FLAVOR MIXING AND MIXING-INDUCED CP VIOLATION IN THE NEUTRAL B SYSTEM

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

Andrew Dean Foland

January 2000

FLAVOR MIXING AND MIXING-INDUCED CP VIOLATION IN THE NEUTRAL B SYSTEM

Andrew Dean Foland, Ph.D.

Cornell University 2000

We describe a new technique for measurements of mixing-induced CP violation in the B_d system, wherein the decay time of only one B meson is measured. We find that excellent decay-point measurement resolution, high-efficiency meson flavor tagging, and precise knowledge of the $B^0-\overline{B^0}$ mixing parameters are required to carry out the experiment. We report on each of these aspects using the data collected by the CLEO experiment at the Cornell Electron Storage Ring. We present methods for assessing the performance and systematic uncertainties of a silicon vertex detector. We present assessments of the detector installed in the CLEO experiment. We present a method for high-efficiency B flavor tagging, and a measurement of its power. Using hadronic decays based on a method introduced by Nelson, Nelson, and Witherell, we present a new measurement of $B^0-\overline{B}^0$ mixing. This method has much reduced systematic uncertainties compared to previous measurements.

BIOGRAPHICAL SKETCH

Andrew Foland was born at the corner of Peach Street and Peach Avenue in Atlanta, Georgia on September 28, 1972. He attended elementary school in Greenville, Mississippi, before moving to Champaign, Illinois. This is where Andrew grew up to become a Midwesterner. He attended high school at University High School in Urbana, Illinois, and graduated in 1989. He received his *B.S.* in physics at the University of Illinois in Urbana-Champaign, graduating *magna cum laude* in May, 1994, after two years of exciting work with Professor Gary Gladding. He pursued his interest in high-energy experimental physics at Cornell University's Laboratory of Nuclear Studies as a National Science Foundation Graduate Fellow in the fall of 1994. He received an *M.S.* in physics from Cornell University in May, 1997.

For my parents, Barbara Hudgings and Ed Foland

ACKNOWLEDGEMENTS

I have been on the CLEO experiment for just over five years, yet some of the data used in this thesis is itself nearly ten years old. Some portions of the detectors used in the thesis are older still. Most of the software I used to analyze the data predates me. And even while I was here, by far the vast majority of the work which made this thesis possible was not carried out by me. This humbling litany of dependence on the work of others is itself a tribute to the collegial spirit to be found among the scholars who call themselves high energy physicists, especially those on the CLEO collaboration. Through their detectors, their data, and their software, as well as their constant insight, enthusiasm, and helpfulness, this thesis has been made possible. I would like to thank these collaborators, not only the 222 who appear on CLEO author lists, but all those who have been on the author lists in the past. A few I would especially like to thank in this capacity are Dan Peterson, Bernie Gittelman, Giancarlo Moneti, Ed Thorndike, Boyce McDaniels, Tomasz Skwarnicki, David Cassel, and Brian Heltsley. I would also like to thank the CESR staff, who puts soup in our bowls, and the LNS technical staff, who make the bowls leakproof.

It has been a pleasure to have as an advisor someone who knows where you can get a rubber chicken in Ithaca. Persis Drell has given me free reign when I needed it without losing track of me. She has provided physics direction and insight in this and many other projects. She has also provided an inspiring example of how to lead a life in physics. I hope to be like Persis when I grow up. I would also like to thank her for her helpful comments on this and numerous other manuscripts.

There are a number of people with whom I shared tasks in the effort of installing and commissioning the CLEO II.V silicon detector. All of them were dedicated, hard-working, and clever people, from whom I both learned a great deal and with whom I had a lot of fun. I would like to thank David Asner, Jesse Ernst, Tony Hill, David Jaffe, Mark Palmer, Soeren Prell, and Bill Ross. And I especially want to pay homage to "The Master", Craig Prescott. Craig has rendered extraordinary service to the collaboration, putting in unbelievable hours, serving us with his extensive knowledge, all the while being helpful and cheerful. That any of our papers make it to publication is a testament to the work Craig has done. For this, every CLEOn owes Craig a beer. May the BCS repay our debt to the University of Florida!

I also had a lot of fun in analysing the CLEO data, often in concert with others. I found that production goes up faster than linearly, and drudgery decreases faster still, when working with Alexey Ershov, Dan Bliss, Pablo Hopman, and Tony Rubiera.

I would like to thank Lawrence Gibbons for his suggestions and useful comments on the single-time measurement technique described in chapter 1, Frank Wuerthwein for his suggestion concerning the $\sin 2\alpha$ measurement, and Alexey Ershov for providing invaluable aid in sensitivity estimation.

The original version of the tagger was written after inspiring discussions with

Pablo Hopman, who has evinced interest and encouragement along the way. I would like to thank Alexey Ershov for his input and excitement along the way, and the high work standard he set while studying the feasibility of a measurement of $\sin 2\beta$.

I'd like to thank David Cinabro, Hitoshi Yamamoto, and Jeff Cherwinka for their help with beampipe dimensions. David Kreinick provided help understanding properties of beam-wall events. Thanks also to the members of the Tracking Group for their helpful suggestions and general encouragement.

I cannot say enough good things about my officemate, Pablo Hopman, but that does not excuse an attempt. Pablo performs inspiring feats of detector-saving. Without exaggeration I can say he taught me everything I know about silicon detector hardware. He has always provided a thoughtful point of view on lab happenings. He has also performed excellent analyses, and always evinces enthusiasm for the analyses I am working on. I should also thank my other officemates during my tenure, Andreas Warburton and Abi Soffer, with whom I did not have the good fortune to spend as much time, for their help on topics from particle ID to CDF software.

I would also like to thank a several CLEOns with whom I became close friends. Dan Bliss and I giggled our way through more late-night CBX revisions than I can adequately recall. Dan was also a great friend and advisor at many stages in my life; I was lucky he left his job at General Dynamics to work at CLEO for a few years. Ken Bloom inspired me to write a skit scene which included the line "Phonons are like the foot cuffs, and the warden is quantum mechanics". Ken also provided thoughtful comment and advice on subjects ranging from burrito restaurants to presidential elections; the absence after his graduation was never filled. I wish the best to Veronique Boisvert, with whom I shared some special times.

The average American household annually contributes a considerable sum of money to high-energy physics research. This research has been made possible by their generosity and curiosity, and it is our duty to handle their contribution responsibly. I would like to thank the taxpayers of the United States for making this research possible.

Finally, I must thank my very close friends and family. Joel Ullom provided sailing lessons and laser tag opponents in Boston. Bao Wen has given me essential love and support through the years, whether I deserved it or not. I was lucky Marc didn't want to leave until Tuesday. My sister provided a smile and a haven in Happy Valley, not to mention the occasional pecan pie. My parents, to whom I dedicate this thesis, instilled in me at an early age a combined love both of learning and of the real world we live in.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION TO THE NEUTRAL B SYSTEM

1.1 Introduction

1.1.1 Who is this thesis for?

This thesis is not written to be understood by the general non-physics reader. This is not to antagonize such readers because of bad childhood experiences, but simply because had it been, the thesis would have become useless to the daily practitioner of physics.

This introduction is not even meant to be particularly pedagogical. Instead, it is to remind the reader of the important points, and to define the notation used in the rest of the thesis.

1.1.2 Offer

As an inheritor of the ancient University tradition, it might be expected that I write a thesis of general use to scholars throughout the University, whether they study physics or something else. Therefore, to scholars (and others) in disciplines other than my own, I make the following offer: I will be happy to spend an hour in a cafe explaining the work in my thesis to you. I will even pay for your coffee.

1.1.3 Overview of the Thesis

This thesis focuses on exploration of the phenomenon of CP violation. As physics is most compactly and deeply understood in terms of the symmetries its laws respect, it is compelling to study the whys and hows of how symmetries come to be broken. Symmetry under spatial reflection, known as "P" (for "parity") symmetry, is expected by physical intuition. Nonetheless it is violated almost 100% by the weak interactions [1]. Symmetry of physical laws under the exchange of "+" and "-" labels of electrical charges, known as "C" (for "charge") symmetry, is also expected on reasonable intuitive grounds, but is also not respected by the weak interactions. The combined operation, "CP", which consists of simultaneously exchanging all charge signs and spatially reflecting the system, was expected to be a respected symmetry even though the separate "P" and "C" symmetries were violated. It came as a surprise in 1964 [2] to find that this symmetry, too, is violated.

CP violation is welcomed on cosmological grounds because it paves the theoretical road to our existence. It has been shown that to generate an excess of matter over anti-matter in our universe requires CP violation [3]. In turn, in order to generate creatures such as ourselves, there must not be equal amounts of matter and anti-matter, lest the universe become a sterile cosmos full of photons.

CP violation is generally studied in the context of a meson system consisting of a flavored, neutral meson and its anti-meson. CP violation is small, but most well studied, in the $K^0 - \overline{K^0}$ system, is expected to be very small in the $D^0 - \overline{D^0}$ system, and is expected to be $\mathcal{O}(1)$ in the $B^0 - \overline{B^0}$ system.

Such flavored neutral mesons may spontaneously transform into their own anti-mesons, in a process known as "mixing". Mixing is generally induced by interactions described by box diagrams, as in figure 1.1.

Like any two state system [4], once transitions between the states, such as mixing, are allowed, the eigenstates of the Hamiltonian no longer coincide with the original states. We write the Hamiltonian of a meson system as

$$\begin{aligned} i\frac{d}{dt}\begin{pmatrix} a \\ b \end{pmatrix} &= (M - i\frac{\Gamma}{2})\begin{pmatrix} a \\ b \end{pmatrix} \\ &= \begin{pmatrix} M_{11} - i\frac{\Gamma_{11}}{2} & M_{12} - i\frac{\Gamma_{12}}{2} \\ M_{12}^* - i\frac{\Gamma_{12}^*}{2} & M_{11} - i\frac{\Gamma_{11}}{2} \end{pmatrix}, \end{aligned} \tag{1.1}$$

where a and b standard for amplitudes to be a meson or anti-meson, respectively. Note that this Hamiltonian is not hermitian. That is not required, as the mesons can decay.

We can solve for the eigenstates of this equation, which we write as:

$$|B_+^0\rangle = p|B^0\rangle + q|\overline{B^0}\rangle \tag{1.2}$$

$$|B_-^0\rangle = p|B^0\rangle - q|\overline{B^0}\rangle \tag{1.3}$$

where p and q can be solved for in terms of M and Γ. Defining r and κ by

$$re^{i\kappa} = \sqrt{\frac{M^*_{12} - i\frac{\Gamma^2_{12}}{2}}{M^*_{12} - i\frac{\Gamma^2_{12}}{2}}} \tag{1.4}$$

we find

$$p = \frac{2}{(1+re^{i\kappa})\sqrt{2(1+\frac{1-q}{1+q}^2)}} \tag{1.5}$$

$$q = \frac{2re^{i\kappa}}{(1+re^{i\kappa})\sqrt{2(1+\frac{1-q}{1+q}^2)}}. \tag{1.6}$$

Notice that only if either M_{12} or Γ_{12} has a complex phase can $re^{i\kappa}$ be different from 1. When $re^{i\kappa} = 1$ both p and q are 1.

A common approximation for M_{12} and Γ_{12} nearly real (r near 1, κ small) is to define a small parameter ϵ

$$\epsilon = \frac{1-re^{i\kappa}}{1+re^{i\kappa}} \tag{1.7}$$

$$p = \frac{1+\epsilon}{\sqrt{1+\epsilon^2}} \tag{1.8}$$

$$q = \frac{1-\epsilon}{\sqrt{1+\epsilon^2}} \tag{1.9}$$

The eigenvalues of the matrix may also be solved for, giving

$$M_{1,2} = M \pm \Re(\sqrt{(M_{12} - i\frac{\Gamma_{12}}{2})(M^*_{12} - i\frac{\Gamma^*_{12}}{2})}) \tag{1.10}$$

$$\Gamma_{1,2} = \Gamma \mp 2\Im(\sqrt{(M_{12} - i\frac{\Gamma_{12}}{2})(M^*_{12} - i\frac{\Gamma^*_{12}}{2})}), \tag{1.11}$$

and leading to

$$\Delta M = \Re(\sqrt{(M_{12} - i\frac{\Gamma_{12}}{2})(M_{12}^* - i\frac{\Gamma_{12}^*}{2})}) \quad (1.12)$$

$$\Delta\Gamma = -4\Im(\sqrt{(M_{12} - i\frac{\Gamma_{12}}{2})(M_{12}^* - i\frac{\Gamma_{12}^*}{2})}) \quad (1.13)$$

Note that whether B_+ or B_- is the heavier state depends on the sign of $\Re(\sqrt{(M_{12} - i\frac{\Gamma_{12}}{2})(M_{12}^* - i\frac{\Gamma_{12}^*}{2})})$, which is an experimental question.

There are three physically distinct ways in which CP violation can manifest itself in a two-state neutral meson system.

- "Direct" CP violation
- "Indirect" CP violation
- "Mixing-induced" CP violation

The longest-standing, and until recently, only observed form of CP violation was *indirect* CP violation. It is also occasionally known as "CP violation in the mass matrix". [1] This kind of CP violation occurs when the mass matrix given in equation 1.1 is CP violating in that $\frac{|p|}{|q|} \neq 1$; that is, the eigenstates themselves each contain unequal mixtures of CP + and CP - states. This is the kind of CP violation that is primarily responsible for the observed CP violation in the $K^0 - \overline{K^0}$ system.

Recently, *direct* CP violation has been observed in the $K^0 - \overline{K^0}$ system [5]. This form of CP violation arises when the decay matrix elements do not observe CP symmetry; that is,

[1] Just to confuse things, some authors even refer to *this* kind of CP violation as mixing-induced; or refer to mixing-induced CP violation as indirect.

$$|A(M \to f)| \neq |\overline{A}(\overline{M} \to \overline{f})| \tag{1.14}$$

Though observed in the $K^0 - \overline{K^0}$ system, this kind of CP violation is not limited to neutral meson systems. For instance, a difference in the decay rates $B^+ \to \pi^+\pi^0$ and $B^- \to \pi^-\pi^0$ would be a signal for direct CP violation. This kind of asymmetry can arise only when both a CP-violating weak phase and CP-conserving strong phase are present in the decay amplitude.

Finally, CP violation can arise because the matrix element for mixing has an imaginary phase. Note that it can arise even if $|\frac{p}{q}| = 1$, so that it is distinct from indirect CP violation. This "mixing-induced" CP violation has never been observed, but may be large in B^0 decays. In decays to CP eigenstates, the contributions from the meson and antimeson enter with opposite phases, where the relative amplitudes for the decay from meson and antimeson are governed by the flavor-mixing oscillations. Uniquely in the B system, both the phase (β) and mixing are expected to be large, and therefore mixing-induced CP violation is expected to be $\mathcal{O}(1)$.

This thesis focuses on mixing-induced CP violation in the B_d system. I present two main areas of inquiry: how might one measure such CP violation and which experimental techniques must be developed; and what is the magnitude of flavor mixing in the B_d system?

In the remainder of this introductory chapter, I will *very* briefly introduce theoretical expectations for the CKM matrix, where the large magnitude and phase of mixing originate. After that, I will review the formalism for measurements in the neutral B system, and particularly for measurements of mixing-induced CP violation.

As the reader will learn, there are three necessary experimental requirements in order to measure mixing-induced CP violation:

- Excellent vertex position resolution (i.e. at what time did the B decay?)
- High-efficiency B flavor tagging (i.e. was it a B or a $\overline{B}$?)
- High-efficiency reconstruction of CP eigenstates

In addition, the magnitude of mixing in the B system must also be known.

High-precision position information is required for the measurement. To acheive such precision, future B-factories plan to employ silicon vertex detectors. In the second chapter, I report on the efficiency, resolution, and systematic reliability of the silicon vertex detector that has been operating for over two years in the CLEO II.V detector.

In the third chapter, I report on a method for high-efficiency tagging of B flavor. This method makes use primarily of leptons, kaons, and uses several other pieces of information. The method uses a neural net to provide the optimal separation between B and $\overline{B}$. I report on construction, tests, and data performance of the tagging method.

In the fourth chapter, I describe a new measurement of the magnitude of $B-\overline{B}$ mixing, taking advantage of a partial hadronic reconstruction method described by

Nelson, Nelson, and Witherell [6]. This method has smaller and better understood systematic uncertainties compared to previous measurements.

In the final chapter, I give the results of the measurement and discuss the implications for measurements of $\sin 2\beta$, limits on CKM elements, and the measurement of V_{td}.

1.2 The CKM Matrix

As far as we know experimentally, all interactions in the physical world aside from gravity can be described by the so-called Standard Model. This is not to say we know it to be fully correct, merely that we have not yet found any experiments inconsistent with it. At higher energies, the Standard Model predicts nonsense (interaction probabilities greater than 1) and we therefore know that experiments at these higher energies will find failures in the Standard Model. We also know that the Standard Model fails to describe gravity, and may not even provide *enough* CP violation to allow our existence.

Nonetheless, the Standard Model describes all experimental data gathered to date, and guides us in expectations for future experiments. Though the full Standard Model Lagrangian of course contains many terms (gauge bosons, gauge interactions, etc.) we focus in this section on a single term of the Lagrangian: the term describing interactions of quarks and the charged weak current.

1.2.1 The Flavor Sector of the Standard Model

In the standard model, quarks are arranged into SU(2) doublets of weak isospin according to

$$\begin{pmatrix} u \\ d \end{pmatrix}_L \begin{pmatrix} c \\ s \end{pmatrix}_L \begin{pmatrix} t \\ b \end{pmatrix}_L \tag{1.15}$$

where the subscript "L" denotes the left-handed helicity component of the relevant Dirac spinor. Here the flavors are designated according to their production by the strong force. When we want to refer to any of the doublets we use the notation

$$\begin{pmatrix} u \\ d \end{pmatrix}_L^i, \tag{1.16}$$

where i=1,2,3 corresponds to $\begin{pmatrix} u \\ d \end{pmatrix}_L \begin{pmatrix} c \\ s \end{pmatrix}_L$ and $\begin{pmatrix} t \\ b \end{pmatrix}_L$ doublets, respectively. We take ψ_i to stand for a left-handed quark doublet. Interactions within a doublet are governed by the term in the Lagrangian, deriving from the charged weak current, that looks like [22]

$$\begin{aligned} \mathcal{L}_{W-q} &= -\tfrac{g}{2\sqrt{2}}(\overline{\psi_i}\gamma^\mu V^{ij}[T^+W_\mu^+ + T^-W_\mu^-]\psi_j) \\ &= -\tfrac{g}{2\sqrt{2}}[\overline{u}^i\gamma^\mu(1-\gamma_5)V^{ij}d^jW_\mu^+ + \overline{d^i}\gamma^\mu(1-\gamma_5)(V^{-1})^{ij}u^jW_\mu^-] \end{aligned} \tag{1.17}$$

where $T^\pm$ are isospin raising and lowering operators, $W^\pm$ is the W gauge boson field, and where V can be shown to be unitary [23]. Though the simplest imaginable case is $V=\mathbf{1}$(corresponding to flavor-conserving W interactions), no symmetry principle places any constraints on V. We know from experiments that indeed $V \neq \mathbf{1}$. Put another way, though we designate the flavors by strong production, the weak interaction can see them in a different basis–one described by V. V is known as the Cabbibo-Kobayashi-Maskawa (CKM) flavor mixing matrix [7],[8].

1.2.2 CKM Physics

The CKM matrix V describes the coupling from quarks with weak isospin $+\frac{1}{2}$ to quarks with weak isospin -$\frac{1}{2}$. When there are 3 generations, the CKM is a 3×3 unitary matrix:

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{1.18}$$

Because it is unitary, and because overall phases are not detectable, rather than measuring 9 complex numbers, we can fully parameterise the CKM matrix with only 4 numbers. One of these numbers is a complex phase, which can generate CP violation[2]. By convention, this phase is usually assigned to the V_{td} and V_{ub} elements. There are many parameterization schemes, but for following CP violation literature, the most useful is probably the Wolfenstein parameterization [24], though we caution the reader that the parameterization is only approximate (it violates unitarity at λ^4):

[2] It is worth considering that we may owe our existence to a property of 3×3 Unitary matrices

$$\begin{pmatrix} 1-\frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ V_{-\lambda} & 1-\frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} \tag{1.19}$$

An additional prominent parameterization is to define α, β, and γ through

$$V_{td} \equiv |V_{td}|e^{-i\beta} \tag{1.20}$$

$$V_{ub} \equiv |V_{ub}|e^{-i\gamma} \tag{1.21}$$

$$\alpha \equiv \pi - \beta - \gamma \tag{1.22}$$

Note that these imaginary phases are associated with the third generation (V_{td} and V_{ub}), and are believed from our current understanding of the CKM matrix to be large. This leads to our expectation that CP violation will be large in the third-generation B system.

1.2.3 The Unitarity Triangle

Unitarity implies several sets of relations among the elements of the CKM matrix. Three of them are independent:

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \tag{1.23}$$

$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0 \tag{1.24}$$

$$V_{ud}V_{us}^* + V_{cd}V_{cs}^* + V_{td}V_{ts}^* = 0 \tag{1.25}$$

$$\tag{1.26}$$

One can see that these relations require that the “phasor diagram” of the relevant equations must close. These phasor diagrams look like traingles, hence they are known as “unitary triangles”. When referring a “the” unitary triangle, equation 1.23 is meant; it is the only relation among the three which does not produce one short side combined with two long, skinny sides. The other triangles produce extremely imprecise determinations of the Jarlskog CP-quantity

$$J_{CP} = \Im(V_{cb}V_{td}V_{cd}^{*}V_{tb}^{*}) = \lambda(1 - \lambda^{2}/2)Im(V_{td}) \tag{1.27}$$

to which all CP-violating quantities in the Standard Model are proportional. As pointed out in [19], mathematically J_{CP} can be as large as $\frac{1}{6\sqrt{3}} \simeq 0.1$ for a unitary matrix, but in our universe is found to be near $4{\times}10^{-5}$.

Since the triangle 1.23 is sensitive both the the area J_{CP} and the angles α, β, and γ, it has excited much interest. It is conventional to divide by $V_{cd}V_{cb}^{*}$ to obtain a normalized equation

$$1 + \frac{V_{ud}V_{ub}^{*}}{V_{cd}V_{cb}^{*}} + \frac{V_{td}V_{tb}^{*}}{V_{cd}V_{cb}^{*}} = 0 \tag{1.28}$$

We can construct the unitarity triangle simply by knowing the sizes of the three sides of the triangle. This determines the angles of the unitarity triangle. However, these angles have interpretations, as shown in figure 1.2, as the physically meaningful complex phases of the CKM matrix, as defined in equations 1.21, 1.22, and 1.22.

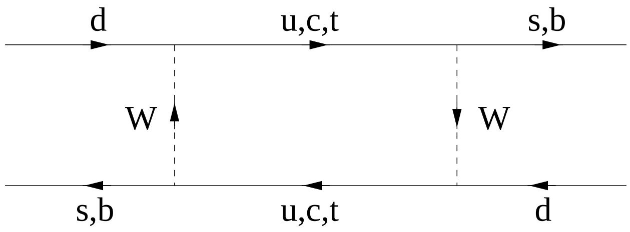


Figure 1.1: Box diagram for $K^0 - \overline{K^0}$ and $B^0 - \overline{B^0}$ mixing.

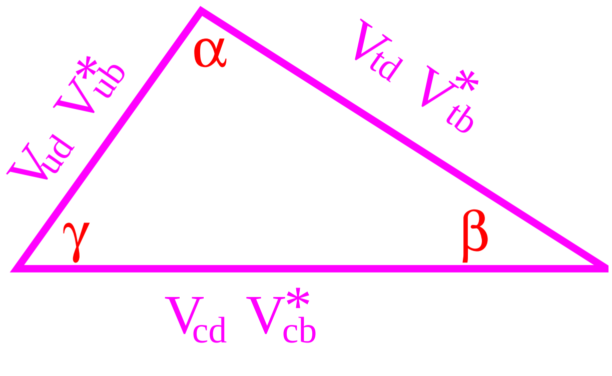


Figure 1.2: The Unitarity Triangle. The angles of the triangle α, β, and γ correspond to physically meaningful phases of the CKM matrix.

As physical quantities, these phases, too, can be measured. There are then two (related) experimental questions. Are the directly measured values of these phases the same as that deduced from the sizes of the sides of the unitarity triangle? And is the sum of the three angles 180^o?

The first goal of a B factory is to measure the angle β. In a subsequent section I demonstrate the formalism for any such phase measurement using neutral mesons, and explain how it may be applied for executing the experiment.

1.2.4 On V_{td}

One of the most poorly known CKM elements is V_{td}. Not only is the phase unmeasured, not even its magnitude is well-constrained. In fact, this is not for lack of experimental input. The best determination of V_{td} is expected to come from the Δm parameter in B^0-$\overline{B^0}$ mixing:

$$\Delta m_d = \frac{G_F^2}{6\pi^2} f_B^2 B_B^2 m_B m_W^2 \eta_{QCD} S_0(m_t) |V_{tb}^* V_{td}|^2 \tag{1.29}$$

Unfortunately, the parameters f_B and B_B are not well-known. They are the psuedoscalar matrix elements and "bag parameters" which must be determined by the theory of strong dynamics. Though lattice QCD has made great progress in calculating these quantities, large uncertainties remain.

There are several ways to circumvent this problem. One is to measure, instead of B mixing, the ratio of B_d and B_s mixing. This gives the ratio $\frac{V_{td}}{V_{ts}}$. The hadronic uncertainties in the ratio $\frac{f_{B_d}\sqrt{B_{B_d}}}{f_{B_s}\sqrt{B_{B_s}}}$ are generally believed to cancel, as the uncertainties arise primarily from the "quenched" approximation in QCD. In this approximation, the quark loops do not modify the bare gluon propagator. The effect of this approximation is expected to be the same for B_d and B_s mesons. As V_{ts} is well-known, this leads to an accurate determination of V_{td}.

Another method is to measure the rate $B^{\pm} \to \tau\nu$, which measures f_B and V_{ub} simultaneously. Other methods include measuring the ratio $\frac{b\to d\gamma}{b\to s\gamma}$ and the rate $K_L \to \pi^0\nu\overline{\nu}$.

In chapter 4, we will measure the value of Δm in B^0-$\overline{B^0}$, and in chapter 5 we present implications for V_{td}.

1.3 Mixing Induced CP Violation

1.3.1 Introduction

Observed CP violation in the neutral K system [10] is explained in the Standard Model by the Cabbibo-Kobayashi-Maskawa (CKM) mechanism [7],[8]. Despite 35 years of study, no other system has been found to display CP violation, profoundly limiting the ability to test the CKM model. One of the most important proving grounds for this test is the neutral B system, which is expected to exhibit large CP asymmetries due to the interference of mixed and direct decays to non-flavor-specific CP eigenstates (henceforth referred to as mixing-induced CP violation). Upcoming asymmetric B factories, operating at the $\Upsilon(4S)$, propose to measure CP asymmetries. These experiments measure flavor asymmetries as a function of the difference of decay times of the two B mesons from the decay of an $\Upsilon(4S)$.

This chapter examines a new method for use at the $\Upsilon(4S)$, in which the decay time of only one B is observed.

Once the Υ(4S) has decayed to a pair of neutral B mesons, the $B^0\overline{B}^0$ exhibit coherent flavor oscillations until the time of the first meson's decay. Afterwards the second B meson oscillates freely,with angular frequency Δm, from its state at the time of the first B decay, until its own decay. When one B decays to a CP eigenstate, CP violation can be established from the other B meson via a flavor asymmetry in flavor-specific decays. Because the oscillations are coherent, the flavor excess depends on the decay times. Because the meson pair is in a $C = -1$ state, the excess vanishes if no time information is used. This leads one naturally to consider $\Upsilon(4S)$ production with asymmetric colliders, where the large boost allows time information to be easily observed.

In the existing literature on $\Upsilon(4S)$ measurements, observation of the time difference Δt between the two B decay times is often stated as a necessity. It is not uncommon to find such statements as "to this end [measuring CP violation], one needs to determine the time interval between the two B-meson decays" [11], "only if we can observe the decay time difference $T = t - t'$ [Δt] can we measure a nonzero asymmetry in the decay of a $C = -1$ $B\overline{B}$ state" [12], or "a determination of Δt is required for the observation of a CP asymmetry in experiments at the $\Upsilon(4S)$" [13]. Several methods have been proposed to measure CP violation using the observed time difference [14, 15]. Some work has addressed the case in which the time-ordering is observed but the time difference is not [16], and other work has addressed the case in which the b quarks are produced incoherently, *e.g.* in $p\overline{p}$

collisions or at the Z^0 [17, 18].

In this section, we show that for interference between mixed and direct decays to CP eigenstates (such as the "golden mode" $B \to \psi K_S$), neither time ordering nor time-differences are required in order to be sensitive to CP violation at the $\Upsilon(4S)$. Here we focus on the case for a coherent $B\overline{B}$ pair when one decay time is measured while the other decay time, and hence even the relative ordering, is unobserved. Specifically, we concentrate on the case in which it is the decay time of the B decay to CP-eigenstate that is observed. Typically, this final state contains a ψ which provides accurate decay time information. Often, the other (flavor-tagging) decay time is difficult to measure accurately; this is the case, for example, when the flavor tag is provided by a charged K, rather than a primary lepton. This is because K's originate from charm mesons, which decay after travelling an average distance of nearly 100μm from the B decay point.

1.3.2 Υ Decay Formalism

I begin with the expression for the decay rate of the correlated $B^0\overline{B^0}$ state into any pair of final states f_1 at time t_1 and f_2 at time t_2:

$$
\begin{aligned}
R(t_1,t_2) \quad = \quad & Ce^{-\Gamma(t_1+t_2)}\{(|A_1|^2+|\overline{A}_1|^2)(|A_2|^2+|\overline{A}_2|^2) - \\
& 4Re(\tfrac{q}{p}A_1^*\overline{A}_1)Re(\tfrac{q}{p}A_2^*\overline{A}_2) \\
& -\cos(\Delta m(t_1-t_2))[(|A_1|^2-|\overline{A}_1|^2)(|A_2|^2-|\overline{A}_2|^2) - \\
& 4Im(\tfrac{q}{p}A_1^*\overline{A}_1)Im(\tfrac{q}{p}A_2^*\overline{A}_2)]
\end{aligned}
$$

$$+2\sin(\Delta m(t_1 - t_2))[Im(\tfrac{q}{p}A_1^*\overline{A}_1)(|A_2|^2 - |\overline{A}_2|^2) -$$
$$(|A_1|^2 - |\overline{A}_1|^2)Im(\tfrac{q}{p}A_2^*\overline{A}_2)]\}, \tag{1.30}$$

where we use the notation found in [19], with the standard [20] meanings for λ, p, and q: p and q are the coefficients for the quantum-mechanical admixture of B^0 and $\overline{B}^0$ which comprises the B mass eigenstates; $\lambda = \frac{q}{p}\frac{\overline{A_{f_{CP}}}}{A_{f_{CP}}}$. A_i is the amplitude for a B^0 to decay to f_i, and $\overline{A}_i$ is the amplitude for $\overline{B^0}$ to decay to the same state f_i. As the B and $\overline{B}$ have few decay final states in common, $\frac{\Delta\Gamma}{\Gamma}$ is expected to be small, and has been set to 0 here.

To measure CP violation, one observes a decay at time t_1 to a CP eigenstate f_{CP} and a decay at time t_2 for a flavor-tagging state f_{tag}. In the case of the flavor-tagging decay of a B^0, A_2=A_{tag},$\overline{A_2}$=0:

$$\begin{aligned} R_+(t_{tag}, t_{CP}) &= Ce^{-\Gamma(t_{tag}+t_{CP})}|\overline{A}_{tag}|^2||A_{CP}|^2 \\ &\quad \times\{1 + |\lambda_{CP}|^2 - \cos[\Delta m(t_{CP} - t_{tag})](1 - |\lambda_{CP}|^2) \\ &\quad +2\sin[\Delta m(t_{CP} - t_{tag})]Im(\lambda_{CP})\} \end{aligned} \tag{1.31}$$

where $x \equiv \frac{\Delta m}{\Gamma}$, and we have set $|\frac{p}{q}| = 1$, again as expected in the Standard Model. When the flavor tag indicates the decay of the $\overline{B^0}$ (as opposed to B^0) the expression for the rate becomes

$$\begin{aligned} R_-(t_{tag}, t_{CP}) &= Ce^{-\Gamma(t_{tag}+t_{CP})}|\overline{A}_{tag}|^2||A_{CP}|^2 \\ &\quad \times\{1 + |\lambda_{CP}|^2 + \cos[\Delta m(t_{CP} - t_{tag})](1 - |\lambda_{CP}|^2) \end{aligned}$$

$$-2\sin[\Delta m(t_{CP}-t_{tag})]Im(\lambda_{CP})\} \quad (1.32)$$

that is, the signs of the cos and sin terms have flipped relative to the case of a B^0 tag. The subscripts on $R_\pm$ serve to denote the flavor tag $B^0(\overline{B}^0)$.

By integrating over all times t_{tag}, while retaining the sign ($\pm$) of the flavor tag, one obtains an expression for the decay rate, $R_\pm$, to a CP eigenstate, as a function of the CP decay time only:

$$\begin{aligned}
R_\pm(t_{CP}) &= \int_0^\infty dt_{tag} R(t_{tag}, t_{CP}) \\
&= \int_0^\infty dt_{tag} C e^{-\Gamma(t_{tag}+t_{CP})} |\overline{A}_{tag}|^2 |A_{CP}|^2 \\
&\quad \times\{1+|\lambda_{CP}|^2 \mp \cos[\Delta m(t_{CP}-t_{tag})](1-|\lambda_{CP}|^2) \\
&\quad \pm 2\sin[\Delta m(t_{CP}-t_{tag})]Im(\lambda_{CP})\} \\
&= \frac{C|\overline{A}_{tag}|^2|A_{CP}|^2}{\Gamma(1+x^2)} e^{-\Gamma t_{CP}} \{(1+x^2)(1+|\lambda_{CP}|^2) \\
&\quad \pm(1-|\lambda_{CP}|^2)[\cos(\Delta m t_{tag}) + x\sin(\Delta m t_{tag})] \\
&\quad \pm 2 Im(\lambda_{CP})[x\cos(\Delta m t_{CP}) - \sin(\Delta m t_{CP})]\} \qquad (1.33)
\end{aligned}$$

Taking $|\lambda_{CP}|$=1 (i.e., neglecting direct CP violation), the expression for the distribution of decay times to a CP eigenstate is of the form,

$$\begin{aligned}
R_\pm(t_{CP}) &= \frac{2C|\overline{A}_{tag}|^2|A_{CP}|^2}{\Gamma(1+x^2)} e^{-\Gamma t_{CP}} \{(1+x^2) \pm Im(\lambda_{CP})[x\cos(\Delta m t_{CP}) \qquad (1.34) \\
&\quad - \sin(\Delta m t_{CP})]\} \\
&= \frac{2C|\overline{A}_{tag}|^2|A_{CP}|^2}{\Gamma} e^{-\Gamma t_{CP}} \{1 \mp \frac{Im(\lambda_{CP})}{\sqrt{1+x^2}}[\sin(x(\frac{t_{CP}}{\tau} - \frac{\tan^{-1}x}{x}))]\} \qquad (1.35)
\end{aligned}$$

This distribution in shown in figure 1.3.

Despite having integrated over t_{tag}, CP asymmetry information is still encoded in the distribution $R_{\pm}(t_{CP})$. (Note that the integral $\int_0^\infty dt_{CP} R_{\pm}(t_{CP})$ does equal 0, as expected). *Hence even though the decay ordering and time difference are not observed, the decay distributions are* **different***for events that are tagged as B^0 or $\overline{B^0}$.* This leads to an asymmetry in $B^0(\overline{B^0})$ tags as a function of the CP eigenstate decay time,

$$\begin{aligned} A_{CP}(t_{CP}) &= \frac{R_+(t_{CP}) - R_-(t_{CP})}{R_+(t_{CP}) + R_-(t_{CP})} \\ &= Im(\lambda_{CP}) \frac{\sin(x(\frac{t_{CP}}{\tau} - \frac{\tan^{-1} x}{x}))}{\sqrt{1+x^2}} \end{aligned} \tag{1.36}$$

this distribution is shown in figure 1.4.

1.3.3 Experimental Implications

For the CP eigenstate ψK_s, $Im(\lambda_{CP})$ is equal to $\sin 2\beta$, where β is the usual angle of the unitarity triangle [19], and is nearly equal to $\arg(V_{td})$. One measurement method proceeds as follows. The experimenter selects decays in the $B \to \psi K_S$ mode. The $\psi \to \ell^+\ell^-$ vertex provides good decay time information when the beam position is well-known. The information is preserved, though diluted, even when only one dimension of the beam is well-determined, as is generally the case in the y dimension of an e^+e^- collider. The other B decay, with decay time unobserved, provides a tag as either a B^0 or $\overline{B^0}$. The decay time distributions for the ψK_S candidates are formed for the two tag cases, and fit to the form of equation 1.35. The lifetimes and mixing parameter x have been well measured and may be fixed in the fit. The experimenter then fits only for a normalization and the value of $\sin 2\beta$. This method is preferable to fitting the asymmetry of equation 1.36 directly, as resolution effects are more easily incorporated. At a symmetric B factory, resolutions are expected to be comparable to the lifetime, and therefore important.

Even in the limit of poor resolution, where the oscillation modulations are not directly observable, one may simply measure the mean decay time of the B CP eigenstate. This mean decay time will be different for events in which the other B is tagged as a $\overline{B^0}$ or B^0 decay:

$$\bar{t}_{\pm} = \tau\left(1 \mp \tfrac{x \sin 2\beta}{(1+x^2)^2}\right) \tag{1.37}$$

The difference in the mean times is proportional to $\sin 2\beta$:

$$\bar{t}_{+} - \bar{t}_{-} = \tau \tfrac{2x \sin 2\beta}{(1+x^2)^2} \tag{1.38}$$

For $x = 0.7$, $\Delta\bar{t} = 0.63\ \tau \sin 2\beta$. In the Υ rest frame this is a difference of 18 $\times \sin 2\beta$ μm.

The technique of integrating over one decay time may also be applied to the CP side. That is, CP violation may be established by observing the flavor-tagging B's decay time while integrating over the CP tag decay time. This makes it possible to observe indirect CP violation in all-neutral final states. Previously it was often assumed that these modes could be useful only for measuring decay rates and direct CP violation. Examples of such modes include $\pi^0\pi^0$, $K_S\pi^0$, or K_SK_S.

These modes are expected to have very small branching fractions, but may be helpful in understanding penguin contributions to the measurement of $\sin 2\alpha$.

Many of the remarks and conclusions concerning CP asymmetries apply also to a measurement of Δm. To measure Δm, the two decays observed are both flavor-tagged. As with CP violating parameters, Δm may be measured using events in which one decay time is observed and the other unobserved. Equation 1.30 is again the starting point. For events observed to be unmixed (two flavor tags of opposite flavor), $A_1 = A$, $\overline{A_1} = 0$, $\overline{A}_2 = 0$, and $A_2 = A^*$. Performing the integration over one of the times gives

$$R_{unmixed}(t) = 2C|A|^4 \frac{e^{-\Gamma t}}{\Gamma(1+x^2)}\{1 + x^2 + [\cos(\Delta m t) + x\sin(\Delta m t)]\} \quad (1.39)$$

For the rate to mixed events (both flavor tags of the same sign), $A_1 = A$, $\overline{A_1} = 0$, $\overline{A}_2 = A^*$, and $A_2 = 0$:

$$R_{mixed}(t) = 2C|A|^4 \frac{e^{-\Gamma t}}{\Gamma(1+x^2)}\{1 + x^2 - [\cos(\Delta m t) + x\sin(\Delta m t)]\} \quad (1.40)$$

The distributions $R_\pm$ are shown in figure 1.5.

The resulting asymmetry of mixed to unmixed events is thus

$$\begin{aligned} A_{unmixed/mixed}(t) &= \frac{\cos(\Delta m t) + x\sin(\Delta m t)}{1+x^2} \\ &= \frac{\cos(x(\frac{t}{\tau} - \frac{\tan^{-1} x}{x}))}{\sqrt{1+x^2}} \end{aligned} \quad (1.41)$$

This asymmetry, along with the CP asymmetry, is shown in figure 1.4. Again, the average decay times are different for the two cases of mixed and unmixed events,

$$\bar{t}_{u/m} \quad = \quad \tau(1 \pm \tfrac{x^2}{(1+x^2)(1 \mp 1+x^2)}) \tag{1.42}$$

and mean decay time difference

$$\Delta\bar{t} \quad = \quad \tau\tfrac{2}{1+x^2} \tag{1.43}$$

The form of the asymmetry in equation 1.41 provides an interpretation of the physics. The more familiar form for the asymmetry 1.41, applicable when the B decay is incoherent with the other B in the event, is given by

$$A_{unmixed/mixed}(t) \quad = \quad \cos(\Delta mt) \tag{1.44}$$

The asymmetry 1.41 is offset from the familiar $\cos(\Delta mt)$ form by a time phase of $\frac{\tan^{-1} x}{x}$ lifetimes, which is nearly 1 for small values of x [3]. That is, the mixing oscillation proceeds nearly as if the other B had decayed at time $t \simeq \tau$. A slight ($\simeq$ 20%) dilution arises from the $\frac{1}{\sqrt{1+x^2}}$ term. This also explains the form of the equations describing mixing-induced CP violation. Because it arises from interference between mixing and direct decays, CP violation carries the same phase.

In conclusion, we have shown that CP violation and mixing are both observable at the $\Upsilon(4S)$ without the need for time ordering of the B decays or measurement of

[3]The time phase is 0.87 lifetimes for x=0.7.

Δt. The necessary experimental ingredients are precision vertexing information, efficient flavor tagging, and high-efficiency reconstruction of CP eigenstates. A necessary experimental input is also the value of the mixing parameter x_d.

In the following four chapters we elaborate on these requirements. We have used data collected by the CLEO experiment at the Cornell Electron Storage Ring (CESR) to measure the effectiveness of a silicon vertex detector, the power of a flavor tagger, and to make a precision measurement of the value of χ_d. The reconstruction of CP eigenstates is already well-established, and is not discussed here in detail.

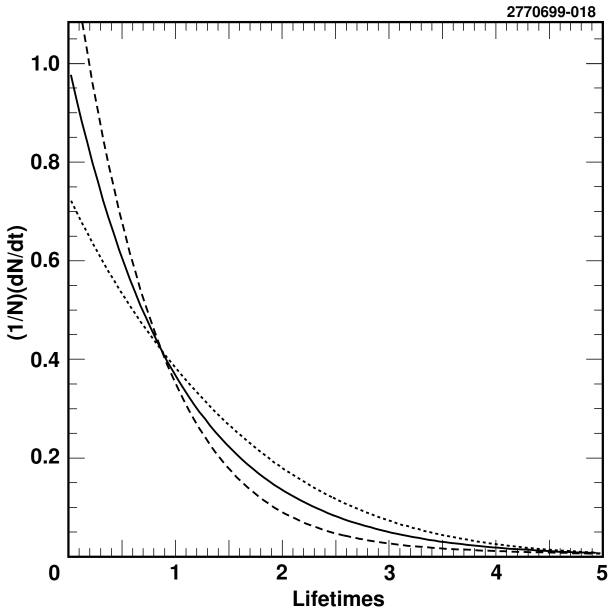


Figure 1.3: Effect of CP violation on decay time distributions of ψK_S final state. The solid line shows the case $\sin 2\beta$=0; the dotted(dashed) line shows the distribution of the decay times when the opposite B is tagged as a $B^0(\overline{B^0})$, for $\sin 2\beta$=0.7.

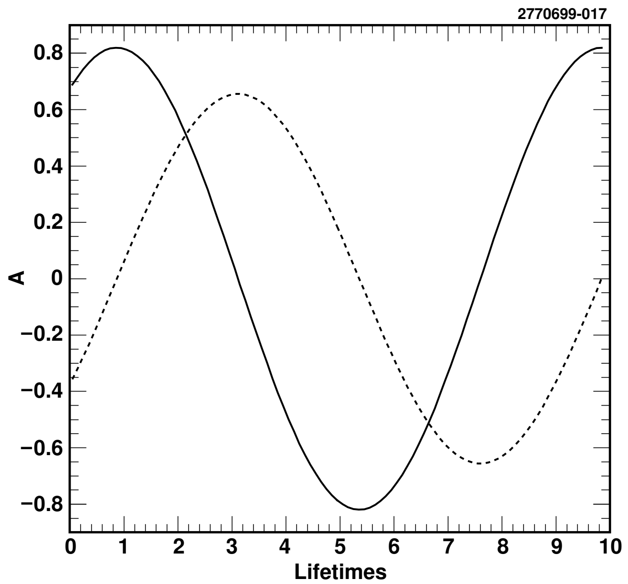


Figure 1.4: The asymmetries for A_{CP} (dashed) and $A_{mixed/unmixed}$ (solid) and CP tags (dashed) of the measured decay time. The plots are generated setting x=0.7, $\sin 2\beta$=0.7.

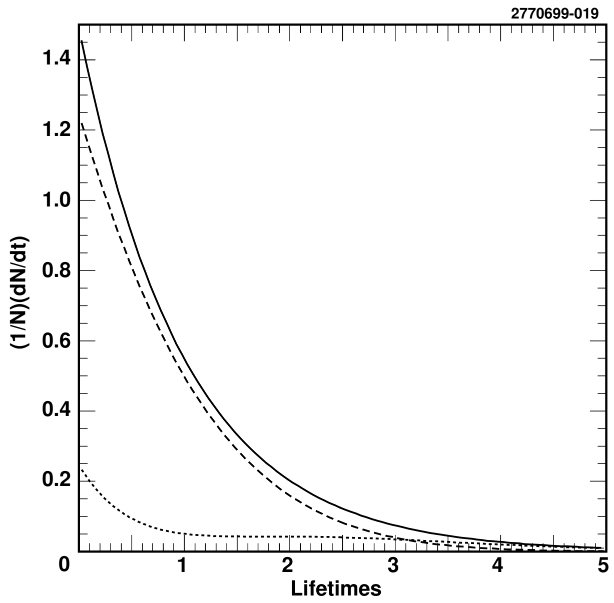


Figure 1.5: Effect of $B-\overline{B}$ mixing on decay time distributions on a flavor-tagging final state. The solid line shows the exponential decay in the case of no mixing ($x_d = 0$). The dotted(dashed) line shows the distribution of the decay times when the opposite B is tagged as a mixed(unmixed), for x_d=0.7. The large average decay time of mixed events is readily apparent from the plot.

CHAPTER 2

POSITION RESOLUTION: PERFORMANCE OF A SILICON VERTEX DETECTOR

2.1 Introduction

As noted in the introduction, precise spatial information is required in order to measure mixing-induced CP violation at the $\Upsilon(4S)$. This was further seen in the development of the formalism of Υ decays to $B\overline{B}$, where it was crucial to observe time-dependent information. Temporal and spatial information are, of course, related by $\beta\gamma c = 0.056c$. To provide this accurate spatial information, CLEO and the upcoming B-factories will use silicon vertex detectors surrounding the e^+e^- interaction point. In this chapter, we study the precision and abilities of the CLEO II.V silicon vertex detector (SVX) as a prototype of these future detectors. This is the first such detector to be operated at energies as low as the $\Upsilon(4S)$, where multiple scattering and range-out are much greater issues than at high-energy accelerators, such as LEP or the Tevatron.

The CLEO II.V SVX detector was installed in the CLEO detector at the Cornell Electron Storage Ring in November, 1995. It was operated until February 1999, taking over 9 fb^{-1} of e^+e^- data. This detector, with approximately 12 μm per-hit resolution, is intended to provide precision position information, allowing the reconstruction of particle lifetimes. Indeed, CLEO has already published a lifetime measurement using the detector [21].

In this chapter, we present studies of the silicon detector's efficiency, resolution, and noise. We also present a study of the systematic accuracy of the length scale in the detector.

2.2 The CLEO II.V SVX Detector

For a fuller description of the detector, we refer the reader to the notes [25],[26]. As a brief reminder, the CLEO II.V SVX detector is a device consisting of three layers of 300 μm double-sided SI wafers. This means that both $r-\phi$ and $r-z$ information is read out for each wafer. The wafers in a layer overlap each other in about 5% of the detector, so that it is possible for two wafers to be hit in the same layer.

Aside from the overlap regions, a perfect detector with perfect reconstruction software would produce three $r-\phi$ and three $r-z$ hits for each track. In addition, on each individual wafer, the $r-\phi$ and $r-z$ charge depositions would be equal[1]. The detector is of course known to be imperfect (due mainly to known bad channels) and to have fairly poor resolution ($\simeq$25%) on the charge deposition measurement.

[1] Up to backplane capacitance on the $r-\phi$ side.

The pitch of the readout strips in the detector is roughly (depending on layer) 112 μm in $r-\phi$ and 105 μm in $r-z$. The deposited charge is distributed among multiple strips, an average of 2.2 strips on the $r-\phi$ side and 2.8 strips on the $r-z$ side. This leads us to "cluster" the raw hits. The raw hits on the channels are grouped together into "analyzed hits" by the routine `svrtoa`. It is these analyzed hits which are subsequently used by the tracking packages.

2.3 SVX Tracking Software and Algorithms

It is not the intention here to present a complete description of the tracking software in use in CLEO II.V. Such a description is to be written up in a separate note. The intention in this section is to present an overview of the algorithms.

The collaboration has two independent tracking packages for the silicon, `DUET` (also unofficially known, in its silicon incarnation, as "xDUET") and `SVTF`. `DUET` is based on extending the DUET pattern recognition to work using silicon hits. This package is based on tracks already found using `DUET`, so is in some sense "outside-in". The `SVTF` processor lists all possible combinations of silicon hits that form track segments, then attempts to match them to DUET outer-chamber tracks. In this sense SVTF is an "inside-out" package.

Though each package is designed to be useful on its own, we obtain the best overall results using a combination of the two, often called the "hybrid tracker". The hybrid tracker is now the default tracker run in the `PASS2` of CLEO II.V data.

2.3.1 `DUET` with Silicon

For an overall reference to how `DUET` worked in CLEO II, we refer the reader to the excellent notes by Rob Kutschke [27][28]. Here we recap the changes relevant for silicon.

The `DUET` routine `tfxfit` is responsible for adding silicon hits to tracks. This routine is also responsible for adding hits from all chambers in order to fill out the hit list received from the `linktree` algorithm. `tfxfit` is run twice, first in the $r-\phi$ view and then in the $s-z$ view. In each view, it adds hits to tracks in an iterative series of passes. First comes a series of "local ambiguity" passes, in which hits can only be added to the track in groups of neighboring hits. The first pass adds only DR "inner" (first 30 axial layers) hits. The second pass considers VD hits. Pass three adds hits from the "outer" DR layers (the last 10 axial layers). After 3 full passes through the outer chambers, silicon hits are considered.

When silicon hits in the view are considered for addition to a track, hits from different layers are grouped together by links. (This is taken from the drift chamber code, in which such groups are used to locally determine the drift sign ambiguities from the layer-to-layer stagger.) Groups of 2 (from layers 1 and 2, 2 and 3, or 1 and 3) and groups of 3 (from all three layers) are considered. (Hence, in each view, DUET is effectively doing SV-only pattern recognition). For each group, the mean residual from the track, and the spread of residuals within the group, is considered. The group with the best mean and RMS is taken.

After the four local ambiguity passes, hits continue to be added and deleted in a series of "standard" passes. The standard pass acts as a clean-up pass, adding those hits which fall unambiguously near the track, regardless of whether neighboring hits are grouped with it. The standard pass is reiterated until no more hits are found.

Note that this scheme does not take full advantage of the physical correlations built into the silicon detector. For instance, since East-West information is not available when the $r-\phi$ hits are added, it is logically possible to add hits in $r-\phi$ that come from the wrong half of the detector. Also note that this scheme does not take advantage of charge-correlations between the ϕ and z sides of the silicon wafer. In order to address these shortcomings, a routine called `sv_postop.F` exists in order to ensure that the $r-\phi$ and z hits are consistent with each other.

2.3.2 `SVTF`, the Standalone Tracker

The other tracking software used with the SVX, `SVTF`, is based on a philosophy of matching silicon track segments to DR tracks. The `SVTF` processor makes a list of all physically possible combinations of silicon hits into short track segments, taking into account the physical correlations between $r-\phi$ and $r-z$ hits. In addition, VD and DR cathodes are used to limit the possible combinations. For a typical event with 250 hits, this yields around 400 track segments. For each segment, track parameters and an error matrix are calculated by a helix fit.

The segments can later be used in the `DUET` processor in `pass2`, by matching drift chamber tracks to SVX track segments. After `DUET` finds a track, the track is fit and extrapolated for χ^2 comparison to the standalone track parameters. If this χ^2 is sufficiently small, the SV hits in the segment are added to the track and the track is refit with the hits.

In order to save computing time in processing, `SVTF` does not generate track segments in every event. Rather, it is run on those events which have either the Level 3 Hadron bit set, or those events in which `TRIO` finds 3 tracks. In particular, `SVTF` is not necessarily run on two-photon events.

2.3.3 The Hybrid Tracker

`DUET` and `SVTF` both have strengths and weaknesses, and to take maximum advantage of both packages, they run in sequence. If, after the DUET pattern recognition has run, it has not found at least 2 hits in each view then the χ^2 matching of the `SVTF` segments is invoked in order to add SV hits to the track. Of course, if `SVTF` did not generate segments for this event because it was not a hadronic event, there will be no matching segments.

The impact of the hybrid upon performance is most apparent at low momentum, below 300 MeV. This is shown in figure 2.1, which shows the fraction of tracks with silicon hits in which the silicon is contributed by the stand-alone tracker.

2.4 Using Physics to Measure SVX Performance

Inclusive measures of the efficiency and noise in the SVX can be very misleading. There are many ways for the tracking software to find tracks which should not have silicon hits on them. These can be δ rays, long-lived K_S/Λ, γ conversions, in-flight decays, curler back halves, hard scatters, and ghosts. These can contribute up to

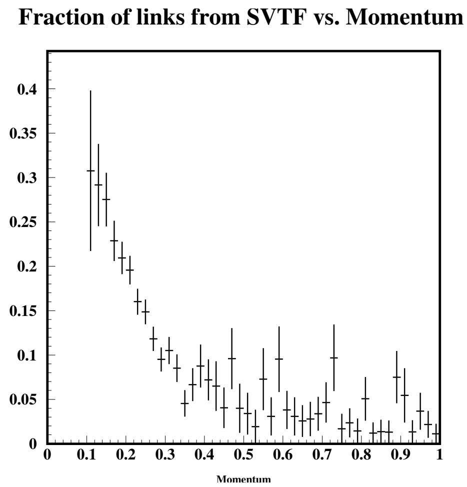


Figure 2.1: Fraction of silicon hits on tracks that were added by `SVTF`.

20% to inclusive tracks, artificially lowering the apparent efficiency.

This makes it critical to use physics processes to study these questions. This study makes use of Full Roar data from the official CLEO II.V `PASS2` of the 4sL dataset. It uses μ-pairs, $D^{*+} \to D^0\pi^+$, $\gamma\gamma \to e^+e^-$, and $\gamma\gamma \to 4\pi$ events, whose selection is described below.

The question of alignment of the detector is not addressed here. Instead, there will be an upcoming note in which alignment studies are outlined in detail. We note that the detector is without question sufficiently well-aligned as to have negligible effects on the conclusions presented in this chapter.

2.4.1 Selection of Events

For this study, all events come from the end (Full Roar tapes 70-80) of the 4sL dataset. It was necessary to run on the Full Roar because much silicon-related information is not contained in the hadron skims.

The Kalman Material

After the 4sL was fully processed through `pass2`, there was found an error in the implementation of the Kalman material description file for much of the 4sL, and in particular its description of the 900 μm of silicon, was found. This primarily affects the error matrices, and so for the most part does not change the conclusions of this paper. However, at the few percent level, this can have two effects.

First, the Kalman package deletes hits that are 4σ inconsistent with the track fit. Because the silicon material was improperly accounted for, silicon hits may have been deleted somewhat more often than they should have been. Secondly, due to the improper relative weighting of the SV and DR, the actual resolution can be slightly changed. We believe the first effect is negligible, so that most of the studies in this analysis were not remeasured after the bug was found. To account for the second effect, we remeasure the impact parameter resolution only with data for which this bug has been fixed.

Zero-lifetime Samples

In order to carry out resolution studies and systematic studies of the alignment, we need a sample which is known, without reference to the silicon, to have no lifetime component. μ pairs and two photon events can provide such samples. Unfortunately, two photon events can also have large backgrounds due to beam-gas and beam-wall events. In order to eliminate these backgrounds, we additionally require tight particle ID.

At the time of this writing, `DIFIX` and `TFFIX` were not available for CLEO II.V. Instead, we make use of the variables `BETATF` and `DE50DI`. Both of these variables have both sufficient resolution and sufficiently controlled systematic variation to be useful for these studies. In figure 2.2 we show the measured values of `BETATF` for electrons ($\langle$`BETATF`$\rangle = 1$). For 300 MeV π's $\langle$`BETATF`$\rangle = 0.90$, as it should. In figure 2.3 we show the `DE50DI` distribution for electrons as a function of momentum[2].

It is also important to know whether tracking efficiencies are adversely affected by the event environment. In order to study this, we also select a sample of $D^{*+} \to D^0\pi^+$ decays.

μ-pairs

μ pairs for this study are selected by requiring `KLASGL`=5.

$\gamma\gamma \to e^+e^-$

Electron selection can be made very clean, even for very low momentum electrons. The combination of $\frac{dE}{dx}$, `BETATF`, and $\frac{E}{p}$ can reject πs by a factor of several hundred even in the worst momentum region at 400 MeV. To select electrons, we require

[2] After making this plot it came to my attention that `DE50DI` has corrections applied to it that are not appropriate to electrons. The corrections have little impact for the purpose of this note

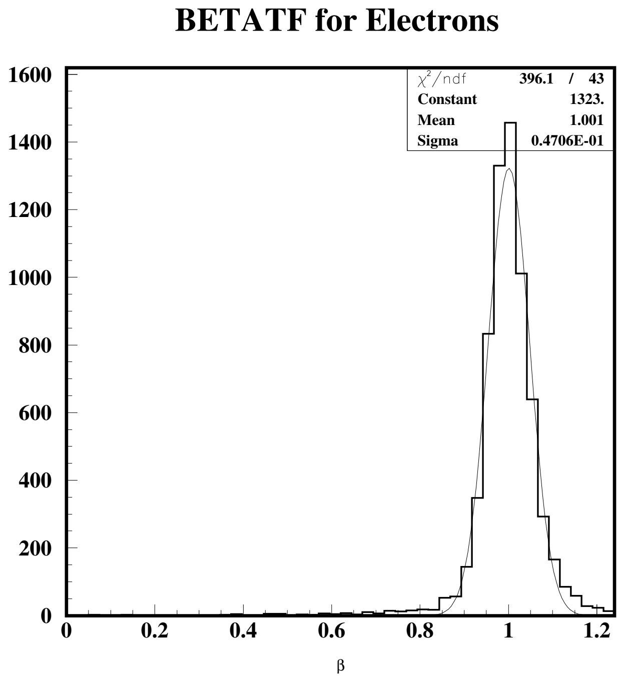


Figure 2.2: `BETATF` for electrons

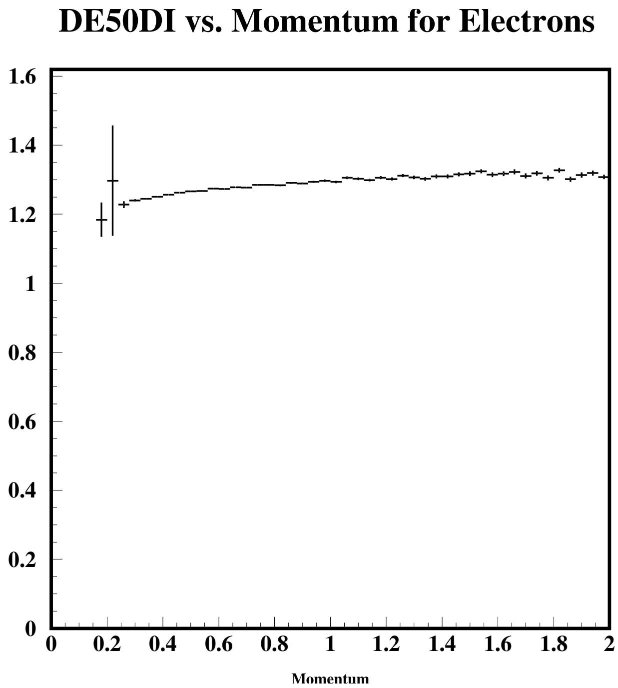


Figure 2.3: DE50 for electrons

- $\frac{ENTRSH}{MOMENTUM}$ >0.85
- BETATF>0.95
- $\frac{DE50DI-\langle DE50DI(MOMENTUM,CZCD)\rangle}{\sigma_{DE50DI}(MOMENTUM)}$ <2.5

We require the event to have

- KLASGL = 4
- NTRKCD = 2 or # of TNG .GE. 0 tracks = 2
- Total charge of event = 0
- $|\sum_{tracks}\vec{p_t}|$ <200 MeV
- 2 electrons as defined above

This skim selects about 500 events per pb^{-1}. It has less than 0.1% contamination from τs, as determined from generic τ Monte Carlo.

$\gamma\gamma \to 4\pi$

We require the event to have

- No particles consistent with being protons based on DE50DI
- No particles more likely to be K than π based on DE50DI
- KLASGL = 4
- NTRKCD = 4 or # of TNG .GE. 0 tracks = 4

- Total charge of event = 0
- $|\sum_{tracks} \vec{p_t}|$ <200 MeV
- 4 pions as defined above
- Some partitioning of $\pi^+\pi-$ pairs such that both pairs are consistent (within 150 MeV) with the ρ mass (770 MeV).

The proton veto rejects beam-wall and beam-gas events. The K veto rejects events which may contain a K_S. These criteria select about 50 events per pb-1. It has less than 0.1% contamination from τs, as determined from generic τ Monte Carlo.

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

Events must have `KLASGL`=10.

Candidates are selected by requiring

- $|m_{K\pi} - m_D|$ <15 MeV (2 σ)
- $|m_{K\pi\pi_s} - m_{K\pi}|$ <1 MeV (2-4 σ)
- $|\vec{p_{\pi_s}}|$ >100 MeV
- $|\vec{p_D}|$ >2.3 GeV

2.4.2 Good Tracks

In order to compare apples to apples, the silicon group has defined a set of good tracks criteria. These are **not intended for use in analyses**, but rather as a standard tracking candle. All tracks used in the efficiency, noise, and resolution analysis must pass the good track criteria, which are as follows:

- >20 VD+DR hits
- >5 VD hits
- `TFBNCF`>0.99
- `KINCD`=0
- `TNG` .GE. 0
- Not dredge
- Not z-escape

2.4.3 Fiducial Coverage of SVX

Figures 2.4 through 2.6 show, for each layer, the projection of particular μ-pair tracks in ϕ and z onto that layer. These are the tracks which contain hits in the other two layers, but have no hits in the layer under consideration. That is, these plots show track projections where hits should be, but aren't.

These plots show, taken together, four striking features:

- There is an obvious "crack" at Z=0, which corresponds to where the two halves of the detector meet.
- On layer 3, there are obvious secondary cracks at about z=6 cm; these are due to where the inner and outer detectors on layer 3 meet.

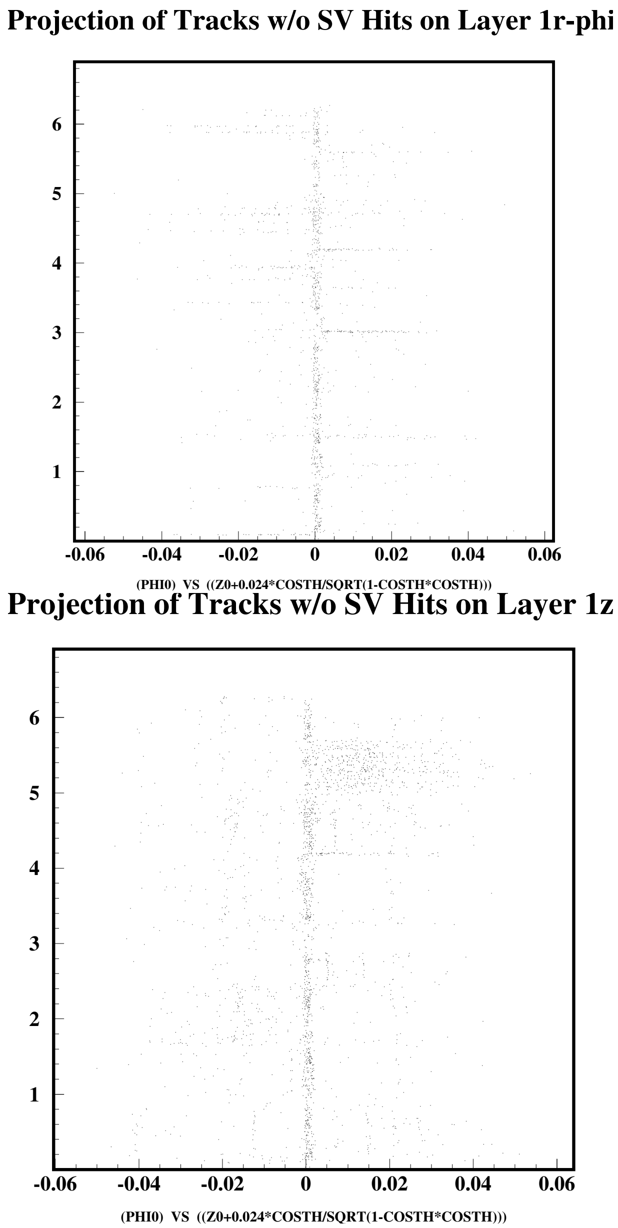
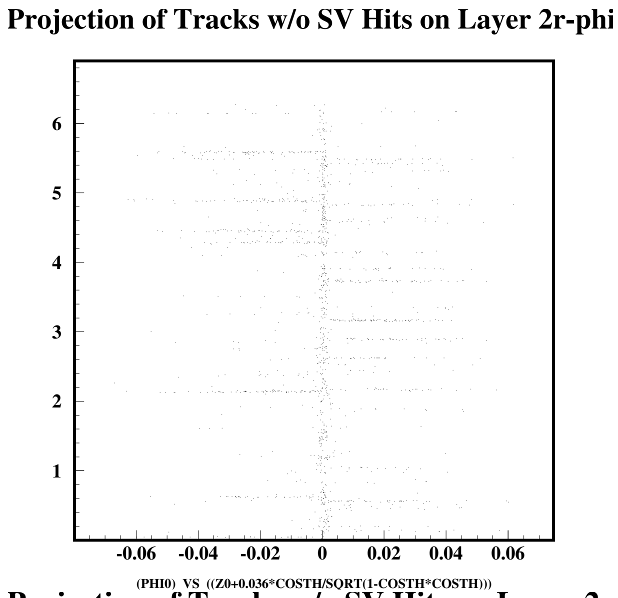


Figure 2.4: The plots show the detector "unrolled" in ϕ and z to display the location of tracks without hits on layer 1. Location of missing $r-\phi$ hits is on top, $r-z$ underneath.



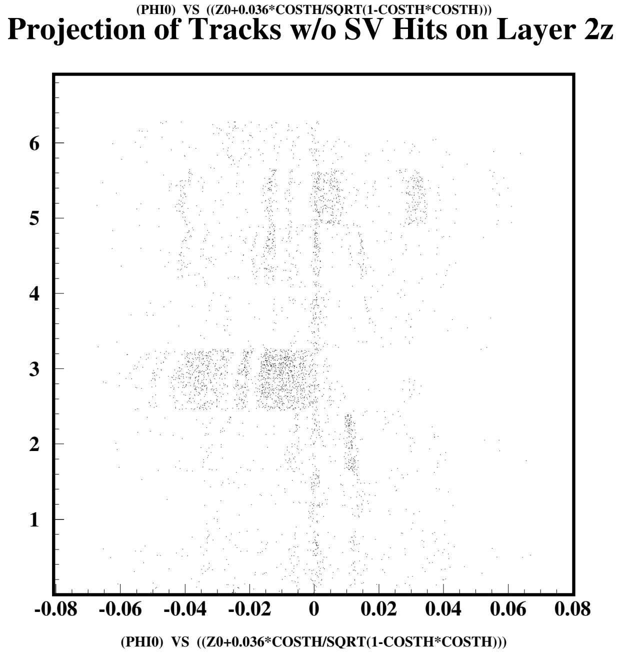
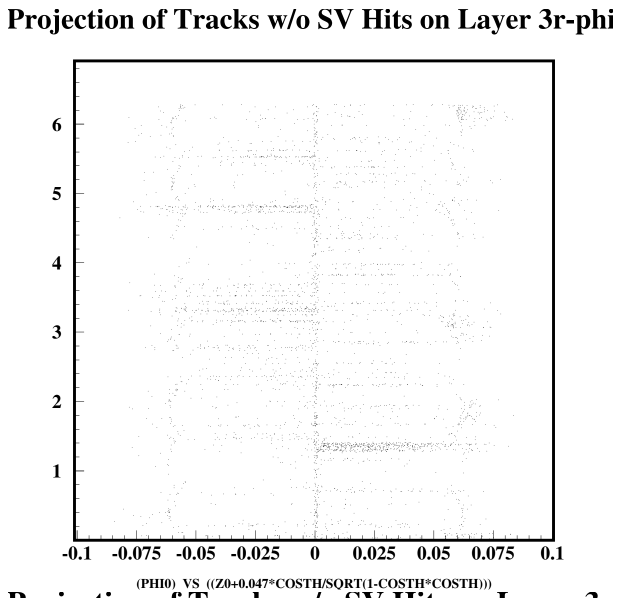


Figure 2.5: Location, in ϕ and z, of tracks without hits on layer 2. Location of missing $r-\phi$ hits is on top, $r-z$ underneath.



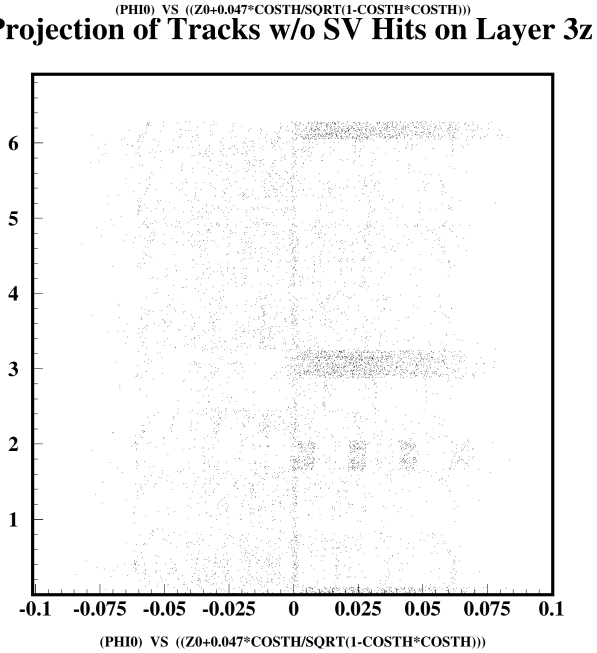


Figure 2.6: Location, in ϕ and z, of tracks without hits on layer 3. Location of missing $r-\phi$ hits is on top, $r-z$ underneath.

- There are obvious stripes, of constant ϕ on the $r-\phi$ side, and constant z on the $r-z$ side, which correspond to known bad strips, camexes, and hybrids.

- There is no flat background of missed hits.

There are two important implications for this study and for tracking in general. To take the more general implication, the fourth item, taken together with the first three, strongly imply that *when a layer does not have a hit on the track, there is a good reason for it.* Either the track went through a bad region or it didn't go through the wafers.

Secondly, in order to compare to hardware efficiencies, we must restrict ourselves to the fiducial volume of the detector. In particular, we need to be well away from the cracks. Note that though the cracks are only 2 mm wide, if a track goes through the cracks, it has only DR z information available to it. This means that due to the DR Z0 resolution (about 2 mm), it will often appear as if the track in fact passed through the silicon. In order to avoid this, we require that the tracks extrapolate at least 5mm away from the central crack on all three layers.

Due to the long size of the beam spot, tracks can also entirely miss the end of SV detector. We avoid this, again, by requiring that the track extrapolate at least 5mm away from the end of the wafers in all three layers.

2.4.4 SV Efficiencies

As mentioned earlier, a perfect detector would have 3 hits each in the $r-\phi$ and $s-z$ view. However, the detector is of course not perfect. The silicon group has taken as its benchmark a quantity known as "the two hit per view efficiency", here abbreviated as $\varepsilon_{2/2}$. This is the fraction of tracks that have at least 2 hits in the $r-\phi$ view and at least 2 hits in the $r-z$ view. This is the efficiency to put SVX hits on a well-found track, *not a track-finding efficiency*!

Note that the fiducial cuts (5 mm away from the central cracks and from the ends) are for evaluating the question, "Is there any unexplained loss in efficiency?" This does not mean that they are useful for answering the question "What is the efficiency for the tracks in my analysis?" In the remainder of the note we will give two efficiencies: one with the fiducial cut, and one without. Remember that the first one is relevant to understanding where efficiency losses are occurring, and the second relevant to understanding the overall efficiency for tracks in an analysis.

Hardware Efficiency

The hardware efficiency of the SVX detector is calculated in a separate job by using bhabha events [29]. The bhabha hardware efficiencies for the beginning of the 4sM, near to the data used in this study, are shown in table 2.2. The efficiencies can also be calculated octant-by-octant, and combined to form an expected value of $\varepsilon_{2/2}$. The two hit per view efficiency should be 96%. This procedure has an estimated systematic uncertainty of roughly 1% on $\varepsilon_{2/2}$ [29].

μ-pair Efficiency

μ pairs, of course, should be well-reconstructed. If the tracking fails on high-momenum, isolated tracks, then we have big problems. Recall, though, that `SVTF`

Table 2.1: $\varepsilon_{2/2}$ for μ pair tracks with and without fiducial cuts

Requirement	Tracks	Tracks with 2/View	$\varepsilon_{2/2}$
Good Track	41790	39205	93.8 $\pm$ 0.1 %
Good Track + Fiducial	32215	30952	96.1 $\pm$ 0.1 %

does not generate track segments for μ pairs, so that the SV tracking really consists only of the DUET SV pattern recognition. On the other hand, recall from figure 2.1 that at these momenta, we expect less than a 1% contribution from `SVTF`.

To find the total SV tracking efficiency on μ tracks, we select all the tracks that pass the good tracks criteria defined in subsection 2.4.2. The efficiency $\varepsilon_{2/2}$ is defined as

$$\frac{\text{Good tracks w/ 2 hits per view}}{\text{Good tracks}} \tag{2.1}$$

We also calculate the efficiency when fiducial cuts have been imposed on the tracks. This $\varepsilon_{2/2}$ is directly comparable to the hardware efficiency. The results for both sets of tracks are shown in table 2.1. The errors shown are statistical only.

We can also calculate the efficiency layer-by-layer to compare to the hardware layer-by-layer efficiency. For a given layer, we define the efficiency to be

$$\frac{\text{Good tracks w/ 3 hits in view}}{\text{Good tracks w/ 3 hits} + \text{Good tracks w/ hits on other 2 layers}} \tag{2.2}$$

Layer-by-layer efficiency can also be estimated from the bhabha sample, with an estimated systematic uncertainty of 1%. The comparison is shown in table 2.2.

Table 2.2: Total efficiency to find hits in a layer, compared to hardware-only efficiency

Layer	Total ε/track	ε/track from [29]
1rϕ	98.4	98.2%
2rϕ	97.5	97.2%
3rϕ	90.9	93.4%
1z	96.2	95.5%
2z	91.0	88.6%
3z	80.3	84.5%

It is apparent from tables 2.2 and 2.1 that μ pair tracks are essentially 100% efficient relative to the measured hardware efficiencies. Or, put otherwise, there is no evidence for software losses in efficiency for μ pairs greater than 1%.

Low-Momentum Efficiency

If there are efficiency losses at momenta lower than 5.3 GeV, we would like to separate out those losses due to hadronic event environments and those inherent to lower momentum–multiple scattering, curling, etc. In order to accomplish this, we make use of both the $\gamma\gamma \to e^+e^-$ and $\gamma\gamma \to 4\pi$ skims. The momentum spectra of these events are shown in figures 2.7 and 2.8.

It is important to recall that `SVTF` is essentially never run on the $\gamma\gamma \to e^+e^-$ skims, and not always on the $\gamma\gamma \to 4\pi$ skims. We select those $\gamma\gamma \to 4\pi$ events for which `SVTF` was run, independent of whether any of the silicon hits in the event

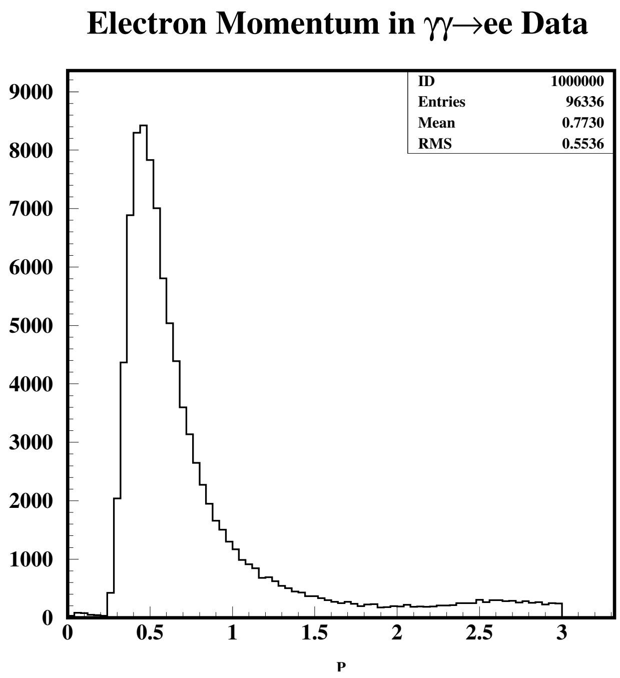


Figure 2.7: Momentum of tracks in the $\gamma\gamma \to e^+e^-$ skim

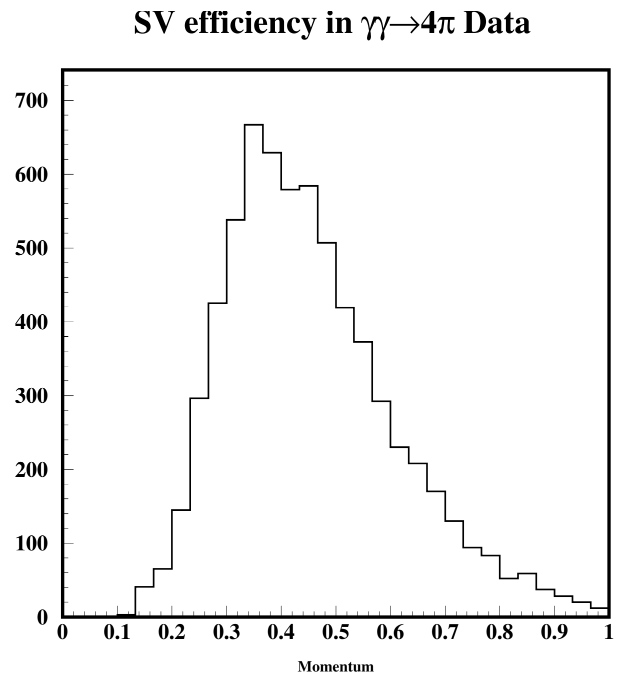


Figure 2.8: Momentum of tracks in the $\gamma\gamma \to 4\pi$ skim

Table 2.3: SV efficiency in $\gamma\gamma \to e^+e^-$ events

Requirement	Tracks	Tracks with 2/View	$\varepsilon_{2/2}$
Good Track	88388	81070	91.7 $\pm$ 0.1 %
Good Track + Fiducial	66124	62212	94.1 $\pm$ 0.1 %

Table 2.4: SVX efficiency in $\gamma\gamma \to 4\pi$ events.

Requirement	Tracks	Tracks with 2/View	$\varepsilon_{2/2}$
Good Track	9102	8146	89.5 $\pm$ 0.4 %
Good Track + Fiducial	6748	6287	93.1 $\pm$ 0.4 %

were actually attached by `SVTF`.

We repeat the analysis that we undertook for the μ pairs for total efficiencies. The results are shown in table 2.3. In addition, for $\gamma\gamma \to e^+e^-$ events we can calculate $\varepsilon_{2/2}$ as a function of momentum from 500 MeV to 3 GeV (below this it becomes important that `SVTF` was not run). As figure 2.9 shows, the SV efficiency is quite flat as a function of momentum down to 500 MeV.

We perform the same analysis on $\gamma\gamma \to 4\pi$ events, whose momentum spectrum peaks at 500 MeV. Table 2.4 shows the results.

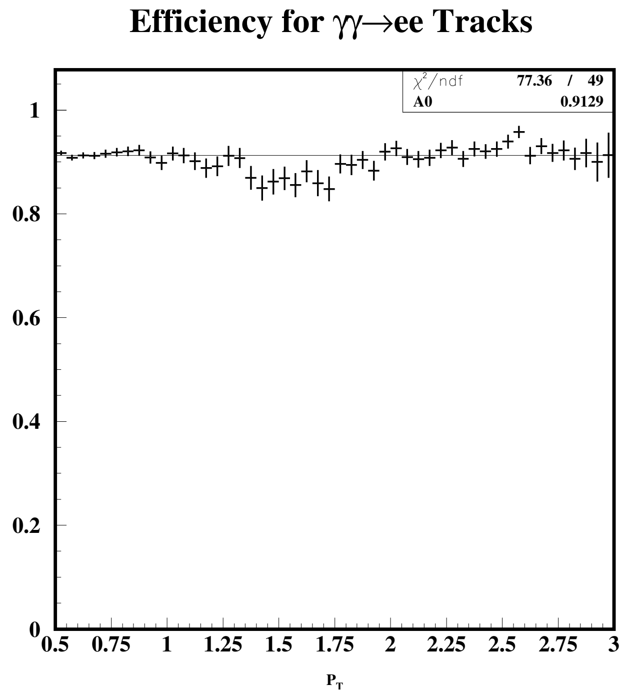


Figure 2.9: $\varepsilon_{2/2}$ vs. p_T for $\gamma\gamma \rightarrow e^+e^-$ tracks. There is no fiducial volume cut. The fit line is 91%.

Table 2.5: $\varepsilon_{2/2}$ for tracks from the decay daughters of the D^0 in $D^{*+} \to D^0\pi^+$. The momentum spectrum is flat from 500 MeV to 3 GeV

Requirement	Tracks	Tracks with 2/View	$\varepsilon_{2/2}$
Good Track	3083	2822	91.5 $\pm$ 0.5 %
Good Track + Fiducial	2419	2270	93.8 $\pm$ 0.5 %

Hadronic Environment

Finally we come to the the question of silicon efficiency in hadronic events, which is the quantity that matters to most analyses. For this purpose we use continuum $D^{*+} \to D^0\pi^+$, $D^0 \to K\pi$ candidates. $D^{*+} \to D^0\pi^+$ candidates produce two sets of kinematically distinct tracks: soft pions, with momentum from 0.1 to 0.4 MeV, and π and K daughters from D^0 decay, which have momenta flatly distributed from 0.4 to 3 GeV. We treat these two sets of tracks separately, in order to specifically extract values of $\varepsilon_{2/2}$ for soft and hard tracks in hadronic events. Once again we calculate $\varepsilon_{2/2}$ with and without fiducial cuts. The results are shown in tables 2.5 and 2.6. We can combine the π_s and π/K samples to calculate the efficiency in the fiducial volume as a function of transverse momentum (the quantity relevant for curling). This is shown in figure 2.10. As the figure shows, the SV efficiency is very flat down to 150 MeV, tailing down to 80% at 100 MeV.

Note in particular that the efficiency for the D^0 daughter tracks is almost the same as that of the $\gamma\gamma \to e^+e^-$ events between 500 MeV and 3 GeV. This allows

Table 2.6: $\varepsilon_{2/2}$ for π_s tracks in $D^{*+} \to D^0\pi^+$. The momentum spectrum reaches from 100 to 400 MeV, peaked at 250.

Requirement	Tracks	Tracks with 2/View	$\varepsilon_{2/2}$
Good Track	2650	2362	89.1 ± 0.7 %
Good Track + Fiducial	2059	1875	91.0 ± 0.7 %

us to measure the environmental effect in this momentum range : 0.4±0.6%. The comparison below 500 MeV is best made with $\gamma\gamma \to 4\pi$ events, as the $\gamma\gamma \to e^+e^-$ events have not had `SVTF` run on them. Here we measure the environmental effect on SV efficiency to be 2.1±0.8 %.

2.4.5 Resolution

We can measure the resolution of the detector using events which we know to have no lifetime component. Our track samples provide us with three such event types:

- μ pairs
- $\gamma\gamma \to e^+e^-$
- $\gamma\gamma \to 4\pi$

Recall that impact parameter resolutions, *for fixed intrinsic hit resolution*, are expected to be of the general form

$$a \oplus \frac{b}{p\beta \sin^{\frac{3}{2}}\theta}(r-\phi) \tag{2.3}$$

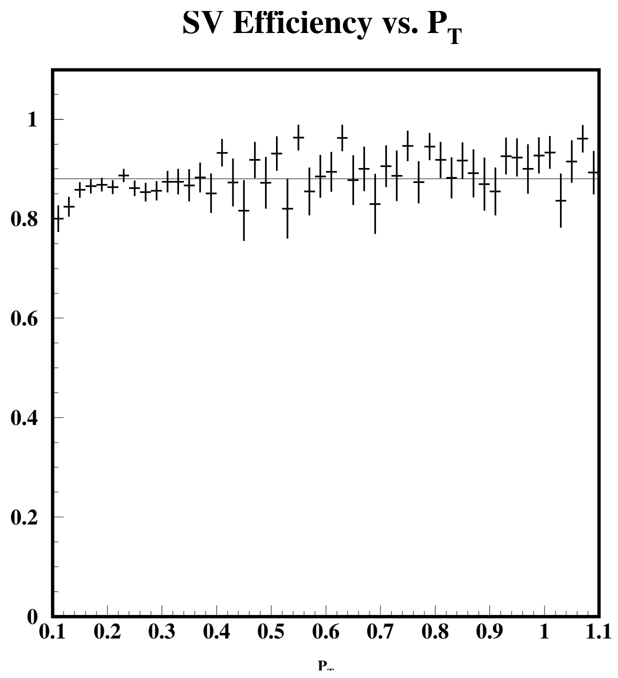


Figure 2.10: $\varepsilon_{2/2}$ vs. p_T for D^* tracks. There is no fiducial volume cut. The fit line is 89%.

$$c \oplus \frac{b}{p\beta \sin^{\frac{5}{2}}\theta}(r-z) \tag{2.4}$$

where θ is the dip angle of the track. Note that this reduces to the familiar form at normal incidence; the reason for using the formulas is to improve statistics by using non-normal tracks. However, the formulas do not hold in z, as the detector resolution is itself a function of θ. Because of this, we use only normally incident tracks (sinθ >0.96) for measuring the z-side impact parameter resolution.

In order to measure the impact parameter with the track-pair events (μ pairs and $\gamma\gamma \to e^+e^-$), we calculate the value of $p\beta\sin^{\frac{3}{2}}\theta$ for the $r-\phi$ impact parameter measurement, and simply p for the z measurement, for both tracks in the event. These we call "psuedo-momentum". If they are similar (within 5%) then we combine the two tracks and the known beam position for the run to form a vertex position. We then calculate the distance between the two tracks at the vertex in the $r-\phi$ and $r-z$ planes. This, divided by $\sqrt{2}$, is the miss distance for the pair. The miss distance is recorded as a function of the psuedo-momentum relevant for the view.

For the $\gamma\gamma \to 4\pi$ events, one track is selected as the "probe" track. We select the first positively charged track in the list. The other three tracks are combined, along with the known beamspot for the run, to form a vertex. The other three tracks are required to have 2 SV hits in each view, the χ^2 of the vertex fit is required to have probability greater than 10%, and the error on the $r-\phi$ position of the vertex is required to be less than 30 μm, and the z less than 100 μm. We then calculate the miss distance of the probe track with respect to the vertex. This is also recorded as a function of the psuedo-momentum relevant to the view.

Figures 2.11 and 2.12 show the results of this procedure, where miss distances from all three event classes have been used.

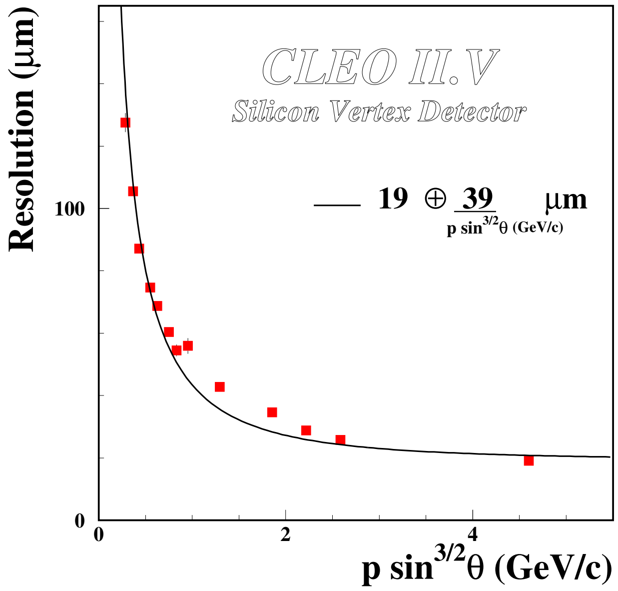


Figure 2.11: The $r - \phi$ impact parameter resolution

2.4.6 Noise

By "noise" we mean here the fraction of hits that are put on tracks that do not belong on them. Of course, defining "do not belong" can be a slightly subtle business. For the course of this discussion, an analyzed hit "belongs" on a track if any of the raw hits it is clustered from truly come from the track. This is the definition used both by the Monte Carlo tagger and the hardware noise measurement.

How can we discover the fraction that "truly" come from the track when all we have are residuals? Note especially that in this definition of "noise" hits, many

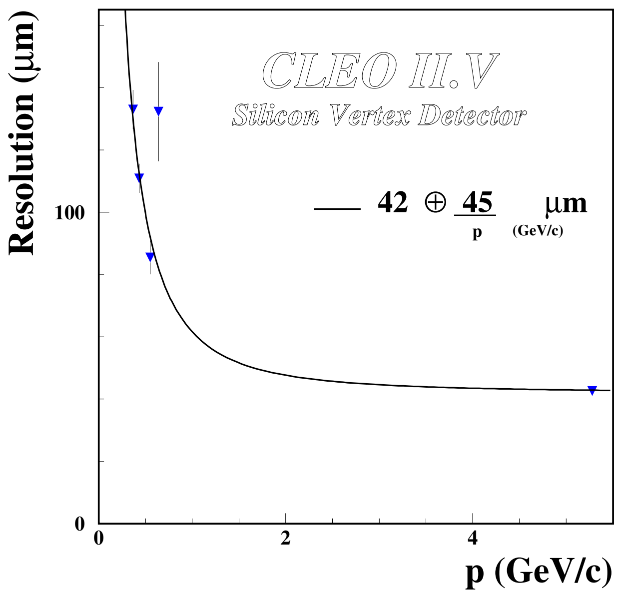


Figure 2.12: The $r - z$ impact parameter resolution

non-noise hits can look very bad. For instance the following categories of mishaps are not noise, but generate large residuals:

- Tracks with a hard scatter
- Analyzed clusters containing raw hits from a hard δ ray
- Analyzed clusters containing both true hits and noise hits
- Analyzed clusters adjacent to a dead channel which should have fired but didn't
- Analyzed clusters in which a true hit has been rejected due to low signal-to-noise

In all of these cases residuals can look bad even though everything was done "right". To put this in perspective, the rate for any of these things happening in an individual hit is estimated to be 5% [30].

To finesse the question, we need to find a sample of hits which we know independently "have no noise". Then we can attribute the large-residual fraction of this sample to the mishaps above. Then we can move to other samples, and attribute any difference in large-residual fraction to noise. In $r-\phi$, given the measured efficiencies and electronic noise rates, we can calculate that 0.07% of hits in 3-hit tracks are due to noise. In $r-z$, we calculate that 0.1% of hits on 3-hit tracks are due to noise. These fractions are sufficiently low that we take 3-hit tracks as our "no noise" sample.

We define the following measure of noise: for the $\gamma\gamma \to 4\pi$ "probe" tracks, we will measure the fraction of tracks which miss the vertex by more than 10σ, where σ is the reported fit error times that measured error inflation scale factor as determined from D^0 lifetime studies [3] The results are shown in table 2.7 below.

The table shows two things clearly. First, the tracks with 3 z hits are worse than the tracks with 3 $r-\phi$ hits. As $\frac{S}{N}$ is much lower on the z side, and the dead channel rate much higher, this is not surprising. It will happen more often that there is a clustering error (raw hits either not found or rejected due to low S/N).

The table also shows that there is really no strong evidence that when there are 2 z hits the hits are more likely to contain noise. The difference between 3 and

[3] At the time of the writing, this was 1.8 for $r-\phi$, 1.4 for $r-z$. A reconstruction bug was found to be responsible for this and scale factors in the data are now near to 1.

Table 2.7: The fraction of $\gamma\gamma \rightarrow 4\pi$ tracks that miss the event vertex by more than 10 σ in $r-\phi$ and z, broken down by the number of SV hits in the appropriate view.

View	SV Hits on Track	Total Tracks	Tracks beyond 10 σ	10 σ Fraction
$r-\phi$	3	1260	64	5.0±0.6 %
$r-\phi$	2	346	26	7.5±1.5 %
$r-z$	3	850	95	11.2±1.1%
$r-z$	2	652	64	9.8±1.2%

2 hit tracks is 2.5±1.6 % in the $r-\phi$ view, and -1.4±1.7% in the $r-z$ view. Of course, the measurement statistics are not great. However, it seems we can set a 95% (one-sided 1.6σ) upper limit on the fraction of tracks with a noise hit in $r-\phi$ and $r-z$ of 1.3% and 0.9%.

2.4.7 Systematic Uncertainties

All errors contained in this note have been statistical errors only. As this is a measurement of detector performance, rather than of physics, very little unfolding of variables is required. Systematic uncertainties enter predominantly in the choice of what to measure, rather than in the value of any particular measurement. The reader must judge for him or herself whether the quantities measured here are the quantities they would like to know the values of.

Choice of Dataset

There is one overall systematic effect for all of the measurements: they all use the end of the 4sL, rather than the entire CLEO II.V dataset. The silicon detector during its first year of commisioning was a very dynamic system. The hardware efficiency is known to vary dataset to dataset; the variation in tracking software efficiency is not known. To address this uncertainty, we evaluated the total efficiency for μ pairs in the 4sK dataset, run 83276. Tony Hill calculated the expected value of $\varepsilon_{2/2}$ for this dataset, finding 93% with an uncertainty of 2%. For the tracks within the fiducial volume, we find for μpairs $\varepsilon_{2/2}$ of 89.1 $\pm$ 1.0 %. This tells us that to 4$\pm$2%, the total efficiency can be explained by the hardware efficiency in 4sK as well as 4sL.

Systematics of the Efficiency measurements

Systematic uncertainties enter into the values of the efficiencies in the following ways:

- Purity of the samples
- Whether `SVTF` was run on the samples

Systematic Uncertainty: Sample Purity

For instance, $\pi^+\pi^-$ contamination of the $\gamma\gamma \to e^+e^-$ sample can be measured in the data (by beating time-of-flight against $\frac{dE}{dx}$) to be smaller than 1 part in 1,000.

τ contamination can be measured in generic τ Monte Carlo as less than 0.1%. Radiative bhabha contamination is unknown but also unimportant, as such events also originate from the IP.

τ contamination of the $\gamma\gamma \to 4\pi$ is also known from MC to be less than 0.1%. Hadronic annihilation contamination is small due to the low (6 GeV) energy of the selected events.

From the sidebands of the $D^{*+} \to D^0\pi^+$ sample, we estimate the contamination in our sample to be less than 1%.

Systematic Uncertainty: Running `SVTF`

Strictly speaking, this is not a systematic uncertainty: it is well-known, for a given event, whether `SVTF` generated track segments for the event. Nonetheless, it is a potential route for discrepancies to arise, so we list it as a systematic primarily as a warning to others.

Systematics of the Resolution Measurements

Two important effects come into this measurement, but we believe that they are small. This belief is bolstered by the recognition that data from all three data samples–$\gamma\gamma \to e^+e^-$,$\gamma\gamma \to 4\pi$, and μ pairs– lie on the same curves.

- Size and uncertainty of the IP in 4π events
- Choice of where to evaluate the Z miss distance in pair events

The IP formed from the "other 3" tracks in 4π events can have a rather large error, expecially in the Z view. In the $r - \phi$ view it is a rather small ellipse, due to the smallness of the beam in that plane. Therefore, we believe that the z miss distance in these events is affected at less than the 10% level–on average, the size of the IP error in z is less than 1/2 that of the track error.

We evaluate the Z miss distance by finding each track's 2-D point of closest approach to the vertex formed by the two tracks and the beam. Pablo Hopman has evaluated different schemes for choosing the point at which to evaluate them, and finds differences only at the μm scale [31].

Noise

We have arbitrarily chosen 10σ as the "noise" portion. The fraction obviously changes as a function of this choice, but the reader knows where the cut is. The greater uncertainty is not in the value of the measurements, but rather in interpreting them. For instance, the reader may not agree with our definition of noise. We must especially warn the reader to be carefully suspicious of the claim that "3 hit tracks are perfect", though we have found no good reason to believe otherwise.

2.5 Hadronic Interactions at the Beampipe

One question that came up in the course of the D lifetime analysis [39] was, as Harry Nelson put it, "Is our definition of a micron 1% different than the NIST micron?". In order to explore issues of the length scale in the inner detector, we examine two data samples. We select a sample of $\gamma\gamma \to 4\pi$ events and a sample of hadronic nuclear interactions with the detector material. The $\gamma\gamma \to 4\pi$ events quantify our ability to measure something at distance 0 from the beamspot. The hadronic interactions, as they occur at the beampipe, measure our ability to reconstruct lengths at distance 2 cm from the beamspot.

Most hadronic interactions occur where the detector has the most hadrons, that is, the beampipe and silicon. As we require 2 SV hits, we only have high efficiency for layer 1 of the SV. Our plan is to select a sample of events which we can conclusively demonstrate (through the $p/\overline{p}$ ratio) are due to interactions with the detector material, examine where these interactions occur, then compare to known properties of the beampipe.

2.5.1 Selection

We select hadronic interactions by searching for protons that miss the beamspot[4]. We use the `hadsel` routine, modified to take species multiplicity into account, to identify protons. We look for those tracks most likely to be protons, and require at least a 1% probability to be protons. We do not make any requirement on the charge of the proton, and through this section, "proton" will be taken to mean proton or anti-proton. We require `DBCD` greater than 5 mm, at least 20 DR hits, `KINCD .GE. 0`, and 2 SV hits per view.

If we find a proton, we attempt to find additional tracks, not required to be

[4] As there is very little Q value in Λ decay, the protons and anti-protons from these decays nearly always point at the beamspot, and are not generally selected by the algorithm

identified as protons, which form a good vertex (probability of χ^2 at least 0.1%) with the proton. For these additional tracks, we impose the same track quality requirements, except in relaxing the `DBCD` cut to 1 mm.

We list all tracks which satisfy these requirements. We form a vertex from the proton and this list of tracks. On average, the proton and 1.3 additional tracks are used.

We require the vertex so formed to have at least a 5% probability of χ^2. In figure 2.13 we show the x and y locations of all vertices formed in the 4sM data set, requiring the reported radial resolution to be less than 200 μm. In figure 2.14 we show the r-z locations of the vertices, where r carries the sign of y. The beampipe, the first layer of silicon, and the CAMEX stainless steel shielding ring (described in [32]) are all apparent in these plots.

2.5.2 Anti-protons

At this point we check how many of our events are protons and how many anti-protons. We find 167 anti-protons and 3186 protons. This demonstrates that our technique selects interactions with the detector, rather than "physics". We require from now on that "protons" have charge +1, with confidence that there is little background.

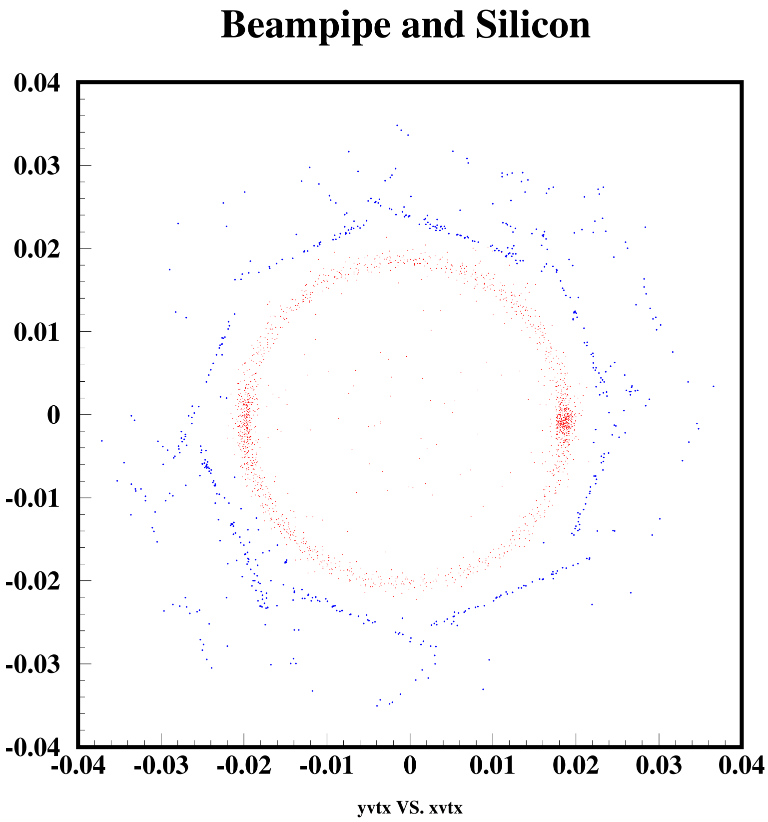


Figure 2.13: $x - y$ location of selected vertices. Clearly visible are the circular beampipe and the pinwheeled wafers of layers 1 of the SVX. A ghostly layer 2 is also visible.

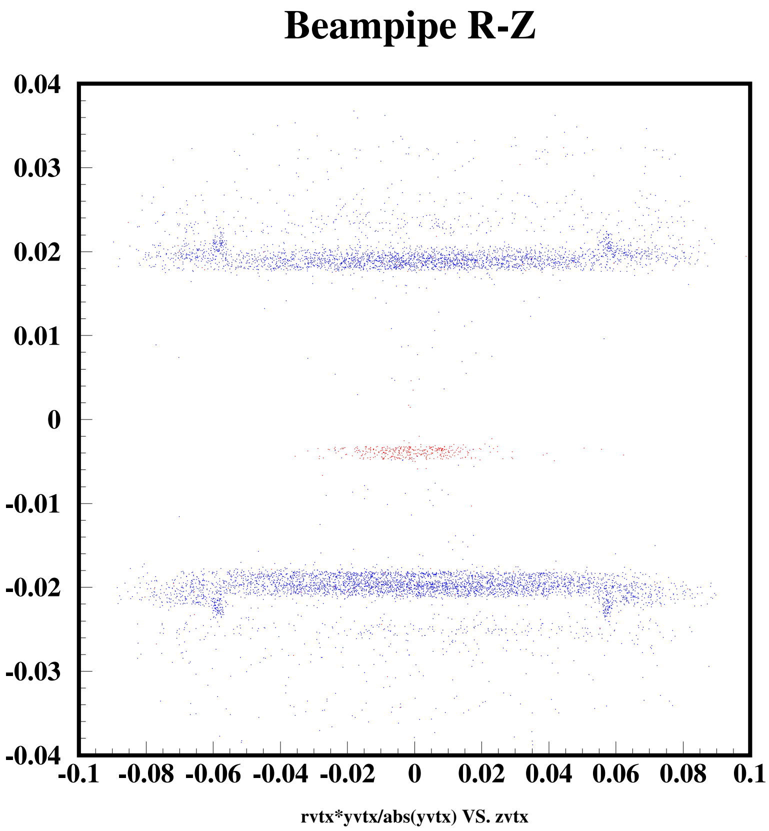


Figure 2.14: $r - z$ location of selected vertices. r is signed by y. The luminous region is shown in the center; the "knobs" located at ± 5.6 cm in z are the CAMEX shielding rings.

2.5.3 What are these events?

Table 2.8 shows what kinds of events these are. Three-quarters of the events—classified by EVCL as two-photon, junk, bad z vertices, or beam gas—are classic beam-wall events. That is, these events contain no activity other than the hadronic interaction in the detector material. These events show a strong angular distribution in the detector, as shown in figure 2.15. The remaining quarter of the events come from $e^+e^- \to$ hadrons, where one of the hadrons initiates the interaction in the detector material. KLASGL = 10 has been subdivided into two subclasses–those in which the vertices formed have 2 tracks, and those in which there are 3. From scanning, events with 3 tracks in a vertex are often only a single beam-wall event, which have been misclassified as KLASGL=10. Those events with two tracks are generally truly produced by a hadron from e^+e^- annihilation. The interactions initiated by hadrons from $e^+e^- \to$ hadrons are, as expected, produced uniformly in azimuth.

In order to avoid a bias toward a particular part of the detector (which would spoil azimuthal averages), in the remainder of this note we require the reconstructed vertex ϕ to be at least 0.5 radians away from 0 or π.

Two typical events are shown in figures 2.16 and 2.17.

2.5.4 Centering and Squash

For all measurements of centering, squash, and radii of the beampipe, we require the z-coordinate of the vertex to be within 5 cm of the origin. This eliminates

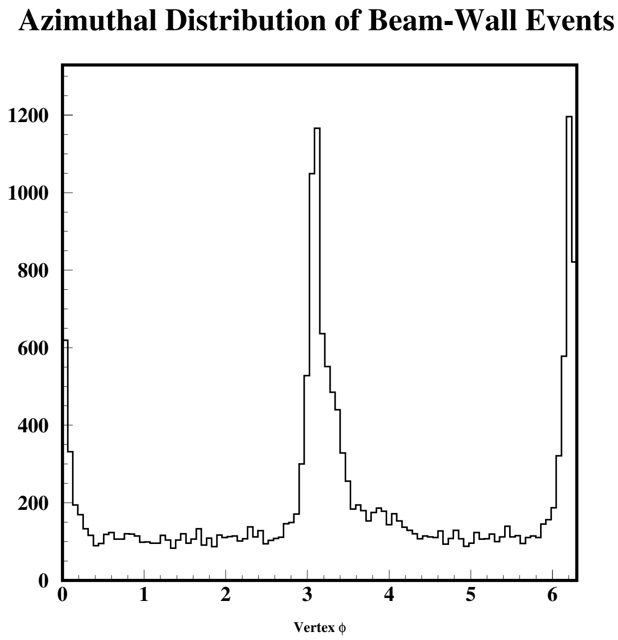


Figure 2.15: ϕ of beamwall vertices.

Table 2.8: EVCL classification of selected events.

KLASGL	Type	Events	Fraction
1	Junk	899	8%
2	Cosmic	65	0.5%
3	Beam Gas	1885	16%
4	Two photon	4108	36%
5	e pair	8	-
6	mu pair	3	-
7	gamma gamma	0	-
8	ee gamma	34	-
9	Unknown	0	-
10(2 tracks in vertex)	Hadrons	3177	27%
10(3 tracks in vertex)	Hadrons	633	5%
11	Bad z hadrons	710	6%

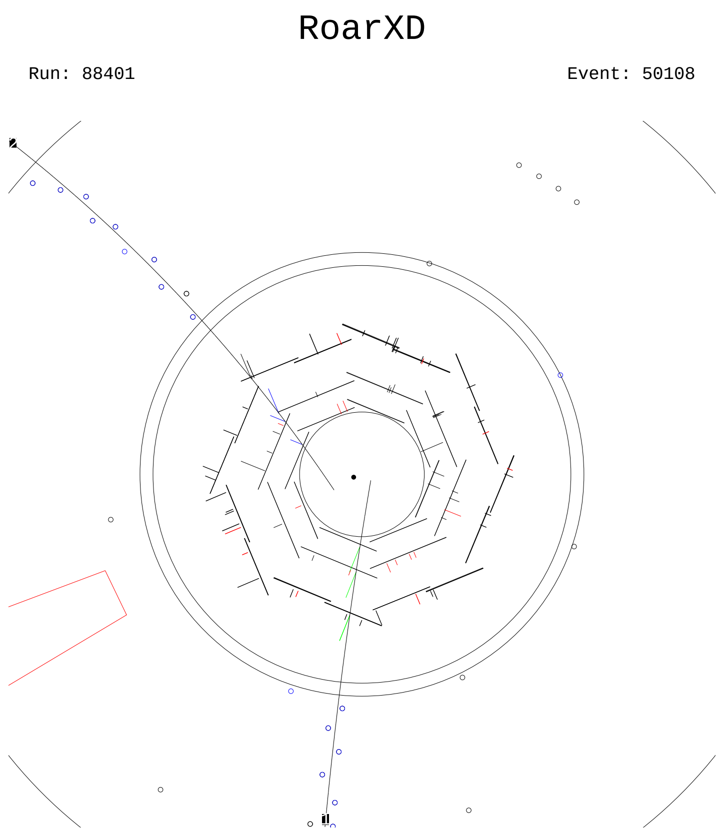


Figure 2.16: A beam-wall event. Both tracks are identified as protons.

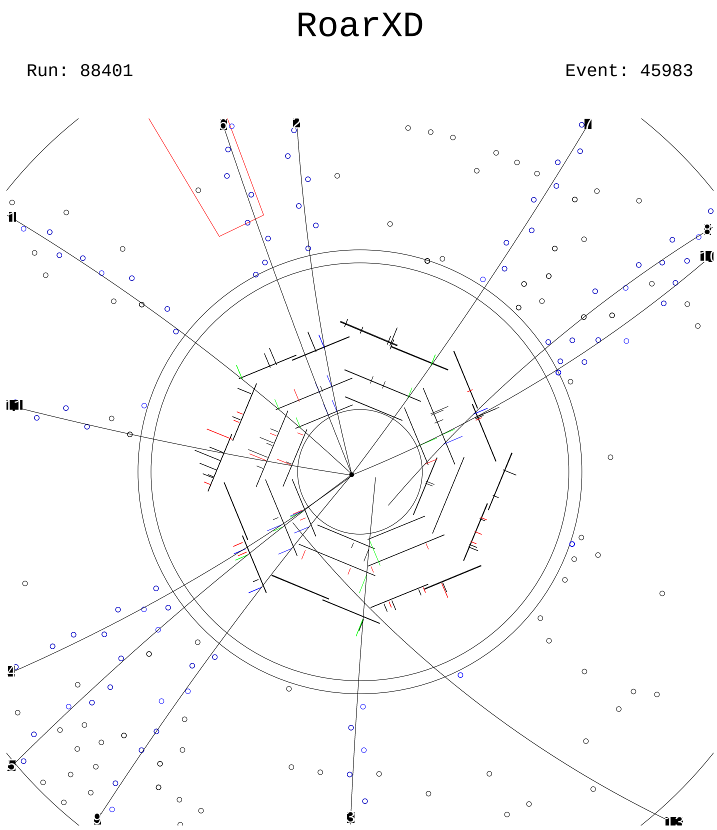


Figure 2.17: A hadronic interaction embedded in real annihilation event. The hadronic interaction occurs near 6 o'clock on the beampipe; the proton, labelled as #3, is traveling toward 6 o'clock, the other track, labelled as#8, travels toward 2 o'clock.

those vertices due to interactions in the CAMEX shielding ring on the radiation shields.

First we determine the center of the circle formed by the beampipe. To do so, we make a PAW profile plot of r vs. ϕ for all vertices for which $0.17 < r < 0.215$. Though we don't really know what this average corresponds to (density-average? radiation length average? nuclear-cross-section average? etc.), we make the assumption that the correspondence is not a function of ϕ. The result is shown in figure 2.18. We find the center displaced by 810 μm in y and 495 μm in x.

Next we check for a period π oscillation in r, which would indicate a squash. We show this in figure 2.19. We find a squash magnitude of 67$\pm$29 μm. We conclude the detector is somewhat, but not severely, out of round. This magnitude is consistent with the magnitude of ovality reported in [33], which is 42 μm.

2.5.5 Z tilt

We can measure the alignment of the beampipe with respect to the z axis. A crude measure is to take events in each of four quadrants corresponding to x or y greater or less than 0, and plotting the mean radius against z. We show these plots in figure 2.20. We see no strong evidence for a tilt. Note that though quadrants 1 and 3 appear to have the same 2σ significant slope, in fact there should be a relative - sign between them for a tilt. If taken seriously, these plots show a conical beampipe shape (at the 50 μm level).

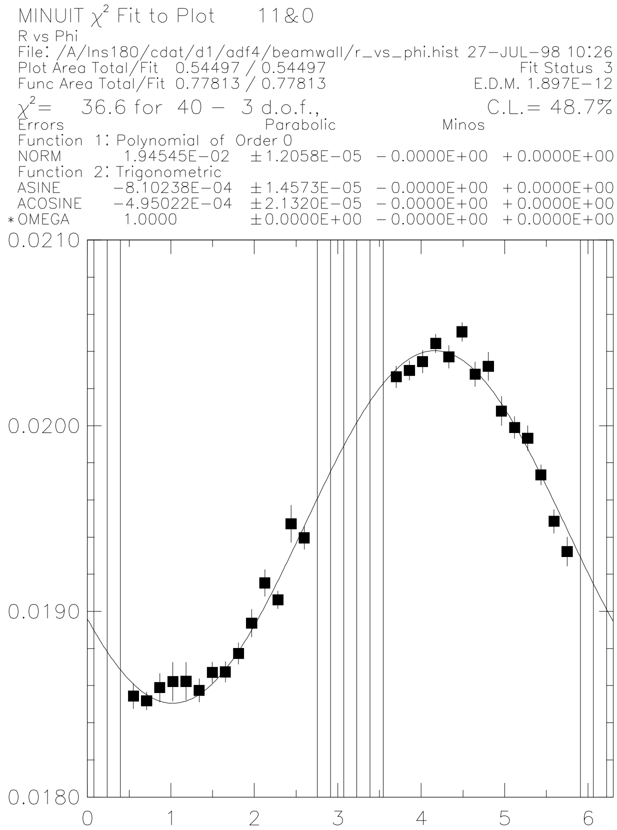


Figure 2.18: r (measured from the CLEO origin) as a function of ϕ. Shown is the fit to a period 2π oscillation.

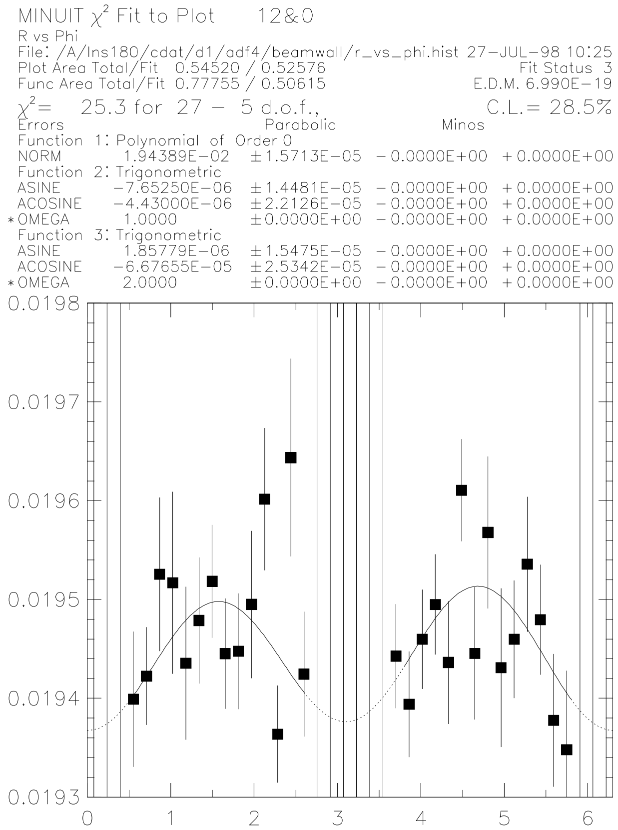


Figure 2.19: r (measured from the beampipe center in figure 2.18) as a function of ϕ, with fits to period 1 and period 2 oscillations

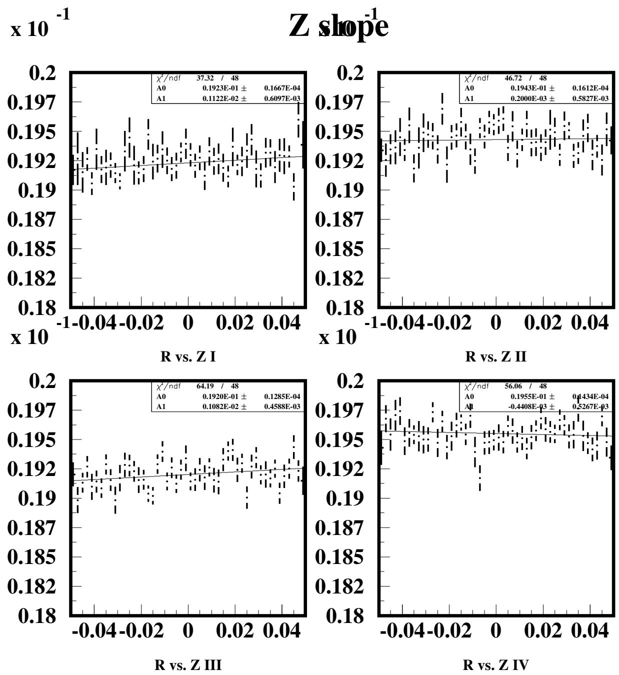


Figure 2.20: r (measured from the beampipe center in figure 2.18) as a function of z of the vertex, for the four quadrants in the $x-y$ plane. Shown are linear fits to the plots.

2.5.6 Beampipe Radii

Having found the center and verified that the beampipe is round, we need a more sophisticated method than the PAW profile plot in order to find the inner and outer radius. We must be sure that we understand clearly what quantity we are measuring.

We plot the radius of all vertices between 17 and 22.5 mm from the determined center point of the beampipe. As expected, we find the beampipe. We attempt to fit this distribution of vertex radii to a beampipe shape.

We write a fitting function for MN_FIT. The fit consists of a beampipe signal shape and a background/mismeasurement gaussian "hump". The beampipe signal shape has the following fit parameters:

- Overall normalization
- Inner Radius
- Outer Radius
- σ of a resolution gaussian function which is convolved with box shape.
- Normalization of Be, relative to H_20
- Normalization of Au, relative to H_20

In addition to the floating inner and outer radii parameters, the beampipe is divided in the ratio 0.1:4:5:3 to be of density Au:Be:1.0:Be, that is, Au:Be:H_20:Be.

To review, the signal shape is a box shaped as in figure 2.21 convolved with a gaussian. This gaussian accounts mainly for vertex resolution, and also for additional slop such as residual miscentering and out-of-roundness.

Recall that this all sits on top of a gaussian background hump. In reality this hump is probably the wide part of a double gaussian which we should convolve with the signal shape, but this seems good enough.

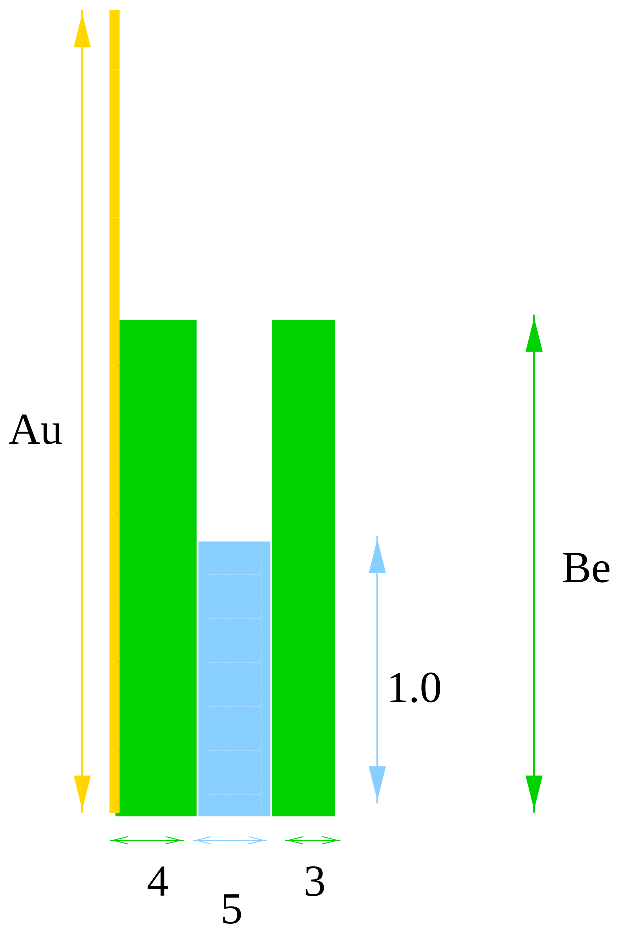


Figure 2.21: The beampipe shape used in the fit. The gold inner coating, shown on the furthest left, is assigned a thickness of 0.1.

Table 2.9: The fitted beampipe parameters, for vertices with small reported uncertainties.

Parameter	Fitted Value	Uncertainty	Units
σ	128	16	μm
R_{in}	18,865	3	μm
R_{out}	20,106	5	μm
Bkgd. Norm.	592	258	Events
Bkgd. Mean	19,453	40	μm
Bkgd. σ	532	59	μm
$\frac{\sigma_{Be}}{\sigma_{H_20}}$	2.2	0.6	–
$\frac{\sigma_{Au}}{\sigma_{H_20}}$	8.2	3.5	–

2.5.7 Results

We perform the measurement 2 ways, in order to probe systematic uncertainties. In the primary method, we use only vertices for which the reported radial resolution is less than 200 μm. For these events the average uncertainty is 99 μm. In the second method, we use all vertices; for these events, the average uncertainty is 200 μm, with considerable (up to 1 mm) spread.

The result of the fit is shown in figure 2.22 and the fitted parameters are shown in table 2.9. The $\chi^2/d.o.f.$ is 21.4 / 26, for a C.L. of 71.9%.

We find a reasonable value for the smearing gaussian, 128±14 μm. This value

Table 2.10: The fitted beampipe parameters for all vertices, without accounting for gold, and without cuts on the vertex resolution.

Parameter	Fitted Value	Error	Units
σ	207	21	μm
R_{in}	18,873	23	μm
R_{out}	20,139	15	μm
Bkgd. Norm.	654	119	Events
Bkgd. Mean	19,408	54	μm
Bkgd. σ	920	116	μm

is consistent with the average reported radial uncertainty from the vertex fit, 99 μm, combined with a 63 μm spread due to the beampipe squash.

We also fit using all vertices of any reported resolution, and without using a separate term for the gold coating. The results, which are consistent with the more precise vertices, are shown in table 2.10.

2.5.8 Resolution Sanity Check

The fit finds the width of the resolution (convolved) gaussian to be 128$\pm$16 and 207$\pm$21 μm in the two cases considered above, that is, with and without requirements on the reported radial vertex resolution. The average reported resolution was 99 and 200 μm, respectively, for the two cases. That the gaussian widths agree well with the average reported resolution confirms that we understand the

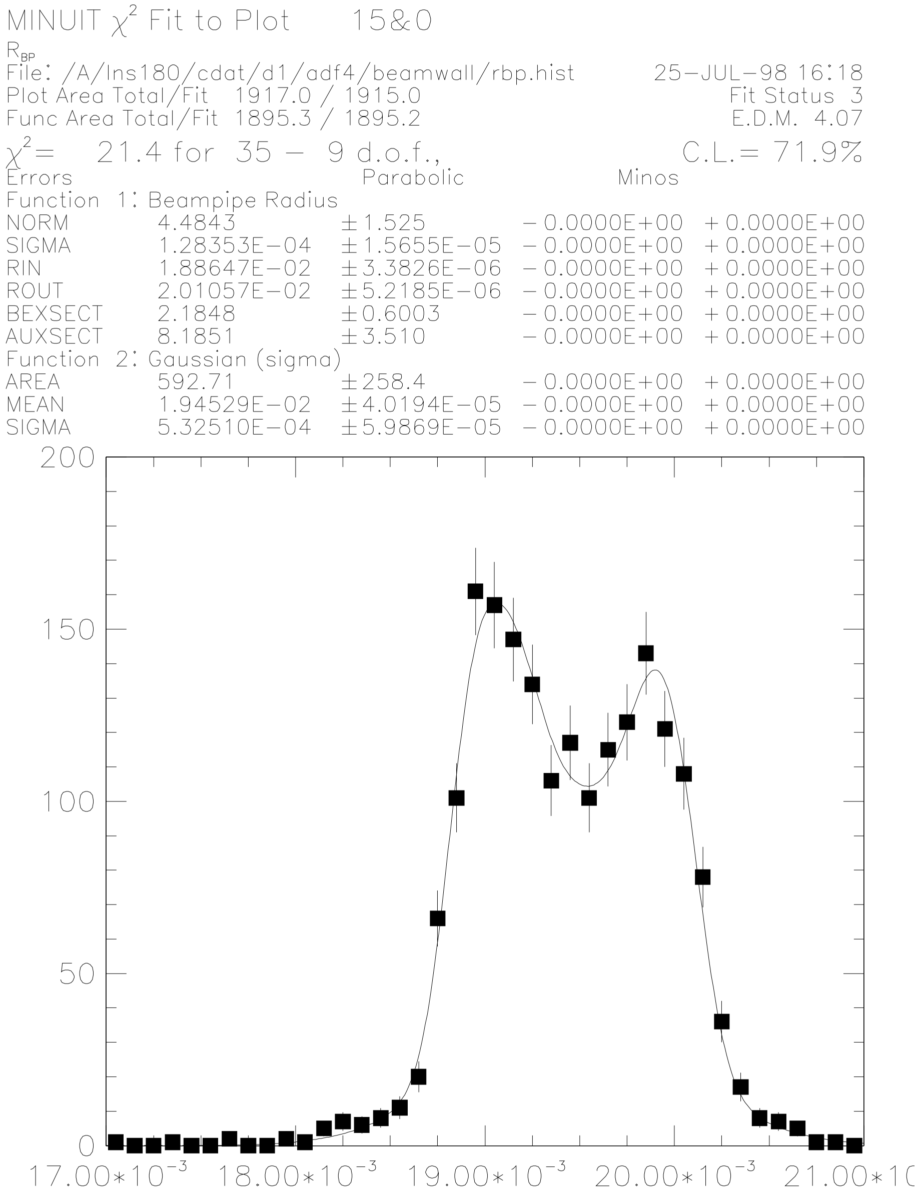


Figure 2.22: The `MN_FIT` fit to beampipe radius.

origin of the gaussian width, and incidentally that our error matrices are very well normalized.

2.5.9 The Ruler Check

The nice thing about the beampipe is that it has been measured independently. From a survey of the beampipe [33] the average outer radius is reported to average 20,112 μm. In the shield writeup [32] the outer radius is quoted as 2.0120 cm. We believe the former to be more accurate than the latter, to a few microns. A layer of epoxy [5] , of unmeasured thickness, was also added, believed to be 5 μm thick. There is also a layer “about” 55 μm thick of PEEK plastic. As PEEK has a density of 1.3 $\frac{g}{cc}$[6], this is the equivalent of 39 μm of Be. We assign an additional uncertainty of 10 μm for the PEEK thickness. We assign total uncertainty of 15 μm to the outer diameter. From surveys [33] we know that the inner radius of the spare beampipe was 1.889 cm (nominal of 18,860). We take 18,860 $\pm$ 30 μm as the inner diameter value. These are compared to the fitted values in table 2.11.

We find that the lengths have been reconstructed correctly to one part in 500, the limit of our resolution. We do not take seriously the 3 μm error on the inner radius; we estimate systematic effects from the gold to be 10 μm.

It is also fair to check the thickness reported by the fit. We find 1237 $\pm$ 16 μm.

[5] This epoxy, incidentally, consists of about 1% heavy metals. The exact composition is a proprietary secret; however, with the aid of Jeff Cherwinka [34] we confirmed that the metals are of negligible additional cross section.

[6] For reasons of national security, PEEK is a controlled material under the dual-use provisions of the 1996 Wassenaar Arrangement. We had a difficult time finding its density.

Table 2.11: Comparison of the fitted beampipe parameters to their surveyed values.

Radius	Fitted Value (μm)	Survey (μm)	Fractional Difference	Source of survey
Inner	18,865 $\pm$ 3	18,860$\pm$30	0.03$\pm$0.17%	[33]
Outer	20,106 $\pm$ 5	20,156$\pm$15	0.23$\pm$0.08%	[32],[33]

The true thickness from [35] is 1250 μm. Hence from the thickness alone we can limit the length scale to be correct to 1.0$\pm$1.3%!

2.6 Systematic Checks of the Length Scale

2.6.1 $r-\phi$ Only Fit

We re-measure the beampipe parameters using only an $r-\phi$ fit to the vertex; that is, ignoring z information. This allows us to examine whether 2-D fits to the radial position are systematically different from 3-D fits.

We perform the test on a subset of data[7], the 4sM tapes 1-12. On this subset, we form vertices in 2 or 3 dimensions, then fit as above for the beampipe parameters. As we are fitting (largely) the same data, differences are (largely) due to systematic shifts. Because a 2-D fit to 2 tracks has no degrees of freedom, it is not possible to form a probability of χ^2 for the vertex; for this reason, selection of the vertices to be used in the sample may differ slightly, and introduce small statistical differences.

[7]Only a subset due to disk space constraints on long-term storage.

The results are shown in Fig. 2.23 and Fig. 2.24.

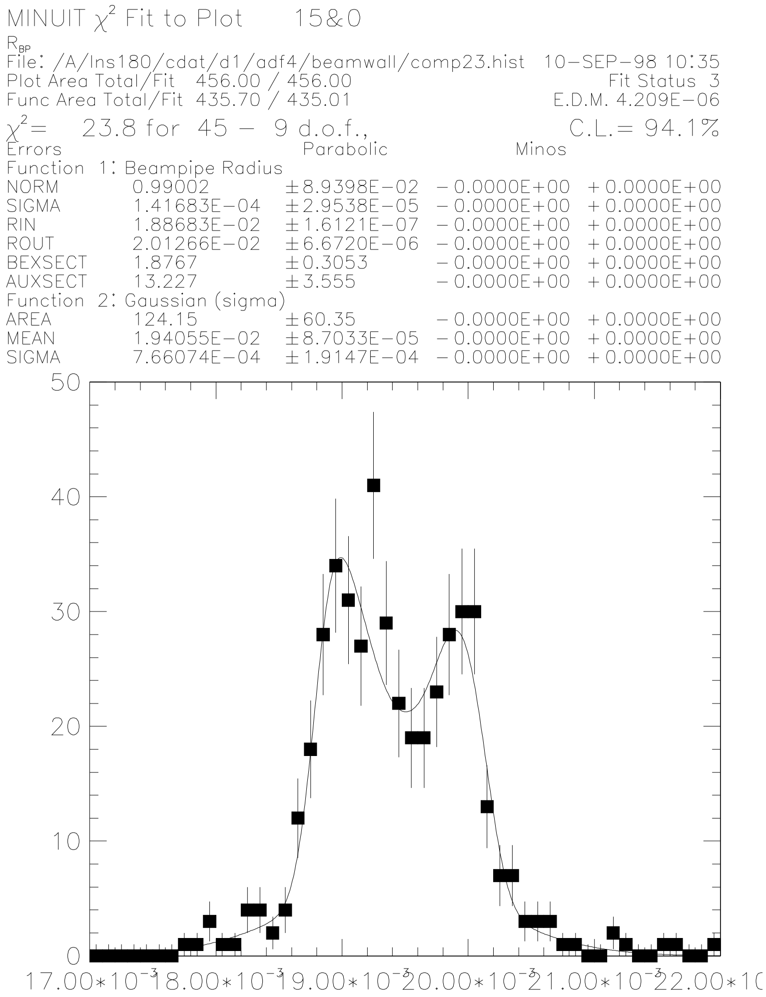


Figure 2.23: Fits to tapes 4sM1-12 in 2 dimensions.

As the figure shows, there is a shift inward of the inner radius of 1μm, and a shift outward of the outer radius of 19 μm. We assign an average shift of 10μm as an estimate of the systematic uncertainty associated with choosing to do the fit in 2 or 3 dimensions.

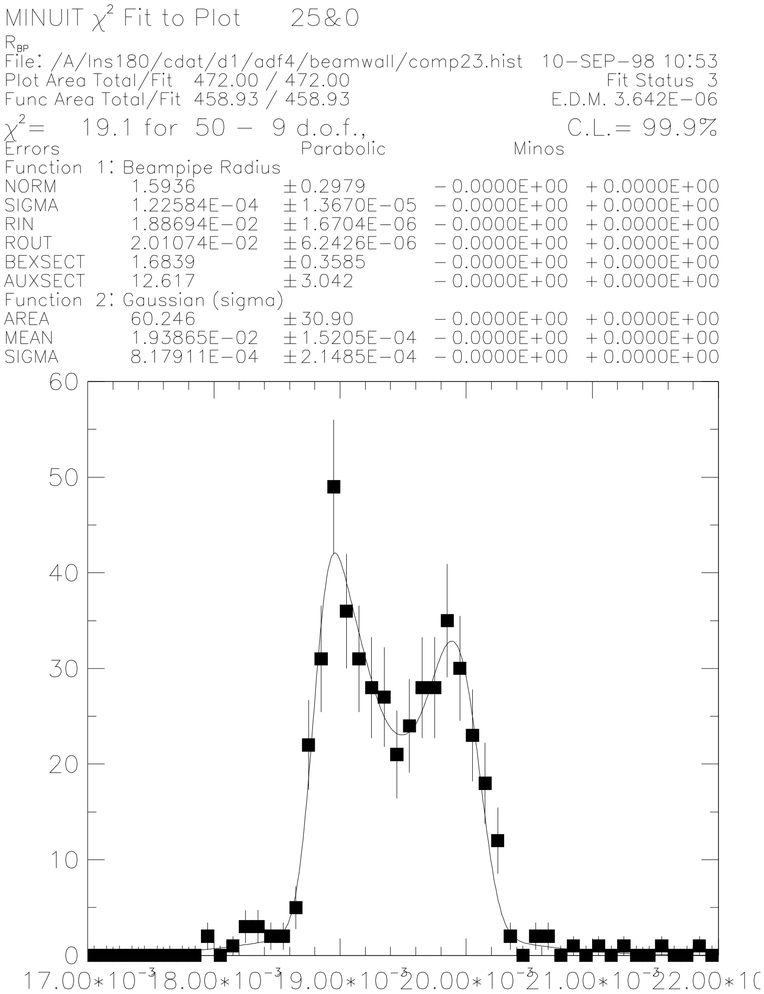


Figure 2.24: Fits to tapes 4sM1-12 in 3 dimensions.

2.6.2 Z-Weights Bug Fix

The data used in this analysis suffered from a reconstruction mistake known simply as “the z-weight bug” [36]. This might in principle induce systematic uncertainty in the $r - \phi$ results through tracking parameter correlations. It is nonetheless believed that the z-weights bug has no large effect upon r-ϕ only measurements. To test the effect, we perform the 2-D fit with and without the z-weights bug fixed.

We overlay the beampipe shape found with and without the z-weights bug fix in figure 2.25. As the figure shows, the z-weights bug makes very little difference to the r-ϕ measurements, as expected. The fit finds a shift in R_{in} of 1 μm and in R_{out} of 6 μm; we assign 5 μm as the systematic uncertainty due to the z-weights bug.

2.6.3 Images of the SVX Detector

For entertainment purposes only, we can redraw figure 2.13, but with the CLEO II.V dataset, from 4sH-4sR. This is shown in figure 2.26 and 2.27. Because tracks that originate from interactions in the third layer of silicon do not have silicon hit information, resolution for these vertices is much worse. In order to enhance the third layer, every point at radii larger than 4 cm is reflected eight times, corresponding to the eight-fold symmetry of the third layer.

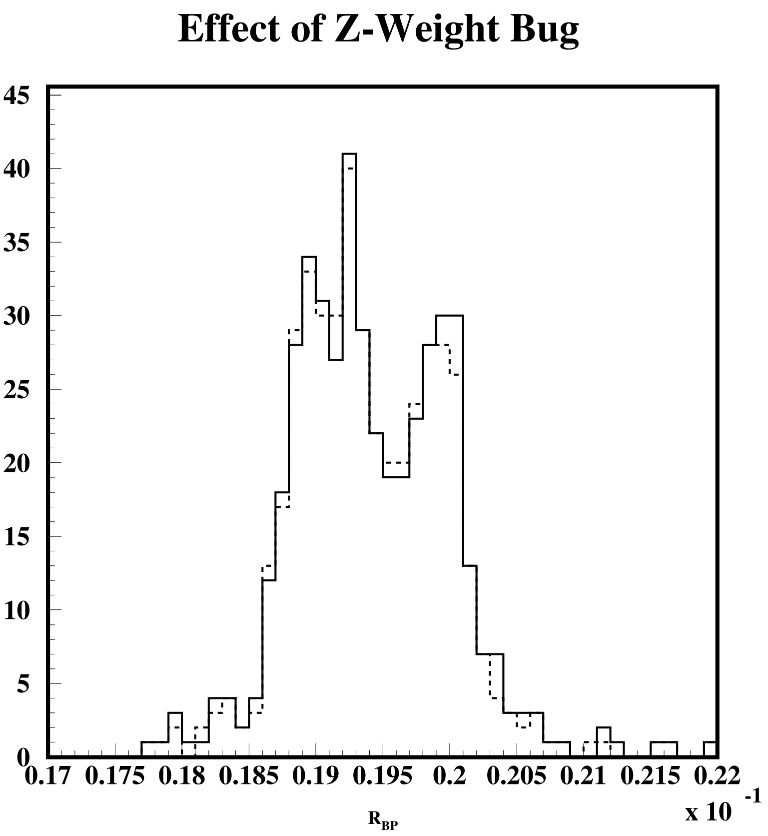


Figure 2.25: The beampipe shape found with and without the z-weights bug fixed. The solid line is a 2-D fit without the bug fixed, and the dashed line is with the fix.

Table 2.12: Summary table of systematic uncertainties

Source	Assigned Uncertainty	Assignment Method
Gold	10 μm (inner only)	Au thickness
Reported Vtx.	10 μm	Changed requirement
Uncertainty		from 200 μm to 1 mm
2-D vs. 3-D	10 μm	Average radial shift
Z-weights	5 μm	Average radial shift
Total	18/15 μm (in/out)	Add in quadrature

2.6.4 Summary of Systematic Uncertainties

In Table 2.12 we summarize the systematic uncertainties in the measurement of the beampipe radius, along with how we assigned the uncertainty. We do not pretend that this is a complete list of the uncertainties, but believe it is indicative of the most important questions that can come up in evaluation of the length scale using this method. The final comparisons of beampipe parameters are shown in Table 2.14.

2.7 Two-Photon Events: A Zero-Lifetime Sample

We know that $\gamma\gamma$ events originate at the origin, and therefore we can use them to nail down the "0" end of our scale. We have already described a selection of $\gamma\gamma \to 4\pi$ events in the previous chapter. The location of two-track vertices from this sample has been examined in the excellent note on the D lifetime [39].

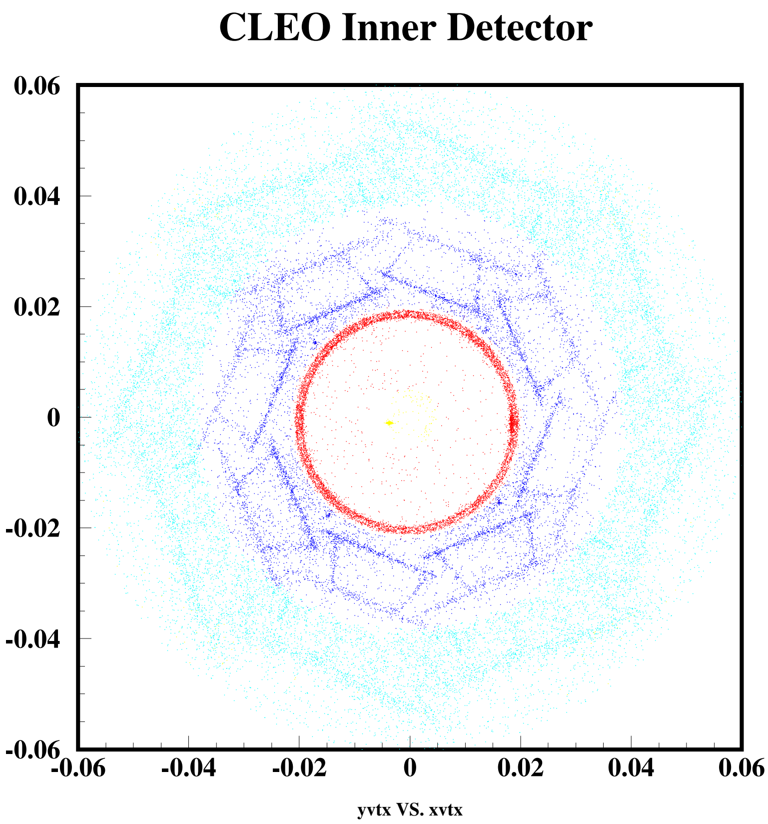


Figure 2.26: Location of vertices in the $x-y$ plane. In order to image the SVX detector, here we use the entire CLEO II.V data set, and loosen our vertex resolution requirement. Layer three has been folded 8 times in order to enhance its visibility. The support structures are clearly visible in this plot, as are the shield cables. The cables are visible as circular smudges nestled between the beampipe and layer 1 under the overlap regions.

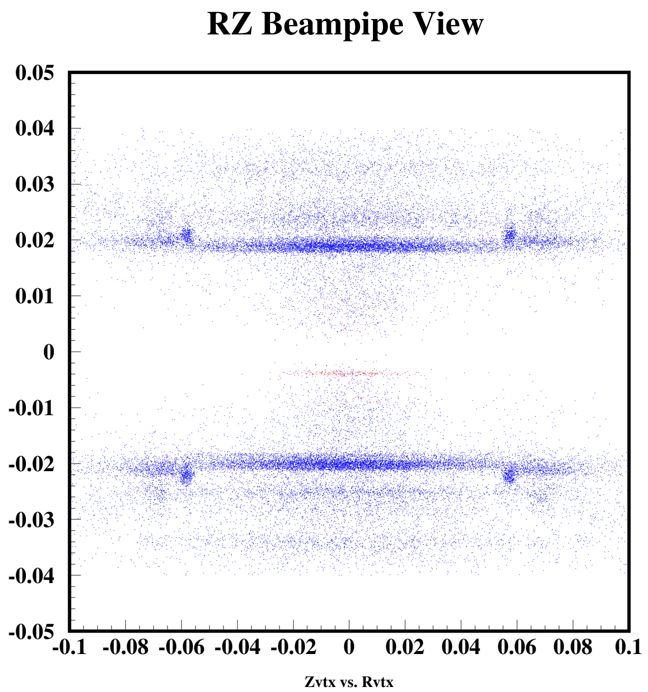


Figure 2.27: Location of vertices in the $r - z$ plane. In order to image the SVX detector, here we use the entire CLEO II.V data set, and loosen our vertex resolution requirement. The luminous region is shown in the center; the "knobs" located at ± 5.6 cm in z are the CAMEX shielding rings. The thick tungsten sheilds are just discernible behind the CAMEX shielding rings.

Note that though this measures “0” lifetime, it is measuring a distance 2 cm away from the nearest measurement point. In [39], the authors found a net lifetime bias of roughly 1 μm in flight distance.

2.8 Conclusion

Using the physics processes $\gamma\gamma \rightarrow e^+e^-$,$\gamma\gamma \rightarrow 4\pi$, $D^{*+} \rightarrow D^0\pi^+$, and nuclear interactions with the detector material, we have studied the hardware efficiency, tracking efficiency, resolution, hit purity, and distance scale accuracy of the CLEO II.V SVX detector. The results for the tracking efficiency $\varepsilon_{2/2}$ are summarized in table 2.13. Figures 2.9 and 2.10 show the dependence of efficiency on momentum, which is very flat down to 150 MeV. Hit purities seem to be high, less than 1.4% and 0.9% of tracks having noise hits in $r-\phi$ and $r-z$, respectively. The resolutions in $r-\phi$ and $r-z$ have been characterized as

$$19 \oplus \frac{39}{p\beta \sin^{\frac{3}{2}}\theta}(r-\phi) \tag{2.5}$$

$$42 \oplus \frac{45}{p\beta}(r-z) \tag{2.6}$$

The results shown in table 2.14 using hadronic interactions show that we can measure very long distances from the beam ($<$ 1 cm from SV) without bias. The $\gamma\gamma \rightarrow 4\pi$ results from [39] show that we can also measure 0 lifetime from the beam (2 cm from the SV) with little systematic bias. The overall distance scale is correct within our resolution, that is, to roughly 1 part in 500.

Table 2.13: A summary of SV efficiencies for the different types of events. The second column is a measure of the reconstruction software efficiency, the third is the total SV efficiency for all tracks.

Track Type	Momentum Range (GeV)	$\frac{\varepsilon_{2/2}(fiducial)}{\varepsilon_{2/2}(hardware)}$	$\varepsilon_{2/2}$
μ Pair	5.2	99.3±0.1%	93.8±0.1%
$\gamma\gamma \rightarrow e^+e^-$	0.5-3	97.1±0.2%	91.7±0.1%
D^0 Daughter	0.5-3	96.9±0.5%	91.5±0.5%
$\gamma\gamma \rightarrow 4\pi$	0.1-1.0	96.0±0.4%	89.5±0.4%
π_s	0.1-0.4	94.3±0.7%	89.1±0.7%

Table 2.14: Summary table of results for length scale.

Radius	Fitted Value (μm)	Survey (μm)	Fractional Difference
0 (beamspot)	1.0 ± 1.8	N/A	N/A
Inner	18,865 ± 3 ±18	18,860±30	0.03±0.18%
Outer	20,106 ± 5 ±15	20,156±15	0.23±0.11%

CHAPTER 3

TAGGING THE B FLAVOR

3.1 Introduction

As noted in the first chapter, a CP-violation experiment must record not only temporal information concerning one B decay time, but also be able to determine whether the other B in the event decayed as a B or $\overline{B}$. The experiment must also manage to do this with high efficiency. We would like to find the most accurate and efficient way to separate B from $\overline{B}$. Obviously a high-momentum lepton, signaling a primary decay of the type $b \to c\ell\nu$, is one accurate but inefficient way to tag the flavor. On the other hand, kaons from the decay chain $b \to cX \to sX$ compete with kaons from decays chains such as $b \to c\bar{c}s$ and $b \to cX \to s\bar{s}sX$ or $b \to cX \to s\bar{s}X$, to provide an efficient but less accurate way to tag the flavor.

The idea of flavor tagging that takes advantage of more than just leptons has been floating around for some time; both Babar and Belle plan to use such taggers for their studies of CP violation. ARGUS has published $K^{\pm}$ flavor correlations [40], CLEO has done some studies [12], and also Babar has done a great deal of study of inclusive flavor tagging [11]. CDF and the LEP experiments have both used a number of techniques to tag B flavors, however, no proven flavor tagger exists for events at the $\Upsilon(4S)$. In this note, we describe a B flavor tagger we have written, and its evaluation with the CLEO data.

In essence, there are three steps to writing a flavor tagger:

- Identifying a set of quantities that discriminate between B and $\overline{B}$ decays
- Optimally combining the information from this (possibly large) set of quantities
- Measuring the performance of the tagger with data

Note that because of step 3, we are more or less free to combine any set of variables in any way we like, without introducing dependence on the Monte Carlo or assumptions about the data. Because of this, we utilize a neural net to squeeze out the last few percent of tagging ability.

This chapter proceeds as described below:

- Framework for flavor estimation: dilution
- Description of useful discriminators
- Parameterization of dilutions with Monte Carlo
- Optimization of a neural net with Monte Carlo
- Evaluation of tagger performance with $D^{*\pm}l^{\mp}\nu$ data

In an appendix we describe our algorithm for identifying $K^{\pm}$.

As the simplest (and most important case) we consider here the case of tagging the flavor of a single remaining neutral B in an event when a signal B has been fully reconstructed, and its tracks and showers removed from consideration in the flavor tag. A fully reconstructed signal B is the circumstance when using flavor-tagging for CP-violation measurements, such as with ψK_S, $\pi\pi$, or $\rho\pi$.

We concentrate on the case of tagging neutral B's. We remind the reader that a charged B's flavor may be reconstructed by simply summing the charge. We find that there are in fact substantial differences between flavor tagging charged and neutral B's. This is for three easily understood reasons:

- In B^+ decays in a perfect detector, the charge sum tells you the flavor. That is, every single track carries some flavor information. In B^0 decays, most tracks do not. This is the most noticeable effect.

- Because the $D^\pm$ and D^0 semileptonic BR are different by a factor of 2.5, and because B^0 and $B^\pm$ contain differing amounts of $D^\pm$ and D^0, the cascade spectrum and normalizations are slightly different.

- $B^\pm$ and B^0 have different $D^\pm$ and D^0 fractions, which in turn have different K_S and $K^\pm$ fractions. As K_S are useless for tagging, this causes differences in the ability to tag B^0 and $B^\pm$ decays.

3.1.1 Dilution and Flavor Estimation

Any tagging procedure considers the electric charge of various particles within the event, and attempts to correlate them to the charge of the b or $\overline{b}$ quark. For instance, one might select leptons with momentum above 1.5 GeV, and assign the b quark charge sign to be equal to the lepton charge sign. Or one might consider $K^{\pm}$ in the event, and assign the b quark sign to equal the K charge sign. Some tagging particles have sign opposite to the b quark. For instance, one could select $D^{*\pm}$ and assign the b sign to be opposite the D^* sign. Very low momentum leptons (below 900 MeV) are more likely to be from cascade c decays than primary b decays, and therefore have sign opposite to the b quark.

We utilize a generic approach for keeping track of all the possible correlations or anti-correlations and combining all the information available in the event. We require to know, for each event, whether the flavor prediction is very reliable (high-momentum lepton), somewhat reliable (K) or pretty shaky (when neither is found in the event.) We describe the expected dilution $\mathcal{D}$ of a tag, and combine dilution information from all tags into an event-by-event flavor estimator $\hat{\mathcal{B}}$.

The flavor tagger we describe takes advantages of two methods to combine the information from several tags: a dilution-based method and a neural net method. The dilution-based method, is more transparent, but cannot account for correlations between observables, and therefore loses information. The result from this method is then used as a primary input to a neural net, which can account for the correlations.

Dilution

A most useful way to think about flavor tagging is in terms of the dilution, $\mathcal{D}$, of the tagging procedure. $\mathcal{D}$ is defined as

$$\begin{aligned}\mathcal{D} &= \tfrac{\text{Right sign} - \text{Wrong sign}}{\text{Right sign} + \text{Wrong sign}} \\ |\mathcal{D}| &= 2P - 1\end{aligned} \tag{3.1}$$

where P is the purity, and the sign of $\mathcal{D}$ is meaningful.

One can define $\mathcal{D}_\ell$, $\mathcal{D}_K$, $\mathcal{D}_{D^*}$, etc. For instance, leptons above 1.4 GeV are 96.5% from primary decays, and 3.5% from cascade decays (wrong-sign). From equation 3.2, we see that this corresponds to $\mathcal{D}$ of 0.93. Leptons from 600-800 MeV are 60% from secondary (wrong-sign) decays and 40% from primary (right-sign) decays; this corresponds to $\mathcal{D}$ of -0.2. The negative sign is important–it tells us to flip the sign of the electric charge.

For leptons, $\mathcal{D}$ is a function of momentum. In general, for any tagging category, we can make $\mathcal{D}$ a function of any number of variables, though usually momentum is the most useful. The sign of $\mathcal{D}$ can even change as a function of these variables, as demonstrated by the lepton case.

Flavor Estimation

$\mathcal{D}$ is also useful because it allows us to combine information from several different tags. For instance, if an event contains both a K and a lepton, we would like to use them both in order to improve our confidence over using only the lepton. We can build a flavor estimator for each tag category,

$$\hat{\mathcal{B}}_i = \mathcal{D}_i Q_i \tag{3.2}$$

where i is, for instance, ℓ, K, or D^*; and Q_i is the electric charge of the tagging track. If $\hat{\mathcal{B}}$ is greater than 0, we call it a B meson, if less, a $\overline{B}$ meson. For a single tag category i, the purity of the estimator $\hat{\mathcal{B}}_i$ is clearly the purity of $\mathcal{D}_i$, that is,

$$P(\text{Correct Tag}) = 0.5 + \tfrac{\hat{\mathcal{B}}}{2} \tag{3.3}$$

However, we can combine estimators from all tag categories in an event to find a more powerful estimator. It is a result [44] that if the estimators are uncorrelated, the relation 3.3 is preserved if the individual estimators (dilutions) are summed like relativistic velocities,

$$\hat{\mathcal{B}}' = \tfrac{\hat{\mathcal{B}}_1 + \hat{\mathcal{B}}_2}{1 + \hat{\mathcal{B}}_1 \hat{\mathcal{B}}_2} \tag{3.4}$$

For instance, for an event with both a high-momentum lepton and a $K^{\pm}$ in the event,

$$\mathcal{B} = \tfrac{\mathcal{D}_\ell(p_\ell)Q_\ell + \mathcal{D}_K(p_K)Q_K}{1 + \mathcal{D}_\ell(p_\ell)Q_\ell \mathcal{D}_K(p_K)Q_K} \tag{3.5}$$

In the (actual) case that the individual estimators are correlated, the relativistic-velocity combination is not optimal, and does not preserve the relation 3.3. To account for these correlations motivates us to use a neural net.

We mentioned earlier that for a given tag species there may be several variables of interest x_j on which $\mathcal{D}$ depends. For instance, $\mathcal{D}$ for leptons is found to depend usefully not only on p, but also m_{vis}, the invariant mass of the remaining tracks and showers in the event, and cosθ_{near}, the angle to the nearest track. We assume that the function $\mathcal{D}(p, m_{vis}, \cos\theta_{near})$ factorizes into three independent dilutions, $\mathcal{D}(p)$, $\mathcal{D}(m_{vis})$, and $\mathcal{D}(\cos\theta_{near})$. In this case, for a given lepton track, we combine the three flavor estimators as described in equation 3.4.

3.1.2 Road Map

To make a dilution-based tagging package, we first choose a set of tag categories i (leptons, K's, etc.), then map out $\mathcal{D}_i(x_j)$ in each category using Monte Carlo. Once we have this set of $\mathcal{D}$ functions, we run the tagger over each event. In each event, we sum all $D_i(x_j)Q_i$ according to equation 3.4 to form $\hat{\mathcal{B}}$ for the event. To check that the MC is consistent and the variables reasonably uncorrelated, we plot the resulting true (from Monte Carlo) dilution $\mathcal{D}$ of the package as a function of the reported $\hat{\mathcal{B}}$ of the event; that is, in each bin of $\hat{\mathcal{B}}$, subtract correct tags from incorrect tags. If all variables are truly uncorrelated, $\hat{\mathcal{B}}$=$\mathcal{D}$; deviations from a straight line indicate correlations. Finally, once that is done, we test it on data, for instance, a fully-reconstructed B sample, where the flavor is known.

3.2 Choosing Tag Categories

Which categories are actually useful? Obvious categories of tracks are e, μ, and $K^{\pm}$. Less obvious categories are the highest-momentum track in the event (which, if the momentum is high, can indicate a two-body or psuedo-two body decay such as $D^*\pi$ or Da_1) or the lowest-momentum track in the event (which often is the soft pion from a D^*).

We also want to know, for each category, which variables most usefully parameterize the variation in the tagging power of tracks in the category. For instance, the momentum of a lepton indicates whether the tag is reliable (high momentum) or suspect (low momentum, which may be from a secondary charm decay).

We use the following categories and variables:

- Track categories
  - `R2ELEC` > 3
  - `DPTMHU` > 5 .OR. ($3<$ `DPTMHU` < 5 .AND $p < 1.4$)
  - $L_K > 0.9$ .AND. `R2ELEC`< 0 (L_K is described in the appendix on $K^{\pm}$ identification.)
  - The lowest momentum track
  - The highest momentum track
  - All other tracks
- Quantities which parameterize above cases
  - p (all categories)
  - Visible m_{recoil} (soft ℓ only)
  - Angle to most nearby track (soft ℓ only)
  - Number of K_S in the event ($K^{\pm}$ only)

To generate the distributions of $\mathcal{D}_i(x_j)$, we plot the dilution (asymmetry) in the charge Q of each track, as a function of momentum p, for each of the following track categories. We subdivide the above list into the following types of tags:

- High-momentum (above 1.4 GeV) leptons (p)
- Soft Leptons (p, m_{vis}, $\cos\theta_{near}$)
- $K^{\pm}(p)$ in events with no K_s
- $K^{\pm}(p)$ in events with some K_s
- If applicable, the only $\pi(p)$ in the event
- The lowest-momentum $\pi(p)$ in the event, if different from above.
- All other tracks(p)

For leptons below 1.4 GeV, we consider two additional quantities in the event. The first, $\cos\theta_{near}$, is the cosine of the angle to the nearest track in the event. The second, m_{vis}, is the invariant mass of all additional tracks and showers in the event. The expected dilution as a function of these variables are added to the expected dilution as a function of the momentum as in equation 3.4, according to the principle of factorization. In reality, $\cos\theta_{near}$ is probably correlated to m_{vis}, so this is not the optimal procedure.

For $K^{\pm}$, the dilutions are very different depending on whether there is a K_S in the event. When there is an additional K_S in the event, this is usually a sign either that (a) a particle ID error occurred (it isn't really a $K^{\pm}$), or (b) there was upper vertex production in the event. In either case the K charge is a less accurate estimator and we must reduce its assigned weight.

3.2.1 Flavor Estimation

The actual estimation of the flavor in an event proceeds as follows. We loop through all trackman-approved tracks, and for each track, we determine whether it falls into any of the categories listed in the previous section. We then look up the expected dilution $\mathcal{D}$ (from the following section) for a track in this category with the measured momentum (and, for some tag categories, additional event quantities.) We multiply $\mathcal{D}$ by the charge of the track as in equation 3.2, then sum the expected dilutions from each track as described by equation 3.4.

3.3 Construction of the Dilution

We have made plots, from the Monte Carlo, of the expected dilution for each of the tag categories listed above for B^0 decays.

Note that these distributions are generated from Monte Carlo, and therefore directly affect the reported value of $\mathcal{D}$ for each event. However, because we will use a data sample to map from reported dilution $\mathcal{D}$ to true dilution, there will be no direct Monte Carlo dependence in the final result. However, the closer Monte Carlo matches data, the more *optimal* the tagger.

3.3.1 On QQ

For the plots in this section, we generated single B Monte Carlo. That is, one B decays to $\nu\overline{\nu}$ to simulate the removal of a fully reconstructed signal B. The single `B0B` or `B`$^-$ was allowed to decay generically. We verified that the version of QQ used had upper-vertex decay consistent with the CLEO measurement of CLNS 97/1516.

3.4 Results for the Dilution-Based Tagger

3.4.1 Distributions in MC

Having parameterized the dilutions of the various tags, we now run over the generic MC to see how it did. We will be interested in three quantities: the actual distribution of $\hat{\mathcal{B}}$, the relation between the reported $\hat{\mathcal{B}}$ and the true dilution $\mathcal{D}$ of the events, and the effective efficiency, εD^2, of the tagging package.

$\hat{\mathcal{B}}$ Distribution

These distributions, shown in figure 3.1 look as expected. We see a sharp peak at 1, corresponding to events with a found lepton, a broad lump around 0.7 corresponding to a kaon, and a dribble to lower values indicating less reliable information when these two are not available. The wrong-sign distribution peaks near 0 and tails off very quickly.

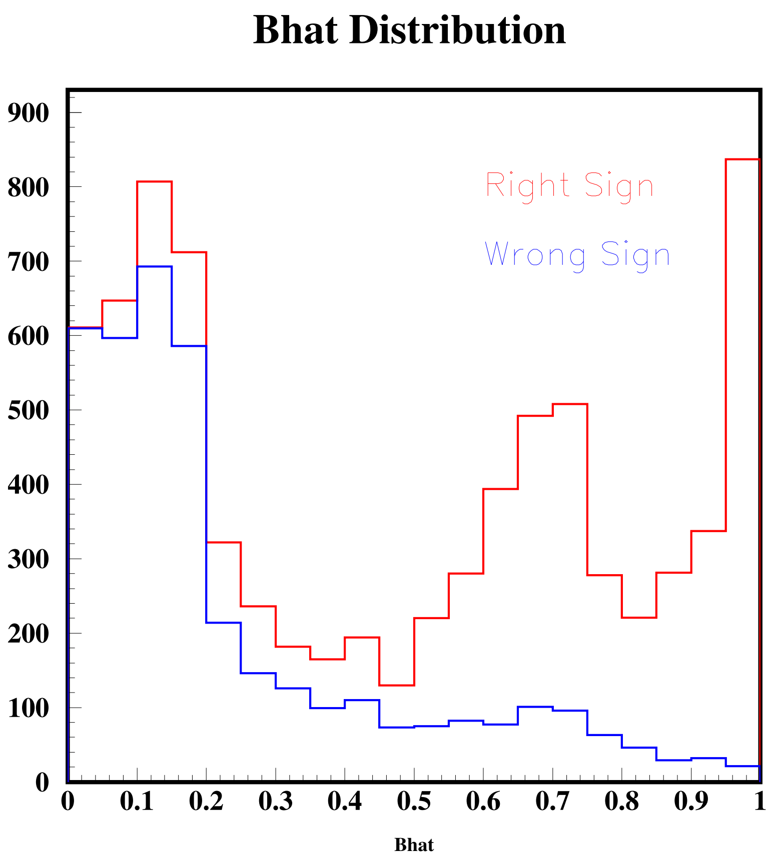


Figure 3.1: Monte Carlo distribution of the flavor estimator B for generic single B^0 events.

Does $\hat{\mathcal{B}} = \mathcal{D}$?

The reason for constructing the dilutions and their combinations as we have is to preserve $\hat{\mathcal{B}}$=$\mathcal{D}$, that is, so that the reported $\hat{\mathcal{B}}$ estimator is also a measure of the expected dilution. This is true in the absence of correlations. We therefore expect that in any bin of $\hat{\mathcal{B}}$, the true dilution $\mathcal{D}$ of the events is $\hat{\mathcal{B}}$. As is shown in figure 3.2, this expectation is rather well-met in B^0 events.

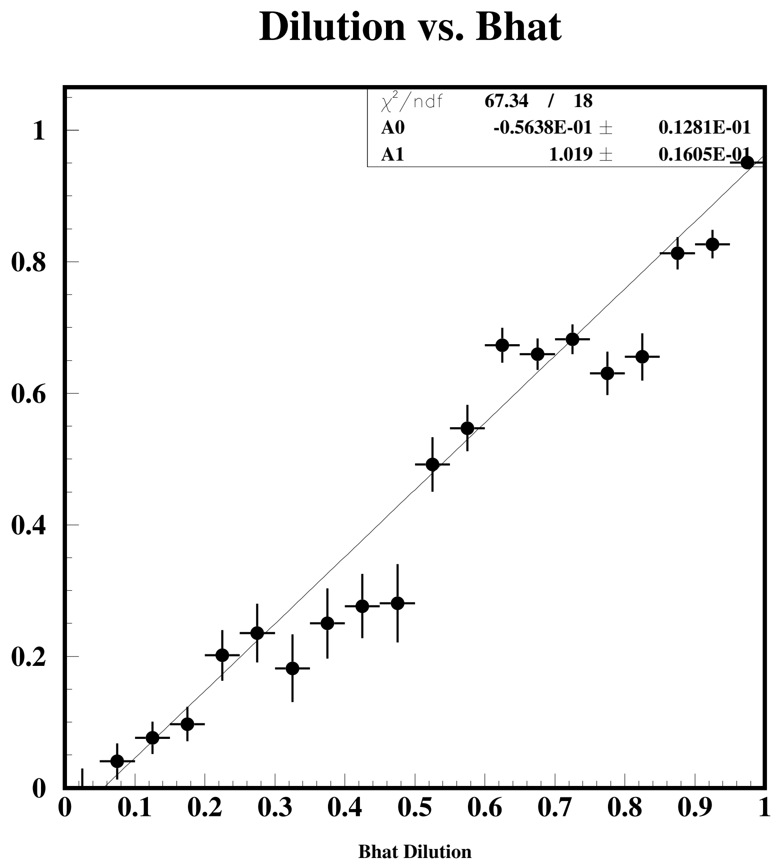


Figure 3.2: Monte Carlo comparison of B estimator (x axis) to true dilution in B bin (y-axis), for generic single B^0 events (left).

Calculating $\varepsilon\mathcal{D}^2$

We would like to know the separation between B and $\overline{B}$ events we can acheive using this tagger. In trying to measure a property of B mesons as opposed to $\overline{B}$ mesons, in a given bin of dilution, the statistical uncertainty of the measurement is proportional not to $\sqrt{N}$, but rather $\frac{\sqrt{N}}{\mathcal{D}}$. This is often accounted for by describing an effective efficiency that is simply decreased by $\mathcal{D}^2$. Therefore, the overall effective efficiency for measuring properties of B mesons of a specific flavor, also known as the separation, is given by integrating over all bins, and is often notated by s or εD^2. Sometimes the experimenter may wish only to consider events with a minimum value of $|\hat{\mathcal{B}}|$. If imposing a cut of $\hat{B}$, the separation is given by

$$\begin{aligned} s(\hat{B}) &= \int_{\hat{B}}^{1} dB' \frac{d\varepsilon}{dB'} \hat{\mathcal{D}}^2(B') \\ &\simeq \int_{\hat{B}}^{1} dB' \frac{d\varepsilon}{dB'} \hat{\mathcal{B}}^2(B') \end{aligned} \tag{3.6}$$

The separation when the cut is set to 0 is the maximal possible effective efficiency. The distribution of $s(\hat{B})$ for Monte Carlo is shown in 3.3.

As figure 3.3 shows, the events with B less than about 0.4 contribute quite little to εD^2. Most of the power comes from events which are well-tagged.

We note that Babar[43] and Belle[48] expect to find a value of 0.30 for their detectors without RICH or DIRC, but with $\frac{dE}{dx}$. Babar expects to find 86% efficiency for a $|B|$ cut of 0.1; we find 80%. This gives us some confidence we are not making obviously stupid mistakes, and also some warning that there is not necessarily much more to be squeezed out.

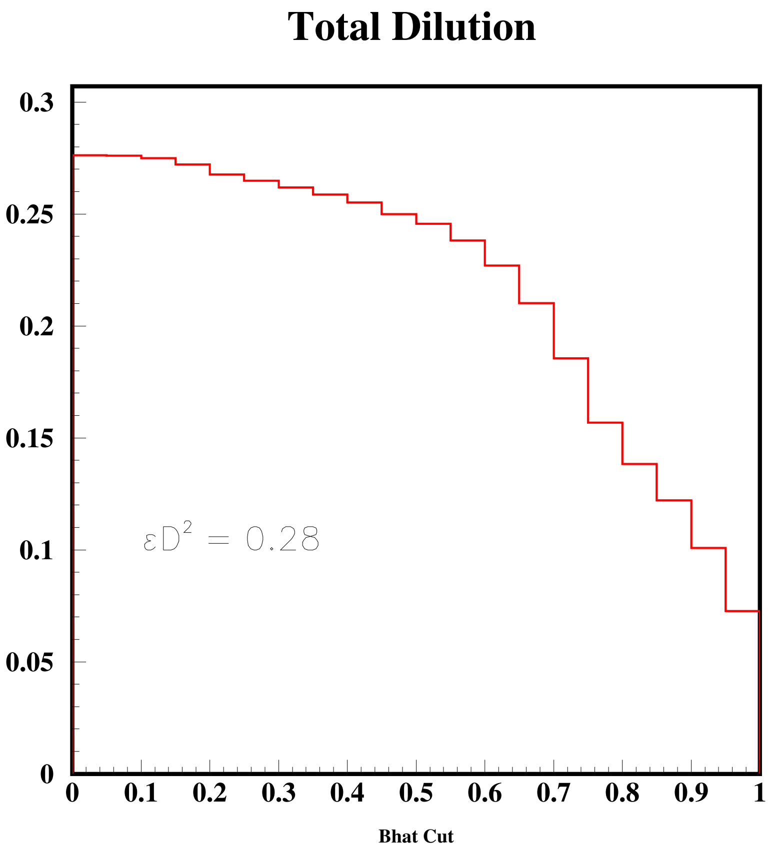


Figure 3.3: “Effective efficiency”, εD^2, as a function of cut on the flavor estimator B, in Monte Carlo. Generic single B^0 events (left) show an asymptotic value of 0.28.

3.5 Getting Fancy: the Neural Net

As mentioned earlier, the dilution-based approach loses ground when input variables are correlated. This limitation also prevents us from adding certain kinds of variables to the dilution, because they are strongly correlated with others. For example, the sum over all tracks $\sum_{tk} p_{tk} q_{tk}$ for an event shows some useful discriminating power, as shown in figure 3.4. However, it is strongly correlated with all of the other variables, such as momentum of the highest-momentum track, or lepton momentum. Hence we do not use it in the dilution-based approach.

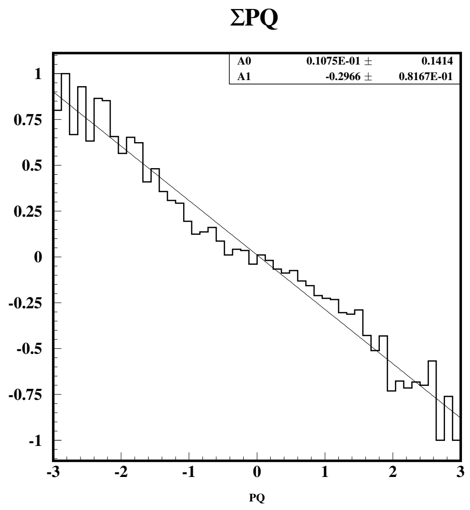


Figure 3.4: The $B - \overline{B}$ asymmetry as a function of the sum over all tracks $\sum p_{tk} q_{tk}$, in generic Monte Carlo.

With a neural net, we can overcome these disadvantages. There are reports in the literature of substantial gains with the use of neural nets [48] for flavor tagging. We use the DRAGNET package, which is described in [52, 53]. The package also usefully identifies useful variables, by examining $\frac{\partial \text{Net Output}}{\partial x_i}$ for many input variables x_i.

A neural net is a series of "nodes", each of which takes a set of inputs. Each node's output is determined from it inputs; the output of node j for a set of inputs x_i is

$$r_j = \tanh(\sum_i W_{ji} x_i - T_j) \tag{3.7}$$

"Training" a neural net consists of finding the proper weights W_{ji} and thresh-holds T_j for the nodes. Once these have been set, the output of a neural net is *completely deterministic*.

Nets are generally more powerful when outputs of some nodes are used as inputs to others. How the nodes are hooked up to each other is called the "architecture" of the net. We use the 7×5×1 architecture shown in figure 3.5.

The words "neural net" rightly cause fears of overtuning, non-understood dependences, and possibly pathological behavior. We share in these fears of neural nets and have gone to lengths to avoid them. We avoid them in five ways:

- We use the transparent dilution-based estimator as an input
- We minimize the size of the net
- We minimize the number of inputs, using only the most powerful and sensible inputs
- We vary the weights around their trained values

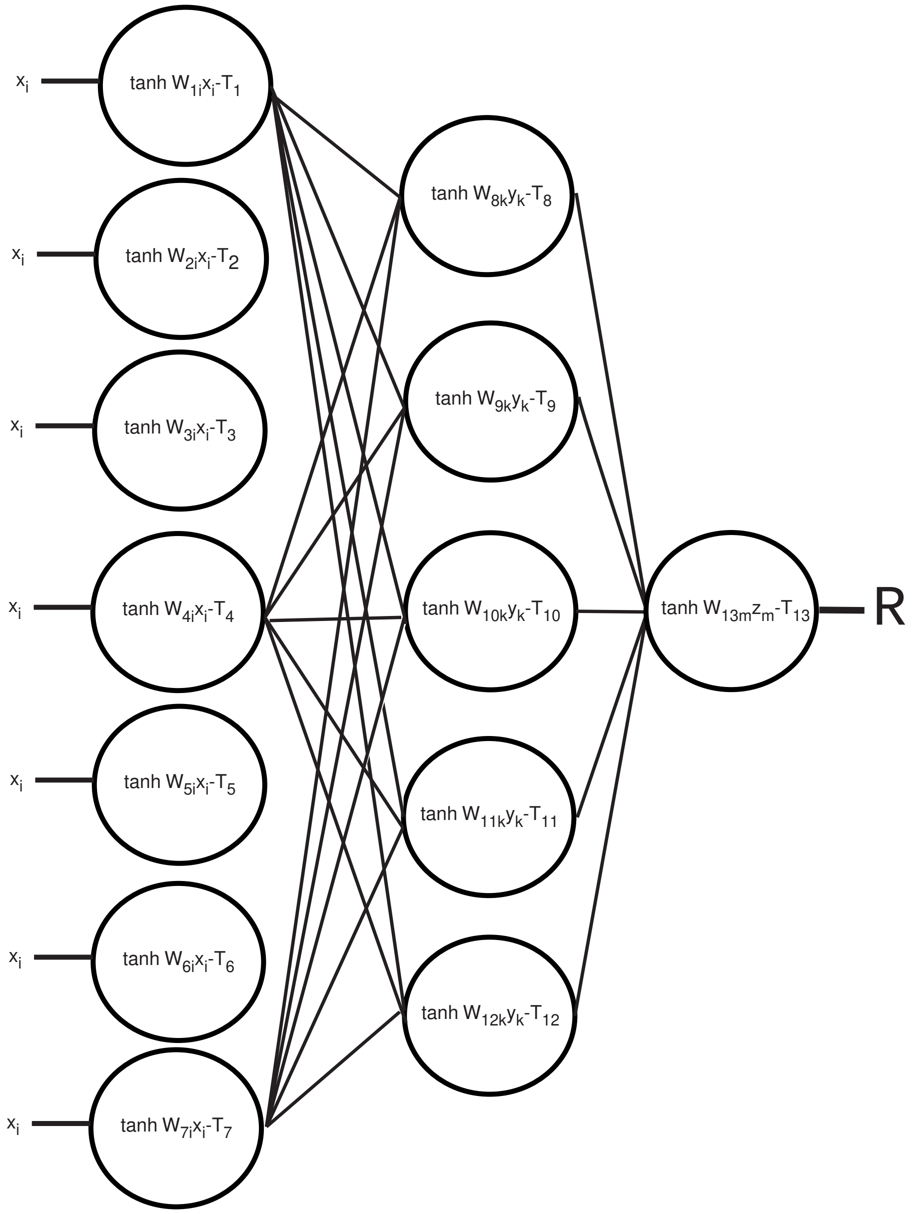


Figure 3.5: The neural net architecture used for the flavor tagger. The outputs from nodes 2,3,5, and 6, each of which are used as inputs to the second-layer nodes, are suppressed for clarity. x_i are the event inputs; y_k are the first-layer outputs, and z_m are the second-layer outputs.

- We test it all on data (see below)

The dilution-based estimator $\hat{\mathcal{B}}$ is indeed one of the inputs the nets finds to be the most powerful, and therefore which the net weights most heavily. This immediately limits the amount of pathological behavior possible, since we effectively tell the net the right answer to start with. By minimizing the size of the net and the number of inputs, we limit the number of free parameters in the training. By varying the weights from their tuned values, we can insure that the weights have not been fine-tuned to particular features in the events. Finally, we test it all on data, and measure all properties of the tagger in data, rather than relying on the Monte Carlo.

In this section, we describe the selection of input parameters and of a net architecture. In the next section we describe the final performance of the neural net tagger in Monte Carlo. In the subsequent section we discuss performance in the data. We refer henceforth to the neural net output as F_{NN}.

Selection of input parameters and of the architecture is executed upon 400k `trksim` events. Once these were selected, the nets were trained (i.e., their weights optimized) upon these samples. After that, the nets were trained for one more cycle on 20,000 CLEOG events.

3.5.1 Selection of Input Parameters

We end up with the following set of variables for the event; in parentheses is a measure of relative importance.

- $\sum Q_K$ (0.37)
- $\sum Q_\pi P_\pi$ (0.30)
- $\sum Q_\ell P_\ell$ (0.23)
- $\hat{\mathcal{B}}$ (dilution estimator)(0.15)
- $Q_{hardest}$ (0.13)
- $\sum P_{tk} Q_{tk}$ (0.13)
- $Q_{hard} m_{recoil}$ against hardest track (0.12)
- $Q_\ell(E_\ell + |p_{miss}|)$ (0.11)
- Q_2 (0.08)
- $E_{K_S} Q_K$ (0.08)
- Q_{soft} (0.07)
- $\sum Q_\ell$ (0.06)
- $Q_\ell m_{recoil}$ against lepton (0.06)
- $Q_{soft} m_{recoil}$ against softest track (0.05)
- $Q_\ell m \ell K ``\nu''$ (0.05)
- $Q_1 m_{12}$ (0.04)
- $Q_{hardest} p_{hardest}$ (0.03)

- $\sum Q_\Lambda P_\Lambda$ (0.001)

All variables are signed by an appropriate charge. This ensures that all variables are mirror reflections for B and $\overline{B}$.

One should not be alarmed that $\hat{\mathcal{B}}$ is not the apparently most important variable. It is strongly correlated with the other powerful variables, and if they were removed, would itself move up in importance to replace them.

3.5.2 CP Properties of the Tagger

Because we want to measure CP violation, we must ensure that the tagging procedure is not itself CP asymmetric. Such an asymmetry could arise either from an asymmetry in the neural net, or an asymmetry arising from the detector. An example of the latter could be charge-dependent differential lepton detection asymmetries or particle identification asymmetries.

CP Asymmetry of the Neural Net

When we came to the final tuning, we encountered a "feature" of neural nets, which is the following. A neural net is designed to maximize the separation between two classes (i.e. minimize the overlap). Suppose the net output has separation s^2 with a particular output distribution. One can easily convince oneself that any monotonic transformation of the output distribution leaves the overlap integral, and therefore the separation, unchanged.

All monotonic transformations are not symmetric around 0, and because they are degenerate in separation, the neural net distributions will not, in general, be symmetric about 0. However, precisely because of this degeneracy, we are free to enforce some conditions on the threshholds T_i of the nodes. We enforce T_i=0. This means that because all inputs are C-odd, and because the tanh functions are C-odd, all outputs are odd under simultaneous change of all signs in the event (i.e. $B \to \overline{B}$.) Hence the final output of the tagger, all inputs being signed by the charge, and all tagger functions being odd in the charge, has definite "parity" with respect to C. If all charges are simultaneously flipped in sign, and all other quantities (momenta, etc) remain unchanged the tagger will necessarily produce a result exactly opposite in sign.

The above discussion is true in the case of perfect execution of enforcement of $T_i = 0$. In the presence of typos or bugs it need not be the case. We explicitly test this CP invariance as follows. We run the tagger over 10k B events. We then flip all charges in the event, changing nothing else. We rerun the tagger on this modified event. If our tagger has been correctly written, these distributions must be exactly (not approximately or statistically) the opposite. We find that they are, event by event, for all 10k events, and therefore that the tagger is indeed C-symmetric.

In some sense, this is an aesthetic enforcement, as C-asymmetric distributions are not necessarily CP violating (as long as the "$\overline{B}$-like" component of the B sample looks like true "$\overline{B}$", and vice-versa), nor need they have have different separation. It is also comforting that without enforcing the threshhold condition, 11 of the 13 threshholds were less than 0.02, the remaining 2 being 0.15 and 0.11. That is, this is the sort of solution that the net finds anyway.

After enforcing this condition, we train on the final sample of `cleog` events, beginning with the weights we found in the `TRKSIM` training, with the T_i adjusted (and subsequently pinned) to 0. This final cycle in fact causes very little change in εD^2–the addition of new K_S and Λ information, which were not in the `TRKSIM` simulation, are cancelled by resolution effects, which are more accurately accounted for in `cleog`,

The final results on an independent sample of 4000 `cleog`'ed events are shown in figure 3.6. The neural net output very accurately reflects the actual dilution, as shown in the upper right plot. We find εD^2=0.28 in this sample. Figure 3.7 also shows the correlation between the dilution-based tagger and the neural net output. Figure 3.8 shows the difference. As expected, this plot shows a net shift to the right, indicating that the neural net tagger generally reports a higher expected dilution.

CP-violating Tagging Effects in the Detector

We use fully `cleog`'ed generic single B events where, as noted earlier, we have taken pains to verify that the $c\overline{s}$ production is as correct as possible. We use a sample of 20k events *independent* of all training and evaluation cycles above.

We have generated distributions for all of the input quantities of the neural net. We show these distributions in an appendix. In each plot, we show distributions of the input quantity for B events, on which we superimpose the $-\overline{B}$ distributions. If QQ, and Monte Carlo are CP-symmetric, these distributions should be degenerate. QQ is known to be CP symmetric. Hence deviations from degeneracy arise from

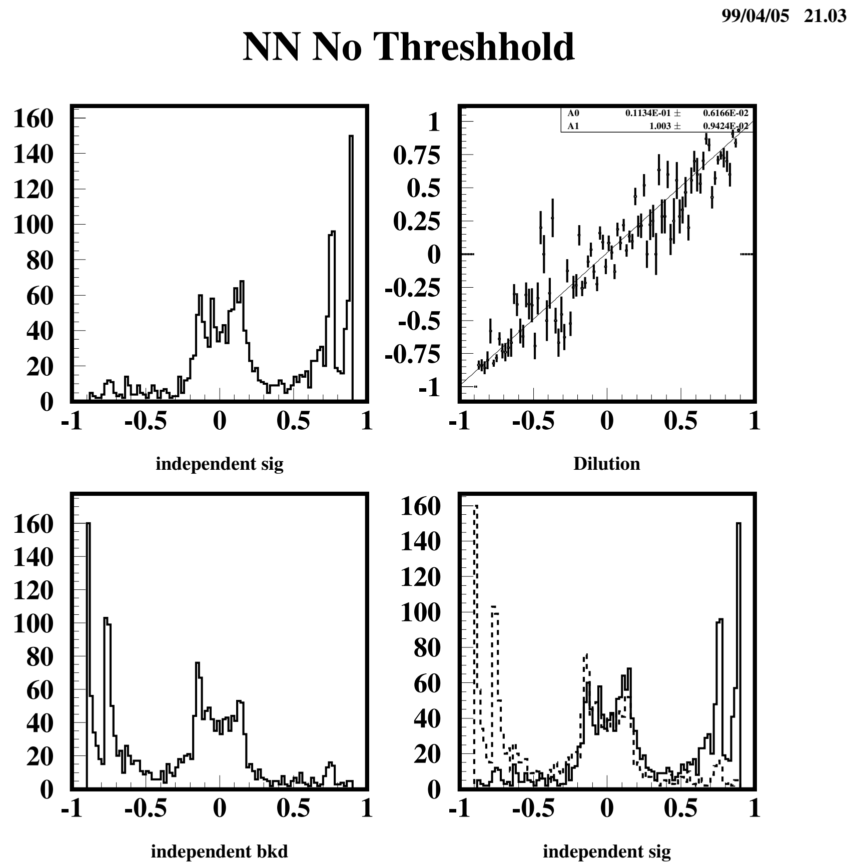


Figure 3.6: Output of the neural net tagger on 4000 `cleog` events not used in its training. On the left are the outputs for B (top) and $\overline{B}$ (bottom). The upper right shows the relation between true dilution in the bin and neural net output. F_{NN}. The lower right simply superimposes the B and $\overline{B}$ distributions for ease of comparison.

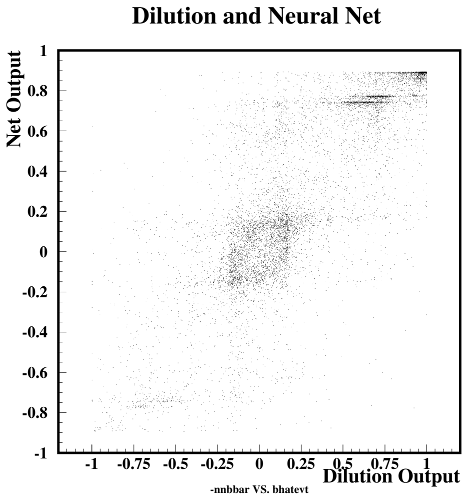


Figure 3.7: Comparison of the dilution based output $\hat{\mathcal{B}}$ to the neural net output F_{NN} (which uses $\hat{\mathcal{B}}$).

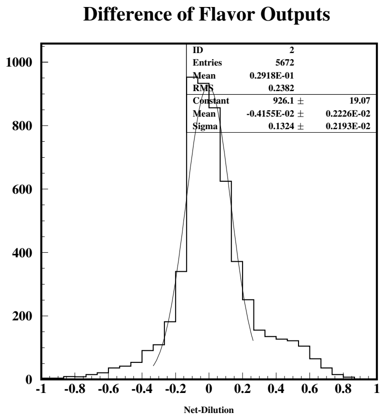


Figure 3.8: Difference of the tagger outputs $|f_{NN}|$ and $|\hat{\mathcal{B}}|$ for CLEO II data. The net skew to the right indicates that the neural net tagger, on average, reports a higher and more reliable dilution.

detector CP asymmetries, as simulated by the Monte Carlo.

The plot figure 3.9, showing a CP comparison of the final tagger output, shows no significant asymmetry. Mathematically, the tagger is manifestly CP symmetric, as discussed above. Hence this plot probes the MC prediction for detector CP asymmetries. We find that the Monte Carlo simulation of the detector indicates that there is no significant CP-tagging asymmetry at the 1% level.

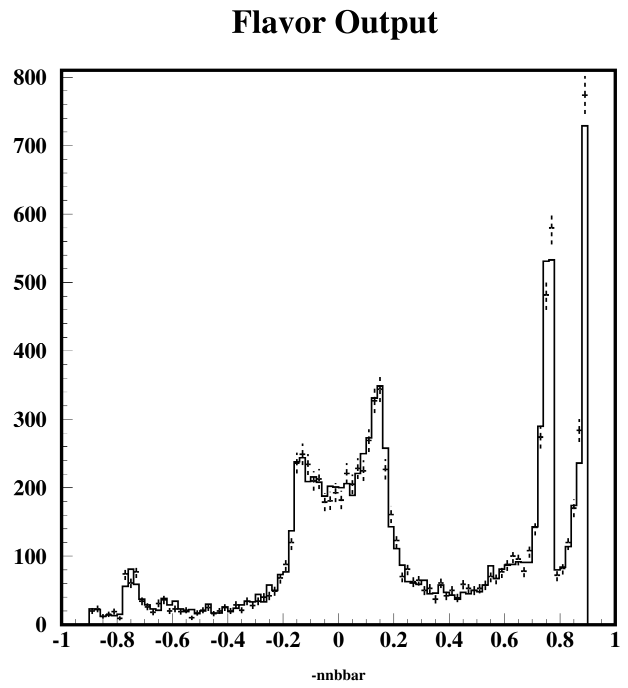


Figure 3.9: F_{NN} for 10k B events, and -F_{NN} of the tagger for 10k $\overline{B}$ events. This plot probes CP violation in the detector simulation.

Since this is an independent sample of events, we can also examine the tagger separation. The relation between reported and true dilution is well-behaved, having a slope of 0.99. We find εD^2=0.265 in this sample.

3.6 Measuring Performance in Data

There are three approaches to measuring the performance of a neutral B flavor tagger. The first, and most practical, is to give up any hope of ever making a time-integrated mixing measurement, and use the known world average to test the tagger. The second is to measure time-dependent mixing, where χ_d and $\mathcal{D}$ may be determined separately. The third, and least attractive, is to gain confidence in the Monte Carlo by using tests on the $D^*\ell\nu$ and $B^{\pm}$ data samples, and then to believe the Monte Carlo. We rely on the first method.

3.6.1 Using Time-Integrated Mixing of Fully Reconstructed Events

In this method, we must assume the world average for mixing, but we get to test the tagger directly on data and eliminate any dependence on Monte Carlo. Or, as Alexey Ershov put it, "You sacrifice χ_d, and all your sins are forgiven" [56].

The idea is to find in the data a sample of fully reconstructed, flavor-tagged neutral B's. Examples are $D^*\pi$ or $D^*\ell\nu$. We then measure "mixing" in this sample. Of course, because the flavor tagger is not perfect, we do not observe the true value of mixing. We reverse-engineer the tagging purity from the observed χ_d and the well-known world average.

We count the number of events tagged as same sign (apparently mixed) and opposite sign (apparently unmixed). We call this $\chi_{apparent}$. If w is the mistag fraction (1-P),

$$\chi_{apparent} = \chi(1-w) + (1-\chi)w \tag{3.8}$$

From eqn 3.2 we find

$$\mathcal{D} = 1 - 2w \tag{3.9}$$

and, after algebra we find

$$\mathcal{D} = 1 - 2\tfrac{\chi_{apparent} - \chi}{1-2\chi} \tag{3.10}$$

With enough events, this procedure may be followed bin-by-bin in expected dilution, to yield a map of expected dilution versus true dilution directly from the data.

This measurement was carried out, using reconstructed $D^*\ell\nu$ events, and is described in the flavor-tagging note [55]. The results are summarized in tables 3.1 and 3.2. We find the data results are consistent with what our results from the Monte Carlo, and that the tagger has εD^2=0.22$\pm$0.04 in the CLEO II data (for which it was optimized.)

3.6.2 Time-Dependent Mixing

One may measure time-dependent mixing in the CLEO II.V detector in a manner very similar to that described for $\sin 2\beta$ in [57]. In the analysis, one observes

Table 3.1: The slopes found in the dilution versus tagger output plots and the implied effective efficiency, ϵD^2 (corrected for the slope and B^0-$\overline{B}^0$ mixing) for the two tagging algorithms, in CLEO II data

	Signal MC (no mixing)		Generic MC		CLEO II Data	
Tag	*slope*	ϵD^2	*slope*	ϵD^2	*slope*	ϵD^2
$\widehat{B}$	0.93 ± 0.01	0.22 ± 0.01	0.93 ± 0.05	0.21 ± 0.02	0.93 ± 0.09	0.21 ± 0.04
F_{NN}	1.00 ± 0.01	0.27 ± 0.01	0.97 ± 0.05	0.25 ± 0.02	0.93 ± 0.09	0.22 ± 0.04

the flight distance of the fully-reconstructed B decay (such as $D^*\ell\nu$ or $D^*\pi$) as a function of the flavor tag output. The decay time distributions will be different for events tagged as mixed or unmixed, according to the distributions found in 1.39:

$$R_{unmixed}(t) \quad = \quad 2C|A|^4 \frac{e^{-\Gamma t}}{\Gamma(1+x^2)}\{1 + x^2 + [\cos(\Delta m t) + x\sin(\Delta m t)]\} \tag{3.11}$$

For the rate to mixed events (both flavor tags of the same sign), $A_1 = A$, $\overline{A_1} = 0$, $\overline{A}_2 = A^*$, and $A_2 = 0$:

$$R_{mixed}(t) \quad = \quad 2C|A|^4 \frac{e^{-\Gamma t}}{\Gamma(1+x^2)}\{1 + x^2 - [\cos(\Delta m t) + x\sin(\Delta m t)]\} \tag{3.12}$$

The resulting asymmetry of mixed to unmixed events is thus

$$\begin{aligned} A_{unmixed/mixed}(t) &= \frac{\cos(\Delta m t) + x\sin(\Delta m t)}{1+x^2} \\ &= \frac{\cos(x(\frac{t}{\tau} - \frac{\tan^{-1} x}{x}))}{\sqrt{1+x^2}} \end{aligned} \tag{3.13}$$

The theoretical form of the decay times is shown in 1.5; the asymmetry is shown in 1.4.

In figure 3.10 we show the asymmetry in decay time distributions for mixed and unmixed events in Monte Carlo. It differs from the form shown in 1.4 due to significant decay length resolution effects. The mixing parameter Δm may be determined from the period of oscillation (only one oscillation is clearly visible in the plot), and the tagging purity from the amplitude of the oscillation. In this way, one could eliminate dependence on the world average for the value of mixing, as mixing and the tagging purity are measure simultaneously. However, this is an entire analysis on its own which we do not undertake here.

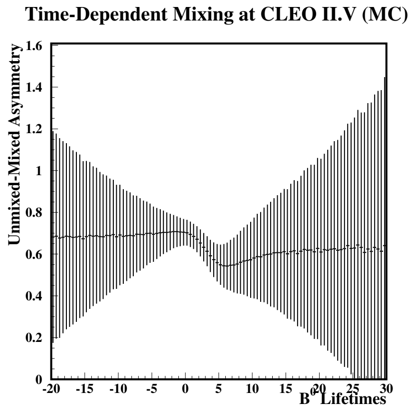


Figure 3.10: Asymmetry in mixed-unmixed distributions as a function of decay time of the B. The error bars correspond to a dataset of 10,000 background-free fully reconstructed B decays.

Table 3.2: The slopes and dilutions found in the CLEO II.5 data. The first column's values are directly from the D vs tag plot (ϵD^2 assumes $tag = D$). Assuming $\chi_d = 0.172$ we determine the true (corrected) values shown in the last column.

Tag	Raw Fit Values		Corrected	
	$slope$	ϵD^2	$slope$	ϵD^2
$\widehat{B}$	0.542 ± 0.090	0.259	0.83 ± 0.14	0.178 ± 0.060
F_{NN}	0.578 ± 0.089	0.267	0.88 ± 0.14	0.207 ± 0.065

CHAPTER 4

MEASURING $B^0 - \overline{B^0}$ MIXING

4.1 Introduction to $B^0 - \overline{B^0}$ Mixing

As we discussed in chapter one, flavor-mixing among the neutral B mesons allows mesons originally produced as B's to decay with non-zero probability as $\overline{B}$ mesons, and vice versa. The probability of such a flavor-mixing depends on the time between production and decay of the meson. We also saw that at the $\Upsilon(4S)$, the time-dependence of the decay flavor probabilities of the initially produced B and $\overline{B}$ mesons are entangled by Bose-Einstein statistics. This leads to the following form for the probability for an event to emerge unmixed (exactly one B^0 and exactly one $\overline{B^0}$) or mixed (exactly two B^0's or exactly 2 $\overline{B)}$'s), as a function of the two B decay times t_1 and t_2:

$$R_u(t_1, t_2) = Ce^{-\Gamma(t_1+t_2)}[1 + \cos(\Delta m(t_1 - t_2))] \tag{4.1}$$

$$R_m(t_1, t_2) = Ce^{-\Gamma(t_1+t_2)}[1 - \cos(\Delta m(t_1 - t_2))] \tag{4.2}$$

At a hadron, Z^0, or asymmetric machine, or with a precise vertex detector[58], one may attempt to observe the oscillations, whose frequency leads to a direct determination of Δm. However, one may integrate these expressions over all decay times of the two B's to find the total probability that an event emerges as mixed:

$$P_{mixed} = \frac{x^2+y^2}{2(1+x^2)} \tag{4.3}$$

where $x \equiv \frac{\Delta m}{\Gamma}$ and $y \equiv \frac{\Delta\Gamma}{2\Gamma}$. Please note the factor of 2 in the denominator of the y definition.

One may determine whether the event is mixed or unmixed by tagging the decay flavors of each of the two mesons in the event. The fraction of mixed events at the $\Upsilon(4S)$ is equal to χ_d, the total probability that a single meson produced as a B decays as $\overline{B}$, or *vice versa*.

4.2 Analysis Overview

The analysis compares the electric charge of a high-momentum π or ρ (henceforth, h_W) from one B meson in an event to the charge of a high-momentum lepton from the other B in the event. By requiring the h_W to be consistent with having come from the decay $B^0 \to D^{*\pm}h_W^{\mp}$, we reject most charged B decays. The remaining charged B backgrounds are primarily of a particular set–namely, 2-body decays–which are well-measured and easily modeled.

In order to determine whether the h_W is consistent with the decay $B^0 \to D^{*\pm}h_W^{\mp}$, we employ a partial reconstruction method described by Nelson, Nelson

, and Witherell in [6], and already used to great effect with $D^{*+}\pi$ decays in the thorough note [62]. The technique is an extension of that described by Kinoshita in [63], using current knowledge of the B mass. We look for the products of the decay chains:

$$\begin{tikzcd}[column sep=small]
\Upsilon(4S) \arrow[r] & B^0 \arrow[d] & \overline{B}^0 \arrow[r] & D^{*+} h_W^- \arrow[d] \\
& X\ell\nu & & D^0\pi_s
\end{tikzcd}$$

We partially reconstruct the D^* decay chain using only the detected π_s and h_W. The method of partial reconstruction is described in more detail in a following section. Once reconstructed, the analysis counts the number of $h_W^+\ell^+$, $h_W^-\ell^-$, and $h_W^{\pm}\ell^{\mp}$ event combinations, in the on-resonance and off-resonance data. After a continuum subtraction, the quantity χ_{raw} is defined as

$$\chi_{raw} = \frac{N_{h^+\ell^+} + N_{h^-\ell^-}}{N_{h^+\ell^+} + N_{h^-\ell^-} + N_{h^{\pm}\ell^{\mp}}} \tag{4.4}$$

However, the counts N are known to be incorrect due to flavor mistags. The counts are sequentially corrected for three effects: mistagging leptons, mistagging h_W, and charged B contamination. The intricacy of the analysis consists in nailing the size and uncertainty of the last two corrections. To set the scales, we expect to find around 2000 events, and we find in the end that the leptons have a mistag rate of 3.7%, hadrons 3.1%, and that the $B^{\pm}$ contamination is 13%.

A note on the charge conjugation convention. For notational simplicity, we refer to all neutral B's as B^0, and all charged B's as B^+. Throughout this note, charge conjugation for individual particles or decay chains is implied.

4.3 The Lepton Tag

B flavor determination using lepton tagging is a well-known technique. In most respects the leptonic portion of this analysis presents nothing unusual. The electrons are identified primarily through use of the calorimeter, and the muons primarily through the muon chambers. The lepton tag is subject to several flavor-mistagging effects which we must correct for. Some are physics processes–cascades, J/ψ, and photon conversions–and others are due to detector imperfections–misidentified hadrons.

The lepton tag in this analysis actually serves a dual purpose. The lepton tags the decay flavor of the B meson. However, the presence of a hard lepton also helps to suppress continuum $e^+e^- \to q\overline{q}$ events. This suppression is especially important, as otherwise the D^*h_W tagging is contaminated by continuum backgrounds.

4.3.1 Selection Criteria

We select leptons from among well-measured tracks. We require lepton identification by `R2ELEC`>3 or `DPTHMU`>5. We require the lepton to have momentum at least 1.4 GeV. We require the cosine of the angle between the lepton and the π_s candidate in the event to be less than 0.75 to reject secondary leptons from the decay of the unreconstructed D^0. If more than one lepton in an event satisfies these criteria, we select the lepton of the highest momentum. If the lepton forms an invariant mass within 50 MeV of the J/ψ mass with any track in the event other than the h_W candidate, the lepton is rejected. If the event has been reconstructed

Table 4.1: Lepton selection criteria.

Species	Requirement
e	$\texttt{R2ELEC} > 3$
μ	$\texttt{DPTHMU} > 5$
e or μ	$\texttt{KINCD} = 0$
e or μ	$\texttt{TNG} \geq 0$
e or μ	$\lvert\texttt{Z0CD}\rvert$ <3 cm
e or μ	$\lvert\texttt{DBCD}\rvert$ <2 mm
e or μ	$\frac{\mathrm{Hitsfound}}{\mathrm{HitsExpected}}$ >0.2
e or μ	p_ℓ >1.4 GeV
e or μ	$\cos\theta_{\ell-\pi_s}$ <0.75
e or μ($D^*\pi$ only)	$\lvert\cos\theta_{\ell-\mathrm{thrust}}\rvert$ <0.8
e or μ	$\lvert m_{\ell-\mathrm{track}} - m_{J/\psi}\rvert$ >50 MeV
e or μ	Highest-momentum lepton in event

as a $D^*\pi$, we additionally require that the lepton is not near the thrust axis of the event, in order to reject continuum backgrounds. For the $D^*\rho$ mode, the signal to noise is not appreciably increased by this requirement and it is not imposed. These criteria are summarized in table 4.1.

4.3.2 Physics Backgrounds to the Lepton Tag

There are three important sources of identified leptons which do not originate from the primary semileptonic decay of a B meson: true leptons (e/μ) from secondary charm decays, true leptons (e/μ) from J/ψ decays, and misidentified hadrons (mostly μ). Less important sources of background leptons include δ rays (e), γ conversions (e), Dalitz decays of the π^0 (e), and in-flight decays of πs and Ks (mostly μ). Each of these sources of leptons is addressed in turn in the following sections.

Cascades

Cascade leptons, also known as "secondary leptons", are true leptons that come from the semileptonic decay of charmed mesons produced in B decays. The decay chain $b \to c \to s\ell^-\nu$ produces leptons of sign opposite to the primary decay $b \to c\ell^+\nu$. In this sense, secondary leptons are biased in that the leptons have a definite charge with respect to the true flavor of the B.

There are two cuts in the analysis to discriminate against cascade decays. The first is a straightforward momentum cut, and the second an angular cut with respect to the soft pion direction.

As fig 4.1 shows, the momentum spectrum of generic secondary charm leptons lies almost entirely below 1.4 GeV. Leptons above this momentum are very purely (>96%) from primary B decays. A positively charged lepton indicates a B meson, and a negatively charged lepton indicates a $\overline{B}$ meson.

We require the momentum of the lepton to be above 1.4 GeV. From the semilep-

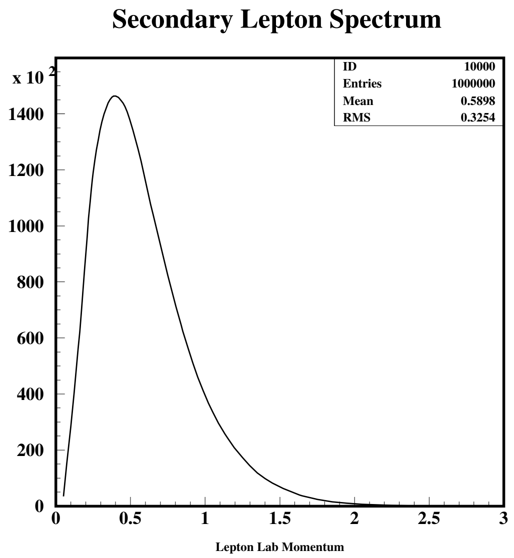


Figure 4.1: The spectrum of secondary leptons, from toy Monte Carlo combining D semileptonic and D momentum measurements.

tonic B Monte Carlo, this criterion retains 42±2% of the signal. To estimate the fraction of cascade leptons retained by this cut, we have reevaluated the cascade lepton spectrum, using the recently measured D^0 and D^+ momentum spectra[65], and the different semileptonic branching fractions of the D^+ and D^0, and the charm semileptonic spectrum.

We generate our own charm semileptonic spectrum using measured values of the form factors and D branching fractions. We use the review [66] to obtain values of the semileptonic branching fractions and form factors, as listed in tables 4.2 and 4.3. We combine these form factors and branching fractions with the equations for semileptonic decay widths to generate a D semileptonic spectrum. We can then combine the spectrum of leptons from D decay with the CLEO measurements of the D momentum spectrum to form the secondary spectrum.

The equations [1] to generate the semileptonic spectrum for the decay modes are

$$\frac{d\Gamma_{scalar}}{dq^2 d\cos\theta_l} \quad \alpha \quad \sin^2\theta_l p_K^3 |f_+(q^2)|^2 \tag{4.5}$$

$$\begin{aligned}
\frac{d\Gamma_{vector}}{dq^2 d\cos\theta_l d\cos\theta_V d\cos\chi_l} \quad & \alpha \quad p_V q^2 [(1+\cos\theta_l)^2 \sin^2\theta_V |H_+(q^2)|^2 \\
& + \quad (1-\cos\theta_l)^2 \sin^2\theta_V |H_-(q^2)|^2 \\
& + \quad 4\sin^2\theta_l \cos^2\theta_V |H_0(q^2)|^2 \\
& + \quad \sin\theta_l (1+\cos\theta_l) \sin\theta_V \cos\theta_V \cos\chi H_+(q^2) H_0(q^2) \\
& - \quad \sin\theta_l (1-\cos\theta_l) \sin\theta_V \cos\theta_V \cos\chi H_-(q^2) H_0(q^2) \\
& - \quad 2\sin^2\theta_l \sin^2\theta_V \cos 2\chi H_+(q^2) H_-(q^2)]
\end{aligned} \tag{4.6}$$

where we use the values of $A_1(0), A_2(0), V(0), f_+(0)$, and α listed in table 4.3. For the q^2 dependence of the form factors we use $\exp^{-\alpha q^2}$ for the scalar case and $\frac{1}{1-\frac{q^2}{m_V^2}}$ for the vector case. The results are shown in table 4.4, and the expected fraction in 4.1. The uncertainties take into account the uncertainties due to the measurements of the D spectrum, the semileptonic branching fraction measurements, and the form factor measurements, as those uncertainties are given either by the PDG or [65].

[1] We warn the reader against blind duplication of these formulas. We found typos in 3 sources, including the PDG! Also note that the formulas are given here for $c \to s$ transitions and are not directly applicable to $b \to c$ transitions.

Table 4.2: D semileptonic branching fractions. Taken from [66].

Mode	Measurement (%)	Error (%)
$D^0 \to Kl\nu$	3.46	$\pm$ 0.16
$D^0 \to K^*l\nu$	2.01	$\pm$ 0.33
$D^0 \to \pi l\nu$	0.4	$\pm$ 0.1
D^0 exclusive sum	5.87	$\pm$ 0.46
$D^0 \to Xl\nu$	7.1	$\pm$ 0.3
Difference	0.83	$\pm$ 0.58
$D^+ \to Kl\nu$	6.7	$\pm$ 0.8
$D^+ \to K^*l\nu$	4.7	$\pm$ 0.4
$D^+ \to dl\nu$	0.7	$\pm$ 0.1
D^+ exclusive sum	13.1	$\pm$ 0.9
$D^+ \to Xl\nu$	17.2	$\pm$ 1.9
Difference	3.9	$\pm$ 2.1

Table 4.3: D semileptonic form factors. Taken from [66].

Form factor	Measurement	Error	Experiment(s)
$A_1(0)$	0.55	0.03	(E687/E691)
$A_2(0)$	0.40	0.08	(E687/E691)
$V(0)$	1.0	0.2	(E687/E691)
$f_+(0)$	0.74	0.03	(CLEO II)
α_{scalar}	0.29	0.07	(CLEO II)

We find that f_c, the fraction of selected leptons which result from the decay of charm cascades, is consistent with those in the previous CLEO studies, with greatly reduced uncertainties, due to the improved knowledge of the D spectrum. The resulting secondary spectrum is plotted in fig. 4.1. We obtain uncertainties from statistical errors that result from running 1000 experiments with D and $D \to Xl\nu$ spectra generated according to the central values and Gaussian errors of the CLEO D momentum spectrum and varying the D semileptonic form factors and branching fractions about their uncertainties. We note that this shape and its normalization are not the same for $B^0\overline{B^0}$ events as for B^+B^-, as the relative fraction of D^0 and D^+ (with their very different semileptonic branching fractions) is different in the two cases. We have found that the fraction of charm leptons which satisfy the minimum momentum requirement of 1.4 GeV is 2.4$\pm$0.4%.

We may also use the DELCO leptonic momentum spectrum[64] instead of the

Table 4.4: Fraction f_c of selected leptons that originate from cascade decays, for several lepton momentum cuts. The fractions and errors are compared to the study done for the previous CLEO B mixing measurement in [60].

Momentum Cut (GeV)	Fraction, this work (%)	Uncertainty, this work (%)	Previous Fraction (%)	Previous Uncertainty (%)
1.0	11.4	0.6		
1.2	7.7	0.5	5.4	1.6
1.4	3.4	0.3	2.8	0.8
1.5	2.1	0.2	2.0	0.6
1.6	1.3	0.2	1.4	0.4

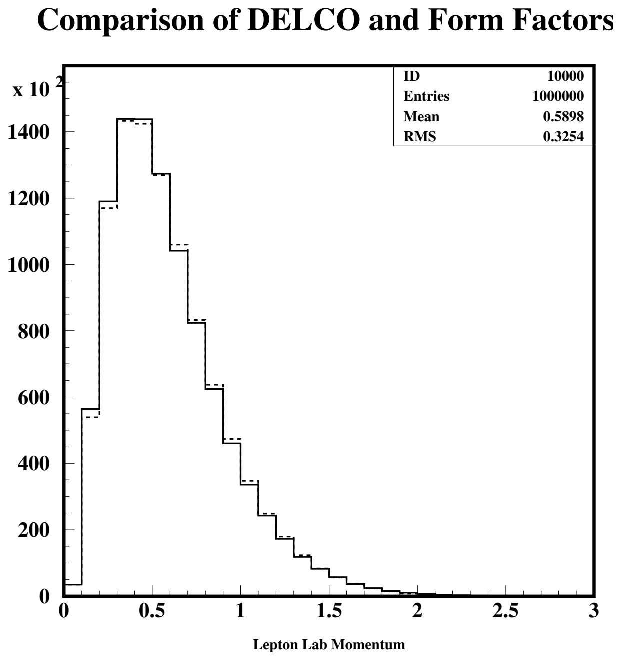


Figure 4.2: The secondary lepton ($b \to c \to sl\nu$) spectrum in CLEO. Shown are the values of the branching fraction per 100 MeV, averaged over many toy Monte Carlo runs. The dashed histogram uses the DELCO measurement, the solid histogrammed is generated from form factors.

form factors and D semileptonic branching fractions. A comparison of the results using the DELCO measurements and form factors is shown in figure 4.2.

The second cascade cut rejects leptons near the soft pion direction. In the $D^*\pi$ or $D^*\rho$ decay chains, the unreconstructed D comes out with momentum at nearly the kinematic limit for D decays– around 2.3 GeV. The cascade leptons from this D decay tend to be much more highly boosted than generic cascade leptons, and hence more likely to satisfy the momentum requirement of the analysis. The situation is shown in figure 4.3. To reduce this source of background, the cosine of the angle between the lepton and the soft pion is required to be less than 0.75. The requirement retains 82% of signal (it nearly amounts to a fiducial acceptance cut) while eliminating 95% of this cascade background.

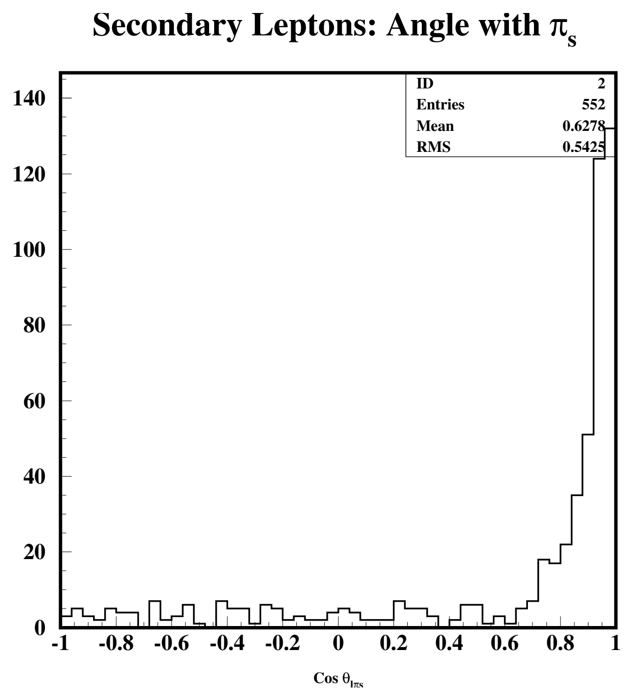


Figure 4.3: The cosine of the angle between secondary leptons and the π_s direction. The sharp peak is due to semileptonic decays of the fast D^0 in the partial reconstruction decay chain.

The total cascade fraction f_c after both cuts is found to be 3.4±0.3%, where the error comes from a combination of the systematic error due to the shape of the cascade spectrum (originating from uncertainties in the D spectrum and form factors) and the statistical error in the Monte Carlo.

We note that the D momentum spectrum and branching fraction measurement we use from [65] is not flavor-specific; we find that the spectra for leptons originating from $B \to D$ and $B \to \overline{D}_{(s)}$ are broadly similar in generic Monte Carlo. We take the so-called "upper-vertex" (correct-sign) lepton branching fraction to be 20% of the lower-vertex fraction, as indicated by the Monte Carlo, of which 10% is from wrong-sign D decays and 10% from D_s decays. This fraction is also supported by measurements of the wrong sign D and D_s spectra [82, ?]. Leptons from upper vertex production $b \to c\overline{c}s$, $\overline{c} \to \overline{s}\ell^- \nu$ correctly tag the B flavor, and therefore reduce the error due to secondary leptons.

J/ψ Decays

Leptons arising from J/ψ decays might seem innocuous, as they are, of course, half positive and half negative. However, it is better to think of this background as half mixed and half unmixed[2]. Hence it can bias the measured value of the mixed fraction toward 0.50.

To reduce J/ψ backgrounds, we form the invariant mass of all combinations of leptons (e or μ) with oppositely-charged tracks. If any such combination is within 50 MeV of the known ψ mass, the lepton is rejected. This veto is 70±6% efficient at eliminating J/ψ in Monte Carlo. We expect 16±5 ψ events to escape the veto in the full data set.

[2]The effect is more subtle in dilepton analyses, where both leptons may be counted.

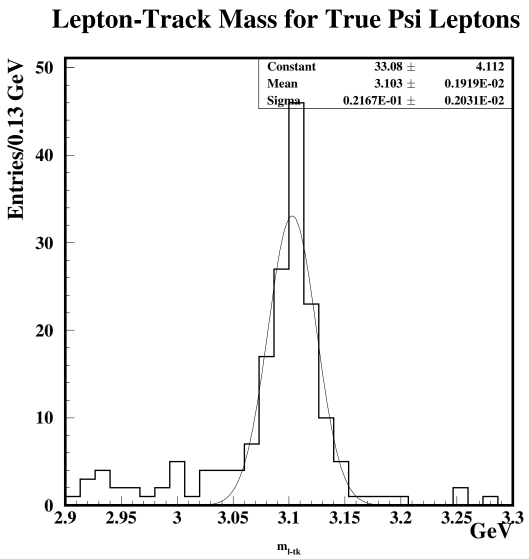


Figure 4.4: The mass spectrum of the best lepton-track combinations for leptons from ψ decays. The analysis vetoes leptons for which the best combination is within 50 MeV of m_{ψ}.

Photon Conversions

As with J/ψ decays, electrons from photon conversion are half mixed and half unmixed. We expect only one event from γ conversion in the full data set.

π^0 Dalitz Decays

Electrons from π^0 Dalitz decays are half mixed and half unmixed. In 20.1 fb^{-1} of Monte Carlo, we find one electron from this source. We take this correction directly from the Monte Carlo. The small size of this correction does not surprise us. π^0 decay is a three-body process, and π^0s from B decays are limited to momenta of

2.45 GeV or less.

δ Rays

We mention δ rays only for completeness. Knock-ons essentially never reach 1.4 GeV; no δ rays were selected in the generic Monte Carlo. The fact that they are always negative affects mixing in the same way as J/ψ decays or γ conversions. However, in contrast to those backgrounds, they could in principle generate a false CP violation signal.

In-flight Decays

In-flight decays $\pi \to \mu\nu$ or $K \to \mu\nu$ can have a specific charge correlation to the decay flavor of the parent B meson, since for instance K's generally tag the B decay flavor correctly. However, no events were seen in 20 fb^{-1} of Monte Carlo.

4.3.3 Hadronic Fakes

In principle hadronic tracks which are falsely identified as leptons can have a specific charge relation to the decay flavor of their B ancestor, as charged K's have a higher fake rate and generally good correlation to the decaying B flavor. In practice, however, there being so few hadronic fakes, this effect is negligible.

Once the essential R2ELEC and DPTHMU cuts have been made, there is not much to separate hadronic fakes from electrons and muons, although the muon fake rate is higher than that of electrons.

For the average hadron in the momentum region applicable to this analysis, the total hadronic fake rate (π + K) for electrons and μs is 0.15%/tk and 1.2%/tk, respectively, as determined by the lepton Systematics Advisory Committee [67]. As we will see presently, the correction for and uncertainty in the fake rate can be obtained from data, without reliance on the Monte Carlo spectrum of unassigned tracks or the fake rates determined in [67].

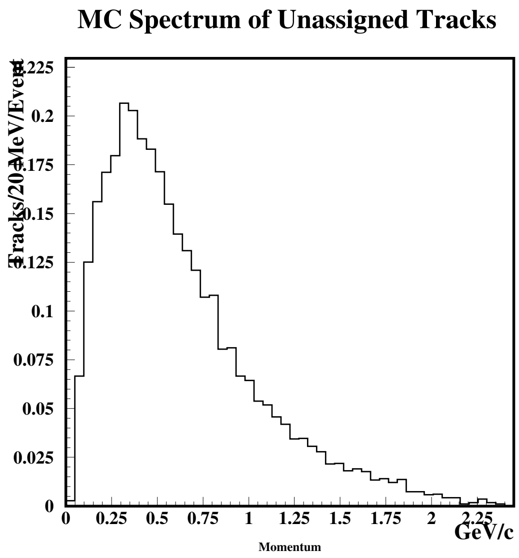


Figure 4.5: The spectrum of unassigned (i.e., not π_s,h_W, or identified e/μ) tracks in 2-body $B\overline{B}$ events selected by either the $D^*\pi$ or $D^*\rho$ selection criteria.

4.3.4 Evaluation of Lepton Fake Rates

We use two data samples in order to evaluate lepton fake rates. The first is a sample of $K_s \to \pi\pi$, the other $D^{*+} \to D^0\pi_s\,, D^0 \to K^-\pi^+$. These samples provide clean sources of π's and K's in order to measure the fake rates. In all studies, the leptons must satisfy all the requirements of table 4.1, except for those related to the hadronic tag ($\cos\theta_{\ell-\pi_s}$ <0.75 and $|\cos\theta_{\ell-\mathrm{thrust}}|$ <0.8).

Selection of $K_S \to \pi^+\pi^-$

We select vertices from `knvfc100` with the following requirements:

Table 4.5: Selection requirements for K_S used in the fake rate study.

Variable	Requirement
`IDVO`	= `KSHHYP`
`CHIVTX`	< 10
`FVSGNF`	> 5
`RBMTX`	Between 1 and 5 cm
`VETO_VO`	Not also a Λ
subroutine	or γ candidate

In addition, at least one of the tracks must be in the momentum region of interest (above 1.4 GeV).

We do not wish to use K_S that originate far out in the drift chamber. Such tracks have poorer $\frac{dE}{dx}$ resolution and hence are more likely to overlap with the electron $\frac{dE}{dx}$ band than generic tracks from the origin, which we expect to be the main component of our background. Such tracks are also more likely to suffer from systematic entrance-angle miscalculation in the `dedr` code [3]. This is why restrict

[3]The calculation is known to be incorrect for tracks with significant impact

ourselves to vertices between 1 and 5 cm.

Background Subtraction for K_S

There is a large background from $\gamma \to ee$ events. When checking the fake rate for one track we require the other track to not be identified as an electron (`R2ELEC`<-10). We require the vertex mass to be within 6 MeV of 497.6 MeV. We also perform a sideband subtraction, using vertices from 6 to 12 MeV from the peak.

We find that the background from conversions is not symmetric about m_K, and in particular that it peters out at around 510 MeV. Previous studies may have systematically misestimated the size of this background by assuming that a sideband subtraction properly accounted for it. The background shape and normalization also have a strong dependence on $\cos\theta$. We perform two measurements, one restricting our range to $|\cos\theta|$ <0.85, and the other out to 0.93. We find that the solid-angle-weighted average fake rate is 0.16±0.10%. The fake rate results are shown in table 4.6.

Selection of D^* Candidates

We require `DBCD` for π_s tracks to be less than 3 mm, and less than 2 mm for D tracks. We require all tracks to have `KINCD`=0.

To select D candidates, we require that the invariant mass of track pairs be within 80 MeV of the D mass, when one track is assigned the K mass and the other the π mass.

---

parameter

Table 4.6: Summary of fake rate results for π's from K_S decays. Two previous measurements are provided for comparison.

Dataset/Msmnt	$\pi \to e$	$\pi \to \mu$
CLEO II	0.16±0.10%	1.46±0.07%
CLEO II.V	0.03±0.20%	1.5±0.1%
Lepton SAC	0.1%	
Bellerive		1.3%

A D^* candidate is selected if any additional track in the event, with the same charge as the π daughter of the D^0, when assigned the π mass, forms a mass difference $m_{K\pi\pi_s} - m_{K\pi}$ less than 150 MeV.

We select, finally, those D^* candidates for which Δm is less than 1.5 MeV from the nominal value. When evaluating the π fake rate, we require that the K PID is consistent with the K hypothesis at the 0.1% level, and that the K hypothesis is more probable than the π hypothesis, using "level cuts" weighted by species multiplicity.

We evaluate the fake rates using $D*$ candidates for which the D mass is within 25 MeV of 1.865 GeV. We perform a sideband subtraction using events from D masses 1.79 to 1.815 GeV and 1.905 to 1.93 GeV. We first count the number of background-subtracted D^0 candidates in the data. To evaluate the fake rates, we then count the number of background-subtracted D^0 candidates in the data, requiring in turn that the π or K daughter be identified as an e or μ. The results are shown in table 4.7.

Table 4.7: Summary of fake rate results for π's and K's from D^* decays. Two previous measurements are provided for comparison.

Dataset	$\pi \to e$	$\pi \to \mu$	$K \to e$	$K \to \mu$
CLEO II	0±0.2%	0.8±0.3%	0.2±0.1%	3.5±0.6%
CLEO II.V	0.3±0.2%	1.0±0.4%	0±0.2%	4.9±0.8%
Lepton SAC	0.1%		0.05%(??)	
Bellerive		1.3%		2.5%

Combined Results for Fake Rates

We combine fake rates from K_S, D^*, CLEO II, and CLEO II.V; the small discrepancies are covered by large systematic uncertainties of 50The results are shown in table 4.8.

Table 4.8: The fake rates we use for the mixing analysis. The errors are statistical only. The systematic variation of 50% we employ is given in table 4.28.

$\pi \to e$	$\pi \to \mu$	$K \to e$	$K \to \mu$
0.13±0.08%	1.4±0.2%	0.2±0.1%	4.0±0.5%

4.3.5 Data Determination of Fake Backgrounds

In order to determine the level of fake backgrounds, we analyze the data requiring the "lepton" to be **not** identified as an e or μ. All other selection criteria of the analysis are imposed. We analyze the 4s4, 4s8, 4s9, and 4sG, and 4sH, 4sM, 4sN, and 4sP datasets. We find the event counts listed in table 4.9. We multiply by the ratio of luminosities to the full dataset. We use the fake rates of π and K in the ratio 2:1, the average multiplicity ratio in Monte Carlo for tracks of momentum above 1.4 GeV satisfying the analysis criteria. Altogether we find a fake rate of 2.3%. We use this fake rate to determine the number of fakes we expect in the data. As shown in table 4.9, we find that we should expect around 5 hadrons misidentified as leptons in both the mixed and unmixed samples of the analysis.

Table 4.9: Measurement of the magnitude of misidentified hadron backgrounds in the analysis.

Data	Mixed Fakes	Unmixed Fakes	$\mathcal{L}$
ON	55	94	1.98 fb^{-1}
OFF	5	25	1.16 pb^{-1}
ON-OFF	46.5±8.3	51.3±12.9	1.98 fb^{-1}
Full data, e	0.5±0.2	0.5±0.2	9.1 fb^{-1}
Full data, μ	4.9±0.9	5.4±1.4	9.1 fb^{-1}

4.3.6 Lepton Summary

In table 4.10 we summarize the expected level of electrons from sources other than primary B decays, as a function of the momentum requirement on the lepton. In table 4.11 we summarize the same information for muons. Table 4.12 provides a

Table 4.10: Summary of electron backgrounds: percent of electrons selected from non-primary-$B\overline{B}$ sources, as a function of electron momentum. Though hadronic fakes are tabulated in table 4.9 by mixed or unmixed, the sum appears in this table.

Lepton Source	Secondary Leptons	J/ψ Leptons	Hadronic Fakes	Others (In-flight, γ,π^0)
Momentum Cut				
1.0 GeV/c	11.3±0.6	0.2±0.1	0.2±0.05	0.8±0.3
1.1 GeV/c	9.5±0.6	0.2±0.1	0.2±0.05	0.6±0.3
1.2 GeV/c	7.6±0.5	0.3±0.2	0.1±0.03	0.3±0.2
1.3 GeV/c	5.5±0.4	0.3±0.2	0.1±0.03	0.2±0.2
1.4 GeV/c	3.4±0.3	0.4±0.2	0.1±0.03	0.2±0.2
1.5 GeV/c	2.1±0.2	0.4±0.2	0.1±0.03	0.3±0.3
1.6 GeV/c	1.3±0.2	0.5±0.3	0.1±0.03	0.3±0.3

summary for both lepton species. By adding 90% of the secondary fraction (to account for 10% upper vertex D decays), and 50non-primary lepton sources, we find mistag rates of 3.6% for electrons and 3.8% for muons.

In figure 4.6 we show the expected lepton components as a function of lepton momentum, overlaid with the lepton spectrum of events selected from the data.

Table 4.11: Summary of μ backgrounds: percent of muons selected from non-primary-$B\overline{B}$ sources, as a function of muon momentum. Though hadronic fakes are tabulated in table 4.9 by mixed or unmixed, the sum appears in this table.

Lepton Source	Secondary Leptons	J/ψ Leptons	Hadronic Fakes	Others (In-flight, γ,π^0)
Momentum Cut				
1.3 GeV/c	5.5±0.4	0.7±0.4	0.7±0.2	0.3±0.3
1.4 GeV/c	3.4±0.3	0.7±0.4	0.7±0.2	0.3±0.3
1.5 GeV/c	2.0±0.2	0.7±0.4	0.7±0.2	0.3±0.3
1.6 GeV/c	1.2±0.2	0.9±1.0	0.6±0.3	0.4±0.4

4.4 The Hadronic Tag

In the note [6], Nelson, Nelson, and Witherell describe a method for partial reconstruction of $D^* n\pi$ decays of the B meson, when only the $n\pi$ and π_s are detected while the D decay goes unreconstructed. The method depends on the special kinematics of a series of 2-body decays. The technique results in a reconstruction of the three decay angles describing the decay chain.

We apply this technique to select a sample of flavor-tagged B^0 decays. Being a partial reconstruction, the analysis does select a non-negligible fraction of events other than true signal decays. This means that some events are not B^0 decays,

Table 4.12: Summary of lepton yields and backgrounds, for signal Monte Carlo with reconstructed with a good D^*h_W candidate. Primary leptons are those that arise from $B \to X\ell\nu$ decay.

Momentum Cut	Primary e/fb^{-1}	Background e/fb^{-1}	Primary μ/fb^{-1}	Background μ/fb^{-1}
1.0 GeV/c	137±2.6	17.1±1.0	66±1.8	4.2±0.4
1.1 GeV/c	124±2.4	13.0±0.9	66±1.8	4.2±0.4
1.2 GeV/c	110±2.3	9.1±0.7	66±1.8	4.2±0.4
1.3 GeV/c	98±2.2	5.9±0.5	66±1.8	4.2±0.4
1.4 GeV/c	83±2.0	3.4±0.3	64±1.8	3.2±0.4
1.5 GeV/c	69±1.8	2.0±0.3	59±1.7	2.1±0.3
1.6 GeV/c	55±1.6	1.2±0.2	50±1.5	1.9±0.4

and others have not correctly tagged the B flavor.

In order to analyze the charged B fraction and hadronic mistag rates, we subject the selected hadronic tags to fits in two of the reconstruction variables. In these fits, we determine the level and compositions of non-$D^{*\pm}h_W^{\mp}$ backgrounds. Together with Monte Carlo estimates of the mistag rates for each background, we learn both the total hadronic mistag rate and the charged B fraction. Here "mistag rate" means the fraction of events for which the beauty quantum number B is mismeasured. The uncertainties from the fit provide uncertainties for the mistag rates and $B^{\pm}$ fraction.

The fits are carried out independently of whether events are mixed or unmixed. They also cover a much greater space than we will ultimately use to measure χ_d, in order to provide a better uncertainty on the background levels. We simultaneously fit two 2-dimensional plots, where the axes are two decay angles describing the decay chain. The two plots are differentiated by the angle between the hard hadron and the π_s; one is a signal fitting region and the other is a background-dominated sideband region useful for normalizing the background levels.

In the following sections, we describe the partial reconstruction method and selection of the events for the fits and measurement of mixing. We then describe the physics content of each of the shapes and their simulation. We then describe the fitting procedure, and the fit results.

4.4.1 Partial Reconstruction Method

In the partial reconstruction we focus on the $D^*\pi$ and $D^*\rho$ decays of the B meson. The π_W in B decays comes out at 2.3 GeV, and the ρ_W at 2.15 GeV. Both spectra are smeared by the 300 MeV/c momentum of the B. The π_s from the decay of the D^* has a momentum between 80 and 220 MeV.

The reconstruction proceeds as follows. We measure the masses and momenta of the π_s and the h_W. From conservation of four momenta we know that

$$\begin{aligned} E_D &= E_{D^*} - E_{\pi_s} \\ &= E_B - E_{h_W} - E_{\pi_s} \end{aligned} \tag{4.7}$$

$$\vec{p}_{D*^+} = \vec{p}_D + \vec{p}_{\pi_s} \tag{4.8}$$

$$\vec{p}_{D*^+} = \vec{p}_B - \vec{p}_{h_W} \tag{4.9}$$

We also know

$$|\vec{p}_B| = \sqrt{E_{beam}^2 - m_B^2} \tag{4.10}$$

$$|\vec{p}_D| = \sqrt{E_D^2 - m_D^2} \tag{4.11}$$

With $|\vec{p}_D|$ known from (4.7) and (4.11), the equation (4.8) describes a cone on which the D^* momentum must lie. Similarly, with $|\vec{p}_B|$ known from (4.10), the equation (4.9) also describes a cone on which the D^* momentum must lie. These cones must intersect for events which truly originate from the hypothesized decay chains, and in general will do so in two places. Figure 4.7 shows these cones.

The entire decay chain can be described by the use of three decay angles. We use the two angles $\cos\theta_B$ and $\cos\theta^*_{D*}$, defined by the angles of the B and D^* decays with respect to their flight directions in the lab, and the angle ϕ. ϕ is defined as shown in figure 4.7; it is the half-angle subtended by the intersection points. In the limit that the Q values of all decays goes to 0, the cones will be tangent to one another, and ϕ is a delta function. We are actually quite close to this limit, the values of Q being 20 and 4.5 MeV in the decays of the Υ and D^*, respectively. The smearing is not too serious, and we see a narrow peak in cosϕ near 1. The three cone angles are related to the measured momenta by formulas given in [6]. At the end of the day we can also solve for the two allowed D momentum vectors.

The power of this technique comes from the fact that by a stupendous margin, random pairs of hard and soft πs are in general flatly inconsistent with the hypothesis that they come from the decay chains above. This inconsistency is manifested by failure of one of the angles to have a cosine between -1 and 1. Even decay chains closely related to $B^0 \to D^{*+}\pi^-$, such $B^0 \to D^{**+}\pi^-$, have rather broad structures in cosϕ.

We exploit the fact that signal and background events have differing distributions in the reconstructed decay angles. Most striking is the difference in $\cos\theta_{D^*}$; because signal events have extremely high longitudinal polarization (100% and 87% for $D^*\pi$ and $D^*\rho$, respectively), they exhibit a nearly pure $\cos^2\theta_{D^*}$ distribution, whereas the backgrounds are distributed quite broadly. Signal events also peak near 0.6 in $\cos\phi$, with backgrounds exhibiting smoother distributions.

We also exploit a sideband in the measured angle between the π_s and h_W. The sideband region extends from -0.80 to -0.90 in the cosine of this angle. There is very little true signal in the sideband region, and it therefore provides an excellent means for normalizing the background levels.

4.4.2 The Analysis Angles

There are three angles of interest that can be reconstructed in the partial reconstruction:

cosθ_B : the angle, in the rest frame of the B, between the D^* and the negative of the direction of the lab frame

cosθ^*_{D*}: the angle, in the rest frame of the D^*, between the π_s and the negative of the direction of the lab frame

cosϕ: the cone intersection angle

These three variables have distinctive distributions for signal and backgrounds.

The cosθ_B angle, of course, must be flat for signal, as the B is a scalar. It is primarily a measure of the momentum of the h_W. Low values of cosθ_B correspond to lower momenta.

The distribution of cosθ^*_{D*} in the $D^*\pi$ mode shows the 100% polarization of the D^* that is required in a $P \to PV$ decay, and the 87% polarization that has been measured in the $D^*\rho$ mode. The angle $\cos\theta^*_{D*}$ is primarily determined by the measurement of the π_s momentum. Low values of cos$\theta^*_{D^*}$ correspond to lower π_s momenta. The lower lobe of the distribution is not as populated as the upper one because track finding efficiency is lower for softer π_s.

The variable cosϕ is a combination of both the π_s and h_W momenta, and of the angle between them. For most non-signal modes, it reconstructs outside the physical region between -1 and 1, whereas it is well-contained for the signal modes.

Figure 4.8 shows these three variables for signal modes for both $D^*\pi$ and $D^*\rho$ modes.

4.4.3 Hadronic Selection Criteria

In this section we describe the requirements we impose in order to select π_s-h_W combinations. First we describe a loose selection of candidates. We then describe tighter requirements which these events must satisfy in order to enter the signal or sideband fitting regions for the determination of sample composition. In order to measure the value of χ_d, we impose yet tighter restrictions on the events which enter the signal fitting region. We count, in the restricted signal region, the number of events for which the the lepton and h_W charges agree (mixed) and disagree (unmixed.)

Selection of Candidate π_s-h_W Combinations

We select soft tracks as potential D^* decay products. We select well-measured tracks of momentum less than 300 MeV. To form a potential h_W, we select either fast pions or fast ρ's. We require both π_f and the charged π for ρ decay to be well-measured, and to not be identified as leptons. We require a π^0 for a ρ be formed of a $\gamma\gamma$ combination consistent with the π^0 hypothesis. We require the $m_{\pi^\pm\pi^0}$ mass to be near the ρ mass. These criteria are described in table 4.13.

To form a possible π_s+h_W combination we require the cosine of the angle between the π_s and h_W, $\cos\theta_{h_W-\pi_s}$, is less than -0.8. We calculate the angular reconstruction quantities described above and require $|\cos\phi|$ <7 and $|\cos\theta_B| < 1$. The criteria are described in table 4.14. In 23% of decays, there is more than one possible combination which satisfies these requirements. When this occurs, we accept the "best" solution as described in order of descending preference in table

4.15. We prefer "right sign" ($q_s = -q_f$) solutions to "wrong sign" ($q_s = q_f$), and $D^*\pi$ to $D^*\rho$. If there is still more than one preferred solution, the one nearest ϕ of 0.6 is selected.

Selection of π_s-h_W Combinations for Fit

Once we have selected a set of candidate π_s-h_W combinations, we wish to impose yet stricter requirements in order to get a systematically reliable estimate of χ_d. At the same time, we wish to retain many of the rejected combinations in order to gain information concerning the composition of the selected sample.

To this end, we form two subsets of events. One subset coincides largely with the final selection of events for measurement of χ_d. The second is a set of events in the $\cos\theta_{h_W-\pi_s}$ sideband which aids in the determine the composition of the signal sample, and in particular the normalization of the random combinatoric background.

In order to select events the signal fitting region, we require that the h_W and π_s are back to back. We require that the h_W and π_s have the opposite charge sign, as they must in the hypothesized decay chain. We require that the reconstructed decay angles are nearly consistent with the decay chain $B^0 \to D^*h_W$. We require the π_s is in the momentum range allowed for the decay chain $B^0 \to D^*h_W$. The criteria are summarized in table 4.16. The analysis selects 2166 events in the signal fitting region.

In order to select events the sideband fitting region, we require that the h_W and π_s are not as back-to-back as the signal region. We require that the h_W and

Table 4.13: Soft pion and h_W selection criteria.

π_s **Requirements**
`KINCD` $= 0$ or 2
`TNG` ≥ 0
$\vert$`ZOCD`$\vert$ <5 cm
$\vert$`DBCD`$\vert$ <2 mm
$\frac{\text{Hitsfound}}{\text{HitsExpected}}$ >0.2
p_{π_s} <300 MeV
π_f **Requirements**
`KINCD` $= 0$
`TNG` ≥ 0
$\vert$`ZOCD`$\vert$ <3 cm
$\vert$`DBCD`$\vert$ <2 mm
$\frac{\text{Hitsfound}}{\text{HitsExpected}}$ >0.2
`DPTHMU`< 5
`R2ELEC`< 0
ρ **Requirements**
`CHSQ0`< 20
$\vert m_{\pi_f \pi^0} - m_\rho \vert$ <200 MeV
h_W **Requirement**
Found π_f or ρ_f
$1.9 < p_{h_W} < 2.5$

Table 4.14: Selection criteria for partial reconstruction with π_s-h_W.

Requirement
Found π_s
Found h_W
$\cos\theta_{h_W-\pi_s}$ <-0.8
$\lvert\cos\phi\rvert$ <7
$\lvert\cos\theta_B\rvert$ <1
"Best" π_s-h_W if multiple combinations

Table 4.15: Partial reconstruction solution preference order. The first two classes are "right-sign", the second two "wrong-sign".

$\pi_s\pi_f$, $q_f = -q_s$ (right-sign)
$\pi_s\rho_f$, $q_f = -q_s$
$\pi_s\pi_f$, $q_f = q_s$ (wrong-sign)
$\pi_s\rho_f$, $q_f = q_s$

Table 4.16: Selection of events for the signal region fit.

Requirement
$\cos\theta_{\pi_s - h_W}$ <-0.95
$\lvert\cos\phi\rvert$ <2
$\lvert\cos\theta^*_{D^*}\rvert$ <2
p_{π_s} <250 MeV
$q_{\pi_s} = -q_W$

Table 4.17: Selection of events for the sideband region fit.

Requirement
-0.90< $\cos\theta_{\pi_s - h_W}$ <-0.80
$q_{\pi_s} = -q_W$

π_s have the opposite charge sign. The criteria are summarized in table 4.17. The analysis selects 2131 events in the sideband fitting region.

Selection of Events for Measurement of χ_d

In order to measure mixing, we impose one final requirement on the events in the signal fitting region, requiring $\lvert\cos\theta^*_{D^*}\rvert$ <1. We find that because many backgrounds retain good charge information concerning the B decay flavor we do not have to restrict ourselves to the physical region $\lvert\cos\phi\rvert < 1$, and instead retain the more relaxed requirement $\lvert\cos\phi\rvert < 2$. This is displayed in table 4.18. The analysis selects 1982 events for the measurement of mixing in the on-resonance data, and

Table 4.18: Selection of events for the measurement of mixing.

Requirement
Selected for signal fit region
$\lvert\cos\theta^*_{D^*}\rvert < 1$

58 events in the off-resonance data.

4.4.4 The Fit Components

As mentioned previously, in addition to true signal events, we select events from several background sources. We divide the selected events into five categories. Four categories are from B decays; the fifth accounts for contamination from continuum $q\overline{q}$ contamination. The categories, which will be described in more detail, are

- Signal ($D^{*\pm}\pi^{\mp}, D^{*\pm}\rho^{\mp}$)
- Other 2-body B^0 decays
- 2-body $B^{\pm}$ decays
- Random combinatoric (non-2-body) B decays
- Continuum backgrounds

We perform fits in order to determine the contributions of each of these backgrounds to the events selected to measure mixing. The fits are performed simultaneously to the two-dimensional distributions of $\cos\phi$ vs. $\cos\theta^*_{D^*}$ in the signal and sideband fitting regions. The first region, a signal-dominated region, consists of events for which $\cos\theta_{h_W-\pi_s}$ is less than -0.95. The second, a background-dominated sideband fitting region, requires $\cos\theta_{h_W-\pi_s}$ to be between -0.8 and -0.9.

We fit the signal and sideband plots to a sum of five two-dimensional distributions in $\cos\phi$ and $\cos\theta^*_{D^*}$, one distribution for each category of selected events. Each component contributes events to both the signal and sideband fitting regions, though with differing ratios. The normalizations of the B shapes float in the fit; the normalization of the continuum background is set by the luminosity. The two plots (signal and sideband) are fit simultaneously and their joint likelihood maximized.

We have confidence in the Monte Carlo's ability to simulate 2-body B decays, and therefore trust the relative efficiencies and shapes the Monte Carlo gives for each individual 2-body decay mode. Therefore, the two-dimensional distributions for the following categories are taken from Monte Carlo:

- Signal
- Non-signal 2-body B^0 decays
- Non-signal 2-body $B^\pm$ decays

The distributions for the random combinatoric and continuum backgrounds may not be well-modeled by the Monte Carlo. We therefore generate the two-dimensional distributions for these categories from the data, as will be described.

Below we describe the relevant features of each category, its shape, and how we obtain the two-dimensional distributions of $\cos\phi$ and $\cos\theta^*_{D^*}$ for each category.

Signal

As described earlier, signal events exhibit striking features in the reconstruction angles, especially $\cos\theta^*_{D^*}$. The primary uncertainty concerning the physics modeling of these decays is the longitudinal polarization of $D^*\rho$ decays. We use the measurement of 87$\pm$5% from [74]. Signal decays populate the signal fitting almost exclusively; the ratio of signal events in the signal to sideband region is 1:0.07.

2-body B Decays

We define 22 B decay modes as "2-body" modes. These modes include four semileptonic modes which, though not 2-body, are thoroughly measured and whose kinematics are well-known. In table 4.20 we list each of these decay modes, its expected branching fraction, and the efficiency we find in the Monte Carlo.

2-body decays consist primarily of real hard π/ρ's, and consist of differing levels of real π_s depending on whether it is a B^0 2-body decay (many $D^{*\pm}$) or $B^{\pm}$ (many D^{*0}). The spectrum of π_s's not from D^* decays is similar to the inclusive track spectrum for low-momentum tracks. We expect the decays to be mostly back-to-back for π_s from real D^*'s and not otherwise. The distributions are displayed graphically in figures 4.9 and 4.10, and summarized in table 4.19.

These decays populate both the signal and sideband fitting regions. From the Monte Carlo, the ratio of signal events in the signal to sideband region is 1:0.80 for 2-body decays of neutral B's and 1:1.75 for 2-body decays of charged B's.

Table 4.19: Qualitative descriptions of the expected distributions for the three principle quantities that go into forming the angles $\cos\phi$, $\cos\theta^*_{D*}$, and $\cos\theta_B$, for 2-body decays.

Event Type	π_s	π/ρ	Angle
2-body B^0, real D^* decay	Peaked	Peaked	Back-to-back
2-body B^0, no D^* decay	Generic	Peaked	Flat
2-body $B^\pm$	Mostly Generic	Peaked	Flat

Combinatoric Backgrounds

All decays other than signal and 2-body decays are defined as random combinatoric backgrounds. We found no useful distinguishing features of these events in event dumps.

For combinatoric decays, we expect inclusive spectra for both the π_s and hard hadron, and flat angles between them. To test this expectation, we select well-measured tracks from events with a lepton of momentum greater than 1.4 GeV. We call these tracks the "inclusive" sample. We have overlaid the inclusive spectrum over the spectrum of tracks from random combinatoric backgrounds in figure 4.11. Random combinations of tracks were required to satisfy the same criteria as tracks selected for the fits. The normalizations are to the number of events in the plots. Agreement of the momentum spectra is satisfactory. The slope of the linear fit to the angle $\cos\theta_{h_W-\pi_s}$ is consistent with 0.

As the $B\overline{B}$ generic Monte Carlo is statistics-limited, we mock up a synthetic

Table 4.20: Yields in generic BB MC; 5.8 M events each of neutral and charged B. Recall efficiency also includes lepton ID, lepton angular cuts, and plep>1.4.

Decay Mode	Events accepted	# $B\overline{B}$ (plep>1.4)	BF Used (%)	Efficiency or 95% U.L.
B^0 Decays				
$D^{*-}\pi^+$	357	865 k	0.27	15.3 ± 1.2
$D^{*-}\rho^+$	542	865 k	0.70	9.0 ± 0.4
Signal	899			
$D^{*-}\pi^+\pi^0$	88	865 k	0.7	1.4±0.15
$D^{*-}a_1^+$	28	865 k	1.2	0.3±0.1
$D^{*-}\mu\nu$	41	865 k	5.4	0.1±0.02
$D^-\pi^+$	8	865 k	0.26	0.3±0.1
$D^-\rho^+$	17	865 k	0.73	0.3±0.1
$D^-\pi^+\pi^0$	0	865 k	0.10	0.2±0.2
$D^-a_1^+$	3	865 k	0.85	0.1±0.05
$D^-\mu\nu$	1	865 k	2.0	0.01±0.01
$D^{**}-\pi^+$	26	865 k	0.21	1.4±0.3
$D^{**}-\rho^+$	12	865 k	0.58	0.2±0.1
2-Body	224			
B^+ Decays		865 k		
$\overline{D^0}\pi^+$	11	865 k	0.47	0.3±0.1
$\overline{D^0}\rho^+$	16	865 k	0.98	0.2±0.05
$\overline{D^0}\pi^+\pi^0$	1	865 k	0.05	0.2±0.2
$\overline{D^0}\mu\nu$	1	865 k	2.0	0.005±0.005
$\overline{D^0}a_1^+$	1	865 k	0.94	0.01±0.01
$\overline{D}^{*0}\pi^+$	17	865 k	0.50	0.4±0.1
$\overline{D}^{*0}\rho^+$	35	865 k	1.49	0.3±0.1
$\overline{D}^{*0}a_1^+$	3	865 k	1.87	0.01±0.01
$\overline{D}^{*0}\pi^+\pi^0$	6	865 k	0.05	1.4±0.7
$\overline{D}^{*0}\mu\nu$	9	865 k	5.4	0.01±0.003
$D^{**}0\pi^+$	63	865 k	0.33	2.2±0.3
$D^{**}0\rho^+$	38	865 k	0.92	0.5±0.1
2-Body	201			
B^0 Comb.	44	865 k	–	
B^+ Comb.	20	865 k	–	
Total				

distribution of combinatoric decays from the inclusive track data. We fit the inclusive p_{π_s} and p_{π_f} spectra to a 5th and 1st order polynomial, respectively. We then use the polynomials to generate p_{π_s} and p_{π_f} spectra. We generate synthetic combinations by throwing the angle between them to have flat cosine from -0.8 to -1.0. We calculate the resulting values of $\cos\phi$,$\cos\theta_B$, and $\cos\theta^*_{D*}$. We use these synthetically generated angular distributions as the fitting shape for combinatoric backgrounds. As shown in the following figures 4.12, 4.13, and 4.14, when the synthetic shapes are generated from the inclusive Monte Carlo spectrum, they are consistent with the generic Monte Carlo combinatoric shape.

These decays populate primarily the sideband fitting regions. The ratio of signal events in the signal to sideband region is 1:3.2±0.4 in the generic sample, and 1:3.4±0.1 in the synthetic sample.

Continuum Events

We use the off-resonance data as the fitting shape for $e^+e^- \rightarrow q\overline{q}$ backgrounds. We fix the normalization to 2.03 ($0.99 \times \frac{\mathcal{L}_{ON}}{\mathcal{L}_{OFF}}$) in the fits. When calculating the reconstruction angles in off-resonance data, we shift the B mass downwards by 5.29-E_{beam}.

These decays populate primarily the sideband fitting regions. The ratio of signal events in the signal to sideband region is 1:4.6.

4.4.5 The Fit Procedure

Having determined the two-dimensional distributions of each of the categories in both the signal and sideband regions, we fit the data to find the best component normalizations. We fit both the signal and sideband two-dimensional plots simultaneously to obtain the normalizations by maximizing the likelihood. These plots, their binning, and the likelihood function are described in the following sections.

The Plots We Are Fitting

We fit the two two-dimensional distributions shown as lego plots in figure 4.15 to the sum of the six component shapes described in section 4.4.4. The x-axis of each plot is $\cos\phi$ and the y-axis $\cos\theta^*_{D*}$. The signal and sideband fitting region plots are shown in figure 4.15 as 20×20 bins. However, as described presently, we do not use this binning in the fit.

Binning

In order to reduce uncertainties associated with finite Monte Carlo samples, we do not fit all 400 bins of each region individually. Instead, we have variable-sized bins in different areas in the plot. This is in order to preserve sensitivity in areas of the plot with fine structure (for instance, in the peak near $\cos\phi = 1, \cos\theta_{D^*} = 1$) while achieving more Gaussian-like statistics in the less populated regions. Some regions are not kinematically allowed and we do not consider these regions in the fit. The variable bin sizes we select are equivalent to 4×4 bins, 2×2 bins, or just 1 bin on the plots shown in figure 4.15. The assignment of the variable bin sizes is shown in figure 4.16, overlaid on the data.

Fitting with Finite Monte Carlo Samples

In [83] an exact solution to the problem of Poisson likelihood fitting with finite Monte Carlo statistics is obtained. In summary, it amounts to replacing the usual Poisson likelihood probability given in 4.12 with that given in 4.13, where there are n bins i and m Monte Carlo shapes j; d_i is the number of data events in bin i, a_{ij} the number of Monte Carlo events from component j in bin i, and p_j the Monte Carlo scaling parameter which we wish to find by maximizing the likelihood. The method notes that for finite Monte Carlo samples, $a_{ij} \neq A_{ij}$, where A_{ij} is the true expectation for the number of events in bin i from component j. The method consists of solving for the A_{ij}, then maximizing a joint probability to observe a_{ij} events in the Monte Carlo and d_i in the data.

$$P_i(d_i, a_{i1}..a_{im}) = \frac{(\sum_j p_j a_{ij})^{d_i} e^{-\sum_j p_j a_{ij}}}{d_i!} \tag{4.12}$$

$$P_i(d_i, a_{i1}..a_{im}) = \frac{(\sum_j p_j A_{ij})^{d_i} e^{-\sum_j p_j A_{ij}}}{d_i!} \prod_j \frac{A_{ij}^{a_{ij}} e^{-A_{ij}}}{a_{ij}!} \tag{4.13}$$

However, statistical uncertainty due to finite Monte Carlo statistics is not a significant contributor to the statistical uncertainty in the analysis. The statistical uncertainty in the analysis is dominated by the statistical uncertainty on the yields of mixed events. The fit uncertainty is not even a large uncertainty in the systematic determination of the hadronic mistag rate, which uncertainty is dominated by the assigned combinatoric mistag rate.

4.5 Fit Results

The fit returns a confidence level (defined as in PDG for likelihood fits) of 129.7 for 151 degrees of freedom. This corresponds to a confidence level of 91.4%. For the individual regions, the fit gives 75.1 for 85-4 d.o.f. (66% C.L.) and 54.6 for 70-4 d.o.f.(84% C.L.). The correlation matrix for the fit is shown below. Parameters 1-4 are Signal, B^0 2-body, $B^{\pm}$ 2-body, and combinatoric.

```
PARAMETER  CORRELATION COEFFICIENTS
      NO.  GLOBAL      1      2      3      4
        1  0.59062  1.000-0.493 0.077 0.105
        2  0.88154 -0.493 1.000-0.577 0.068
        3  0.94122  0.077-0.577 1.000-0.771
        4  0.90484  0.105 0.068-0.771 1.000
```

In table 4.21 we show the fit results for the composition of the selected events. In figure 4.20 we show the residuals of the fit in both the signal and sideband regions. In figure 4.17 we show the projections of the signal-region fit in the two decay-chain angles. In figure 4.18 we show the projections of the sideband-region fit in the two decay-chain angles. In figure 4.19 we show the projections for those events that are selected for the measurement of mixing after imposition of the final $\cos\theta^*_{D^*}$ requirement. We also show the prediction for the $\cos\theta_B$ distribution,

Table 4.21: The composition of the sample selected for measuring mixing, as determined by the simultaneous fit to signal and sideband fitting regions.

Event Type	Fit	Fraction
Signal	1244±52	66±3%
2-body B^0	267±60	14±2%
2-body $B^{\pm}$	174±45	9±2%
Random Combinatoric	184±21	10±1%
Sum	1881±63	100±3%
Data B	1876	100 %

which is uninspiring in agreement. We discuss this disagreement in the section on systematic uncertainties.

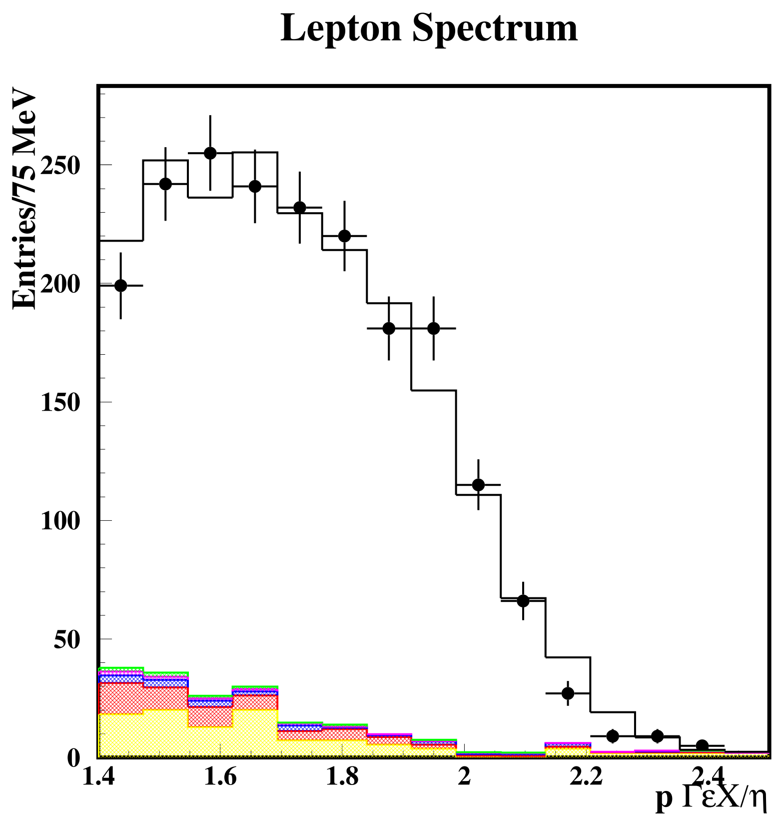


Figure 4.6: The components of the Monte Carlo-simulated lepton spectrum (histograms) and leptons selected in data (points). The unshaded histogram shows primary leptons from semileptonic B decay. The largest background components are continuum leptons (lightly shaded), cascade leptons (more darkly shaded); small backgrounds also shown are leptons from ψ decays, fakes, and non-primary leptons from all other sources.

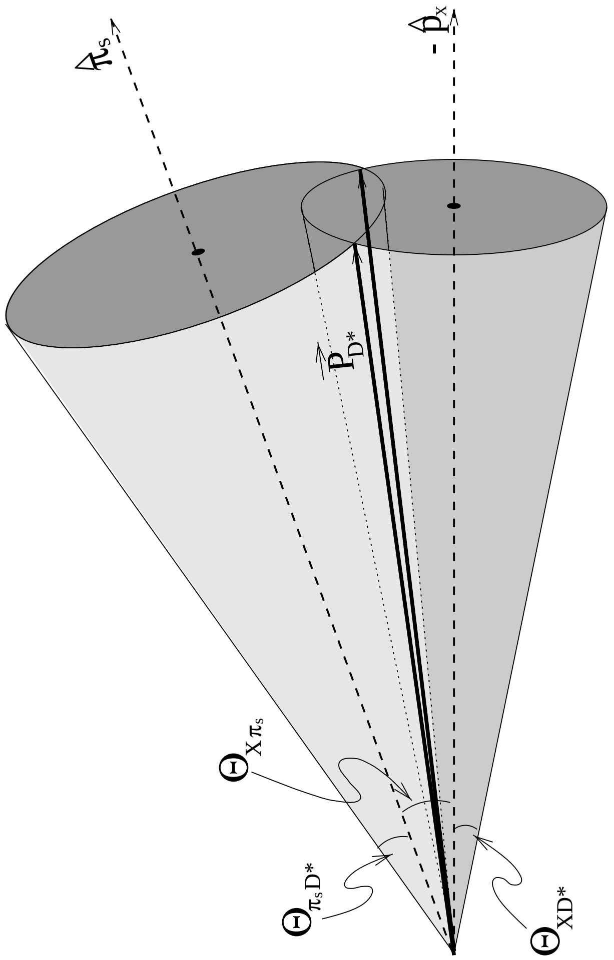


Figure 4.7: The momentum cones used in the partial reconstruction tag. From [6]. The radius of the upper cone is $|\vec{p}_D|$, and of the lower is $|\vec{p}_B|$.

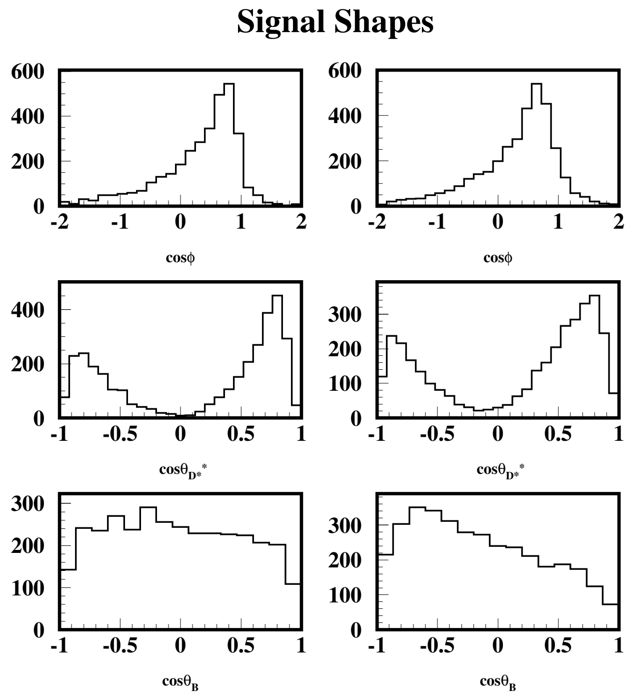


Figure 4.8: Distributions of the angle cosines for signal modes in the analysis. The angles are described in the text. The $D^*\pi$ mode is on the left, $D^*\rho$ on the right.

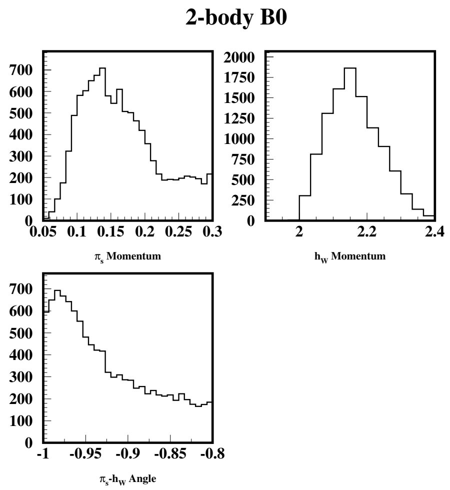


Figure 4.9: The π_s, h, and angle distributions in 2-body B^0 decays.

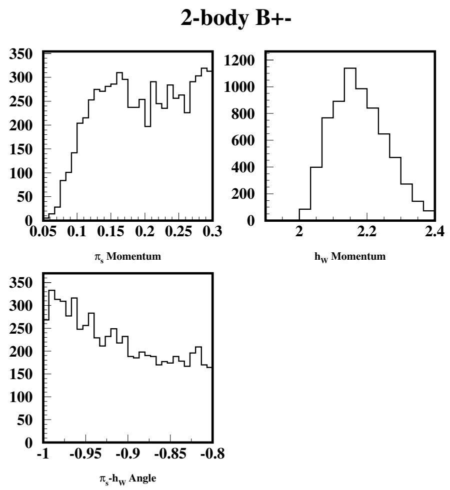


Figure 4.10: The π_s, h, and angle distributions in 2-body B^+ decays.

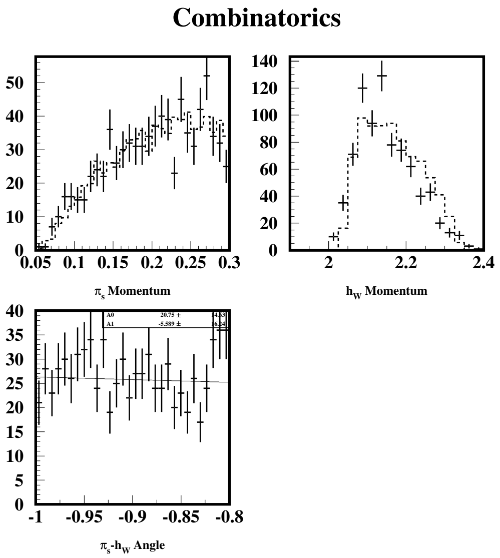


Figure 4.11: The π_s, h, and angle distributions for combinatoric background Monte Carlo events selected by the analysis for the fits. Overlaid are inclusive track spectra from the data.

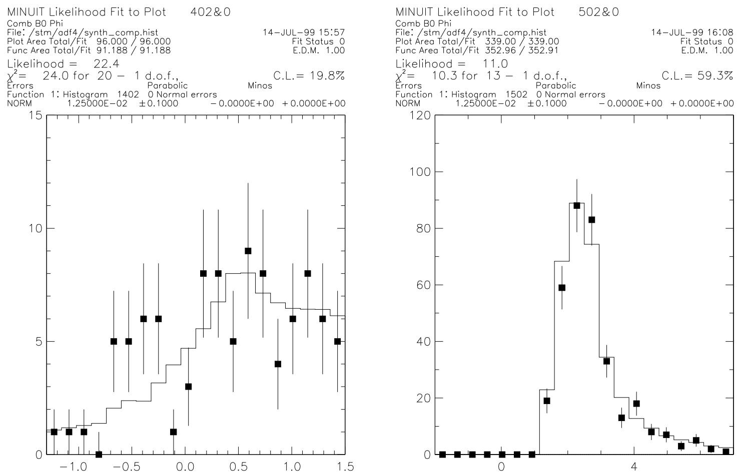


Figure 4.12: Distribution of $\cos\phi$ for combinatoric backgrounds in generic Monte Carlo (points) and in synthetic simulation (histogram). On the left is the distribution in the signal fitting region; on the right the distribution in the sideband fitting region.

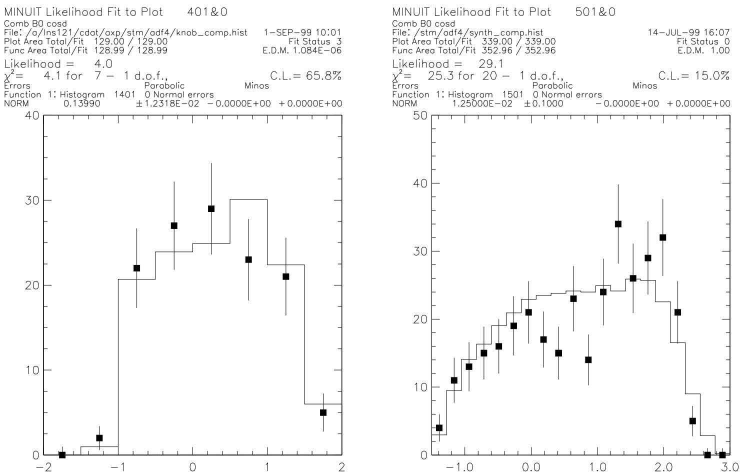


Figure 4.13: Distribution of $\cos\theta_{D*}$ for combinatoric backgrounds in generic Monte Carlo (points) and in synthetic simulation (histogram). On the left is the distribution in the signal fitting region; on the right the distribution in the sideband fitting region.

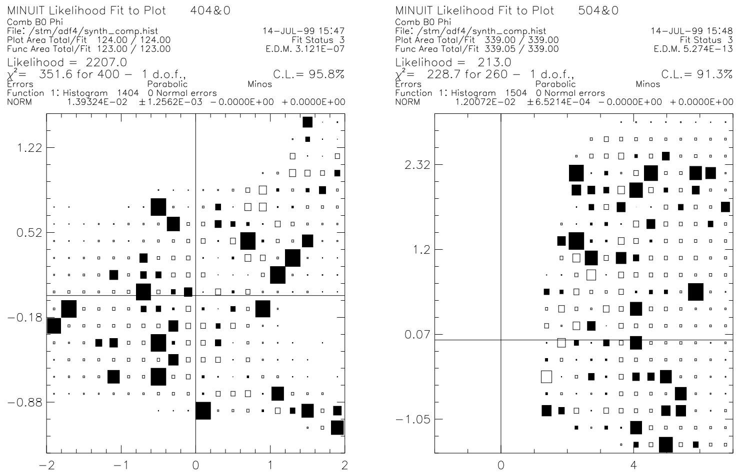


Figure 4.14: Residuals of comparison (equal areas) of 2-D distribution of $\cos\phi$ vs. $\cos\theta_{D*}$ for combinatoric backgrounds in generic Monte Carlo and synthetic simulation generated from the Monte Carlo inclusive spectrum. On the left is the signal region, and the sideband is on the right.

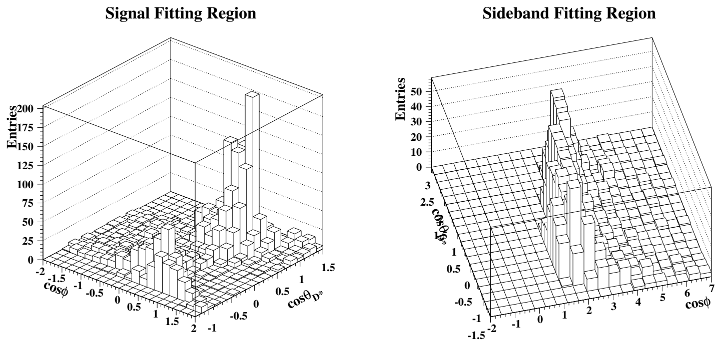


Figure 4.15: Lego plots of the data distribution of $\cos\phi$ vs. $\cos\theta_{D^*}$, in the signal (left) and sideband (right) fitting regions.

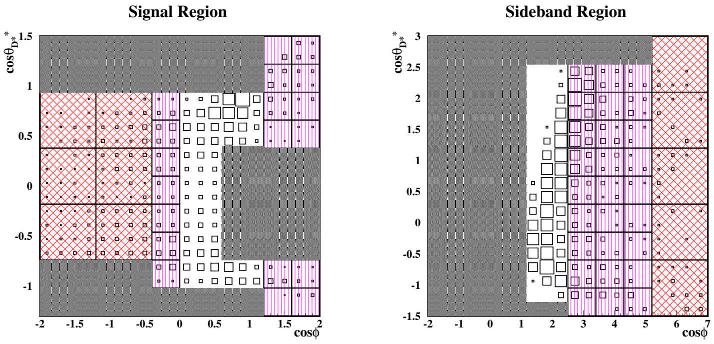


Figure 4.16: Implementation of variable-sized bins in $\cos\phi$ vs. $\cos\theta_{D^*}$ for signal (left) and sideband (right) fitting regions, overlaid on the on-resonance data. The fine 20×20 binning is equivalent to that in figure 4.15. Dark shaded regions are not considered in the fit. Cross-hatched regions are equivalent to 4×4 bin regions in figure 4.15, vertically hatched to 2×2, and unshaded bins are fit individually.

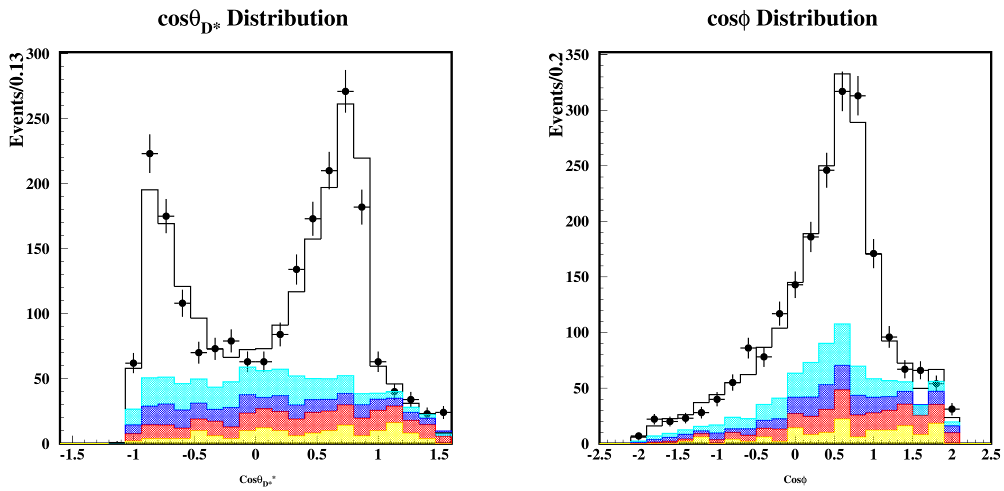


Figure 4.17: Projections of the fit to the signal region, onto the $\cos\theta_{D^*}$(upper left) and $\cos\phi$ axes (upper right). In each plot, the points are the data and the histograms are the best-fit shapes from the Monte Carlo. The white histogram is the signal component, the light blue (uppermost shaded) is the 2-body B^0 component, the dark blue (darkest shaded component) is 2-body $B^{\pm}$, the red (lighter shaded) is the combinatoric background, and the yellow (lowermost shaded component) is the off-resonance data.

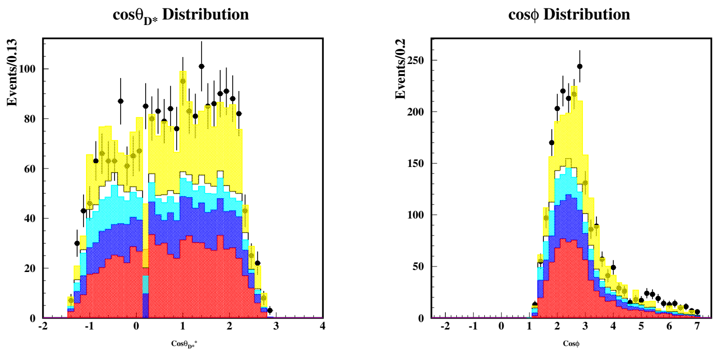


Figure 4.18: Projections of the fit to the sideband region. In each plot, the points are the data and the histograms are the best-fit shapes from the Monte Carlo. The white histogram (of which there is very little) is the signal component, the yellow (uppermost shaded component) is the off-resonance data, the light blue (second shaded component) is the 2-body B^0 component, the dark blue (darkest shaded component) is 2-body $B^{\pm}$, and the red (lowermost shaded component) is the combinatoric background. The projections are on the $\cos\theta_{D^*}$ (upper left) and $\cos\phi$ axes (upper right).

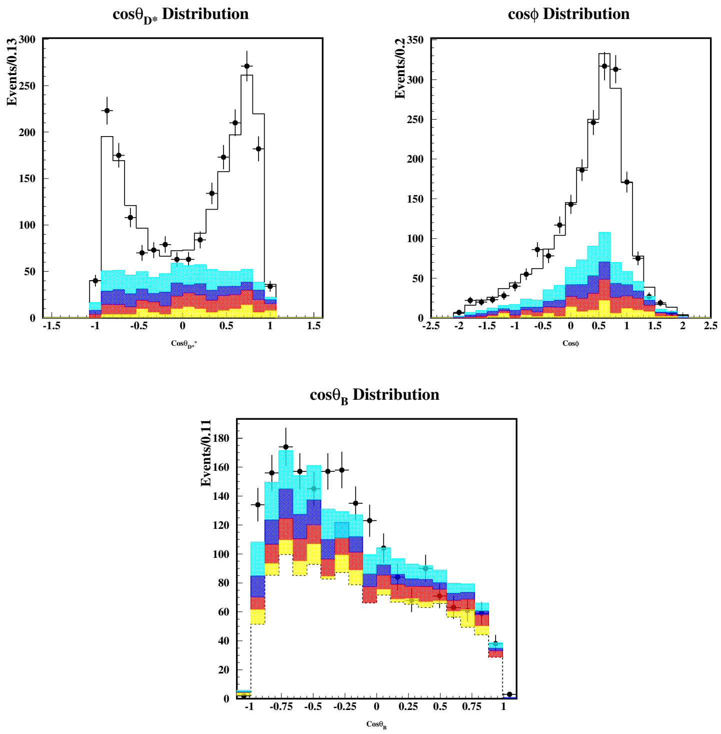


Figure 4.19: Projection of the fit, onto the $\cos\theta_{D^*}$(left) and $\cos\phi$ axes (right), for events that are used to measure the value of mixing. Below: fit prediction for $\cos\theta_B$.

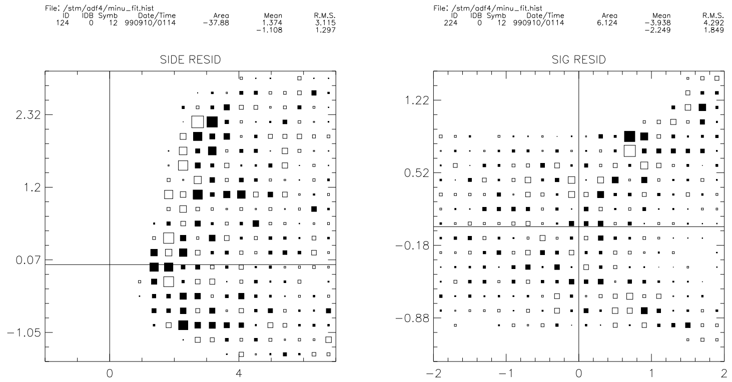


Figure 4.20: `mn_fit` two-dimensional “speckle plots” of the residuals. The x axis is $\cos\phi$, the y axis $\cos\theta_{D^*}$. The box sizes correspond to the size of the residual value; in black boxes the fit exceeds the data, in white boxes the data exceed the fit.

Table 4.22: The determination of the charged B fraction. The total random combinatoric background is provided as a reminder. The uncertainties shown are solely those that arise from the statistical uncertainties of the fit.

Event Type	Fit	Fraction
Events	1881	100%
Combinatoric B^0 and B^+	184±21	10±1%
Comb. $B^{\pm}$	74±8	4±1%
2-body $B^{\pm}$	173±45	9±2%
Charged B Sum	247±47	13±3%

4.5.1 The B^+ Fraction

To determine the charged B fraction, we add the number of 2-body B^+ decays to the number of combinatoric B^+ decays, divided by the total number of B events in the fit. As per the Monte Carlo, we assign 40% of the combinatoric background as charged B. As shown in table 4.22, we obtain a charged B fraction of 13±3%, where the uncertainties are solely those that arise from the statistical uncertainties of the fit.

4.5.2 The Hadronic Mistag Rate

To obtain the hadronic mistag rate, we determine the number of mistags that arise from each of the event categories. We sum the mistags and divide by the total number of events. As per our measurement (described in section 4.6.2) we

Table 4.23: Determination of the mistag fraction from the fit. The errors shown are *statistical only*, whether from the fit to the data (first column) of the size of the Monte Carlo sample (second column). The mistag rate we determine is 3.1%.

Event Type	Fit	Mistag Fraction	Mistagged Events
Signal	1244±52	0	0
2-body B^0	267±60	2±0.3%	6±1
2-body $B^\pm$	173±45	5±0.7%	9±1
Comb. B^0	111±13	21±2%	25±3
Comb. $B^\pm$	73±8	21±2%	16±2
Sum	1870	**3.1±0.2%**	56±4

assign a mistag rate of 21% for the combinatoric backgrounds, and use this rate to determine the mistag rate shown in table 4.23. We expect 56 events to be mistagged in the data sample; with 1870 B events in the sample this results in a mistag rate of 3.1±0.2%, where the uncertainty is only the statistical uncertainty propagated from the fit.

4.5.3 Mixing

In the table 4.24 we show the recorded counts for each lepton species (since we make separate lepton mistag corrections for them).

Table 4.24: Data event counts. χ_{raw} is calculated for the convenience of the reader. Because e and μ mistag corrections are not the same the two values of χ_{raw} need not be the same.

Mode	$h^+\ell^+/h^-\ell^-$	$h^\pm\ell^\mp$	Raw χ
e ON	243	868	
e OFF	16	15	
e Subtracted	210$\pm$17	838 $\pm$ 30	20.0 $\pm$1.5 %
μ ON	215	656	
μ OFF	12	15	
μ Subtracted	191$\pm$16	626 $\pm$ 26	23.4 $\pm$1.7 %

Using the Mistag Rate

We use a set of 2×2 matrices in order to unfold the effects of mistagging on the result. We reconstruct events in the vector $\begin{pmatrix} \text{Reconstructed Mixed} \\ \text{Reconstructed Unmixed} \end{pmatrix}$ and would like to recover the vector $\begin{pmatrix} \text{True Mixed} \\ \text{True Unmixed} \end{pmatrix}$ accounting mainly for the effects of lepton mistagging, hadronic mistagging, and B^+ contamination.

To do this accounting we use the results of the Monte Carlo studies, lepton fake rates, etc. to construct the 2×2 matrices which take true → reconstructed, that is

$$\mathbf{M} \begin{pmatrix} \text{True Mixed} \\ \text{True Unmixed} \end{pmatrix} = \begin{pmatrix} \text{Reconstructed Mixed} \\ \text{Reconstructed Unmixed} \end{pmatrix} \tag{4.14}$$

These matrices $\mathbf{M}$ depend only on the various fake rates, and do not depend on the actual value of χ_d assumed in the Monte Carlo. If we allow non-symmetric matrices, we can account for any possible differences in tagging purity between mixed and unmixed events. We find no evidence for differing mistag rates in the generic Monte Carlo, and we therefore use symmetric matrices. The matrices also contain individual pieces of the errors in the values of the elements themselves, simplifying and organizing the analysis of the systematic uncertainties.

We invert these matrices and apply them to the data to estimate the true number of mixed and unmixed events:

$$\mathbf{M}^{-1} \begin{pmatrix} \text{Reconstructed Mixed} \\ \text{Reconstructed Unmixed} \end{pmatrix} = \begin{pmatrix} \text{True Mixed} \\ \text{True Unmixed} \end{pmatrix} \tag{4.15}$$

We use three matrices in sequence: a lepton mistagging matrix, a hadron mistagging matrix, and a charged-B contamination matrix. Which lepton mistag matrix we use depends on the lepton species e or μ. The lepton and hadron matrices commute, but they do not commute with the charged B matrix; this means the order of matrix application is important. For either lepton species, the estimate of the true number of events in the mode is:

$$\mathbf{F_{B^+}}\mathbf{M_{had}}^{-1}\mathbf{M_{lep}}^{-1}\begin{pmatrix}\text{Reconstructed Mixed}\\ \text{Reconstructed Unmixed}\end{pmatrix} = \begin{pmatrix}\text{True Mixed}\\ \text{True Unmixed}\end{pmatrix} \tag{4.16}$$

where $\mathbf{F_{B^+}}$ is the matrix

$$\begin{pmatrix}1 & 0\\ -f_{B^+} & 1-f_{B^+}\end{pmatrix} \tag{4.17}$$

and $\mathbf{M_{had}}$ and $\mathbf{M_{lep}}$ are the matrices describing hadronic and leptonic mistagging, respectively. The mistag rate used for $\mathbf{M_{had}}$ is given by the result of table 4.23.

$$\mathbf{M_{had}} = \begin{pmatrix}0.969\pm0.016 & 0.031\pm0.016\\ 0.031\pm0.016 & 0.969\pm0.016\end{pmatrix}$$

$$\mathbf{M_e} = \begin{pmatrix}0.964\pm0.003 & 0.036\pm0.003\\ 0.036\pm0.003 & 0.964\pm0.003\end{pmatrix} \mathbf{M}_\mu = \begin{pmatrix}0.962\pm0.005 & 0.038\pm0.005\\ 0.038\pm0.005 & 0.962\pm0.005\end{pmatrix}$$

$$\mathbf{F_{B^+}} = \begin{pmatrix}1 & 0\\ 0.13\pm0.03 & 0.87\pm0.03\end{pmatrix} \tag{4.18}$$

The method assumes that hadronic and leptonic mistags are uncorrelated. We address this assumption in the following section.

Correlated Mistags

If either the hadron or lepton incorrectly tags the flavor of a B, is the other tag more likely to misreconstruct the other B flavor? We find that a hadron mistag does not necessarily cause the lepton mistag rate to increase in generic Monte Carlo:

- If hadronic tag side is true signal, fraction of non-primary leptons on other side: 4.9 ± 0.5 %
- If hadronic tag side is not true signal, fraction of non-primary leptons on other side: 3.7 ± 0.8%

Similarly, a lepton mistag does not cause a significant increase in the hadron mistag rate in generic Monte Carlo:

- If lepton is truly primary, fraction of hadronic tags which incorrectly tag other side: 2.5 ± 0.3 %
- If lepton is truly non-primary[4], fraction of hadronic tags which incorrectly tag other side: 5.7 ± 3.2%

We evaluate the double mistag rate in Monte Carlo as follows. We require a lepton from a non-primary B decay which incorrectly tags the flavor of that B which is its ultimate ancestor. We then examine whether the hadronic reconstruction incorrectly tags the decay flavor of the *other* B in the event. We find no events

[4] Including upper-vertex secondaries.

in which this happens in the generic Monte Carlo. In the 3 events which contained both a secondary lepton and a mistagging hadronic decay, all three leptons originated from upper vertex decay, correctly tagging the flavor of their ancestors. We therefore find the fraction of "doubly mistagged" events is <0.13%. If single mistags are uncorrelated the mistag rate should be 0.12±0.02%. Since the rate of double mistags is around 10^{-3} and there is no evidence for high correlation, we neglect corrections for correlated mistags.

Result: Count Corrections Using Mistag Rates

In table 4.25 we show the effects of the sequential corrections, to obtain the final results. With statistical uncertainties *only*, we find $\chi_d = 19.9 \pm 1.3\%$. As we discuss in the $\cos\theta_B$ consistency check, we have reason to believe the beam energy is off by 400 keV, and we shift the result by 0.2% to account for this, giving us a central value,

$$\chi_d = 19.7 \pm 1.3\% \tag{4.19}$$

4.5.4 A_{CP}

We prefer to measure the CP asymmetry as an asymmetry in χ and not simply an asymmetry in the number of like-signs. We define

$$\chi_d(\ell^\pm) = \frac{h^\pm \ell^\pm}{h^\mp \ell^\pm + h^\pm \ell^\pm} \tag{4.20}$$

Table 4.25: Data event counts, sequentially subjected to the corrections of the analysis.

Mode	$h^+\ell^+/h^-\ell^-$	$h^\pm\ell^\mp$	Raw χ
e subtracted and corrected	186±18	861±31	17.7±1.7 %
μ subtracted and corrected	172±16	643±27	21.1±1.8 %
Total, subtracted and corrected	359±24	1505±42	19.2±1.2 %
Corrected for hadron mistags	322±25	1538±43	17.3±1.2 %
Corrected for $B^\pm$	322±25	1296±38	19.9±1.3 %
Beam Energy Correction			19.7±1.3 %

We have studied χ more thoroughly, and are therefore more confident in canceling the systematics. This method also cancels detector CP asymmetries in the lepton identification. In principle there may still be asymmetries left due to CP violating detector effects on the charged hadronic tracks. However, these effects have been studied throughly in many places [75–77, 84] and found generally to be less than 1%, even at low momentum [75, 76, 84]. We therefore find A_{CP} as follows:

$$A_{CP} = \frac{\chi_d(\ell^+)-\chi_d(\ell^-)}{\chi_d(\ell^+)+\chi_d(\ell^-)} \tag{4.21}$$

We rerun the fit twice, once for each lepton sign. We find all components in the fits are consistent between positive and negative leptons. The results for χ_d from each fit are shown in table 4.27. Using *statistical* uncertainties only, we find

Table 4.26: Data event counts, separated by lepton charge

Mode	Mixed	Unmixed	Raw χ
e^+ ON	126	447	
e^+ OFF	8	7	
e^+ Subtracted	111±12	434 ± 22	20.4 ±1.5 %
μ^+ ON	106	331	
μ^+ OFF	4	8	
μ^+ Subtracted	100±11	316 ± 19	24.0 ±1.7 %
e^- ON	117	421	
e^- OFF	8	8	
e^- Subtracted	102±12	406 ± 21	20.1 ±1.5 %
μ^- ON	109	325	
μ^- OFF	8	7	
μ^- Subtracted	94±12	312 ± 18	23.1 ±1.7 %

Table 4.27: Measured quantities for the A_{CP} limit.

Quantity	Measurement	Statistical Uncertainty
$\chi(\ell^+)$	19.6 %	1.8%
$\chi(\ell^-)$	20.3 %	1.9%

$$A_{CP} = 0.017 \pm 0.070 \tag{4.22}$$

Since for small A_{CP}, $4\Re(\epsilon_B) = A_{CP}$, this value of A_{CP} corresponds to

$$\Re(\epsilon_B) = 0.4 \pm 1.8\% \tag{4.23}$$

The limit on $|\Re(\epsilon_B)|$ at the 95% C.L. is $|\Re(\epsilon_B)| < 3.3\%$.

4.6 Systematic Uncertainties

The counting procedure is subject to a systematic uncertainty in the relative efficiency for mixed and unmixed events to be selected by the analysis, and in possible differential mistagging rates for mixed and unmixed events. However, the most considerable systematic uncertainties are associated with the corrections we apply. These are most usefully categorized separately as lepton and hadron uncertainties.

4.6.1 Lepton Uncertainties

Most of the uncertainties associated with the lepton mistagging have already been discussed. In addition, we vary the upper vertex branching fraction by 50% of itself. In table 4.28 we summarize the uncertainties. The total lepton uncertainty includes 100% correlations between electron and muon uncertainties in the secondary fraction, hadron fake measurement statistical uncertainties, J/ψ decays, and upper vertex decays.

Table 4.28: A list of the quantities that contribute systematic uncertainty to the lepton tagging portion of the analysis. The total lepton uncertainty accounts for sources of error that are correlated between e and μ.

Source of Error	Assigned Source Uncertainty	$\Delta\chi_d$	Reason for assigned uncertainties
e **Tag**			
Secondary Fraction f_c	$\pm$ 10%	$\pm$0.1%	PDG Form Factors, B.F. Measurements, primary spectrum
Fake Rate	$\pm$50%	$\pm$0.04%	Fake rate measurements
Hadron Fakes	$\pm$25%	$\pm$0.02%	Data statistics
J/ψ Decays	$\pm$3 events	$\pm$0.07%	MC statistics
γ Conversions	$\pm$ 1 event	$\pm$0.03%	MC Statistics
π^0 Dalitz Decays	$\pm$ 1 event	$\pm$0.03%	MC Statistics
δ Rays	<2.3 events	<0.08%	MC Statistics
Upper Vertex Decays	$\pm$50%	$\pm$0.05	
μ **Tag**			
Secondary Fraction f_c	$\pm$ 10%	$\pm$0.1%	PDG Form Factors, B.F. measurements, primary spectrum
Fake Rate	$\pm$35%	$\pm$0.08%	Fake rate measurements
Hadron Fakes	$\pm$25%	$\pm$0.06%	Data statistics
J/ψ Decays	$\pm$3 events	$\pm$0.07%	MC statistics
In-flight decays	<2.3 events	<0.03%	MC Statistics
Upper Vertex Decays	$\pm$50%	$\pm$0.05%	
Total Uncertainty		$\pm$0.29%	

4.6.2 Hadron Uncertainties

The two most important uncertainties arise in estimation of the charged B fraction and of the mistag fraction. However, there are also uncertainties associated with the fitting distributions of each event category; such uncertainties propagate into the final result because they can change the results of the fit, changing in turn the charged B and mistag fractions.

Charged B Fraction of the background

The charged B fraction uncertainty returned by the fit is pretty small–2.5%–because the negative correlation coefficient between the combinatoric and 2-body $B^{\pm}$ keeps the charged B fraction sum pretty constant. One procedure would be to use this statistical fit uncertainty as the systematic uncertainty on the charged B fraction. This is naive because we don't really know (as a question of physics) what fraction of combinatoric B events are from charged B's. Though Monte Carlo indicates that 40% of them are, there is no particularly compelling reason to believe it. Such an uncertainty is not reflected in the fit. Similarly naive is to ignore the fit altogether and take the charged B fraction from the Monte Carlo. This procedure is naive not only because it ignores the physics uncertainty in the fraction that is from charged B's in the combinatorics, but because it ignores $\frac{f_{\pm}}{f_0}$; the Monte Carlo is generated with $\frac{f_{\pm}}{f_0} = 1.000$.

From a physics point of view, we don't really know anything about these events. So the only error we can fully justify is the assumption that combinatorics could be anything between 0% and 100% charged B–giving $\frac{1}{\sqrt{12}} = 29\%$ uncertainty on

Table 4.29: Summary table of possible charged B fraction systematic uncertainties. We assign the top line in this table as the charged B systematic uncertainty.

Systematic Uncertainty	$\frac{\Delta f_{B^{\pm}}}{f_{B^{\pm}}}$	$\Delta\chi_d$	Set by
$B^{\pm}$ Fraction	$\pm 29\%$	$\pm 0.74\%$	Fit results $\oplus$ combinatoric
			$B^{\pm}$ fraction flat 0-1
$B^{\pm}$ Fraction	$\pm 19\%$	$\pm 0.48\%$	Fit results, statistical only
$B^{\pm}$ Fraction	$\pm 3.3\%$	$\pm 0.10\%$	MC blind faith $\oplus$ MC statistics

the combinatoric charged B fraction. This is the systematic uncertainty we assign. A spectrum of other possible uncertainty assignments is shown in table 4.29.

It is worth noting that because the normalization of the charged B components in taken from data, there is *no* assumption concerning $\frac{f_{\pm}}{f_0}$ in the analysis, and therefore no dependence on its value.

Measurement of combinatoric mistag rates using $D^*\ell\nu$

Rather than believe the Monte Carlo prediction of the mistag rate of random combinatoric backgrounds, we would like to establish the mistag rate using data. We do this by measuring the charge-flavor correlation of high-momentum tracks to their parent B mesons. We then use the less strong assumption (and assign substantial uncertainty) that the Monte Carlo can correctly relate the single-track correlation to the correlation of a random h_W-π_s combination.

To measure the single-track charge-flavor correlation we use Mark Dickson's $D^*\ell\nu$ reconstruction code on Karl Ecklund's Recompress skim. The fully-reconstructed $D^*\ell\nu$ tells us (up to mixing) the flavor of the other B in the event, from which originate the other tracks in the event. Having reconstructed $D^*\ell\nu$, we look for additional tracks in the event which are over 1.3 GeV, and for which there is a soft track of opposite sign. We then compare the hard track sign with the known parent B flavor. We must then account for mixing and for the fact that many high-momentum tracks originate from 2-body B decays rather than random combinations.

In the data, we find 97 tracks above 1.3 GeV in $D^*\ell\nu$ decays. 39 have the same sign as the lepton, 58 have the opposite. We must correct the counts for mixing, yielding 33.2$\pm$2.0$\pm$1.6 of the same sign, and 63.8$\pm$2$\pm$1.6 of the opposite. The second uncertainty comes from varying the value of χ_d by 18% (3σ).

We subtract the absolute number of right and wrong sign tracks expected from 2-body and signal decays. The Monte Carlo expectation is 0.19$\pm$0.02 same-sign and 0.39$\pm$0.02 opposite-sign decays per $D^*\ell\nu$ event. After subtracting the two-body decays, 14.8$\pm$3.2 and 26$\pm$3.2 events of same and opposite sign, respectively. We find the mistag rate for high-momentum tracks is $\frac{14.8}{14.8+26}$ =36.3$\pm$12.1$\pm$7.8%. This rate should be compared to the direct Monte Carlo expectation for single high-momentum tracks *not originating from a 2-body decay* of 44.4$\pm$2.6%. We find that this measurement is consistent with the Monte Carlo, with a difference of 8.1$\pm$14.6%.

According to the Monte Carlo, while the flavor mistag rate for tracks above 1.3

GeV is 44.4±2.6%, the mistag rate for combinatorics is 28.4±4.6%. Thus we find in the Monte Carlo that the relation between the single-tag mistag rate and the combination mistag rate is -16±5.3%.

Since we found that the Monte Carlo predicts that the combinatoric mistag rate is different from the single-track mistag rate, we subtract 16±5.3% from the single-track mistag rate to obtain a combinatoric mistag rate of 20±15.5%. We compare this rate to the Monte Carlo mistag rate of 28.4±4.5%. We use the data-based mistag rate in the fits.

Because we cannot pinpoint precisely the origin of the 16% shift for using lower momentum tracks, we assign two additional systematic uncertainties: one set arbitrarily to $\frac{1}{\sqrt{12}}$ of the 16% shift, i.e. 5%. We also add in quadrature the difference between the data and Monte Carlo mistag rates, which is 8.4%. We therefore assign a total systematic uncertainty of 9.4% to the subtraction.

Since a significant portion of the probability is in the unphysical region of negative mistag rates, we evaluate the uncertainty by evaluating the variance of only the portion of Gaussian probability that falls in the physical region from 0 to 1. We find an uncertainty of 13.0% on the mistag rate, corresponding to a systematic uncertainty of 0.9% on χ_d.

In table 4.30, I show a series of possible choices for the assignment of a systematic uncertainty for the combinatoric mistag rate. We use ±13% as our best estimate of the uncertainty.

Table 4.30: Summary table of possible hadronic mistag systematics. We use line 4 in the final evaluation.

Systematic Uncertainty	ΔM	$\Delta\chi_d$	Set by
Combinatoric Mistag Rate	28%	1.8%	Flat from 0 to 1
Combinatoric Mistag Rate	16%	1.1%	$D^*\ell\nu$ uncertainty
			(including unphysical region)
Combinatoric Mistag Rate	14%	1.0%	Flat from 0 to 0.5
Combinatoric Mistag Rate	13%	0.9%	Physical Gaussian $\oplus$ 9% systematic
Combinatoric Mistag Rate	4.5%	0.3%	Blind MC Faith + MC Statistics

Combinatoric shape systematic

We generate a synthetic combinatoric shape as described above. We perform two consistency checks. First, we simply look at the shift in χ_d which results when we no longer fit using the generic combinatoric MC shape and instead use the synthetic shape. The shift is 0.1%. Second, we verify that the synthetic Monte Carlo correctly models the ratio of combinatoric events in the signal and sideband regions. We find that the ratio of signal:sideband is 1:3.8±0.4 in the generic combinatoric Monte Carlo, and 1:4.1±0.2 in the synthetic sample.

To evaluate a systematic uncertainty, we add in quadrature three effects:

- The shift in the result when the shape of the high-momentum inclusive spectrum is distorted by a linear term

- The shift in the result when the shape of the low-momentum inclusive spectrum is distorted by a linear term

- The shift in the result when the shape of the low-momentum inclusive spectrum is distorted by a term for excess D^*'s.

For each of the effects, we intentionally distort the shapes and watch the change in χ_d. We must vary each shape by 1σ and observe the effect on the result; however, it is not immediately obvious what 1σ means in a shape. We set the size of the variation by the following means for all three cases.

We must quantify how many "σ" our shape distortions correspond to. We generate synthetic shapes from fits to the inclusive Monte Carlo track spectrum (as opposed to the inclusive *data* track spectrum). We distort these shapes in the same way that we distort the data spectrum. We then treat the combinatoric events from the generic Monte Carlo as "data", and perform the same 2-D fit to these data as to the real mixing data. We perform a one-parameter fit of the normalization to the combinatoric "data" using the Monte Carlo-based synthetic shapes. We then observe the change in the log likelihood from the nominal fit and use that to set the size of σ.

To assign the systematic uncertainty, we then observe the shift in χ_d due to the change. Since we generally vary the shape by more than 1σ, we divide the shift by the number of σ the shape change corresponded to.

First, we vary the linear terms of the parameterizations that describe the fits. We chose the linear term because it is the important contributor to both shapes.

We keep the other terms fixed as we vary the linear term. The resulting spectra are shown in figure 4.21. For the hard pion linear distortion the log likelihood shifts 3.8σ; the soft pion by 1.9 σ. The result for χ_d shifts by -0.28 and -0.28 %, respectively.

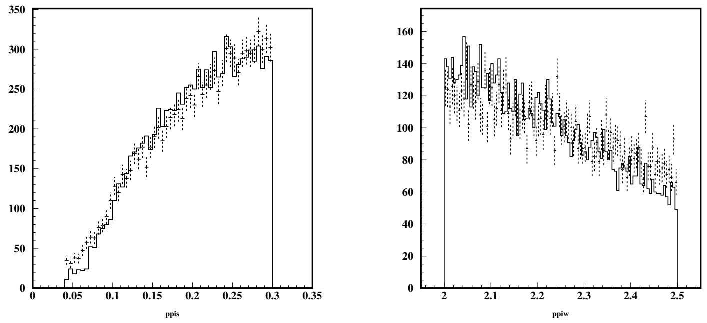


Figure 4.21: Nominal (histogram) and linearly shifted (points) shapes for the π_s (right) and h_W (left) momentum spectrum.

Next we vary the shape to account for differences between the combinatoric and inclusive samples in the fraction of soft tracks which come from D^*'s. 15.5±1.7% of combinatoric events in both the signal and sideband combinatoric MC are from D*'s, where 9.1% of low-momentum tracks in the inclusive Monte Carlo are from D^* decays. Below I show the plots for the sum of signal and sideband regions in ϕ, $\cos\theta_D$, and momentum, for the two cases when a pion does or does not come from a D^*.

To address the D^* fraction, we add a Gaussian lump of mean 140 MeV and width 42 MeV (to crudely simulate the D^* soft π shape) to the synthetic Monte

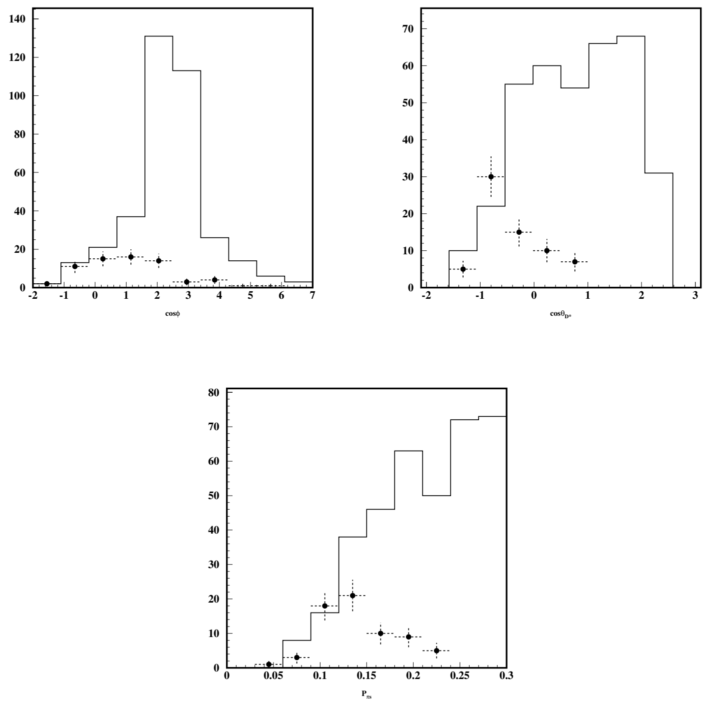


Figure 4.22: Shapes in $\cos\phi$, $\cos\theta_{D*}$, and p_π when the soft pi does (points) and does not (histogram) come from a D^*.

Carlo. All other terms remain fixed. We choose a normalization to give an additional 9% of events in this region, corresponding to 1.5 $\times$ the difference between 9 and 15%. We show the resulting momentum distribution on the right in figure 4.23 We find the combinatoric Monte Carlo *does not* prefer this shape, but rather prefers the inclusive shape, at the 2.9σ level. We find that χ_d shifts downwards by 0.25% when fitting with the D^*-modified synthetic shape. The resulting systematic uncertainty is $\frac{0.25}{3} = 0.08\%$.

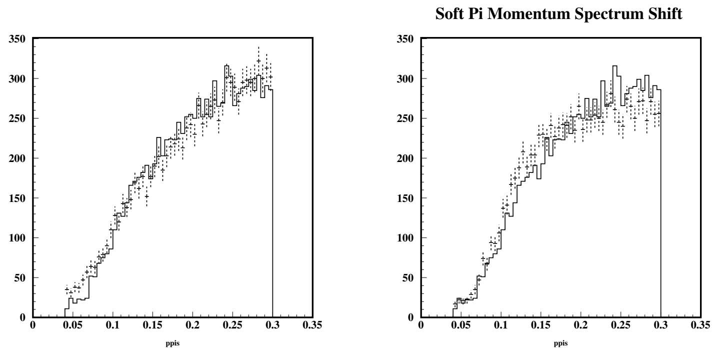


Figure 4.23: Nominal (histogram) and linearly shifted (points) shapes for the π_s (right) momentum spectrum; nominal (histogram) and D^*-shifted (points) shapes for the spectrum(left).

Each effect, the size of the χ_d shift, and the likelihood change for the fits, is summarized in table 4.31. The total systematic uncertainty is summarized in table 4.32.

Table 4.31: Uncertainty in χ_d due to distortions of the combinatoric shape. σ has been artificially decreased by 10% in all cases to reflect uncertainty due to the different number of fit bins.

Shape	MC Goodness of fit	"σ"	$\Delta\chi_d$	Syst. shift
Nominal	198.9/150	0	0	0
Linear shift in π_s	202.8/150	1.8	-0.28	-0.15%
Gaussian D^* shift in π_s	209.3/150(206.5/148)	2.9	-0.25	-0.09%
Linear shift in π_w	204.1/150	2.1	-0.28	-0.13%

Table 4.32: Systematic uncertainty in χ_d due to distortions of the combinatoric shape.

Systematic Uncertainty	$\Delta\chi_d$	Set by
Combinatoric Shape	-0.13%	Additional linear term in h_W shape
Combinatoric Shape	-0.15%	Additional linear term in π_s shape
Combinatoric Shape	-0.09%	Additional D^*'s
Combinatoric Shape	0.22%	Sum in quadrature

A physics shape systematic for signal and 2-body decays

There are primarily two ways in which the 2-body shapes may be systematically misestimated:

- Shapes may differ mode-by-mode in the individual modes which are summed to form signal or 2-body decays. Each mode has uncertainties in the relative size of its contribution.

- The detector tracking efficiency may differ from what is assumed, so that the odds for π to be lost or found will be wrong, changing each mode's shapes.

Here we address the first; the second is addressed in the next section. We reweight each B decay mode, with the relative reweighting given by the relative uncertainty in its total yield. For the D^{**} modes, we vary the yield by 50%, rather than the reported uncertainties. We refit using the new shape and observe the change in χ_d. We then add all the changes together in quadrature. We assess a total uncertainty of 0.5% due to 2-body mode-by-mode shape variation.

Tracking Efficiency

We reweight the Monte Carlo events as a function of momentum, reweighting according to the uncertainty at each momentum. We use the uncertainties given in Bruce Berger's study of `prelib` efficiency [78], combining in quadrature the shifts and uncertainties from table 8 of that note. We fit Bruce's numbers as a function of momentum to get the reweighting as a function of momentum for

Table 4.33: Variation of B modes. The nominal fit yields χ_d=19.67%. The uncertainties for D^{**} modes reflect a 50% variation on their total yields, rather than the quoted uncertainties.

Decay Mode	B.F. (%)	Efficiency (%)	Yield / 1M$B\overline{B}$	$\Delta\chi_d$ (%)
B^0 Decays				
$D^{*-}\pi^+$	0.28±0.02[9]	14.7±0.8	51±2.5	0.02
$D^{*-}\rho^+$	0.73±0.06[72]	9.1±0.3	83±4	-0.05
$D^{*-}\pi^+\pi^0$	0.7±0.5[9]	1.4±0.1	12.3±8.7	0.24
$D^{*-}a_1^+$	1.3±0.3[9]	0.2±0.05	3.3±1.1	0.03
$D^{*-}\mu\nu$	4.60±0.3[9]	0.08±0.01	4.6±0.5	-0.16
$D^-\pi^+$	0.30±0.04[9]	0.3±0.07	1.1±0.6	0.03
$D^-\rho^+$	0.79±0.14[9]	0.17±0.07	1.7±0.8	0.27
$D^-\pi^+\pi^0$	0.05±0.05[71]	0.2±0.2	0.1±0.1	0.00
$D^-a_1^+$	0.6±0.3[9]	0.2±0.1	1.2±0.1	0.06
$D^-\mu\nu$	2.0±0.3[9]	0.01±0.01	0.2±0.2	-0.02
$D^{**}-\pi^+$	0.20±0.05[73]	1.2±0.2	3.0±0.8	0.02
$D^{**}-\rho^+$	0.57±0.13[73]	0.3±0.07	2.1±0.7	0.02
B^+ Decays				
$\overline{D^0}\pi^+$	0.53±0.05[9]	0.2±0.1	1.3±0.6	0.07
$\overline{D^0}\rho^+$	1.3±0.2[9]	0.3±0.1	4.9±1.6	-0.03
$\overline{D^0}\pi^+\pi^0$	0.05±0.05[71]	<1.0		-0.01
$\overline{D^0}\mu\nu$	1.9±0.3[9]	0.02±0.01	0.4±0.2	-0.26
$\overline{D^0}a_1^+$	0.5±0.4[9]	0.02±0.02	0.1±0.1	0.02
$\overline{D}^{*0}\pi^+$	0.46±0.04[9]	0.1±0.05	0.5±0.3	0.02
$\overline{D}^{*0}\rho^+$	1.6±0.3[9]	0.25±0.07	5.0±1.5	0.10
$\overline{D}^{*0}a_1^+$	1.9±0.5[9]	0.04±0.04	0.8±0.8	Failed
$\overline{D}^{*0}\pi^+\pi^0$	0.05±0.05[71]	3.2±1.7	2.0±2.0	-0.29
$\overline{D}^{*0}\mu\nu$	5.3±0.8[9]	0.015±0.007	1.0±0.5	-0.00
$D^{**}0\pi^+$	0.33±0.08[73]	2.2±0.3	9.1±2.3	-0.28
$D^{**}0\rho^+$	0.92±0.22[73]	0.45±0.14	5.1±1.9	0.00
Total (quadrature)				0.61
Average				-0.01

CLEO II; the result is shown in figure 4.24. We also set to 0 the efficiency for any track under 52 MeV.

For CLEO II.V I double Bruce's uncertainties and shift them upwards by 7 MeV. (In CLEO II.V, the average momentum required for a π to penetrate the silicon is 7 MeV higher than in CLEO II.) I set to 0 the efficiency for any track under 63 MeV. The reweighting is shown in figure 4.24. This procedure results in a net 4% change in the total number of Monte Carlo events. The projections of the fits to the data after the reweighting are shown in 4.25.

The answer shifts by 0.015%. Actually, this tiny dependence is not a surprise. To first order, efficiency cancels because mixing is a ratio. Tracking efficiency can only affect the answer if the inefficiency *shape* is not flat over the region of π_s we are looking for, *and* if that shape difference causes systematic shifts in the fit components. As I show in figure 4.26, the soft pions in this analysis are all over 90 MeV, and almost all over 100 MeV–in the region where the efficiency shape is pretty flat in any case.

This procedure correctly accounts for systematic uncertainties due to tracking efficiency mismodeling if events do not have additional possible solutions with different tracks. When there are additional solutions, there is an another possible effect. It is possible for the chosen solution to change when the efficiency changes (due to the original solution no longer being found). We find the expected magnitude of this effect is 0.0002% and we neglect it.

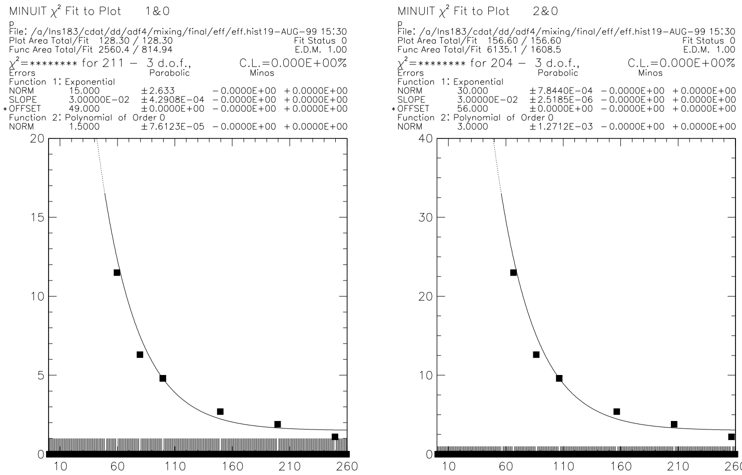


Figure 4.24: Parameterization of the tracking efficiency uncertainty, as a function of momentum (MeV). The values of the points are from Table 9 of [78].

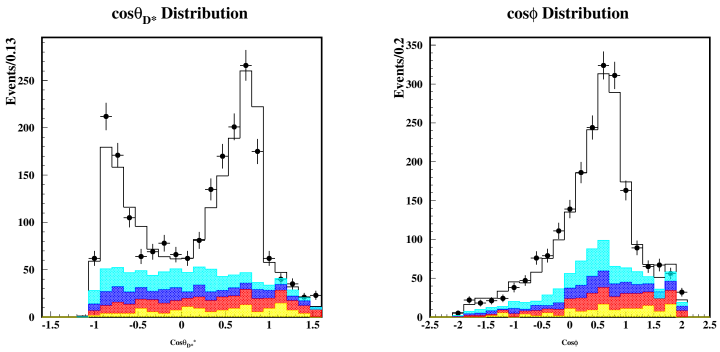


Figure 4.25: Projections of the fits after reweighting by tracking efficiency.

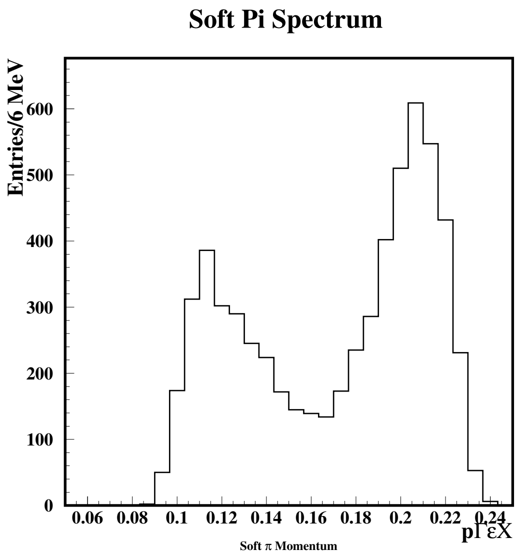


Figure 4.26: The soft track momentum spectrum.

Beam Energy Distribution

We believe that the beam energy distribution affects primarily the (narrowly peaked) signal shape, and the others to a much lesser degree. What matters really is the B momentum. This is because a shift in the absolute energy scale that is coherent in E_{beam} and m_B will effect only p_D of equation 11, and not p_B of equation 10. We tune the beam energies to give the p_B observed in data. Since p_D averages over 2 GeV, a 2 MeV shift is negligible. So for this systematic we studied the signal events only. In figure 4.27 we show that the mean beam energies of signal MC and data differ by 100$\pm$400 keV; the RMS's are 10.7 and 10.9 MeV.

We directly use two different subsamples of signal Monte Carlo generated at different beam energies. When we use the signal shapes from two subsamples of the signal MC which differ in average energy by 1.6 MeV (each 800 keV above or below the central value), the value of χ_d changes by a total of 0.8%, with negligible statistical uncertainty. Since the real uncertainty is 400 keV and not 1.6 MeV, this shift corresponds to a systematic uncertainty of 0.20%.

As an alternative, we can we reweight the Monte Carlo to reflect those shapes that result from a different distribution of beam energies, and refit with the new Monte Carlo shapes.

To evaluate the systematic uncertainty, we split the Monte Carlo into two sets which differ in energy by 1.7 MeV (5σ). A comparison of the $\cos\phi$ and $\cos\theta_{D*}$ are shown in figure 4.28. The first differs significantly in the peak region, the second does not. The mean value of the peak, as determined by Gaussian fits to the central region (0 to 1.5) shifts by 0.03, but the RMS remains the same.

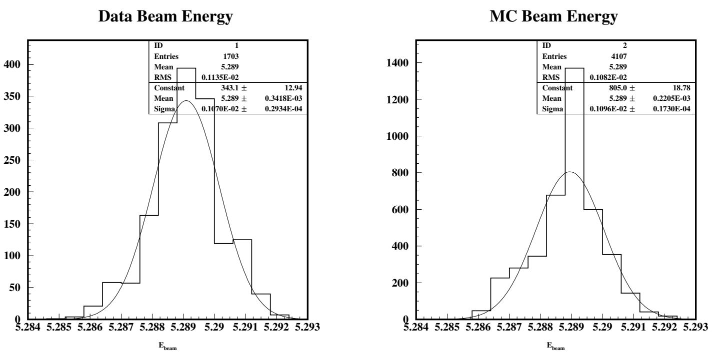


Figure 4.27: Distribution of beam energy in signal Monte Carlo and data.

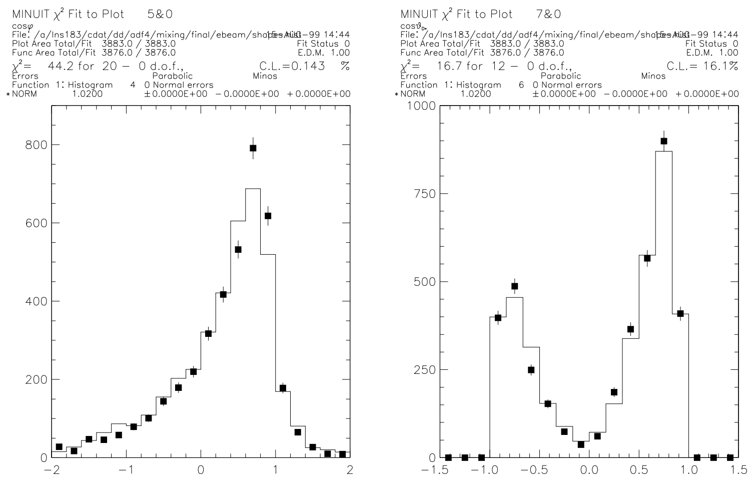


Figure 4.28: Distributions of $\cos\phi$ (left) and $\cos\theta_{D^*}$ in signal Monte Carlo. The points with errors are from Monte Carlo generated with average beam energy 800 keV above the average beam energy in data, the histogram from Monte Carlo 800 keV below the data. The histograms have uncertainties comparable to the data points.

Table 4.34: Summary table of possible beam energy systematics. We use line 2 in the final evaluation.

Systematic Uncertainty	ΔE	$\Delta\chi_d$	Set by
Beam Energy	400 keV	0.20%	Two MC subsamples
Beam Energy	400 keV	0.03%	Reweighted Monte Carlo

We reweight the Monte Carlo to produce a shift of 0.03 MeV in the peak region from 0 to 1.5 but to keep the same distribution below 0. This would be an appropriate reweighting if the distributions differed by 1.7 MeV, but is four times too big given that the data and MC agree with uncertainty 400 keV. So we refit with the new reweighted shape and divide the shift in χ_d by four (since it is a 4σ shift). The fit shifts the χ_d result downwards by 0.1%, so we assign a systematic uncertainty of 0.025%.

Table 4.34 summarizes the beam-energy systematic uncertainty studies.

Multiple entries

In table 4.15 we describe how we choose a solution when there are multiple solutions in an event. Because we are making a preference-based choice, the shapes of the distributions have some dependence on the accuracy of the Monte Carlo predictions for the multiplicity of additional solutions.

The Monte Carlo predicts 0.217 additional solutions per event; the data have 0.223 ± 0.013 additional solutions. I reweight the signal monte carlo with more than one solution to produce an upwards shift in the average number of solutions of of 0.04 (3σ). We find that the χ_d results shifts upwards 0.2%. Since we do have access to number of solutions for events in the data, we should in principle use a shift of $0.2\%*\frac{0.223-0.217}{0.04} = 0.03\%$ to the central value as a correction; 0.03% is negligible. We assign one-third the shift (since it was 3σ) as the systematic uncertainty.

Polarization of $B^0 \to D^*\rho$

We vary the polarization up one σ, from 0.87 to 0.92. Repeating the analysis, χ_d changes by 0.12%. We see no significant change in the distribution of $\cos\theta^*_{D^*}$, as shown in figure 4.29.

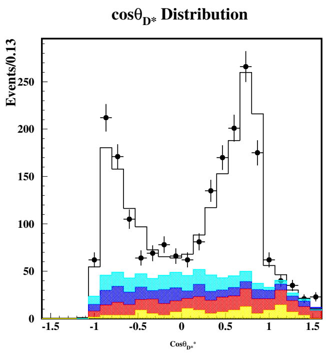


Figure 4.29: Projection of fit for $\cos\theta_{D*}$ when the ρ polarization is varied by 1 σ.

Asymmetric Mistag Rates

We find in the generic Monte Carlo that the overall mistag rate is 3.2±0.5%. Dividing the sample into truly mixed and truly unmixed events, we find mistag rates of 3.4±1.3% and 3.2±0.5%, respectively. We find the difference in mistag rate between mixed and unmixed events is 0.2±1.4%. We assign an uncertainty for allowing the mixed mistag rate to be up to 1.4% different from the average mistag rate. This results in a 0.3% uncertainty in the value of χ_d.

4.6.3 Mixed to Unmixed Efficiency

For the mixed to unmixed efficiency, there is only one method: direct comparisons of the efficiency in Monte Carlo. We made four independent measurements shown below in table 4.35. Combining the measurements, we see no evidence for differential efficiency within the 2% uncertainty of the study. This corresponds to an uncertainty of 0.4% in χ_d.

4.6.4 $\cos\theta_B$

One of our cross-checks, the prediction from the fit for the distribution of $\cos\theta_B$, shown in figure 4.19, does not appear to agree well with the data. The fit is high for low values of $\cos\theta_B$ and low for high values. The plot and residuals are shown in figure 4.30.

We are forced to ask two questions. First, does this indicate a systematic uncertainty in the result for which we have not already accounted? As we will see,

Table 4.35: Mixed:Unmixed efficiency systematics.

Systematic Uncertainty	$\frac{\varepsilon_{mixed}}{\varepsilon_{unmixed}}$	Set by
Mixed:Unmixed Efficiency	1.04±0.04	`prelib` generic MC
Mixed:Unmixed Efficiency	0.92±0.07	Recompress generic MC
Mixed:Unmixed Efficiency	1.01±0.03	Signal MC
Mixed:Unmixed Efficiency	0.90±0.05	More Signal MC
Mixed:Unmixed Efficiency	1.00±0.02	Combined

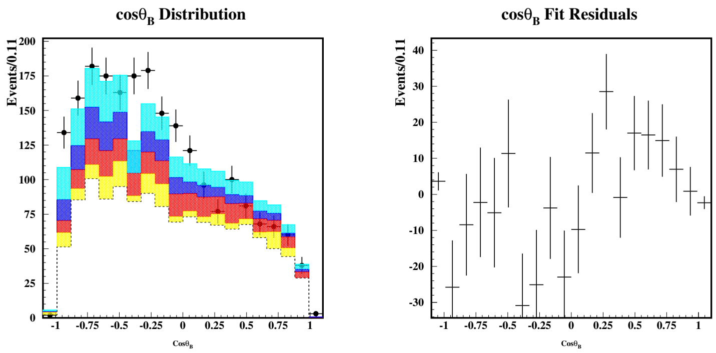


Figure 4.30: Residual of the 1-d cosθ_B signal shape.

the answer will be no: the disagreement is can be attributed to a 1σ variation in the beam energy. Second, does it suggest that we may more obtain a more accurate result by shifting the beam energy used in the analysis within its uncertainties? We conclude that the answer is yes.

We quantify and probe this discrepancy in the $\cos\theta_B$ plot in several ways. First we inform the reader that the χ^2 of the nominal fit is 31.1 for 22 d.o.f. This corresponds to a confidence level of 9.6%. Secondly, we find that the slope of the residuals is significant at the 2.4σ level.

Beam Energy Effect on $\cos\theta_B$

We have already assigned a systematic uncertainty for E_{beam} (which enters $\cos\theta_B$ primarily through p_B.) Is it the right size, and is the effect consistent with being due to the beam energy? To quantify the effect of E_{beam} on the $\cos\theta_B$ shape, we consider the mean value of $\cos\theta_B$ for signal events. In the Monte Carlo, we find that a shift of 500 keV in beam energy corresponds to a change in of $\cos\theta_B$ 0.06+-0.02; lower values of the energy cause an upwards shift in the mean value of cos theta B.

The discrepancy we see in the data can be quantified in a similar manner. We find that the mean value of the Monte Carlo is 0.042 higher than in the data, corresponding to a mean beam energy different by 350 keV. E_{beam} has already been considered as a systematic uncertainty, where we considered a variation of 450 keV.

If we actually reconstruct the MC with an energy 450 keV above its mean value, the fit result for $\cos\theta_B$ is shown in figure 4.31. I find the fit χ^2 is 16.9 / 20 d.o.f. (66%). The residual slope is greatly reduced, from the nominal 11.7 to 7.1, as shown in figure 4.32.

This suggests that the beam energy uncertainty introduces a systematic uncertainty of 4.6 to the slope; this in turn suggests we should think of the slope as 11.7±5.0±2.2±4.6 (errors are statistical, branching-fraction, and beam energy), which in sum corresponds to 11.7±7.1.

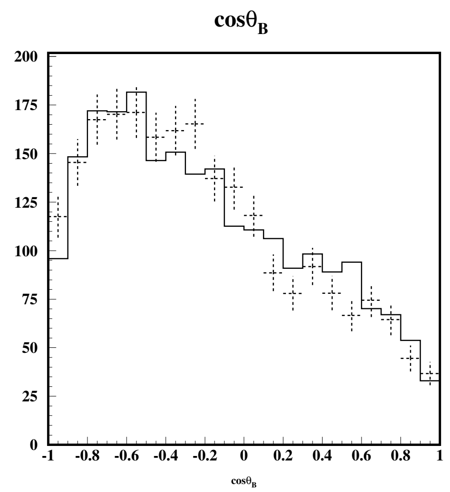


Figure 4.31: Distribution of $\cos\theta_B$ in data (points) and Monte Carlo (histogram), when Monte Carlo is reconstructed 450 keV above its nominal value.

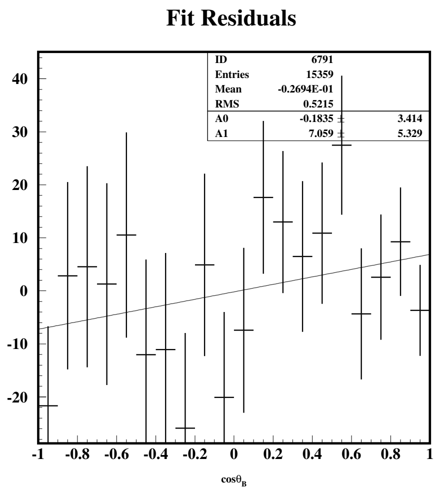


Figure 4.32: Distribution of $\cos\theta_B$ in data (points) and Monte Carlo (histogram), when Monte Carlo is reconstructed 450 keV above its nominal value.

4.6.5 Other effects

We explored a variety of other possible effects to explain the $\cos\theta_B$ problem. We find that the value of mixing, as measured in high $\cos\theta_B$ and low $\cos\theta_B$ regions, is consistent within large uncertainties. We find that fitting to the plot of $\cos\theta_B$ only does not relieve the problem, as it produces a preposterously low signal level. We find that by varying the fit components within their uncertainties, we cannot make the residual slope go away. We cannot make the slope go away by varying the 2-body branching fractions within their uncertainties. The momentum distribution of the h_W tracks looks beautiful and cannot account for discrepancy. The fractional E_{beam} error enters the calculation of $\cos\theta_B$ at a negligible level (0.04%).

Other Effects and Conclusion

We find an E_{beam} uncertainty of 400 keV, can indeed cause the effect we see, and has already been accounted for in the tally of systematic uncertainties. We find no evidence, within large uncertainties, that the measured values of χ_d are different for high and low values of $\cos\theta_B$. When we fit the $\cos\theta_B$ distribution alone, the fit value for the total signal level is patent nonsense.

Given that the other distributions are in good agreement (and that one of them has little effect of $\cos\theta_B$), and given that several other tests (high vs. low $\cos\theta_B$, fitting only in $\cos\theta_B$, varying 25 branching fractions) have revealed nothing, and given the significant residual slope reduction we find by varying the beam energy 1σ, we believe we are the vicitims of beam energy uncertainty. We believe this is already covered in the systematic table.

Because we believe there may be a real effect due to the beam energy, we shift the χ_d result by 1 systematic σ downwards, to 0.197.

4.6.6 Hadronic $\overline{B^0} \to \pi^+ D^{*-}$ Decays

A correctly-reconstructed event may in principle have a π and lepton of the same charge even in the absence of $B - \overline{B}$ mixing. This is because there is a tree-level $\overline{B^0} \to \pi^+ D^{*-}$ decay, which arises from a $b \to u$, followed by upper-vertex Cabbibo-suppressed $c\overline{d}$ decay. The rate for such a process is suppressed by a factor of 10^{-3} relative to the mixed decay and may safely be neglected. However, there can be interference between the mixed and direct decays. In [19], it is proposed to exploit this interference for detection of CP violating quantities. In the present context, however, we must account for whether this interference can change the apparent mixing rate. Using [79], we find that the interference term causes an apparent χ to shift relative to the true χ by the fractional amount Δ_χ

$$\Delta_\chi = \frac{\frac{b}{a}\lambda^2 \rho \cos\delta}{1+x^2} \tag{4.24}$$

where x is the mixing parameter, λ and ρ are Wolfenstein parameters, δ is the strong phase shift, and $\frac{b}{a}$ is ratio of matrix elements, expected [79] to be near 0.5. Inserting Standard Model values and uncertainties from [79] and [80] we find the fractional shift is a negligible 0.4%.

4.6.7 Systematic Uncertainty Tables for χ_d

In table 4.36 we show a detailed list of the systematic uncertainties in the analysis. In table 4.37 we show a summary table.

4.6.8 A_{CP}

Because we measure A_{CP} with χ according to the lepton charge, CP-violating detection efficiency differences for the leptons cancel. However, they do not cancel for the hadronic portion of the analysis, or for lepton fakes. We address these systematic uncertainties below. In table 4.38 we show the final assigned systematic uncertainty.

Hadronic CP-Violating Detector Asymmetries

As mentioned earlier, there is strong evidence for detector CP symmetry of reconstruction of generic tracks [75–77, 84] and .

Here we make a study of low-momentum tracks that is more specfic to our analysis.

We choose an inclusive sample of low-momentum (less than 300 MeV) tracks in the data which are associated with a high-momentum (above 1.4 GeV) lepton in the event. We skim these tracks out of the 4sA and 4sK data sets. We require these tracks to pass the same track-quality cuts as tracks in this analysis. In total we find 251,141 tracks below 300 MeV. 125,505 are of positive charge, 125,636 of negative, giving a CP asymmetry of -0.05 $\pm$ 0.10 %. The asymmetry is flat as a function of momentum, as shown in figure 4.33.

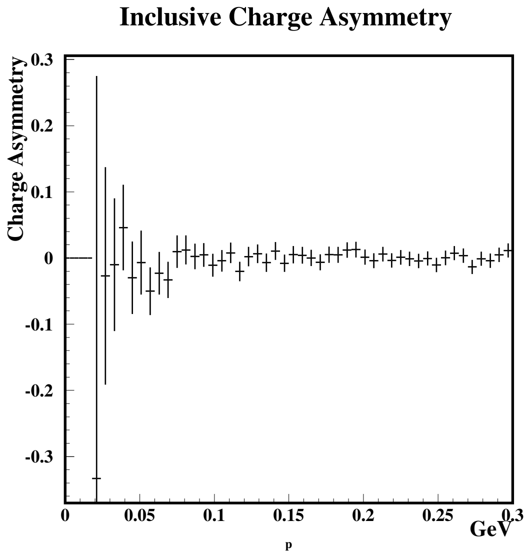


Figure 4.33: Track charge asymmetry ($\frac{N_+ - N_-}{N_+ + N_-}$) as a function of track momentum, for tracks that satisfy the track quality cuts of the analysis, in events with a lepton above 1.4 GeV. The measured net asymmetry is -0.05 $\pm$ 0.10 %.

Under the assumption that real CP violation may be neglected at these momenta, we have set a direct limit on detector-induced CP asymmetry in the reconstruction of soft tracks in this analysis. If there is real CP violation such violation might conspire with a real detector asymmetry to produce the smallness of the observed asymmetry.

To test this possibility, we examine the track asymmetry for all events, with or without leptons, for which the soft pions still must pass our good track quality cuts. We remind the reader that most hadronic tracks originate from continuum *udsc* events, while events with high-momentum leptons are primarily B decays. We postulate that a fixed detector asymmetry would not be cancelled out by a conspiracy of true CP violation in two such different samples.

We find 310,430 low-momentum positively charged tracks, and 308,608 negatively charged, giving an asymmetry of 0.3±0.1%.

As a third check, we employ lepton events in the off-resonance data of 4sH. We expect no CP-violating physics in the continuum data. We find 14977 positive and 15011 negative well-tracked low-momentum tracks. This yields an asymmetry of -0.1±0.6%.

We assign a systematic uncertainty by adding in quadrature the magnitudes and uncertainties of all three track samples, resulting in a total uncertainty of 0.7%. The total systematic uncertainty may be found in 4.38.

Finally, let us note that at low momentum, CP-violating asymmetries can be apparently large due to nuclear interactions which produce more p than $\overline{p}$. This does not really affect us because the track quality cuts, which include a `DBCD` cut, reject most tracks from such interactions. As a cross check, we release the track quality cuts on the second sample (no lepton required) and measure the CP asymmetry. We find 2.8±0.1%.

Hadronic Fakes

Assuming that all fakes are of a single sign shifts the asymmetry by 1.2%. This is a dramatic assumption but clearly a limit on the systematic uncertainty. We use this 1.2% shift as a systematic uncertainty.

4.7 Consistency Checks

4.7.1 Fit procedure variations

1-D Fits

We also perform 1-D fits to the $\cos\phi$ and $\cos\theta_D$ distributions. These are shown in figures 4.34 and 4.35. In both cases the parameters are consistent with those found by the 2-D fit, with χ^2 of 1.17 and 0.8 for 4 d.o.f. Of course, the fits are strongly correlated. If we undertook to measure χ_d with these fits, we would find values of 19.2$\pm$2.4% and 21.2$\pm$2.0%, compared to the nominal value of 19.9%. The assigned uncertainties are due only to the increased statistical uncertainty from the fit, since all other systematics and statistics are correlated.

Signal and sideband

We also compared the result if we do not simultaneously fit the two-dimensional signal and sideband plots, but fit each independently and compare the parameters. We find the parameter agreement has a χ^2 of 4.2 for 4 d.o.f., corresponding to a confidence level of 38%.

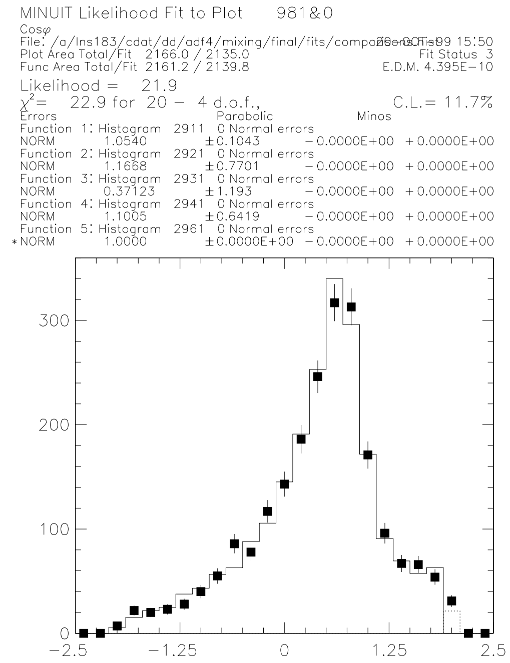


Figure 4.34: Fit to the data using only the 1-d cosϕ signal shape.

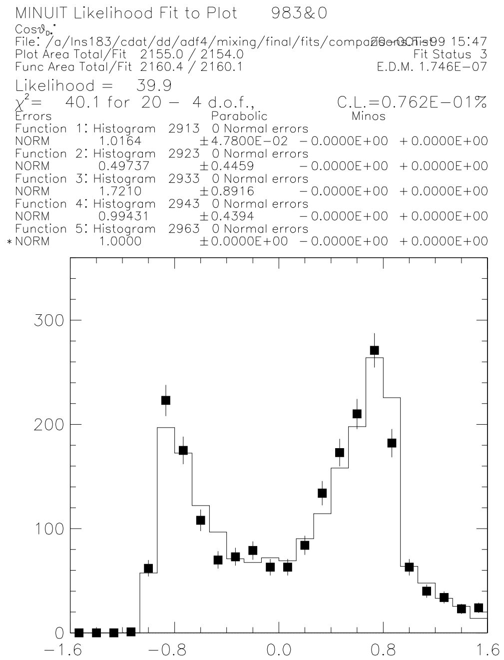


Figure 4.35: Fit to the data using only the 1-d $\cos\theta_{D^*}$ signal shape.

4.7.2 1-0 Plots

In order to verify that our fitter is unbiased and check for any overlooked mistakes, we perform the following check. We generate 8 artificial Monte Carlo "data samples", and fit for the value of mixing and charged B fraction in the samples, using the same code, scripts, and kumacs as the data. We plot measured vs. input χ and measured vs. input $B^{\pm}$; the points should lie on a line of slope one and intercept zero (hence "1-0").

The Artificial Data Samples

We generate eight different data samples, for the most part independent of each other and of the fitting functions. In each case, we combine known amounts of signal, 2-body, and synthetic Monte Carlo. We make datasets which cover a range of different possible scenarios, from low to high mixing, low to high charged B fraction. etc. The compositions of each dataset are shown in table 4.40.

For all eight cases, the signal and synthetic combinatoric samples are generated entirely independently of the fitting shapes, though of course drawn from the same distribution. Seven of the eight samples are statistically independent of each other; the "data-like" and "small" samples contain the same signal events. The 2-body B^0 shapes may contain 25% of the fitting sample, and the 2-body $B^{\pm}$ may contain up to 40% of the fitting sample.

Results

The results are shown in tables 4.41 and 4.42. The plots of reconstructed vs. true are shown in figure 4.36 with a fit to a line. Both lines are consistent with slope 1 and offset 0. At the fit reconstructed values of χ=19.9% and $f_{B^\pm}$=13.1%, the fit predicts biases of -0.4±0.4% and +0.3±0.4%, respectively.

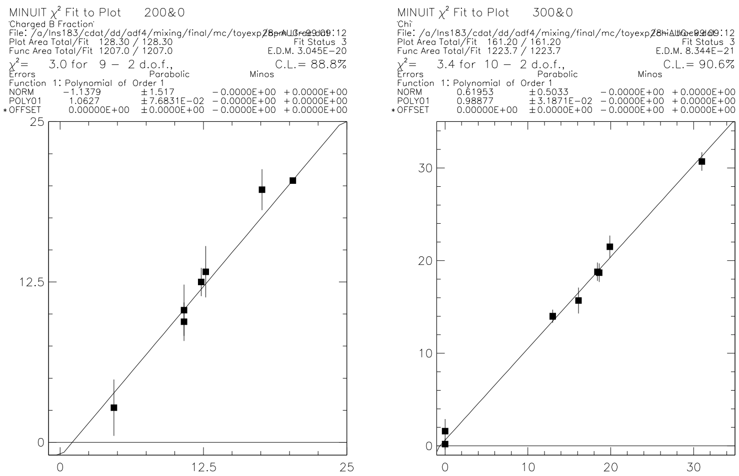


Figure 4.36: Reconstructed vs. generated values of the charged B fraction (left) and χ (right) for 8 artifical data samples.

4.7.3 Subsamples

We have measured χ_d using three different pairs of subsamples, and also dataset-by-dataset. We see no significant effects in comparing e to μ, π to ρ, or CLEO II to II.V, as seen in table 4.43. We show the dataset-by-dataset variation in figure 4.37. We see no significant dataset-variations, except for the 4sG; the χ^2/d.o.f. for the datasets is 27.8 / 25.

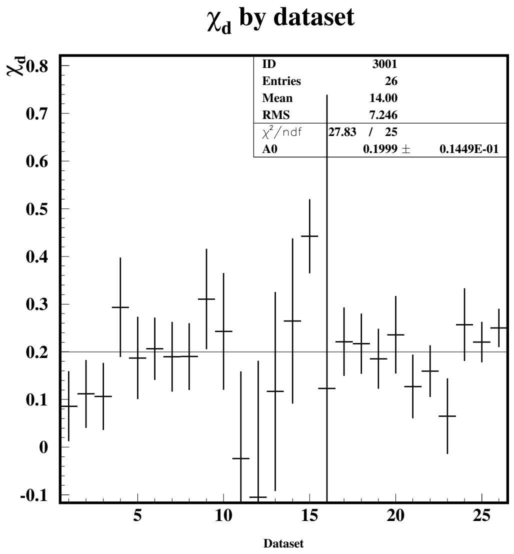


Figure 4.37: The measured value of χ_d, by dataset. Error bars are statistical only. The outlier is 4sG.

4.7.4 Variation with Binning

In principle the binning should not make a difference. We use several different binning procedures besides the nominal variable-bin size procedure. We see nothing disturbing as we vary the binning, as shown in table 4.44.

4.7.5 Absolute Yield Prediction from Monte Carlo

The Monte Carlo should predict the absolute yields of signal, 2-body B^0 decays, and 2-body $B^{\pm}$ decays. The predictions and fit are shown in table 4.45. We conclude that the Monte Carlo does a poor job of predicting the absolute normalization of the combinatoric background; fear of such a problem was precisely why we took the normalization in the analysis from the fit.

4.7.6 p_s and $cos\theta_{h-\pi_s}$

Naturally, our experience with $\cos\theta_B$, in which p_h gave considerably better agreement than $\cos\theta_B$, gave us pause with respect to the other two "experimental primitives" of the analysis, p_s and $cos\theta_{h-\pi_s}$. Their distributions and comparison the Monte Carlo, are shown in figure 4.38 and 4.39. There appears to be no cause for concern.

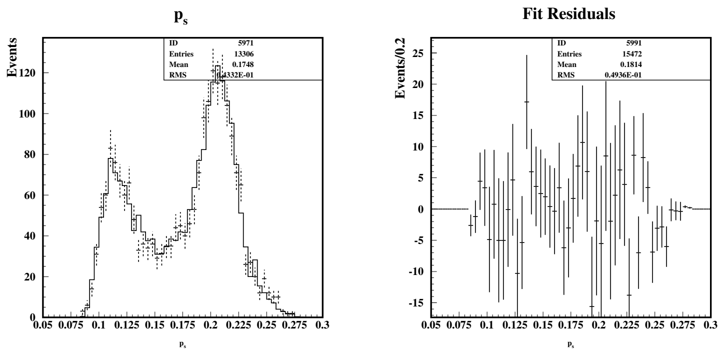


Figure 4.38: Data (points) and Monte Carlo prediction (histogram) for p_s (left); residual (right).

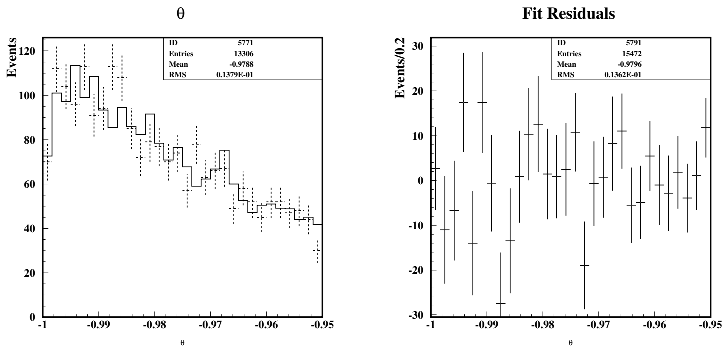


Figure 4.39: Data (points) and Monte Carlo prediction (histogram) for $cos\theta_{h-\pi_s}$ (left); residual (right).

Distribution of m_h

The m_h distribution in data, shown in figure 4.40, is consistent with the Monte Carlo (χ^2=32/27 d.o.f., C.L. = 23%), and has no significant slope. If we reweight to obtain better m_h agreement in the ρ width, we find a χ_d shift of 0.2%. We believe that as the data appears consistent with the Monte Carlo, and as we expect some m_h disagrement to arise as a result for 2-body branching fraction variation, this uncertainty is already covered by the 2-body shape uncertainty.

4.7.7 A_{CP} Cross-checks

As a cross-check, we calculate A_{CP} using the more usual method of like sign number asymmetries:

$$A_{CP} = \frac{N_{h^+\ell^+} - N_{h^-\ell^-}}{N_{h^+\ell^+} + N_{h^-\ell^-}} \tag{4.25}$$

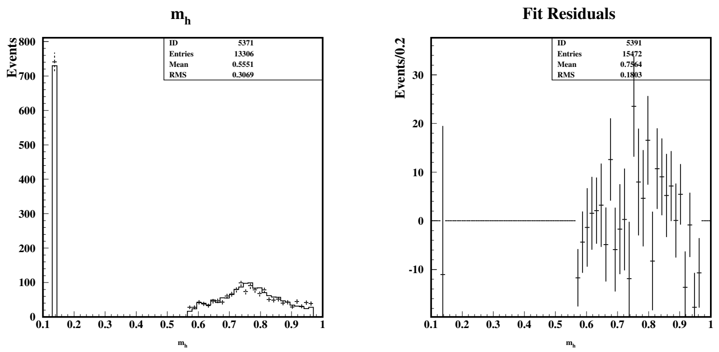


Figure 4.40: Data (points) and Monte Carlo prediction (histogram) for m_h (left); residual (right).

The counts are shown in table 4.46. We find (statistical errors only)

$$A_{CP} = 0.035 \pm 0.06 \tag{4.26}$$

which corresponds to

$$\Re(\epsilon_B) = 0.9 \pm 1.5\% \tag{4.27}$$

We expect the systematic uncertainties to be considerably worse than the 0.3% estimated for our method of choice.

4.8 Conclusion

We have described a measurement of χ_d and $\Re(\epsilon_B)$. The measurement is made by comparing the charge of a high-momentum lepton to that of a high-momentum π or

ρ from the decay $B^0 \to D^*\pi$ or $B^0 \to D^*\rho$ decay. We determine the composition of our mixing sample hadron tags by fits to signal and sideband regions of the angle between the soft pion from the D^* decay and the high-momentum hadron. We find small charged B contamination (13.1±2.5%) and small random background contamination (10.0±1.1%). We have also described a measurement of $\Re(\epsilon_B)$. We find that CP-violating detector asymmetries are small in the data. We find:

$$\chi_d = 19.7 \pm 1.3 \pm 1.4\% \tag{4.28}$$

$$\begin{aligned} \Re(\epsilon_B) &= 0.4 \pm 1.8 \pm 0.3\% \\ |\Re(\epsilon_B)| &< 3.3\% @ 95\% \text{C.L.} \end{aligned} \tag{4.29}$$

Table 4.36: Detailed table of systematics.

Systematic Uncertainty	$\Delta\chi$	Set by
e **Tag**		
Secondary Fractionf_c	$\pm$ 0.1%	PDG Form Factors, CLEO D Spectrum
Fake Rate	$\pm$0.02%	Fake rate + hadronic data
J/ψ Decays	$\pm$0.07%	MC statistics
γ Conversions	$\pm$0.03%	MC Statistics
π^0 Dalitz Decays	$\pm$0.03%	MC Statistics
μ **Tag**		
Secondary Fractionf_c	$\pm$ 0.1%	PDG Form Factors, CLEO D Spectrum
Fake Rate	$\pm$0.09%	Fake rates + hadronic data
J/ψ Decays	$\pm$0.07%	MC statistics
Upper Vertex Leptons	$\pm$0.2%	50% variation of BF
$B^{\pm}$ Fraction	$\pm$0.54%	Combinatoric B^+ fraction flat 0 to 1
$B^{\pm}$ Fraction	$\pm$0.48%	Propagated uncertainty
Combinatoric Mistag Fraction	$\pm$0.9%	$D^*\ell\nu$ data flavor-tagging study
Hadronic Mistag Fraction	$\pm$0.17%	Component relative sizes
Non-combinatoric shapes	$\pm$0.6%	Mode-by-mode BF+ efficiency variation
Combinatoric Shape	$\pm$0.23%	Generic MC constraint on parameterization
Beam Energy	$\pm$0.18%	$\pm$0.35 MeV
$D^*\rho$ Polarization	$\pm$0.12%	Uncertainty from CBX 98-47
mh	$\pm$0.25%	Reweight by m_h
Solution Multiplicity	$\pm$0.07%	Data statistics
Tracking Efficiency	$\pm$0.02%	Reweighting vs. p_π per CBX 98-36
Mixed:Unmixed Efficiency	$\pm$0.4%	Signal MC studies
Mixed:Unmixed Differential Mistag Rates	$\pm$0.3%	Generic MC studies
Total	$\pm$1.4%	Sum in quadrature

Table 4.37: Summary table of systematics for χ_d.

Systematic Uncertainty	Size	Set by
Lepton Mistags	$\pm 0.3\%$	Fake rate $\oplus$ cascades $\oplus$ real other leptons $\oplus$ UV
Hadronic Mistag Fraction	$\pm 0.9\%$	$D^* \ell \nu$ data flavor-tagging study, fit results
$B^{\pm}$ Fraction	$\pm 0.7\%$	Combinatoric B^+ fraction flat 0 to 100%, fit uncertainty
Mixed:Unmixed Efficiency	$\pm 0.4\%$	Signal MC studies
Mixed:Unmixed differential mistag rate	$\pm 0.3\%$	Generic MC studies
Signal Shape	$\pm 0.3\%$	$E_{beam} \oplus$ Polarization$\oplus$ Tracking$\oplus$ Multiplicity
2-body shapes	$\pm 0.6\%$	Mode-by-mode BF+ efficiency variation
Combinatoric Shape	$\pm 0.2\%$	Generic MC constraint on parameterization
Total	$\pm 1.4\%$	Sum in quadrature

Table 4.38: Summary table of systematics for A_{CP}.

Systematic Uncertainty	Size	Set by
Lepton Fakes	1.2%	100% of one sign
Soft Hadronic Tracking	0.7%	Inclusive soft track data
High-momentum Hadronic Tracking	0.03%	[84]
Total	$\pm 1.4\%$	Sum in quadrature

Table 4.39: Summary table of track asymmetry samples.

Sample	Positive Tracks	Negative Tracks	Asymmetry
Good track + lepton	125,505	125,636	-0.05±0.1 %
Good track	310,430	308,608	0.3±0.1%
Good track + lepton + OFF	14,977	15,011	-0.1±0.6%
Tracks	1,181,712	1,113,339	3.0±0.1%

Table 4.40: Description of the artificial “data sets”.

Mnemonic	Signal	2-body B^0	2-body $B^{\pm}$	Synthetic	χ (%)
CLEO II Mixless	385	104	59	0	0.
CLEO II.V Mixless	320	77	58	0	0.
Data-like	1378	178	359	150	19.9
Large Sample	3327	180	359	211	18.6
Small Sample	1584	145	59	75	18.4
Clean	2473	177	0	0	12.9
Dirty	2177	281	418	414	16.1
High Mixing	2974	281	418	139	31.0

Table 4.41: Results for measurement of χ. The uncertainty reflects both statistical and systematic uncertainties.

Mnemonic	$\chi_{true}(\%)$	Reconstructed $\chi(\%)$
CLEO II Mixless	0.0	0.2±0.6
CLEO II.V Mixless	0.0	1.6±1.4
Data-like	19.9	21.5±1.2
Large Sample	18.6	18.7±1.0
Small Sample	18.4	18.8±1.0
Clean	13.0	14.0±0.7
Dirty	16.1	15.7±1.4
High Mixing	31.0	30.7±1.0

Table 4.42: Results for measurement of the charged B fraction. The uncertainty reflects only statistical uncertainty from the fits.

Mnemonic	True $B^{\pm}$ fraction(%)	Reconstructed $f_{B^{\pm}}$ (%)
CLEO II Mixless	10.8	9.4±1.4
CLEO II.V Mixless	12.7	13.3±2.0
Data-like	20.3	20.4±0.2
Large Sample	10.8	10.3±2.0
Small Sample	4.7	2.7±2.2
Clean	0	0±0
Dirty	17.6	19.7±1.6
High Mixing	12.3	12.5±1.1

Table 4.43: The measured value of χ for several different subdivisions of the data.

Division	Measured 1	Measured 2	Δ
II / II.V	18.8±2.3	21.3±1.7	2.5±2.9%
e/μ	19.0±1.9	21.1±2.2	2.1±2.9%
π/ρ	18.9±1.9	20.7±2.0	1.8 ±2.9%

Table 4.44: Variation of the result with binning.

Binning Variable	$\Delta\chi_d$
20 $\times$ 20	0.2%
10$\times$10	0.2%
15$\times$15	0.4%
30$\times$30	0.2%
7$\times$7	-0.3%

Table 4.45: The composition of the selected sample, as determined by the fit. Also shown are absolute predictions for several subsamples from table 4.20.

Event Type	Fit	Fraction	Absolute Prediction
Signal	1244$\pm$52	68$\pm$3%	1219$\pm$61
2-body B^0	268$\pm$60	13$\pm$2%	269$\pm$80
2-body $B^\pm$	173$\pm$45	16$\pm$2%	275$\pm$40
Comb. B^0	116$\pm$13	5$\pm$1%	55 $\pm$ 5
Comb. $B^\pm$	78$\pm$8	3$\pm$1%	38 $\pm$ 5
Sum	1808$\pm$63	99$\pm$3%	1856
Continuum-Subtracted Data	1824	100 %	

Table 4.46: Measured quantities for the A_{CP} limit.

Quantity	Measurement	Statistical Uncertainty
$N_{h^-\ell^-}$	196	17
$N_{h^+\ell^+}$	211	16

CHAPTER 5

RESULTS OF $B - \overline{B}$ MIXING

5.1 Results for Δm

5.1.1 Comparison to other Measurements

In figure 5.1 we show results for χ_d from several other experiments. Many experiments measure Δm; we have translated them into χ_d measurements using $\tau_{B^0} = 1.56 \pm 0.06$ ps and setting y_d to 0. We find the present experiment is the most precise measurement of χ_d.

We may also translate the time-integrated measurements of χ_d into values of Δm. This is shown in table 5.2. When we do this, we find that only one experiment exceeds this result in sensitivity.

This experiment is statistically uncorrelated with all other reported experiments. It is also almost entirely systematically decoupled from all other experiments. Uncertainty arising from the ratio $\frac{f_{+-}}{f_{00}}$, which constituted a large systematic uncertainty in previous time-integrated results, does not affect this result. The secondary lepton fraction used in this result is derived from sources independent of those used in previous experiments.

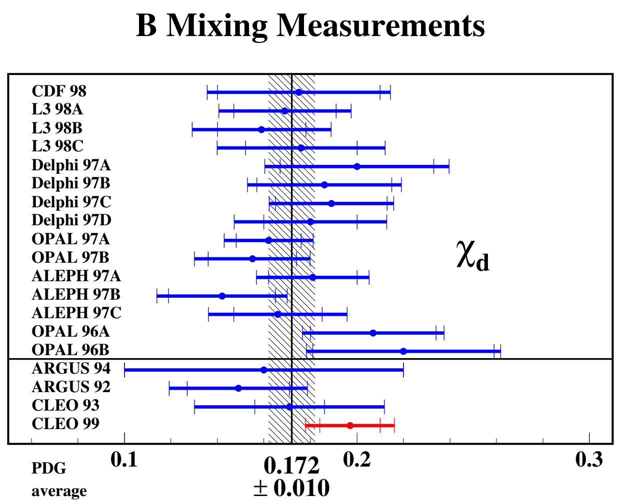


Figure 5.1: Summary of experimental results for χ_d, as given in the PDG [9]. The present work is shown as the lowermost point, labeled "CLEO 99".

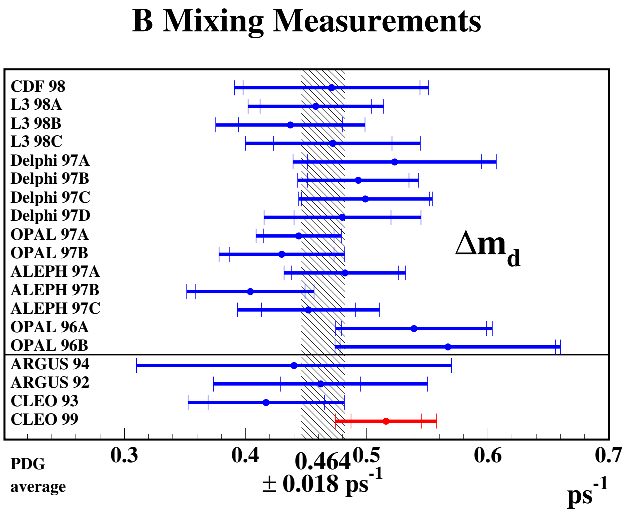


Figure 5.2: Summary of experimental results for Δm_d, as given in the PDG [9]. The present work is shown as the lowermost point, labeled "CLEO 99".

5.1.2 Implications for y_d

Time-dependent experiments typically measure only x_d and explicitly set y_d to 0 in their measurement. This allows us to set a limit on y_d by comparing time-integrated results (necessarily sensitive to y) to time-dependent ones. The average of all time-dependent results, given by the PDG [?], is a value of $x_d = 0.722$, corresponding to the value χ_d of 17.5±1.0%.

The PDG gives an average over previous χ_d experiments of 15.9±2.4% [?]. This experiment, is both statistically and systematically independent of those experiments. When combined with these previously published time-integrated rates, we find a result of χ_d=18.2±1.5%. We take the difference between the two results and find the one-sided limit on the difference in χ. A significant difference would be attributable to y_d. We use the formula [1] for χ:

$$\chi_d = \frac{x_d^2+y_d^2}{2(1+x_d^2)} \tag{5.1}$$

The difference is 0.7±1.8%, which gives a one-sided upper limit of 3.6% . This limits y <0.33 at the 95% confidence level.

5.1.3 Implications for the CKM Matrix

From Heavy Flavors 8 [85], the current best estimates for hadronic matrix elements are given in table 5.1. The results given are *unquenched*. It is worth noting that

[1] This differs substantially from the form given in the PDG on page 557. We believe the formula given in this thesis is correct. It may be obtained by integrating formula 1.39 over all times.

Table 5.1: Estimates for hadronic psuedoscalar decay constants and their ratios, as given by the Lattice 99 summary [85] and the MILC collaborations[86]

Decay Constant	Value	Reference
f_B	210±20	[85]
f_{B_s}	245±20	[85]
$\frac{f_B}{f_{B_s}}$	1.16 ± 0.04	[85]
f_{D_s}	210±30	[86]
$\frac{f_B}{f_{D_s}}$	0.73±0.10	[86]

the quenched and unquenched values for f_B and f_{B_s} differ by 20%, but the ratio differs by only 1%. This is a direct test of the claim that hadronic uncertainties cancel in the ratio of matrix elements.

The experimental value for f_{D_s} is in fact reasonably well-known. From the summary at the 1999 DPF meeting [88], the world average for f_{D_s} is 260±41 MeV.

We can combine the experimental results for f_{D_s} with the lattice results for the ratio $\frac{f_B}{f_{D_s}}$ from [86], and expect to have very little theoretical uncertainty. We find

$$f_B = 189 \pm 42\mathrm{MeV} \tag{5.2}$$

where the uncertainty comes from the combination of experimental uncertainty on f_{D_s} and the uncertainty quoted in [86] on the ratio $\frac{f_B}{f_{D_s}}$. We can take the expression for x_d from equation 1.29:

$$x_d = \frac{G_F^2}{6\pi^2\tau_B} f_B^2 B_B^2 m_B m_W^2 \eta_{QCD} S_0(m_t) |V_{tb}^* V_{td}|^2 \tag{5.3}$$

We take $\sqrt{B_B}$=0.90±0.05 [85] and m_t of 170 GeV to find, from this single experiment,

$$|V_{td}V_{tb}| = 0.107 \pm 0.024 \tag{5.4}$$

5.1.4 Implications for Supersymmetry

We also wish to bring the reader's attention to a point raised by M. Mangano [87]. Once B_S mixing has been measured, the absolute value of B_d mixing becomes a direct search for SUSY! This is because the ratio $\frac{x_{B_s}}{x_{B_d}}$ is insensitive to most new physics, and therefore allows one to measure the ratio $\frac{|V_{ts}|}{|V_{td}|}$. $|V_{ts}|$ is already known (since it is nearly equal to $|V_{cd}|$). Therefore the ratio of B_S mixing to B_d mixing becomes a measurement of V_{td}. Then, having measured V_{td}, we could ask, "Is the absolute value of B_d mixing consistent with coming entirely from Standard Model box diagrams?"

As a numerical example, let us take Δm=14±2 ps^{-1}. This central value may be intriguingly indicated by LEP and SLC data, and the uncertainty is one that might be acheived within a short time. Then

$$\frac{\Delta m_s}{\Delta m_d} = \frac{m_{B_s}}{m_B} |\frac{V_{ts}}{V_{td}}|^2 \frac{f_{B_s}\sqrt{B_{B_s}}}{f_{B_d}\sqrt{B_{B_d}}} \tag{5.5}$$

and $|\frac{V_{ts}}{V_{td}}|^2 = 23.9 \pm 3.8$.

Using $|V_{ts}| = 0.039 \pm 0.002$, we would find $|V_{td}|=8.0 \pm 0.7 \times 10^{-3}$. The difference between V_{td} determined in this way, and V_{td} determined from equation 1.29 can indicate SUSY contributions. For instance, [87] has explored the limits for stop and charginos masses for large $\tan\beta$, for higgsino-like charginos. For the hypothetical value of Δm=14$\pm$2 ps^{-1}, the present measurement could rule out stop and chargino masses less than 100 GeV.

5.2 Sensitivity Estimate for $\sin 2\beta$

In order to estimate the experimental sensitivity to the CP violation parameter $\sin 2\beta$ at a symmetric B factory, one must make assumptions about decay resolutions, reconstruction efficiencies, and effective flavor tagging efficiency. Expected decay lengths for $B^0(\overline{B^0})$ tagged $B \to \psi K_S$ decays are 34(22)μm at a symmetric $\Upsilon(4S)$ machine, for $\sin 2\beta = 0.7$ and $c\tau_B = 468\mu$m. The average y projections of these decay lengths are 19(12) μm. With an integrated luminosity of 30 fb^{-1}, effective effective flavor-tagging efficiency of 0.25, and average y-vertex resolution of 25 μm, Monte Carlo studies indicate that the ψK_S mode provides a measurement of $\sin 2\beta$ with statistical uncertainty $\pm$0.43. The CLEO II.V detector, in its measurements of the D lifetimes [21], has demonstrated the ability to control systematic uncertainties to a few μm. Due to the small decay lengths, sensitivity is nearly linear in the vertex resolution, in contrast to an asymmetric B factory.

APPENDIX A

IDENTIFICATION OF CHARGED K'S IN THE CLEO DETECTOR

A.1 K Identification

Using tracks tagged to QQ, we generate a spectrum showing the relative abundances of π, K, p in $B\overline{B}$ generic Monte Carlo. This plot is shown in figure A.1.

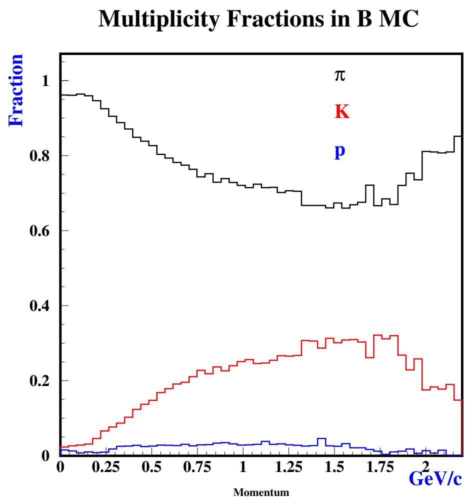


Figure A.1: Relative track abundances in $B\overline{B}$ generic MC. Black (most abundant) is π, red is K, blue (least abundant) is p.

We calculate a probability to be π,K,p as follows, where XX = PI, KA, PR, and DV = DI, TF, and $F_{XX}(p)$ is the momentum-dependent multiplicity fraction shown in figure A.1:

$$P_{XX} = F_{XX}(p)\mathrm{Prob}(\sum_{DV} \mathtt{SGXXDV}^2, N_{DV}) \tag{A.1}$$

From this we form $\mathtt{XXLEV} = \frac{P_{XX}}{P_{PI}+P_{KA}+P_{\mathrm{PR}}}$

Over most of the momentum range for $K^{\pm}$ from B decays, KALEV provides excellent identification of K and rejection of π; see figure A.2. We require KALEV to be greater than 0.9.

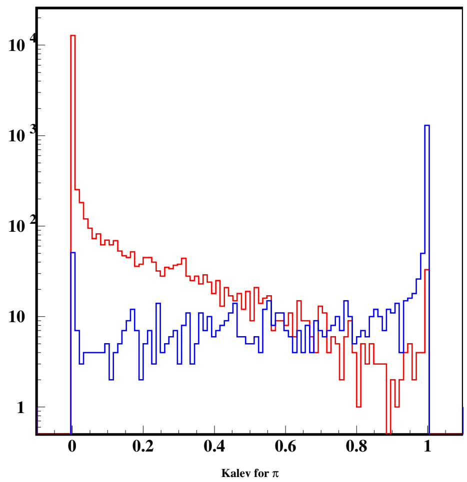


Figure A.2: Blue is true Kaons, red is everything else, requiring `R2ELEC` $<$ 0 and `DPTHMU`. Note the log scale.

A.2 K Yield

A.2.1 Yields and Fakes From QQ

From tracks tagged to QQ parents, we have made plots of the efficiency and fake rate of the Kaon ID procedure outlined in section A.1. The plot is shown in figure A.3. As the plot shows, for kaons from B decays, the particle ID is both efficienct and pure. The figure also compares 4sC and 4sM; there is almost no difference between them.

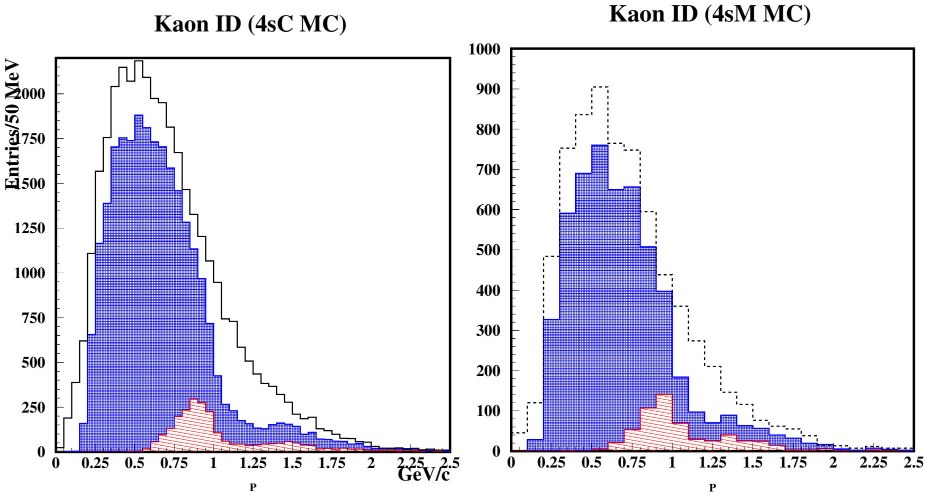


Figure A.3: Efficiency and purity of the K id using `KALEV`. The red area is π selected as K, the blue is K selected as K, and the solid histogram is the total number K tracks. There is essentially no difference between 4sM and 4sC.

We have also generated the fake rates and efficiencies as a function of momentum, these are shown in figure A.4.

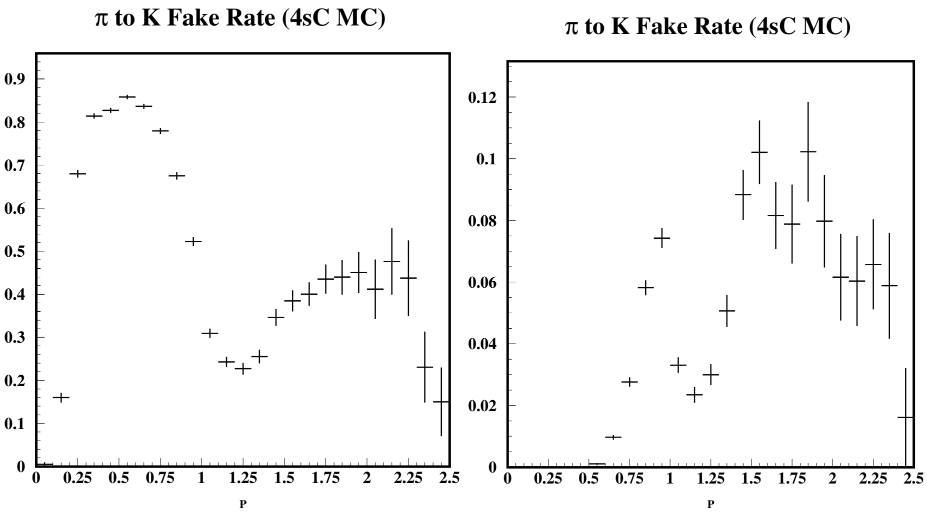


Figure A.4: Efficiency (left) and $\pi \to K$ fake rate (right) of the K id using `KALEV`.

A.2.2 Efficiencies and Fake Rates in Data

In figure A.5 we have selected D^* events in the data. The sideband-subtracted results are shown for the 4sC and 4sM, and compared to a combination of 4sC and 4sM Monte Carlo. We find the agreement is quite good.

A.2.3 Data-MC Production Comparisons

Argus has published species production per Υ as a function of momentum in [51]. In figure A.6 below, we show a comparison of the ARGUS data and CLEO Monte Carlo. The agreement is not good at low momenta. Note that the integrated K yield is off by about 10%.

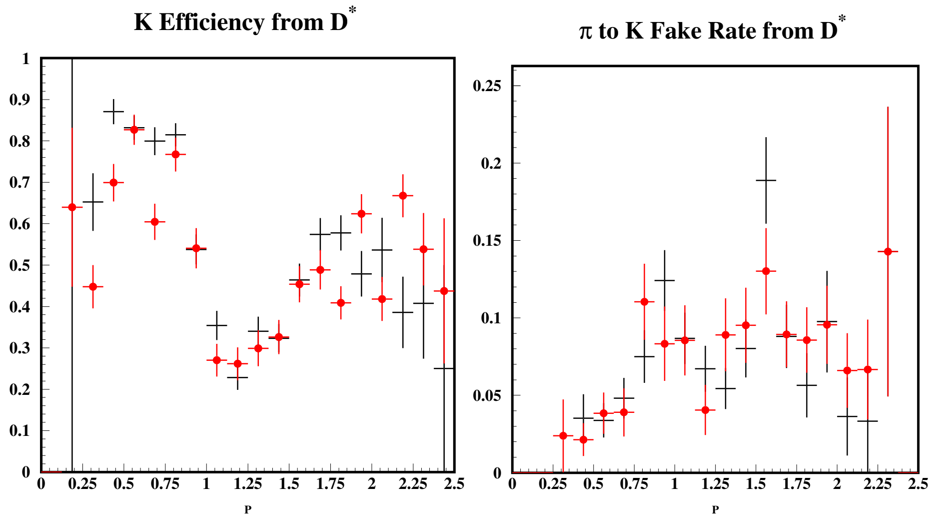


Figure A.5: Efficiency (left) and $\pi \to K$ fake rate (right) of the K id using `KALEV`. The data (red) are taken from D^* events in the 4sM and 4sC, the Monte Carlo(black) from also 4sC and 4sM.

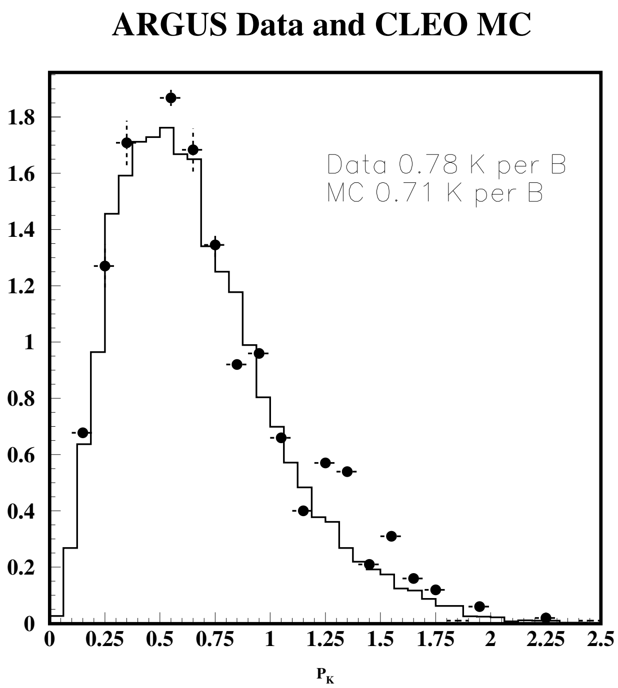


Figure A.6: Comparison of ARGUS data (from [51]) and the CLEO QQ spectrum.

APPENDIX B

DISTRIBUTIONS OF BASIC FLAVOR-TAGGING QUANTITIES

B.1 Parameterization of Dilutions

B.1.1 B^0 Dilutions

The plots in this section show the fits that are used to parameterize the dilutions. In each figure, the upper left-hand plot shows the distribution for negative (blue/black) and positive (red/gray) tracks, and the upper right-hand plot shows the bin-by-bin asymmetry. The two lower plot(s) show the fit(s) to the asymmetries. Sometimes the fits are to two separate regions in the plot. It is from these fits that we calculate the expected dilution for a track.

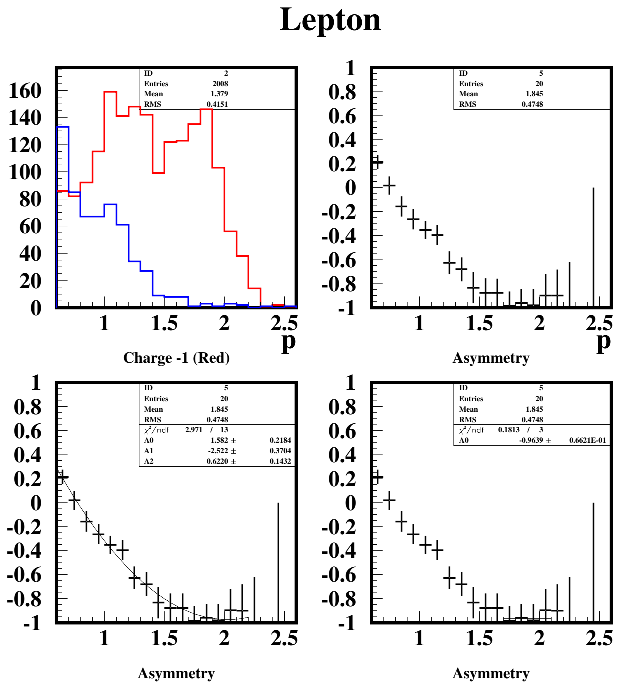


Figure B.1: Dilution in B^0 events as function of momentum for leptons. In this and the following figures, the upper left-hand plot shows the distribution for negative (blue/black) and positive (red/gray) tracks, and the upper right-hand plot shows the bin-by-bin asymmetry. The two lower plot(s) show the fit(s) to the asymmetries. Sometimes the fits are to two separate regions in the plot.

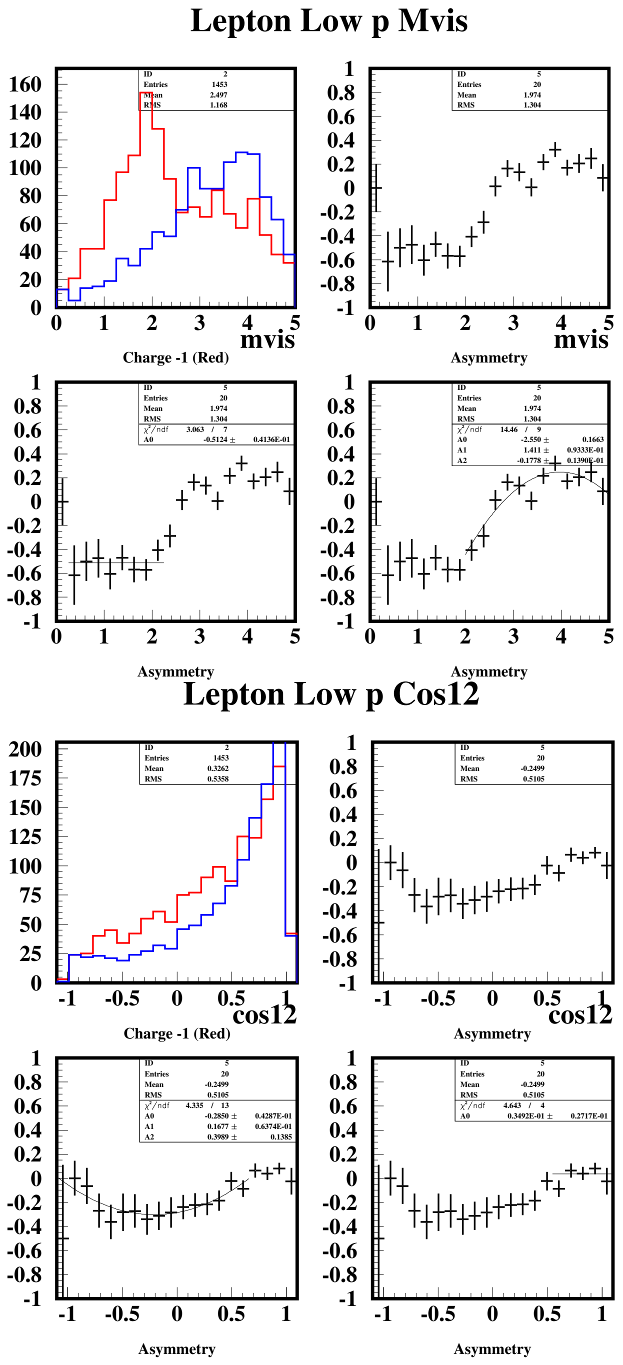


Figure B.2: Dilution in B^0 events as function of the visible energy (left) and angle to nearest track (right) for leptons below 1.4 GeV.

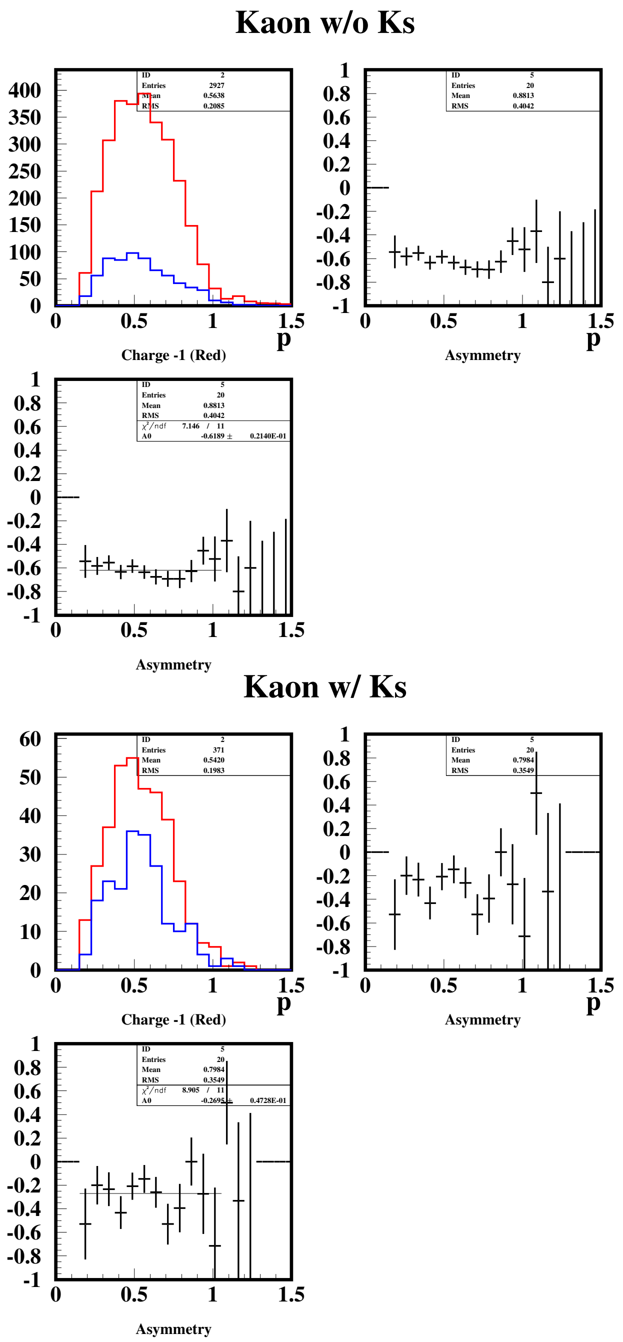


Figure B.3: Dilution in B^0 events as function of momentum for $K^{\pm}$, when there are no K_s in the event (left) and at least one K_s (right).

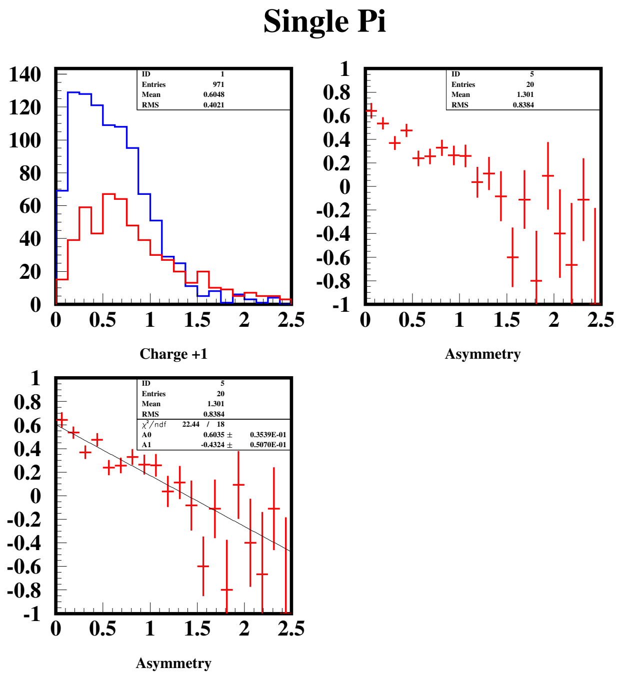


Figure B.4: Dilution in B^0 events as function of momentum for π when the track is the only identified π in the event.

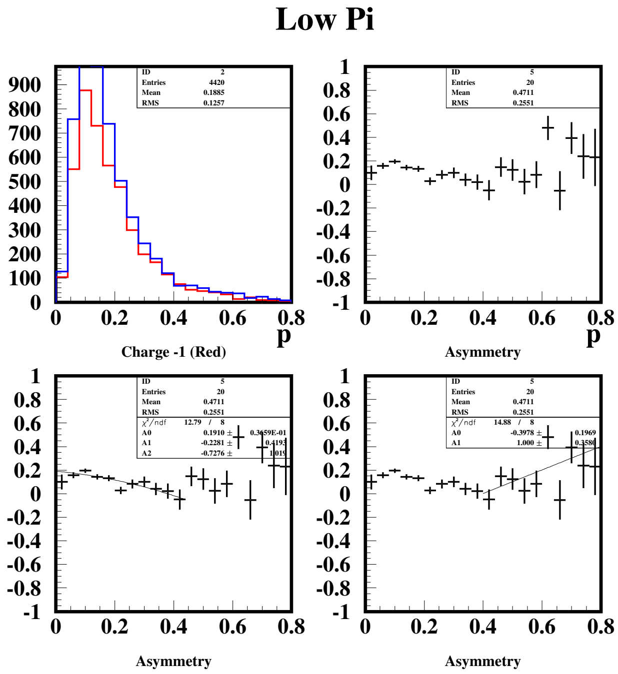


Figure B.5: Dilution in B^0 events as function of momentum for π when it is the softest π in the event, and is not the only π.

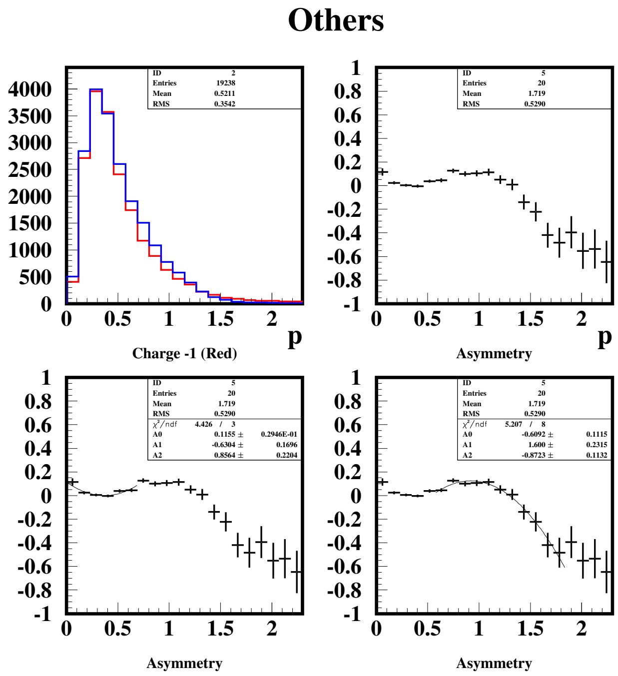


Figure B.6: Dilution in B^0 events as function of momentum for all other π's.

APPENDIX C

DISTRIBUTIONS OF NEURAL NET INPUTS

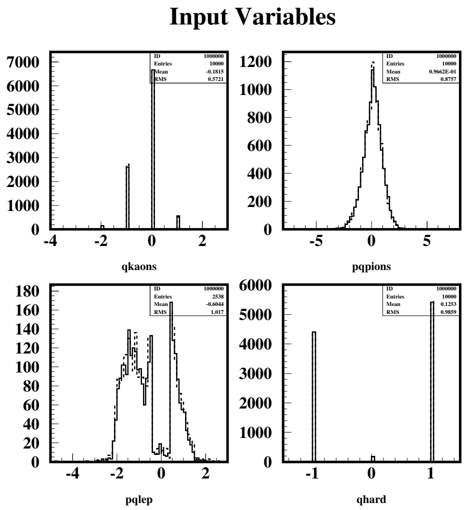


Figure C.1: Input variables for B(solid) and $\overline{B}$(dashed) Monte Carlo events. Clockwise from the upper left, the plots show the total charge of $K^{\pm}$ in each event; the sum over all pions in the event of $p_{\pi}Q_{\pi}$; the sum over all leptons of $p_{\ell}Q_{\ell}$; and the charge of the hardest track in the event.

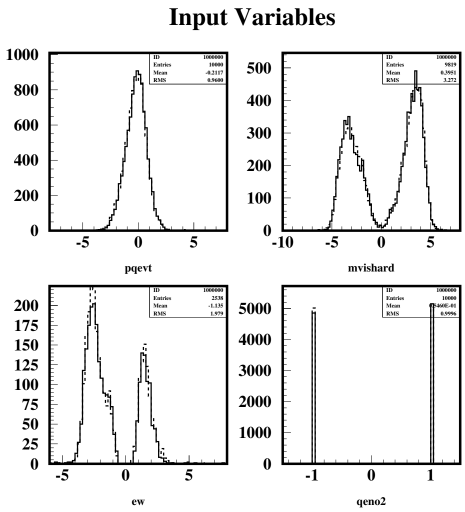


Figure C.2: Input variables for B(solid) and $\overline{B}$(dashed) Monte Carlo events. Clockwise from the upper left, the plots show the invariant mass of a K-*lepton*-p_{miss} combination; the invariant mass of the hardest and second hardest track; pq for the hardest track; and the total energy of K_S's in the event.

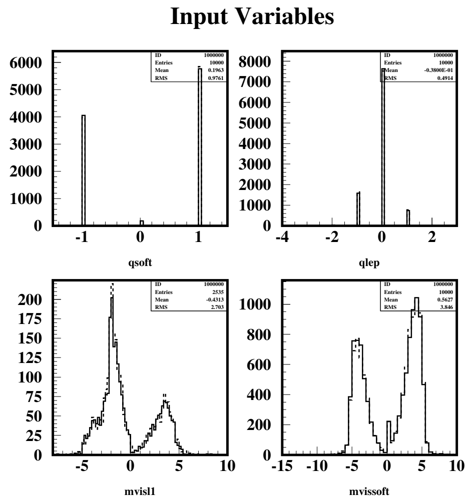


Figure C.3: Input variables for B(solid) and $\overline{B}$(dashed) Monte Carlo events. Clockwise from the upper left, the plots show the charge of the softest track in each event; the total charge of all leptons in the event; the visible recoil mass against a lepton, multiplied by the lepton charge; and the visible recoil mass against the softest track in an event, multiplied by the soft track's charge.

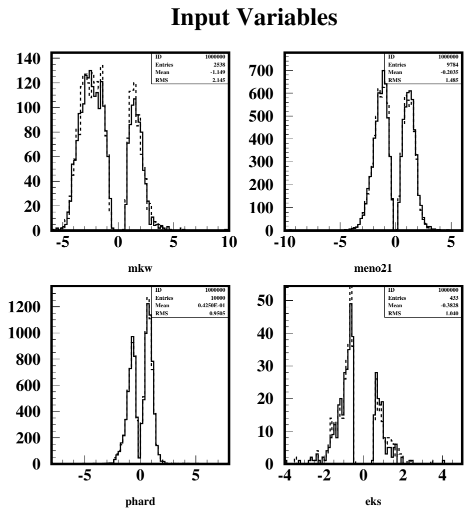


Figure C.4: Input variables for B(solid) and $\overline{B}$(dashed) Monte Carlo events. Clockwise from the upper left, the plots show the total charge of pq over all tracks in each event; the visible recoil mass against the softest track in an event, multiplied by the soft track's charge; the sum of lepton momentum and magnitude of the missing energy, multiplied by the lepton charge; and the charge of the second-hardest track in the event.

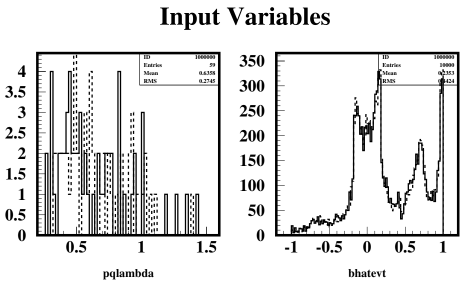


Figure C.5: Input variables for B(solid) and $\overline{B}$(dashed) Monte Carlo events. Clockwise from the upper left, the plots show the sum of the proton charge of Λ's, multiplied by the Λ momentum; and the dilution-based flavor estimator.

APPENDIX D

COMBINING FLAVOR TAGGING AND MIXING

D.0.2 Using the Flavor Tagger in a Mixing Measurement

We find that the leptons in the mixing analysis of [45], with a momentum cut of 1.4 GeV, have a reported dilution of 0.84. We can acheive the same dilution by selecting events with leptons of all momenta, while placing a requirement on the output of the flavor tagger. We find that by requiring event with leptons to have $|\hat{\mathcal{B}}| >$0.37, we acheive the same dilution of 0.84. This includes many leptons with momentum below 1.4 GeV. In fact, as shown in table D.1, we find that in our sample of single B^0 events, there are more than twice as many tagging events as before!

Table D.1: Comparison of two possible lepton tagging strategies: simple momentum cut, as in [45], and a dilution-based cut. As the table shows, this flavor-tagging method results in an efficiency increase of **2.2** while keeping purity fixed.

Tagging Requirement	Reported Dilution	Events (of 10000 B^0)
p >1.4 GeV	0.841	961
$\|B\|$ >0.37	0.841	2114

APPENDIX E

A GLOSSARY OF CLEO-SPECIFIC TERMS

E.1 A Glossary of CLEO-Specific Terms

`anlcp0`: a FORTRAN common block containing information about recostructed π^0's

`BETATF`: the measured velocity, in units of the speed of light c.

`CHIVTX`: the χ^2 of a K_S vertex fit

`CHSQ0`: the χ^2 of the consistency of a $\gamma\gamma$ pair with the π^0 mass

`cleog`: the CLEO version of GEANT, which is used to carefully model the detector

`DBCD`: the impact parameter at the point of closest approach to the beam axis.

`DE50DI`: the mean value of the lowest 70% of specific ionization measurements in the drift chamber. It is named DE50DI because it was historically defined as the lowest 50%.

`dedr`: the processor (see below) which measures the specific ionization of particles in the tracking volume, which is useful for particle identification.

`difix`: the post-processing tweak of the ionization measurement data

DPTHMU: the number of nuclear interaction lengths penetrated by a track, as detected by the muon chambers.

DUET: the tracking processor

FVSGNF: the significance ($\frac{l}{\sigma}$) of the flight distance of a K_S candidate.

hadsel: a routine common in CLEO to determine the probability of different particle identification hypotheses given measurements from time-of-flight and/or specific ionization.

IDV0: an identification code given to secondary vertices to identify γ conversions, Λ's, and K_S's.

KINCD: a track-quality flag. Well-measured tracks have KINCD greater than or equal to 0.

KLASGL: event classification. Event type 10 is hadronic annihilation, as determined by the number of tracks and calorimeter energy.

linktree: the pattern recognition algorithm for tracks. linktree is a part of DUET.

mn_fit: a fitting/maximum likelihood package commonly used at CLEO. It is basedon the popular MINUIT package from CERN.

PASS2: the program we run to process the raw data into physically meaningful quantities.

processor: A subprogram of PASS2 responsible for one modular physics task. Examples include dedr (ionization), duet (tracking), or evcl (event classification).

R2ELEC: An electron-identification variable which takes into consideration distributions of $\frac{E}{p}$ in the calorimeter, time-of-flight, ionization, and transverse energy deposition in the calorimeter.

`RBMTX`: the radius from the beam axis at which a K_S vertex is found.

`svrtoa`: a program which clusters raw silicon hits into clusters of analyzed hits.

`SVTF`: a processor which generates track segments, which may later be added to tracks.

`TFBNCF`: because CLEO runs in multibunch mode, the bunch number of a particular event must be determined in order to get the correct t_0's for the drift chamber. Once the determination has been made, `TFBNCF` is a measure of the confidence of the determiniation.

`tffix`: a post-pass2 tweak of time-of-flight information.

`tfxfit`: a main routine in the DUET tracking reconstruction.

`TNG`: the "Trackman" variable, which determines the overall quality of the track. Well-measured tracks have TNG greater than 0.

`TRIO`: a tracking processor which finds track seeds for DUET in pass2. TRIO is fast enough to be used for on-line tracking; DUET cannot run sufficiently fast for this.

`trksim`: a fast Monte Carlo used for simple detector and physics modeling.

`VETO_V0`: a routine which eliminates K_S candidates which may in fact be due to misidentified Λ decays or γ conversions.

`Z0CD`: the z-coordinate of the point of closest approach of a track to the z axis.

REFERENCES

[1] C.S. Wu, E. Ambler, R.W. Hayward, D.D. Hoppes, R.P. Hudson, *Phys. Rev.* **105**, 1413 (1957).

[2] J. H. Christenson, J. W. Cronin, V. L. Fitch, R. Turlay, *Phys. Rev. Lett* **13**, 138 (1964).

[3] A. D. Sakharov, *Sov. Phys. -JETP Lett.* **5**, 24 (1967).

[4] R. Feynman, "The Feynman Lectures on Physics, Vol. III", Addison-Wesley (1965).

[5] A. Alavi-Harati et al., hep-ex/9905060 (1999).

[6] Nelson, Nelson, and Witherell, CBX 95-8.

[7] N. Cabbibo, Phys. Rev. Lett **10** 531 (1963)

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

[9] R.M. Barnett *et al.* (Particle Data Group), Eur. Phys. J. **3**, 1-4 (1998).

[10] J.E. Christenson *et al.*, Phys. Rev. Lett. **13**, 138 (1964).

[11] *Babar Technical Design Report*, SLAC-R-95-457, Stanford, California (1995).

[12] *Physics Rationale for a B Factory"*, Lingel, *et al* (CESR B Working Group). Ithaca, NY (1993).

[13] *Belle Technical Design Report*, KEK Report 95-1, Tsukuba, Japan (1995).

[14] R. Aleksan, J. Bartelt, P.R. Burchat, and A. Seiden, Phys. Rev. D **39** 1283 (1989).

[15] K. Berkelman , Mod.Phys.Lett.A**10** 165 (1995).

[16] N. Sinha and R. Sinha, "CP Violation and Lifetime Differences of Neutral B Mesons from Correlated $B^0\overline{B^0}$ Pairs", hep-ph/9805501.

[17] I. Dunietz and J. Rosner, Phys.Rev.D34, 1404 (1986) .

[18] G.V. Dass, K.V.L. Sarma, Phys.Lett.B **421** 289 (1998) .

[19] P.F. Harrison, H.R Quinn, ed. *The Babar Physics Book*, 1998.

[20] A.B. Carter and A.I. Sanda, Phys.Rev.Lett. **45** 952 (1980).

[21] Bonvicini, et al. *Measurement of Charm Meson Lifetimes* Submitted to Phys. Rev. Lett.

[22] C. Quigg, "Gauge Theories of the Strong, Weak and Electromagnecit Interactions", Addison-Welsely (1993).

[23] M. Peskin, D. Schroeder, "An Introduction to Quantum Field Theory", Addison-Wesley (1997).

[24] L. Wolfenstein, *Phys. Rev. Lett.* **51**, 1945 (1983).

[25] Tim Nelson, Harry Nelson, Jeff Gronberg CBX 93-72. "Geometry of the Silicon Vertex Detector".

[26] Jim Alexander, *et al*, CBX 95-102. "The Silicon System Manual"

[27] Rob Kutschke, CSN 94/335. "DUET Flow Control".

[28] Rob Kutschke, CSN 94/334. "How and Why Wonder Book of DUET Conventions".

[29] Tony Hill, private communication.

[30] Tony Hill, private communication.

[31] Pablo Hopman, private communication.

[32] D. Dumas, "Retractable Radiation Shields for the CLEO II Silicon Vertex Detector", CBX 95-106.

[33] H. Yamamoto, private communication, from survey on 1/30/95.

[34] J. Cherwinka, private communication.

[35] D. Cinabro, private communication.

[36] J. Urheim, report to CLEO Collaboration, summarized in CBX 98-43, "Summary of the June CLEO PhysicsFest".

[37] A. Eppich, H. Nelson, "Survey Corrections to the Silicon Vertex Detector Geometry", CBX 95-94.

[38] A. Foland, "Studies of Tracking with the CLEO II.V Silicon Vertex Detector", CBX 98-4.

[39] S. Prell, V. Sharma, H. Paar, "Precise Measurements of Charm Meson Lifetimes", CBX 98-13.

[40] H. Albrecht, et al. Z.Phys.C. **62** 371 (1994).

[41] CESR B Factory Proposal

[42] F. Abe et al, CDF Collaboration, "Measurement of the $B^0_d - \bar{B}^0_d$ flavor oscillation frequency and study of same side flavor tagging of B mesons in $p\overline{p}$ collisions", hep-ex/9806026, Fermilab 98/188-E

[43] R. Waldi, Babar Note 204, "Flavor Tagging Studies for the Babar TDR".

[44] R. Waldi, "Lectures on Flavor Oscillation and CP Asymmetry in B Meson Decays", given at *Herbstschule Maria Laach*, TUD-IKTP/97-03,

[45] A. Foland, "Measuring B0- B0 Mixing Using Partial Hadronic Reconstruction ", CLEO CBX 98-14.

[46] F. Abe et al, CDF Collaboration, "Measurement of the $B^0 - \bar{B}^0$ oscillation frequency using $\pi - B$ meson charge-flavor correlations in $p\overline{p}$ collisions at $\sqrt{s}$=1.8 TeV" from CDF website

[47] S. Plaszczynski, "CP Violation with BaBar", LAL 98-16, talk at XIIth La Thuile Conference March 1998.

[48] R. Enomoto, "New Tagging Method of B Flavor of Neutral B Meson in CP Violation Measurement in Asymmetric B-Factory Experiment", Journal of the Physical Society of Japan, Vol 63 No. 10, Oct. 1994, pp. 3542-3545.

[49] J. Rodriguez et al, "Exclusive Hadronic $B \to D^{(*)}(n\pi)^-$ Decays", CBX 96-64

[50] C. J. Jessop et al, "Measurement of the Decay Amplitudes and Branching Fractions of $B \to J/\psi^*$ and $B \to \psi K$ Decays". CLNS 96/1455.

[51] H. Albrecht, et al. Z. Phys. C **58** 191 (1993).

[52] J. Ernst, CSN Note, "Dragnet: A Neural Network Facility"

[53] J. Ernst, Ph.D. Thesis, "A Measurement of $b \to s\gamma$."

[54] S. Versille, F. Le Diberder, "The Kin Variable", Babar Note #406.

[55] A. Foland, M. Dickson, J. Ernst, "A Flavor-Tagging Package for CLEO", CBX 99-19.

[56] A. Ershov, personal communication.

[57] A. Ershov, A. Foland, "Can CLEO Measure $\sin 2\beta$?", CBX 99-16.

[58] A. Foland, "Measurement of Mixing-Induced CP Violation at the $\Upsilon(4S)$ without Time Ordering or Δt" CLNS 99-1716.

[59] CLEO Collaboration, J. Bartelt *et al.* Phys. Rev. Lett. **71** (1993) 1680.

[60] Thorndike, Kroha and Wang CBX 92-76.

[61] Mike Saulnier, Ph.D. Thesis.

[62] Gronberg, Nelson CBX 96-69.

[63] Kinoshita, CBX 84-12.

[64] DELCO collaboration, R.M. Baltrusaitis *et al.* Phys. Rev. Lett. **54** (1985) 1976.

[65] Moneti, Xing, CLEO CBX 95-77.

[66] Burchat and Richman, Review of Particle Properties.

[67] CLEO Lepton SAC Report CBX 95-35.

[68] Nelson and Nelson, CBX 96-66.

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

[70] A. Foland, CBX 98-14.

[71] J. Rodriguez, personal communication to QQ czar; recorded in decay.dec.annotated

[72] CLEO-CONF 97-01, "Exclusive Reconstruction of B to D(*) n pi Decays".

[73] Tim Nelson, CBX 98-62, 96-66, CLEO-CONF 96-25; ρ B.F.'s come from multiplying by 2.78 as in decay.dec.annotated; B^0 B.F's from assuming factor of 0.62 as in decay.dec.annotated.

[74] L. Perera, CBX 98-47. “Angular Distributions in $B^0 \to D^{*-}$ and $B^+ \to D^{*0}\rho^+$.

[75] P. Rankin, K. Cho, CBX 94-43. “Seach for CP Violation in D^0 Decay”.

[76] T. Hill, personal communication; from studies done for $D^0 - \overline{D^0}$ mixing.

[77] C. Jessop, CBX 97-106. “Search for CP Violation in τ Decay”.

[78] B. Berger, CBX 98-35. “Slow Track Embedding Studies I”.

[79] M. Neubert, hep-ph/9705292, 1997.

[80] F. Parodi et al, hep-ph/9903063, 1999.

[81] Y. Liu and R. Patterson, “A Measurement of the $B \to D\ell\nu$ Branching Fraction and a Study of its Form Factor”, CBX 98-20.

[82] Y. Kwon, et al, CBX 97-65. “Lepton-Charm Correlations Revisited”.

[83] R. Barlow, C. Beeston, . *Comp. Phys. Comm.* **77** 219 (1 993).

[84] J. Alexander, et al, draft CBX, “Charge-Correlated Systematic Effects for High-Momentum Tracks”.

[85] S. Hashimoto, “Summary of lattice results for decay constants, mixing, etc.”, plenary talk given at Heavy Flavors 8, 1999.

[86] C. Bernard et al, hep-ph/9806412, 1998.

[87] M. Mangano, “Results and Perspectives in HEP Phenomenology, vis-a-vis Lattice”, plenary talk at Lattice 99, 1999.

[88] R. Patterson, “Heavy Flavor Decays”, Plenary talk given at DPF 1999.