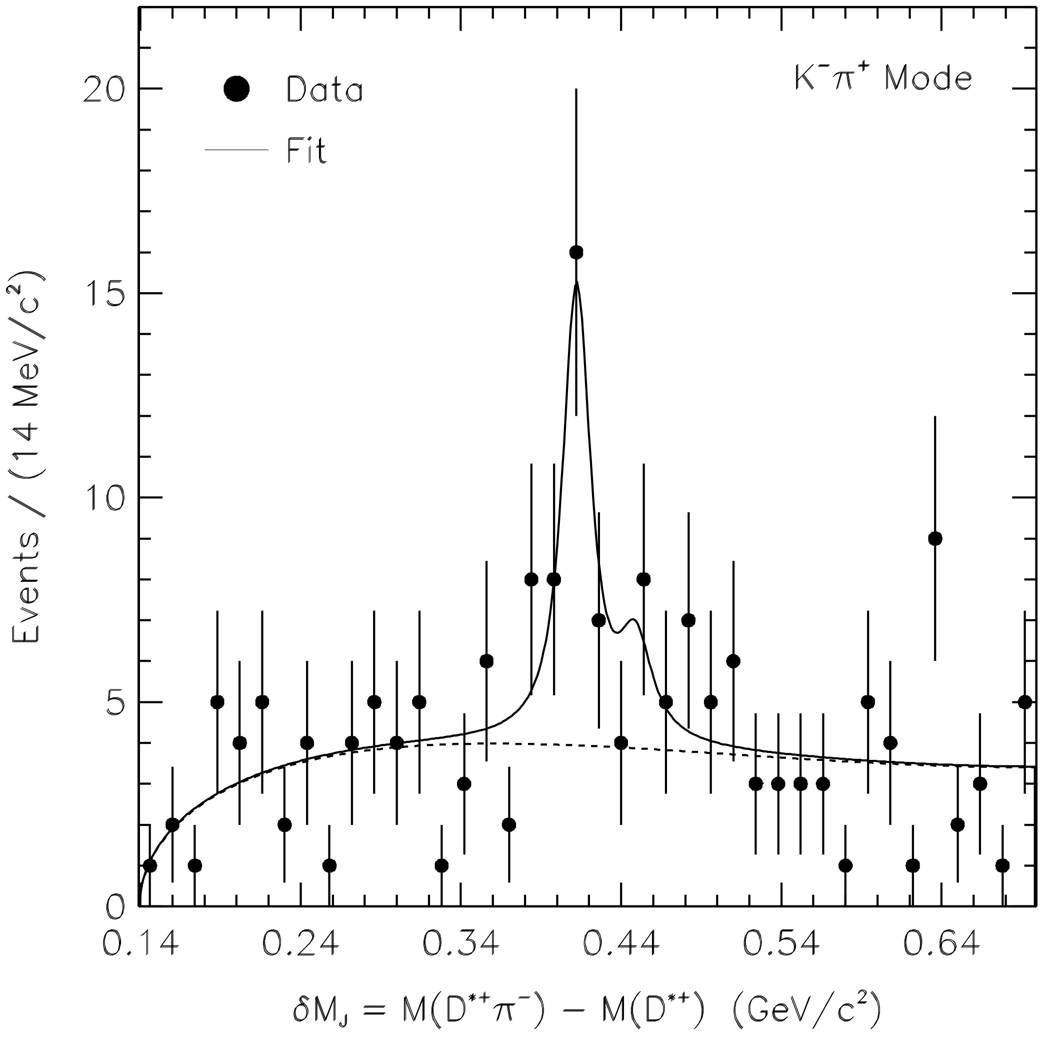


Figure 7.1: The δM_J distribution from the ON $\Upsilon(4S)$ Resonance data for $B^- \to D_1^0\ell^-\bar{\nu}_\ell$ and $B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell$ ($\ell = e$ and μ) candidates for the $D^0 \to K^-\pi^+$ mode. The dashed curve illustrates the background function, whereas the solid line shows the sum of the background and signal functions. The data selection used is summarized in Section 7.1 and the fitting procedure is described in Section 7.2.

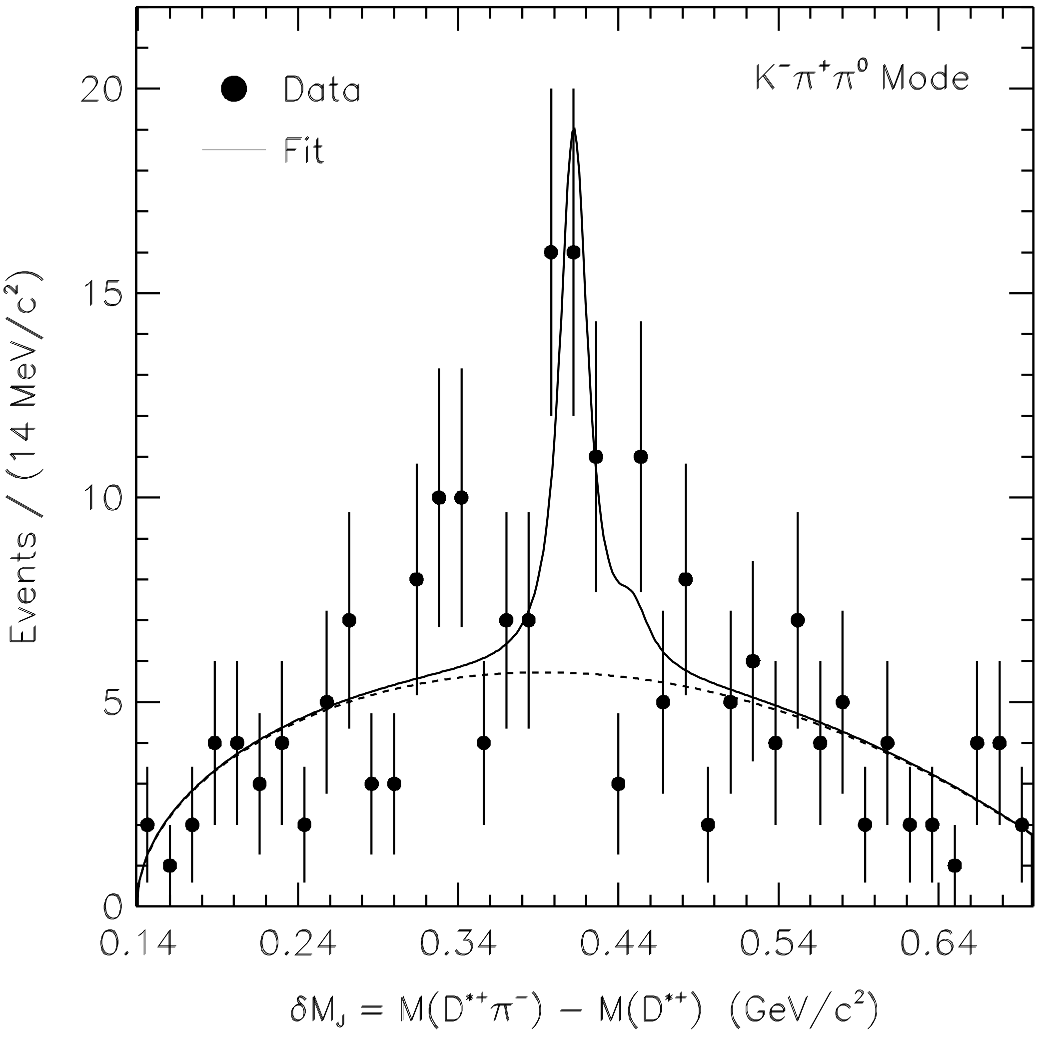


Figure 7.2: The δM_J distribution from the ON $\Upsilon(4S)$ Resonance data for $B^- \to D_1^0 \ell^- \bar{\nu}_\ell$ and $B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell$ ($\ell = e$ and μ) candidates for the $D^0 \to K^-\pi^+\pi^0$ mode. The dashed curve illustrates the background function, whereas the solid line shows the sum of the background and signal functions. The data selection used is summarized in Section 7.1 and the fitting procedure is described in Section 7.2.

7.1 Selecting The Best Candidate

After all the cuts have been applied, 46% of the events in the $D^0 \to K^-\pi^+$ mode and 41% of the events in the $D^0 \to K^-\pi^+\pi^0$ mode have two or more $D_J^0\,\ell^-$ candidate combinations. This multiple counting is mainly due to extra combinations created with random soft particles. It is important to select one signal candidate per event, otherwise the statistical errors on the signal yields may be underestimated. To avoid these complications, the best combination in each event is selected based on a confidence level or probability calculated using M(π^0), M(D^0), $\delta m = M(D^0\pi^+) - M(D^0)$, and the missing mass squared of the neutrino, $M^2(\bar{\nu}_\ell)$.

First, we look at the independent observables M(π^0), M(D^0), and δm. These observables have approximately symmetric and Gaussian distributions. We can therefore easily construct χ^2s, and then compute confidence levels.

$$\chi^2(M_{\pi^0}) = \chi^2(\vec{\eta},\lambda) \text{ of the } \pi^0 \text{ kinematic fit} \tag{7.1}$$

$$\chi^2(M_{D^0}) = [M(K\pi(\pi^0)) - m_D]^2/\sigma^2_{M_{D^0}} \tag{7.2}$$

$$\chi^2(\delta m) = [\delta m - 0.14542]^2/\sigma^2_{\delta m}\,. \tag{7.3}$$

The errors on the invariant mass of the D^0 meson ($\sigma_{M_{D^0}}$) and on the mass difference ($\sigma_{\delta m}$) are calculated from the actual error matrices of the tracks and showers used for the candidate. The quality of each track and shower is then correctly taken into account for every possible combination. Then we construct:

$$\chi_1^2(K\pi) = \chi^2(M_{D^0}) + \chi^2(\delta m), \tag{7.4}$$

$$\chi_1^2(K\pi\pi^0) = \chi^2(M_{\pi^0}) + \chi^2(M_{D^0}) + \chi^2(\delta m). \tag{7.5}$$

From χ_1^2 with n_1 degrees of freedom ($n_1 = 2$ for the $K\pi$ mode and $n_1 = 3$ for the $K\pi\pi^0$ mode), we compute a confidence level

$$\alpha_1 = \mathrm{CL}(\chi_1^2, n_1). \tag{7.6}$$

Secondly, we look at $M^2(\bar{\nu}_\ell)$ of the B meson candidate. In a decay $B^- \to D_J^0\ell^-\bar{\nu}_\ell$, one can calculate the mass of the particle recoiling against the $D_J^0\,\ell^-$ system under the assumption that the observed D_J^0 and lepton are produced from a single B meson. In this case

$$\begin{aligned} P^2_{\rm miss} &= (P_B - P_{D_J\ell})^2 \\ &= m_B^2 + M^2(D_J\ell) - E_B E(D_J\ell) + 2|\mathbf{p}_B||\mathbf{p}_{D_J\ell}|\cos\theta_{B-D_J\ell}\,. \end{aligned} \tag{7.7}$$

For true $B^- \to D_J^0\ell^-\bar{\nu}_\ell$ events, the missing mass squared of the neutrino is $M^2(\bar{\nu}_\ell) = P^2_{\rm miss} = 0$. If additional particles are produced in the semileptonic B decay, then $M^2(\bar{\nu}_\ell)$ increases. If the neutrino comes from $\bar{B} \to D^{(*)}\ell\bar{\nu}_\ell$, instead of $B^- \to D_J^0\ell^-\bar{\nu}_\ell$, then $M^2(\bar{\nu}_\ell)$ decreases.

Despite the fact that the neutrino is massless, there is not enough information in the event to calculate $M^2(\bar{\nu}_\ell)$ exactly because the directions of the B mesons from $\Upsilon(4S) \to B\bar{B}$ are not known. However, we know that the B mesons momenta are fairly small ($|\mathbf{p}_B| \simeq 300$ MeV$/c$), so to a good approximation, we can write:

$$M^2(\bar{\nu}_\ell) \simeq m_B^2 + M^2(D_J\ell) - 2E_B E(D_J\ell). \tag{7.8}$$

The RMS width of the $M^2(\bar{\nu}_\ell)$ distribution is then dominated by the neglect of $|\mathbf{p}_B|$ in the calculation. The resolution of $M^2(\bar{\nu}_\ell)$ for signal events is about 380 MeV$/c$ before the angular cut $\cos\theta_{B-D^*\ell}$ is applied, as shown in Figure 7.3(a). Unfortunately, the $\cos\theta_{B-D^*\ell}$ cut, which is meant to suppress $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$ backgrounds, makes the $M^2(\bar{\nu}_\ell)$ distribution asymmetric and no longer Gaussian. One can see the effect of the $\cos\theta_{B-D^*\ell}$ cut in Figure 7.3(a). For this reason, we are unable to construct a χ^2 for $M^2(\bar{\nu}_\ell)$ like we did for the other observables. Instead we will use the $M^2(\bar{\nu}_\ell)$ distribution as the probability density function $f(x) \equiv f(M^2(\bar{\nu}_\ell))$, and we will compute a confidence level as follows:

$$\alpha_2 = F(x_{\rm meas}) = \begin{cases} 2\displaystyle\int_{-\infty}^{x_{\rm meas}} f(x)\,dx & \text{if } x_{\rm meas} \le x_{\rm median} \\ 2\displaystyle\int_{x_{\rm meas}}^{+\infty} f(x)\,dx & \text{if } x_{\rm meas} > x_{\rm median} \end{cases} \tag{7.9}$$

where $x_{\rm meas}$ is the measured missing mass squared, $M^2(\bar{\nu}_\ell)$, of the candidate, and $F(x_{\rm meas})$ is the cumulative distribution of $M^2(\bar{\nu}_\ell)$ with $F(x_{\rm meas} = x_{\rm median}) = 100\%$. We should note however that in our case: $x_{\rm median} \simeq \bar{x}$. Since the resolution on $M^2(\bar{\nu}_\ell)$

depends on q^2, the probability density function $f(x)$ is determined from signal Monte Carlo for five bins of q^2. The definition of a confidence level is not unique, but Monte Carlo studies have shown that this technique is the best of several possible alternatives.

Because the direction of the neutrino is not known, q^2 cannot be calculated exactly. The missing information is in fact the azimuthal angle of the B momentum around the $D_J\,\ell$ system. The projection of the B momentum on the $D_J\,\ell$ system is known, being the quantity $\cos\theta_{B-D_J\ell}$ defined in Figure 6.11; but the direction of the B meson with respect to the $D_J\,\ell$ system is unknown. In the computation of q^2, we consider two extreme configurations. The first one is when the opening angle between the B and the D_J vectors is minimal, corresponding to the maximum q^2 for a given $\cos\theta_{B-D_J\ell}$. The second is when the opening angle between the B and the D_J vector is maximal, corresponding to the minimum q^2 for a given $\cos\theta_{B-D_J\ell}$. In both cases, the q^2 is restricted to be physical. The average of the maximum and minimum possible value of q^2 for each event provides good agreement between generated and reconstructed values in MC simulation, as shown in Figure 7.3(b-c).

Finally, the significance level of the two independent tests applied on a given $D_J^0\,\ell^-$ candidate is [104]:

$$\alpha = \int_{\alpha_1'\alpha_2'<\alpha_1\alpha_2} d\alpha_1'\,d\alpha_2' = \alpha_1\alpha_2[1-\ln(\alpha_1\alpha_2)]. \tag{7.10}$$

Only the combination with the largest α is kept[1]. The main source of double counting comes from real or fake D^{*+} paired with random pions. The second source comes from multiple D^{*+} candidates. The other two sources are multiple D^0 candidates and multiple lepton candidates respectively. The contribution of multiple D^0 candidates is somewhat higher in the $K^-\pi^+\pi^0$ mode due to soft photons combinatorics.

To investigate the behavior of the estimator on signal MC events, we used MC generator level information to tag the reconstructed combinations. In the remainder of this discussion, we refer to the events which are correctly (incorrectly) reconstructed as tagged (untagged) MC signal events. For signal $B^-\to D_1^0\ell^-\bar{\nu}_\ell$ MC events, the

[1]If the missing mass squared $M^2(\bar{\nu}_\ell)$ was a Gaussianly distributed observable, extracting a confidence from $\chi^2=\chi_1^2+\chi_2^2$ with $n=n_1+1$ degrees of freedom would be the equivalent to Equation (7.10).

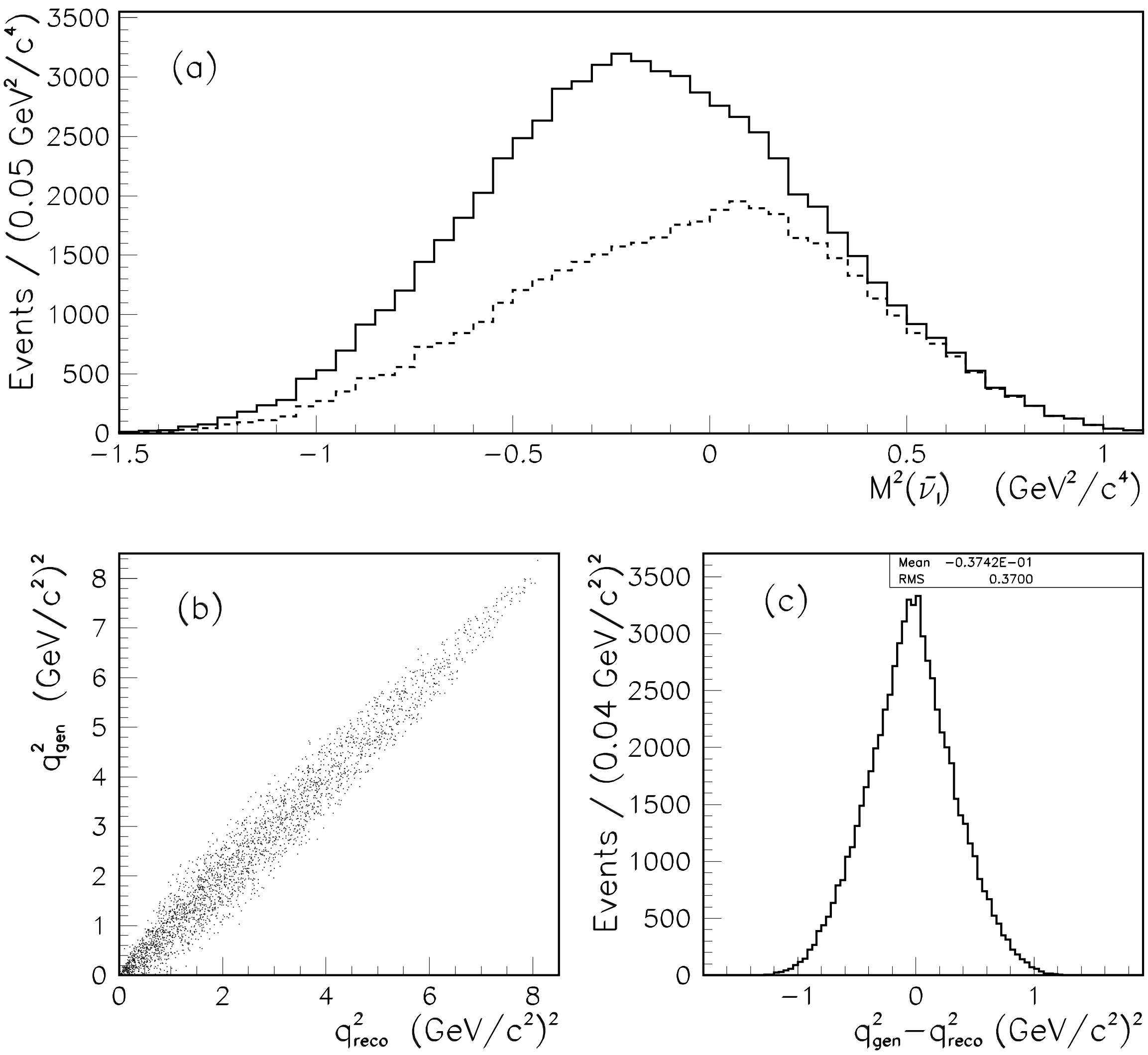


Figure 7.3: The distributions of the missing mass squared $M^2(\bar{\nu}_\ell)$ and q^2 for signal MC events. In (a), the missing mass squared $M^2(\bar{\nu}_\ell)$ of signal MC events for all values of q^2 is shown. The solid line is the $M^2(\bar{\nu}_\ell)$ distribution without the $\cos\theta_{B-D^*\ell}$ cut, and the dashed line the $M^2(\bar{\nu}_\ell)$ is after the $\cos\theta_{B-D^*\ell}$ cut is applied. In (b) and (c), the reconstructed and generated q^2 for signal MC events are shown.

reconstructed distribution for δM_J, along with the corresponding untagged contribution, are shown in Figures 7.4 and 7.5. As one can see, the untagged component is suppressed by selecting the best combination. Misreconstructed D_1^0 candidates do not peak in the signal region for the $K^-\pi^+$ mode; but due to soft photon combinatorics, there is some peaking for the $K^-\pi^+\pi^0$ mode, which is also reduced by selecting the best combination. Furthermore, there is always a better agreement between the generated and reconstructed widths of the narrow D_J^0 resonances in MC events after suppressing multiple combinations.

A study of the background in generic $B\bar{B}$ MC events allows us to identify the contributions to the background yield after the selection of the best $D_J^0\,\ell^-$ candidates. For the $D^0 \to K^-\pi^+$ mode, real D^{*+}s make up to 86% of the combinatoric background, while fake D^{*+}s contribute about 11%. For the $D^0 \to K^-\pi^+\pi^0$ mode, real D^{*+}s make up to 56% of the combinatoric background, while fake D^{*+} contribute about 42%. In both modes, 95% of the fake D^{*+} mesons come from fake D^0 mesons.

In our generic $B\bar{B}$ MC sample, no trace of a resonance in the δM_J spectrum was found for the second best candidates. The study of generic $B\bar{B}$ events shows no improvement in statistical significance of the signal over the background when we keep one candidate per event (but we know that multiple counting in an event leads to an underestimation of the statistical error). Nevertheless, the procedure of keeping the best candidate based on the calculation of the confidence level α improves our understanding of the combinatoric background, and therefore of the reconstructed efficiencies. The background shape of the δM_J distribution was checked with generic $B\bar{B}$ MC events. It turns out to be more symmetric around the world average $D_1^0 - D^{*+}$ mass difference when we suppress multiple entries in the event. As a result, the backgrounds for both modes have similar shapes for the best combination in each event, in contrast to the case with multiple combinations.

The level of multiple counting in data is reproduced in generic $B\bar{B}$ MC. Our MC study shows that about 42% of the events in the $K\pi$ mode and 46% of the events in the $K\pi\pi^0$ mode have two or more entries, which is in agreement with the observation in data. After sorting a sample of signal MC events with the method described above, the correct combination is kept 84% of the time for the $D^0 \to K^-\pi^+$

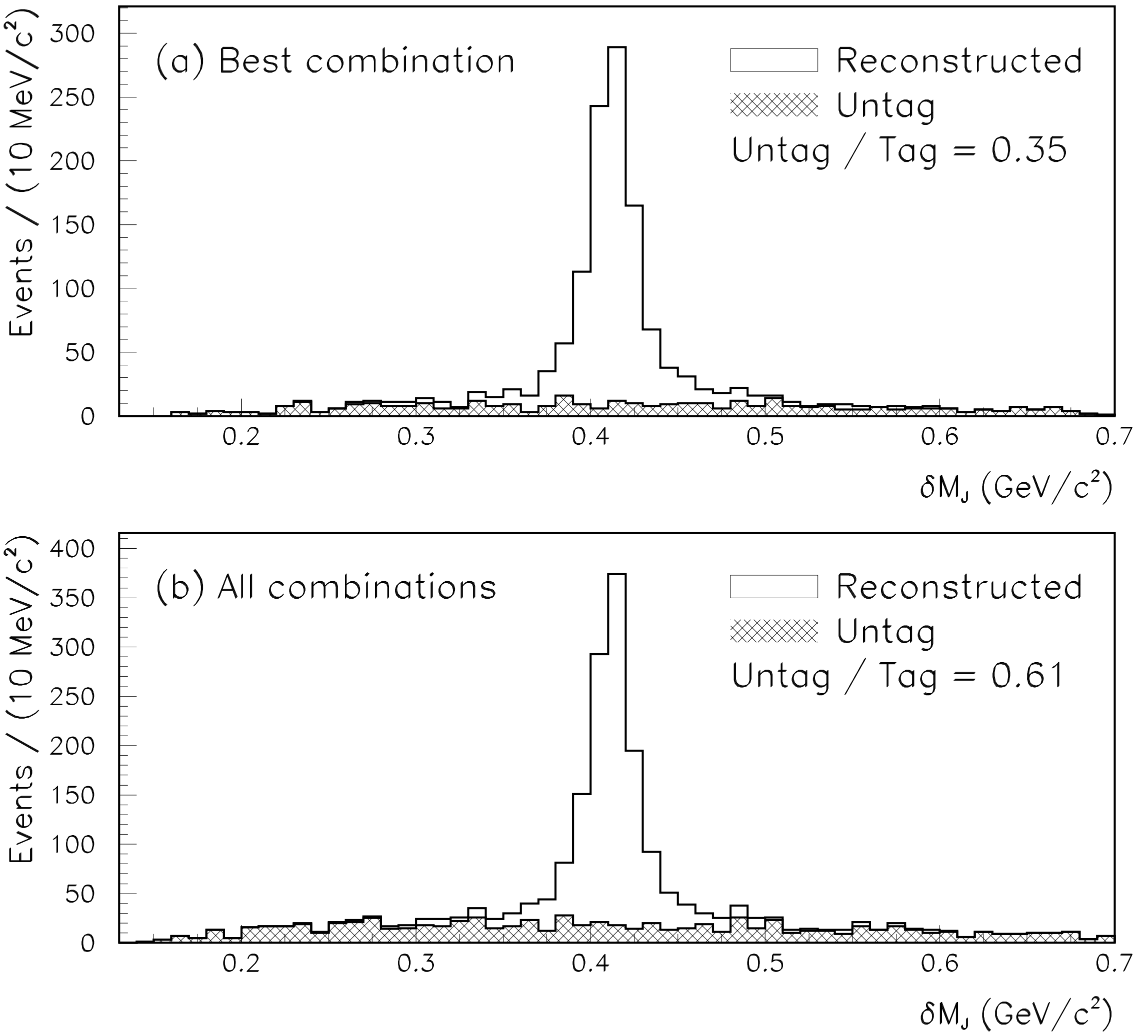


Figure 7.4: The δM_J distributions for (a) the best combination and (b) all combinations in signal MC events for the $D^0 \to K^-\pi^+$ mode. The hatched histogram overlaid on the reconstructed distribution is the untagged component.

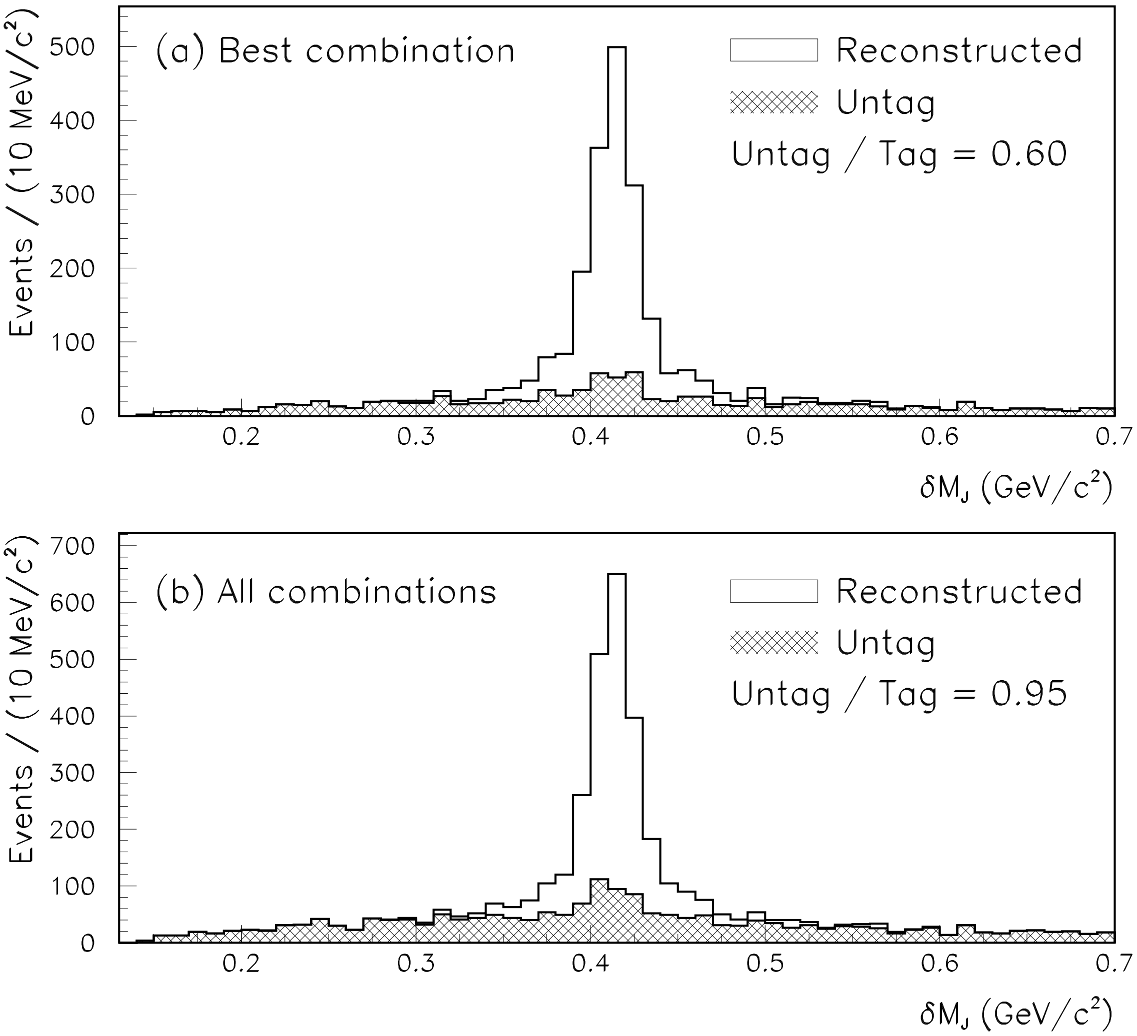


Figure 7.5: The δM_J distributions for (a) the best combination and (b) all combinations in signal MC events for the $D^0 \to K^-\pi^+\pi^0$ mode. The hatched histogram overlaid on the reconstructed distribution is the untagged component.

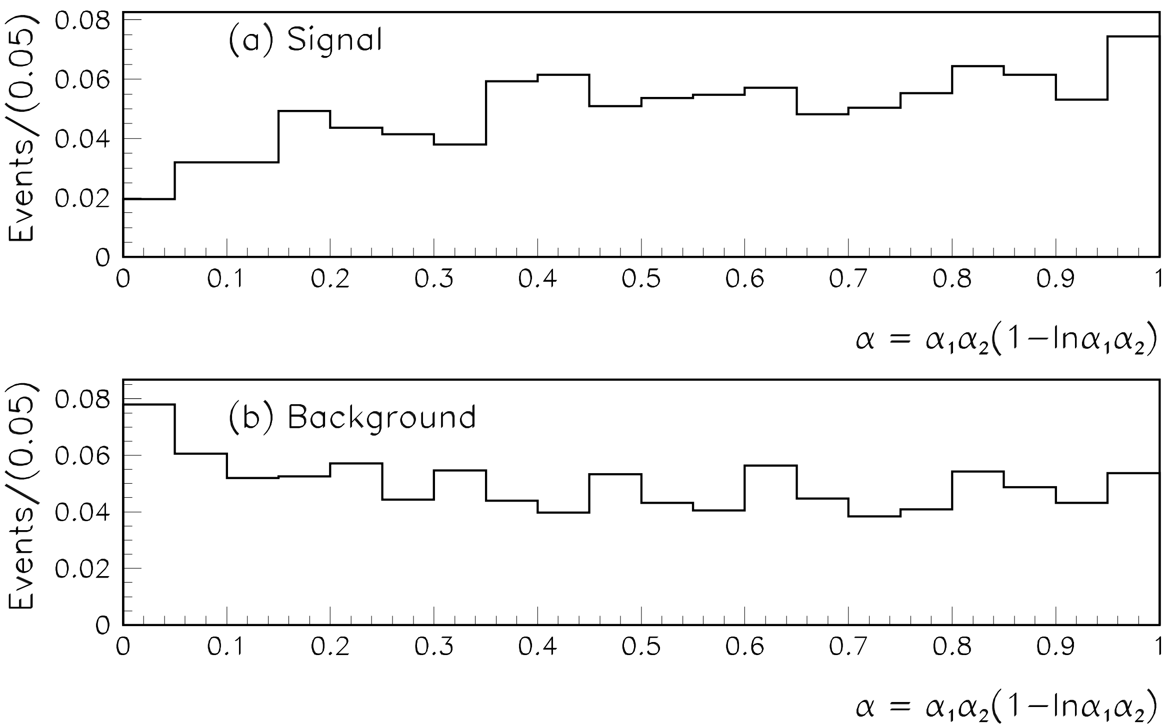


Figure 7.6: The estimator α for (a) tagged signal and (b) background events. The positive slope in (a) and the negative slope in (b) insures that the estimator α has some ability to separate signal and background.

mode and 80% of the time for the $D^0 \to K^-\pi^+\pi^0$ mode. Furthermore, for signal MC events, about 50% of the unwanted combinatoric background events (or untagged signal) are rejected. By keeping one candidate per event, we remove much of the combinatoric background events in the tail of the δM_J distribution. Figures 7.4 and 7.5 show the δM_J distributions for signal MC events in the $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^0$ modes before and after we require one candidate per event.

To check that the estimator for the best $D_J^0\,\ell^-$ candidates is not biased, we have computed the confidence level for signal and background MC events. Figure 7.6 illustrates that the estimator α has some ability to separate signal and background, although we do not use it for this purpose.

State	Mass (MeV/c^2)	Width (MeV/c^2)
D_1^0	2422.2 ± 1.8	$18.9\ ^{+4.6}_{-3.5}$
D_2^{*0}	2458.9 ± 2.0	23.0 ± 5.0

Table 7.1: Mass and width of the narrow D_J^0 states from PDG96 [6]. The central values are used in the fit as the Breit-Wigner parameters.

7.2 The Fitting Function

After applying all the requirements described above, a histogram of the mass difference δM_J is made for all the best $D_J^0\,\ell^-$ candidates. The δM_J distribution obtained by combining the two decay modes of the D^0 meson is shown in Figure 7.7. To address the problems caused by statistical fluctuation in small event samples and to make sure that the choice of bin size used for the histograms of δM_J does not bias our measurements, an unbinned likelihood fit, in which each event is weighted equally, is performed. The δM_J distribution is fit using a background function plus two signal functions corresponding to the lineshapes of the two narrow D_J^0 resonances with their respective masses and widths fixed [6].

The signal functions are two nonrelativistic Breit-Wigner resonance functions of the form:

$$y_{\mathrm{S}}(\delta M_J, \delta M_0, \Gamma_0) = \frac{\Gamma_0}{\pi} \frac{1}{(\delta M_J - \delta M_0)^2 + (\Gamma_0/2)^2}, \tag{7.11}$$

where δM_0 is the $D_J^0 - D^{*+}$ mass difference and Γ_0 is the width of the D_J^0 resonance, as previously determined from inclusive measurements [6]. The parameters of the Breit-Wigner resonance functions are shown in Table 7.1. Each Breit-Wigner function is convoluted with a Gaussian whose width accounts for detector resolution. Table 7.2 shows the estimated shifts in the mean and the resolutions for the two D^0 modes as found by Monte Carlo simulation. Although the MC results show no significant bias, we use the slight shift in mean values shown during the fit. When we combine the $K^-\pi^+$ and the $K^-\pi^+\pi^0$ modes, the values used for the bias and the resolution are 0.0 MeV/c^2 and 2.8 MeV/c^2 respectively.

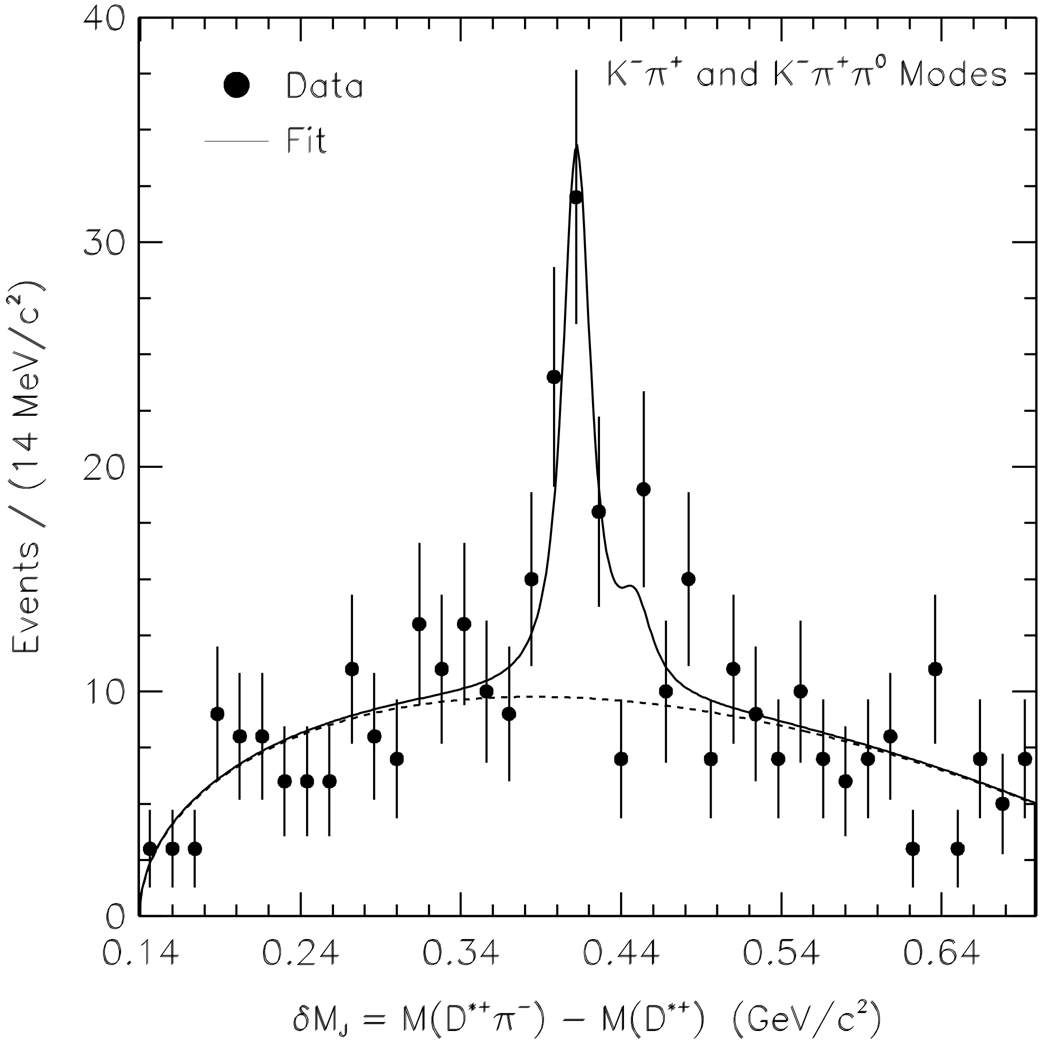


Figure 7.7: The δM_J distribution from the ON $\Upsilon(4S)$ Resonance data for $B^- \to D_1^0 \ell^- \bar{\nu}_\ell$ and $B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell$ ($\ell = e$ and μ) candidates obtained by combining both the D^0 decay modes. The dashed curve illustrates the background function, whereas the solid line shows the sum of the background and signal functions.

Mode	Bias (MeV/c^2)	Resolution (MeV/c^2)
$D^0 \to K^-\pi^+$	+0.10 $\pm$ 0.09	2.68 $\pm$ 0.12
$D^0 \to K^-\pi^+\pi^0$	-0.10 $\pm$ 0.17	2.90 $\pm$ 0.16

Table 7.2: Bias and resolution in the measurement of δM_J, as determined with MC signal events generated with a fixed D_J^0 mass.

The background is described by a smooth function parameterized using a polynomial with a threshold factor of the form

$$y_{\mathtt{B}}(\delta M_J) = a_1\sqrt{\delta M_J - m_{\pi^-}}\,\left[1 + a_2(\delta M_J - m_{\pi^-}) + a_3(\delta M_J - m_{\pi^-})^2\right], \qquad (7.12)$$

where the a_i are free parameters. Since there is no way to distinguish broad or non-resonant $B^- \to D^{*+}\pi^- X\,\ell^-\bar{\nu}_\ell$ contributions from background in the δM_J distribution alone, the background function incorporates both the combinatoric background and the possible contribution from broad and nonresonant $D^{*+}\pi^- X$ states.

To check the plausibility of the shape of the background function, we use a wrong-sign sample of $D^{*+}\pi^+\,\ell^-$ from data. This wrong-sign sample models random combinatorics of real and fake D^{*+} with random pions, which is indeed our main background. In Figure 7.8, the δM_J distributions from data are shown for the right-sign and wrong-sign samples, and the extracted wrong-sign and right-sign background functions are compared. This illustrates that the parameterization of the background function is reasonable.

In Figure 7.1 and Figure 7.2, the results from the unbinned likelihood fit for the $K\pi$ and the $K\pi\pi^0$ modes in data are overlaid on their respective δM_J distributions. For clarity, the δM_J distributions are presented as binned histograms. The corresponding yields are: 25.4 $\pm$ 7.8 D_1^0 candidates and 6.9 $\pm$ 6.6 D_2^{*0} candidates for the $K\pi$ mode; and, 31.5 $\pm$ 8.9 D_1^0 candidates and 3.9 $\pm$ 6.7 D_2^{*0} candidates for the $K\pi\pi^0$ mode.

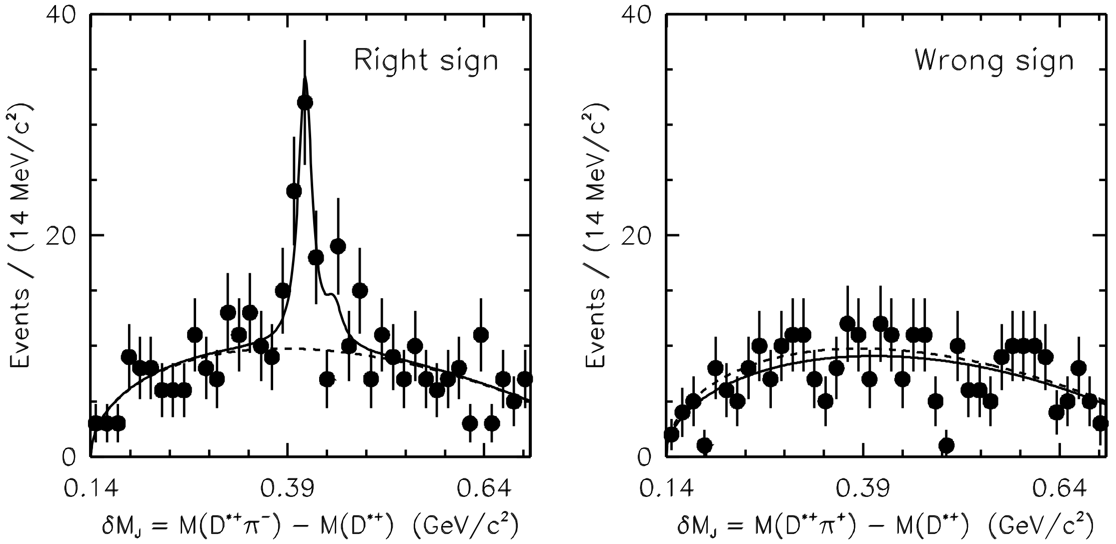


Figure 7.8: The right-sign and wrong-sign δM_J distributions from the ON $\Upsilon(4S)$ Resonance data for $B^- \to D_1^0 \ell^- \bar{\nu}_\ell$ and $B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell$ ($\ell = e$ and μ) candidates obtained by combining both the $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^0$ modes. In (a), the solid curve is the result of the fit to the right-sign data described in the text, while in (b), the solid curve is the result of the fit to the wrong-sign data in which only the background function was used. In (a) and (b), the dashed curve describes the right-sign background function.

7.3 Branching Fractions for $B^- \to D_J^0 \ell^- \bar{\nu}_\ell$

Because the branching fractions for $D_J^0 \to D^{*+}\pi^-$ have not yet been measured, we determine only the product of branching fractions for the narrow states $\mathcal{P}(D_J^0) = \mathcal{B}(B^- \to D_J^0 \ell^- \bar{\nu}_\ell)\, \mathcal{B}(D_J^0 \to D^{*+}\pi^-)$

The values for $\mathcal{P}(D_J^0)$ are obtained by dividing the net signal yields n_{D_J} by the total numbers of $B\bar{B}$ events in our data sample and the sum of the products of the efficiency, the D^{*+} branching fraction, and D^0 branching fraction for each mode used. The yields n_{D_J} are obtained from the δM_J fit after correcting for continuum and fake lepton backgrounds. The number of $B\bar{B}$ events is $N_{\Upsilon(4S)} = (3.29 \pm 0.07) \times 10^6$. In the next sections, the signal yields, the reconstruction efficiencies, and then the results for $\mathcal{P}(D_1^0)$ and $\mathcal{P}(D_2^{*0})$ are presented.

7.3.1 The $D_J^0\,\ell^-$ Yields

The δM_J distribution obtained by combining the two decay modes of the D^0 meson is shown in Figure 7.7. The fitted yields are 56.6 ± 11.9 events in the D_1^0 peak and 10.3 ± 9.4 in the D_2^{*0} peak. We have decided to use the combined fit to compute the branching fractions in order to reduce the systematic errors due to the uncertainties on the D_1^0 lineshape.

If the mass and the width of the D_1^0 resonance are allowed to float, the fitted values obtained are 2420 ± 4 MeV/c^2 and 23 ± 9 MeV/c^2 respectively. The area of the D_1^0 peak becomes 62.5 ± 16.7 and the area of the D_2^{*0} peak becomes 10.5 ± 9.8. The fitted mass and width agree well with the PDG96 averages listed in Table 7.1. Because the masses and widths of the narrow D_J^0 resonances are known from inclusive measurements, and because statistical fluctuations in the signal and in the background levels are a concern, the yields of the fit obtained with fixed δM_0 and Γ_0 are used for the calculation of the branching fractions.

To check that the data are consistent with the presence of a signal, we fit the δM_J distribution with only the smooth background function. The difference between the logarithm of the likelihood of the fit with the signal plus the background functions and the logarithm of the likelihood with only the background function is 18.7. Assuming Gaussian statistics, this corresponds to a 6.1σ statistical significance for the signal.

7.3.2 Continuum

To estimate the contamination of our sample by non-$B\bar{B}$ events, the same analysis is performed on continuum data (OFF Resonance). In subtracting this contribution, the continuum data must be scaled to account for the difference of luminosity and cross-section between the ON and OFF $\Upsilon(4S)$ data sets. Since the cross section for $e^+e^- \to q\bar{q}$ is proportional to the inverse of the squared of center-of-mass energy (s), the scale factor for the continuum data is given by:

$$\mathrm{f}_{q\bar{q}} = \frac{\mathcal{L}_{\mathrm{ON4S}}}{\mathcal{L}_{\mathrm{OFF4S}}}\frac{\bar{s}_{\mathrm{OFF4S}}}{\bar{s}_{\mathrm{ON4S}}} = 1.92\,. \tag{7.13}$$

The scaled continuum yields are: 2.3 ± 2.7 D_1^0 candidates and 1.5 ± 2.8 D_2^{*0} candidates when the $K\pi$ and $K\pi\pi^0$ modes are combined. The quoted errors on the yields are statistical only.

7.3.3 Fake Leptons

Fake leptons are hadrons misidentified as leptons and the fake rates are the probabilities of misidentification. To determine the average misidentification probabilities for hadrons in B decays, we need to know the individual misidentification probabilities for pions, kaons, and protons, and their relative abundances. The hadron abundances (Y_i with $i = \pi, K$, or p) are given in Appendix E. The misidentification probabilities for pions ($\mathcal{F}_\pi$), kaons ($\mathcal{F}_K$), and protons ($\mathcal{F}_p$) are taken from Reference [35] for electrons and from Appendix E for muons. The hadron abundances and the misidentification probabilities are generally momentum and charge dependent. The lepton fake rate is then expressed as follows

$$\mathcal{F}_\ell^\pm = \sum_{i=\pi,K,p} Y_i^\pm \times \mathcal{F}_i^\pm. \tag{7.14}$$

The number of fake leptons that combine with a D_J^0 in the event is estimated by performing the same analysis using tracks that are not leptons, and then scaling the yields by the misidentification lepton probability $\mathcal{F}_\ell$.

Fake Electrons

The fake rates for electrons is around 0.1% to 0.2%. After all of the analysis cuts, the number of fake electrons from misidentified hadrons is in fact consistent with zero. Converted photons are also a source of fake electrons. We estimate the number of converted photons with Monte Carlo simulation and find no significant contribution to the yield for fast electrons with momenta between 0.8 GeV/c and 2.0 GeV/c. Thus, no contribution from fake electrons (misidentified hadrons and converted photons) is subtracted from the fitted yields.

Fake Muons

Apart from random matches of noise hits in the muon chambers with extrapolated trajectories of particles seen in the tracking system, there are three main sources of fake muons from hadrons:

1. Sail through: A hadron coming from the CD which does not interact with the iron absorber.
2. Punch through: A hadron which interacts in the detector material with a reaction product penetrating to the muon chambers.
3. Real muon: A hadron which decays in flight into a muon; the later then continues to the muon counters.

The number of fake muons is estimated to be small but not entirely negligible. A full fake muon study is presented in Appendix E. After computing $\mathcal{F}_\mu$, we estimate the number of fakes to be: 0.8 ± 0.6 D_1^0 and 0.0 ± 0.3 D_2^{*0} for the $K\pi$ and $K\pi\pi^0$ modes combined. The quoted errors on the yields are statistical only.

7.3.4 Reconstruction Efficiencies

Our event selection efficiencies were obtained using Monte Carlo events generated with either $B^- \to D_1^0 \ell^- \bar{\nu}_\ell$ or $B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell$ and containing one signal decay per event. The Monte Carlo events were generated according to the ISGW2 model [51] and were passed through the GEANT [72] based simulation of the CLEO II detector. The entire procedure applied to the data was repeated on this Monte Carlo sample and the final signal yield divided by the number of events generated was interpreted as the efficiency. The δM_J distributions for signal Monte Carlo events are shown in Figure 7.9, where the fits shown are made following the same procedure as the fits to the ON resonance data. The efficiencies are given in Table 7.3. The quoted errors on the efficiencies are statistical only.

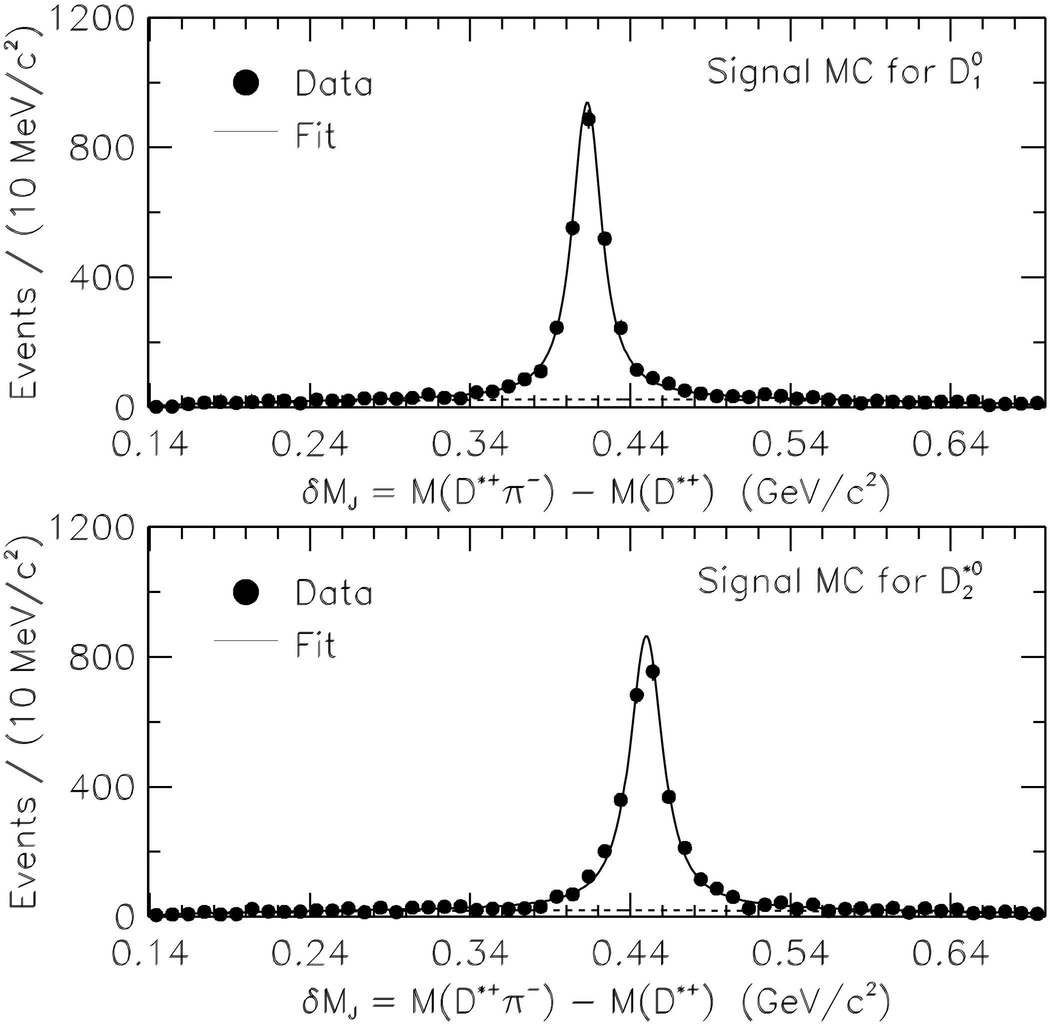


Figure 7.9: The δM_J distributions for signal MC events: (a) $B^- \to D_1^0 \ell^- \bar{\nu}_\ell$ and (b) $B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell$. The dashed curve describes the background function, whereas the solid line is the sum of the background and signal functions. The fitted yields were used to compute the reconstruction efficiencies ε_{D_1} and $\varepsilon_{D_2^*}$.

	D^0 Decay Mode	
Efficiency	$D^0 \to K^-\pi^+$	$D^0 \to K^-\pi^+\pi^0$
$\varepsilon_{D_1} \equiv \varepsilon(B^- \to D_1^0 \ell^- \bar{\nu}_\ell)$	(4.37 ± 0.09) %	(1.09 ± 0.02) %
$\varepsilon_{D_2^*} \equiv \varepsilon(B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell)$	(4.61 ± 0.09) %	(1.10 ± 0.02) %

Table 7.3: Reconstruction efficiencies ε_{D_J} for $B^- \to D_1^0 \ell^- \bar{\nu}_\ell$ and $B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell$ decays for each of the D^0 decay modes used in the reconstruction ($\ell = e$ and μ).

7.3.5 Results

As mentioned in Section 7.3.1, the ON resonance D_1^0 and D_2^{*0} yields obtained from the fit are 56.6 ± 11.9 and 10.3 ± 9.4 respectively. The sum of the continuum and fake lepton backgrounds are subtracted from the ON Resonance yields as indicated in Table 7.4. This leads to final yields of $n_{D_1} = 53.5 \pm 12.2$ and $n_{D_2^*} = 8.8 \pm 9.8$. From the final yields and the reconstruction efficiencies, the product branching fractions are then computed using:

$$\mathcal{P}(D_J^0) = \frac{n_{D_J}/\varepsilon_{D_J}}{4\, N_{\Upsilon(4S)}\, f_{+-}\, \mathcal{B}(D^{*+} \to D^0\pi^+)\, \mathcal{B}(D^0 \to K^-\pi^+(\pi^0))}. \quad (7.15)$$

The values of the D^{*+} and D^0 branching fractions that we use are [6]:

$$\begin{aligned} \mathcal{B}(D^{*+} \to D^0\pi^+) &= (68.3 \pm 1.4)\%, & (7.16)\\ \mathcal{B}(D^0 \to K^-\pi^+) &= (3.83 \pm 0.12)\%, & (7.17)\\ \mathcal{B}(D^0 \to K^-\pi^+\pi^0) &= (13.9 \pm 0.9)\%. & (7.18) \end{aligned}$$

With the assumption that the branching fractions of $\Upsilon(4S)$ to charged and neutral $B\bar{B}$ pairs are $f_{+-} = f_{00} = 0.5$, the product branching fractions are obtained from the yields summed over the two D^0 modes and the appropriate sum of efficiencies times branching fractions:

$$\begin{aligned} \mathcal{P}(D_1^0) &= (0.373 \pm 0.085)\,\%, & (7.19)\\ \mathcal{P}(D_2^{*0}) &= (0.059 \pm 0.066)\,\%, & (7.20) \end{aligned}$$

where the errors are statistical only. The experimental systematic and theoretical uncertainties on these results are presented in the next sections.

	D_1^0	D_2^{*0}
ON Resonance Yield	56.6 ± 11.9	10.3 ± 9.4
Subtracted Yield	3.1 ± 2.8	1.5 ± 2.8
Final Yield n_{D_J}	53.5 ± 12.2	8.8 ± 9.8
$\mathcal{P}(D_J^0)$	$(0.373 \pm 0.085)\,\%$	$(0.059 \pm 0.066)\,\%$

Table 7.4: Yields and product branching fractions. The "Final Yields" are given after the subtraction of the OFF Resonance and fake lepton from the ON Resonance fits. The error on the yields and on the product branching fractions is statistical only.

7.3.6 Systematic Uncertainties

In this section, we estimate the impact of systematic errors on our measurements of $\mathcal{P}(D_1^0)$ and $\mathcal{P}(D_2^{*0})$. A list of the sources of systematic uncertainties for $\mathcal{P}(D_1^0)$ and $\mathcal{P}(D_2^{*0})$ is given in Table 7.5. The total systematic uncertainties on $\mathcal{P}(D_1^0)$ and $\mathcal{P}(D_2^{*0})$ are 14.0% and 17.3% respectively. We added the contributions in quadrature for the total systematic errors. Details on the estimates of the systematic uncertainties are presented below.

Fitting Function

To estimate the uncertainties from the input mass and width of the D_1^0 resonance, the PDG96 values are varied within their uncertainties [6]. This leads to systematic errors on the yield of 1% due to the mass and 10% due to the width. The uncertainty on the background level is determined by varying the shape of the background function. We used alternative functions of the form

$$\begin{aligned} y_{\mathsf{B}}(\delta M_J) &= a_1\sqrt{\delta M_J - m_{\pi^-}}\left[1 + \sum_{i=2}^{4} a_i(\delta M_J - m_{\pi^-})^{i-1}\right], \\ y_{\mathsf{B}}(\delta M_J) &= a_1\sqrt{\delta M_J - m_{\pi^-}}\ \exp[a_2(\delta M_J - m_{\pi^-})], \\ y_{\mathsf{B}}(\delta M_J) &= a_1\,(\delta M_J - m_{\pi^-})^{a_2}\ \exp[a_3(\delta M_J - m_{\pi^-})] \end{aligned}$$

in repeating the fit. This leads to a systematic error of 4% on the yield for the D_1^0 meson.

Source of Systematic Error	$\mathcal{P}(D_1^0)$	$\mathcal{P}(D_2^{*0})$
M_{D_J}	1.0%	1.1%
Γ_{D_J}	10.0%	14.0%
Background Function	4.0%	5.0%
Uncorrelated Background	0.5%	0.4%
Lepton Fake	1.0%	1.0%
Lepton ID	1.3%	1.3%
MC Statistics	1.5%	1.5%
$\mathcal{B}(D^{*+} \to D^0\pi^+)$	2.0%	2.0%
$\mathcal{B}(D^0 \to K^-\pi^+(\pi^0))$	3.5%	3.5%
Tracking Efficiency	4.0%	4.0%
Slow π Efficiency	5.0%	5.0%
π^0 Reconstruction	2.4%	2.4%
Dalitz Weight	1.9%	1.9%
Multiple Counting	1.4%	1.4%
Hadron Identification	1.0%	1.0%
$B\bar{B}$ Cross-Section	2.0%	2.0%
Total	14.0%	17.3%

Table 7.5: Experimental systematic errors on the product branching fractions $\mathcal{P}(D_1^0)$ and $\mathcal{P}(D_2^{*0})$. Tracking uncertainties are for all charged particles other than the slow pion.

It is much harder to quote a fitting systematic error for the D_2^{*0} because there is no significant contribution from $B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell$ in our data sample. Thus, the fitting uncertainties on the mass and the width for the D_2^{*0} are obtained by scaling the systematic errors of the D_1^0 by the ratio of the experimental uncertainties on the lineshape of the D_2^{*0} to that of the D_1^0. This leads to a 1.1% uncertainty contribution due to the mass and a 14% uncertainty due to the width. For the uncertainty on the background level, we find a variation of 5% on the D_2^{*0} yield when we change the shape of the background function.

Another source of uncertainty on the fitting function is the form of the Breit-Wigner resonance function. When D-wave relativistic Breit-Wigner resonance functions are used to describe the narrow D_J^0 resonances, the fitted yields change by about 1%, which is negligible compared to the uncertainties on the D_J^0 widths.

Correlated Background

We are not able to subtract contributions from processes such as $\bar{B} \to D_J^0 X \ell^- \bar{\nu}_\ell$ decays. The $\cos\theta_{B-D_J\ell}$ cut is effective in removing correlated backgrounds, but not 100% efficient. Fortunately, these higher multiplicity semileptonic $\bar{B}$ decays are predicted to be tiny. First, the rates $B^- \to D^{(*)\prime}\ell^-\bar{\nu}_\ell$ are expected to be small [42, 51, 55]; and moreover, $D^{(*)\prime} \to (D_1^0 \text{ or } D_2^{*0})X$ is believed to be suppressed [91]. We therefore neglect such background on the basis that processes which contribute to the correlated background are small.

Uncorrelated Background

To study uncorrelated background, we used our generic $B\bar{B}$ MC. The main source of uncorrelated background is when

$$\begin{aligned}
B^- &\longrightarrow D_J^0\,\pi^- \\
&\quad\hookrightarrow D^{*+}\pi^- \\
&\quad\quad\hookrightarrow D^0\pi^+ \\
&\quad\quad\quad\hookrightarrow D^0 \to K^-\pi^+ \text{ or } D^0 \to K^-\pi^+\pi^0
\end{aligned}$$

$$\begin{array}{l} B^{+} \longrightarrow X \\ \qquad\quad \hookrightarrow Y\,\ell^{-} \end{array}$$

Based on recent CLEO measurements [105] of $B^- \to D_1^0\pi^-$ and $B^- \to D_2^{*0}\pi^-$; we assumed $\mathcal{B}(B^- \to D_1^0\pi^-) = 0.12\%$ and $\mathcal{B}(B^- \to D_2^{*0}\pi^-) = 0.21\%$. We find negligible contributions from such processes to the $D_J^0\,\ell^-$ yields in our generic $B\bar{B}$ MC. Therefore, we do not subtract any contribution from uncorrelated background. Our insensitivity to uncorrelated backgrounds is mainly due to the lower limit on the lepton momentum: $|\mathbf{p}_e| > 0.8$ GeV/c and $|\mathbf{p}_\mu| > 1.0$ GeV/c. We nevertheless quote a small systematic uncertainty of 0.5% and 0.4% on $\mathcal{P}(D_1^0)$ and $\mathcal{P}(D_2^{*0})$ respectively.

Lepton Identification and Fake Rates

The systematic errors associated with the lepton detection efficiency are taken from a CLEO study [98]; these are 2% for electrons and 1% for muons respectively. The uncertainty on electron identification is estimated from the difference in the efficiency for finding electrons from radiative Bhabha events before and after they are embedded in hadronic events. Muon identification is modeled quite well at CLEO II. The program CLEOG has been tuned to simulate the complicated geometry and chamber response with great accuracy, achieving 0.5% (1.0%) precision in the barrel (endcap) region [106].

We estimated the systematic uncertainty on the fake electrons (hadron fakes and converted photons) to be 0.2%. To be conservative, the electron fake and the converted photon rates were doubled in the calculation of the electron fake uncertainty. The systematic uncertainty on the muon fake rates is larger than that for electrons. Based on the study described in Appendix E, a systematic error of 25% should be used. We attribute a very conservative systematic error of 2% on the fake muon yields, which corresponds to a ±50% systematic uncertainty on the muon misidentification probabilities.

When the electron and muon modes are combined, the uncertainty on the lepton detection efficiency becomes 1.3% and the uncertainty on the fake lepton rates becomes 1.0%.

MC Statistics

The statistical errors on the efficiencies listed in Table 7.3 are propagated in the calculation of the yields. This leads to a 1.5% uncertainty on the branching fractions.

D^{*+} and D^0 Branching Fractions

The experimental errors on the $D^{*+} \rightarrow D^0\pi^+$ and $D^0 \rightarrow K^-\pi^+(\pi^0)$ branching fractions lead respectively to a 2% and 3.5% uncertainty on the yields of the narrow D_J^0 mesons. The experimental errors on the D^{*+} and D^0 branching fractions are given in Equations (7.16)-(7.18).

Detector Efficiency

The Monte Carlo simulation of the passage of particles from the interaction point out through the passive and active detector elements contributes to the modeling of the reconstruction efficiencies for $B^- \rightarrow D_1^0\ell^-\bar{\nu}_\ell$ and $B^- \rightarrow D_2^{*0}\ell^-\bar{\nu}_\ell$. The program CLEOG is tuned to reproduce the measured efficiency for each data set. This procedure takes into account effects such as fluctuations in running conditions and detector aging. We divide the systematic uncertainty on the MC detector reconstruction efficiency into three categories: track, slow pion, and neutral pion reconstruction.

The track selection criteria described in Section 6.5 did not include the slow pion selection criteria. Thus, the tracking uncertainty described here applies to all charged particles other than the slow pion. Based on numerous tracking studies carried by the tracking group and various individuals in CLEO [107, 108], the uncertainty on the track reconstruction efficiency is taken to be 1% per track. Our signal events contains a total of four charged tracks and one slow track. Thus, we attribute a fractional error of 4% to the tracking uncertainty.

Since the slow pions have very low momentum near the edge of acceptance, care must be taken to determine their detection efficiency separately. In the high magnetic field of CLEO II, the slow pions are often restricted to the inner tracking chambers and are subject to curl. The slow pion spectrum in data and MC is shown in Figure 7.10 for the D^{*+} candidates. Extensive studies [95, 109] have been performed

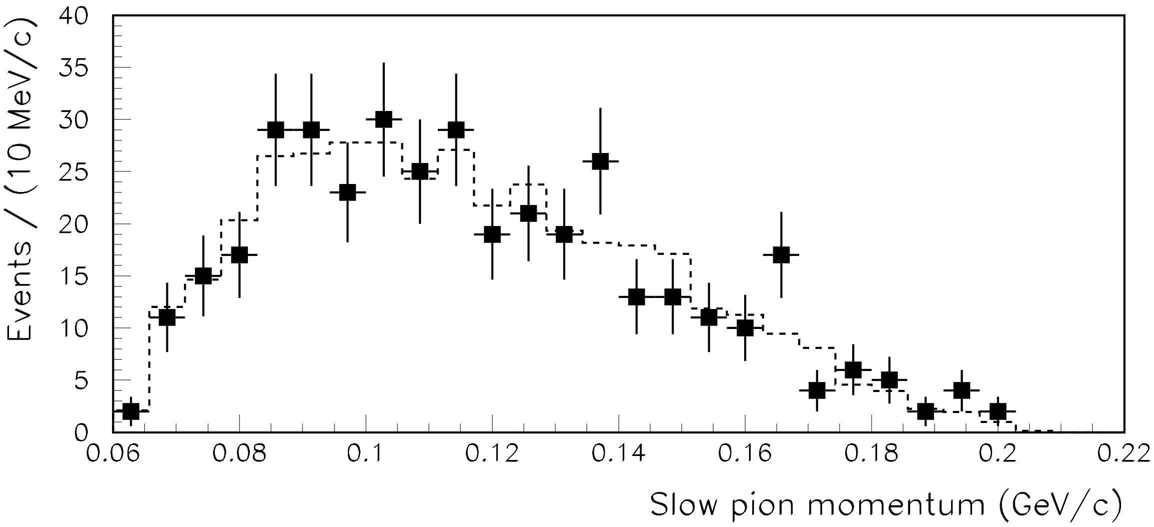


Figure 7.10: The slow pion momentum spectrum in data and MC. The dashed line is the MC and the data points are indicated with the solid squares. Both spectra are normalized to equal areas.

to investigate the uncertainty on the reconstruction efficiencies due to slow tracks. The slow pion reconstruction efficiency has been measured in both data and MC. The agreement between data and MC suggests a 5% systematic error on the slow pion reconstruction efficiency.

The Dalitz weight cut naturally puts a lower energy bound of 250 MeV on the neutral pion candidates. Past studies [107, 109] indicate that the MC simulation is accurate to within 5% in estimating the absolute efficiency for π^0 reconstruction. When the $K^-\pi^+$ and $K^-\pi^+\pi^0$ modes are combined, the 5% uncertainty on the neutral pion reconstruction implies a 2.4% uncertainty on the branching fractions. We have investigated the impact of varying the photon and the π^0 energy thresholds in the range of 30 MeV to 50 MeV and 250 MeV to 350 MeV respectively. This leads to a fluctuation on the branching fractions of less than 2%. To be conservative, the higher systematic uncertainty of 2.4% is used.

Dalitz Weight

To estimate the systematic uncertainty on the Dalitz weight cut described earlier, we compute the efficiency according to E691 [103], E687 [110], and Mark III [111] measurements of the Dalitz decay $D^0 \to K^-\pi^+\pi^0$. We quote 1.9% as the Dalitz systematic uncertainty, which is the largest variation in efficiency among E691, E687, and Mark III.

Multiple Counting

The good match between data and MC on the absolute level of multiple counting (see Section 7.1) suggests a small systematic error on the selection of the best $D_J^0\,\ell^-$ candidate. We estimated the systematic uncertainty associated with the multiple counting to be 1.4%. Recent measurements at CLEO confirmed the good agreement between data and MC on the charged multiplicity in B semileptonic decays [112].

Hadron Identification

The hadron identification procedure described in Section 6.7 is expected to be well modeled in the GEANT based simulation of the CLEO II detector. In the analysis, we require that the measured dE/dx value for the pions (kaons) be consistent with the expected value within 3.0 (2.5) standard deviations. These selection criteria are intentionally made loose to limit the systematic effects in the modeling of the efficiencies in the MC. It has determined that the ratio of the efficiency per track in data and in MC is one with good precision [113]. We therefore quote a 1% systematic uncertainty for the hadron identification.

$B\bar{B}$ Cross-Section

The CESR luminosity and the $B\bar{B}$ cross-section play a crucial role in the present analysis. The CLEO collaboration has presented [114] a careful determination of the CESR luminosity. A precise measurement of the number of $\Upsilon(4S) \to B\bar{B}$ events is needed for the extraction of B meson branching fractions. At CLEO, the number $N_{\Upsilon(4S)}$ is determined from the measurement of the $B\bar{B}$ cross-section and the integrated luminosity. A 2% systematic uncertainty is assigned to $N_{\Upsilon(4S)}$. This normalization error is dominated by the run-to-run variation of the $B\bar{B}$ cross-section within each data set [115].

7.3.7 Model Dependence

The reconstruction efficiency is based on the CLEO II acceptance and on the selection criteria used to optimize the statistical significance of the signal. The reconstruction efficiency is then sensitive to: the lepton efficiency, which depends on the shape of the lepton energy spectrum; and the D_J^0 efficiency, which depends mostly on the detection efficiency for the slow pion. The slopes of the form factors are strongly correlated with q^2, and thus with the lepton and slow pion momenta. Another less intuitive dependence of the efficiency on the Monte Carlo simulation is the angular distribution of the decay. The efficiencies of the $\cos_{B-D_J\ell}$ and $\cos_{B-D^*\ell}$ angular cuts depends on the angular correlation between the decay products of the B meson. The cut on $\cos_{B-D_J\ell}$ is based on kinematics, and is not sensitive to the model used to describe $B^- \to D_J^0 \ell^- \bar{\nu}_\ell$. On the other hand, the $\cos_{B-D^*\ell}$ distribution is strongly correlated with the D_J^0 polarization and the modeling of the B decay. In summary, any variation of the model will affect the lepton energy spectrum, the D_J^0 energy spectrum, the angular correlation of the decay products, and therefore the overall efficiency to reconstruct $B^- \to D_J^0 \ell^- \bar{\nu}_\ell$.

It is then obvious that the calculation of $\mathcal{P}(D_1^0)$ and $\mathcal{P}(D_2^{*0})$ depends on the model used to compute the efficiencies. As mentioned earlier, we used the ISGW2 model for this task. In our MC simulation, we assume the D_1 and the D_2^* decay via pure D-wave with helicity angle distributions proportional to $1 + 3\cos^2\alpha$ and $\sin^2\alpha$ respectively. The systematic error on the yields due to the uncertainty on the helicity angle α is neglected. According to Figure 3.4 and some more recent CLEO measurements [105], our assumption on the forms of the helicity distributions of the D_1 and the D_2^* in the MC is adequate. Any possible mixing between the D_1 and the D_1^* is also not considered. Hence, the model dependence presented here only deals with the theoretical prediction underlying the semileptonic decay of the B meson to

P-wave charm mesons.

A detailed model dependence study would require extensive coding of all the theoretical models described in Chapter 3 in the MC generator. We took a different approach. We intend to rely on HQET and HQS for the description on the dynamics of $B^- \to D_J^0 \ell^- \bar{\nu}_\ell$. ISGW2 incorporates most of the phenomenology of HQET. In fact, the ISGW2 model is presently the state-of-the-art model in describing form factors for semileptonic meson decays. It provides a very good match to most of the experimental measurements in charm and bottom semileptonic decays [51]. Nevertheless, the ISGW2 has a number of free parameters that one can vary. The theoretical uncertainties associated with the model dependence of the efficiency can be obtained by varying the parameters and the form factors used in the ISGW2 model. Therefore, as described in more detail below, we studied the model dependence of the reconstruction efficiencies by:

1. Varying the β_B and β_X parameters.

2. Varying the slope of the form factors (namely r^2_{BX}).

3. Changing the relative strength of the form factors predicted by ISGW2.

Certain exclusive models, like the Körner and Schuler (KS) model [116], and the Bauer, Stech, and Wirbel (BSW) model [117] give predictions for the q^2 dependence of the form factors for $\bar{B} \to D^{(*)}\ell\bar{\nu}_\ell$, but do not give any prediction for $\bar{B} \to D^{**}\ell\bar{\nu}_\ell$. Those models extrapolate the q^2 dependence of the form factors from $q^2_{\rm min}$, which is less reliable according to HQET. For these reasons, the BS and BSW models were not considered in this analysis.

The ACCMM free quark model also gives predictions for the lepton energy spectrum for $b \to c\,\ell\bar{\nu}_\ell$. As described below, we computed the reconstruction efficiencies with the ACCMM model as a consistency check.

ISGW2

In ISGW2, the parameters β_B and β_X are used to describe the quark position wave function. A change of the parameters β_X and β_B is equivalent to a variation of the wave functions which describe the heavy meson states. A variation of $\pm 25\%$ of β_B and β_X leads to a systematic uncertainty of 2% on the reconstruction efficiencies.

Heavy Quark Symmetry determines various aspects of the behavior of the form factors. The form factors used to describe the hadronic currents in ISGW2 are related to the universal function F_5^α of Equation (3.26). For $\bar{B} \to X_{c\bar{q}}\ell^-\bar{\nu}_\ell$, where $X_{c\bar{q}}$ is a D, a D^*, a D_1, or a D_2^* meson, $r_{BX}^2 \simeq 0.2$. Varying r_{BX}^2 between 0.14 and 0.40 leads to a systematic uncertainty of 2% – 3% on the efficiency[2].

It is surprising that a large variation of the slope did not affect our efficiency by more than 2% – 3%. Changing the slope of the form factors mainly affects the magnitude of the lepton and the slow pion momenta. A faster lepton implies a softer slow pion, and vice versa. It turns out that by changing ρ^2, the overall efficiency stays approximately constant because the event efficiency decreases with the momentum of the lepton and the slow pion efficiency increases with its momentum. Figures 7.11 illustrates the event efficiency as a function of the lepton momentum. After a closer study, we found that the reconstruction efficiency is much more sensitive to the polarization of the D_J^0, and thus the angular distribution of the semileptonic B decay.

The sensitivity of the reconstruction efficiency to the polarization of the D_J^0 meson is due to the requirement $\cos_{B-D^*\ell} < -1$, which rejects the largest background, namely $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$. The angular correlation between the D_J^0 and the D^{*+} mesons directly affects the value of $\cos_{B-D^*\ell}$. Figure 7.11 shows the efficiency for the $B^- \to D_1^0\ell^-\bar{\nu}_\ell$ signal MC events as a function of the lepton momentum, before and after the $\cos_{B-D^*\ell}$ cut has been applied. One should notice that when we select candidates with $\cos_{B-D^*\ell} < -1$, the reconstruction efficiency decreases as the lepton momentum increases. Near $|\mathbf{p}_\ell|_{\max}$, $|\mathbf{p}_{\bar{\nu}_\ell}| \to 0$ and thus $\cos_{B-D_J\ell} \to 1$. This also implies that $\cos_{B-D^*\ell} \to 1$ since $q^2 \to 0$ as $|\mathbf{p}_\ell| \to |\mathbf{p}_\ell|_{\max}$. For high lepton momentum, the $\cos_{B-D^*\ell}$ distribution tends to peak toward positive values, and more signal events are discarded.

[2] A value of $r_{BX}^2 \simeq 0.2$ for $\bar{B} \to D^*\ell\bar{\nu}_\ell$ corresponds to $\rho^2 \simeq 0.74$. Note that ρ^2 is not $\rho_{A_1}^2$ measured by CLEO. In ISGW2, $\rho_{A_1}^2 \simeq 0.53$ [118], which is lower than the measured value $\rho_{A_1}^2 = 0.92 \pm 0.12 \pm 0.06$ [48]. The variation of r_{BX}^2 described above corresponds to a variation of ρ^2 between 0.52 and 1.48, or a variation of $\rho_{A_1}^2$ between 0.38 and 1.06 for $\bar{B} \to D^*\ell\bar{\nu}_\ell$.

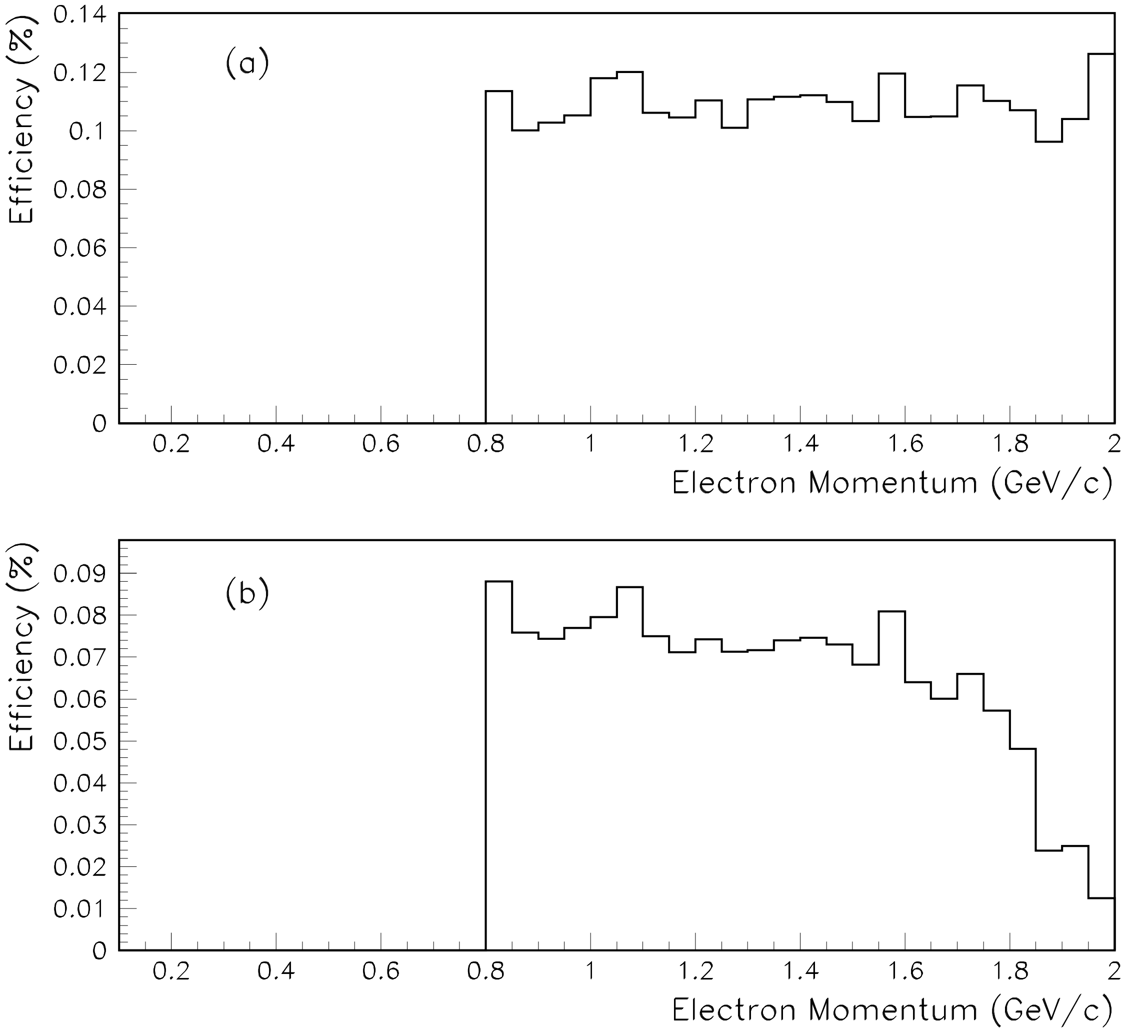


Figure 7.11: The event efficiency for $B^- \to D_1^0 e^- \bar{\nu}_e$ ($K\pi$ mode) as a function of the electron momentum: (a) before the cut $\cos\theta_{B-D^*\ell}$ is applied, and (b) after the cut $\cos\theta_{B-D^*\ell}$ is applied. All the other selection criteria were applied on the $D_J^0\,\ell^-$ candidates in (a) and (b). One can see that for the $D^0 \to K^-\pi^+$ mode the reconstruction efficiency is around 7% above the electron acceptance threshold of 0.8 GeV/c used in this analysis. The reconstruction efficiency for the $D^0 \to K^-\pi^+\pi^0$ mode is about a factor of four smaller. The $K^-\pi^+$ and $K^-\pi^+\pi^0$ yields are nevertheless comparable because $\mathcal{B}(D^0 \to K^-\pi^+\pi^0) \simeq 3.5\,\mathcal{B}(D^0 \to K^-\pi^+)$.

In ISGW2, the D_J^0 decay angle distributions are predicted to be:

$$dN/d\cos\theta_J = \begin{cases} 0.32 + 0.55\cos^2\theta_J & (D_1) \\ 0.20 + 1.95\cos^2\theta_J - 1.75\cos^4\theta_J & (D_2^*) \end{cases} \tag{7.21}$$

The polarization of the D_J^0 is entirely constrained by the relative strength and q^2 dependence of the form factors. An extreme approach would be to generate signal Monte Carlo with a flat D_J^0 decay angle distribution and see how it affects the event efficiency. By doing so, we found that our efficiency changed by 10%. Since the polarization is correlated to the relative strength of the from factors, a more reasonable approach is to vary $\bar{A}_{\rm POL}$ and $\bar{A}_{\rm FB}$, and recompute the event efficiency. By varying the relative weight of the form factors and by varying the form factor ratios, one can induce a variation on $\bar{A}_{\rm POL}$ and $\bar{A}_{\rm FB}$. A variation of $\pm 25\%$ on $\bar{A}_{\rm POL}$ and $\bar{A}_{\rm FB}$ leads to a systematic uncertainty of 5% – 6% on the reconstruction efficiency[3].

ACCMM

In the CLEO implementation of the fragmentation for the ACCMM model, the D_J^0 meson from the decay $B^- \to D_J^0 \ell^- \bar{\nu}_\ell$ is unpolarized. In other words, the decay angle distribution follows

$$dN/d\cos\theta_J = 1/2\,. \tag{7.22}$$

When the ACCMM model is used to compute the reconstruction efficiencies, the net yields for n_{D_J} shift by 12%. The model parameters used are extracted from the inclusive single-lepton analysis performed by CLEO [35]: $p_F = 265$ MeV/c, $m_c = 1.670$ GeV/c^2, and $m_{\rm sp} = 150$ MeV/c^2. We did not study the effect of varying p_F and $m_{\rm sp}$. As one can see, the predictions of ACCMM and ISGW2 for an unpolarized D_J^0 meson are quite similar.

[3]Following the notation of Reference [3], one can define the form factor ratios R_1 and R_2 for $\bar{B} \to D^* \ell \bar{\nu}_\ell$. The observable $\bar{A}_{\rm POL}$ and $\bar{A}_{\rm FB}$ are very sensitive to R_1 and R_2 in $\bar{B} \to D^* \ell \bar{\nu}_\ell$. CLEO measured $R_1 = 1.24 \pm 0.26 \pm 0.12$ and $R_2 = 0.72 \pm 0.18 \pm 0.07$ [48], which is in good agreement with the ISGW2 predictions: $R_1 \simeq 1.26$ and $R_2 \simeq 1.03$.

Model Dependence – Summary

The total systematic uncertainty on the theoretical model is taken to be 6.5%. We added in quadrature the effects due to the variation of the βs, the slope of the form factors, and the D_J^0 polarization in ISGW2. Table 7.6 summarizes the model dependence uncertainties for each lepton and D^0 decay mode. The reconstruction efficiencies obtained with the ACCMM and the ISGW2 models for an unpolarized D_J^0 agree within 2%. Both models predict an increase of about 10% in the reconstruction efficiencies when the angular correlation between the D_J^0 and the W^- are not properly taken care of in the decay $B^- \to D_J^0 W^-$. The reasonable agreement between ACCMM and ISGW2 in this particular case gives us confidence that a model uncertainty of 6.5% is reasonable.

Source	$K\pi\ e$	$K\pi\ \mu$	$K\pi\pi^0\ e$	$K\pi\pi^0\ \mu$
β_X and β_B	2%	2%	2%	2%
Slope of the Form Factors	2%	3%	2%	3%
Polarization of the D_J^0	5%	5%	6%	6%
Total	6%	6%	7%	7%

Table 7.6: Systematic errors on $\mathcal{P}(D_1^0)$ and $\mathcal{P}(D_2^{*0})$ associated with the variation of the ISGW2 parameters used to compute the efficiencies. The model dependence for the $K\pi\pi^0$ mode is slightly higher because of the momentum cut on the D meson candidate ($|\mathbf{p}_D| > 0.8$ GeV/c).

7.3.8 Checking the Measurements

As a consistency check, the product branching fractions were also computed separately for each D^0 mode in the data. Yields and branching fractions are obtained by fitting the $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^0$ data separately (see Figures 7.1 and 7.2). The product branching fractions obtained by averaging the results from separate fits agree very well with the results of the simultaneous fit to both modes. The tiny difference in the fitted yields can be attributed to expected statistical variation,

$\mathcal{B}(B^- \to D_1^0 \ell^- \bar{\nu}_\ell)\, \mathcal{B}(D_1^0 \to D^{*+}\pi^-)$		
	$K^-\pi^+$	$K^-\pi^+\pi^0$
ON Resonance Yield	25.4 ± 7.8	31.5 ± 8.9
Subtracted Yield	1.8 ± 2.0	1.3 ± 1.9
Final Yield n_{D_1}	23.6 ± 8.1	30.2 ± 9.1
$\mathcal{P}(D_1^0)$	(0.314 ± 0.108) %	(0.444 ± 0.134) %
Average $\mathcal{P}(D_1^0)$	(0.365 ± 0.084) %	

$\mathcal{B}(B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell)\, \mathcal{B}(D_2^{*0} \to D^{*+}\pi^-)$		
	$K^-\pi^+$	$K^-\pi^+\pi^0$
ON Resonance Yield	6.9 ± 6.6	3.9 ± 6.7
Subtracted Yield	0.0 ± 1.9	1.9 ± 1.9
Final Yield $n_{D_2^*}$	6.9 ± 6.9	2.0 ± 7.0
$\mathcal{P}(D_2^{*0})$	(0.092 ± 0.092) %	(0.029 ± 0.102) %
Average $\mathcal{P}(D_2^{*0})$	(0.064 ± 0.068) %	

Table 7.7: Yields and product branching fractions for both D^0 decay modes. The errors are statistical only. The agreement is excellent between the results obtained with a weighted average versus the combined fit.

so that no additional systematic error is associated with this source. The results are summarized in Table 7.7. Special attention was taken to see if an event could possibly contribute a $D_J^0\, \ell^-$ candidate in the $K^-\pi^+$ and $K^-\pi^+\pi^0$ modes simultaneously. We found no event shared between both D^0 decay modes. The advantage of combining the $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^0$ mode is that it reduces the fitting systematic uncertainties. Likewise, the branching fractions were computed separately for each lepton mode in data. No discrepancy was found.

Monte Carlo experiments were used to verify the analysis algorithm. The entire analysis procedure applied to data was carried out on the generic $B\bar{B}$ MC sample. The generated branching fractions for our signal sample were $\mathcal{B}(B^- \to$

$D_1^0\ell^-\bar{\nu}_\ell) = 0.74\%$ and $\mathcal{B}(B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell) = 0.43\%$. The δM_J distributions for the generic $B\bar{B}$ MC is shown in Figure 7.12 for both D^0 decay modes combined. One can clearly see the contribution from $B^- \to D_1^0\ell^-\bar{\nu}_\ell$ and $B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell$. All the events above the dashed line are the D_1^0 and D_2^* candidates. The measured values for the branching fractions are $\mathcal{B}(B^- \to D_1^0\ell^-\bar{\nu}_\ell) = (0.72 \pm 0.05)\%$ and $\mathcal{B}(B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell) = (0.46 \pm 0.14)\%$, where the errors are statistical only. As one can see in Figure 7.12, the background shape of the δM_J distribution in generic $B\bar{B}$ MC is quite similar to the background function extracted from data. This check shows that the fitting functions used to describe the signal and the background are appropriate and reliable. Both decay modes of the D^0 were checked separately and similar results were obtained.

By looking closely at Figure 7.12, one might note a gap between the dashed line and the histogram. The dashed line describes the background function and the histogram is the tagged combinatoric background. As mentioned before, the smooth background function describes the combinatoric background and the possible contribution from broad and nonresonant $D^{*+}\pi^-$ states. Hence, the histogram does not include any real $D^{*+}\pi^-$ candidates correctly reconstructed as $B^- \to D^{*+}\pi^-\ell^-\bar{\nu}_\ell$. Nevertheless, the possible presence of broad or nonresonant $D^{*+}\pi^-$ states does not significantly change the shape of the background function and we believe that it does not bias the measurement of $\mathcal{B}(B^- \to D_1^0\ell^-\bar{\nu}_\ell)$ and $\mathcal{B}(B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell)$.

To further ensure that the combinatoric background is properly modeled in MC, the wrong-sign sample from data was compared with the wrong-sign sample from generic $B\bar{B}$ MC. The wrong-sign MC sample was first scaled down to match the number of $\Upsilon(4S)$ in data. The continuum and the fake lepton contaminations to the wrong-sign sample in data have been removed. In the generic $B\bar{B}$ MC, the fake leptons are discarded so that we can compare the combinatoric background in data and in MC with no contribution from other backgrounds. The comparison between the wrong-sign combination from data and MC is shown in Figure 7.13, where both distributions are normalized to equal area. The normalization factor is given by

$$\mathcal{N} = \frac{\text{Area of the wrong-sign distribution in data}}{\text{Area of the wrong-sign distribution in MC}} = 0.96 \pm 0.06\,, \tag{7.23}$$

where the error on $\mathcal{N}$ is statistical only. The overall normalization between data

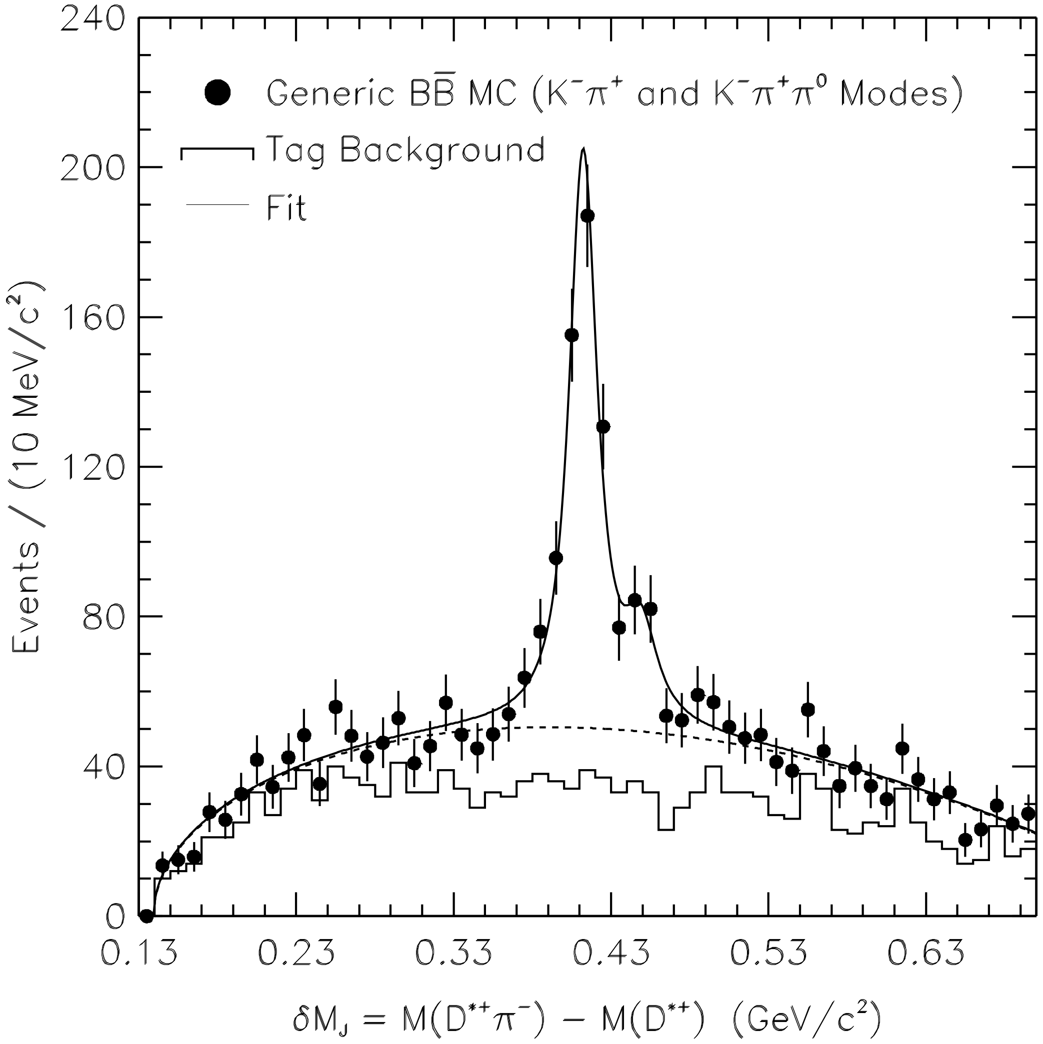


Figure 7.12: The δM_J distribution for generic $B\bar{B}$ MC events. The points are the MC data and the histogram is the tagged combinatoric background. The dashed curve describes the background function, whereas the solid line is the sum of the background and signal functions.

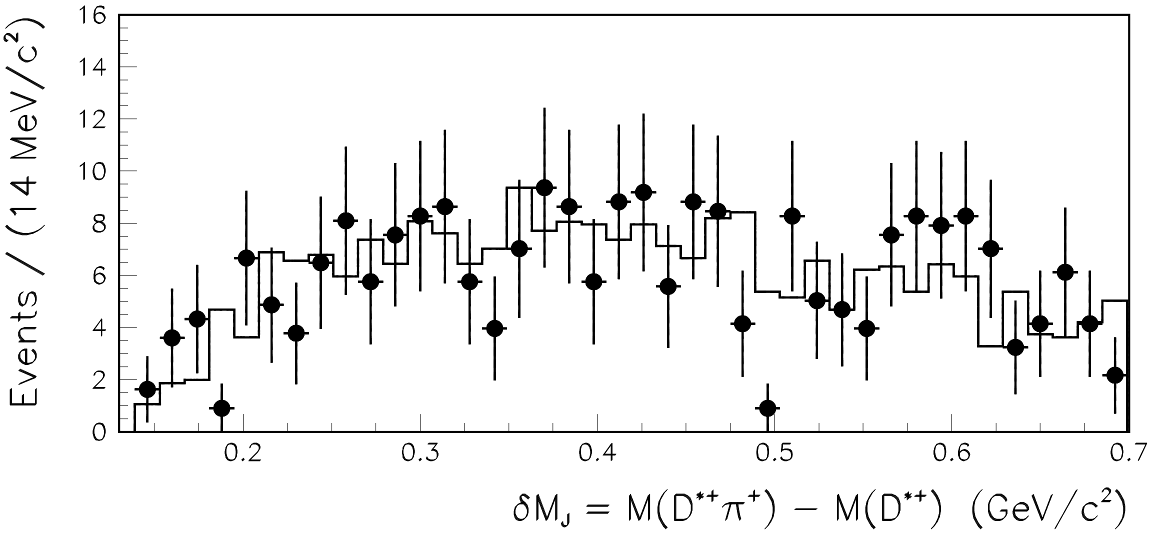


Figure 7.13: The wrong-sign sample in data (solid circles) and MC (histogram). The area of the histogram is normalized to match the area of the distribution in data. The normalization factor is $\mathcal{N} = 0.96 \pm 0.06$.

and MC is consistent with unity; but more importantly, the shape of the wrong-sign distribution is well reproduced in the MC simulation.

7.3.9 Summary

Having evaluated the experimental and theoretical uncertainties, the product branching fractions are then:

$$\mathcal{P}(D_1^0) = (0.373 \pm 0.085 \pm 0.052 \pm 0.024)\,\%, \quad (7.24)$$

$$\mathcal{P}(D_2^{*0}) = (0.059 \pm 0.066 \pm 0.010 \pm 0.004)\,\% < 0.16\,\%\ (90\%\ \text{C.L.}), \quad (7.25)$$

where the errors are statistical, systematic, and theoretical, respectively. For the quoted upper limit, we add the experimental systematic and the theoretical uncertainties in quadrature, and add the result to the upper limit computed with the statistical error only. The upper limit is at the 90% confidence level.

The uncertainties on the width of the D_J^0 resonances turn out to introduce the largest systematic error. Fortunately, the dependence of $\mathcal{P}(D_1^0)$ on the input width of the D_1^0 can be parameterized, as shown in Figure 7.14. Consequently, we are able to determine the product branching fraction $\mathcal{P}(D_1^0)$ as a function of the width of the D_1^0 resonance:

$$\mathcal{P}(D_1^0) = (\mathcal{P}_0 + \frac{d\mathcal{P}}{d\Gamma}\Delta\Gamma)\,\%, \tag{7.26}$$

where $\mathcal{P}_0$ is the central value quoted in Equation (7.24), and $\Delta\Gamma = \Gamma - \Gamma_0$. The width of the D_1^0 used to compute $\mathcal{P}_0$ is $\Gamma_0 = 18.9$ MeV$/c^2$, as indicated in Table 7.1. The value of the slope $d\mathcal{P}/d\Gamma$ is extracted from a linear fit of $\mathcal{P}(D_1^0)$ versus $\Delta\Gamma$ (see Figure 7.14). We find $d\mathcal{P}/d\Gamma = 9.25 \times 10^{-2}$ MeV$^{-1}c^2$. The quadratic component is negligible. To first order $\Delta\mathcal{P}(D_1^0) \simeq \Delta\mathcal{P}_0$. Thus, the product branching fraction can be written as:

$$\mathcal{P}(D_1^0) = [(0.373 + 9.25 \times 10^{-2}\,\Delta\Gamma) \pm 0.085 \pm 0.037 \pm 0.024]\,\%, \tag{7.27}$$

where the errors are statistical, systematic, and theoretical respectively. Of course, the systematic uncertainty does not include the uncertainty on the D_1^0 width.

In order to estimate the contribution of the decays $B^- \to D_1^0\ell^-\bar{\nu}_\ell$ and $B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell$ to the total semileptonic B meson branching fraction, we need to make some assumptions about the branching fractions of the D_J^0 mesons. We assume that $\mathcal{B}(D_1 \to D^*\pi) = 100\%$ and $\mathcal{B}(D_2^* \to D\pi + D^*\pi) = 100\%$. Decay modes such as $D_J \to D\rho$ or $D\eta$ and $D_J \to D^*\rho$ or $D^*\eta$ are kinematically disfavored, and therefore not considered. We also neglect any possible contributions from $D_J \to D^{(*)}\pi\pi$. Several theoretical predictions [91] and recent measurements from the DELPHI collaboration [119] seem to confirm our assumptions. Isospin symmetry suggests $\mathcal{B}(D_J^0 \to D^{*+}\pi^-)/\mathcal{B}(D_J^0 \to D^{*0}\pi^0) = 2$, leading to

$$\begin{aligned}\mathcal{B}(D_1^0 \to D^{*+}\pi^-) &= \frac{2}{3}, \\ \mathcal{B}(D_2^{*0} \to D^{*+}\pi^-) &= \frac{2}{3}\left(\frac{1}{1+R}\right),\end{aligned} \tag{7.28}$$

where $R = \mathcal{B}(D_2^* \to D\pi)/\mathcal{B}(D_2^* \to D^*\pi)$. Using the result $R \simeq 2.2$ quoted in Equation (3.15), $\mathcal{B}(D_2^{*0} \to D^{*+}\pi^-) \simeq 20.9\%$. We are conservative and use $\mathcal{B}(D_2^{*0} \to D^{*+}\pi^-) = 20\%$ since we quote an upper limit for $B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell$. Therefore, using

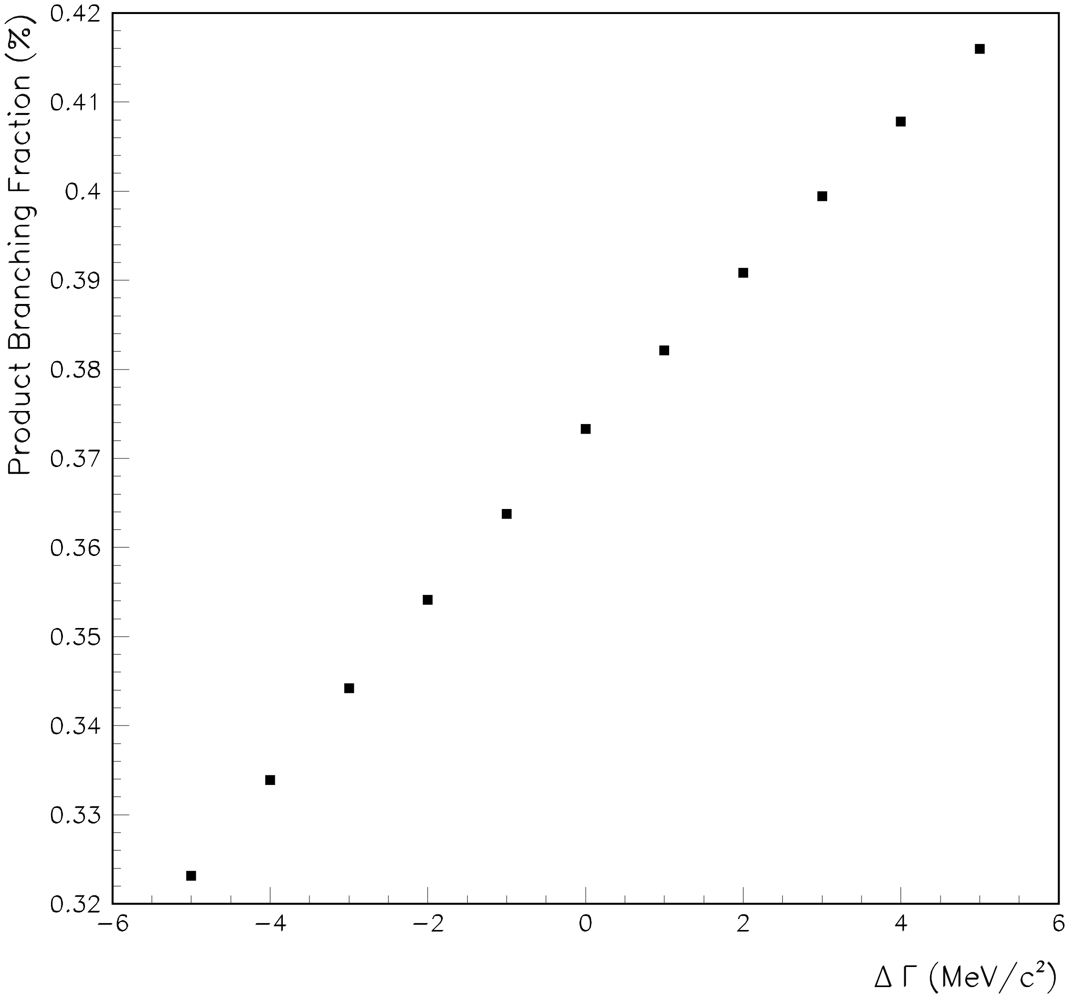


Figure 7.14: Dependence of the product branching fraction $\mathcal{P}(D_1^0)$ on $\Delta\Gamma = \Gamma - \Gamma_0$. The parameter Γ is the width of the D_1^0 resonance, and Γ_0 is the value used in the fit described in the text. It clearly shows the 10% systematic error on $\mathcal{P}(D_1^0)$ caused by the uncertainty on the D_1^0 width. The slope of the distribution is used to parameterize $\mathcal{P}(D_1^0)$ as a function of $\Delta\Gamma$.

	D_1^0
ON Resonance Yield	56.6 ± 11.9
Subtracted Yield	3.1 ± 2.8
Final Yield n_{D_1}	53.5 ± 12.2
$\mathcal{P}(D_1^0)$	$(0.373 \pm 0.085 \pm 0.052 \pm 0.024)\%$
$\mathcal{B}(B^- \to D_1^0 \ell^- \bar{\nu}_\ell)$	$(0.56 \pm 0.13 \pm 0.08 \pm 0.04)\%$

	D_2^{*0}
ON Resonance Yield	10.3 ± 9.4
Subtracted Yield	1.5 ± 2.8
Final Yield $n_{D_2^*}$	8.8 ± 9.8
$\mathcal{P}(D_2^{*0})$	$< 0.16\,\%$ (90% C.L.)
$\mathcal{B}(B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell)$	$< 0.8\ \ \%$ (90% C.L.)

Table 7.8: Summary of the yields and branching fractions for $B^- \to D_J^0 \ell^- \bar{\nu}_\ell$. The error on the yields is statistical. The first error on the branching fractions is statistical, the second is experimental systematic, and the third is theoretical. The confidence level for the upper limits is 90%.

the estimates

$$\begin{aligned} \mathcal{B}(D_1^0 \to D^{*+}\pi^-) &= 67\%, \qquad (7.29) \\ \mathcal{B}(D_2^{*0} \to D^{*+}\pi^-) &= 20\%, \end{aligned}$$

we find

$$\mathcal{B}(B^- \to D_1^0 \ell^- \bar{\nu}_\ell) = (0.56 \pm 0.13 \pm 0.08 \pm 0.04)\,\%, \qquad (7.30)$$

$$\mathcal{B}(B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell) < 0.8\,\% \quad (90\%\ \text{C.L.}), \qquad (7.31)$$

where no attempt has been made to estimate the systematic uncertainties due to the $D_J^0 \to D^{*+}\pi^-$ branching fractions. The results are summarized in Table 7.8.

Source of Systematic Error	$\mathcal{R}$
M_{D_J}	1.7%
Γ_{D_J}	11.4%
Background Function	5.4%
Uncorrelated Background	0.7%
Lepton Fake	1.4%
MC Statistics	2.1%
Total	13.0%

Table 7.9: Experimental systematic errors on $\mathcal{R}$.

Finally, the ratio of branching fractions $\mathcal{R}$ is:

$$\mathcal{R} = \frac{\mathcal{B}(B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell)}{\mathcal{B}(B^- \to D_1^0\ell^-\bar{\nu}_\ell)} = \frac{(0.30 \pm 0.33)\,\%}{(0.56 \pm 0.13)\,\%}, \tag{7.32}$$

where the errors on the branching fractions are statistical only. In the ratio of the two branching fractions for $B^- \to D_1^0\ell^-\bar{\nu}_\ell$ and $B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell$, all the common systematic uncertainties cancel out. The remaining systematic errors are enumerated in Table 7.9. The theoretical uncertainty is taken to be 9%. Then,

$$\mathcal{R} = \frac{\mathcal{B}(B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell)}{\mathcal{B}(B^- \to D_1^0\ell^-\bar{\nu}_\ell)} = 0.54 \pm 0.60 \pm 0.07 \pm 0.05, \tag{7.33}$$

where the first error is statistical, the second is systematic, and the third is theoretical. In the computation of the errors on the ratio, the correlation between n_{D_1} and $n_{D_2^*}$ has been properly taken into account. This leads to an upper limit of

$$\mathcal{R} = \frac{\mathcal{B}(B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell)}{\mathcal{B}(B^- \to D_1^0\ell^-\bar{\nu}_\ell)} < 1.5 \text{ (90\% C.L.)}. \tag{7.34}$$

7.4 Other Contributions to $B^- \to D^{*+}\pi^-\ell^-\bar{\nu}_\ell$

The right-sign sample in the data, after the subtraction of the continuum and fake leptons backgrounds, is shown in Figure 7.15. The solid line is the result of the unbinned likelihood fit using the sum of the background and signal functions. The dashed curve illustrates the background function, while the histogram is the estimate of the combinatoric background from MC. The amount of combinatoric background was obtained from the tagged background in generic $B\bar{B}$ MC, which was then scaled by $\mathcal{N}$ in order to reproduce the amount of combinatoric background in data. The normalization factor $\mathcal{N}$ was calculated in Section 7.3.8 (see Equation (7.23)).

The difference in yield between the dashed line and the histogram could be interpreted as a contribution from broad or nonresonant $D^*\pi$ in B semileptonic decays. As one can see, we are unfortunately not directly sensitive to the broad and nonresonant $D^{*+}\pi^-$ states, which instead would only be seen as an overall normalization difference.

The statistical uncertainty on the background yield from the fit is considerable ($\sim 16\%$). Moreover, the amount of $D^{*+}\pi^-$ is negligible within the errors (statistical and systematic) in the histogram. The systematic uncertainty on the scaling of the combinatoric background is of the order of ${}^{+20\%}_{-13\%}$. The systematic uncertainty was obtained by varying the contribution of the narrow D_J states within the error on n_{D_J} and by varying the contribution of the broad and nonresonant $D^{*+}\pi^-$ states to the combinatoric background in our generic $B\bar{B}$ MC. A similar conclusion is obtained when we use wrong-sign data to estimate the level of combinatoric background (see Figure 7.8).

In principle, one might be able to estimate the contribution of the broad and nonresonant states to $B^- \to D^{*+}\pi^-\ell^-\bar{\nu}_\ell$. In practice, the scaling of the background is rather difficult to estimate since the states which decay to $D^{*+}\pi^-$ contribute to the combinatoric background. The complications for the measurement of $\mathcal{B}(B^- \to D^{*+}\pi^-\ell^-\bar{\nu}_\ell)$ arise from the fact that the shape of the combinatoric background resembles the shape of the signal for the broad and nonresonant $D^{*+}\pi^-$ states. This leaves only the possibility of distinguishing the signal from the background on the basis of overall normalizations. Such a measurement then requires a large sample of $B\bar{B}$ and an excellent understanding of the level of all the combinatoric backgrounds.

Therefore, no attempt was made to estimate the contribution of the broad

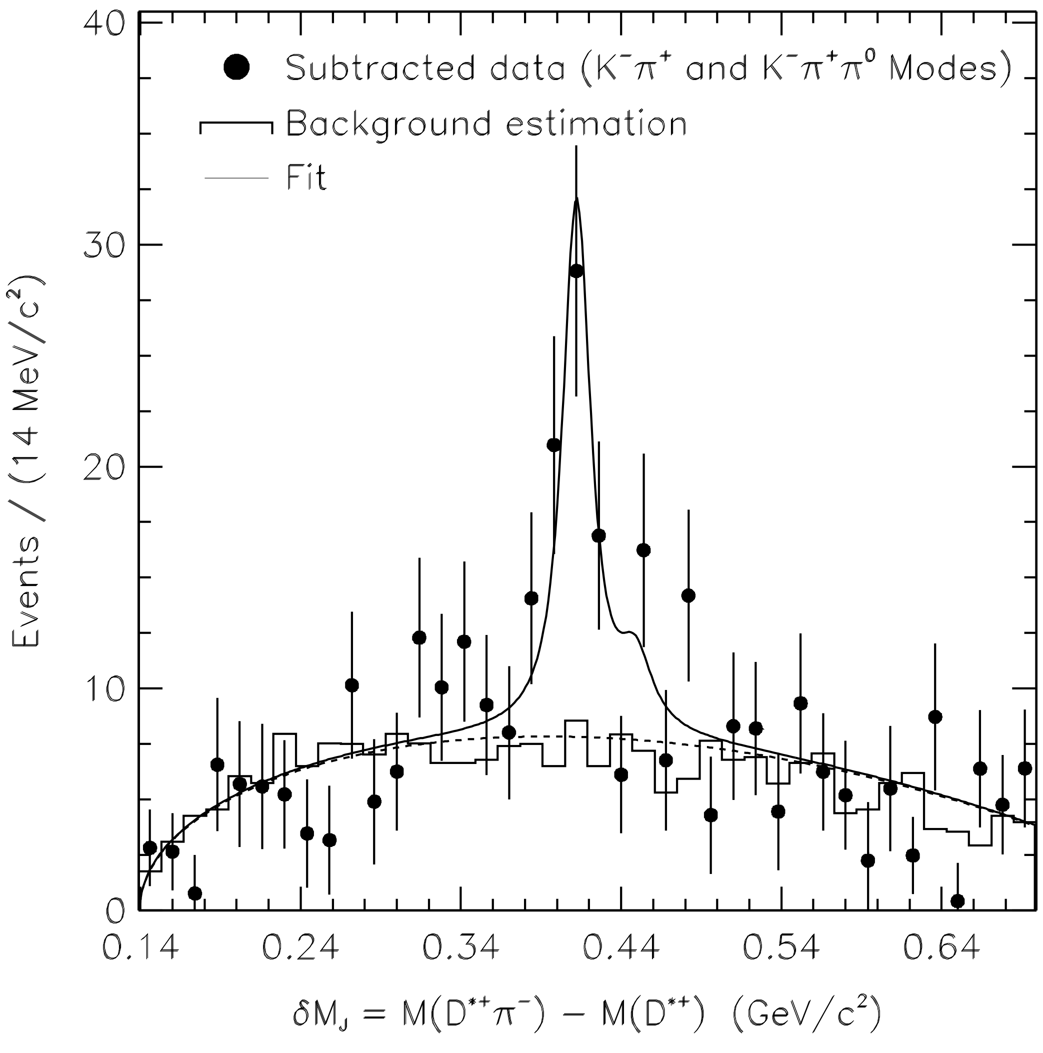


Figure 7.15: The δM_J distribution from data after the subtraction of the continuum and fake lepton backgrounds The distribution is obtained by combining both D^0 decay modes. The solid line is the result of the fit with the sum of the background and signal functions. The dashed curve shows the background function from the fit, while the histogram is the background estimation from MC.

and nonresonant $D^{*+}\pi^-$ states in semileptonic decays of the B meson. Such a measurement would be difficult and highly model dependent because:

1. The broad D_J and the pure nonresonant $D^*\pi$ states have not yet been established experimentally; thus, the estimation of the relative amount of each $D^*\pi$ states in B decays relies on theoretical predictions.

2. The detection efficiency also depends on the predicted $D^*\pi$ invariant mass for each state and on the decay dynamics of the B semileptonic decay within a given theoretical framework.

7.5 q^2 Spectrum for $B^- \to D_1^0 \ell^- \bar{\nu}_\ell$

Despite the fact that the analysis is statistically limited, we are nevertheless able to study the q^2 spectrum for $B^- \to D_1^0 \ell^- \bar{\nu}_\ell$. The q^2 spectrum is extracted by fitting the δM_J distribution in four bins of q^2, keeping the mass and the width of the D_1^0 fixed (see Table 7.1). In each bin, the appropriate continuum and fake lepton yields are subtracted from the fitted yield. The final or net yield $n_{D_1}(q^2)$ is then corrected by the reconstruction efficiency $\varepsilon_{D_1}(q^2)$, which was computed for the same q^2 bin. Following Equation (7.15), the q^2 spectrum is then the differential decay rate:

$$\frac{d\Gamma_1}{dq^2} = \frac{n_{D_1}(q^2)/\varepsilon_{D_1}(q^2)}{2\,\tau_{B^-}\, N_{\Upsilon(4S)}\, \mathcal{B}(D_1^0 \to D^{*+}\pi^-)\, \mathcal{B}(D^{*+} \to D^0\pi^+)\, \mathcal{B}(D^0 \to K^-\pi^+(\pi^0))}. \quad (7.35)$$

The B^- lifetime is taken to be $\tau_{B^-} = (1.62 \pm 0.06)$ ps [6]. The D^{*+} and D^0 branching fractions used are given in Equations (7.16)-(7.18). We assumed $\mathcal{B}(D_1^0 \to D^{*+}\pi^-) = 67\%$ and $f_{+-} = 0.5$. The resulting q^2 spectrum is shown in Figure 7.16. The dashed line in Figure 7.16 is the prediction from the ISGW2 model and the histogram is the result of the fit described below. One can then make the change of variable $q^2 \to w$ (*c.f.*, Equation (2.8)) and look at the w distribution:

$$\begin{aligned} \frac{d\Gamma_1}{dw} &= \frac{G_F^2 |V_{cb}|^2 m_B^5 r_1^3}{48\pi^3} (w^2-1)^{3/2}\, \mathcal{F}_{D_1}^2(w) \times \\ & \frac{2}{3}(w+1) \left[(w-1)(1+r_1)^2 + w(1-2wr_1+r_1^2)\right] , \end{aligned} \quad (7.36)$$

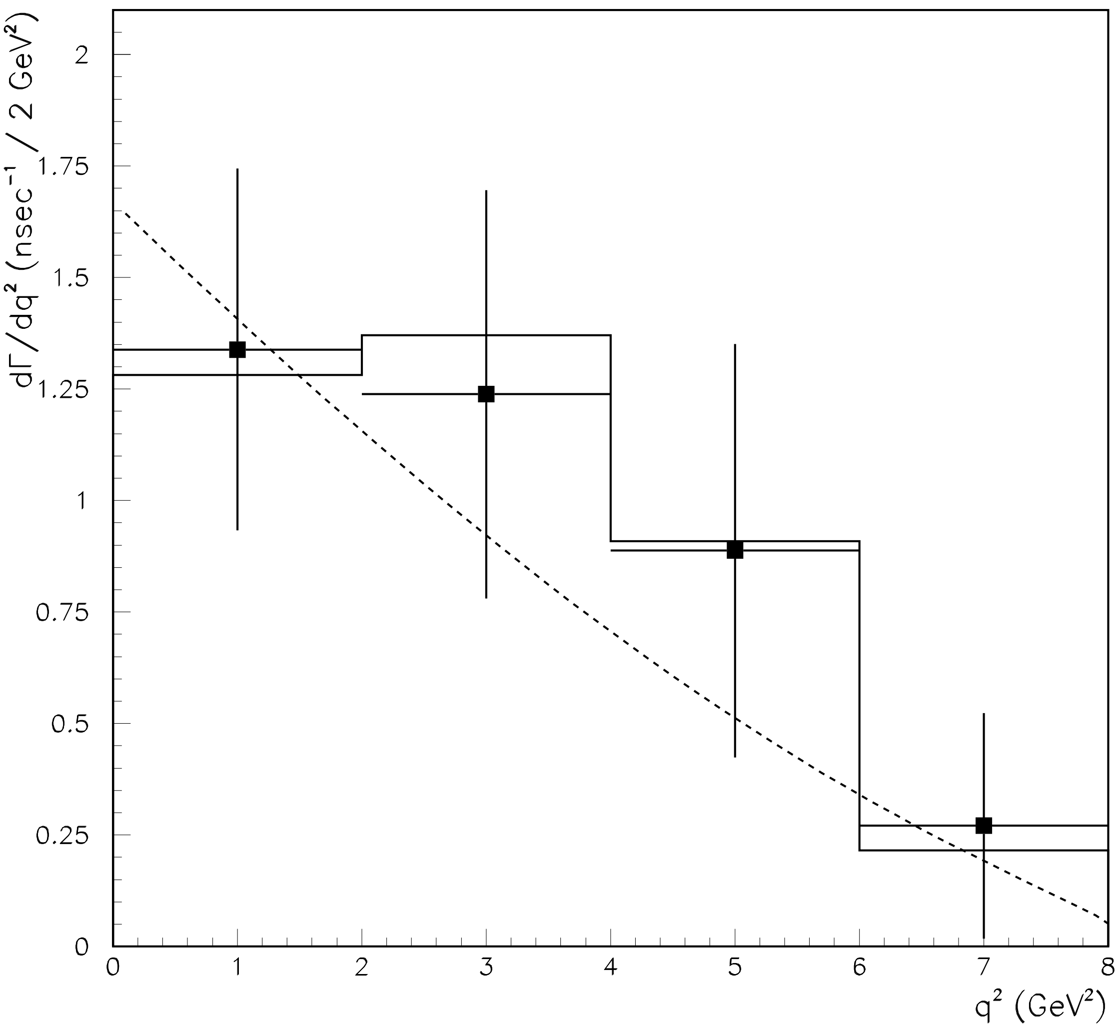


Figure 7.16: The q^2 spectrum for $B^- \to D_1^0 \ell^- \bar{\nu}_\ell$ data after background subtraction and efficiency correction. The error bar on each data point is statistical only. The dashed line is the prediction from the ISGW2 model and the histogram is the result of the fit to the distribution. The details of the fit are discussed in the text.

Source of Systematic Error	$\mathcal{F}_{D_1}(1)$	$\rho^2_{D_1}$
$n_{D_J}(q^2)$	6.6%	6.9%
$\varepsilon_{D_1}(q^2)$	4.5%	4.7%
$\mathcal{B}(D^{*+} \to D^0\pi^+)$	1.0%	–
$\mathcal{B}(D^0 \to K^-\pi^+(\pi^0))$	1.8%	–
$B\bar{B}$ cross-section	1.0%	–
B^- lifetime	2.3%	–
$\lvert V_{cb}\rvert$	4.3%	–
q^2 resolution	10.0%	10.0%
Total	13.9%	13.0%

Table 7.10: Experimental systematic errors on $\mathcal{F}_{D_1}(1)$ and $\rho^2_{D_1}$.

The decay rate $d\Gamma_1/dw$ is related to $\mathcal{F}^2_{D_1}(w)\,|V_{cb}|^2$. The measured $\mathcal{F}_{D_1}(w)\,|V_{cb}|$ distribution is shown in Figure 7.17 together with a fit to Equation (7.36) using the functional form:

$$\mathcal{F}_{D_1}(w) = \mathcal{F}_{D_1}(1)\left\{1 - \rho^2_{D_1}(w-1) + \mathcal{O}[(w-1)^2]\right\}. \tag{7.37}$$

At this point, one can use the world average for the CKM element $|V_{cb}| = (39.6 \pm 1.7) \times 10^{-3}$ [15] to compute $\mathcal{F}_{D_1}(1)$ and $\rho^2_{D_1}$. We obtain $\mathcal{F}_{D_1}(1) = 0.38 \pm 0.25$ and $\rho^2_{D_1} = 2.1 \pm 1.4$. A linear fit to $\mathcal{F}_{D_1}(w)\,|V_{cb}|$ is crude but we do not have the statistical power to add a quadratic term in the fit. Therefore, we neglect the possible curvature of the function $\mathcal{F}_{D_1}(w)$. This effect can be quite significant, as shown by the ISGW2 prediction for $\mathcal{F}_{D_1}(w)$ in Figure 7.17.

The precision of $\mathcal{F}_{D_1}(1)$ is determined primarily by the data point at the lowest w value, where, unfortunately, the statistical precision is poor. Each variation of the $n_{D_1}(q^2)$, $\varepsilon_{D_1}(q^2)$, $N_{\Upsilon(4S)}$, $\mathcal{B}(D^{*+} \to D^0\pi^+)$ and $\mathcal{B}(D^0 \to K^-\pi^+(\pi^0))$ generates a new q^2 spectrum and gives new values for $\mathcal{F}_{D_1}(1)$ and $\rho^2_{D_1}$. The systematic uncertainties associated with the measured yield, the $B\bar{B}$ cross-section, the branching

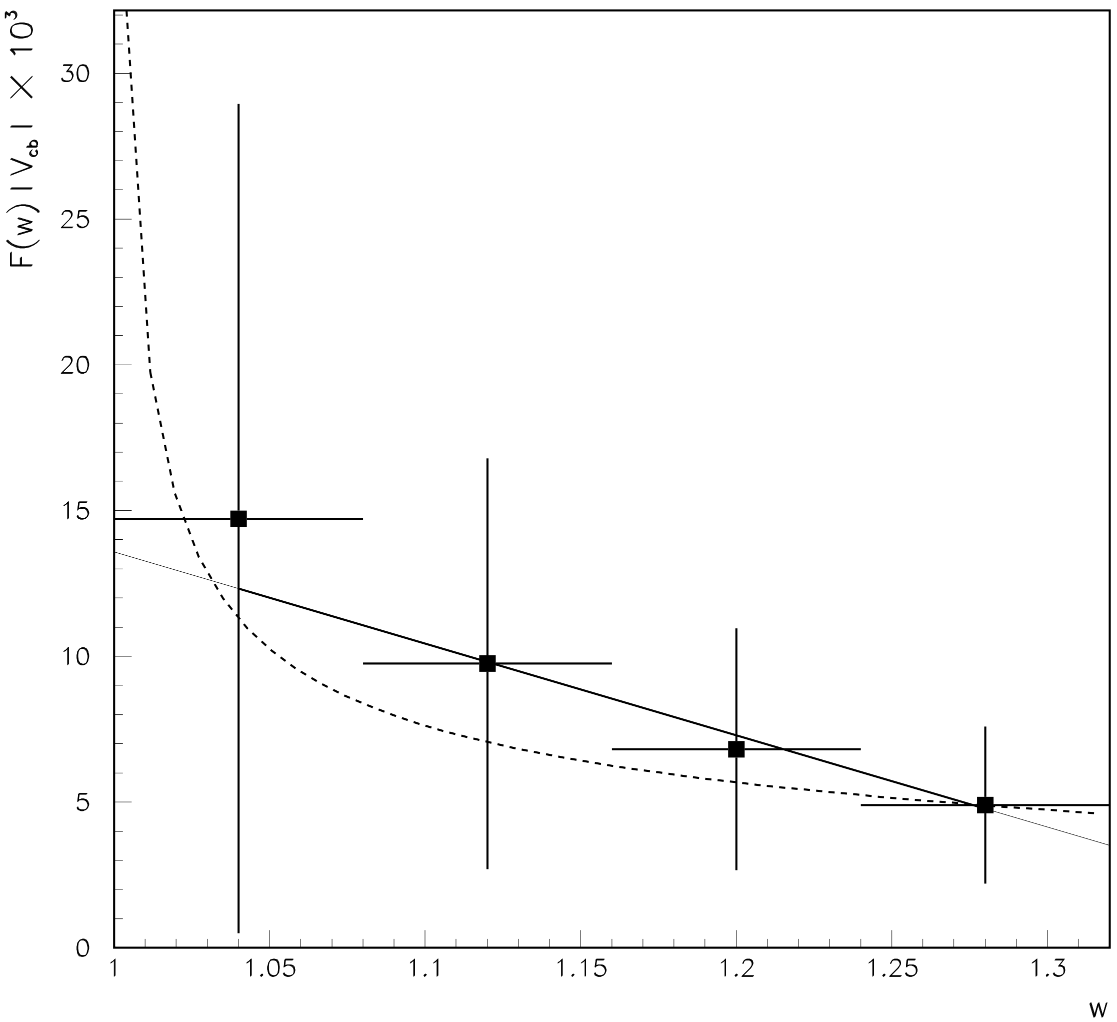


Figure 7.17: Measured values of $\mathcal{F}_{D_1}(w)\,|V_{cb}|$ from data. The result of the fit described in the text is the solid line, and the dashed line is the prediction for $\mathcal{F}_{D_1}(w)\,|V_{cb}|$ in the ISGW2 model with $|V_{cb}| = 39.6 \times 10^{-3}$.

fractions, and the reconstruction efficiency are propagated in the calculation of the error on $\mathcal{F}_{D_1}(w)$. They are listed in Table 7.10. The systematic uncertainties associated with n_{D_1} and ε_{D_1} are described in Section 7.3.6. The uncertainty on the q^2 resolution is about 10% (see Figure 7.3(c)). Other systematic errors, such as the experimental errors on τ_{B^-} and $|V_{cb}|$ lead to a systematic uncertainty of 2.3% and 4.2% on $\mathcal{F}_{D_1}(1)$ respectively. The slope $\rho^2_{D_1}$ is not affected by an overall normalization caused by the uncertainties on the $B\bar{B}$ cross-section, the branching fractions, the B^- lifetime, and $|V_{cb}|$. We add the systematic errors in quadrature for a total uncertainty of 13.9% on $\mathcal{F}_{D_1}(1)$ and 13.0% on $\rho^2_{D_1}$, as summarized in Table 7.10. The zero-recoil curvature and the higher derivative of the Isgur-Wise function are dropped from our description of $\mathcal{F}_{D_1}(w)$ and no attempt has been made to include the theoretical uncertainty due to the shape of $\mathcal{F}_{D_1}(w)$. The model dependence of $\rho^2_{D_1}$, due to the q^2 dependence of the form factors, is also neglected. Together with the statistical and systematic uncertainties, the fitted parameters are:

$$\mathcal{F}_{D_1}(1) = 0.38 \pm 0.25 \pm 0.05 \tag{7.38}$$

$$\rho^2_{D_1} = 2.1 \pm 1.4 \pm 0.3 \tag{7.39}$$

The result for $\mathcal{F}_{D_1}(1)$ can be converted to an upper limit since we expect $\mathcal{F}(1)$ to be greater than zero.

$$\mathcal{F}_{D_1}(1) < 0.8 \text{ (90\% C.L.).} \tag{7.40}$$

For the upper limit calculation, we add the experimental systematic uncertainty to the upper limit computed with the statistical error only.

Chapter 8

Interpretations and Conclusion

8.1 Experimental Results

This thesis presents an investigation of the production of the narrow P-wave charm mesons in semileptonic B decays using data collected by the CLEO II detector. We measure the product branching fractions $\mathcal{P}(D_J^0) = \mathcal{B}(B^- \to D_J^0 \ell^- \bar{\nu}_\ell)\, \mathcal{B}(D_J^0 \to D^{*+}\pi^-)$ to be:

$$\begin{aligned} \mathcal{P}(D_1^0) &= (0.373 \pm 0.085 \pm 0.052 \pm 0.024)\,\%, && (8.1)\\ \mathcal{P}(D_2^{*0}) &= (0.059 \pm 0.066 \pm 0.010 \pm 0.004)\,\% < 0.16\ \%\ (90\%\ \text{C.L.}), && (8.2) \end{aligned}$$

where the errors are statistical, systematic, and theoretical, respectively. The dependence of $\mathcal{P}(D_1^0)$ on the input width of the D_1^0 can be parameterized. This gives

$$\mathcal{P}(D_1^0) = [(0.373 + 9.25 \times 10^{-2}\, \Delta\Gamma) \pm 0.085 \pm 0.037 \pm 0.024]\,\%. \quad (8.3)$$

In order to estimate the contribution of $B^- \to D_1^0 \ell^- \bar{\nu}_\ell$ and $B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell$ to the total semileptonic B meson branching fraction, we need to make some assumptions about the branching fractions of the D_J^0 mesons. Isospin conservation and CLEO measurements of the decays of the D_J^0 mesons suggest that $\mathcal{B}(D_1^0 \to D^{*+}\pi^-) = 67\%$ and $\mathcal{B}(D_2^{*0} \to D^{*+}\pi^-) = 20\%$. Using these estimates we find

$$\begin{aligned} \mathcal{B}(B^- \to D_1^0 \ell^- \bar{\nu}_\ell) &= (0.56 \pm 0.13 \pm 0.08 \pm 0.04)\,\%, && (8.4)\\ \mathcal{B}(B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell) &< 0.8\ \%\ \ (90\%\ \text{C.L.}), && (8.5) \end{aligned}$$

which leads to an upper limit on the ratio of branching fractions

$$\mathcal{R} = \frac{\mathcal{B}(B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell)}{\mathcal{B}(B^- \to D_1^0\ell^-\bar{\nu}_\ell)} < 1.5 \text{ (90\% C.L.)}. \tag{8.6}$$

Furthermore, we measure the q^2 spectrum for $B^- \to D_1^0\ell^-\bar{\nu}_\ell$. Under strict assumptions on the form of $\mathcal{F}_{D_1}(w)$, we determine the parameters $\mathcal{F}_{D_1}(1)$ and $\rho^2_{D_1}$ to be

$$\mathcal{F}_{D_1}(1) < 0.8 \text{ (90\% C.L.)}, \tag{8.7}$$

$$\rho^2_{D_1} = 2.1 \pm 1.4 \pm 0.3\,, \tag{8.8}$$

where the first error is statistical and the second is the experimental systematic uncertainty.

8.2 Other Experimental Results

In this section, we compare our experimental results for $\mathcal{B}(B^- \to D_1^0\ell^-\bar{\nu}_\ell)$ and $\mathcal{B}(B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell)$ with other experimental results and we describe similar measurements done elsewhere. Here and throughout the remainder of this thesis, we use $\sum \bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ to denote all semileptonic decays of the B meson to orbitally and radially excited D^{**} and nonresonant states D^*X.

Early measurements of $\sum \mathcal{B}(\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell)$ were performed by the ARGUS [27] and the CLEO [28] collaborations. The ARGUS collaboration used the $D^*\ell$ recoil mass squared, $M^2(D^*\ell)$, to search for $\bar{B} \to D^*\pi\ell^-\bar{\nu}_\ell$. They used a MC simulation based on the ISGW model to predict the rate, the shape of the $M^2(D^*\ell)$ spectrum, and the detection efficiencies for all the possible states which contribute to $\sum \mathcal{B}(\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell)$. Isospin symmetry was employed to estimate the branching fraction for $D^{**} \to D^*X$. The signal process $\bar{B} \to D^*X\ell^-\bar{\nu}_\ell$ is shifted to positive $M^2(D^*\ell)$ values as expected from the non-vanishing $\bar{\nu}_\ell X$ invariant mass. Hence, the yield for $\sum \mathcal{B}(\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell)$ is obtained by fitting the shoulder of $M^2(D^*\ell)$. The ARGUS measurement is [17]:

$$\sum \mathcal{B}(\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell) = (2.7 \pm 0.5 \pm 0.5)\,\%, \tag{8.9}$$

where the fist error is statistical and the second is systematic. The CLEO result extended the recoil mass squared method to include the lepton momentum information because a process like $\bar{B} \to D^* \ell^- \bar{\nu}_\ell$ has a harder lepton momentum spectrum than $\bar{B} \to D^* \pi \ell^- \bar{\nu}_\ell$. Using the ISGW model to obtain an average efficiency for $\sum \bar{B} \to D^{**} \ell^- \bar{\nu}_\ell$, the CLEO collaboration obtains:

$$\sum \mathcal{B}(\bar{B} \to D^{**} \ell^- \bar{\nu}_\ell) < 2.8\,\% \text{ (95\% C.L.)}, \tag{8.10}$$

which is consistent with the ARGUS results. The ARGUS and the CLEO measurements do not include the uncertainty of the ISGW estimate of the relative D^{**} abundances. In fact, no model dependence is included in their systematic uncertainties. Both CLEO and ARGUS neglected the possible contribution from nonresonant $D^* \pi$ since the ISGW model does not incorporates such states. Therefore, it is difficult to identify the decay chain that produced the excess of $D^* X$ events.

The ALEPH, OPAL, and DELPHI collaborations have performed a direct search for $D^{(*)} X$ production in B semileptonic decays. The ALEPH, OPAL, and DELPHI detectors are located at the LEP collider at CERN, which operated at $\sqrt{s} \simeq 92$ GeV during the LEP1 runs. At LEP1, the B mesons were produced back-to-back from the decay of the Z^0 to a pair of $b\bar{b}$ quarks. The fragmentation process of the bottom quark gives rise to jet-like events with distinctive vertex topologies. The B mesons are highly boosted ($\beta\gamma \simeq 6$) in the laboratory frame and typically travel a few centimeters before they decay. Silicon vertex detector information is therefore extremely helpful in associating tracks from B and D meson decays.

The ALEPH and OPAL analysis searched for B semileptonic decays with a narrow, a broad, or a nonresonant $D^{(*)} \pi$ pair in the final state. Such processes have interesting vertex topology due to the large boost of the B meson, as shown in Figure 8.1. Another desirable consequence of the B meson boost is that the decay products from $\bar{B} \to D^{(*)} \pi \ell^- \bar{\nu}_\ell$ are also boosted; which reduces the model dependence of the reconstruction efficiencies because larger fractions of the momentum spectra are sampled. Hence, the detection efficiency at LEP1 is less model dependent than at the $\Upsilon(4S)$. Yet, the large boost at LEP1 implies a lack of knowledge of the energy of the B meson. As a consequence, the LEP experiments always quote branching fractions

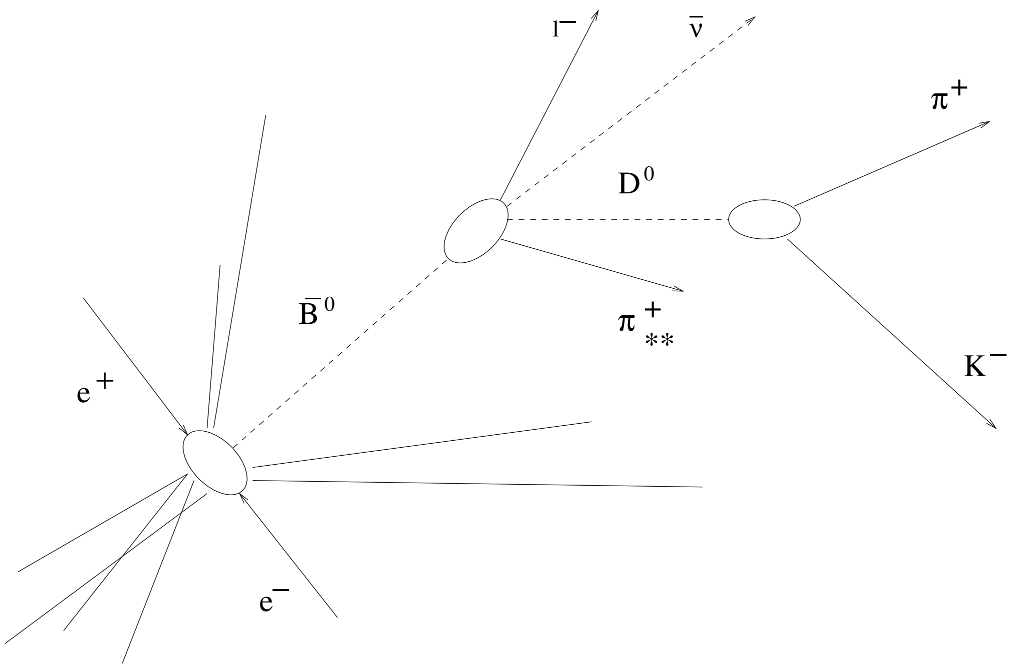


Figure 8.1: Vertex topology at LEP1 for a semileptonic $\bar{B}^0$ decay to a D^{**+} which then decays to a $D^0\pi^+_{**}$. A semileptonic B^- to a four-body $D^{*+}\pi^-\ell^-\bar{\nu}_\ell$ state would have a similar topology [29].

for $\bar{B} \to D^{(*)}\pi X\,\ell^-\bar{\nu}_\ell$. Here, we assume that most of the $D^{(*)}\pi$ states are produced in B semileptonic decays with no additional particles. In Table 8.1, our results are compared with the results from the LEP experiments. The ALEPH results for the narrow D^0_J states are [29, 90]:

$$\mathcal{B}(B^- \to D_1^0\ell^-\bar{\nu}_\ell) = (0.70 \pm 0.15 \pm 0.12)\,\% \quad (8.11)$$

$$\mathcal{B}(B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell) < 1.7\,\%\ (95\%\ \text{C.L.}). \quad (8.12)$$

The corresponding OPAL results are [30]:

$$\mathcal{B}(B^- \to D_1^0\ell^-\bar{\nu}_\ell) = (2.14 \pm 0.53 \pm 0.50)\,\% \quad (8.13)$$

$$\mathcal{B}(B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell) = (0.93 \pm 0.37 \pm 0.19)\,\%. \quad (8.14)$$

Independent topological searches from ALEPH [29] and DELPHI [31] yield results for inclusive $D^{(*)}X$ production in B semileptonic decay. ALEPH reported

$$\mathcal{B}(\bar{B} \to D\pi\ell^-\bar{\nu}_\ell) + \mathcal{B}(\bar{B} \to D^*\pi\ell^-\bar{\nu}_\ell) = (2.26 \pm 0.29 \pm 0.33)\,\%, \quad (8.15)$$

	$\mathcal{B}(B^- \to D_1^0 \ell^- \bar{\nu}_\ell)$	$\mathcal{B}(B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell)$	Reference
ALEPH	$(0.70 \pm 0.15 \pm 0.12)\%$	$< 1.7\%$ (95% C.L.)	[29, 90]
OPAL	$(2.14 \pm 0.53 \pm 0.50)\%$	$(0.93 \pm 0.37 \pm 0.19)\%$	[30]
CLEO	$(0.56 \pm 0.13 \pm 0.09)\%$	$< 0.9\%$ (95% C.L.)	This Analysis

Table 8.1: Comparison of experimental values for the branching fractions. The first error listed is statistical and the second systematic. Our systematic uncertainty is taken as the sum in quadrature of the experimental systematic and theoretical uncertainties. We quote a 95% confidence level in the table for $\mathcal{B}(B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell)$ in order to compare with the LEP measurements. We used $\mathcal{B}(b \to \bar{B}) = 37.8\%$ [6], $\mathcal{B}(D_1^0 \to D^{*+}\pi^-) = 67\%$ and $\mathcal{B}(D_2^{*0} \to D^{*+}\pi^-) = 20\%$ to extract the ALEPH and OPAL numbers from the product branching fractions that they quote. The LEP experiments actually measure $\mathcal{B}(\bar{B} \to D_J^0 X\, \ell^- \bar{\nu}_\ell)$, but we neglect any contribution from additional particles X for the comparison. ALEPH and OPAL also quote results for other charged modes.

and DELPHI quoted a value for the ratio

$$\frac{\mathcal{B}(B \to D^{*-} X\, \ell^+ \nu_\ell)}{\mathcal{B}(B \to D^{*-} X\, \ell^+ \nu_\ell) + \mathcal{B}(B^0 \to D^{*-} \ell^+ \nu_\ell)} = 0.19 \pm 0.10 \pm 0.06, \quad (8.16)$$

where X represents neutral or charged particles. In all LEP1 measurements, the first error is statistical and the second is the experimental systematic uncertainty. The ALEPH and OPAL results for the narrow D_J^0 production in B semileptonic decays are consistent with ours. The OPAL measurement for $B^- \to D_1^0 \ell^- \bar{\nu}_\ell$ is high, but still consistent because of its large uncertainties.

8.3 Theoretical Predictions

As outlined in Chapter 3, several theoretical models make predictions for the decay rate of exclusive semileptonic decays of the B meson to excited charm mesons. In general, the theoretical models can be divided in two classes. The first class (Class I) includes the models which consider the charm and bottom quarks heavy enough to neglect higher order Λ_{QCD}/m_Q corrections beyond the HQS prescriptions. In Table 8.2,

Experimental Results		
	$\mathcal{B}(B^- \to D_1^0 \ell^- \bar{\nu}_\ell)$	$\mathcal{B}(B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell)$
CLEO II	$(0.56 \pm 0.13 \pm 0.08 \pm 0.04)\%$	$< 0.8\%$ (90% C.L.)

Theoretical Predictions (Class I)		
	$\mathcal{B}(B^- \to D_1^0 \ell^- \bar{\nu}_\ell)$	$\mathcal{B}(B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell)$
SISM	0.088%	0.125%
VO	0.281%	0.448%
CNP	0.131%	0.263%
SHJL	0.178%	0.264%

Theoretical Predictions (Class II)		
	$\mathcal{B}(B^- \to D_1^0 \ell^- \bar{\nu}_\ell)$	$\mathcal{B}(B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell)$
ISGW2	0.457%	0.229%
SHJL	0.297%	0.193%

Table 8.2: Experimental results and theoretical predictions for the branching fractions for $B^- \to D_1^0 \ell^- \bar{\nu}_\ell$ and $B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell$. We used $|V_{cb}| = 39.6 \times 10^{-3}$ and $\tau_{B^-} = 1.62$ ps to compute the theoretical branching fractions.

our experimental results for $\mathcal{B}(B^- \to D_1^0 \ell^- \bar{\nu}_\ell)$ and $\mathcal{B}(B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell)$ are compared with the theoretical predictions obtained in the infinite heavy quark mass limit. The second class (Class II) includes the models which account for possible effects which break HQS. They are also summarized in Table 8.2.

The result for $\mathcal{B}(B^- \to D_1^0 \ell^- \bar{\nu}_\ell)$ presented in this thesis disfavor all the theoretical predictions which use the infinite heavy quark mass limit. The ISGW2 model is the one which is most consistent with both decay modes.

It is convenient to compare the ratio of branching fraction for B semileptonic decays into members of the (D_1, D_2^*) doublet. The theoretical prediction for both classes of models are listed in Table 8.3. The LLSW model gives the ratio $\mathcal{R}$ as

Experimental Results	
	$\mathcal{R}$
CLEO II	< 1.5 (90% C.L.)

Theoretical Predictions (Class I)	
	$\mathcal{R}$
SISM	1.4
VO	1.6
CNP	2.0
SHJL	1.5
LLSW	$1.77 + 0.51\hat{\tau}'$

Theoretical Predictions (Class II)	
	$\mathcal{R}$
ISGW2	0.50
SHJL	0.65
LLSW (Approx. A)	$0.85 + 0.27\hat{\tau}'$
LLSW (Approx. B)	$0.67 + 0.13\hat{\tau}'$

Table 8.3: Experimental results and theoretical predictions for the ratio $\mathcal{R}$.

a function of $\hat{\tau}' \equiv \tau'(1)/\tau(1)$ [120]. Based on HQET predictions, the slope of the leading Isgur-Wise function is believed to be negative. Therefore, $\hat{\tau}' < 0$. LLSW estimate $\hat{\tau}' \simeq -1.5$.

Our result for $\mathcal{R}$ seems to be more consistent with the second class of models. Unfortunately, our upper limit barely disfavors some models which consider the limit $m_Q \to \infty$. It becomes clear however that the discrepancy between the two classes of models may be explained by the $\Lambda_{\rm QCD}/m_Q$ effects.

The ISGW2, SHJL, and LLSW models provide us with predictions for the intercept $\mathcal{F}_{D_1}(1)$. They are listed in Table 8.4. It must be noted that LLSW use the

Experimental Results	
	$\mathcal{F}_{D_1}(1)$
CLEO II	< 0.8 (90% C.L.)

Theoretical Predictions (Class I)	
	$\mathcal{F}_{D_1}(1)$
SHJL	0.29
LLSW	0.88 - 0.02$\hat{\tau}'$

Theoretical Predictions (Class II)	
	$\mathcal{F}_{D_1}(1)$
ISGW2	0.82
SHJL	0.23
LLSW (Approx. A)	0.61 - 0.01$\hat{\tau}'$
LLSW (Approx. B)	0.55 - 0.04$\hat{\tau}'$

Table 8.4: Comparison of the experimental results and theoretical predictions for the intercept $\mathcal{F}_{D_1}(1)$.

assumption that $\mathcal{B}(B^- \to D_1^0 \ell^- \bar{\nu}_\ell) = 0.6\%$ in their calculation of $\mathcal{F}_{D_1}(1)$. Because of the large statistical uncertainty on the experimental $\mathcal{F}_{D_1}(w)\,|V_{cb}|$ distribution, a fit was performed with the assumption of a linear form for $\mathcal{F}_{D_1}(w)$. The statistical precision on $\mathcal{F}_{D_1}(1)$ is driven by the data near zero recoil. Unfortunately, the error bar on the data point with the lowest w is large. Therefore, it is difficult to draw a definite conclusion based on our measurement of the intercept $\mathcal{F}_{D_1}(1)$. Although the curvature might be significant at zero recoil, it seems that $\mathcal{F}_{D_1}(1)$ is smaller than unity, which is a conservative statement considering the measured upper limit for $\mathcal{F}_{D_1}(1)$ presented here.

8.4 Interpretations

Recent evidence of orbitally excited charm mesons production in B semileptonic decays opens unexplored experimental and theoretical territories. Only a few years ago, the contribution of the P-wave charm mesons represented a poorly understood part of the B semileptonic rate. To comprehend the whole of the inclusive and exclusive branching fractions of the B semileptonic decays, it is essential to study the higher resonance contributions.

Several implications stand out from the investigations of $D^{(*)}\pi$ production at CESR and LEP. The ALEPH result based on the topological vertex study (see Equation (8.15)) suggests that the $D^{(*)}\pi$ states make up $22.2 \pm 4.3\%$ of the B semileptonic decays. The DELPHI search for D^*X is consistent with the ALEPH result. The measurements presented in this thesis imply that the sum of the semileptonic decays which produce a $D_1(2420)$ or a $D_2^*(2460)$ meson in the final state accounts for at least 13% of $\mathcal{B}_{\mathrm{SL}}$. Therefore, our results indicate that a substantial fraction ($\gtrsim$18%) of the inclusive $\bar{B}$ semileptonic rate is from modes other than $D\ell\bar{\nu}_\ell$, $D^*\ell\bar{\nu}_\ell$, $D_1\ell\bar{\nu}_\ell$, and $D_2^*\ell\bar{\nu}_\ell$, as summarized in Table 8.5. A clear picture of the contribution of the broad and nonresonant $D^{(*)}\pi$ to the total B semileptonic rate has not yet emerged from the recent experimental efforts. It should be noted that these interpretations hold under specific assumptions: we assume the contribution of three body, ρ, and η decays of the narrow D_J to be negligible.

On the theoretical side, considerable effort has been devoted to understanding the dynamics of heavy quark mesons. Great interest has been given to the description of $\bar{B} \to D^{(*)}\ell\bar{\nu}_\ell$ semileptonic decays in the framework provided by HQET. The heavy quark symmetry enormously simplifies the analysis of $b \to c$ transitions. In HQET, the universal Isgur-Wise functions embody details of low-energy QCD effects. As the accuracy of the experimental measurements increase, other open questions related to the form of the Isgur-Wise functions must be addressed. A precise determination of $|V_{cb}|$ therefore relies on profound theoretical understanding of nonperturbative strong interaction physics underlying the decay of the B meson.

The decays of the B meson to orbitally excited charm mesons offer a promising

Decay Mode	Branching Fraction
$\bar{B} \to D\ell\bar{\nu}_\ell$	$(1.94 \pm 0.26)\%$
$\bar{B} \to D^*\ell\bar{\nu}_\ell$	$(5.05 \pm 0.25)\%$
$\bar{B} \to D_1\ell\bar{\nu}_\ell$	$(0.56 \pm 0.16)\%$
$\bar{B} \to D_2^*\ell\bar{\nu}_\ell$	$< 0.8\%$
$\mathcal{B}_{\rm SL}$	$(10.18 \pm 0.40)\%$
Inclusive - Exclusive	$\gtrsim 1.83\%$

Table 8.5: Contribution of $B^- \to D_1^0\ell^-\bar{\nu}_\ell$ and $B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell$ to the inclusive B semileptonic rate based on our measurements. The quoted error for $\mathcal{B}(\bar{B} \to D_1\ell\bar{\nu}_\ell)$ is the sum in quadrature of the statistical, experimental systematic, and theoretical uncertainties. The confidence level for the upper limit on $\mathcal{B}(\bar{B} \to D_2^*\ell\bar{\nu}_\ell)$ is 90%.

capability to understand the level of the heavy quark symmetry breaking to order of $\Lambda_{\rm QCD}/m_Q$. Our measurement of the rate for $B^- \to D_1^0\ell^-\bar{\nu}_\ell$ disfavors all the theoretical predictions that advocate small $\Lambda_{\rm QCD}/m_Q$ corrections beyond HQS for semileptonic decays of the B meson to P-wave charm mesons (see Table 8.2 for details).

As mentioned earlier, the zero recoil matrix elements of the weak currents between the B meson and any excited charm meson vanish in the $m_Q \to \infty$ limit. However, at order of $\Lambda_{\rm QCD}/m_Q$, these matrix elements are not necessarily zero and the rates might be enhanced because most of the available phase space for $B^- \to D_J^0\ell^-\bar{\nu}_\ell$ is near zero recoil. It turns out that the $\Lambda_{\rm QCD}/m_Q$ corrections are more important for the spin one ($J = 1$) member of the $j = 3/2$ doublet since the matrix element $\langle D_J(v',\epsilon)|(V^\mu - A^\mu)|B(v)\rangle$ near zero recoil can only be nonzero for spin zero or spin one charm mesons. This argument then explains why the $\Lambda_{\rm QCD}/m_Q$ corrections lead to a suppression of the ratio $\mathcal{R}$. The measured upper limit in Equation (8.6) suggests that $B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell$ is suppressed in comparison to $B^- \to D_1^0\ell^-\bar{\nu}_\ell$.

In HQET, the differential decay rate for $B^- \to D_1^0\ell^-\bar{\nu}_\ell$ provides information about the shape of the leading Isgur-Wise function $\tau(w)$, which then allows insight into nonperturbative QCD effects. In this thesis, we presented the first measurement of the differential decay rate for $B^- \to D_1^0\ell^-\bar{\nu}_\ell$. Conclusions and interpretations based on our measurement of $\mathcal{F}_{D_1}(w)$ are hard to draw since the statistical precision is poor. Nevertheless, our measurement opens new ground for the study of the shape of the Isgur-Wise function and thus on nonperturbative strong physics in B semileptonic decays.

8.5 Future Prospects

Accurate experimental studies of semileptonic B decays to P-wave charm mesons will be possible in the future. Near term measurements are most likely to come from the CLEO collaboration. The actual CLEO II data set consists of approximately 5.5 pb^{-1} on the $\Upsilon(4S)$ resonance and 2.8 pb^{-1} below the $\Upsilon(4S)$ resonance. At CLEO, the use of the recompress data and the installation of a silicon vertex detector in July 1995 should improve the track-finding efficiency on the slow pion from the D^* and therefore lead to better measurements of $\mathcal{B}(B^- \to D_1^0\ell^-\bar{\nu}_\ell)$ and $\mathcal{B}(B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell)$. Furthermore, precise measurements of the masses and widths of the narrow D_J^0 states from an inclusive analysis should reduce the uncertainty on the shape of the resonances; which will imply a significant reduction of the main systematic error in any study of narrow P-wave charm mesons production in B decays.

Long term measurements will probably take place at CLEO III, as well as at the SLAC and KEK asymmetric B-Factories. The peak luminosity designed for future e^+e^- colliders operating at a center-of-mass energy near the $B\bar{B}$ threshold is about 3×10^{33} cm^{-2}s^{-1}. Achieving this luminosity would yield approximatively 10^7 $B\bar{B}$ pairs per year. With the large B meson data samples expected from these experiments, the opportunities for precise studies of B semileptonic decays will be greatly expanded. It will be possible to investigate the full dynamics of $B^- \to D_1^0\ell^-\bar{\nu}_\ell$ and $B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell$ by measuring the form factors governing such decays. However, much work remains in improving the measurements of the semileptonic branching fraction of the B mesons to the elusive broad and nonresonant $D^{(*)}\pi$ states. A detailed understanding of their contribution to the inclusive B semileptonic rate will certainly require advanced analysis techniques.

Analysis at asymmetric B-Factories might be able to reduce the combinatoric background by using separate B-vertex constraints. Other methods, such as the neutrino reconstruction technique developed at CLEO to reconstruct $\bar{B} \to \pi \ell \bar{\nu}_\ell$, should provide a new experimental tool for reducing the background level in future exclusive B measurements at high luminosity machines.

In summary, it is expected that the global effort in studying the semileptonic B decays to charm mesons will provide a better understanding of the fundamental interactions which govern heavy quark decays. It is clear that there are a large number of interesting phenomena in B physics that require a large data sample. Future experimental facilities operating near the threshold of open beauty production will provide the opportunity to investigate the weak properties of the bottom quark in much more detail. This will open the possibility of many new tests of the Standard Model and widen the search for the origin of CP violation.

8.6 Conclusion

Since the early 20^{th} century, tremendous progress has been made in particle physics. The considerable achievement in establishing that matter is made of quarks and leptons interacting via gauge bosons is due to great experimental and theoretical efforts. The realization that the dynamics of elementary particles can be described by quantum field theories possessing local gauge symmetry represents indeed a remarkable breakthrough in understanding the distinct fundamental forces that govern quarks and leptons. The description of how matter interacts through the exchange of gauge field quanta such as photon, gluons, and weak bosons proved the success of the modern framework provided by the Standard Model. In recent years, the Standard Model has succeeded remarkably well in describing the fundamental constituents of the microphysical world. Nevertheless, many questions remain unanswered and there are certainly discoveries waiting beyond the Standard Model. For instance, such speculation inevitably leads to the possibility that quarks and leptons have substructure themselves. Decisive searches for the ultimate building blocks of nature and their interactions will take place at the new colliders of the next century.

Appendix A

CLEO Collaboration

This list of collaborators includes all the members of the CLEO collaboration for the publication of the present analysis in the Physical Review Letters [34].

A. Anastassov,[1] J. E. Duboscq,[1] D. Fujino,[1,1] K. K. Gan,[1] T. Hart,[1] K. Honscheid,[1] H. Kagan,[1] R. Kass,[1] J. Lee,[1] M. B. Spencer,[1] M. Sung,[1] A. Undrus,[1,2] R. Wanke,[1] A. Wolf,[1] M. M. Zoeller,[1] B. Nemati,[2] S. J. Richichi,[2] W. R. Ross,[2] P. Skubic,[2] M. Bishai,[3] J. Fast,[3] J. W. Hinson,[3] N. Menon,[3] D. H. Miller,[3] E. I. Shibata,[3] I. P. J. Shipsey,[3] M. Yurko,[3] S. Glenn,[4] S. D. Johnson,[4] Y. Kwon,[4,3] S. Roberts,[4] E. H. Thorndike,[4] C. P. Jessop,[5] K. Lingel,[5] H. Marsiske,[5] M. L. Perl,[5] V. Savinov,[5] D. Ugolini,[5] R. Wang,[5] X. Zhou,[5] T. E. Coan,[6] V. Fadeyev,[6] I. Korolkov,[6] Y. Maravin,[6] I. Narsky,[6] V. Shelkov,[6] J. Staeck,[6] R. Stroynowski,[6] I. Volobouev,[6] J. Ye,[6] M. Artuso,[7] A. Efimov,[7] M. Goldberg,[7] D. He,[7] S. Kopp,[7] G. C. Moneti,[7] R. Mountain,[7] S. Schuh,[7] T. Skwarnicki,[7] S. Stone,[7] G. Viehhauser,[7] X. Xing,[7] J. Bartelt,[8] S. E. Csorna,[8] V. Jain,[8,4] K. W. McLean,[8] S. Marka,[8] R. Godang,[9] K. Kinoshita,[9] I. C. Lai,[9] P. Pomianowski,[9] S. Schrenk,[9] G. Bonvicini,[10] D. Cinabro,[10] R. Greene,[10] L. P. Perera,[10] G. J. Zhou,[10] B. Barish,[11] M. Chadha,[11] S. Chan,[11] G. Eigen,[11] J. S. Miller,[11] C. O'Grady,[11] M. Schmidtler,[11] J. Urheim,[11] A. J. Weinstein,[11] F. Würthwein,[11] D. W. Bliss,[12] G. Masek,[12] H. P. Paar,[12] S. Prell,[12] V. Sharma,[12] D. M. Asner,[13] J. Gronberg,[13] T. S. Hill,[13] D. J. Lange,[13] S. Menary,[13] R. J. Morrison,[13] H. N. Nelson,[13] T. K. Nelson,[13] C. Qiao,[13] J. D. Richman,[13] D. Roberts,[13] A. Ryd,[13] M. S. Witherell,[13] R. Balest,[14]

B. H. Behrens,[14] W. T. Ford,[14] H. Park,[14] J. Roy,[14] J. G. Smith,[14] J. P. Alexander,[15] C. Bebek,[15] B. E. Berger,[15] K. Berkelman,[15] K. Bloom,[15] D. G. Cassel,[15] H. A. Cho,[15] D. S. Crowcroft,[15] M. Dickson,[15] P. S. Drell,[15] K. M. Ecklund,[15] R. Ehrlich,[15] A. D. Foland,[15] P. Gaidarev,[15] L. Gibbons,[15] B. Gittelman,[15] S. W. Gray,[15] D. L. Hartill,[15] B. K. Heltsley,[15] P. I. Hopman,[15] S. L. Jones,[15] J. Kandaswamy,[15] P. C. Kim,[15] D. L. Kreinick,[15] T. Lee,[15] Y. Liu,[15] N. B. Mistry,[15] C. R. Ng,[15] E. Nordberg,[15] M. Ogg,[15,5] J. R. Patterson,[15] D. Peterson,[15] D. Riley,[15] A. Soffer,[15] B. Valant-Spaight,[15] C. Ward,[15] M. Athanas,[16] P. Avery,[16] C. D. Jones,[16] M. Lohner,[16] C. Prescott,[16] J. Yelton,[16] J. Zheng,[16] G. Brandenburg,[17] R. A. Briere,[17] A. Ershov,[17] Y. S. Gao,[17] D. Y.-J. Kim,[17] R. Wilson,[17] H. Yamamoto,[17] T. E. Browder,[18] Y. Li,[18] J. L. Rodriguez,[18] T. Bergfeld,[19] B. I. Eisenstein,[19] J. Ernst,[19] G. E. Gladding,[19] G. D. Gollin,[19] R. M. Hans,[19] E. Johnson,[19] I. Karliner,[19] M. A. Marsh,[19] M. Palmer,[19] M. Selen,[19] J. J. Thaler,[19] K. W. Edwards,[20] A. Bellerive,[21] R. Janicek,[21] D. B. MacFarlane,[21] P. M. Patel,[21] A. J. Sadoff,[22] R. Ammar,[23] P. Baringer,[23] A. Bean,[23] D. Besson,[23] D. Coppage,[23] C. Darling,[23] R. Davis,[23] N. Hancock,[23] S. Kotov,[23] I. Kravchenko,[23] N. Kwak,[23] S. Anderson,[24] Y. Kubota,[24] S. J. Lee,[24] J. J. O'Neill,[24] S. Patton,[24] R. Poling,[24] T. Riehle,[24] A. Smith,[24] M. S. Alam,[25] S. B. Athar,[25] Z. Ling,[25] A. H. Mahmood,[25] H. Severini,[25] S. Timm,[25] and F. Wappler[25]

[1]Ohio State University, Columbus, Ohio 43210

[2]University of Oklahoma, Norman, Oklahoma 73019

[3]Purdue University, West Lafayette, Indiana 47907

[4]University of Rochester, Rochester, New York 14627

[5]Stanford Linear Accelerator Center, Stanford University, Stanford, California 94309

[6]Southern Methodist University, Dallas, Texas 75275

[7]Syracuse University, Syracuse, New York 13244

[8]Vanderbilt University, Nashville, Tennessee 37235

[9]Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061

[10]Wayne State University, Detroit, Michigan 48202

[11]California Institute of Technology, Pasadena, California 91125

[12]University of California, San Diego, La Jolla, California 92093

[13]University of California, Santa Barbara, California 93106
[14]University of Colorado, Boulder, Colorado 80309-0390
[15]Cornell University, Ithaca, New York 14853
[16]University of Florida, Gainesville, Florida 32611
[17]Harvard University, Cambridge, Massachusetts 02138
[18]University of Hawaii at Manoa, Honolulu, Hawaii 96822
[19]University of Illinois, Champaign-Urbana, Illinois 61801
[20]Carleton University, Ottawa, Ontario, Canada K1S 5B6
and the Institute of Particle Physics, Canada
[21]McGill University, Montréal, Québec, Canada H3A 2T8
and the Institute of Particle Physics, Canada
[22]Ithaca College, Ithaca, New York 14850
[23]University of Kansas, Lawrence, Kansas 66045
[24]University of Minnesota, Minneapolis, Minnesota 55455
[25]State University of New York at Albany, Albany, New York 12222

Appendix B

CLEO Terminology

Detector

PTL	Precision tracking layers
VD	Vertex chamber
DR	Main drift chamber
CD	Central drift chambers (PTL+VD+DR)
CC	Electromagnetic calorimeter
TF	Time-of-flight counters
MU	Muon chambers

General

KLASGL	Event class: < 10 – Bhabhas, mupairs, cosmic and $\gamma\gamma$ events, $= 10$ – Hadronic events, $= 11$ – Beam gas events.
R2GL	Ratio of Fox-Wolfram moments.
EBEAM	Beam energy.
DUET	Tracking program.
CDFT	Calibration program.

Lepton Identification

DPTHMU	Number of nuclear absorption lengths traveled in MU.
MUQUAL	Track quality matching flag between MU and CD.
R2ELEC	Logarithm likelihood ratio for electron.

Track

TRKMAN	Track quality flag.
KINCD	Track identifier:
	$= 0$ – Track from primary vertex,
	$= 2$ – Track from secondary vertex.
DBCD	Distance of closest approach to the interaction point ($r - \phi$ plane).
Z0CD	Z coordinate at the point of closest approach to the origin.
CZCD	$\cos\theta$ of the track ($\theta \equiv$ polar angle).

Vertex

VFINDR	The CLEO vertex finder.
RBMTX	Vertex displacement from the interaction point.
CHITX	χ^2 of vee pointing to main vertex.
COSV0	$\cos(\mathbf{p}, \mathbf{p}_{\mathrm{vertex}})$.
IDTX	Vertex type:
	$= 0$ – Unidentified,
	$= 1$ – γ,
	$= 2$ – K_s^0,
	$= 4$ – Λ,
	$= 8$ – $\bar{\Lambda}$.

Appendix C

Hybrid of the Goity and Roberts Model

In this appendix we give the form factors of the G&R hybrid model used at CLEO. The hybrid model coded in EvT [54] is meant to describe the nonresonant decay $\bar{B} \to D^{(*)}\pi\ell\bar{\nu}_\ell$. We propose a simple extension of the standard G&R model: to remove the $\tilde{D}$ contributions to the rates, we take out from the form factors all the terms with a $\tilde{D}$ propagator.

In the next sections, we describe one by one the differences between the form factors of the standard G&R model and the EvT hybrid. The formulae of Appendix C of Reference [60] are rewritten as described in Section 3.5.5. We refer continually to this publication for equation numbers. The reader should consult it for details on the explicit definition of each form factor. We will therefore employ exactly the same notation as used by Goity and Roberts in their paper.

C.1 The Form Factors for $\bar{B} \to D\pi\ell\bar{\nu}_\ell$

The form factors for $\bar{B} \to D\pi\ell\bar{\nu}_\ell$ are divided in two categories: the nonresonant (NR) and the resonant (R) contributions. Equation (C3) in [60] is replaced by:

$$h_{NR} = \frac{g}{2F_0}\frac{\xi(\nu)}{M_B M_D}\left(\frac{1}{p_\pi\cdot v + \delta m_B - i\epsilon}\right),$$

$$
\begin{aligned}
A_{1\,NR} &= -\frac{g}{2F_0}\,\xi(\nu)\,(1+\nu)\left(\frac{1}{p_\pi\!\cdot\! v+\delta m_B-i\epsilon}\right),\\
A_{2\,NR} &= \frac{g}{2F_0}\,\frac{\xi(\nu)}{M_B}\,\left(\frac{p_\pi\!\cdot\! v+p_\pi\!\cdot\! v'}{p_\pi\!\cdot\! v+\delta m_B-i\epsilon}\right),\\
A_{3\,NR} &= 0.
\end{aligned}
\tag{C.1}
$$

And Equation (C4) becomes:

$$
\begin{aligned}
h_R &= \frac{\alpha_2\;\rho_2(\nu)}{6F_0M_BM_D}\;(\nu-1)\left(\frac{1}{p_\pi\!\cdot\! v+\delta\tilde{m}_2}\right)\\
&+ \frac{\alpha_3}{2F_0}\,\frac{\xi^{(1)}(\nu)}{M_BM_D}\,\left(\frac{1}{p_\pi\!\cdot\! v+\delta\tilde{m}_3}\right),\\
A_{1\,R} &= -\frac{\alpha_2\;\rho_2(\nu)}{6F_0}\,(\nu^2-1)\,\left(\frac{1}{p_\pi\!\cdot\! v+\delta\tilde{m}_2}\right)\\
&- \frac{\alpha_3\,\xi^{(1)}(\nu)}{2F_0}\,(1+\nu)\left(\frac{1}{p_\pi\!\cdot\! v+\delta\tilde{m}_3}\right),\\
A_{2\,R} &= \frac{\alpha_1\;\rho_1(\nu)}{2F_0M_B}\left(\frac{p_\pi\!\cdot\! v}{p_\pi\!\cdot\! v+\delta\tilde{m}_1}\right)\\
&+ \frac{\alpha_2\;\rho_2(\nu)}{F_0M_B}\,\left\{\frac{1}{p_\pi\!\cdot\! v+\delta\tilde{m}_2}\left[\frac{1}{6}(\nu\;p_\pi\!\cdot\! v'-p_\pi\!\cdot\! v)+\frac{1}{3}(p_\pi\!\cdot\! v'-\nu\;p_\pi\!\cdot\! v)\right]\right\}\\
&+ \frac{\alpha_3}{2F_0}\,\frac{\xi^{(1)}(\nu)}{M_B}\,\left(\frac{p_\pi\!\cdot\!(v+v')}{p_\pi\!\cdot\! v+\delta\tilde{m}_3}\right),\\
A_{3\,R} &= -\frac{\alpha_1\;\rho_1(\nu)}{2F_0M_D}\left(\frac{p_\pi\!\cdot\! v}{p_\pi\!\cdot\! v+\delta\tilde{m}_1}\right)\\
&- \frac{\alpha_2\;\rho_2(\nu)}{F_0M_D}\,\left\{\frac{1}{2}\,\frac{1}{p_\pi\!\cdot\! v+\delta\tilde{m}_2}(p_\pi\!\cdot\! v'-\nu\;p_\pi\!\cdot\! v)\right\}.
\end{aligned}
\tag{C.2}
$$

C.2 The Form Factors for $\bar{B} \to D^*\pi\ell\bar{\nu}_\ell$

For $B \to D^*\pi\ell\bar{\nu}_\ell$, Equation (C7) with the NR contribution becomes

$$
\begin{aligned}
h_{1\,NR} &= -\frac{g\xi(\nu)}{F_0M_BM_D}\;\frac{p_\pi\!\cdot\! v}{p_\pi\!\cdot\! v+\delta m_B-i\epsilon},\\
h_{2\,NR} &= -\frac{g\xi(\nu)}{F_0M_B}\;\frac{1}{p_\pi\!\cdot\! v+\delta m_B-i\epsilon},
\end{aligned}
$$

$$
\begin{aligned}
h_{3\,NR} &= -\frac{g\xi(\nu)}{F_0 M_D}\left(\frac{1}{p_\pi\!\cdot\! v+\delta m_B - i\epsilon} - \frac{1+\nu}{p_\pi\!\cdot\! v' + i\epsilon}\right),\\
f_{1\,NR} &= -\frac{g\xi(\nu)}{2F_0 M_B}\left(\frac{1}{p_\pi\!\cdot\! v+\delta m_B - i\epsilon} - \frac{1}{p_\pi\!\cdot\! v' +\delta m_D + i\epsilon}\right),\\
f_{2\,NR} &= \frac{M_B}{M_D}\, f_{1\,NR},\\
f_{3\,NR} &= f_{4\,NR} = 0,\\
f_{5\,NR} &= \frac{g\xi(\nu)}{2F_0 M_B M_D}\left(1+\frac{p_\pi\!\cdot\! v}{p_\pi\!\cdot\! v+\delta m_B - i\epsilon}\right),\\
f_{6\,NR} &= \frac{g\xi(\nu)}{2F_0 M_B}\left(\frac{1}{p_\pi\!\cdot\! v+\delta m_B - i\epsilon} - \frac{1}{p_\pi\!\cdot\! v' + i\epsilon}\right),\\
k_{NR} &= \frac{g\xi(\nu)}{2\,F_0}\left(\frac{p_\pi\!\cdot\! v' - \nu\, p_\pi\!\cdot\! v}{p_\pi\!\cdot\! v+\delta m_B - i\epsilon} + \frac{p_\pi\!\cdot\! v - \nu\, p_\pi\!\cdot\! v'}{p_\pi\!\cdot\! v' + i\epsilon}\right),\\
g_{1\,NR} &= 0,\\
g_{2\,NR} &= 0,\\
g_{3\,NR} &= 0,\\
g_{4\,NR} &= \frac{g\xi(\nu)}{F_0\, M_D}\;\frac{1}{p_\pi\!\cdot\! v' + i\epsilon},\\
g_{5\,NR} &= 0.
\end{aligned}
\tag{C.3}
$$

As one may note, (C7) in [60] is indeed unchanged because it contains no pole in any of the form factors. Finally, Equation (C8) is replaced by

$$
\begin{aligned}
h_{1\,R} &= \frac{\alpha_1\;\rho_1(\nu)}{F_0 M_B M_D}\left(\frac{p_\pi\!\cdot\! v}{-p_\pi\!\cdot\! v - \delta\tilde m_1}\right)\\
&+ \frac{\alpha_2\;\rho_2(\nu)}{3F_0 M_B M_D}\left(\frac{p_\pi\!\cdot\! v\,(1+2\nu) - p_\pi\!\cdot\! v'}{p_\pi\!\cdot\! v + \delta\tilde m_2}\right)\\
&- \frac{\alpha_3\;\xi^{(1)}(\nu)}{F_0 M_B M_D}\frac{p_\pi\!\cdot\! v}{p_\pi\!\cdot\! v + \delta\tilde m_3},\\
h_{2\,R} &= \frac{\alpha_2\;(1+\nu)\;\rho_2(\nu)}{3F_0 M_B}\frac{1}{-p_\pi\!\cdot\! v - \delta\tilde m_2} - \frac{\alpha_3\;\xi^{(1)}(\nu)}{F_0 M_B\;(p_\pi\!\cdot\! v + \delta\tilde m_3)},\\
h_{3\,R} &= \frac{\alpha_2\;\rho_2(\nu)}{3F_0 M_D}\left(\frac{1+\nu}{p_\pi\!\cdot\! v + \delta\tilde m_2}\right)\\
&- \frac{\alpha_3\;\xi^{(1)}(\nu)}{F_0 M_D}\left(\frac{1}{p_\pi\!\cdot\! v + \delta\tilde m_3}\right),
\end{aligned}
$$

$$
\begin{aligned}
f_{1\,R} &= -\frac{\alpha_2\ \rho_2(\nu)\,(\nu-1)}{6F_0 M_B}\ \left(\frac{1}{p_\pi\!\cdot\! v+\delta\tilde{m}_2}\right) \\
&- \frac{\alpha_3\ \xi^{(1)}(\nu)}{2F_0 M_B}\ \left(\frac{1}{p_\pi\!\cdot\! v+\delta\tilde{m}_3}\right), \\
f_{2\,R} &= \frac{M_B}{M_D}\ f_{1\,R}, \\
f_{3\,R} &= f_{4\,R} = 0, \\
f_{5\,R} &= \frac{\alpha_1\ \rho_1(\nu)}{2F_0 M_B M_D}\ \left(\frac{p_\pi\!\cdot\! v}{p_\pi\!\cdot\! v+\delta\tilde{m}_1}\right) \\
&+ \frac{\alpha_2\ \rho_2(\nu)}{2F_0 M_B M_D}\ \left(\frac{p_\pi\!\cdot\! v' - \frac{1}{3}p_\pi\!\cdot\! v\,(1+2\,\nu)}{p_\pi\!\cdot\! v+\delta\tilde{m}_2}\right) \\
&+ \frac{\alpha_3\ \xi^{(1)}(\nu)}{2F_0 M_B M_D}\ \left(\frac{p_\pi\!\cdot\! v}{p_\pi\!\cdot\! v+\delta\tilde{m}_3}\right), \\
f_{6\,R} &= \frac{\alpha_2\ \rho_2(\nu)\,(\nu-1)}{6F_0 M_B}\ \left(\frac{1}{p_\pi\!\cdot\! v+\delta\tilde{m}_2}\right) \\
&+ \frac{\alpha_3\ \xi^{(1)}(\nu)}{2F_0 M_B}\ \left(\frac{1}{p_\pi\!\cdot\! v+\delta\tilde{m}_3}\right), \\
k_R &= -\frac{\alpha_1\ \rho_1(\nu)\,(\nu-1)}{2F_0}\ \left(\frac{p_\pi\!\cdot\! v}{p_\pi\!\cdot\! v+\delta\tilde{m}_1}\right) \\
&- \frac{\alpha_2\ \rho_2(\nu)\,(\nu-1)}{3F_0}\ \left(\frac{(p_\pi\!\cdot\! v' - \nu\ p_\pi\!\cdot\! v)}{p_\pi\!\cdot\! v+\delta\tilde{m}_2}\right) \\
&+ \frac{\alpha_3\ \xi^{(1)}(\nu)}{2F_0}\ \left(\frac{(p_\pi\!\cdot\! v' - \nu\ p_\pi\!\cdot\! v)}{p_\pi\!\cdot\! v+\delta\tilde{m}_3}\right), \\
g_{1\,R} &= 0, \\
g_{2\,R} &= 0, \\
g_{3\,R} &= 0, \\
g_{4\,R} &= \frac{\alpha_2\ \rho_2(\nu)}{3F_0\,M_D}\ \left(\frac{2}{p_\pi\!\cdot\! v+\delta\tilde{m}_2}\right), \\
g_{5\,R} &= 0.
\end{aligned}
\tag{C.4}
$$

Appendix D

Piecewise Linear Fit

The piecewise linear method is used to fit a continuous function or distribution with a variable number of connected straight line segments of equal width. Each segment is parameterized by a slope and a y-intercept. The slope and the intercept are calculated with a weighted least-squares fit of the data points within the bin they parameterize. Briefly, the piecewise method may be described as follows: Consider a distribution of n data points $(x_1, y_1 \pm \sigma_1), (x_2, y_2 \pm \sigma_2), \ldots, (x_n, y_n \pm \sigma_n)$ that we want to fit into k $(k < n)$ connected line segments with $m = n/k$ points per segment. The condition that the segments have to be connected reduces the number of free parameters to just the slopes of the line segments and the y-intercept of the first segment. All the other intercepts can be obtained from the first intercept and the preceding slopes. Let L be the width of each segment. The functional dependence can be written as:

$$f = f(y_0, \theta_1, \theta_2, \ldots, \theta_k; x) \tag{D.1}$$

where

$$f = \begin{cases} y_0 + x\theta_1 & \text{if } 0 \leq x \leq L \\ y_0 + L\theta_1 + (x - L)\theta_2 & \text{if } L \leq x \leq 2L \\ \cdots & \\ y_0 + L(\theta_1 + \theta_2 + \cdots + \theta_{k-1}) + (x - (k-1)L)\theta_k & \text{if } (k-1)L \leq x \leq kL \end{cases}$$

According to the least squares principle, the best values of the unknown parameters are those which minimize

$$\chi^2 = \sum_{i=1}^{n} \{y_i - f(x_i)\}/\sigma_i^2 \tag{D.2}$$

In matrix notation

$$\vec{f} = A\vec{\Theta} \tag{D.3}$$

where A is a $(k+1) \times n$ matrix and $\vec{\Theta}$ is a $(k+1)$ dimensional vector.

$$\vec{\Theta} = \begin{pmatrix} y_0 \\ \theta_1 \\ \theta_2 \\ \vdots \\ \theta_k \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & x_1 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_m & 0 & 0 & \cdots & 0 \\ 1 & L & x_{m+1} - L & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & L & x_{2m} - L & 0 & \cdots & 0 \\ 1 & L & L & x_{2m+1} - 2L & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & L & L & L & \cdots & x_n - (k-1)L \end{pmatrix}$$

The solution for $\vec{\Theta}$ is:

$$\vec{\Theta} = (A^T V^{-1} A)^{-1} A^T V^{-1} \vec{y} \tag{D.4}$$

with

$$\vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

and V = Diag(σ_i^2). The final solution for the slopes and the intercepts of the piecewise fit is $\vec{\Theta}$.

Appendix E

Fake Muon Study

The experimental study of B semileptonic decays at CLEO II is of basic importance in understanding the weak interaction in the framework of the Standard Model. Muon detection and identification is a key factor in such studies. In this appendix, we present the results of muon fake rates using the MUTR package. This study was essential in the determination of the fake muon background yields for the measurements of $\mathcal{B}(B^- \to D_1^0 \ell^- \bar{\nu}_\ell)$ and $\mathcal{B}(B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell)$.

E.1 Data Sample for the Fake Muon Study

The full data set available for this fake rate study correspond to an integrated luminosity of 3.11 fb^{-1} on the $\Upsilon(4S)$ resonance, and 1.61 fb^{-1} at a center-of-mass energy $\sim$ 55 MeV below the $\Upsilon(4S)$ resonance (*i.e.*, pre-recompress 4S2 through 4SG). This large sample, collected with the CLEO II detector, allows us to use tight requirements for our hadron selection, thereby providing clean hadron samples with high statistics.

E.2 Hadron Abundances and Fake Rates

Fake muons are hadrons misidentified as muons. The fake rates are the probabilities of misidentification. The fake rate probabilities for muons are much higher than for electrons. The fake rates for pions ($\mathcal{F}_\pi$), kaons ($\mathcal{F}_K$), and protons ($\mathcal{F}_p$) depend on the momentum, the charge, and the penetration depth of the particle. To determine the overall background from muon fakes from charged particles at the $\Upsilon(4S)$ resonance, we need to know the individual misidentification probabilities for pions, kaons and protons, and their relative abundances.

We determined the hadron abundances, Y_i with $i = \pi, K$ or p, from Monte Carlo simulation. The generic $B\bar{B}$ Monte Carlo has been adjusted to simulate adequately the abundances for every particle species. The average hadron abundances, for positively and negatively charged hadrons produced at the $\Upsilon(4S)$, is shown in Figure E.1 as a function of the hadron momentum.

Fake rate probabilities are poorly modeled [98] in the GEANT based simulation of the CLEO II detector. Hence, we need to develop a method for selecting pure samples of hadrons from data. To determined the fake rates ($\mathcal{F}_i$), pions from $K_s^0 \to \pi^+\pi^-$, kaons from $D^0 \to K^-\pi^+$, and protons from $\Lambda^0 \to p\pi^-$, are selected. Depending on their momentum, both the K_s and Λ particles can travel several centimeters into the central detector before decaying to the observable final state. This makes it possible to reconstruct their secondary vertices by pairing two opposite charged particles reconstructed in the CLEO II tracking system. For the D^0, we use the usual D^{*+} trick, where $\delta m = M(D^*) - M(D)$ is well known. By reconstructing the decay chain $D^{*+} \to D^0\pi^-$ with $D^0 \to K^-\pi^+$ [19], the charge of the pion from the D^{*+} uniquely tags the charged tracks from the D^0 meson.

E.3 Hadron Selection

The method developed to measure the fake rate probabilities uses samples of pions, kaons, and protons with a small contamination from other particles. In this section, we will first give some general track selection criteria, and then enumerate the individual selection criteria for our pion, kaon, and proton samples. The CLEO terminology can be found in Appendix B.

E.3.1 General Track Selection Criteria

We obtain $B\bar{B}$ events by selecting hadronic events (*i.e.*, KLASGL=10) with R2GL < 0.4 and with at least four good tracks. Each charged track used for the fake analysis must have:

- Passed the TRKMAN criteria
- $|\mathrm{CZCD}| < 0.85$ for 1.0 GeV/c $\leq |\mathbf{p}| <$ 1.5 GeV/c
- $|\mathrm{CZCD}| < 0.82$ for 1.5 GeV/c $\leq |\mathbf{p}| \leq$ 2.0 GeV/c

For each hadron hypothesis i (with $i = \pi, K$ or p), we compute for each charged track the probability, $P(\chi_i^2)$, and the likelihood, ℓh_i.

$$\chi_i^2 = (\chi_i^{dE/dx})^2 + (\chi_i^{\mathrm{TOF}})^2. \tag{E.1}$$

The calculated values of χ_i^2 are converted into the probability $P(\chi_i^2)$ which is based on either one or two independent PID measurements, depending whether the track has dE/dx and/or time-of-flight information. Then,

$$lh_i = \frac{n_i P(\chi_i^2)}{\sum n_j P(\chi_j^2)}, \tag{E.2}$$

where the factors n_j are the relative abundances for each hadron. We use the approximation $n_\pi = 0.78$, $n_K = 0.20$, and $n_p = 0.02$. We force every track under a given hypothesis to have $lh_i > 0.01$.

Here we rely on both the dE/dx and TOF information for the hadron identification since this study does not depend on the MC simulation of the TOF counters. Hadron identification based on dE/dx is described in Section 6.7. The hadron identification capability of the TOF system is summarized in Section 4.4.

E.3.2 Pion sample

We identify pions from $K_s^0 \to \pi^+\pi^-$ secondary vertices using the CLEO vertex finder VFINDR. We require the invariant mass difference $|M(\pi^+\pi^-) - m_{K_s}|$ to be less than 8 MeV/c^2. The vertex and track criteria for selecting pions are:

- IDTX = 2
- KINCD = 2
- DBCD $\geq$ 0.001 m
- CHITX $\leq$ 3
- RBMTX $\geq$ 0.01 m
- COSV0 $\geq$ 0.95

After applying all the selection criteria, the combinatoric background is negligible, as can be seen in Figure E.2(a), so that the contamination of the pion sample is negligible.

E.3.3 Proton Sample

Protons are selected from the decay of the long-lived Λ particles, which decay 63.9% of the time to a proton and a pion. We require the invariant mass difference $|M(p\pi^-) - m_\Lambda|$ to be less than 3 MeV/c^2. The vertex and track criteria for selecting $p\pi$ pairs are:

- IDTX = 4 or 8
- KINCD = 2
- DBCD $\geq$ 0.001 m
- CHITX $\leq$ 3
- RBMTX $\geq$ 0.01 m
- COSV0 $\geq$ 0.95

As for the K_s, the detached vertex of the Λ provides a clean sample of protons. The invariant mass distribution for our Λ candidates is shown in Figure E.2(b).

E.3.4 Kaon Sample

We select kaons in the decay chain $D^{*+} \to D^0\pi^-_{\text{slow}}$, where $D^0 \to K^-\pi^+$. First, the $K^-\pi^+$ combination is required to have an invariant mass within 16 MeV/c^2 ($\sim 2\sigma$) of the nominal D^0 mass. The reconstructed mass difference $\delta m = M(D^0\pi^+) - M(D^0)$ is required to be within 1.5 MeV/c^2 ($\sim 1.5\sigma$) of the known $D^{*+} - D^0$ mass difference. Besides the mass cuts, we require good PID on the pion from the D^{*+}. The momentum of the pion from the D^{*+} is in the range where PID provides good separation. Thus, we require $\ell h_{\pi_{\text{slow}}} > 0.1$. The vertex and track quality cuts for the D^{*+} daughters are:

- KINCD = 0
- DBCD $\leq$ 0.005 m and Z0CD $\leq$ 0.05 m

By requiring the cuts listed above on the D^0 candidates, we suppress the combinatoric background sufficiently (see Figures E.2(c) and E.2(d)).

E.4 Results

For a given hadron hypothesis, the fake rate is defined as the ratio of the number of tracks identified as a muon in the fiducial volume of the muons chambers ($N_{i\to\mu}$) over the total number of tracks in the same fiducial volume (N_i):

$$\mathcal{F}_i = N_{i\to\mu}/N_i. \tag{E.3}$$

The fake rates are calculated in different momentum bins for positively and negatively charged particles separately. Hadron are selected with the criteria described in the previous section. Fake muons are those particles which pass the hadron track criteria and the muon acceptance cuts:

- MUQUAL = 0
- For 1.0 GeV/$c \leq |\mathbf{p}| <$ 1.5 GeV/c: $3 \leq$ DPTHMU < 5
- For 1.5 GeV/$c \leq |\mathbf{p}| \leq$ 2.0 GeV/c: DPTHMU ≥ 5

The individual muon misidentification probabilities are shown in Figures E.3, E.4, and E.5. In each figure there are two plots, one for positively charged particles, and one for the negatively charged particles. The errors on the fake probabilities are statistical only.

E.5 Consistency Checks

A similar fake rate study for pions and kaons was performed in the early stage of CLEO II [121]; the two results are in very good agreement. As a consistency check, we computed the muon fake rates for pions with the D^0 sample. The agreement between the two measurements is reasonable. Of course the results from the K_s sample have smaller statistical errors. In Figure E.6, the fake rates are overlaid for both pion samples.

As expected, we observe a significant difference between the fake rates for K^+ and K^-. The asymmetry is due to the cross-section difference, $\sigma(K^- p) > \sigma(K^+ p)$, and to the larger nuclear capture probability for negatively charged kaons. Positively charged pions also have a somewhat higher fake probabilities than negatively charged pions.

Other studies of fake muons from pions have been done by other collaborators [98, 122]. They obtained a clean sample of pions from the decay chain $\tau^+ \to \rho^+\nu$, with $\rho^+ \to \pi^+\pi^0$ ($\mathcal{B}(\tau^+ \to \rho^+\nu) \simeq 25\%$). Tau pair candidates are tagged by selecting events in which one of the tau decays into a muon plus neutrinos. The π^+ is tagged by first reconstructing the π^0. Their fake rates are larger than ours ($\Delta\mathcal{F}_\pi/\mathcal{F}_\pi \lesssim 10\% - 15\%$), which could be explained by an admixture of kaons. They quote a 2% admixture of kaons in their pion sample due to the Cabibbo suppressed decay $\tau^+ \to K^{*+}\nu$ when $K^{*+} \to K^+\pi^0$. We estimate our overall contamination to be less than 1%. The agreement is satisfactory when we consider the difference in topology between $\tau^+\tau^-$ and $B\bar{B}$ events.

The contamination of our kaon sample from combinatoric background is somewhat larger than the contamination in our pion and proton samples (see Figure E.2). As a cross-check, we estimated the contamination using a D^0 sideband and re-calculated the fake rates $\mathcal{F}_K$. No major discrepancy was found between the $\mathcal{F}_K$ calculated with a D^0 sideband subtraction and $\mathcal{F}_K$ given in Figure E.4.

E.6 Systematic Uncertainty

The determination of the fake rates using the method described in the previous sections deals only approximately with kaons or pions decaying in flight to muons within the tracking chambers. Based on a study described in [121], the discrepancies are expected to be very small however. Another effect that we neglected in the present study is the dependence of the fake probabilities on the polar angle of the track. This question was raised by the Systematic Advisory Committee (SAC) [98] at CLEO. The fake probabilities show no dependence within the barrel region; outside the barrel region, they are somewhat lower, since there is more shielding. This is believed to be a small effect because the distribution of polar angle of the non-leptons used for the fake background study is very similar to that of the lepton in the data events. A conservative systematic error of $\pm 25\%$ on the fake rates should be used.

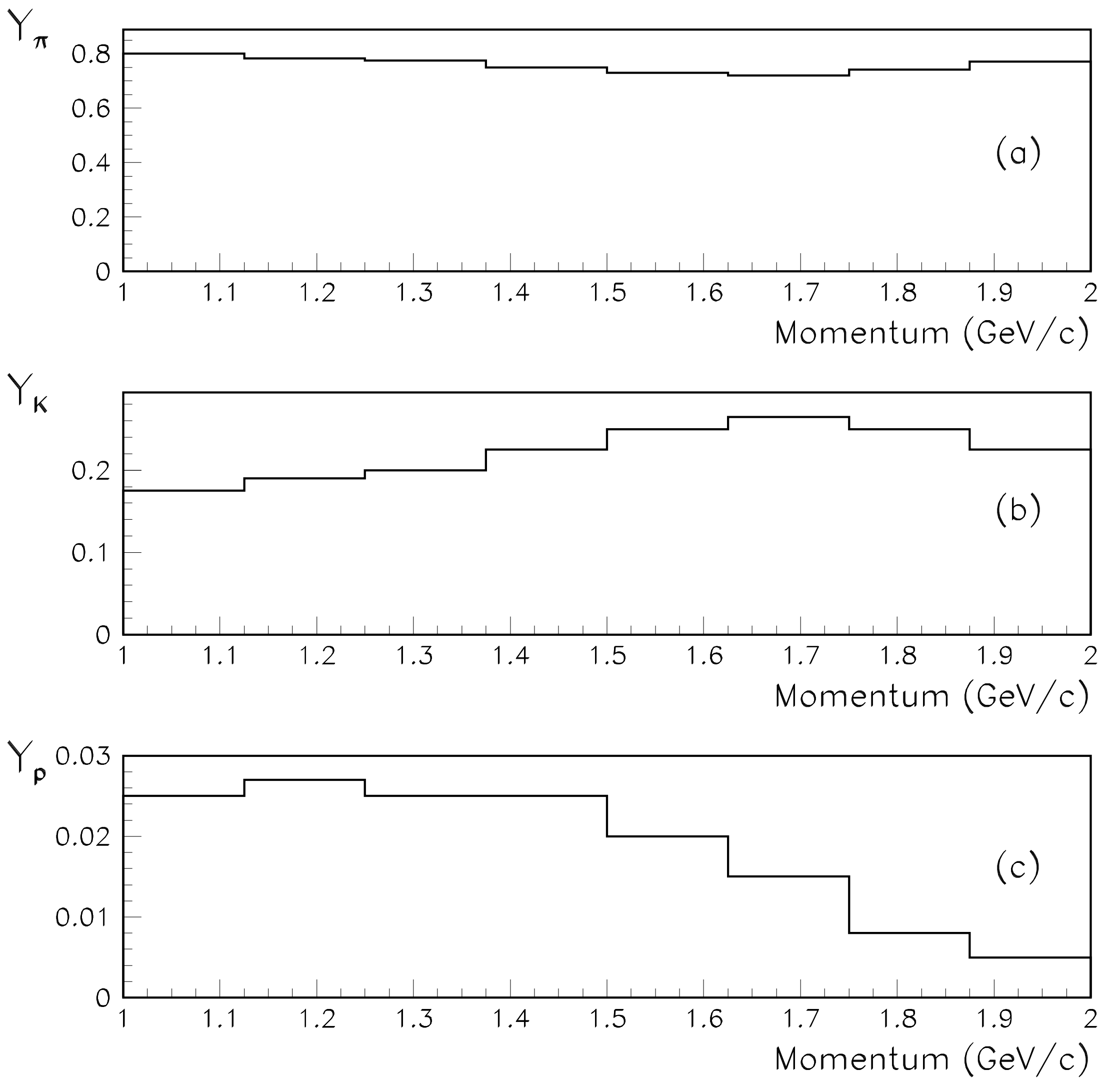


Figure E.1: Particle abundances at the $\Upsilon(4S)$ as a function of the momentum. The distributions are from MC simulation. We assume $Y_i^+ = Y_i^-$.

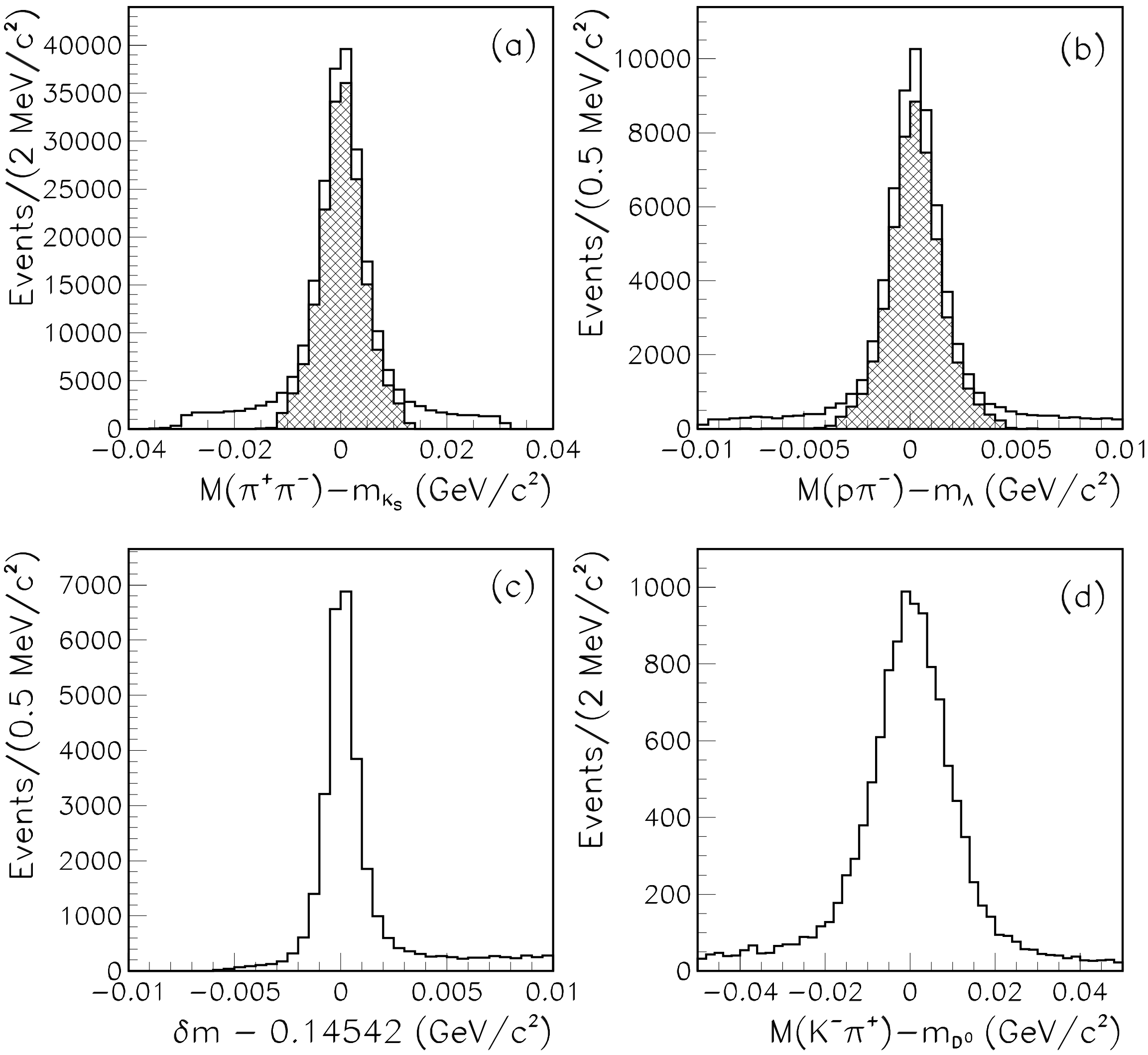


Figure E.2: Distribution of invariant mass differences: (a) $M(\pi^+\pi^-) - m_{K_s}$ and (b) $M(p\pi^-) - m_\Lambda$ before (unshaded) and after (hatched) the application of the cut for a detached vertex (KINCD = 2). The requirement that K_s and Λ must originate from a secondary vertex leads to negligible background. The mass difference δm is shown in (c), and the invariant mass of our D^0 candidates is shown in (d). Again, there is little combinatoric background under the mass peaks.

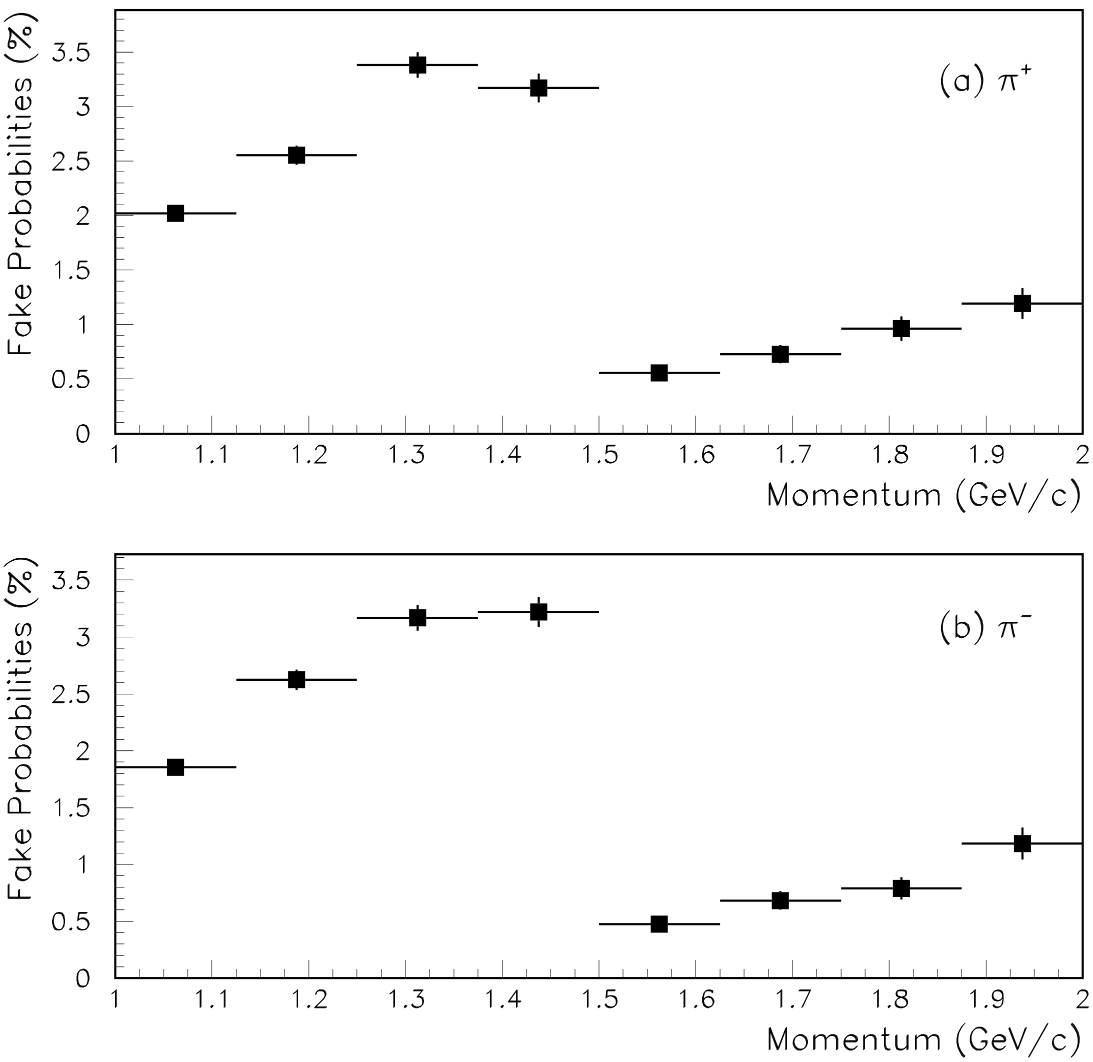


Figure E.3: Fake probabilities for misidentifying pions as muons as a function of momentum. Fake muons must have $3 \leq$ DPTHMU < 5 for 1.0 GeV/$c \leq |\mathbf{p}| < 1.5$ GeV/c, and DPTHMU ≥ 5 for 1.5 GeV/$c \leq |\mathbf{p}| \leq 2.0$ GeV/c.

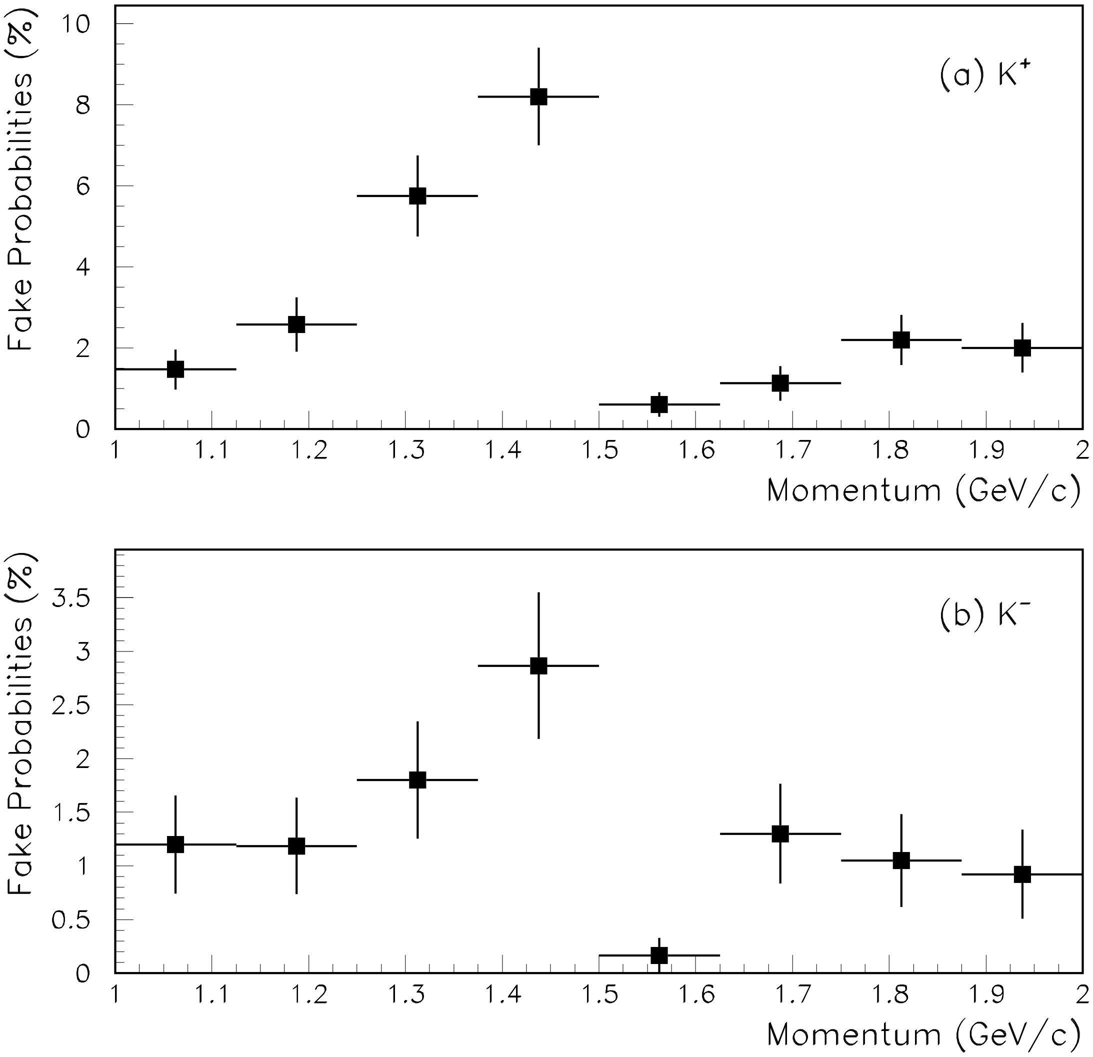


Figure E.4: Fake probabilities for misidentifying kaons as muons as a function of momentum. Fake muons must have $3 \leq$ DPTHMU < 5 for 1.0 GeV/$c \leq |\mathbf{p}| < 1.5$ GeV/c, and DPTHMU ≥ 5 for 1.5 GeV/$c \leq |\mathbf{p}| \leq 2.0$ GeV/c.

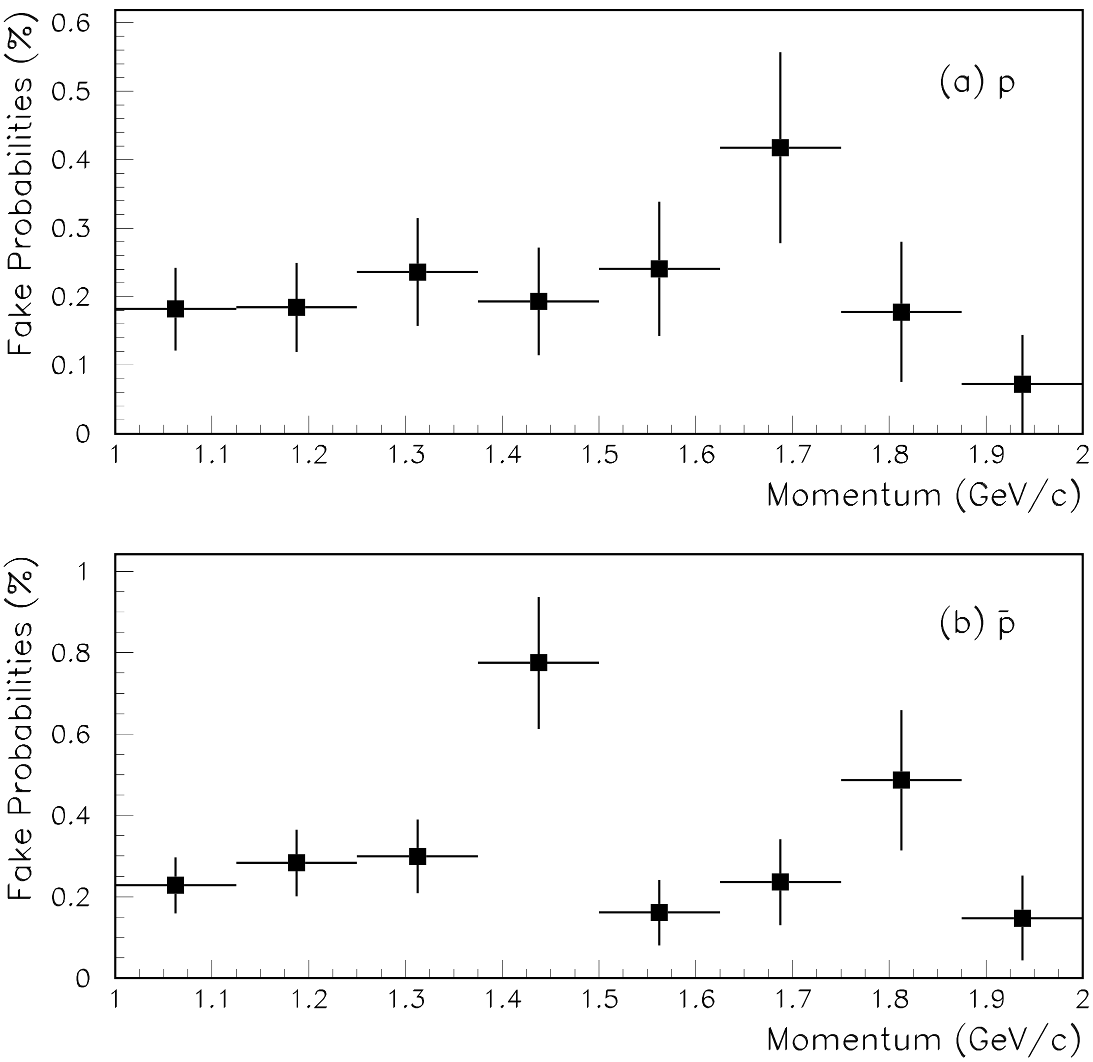


Figure E.5: Fake probabilities for misidentifying protons as muons as a function of momentum. Fake muons must have $3 \leq$ DPTHMU < 5 for 1.0 GeV/$c \leq |\mathbf{p}| < 1.5$ GeV/c, and DPTHMU ≥ 5 for 1.5 GeV/$c \leq |\mathbf{p}| \leq 2.0$ GeV/c.

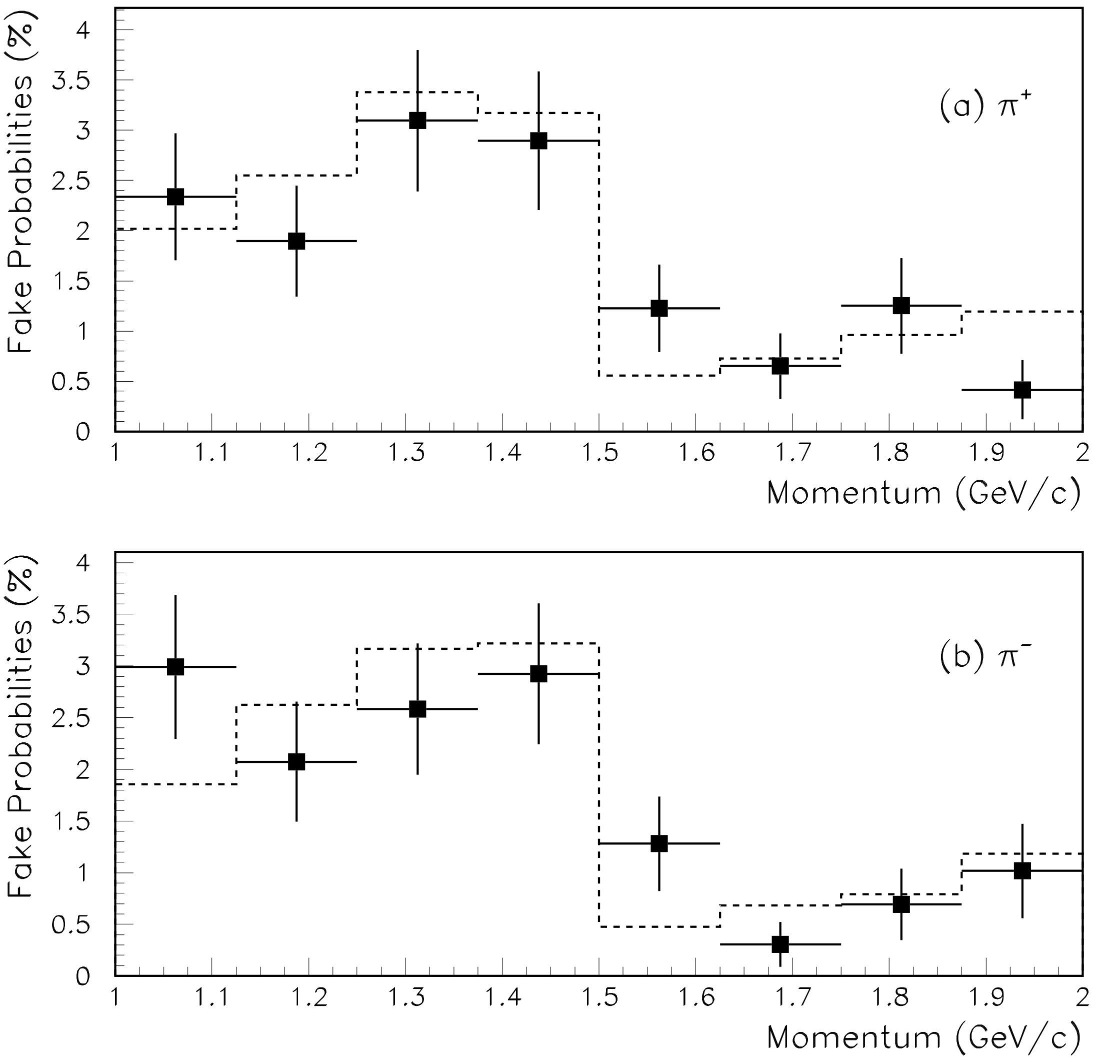


Figure E.6: Consistency check for pion fake rates. Fake probabilities for misidentifying pions as muons as a function of momentum. Fake muons must have $3 \leq$ DPTHMU $<$ 5 for 1.0 GeV/$c \leq |\mathbf{p}| <$ 1.5 GeV/c, and DPTHMU ≥ 5 for 1.5 GeV/$c \leq |\mathbf{p}| \leq$ 2.0 GeV/c. The solid squares are from the D^0 sample and the dashed lines are from the K_s^0 sample. The difference between the two samples leads to $\chi^2/d.o.f = 1.3$ (20% probability) for both (a) and (b). $\chi^2/d.o.f = \frac{1}{N}\sum[\mathcal{F}_\pi(D) - \mathcal{F}_\pi(K_s)]^2/\sigma^2(D)$, where N is the number of data points per histogram.

Bibliography

[1] A general introduction to particle physics can be found in:
D. H. Perkins, "Introduction to High Energy Physics", Addison Wesley (1987);
B. R. Martin and G. Shaw, "Particle Physics", John Wiley & Sons (1992).

[2] V. V. Ezhela *et al.*, "Particle Physics: One Hundred Years of Discoveries", Springer (1996).

[3] For a marvelous review on HQET consult:
M. Neubert, Phys. Rep. **245**, 259 (1994).

[4] A more advanced description of the fundamental concepts in modern particle physics can be found in:
F. Halzen and A. D. Martin, "Quarks and Leptons", John Wiley & Sons (1984);
F. Mandl and G. Shaw, "Quantum Field Theory", John Wiley & Sons (1984).

[5] S. Weinberg, A. Salam, and S. Glashow, Rev. Mod. Phys. **52**, 515 (1980);
S. L. Glashow, Nucl. Phys. **22**, 579 (1961);
A. Salam and J. C. Ward, Phys. Lett. **13**, 168 (1964);
S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967).

[6] Particle Data Group, R.M. Barnett *et al.*, Phys. Rev. D **54**, 1 (1996).

[7] The limits on the neutrino masses are: $m_{\nu_e} < 15$ eV, $m_{\nu_\mu} < 170$ KeV, and $m_{\nu_\tau} < 24$ MeV from Reference [6].

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

[9] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963).

[10] L. Wolfenstein, Phys. Rev. Lett. **51**, 1945 (1983);
L. Wolfenstein, Ann. Rev. Nucl. Part. Sci. **36**, 137 (1986).

[11] For a review on B meson decays consult:
"B decays", 2^{nd} Edition, edited by S. Stone, World Scientific, Singapore (1993).

[12] For a review on CP violation consult:
"CP Violation", edited by C. Jarlskog, World Scientific, Singapore (1989).

[13] C. Dib, I. Dunietz, F. Gilman, and Y. Nir, Phys. Rev. D **41**, 1522 (1990);
F. Gilman and Y. Nir, Ann. Rev. Nucl. Part. Sci. **40**, 213 (1990);
D. London, Mod. Phys. Lett. A **5**, 337 (1990).

[14] For a recent review consult:
A. Buras, Proceedings of "Beauty 95" meeting, Oxford, hep-ph/9509329.

[15] D. G. Cassel, in the Proceedings of the Frontiers in Contemporary Physics, Nashville, TN, 1997 (unpublished).

[16] P. S. Drell, "Experimental Aspects of the Standard Model: A Short Course for Theorists", Cornell preprint CLNS 96-1453, Lectures given at Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 96), Boulder, Colorado, June 1996.

[17] A great review of leptonic and semileptonic decays is giving in:
J. D. Richman and P. R. Burchat, Rev. Mod. Phys. **67**, 893 (1995).

[18] Here and throughout this paper, the symbol for lepton (ℓ) refers to electron (e) or muon (μ) unless otherwise specified.

[19] Charged conjugate states are implied if not stated otherwise.

[20] P. S. Drell, in the Proceedings of the XVIII International Symposium on Lepton-Photon Interactions, Hamburg, 1997 (unpublished).

[21] G. Altarelli, N. Cabibbo, G. Corbò, L. Maiani, G. Martinelli, Nucl. Phys. B **208**, 365 (1982).

[22] N. Isgur, D. Scora, B. Grinstein, M. B. Wise, Phys. Rev. D **39**, 799 (1989);
N. Isgur and M. B. Wise, Phys. Rev. D **41**, 151 (1990).

[23] ISGW** is simply the ISGW model in which the amount of excited charm mesons D^{**} is allowed to float.

[24] CLEO Collaboration, B. Barish *et al.*, Phys. Rev. Lett. **76**, 1570 (1996).

[25] ARGUS Collaboration, H. Albrecht, Phys. Lett. B **318**, 397 (1993).

[26] Theoretical review on the subject can be found in:
A. J. Buras, "CKM (Present and Future)", in the Proceedings of the 7^{th} International Symposium on Heavy Flavor Physics, Santa Barbara, 1997 (unpublished).
An experimental review and physics potential of future e^+e^- facilities can be found in the following references:
CLEO Collaboration, Cornell preprint CLNS 94-1277 (1994);
BaBar Collaboration, SLAC report SLAC-443 (1994);
Belle Collaboration, KEK report 94-2 (1994).

[27] ARGUS Collaboration, Z. Phys. C **57**, 533 (1997).

[28] CLEO Collaboration, B. Barish *et al.*, Phys. Rev. D **51**, 1014 (1995).

[29] ALEPH Collaboration, D. Buskulic *et al.*, Z. Phys. C **73**, 601 (1997);
ALEPH Collaboration, D. Buskulic *et al.*, Phys. Lett. B **345**, 103 (1995).

[30] OPAL Collaboration, R. Akers *et al.*, Z. Phys. C **69**, 57 (1995).

[31] DELPHI Collaboration, P. Abreu etal, Z. Phys. C **71**, 539 (1996).

[32] CLEO Collaboration, J. P. Alexander *et al.*, CLEO-CONF 95-30, EPS-0168 (1995).

[33] CLEO Collaboration, T. E. Browder *et al.*, CLEO-CONF 96-02, ICHEP96 pa05-077 (1996).

[34] CLEO Collaboration, A. Anastassov *et al.*, Cornell preprint CLNS 97-1501, CONF 97-21, EPS-0341 (1997).

[35] R. Wang, Ph.D. thesis, University of Minnesota (1994).

[36] A. K. Leibovich, Z. Ligeti, I. W. Stewart, and M. B. Wise, Phys. Rev. Lett. **78**, 3995 (1997).

[37] E. V. Shuryak, Nucl. Phys. B **93**, 134 (1980);
M. B. Voloshin and M. A. Shiftman, Sov. J. Nucl. Phys. **47**, 511 (1988).

[38] N. Isgur and M. B. Wise, Phys. Lett. B **232**, 113 (1989);
N. Isgur and M. B. Wise, Phys. Lett. B **237**, 527 (1990);
A. F. Falk, H. Georgi, B. Grinstein, and M. B. Wise, Nucl. Phys. B **343**, 1 (1990).

[39] CLEO Collaboration, P. Avery *et al.*, Phys. Lett. B **331**, 236 (1994); Erratum *ibid.* Phys. Lett. B **342**, 453 (1995).

[40] N. Isgur and M. B. Wise, Phys. Rev. Lett. **66**, 1130 (1991).

[41] M. Lu, M. B. Wise, and N. Isgur, Phys. Rev. D **45**, 1553 (1992).

[42] S. Veseli and M. G. Olsson, Phys. Rev. D **54**, 886 (1996).

[43] A. Falk, M. Neubert, and M. Luke, Nucl. Phys. B **388**, 3363 (1992);
T Mannel, Phys. Rev. D **50**, 428 (1994).

[44] M. E. Luke, Phys. Lett. B **252**, 447 (1990).

[45] I. Caprini and M. Neubert, Phys. Lett. B **380**, 376 (1996).

[46] ALEPH Collaboration, D. Buskulic *et al.*, Phys. Lett. B **395**, 373 (1997).

[47] CLEO Collaboration, M. Athanas *et al.*, Cornell preprint CLNS CLNS 97-1486, CONF 97-02, EPS-0342 (1997).

[48] CLEO Collaboration, J. E. Duboscq *et al.*, Phys. Rev. Lett. **76**, 3898 (1996); CLEO Collaboration, A. Anastassov *et al.*, CLEO-CONF 96-08, ICHEP96 pa05-079 (1996).

[49] "Gauge Bosons and Heavy Quarks", SLAC report SLAC-378, in the Proceedings of the 18^{th} SLAC Summer Institute of Particle Physics, Stanford, 1990, edited by J. F. Hawthorne.

[50] N. Isgur and M. B. Wise, Phys. Rev. D **43**, 819 (1991).

[51] D. Scora and N. Isgur, Phys. Rev. D **52**, 2783 (1995).

[52] A. K. Leibovich, Z. Ligeti, I. W. Stewart, and M. B. Wise, CALT 68-2120 (hep-ph 9705467), 1997.

[53] See Appendix in Reference [55] for the transformation formulae.

[54] D. Lange, A. Ryd, V. Boisvert, and A. Bellerive, CLEO internal report CBX (in progress).

[55] T. B. Suzuki, T. Ito, S. Sawada, and M. Matsuda, Prog. Theor. Phys. **91**, 757 (1994).

[56] P. Colangelo, G. Nardulli, and N. Paver, Phys. Lett. B **293**, 207 (1992).

[57] For an overview on QCD sum rules consult:
"Vacuum Structure and QCD Sum Rules", edited by M. A. Shifman, North Holland, Amsterdam, 1992.
For a more detailed description see: M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, Nucl. Phys. B **147**, 385 (1979) and Nucl. Phys. B **147**, 448 (1979); L. J. Reinders, H. R. Rubinstein, and S. Yazaki, Phys. Rep. **127**, 1 (1985).

[58] M. Sutherland, B. Holdom, S. Jaimungal, and R. Lewis, Phys. Rev. D **51**, 5053 (1995).

[59] B. Holdom and M. Sutherland, Phys. Lett. B **313**, 447 (1993);
B. Holdom and M. Sutherland, Phys. Rev. D **48**, 5196 (1993).

[60] J. L. Goity and W. Roberts, Phys. Rev. D **51**, 3459 (1995).

[61] T.-M. Yan *et al.*, Phys. Rev. D **46**, 1148 (1992); Erratum *ibid.* Phys. Rev. D **55**, 5851 (1997).

[62] T.-Y. Cheng *et al.*, Phys. Rev. D **48**, 3204 (1993).

[63] C. L. Y. Lee, M. Lu, and M. B. Wise, Phys. Rev. D **46**, 5040 (1992).

[64] Private communication with W. Roberts.

[65] CLEO Collaboration, Y. Kubota *et al.*, Nucl. Instrum. Methods A **320**, 66 (1992).

[66] The term BTF (BTF) refers to the barrel (endcap) time-of-flight. Similarly, BCC (ECC) stands for barrel (endcap) calorimeters.

[67] D. G. Cassel *et al.*, Nucl. Instr. and Meth. A **252**, 325 (1986).

[68] D. Bortoletto *et al.*, Nucl. Instrum. Methods A **320**, 114 (1992).

[69] C. Bebek *et al.*, Nucl. Instrum. Methods A **302**, 261 (1991).

[70] C. Bebek *et al.*, Cornell preprint CLNS 93-1241(1993).

[71] T. Sjöstrand, "PYTHIA 5.6 and JETSET 7.3: Physics and manual", CERN TH 6488-9 (1992).

[72] Applications Software Group, Computing and Networks Division, CERN, "GEANT, Detector Description and Simulation Tool," CERN Program Library Long Writeup Q123.

[73] D. G. Cassel and M. Ogg, CLEO software report CSN 83-333 (1983);
R. Kutschke, CLEO software report CSN 94-334 (1994) and R. Kutschke, CLEO software report CSN 94-335 (1994).

[74] K. Berkelman, CLEO software report CSN 95-342 (1995);
K. Berkelman, CLEO software report CSN 87-261 (1987).

[75] This analysis uses the pre-recompress data. The pre-recompress data were calibrated between winter 1991 and summer 1996. The drift chambers calibration procedure described in Chapter 5 was used for the pre-recompress data. The recompress data contained many new features such as: a KALMAN filter, and improved calibration and MC constants of all the sub-detectors. The recompress data were unfortunately not available for the analysis presented in this thesis.

[76] A. Bellerive, S. L. Jones and M. G. Sciveres, CLEO internal report CBX 94-39 (1994).

[77] A. Bellerive and S. L. Jones, CLEO software report CSN 94-336 (1994).

[78] A. Bellerive, CLEO internal report CBX 95-45 (1995).

[79] Brian P. Geiser, Ph.D. Thesis, Cornell University (1994).

[80] D. Perticone, CLEO software report CSN 89-285 (1989).

[81] R. L. Gluckstern, Nucl. Instr. and Meth. **24**, 381 (1963);
W. R. Innes, Nucl. Instr. and Meth. A **329**, 238 (1993).

[82] U. Binder *et al.*, Nucl. Instr. and Meth. **217**, 285 (1983);
V. Palladino and B. Sadoulet, Nucl. Instr. and Meth. **128**, 232 (1975).

[83] The standard deviation σ is the spatial resolution. Taking the inverse of the intrinsic resolution as the weight would only be correct for a Gaussian distribution. Our spatial residual distribution contains non-Gaussian tails, which means that the spatial resolution may depend on where the tails are cut. The actual resolution is worse than what is shown in Figure 5.16. Since DUET needs the distribution of residuals to be Gaussian in order to correctly compute the error matrices, the residual distribution are fit to a Gaussian by excluding the tails. The sigma (σ) of the fit determines the weight. It means that DUET have to cut on residual distribution tails to come up with a decent error matrix.

[84] Private communication with D. P. Peterson.

[85] Profile histograms are used to display the mean value of Y and its RMS for each bin in X. The mean is $\bar{y} = \sum(Y)/n$ and the error on the mean is RMS $= \sqrt{\{\sum(Y^2)/n - \bar{y}^2\}/n}$, where n is the number of entries.

[86] The Central Detector (CD) numbering scheme numbers all the anodes and after that, all the cathodes. The first PTL layer is CD layer 1, the first VD layer is CD layer 7, the first DR layer is CD layer 17 and the last DR layer is CD layer 67. The first VD cathode is CD layer 68 and the last DR cathode is CD layer 71.

[87] A better algorithm was developed and used in the recompress of all the data sets. See [75] for detail.

[88] I. Kravchenco, CLEO internal report CBX 97-10 (1997).

[89] See Appendix C for detail.

[90] Note that the ALEPH Collaboration did not use the assumption $\mathcal{B}(D_2^{*0} \to D^{*+}\pi^-) = 20\%$ in Reference [29] where they quote upper limits for product branching fractions. When $\mathcal{B}(D_2^{*0} \to D^{*+}\pi^-) = 20\%$ is used, their upper limit becomes $\mathcal{B}(B^- \to D_2^{*0}\ell^-\bar{\nu}_\ell) < 1.7\%$ (95% C.L.).

[91] J.L. Rosner, Comments on Nuclear and Particle Physics **16**, 109 (1993);
A.F. Falk and M. Luke, Phys. Rev. Lett. **292**, 119 (1992);
S. Godfrey and R. Kokoski, Phys. Rev. D **43**, 1679 (1991);
S. Godfrey and N. Isgur, Phys. Rev. D **32**, 189 (1985);
E. Eichten, K. Gottfried, T. Kinoshita, K.D. Lane, and T.M. Yan, Phys. Rev. D **21**, 203 (1980).

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

[93] M. M. Zoeller, Ph.D. thesis, SUNY Albany, 1994.

[94] S. E. Roberts, Ph.D. Thesis, University of Rochester (1997);
S. E. Roberts, M. Battle and E. Thorndike, CLEO internal report CBX 93-113 (1993).

[95] S. Menary, CLEO internal report CBX 93-103 (1993).

[96] C. P. O'Grady and B. Gittelman, CLEO internal report CBX 91-82 (1991).

[97] C. P. O'Grady, Ph.D. Thesis, Cornell University (1994).

[98] Brian Heltsley *et al.*, CLEO internal report CBX 95-35 (1995).

[99] The endcap TOF were not calibrated in the our data sample, only the barrel was calibrated. The data/MC match incorporates a 3% systematic uncertainty. The TOF calibration is another improvement of the recompress data.

[100] W. R. Leo, "Techniques for Nuclear and Particle Physics Experiments", Springler-Verlag (1987).

[101] The reconstructed invariant mass of the particles included in a candidate for the state X is denoted by $M(X)$; while the nominal mass or book value of a candidate for the state X is denoted by m_X.

[102] A. Frodesen, G. Skjeggestad, and H. Tøfte, "Probability and Statistics in Particle Physics", Universitetsforlaget (1993).

[103] E691 Collaboration, J.C. Anjos *et al.*, Phys. Rev. D **48**, 56 (1993).

[104] W. T. Eadie, D. Dryard, F. E. James, M. Ross, and B. Sadoulet, "Statistical Methods in Experimental Physics", North Holland, Amsterdam (1971).

[105] CLEO Collaboration, J. Gronberg *et al.*, CLEO-CONF 96-25, ICHEP96 pa05-069 (1996).

[106] G. Wei, T. Skwarnicki, R. Stroynowski, CLEO internal report CBX 94-72 (1994).

[107] D. Besson, CLEO internal reports CBX 93-01 (1993) and CBX 93-96 (1993).

[108] D. Besson, CLEO internal reports CBX 95-46 (1995), CBX 96-01 (1996), and CBX 96-75 (1996);
J. R. Patterson and J. G. Smith, CLEO internal report CBX 95-39 (1995);
K. Bloom, R. Patterson and P. Drell, CLEO internal report CBX 96-101 (1996).

[109] M. G. Sciveres, Ph.D. thesis, Cornell University (1994).

[110] E687 Collaboration, P. L.Frabetti *et al.*, Phys. Lett. B **331**, 217 (1994).

[111] Mark III Collaboration, J. Adler etal, Phys. Lett. B **196**, 107 (1987).

[112] CLEO Collaboration, P. Avery *et al.*, CONF 96-28, ICHEP96 pa05-075 (1996).

[113] Private communication with K. Bloom.

[114] CLEO Collaboration, G. Crawford *et al.*, Nucl. Instrum. Methods A **345**, 429 (1994).

[115] Private communication with D. Besson.

[116] J. G. Körner, G. A. Schuler, Z. Phys. C **38**, 511 (1988).

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

[118] Private communication with R. Patterson.

[119] DELPHI Collaboration, DELPHI preprint 97-102, CONF 84, EPS-0452 (1997).

[120] Z. Ligeti, "Semileptonic B Decays to Excited Charm Mesons", in the Proceedings of the 7^{th} International Symposium on Heavy Flavor Physics, Santa Barbara, 1997 (unpublished).

[121] F. Muheim, CLEO internal report CBX 92-104 (1992).

[122] M. Chadha, CLEO internal report CBX 93-27 (1993).