MEASUREMENT OF TWO BODY B MESON DECAYS TO PIONS AND KAONS WITH THE CLEO III DETECTOR

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Alan Jay Magerkurth

January 2003

MEASUREMENT OF TWO BODY B MESON DECAYS TO PIONS AND KAONS WITH THE CLEO III DETECTOR

Alan Jay Magerkurth, Ph.D.

Cornell University 2003

We have measured the branching ratios of the decays of B mesons to pseudoscalar kaons and pions using data from the CLEO III detector. We combine our results with the CLEO II measurement and find the branching ratios for the decays $BR(B^0 \to K^+\pi^-) = 18.0^{+2.3+1.2}_{-2.1-0.9} \times 10^{-6}$, $BR(B^+ \to K^+\pi^0) = 12.9^{+2.4+1.2}_{-2.2-1.1} \times 10^{-6}$, $BR(B^+ \to K^0\pi^+) = 18.8^{+3.7+2.1}_{-3.3-1.8} \times 10^{-6}$, $BR(B^0 \to K^0\pi^0) = 12.8^{+4.0+1.7}_{-3.3-1.4} \times 10^{-6}$, and $BR(B^0 \to \pi^+\pi^-) = 4.5^{+1.4+0.5}_{-1.2-0.4} \times 10^{-6}$, $BR(B^0 \to \pi^+\pi^0) = 4.6^{+1.8+0.6}_{-1.6-0.7} \times 10^{-6}$, where the first error is statistical and the second is systematic. We set limits on the branching ratios $BR(B^0 \to \pi^0\pi^0) < 4.4 \times 10^{-6}$, $BR(B^0 \to K^+K^-) < 0.8 \times 10^{-6}$, $BR(B^+ \to K^+\bar{K}^0) < 3.3 \times 10^{-6}$, and $BR(B^0 \to K^0\bar{K}^0) < 3.3 \times 10^{-6}$ at the 90% confidence level.

We also discuss the possibility of constraining the (ρ, η) plane from the measurement of these branching ratios.

BIOGRAPHICAL SKETCH

Alan Jay Magerkurth was born December 9, 1976 in Geneseo, Illinois. He attended school in Geneseo through his first year of high school, and then had the good fortune of attending the Illinois Mathematics and Science Academy in Aurora, Illinois for the remaining three years of high school. After many classes taught by excellent teachers, physics emerged as his favorite discipline among the math and sciences. He graduated from IMSA in June 1994.

After high school, he attended the University of Illinois in Urbana-Champaign to pursue a degree in physics. He spent very enjoyable summers working first at the $g-2$ experiment at Brookhaven National Laboratory, and then studying amorphous diamond films at Sandia National Laboratory. He graduated *magna cum laude* from the University of Illinois with a B.S. in physics in December 1997. He returned to Sandia briefly, before beginning his Ph.D. studies at Cornell University in August 1998. There he did his research on the CLEO experiment, developing software for the CLEO III detector and studying rare B meson decays, and receiving a M.S. in physics May 2002.

To Christy, who's waited a long time for this.

ACKNOWLEDGMENTS

There are many people who have helped me get to where I am today, and it is a great pleasure to have the chance to recognize all of their contributions.

First, I want to thank my parents. They have always encouraged me to do my best, fostered my curiosity, and supported me and guided me along the way. I hope they know how much I appreciate all that they have done for me.

I would like to thank Yannis Semertzidis, my advisor at Brookhaven while I worked on the $g-2$ experiment. My first experience with real world science was a thoroughly enjoyable one, even when we were working with the magnet until the break of dawn. Many thanks to my supervisors at Sandia, Tom Friedman and John Sullivan. They are both talented scientists and great people to work with. Special thanks to John, who was an excellent companion on all those hikes up mountains, down canyons, into caves and across lava fields. I also want to thank Scott Willenbrock for the time he spent working with me on our Υ spectroscopy project my final semester at Illinois.

CLEO would not be the successful experiment it is today without the contributions of the dedicated and incredibly talented members of its collaboration. Numerous people were instrumental in helping me learn what we do and how we do it here at CLEO. In particular, thanks to Mike Marsh for his patience while I learned the CLEO software system, Peter Gaidarev for sharing his expertise on the CLEO III silicon and CLEO in general, Chris Jones for teaching me almost everything I know about C++, and Frederic Blanc and Jean Duboscq for their physics insight and good advice.

I also want to recognize the contributions of my fellow Rare B group members, both past and present: Chris Anderson, Kamal Benslama, Eric Eckhart, Ahren Sadoff, Matt Shepherd, David Urner and Michael Weinberger.

My stay in Ithaca has been much more enjoyable because of the time that I have spent with Chris Stepaniak, Gregg Thayer and Pete Zweber, whether it was softball, poker, football or just shooting the breeze in someone's office.

Jim Alexander has been a superb advisor over the past few years. He has always had good ideas, a grasp of the big picture and an enthusiasm for the work that we do here. His help in performing the analysis and preparing this thesis has been invaluable. He always found time in his busy schedule to chat, and it has always been a pleasure to work with him.

Thanks to the members of my special committee over the past few years, Karl Berkelman, Lawrence Gibbons, David Lee and Henry Tye. Special thanks to Lawrence for suggesting I go to the summer school in Scotland. That was the highlight of my graduate career and I will always remember the many wonderful experiences I had there.

Finally, I want to thank Christy, who has been by my side through it all, even when she was 200 miles away. I have never known anyone as wonderful and caring, and her love and support have been invaluable.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Particle physics experiments are aimed at answering some of the most fundamental questions about the physical world. What is the source of mass? What is the origin of the universe? Are the forces we see really all different manifestations of one unified force? These are the questions that motivate physicists to study the wide variety of phenomenon exhibited by the most fundamental particles that we know of. The analysis presented in this thesis is one piece in the puzzle to answer the question, why do we live in a matter-dominated world?

1.1 CP Violation and the The Standard Model

One of the most striking observations about our universe is perhaps one of the most basic, that we live in a world comprised of "ordinary" matter. More precisely, our world is composed of baryons, but as far as we know, essentially no anti-baryons [1]. Given our current understanding of the universe, we would expect baryons and anti-baryons to have been produced in equal numbers, and also to be annihilated in equal numbers. As the universe cools, the annihilations should cause all baryons and anti-baryons to disappear, leaving no matter behind. The question then becomes whether or not our current theory is consistent with the existence of a world composed of matter.

There are three ingredients necessary to generate a baryon asymmetry [2]:

- existence of processes that do not conserve baryon number,
- periods of thermodynamic disequilibrium with respect to the baryon number violating processes when these processes can occur,
- existence of charge conjugation-parity (CP) symmetry violating processes.

For our discussion, we will assume the first two items as given, and examine the third point. In order to do this, we need to know the origins of CP-violating processes in the current theory, and also, we need to know if the magnitude of these processes is sufficient to account for the observed baryon asymmetry.

1.1.1 Discrete Symmetries

There are two main types of transformation common in physics. Continuous transformations are those that can be performed as a number of small steps. Translation is one example of a continuous transformation, e.g. we can translate our coordinate system by one meter by making a series of very small translations. Rotation and time translation are other examples of continuous transformations.

Discrete transformations, on the other hand, are those that cannot be accomplished by a series of smaller steps. A CP transformation is the composition of two of the three basic discrete transformations of particle physics: parity (P), charge conjugation (C) and time reversal (T). The parity operation changes the sign of all coordinates, charge conjugation replaces particles with anti-particles, and time reversal plays an interaction backwards in time. CP is the application of both parity and charge conjugation, and is thus the inversion of the coordinate system followed by the exchange of particles and anti-particles. At first, we might expect that all microscopic processes would be invariant (symmetric) under any of these transformations, however experiment has shown that the weak interaction is neither parity nor charge conjugation invariant.

The textbook example of parity and charge conjugation violation in the weak interactions is the helicity of neutrinos. The helicity of a particle is determined by whether the projection of its spin along its momentum is parallel or anti-parallel to its momentum. For massive particles, we can always transform to a frame that switches the direction of the momentum vector without changing the spin, and thus change the helicity of the particle. For a massless neutrino, this is not possible, and we can unambiguously determine its helicity, and experiments find that there are only left-handed neutrinos and right-handed anti-neutrinos (a right-handed particle has its spin projection in the same direction as its momentum, while a left-handed particle has its spin projection opposite its momentum). If we examine the effect of the parity and charge conjugation operations on neutrinos, we see that each operation transforms physical states into unphysical states. However, if we perform the combined C$\times$P operation, we find that physical states are transformed into physical states (see Figure 1.1). Therefore the weak interaction, which is the only force that couples to neutrinos, violates both symmetries individually, but we might expect it to obey the stronger requirement of CP symmetry.

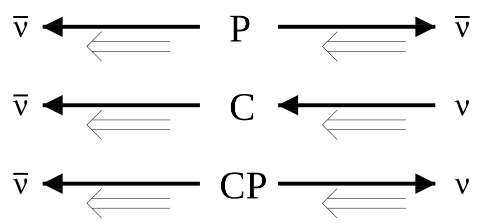

Figure 1.1: Effects of discrete symmetry operations on neutrinos. The heavy arrow indicates momentum direction, while the double arrow indicates spin.

1.1.2 Standard Model Overview

Before we can discuss CP violation in any detail, it is necessary to know something about the theory that has emerged to explain most particle physics phenomena, the Standard Model (SM).

The SM can be minimally described by four features, namely its:

- gauge symmetries,
- fermions and scalars,
- Lagrangian, and
- symmetry breaking.

The gauge symmetry of the SM is $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ (the notation for this section follows that of [3]), where the subscripts denote the color symmetry of the strong interactions, weak isospin, and hypercharge respectively. The SU(2)×U(1) symmetry is spontaneously broken via the Higgs mechanism so that $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ goes to $\mathrm{U}(1)_{EW}$.

There are three families of fermions, each consisting of

$$Q_L(3,2,\tfrac{1}{6}), \quad U_R(3,1,\tfrac{2}{3}), \quad D_R(3,1,-\tfrac{1}{3}), \quad L_L(1,2,-\tfrac{1}{2}), \quad E_R(1,1,-1) \tag{1.1}$$

The first number indicates whether or not the fermion is an SU(3) triplet or singlet, the second indicates whether or not the fermion is an SU(2) doublet or singlet, and the last indicates the hypercharge of the fermion. The L and R subscripts refer to the helicity of the particle.

The Lagrangian can be considered to be the sum of three pieces, a kinetic term, a Higgs term, and a Yukawa term. The kinetic terms are all of the form

$$\mathcal{L} = i\bar{F}\gamma_\mu D^\mu F, \tag{1.2}$$

where F refers to the fermion fields, and D^μ is the covariant derivative,

$$D^\mu = \partial^\mu + ig_s G^\mu_a L_a + igW^\mu_b T_b + ig' B^\mu Y. \tag{1.3}$$

G^μ_a, W^μ_b, and B^μ are the gluon, weak, and hypercharge fields, respectively, and L_a, T_b, and Y are the corresponding generators. The Higgs term in the Lagrangian is the scalar self-interaction term

$$\mathcal{L} = \mu^2\phi^\dagger\phi - \lambda^2(\phi^\dagger\phi)^2. \tag{1.4}$$

Finally, the Yukawa term describes the interaction of fermions and the scalar, and are of the form

$$\mathcal{L} = -Y^d_{ij}\bar{Q}^I_{Li}\phi D^I_{Rj} - Y^u_{ij}\bar{Q}^I_{Li}\tilde{\phi}U^I_{Rj} + \text{h.c.} \tag{1.5}$$

where

$$\tilde{\phi} = i\tau^2\phi^*, \tag{1.6}$$

for instance.

This is of course a vast oversimplification, and it tends to minimize the tremendous amount of work to develop the theory. It also masks the rich phenomenology seen in a wide variety of physical processes. Calculating physical quantities from these first principles can often be extraordinarily difficult. This framework will, however, provide a convenient starting point for our discussion.

Of particular interest to this analysis is the role of the quark sector in the kinetic and Yukawa Lagrangian terms. The fermions in the previous paragraph refer to interaction eigenstates, while the quarks in the Yukawa term of the Lagrangian are most naturally thought of as mass eigenstates. The relationship between interaction and mass eigenstates is

$$q^M_{iH} = (V_{qH})_{ij} q^I_{jH} \tag{1.7}$$

where $H = L, R$ and $q = u, d$. The term in the Lagrangian that couples quarks to $W^{\pm}$, when written in the mass basis, becomes

$$\mathcal{L} = -\frac{g}{\sqrt{2}} u_{Li}\gamma^{\mu}(V_{uL}V^{\dagger}_{dL})_{ij} d_{Lj} W^{+}_{\mu} + \text{h.c.}. \tag{1.8}$$

The combination $V_{uL}V^{\dagger}_{dL} = V_{\text{CKM}}$ is the only source of CP violation in the SM[1]. As a matter of fact, all CP violating processes in the SM are related to one

[1]Assuming $\theta_{QCD} \ll 1$, as implied by neutron electric dipole measurements, and as long as massive neutrinos, extended Higgs sectors, etc. are considered extensions to the SM.

parameter of V_{CKM}, and as such, are all very strongly related. Indeed, independent measurements of this parameter from various sources are quite consistent (see Figure 1.3). One of the problems of the SM is that, despite its success in describing CP violation in meson decays, it falls dramatically short in predicting another CP observable, the baryon number of our universe [4]. The baryon asymmetry as predicted by the SM is many orders of magnitude lower than what we observe. Thus there is great interest in finding new sources of CP violation or inconsistencies in the SM picture of CP violation that might point to new physics.

1.1.3 The CKM Matrix

One of the favorite hunting grounds for inconsistencies in the SM is the Cabibbo-Kobayashi-Maskawa (CKM) matrix [5], [6]. Shown in Eq. 1.9, it essentially tells us the probability of an up-type quark to couple to a down-type quark in an interaction involving a W boson. In principle, it is described by nine parameters: three angles and six phases. However, it is possible to to transform V with two diagonal phase matrices and eliminate five of the six phases without affecting the physics of the SM. V is then described by four parameters: three angles and one phase that cannot be transformed away. It is this irreducible phase that leads to CP violation in the SM.

$$V \equiv \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{1.9}$$

One very popular parameterization of the CKM matrix is that of Wolfenstein

[7]. This parameterization is an expansion in the parameter $\lambda = |V_{us}|$, where $\lambda = 0.22$. Corrections of order λ^4 have been neglected.

$$V \simeq \begin{pmatrix} 1-\frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} \tag{1.10}$$

Unitarity of V requires (among other things)

$$\sum_{i=u,c,s} V_{ij}V^*_{ik} = 0 \tag{1.11}$$

where j and k are two different down-type quarks. These relations can be pictured as closed triangles in the complex plane. When $j = d$, $k = s$ or $j = s$, $k = b$, these triangles have two sides that are much longer than the third side. This is evident by looking at Eq. 1.10, where we see that for $j = d$, $k = s$, one side is down by a factor λ^4, and for $j = s$, $k = b$, one side is down by a factor λ^2, whereas for $j = d$, $k = b$, all sides are $\sim \lambda^3$. It is this third triangle that is the familiar Unitarity Triangle (Figure 1.2). It is typical to choose the phases of the CKM matrix such that $V_{cd}V^*_{cb}$ is real, and to divide all three sides by $|V_{cd}V^*_{cb}|$ so that the triangle has vertices at (0,0), (1,0), and (ρ,η).

One strategy for looking for physics beyond the SM is to measure the sides and angles of the Unitarity Triangle so that constraints are put on ρ and η. If processes are found that are inconsistent with the CKM framework, it could give important clues as to the nature of physics beyond the SM.

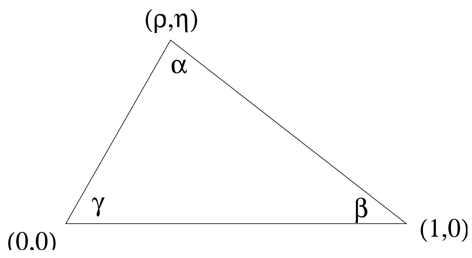

Figure 1.2: The Unitarity Triangle.

1.1.4 Status of CKM Parameters

At this point, it is important to discuss what constraints currently exist on the (ρ, η) plane. We want to make sure that we can measure (ρ, η) through rare B decays well enough to find discrepancies with other measurements, otherwise there would be much less motivation to measure them.

Four constraints on the (ρ, η) plane are shown in Figure 1.3. These include measurements of semi-leptonic B decays ($|V_{ub}/V_{cb}|$), CP violation in kaon decay ($|\epsilon_K|$), B mixing (Δm_d and $\Delta m_s/\Delta m_d$), and CP violation in $b \to c\bar{c}s$ systems ($\sin 2\beta$). The favored region is designated by the ellipsoidal regions near (0.25, 0.35). While all of these measurements will certainly improve over time, it would be good to have another source of information that could tell us about the angle γ. As we will see, the decays of B mesons to two charmless pseudoscalar mesons have the potential to do just that.

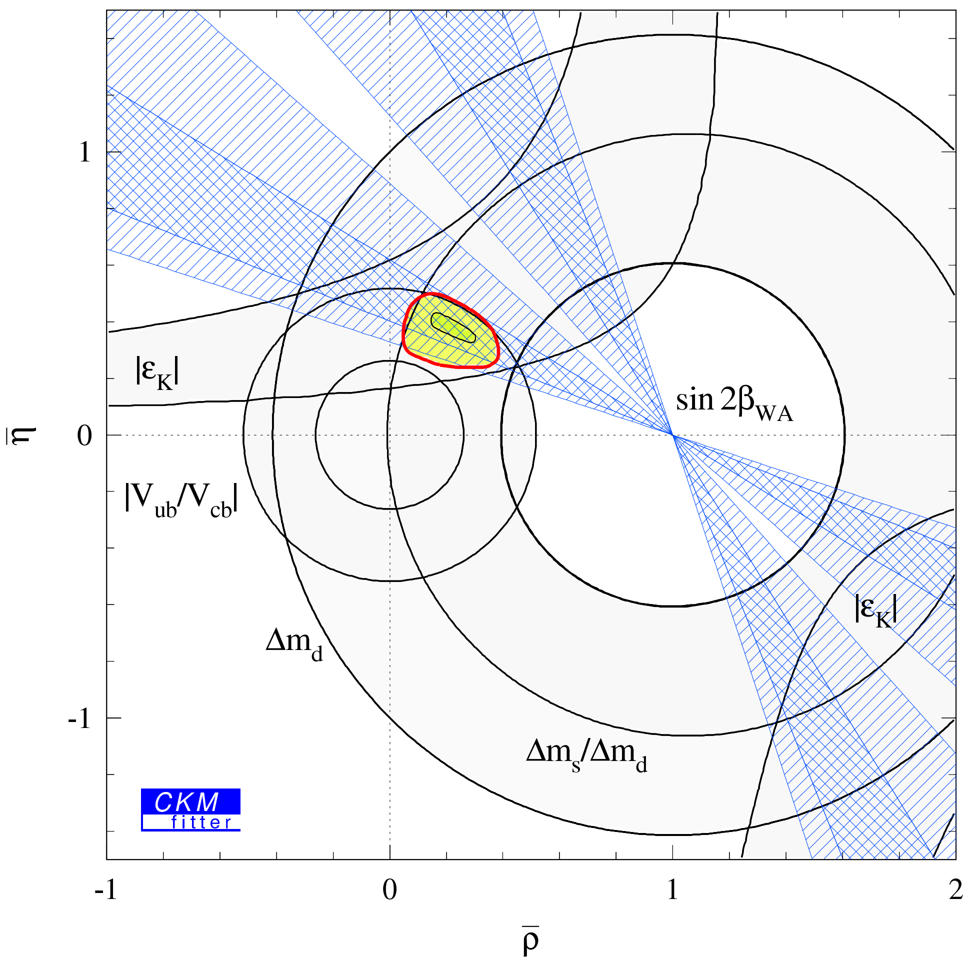

Figure 1.3: Constraints on the (ρ, η) plane (taken from [8]).

1.2 Theory of Rare B Decays

The decays of B mesons ($b\bar{u}, b\bar{d}$ in particular, and to a lesser extent $b\bar{s}$, and their charge conjugates, for this discussion) to two light mesons are of great theoretical and experimental interest. Theoretical advances in approaches such as QCD factorization, along with improved experimental techniques and high luminosity colliders, have made these rare ($\sim 10^{-5} - 10^{-7}$) B decays candidates for measuring angles of the Unitarity Triangle.

Feynman Diagrams

The first step in understanding rare B decays is to discuss the various types of decays that contribute to $B \to K\pi, \pi\pi$. For completeness, there are six types of diagrams that can contribute, but three of these will dominate. Feynman diagrams are shown in Figure 1.4.

Diagram (A) is a tree diagram, where the b quark emits a W boson that subsequently decays to a $q\bar{q}'$ system. Diagram (B) is similar, but here the amplitude is color-suppressed since the quark from the W pairs with the spectator anti-quark. Diagram (C) is the penguin diagram, where an up-type quark mediates the W loop and emits a gluon. There also exist similar diagrams where instead of a gluon, a photon or Z boson is emitted. Diagrams involving a photon or Z boson are commonly grouped together and referred to as electroweak penguins. These three diagram (A-C) are the most important for our discussion.

In diagram (D), a W boson is exchanged between the two quarks of the B

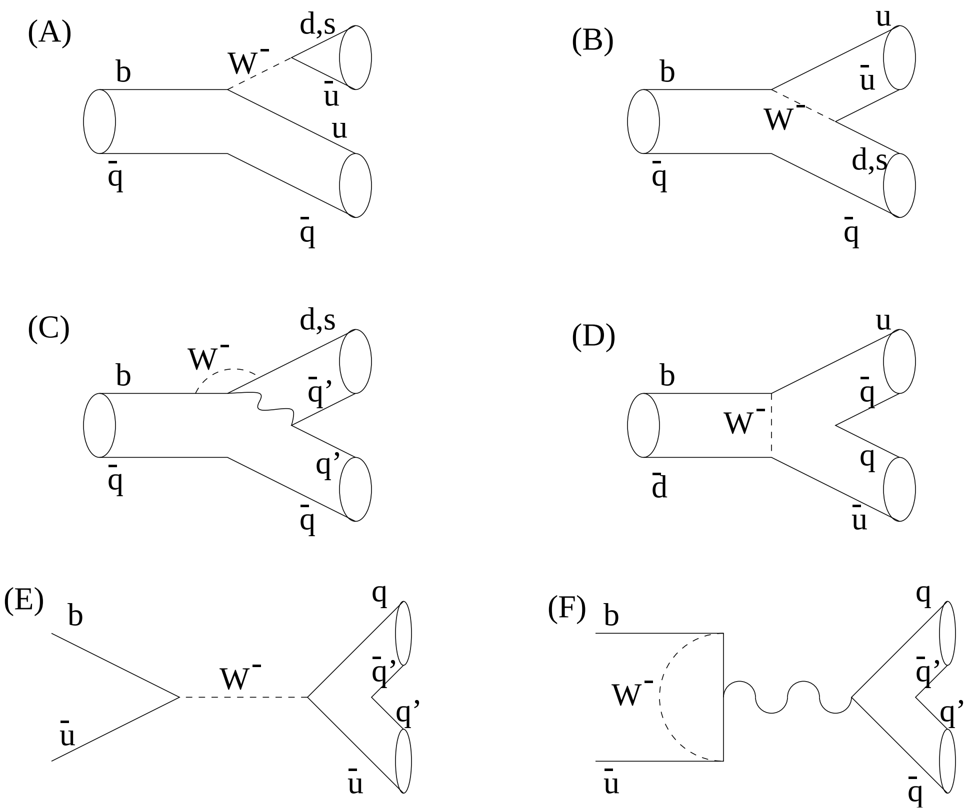

Figure 1.4: Feynman diagrams for the various $B \rightarrow K\pi, \pi\pi$ decays.

meson, and a quark-antiquark pair is created. This type of decay is most important for decays like $B^0 \to K^+K^-$, which has no contributions from tree or penguin decays. Finally, diagrams (E) and (F) are annihilation and penguin annihilation diagrams. All three of these diagrams are suppressed relative to the tree and penguin diagrams by a factor of f_B/m_B ($\mathcal{O}(5\%)$).

Table 1.1 shows which amplitudes contribute to the various $B \to K\pi, \pi\pi$ decays. All of these decays can proceed through (sometimes color-suppressed) electroweak penguins. These diagrams can be arranged [9] in a hierarchy that roughly indicates their relative size, which serves as a useful guide when deciding what amplitudes are important for a particular decay (Table 1.2). The use of the parameter λ is not strictly related to the CKM parameter, but is used only for bookkeeping.

1.2.1 QCD Factorization

Calculating the amplitudes of these diagrams is notoriously difficult. Since both decay products are hadronic, gluons can be exchanged between the parent B meson and either daughter and between the daughters themselves. Faced with these difficulties, theorists have tried numerous techniques to make sense of these decays. Here we will briefly discuss one of these techniques, known as factorization.

Factorization refers to the separation of short distance and long distance scales. Short distance effects can be calculated in perturbation theory, while long distance effects may be measurable (form factors and decay constants, for example), or may have to be estimated from the theory.

The starting point for two-body hadronic B decays is the weak effective Hamil-

Table 1.1: Diagrams that contribute to rare B decays.

Mode	Contributions
$\pi^+\pi^-$	T,P,E,PA
$\pi^+\pi^0$	T,C
$\pi^0\pi^0$	C,P,E,PA
$K^+\pi^-$	T,P
$K^+\pi^0$	T,C,P,A
$K^0\pi^-$	P,A
$K^0\pi^0$	C,P
K^+K^-	E,PA
$K^+\overline{K^0}$	P,A
$K^0\overline{K^0}$	P,PA

tonian,

$$H = \frac{G_F}{\sqrt{2}} \sum \lambda_i C_i(\mu) O_i(\mu) \tag{1.12}$$

where G_F is the Fermi constant, λ_i are the appropriate CKM factors, and C_i are Wilson coefficients. The decay amplitudes are then

$$A(B \to P_1 P_2) = \frac{G_F}{\sqrt{2}} \sum \lambda_i C_i(\mu) \langle P_1 P_2 | O_i | B \rangle (\mu) \tag{1.13}$$

The decay amplitudes of the different B decays can also be parameterized in a form that emphasizes the contribution of the different Feynman diagrams. For

Table 1.2: Approximate hierarchy of diagrams for $b \to d$ and $b \to s$ transitions.

Order	$b \to d$	$b \to s$
1	T	P
λ	C,P	T,P$_{EW}$
λ^2	E,A,P$_{EW}$	C,PA,P$^C_{EW}$
λ^3	PA,P$^C_{EW}$	E,A

instance, following the notation of [10],

$$A(B^- \to K^0\pi^-) = P(1 + \epsilon_a e^{i\phi_a} e^{-i\gamma}) \tag{1.14}$$

$$-\sqrt{2}A(B^- \to K^-\pi^0) = P(1 + \epsilon_a e^{i\phi_a} e^{-i\gamma} - \epsilon_{3/2} e^{i\phi}(e^{-i\gamma} - q e^{i\omega})) \tag{1.15}$$

$$-A(\bar{B}^0 \to K^-\pi^+) = P(1 + \epsilon_a e^{i\phi_a} e^{-i\gamma} - \epsilon_T e^{i\phi_T}(e^{-i\gamma} - q_C e^{i\omega_C})). \tag{1.16}$$

Here P is the magnitude of the penguin contribution (where penguin means all diagrams not proportional to $e^{-i\gamma}$ in the $K^0\pi^-$ amplitude), the ϕ's and ω are strong phases, $\epsilon_{3/2}$ is the ratio of tree to penguin diagrams, ϵ_a is the ratio of rescattering to penguin contributions, and q is the ratio of electroweak penguin to penguin diagrams. Both ϵ_a and ω are small ($\mathcal{O}(1\%)$) parameters. The amplitudes for the charge conjugate processes are the same, except the weak phase γ changes sign everywhere.

Each of the terms in the parameterizations of eqs. 1.14-1.16 can be solved in terms of Wilson coefficients and parameters such as form factors, decay constants and CKM quantities. These parameterizations can then be used to calculate ratios of branching fractions and CP asymmetries. With fortuitous values of experimental quantities, bounds can be placed in a model-independent way on CKM parameters (see, for example, eqs. 1.26 and 1.32), or by introducing some model-dependence, CKM parameters can be directly obtained regardless of the particular values of the experimental results.

A comment about nomenclature is in order at this point. There is some ambiguity about the term factorization as it is used in the literature. Factorization as described above could be called QCD factorization, where the factorization refers to the separation of distance scales. Here, hard gluon exchange is a perturbative effect and can be incorporated into the theory and is therefore “factorizable”. Another approach to factorization is to consider the four quark operator for a decay to be the product of two bilinear operators. This could be more precisely termed naive factorization. Here factorization refers to the separation of the matrix elements into two terms. In this sense, any exchange of gluons, for instance, regardless of its momentum is a non-factorizable contribution. Further confusing the issue is that when using the heavy quark limit ($m_b \gg \Lambda_{QCD}$), one can obtain the expression[11]

$$\langle P_1 P_2 | O_i | B \rangle = \langle P_1 | (q_1 q_2) | 0 \rangle \langle P_2 | (q_3 q_4) | B \rangle [1 + \sum r_n \alpha_s^n + \mathcal{O}(\Lambda_{QCD}/m_b)]. \quad (1.17)$$

When corrections of order α_s and Λ_{QCD}/m_b are neglected, the naive factorization formula is recovered from a QCD factorization starting point.

We now focus on the methods of constraining CKM parameters using the measurements of branching ratios and CP asymmetries of the rare B decays.

Fleischer and Mannel Method

Starting with methods that are the most feasible experimentally, we look at proposals that use CP-averaged branching ratios to constrain the CKM parameters. One of the first of these methods was proposed by Fleischer and Mannel [12]. They look at the ratio of the CP-averaged branching fractions of $B^0 \to K^+\pi^-$ and $B^+ \to K^0\pi^+$. Both amplitudes have contributions from gluonic and electroweak penguin diagrams, and the B^0 mode has an additional tree level component. The annihilation diagram that contributes to the B^+ decay is considered negligible and ignored. The amplitudes for these two decays can be written

$$A(B^0 \to K^+\pi^-) = -[(P_{QCD} + c_u P_{EW}) + T] \tag{1.18}$$

$$A(B^+ \to K^0\pi^+) = \mathcal{P}_{QCD} + c_u \mathcal{P}_{EW}, \tag{1.19}$$

where $c_u = +2/3$ is the up-quark charge. They then write the amplitudes for the tree and penguin pieces as

$$T = e^{i\gamma} e^{i\delta_T} |T| \tag{1.20}$$

and

$$P = e^{i\pi} e^{i\delta_P} |P| \tag{1.21}$$

where P is the sum of the gluonic and electroweak penguin amplitudes, γ and π are weak phases, and δ_i are strong phases. Then the amplitude for the decay $B^0 \to K^+\pi^-$ can be written

$$A(B^0 \to K^+\pi^-) = e^{i\delta_P} |P| (1 - e^{i\gamma} e^{i\delta} r) \tag{1.22}$$

where

$$r \equiv \frac{|T|}{|P|},\ \delta \equiv \delta_T - \delta_P. \tag{1.23}$$

This leads immediately to an expression for the CP-averaged decay amplitude for the B^0:

$$|A(B \to K^{\pm}\pi^{\mp})|^2 = |P|^2(1 - 2r\cos\gamma\cos\delta + r^2). \tag{1.24}$$

They then assume that the electroweak penguins are negligible (supported by estimates that they make that this is a $\mathcal{O}(1\%)$ correction) and use isospin symmetry to justify the assertion that the gluonic penguin contribution for these two decays are identical. Thus, the magnitude of the penguin diagram drops out in the ratio of the branching fractions, giving

$$R = 1 - 2r\cos\gamma\cos\delta + r^2 = \frac{B(B \to \pi^{\pm}K^{\mp})}{B(B^{\pm} \to \pi^{\pm}K^0)}. \tag{1.25}$$

Rewriting Eq. 1.25 as

$$\cos\gamma\cos\delta = \frac{1-R}{2r} + \frac{r}{2} \tag{1.26}$$

we see that $\cos\gamma$ is constrained if $R < 1$, since the right hand side of eq. 1.26 is then positive and $|\cos\delta| < 1$ (also see the curves for $R < 1$ in Figure 1.5). It should also be noted that the ratio of branching fractions they use ignores phase space factors. The main correction is then due to the lifetimes of the charged and neutral B's. This implies that R in Eq. 1.25 should be replaced by

$$R' = R\frac{\tau_{B^+}}{\tau_{B^0}}. \tag{1.27}$$

Since this ratio of B lifetimes is 1.08 ± 0.02, this weakens the constraint slightly.

More CP-Averaged Constraints

The Fleischer and Mannel bound was one of the first methods to use CP-averaged branching ratios to explore CKM parameters. Much work has been done since then to derive more sophisticated bounds of this type, motivated by the fact that these types of measurements are the easiest experimentally.

Perhaps the simplest thing that can be done is to estimate the value of r in Eq. 1.25. Calculations have been performed [10] that put r at approximately 20%. Depending on the particular value of R found from experiment, one can see from Figure 1.5 that fixing r can improve the bound on γ.

A more sophisticated analysis involves the ratio of the CP-averaged branching fractions of $B^+ \to K^0\pi^+$ and $B^+ \to K^+\pi^0$ [10]. The amplitudes for these two decays were parameterized in eqs. 1.14 and 1.15.

With these amplitudes, we can calculate the ratio of branching fractions

$$R_* = \frac{B(B^+ \to K^0\pi^+) + B(B^- \to \bar{K}^0\pi^-)}{2[B(B^+ \to K^+\pi^0) + B(B^- \to K^-\pi^0)]}. \tag{1.28}$$

Using the amplitudes given by Eqs. 1.14 and 1.15, we can calculate R_*^{-1}, up to first order in the small parameters ω and ϵ_a.

$$\begin{aligned} R_*^{-1} &= 1 + 2\bar{\epsilon}_{3/2}\cos\phi(q - \cos\gamma) + \bar{\epsilon}_{3/2}^2(1 - 2q\cos\gamma + q^2) \\ &\quad -2\bar{\epsilon}_{3/2}\epsilon_a[\sin^2\gamma\cos\phi\cos\phi_a + (1 - q\cos\gamma)\sin\phi\sin\phi_a] \\ &\quad -2\bar{\epsilon}_{3/2}q\sin\phi\sin\omega \end{aligned} \tag{1.29}$$

where

$$\bar{\epsilon}_{3/2} = \frac{\epsilon_{3/2}}{\sqrt{1 + 2\epsilon_a\cos\phi_a\cos\gamma + \epsilon_a^2}} = R_{th}\epsilon_{exp} \tag{1.30}$$

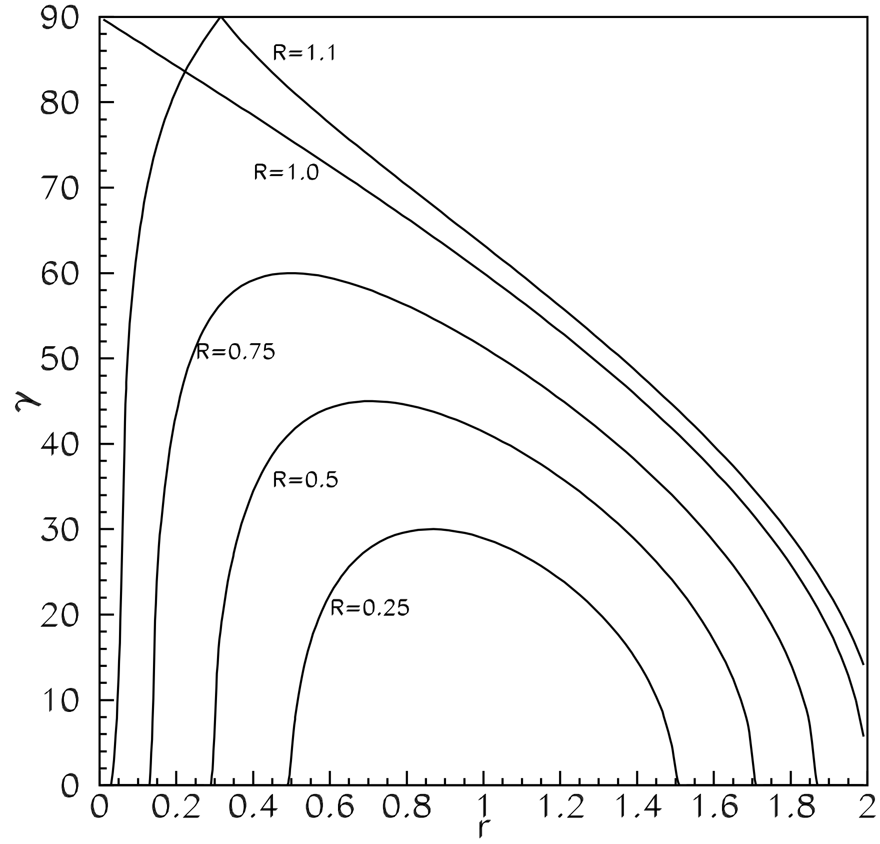

Figure 1.5: Maximum value of γ as a function of r for various values of R.

The advantage of trading $\epsilon_{3/2}$ for $\bar{\epsilon}_{3/2}$ is that R_{th}, which is essentially a ratio of Wilson coefficients, can be determined to the 5% level, and ϵ_{exp} (Eq. 1.31) can be measured experimentally.

$$\epsilon_{exp} = \tan\theta_C \frac{f_K}{f_\pi}\left[\frac{2(B(B^+\to\pi^+\pi^0)+B(B^-\to\pi^-\pi^0))}{B(B^+\to K^0\pi^+)+B(B^-\to\bar{K}^0\pi^-)}\right]^{1/2}. \tag{1.31}$$

Allowing the strong phases to take on any value, Eq. 1.29 can be converted into the bound

$$R_*^{-1} < (1+\bar{\epsilon}_{3/2}|q-\cos\gamma|)^2 + \bar{\epsilon}_{3/2}(\bar{\epsilon}_{3/2}+2\epsilon_a)\sin^2\gamma \tag{1.32}$$

With R_*^{-1} large enough, this leads to a constraint that depends strongly on γ (q itself is equal to $\hat{q}/R_u$, with $R_u = \frac{1-\lambda^2}{\lambda}\left|\frac{V_{ub}}{V_{cb}}\right|$, and thus depends on ρ and η).

To go from a bound to a measurement of γ, some model dependence is necessary. Eq. 1.29 can be solved for $\cos\gamma$,

$$\cos\gamma = q + \frac{1-R_*^{-1}+\bar{\epsilon}_{3/2}^2(1-q^2)-2\bar{\epsilon}_{3/2}\delta}{2\bar{\epsilon}_{3/2}(\cos\phi+\bar{\epsilon}_{3/2}q)} \tag{1.33}$$

with

$$\delta = \epsilon_a[\sin^2\gamma\cos\phi\cos\phi_a + (1-q\cos\gamma)\sin\phi\sin\phi_a] + q\sin\phi\sin\omega \tag{1.34}$$

and then $\hat{q}$, $\bar{\epsilon}_{3/2}$, ϕ and ϕ_a can be scanned over a plausible range of values to give an allowed region in the (ρ,η) plane. Where Eq. 1.32 is only useful if $R_*^{-1} > 1$, any value of R_*^{-1} can provide (model-dependent) information about the CKM parameters by use of Eq. 1.33.

Time-Independent CP Asymmetries

In general, the decay of a particle and its charge conjugate can exhibit direct CP-violation if there is interference between two decay channels that have non-zero relative strong and weak phases. The strong phase is the same for both the particle and its anti-particle, while the weak phase changes sign. When calculating the branching ratio from the squared matrix elements, the interference between the two diagrams will be different, leading to a difference between the decay rates for particle and anti-particle. Several strategies exist for determining γ using measurements of CP asymmetries in rare B decays. We will review two of them here.

One method [13] uses SU(3) symmetry to represent three decay amplitudes as

$$A(B^+ \to \pi^+\pi^0) = -\frac{1}{\sqrt{2}}(T + C) \tag{1.35}$$

$$A(B^+ \to K^0\pi^+) = (P' + A') \tag{1.36}$$

$$A(B^+ \to K^+\pi^0) = -\frac{1}{\sqrt{2}}(T' + C' + P' + A') \tag{1.37}$$

where T, C, P, and A represent tree, color-suppressed, penguin, and annihilation topologies. Primed quantities are $|\Delta S| = 1$ processes. Since primed and unprimed quantities are the same in SU(3) up to CKM factors, their ratios are all equal.

$$T'/T = C'/C = A'/A = V_{us}/V_{ud} \tag{1.38}$$

From Eqs. 1.35-1.37, we can construct a triangle relation in the complex plane.

$$\sqrt{2}A(B^+ \to K^+\pi^0) + A(B^+ \to K^0\pi^+) = \frac{V_{us}}{V_{ud}}\sqrt{2}A(B^+ \to \pi^+\pi^0) \tag{1.39}$$

In the limit that the annihilation diagrams can be ignored, the weak phases can be determined for two of the decays. Dominance of the top quark in the penguin loop implies that the weak phase of P' will be $V_{tb}^*V_{ts}$ from the two W vertices, which gives a weak phase of π. Note that $A(B^+ \to K^0\pi^+) = A(B^- \to \bar{K}^0\pi^-)$.

For $\pi^+\pi^0$, the weak phase again comes from two W vertices. This time, the two factors are $V_{ub}^*V_{ud}$, and when multiplied by the ratio V_{us}/V_{ud} to relate it to the $|\Delta S| = 1$ decays, gives $V_{ub}^*V_{us}$, which has the phase γ. To account for the SU(3) violation between $|\Delta S| = 0, 1$ decays, a factor of f_K/f_π can be included.

If we now consider the triangle relation for the charge conjugate processes to those in Eq. 1.39, we see that because of its trivial weak phase, the amplitude is unchanged for $K^0\pi^+$ and its conjugate. Hence the two triangles share a common side. Since the $\pi^\pm\pi^0$ decays have the weak phase γ, they will be rotated relative to each other by 2γ. The $K^\pm\pi^0$ decays may exhibit CP-violation and have different branching fractions, and thus the sides may have different lengths between the two triangles. This gives the picture shown in Figure 1.6, where 2γ is determined up to an eight-fold ambiguity.

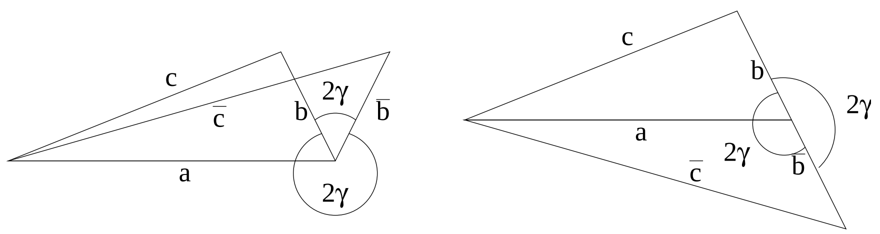

Figure 1.6: Triangle diagram for rare B decays and their conjugates. Side a is $A(K^0\pi^\pm)$, b and $\bar{b}$ are $\sqrt{2}\frac{f_K}{f_\pi}\frac{V_{us}}{V_{ud}}A(\pi^\pm\pi^0)$, and c and $\bar{c}$ are $\sqrt{2}A(K^\pm\pi^0)$.

This method is appealing in its simplicity. It requires the measurement of three decays for $B^\pm$, only one of which is expected to exhibit a CP-asymmetry. All three decay modes are also self-tagging. It would seem that this method is perfect for measuring γ, however, there are difficulties. While some SU(3) breaking effects can be accounted for in the ratio f_K/f_π, others are non-factorizable and introduce theoretical errors to this method. Also, electroweak penguins are, somewhat surprisingly, not negligible when compared to gluonic penguins. Since the magnitude of the electroweak penguins is unknown but thought to be of the same order as the tree level amplitudes, it introduces a large theoretical error that ruins the measurement of γ. Finally, the large number of discrete ambiguities makes it hard to rule out regions of the (ρ, η) plane. All of these effects combine to make this particular triangle construction difficult to interpret.

It should be noted that there are other triangle constructions, similar to the one above, that are also affected by electroweak penguins. Some of these triangles are more difficult experimentally, as they require the measurement of CP-violation in modes such as $\pi^+\pi^-$, which are not self tagging. Experimentally, this means that the B must be tagged by identifying the flavor of the other B meson produced, which makes a difficult analysis even harder. Other experimentally difficult aspects of these triangles are the measurement of modes like $\pi^0\pi^0$, which are expected to have $\mathcal{O}(10^{-7})$ branching fractions.

Perhaps the biggest problem plaguing the triangle construction above is the contamination from electroweak penguins. In order to deal with this problem, more elaborate constructions have been proposed that involve the construction of amplitude quadrangles.

Since B mesons are isospin 1/2 particles, and the effective weak Hamiltonian has both $\Delta I = 0, 1$ pieces, there are both $I = 1/2$ and $3/2$ final states. The set of four $B \to K\pi$ decays are described by the sum of the two isospin pieces, and can be arranged to give the following relation [9].

$$A(K^0\pi^+) + \sqrt{2}A(K^+\pi^0) = A(K^+\pi^-) + \sqrt{2}A(K^0\pi^0) = A_{3/2} \tag{1.40}$$

With the appropriate choice of phase convention, this equation can be drawn in the complex plane as seen in Figure 1.7. One side, f, is the amplitude $A_{3/2}$. The other side, e, is given by

$$e = -C' - (c_u - c_d)P'_{EW} - A' \tag{1.41}$$

which, neglecting small exchange and annihilation diagrams, is the same (to a factor of $\sqrt{3}$) as the amplitude for $B_s \to \pi^0\eta$ decays (Eq. 1.42).

$$\sqrt{3}A(B_s \to \pi^0\eta) = -C' - (c_u - c_d)P'_{EW} - E' \tag{1.42}$$

With all of this information, the quadrilateral is fixed, up to discrete ambiguities. The corresponding charge conjugate quadrilateral can be constructed in the same way. The orientation is fixed by the decay $B^+ \to K^0\pi^+$, which is expected to have a trivial weak phase. Here, the diagonal f is the conjugate amplitude $\bar{A}_{3/2}$. Since the $I = 3/2$ piece has two components, one from the tree level diagrams, and one from electroweak penguins, it can be written

$$A_{3/2} = |A^T|e^{i\gamma}e^{i\delta_T} - |A^{EWP}|e^{i\delta_{EWP}} \tag{1.43}$$

$$\bar{A}_{3/2} = |A^T|e^{-i\gamma}e^{i\delta_T} - |A^{EWP}|e^{i\delta_{EWP}} \tag{1.44}$$

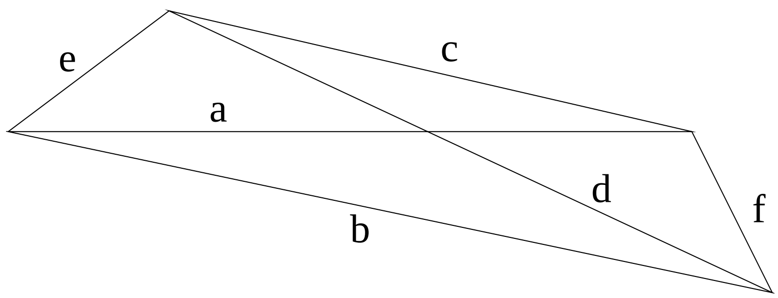

Figure 1.7: Quadrilateral diagram for $B \to K\pi$ decays. Side a is $A(B^+ \to K^0\pi^+)$, b is $\sqrt{2}A(B^+ \to K^+\pi^0)$, c is $\sqrt{2}A(B^0 \to K^0\pi^0)$, d is $A(B^0 \to K^+\pi^-)$, e is $\sqrt{3}A(B_s \to \eta\pi^0)$, and f is the amplitude $A_{3/2}$.

where the δ's are strong phase differences relative to $K^0\pi^+$.

Thus the difference, by eliminating the electroweak penguin contribution, tells us about $\sin\gamma$,

$$A_{3/2} - \bar{A}_{3/2} = |A^T| 2i \sin\gamma e^{i\delta_T} \tag{1.45}$$

To good approximation, the quantity $|A^T|$ can be obtained from

$$|A^T| = \lambda(f_K/f_\pi)\sqrt{2}|A(B^+ \to \pi^+\pi^0)| \tag{1.46}$$

Thus, a measurement of $\sin\gamma$ is possible. Additionally, the strong phase δ_T can also be determined with this method.

While this method addresses electroweak penguin contributions, it is experimentally quite challenging. First, the B_s meson is not produced in the typical running of the e^+e^- machines, so the branching ratio may not be measured in the near future. Furthermore, this branching ratio is estimated to be $\mathcal{O}(10^{-7})$, making observation difficult. Also, this method requires measuring the branching fractions for B and $\bar{B}$ mesons separately. Like the triangle construction above, this method also has several discrete ambiguities, making it harder to rule out values of $\sin\gamma$.

Again, it should be noted that there are other quadrilateral diagrams that one can construct to determine angles of the CKM matrix. They are generally of similar complexity to the one shown here, usually requiring the measurement of B_s meson branching ratios to quantify electroweak penguin contributions.

Time-Dependent CP Asymmetries

There are a number of time-dependent measurements of rare B decays that can give information about the CKM parameters. Here we will discuss the measurement of $\sin 2\alpha$ using the $B \to \pi\pi$ decays.

We start by defining the usual parameters of CP-violation in mixing. The light and heavy mass eigenstates of the neutral B mesons can be written

$$|B_L\rangle = p|B^0\rangle + q|\bar{B}^0\rangle \tag{1.47}$$

$$|B_H\rangle = p|B^0\rangle - q|\bar{B}^0\rangle. \tag{1.48}$$

In addition, the amplitudes A_f and $\bar{A}_f$ are defined

$$A_f = \langle f|H|B^0\rangle \tag{1.49}$$

$$\bar{A}_f = \langle f|H|\bar{B}^0\rangle \tag{1.50}$$

and the parameter λ is

$$\lambda = \frac{q}{p}\frac{\bar{A}_f}{A_f}. \tag{1.51}$$

By analyzing the time evolution of the $\Upsilon(4S)$ state, it can be shown that the time-dependent CP-asymmetry is given by

$$a_{CP}(t) = \frac{(1-|\lambda|^2)\cos(\Delta m_B t) - 2\mathrm{Im}(\lambda)\sin(\Delta m_B t)}{1+|\lambda|^2} \tag{1.52}$$

For the decay $B^0 \to \pi^+\pi^-$, we start by assuming that the penguin contribution is negligible. In that case,

$$\frac{\bar{A}_{\pi\pi}}{A_{\pi\pi}} = \eta_{\pi\pi}\frac{V_{ub}V_{ud}^*}{V_{ub}^*V_{ud}}, \tag{1.53}$$

$$\frac{q}{p} = \frac{V_{tb}^*V_{td}}{V_{tb}V_{td}^*} \tag{1.54}$$

which leads to

$$\lambda_{\pi\pi} = \frac{V_{tb}^*V_{td}}{V_{tb}V_{td}^*}\frac{V_{ub}V_{ud}^*}{V_{ub}^*V_{ud}} \tag{1.55}$$

First, we notice that $|\lambda_{\pi\pi}| = 1$, so that the expression for the time-dependent asymmetry simplifies to

$$a_{CP} = -2\mathrm{Im}(\lambda)\sin(\Delta m_B t) \tag{1.56}$$

and $\mathrm{Im}(\lambda_{\pi\pi}) = \sin 2\alpha$.

This implies that a time-dependent measurement of the rates for $B^0 \to \pi^+\pi^-$ and $\bar{B}^0 \to \pi^+\pi^-$ would give us α. However, we have neglected the penguin contribution to the amplitude. To extract α, we need to separate the penguin and tree amplitudes. We will use a triangle construction similar to the one used for $K\pi$ above.

Again, the isospin relation between several amplitudes is the key to this method. For the $\pi\pi$ system, the amplitude equation is

$$A^{+-} + \sqrt{2}A^{00} = \sqrt{2}A^{+0} \tag{1.57}$$

with $A^{ij} = A(B \to \pi^i \pi^j)$. The charge conjugate relation is similar.

$$\bar{A}^{+-} + \sqrt{2}\bar{A}^{00} = \sqrt{2}\bar{A}^{-0} \tag{1.58}$$

Since the $\pi^+\pi^0$ final state must be $I = 2$, it can only occur through the $I_{3/2}$ tree diagram. Thus $|A^{+0}| = |\bar{A}^{-0}|$, and the two triangles have a common side. This then gives the picture of Figure 1.8.

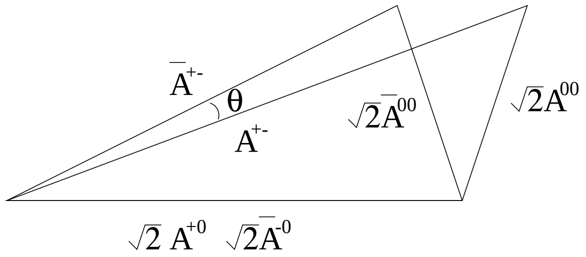

Figure 1.8: Triangle diagram for the decays of B and $\bar{B} \to \pi\pi$.

The measurement of the time-dependent asymmetry of $B \to \pi^+\pi^-$, along with the CP-averaged measurements of $B \to \pi^+\pi^0$ and $B \to \pi^0\pi^0$ allows us to measure the magnitude and relative weak phase of $A^{+-}/\bar{A}^{+-}$. Since the observable CP-asymmetry is given by

$$\mathrm{Im}(\lambda_{\pi\pi}) = \mathrm{Im}(e^{-2i\alpha}\frac{A^{+-}}{\bar{A}^{+-}}) \tag{1.59}$$

we can now extract α cleanly.

One theoretical problem with this method is that it does not include the effect of electroweak penguins. Since electroweak penguins can have an $I_{3/2}$ component, the amplitudes A^{+0} and $\bar{A}^{-0}$ can no longer necessarily be rotated by a strong phase to have a common base. Hence there is an uncertainty due to the magnitude of electroweak penguin contributions that would require theoretical or experimental input to estimate.

1.3 Summary

The measurement of the CKM parameters ρ and η is of fundamental importance in particle physics. We have shown that there many ways to use two-body, charmless pseudoscalar-pseudoscalar B decays to gain insight into the CKM mechanism, using an array of theoretical approaches. These techniques require the measurement of a variety of experimental quantities, ranging from CP-averaged branching ratios to time-dependent CP-asymmetries. As both theoretical methods and experimental results improve, we expect these rare B decays to provide significant constraints on the (ρ, η) plane, and perhaps point the way to physics beyond the Standard Model.

CHAPTER 2

THE EXPERIMENT

In order to truly understand how any particle physics analysis is done, and in particular the capabilities and limitations of the analysis, it is essential to have some knowledge of the experimental apparatus used to collect the data. This analysis relies on two major pieces of equipment, the accelerator used to generate B mesons, and the detector used to observe their decays. In this chapter, an overview of both of these components will be given.

2.1 CESR

The Cornell Electron Storage Ring (CESR) produces the B mesons that CLEO studies. Understanding how CESR operates helps to understand how CLEO does its job.

While the acronym CESR refers to a storage ring, the CESR facility is actually three separate accelerator components: the linear accelerator (LINAC), the synchrotron, and the storage ring. A schematic depiction of CESR is shown in Figure 2.1.

The LINAC generates its electrons by heating a filament to high temperatures.

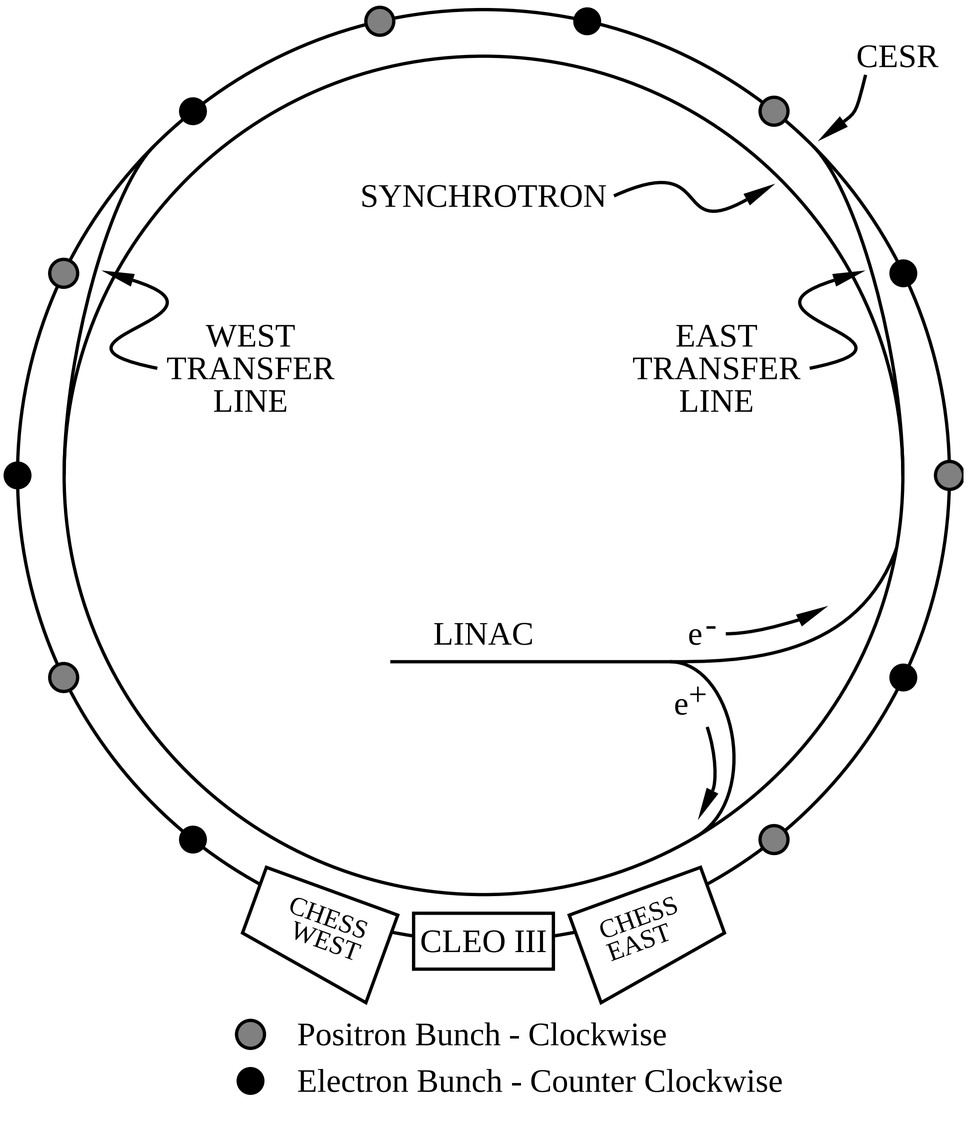

Figure 2.1: The CESR facility.

The electrons have sufficient energy to escape from the filament. The electrons then pass through a prebuncher that compresses the electrons into packets small enough to be accelerated by the accelerating sections of the LINAC. Oscillating electric fields in the accelerating sections are synchronized so that as each electron bunch passes through, the electric field points in the proper direction to accelerate the electrons. By the time the electrons reach the end of the LINAC, they have been accelerated to an energy of about 300 MeV.

Positrons, not found in typical matter, have to be created in the LINAC. Electrons are accelerated in the first part of the LINAC, then hit a heavy metal target. Here, electrons, positrons and photons are created in a shower of particle debris. The positrons are separated from the electrons and accelerated down the remainder of the LINAC.

Once the positrons or electrons reach the end of the LINAC, they enter the synchrotron. Here they are accelerated to the desired beam energy, usually between 5.27 and 5.3 GeV. Accelerating sections in the synchrotron increase the energy of the particles at four stations around the ring. In order to keep the particles in the same orbit as they gain energy, an increasing magnetic field is required. The magnets of the synchrotron are wire coils that carry an alternating current at a frequency of 60 Hz. At the beginning of every sinusoidal oscillation, the current in the coils, and therefore the magnetic field in the synchrotron, increases nearly linearly with time. This portion of the 60 Hz oscillation is the time used by the synchrotron to provide an increasing magnetic field while the particles are accelerated by four 3 meter long linear accelerators to their final energy. The particles are then transferred to the storage ring to be used by CLEO. This process can be repeated sixty times per second until the beams in the storage ring have accumulated the desired current.

Once the electrons and positrons are transferred to the storage ring, they are stored there until the beams have lost enough current to warrant refilling the machine. As the particles are steered around the ring by dipole magnets, they lose energy to synchrotron radiation. This energy is replaced by superconducting radio frequency cavities. Both the positrons and the electrons in the storage ring travel in nine well separated trains, with each train containing as many as five bunches. Electrostatic separators in the storage ring pull the orbits of the electrons and positrons apart so that the trains of electrons can be steered around the trains of positrons in the so-called pretzel pattern. This allows the electron and positrons to collide only at the interaction point, which improves luminosity of the beam. At the interaction point, the beams do not collide head on, but with a small crossing angle. This allows the first bunch in the electron train to collide only with the first bunch of the positron train, the second electron bunch to collide only with the second positron bunch, and so on. This allows a significant increase in luminosity of the beams. The timing of the bunch crossings is also critical for CLEO, since it provides the clock by which the electronic read out of some of the sub-detectors are timed.

2.2 The CLEO Detector

Once the electrons and positrons have been accelerated to high energies and brought into the storage ring, they collide at the interaction point, where the CLEO detector is situated. Regardless of the type of interaction between the electrons and positrons, all of the particles produced will eventually decay to final states made up of relatively long-lived particles. Some of these particles can be detected (charged particles, photons) while others are either impossible or very difficult to detect (neutrinos, neutrons, K_L^0's). Here we will briefly describe how we measure some of the properties of these particles, focusing on the aspects that are most relevant to two body hadronic B decays.

The CLEO detector is a multi-purpose detector with several sub-detector components. The entire detector is essentially cylindrically symmetric, with its axis coinciding with that of the positron and electron beams. The detector is also very hermetic, with some elements covering 93% of the solid angle. A quarter section view is shown in Figure 2.2.

2.2.1 Silicon Detector

The innermost device is the silicon detector, which is a four layer device that measures the position of charged particles with high resolution. The end view and side view are shown in the top and bottom of Figure 2.3, respectively. The basic component of the silicon detector is a wafer 1 inch wide by 2 inches wide by 300 μm thick. When charged particles pass through the silicon, they ionize the atoms

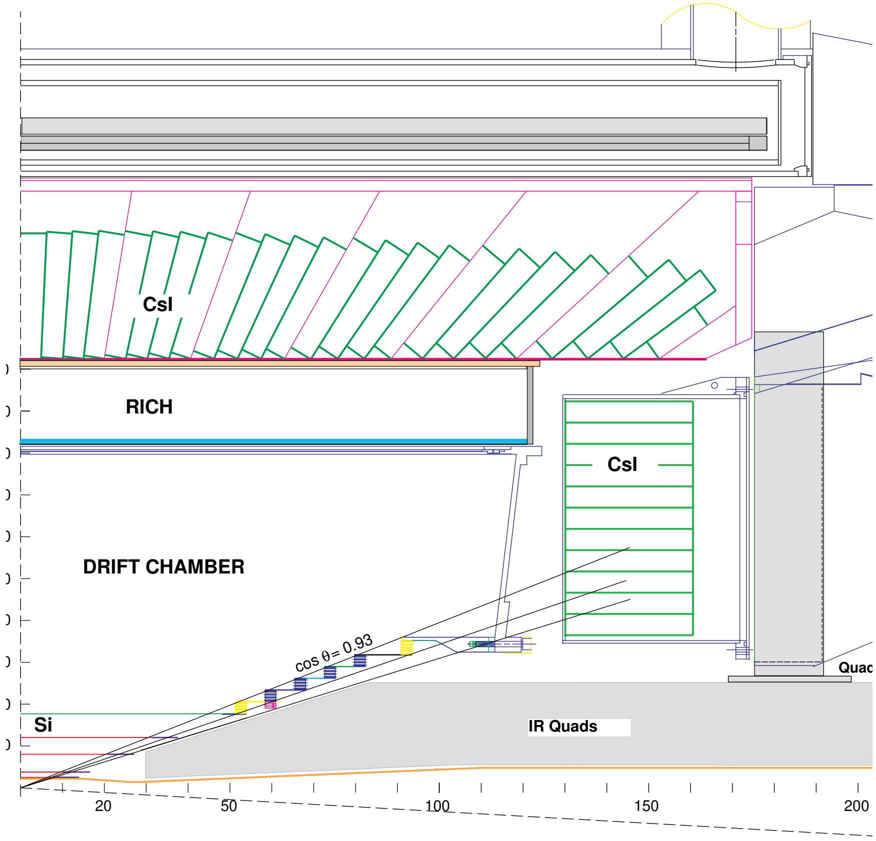

Figure 2.2: Cross-section of the CLEO III detector in the $r-z$ plane.

in the material, creating electron-hole pairs. An electric field sweeps electrons to one side, and the holes to the other. Traces running the length of the wafer on one side and running the width of the other side measure the amount of ionization caused by the passage of the charged particle. This information can be used to measure the $r-\phi$ and $r-z$ position of the particle quite precisely.

The silicon detector is crucial to any analysis that needs excellent vertex resolution of particles that decay near the beam spot. None of the decay modes presented in this analysis make use of this vertexing capability, but the silicon is the only source of z information near the beam axis, which does help improve momentum resolution.

2.2.2 Drift Chamber

Immediately outside the silicon detector is the drift chamber, which is the main tracking chamber. It consists of 47 layers of cells, each cell consisting of one sense wire surrounded by field wires, aligned parallel to the beam axis. The sense wires are held at a large (>2000V) positive voltage relative to the field wires. The chamber is filled with a helium-propane gas mixture, and when charged particles pass through the gas, they ionize the atoms. The voltage between the sense wires and field wires produces a very non-uniform electric field that attracts the electrons to the sense wires. The fields near the sense wire are large enough to cause the electrons to ionize more atoms on the gas, which in turn ionize yet more atoms, causing an avalanche of electrons to be collected at the sense wire. The time of arrival of this pulse can be converted into a distance of closest approach to the

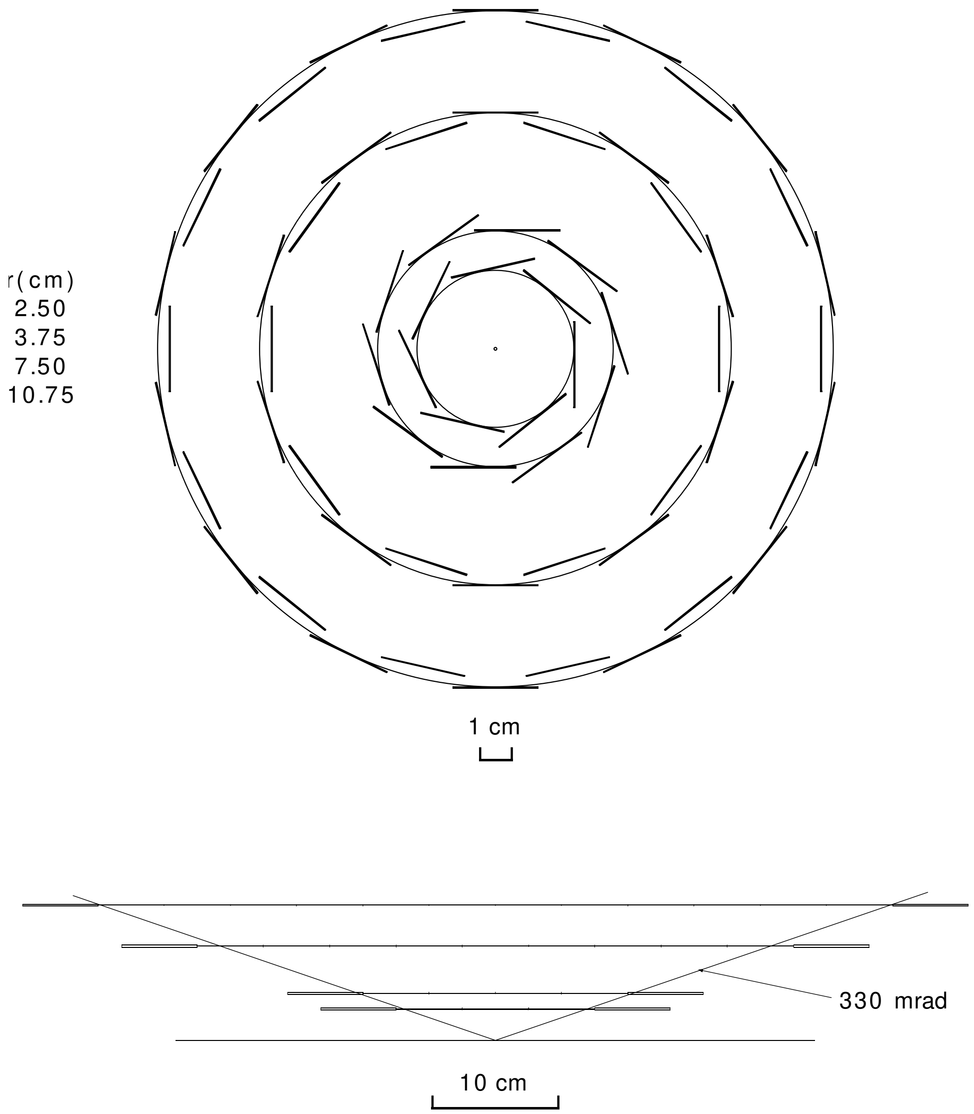

Figure 2.3: The silicon detector. The top drawing is a cross-section in the $r-\phi$ plane, and the bottom drawing is a cross-section in the $r-z$ plane.

wire by knowing how fast the electrons drift through the various regions of the drift cell. The precisely known timing structure of the bunches in the storage ring provides the timing needed for the drift chamber measurements. Measuring the drift distance gives information about the trajectory of the charged particle in the $r - \phi$ plane. The outer 31 layers of the drift chamber are not aligned exactly parallel to the beam axis, but are arranged in groups of four layers that have alternating stereo angles, meaning that the ϕ coordinate at the two ends of the wire are not the same. This small stereo angle gives additional spatial information in the z direction. Pattern recognition software attempts to identify all of the hits associated with a particular charged particle trajectory. Since the entire drift chamber is inside a solenoidal magnetic field, all charged particles will have curvature in the $r - \phi$ plane. Precisely reconstructing the trajectory of the particle yields information about the momentum of the particle.

A variety of secondary effects cause the behavior of the charged particle trajectory to deviate from the ideal described above. As the particle ionizes the material through which it passes, it loses energy. This makes the curvature of the particle increase as it moves outward. Multiple scattering induces small random deviations in the path of the particle. A second piece of software, the Kalman fitter, attempts to account for these effects to get the best measure of the track momentum as possible. It also takes into account things like entrance angle corrections, finite signal propagation time along the sense wires, and time of flight corrections.

The drift chamber provides almost all of the momentum information about charged tracks. The momentum resolution is a function of the spatial resolution of the individual hits and the lever arm over which the trajectory is measured. The CLEO III drift chamber has performed beautifully, giving momentum resolution comparable to that of its predecessor in CLEO II, despite the smaller tracking volume to accommodate a particle identification system. This momentum resolution is important for the rare B analyses for two reasons. First, it provides a sharper peak to look for amongst the large backgrounds. Second, it provides kinematic separation between decay modes that only differ by the strangeness of a charged track. The momentum resolution of 2.6 GeV/c tracks ranges between 13 and 16 MeV/c.

The drift chamber also serves a second purpose. The amount of energy that a charged particle loses as it passes through material is a function of its velocity. The charge collected by a sense wire is a measure of the energy loss of a particle, and the energy loss per unit length (dE/dx) can be compared to that which would be expected for particles with pion and kaon mass hypotheses. Energy loss is a random process, and as such, subject to statistical fluctuations and therefore is not a perfect means to identify particle type, but it does provide valuable particle identification information.

2.2.3 RICH Detector

After the drift chamber is the Ring Imaging Cherenkov (RICH) detector. It is the primary particle identification device for high momentum hadrons, and the most notable improvement over the CLEO II detector. The main physics principle behind the operation of a RICH detector is that when the speed of light in a material is less than the speed of a charged particle passing through that material, it will emit a cone of photons. The apex angle of this cone is given by Eq. 2.1

$$\cos\alpha = 1/(n\beta), \tag{2.1}$$

where n is the index of refraction of the material, and β is the speed of the particle in units of c, the speed of light in vacuum. A measurement of the Cherenkov light gives the speed of the particle, and combined with a measurement of the momentum of the particle, gives its mass, which tells us which kind of charged particle we have.

The RICH detector in CLEO III covers the region $|\cos\theta| < 0.8$. At a radius of about 82 cm are the radiators, which are LiF planes 1 cm thick with an index of refraction of 1.5. Near $\cos\theta = 0$, the LiF crystals have a sawtooth geometry so that not all of the photons are lost to internal reflection. Beyond the radiators is a nitrogen expansion volume. This allows the photons to spread out, allowing the apex angle of the Cherenkov light to be accurately determined.

Cherenkov photons are detected using photosensitive triethylamine (TEA) gas. In a manner much like the drift chamber, the photoelectrons are accelerated towards the sense wires of a multi-wire proportional chamber (MWPC). The charge is recorded by a two-dimensional array of cathode pads, which gives the position of the photons.

For every charged particle, the trajectory is extrapolated to the LiF radiators so that Cherenkov angles can be calculated. Resolution on the z coordinate of the intersection point can be improved if a cluster in the MWPC can be associated with the passage of the charged particle. Photons within $\pm 5\sigma$ of the expected Cherenkov angle for a given mass hypothesis are used in the particle identification. For each mass hypothesis, a likelihood is calculated that takes into account both the Cherenkov angle and the photon yield. For the high momentum tracks of interest in this analysis, the expected number of photons is essentially the same, so the separation comes mainly from the measurement of the Cherenkov angle. The difference of $-2\ln\mathcal{L}$ for the kaon and pion hypotheses should be akin to a $\Delta\chi^2$ quantity, and the cut on this value can be chosen to optimize the efficiency and fake rates.

The resolution of the RICH detector near $\cos\theta = 0$ is dominated by the chromatic error on the photons, and at large values of $|\cos\theta|$ by the track extrapolation and emission point errors. For high momentum pions and kaons, we find fake rates less than 10% when the efficiency is required to be 90%.

2.2.4 Crystal Calorimeter

Photon detection is performed by the crystal calorimeter, which is the next subdetector in CLEO. The calorimeter is divided into two sections, the barrel region which is immediately outside the RICH detector, and the endcap region which is beyond the ends of the drift chamber. The calorimeter is built from 7784 cesium iodide crystals, doped with thallium. Each crystal is 30 cm long with a 5 cm by 5 cm face. Crystals in the endcap are parallel to the beam axis, while the crystals in the barrel region are tilted so that they point towards the intersection point, but with a small offset to minimize inefficiencies from photons passing between crystals.

When photons pass through the crystals, they produce a shower of electrons and photons that excite the atoms in the crystal. As the atoms de-excite, they emit light. This light is collected by photodiodes attached to the back of the crystals, and the signal can be converted into an energy. Photons will typically deposit energy in several crystals, and the energy from these crystals is clustered together to measure the photon energy. Energy resolution of the calorimeter is affected by statistical fluctuations in the shower, electronic noise, non-uniformities in light collection, and leakage of energy out of the crystals. Photons in the transition region between the barrel and endcap portions of the calorimeter, along with photons in the crystals nearest the beampipe are not used to form π^0 candidates.

Charged particles and neutral hadrons can also deposit energy in the calorimeter. The trajectories of charged particles are extrapolated to the calorimeter and matched to showers. Only unmatched showers will be considered as potential photon candidates. Neutral hadrons also interact in the calorimeter, and they can be rejected by examining the lateral shape of the showers. Electrons will usually deposit all of their energy in the calorimeter, and the ratio of of energy over momentum (E/p) can be used as particle identification.

2.2.5 Muon Detector

The muon detectors are the last element in CLEO. They are situated outside the solenoid coils and are interleaved between the layers of iron that serve as the flux return for the magnet. Since muons are the only charged particle that can readily penetrate these thicknesses of iron, detection of charged particles in this region can be used for particle identification. Muon identification plays only a very minor role in this analysis.

2.3 The CLEO III Data

This analysis uses most of the CLEO III $\Upsilon(4S)$ data. An initial pass2 of the data was done in 2001, and a subsequent repass2 is in progress. These two versions of pass2 will be referred to as old and new pass2 respectively. The tracking resolution in the new pass2 is considerably better than that of the old pass2. The list of datasets used in this analysis, their approximate size, and version of pass2 is listed in Table 2.1.

Another major difference between the old and new pass2 versions is the event filtering that was used. The old pass2 used tighter event filters than the new pass2, resulting in lower efficiencies for the rare B modes. To account for this difference in efficiency, we used signal Monte Carlo from last summer that used the old pass2 and its filters to get efficiencies for old pass2 data, and generated new Monte Carlo to find the efficiency in the new pass2.

2.3.1 Number of $B\bar{B}$

At the time of this analysis, no accurate determination of the number of B mesons produced during the CLEO III running had been made. We therefore needed to measure $N_{B\bar{B}}$ and also quantify the error on this number.

Table 2.1: Summary of the datasets used in this analysis.

Dataset	On (pb^{-1})	Off (pb^{-1})	Pass2
data6	630	-	new
data7	780	-	new
data8	880	-	new
data9	680	220	old
data10	470	530	old
data11	660	270	new
data12	470	640	old
data13	830	240	old
data14	780	340	old

We have explored three different methods for counting $N_{B\bar{B}}$: an R_2 subtraction, counting B decays to exclusive final modes with relatively large branching ratios, and a luminosity times cross-section method.

R_2 Method

With the R_2 method, we plot the R_2 distribution for all hadronic events in the on- and off-resonance data, then subtract the off-resonance distribution from the on-resonance distribution, using the appropriate weighting. One advantage of this method is that the R_2 distribution for $B\bar{B}$ events goes to zero for values near one. If the subtraction has been done properly, the R_2 distribution should be essentially zero from 0.6 to 1.0.

This subtraction can be performed for a particular data set. For example, Figure 2.4 shows R_2 distributions for data14. The scale factor used for the off-resonance data was the ratio of on-resonance luminosity to off-resonance luminosity. There is clearly excellent agreement between the two curves for large values of R_2. Varying the endpoint of the integration between 0.6 and 0.9 gives a variation in $N_{B\bar{B}}$ of 1%.

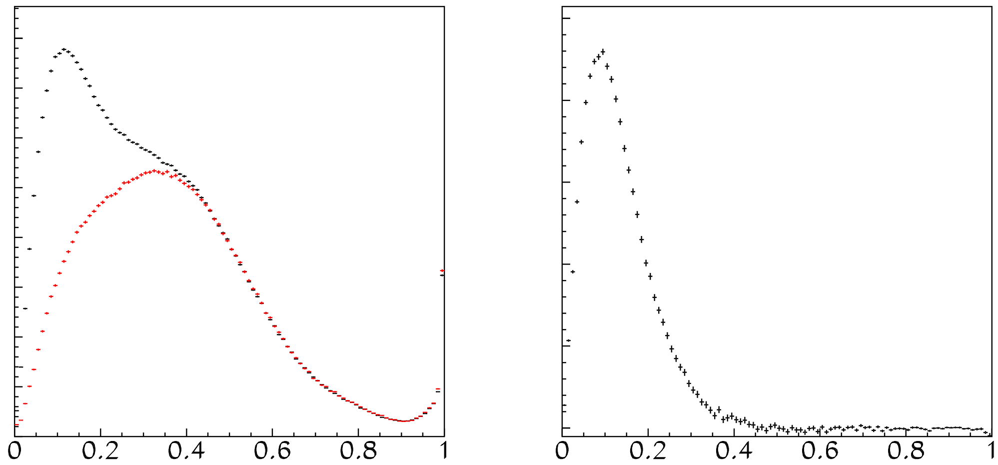

Figure 2.4: On the left, data14 R_2 distribution for on-resonance data and luminosity-scaled off-resonance data. The plot on the right shows the result of the subtraction.

The difficulty with this method is that there was no off-resonance data taken for datasets 5 through 8. Any attempts to do subtractions using off-resonance data from later datasets give unsatisfactory results. Some of this effect can be attributed to the different event filters used in old and new pass2 (see Figure 2.5). It is impossible, however, to get a subtraction that is good enough to determine $N_{B\bar{B}}$. The R_2 distributions and luminosity-scaled subtractions are shown in Figures 2.6 and 2.7. For data5-8, where there is no off-resonance data, we use data11 as the source of our off-resonance data. The $1/s$ cross-section dependence has been accounted for in the luminosity scaling.

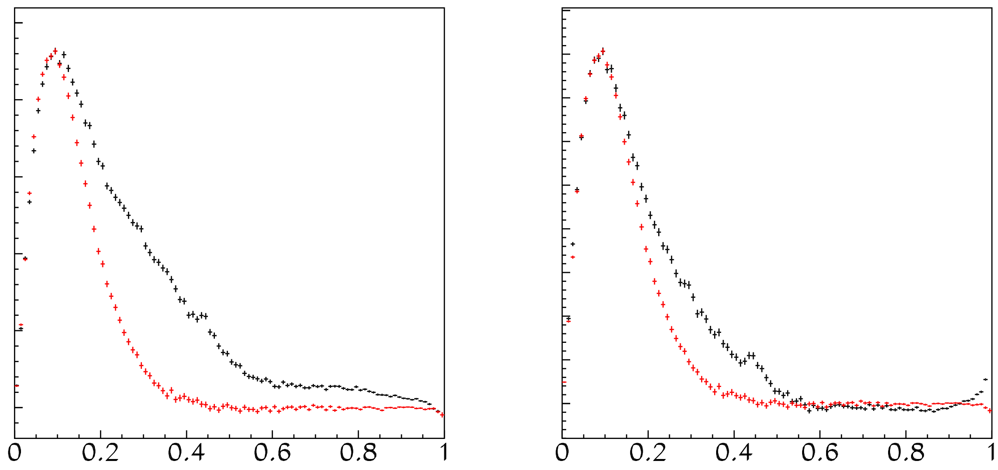

Figure 2.5: Results of luminosity-scaled R_2 subtractions for data5 (new pass2) on-resonance data. On the left, we use data9 off-resonance data (old pass2), and on the right we use data11 off-resonance data (new pass2). The $B\bar{B}$ shape from data14 (lower curve) is shown in both plots for comparison.

We can attempt to estimate $N_{B\bar{B}}$ by integrating the R_2 distribution up to some value, then dividing by the efficiency for that choice of R_2 cut. We see deviations between MC and data R_2 distributions (Figure 2.8), so we did the calculation once

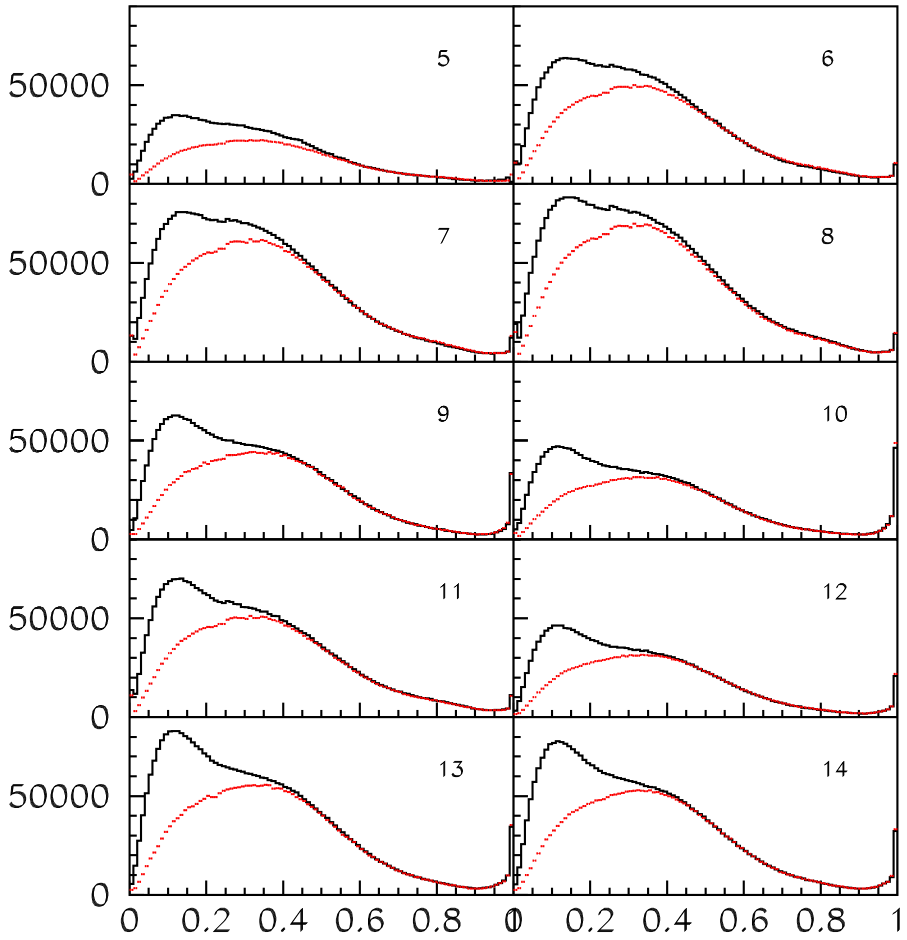

Figure 2.6: On-resonance R_2 distribution and luminosity scaled off-resonance R_2 distribution for each dataset.

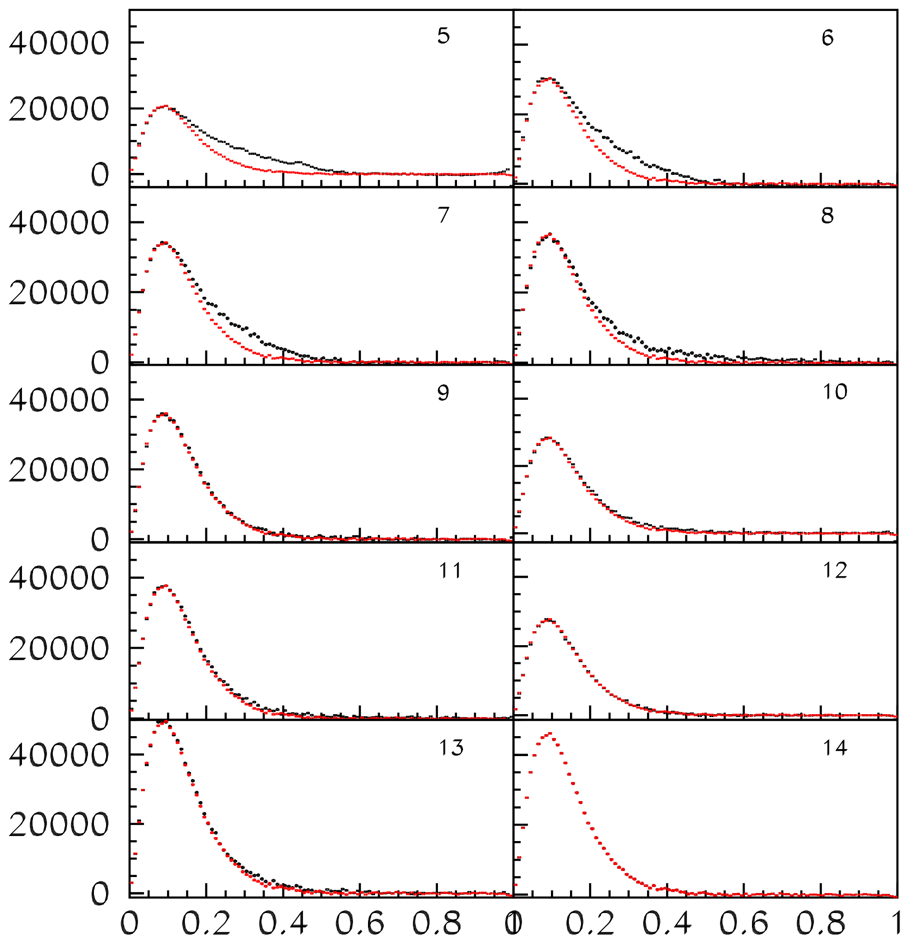

Figure 2.7: Luminosity-scaled R_2 subtractions for each dataset. For data5-8, the upper curve is the result of the subtraction, and the lower curve is the shape from data14, which we take to represent the $B\bar{B}$ R_2 distribution.

using the efficiency as found in MC, and once using the efficiency as found in data. The results are shown in Tables 2.2 and 2.3.

Table 2.2: $N_{B\bar{B}}$ (1000's) by dataset, integrating the R_2 distribution to the indicated point and using MC to calculate the efficiency.

Dataset	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
5	318	364	414	452	476	481	482	481	479	479
6	482	539	597	634	647	648	643	637	632	630
7	535	595	651	688	700	700	698	695	690	689
8	569	625	665	689	707	718	727	732	735	735
9	566	611	627	630	630	631	631	630	630	628
10	444	483	500	507	511	513	514	515	516	515
11	597	644	664	672	676	679	681	684	685	686
12	434	466	476	477	475	473	472	471	471	469
13	784	844	866	876	880	883	884	884	885	884
14	727	777	793	792	790	788	787	787	786	785

It is also interesting to compare the result of the R_2 subtraction in data11 where we have both old and new pass2. The subtracted R_2 distributions are shown in Figure 2.9. There is a 3% decrease in $N_{B\bar{B}}$ going from new pass2 to old pass2. Of this, 1.5% can be attributed to differences in which runs we processed in the two samples, which means that the other 1.5% difference is likely due to the looser event filters used in the new pass2. These differences have errors of order 1%. It is

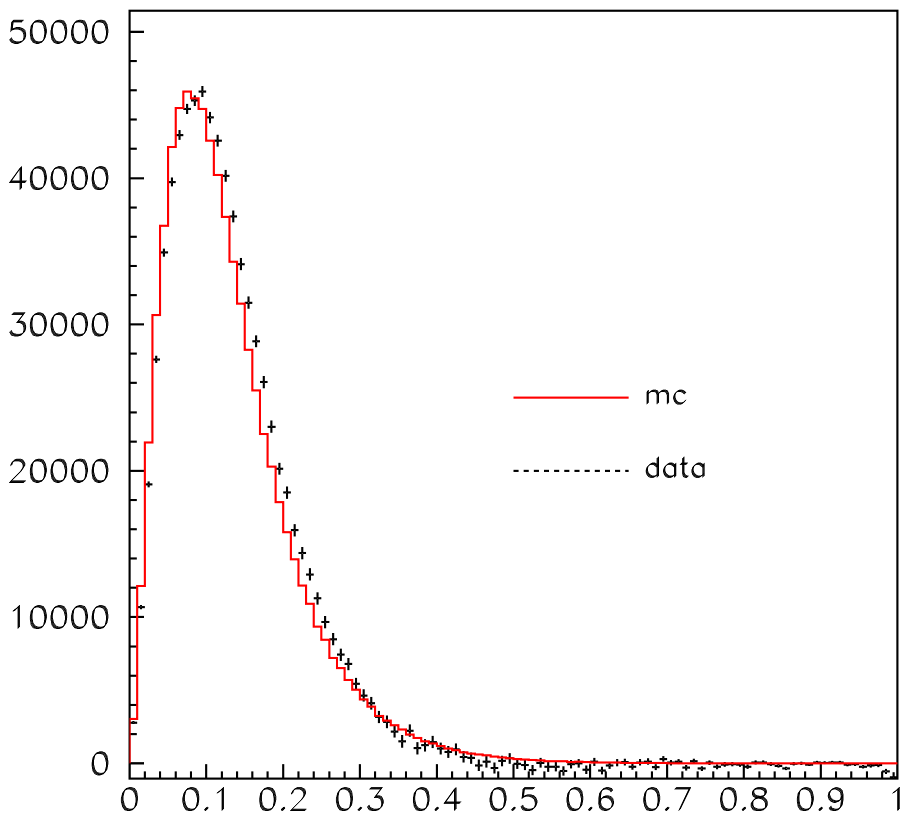

Figure 2.8: R_2 distribution for $B\bar{B}$ events as determined in data (points) and MC (line).

Table 2.3: $N_{B\bar{B}}$ (1000's) by dataset, integrating the R_2 distribution to the indicated point and using data to calculate the efficiency.

5	357	378	418	453	476	481	482	481	479	479
6	541	559	602	636	647	647	643	637	632	630
7	600	617	658	690	700	699	697	695	690	689
8	638	648	672	690	707	718	727	732	735	735
9	635	634	633	631	630	631	631	630	630	628
10	498	501	505	508	511	513	514	515	516	515
11	670	668	671	674	676	679	681	684	685	686
12	487	483	481	478	475	473	472	471	471	469
13	880	875	875	878	880	882	884	884	885	884
14	816	806	801	794	790	787	787	787	786	785

also of note that the fraction of $B\bar{B}$ MC events that satisfy the criteria necessary to make it into the hadron datasets for new pass2 is 99.85%.

$B^- \to D^0\pi^-$ Method

Another choice for calculating $N_{B\bar{B}}$ is to use a B decay mode with a well known and relatively large branching ratio. In this case, we have chosen $B^- \to D^0\pi^-$, with $D^0 \to K^-\pi^+$, $K^-\pi^+\pi^0$, $K^-\pi^+\pi^-\pi^+$. Our requirements for B candidates are

- $|\Delta E| < 50$ MeV

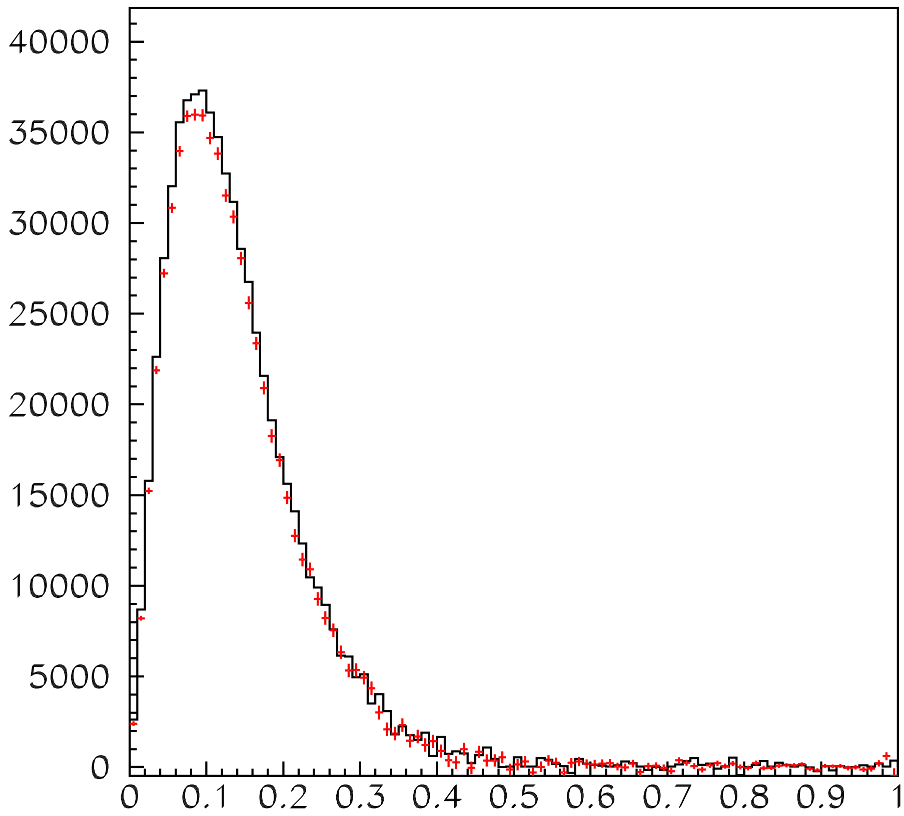

Figure 2.9: R_2 subtractions for data11 new (line) and old (error bars) pass2.

- $|\cos\theta_{sp}| < 0.8$
- $|M_{D^0} - 1864.5 \text{ MeV}/c^2| < 30 \text{ MeV}/c^2$
- $|M_{\pi^0} - 135 \text{ MeV}/c^2| < 30 \text{ MeV}/c^2$

These cuts are chosen to be loose enough that they do not cut significant amounts of signal in data or MC (we find significant tails in the data that are not present in the MC, as in Figure 2.10). We use signal MC to determine our reconstruction efficiencies. $B^- \to D^0 K^-$ background is accounted for by using signal MC to determine the efficiency for selecting DK events as $D\pi$, multiplied by the ratio of branching fractions from the PDG:

$$\frac{BR(B^- \to D^0 K^-)}{BR(B^- \to D^0 \pi^-)} = 0.055 \tag{2.2}$$

To measure the yield, we count the number of candidates whose M_B is within 10 MeV/c^2 of its nominal value. We estimate the amount of background by counting the number of candidates with $5.24\text{GeV}/c^2 < M_B < 5.26\text{GeV}/c^2$, and scale by a factor of 0.63. The scale factor was determined by fitting the off-resonance M_B distribution to an ARGUS shape and integrating over the signal and background regions. We can estimate the effects of any $B\bar{B}$ backgrounds (excluding DK, which we have accounted for) by looking at the M_B distributions for B candidates found in the CLEO II generic $B\bar{B}$ MC [15]. For all three D^0 submodes, the amount of B backgrounds is less than 0.5%. If there was a 100% difference in the efficiency for these backgrounds between CLEO II and CLEO III (either the backgrounds found in CLEO II were absent in CLEO III, or were twice as large in

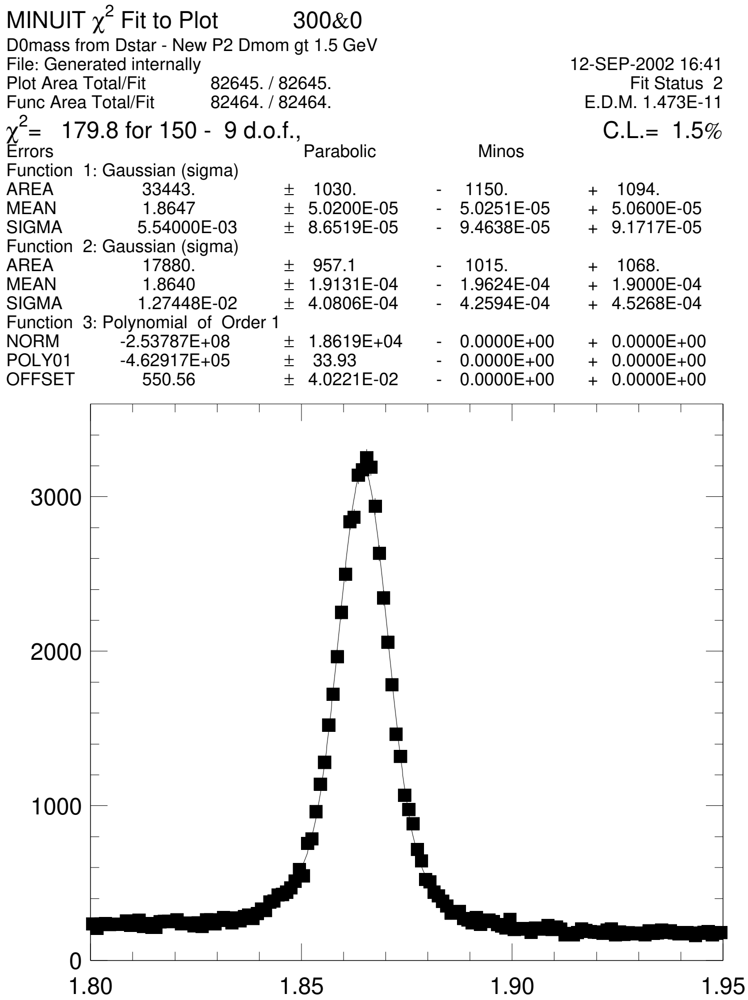

Figure 2.10: D^0 mass peak, using D^* tag. Note the amount of signal in the second gaussian, which is not seen in the MC.

CLEO III as in CLEO II), this would result in a 0.5% offset. If the efficiencies for B backgrounds are similar for CLEO II and CLEO III, we would find similar amounts of background in both data samples, and this already small effect would largely cancel. The mass plots for the three D^0 modes are shown in Figures 2.11-2.13.

To calculate $N_{B\bar{B}}$ in CLEO III, we use the formula

$$N_{B\bar{B}3} = \frac{Y_3 \times \epsilon_2}{\epsilon_3 \times Y_2} \times N_{B\bar{B}2}, \tag{2.3}$$

where the subscripts refer to CLEO II and CLEO III, Y represents the raw yield for a given D^0 submode, and ϵ is the efficiency for that mode. This method has the benefit of removing the error on any D branching ratios from our measurement. We can then consider the three different D^0 submodes to be three separate measurements of $N_{B\bar{B}}$. We combine these three measurements, taking into account any correlations between them (see [16] for the general method for dealing with correlations).

We use the recent measurement in CLEO II of $B^- \to D^0\pi^-$[15] as our yardstick. The yields for CLEO II and CLEO III are shown in Table 2.4. Using our efficiencies and those quoted in the CLEO II CBX, we can calculate $N_{B\bar{B}}$.

To calculate the error on $N_{B\bar{B}}$, we construct a table of all the errors that enter into our calculation (Table 2.5). Each mode also has a 1.4% error for CLEO III MC statistics that is uncorrelated among the modes, and a 2% $N_{B\bar{B}}$ and a 1% DK background error for CLEO II that are correlated among the modes. We can use these errors and the knowledge of the correlations to combine the three D^0 modes to get a single measurement of $N_{B\bar{B}} = (5.73 \pm 0.47) \times 10^6$. The relative weights

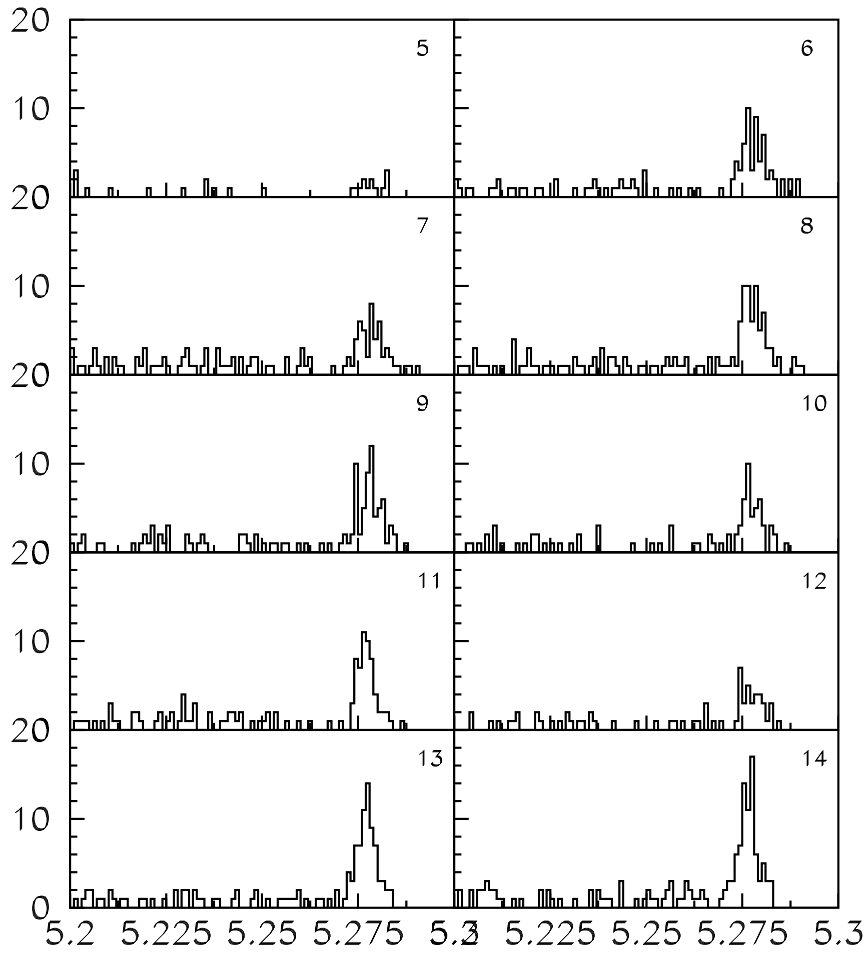

Figure 2.11: M_B distributions for $B^- \rightarrow D^0\pi^-$, with $D^0 \rightarrow K^-\pi^+$, by dataset.

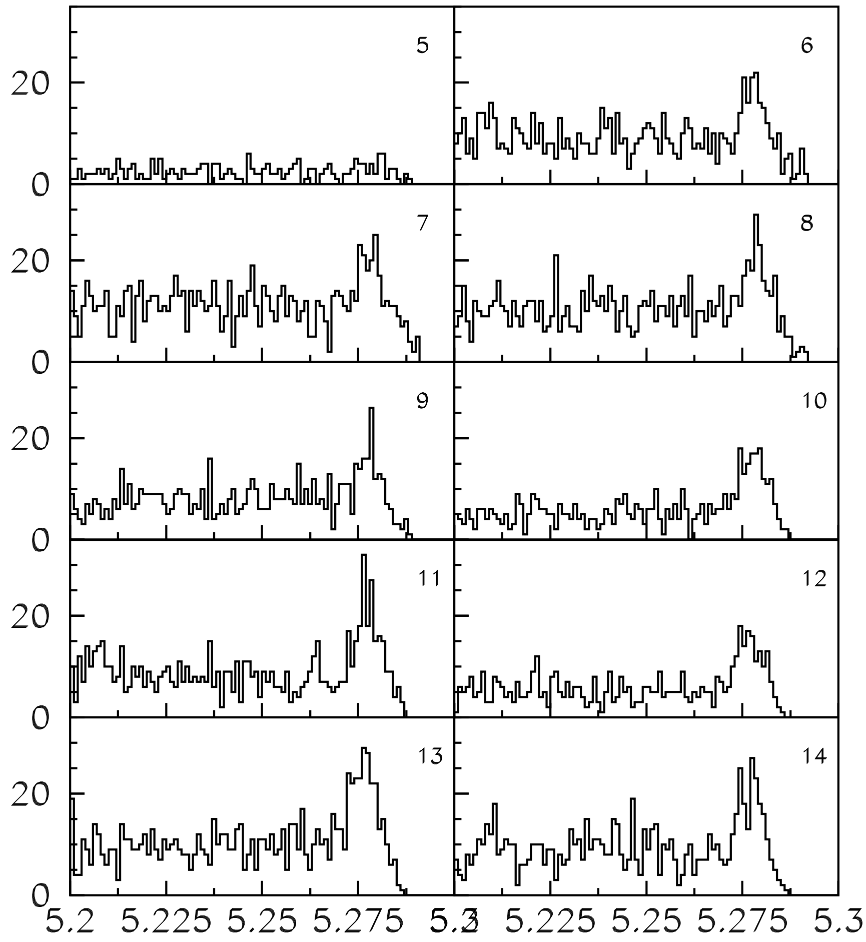

Figure 2.12: M_B distributions for $B^- \to D^0\pi^-$, with $D^0 \to K^-\pi^+\pi^0$, by dataset.

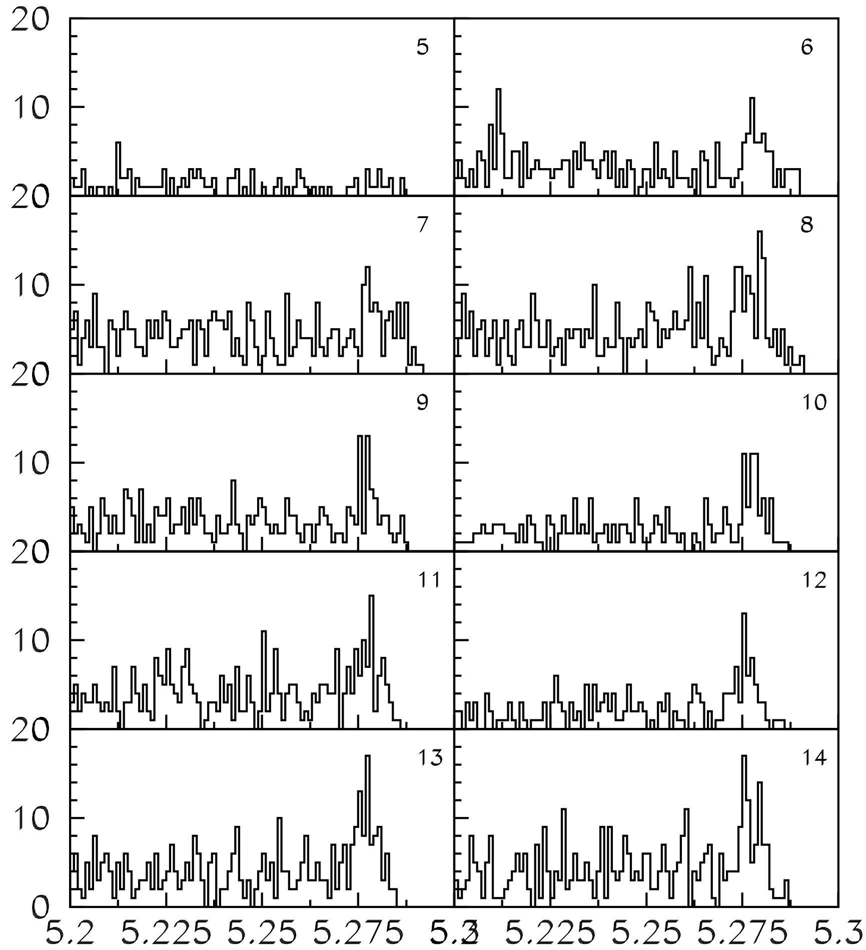

Figure 2.13: M_B distributions for $B^- \to D^0\pi^-$, with $D^0 \to K^-\pi^+\pi^-\pi^+$, by dataset.

Table 2.4: Raw $B^- \to D^0\pi^-$ yields and efficiencies (%) for the three D^0 modes.

Mode	CLEO II	CLEO II.V	CLEO III
$K^-\pi^+$	281±18, 44.5	540±25, 45.9	457±24, 45.4
$K^-\pi^+\pi^0$	448±27, 18.3	751±35, 16.9	963±52, 22.2
$K^-\pi^+\pi^-\pi^+$	257±20, 20.0	483±26, 21.9	414±34, 17.4

of the three D^0 modes are 61.9% for $K^-\pi^+$, 13.6% for $K^-\pi^+\pi^0$, and 24.5% for $K^-\pi^+\pi^-\pi^+$.

We can verify that these numbers are reasonable by comparing the value we get using the branching ratio method to those we get using the R_2 method for those datasets where we have off-resonance data and the R_2 subtraction can be performed. A quick calculation of $N_{B\bar{B}}$ using $B \to D\pi$ for data9-14 gives 3.75±0.16 million B's (statistical error only), while the R_2 method gives 3.94 million B's. The R_2 values are obtained by integrating the subtracted R_2 curves for each dataset into the plateau at high R_2 values (and thus are not identical to the values obtained in Table 2.2 or 2.3). The favorable comparison also holds dataset by dataset (see Table 2.6). This gives us some degree of confidence that the $D\pi$ method gives sensible results. Since there is such a large discrepancy between the two methods for data5, and because data5 is such a small dataset, we have decided to exclude it from this analysis.

Table 2.5: Errors (%) for the three D^0 modes. Errors to the left of the triple line are for CLEO III, those to the right for CLEO II. The X's indicate the errors that are correlated among all the modes.

Mode	Stat	Track	π^0	Stat	Track	dE/dx	Cuts	Bkg	π^0
$K\pi$	5.3	3	0	3.7	3	1.1	2.5	2	0
$K\pi\pi^0$	5.4	3	3	4.7	3	2.1	2.5	2.5	2
$K3\pi$	8.2	5	0	5.3	5	5.2	2.5	3	0
		X	X		X	X			X

Lineshape Method

Another possibility for estimating the number of $\Upsilon(4S)$ in the data set is to predict it given the luminosity accepted for each run in the dataset, and given the line shape as a function of beam energy.

The $\Upsilon(4S)$ line shape should experimentally be described by a convolution of the inherent physics shape of the $\Upsilon(4S)$ and the beam energy envelope of CESR, including radiation effects. The beam energy envelope of CESR is thought to be about 2.5 MeV, much narrower than the reported width of the $\Upsilon(4S)$ in the PDG. However, radiation effects will significantly distort the line shape, causing a drop in effective cross section at low reported CESR beam energies, and an increase at energies above the $\Upsilon(4S)$ giving the final lineshape a manifestly asymmetric shape. In addition, although one naively might guess that a Breit-Wigner could describe the inherent shape of the $\Upsilon(4S)$, there is no reason that this has to be the case -

at the very least threshold effects will distort the low end of the shape.

We attempted to use the line shape estimated from CLEO II.V energy scans (see Figure 2.14) to estimate the number of $B\overline{B}$, letting the beam energy offset float to allow for any offset in energy program calibration. Apart from this possible calibration offset, the CLEO III line shape is expected to be similar except for a small difference in beam spread. In this calculation, we correct the beam energies for each dataset to account for the M_B offset, as determined in Section 3.4.2 by examining the B mass shift from dataset to dataset. The convoluted peak cross-section of the line shape is normalized to 1.07 nb for this calculation. The results are displayed in Table 2.7. Note the general good agreement between all three methods. Figure 2.15 illustrates where each dataset was taken relative to the $\Upsilon(4S)$ line shape.

We also thought about using the various cross sections reported directly in the Pass2 job to obtain the cross section as a function of beam energy. Unfortunately, the lack of a beam spot during Pass2 and the unknown dataset to dataset variation in beam conditions made this exercise too difficult to pursue.

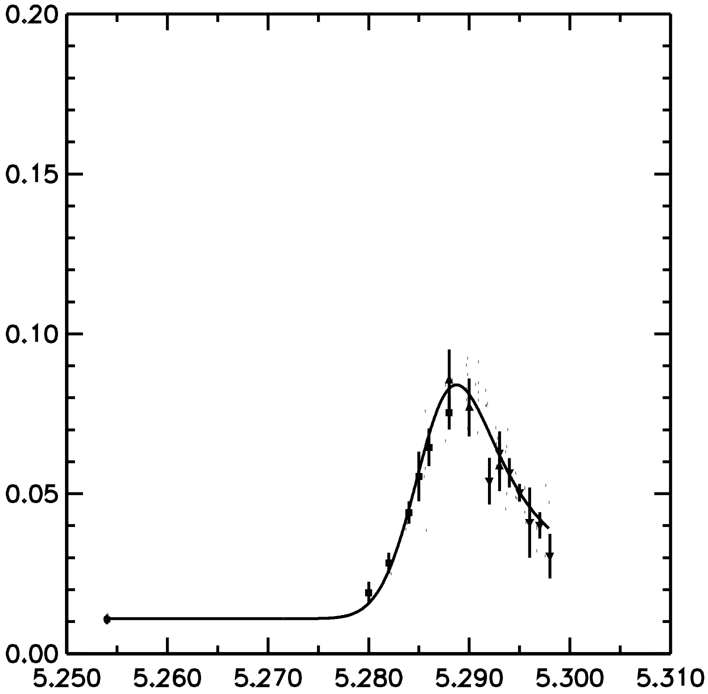

Figure 2.14: The CLEO II.V line shape, with a superimposed fit to a Breit-Wigner convoluted with a radiation kernel.

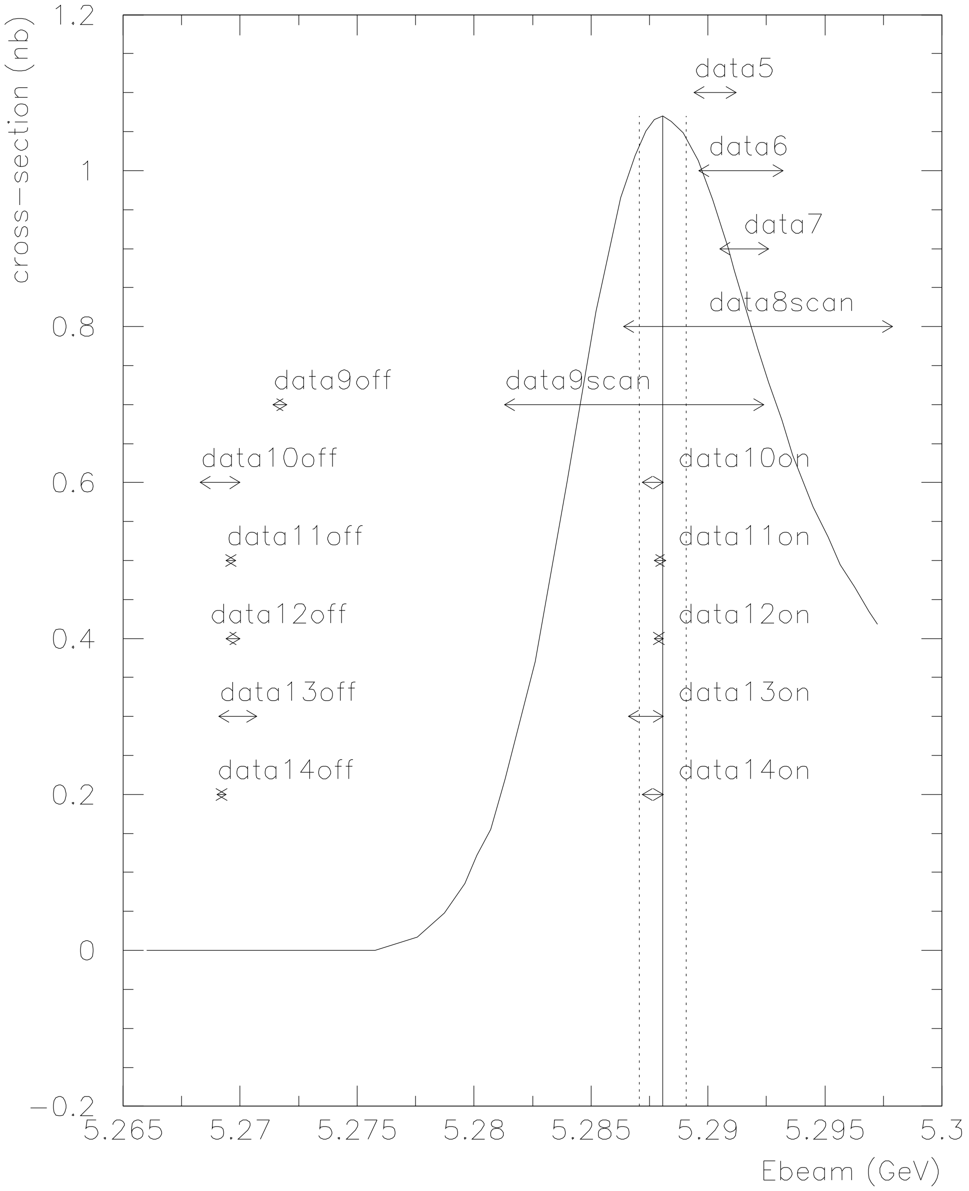

Figure 2.15: The CLEO II.V line shape and beam energies for different datasets. The solid vertical line indicates the peak of the cross-section, while the dotted lines are 1 MeV away.

Table 2.6: Comparison of $N_{B\bar{B}}$ (1000's) as determined by the R_2 method and the $D\pi$ method, by dataset. Errors for the $D\pi$ values are statistical only. Dashes indicate the datasets where the R_2 method is less reliable (see Tables 2.2 and 2.3 for a range of possible values).

Dataset	$B \to D\pi$	R_2
data6	631 ± 63	-
data7	598 ± 63	-
data8	748 ± 70	-
data9	606 ± 66	624
data10	558 ± 55	509
data11	607 ± 63	690
data12	448 ± 49	466
data13	791 ± 70	874
data14	744 ± 73	777
Sum(9-14)	3754 ± 155	3940

Table 2.7: Breakdown by dataset of $N_{B\bar{B}}$ (1000's) using three different methods and the associated χ^2. The two major columns labeled with δE_{Beam} show the number of $N_{B\bar{B}}$ as estimated from the luminosity and a line shape as a beam energy scale shift is applied. For other beam energy shifts, the χ^2 is worse. All errors shown are statistical only, and only the N($D\pi$) error is used in the χ^2 determination.

				$\delta E_{Beam} = 0.0$		$\delta E_{Beam} = 0.7$	
	N($D\pi$)	R_2	χ^2	N(Υ)	χ^2	N(Υ)	χ^2
5	222 ± 51	392	11.11	259	0.53	247	0.24
6	631 ± 63	618	0.04	512	3.51	470	6.45
7	598 ± 63	700	2.62	669	1.34	614	0.06
8	748 ± 70	750	0.00	757	0.02	698	0.51
9	606 ± 66	624	0.08	635	0.19	625	0.08
10	558 ± 55	509	0.79	481	1.96	501	1.07
11	607 ± 63	690	1.74	667	0.91	664	0.82
12	448 ± 49	466	0.13	494	0.88	481	0.45
13	791 ± 70	874	1.41	882	1.69	866	1.15
14	744 ± 73	777	0.20	825	1.23	810	0.82
All	5953 ± 240	6400	18.12	6183	12.26	5975	11.65

CHAPTER 3

DATA ANALYSIS

3.1 Background Suppression

The most important background for all of the charmless pseudoscalar-pseudoscalar B decays is $e^+e^- \to q\bar{q}$ events, with $q = u, d, s, c$. Since the mass of these quarks is significantly less than the energy of the beams, these quarks have relatively high momentum. The particles that form as the quarks hadronize tend to move along the direction of the initial light quark momentum. Hence, these events exhibit a jetty structure, with two groups of particles roughly back to back. A high momentum particle from each of these jets can combine to form a B candidate consistent with our M_B and ΔE requirements. We can take advantage of the differences between continuum decays and B decays to separate signal from background.

Cosine sphericity cut

The first cut made to suppress the $q\bar{q}$ background exploits the fact that the rest of the particles in the event tend to travel in the same direction as the candidate particles. The sphericity axis of the candidate is constructed, along with the sphericity axis of the rest of the tracks and unmatched showers in the event. The cosine of the angle between these two axes is required to be less than 0.8. Since B mesons are produced nearly at rest, the decay particles should be distributed uniformly in space, and the direction of the sphericity axis for each of the two B mesons should be independent. Therefore the cosine of the sphericity angle should be uniformly distributed. For continuum events, the sphericity axis will tend to be aligned with the sphericity axis of the rest of the particles in the event, and will therefore be strongly peaked near 1. Figure 3.1 shows these distributions for signal MC and off-resonance candidates.

Clearly, requiring that $|\cos(\theta_{sp})|$ be less than 0.8 is a powerful cut. Almost 92% of the background is eliminated, while nearly 80% of the signal passes the cut.

While the $|\cos(\theta_{sp})|$ cut is a very effective background suppression tool, further separation between signal and background is desirable. Once the $|\cos(\theta_{sp})|$ cut has been applied, no one variable provides sufficient discrimination between signal and background to justify applying further cuts. However, a number of variables have enough resolving power to be of use.

Cosine of B flight direction

We define $\cos(\theta_B)$ to be the angle between the B momentum vector and the beam direction. The B mesons produced at CESR come from a virtual photon at the $\Upsilon(4S)$ resonance, and since the B is a spin 0 particle, the distribution of $\cos(\theta_B)$ will have a $1-\cos^2\theta_B$ shape. The momentum vector of B candidates formed from $q\bar{q}$ background will be randomly distributed, and therefore the $\cos(\theta_B)$ variable for $q\bar{q}$ events should be flat. This behavior is shown in Figure 3.2.

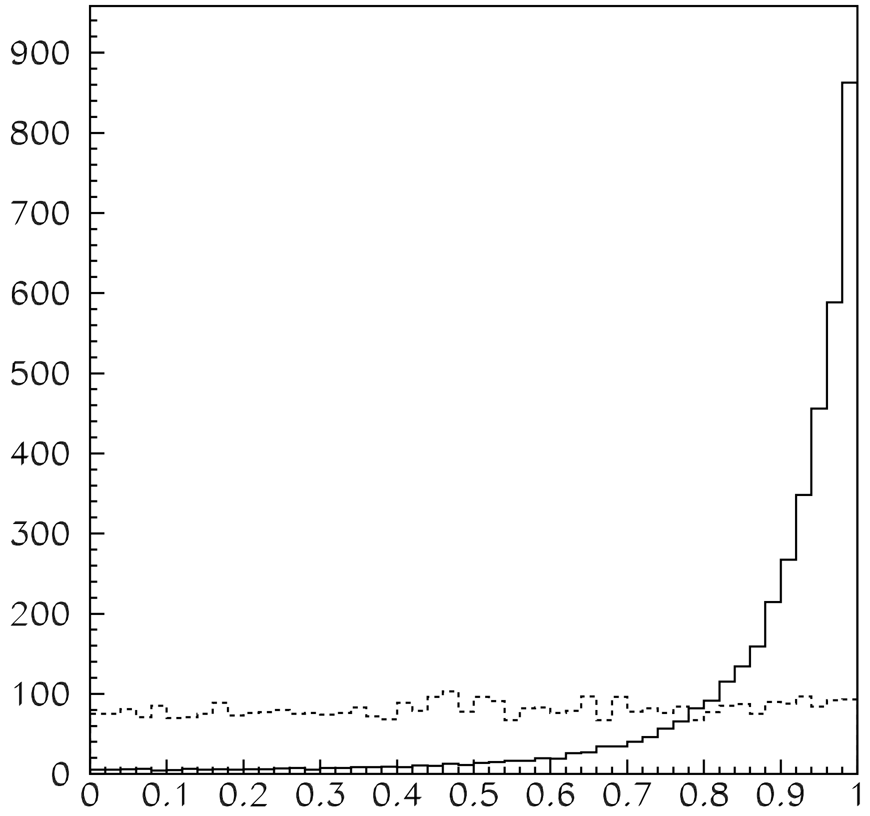

Figure 3.1: Distribution of $|\cos(\theta_{sp})|$ for signal MC (dashed) and off-resonance (solid) $B^0 \to h^+h^-$ candidates.

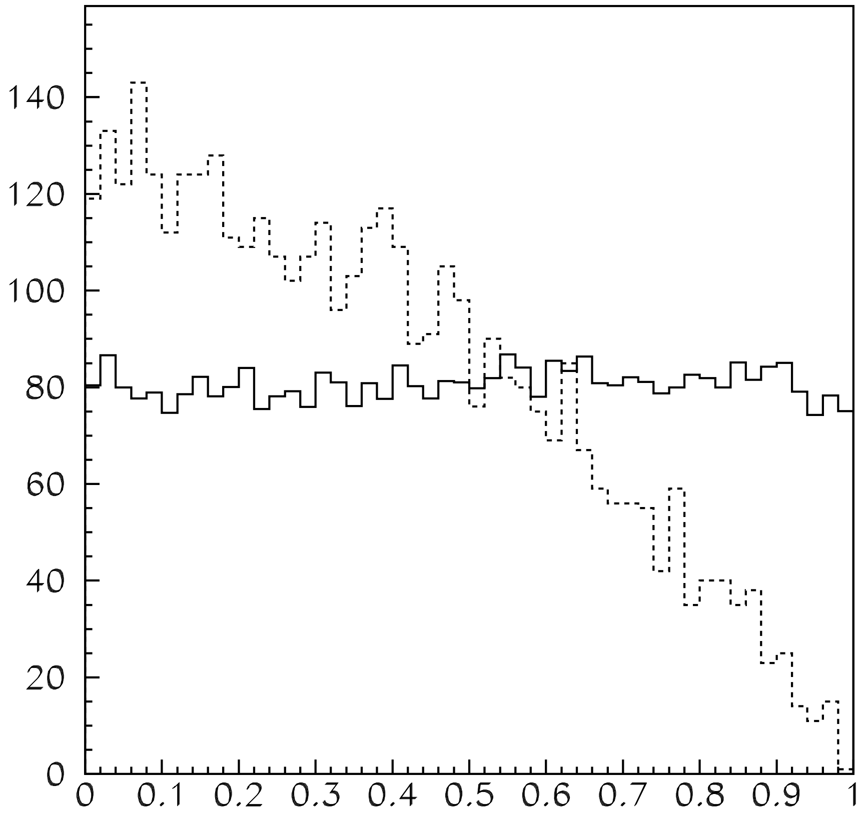

Figure 3.2: Distribution of $|\cos(\theta_B)|$ for signal MC (dashed) and off-resonance (solid) $B^0 \to h^+h^-$ candidates.

Cosine of B thrust axis

Another variable, $\cos(\theta_{thr})$, is defined as the cosine of the angle between the candidate thrust axis and the beam direction. The thrust axis will be along the direction of the two particles forming the B candidate. Since the B meson is a spin 0 particle with very little momentum, the thrust axis should be randomly distributed for true B decays. The $q\bar{q}$ background, however, will have its thrust axis aligned with the initial quark momenta. Because the quarks are spin-$\frac{1}{2}$ fermions, they will be produced with a $1 + \cos^2\theta_{thr}$ distribution. Unlike the $\cos(\theta_B)$ distributions, the $\cos(\theta_{thr})$ distribution for both signal and background will be cut off near ± 1, which is the area outside the detector volume. Figure 3.3 shows these distributions for signal and background.

Second Fox-Wolfram Moment

The second Fox-Wolfram moment (R_2) measures the shape of an event. Events with a uniform distribution of particles have R_2 values closer to 0, while events with a majority of particles along a given axis have R_2 values near 1. B decays tend to have their decay products more uniformly distributed than the jettier $q\bar{q}$ events, and thus the R_2 distribution for signal events will be lower than the distribution for background events. Figure 3.4 displays the R_2 shapes for signal and background.

Virtual Calorimeter

While the $|\cos(\theta_{sp})|$ and R_2 cuts use some event shape information to discriminate between signal and background, more separation can be achieved by looking at the

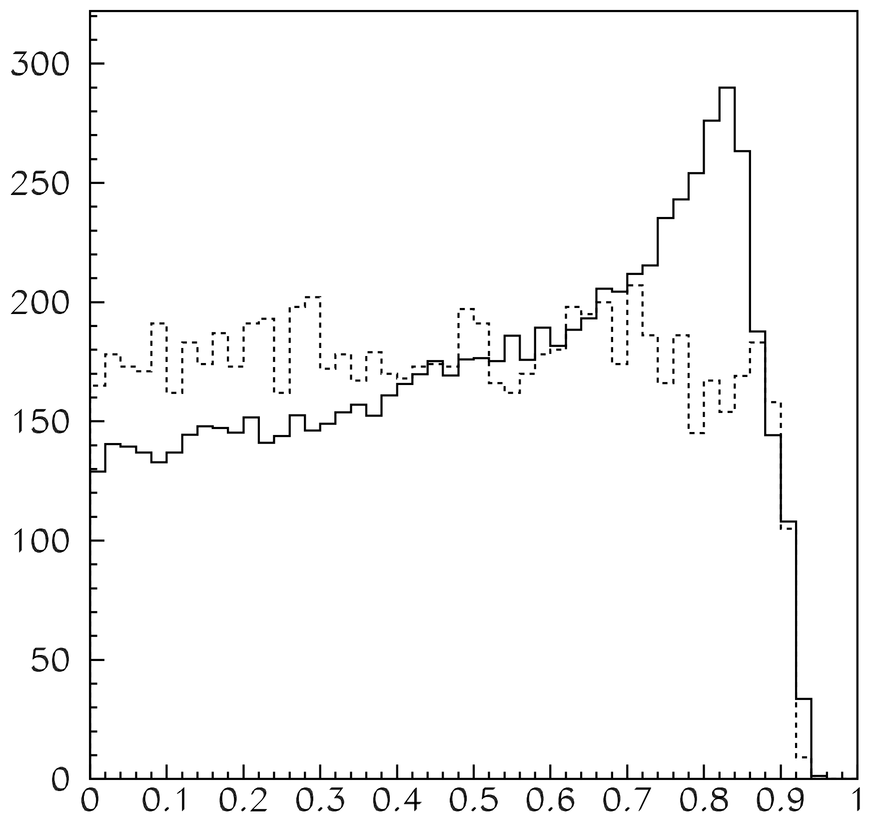

Figure 3.3: Distribution of $|\cos(\theta_{thr})|$ for signal Monte Carlo (dashed line) and off-resonance (solid line) $B^0 \to h^+h^-$ candidates.

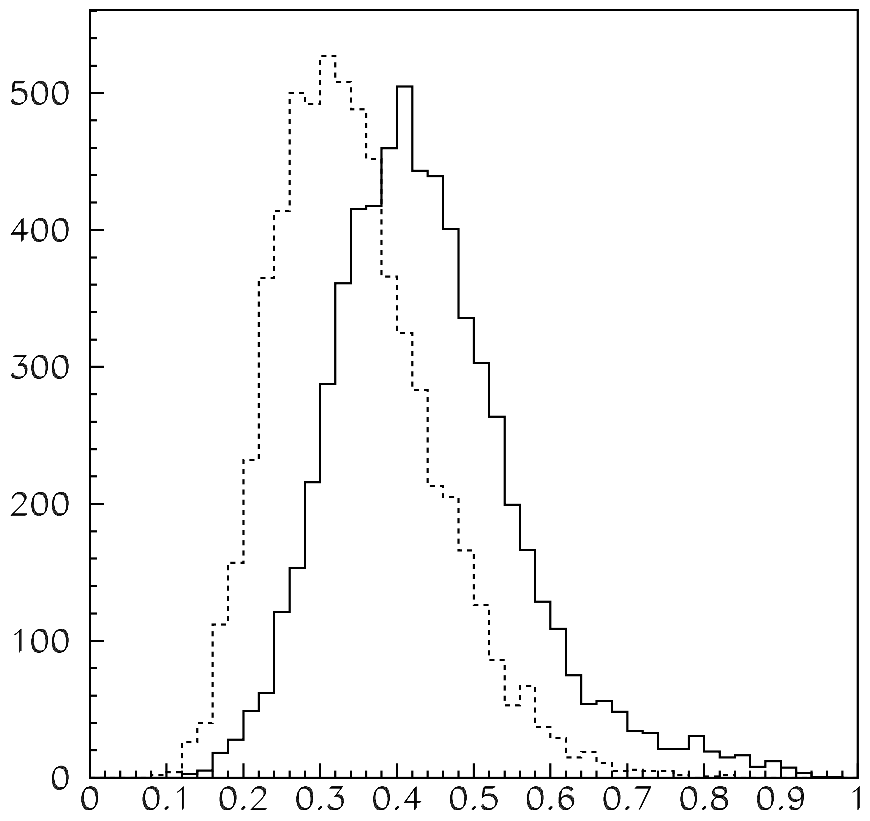

Figure 3.4: Distribution of R_2 for signal Monte Carlo (dashed line) and off-resonance (solid line) $B^0 \to h^+h^-$ candidates. All candidates are required to have $|\cos\theta_{sp}| < 0.8$.

detailed distribution of particles relative to the B candidate.

First, the thrust axis of the B candidate is calculated. This defines the axis of the "virtual calorimeter". The detector volume is divided up into cones 10° wide around the thrust axis. Every track and unmatched shower that is not part of the B candidate and that passes standard cuts is binned in the virtual calorimeter. The momentum of every particle is summed in each bin, and the forward and backward cones are added together. This gives nine variables in the virtual calorimeter. Background events typically have higher values for the bins near the thrust axis than signal events. This can be seen in Figure 3.5, which shows the distribution of the nine virtual calorimeter bins for signal and background.

Particle Identification

One more feature that we expect to distinguish the B signal from $q\bar{q}$ background is the presence of charged tracks other than $\pi^{\pm}$. We attempt to identify electrons, muons, kaons and protons in the tracks that do not make up our B candidate.

For signal events, two B mesons are produced in the decay of the $\Upsilon(4S)$. One B is our candidate, while the other B decays independently into some final state. The frequency of e, μ, K and p from this second B should be different than that found in the rest of the tracks from a $q\bar{q}$ event. For instance, leptons will be produced in B decays directly in semi-leptonic decays, and also indirectly in semi-leptonic D decays. However, $q\bar{q}$ events with $q = u, d, s$ are expected to produce relatively few leptons, and $c\bar{c}$ events will tend to only produce leptons from the semi-leptonic D decays. Therefore, we should find more leptons in signal events than background

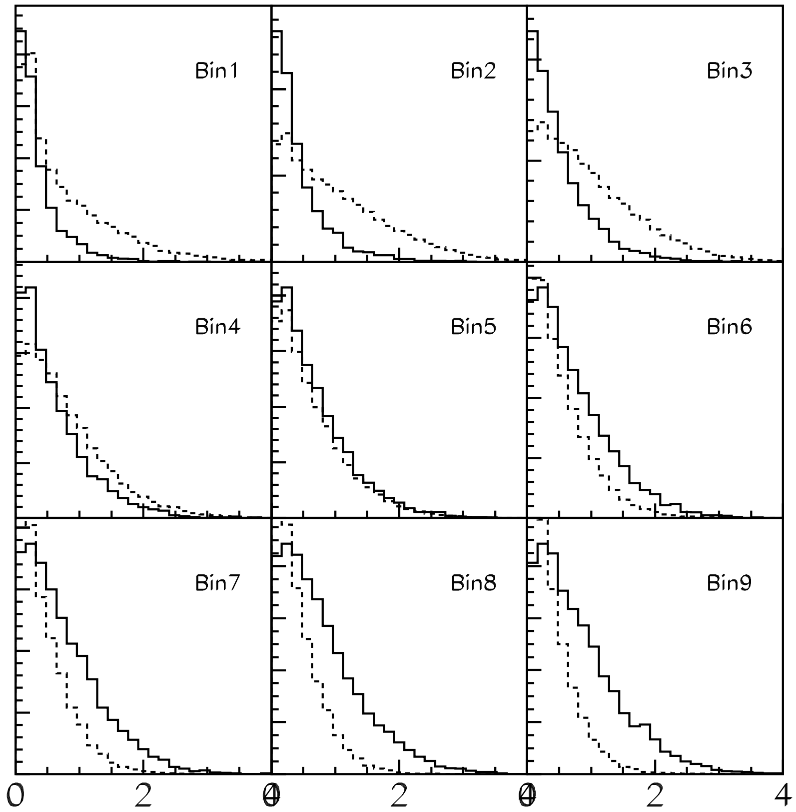

Figure 3.5: Distribution of the nine virtual calorimeter bins for signal Monte Carlo (solid) and off-resonance (dashed) $B^0 \to h^+h^-$ candidates. Virtual calorimeter bins with a value of zero have been omitted.

events. Conversely, B decays to baryons are less common, while hadronization in $q\bar{q}$ events produces a number of baryons that can eventually decay to protons.

To assign a hypothesis to a particular track, we use the following algorithm. First, we see if the track satisfies electron identification criteria. The energy of the matched showers are summed, and the ratio of this energy and the momentum of the track is calculated. This ratio is required to be between 0.85 and 1.1. Also, the sigma value from dE/dx for the electron hypothesis must be greater than -2. All tracks that pass these two requirements are labeled electrons.

Muon identification is done by requiring that the track have momentum of at least 1 GeV/c, a matched shower energy between 0.1 and 0.6 GeV, and a sufficient value of the variable depthmu. For tracks below 2.0 GeV/c, depthmu must be greater than 3.0, for higher momentum tracks, depthmu must be greater than 5.0. Any tracks not identified as electrons that pass these criteria are called muons.

If a track has not yet been identified, we turn next to dE/dx. Tracks with momentum less than 1 GeV/c that are within three sigma of the proton hypothesis are called protons, and tracks that are less than 750 MeV/c within three sigma of the kaon or pion hypothesis are identified as kaons or pions, respectively. Any remaining tracks that have a greater likelihood of being a kaon than a pion, as determined by the RICH, and that have at least four photons associated with the kaon hypothesis in the RICH, are labeled kaons. All remaining unidentified tracks are called pions.

For every track that is not part of the B candidate, we apply this algorithm. We define four variables, one for each of the non-pion hypotheses, that is the momentum of the highest momentum track of a given species. If no track of a particular species is found, the variable for that track is set to 0.

The distributions for these four variables are shown in Figure 3.6. Entries of 0 (those events that do not have any track identified under the given hypothesis) have been left out of the histograms. The kaon distribution provides little information to separate signal from background, but the other three particle species are more useful. Approximately one fifth of all events contain a track identified as a charged lepton or proton.

3.1.1 Fisher Discriminant

With all of these assorted variables that provide limited discriminating power, we need to decide how to most effectively use this information. To do this, we use the Fisher discriminant [14]. Given a set of N variables X_i, the Fisher discriminant ($\mathcal{F}$) is defined as

$$\mathcal{F} = \sum_{i=1}^{N} \alpha_i X_i, \tag{3.1}$$

where the α_i are coefficients calculated to maximize separation between the signal and background samples. These coefficients are calculated using

$$\alpha_i = \sum_{j=1}^{N} (C^S + C^B)^{-1}_{ij} \times (\mu_i^B - \mu_i^S), \tag{3.2}$$

where C is the covariance matrix, and μ is the mean of the X_i for signal and background. One advantage of using the Fisher discriminant is that it takes

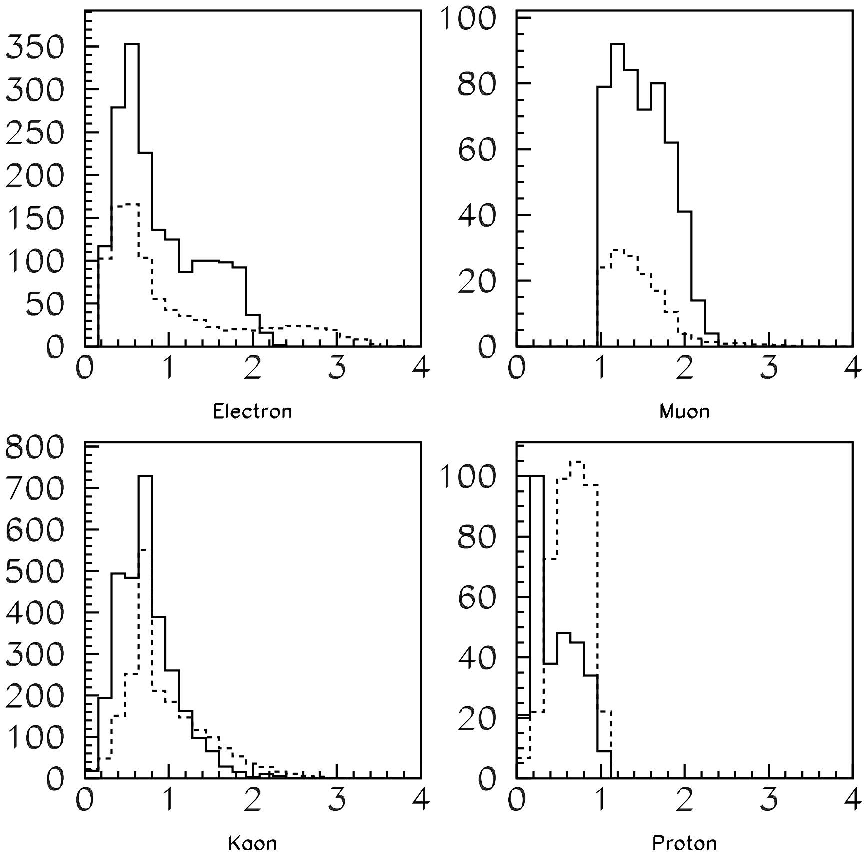

Figure 3.6: Distribution of the electron, muon, kaon and proton momentum bins for signal Monte Carlo (dashed) and off-resonance (solid) $B^0 \rightarrow h^+h^-$ candidates.

advantage of correlations among the X_i.

The initial analysis of $B^0 \to h^+h^-$ in CLEO III used the α_i calculated for CLEO II. Of particular interest was the fact that the coefficient for R_2 was negative. By construction, signal events should have lower values of $\mathcal{F}$ than background events. Since the R_2 values of signal events are generally lower than those for background, we would expect this coefficient to be positive. This, among other things, provided the motivation to study the Fisher discriminant anew for CLEO III.

There were several goals for a new Fisher study. One objective was to recalculate the coefficients to maximize the separation between signal and background. Another goal was to see if the old variables could be improved or if new variables could be added to the Fisher that would significantly improve separation. The third aim was to verify that the coefficients used for $B^0 \to h^+h^-$ could be used for other $B \to P_1P_2$ modes, where P_i represents any charmless pseudoscalar meson. The sample of background events used to calculate the Fisher coefficients and the shapes of the resulting Fisher distributions come from off resonance data, and from events in the M_B vs. ΔE grand sideband. For several of the of the PP modes, particularly those with two neutral mesons, there are very few events available with which to do the necessary studies. It would therefore be very beneficial for these analyses to use a Fisher that could be shown to be reasonably consistent for all ten PP modes so that the relatively high statistics $B^0 \to h^+h^-$ background sample could be used to determine the coefficients and shapes for all modes. With these three goals in mind, a number of studies were done.

We tried to use a virtual calorimeter that used the energy flow into each bin instead of the momentum flow into each bin. This strategy depends on accurately determining the particle species for each track to determine its energy. Unfortunately, the highest momentum tracks are the ones that drive the value of $\mathcal{F}$, but the energy of these tracks are the least affected by the mass hypothesis assigned to them. No increase in separation was found for this strategy.

Another approach was to calculate the missing momentum of an event. In this method, we calculate the vector sum of the momenta of all the tracks and unmatched showers. The missing momentum in the event is then calculated, and this vector is added to the virtual calorimeter like any other track or unmatched shower. If any particle has not been reconstructed, we hope to add it back in at this stage to get a more accurate picture for the virtual calorimeter. However, the combination of inefficiencies in the detector, less than complete solid angle coverage, and undetectable particles, such as neutrinos and K_L^0's, is big enough that we often have multiple undetected particles, and the missing momentum calculated is not representative of where these particles went. Therefore this method also showed no improvement in the separation.

To verify that the coefficients given by eq. 3.2 do provide the maximum separation, we calculated the separation for the initial α_i, then varied each coefficient one at a time. At each point, we recalculated the separation, then plotted this separation as a function of the coefficient. For all the variables, the separation did indeed peak at the value given by the Fisher prescription.

Another issue is to decide what variables should be in the Fisher discriminant and what variables should be used in the maximum likelihood fit. Variables that have significant correlations with other Fisher inputs should remain in the Fisher discriminant to take advantage of the fact that correlations are taken into account when calculating the α_i. We also need to make sure that no two variables in our maximum likelihood fit are correlated, or else we will get biased results. Furthermore, any variable in the maximum likelihood fit needs a functional parameterization for both signal and background, therefore any input that cannot be easily fit is best left in the Fisher discriminant.

Of the possible input variables, only $\cos(\theta_B)$ can be fit with a relatively simple function for both signal and background. To show that it can be removed from the Fisher discriminant and put in the fit, we should look at its correlations to the other inputs for signal and background, and verify that they are small. Table 3.1 shows the correlation matrix elements for $\cos(\theta_B)$ for signal and background. With the exception of the background $\cos(\theta_{thr})$ element, all values are at the percent level. These correlations are so small that there is no loss in separation if $\cos(\theta_B)$ is removed from the Fisher discriminant and put in the maximum likelihood fit.

To determine what variables could be used in the Fisher discriminant while requiring that the resulting shapes for signal and background be mode independent, we calculated the α_i for the eight combinations of $\cos(\theta_B)$, $\cos(\theta_{thr})$, and R_2 using $B^0 \to h^+h^-$. We used these eight sets of coefficients to make the Fisher distributions for $B^0 \to h^+h^-$ and $B^+ \to K^+\pi^0$ signal and background samples and fit them. We then calculated the signal and background efficiency for $B^+ \to K^+\pi^0$ as a function of $\mathcal{F}$, and plotted background efficiency as a function of signal efficiency. Next, we repeated the background efficiency versus signal efficiency pro-

Table 3.1: Correlation matrix coefficients for $\cos(\theta_B)$ in signal and background.

Variable	Signal	Background
R_2	0	0.03
$\cos(\theta_{thr})$	0.01	0.08
VC1	0	0
VC2	0	0
VC3	-0.01	-0.01
VC4	0	0
VC5	0	0
VC6	0.02	0
VC7	-0.02	-0.01
VC8	0	0
VC9	0	-0.02
Elec	0	0
Muon	0	0
Kaon	0.01	0
Proton	-0.01	-0.02

cedure, only this time we used the $B^0 \to h^+h^-$ background shape instead of the $B^+ \to K^+\pi^0$ background shape. We compared the two background versus signal efficiency curves. If the two curves are the same, that indicates that the Fisher discriminant is insensitive to the mode under consideration. Alternatively, one can choose a signal efficiency and consider the difference in the background efficiency to be a measure of the systematic error from using the same Fisher shape parameterization for all $B \to P_1P_2$ modes. Since we have already established that $\cos(\theta_B)$ can be removed from the Fisher discriminant, only the four pairs of efficiency curves that do not include $\cos(\theta_B)$ are shown in Figure 3.7. Clearly R_2 introduces a systematic shift between the modes, whereas $\cos(\theta_{thr})$ does not. For instance, at a signal efficiency of 75%, the difference in background efficiency is less than 1%.

Having studied the possible inputs to the Fisher discriminant in detail, we conclude that the nine virtual calorimeter bins, the four particle identification momentum variables, and $\cos(\theta_{thr})$ should be used to calculate $\mathcal{F}$, $\cos(\theta_B)$ should be a component of the maximum likelihood fit, and R_2 should be omitted entirely from the analysis. The α_i for these fourteen variables are listed in Table 3.2.

3.2 Rare B Standards

All tracks and showers used in these analyses were required to pass loose selection cuts.

Tracks:

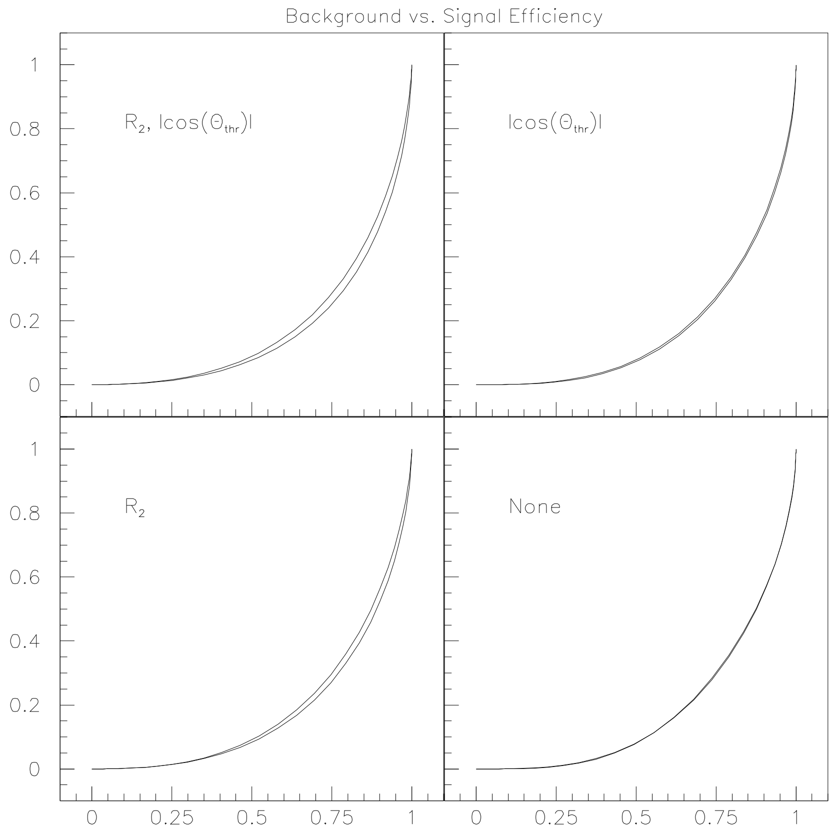

Figure 3.7: Comparison of signal versus background efficiency curves using $B^0 \to h^+h^-$ and $B^+ \to K^+\pi^0$ background parameterizations. For each plot, $\mathcal{F}$ includes the virtual calorimeter, particle identification variables plus the variables indicated on the plot.

Table 3.2: Coefficients used to calculate Fisher discriminant.

Variable	Coefficient
$\cos(\theta_{thr})$	0.631
VC1	0.323
VC2	0.509
VC3	0.651
VC4	0.460
VC5	0.262
VC6	0.022
VC7	-0.242
VC8	-0.364
VC9	-0.470
Electron	-0.496
Muon	-0.487
Kaon	-0.204
Proton	1.189

- $|d_0| < 5$mm
- $|z_0| < 5$cm
- $0.5 <$hit fraction< 1.2
- $\chi^2/n < 20$
- $p < 5.3$ GeV/c

Both impact parameters are measured with respect to (0,0,0).

Showers:

- Unmatched
- 50 MeV $< E_\gamma <$ 5 GeV

π^0s:

- $E_\gamma >$ 100MeV

In addition, we inherit the implicit cuts of the π^0 finder, which excludes bad endcap photons and transition region photons (rings 1, 2, 9-13 and the corresponding rings on the other side). Both photons must also pass E9/E25 shape cuts.

3.3 Event Selection

All of the modes use the same basic method to select events. For a candidate, we construct the beam constrained B mass $M_B = \sqrt{(E_b^2 - p_B^2)}$ where E_b is the beam energy, and p_B is the momentum of the candidate, and $\Delta E = E_{cand} -$

E_b. Any candidate with $|\Delta E| < 400$ MeV and $M_B > 5.2$ GeV/c^2 is kept. The largest background in all of these modes is from continuum $q\bar{q}$ processes. A cut on $\cos\theta_{sp}$, the cosine of the angle between the sphericity axis of the candidate and the sphericity axis of the rest of the event, is a powerful tool in rejecting background. All modes use a cut of $|\cos\theta_{sp}| < 0.8$.

For all candidates, we calculate $\cos\theta_B$, the cosine of the angle between the B candidate momentum direction and the z-axis, and a Fisher discriminant ($\mathcal{F}$) that is described in detail in Section 3.1.1. The four variables M_B, ΔE, $\cos\theta_B$, and $\mathcal{F}$ are used as inputs for a maximum likelihood fitter.

3.3.1 Particle Identification

One of our objectives for this analysis was to combine the particle identification information from dE/dx and the RICH detector. We tried two methods to accomplish this goal. One approach was a combined likelihood method, and the other was a neural net. Both methods used a sample of $D^{*+} \rightarrow D^0\pi^+$ candidates to kinematically identify kaons and pions. Efficiencies and fake rates could be measured by fitting the D^0 mass peak before and after particle identification cuts and taking the ratio of the yields.

The first approach was to form a combined likelihood difference. The RICH analysis code gives the user the $-2\ln\mathcal{L}$ value for each of the five mass hypotheses. The quantity of interest to the user is then the difference between $-2\ln\mathcal{L}$ for two different mass hypotheses. If we define l_X as $-2\ln\mathcal{L}$ for mass hypothesis X, kaons could be separated from pions by requiring $l_K - l_\pi < 0$ (i.e. - the track is more likely to be a kaon than a pion). This value is essentially a χ^2 difference. The equivalent quantity for dE/dx would be $\sigma_K^2 - \sigma_\pi^2$. This assumes that the dE/dx gives unit gaussians for the momentum range of interest. If the gaussians have a non-zero mean, or if the width of the gaussians are not one, the reported σ's would need to be corrected.

Since most of the separation between kaons and pions at 2.5 GeV comes from the RICH, it is important that it provide us with reliable information. Therefore, we require that the RICH uses at least four photons to calculate l_X (i.e. - at least four photons used to calculate l_K when we identify kaons, or at least four photons used to calculate l_π when we identify pions). Using our sample of D^{*+} candidates, we first measured the mean and width of the dE/dx distribution for kaons and pions. We found that the means for both hypotheses were shifted by 0.14σ. The width of the kaon peak was 0.800, and the width of the pion peak was 0.854. To identify kaons, we require that

$$l_K - l_\pi + \left(\frac{\sigma_K - 0.14}{0.800}\right)^2 - \left(\frac{\sigma_\pi - 0.14}{0.854}\right)^2 < \delta_K, \tag{3.3}$$

where δ_K is a cut value that can be chosen to give the optimum efficiency and fake rate. Likewise, to identify pions we require that

$$l_\pi - l_K + \left(\frac{\sigma_\pi - 0.14}{0.854}\right)^2 - \left(\frac{\sigma_K - 0.14}{0.800}\right)^2 < \delta_\pi, \tag{3.4}$$

where δ_π is the analogous pion identification cut. All analyses that need particle identification use $\delta_K = 2.0$, and $\delta_\pi = -0.5$. These cut values correspond to 90% efficiency for both kaon and pion identification, with fake rates less than

10%. Statistical errors on our D^* sample lead to a systematic error on particle ID efficiency of 3%.

The neural net used the likelihood values from the RICH, the number of photons in the RICH, dE/dx and momentum as inputs. We chose our cuts to give 90% efficiency for both methods, and over the narrow range of momentum of interest to the rare B group, the two methods were found to have similar fake rates, so we chose the simpler combined likelihood method for PID.

3.4 Maximum Likelihood Fit

In order to separate our signal events from the substantial $q\bar{q}$ background efficiently, we use a maximum likelihood fit. We will briefly discuss the principles of maximum likelihood fits, and then discuss the application to rare B decays.

3.4.1 Fundamentals of Maximum Likelihood Fits

To determine the number of signal events present in our data samples, we perform an extended maximum likelihood fit (EML). The EML method differs from the standard maximum likelihood (ML) method in that it allows for Poisson fluctuations in the number of events. Details on ML and EML fits can be found in [17], [18]. The idea is to parameterize the likelihood as a function of a set of parameters, which in our case is a signal yield and one or more background yields, and maximize it. In practice, it is much easier to minimize the quantity $-2\ln\mathcal{L}$.

For this analysis we have four quantities that distinguish between the signal and background components. For each variable of every component, we determine a probability density function (PDF) $P_{ij}(x_i)$, where i refers to the variable, j refers to the component of the fit (signal or background), and x_i represents the vector of the four quantities that we calculate for every event. The likelihood that a particular event is signal or background is then

$$P_j = \prod_i P_{ij}(x_i). \tag{3.5}$$

We assume that there are S signal events and B background events, in which case the total PDF becomes $SP_S + BP_B$. The case when there are additional background components generalizes in a straight-forward manner. The likelihood for a sample of N data points is then

$$\mathcal{L} = e^{-(S+B)} \prod_{i=1}^{N} [SP_S + BP_B](\vec{x}) \tag{3.6}$$

It is computationally much simpler to minimize $-2\ln\mathcal{L}$, which is

$$-2\ln\mathcal{L} = -2\left(S + B + \sum_{i=1}^{N} \ln([SP_S + BP_B](\vec{x}))\right) \tag{3.7}$$

Another advantage of the $-2\ln\mathcal{L}$ variable is that is a χ^2-like quantity, which means that we can easily calculate one σ errors and significances.

Our likelihood fits are performed using the FELIX package, which uses MINUIT to do the minimizations.

3.4.2 Common PDF's

Several of the PDF's are common to all modes and we describe them here.

Cosine of B Thrust Angle

Our B mesons come from the $\Upsilon(4S)$ resonance, and are therefore produced with a $1 - \cos^2\theta$ distribution. The shape of the $\cos\theta_B$ PDF for our signal is then simply $1-\cos^2\theta$. For the continuum background, we are choosing two essentially random tracks to form a B candidate, and the vector sum of their momenta should be randomly distributed in space. Therefore we should have a flat $\cos\theta_B$ distribution in our background sample. Both of these observations are confirmed by looking at signal MC and off resonance data, respectively.

Background $\mathcal{F}$

In previous rare B analyses, modes that have low statistics in the off resonance data have had large systematic errors due to the inability to accurately determine the background Fisher shape. To address this issue, we wanted to use a Fisher discriminant that had the same shape for all modes. The relatively high statistics h^+h^- and $h^+\pi^0$ samples were used to verify that the off resonance Fisher shapes were consistent for different modes. The R_2 variable was removed from the Fisher to make the distributions mode-independent. The Fisher shape is then determined for all modes by combining the Fisher values for all of the rare B topologies and doing one fit. Care was taken to make sure that the tails of the Fisher distribution on the signal side were well modeled. We reconstructed B candidates to a three charged track final state and examined both off-resonance data and the grand side-band region of the on-resonance data, avoiding the substantial B signal from, for instance, $B^- \to D^0\pi^-$, $D^0 \to K^-\pi^+$, among others. We found that the tail of the

Fisher distribution for the three track sample, which had approximately 10 times more candidates than the two track sample, was well modeled by a Breit-Wigner distribution. We then assume that the Breit-Wigner adequately describes the tail of the Fisher distribution for the two body decay modes. The final background Fisher shape was the sum of a bifurcated gaussian and a Breit-Wigner forced to have the same mean, plus another gaussian. The values of the parameters and the errors used for PDF variations are given in Table 3.3. The distribution of $\mathcal{F}$ and the fit results are shown in Figure 3.8. Since we use off-resonance track-track candidates to calculate our Fisher coefficients, and then also use track-track off-resonance candidates (plus the other five rare B topologies) to calculate the PDF's, there is in principle the potential for a bias. However, since the systematic error from the variation of the mean of the background Fisher is much smaller than the largest PDF variations, any bias is entirely negligible (see Tables 3.13 and 3.14).

Signal M_B

To obtain a value for the mean of our signal M_B distribution, we used a sample of $B^- \to D^0\pi^-$ events, with the D^0 reconstructed in three submodes: $K^-\pi^+$, $K^-\pi^+\pi^0$, and $K^-\pi^+\pi^-\pi^+$. We cut on the D^0 mass, RICH likelihood of the kaon, ΔE, $|\cos\theta_{sp}|$, and on the π^0 mass. Adding the M_B histograms from the three D^0 submodes for each dataset, we find a significant drift in the fitted M_B peak, indicating a change in the difference between the actual beam energy and the value reported by the energy program. Figure 3.9 shows the mean of the M_{B^+} peak as a function of dataset. We apply a correction to the beam-constrained B candidate

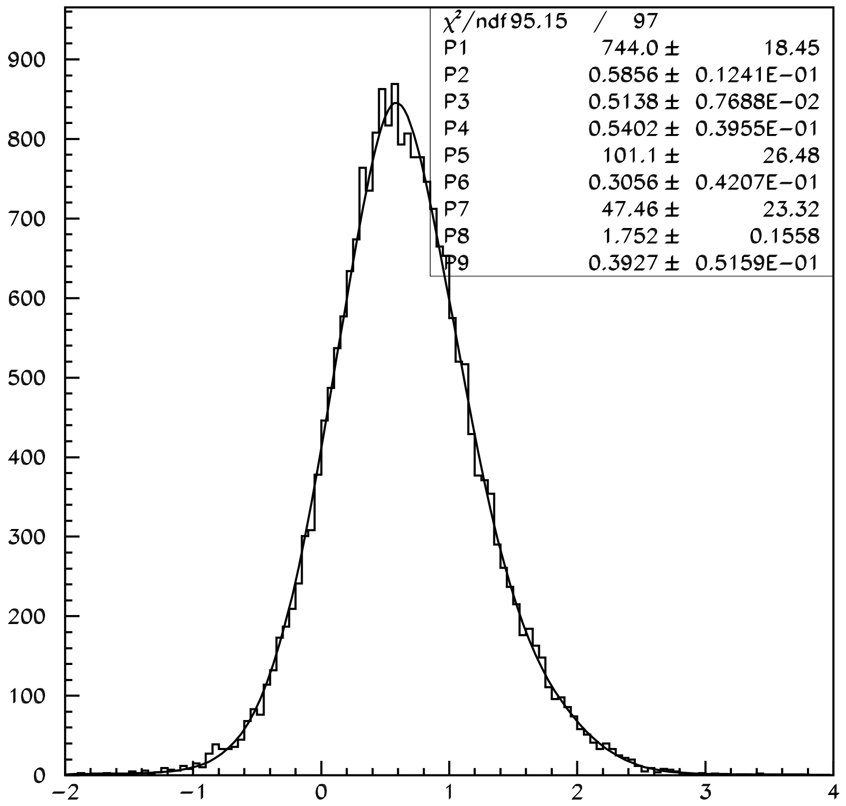

Figure 3.8: Combined background Fisher shape with fit and errors. There is a factor of two difference in the definition of the width of the Breit-Wigner used in the fit and that used by FELIX, hence the $0.305 \rightarrow 0.611$.

Table 3.3: Background Fisher parameters and their errors.

μ	0.5856 ± 0.0124
σ_L	0.5138 ± 0.0077
σ_R	0.5402 ± 0.0396
σ_{BW}	0.6112 ± 0.0841
μ_G	1.752 ± 0.1558
σ_G	0.3927 ± 0.0516
Area_{BW}	0.086 ± 0.023
Area_G	0.042 ± 0.021
Area_{BG}	$1-\text{Area}_{BW}-\text{Area}_G$

mass for each dataset so that M_{B^+} is at its nominal PDG value of 5279.0 MeV. The error on the mean from our fit is 0.1 MeV(Figure 3.10). We use the PDG value of $M_{B^0} - M_{B^+} = 0.33 \pm 0.28$ MeV to fix the mean of the M_{B^0} distribution at 5279.3±0.3 MeV.

Background M_B

In general, we model our continuum M_B shape with an ARGUS [19] function. Two parameters describe an ARGUS shape, Efact and the endpoint value E_b. Efact is measured for each mode by using off resonance data scaled to a common beam energy. E_b is the endpoint of the ARGUS distribution, and is determined by the beam energy for a given run. If all data were taken at one beam energy, one value

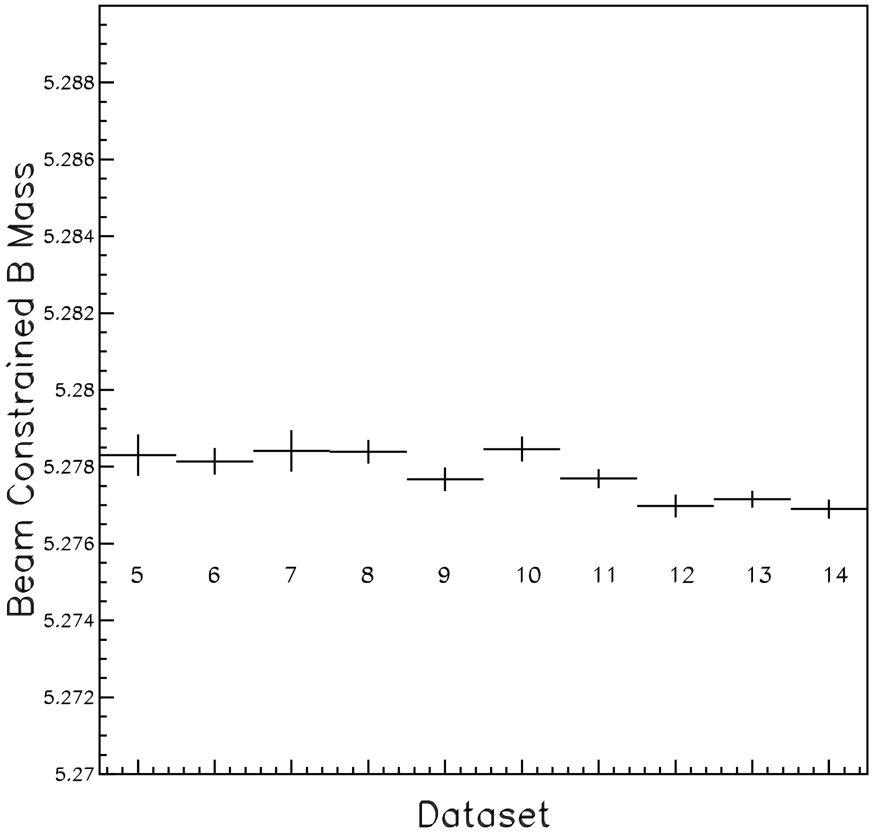

Figure 3.9: Beam constrained B mass as a function of dataset using $B^- \to D^0\pi^-$ decays, with the D^0 reconstructed in $K^-\pi^+$, $K^-\pi^+\pi^0$ and $K^-\pi^+\pi^-\pi^+$.

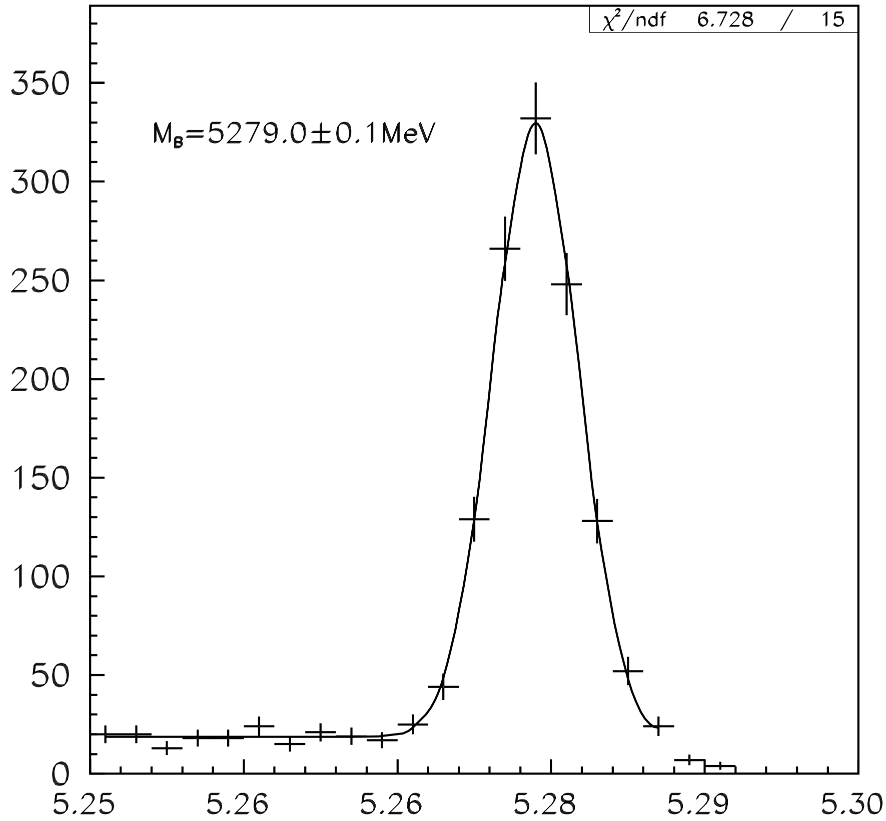

Figure 3.10: Beam constrained B mass using $B^- \to D^0\pi^-$ decays, with the D^0 reconstructed in $K^-\pi^+$, $K^-\pi^+\pi^0$ and $K^-\pi^+\pi^-\pi^+$, after applying dataset-dependent beam energy shifts.

of E_b would be sufficient. However, the CLEO III data set consists of data taken around several values of the beam energy. Moreover, there is a spread in reported beam energy around each of these three points (this spread is not to be confused with the event to event beam energy spread). Finally, we see clear evidence of a change in the difference between the value reported by the beam energy program and the actual beam energy. This forces us to correct all M_B values, which will require further corrections to the endpoint region of the continuum M_B shape.

To model these features, we looked at the number of candidate events found in the h^+h^- mode at each of the beam energies. We then measured the spread in beam energy about each of these points after applying our dataset-dependent energy shifts. For each of the ranges of beam energies, we construct an ARGUS shape convolved with a gaussian of width equal to the corresponding spread. Finally, these convoluted ARGUS shapes are added together, weighted by the relative number of events at each beam energy value. Since we fit old pass2 and new pass2 separately, this technique was performed separately for both versions of pass2. Every mode has its own value of Efact, determined from the off resonance data, which is inserted into this formulation for M_B. The distributions of beam energies and fits to determine the spread are shown for old and new pass2 in Figure 3.11 and Figure 3.12. To illustrate the result of this prescription, a plot of the endpoint region for the h^+h^- mode is shown in Figure 3.13. The values used for old pass2 and new pass2 are given in Table 3.4.

Another method for handling multiple beam energies would be to define a fifth parameter for the maximum likelihood fit, for example $x = M_B - E_b + E_0$. E_0 is

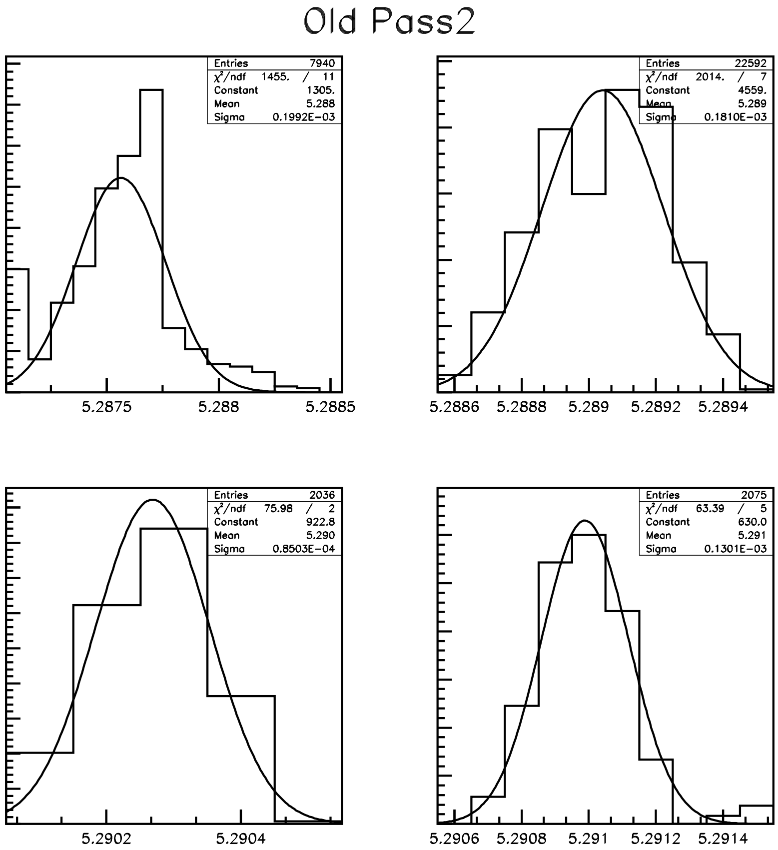

Figure 3.11: Fits to the beam energy distribution for the four energy ranges of old pass2.

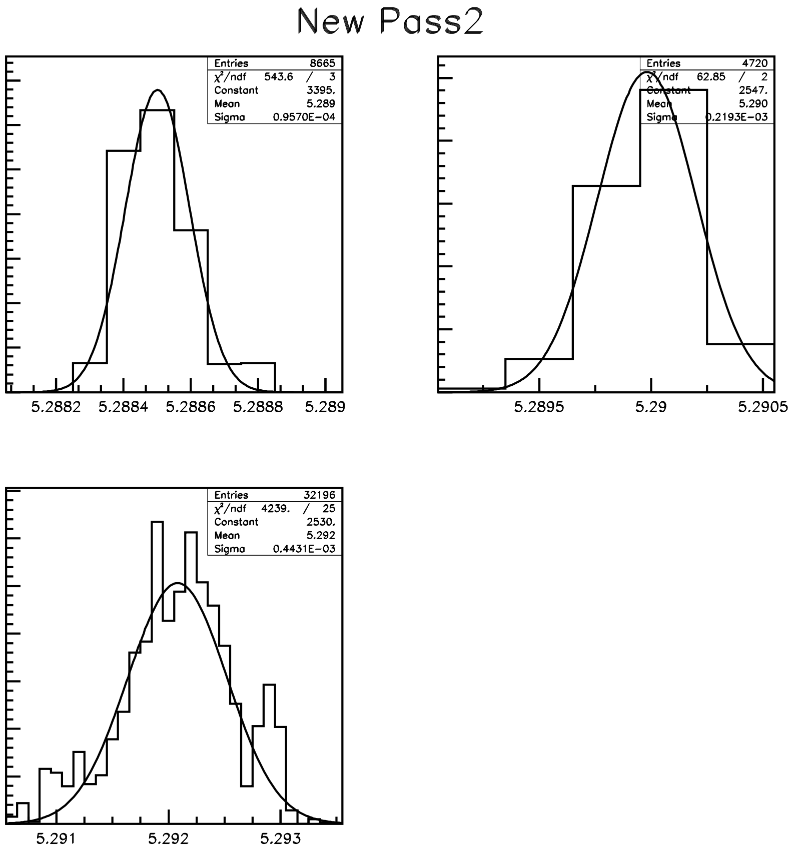

Figure 3.12: Fits to the beam energy distribution for the three energy ranges of new pass2.

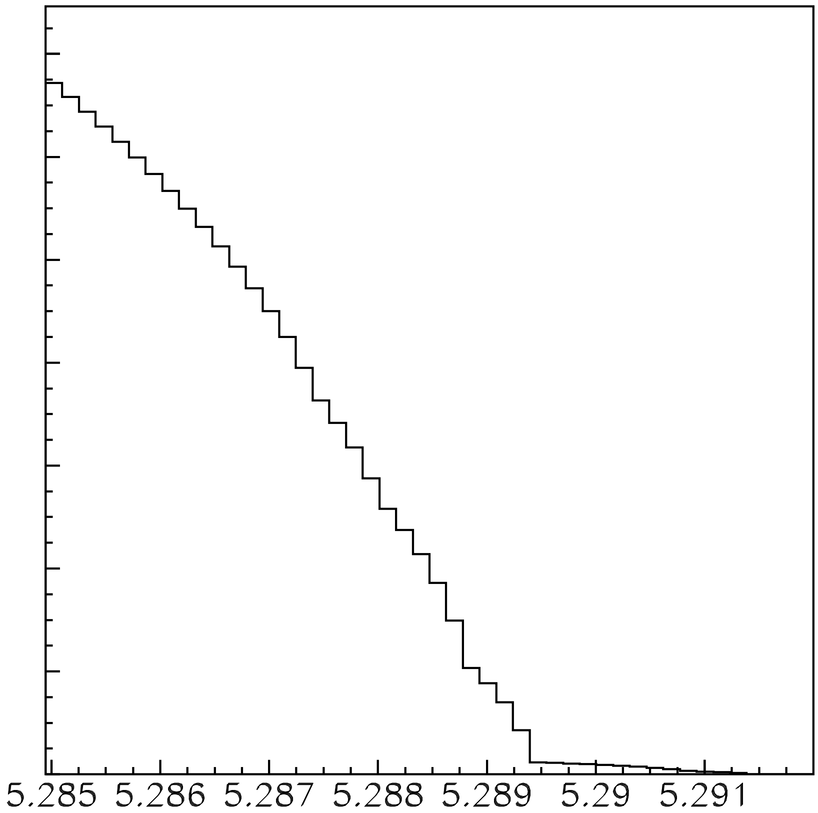

Figure 3.13: Endpoint region of the h^+h^- background M_B shape for old pass2, consisting of four convoluted ARGUS shapes.

Table 3.4: Endpoint M_B prescription for old pass2 and new pass2.

Pass2	E_b (GeV)	Spread (MeV)	Weight (%)
Old	5.2876	0.20	23
	5.2890	0.18	65
	5.2903	0.09	6
	5.2910	0.13	6
New	5.2885	0.10	20
	5.2900	0.22	10
	5.2921	0.44	70

a constant energy offset so that x behaves like the usual ARGUS distribution for background candidates with an endpoint near typical on-resonance beam energies. In this case, the PDF for signal M_B would remain unchanged, for instance a gaussian $G(M_B)$. The PDF for background x would be an ARGUS shape with a fixed endpoint at E_0, called $A(x)$. Both the signal x and background M_B PDF's would be flat. Since the PDF's for each component must be normalized to unit area, the functional form for the two constant PDF's will be $1/\Delta_i$, where Δ_i is the range between the lower and upper limits for M_B and x, respectively (eq. 3.8)

$$\mathcal{L}_1 = \prod [sG(M_B)\left(\frac{1}{\Delta_x}\right) + bA(x)\left(\frac{1}{\Delta_{M_B}}\right)]. \tag{3.8}$$

It is imperative that $\Delta_{M_B} = \Delta_x$. This makes the likelihood for signal x equal to the likelihood for background M_B, and therefore these constants can be factored out of the likelihood function and simply yield a trivial overall multiplicative constant (eq. 3.9).

$$\mathcal{L}_1 = \Delta^{-N} \prod [sG(M_B) + bA(x)] \tag{3.9}$$

Eq. 3.9 can be compared to eq. 3.10, which shows the likelihood function for a one dimensional fit to only M_B, where data was taken at two different beam energies.

$$\mathcal{L}_2 = \prod [sG(M_B) + bA(M_B)] = \prod [sG(M_B) + b\{A_1(M_B) + A_2(M_B)\}] \tag{3.10}$$

This method would not discard any information and would, in principle, be a more accurate representation of the data. However, we expect any effect to be small, and we have not gone back and refit all of the data.

Signal ΔE

To measure the mean of the ΔE distribution, we examine $B^- \to D^0\pi^-$, with the D decaying to $K^-\pi^+, K^-\pi^+\pi^0, K^-\pi^+\pi^-\pi^+$. Candidates must pass the following requirements:

- D^0 candidate mass within 15 MeV/c^2 of nominal D^0 mass
- $l_K - l_\pi < 0$ for kaon from D^0 candidate
- $|\cos\theta_{sp}| < 0.8$
- Beam constrained B mass within 10 MeV/c^2 of nominal mass
- π^0 candidate mass within 10 MeV/c^2 of nominal mass
- $l_\pi - l_K < 0$ for fast pion

The PID requirement on the fast pion eliminates most of the $B^- \to D^0 K^-$ background. The ΔE distributions for the three D^0 submodes are shown in Figure 3.14. The value of the mean for the three D^0 modes are 1.4±1.2 MeV, -0.3±1.5 MeV, and -0.5±1.5 MeV respectively. We conclude that this is consistent with no systematic offset in the ΔE distribution, and we assign a systematic uncertainty of 1.5 MeV on the mean.

To determine the width of the ΔE distribution in our signal modes, we employed a procedure based on the ΔE distribution measured in $B^- \to D^0\pi^-$ ($D^0 \to K^-\pi^+$) decays and the momentum distribution of muons measured in $e^+e^- \to \mu^+\mu^-$ events. We characterize the momentum resolution by a simple model,

$$\left(\frac{\sigma_p}{p}\right)^2 = A^2 + B^2 p^2, \tag{3.11}$$

averaging over the detector acceptance. The D^0 decay daughters have a wide spectrum of momenta up to about 2 GeV, the primary π^- lies between 2.1 and 2.5 GeV, and the muons have momentum near 5.3 GeV. The observed $D^0\pi^-$ ΔE spread and the observed $\mu^+\mu^-$ momentum spread can be reproduced assuming the model in Eq. 3.11 and fitting for suitable values of A and B. The resulting values differ slightly between old and new Pass2, as indicated in Table 3.5.

Using these values of A and B we calculate the contribution of tracking to ΔE width. We test the method on Monte Carlo and find that it predicts ΔE for the $K\pi$ mode correctly. We also find that we get a satisfactory description if we use this recipe on the two pions from K^0_S decays. The π^0 contribution to ΔE width is taken from MC.

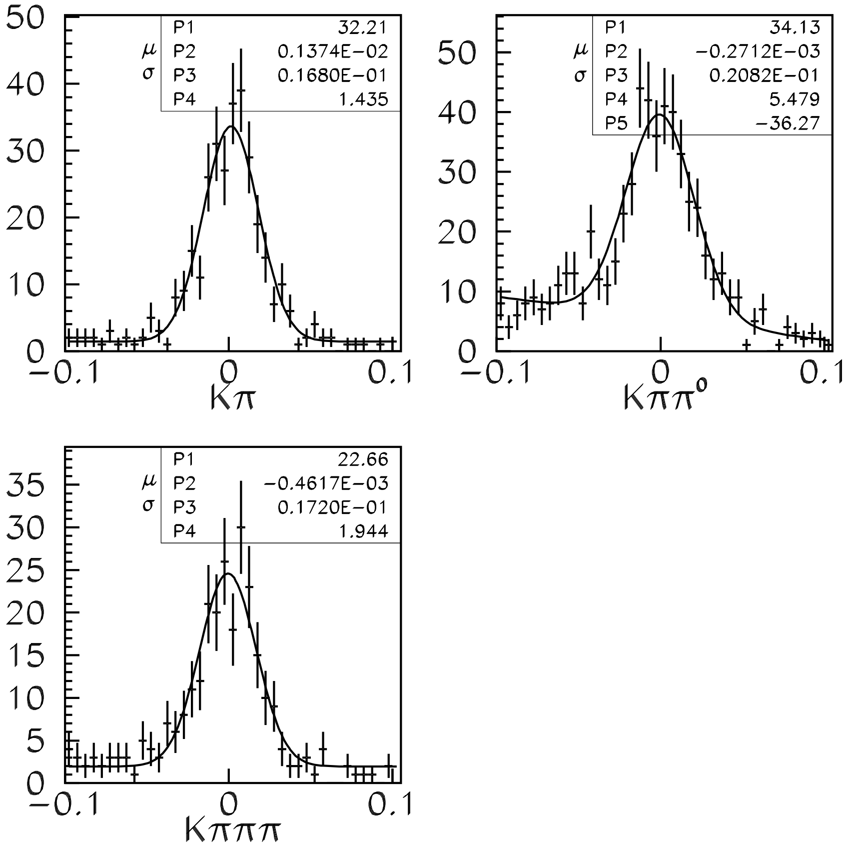

Figure 3.14: ΔE distribution for $B^- \to D^0\pi^-$, with the D^0 reconstructed in the indicated mode.

Table 3.5: A and B coefficients in Old and New Pass2

Quantity	Old Pass2	New Pass2
'A' Coefficient	0.0055	0.0044
'B' Coefficient (GeV^{-1})	0.0011	0.0010

3.4.3 Mode Dependent PDF's and Cuts

The parameters that describe the mode-dependent PDF's, and the error on those parameters are shown in Tables 3.6 and 3.7. Here G, BG and BW refer to a gaussian, bifurcated gaussian and Breit-Wigner, respectively. When the PDF for a particular mode uses only one function where the sum of two functions is indicated, no values are listed for the unused function. When only σ_L is given for a bifurcated gaussian, this indicates that a gaussian was used instead of a bifurcated gaussian for the PDF for that particular mode.

h^+h^-

All three h^+h^- modes $(\pi^+\pi^-, K^+\pi^-, K^+K^-)$ require both tracks to be positively identified according to the particle identification method described above. For $\pi^+\pi^-$ and K^+K^- signal fits, $K^+\pi^-$ is included as a background. The backgrounds were chosen by calculating relative yields, given the particle identification efficiencies and fake rates and the order of magnitude of relative branching ratios. The ΔE variable is calculated for all three modes assuming that both tracks are pions. Therefore the mean of the ΔE distribution for $K^+\pi^-$ and K^+K^- is offset by 42.3

MeV and 86.3 MeV, respectively.

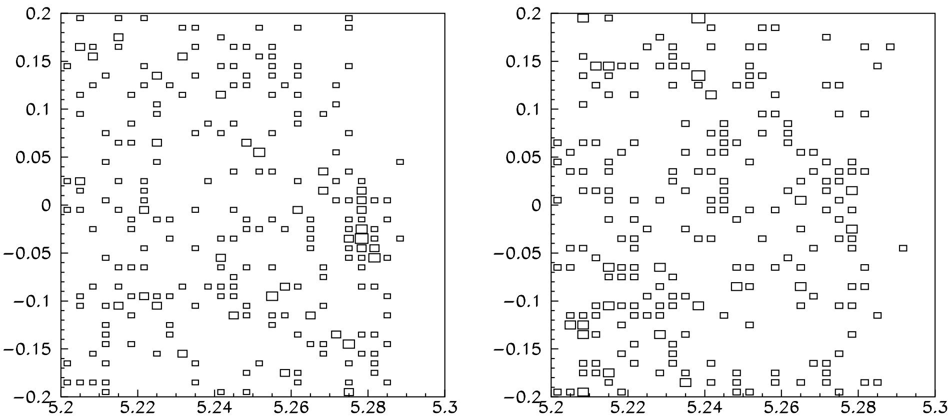

Figure 3.15: ΔE vs. M_B for $K^+\pi^-$ and $\pi^+\pi^-$, with a cut on $\mathcal{F}$.

K^0K^0

For the K^0K^0 mode, we use several additional cuts. The invariant mass of the K^0 candidate is required to be within 10.4 MeV of the nominal mass. The second K^0 candidate is required to be within 13.4 MeV of the nominal mass. The vertex position in the $x - y$ plane is required to be greater than 5 mm. Also, the K^0 is required to have the cosine of the angle between the momentum vector and the vector pointing from the origin to the vertex be greater than 0. This ensures that the K^0 momentum vector projects somewhere near the origin.

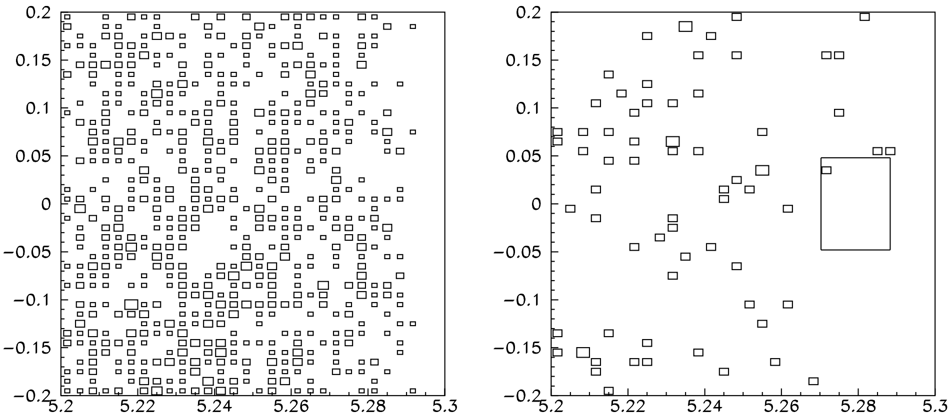

Figure 3.16: ΔE vs. M_B for K^+K^- and K^0K^0 modes, with no additional cuts. The box indicates the 3σ signal region.

$h^+\pi^0$

In the $h^+\pi^0$ modes, the track is required to pass the standard particle identification cuts. In addition, the π^0 is required to be within 2.5σ of the nominal mass. In the $\pi^+\pi^0$ mode, $K^+\pi^0$ is included as a background component in the fit.

h^+K^0

The h^+K^0 mode requires that the K_S^0 be within 10.4 MeV of the nominal mass, that the radial vertex position be greater than 5 mm from the origin, and that the charged track pass particle ID requirements. Also, the K_S^0 is required to have the cosine of the angle between the momentum vector and the vector pointing from the origin to the vertex be greater than 0. This ensures that the K_S^0 momentum vector

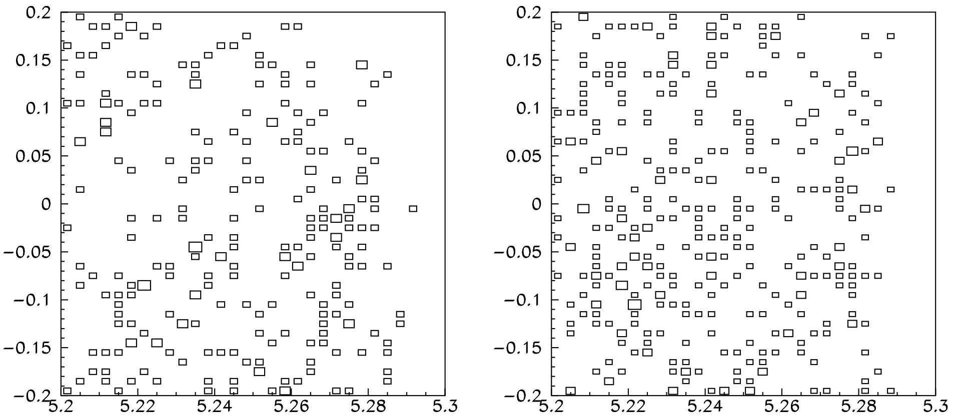

Figure 3.17: ΔE vs. M_B for $K^+\pi^0$ and $\pi^+\pi^0$, with a cut on $\mathcal{F}$.

projects somewhere near the origin. The K^+K^0 mode does not include π^+K^0 as a background in the likelihood fit.

$K^0\pi^0$

K_S^0 candidates for the $K^0\pi^0$ mode are required to pass the same criteria as the K_S^0's used in the h^+K^0 mode described above. The π^0 mass is required to be within 2.5σ of the nominal π^0 mass.

$\pi^0\pi^0$

For $\pi^0\pi^0$, both π^0s are required to be within 2.5σ of the nominal mass. $B \to \rho^+\pi^0$ is included as a background component in the likelihood fit. PDF's for the $\rho^+\pi^0$ component are

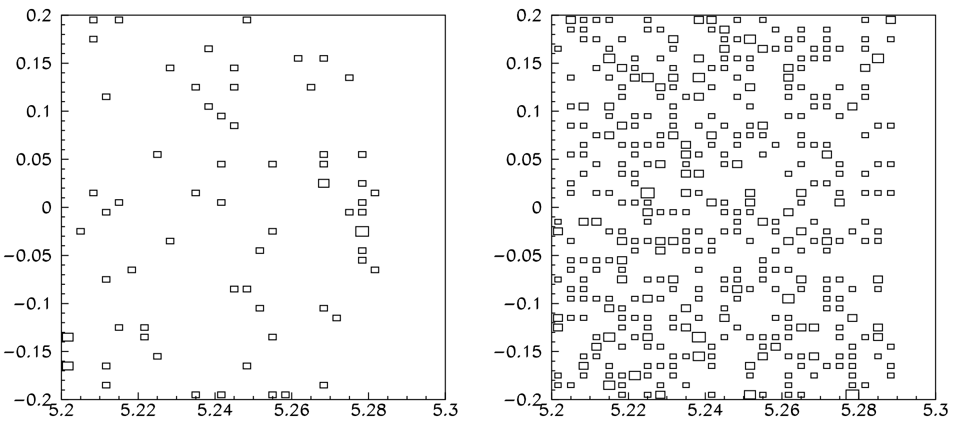

Figure 3.18: ΔE vs. M_B for $K^0\pi^+$, with a cut on $\mathcal{F}$, and K^0K^+, with no additional cuts.

- M_B: Bifurcated gaussian, $\mu = 5279.3 \pm 0.3$MeV, $\sigma_L = 8.4 \pm 0.2$MeV, $\sigma_R = 3.2 \pm 0.2$MeV
- ΔE: Bifurcated gaussian, $\mu = -261 \pm 1.5$MeV, $\sigma_L = 132 \pm 3.7$MeV, $\sigma_R = 69 \pm 3.7$MeV
- Fisher: Bifurcated gaussian, $\mu = -0.073 \pm 0.116$, $\sigma_L = 0.681 \pm 0.091$, $\sigma_R = 0.354 \pm 0.105$

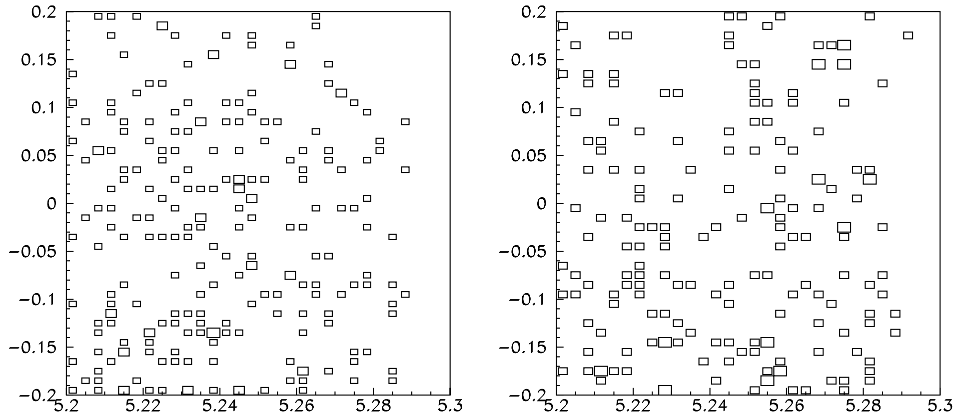

Figure 3.19: ΔE vs. M_B for $\pi^0\pi^0$, with no additional cuts, and $K^0\pi^0$, with a cut on $\mathcal{F}$.

Table 3.6: PDF parameters for the various modes. The variable Q is the ΔE offset described in Section 3.4.3 for the three h^+h^- modes.

PDF		h^+h^-	$K^+\pi^0$	$\pi^+\pi^0$	$\pi^0\pi^0$
	σ	2.7±.15	3.15±.15	2.6±.15	-
Signal M_B	σ_L	-	9.6±.15	5.5±.15	4.51±.15
G+BG	σ_R	-	.57±.15	2.2±.15	2.67±.15
	BG/G	-	.227	1.59	-
Bkg. M_B	Efact	-29.6±1.6	-16.8±1.1	-16.8±1.2	-7.0±1.1
	μ_1	Q±1.5	-4.5±1.5	-2.4±1.5	-11.3±1.5
	σ_L	18.9±1.9	31.7±3.0	42.6±3.0	54.4±3.7
Signal ΔE	σ_R	-	30.2±3.0	31.1±3.0	32.2±3.7
	μ_1	-	-79.6±1.5	7.6±1.5	83.6±1.5
BG+BG	σ_L	-	42.9±3.0	139.2±3.0	103±3.7
	σ_R	-	77.8±3.0	42.9±3.0	89.0±3.7
	1/2	-	.435	.399	1.82
Bkg. ΔE	m	-.895±.045	-.95±.11	-1.04±.12	-.63±.11
Signal $\mathcal{F}$	μ	.02±.09	.014±.055	.074±.051	-.015±.057
BG	σ_L	.777±.013	.802±.033	.836±.033	.775±.037
	σ_R	.547±.013	.549±.052	.471±.043	.527±.053

Table 3.7: PDF parameters for the various modes.

PDF		$K^0\pi^-$	$K^+\overline{K^0}$	$K^0\pi^0$	$K^0\overline{K^0}$
	σ	-	-	3.07±.15	-
Signal M_B	σ_L	3.0±.15	2.6±.15	8.5±2.8	3.2±.15
G+BG	σ_R	2.5±.15	2.7±.15	3.0±1.0	2.7±.15
	BG/G	-	-	.162±.062	-
Bkg. M_B	Efact	-23.2±2.1	-23.2±2.3	-15.7±3.6	-26.0±6.0
	μ_1	4.2±1.4	4.4±.5	6.9±2.0	6.6±1.5
Signal ΔE	σ	20.2±1.9	20.2±1.9	31.2±1.8	16.0±1.0
	σ_L	35.3±6.4	30.0±3.4	118±10	45.3±3.7
G+BG	σ_R	-	-	54.5±4.0	-
	BG/G	.174±.024	.212±.068	.405±.05	.238±.045
Bkg. ΔE	m	-1.51±.25	-.83±.83	-1.23±.42	-1.2±.1
	μ	.007±.051	-.026±.041	-.095±.048	0.±.04
Signal $\mathcal{F}$	σ_L	.76±.035	.77±.032	.82±.045	.755±.02
	σ_R	.62±.032	.62±.033	.69±.06	.56±.015
BG+BW	σ_{BW}	-	.47±.32	.782±.238	-
	BW/BG	-	.081±.062	.261±.077	-

3.5 Yields and Efficiencies

Table 3.8: Fit results for old (left) and new (right) pass2, with statistical errors. Efficiencies include any relevant PID efficiencies.

Mode	Fit Yield	Efficiency(%)	Mode	Fit Yield	Efficiency(%)
$\pi^+\pi^-$	$7.8^{+5.5}_{-4.5}$	39.0	$\pi^+\pi^-$	$4.3^{+4.1}_{-3.1}$	45.3
$\pi^+\pi^0$	$2.8^{+3.3}_{-1.9}$	34.9	$\pi^+\pi^0$	5.7 ± 5.9	37.5
$\pi^0\pi^0$	$2.2^{+2.5}_{-1.5}$	22.1	$\pi^0\pi^0$	$0.4^{+2.7}_{-3.4}$	22.4
$K^+\pi^-$	$28.1^{+6.8}_{-6.0}$	37.9	$K^+\pi^-$	$19.1^{+5.3}_{-4.6}$	45.3
$K^0\pi^-$	$12.1^{+4.4}_{-3.7}$	12.3	$K^0\pi^-$	$2.9^{+1.8}_{-2.7}$	12.8
$K^+\pi^0$	$16.7^{+6.2}_{-5.3}$	32.6	$K^+\pi^0$	$10.8^{+5.1}_{-4.1}$	35.3
$K^0\pi^0$	$3.5^{+2.8}_{-1.9}$	9.6	$K^0\pi^0$	$2.9^{+2.4}_{-1.6}$	10.5
K^+K^-	$2.3^{+3.7}_{-2.9}$	35.2	K^+K^-	0.0 ± 0.7	42.1
K^0K^-	0.0 ± 0.6	13.8	K^0K^-	0.0 ± 0.6	13.0
K^0K^0	0.0 ± 0.5	8.1	K^0K^0	0.0 ± 0.5	8.0

3.6 Systematics

Certain systematic errors in the determination of branching ratios are common to many or all modes. Here we describe the results of our studies for these systematics.

3.6.1 $N_{B\bar{B}}$

The determination of $N_{B\bar{B}}$ is discussed in detail in Section 2.3.1. We repeat the result here: $N_{B\bar{B}} = (5.73 \pm 0.47) \times 10^6$. This corresponds to an 8.2% error.

3.6.2 MC

For our efficiency and fitting systematic error we use the statistical errors on our signal MC. In all cases this is a small ($\mathcal{O}(1\%)$) error.

3.6.3 Tracking

For the tracking error, we use Feng Liu's study [20] of tau decays in data and MC to establish a 1% track-finding systematic error. There is also a particle ID systematic error from the finite statistics of our D^* sample of 3% per track, which gives an overall track efficiency error of 3.2% per track.

3.6.4 π^0's

We estimate the systematic error associated with π^0 reconstruction in hadronic environments by examining the ratio of measured branching ratios

$$R = \frac{BR(D^0 \to K^-\pi^+\pi^0)}{BR(D^0 \to K^-\pi^+)}. \tag{3.12}$$

By measuring the relative efficiency corrected number of D decays into the $K^-\pi^+\pi^0$ and $K^-\pi^+$ final states, and accounting approximately for known deviations from uniform three-body phase space decay, we can estimate the relative Monte Carlo and data π^0 finding efficiency:

$$\frac{\epsilon_{DATA}(\pi^0)}{\epsilon_{MC}(\pi^0)} = \frac{R_{CLEO}}{R_{PDG}} \tag{3.13}$$

The value of R has been previously measured at CLEO II [21], with a value of

$$\frac{BR(D^0 \to K^-\pi^+\pi^0)}{BR(D^0 \to K^-\pi^+)} = 3.78 \pm 0.071 \pm 0.26. \tag{3.14}$$

The CLEO value dominates the PDG02[22] average, which has a value of 3.44 ± 0.22. These two numbers indicate that in the absence of any other systematic error, without redoing the subcomponent analysis, the best we could hope for from this technique is a 7% error. The full Dalitz analysis is done in [23] by Bergfeld and Selen. They also find that the dominant subcomponent in the three-body decay is the $K^-\rho^+$ component, which has a reported fit fraction of:

$$\frac{\int d\Gamma |\mathcal{A}_{K^-\rho^+}|^2}{\int d\Gamma |\sum \mathcal{A}|^2} = 0.79 \pm 0.05 \tag{3.15}$$

Rather than use the full matrix element for this study, we have assumed that $79 \pm 10\%$ of the time, the D decays via the $K^-\rho^+$ and decays via phase space the rest of the time. The other components in the fit have fit fraction similar in size to the non-resonant component, and phases that tend to cancel. Table 3.9 also summarizes the results of varying the proportion of the $K^-\rho^+$ by a generous amount (± 0.10)

Tracks selected for this study are all required to pass the rare B standards. In addition, all charged kaon and pion candidates, as well as π^0's, are required to have more than 300 MeV/c of momentum. The lower (higher) energy photon from the π^0 must have energy above 100 MeV (500 MeV). The resulting D is also required to have a momentum above 2 GeV/c. The π^0 mass is required to be between 120 MeV/c^2 and 145 MeV/c^2. To avoid additional systematic uncertainties, we have not used any charged particle identification.

The data used were the same combination of old and new pass2 as in the rest of the rare B analysis. The Monte Carlo used to assess efficiency consists of 5000 generated events in the mode $D^0 \to K^-\pi^+$, 10000 in the $D^0 \to K^-\pi^+\pi^0$ non-resonant mode, and 20000 in the $D^0 \to K^-\rho^+$ mode. We have not attempted to model the smaller $K^*\pi$ resonances.

The data and Monte Carlo surviving after the above cuts are shown in Figures 3.20, 3.21 and 3.22.

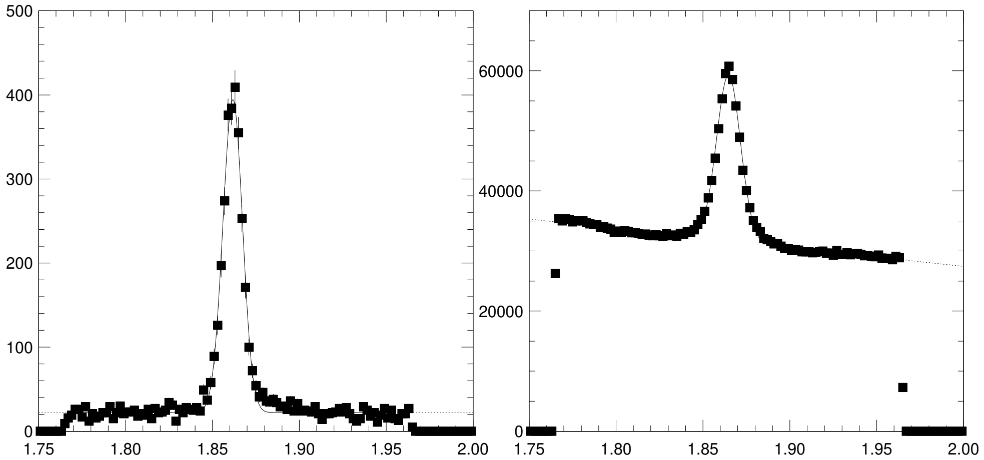

Figure 3.20: $K^-\pi^+$ Monte Carlo (left) and data (right). The MC peak has 2636±36 events, and the data peak has 238,200±1000 events.

The resulting histograms were fit in three different ways to assess the dependency on fitting formulation. Figures 3.20- 3.22 use simple fits to a sum of a

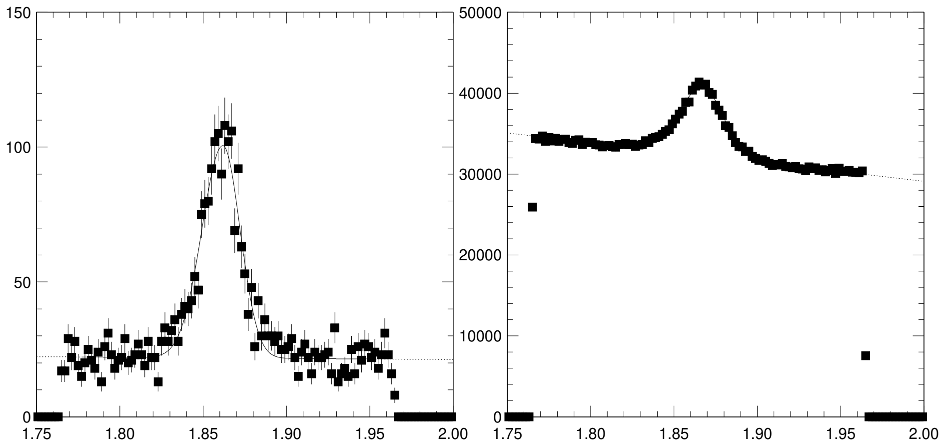

Figure 3.21: $K^-\pi^+\pi^0$ Monte Carlo (left) and Data (right). The Monte Carlo is non-resonant $D^0 \to K^-\pi^+\pi^0$. The MC peak has 1131±48 events, and the data peak has 139,500±1500 events.

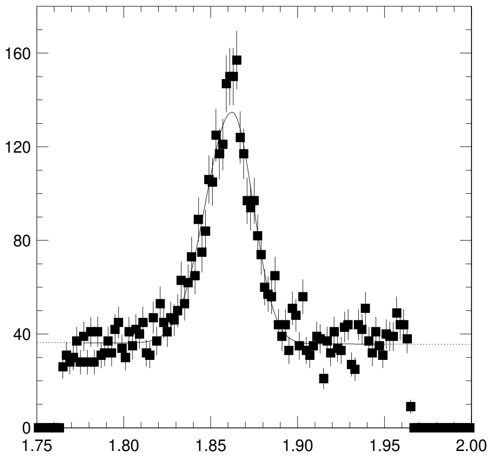

Figure 3.22: Monte Carlo for $D^0 \to K^-\rho^+$. The MC peak has 1701±66 events.

bifurcated Gaussian and a first order polynomial. Fits were also done to the sum of a bifurcated Gaussian, a symmetric Gaussian and a first order polynomial. Finally, the signal size was estimated by side band subtraction, considering the side bands as $(1.77, 1.82) \bigcup (1.90, 1.95)$ and the signal region as $(1.83, 1.89)$. In all cases, the fit to a single Gaussian did not give a very good confidence level, and the fit to a sum of Gaussians did. Fortunately, however, we are insensitive to this. Table 3.9 summarizes the results of the fits. The Monte Carlo used in the fits for the $K^-\pi^+\pi^0$ is a mixture of pure phase space Monte Carlo and $K^-\rho^+$ Monte Carlo.

The results of this analysis support the assertion that the CLEO III Monte Carlo models the data π^0 finding efficiency for momenta above 300 MeV/c at the 10 % level or better.

Additionally, we studied the following η decays: $\eta \to \pi^+\pi^-\pi^0$, and $\eta \to 3\pi^0$ (similar to the CLEO II π^0 analysis [24]). We ran over all new pass2 (data5-8 and data11) data and most of the generic $q\bar{q}$ MC which was generated for data7 and data8. The cuts used were the same as for the rest of the analysis and are explained in Section 3.2. In addition we used a minimum cut of 0.2 GeV/c for π^0's and charged pions. This should make the comparison between $\pi^+\pi^-\pi^0$ and $3\pi^0$ more accurate, since the efficiency for low momentum charged pions is not constant and perhaps not very well understood. All the fits were done using the following method. First we fit the background excluding the η signal with a polynomial of the lowest possible order to get a good fit. The same order polynomial was used for Monte Carlo and data. All the parameters of the polynomial were fixed with the exception of an overall constant, which was left floating, and a gaussian, bifurcated

Table 3.9: Summary of R results. The columns G+G and SBS denote fits to a sum of two Gaussians and a straight line and a side band subtraction. The last line shows the estimated π^0 systematic including the propagated Monte Carlo and Data statistical errors, as well as the Dalitz model variation.

	G+G	SBS
$K^-\pi^+$ Data	$2.69 \times 10^5 \pm 1345$	$2.49 \times 10^5 \pm 1332$
$K^-\pi^+$ MC	57.8 ± 1.5	55.8 ± 1.2
$K^-\pi^+\pi^0$ Data	$1.47 \times 10^5 \pm 1774$	$1.37 \times 10^5 \pm 1301$
$K^-\pi^+\pi^0$ MC	12.4 ± 0.7	11.6 ± 0.5
$K^-\rho^+$ MC	8.7 ± 0.4	8.3 ± 0.3
R($K^-\rho^+$ at 69%)	3.20 ± 0.14	3.30 ± 0.12
R($K^-\rho^+$ at 79%)	3.32 ± 0.15	3.42 ± 0.13
R($K^-\rho^+$ at 89%)	3.46 ± 0.16	3.55 ± 0.14
$\epsilon_{DATA}(\pi^0)/\epsilon_{MC}(\pi^0)$	$0.97 \pm 0.08 \pm 0.04$	$1.00 \pm 0.08 \pm 0.02$

gaussian, or double gaussian which was used to fit the signals. The results of the fits are given in Table 3.10. We use the PDG values for the branching fractions of the η decays. The π^0 efficiency is estimated by

$$\rho = \frac{\epsilon(\pi^0)}{\epsilon(\text{track})} = \sqrt{\frac{N(3\pi^0)}{N(\pi^+\pi^-\pi^0)} \cdot \frac{BR(\eta \to \pi^+\pi^-\pi^0)}{BR(\eta \to 3\pi^0)}} \tag{3.16}$$

The ratio $\frac{\rho_{data}}{\rho_{MC}}$ is a measure of how well the MC agrees with the data. Note that we believe we know the tracking efficiency quite well, and that this error does not contribute significantly to our overall π^0 error.

Table 3.10: CLEO III π^0 efficiency studies with different cuts on η and π^0 momenta. The last column would indicate that in CLEO III there are fewer π^0's found in the data than in the Monte Carlo.

cut	ϵ_{data}	ϵ_{MC}	$\frac{\epsilon_{data}}{\epsilon_{MC}}$
p_η >1 GeV/c	0.422±0.022	0.543±0.030	0.78±0.06
p_η >2 GeV/c	0.239±0.015	0.307±0.022	0.78±0.07
p_{π^0} >1 GeV/c	0.504±0.025	0.607±0.027	0.83±0.06
0.5< p_{π^0} <1 GeV/c	0.545±0.019	0.624±0.019	0.87±0.04

It should be noted that this study has some potential drawbacks. The environment in η decays is different than in our rare B decays. η's have three π^0's all close to each other, with potentially six photons all in the same area of the calorimeter, while for the single π^0 in a hemisphere for the rare B decays there is a minimum separation of the photons from kinematics. Toy MC of 2 GeV η's decaying to phase space $3\pi^0$ shows that a small change in the spatial separation required to resolve two nearby photons can cause large changes in efficiencies. If we assume that there is some minimum spatial separation required between the photons for them to be identified as two distinct photons, and that the MC had this value at 10 cm, but the value of this parameter in data is 11 cm, the effect is a 10% decrease in efficiency. This effect is solely due to the overlap of photons from two different π^0's, i.e. an η effect, not a π^0 effect. This is a naive model, but it illustrates that this method is sensitive to features other than those which we wish to measure. This is much less of a concern in the D^0 decays, so we use the D^0 study to determine the systematic error. We conclude that a 15% systematic error on π^0 efficiency is a conservative estimate, while not being so large as to be a dominant contribution to the overall error on the branching ratio.

Since we use the cleo3_Apr19_2002 software release we did not benefit from the work done to suppress noisy crystals. In order to prove that this does not change the number of high momentum π^0's, we conducted the following test. We ran over data5 with the cleo3_Apr19_2002 version and counted 82840 π^0 events between 2.1 and 2.6 GeV on a background of 260 events/MeV. Running over the same data with a release which suppresses the noisy crystals we find 82410 π^0's on a background of 240 events/MeV. From this we conclude that noisy crystals do not produce a significant number of fake π^0's in our momentum range. Data5 is by far the data set with the largest number of noisy crystals, and therefore the worst case.

3.6.5 K_S^0

We measured the K_S^0 efficiency in data using $D^+ \to K^-\pi^+\pi^+$ compared to $D^+ \to K_S^0\pi^+$. We used tracks selected by rare B standards as described in Section 3.2, and we look at high momentum K_S^0. We used the following additional cuts:

- $p_D > 2.5$ GeV
- radial distance from origin > 0.005 m
- $\vec{p_{K_S^0}} \cdot \vec{v_{K_S^0}} > 0$, where $\vec{v_{K_S^0}}$ is the vector pointing from the origin to the vertex.

We measured the following ratio,

$$\frac{D^+ \to K^0_S \pi}{D^+ \to K\pi\pi} \times \frac{\epsilon(D^+ \to K\pi\pi)}{\epsilon(D^+ \to K^0_S \pi)} \times \frac{BR(D^+ \to K\pi\pi)}{BR(D^+ \to K^0_S \pi)} = 1.013 \pm 0.065, \qquad (3.17)$$

and conclude that there is no statistically significant difference between data and MC. We take the error on this ratio to be our systematic uncertainty on K^0_S reconstruction and we therefore assign a 6.5% systematic error. Plots from data and MC for both D^+ modes are shown in Figures 3.23-3.26.

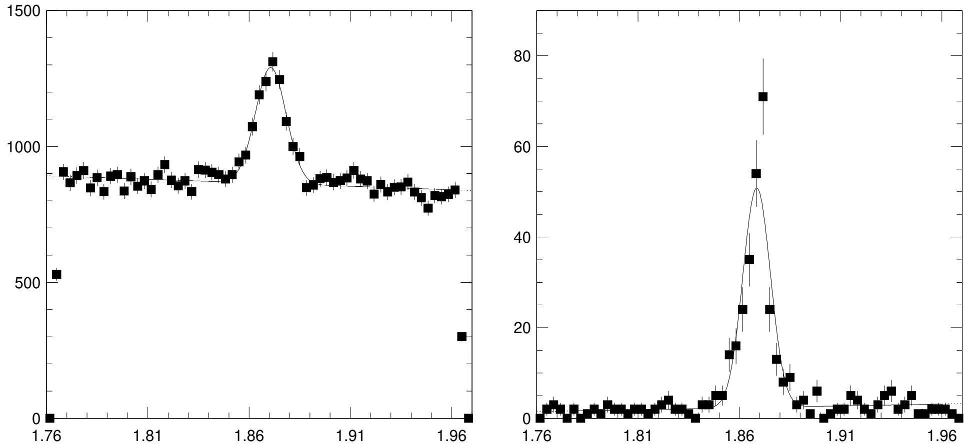

Figure 3.23: $D^+ \to K^0_S \pi^+$ yield for data 5, 6, 7, 8, 11 (new pass2) (left) and the corresponding 2500 MC events (right).

A second study was performed to cross check the K^0_S result, albeit with less accuracy. We reconstruct Λ_c candidates in both the $pK\pi$ and pK^0_S modes, and compare the efficiency corrected yields for these two decay modes to the measured branching ratios. The study is based on the datasets 6, 7, 8, and 11 (3.3 fb^{-1}) for

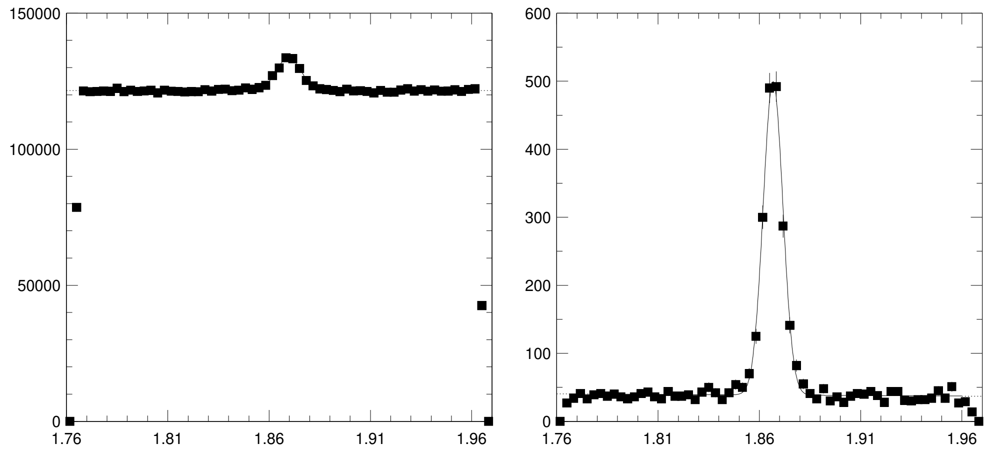

Figure 3.24: $D^+ \to K^-\pi^+\pi^+$ yield for data 5, 6, 7, 8, 11 (new pass2) (left) and the corresponding 2500 MC events (right).

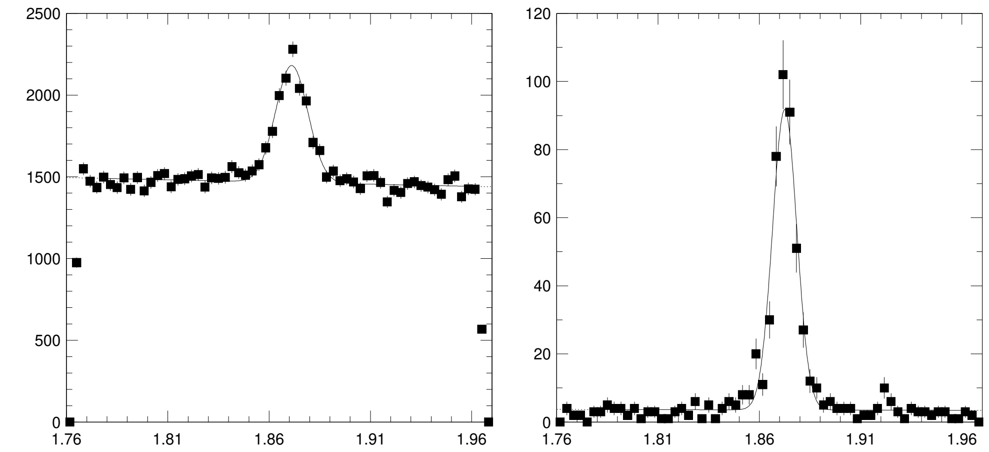

Figure 3.25: $D^+ \to K^0_S\pi^+$ yield for data 9, 10, 12, 13, 14 (old pass2) (left) and the corresponding 8892 MC events (right).

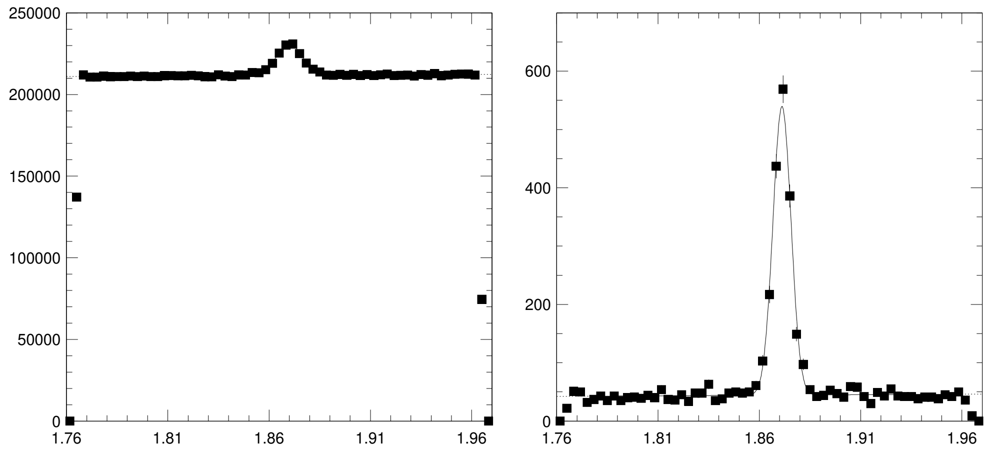

Figure 3.26: $D^+ \to K^-\pi^+\pi^+$ yield for data 9, 10, 12, 13, 14 (old pass2) (left) and the corresponding 5651 MC events (right).

$x_p > 0.5$ (eq. 3.18).

$$x_p = \frac{P}{\sqrt{E_{beam}^2 - M_{\Lambda_c}^2}}. \tag{3.18}$$

Table 3.11: CLEO III Monte Carlo efficiencies and mass peak widths for $x_p > 0.5$.

Λ_c mode	Efficiency (%)	Width (MeV/c^2)
$pK^-\pi^+$	47.1 ± 1.4	3.65
pK_s	25 ± 1.0	5.6

Monte Carlo efficiencies and data yields for the Λ_c modes are given in Table 3.11 and Table 3.12, respectively. As in the D^+ study above, we measure the ratio of yields divided by the ratio of branching fractions and find

Table 3.12: CLEO III data Λ_c yield for $x_p > 0.5$.

Λ_c mode	Yield	Corrected Yield
$pK^-\pi^+$	8627 ± 172	18315 ± 656
pK_s	1161 ± 40.0	4643 ± 243

$$\frac{\Lambda_c \to pK_S^0}{\Lambda_c \to pK\pi} \times \frac{\epsilon(\Lambda_c \to pK\pi)}{\epsilon(\Lambda_c \to pK_S^0)} \times \frac{BR(\Lambda_c \to pK\pi)}{BR(\Lambda_c \to pK_S^0)} = 1.08 \pm 0.11. \tag{3.19}$$

We find that the ratio is again consistent with no systematic offset in K_S^0 reconstruction efficiency between data and MC. We use this result as a cross check for the systematic error found using the $D^+ \to K^-\pi^+\pi^+$ and $K_S^0\pi^+$ decays.

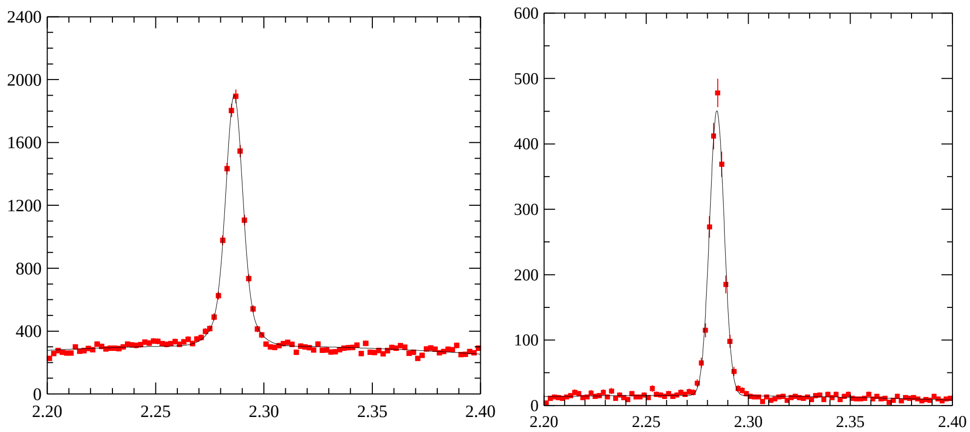

Figure 3.27: $\Lambda_c \to pK\pi$ yield for data 6, 7, 8, 11 (new pass2) (left) and the corresponding MC (right).

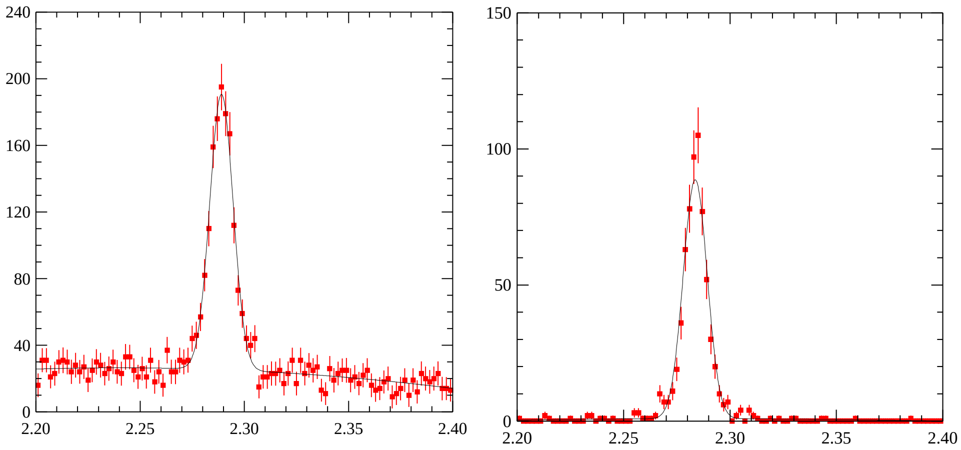

Figure 3.28: $\Lambda_c \to pK_s$ yield for data 6, 7, 8, 11 (new pass2) (left) and the corresponding MC (right).

3.6.6 PDF's

A common strategy for evaluating systematic errors due to the uncertainty in PDF parameters is to vary a single parameter by its error, refit the data, and consider the change in yield to be the systematic error for that parameter. All these errors are added in quadrature and this number is then divided by the efficiency and the number of B's to give the PDF systematic error on the branching ratio. Things are not quite as simple in this analysis since we do two separate fits. To add the systematic errors for old and new pass2, we do the standard parameter variation study on both old and new pass2, and calculate the change in yield due to each variation. We then divide this change in yield in the old pass2 fit by the old pass2 efficiency, and divide the change in yield in the new pass2 fit by the new pass2 efficiency. Since the errors associated with a given parameter variation (for example, moving the mean of the M_B distribution) should be completely correlated between old and new pass2, we add the change in yield for each parameter variation linearly (not in quadrature) to find the efficiency corrected change in yield for a given systematic variation. Tables 3.13 and 3.14 show the results of the systematic variations for the $\pi^+\pi^-$ and $K^+\pi^-$ modes. Each pair of entries represents the negative and positive fluctuations, respectively. The old and new variation columns are the changes in the raw yield, and they are efficiency corrected and added to get the variation in the last column. Signal parameters are listed in the top half of the tables, and background parameters in the bottom half. In all cases, the PDF systematic error is small compared to the multiplicative and/or the statistical errors. Since all other modes exhibit the same general behavior as $\pi^+\pi^-$ and $K^+\pi^-$ with respect to systematic variations, we have omitted the PDF variation table for the rest of the modes.

Table 3.13: Summary of PDF variations for $\pi^+\pi^-$.

Parameter	Old Var.	New Var.	Combined Var.
μ_{M_B}	0.321, -0.369	0.269,-0.290	1.423, -1.594
σ_{M_B}	-0.427, 0.377	-0.078,0.054	-1.270, 1.090
$\sigma_{\Delta E}$	-0.413, 0.321	-0.161, 0.129	-1.419, 1.110
$\mu_{\Delta E}$	-0.043, 0.031	0.210, -0.223	0.353, -0.412
$\mu_{\mathcal{F}}$	-0.147, 0.144	-0.054, 0.054	-0.499, 0.490
$\sigma_{L_{\mathcal{F}}}$	0.063, -0.064	0.041, -0.041	0.255, -0.255
$\sigma_{R_{\mathcal{F}}}$	-0.053, 0.047	-0.041, 0.042	-0.228, 0.216
Efact	-0.221, 0.223	-0.108, 0.110	-0.809, 0.815
Slope	0.015, -0.015	0.004, -0.004	0.048, -0.048
$\mu_{\mathcal{F}}$	-0.172, 0.173	-0.110, 0.111	-0.686, 0.691
$\mu_{G_{\mathcal{F}}}$	-0.016, 0.009	-0.015, 0.013	-0.076, 0.052
$\sigma_{L_{\mathcal{F}}}$	0.125, -0.123	0.082, -0.079	0.503, -0.493
$\sigma_{R_{\mathcal{F}}}$	-0.287, 0.279	-0.112, 0.119	-0.985, 0.981
$\sigma_{BW_{\mathcal{F}}}$	0.010, -0.010	0.024, -0.020	0.080, -0.071
$\sigma_{G_{\mathcal{F}}}$	0.006,-0.006	0.002, -0.002	0.0220, -0.021
Area_{BW}	0.050, -0.047	-0.034, 0.035	0.052, -0.042
Area_{G}	-0.198, 0.200	-0.092, 0.096	-0.713, 0.726

Table 3.14: Summary of PDF variations for $K^+\pi^-$.

Parameter	Old Var.	New Var.	Combined Var.
μ_{M_B}	0.090, -0.172	-0.178, 0.144	-0.154, -0.136
σ_{M_B}	-0.285, 0.240	-0.297, 0.280	-1.411, 1.253
$\sigma_{\Delta E}$	-0.657, 0.588	-0.503, 0.434	-2.846, 2.511
$\mu_{\Delta E}$	-0.446, 0.427	-0.049, 0.026	-1.285, 1.188
$\mu_{\mathcal{F}}$	-0.347, 0.344	-0.055, 0.051	-1.037, 1.022
$\sigma_{L_{\mathcal{F}}}$	0.104, -0.104	0.004, -0.009	0.284, -0.295
$\sigma_{R_{\mathcal{F}}}$	-0.271, 0.265	-0.045, 0.043	-0.818, 0.797
Efact	-0.299, 0.302	-0.137, 0.138	-1.094, 1.103
Slope	-0.016, 0.016	-0.010, 0.010	-0.065, 0.065
$\mu_{\mathcal{F}}$	-0.100, 0.105	-0.120, 0.121	-0.531, 0.545
$\mu_{G_{\mathcal{F}}}$	-0.049, 0.046	-0.014, 0.012	-0.163, 0.148
$\sigma_{L_{\mathcal{F}}}$	0.048, -0.046	0.106, -0.104	0.363, -0.355
$\sigma_{R_{\mathcal{F}}}$	-0.211,0.240	-0.139, 0.143	-0.866, 0.950
$\sigma_{BW_{\mathcal{F}}}$	0.028, -0.021	0.070, -0.064	0.230, -0.198
$\sigma_{G_{\mathcal{F}}}$	0.004, -0.004	0.002, -0.002	0.017, -0.018
Area_{BW}	0.008,-0.006	0.080, -0.075	0.199, -0.183
Area_{G}	-0.193, 0.206	-0.106, 0.109	-0.744, 0.787

3.7 Combining Old and New Pass2

Once the likelihood fits for old and new pass2 have been performed, we use the likelihood shapes from FELIX to combine the fit results. We use toy MC to randomly select yields according to the likelihood distribution for old and new pass2, and divide by the efficiency and $N_{B\bar{B}}$ for old and new pass2, respectively. This gives a likelihood function for old and new pass2 in terms of branching ratios. We then multiply the likelihood functions bin by bin to get the CLEO III likelihood function. We convert this plot to an effective χ^2 plot by taking $-2\ln\mathcal{L}$, from which we determine the asymmetric statistical errors for CLEO III.

Once we have the statistical errors measured, we smear the likelihood function according to our systematic errors. There are two types of systematic errors involved in this process. We have additive systematic errors (σ_A) from our PDF variations. These errors are in general asymmetric, specified by widths σ_L and σ_R, so we apply a bifurcated smearing. The additive systematic smearing is therefore generated according to the distribution given in Eq. 3.20.

$$p(\sigma_A) = \begin{cases} e^{-x^2/2\sigma_L^2}, x < 0 \\ e^{-x^2/2\sigma_R^2}, x > 0 \end{cases} \tag{3.20}$$

The other type of error in our analysis is multiplicative errors. This includes errors on $N_{B\bar{B}}$, tracking efficiency, π^0 efficiency, etc. The multiplicative errors are added in quadrature to obtain an overall multiplicative error(σ_M) for a given mode. The multiplicative systematic smearing is generated according to the distribution shown in Eq. 3.21.

$$p(\epsilon) = e^{-(x-1)^2/2\sigma_M^2} \tag{3.21}$$

When we generate our toy MC to model the systematic errors, we choose a yield Y according to our likelihood distribution, and smear it according to Eq. 3.22.

$$Y' = Y/\epsilon + \sigma_A \tag{3.22}$$

After this smearing, we have a likelihood function that includes all of the errors in our measurement. From this we find the minimum of the $-2\ln\mathcal{L}$ distribution to measure our mean, and find the one σ intersections to measure our errors. Since this is the total error, we calculate the systematic error by subtracting in quadrature the statistical error from the total error. We also calculate upper limits by integrating the likelihood function over positive yields and calculate significances by looking at the zero yield value of the $-2\ln\mathcal{L}$ distribution.

While generating our toy MC, it is possible to obtain a negative yield if the additive or multiplicative systematic smearing has a relatively large negative fluctuation. In these cases, we do not plot the negative-yield value, nor do we regenerate smearings for that event. This prevents us from biasing our likelihood function at low signal yields.

3.8 Combining with CLEO II Results

Our ultimate goal is to have branching ratio measurements for the entire CLEO dataset, combining CLEO II and CLEO III. To do this, we have reconstructed the CLEO II likelihood shapes [25], [26]. Some of the likelihood plots are two dimensional, so we project the likelihood function onto one axis to get a likelihood shape for a given mode. These likelihood shapes are in terms of raw yield, so we divide by the efficiency and $N_{B\bar{B}}$ to get a likelihood function in terms of branching ratios. We perform a similar procedure for CLEO III. Armed with likelihood shapes for CLEO II and CLEO III, we multiply the likelihood shapes bin by bin to get the total CLEO likelihood distribution. From this we measure the statistical error for the entire CLEO data sample.

Next, we smear the CLEO II likelihood shape with its additive and multiplicative systematic errors in the same way as described in section 3.7. We then multiply the likelihood shapes for CLEO II and CLEO III, each smeared with its respective systematic errors. The resulting likelihood shape gives our best estimate of the mean, as well as the total error. Again, we subtract in quadrature the statistical error from the total error to calculate the total CLEO systematic error.

3.8.1 $-2\ln\mathcal{L}$ and Likelihood Distributions

For those modes where there is a significant amount of signal, we show the $-2\ln\mathcal{L}$ distribution before and after systematic errors have been included. For the modes where we do not see a significant signal, we show the likelihood distribution after systematic error smearing, with the shaded region indicating the 90% confidence level upper limit integration region. Any peaking in these likelihood distributions is an artifact of the smearing process, and should not be interpreted as signal (with the exception of $\pi^0\pi^0$, which has a marginal yield). Any bias will be an upward bias, and will lead to more conservative upper limits. The $-2\ln\mathcal{L}$ curves for the $\pi\pi$ and $K\pi$ modes were fit with a sixth order polynomial, and the parameters for those fits are given in Table 3.15 and 3.16.

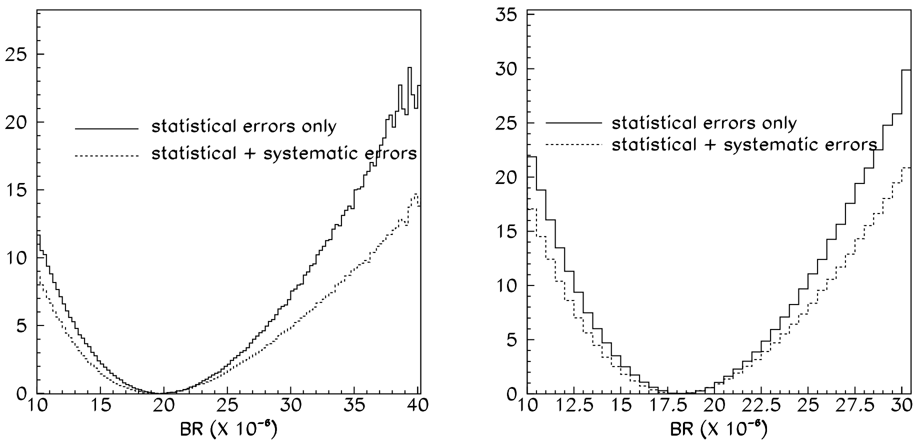

Figure 3.29: $-2\ln\mathcal{L}$ distribution for $K^+\pi^-$ for CLEO III and for CLEO II and CLEO III combined.

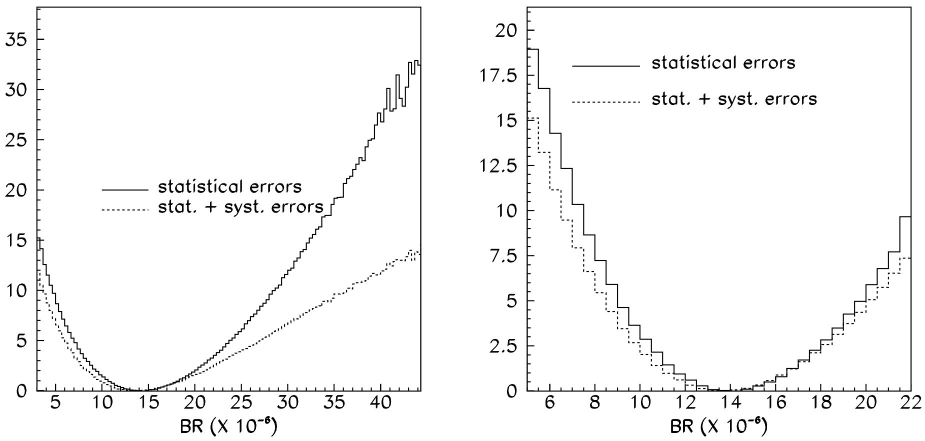

Figure 3.30: $-2\ln\mathcal{L}$ distribution for $K^+\pi^0$ for CLEO III and for CLEO II and CLEO III combined.

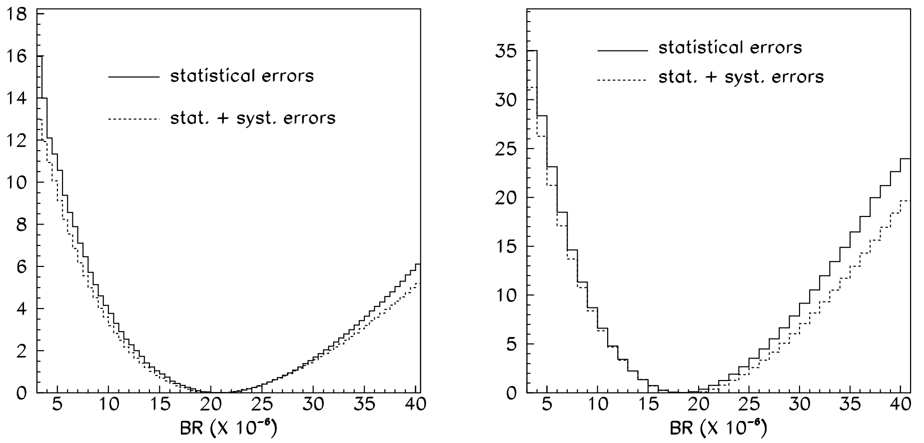

Figure 3.31: $-2\ln\mathcal{L}$ distribution for $K^0\pi^+$ for CLEO III and for CLEO II and CLEO III combined.

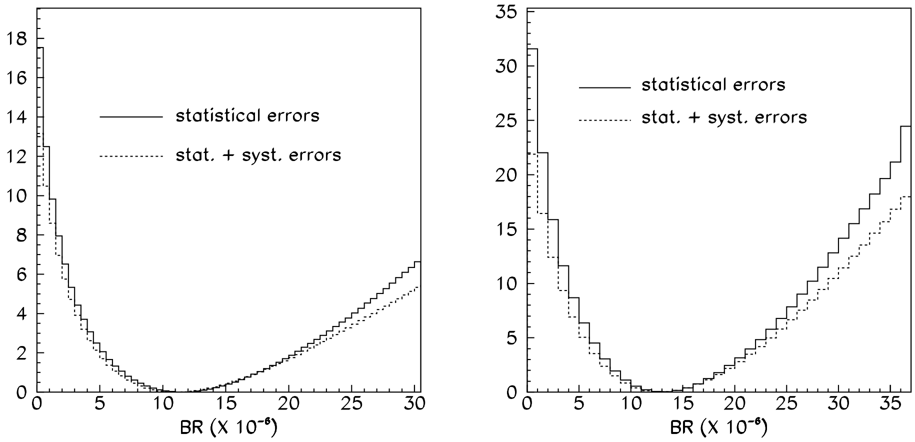

Figure 3.32: $-2\ln\mathcal{L}$ distribution for $K^0\pi^0$ for CLEO III and for CLEO II and CLEO III combined.

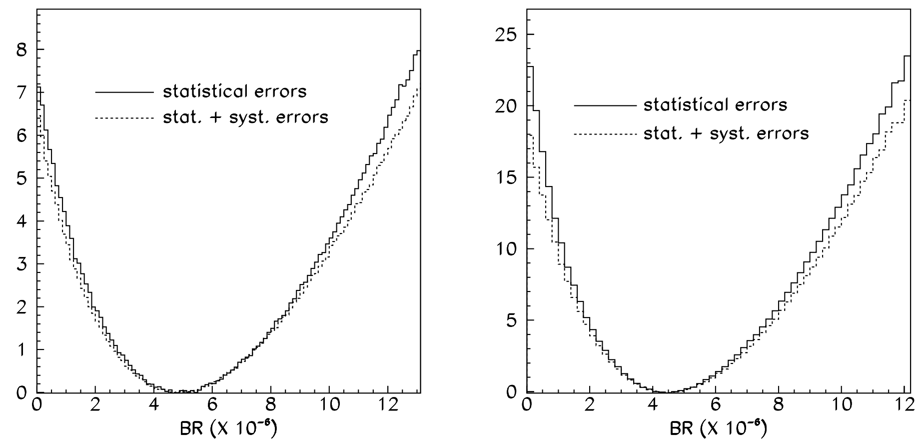

Figure 3.33: $-2\ln\mathcal{L}$ distribution for $\pi^+\pi^-$ for CLEO III and for CLEO II and CLEO III combined.

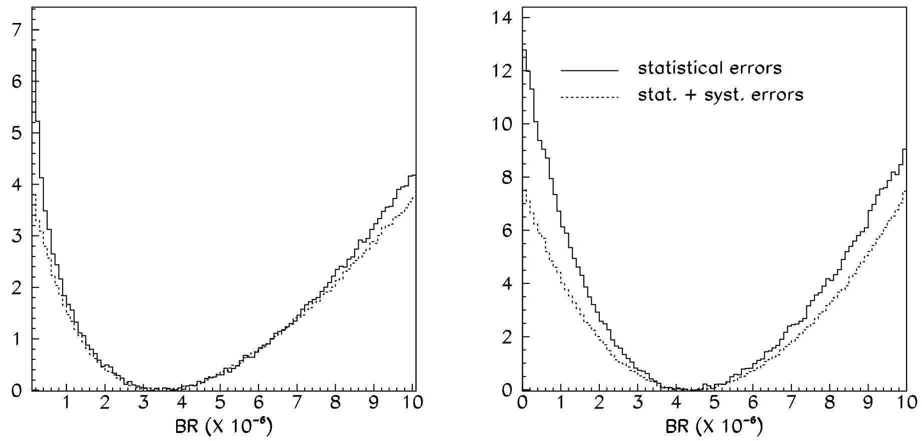

Figure 3.34: $-2\ln\mathcal{L}$ distribution for $\pi^+\pi^0$ for CLEO III and for CLEO II and CLEO III combined.

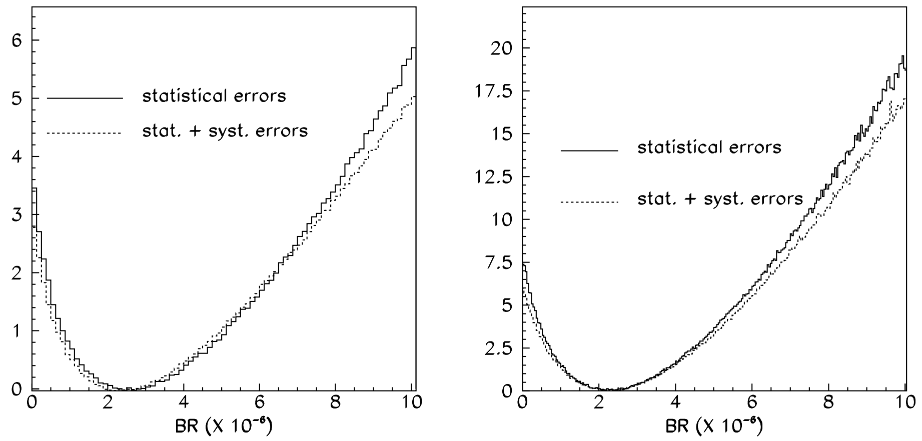

Figure 3.35: $-2\ln\mathcal{L}$ distribution for $\pi^0\pi^0$ for CLEO III and for CLEO II and CLEO III combined.

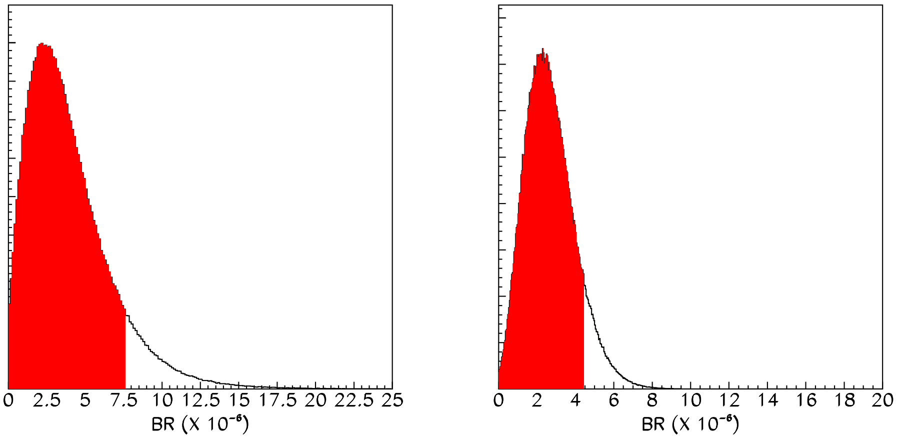

Figure 3.36: Likelihood distribution for $\pi^0\pi^0$ for CLEO III and for CLEO II and CLEO III combined. Shaded region is the 90% C.L. upper limit integration.

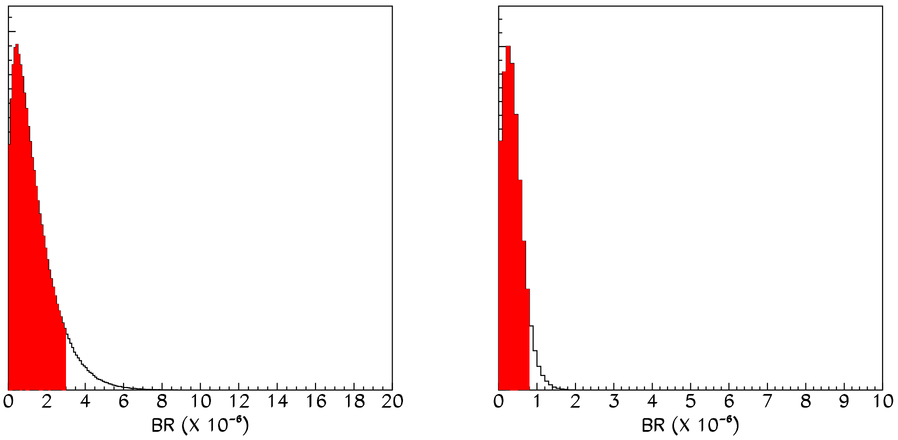

Figure 3.37: Likelihood distribution for K^+K^- for CLEO III and for CLEO II plus CLEO III. Shaded region is the 90% C.L. upper limit integration.

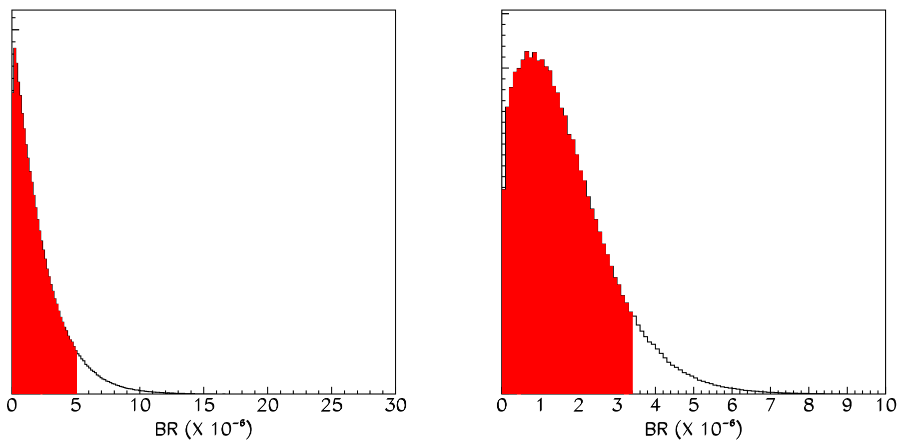

Figure 3.38: Likelihood distribution for K^0K^+ for CLEO III and for CLEO II plus CLEO III. Shaded region is the 90% C.L. upper limit integration.

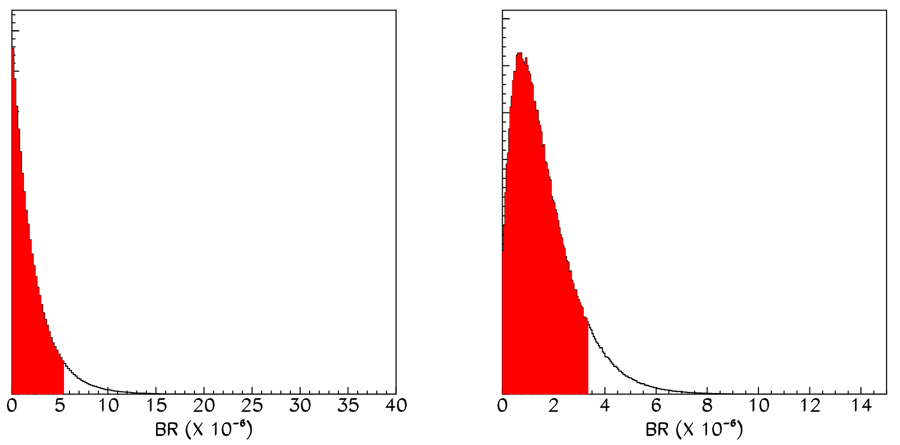

Figure 3.39: Likelihood distribution for K^0K^0 for CLEO III and for CLEO II plus CLEO III. Shaded region is the 90% C.L. upper limit integration.

Table 3.15: Parameters for fits of $-2\ln\mathcal{L}$ for $K\pi$ modes.

	$K^+\pi^0$	$K^0\pi^0$	$K^0\pi^+$	$K^+\pi^-$
x^0	65.804	23.649	57.427	158.98
x^1	-19.003	-5.6752	-9.5346	-27.207
x^2	2.5692	0.55734	0.65651	1.9103
x^3	-0.21491	-0.30791E-01	-0.26219E-01	-0.76542E-01
x^4	0.11134E-01	0.10587E-02	0.67383E-03	0.19660E-02
x^5	-0.31364E-03	-0.19847E-04	-0.97622E-05	-0.28677E-04
x^6	0.36586E-05	0.15360E-06	0.59891E-07	0.17692E-06

Table 3.16: Parameters for fits of $-2\ln\mathcal{L}$ for $\pi\pi$ modes.

	$\pi^0\pi^0$	$\pi^+\pi^0$	$\pi^+\pi^-$
x^0	5.9296	11.708	18.821
x^1	-7.1573	-8.5274	-11.212
x^2	3.1996	2.9179	2.5696
x^3	-0.71879	-063172.	-0.33428
x^4	0.10121	0.85308E-01	0.31620E-01
x^5	-0.76540E-02	-0.61164E-02	-0.17484E-02
x^6	0.23519E-03	0.17902E-03	0.41071E-04

CHAPTER 4

SUMMARY AND DISCUSSION

Here we summarize the results of our analysis, and examine the potential to constrain the (ρ, η) plane using rare B decay results.

4.1 Experimental Results

We have observed all four $K\pi$ decays in the CLEO III data, and with the addition of the CLEO II data, also observe the $\pi^+\pi^-$ channel. We also see evidence (3.6σ) for the $\pi^+\pi^0$ channel. There is only marginal evidence of the $\pi^0\pi^0$ decay, and there is no indication of signal in any of the three KK modes. The results for CLEO III and CLEO II and CLEO III combined are shown in Table 4.1. There are currently measurements of rare B modes from BaBar [27][28][29][30] and Belle [31]. The most recent published results are shown in Table 4.2.

We can compare the CLEO results to the combined results of BaBar and Belle, as shown in Figure 4.1.

We combine the CLEO branching ratios with the branching ratios in Table 4.2. All measurements are assigned a bifurcated gaussian likelihood function, then multiplied together to give the combined likelihood function. Mean values and asym-

Table 4.1: Experimental results for CLEO III and CLEO II and CLEO III combined. Significances are statistical significance, and significance with systematic errors included. Upper limits are 90% confidence level.

Mode	CLEO III Significance(σ)	CLEO III BR($\times 10^{-6}$)	CLEO II+III Significance(σ)	CLEO II+III BR($\times 10^{-6}$)
$\pi^+\pi^-$	2.7, 2.6	$4.8^{+2.5+0.8}_{-2.2-0.5}$	5.0, 4.4	$4.5^{+1.4+0.5}_{-1.2-0.4}$
$\pi^+\pi^0$	2.6, 2.1	$3.4^{+2.8+0.8}_{-2.0-0.3}$	4.0, 3.5	$4.6^{+1.8+0.6}_{-1.6-0.7}$
$\pi^0\pi^0$	2.0, 1.8	(< 7.6)	2.8, 2.5	(< 4.4)
$K^+\pi^-$	> 7	$19.5^{+3.5+2.5}_{-3.2-1.6}$	> 7	$18.0^{+2.3+1.2}_{-2.1-0.9}$
$K^0\pi^+$	5.2, 4.6	$20.5^{+7.1+3.0}_{-5.9-2.1}$	> 7	$18.8^{+3.7+2.1}_{-3.3-1.8}$
$K^+\pi^0$	6.1, 5.0	$13.5^{+4.0+2.4}_{-3.5-1.5}$	> 7	$12.9^{+2.4+1.2}_{-2.2-1.1}$
$K^0\pi^0$	4.6, 3.8	$11.0^{+6.1}_{-4.6} \pm 2.5$	6.2, 5.0	$12.8^{+4.0+1.7}_{-3.3-1.4}$
K^+K^-	-	(< 3.0)	-	(< 0.8)
K^+K^0	-	(< 5.0)	-	(< 3.3)
K^0K^0	-	(< 5.2)	-	(< 3.3)

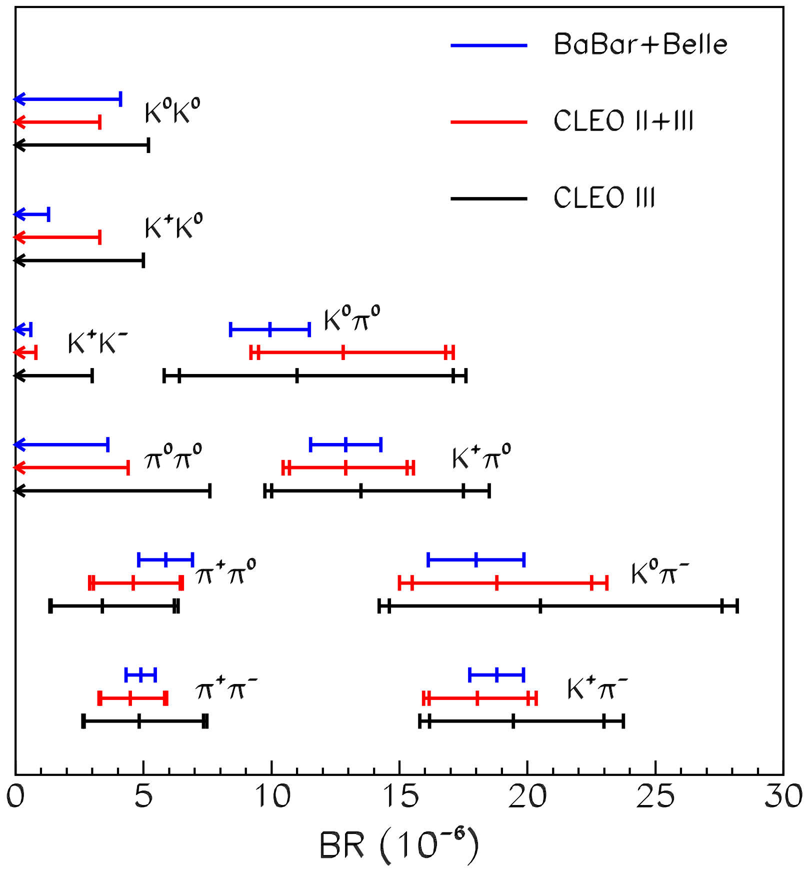

Figure 4.1: Comparison of the CLEO branching ratio numbers and the average of BaBar and Belle for the rare meson modes.

Table 4.2: Most recent branching ratios (upper limits)[10^{-6}] of the 10 $B \to PP$ modes, with statistical and systematic errors.

Mode	BaBar	Belle
$\pi^+\pi^-$	$4.7 \pm 0.6 \pm 0.2$	$5.4 \pm 1.2 \pm 0.5$
$\pi^+\pi^0$	$5.5^{+1.0}_{-0.9} \pm 0.6$	$7.4^{+2.3}_{-2.2} \pm 0.9$
$\pi^0\pi^0$	(< 3.6)	(< 6.4)
$K^+\pi^-$	$17.9 \pm 0.9 \pm 0.7$	$22.5 \pm 1.9 \pm 1.8$
$K^0\pi^-$	$17.5^{+1.8}_{-1.7} \pm 1.3$	$19.4^{+3.1}_{-3.0} \pm 1.6$
$K^+\pi^0$	$12.8^{+1.2}_{-1.1} \pm 1.0$	$13.0^{+2.5}_{-2.4} \pm 1.3$
$K^0\pi^0$	$10.4 \pm 1.5 \pm 0.8$	$8.0^{+3.3}_{-3.1} \pm 1.6$
K^+K^-	(< 0.6)	(< 0.9)
K^+K^0	(< 1.3)	(< 2.0)
K^0K^0	-	(< 4.1)

metric errors are found by fitting the combined likelihood distribution. All errors are assumed to be uncorrelated. For measurements where only an upper limit is shown, we have used the signal yields and efficiencies to calculate a branching fraction even if the signal is not statistically significant. The combined branching ratios are shown in Table 4.3.

Table 4.3: Results of combining the CLEO, BaBar and Belle branching ratio measurements.

Mode	Combined BR (10^{-6})
$\pi^+\pi^-$	4.8±0.5
$\pi^+\pi^0$	5.6±0.9
$\pi^0\pi^0$	$2.0^{+0.7}_{-0.6}$
$K^+\pi^-$	18.5±1.0
$K^+\pi^0$	12.9±1.2
$K^0\pi^-$	$18.2^{+1.7}_{-1.6}$
$K^0\pi^0$	10.4±1.4

4.2 Measuring the CKM Parameters

Several interesting quantities can be calculated using the measured branching ratios. These are summarized in Table 4.4. We use the values $\frac{\tau_{B+}}{\tau_{B^0}} = 1.083 \pm 0.017$, $f_\pi = 131$ MeV, $f_K = 160$ MeV, $|\frac{V_{us}}{V_{ud}}| = 0.23$ and $\tan\theta_C = 0.22$.

From Eqs. 1.25 and 1.26, we can attempt to bound $\cos\gamma$. The current experimental value of R' (which accounts for phase space factors) is 1.11±0.12. Without any information about the phase δ or the relative tree to penguin strength, we cannot rule out any value of γ. It should be noted that when this method was first proposed, the value of R' was 0.69 (with large errors), so that it seemed more likely that γ could be constrained with this method.

We can use Eq. 1.33 to give allowed regions in the (ρ, η) plane for a range of

Table 4.4: Quantities derived from branching ratio measurements. Sides a-c refer to the triangle shown in Figure 1.6.

Variable	Definition	CLEO value	World average
R'	$\frac{B(B\to\pi^{\pm}K^{\mp})}{B(B^{\pm}\to\pi^{\pm}K^{0})}\frac{\tau_{B^+}}{\tau_{B^0}}$	1.04±0.26	1.10±0.11
R_*^{-1}	$\frac{2B(B^{\pm}\to K^{\pm}\pi^{0})}{B(B^{\pm}\to K^{0}\pi^{\pm})}$	1.37±0.40	1.42±0.18
ϵ_{exp}	$\tan\theta_C\frac{f_K}{f_\pi}\left[\frac{2B(B^{\pm}\to\pi^{\pm}\pi^{0})}{B(B^{\pm}\to K^{0}\pi^{\pm})}\right]^{1/2}$	0.19±0.04	0.21±0.02
Side a	$\sqrt{B(B^{\pm}\to K^{0}\pi^{\pm})}$	4.34±0.46	4.27±0.19
Side b	$\frac{f_K}{f_\pi}\frac{V_{us}}{V_{ud}}\sqrt{2B(B^{\pm}\to\pi^{0}\pi^{\pm})}$	0.85±0.17	0.90±0.07
Side c	$\sqrt{2B(B^{\pm}\to K^{\pm}\pi^{0})}$	5.08±0.51	5.08±0.24

values of R_*^{-1}. Figure 4.2 shows the allowed regions for the central value and for the 68% and 95% confidence level bands. For the central value of R_*^{-1}, we scan all unknown parameters over what is thought to be a generous range of values as suggested by [10], namely $0 < \delta < 0.05$, $0.197 < \hat{q} < 0.247$, $0.93 < R_{th} < 1.03$, and $0.8 < \cos\phi < 1.0$. We plot any value of ρ and η that satisfies Eq. 1.33 within these variations of parameters. To find the 68% and 95% confidence bands, we also vary R_*^{-1} over its 68% and 95% confidence level range respectively. The X shown in Figure 4.2 corresponds to the preferred value of (ρ, η) from a fit to the standard CKM inputs (Figure 1.3). There is already substantial disagreement between the preferred region from the standard CKM fit and the region obtained from the rare B decays. It is still too early to say that there is definitive evidence for an inconsistency between the two methods of determining the CKM parameters, but this could be an indication of physics beyond the Standard Model.

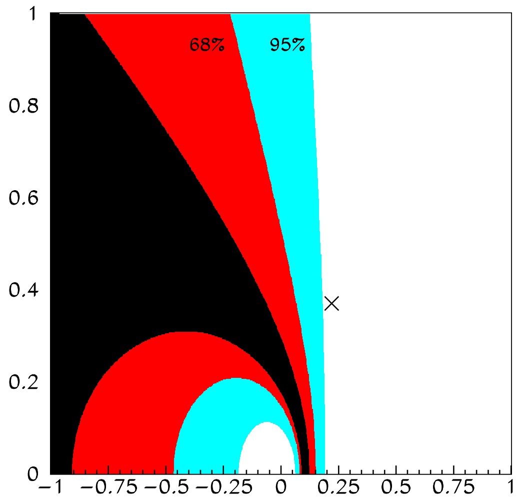

Figure 4.2: Allowed regions of (ρ, η) plane for R_*^{-1}. The darkest band is the allowed region for the central value with the current theoretical errors. The lighter bands are the 68% and 95% confidence level bands. The constraints are symmetric about the $\eta = 0$ axis.

Next we turn to our triangle construction (Figure 1.6). The lengths of the three sides are (in arbitrary units) $a = 4.27 \pm 0.19$, $b = 0.90 \pm 0.07$, and $c = 5.08 \pm 0.24$. We can immediately see that variations within the errors can lead to triangles that do not close (for instance a 1σ upward variation in c gives $5.32 \not< 4.27 + 0.90$). Small CP-asymmetries ($\sim$10%) in the $K^{\pm}\pi^0$ decays would require large asymmetries ($\sim$50%) in the $\pi^{\pm}\pi^0$ decays for the triangles to close. This makes it quite difficult to draw any conclusions using this method, but for illustration, we will look at the case where there is a 2% CP-asymmetry in the $K^{\pm}\pi^0$ decays for the current values of the branching ratios. This scenario is illustrated in Figure 4.3. The CP-asymmetry in the $K^{\pm}\pi^0$ decay makes the sides c and $\bar{c}$ inequal, giving different shapes for the B^+ and B^- amplitude triangles. Since we do not know which reflection about the $K^0\pi^{\pm}$ base of the amplitude triangles to use, we can only determine γ to an eight-fold ambiguity. For this exmaple, $\gamma = \pm 27°$, $\pm 50°$, $\pm 130°$, or $\pm 153°$.

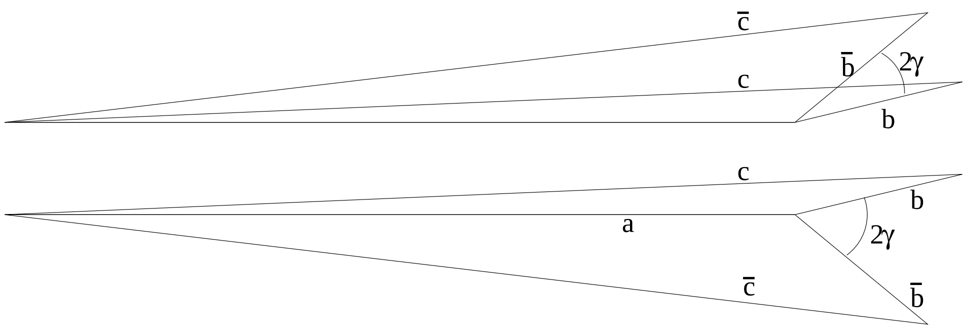

Figure 4.3: Triangle diagrams assuming a 3% CP-asymmetry in $B^{\pm} \to K^{\pm}\pi^0$ decays.

Both BaBar and Belle have reported measurements of the CP-violating parameters $S_{\pi\pi}$ and $C_{\pi\pi}$ in $\pi^+\pi^-$ decays. These parameters are the coefficients of the $\sin(\Delta m_B t)$ and $\cos(\Delta m_B t)$ terms in eq. 1.52, respectively. The values from BaBar are $S_{\pi\pi} = 0.02 \pm 0.34 \pm 0.05$ and $C_{\pi\pi} = -0.30 \pm 0.25 \pm 0.04$, while Belle finds $S_{\pi\pi} = -1.21^{+0.38+0.16}_{-0.27-0.13}$ and $C_{\pi\pi} = -0.94^{+0.25}_{-0.31} \pm 0.09$.

The measurement of these parameters is, in theory, sufficient to determine $\sin 2\alpha$, however there are difficulties in determining α. The source of the difficulty is that there is penguin pollution in the $\pi^+\pi^-$ decays. To disentangle the contribution from penguin decays, we need to construct another amplitude diagram for the $\pi\pi$ decays (Figure 1.8). CP-asymmetries in the $\pi^0\pi^0$ mode could be large, so measurements of both $B^0 \to \pi^0\pi^0$ and $\bar{B}^0 \to \pi^0\pi^0$ are also necessary. Additionally, to determine the ratio of the amplitudes of B^0 and $\bar{B}^0$ decays to $\pi^+\pi^-$ needed for eq. 1.59, we need to know the relative strength and relative strong phases of the tree and penguin diagrams for $\pi^+\pi^-$. Given the large errors and disagreement between the two measurements of $S_{\pi\pi}$ and $C_{\pi\pi}$, and the fact that there is yet to be a statistically significant measurement of the CP-averaged branching ratio $B^0 \to \pi^0\pi^0$, we cannot expect reasonable constraints on $\sin 2\alpha$ at this time.

In conclusion, we have measured the branching ratios of the ten charmless two-body pseudoscalar B meson decays using the CLEO III detector. We have also combined the CLEO II and CLEO III measurements to produce the final CLEO results for these modes. In addition, we compare our results with those of the B factories, and comment on our current ability to constrain the (ρ, η) plane using the global averages of these quantities. With improved experimental accuracies, we find that these measurements are on the verge of providing unique bounds on the CKM parameters.

REFERENCES

[1] S. Burles et al, Astrophysics Journal **552**, L1 (2001).

[2] Chris Quigg, *Gauge Theories of the Strong, Weak, and Electromagnetic Interactions*, Addison Wesley, 1983.

[3] Y. Nir, in *Heavy Flavour Physics, Theory and Experimental Results in Heavy Quark Physics*, edited by C. Davies and S. Playfer, 147 (2002).

[4] G. Farrar, M. Shaposhnikov, Phys. Rev. D **50**, 774 (1994).

[5] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963).

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

[7] L. Wolfenstein, Phys. Rev. Lett. **55**, 1039 (1983).

[8] A. Höcker *et al.*, *Proceedings of the 9th International Symposium on Heavy Flavour Physics*, September 2001; hep-ph/0112295.

[9] M. Gronau, O. Hernandez, D. London and J. Rosner, Phys. Rev. D **52**, 6374 (1995).

[10] M. Beneke, G. Buchalla, M. Neubert, and C.T. Sachrajda, Nucl.Phys. B **606**, 245 (2001); hep-ph/0104110.

[11] M. Beneke, G. Buchalla, M. Neubert, and C.T. Sachrajda, Phys. Rev. Lett. **83**, 1914 (1999).

[12] R. Fleischer and T. Mannel, Phys. Rev. D **57**, 2752 (1998).

[13] M. Gronau, O. Hernandez, D. London and J. Rosner, Phys. Rev. Lett. **73**, 21 (1994).

[14] R.A. Fisher, Annals of Eugenics, **7**, 179 (1936).

[15] T. Pedlar *et al.*, *Measurement of* $\mathcal{B}(B^- \to D^0\pi^-)$ *and* $\mathcal{B}(B^0 \to D^+\pi^-)$ *and Isospin Analysis of* $B \to D\pi$ *Decays*, CBX 02-15.

[16] P. Avery, *Combining Measurements with Correlated Errors*, CBX 95-55.

[17] W. Ford , *Poisson Likelihood Fits and Rare B Branching Fraction Measurements*, CBX 98-6.

[18] R. Barlow, *Statistics*, Wiley (1989).

[19] H. Albrecht *et al.*, ARGUS Collaboration, Phys. Lett. B**241** 278 (1990); **254** 288 (1991).

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

[21] T. Bergfeld and M. Selen, *A Measurement of* $\mathcal{B}(D^0 \to K^-\pi^+\pi^0)/\mathcal{B}(D^0 \to K^-\pi^+)$, CBX 95-59.

[22] Particle Data Group, Phys. Rev. D **66**, 010001 (2002).

[23] T. Bergfeld and M. Selen, *Dalitz Analysis of the Decay* $D^0 \to K^-\pi^+\pi^0$, CBX 00-62.

[24] A. Gritsan, $\Gamma(\eta \to 3\pi^0)/\Gamma(\eta \to 2\gamma)$ *and* $\eta/\pi^0/\gamma$ *Systematics*, CBX 99-72, .

[25] D. Cronin-Hennessy *et al.*, CLEO Collaboration, Phys. Rev. Lett. **85**, 515 (2000), D. Asner *et al.*, CLEO Collaboration, Cornell Report No. CLNS 01/1718, CLEO 01-02 (2001).

[26] D. M. Asner *et al.*, CLEO Collaboration, Phys. Rev. D **65**, 031103 (2002).

[27] B. Aubert *et al.*, BABAR Collaboration, hep-ex/0207055.

[28] B. Aubert *et al.*, BABAR Collaboration, hep-ex/0207063.

[29] B. Aubert *et al.*, BABAR Collaboration, hep-ex/0207065.

[30] B. Aubert *et al.*, BABAR Collaboration, hep-ex/0206053.

[31] B. Casey *et al.*, Belle Collaboration, Phys. Rev. D **66**, 092002 (2002); hep-ex/0207090.