UNIVERSITY of CALIFORNIA
Santa Barbara

Measurements of Hadronic B Decays to Excited Charm Mesons, Observation of a New Charm Resonance and Construction of a Silicon Vertex Detector for CLEO II.V

A dissertation submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy

in

Physics

by

Timothy Knight Nelson

Committee in charge:

Professor Harry N. Nelson, Chair
Professor Michael S. Witherell
Professor Joseph Polchinski

December 1998

The dissertation of Timothy Knight Nelson is approved:

Chair

December 1998

Measurements of Hadronic B Decays to Excited Charm Mesons,
Observation of a New Charm Resonance
and
Construction of a Silicon Vertex Detector for CLEO II.V

Copyright 1998

by

Timothy Knight Nelson

To Jim

For teaching me about the truth

To Molly

For her love and encouragement

To John

For sharing his humanity

Acknowledgements

I would like to thank Harry Nelson, my advisor, for the opportunities to work on these wonderful projects. It is hard to believe how much I have grown under his tutelage. I would also like to thank Mike Witherell and Rollin Morrison for their help in developing the reconstruction method and sorting out the spectroscopy of the D_J.

None of the SVX construction would have been possible without the terrific team of Kirk Arndt and Susanne Kyre. They both deserve credit for their expertise and for putting up with all of our nonsense. I would also like to thank the SVX team at Cornell, especially Tony Hill and Mark Palmer, for all of the hard work that went into their plots of the detector performance shown in chapter eight.

Thanks go to Jeff Gronberg for helping keep me on the straight-and-narrow, Dave Hale for throwing me off it from time-to-time, and Chris Korte for being there during both. I would especially like to thank John Staren for helping me to keep things in perspective. Go, John!

Thanks also to all of those who have supported me in ways large and small during the completion of this work, in particular: David Lange, Scott Menary, Rob Kutschke, Roy Briere, David Asner, Debbie Ceder, Roberta Milatello and Judy Young.

Finally, I must thank my wife, Molly Young, and my family, James, Elizabeth and Sarah Nelson, for standing by me in everything I do.

CURRICULUM VITAE

November 18, 1967 — Born — Dayton, Ohio

1989 — B.A. Physics with Honors, University of Chicago

1993 — M.S. Physics, University of California at Santa Barbara

1998 — Ph.D. Physics, University of California at Santa Barbara

PUBLICATIONS

"Studies of Double Sided Double Metal Silicon Strip Detectors." J. Alexander, *et al.*, Nucl. Inst. Meth. **A342** (1994) 282-286.

"Measurement of $BR\left(B^- \to D_J^0\,\pi^-\right)$ Using Partial Reconstruction." J. Gronberg *et al.*, CLEO CONF 96-25. Presented at the 28th International Conference on High Energy Physics, Warsaw, 1996.

"Measurements of $B^- \to D_J^0\,\pi^-$". CLEO Collaboration. Work in progress.

Abstract

Measurements of Hadronic B Decays to Excited Charm Mesons, Observation of a New Charm Resonance

and

Construction of a Silicon Vertex Detector for CLEO II.V

by

Timothy Knight Nelson

We describe measurements of the branching ratios

$$\mathcal{B}\left(B^- \to D^{\star +}\,\pi^-\,\pi^-\text{total}\right) = (29.2 \pm 4.5 \pm 3.8 \pm 3.1)\times 10^{-4}$$

$$\mathcal{B}\left(B^- \to D^{\star +}\,\pi^-\,\pi^-\text{non-res}\right) = (9.7 \pm 3.6 \pm 1.5 \pm 1.9)\times 10^{-4}$$

$$\begin{aligned}
\mathcal{B}\left(B^- \to D_1(2420)^0\,\pi^-\right)\mathcal{B}\left(D_1(2420)^0 \to D^{\star +}\,\pi^-\right) &= (6.9\,^{+1.8}_{-1.4} \pm 1.1 \pm 0.4)\times 10^{-4}\\
\mathcal{B}\left(B^- \to D_1^0(j=\tfrac{1}{2})\,\pi^-\right)\mathcal{B}\left(D_1^0(j=\tfrac{1}{2}) \to D^{\star +}\,\pi^-\right) &= (10.6 \pm 1.9 \pm 1.7 \pm 2.3)\times 10^{-4}\\
\mathcal{B}\left(B^- \to D_2^\star(2460)^0\,\pi^-\right)\mathcal{B}\left(D_2^\star(2460)^0 \to D^{\star +}\,\pi^-\right) &= (3.1 \pm 0.84 \pm 0.46 \pm 0.28)\times 10^{-4},
\end{aligned}$$

using data collected by the CLEO II detector. These measurements provide the first observation of the $D_1^0(j = \frac{1}{2})$ with a mass and width of $2.461^{+0.053}_{-0.049}$ GeV and 290^{+110}_{-91} MeV respectively. The mixing angles between the partial waves and strong phase shifts among the resonances are also measured assuming one possible parameterization of the amplitude. A method allowing full reconstruction of the signal without reconstruction of the D meson in the final state is used. The measurements are extracted using an four-dimensional, unbinned, maximum-likelihood fit to the distributions of the $D^{\star +}\pi^-$ mass and the decay angles.

The primary element of the CLEO II.V upgrade was the installation of a three-layer Silicon Vertexing Detector. The design and construction of this detector are described in detail.

Contents

List of Figures

List of Tables

Part I

Measurements of Hadronic B Decays to Excited Charm Mesons and Observation of a New Charm Resonance

Chapter 1

Introduction

1.1 The Standard Model

The standard model provides a description of the elementary nature of matter that has been remarkably flexible in accounting for several decades of experimental discoveries. The theory resulting from this evolution is self-consistent, but fundamental properties have no explanation. These include the relationships between quarks and leptons, and between the strong, electroweak and gravitational interactions. This consistency without explanation has made it unattractive as a 'theory of everything.'

The success of the standard model makes the development of a more unified theory a daunting task. There are no spectacular failures to focus the search and a large number of these theories match the limited experimental evidence available. None of the alternatives can equal the predictive power of the standard model after decades of refinement.

This gives the experimentalist two tasks, searching for inconsistencies with the standard model and further refining our knowledge of the pieces we are trying to unite. We must first decide where we are likely to encounter a contradiction and where more data will differentiate among the candidates. One of these frontiers is the study of how the quarks couple to the weak interaction. Another is the study of how the strong interaction determines the spectra and decays of the observed bound states of quarks.

1.1.1 The Weak Interaction

One of the puzzles left unanswered by the standard model is the relationship between leptons and quarks. There are six of each arranged in three generations:

$$\begin{array}{ccc} \text{Leptons} & \qquad & \text{Quarks} \\ \begin{pmatrix} e & \nu_e \\ \mu & \nu_\mu \\ \tau & \nu_\tau \end{pmatrix} & \qquad & \begin{pmatrix} u & d \\ c & s \\ t & b \end{pmatrix} \end{array}$$

The standard model offers no explanation for this apparent symmetry. Adding to this mystery are the striking differences between their properties. Quarks interact strongly while the leptons lack any strong interactions. All of the quarks have mass and fractional charge while the leptons either have mass and unit charge or (apparently) neither property. Finally, leptons and quarks have very different couplings to the weak interaction: the mass eigenstates of the quarks are not the weak eigenstates, leading to weak decays between generations. This fact is incorporated into the standard model by replacing the charge -1/3 quarks, (d,s,b), by mixed quarks in the weak coupling:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.1.1}$$

The matrix V_{ij} is the Cabbibo-Kobayashi-Maskawa (CKM) mixing matrix [1]. Assuming unitarity of this mixing, the matrix can be uniquely described by four parameters — three real angles and a complex phase. The angles are essentially mixing angles between the quark generations. A non-zero phase gives rise to CP violation. This is the only possible explanation for CP violation in the standard model.

The relationships resulting from unitarity are experimentally testable. One such expression is derived from application of the unitarity constraints to the first and third columns:

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \tag{1.1.2}$$

To simplest approximation, the CKM matrix is the identity matrix, allowing the simplification

$$V_{ub}^* + V_{cd}V_{cb}^* + V_{td} = 0. \tag{1.1.3}$$

This can be visualized as a triangle in the complex plane, as shown in Figure 1.1. Standard model CP-violating effects are proportional to the area of this triangle.

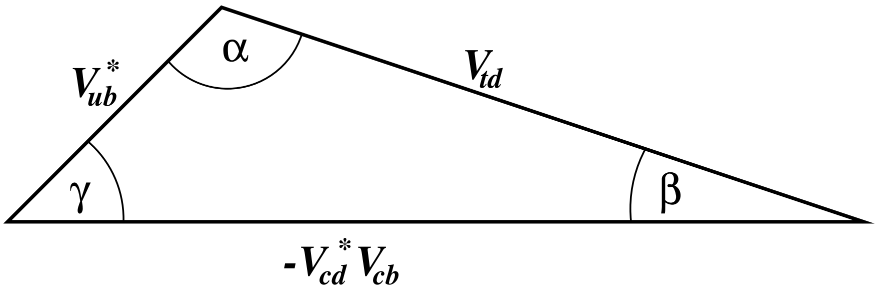


Figure 1.1: One of the triangles obtained by application of the unitarity constraint to the CKM matrix. The relationships between sides and angles enforced by closure of the triangle mirror those between the magnitudes and complex phases of the CKM matrix enforced by unitarity.

We can overconstrain this triangle by independently measuring both the lengths of its sides and the angles between them. If these measurements are consistent, then we have strong confirmation of the CKM matrix as the source for CP violation. If not, other mechanisms outside of the standard model must be invoked in order to explain the data. Discovery of such an inconsistency would produce a revolution in elementary particle physics and help identify the correct route to a more fundamental theory.

Since V_{cd} is already well determined, the measurements of the sides we still need are

- V_{cb} — from $b \to cl\nu_l$ decays
- V_{ub} — from $b \to ul\nu_l$ decays
- V_{td} — from B^0-$\overline{B^0}$ mixing.

Quark-level diagrams for these decays are shown in Figure 1.2. To determine the angles of this triangle, we need

- α — asymmetries in $b \to u\bar{u}d$ from B^0 decays
- β — asymmetries in $b \to c\bar{c}d$ from B^0 decays
- γ — asymmetries in $b \to u\bar{u}d$ from B_s^0 decays.

The use of semileptonic decays simplifies the extraction of V_{cb} and V_{ub} since there are no interactions between the leptons and quarks in the final state. As the largest single branching ratio of the B meson, the decay $B \to D^\star\, l\, \nu_l$ provides the most sensitive measurement of V_{cb}. This decay mode is also important to measurements of B^0-$\overline{B^0}$ mixing(V_{td}) and decay asymmetries in B^0 decays (α, β). One must 'tag' the species of one B^0 while reconstructing the decay of the other. Measuring the sign of the primary lepton from $\overline{B^0} \to D^{\star +}\ell^- \overline{\nu}$ is the simplest and single most efficient tagging method because of the large branching ratio of this mode ($4.60 \pm 0.27\%$ [2]).

Only the decays to the $L = 0$ charm mesons, the D and the $D^\star$, are well measured. The contributions from decays to the $L = 1$ charm mesons, the D_J, are a significant uncertainty in interpretation of $B \to D^\star\, l\, \nu_l$ decays in measuring of V_{cb} and tagging B mesons [3, 4]. As a result, improved measurements of $B \to D_J\, l\, \nu_l$ will be important for reducing the errors on four of the six parameters of the unitarity triangle; V_{cb}, V_{td}, α and β.

Unfortunately, these decays are very difficult to reconstruct. With smaller branching ratios than $B \to D^\star\, l\, \nu_l$ and a lack of kinematic constraints due to the undetected neutrino, it is difficult to obtain a significant signal. It is simpler to measure the hadronic decay, $B \to D_J\, \pi$, where the virtual W decays to a pair of quarks rather than leptons as shown in Figure 1.3. Without the missing neutrino to contend with, there are considerably more constraints.

The primary difficulty in this approach is converting a measurement of $B \to D_J\, \pi$ into meaningful information about $B \to D_J\, l\, \nu_l$. The interactions between the quarks in the final state and the possibility of internal spectator decays complicate the hadronic decays considerably. However, the high q^2 of this decay may allow us to ignore some of these effects. If we assume no final state interactions due to the small time-integrated overlap of the wavefunctions of the outgoing mesons, we have the relationship

$$\frac{\Gamma(B^- \to D_J^0\, \pi^-)}{\frac{d\Gamma}{dq^2}(B^- \to D_J^0\, \ell^-\, \overline{\nu}_\ell)|_{q^2=m_\pi^2}} = 6\pi^2 f_\pi^2 |V_{ud}|^2 \frac{\Gamma(B^- \to D_J^0\, \pi^-)}{\Gamma(\overline{B^0} \to D_J^+\, \pi^-)}. \tag{1.1.4}$$

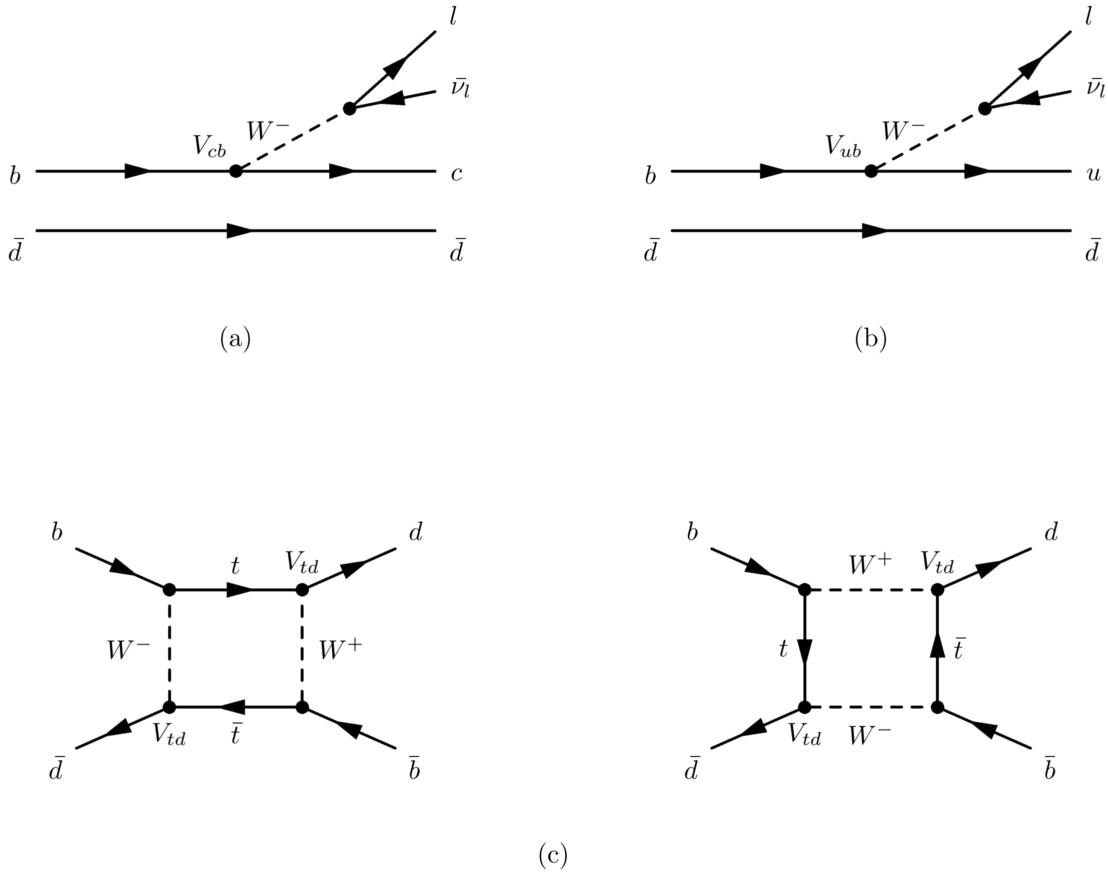


Figure 1.2: Quark level diagrams of the decays used to determine the magnitudes of (a) V_{cb}, (b) V_{ub} and (c) V_{td}. The semileptonic decays, (a) and (b) are chosen to avoid strong interactions between the mesons in the final state. While both u and c quarks also contribute to the box diagrams (c), the top quark diagram dominates.

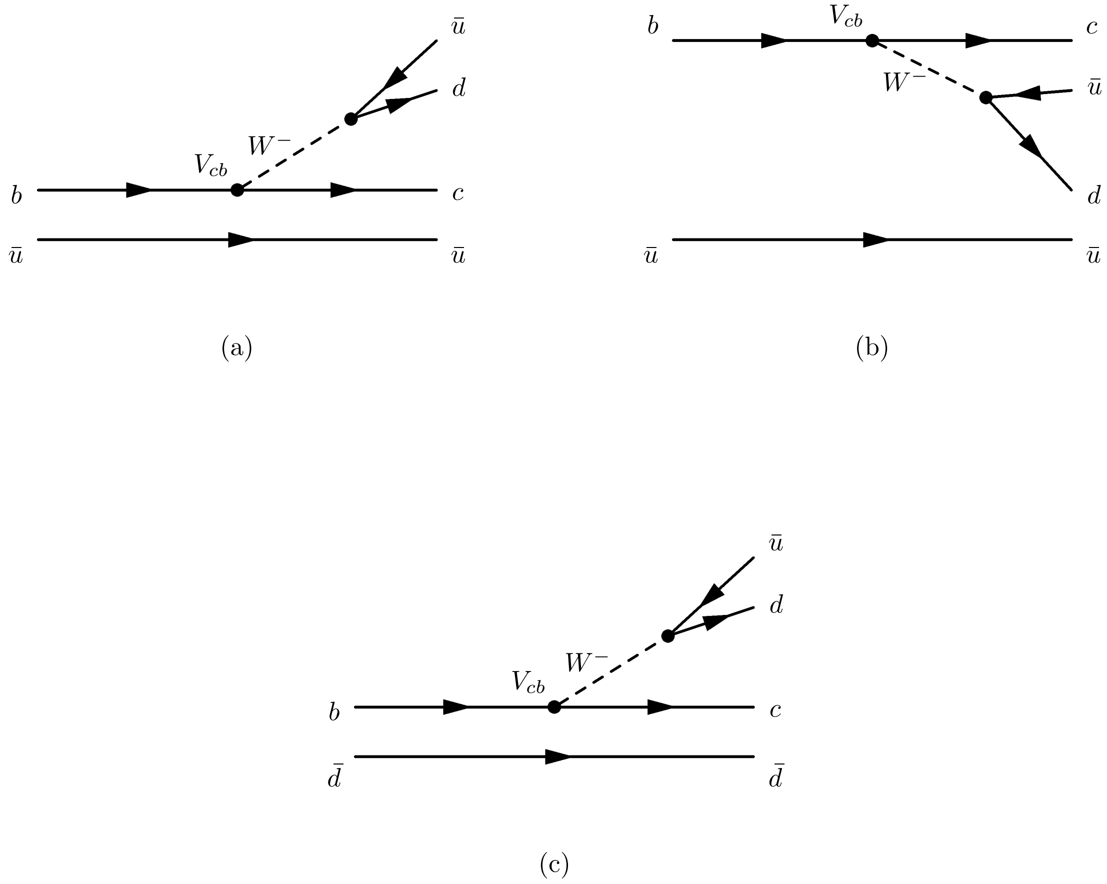


Figure 1.3: Quark level diagrams of hadronic B decays. The decay $B^- \to D_J^0\,\pi^-$ can proceed through both external (a) and internal (b) spectator diagrams with competing amplitudes a_1 and a_2. Only the external spectator decay is possible for $\overline{B^0} \to D_J^+\,\pi^-$ (c).

This type of approximation has worked well for other B decays with q^2 similar to $B^- \to D_J^0\, \pi^-$ [5]. Thus, careful measurements of $B \to D_J\, \pi$ decays may be very useful in the large scale effort to measure the unitarity triangle.

1.1.2 The Strong Interaction

At low energies, the strong coupling constant, α_s, is larger than unity. This explains the confinement of quarks and gluons inside hadrons and makes calculating the properties of those hadrons difficult since higher order corrections in QCD contribute as powers of α_s. This limit of QCD is the least understood part of strong interaction phenomenology. It is reasonable to hope that a better understanding of these bound states will lead to a better understanding of the strong interaction and help us unite QCD with the electroweak theory.

In order to make predictions in this regime, there are some approximate symmetries we may employ. The classic example is the isospin symmetry between the u and d quarks under the assumption that $m_u - m_d << \Lambda_{QCD}$. Recently, it has been recognized that there is another approximate symmetry for the heaviest quarks, Q, under the assumption that $m_Q >> \Lambda_{QCD}$ [6, 7]. This symmetry gives rise to heavy quark effective theory (HQET) which makes many predictions about the properties and decays of hadrons with one heavy quark by expanding in corrections around this limit. These include predictions about hadronic decays of the B and the spectroscopy of the D_J. As will be shown below, detailed measurements of the decay sequence $B \to D_J\, \pi, D_J \to D^\star\, \pi, D^\star \to D\, \pi$ can test these predictions, providing useful information about the strong interaction.

1.1.3 D_J Spectroscopy

The simplest D_J mesons consist of a light quark and a charm quark in a relative (orbital angular momentum) $L = 1$ state. The total parity is positive, because it is the product of negative relative parity of the quark and antiquark, and the negative parity of the $L = 1$ state. The $L = 1$ combines with the total spin of $S = 0$ or 1 to produce a set of four states. In the limit that $m_c \ll \Lambda_{QCD}$, we expect a singlet with $J = 0$ and $S = 0$, and a triplet with $J = 0, 1, 2$ and $S = 1$ as shown on the left of Figure 1.4. The splitting between the triplet from the singlet would be caused by a spin-spin interaction, and the splitting among the triplet levels by a spin-orbit interaction. The actual magnitude of the splitting between the singlet and triplet would probably be less than illustrated, because spin-spin interactions are suppressed in an $L\!=\!1$ state, at least relative to an $L\!=\!0$ state.

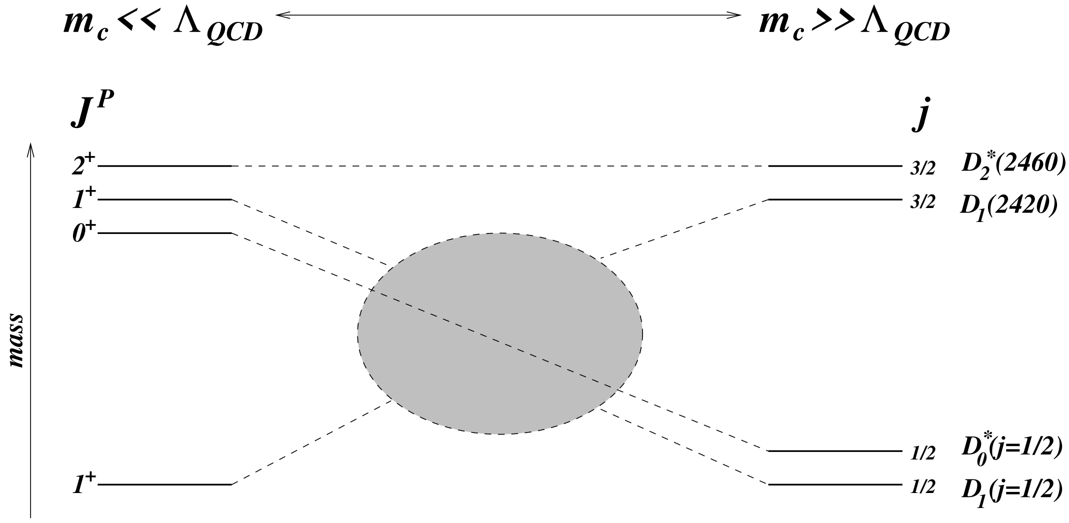


Figure 1.4: The spectroscopy of the $L = 1$ D_J mesons.

However, the limit $m_c \ll \Lambda_{QCD}$ is a poor approximation: the limit of heavy quark symmetry, $m_c \gg \Lambda_{QCD}$, should better describe these mesons. In this limit, the spin of the heavy quark decouples from the other angular momenta, and we expect the two doublets shown on the right in Figure 1.4. One doublet has $j\!=\!\frac{1}{2}$ and the other $j\!=\!\frac{3}{2}$, where j is the total angular momentum excluding the spin of the heavy quark. Near this limit, the splitting between the two doublets is caused by a spin-orbit interaction of the light quark, and the splitting between the two levels in the doublet is caused by a spin-spin interaction.

The $J = 0$ state must have $j = \frac{1}{2}$ and the $J = 2$ state, $j = \frac{3}{2}$. However, the two $J^P\!=\!1^+$ states may be a mixture of $j\!=\!\frac{1}{2}$ and $j\!=\!\frac{3}{2}$. The details of this mixing probe the breaking of the heavy quark symmetry due to the finite mass of the charm quark [7, 8].

Conservation of parity and angular momentum restrict the final states and partial waves that are involved in the decays of the various D_J mesons. Table 1.1 summarizes the possibilities for the $L = 1$ D_J^0 mesons. While the 2^+ may decay via a **d**-wave alone, the 1^+ states may decay via either a **d**-wave or an **s**-wave. In the limit of heavy quark symmetry, the $j = \frac{3}{2}$ decays via a **d**-wave alone, and is therefore a narrow resonance; while the $j = \frac{1}{2}$ decays via the **s**-wave alone, and is therefore a broad resonance. Our four-dimensional fit, which constitutes a partial-wave analysis of the D_J^0 decays, provides a measurement of the HQET-breaking mixing between the 1^+ states, as manifested by the presence of the **s**-wave in the decay distribution of the narrow 1^+ state, and the presence of a much smaller amount of **d**-wave in the decay distribution of the broad 1^+ state.

Table 1.1: The $L = 1$ D_J^0 mesons and their allowed decays. Neither the $D_0^\star(j=\frac{1}{2})^0$ nor the $D_1^0(j=\frac{1}{2})$ have been observed, prior to this work. In the limit of heavy quark symmetry, the $D_1(2420)^0$ decays only by a **d**-wave, and $D_1^0(j=\frac{1}{2})$ only by an **s**-wave. Masses and widths here are taken from the 1998 Review of Particle Physics [2].

State	J^P	Mass (MeV)	Width (MeV)	Allowed decays	Partial waves	HQET allowed
$D_0^\star(j=\frac{1}{2})^0$	0^+	—	—	$D\pi$	**s**	**s**
$D_1(2420)^0$	1^+	2422.2 ± 1.8	$18.9^{+4.6}_{-3.5}$	$D^\star\pi$	**s**,**d**	**d**
$D_1^0(j=\frac{1}{2})$	1^+	—	—	$D^\star\pi$	**s**,**d**	**s**
$D_2^\star(2460)^0$	2^+	2458.9 ± 2.0	23 ± 5	$D^\star\pi$, $D\pi$	**d**	**d**

From Table 1.1, the three states $D_1(2420)^0$, $D_1^0(j=\frac{1}{2})$, and $D_2^\star(2460)^0$ are able to decay to $D^{\star+}\pi^-$. These are the three states that we characterize with the four-dimensional fit to the variables describing $B^- \to D^{\star+}\pi^-\pi^-$ decay.

1.1.4 $B^- \to D_J^0\,\pi^-$ Decay Amplitudes

The full description of the $B^- \to D^{\star+}\pi^-\pi^-$ decay is rather involved, and we make the description in five steps. First, we start with a review of the part of the angular distribution that has been used until now to extract information on the D_J: the helicity angle of the $D^{\star+} \to D^0\,\pi^+$ decay. Second, we incorporate the additional two decay angles that describe the $B^- \to D^{\star+}\pi^-\pi^-$ decay. Third, we write down the 'grand amplitude' in Eq. 1.1.23, which includes contributions from the various resonances, and the non-resonant decays. Fourth, we discuss the mixing of the **s**-wave and **d**-wave contributions to the decays of the $J^P = 1^+$ states, and the complex phases. Last, we address the complex-valued resonant amplitudes, described by Breit-Wigner functions.

The angle in the $D^{\star+}$ rest frame of the D^0 with respect to the direction of the boost that puts the $D^{\star+}$ back into the D_J^0 rest frame is denoted here by θ_3, and is portrayed in Fig. 1.5. The angle θ_3 can be used to analyze the partial wave content of the interaction

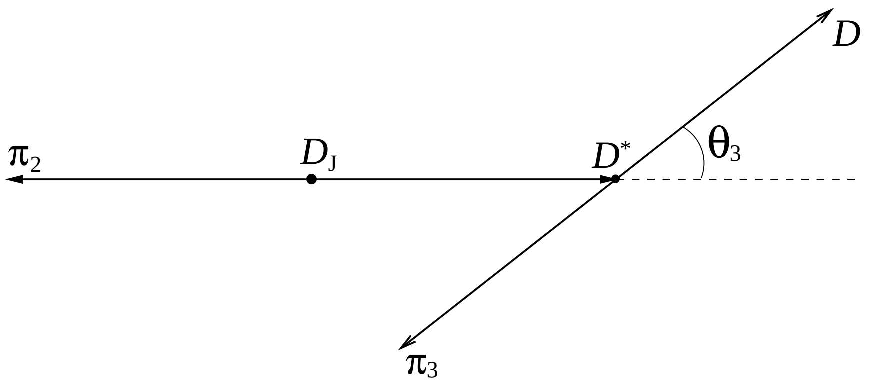


Figure 1.5: Portrayal of the D_J decay, and the variable θ_3. The D_J decays to $D^\star\pi$, with the resulting pion denoted here by π_2. The helicity angle θ_3 is defined in the rest frame of the $D^\star$, as the angle between the boost that would take the $D^\star$ back into the D_J rest frame, and the direction of the daughter D; the pion from the $D^\star \to D\pi$ decay is denoted here by π_3. The angle θ_3 can be used to deduce the partial waves that contribute to the D_J decay, *independent* of the initial polarization state of the D_J.

that mediates the D_J^0 decay, *independent* of the initial state of the D_J^0 (for those D_J^0 that are able to decay to $D^\star\pi$). The expected distributions for the various cases are [9]

$$\frac{dN}{d\cos\theta_3} \propto \begin{cases} \sin^2\theta_3 & (\text{ pure } 2^+ \ \mathbf{d}\text{-wave decay }) \\ 1+3\cos^2\theta_3 & (\text{ pure } 1^+ \ \mathbf{d}\text{-wave decay }) \\ 1 & (\text{ pure } 1^+ \ \mathbf{s}\text{-wave decay }). \end{cases} \tag{1.1.5}$$

The variable θ_3 provides information on the partial wave of the decay because the direction of the $D^{\star+}$, in the D_J^0 rest frame, must be orthogonal to the direction of the orbital angular momentum of the partial wave. Then, the spin of the $D^{\star+}$ must combine with the orbital angular momentum of the partial wave to add up to the D_J total J. So, for example, in the **d**-wave decay of the 2^+, the helicity-0 state of the $D^{\star+}$ is forbidden, because the helicity-0 state would spoil angular momentum conservation; a distribution in $D^{\star+}$ helicity angle characteristic of the helicity ±1 states, $\sin^2\theta_3$ results. For the **s**-wave decay, the $D^{\star+}$ carries off the angular momentum of the D_J^0, but since the **s**-wave is uniformly distributed, no correlation between that angular momentum and the direction of the D_J^0 survives, yielding the flat distribution. For the **d**-wave decay of the 1^+ state, a few Clebsch-Gordan coefficients can be used to obtain the $1+3\cos^2\theta_3$ distribution, which is very nearly equivalent to a pure

$\cos^2\theta_3$ distribution, characteristic of a helicity-0 $D^{\star +}$ decay.

The $D_1(2420)^0$ has been observed with a width and distribution in θ_3 consistent with the **d**-wave decay of the 1^+ state [9, 10, 11, 12]. Thus, the $D_1(2420)^0$ has been identified as predominantly $j = \frac{3}{2}$ and the implication is that the other 1^+ state will be mostly $j = \frac{1}{2}$. The expected **s**-wave decay of the $j = \frac{1}{2}$ state suggests a width exceeding 100 MeV, and a flat distribution in $\cos\theta_3$. The flat distribution makes detection of the $D_1^0(j = \frac{1}{2})$ a challenge when the polarization of the D_J is unconstrained.

More information is available in the angular distributions of D_J^0 mesons produced by the two body decay $B^- \to D_J^0\,\pi^-$. Because the B and π are both pseudoscalars, the D_J^0 mesons are produced in a helicity-0 state. This constraint on the spin-density matrix of the D_J^0 states provides two more variables that distinguish among the D_J^0 and the partial waves that mediate their decays. The first, which we shall call θ_2, is analogous to the helicity angle of the $D^{\star +}$ decay, but for the D_J^0: θ_2 is the polar angle in the D_J^0 rest frame between the boost back to the B rest frame, and the $D^{\star +}$ direction. The second is the angle between the plane defined by the $D_J \to D^\star\,\pi$ decay and the similar plane for the $D^{\star +} \to D^0\,\pi^+$ decay. This angle, χ, is the difference between the azimuthal angles of the decays, $\phi_2 - \phi_3$. The relationships among the two body decays and these variables are shown in Figure 1.6.

We calculate the amplitude, as a function of the decay angles, that results from the decay of the helicity-0 D_J^0 through each partial wave by use of the helicity formalism [13, 14]:

$$\mathcal{A}_J(\Omega_2, \Omega_3) \propto \sum_{\lambda_{D^\star}} D^{J\star}_{0\lambda_{D^\star}}(\Omega_2) A_{\lambda_{D^\star}\,0} D^{1\star}_{\lambda_{D^\star}\,0}(\Omega_3), \tag{1.1.6}$$

where

$$\begin{aligned} \Omega_2 &= (\theta_2, \phi_2) && (1.1.7)\\ \Omega_3 &= (\theta_3, \phi_3), && (1.1.8) \end{aligned}$$

and the sum is over the $D^\star$ helicity states, $\lambda_{D^\star} = \{-1, 0, 1\}$. The functions $D^j_{m'm}(\Omega_i)$ are given by

$$D^j_{m'm}(\Omega_i) = e^{-i\phi_i m'} d^j_{m'm}(\theta_i) e^{i\phi_i m}, \tag{1.1.9}$$

where the $d^j_{m'm}(\theta_i)$ are the usual d-functions [2]. The $A_{\lambda_{D^\star}\,0}$ come from the Jacob-Wick transformations that allow one to translate from the partial wave basis into the helicity basis of the D functions [15]. For decays of the D_1 we have

$$D^{1\star}_{0-1}(\Omega_2) D^{1\star}_{-10}(\Omega_3) \quad = \quad -\frac{1}{2} e^{i\chi} \sin\theta_2 \sin\theta_3 \tag{1.1.10}$$

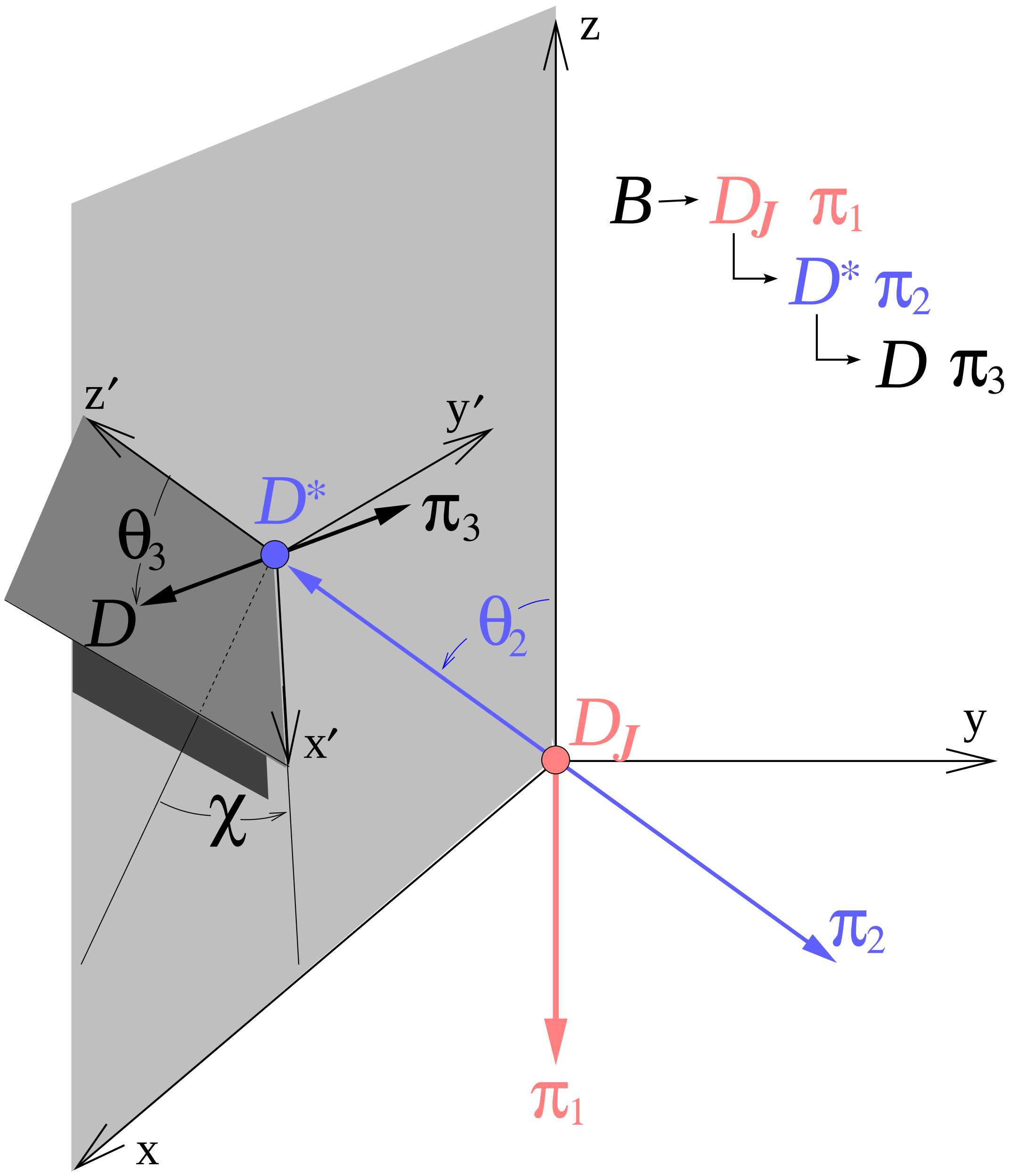


Figure 1.6: Decay angles in the two body decay sequence $B \to D_J\,\pi$, $D_J \to D^\star\,\pi$, $D^\star \to D\,\pi$. The polar decay angles θ_2 and θ_3 are defined in the rest frames of the D_J and $D^\star$ respectively. This diagram is portrayed in the rest frame of the D_J.

$$D^{1\star}_{00}(\Omega_2)D^{1\star}_{00}(\Omega_3) = \cos\theta_2\cos\theta_3 \tag{1.1.11}$$

$$D^{1\star}_{01}(\Omega_2)D^{1\star}_{10}(\Omega_3) = -\frac{1}{2}e^{-i\chi}\sin\theta_2\sin\theta_3. \tag{1.1.12}$$

The Jacob-Wick transformations for decay from a $J^P = 1^+$ state are

$$A_{10} = A_{-10} = \sqrt{\frac{1}{3}}\,|s\rangle + \sqrt{\frac{1}{6}}\,|d\rangle \tag{1.1.13}$$

$$A_{00} = \sqrt{\frac{1}{3}}\,|s\rangle - \sqrt{\frac{2}{3}}\,|d\rangle\,, \tag{1.1.14}$$

where **s** and **d** are the amplitudes of the **s**-wave and **d**-wave contributions to the decay. Insertion of equations (6-10) into equation (2) gives the decay amplitude of a 1^+ in terms of the **s**-wave and **d**-wave components:

$$\mathcal{A}_J(\Omega_2,\Omega_3) \propto a_{1s}(\cos\theta_2,\cos\theta_3,\chi)\,|s\rangle + a_{1d}(\cos\theta_2,\cos\theta_3,\chi)\,|d\rangle\,. \tag{1.1.15}$$

Thus, the **s**-wave and **d**-wave angular amplitudes for the decay of a 1^+ state are (ignoring leading constants)

$$\begin{aligned} a_{1s}(\cos\theta_2,\cos\theta_3,\chi) &= \cos\theta_2\cos\theta_3 - \sin\theta_2\sin\theta_3\cos\chi \\ a_{1d}(\cos\theta_2,\cos\theta_3,\chi) &= -2\cos\theta_2\cos\theta_3 - \sin\theta_2\sin\theta_3\cos\chi. \end{aligned} \tag{1.1.16}$$

The relative $(-)$ sign between a_{1s} and a_{1d} when $\theta_2 = \theta_3 = 0$ is a result of the conventions of the helicity formalism. The results of the fit to the data, described in Table 5.2, in effect reverse this convention, through a strong phase $\delta_1 \approx -\pi$, so the data favor constructive interference between the two partial waves when $\theta_2 = \theta_3 = 0$. Recovery of two of the angular distributions of Eq. 1.1.5 can by obtained by integrating the appropriate $|a|^2$ over θ_2 and ϕ_2.

For the 2^+ D_J meson decay we have

$$D^{2\star}_{0-1}(\Omega_2)D^{1\star}_{-10}(\Omega_3) = -\frac{\sqrt{3}}{2}e^{i\chi}\sin\theta_2\cos\theta_2\sin\theta_3 \tag{1.1.17}$$

$$D^{2\star}_{00}(\Omega_2)D^{1\star}_{00}(\Omega_3) = \frac{1}{2}\left(3\cos^2\theta_2 - 1\right)\cos\theta_3 \tag{1.1.18}$$

$$D^{2\star}_{01}(\Omega_2)D^{1\star}_{10}(\Omega_3) = -\frac{\sqrt{3}}{2}e^{-i\chi}\sin\theta_2\cos\theta_2\sin\theta_3. \tag{1.1.19}$$

The Jacob-Wick transformations for the 2^+ decay are

$$A_{10} = -A_{-10} = -\sqrt{\frac{1}{2}}\,|d\rangle \tag{1.1.20}$$

$$A_{00} = 0. \tag{1.1.21}$$

So for the angular amplitude of the 2^+ decay, defined in a manner analogous to the angular amplitudes a_{1s} and a_{1d}, we have (ignoring leading constants)

$$a_2(\cos\theta_2, \cos\theta_3, \chi) = -i \sin\theta_2 \cos\theta_2 \sin\theta_3 \sin\chi. \tag{1.1.22}$$

The leading factor of $-i$ is also the result of the conventions of the helicity formalism; in the fit to the data, described in Table 5.2, the strong phase δ_2 takes a value that nearly cancels this $-i$. Integration over θ_2 and ϕ_2 gives the $\sin^2\theta_3$ distribution of Eq. 1.1.5.

Figure 1.7 shows the complete set of two-dimensional projections of the resulting angular distributions for Monte Carlo simulated signal. These plots include the effects of acceptance and smearing in our reconstruction method. The area of each square is proportional to the number of reconstructed and accepted events in that bin.

Before squaring these amplitudes to obtain the angular distributions for the fit, we must multiply by the appropriate, complex Breit-Wigner amplitudes. We denote these as A_{1n}, A_{1b} and A_2 for the narrow 1^+, broad 1^+ and 2^+ states respectively. We adopt the convention that there is no strong phase associated with the narrow 1^+ resonance, and then allow a strong phase δ_1 for the broad 1^+, and additional strong phases δ_2 and δ_n for the 2^+ and non-resonant components. The complete expression for the 'grand' decay amplitude $\mathcal{A}$ is then

$$\begin{aligned} \mathcal{A} \quad = \quad & \alpha_{1n}\, A_{1n}\, \left(a_{1d}\cos\beta + a_{1s}e^{i\phi}\sin\beta\right) \\ + \quad & \alpha_{1b}\, A_{1b}\, \left(a_{1s}\cos\beta - a_{1d}e^{i\phi}\sin\beta\right)\, e^{i\delta_1} \\ + \quad & \alpha_2\, A_2\, a_2\, e^{i\delta_2} \\ + \quad & \alpha_{non-res}\, e^{i\delta_n}. \end{aligned} \tag{1.1.23}$$

The coefficients α characterize the amount of amplitude that each resonance contributes. In the case of no interference, the yield that can be attributed to each component is proportional to the respective α_i^2. The angles β and ϕ (both real) characterize mixing between the $J^P = 1^+$ states.

We can also parameterize the mixing through the off-diagonal terms in the mass and decay matrices as follows. Call state #1 that superposition of $J^P = 1^+$ states which may directly decay via only a **d**-wave intermediate state to $D^{\star+}\pi^-$, and state #2 that superposition of 1^+ states which may directly decay via only an **s**-wave intermediate state to $D^{\star+}\pi^-$. Write the representation of the 2 by 2 hamiltonian of the system as:

$$H = (\overline{M} - \frac{i}{2}\overline{\Gamma})\mathbf{I} + \frac{1}{2}(\Delta M - \frac{i}{2}\Delta\Gamma)\sigma_{\mathbf{z}} + (M_{12} - \frac{i}{2}\Gamma_{12})\sigma_{\mathbf{x}}, \tag{1.1.24}$$

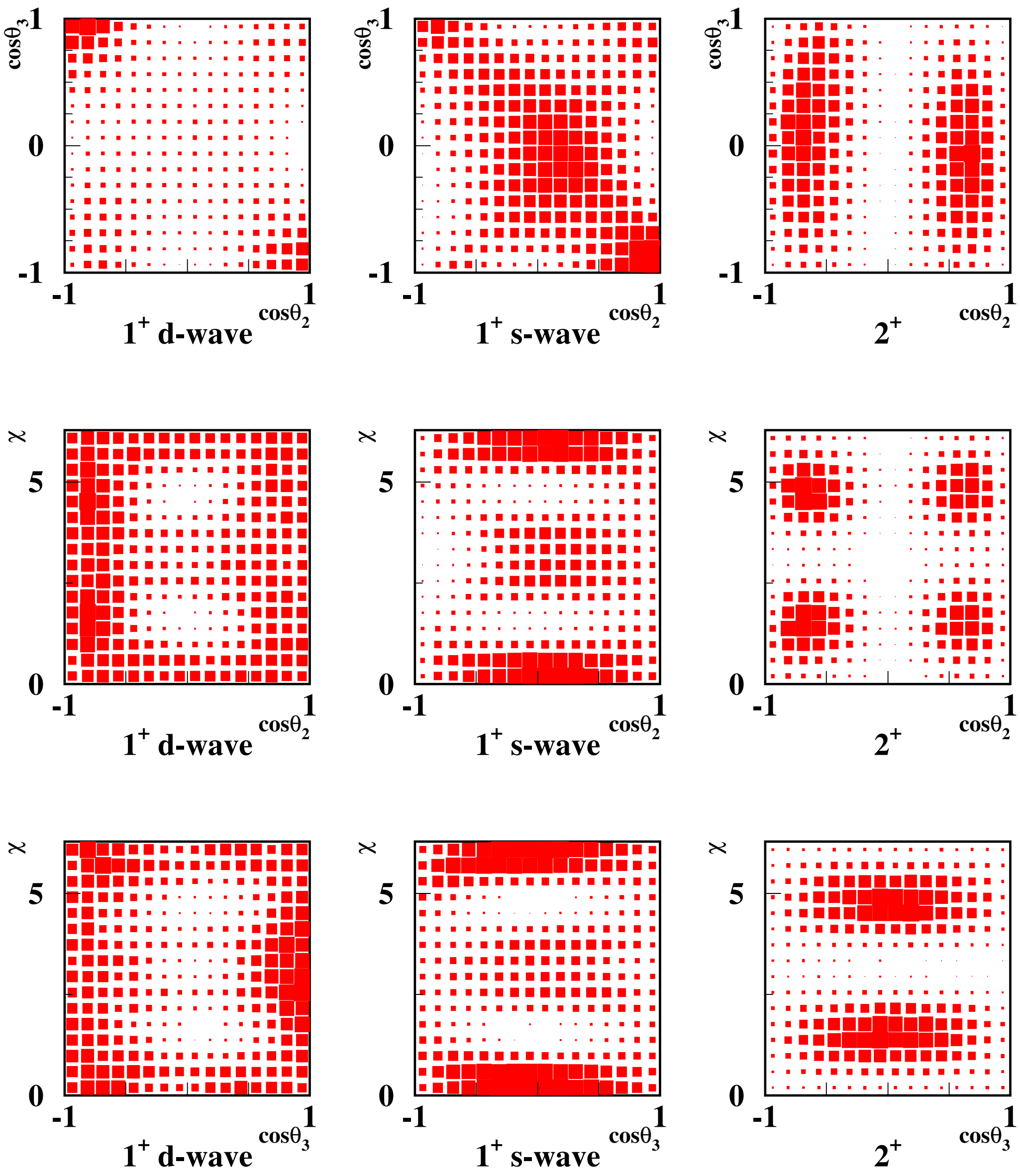


Figure 1.7: The reconstructed angular distributions for Monte Carlo of $B^- \to D_J^0\,\pi^-$, $D_J^0 \to D^{\star +}\,\pi^-$, $D^\star \to D\,\pi$ decays through the 1^+ **d**-wave, 1^+ **s**-wave and 2^+ D_J^0. Each row shows the projection of the three-dimensional angular distribution onto two of the three angles. Acceptance and smearing are included: the area of each square is proportional to the number of reconstructed and accepted events in that bin.

where $\overline{M} = \frac{1}{2}(M_1 + M_2)$ and $\overline{\Gamma} = \frac{1}{2}(\Gamma_1 + \Gamma_2)$ are the mean mass and width of the two states, the differences are $\Delta M(\Gamma) = M_1(\Gamma_1) - M_2(\Gamma_2)$, and M_{12} and Γ_{12} characterize mixing. $\mathbf{I}$, $\sigma_{\mathbf{z}}$, and $\sigma_{\mathbf{x}}$ are the identity and Pauli-spin matrices.

Physically, M_{12} might arise from a spin-spin interaction of the charm and light quark, or from off-shell transitions between the states #1 and #2. Γ_{12} might arise from on-shell transitions between the states, perhaps through intermediate states like $D\pi\pi$. Given the definitions of states #1 and #2, on-shell transitions through $D^{\star +}\pi^-$ itself do not occur, because, from Eq. 1.1.16, the angular functions are orthogonal when integrated over the three decay angles θ_2, θ_3, and χ.

The masses and widths of the eigenstates, i.e. the observed resonances, follow directly from Eq. 1.1.24 and the analogy with spin up and down in a magnetic field:

$$\hat{M}_\pm - \frac{i}{2}\hat{\Gamma}_\pm = \overline{M} - \frac{i}{2}\overline{\Gamma} \pm \left\{ \frac{1}{4}(\Delta M - \frac{i}{2}\Delta\Gamma)^2 + (M_{12} - \frac{i}{2}\Gamma_{12})^2 \right\}^{1/2}. \tag{1.1.25}$$

The mixing angles β and ϕ can be obtained with the help of the intermediate complex number z:

$$\tan z \equiv \frac{M_{12} - \frac{i}{2}\Gamma_{12}}{\frac{1}{2}(\Delta M - \frac{i}{2}\Delta\Gamma)} \tag{1.1.26}$$

$$e^{i\phi}\tan\beta \equiv \tan\frac{z}{2}. \tag{1.1.27}$$

From Eqs. 1.1.26 and 1.1.27, one can infer that a zero value for ϕ means that $\Delta M - \frac{i}{2}\Delta\Gamma$ and $M_{12} - \frac{i}{2}\Gamma_{12}$ would share the same complex phase. The sign convention in Eq. 1.1.16 could be changed to a relative + sign, between the **s**-wave and **d**-wave components at $\theta_2 = \theta_3 = 0$, with the consequence that β would need to be changed to $-\beta$, and all the strong phases would be advanced (or retarded) by π.

The two narrow resonances are described by non-relativistic Breit-Wigner amplitudes (A_{1n} and A_2):

$$A_{\text{nrel}} = \frac{\Gamma/2}{M - E - i\Gamma/2}, \tag{1.1.28}$$

where $E = M(D^{\star +}\pi^-)$ and M and Γ are the mass and width of the resonance. However, because the 1^+ state is broad, it is important to use the relativistic Breit-Wigner and include the energy dependence of the width. The form for the amplitude under these circumstances is [2, 16]

$$A_{\text{rel}} = \frac{\sqrt{E\Gamma(E)}}{M^2 - E^2 - iM\Gamma(E)}. \tag{1.1.29}$$

The energy-dependent width is given by [16]

$$\Gamma(E) = \Gamma_0 \frac{p^\star}{p_0^\star} \frac{M}{E}. \tag{1.1.30}$$

The first factor is the phase space factor for an **s**-wave decay, where $p^\star$ is the decay three-momentum in the rest frame of the resonance and the $p_0^\star$ is the on-shell value. The second is the barrier penetration factor $\frac{\rho(E)}{\rho(M)}$ with $\rho(E) = E$.

Chapter 2

Reconstruction Method

2.1 Introduction

Product branching ratios for the decay chain $B^- \to D_J^0\,\pi^- D_J^0 \to D^{\star +}\,\pi^- D^{\star +} \to D^0\,\pi^+$ are expected to be on the order of 10^{-4}–10^{-3}. Yet, with millions of B mesons in the CLEO II data, assembling the large sample of these decays required for the partial-wave analysis is very difficult. The primary reason for this is the low reconstruction efficiency for the D meson. This limitation is common to the study of all B decays with a D meson in the final state.

With only weak decay channels and a relatively large mass, the D meson has many decay modes and most of the width decays to high multiplicity final states. Selecting the largest, low multiplicity modes that can be fully reconstructed (see Table 2.1) typically results in reconstruction of $\approx 10\%$ of the width.

It is possible to avoid this limitation. Consider reconstruction of the decay sequence

$$B \to D^\star X\ ,\ D^\star \to D\pi_{\mathrm{slow}}, \tag{2.1.1}$$

where the B is produced with known energy at the $\Upsilon(4S)$. As will be shown below, it is possible to reconstruct the kinematics of the decay sequence by reconstructing only the pseudo-particle X and the π_{slow}: no reconstruction of the D is necessary.

By performing this reconstruction with $X = \pi^-\pi^-$, we may obtain the high statistics sample of $B^- \to D^{\star +}\,\pi^-\,\pi^-$ we need for the partial-wave analysis of the $D^{\star +}\pi^-$ reso-

Table 2.1: Modes typically used for full reconstruction of the D^0 meson. The acceptance is shown for typical reconstruction efficiencies.

Mode	$\mathcal{BR}$	ϵ_{recon}	acceptance
$D^0 \to K^-\pi^+$	0.04	0.7^2	0.02
$D^0 \to K^-\pi^+\pi^0$	0.14	$0.7^2 \times 0.5$	0.035
$D^0 \to K^-\pi^+\pi^+\pi^-$	0.08	0.7^4	0.02
$D^0 \to K^-\pi^+\pi^+\pi^-\pi^0$	0.04	$0.7^4 \times 0.5$	0.005
Total			$\approx 10\%$

nances described in Chapter 1.

2.2 Variables and Constraints

There are five particles in the sequence described by Eq. 2.1.1 (B,$D^\star$,X,D,π). One needs four variables (a 4-momentum) to describe each particle. A total of $20 = 5{\times}4$ variables are then needed to completely describe the sequence described in Eq. 2.1.1. The constraints we exploit are

1. energy-momentum conservation in the two decays, $4 \times 2 = 8$ constraints,

2. the B mass of $5279 \pm 2\,\text{MeV}$ (1 constraint),

3. the B energy, or the beam energy of CESR at the $\Upsilon(4S)$, $E_B = 5290 \pm 3\,\text{MeV}$. (In practice, we use the CESR beam energy for the event undergoing reconstruction) (1 constraint),

4. the $D^\star$ mass: $2010.1 \pm 0.6\,\text{MeV}$ for the $D^{\star+}$(1 constraint),

5. the four reconstructed variables that describe the X. These may be taken as the mass of X, which may itself be reconstructed from its decay products, and the three momentum components of X(4 constraints),

6. the four reconstructed variables that describe the π_{slow}, which may be taken as the π mass, $139.57\,\text{MeV}$ for the π^+, and the three measured momentum components. (4 constraints)

7. the mass of the D: $1864.5 \pm 0.5\,\text{MeV}$ for the D^0(1 constraint),

There are $8+1+1+1+4+4+1 = 20$ constraints. This is just enough to cover the variables and produce a complete solution, a '0C' fit, for the kinematics. However, the solution is quadratic, meaning that:

1. the solutions are, in general, complex-valued. When a non-zero imaginary part is required, there is no physically acceptable solution.

2. for most cases with a physically acceptable (real) solution, there are two equally acceptable solutions due to a "quadratic ambiguity."

With no extra constraints, one cannot make the 'usual' plots showing evidence of a signal, like $\Delta E = E_{\text{rec}} - E_B$ and $M_{\text{rec}} - m_B$. There are still a number of methods to establish a signal:

1. use of characteristic shapes in some distributions of kinematic variables, such as $\cos\phi$ (ϕ defined in Figure 2.3),

2. use of angular distributions in the decay of the $D^\star$,

3. use of Dalitz plots to search for resonances between the $D^\star$ and X, when X is not a single body.

The latter two of these methods are used our analysis of the $D^{\star +}\pi^-\pi^-$ final state.

2.3 Physical Constraints and Reconstruction

Eliminating the cases with no physically acceptable (real) solution for the kinematics is the most powerful tool in rejecting backgrounds. To master this task, the essence of these cases must be understood.
Given

1. a measured $\vec{p}_\pi$, hypothesized to be the slow pion,

2. a measured $\vec{p}_X$and m_X, hypothesized to come from $B \to D^\star X$,

3. the constraints to the hypothesized decay sequence $B \to D^\star X$, $D^\star \to D\,\pi$;

(a) E_B,

(b) m_B,

(c) $m_{D^\star}$,

(d) m_D,

(e) m_π;

there are three types of inconsistency with the hypothesized decay sequence. These are as follows:

1. Given the measured $\vec{p}_X$ and m_X, 4-momentum cannot be conserved in the decay $B \rightarrow D^\star X$.

2. Given the measured E_X and $\vec{p}_\pi$, 4-momentum cannot be conserved in the decay $D^\star \rightarrow D\,\pi$.

3. Even when 4-momentum can be conserved in both (1) and (2), no $D^\star$ direction exists that is consistent with both (1) and (2).

We shall discuss the rejection of these cases and the solution for the kinematics of the decay sequence in the following sections.

2.3.1 The Decay $B \rightarrow D^\star X$

Here, we are given

1. m_B, E_B(assumed), and thus $|\vec{p}_B|$ as well,

2. $m_{D^\star}$(assumed),

3. $\vec{p}_X$and m_X(reconstructed).

By requiring energy conservation we can calculate the energy of the $D^\star$ given by $E_{D^\star} = E_B - E_X = E_B - \sqrt{m_X^2 + |\vec{p}_X|^2}$. The most basic constraint is that $E_{D^\star}$ should exceed $m_{D^\star}$:

$$E_{D^\star} > m_{D^\star}. \tag{2.3.1}$$

This is a weak constraint. If true, one can compute the magnitude of the momentum of the $D^\star$ in the lab frame:

$$|\vec{p}_{D^\star}| = \sqrt{E_{D^\star}^2 - m_{D^\star}^2}. \tag{2.3.2}$$

Once $|\vec{p}_{D^\star}|$ has been computed using conservation of energy, we can examine whether the hypothesized decay can conserve the spacelike components of momentum as well: $\vec{p}_{D^\star} = \vec{p}_B - \vec{p}_X$. The three momentum vectors must satisfy the triangle inequality: None of the momenta can be greater than the sum of the other two. The power of this requirement comes from the relatively small size of $|\vec{p}_B|$. A simple way of parameterizing this is to calculate the cosine of an interior angle of the triangle formed by $|\vec{p}_{D^\star}|$, $|\vec{p}_B|$ and $-\vec{p}_X$. We shall choose the angle between $-\vec{p}_X$ and the $D^\star$ directions as shown in Figure 2.1.

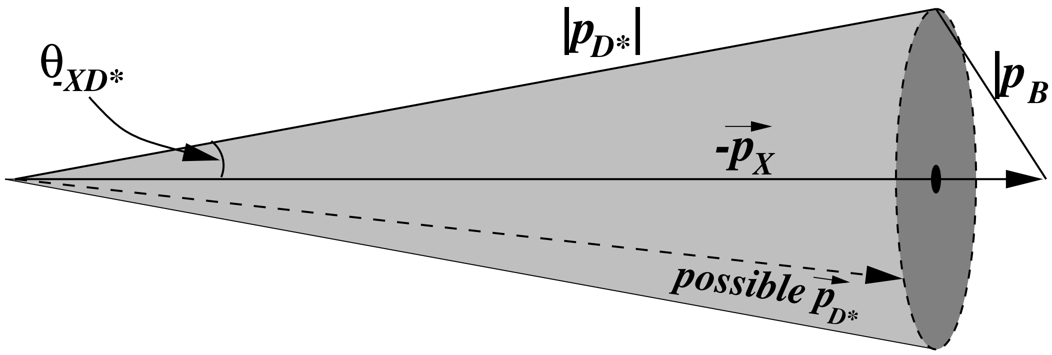


Figure 2.1: The $-X$–$D^\star$ cone angle as described in the text. Notice that because $|\vec{p}_B| \ll |\vec{p}_{D^\star}|$ or $|\vec{p}_X|$ the $D^\star$ must be nearly back to back with X for signal.

Since we know the direction of the X system, we have the $D^\star$ direction constrained to a cone about the $-\vec{p}_X$ direction. The cosine of the opening angle of this cone, in terms of lab quantities is given by

$$\cos\theta_{XD^\star} = \frac{m_B^2 - m_{D^\star}^2 - m_X^2}{2|\vec{p}_{D^\star}||\vec{p}_X|} - \frac{1}{\beta_{D^\star}\beta_X}. \tag{2.3.3}$$

For cases where the triangle does not close this gives $|\cos\theta_{XD^\star}| > 1$. The requirement for a physical solution, or "physicality," is then that $|\cos\theta_{XD^\star}| \leq 1$. For signal, typically, $\theta_{XD^\star} \approx 0.1$. The lower limit on $\cos\theta_{XD^\star}$ depends on m_X and $|\vec{p}_X|$. To account for variability, it is convenient to work in the B rest frame. With $\gamma_B = E_B/m_B$, $\beta_B = \sqrt{1 - 1/\gamma_B^2}$, and using the quantities that describe the $D^\star$ and the X in the B rest frame,

$$E_{D^\star}^\star = (m_B^2 + m_{D^\star}^2 - m_X^2)/2m_B \tag{2.3.4}$$

$$E_X^\star = (m_B^2 + m_X^2 - m_{D^\star}^2)/2m_B \tag{2.3.5}$$

$$|\vec{q}_X| = |\vec{q}_{D^\star}| = \lambda(m_B^2, m_X^2, m_{D^\star}^2)/4m_B^2 \tag{2.3.6}$$

$$\lambda(a,b,c) \quad = \quad a^2+b^2+c^2-2ab-2ac-2bc, \tag{2.3.7}$$

one can reconstruct the decay zenith of the X. This angle, $\theta^\star_X$, is given by

$$\cos\theta^\star_X = -\frac{\beta_B(E^\star_X - E^\star_{D^\star})}{2|\vec{q}_X|} + \frac{|\vec{p}_X|^2 - |\vec{p}_{D^\star}|^2}{2\gamma_B^2\beta_B m_B|\vec{q}_X|}. \tag{2.3.8}$$

The physical constraint is then, for all m_X and $|\vec{p}_X|$,

$$-1 \leq \cos\theta^\star_X \leq 1. \tag{2.3.9}$$

This is effectively a $m_X$dependent requirement on $|\vec{p}_X|$. As noted above, one benefits from the small value of β_B, which yields a small range of allowed $|\vec{p}_X|$.

2.3.2 The Decay $D^\star \to D\,\pi$

Here, we are given

1. $m_{D^\star}$, $E_{D^\star}$(calculated in the last section),
2. m_{D^0}(assumed),
3. $\vec{p}_\pi$(reconstructed) and m_π(assumed).

We have the same information as above in the decay $B \to D^\star X$. Similarly, we have $E_\pi = \sqrt{m_\pi^2 + |\vec{p}_\pi|^2}$ and

$$E_{D^0} \quad = \quad E_B - E_X - E_\pi \tag{2.3.10}$$

$$E_{D^0} \quad > \quad m_{D^0} \tag{2.3.11}$$

$$|\vec{p}_{D^0}| \quad = \quad \sqrt{E_{D^0}^2 - m_{D^0}^2}. \tag{2.3.12}$$

Having required energy conservation to derive $|\vec{p}_{D^0}|$, we check that 3-momentum can be conserved in the decay by asking that a triangle formed by the three momenta can close. The power of this physicality requirement again comes from the relatively small value of $\vec{p}_\pi$ compared to the other two. We choose the cosine of the angle between the $D^\star$ and the π_{slow} directions to parameterize the closure of this triangle as in Figure 2.2. The angle of this cone in terms of lab quantities is given by

$$\cos\theta_{\pi D^\star} = -\frac{m_{D^\star}^2 + m_\pi^2 - m_{D^0}^2}{2|\vec{p}_{D^\star}||\vec{p}_\pi|} + \frac{1}{\beta_{D^\star}\beta_\pi}. \tag{2.3.13}$$

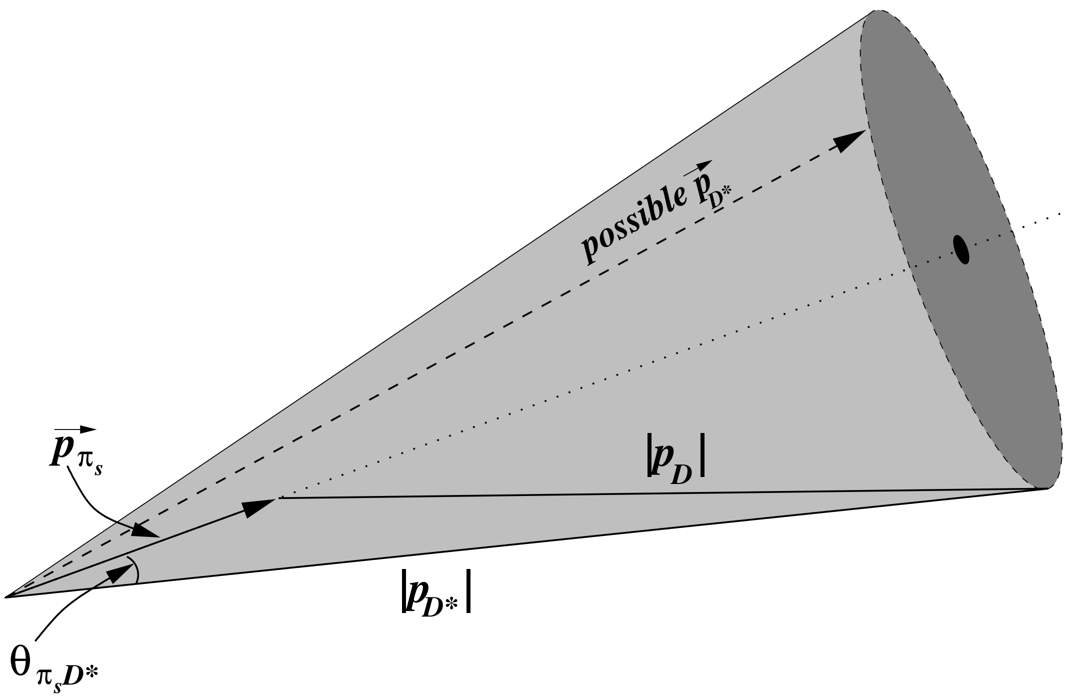


Figure 2.2: The $\pi_{\rm slow}$–$D^\star$ cone angle as described in the text. Notice that the $D^\star$ will be moving in a similar direction to the $\pi_{\rm slow}$ for signal.

This angle is typically larger than the angle $\theta_{XD^\star}$. By knowing the direction of the $\pi_{\rm slow}$ we have the $D^\star$ direction constrained to a second cone of angle $\theta_{\pi D^\star}$ about the $\pi_{\rm slow}$. As in the case of $B \to D^\star X$, $\cos\theta_{\pi D^\star}$ has a lower limit that varies event by event and we account for this by calculating the decay zenith of the $\pi_{\rm slow}$ with respect to the $D^\star$ lab direction in the $D^\star$ rest frame. This yields

$$\cos\theta^\star_\pi = -\frac{\beta_{D^\star}(E^\star_\pi - E^\star_{D^0})}{2|\vec{q}_\pi|} + \frac{|\vec{p}_\pi|^2 - |\vec{p}_{D^0}|^2}{2\gamma^2_{D^\star}\beta_{D^\star}m_{D^\star}|\vec{q}_\pi|}, \tag{2.3.14}$$

and we require,

$$-1 \leq \cos\theta^\star_\pi \leq 1. \tag{2.3.15}$$

This E_X-dependent requirement on $|\vec{p}_\pi|$ is effective because of the small value of $\beta^\star_\pi$ that gives $|\vec{p}_\pi|$ in a narrow range.

2.3.3 Solutions for $D^\star$ Direction

The possible $D^\star$ momenta must have magnitude defined by Eq. 2.3.2, and must lie simultaneously on the cone defined in Figure 2.1, and the distinct cone defined in Figure 2.2. The vertices of the cones are both at the origin on momentum space. There are two possibilities:

1. these cones may not cross (no real solution for $\vec{p}_{D^\star}$).
2. if they do cross, there will generally be two different but equally viable solutions (a quadratic ambiguity).

In order to parameterize the crossing of the two cones, both to reject cases where they do not cross as well as for the reconstruction of the $D^\star$ direction in the lab, we shall define an angle ϕ, as shown in Figure 2.3. We define ϕ as the azimuthal angle about the $-X$–$D^\star$ cone where the two cones cross. In order that $\pm\phi$ describe both solutions, $\phi = 0$ is defined so that it lies in the $-X$–$\pi_{\rm slow}$ plane. It will be helpful if we have a method for calculating ϕ which continues smoothly when the cones fail to intersect. We can compute phi as

$$\cos\phi = \frac{\cos\theta_{\pi D^\star} - \cos\theta_{XD^\star}\cos\theta_{X\pi}}{\sin\theta_{X\pi}\sin\theta_{XD^\star}}. \tag{2.3.16}$$

When the two cones do not intersect, $|cos\phi| > 1$. When they do, we have both $\pm\phi$ as solutions. The directions of the two $D^\star$ solutions in the lab can be reconstructed with the help of three special unit vectors:

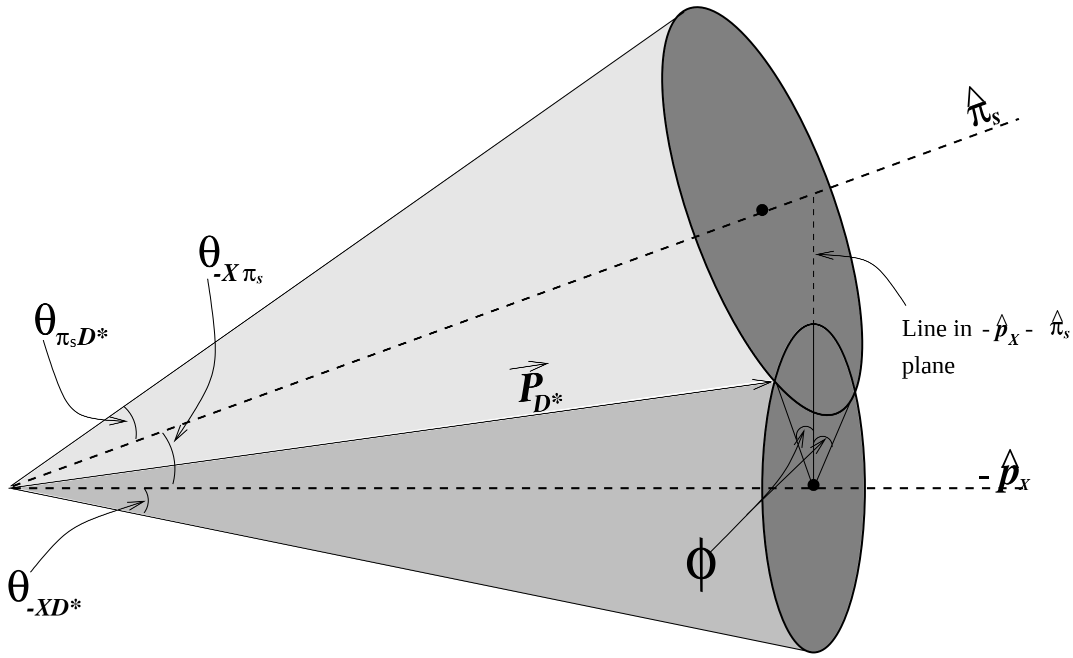


Figure 2.3: The angle ϕ which defines the possible solutions for the $D^{\star}$ direction as described in the text.

1. the first is in the direction of $-\vec{p}_X$,

$$\hat{n}_1 \equiv -\frac{\vec{p}_X}{|\vec{p}_X|};$$

2. the second is perpendicular to $-\vec{p}_X$, and in the plane of $-\vec{p}_X$ and $\vec{p}_\pi$, and at $\phi = 0$,

$$\hat{n}_2 \equiv \frac{\vec{p}_\pi - (\vec{p}_\pi \cdot \hat{n}_1)\hat{n}_1}{|\vec{p}_\pi - (\vec{p}_\pi \cdot \hat{n}_1)\hat{n}_1|};$$

3. the third is perpendicular to the first two, easily formed by taking the cross product of $\hat{n}_1$ on $\hat{n}_2$,

$$\hat{n}_3 \equiv \hat{n}_1 \times \hat{n}_2.$$

Then, we get for the momentum of the $D^\star$ in the lab frame,

$$\vec{p}_{D^\star} = |\vec{p}_{D^\star}|(\cos\theta_{XD^\star}\hat{n}_1 + \sin\theta_{XD^\star}[\cos\phi\,\hat{n}_2 \pm \sin\phi\,\hat{n}_3]).$$

With the solutions for $\vec{p}_{D^\star}$, it is possible to reconstruct $\vec{p}_B$ and $\vec{p}_{D^0}$.

2.4 $B^- \to D^{\star+}\pi^-\pi^-$ Signal Distributions

With a complete solution for the kinematics of the decay sequence $B^- \to D^{\star+}\pi^-\pi^-$, $D^{\star+} \to D^0\pi^+$ we can reconstruct the decay angular distributions discussed in Chapter 1 ($cos\theta_2$,$\cos\theta_3$,χ) and create a Dalitz plot of the $B^- \to D^{\star+}\pi^-\pi^-$ three-body phase space. With no resonances of two like-sign pions, we do not study the Dalitz plot as such: we only plot the mass of each $D^{\star+}\pi^-$ combination. Since $M^2(\pi^-\pi^-) \propto \cos\theta_2$ we have already captured the information in this distribution.

Before we can proceed with the data, we must dispose of two more issues. The first is deciding which $D^{\star+}\pi^-$ combination to study. Although both combinations may be plotted, this makes the statistical interpretation of the data much more complicated. For the general case of $B^- \to D^{\star+}\pi^-\pi^-$, the pions from the B decay are indistinguishable. However, for the resonant components $B^- \to D_J^0\pi_1^-$,$D_J^0 \to D^{\star+}\pi_2^-$, $D^{\star+} \to D^0\pi_3^+$ it is important to remember that the π_1 and π_2 play different roles: for $D_1(2420)^0$, $D_2^\star(2460)^0$ and most $D_1^0(j=\frac{1}{2})$ decays, their momentum spectra do not overlap as shown in Figure 2.4. We may use this fact to distinguish them by considering only the $D^{\star+}\pi^-$ combination with the lower mass. This requirement keeps all signal decays with $M(D^{\star+}\pi^-) < 3.0\,\text{GeV}$ and most of the phase space

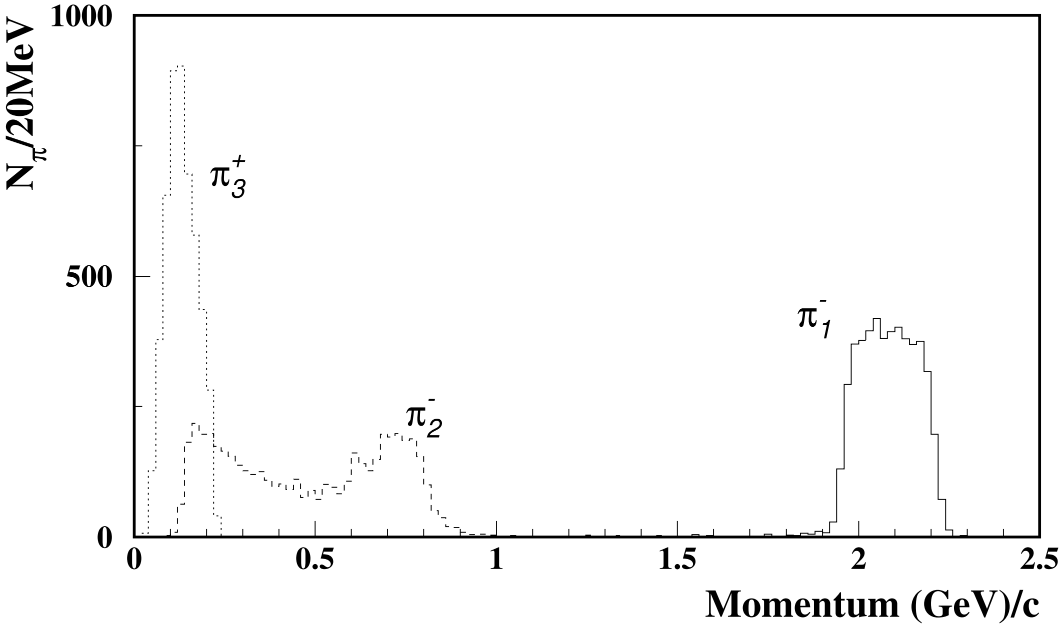


Figure 2.4: The reconstructed spectra of the three pions produced in the decay $B^- \to D_1(2420)^0\,\pi_1^-$, $D_1(2420)^0 \to D^{\star+}\,\pi_2^-$, $D^{\star+} \to D^0\,\pi_3^+$. Note the peaking in p_{π_2} due to the angular distribution in the $D_J^0 \to D^{\star+}\,\pi^-$ decay.

in the region $3.2\,\text{GeV} < M(D^{\star +}\pi^-) < 4.0\,\text{GeV}$, giving very high acceptance for decays to all three D_J.

The second issue is deciding which quadratic solution from the reconstruction method will be used. Since the cones are typically small, the two solutions are similar, producing only a small smearing in the decay angles and $M(D^{\star +}\pi^-)$ when the wrong solution is used. This smearing is most serious for the $M(D^{\star +}\pi^-)$ distributions of the narrow $D_1(2420)^0$ and $D_2^\star(2460)^0$. We may use the fact that the two quadratic solutions for $\vec{p}_B$ are at different angles with respect to the beamline to select one of them. Because B production at the $\Upsilon(4S)$ has a distribution of $dN/d\cos\theta_B = \sin^2\theta_B$ with respect to the beamline, we select the quadratic solution with smaller $|\cos\theta_B|$. This enhances the likelihood of selecting the right solution by $\approx 60\%$ over selecting one of the two solutions randomly.

Chapter 3

The Experiment and Data

3.1 The CLEO II Experiment

The data used in the analysis of $B \rightarrow D_J\,\pi$ were collected by the the CLEO II experiment, located at the Cornell Electron Storage Ring (CESR) in Ithaca, New York. CESR runs at center of mass energies near the $\Upsilon(4S)$ providing data suitable for the kinematic reconstruction described in the Chapter 2. Figure 3.1 shows the cross section for $e^+e^- \rightarrow$ hadrons as a function of center of mass energy near the $\Upsilon(4S)$.

The maximum cross-section for $e^+e^- \rightarrow \Upsilon(4S)$ is ≈ 1.07 nb, while the cross-section for continuum production of quarks is ≈ 3 nb. Data collected at a center of mass energy 50 MeV below the $\Upsilon(4S)$ peak are used to model the continuum background under the $\Upsilon(4S)$.

The CLEO II detector, shown in Figure 3.2, measures both charged and neutral particles with good resolution and efficiency. Charged particles are measured in a 1.5 Tesla magnetic field through three nested, cylindrical drift chambers covering 94% of the solid angle. These chambers are followed by time-of-flight (TOF) counters. Particle identification is accomplished by combining specific ionization (dE/dx) information from the drift chambers with TOF information. An electromagnetic calorimeter and muon chambers are used to help reject electrons and muons as $\pi^{\pm}$ candidates.

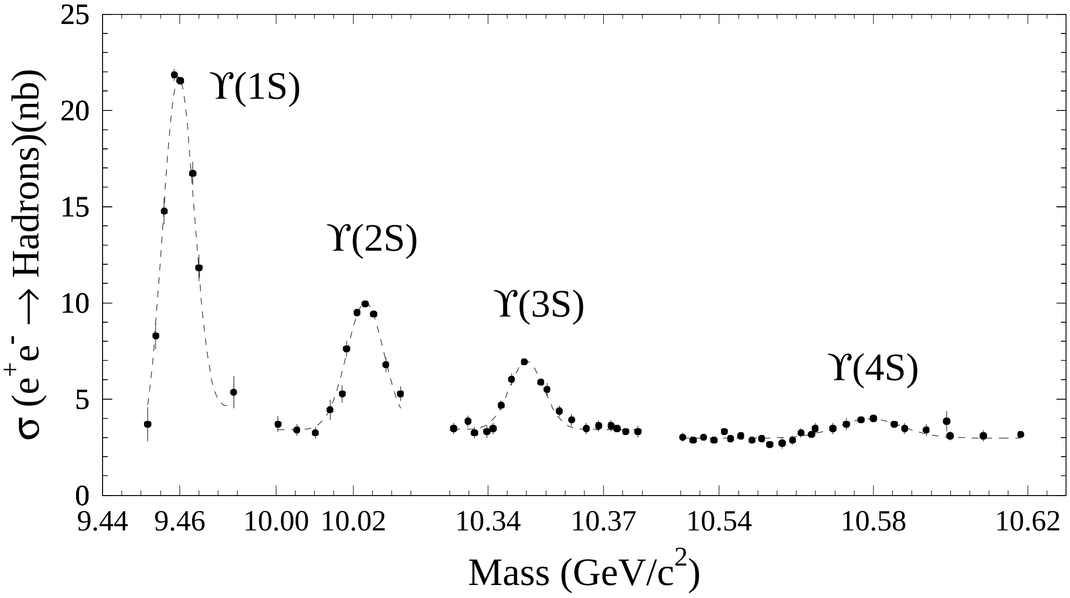


Figure 3.1: The cross section for $e^+e^- \rightarrow$ hadrons near the $\Upsilon(4S)$.

3.1.1 Tracking System

The tracking system consists of three separate drift chambers. The precision tracking layer (PTL) is the innermost chamber. This straw tube chamber has six layers with 64 axial wires per layer and provides the best measurements in the $r-\phi$ plane. The PTL was intended to make z measurements using charge sharing information between the readouts at both ends of the straw tubes. These efforts have been largely unsuccessful due to the short length of the PTL. The replacement of the PTL with a silicon vertexing detector capable of making precision measurements in both $r-\phi$ and z will be discussed in Part II.

Outside of the PTL is the vertex detector (VD). Once the innermost tracking device, this ten-layer wire chamber contains 800 axial sense wires to measure trajectories in $r-\phi$ as well as segmented mylar cathodes for measuring z at the inner and outer radii of 8.1 cm and 16.4. These are the innermost z measurements available in the CLEO detector.

The main drift chamber (DR) is a wire chamber with inner radius of 17.5 cm and outer radius of 95 cm. There are 40 axial wire layers and eleven layers with stereo angles of $\approx$ 5% . The stereo layers, as well as cathodes at the inner and outer radii, give the

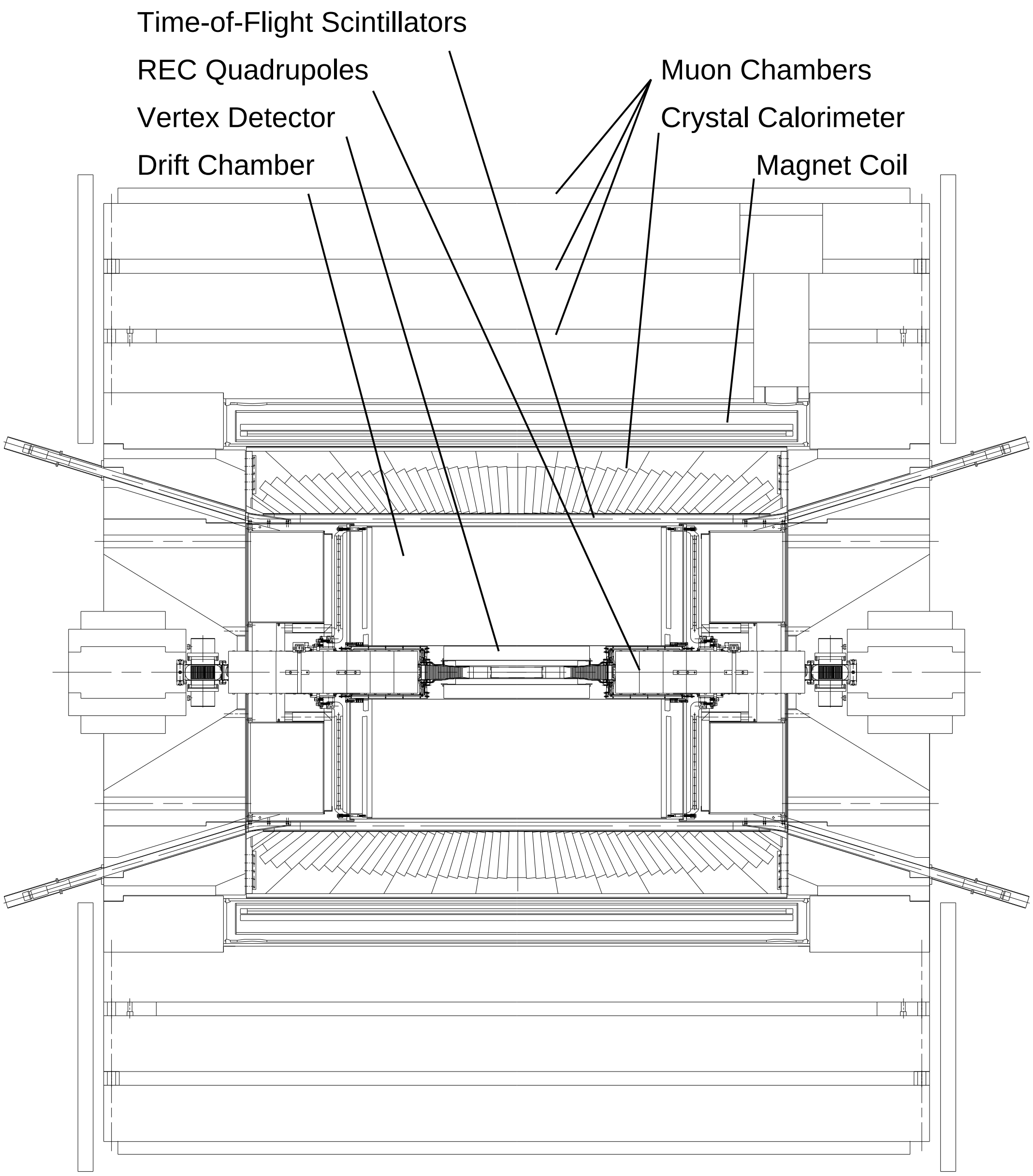


Figure 3.2: A side view of the CLEO II detector.

best z measurements available in CLEO II. The large range in radius allows for precise measurements of track curvature giving good momentum resolution. The large path length also provides a good measurement of specific ionization (dE/dx) which is used to help identify pions in this analysis.

3.1.2 Time of Flight Counters

At the outside of the drift chamber are time of flight counters consisting of 64 scintillators read out by phototubes. With a time resolution of 150 ps these counters provide particle identification used in selection of pions. Together with dE/dx information, there is good separation between pions and kaons with momenta between 150 and 1000 MeV.

3.1.3 Electromagnetic Calorimeter

Inside of the superconducting magnet is an electromagnetic calorimeter consisting of 7800 thallium-doped CsI crystals read out by photodiodes. These crystals measure the energy of electromagnetically interacting particles, providing excellent identification of electrons by comparing the energy deposited in the calorimeter with the momentum of the corresponding track in the tracking system. For electrons one expects the ratio $E/P = 1$. This information is helpful in rejecting high momentum primary electrons from B decays as pion candidates.

3.1.4 Muon Detectors

The return yoke of the superconducting magnet serves as the absorber for the muon identification system. There are three layers of proportional counters at depths of 36, 72 and 108 cm. The total thickness is between seven and ten nuclear interaction lengths depending on the direction of the track. The muon system is used in this analysis to reject high momentum primary muons from B decays as pion candidates.

3.2 Event Selection

3.2.1 Data Set

The data used in this analysis were collected from November 1990 to April 1995. The data are divided into fifteen datasets distinguished by minor changes in detector configuration. These changes lead to shifts in the calibration of the beam energy. As an important constraint in our reconstruction method, we are sensitive to these changes. Use of the incorrect beam energy in the method leads to systematic errors in reconstruction efficiency.

These energy calibration changes are described in terms of shifts in the apparent B mass which should remain constant. The apparent B mass is calculated by measuring the momentum of B mesons in fully reconstructed hadronic decays and using $m_B = \sqrt{(\frac{E_{beam}}{2})^2 - |\vec{p}_B|^2}$. When performing our partial reconstruction on data from each dataset, the energy of the beam is corrected for this miscalibration. Information about these datasets, including these 'B mass shifts' is summarized in Table 3.1.

Table 3.1: A summary of the CLEO II datasets used in this analysis.

Dataset	Run Period	$\int \mathcal{L}_{\mathrm{on}\Upsilon(4\mathrm{S})}$ (pb^{-1})	$\int \mathcal{L}_{\mathrm{off}\Upsilon(4\mathrm{S})}$ (pb^{-1})	$N_{B\overline{B}}$ (thousands)	Reconstructed m_B(GeV)
4s2	11/90 – 6/91	463	197	501	5.2796
4s3	7/91 – 2/92	436	209	458	5.2802
4s4	4/92 – 5/92	214	101	231	5.2783
4s5	7/92 – 10/92	216	105	231	5.2783
4s6	11/92 – 1/93	233	85	248	5.2779
4s7	3/93 – 7/93	285	177	313	5.2798
4s8	8/93 – 9/93	189	94	203	5.2776
4s9	11/93 – 1/94	230	117	251	5.2790
4sA	1/94 – 2/94	138	54	144	5.2783
4sB	3/94 – 5/94	85	64	87	5.2786
4sC	6/94 – 8/94	115	36	115	5.2803
4sD	9/94 – 10/94	53	50	54	5.2801
4sE	10/94 – 11/94	71	62	66	5.2786
4sF	11/94 – 11/94	89	66	91	5.2793
4sG	1/95 – 4/95	293	192	301	5.2796
Total	11/90 – 4/95	3111	1608	3295	—

3.2.2 Global Event Cuts

Since the B mesons are produced almost at rest, their decay products are, on average, isotropically distributed throughout the detector. This produces events that tend to be 'spherical' in shape. Continuum events are more jetlike in shape. In order to select events that are consistent with the hypothesis of $B\overline{B}$ production, we choose events that are spherical using the shape variable R_2. This variable is the ratio of the second Fox-Wolfram moment to the zeroth moment [17] and is useful in discriminating between $B\overline{B}$ and continuum events. Only events with $R_2 < 0.25$ are used in this analysis.

We also require that the events we use have no charged tracks with momentum greater than 2.45 GeV. This requirement eliminates a small amount of continuum background at the cost of no signal.

3.2.3 Track Selection

The decay sequence $B^- \to D^{\star +}\,\pi^-\,\pi^-,\ D^\star \to D\,\pi$ is reconstructed using only the three pions. All candidate pions are required to be consistent with production at the primary event vertex. In addition, we use only tracks with $|\cos\theta| < 0.91$ with respect to the beamline so that they pass through some portion of the central drift chamber. The detector acceptance for charged tracks below 50 MeV is negligible. We restrict the candidate π_3 to have momentum greater that 50 MeV, discarding a small fraction of π_3 from signal events.

Consistency with the pion hypothesis is required for all pion candidates by demanding that the dE/dx and time of flight information are within three standard deviations of the expected values when the information is available. When unavailable, no cut is made.

We use information for the muon counters and the calorimeter to reduce the number of muons and electrons which can fake our π_1 and π_2. These vetoes are only applied when the information from these detector elements is available and are not applied to π_3 candidates.

The efficiency of these loose tracking and pion consistency cuts for the three pions from $B^- \to D_1(2420)^0\,\pi^-,\ D_1(2420)^0 \to D^{\star +}\,\pi^-,\ D^{\star +} \to D^0\,\pi^+$ are 0.75, 0.69, and 0.51 respectively.

3.2.4 Selecting a Three-π Combination

Reconstruction of $B^- \to D^{\star +}\,\pi^-\,\pi^-$ is attempted with all combinations of three pions that meet the above requirements. This often produces more than one acceptable signal combination per signal event. To choose a three pion combination for each event, we note that entries where a random track successfully substitutes for a real π in a signal event are dominated by fake π_2 from the 'other B.' The momentum spectrum of these fakes is roughly the same as the charged particle spectrum at the $\Upsilon(4S)$: they are mostly low-momentum tracks (65-300 MeV/c). For this reason, we accept only the three-π combination in an event with the highest p_{π_2}. This enhances the likelihood of selecting the signal combination by 50% over selecting a combination at random. Mis-reconstruction of signal due to selection of the wrong three-π combination constitutes a significant source of background.

3.3 Backgrounds

Typical events at the $\Upsilon(4S)$ have several candidate pions that satisfy all of our pion selection cuts. The leads to tens or hundreds of candidate three-π combinations for each event. Rejection of this large combinatoric background relies upon the impositions of "physicality" inherent in the reconstruction method. These requirements enforce a specific kinematic relationship among the candidate tracks: the three charged tracks must have momenta and angles that are physically possible for the decay chain we seek. For some background sources, the kinematic relationship is achieved by accident: we call these sources 'random.' For others, the kinematic relationship is closely approximated by a decay that is similar to $B^- \to D^{\star +}\,\pi^-\,\pi^-$: we call these sources 'semi-correlated feedthrough.' The expected contributions from these backgrounds are shown along with the reconstructed data in Figure 3.3. The top two backgrounds in the legend of that figure are 'random,' and the bottom three are 'semi-correlated feedthrough' of various types.

3.3.1 Random Combinatoric Background

Random background is produced by three-π combinations of random tracks in the detector. These tracks are successfully reconstructed to create the candidate decay chain but have no special kinematic relationship to one another. The distribution of this background is dictated by the momentum spectrum of candidate pions and their distribution in the detector, and tends to be smooth in $M(D^{\star +}\pi^-)$, compared to the distributions for our resonant signal modes.

The background from continuum is of this type since the constraint for $B \to D^\star X$ has no meaning for these events. With samples of data collected 'off-resonance' it is a simple matter to model the contribution from continuum. $B\overline{B}$ events also contain random background. However, $B\overline{B}$ events also contain a background of a second type, which we refer

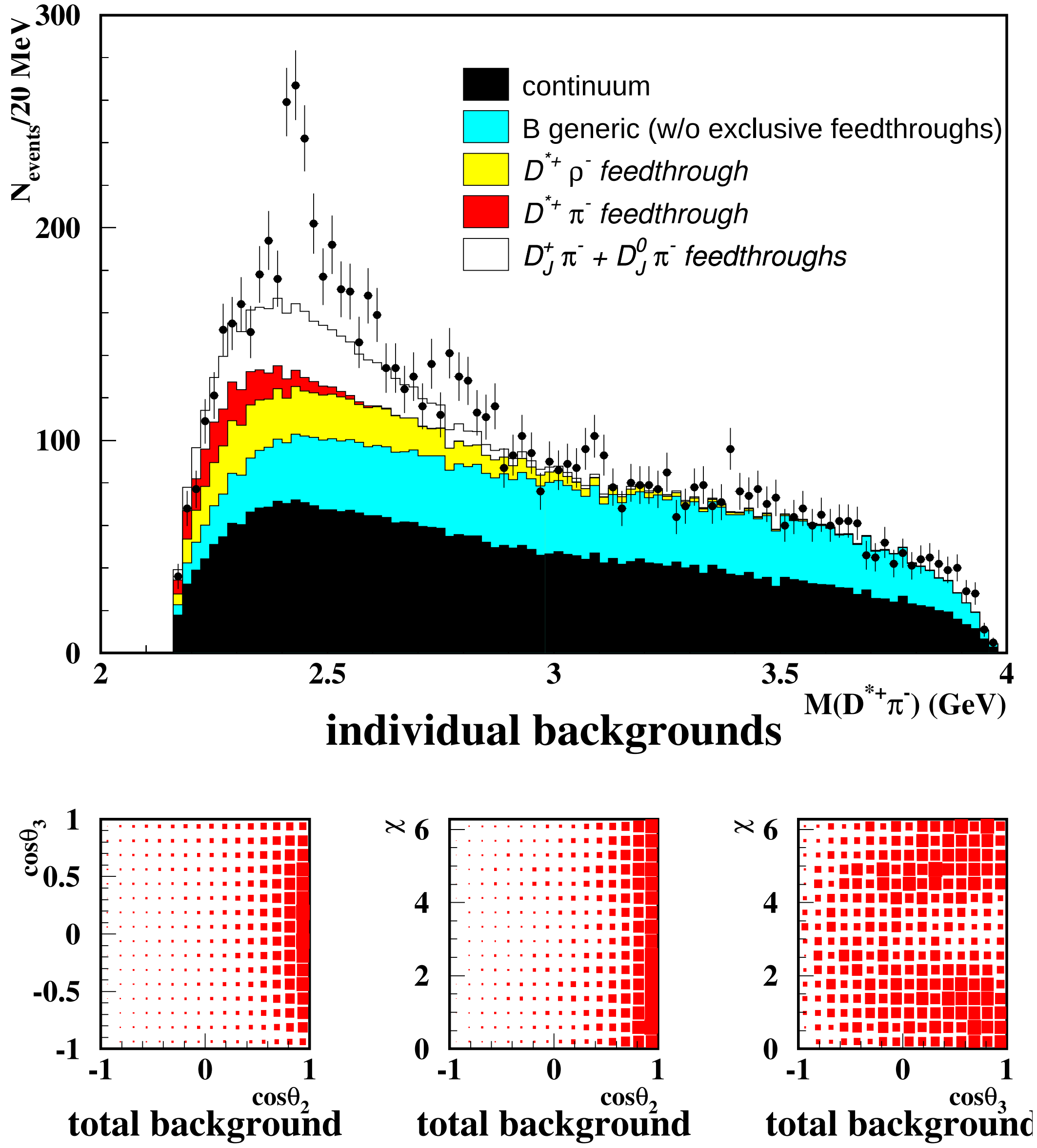


Figure 3.3: The distribution of backgrounds in our observables. The upper plot shows the mass distributions of the individual expected backgrounds adding to the total background. The continuum background shown is from off-resonant data. The others are from the Monte Carlo samples described later. The lower plots show the projected two-dimensional angular distributions of the total background. A $\cos\theta_2$ near one corresponds to a slow π_2.

to as 'semi-correlated feedthroughs.'

3.3.2 Semi-correlated Feedthroughs

Some $B\overline{B}$ decays mimic the kinematics of the signal closely enough to enter the signal plots with relatively large efficiency. Investigation of the $B\overline{B}$ background populating the signal plot reveals a pattern: modes that are kinematically similar to the signals have higher feedthrough rates and peak more strongly in $M(D^{\star+}\pi^-)$.

Based on this observation, Monte Carlo studies of nineteen potential background modes was undertaken. The efficiency for events from these modes to enter into our signal plots is shown in Table 3.2.

The non-signal modes which have high efficiency to enter our signal plot are of the type $B \to D^\star X\pi$, $D^\star \to D\pi_{\rm slow}$ (here, and in further discussion, we denote the π_3 of Fig. 1.6 by $\pi_{\rm slow}$) where X consists of either zero or one light particle. In these events, the B decay provides a pion in the correct momentum region for a π_1 candidate, while the $D^\star$ decay provides a slow pion traveling in the proper direction for a $\pi_{\rm slow}$ candidate. The fake π_2 is typically from the decay of the other B in the event. Only the π_1 and $\pi_{\rm slow}$ are 'correlated,' while the π_2 is 'random.' This is the same mechanism responsible for multiple entries in signal events.

Since these random π_2 have the momentum distribution of charged tracks at the $\Upsilon(4S)$, the largest feedthrough efficiencies are for modes where the reconstruction requires a low momentum π_2 candidate. These modes are characterized by having a fast π_1 (≈ 2 GeV/c). Modes where this π_1 peaks strongly in momentum also peak in our $M(D^{\star+}\pi^-)$ distribution.

There are four $B\overline{B}$ decay modes that are of particular concern due to their tendency to peak near the signal region in $M(D^{\star+}\pi^-)$. These are

1. mis-reconstructed $B^- \to D^0_J\,\pi^-$, $D^0_J \to D^{\star+}\,\pi^-$, $D^{\star+} \to D^0\,\pi^+$;
2. $\overline{B^0} \to D^+_J\,\pi^-$, $D^+_J \to D^{\star+}\,\pi^0$, $D^{\star+} \to D^0\,\pi^+$;
3. $\overline{B^0} \to D^{\star+}\,\pi^-$, $D^{\star+} \to D^0\,\pi^+$;
4. $\overline{B^0} \to D^{\star+}\,\rho^-$, $D^{\star+} \to D^0\,\pi^+$.

Table 3.2: Background Efficiency. The typical overall efficiency (including acceptance) for the signal modes and off-resonant is given first, followed by the four most prominent 'semi-correlated feedthroughs,' which tend to peak in $M(D^{\star +}\pi^-)$. The next fifteen feedthroughs tend not to peak in $M(D^{\star +}\pi^-)$, and are similar to the random background that is contributed by the B decay modes that are not entered in this table (so-called B-generic modes).

Mode	Efficiency
Reference modes	
$B^- \to D_J^0\,\pi^-$(signal)	≈ 0.15
Off-resonant	2.1×10^{-4}
Backgrounds peaking in $M(D^{\star +}\pi^-)$	
$\overline{B^0} \to D_J^+\,\pi^-$	≈ 0.09
$B^- \to D_J^0\,\pi^-$(mis-reconstructed)	≈ 0.05
$\overline{B^0} \to D^{\star +}\,\pi^-$	0.013
$\overline{B^0} \to D^{\star +}\,\rho^-$	0.0087
Flat backgrounds	
$\overline{B^0} \to D^{\star +}\,\pi^-\,\pi^0$	2.4×10^{-3}
$B^- \to D^{\star +}\,\pi^-\,\pi^-$(mis-reconstructed)	1.9×10^{-3}
$\overline{B^0} \to D_J^+\pi^-\pi^0$	1.9×10^{-3}
$\overline{B^0} \to D_J^+\rho^-$	1.1×10^{-3}
$\overline{B^0} \to D^+\pi^-$	1.1×10^{-3}
$\overline{B^0} \to D^{\star +}\pi\pi\pi^0$	8.4×10^{-4}
$\overline{B^0} \to D^{\star +}\pi^-\pi^-\pi^+$	8.2×10^{-4}
$\overline{B^0} \to D^{\star +}e^-\overline{\nu}_e$	8.0×10^{-4}
$\overline{B^0} \to D^+\rho^-$	6.2×10^{-4}
$\overline{B^0} \to D^\star\mu^-\overline{\nu}_\mu$	4.9×10^{-4}
$\overline{B^0} \to D^+\pi^-\pi^-$	4.0×10^{-4}
$\overline{B^0} \to D^+\rho^0\pi^-$	3.5×10^{-4}
$\overline{B^0} \to D^+\pi^-\pi^-\pi^+$	3.1×10^{-4}
$\overline{B^0} \to D^+e^-\overline{\nu}_e$	7.3×10^{-5}
$\overline{B^0} \to D^+\mu^-\overline{\nu}_\mu$	6.1×10^{-5}

All four have relatively high feedthrough rates and peak somewhat in our signal plots as shown in Figure 3.3. Because of this, these modes are removed from the sample of $B\overline{B}$ generic Monte Carlo and separate samples are generated to model their contribution to the fit more carefully.

Chapter 4

Fit Technique

4.1 Introduction

We must fit the four-dimensional distribution of the data in the observables $\{M(D^{\star +}\pi^-), \cos\theta_2, \cos\theta_3, \chi\}$ in order to perform the partial-wave analysis of the decay sequence $B^- \to D_J^0\,\pi^-$,$D_J^0 \to D^{\star +}\,\pi^-$, $D^\star \to D\,\pi$ and measure the $B^- \to D_J^0\,\pi^-$ branching ratios. The essential ingredient in this fit is a smooth model for the four-dimensional distribution that includes the acceptance and smearing inherent in the experiment and reconstruction method. We use a method of Monte Carlo interpolation to create a model of the data and perform our parameter estimation with an unbinned maximum-likelihood fit.

Unbinned maximum-likelihood fits have become a popular method of estimating the parameters that describe large multi-dimensional datasets. Using a sample of Monte Carlo to derive the probability distribution function allows one to capture all of the acceptance and smearing effects in the simulation. While such fits have become common in estimating the parameters that describe the shape of such distributions, some new problems must be addressed in order to properly extract branching ratios from these fits. These problems include proper treatment of the error on the total yield and the extraction of efficiencies needed to correct the yields. In addition, determination of the confidence level of such fits is always difficult.

The following sections describe the mathematical foundations of unbinned maximum-likelihood along with the technical details of performing these fits using Monte Carlo samples to estimate the probability distribution function. We will present solutions to the problems of extracting proper yields and efficiencies, as well as a simple method for measuring the goodness-of-fit. Finally, we discuss the details of the method used to perform our four-dimensional fit to the data and the results of tests that confirm our ability to properly estimate the parameters that describe our data sample.

4.2 Unbinned Likelihood

4.2.1 The Parent Distribution

In any experiment there are some number of observables, n, containing information about the phenomena under investigation. We may write that set as a vector, $\vec{x}$, in a space of dimension n. For a sample of N data from the experiment, we have a set of vectors $\{\vec{x}_i\}; i = 1, N$.

The function that describes the frequency of occurrence of data at $\vec{x}$ is the probability distribution function (*pdf*), $p(\vec{x}, \vec{a}_0)$, where $\vec{a}_0$ is a set of m parameters that describe the shape of the parent distribution in the space of observables.

For the purpose of estimating $\vec{a}_0$ with our fit, we define a more general vector of parameters, $\vec{a}$ which allows us to vary the shape of the *pdf* in the vicinity of the true shape.

In the 'classical' maximum-likelihood technique, the *pdf* is assumed to be normalized for all $\vec{a}$:

$$\int_V d^n x p(\vec{x}, \vec{a}) = 1, \quad \forall \vec{a}. \tag{4.2.1}$$

A technique known as 'extended maximum-likelihood' [18] can be used to avoid performing the normalization integral of $p(\vec{x}, \vec{a})$ and is more appropriate to the case where the actual number of events is not known and yields are being measured. The 'extended *pdf*' is given by

$$P(\vec{x}, \vec{A}) = \mathcal{N} p(\vec{x}, \vec{a}), \tag{4.2.2}$$

so that

$$\int_V d^n x P(\vec{x}, \vec{A}) = \mathcal{N}. \tag{4.2.3}$$

Notice that the normalization has been absorbed into the list of parameters, $\vec{A}$:

$$\vec{A} = \{\mathcal{N}, \vec{a}\}. \tag{4.2.4}$$

4.2.2 The Likelihood Function

To derive the extended likelihood, consider drawing N events according to $p(\vec{x}, \vec{a}_0)$. For a particular 'draw' of events, one gets a list of the N n-component vectors $\vec{x}$, labeled by the index i in the set $\{\vec{x}_i\}, i = 1, N$.

In each bin, $\Delta^n x$, of the n-dimensional space of independent variables, one will find some number of the N events from the draw. Imagine averaging over many 'draws' of N events, and keeping track of the mean number of events, $\mu(\vec{x})$, in the bin $\Delta^n x$ located at $\vec{x}$:

$$\mu(\vec{x}) = P(\vec{x}, \vec{A}_0)\Delta^n x. \tag{4.2.5}$$

The unbinned limit corresponds to considering the limit $\Delta^n x \to d^n x$ so that all bins contain either zero or one event. From Poisson statistics,

$$\begin{aligned} \text{Probability of Zero Events in bin} &= e^{-\mu(\vec{x})} &&= e^{-d^n x P(\vec{x}, \vec{A}_0)} \\ \text{Probability of One Event in bin} &= \mu(\vec{x}) e^{-\mu(\vec{x})} &&= d^n x P(\vec{x}, \vec{A}_0) e^{-d^n x P(\vec{x}, \vec{A}_0)}. \end{aligned} \tag{4.2.6}$$

For a given draw of N events, the extended likelihood $\mathcal{L}_N$ of the draw is proportional to the product of the probabilities for all the bins with zero events in them, and of the probabilities for all the bins with one event in them:

$$\mathcal{L}_N \propto \left[\prod_{\text{all bins}} e^{-d^n x P(\vec{x}, \vec{A}_0)} \right] \left[\prod_{i=1}^{N} d^n x P(\vec{x}_i, \vec{A}_0) \right]. \tag{4.2.7}$$

The first product in square brackets is independent of the particular set $\{\vec{x}_i\}, i = 1, N$ that is drawn and contributes a factor of $e^{-\mathcal{N}}$, which is retained only for future use in the maximum-likelihood fit for $\vec{A}$. The second term does depend on the events drawn. Therefore, variations in $\mathcal{L}_N$ all come from the second term. It is the real focus of likelihood analysis.

To dispose of the differential factor in the second term, one argues that the product of all $d^n x$ is a normalization factor that is independent of either the parameters $\vec{a}$ or of the particular draw; one can re-normalize it away in the precise definition of $\mathcal{L}_N$.

Taking logarithms, we have

$$\ln \mathcal{L}_N = \sum_{i=1}^{N} \ln P(\vec{x}_i, \vec{A}_0) - \mathcal{N}. \tag{4.2.8}$$

The first term above is similar in form to what appears in the 'classical' expression for the log-likelihood:

$$\ln \mathcal{L}_N = \sum_{i=1}^{N} \ln p(\vec{x}_i, \vec{a}_0). \tag{4.2.9}$$

The inclusion of the second term, $\mathcal{N}$, is what is implied by the phrase 'extended' log-likelihood: the 'classical' definition of the likelihood is 'extended' to take the overall normalization into account. It turns out that maximization of the extended likelihood with respect to $\vec{A}$ will result in $\mathcal{N} = N$. In essence, one employs the maximization program (such as `Minuit`) to do the numerical analysis of the normalization.

4.2.3 Goodness-of-fit

Consider the distribution of $\ln \mathcal{L}_N$, for many 'draws' of N data points, each draw described by a set of vectors $\{\vec{x}_i\}, i = 1, N$. In the equation

$$\ln \mathcal{L}_N = \sum_{i=1}^{N} \ln P(\vec{x}_i, \vec{A}_0) - \mathcal{N} \tag{4.2.10}$$

the second term, $\mathcal{N}$, is independent of the particular draw of data points. The first term provides all the variation.

The mean of the extended log-likelihood, denoted $\langle \ln \mathcal{L}_N \rangle$, can be found by by convoluting $\ln P(\vec{x}, \vec{A}_0)$ N times:

$$\begin{aligned} \langle \ln \mathcal{L}_N \rangle &= N \langle \ln P(\vec{x}, \vec{A}_0) \rangle - \mathcal{N} \\ &= \tfrac{N}{\mathcal{N}} \int_V d^n x P(\vec{x}, \vec{A}_0) \ln P(\vec{x}, \vec{A}_0) - \mathcal{N}. \end{aligned} \tag{4.2.11}$$

As $N \to \infty$, the central limit theorem implies that $\ln \mathcal{L}_N$ will be distributed about $\langle \ln \mathcal{L}_N \rangle$ in a Gaussian distribution, with variance

$$\mathrm{Var}(\ln \mathcal{L}_N) = N \langle (\Delta \ln P(\vec{x}, \vec{A}_0))^2 \rangle, \tag{4.2.12}$$

where

$$\begin{aligned} \langle (\Delta \ln P(\vec{x}, \vec{A}_0))^2 \rangle &= \tfrac{1}{\mathcal{N}} \left\{ \int_V d^n x P(\vec{x}, \vec{A}_0) (\ln P(\vec{x}, \vec{A}_0) - \langle \ln P(\vec{x}, \vec{A}_0) \rangle)^2 \right\} \\ &= \tfrac{1}{\mathcal{N}} \left[\int_V d^n x P(\vec{x}, \vec{A}_0) (\ln P(\vec{x}, \vec{A}_0))^2 \right] - \langle \ln P(\vec{x}, \vec{A}_0) \rangle^2. \end{aligned} \tag{4.2.13}$$

This result can be used to form an approximate goodness-of-fit test, when $m << N$: one replaces $\vec{A}_0$ with the best fit parameters $\vec{A}$, and computes log-likelihood. Then, one considers the probability that the Gaussian described above is consistent with the computed log-likelihood. We may then compute the confidence level of a fit that has a maximized log-likelihood of $\ln \mathcal{L}_{max}$ as

$$\mathrm{C.L.} = \frac{1}{2\sqrt{2\pi}} \int_{-\infty}^{\mathrm{x}_0} \mathrm{e}^{\frac{-\mathrm{x}^2}{2}} \mathrm{dx}, \tag{4.2.14}$$

where

$$x_0 = \frac{\ln \mathcal{L}_{max} - \langle \ln \mathcal{L} \rangle}{\sigma_{(\ln \mathcal{L})}}. \tag{4.2.15}$$

When the number of parameters is not negligible, the resulting fit pulls alter the expected distribution of likelihood. We may account for this change in the number of degrees-of-freedom by replacement of N in the above formulae with the factor $N - m$.

4.3 Performing Unbinned Maximum-likelihood Fits

4.3.1 The Probability Distribution

The first step in applying the above prescription is the estimation of the *pdf* at each datapoint, $p(\vec{x}_i, \vec{a})$. In the simplest case one has a physically motivated analytic function for $p(\vec{x}, \vec{a})$ for which the normalization can be easily calculated. In practice it is much more common to have smearing and acceptance effects which cause the observed probability distribution $p(\tilde{\vec{x}}, \vec{a})$, a function of the smeared measured variables $\tilde{\vec{x}}$, to differ from the underlying physical distribution $\rho(\vec{x}, \vec{a})$. The relationship between these two distributions can be expressed as

$$p(\tilde{\vec{x}}, \vec{a}) = \int A(\tilde{\vec{x}}, \vec{x}) S(\tilde{\vec{x}}, \vec{x}) \rho(\vec{x}, \vec{a}) dx, \tag{4.3.1}$$

where $A(\tilde{\vec{x}}, \vec{x})$ is the acceptance function and $S(\tilde{\vec{x}}, \vec{x})$ is the smearing function. Naturally, it is this function of the reconstructed, smeared observables that we need for our likelihood calculation.

In some cases it may be possible to parameterize the smearing and acceptance functions and calculate $p(\tilde{\vec{x}}, \vec{a})$ analytically. One would typically use a Monte Carlo simulation of the the experiment to find these functions. This can be very difficult when there are a large number of correlated observables.

Fortunately, it is possible to use the Monte Carlo data directly to evaluate $p(\tilde{\vec{x}}, \vec{a})$ at every datapoint [19]. To do this we create a volume, V_i, centered about the $i-th$ datapoint,

$\tilde{\vec{x}}_i$, and count the number of Monte Carlo events, N_{V_i}, inside that volume. The probability that the datapoint was drawn from the same parent as the Monte Carlo is then

$$p(\tilde{\vec{x}}_i, \vec{a}) \propto \frac{\sum_j^{N_{V_i}} 1}{V_i} = \frac{N_{V_i}}{V_i}. \tag{4.3.2}$$

To enforce the normalization condition for 'classical' log-likelihood,

$$\int_V d^n x p(\tilde{\vec{x}}, \vec{a}) = 1, \quad \forall \vec{a}, \tag{4.3.3}$$

we can define

$$p(\tilde{\vec{x}}_i, \vec{a}) = \frac{1}{N_{TOT}} \frac{N_{V_i}}{V_i}, \tag{4.3.4}$$

where N_{TOT} is the total number of Monte Carlo events in the sample. Multiplication by a factor of $\mathcal{N}$ gives us the extended probability, $P(\tilde{\vec{x}}_i, \vec{A})$.

If there are several samples of Monte Carlo representing N_{comp} different hypothesized components of the data we have for the extended probability at the $i-th$ datapoint

$$P(\tilde{\vec{x}}_i, \vec{A}) = \sum_{k=1}^{N_{comp}} \mathcal{N}_k p_k(\tilde{\vec{x}}_i, \vec{a_k}), \tag{4.3.5}$$

where the sum is over the various components and the probability distribution for each component at the $i-th$ datapoint is

$$p_k(\tilde{\vec{x}}_i, \vec{a_k}) = \frac{1}{N_{TOT_k}} \frac{N_{V_{ik}}}{V_{ik}}. \tag{4.3.6}$$

Note that the sizes and shapes of the volumes V_{ik} may be chosen differently for each datapoint and sample of Monte Carlo. Our only concern is proper evaluation of the *density*.

4.3.2 Estimation of Shape Parameters

The above prescription is complete if we are only extracting the yield, $\mathcal{N}$, of each Monte Carlo component. However, we would often like to fit other parameters, $\vec{a}$, that describe the shape of some component of the data for an estimate of their 'true' values, $\vec{a_0}$. Examples are the mass and width of a resonance, or the decay angular distributions in a decay sequence. In this case we must model a continuum of hypotheses in order to extract the relevant parameters from the fit. When one has an analytic function for $p(\tilde{\vec{x}}, \vec{a})$ this is simple. When Monte Carlo data is used to generate $p(\tilde{\vec{x}}_i, \vec{a})$ this is more difficult.

What we need is Monte Carlo for each hypothesized shape of $p(\tilde{\vec{x}},\vec{a})$ required in the process of minimizing the likelihood function. Clearly this is unrealistic, but it is possible to weight the events in a Monte Carlo sample to simulate all possible hypotheses [19]. In general, we need $p(\tilde{\vec{x}},\vec{a})$ for each possible, physically meaningful $\rho(\vec{x},\vec{a})$. For this we may take a set of Monte Carlo events generated according to $\rho(\vec{x},\vec{a})$ and reweight them with the function

$$W(\vec{x},\vec{a},\vec{a}\,') = \frac{\rho(\vec{x},\vec{a}\,')}{\rho(\vec{x},\vec{a})}, \tag{4.3.7}$$

where $\vec{a}$ are the parameters used to generate the Monte Carlo and $\vec{a}\,'$ are the parameters we would like to 'simulate'. For the observed probability distribution, we have

$$p(\tilde{\vec{x}},\vec{a}\,') \propto \frac{\sum_j^{N_V} W(\vec{x}_j,\vec{a},\vec{a}\,')}{V}. \tag{4.3.8}$$

The final issue is normalization. The properly normalized *pdf* for each sample at the $i-th$ datapoint is

$$p(\tilde{\vec{x}}_i,\vec{a}) = \frac{1}{\sum_j^{N_{TOT}} W(\vec{x}_j,\vec{a},\vec{a}\,')} \frac{\sum_j^{N_{V_i}} W(\vec{x}_j,\vec{a},\vec{a}\,')}{V_i} = \frac{1}{W_{TOT}}\frac{W_{V_i}}{V_i}, \tag{4.3.9}$$

where W_{V_i} is the total weight inside the volume V_i and W_{TOT} is the total weight of all Monte Carlo events. Note that this is just a generalization of the earlier case where the weight of each Monte Carlo event was unity.

It is important to remember that the weight is a function of the generated, physically meaningful location of each point in the space, $\vec{x}$ and not the observed location, $\tilde{\vec{x}}$. This requires us to keep track of both the generated location and the observed location of each Monte Carlo event.

4.3.3 Calculation of Efficiencies

To calculate the number of events produced by our experiment, the efficiency for each hypothesized sample must be calculated. When the only parameter for a sample in the fit is the yield this is quite simple. Using our Monte Carlo we may take the efficiency as

$$\epsilon = N_{observed}/N_{generated}. \tag{4.3.10}$$

However, when we weight some generic Monte Carlo sample with a function as described in the previous section, we must remember that the efficiency may be a function of the parameters:

$$\epsilon = \epsilon(\vec{a}\,'). \tag{4.3.11}$$

We must be prepared to calculate the efficiency for any set of parameters, which is given by

$$\epsilon = \frac{\sum_{j}^{N_{pass}} W(\vec{x}_j, \vec{a}, \vec{a}\,')}{\sum_{j}^{N_{TOT}} W(\vec{x}_j, \vec{a}, \vec{a}\,')}, \tag{4.3.12}$$

where N_{TOT} is the number of physical events generated at some $\vec{x}$, regardless of whether they were 'observed' and N_{pass} is the number of events actually reconstructed and observed at some $\tilde{\vec{x}}$. Note that performing this calculation means keeping track of $\vec{x}$ for every Monte Carlo event, regardless of whether our experiment would have observed that event or not. In experiments where the efficiency is very low, this may create the need to handle very large data structures. Fortunately this step is not neccessary to the minimization procedure and need only be performed once for each fit.

4.3.4 Calculating Goodness-of-fit

In section 2.3 we described a method for estimating goodness-of-fit for unbinned maximum likelihood. This method requires integrating functions of the *pdf* over the space of observables. When using Monte Carlo directly, the *pdf* is not easily integrable because we only evaluate it at each datapoint, $p(\tilde{\vec{x}}, \vec{a})$. Even when it is possible to parameterize the *pdf* with an analytic function, that function is rarely simple enough to perform the integrals that we need. In most cases, it is necessary to use a Monte Carlo integration technique.

This technique involves sampling the *pdf* at a large number ($N_{thrown} >> N$) of random points in the allowed region for our observables, V_{tot}. By calculating $p(\tilde{\vec{x}}, \vec{a})$ at each point, we may integrate any function of the *pdf* using

$$\int_V d^n x F(pdf) = \frac{V_{tot}}{N_{thrown}} \sum_{i}^{N_{thrown}} F(pdf). \tag{4.3.13}$$

Specifically, we need to evaluate

$$\langle \ln \mathcal{L}_N \rangle = N \langle \ln P(\vec{x}, \vec{A}_0) \rangle - \mathcal{N} \tag{4.3.14}$$

$$\sigma_{(\ln \mathcal{L}_N)} = \sqrt{N\left[\langle(\ln P(\vec{x},\vec{A}_0))^2\rangle - \langle \ln P(\vec{x},\vec{A}_0)\rangle^2\right]}, \tag{4.3.15}$$

where

$$\langle \ln P(\vec{x},\vec{A}_0)\rangle = \frac{1}{\mathcal{N}}\sum_i^{N_{thrown}} P(\vec{x}_i,\vec{A})\ln P(\vec{x}_i,\vec{A}) \tag{4.3.16}$$

$$\langle(\ln P(\vec{x},\vec{A}_0))^2\rangle = \frac{1}{\mathcal{N}}\sum_i^{N_{thrown}} P(\vec{x}_i,\vec{A})(\ln P(\vec{x}_i,\vec{A}))^2. \tag{4.3.17}$$

The time needed to evaluate the *pdf* at such a large number of points will be the limiting factor in the speed of the fitter. While it is necessary that $N_{thrown} >> N$, clever sampling techniques can reduce the number of points required for a given precision by orders of magnitude.

If truly random points are used, the integral converges as $\frac{1}{\sqrt{N_{samp}}}$. There are two methods we use to improve upon this. First, there are a number of sampling schemes that use 'quasi-random' sequences. One of these, Sobel's sequence, converges almost as rapidly as $\frac{1}{N_{thrown}}$, and is easily generalized to any number of dimensions [20]. Second, if we can place our sampling points according to the value of the function we are integrating, we achieve greater precision with fewer points. Since the goal is to sample the *pdf* better than it is sampled by the data, we throw our integration points according to the density of data in our space of observables. By combining these two schemes, we are able to evaluate the integral with high precision in a reasonable time.

4.3.5 Systematic Errors

There are two sources of systematic error associated with the use of Monte Carlo to model the *pdf*:

1. limited Monte Carlo statistics in the volume about each datapoint will lead to large statistical errors in the fit shape.

2. nonlinearity in $p(\tilde{\vec{x}},\vec{a})$ over the volume about each datapoint introduces systematic errors in $p(\tilde{\vec{x}},\vec{a})$.

The former is a matter of simple statistics. A lack of events to effectively sample the *pdf* produces a poor estimate of its true value. The latter is due to the fact that our prescription for calculating the probability density at each datapoint,

$$p_i(\tilde{\vec{x}}, \vec{a}) = \frac{1}{N_{TOT}} \frac{N_{V_i}}{V_i}, \tag{4.3.18}$$

assumes that the value at the datapoint is the same as the average value of $p(\tilde{\vec{x}}, \vec{a})$ in the vicinity of that point. In the case that the datapoint is at the center of the volume, this is the same as assuming that even terms in the distribution of Monte Carlo points are small in the range of V_i.

The only way to reduce the statistical error on the *pdf* is by increasing the number of points in the sampling volumes. While one can simply generate more Monte Carlo, the resources for this may be prohibitive. In other cases where sidebands or off-resonant data can be used instead of Monte Carlo, the available data must suffice. In these situations we must increase the size of our sampling volumes, V_i. Unfortunately, this increases the non-linearities across the box, producing larger systematic shifts in the *pdf*. These shifts affect the fit in two ways:

- systematic over or underestimation of $p(\vec{x}, \vec{a})$ can violate the normalization condition $\int_V d^n x p(\tilde{\vec{x}}, \vec{a}) = 1$, biasing fits with more than one component.
- errors in the shape of $p(\vec{x}, \vec{a})$ can produce incorrect fit results.

In general, the first of these is the more serious condition. Methods of correcting for these problems will be discussed in the following sections.

4.3.6 Corrections to $p(\vec{x}, \vec{a})$

Ensuring Proper Normalization of $p(\vec{x}, \vec{a})$

Proper normalization of the *pdf* for each hypothesized component of the data is critical to the extraction of accurate fit parameters. An error in the normalization upsets the balance between the first and second terms of equation for that component. As a result, one component will be favored over another in maximizing the likelihood.

Previously, we wrote the probability density as derived from Monte Carlo as

$$p(\tilde{\vec{x}}, \vec{a}) = \frac{1}{W_{TOT}} \frac{W_V}{V}. \tag{4.3.19}$$

If we were to use as the $p(\tilde{\vec{x}}, \vec{a})$ value of the probability throughout volume V, as in a binned fit, then the proper normalization of $p(\tilde{\vec{x}}, \vec{a})$,

$$\int_V d^n x p(\tilde{\vec{x}}, \vec{a}) = 1, \quad \forall \vec{a}, \tag{4.3.20}$$

is guaranteed. However, that is not what our unbinned prescription calls for: the value of $p(\tilde{\vec{x}}, \vec{a})$ at each datapoint is taken in the same location in each volume, typically in the center. Because of this, the method may introduce an error in the normalization if $p(\tilde{\vec{x}}, \vec{a})$ is systematically over or underestimated. This can easily occur in the case where large volumes are required to sample the Monte Carlo shape.

We solve this problem by calculating the normalization integral via Monte Carlo integration. We then take this 'brute force' normalization and renormalize $p(\tilde{\vec{x}}, \vec{a})$ with it. This can be done with same numerical integration method as we use for calculating the integrals for our goodness-of-fit test,

$$\begin{aligned} \text{Norm} &= \int_V d^n x p(\tilde{\vec{x}}, \vec{a}) \\ &= \frac{V_{tot}}{N_{thrown}} \sum_i^{N_{thrown}} .p(\tilde{\vec{x}}, \vec{a}) \end{aligned}$$

The normalization calculated in this way may differ from unity due to the systematic effects discussed above. We may then renormalize $p(\tilde{\vec{x}}, \vec{a})$ by taking

$$p(\tilde{\vec{x}}, \vec{a})_{renormalized} = \frac{1}{\text{Norm}} p(\tilde{\vec{x}}, \vec{a}). \tag{4.3.21}$$

Projecting onto Orthogonal Polynomials

The simplest method of calculating our *pdf* to higher order is expansion of $p_i(\tilde{\vec{x}}, \vec{a})$ about each datapoint in sum of multi-dimensional Legendre polynomials:

$$\begin{aligned} p_i(\tilde{\vec{x}}, \vec{a}) &= \sum_{i,j,k\ldots N_{dim}} a_{ijk\ldots N_{dim}} P_{ijk\ldots N_{dim}} \\ &= \sum_{i,j,k\ldots N_{dim}} a_{ijk\ldots N_{dim}} \tfrac{2n_i+1}{2} P_i \tfrac{2n_j+1}{2} P_j \tfrac{2n_k+1}{2} P_k \ldots \tfrac{2n_{N_{dim}}+1}{2} P_{N_{dim}}, \end{aligned} \tag{4.3.22}$$

where the $P_n(x)$ are the one-dimensional Legendre polynomials,

n	$P_n(x)$
0	1
1	x
2	$\frac{1}{2}(3x^2-1)$
$\vdots$	$\vdots$

and the coefficients are calculated via projection of the Monte Carlo:

$$a_{ijk\ldots N_{dim}} = \int_V d^n x P_{ijk\ldots N_{dim}} \sum_{1}^{N_V} \delta(\tilde{\vec{x}}_i - x_{center}) W_i(\vec{x}, \vec{a}, \vec{a}\,') \tag{4.3.23}$$

$$= \sum_{1}^{N_V} P_{ijk\ldots N_{dim}}(\tilde{\vec{x}}_i - x_{center}) W_i(\vec{x}, \vec{a}, \vec{a}\,'). \tag{4.3.24}$$

Though this method may be used to any order, it has some shortcomings. Enlargement of the sampling volumes increases the number Monte Carlo points in each volume, but more are required to produce stable solutions for the $a_{ijk\ldots N_{dim}}$. Fluctuations in these solutions can cause the $p_i(\tilde{\vec{x}}, \vec{a})$ calculated in this way to be negative. This is especially true near the edges of the space. There, $p(\vec{x}, \vec{a})$ must be evaluated at the edges of our sampling volumes where the Legendre expansion is not convergent.

To handle the above problems we need a way to deal with negative probability densities, especially near the edges of our multi-dimensional fit space. A small number of problematic datapoints may simply be removed from our fit. If there are large regions of the multi-dimensional space, e.g. near the edges, where $p_i(\tilde{\vec{x}}, \vec{a})$ is computed to be negative, cutting the data in these regions can be a useful approach.

In achieving stability, it is important to calculate only those corrections that appear to be significant using goodness-of-fit as a guide. Because of the added complexity, it is best to avoid making any corrections at all when possible. In many cases, the above method provides only a modest improvement over the linear approximation with carefully tuned volumes. For our fits to the data, we model the *pdf* to linear order, and cut away the edges of the fit space to ensure stability, as described in the following section.

4.4 Fitting the Data for $B^- \to D^{\star +}\, \pi^-\, \pi^-$

4.4.1 Monte Carlo Samples

Large data samples are required to model the probability density function (*pdf*) in four dimensions for the fit. The contributions to the fit are from

- $B^- \to D^0_J\, \pi^-$ signal,
- $B^- \to D^{\star +}\, \pi^-\, \pi^-$ signal,
- mis-reconstructed $B^- \to D^0_J\, \pi^-$,
- $\overline{B^0} \to D^+_J\, \pi^-$,
- $\overline{B^0} \to D^{\star +}\, \pi^-$,
- $\overline{B^0} \to D^{\star +}\, \rho^-$,
- other $B\overline{B}$ generic decays,
- continuum.

The samples used to model these fit components are summarized in Table 4.1. The $B^- \to$

Table 4.1: The samples used to model the *pdf* of each component in the fit. The tagger was used to remove contributions from the exclusive $B\overline{B}$ modes from the $B\overline{B}$ generic Monte Carlo sample.

Component	Sample type	Amount
$B^- \to D^0_J\, \pi^-$ and $B^- \to D^{\star +}\, \pi^-\, \pi^-$	$D^{\star +}\pi^-\pi^-$ MC	3,000,000 events
$\overline{B^0} \to D^+_J\, \pi^-$	$D^{\star +}\pi^-\pi^0$ MC	1,000,000 events
$\overline{B^0} \to D^{\star +}\, \pi^-$	$\overline{B^0} \to D^{\star +}\, \pi^-$ MC	400,000 events
$\overline{B^0} \to D^{\star +}\, \rho^-$	$\overline{B^0} \to D^{\star +}\, \rho^-$ MC	800,000 events
Other $B\overline{B}$ background	$B\overline{B}$ generic MC	$\approx 2 \times 10^7$ events
Continuum background	Off $\Upsilon(4S)$ data	1.6 fb^{-1}

$D^{\star +}\, \pi^-\, \pi^-$ Monte Carlo was used for modeling both the resonant and non-resonant signal. This sample was generated with a flat distribution in Dalitz space and each of the angular distributions, multiplied by a factor producing more Monte Carlo in the $M(D^{\star +}\pi^-)$ region where we know the D_J lie. Monte Carlo for modeling the contributions from $\overline{B^0} \to D_J^+ \pi^-$ feedthroughs are generated with the same distribution.

The shapes of the $B^- \to D_J^0 \pi^-$ and $\overline{B^0} \to D_J^+ \pi^-$ signals that contribute to background from mis-reconstruction are varied in unison with the changes in the resonant signal shape. The resonant signal shape is assumed to be the same for $B^- \to D_J^0 \pi^-$ and $\overline{B^0} \to D_J^+ \pi^-$. The yields of these backgrounds are allowed to float in the fit.

Because the sample of off-resonant data has limited statistics, an additional trick is used. Both quadratic solutions from the partial reconstruction are used to model the *pdf*. Since the constraints of the reconstruction have no meaning without a real B decay, there is no difference between the distributions of the two solutions for continuum background. Since the two solutions do not typically fall inside the same volume used to estimate the *pdf*, the statistical significance of the sample in each volume is enhanced.

The contribution from continuum was allowed to float within the gaussian errors due to reconstruction efficiency for off-resonant events. The contributions from $\overline{B^0} \to D^{\star +} \pi^-$ and $\overline{B^0} \to D^{\star +} \rho^-$ were allowed to float within the error due to their branching ratios and the reconstruction efficiencies for those samples. The expected yields from these sources are shown in Table 4.2. The contribution from other $B\overline{B}$ decays was allowed to float without constraint.

Table 4.2: The expected yields and errors for known backgrounds in the data sample. These yields were allowed to float within their gaussian errors for the fit.

Component	Expected Yield
Continuum background	4007 ± 281
$\overline{B^0} \to D^{\star +} \pi^-$	234 ± 19
$\overline{B^0} \to D^{\star +} \rho^-$	614 ± 307

4.4.2 Tests of Fit for $B^- \to D^{\star +} \pi^- \pi^-$ Decays

The fit technique was tested on eight samples of simulated data. Data for the expected background were drawn from the Monte Carlo samples described in the previous section. Continuum Monte Carlo was used to draw samples representing the off-resonant data. Eight signal samples were generated with different values for each of the parameters.

The set of parameters for each test was randomly selected from predetermined pools for each parameter.

All components of the fit are modeled to linear order by expansion in Legendre polynomials. This eliminates effects at the edges of the four-dimensional space of $M(D^{\star +}\pi^-)$ and the decay angles that are inherent in our Monte-Carlo based technique for estimation of the *pdf*. Similar effects at the edges of the Dalitz space were found to cause systematic shifts during testing with Monte Carlo samples. Cutting the data at $cos\theta_{Dalitz} < 0.99$ in all $D^{\star +}\pi^-\pi^-$ Dalitz angles eliminates these errors. This cut is 96% efficient for signal and 89% efficient for background.

Figures 4.1 and 4.2 show plots of the reconstructed versus generated value for each parameter. No significant systematic shifts are apparent in these results. These tests also revealed that our fit is not sensitive to the strong phase between the resonant and non-resonant contributions to $D^{\star +}\pi^-\pi^-$. This phase was removed from our parameterization of the fit function for the results of Figures 4.1 and 4.2 and for our fit to the data. The Monte Carlo tests show that neglecting this phase does not introduce any significant systematic shifts in the parameter estimates.

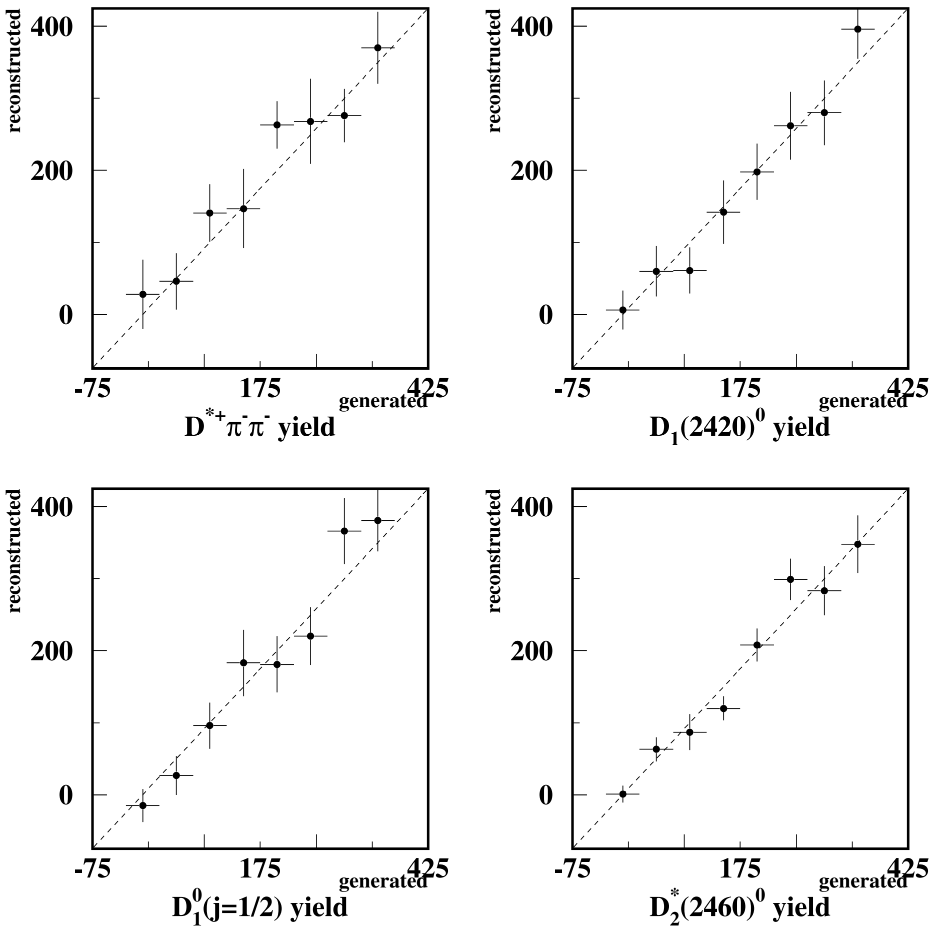


Figure 4.1: Reconstructed versus generated yields for eight Monte Carlo tests with different signal distributions. The dashed line represents the perfect fit.

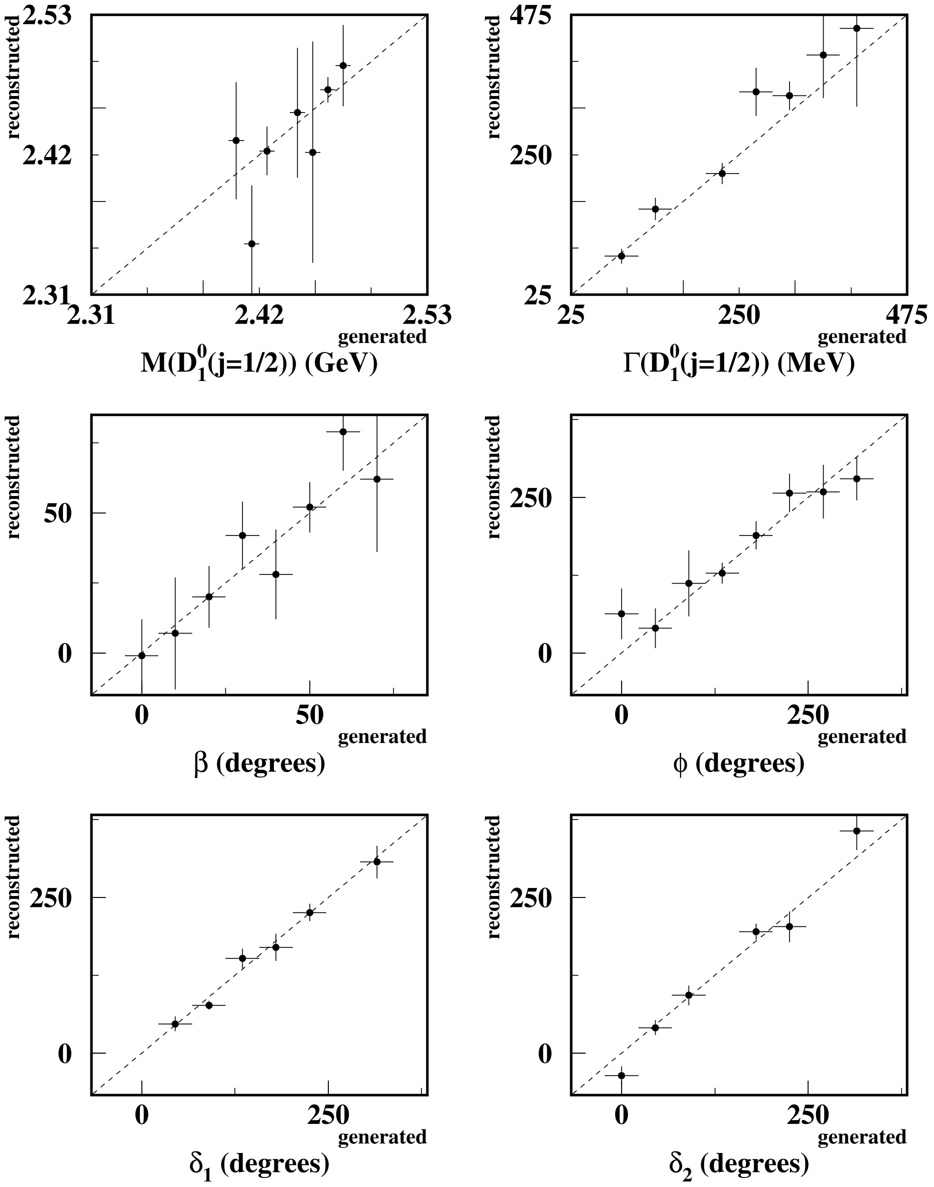


Figure 4.2: Reconstructed versus generated parameters for eight Monte Carlo tests with different signal distributions. The dashed line represents the perfect fit. Missing points correspond to cases where a non-existent resonant yield leads to large errors or no clear minimum for that parameter.

Chapter 5

Results

The fit to the data was performed exactly as described for the Monte Carlo tests described in Chapter 4. The results of the four-dimensional fit to the data are summarized in Tables 5.1 and 5.2. For this fit, the mass and width of the $D_1^0(j=\frac{1}{2})$ are allowed to float, while the masses and widths of the $D_1(2420)^0$ and $D_2^\star(2460)^0$ are fixed to their known values from Table 1.1 [2].

Table 5.1: Results of the fit to the three body phase space for $B^- \to D^{\star +}\,\pi^-\,\pi^-$ using 3.1 fb^{-1} of data collected at the $\Upsilon(4S)$. The yields of all components with errors are shown. There are 9014 reconstructed events in the data sample.

Component	Yield
$B^- \to D_J^0\,\pi^-$ signal	631.5 ± 80
$B^- \to D^{\star +}\,\pi^-\,\pi^-$ non-res signal	199 ± 74
Continuum background	4025 ± 157
$\overline{B^0} \to D^{\star +}\,\pi^-$	222 ± 17
$\overline{B^0} \to D^{\star +}\,\rho^-$	714 ± 77
mis-reconstructed $B^- \to D_J^0\,\pi^-$	308 ± 78
$\overline{B^0} \to D_J^+\,\pi^-$	470 ± 96
other $B\overline{B}$ background	2448 ± 151
Total	9018 ± 95

The net (integrated) interference among the resonances is found to be destructive.

Table 5.2: Details of the fit to the data for the signal distribution. Yields in the resonant sub-channels are corrected for their relative efficiencies. The statistical significance with which each parameter is determined to have a non-trivial value is also listed. The significance for δ_1 is measured relative to both 0 and (π), while that for δ_2 is measured relative to both 0 and ($\pi/2$), as described in the text.

Parameter	Central value	Statistical error	Significance $\left(\sqrt{2(\Delta \ln \mathcal{L})}\right)$
$D^{\star}\pi^{-}\pi^{-}$ Yields			
$B^{-} \to D_1(2420)^0\, \pi^{-}$	223.5	$^{+58}_{-44}$	$4.7\,\sigma$
$B^{-} \to D_1^0(j=\frac{1}{2})\, \pi^{-}$	344.4	± 63	$6.1\,\sigma$
$B^{-} \to D_2^{\star}(2460)^0\, \pi^{-}$	100.1	± 27	$5.0\,\sigma$
$B^{-} \to D^{\star +}\, \pi^{-}\, \pi^{-}$ non-res	199	± 74	$2.9\,\sigma$
$D_1^0(j=\frac{1}{2})$ lineshape			
$M_{D_1^0(j=\frac{1}{2})}$ (GeV)	2.461	$^{+0.041}_{-0.034}$	—
$\Gamma_{D_1^0(j=\frac{1}{2})}$ (GeV)	0.290	$^{+0.101}_{-0.079}$	—
Mixing and interference			
β	0.368	± 0.100	$2.9\,\sigma$
ϕ	-0.244	± 0.42	$< 1\,\sigma$
δ_1	2.62	± 0.30	$4.8(1.7)\,\sigma$
δ_2	2.10	± 0.34	$2.3(1.6)\,\sigma$

The sum of the individual product branching ratios would yield 1.058 ± 0.063 times the observed total resonant contribution to the $D^{\star+}\pi^-\pi^-$ final state.

The statistical significance of the signal yields, mixing and interference are also shown in Table 5.2. The statistical significance of each fit parameter is $\sqrt{2\Delta \ln \mathcal{L}}$, where the change in log-likelihood is obtained by performing a fit with that parameter fixed to zero or another trivial value. Because the estimated value of δ_1 is close to π where it has no imaginary part, the significance of δ_1 is determined relative to both 0 and π. Similarly, δ_2 is found to be close to $\pi/2$ where the imaginary part cancels the i in the angular amplitude of the 2^+ decay. The significance of δ_2 is listed relative to both 0 and $\pi/2$.

With the help of Eqs. 21-23 and the covariance matrix from the fit, we may convert our measurements of $M_{D_1^0(j=\frac{1}{2})}$, $\Gamma_{D_1^0(j=\frac{1}{2})}$, β and ϕ into measurements of ΔM, $\Delta\Gamma$, M_{12} and Γ_{12}. The results for the mixing parameters are shown in Table 5.3.

Table 5.3: Estimates of the mixing parameters for the 1^+ D_J. The errors on these values are computed using the covariance matrix from the fit.

Mixing parameter	Estimate (MeV)
ΔM	-9.6 ± 34.3
$\Delta\Gamma$	-190 ± 76
M_{12}	22.2 ± 21.2
Γ_{12}	98.9 ± 33.7

The fit was also performed with the masses and widths of the $D_1(2420)^0$ and $D_2^\star(2460)^0$ floating within the errors on their known values [2]. The resulting estimates of these parameters were all within one standard deviation of the known values. Finally, these constraints on the lineshapes were removed entirely. This yields a consistent fit as well, but with much larger errors.

A projection of the fit onto the $M(D^{\star+}\pi^-)$ axis is shown in Figure 5.1. The demonstration that the the resonant subcomponents of the fit are present in the data is challenging because of substantial overlaps of the angular functions. To attempt a meaningful comparison, we have made projections onto the $M(D^{\star+}\pi^-)$ axis by weighting the events, where the weights are determined by the *decay angles* alone. If, when we weight the events based only on the expected decay angular distribution for the 2^+ resonance, we observe a peak with the characteristics of the $D_2^\star(2460)^0$, then we have meaningfully tested our fit.

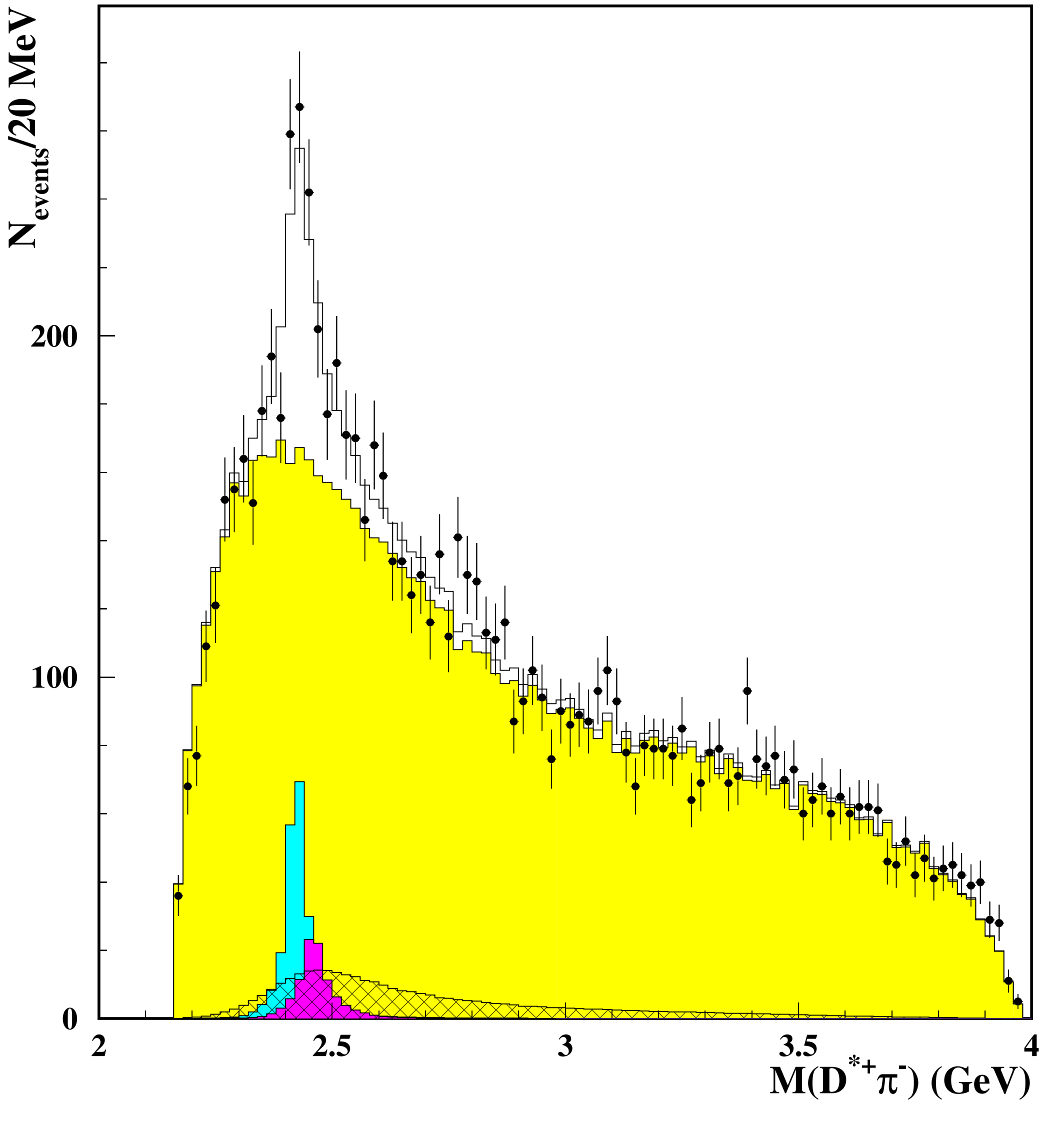


Figure 5.1: A projection of the four-dimensional fit onto the $M(D^{\star +}\pi^{-})$ axis. The points with errors are the data. The expected background and the fit are shown, along with the fractions of $D_1(2420)^0$ (light gray), $D_2^{\star}(2460)^0$ (dark gray) and $D_1^0(j=\frac{1}{2})$ (cross-hatched), where interference has been neglected.

To make these projections, four weights are constructed based on the expected angular distributions of the various partial waves and the background. These weights are normalized to have the same maximum value and are proportional to

$$\begin{aligned} w_{1d} &\propto |a_{1d}|^2 && (5.0.1)\\ w_{1s} &\propto |a_{1s}|^2 && (5.0.2)\\ w_2 &\propto |a_2|^2 && (5.0.3)\\ w_{back} &\propto f(\cos\theta_2, \cos\theta_3, \chi), && (5.0.4) \end{aligned}$$

where the a are the angular amplitudes of section 1.2 and w_{back} is a background weight constructed by fitting the one-dimensional distributions of the backgrounds in the signal region (2.4 GeV $< M(D^{\star+}\pi^-) <$ 2.9 GeV) to analytic functions. From these we construct four more weights that emphasize the $S^2/(S+B)$ of each partial wave:

$$\begin{aligned} W_{1d} &= \frac{w_{1d}^2}{w_{tot}} && (5.0.5)\\ W_{1s} &= \frac{w_{1s}^2}{w_{tot}} && (5.0.6)\\ W_2 &= \frac{w_2^2}{w_{tot}} && (5.0.7)\\ W_{back} &= \frac{w_{back}^2}{w_{tot}}, && (5.0.8) \end{aligned}$$

where

$$w_{tot} = w_{1d} + w_{1s} + w_2 + w_{back}. \quad (5.0.9)$$

As defined, these weights are not sensitive to the *level* of any resonant sub-component that contributes to the fit.

Using these weights, we make $M(D^{\star+}\pi)$ plots where events are weighted by the 1^+ **d**-wave, 1^+ **s**-wave, 2^+ and background angular information. These plots are displayed in Figure 5.2. The data show the enhancements expected due to the three resonances. Subtracting the backgrounds from these plots, as in Figure 5.3, shows the differences between the data for the three signal weights more clearly.

The confidence level of the fit was evaluated using the method described in CBX 98-61. The expected distribution $\ln\mathcal{L}_{\max}$ is gaussian, and is given specifically by:

$$\exp\left(\frac{-(\ln\mathcal{L}_{\max} - 43841.73)^2}{2\times 68.30^2}\right). \quad (5.0.10)$$

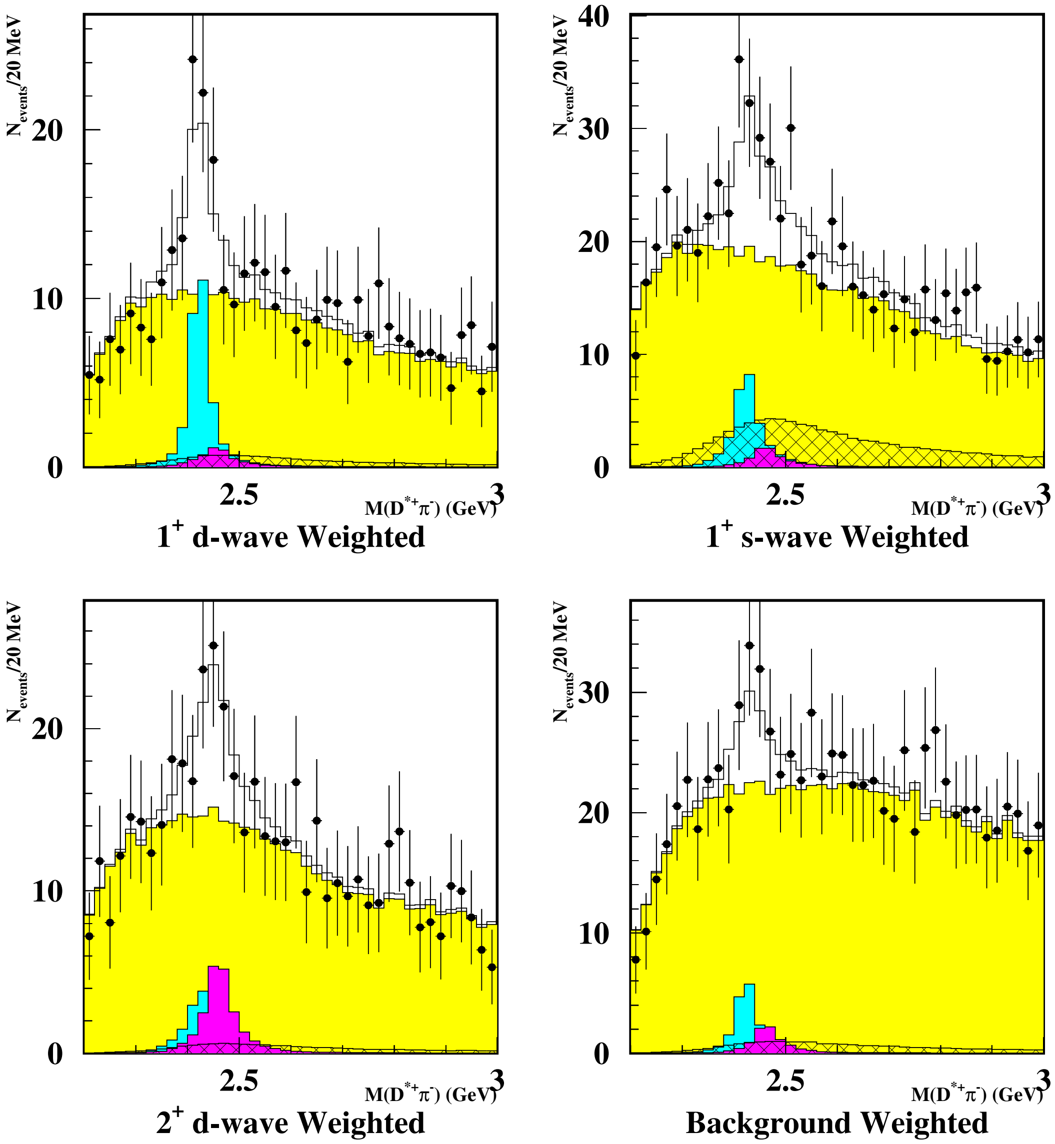


Figure 5.2: Weighted projections of the four-dimensional fit onto the $M(D^{\star+}\pi^-)$ axis. Each event is weighted according to the expected angular distributions for the three resonant partial waves and the background. The points with errors are the data. The expected background and the fit are shown, along with the un-interfered fractions of $D_1(2420)^0$ (light gray), $D_2^\star(2460)^0$ (dark gray) and $D_1^0(j=\frac{1}{2})$ (cross-hatched).

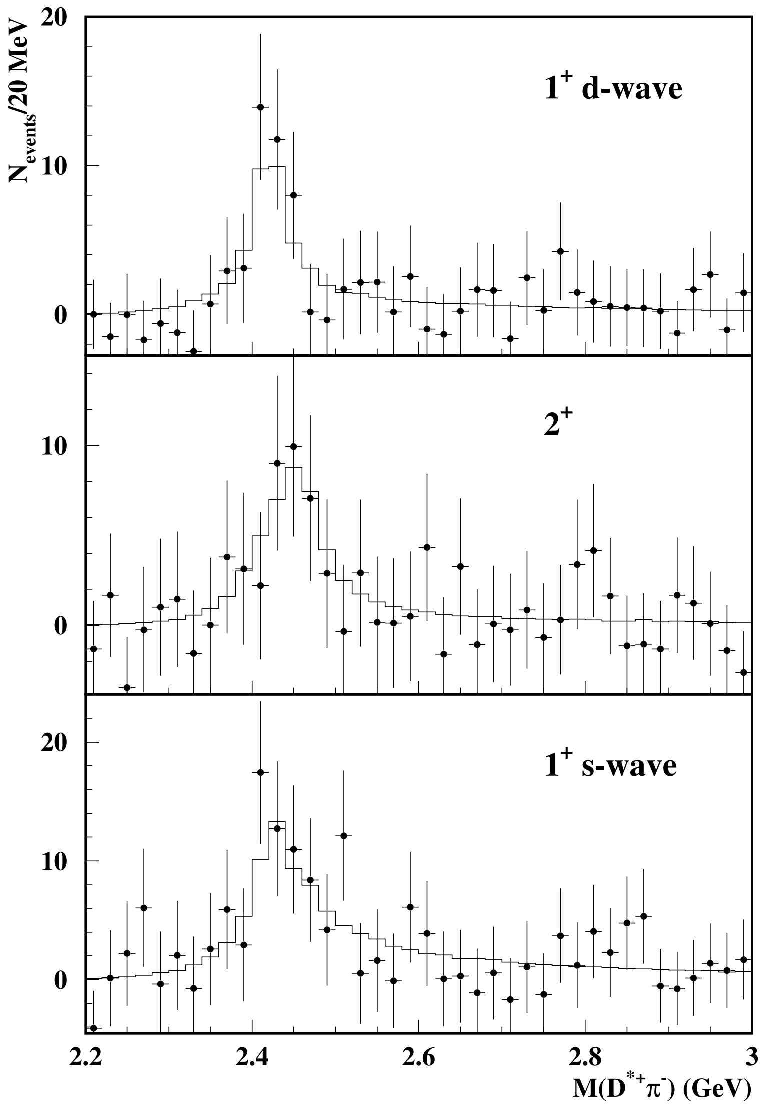


Figure 5.3: Background-subtracted weighted projections of the four-dimensional fit onto the $M(D^{\star+}\pi^-)$ axis. The differences among the distributions are more clearly visible than in Figure 5.2.

The log-likelihood of our fit is 43916.10; thus, the fit is about one sigma more likely than the typical fit. The relationship between the log-likelihood of this fit and the expected distribution is shown graphically in Figure 5.4. Evaluation of the probability integral results in a confidence level of 86% for the fit.

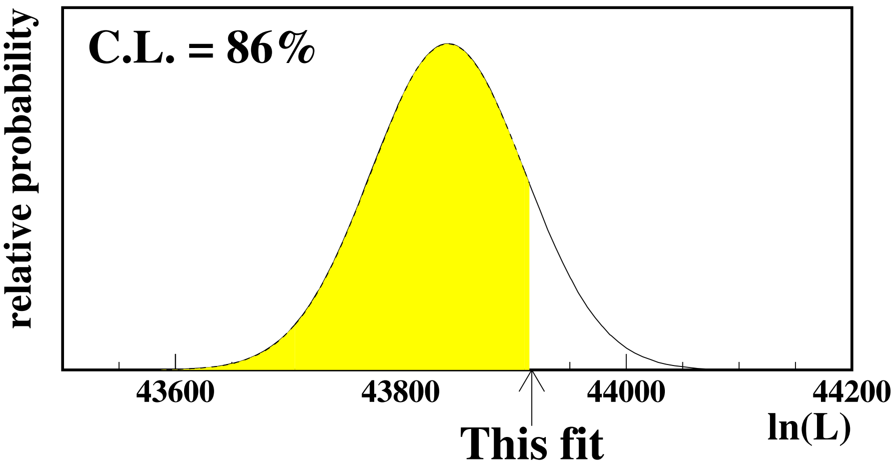


Figure 5.4: The expected distribution of $\ln \mathcal{L}$ and the maximum $\ln \mathcal{L}$ for the fit to the data.

5.1 Consistency Checks

Two important consistency checks have been performed. The first involves generation of Monte Carlo samples according to the parameters found by the fit to the data. Ten complete simulations of the data were drawn from the Monte Carlo samples shown in Table 4.1. CLEO continuum Monte Carlo was used to draw samples for the continuum background. The results of these fits are summarized in Table 5.4. A plot showing the distribution of $\ln \mathcal{L}_{\max}$ for these fits along with the expected and actual $\ln \mathcal{L}_{\max}$ for the fit to the data is shown in Figure 5.5. The apparent shift in likelihood distributions is attributed to small differences in the observed distributions for continuum Monte Carlo and the off-resonant data.

Table 5.4: Results of fits to ten Monte Carlo samples generated with the estimated parameters from the fit to the data. N_g is the generated value for each parameter. $\overline{N_f}$ and $\overline{\sigma_f}$ are the mean estimated value and error for each parameter from the fit. $\sigma_{\overline{N_f}}$ is the error on the mean estimated value for the ten fits. $\overline{\chi^2}$ is average of the chi-squared discrepancy for each parameter over the ten fits.

Parameter	N_g	$\overline{N_f}$	$\overline{\sigma_f}$	$\sigma_{\overline{N_f}}$	$\overline{\chi^2}$
Background yields					
Continuum yield	4025	4077	165	53	0.96
$\overline{B^0} \to D^{\star +}\,\pi^-$	222	229	16	5.7	0.89
$\overline{B^0} \to D^{\star +}\,\rho^-$	714	687	66	22	0.99
$B^- \to D^0_J\,\pi^-$ mis-recon	308	296	75	24	1.13
$\overline{B^0} \to D^+_J\,\pi^-$	470	420	102	35	1.24
Other $B\overline{B}$ background	2448	2420	142	50	1.08
$B^- \to D^0_J\,\pi^-$ signal	632	640	72	22	1.10
Signal yields					
$B^- \to D_1(2420)^0\,\pi^-$	223	225	53	16	1.08
$B^- \to D^0_1(j=\frac{1}{2})\,\pi^-$	344	330	65	22	0.97
$B^- \to D^\star_2(2460)^0\,\pi^-$	100	102	24	8.2	0.87
$B^- \to D^{\star +}\,\pi^-\,\pi^-$ non-res	199	227	62	18	1.17
$D^0_1(j=\frac{1}{2})$ lineshape					
$M_{D^0_1(j=\frac{1}{2})}$ (GeV)	2.461	2.477	0.039	0.017	1.06
$\Gamma_{D^0_1(j=\frac{1}{2})}$ (GeV)	0.290	0.279	0.085	0.025	0.92
Mixing and interference					
β	0.368	0.361	0.093	0.03	1.00
ϕ	-0.244	-0.268	0.53	0.18	1.07
δ_1	2.62	2.58	0.32	0.10	0.82
δ_2	2.10	2.08	0.32	0.13	0.96

The second check was a fit to the wrong-sign distribution in data corresponding to the final state $D^{\star +}\pi^+\pi^-$. Since B^+ decays to this final state are highly suppressed, the reconstructed mass of the candidate $D^\star\pi$ produces distributions that are similar to those for right-sign background. Fits were performed to this sample using the expected wrong-sign distributions for the backgrounds and the right-sign distributions for the signals. The fit was

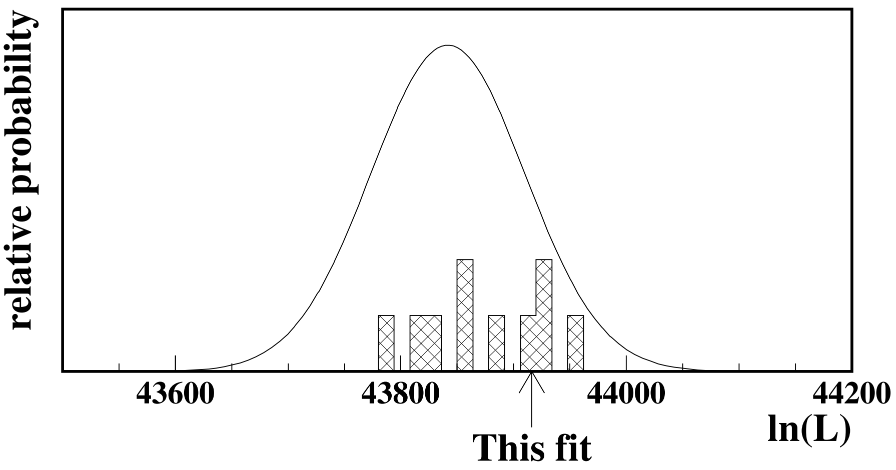


Figure 5.5: A histogram of $\ln \mathcal{L}_{\max}$ from fits to ten Monte Carlo samples generated with the estimated parameters from the fit to the data. The expected distribution and the $\ln \mathcal{L}_{\max}$ for the fit to the data are also shown. The small shift between the two distributions is attributed to the observed difference between the off-resonant data and the continuum Monte Carlo.

first minimized with the signal distribution fixed to the right-sign result. The fit was also minimized with the signal distribution floating as it was for the right-sign fit. The estimated background yields were consistent with those from the right-sign fit while the yields in the signal distributions were consistent with zero for both fits.

With the signal distribution floating from various initial configurations, three different local minima were found with similar $\ln \mathcal{L}$. No convincing minima were found for the mixing and interference parameters. Table 5.5 shows the estimated parameters for the most likely wrong-sign fit. The confidence level of this fit is 95%. The $(\Delta\chi^2/DOF)$ between the expected fit and the most likely fit is 1.16. A projection of this fit onto the $M(D^\star\pi)$ axis is shown in Figure 5.6.

Table 5.5: The results of the fit to the the wrong-sign data as described in the text.

Parameter	Expected value	Estimated value
Background yields		
Continuum yield	9232 ± 380	8863 ± 256
$\overline{B^0} \to D^{\star +}\pi^-$	222 ± 17	246 ± 18
$\overline{B^0} \to D^{\star +}\rho^-$	725 ± 78	917 ± 88
$B^- \to D_J^0\pi^-$ mis-recon	862 ± 218	908 ± 116
$\overline{B^0} \to D_J^+\pi^-$	495 ± 101	422 ± 107
Other $B\overline{B}$ background	4399 ± 241	4771 ± 260
$B^- \to D_J^0\pi^-$ signal	0	52 ± 67
$D^\star\pi^-\pi^-$ Yields		
$B^- \to D_1(2420)^0\pi^-$	0	6 ± 13
$B^- \to D_1^0(j=\frac{1}{2})\pi^-$	0	43 ± 65
$B^- \to D_2^\star(2460)^0\pi^-$	0	4 ± 9
$B^- \to D^{\star +}\pi^-\pi^-$ non-res	0	94 ± 90
$D_1^0(j=\frac{1}{2})$ lineshape		
$M_{D_1^0(j=\frac{1}{2})}$ (GeV)	—	2.907 ± 0.190
$\Gamma_{D_1^0(j=\frac{1}{2})}$ (GeV)	—	0.148 ± 0.075

It is worth noting that ARGUS used this wrong-sign distribution to model the background shape in a similar analysis of the narrow resonances with low statistics [21]. Examination of the right- and wrong-sign distributions for off-resonant data shows that

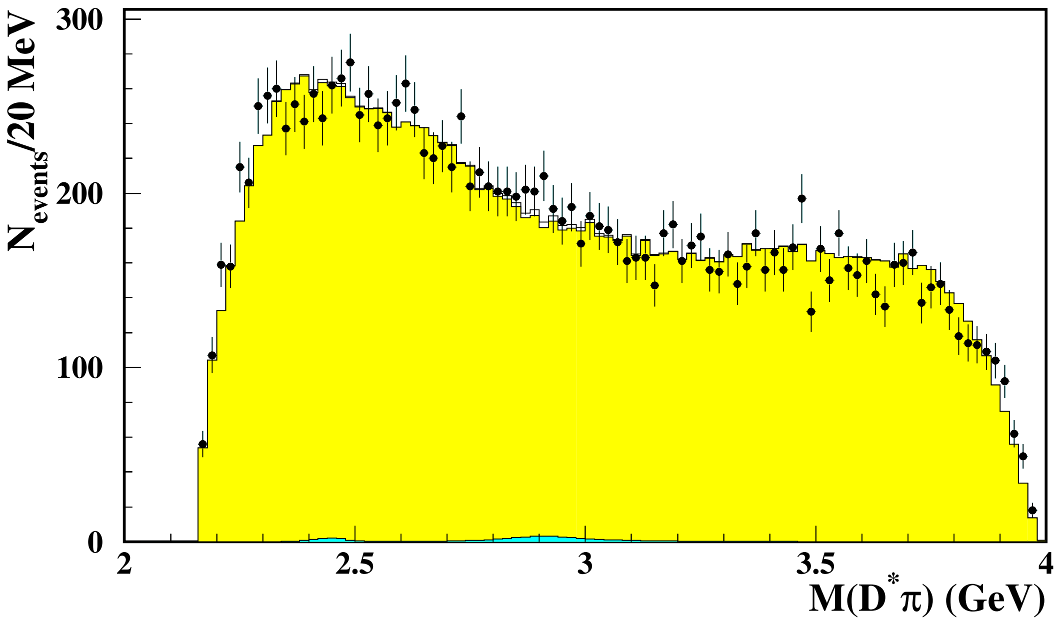


Figure 5.6: A projection of the wrong-sign fit onto the $M(D^{\star}\pi)$ axis. The small estimated contribution from the resonant signal lineshape is shown.

charge correlations among the pions produces somewhat different distributions. A much larger effect is caused by the fact that random backgrounds scale up by a factor of two from right-sign to wrong-sign while correlated feedthroughs such as $\overline{B^0} \to D^{\star +} \pi^-$ and $\overline{B^0} \to D^{\star +} \rho^-$ contribute almost equally in both signs. Had we used this distribution to model our background shape in the fit, the yield of the broad resonance would have been much larger as can be readily seen by comparison of Figures 5.1 and 5.6.

5.2 Systematic Errors

Systematic errors from several sources must be considered. Estimates of these errors are summarized in Tables 5.6, 5.7 and 5.8.

Table 5.6: Systematic errors on the total and non-resonant branching fractions for $B^- \to D^{\star +} \pi^- \pi^-$. The first six errors alter the branching fraction through changes in efficiency. The remaining seven effect the estimated yields through changes in the signal and background shapes.

	$D^{\star+}\pi^-\pi^-$	$D^{\star+}\pi^-\pi^-$ non-resonant
Luminosity	2.0%	2.0%
$\mathcal{B}(D^{\star +} \to D^0 \pi^+)$[22]	2.0%	2.0%
R_2 cut	3.1%	2.8%
Beam energy	4.6%	3.7%
π_{slow} parameters	1.0%	1.0%
$\pi^\pm$ reconstruction	8.0%	8.0%
$\pi^\pm$ ID	4.0%	4.0%
$\epsilon_{\pi_{\mathrm{slow}}}$ slope	3.1%	7.7%
Background shape	4.1%	7.2%
$M_{D_1(2420)^0}$	0.1%	0.5%
$\Gamma_{D_1(2420)^0}$	0.6%	0.5%
$M_{D_2^\star(2460)^0}$	0.1%	0.5%
$\Gamma_{D_2^\star(2460)^0}$	0.4%	1.0%
Fit method	2.4%	4.7%
Total	12.4%	15.7%

An error in reconstruction efficiency may be introduced by poor modeling of the event shape in the Monte Carlo. The variation in fit results for different R_2 cuts in the range

Table 5.7: Systematic errors on the product branching fractions for $B^- \to D_J^0\,\pi^-$, $D_J^0 \to D^{\star+}\,\pi^-$. The first six errors alter the branching fraction through changes in efficiency. The remaining seven effect the estimated yields through changes in the signal and background shapes.

$B^-\to D_J^0\,\pi^-, D_J^0\to D^{\star+}\,\pi^-$			
	$D_1(2420)^0$	$D_1^0(j=\frac{1}{2})$	$D_2^\star(2460)^0$
Luminosity	2.0%	2.0%	2.0%
$\mathcal{B}(D^{\star+}\to D^0\,\pi^+)$[22]	2.0%	2.0%	2.0%
R_2 cut	3.5%	2.3%	3.2%
Beam energy	3.2%	4.8%	3.5%
$\pi_{\rm slow}$ parameters	1.0%	1.0%	1.0%
$\pi^\pm$ reconstruction	8.0%	8.0%	8.0%
$\pi^\pm$ ID	4.0%	4.0%	4.0%
$\epsilon_{\pi_{\rm slow}}$ slope	5.4%	6.8%	4.7%
Background shape	4.6%	7.1%	4.6%
$M_{D_1(2420)^0}$	2.2%	4.7%	1.0%
$\Gamma_{D_1(2420)^0}$	10.3%	1.7%	2.0%
$M_{D_2^\star(2460)^0}$	1.3%	0.3%	1.0%
$\Gamma_{D_2^\star(2460)^0}$	0.5%	0.6%	7.0%
Fit method	2.1%	4.2%	2.0%
Total	16.6%	16.0%	14.6%

Table 5.8: Systematic errors on the parameters describing the lineshape, mixing and interference of the D_J.

	$M_{D_1^0(j=\frac{1}{2})}$	$\Gamma_{D_1^0(j=\frac{1}{2})}$	β	ϕ	δ_1	δ_2
$\pi_{\rm slow}$ efficiency slope	1.1%	2.9%	3.0%	4.2%	0.6%	0.9%
Background shape	0.3%	6.7%	5.4%	23%	1.4%	3.0%
$M_{D_1(2420)^0}$	0.2%	2.8%	6.8%	48%	2.2%	5.5%
$\Gamma_{D_1(2420)^0}$	0.5%	1.7%	4.1%	27%	1.7%	1.9%
$M_{D_2^\star(2460)^0}$	—	0.3%	0.6%	2.9%	0.4%	2.9%
$\Gamma_{D_2^\star(2460)^0}$	—	2.1%	0.6%	9.0%	—	0.3%
Fit method	3.6%	3.7%	2.7%	12.5%	5.2%	4.3%
Total	3.8%	9.1%	10.5%	62%	6.1%	8.4%

$0.20 < R_2 < 0.30$ provides the estimates of errors from this cut shown in Tables 5.6 and 5.7.

Mis-estimation of the beam energy and the $\pi_{\rm slow}$ tracking parameters produce uncertainties in the efficiency for our reconstruction method. The error in signal efficiency from E_{beam} is estimated by variation within known errors for each dataset. The loss of efficiency due to mis-measurement of the slow pion is 4%. Assuming this loss is modeled correctly to 25% in the Monte Carlo, we estimate the error from this source at 1%.

The correlated errors on efficiency for reconstruction of the three pions are roughly 1%, 2% and 5% for the fast, middle and slow pions from $B^- \to D_J^0\,\pi^-$ decays. The sum of 8% is also an accurate estimate for non-resonant $B^- \to D^{\star +}\,\pi^-\,\pi^-$ decays. The pion consistency cuts also introduce efficiency errors that add linearly. The error due to these cuts is roughly 1.3% for each of the three pions.

There is an uncertainty in the slope of the reconstruction efficiency for charged tracks between 50 and 150 MeV. It is estimated that the error on the relative efficiency at 50 and 150 MeV is 10%. The effect of this uncertainty was modeled for the extreme cases by re-weighting the off-resonant and Monte Carlo samples with the functions

$$F(p) = 1 \pm 0.1\frac{150 - p}{100} \tag{5.2.1}$$

applied to all pions in this momentum region and refitting the data.

Uncertainty in the lineshapes for the narrow $D_1(2420)^0$ and $D_2^\star(2460)^0$ are an important source of systematic error. Both the mass and width of these resonances were varied within their known errors [2] to estimate the uncertainty from this source.

Systematic error due to uncertainty in the background shapes is estimated by variation of each fraction within the known limits, based on errors in reconstruction efficiencies and branching fractions. The systematic error from the fit technique is estimated by varying the number of Monte Carlo points used to estimate the *pdf* by a factor of two.

An attempt was made to include the effects of these systematic errors in the measure of statistical significance. The factors responsible for systematic changes in the *pdf* were varied coherently to minimize the significance of an individual parameter. The significance of that parameter was then retested. Repeating this process for each parameter produces the results in Table 5.9.

Table 5.9: Overall significances of the yields, and the mixing and interference parameters of the resonances.

Parameter	Significance
Yields	
$B^- \to D_1(2420)^0\,\pi^-$	4.6σ
$B^- \to D_1^0(j = \frac{1}{2})\,\pi^-$	5.7σ
$B^- \to D_2^\star(2460)^0\,\pi^-$	4.6σ
$B^- \to D^{\star +}\,\pi^-\,\pi^-$ non-res	2.1σ
Mixing and interference	
β	1.7σ
ϕ	$< 1\sigma$
δ_1	$4.0(1.2)\sigma$
δ_2	$2.0(1.3)\sigma$

5.3 Spin-parity Assignment for $D_1^0(j = \frac{1}{2})$

We have tested the $J^P = 1^+$ hypothesis for the spin-parity of the broad resonance. The 0^+ assignment is forbidden by parity conservation, but the sequence of assignments 0^-, 1^- and 2^- could all decay to $D^{\star +}\pi$ via a **p**-wave intermediate state. Additionally, the 2^+ assignment could decay via the **d**-wave intermediate state. The respective angular amplitudes in the decay angles for these cases are (**f**-wave contributions are neglected)

$$\begin{aligned}
a_{0^-}(\cos\theta_2, \cos\theta_3, \chi) &= -\ \cos\theta_3 \\
a_{1^-}(\cos\theta_2, \cos\theta_3, \chi) &= i\ \sin\theta_2 \sin\theta_3 \sin\chi \\
a_{2^-}(\cos\theta_2, \cos\theta_3, \chi) &= (3\cos^2\theta_2 - 1)\cos\theta_3 \\
&\quad -3\ \sin\theta_2 \cos\theta_2 \sin\theta_3 \cos\chi \\
a_{2^+}(\cos\theta_2, \cos\theta_3, \chi) &= -i\ \sin\theta_2 \cos\theta_2 \sin\theta_3 \sin\chi.
\end{aligned} \tag{5.3.1}$$

The data were fit again using these angular amplitudes for the broad resonance. When testing one of these amplitudes, no mixing with the $D_1(2420)^0$ was allowed, although the $D_1(2420)^0$ *was* allowed to decay through both **d**-wave and **s**-wave intermediate states. The mass and width parameters that describe the broad resonance were allowed to float freely in the fits.

The statistical significance ($\sqrt{2\Delta \ln \mathcal{L}}$) with which the 1^+ assignment is favored over each of these possibilities is listed in Table 5.10. By testing the effects of our systematic uncertainties on the ($\ln \mathcal{L}_{\mathrm{max}}$) of each spin-parity assignment, we may include these effects in a measure of overall significance. This overall significance for the 1^+ assignment is also listed in Table 5.10. These results strongly support the assignment of 1^+ spin-parity to the broad resonance.

Table 5.10: The maximum log-likelihood for each tested assignment of spin-parity. The significance of the 1^+ assignment relative to each alternative is listed. The overall significance was tested as described in the text.

J^P	$\ln \mathcal{L}_{max}$	Statistical significance	Overall significance
$\mathbf{1^+}$	**43916.10**	—	—
0^-	43912.98	2.5σ	2.0σ
1^-	43909.60	3.6σ	3.3σ
2^-	43902.37	5.2σ	4.9σ
2^+	43905.45	4.6σ	3.5σ

5.4 Alternate Parameterizations

It is not clear whether the parameterization of Eq. 1.1.23 is optimal, or even appropriate for the mixing and strong phases among the resonances. One possibility is the 'grand' amplitude

$$\begin{aligned} \mathcal{A} \;=\; & \quad \alpha_{1n}\, A_{1n}\, \left(a_{1d}\cos\beta + a_{1s}e^{i\phi}\sin\beta e^{i\delta_0}\right) \\ & + \; \alpha_{1b}\, A_{1b}\, \left(a_{1s}\cos\beta e^{\delta_0} - a_{1d}e^{i\phi}\sin\beta\right) \\ & + \; \alpha_2\, A_2\, a_2 \\ & + \; \alpha_{non-res}. \end{aligned} \tag{5.4.1}$$

Here, we assume that the only significant phase is between the **s**-wave and **d**-wave channels of the 1^+ resonances: the phases among the mass eigenstates are assumed to be negligible. Using this convention to fit the data yields the results shown in Table 5.11, to be compared with Table 5.2.

Table 5.11: Details of the fit to the data for the signal distribution with the parameterization shown in Eq. 5.4.1. Yields in the resonant sub-channels are corrected for their relative efficiencies. The yields in the $D_1(2420)^0$, $D_2^\star(2460)^0$ and non-resonant channels are only slightly different from the results in Table 5.2. The same is true for the mass and width of the $D_1^0(j=\frac{1}{2})$. The change in yield for the broad resonance is slightly more than 1σ.

Parameter	Central value	Statistical error
$D^\star\pi^-\pi^-$ Yields		
$B^- \to D_1(2420)^0\,\pi^-$	237.1	± 42
$B^- \to D_1^0(j=\frac{1}{2})\,\pi^-$	420.0	± 41
$B^- \to D_2^\star(2460)^0\,\pi^-$	109.5	± 26
$B^- \to D^{\star +}\,\pi^-\,\pi^-$ non-res	160	± 71
$D_1^0(j=\frac{1}{2})$ lineshape		
$M_{D_1^0(j=\frac{1}{2})}$ (GeV)	2.492	± 0.026
$\Gamma_{D_1^0(j=\frac{1}{2})}$ (GeV)	0.326	± 0.086
Mixing and interference		
β	0.764	± 0.101
ϕ	2.66	± 0.12
δ_0	1.98	± 0.16

As in the previous fit, the net interference is negative, so that the sum of the individual product branching ratios would yield 1.103 ± 0.051 times the observed total resonant contribution to the $D^{\star +}\pi^-\pi^-$ final state (697 events). The log-likelihood of this fit is 43919.47. Thus, with *fewer* free parameters, the parameterization of Eq. 5.4.1 is favored by 2.6σ over that of Eq. 1.1.23. Comparison with Table 5.2 shows that the errors on the fit parameters are similar or smaller than in the previous fit.

Neither the yields of the D_J^0 nor the parameters of the broad resonance are severely effected by this change. However, the values of the mixing parameters are altered profoundly. This fit predicts equal parts of **s**-wave and **d**-wave in the decays of the 1^+ states. Meanwhile, the contribution of the mixing phase gains an overall (-) sign.

Subsequently, several suggestions have been made regarding possible improvements to our 'grand' amplitude. These changes affect both the degrees of freedom for the strong phases and the modeling of the mixing and lineshapes for the partial waves of the 1^+ states. It is not clear at this time what parameterization is necessary to accurately capture the underlying physics of these decays.

Thus, we use the central values from our original fit and ascribe a systematic uncertainty on the branching fractions and the lineshape of the $D_1^0(j=\frac{1}{2})$ due to our parameterization. We base these estimates on the results of the fit shown in Table 5.11. In addition, we emphasize that the results obtained for the mixing parameters and strong phases must be treated with great caution. Further investigation will be necessary to finalize those measurements and will be reported upon in the near future.

5.5 Conclusions

The four-dimensional fit to the data provides convincing evidence for $B^- \to D_J^0\,\pi^-$ decays to all three D_J that can decay to $D^{\star +}\pi^-$. Converting these yields to branching ratios gives the measurements,

$$\mathcal{B}\left(B^- \to D^{\star +}\,\pi^-\,\pi^-\mathrm{total}\right) = (29.2 \pm 4.5 \pm 3.8 \pm 3.1)\times 10^{-4} \tag{5.5.1}$$

$$\mathcal{B}\left(B^- \to D^{\star +}\,\pi^-\,\pi^-\mathrm{non\text{-}res}\right) = (9.7 \pm 3.6 \pm 1.5 \pm 1.9)\times 10^{-4} \tag{5.5.2}$$

$$
\begin{aligned}
\mathcal{B}\left(B^- \to D_1(2420)^0\,\pi^-\right)\mathcal{B}\left(D_1(2420)^0 \to D^{\star+}\,\pi^-\right) &= (6.9\,^{+1.8}_{-1.4} \pm 1.1 \pm 0.4)\times 10^{-4} \\
\mathcal{B}\left(B^- \to D_1^0(j=\tfrac{1}{2})\,\pi^-\right)\mathcal{B}\left(D_1^0(j=\tfrac{1}{2}) \to D^{\star+}\,\pi^-\right) &= (10.6 \pm 1.9 \pm 1.7 \pm 2.3)\times 10^{-4} \\
\mathcal{B}\left(B^- \to D_2^\star(2460)^0\,\pi^-\right)\mathcal{B}\left(D_2^\star(2460)^0 \to D^{\star+}\,\pi^-\right) &= (3.1 \pm 0.84 \pm 0.46 \pm 0.28)\times 10^{-4}.
\end{aligned}
\tag{5.5.3}
$$

The final error is due to the uncertainty in the parameterization of the strong phases. The sum of the contributions to the $D^\star\pi^-\pi^-$ final state is greater than the total observed due to the destructive interference between the resonant contributions. Using the D_J branching fractions that are assumed in much of the theoretical literature [23]:

$$\mathcal{B}\left(D_1(2420)^0 \to D^{\star+}\,\pi^-\right) = 2/3 \tag{5.5.4}$$

$$\mathcal{B}\left(D_1^0(j=\tfrac{1}{2}) \to D^{\star+}\,\pi^-\right) = 2/3 \tag{5.5.5}$$

$$\mathcal{B}\left(D_2^\star(2460)^0 \to D^{\star+}\,\pi^-\right) = 0.20, \tag{5.5.6}$$

we obtain the absolute $B^- \to D_J^0\,\pi^-$ branching fractions,

$$\mathcal{B}\left(B^- \to D_1(2420)^0\,\pi^-\right) = \left(1.04\,^{+0.27}_{-0.21} \pm 0.17 \pm 0.07\right) \times 10^{-3} \tag{5.5.7}$$

$$\mathcal{B}\left(B^- \to D_1^0(j=\tfrac{1}{2})\,\pi^-\right) = (1.59 \pm 0.29 \pm 0.26 \pm 0.35) \times 10^{-3} \tag{5.5.8}$$

$$\mathcal{B}\left(B^- \to D_2^\star(2460)^0\,\pi^-\right) = (1.55 \pm 0.42 \pm 0.23 \pm 0.14) \times 10^{-3}. \tag{5.5.9}$$

This measurement provides evidence that establishes the existence of a broad $D^{\star+}\pi^-$ resonance with a significance of 5.7σ. Tests of the J_P assignments, 0^-, 1^-,1^+, 2^-, 2^+ favor the assignment of $J^P = 1^+$ to this state by 2σ relative to the next best alternative, which is 0^-. The mass and width of this state are determined to be

$$M_{D_1^0(j=\frac{1}{2})} = 2.461^{+0.041}_{-0.034} \pm 0.010 \pm 0.032\text{GeV} \tag{5.5.10}$$

$$\Gamma_{D_1^0(j=\frac{1}{2})} = 290^{+101}_{-79} \pm 26 \pm 36\text{MeV}, \tag{5.5.11}$$

and the mixing angle and mixing phase between the two 1^+ states are found to be

$$\beta = 0.368 \pm 0.100 \pm 0.040 \tag{5.5.12}$$

$$\phi = -0.244 \pm 0.42 \pm 0.15. \tag{5.5.13}$$

We measure the strong phases that cause interference among resonant decays to be

$$\delta_1 = 2.62 \pm 0.30 \pm 0.16 \tag{5.5.14}$$

$$\delta_2 = 1.04 \pm 0.34 \pm 0.09, \tag{5.5.15}$$

where these phases are measured relative to that of the $D_1(2420)^0$. The imaginary parts of these phase shifts have a significance of 1.6σ and 1.7σ respectively. Our fit is not sensitive enough to determine the strong phase between the resonant and non-resonant contributions to the $D^{\star+}\pi^-\pi^-$ final state. Due to the uncertainties in the modeling of the phases and mixing, we urge great caution in the interpretation of the results for β, ϕ, δ_1 and δ_2.

CLEO, ARGUS and E687 have published the most recent results for the decays of the D_J^0[11, 10, 9, 12, 24]. Partial wave analysis of the $D_2^\star(2460)^0$ led to the spin-parity assignment of 2^+. The lack of $D_1(2420)^0$ decays to the $D^+\pi^-$ final state, and decay angular distributions consistent with the d-wave decay of the 1^+ leads to the assumption that the $D_1(2420)^0$ has spin-parity of 1^+ and that the mixing angle, β, is small.

Both CLEO and ARGUS have also obtained results for the decay $B^- \rightarrow D_1(2420)^0\,\pi^-$[25, 21]. Earlier results of the present analysis based on a two-dimensional fit reported a measurement for the decay to the $D_2^\star(2460)^0$ as well. A summary of these results is presented in Table 5.12. One puzzle created by our measurement is the larger

Table 5.12: A summary of existing measurements of the decays $B^- \rightarrow D_J^0\,\pi^-$, including the earlier results of this analysis using a two-dimensional fit.

	$\mathcal{B}(B^- \rightarrow D_J^0\,\pi^-) \times \mathcal{B}(D_J^0 \rightarrow D^{\star+}\,\pi^-)$	
	$D_1(2420)^0$	$D_2^\star(2460)^0$
ARGUS(1994)[21]	$(16.8 \pm 4.7 \pm 4.0) \times 10^{-4}$	$< 6.9 \times 10^{-4}$
CLEO(1994)[25]	$(7.4 \pm 3.3 \pm 1.3) \times 10^{-4}$	$< 5.6 \times 10^{-4}$
CLEO(1996)[26]	$(7.8 \pm 1.6 \pm 1.0) \times 10^{-4}$	$(4.2 \pm 1.6 \pm 0.6) \times 10^{-4}$
This measurement	$(6.9\,^{+1.8}_{-1.4} \pm 1.1) \times 10^{-4}$	$(3.1 \pm 0.84 \pm 0.45) \times 10^{-4}$

than expected branching ratio for $B^- \rightarrow D_2^\star(2460)^0\,\pi^-$[27]. A number of theoretical and experimental factors may be responsible for this discrepancy. Among them are the CLEO measurement of the ratio

$$\frac{\Gamma\left(B^- \rightarrow D_2^\star(2460)^0\ell^-\overline{\nu}_\ell\right)}{\Gamma\left(B^- \rightarrow D_1(2420)^0\ell^-\overline{\nu}_\ell\right)}, \tag{5.5.16}$$

and of the D_J branching fractions that are also dominated by CLEO measurements. More precise measurements of the D_J branching fractions will be important in unraveling this mystery.

There are theoretical predictions for the mass and width of the $D_1^0(j=\frac{1}{2})$ [8] as well as the mixing of the 1^+ states [8, 28]. The most specific predictions for the spectroscopy of the D_J are due to Godfrey and Kokoski and are shown in Table 5.13 along with our results. The measured value of the 1^+ mixing angle is somewhat larger than expected. The

Table 5.13: The predictions of Godfrey and Kokoski compared to the current results.

	$M_{D_1^0(j=\frac{1}{2})} - M_{D_1(2420)^0}$	$\Gamma_{D_1^0(j=\frac{1}{2})}$	β
G & K	$+10$ MeV	250–1000 MeV	0.17
Measured	$+39.2\,^{+42}_{-35}$ MeV	$290\,^{+104}_{-83}$ MeV	0.368 ± 0.11

theoretical estimate of β depends upon many quantities including the width of the $D_1^0(j=\frac{1}{2})$. Based upon the larger than expected width of the $D_1(2420)^0$, others have suggested that the mixing might be somewhere in the neighborhood of 0.20 [28].

There are also predictions for the $B^- \to D_J^0\,\pi^-$ branching ratios that use the same assumptions as we have for the $D_J^0 \to D^{\star+}\,\pi^-$ branching ratios. These predictions are [29]

$$\mathcal{B}\left(B^- \to D_1(2420)^0\,\pi^-\right) = 4 \times 10^{-4} \tag{5.5.17}$$

$$\mathcal{B}\left(B^- \to D_1^0(j=\tfrac{1}{2})\,\pi^-\right) = 6 \times 10^{-4} \tag{5.5.18}$$

$$\mathcal{B}\left(B^- \to D_2^\star(2460)^0\,\pi^-\right) = 6 \times 10^{-4}. \tag{5.5.19}$$

Our results appear to disagree with these absolute branching fractions, however, the agreement with the pattern of relative branching fractions is striking.

Part II

Construction of a Silicon Vertex Detector for CLEO II.V

Chapter 6

SVX Design

6.1 Introduction

Distinguishing the decay vertexes in complicated reactions has often been the key to new discoveries in particle physics. These include the first observation of the pion [30], early evidence for the neutrino [31], and the discovery of the kaon [32] which led ultimately to the quark model.

The task performed by the cloud chamber in the discovery of the Kaon was simple by modern standards. The observed K^0_{S} has a lifetime of about 10^{-10} sec. and travels a few centimeters before decaying into a pion and a proton. The B and D mesons we now wish to study in detail are more massive and have greater phase space in which to decay. As a result, they have lifetimes on the order of 10^{-12} sec. and decay lengths measured in tens or hundreds of microns. Fortunately, solid state devices capable of detecting the passage of charged particles have made such precise measurements possible. These detectors have been successfully used in a number of recent fixed target and collider experiments.

A silicon vertex detector using this technology has been constructed for the CLEO II experiment at the CESR e^+e^- collider. This device, the SVX, implements the latest advances in silicon microstrip detectors and mechanical design to identify detached decay vertexes from charm mesons, tau leptons and other long-lived particles. This detector and its descendants will lead the way to a more complete understanding of heavy quarks, the strong interaction that binds them together and the weak interaction responsible for their decays.

6.2 Objectives

The primary objective of the SVX design must be reconstruction of decay vertexes. For the study of CP violation in the B system, we would like to measure the flight paths of the B mesons before decaying to specific final states. At a symmetric $\Upsilon(4S)$ machine such as CESR, the B lifetime of 1.6 ps corresponds to a flight path of only ≈ 30 μm. Meanwhile, RMS of the beam spot in the CLEO interaction region is roughly 350 μm (horizontal) $\times$ 7 μm (vertical) $\times$ 1.8 cm (Z) [33] and typical single hit resolutions for the current generation of silicon detectors is 10-20 μm. Clearly, reconstructing separated vertexes from B decays at CLEO is beyond the reach of even the most optimistic scenario. Future 'asymmetric B factories' will address this problem by using colliding beams of different energies to produce $\Upsilon(4S)$ that are boosted in the lab frame.

The D mesons have shorter lifetimes still. However, the D momentum B decay results in flight paths on the order of 300 μm for D^+ and 120 μm for D^0. D mesons from continumm production have even longer flight paths. By using the SVX to identify these decay vertexes, the D mesons may be studied in greater detail. Since the majority of B mesons decay to a final state including a D meson, this also enables most of the B decay width to be studied with higher precision.

The situation for τ lepton decays at CLEO is comparatively straightforward. Tau leptons at CLEO can be produced in the reaction $e^+e^- \rightarrow \tau^+\tau^-$ giving $\beta\gamma c\tau \approx 270$ μm. Large, clean samples of τ decays should become possible.

A secondary objective of the SVX design is improved tracking of charged particles. The CLEO II detector described in Part I, provides good resolution for measurements in the r-ϕ plane. However, only tracks that reach the outermost drift chamber are well measured in z. This results in poor determination of θ and the z impact parameter, especially for low-momentum tracks. The replacement of the PTL with an SVX capable of making precise z measurements will significantly improve tracking in the CLEO detector.

These objectives define a starting point for the SVX design. To provide vertexing in three dimensions and improve low-momentum tracking, we need at least two precision measurements in both r-ϕ and z as close to the interaction region as possible. A third layer adds redundancy and the capability for stand-alone tracking.

6.3 Constraints

Several constraints shape the design and construction of the SVX. Primary among them are

- *time* — The time allocated for construction of the SVX is minimal compared to other similar detectors which have been built. The SVX design and construction process must stress simplicity to achieve high yields and rapid assembly.

- *space* — Since the VD will not be replaced, the SVX must fit entirely within the space currently occupied by the PTL. This creates a number of heat removal and cabling problems for a three-layer device.

- *material* — There are tight constraints on the amount of material that may be used in the SVX. The PTL presents about 0.5% of a radiation length to tracks before they reach the VD or DR. Significant additional material will increase multiple scattering for low-momentum tracks. At 0.3% per 300 μm layer, a three-layer device doubles the material without including a support structure. Considerable importance must be placed on the lightest possible supports, compounding the problems of heat removal and dimensional precision.

- *alignment* — The resolution of silicon detectors far exceeds that of any other element of the CLEO detector. This renders the VD and DR incapable of providing the information necessary to understand the alignment of the individual silicon detectors within the SVX. Experience at ALEPH has shown the importance of overlap between adjacent detectors in performing this alignment. This overlap magnifies the difficulty of designing the SVX within the available space.

- *high luminosity* — The CESR Phase II design luminosity is $6 \times 10^{32}\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The SVX must be capable of fast readout while distinguishing signal from considerable backgrounds. All components must be sufficiently radiation-hard to withstand the significant dosage expected over the lifetime of CLEO II.V.

Table 6.1: General specifications of the SVX detector.

	Layer 1	Layer 2	Layer 3	Total
Radius	23 mm	32 mm	48 mm	—
Active area	201.6 cm^2	351.0 cm^2	806.6 cm^2	1359.2 cm^2
Angular coverage	92%	92%	92%	92%
# silicon detectors	16	16	64	96
# sense strips	12096	16128	42336	70560
# readout channels	6048	8064	12096	26208
Heat load	24W	32W	48 W	104 W

6.4 Overview

The SVX, pictured in Figure 6.1 and summarized in Table 6.1, consists of 96 silicon microstrip detectors assembled in eight octants providing 92% coverage of the solid angle. The octants are placed inside a carbon fiber composite tube with a 5% overlap between the azimuthal coverage of adjacent octants in all three layers.

The twelve detectors in each octant are arranged to form three layers at the approximate radii of 23mm, 32mm and 48mm, which we call layers one, two and three respectively. Layers one and two of each octant consist of two detectors each, one on either side of the ideal interaction point in z (East/West along the beamline). Layer three consists of eight detectors in each octant. The coverage in azimuthal angle, ϕ, is split into two 'sub-layers,' one at 47mm and one at 48mm. Each of these sub-layers has four detectors, two on either side of the IP which are glued end to end and connected in a daisy-chained fashion, giving the effect of one wafer as in layers one and two. Each of these detectors or detector assemblies is glued to a pair of beryllium oxide (BeO) hybrid circuit boards. These hybrids contain the pre-amplifiers and readout electronics for the silicon detectors and are connected to the detectors by wirebonds.

These detector-hybrid assemblies are glued to a support structure consisting of Kevlar®[a] fiber composite beams glued to BeO ceramic endpieces. The surfaces of the silicon detectors are glued directly to the Kevlar® beams while the hybrids are sandwiched between the BeO endpieces, forming a passive cooling path for the electronics on the hybrids. Layers one and two are glued to either side of a beam assembly consisting of a single Kevlar® beam

[a]Kevlar® Kapton® and Teflon® are registered trademarks of the DuPont Corporation

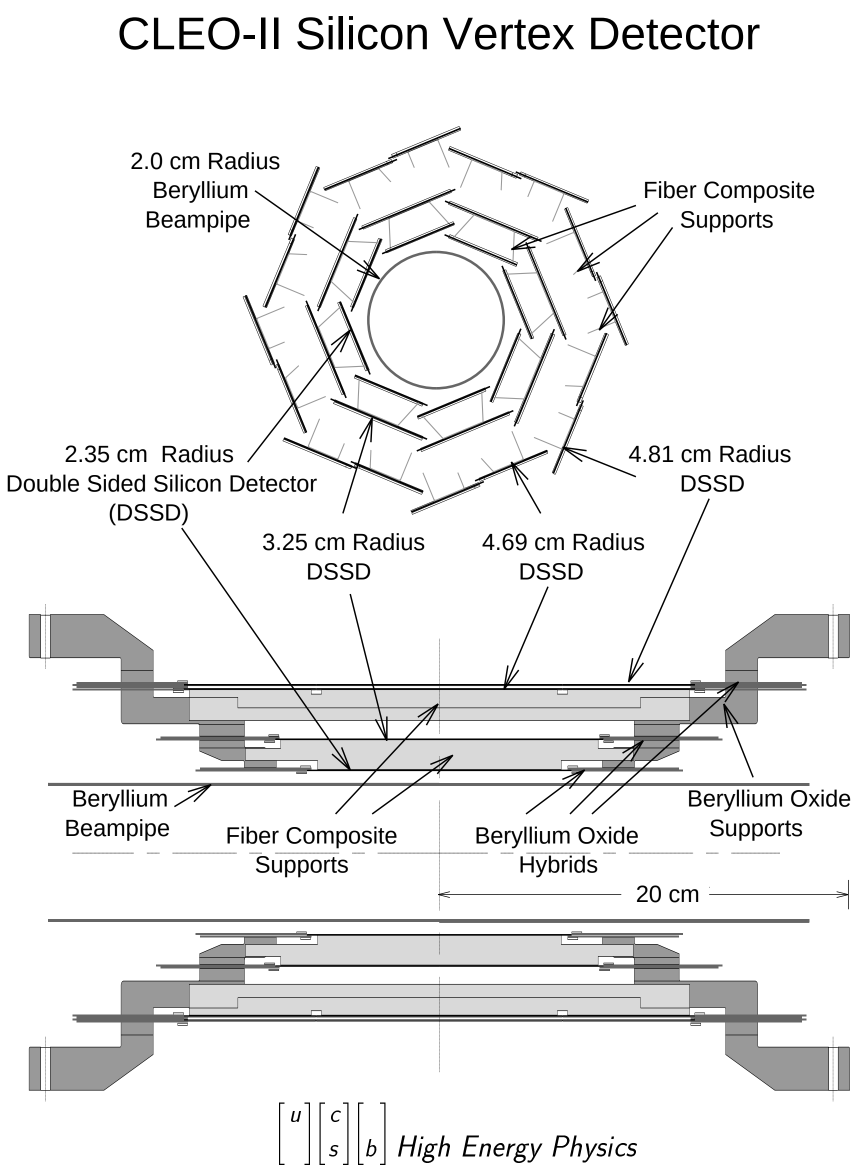


Figure 6.1: An overview of the SVX design. The view of the r-phi plane (top) shows the overlapping octants around the Be beampipe, supported by Kevlar® fiber U-channels. The view of the r-z plane (bottom) shows the BeO hybrids with CAMEXes sandwiched between the BeO endpieces. The vertical dimension (r) is exaggerated somewhat in the r-z view for clarity.

and two endpieces. The beam assembly for layer three consists of two Kevlar® beams, one for each sub-layer, and a common BeO endpiece on either end that unifies the two sub-layers into a single unit. The inner beam assembly of layers one and two is then glued to the outer beam assembly of layer three at the endpieces, forming a complete octant.

Completed octants are placed in sets of four inside two semi-cylindrical carbon fiber composite housings. Copper tubing for cooling fluid fits inside a channel in the housings, drawing the heat load of each octant through the endpieces for the outer beam assembly. There are shorter, thicker carbon fiber extensions that house electronics to drive the signals through the long cables to the data acquisition electronics. Stainless steel support rings laid into the outer ends of the extensions are the mounting point for the SVX as a whole.

6.5 Silicon Detectors

The 96 silicon detectors used in the SVX are of five varieties, whose properties are summarized in Table 6.2. There are two designs in layer one, one for each side of the interaction point (IP) in z, that are mirror images of one another. The same is true in layer two. In layer three, a fifth symmetric design is glued edge-to-edge with layer one detectors to produce the longer detector assemblies required. This arrangement is dictated by re-use of the design for layer one and the constraints of fabricating the detectors from four-inch diameter silicon wafers.

To enable the use of pre-existing 64-channel readout electronics, the number of channels on either side of each detector is a multiple of 63. A spare was left so that defective amplifier channels could be skipped when necessary. Wirebonds are used to connect the detectors to readout electronics. The pitch of the wirebonding pads on the detectors must match the 97 μm pitch of pads on the readout chips, so a 'fan-in' of the readout strips occurs at the ends of the detectors. Layer one detectors have wirebond pads at both ends to facilitate their use in layer three, where they are wirebonded to the extensions at one end and the hybrid electronics at the other.

The detectors operate on relatively straightforward principles. The bulk silicon is slightly n-doped, giving it an excess of negative charge carriers. Strips of p-type material are implanted on one side, the 'p-side', with an excess of positive charge carriers. On the other side, the 'n-side', strips of n-type material are implanted perpendicular to the p-side

Table 6.2: A summary of the properties of the silicon detectors. A layer three detector assembly consists of a layer one detector and a layer three extension daisy-chained together.

	Layer 1	Layer 2	Layer 3 extensions
# designs	2, East & West	2, East & West	1
Full depletion voltage	60 V	60 V	60 V
Minimum ionizing signal	22,000 e^-	22,000 e^-	22,000 e^-
Length	65.580	81.270	61.045
Width	22.568	30.128	22.568
Active length	59.535 mm	76.356 mm	59.535 mm
Active width	21.168 mm	28.614 mm	21.168 mm
p-side (r-ϕ) properties			
# implants	753	1005	753
Implant width	10 μm	10 μm	10 μm
Implant pitch	28.0 μm	28.5 μm	28.0 μm
# instrumented strips	189	252	189
# readout channels	189	252	189
Readout pitch	112 μm	114 μm	112 μm
Theoretical resolution	12 μm	12 μm	12 μm
n-side (z) properties			
# implants	567	756	567
Implant width	20 μm	20 μm	20 μm
Implant pitch	105 μm	101 μm	105 μm
p-barrier width	69 μm	65 μm	69 μm
# instrumented strips	753	1005	753
# readout channels	189	252	189
Readout capacitance	22 pf	28 pf	22 pf
Theoretical resolution	30 μm	30 μm	30 μm

implants. A p-n junction is formed at the interface between the p-strips and the bulk n-type silicon. The mobile positive and negative charge carriers diffuse through the material and combine near this interface, leaving a layer devoid of free charge called the depletion layer. The lattice charges that remain create an electric field pointing toward the p-side in this region.

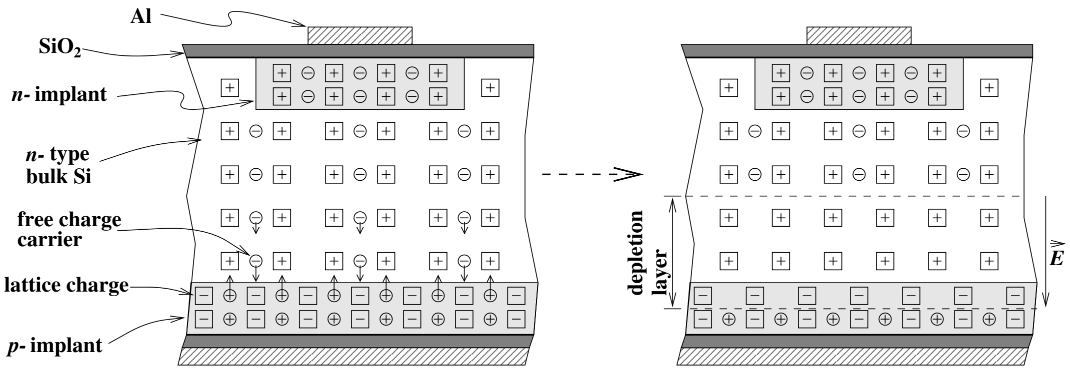


Figure 6.2: A diagram showing formation of the depletion layer in an unbiased silicon detector.

When an ionizing particle passes through the silicon, it excites electron-hole pairs into the conduction band. A typical minimum ionizing particle will produce approximately 22,000 such pairs of free charge carriers in passing through the 300 μm thick layer of silicon.

The electric field in the depletion layer causes the electron-hole pairs produced there to drift toward the p and n-sides respectively. Those produced elsewhere in the silicon diffuse and recombine. Thus, only a small fraction of the charge can be collected on the p-side strips and none collected on the n-side strips. By placing a voltage across the detector so that the p-n junction is reverse-biased, the conduction electrons in the rest of the bulk n-type silicon are removed, enlarging the depletion layer. When the bulk silicon is devoid of free charge carriers, the device is said to be fully depleted. The larger depletion region allows nearly all of the electrons and holes to drift the full thickness of the silicon and be collected on the orthogonal p-side and n-side strips. The increased electric field increases the drift velocity, allowing faster charge collection. The voltage required to fully deplete the silicon is approximately 60V.

On the n-side there is an additional complication. A thin layer of mobile charge carriers exists at the n-side surface of the bulk silicon even when the device is fully depleted.

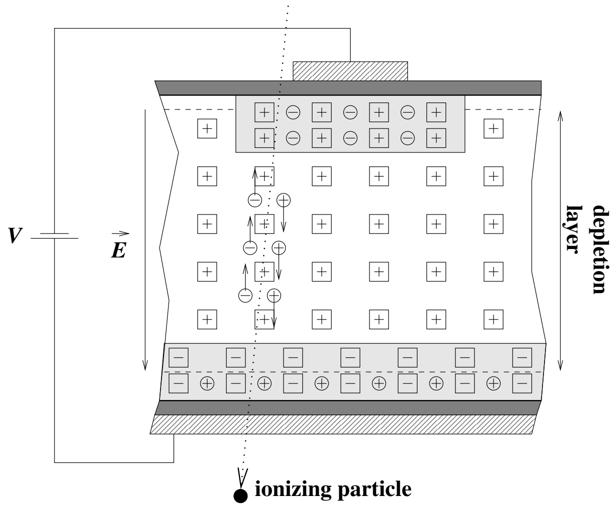


Figure 6.3: A diagram of a silicon detector fully depleted by application of bias voltage. Full depletion allows all of the electron-hole pairs produced by the passage of particles to be collected.

These electrons cause the n-strips to be shorted together, making it impossible to collect charge on a single strip. A wide p-type strip is implanted between each pair of n implants to drive away these electrons and isolate the n-strips.

The silicon is read out via aluminum sense strips placed above the implants. Only every fourth strip is read out on the p-side. The remaining sense strips contribute to the resolution of the device by capacitive charge sharing. Every strip on the n-side is read out. Instead of being directly connected to the implants ('DC coupled'), the strips are deposited on top of a 1μm layer of silicon dioxide (SiO_2). This creates a capacitor between each charge-collecting implant and its aluminum sense strip ('AC coupled'), with the SiO_2 acting as the dielectric. As a result, the readout electronics can be operated from a common ground or other convenient reference, rather than needing to be grounded at the bias voltage. Defects in the SiO_2 layer, called pinholes, can cause breakdown of the capacitor, rendering the channel inoperable.

Both sides of the detector are read out at the end of the device to save space and material in the tracking volume. While sense strips on the r-ϕ side naturally terminate at the end of the detector, the z-side strips run width-wise across the detectors. A second layer

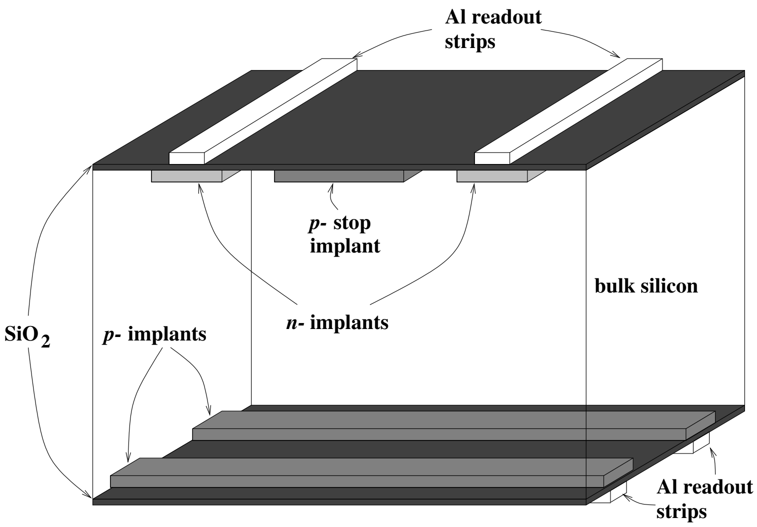


Figure 6.4: A diagram showing the orthogonal strips of a silicon detector and the p-stops necessary to isolate the n-implants.

of readout traces orthogonal to the sense strips carries the signals to the ends of the detectors on the z-side. These traces are isolated from the sense strips by a 2μm layer of polyamide insulator. Due to the length-to-width ratio of the detectors, each z-side readout trace is connected to three sense strips in layers one and two; six in layer three where two detectors are bonded end-to-end. This leads to a three-fold ambiguity in z position in layers one and two, and a six-fold ambiguity in layer three. Another effect of the double-metal design is increased capacitance of the readout traces on the z side.

The bias voltage is applied to the implants through resistors from aluminum bias rings around the sense region. Ten MΩ polysilicon resistors are used on the n-side. On the p-side the resistance is from the bulk silicon itself. These resistances range from 10-100 MΩ over the voltage range in which the detectors are operated and are referred to as 'punch through' resistors.

A one μm layer of SiO_2 passivation coats both sides of the detectors. This layer protects the detectors from physical damage that can result in pinholes and electrically isolates the sense strips from the environment. To improve this protection, later deliveries of detectors employed an additional layer of silicon nitride (Si_3N_4), an extremely hard ceramic

Table 6.3: Specifications for the CAMEX and JAMEX chips used in the SVX

	CAMEX	JAMEX
Die size (mils)	$250 \times 200 \times 19$	$250 \times 200 \times 19$
# input channels	64	64
Input pitch	97 μm	97 μm
Gain	14 mV/fC	N/A
# used	320	96

material. These detectors proved more resistant to damage-induced pinholes.

6.6 Readout Electronics

The task of amplifying the charge on each readout strip is performed by the CAMEX pre-amplifier chip. The CAMEX is a 64-input, VLSI, charge-to-voltage amplifier with multiplexed output. The device was originally developed by MPI-Munich for the ALEPH vertex detector and was chosen for its performance and the simplicity of using a pre-existing device [34]. Table 6.3 summarizes some essential features of the CAMEX.

The CAMEX operates by sampling the charge on the readout strip four times before and after the passage of an ionizing particle and taking the difference. A schematic of a single CAMEX channel is shown in Figure 6.5. The switches R1 and R2, and S1-S4 are controlled by signals from the data acquisition system and perform the above task as follows. When R1 is open, charge at the input is converted into a voltage at the output of CSA1. Before a beam crossing, R2 is kept closed while S1-S4 are sequentially closed and reopened. The amplifier CSA2 brings charge to its input to zero the voltage there. Operating continuously in this way, the sampling capacitors, C1-C4, always contain four recent samples of the input.

After a beam crossing, R2 is opened and S1-S4 are closed and reopened again. This places charge on Cf2 equal to the new charge needed at the sampling capacitors since the last set of samples before beam crossing. The output voltage is then proportional to the difference between the input before and after the beam crossing. This scheme gives a four-fold sampling of the charge collected by the readout strip, enhancing signal to noise since the signal adds coherently but noise does not. The delay required for the CLEO trigger precludes collection of the first sample after an event. The cost in signal/noise of having three samples instead

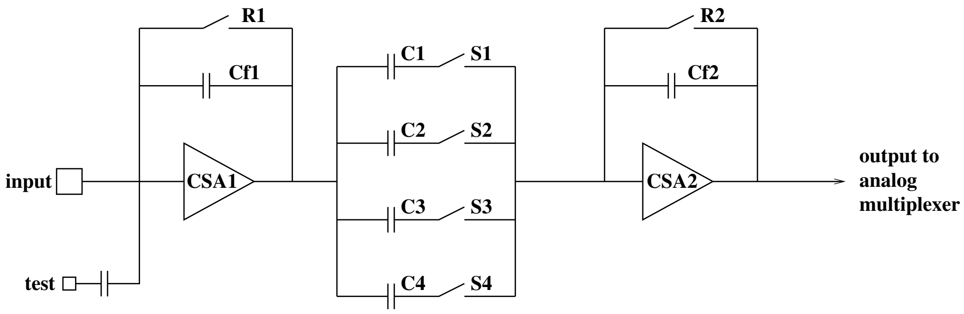


Figure 6.5: A schematic showing a single CAMEX amplifier channel.

of four is a factor of $\sqrt{3/4}$.

The equivalent noise charge (ENC) when reading out a detector with the CAMEX using all four samples is the incoherent sum of several sources [35]:

$$(NOISE)^2 = \left[(340 + 31C_D)^2 + \left(\frac{3.6kT}{R} + 1.8qI\right)\tau + \alpha^2\right] e^- ENC. \qquad (6.6.1)$$

The first term is the intrinsic noise of the CAMEX attached to a detector channel with capacitance C_D (in pF). The second term is thermal noise due to the detector resistance R and shot noise from leakage current, I. The contribution from these sources depends upon the sampling time, τ, and is approximately 200 e^- for our standard sampling time. The last term is the additional noise from external sources, typically 400–600 e^-.

Noise on the n-side is twice that on the p-side, primarily due to the higher capacitance of the double-metal readout. Using $C_D = 22$ pF from Table 6.2, we expect roughly 1400 e^- ENC on the n-side of layer one. This noise is doubled again in layer three where readout strips from two detectors are connected to each CAMEX input. With a signal of 22,000 e^- this gives $S/N \approx 8$ before subtracting 15% from loss of the first sample.

To address this problem, CAMEX chips rendered in a lower-noise CMOS process are used on the n-side of layer three. This device, otherwise known as the JAMEX, is identical to the CAMEX in operation, but with one half of the intrinsic noise and noise slope of the CAMEX. Specifications for the JAMEX are also shown in Table 6.3.

Neither of these chips is radiation-hard. Because of the shielding effect of the SVX itself, only the CAMEX chips on layer one receive significant dosage, primarily from synchrotron radiation. Detailed studies [36] have shown that CAMEXes under power begin

Table 6.4: Properties of the BeO hybrids used in the SVX.

	Layer 1&3	Layer 2
# CAMEX/JAMEX	3	4
Design thickness	20 mils	20 mils
# used	48	16
Metallization & wirebonds	Gold	Gold

to experience significant channel losses at 20-30 Krad due to loss of gain. When unpowered, as during CESR injection, the CAMEX can withstand between 60 and 120 Krad before noise increases significantly ($> 2\times$), or gain fails for all channels. These levels are well within the designed lifetime of the SVX, which will be replaced with the installation of CLEO III.

6.7 BeO Hybrids

The readout electronics are mounted on hybrid circuit boards glued directly to the silicon detectors. These hybrids must conduct the heat from their CAMEX chips (0.25 W/CAMEX) to the thermally conductive support structure. Good contact between the hybrids and the supports is accomplished by sandwiching the hybrids between the three tiers of endpieces. This makes them part of the passive cooling path and dimensionally critical components of the detector.

The double-sided silicon detectors require hybrids with CAMEXes on both sides. Difficulties fabricating such boards led to a simpler design of two single-sided hybrids glue-bonded back-to-back. This places great importance on the thermal conductivity and thickness of the glue layer that joins the hybrids.

Basic properties of the hybrids used in the SVX are summarized in Table 6.4. The hybrids are fabricated on substrates of beryllium oxide ceramic. BeO was chosen for its unusual properties as an electrically insulating thermal conductor with low Z. The thermal conductivity of BeO is (260W/(m-°K)), comparing favorably to metals such as aluminum (234W/(m-°K)).

Signals travel between the hybrids and the I/O buffer boards in the housing extensions via Kapton® flex circuits with copper traces. Although there are only two hybrid

Table 6.5: Specifications for the U-channels. The finished weight is the weight after machining.

	Inner beam	Outer beam #1	Outer beam #2
Detectors supported	layers 1&2	layer 3, inner sub-layer	layer 3, outer sub-layer
Design thickness	12 mils	12 mils	12 mils
Finished thickness	14–17 mils	13–16 mils	10–15 mils
Finished weight	1.8 g	2.2 g	2.0 g

designs (three including those with JAMEX chips), there are eight different Kapton® cables that distinguish the hybrids once they are attached.

6.8 U-channels

The silicon detectors are glued to Kevlar® composite supports of U-shaped cross section. These supports must use the smallest amount of material possible while providing structural integrity and dimensional stability to the silicon. Three types of U-channels are used in the detector. Their properties are summarized in Table 6.5.

Kevlar® is a non-conductive aramid fiber with somewhat lower modulus than carbon fiber. Because the design calls for the detector to be glued directly to the support beams, carbon fiber could not be used. U-channels fabricated using carbon fiber composite have a resistance of $\approx 10\Omega$ from end-to-end and act as antennas. Signals picked up by these supports generate high noise on adjacent readout traces. Careful grounding of the carbon fiber failed to reduce this noise to acceptable levels.

A drawback of molded composites is difficulty in producing parts to a high degree of dimensional accuracy. Finished parts were 10-20% thicker than designed. Windows in non-critical areas of the beams were added to offset this excess. In addition to windows, notches are machined in the layer three beams for clearance of the wirebonds connecting the two detectors in each assembly. Because of these windows and the lower modulus of Kevlar®, the silicon detectors are an important structural member of the finished assemblies. In particular, layers one and two are a sandwich with silicon skins.

Table 6.6: Specifications of the BeO endpieces.

	Inner endpiece	Middle endpiece	Outer endpiece
Finished weight	6.5 g	39.5 g	N/A
Heat load	1.5 W	3.5 W	6.5 W
Minimum cross-section	0.41 cm^2	0.38 cm^2	0.43 cm^2

6.9 BeO Supports

The support structure for an octant is completed by BeO endpieces glued to the ends of the Kevlar® U-channels. Just as the detectors glue to the faces of the U-channels, their hybrids are sandwiched between the three tiers of endpieces. Together with the hybrids, they form the primary mechanical support and the passive heat path to active cooling at the ends of the carbon fiber housings. This arrangement saves the space and material of active cooling inside the tracking volume, but places great emphasis on the thermal conductivity of the glue jonts between the hybrids and these supports. The hybrid electronics at each end of an octant produce approximately six Watts, all of which must pass through these supports. As with the hybrids, BeO ceramic was chosen for its excellent thermal conductivity, low mass and low Z. There are six types of endpieces in each octant, three on each side in z that are mirror images. Some basic properties of the BeO supports are summarized in Table 6.6.

The innermost supports are glued to the inner Kevlar® U-channels. These beam assemblies are then sandwiched between the layer one and layer two detector-hybrid assemblies. The second pair attach to the Kevlar® beams supporting layer three and sandwich between the hybrids for layer two and layer three. The third pair is glued to the outer sides of the layer three hybrids and attaches to the carbon fiber housing. Four screws and two stainless positioning pins fix these supports to the carbon fiber housing. Precision brass inserts glued into the outermost endpieces provide the points of attachment.

6.10 Carbon Fiber Housings

The octants are supported inside of two half-cylinders of carbon fiber composite. This allows the finished SVX to be assembled around the new beryllium beampipe before installation of both as a unit. The central section of this cylinder is only twenty mils thick, adding little material to the tracking volume. The housings are much thicker at the ends, to provide for attachment of the two halves, solid mounting for the octants and mounting of the central housings to the much heavier housing extensions. A copper tube fits inside a groove at each end of the housing to deliver cooling fluid to the outer BeO supports at the ends of the octants.

The semi-cylindrical housing extensions support the receiver boards that buffer I/O to the hybrids and have stainless steel rings laid into the ends to mount the SVX the the rest of the CLEO detector. These parts must be very strong, as they support the weight of the SVX, the beampipe, and two 40Kg tungsten shields during insertion of the detector into CLEO.

Chapter 7

SVX Construction

7.1 Overview

The SVX was assembled at UCSB between August 1993 and September 1995. Among the major challenges were

- assembly to the required precision,
- production of 48,000 wirebonds,
- achieving high thermal conductivity of the support structure,
- maintaining high yields.

7.1.1 Precision Assembly

Dimensional precision was achieved using machined aluminum fixtures attached to pairs of 'platens' for positioning parts during glue-bonding. A set of platens is depicted in Figure 7.1. Each platen consisted of a flat plate of precision-machined aluminum with a grid of reamed positioning holes and tapped bolt holes. The fixtures holding the two parts to be mated were positioned using stainless steel pins and bolted onto the platens. The bottom platen had cylindrical steel posts at each corner, perpendicular to the surface. The top platen had precision bearings that slid on these rails. The vertical spacing between them was set using four aluminum standoffs bolted to the bottom platen. By clamping the parts to the

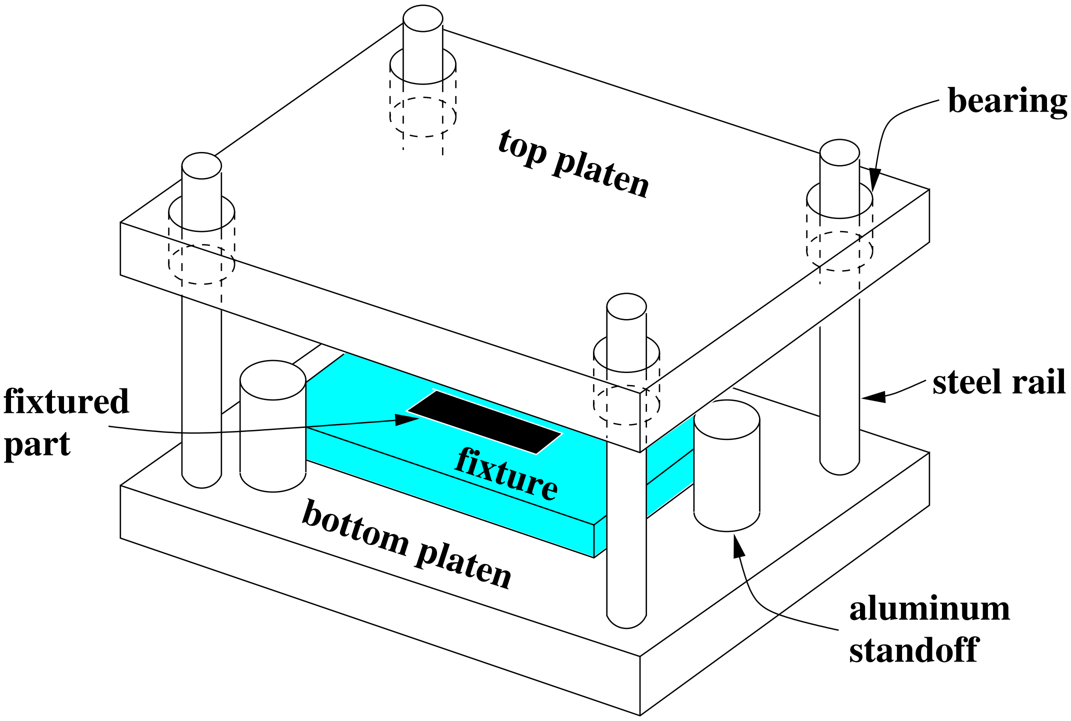


Figure 7.1: A drawing of the platens used to achieve high precision for SVX assembly.

appropriate fixtures on each platen and lowering one platen onto another, two parts could be placed with a relative precision of $\approx 10\mu$m.

The fixtures were also machined on a computer controlled mill from billet aluminum. Stainless steel pins were used to position the parts on the fixtures. A combination of mechanical clamps and vacuum chucks were used to attach the parts to the fixtures. Only vacuum chuck was used for fixturing silicon detectors.

A precise survey of the final positions of the silicon detectors was even more critical than precision in the assembly process. An optical microscope with three-dimensional position readout was used to survey the detector. The precision of this microscope with the highest-power objective was approximately 0.5μm in x and y and 2μm in z. The microscope had a serial connection to a VaxStation 3100, allowing automated readout and analysis of predetermined survey fiducials. The positions of the sense strips on all detectors were catalogued as the SVX was assembled. With single-measurement precisions of $< 1\mu$m in x and y and $< 3\mu$m in z for a skilled operator, the survey of each octant was accurate to a few microns. This information was used to generate constants as a starting point for the alignment of the SVX with charged tracks after installation.

7.1.2 Wirebonding

The task of wirebonding silicon detectors is often contracted to electronics manufacturers with large wirebonding facilities. However, a silicon detector is very different from the components such facilities usually handle. A single misplaced wirebond has the potential to destroy an entire piece of silicon. For small, inexpensive computer chips with tens of wirebonds each, the cost of destroying parts with a single bad wirebond is small. The parts are simply discarded. A silicon detector is worth at least a factor of ten more and requires hundreds of wirebonds. Even one fatal wirebond in $1,000$ is unacceptable. To maintain the highest possible yields, a facility was created at UCSB for production of the wirebonds to the detectors.

All wirebonding was performed on a Kulicke & Soffa 4123 manual wirebonder[a]. This wirebonder delivers a short pulse of ultrasonic energy while the bonding wedge presses the wire into the bond pad with some force. The wirebonder was situated in a class-10 clean area along with the surveying microscope. Training on the use of the wirebonder was

[a]Kulicke & Soffa Corp., Willow Grove, Pa.

Table 7.1: Wires tested for the wirebonding of the silicon detectors.

Wire	Composition	Tensile strength (g)	Elongation (%)
SPM	Al-1%Si	14-16	1-3
CFW	Al-1%Si	14-16	1-3
Hydrostatics	Al-1%Si	15-17	1-4

provided as a personal and professional courtesy by Todd Calvert of the Ferro Corporation[b].

To ensure reliable wirebonds, a pull-tester was purchased and all aspects of the procedure were evaluated. The main variables were wire, bonding wedges, fixtures and wirebonder settings (power, time, force).

The metallization on both the CAMEX and the silicon detectors is aluminum, prompting use of aluminum wire for wirebonding. For the 3-4 mil wide wirebonding pads at both ends, 1 mil wire is the largest recommended gauge. Almost all fine aluminum wire is alloyed with a small percentage of either silicon or magnesium to produce a harder wire. Pure aluminum cannot be easily stretched into such fine wire. A disadvantage of these alloys is that the impurities tend to be unevenly distributed along the length of the wire due to the low solubility of these elements in aluminum. At a site where the concentration of impurities is high, there is a weak spot in the wire. The weakness at these points increases as the wire ages.

Three types of aluminum wire were tested, whose properties are summarized in Table 7.1. The wire made by Hydrostatics[c] was recommended by Promex[d], the contractor for wirebonding of the original SVX for CDF[37]. This wire is produced by high-temperature, high-pressure extrusion. It is claimed that this results in a more even distribution of silicon. While the wirebonds produced with this wire were of average strength, the surface irregularity caused by the extrusion process was problematic. The wire frequently became caught or otherwise jammed along the wire feed path of the bonder.

The other two wires were essentially indistinguishable. The wire from California Fine Wire[e] over that from SPM[f] due to its availability in relatively small quantities, ensuring

[b] Ferro Corp., Electronic Materials Division, Santa Barbara, Ca.
[c] Hydrostatics Corp., Bethlehem, Pa.
[d] Promex Inc., Santa Clara, Ca.
[e] California Fine Wire Corp.,
[f] SPM

that fresh wire could always be used for wirebonding. The wire was stored in a cabinet with a dry-air supply to prevent oxidation.

Only two wirebonding wedges were tested. The first was the standard wedge for the 1423 with 1 mil Al wire (# 44129-1520-152) manufactured by Microswiss, a subsidiary of Kulicke & Soffa. All attempts to feed wire through these wedges were unsuccessful. It was concluded that the wedges were either not the indicated part for 1 mil wire or that the machining of the tool was defective. After difficulty obtaining any helpful information from Microswiss, arrangements were made to return the wedges.

Wedges were subsequently ordered from the Gaiser Tool Corporation[g]. The smaller size of this firm was better suited to the custom work required for the SVX. To facilitate bonding with our large vertical drop and small pitch, custom tungsten-carbide wedges were designed in cooperation with Gaiser engineers. The part number for these custom wedges is Gaiser #4560-1520-3/4-CBR(.001)-DSR(.005×.020). These wedges had extra material (double side relief) cut away from the sides of the wedge near the foot and performed extremely well in our application. The lifetime of a wedge was found to be $5,000$–$10,000$ wirebonds. As the wedge wore down the foot became chipped, taking on a porous appearance. The irregularity of this surface retained aluminum and the foot of the wirebond began to adhere to the wedge. When the bonding wedge began to lift wirebonds, replacement of the wedge was necessary.

Development of fixtures to hold the detector-hybrid assemblies during wirebonding was a key element in improving the process. Rigid support for the silicon was critical to consistent bonding. Without adequate support, the wafer vibrated, dissipating the ultrasonic energy used to create the bond. With early fixtures, ‘cold spots’ where good bonds were difficult to achieve were apparent. The fixtures were redesigned to place support as close to the bond pads as possible: support directly underneath was not possible due to the presence of wirebonds on both sides of the device. The fixtures used a mechanical clamp to hold the hybrid and a vacuum chuck to hold the silicon detector. Since positioning in these fixtures was not critical, the positioning pins and supports of these fixtures were machined from Teflon® to eliminate the possibility of damage.

The pull test stand was also used to test the space of wirebonding parameters. The three parameters were the length of the ultrasonic pulse, the power of that pulse and the

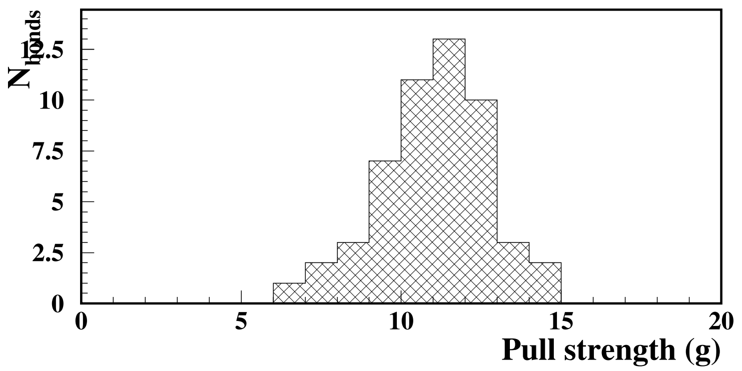
[g] Gaiser Tool Corp., Ventura, Ca.

Table 7.2: The standard machine parameters used for wirebonding of the silicon detectors.

Parameter	Optimum setting
Power	3.30
Time	2.0
Force	2.5

force with which the wire is pressed into the bond pad. In general, too little power, force or time resulted in a poor weld between the wire and the bond pad. Too much resulted in destruction of the bond pad or weakening of the wire adjacent to the bond (the 'neck'). Consistent pull strengths of more than six grams with breakage at the neck of the bond were achieved using the parameters shown in Table 7.2. Figure 7.2 shows a histogram of pull strengths for these settings.

Figure 7.2: Results of pull tests with the standard settings used for all SVX wirebonding.



Seven to nine grams is considered good pull strength for 1 mil wire. The observed failure mode is normal for wirebonds of good quality. These optimized settings were used for all wirebonding. It would be dangerous to assume that these parameters are universally optimized, even using identical parts on a similar wirebonder: significant differences exist among well-calibrated wirebonders of the same model.

A small decrease in both power and force ($P = 2.8$, $F = 2.0$) actually produced wirebonds with higher median strength. The strongest of these bonds broke at the pull-testing hook indicating that the tensile strength of the wire was the limiting factor. While these settings initially appeared to produce stronger bonds, consistency suffered. A small percentage of these wirebonds had very low pull strength, with the foot of the bond lifting from the bond pad. A histogram of pull strengths for these settings is shown in figure 7.3.

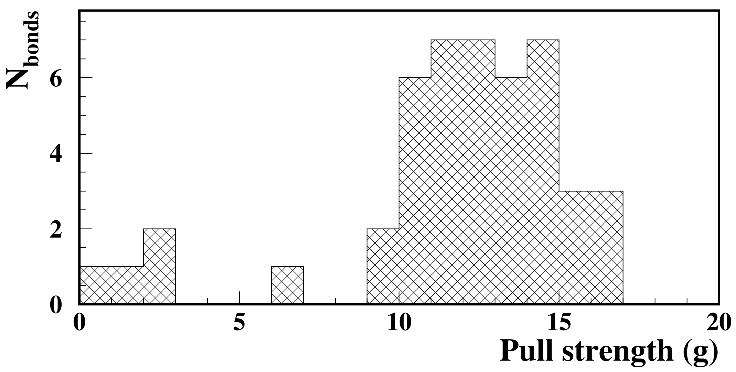


Figure 7.3: Results of pull tests with settings only slightly different from the standard settings.

The large vertical drop ($\approx$ 40 mils) from the CAMEX to the detector on the n-side was near the maximum vertical range of the wirebonder. This posed a serious threat to the operation of the SVX. If the wirebonds were 'pinched' at the edge of the CAMEX chip, electrical contact between the wire and the CAMEX substrate created inoperable channels. The wirebonding protocol was altered to ensure that all wirebonds were clear of the CAMEX. Optimally adjusted, the clearance between the wirebonds and the edges of the CAMEX on the z-side ranged from 25 to 30 microns.

The rate of successfully connected channels was in excess of 99.9%. No detectors were destroyed as a result of failed wirebonds. As far as we are aware, not a single channel in the SVX has experienced failure due to faulty wirebonds to the detectors.

7.1.3 Thermal Conductivity

Good thermal conductivity of the support structure is critical to the performance of the SVX for two reasons. The first is that the epoxies used to hold the SVX together begin to soften at 70-80°C. The second is that the leakage current and noise of the silicon detectors increase as a function of temperature.

The BeO used in the support structure has very high thermal conductivity compared to any adhesive. With eight glue joints along the heat path, minimizing the amount of adhesive used and maximizing its thermal conductivity were important.

A test stand was built to verify the thermal conductivity of EG765 from A.I Tech and explore the possibility of producing our own diamond-filled epoxy to avoid the high cost of EG765. This stand had test units for parallel testing of four samples. A single test unit is depicted in Figure 7.4 and consisted of two copper blocks glued together with a gap of known thickness and cross section. A 10W resistor applied known power to one block while

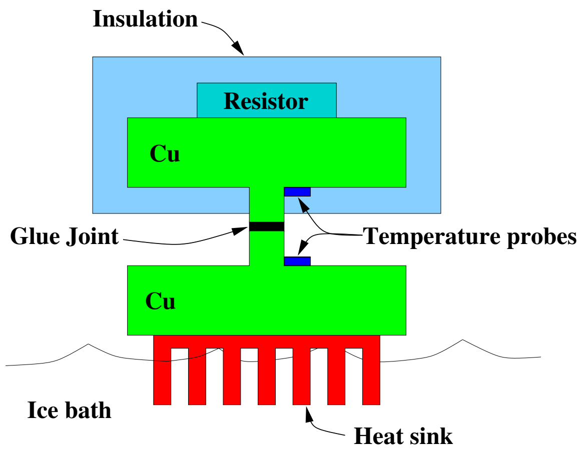


Figure 7.4: A drawing of the apparatus used to test the thermal conductivity of various adhesives.

the other block was placed in an ice bath. Measuring the temperature drop across the gap allowed a calculation of thermal conductivity to a precision of 10-25%, depending on the properties of the adhesive and the thickness of the glue layer.

During testing, it was discovered that the stated design specification of 11W/(m-°K) for EG765 could only be met at glue gaps similar to the size of the diamond dust used to fill the epoxy. At such gaps, all heat flows through diamond particles bridging the entire gap. At the larger gaps called for in the support structure thermal conductivity approached 1W/(m-°K).

Thermal conductivities of 1.5W/(m-°K) were achieved filling West Systems epoxy (resin 105, hardener 206)[h] with large amounts of diamond dust. However, strength suffered at the higher filling ratios necessary ($\approx$ 30%). A commercially available boron-nitride filled epoxy achieved the advertised conductivity of 3.3W/(m-°K) at a wide variety of glue gaps, but the paste-like consistency of this adhesive, with large chunks of boron nitride, was unacceptable for gluing hybrids together. Slightly over-thick glue layers caused hybrid cracking when the fixture (described later) was tightened to achieve the proper overall thickness. Silver epoxy achieved 1.5W/(m-°K) and was thin enough to use for the large area glue joints between hybrids. This layer also formed a ground plane for each hybrid pair.

It was not possible to use the silver epoxy for hybrid-to-endpiece glue joints. The conductive epoxy would have shorted exposed traces on the surfaces of the hybrids. By filling West Systems epoxy with boron nitride powder, we produced our own electrically insulative epoxy with a thermal conductivity in excess of 2W/(m-°K).

7.1.4 High-yield Assembly Process

The planned rate of spare, good-quality detectors and hybrids was 25%. By keeping the yields in the assembly process high, we hoped to assemble ten functioning octants in the case that spares were needed for testing or replacement. Since the typical rate of spare detectors for construction of a silicon vertexing detector is closer to 50%, it was clearly a challenge to achieve this high yield. The primary ingredient to success was the strict testing regimen and the modular approach of the construction process.

The octant is the main modular unit of the SVX. The hierarchy of sub-units within an octant is shown in Figure 7.5. This not only simplified and accelerated construction, but allowed for careful electronic and mechanical testing of each module before committing it to other modules and further processing. The following sections summarize each step in this process, from the assembly of individual parts to the insertion of the octants into the carbon-fiber housings.

[h] Gougeon Brothers Inc., Bay City, Mi.

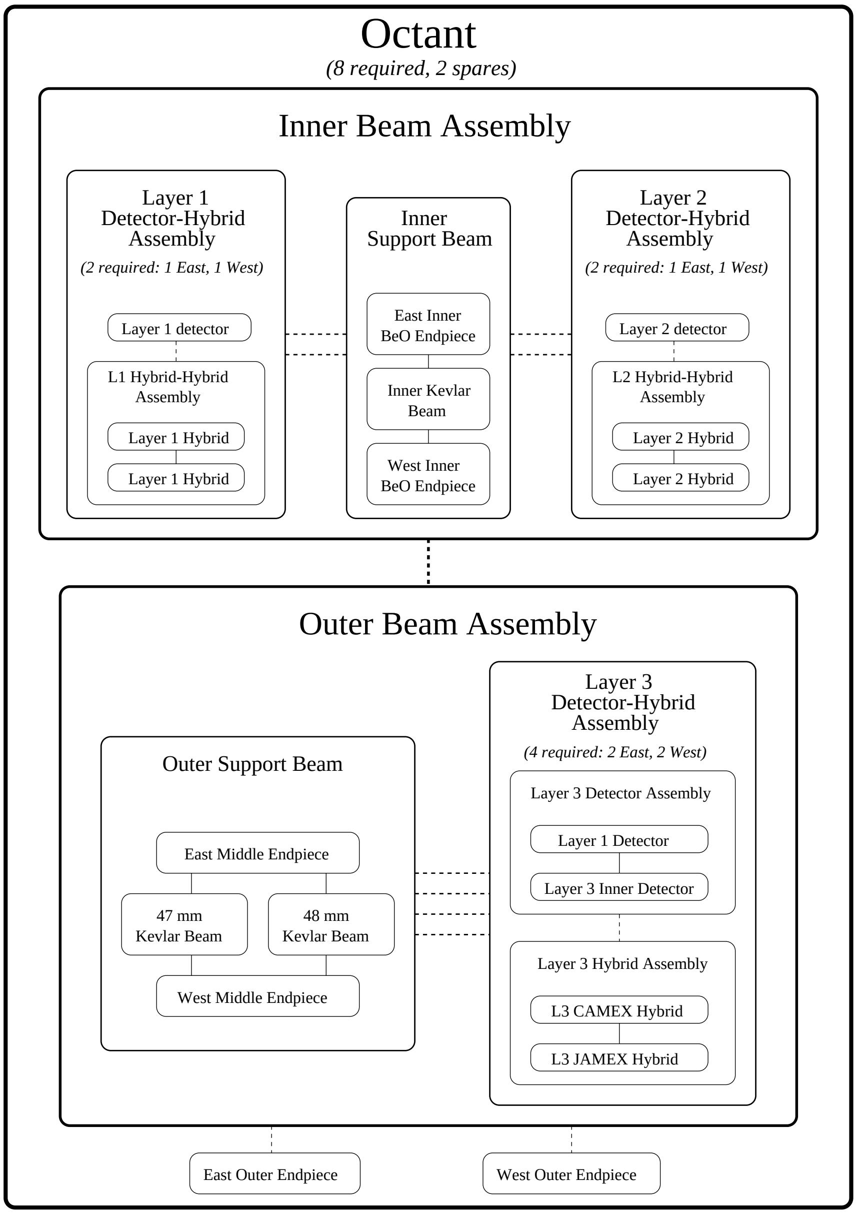


Figure 7.5: The hierarchy of modules within an octant. Each outlined module underwent testing before bonding to other modules. Glue bonds are shown as lines between modules. Those shown as dashed lines used the platens for positioning, while those requiring special fixtures are shown as solid lines.

7.2 Detector-hybrid Assemblies

The detector-hybrid assembly is basic functional sub-unit of the SVX. Each consists of one or two silicon detectors and a pair of hybrids. The steps involved in producing a detector-hybrid assembly were:

- testing and assembly of detectors,
- testing and assembly of hybrids,
- detector-hybrid glue bonding,
- detector-hybrid wirebonding,
- detector-hybrid testing and burn-in.

The following subsections describe the details of these steps.

7.2.1 Silicon Detectors

The silicon detectors were received from Hamamatsu Photonics along with lists of pinholes for each detector. Detectors then underwent the following set of processes:

- inspection and mechanical survey,
- electronic testing (at Cornell),
- edge bonding of layer three detectors.

Inspection and mechanical survey

Each detector was inspected and surveyed using the optical microscope. The purpose of the mechanical survey was twofold. The first was to check for obvious defects such as broken traces or defective passivation. In some cases bias traces were broken, a defect rendering the detector inoperable. In other cases, defects in passivation indicated substandard processing, which could lead to higher noise and pinhole rates. The second purpose was to measure the position of the sense strips relative to the edges of the detectors. The positions of the sense strips on all detectors were catalogued as the SVX was built up by surveying the visible edges of the detectors after each assembly step.

Table 7.3: A small subset of survey results from layer one east and layer two silicon detectors. Note that the 'active centers' are similarly displaced for the ten surveyed detectors of each type.

Parameter	Nominal(μm)	(Measured−Nominal)(μm)
Thickness	300μm	-1 ± 2 μm
Layer 1		
r-ϕactive length	21,168	-1.8 ± 0.8
r-ϕactive center	11,284	12.2 ± 2.2
r-ϕangle	0	$-7 \pm 5\mu$rad
z active length	59,535	-9.2 ± 1.6
z active center	35,032.5	-11.5 ± 2.2
z angle	0	$50 \pm 34\mu$rad
Layer 2		
r-ϕactive area	28,614	-1.7 ± 0.5
r-ϕactive center	15,064	16.4 ± 3.5
r-ϕangle	0	$9.2 \pm 58\mu$rad
z active length	76,356	-5.2 ± 1.7
z active center	42,392	-12.6 ± 3.3
z angle	0	$12 \pm 109\mu$rad

These surveys showed that the pitch of the sense strips accurately reproduced the design specification to within several microns over the entire length and width of the sense regions. Meanwhile, the sense areas were similarly shifted on all detectors of the same type, allowing adjustment of the fixtures to more accurately place the sense regions. Table 7.3 shows typical values for the misplacement of the sense areas of the detectors. All features on the detectors were measured relative to a 'reference corner' that was defined by the fixturing method used for the silicon detectors and detector-hybrid assemblies. This method and the resulting reference corner are depicted in Figure 7.6. Using this corner as the origin for all surveys provided the highest possible precision of the assembly and surveying process.

Two interesting details of the silicon detectors were discovered during surveying. The first was a bevel in the cut edges of the silicon. This bevel resulted in the appearance that the r-ϕ side was shorter and narrower than the z side by $\approx 5\mu$m when held flat by the vacuum chuck on the microscope stage. The second was that the silicon detectors were not flat. The

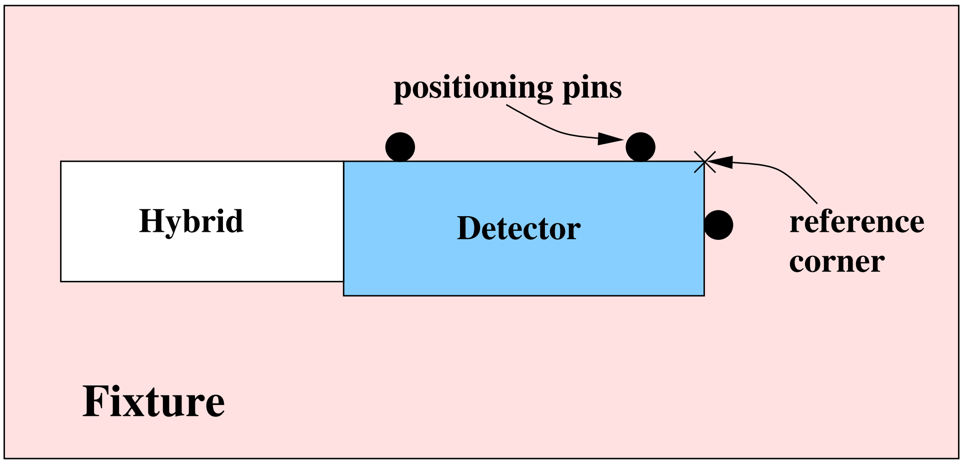


Figure 7.6: A diagram showing the method used to fixture the silicon detectors for glue processes. The three positioning pins defined a natural 'reference corner,' which served as the origin for all assembly and surveying of the detectors.

lengthwise sagitta of a layer one (layer two) detector was approximately $17 \pm 2\mu$m($31 \pm 5\mu$m). The widthwise sagitta of the detectors was consistent with zero.

Electronic testing

At Cornell University, each silicon detector underwent a series of tests known as the sidev-battery, or 's-battery.' These included a cross-check of the pinhole lists and tests of readout strip capacitance. The latter was capable of identifying discontinuities in the readout and sense strips with no bias voltage applied, as well as the onset of full depletion, by examination of the capacitance vs. bias voltage curve. Capacitance increases with bias voltage until full depletion occurs, after which capacitance remains constant despite increases in bias voltage. Typical full depletion voltage for the detectors was 60V.

Edge bonding of layer three detectors

Two detectors were bonded end-to-end with no overlap to produce the longer detector assemblies for layer three. Because both detectors are read out through the same strips, a bad channel in the outer detector renders a perfectly functional channel on the inner detector useless. A program was written to find the optimal set of detector pairs, the one with the maximum number of matching dead channels. Approximately 40 layer three inner detector channels were salvaged.

Several epoxies were tested for strength and purity. Strength testing was performed by edge bonding two glass microscope slides. The best epoxies produced a bond stronger than the slides themselves. Purity testing involved both laboratory testing by an independent laboratory and application of various glues to the surface of an operational detector. Noise on the channels underneath each sample was monitored carefully. Any increase in noise of the adjacent channels indicated leaching of ionic impurities into the bulk silicon. TRA-DUCT 2112, a standard electronics adhesive, was used used for all glue joints involving the silicon detectors.

A controlled amount of epoxy was deposited using a machined aluminum block with a shallow groove. Epoxy was applied to the block and excess scraped off with a blade. This left a controlled bead in the groove that could be transferred to the edge of a detector. Assembly took place using an aluminum fixture with two halves whose separation was controlled with a micrometer knob. Epoxy was applied to one detector, both were held to the fixture by vacuum chuck and brought into contact under optical microscope to the proper gap of 3 mils.

Scratches caused by fixturing of the silicon were found to create new pinholes. Teflon® tape was applied to the assembly fixtures to protect the silicon. The detectors with an extra layer of silicon nitride passivation were much less susceptible to damage-induced pinholes.

After glue-bonding, the readout channels of the detector assemblies were wire-bonded together and RTV adhesive was added to protect the wirebonds.

7.2.2 BeO Hybrids

The BeO hybrids were received at Cornell University where initial testing was done and Kapton® flex circuits were attached. The Kapton® was glued to the hybrids using RTV adhesive. Wirebonds between the copper traces on the Kapton® and the gold bond pads on the hybrids were made using 1 mil aluminum wire. Wirebonds were then potted with RTV adhesive for protection.

After delivery to UCSB, the hybrids underwent the following processes:

- mechanical survey,

- electronic testing,
- hybrid-hybrid assembly.

Mechanical Survey

Mechanical survey of the hybrids verified two critical physical dimensions. The first was the thickness of the hybrid in the central glue pad, where the hybrids attach to the BeO support structure. Maintaining the design thickness of the hybrid assemblies was critical since excess thickness anywhere in the support structure might use up the minute clearance ($<$ 1mm) between the layer one wirebonds and the beampipe. On average, the hybrids were found to be 1–2 mils thicker than specified. Thicker hybrids were paired with thinner ones to keep the glue gaps as consistent as possible. The second critical dimension was the position of the CAMEX chips. Badly misplaced CAMEX chips make wirebonding between the detectors and the CAMEXes difficult. In general, the precision achieved by the hybrid manufacturer in the placement of the CAMEX was sufficient, although not to the specifications of the design. One standard deviation in lateral CAMEX placement was approximately 3 mils, or almost one bond pad of offset.

Electronic Testing

The most important testing of the hybrids was of the electronic performance. The standard battery of tests was referred to as pedestals-gains-noise, or PGN, testing. The hybrids were attached to data acquisition electronics and voltage was applied at the test inputs of the CAMEX chips. The purpose was to produce a list of inoperable channels and provide a baseline for the PGN of all CAMEX channels. Bad channels were not bonded to the detector and changes in PGN tests were tracked carefully for each CAMEX as the detector was assembled.

Hybrid-hybrid Assembly

There were two critical properties of the glue joint between hybrids pairs, thickness and thermal conductivity of the bond. The fixture used for gluing hybrids consisted of two pieces of machined aluminum that are bolted together to enforce the specified overall thickness of the hybrid assembly. To achieve high thermal conductivity, a bond of largest possible surface area was required. This dictated careful deposition of a thin glue layer to meet specified thickness without bubbles or excessive overflow. The solution was machining rectangular windows in brass shim stock of varying thickness to create a stencil. Glue was applied to the hybrid inside the stencil and a blade was used to level the glue layer to the shim thickness.

7.2.3 Detector-hybrid Glue Bonding

Glue bonding of detectors to hybrids was performed using fixtures attached to the platens. Both fixtures used vacuum chuck to hold the parts. Three precision stainless pins, two along the side of the detector and one at the end, positied the silicon exactly as was shown in Figure 7.6. Early attempts to use these fixtures resulted in damage caused by the aluminum support pad and stainless positioning pins. These problems were eliminated by applying Teflon® tape to critical surfaces of the fixtures.

7.2.4 Wirebonding

Wirebonding of the readout channels to the CAMEX inputs was the final step in producing working detector-hybrid assemblies. Wirebonds between detectors and hybrids, and between detectors in layer three were produced at UCSB. After setting up the wirebonder according to a standard worksheet, all channels without known pinholes were wirebonded. The typical time to wirebond a single CAMEX (63 channels) to the detector was twelve to twenty minutes depending upon the operator.

7.2.5 Testing and Burn-in

After assembly was complete, each detector-hybrid was subjected to careful testing. PGN testing was repeated to provide a baseline for each assembly. If new pinholes were discovered, the part was sent back to the wirebonding station where the bonds to pinhole channels were removed.

A computer controlled test stand was developed to test the response of each channel to infrared light. A pulsed infrared LED was mounted on a platform positioned by stepper motors. A programmed sequence pulsed each channel with penetrating infrared light, measuring the response of each channel on both sides of the detector.

Detector-hybrid assemblies were stored with CAMEX power and bias voltage applied until assembled with the support structure. Special fixtures and cabinets were built to provide power, monitor voltages and currents and keep the assemblies in a dry, dark environment. Pinholes resulting from this burn-in were identified by a final PGN test and disconnected from the CAMEXes at the wirebonding station. After all testing was completed, wirebonds were potted for protection using an RTV adhesive.

7.3 Octant Assembly

Completed detector-hybrid assemblies were mated to the support structure to create the octants to be inserted into the carbon fiber housings. The steps in completing an octant were

- assembly of support beams (inner and outer),
- attachment of detectors to support beams (inner and outer),
- final octant assembly.

7.3.1 Support Beam Assembly

Before attachment of the detector-hybrid assemblies, the BeO endpieces were mated with the Kevlar U-channels to produce finished support beams. The steps in assembly of the support beams were

- production and surveying of BeO endpieces,
- production and surveying of Kevlar® U-channels,
- glue-bonding of endpieces to U-channels.

BeO Endpieces

The fabrication of the BeO supports was contracted to the Ceradyne Corporation[i]. Following delivery, all BeO endpieces were subjected to careful surveying of critical dimensions. As with the hybrids, the radial thickness is critical due to the tight clearances between

[i]Ceradyne Corp., Costa Mesa, Ca.

the layer one wirebonds and the beampipe. Although all pieces were within tolerances, some were returned to Ceradyne for machining. This was done because the hybrids tended to be one to two mils thicker than designed. The brass sleeves were glued into the outer endpieces to accept the pins and screws that attach the octants to the carbon fiber housings.

U-channels

All Kevlar® U-channels were produced at UCSB. Molds were machined from aluminum and the layup was three layers; two layers of woven Kevlar® fabric sandwiched around a thin Kevlar® mat. West Systems epoxy completed the composite.

The beams were then machined to remove excess material and machine the windows added for wirebond clearance and material reduction. The abrasiveness and high tear strength of Kevlar® is very destructive to conventional machine tools and produces a raggedly finished edge. Laser machining of the beams by LaserFab[j] was necessary to produce cleanly cut edges.

The removal of significant material relieved internal stresses inherent in the U-channels, causing warping of the beams. To reduce this warping, the parts were placed in the molds and recured. Slow, controlled heating and cooling allowed the epoxy to soften and recure in the desired shape.

Support Beam Glue-bonding

The inner and outer support beams were glue bonded to the BeO endpieces using special aluminum fixtures designed for this task. An inner support beam consisted of two inner endpieces and an inner U-channel. An outer support beam consisted of two middle endpieces and two U-channels, one for each outer sub-layer. In both cases, all parts were fixtured and glued simultaneously.

7.3.2 Attachment of Detector-hybrid assemblies

Detector-hybrid assemblies were attached to the support beams in east-west pairs. The detectors were glued to the U-channels with TRA-DUCT 2112 and the hybrids were glued to the BeO endpieces with our thermally conductive adhesive. Only small dots of glue were placed at 2mm intervals along the U-channel to attach to the silicon. This ensured the

[j] LaserFab Corp., Concord, Ca.

smallest possible amount of material, the lowest exposure of the silicon to the glue and the possibility of removing a detector-hybrid assembly if necessary. After each set of detectors was attached, the positions of the detectors were measured relative to one another to build the survey of the finished octant.

For the inner beam assembly, layer one detector-hybrids were attached to the support beam first. This assembly remained in the fixture for the layer one detector-hybrids, and the layer two detectors were attached. Since the layer one detectors were never removed from their fixture, very high relative positioning of layer one and layer two was possible. After layer two detector-hybrids were attached, the assembly again remained in the fixture that initially held the layer one detectors.

For the outer beam assembly, the inner sub-layer for layer three was attached first. This assembly remained in the fixture for the support beam. The outer sublayer was then attached. Finally, the outer support beam was mated with the outer endpieces, with the assembly remaining in the fixture for the endpieces. This fixture ultimately held the finished octant.

7.3.3 Octant Assembly

Once all detectors were attached to the support beams the inner and outer beam assemblies were joined together. They connect only at the support structure, simplifying this step.

All visible corners and edges of the detectors were measured. Although some reference corners and edges were obscured during surveying, those still visible were sufficient to completely constrain the positions of all sensitive regions of the detectors.

Each detector in the octant was subjected to the standard PGN tests to form a baseline for the performance of all CAMEX channels in that octant. When this testing was complete the octant was ready for insertion into the carbon fiber housing

7.4 Final Assembly and Transport

Complete octants were placed, four each, inside the two semi-cylindrical carbon fiber housings. The housings, with extensions mounted, were mated around the new Be beampipe at Cornell, and the entire assembly was then inserted through the rest of the CLEO detector from the end. Before commitment of the octants to the housings, the strength of these pieces were carefully tested. Finite element analysis was used to estimate the deformation of the completed housing under load. This unit was also tested under an actual load to ensure that it was capable of supporting the shields.

The overlaps between detector octants created very tight clearances between detectors of adjacent octants. To perform the delicate task of positioning the octants in the CF housings, a motorized assembly machine was built. This machine held the semi-cylindrical carbon-fiber shells on a motorized, rotating mount. By rotating and lowering the carbon fiber, it could be positioned for vertical insertion of each octant.

The octant to be placed was held stationary and the housing is lowered from above and into position against the octant. The octant was attached by hand with the positioning pins and screws. Multiple lowering and rotating steps were required insert an octant with an already-installed overlapping neighbor.

The detector was crated for transport via Emery Air Freight. The crate was instrumented with several shock detectors and a chart recorder to measure acceleration during transport. Within minutes of leaving UCSB, the SVX experienced a severe shock. It is most probable that this was due to the crate being improperly tied down in the delivery truck. Upon inspection at Cornell, it was discovered that a layer one detector became separated from the U-channel. A small chip was also missing from the detector. No electronic failures or mechanical instabilities appear to have resulted from this damage.

Chapter 8

Detector Performance

The detector was installed during October 1995 and first took data on October 29, 1995. An event display of an early Bhabha ($e^+e^- \rightarrow e^+e^-$) event is shown in Figure 8.1. As of October 6, 1998, the detector has collected 7.5 fb^{-1} of total luminosity.

The adjusted total radiation dose through August 9, 1998 is shown in Figure 8.2. The adjusted dose represents an attempt to estimate the damage seen by layer one CAMEX chips and is defined as

$$\text{Adjusted dose} = (\text{Power-on dose}) + 0.3 \times (\text{Power-off dose}). \tag{8.0.1}$$

The doses at which we expect performance degradation and complete failure are shown. In July, some CAMEX channels in layer one were observed with increased noise and decreased efficiencies. It is thought that this degradation indicates the onset of radiation damage to the SVX.

After much work in eliminating noise in the SVX, the S/N ratio in all detectors are within acceptable limits. The S/N on the r-ϕ (p) and z (n) sides of the SVX detectors are shown in Figure 8.3. The pulse height distributions shown in Figure 8.4 show the pulse heights from tracks at normal incidence in the detector. The distribution, a Landau convoluted with the incoherent noise, shows that the detectors efficiently collect the electron-hole pairs produced by the passage of charged particles.

There are three main classes of bad channels in the SVX. The first is non-functioning CAMEX channels. Most of these channels were non-operational at the time of installation.

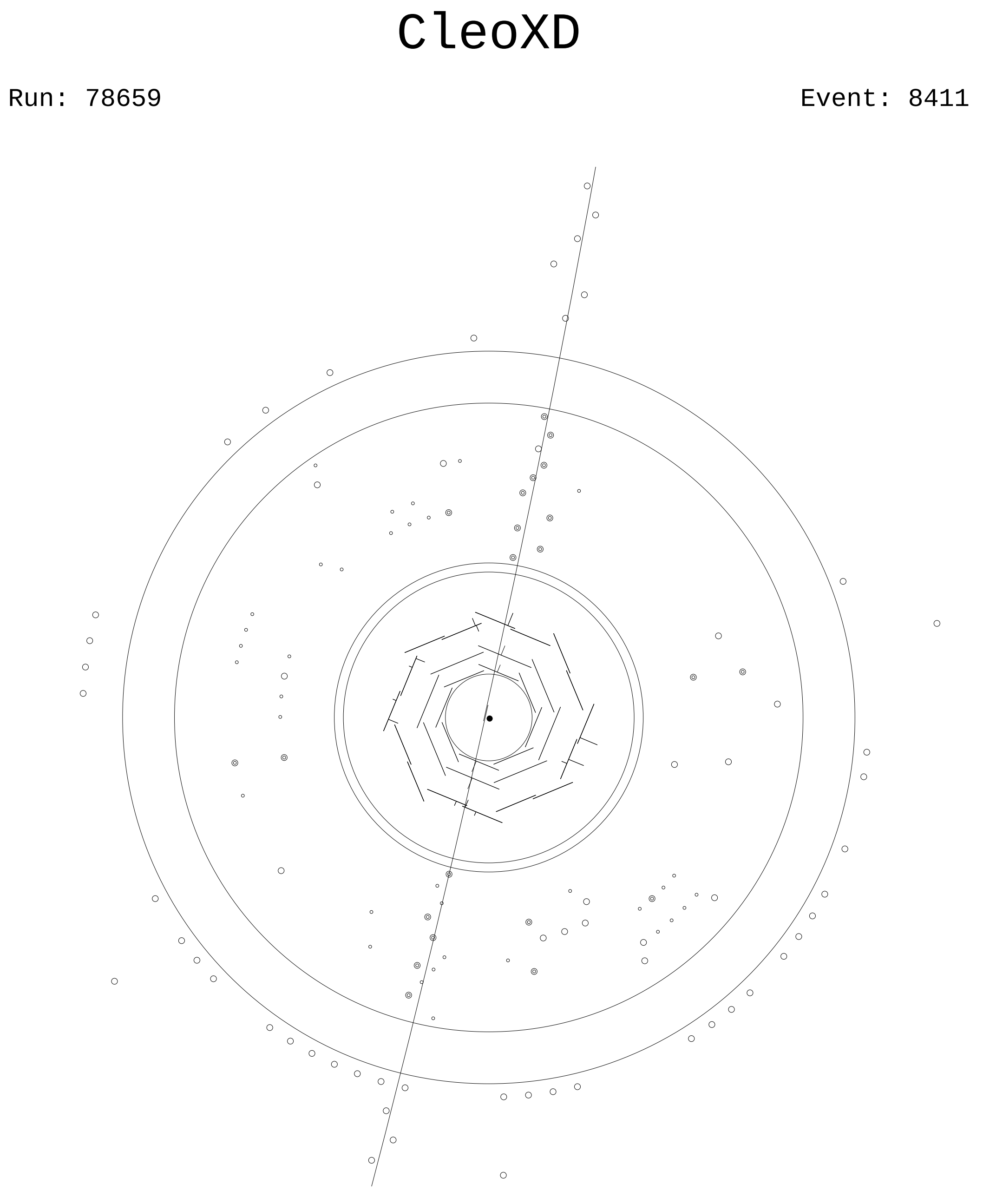


Figure 8.1: A display of an early e^+e^- event in the SVX. The hits in the SVX are shown but were not used to fit the tracks shown. Some misalignment of the entire SVX with respect to the VD is apparent.

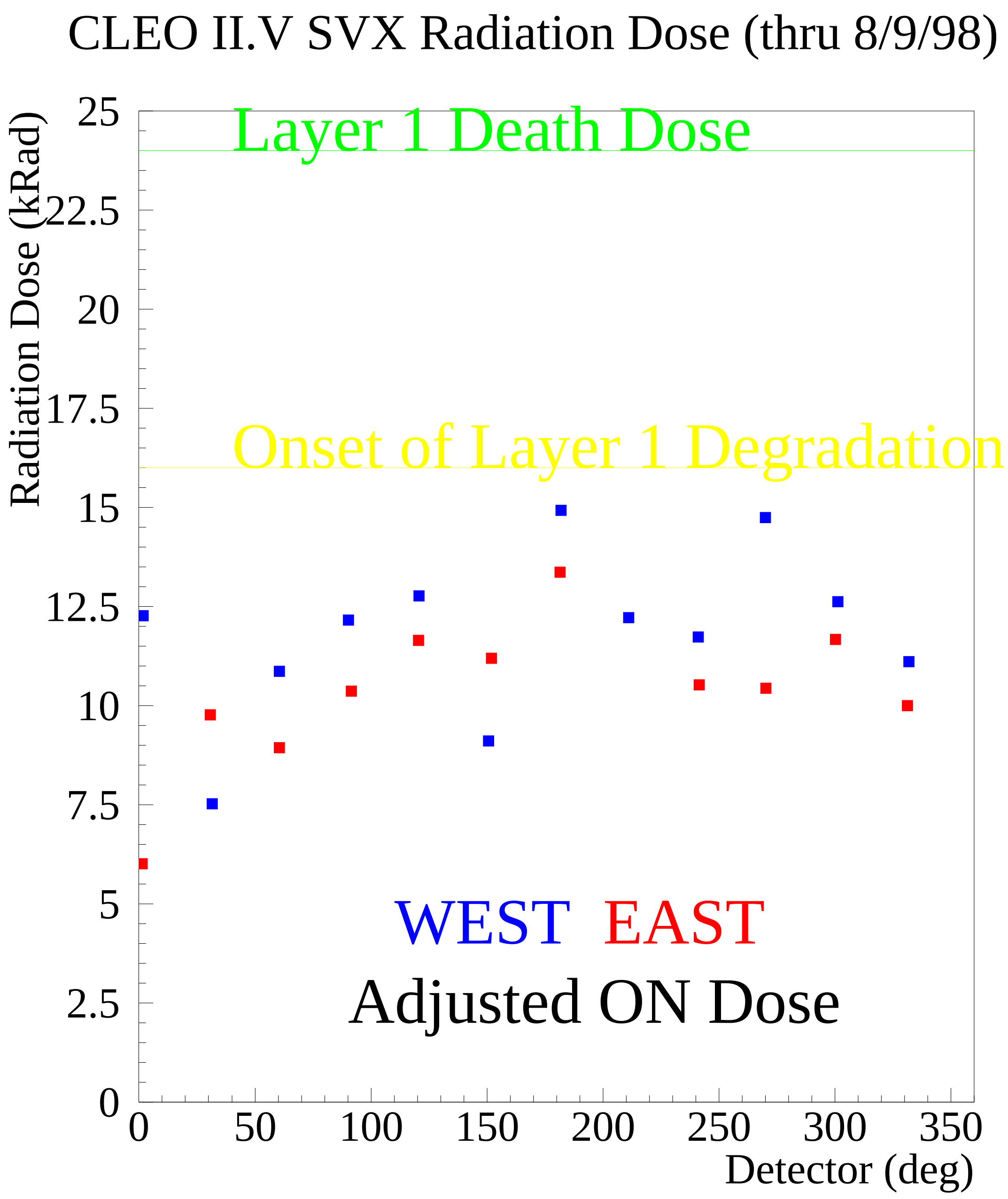


Figure 8.2: A plot of the adjusted radiation doses of the layer one detectors. Adjusted dose, defined in the text, provides a rough measure of the expected lifetime of the SVX. Radiation is measured by monitors mounted near the layer one CAMEX chips of each octant.

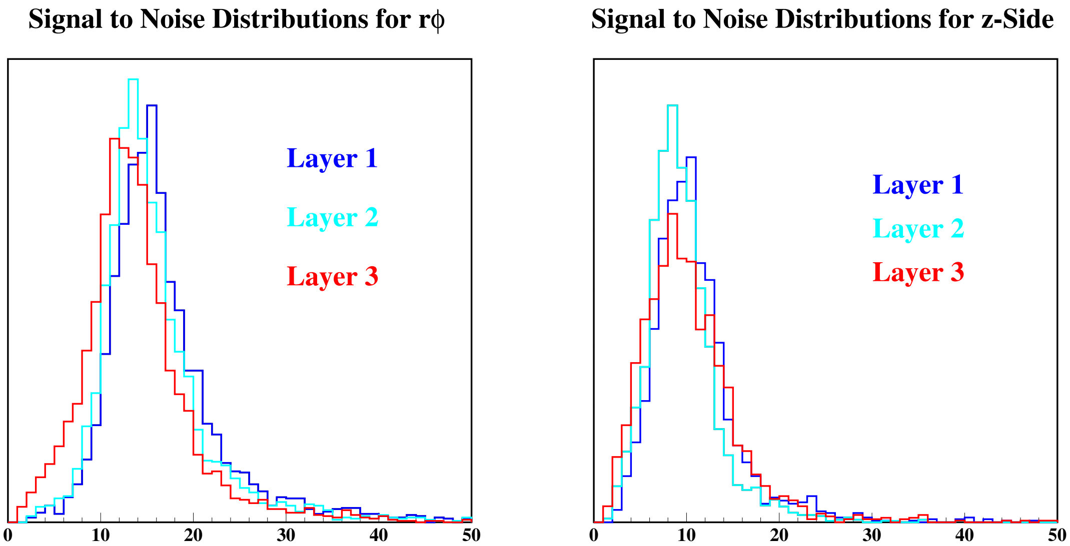


Figure 8.3: A plot of the signal to noise ratios for the r-ϕ and z sides of the three SVX layers. For most hits the noise is within acceptable limits.

However, some recent failures are thought to be due to radiation damage. The second class are pinhole channels. Again, most of these were faulty at installation, although some new pinholes have developed over time. The third class is entire detectors that have become non-functional. It is thought that these failures are due to the wirebonds between the hybrids and the Kapton® flex circuits.

These wirebonds, produced at Cornell with 1 mil aluminum wire, are extremely difficult to create reliably, but critical to the operation of the hundreds of channels on each detector. Both the Kapton® itself and the RTV used to glue the flex cable to the hybrid are very soft. Our experience at UCSB shows that this substrate tends to absorb much of the ultrasonic energy that creates the wirebond. In addition, the metallization on the flex cables are copper: creating a good weld to copper with the softer aluminum wire is quite difficult. In retrospect, redundant wirebonds between the Kapton® cables and the hybrids would have been the best solution. The large size of the bond pads would easily have allowed for several connections for each signal to the hybrid.

Regardless of the unusable channels, the efficiency for detecting the passage of a charged track is quite high. The inefficiencies are shown in Figure 8.5 and are roughly

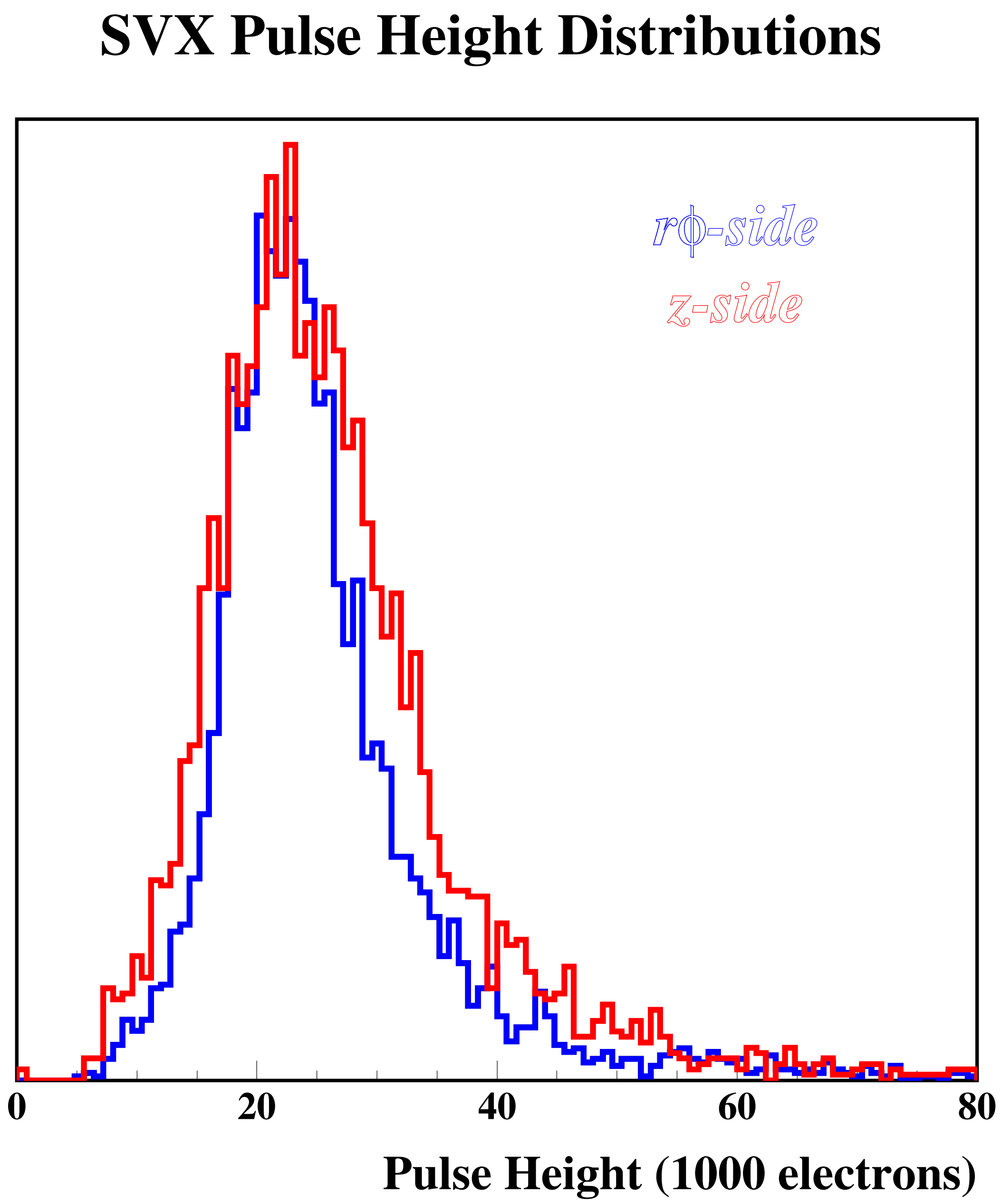


Figure 8.4: A plot of pulse heights on the r-ϕ and z sides of the silicon detectors. The Landau distribution is convoluted with the gaussian distribution of white noise. As expected the noise is larger on the z-side.

consistent with those expected from the number of bad channels.

The intrinsic resolutions (no multiple scattering) for tracks as a function of angle of incidence are shown in Figure 8.6. The resulting resolutions on the position of the CLEO interaction point for high momentum tracks are 39 and 45 μm in r-ϕ and z respectively. Figure 8.7 shows clearly that the SVX can be used produce cleaner samples of D decays than was previously possible with the CLEO detector. As our understanding of the subtleties of the CLEO II.V datasets improves, the data will produce some striking new results in the physics of B and D mesons and τ leptons.

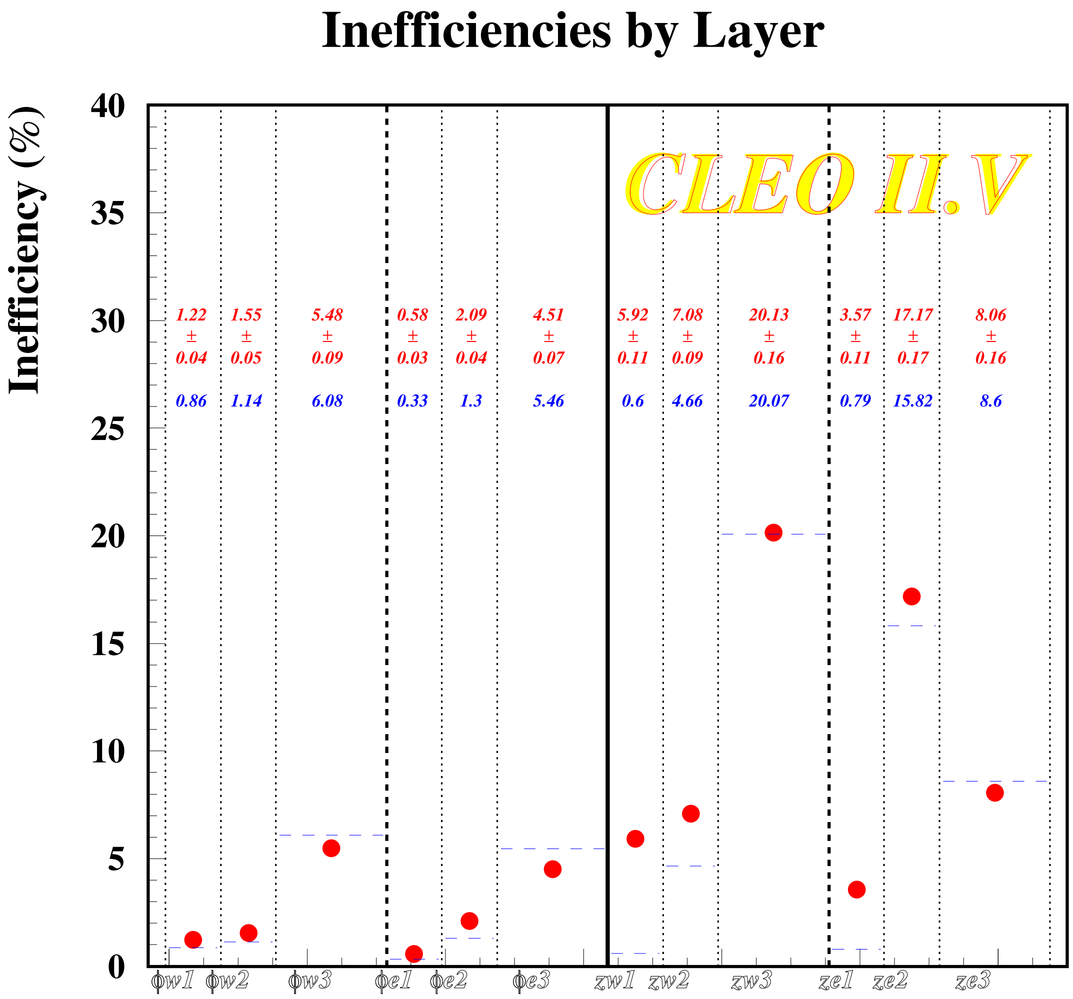


Figure 8.5: Plots of the efficiencies for detection of charged tracks by the various segments of the SVX. The notations w and e refer to the detectors on either side of the interaction point in z. The points (top numbers with errors) are the calculated efficiencies while the histogram (bottom numbers) are the expectations based on the rate of unused channels.

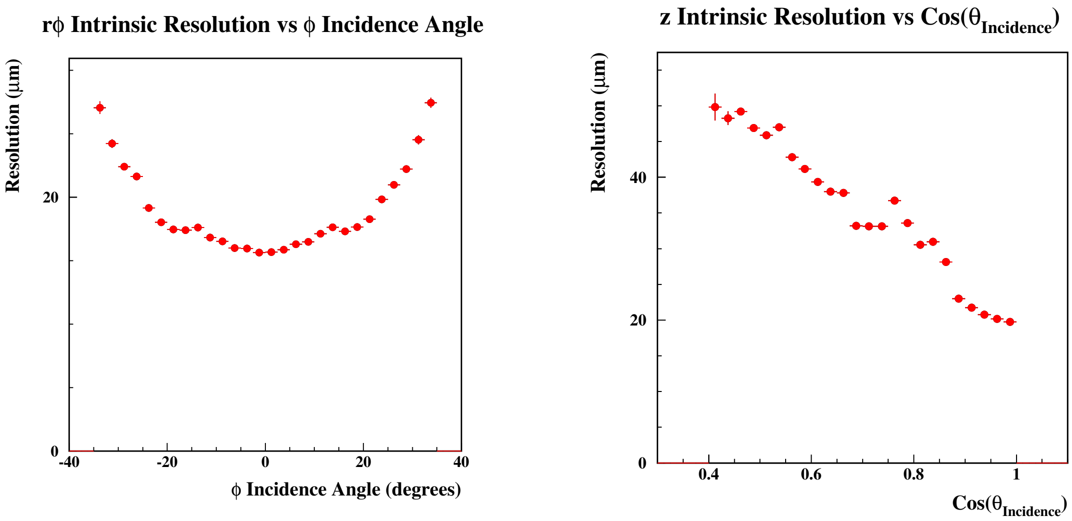


Figure 8.6: A plot of the single-hit resolutions for the SVX as a function of incidence angle for high-momentum tracks. Tracks with higher angles of incidence spread their charge over a larger number of readout strips.

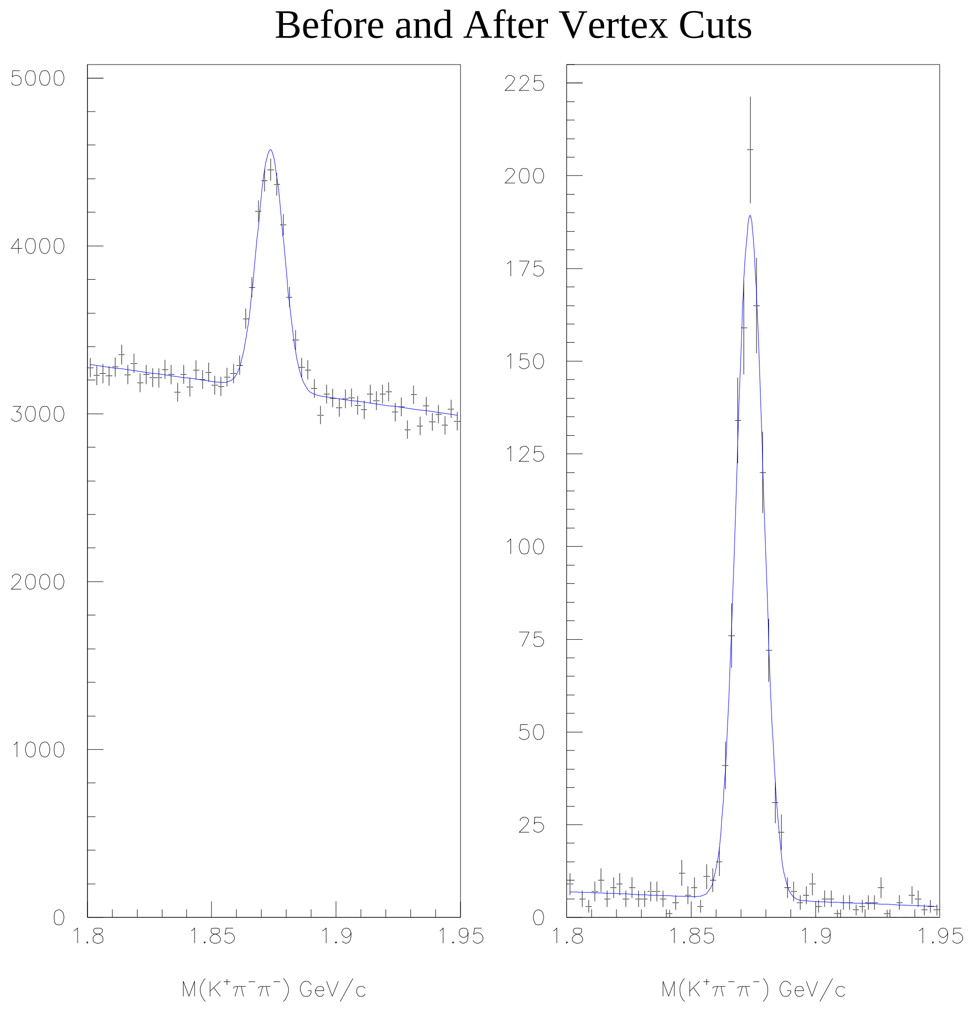


Figure 8.7: A plot of $K^+\pi^-\pi^-$ mass with (right) and without (left) the requirement of a separated vertex using the SVX. The bins are 2.5 MeV in width.

Bibliography

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

[2] C. Caso *et al.*, European Physical Journal C **3** (1998) 1.

[3] A. Bean, C. Darling and K. Kinoshita, CLEO-CBX 98-33a (unpublished).

[4] CDF Collaboration, F. Abe *et al.*, FERMILAB-PUB-98/188-E (1998).

[5] CLEO Collaboration, Phys. Rev. Lett. **65** (1990) 2951.

[6] N. Isgur and M. Wise, Phys. Lett. B **232** (1989) 113.

[7] J. Rosner, Comments Nucl. Part. Phys. **16** (1986) 109.

[8] S. Godfrey and R. Kokoski, Phys. Rev. D **43** (1991) 1679.

[9] ARGUS Collaboration, H. Albrecht *et al.*, Phys. Lett. B **232** (1989) 398.

[10] ARGUS Collaboration, H. Albrecht *et al.*, Phys. Lett. B **221** (1989) 422.

[11] CLEO Collaboration, P. Avery *et al.*, Phys. Lett. B **331** (1994) 236.

[12] E687 Collaboration, P.L. Frabetti *et al.*, Phys. Rev. Lett. **72** (1994) 324.

[13] R. Kutschke, CLEO-CBX 98-65 (unpublished).

[14] J. Richman, CALT-68-1148 (1984).

[15] M. Jacob and G.C. Wick, Ann. Phys. (N.Y.) **7** (1959) 404.

[16] J.D. Jackson, Il Nuovo Cimento, XXXIV **6** (1964) 1644.

[17] G. C. Fox and S. Wolfram, Phys. Rev. Lett. **41** (1978) 1581.

[18] Roger Barlow, Nucl. Instr. and Meth. **A297** (1990) 496.

[19] D. Schmidt, R. Morrison and M. Witherell, Nucl. Instr. and Meth. **A328** (1993) 547.

[20] Numerical Recipes Online, http://www.ul.cs.cmu.edu/books/numerical_recipes/

[21] ARGUS Collaboration, H. Albrecht *et al.*, Phys. Lett. B **335** (1994) 526.

[22] CLEO Collaboration, F. Butler *et al*, Phys. Rev. Lett. **69** (1992) 2041.

[23] A. Falk and M. Luke, Phys. Lett. B **292** (1992) 119.

[24] J. Bartelt and S. Shukla, Annu. Rev. Nucl. Part. Sci. **45** (1995) 133.

[25] CLEO Collaboration, M.S. Alam *et al.*, Phys. Rev. D **50** (1994) 43.

[26] CLEO Collaboration, J. Gronberg *et al.*, CLEO CONF 96-25 (1996).

[27] M. Neubert, CERN-TH/97-240 (1997).

[28] A. Falk and M. Peskin, Phys. Rev. D **49** (1994) 3320.

[29] P. Colangelo, F. De Fazio and G. Nardulli, Phys. Lett. B **303** (1993) 152.

[30] C. Lattes *et al.*, Nature **159,**(1947) 694.

[31] R. Brown *et al.*, Nature **163,**(1949) 82.

[32] G. Rochester *et al.*, Nature **160,**(1947) 855.

[33] CLEO Collaboration, CLNS96-1398 (1996).

[34] H. Becker, *et al.*, IEEE Trans. Nucl. Sci. **36** (1988), 246.

[35] J.P. Alexander *et al.*, Nucl. Instr. and Meth. **A342** (1994) 282.

[36] J.P. Alexander *et al.*, Nucl. Instr. and Meth. **A337** (1993) 171.

[37] J. Sharka *et al.*, Conference Record of the 1991 IEEE Nuclear Science Symposium.