C2

UNIVERSITY OF CALIFORNIA

Santa Barbara

Measurement of Charm Semileptonic Form Factors

A Dissertation submitted in partial satisfaction

of the requirements for the degree of

Doctor of Philosophy

in

Physics

by

David Michael Schmidt

Committee in charge:

Professor Rollin Morrison, Chairperson

Professor Michael Witherell

Professor Gary Horowitz

November 1992

The dissertation of David Michael Schmidt
is approved:

Gary Horowitz

Michael Witbrock

Walter J. Karplus

Committee Chairperson

November 1992

© David Michael Schmidt 1992

ALL RIGHTS RESERVED

ACKNOWLEDGEMENTS

I would like to thank my advisers, Rolly Morrison and Mike Witherell, for their guidance and patience. I am especially grateful to them for granting me the opportunity to work on such an interesting project.

This work would not have been possible without the efforts of all those in the Fermilab E691 collaboration; thank you for allowing me to share in the fruits of your labor. I am especially grateful to Alice Bean, Jean Duboscq, Dan Sperka, and Audrius Stundžia for their help and companionship.

At Cornell, I would like to thank Jim Alexander, Marina Artuso, Brian Busch, John Dobbins, Chris Jones, and Frank Würthwein for their instruction and company during the long hours in the lab.

I thank all the staff at CLEO and UCSB, particularly Debbie Ceder.

Finally, I would like to thank my parents and family for their guidance and support over the years. And to my wife, Bert, who has been beside me the whole way, I give my most heartfelt appreciation for her counsel, love, and patience.

CURRICULUM VITAE

Born:	March 2, 1964, Columbus, OH	
Education:	Denison University, Granville, OH	1982–1986
	B.S. Physics with High Honors	June 1986
	University of California, Santa Barbara	1986–1992
	Ph.D. Physics	November 1992

PUBLICATIONS

"Measurement of the Form Factors in the Decay $D^+ \rightarrow \bar{K}^{*0} e^+ \nu_e$." J.C. Anjos *et al.*, *Phys. Rev. Lett.* **65** (1990), 2630.

"Studies of Performance and Resolution of Double Sided Silicon Strip Detectors." J.P. Alexander *et al.*, *IEEE Trans. Nucl. Sci.* **38** (1991), 263.

Measurement of Charm Semileptonic Form Factors

by

David Michael Schmidt

ABSTRACT

We report on the first measurement of the three form factors governing the charm semileptonic decay, $D^+ \rightarrow \bar{K}^{*0} e^+ \nu_e$. This measurement attempts to utilize all the form factor information contained in the data, observed by Fermilab photoproduction experiment E691, by performing a maximum likelihood fit using the complete five-dimensional decay distribution for this decay mode. The presence of significant smearing and acceptance effects prompted us to develop a generalized maximum likelihood fitting technique, which is also presented. In addition, we compare the measured form factors with those from predictions of various theoretical models and discuss the relationship between the charm form factors measured here and those that are needed in determining important elements of the Kobayashi-Maskawa mixing matrix.

TABLE OF CONTENTS

Chapter 1 Introduction 1

1.1 The Kobayashi-Maskawa Matrix 2

1.2 Beta Decay 2

1.3 Charm 4

1.4 Bottom and HQET 5

Chapter 2 Kinematics 9

2.1 Form Factors 9

2.2 Matrix Element 12

2.3 Helicity Amplitudes 14

2.3.1 Partial Waves 16

2.4 Decay Rate 17

Chapter 3 The Experiment and Data 21

3.1 Brief Description of Experiment 21

3.2 The SMDs 23

3.3 Extraction of Signal 27

3.3.1 Backgrounds 29

3.3.2 Magnitude of Background 30

3.4 Determining Kinematic Variables 31

Chapter 4 Information and Minimum Variance 33

4.1 Form Factor Parametrization 33

4.2 Information on Helicity Amplitudes 34

4.3 Information on Form Factor Ratios 37

Chapter 5 Fit Technique 41

5.1 Smearing and Acceptance 42

5.1.1 Smearing 42

5.1.2 Acceptance 44

5.2 Description of Technique 45

Chapter 6 Fit Results 52

6.1 Three-Dimensional Fit 52

6.2 Four-Dimensional Fit 58

6.2.1 Summary 62

6.3 Branching Ratio 63

6.4 The Normalized Form Factors 68

Chapter 7 Conclusions 71

7.1 Other Experiments 71

7.2 Comparison with Theory 73

Appendix A Cleo-II Silicon Detectors 76

A.1 Operation and Gain of the CAMEX 78

A.2 Noise Studies 81

A.2.1 CAMEX Without Detector 81

A.2.2 CAMEX With Detector 86

A.3 Noise as Information 88

A.3.1 Theory 89

A.3.2 Effect of CAMEX on Noise Spectrum 91

A.3.3 Testing CAMEX Noise Weighting Function 92

A.3.4 Evaluation of Total Noise Integral 93

A.3.5 Comparison With P-Side Noise 95

A.3.6 Comparison With N-Side Noise 96

A.4 Summary of Main Noise Results 97

REFERENCES 99

LIST OF FIGURES

2.1 Spectator diagram of the semileptonic decay, $D^+ \rightarrow \bar{K}^{*0} e^+ \nu_e$. 10

3.1 Illustration of the E691 spectrometer. 22

3.2 Cross section of an E691 SMD. 24

3.3 Top view of SMD telescope showing the relative location and orientation (x, y, or v) of each SMD plane. 25

3.4 Spectrum of $K^{\mp}\pi^{\pm}e^{\pm}$ masses for the standard cuts. The wrong-sign ($K^{\mp}\pi^{\pm}e^{\mp}$) distribution is superimposed (dashed line). 29

3.5 The $K\pi$ mass spectrum with right-sign (solid) and wrong-sign (dashed) combinations. 31

4.1 Illustration of where the different helicity rates are dominant in $\cos\theta_e$, $\cos\theta_v$ space. 36

4.2 Helicity rates as a function of t for $R_2 = 0.7$ and $R_V = 2.0$. Top insert shows the derivative of the helicity rates with respect to R_2 and R_V as a function of t. 37

4.3 Information on R_2 and R_V plotted as a function of $\cos\theta_v$,$\cos\theta_e$ and t,$\cos\theta_v$. 39

5.1 Variance of smeared t/t_{max}, $\cos\theta_v$, $\cos\theta_e$, and χ about the true values as a function of the true values. 43

5.2 Shape of acceptance as a function of the true variables for t/t_{max}, $\cos\theta_v$, $\cos\theta_e$, and χ. 45

6.1 Projection of the data (dots with error bars) and Monte Carlo (solid histogram) events onto $\cos\theta_e$, $\cos\theta_v$, and t/t_{max}. The Monte Carlo events are weighted using the measured form factor ratios. 58

6.2 Projection of the data (dots with error bars) and Monte Carlo (solid histogram) events onto χ. In (a) events are weighted with the coefficient which multiplies the $\cos\chi$ term in the decay rate. This brings out the $\cos\chi$ dependence. (b) is a standard projection which should have a small $\cos 2\chi$ dependence. 60

6.3 (a) S-wave Monte Carlo data (dots with error bars) and fitted Gamma distribution (solid curve). (b) wrong-sign data (dots with error bars) and its fitted Gamma distribution. 64

6.4 Histogram of $K\pi$ mass distribution together with that of the fitting function. The shape of the background used in the fit is also shown. 66

A.1 Schematic of a double sided silicon detector. 77

A.2 Schematic of a single channel of the CAMEX. 78

A.3 Plot of Gain vs. time between charge injection and closing of the next S switch. 79

A.4 Gain vs. width of S switch. 80

A.5 Total noise (a) and noise after subtracting the common mode component (b) vs. channel number. This is with all known noise-reducing techniques being used. Total noise (a') and noise after subtracting the common

mode component (b') vs. channel number. This is from a CAMEX without the testpulse lines grounded. 82

A.6 Average pedestal level vs. time between opening of R1 and opening of R2. 85

A.7 Noise squared plotted as a function of the sampling period, τ_s, for the p-side of a detector. The line is a fit to the data using a second order polynomial. 88

A.8 Noise squared plotted as a function of the sampling period, τ_s, for the n-side of the detectors. The line is a fit to the data using a first order polynomial. 89

A.9 Diagram of RC circuit showing thermal noise as a parallel source of current noise, I. 90

A.10 Calculated noise in ADC units on the output of a CAMEX with a sampling period of $5.8\mu s$, for an input sine wave voltage of $1mV$ through a capacitor of $4.7pF$. This is plotted vs. frequency, in MHz, of the sine wave. 94

A.11 Plot showing the theoretical shape of noise squared as a function of $\frac{\tau_s}{R(C+C_i)}$. 95

A.12 Fit of noise from the n-side of a detector. 97

LIST OF TABLES

4.1 Minimum errors on R_2 and R_V. 40

6.1 Results of fits for different volume dimensions. 55

6.2 Result of fits with 4 independent Monte Carlo samples. 56

6.3 Summary of results on R_2 and R_V and comparison with theoretical minimum errors. 62

7.1 Results from this work compared with those of other experiments. . 72

7.2 Comparison of measured form factors with theoretical predictions. . 73

1 Introduction

The Standard Model has been very successful in describing the phenomena of particle physics, both qualitatively as well as quantitatively. Although this model is fairly specific, there are many free parameters as well as lingering questions about the origin of its structure. For example, the Standard Model makes no prediction for the mass of the quarks and leptons nor does it explain why there are three generations. An interesting phenomenon that the model can incorporate, but not require, is the fact the the weak interactions couple to a mixed set of quark mass eigenstates. The weak decay of a quark occurs by virtual W emission just as for a lepton—in fact the coupling is exactly the same. The W couples only to leptons or quarks of the same family. If there were no difference between the weak quark eigenstates and the mass eigenstates, then the s quark would not decay. However, the s quark does decay (we observe the decay of the K meson) and we say that this is because it is part of the weak d quark eigentstate.

1.1 The Kobayashi-Maskawa Matrix

This is incorporated into the Standard Model by replacing the charge $-1/3$ quarks (d, s, b) in a weak interaction with the mixed quarks, (d', s', b'), such that

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.1}$$

Here V_{ij} are elements of the unitary Kobayashi-Maskawa mixing matrix which can be parametrized in terms of three real angles and one phase.[1] In principle, the elements of this matrix can be determined by comparing the decay rate of real W bosons to quarks versus leptons. This, however, is experimentally difficult since the the quarks manifest themselves as jets of hadrons so that the flavor of the initial quark is difficult to identify. The information that we do have on the Kobayashi-Maskawa matrix comes mostly from beta decay of nucleons or semileptonic decays of hadrons. This dissertation concerns itself with a semileptonic decay of the D^+ charm meson which is important, although indirectly, for the determination of some of the Kobayashi-Maskawa matrix elements.

1.2 Beta Decay

As a useful example, consider the determination of the Kobayashi-Maskawa matrix element, V_{ud}, by nuclear beta decay. Here a d quark in an initial nucleus, N, decays into a u quark and forms a final nucleus, n, with the emission

of an electron and a neutrino. The effective matrix element for this process is $\frac{G_F}{\sqrt{2}} V_{ud} h^\mu l_\mu$ where G_F is the Fermi coupling constant which can be determined by muon decay, $l_\mu = \bar{e}\gamma_\mu(1+\gamma_5)\nu_e$ is the leptonic current, and $h^\mu = \bar{u}\gamma^\mu(1+\gamma_5)d$ is the hadronic current. If the quarks were not bound in nuclei then the decay rate could be calculated directly from this matrix element. However, the quarks are bound and this can affect the interaction amplitude. To first order, the amplitude for the decay with bound quarks is

$$\frac{G_F}{\sqrt{2}} V_{ud} \langle n| \bar{u}\gamma^\mu(1+\gamma_5)d |N\rangle \cdot l_\mu \tag{1.2}$$

where $|N\rangle$ and $|n\rangle$ are the quantum states of the initial and final nuclei, respectively. Even though we can not calculate this matrix element perturbatively, we can use the fact that isospin is a very good symmetry of the strong interactions to estimate the matrix element between nuclei of the same isospin doublet.

If the initial and final nuclei have spin zero then the most general form of the hadronic current is

$$\langle n(p)| h^\mu |N(P)\rangle = f_+(t)(P^\mu + p^\mu) + f_-(t)(P^\mu - p^\mu) \tag{1.3}$$

where $t = (P^\mu - p^\mu)^2$ is the squared momentum transfer or the squared mass of the electron-neutrino system. This form follows from Lorentz invariance. The parameters, $f_\pm$, are functions of the possible Lorentz scalars, of which t may be taken as the only dynamical scalar. These parameters are known as form factors and completely determine the decay. By comparing the behavior

of both sides of eq. (1.3) under a parity transformation one can see that only the vector term in the hadronic current survives; hence these form factors are called vector form factors. In order to determine V_{ud} from the measured decay rate the value of the form factors must be known.

In the limit that isospin is an exact symmetry, the u and d quarks form an isospin doublet and the hadronic current, h^{μ}, is conserved. The matrix elements of the conserved charge, $Q = \int d^3x h^0(x)$, between the states $|N\rangle$ and $|n\rangle$, which are isospin eigenstates, can then be determined by the structure of the isospin group, $SU(2)$.[2] In this limit one finds that $f_+(t_{max})$ is unity and $f_-(t_{max})$ is zero. These values are obtained at $t = t_{max}$ since here the final nucleus is stationary with respect to the initial nucleus. Corrections to these values are expected to be on the order of $(M(d)-M(u))/\Lambda_{QCD}$ which is only a few percent. Basically the symmetry is insuring that the wave functions for the initial and final state nuclei are identical so that the matrix element should be large—in this case unity. Such decays are known as superallowed Fermi transitions. The value given by the Particle Data Group,[3] $|V_{ud}| = 0.9744 \pm 0.0010$, is obtained from these decays. Without isospin symmetry the form factors would be much less certain and the error on V_{ud} correspondingly larger.

1.3 Charm

Such is the case for the element V_{cs}. The relevant decay here is the semileptonic decay, $D \to Ke\nu_e$. The same matrix element applies as for beta decay, with the nuclei wave functions being replaced by those of the D and the K.

In this case isospin does not help. Instead one calculates values for the form factors using models; such as treating the mesons as non-relativistic bound states of an effective potential. One can also use QCD sum rules and lattice QCD. All of these methods, however, are only good to about 20%. Thus the value given by the Particle Data Group[3] $V_{cs} = 1.0 \pm 0.2$ is limited by the theoretical uncertainties rather than the experimentally determined decay rate which gives $|V_{cs}|^2|f_+(0)|^2 = 0.49 \pm 0.10$.

With the constraint of unitarity within three generations, V_{cs} is determined much more precisely to be within [0.9735, 0.9751] at a 90% confidence limit.* Thus we can turn the argument around and *measure* the form factors in charm decay using this precise value of V_{cs}. This provides a sensitive test of the models which would be used in determining the very important and least known values, V_{cb} and V_{ub}. The form factors measured in $D^+ \to \bar{K}^0 e^+ \nu_e$ are consistent with those predicted by most models. However, for the form factors in $D^+ \to \bar{K}^{*0} e^+ \nu_e$ (and $D^- \to K^{*0} e^- \bar{\nu}_e$),† which is the subject of this dissertation, it will be shown that the models fail.

1.4 Bottom and HQET

The failure of the models in charm decay makes their reliability in calculating bottom decay form factors suspect. This in turn implies there will be large

---

⋆ Recent measurements of the Z cross section indicate that there are indeed only 3 generations.[4]

† CP-conjugate states are implied throughout unless otherwise noted.

uncertainties in the values of V_{cb} and V_{ub} which ultimately determine the complete Kobayashi-Maskawa matrix. Recently, however, a new set of symmetries for heavy quarks in QCD has been discovered which puts the calculation of the bottom form factors on much firmer ground. The situation is very similar to that for the superallowed transitions of beta decay. When the heavy quark mass is much greater than the typical QCD energy scale ($M_Q \gg \Lambda_{QCD}$) the value of the quark mass decouples from the low energy effective Lagrangian for QCD.[5] In this limit the heavy quark acts as a static source of color. Thus to the extent that the b and c quark masses are large, the B and D pseudoscalar mesons are degenerate and form a $SU(2)$ doublet. Then the form factors in the decay $B \to De\nu_e$ can be calculated in the same manner as for the superallowed nuclear beta decays—again at $t = t_{max}$ in order to satisfy the static requirement. There is, however, a small correction due to the running of the coupling constant, α_S, between the c and b mass scales.

There is an additional symmetry in this heavy quark effective theory (HQET) due to the fact that the spin of the heavy quark also decouples from the Lagrangian.[5] This allows one to relate the form factors in pseudoscalar to vector meson decay to those of pseudoscalar to pseudoscalar decay. The result is that there is one universal form factor, $\xi(t_{max} - t)$ with normalization $\xi(0) = 1$ from symmetry, from which all other form factors can be determined. Corrections to these predictions are expected at the level of Λ_{QCD}/M_Q. For B to D decays this implies corrections of the order of 5% and 20% respectively. Fortunately, it has been shown that the order Λ_{QCD}/M_Q corrections vanish

at the zero recoil limit of $t = t_{max}$; away from this point the corrections do not vanish but have been calculated using QCD sum rules.[6] Since sum rule calculations have a typical uncertainty of about 20%, the uncertainty on the form factors—and therefore on V_{cb}—should be on the order of a few percent.

The situation is not as clean for determining V_{ub} since the u quark is certainly not heavy in the HQET sense. Nevertheless, the form factors in $c \to d$ decay can be related to those of $b \to u$ decay near $t = t_{max}$* using both HQET and isospin symmetries. And the $c \to d$ form factors can be measured using the value of V_{cd} which is known to a few percent, assuming three generation unitarity. Thus it appears likely that a precise determination of V_{cb} and V_{ub} can be made using the symmetries of HQET.

It is important to verify the validity of the HQET symmetries and the extent to which corrections can be estimated. One way to do this is to measure the ratio of form factors governing the decay $B \to D^* l \nu_l$. This dissertation presents (among other things) a measurement of the form factor ratios in $D^+ \to \bar{K}^{*0} e^+ \nu_e$ which can serve as a prototype for measuring the B form factor ratios. Moreover, the D form factors are interesting in themselves since they can be used to test the factorization hypothesis in charm hadronic decays, related to the form factors in $B \to K^* e^+ e^-$ using the symmetries of HQET,[7] and serve as a significant test for any method which attempts to calculate

* The matching condition is that the velocity of the daughter mesons, in the rest frame of the parent meson, must be equal. This means, for example, that the $D \to \rho l \nu_l$ form factors, over the full range of t, only map into the last 17% of the t range of the $B \to \rho l \nu_l$ form factors.

heavy quark semileptonic form factors. The extent to which HQET can be tested by the measured charm form factors will be discussed in the concluding chapter.

2 Kinematics

This chapter defines the kinematics involved in the decay $D^+ \to \bar{K}^{*0} e^+ \nu_e$. First the form factors are defined from the hadronic part of the matrix element. Then the matrix element is evaluated within the helicity formalism. Finally, the differential decay rate is obtained which will be used in determining the form factors from the the data distribution.

2.1 Form Factors

The diagram representing the decay $D^+ \to \bar{K}^{*0} e^+ \nu_e$ is shown in Fig. 2.1. This is known as a spectator decay since the light quark does not participate and is merely a spectator in the interaction. In analogy to the beta decay amplitude discussed in the introduction, the amplitude for this process is

$$\frac{G_F}{\sqrt{2}} V_{cs} \langle K^*, \xi | \bar{s}\gamma^\mu (1+\gamma_5) c | D \rangle \; l_\mu \tag{2.1}$$

where G_F is the Fermi coupling constant, ξ is the polarization vector of the $\bar{K}^{*0}$, and l_μ is the lepton current, $\bar{e}\gamma^\mu(1+\gamma_5)\nu_e$. The weak and strong interactions are not coupled, so the matrix element, or amplitude, factorizes into a

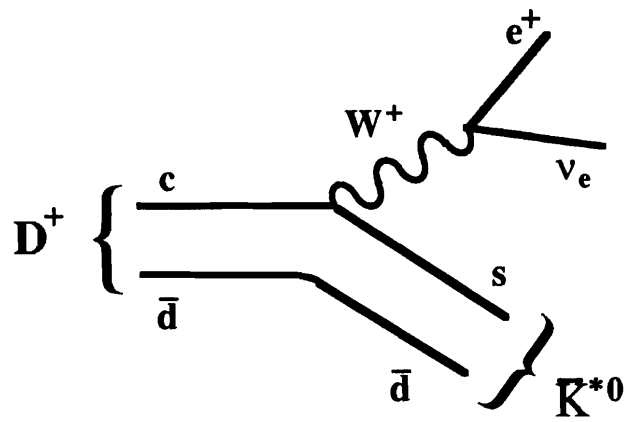

Figure 2.1 Spectator diagram of the semileptonic decay, $D^+ \to \bar{K}^{*0} e^+ \nu_e$.

hadronic current and a leptonic current. As discussed in the introduction, we do not know how to calculate the hadronic part of the matrix element so we parametrize it in terms of form factors. We simply ask how many ways can we form a vector out of the momentum vectors of the D and the $\bar{K}^{*0}$ and the polarization vector of the $\bar{K}^{*0}$. The most general form can be written

$$\begin{aligned}\langle K^*(k), \xi | h_\mu | D(p) \rangle = (M_D + M_{K^*}) A_1(t) \xi^*_\mu - \frac{A_2(t)}{M_D + M_{K^*}} \xi^* \cdot p \, (p_\mu + k_\mu) \\ - \frac{A_3(t)}{M_D + M_{K^*}} \xi^* \cdot p \, q_\mu - i \frac{2V(t)}{M_D + M_{K^*}} \epsilon_{\mu\nu\rho\sigma} \xi^{*\nu} p^\rho k^\sigma\end{aligned} \tag{2.2}$$

where $q_\mu = p_\mu - k_\mu$ is the momentum transfer and $t = q^2$. This form follows from Lorentz invariance and the requirement that each term be linear in the polarization vector. The form factors, $A_i(t)$ and $V(t)$, are functions of the only dynamical scalar, which we take to be t (others are possible but they can be written in terms of t and the masses).

By examining how each side of eq. (2.2) behaves under a parity transformation one can identify the A_i with the axial vector part of h_μ and V with the vector part of h_μ. Since the D is a pseudoscalar, the hadronic matrix element is odd under parity reversal for the axial vector part of h_μ and even for the vector part. The polarization vector is odd under parity reversal, hence the A_i must be axial vector form factors. The totally anti-symmetric tensor in the final term of eq. (2.2) produces a axial vector which is even under parity reversal; hence the identification of V as a vector form factor.

In addition, by examining how each side behaves under parity *and* time reversal one can see that the form factors must be real if the interaction is time-reversal invariant. This transformation interchanges bras and kets and takes operators into their adjoints. Time reversal changes the signs of momenta and spins, but then parity reversal changes the momenta back. Thus,

$$\begin{aligned} \langle K^*(k,s)|\, h_\mu \,|D(p)\rangle^{PT} &= \pm \langle D(p)|\, h_\mu^\dagger \,|K^*(k,-s)\rangle \\ &= \pm \langle K^*(k,-s)|\, h_\mu \,|D(p)\rangle^* \end{aligned} \tag{2.3}$$

where the $+(-)$ sign is for the axial vector (vector) parts of h_μ. Since $\xi(-s) = \xi^*(s)$, it follows that all the form factors are real if the interaction is time reversal invariant. This is why the imaginary factor of i is included in the last term of eq. (2.2).

2.2 Matrix Element

The $\bar{K}^{*0}$ decays strongly to $K^-\pi^+$ and $\bar{K}^0\pi^0$ with a branching ratio of 2/3 and 1/3 respectively. What we observe in the detector is the decay $D^+ \to K^-\pi^+e^+\nu_e$. The $\bar{K}^{*0}$ can be identified, statistically, by a peak in a plot of the mass of the $K^-\pi^+$ system, labeled $M_{K\pi}$, at the mass of the $\bar{K}^{*0}$.

The matrix element for the observed process is the product of the matrix elements for $D^+ \to \bar{K}^{*0}e^+\nu_e$ and for $\bar{K}^{*0} \to K^-\pi^+$ and an additional factor for the propagator of the $\bar{K}^{*0}$. Thus the complete matrix element, $\mathcal{M}$, is

$$\mathcal{M} = \sum_{\lambda} \mathcal{M}_\lambda(\bar{K}^{*0} \to K^-\pi^+)\ \mathcal{P}_\lambda(\bar{K}^{*0})\ \mathcal{M}_\lambda(D^+ \to \bar{K}^{*0}e^+\nu_e) \tag{2.4}$$

where $\mathcal{P}_\lambda(\bar{K}^{*0}) = [M^2_{K^*} - M^2_{K\pi} - iM_{K^*}\Gamma(M_{K\pi})]^{-1}$ represents the $\bar{K}^{*0}$ propagator and the sum is over the helicities of the $\bar{K}^{*0}$, $\lambda \in \{-1, 0, +1\}$.

In calculating this matrix element it is convenient to break the decay, $D^+ \to \bar{K}^{*0}e^+\nu_e$, into a series of two-body decays, $D^+ \to \bar{K}^{*0}W^+$ and $W^+ \to e^+\nu_e$. Then

$$\begin{aligned} \mathcal{M} = \sum_{\lambda_1,\lambda_2} \mathcal{M}(D^+ \to \bar{K}^{*0}_{\lambda_1}\ W^+_{\lambda_2}) \times \\ \mathcal{M}_{\lambda_1}(\bar{K}^{*0} \to K^-\pi^+)\ \mathcal{P}_{\lambda_1}(\bar{K}^{*0})\ \mathcal{M}_{\lambda_2}(W^+ \to e^+\nu_e)\ \mathcal{P}_{\lambda_2}(W^+). \end{aligned} \tag{2.5}$$

The virtual W^+ would in general have a fourth time-like helicity component. However, this goes to zero in the limit that the electron is massless, which is a very good approximation at the energy scales for this decay. It will be shown

below that

$$\mathcal{M}(D^+ \to \bar{K}^{*0}_{\lambda_1}\, W^+_{\lambda_2}) = C(t) H_\lambda \delta_{\lambda_1,\lambda_2} \tag{2.6}$$

where the H_λ are called helicity amplitudes and are functions of the form factors.

The evaluation of $\mathcal{M}_{\lambda_2}(W^+ \to e^+\nu_e)$ is particularly simple since, if the electron is massless, we know that it must have positive helicity and the neutrino must have negative helicity. This follows from the $V - A$ form of the leptonic current. Since the helicities of the parent and daughter particles are known, the W^+ decay matrix element can therefore be written in terms of the appropriate Wigner d-function.[8] The result is,

$$\mathcal{M}_\lambda(W^+ \to e^+\nu_e) = \begin{cases} (1+\cos\theta_e), & \lambda = 1 \\ (1-\cos\theta_e)\, e^{i2\chi}, & \lambda = -1 \\ -\sqrt{2}\sin\theta_e\, e^{i\chi}, & \lambda = 0 \end{cases} \tag{2.7}$$

where any proportionality constants are to be included in $C(t)$ in eq. (2.6). Here, θ_e is the angle between the e^+ and the direction opposite that of the $\bar{K}^{*0}$ in the rest frame of the W^+. The axial angle χ is the angle between the decay planes of the $e^+\nu_e$ and the $K^-\pi^+$ systems in the D^+ rest frame.

A similar analysis can be applied to the matrix element for the $\bar{K}^{*0}$ decay yielding:

$$\mathcal{M}_\lambda(\bar{K}^{*0} \to K^-\pi^+) = B(M_{K\pi}) \begin{cases} \pm\sin\theta_v, & \lambda = \pm 1 \\ \sqrt{2}\cos\theta_v, & \lambda = 0 \end{cases} \tag{2.8}$$

where $B(M_{K\pi})$ is a proportionality constant. Here, θ_v is the angle between the K^- and the direction opposite that of the W^+ measured in the rest frame

of the $\bar{K}^{*0}$. The propagator of the W^+ just generates the Fermi constant. Putting everything together and taking the absolute square yields:

$$
\begin{aligned}
|\mathcal{M}|^2 = & \frac{G_F^2 |V_{cs}|^2 |C(t)|^2 \, |B(M_{K\pi})|^2}{(M_{K\pi}^2 - M_{K^*}^2)^2 + M_{K^*}^2 \Gamma^2(M_{K\pi})} \times \\
& \Big\{ \Big[(1+\cos\theta_e)^2 |H_+(t)|^2 + (1-\cos\theta_e)^2 |H_-(t)|^2 \Big] \sin^2\theta_v \\
& + 4\sin^2\theta_e \cos^2\theta_v |H_0(t)|^2 - 2\sin^2\theta_e \sin^2\theta_v \; \Re(e^{i2\chi} H_+^* H_-) \\
& - 4\sin\theta_e \, (1+\cos\theta_e) \sin\theta_v \cos\theta_v \; \Re(e^{i\chi} H_+^* H_0) \\
& + 4\sin\theta_e \, (1-\cos\theta_e) \sin\theta_v \cos\theta_v \Re(e^{i\chi} H_-^* H_0) \Big\}.
\end{aligned}
\tag{2.9}
$$

This structure follows simply from angular momentum conservation and the $V-A$ character of the W^+ decay; the interesting physics issues are contained in the helicity amplitudes.

2.3 Helicity Amplitudes

The helicity amplitudes, as given in eq. (2.6), are defined in terms of the matrix element, $\mathcal{M}(D^+ \to \bar{K}^{*0}_{\lambda_1} \, W^+_{\lambda_2})$. Since the hadronic and leptonic parts of this element factorize, we can write

$$
\mathcal{M}(D^+ \to \bar{K}^{*0}_{\lambda_1} \, W^+_{\lambda_2}) = \langle W^+(q), \epsilon_{\lambda_2} | L^\mu \, |0\rangle \; \langle K^*(k), \xi_{\lambda_1} | \, h_\mu \, | D(p)\rangle \tag{2.10}
$$

where L^μ is the effective leptonic current and ϵ_{λ_2} is the polarization vector of the W^+. The hadronic amplitude was defined in terms of the form factors in eq. (2.2) using the requirements of Lorentz invariance and linearity in the

polarization. The same requirements applied to the leptonic part of eq. (2.10) yields the most general form,

$$\langle W^+(q), \epsilon_{\lambda_2} | L^\mu | 0 \rangle = C(t)\ \epsilon^{*\mu}_{\lambda_2} \tag{2.11}$$

where $C(t)$ is a possible proportionality constant which can depend on t. We have used the fact that for the spatial polarization vectors, $\epsilon_\lambda \cdot q = 0$; this is not true for the time-like polarization vector, but this term is negligible since the electron may be considered massless.

We are now in position to evaluate the matrix element in eq. (2.10). To be definite, it is useful to explicitly write out the polarization vectors. We chose a coordinate system, in the rest frame of the D^+, in which the positive z-axis is in the direction of the $\bar{K}^{*0}$. Then,

$$\begin{aligned} \xi^\mu_\pm &= \frac{1}{\sqrt{2}}(1, \pm i, 0, 0) \qquad & \xi^\mu_0 &= \frac{1}{M_{K\pi}}(0, 0, E_{K^*}, K) \\ \epsilon^\mu_\pm &= \frac{1}{\sqrt{2}}(1, \mp i, 0, 0) \qquad & \epsilon^\mu_0 &= \frac{1}{\sqrt{t}}(0, 0, -E_W, K) \end{aligned} \tag{2.12}$$

where K is the magnitude of the spatial momentum of the $\bar{K}^{*0}$ (or the W^+) in the rest frame of the D^+. In eq. (2.12) the fourth component of each vector is the time-like component. These helicity vectors satisfy the orthonormality relation (for both ξ and ϵ), $\xi^*_{\lambda_1} \cdot \xi_{\lambda_2} = \delta_{\lambda_1,\lambda_2}$. Putting everything together one finds that $\mathcal{M}(D^+ \to \bar{K}^{*0}_{\lambda_1}\ W^+_{\lambda_2}) = C(t) H_\lambda \delta_{\lambda_1,\lambda_2}$ where the H_λ, determined by

the relation $H_\lambda = \epsilon_\lambda^{*\mu} \langle K^*(k), \xi_\lambda | \, h_\mu \, | D(p) \rangle$, are:

$$
\begin{aligned}
H_\pm(t) &= (M_D + M_{K\pi})A_1(t) \mp 2\frac{M_D K}{M_D + M_{K\pi}} V(t) \\
H_0(t) &= \frac{1}{2M_{K\pi}\sqrt{t}} \Bigg[(M_D^2 - M_{K\pi}^2 - t)(M_D + M_{K\pi})A_1(t) \\
&\qquad - 4\frac{M_D^2 K^2}{M_D + M_{K\pi}} A_2(t) \Bigg].
\end{aligned}
\tag{2.13}
$$

The A_3 form factor drops out since it is proportional to q_μ and $\epsilon_\lambda^* \cdot q = 0$. This occurs because we have dropped the time-like helicity component of the virtual W^+. The A_3 form factor couples to this helicity component which is non-negligible only if the charged lepton is fairly massive, such as with a τ^+.

2.3.1 Partial Waves

At this point it is useful to change from a helicity point of view to a partial wave point of view and ask which form factors are present in which partial waves. This decay is from a pseudoscalar to two vector particles which can occur via. S, P, and D waves. For two-body decays the amplitude must be proportional to K^l where K is the magnitude of the momentum of the daughter particles in the parent rest frame, and l is the orbital angular momentum. Since we have shown that the amplitude for $D^+ \to \bar{K}^{*0} W^+$ is proportional to the helicity amplitudes, we can simply look at eq. (2.13) and identify $A_1(t)$ with the S-wave, $V(t)$ with the P-wave, and $A_2(t)$ with the D-wave.

2.4 Decay Rate

The differential decay rate is given by the formula

$$d\Gamma = \frac{(2\pi)^4}{2M_D} \, |\mathcal{M}|^2 \, d\Phi_4(D; K, \pi, e, \nu) \tag{2.14}$$

where $d\Phi_4(D; K, \pi, e, \nu)$ is the four-body phase space element. In evaluating this it is useful to break the decay into a series of two-body decays using the recursive relation

$$d\Phi_4(S; p_1, p_2, p_3, p_4) = d\Phi_3(D, q, p_3, p_4) \, d\Phi_2(q; p_1, p_2) \, (2\pi)^3 dq^2 \tag{2.15}$$

where $q^\mu = p_1^\mu + p_2^\mu$. The two-body phase space element is given by

$$d\Phi_2(S; p_1, p_2) = \frac{1}{4(2\pi)^6} \frac{P}{\sqrt{S}} \, d\phi \, d\cos\theta \tag{2.16}$$

where S is the squared mass of the parent particle and P is the magnitude of the spatial momentum of the daughter particles in the rest frame of the parent. The angles are the usual polar angles and are measured in the rest frame of the parent. Using the angles defined above for the matrix element, one finds that

$$\frac{d\Gamma}{dM_{K\pi}^2 \, dt \, d\cos\theta_v d\cos\theta_e d\chi} = \frac{K P_k}{2(4\pi)^6 M_D^2 M_{K\pi}} \, |\mathcal{M}|^2 \tag{2.17}$$

where P_k is the magnitude of the momentum of the K^- in the rest frame of the $\bar{K}^{*0}$.

All that remains is to evaluate the two proportionality functions, $C(t)$ and $B(M_{K\pi})$. $B(M_{K\pi})$ can be expressed in terms of the total decay rate for the $\bar{K}^{*0}$ since

$$\Gamma(M_{K\pi}) = \int \frac{(2\pi)^4}{2M_{K^*}} \frac{1}{3} \left| \sum_{\lambda} \mathcal{M}_{\lambda}(\bar{K}^{*0} \to K^-\pi^+) \right|^2 d\Phi_2(K^*; K, \pi) \tag{2.18}$$

and the matrix element, $\mathcal{M}_{\lambda}(\bar{K}^{*0} \to K^-\pi^+)$, was given in terms of $B(M_{K\pi})$ in eq. (2.8). One finds that

$$|B(M_{K\pi})|^2 = \frac{24\pi M_{K\pi} M_{K^*} \Gamma(M_{K\pi})}{P_k}. \tag{2.19}$$

Here, the total rate, or width, of the $\bar{K}^{*0}$ is written as a function of the mass, $M_{K\pi}$. Since the $\bar{K}^{*0}$ is a vector particle and decays into two scalar particles, the daughter particles must be in a relative P-wave. The amplitude for the decay must be therefore be proportional to P_k. In addition, the phase space element for this decay contributes another factor of $P_k/M_{K\pi}$. Thus, the total decay rate is proportional to $P_k^3/M_{K\pi}$ and P_k is a function of $M_{K\pi}$. For this reason we parametrize $\Gamma(M_{K\pi})$ in the following way:

$$\Gamma(M_{K\pi}) = \Gamma(\bar{K}^{*0}) \frac{P_k^3(M_{K\pi}) M_{K^*}}{P_k^3(M_{K^*}) M_{K\pi}} \tag{2.20}$$

where $\Gamma(\bar{K}^{*0})$ is the measured, average width. The effect of this mass dependence is to skew the otherwise symmetric Breit-Wigner distribution towards higher mass values.

The $C(t)$ function can be calculated by expanding out the original matrix element, eq. (2.1). One finds simply that

$$|C(t)|^2 = \frac{t}{2}. \tag{2.21}$$

With this, all the unknown factors have been determined and we can present the full differential decay rate in all of its glory:

$$\begin{aligned} d\Gamma = G_F^2|V_{cs}|^2 \frac{3}{2(4\pi)^5 M_D^2} \frac{M_{K^*}\,\Gamma(M_{K\pi})}{(M_{K\pi}^2 - M_{K^*}^2)^2 + M_{K^*}^2\Gamma^2(M_{K\pi})} Kt \times \\ \Big\{ \Big[(1+\cos\theta_e)^2|H_+(t)|^2 + (1-\cos\theta_e)^2|H_-(t)|^2\Big] \sin^2\theta_v \\ +4\sin^2\theta_e\cos^2\theta_v|H_0(t)|^2 - 2\sin^2\theta_e\sin^2\theta_v\ \Re(e^{i2\chi}H_+^*H_-) \\ -4\sin\theta_e\,(1+\cos\theta_e)\sin\theta_v\cos\theta_v\ \Re(e^{i\chi}H_+^*H_0) \\ +4\sin\theta_e\,(1-\cos\theta_e)\sin\theta_v\cos\theta_v\Re(e^{i\chi}H_-^*H_0)\Big\} \\ \times\ dM_{K\pi}^2\,dt\,d\cos\theta_v d\cos\theta_e d\chi. \end{aligned} \tag{2.22}$$

The only free parameters are the form factors which are buried in the helicity amplitudes (eq. (2.13)). By comparing the distribution of the experimental data to this differential decay rate we can estimate the value of the form factors.

In calculating the total decay rate the integrals over the angles are simple and yield

$$\begin{aligned} \Gamma = \int \frac{G_F^2|V_{cs}|^2}{96\pi^4 M_D^2} \frac{M_{K^*}\,\Gamma(M_{K\pi})}{(M_{K\pi}^2 - M_{K^*}^2)^2 + M_{K^*}^2\Gamma^2(M_{K\pi})} \times \\ Kt\ \left(|H_+(t)|^2 + |H_-(t)|^2 + |H_0(t)|^2\right) dM_{K\pi}^2\ dt \end{aligned} \tag{2.23}$$

where the integral over t extends from the squared mass of the electron (effectively zero) to $t_{max} = (M_D - M_{K\pi})^2$. In the limit that the $\bar{K}^{*0}$ width is

negligible then the Breit-Wigner term becomes

$$\frac{M_{K^*}\,\Gamma(M_{K\pi})}{(M_{K\pi}^2 - M_{K^*}^2)^2 + M_{K^*}^2\Gamma^2(M_{K\pi})} \Longrightarrow \pi\ \delta(M_{K\pi}^2 - M_{K^*}^2). \tag{2.24}$$

This is only good to about 10% since the $\bar{K}^{*0}$ is not a narrow resonance.

3 The Experiment and Data

Fermilab photoproduction experiment E691 has been described in detail previously.[9] This chapter describes those aspects of the experiment pertinent for observing the decay $D^+ \rightarrow \bar{K}^{*0} e^+ \nu_e$ and then summarizes the cuts used in the extraction of the signal.

3.1 Brief Description of Experiment

The experiment was based at Fermilab and took data from April through August of 1985. It was a fixed-target experiment with a tagged photon beam at an average energy of $120\,GeV$ incident on a 5 cm long piece of Beryllium. The spectrometer, illustrated in Fig. 3.1, consists of silicon microstrip detectors (SMD), drift chambers around two magnets, Čerenkov counters, calorimeters, and a muon detector at the very back. The SMDs (to be described in more detail below) are all within 25 cm downstream of the target and with 16 μm hit resolution, they provide precise track coordinates for charged particles close to the interaction region. This is crucial in being able to isolate decay vertices from production vertices.

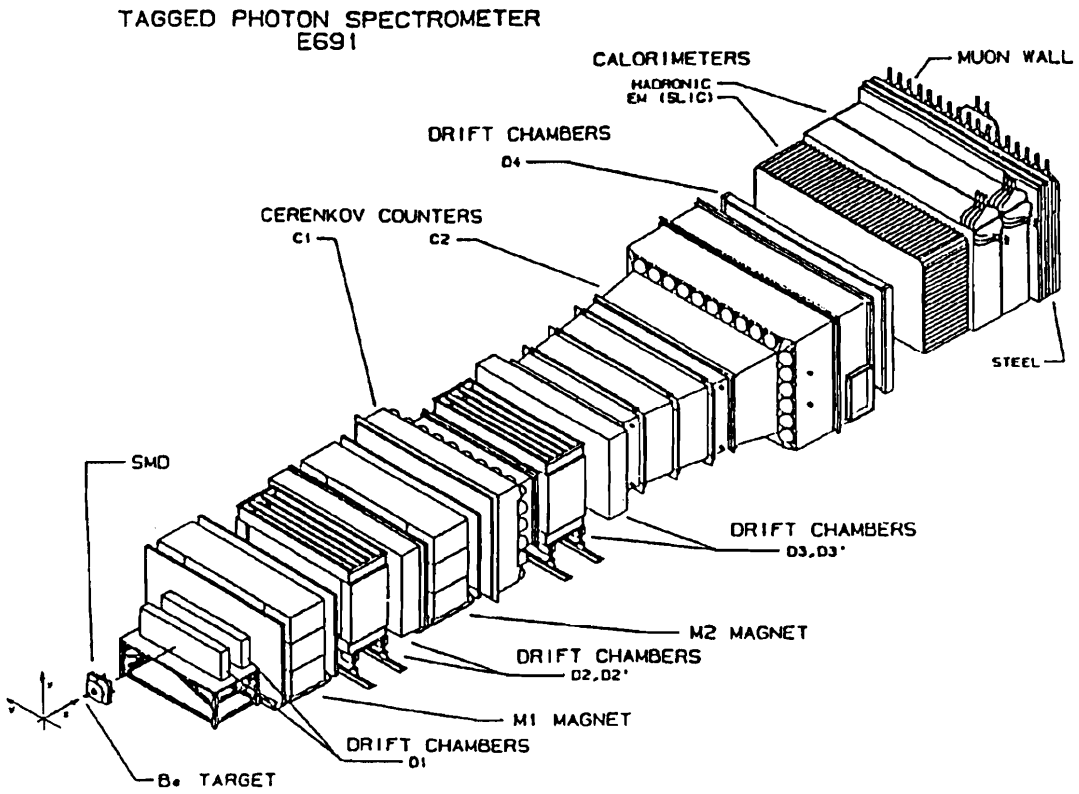

Figure 3.1 Illustration of the E691 spectrometer.

The drift chambers, being located upstream and downstream of the magnets, allow one to measure the momentum and sign of the charge of particles from the deviation that occurred within the regions of the magnets. The track measurements made by the drift chanbers are also matched with those observed in the SMDs. Together with the momentum estimate, the Čerenkov detectors provide information about the mass or identification of the charged particle. Electrons and muons are further identified using the electromagnetic calorimeter and the muon detector, respectively. The hadronic calorimeter allows identification of neutral hadrons. In summary, the SMDs provide geometrical information while the rest of the spectrometer provides momentum, energy, and charge information.

3.2 The SMDs

The key feature of the SMDs which makes them so important in isolating charm events is their ability to measure the position of a charged track very precisely. To illustrate how this is accomplished, consider the schematic of a typical E691 SMD shown in Fig. 3.2. It consists of a 290 μm thick layer of slightly n-doped (an excess of negatively charged carriers) silicon with a thin layer of n-type material (Arsenic) on the bottom and strips of p-type (an excess of positively charged carriers) material (Boron) on the top. Aluminum is placed in contact with the Arsenic below and with the Boron above, in order to provide for a good ohmic connection to outside electronics.

The interface between the p-strips and the bulk n-type silicon forms a

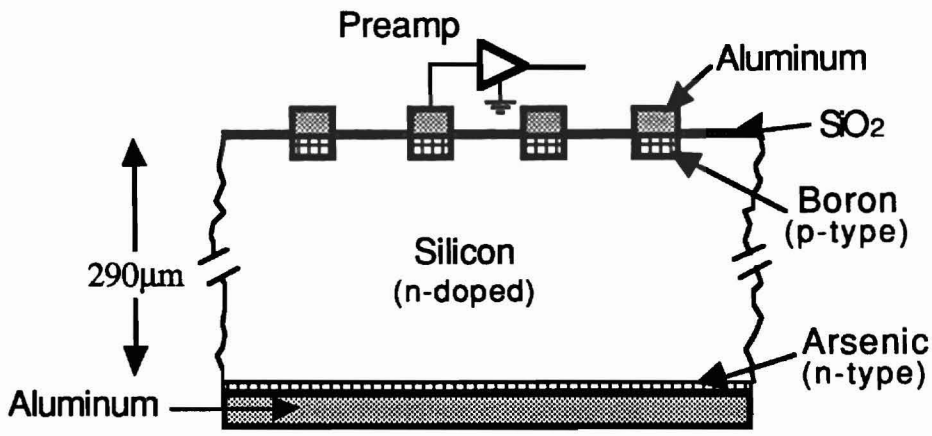

Figure 3.2 Cross section of an E691 SMD.

p-n junction. The mobile positive and negative charges of the p- and n-type regions, respectively, diffuse and cancel each other when they meet near the interface. This ionizes the fixed lattice charges near the interface leaving an excess of negative charge on the p-side and positive charge on the n-side. An electrical potential thus forms replacing the kinetic energy of the mobile carriers. This small region of depleted charge carriers is called the depletion region.

If a charged particle traverses the silicon, electron-hole pairs are produced. Typically there are about 21000 pairs for 290 μm of silicon. This is the signal which we would like to detect. However, these electrons and holes will recombine as they diffuse through the silicon. If the electron-hole pairs were generated within the depletion region, though, then the potential in this region would seperate the electrons and holes, driving the electrons towards the p-side and the holes towards the n-side. The depeletion region is enlarged by

applying a reverse bias voltage, V_b, between the aluminum of the two sides of the silicon. Usually between 40 V and 100 V is required to fully deplete the 290 μm of bulk silicon depending on its resistivity ($V_{depleted} \approx 1/\sqrt{\rho}$). In this way the cloud of holes generated by a charged particle is forced towards the p-type strips so that the position of the particle is identified by which strip produces an excess current.

The strips are about 30 μm wide and have a pitch (separation) of 50 μm. Thus, the SMDs should be able to locate the transverse position of the charged track to within 50 μm, or with a resolution characterized by a standard deviation of 50 $\mu m/\sqrt{12} \approx 14$ μm. In practice E691 SMDs had a resolution of about 16 μm.

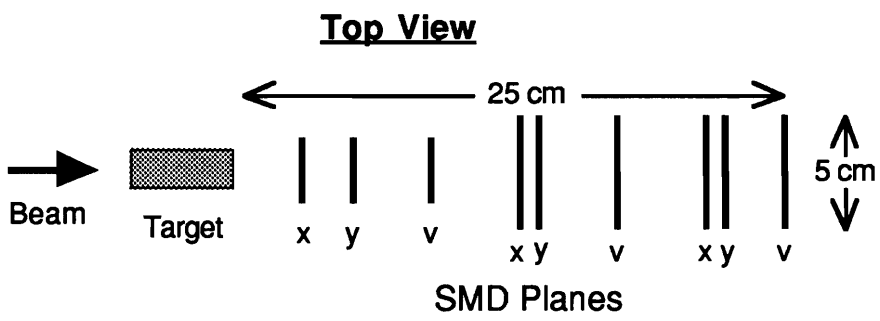

Figure 3.3 Top view of SMD telescope showing the relative location and orientation (x, y, or v) of each SMD plane.

Each SMD, however, can only measure the transverse position of a charged track in one dimension. Ultimately, one would like to know the position *and* direction of a track. To this end a telescope was formed using 9 planes of SMDs measuring positions in x, y, and v (rotated about 20° from x) views as

shown in Fig. 3.3. The angular acceptance of this telescope was ± 100 *mrad* and the per-plane efficiency was about 95%. In order to not waste the precision of each SMD, the mounting of each was carefully controlled and monitored to maintain a relative rotational offset about the beam axis of less than 0.8 *mrad*.

To measure the small charge of 21000 electrons, the p-type strips were wire bonded to strips on a printed circuit fan-out board which were then connected to current-sensitive preamplifiers. This method is in contrast to the one presented in the Appendix, where the preamplifiers are integrated into a small chip which can be bonded directly to the strips. The preamplifiers had a gain of about 200 and an equivalent noise charge of about 1600 e^- to be compared to the expected signal size of about 21000 e^-.* The signals from the preamplifiers were then sent to a set of discriminators which determined which strip had registered a "hit". The criteria for a hit was whether the measured signal was larger than 0.4 times that which was expected from a minimum ionizing particle. This threshold was reasonably efficient for large angle tracks, but high enough to keep noise rates down to a manageable level (typically 1 noise hit per plane per trigger). A more detailed description and analysis of a silicon detector amplification system (to be used in the CLEO-II experiment) is presented in the Appendix.

---

* For a definition of equivalent noise charge please see the Appendix.

3.3 Extraction of Signal

The key element in extracting the decay $D^+ \rightarrow \bar{K}^{*0} e^+ \nu_e$, $\bar{K}^{*0} \rightarrow K^- \pi^+$ from the general data set is the ability to require a well-isolated decay vertex of the proper topology using the high-precision silicon microstrip detector. Three-particle $K\pi e$ combinations are selected using only tracks which are well defined in the microstrip and drift-chamber systems. Electrons are identified by the pattern of energy deposition in the electromagnetic calorimeter. We identify tracks as electrons with an electron probability corresponding to a typical efficiency for electrons and pions of 61% and 0.3%, respectively. A cleaner electron sample with electron and pion efficiencies of 44% and 0.13% is also used. Furthermore, only electrons with a lab energy of $12\,GeV$ or more are kept and those consistent with beam pair conversions are eliminated. To suppress electrons from pair conversions from π^0 decays we also eliminate electrons for which the other member of the pair is seen in the spectrometer.

Tracks are identified as K^- or π^+ by requiring the Čerenkov probability for both the kaon and pion to be greater than 0.4. The pion is required to pass through at least the first magnet while the electron and kaon must pass through both magnets. Only neutral $K\pi$ combinations are used. Combinations in which the kaon and the electron have opposite (same) charges are labeled right- (wrong-) sign combinations.

The selected $K\pi e$ tracks are required to emanate from a decay vertex separated from the primary event vertex by at least 10 standard deviations plus an additional distance corresponding to $0.2\,ps$ proper decay time. In

addition, the χ^2 per degree of freedom of the decay vertex is required to be less than 3.5. Another vertex cut requires the D line of flight to point back to the primary vertex. The maximum neutrino momentum, transverse to that of the $K\pi e$ momentum, is given by equating the mass of the $K\pi e\nu$ system (assuming the neutrino is transverse to the $K\pi e$ direction) to the mass of the D. With this we can obtain the maximum displacement of the $K\pi e$ line of flight at the primary vertex. We require that the $K\pi e$ line of flight point back to the primary vertex to within this distance plus 2.5 times the resolution. Two other vertex cuts require that all of the decay tracks pass closer to the decay vertex than to any other possible vertex, and that no other tracks pass within 65 μm of the decay vertex.

To remove some known backgrounds, we identified and removed seven events in which the electron could have been a misidentified pion from the dominant $D^+ \rightarrow K^-\pi^+\pi^+$ decay. Possible $D^{*+} \rightarrow \pi^+ D^0$, $D^0 \rightarrow K^- e^+ \nu_e$ events in which the bachelor pion appears to come from the decay vertex due to the low Q value of the D^{*+} decay are also removed. And finally, a few events with $K\pi e$ masses greater than the D^+ mass or with decay times longer than 4 D^+ lifetimes are also eliminated.

The $K\pi e$ mass spectrum of the remaining 318 right-sign and 66 wrong-sign combinations is shown in Fig. 3.4. A majority of the events lie below the D^+ mass due to the missing neutrino. Our Monte Carlo simulation is in good agreement with the shape of this spectrum, with the wrong-sign background subtracted. The decay-time distribution of these events is also in good

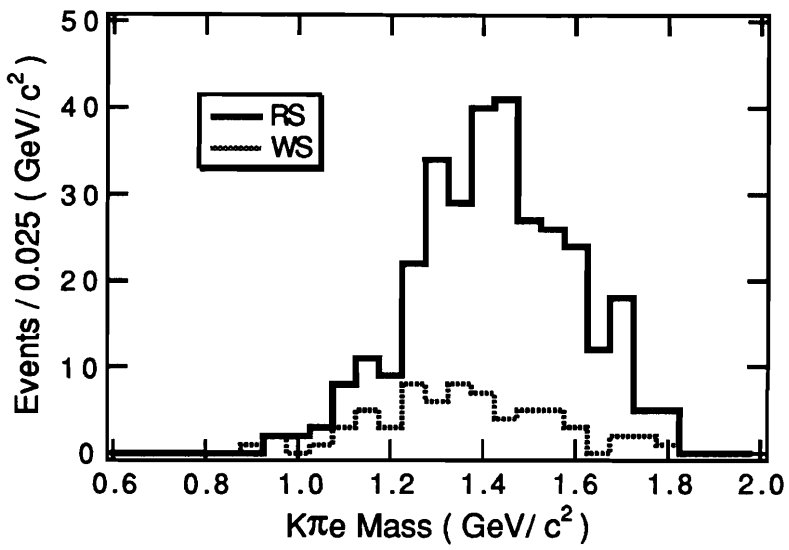

Figure 3.4 Spectrum of $K^{\mp}\pi^{\pm}e^{\pm}$ masses for the standard cuts. The wrong-sign ($K^{\mp}\pi^{\pm}e^{\mp}$) distribution is superimposed (dashed line).

agreement with that of the Monte Carlo simulation.

3.3.1 Backgrounds

Conversion electrons and misidentified pions are the major sources of background; both of which lead to charge-symmetric backgrounds. The first, which is the largest background, is manifestly charge symmetric; the charge symmetry of the second was checked using backgrounds for hadronic decay modes. The most serious non-charge-symmetric background is due to the decay $D^+ \to K^-\pi^+\pi^+\pi^0$ where one of the charged pions is misidentified as an electron and the π^0 is undetected. To estimate the magnitude of this feed-through a Monte Carlo simulation of this decay was performed and the events were passed through the same analysis code as for the signal. For the K^* and nonresonant cases, it was found that a possible 1 ± 1 and 3 ± 3 events could

be present in the data sample. Smaller backgrounds are expected for other D^+ decay modes. Thus the magnitude and shape of the background should be represented well by the wrong-sign data.

An additional check that wrong-sign events are a good measure of the background is obtained by studying the data with a variety of tighter cuts. By requiring tighter electron and kaon identification probabilities and a χ^2 per degree of freedom of 1.75 for the decay vertex, a sample of 169 right-sign and 14 wrong-sign events is obtained. The background has been reduced by a factor of 5, while the signal is smaller by the ratio of efficiencies, within the statistical errors.

3.3.2 Magnitude of Background

The $K\pi$ mass spectrum for the right- and wrong- sign events, using the standard cuts, is shown in Fig. 3.5. Clearly the K^* resonance is dominant. To estimate the amount of non-resonant $K\pi$ in the data a Monte Carlo simulation of $D^+ \rightarrow K^-\pi^+e^+\nu_e$ was performed using a phase-space mass distribution with the $K\pi$ system in an S-wave state. The resulting distribution of events was used to fit the $K\pi$ mass spectrum of the real data, together with the Breit-Wigner of the K^*, and the wrong sign data events. The number of observed K^*'s from the fit is 227 ± 20 and the number of non-resonant events is 25 ± 18. To get the final branching ratios the fit will be performed later using a more nearly correct shape for the K^*. But for now this gives us a good estimate of the non-resonant component of the data, which together with the charge-symmetric background, make up the total background for the K^* signal.

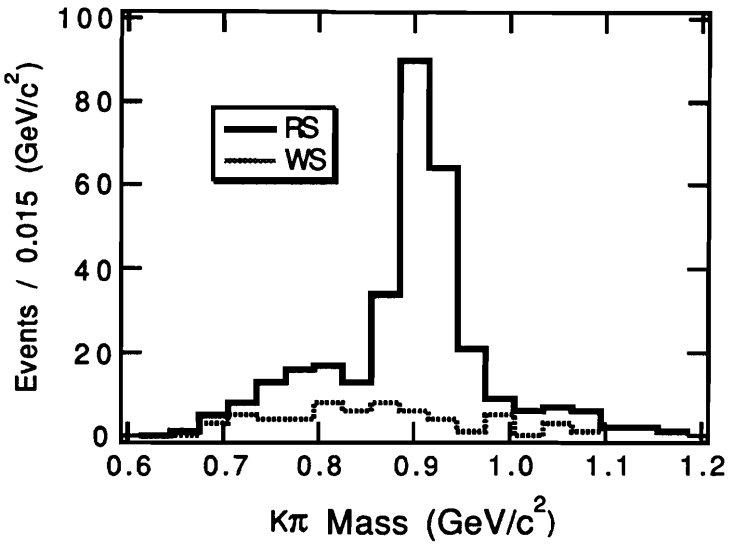

Figure 3.5 The $K\pi$ mass spectrum with right-sign (solid) and wrong-sign (dashed) combinations.

3.4 Determining Kinematic Variables

In order to measure the form factors, the kinematic variables, $\cos\theta_v$, $\cos\theta_e$, χ, and t defined in Chpt. 2 need to be determined for each event. This is complicated because of the missing neutrino information. The direction of the D is taken to be the line between the primary and decay vertices. The component of the neutrino momentum transverse to this direction can be determined by equating it to the transverse momentum of the $K\pi e$ system. The longitudinal neutrino momentum, however, is unknown since the total momentum of the D is not known. However, energy and momentum conservation, using the known mass of the D^+, determine the longitudinal momentum to within a quadratic ambiguity. Since the momentum spectrum of the D is expected to be dominated by lower rather than higher momentum, the solution with the lower D

momentum is taken. A Monte Carlo simulation demonstrated that this was the correct solution most often; however, making the wrong choice for some fraction of the events results in smearing. The chosen values of the kinematic variables are not necessarily the true or correct values. Smearing from this two-fold ambiguity adds to the smearing from limited resolution and poses an important problem for measuring the form factors. This will be addressed in more detail in Chapter 5. Nevertheless, by making a choice for the longitudinal momentum of the neutrino all kinematic variables are uniquely determined.

4 Information and Minimum Variance

The form factor ratios are buried in the complicated differential decay distribution and it is not at all obvious how their values can be estimated from the data. It would be useful to know where, in the 4–dimensional space of observable variables, the information on the form factors is large and which variables are most important for their estimation. This chapter analyzes the properties of the decay distribution relative to the form factors and obtains expected minimum variances using the Cramér-rao theorem.

4.1 Form Factor Parametrization

Since there are only about 200 events in the data sample it is not possible to obtain the t-dependence of the form factors directly from the data. Instead, we choose to parametrize the t-dependence of the form factors using a pole-dominant form such that $A_i(t) = A_i(0)/(1 - t/M_A^2)$ and $V(t) = V(0)/(1 - t/M_V^2)$, where $M_A = 2.5\ GeV$ and $M_V = 2.1\ GeV$ represent the masses of the lowest mass $c\bar{s}$ states with $J^P = 1^+$ and 1^- respectively. This form is consistent with those used by the authors of the various models which predict

form factor values at a particular value of t and then extrapolate to other values of t using a parametrization. Nevertheless, this is an assumption and the dependence of the results on this assumption will be discussed later.

Having made this parametrization, the differential decay rate now depends only on the three constants $A_1(0)$, $A_2(0)$, and $V(0)$. Since the $A_1(0)$ parameter is present in all three helicity amplitudes it serves as a normalization constant and may be brought out front in the differential decay rate formula. Thus the angular and t-dependence of the data will determine the form factor ratios, $R_2 = A_2(0)/A_1(0)$ and $R_V = V(0)/A_1(0)$, while the measured branching ratio will determine the normalization, $A_1(0)$.

4.2 Information on Helicity Amplitudes

It is easy to see how one can extract the relative amplitudes of the helicities as compared to those of the form factor ratios. The decay distribution, after integrating over χ and neglecting constants and the dependence on $M_{K\pi}$, is

$$\begin{aligned}\Gamma(t, \cos\theta_v, \cos\theta_e) = \Big[\Gamma_+(t)(1+\cos\theta_e)^2 + \Gamma_-(t)(1-\cos\theta_e)^2\Big]\sin^2\theta_v \\ + 4\sin^2\theta_e\cos^2\theta_v\Gamma_0(t)\end{aligned} \tag{4.1}$$

where $\Gamma_i(t) = KtH_i^2(t)$ are the helicity rates and from eq. (2.13),

$$H_{\pm}(t) = \frac{(M_D + M_{K\pi})}{(1 - t/M_A^2)} \mp 2\frac{M_D K}{M_D + M_{K\pi}} \frac{R_V}{(1 - t/M_V^2)}$$

$$H_0(t) = \frac{1}{2M_{K\pi}\sqrt{t}} \left[(M_D^2 - M_{K\pi}^2 - t)\frac{(M_D + M_{K\pi})}{(1 - t/M_A^2)} - 4\frac{M_D^2 K^2}{M_D + M_{K\pi}} \frac{R_2}{(1 - t/M_V^2)} \right]. \tag{4.2}$$

From this distribution one can readily tell that the longitudinal term, Γ_0, will be dominant near $\cos\theta_v = \pm 1$ whereas the transverse terms, $\Gamma_{\pm}$, will be dominant near $\cos\theta_v = 0$. Furthermore, Γ_- and Γ_+ are separated in $\cos\theta_e$. This is illustrated in Fig. 4.1. Hence, merely projecting the data onto $\cos\theta_v$ and $\cos\theta_e$ would allow one to determine the relative amplitudes of the helicity rates. Since it is the form factor ratios which determine the magnitude of the relative helicity amplitudes, it must also be possible to separate and determine the form factor ratios themselves from this two-dimensional projection of the data.

The different helicity amplitudes have different shapes over t. This yields additional information about the relative magnitudes of the helicities and therefore the form factor ratios. The functions $\Gamma_i(t)$ are plotted in Fig. 4.2 for $R_2 = 0.7$ and $R_V = 2.0$ which are typical of model predictions.* Near $t = 0$ the longitudinal amplitude is clearly dominant. The transverse terms grow as t increases till at about $t = 0.7\,GeV^2$ where Γ_- surpasses Γ_0. Near $t = t_{max}$ all three amplitudes are equal but diminishing towards zero. Here the $\bar{K}^{*0}$ has

* The values $R_2 = 0.7$ and $R_V = 2.0$, will be used throughout this chapter unless explicitly stated otherwise.

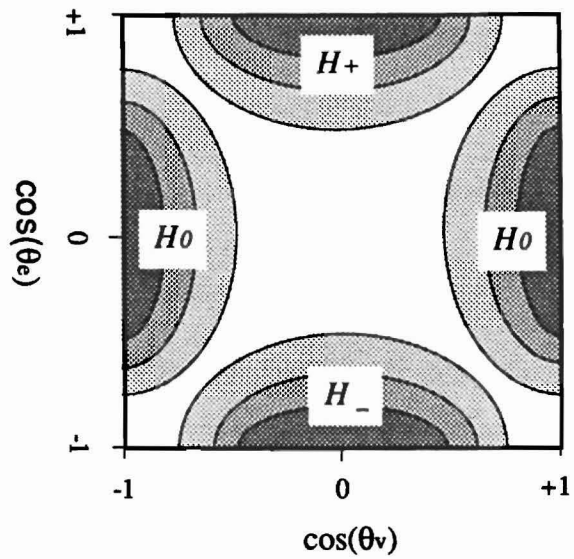

Figure 4.1 Illustration of where the different helicity rates are dominant in $\cos\theta_e$, $\cos\theta_v$ space.

zero recoil momentum so that helicities are ill-defined and whence all three must be equal in magnitude.

The top insert of Fig. 4.2 shows the derivative of the helicity rates with respect to the form factor ratios. The longitudinal rate is a function only of R_2 whereas the transverse rates are functions only of R_V. This graph indicates that the total rate near $t = 0$ is a strong function of R_2 and that an excess of events at this point would yield a lower value for R_2. The transverse terms are not so well separated in t; most of the information on the relative magnitudes of these rates—and therefore on R_V—comes from the $\cos\theta_e$ dependence.

The above analysis yields a qualitative description of information on the helicity rates available within the three-dimensional decay distribution. This translates into information about the form factor ratios since the helicity rates

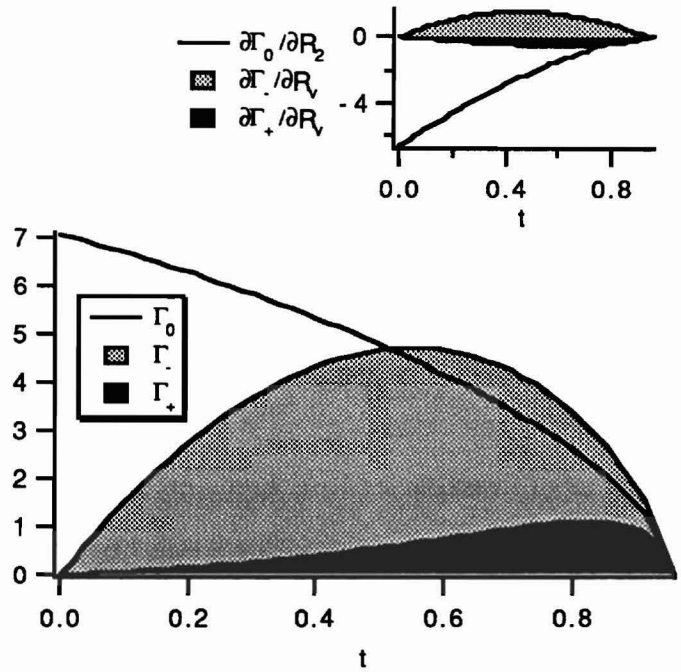

Figure 4.2 Helicity rates as a function of t for $R_2 = 0.7$ and $R_V = 2.0$. Top insert shows the derivative of the helicity rates with respect to R_2 and R_V as a function of t.

are functions of them. Nevertheless, it would be useful to be able to quantitatively analyze the information available on the form factor ratios themselves. Then it would be possible to estimate the expected errors on our measurement of R_2 and R_V as well as how much information is gained by including χ in the fit.

4.3 Information on Form Factor Ratios

The Cramér-rao theorem relates the minimum variance of an estimate of a parameter to the information content of this parameter contained in its probability distribution.[10] For a parameter, μ, with (normalized) probability

distribution, $P(x;\mu)$, over the observable variable, x, the information per event is:

$$I(\mu) = \int \left(\frac{\partial P}{\partial \mu}\right)^2 \frac{1}{P}\, dx. \tag{4.3}$$

The Cramér-rao theorem states that the minimum variance on an estimate of μ is given by

$$V_{min}(\mu) = \frac{1}{nI(\mu)} \tag{4.4}$$

where n is the number of events used in the estimate. If there is more than one parameter to be estimated then the information becomes a matrix with elements,

$$I_{ij}(\bar{\mu}) = \int \left(\frac{\partial P}{\partial \mu_i}\frac{\partial P}{\partial \mu_j}\right) \frac{1}{P}\, dx \tag{4.5}$$

and the minimum variance matrix is equal to the inverse of the information matrix divided by the number of events. Note that diagonal elements of the variance matrix may be much larger than the inverse of the diagonal elements of the information matrix if there are large off-diagonal elements.

One can identify the locations of the minimum or maximum information of a parameter simply by looking at a plot of the information distribution, $(\partial\Gamma/\partial\mu)^2/\Gamma$, over the observable variables. Figure 4.3 shows this for μ replaced by R_2 and R_V, over the 2–dimensional spaces in $\cos\theta_v$,$\cos\theta_e$ and t,$\cos\theta_v$. For this calculation, the total differential distribution was integrated over all other variables except the two to be used in the plot, before calculating the information distribution. The maximum information on R_2 occurs at $t = 0$,

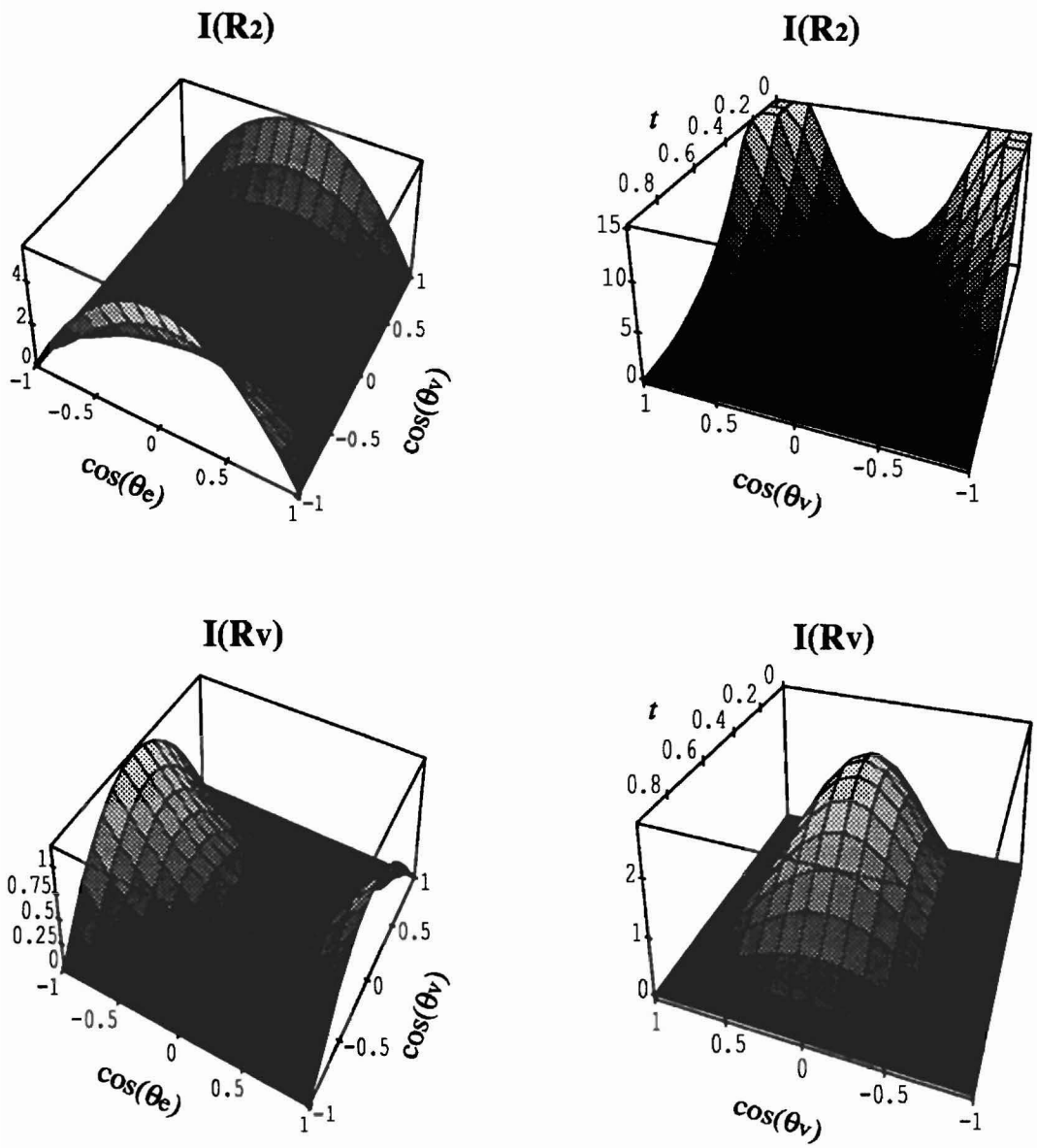

Figure 4.3 Information on R_2 and R_V plotted as a function of $\cos\theta_v$,$\cos\theta_e$ and t,$\cos\theta_v$.

Table 4.1 Minimum errors on R_2 and R_V.

Space Variables	$\sqrt{n}\ \sigma_{min}(R_2)$	$\sqrt{n}\ \sigma_{min}(R_V)$	ρ
$t, \cos\theta_v, \cos\theta_e, \chi$	3.53	4.64	-.30
$t, \cos\theta_v, \cos\theta_e$	3.66	6.10	-.32
$\cos\theta_v, \cos\theta_e$	5.70	7.10	-.49

$\cos\theta_e = 0$, and $\cos\theta_v = \pm 1$. There is little information about R_2 at large t and $\cos\theta_v$ near zero. The compliment is true for R_V except near $t = t_{max}$ where there is no information on either parameter. In addition, it can be seen that overall there is more information on R_2 due to the large peaks near $t = 0$.

The information matrix has been calculated for R_2 and R_V for different combinations of observable variables in two through four dimensions. Table 4.1 presents the results in terms of the minimum sigma, $\sigma_{min}(i) = \sqrt{V_{i,i}}$, and the correlation coefficient, $\rho = V_{i,j}/\sqrt{V_{i,i}V_{j,j}}$. Adding t to the fit significantly decreases the error on R_2 corresponding to a doubling of the information. The error on R_V is also decreased but not by as much. Adding χ to the fit does not affect the precision of R_2 but does decrease the error on R_V by about 25%. These calculations have not taken into account the affects of smearing acceptance or background. The minimum errors would rise with these effects included, with the error on R_2 rising strongly if the acceptance is low near $t = 0$ where the majority of the information on R_2 is obtained.

5 Fit Technique

A common technique for determining physical parameters from a data set is the maximum likelihood method in which one finds the values of the parameters which maximize the likelihood of the data set. The likelihood is calculated from an analytical probability distribution parametrized by the physical variables to be estimated. Thus one might expect that to estimate the form factor ratios all that needs to be done is to maximize the likelihood of the data using the full differential decay rate formula, eq. (2.22). The effects of acceptance and smearing, however, distort the data from the underlying physical distribution. If these effects are large then an analytical form for the distorted distribution must be found in order to perform a maximum likelihood fit. It may be very difficult to obtain an analytical form for the distorted distribution, however, especially over many dimensions. This is the situation we are confronted with. We wish to make a multi-dimensional maximum likelihood fit in order to use as much of the available information as possible; however, smearing due to incomplete information about the neutrino is significant and very difficult to parametrize over the multi-dimensional space of observables. This chapter first

details the acceptance and smearing problem and then explains our solution for making a maximum likelihood fit over many dimensions in the presence of such effects.

5.1 Smearing and Acceptance

5.1.1 Smearing

As described in section 3.4 the quadratic ambiguity in the longitudinal momentum of the neutrino causes smearing so that the calculated values of the kinematic variables (i.e. $\cos\theta_v$, $\cos\theta_e$, t, and χ) are sometimes different from their *true* values. Another cause of smearing is limited resolution in the direction of the D. This is taken to be the direction between the reconstructed production and decay vertices. However, the resolution on each of these vertices is limited, thereby limiting the resolution on the direction of the D.

One way to see the magnitude of the smearing is to take the Monte Carlo events and to calculate the variance of the smeared values around the true values. This was done as a function of the true variables so that one could see how the smearing changes over the full range of values. The results are shown in Fig. 5.1 for all four variables. The smearing corresponds to a standard deviation of between 10% to 30% of each variables range. Furthermore, the largest smearing occurs near $t = 0$ and $\cos\theta_e = -1$ which, as discussed in the previous chapter, is where most of the information on R_2 and R_V, respectively, is contained. This means that the information contained in the data is much less than that calculated in Chapter 4; thus, the errors on R_2 and R_V will be

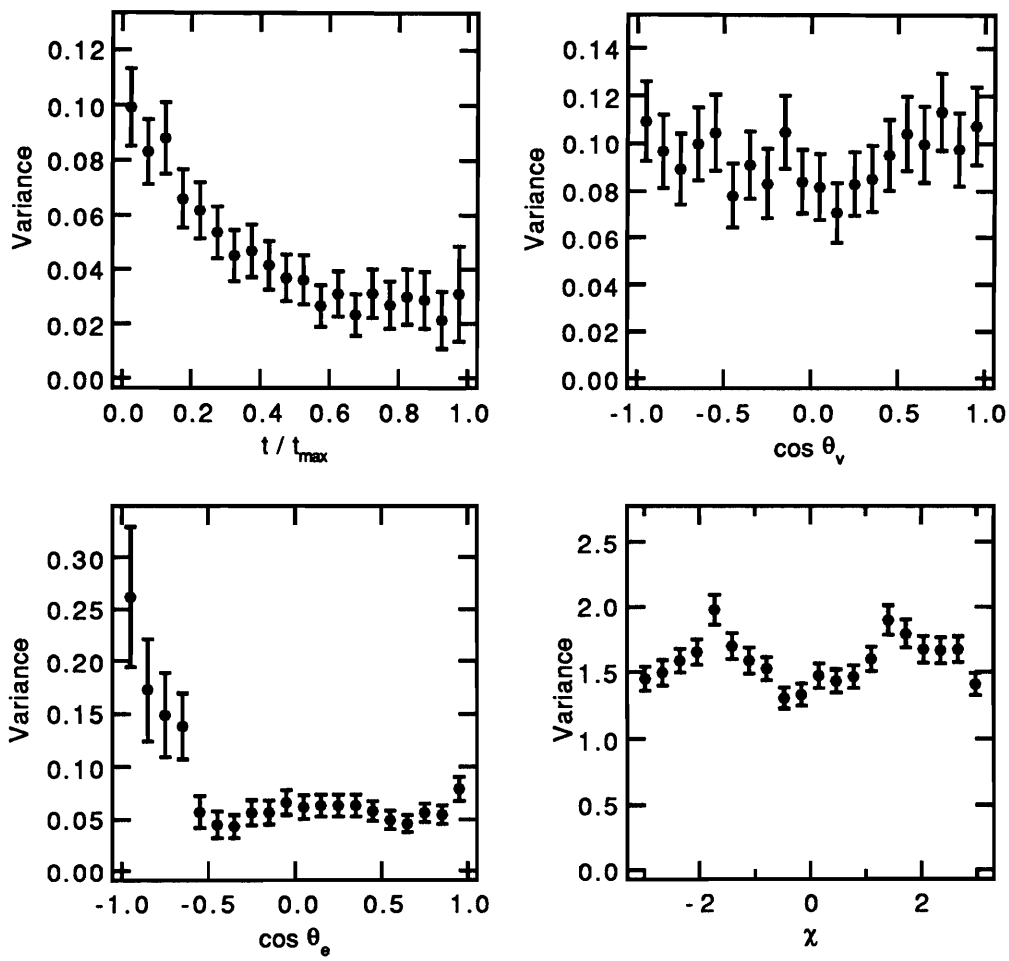

Figure 5.1 Variance of smeared t/t_{max}, $\cos\theta_v$, $\cos\theta_e$, and χ about the true values as a function of the true values.

larger than the minimum errors calculated from the Cramér-Rao theorem.

The effect of the smearing on the shape of the differential decay distribution is to flatten it out so that any peaks or valleys are less prominent. In addition, the large smearing at the edge of the range of t and $\cos\theta_e$ skews the decay distribution towards higher t and $\cos\theta_e$. Both these effects are significant and need to be included in measuring the form factor ratios.

5.1.2 Acceptance

The major effect of the acceptance is due to the requirement that the electron laboratory energy be greater than $12\,GeV$ which causes a low efficiency for decays with low t and $\cos\theta_e$ near -1. Again the Monte Carlo events are very useful for demonstrating this. Having generated the Monte Carlo with a known distribution over the kinematic variables, the resulting shape of the accepted Monte Carlo events can be compared and a shape for the acceptance calculated. The shape of the acceptance for all four variables is shown in Fig. 5.2. The acceptance drops about 90% from $\cos\theta_e = 1$ to $\cos\theta_e = -1$. This also tends to bias the observed decay distribution towards high $\cos\theta_e$ thus hiding the dominant H_- term which populates the area near $\cos\theta_e = -1$.

The effects of smearing and acceptance are significant and need to be included in the measurement of the form factors. We can use the Monte Carlo to estimate the shape of the smearing and acceptance functions, however, it is very difficult to parametrize them over the full four-dimensional space. For this reason a new fitting technique was developed which is described in the next section.

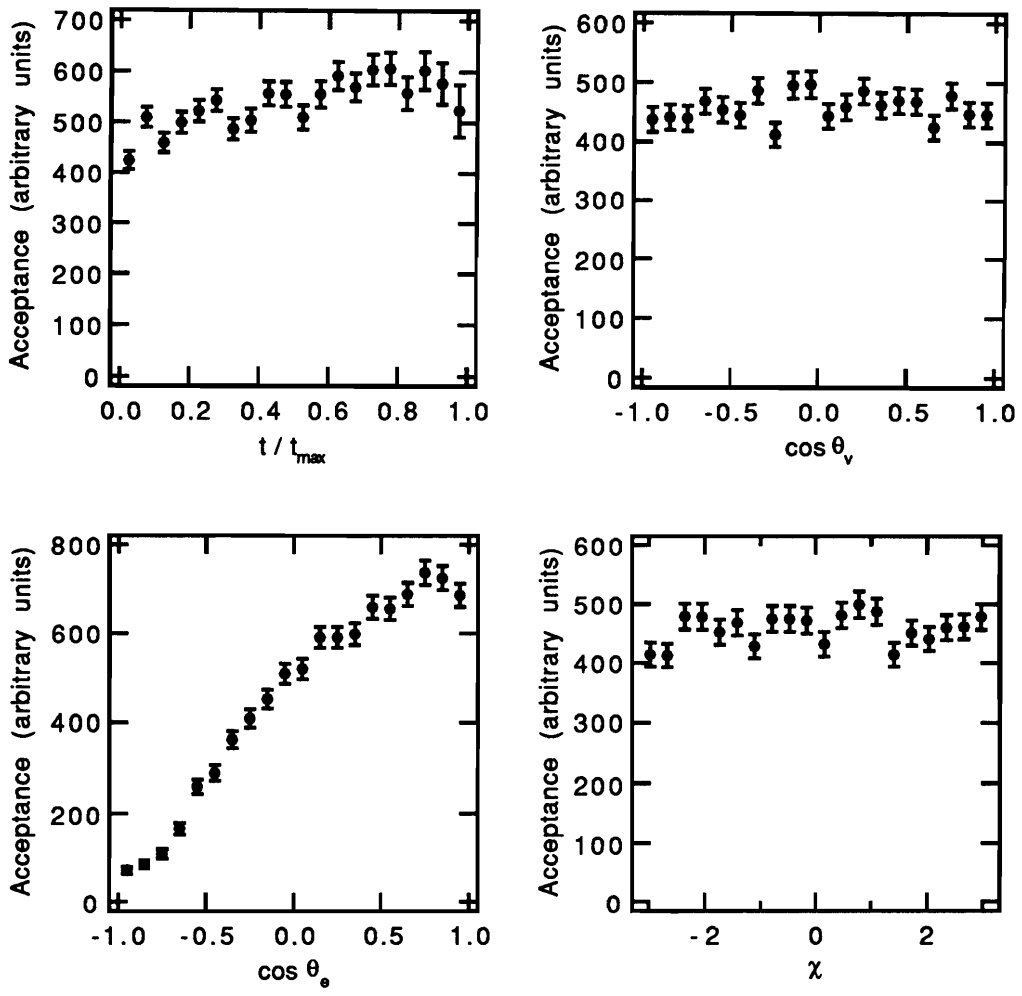

Figure 5.2 Shape of acceptance as a function of the true variables for t/t_{max}, $\cos\theta_v$, $\cos\theta_e$, and χ.

5.2 Description of Technique

Consider the maximum likelihood estimate of the polarization of a particle with differential decay distribution given by $\Gamma(x,\mu) = (1+\mu x^2)$. Here x is the cosine of the decay angle and the parameter μ, which is to be estimated with a maximum likelihood fit, is related to the polarization. The data consists of a set of n events with values of x, $\{x_i\}$. The log-likelihood of these events is given by

$$\ln L = \sum_{i=1}^{n} \Big(\ln \Gamma(x_i,\mu) - \ln N(\mu) \Big) \tag{5.1}$$

where $N(\mu)$ is the normalization, $N(\mu) = \int \Gamma(x,\mu)\, dx$. The maximum likelihood estimate of the parameter μ is that value of μ which maximizes the log-likelihood of the data. This is the standard maximum likelihood method of estimation of a parameter given an analytical formula for the probability distribution, $\Gamma(x,\mu)$.

In our case, $\Gamma(x,\mu)$ is the full differential decay distribution given in eq. (2.22); x represents the variables $\cos\theta_v$, $\cos\theta_e$, t, and χ; and μ represents R_2 and R_V. In addition, the effects of acceptance and smearing distort the data so that it is no longer distributed according to $\Gamma(x,\mu)$ but rather according to some other distribution. This modified, observable distribution, $F(\tilde{x},\mu)$, which is a function of the smeared, measured variables, $\tilde{x}$, is related to the underlying physical distribution by the equation:

$$F(\tilde{x},\mu) = A(\tilde{x}) \int S(\tilde{x},x)\Gamma(x,\mu)\, dx \tag{5.2}$$

where $A(\tilde{x})$ is the acceptance function, and $S(\tilde{x}, x)$ represents the smearing. It is this function, $F(\tilde{x}, \mu)$, which should be used in the calculation of the likelihood,

$$\ln L = \sum_{i=1}^{n} \Big(\ln F(\tilde{x}_i, \mu) - \ln N(\mu) \Big) \tag{5.3}$$

with $N(\mu) = \int F(\tilde{x}, \mu)\, d\tilde{x}$. However, in order to get an analytical form for $F(\tilde{x}, \mu)$, one must have an analytical form for the acceptance and smearing functions. Often this is accomplished by using a Monte Carlo simulation of the experiment. This scheme becomes very difficult, however, when the space of observables is multi-dimensional. Information about the shape of the acceptance and smearing functions is still contained in the Monte Carlo events; it is just difficult to parametrize.

Our method uses Monte Carlo events directly in the log-likelihood calculation. If our Monte Carlo simulation of the experiment properly includes the effects of smearing and acceptance, then Monte Carlo events that at the generated level are generated according to the distribution $\Gamma(x, \mu)$, will at the observable level be distributed according to $F(\tilde{x}, \mu)$. With a large enough set of Monte Carlo events the shape of $F(\tilde{x}, \mu)$ can be determined well enough so that the log-likelihood of the data can be calculated for a given value of μ. In order to perform a maximum likelihood fit using Monte Carlo events directly there are two main problems that must be solved: (1) generating large amounts of Monte Carlo events for different values of μ efficiently, and (2) calculating the log-likelihood from the Monte Carlo events even though there is not an analytical form for $F(\tilde{x}, \mu)$.

The first problem can be solved by using the same set of Monte Carlo events and weighting each event for a given value of μ. First, generate a set of m Monte Carlo events which at the generator level are distributed uniformly over the space of unsmeared observables, x. Let y_j (we use the letter y to distinguish Monte Carlo events from data events) be the generator-level values of the unsmeared observables for the j'th Monte Carlo event, and $\tilde{y}_j$ be the values of the smeared observables, $\tilde{x}$, for that same Monte Carlo event. Record the set $\{\tilde{y}_j, y_j\}$ for each Monte Carlo event. Then by weighting each Monte Carlo event with the value of $\Gamma(y_j, \mu)$ using the unsmeared values, the set $\{\tilde{y}_j\}$ of smeared values will be distributed according to $F(\tilde{x}, \mu)$. In this way the same set of Monte Carlo events can be used to represent $F(\tilde{x}, \mu)$ for any value of μ.

Generating the Monte Carlo events distributed uniformly might be inefficient. More generally, one can factorize the function, $\Gamma(x, \mu)$, into two terms: one which depends on μ, $W(x, \mu)$, and one which is independent of μ, $P(x)$. Then one generates Monte Carlo events according to $P(x)$ and weights them with $W(x, \mu)$. We chose $P(x)$ to be the phase space distribution, including the Breit-Wigner term, and $W(x, \mu)$ the matrix element squared.

The solution to the second problem comes from the fact that a histogram of weighted Monte Carlo events will asymptotically become indistinguishable from a plot of $F(\tilde{x}, \mu)$ as the number of Monte Carlo events increases and the size of the histogram bins decreases. One can estimate $F(\tilde{x}_i, \mu)$ for each data event by forming a small volume in $\tilde{x}$ space centered about $\tilde{x}_i$, summing the

weights of the Monte Carlo events whose smeared observable values, $\tilde{y}_j$, lie in this volume, and dividing by the size of the volume. By centering the volume around the point of interest a linear change in F over the volume does not contribute any error to the estimate of $F(\tilde{x}_i, \mu)$; only non-linear variations of F over the volume contributes to this systematic error. In addition, each Monte Carlo event may be in more than one volume. Normalization is accomplished by demanding that the sum of the weights of all Monte Carlo events is unity—for each value of μ.

With this technique the log-likelihood of the set of data events is given by

$$\ln L = \sum_{i=1}^{n} \ln\Big(\frac{\sum_{\tilde{y}_j in V_i} W(y_j, \mu)}{C(\mu) V_i}\Big) \tag{5.4}$$

where V_i is the volume centered around $\tilde{x}_i$ and $C(\mu) = \sum_{j=1}^{m} W(y_j, \mu)$ is the normalization.

There is also some amount of background present in the data set. Having estimated that out of a total of n events there are N_B background events distributed according to a normalized distribution, $P_B(\tilde{x})$, then the log-likelihood is given by

$$\ln L = \sum_{i=1}^{n} \ln\Big\{(n - N_B)\frac{\sum_{\tilde{y}_j in V_i} W(y_j, \mu)}{C(\mu) V_i} + N_B P_B(\tilde{x}_i)\Big\}. \tag{5.5}$$

Using this method for calculating the log-likelihood one can make a maximum likelihood fit to the data over any number of dimensions even with large smearing and acceptance effects.

The systematic errors associated with this fitting technique are due to: a) any non-linearity in F over the small volumes, V_i and b) limited Monte Carlo statistics in each volume. To minimize the former, volume size should be chosen small enough that F varies only linearly over the volume. However, with smaller volumes, fewer Monte Carlo events will be included in the calculation of $F(\tilde{x}_i, \mu)$, which will result in an increase of the systematic error of the second type.

One can estimate the magnitude of the first type of systematic error by redoing the fit with different volume sizes. For very small volumes the resulting values of μ will vary greatly due to limited Monte Carlo statistics but will be centered around the true μ. As the volume size is increased, the variation in the resulting values of μ for different fits will decrease but the average should remain centered around the true value if non-linear effects are not large at this volume size. But as the volume size increases to where non-linear effects are important, then the resulting values of μ will start to diverge from the true value. If there are enough Monte Carlo events there will be a range of volume sizes where the resulting values of μ are fairly stable. The standard deviation of these values within this nearly stable region is an estimate of the systematic error of the first type.

By holding the volumes fixed and redoing the fit a number of times with independent sets of Monte Carlo events, one can estimate the magnitude of the systematic error of the second type. A simple way of doing this is to divide the full Monte Carlo sample into 4 separate files and to perform the fit with

these 4 samples. The standard deviation of the resulting values of μ divided by 2 is then an estimate of the systematic error of the second type, assuming that this varies as the square root of the number of Monte Carlo events used. If the magnitude of these systematic errors is found to be unacceptably large then one can always reduce them by generating more Monte Carlo events.

6 Fit Results

Historically, we first measured the form factor ratios using only $\cos\theta_v$ and $\cos\theta_e$. After being satisfied that things were working well in two-dimensions, the fit was extended to include first t and then χ. For this dissertation, the fit will be described in detail for just the three-dimensional case and for the addition of χ since it was treated in a special way. Nevertheless, results from the two- through four-dimensional cases are shown. Having determined the form factor ratios, the branching ratio is measured using the measured values of the form factor ratios to estimate the acceptance. Then using the branching ratio value, the absolute value of $A_1(0)$ is calculated followed by that of $A_2(0)$ and $V(0)$.

6.1 Three-Dimensional Fit

Monte Carlo events were generated distributed according to phase space plus the Breit-Wigner of the $\bar{K}^{*0}$,

$$P(t, \cos\theta_v, \cos\theta_e; R_2, R_V) = \frac{K\Gamma(M_{K\pi})}{(M_{K\pi}^2 - M_{K^*}^2)^2 + M_{K^*}^2\Gamma^2(M_{K\pi})}, \tag{6.1}$$

where K is the momentum of the $\bar{K}^{*0}$ in the rest frame of the D^+. To get the distorted, observed distribution, $F(\tilde{t}, \widetilde{\cos\theta_v}, \widetilde{\cos\theta_e}; R_2, R_V)$, the Monte Carlo events are weighted with the matrix element squared (after integrating over χ),

$$\begin{aligned} W(t, \cos\theta_v, \cos\theta_e; R_2, R_V) = t\Big\{ & 4\sin^2\theta_e \cos^2\theta_v |H_0(t)|^2 \\ & + \left[(1+\cos\theta_e)^2|H_+(t)|^2 + (1-\cos\theta_e)^2|H_-(t)|^2\right]\sin^2\theta_v \Big\} \end{aligned} \tag{6.2}$$

using the generated, unsmeared values of the dynamical variables. The helicity amplitudes are parametrized by two form factor ratios (R_2, R_V) which are the quantities to be estimated (the μ parameters of the previous Chpt.);

$$\begin{aligned} H_\pm(t) &= \frac{(M_D + M_{K\pi})}{(1 - t/M_A^2)} \mp 2\frac{M_D K}{M_D + M_{K\pi}} \frac{R_V}{(1 - t/M_V^2)} \\ H_0(t) &= \frac{1}{2M_{K\pi}\sqrt{t}} \left[(M_D^2 - M_{K\pi}^2 - t)\frac{(M_D + M_{K\pi})}{(1 - t/M_A^2)} \right. \\ & \qquad \left. - 4\frac{M_D^2 K^2}{M_D + M_{K\pi}} \frac{R_2}{(1 - t/M_V^2)} \right]. \end{aligned} \tag{6.3}$$

Here $M_A = 2.53\ GeV$ and $M_V = 2.11\ GeV$ are the values of the poles used in assuming a single pole dominance for the t-dependence of the form factors. A total of 9000 accepted (out of one million generated) Monte Carlo events are used.

In calculating the log-likelihood, Monte Carlo events are binned into small three-dimensional volumes centered about each data point. As can be seen by the plots of the data in Fig. 6.1 at the end of this section, the distributions are

fairly linear so that bins of the order of 1/10 the range of each variable per side should be mostly free of non-linear effects which can increase the systematic error of the fit. This then will be a good starting volume size although different volume sizes will need to be used in order to establish the stability of the fit and to estimate systematic errors. We use the same volume size throughout the three-dimensional space with the exception of data events which occur near the boundary. For these events we shrink the volume so that the data event remains in the center. In addition we omit from the fit any data events whose Monte Carlo volumes did not contain at least 4 Monte Carlo events. This only occurs near the boundaries where volumes have been decreased a great deal and does not significantly affect the results.

The code which performs the fit consists of an initialization and a minimization section. The initialization section reads in the data and Monte Carlo events, calculates the volume dimensions for each data point, and identifies the Monte Carlo events which fall into each data event volume. Data events which do not contain at least 4 Monte Carlo events in their volumes are eliminated here. The minimization section consists of varying the values of R_2 and R_V until the negative log-likelihood, calculated using equation (5.5), is minimized.

In order to minimize the amount of background, the fit is performed using data events which are within the $\bar{K}^{*0}$ mass range of [0.8408, 0.9434] GeV. Only those Monte Carlo events whose true value of $M_{K\pi}$ is within this range are also used. This is important since the value of $M_{K\pi}$ is used in calculating the weights. In addition, since the maximum of t depends on the value of $M_{K\pi}$ we

Table 6.1 Results of fits for different volume dimensions.

Volume ($\times 10^{-3}$)	$\Delta \cos\theta_v$	$\Delta \cos\theta_e$	Δt	R_2	$\ln R_V$
0.004	0.033	0.027	0.05	-0.1 ± 0.5	-0.7 ± 0.7
0.36	0.067	0.053	0.1	0.2 ± 0.5	0.6 ± 0.4
0.81	0.09	0.09	0.1	0.1 ± 0.5	0.6 ± 0.4
1.2	0.1	0.08	0.15	0.0 ± 0.5	0.6 ± 0.4
4.0	0.2	0.1	0.2	-0.2 ± 0.5	0.7 ± 0.4
4.05	0.15	0.12	0.225	-0.1 ± 0.4	0.6 ± 0.4
32.4	0.3	0.24	0.45	-0.4 ± 0.5	0.8 ± 0.4

employ the normalized variable, $y = t/t_{max}$, when comparing smeared data to smeared Monte Carlo events in the fit. From our estimate of the total amount of background over the full $M_{K\pi}$ range given in Chapter 2, and the assumed shape of the background distribution, we estimate that there are 21 ± 5 background events within the narrow $M_{K\pi}$ range. For the background shape over the three-dimensional space of the fit, we want to use a parametrization which is consistent with that of the wrong-sign data. There are not many events in the wrong-sign sample so this information is very limited. Nevertheless, it appeared that a distribution that is uniform in $\cos\theta_v$ and $\cos\theta_e$ and linear in t was appropriate. We fit the y-distribution of the wrong-sign data to a linear function and used this, normalized over the full three-dimensional space, as the background shape. The result is $P(y, \cos\theta_v, \cos\theta_e) = (3y + 0.2)/3.4$ for the background function.

With this method it is important to perform the fit a number of times, varying the volume dimensions and number of Monte Carlo events used, in

Table 6.2 Result of fits with 4 independent Monte Carlo samples.

R_2	$\ln R_V$
0.1 ± 0.5	0.5 ± 0.4
0.0 ± 0.5	0.4 ± 0.5
0.4 ± 0.4	0.6 ± 0.4
0.1 ± 0.4	0.5 ± 0.4

order to insure the fit is stable and to estimate the magnitude of systematic errors. Table 6.1 shows the results of fits for different volume sizes using the same number of Monte Carlo events. Here we have expressed the results in terms of $\ln R_V$ which has a distribution that is more nearly Normal. The errors shown in the table are the statistical errors obtained by locating where the log-likelihood decreases by 0.5 from its maximal value. The results are fairly stable over a wide range of volumes with the exception of the very largest and the very smallest volumes. Neglecting these extreme cases, the average of the remaining values becomes our best estimate of the central values. Thus our measurement is $R_2 = 0.0 \pm 0.5$, $\ln R_V = 0.6 \pm 0.4$ where the errors shown are only the statistical errors from the data.

To estimate the systematic error of the first type, we take the standard deviation of the fit values from the trials within the nearly stable region. This yields $\sigma_1 = 0.16$ and $\sigma_1 = 0.04$ for R_2 and $\ln R_V$ respectively.

To estimate the systematic error of the second type, we hold the volume dimensions fixed at the ones shown in the forth row of Table 6.1 which yield the best central values, divide the Monte Carlo events into 4 files, and redo

the fit with these 4 independent sets. The results are shown in Table 6.2. The standard deviation of the values for R_2 and $\ln R_V$ is taken as an estimate of the second type of systematic error for a quarter of the Monte Carlo events. Assuming this error goes as $1/\sqrt{N}$, the error for the full Monte Carlo sample should be half these values. This yields $\sigma_2 = 0.09$ and $\sigma_2 = 0.08$ for R_2 and $\ln R_V$ respectively. The total systematic error we take to be the sum (in quadrature) of these two types of systematic errors. Thus our result is $R_2 = 0.0 \pm 0.5 \pm 0.2$, $\ln R_V = 0.6 \pm 0.4 \pm 0.1$, or $R_V = 1.8 \pm 0.7 \pm 0.2$, where the first error is statistical associated with the data and the second is systematic associated with the fit method. There is also the systematic errors associated with the data and with the uncertainty in the background; these will be discussed for the final four-dimensional fit in the next section.

We looked at many plots to check the integrity of the fit. Figure 6.1 compares the data and Monte Carlo distributions over various slices through the three-dimensional space. In all cases the data and Monte Carlo follow each other quite well. Also, one can see that the deviation of the Monte Carlo around a smooth curve is much smaller than that of the data. This indicates that the systematic error due to the statistics of the Monte Carlo is small compared to that of the data.

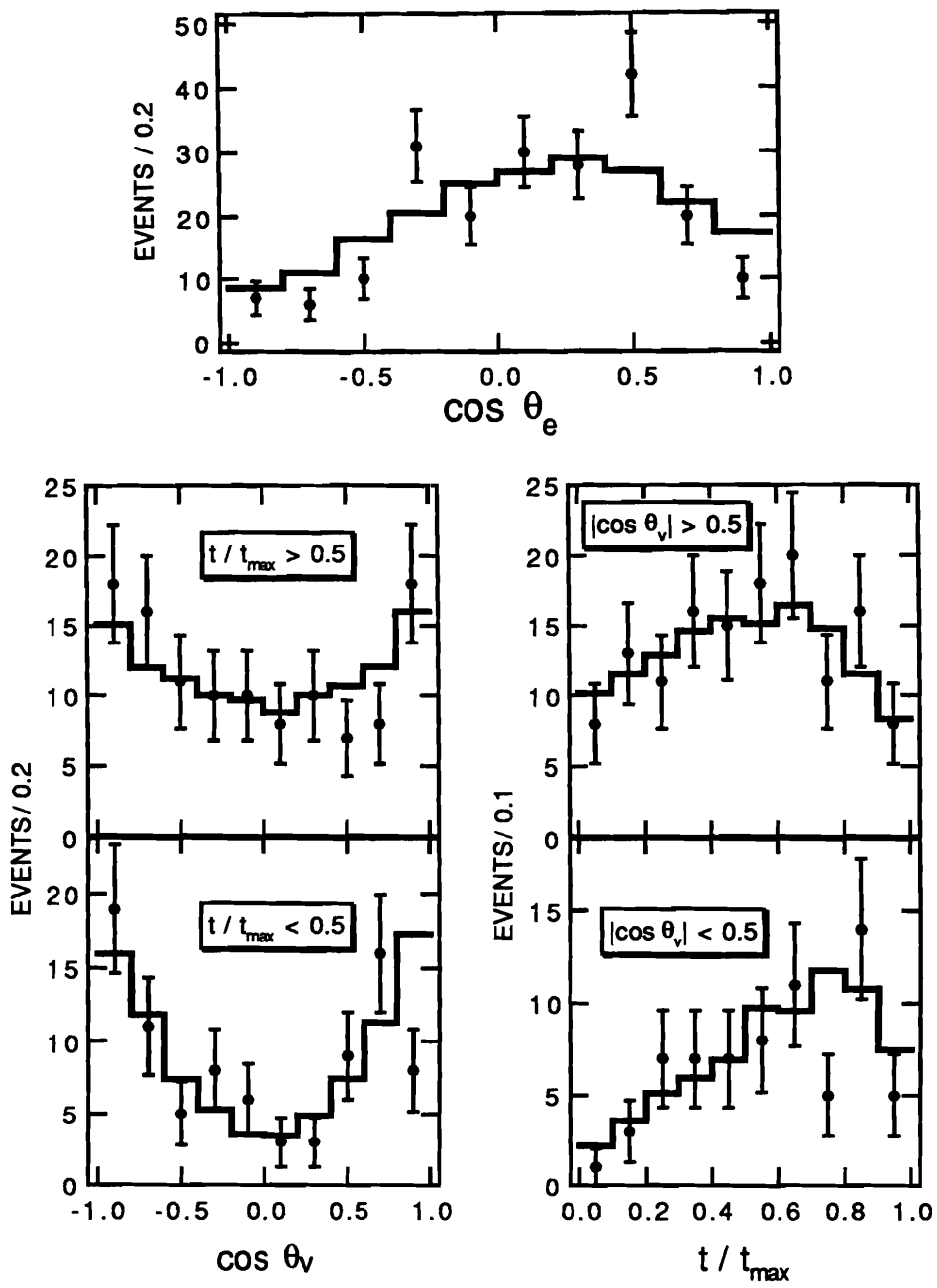

Figure 6.1 Projection of the data (dots with error bars) and Monte Carlo (solid histogram) events onto $\cos\theta_e$, $\cos\theta_v$, and t/t_{max}. The Monte Carlo events are weighted using the measured form factor ratios.

6.2 Four-Dimensional Fit

In addition to the longitudinal and transverse terms in the matrix element, there are also three interference terms which depend on the angle between the decay planes of the $e^- \nu_e$ and the $K^- \pi^+$, labeled χ,

$$\begin{aligned} &-2\sin^2\theta_e \sin^2\theta_v \cos 2\chi \, H_+ H_- \\ &-4\sin\theta_e\,(1+\cos\theta_e)\sin\theta_v \cos\theta_v \cos\chi \, H_+ H_0 \\ &+4\sin\theta_e\,(1-\cos\theta_e)\sin\theta_v \cos\theta_v \cos\chi \, H_- H_0. \end{aligned} \tag{6.4}$$

These terms integrate to zero and were not used in the previous three-dimensional fit. The $H_- H_0$ term should be largest but only over a small region of phase space. We have plotted the data and Monte Carlo events as a function of χ in Fig. 6.2 using the values of the form factor ratios found above. One can see the expected $\cos\chi$ dependence in (a); however, the small $\cos 2\chi$ dependence that is expected in (b) is lost in the limited statistics of the data. The fact that the data and Monte Carlo projections match each other fairly well over χ provides an independent check that the data and Monte Carlo agree since nowhere was information about χ used in the previous fit.

Including χ in the fit is complicated by the large smearing that occurs when χ is ill-defined. This happens when the $K^- \pi^+$ or the $e^- \nu_e$ lie in a line with the direction of the D^+. Fortunately this is where the interference terms go to zero. Rather than bin the Monte Carlo into four-dimensional bins, which would have decreased the number of Monte Carlo events per bin, we decided to do the following. We bin the Monte Carlo events into three-dimensional boxes

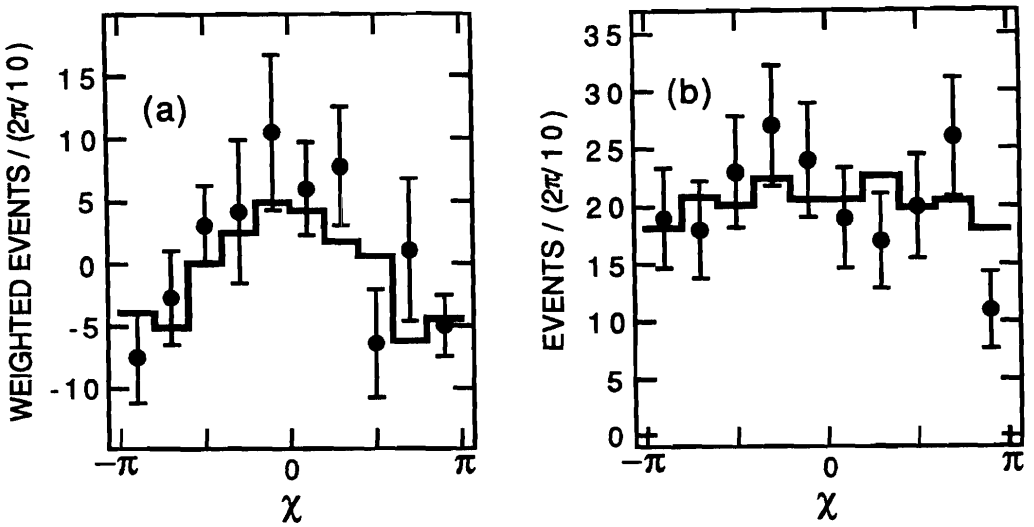

Figure 6.2 Projection of the data (dots with error bars) and Monte Carlo (solid histogram) events onto χ. In (a) events are weighted with the coefficient which multiplies the $\cos\chi$ term in the decay rate. This brings out the $\cos\chi$ dependence. (b) is a standard projection which should have a small $\cos 2\chi$ dependence.

as before, and when adding the weights around each data event include the interference weights but evaluated with $\chi_{data} + \chi_{observed} - \chi_{generated}$ replacing $\chi_{generated}$. When the smearing is small, generated and observed χ are close so one gets the weight at the χ of the data. When the smearing is large then the weights fluctuate and add up to zero. Thus by rotating the χ that goes into the weights in this way, one is effectively getting out the weight at the χ of the data event. This works because the amount of smearing in χ is largely determined solely by the values of $\cos\theta_e$ and $\cos\theta_v$.

Putting this into the fit we find $R_2 = 0.0\pm0.5\pm0.15$ and $R_V = 2.0\pm0.6\pm0.3$ where the first error is due to the statistics of the data and the second comes from the systematics of the fitting method calculated in the way described for

the three-dimensional fit. In making this fit we assumed that 21 out of the 204 data events were background and were distributed in a manner consistent with the wrong-sign data. The choice of 21 had a standard deviation of 5, also found from the wrong-sign data. By varying 21 by 5 either way we found the error due to this uncertainty was .1 for R_2 and .04 for R_V. The choice for possible background distributions also led to a similarly small error.

We felt the greatest error in the Monte Carlo itself was in the electron identification efficiency. We estimated this could be off by at most 7% and modified each Monte Carlo event weight by $1 - .07\frac{12}{E}$ where E is the energy (in GeV) of the electron in the lab frame and the 12 comes from the fact that we have a 12 GeV electron energy cutoff. We used a linear dependence since a constant term would go out in the normalization. The resulting deviation of R_2 and R_V was less than 10% of the error due to that of the fitting method.

Adding these systematic errors in quadrature with those due to the fitting method yields our final result:

$$R_2 = 0.0 \pm 0.5 \pm 0.2 \quad R_V = 2.0 \pm 0.6 \pm 0.3 \tag{6.5}$$

with correlation coefficient, $\rho_{R_2,R_V} = -0.23$. With these values we can calculate the ratio of longitudinal to transverse polarization and compare it to the value measured earlier[11] using this same data set. That measurement found $\frac{\Gamma_L}{\Gamma_T} = 2.4^{+1.7}_{-0.9} \pm 0.2$ and only used the $\cos\theta_v$ projection of the data. The value using the measured form factor ratios from the full four-dimensional distribution of the data is $\frac{\Gamma_L}{\Gamma_T} = 1.8^{+0.6}_{-0.4} \pm 0.3$. This is consistent with the previous

Table 6.3 Summary of results on R_2 and R_V and comparison with theoretical minimum errors.

Space Variables	Fit $R_2 \pm \sigma_{(R_2)}$	Fit $R_V \pm \frac{\sigma_{(R_V)}}{R_V}$	$min\ \sigma_{(R_2)}$	$\frac{min\ \sigma_{(R_V)}}{R_V}$
$t, \cos\theta_v, \cos\theta_e, \chi$	0.0 ± 0.5	2.0 ± 0.30	0.26	0.17
$t, \cos\theta_v, \cos\theta_e$	0.0 ± 0.5	1.8 ± 0.39	0.27	0.23
$\cos\theta_v, \cos\theta_e$	$-0.4^{+1.0}_{-1.6}$	3.0 ± 0.55	0.42	0.26

result, however, the errors are smaller by more than a factor of two. This gain in precision comes about mainly from adding t to the fit. The transverse helicity amplitudes are proportional to t and die at $t = 0$, while the longitudinal helicity amplitude has no suppression at this point. Thus one obtains very good separation.

6.2.1 Summary

It is useful to compare the results of the fits in two through four dimensions and with the minimum errors calculated in Chapter 4. This is shown in Table 6.3. The minimum errors were calculated assuming the values of $R_2 = 0.7$ and $R_V = 2.0$. With different values for R_2 and R_V the minimum errors will change. We have attempted to account for this by tabulating the ratio of the error to the central value for R_V. This is not useful for R_2 since the central values are close to zero. Nevertheless, it appears from Table 6.3 that the measured errors follow the pattern of the minimum ones quite well. The two-dimensional fit yielded rather large errors; adding t to the fit significantly decreased the error on R_2. Adding χ to the fit reduced the error on R_V. In general, there is about a factor of two difference between the measured errors

and the minimum possible errors. This means that the effects of smearing and acceptance eliminated about three quarters of the possible information. This was expected given the poor acceptance and large smearing near $t = 0$ and $\cos\theta_e = -1$ which is where the information on R_2 and R_V, respectively, is greatest.

6.3 Branching Ratio

Having obtained the form factor ratios, we re-analyze the branching ratio. The previous measurement of the branching ratio[11] used a value for $|H_+|/|H_-|$ derived from a model[12] and the value of Γ_L/Γ_T derived from the observed $\cos\theta_v$ dependence. The effective form factor ratios implicit in these values are different from those we have measured, which affects the estimate of the efficiency used to obtain the branching ratio. The method is to first fit the data projected onto $M_{K\pi}$ to obtain the number of $D^+ \to \bar{K}^{*0} e^+ \nu_e$ events present. Then we calculate the efficiency for this mode and divide the found number of events by this efficiency to obtain the total number of $D^+ \to \bar{K}^{*0} e^+ \nu_e$ events present in the full data set. Finally we compare this number to that for the mode, $D^+ \to K^- \pi^+ \pi^+$, and use a previously measured value of the branching ratio for this mode to get the branching ratio of $D^+ \to \bar{K}^{*0} e^+ \nu_e$.

To fit the data projected onto the $M_{K\pi}$ axis we need the functional form of the signal as well as that for the background. For the signal, we again use the Monte Carlo events directly, weighting them using the measured form factor ratios. We do this because phase space and the matrix element alter

the shape of the signal when projected onto the $M_{K\pi}$ axis so that it is not just that of the Breit-Wigner. The previous analysis of the branching ratio[11] used a plain Breit-Wigner shape for the signal; thus the present analysis differs in not only using a different shape for the signal, but also a different efficiency from different form factor ratios. We also include radiative corrections to the electron spectrum as outlined in Ref. 13 (this only has a 0.7% effect).

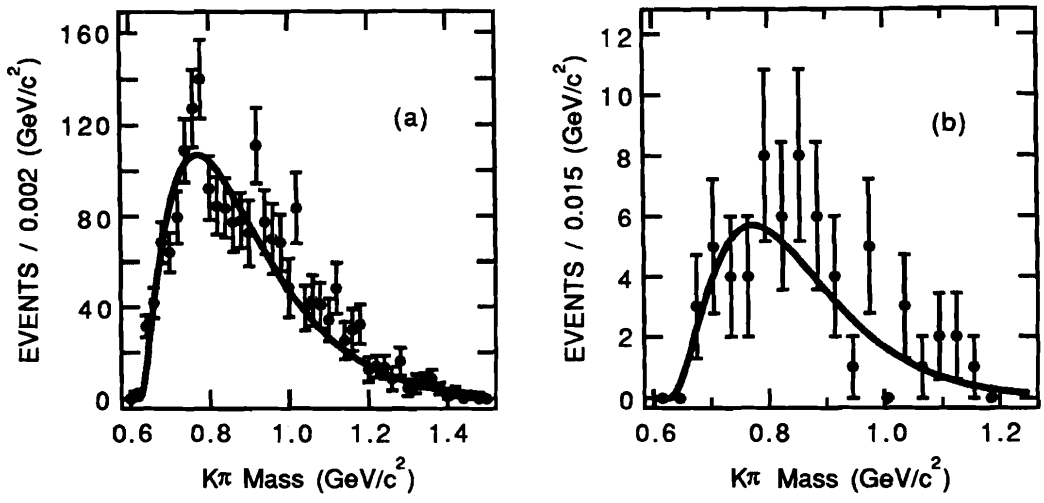

Figure 6.3 (a) S-wave Monte Carlo data (dots with error bars) and fitted Gamma distribution (solid curve). (b) wrong-sign data (dots with error bars) and its fitted Gamma distribution.

The background consists of nonresonant $K^-\pi^+$ signal plus true background events. To estimate the shape of the nonresonant signal, we generated Monte Carlo events in which the $K^-\pi^+$ are in a relative S-wave state. That distribution is shown in Fig. 6.3 (a) along with a fit to a Gamma distribution of the form, $x^\alpha e^{-\beta x}$. Here, $x = (M_{K\pi} - M_0)$, α and β, are parameters to be determined and M_0 is set to $0.63\,GeV$. The estimated values of α and

β are 1.2 ± 0.1 and 8.7 ± 0.4 respectively. For the true background we use the wrong-sign data. Its distribution is shown in Fig. 6.3 (b) along with its Gamma distribution. Here the values of α and β were found to be 1.8 ± 0.8 and 14 ± 5 respectively. So there is not much difference between the shape of the nonresonant signal and the true background. Thus we simply use one distribution for the background; that of a gamma distribution whose parameters, α and β are consistent with those shown above. These parameters will be varied within the errors shown above and the fit performed a number of times in order to estimate the systematic error associated with not knowing the exact shape of the background.

We use a binned maximum likelihood fit with the negative log-likelihood given by:

$$-\ln(L) = N_S + N_B - \sum_{i=1}^{N_{bins}} N_i \ln\left[N_S P_S(i) + N_B P_B(i)\right] \tag{6.6}$$

where N_i is the number of data events in the i'th bin, N_S and N_B are the number of signal and background events, respectively, to be determined, and $P_S(i)$ and $P_B(i)$ are the signal and background probabilities at the $M_{K\pi}$ value of the i'th bin. This form of the log-likelihood includes the Poisson probability of finding the total number of events when expecting $(N_S + N_B)$ events so that one can extract the actual values of N_S and N_B from the fit instead of just their ratio. For $P_S(i)$ we add up the weights of the Monte Carlo events which lie in the i'th bin, divide by the bin width, and divide by the sum of weights of all the Monte Carlo events to insure proper normalization.

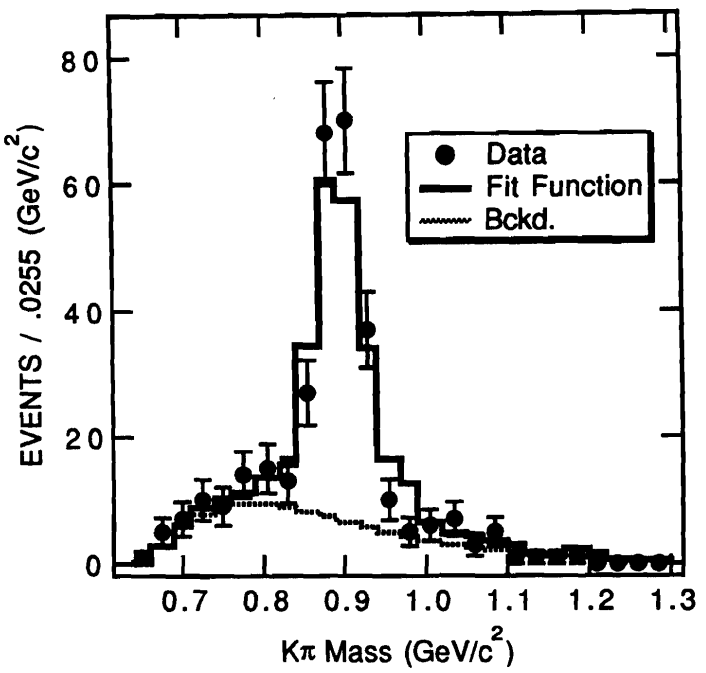

Figure 6.4 Histogram of $K\pi$ mass distribution together with that of the fitting function. The shape of the background used in the fit is also shown.

For $P_B(i)$, we simply take the value of the normalized Gamma distribution at the central value of $M_{K\pi}$ for the i'th bin. The fit is made in the $M_{K\pi}$ mass range, $[0.6371, 1.4021]\,GeV$ where there are 317 data events. We find $N_S = 209 \pm 18 \pm 9$ and $N_B = 108 \pm 17 \pm 9$. The systematic errors are 2.7% from the choice of binning, 3.3% from the uncertainty of the background shape, and 0.7% due to the errors of R_2 and R_V. The resulting signal plus background shape together with that of the data is shown in Fig. 6.4. The previous analysis found 227 signal events. The difference is due to the modifications of the Breit-Wigner shape described earlier.

The efficiency is calculated by summing the weights of the accepted Monte Carlo events and dividing by the sum of the weights of the generated Monte

Carlo events. Again, the measured values of the form factor ratios are used when calculating the weight for an event. With these form factors, the efficiency is found to be $(1.32 \pm 0.03 \pm 0.09)\%$. Here the systematic error is dominated by the 7% error on the electron identification efficiency. The efficiency corrected number of $D^+ \to \bar{K}^{*0} e^+ \nu_e$ events is thus $15870 \pm 1330 \pm 1320$ where the errors on the efficiency are treated as systematic. To correct for a possible contamination from $D^+ \to K^- \pi^+ \pi^+ \pi^0$ a Monte Carlo simulation of this mode was performed. The efficiency for a $D^+ \to K^- \pi^+ \pi^+ \pi^0$ event to simulate a $D^+ \to \bar{K}^{*0} e^+ \nu_e$ event was found to be 3×10^{-5}. Assuming a branching ratio for this mode of 7%, we expect about 0.5% of the data events to be from this mode. To correct for this we lower the efficiency corrected number of events by 0.5% to obtain 15790 events. This number, however, only includes $\bar{K}^{*0} \to K^- \pi^+$. A third of the time, $\bar{K}^{*0}$ decays to $K^0 \pi^0$. Thus the actual efficiency corrected number of $D^+ \to \bar{K}^{*0} e^+ \nu_e$ is $23690 \pm 1990 \pm 1970$.

To obtain the branching ratio we compare to the mode $D^+ \to K^- \pi^+ \pi^+$ such that

$$\frac{Br(D^+ \to \bar{K}^{*0} e^+ \nu_e)}{Br(D^+ \to K^- \pi^+ \pi^+)} = \frac{\mathcal{N}(\bar{K}^{*0})}{\mathcal{N}(K^- \pi^+ \pi^+)} \frac{\mathcal{L}(K^- \pi^+ \pi^+)}{\mathcal{L}(\bar{K}^{*0})} \tag{6.7}$$

where $\mathcal{N}$ is the efficiency corrected number of events and $\mathcal{L}$ is the effective luminosity for each mode. It turns out that this ratio of luminosities is not one but 1.045 ± 0.01 due to some runs being missed in the stripping process. The $D^+ \to K^- \pi^+ \pi^+$ mode has been standardized within E691 since it is used as a normalization mode. The efficiency corrected number of $D^+ \to K^- \pi^+ \pi^+$

events is 51870 ± 1820. From equation (6.7) we obtain a branching ratio for $D^+ \to \bar{K}^{*0}e^+\nu_e$ of $(0.48 \pm 0.04 \pm 0.04)Br(D^+ \to K^-\pi^+\pi^+)$. Using the measured value of $Br(D^+ \to K^-\pi^+\pi^+) = (9.11 \pm 1.3 \pm 0.4)\%$ obtained by the Mark III collaboration[14] and treating the errors in this measurement as systematic, yields

$$Br(D^+ \to \bar{K}^{*0}e^+\nu_e) = (4.4 \pm 0.4 \pm 0.8)\%. \tag{6.8}$$

This is not much different from the 4.5 value found in the previous analysis[11] because although the efficiency went down, so did the number of observed events.

6.4 The Normalized Form Factors

We calculate $A_1(0)$ by equating the total decay rate of this mode to the measured branching ratio divided by the lifetime of the D^+. The total decay rate is given by

$$\begin{aligned}\Gamma = A_1^2(0) \int_0^\infty dM_{K\pi} \int_{M_{e^-}^2}^{t_{max}} dt \; \Bigg\{ G_F^2 |V_{cs}|^2 \frac{8}{3(4\pi)^4 M_D^2} \; Kt \\ \times \frac{M_{K^*}\,\Gamma(M_{K\pi})}{(M_{K\pi}^2 - M_{K^*}^2)^2 + M_{K^*}^2 \Gamma^2(M_{K\pi})} \; (H_0^2 + H_-^2 + H_+^2) \Bigg\}\end{aligned} \tag{6.9}$$

where the helicity amplitudes, given in eq. (6.3), depend on the values of R_2 and R_V. We use the measured values of the form factor ratios, the branching

ratio, and the lifetime[15] $\tau_{D_+} = 1.090 \pm 0.030 \pm 0.025\ ps$, all obtained from the E691 experiment. With $V_{cs} = 0.975$ we find

$$A_1(0) = 0.46 \pm 0.05 \pm 0.05 \quad A_2(0) = 0.0 \pm 0.2 \pm 0.1 \quad V(0) = 0.9 \pm 0.3 \pm 0.1.$$

The statistical error for $A_1(0)$ comes from a statistical error of 0.04 and 0.02 for the form factor ratios and the branching ratio, respectively. The $A_1(0)$ systematic error is built out of systematic errors of 0.02 and 0.04 for the form factor ratios and the branching ratio, respectively. Only the uncertainty of R_2 and R_V contributes significantly to the errors of $A_2(0)$ and $V(0)$.

The only assumption that is made in measuring the form factors is the dependence on t. Because of the limited range of t ($t_{max}/M_v^2 = 0.2$, where M_v is the value of the vector pole mass), the results are only sensitive to the slope of the form factor near $t = 0$, and only mildly sensitive to that. In some models, other forms for the t dependence are used, but they are effectively equivalent at the present level of the measurement. The only quantitative information on the t dependence of charm form factors comes in the $D^0 \to K^- e^+ \nu_e$ mode, in which the pole mass is measured directly.[16] That result can be expressed as $f(t) = f_p(t)(1 + St)$, where f_p is the simple pole form with the vector mass and S is a measure of the correction to the slope, which was measured to be 0.00 ± 0.07. A similar uncertainty in the pole form used for the present analysis would cause an additional error of ± 0.015 in $A_1(0)$, ± 0.04 in $A_2(0)$, and ± 0.02 in $V(0)$. When these errors are added in quadrature to other systematic errors, the total systematic error does not change at the level of precision quoted; those

systematic errors are in turn smaller than the statistical errors. Thus there is no significant change in the form factors unless the t dependence is both outside the range of models and in complete disagreement with that measured in $D^0 \to K^- e^+ \nu_e$.

7 Conclusions

Having determined the form factor values we compare these results with those of other experiments and predictions from theory. Finally we look to the future and examine how these and future measurements of charm form factors will effect the determination of the Kobayashi-Maskawa matrix elements, V_{cb} and V_{ub}.

7.1 Other Experiments

After we published the measurement of the form factors[17] two other experimental groups subsequently made their own measurements: Fermilab experiments E653[18] and E687.[19] Both these experiments observe the semi-muonic decay $D^+ \rightarrow \bar{K}^{*0}\mu^+\nu_e$. The same form factors govern this decay.* The only significant effect of the muon is to increase the lower limit of t from the electron mass squared to that of the muon mass squared. This effect was included in the measurement by E653 and they were able to demonstrate that the mea-

* A fourth form factor which is always proportional to the lepton mass in the decay rate is still not important for muons.

Table 7.1 Results from this work compared with those of other experiments.

	E691	E653[18]	E687[19]	Wtd. Avg.
R_2	$0.0 \pm 0.5 \pm 0.2$	$0.8 \pm 0.2 \pm 0.1$	0.5 ± 0.4	0.65 ± 0.1
R_V	$2.0 \pm 0.6 \pm 0.3$	$2.0 \pm 0.3 \pm 0.2$	2.1 ± 0.6	2.0 ± 0.3
$A_1(0)$	$0.46 \pm 0.05 \pm 0.05$	0.49 ± 0.07		0.475 ± 0.05
$A_2(0)$	$0.0 \pm 0.2 \pm 0.1$	0.4 ± 0.14	0.2 ± 0.2	0.26 ± 0.1
$V(0)$	$0.9 \pm 0.3 \pm 0.1$	1.0 ± 0.3	1.0 ± 0.3	1.0 ± 0.2

sured form factor ratios were stable with respect to a cut on the minimum allowed value of t/t_{max}.

Table 7.1 compares the results of this work with those of other experiments. The weighted average (as used by the Particle Data Group) of the results for each quantity is also listed. The E687 group has only presented results on the form factor ratios. We use the weighted average of $A_1(0)$ to compute the values of $A_2(0)$ and $V(0)$ from their measured form factor ratios. Both the E653 and the E687 results are consistent with ours with the greatest discrepancy occurring for $A_2(0)$ and R_2 where there is about a two sigma difference.

Perhaps the most surprising aspect of the E653 result is that their errors for the form factor ratios are so small; they are a factor of two smaller than ours. Yet, they do not have much more data than we do; we have 204 events with 21 as background whereas they have 305 events with about 20 as background. They do have better vertex resolution and acceptance, however. One can check whether it is actually possible to obtain such small errors by using the results presented in Chpt. 4 on the minimum possible errors. Those results, scaled for

Table 7.2 Comparison of measured form factors with theoretical predictions.

	Group	$A_1(0)$	$A_2(0)$	$V(0)$
Experiment	Wtd. Avg.	0.475 ± 0.05	0.26 ± 0.1	1.0 ± 0.2
Models	ISGW[20]	0.8	0.8	1.1
	AW/GS[21]	0.8	0.6	1.5
	BSW[22]	0.9	1.2	1.3
	KS[23]	1.0	1.0	1.0
Sum Rules	BBD[24]	0.5 ± 0.15	0.6 ± 0.15	1.1 ± 0.25
	AOS[25]	0.9 ± 0.2	0.8 ± 0.3	1.7 ± 0.6
Lattice	BKS[26]	$0.8 \pm 0.14 \pm 0.3$	$0.6 \pm 0.14 \pm 0.2$	$1.4 \pm 0.5 \pm 0.5$
	LMS[27]	$0.52 \pm (0.07 \times 3)$	$0.1 \pm (0.35 \times 3)$	$0.85 \pm (0.08 \times 3)$

285 events are: $\sigma_{min}(R_2) = 0.21$ and $\sigma_{min}(R_V) = 0.28$ to be compared with E653's measured statistical errors of 0.23 and 0.33 respectively. Thus their errors are right at the minimum possible edge. This is surprising given the existence of acceptance and smearing effects plus background in the real data which are not included in the calculation of the minimum possible errors.

7.2 Comparison with Theory

To expedite comparisons, the weighted average of the measured form factors are listed together with theoretical predictions in Table 7.2. All of the models fail to predict the low measured values of the two axial vector form factors. This observation is even more striking when one takes into account the fact that there is generally good agreement between predicted and measured values of the vector form factor governing the decay $D \to Ke\nu_e$. Thus

it appears that the models systematically fail with axial vector form factors but succeed with vector form factors.

After the discrepancy between the models and measured values became known, BSW reanalyzed their model to see if this discrepancy could be reconciled.[28] They found that the form factors had a strong dependence on the value of a quarkmass-dependent integral over the meson wave functions. They therefore let the value of this integral be a free parameter and listed the value of the form factors as a function of this parameter. However, this modification left the ratio, R_V, constant at 1.4 which is about two sigmas away from the weighted average of the measured values, 2.0 ± 0.3. Similarly, ISGW tried to reconcile the non-relativistic quark model but found that only a 20% change could be affected—which is not enough.[29]

Unfortunately, the QCD Sum Rules and Lattice QCD results, which have a stronger grounding in the Standard Model, are at present not known very precisely so that they are consistent, at about the two sigma level, with both the measured values and the model values. Nevertheless, the central values from these two approaches tend to be going in the right direction, towards lower values for the axial vector form factors.

The application of HQET symmetries to this decay is difficult since the corrections are of the order Λ_{QCD}/M_s which is between 70 and 200%. So clearly, at least the first order corrections would need to be calculated in order to even guess at the form factors from an HQET point of view. The first order corrections have been calculated for $b \to c$ decay using QCD sum rules.[6]

The form factor ratios, R_2 and R_V, are defined for $B \to D^* l \nu_l$ decay in the same way as for charm decay and in the limit of exact HQET symmetry these quantities are unity. It is interesting to note that the first order corrections for the b decay form factors find R_2 decreasing by 10% and R_V increasing by about 30%.

Although the $D^+ \to \bar{K}^{*0} e^+ \nu_e$ form factors appear to be difficult to calculate theoretically, fortunately, they can be measured and this information is useful in itself. For example the measured charm factors can be used to test the factorization hypothesis in charm hadronic decays, be related to the form factors in $B \to K^* e^+ e^-$ using the symmetries of HQET,[7] and serve as a significant test for any future theory which attempts to calculate heavy quark semileptonic form factors. In addition, the method developed here for measuring form factors from the complete differential decay distribution serves as a prototype for similar measurements in the decays $B \to D^* l \nu_l$ and $D \to \rho l \nu_l$ which are needed to determine the important Kobayashi-Maskawa matrix elements, V_{cb} and V_{ub}.

APPENDIX A

Cleo-II Silicon Detectors

Silicon vertex detectors have been successful in identifying charm meson decay vertices in fixed target experiments such as E691 and are now finding use in collider experiments for tagging events with heavy quarks and measuring lifetimes. A silicon microstrip detector is presently being constructed for the CLEO-II experiment at the CESR e^+e^- collider, operating at center of mass energies close to the $\Upsilon(4S)$. This detector will permit the identification of detached decay vertices from charmed mesons, tau leptons, and other long-lived particles. Since the momentum spectrum of tracks at the $\Upsilon(4S)$ is soft the detector must be as low mass as possible; in addition, it must provide z measurements in order to reconstruct vertices in 3 dimensions. These considerations have led to the use of double-sided silicon microstrip detectors.

Double sided silicon detectors operate in the same manner as for the single sided detectors described in Chpt. 3. The main difference is that instead of having a uniform layer of n-type material on the back, there are a set of strips of n-type material. This is shown in Fig. A.1. These strips, when biased, attract the cloud of electrons that are generated by an ionizing particle, so that the position of the ionizing particle can be determined on this n-side in the same manner as for the p-type strips on the opposite side of the detector. By orienting the n-strips perpendicular to those on the p-side, both the x *and* y positions of the ionizing particle can be measured. There is an additional feature that makes the n-side a bit more complicated than the p-side. An

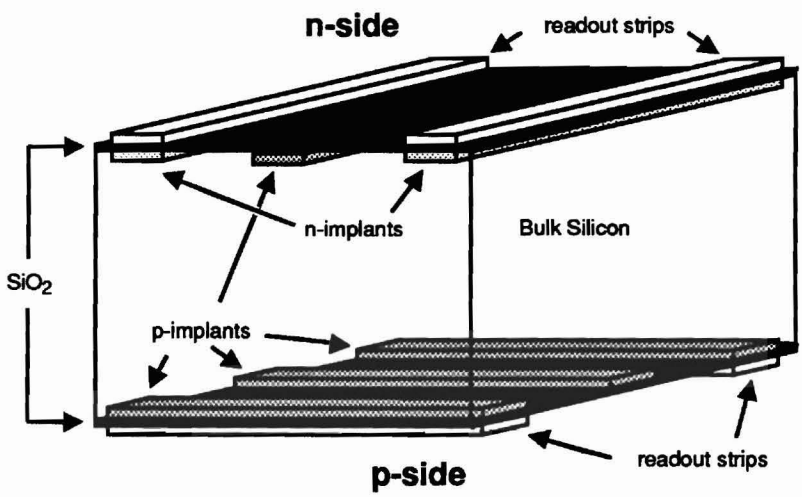

Figure A.1 Schematic of a double sided silicon detector.

accumulation layer of negative charge carriers exists at the interface of the SiO_2 and the bulk silicon which is highly conductive. This would serve to short all the strips on the n-side. To isolate the n-strips, p-type strips are deposited between the n-strips which drive away the electrons in the accumulation layer. This is also shown in Fig. A.1.

Instead of using a fan-out printed circuit board to connect the sense strips to the (rather large) preamplifiers as was done in E691, here the preamplifiers for a whole set of strips (64 to be exact) are formed together into one chip which can be bonded directly to the silicon detector. The CAMEX multiplexing amplifier, developed at the MPI-Munich[30] is used as the preamplifier for both the n- and p-strips. The rest of this appendix describes the results of studies done by the author for high signal/noise operation of the silicon microstrip detectors using the CAMEX.

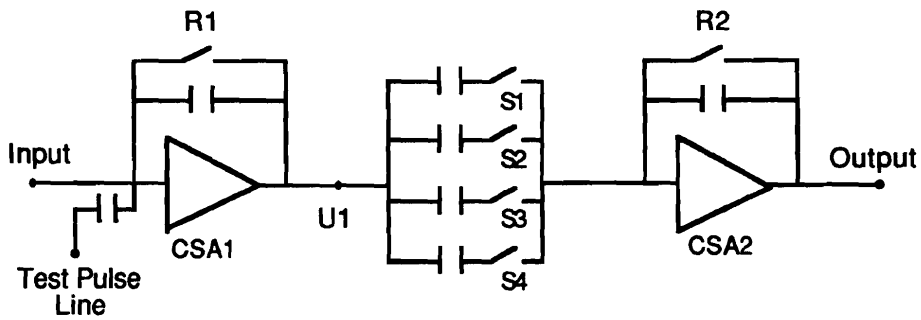

Figure A.2 Schematic of a single channel of the CAMEX.

A.1 Operation and Gain of the CAMEX

A schematic diagram of a single channel of the CAMEX is shown in Fig. A.2. The principle of operation is to sample the charge present on the sense strip four times before and after the passage of an ionizing particle and to take the difference. The series of events which allows this to happen is given here. Once $R1$ is opened, charge at the input is converted into voltage at $U1$. The switches, $S1$ through $S4$, are sequentially closed and reopened. If $R2$ is closed, the closing of each S switch will bring charge into the right side of the four capacitors to keep the voltage zero on that side. Then when $R2$ is opened, the closing of each S switch will put charge on the capacitor across the second amplifier equal to the amount of new charge that is needed at each sampling capacitor. The net result is that the output voltage, at the closing of each S switch, will be proportional to the charge difference at the input between when that S switch was closed (before $R2$ was opened) and when that S switch was closed (after $R2$ was opened). The output of each sampling capacitor is summed to produce the final output. So ideally this results in a

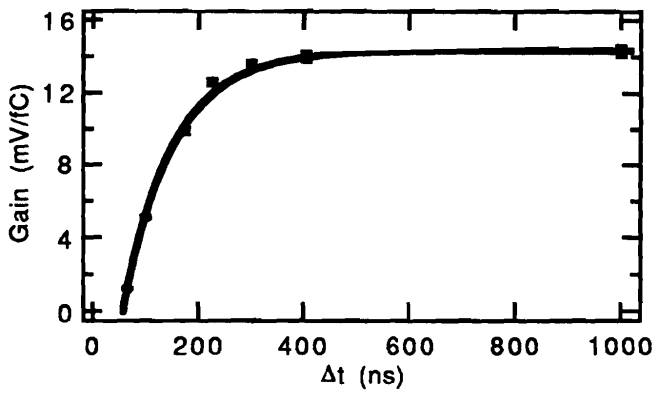

Figure A.3 Plot of Gain vs. time between charge injection and closing of the next S switch.

four fold sampling before and after the signal arrives.

In practice, however, signal can be degraded due to rise times of the amplifiers in the CAMEX. One would like to know the time it takes for charge on the input to become fully amplified at point $U1$ as well as how long the S switches must be left open in order to assure full gain. In order to measure these quantities, a precision capacitor of $18 \pm 0.5\ pF$ was bonded to the input of a CAMEX channel and a voltage pulse of known magnitude applied to this capacitor. This results in charge being injected into the CAMEX. By monitoring the relative times of the injected charge and the sequencing of the switches on the CAMEX one can obtain the desired rise times.

It was observed that the gain, or output level, depends on the time between charge injection and the closing of the first S switch. A plot of this is shown in Fig. A.3. This curve was fit to a rising exponential and a time constant of 100 ns was found. This then is effectively the rise time constant of the first amplifier.

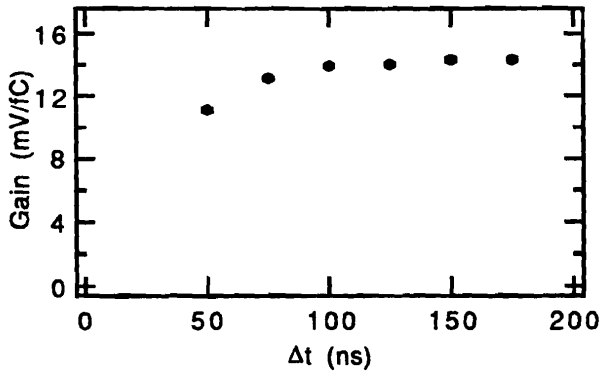

Figure A.4 Gain vs. width of S switch.

Since the values of the capacitance and applied voltage were measured, one can obtain the absolute gain of the CAMEX by the relation $V_{out} = GCV_{in}$ where G is the Gain. From the value of V_{out} when the output has risen to its maximum value one obtains a gain of $G = 14.3 \pm 0.3\ mV/fC$ for a total of four samples. For this paper, a Gain of 14 mV/fC will be assumed throughout unless explicitly stated otherwise.

It was also observed that the output level decreases if the time that the S switch is open decreases below a certain value. With the same setup as described above, the width of the S switch was varied, making sure that the switch did not open till at least 500 ns after charge was injected. The resulting plot of effective gain vs. switch open time is shown in Fig. A.4. The switches must be open for no less than 100 ns in order to achieve full gain.

A.2 Noise Studies

In general, operation of the CAMEX requires considerable care in order to minimize noise. Adding a detector only compounds the problem. This section presents the results of some studies which sought to identify sources of noise and what can be done to reduce or eliminate their effects.

A.2.1 CAMEX Without Detector

Much can be learned by studying the dependence of noise on the time between S switches. We define the sampling time of the CAMEX, τ_s, as the time between the falling edge of successive S switches. For a CAMEX without a detector connected, the noise has been found to be independent of this sampling time. The best noise value observed, using all the known noise-reducing techniques listed below, is an equivalent noise charge (ENC) of 290 e^-.* In addition to the incoherent noise stated above there is a coherent noise (common to all channels on the CAMEX) that has been subtracted in software. This noise is usually between 100 and 200 e^- (added in quadrature), again using all known noise-reducing techniques. Furthermore, both the coherent and incoherent noise are fairly uniform across all channels. Fig. A.5 (a and b) shows the total noise for each channel and the noise after subtracting the common mode component.

The most drastic of the techniques used for reducing noise is the grounding of test pulse lines at the hybrid. As shown in Fig. A.5 (a' and b') the noise

* ENC is defined as that charge, Q, on the input of the CAMEX which would result in a signal whose magnitude is numerically the same as that of the σ of the noise.

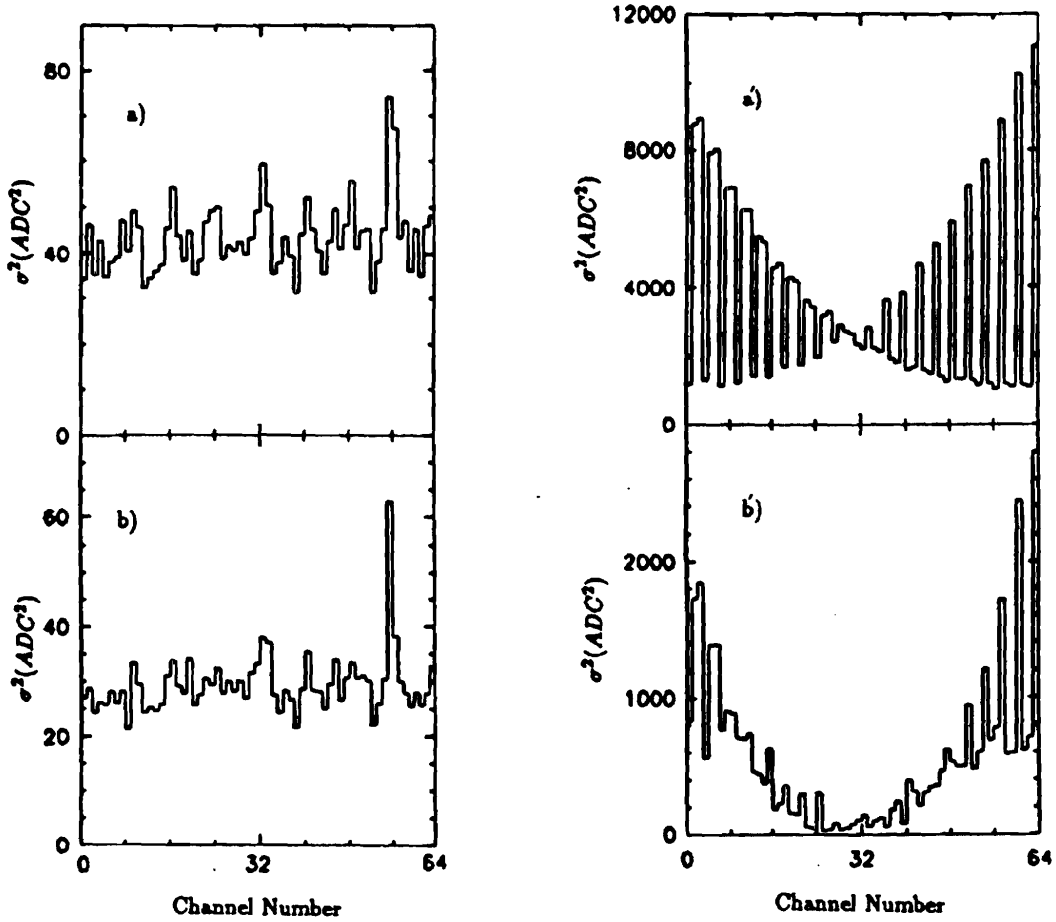

Figure A.5 Total noise (a) and noise after subtracting the common mode component (b) vs. channel number. This is with all known noise-reducing techniques being used. Total noise (a') and noise after subtracting the common mode component (b') vs. channel number. This is from a CAMEX without the testpulse lines grounded.

without the testpulse lines grounded is very large and there is a distinct pattern for every third channel. This pattern, in fact, was a clue that the testpulse lines were involved since the three testpulse lines connect every third channel and one of the testpulse lines starts at the high channel number end. The other two testpulse lines start at the low channel number end. It was discovered that charge injection effects from the S switches opening and closing were much larger than on the CAMEXes from an older batch which did not show this noise pattern. It is believed that these charge injection effects which appear as variations on the output line were being picked up by the testpulse lines along the strip cable which contains them both. The testpulse lines connect to capacitors which inject charge into the input of the CAMEX. The testpulse capacitors have been measured to be 0.63 pF so that noise on the order of a thousand electrons could result from voltage variations of only a millivolt. This problem was compounded by the fact that the output line driver on the hybrid was amplifying signals with frequencies of 1 MHz and above and the charge injection occurred at the sampling frequency which was 2.5 MHz. This latter problem was easily eliminated by shorting out the line driver, but the bulk of the noise still remained. Various grounding and shielding techniques were tried but the only way that was found that could eliminate the problem was to ground the testpulse lines at the hybrid.

The problem with grounding the testpulse lines is that one can not use them for testing gains on the CAMEX. In order to try to preserve the ability to test gains on the CAMEX a voltage divider was added to the testpulse lines at

the hybrid, with all three testpulse lines connected together. It was found that a factor of 100 in attenuation was sufficient to suppress the charge injection effects to an acceptable level. With this attenuation, an ENC of 340 e^- was obtained but with a large common mode noise of 1000 e^-. The large common mode noise appears the same as when the testpulse line pads on the CAMEX are left unbonded and therefore ungrounded. So it appears that the resistor used in the voltage divider (resistors as low as 10 Ω have been tried) leave the testpulse lines effectively ungrounded. At this point in time the effects of attenuating each test pulse line separately have not been investigated, nor have the effects of using a solid state switch.

In addition to noise resulting from pickup on the testpulse lines, a number of other sources have been identified and are listed together with their magnitude and remedy below:

1. The average pedestal level, or nominal output level of a CAMEX, changes as a function of the time between the last reset of R1 and the trigger (opening of R2). Fig. A.6 shows the average pedestal level, in electrons, for the 64 channels of a CAMEX plotted against the time difference between the reopening of R1 and the opening of R2 (arrival of trigger). One can see that for times up to about 160 μs the average pedestal level is changing by many thousands of electrons. This is probably due to charge injection effects from the reopening of R1 being propagated through the CAMEX electronics. If one operates in such a way that triggers arrive at various time intervals after resetting R1, and with at

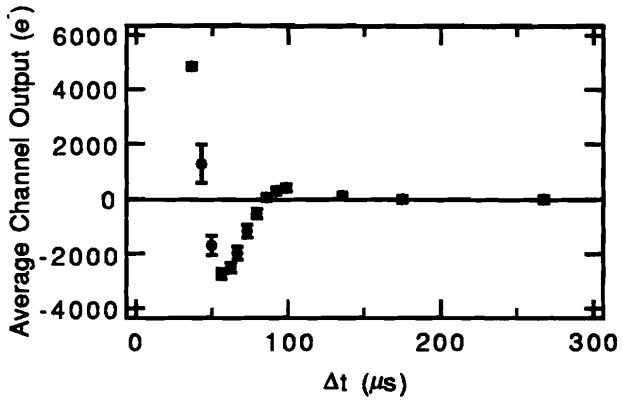

Figure A.6 Average pedestal level vs. time between opening of R1 and opening of R2.

least some triggers arriving within 160 μs, then the resulting variations in pedestals will appear as common mode noise. The magnitude of this noise could be as large as thousands of electrons. One could, in principle, eliminate this noise by operating in such a way that the trigger arrives at precisely the same interval. However, the slope of this plot is sufficiently large that one would need to insure that the interval was constant to a high degree of precision. The other alternative is to simply veto all triggers until 160 μs after the reopening of R1. It is expected that it will be possible to run in this manner at CLEO without much increase in dead time.

2. It has been found that noise, both coherent and incoherent, is reduced by 100 to 200 e^- by shielding the cable carrying all the input and output lines to the CAMEX, *and* by not connecting this shield to ground at either end. The reason for this latter part is not understood, but for the former it is believed that shielding reduces the effects of pickup

among the various input and output lines within the cable. Obviously, the amount of pickup will depend on the particular circumstances of any given setup, nevertheless, it will remain true that it is important to shield and isolate both input and output lines as well as possible.

3. There can be variations between the ground of the CAMEX and the ground of the ADC electronics. These variations will appear as noise on the digitized output. An effective way to reduce this is to amplify the output signal close to the CAMEX. This not only reduces the effects of changing grounds, but also reduces the effects of any pickup the might occur along the cable connecting the output to the ADC. For the present setup, it was found that by amplifying the output by a factor of 10, reduced the common mode noise by about 50 e^- and the incoherent noise by bout 100 e^- (in quadrature).

A.2.2 CAMEX With Detector

With the addition of a detector noise increases due to the capacitance on the input of the CAMEX as well as noise from the detector itself. In addition, it has been found necessary to filter the bias line supplying the bias voltage for the detectors with a simple RC filter. This is necessary, in that if the filter is not present then noise, both coherent and incoherent, will be present at the level of the expected signal size. What has been done is to connect the bias line to the bias pad on one side of the detector through a $1M\Omega$ resistor and then connect the ground from the power supply to the bias pad on the other side of the detector, with a capacitor across the two sides of the detector. For

the detectors and power supplies used thus far, a value of $RC \geq 5 \times 10^{-2} s$ has been found to be sufficient. It is not clear how universal this number is; one may need a larger value depending on the particular detector or power supply used. Furthermore, by measuring the voltage drop across the $1M\Omega$ resistor, one can monitor the leakage current of the detector.

The noise resulting from the capacitance that the detector adds to the input of the CAMEX is actually due to the series voltage noise from the first amplifier in the CAMEX. The capacitor turns this into charge noise at the input of the CAMEX. According to the CAMEX spec. sheet, the total noise increases roughly linearly with input capacitance with a slope of about $30\ e^-/pF$.

With the addition of a detector, the noise is also found to increase with increasing sampling period, τ_s. This was not the case for a pure CAMEX so this noise must be from the detector. A plot of noise squared versus τ_s is shown in Fig. A.7 for the p-side of the detectors.* The noise grows rapidly with τ_s so that at above 1 μs the detector noise dominates. But at 400 ns, which is close to the sampling period which is expected to be used in CLEO, this noise is minimal. Fitting this plot to a quadratic yields

$$ENC^2(\tau_s) = (1.1 \pm 0.1) \times 10^5 + (2.9 \pm 1.8) \times 10^4\, \tau_s + (4.4 \pm 0.4) \times 10^4\, \tau_s^2 \quad \text{(A.1)}$$

where τ_s is in μs and ENC is in electrons. The constant term implies an ENC of $332 \pm 15\ e^-$.

---

* The conversion between electrons and ADC units is $1\ ADC = 58.3\ e^-$.

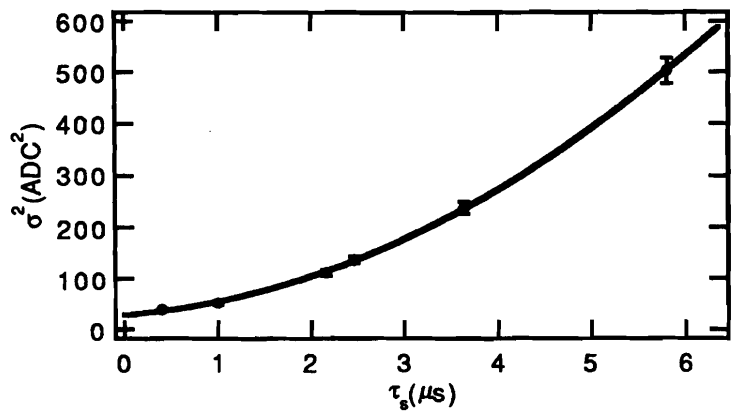

Figure A.7 Noise squared plotted as a function of the sampling period, τ_s, for the p-side of a detector. The line is a fit to the data using a second order polynomial.

For the n-side of the detectors with polysilicon bias resistors and wide p-barriers the noise squared is shown in Fig. A.8. As can be seen, the noise is approximately linear in this case with the parametrization being given by

$$ENC^2(\tau_s) = (2.7 \pm 0.3) \times 10^5 + (1.26 \pm 0.07) \times 10^5\, \tau_s. \qquad \text{(A.2)}$$

The constant term is much bigger here and translates into an ENC of $520 \pm 29\ e^-$. Attempting to understand these plots in terms of the expected sources of noise in the detector is the subject of the next section.

A.3 Noise as Information

It is important to try to understand the source of noise intrinsic to the detector. As was seen in the previous section, this noise is different on the p-side than on the n-side. If one could identify the source of this difference, one might be better able to reduce the noise. The expected sources of noise are the

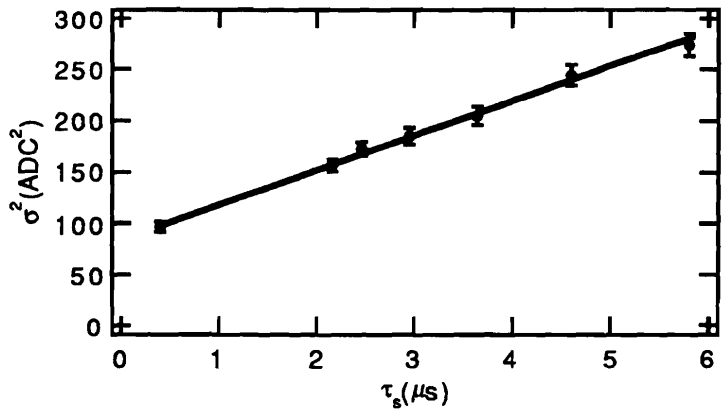

Figure A.8 Noise squared plotted as a function of the sampling period, τ_s, for the n-side of the detectors. The line is a fit to the data using a first order polynomial.

thermal or Johnson noise from the bias resistor, and shot noise associated with the leakage current. This section is devoted to the attempt at understanding the observed noise results in terms of these sources.

A.3.1 Theory

The relationship between observed noise spectrum and noise source is given by

$$\sigma^2 = \frac{1}{\pi} \int_0^\infty N(\omega)\, |H(\omega)|^2\, d\omega \tag{A.3}$$

where $N(\omega)$ is the spectral power density of the noise source and $H(\omega)$ is the transfer function of the electronics connecting the point of origin of the noise with the output.[31] As an example, consider the case of a RC circuit where the noise source is the Johnson noise of the resistor and the output is the charge on the capacitor. The Johnson noise may be considered to be a parallel source

of current with a spectral power density of[32]

$$N(\omega) = \frac{2kT}{R} = \frac{d\langle I^2 \rangle}{d\omega} \tag{A.4}$$

This noise spectrum is constant for all frequencies and is therefore also called white noise.

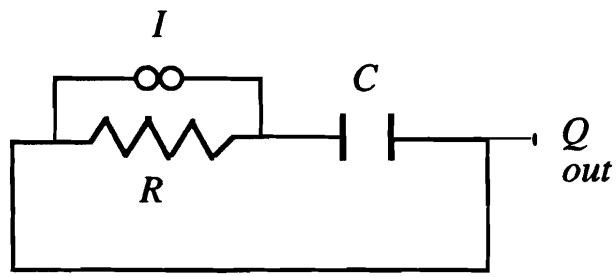

Figure A.9 Diagram of RC circuit showing thermal noise as a parallel source of current noise, I.

The equivalent circuit with current source, I, is shown in Fig. A.9. The transfer function between the current source and the output is $H(\omega) = \frac{RC}{1-i\omega RC}$. The mean square charge noise on the output is then

$$\sigma^2 = \frac{1}{\pi} \int_0^\infty \frac{2kT}{R} \frac{R^2C^2}{1+\omega^2R^2C^2} d\omega = kTC \tag{A.5}$$

The effect of the electronics, which takes the current source to the charge sensitive output, is to change the frequency shape of the noise from one that is constant to one that decreases as $\frac{1}{\omega^2}$ for $\omega \gg \frac{1}{RC}$. In this case, the total noise on the output is independent of R.

For silicon detectors, a similar circuit applies. Here, the resistor is the bias resistor and the capacitor represents the AC coupling capacitor. The charge at the output goes to the input of the CAMEX. In addition, there is a parallel capacitance which is dominated by the interstrip capacitance. This makes the transfer function $\frac{RC}{1-i\omega R(C+C_i)}$.

Shot noise from the leakage current of the detectors is also present. This noise has a white power spectrum as well, with amplitude qI where q is the charge of the electron, and I is the leakage current.[33] The total noise power spectrum is then the sum of the Johnson and shot noise power spectra.

A.3.2 Effect of CAMEX on Noise Spectrum

The CAMEX has its own effective transfer function which will modify the noise spectrum and it is necessary to know this function if one is to understand the observed noise. To obtain the total noise in an extended series of electronics one simply multiplies the absolute square of the separate transfer functions together and integrates this with the total noise spectral power density. With $H(\omega)$ being the transfer function of the CAMEX, the total mean square charge noise on the output of each channel of a CAMEX due to the Johnson and shot noise is given by

$$\sigma^2 = \frac{1}{\pi}\int_0^\infty \left(\frac{2kT}{R} + qI\right)\frac{R^2C^2}{1+\omega^2R^2(C+C_i)^2}\,|H(\omega)|^2\,d\omega \qquad \text{(A.6)}$$

The transfer function of the CAMEX may be obtained by taking the Fourier transform of the time-dependent impulse response function of the

CAMEX; this is the response of the CAMEX to a delta-like charge pulse. For an ideal CAMEX, with zero rise and fall times, the quadruple sampling produces the impulse response function

$$h(\tau) = \frac{G}{4}\Big\{\delta(\tau - \frac{\tau_s}{2}) + \delta(\tau - \frac{3\tau_s}{2}) + \delta(\tau - \frac{5\tau_s}{2}) + \delta(\tau - \frac{7\tau_s}{2}) \\ -\delta(\tau + \frac{\tau_s}{2}) - \delta(\tau + \frac{3\tau_s}{2}) - \delta(\tau + \frac{5\tau_s}{2}) - \delta(\tau + \frac{7\tau_s}{2})\Big\} \tag{A.7}$$

where τ_s is the sampling time and G is the gain for four samples, 14 mV/fC. Taking the absolute square of the Fourier transform of this function yields

$$|H(\omega)|^2 = \frac{G^2}{4}\Big[\sin(\frac{1}{2}\omega\tau_s) + \sin(\frac{3}{2}\omega\tau_s) + \sin(\frac{5}{2}\omega\tau_s) + \sin(\frac{7}{2}\omega\tau_s)\Big]^2 \tag{A.8}$$

This function, the noise weighting function of the CAMEX, is periodic and does not decrease for large ω. For high frequencies, the rise time of the CAMEX comes into play, reducing the effects of the high frequency noise. It was shown in the first section that the effective rise time constant of the CAMEX is 100 ns. One may include this effect by multiplying an additional factor of $\frac{1}{1+\omega^2\tau_R^2}$, where τ_R is the rise time constant, to the above noise weighting function.

A.3.3 Testing CAMEX Noise Weighting Function

It is important to be confident that this weighting function, and the procedure for obtaining the total noise, is correct. One piece of evidence is that the weighting function for one sample, obtained in the above manner, agrees with the one obtained in an early paper on the CAMEX.[34] Another check is to compare the noise predicted using the above weighting function to the

measured noise from a known source. What was done was to inject charge into the CAMEX with a sine wave of known frequency and amplitude. A function generator was connected to a capacitor of $C = 4.7\ pF$ which was bonded to the input of a channel on a CAMEX. Since the input is of a single frequency, the total noise on the output is simply given by

$$\sigma = \frac{G}{4}|H(\omega)|C\frac{A}{\sqrt{2}} \tag{A.9}$$

Here $C\frac{A}{\sqrt{2}}$ is the power of the input noise of angular frequency, ω, where A is the amplitude of the input voltage sine wave. This is plotted in ADC units per mV of input voltage as a function of the frequency of the sine wave in Fig. A.10. By varying the frequency of the function generator, this same pattern was observed on the output of the CAMEX. The maximum and minimum occurred at the frequencies expected. Moreover, the sigma of the output was measured for frequencies of 189 kHz and 108 kHz with amplitudes of 5 mV and 10 mV respectively. For the former case, the observed sigma was 260 and the expected value using the above weighting function is 265. For the latter frequency, the observed sigma was 198 and the expected value was 200. Taking into account the uncertainties involved (about 5% just for the capacitance) there is good agreement between the predictions of this model and the observed results.

A.3.4 Evaluation of Total Noise Integral

Confident that the weighting function is correct, let us turn to the main problem of calculating the expected noise from the Johnson and shot noise sources. At this point, this amounts to performing an integration. One can

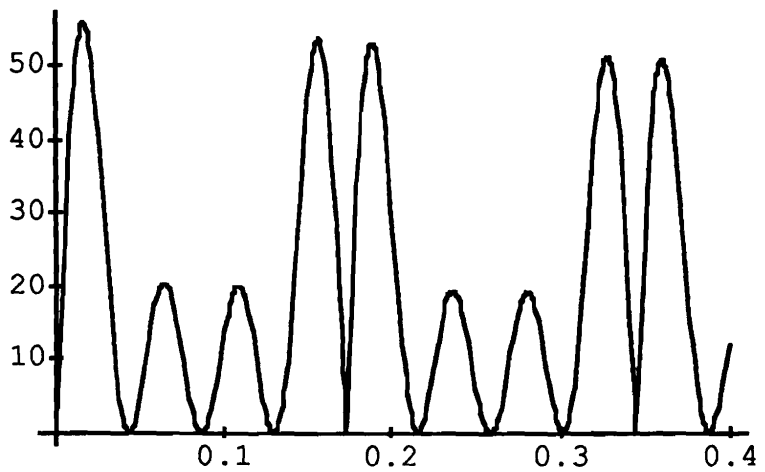

Figure A.10 Calculated noise in ADC units on the output of a CAMEX with a sampling period of $5.8\mu s$, for an input sine wave voltage of $1mV$ through a capacitor of $4.7pF$. This is plotted vs. frequency, in MHz, of the sine wave.

simplify this integral by noting that the effect of the rise time of the CAMEX can be neglected without much error as long as the sampling time is large compared to the rise time. In doing so the noise will be only underestimated, with the amount of this error increasing for smaller sampling times. Also, it is useful to rewrite the noise weighting function so that all terms are of the form, $\sin^2(x)$, using trigonometric identities. Then all terms in the integral are of the form

$$\int_0^\infty \frac{\sin^2(\frac{n}{2}\omega\tau_s)}{(\omega_l^2+\omega^2)}\,d\omega = \frac{\pi}{4}(1-e^{-n\tau_s\omega_l}) \equiv \frac{\pi}{4}f(n) \tag{A.10}$$

where $\omega_l = \frac{1}{R(C+C_i)}$. The total mean square noise is then

$$\begin{aligned}\sigma^2 =&(\frac{2kT}{R}+qI)\,\frac{G^2C^2}{16\omega_l(C+C_i)^2}\times\\ &\Big[-5f(1)-2f(2)+f(3)+4f(4)+3f(5)+2f(6)+f(7)\Big]\end{aligned} \tag{A.11}$$

Or, in terms of the equivalent noise charge, $ENC = \sigma/G$.* A plot of this function vs. $\frac{\tau_s}{R(C+C_i)}$ is shown in Fig. A.11. There is a peak at about 0.5 above which the noise decreases to a constant value. For all of the cases of interest, however, the RC value is much larger than the sampling time. This means that only the part of the curve below the peak is applicable.

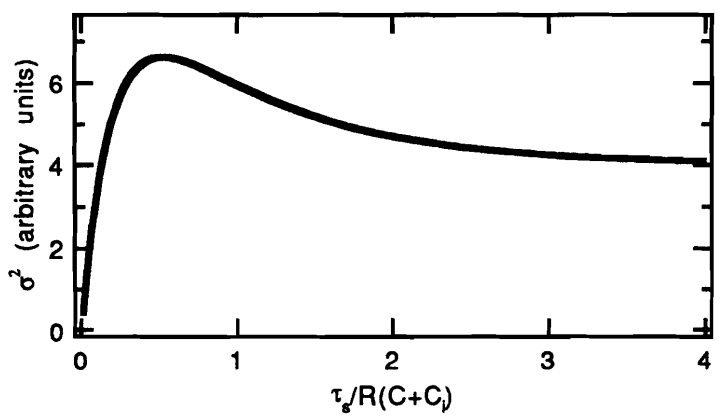

Figure A.11 Plot showing the theoretical shape of noise squared as a function of $\frac{\tau_s}{R(C+C_i)}$.

A.3.5 Comparison With P-Side Noise

It was shown in the previous subsection that the noise on the p-side grows quadratically with sampling period. This cannot be understood in terms of the expected white noise. For large RC, one has at most a linear rise in noise. Expanding out the exponentials gives $ENC^2 = (qI + \frac{2kT}{R})\frac{11C^2}{4(C+C_i)^2}\tau_s$. In units of electrons, this is $ENC(e^-)^2 = \frac{C^2}{(C+C_i)^2}[0.086I(nA)+4.4/R(m\Omega)]\times 10^{-5}\,\tau_s(\mu s)$. Comparing this formula with the fit value shown in the previous section, and

* Recall that ENC is defined as the charge, Q, such that $QG = \sigma$.

using the measured leakage current value of 1 nA per strip plus the expected value of $C/(C + C_i)$ of 0.8, yields a value for the bias resistor of 28^{+40}_{-12} $M\Omega$. This is consistent with an independently measured lower limit of 100 $M\Omega$ at the level of about 2σ because of the highly asymmetric errors. The quadratic component of the noise is not understood; nevertheless, the large RC value implies that the spectral power density of the noise source, whatever it is, must be of the $1/f$ form in order to produce a quadratic dependence.

A.3.6 Comparison With N-Side Noise

For the n-side, the noise was found to be consistent with a linear dependence on τ_s as shown in the previous section. However, the bias resistor was independently measured to be 3.5 $M\Omega$ and the capacitance at about 10 pF so that RC is only about 35 μs. This value of RC is not large enough to justify using only the linear terms in the noise function. For this reason a fit of the data was performed using the full noise function. The parameters to be found were taken to be $R(C + C_i)$, the normalization in front of the above noise function, and a constant representing the noise that is independent of τ_s. The fit is shown in Fig. A.12. The resulting value of RC was 65^{+40}_{-10} μs and for the normalization, a value of $.95 \pm .08$. Taking $C = 11.4$ pF from an independent measurement, and using the measured leakage current of 0.5 nA, one can solve for the bias resistor and $C + C_i$. The results are 5^{+6}_{-2} $M\Omega$ and 13^{+8}_{-2} pF respectively. These results are consistent with the values of 3.5 $M\Omega$ and 12 pF obtained in an independent measurement. The fit value of the constant noise term is 79 ± 5.5 which translates into an ENC of about 500 e^-. If this were due

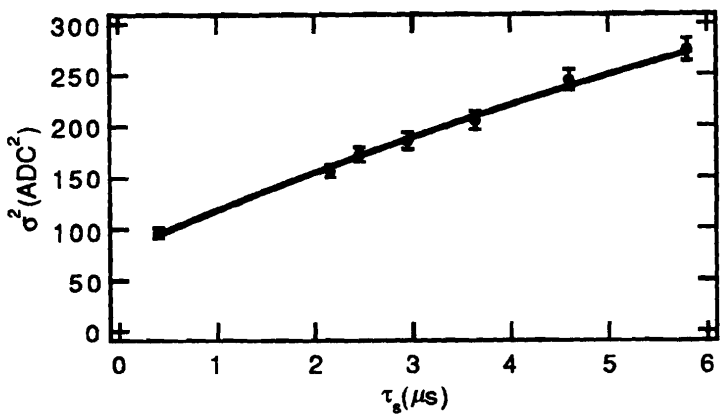

Figure A.12 Fit of noise from the n-side of a detector.

solely to the capacitance of the detector and taking $ENC = 290 + 30C(pF)$, then one obtains $C = 7.6 \pm 0.3\ pF$. This is higher than the expected value of about 1 pF. It is not clear what the source of this noise is. It could be within the detector itself, or it could be that the CAMEX used with this detector (it was from the new batch of CAMEXes) has a capacitance noise slope that is much larger than 30 e^-/pF. This is currently under investigation.

A.4 Summary of Main Noise Results

A number of techniques have been found which reduce the noise. These are outlined here together with the magnitude of their effect on the ENC:

- Have time between R2 and R1 greater than 100 μs. (1000 e^-)
- Leave the shield of strip cables unconnected. (130 e^-)
- Have an $RC > 1\ M\Omega \times 50\ nF = .05\ s$ filter on the bias line to each detector. (2000 e^-)

- Amplify the output before sending it down a long cable. This reduces pickup and the effects of ground differences between the ADC and the CAMEX. (100 e^-)

With these noise reducing techniques the ENC of a blank CAMEX is about 300 e^-. For the new batch of CAMEXes, however, the testpulse lines need to be grounded or attenuated by a factor of at least 100. With a detector connected, the noise on the p-side is about 340 e^- at a sampling time of 400 ns and grows quadratically with sampling time. This quadratic rise is indicative of a $\frac{1}{f}$ current noise source. For the n-side, the noise is about 570 e^- at a sampling time of 400 ns and grows linearly with sampling time. The magnitude of the linear rise is consistent with what would be expected from the Johnson noise associated with the bias resistor. The excess noise on the n-side compared to the p-side could be due to either unknown sources within the detector itself, or an unusually large noise vs. input capacitance slope for the new batch of CAMEXes.

REFERENCES

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

2. H. Georgi, *Weak Interactions and Modern Particle Theory*, Benjamin-Cummings, Menlo Park, CA, 1984.

3. M. Aguilar-Benitez *et al.*, *Phys. Rev.* **D45** (1992), Part 2.

4. D. Decamp *et al.*, *Phys. Lett.* **B235** (1990), 399; P. Abreu *et al.*, *Phys. Lett.* **B241** (1990), 435; B. Adeva *et al.*, *Phys. Lett.* **B237** (1990), 136; M. Z. Akrawy *et al.*, *Phys. Lett.* **B240** (1990), 497.

5. N. Isgur and M. B. Wise, *Phys. Lett.* **B232** (1989), 113, and **B237** (1990), 527; H. Georgi, *Phys. Lett.* **B240** (1990), 447.

6. M.E. Luke, *Phys. Lett.* **B252** (1990), 447; M. Neubert, *Phys. Lett.* **B264** (1991), 455.

7. N. Isgur and M. B. Wise, *Phys. Rev.* **D42** (1990), 2388.

8. J. D. Richman Cal. Tech. Preprint CALT-68-1148, 1984.

9. J. R. Raab *et al.*, *Phys. Rev.* **D37** (1988), 2391; J. R. Raab, thesis, UCSB, 1987.

10. See W. T. Eadie *et al.*, *Statistical Methods in Experimental Physics*, Elsevier, New York, 1988.

11. J. C. Anjos *et al.*, *Phys. Rev. Lett.* **62** (1989), 722.

12. M. Bauer *et al.*, *Z. Phys.* **C34** (1987), 103.

13. The radiative correction has been evaluated by adapting a calculation for $Ke3$ decays given by E. S. Ginsberg, *Phys. Rev.* **142** (1966), 1035.

14. J. Adler *et al.*, *Phys. Rev. Lett.* **60** (1988), 89.

15. J. R. Raab *et al.*, *Phys. Rev.* **D37** (1988), 2319.

16. J. C. Anjos *et al.*, *Phys. Rev. Lett.* **62** (1989), 1587.

17. J. C. Anjos *et al.*, *Phys. Rev. Lett.* **65** (1990), 2630.

18. K. Kodama *et al.*, *Phys. Lett.* **B274** (1992), 246, Carnegie-Mellon Preprint CMU-HEP92-10.

19. Presented at the "XXVI International Conference on High Energy Physics," Dallas, TX, 1992.

20. N. Isgur *et al.*, *Phys. Rev.* **D39** (1989), 799.

21. T. Altomari and L. Wolfenstein, *Phys. Rev.* **D37** (1988), 681; plus F.J. Gilman and R.L. Singleton, *Phys. Rev.* **D41** (1990), 142.

22. M. Wirbel *et al.*, *Z. Phys.* **C29** (1985), 627.

23. J.G. Korner and G.A. Schuler, *Z. Phys.* **C38** (1988), 511.

24. P. Ball *et al.*, *Phys. Rev.* **D44** (1991), 3567.

25. T.M. Aliev *et al.*, Trieste preprint IC-89-382.

26. C.W. Bernard *et al.*, *Phys. Rev.* **D45** (1992), 869.

27. V. Lubicz *et al.*, *Nucl. Phys.* **B356** (1991), 301.

28. M. Bauer and M. Wirbel, *Z. Phys.* **C42** (1989), 671.

29. D. Scora and N. Isgur, *Phys. Rev.* **D40** (1989), 1491.

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

31. E. Gatti and P. F. Manfredi, *Rivista Del Nuovo Cimento* **9** (1986), 1.

32. J. B. Johnson, *Phys. Rev.* **32** (1928), 97.

33. W. Shottky, *Ann. Physik.* **57** (1940), 244.

34. W. Butler et. al., M.P.I. Preprint MPI–PAE/Exp. El. 178, 1987.