STUDIES OF THE SEMILEPTONIC DECAYS OF D MESONS

BY

CHRISTOPHER ALAN CAWLFIELD

B.S., Principia College, 1993

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Physics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2001

Urbana, Illinois

STUDIES OF THE SEMILEPTONIC DECAYS OF D MESONS

Christopher Alan Cawlfield, Ph.D.
Department of Physics
University of Illinois at Urbana-Champaign, 2001
Jim Wiss, Advisor

We present new measurements of the branching ratios and decay distributions for the decays $D^+ \to \overline{K}^{*0} \ell^+ \nu_\ell$. We observe an asymmetry in the $\overline{K}^{*0}$ angular decay distribution that we interpret as evidence for a previously unobserved s-wave $K^-\pi^+$ amplitude that interferes with the $\overline{K}^{*0}$. These measurements are taken from the data collected by the Fermilab charm photoproduction experiment E831(FOCUS) during 1996-97.

To God, with whom "all things are possible."

Acknowledgments

This research is made possible by the inspiration and tireless dedication of all those who worked on the FOCUS experiment, an international collaborative effort of over a hundred scientists and technicians from seventeen different institutions. I could not mention all the people who have contributed vitally to this experiment, and this analysis in particular, in such a brief space. But I would like to give particular recognition to those individuals with whom I've worked most closely.

First and foremost, I thank my advisor Jim Wiss, who has spent hours each week teaching me about all aspects of the experiment and guiding this analysis steadily toward completion. Jim has worked hard to unite his small group into a cohesive team. He has earned great respect in the collaboration for this and his outstanding analytical insights.

Amir Rahimi, a fellow graduate student, has been a vital team member in our group. He has contributed in countless ways to the experiment. He and I would often bounce ideas around, and learn from each other. Our recent post-docs, Rob Gardner, Erik Gottschalk, and Doris Kim, have all played important roles in this analysis too. They have helped with my education and with the tremendous task of fine-tuning the FOCUS reconstruction and Monte Carlo to the level required for high-precision measurements such as this. Fred Cogswell, our machinist, was instrumental to the design and construction of the inner muon arrays.

I would also like to acknowledge those directly responsible for my ability and opportunity to conduct this exciting research near the high-energy frontier. My wife Heather has been a great support to me by providing encouragement, companionship, patience, and faith. My parents and parents-in-law have also been entirely supportive and enthusiastic during this venture. I would also like to thank my undergraduate professors for helping to prepare me for this work, and in particular Dr. David Cornell, whose great enthusiasm and high standards easily rub off on his students.

This research is funded in part by the U.S. Department of Energy under contract DOE F4650.2

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

The field of charm physics is currently open to a lot of interesting high-precision measurements and studies of rare decay processes.

FOCUS, or experiment E831 at Fermilab, collected data during the 1996-1997 fixed target run at Fermilab. The design goal was to be able to fully reconstruct a million charm meson decays. In this way and others, the experiment exceeded expectations. FOCUS is an upgrade from E687, which ran from 1987 through 1991. One of the more important upgrades for this analysis was a new inner muon hodoscope constructed of about 500 scintillators. I participated in the construction, commissioning and software for this array.

Other FOCUS upgrades included a new outer muon detector, four planes of target microstrip detectors, a new hadrometer, and various upgrades and improvements to existing detectors. The trigger and data acquisition systems were some of the more important upgrades.

This thesis is concerned with studying the semileptonic decay mechanism of ground state charm mesons. The ground state charm mesons are states that contain a charm quark and a light $(\bar{u}, \bar{d}, \bar{s})$ anti-quark to form a system with zero spin. Semileptonic decays have a final state involving a daughter hadron, a charged lepton and a neutrino. In this work the daughter hadron will be a spin one (vector) meson. The motivation for such studies is to provide incisive tests of Lattice Gauge theory and quark models.

There are still many important questions in field of semileptonic physics which can be resolved using the FOCUS sample. It is important to test the models used to predict semileptonic form factors since they provide the principal means to compute the hadronic matrix elements required to extract the CKM matrix elements and CP violation in the strange, charm, and beauty sectors.

The goal of this thesis is a study of Cabibbo-allowed vector-ℓ-ν semileptonic decays for charmed mesons such as $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ and $D_s^+ \to \phi\ \mu^+\nu_\mu$. Our original intention was to measure the form factor ratios r_v, r_2, and r_3 in the decays $D^+ \to \overline{K}^{*0}\ell^+\nu_\ell$ and $D_s^+ \to \phi\ \ell^+\nu_\ell$. After exhaustive efforts we found that we were unable to get acceptable quality fits to the $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ form factors. In particular, we observed a statistically significant asymmetry in the $K^-\pi^+$ angular distribution that cannot be accommodated in the model discussed in Chapter 2. This asymmetry could be a result of interference between the K^{*0} and a slowly varying s-wave amplitude. The possibility of interfering amplitudes in $D^+ \to K^-\pi^+\mu^+\nu$ is rather unexpected. In all previous experimental analyses of the decay $D^+ \to K^-\pi^+\ell^+\nu_\ell$ it has been assumed that the decay was completely dominated by $D^+ \to \overline{K}^{*0}\ell^+\nu_\ell$. Indeed there is little if any evidence in the $K^-\pi^+$ invariant mass spectrum in $D^+ \to K^-\pi^+\mu^+\nu$ decays for anything other than a single Breit-Wigner resonance with the known parameters of the $\overline{K}^{*0}$. This is true in this analysis (See Chapter 5) and in previous analyses.[3][10][17].

Chapter 2 discusses the relevant decay intensity functions. Chapter 3 describes the experimental apparatus and the history of the FOCUS data run, with an emphasis on those detectors particularly important to this analysis. Chapter 4 explains a variety of methods used to reconstruct the FOCUS data and to find clean charm semileptonic decay candidates. It also describes some important aspects of the FOCUS simulation program. Chapter 5 presents an analysis of the relative branching ratio between the $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ and $D^+ \to K^-\pi^+\pi^+$ final state. The stability of this branching ratio over a wide range of analysis cuts helps to validate the simulation as well as demonstrating the effectiveness of our methods for suppressing backgrounds and "counting" the $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ final state. Chapter 6 will present the technique used to measure the form factor ratios and assess the goodness of the fit. Chapter 7 describes the form factor fit for $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$, identifies problems with the fit, and then narrows down the possible causes for the problem. Chapter 8 presents a toy model for a broad s-wave amplitude, and shows how the interference of this with the K^{*0} produces the anomalous effects we observe in data. Final results are summarized in the last chapter.

The study of semileptonic decay is particularly interesting with the FOCUS data set, partly because we are able to reconstruct many more events with $D^+ \to \overline{K}^{*0}\ell^+\nu_\ell$ than any previous experiment. In 1998, E791 at Fermilab provided the best previous measurement of these form factor ratios. FOCUS provides roughly a factor of fourteen more events than E791 used, and a factor of nearly thirty more events than E687. This would, in principle, allow us to measure the form factor ratios with unprecedented accuracy, and also provide a better way of comparing the theoretical decay intensity distribution with observation.

1.1 Semileptonic Form Factors

Form factor measurements provide an incisive test of QCD-based calculational techniques. Semileptonic decays are more tractable since there are no final-state interactions between the leptons and the hadrons that would otherwise complicate the situation. An additional simplification is due to the relative "heaviness" of the charm quark on the the QCD scale.

All the QCD effects for semileptonic decays are contained in Q^2-dependent form factors which describe how the hadronic weak current between the parent and daughter is constructed from their four-momenta. Figure 1.1 shows how a semileptonic decay can be visualized. The coupling of the virtual W^+ to the charm quark as it transforms to a strange quark is a well-understood application of the charged weak current interaction involving the CKM matrix element V_{cs} and a V-A current interaction. The probability for the resultant $s\bar{u}$ to materialize as a K^{*0} as a function

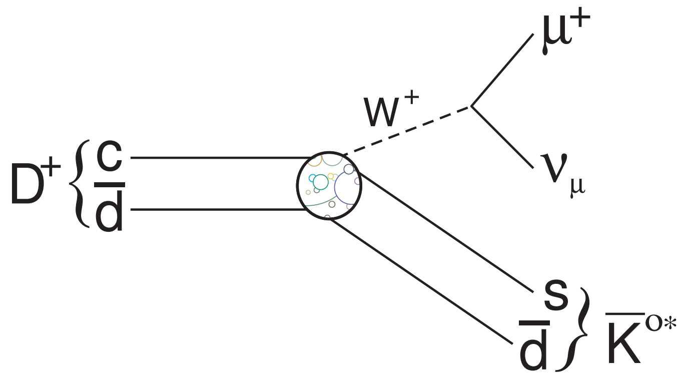

Figure 1.1: Quark-level diagram of the semileptonic decay $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$.

of momentum transfer (the squared invariant mass $\mu\nu$ system in Figure 1.1) can be computed from first principles using non-perturbative QCD methods such as Lattice Gauge Theories or through the use of phenomenological quark model wave functions.

Lorentz invariance, when applied to the case of $D^0 \to K^-\ell^+\nu$ where the daughter is a spinless particle, limits the weak current to a very simple form:

$$< D|J_\mu^\dagger|P > \;=\; f_+(Q^2)\ (P+D)_\mu + f_-(Q^2)\ (P-D)_\mu \tag{1.1}$$

Here D represents the parent charm meson's wavefunction and 4-momentum, and P is the daughter pseudoscalar.

Only the $f_+(Q^2)$ form factor plays an important role in decay intensities since all terms involving $f_-(Q^2)$ multiply the charged lepton mass which is very small. The decay $D^+ \to K^{*-}\mu^+\nu$ is also a quasi-three body decay[1]. But the fact that the $\bar{K}^*$ is a vector particle complicates the Lorentz structure of the weak current $< D^+|J_W|\bar{K}^* >$. Because of $\bar{K}^{*0}$ spin, this current is described by three dominant Q^2-dependent form factors[2] — two axial and one vector ($A_1(Q^2)$, $A_2(Q^2)$, $V(Q^2)$).

[1]Studies of the $K^-\pi^+$ mass distribution suggests that the four-body decay $D^+ \to K^-\pi^+\mu^+\nu$ is strongly dominated by the quasi-three-body process $D^+ \to K^{*-}\mu^+\nu$. By fitting the $M\,(K^-\pi^+)$ versus $M\,(K^-\pi^+\mu^+)$ for the detached events, E687[8] showed that less than 12% of $D^+ \to K^-\pi^+\mu^+\nu$ decays are other than $K^{*-}\mu^+\nu$.

[2]Experimentalists generally analyze their data assuming that the Q^2 dependence of these form factors are of the pole form with the D_s^{*+} and D_{s1}^+ poles.

The three form factors control the helicity state of the virtual W^+. This in turn affects the decay angular distribution in the strong decay $K^{*+} \to K^- \pi^+$, the weak decay $W^+ \to \mu^+ \nu$, and the angle (called "acoplanarity") between the two decay planes. The fourth and final observable, Q^2, is the squared mass of the $\mu\nu$. The decay angles are defined more precisely in Chapter 2, and illustrated in Figure 2.1.

One can measure the ratio of the form factors by fitting the observed $D^+ \to \overline{K}^{*0} \mu^+ \nu_\mu$ decay probability density, which is a function of these four kinematic observables. Traditionally one measures the ratios $r_v = V(Q^2 = 0)/A_1(Q^2 = 0)$ and $r_2 = A_2(Q^2 = 0)/A_1(Q^2 = 0)$. Once the r_v and r_2 shape variables are determined, one can measure $A_1(0)$ from the decay width since the overall decay intensity becomes proportional to $|A_1(0)|^2$. The total decay width follows from the D^+ lifetime and its branching fraction into the $D^+ \to \overline{K}^{*0} \mu^+ \nu_\mu$ final state. This branching fraction can be evaluated using the analysis presented in Chapter 5.

In charm semileptonic decay, the weak link between experimental measurements and theory is the need for experiments to assume a phenomenological Q^2 dependence to fit the data and compare to theory. We will describe this problem in detail below.

For charm semileptonic decays which represent heavy-to-light transitions, the primary theoretical tools have been Lattice Gauge Theory (LGT) and quark models.[3] In LGT[21], form factors are evaluated as the expectation value of 3-point correlation functions describing the parent, current, and daughter. These expectation values involve evaluating integrals by summation over a four-dimensional space-time lattice of size L and spacing a. Naively, computation times would scale as $(L/a)^4$, so there is a computational premium in keeping L as small as possible (but no smaller than the extent of the hadron wave function $\approx 1\ fm$), and a as large as possible (but smaller than the daughter wavelength). To minimize computational time, calculations are typically performed in the long wavelength limit where the daughter is essentially at rest in the parent rest frame. This configuration gives the largest possible Q^2 to the

[3]In b to c decays, Heavy Quark Effective Theory, provides considerable understanding of beauty semileptonic decay physics.

neutrino-lepton system. This is also true of Quark Model calculations which utilize quark wavefunctions. As the daughter momentum increases and Q^2 approaches $Q^2_{\rm min}$, the form factor probes the tail of the wave function, which is the least-accurately-known part.

Unfortunately, the observed decay intensity is peaked near the maximum daughter momentum of $Q^2 \to Q^2_{min}$ where theory has the most difficulty predicting. In order to bridge the gap between Q^2_{min} where the data is measured and Q^2_{max} where the data is predicted, most groups use the pole form *ansatz* given by Equation 1.2.

$$f(Q^2) = \frac{f(0)}{1 - Q^2/m^2_{\rm pole}} \tag{1.2}$$

Expectations are that $m_{\rm pole}$ will be the lowest mass $Q\overline{q}$ resonance with the same spin-parity as the hadronic weak current proportional to V_{Qq}. E687 and other groups have made measurements of the pole mass parameter through fits to the $D^0 \to K^-\ell^+\nu$ decay intensity, and have obtain pole masses that are roughly consistent (within 2 σ) with the mass of the D_s^{*+} which is the expected value. No attempts have been made to measure the Q^2 dependence of the $D^+ \to \overline{K}^{*0}\ell^+\nu_\ell$ form factors.

For the case of pseudoscalar-ℓ^+-ν_ℓ decays, the current of Equation 1.1 gives the differential width expression given by Equation 1.3.

$$\frac{d^2\Gamma}{dQ^2} = \frac{G_F^2|V_{Qq}|^2|\vec{h}^{(D)}|^3}{24\pi^3}\{|f_+(Q^2)|^2 + m_\ell^2\ |f_-(Q^2)|^2 ...\} \tag{1.3}$$

where $\vec{h}^{(D)}$ is the three momentum of the daughter hadron in the D rest frame. These three powers of $|\vec{h}^{(D)}|$ (one from phase space and two from the squared amplitude modulus), severely deplete the intensity near Q^2_{max}.

1.2 Previous Experimental Data

Figure 1.2 compares measurements of the r_2 and r_v form factor ratios to those predicted by quark models and lattice gauge theories for the process $D^+ \to \overline{K}^{*0}\ell^+\nu_\ell$. In general, agreement is fairly good although error bars are large. Figure 1.3 compares measurements of the r_2 and r_v form factor ratios for $D_s^+ \to \phi\ \ell^+\nu_\ell$ to both

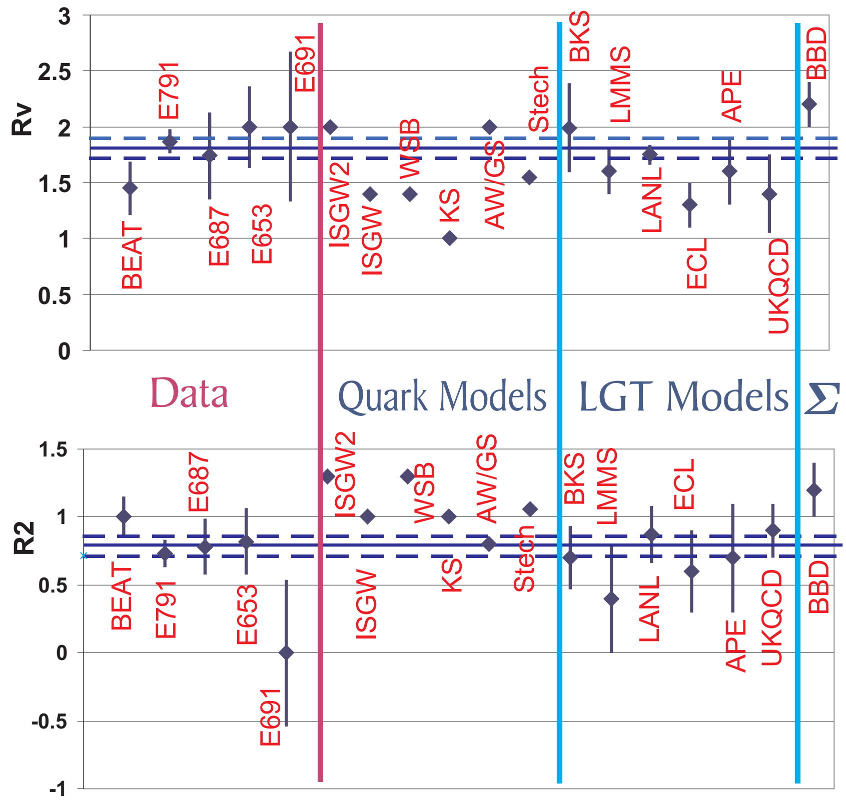

Figure 1.2: Previous measurements of r_2 and r_v from $D^+ \to \overline{K}^{*0}\ell^+\nu_\ell$. Experiments are generally consistent with theoretical predictions. [26][27]

the theoretical predictions and the form factors for $D^+ \to \overline{K}^{*0}\ell^+\nu_\ell$. We present this comparison since it is expected that the form factors for the D_s^+ decay should be equal to those for the D^+ to within about 10%. The present data seems to contradict this by about 2.5σ, although the experimental situation is far from clear given the present statistics of less than 200 events for each experiment.

In FOCUS data we can obtain a clean signal consisting of about 2800 events (See Figure 1.4) in $D_s^+ \to \phi\ \mu^+\nu_\mu$. With a comparable signal expected for $D_s^+ \to \phi\ e^+\nu_e$, we should be able to make a definitive measurement which should either confirm or refute the present experimental anomaly.

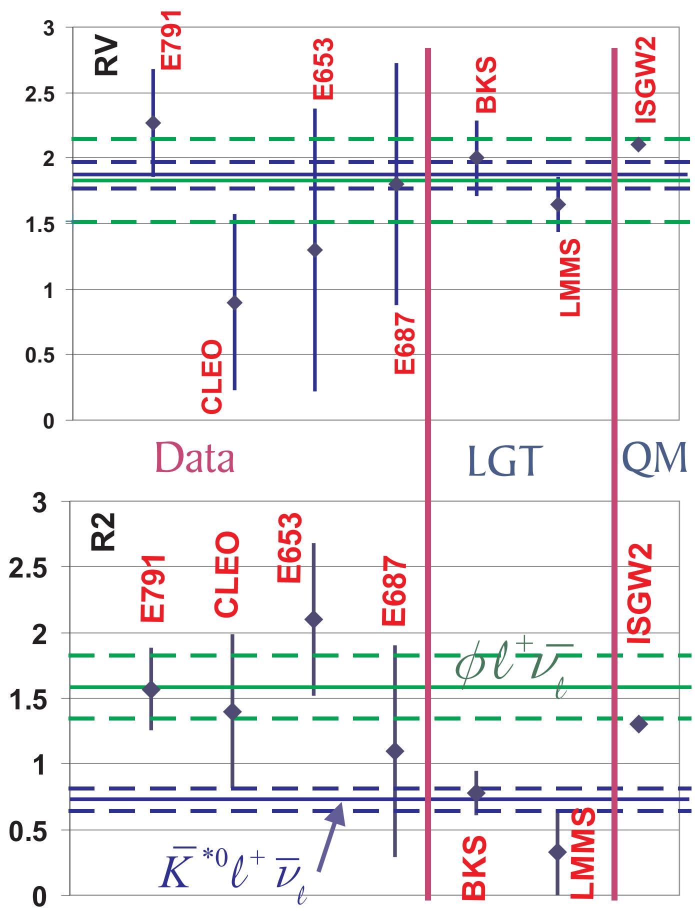

Figure 1.3: Previous measurements of r_2 and r_v from $D_s^+ \to \phi\ \ell^+ \nu_\ell$. There is a 2σ discrepancy in r_2 between the measurements and most models. We expect r_2 and r_v to be approximately the same in the two decay modes.

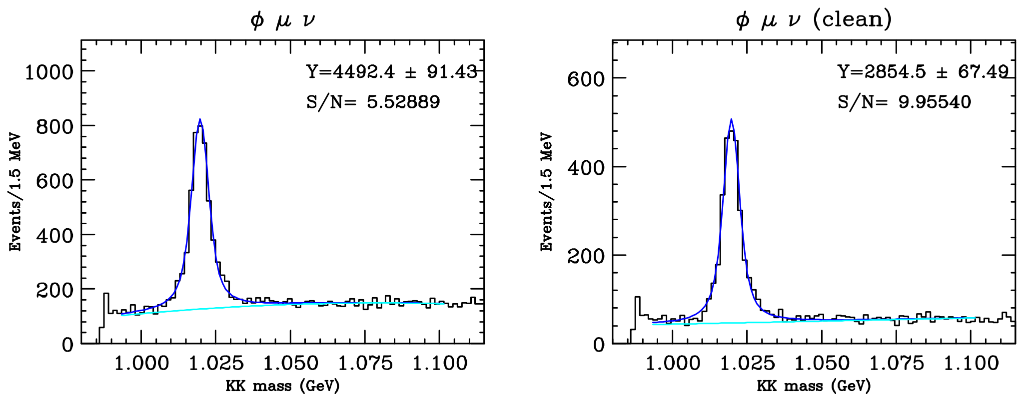

Figure 1.4: ϕ mass for $D_s^+ \to \phi\ \mu^+\nu_\mu$ candidates. The plot on the left has cuts which are about as loose as might be possible for a form factor measurement. These are: DCL>1%, $\ell/\sigma > 5$, Iso2ex<0.1, IMU only, minimum kaonicity>3, the $D \to K^-K^+\pi^+$ normalized mass is $> 2\sigma$ away from the true D^+ and D_s^+ masses, and μ-misspl<2. For the plot on the right, additional cuts are applied: DCL>10%, and CL$\mu > 0.05$

Chapter 2

Semileptonic Decay Intensity

In this chapter we present the differential decay rate for charm mesons decaying into a vector meson, a charged lepton, and its neutrino (Figure 1.1). Although this intensity applies to any decay of the form $D \to V\ell\nu$, we will present it in terms of the specific decay $D^+ \to \overline{K}^{*0}\ell^+\nu$. We begin by a simple heuristic treatment of the decay that describes the polar angular dependence of the $D^+ \to \overline{K}^{*0}\ell^+\nu_\ell$ decay intensity. These two polar angles θ_V and θ_ℓ (see Figure 2.1) describe the decay of the $\overline{K}^{*0}$ and the virtual W^+. We then follow this with an expression that frequently appears in the literature for the full decay amplitude that is differential in all five kinematic variables. This decay intensity can be factorized into phase space factors, a Breit-Wigner line shape for the $\overline{K}^{*0}$ and three types of terms that relate the decay angular dependence to helicity basis form factors. The first of this series of terms describe the polar decay angle dependences, the second describes the acoplanarity angle χ between the decay planes, and the third describe terms that vanish in the limit that the charged lepton mass approaches zero.

2.1 Heuristic Treatment

Before summarizing the formal phenomenology of the semileptonic decay distributions, we give a simplified, heuristic picture of the angular decay distribution which is

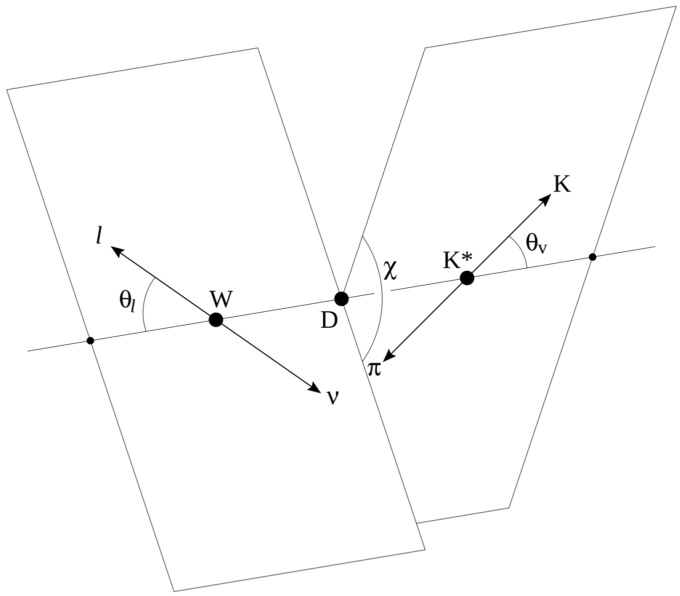

Figure 2.1: Definition of the decay angles θ_V, θ_ℓ, and χ

schematically illustrated in figure 2.2. We wish to show that one expects considerable anisotropy in the decay distribution due to simple consequences of the lefthanded nature of the weak interaction and conservation of angular momentum. For definitiveness, consider the case of the D^+ decaying to $\overline{K}^{*0}\ell^+\nu$ in the $m_\ell \to 0$ limit.

In this simple picture, the virtual W^+ that orients both the $\mu^+\nu$ and $\overline{K}^{*0} \to K^-\pi^+$ decays can exist in any of the three possible helicity states $m = -1\ ,\ 0\ ,\ \text{or}\ +1$ with a probability proportional to a width Γ_m. The helicity states $m = \pm 1$ correspond to transverse W^+'s where the W spin lies along or against its momentum vector (in the D^0 rest frame) (like the spin of a photon); the helicity state $m = 0$ corresponds to the longitudinal W^+. The W^+ momentum vector relative the D^+ will serve as the spin quantization (polar) axis.

Angular momentum conservation applied in the $\overline{K}^{*0}$ rest frame (see Figure 2.2a) tells us that this decay amplitude is proportional to the Wigner D-matrix $d^1_{0m}\left(\cos\theta_v\right)$ where θ_v is the angle of the kaon with respect to the polar axis in the $\overline{K}^{*0}$ rest frame. The fact that the kaon and pion are spinless implies that the W^+ has zero angular momentum along the $\overline{K}^{*0}$ decay axis while simultaneously having an angular

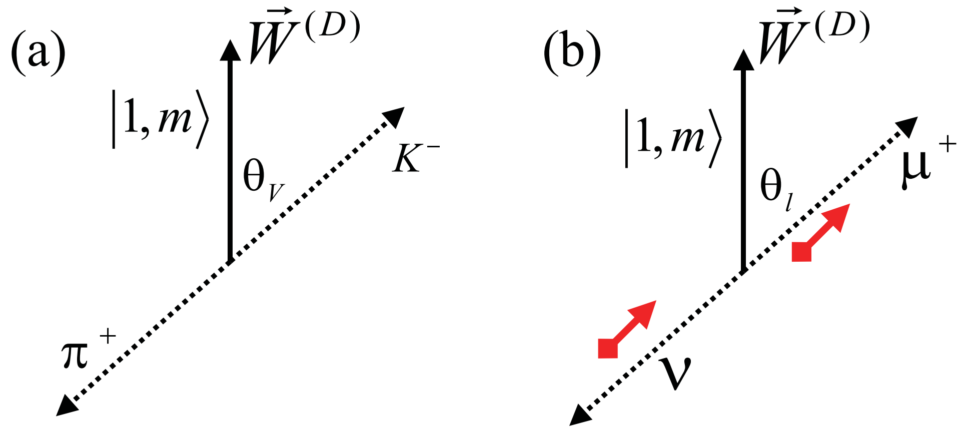

Figure 2.2: (a) We are in the $\overline{K}^{*0}$ frame. The polar axis is the momentum of the W^+ with respect to the D when viewed in this rest frame.(b) We are in the W^+ frame. The polar axis is the momentum of the W^+ with respect to the D when viewed in this rest frame.

momentum of m along the W^+ spin quantization axis. The probability for the $\overline{K}^{*0}$ to be in this spin state with respect to these two axes is just $|d^1_{0m}\ (\cos\theta_v)|^2$ We turn next to the decay of the virtual W^+. The polar axis is the spin quantization axis for the W^+. We now consider the spin state along the $\ell^+\nu$ decay axis. Since the ν must be left-handed while the μ^+ is overwhelmingly right-handed implies that the $W^+ \to \mu^+\nu$ must have a spin component of $+1$ with respect to the decay axis. The virtual W^+ decay amplitude is thus proportional to $d^1_{1m}\ (\cos\theta_\ell)$ where θ_ℓ is the angle of the μ^+ with respect to the spin quantization axis in the W rest frame (see Figure 2.2 b). We then expect:

$$\begin{aligned} \frac{d^2\Gamma}{d\cos\theta_v\ d\cos\theta_\mu} &\propto \sum_{m=-1}^{m=1} \Gamma_m \mid d^1_{1m}\ (\cos\theta_\mu)\ |^2\ |d^1_{0m}\ (\cos\theta_v)\ |^2 \\ &\propto \sin^2\theta_v\ \{(1+\cos\theta_\mu)^2\ \Gamma_+ \ +\ (1-\cos\theta_\mu)^2\ \Gamma_-\} + 4\cos^2\theta_v \sin^2\theta_\mu\ \Gamma_0 \end{aligned} \tag{2.1}$$

where we have averaged over the azimuthal angle between the $\mu^+\nu$ and $K^-\pi^+$ decay planes and ignored any possible Q^2 dependance on the Γ_m's. If all three Γ_m's were equal (not the case in reality!) there would be no net alignment effects and isotropic decay distributions would result. The degree of $\cos\theta_v$ anisotropy is directly related to the ratio of the longitudinal $\Gamma_\ell = \Gamma_0$ and transverse ($\Gamma_t = \Gamma_+ + \Gamma_-$) widths:

$$\frac{d\Gamma}{d\cos\ \theta_v} \propto 1 - \left(2\,\frac{\Gamma_\ell}{\Gamma_t} - 1\right)\ \cos^2\theta_v \tag{2.2}$$

We note in this heuristic argument that, according to Equation 2.2, we can at most have quadratic terms in $\cos\theta_v$. Equation 2.1 shows that the curvature in $\cos\theta_v$ depends heavily on $\sin^2\theta_\mu$, but there are no linear terms that would create a forward-backward asymmetry in the decay angle θ_v. On the other hand, Equation 2.1b tells us that if $\Gamma_+ \neq \Gamma_-$ one expects (and observes) forward-backward asymmetry in $\cos\theta_\mu$. We also anticipate a considerable dependence of the decay angles on Q^2 since the various "helicity" widths are Q^2-dependant.

2.2 The Complete Decay Width Expression

From the Lorentz structure of the weak currents involved in $D^+ \to \overline{K}^{*0}\ell^+\nu_\ell$, four Q^2-dependent form factors are required to fully specify the decay: V, A_1, A_2, and A_3. The form factors can only depend on the square of the virtual W^+ mass or the squared mass of the $\ell^+\nu$ system. We will call kinematic variable $t \equiv Q^2 \equiv M^2(\mu^+\nu)$ We will assume that t dependence of the three axial and the vector form factors are given by:

$$A_i(t) = \frac{A_i(0)}{1 - t/M_A^2} \qquad V(t) = \frac{V(0)}{1 - t/M_V^2} \tag{2.3}$$

The pole parameters M_V and M_A are traditionally taken as the lowest masses of bound state system $c\bar{s}$ system with vector or axial vector quantum numbers. These would be the D_s^{*+} ($M_V \simeq 2.112$ GeV/c^2) and the D_{s1}^+ ($M_A \simeq 2.536$ GeV/c^2).[2] Given this (somewhat questionable) *ansatz* for the t dependence, the shape of the differential decay width is fully specified by 3 form factor ratios (taken at $t = 0$): $r_v \equiv V(0)/A_1(0)$, $r_2 \equiv A_2(0)/A_1(0)$, and $r_3 \equiv A_3(0)/A_1(0)$. The total decay width — proportional to the absolute branching fraction $\mathcal{B}(D^+ \to \overline{K}^{*0}\ell^+\nu_\ell)$ divided by the D^+ lifetime — is then proportional to single constant $|A_1(0)|^2$.

The vector and axial vector form factors are combined (with some rather ugly kinematic factors) in a helicity basis given by Equation 2.4. The "K" appearing in Equation 2.4 is the momentum of the $\overline{K}^{*0}$ in the D^+ rest frame.

$$
\begin{aligned}
H_\pm(t) &= (M_D + M_{K\pi})A_1(t) \mp 2\frac{M_D K}{M_D + M_{K\pi}} V(t) \\
H_0(t) &= \frac{1}{2M_{K\pi}\sqrt{t}}\left[(M_D^2 - M_{K\pi}^2 - t)(M_D + M_{K\pi})A_1(t) - 4\frac{M_D^2 K^2}{M_D + M_{K\pi}} A_2(t)\right] \\
H_t(t) &= \frac{M_D K}{M_{K\pi}\sqrt{t}}\left[(M_D + M_{K\pi})A_1(t) - \frac{(M_D^2 - M_{K\pi}^2 + t)}{M_D + M_{K\pi}} A_2(t) + \frac{2t}{M_D + M_{K\pi}} A_3(t)\right]
\end{aligned}
\tag{2.4}
$$

The squared moduli of the $H_\pm$ and H_0 are essentially the helicity widths in our heuristic expression (equation 2.1) that give the t-dependent probability that the W^+ exists in each of its three possible spin states: $(|H_+|^2\ ,\ |H_0|^2\ ,\ |H_-|^2) \propto (\Gamma_+\ ,\ \Gamma_0\ ,\ \Gamma_-)$. The influence of the H_t form factor disappears as the charged lepton mass approaches zero and only contributes to what we will call "mass" terms. A term analogous to H_t does not appear in our heuristic expressions since we assume that the leptons are completely left handed — an assumption that is violated for $V - A$ weak interactions with a probability $(1 - (v/c))/2 \propto m_\ell^2$ where v is the velocity of the lepton in the D^+ rest frame.

Equation 2.5 gives an expression for the differential decay width that depends on t and the 3 decay angles θ_V, θ_ℓ, and χ defined in Figure 2.1. To our knowledge, the problem of $D^+ \to \overline{K}^{*0}\ell^+\nu_\ell$ has only been worked out in the limit of a stable $\overline{K}^{*0}$. To account for finite width of the $\overline{K}^{*0}$, it is traditional to "graft" on a Breit-Wigner factor making a width expression that is differential in $M_{K\pi}^2$ as well — for a five-fold differential width. We begin with the prefactors and overall structure of the differential width.

$$
\begin{aligned}
\frac{d^5\Gamma}{dM_{K\pi}^2\ dt\ d\cos\theta_\mathrm{V}\ d\cos\theta_\ell\ d\chi} &= G_F^2|V_{cs}|^2\ \frac{3}{2(4\pi)^5}\frac{M_{K^*}}{M_D^2 M_{K\pi}} \\
&\quad \frac{M_{K^*}\Gamma}{(M_{K\pi}^2 - M_{K^*}^2)^2 + M_{K^*}^2\Gamma^2}\ Kt\left(1 - \frac{m_\ell^2}{t}\right)^2 \\
&\quad \left[\text{diagonal terms} + \text{cross terms} + \frac{m_\ell^2}{2t}(\text{mass terms})\right]
\end{aligned}
\tag{2.5}
$$

As befits a weak process, the width is proportional to the square of the Fermi constant G_F and is proportional to the Cabibbo favored CKM coupling $|V_{cs}|^2$. We note a factor of $M_{K\pi}$ in the denominator. We believe this is a remnant of writing our differential as $d\Gamma/dM^2_{K\pi}$, whereas the traditional Breit-Wigner is differential in mass rather than squared mass: $dM^2_{K\pi}/M_{K\pi} = 2dM_{K\pi}$. We also note a power of the $\overline{K}^{*0}$ momentum (K) and two powers of $(1 - m^2\ell/t)$. The factor K and one power of $(1 - m^2\ell/t)$ represent phase space factors. The phase space factors need to be included since equation 2.5 is an expression for the decay rate rather than the squared modulus decay amplitude.[1] We begin with the diagonal terms:

$$\begin{aligned} \text{diagonal terms} \;=\; & \sin^2\theta_{\rm V} \;\Big((1+\cos\theta_\ell)^2\;|H_+(t)|^2 \;+\; (1-\cos\theta_\ell)^2\;|H_-(t)|^2\Big) \\ & + 4\cos^2\theta_{\rm V}\sin^2\theta_\ell\;|H_0(t)|^2 \end{aligned} \tag{2.6}$$

The diagonal terms give the same (acoplanarity averaged) angular decay distributions as those expected from simple angular momentum arguments of our heuristic form — Equation 2.1 with $|H|^2 \propto \Gamma$. The off-diagonal terms give the dependence of the decay width on the acoplanarity angle χ.

$$\begin{aligned} \text{cross terms} \;=\; & -2\sin^2\theta_{\rm V}\sin^2\theta_\ell\cos 2\chi\;\mathcal{R}\mathrm{e}(H_+^*H_-) \\ & - 4\sin\theta_{\rm V}\cos\theta_{\rm V}\sin\theta_\ell(1+\cos\theta_\ell)\cos\chi\;\mathcal{R}\mathrm{e}(H_+^*H_0) \\ & + 4\sin\theta_{\rm V}\cos\theta_{\rm V}\sin\theta_\ell(1-\cos\theta_\ell)\cos\chi\;\mathcal{R}\mathrm{e}(H_-^*H_0) \end{aligned} \tag{2.7}$$

These "cross" terms are proportional to trig functions of χ and disappear when averaging over the acoplanarity angle. Finally, we include "mass" terms that have no influence in the limit of zero mass charged leptons. We note from the form of Equation 2.5 the importance of the mass terms should be largest at low t. [2]

[1] Because of this distinction, one must be careful to remove these factors when simulating these decays using a phase space generator.

[2] We mention that this form for the mass terms agrees with Korner and Schuler, but differs from that of Dr. Ray Culbertson — an alumnus of the E687 Illinois group. Ray had $|M|^2 = ... + \frac{m_\ell^2}{t}$(mass terms), which is too much "mass terms" by a factor of two. Ray also had mistaken one of the mass terms, $8\cos^2\theta_{\rm V}|H_t|^2$, for $8\cos^2\theta_{\rm V}|H_0|^2$ which he merged into the second term. The

$$
\begin{aligned}
\text{mass terms} \quad = \quad & 2\sin^2\theta_{\rm V}\sin^2\theta_\ell(|H_+|^2+|H_-|^2) \\
& + 8\cos^2\theta_{\rm V}\cos^2\theta_\ell|H_0|^2 \\
& + 4\sin^2\theta_{\rm V}\sin^2\theta_\ell\cos 2\chi\ \mathcal{R}\mathrm{e}(H_+H_-^*) \\
& + 2\sin 2\theta_{\rm V}\sin 2\theta_\ell\cos\chi\ \mathcal{R}\mathrm{e}(H_+H_0^*+H_-H_0^*) \\
& + 16\cos^2\theta_{\rm V}\cos\theta_\ell\ \mathcal{R}\mathrm{e}(H_tH_0^*) \\
& + 4\sin 2\theta_{\rm V}\sin\theta_\ell\cos\chi\ \mathcal{R}\mathrm{e}(H_+H_t^*+H_-H_t^*) \\
& + 8\cos^2\theta_{\rm V}|H_t|^2
\end{aligned}
\tag{2.8}
$$

We have written these expressions for the general case of complex form factors. Our understanding is that the form factors must all be relatively real to the extent that $D^+ \to \overline{K}^{*0}\ell^+\nu_\ell$ exhibits negligible CP violation.

first mistake was important, but the second one doesn't make much difference (unless you're fitting for r_3).[3]

Chapter 3

The E831 Apparatus

In this chapter, we review the E831 apparatus and beamline, concentrating on those aspects which bear directly on the analysis presented in this thesis. The E831 experiment studies the production of charmed particles by high energy photons with wavelengths on the order of 1/500 the diameter of a proton. Photoproduction offers significant advantages in the study of charm physics over e^+e^- annihilation and hadroproduction. Charm physics is the principal goal of our experiment, although there are interesting things to learn about the photoproduction process itself which tests perturbative QCD. Photoproduced charm is relatively clean with few particles accompanying the charm-anticharm pair compared to hadroproduction. This ultimately allows a smaller data sample to reconstruct and study, because we can construct a trigger that favors hadronic events, and the events themselves are less complex. Unlike charm studies at e^+e^- annihilation, the charm particles are produced at ultra-relativistic velocities. This means that the charm decay products are produced at very forward angles and our apparatus need only subtend the forward 100 miliradians. Compared to charm hadroproduction, a relatively large fraction of the incident photon energy is given to the charm particles, so acceptance tends to be greater. The high momentum of the charm particles allows us to effectively separate them from a copious background of non-charm events[1] through their short but finite lifetime. The high momentum of the charm secondaries allows us to identify them using relatively inexpensive conventional gas threshold Čerenkov counters.

3.1 Beamline

The FOCUS beamline is unique in the world since it is high intensity, high energy, and is relatively free from neutral hadron contaminants such as K_ℓ^o's and neutrons.[2] The photon beam is classified as a *tertiary* beam since it is the result of three stages.

In the first stage, 800 GeV protons at the rate of about 5×10^{12} per 40 second spill impinge on a liquid deuterium target and produce hadronic secondaries. The charged secondaries are swept out by magnets, leaving photons (which are decay products of π^0's and η's) and other neutral hadrons. These photons then strike the photon converter, a lead sheet 60% of a radiation length thick, and converted to e^+e^- pairs. After being focused with quadrapole magnets, these electrons and positrons are bent around a dump which absorbs the remaining neutral particles. Here the beam momentum is selected by being passed through collimators. For most of the FOCUS run, the beam momentum was chosen to be $250\text{GeV}/c \pm 15\%$. This is a relatively wide momentum bite, hence the designation "Wide Band Photon Lab". On the other side of the beam dump, the electron and positron beams are recombined, and finally encounter a lead foil "radiator", which is 27% of a radiation length thick. Here bremsstrahlung photons are produced, and it is these that strike our experimental target and hopefully produce charm. The recoiling electrons and positrons are swept out by magnets and hit the RESH and POSH scintillator paddles, followed by a beam dump.

[1] As a crude guide, about 1/500 photons create hadronic final states as opposed to e^+e^- pairs and about 1% of the photohadronic cross section contains charmed particles.

[2] It is crucial to reduce hadronic contaminants, since their cross-section into hadronic final states is approximately a hundred times larger than that of photons. These hadronic interactions would create a much higher level of light hadronic backgrounds in the events.

3.1.1 Photon Beam

In the fixed target mode, the proton beam is extracted into a switchyard and there it is directed to three major areas designated as "Proton," "Neutrino", and "Meson." These designations are for historical reasons, for the Tevatron protons are sent to all three areas. The Proton area beam is split again into other beamlines, one of which is Wideband Photon beamline. To create photons, the protons interact with a liquid deuterium target. At this stage the goal is to maximize the number of hadronic interactions producing π^o's and η's. The mean free path for π^0 and η production is given by the interaction length, while the mean free path for photon absorption is given by the radiation length. The target material is chosen to maximize the radiation-length-to-interaction-length ratio by maximizing the number of nucleons relative to the number of protons:

$$\frac{\sigma_{hadronic}}{\sigma_{electromagnetic}} \propto \frac{number\ of\ protons\ and\ neutrons\ (A)}{(number\ of\ protons\ (Z))^2} \tag{3.1}$$

The photons and other neutral hadrons impinge on a 60% radiation length lead foil which creates e^+e^- pairs. Lead with a high absorption-length-to-radiation-length ratio is a good choice for this foil (called a converter), since it creates a generous number of pairs while minimizing the unwanted hadronic interactions of the neutral hadron contaminants.

The converted e^+e^- pairs are transported around the large neutral dump shown in Figure 2.3 using a beamline replete with magnets, collimators and beam tagging components which will be discussed shortly. The primary purpose for the "double-dog-leg" transport was to absorb neutral hadrons (such as neutrons, K_ℓ^o's and Λ_o's) before they could interact in our experimental target. Neutral hadrons could create severe backgrounds to the photoproduction events that the experiment was designed to study, since they have $\approx 100\times$ the interaction cross section with our nuclear target.

The collimators define a central transport momentum for electrons and positrons which ultimately specifies the end point energy of the bremsstrahlung spectrum of our photon beam. The transported beam has a nearly Gaussian momentum distribution

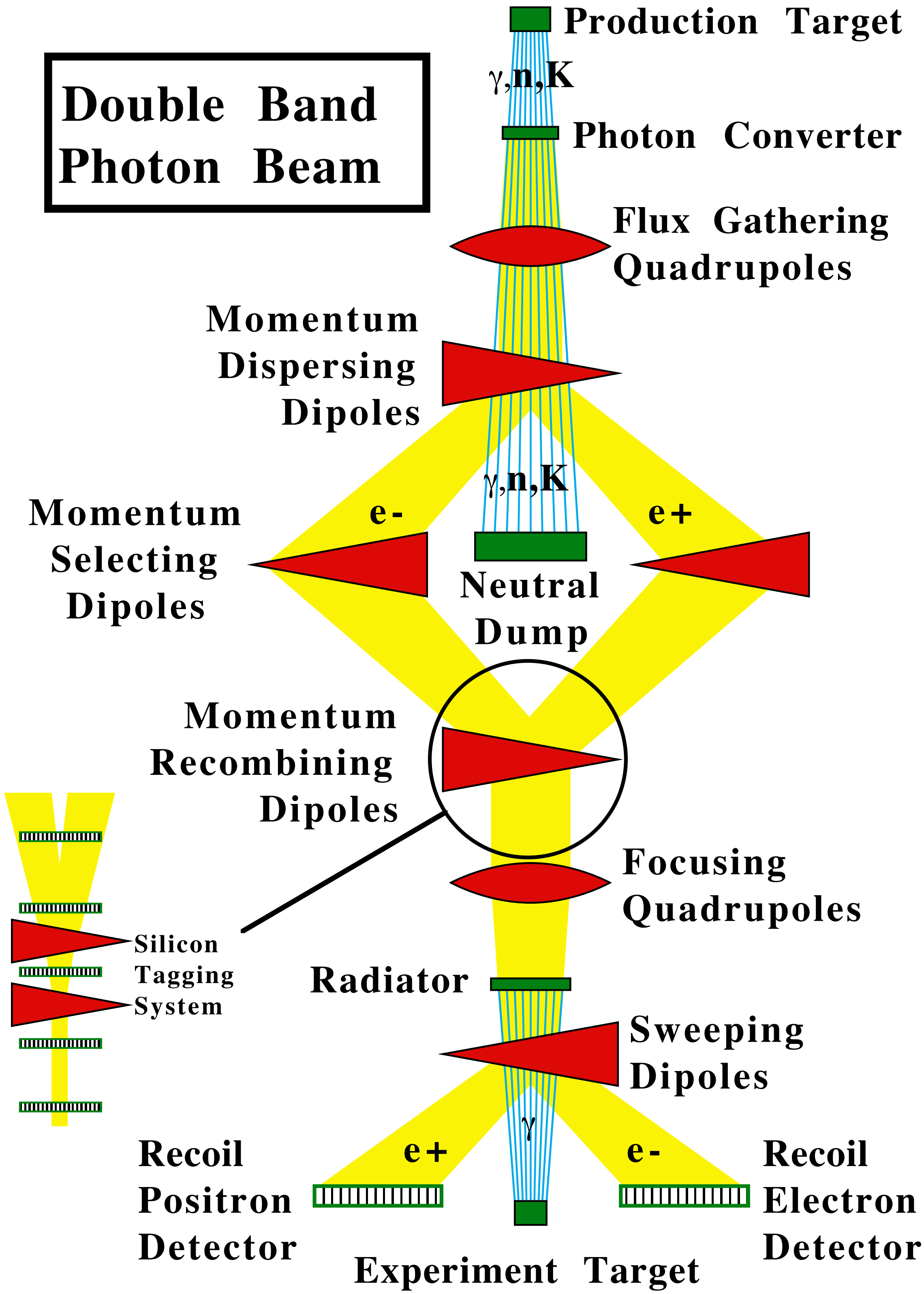

Figure 3.1: Wideband photon beamline schematic (not to scale). On the lower left hand side a more detailed schematic for the tagging system is shown.

with an RMS spread of 15% about the central value. For the bulk of the data reported here, our central transport momentum (*ie* endpoint energy) was adjusted to be ≈ 250 GeV/c. The photons which interact with our experimental target are produced as bremsstrahlung radiation from the secondary electrons and positrons as they pass through the radiator. The radiator is a 27% radiation-length lead foil. The bremsstrahlung photon spectrum is approximately of the form $dN_\gamma/dE_\gamma \propto 1/E_\gamma$ up to the endpoint energy of 250 GeV/c^2, so the bulk of the bremsstrahlung photons are soft. We designed triggers to select events from the upper end of the bremsstrahlung spectrum where our acceptance is best and the cross section for photoproduced charm particles is appreciable. Our typical triggered energy is ≈ 180 GeV.

3.2 Target

The FOCUS experiment uses a segmented target configuration. This increases the number of charm decays occurring in air. Based on the experience from E687, cleaner charm signals are obtained when the secondary vertices form outside the target.[15] We believe this is because a major background to charm are events with multiple interactions which can "fake" detached vertices. For example, a non-charm hadronic photon interaction can produce a secondary which can undergo a subsequent inelastic scattering within the target. The final state will have separated vertices just like a charm event and thus can not be eliminated with a detachment cut. But it can be eliminated by requiring that one vertex is in the gap between target segments.

The FOCUS segmented target was constructed from slabs of beryllium oxide (BeO), totalling 15% of an interaction length. Having a low average Z, BeO has a fairly large radiation-length-to-interaction-length ratio and a fairly high density. The large radiation length minimizes multiple Coulomb scattering within the target, as well as suppressing the production of e^+e^- pairs in the target. The relative high density allows us to use thin target segments. Using thin segments increases the number of charm decays which take place in the air gaps between target segments. This

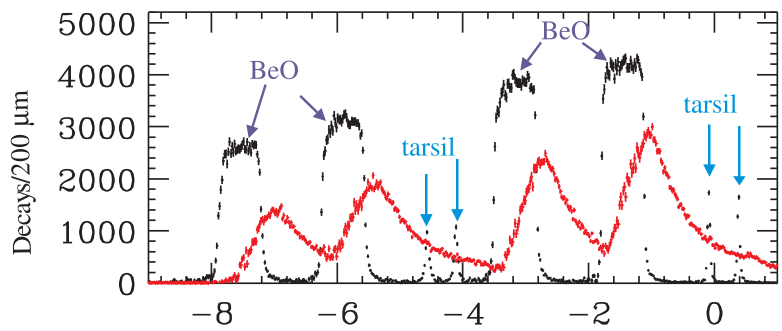

Figure 3.2: Z distribution for the primary and secondary vertices for the background subtracted Golden mode charm. The black points are the positions of the charm production vertices, and red points correspond to the charm decay vertices. The BeO segmented target and target silicon detector planes are evident. This also illustrates the way the segmented target configuration greatly increases the number of charm decays out of the target material.

is demonstrated in Figure 3.2. These "out-of-target" decays are especially clean.

Each target segment is $6.75\,mm$ thick in the beam direction and $25.4 \times 25.4(mm)^2$ in area transverse to the beam. The targets are separated by $10\,mm$. The bulk of the data were taken with this target configuration, though a different configuration was used for the first couple months.

3.3 FOCUS Detector Overview

To reconstruct charged particles, FOCUS uses three general classes of detectors: those which together resolve vertex and line-of-flight information, those which measure momentum and charge, and those which determine the rest mass through particle identification. To reconstruct neutral particles, various calorimeters are used.

FOCUS uses silicon strip detectors (SSD's) in order to resolve particle production and decay vertices, and measure the particle line-of-flight. The other major tracking system in FOCUS is the proportional wire chambers (PWCs). The PWC's allow us to measure the bend in particle trajectories through the two analysis magnets. By linking a track found in the microstrips (having accurate trajectory information) with a track found in the PWC system (having well-measured momentum and charge), we can reconstruct the 4-momentum of the particle. To determine the particle identity, we use Čerenkov detectors and muon detectors.

3.4 Silicon Strip Detectors

The SSD detectors provide a spatial resolution of roughly $7 \rightarrow 14\mu m$ (depending on the plane). They are composed of separate readout strips which collect charge that are freed when ionizing particles traverse through it. The thickness of the detector is usually on the order of $300\,\mu m$. For minimum-ionizing particles, average energy loss in the silicon is about $39\,\mathrm{KeV}c^{-2}/100\,\mu m$ so that about 100 electron-ion pairs/μm are created [16]. Therefore, about 30,000 ion pairs are created in the detector. This implies both high efficiency and a good signal-to-noise ratio to suppress Johnson noise. Each silicon strip is pulse-height analyzed.

Figure 3.4 shows two silicon detector systems, the "SSD" or silicon microstip system, and the "TSSD" or target microstrip system. The microstrip system is more complete in terms of the number of stations and redundancy of views, and it is used to find tracks. The target microstrip system is a later addition, and is used primarily to refine the tracking information.

3.4.1 Target Silicon Detector

Four planes of silicon microstrips are installed in two stations of two planes (X and Y) each, to help improve the vertex resolution.[3] The components of the target silicon consist of four planes of 25μ-pitch 300μ-thick planes of silicon microstrip detector. Each silicon plane has 1024 channels, or about an inch of active area, which covers the majority of the photon beam size at the E831 target. As shown in Figure 1, these target silicon planes were interleaved between the experimental target segments. The target silicon information is present for roughly two thirds of the FOCUS data, as it

[3] The target silicon provides a hit coordinate close to the vertex position before a substantial lever arm for multiple Coulomb scattering gets a chance to develop.

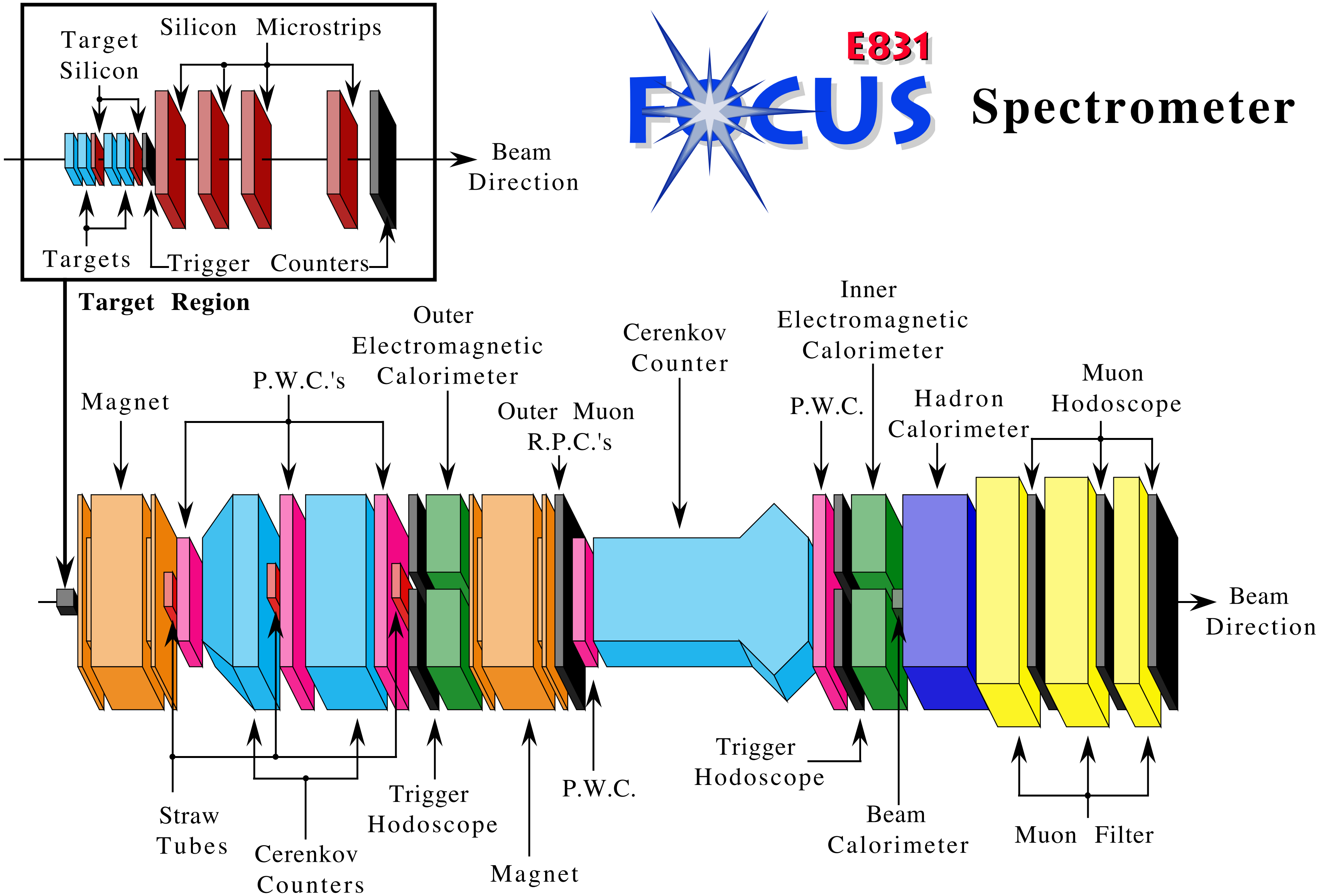

Figure 3.3: FOCUS spectrometer. In the upper left, the target region is shown.

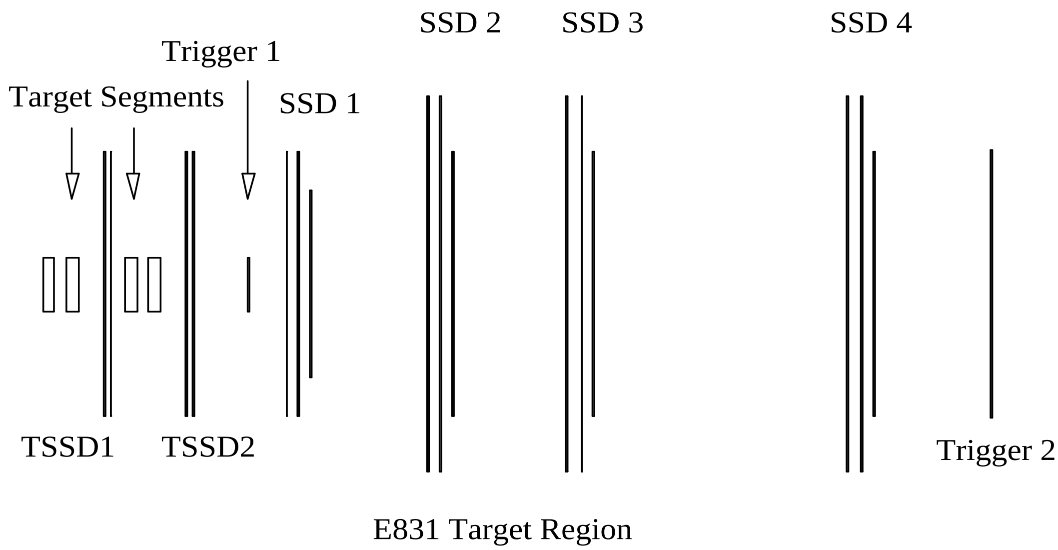

Figure 3.4: Target region and SSD configuration. Target silicon detectors (TSSD) help increase vertex resolution. Trigger 1 (TR1) and Trigger 2 (TR2) are scintillator counters used in the first level trigger.

was installed part-way through the run.

3.4.2 Microstrip Detector

The microstrip detector is located about 5 cm downstream of the most downstream experimental target segment. By measuring the trajectories of charged particles with high precision, one can exploit the lifetime of charm particles (typically 1 cm decay length for D meson) to identify charm signals. Referring to the diagram in Figure 3.4, the microstrip system consists of twelve planes arranged into four stations of three planes each. In each station, one plane measures in the y-direction and two planes are titled in opposite direction from the y-axis by 45^o. Each plane is divided into an inner high-resolution region and an outer lower-resolution region. The strips in the planes of the most upstream station have 25μ pitches in the inner region and 50μ pitches in the outer region. The remaining stations have 50μ and 100μ pitches, respectively. The SSD is described in Table 3.1.

Station	Total Active Area	High Res. Active Area	Strip Pitch
1	$2.5 \times 3.5cm^2$	$1.0 \times 3.5cm^2$	$25\,\mu m$, $50, \mu m$
2	$5.0 \times 5.0cm^2$	$2.0 \times 5.0cm^2$	$50\,\mu m$, $100, \mu m$
3	$5.0 \times 5.0cm^2$	$2.0 \times 5.0cm^2$	$50\,\mu m$, $100, \mu m$
4	$5.0 \times 5.0cm^2$	$2.0 \times 5.0cm^2$	$50\,\mu m$, $100, \mu m$

Table 3.1: SSD specifications. Each station consist of 3 planes in y, u and v views. The u and v views are titled in opposite direction from the y-axis by 45^o.

3.5 The Analysis Magnets

Momentum analysis of charged particles is performed by measuring the deflection in the fields of two dipole magnets M1 and M2. These magnets are used to deflect charge tracks in opposite directions. The transverse momentum kick is 0.4 GeV/c for M1 and 0.85 GeV/c for M2. The "kick" of a magnet is a measurement of the integral of the B-field such that the angular deflection of a charged particle traversing the magnet is given by the kick over momentum (to leading order in 1/momentum). The ratio of the transverse kicks is such that the tracks come back to their original undeflected position toward the downstream end of the spectrometer. This ratio was chosen so that the e^+e^- pairs converted in the experimental target (which are nearly parallel to the beam direction) would tend to focus toward the central electromagnetic calorimeter hole rather than creating additional showers and radiation damage for the lead glass blocks located close to the central median plane.

3.6 Multiwire Proportional Chambers (MWPC's)

The multiwire proportional chamber system consists of twenty signal planes grouped into five stations with four planes per station. Each station has four views. The X view wires run vertically and measured horizontal position. The stereo angle for U and V views is 11.3^o from the Y view. The orientation of the wires is shown in Figure 3.5. This arrangement concentrates the information in the vertical readout

Properties	P0	P1	P2	P3	P4
Aperture (in^2)	30 x 50	60 x 90	60 x 90	30 x 50	60 x 90
Wire Spacing (mm)	2.0	3.0	3.0	2.0	3.3
No. X-view Wires	376	480	480	376	480
No. Y-view Wires	640	704	704	640	704
No. U/V-view Wires	640	768	768	640	768
Gas Used	Argon-Ethane(65/35) Bubbled through 0^o C ethyl alcohol				
Voltage	2.80-3.20 kilovolts				

Table 3.2: PWC specifications.

direction, because it is the vertical change in a particle's trajectory that determines its momentum. The chambers P0, P1 and P2 are located between the two analysis magnets, and P3 and P4 are located downstream of the second analysis magnet. P0 and P3 are smaller than the other chambers because they are located directly downstream of the the analysis magnets M1 and M2. These magnets constrict the aperture of the P0–P2 chambers, and the P3–P4 chambers. Particle traces which pass through both M1 and M2 are called "tracks", while particles passing through M1 but not M2 are called "stubs." The specifications of the chambers are summarized in Table 3.2.

3.7 The Straw Tube System

Three straw tube wire chambers are used to measure tracks in the high-intensity "pair region" — a vertical stripe down the center of the spectrometer — where the conventional MWPC's were deadened. Each chamber has three views, and within each view there are three nested straw layers. All three chambers have a vertical (X-measuring) view and two views angled $\pm 11.33^o$ from vertical. The straw tube system is also capable of providing timing information, since the signals are sent into TDCs. The information from this system has yet to be used to its full potential.

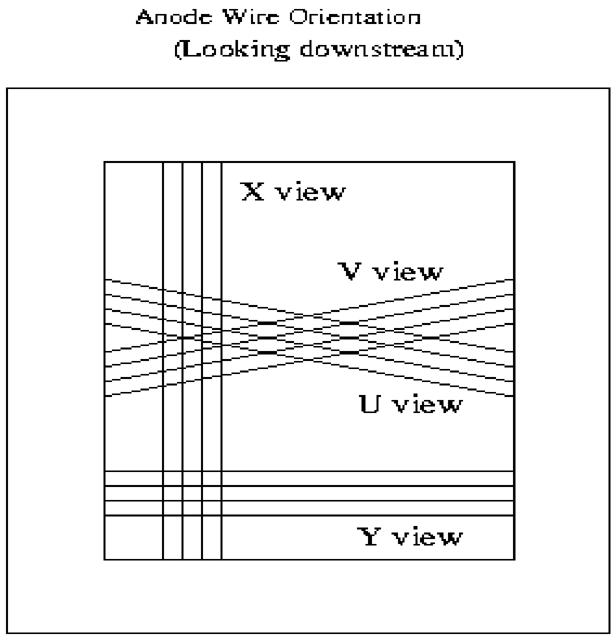

Figure 3.5: PWC orientation looking downstream. There views u and v are tilted 11.3^o with respect to the horizontal.

3.8 Čerenkov System

Čerenkov radiation in a medium occurs when a charged particle travels faster than speed of light in that medium. The speed of light in the medium is given by:

$$\beta = \frac{P}{E} >? \frac{1}{N} \tag{3.2}$$

where N is the index of refraction of the medium. Therefore, a particle emits Čerenkov light when its speed exceeds c/N.

There are three multi-cell threshold Čerenkov detectors in the experiment, referred to as C1, C2 and C3. The detectors are operated at atmospheric pressure and in the threshold mode. The gases are chosen so that the different indices of refraction (i.e. different light velocities) establish different momenta in which pions, kaons, and protons will begin to radiate Čerenkov light (see Table 3.3).

For convenience, we summarize a few useful formulae for Čerenkov counters using

Table 3.3: Čerenkov counter specifications. The momentum threshold for the three charged particles are give for each counter.

counter	Gas	Threshold (GeV/c)			No. of Cells	Ave. Number of Photoelectrons
		pion	kaon	proton		
C2	N_2O	4.5	15.9	30.2	110	$8-11$
C1	$He-N_2$	8.4	29.7	56.5	90	$2.5-3.6$
C3	He	17.4	61.5	117	100	9

a gas with index of refraction $N \equiv 1+\delta$ and the relevant approximate forms when $\delta \ll 1$. The minimum (threshold) momentum for a particle of mass m to emit Čerenkov light is:

$$P_{\text{Threshold}} = \frac{m}{\sqrt{1-1/N^2}} \approx \frac{m}{\sqrt{2\delta}}$$

and the angle of emitted Čerenkov radiation is :

$$\theta_C = \cos^{-1}\frac{1}{N\beta} \approx \frac{m}{P_{\text{Threshold}}}\sqrt{1-\left(\frac{P_{\text{Threshold}}}{P}\right)^2}$$

Finally, the number of Čerenkov photons produced per unit radiator length is proportional $\sin^2(\theta_c)$ which approaches $(m/P_{\text{Threshold}})^2$ as $P \gg P_{\text{Threshold}}$. This requires the high threshold counters to be very long. Many of the relevant Čerenkov properties can be computed by specifying the *pion* threshold or $P_{\text{Threshold}}$ for particles of mass $m = m_\pi \approx 0.140 GeV/C^2$ and the number of photoelectrons recorded for a $\beta = 1$ track. For our system, the three pion thresholds were chosen to be 4.5, 8.4 and 17.4 GeV/c by use of appropriate gas mixtures, and the photo-electron yield ranged from roughly 2.5 to 20 depending on the phototube and Čerenkov counter.

Our Čerenkov algorithm (described in depth in the next chapter) exploits the threshold behavior, size of the Čerenkov cone and number of Čerenkov photons to identify charged particles. Some of the Čerenkov counter specifications are summarized in Table 3.3

3.9 Calorimetry

The inner and outer electromagnetic calorimeters, OE and IE, are not directly used in this analysis because I chose not to study electrons. The hadron calorimeter and inner electromagnetic calorimeter are used in the trigger. They are also major contributors to mulciple Coulomb scattering for muons, so they needed to be modeled for the muon identification and calibration algorithms.

3.9.1 Outer Electromagnetic Calorimeter (OE)

The OE measures the wide angle electromagnetic particles (photons and $e^{\pm}$) that pass outside the acceptance of the second analysis magnet. It is located just downstream of the M2 magnet. The OE is a sampling calorimeter with alternating layers of lead and plastic scintillator, and it consists of x, y and z planes. There are 100 scintillator tiles per plane. The scintillation light in each tile is carried out through fiber optic cables, and directed into photomultiplier tubes.

3.9.2 Inner Electromagnetic Calorimeter (IE)

The IE measures the electromagnetic particles that pass through the second analysis magnet. It is composed of lead glass blocks arranged in a tower geometry. There are two sides to the detector with a central gap. The gap allows the passage of the intense beam of non-interacting photons and e^+e^- pairs. Again, the IE only affected this analysis due to its role in the trigger.

3.9.3 Hadron Calorimeter (HC)

The hadronic calorimeter is constructed with 18 alternate layers of 4.4 cm-thick iron plates and 0.7 mm-thick scintillating tile planes using fiber readout in tower geometry, with smaller tiles in the center and larger ones at the outside region. The layers are arranged into three sections, with the first section having nine layers, the second section having 15 layers and the third section having 4 layers (Figure 3.6). This

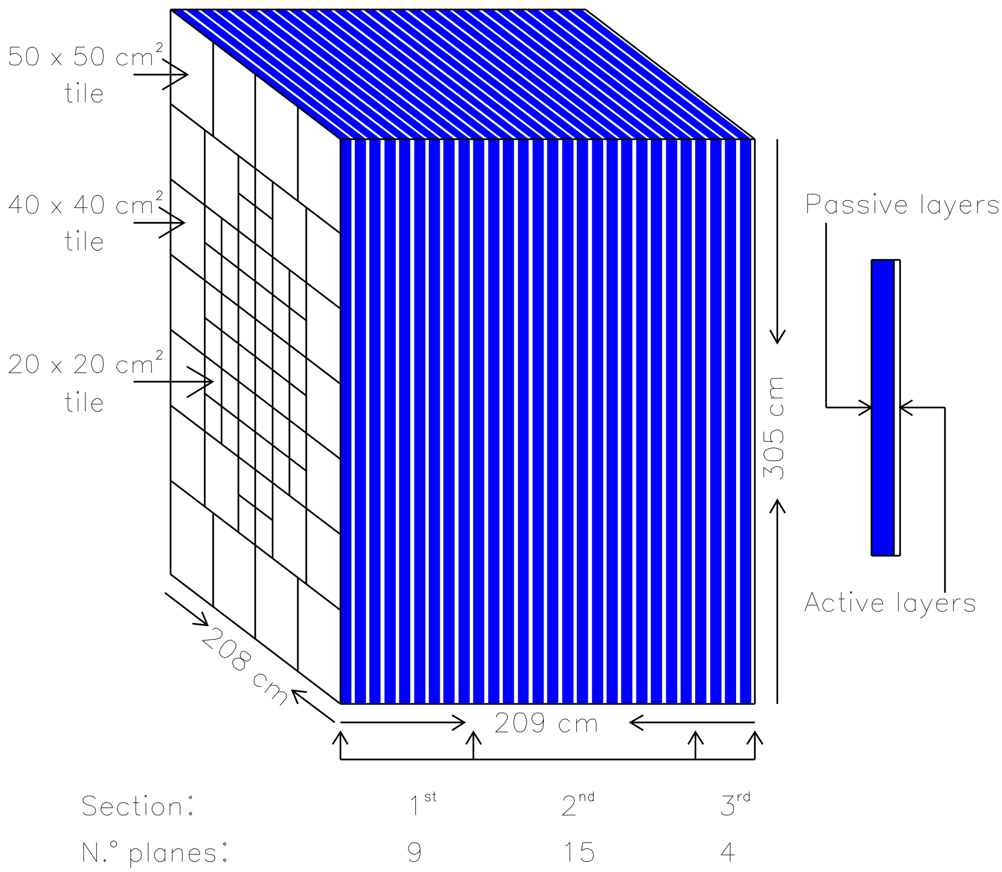

Figure 3.6: HC schematic .

scheme allows for deposited energy to be measured as a function of depth. The scintillators have fast response, enabling the HC to be included in the first level trigger. The HC is useful for triggering events with hadronic final states rather than $e^{\pm}$.

3.10 Muon Detector

Muons do not suffer significant energy loss due to electromagnetic showering because their mass is larger than the electron. They do not undergo hadronic absorption because they don't interact strongly. Therefore, muon detectors are placed downstream of large iron blocks which act as filters to filter out electrons and hadrons and let the muons through. There are two sets of muon detectors. The Outer Muon detector uses a Resistive Plate Chamber system to detect high angle muons. The Inner Muon detector consista of three stations made of scintillating counter arrays called MH1, MH2, and MH3. The MH1 and MH2 stations consist of x and y views and MH3 consist of u and v views. Figure 3.7 shows one view from each station. MH1 and MH2 were built, designed, and commissioned by the University of Illinois group, a project in which I was actively involved.

3.11 Inner Muon Arrays

Figure 3.7 shows the layout of the three inner muon stations. The individual detectors in these arrays are designed to overlap each other for complete coverage. They are shielded from delta-rays by 1/4"-thick aluminum plates mounted directly on the upstream sides of the detectors.

The signal cables for the muon detectors were salvaged from the E687 electromagnetic calorimeter. These cables were repaired, and their lengths were individually measured by "pinging" them with a pulse generator and oscilloscope, set up to measure the round-trip time of voltage pulses. Cables of similar lengths were grouped

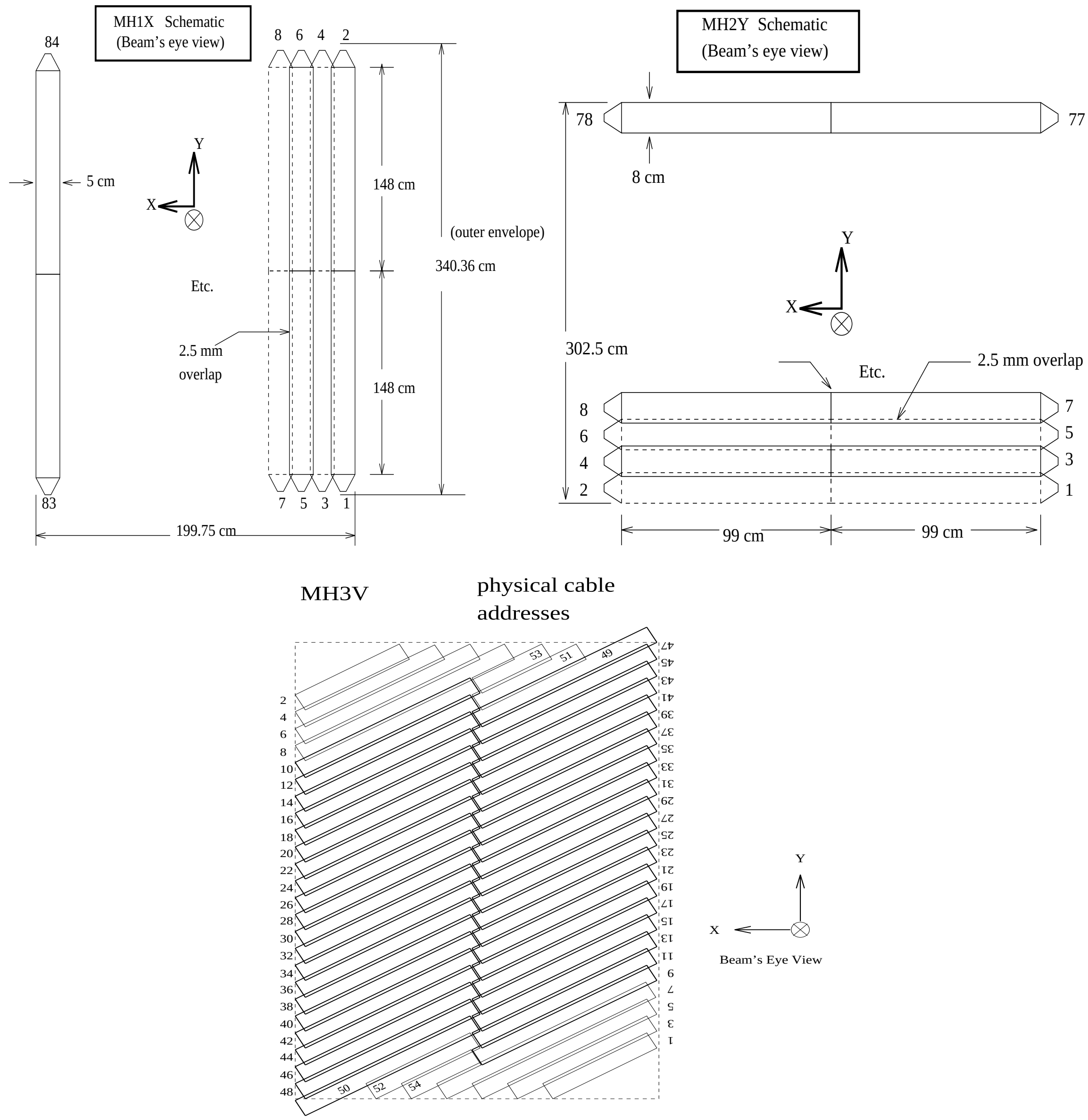

Figure 3.7: Inner Muon array schematics. The schematics for MH1Y, MH2X and MH3U are similar to MH1X, MH2Y and MH3V respectively. The MH3 detectors are each 10 cm wide.

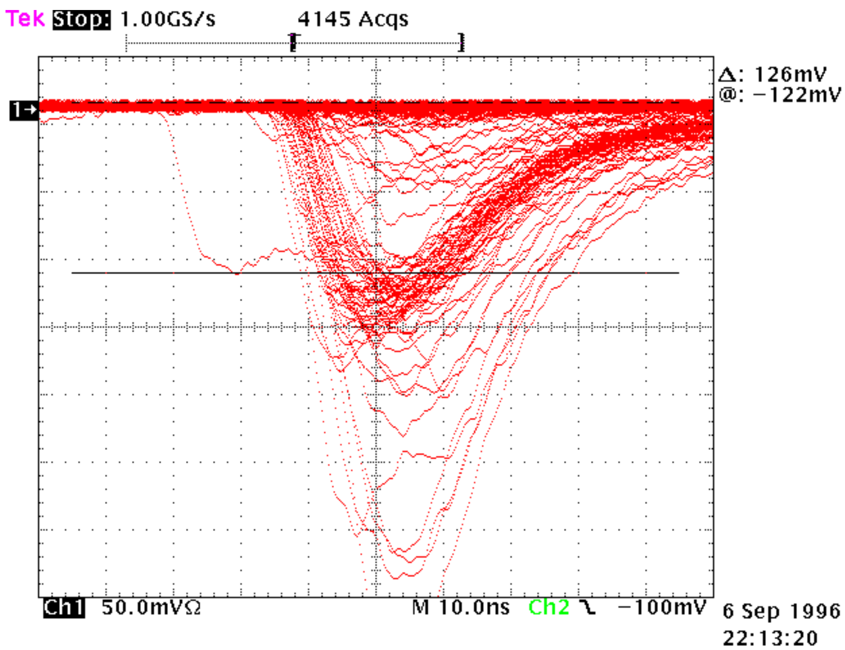

Figure 3.8: Pulses from muons in the inner muon detectors. This is a collection of electronic pulses at the input to the discriminators. There is a main clustering of pulses which represent genuine muon hits. The average height of these pulses were measured and used to tune the high voltage setting for the given PMT. For this detector, the minimum pulse height was roughly 126 mV. The discriminator voltage would have been set to about half this value.

together in bundles, and ran from the detectors into the counting room.

Initial voltages were adopted based on source tests of the detectors, but these were refined by measuring the pulse heights during muon test-beam running. Figure 3.8 shows an example of this measurement process.

The signals from the inner muon PMT's went into discriminators, and then into latches with programmable delays. One common delay is used for each module holding 32 channels. At two or three times during the run, the inner muon array timing was adjusted by stepping through various delays in each latch module and constructing a plot of efficiency versus delay. The efficiency function looked like a plateau, and the center of the plateau was chosen for each latch module. Figure 3.9 shows an example timing plateau for a certain module.

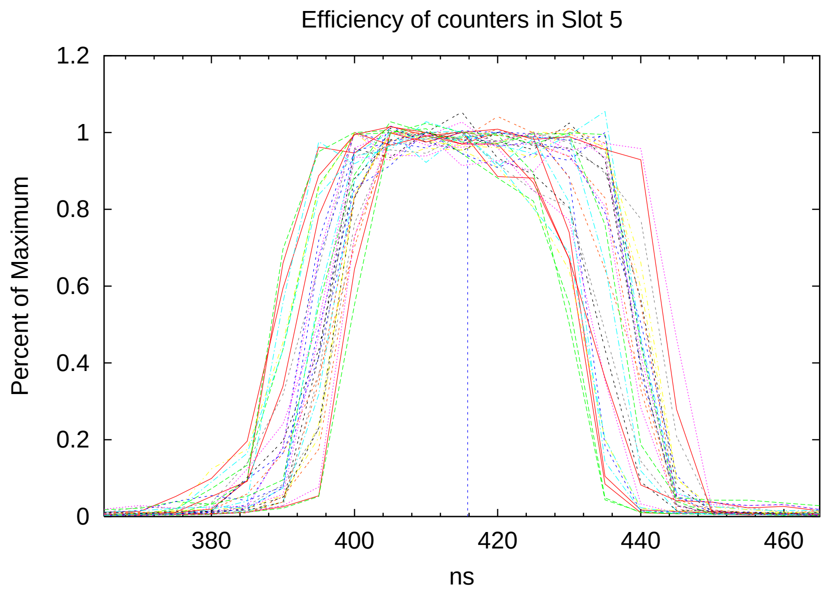

Figure 3.9: Inner muon timing efficiency plateau. The efficiency is a crude estimate based on the number of firings per event, normalized to the peak of each plateau. The dashed line represents the optimal value for the PCOS latch delay. The height of the plateau is approximately 40 ns, which is the width of the latch gate. The different colors represent the 32 channels going into this module.

3.12 Trigger

The electronic logic that is required to select the interesting events from the background is called the trigger. FOCUS typically had about 10^8 interactions per spill (mostly electromagnetic) and triggered on about 30,000 interactions per spill (mostly hadronic). It is known that in high-energy photoproduction interactions, the hadronic interaction rate is about 1/500 of the rate for pair production. The $e^{\pm}$ pairs generally are produced at a small transverse momentum, thus very low angle trajectories, whereas the hadronic interactions produce particles with larger transverse momentum and wider trajectories. The hadrons also deposit larger energies in the hadronic calorimeter. Therefore, the main purpose of the trigger is to trigger on the events with wide-angle tracks and significant energy deposited in the HC.

The first level of main hadronic trigger is called the "Master Gate" (MG). The MG trigger occurs within $200\,ns$ from the time that interaction takes place. It takes about $160\,ns$ to transfer the information from the spectrometer, and the remaining $40\,ns$ is used to make a decision.

If the MG accepts the event, then the readout process and evaluation of the second level trigger begins. The second level trigger decision takes $1.2\,\mu s$. If the event is accepted by the second level trigger, all the detector information for that event is written to magnetic tape for off-line analysis. Afterward, the readout electronics are reset and the process is repeated. The reset takes $1\,\mu s$.

3.12.1 First Level Trigger

The first level trigger checks to ensure that the photon has interacted in the target and that charged particles are emitted from the target. This is achieved by TR1, which is located at the upstream of the first SSD plane (see Figure 3.4). The TR1 counter consists of a scintillator counter and a PMT. The TR2 counters are located downstream of the last SSD plane, and ensures that the charged particles that pass through TR1 also go through the microstrips. TR2 consists of four scintillator coun-

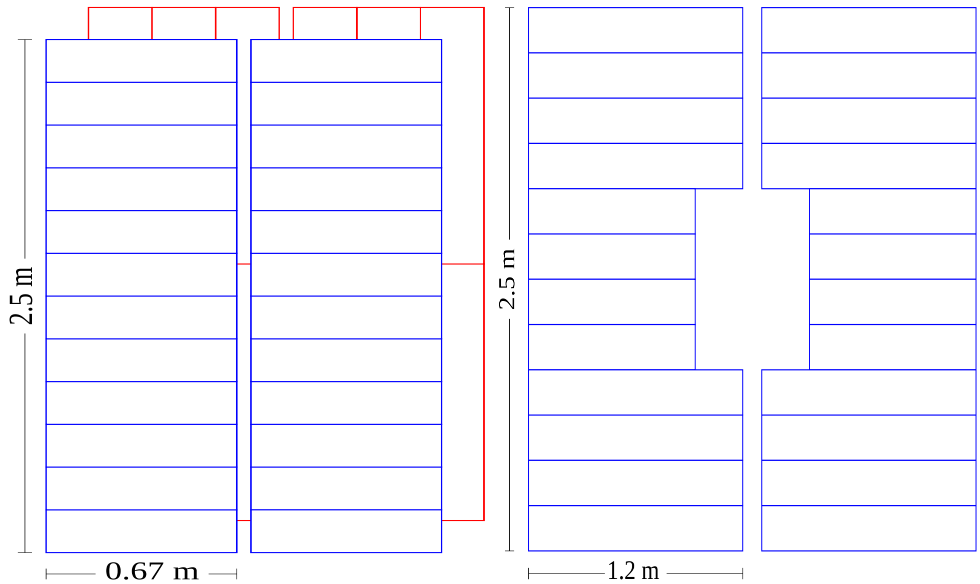

Figure 3.10: $H \times V$ and OH schematic

ters and PMT's.

The requirement that there are wide-angle tracks in the event is ensured by using the $H \times V$ array. These scintilation detectors are located downstream of M2 after the last PWC station. A fast trigger logic module determines if the firing pattern is consistent with one charged particle, $(H \times V)_1$, or more than one, $(H \times V)_2$. The array has a central gap to let the $e^{\pm}$ pairs through without counting them.

Another set of scintillator counters called OH are located upstream of OE to assure passage of at least one particle.

The hadronic MG requirement is:

$$\mathrm{MG1} = \mathrm{TR1} \cdot \mathrm{TR2} \cdot \{(H \times V)_2 + [(H \times V)_1 \cdot \mathrm{OH}_1]\} \cdot E_{HI} \tag{3.3}$$

where E_{HI} ensures the energy deposited in HC by the hadrons is above a high threshold.

3.12.2 Second Level Trigger

The second level trigger requires evidence of a sufficient number of tracks outside the pair region. This is done by evaluating a logic signal with a voltage proportional to the number of hits in each plane that is derived from the PWC read-out module. The information from each plane is combined, and evidence of at least three tracks outside the pair region is required ($MULT4$).

The inner electromagnetic calorimeter (IE) is also used at the second level trigger. The electromagnetic energy deposited in the IE is required to be over a threshold, and at least two hits above threshold in IE is required (E_{IE-2}). This threshold is very low, and is intended to help validate the HC energy requirement.

The hadronic second level trigger requirement is:

$$TRIG1 = \mathrm{MG1} \cdot E_{IE-2} \cdot MULT4 \tag{3.4}$$

Chapter 4

Analysis Techniques in FOCUS

In this section we summarize some critical aspects of our track reconstruction, vertexing, particle identification algorithms and the data reduction techniques used to skim our data down into manageable sets.

To a large extent, tracks are found independently in the microstrip detectors and the MWPC system. An SSD and MPWC track can be "linked" to each other if they are consistent with being due to a single particle interacting with the two systems. Although this strategy was primarily historical (*e.g.* a US group wrote the MWPC reconstruction code and built the chambers, and an Italian group built the SSD system and wrote the SSD tracking code), it turned out to be a wise decision since we discovered relative motions between the systems on the order of a millimeter due to ground motions. By comparing the MWPC and SSD trajectories, we could correct for this motion and thus maintain a high linking efficiency over time.

In general, charm decay products that are long-lived and electrically charged (electrons, muons, pions, kaons, and protons) are required to be *linked tracks.* These tracks have complete, well-measured 4-momenta. Wide-angle tracks, reconstructed by the SSD system but missing links to the PWC, are allowed to be included in the primary vertex. We believe the inclusion of unlinked primary vertex tracks significantly increased the efficiency of our vertex algorithm. We begin with a discussion of MWPC tracking.

4.1 MWPC Tracking

There are three major categories of tracks in FOCUS. The tracks which hit P0, pass through the downstream magnet (M2) aperture, and leave some hits in P3 and P4 are called “5-chamber” tracks.

Tracks with hits only in P0, P1, and P2 are called “stubs” or “3-chamber” tracks. These tracks are not able to traverse the M2 aperture, either because they are wide angle, or they have low momentum and are deflected too much by M1 to traverse the M2 aperture.

The third category are the “recon” tracks which are from decays downstream of the first wire chamber, that do not leave hits in P0. This category includes secondaries from the long lived neutral “vees” decaying into two charged tracks between P0 and P2, and “kinks” (track segments which join with other track segments but with a finite angle between the segments).

It is important to be cognizant of whether a particle trace is a track or a stub, since tracks had significantly better fractional momentum resolution than stubs. Many of the reconstruction algorithms make use of the anticipated tracking momentum error in constructing vertices and computing quantities such as the normalized mass.[1] Also, tracks typically interact with the “inner” systems (IE, HC, and IMU) while stubs interact with the “outer” ones (OE and OMU).

4.1.1 Microstrip Track Reconstruction

The algorithm for SSD track reconstruction is based on projection-finding on the three separate views. Of the four planes in each view, a minimum of three planes are required to form a projection. Projections are formed into tracks if the group of projections pass a test on the hypothesis that they are consistent with a line. If the χ^2 per degree of freedom for the hypothesis is less than 8, the projections are

[1]A normalized mass for say a $D^+ \to K^-\pi^+\pi^+$ candidate would be the difference between the reconstructed mass and the nominal mass divided by the anticipated mass error.

considered to form a track. The parameters of the fit are slope and intercept of the track in the granite block coordinate system.

The class of hits that are not associated with any reconstructed track are used to search for wide angle tracks and for highly multiple-Coulomb-scattered tracks.

The reconstruction efficiency of the SSD tracks is proportional to momentum, with higher momentum tracks having better efficiency. The resolution of the SSD tracks is also a function of momentum, for lower momentums tracks have a larger multiple Coulomb scattering, hence a worse resolution. The resolution on the intercept of a track extrapolated to the center of our target can be expressed as:

$$\sigma_x = 11\ \mu \sqrt{1 + \left(\frac{17.5\ GeV}{P}\right)^2} \quad , \quad \sigma_y = 7.7\ \mu \sqrt{1 + \left(\frac{25.5\ GeV}{P}\right)^2} \tag{4.1}$$

where P is the track momentum. The constants $11\mu m$ and $7.7\mu m$ are the granularity of SSD strips. This equation applies to tracks which traverse the high-resolution portion of each of the SSD microstrips. It includes the multiple scattering contributions in the SSD and TR1. In fact, we often do better than this because of the target silicon.

4.1.2 PWC Track Reconstruction

The algorithm proceeds by reconstructing projections in all views in each station. The projections in the X (non-bend) view are formed by matching the hits on the PWC planes with the seed track extensions in the SSD. The Y, U and V (bend) view projections are formed independently from the PWC hits alone. Then, the X projection is matched to the U, V, and Y projections to form a track. After all the tracks from the SSD-extended projections are used, new projections in the X view are formed using the hits that were not used in the previous steps. These projections are in turn matched to the unused U, V, and Y projections to form more tracks.

The reconstruction algorithm requires that the reconstructed tracks have hits in the first PWC station, P0. It also requires that the tracks can miss a maximum of four hits in all PWC stations, and that they can miss a maximum of two hits per

PWC station.

At this stage, a χ^2 fit is performed on all candidate tracks, where the fit parameters are the slope and intersection of each track in X and Y views in the M2 magnet bend plane. An additional fit parameter for the 5-chamber tracks is the change in slope in the Y view between the track segments upstream and downstream of M2. Finally, a χ^2/DoF cut is applied to select the legitimate tracks from the candidate list.

There are tracks that leave the PWC geometrical acceptance before P2. To try to recover these tracks, the algorithm uses the SSD tack extensions to search for the unused hits in the X view of P0 and P1, which are then combined with the U, V, Y projections to form tracks. The tracks which extend only to P0 are required to have hits in all four views, and the tracks which extend to P1 are required to have at least three hits in each station.

There are higher-order corrections to the linear least-squares fits described above. Because of the finite length of M2, one correction is to account for the sudden bend approximation implied by the above parameterization. There also other magnetic corrections such as counting for the fringe field, and the fact that the components of the $\vec{B}$ field are not constants but are functions of x, y and z.

4.1.3 Linking of SSD and PWC tracks

Linking is performed by extending the SSD track to the M1 center and matching them with the PWC tracks. In addition, for the 5-chamber PWC tracks, bend angles inside the M1 and M2 magnetic field should be consistent. We significantly reduce backgrounds by requiring that the PWC and SSD track segments "link". Those PWC tracks which do not link can be used as possible "vees" and "kinks" (such as K_s, Λ, Ξ^-, and Ω^- decays).

4.1.4 Momentum Determination

The momentum associated with a track is calculated by the bend angle of the track in M1 and M2 magnets. Depending on the type of the track, different methods are used to calculate the momentum. For example, in 5-chamber tracks the bend angle in M2 is calculated using the track parameters upstream and downstream of M2. For the 3-chamber tracks, the SSD and PWC portions of the track are used for calculating the bend angle in M1. The algorithm takes into account the magnetic field corrections. The following expressions give our approximate momentum resolution in M1:

$$\frac{\sigma}{P} = 3.4\% \left(\frac{P}{100\,GeV}\right)\sqrt{1+\left(\frac{17\,GeV}{P}\right)^2} \tag{4.2}$$

and in M2

$$\frac{\sigma}{P} = 1.4\% \left(\frac{P}{100\,GeV}\right)\sqrt{1+\left(\frac{23\,GeV}{P}\right)^2} \tag{4.3}$$

where the second terms inside the square root account for the multiple Coulomb scattering.[2]

4.2 Vertex Reconstruction

The fully-reconstructed charm sample discussed here was both selected and reconstructing using the DVERT vertex finder developed at the University of Illinois. DVERT is a "candidate driven" (rather than "topological") vertex finder which remains efficient at short detachment distances. We illustrate the vertexing and selection philosophy with the state $D^+ \to K^-\pi^+\pi^+$, which is one of the states analyzed in this thesis as a normalizing mode (see Chapter 5).

Rather than searching for a primary and secondary decay vertex using pattern recognition techniques, we begin by looping over all potential three-track combinations having a net charge of ± 1 as possible $D^+ \to K^-\pi^+\pi^+$ candidates. We typically

[2]These are correct expressions in the absence of adjacencies. One gets an adjacency when several adjacent MWPC wires fire for a given track due to the presence of δ-rays, or sharing of the ionization cloud.

require the DCL ("Dee" confidence level) of the $K^-\pi^+\pi^+$ vertex to be larger than 1%. This DCL is computed by selecting an x_o , y_o ,and z_o which minimizes a χ^2. The extrapolation errors σ_{xy} are approximately those of Eq. 4.1.[3] For $K^-\pi^+\pi^+$, this χ^2 would be converted to a confidence level using 3 degrees of freedom.

In order to identify a charm production (or primary) vertex, we employ one of three different algorithms. DVNUCL is used for all fully-reconstructed charm decays, where the decay vertex is defined by two or more tracks from charged daughters. In essence, DVNUCL uses the reconstructed charm momentum as a virtual track, and "nucleates" other tracks about this one, forming the primary vertex. In semileptonic decays, where an undetectable neutrino is present, the D is not fully reconstructed, and we have no D momentum vector to point back with. In these cases, we use the DVFREE algorithm, which simply searches for all possible vertices. We will describe these two alrorithms in more detail below. A third alrogithm, DVNUCG, is used when the charm momentum vector is known, but only one daughter track is available. This is somewhat similar to DVNUCL, except that the vertex is not constrained to be on a line ("seed track"), but rather is constrained to a half-plane.

4.2.1 DVNUCL

In most charm decay studies, the charm state will be fully reconstructed, and we will have multiple tracks in the SSD system which define a decay vertex. In order to locate the primary vertex, a "seed track" is constructed which begins at the decay vertex, and points back opposite to the D momentum vector. The seed track has much better measurement error than any of its constituent tracks because its intercept is averaged. We then search for all intersections of the seed track with any of the unused tracks in the event. The DVNUCL algorithm tries to form the largest-multiplicity primary vertex with a PCL (Primary vertex Confidence Level) in excess of 2%). If a primary vertex is found, we compute the detachment of the three-dimensional sepa-

[3]However, we include information from the target silicon, and weight points appropriately for being in the high versus low resolution region of the SSD system.

ration between the primary and secondary vertex (ℓ) and divide it by the anticipated, event-by-event RMS resolution, σ_ℓ. ℓ/σ is an excellent measure of the detachment of secondary vertex from the primary vertex. Cuts on this detachment improve the signal-to-noise ratio by requiring larger ℓ/σ, at the cost of some signal.

4.2.2 DVFREE

The semileptonic samples reported here were reconstructed using an alternative vertex algorithm called DVFREE. DVFREE is a routine that can find a primary vertex in the case that there is no seed track from a charm candidate. This occurs for decays with ν's, such as $K\pi\mu\nu$ or unreconstructed π^0's, such as $K\pi\pi^0$. Because the daughters are not all found, the D direction is not known, so we cannot create a seed track.

Typically, a set of tracks forming a good vertex are taken as a D candidate. These tracks are entered in the exclusion list, then DVFREE is called. DVFREE returns a set of vertices formed from the remaining tracks in the event and a primary vertex is chosen. Usually the most-upstream vertex or the highest multiplicity vertex are used.

The DVFREE algorithm begins with finding any pair of tracks that form a vertex with a confidence level greater than 1%. It then adds as many tracks as possible to that vertex as long as the confidence level remains above 1%. The tracks are not searched or included in any particular order.

The routine then chooses a track that is not in the first vertex and attempts to cluster other tracks around it. When it is looking for these additional tracks, it considers tracks that might already be in another vertex. This way, a track may be in any number of vertices. This process (of choosing a track that is not in a vertex and clustering around it) is repeated until all tracks are in every vertex they are consistent with, or in no vertex at all.

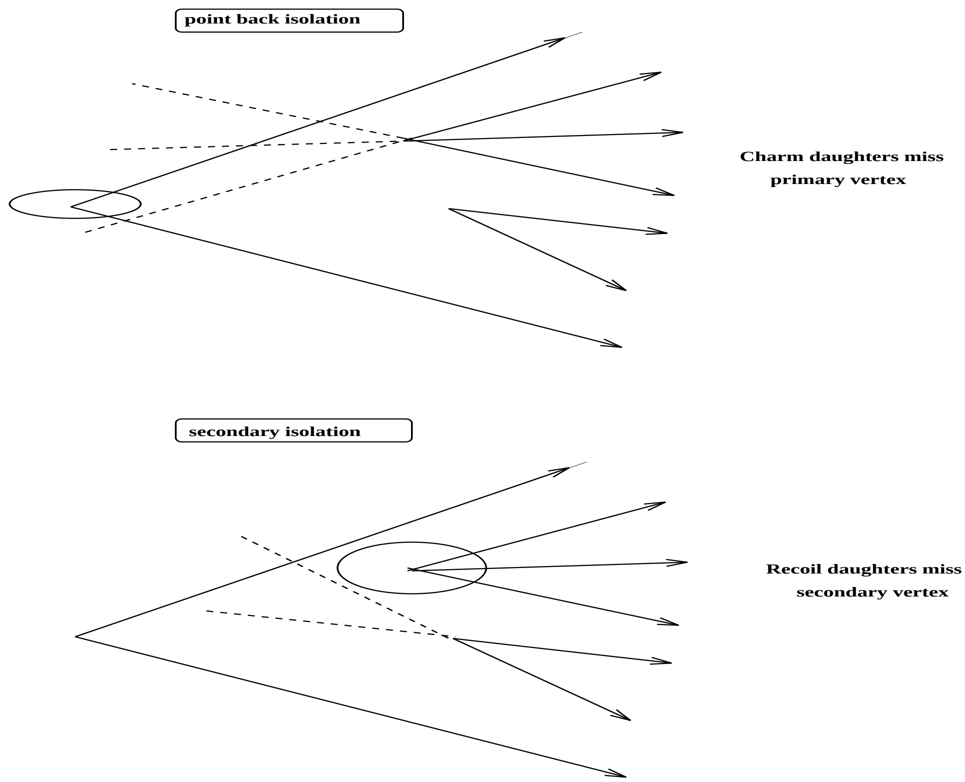

Figure 4.1: Schematic representation of isolation cuts ISOP and ISO2

4.2.3 Vertex Isolation Cuts

Isolation cuts are often used to greatly suppress non-charm and charm reflection backgrounds. The first type of cut, which we will call the primary isolation cut or ISOP, is basically a requirement that no charm-secondary track is consistent with pointing to the primary vertex. For the case of $D^+ \to K^-\pi^+\pi^+$, we find the highest confidence level vertex constructed from the primary vertex tracks and each of the $K^-\pi^+\pi^+$ tracks forced in the vertex one at a time. ISOP is the largest confidence level found, and cuts on this variable require it to be small.

The second type of isolation cut is based on the confidence level of the hypothesis that other tracks are consistent with the decay vertex. We call this the secondary isolation confidence level, or ISO2.

We compute ISO2 by testing all other tracks in the SSD system, searching for the largest confidence level that a track fits in the decay vertex. There are two variants of this technique: one where we exclude primary vetrex tracks from the search list, and one where we consider all tracks (except the tracks used to form the secondary vertex, of course). The purpose of the ISO2 cut is to remove backgrounds from higher-multiplicity charm decays. In principle the primary vertex tracks, which are already accounted for, need not be considered. Thus the secondary isolation cut exclusive of primary vertex, "ISO2EX", is usually prefered. Since ISO2 uses some tracks not used in ISO2EX, it is always true that ISO2EX $\leq$ ISO2. Therefore the ISO2EX cut is more efficient, and inclusive of a ISO2 cut.

4.3 Čerenkov Identification Algorithm

This section describes the algorithm used to identify charged tracks in FOCUS. (I was involved in the development of the FOCUS Čerenkovalgorithm.) We begin by describing the algorithm and conclude by illustrating the algorithm's performance using physics signals.

The Čerenkov algorithm used in FOCUS is given the acronym CITADL (for

Čerenkov Identification of Tracks by an Algorithm using Digital Likelihood). CITADL returns relative likelihoods that the track had a Čerenkov pattern consistent with that expected for the electron, pion, kaon, or proton hypothesis. CITADL constructs a log likelihood variable based on the firing probability for all Čerenkov cells that a given track could potentially affect — all cells within the track's $\beta = 1$ Čerenkovcone. Assume for the moment that a cell only fired in response to Čerenkov light. If the cell fired, and μ photoelectrons were expected, the log likelihood was incremented by $\log(1 - \exp(-\mu))$; while if the cell failed to fire the log likelihood was incremented by $\log(\exp(-\mu))$. Cells which overlapped more than one track's Čerenkov$\beta = 1$ cone were considered "confused" and excluded from the sum. The likelihood returned by CITADL is similar in spirit to the traditional continuous likelihood used in fitting. The only difference is that each event has only two outcomes — on or off. For this reason, we call it a "digital" likelihood.

We found significant improvements in the performance of the Čerenkovalgorithm by including the effects of accidental cell firing due to untracked electromagnetic debris. We determined the accidental firing rate by measuring the fraction of times a Čerenkov cell would fire, even if it were outside of the $\beta = 1$ Čerenkov cone of all observed tracks. The accidental rate varied considerably, and for central cells it was very large. It is easy to incorporate accidental firing rates in the firing probability. The prescription is $P_{fire} = a + (1 - \exp(-\mu)) - a\ (1 - \exp(-\mu))$, where a and μ are the accidental rate and the number of photoelectrons expected for the given cell. We found that a was often proportional to the beam intensity, particularly for cells near the beam axis. The inclusion of accidental rates significantly improved the performance of CITADL.

CITADL returns its identification in the form of χ^2-like variables which we will call W_e, W_π, W_K, and W_P. They are defined by $W_i \equiv -2\sum_j \log P_j$, where P_j is the probability for the observed outcome (on or off) for that cell under each of the 4 particle hypotheses. One would typically require that kaon hopefuls (from charm decay candidates) pass a minimum cut on a likelihood difference variable such as

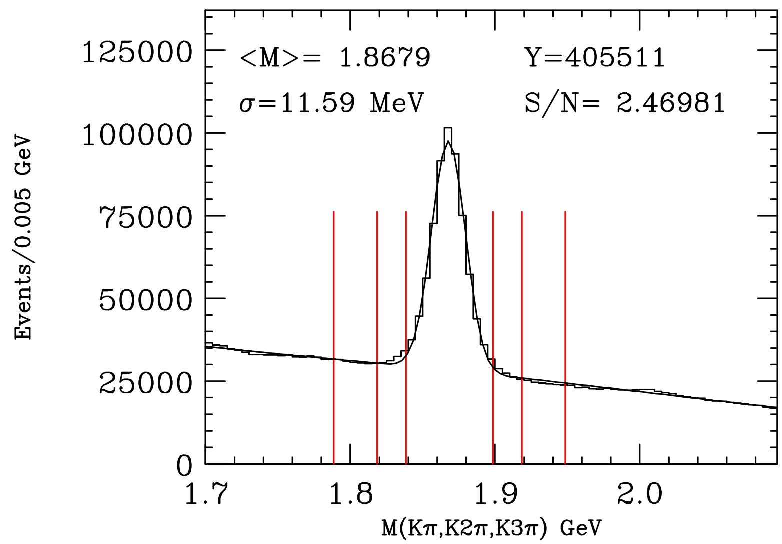

Figure 4.2: Invariant mass plot for three golden mode decays $D^0 \rightarrow K^-\pi^+, K^-\pi^+\pi^+\pi^-$, and $D^+ \rightarrow K^-\pi^+\pi^+$. The reconstructed D^+ mass was shifted by 5 MeV so that its peak will reconstruct in the same place as the peak of the D^0. This data has vertex quality, and kinematic cuts only. No Čerenkov cuts were used. The vertical lines denote signal and sideband regions which will be used to make a background subtraction.

$\Delta W_K \equiv W_\pi - W_K$. A large ΔW_K implies that the kaon hypothesis is significantly favored over the pion hypothesis.

We found that it was possible to use golden mode charm as a monitor of Čerenkov performance. Figure 4.2 shows a 405,000 golden mode charm sample obtained (using about 75% of our data) without any Čerenkovcuts. A selection of stringent cuts on vertex detachment, isolation, the $D^{*+} - D^0$ mass difference, and momentum were used to obtain this reasonably clean sample. Also shown are sideband regions used for background subtraction. Figure 4.3 shows the likelihood difference $\Delta W_K = W_\pi - W_K$ for the kaon and pion daughters of these background-subtracted charm decays, in two ranges of momentum. For convenience, we will call the variable $\Delta W_K \equiv W_\pi - W_K$ "kaonicity". A positive kaonicity implies that a given track is more likely to be a kaon as opposed to a pion.

Figure 4.3(a) shows the kaonicity distribution for charm kaons and pions in a momentum range above the pion threshold of C2 (the lowest threshold counter) but

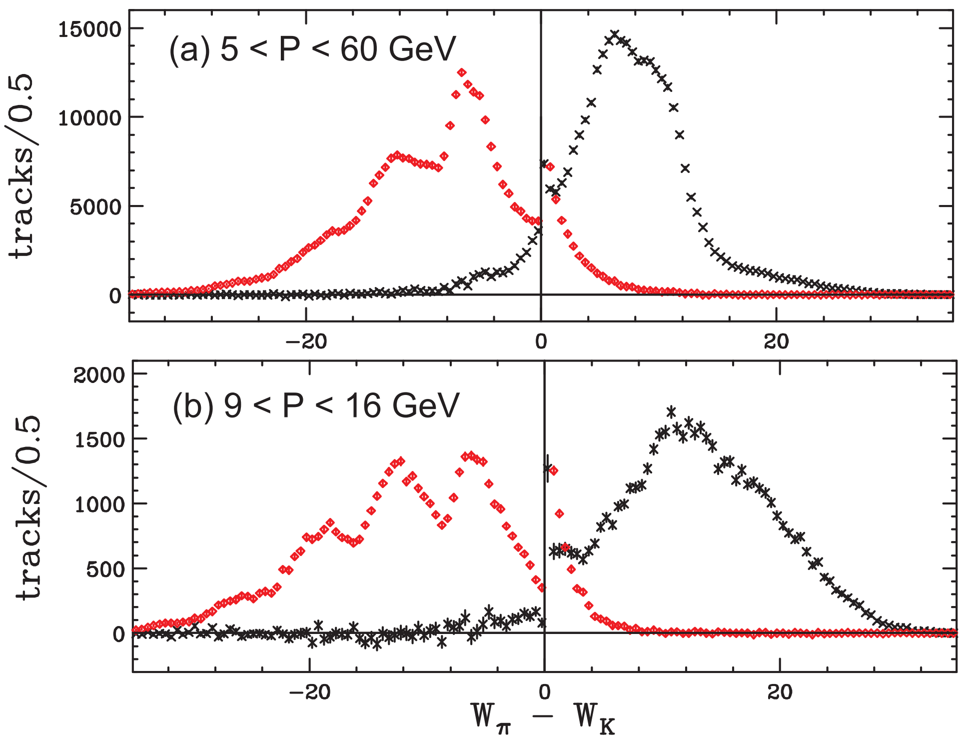

Figure 4.3: The log likelihood difference $W_\pi - W_K$ distribution obtained from background subtracted kaons and pions from golden mode charm signal shown in Figure 4.2. The pion distributions were rescaled to have the same area as the kaon distributions. Plot (A) is for tracks with momenta in the range $5 <$ P < 60 GeV/c. Plot (B) is for tracks with momenta in the range $9 <$ P < 16 GeV/c. There are off-scale spikes in the 0 bin consisting of 20,000 and 4,500 events for Plots (A) and (B).

below the kaon threshold of C3 (the highest threshold counter). Outside of this momentum range, the FOCUS Čerenkov system is incapable of much K-π separation, and the kaonicity distribution is strongly peaked near zero. Figure 4.3(b) shows the kaonicity distribution in the more restricted range from 9 to 16 GeV. In this range kaon-pion discrimination is particularly effective, since it lies above the pion threshold for C1 but below the kaon threshold of C2.

Figure 4.3 shows that even though the likelihoods are constructed from the discrete firings of Čerenkov cells, the kaonicity distribution for kaons is reasonably continuous except near $\Delta W_K = 0$. As Figure 4.3(a) shows, averaged over the accepted charm momentum spectrum, pion backgrounds to kaons can be effectively eliminated while still maintaining high efficiency for charm kaons. A cut just above kaonicity of zero rejects a large fraction of pions. The fraction of background pions then dies away exponentially for $\Delta W_K > 0$ beyond zero. Over the more restricted range from 9 to 16 GeV, where cells from both C1 and C2 discriminate pions from kaons, the ΔW_K distribution for kaons broadens considerably. One can make a very stringent kaonicity cut to suppress pion backgrounds and still maintain good efficiency for real kaons.

As Figure 4.2 shows, it was indeed possible to get reasonably clean charm signals without the use of Čerenkov information. However, many FOCUS analyses employed Čerenkov cuts as an effective way of increasing signal-to-noise, while maintaining reasonable efficiency. Figure 4.4 illustrates the effectiveness of kaon and pion Čerenkov cuts for $D^0 \to K^-\pi^+\pi^+\pi^+$ events selected using an $\ell/\sigma > 9$ detachment cut but without any Čerenkov cuts. The kaon cut is on "kaonicity", or the log likelihood difference $\Delta W_K \equiv W_\pi - W_K$ discussed previously. The pion cut is based on a variable which we will call "piconicity", defined as $\Delta W_\pi \equiv W_{\text{min}} - W_\pi$. The ΔW_π cut is placed on all D decay pions and is meant to insure that no pion being considered as a charm daughter is grossly inconsistent with the pion hypothesis. A cut such as $\Delta W_\pi > -2$ means that none of the other three particle hypotheses are favored over the pion hypothesis by more than a factor of $\exp(2/2) = 2.71$. For the $D^0 \to K^-\pi^+\pi^+\pi^+$ sample, the requirement $\Delta W_K > 0$ preserves 84% of the yield while increasing the

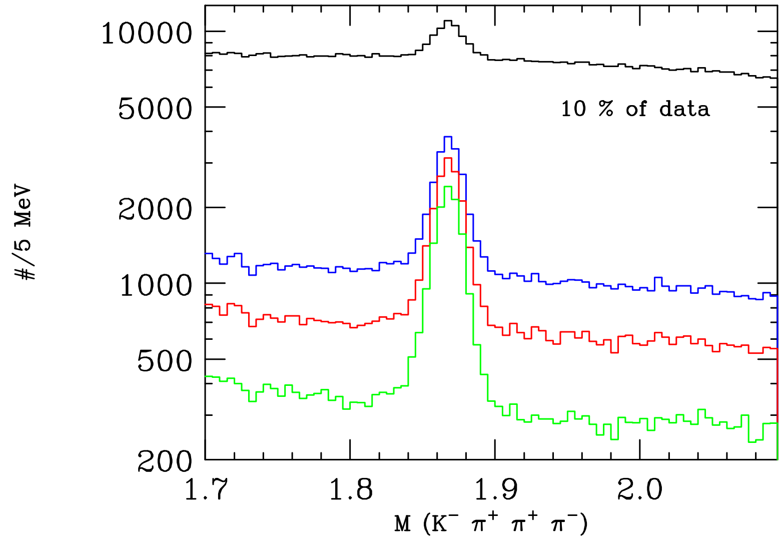

Figure 4.4: Illustration of the effectiveness of Čerenkov cuts in reducing backgrounds to $D^0 \to K^-\pi^+\pi^+\pi^-$. Note the logarithmic scale. The upper curve has no Čerenkov cuts. The second histogram requires $\Delta W_K > 0$. The third requires $\Delta W_K > 2$. The fourth histogram requires $\Delta W_K > 2$ and $\Delta W_\pi > -2$. The fitted signal yields in these plots are 15307, 12783, 11699, and 9996 respectively.

signal to noise by a factor of 6.2. The more stringent $\Delta W_K > 2$ and $\Delta W_\pi > -2$ preserves 75% of the un-cut signal yield but increases the signal-to-noise by a factor of 16.

4.4 Data Reconstruction and Skims

During the 1996-97 run, FOCUS collected about 6.5 billion photon triggers on about 6000 $8mm$ "Exabyte" magnetic tapes. This amounted to nearly 30 TB of data. Because of the large amount of data, the reconstruction and skimming process of data consisted of three stages, as described below.

4.4.1 Pass One

The Pass One reconstruction process involved analyzing the raw data and writing the reconstructed data on another set of 6000 tapes. Pass One reconstruction required

Table 4.1: Super-stream Descriptions

Superstream	Description	Skim Two Institution
1	Semi-leptonic	Puerto Rico
2	Global Vertex and K_s	Illinois
3	EM Neutrals and ϕ Skims	CPBF, Brazil
4	Baryons	Fermilab
5	Diffractive, Leptonic and Out-of-Target	California, Davis
6	Fully Reconstructible Charm Decay Modes	California, Davis

running the reconstruction algorithms for all the detectors.

A Fermilab software product called CPS (Cooperative Process Software) was employed to construct a "farm" consisting of a "server" node and about 10 "worker" nodes. This increased the computing power substantially, allowing the Pass One processing to occur in parallel.

Up to eight computing "farms" at Fermilab consisting of up to 90 worker nodes were employed for the Pass One. The types of workstations used included SGI workstations based on the MIPS R5000 CPU and IBM workstations based on the IBM/Motorola PowerPC CPU. Pass One processing took roughly one year to complete.

4.4.2 Skim One

The purpose of Skim One is to divide the data into smaller data sets, each set containing data based on various classes of physics (see Table 4.1). This skim divided the Pass One data set to six super-streams. Each superstream required 200-500 tapes.

University of Colorado and Vanderbilt University ran the Skim One on two computer clusters of about 4000 MIPS each, which began in October 1998 and finished in February 1999.

4.4.3 Skim Two

Skim Two constitutes the final skimming stage, where the Skim One output tapes were further split into smaller skims with more specific physics topics. Generally, 5–12 sub-streams were written from each super-stream. Skim Two was done in five institutions, listed in Table 4.1. It began in January 1999 and was completed by June 1999. I was primarily responsible for the software and operations of Skim Two of superstream 2.

In the form factor analysis, the data were taken from the semileptonic+meson subskim, FSAA. This was a subskim of superstream 1. The events in this subskim were required to have a muon track, with either an inner-muon or outer-muon confidence level greater than 1%, intersecting with another track. The vertex confidence level must exceed 1%. At least one of the intersecting tracks must be consistent with a kaon, having a kaonicity > 1.

The data for the branching ratio measurement were taken from a special skim called SLEPNRM, described in Chapter 5.

4.5 Performance of the Inner Muon System

In this section, we assess the performance of the inner muon system in terms of efficiency and muon misidentification. Most of our studies use a convenient golden mode ($K\pi$, $K2\pi$, and $K3\pi$) sample designed to study Čerenkov algorithm performance. Our conclusion is that the inner muon system performed extremely well during our FOCUS run with essentially 100% efficiency and nearly no noise.

The efficiency studies form an important validation of the Monte Carlo model used in this thesis to measure form factors and the $D^+ \to \bar{K}^*\mu^+\nu/K^-\pi^+\pi^+$ relative branching fraction. The misidentification studies will form the underpinning of our studies of backgrounds for the $\overline{K}^{*0}\mu^+\nu$ state discussed in Section 7.5.

4.5.1 Efficiency Studies

There are two components to determining muon identification efficiency — the efficiency of the individual muon detectors, and the efficiency of the muon algorithm. We have to study these efficiencies in a rather indirect way since we have no easily accessible signal to supply us with an independent unbiased source of muons, short of the semileptonic decays studied here. The obvious source of muons would be the copious FOCUS sample of photoproduced $J/\psi \rightarrow \mu^- \mu^+$ decays. Unfortunately, we only collect these events using a dimuon trigger that would significantly bias the studies of muon identification.

We use the copious supply of muon halo tracks present in our data to monitor the efficiency of the counters comprising the six inner muon arrays MH1X,MH1Y... MH3Y. A large number of triggered events included an inner muon due to accelerator backgrounds. These halo tracks helped to satisfy our muon triggers and therefore appeared quite often in the data that we wrote to tape. The muon halo rate in our main-line hadrometer-triggered data was much smaller.

Figure 4.5 shows evidence for muon halo present at the $\approx 20\%$ level in a typical run. This very copious supply of halo tracks served as an essentially continuous monitor of the efficiency of the planes of the inner muon array. A "time-line" of the average efficiency of one of the muon arrays is plotted as a function of run number in Figure 4.6. The efficiency for a given plane is based on events where all other muon array planes fire due to a halo muon. The efficiency is then the fraction of times all 6 planes fire. As shown in Figure 4.6, the typical efficiency in the inner muon system was superb — generally in excess of 99%. Although achieving this high efficiency was very satisfying, we expected excellent performance. The inner muon array used 1.5 cm thick high grade (NE110) scintillator. The counter and phototube mounts were designed to collected more than 500 photoelectrons per traversing muon. In bench tests, cosmic rays often produced several-volt signals from our phototubes when running at nominal high voltage.

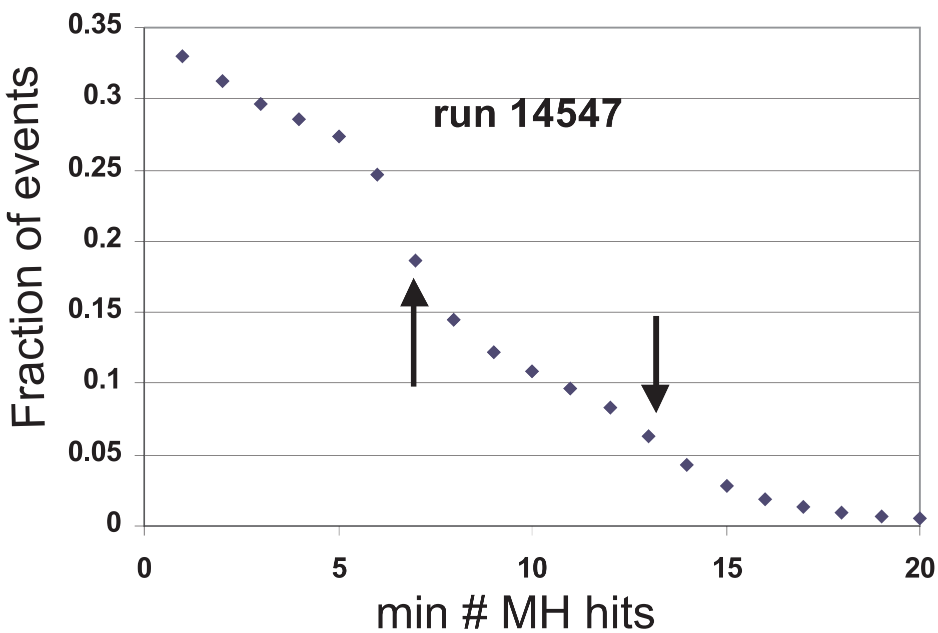

Figure 4.5: We plot the fraction of events in Run 14547 with the number of struck counters in the inner muon system equal to or exceeding the abscissa. A clear "knee" can be seen at 6 or more hits which we attribute to a muon halo track traversing all 6 counter planes that comprise the inner muon system. This knee is rather broad because of range out, wide angle muons that pass wide of the last detector plane, and muons that fire adjacent counters.

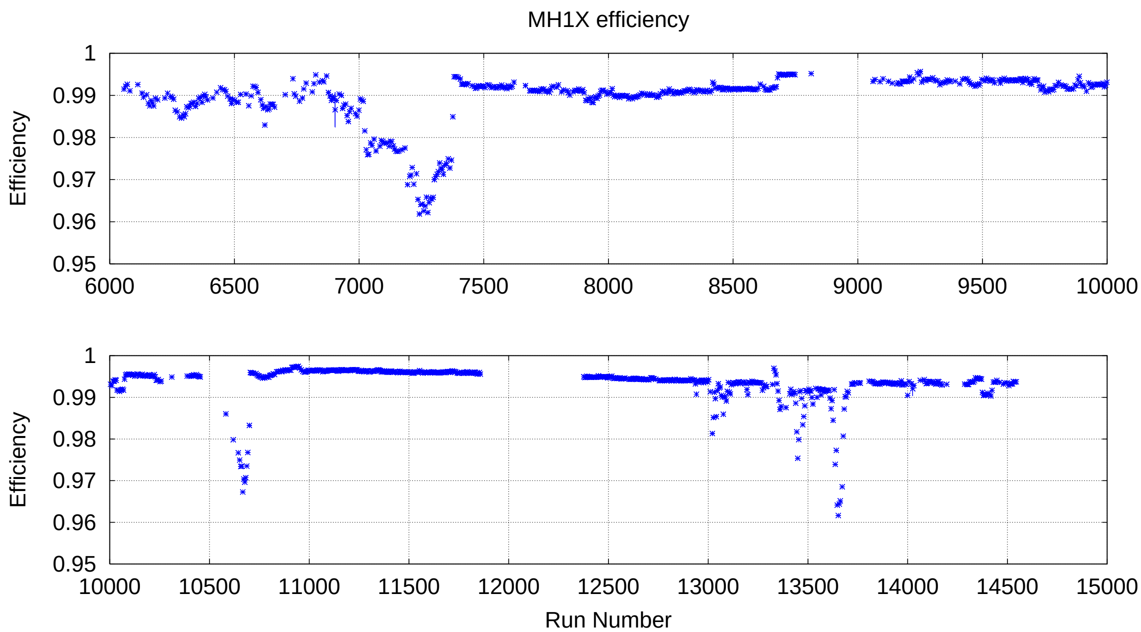

Figure 4.6: The average efficiency of the MH1X array based on the fraction of times this array fires in events where all other muon planes fire. The efficiency of all other planes is very comparable to the efficiency of this plane. The 2% inefficiency dip near 7250 was due to a slight shift in the latch timing. We corrected this timing immediately after discovering the problem — thus restoring the efficiency to its original level.

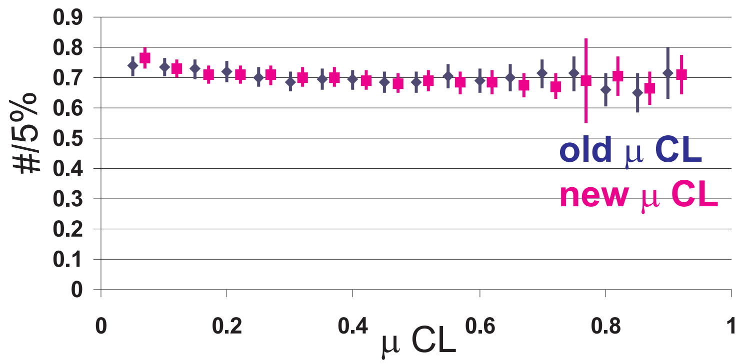

Figure 4.7: Comparison of the ratio of data and Monte Carlo yields for WS subtracted $D^+ \to \bar{K}^* \mu^+ \nu$ events as a function of the muon confidence level. We show results with the original muon confidence level and a more refined confidence muon confidence level incorporating better multiple scattering and granularity errors. The new confidence level points are shifted slightly horizontally for clarity. The vertical scale is arbitrary since it is based on an arbitrary number of MC events.

4.5.2 Analysis Efficiency

In this section, we use $D^+ \to \bar{K}^* \mu^+ \nu$ events to measure the efficiency of our analysis algorithm. The inner muon algorithm provides a confidence level (μCL) that a given track is a muon by matching its trajectory to hits in the inner muon array. Figure 4.7 compares the ratio of the the yield for $D^+ \to \bar{K}^* \mu\nu$ signal events in data over the yield in Monte Carlo, in bins of the returned muon confidence level. The method for "event" counting is discussed in Chapter 5.

We believe that both muon confidence levels are well simulated by the Monte Carlo, since the data-to-MC ratios are nearly independent of either muon confidence level. The most likely source of a mismatch would be a mismodeling of the multiple scattering in the Monte Carlo. Our multiple scattering simulation has been checked using a variety of techniques. We also have explicitly checked, using the techniques described in Chapter 5, that the relative branching fraction for $D^+ \to \bar{K}^* \mu\nu / K^- \pi^+ \pi^+$ is essentially independent of the muon momentum. We compare this relative branching fraction computed for muons with $P\mu > 10\ GeV$ to that obtained in a sample with $P_\mu > 30\ GeV$. Only 32% of the $D^+ \to \bar{K}^*\mu\nu$ of the 10 GeV sample had a momentum exceeding 30 GeV. The two samples had branching fractions that differed by an average of only 1.2 σ over 16 different sets of clean up cuts. We believe that a mismatch between the Monte Carlo multiple scattering simulation and the data would reveal itself in an inconsistency between these different muon momenta samples.

4.5.3 Muon Misidentification Studies

We have studied the misidentification of pions and kaons using a large golden mode charm sample. Figure 4.8 provides a general impression of the fraction of golden mode pions and kaons that are misidentified as muons. For example, the fraction of kaons satisfying an inner muon confidence level cut in excess 1×10^{-4} is about 0.9% averaged over all momenta. The term misidentification is somewhat misleading in this context, since it is probably dominated by legitimate charm secondaries that decay in flight prior to reaching the calorimeters or the muon steel.

We have devised a way of studying the momentum dependence of misidentified pions and kaons. Our approach studies the fraction of golden mode secondary 5 chamber tracks that satisfy muon identification cuts. In order to subtract backgrounds, we make weighted averages of the fraction of times a track is identified as a muon. Our weight is +1 if the candidate mass falls within the signal regions of Figure 4.9, and −1 if the mass falls within either sideband.[4] The signal and sideband regions are illustrated in Figure 4.9.

Figure 4.10 compares the observed muon misidentification versus momentum in data to that in two golden-mode Monte Carlos that differ by their assumed level of muon halo. The data has slightly more misidentification than a "noiseless" simulation (PILEMU = 0), and significantly less than the noisy one (PILEMU = 0.5). *PILEMU* is the Poisson average number of halo muons in each event. The halo muon distribution was, at the time of this study, more of a guess than a measurement, and the

[4]The technique for computing errors of such weighted averages is described in Appendix A.

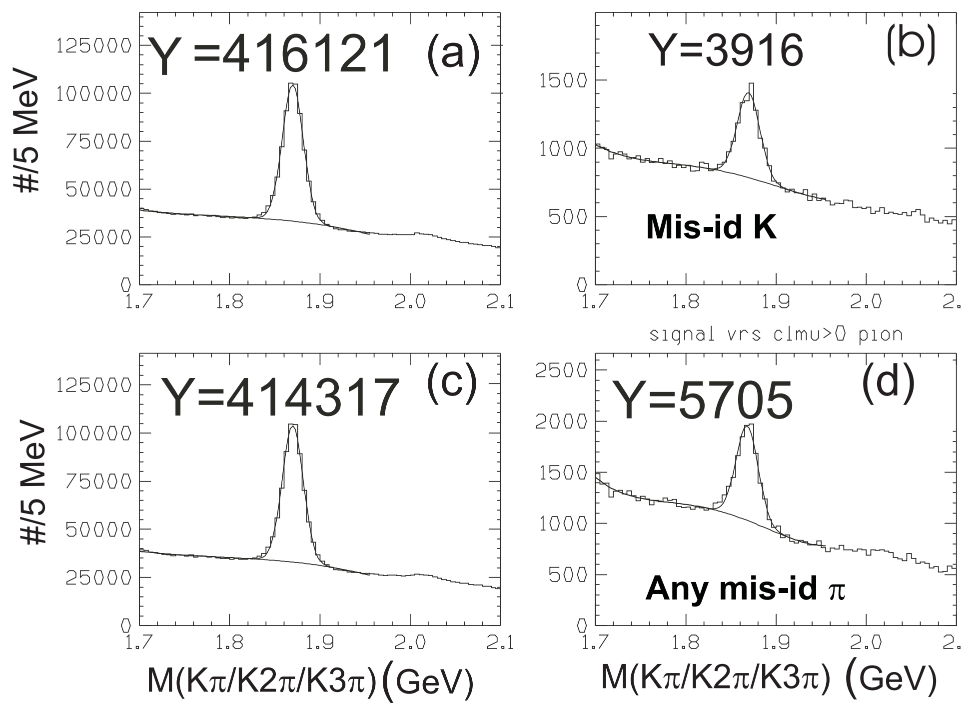

Figure 4.8: Golden mode charm samples used for muon misidentification studies. (a) The golden mode signal when no pion has been misidentified as a muon. (b) The golden mode signal when at least one pion has been misidentified as a muon. (c) The golden mode signal when the kaon track was not misidentified as a muon. (d) The golden mode signal when the kaon track has been misidentified as a muon.

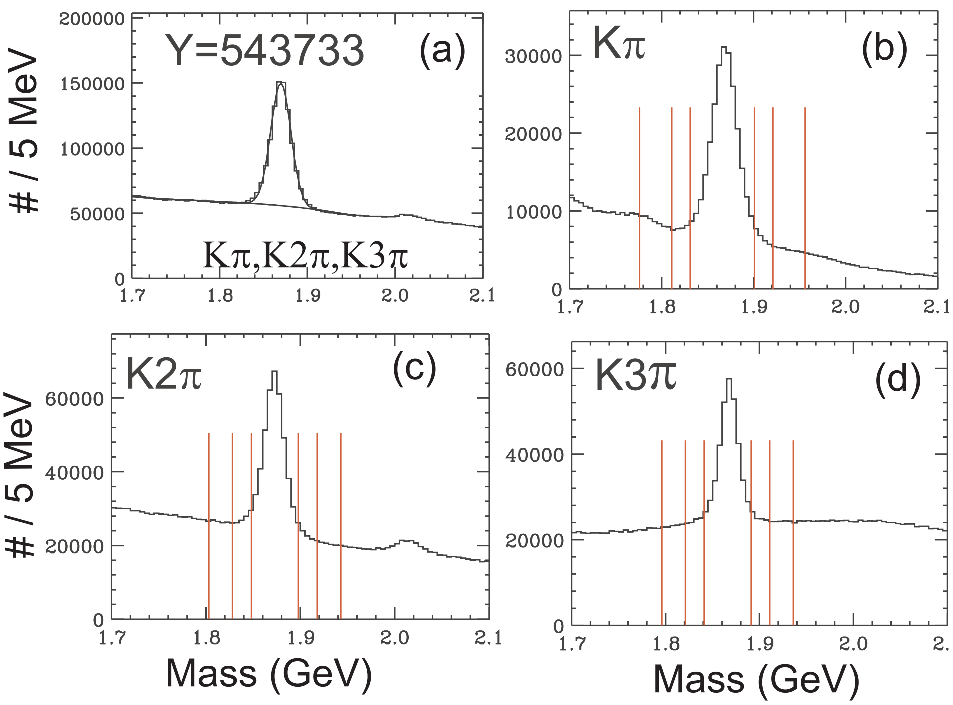

Figure 4.9: We show the golden mode sample and sidebands used in the muon misidentification sample. (a) All three golden modes: $D^0 \to K^-\pi^+, K^-\pi^+\pi^+\pi^-$ and $D^+ \to K^-\pi^+\pi^+$ combined into a single mass plot. The yield of this particular (no Čerenkov cut) selection is about 1/2 million events. (b) The $K^-\pi^+$ mass plot showing the D^0 peak and the signal and two sideband regions used for background subtraction. (c) The $K^-\pi^+\pi^+$ mass plot showing the D^+ peak and the signal and sideband regions. (d) The $K^-\pi^+\pi^+\pi^-$ mass plot showing the D^0 peak and the signal and sideband regions.

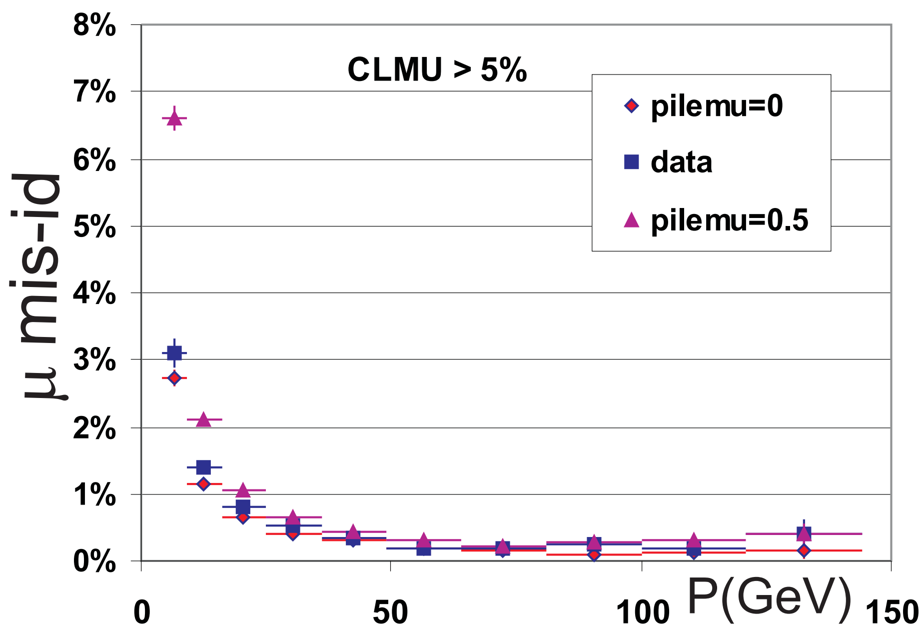

Figure 4.10: Misidentification of golden mode kaons and pions (5 chamber tracks) as inner muons as a function of track momentum. For this plot we use a loose criteria for muons: an inner muon confidence level exceeding 1×10^{-4}. The data is background-subtracted using the sideband method. The data is compared to two Monte Carlos: a zero noise Monte Carlo and one with a much higher muon halo rate than is present in typical charm data.

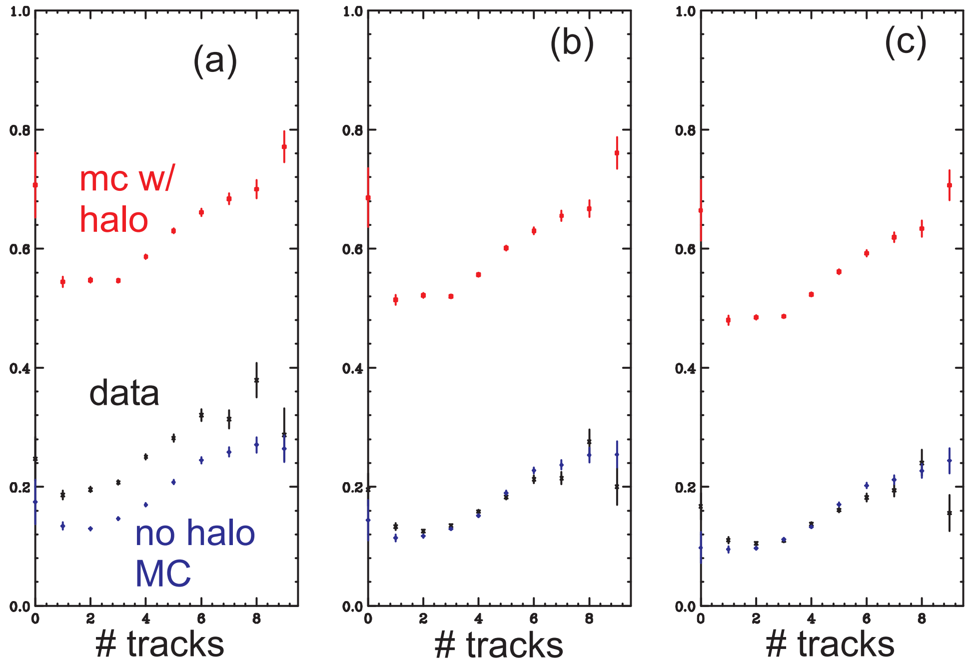

Figure 4.11: The average number of struck inner muon counters in a given array versus the number of reconstructed 5-chamber tracks compared to two Monte Carlos. The first Monte Carlo assumes no halo (PILEMU = 0), the second assumes a halo with a rate specified by (pilemu = 0.5). The three plots are for: (a) MH1X (b) MH2x (c) MH3X. The data is for sideband-subtracted golden mode charm.

rate of 0.5 seemed to match the rate of hits in the muon arrays for some raw events processed in our "expressline analysis".

Figure 4.11 confirms many of these conclusions by comparing the number of struck counters in background-subtracted golden mode data to those in two Monte Carlo simulations. We plot the average number of struck MH counters in three inner muon arrays as a function of the number of reconstructed 5-chamber tracks in an event.

Figure 4.11 has several interesting features worthy of comment. Except for MH1 (Figure 4.11a), the average MH plane multiplicity is an excellent match to a Monte Carlo with no halo and no other source of noise. One might expect more noise in MH1X and MH1Y since these arrays have the least amount of hadronic shielding. We do not simulate hadronic punch through in the Monte Carlo simulation. The average number of struck counters has a component that is proportional to the number of chamber tracks. Presumably this is due to kaons and pions that decay to muons prior to reaching the inner muon steel. The slope of this average number versus track is the same between the two simulations. We are not sure exactly why there appears to be an offset of about 1/10 of a struck counter even when there are no detected 5-chamber tracks in a given event, but we are reassured by the fact that the same offset appears in our noiseless Monte Carlo. Most likely, this represents either a fiducial mismatch between the PWC system and the inner muon muon detector area or a tracking inefficiency.

We believe that the mismatch in charm secondaries that are misidentified between the data and the "noisy" Monte Carlo with PILEMU = 0.5 is likely due to the unrealistically high level of struck counters in this simulation relative to the data. At very low momenta, we use a large "search" radius of muon counters that can be associated with a given track due to an increased multiple scattering smearing. We thus expect the large difference between the misidentification level at low momenta between the halo and no halo Monte Carlos that is clear in the misidentification curves of Figure 4.10.

More recently we have revised the muon halo parameterization in MCFOCUS after a more careful study of the data. The high rate initially seen was due to events accepted through a di-muon trigger. The new halo distribution met the demands of this analysis with an insignificant amount of noise in the inner muon system, yet accounts for the considerable noise in the outer muon RPC's. The current muon rate is 0.15 per event, and the dispersion of the halo has been increased significantly.

4.5.4 Tighter Muon Cuts

Figure 4.12 demonstrates that one can substantially decrease the muon misidentification of golden mode secondaries by tightening the muon identification cuts. We will typically require at most one missing muon plane, increase the inner muon confidence level cut to μCL > 5%, and require $P_\mu > 10\ GeV$. These cuts are 86% efficient for

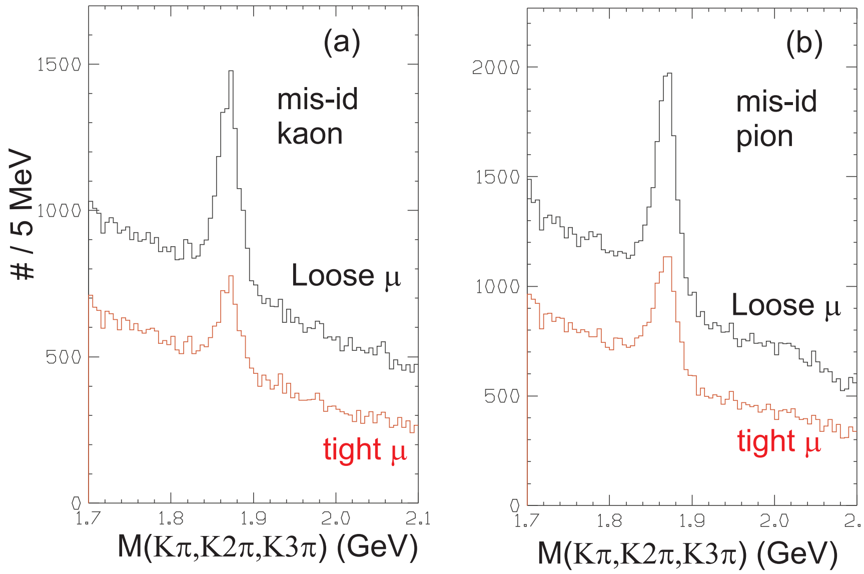

Figure 4.12: We show a comparison of the golden mode misidentification signals with the very loose muon identification $\mu\mathrm{CL} > 1 \times 10^{-4}$ and a tighter muon identification requiring: (1) $\mu\mathrm{CL} > 0.05$ (2) 0 or 1 missing planes (3) momentum > 4 GeV (4) Iso$\mu < 0.1$. Plot (a) has a misidentified kaon and (b) has at least one misidentified pion.

$D^+ \to \bar{K}^* \mu\nu$ signal. As shown in Figure 4.12, somewhat looser cuts than this reduce the fraction of golden mode secondary kaons and pions by a factor of roughly 2.5. Figure 4.12 includes an additional muon isolation cut[5] but has a much weaker (> 4 GeV) momentum cut.

Figure 4.13 compares the background-subtracted charm secondary muon misidentification level with the additional missing plane requirement to the previous misidentification rate in data. These studies show that a variety of cuts can be used to significantly reduce hadron as muon misidentification without substantial efficiency loss.

[5]This is a cut (ISOμ) on the maximum confidence level that another track matches the muon hit pattern associated with the given identified muon

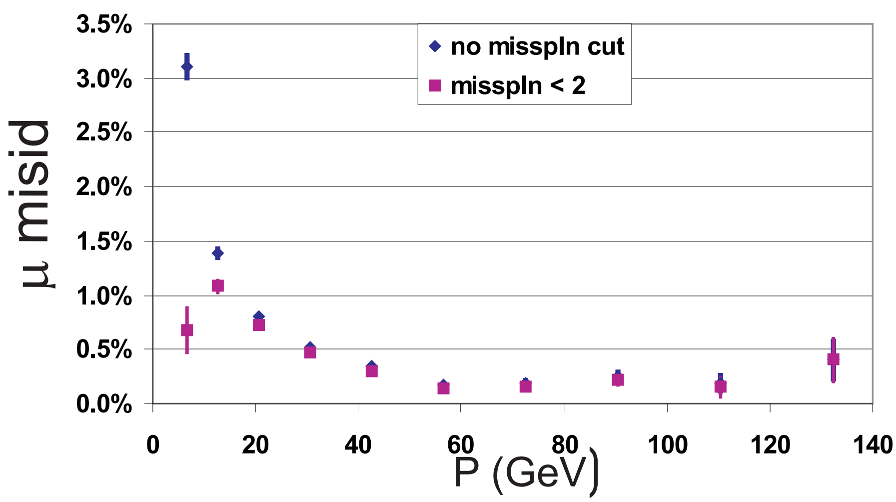

Figure 4.13: Comparison of the charm secondary muon misidentification rate in data for $\mu\mathrm{CL} > 1 \times 10^{-4}$ with and without a cut on the number of missing inner muon planes

4.5.5 Conclusions

We hope this section has supported the following conclusions:

1. The inner muon arrays were extremely ($> 99\%$) efficient throughout the FOCUS run. We will assume that they are 100% efficient in Monte Carlo simulations.

2. We also believe that our analysis cuts primarily based on the muon confidence level (μCL) are well simulated:

   (a) The ratio of background subtracted $D^+ \to K^*\mu^+\nu$ events in data and Monte Carlo are constant as a function of μCL.

   (b) The fraction of $D^+ \to K^*\mu^+\nu$ with $P_\mu > 30\ GeV$ in data is well matched to that in the Monte Carlo.

3. We have performed several studies of muon misidentification of golden mode charm secondaries that help in estimating backgrounds to semileptonic signals. These studies conclude that the data is well simulated using a Monte Carlo with negligible hadronic punch through and no additional noise from muon halo.

   (a) The momentum dependence of the muon misidentification rate for charm secondaries is reproduced by our zero-noise Monte Carlo.

   (b) The zero-noise Monte Carlo reproduces the average number of struck counters in each muon array as a function of the number of observed 5-chamber tracks.

4.6 Neutrino Closure

In this section, we describe the method used to estimate the neutrino momentum and energy using the D^+ line-of-flight. We need this information in order to compute Q^2 and the decay angles that are analyzed in this thesis. Figure 4.14 illustrates our basic technique. In FOCUS, our microstrip system gives us a very good secondary vertex

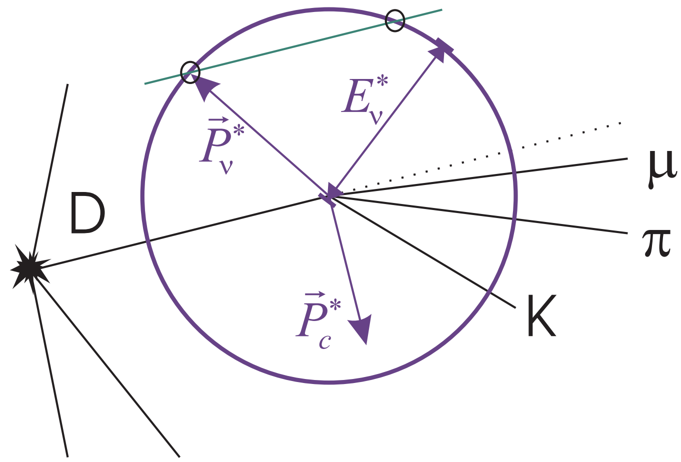

Figure 4.14: A cartoon illustrating the method used to estimate the neutrino momentum from the D^+ line-of-flight

for the $K^-\pi^+\mu^+$ system and for tracks that emanate from the primary vertex. The displacement of the secondary from the primary vertex defines the D^+ line of flight.

Imagine boosting along the D^+ line-of-flight until the $\vec{P}_c = \vec{P}(K^-)+\vec{P}(\pi^+)+\vec{P}(\mu^+)$ or "charged system" momentum 3 vector lies perpendicular to the boost direction. We will call this reference frame the "transverse" frame to differentiate it from the actual D^+ rest frame. By momentum balance we know that the neutrino must be equal and opposite to the boosted $\vec{P}_c$. By a simple invariance argument, we also know the energy of the neutrino in the "transverse" frame. The $P_\perp$ balance gives us the line in Figure 4.14. The neutrino energy constraint gives us the circle. The two intersections of the line with the circle determines the neutrino kinematics in the transverse frame to within a two-fold ambiguity. We can either compute invariants by boosting the $K^-\mu^+\pi^+$ into the transverse frame, or by boosting the neutrino back to the lab.

The transverse frame neutrino energy

Before discussing some of the complications of this method, we will discuss the calculation of the neutrino energy in the transverse frame. We will call the neutrino energy

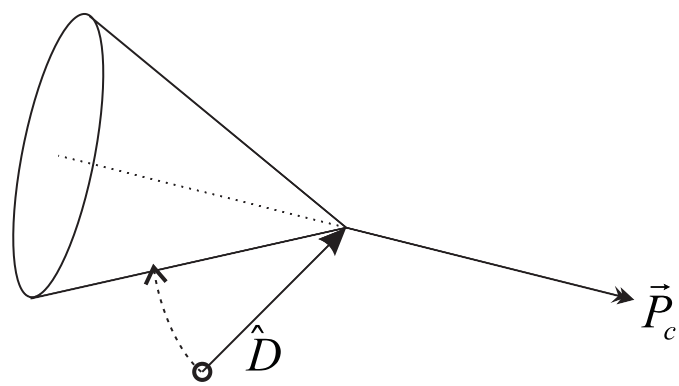

Figure 4.15: The primary vertex cone.

in the transverse frame ω^* and the energy of the charged system E_c^*. We write $\tilde{\nu}$ and $\tilde{c}$ for the 4-vectors of the neutrino and $K^-\mu^+\nu$ system. The D^+ 4-vector is then $\tilde{\nu}+\tilde{c}$.

$$(\tilde{\nu}+\tilde{c})^2 = m_D^2 = m_c^2 + 2\tilde{\nu}\cdot\tilde{c} = m_c^2 + 2\omega^* E_c^* + 2P_\perp^2 \tag{4.4}$$

$$\text{Thus } \omega^* = \frac{m_D^2 - m_c^2 - 2P_\perp^2}{2E_c^*} = \frac{m_D^2 - m_c^2 - 2P_\perp^2}{2m_\perp} \tag{4.5}$$

$$\text{where } m_\perp = \sqrt{m_c^2 + P_\perp^2} \tag{4.6}$$

We note that $E_c^* = m_\perp$ since $E_c^{*2} - P_\perp^2 = m_c^2$ in the $*$ or transverse frame.

Unphysical solutions

Because the DVFREE primary vertex is often badly measured (since it is often pulled by recoil charm tracks that are improperly included), we often get unphysical solutions. These solutions would have $P_\perp > \omega^*$ and correspond to the line lying outside of the circle in Figure 4.14. This happens roughly 40% of the time for the cuts used for our typical $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ signals. The best we can do in these situations is to force such solutions to be physical by moving the primary vertex so that it lies within the physical cone shown in Figure 4.15. The physical cone corresponds to the case where the $P_\perp = \omega^*$ — that is the case where the $P_\perp$ line just touches the circle of Figure 4.14. The apex of this cone passes through the well-determined secondary vertex, its axis is about the well-measured $\vec{P}_c$ vector and it opens upstream of the secondary vertex. The cone 1/2 angle is determined as the maximum lab angle that the charged system can make with respect to the D^+ momentum, given the known D^+ and the well-measured charged-system mass. When the solution is unphysical, the primary vertex lies outside the cone, and we move it the minimal distance so it lies just within the cone boundary. We do this by rotating the unit $\hat{D}$ line-of-flight unit vector about the normal to the plane of Figure 4.15. This rotation axis is along the cross-product of the initial D^+ line-of-flight and the charged momentum. We thus have the following topology of solutions. If the primary vertex naturally lies within the cone, we have two solutions and $\omega^* < P_\perp$. Otherwise, after the recover operation, we will have a single solution with $\omega^* = P_\perp$ and will essentially have a modified boost direction to boost the single neutrino solution back to the laboratory.

Resolving the two-fold ambiguity

In cases where the primary vertex naturally lies within the physical cone, we have the added chore of choosing one of the two solutions. Like many fixed target experiments before us, we chose the "backward" hemisphere solution where the component of the neutrino momentum with respect to the boost direction is negative in the transverse frame. We believe that either solution is equally likely according to the decay distribution, since the D^+ has zero spin, and thus it carries no information about its laboratory momentum. Traditionally, groups use the backward hemisphere solution, since it thrusts the charged system forward in the lab, which is a configuration more likely to be accepted in a forward spectrometer.

Self-criticism of the Closure Technique

Although our method of recontructing the neutrino is essentially what previous fixed-target experiments with microstrip vertexing have done, we suspect there may be many alternative techniques that offer better resolution. In particular, one can imagine exploring many alternative closure methods:

The problem with charm photoproduction is that there are few legitimate primary vertex tracks. We believe the DVFREE vertex used for the primary vertex in semileptonic decays is rather poor compared to the DVNUCL primary vertex used for fully reconstructed charm states. The DVFREE vertex is missing an important constraint: the "D" seed track that passes through the secondary vertex and is directed against its momentum vector. The seed track is the seed about which the other tracks nucleate to create a reasonably reliable primary vertex. Without the luxury of a seed track, DVFREE is much more less constrained and much more susceptible to making false vertices from true primary vertex tracks and tracks from recoil charm particles. It is hard to believe that a constraint that the primary vertex lies within the physical cone would not serve as a powerful tool in arbitrating between potential tracks to include in a DVFREE vertex. Although we made a few attempts to use this information, we were not able to see dramatic improvements in resolution. But we did not make exhaustive studies.

Although the backward hemisphere solution does a somewhat better job at reconstructing the four kinematic invariants involving the neutrino, it is a close contest. One can think of many other aspects to consider for a superior arbitration solution.

1. In cases where the only viable primary vertex is close to the physical cone, one could simply default to the $\omega^* = P_\perp$ solution since the two solutions are close enough. In such a case, we could use the shift between the two solutions as an (inverse) weighting factor in incorporating the event in a fit.

2. In some cases, the shift in kinematics is so large, that the event could not possibly be accepted in the spectrometer for a given hemispheric choice. The unaccepted solution can then be chosen with a high weight.

3. In some cases, one closure solution may produce a D that has an atypical momentum relative to the measured production momentum spectrum for fully reconstructed D^+'s. Where a given event's momentum lies relative to the known spectra can be used as a component to an "arbitration" likelihood.

4. The two solutions will generally produce different proper times for a given event with a given vertex separation. This proper time can be converted to an "arbitration" likelihood contribution from the known D^+ lifetime.

5. Events with little arbitration likelihood separations, but with large kinematic shifts can be either severely down weighted or eliminated from the fit completely.

Although it is difficult to believe that the neutrino kinematics could not be improved upon from some combination of the above ideas, we believe that the resolution we obtain from the traditional method is adequate for the studies summarized in this thesis. The critical issue is whether the Monte Carlo "understands" the resolution of the five kinematic variables describing the $D^+ \to \overline{K}^{*0} \mu^+ \nu_\mu$ decay. This crucial issue is addressed in the next section.

4.7 Studying Neutrino Closure with $D^0 \to K3\pi$

Several of the studies reported in this thesis require that we understand the resolution on Q^2, the decay angles θ_{V}, θ_ℓ, and the acoplanarity angle χ. Because these kinematic quantities depend on an undetected neutrino, they are poorly measured compared to the typical 1/10 mradian resolution that FOCUS has for reconstructed charged tracks. As discussed in Section 4.6, we determine the momentum of the neutrino to a two fold ambiguity using $P_\perp$ balance about the D^+ line-of-flight. This line-of-flight is taken as the ray between the primary and secondary vertex. The secondary vertex is well-measured with understood errors. The FOCUS primary vertex is not as well measured in semileptonic events. For example, tracks from the recoil charm particle are often included with the primary vertex, creating a downstream pull with large non-Gaussian tails. The primary vertex depends on the primary vertex multiplicity and momenta spectra — details that cannot be computed from first principles but rather rely on the Pythia fragmentation model.

4.7.1 The $D^0 \to K3\pi$ Pion Blanking Method

We have developed a method for verifying our Monte Carlo models for the resolution on the four kinematic variables germane to the form factor analysis. Our method uses the abundant reconstructed $D^0 \to K3\pi$ events as a "surrogate" for the semileptonic decay $D^+ \to (K^-\pi^+)\mu^+\nu$. Our basic method will be to "blank" one of the pions in a reconstructed $D^0 \to K3\pi$ event, and use the line-of-flight technique to reconstruct the "blanked" pion momenta. We can then compare the reconstructed decay kinematics using the line-of-flight reconstruction with the very well-measured decay kinematics from fully reconstructed charged tracks. The distribution of the differences of the line-of-flight and a given fully reconstructed variable provides a resolution plot for the given variable. The resolution plot measured directly in "blanking" data can then be compared to that in the Monte Carlo in order to test the fidelity of our simulation.

Since our only reliable simulation is for charm events (rather than non-charm backgrounds), it is necessary to subtract the backgrounds. We subtract these backgrounds using mass sideband weights. Figure 4.16 illustrates the $D^0 \to K3\pi$ sample and signal and sideband regions used for this analysis.

In order to subtract backgrounds, we make weighted histograms of the resolution plots where the weight is $+1$ if the candidate mass falls within the signal regions of Figure 4.16 and -1 if the mass falls within either sideband.

4.7.2 Resolution Studies

Figure 4.17 illustrates the power and limitations of the line-of-flight method by comparing the $\cos\theta_\mathrm{V}$ resolution plots in data and MC for "blanked" and background subtracted $D^+ \to K^-\pi^+\pi^+\pi^-$ events. In this Figure, the "blanked" pion that serves as the neutrino surrogate is the softest of the three pions. We show the distribution for 3 primary vertex choices. Figure 4.17 (a) shows $\cos\theta_\mathrm{V}$ distribution where the DVNUCL algorithm is used to find the primary vertex. This is a cheat in the context

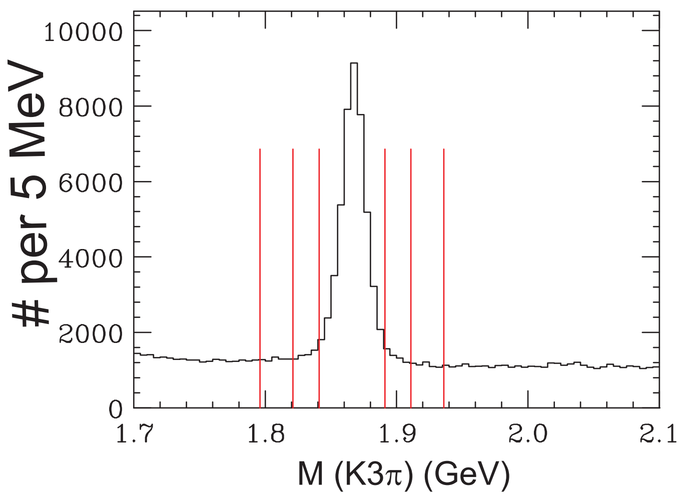

Figure 4.16: The $K^-\pi^+\pi^+\pi^-$ mass plot showing the D^0 peak and the signal and sideband regions used for the resolution study.

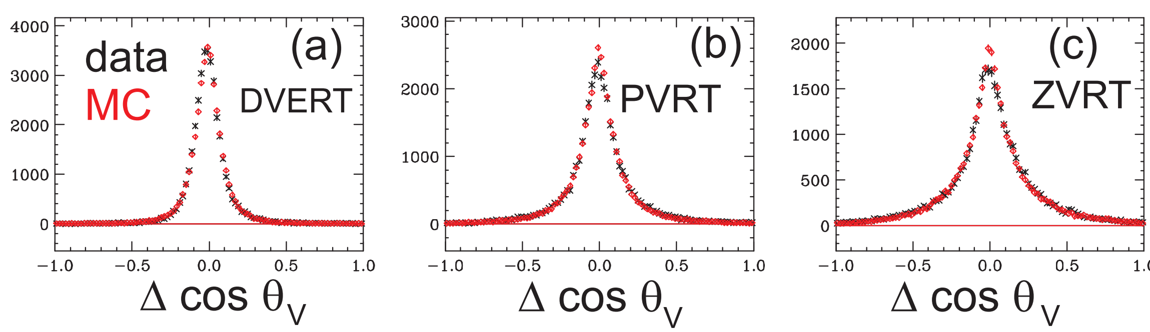

Figure 4.17: We show resolution plots or the distribution of $\Delta\cos\theta_{\rm V}$. We show the difference between $\cos\theta_{\rm V}$ reconstructed using the line-of-flight method on a "blanked" $K3\pi$ sample minus the $\cos\theta_{\rm V}$ computed from a fully reconstructed sample. The "blanked" pion representing the neutrino is the softest pion. Black is data and red is Monte Carlo. The data and Monte Carlo have been background subtracted using the sidebands illustrated in Figure 4.16. The plots use different primary vertices to compute the line of flight: (a) The DVNUCL vertex, (b) the most detached DVFREE vertex, and (c) A ZVRT vertex chosen to give the D^0 a nominal 0^o line-of-flight along the direction of the incident photon.

of neutrino closure, since the DVNUCL primary vertex uses the fully reconstructed D^0 momentum, which is unavailable in legitimate semileptonic events because of the undetected neutrino. DVNUCL and DVFREE are discussed in Section 4.2. Figure 4.17(b) shows the $\cos\theta_{\mathrm{V}}$ distribution where the PVRT algorithm is used to find the primary vertex. The PVRT vertex is the most detached of primary vertices found by the DVFREE algorithm. This is the standard primary vertex used throughout this thesis. As you can see, the realistic PVRT vertex produces a noticeably poorer resolution than one could get by using DVNUCL algorithm were that possible. Finally, Figure 4.17(c) shows what the resolution would be if one used a nominal 0^o line-of-flight for D^0 rather than measuring the D^0 line of flight from the separation of the primary and secondary vertex. We call this the ZVRT or Zero degree VeRTex.

Several conclusions are possible from Figure 4.17. One is clearly able to improve resolution by utilizing the D^0 line-of-flight from the PVRT vertex, but the resolution is compromised by the poorer primary vertex information available using DVFREE. Perhaps the most important conclusion is that in all three vertex cases the resolution plots observed in sideband subtracted data are well matched by the Monte Carlo.

Figure 4.18 shows background-subtracted resolution plots for all 4 kinematic variables. Again, we “blank” the softest (lowest momentum) pion to represent the missing neutrino. We use the most detached primary vertex to compute the D^0 line-of-flight, as usual. The Monte Carlo points are an excellent match to the observed resolution function in data.

The kinematic resolution is a strong function of the momentum of the missing neutrino. Figure 4.19 dramatically illustrates this by showing the four resolution plots where we “blank” the hardest (highest momentum) pion in the $K3\pi$ candidate. The most dramatic degradation in resolution is for $\cos\theta_\ell$. This behavior is also well modeled by the MC.

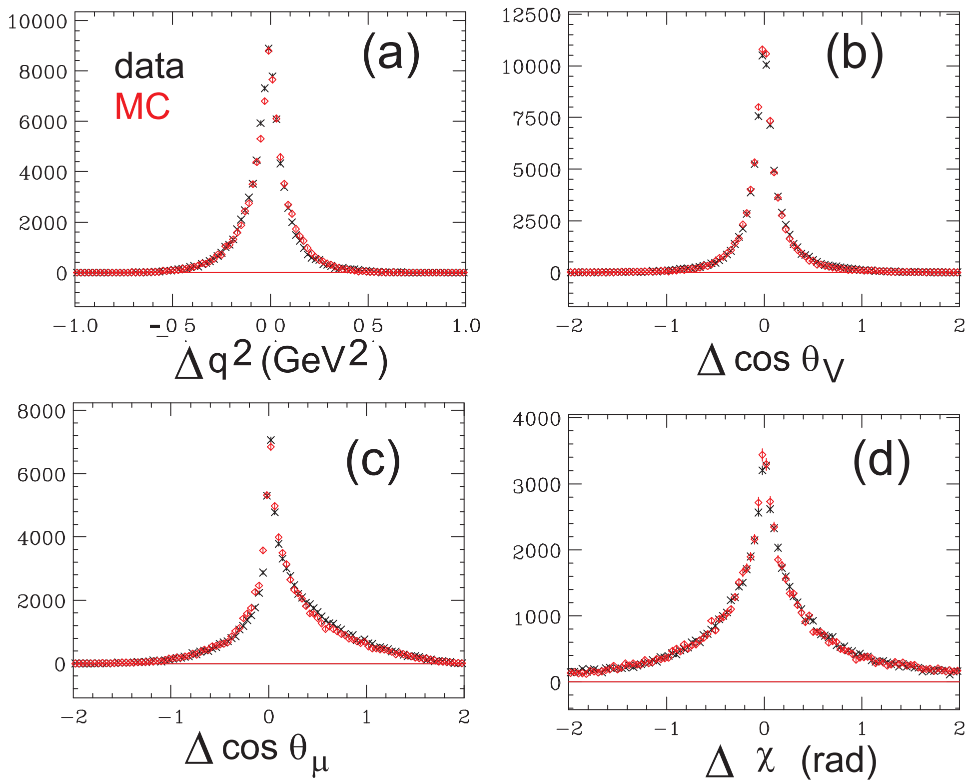

Figure 4.18: We show resolution plots observed in background subtracted data (black) to those in Monte Carlo (red). The most detached vertex is used for the time-of-flight method. The **softest pion** in the $K3\pi$ is "blanked" to represent the missing neutrino. The four plots give the the resolutions of :(a) Q^2, (b) $\cos\theta_{\rm V}$, (c) $\cos\theta_\ell$ and (d) the acoplanarity angle χ.

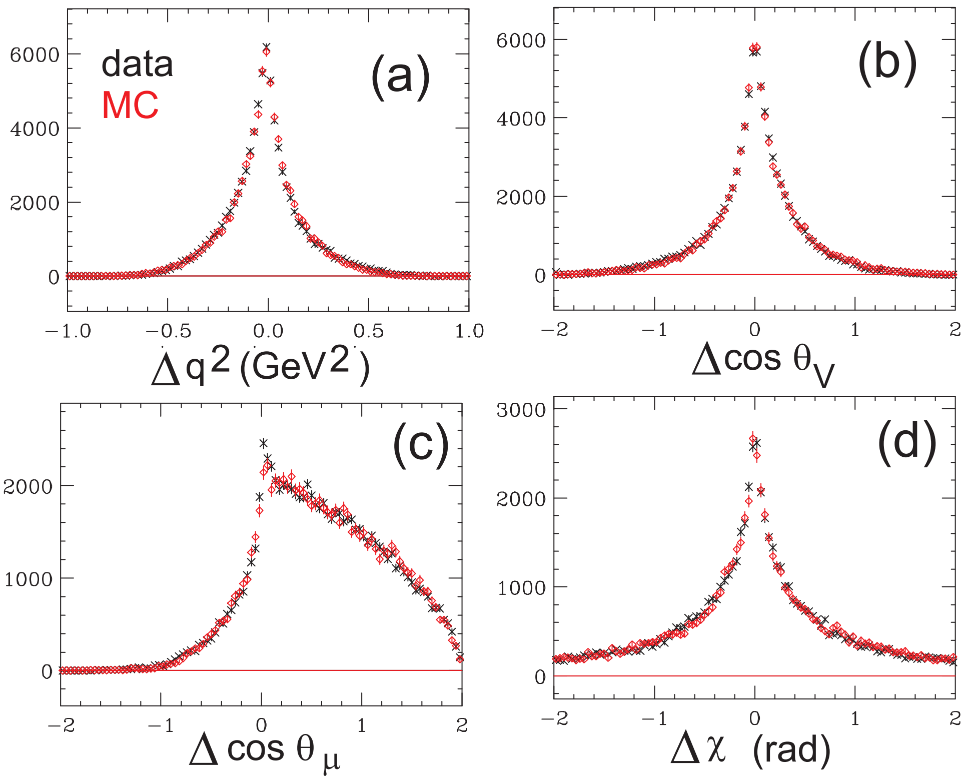

Figure 4.19: We show resolution plots observed in background subtracted data (black) to those in Monte Carlo (red). The most detached vertex is used for the time-of-flight method. The **hardest pion** in the $K3\pi$ is "blanked" to represent the missing neutrino. The four plots give the the resolutions of :(a) Q^2, (b) $\cos\theta_{\rm V}$, (c) $\cos\theta_\ell$ and (d) the acoplanarity angle χ.

4.7.3 Conclusions

We have used a background-subtracted $D^0 \to K3\pi$ sample to study the resolution on the Q^2 and the 3 decay angles using the line-of-flight method to reconstruct the kinematics of the missing neutrino in $D^+ \to K\pi\mu\nu$ events. Our technique involves "blanking" one of the pions and reconstructing its momentum using the line-of-flight technique. We then compare the decay kinematics to that deduced from fully reconstructed $D^+ \to K\pi\mu\nu$ events. We have reached the following conclusions:

1. The line-of-flight technique that uses information on the primary and secondary vertex is indeed an improvement over a method where the D^0 is given a nominal line-of-flight along the incident photon direction.

2. The DVFREE vertex gives poorer resolution than the DVNUCL vertex in these studies. Unfortunately, we need to use the DVFREE vertex in decays with missing neutrinos.

3. The resolution of kinematic variables (especially $\cos\theta_\ell$) depends dramatically on the momentum carried by the neutrino.

4. In all of these studies, the Monte Carlo resolution plots are very well matched to the observed resolution plots for all four kinematic variables.

4.8 MCFOCUS

The FOCUS Monte Carlo, called MCFOCUS, is a complete and reliable simulation of the FOCUS experiment. A complete simulated event is produced in four stages. The first stage is the generation, where the photon-gluon fusion process is simulated and decay modes are chosen as well as masses, lifetimes, and momenta. The generation handles all the particles created at the radiator and the particles produced from the photons at the target. The second stage is the simulation, in which the interaction of the particles with the rest of the experiment is simulated. The primary business of the simulation stage is to simulate the "raw" detector responses. Various types of scattering, both elastic and inelastic, are also simulated as each generated particle is "traced" through the spectrometer. The third stage is the reconstruction, which interprets the detector response in order to reconstruct the event. This includes the track finding algorithms, the Čerenkov and muon identification, and so on. The final stage is the analysis or skimming stage during which, for example, charm decay candidates would be searched for. The same reconstruction code is also used on real data. MCFOCUS is just the event generator and simulation.

In this section, we discuss the (new) organization of decay modes and describe the algorithm allowing multiple decay modes to be requested in a simulation.

4.8.1 Particle Dictionary

The MCFOCUS particle dictionary file contains all the particle properties and decay modes known to MCFOCUS. The particle properties include the mass, lifetime (or width), charge, spin, and parity.

The decay modes for each particle are grouped into ***final states***. A *final state* is the set of quasi-stable particles that are the final decay products of the charm decay. Each final state is reached through one or more ***decay paths***. *Decay paths* will describe the chain of decays, usually involving unstable particles, which lead to the given final state. The distinction between quasi-stable and unstable particles is an important one. It is somewhat arbitrary, but there are practical ramifications. Particles such as the $K^*(890)$ are certainly "unstable", but the K^+ is not. The π^0 is considered quasi-stable, as well as the K^0. K^0 decays are handled as if the particle immediately "decays" into a K_s or K_l, but for the purposes of cataloging final states, the K^0 is stable. The general idea is that no quantum-mechanical interference can arise between different final states. It's generally true that particles which can decay on the strong-interaction time scale would be considered unstable, while others would not.

A *final state* is simply an unordered list of particles, which are the final decay products. But a *decay path* is a more complex object, a tree-like structure where each particle has a specific parent and daughters. In addition, a decay path can specify a "matrix element" for each parent particle in the decay chain. This allows for the simulation of decay dynamics, including conservation of angular momentum and quantum coherence between multiple resonances.

This structure of final states and their corresponding lists of decay paths is very convenient for the event generation, and also allows gradual improvements to take place. For example, the decay $D^+ \to K^- \pi^+ \pi^+$ was originally represented by a series of decay paths. This is equivalent to an incoherent description of the decay dynamics involving various distinct resonances. Later on, the results of a careful Dalitz analysis were applied, and a matrix element was included which described the fully-coherent 3-body decay. The multiple decay paths were replaced by a single one that referenced thsi $K2\pi$ matrix element.

4.8.2 Event Generation

The event generation algorithm in MCFOCUS uses Pythia to produce a list of products from the primary vertex interaction. Imagine this as a black box: give it the 4-momentum of a charm-producing photon, and it returns the 4-momenta and "LUND" particle IDs of the resulting particles. Two of these will be charm particles, and MCFOCUS handles the decay of these and everything else.

A $c\bar{c}$ Monte Carlo is straightforward. We address the more interesting question of what to do when only certain particle decays are requested. The objective is to provide the requested decays in such a way that **the resulting distributions of all event properties are identical to what is produced if one ran $c\bar{c}$ Monte Carlo and immediately discarded all events not containing any of the decays specified**. We now will examine algorithms and evaluate their practicality and whether or not they achieve this objective.

Figure 4.20 shows a schematic for MCFOCUS that illustrates the ideal event generator for requested decays. This ideal model would not usually be practical,

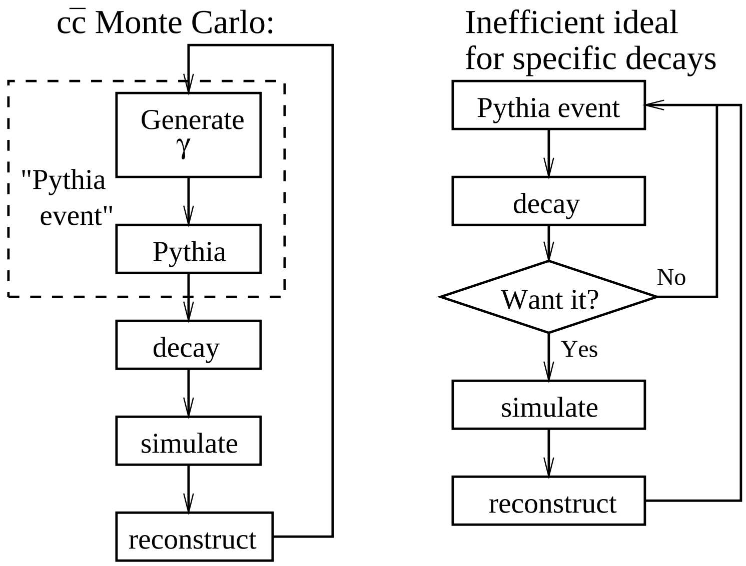

Figure 4.20: Flowcharts for MCFOCUS $c\bar{c}$ mode and the ideal for specified decays.

since it may require hundreds or even millions of Pythia events for just a single event of interest. The Pythia event generator is somewhat time-consuming, so it is necessary to use it more efficiently.

Production Rates

It is helpful to describe the average number of decays of a certain type per Pythia call given a variety of different generation methods.

We will first consider the case where one or more decays were requested for each of two different "stable" charm species. The case of requested decays from a single charm species is degenerate with this one. More complex cases involving excited charm states and more than three species will be briefly discussed later. Consider a minimal set of charm species: A, B, and C. These could represent particles such as D^+, D^0, D_s^+, Λ_c, and so on. Two of these, A and B, will be the species involved in the requested decays. C will represent all other charm species not requested. Each species will have a set of decay modes we are requesting, represented as $A \to \mu$ and

$B \to \nu$, having branching ratios $\mathcal{B}_{A\to\mu}$ and $\mathcal{B}_{B\to\nu}$.

Pythia's production model will be abstracted as a matrix (G) of probabilities for each possible pair of charm and anti-charm particles. G_{AB} will, for example, signify the probability of an $A\bar{B}$ event. In a $c\bar{c}$ Monte Carlo of N_P Pythia events, the expected number of decays $A \to \mu$ and $B \to \nu$ (including charge conjugate modes) are:

$$
\begin{aligned}
\frac{\langle N_{A\to\mu}(c\bar{c})\rangle}{N_P} &= \sum_{i=(A,B,C)} \mathcal{B}_{A\to\mu}(G_{Ai} + G_{iA}) \\
&= \mathcal{B}_{A\to\mu}\left[2G_{AA} + (G_{AB} + G_{BA}) + (G_{AC} + G_{CA})\right] \qquad (4.7) \\
\frac{\langle N_{B\to\nu}(c\bar{c})\rangle}{N_P} &= \mathcal{B}_{B\to\nu}[2G_{BB} + (G_{BA} + G_{AB}) + (G_{BC} + G_{CB})] \qquad (4.8)
\end{aligned}
$$

For the Monte Carlo where $A \to \mu$ and $B \to \nu$ were requested, we would like to be able to generate these requested decays efficiently. We ought to be able to produce these decays with fewer Pythia calls than a $c\bar{c}$ MC. The simplest method for improving efficiency was implemented in the E687 Monte Carlo, *Rogue*. Rogue allowed one decay to be specified. It would generate events through Pythia until either one of the charm particles produced was of the species requested. It would decay this particle in the mode requested, and the other one according to its branching ratio tables. For a particle with a 1% branching ratio, this method is roughly 100 times more efficient. But it is not without drawbacks. We can expand on this to allow multiple requested decays even for different charm species and correct a bias in the recoil species that it produces. Being able to specify multiple decay modes can be a great benefit for many analyses, especially those investigating specific sets of backgrounds, or for those measuring relative branching ratios.

At the beginning of the job, all requested decays modes are organized into groups with common parent charm species. All of the decay modes $A \to \mu$ are in one group, and the modes $B \to \nu$ in another. The sum of the branching ratios of each group are computed and symbolized as follows:

$$S_A = \sum_{\mu} A \to \mu \qquad S_B = \sum_{\nu} B \to \nu$$

Next, each requested species is assigned an acceptance probability, a_A and a_B for

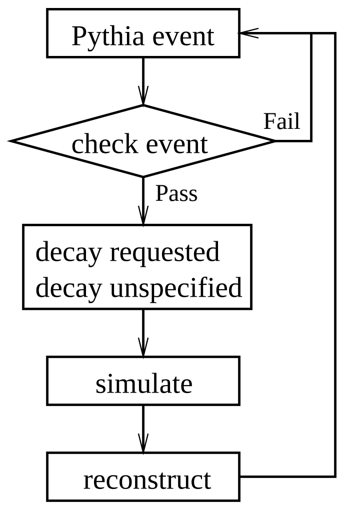

Figure 4.21: Abstract flowchart for methods (1)-(3).

this example, which is proportional to the total branching ratio S of the requested decays from that species. The charm species with the highest total requested branching ratio is assigned an acceptance probability of unity, and all other species' acceptance probabilities are scaled accordingly. For convenience here, we will choose species A to be the one with the highest branching ratio. Thus,

$$S_A \geq S_B \qquad a_A \equiv 1 \qquad a_B \equiv \tfrac{S_B}{S_A}$$

We will now consider various candidate algorithms and evaluate their performance against equations 4.7-4.8.

Method (1): Consider an algorithm that generates an event from Pythia, randomly chooses a charm particle to examine, and compares this to the list of charm species requested. If the chosen charm particle is not any of the requested species, the event is rejected and a new one is drawn from Pythia. If the chosen particle *is* a requested species, a random draw is taken against the acceptance probability a. If this acceptance challenge is passed, the successful charm particle is "forced" to decay as specified. The other (recoil) charm particle decays "naturally" — according to the particle dictionary's branching ratios.

To help clarify this procedure, Figure 4.21 shows the basic code structure of this method. The events from Pythia are tested according to these rules:

1. Decide which particle to attempt a match on. 50% chance to choose c, 50% $\bar{c}$.

2. Is the particle the right type?

    - A — keep always ($a_A = 1$)
    - B — keep with probability a_B
    - C — reject always ($a_C = 0$)

3. Decay matched particle — x decays into μ with probability $\mathcal{B}_{x\to\mu}/S_x$

4. Decay unmatched particle — x decays into γ with probability $\mathcal{B}_{x\to\gamma}$

With this algorithm, the expected number of decays per Pythia event is:

$$\begin{aligned}
\frac{\langle N_{A\to\mu}(\mathrm{M1})\rangle}{N_P} &= G_{AA}\left(\frac{\mathcal{B}_{A\to\mu}}{S_A} + \mathcal{B}_{A\to\mu}\right) + \frac{1}{2}\frac{\mathcal{B}_{A\to\mu}}{S_A}(G_{AC}+G_{CA}) \\
&\quad + (G_{AB}+G_{BA})\left(\frac{1}{2}\frac{\mathcal{B}_{A\to\mu}}{S_A} + \frac{1}{2}a_B\mathcal{B}_{A\to\mu}\right) \\
&= \left(\frac{1}{2S_A}\right)\mathcal{B}_{A\to\mu}\Big[2(1+S_A)G_{AA} \\
&\quad + (1+S_B)(G_{AB}+G_{BA}) + (G_{AC}+G_{CA})\Big] \qquad (4.9)
\end{aligned}$$

$$\begin{aligned}
\frac{\langle N_{B\to\nu}(\mathrm{M1})\rangle}{N_P} &= G_{BB}a_B\left(\frac{\mathcal{B}_{B\to\nu}}{S_B} + \mathcal{B}_{B\to\nu}\right) + \frac{1}{2}\frac{\mathcal{B}_{B\to\nu}}{S_B}(G_{BC}+G_{CB}) \\
&\quad + (G_{BA}+G_{AB})\left(a_B\frac{1}{2}\frac{\mathcal{B}_{B\to\nu}}{S_B} + \frac{1}{2}\mathcal{B}_{B\to\nu}\right) \\
&= \left(\frac{a_B}{2S_B}\right)\mathcal{B}_{B\to\nu}\Big[2(1+S_B)G_{BB} \\
&\quad + (1+S_A)(G_{BA}+G_{AB}) + (G_{BC}+G_{CB})\Big] \qquad (4.10)
\end{aligned}$$

We will consider the first of these factors to see the reasoning behind all of them. $G_{AA}\left(\frac{\mathcal{B}_{A\to\mu}}{S_A} + \mathcal{B}_{A\to\mu}\right)$ is the contribution from those events where $A\bar{A}$ were generated. For each event of this type, one of the charm particles will be chosen and will match. Of those matched particles, $\frac{\mathcal{B}_{A\to\mu}}{S_A}$ will decay into μ. Each of these events will also have an unmatched A which will decay into μ with probability $\mathcal{B}_{A\to\mu}$.

Note the prefactors are identical: $\left(\frac{a_B}{2S_B}\right) = \left(\frac{1}{2S_A}\right)$. Comparing equations 4.9-4.10 with 4.7-4.8, we see two differences. One difference is the universal scaling factor of $\left(\frac{1}{2S_A}\right)$, which actually represents the increase in efficiency over a $c\bar{c}$ Monte Carlo. The other differences are the undesirable factors $(1 + S_A)$ and $(1 + S_B)$. Note that these factors are usually unimportant since S_A and S_C are typically small. Even so, we can exactly compensate for these by introducing an additional rejection step based on the unmatched (recoil) charm particle.

Method (2): We can achieve the right decay populations if we include an additional probability to reject an event based on the recoil charm species. Here, we add a second acceptance challenge for each event. During this new step, the probability to accept an event is $1/(1 + S_x)$, where the recoil particle is x. If the recoil is not a requested parent ($x = C$), the event always passes the second rejection step. $S_C = 0$, so $1/(1+S_C) = 1$. The motivation for this additional rejection lies in the algebra, where the undesirable terms all miraculously cancel. Again, we will only look at $A \to \mu$, since $B \to \nu$ is essentially the same except for a few extra factors of a_B.

$$\begin{aligned}
\frac{\langle N_{A\to\mu}(\mathrm{M2})\rangle}{N_P} &= G_{AA}\frac{1}{1+S_A}\left(\frac{\mathcal{B}_{A\to\mu}}{S_A} + \mathcal{B}_{A\to\mu}\right) + \frac{1}{2}\frac{\mathcal{B}_{A\to\mu}}{S_A}(G_{AC} + G_{CA}) \\
&\quad + (G_{AB} + G_{BA})\frac{1}{2}\left(\frac{\mathcal{B}_{A\to\mu}}{S_A}\frac{1}{1+S_B} + a_B\frac{1}{1+S_A}\mathcal{B}_{A\to\mu}\right) \\
&= \left(\frac{1}{2S_A}\right)\mathcal{B}_{A\to\mu}\Big[2G_{AA} \\
&\quad + (G_{AB} + G_{BA}) + (G_{AC} + G_{CA})\Big]
\end{aligned} \tag{4.11}$$

These additional rejection factors turn out to exactly cancel the unwanted terms in equations 4.9-4.10, leaving the correct expected number of decays for each mode. We are not able to show how every conceivable distribution matches that of a $c\bar{c}$ Monte Carlo, but the fact that this particular distribution matches piece-by-piece provides a great deal of confidence.

To test this algorithm, once implemented, we ran various generation-only Monte Carlos and counted the number of Golden Mode charm decays. One run was a $c\bar{c}$ run with 100,000,000 events, the next two were specific Monte Carlos requesting both

$c\bar{c}$ MC:	$D^+ \to K^-\pi^+\pi^+$	3388485
	$D^0 \to K^-\pi^+$	4475548
	$N_{K^-\pi^+}/N_{K^-\pi^+\pi^+}$	1.32081 ± 0.00095
old specified MC:	$D^+ \to K^-\pi^+\pi^+$	2208493
	$D^0 \to K^-\pi^+$	2956805
	$N_{K^-\pi^+}/N_{K^-\pi^+\pi^+}$	1.3388 ± 0.0012
	difference	11.8σ
new specified MC:	$D^+ \to K^-\pi^+\pi^+$	5812036
	$D^0 \to K^-\pi^+$	7685748
	$N_{K^-\pi^+}/N_{K^-\pi^+\pi^+}$	1.32238 ± 0.00073
	difference	1.3σ

Table 4.2: Monte Carlo event generation test results

$D^+ \to K^-\pi^+\pi^+$ and $D^0 \to K^-\pi^+$ decays. These decay modes are common, easy to count, and have very different branching ratios (9% and 3.85% respectively) which tends to exacerbate any problems. For comparison, we generated one data set with the original algorithm, which is equivalent to method (1) but with an additional step. This extra step tried to match the recoil charm particle to the requested decay if the first chosen charm particle failed to match. This was a bad idea since it introduces some additional biases. Finally, we generated a data set with the new, correct algorithm representing Method (2). The results, summarized in table 4.2, show that the proposed algorithm is consistent with a $c\bar{c}$ Monte Carlo to at least the 0.1% level. The original algorithm was off by about 1.5%.

Chapter 5

The $D^+ \to \overline{K}^{*0} \ell^+ \nu_\ell / K\pi\pi$ Relative Branching Ratio

In this chapter we present new, precise measurements of the relative branching ratio[1] of

$$\mathrm{BR} = \frac{\Gamma(D^+ \to \overline{K}^{*0} \mu^+ \nu \to (K^- \pi^+) \mu^+ \nu)}{\Gamma(D^+ \to K^- \pi^+ \pi^+)}$$

There are several motivations for this work. Besides providing useful "engineering" numbers for other experiments that count charm particles through their semileptonic decays to electrons or muons, the semileptonic width sets the scale of the $a_1(0)$ form factor and therefore provides a further test of Lattice Gauge calculations.[2] Such tests of the overall form factor scale are particularly relevant to the determination of CKM matrix element ratio $|V_{cu}/V_{cs}|^2$, which is related to (for example) $\Gamma(D^+ \to \rho \mu^+ \nu)/\Gamma(D^+ \to \overline{K}^{*0} \mu^+ \nu)$.

[1] We will always report on the relative BR for the $\overline{K}^{*0}$ to decay via $\overline{K}^{*0} \to K^- \pi^+$. One needs to multiply the BR we are reporting by a factor of 3/2 to obtain the full $D^+ \to \overline{K}^{*0} \mu^+ \nu$ including all K^{*0} decays: $\overline{K}^{*0} \to \bar{K}_0 \pi^0$ as well as $\overline{K}^{*0} \to K^- \pi^+$

[2] The absolute $K^- \pi^+ \pi^+$ branching fraction multiplies the $K^{*0} \mu\nu / K\pi\pi$ relative ratio to produce an absolute branching fraction for $K^{*0} \mu\nu$. This can be converted to a decay width by dividing by the known D^+ lifetime. The width is proportional to the CKM matrix element $|V_{cs}|^2$, the $|a_1(0)|^2$ form factor and an expression involving phase space integrals over the intensity shape that depends on the R_V, R_2 and R_3 form factor ratios.

In the context of this thesis, a principal motivation is to demonstrate that our Monte Carlo does an adequate job at simulating the acceptance and analysis cut response for the $D^+ \to \bar{K}^* \mu^+ \nu$. Such a demonstration is a critical ingredient to the form factor ratio measurement and shows that the $\cos\theta_V$ anomaly discussed in Chapter 8 is not an artifact of backgrounds or mismodeled acceptance. In particular, stability of the branching ratio over a wide range of cuts of varying efficiencies and associated background levels will be demonstrated.

5.1 The SLEPNRM skim

The data used for this analysis were taken from a special **S**emi**LEP**tonic **N**o**RM**alizing skim. This skim was designed to select semileptonic decays in the same skim as "normalizing" hadronic decay modes. The normalizing modes are decays with well-measured branching fractions such as $D^+ \to K^-\pi^+\pi^+$, $D_s^+ \to \phi\mu^+\nu$, and $D^0 \to K^-\pi^+$.

This skim was principally designed (by the University of Illinois FOCUS group) to minimize any controllable sources of systematic error. The SLEPNRM skim was a sub-skim of the global vertex skim. By selecting the numerator as well as the denominator sample simultaneously rather than relying on separate skims, one avoids the inevitable complication that the two quotient samples represent slightly different data sets owing to tape mishandling, processing crashes, etc. We also tried to use similar analysis cuts for both samples in order to cancel common systematics to the extent practical. We did not use any cuts requiring a primary vertex, so that neither DVFREE or DVNUCL were needed at the skimming stage. This allowed for the possibility of using either primary vertex algorithm (or a new one altogether) at the time of analysis. For the muon part of the semileptonic states, we simply required that the muon verticized with another linked track with a confidence level exceeding 1×10^{-4}. For the normalizing $D^+ \to K^-\pi^+\pi^+$ mode, we required a 3 -track vertex to the same confidence level and a broad mass cut of $1.7 < M_{K\pi\pi} < 2.1\ GeV$.

We also applied fairly soft Čerenkov cuts to the hadronic normalizing sample: the kaon track passed a kaonicity cut $W_\pi - W_K > 1$, while the pion passed a piconicity cut of $W_{\min} - W_\pi > -7$. In order to reduce number of candidates to a tolerable level, we demanded that the $D^+ \to K^-\pi^+\pi^+$ was either out of the target material by 3 σ or had an isolated secondary vertex.[3] The entire SLEPNRM sample resided on less than 300 GB of tape storage.

5.2 Analysis Cuts

Our lowest-level cuts (baseline cuts) for the $K^*\mu\nu$ sample were the following:

1. A secondary to primary vertex separation of $\ell/\sigma > 5$. The primary vertex was found for both the $K^*\mu\nu$ and $K\pi\pi$ sample using DVFREE.

2. The muon was required to be an inner muon with a confidence level greater than 5%, a momentum greater than 10 GeV/c, and at most two missing hits in the inner muon array.

3. The kaon was required to have a kaonicity exceeding 1; the pion was required to have a pionicity exceeding 1.

4. In order to suppress backgrounds from $D^{*+} \to \pi^+(K^-\mu^+\nu)$ we required $\Delta m = M(K^-\mu^+\nu\pi^+) - M(K^-\mu^+\nu) > 180$ MeV/c^2. The ν kinematics were reconstructed using the line-of-flight closure technique described in Section 4.6. This cut was applied since the D^{*+} background creates a pronounced reflection. This reflection greatly complicates the K^* fit to the $K\pi$ mass spectra.

5. We required that the event was admitted to the sample via the hadron calorimeter trigger rather than less-reliable triggers such as the dimuon trigger.

[3] The ISO2 cut required that no additional track formed a 4 track vertex with the $K^-\pi^+\pi^+$ candidate with a confidence level exceeding 1%.

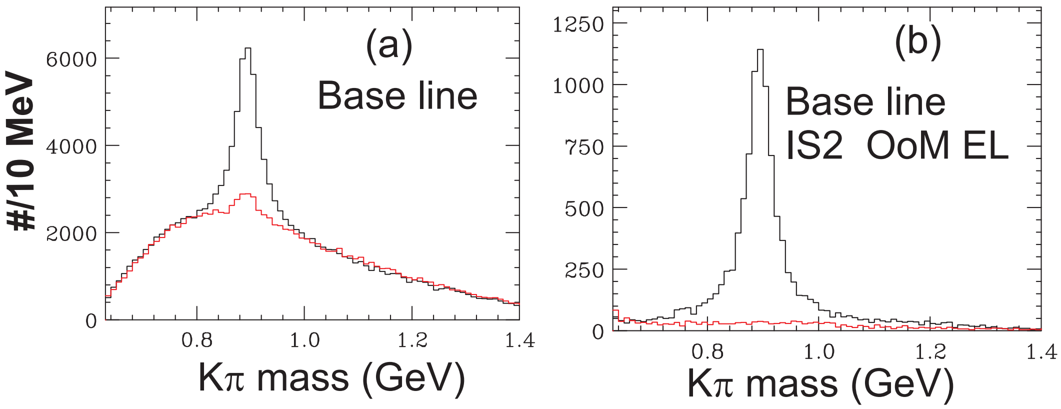

Figure 5.1: We show two $D^+ \to \overline{K}^{*0}\ell^+\nu_\ell$ signals. (a) The $K^-\pi^+$ mass spectra subjected to baseline cuts. (b) The $K^-\pi^+$ mass spectra subjected to baseline, detachment, out-of-material, and secondary isolation cuts. The RS signal is plotted in black and WS signal is plotted in red.

The muon selection cuts were fairly stiff, designed to keep muon misidentification backgrounds below the 0.5% level. Apart from these muon cuts, the baseline requirements on the $D^+ \to K^-\pi^+\pi^+$ normalizing sample were identical to those used for $K^*\mu\nu$. To further reduce systematics by forcing similar decay topologies, we required that one of the pions in the $D^+ \to K^-\pi^+\pi^+$ candidate was a 5-chamber track, since there was a 5-chamber track requirement on the muon in $K^*\mu\nu$ candidates.

Table 5.1: Additional cuts designed to remove backgrounds

Cut	Description	Value
is2	secondary isolation	CL $< 1 \times 10^{-3}$
el	primary secondary detachment	$\ell/\sigma > 20$
OoM	Out of Material	$Z_v - Z_{\text{edge}} > 3\sigma$

In order to assess systematic uncertainties, we analyzed our data using many different additional cut selections designed to probe a wide range of signal purity and specifically search for known, potential systematic problems. Table 5.1 gives a summary of those cuts designed primarily to eliminate backgrounds. Requiring more stringent detachment cuts and/or out-of-material cuts should substantially eliminate

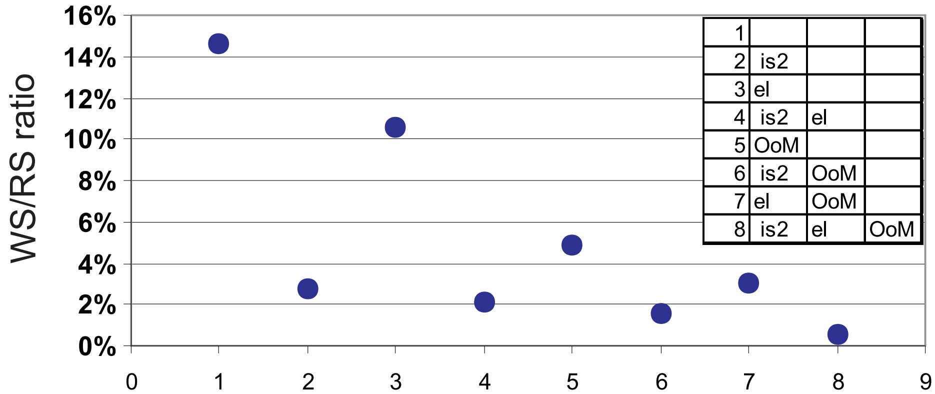

Figure 5.2: This figure shows the relative cleanliness of the $D^+ \to \overline{K}^{*0}\ell^+\nu_\ell$ signal when subjected to the some of the cuts described in Table 5.1. We plot the ratio of the wrong-sign (WS) to right-sign (RS) K^* yield where in RS events the kaon has the opposite charge as the muon, and in WS events the kaon and muon have the same charge.

non-charm backgrounds. Our most effective background reducing cut was a cut on isolation of the secondary vertex. Such a cut should reduce feed down from higher multiplicity charm decays that will generally have extra tracks in the secondary vertex. We are a bit mystified why the cut is as effective as it is in reducing backgrounds in the data. Figure 5.1 shows the baseline signal and the signal after all three additional clean up cuts are applied. Figure 5.1 (b) is particularly clean. There is very little wrong sign signal. When the wrong sign signal is subtracted, there is no evidence for a nonresonant contribution (or any non-$K^*(890)$ resonance) in the $K^-\pi^+$ spectrum in our $D^+ \to K^-\pi^+\mu^+\nu$ candidates.

Figure 5.2 gives an indication of the signal purity by plotting the ratio of the K^* signal yield in WS over RS events. This yield is based on fitting the $K\pi$ mass spectra to a Breit-Wigner line shape over a polynomial background. Our line shape is described in the next section. Figure 5.2 shows that the cut primarily responsible for eliminating WS K^* backgrounds is the secondary isolation cut.

	indiv eff	product	all cuts
baseline	100%		
is2	85%	85.0%	85%
el20	68%	57.8%	58.6%
OoM	55%	31.8%	31.4%
HC	36%	11.5%	11.0%
RN	64%	7.4%	7.1%
Wide	80%	5.9%	5.4%
Etrack	41%	2.4%	2.7%

Table 5.2: This Table shows the efficiency (relative to baseline)for the cuts explored in the $K^*\mu\nu/K\pi\pi$ branching fraction analysis. We show the individual efficiencies as well as the efficiency as each cut is sequentially applied. Finally, we show the accumulated product of the cuts.

5.3 Analysis of the Line Shape

It is important to have an adequate fit to the $K^*(890)$ line shape in order to get an accurate estimate of the $D^+ \to K^*\mu^+\nu$ yield. Rather than worrying about line shapes suggested in the literature, we opted to use the line shape that best fits our experimental data. In particular we fit to line shapes of the relativistic form :

$$\frac{dN}{dm} \propto \frac{\Gamma}{(m^2 - m_o^2)^2 + (m\Gamma)^2} \tag{5.1}$$

As an expedient, we did not correct for potential mass dependent efficiency variation in these fits, since our primary use of the fit is event counting.

A variety of forms appear in the literature concerning the mass dependence of the width Γ. We fit to the form:

$$\Gamma = \Gamma_o \left(\frac{P}{P_o}\right)^N \tag{5.2}$$

where P_o is the momentum of the kaon (or pion) in the K^* rest frame when the $K\pi$ mass is at its resonant value (m_o), and P is the center of mass momentum at a $K\pi$ mass of m. In our fits, m_o, Γ_o and the power N are free parameters.[4] In certain

[4]We also allow for an offset and linear term to model backgrounds that are not eliminated through

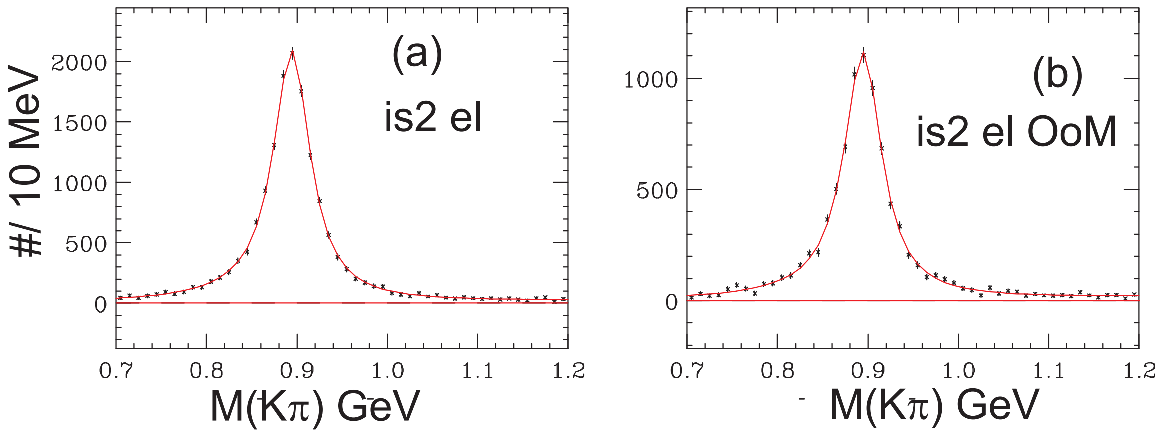

Figure 5.3: This figure compares our observed RS-WS subtracted $K^-\pi^+$ mass spectra to our fit for two different (clean) signal selections.

treatments the power N is assumed to be $2\ell+1 = 3$ for a spin one K^*. The expected values of the other parameters are $m_o = 0.896\ GeV$ and $\Gamma_o = 0.051\ GeV$.

Table 5.3: K^* line shape fit parameters

fit	m_o (MeV)	Γ_o (MeV)	N	Confid Lev
(a)	894.7 ± 0.3	50.1 ± 0.86	-0.14 ± 0.12	97%
(b)	895.0 ± 0.4	51.1 ± 1.17	-0.10 ± 0.17	6.2%

Figure 5.3 and Table 5.3 show the results of our fits to the $K\pi$ mass spectra for two of our clean $K^*\mu\nu$ spectra. In both cases, we obtain good-to-excellent quality fits (judging from the confidence level deduced from the χ^2 of the fits) with width and mass parameters consistent with the known K^* parameters. Figure 5.4 gives "likelihood" contours for the three shape parameters m_o, Γ_o, and the power N. These contours are plots of χ^2 as a function of each fit parameter as all other fit parameters are held constant. In these fits, the contours appear parabolic with a χ^2 change of one unit occurring at roughly 1σ away from the minimizing fit parameter.[5]

the RS - WS subtraction procedure. The illustrated fits results varied slightly (primarily in width) when these "background" terms are left out.

[5]The parabolic contour suggest a nearly Gaussian errors. The shift of one χ^2 unit by varying a

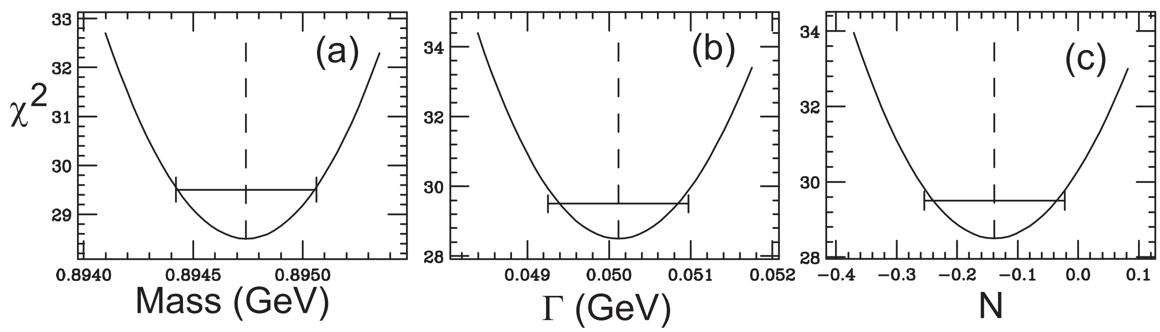

Figure 5.4: This figure shows χ^2 as a function of the mass, width, and width power for the fit shown in 5.3 (a). The horizontal bar is 1 unit from the χ^2 minimum and the flats are located at $\pm 1\sigma$.

Interestingly enough, the power in both cases is consistent with zero — implying a constant rather than energy dependent width. This result is not expected, but no experiment has ever approached our statistics and cleanliness for $\bar{K}^*$'s produced via D^+ decay. It is possible (albeit unlikely) that the true K^* line shape is a p-wave Breit-Wigner with $N = 3$, but $K\pi$ mass dependent efficiency variation or influence of the $D^+ \to K^*\mu\nu$ matrix element transform an essentially p-wave Breit-Wigner into a nearly perfect observed s-wave line shape.[6] Alternatively, the mysterious shape transformation could be due to interference with the same (s-wave?) amplitude responsible for the $\cos\theta_V$ anomaly described in Chapter 8. Although we have no ready explanation for the unexpected simplicity in the K^* line shape, our fits indicate that it is quite adequate for the "counting" purposes used in the present analysis.

given parameter by $\pm 1\sigma$ while holding the other parameters constant implies that it is only weakly correlated with the others.

[6]Certainly the $D^+ \to K^*\mu\nu$ is known and used throughout this work. The problem is that it has always been calculated in the quasi-stable limit and the Breit-Wigner line shape has therefore been "grafted on" as a factor in order to fit experimental data.

5.4 Methodology

In the process of performing this analysis, we learned that the model used for the simulation of the hadron calorimeter trigger produced a major systematic error on the $\Gamma(D^+ \to \bar{K}^* \mu^+ \nu)/\Gamma(D^+ \to K^- \pi^+ \pi^+)$ branching ratio. Fortunately we were able to develop a method (that we will call the *ETD* method) to control this systematic to a level commensurate with other systematic sources. We begin by describing those features of our method common to both approaches before describing the ETD method.

1. A large Monte Carlo sample is generated containing both $D^+ \to K^- \pi^+ \pi^+$ and $D^+ \to \bar{K}^* \mu^+ \nu$ events in their approximate known branching ratio. In both cases the known (highly structured) decay intensity is used for the simulation. The $D^+ \to \bar{K}^* \mu^+ \nu$ decay intensity used at generation employed the previous world average values of the form factor ratios.

2. We then fit the $D^+ \to K^- \pi^+ \pi^+$ mass peak to a Gaussian over a polynomial background, and we separately fit the $K\pi$ spectra for right sign and wrong sign $\bar{K}^* \mu^+ \nu$ combinations to an s-wave Breit-Wigner over a polynomial background. We next form the ratio of RS - WS subtracted K^* yields to the fitted $D^+ \to K^- \pi^+ \pi^+$ yield in both data and MC. These ratios are obtained for every cut set. Apart from the muon identification and mass difference cut, the same cuts are applied to the $K^- \pi^+ \pi^+$ and $\bar{K}^* \mu^+ \nu$ candidates.

3. Our measured value for $\Gamma(D^+ \to \bar{K}^* \mu^+ \nu)/\Gamma(D^+ \to K^- \pi^+ \pi^+)$ was the ratio of $\bar{K}^* \mu^+ \nu / K^- \pi^+ \pi^+$ yields obtained in the data, divided by that in the Monte Carlo sample, times the assumed relative BR used in the Monte Carlo.

Comparison of the $\Gamma(D^+ \to \bar{K}^* \mu^+ \nu)/\Gamma(D^+ \to K^- \pi^+ \pi^+)$ values for the typically 64 cut sets was then used to get an estimate of the systematic error. When this analysis was initially performed with our default shower library based on GEANT simulations (SHW), we obtained remarkably consistent relative branching ratios over a large range of cuts but obtained values that were considerably higher than the world average branching ratio of ≈ 0.36. What was more disturbing was that we found a significant reduction in the branching ratio when an alternative "crude" parameterization of the hadrometer response was used (CRD). The branching fractions are compared for 8 cuts in Table 5.4.

Table 5.4: Initial $\Gamma(D^+ \to \bar{K}^* \mu^+ \nu)/\Gamma(D^+ \to K^- \pi^+ \pi^+)$ measurements

Cut	CRD	SHW
base	0.408±0.008	0.475±0.009
iso2	0.434±0.005	0.508± 0.005
el20	0.426±0.007	0.491±0.008
iso2,el20	0.407±0.006	0.509±0.006
OoM	0.454±0.006	0.48±30.007
OoM,iso2	0.422±0.007	0.505±0.007
OoM,el20	0.417±0.007	0.491±0.008
OoM,iso2,el20	0.426±0.007	0.506±0.008

The shower library parameterization produced relative branching ratios that were typically 15 - 17% higher than those of the crude parameterization. We thought this shift was unacceptably large given that our statistical error bars were typically 1 – 2%.

5.4.1 The ETD method

It is not that surprising that the HC calorimeter trigger simulation is likely to be a major source of errors on $\Gamma(D^+ \to \bar{K}^* \mu^+ \nu/\Gamma(D^+ \to K^- \pi^+ \pi^+)$, since the semileptonic state deposits far less energy in the hadrometer than the hadronic state (since in semileptonic decays neither the neutrino nor the muon will leave much energy). It is not that easy to obtain a precise model for calorimetric response on charm events. It is relatively easy to measure the response on unbiased data, but there is no guarantee that a charm event of a given visible energy will have the same response as unbiased data of the same energy. For example it is conceivable that charm data may have a larger fraction of unobserved K_ℓ^o or neutrons produced in charm events than the bulk of unbiased photoproduced events (that are dominated perhaps by vector meson diffractive dissociation).

The method we developed was to bin our data in bins of a variable that we will call ETD. ETD is the total visible momentum of the D^+ secondaries that are 5 chamber tracks which extrapolate to the calorimeter.

1. We compute the acceptance corrected yield of $\Gamma(D^+ \to \bar{K}^* \mu^+ \nu / \Gamma(D^+ \to K^- \pi^+ \pi^+)$ within each ETD bin to cancel HC systematics.

2. We can then make (inverse variance) weighted averages of these individual bin branching ratios in order to preserve statistical accuracy.

In computing the ETD for the semileptonic state, we have considered two ETD variables: (1) the muon is assigned zero hadronic energy or, (2) a small nominal (constant) dE/dx energy loss.

The Monte Carlo studies summarized in Figure 5.5 serves to illustrate the validity of the ETD method by showing the simulated trigger efficiency for our two models and the two quotient D^+ decay modes. Although the trigger efficiency is substantially higher in the CRD simulation compared to that in the SHW simulation, the trigger efficiencies for $K\pi\pi$ and $K^*\mu\nu$ within the same ETD bin are nearly identical within a given model. One might expect this behavior since the calorimeter responds to the energy striking it from the D^+ secondaries, along with those from the charm recoil and primary vertex fragmentation products. To the extent that the recoil charm and primary vertex tracks have roughly the same momenta for the $D^+ \to K^- \pi^+ \pi^+$ and $D^+ \to \bar{K}^* \mu \nu$ events of equal ET, the trigger efficiency will be the same as evidenced by Figure 5.5.

The hadronic trigger fires in a narrow range of actual deposited energy at a threshold of about 25 GeV — the value where the triggering efficiency is 50%. The triggering efficiency when expressed as a function of ETD has a much more broad distribution (shown in Figure 5.5) because of fluctuations in the recoil charm and primary vertex

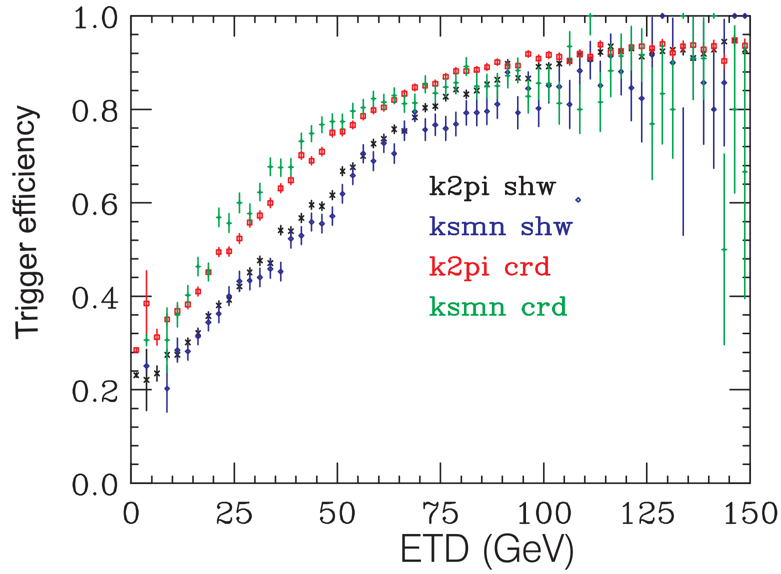

Figure 5.5: We show the HC trigger efficiency as a function of the total track momenta from D^+ secondaries (ETD) striking the calorimeter system for two different HC models and our two states $D^+ \rightarrow K^-\pi^+\pi^+$ and $D^+ \rightarrow \bar{K}^*\mu\nu$ (marked ksmn) in the figure.

cuts, and the large fluctuations in energy lost in the inner electromagnetic calorimeter that lies immediately upstream of the hadrometer.

The ETD method has another more subtle benefit — the elimination of need for "post-hoc" corrections to insure that the MC matches important variables that control the two mode's relative acceptance such as the momentum distribution and multiplicity of the primary vertex. Although our Pythia-based Monte Carlo generation model did a fairly good job at matching momenta, P_T, and the primary vertex multiplicity and primary track momenta spectra, we have not been able to find a perfect tune of Pythia parameters to match all observable distributions. In particular, the tunes that best match momenta spectra tend to predict a higher primary vertex distribution than observed in background subtracted charm events. One can force agreement between (for example) the generated and observed momentum spectra for a given charm particle by a "post hoc" rejection based on keeping events with a probability proportional to the ratio of the observed to simulated momenta spectra. The branching fractions presented in Table 5.4 were subjected to an additional rejection step based on the product of a multiplicity dependent and momentum dependent post-hoc rejection. This adjustment to the simulated momenta and primary vertex distribution shifted the $\Gamma(D^+ \to \bar{K}^* \mu^+ \nu / \Gamma(D^+ \to K^- \pi^+ \pi^+)$ value by about 9% in the SHW simulation and by about 6% in the CRD simulation.

It would be difficult to obtain separate post-hoc corrections in every bin of ETD, but fortunately we found that it was not necessary for variables such as visible momenta. It is easy to understand why. If the Monte Carlo produces the correct decay kinematics in the D^+ rest frame, it should be able to predict the fraction of the D^+ energy in the form of 5-chamber tracks and thus the average ETD for a given bin of total D^+ momenta. Turning this relationship around, for a given bin of ETD, one should have a correct prediction for the average D^+ momentum or any other variables (such as the visible energy) that depends on the D^+ momenta and decay kinematics. Hence to the extent that the decay kinematics are well modeled, one is essentially forced to get the D^+ kinematics correct in each ETD bin. One can then choose to run the Monte Carlo simulations using Pythia parameters that best model the observed primary vertex multiplicity to get all relevant variables correct.

Figure 5.6 demonstrates the validity of this argument by showing the average visible momenta and primary vertex multiplicity in bins of ETD for the $D^+ \to \bar{K}^* \mu^+ \nu$ events for both the CRD and SHW Monte Carlos to the averages observed in background subtracted data for our base line data sample. These plots are obtained by averaging the variables weighted by a factor of $+1$ for RS events and -1 for WS events for events in the K* region ($0.8 < M_{K\pi} < 1.0$ GeV/c^2).

Except for the two lowest ETM bins, where signal to noise is particularly poor in data (see Figure 5.7), there is excellent agreement in the multiplicity and average visible momentum. Agreement in other $D^+ \to \bar{K}^* \mu^+ \nu$ variables such as the total event multiplicity, and the fraction of the D^+ tracks that are 5-chamber (as opposed to 3-chamber), are equally impressive as are the ETD binned averages for (mass) sideband subtracted $D^+ \to K^- \pi^+ \pi^+$ candidates.

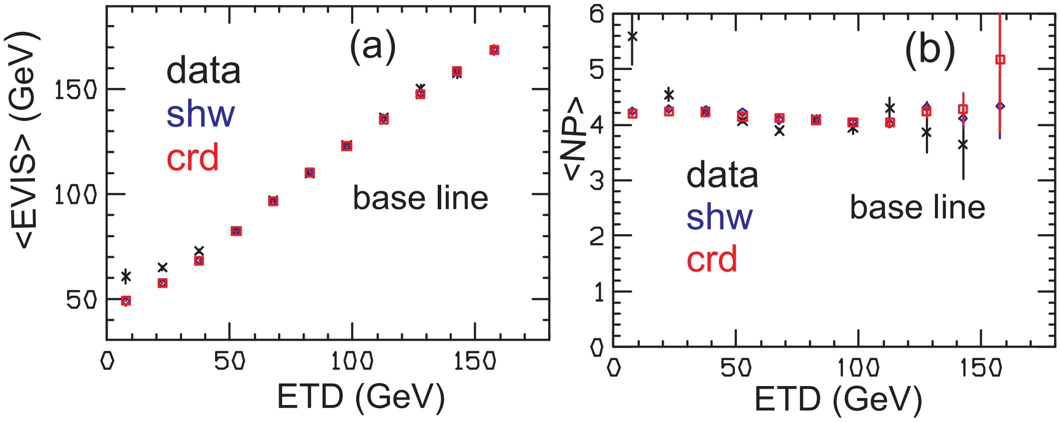

Figure 5.6: We compare average properties of the (base line) $D^+ \to \bar{K}^*\mu\nu$ signal in data and in our two MC models in bins of ETD: (a) Compares the average visible energy (the energy of the kaon pion and muon) (b) Compares the multiplicity of the primary vertex.

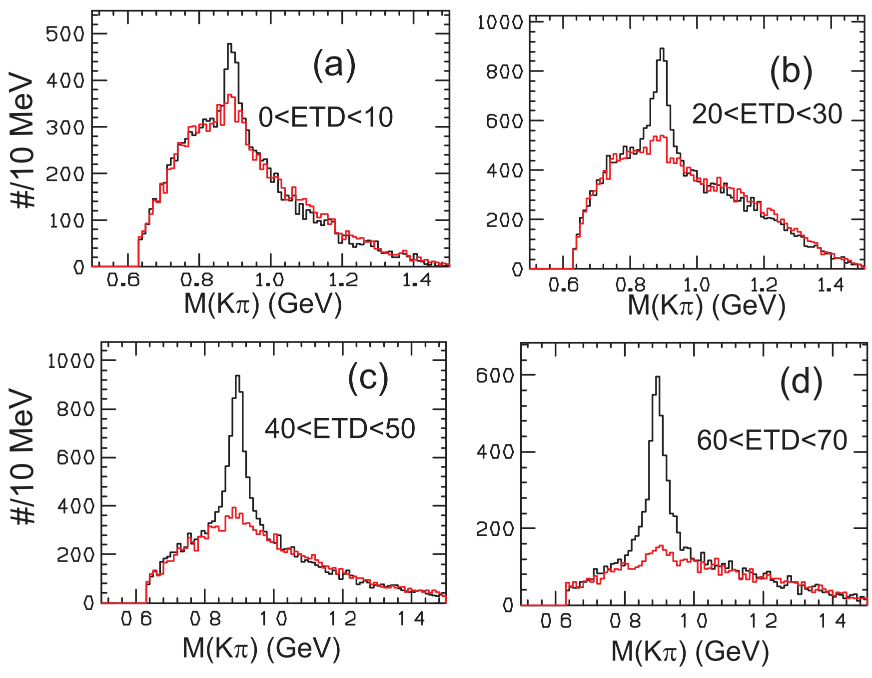

Figure 5.7: We show $K\pi$ mass spectra from 4 $D^+ \to K^*\mu\nu$ samples selected with base line cuts in 4 different ETD ranges. a) $0 < ETD < 10\ GeV$. (b) $20 < ETD < 30\ GeV$ (c) $40 < ETD < 50\ GeV$. (d) $60 < ETD < 70\ GeV$. The RS signal is plotted in black and WS signal is plotted in red.

As shown in Figure 5.7, the signal to background improves dramatically as the ETD bin is increased. The lowest ETD bins are also the most sensitive to the triggering threshold model that is designed to fire at a nominal energy of about 25 GeV. In light of these considerations, we felt the most reliable results were based on the weighted average of the seven 10 GeV bins that start beyond 30 GeV. The choice of averaging range will be considered as a source of systematic error.

5.5 Results and Comparison to Previously Published Data

Figure 5.8 compares our measurements of the $\Gamma(D^+ \to \bar{K}^* \mu^+ \nu)/\Gamma(D^+ \to K^- \pi^+ \pi^+)$ relative branching ratio for 32 cut selections using the ETD method to the five previously published measurements from CLEO[17], the Omega spectrometer [18], ARGUS[19], E687 [7], and E653[20]. The 64 measurements follow by considering all combinations (2^5) of the five cuts summarized in Table 5.5 as well as our two (extreme) HC models: CRD and SHW.

Table 5.5: Additional cuts designed to remove backgrounds

	Cut	Description	Value
1	IS2	secondary isolation	$\mathrm{CL} < 1 \times 10^{-3}$
2	EL20	primary secondary detachment	$\ell/\sigma > 20$
3	OoM	Out of Material	$Z_v - Z_{\mathrm{edge}} > 3\sigma$
4	RN	run number	$9875 < \text{run \#} < 13815$
5	Wide	out of low eff region	$\lvert x\rvert > 4$ cm or $\lvert y\rvert > 4$ cm

The fourth and fifth cuts in Table 5.5 require explanation. The RN cut restricts the run numbers to a period with relatively stable trigger and apparatus. In particular, this data was taken with a functioning and fully installed target microstrip system and before any crystal radiator running was undertaken[7]. The **Wide** cut was a

[7]The crystal radiator was an attempt to stiffen (increase the average photon energy) by taking

cut requiring that no D^+ secondary track intersected the P2 wire chamber within a central 8 cm by 8 cm box. We learned that there was a considerable run-dependent and intensity-dependent efficiency loss in the central area of the PWC system. We went through considerable efforts to model this efficiency loss, presumably due to accumulated polymerization damage by forward Bethe Heitler pairs, and photons interacting in the chamber rather than our experimental target. We include the **Wide** cut, with an 8 cm by 8 cm box which is actually considerably larger than the central chamber damage area, as a systematic check on our efficiency model.

The average of all 64 relative branching ratio estimates is 0.401. A good representative measurement is the CRD version of the third cut set (with IS2 and EL20). This measurement with its statistical error is:

$$\mathrm{BR} = \frac{\Gamma(D^+ \to \bar{K}^* \mu^+ \nu \to (K^- \pi^+) \mu^+ \nu)}{\Gamma(D^+ \to K^- \pi^+ \pi^+)} = 0.406 \pm 0.006$$

We have selected this value as representative since it (1) produces a number close to the average of all 64 variants and (2) is a defendable choice from the standpoint of cleanliness, high statisics, and having a "better" HC model. We believe that the CRD simulation was indeed the better of the two models since the value for cut set 3 using the ETD method was 0.406 ± 0.006 and much closer to the unbinned method CRD MC value of 0.407 ± 0.006 than it was to the unbinned method SHW MC value of 0.509 ± 0.006 shown in Table 5.4.

5.6 Systematics

Calculating systematics errors is more of an art than a science. We have considered several separate sources of systematic error summarized in Table 5.6. Our systematic error analysis concludes that uncertainties due to the K^* line shape, uncertainties as to the ETD averaging range, and variation in the results obtained under different

advantage of coherent bremmstrahlung. It was performed at the end of run primarily as a way of assesssing its advantages for a possible future run of the FOCUS experiment.

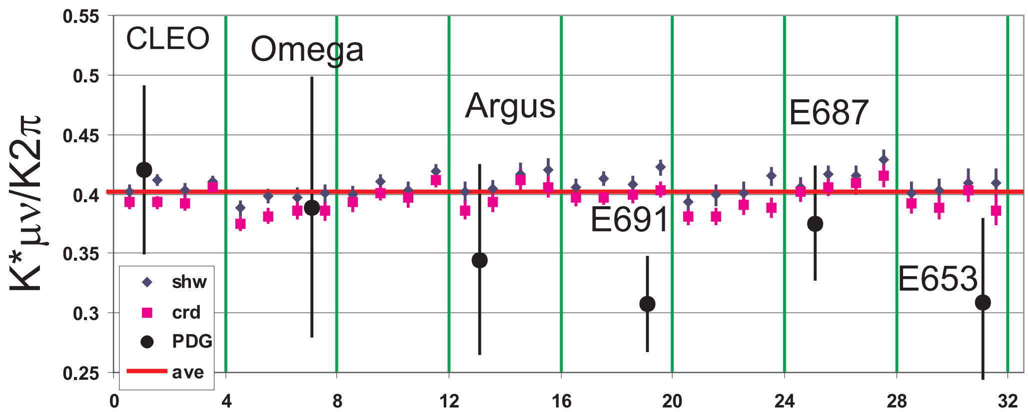

Figure 5.8: Measurements of $\Gamma(D^+ \to \bar{K}^*\mu^+\nu)/\Gamma(D^+ \to K^-\pi^+\pi^+)$ for 32 cut sets and two HC simulations along with previously published data from the 6 indicated experiments. The cut set number is binary coded as $I = \mathrm{IS2} + 2 * \mathrm{EL20} + 4 * \mathrm{OoM} + 8 * \mathrm{RN} + 16 * \mathrm{Wide}$ where (for example) the variable IS2 equals one when the isolation cut is applied and is zero otherwise. The cut acronyms are described in Table 5.1

cuts sets are the dominant sources of systematics. Each of these dominant systematic sources contributed roughly 2.3% errors on our BR of 0.406 for a quadrature summed total error that was about 3 times the statistical error. In this section we will summarize how each systematic estimate was made. We hope to convince the reader that we used conservative estimates of the true systematic error.

Table 5.6: Error sources

Description	Value
Statistical Error	± 0.006
Cut variants	± 0.010
Line shape	± 0.009
ETD averaging range	± 0.009
Form factor	± 0.002
Total sys error	± 0.016

We evaluated the largest systematic error, the cut set variant, by computing the

Figure 5.9: Our estimate of the (fractional) systematics on $\Gamma(D^+ \to \bar{K}^*\mu^+\nu)/\Gamma(D^+ \to K^-\pi^+\pi^+)$ as the cuts are varied. We plot the accumulated *sample variance* of the branching ratio starting from the first two cuts. A roughly constant relative systematic error of about 2.5% is reached after the first four cut variants are considered. Based on this plot, we attribute a 2.5% relative systematic error due to cut variation.

accumulated sample variance over the branching ratio obtained over the 64 cut sets[8]. We use the sample variance under the assumption that any of the three estimates is equally probable. This tends to give a pessimistic estimate of the error since it ignores any statistical fluctuation between the cut variants. By ignoring these fluctuations we are effectively assuming that each of the 64 estimates are fully correlated statistically, while in reality only a small fraction of the sample used to get the 64th measurement are in the first (base line) measurement. The accumulated sample variances are shown in Figure 5.9. After the first eight cuts are included, the sample variance appears to stablize to a value of about 2.5% of our representative relative branching fraction of 0.406.

The line shape systematic of about 2.3% was estimated by comparing the subtracted K^* yield difference fixing the width dependent power to $N = 0$ to that

[8]The 64 cut sets include the 32 cut sets plotted in Figure 5.8 and an additional cut requiring an HC energy deposition greater than 60 GeV. This additional cut was used to further check of the validity of the HC model at the high end of the HC spectra.

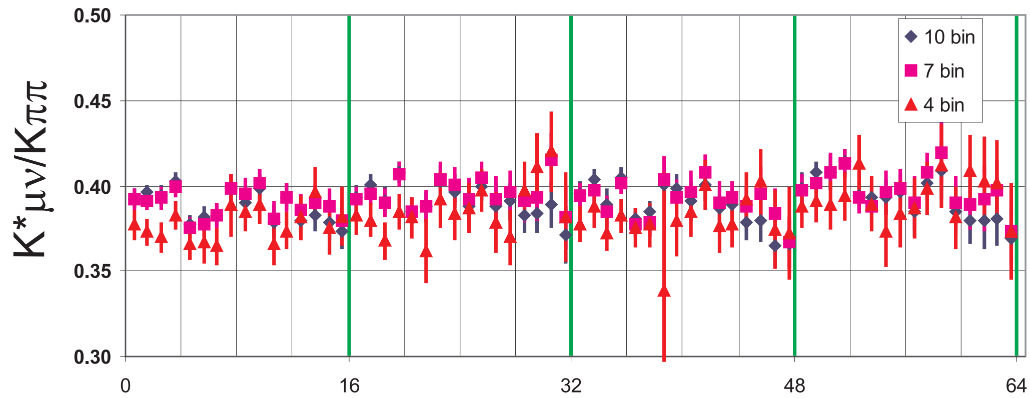

Figure 5.10: We plot $\Gamma(D^+ \to \bar{K}^*\mu^+\nu)/\Gamma(D^+ \to K^-\pi^+\pi^+)$ as the cuts are varied for three different ETD averaging schemes. Based on the sample variance of the 3 averaging schemes averaged over the 64 cut sets considered, we attribute a 2.2% relative systematic error due to averaging variation.

obtained fixing the power to $N = 1$ in Equation 5.2 for a variety of cut sets. Again we believe this somewhat conservative estimate since our fits typically gave width powers in the range from $N = -0.2 \to +0.5$.

The ETD averaging systematic of about 2.2% was estimated by computing the average of the 64 cut variant branching ratios using the CRD HC simulation and three ETD averaging schemes: (a) all ten 10 GeV wide ETD bins (2) the seven highest ETD bins and (3) the four highest ETD bins. The branching ratio obtained using these three ETD average schemes are shown in Figure 5.10.

Our final systematic was obtained by reweighting our Monte Carlo events by intensity-based form factor ratios that were $\pm 3\sigma$ from their world avarge values. Eight such MC reweighting estimates were obtained corresponding to $\pm 3\sigma$ shifts of r_v, $r_2$and r_3. The major effect (that was very small compared to other systematic effects) was obtained by comparing the relative BR obtained using a MC generated with $r_2 = 0.99$ that is 3 σ shifted from its world average value of $artoo = 0.78$. Although a great deal of effort went into this determination of the form factor systematic error, the

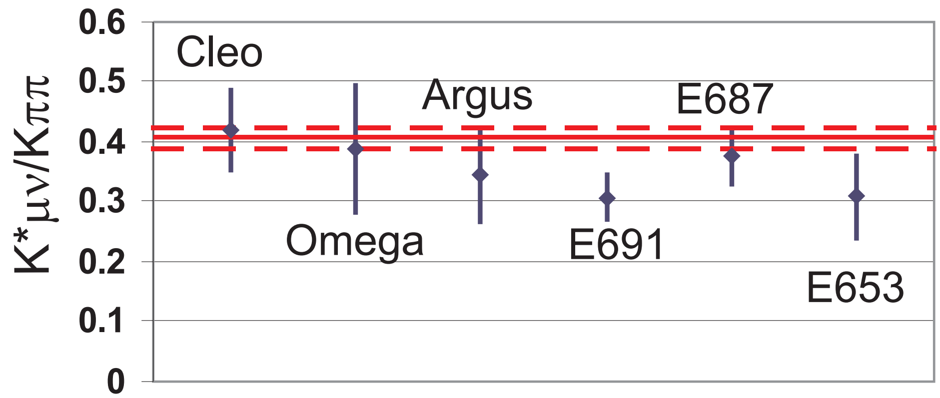

Figure 5.11: Our measurements of $\Gamma(D^+ \to \bar{K}^*\mu^+\nu)/\Gamma(D^+ \to K^-\pi^+\pi^+)$ compared to previously published data from the 6 indicated experiments. Our value including our systematic uncertaintly is given by the solid and dashed horizontal lines. Our (fractional) uncertainty including systematic errors is about 3 times smaller than that obtained in previous experiments.

systematic was negligible even under the conservative assumption of a 3σ shift.

5.7 Conclusions

In this chapter, we have produced a new measurement of

$$\mathrm{BR} = \frac{\Gamma(D^+ \to \bar{K}^*\mu^+\nu \to (K^-\pi^+)\mu^+\nu)}{\Gamma(D^+ \to K^-\pi^+\pi^+)} = 0.406 \pm 0.006\ (\mathrm{stat}) \pm\ 0.016\ (\mathrm{sys})$$

We compare this result to previously published results in Figure 5.11. Our value represents a considerable improvement in the precision of this branching fraction over previous experiments in spite of our rather large systematic error.

Besides obtaining a new measurment of this branching fraction, we believe that this analysis has established several facts germane to the rest of this thesis.

1. The stability of the relative branching ratio as a function of 1st eight cut sets shown in Figure 5.8 underscores the validity of the RS$-$WS background sub-

traction technique. As shown in Figure 5.1 and 5.2, the OoM and EL20 and IS2 cuts dramatically change the ratio of RS to WS events, and yet the resultant branching ratios are very consistent from the very dirty baseline sample 5.1(a) to the very clean sample shown in 5.1(b). We will use the same subtraction technique in the form factor analysis to follow.

2. The stability of the relative branching ratio for the 32 cut sets shown in Figure 5.8 further suggests that the Monte Carlo does an excellent job at computing the relative efficiencies of these various cut sets. This is a worthwhile accomplishment given the factor of 6 difference between efficiency for the loosest and tightest cut sets. In particular, it was re-assuring that the OoM cut results were consistant with the other results, given that this cut was one of the two pseudo-detachment cuts employed by the SLEPNRM skim for the $K^-\pi^+\pi^+$ reference state.

3. We found that the $K^* \to K\pi$ line shape was quite consistant with nearly a perfect s-wave Breit-Wigner. This line shape was that assumed in our MC model.

4. Although we initially found a considerable difference in branching ratios between the CRD and SHW models in our unbinned analysis for the HC simulation, we obtained results that were very close to the (unbinned) CRD simulation after applying the ETD method. Since the ETD bin is designed to be insensitive to the HC model, we believe that the CRD simulation provides a good HC simulation model for subsequent analyses.

Chapter 6

Fitting the Form Factor Ratios

Here we address the problem of fitting the semileptonic decay intensity to measure the form factor ratios.

We choose to use a binned likelihood fit because it is easy to understand and readily provides means for assessing goodness of fit. Our goal is to make a measurement with the least systematic errors possible even if it is at the expense of some statistical accuracy.

6.1 Kinematic Binning

A binning scheme is employed which segregates each event into a *kinematic bin* based on a set of kinematic invariants. The events (decay candidates) in each bin are counted and the distribution of signal events in all bins are thus measured. The principle of this fitting technique is to find what form factor ratios produce the best match to the observed bin populations.

This thesis is concerned with four-body semileptonic decays, and the four-body decay phase space is five-dimensional. We choose one of the dimensions to be the $K\pi$ invariant mass. This is used for our indicator plots — to count the amount of signal present in each kinematic bin. Most of the $K\pi$ mass distribution is the Breit-Wigner, which can be factored out of the decay rate (Equation 2.5). In other words, to a large extent the form factors do not influence the shape of the $K\pi$ mass spectrum. To measure the form factor ratios, we choose four more kinematic invariants that span the remaining four dimensions of phase space. These become the binning variables.

The most obvious kinematic invariants to choose are the ones used in describing the differential decay rate: Q^2, $\cos\theta_V$, $\cos\theta_\ell$, and χ. We will begin by binning in these four variables, though other choices are possible. For convenience, we choose bins of uniform size. $Q^2/Q^2_{\max}$ is used instead of Q^2 because it has a nice upper bound at 1. Since the decay rate only contains even functions of χ, we will bin in $|\chi|$ instead of χ, folding the full range into $0 \leq |\chi| \leq \pi$.

The technique of measuring the signal events in each bin will be discussed more in chapters 5 and 7. But it will either involve performing fits to the vector meson ($K\pi$) mass spectrum or event counting within a certain signal region. For $D^+ \to \overline{K}^{*0}\ell^+\nu_\ell$, a wrong-sign sample can be used for background subtraction.

6.2 Las Vegas Reweighting Technique

The idea of this binned fit is to maximize the agreement between the set of observed bin populations $\{n_i\}$ and the set of expected populations $\{\mu_i\}$ for a given set of form factor parameters. This is accomplished by having the fit vary the form factor ratios to minimize the χ^2 given by Equation 6.1.

$$\chi^2 = \sum_i \frac{(n_i - \mu_i)^2}{\sigma^2_{n_i} + \sigma^2_{\mu_i}} \tag{6.1}$$

The anticipated bin population uncertainties, σ_{μ_i}, are estimated either as the uncertainty in the fitted yield, or by Poisson statistics if we use sideband or wrong-sign subtraction to estimate the yield.

The complication in the procedure is that the expected bin populations for a given set of form factors, $\{\mu_i\}$, must reflect the considerable diffusion of events into and out of a given reconstructed bin due to the sizeable errors in reconstructing the missing neutrino and computing Q^2 and the decay angles. Additionally, we must account for variations from bin to bin in our event acceptance efficiency. In order to properly handle the possible biases and error inflation, we devised a variant of the fitting technique developed by the E691[22] collaboration, which we call the *Las Vegas reweighting method*.

We would like to use the FOCUS simulation, MCFOCUS, to account for the efficiency variation and resolution. The most simple method would be to pick trial values for the form factors, simulate a large data set with these form factors, apply the analysis cuts, and duplicate the data analysis. After normalizing the total signal yield to that of the data, the simulated bin populations become the predicted bin populations. This ideal method is impractical since it is too time consuming. We would like the simulated data set to be much larger (say, 20 times larger) than the data so that counting-statistics in the Monte Carlo do not contribute significantly to the measurement errors. This data set would take at least a day to accumulate with our best computing clusters. Since there are typically about a hundred fit iterations necessary before convergence on the best-fit form factors, this method would take several months to perform one fit. This is why the Las Vegas reweighting method is essential, because it allows all fit iterations to be performed with one simulated data set.

This method exploits the fact that form factors only influence the relative intensity for an event with a given set of kinematic variables and not its acceptance or resolution. Hence as the fit changes its trial form factors, we can compute the new bin populations by *reweighting* the previously generated events rather than regenerating a fresh sample.

We simulate a large data set with no matrix element simulation. For a given trial value of the form factors, each simulated event is given a weight. This weight is the full semileptonic differential decay rate with the given form factors divided by the simulated decay rate. In other words, our decay rate is the product of the phase-space distribution times the matrix element, $|\mathcal{M}|^2$. The matrix element is our weight. The effect of this reweighting is to produce the equivalent to a Monte Carlo sample generated with the given trial form factors.

6.3 Phase Space Monte Carlo

The details of the Monte Carlo generation become important at this stage. Our weighting must be proportional to the semileptonic decay rate divided by the generated decay rate.

We generate the semileptonic decays such as $D^+ \to \overline{K}^{*0}\ell^+\nu_\ell$ in three steps. First, the mass of the resonance (such as K^{*0}) is drawn according to a Breit-Wigner distribution, with the threshold conditions dictated by the masses of the D and final state particles. Next, the three-body decay ($D^+ \to \overline{K}^{*0}\ell^+\nu_\ell$) is drawn uniform in phase space. Finally the two-body decay ($\overline{K}^{*0} \to K^-\pi^+$) is simulated isotropically. This produces the simulated decay rate in Equation 6.2.

$$\frac{d^4\Gamma}{dM^2_{K\pi}\ dt\ d\cos\theta_{\rm V}\ d\cos\theta_\ell\ d\chi} \propto \frac{M_{K^*}}{M_D^2 M_{K\pi}} \frac{M_{K^*}\Gamma}{(M^2_{K\pi} - M^2_{K^*})^2 + M^2_{K^*}\Gamma^2} K \left(1 - \frac{m_\ell^2}{t}\right) \tag{6.2}$$

The weights for each event are given by Equation 2.5, with factors appearing in Equation 6.2 removed along with any constant prefactors.

6.4 Assessing the Goodness of Fits

With this binned technique, there are several ways we can assess the goodness of fit. This is essentially a means of determining if the semileptonic decay model is consistent with the data. We can apply a χ^2 test, examine various conditional projections, and check the consistency of particle and anti-particle samples.

We expect the best-fit χ^2 to follow a chi-square distribution with a number of degrees of freedom equal to $N_{\rm KB} - N_{\rm FFR} - 1$, where $N_{\rm KB}$ is the number of kinematic bins and $N_{\rm FFR}$ is the number of free fit parameters, the form factor ratios. The additional loss of a degree of freedom comes from the normalization condition. This allows us to form a confidence level by comparing the fit's minimum χ^2 with the integrated chi-square distribution of the appropriate number of degrees of freedom. This ought to be a valid test, provided that the statistical fluctuations in the predicted and observed bin populations are Gaussian. As long as the observed number of events in each bin are large, this condition will be nearly true. If the bin populations became small, we could try a minimum likelihood fit instead. But in that case we would have to deal with the statistics of the subtracted wrong-sign distribution as well.

The other consistency checks are simple. Projections of different binning variables can be obtained by summing the bins in the other dimensions, and this will prove to be a valuable tool. Often conditional projections, or the projections of a distribution subjected to a cut on another variable, will be helpful. Since the semileptonic decay rate for particles versus anti-particles has been a tricky subject historically (fraught with mistakes), it will be important to check the projections with these split samples in particular.

6.5 Fitting Simulated Data

We can test our fitting technique for biases by fitting a simulated data set with known values of r_v and r_2. On average, we expect the form factor fit to recover these input values. We use our acceptance Monte Carlo for $D^+ \to \overline{K}^{*0} \mu^+ \nu_\mu$ for both purposes, dividing the data into twenty portions, each roughly the size of one FOCUS data set. One subset is treated like data. Analysis cuts are applied, and events are counted when they fall within the signal region of $0.8 < M(K\pi) < 1.0$ GeV/c^2. Wrong sign events are subtracted. We also give each event a weight based on the matrix element computed with the input values of r_v and r_2. The remaining sets are used as the acceptance Monte Carlo for the fit. This process can be repeated twenty times with each of the subsets treated as data.

Figure 6.1 shows how well the fit results match the input values. There is no significant bias in the fitting method.

The distribution of confidence levels for these fits appears in Figure 6.2. We do not have enough fits to judge the uniformity of the distribution, but there do not appear to be any serious problems.

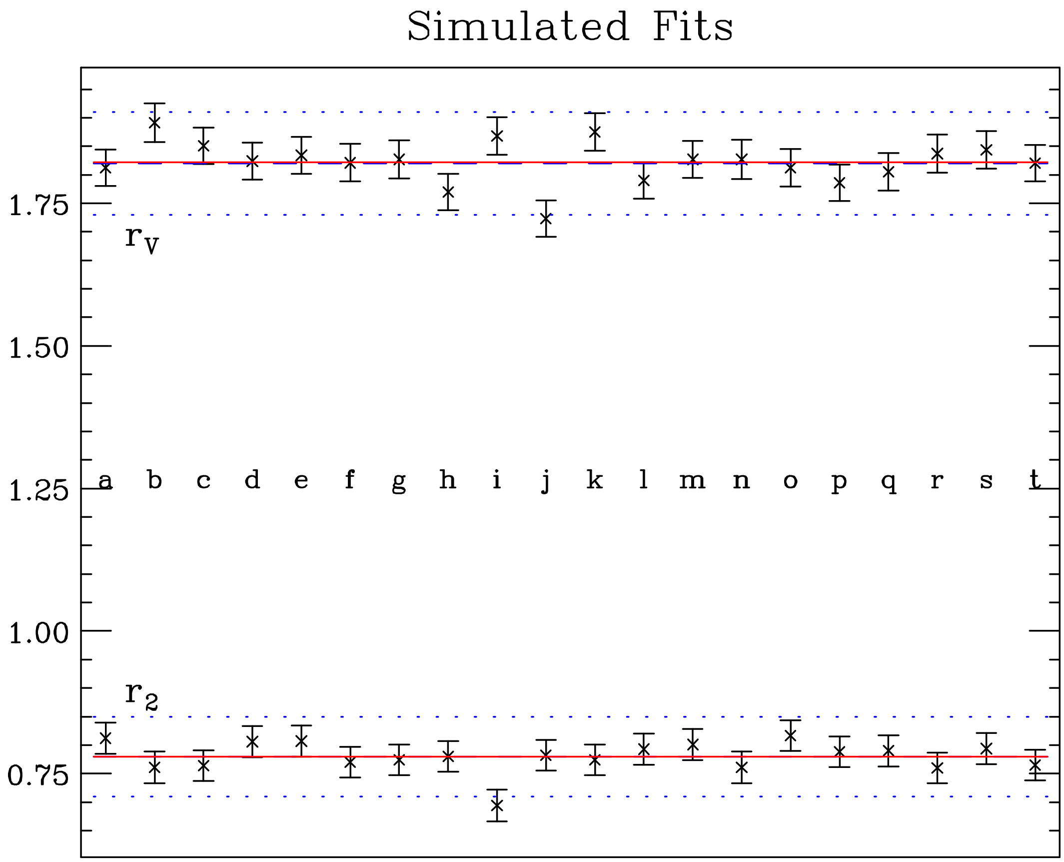

Figure 6.1: Fits of simulated data. Black points are the individual fit results with their errors. The red lines are the averages of these fits, and dashed blue lines are the input values. These are hidden underneath the red line but are centered between the dotted blue lines.

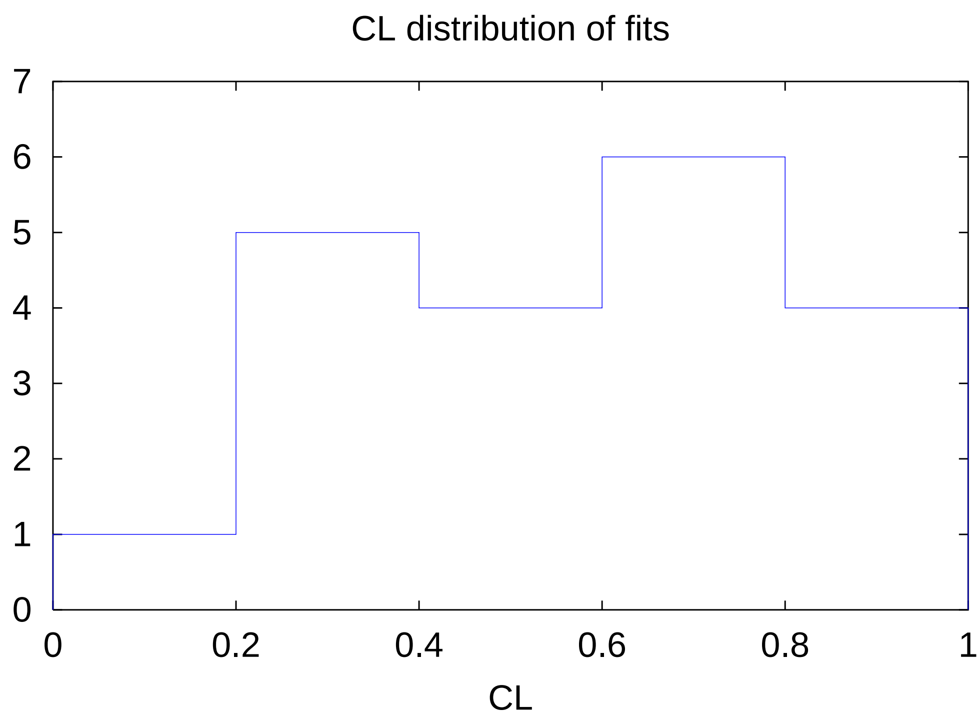

Figure 6.2: Confidence level distribution of fits.

Chapter 7

Analysis of the $D^+ \to \overline{K}^{*0} \ell^+ \nu_\ell$

The decay $D^+ \to \overline{K}^{*0} \ell^+ \nu_\ell \to (K^-\pi^+)\ell^+\nu_\ell$ is the easiest semileptonic decay to observe in the FOCUS data. It makes a natural choice for developing the form factor ratio measurement technique. This decay is relatively abundant as it requires a D^+ (one of the most common charm hadrons) and has a branching ratio of about 3%.

Previous measurements of the form factor ratios for this decay mode are consistent with each other and agree with recent theoretical predictions. The world average values for r_v and r_2 have comparable errors to recent LGT predictions, and have much smaller errors than the form factor ratios for $D_s^+ \to \phi\ \ell^+\nu_\ell$. However, it is still important to perform this measurement with the FOCUS data. The most recent LGT predictions are a few years old now, and can be significantly improved upon at this time. But more importantly, the high level of statistics available for this decay mode allow for unprecedented consistency checks of the differential decay rate model and the measurement technique. The emphasis in this case will be on making a measurement with many consistency checks and hopefully low systematic errors, rather than making the most accurate measurement possible. This emphasis will influence the method of the fit as well as the cuts that are chosen.

In this chapter we focus exclusively on the muon decay mode. The electron sample is likely to be more difficult to understand, and at present the electron particle identification algorithms and calorimeter simulations have not been as exhaustively studied as the muon systems.

We first present a set of cuts used in this analysis, establish a somewhat arbitrary baseline set of cuts, and present the results of a fit with these cuts. This fit shows many problems, the most notable being an asymmetry in $\cos\theta_{\mathrm{V}}$ that is not predicted by the decay rate model or the FOCUS Monte Carlo. We then examine some noteworthy backgrounds, show some fit variants and split samples, and look in more detail at cut responses and other event properties. Throughout all this, our primary objective is to understand the origin of this discrepancy in the fit. The key to unlocking the puzzle turns out to be the way in which the $\cos\theta_{\mathrm{V}}$ asymmetry varies with the $K\pi$ invariant mass. Next we investigate some other possible backgrounds which turn out not to contribute significantly. At last we come up with a toy model that represents a broad scalar resonance or non-resonant contribution to the $K\pi\mu\nu$ final state. The interference between the scalar contribution and the vector resonance can produce exactly the kind of asymmetry that we observe in $\cos\theta_{\mathrm{V}}$.

7.1 $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ Signal

For measuring the $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ form factors, the "indicator" plot we use is the invariant $K\pi$ mass, in which the K^{*o} appears as a Breit-Wigner over a slowly-varying background. We call this an indicator plot because the distribution of the signal events is easily distinguishable from the background distributions, and thus the signal can be counted. The basic technique for measuring the signal is to ensure clean particle identification and vertex quality by a set of analysis cuts, and then count the number of K^{*0}s seen in the $K\pi$ mass spectrum.

7.1.1 Wrong Sign

The strangeness of the K^{*0} is evidenced in the charge of its decay products, and this can be compared with the sign of the muon. In this semileptonic decay, the c quark in the D meson changes into an s quark, emitting a W^+. This decay relates the charge of the muon to the strangeness of the K^{*0}. If the charge of the muon is opposite that of the charged kaon, is is a *right-sign* candidate, otherwise it is a *wrong-sign* candidate. The wrong-sign population is representative of most of the backgrounds present in the right sign sample, in particular those from events where the muon and K^{*0} did not originate from the same decay. This class of backgrounds includes lighter (non-charm) hadronic events and cases where a muon from the recoil charm decay was confused with the K^{*0} vertex. By subtracting the wrong-sign sample from the right-sign, we obtain a rather clean signal. The extent to which various backgrounds are subtracted out with this method will be explored later.

7.1.2 Skim

We use the semi-muonic subskim to search for this signal (see section 4.4.3). This is efficient since any reasonable analysis cuts will be tighter than the cuts used in this skim. The rare exceptions are cases when the 3-track $K\pi\mu$ vertex has a much better confidence level than the $K\mu$ vertex. For the baseline cuts used in this thesis, the skim cuts are about 99.1% efficient.

7.1.3 Baseline Cuts

Various types of vertexing cuts are important. The cut on the $K\pi\mu$ vertex confidence level (DCL) is of primary importance, and is effective even at 1%. Raising this cut to 5% further reduces backgrounds from various processes including those where the pion was really from a D^* decay. The fidelity of the simulation also seems to improve somewhat at larger vertex confidence levels. Cuts demanding isolation of the secondary vertex (ISO2 and ISO2ex) are also important to eliminate contamination from decay modes with higher multiplicity, particularly in this analysis where we do not fully reconstruct the D. Although there is no evidence of higher multiplicity semileptonic decays, there are many 4+-body fully hadronic charm decays, and in these are many possibilities for kaons and pions to be mistaken for a muon. Enforcing primary vertex isolation (ISOP) helps to reduce contamination from decays such as $D^{*+} \to D^0\pi^+ \to (K^-\mu^+\nu_\mu)\,\pi^+$. The D^* decays quickly, so the pion from the D^* should be consistent with coming from the primary vertex. Explicit cuts on the D^*-D mass difference are more efficient and effective at excluding these backgrounds, so primary vertex isolation cuts are largely redundant.

Čerenkov cuts are used on the kaon and pion tracks. Because kaons are more rare in FOCUS data, a cut on the kaonicity of the kaon is extremely effective, and used as a baseline cut. A loose pionicity cut can also be helpful, but does not seem absolutely necessary. Without any pionicity cut, one can observe a small distortion to the $K\pi$ mass spectrum from $D_s^+ \to \phi\ \ell^+\nu_\ell$. This appears equally in the wrong-sign sample, so it is easily subtracted away, but it may be a problem when fitting the individual right-sign and wrong-sign samples.

The main difficulties in this analysis are that clean muon identification is essential, and that the K^{*0} is a fairly broad resonance. The most obvious and threatening backgrounds are from fully hadronic charm decays where a pion is misidentified as a muon. Cabbibo-favored decays, such as $D^+ \to K^-\pi^+\pi^+$ or $K^-\pi^+\pi^+\pi^0$ are always right-sign events when a pion is misidentified. Only Cabbibo-suppressed decays would contribute to the wrong-sign. To make matters worse, typical charm decays include a K^{*0} resonant component which will be indistinguishable from the semileptonic signal. The fact that pions naturally decay into muons means that in principle we cannot completely eliminate muon misidentification backgrounds. In practice, we can reduce these backgrounds to an extremely low level and show that they do not have much effect on the analysis. Cuts demanding a minimum muon momentum, good muon ID confidence level, minimum missing planes, and good track fit confidence level are all helpful and discussed earlier in Chapter 4.

A variety of cuts based on kinematics are helpful. The minimum D mass (see section 4.6) is somewhat useful for reducing a wide variety of backgrounds. Cuts on the minimum visible ($K\pi\mu$) momentum helps the signal-to-noise a great deal, perhaps to reduce backgrounds from light hadron events. An important cut used in this analysis is based on the $M(D^{*0}) - M(D^+)$ mass difference. To compute this difference, we reconstruct the neutrino and suppose that the $K\pi\mu\nu$ combination is really from the decay chain $D^{*+} \to D^0\pi^+ \to (K^-\mu^+\nu_\mu)\,\pi^+$. Demanding that this mass difference is greater than 180 MeV/c^2 almost completely eliminates this class of events, with very little loss of signal. It is also possible to explicitly reject backgrounds from $D^+ \to K^-\pi^+\pi^+$ by assuming the muon track has the mass of a pion, and rejecting candidates that reconstruct to within 2 or 3 σ of the D^+ mass. It is surprising how little of this background makes it through the other cuts, but since this particular cut is so efficient, we use it mostly on principle. Finally, a cut requiring that the invariant mass of the visible daughters be less than the D^+ mass is necessary to close the neutrino kinematics properly. This turns out to be a very permissive cut, and can practically be ignored.

The *baseline* cuts used for this analysis are summarized in Table 7.1. Signal plots with these cuts applied are shown in Figure 7.1

DVERT CL	$> .05$
ℓ/σ	> 8
ISO2ex	$< .001$
Kaonicity	> 2
Pionicity	$>$ -2
CL$\mu_{\text{(new)}}$	$> .05$
missing muon planes	≤ 1
$\lvert P_\mu \rvert$	> 8 GeV/c
$\lvert M(K\pi\pi) - M(D^+) \rvert$	$> 3\sigma$
$M(D^{*+} - D^0)$	> 180 GeV/c^2
$\lvert P_{K\pi\mu} \rvert$	> 30 GeV/c
M(charged daughters)	< 1.869 GeV/c^2

Table 7.1: Baseline cuts for $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$

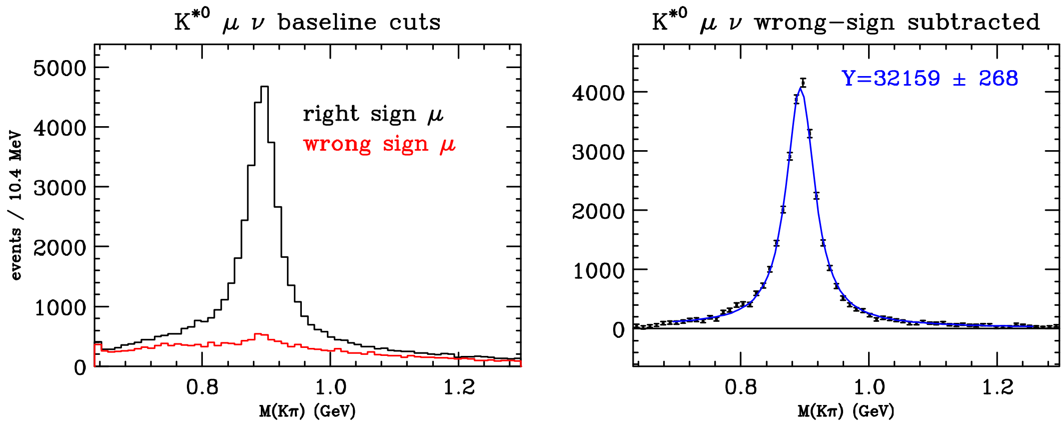

Figure 7.1: $D^+ \to \overline{K}^{*0}\ell^+\nu_\ell$ signal with baseline cuts. The plot on the right is wrong-sign subtracted, and fitted with a Breit-Wigner (non-relativistic spin-less) convoluted with a 5 MeV-wide Gaussian plus a linear background. The fitted signal yield is 32159 ± 268 events.

7.2 The Form Factor Fit

Before discussing further details of the $D^+ \to \overline{K}^{*0}\ell^+\nu_\ell$ signal, it is helpful to show the results of the form factor ratio measurements. We will observe some severe problems with the fit. This provides a context in which to explore possible sources of difficulty such as the fidelity of the Monte Carlo, the semileptonic decay model, and background contamination.

7.2.1 Baseline

For the baseline fit, we use three bins in $\cos\theta_{\rm V}$, $\cos\theta_\ell$, $|\chi|$, and $Q^2/Q^2_{\rm max}$. This makes a total of 81 bins. Three bins in each of these distributions is just enough to capture slope and curvature, while keeping the event count in each bin large enough for the χ^2 fit technique. We count signal by subtracting wrong-sign events from right sign in a K^{*0} mass window from 0.8 – 1.0 GeV/c^2. The results of the fit are summarized in Table 7.2.

These results are difficult to believe, because with a fit confidence level of 10^{-26}, clearly the model is not fitting the data. The most obvious test is whether or not a

r_2	$=$	0.754 ± 0.034
r_v	$=$	1.694 ± 0.051
χ^2/DOF	$=$	3.727
confidence level	$=$	2.9×10^{-26}

Table 7.2: Form factor measurement results for baseline fit

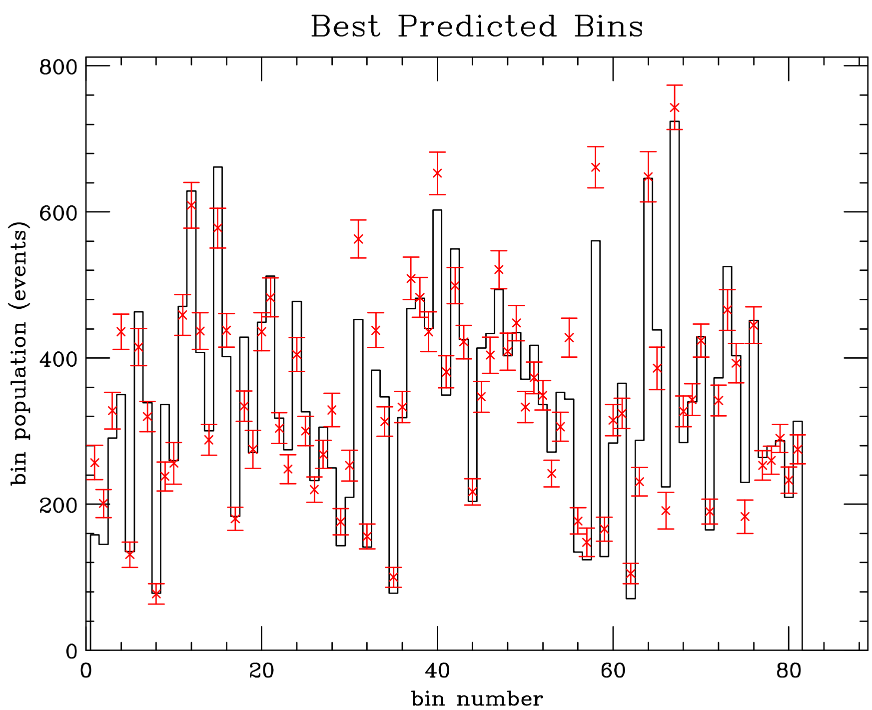

Figure 7.2: Bin populations and predictions for baseline fit. 3 bins in $\cos\theta_V$, $\cos\theta_\ell$, $Q^2/Q^2_{\max}$, and $|\chi|$ for a total of 81 bins. The red points with error bars are the observed bin populations. The black lines are the best-fit bin predictions.

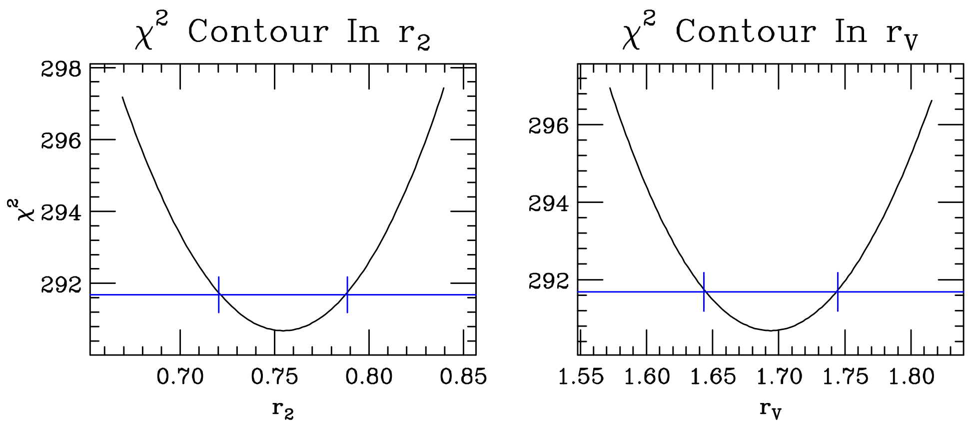

Figure 7.3: χ^2contours for the baseline fit. The blue cross-hairs are one unit above the fit minimum, and 1σ to the left and right. If the contour is truly parabolic, and the fit parameters are not strongly correlated, it will intersect at these points. It evidently does. According to the fitter, there is a 21% anti-correlation between r_v and r_2, which is not severe.

valid minimum was found in the r_v–r_2 plane. Figure 7.3 shows that the χ^2 fit function is well behaved (smooth and parabolic). This can be expected since the predicted bin populations are sums of weights, and the weights have at most quadratic dependencies on the form factors. We conclude that the fit is technically correct insofar as a unique and stable minimum was found, but the height of this minimum indicates that the model is unable to represent the data with any choice of r_v and r_2.

7.2.2 Projections

We now address the question "Where in phase space is the mismatch between the model and data the most severe?" Since the four-dimensional phase space is unwieldy, we must make projections. The most simple way to form these projections is to add up sets of bins used in the form factor fit. These are shown in figure 7.4. With only 3 bins in each dimension, this gives us poor resolution. We also miss a lot of the behavior of the decay intensity because these projections fail to show the correlations between the variables. Even so, one severe mismatch is evident here in $\cos\theta_{\mathrm{V}}$.

Figure 7.5 shows a more detailed set of conditional projections with the data over-

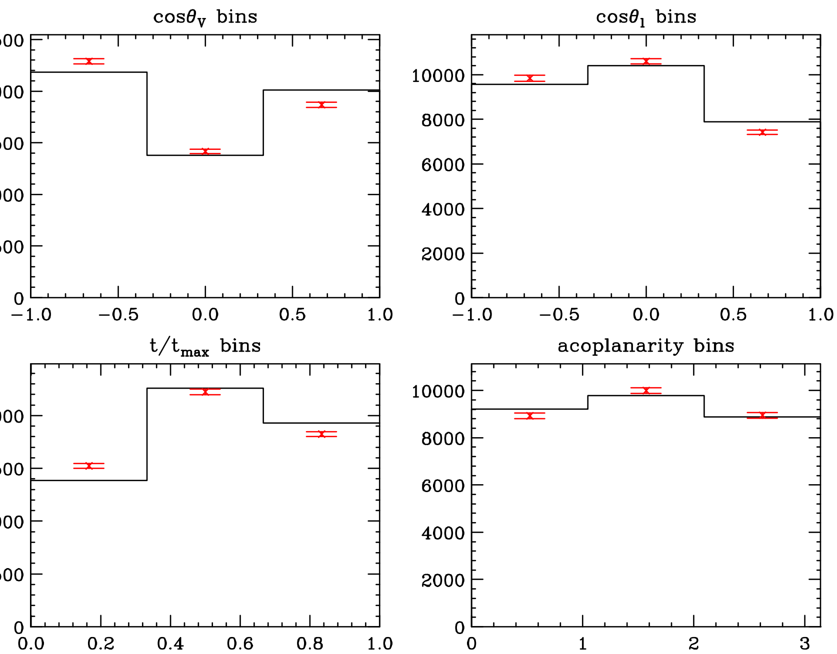

Figure 7.4: Projections formed by integrating over bins orthogonal to the variable of interest. Agreements between the best-fit prediction and the data in $\cos\theta_\ell$, $Q^2/Q^2_{\rm max}$, and χ seem almost reasonable. But there is an asymmetry in $\cos\theta_{\rm V}$ that is observed but not correctly predicted.

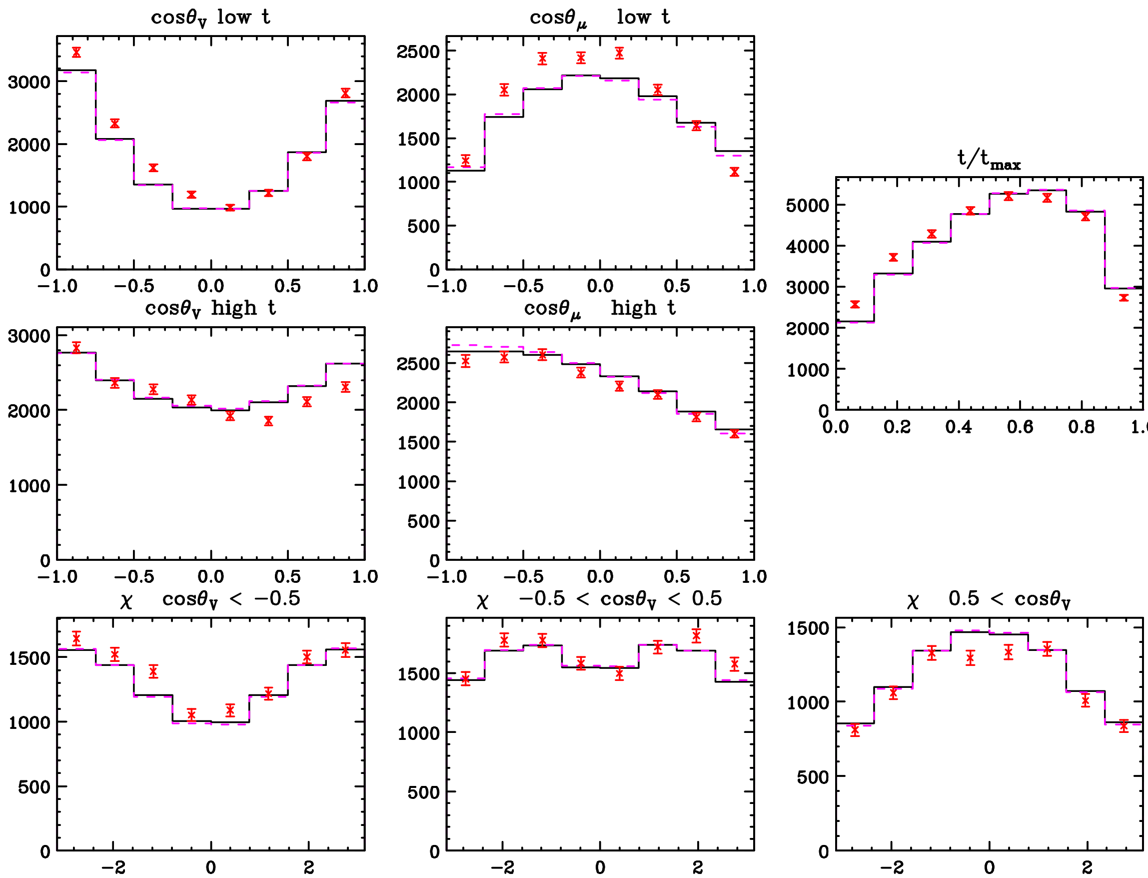

Figure 7.5: Projections with finer binning. The red points are the observations, and black lines are predictions with best-fit form factors. Dashed magenta lines indicate the prediction with the 2000 world average form factor ratios: $r_v = 1.82$, $r_2 = 0.78$, and $r_3 = 0$.[2] $\cos\theta_{\rm V}$ and $\cos\theta_\ell$ projections are broken into events with $Q^2/Q^2_{\rm max} < 1/2$ and $Q^2/Q^2_{\rm max} > 1/2$. The χ projection is divided into three regions of $\cos\theta_{\rm V}$. The predicted distributiosn are normalized to the data yield, and this normalization factor is preserved in the splits.

laid with two predictions — the fitted form factor ratios, and the 1998 world average values. Here we see a variety of interesting features. There is a lot of information in these projections.

All of the distributions exhibit behavior that is strikingly similar between the data and the prediction. This gives us a fair amount of confidence in the decay rate expression, and our interpretation of the internal angles. It is particularly reassuring to see the primary features in the acoplanarity distributions matched so well by our model. Note that although the *integrated* acoplanarity distribution in Figure 7.4 contained little variation, these *conditional* projections show how the decay rate truly is a strong function of acoplanarity. All the variations and correlations will be "felt" by the fitter using 81 bins.

By comparing the current world-average [2] predictions with the best-fit predictions, we get some crude sense of how the form factor ratios can influence these distributions. Our measured values are very close to the world average values, so the differences are subtle. The $\cos\theta_\ell$ distribution seems to have the most influence over the form factor measurement, and so the fitter has the most flexibility to match this distribution. The degree of concavity of $\cos\theta_{\mathrm{V}}$ can be affected by the form factor ratios. But any asymmetry in this distribution can only be accounted for by experimental acceptance effects. This leads to the conclusion that our problems in the fit are best assessed by looking at the $\cos\theta_{\mathrm{V}}$ distributions. It is possible that these same problems are affecting $\cos\theta_\ell$ in ways that are masked by the fit, but which bias the results. For this reason that we cannot trust the form factor ratios from this fit.

We will concentrate on the observed asymmetry in $\cos\theta_{\mathrm{V}}$ since it cannot be accounted for by a shift in the form factors. There are also imperfections in $\cos\theta_\ell$ at low Q^2, and the Q^2 distribution is also poorly modeled. All of these plots share common normalization factors, which explains why the χ plots do not appear to match. Since they are divided into regions of $\cos\theta_{\mathrm{V}}$, the asymmetric deviation in $\cos\theta_{\mathrm{V}}$ upsets the proper normalization within each χ projection.

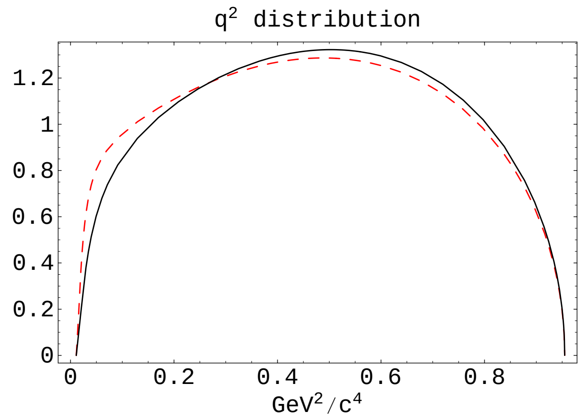

Figure 7.6: Integrated Q^2distribution with and without mass terms. The solid black curve represents this decay rate with no mass terms, and the dashed red curve includes the mass terms. Both are normalized to unit area. The decay rate is numerically integrated in χ, $\cos\theta_V$, and $\cos\theta_\ell$. The $K\pi$ mass is fixed at the K^{*0} pole.

7.3 Mass Terms, r_3, and Pole Masses

In the baseline fit, we included the mass terms although we assumed the A_3 form factor was zero. Is it possible that a particular value of r_3 could improve the goodness of fit? Here we address the issues of how the mass terms affect the decay rate, how sensitive we are to r_3, and whether or not we can measure the axial and vector pole masses assumed in $f(Q^2)$.

Figure 7.6 shows the decay rate (Equation 2.5) as a function of Q^2 ($t \equiv Q^2$), with and without the mass terms. These terms cause an increase at low Q^2, and are significant in this analysis. One might initially guess that the mass terms, proportional to $m_\ell^2/2t$, would constitute only 1% of the total decay rate since the muon mass is $\approx .1\mathrm{GeV}/c^2$ while $2t \approx 1(\mathrm{GeV}/c^2)^2$. Yet the mass terms are numerous, and have large multiplicative factors. They amount to 4.5% of the integrated decay rate.

In order to measure our sensitivity to Q^2, we repeat the studies performed in Section 6.5, allowing all three form factor ratios to be fit simultaneously. For additional fidelity, instead of weighting the input simulation data, we use rejection to sculpt the matrix element.

In these fits, the reported uncertainty in r_2 and r_3 are within 20% of the uncertainties reported in the baseline fit, so our Monte Carlo sample is roughly comperable to the FOCUS data. According to the simulated fits, we find that r_3 is measured without bias, but with an uncertainty of ± 3.1. This seems reasonable. A_3 multiplies roughly one third of H_t, which itself multiplies only about a third of the mass terms. Since the mass terms account for about 5% of the total decay rate, we expect that A_3 multiplies about 0.5% of the rate, whereas V and A_2 will multiply perhaps half of the rate. Therefore changes in r_3 should have only 1% of the effect that changes in r_v and r_2 would have. Our predicted error in r_3 is in fact about 100 times larger than our errors in r_v and r_2. It is difficult to understand how E791, with about 1/14 of the $\overline{K}^{*0}\mu^+\nu_\mu$ signal that FOCUS has, is reportedly able to measure r_3 to a statistical accuracy of ± 0.33 (only 3 times their statistical uncertainty in r_v)[10].

In addition to this, we predict that the correlation between r_3 and r_2 is large: 86% positive correlation. By including r_3 in the fit, we double the measurement error of r_2. For this reason, and because we are not getting a consistent fit, we do not attempt to measure r_3 in this analysis. Incidentally, when we do attempt this we obtain values of r_3 in the neighborhood of -20, and the goodness of fit is not improved. This is clearly unreasonable, and must be caused by the fitter struggling to compensate for a model that simply cannot match the observed data.

It would also be valuable if we could measure the vector and axial pole masses used in $f(Q^2)$. Our sensitivity to the pole masses was predicted *a-priori* to be around 0.16–0.38 GeV/c^2 in a 1995 conference [12]. Again using our Monte Carlo, we predict that we could measure the vector pole mass to $\pm.56$, and the axial pole mass to ± 0.27 GeV/c^2. With 81 bins, we are only splitting Q^2 into 3 regions, and we might be able to do a little better with finer binning. Even so, our sensitivity is not good enough to make this a compelling measurement, and there are severe correlations to the form factor ratios. r_v and M_V are 91% correlated, and r_2 and M_A are 85% correlated using this 81-bin fit. It is best to leave measurements of $f(Q^2)$ to the simpler pseudoscalar-ℓ-ν decays.

Label	Cut	Yield	RS/WS	χ^2/DOF
tb	baseline	27861	4.84	3.73
dcl1	DCL > .01	30036	4.35	3.29
dcl2	DCL > .25	22040	5.34	3.55
ip1	ISOP < .1, no $M(D^{*+} - D^0)$ cut	24795	5.26	4.15
ip2	ISOP < .1	24701	5.29	4.16
is1	ISO2ex < .01	28853	4.58	3.68
is2	ISO2inc < .0001	20587	7.84	4.21
ka1	Kaonicity > 6	20420	5.99	3.54
pi1	no Pionicity cut	28530	3.78	3.66
pi2	Pionicity > 2	20204	5.22	3.52
kpp1	no $M(K\pi\pi)$ cut	28060	4.83	3.69
mu1	μ TRKFIT CL > 1%	23606	7.48	3.41
mu2	mu1, CL$\mu_{\text{(new)}}$ > .15	20948	8.06	3.19
mu3	mu1, CL$\mu_{\text{(new)}}$ > .20, $\lvert P_\mu \rvert > 12$ GeV/c^2	17907	9.07	2.65
ls1	$L/\sigma > 12$	23851	5.97	4.05
ls2	$L/\sigma > 16$	20441	6.57	3.87
ls3	$L/\sigma > 20$	17570	6.87	3.82
mdm1	1.4 < min D mass < 2.2 GeV/c^2	23139	8.17	4.17
mdm2	1.6 < min D mass < 2.0 GeV/c^2	17612	11.17	3.59
mv1	no M(charged daughters) cut	27897	4.75	3.70
pv1	no $\lvert P_{K\pi\mu} \rvert$ cut	28173	4.48	3.72
pv2	$\lvert P_{K\pi\mu} \rvert > 50$ GeV/c	23444	6.34	3.67

Table 7.3: Fit variants for $D^+ \to \overline{K}^{*0} \mu^+ \nu_\mu$.

7.4 Fit Variants, Split Samples

Perhaps the best way to investigate the problems with this baseline fit is to explore different sets of cuts. We will vary certain cuts, and we will also make several split samples. If the source of the problem is a distinct background or a performance mismatch in the Monte Carlo, we ought to be able to identify it with these studies.

Table 7.3 introduces a series of cut variants. These variants loosen and/or tighten a variety of cuts with respect to the baseline. The minimum χ^2 for these fits does not improve significantly for any of these fit variants. The labels for the cut variants appear in Figures 7.7 and 7.8.

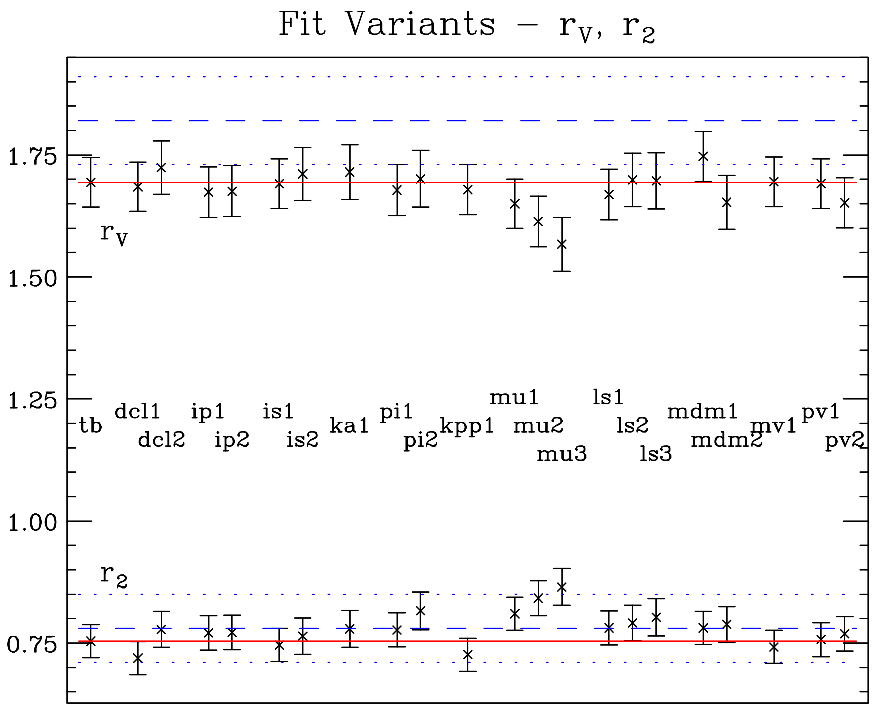

Figure 7.7: r_v and r_2 for fit variants. The blue dashed and dotted lines indicate the world average values.[2]

In Figure 7.7 we see the stability of the form factor ratios for these variants. These results are somewhat stable. The minimum D mass cuts and muon cleanup cuts appear to shift the results noticeably. Additional detachment (ℓ/σ) cuts also appear to affect r_2.

Figure 7.8 shows the difference in the average value of $\cos\theta_{\mathrm{V}}$ between the data and the model. This quantifies the asymmetry effect seen in Figure 7.5. The predicted asymmetry is entirely due to experimental acceptance. We now see how significant this difference is: 8σ with the baseline fit!

The most obvious feature here is the severe asymmetry introduced by a tighter pionicity cut. In contrast with the baseline cuts (which require a pionicity greater than -2), this requires positive pion identification in the Čerenkov system. Moving from the cut at -2 to +2, we cross the discontinuity at 0 and reject pions that are indistinguishable from kaons. If the pion momentum is high enough, both kaons and pions would be fast enough to emit Čerenkov light, so at high momentum, kaons and

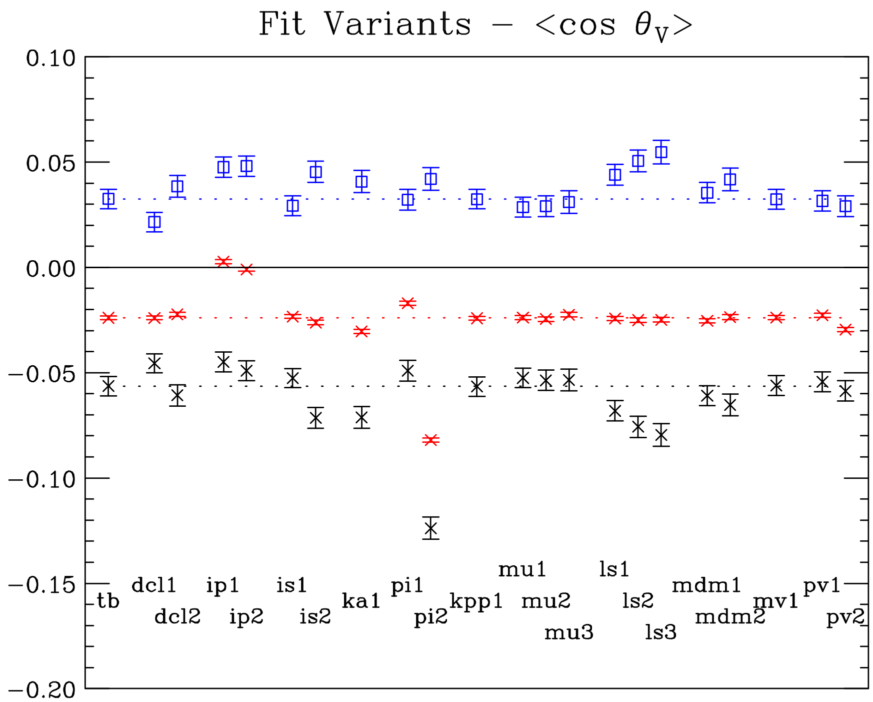

Figure 7.8: $\langle \cos\theta_{\mathrm{V}} \rangle$ for fit variants. Black points are the data, red points are the fit predictions, and the blue points are the differences between the two.

pions are indistinguishable in the Čerenkovdetectors. So this positive pionicity cut effectively sculpts the distribution of $\cos\theta_{\mathrm{V}}$. The Monte Carlo simulates this perfectly, and we see that the difference between data and the predicted asymmetry is stable under these cut variants. The primary isolation cuts also behave in a similar way. If the pion were nearly at rest in the D^+ rest frame, θ_{V}would be roughly π, and in the lab frame, the pion track would point back to the production vertex. This part of phase space would be excluded by the primary isolation cut.

It appears that the only variant which tends to alleviate the problem with the $\cos\theta_{\mathrm{V}}$ asymmetry is dcl1, which relaxes the decay vertex requirement from 5% to 1%. By loosening this cut, we may be allowing in more backgrounds such as $D^{*+} \to D^0\pi^+ \to (K^-\mu^+\nu_\mu)\,\pi^+$. In these events, the pion originates very near the production vertex. But since the D^{*+} decay is so near threshold, the pion will not have much energy with respect to the D^0, and so the pion track will come very close to the D^0 vertex. However, the pion track will not usually be consistent with the secondary

Label	Cut	Yield	RS/WS	χ^2/DOF
tb	baseline	27861	4.84	3.73
dc1	D^-	14953	5.43	3.06
dc2	D^+ only	12908	4.32	1.83
tm1	$Q^2/Q^2_{\rm max} < 0.5$	12926	4.72	3.54
tm2	$Q^2/Q^2_{\rm max} > 0.5$	14935	4.95	1.78
km1	$M(K\pi) < .9\,{\rm GeV}/c^2$	16558	5.17	4.62
km2	$M(K\pi) > .9\,{\rm GeV}/c^2$	11303	4.48	1.69
ph1	physical ν closure	11779	3.66	2.04
ph2	unphysical ν closure	16082	6.69	3.19
om1	decay vtx in material	5354	2.41	1.29
om2	decay vtx out of material	22507	7.53	3.82
pim1	primary vtx in material	23195	5.69	3.71
pim2	primary vtx out of material	4666	3.02	1.00
np1	≤ 4 tracks in primary	19271	4.68	2.81
np2	≥ 5 tracks in primary	8590	5.25	2.10
rn1	run number ≤ 9475	7461	4.73	1.51
rn2	run number ≥ 9476	20400	4.88	3.14
io2	outer muons	5069	3.20	1.28

Table 7.4: Split samples for $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$.

vertex since it does not truly originate there. These types of backgrounds have a highly positive asymmetry which tends to mask the real problem.

Next we investigate several split samples. These are described in Table 7.4, and the fit results are shown in Figures 7.9 and 7.10. The unphysical neutrino closure refers to events where the primary vertex was outside of the physically-allowed cone, and was adjusted to be on the cone for purposes of determining the neutrino's momentum (see Section 4.6). In general, the results are again quite stable but this time there are some exceptions.

r_v changes somewhat when we split on particle versus anti-particle, but this could be a fluke. What we are most concerned with is whether or not our conventions for defining the angles are correct, and whether or not there should be some sign-changes in the decay rate expression. Historically, this has been the subject of much confusion

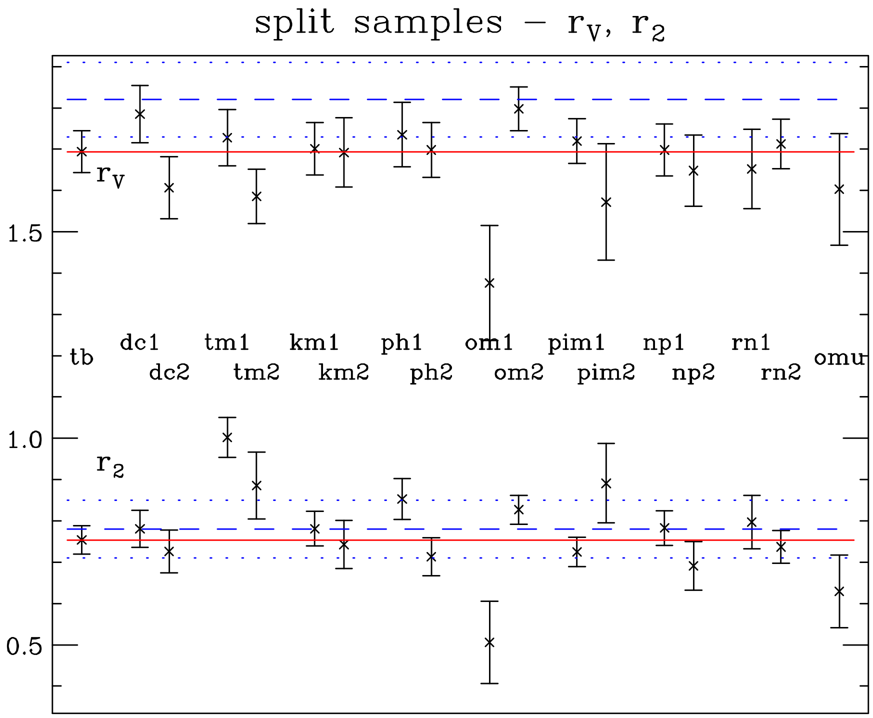

Figure 7.9: r_v and r_2 for split samples.

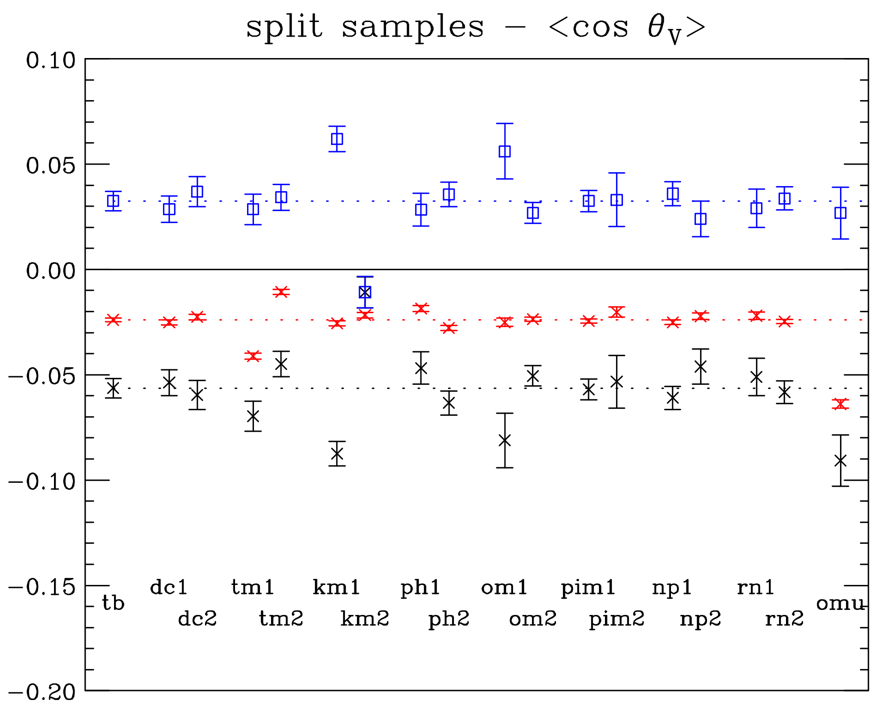

Figure 7.10: $\langle \cos\theta_V \rangle$ for split samples.

and errors, and we would see a terrible χ^2/DOF for one of the fits if we had interpreted any conventions incorrectly or made a sign error. We see no evidence of any severe problems like that, and the conditional projections look generally reasonable (not shown). Curiously, the D^+ sample provides a much better fit. The Q^2 distribution is a much better match for the D^+.

The high-Q^2 sample provides a better fit, for reasons we will understand later. Looking at the table, we see that it may be worthwhile demanding the primary vertex be within some material (target segments or silicon microstrips) and the secondary vertex out of the material. These could be useful clean-up cuts. We get quite inconsistent results for the events that decay inside material, so there may be some significant background contamination here still. Other split samples produce consistent results.

The most remarkable split is on the $K\pi$ mass, roughly split above and below the nominal K^{*0} mass. Although r_v and r_2 do not change significantly, the $\cos\theta_{\mathrm{V}}$ asymmetry completely disappears at high $K\pi$ mass! This is a surprising and helpful result, since it eliminates a lot of potential explanations for the effect and suggests some sort of final state interference phenomenon which we discuss in Chapter 8.

The signal in these fits is measured by subtracting wrong-sign events from right sign ones, when $0.8 < M(K\pi) < 1.0\,\mathrm{GeV}/c^2$. This is a somewhat narrow band of mass, and yet we see a large variation within this band. None of the cuts we use sculpt this $K\pi$ mass spectrum noticeably, so any hypothetical "ill effects" from the mismodeling of cuts by our Monte Carlo are now all but eliminated from consideration.

7.5 Some Known Backgrounds to $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$

There are a few charm decays which are prominent sources of backgrounds in this analysis. Understanding the nature of these as backgrounds will be important.

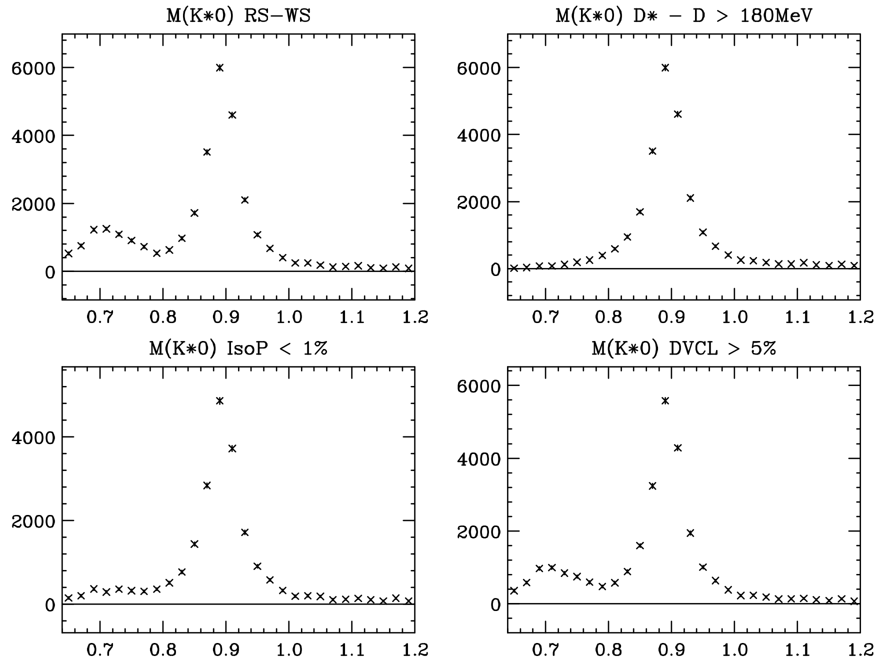

Figure 7.11: Evidence of the background from $D^{*+} \to D^0\pi^+ \to (K^-\mu^+\nu_\mu)\,\pi^+$ in data subject to three additional cuts. The upper-left plot is a loose-cut sample with no effective cuts against this background.

7.5.1 $D^{*+} \to D^0\pi^+ \to (K^-\mu^+\nu_\mu)\,\pi^+$

The decay chain $D^{*+} \to D^0\pi^+ \to (K^-\mu^+\nu_\mu)\,\pi^+$ is a fairly common one, and a significant background for $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu \to (K^-\pi^+)\,\mu^+\nu_\mu$. The D^* decay has a small phase space, so in the laboratory frame the D^0 and π^+ trajectories are nearly collinear. This makes it easy to mistakenly associate the pion track with the D^0 decay vertex. Since the pion is nearly at rest in the D^0 rest frame, and the kaon is also slow in that frame being the most massive decay product, the $K\pi$ invariant mass is low, near threshold. This gives us a large right-sign contamination below the K^{*0} mass peak in the $K\pi$ mass plots.

Figure 7.11 shows clear evidence for this background present in the data as a bump in $K\pi$ mass, in between threshold and the K^{*0} resonance. Three different cuts are applied which are more or less effective at reducing this background. The explicit $M(D^{*+}) - M(D^0)$ mass difference cut is apparently the most effective at reducing the

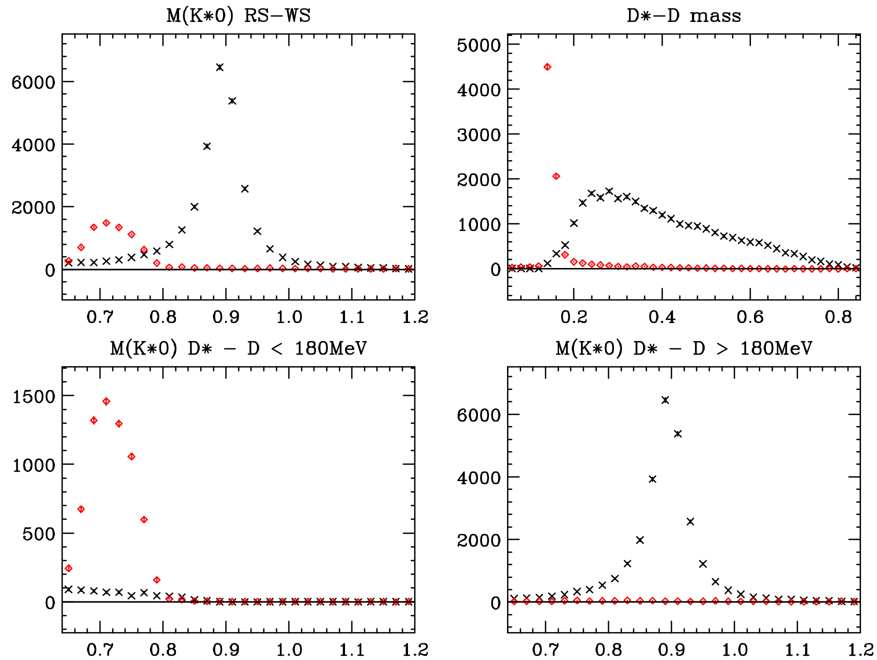

Figure 7.12: Response of the simulated signal and background to an explicit cut on $M(D^{*+}) - M(D^0)$. The distributions in black (X's) are wrong-sign subtracted plots of the simulated $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$decay. The red distribution is for the simulated $D^0 \to K^-\mu^+\nu_\mu$ background process. In the signal sample, events having a D^{*+} in the recoil state are removed.

reflection peak, which would present difficulties when measuring the K^{*0} yield with a fit.

In this analysis, we generally measure the K^{*0} yield by counting right-sign events and subtracting wrong-sign events in the $K\pi$ mass window from 0.8–1.0 GeV/c^2. Figure 7.12 shows simulated $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ signal and the $D^{*+} \to D^0\pi^+ \to (K^-\mu^+\nu_\mu)\,\pi^+$ background separately. Here we see the distribution in $K\pi$ mass for the background, and how the $M(D^{*+}) - M(D^0)$ mass difference cut affects it. Although it would be more apparent with looser cuts than the ones used here, there is a small tail of this background that survives the mass difference cut.

If we relax the DCL and ℓ/σ cuts from the baseline, and look entirely in the $K\pi$ mass window from 0.8–1.0 GeV/c^2, we see a trace of background from the $D^0 \to K^-\mu^+\nu_\mu$ decay. Comparing the red and blue points in Figure 7.13, we see that the

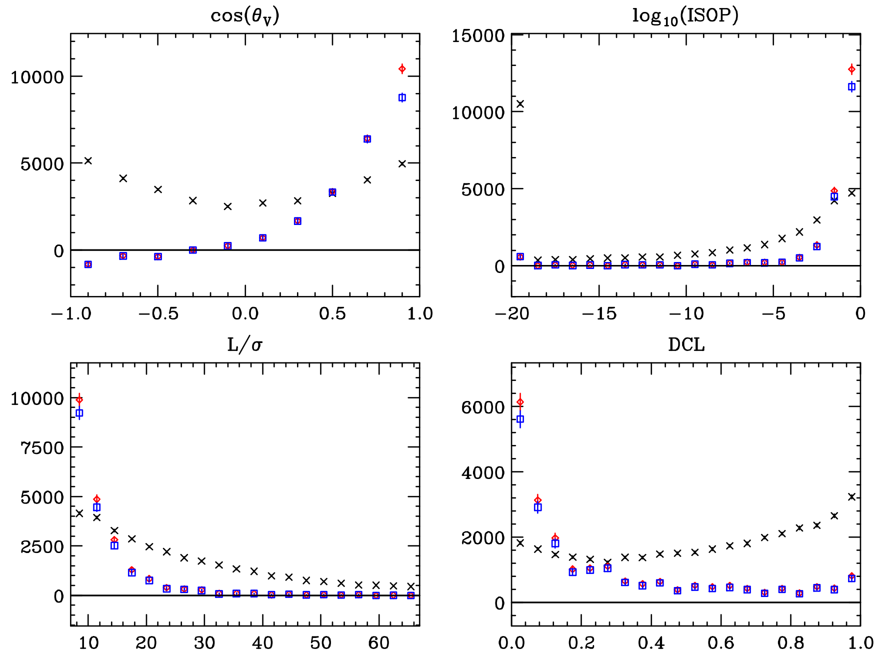

Figure 7.13: Properties of the $D^{*+} \to D^0\pi^+ \to (K^-\mu^+\nu_\mu)\,\pi^+$ background in the K^{*0} mass region. We show simulated signal and $D^{*+} \to D^0\pi^+ \to (K^-\mu^+\nu_\mu)\,\pi^+$ background in four distributions where the difference is most striking. The cuts are relaxed from the baseline cuts: DCL $> 1\%$, $\ell/\sigma > 7$, and no $D^* - D$ mass difference cut. The black distribution is simulated $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ signal, and the blue and red distributions are this background, amplified 12 times with respect to the signal. The blue distribution has the additional $M(D^{*+}) - M(D^0) > .180\mathrm{GeV}/c^2$ cut applied.

mass difference cut does little to the background directly under the K^{*0} resonance. So when we are not concerned with fitting the K^{*0} peak, the mass difference cut becomes largely cosmetic. Cuts on DCL, ℓ/σ, and primary vertex isolation (ISOP) will all help to control this part of the background.

It is important to control this background and backgrounds due to muon misidentification in hadronic D^0 decays. Because the pion from the D^{*+} is so slow in the supposed D^0 rest frame, in the K^{*0} rest frame the pion is always heading in the direction of the D^0. This causes $\cos\theta_V$ to peak toward 1 (see Figure 7.13). Since this creates an asymmetry in $\cos\theta_V$ opposite to the anomaly we are concerned with, this background actually tends to mask the asymmetry problem.

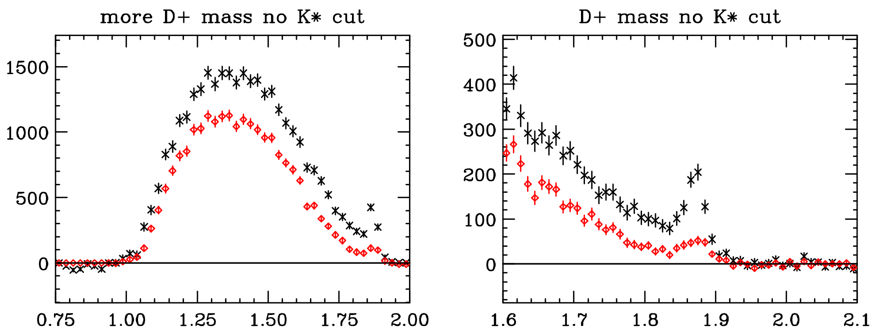

Figure 7.14: $M(K\pi\pi)$ mass spectrum. The background contamination from $D^+ \to K^-\pi^+\pi^+$ is evident as a small bump on the right side. In black is a wide, skim-level K^{*0} mass cut requiring $M(K\pi) < 1.3$ GeV/c^2. The red distribution primarily isolates the K^{*0} resonant part of the spectrum by demanding $0.8 < M(K\pi) < 1.0 GeV/c^2$. The particular muon cuts used are relevant here: CL$\mu_{\text{(new)}} > .05$, missing μ planes ≤ 1, and $|P_\mu| > 10$ GeV/c.

7.5.2 $D^+ \to K^-\pi^+\pi^+$

An obvious source of background contamination is $D^+ \to K^-\pi^+\pi^+$. If either of the two pions are misidentified as muons, this could appear as a right-sign $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ decay. This ought to give us a broad contribution the the $K\pi$ mass spectrum as well as a legitimate K^{*0} bump. According to the 2000 PDG[2], the K^{*0} resonance contributes to about 14% of the $K^-\pi^+\pi^+$ decay mode.

We can deal with this simply by reconstructing the mass of the D^+ under the assumption that there are no other neutral particles such as neutrinos, and that the muon track really has the mass of a pion. Because pion in-flight decay is a common mechanism causing mis-identification as a muon, we expect that the pion momentum may be poorly measured. The pion track may have "kinked" while in between PWC chambers. We expect more than the usual amount of smearing of the D^+ mass peak because one of the pions probably has a bad momentum measurement.

It turns out that we see some, but surprisingly little, evidence of this particular background. Figure 7.14 shows the $M(K\pi\pi)$ mass spectrum in FOCUS data. It is apparent from the figure that the level of this background is small, and also that it can be easily removed with cuts on this mass.

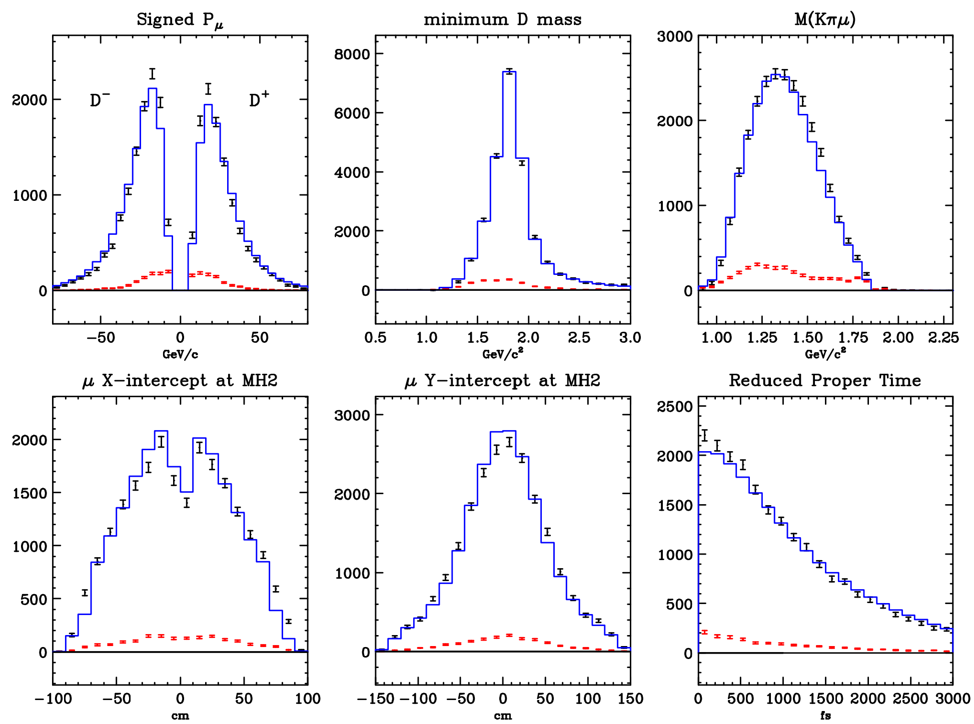

Figure 7.15: Distributions of some well-behaved cut variables and event properties. The black points are data, blue lines are predictions from a signal Monte Carlo, and red points are predictions from a background Monte Carlo.

7.6 Cut Response and Event Properties

It will be helpful at this stage to examine in detail the behavior of predicted and observed yields, and $\langle \cos\theta_{\rm V} \rangle$ as a function of various cut variables and event properties.

Figure 7.15 shows the distributions of some event properties which are matched well with the FOCUS Monte Carlo. This figure shows three samples: one from real FOCUS data, one from a signal simulation, and one from simulated backgrounds. The signal simulation is meant to indicate the response of pure $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ signal, although there may be some impurity. Each of the events in this simulated data set contain exactly one $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ decay. But there is no track-by-track match to the simulation's "God's Block", so whether each candidate found in these events represents the correct choice of well-measured tracks or not is indeterminate. Regardless, this sample is considerably more clean than the real data.

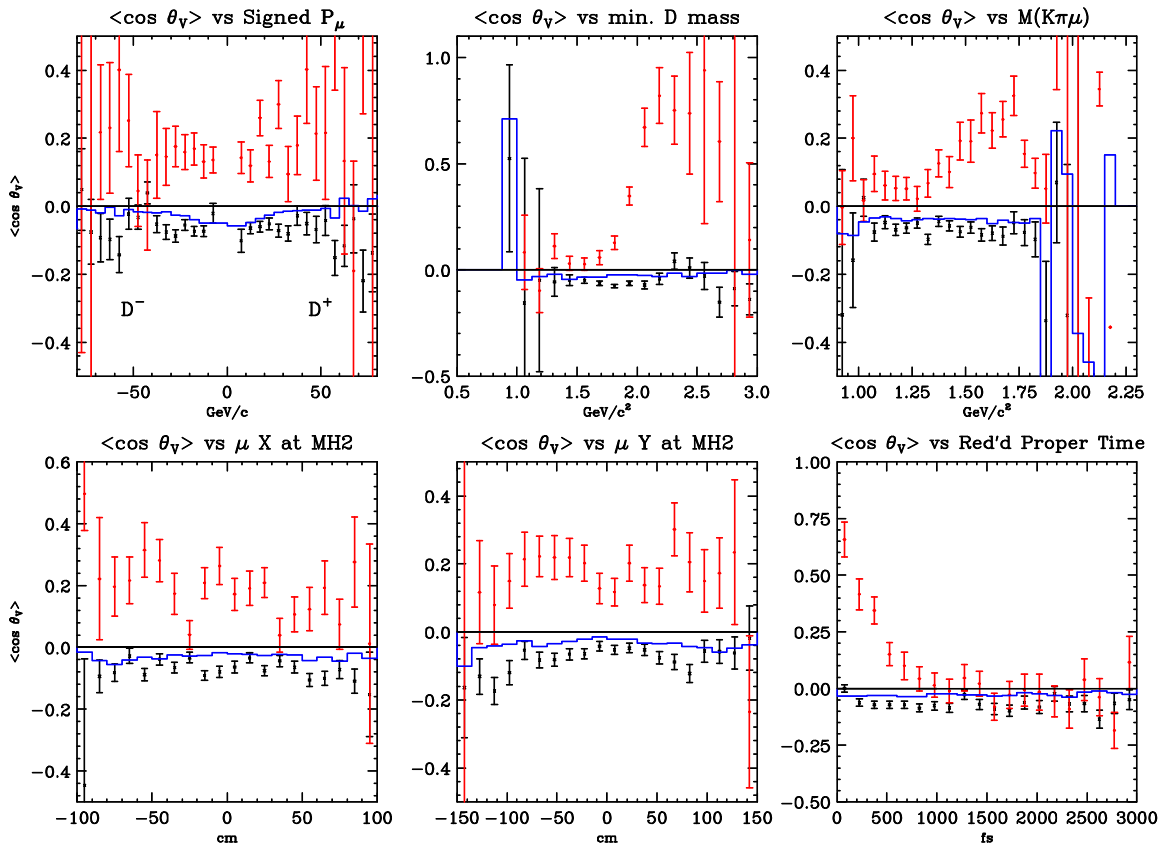

Figure 7.16: $\langle \cos\theta_V \rangle$ as a function of well-behaved event properties. Again black, blue, and red are data, signal Monte Carlo, and background Monte Carlo, respectively.

It's worth mentioning that in previous experience, matching reconstructed tracks to the particles in the "God's Block" introduces significant biases. The track-matching algorithm is not 100% efficient, and this efficiency varies with the kinematics of the underlying event. For this reason we do not use track matching in this analysis, but instead rely upon our analysis cuts and wrong-sign subtraction to remove all backgrounds from a pure-signal Monte Carlo.

The background simulation in Figures 7.15–7.18 is a $c, \bar{c}$ Monte Carlo, where roughly three times the FOCUS data were simulated. The background events are all those where the decay of interest, $D^+ \to \overline{K}^{*0} \mu^+ \nu_\mu$, was not simulated. This sample is normalized to the data by a factor which equalizes the measured yields of the data and full $c, \bar{c}$ Monte Carlo.

The results in Figure 7.15 are encouraging. There is a dip evident in the center of the muon X-intercept distribution. This is due to an inefficiency in the Proportional Wire Chambers (PWC's) near the center — the e^+/e^- pair region. The inefficiency is thought to be caused by combined effects such as polymerization of the wires and voltage sagging where the hit rate is high. If a muon fails to fire enough of the wires, its track will not be reconstructed. These effects have been included in the simulation, as is apparent from the plot.

There is still a considerable amount of background present in the wrong-sign-subtracted signal, and it is evidently difficult to remove it with any of these cuts. The distribution of the right-sign-excess background events is noticeably different in the $|P_\mu|$ and $M(K\pi\mu)$ distributions but broad in both cases.

Looking at the way $\langle \cos\theta_{\rm V} \rangle$ varies as a function of these event properties, we learn more about the anomaly. Figure 7.16 shows these behaviors. The first thing to note is how $\langle \cos\theta_{\rm V} \rangle$ is, on average, positive for the simulated backgrounds. From this we learn that our anomalous $\cos\theta_{\rm V}$ asymmetry is an even greater effect than is observed with the baseline cuts, since it is partly canceled out by backgrounds!

In the reduced-proper-time plots, we see some variation in the background that suggests two contributions. One background source must be short-lived and have a large positive $\cos\theta_V$ asymmetry, while another background contribution is longer-lived and has little asymmetry. We also see a variation of this background in the minimum D mass distribution.

Figure 7.17 shows distributions of event properties that are apparently not well-simulated. The ℓ/σ_ℓ (detachment) distribution does not match the predicted distribution below 20, which is probably due to a more severe mismatch in the expected detachment uncertainty, σ_ℓ. The poor σ_ℓ modeling is likely to be the result of a severe mismatch in the primary vertex multiplicity distribution. The primary vertex multiplicity is difficult to simulate accurately, because it is controlled by production dynamics which are not entirely understood. The present distribution is about the best that can be achieved by adjusting free parameters in the Pythia production model, while maintaining a reasonable match to the D momentum spectra. Looking at the $\langle\cos\theta_V\rangle$ plots, we see that for σ_ℓ and the primary vertex multiplicity, the signal Monte Carlo and data track each other well (with a constant offset). This implies that while the failures in these aspects of the FOCUS Monte Carlo may contribute slightly to the systematics of the form factor measurement, they have nothing to do with the $\cos\theta_V$ anomaly. By examining the plot of $\langle\cos\theta_V\rangle$ versus ℓ/σ, we can better understand where some of this background is coming from. Apparently the anti-D^* cut did not completely eliminate backgrounds from the decay $D^{*+} \to D^0\pi^+ \to (K^-\mu^+\nu_\mu)\,\pi^+$, as this short-lived background with high $\langle\cos\theta_V\rangle$ is plainly evident in the background Monte Carlo and in the data. This also explains the behavior in Figure 7.8, starting with the baseline fit and increasing ℓ/σ to 12, 16, and 20 in ls1, ls2, and ls3.

The DCL distribution shows that, although the distribution is not perfectly simulated, the simulated backgrounds follow the signal distribution well when DCL $>$ 5% or so. One should expect the backgrounds from $D^{*+} \to D^0\pi^+ \to (K^-\mu^+\nu_\mu)\,\pi^+$ to tend toward low DCL, since in these events the pion originates in the production vertex, not the D^+ decay vertex.

The poor match to the muon CL distribution is unfortunate, but this does not create a bias in the analysis, because the CL distribution is quite uniform across the

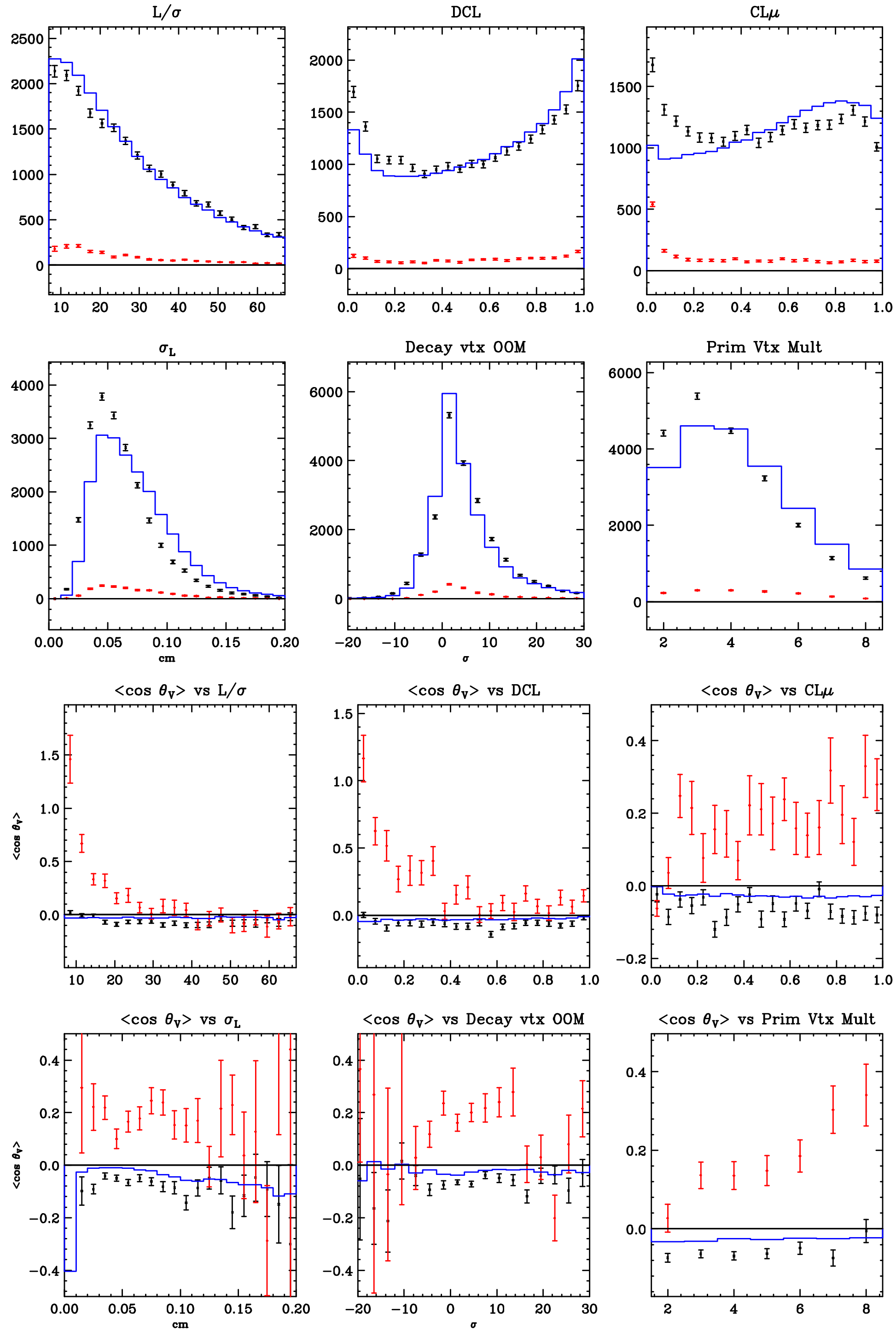

Figure 7.17: Distributions of some problematic event properties. "Decay vtx OOM" is the distance between the decay vertex and the nearest upstream face of material. If this is negative, the decay occured inside the material. $\langle\cos\theta_V\rangle$ versus these properties is also shown.

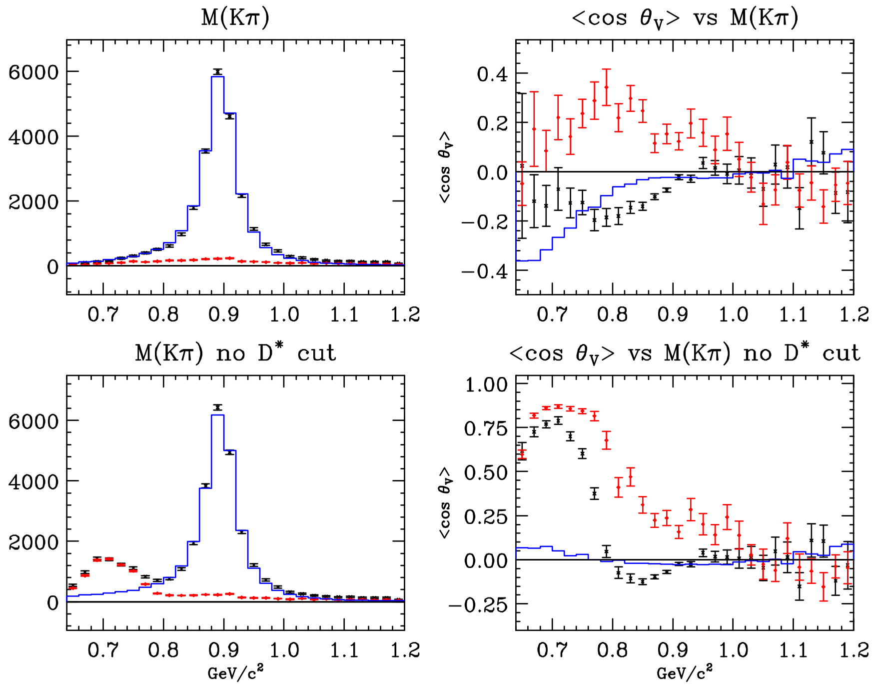

Figure 7.18: $K\pi$ mass distributions and $\cos\theta_{\rm V}$ asymmetry as a function of $K\pi$ mass. The "no D^* cut" plots have no cut on the $M(D^{*+}) - M(D^0)$ mass difference and only a 1% DCL cut as opposed to the 5% cut of the baseline. This allows backgrounds from $D^{*+} \to D^0\pi^+ \to (K^-\mu^+\nu_\mu)\,\pi^+$ to come in at low $K\pi$ mass. The black points are data, blue lines are signal Monte Carlo, and the red is a background Monte Carlo.

entire aperture in both data and Monte Carlo. We see some fully-hadronic backgrounds at low CLμ, below about 5%.

Figure 7.18 is one of the best tools with which to understand the anomalous $\cos\theta_{\rm V}$ asymmetry. The data shows the rapid development of a negative asymmetry beginning at the K^{*0} pole. The prediction in blue seems to exhibit this same trend except it starts at a much lower $K\pi$ mass. This turns out to be misleading, since at this point the prediction is only showing the effects of the $M(D^{*+}) - M(D^0)$ mass difference cut. Without this cut, the data still dips down, yet the prediction becomes flat at nearly zero asymmetry. With no anti-D^* cut, both data and background Monte Carlo show a significant bump to the left of the K^{*0} resonance, where $\langle\cos\theta_{\rm V}\rangle$ approaches $+1$. Section 7.5.1 explained this background in detail.

With these studies, combined with the cut variant and split sample fits, it is clear that the asymmetric anomaly is not an artifact of the spectrometer acceptance or unmatched sculpting due to a mis-modeled cut. We also see that no known background processes are responsible.

The remaining possible explanations for our $\cos\theta_V$ asymmetry are either a failure with the decay rate model, or a new (and not rare) background. Since the anomaly is stable against tightening particle ID cuts and lifetime, this mystery background must be a new D^+ decay with a kaon, pion, muon, and neutral daughters.

7.7 $D^+ \to \overline{K}_1^0(1270)\mu^+\nu_\mu$

The decay $D^+ \to \overline{K}_1^0(1270)\mu^+\nu_\mu$ can reach the final states $K^-\pi^+\pi^0\mu^+\nu_\mu$ and $K^-K^0\pi^+\mu^+\nu_\mu$. A significant fraction of these final states will contain the K^{*0} resonance. The neutral meson will not be detected, and so these events would pass all of the usual vertexing and particle ID cuts. We should expect the visible mass distribution of these events to be lower than the signal events because of the missing momentum.

There are no experimental observations of $D^+ \to \overline{K}_1^0(1270)\mu^+\nu_\mu$, but two limits have been measured. One is $\frac{\Gamma((K^-\pi^+)\pi^0\mu^+\nu_\mu)}{\Gamma(K^-\pi^+\mu^+\nu_\mu)} < 0.052$ at 95% CL assuming a phase-space $D^+ \to \overline{K}^{*0}\pi^0\mu^+\nu_\mu$ decay [8][3]. The other is $\frac{\Gamma(\overline{K}_1^0(1270)\mu^+\nu_\mu)}{\Gamma(\overline{K}^{*0}\mu^+\nu_\mu)} < 0.78$ at 95% CL [6]. The first of these roughly implies a limit of $\frac{\Gamma(\overline{K}_1^0(1270)\mu^+\nu_\mu)}{\Gamma(\overline{K}^{*0}\mu^+\nu_\mu)} < 0.873$ because $\mathrm{BR}(K_1^0(1270) \to (K^+\pi^-)\pi^0) = .045 \pm .011$ [2].[1] One theoretical prediction from quark wave-function models is that $\frac{\Gamma(\overline{K}_1^0(1270)\mu^+\nu_\mu)}{\Gamma(\overline{K}^{*0}\mu^+\nu_\mu)} \approx 0.06$ [9].

We consider this as a background because it appears to have the potential to create a bias in the form factor measurement. In fact, our simplistic models of the decay show that it would contribute a small negative bias in $\cos\theta_V$, so it has the potential to account for the observed anomaly. We present here a search for the decay mode $D^0 \to K_1^-(1270)\mu^+\nu_\mu \to (K^-\pi^+\pi^-)\mu^+\nu_\mu$ and infer from this an upper

[1] The E687 measurement at 1σ is $0 \pm .026$. A Baysian 95% confidence interval contains branching ratios less than 0.873.

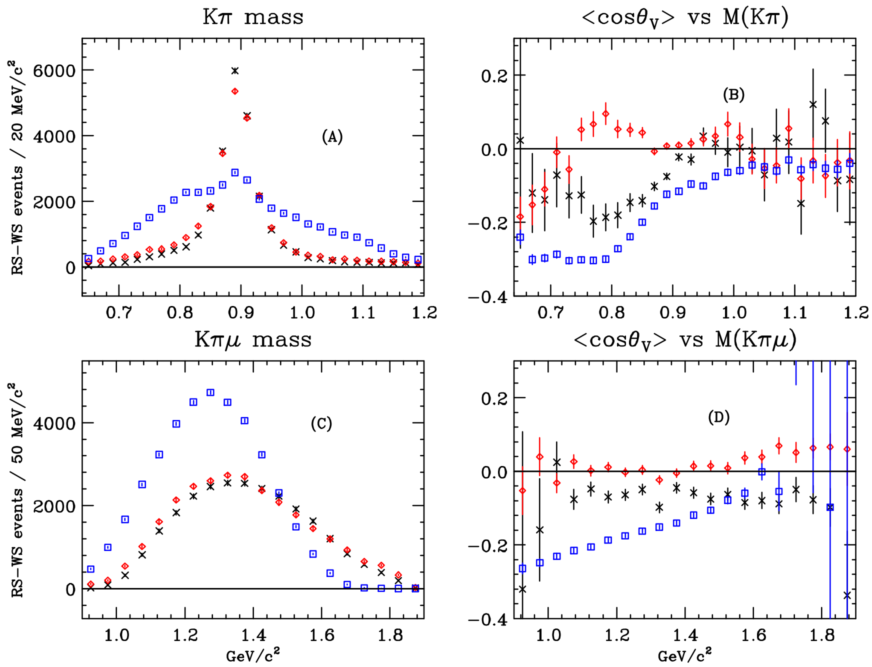

Figure 7.19: Some properties of simulated $D^+ \to \overline{K}_1^0(1270)\mu^+\nu_\mu$ background. The black crosses represent the data, red diamonds a full $c, \bar{c}$ Monte Carlo, and blue squares represent the $D^+ \to \overline{K}_1^0(1270)\mu^+\nu_\mu$ Monte Carlo. Both Monte Carlos are normalized to equal yield in the K^{*0} resonance region.

limit on the branching ratio for $D^+ \to \overline{K}_1^0(1270)\mu^+\nu_\mu$. This limit excludes it from further consideration as a background.

7.7.1 Simulated Signal

We simulate the inclusive decays $D^+ \to \overline{K}_1^0(1270)\mu^+\nu_\mu$, and apply our analysis cuts for $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$. This produces the signature of $D^+ \to \overline{K}_1^0(1270)\mu^+\nu_\mu$ as a background. Because we do not have a proper model for this axial-meson semileptonic decay, we simply draw a $K_1^0(1270)$ mass according to a (spinless) Breit-Wigner, draw the 3-body decay in uniform phase space, and subsequently decay the $K_1^0(1270)$ according to its measured sub-structure and branching ratios.

Figure 7.19 shows some distinguishing properties of this simulated background compared with the observed data and a Monte Carlo (without any $K_1^0(1270)$). The

$K\pi$ mass distribution shows a small K^{*0} resonant component, but is predominantly a more broad distribution. If a small amount of this background were present in the data, the resulting distortion to the line shape may go unnoticed. The more alarming distribution is $\langle\cos\theta_{\mathrm{V}}\rangle$ as a function of $K\pi$ mass. Here we see a behavior that closely mimics the data, although the asymmetry is not much stronger than the observed one. At this point the $K_1^0(1270)$ decay modes become candidates to explain the anomaly, and must be studied further.

The visible mass distribution is essentially the only other property that distinguishes the $K_1^0(1270)$ background from the actual signal. (The minimum D mass response is similar, since the two are highly correlated.) Here we see a background distribution that quite dissimilar from what is seen in the data, and this creates initial doubts about the existence of a significant contamination from this background.

We can go one step further and use the predicted $\langle\cos\theta_{\mathrm{V}}\rangle$ as a function of $K\pi$ mass, in conjunction with the observed $\langle\cos\theta_{\mathrm{V}}\rangle$ discrepancy, to predict how much $D^+ \to \overline{K}_1^0(1270)\mu^+\nu_\mu$ must be present to fully explain the anomaly. We do this in bins of $K\pi$ mass, and re-create what the mass distribution for the background would be, then compare it with the distribution in Figure 7.19 A.

Alternately, we can rely instead on the simulated $K\pi$ mass distribution of the $K_1^0(1270)$ background. For various trial values of the $D^+ \to \overline{K}_1^0(1270)\mu^+\nu_\mu$ branching ratio we can use this to reconstruct what the $\langle\cos\theta_{\mathrm{V}}\rangle$ as a function of $K\pi$ mass ought to be. Next, this can be compared to Figure 7.19 B.

The mathematics behind this is straight forward. Let A be $\langle\cos\theta_{\mathrm{V}}\rangle$, and Y be the yield in each bin of $K\pi$ mass. We will use the subscripts O, S, and B to refer to the observation (data), K^{*0} pure signal, and background (from $K_1^0(1270)$), respectively. Then:

$$A_O = \frac{A_S Y_S + A_B Y_B}{Y_B + Y_S} = A_B \frac{Y_B}{Y_O}$$

No asymmetry in $\cos\theta_{\mathrm{V}}$ is predicted for pure signal, so $A_S = 0$. This relation can be solved for either Y_B or A_B.

Figures 7.20 and 7.21 show the first cross-check (using $Y_B = Y_O \frac{A_O}{A_B}$), which is

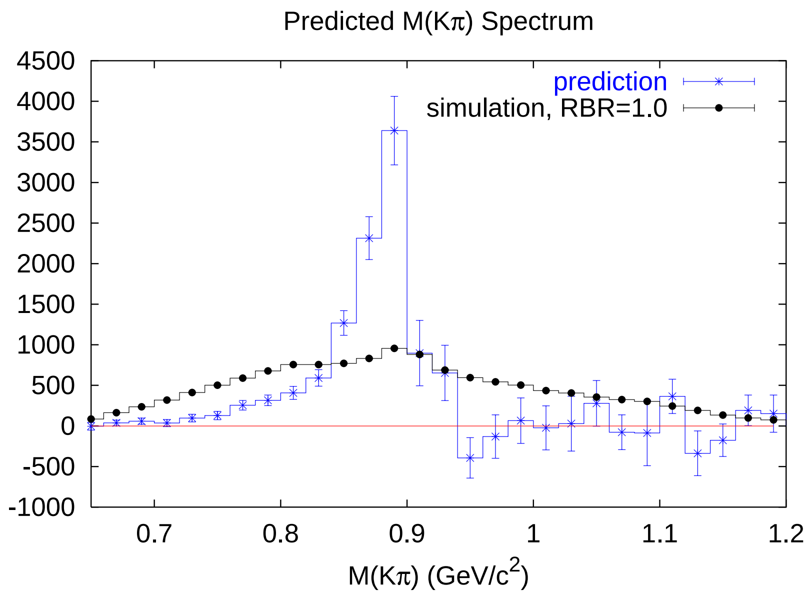

Figure 7.20: On the left is the predicted and simulated $K\pi$ mass distribution for $K_1^0(1270)$. The simulation is normalized to represent the expected yield if the $D^+ \to \overline{K}_1^0(1270)\mu^+\nu_\mu$ inclusive decay mode had the same branching fraction as the $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ decay mode.

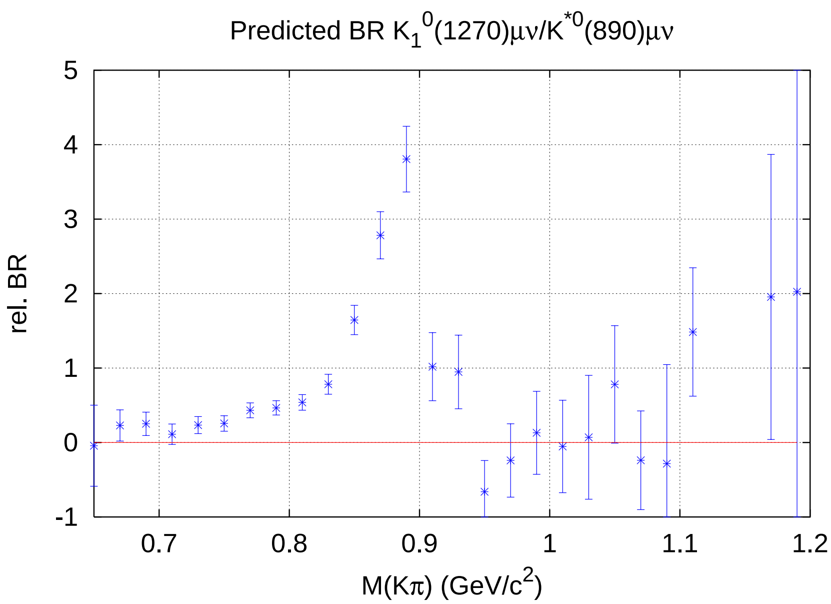

Figure 7.21: The ratio of the prediction over the simulation in Figure 7.20 produces a bin-by-bin predicted relative branching ratio of the two decay modes.

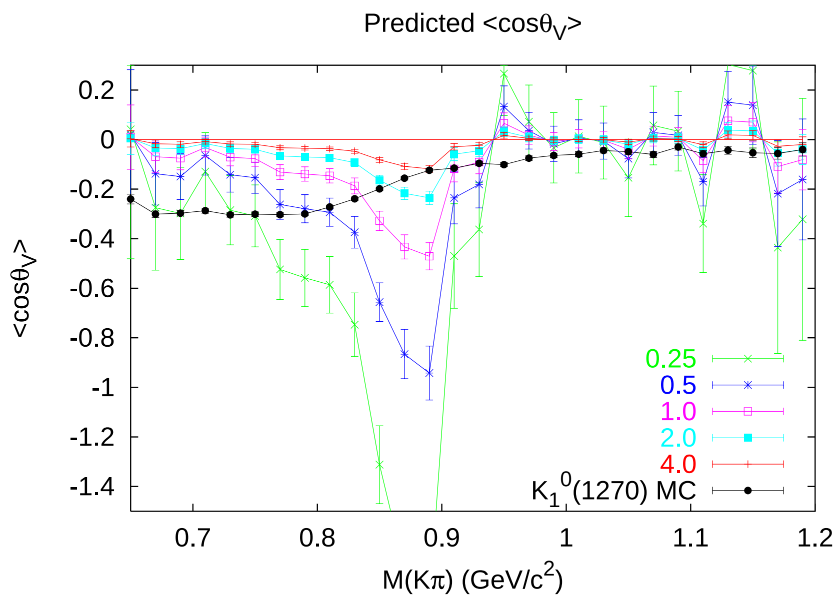

Figure 7.22: The prediction of $\langle \cos\theta_{\rm V}\rangle$ if the background from $D^+ \to \overline{K}_1^0(1270)\mu^+\nu_\mu$ would completely account for the observed $\langle \cos\theta_{\rm V}\rangle$ discrepancy. Branching ratios $\left(\frac{\Gamma(D^+\to\overline{K}_1^0(1270)\mu^+\nu_\mu)}{\Gamma(D^+\to\overline{K}^{*0}\mu^+\nu_\mu)}\right)$ are assumed for the various curves of 0.25, 0.5, 1, 2, and 4. The black curve is for reference, the results from the simulation (also shown in Figure 7.19).

again predicated on the validity of $\langle \cos\theta_{\rm V}\rangle$ as a function of mass for this background. Because the $D^+ \to \overline{K}_1^0(1270)\mu^+\nu_\mu$ decay has been simulated without taking into account any form factors or angular momentum correlations, we have some reason to suspect inaccuracies in the $\langle \cos\theta_{\rm V}\rangle$. But if there is any credibility to our simulated $\langle \cos\theta_{\rm V}\rangle$ versus $K\pi$ mass at all, then it is obvious that this decay mode alone cannot hope to account for the anomaly. Not only does the inferred $k\pi$ mass distribution look entirely different from the simulated one, but the amount of $D^+ \to \overline{K}_1^0(1270)\mu^+\nu_\mu$ needed to entirely explain the anomaly is outrageous. As mentioned earlier, previous searches for this decay mode rule out the possibility of branching ratios in excess of ≈ 1.

Because our $D^+ \to \overline{K}_1^0(1270)\mu^+\nu_\mu$ simulation does not include any of the dynamics of angular momentum conservation or form factors, it is easier to believe that the $K\pi$ mass distribution is accurately simulated, while $\cos\theta_{\rm V}$ distributions are not. We

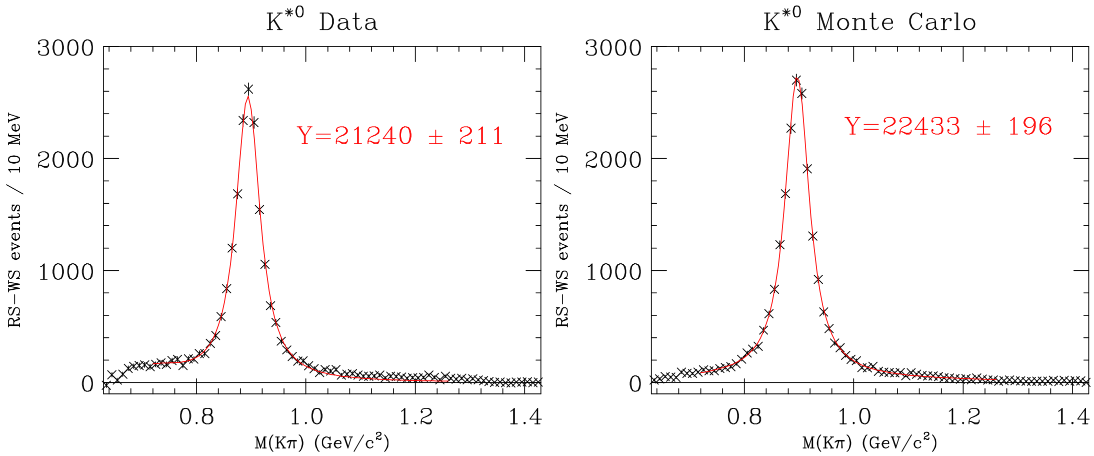

Figure 7.23: Normalizing fits for $D^0 \to K_1^-(1270)\mu^+\nu_\mu \to (K^-\pi^+\pi^-)\mu^+\nu_\mu$ search. These are measurements of the yield of $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ for data and for Monte Carlo. Their ratio determines the proper normalization of the $D^+ \to \overline{K}_1^0(1270)\mu^+\nu_\mu$ Monte Carlo.

can use this mass distribution to predict $\langle\cos\theta_{\rm V}\rangle$ for our mystery background by the relation $A_B = A_O\frac{Y_O}{Y_B}$. Regardless of the internal dynamics of the $D^+ \to \overline{K}_1^0(1270)\mu^+\nu_\mu$ decay, $\langle\cos\theta_{\rm V}\rangle$ cannot be larger than 1. According to the figure, we require that $\frac{\Gamma(\overline{K}_1^0(1270)\mu^+\nu_\mu)}{\Gamma(\overline{K}^{*0}\mu^+\nu_\mu)} > 0.5$.

7.7.2 Search for $D^0 \to K_1^-(1270)\mu^+\nu_\mu \to (K^-\pi^+\pi^-)\mu^+\nu_\mu$

Our most conservative branching ratio requirement of $\frac{\Gamma(\overline{K}_1^0(1270)\mu^+\nu_\mu)}{\Gamma(\overline{K}^{*0}\mu^+\nu_\mu)} > 0.5$ is not favored by quark model predictions, but is not quite ruled out by previous measurements. Here we present a somewhat informal search for the related decay $D^0 \to K_1^-(1270)\mu^+\nu_\mu$ in FOCUS data. This decay mode will be easier to search for, and is related to $D^+ \to \overline{K}_1^0(1270)\mu^+\nu_\mu$ by isospin. The $K_1^-(1270)$ can decay into the all-charged final state $K^-\pi^+\pi^-$.

The technique used here is to search for both $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu \to (K^-\pi^+)\mu^+\nu_\mu$ and $D^0 \to K_1^-(1270)\mu^+\nu_\mu \to (K^-\pi^+\pi^-)\mu^+\nu_\mu$ in FOCUS data and a Monte Carlo. In the Monte Carlo, we simulate both inclusive decay modes $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ and $D^0 \to K_1^-(1270)\mu^+\nu_\mu$ with equal absolute branching fractions. By isospin symmetry,

DCL	$> .01$
ℓ/σ_ℓ	> 8
ISO2ex	$< .001$
$\lvert P_\mu \rvert$	> 10 GeV/c^2
CL$\mu_{(\mathrm{new})}$	$> .05$
missing muon planes	≤ 1
Kaonicity	> 2
Pionicity (both pions)	> 1

Table 7.5: Cuts used in the search for $D^0 \rightarrow K_1^-(1270)\mu^+\nu_\mu \rightarrow (K^-\pi^+\pi^-)\mu^+\nu_\mu$

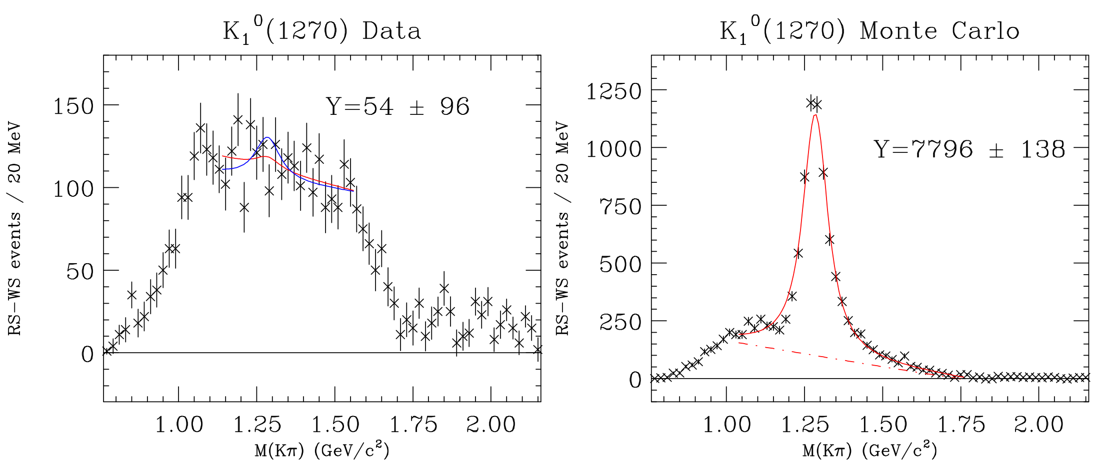

Figure 7.24: $D^0 \rightarrow K_1^-(1270)\mu^+\nu_\mu \rightarrow (K^-\pi^+\pi^-)\mu^+\nu_\mu$ fitted signals in data and Monte Carlo. The solid red lines in both plots are fits to the signal. The broken line underneath the signal in the Monte Carlo fit represents the background determined by the fit, which appears reasonable. The blue curve in the data represents a best-fit linear background plus a Breit-Wigner with a fixed yield of 192 events.

we expect the widths of these decay modes to be equal, but the absolute branching fractions will differ by the ratio of the D^+ and D^0 widths. Present measurements of the lifetimes give us the ratio $\Gamma(D^0)/\Gamma(D^+) \simeq 1051\,fs/413\,fs \simeq 2.54$.

Figure 7.23 shows the fits to the normalizing K^{*0} signal. The Monte Carlo sample happens to be slightly larger than the FOCUS data set, by a factor of 1.056.

Figure 7.24 shows the $D^0 \to K_1^-(1270)\mu^+\nu_\mu \to (K^-\pi^+\pi^-)\mu^+\nu_\mu$ signal, which we do not see in the data. The cuts we use are shown in Table 7.5. Our sensitivity at 2σ (about 95% CL) is 192 events. If the absolute branching fractions were the same $\left(\frac{\Gamma(D^0 \to K_1^-(1270)\mu^+\nu_\mu)}{\Gamma(D^0)} = \frac{\Gamma(D^+ \to \overline{K}^{*0}\mu^+\nu_\mu)}{\Gamma(D^+)}\right)$, we would expect to see $7796/1.056 = 7383$ events. Instead we can set a limit:

$$\frac{\Gamma(K_1^-(1270)\mu^+\nu_\mu)}{\Gamma(D^0)} \simeq \frac{\Gamma(\overline{K}_1^0(1270)\mu^+\nu_\mu)}{\Gamma(D^+)}\frac{\Gamma(D^+)}{\Gamma(D^0)} < .026\frac{\Gamma(\overline{K}^{*0}\mu^+\nu_\mu)}{\Gamma(D^+)}$$

$$\frac{\Gamma(\overline{K}_1^0(1270)\mu^+\nu_\mu)}{\Gamma(\overline{K}^{*0}\mu^+\nu_\mu)} < 0.066 \quad (95\%\ \mathrm{CL})$$

We call this an informal limit because we have not performed any analysis of systematics. The fit technique for the signal could be easily improved. For example, one could first fit a background shape only, ignoring the signal region. Then one would freeze the background shape and fit exclusively for the yield. Other statistical techniques could also be used with less bias and more reliable confidence intervals. These refinements could perhaps change the limit by as much as $\pm 50\%$, but for the limited application of the analysis here, the conclusion would remain the same.

This limit of 0.066 is far below the required level of 0.5 at which backgrounds from $K_1^0(1270)$ could conceivably produce the $\cos\theta_\mathrm{V}$ anomaly. Our informal limit is an order of magnitude better than the present limit, and we were tantalizingly close to an observation if this branching ratio prediction of 0.06 is true. But our single-minded conclusion is that the $\overline{K}^{*0}\mu^+\nu_\mu$ decay is not a relevant background in the form factor ratio fits.

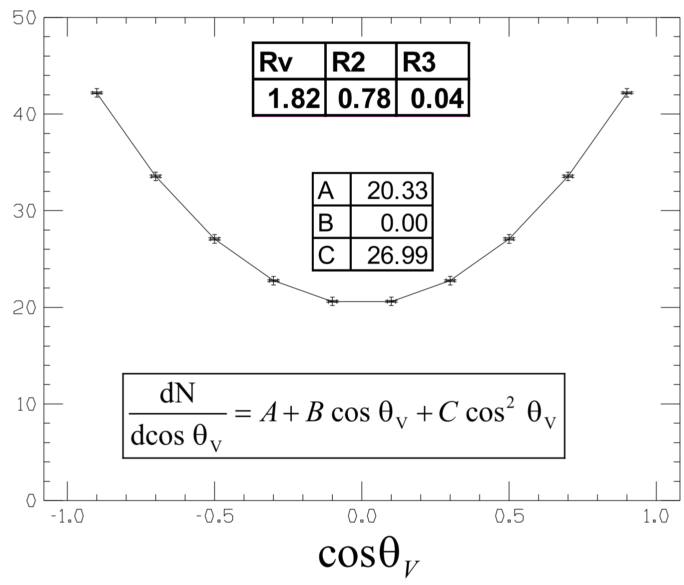

Figure 7.25: We plot the $\cos\theta_{\mathrm{V}}$ projection expected for the world average form factor ratios displayed in the figure. This is overlayed with a parabolic fit.

7.8 Deconvolution of the $\cos\theta_{\mathrm{V}}$ Distribution

As a way of exploring our problems with the form factor fits, we developed a method to correct the observed $D^+ \to \bar{K}^*\mu^+\nu$ decay distributions for both acceptance and resolution smearing. In particular, we are interested in whether or not a corrected $\cos\theta_{\mathrm{V}}$ would be an even function as theoretically expected for any set of input form factors. Figure 7.25 illustrates the expected $\cos\theta_{\mathrm{V}}$ for the World Average form factors. The distribution is well represented by a parabola with no term that is linear in $\cos\theta_{\mathrm{V}}$.

We will show that the corrected $\cos\theta_{\mathrm{V}}$ distributions obtained in our background-subtracted data have a statistically strong linear term in $\cos\theta_{\mathrm{V}}$ that contradicts the expected decay distributions. We have spent a great deal of time making sure that this effect is not due to Monte Carlo or background problems. Perhaps our most incisive demonstration is the stability of the $D^+ \to \bar{K}^*\mu^+\nu/K^-\pi^+$ branching ratio discussed in Chapter 5.

The nature of this linear term will be further explored in Chapter 8, where possible explanations in terms of interesting new physics will be offered.

7.8.1 Deconvolution Method

In this section, we describe a method for using a Monte Carlo simulation to "unfold" the measured distribution and obtain the underlying distribution that one would see in the absence of acceptance or resolution smearing. We consider the example of deconvolving the projection describing $d\Gamma/d\cos\theta_v$, where one integrates over the other kinematic variables such as Q^2, $\cos\theta_\ell$, and the coplanarity angle χ. We would begin by running a Monte Carlo simulation based on a default set of form factor ratios. With this we construct the correlation matrix, illustrated by Equation 7.1, that correlates the number of events N_{Mi}^{Gj} *generated* in the jth $\cos\theta_{\rm V}$ bin to the events that appear in the ith *measured* $\cos\theta_{\rm V}$ bin.

$$\begin{pmatrix} N_{M1}^{G1} & N_{M1}^{G2} & N_{M1}^{G3} \\ N_{M2}^{G1} & N_{M2}^{G2} & N_{M2}^{G3} \\ N_{M3}^{G1} & N_{M3}^{G2} & N_{M3}^{G3} \end{pmatrix} \tag{7.1}$$

For notational simplicity, we illustrated this matrix with a 3×3 form but will use more bins when confronting actual data.

$$\sigma_j = \frac{d\Gamma}{d(\cos\theta_{\rm V})_j}\Delta\cos\theta_{\rm V} \tag{7.2}$$

Let us assume that the matrix given by Equation 7.1 was predicated on a Monte Carlo based on the underlying partial width $\cos\theta_{\rm V}$ distribution that we write using the notation of Equation 7.2. The σ_j factors are valid for a given set of assumed form factor ratios. If the true integrated partial widths were $\tilde{\sigma}_j$ rather than the σ_j values assumed in the Monte Carlo, we would expect via a simple scaling argument the measured bin populations $\tilde{M}_j$ given by Equation 7.3.

$$\begin{pmatrix} \tilde{M}_1 \\ \tilde{M}_2 \\ \tilde{M}_3 \end{pmatrix} = \begin{pmatrix} N_{M1}^{G1}/\sigma_1 & N_{M1}^{G2}/\sigma_2 & N_{M1}^{G3}/\sigma_3 \\ N_{M2}^{G1}/\sigma_1 & N_{M2}^{G2}/\sigma_2 & N_{M2}^{G3}/\sigma_3 \\ N_{M3}^{G1}/\sigma_1 & N_{M3}^{G2}/\sigma_2 & N_{M3}^{G3}/\sigma_3 \end{pmatrix} \begin{pmatrix} \tilde{\sigma}_1 \\ \tilde{\sigma}_2 \\ \tilde{\sigma}_3 \end{pmatrix} \tag{7.3}$$

There is also an implicit assumption that the resolution and efficiency essentially "factorizes", in the sense that the matrix given by Equation 7.3 is independent of the assumed form factor. Presumably a necessary condition for this is that we are in the limit of sufficiently small $\Delta\cos\theta_{\mathrm{V}}$ such that the efficiency and resolution do not appreciably change over the bin width. But perhaps additional assumptions are required to insure the validity of Equation 7.3, and so we will want to test the method using pseudo-"data" and "MC" samples generated with disparate form factors.

If the matrix of Equation 7.3 is indeed independent of form factor assumptions to a precision comparable to our statistical error, we can simply invert it as shown in Equation 7.4 to reconstruct the underlying $\cos\theta_{\mathrm{V}}$ distribution from the measured bin populations.

$$\begin{pmatrix} \tilde{\sigma}_1 \\ \tilde{\sigma}_2 \\ \tilde{\sigma}_3 \end{pmatrix} = \begin{pmatrix} N^{G1}_{M1}/\sigma_1 & N^{G2}_{M1}/\sigma_2 & N^{G3}_{M1}/\sigma_3 \\ N^{G1}_{M2}/\sigma_1 & N^{G2}_{M2}/\sigma_2 & N^{G3}_{M2}/\sigma_3 \\ N^{G1}_{M3}/\sigma_1 & N^{G2}_{M3}/\sigma_2 & N^{G3}_{M3}/\sigma_3 \end{pmatrix}^{-1} \begin{pmatrix} \tilde{M}_1 \\ \tilde{M}_2 \\ \tilde{M}_3 \end{pmatrix} \tag{7.4}$$

Figure 7.26 illustrates four of the distributions used to obtain the matrix given in Equation 7.1. It shows the population in ten measured bins in $\cos\theta_{\mathrm{V}}$, for events in the four indicated slices of generated $\cos\theta_{\mathrm{V}}$. The smearing in the distribution is quite evident. This correlation matrix is then converted to the 10×10 matrix (illustrated by Equation 7.3) by dividing the elements by the integral of $d\Gamma/d\cos\theta_{\mathrm{V}}$ over each of the ten $\cos\theta_{\mathrm{V}}$ bins. The inverse of this matrix forms the deconvolution matrix that can be used to deconvolve experimental data.

Figure 7.27 provides a Monte Carlo test of the deconvolution method by showing a deconvolution of simulated $\cos\theta_{\mathrm{V}}$ and $\cos\theta_\ell$ distributions. In this test, the deconvolution matrix was performed using a substantially different set of form factors than those used for the simulated "data". The two overlayed points represent deconvolutions of the distributions for a "loose" set and "tight" set of analysis cuts. The tighter cuts eliminate roughly two-thirds of the event sample passing the looser cuts. The cut sets are compared in Figure 7.28.

Figure 7.27 confirms the method in several important ways. Both the simulated tight and loose cut sets unfold to the consistent distributions for both $\cos\theta_{\mathrm{V}}$ 7.27 (Figure 7.27 (a)) and $\cos\theta_\ell$ (Figure 7.27 (b)), even though there is a large difference

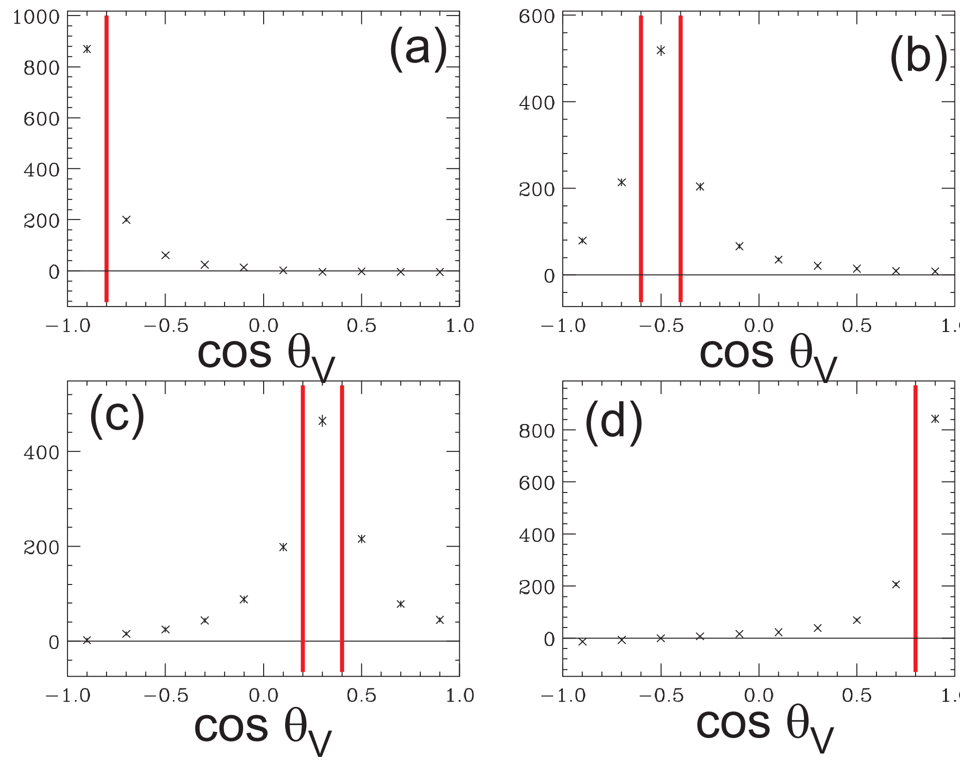

Figure 7.26: Here are four distributions that show the number of MC events that reconstruct in a given 0.2 width bin of $\cos\theta_{\rm V}$. (a) shows the reconstructed $\cos\theta_{\rm V}$ distribution for events that were generated with $\cos\theta_{\rm V} < -0.8$. (b) shows the reconstructed distribution for events generated with $-0.6 < \cos\theta_{\rm V} < -0.4$. (c) shows the distribution for events generated with $+0.4 < \cos\theta_{\rm V} < +0.6$. (d) shows the reconstructed distribution for events generated with $+0.8 < \cos\theta_{\rm V}$.

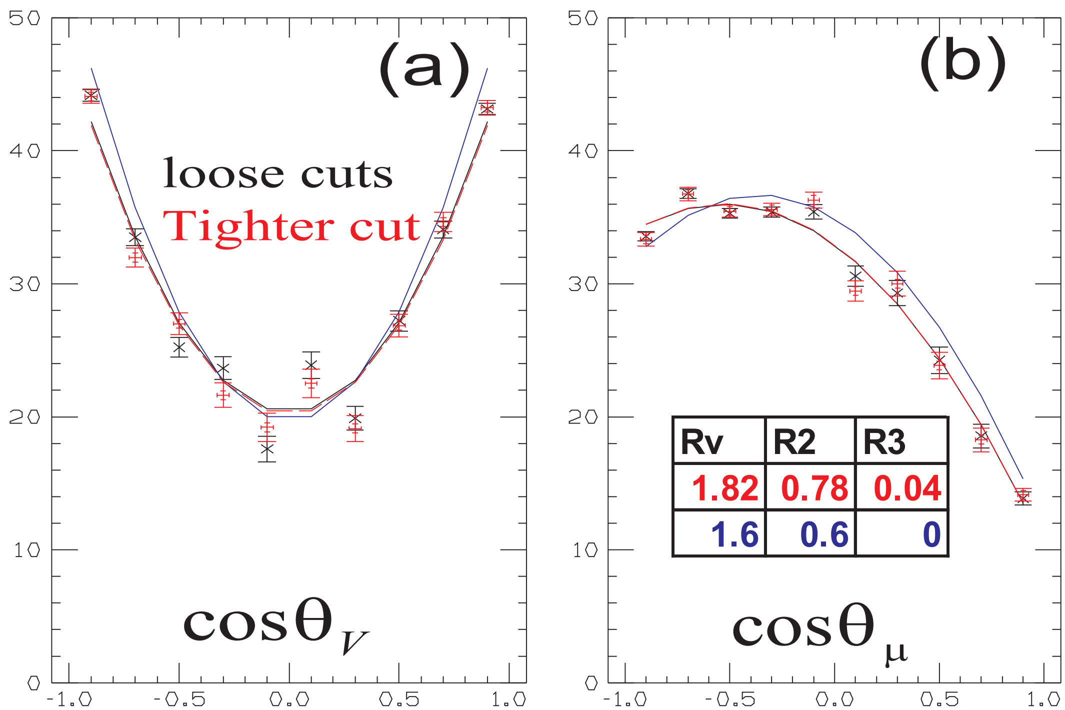

Figure 7.27: Monte Carlo tests of the deconvolution method for both (a) $\cos\theta_{\rm V}$ and (b) $\cos\theta_{\ell}$. The black points with error bars are unfolds using simulated data subjected to loose cuts. The red points with error bars are unfolds using simulated data subjected to much tighter cuts. The red curves are the generation level distributions with the same form factors used in the simulated data. The blue curves are the generation level distributions used to generate the (deliberately mismatched) MC simulation used to unfold the data. The form factor ratio used for the "data" and the unfold MC are given in the table in Figure (b).

	loose	tight		loose	tight
Hadronic	yes	yes	Not K 2pi	yes	yes
FSAA sub	yes	yes	No D*	yes	yes
L/sigma	>10	>12	clmunew		>10%
DCL	>5%	>10%	misspl		<1
iso2ex	<1%	<0.1%	IMU only	yes	yes
iso1	<10%	<10%	min parent		1.6-2
kaonicity	>2	>4	Pvis		>30
picon	>-5		Pmu		>10
pionicity		>1			
Eff wrt skim	53.5%	17.7%			

Figure 7.28: This table gives a summary of the cuts used for the tight and loose cut sets. The bottom of this table gives the efficiency of the two cuts relative to data passing the FSAA subskim.

in the underlying acceptances. In both cases, the unfolded data is consistent with the $d\Gamma/d\cos\theta_{\rm V}$ and $d\Gamma/d\cos\theta_\ell$ projections appropriate to the form factors consistent with the form factors used for the simulated "data" rather than the significantly different set used for the correction MC.

7.8.2 Deconvolution Results

For the data, we perform the deconvolution given by Equation 7.4 using measured yields $\tilde{M}_i$, determined by subtracting the WS event yield with $0.8 < M(K^-\pi^+) < 1\ GeV$ from the RS yield over this mass range. [2]

Figure 7.29 shows the results of the $\cos\theta_{\rm V}$ deconvolution applied to data for the loose cut (a) and tight cut (b) described in Figure 7.28 along with a parabolic χ^2 fit of the result. The data is wrong-sign subtracted and deconvoluted using a Monte Carlo with world average form factor ratios. The table gives the coefficient and error of the term linear in $\cos\theta_{\rm V}$ along with the confidence level of the fit. The fit finds a

[2] There is also a nifty way cast Equation 7.4 as a weighting factor and perform the deconvolution by fitting deconvolution weighted histograms. We did feel it was necessary to use this extention for this analysis.

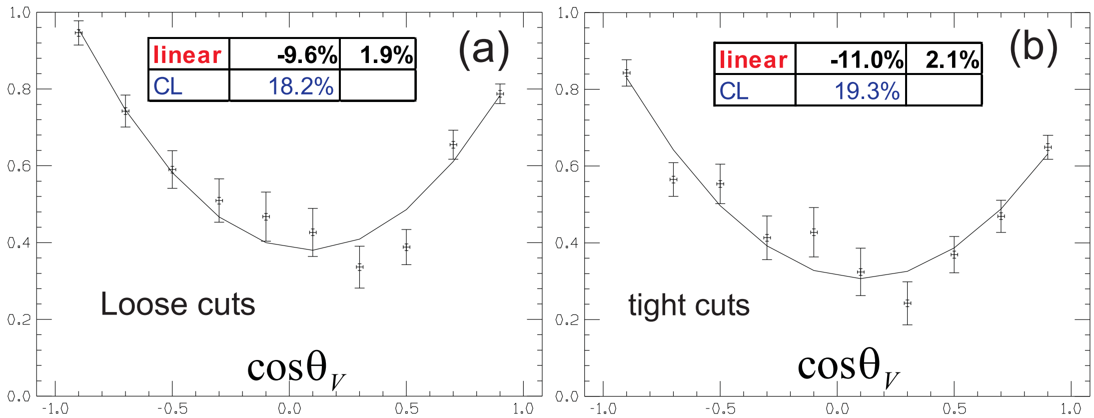

Figure 7.29: We show the result of the deconvolution applied to (a) loosely cut and (b) tightly cut data along with a parabolic fit. In both cases there is a term linear in $\cos\theta_{\mathrm{V}}$ at roughly the 5 σ level. As shown in Figure 7.25, the $\cos\theta_{\mathrm{V}}$ distribution should be an even function irrespective of the form factor ratio. The tables give the linear term, its error, and the confidence level of the parabolic fit.

linear term (not expected in the theory) at the $\approx 5\sigma$ level in both cut selections. We also note that the two cut sets do not unfold to exactly the same level. We believe this is due to the problems with the HC simulation that are discussed in Chapter 5.

The deconvolution was done with the default HC simulation at the time (the so-called SHW simulation). We now believe that the CRD HC model is a much better fit to the data. The tight sample that includes a cut on visible momentum will respond incorrectly relative to the loose sample in this MC run with this rather poor HC model. But the HC simulation model is expected to have very little influence in the $\cos\theta_{\mathrm{V}}$ shape which is consistent between Figure 7.29(a) and (b).

7.8.3 Summary

We have presented a method for unfolding the observed angular distributions in data for efficiency and the poor resolution in the variables due to neutrino closure. This method worked well according to several MC tests. When applied to data, we obtained strong evidence for a term linear in $\cos\theta_{\mathrm{V}}$ at about the 5σ level. No such linear term is expected theoretically for the decay $D^+ \to \bar{K}^* \mu^+ \nu$. This anomoly will be further explored in Chapter 8.

Chapter 8

A Model for the $\cos\theta_{\mathbf{V}}$ Asymmetry

In this chapter we present a possible explanation for our inability to get a satisfactory fit for the $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ decay intensity and our observation for an unexpected linear term in the $\cos\theta_{\mathrm{V}}$ distribution and its strong dependence on the $K^-\pi^+$ mass. We believe that our observations can be explained by including an s-wave $K^-\pi^+$ amplitude that interferes coherently with the $\overline{K}^{*0}$ in the $D^+ \to K^-\pi^+\mu^+\nu$ final state. Presumably this amplitude must be fairly small since the $K^-\pi^+$ mass distribution as discussed in Chapter 5 is nearly a perfect fit to a single $\overline{K}^{*0}$ Breit-Wigner.

8.1 The Toy Model

We have not done a complete explicit calculation of the full five fold differential decay rate including such an interfering amplitude, but rather simply extended the heuristic treatment presented in equation 2.1. We will refer to this as a "toy" model since our treatment is far from complete. For example, it ignores any dependence of the decay rate on the acoplanarity angle and neglects the mass terms.

$$\begin{aligned}\frac{d^2\Gamma}{d\cos\theta_v\ d\cos\theta_\mu} \ \propto\ & \mid BW|^2 \sum_{m=-1,+1} \Gamma_m \mid d^1_{1m}\ (\cos\theta_\mu)\ |^2\ |d^1_{0m}\ (\cos\theta_v)\ |^2 \\ & +\Gamma_0\ \mid d^1_{10}\ (\cos\theta_\mu)\ |^2 \mid d^1_{00}\ (\cos\theta_v)\ BW + Ae^{i\delta}|^2\end{aligned}$$

$$\text{Where } BW = \frac{\sqrt{M_o\Gamma}}{(M^2 - M_o^2) + i\ \Gamma\ M_o} \tag{8.1}$$

What we have done in Equation 8.1 is to add an s-wave constant amplitude contribution to the (Γ_0) piece describing the decay of the virtual W^+ when in its $|1,0>$ state. Such an s-wave contribution cannot conserve angular momentum with a W^+ in any other helicity state. The magnitude A and phase δ of this constant amplitude are left as unknown parameters.

To be specific, we will assume $(\Gamma_+\ ,\ \Gamma_0\ ,\ \Gamma_-) \propto (|H_+|^2\ ,\ |H_0|^2\ ,\ |H_-|^2)$ where the helicity based form factors will be the same forms assumed in Chapter 2.

$$\begin{aligned}
H_\pm(t) &= (M_D + M_{K\pi})A_1(t) \mp 2\frac{M_D K}{M_D + M_{K\pi}}V(t) \\
H_0(t) &= \frac{1}{2M_{K\pi}\sqrt{t}}\left[(M_D^2 - M_{K\pi}^2 - t)(M_D + M_{K\pi})A_1(t) - 4\frac{M_D^2 K^2}{M_D + M_{K\pi}}A_2(t)\right] \\
&\quad \text{Where } A_i(t) = \frac{A_i(0)}{1 - t/M_A^2} \quad V(t) = \frac{V(0)}{1 - t/M_V^2} \\
&\quad \text{and } M_A = 2.5\ GeV \text{ and } M_V = 2.1\ GeV
\end{aligned} \tag{8.2}$$

We will do our toy model calculations assuming $r_v \equiv V(0)/A_1(0)$ and $r_2 \equiv A_2(0)/A_1(0)$ are at their world average measured values. After evaluating the Wigner D-matrices, we have the fully explicit expression given by Equation 8.3.

$$\begin{aligned}
\frac{d^2\Gamma}{d\cos\theta_v\ d\cos\theta_\mu} \propto\ & \sin^2\theta_v\ \{(1+\cos\theta_\mu)^2\ \Gamma_+\ +\ (1-\cos\theta_\mu)^2\ \Gamma_-\}|BW|^2 \\
& +4\ |\cos\theta_{\rm V}\ BW + Ae^{i\delta}|^2 \sin^2\theta_\mu\ \Gamma_0 \\
& \text{Where } BW = \frac{\sqrt{M_o\Gamma}}{(M^2 - M_o^2) + i\ \Gamma\ M_o}
\end{aligned} \tag{8.3}$$

8.2 Mass Dependence of the $\cos\theta_{\rm V}$ Asymmetry in the Toy Model

We see from Equation 8.3 the linear $\cos\theta_{\rm V}$ dependence of the decay rate — presumably responsible for the forward-backward $\cos\theta_{\rm V}$ asymmetry we observed — comes from the "interference" cross term piece of the $|\cos\theta_{\rm V}\ BW + Ae^{i\delta}|^2$. Explicitly the linear

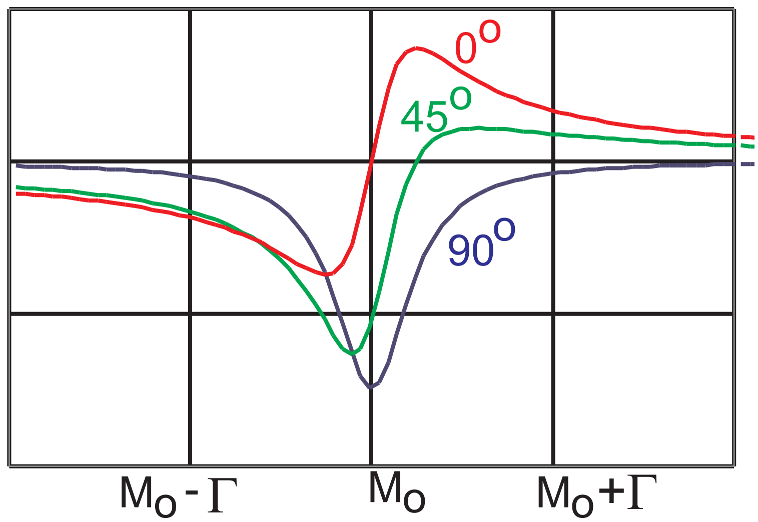

Figure 8.1: The interference term in the toy model. We plot the mass dependence of the linear $\cos\theta_{\rm V}$ term as a function of mass for three different phase choices.

$\cos\theta_{\rm V}$ piece is given by:

$$8\sin^2\theta_\mu\ \Gamma_0\ \frac{A\ \sqrt{M_o\Gamma}}{(M^2-M_o^2)^2+(\Gamma\ M_o)^2}\ \{(M^2-M_o^2)\ \cos\delta-\Gamma\ M_o\ \sin\delta\}\times\cos\theta_{\rm V} \quad (8.4)$$

We see from Equation 8.4 that the linear $\cos\theta_{\rm V}$ term has a t-dependence through $\Gamma_0 = |H_o(t)|^2$ as a well as dependence on a $\sin^2\theta_\mu$. The mass dependence is very sensitive to the unknown phase δ because of the characteristic mass dependence of the real and imaginary parts of a Breit-Wigner. This behavior is graphed in Figure 8.1.

For a relatively real phase $\delta = 0$ we expect that the asymmetry will reverse sign as one passes through the $\overline{K}^{*0}$ pole. For a relatively imaginary phase $\delta = \pi/2$, the $\cos\theta_{\rm V}$ asymmetry will have the same sign on either side of the pole. For $\delta = \pi/4$, we come very close to the observed behavior of the $\cos\theta_{\rm V}$ asymmetry as a function of $M_{K\pi}$ — both showing a strong negative asymmetry below the pole mass, and a small asymmetry above the pole mass.

8.3 Comparing the Asymmetries in the Toy Model to the Data

To further check the model, we wrote a simple program to numerically integrate Equation 8.3 in order to produce "conditional" projections. These projections are designed to crudely test the mass, Q^2, and $\cos\theta_\ell$ dependence of the $\cos\theta_{\mathrm{V}}$ distribution. The model and observed projections are compared in Figure 8.2 for events with $\sin^2\theta_\mu > 0.75$.

The model in the upper frame of Figure 8.2 is computed using an amplitude of $A = 0.3$ and a phase of $\delta = \pi/4$, as defined in Equation 8.3. This amplitude was chosen to roughly match the observed size of the asymmetry and its $K^-\pi^+$ mass dependence displayed for data in the lower frame. We show the predicted $\cos\theta_{\mathrm{V}}$ projections for events with $K^-\pi^+$ masses both below and above the $\overline{K}^{*0}$ pole and with Q^2 values both below and above 1.3 of its kinematic maximum of $Q^2_{\mathrm{max}} = (M_D - M_{K^-\pi^+})^2$.[1] The red points show the predictions for the case of zero amplitude. For the upper frame, the curves show the projection integrals where the amplitude is set to $A = 0$. For the data, the red points show the result of a MC run with no s-wave interference and world average form factor ratios.

Figure 8.2 shows that the asymmetry pattern in the data is remarkably reproduced by the toy model with the interfering s-wave $0.3\exp(i\pi/4)$ amplitude. There is essentially no asymmetry of the data relative to the MC above the pole mass — all of the asymmetry lies below the pole mass. The observed degree of the asymmetry is roughly the same for the high Q^2 and low Q^2 split and is in rough agreement with the amount in the toy model.

Figure 8.3 repeats the comparison between the toy model and observed asymmetry for events with $\sin^2\theta_\mu < 0.75$. Again, the toy model with the interfering s-wave $0.3\exp(i\pi/4)$ amplitude does a reasonably good job at reproducing the asymmetry pattern observed in the data.

[1] The Q^2 and $\sin^2\theta_\mu$ cuts were roughly chosen to produce equal yield on either side of the splits.

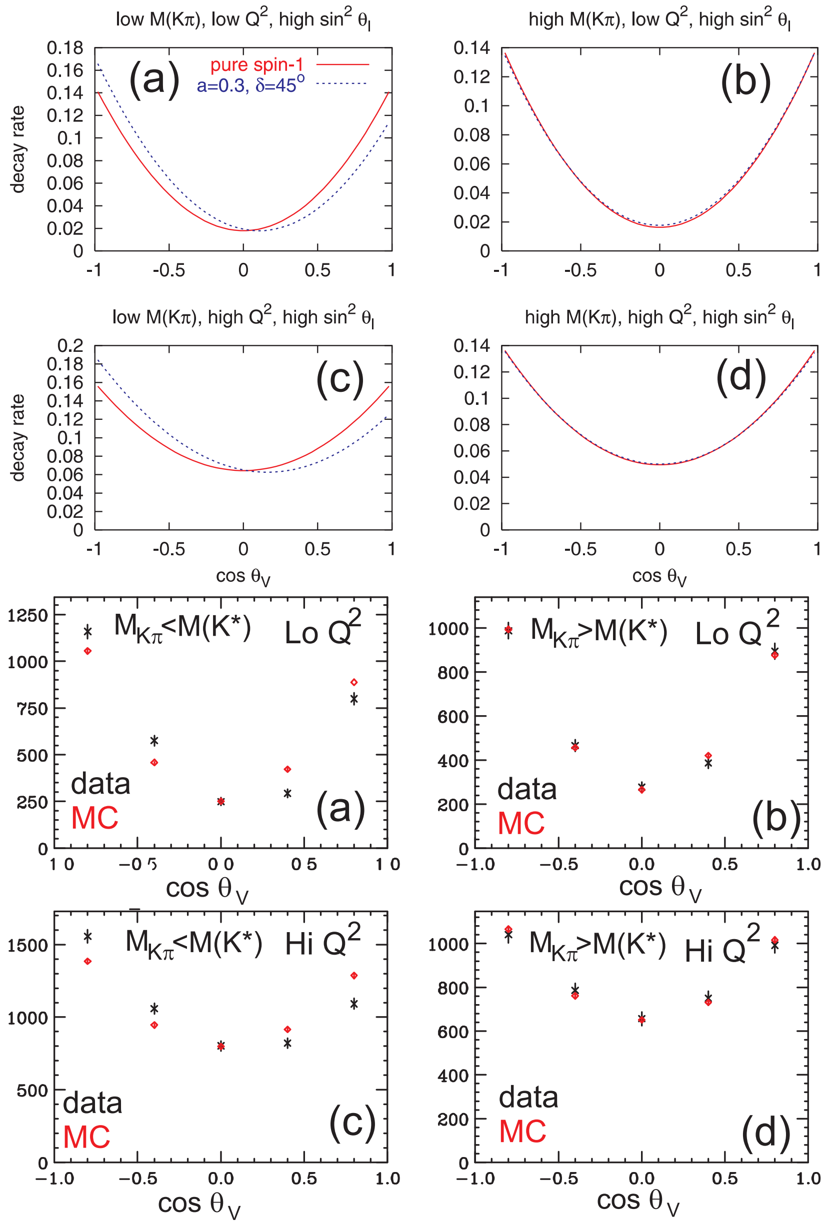

Figure 8.2: Toy model and observed $\cos\theta_V$ projections for $\sin^2\theta_\mu > 0.75$. The upper frame shows the toy model projections for an s-wave amplitude of $0.3\exp(i\pi/4)$. The symmetric red curves are for zero s-wave amplitude. The various projections are (a) $Q^2 < Q^2_{\max}$ and $M_{K^-\pi^+} < M_{K^*}$, (b) $Q^2 < Q^2_{\max}$ and $M_{K^-\pi^+} > M_{K^*}$, (c) $Q^2 > Q^2_{\max}$ and $M_{K^-\pi^+} < M_{K^*}$, (d) $Q^2 > Q^2_{\max}$ and $M_{K^-\pi^+} > M_{K^*}$. The lower frame shows the background subtracted $\cos\theta_V$ distributions observed in the data compared to a MC run (red points) with standard form factors and no interfering s-wave amplitude. The differences between the upper and lower frame represent the effects of smearing and acceptance.

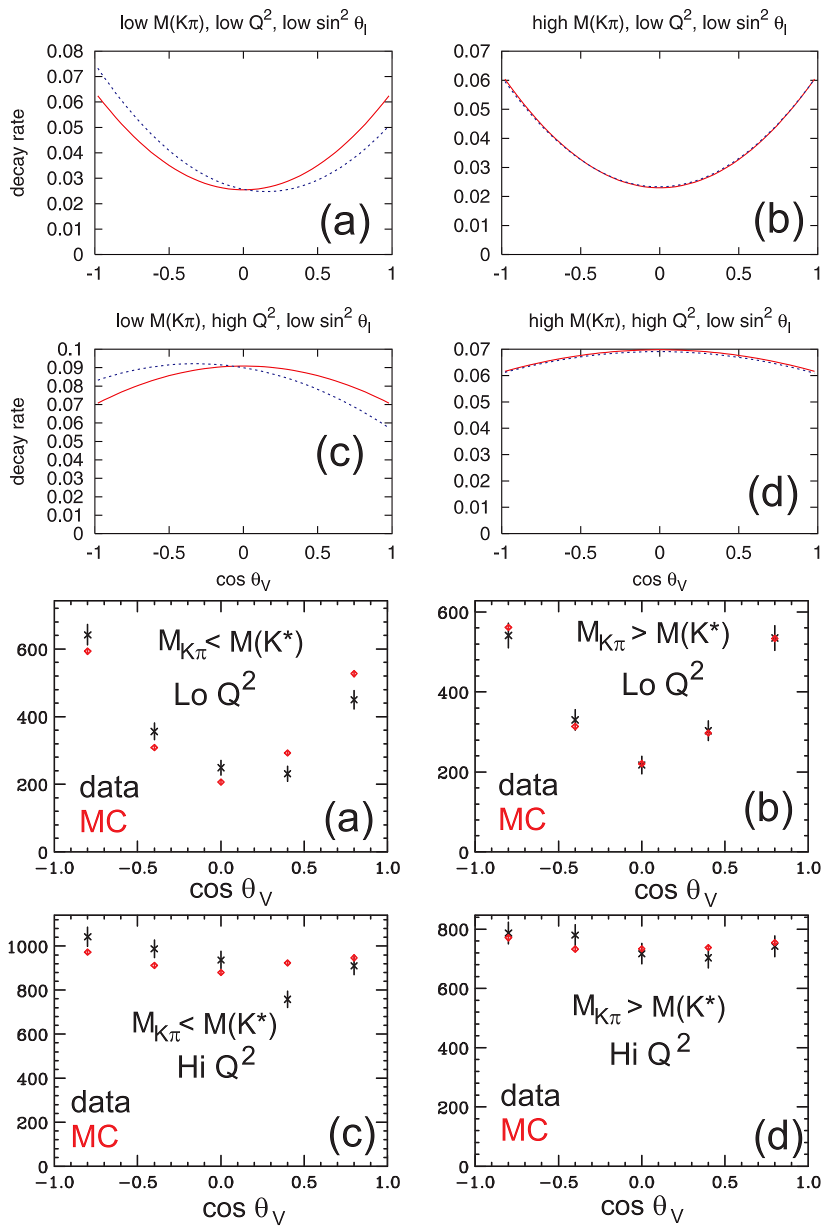

Figure 8.3: Toy model and observed $\cos\theta_V$ projections for $\sin^2\theta_\mu < 0.75$. See Figure 8.2 for description.

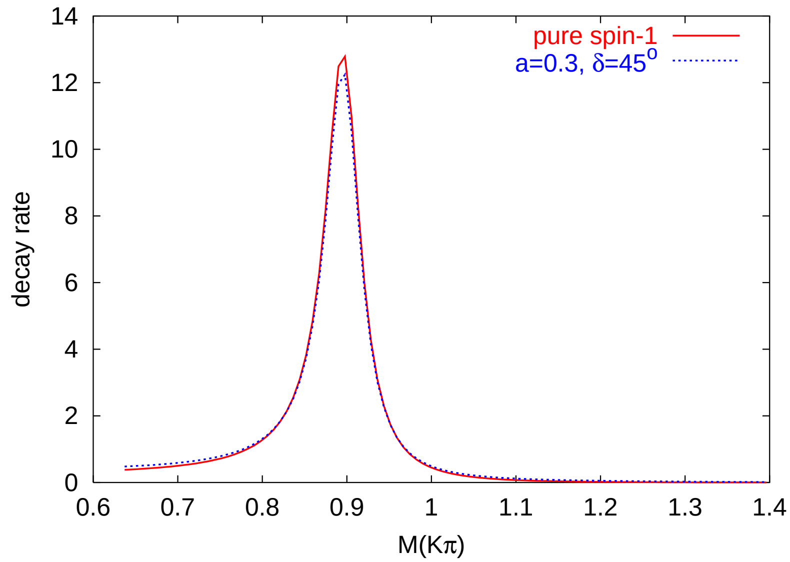

Figure 8.4: Toy model predictions for the $K^-\pi^+$ invariant mass. The blue curve has the interfering amplitude with $0.3\exp(i\pi/4)$ and the red curve has no s-wave interfering amplitude.

8.4 Distortions to the Line Shape

Although the toy model did a good job at reproducing the observed pattern of $\cos\theta_V$, it may have created an unacceptably large distortion to the $K^-\pi^+$ mass distribution. Figure 8.4 addresses this important issue by comparing the line shape with no additional s-wave amplitude and with $0.3\exp(i\pi/4)$. The interfering amplitude causes a very small distortion to the $K^-\pi^+$ mass distribution with this value. There is a noticeable shape asymmetry present in the mass distribution, which is possibly due to the matrix element and/or phase space factors.

The distortions due to the matrix element are very small for accepted events when the events are generated with our standard (no interference) Monte Carlo as shown in Figure 8.5. This Monte Carlo generated the events using a non-relativistic s-wave Monte Carlo with a constant width but with all of the distortions due to phase space and the matrix element. Although the χ^2 for the fit is good, there is a region where a hint of line shape distortion due to matrix element and phase space is evident.

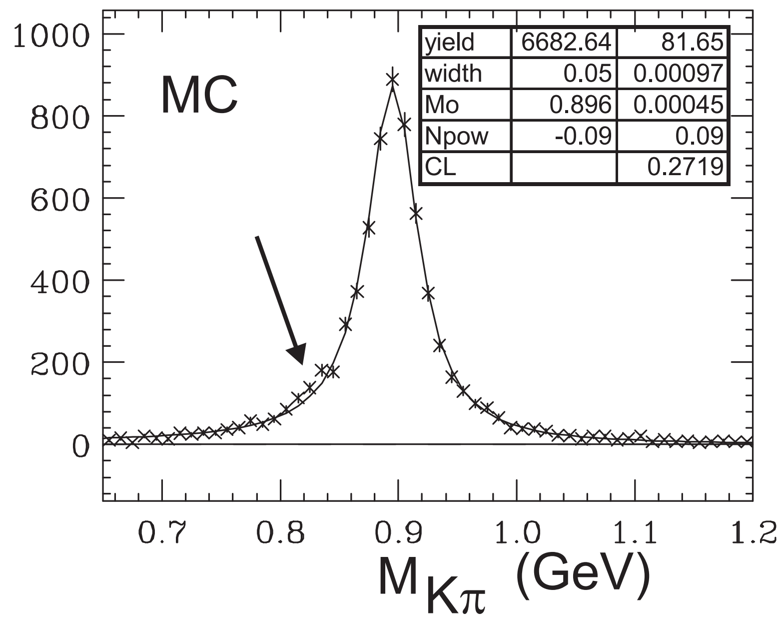

Figure 8.5: The $K^-\pi^+$ invariant mass for s-wave MC events passing analysis cuts. The Monte Carlo events have full phase space and matrix element distortions present in Figure 8.4. The shape is consistent (with a 27% confidence level) with a constant width Breit-Wigner width. This fit, described in Chapter 5, allows the width to vary as $\Gamma = \Gamma_o \left(\frac{P}{P_o}\right)^N$. For this fit, $N = -0.09 \pm 0.09$. The fit to the MC produces essentially the same shape parameters as the fit to the data. The arrow points to the region where one expects matrix element and phase space distortions.

8.5 Could the S-wave Amplitude be Resonant?

We have shown that a constant amplitude with a phase of $\delta = 45^o$ relative to the $\overline{K}^{*0}$ can explain the pattern of the $\cos\theta_{\rm V}$ asymmetries that we observe in the data. Our previous experience in fitting the Dalitz plots describing three-body charm decays shows that one rarely requires non-resonant (constant) amplitudes to fit the data. Even in those cases (such as $D^+ \to K^-\pi^+\pi^+$) where a substantial non-resonant contribution is included [23], the non-resonant contribution may well be mimicked by several broad, poorly understood resonances.

The presence of a constant amplitude with a phase of 45^o is a somewhat unsatisfactory explanation for the $\cos\theta_{\rm V}$ asymmetry from the standpoint of Watson's Theorem. Watson's Theorem contends that in the absence of final state interactions, all amplitudes involved in the decay of a particle should be relatively real. Our (limited) understanding of this very powerful result is that generally the tree level Feynman graphs describing a decay all have the same phase. Additional phase factors require loop contributions. In the context of a weak decay, the only substantial loops are the gluon loops that can be thought of as a rescattering of the final state hadrons. An example might be rescattering of the $\overline{K}^{*0}$ and π^+ in the decay $D^+ \to \overline{K}^{*0}\pi^+ \to K^-\pi^+\pi^+$. Large phase shifts (often around 90^o) are often observed between the amplitudes in three-body hadronic charm decays.[24]

In the semileptonic decay $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$, one cannot get the same sort of final state interaction observed in $D^+ \to K^-\pi^+\pi^+$ since the $\overline{K}^{*0}$ is produced against leptons and thus has no hadron partner to rescatter from. In principle, an s-wave $D^+ \to K^-\pi^+\mu^+\nu$ could acquire a phase relative to $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ by having the K^- rescatter from the s-wave π^+. An alternative way of getting a phase shift would be a resonant s-wave $D^+ \to K^-\pi^+\mu^+\nu$ contribution with a real phase (0^o or 180^o) relative to $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$. Figure 8.6 explores this possibility by comparing the linear $\cos\theta_{\rm V}$ coefficient expected from the interference of a $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ contribution with a resonant s-wave $D^+ \to K^-\pi^+\mu^+\nu$ contribution for two hypothetical resonances to

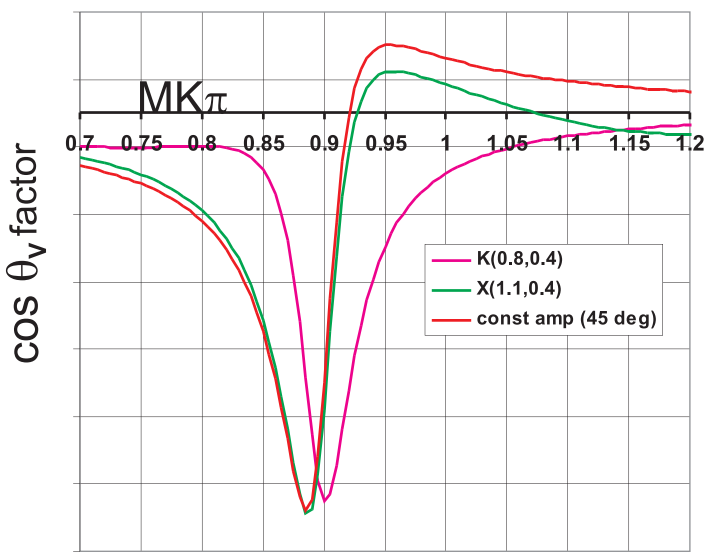

Figure 8.6: The interference term in the toy model. We plot the mass dependence of the linear $\cos\theta_V$ term as a function of mass for a constant amplitude and two hypothetical resonances that interfere with a 180^0 phase with respect to the $\overline{K}^{*0}$. We indicate the mass and width in the legend.

that expected from a constant amplitude with a 45^o phase relative to $D^+ \to \overline{K}^{*0} \mu^+ \nu_\mu$. Equation 8.5 gives the linear $\cos\theta_{\rm V}$ term expected from interference of the $D^+ \to \overline{K}^{*0} \mu^+ \nu_\mu$ with an s-wave resonance in a form reminiscent of Equation 8.4. Here, M_o and Γ are the mass and width of the $\overline{K}^{*0}$, while M_r and Γ_r are the mass and width of the hypothetical s-wave $K^-\pi^+$ resonance.

$$8\sin^2\theta_\mu \; \Gamma_0 \; \frac{A\;\sqrt{M_o \Gamma}}{(M^2 - M_o^2)^2 + (\Gamma \; M_o)^2} \; \{(M^2 - M_o^2) \; \cos\delta - \Gamma \; M_o \; \sin\delta\} \times \cos\theta_{\rm V}$$

$$\cos\delta = \frac{(M^2 - M_r^2)}{(M^2 - M_r^2)^2 + (\Gamma_r \; M_r)^2} \quad , \quad \sin\delta = -\frac{(\Gamma_r \; M_r)}{(M^2 - M_r^2)^2 + (\Gamma_r \; M_r)^2} \tag{8.5}$$

Figure 8.6 plots the $\cos\theta_{\rm V}$ coefficient as a function of $K\pi$ mass for three different models: our uniform amplitude, a Breit-Wigner resonance above the K^{*0} pole, and a resonance below the K^{*0}. While the hypothetical high-mass resonance appears to be a successful match to our observed trend, the low-mass resonance does not.

This conclusion is particularly interesting to some members of the FOCUS collaboration in light of previous evidence for a hypothetical $K^-\pi^+$ resonance called the $\kappa(900)$ with a width of about $\Gamma_r \approx 400$ MeV. This resonance has been proposed in part since its presence significantly improves the fits to the $D^+ \to K^-\pi^+\pi^+$ according to the analysis of the E791 collaboration.[25]. According to Figure 8.6, interference with the $\kappa(900)$ would create a $\cos\theta_{\rm V}$ asymmetry primarily above the K^{*0} pole, whereas we observe the dominant asymmetry below the K^{*0} pole.

8.6 Summary and Conclusions

We believe we have presented a plausible explanation for the the asymmetry in $\cos\theta_{\rm V}$ that we observed in data for $D^+ \to \overline{K}^{*0} \mu^+ \nu_\mu$ decay that had frustrated our attempts to get reasonable quality form factor fits. We observed an asymmetry favoring negative $\cos\theta_{\rm V}$ for events with $K^-\pi^+$ masses below the K^{*0}pole and essentially no asymmetry for events above the K^{*0}pole. As a result, we saw a slight shift in form factors when we split the sample relative to the K^{*0}mass. Although this shift was small, the observed asymmetry could have been an indicator of other, undiscovered effects that bias the result.

We did an extensive search for possible non-physics explanations for the $\cos\theta_{\rm V}$ asymmetry such as backgrounds, or problems in correcting for the acceptance or resolution smearing. We found that the form factor fit was very stable over a wide range of cut sets with both significantly different acceptances and vastly different RS to WS ratios. We showed using the golden mode data with a deliberately blanked track to represent the neutrino, that the neutrino closure technique was very well modeled by our Monte Carlo. Also, we were able to get consistent branching fractions of $D^+ \to \overline{K}^{*0}\mu^+\nu$ relative to $D^+ \to K^-\pi^+\pi^+$ over an exhaustive set of cut sets with vastly different efficiencies and background rejections. Finally, we eliminated specific (incoherent) potential semileptonic decays involving decays into higher mass K^*.

The dramatic dependence of the $\cos\theta_{\rm V}$ asymmetry above and below the pole strongly suggested to us an interference effect was at work since it is well known that the real part of a Breit-Wigner amplitude reverses sign when one passes through the pole. By having the $\overline{K}^{*0}$ contribution that carries a $\cos\theta_{\rm V}$ factor in its amplitude interfere with an s-wave amplitude, one can get a linear $\cos\theta_{\rm V}$ piece in the interference term creating a noticeable $\cos\theta_{\rm V}$. We don't believe this possibility has been considered in the literature, since the observed $K^-\pi^+$ spectrum in previous experiments (and ours) is such a perfect match to a $\overline{K}^{*0}$ Breit-Wigner.

To progress further, we extended the heuristic treatment of Chapter 2 to include a constant or resonant s-wave amplitude contribution that couples to the virtual W^+ in its zero helicity state. This allowed us to form a fully explicit model (except for acoplanarity averaging and lepton mass terms). In this model, the size of the $\cos\theta_{\rm V}$ is proportional to the s-wave amplitude, and its mass dependence depends on the phase. The explicit model was used to compare various $\cos\theta_{\rm V}$ projections that could be compared to the pattern observed in the data. The observed asymmetry pattern was well reproduced by our s-wave interference model with parameters that created a negligible distortion to the $K^-\pi^+$ line shape.

If our interpretation is correct, it seriously calls into question previous attempts to test quark models and lattice gauge calculations using charm semileptonic decay, since the data has an asymmetry that cannot be accommodated in the decay form used to fit for the form factors. We believe that our experiment is the first experiment to accumulate enough statistics and signal cleanliness to observe this asymmetry.

8.7 Comparison with Other Experiments

FOCUS is not the first experiment to measure the form factor ratios for the decay $D^+ \to \overline{K}^{*0}\ell^+\nu_\ell$. Is there evidence in these other experiments of an asymmetry in $\cos\theta_\mathrm{V}$, or are we the only ones with enough data to observe this effect clearly? This is important to establish. We claim that our asymmetry is not an artifact of some improperly-modeled aspect of the FOCUS spectrometer, but rather is a physical process that should affect every analysis of this decay.

We will only look at the three most recent measurements, since earlier ones will certainly not have enough data to probe the asymmetry.

8.7.1 E687

E687 was the predecessor experiment to FOCUS, and in a sense this thesis is a logical extension of the earlier work done by the E687 group. They obtained a $D^+ \to \overline{K}^{*0}\ell^+\nu_\ell$ signal of about 875 events [7], with a higher signal-to-noise ratio than our FOCUS sample (with our baseline cuts). This number is the total yield of the best-fit Breit-Wigner. When we use the same fitting function to measure the yield in FOCUS data passing our baseline cuts, we obtain a yield of 32159 events.

The most natural question to address is whether or not the FOCUS result is consistent with E687. They measured $r_v = 1.74 \pm .27 \pm .28$ and $r_2 = 0.78 \pm .18 \pm .10$. Our results for the baseline cuts are consistent with these values. Since E687 and FOCUS are alike in so many respects, we can go further by duplicating all of E687's cuts and bin the data into the same 18 bins they use. We do not use the same Čerenkov identification algorithm, but we can mimic the general flavor of the Čerenkov cuts used.

Doing this, we measure $r_v = 1.734 \pm .071$, $r_2 = 1.171 \pm .047$, and our best-fit $\chi^2/\mathrm{DOF} = 17.9$ for 15 degrees of freedom. Here we match the E687 result for r_v, but our r_2 is 1.8σ too high. Looking at the K^{*0} yield fits in each bin, it is obvious that egregious errors are made in certain bins that are overwhelmed with contamination from $D^{*+} \to D^0\pi^+ \to (K^-\mu^+\nu_\mu)\,\pi^+$. E687 did not use any cuts which would be effective at reducing this reflection, and the contamination is obvious in their mass plots. The behavior of the mass fits will depend on a lot of factors which cannot be duplicated by FOCUS, including event statistics and the range in $K\pi$ mass over which the fit is performed. If we apply an additional cut on the $D^{*+}-D^0$ mass difference, our result changes dramatically, and we measure $r_v = 2.113 \pm .083$ and $r_2 = 0.780 \pm .051$, with $\chi^2/\mathrm{DOF} = 9.8$. Now we are duplicating E687's measurement of r_2, and are only 0.94σ higher with r_v. It is likely that E687 paid more attention to their fits than we have here, and managed to avoid most, but not all, of the damage caused by the rampant D^* reflection. They did not apparently understand the mechanism of this contamination, or realize the way it falls off so quickly in $K\pi$ mass (see Figure 7.12). This error would most likely have caused them to overestimate the background under the K^{*0} peak, and so underestimate the yield. Since the D^* contamination occurs at large $\cos\theta_{\mathrm{V}}$, this effect would serve to exaggerate the asymmetry.

In Figure 8.7, we show the difference between the E687 data and expectation, as a function of $\cos\theta_{\mathrm{V}}$. The values are taken from the best-fit and observed bin populations in their 18-bin form factor fit. By dividing the slope of the best-fit line by the total signal, we obtain a kind of fractional asymmetric contribution of $.053 \pm .017$. This degree of asymmetry is greater than what we see in FOCUS, which is about .01. The increase in the asymmetry found by E687 is most likely due to the mis-handled D^* reflections.

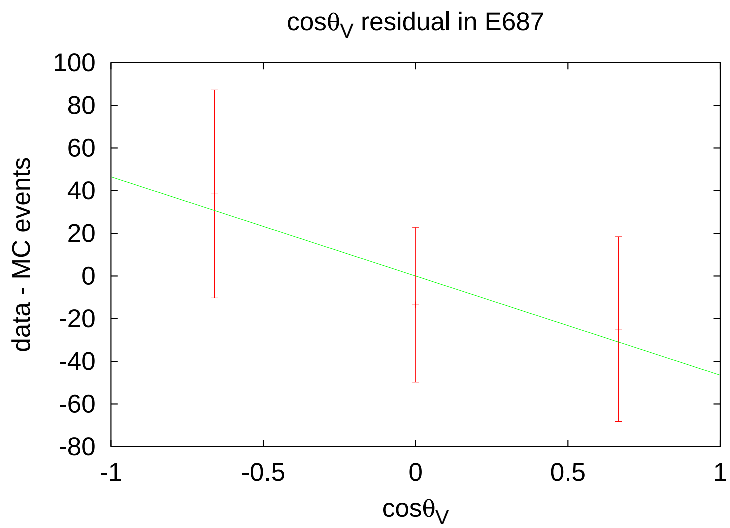

Figure 8.7: E687 $\cos\theta_\mathrm{V}$residual asymmetry. The line is a best-fit straight line, having a slope of -46 ± 15. The significance of this asymmetry is most likely due to contamination from $D^{*+} \to D^0\pi^+ \to (K^-\mu^+\nu_\mu)\,\pi^+$.

8.7.2 E791

E791 is a fixed-target experiment that collected data during the 1990-1991 fixed target run at Fermilab. E791 was a charm hadroproduction experiment, using a 500 GeV/c π^- beam. They used a tighter $K\pi$ mass cut than we do, counting (with wrong-sign subtraction) those events where $0.85 < M_{K\pi} < 0.94\mathrm{GeV}/c^2$. They observed 3034 RS-WS events with a RS/WS ratio of 6.1 [10]. This is a *combined* yield of semielectronic and semimuonic decays. To compare this to the FOCUS muon sample, our baseline cuts gives us 22075 events within this band and a RS/WS ratio of 6.8.

E791 chose to use an unbinned maximum-likelihood technique for their analysis, but as a means of confirmation they also performed a binned fit. This fit gave them a reasonable χ^2.

Since their article includes figures showing the predicted and anticipated $\cos\theta_\mathrm{V}$ distribution (Figure 2 in [10]), we can search for an asymmetry and directly compare this to FOCUS data. Figure 8.8 shows the E791 data and Monte Carlo results, carefully reconstructed from the original figure. The two plots, split in Q^2, are first

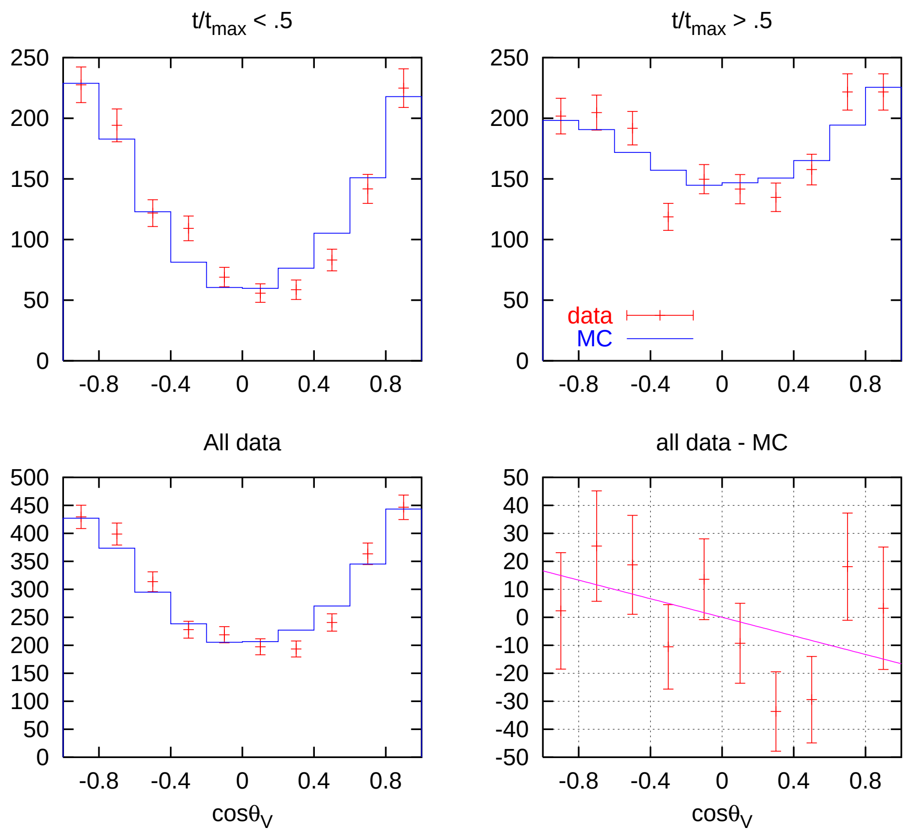

Figure 8.8: E791 $\cos\theta_{\mathrm{V}}$ distributions. Red points are data, and the blue lines are the prediction. The upper plots are reconstructed by visual inspection of the original figure in reference [10]. A straight line is fit to the residual. The slope of this line, divided by the total yield, gives a fractional asymmetric contribution of $0.0055 \pm .0041$. This is a 1.3σ effect, in the same direction as the FOCUS data.

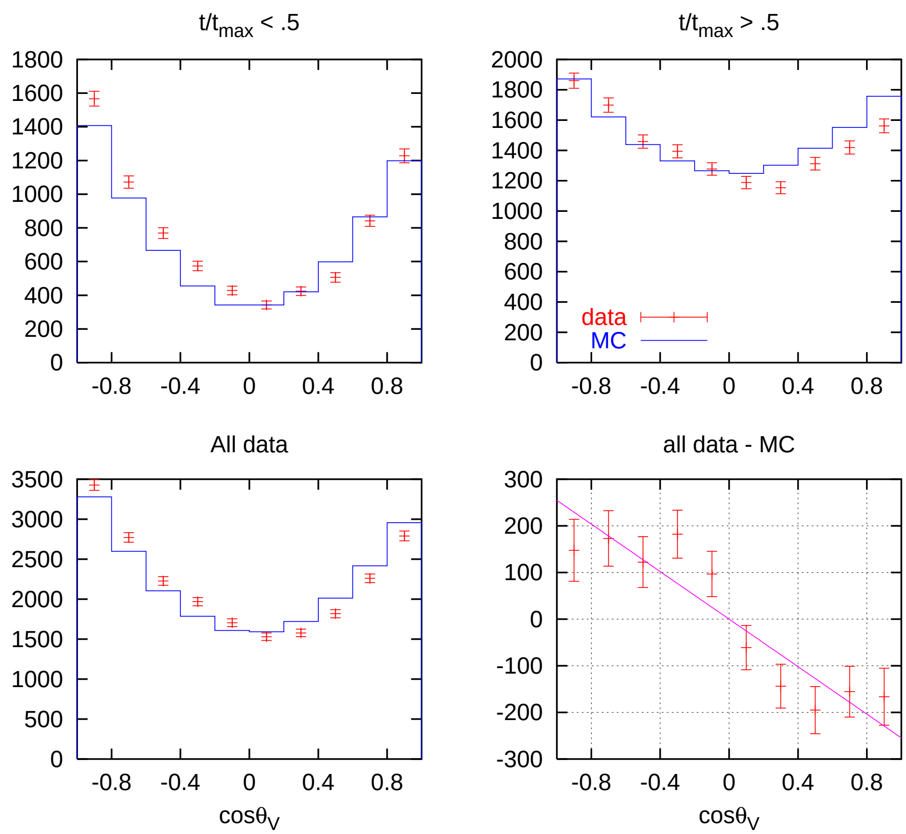

Figure 8.9: FOCUS $\cos\theta_{\rm V}$ distributions. These distributions are most suitable for comparison with the E791 result (Figure 8.8), since they include the same $0.85 < M_{K\pi} < 0.94 \mathrm{GeV}/c^2$ mass cut that E791 uses. A straight line is fit to the residual. The slope of this line, divided by the total yield, gives a fractional asymmetric contribution of $0.012 \pm .002$.

added together. Then their Monte Carlo prediction is subtracted from the data, and a straight line is fit to the difference. By dividing the slope of this line by the total signal, we obtain a fractional asymmetric contribution. Figure 8.9 shows the equivalent analysis in FOCUS data. The FOCUS result shows the asymmetric component in the same direction as E791, but it is about 1.4σ larger. Curiously, in both data sets the central region of $\cos\theta_{\mathrm{V}}$ shows the greatest amount of slope. This may be an artifact of backgrounds and/or the narrow mass cut, and the effect is not necessarily significant in E791.

8.7.3 Beatrice

Prior to now, the most recently published measurement of the form factor ratios of the decay $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ was performed by the BEATRICE collaboration in 1998 [13]. They use an unbinned maximum likelihood technique, but to asses the goodness of the fit, they report on the confidence levels of the projections in the four kinematic variables. They use a sample of 763 events in the analysis. Their mass window, $0.846 < M_{K\pi} < 0.946\mathrm{GeV}/c^2$, is close to the one used by E791.

Figure 8.10 is a reproduction of the $\cos\theta_{\mathrm{V}}$ distribution measured by BEATRICE. There is little evidence for an asymmetry here, and their prediction is well-matched by the data. FOCUS observes about 29 times more events, and this $K\pi$ mass window is small. In the FOCUS data within that mass window, we see the asymmetry at about the 6σ level. If this scales with $\sqrt{N}$, BEATRICE should observe it at about 1σ.

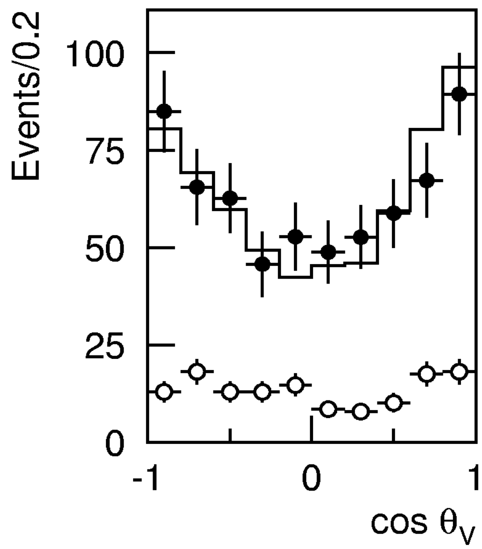

Figure 8.10: BEATRICE $\cos\theta_{\mathrm{V}}$ distribution. The solid points are data, the lines are from Monte Carlo, and the open points are a background estimate. This distribution shows some faint signs of an asymmetry in the outer bins. Their data matches the prediction to confidence level of not less than 33%.

Chapter 9

Summary and Conclusions

This thesis is concerned the charm semileptonic decay process $D^+ \to K^-\pi^+\mu^+\nu$. We study this decay using the copious charm sample collected by the FOCUS fixed-target photoproduction experiment at Fermilab in 1996-1997. This is probably the most experimentally accessible charm semileptonic decay for ours and other experiments owing to its relatively large branching ratio and highly favorable decay topology. Studies of the $D^+ \to K^-\pi^+\mu^+\nu$ decay are of considerable phenomenological interest and have a rich experimental history. Charm semileptonic decays provide incisive tests of non-perturbative QCD.

Previous experiments revealed a rather remarkable simplification in the four body process $D^+ \to K^-\pi^+\mu^+\nu$. In reality, to very good approximation, this four body process is actually a quasi-three body process: $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$, with the $\overline{K}^{*0}$ very rapidly decaying via $\overline{K}^{*0} \to K^-\pi^+$. Prior to this work, there was essentially no experimental evidence for anything other than $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ contributing to $D^+ \to K^-\pi^+\mu^+\nu$. For example, our previous experiment, E687, put a 90% upper limit on a non-resonant contribution to $D^+ \to K^-\pi^+\mu^+\nu$ of less than 12%.[7]

The observation that $D^+ \to K^-\pi^+\mu^+\nu$ is actually $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ leads to considerable theoretical simplification. The decay can be visualized as $D^+ \to W^+\overline{K}^{*0} \to \overline{K}^{*0}\mu^+\nu$, where the W^+ is a highly virtual version of the ≈ 100 GeV mass particle responsible for mediating the charged current weak interaction. The underlying weak interactions that describe the quark couplings to the final state leptons are well understood, and can be reliably computed using well-tested and straightforward perturbation theory. However, non-perturbative techniques must be used to compute how the final state quarks, created from the decay of the charm quark, rearrange themselves to form a $\overline{K}^{*0}$, for each of the three polarization states of the virtual (spin 1) W^+, by exchanging gluons. The coupling of the final state to each of the three polarization states of the W^+ are described by three (helicity basis) form factors that are functions of t or the squared mass of virtual W^+. The helicity basis form factors are linear combinations of a vector and two axial form factors. In particular, the ratios of the three form factors describe the t-dependence of the three decay angles. These deecay angles describe the 4 body final state: the angle of the kaon in the $\overline{K}^{*0}$ reference frame ($\cos\theta_{\mathrm{V}}$), the angle of the μ in the virtual W^+ frame ($\cos\theta_\ell$), and the acoplanarity angle between the $\overline{K}^{*0}$ and W^+ decay planes (χ).

Our original intention was to fit the kinematic dependence of the $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ decay intensity to measure the axial and vector form factors controlling this decay. Our measurement could then be compared to the predictions of Lattice Gauge, QCD Sum Rules, or Relativistic Quark Models. In particular, the Lattice Gauge calculations, although computationally intensive, use a relatively minimal number of simplifying assumptions and should thus be reliable. Similar Lattice Gauge calculations are used to unravel the QCD corrections used in understanding some very topical issues such as direct CP violation in the B and kaon system.

We then intended to compare the $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ form factor ratios to those measured for our sample of the related (but much rarer) decay $D_s^+ \to \phi\ \mu^+\nu_\mu$. A principle motivation for this work would have been to confirm or refute the claim of several standard deviation differences between the two form factor sets made by some previous experiments. A large difference between the $D_s^+ \to \phi\ \mu^+\nu_\mu$ and $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ form factor is unexpected from the perspective of Lattice Gauge and quark model theory. Since we have reconstructed about 1500 $D_s^+ \to \phi\ \mu^+\nu_\mu$ events — nearly a factor of 500 more than any previously published sample — our data would have easily settled this interesting issue.

Although the $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ is one of the most accessible and easily studied charm semileptonic decays, the presence of a final state neutrino makes it considerably more difficult to study in our experiment than the fully reconstructable hadronic decays such as $D^+ \to K^-\pi^+\pi^+$. Here are two reasons:

1. Because we do not reconstruct the complete $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ final state as a single narrow mass peak, we may be contaminated by backgrounds from both known and previously unobserved charm decays as well as non-charm, photo-hadronic backgrounds.

2. Because we don't measure the final state neutrino, we must infer its momentum vector using $P_\perp$ balance about the D^+ line-of-flight (or the displacement of the secondary vertex to the primary vertex). Even with perfect measurement of this direction, we can still only determine the neutrino momentum to within a two-fold ambiguity. In photoproduction, the primary vertex consists of at most a few tracks, and it can be easily pulled by mistakenly including tracks from the recoil charm particle. As a result, the momentum transfer (or virtual W^+ mass) and all three decay angles are very poorly measured. For this reason, complicated techniques must be used to fit the decay intensity, and accurate fits require correct knowledge of both our acceptance and kinematic resolution. The properties of the photoproduced primary vertex cannot be computed from first principles. Rather our resolution is predicted by a Pythia-based Monte Carlo that depends on an specific string fragmentation model.

In addition to these experimental challenges, the expression for the complete decay rate is complicated. Each of the three angular variables is defined in its unique rest frame. The sign conventions between charm and beauty are different, and the conventions in going from the D^+ to D^- were unclear in the early literature, and these have confounded many of the workers in the past.

In spite of these challenges, we set out to measure the $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ form factor ratios using a binned version of the Los Vegas reweighting technique described in Chapter 6. The use of this new fitting technique allowed us to obtain both high precision measurements of the form factors and provided a χ^2 statistic used to test the goodness of the overall fit. As is traditional, we measured the ratios of axial and vector form factors taken at $t = 0$: $r_v \equiv V(0)/A_1(0)$ and $r_2 \equiv A_2(0)/A_1(0)$.

As detailed in Chapter 7, we tried many fits with significantly different analysis cuts, different binning, and different sample splits. In all of our initial work, we obtained reasonably consistent form factor ratios with typical fits and fit ranges of:

$$\text{Baseline fit } r_v = 1.694 \pm 0.051 \quad r_v \text{ range} : 1.60 \rightarrow 1.75$$

$$\text{Baseline fit } r_2 = 0.754 \pm 0.034 \quad r_2 \text{ range} : 0.72 \rightarrow 0.85$$

Figure 9.1 compares our values to the r_v and r_2 values measured in previous experiments. Using the spread of fit results using different cut sets and binning, we conservatively estimate systematic errors on our form factors of $\sigma_{\text{sys}}(r_v) = 0.075$ and $\sigma_{\text{sys}}(r_2) = 0.065$. These systematic errors are about a factor of two larger than our statistical errors. In Figure 9.1, the statistical and systematic errors are added in quadrature.

The overall consistency of all D^+ form factor data was excellent. The largest discrepancy is the 1.66 σ discrepancy between our r_v measurements and that of the E791 collaboration. The most disturbing feature of our data is that essentially all of our binned form factor fits have very poor χ^2 — often with confidence levels below 1×10^{-25}. For this reason, we are not sure how seriously one should take the form factor comparison. After an extensive investigation, we believe that the poor confidence level of our fit reflects new physics rather that problems understanding backgrounds, resolutions, or acceptance. We will try to review our case for new physics in the rest of this summary.

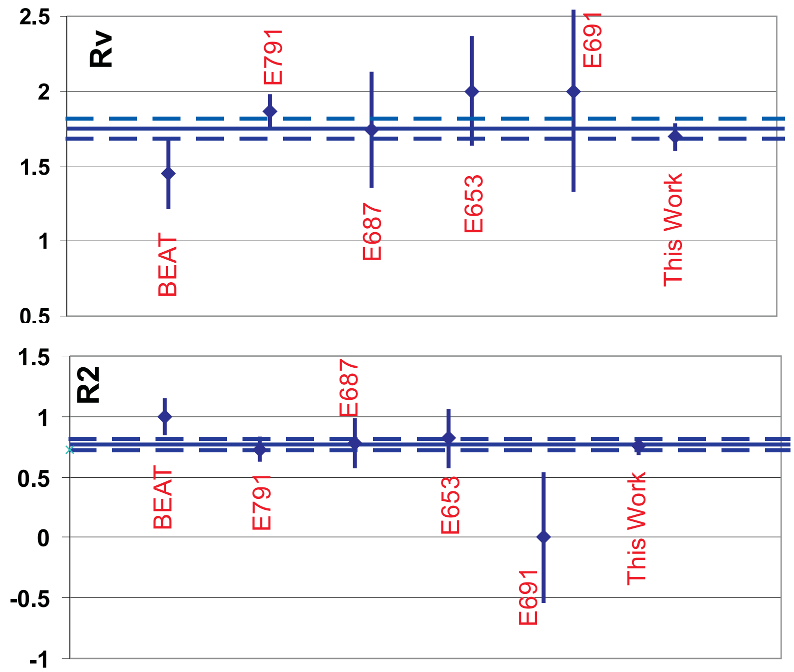

Figure 9.1: Experimental data on r_v and r_2. Horizontal lines indicate the average values of the form factor ratios including this measurement along with their errors. The weighted average values are $r_v = 1.754 \pm 0.064$ and $r_2 = 0.773 \pm 0.052$. The confidence level that all six r_v measurements are consistent is 58%. The confidence level that all six r_2 measurements are consistent is 48%.

9.1 The Usual Suspects

The overall complexity of this problem along with all of the problems associated with the missing neutrino enumerated above could, in principle, be responsible for our inability to get a high quality fit. Since our $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ data set is nearly an order of magnitude larger than that of our closest competitor (E791), we should be much more sensitive to any systematic problems in the fit, since their effect on the fit-χ^2 would be multiplied by a factor of 10 as well.

We checked the validity of our fitting technique by applying it to Monte Carlo samples generated with known form factors. In this study, discussed in Chapter 6, an independent Monte Carlo sample was used to correct the simulated "data". We observed essentially no bias in r_v or r_2.

We made a number of checks to confirm our understanding of the decay width expression and conventions. Our large $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ sample allowed us to make detailed comparisons with the model through the use of the "conditional" projections shown in Figure 7.5 of Chapter 7. For example, our data tracked the expected sign change in the concavity of the $\cos\theta_{\mathrm{V}}$ projection for low versus high t events. The change in the relative frequency components of the acoplanarity distribution in the three different regions of $\cos\theta_{\mathrm{V}}$ was also well tracked by our data. Finally, the form factors measured separately for the D^+ were consistent with those found for the D^-.

In Chapter 4, Section 4.7, we tested our kinematic resolution by "blanking" one of the pions in the fully reconstructible final state $D^0 \to K^-\pi^+\pi^+\pi^-$ and reconstructing the "missing" pion using the same $P_\perp$ balance technique used to reconstruct the missing neutrino in $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ decay. We obtained resolution functions for the mimic decays that were an excellent match to those predicted by our Monte Carlo.

As a way of assessing the reliability of our acceptance, cut efficiency, and non-charm or misidentification backgrounds, we measured the branching ratio of $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ relative to $D^+ \to K^-\pi^+\pi^+$ as discussed in Chapter 5. We were able to get

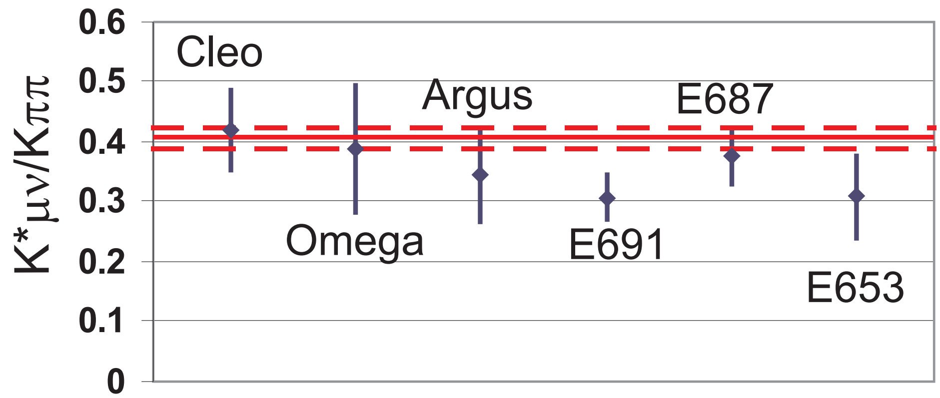

Figure 9.2: Our measurements of $\Gamma(D^+ \to \bar{K}^*\mu^+\nu)/\Gamma(D^+ \to K^-\pi^+\pi^+)$ compared to previously published data from the six indicated experiments. Our value including our systematic uncertainty is given by the solid and dashed horizontal lines. Our (fractional) uncertainty including systematic errors is about three times smaller than that obtained in previous experiments.

a very stable value of:

$$\mathrm{BR} = \frac{\Gamma(D^+ \to \bar{K}^*\mu^+\nu \to (K^-\pi^+)\mu^+\nu)}{\Gamma(D^+ \to K^-\pi^+\pi^+)} = 0.406 \pm 0.006\ (\mathrm{stat}) \pm\ 0.016\ (\mathrm{sys})$$

using a wide range of analysis cuts that sampled a large range of signal-to-noise. We would expect considerable variation in the branching ratio if our simulation did not properly predict the efficiency of our analysis cuts, or if more stringent analysis cuts eliminated non-charm or particle misidentification reflection backgrounds present in our baseline sample. Figure 9.2 compares our relative branching fraction measurement to previously published numbers. Our results are consistent with but more precise than previous measurements.

Other backgrounds to $D^+ \to K^-\pi^+\mu^+\nu$ include charm decay into the same visible $K^-\pi^+\mu^+$ final state. These include $D^{*+} \to D^0\pi^+ \to (K^-\mu^+\nu_\mu)\,\pi^+$ and $D^+ \to \overline{K}_1^0(1270)\mu^+\nu_\mu$ with the $\overline{K}_1^0(1270)$ decaying via $K^-\pi^+\pi^0$. We found that without special cuts, the $D^{*+} \to D^0\pi^+ \to (K^-\mu^+\nu_\mu)\,\pi^+$ background is very serious and can significantly distort the expected decay distribution for $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$. Fortunately, it is easy to eliminate this background through a cut on the $D^{*+} - D^0$ mass difference using our neutrino closure technique.

The $D^+ \to \overline{K}_1^0(1270)\mu^+\nu_\mu$ decay could also significantly affect the decay distribution. This is a particularly interesting background since it could mimic the $\cos\theta_{\rm V}$ anomaly. In Section 7.7, we show that this background is essentially ignorable by making an explicit search for the related decay $D^0 \to K_1^-(1270)\mu^+\nu_\mu \to (K^-\pi^+\pi^-)\mu^+\nu_\mu$ and using an isospin argument to severely limit backgrounds due to $D^+ \to \overline{K}_1^0(1270)\mu^+\nu_\mu$. We thus eliminate it as a possible explanation for the $\cos\theta_{\rm V}$ anomaly.

9.2 New Physics (?)

In the process of investigating the cause of our poor quality form factor fits, we made a large number of two-fold "split-sample" studies where we would compare the form factors obtained in one piece of the data to those obtained in the compliment. These included splits of particle versus anti-particle, low run numbers versus high run numbers, low momentum transfer versus high momentum transfer, etc. We monitored the consistency of the form factors between the two samples, as well as the χ^2 of the fit in either sample. The most interesting split-sample test involved a split of events with $K^-\pi^+$ masses below and above the K^{*0}pole mass. The χ^2 of the fit of events with $M(K^-\pi^+) > M(K^{*0})$ was dramatically smaller than the fit of events with $M(K^-\pi^+) > M(K^{*0})$. Comparisons of the distributions of kinematic quantities of events above and below the K^{*0} pole showed that the events below the pole had a very pronounced asymmetry in $\cos\theta_{\rm V}$ that was not matched by our Monte Carlo; while the observed $\cos\theta_{\rm V}$ asymmetry was nearly a perfect match to the Monte Carlo in the $M(K^-\pi^+) > M(K^{*0})$ events. The expected $D^+ \to \overline{K}^{*0}\mu^+\nu_\mu$ intensity is even in $\cos\theta_{\rm V}$, so any asymmetry would only come through efficiency variation.

In Chapter 8, we propose a model that explains the $\cos\theta_{\rm V}$ asymmetry pattern in the data. We are proposing the existence of a previously unobserved s-wave $K^-\pi^+$

contribution that coherently interferes with a dominant p-wave $\overline{K}^{*0}$ contribution to $D^+ \to K^-\pi^+\mu^+\nu$. Simplifying a bit, the $\cos\theta_{\rm V}$ asymmetry comes about through the cross term between the $\overline{K}^{*0}$ Breit-Wigner (that carries a $\cos\theta_{\rm V}$ factor since the $\overline{K}^{*0}$ is spin 1) interfering with a slowly varying, zero spin amplitude (that carries no $\cos\theta_{\rm V}$ factor).

$$\begin{aligned}&\left|\frac{\cos\theta_{\rm V}\ \ \sqrt{M_o\Gamma}}{(M^2-M_o^2)+i\ \Gamma\ M_o}+Ae^{i\delta}\right|^2 = \\ &...+2\frac{A\ \sqrt{M_o\Gamma}}{(M^2-M_o^2)^2+(\Gamma\ M_o)^2}\ \{(M^2-M_o^2)\ \cos\delta-\Gamma\ M_o\ \sin\delta\}\times\cos\theta_{\rm V}+...\end{aligned} \tag{9.1}$$

We found that a phase of $\delta \approx \pi/4$ produces an $< \cos\theta_{\rm V} >$ versus $M(K^-\pi^+)$ curve that is in excellent agreement with the dependence that we observed in the data. Furthermore, the modulus of the s-wave amplitude that best matched the level of the asymmetry ($|A| \approx 0.3$) created a very minor distortion to the $K^-\pi^+$ line shape when integrated over all $\cos\theta_{\rm V}$, since its asymmetry contributions are first order in $|A|$ while line shape modulation would be second order. We believe this is the reason that the additional s-wave amplitude has eluded previous investigators.

In Chapter 8, we work out an explicit $\cos\theta_\ell$ and t- dependence for the $\cos\theta_{\rm V}$ term that also appear to match the pattern observed in the data. Our model is not unique. A crucial ingredient is that there is a $\approx 45^o$ phase shift between the s-wave amplitude and the Breit-Wigner at the K^{*0} mass. This can be done either through a constant amplitude or alternatively through a possible broad s-wave Breit-Wigner resonance with a pole and width chosen to produce an $\approx 45^o$ phase shift at the K^{*0} pole. For example, we show that an s-wave resonance with a mass of 1.1 GeV/c^2 and a width of $\Gamma = 0.4$ GeV/c^2 inserted with a 180^o phase relative to the $\overline{K}^{*0}$ will produce the same pattern that is observed in the data. Interestingly enough, the $\kappa(800)$ with a width of $\Gamma = 0.4$ GeV/c^2 that has been proposed to explain problems in fitting the $D^+ \to K^-\pi^+\pi^+$, produces a phase of 63^0 at the K^{*0} pole but evolves in mass in a way inconsistent with our observed asymmetry pattern. Reference [9] discusses calculations of the partial widths of possible additional resonances that might contribute to the $D^+ \to K^-\pi^+\mu^+\nu$ final state.

It might be of interest to redo our form factor fit including the effects of the new s-wave amplitude. Unfortunately, we don't have a complete model of the decay intensity, since we worked out the expression using a simple symmetry technique that averaged over the acoplanarity angle between the two decay planes and neglected the "mass" terms proportional to the squared charged lepton mass.

We have reviewed previous $D^+ \to \overline{K}^{*0}\ell^+\nu_\ell$ measurements with an eye to looking for hints of the $\cos\theta_{\rm V}$ anomaly. The E687 data is consistent with a linear $\cos\theta_{\rm V}$ term with the same sign as ours at about the 3 σ level. Their overall fit quality was also rather poor with a χ^2 per degree of freedom of about 2.48 for 16 degrees of freedom. A $\cos\theta_{\rm V}$ asymmetry with the same sign as ours of about 1.3 σ significance appears in the published $\cos\theta_{\rm V}$ projections by the E791 Collaboration. Their overall fit quality, however, is good. Finally, there is a slight hint of our asymmetry effect in the BEATRICE data.

If our anomaly and interpretation is borne out by future high statistics charm experiments, it may prove interesting to revisit the comparison of experimental form factors with those predicted by Lattice Gauge Theories. Another important ingredient to a truly incisive comparison will be additional experimental and theoretical work to unravel the t dependence of the vector and axial form factors. Thus far, we and all other groups have basically assumed a plausible but untested form for this dependence.

Appendix A

Weighted Histograms and Averages

Throughout this thesis, we made significant use of weighted averages and histograms. We discuss the method used to compute errors on weighted histograms and averages since the method (applied to averages) is generally not widely known. In our data reduction program, the errors are obtained by accumulating various moments of the weights as one increments the weighted histogram or averaging plot.

A.1 Weighted Histogram

This is the simplest example which illustrates the general principle. A given bin of weighted histogram has a sum (or weighted yield) given by $S = \sum W$. We are looking for a way of computing $\sigma^2(S)$ due to counting statistic fluctuations of the data. Usually the weight for a given entry will depend on some well-measured property of the event such as a kinematic property. Often the weight is a continuous variable, but in order to make the problem tractable, we "cheat" and consider a discrete set of possible weights, $W \in \{W_\alpha\}$, where α is an integer which we will call the "weight class". Let n_α be the entries in the weight class α. The weighted yield is then:

$$S = \sum_\alpha n_\alpha W_\alpha \tag{1}$$

The advantage of using a discrete weight set $\{W_\alpha\}$ is that we can attribute fluctuations in the sum S to Poisson fluctuations of the number of events in each weight class n_α. We assume that these Poisson fluctuations obey:

$$< \ \delta n_\alpha \ \delta n_\beta \ \ >= \delta_{\alpha\beta} \ n_\alpha \tag{2}$$

which means that different bins are uncorrelated and the usual counting statistics applies: $\sigma = \sqrt{n}$. To calculate the variance on S we autocorrelate the fluctuations on δS due to fluctuations on the number of weight class entries δn_α

$$\delta S = \sum_\alpha \ W_\alpha \ \delta n_\alpha \quad , \quad \sigma^2(S) = \sum_\alpha \sum_\beta W_\alpha \ < \ \delta n_\alpha \ \delta n_\beta \ > \ W_\beta$$

$$\sigma_S^2 = \sum_\alpha \ W_\alpha^2 n_\alpha = \sum \ W^2 \tag{3}$$

The last equality says that sum of squared of each weight when multiplied by the number of events in each weight class is equivalent to just accumulating squared weights for each entry one at a time. We thus have the well known result:

$$\sigma_S = \sqrt{\sum \ W^2} \tag{4}$$

If all the weights are one (such as in an unweighted histogram), Equation (4) says that error is just $\sqrt{N}$. In weighted histogrammers, one simply increments $S \to S + W$ and $\sigma_S^2 \to \sigma_S^2 + W^2$ as each entry is added to each bin.

A.2 How Much Does Weight Variation Inflate Errors?

Let us define the average weight, $< \ W \ >$, and weight variance σ_W^2 through the equations:

$$< W >= \frac{\sum \ W}{N} = \frac{\sum_\alpha \ n_\alpha \ W_\alpha}{N} = \frac{S}{N}$$

$$\sigma_W^2 = \frac{\sum \ (W- < W >)^2}{N} = \frac{\sum_\alpha \ n_\alpha \ (W_\alpha - < W >)^2}{N} \tag{5}$$

where $N = \sum_\alpha \ n_\alpha$. It is straighforward to show and well known that $\sigma_W^2 =< W^2 > - < W >^2$ where

$$< W^2 >= \frac{\sum_\alpha \ n_\alpha \ W^2}{N} = \frac{\sum W^2}{N} = \frac{\sigma_S^2}{N} \tag{6}$$

We can manipulate Equations (5) and (6) to obtain:

$$\frac{\sigma_S}{S} = \sqrt{\frac{N\sigma_W^2}{N^2 < W >^2} + \frac{N \ \ < W >^2}{N^2 < W >^2}} = \frac{1}{\sqrt{N}} \sqrt{1 + \left(\frac{\sigma_W}{< W >}\right)^2} \tag{7}$$

Equation (7) says that there will be an inflation in the relative error of a weighted histogram beyond the $1/\sqrt{N}$ errors due to counting statistics on number of events in a given bin which is related to the rms spread of the contributing weights. Our experience is that often this inflation factor is fairly small **unless** one has negative weights. When a histogram has nearly canceling positive and negative weights due to *eg* severe background subtraction, or severely smeared deconvolution, one can have $\sigma_W \ \ \gg \ \ < W >$. The fractional error on such weighted averages can greatly exceed the fractional error on an unweighted histogram.

A.3 Errors on a Weighted Average

Consider the weighted average $\overline{X}$ of a quantity X subject to a weight W:

$$\overline{X} = \frac{\sum \ W \ X}{\sum \ W} = \frac{\sum_\alpha \sum_i \ n_{\alpha i} \ W_\alpha \ X_i}{\sum_\alpha \sum_i \ n_{\alpha i} \ W_\alpha} = \frac{A}{B} \tag{8}$$

A few notational comments are in order. We have introduced both a weight class α and X class i and are again thinking of discrete outcomes $W \in \{W_\alpha\}$ and $X \in \{X_i\}$. The variables $A \equiv \sum \ W \ X$ and $B \equiv \sum \ W$ are shorthand for the numerator and denominator of Equation (8). In this picture, the source of fluctuations on $\overline{X}$ are the Poisson fluctuations on $n_{\alpha i}$ — the number of events in both the weight class α and the X class i.

Unlike the case of the weighted histogram, we must Taylor expand the fluctuation $\delta\overline{X}$ to first order in fluctuation $\delta n_{\alpha i}$:

$$\delta\overline{X} = \left(\frac{1}{B}\frac{\partial A}{\partial n_{i\alpha}} - \frac{A}{B^2}\frac{\partial B}{\partial n_{i\alpha}}\right) \ \delta \ n_{i\alpha} = \left(\frac{W_\alpha \ X_i}{B} - \frac{A \ W_\alpha}{B^2}\right) \delta \ n_{i\alpha} \tag{9}$$

Autocorrelating and averaging, we have:

$$< \ \delta\overline{X}\delta\overline{X} \ \ >= \sum_{\alpha i\beta j} \ \left(\frac{W_\alpha \ X_i}{B} - \frac{A \ W_\alpha}{B^2}\right) \ \ < \ \delta n_{i\alpha}\delta n_{j\beta} \ \ > \ \ \left(\frac{W_\beta \ X_j}{B} - A \ \frac{W_\beta}{B^2}\right) \tag{10}$$

Invoking independent Poisson fluctuations

$$< \ \delta n_{i\alpha} \delta n_{j\beta} \ > \ = n_{i\alpha} \ \delta_{\alpha\beta} \ \delta_{ij} \tag{11}$$

we have:

$$\sigma_{\overline{X}}^2 = \frac{1}{B^2} \ \sum \ (WX)^2 + \frac{A^2}{B^4} \sum \ W^2 - \frac{2A}{B^3} \ \sum \ W^2 X$$

$$\sigma_{\overline{X}}^2 = \frac{\sum \ W^2 X^2}{\left(\sum \ W\right)^2} + \frac{\left(\sum \ WX\right)^2 \ \sum \ W^2}{\left(\sum \ W\right)^4} - 2 \ \frac{\left(\sum \ WX\right) \ \sum \ W^2 X}{\left(\sum \ W\right)^3} \tag{12}$$

In practice, Equation (12) is quite easy to implement. One merely must accumulate five moments for each desired weighted average. We have tested this formalism extensively using Monte Carlo techniques, and weighted averages are used extensively in FOCUS software.

References

[1] *Review of Particle Physics,* C. Caso *et al.,* The Europearn Physical Journal **C3** (1998) 1

[2] *Review of Particle Physics,* C. Caso *et al.,* The Europearn Physical Journal **C3** (2000) 1

[3] Ray Culbertson, *Four-body Semileptonic Decays of D Mesons,* Ph.D. thesis, University of Illinois at Urbana-Champaing, 1993

[4] J.G. Korner and G.A. Schuler, Z. Phys. C 46 (1990) 93.

[5] F.J. Gilman and R.L. Singleton, Phys. Rev. D 41 (1990) 142.

[6] F. Abe *et al.*, Phys. Rev. D **60**, 092005 (1999)

[7] P.L. Frabetti *et al.*, Phys. Lett. B **307**, 262 (1993)

[8] P.L. Frabetti *et al.*, Phys. Lett. B **307**, 262 (1993)

[9] N. Isgur, D. Scora, Phys. Rev. D **52**, 2783 (1995)

[10] E.M. Aitala *et al.*, Phys. Lett. B **440**, 435 (1998)

[11] E.M. Aitala *et al.*, Phys. Lett. B **450**, 294 (1999)

[12] AIP Conference Proceeding **349**: *Workshop on the Tau/Charm Factory* (1995)

[13] M. Adamocivh *et al.*, Eur. Phys. J. **C6**:35 (1999)

[14] Fermi National Lab, http://www.gnal.gov/

[15] E687 Collab., P.L. Frabetti *et al.*, Phys. Lett. B **407**, 79 (1997)

[16] *Techniques for Nuclear and Particle Physics Experiments,* W.R. Leo

[17] CLEO Collaboration, A. Bean *et al.*, Phys. Lett. B **317**, 647 (1993)

[18] M. Adamocivh *et al.*, Phys. Lett. B **268**, 142 (1991)

[19] ARGUS Collab., H. Albrecht *et al.*, Phys. Lett. B **255**, 634 (1991)

[20] K. Kodama *et al.*, Phys. Lett. B **286**, 187 (1992)

[21] See for example, Anna Hasenfratz and Peter Hasenfratz, Ann. Ref. Nucl. Part. Sci 35 (1985) 559-604.

[22] D. M. Schmidt, R. J. Morrison, and M. S. Witherell, Nucl. Instrum. Methods, A328 (1993) 547.

[23] P. Frabetti *et al.*,Phys. Lett.B331(1994)217

[24] Thomas E. Browder, Klaus Honscheid, and Daniele Pedrini, Ann. Rev. Nucl. Part. Sci. 46 (1997) 395

[25] E791 Collaboration, e-Print Archive: hep-ex/0012009

[26] James N. Simone, *Semileptonic Decays: an Update Down Under*, plenary talk presented at Lattice '95 Melbourne, Australia, Fermilab-Conf-96.017-T.

[27] D. Scorna and N. Isgur, Phys. Rev. D62 (1995) 2783

Vita

Christopher Alan Cawlfield was born and raised in St. Louis, Missouri. His middle school education was at The College School of Webster Groves, an experiential learning school that encouraged his enthusiasm for the sciences. He attended Principia College in Elsah, Illinois, a small liberal arts college, where he recieved his Bachelor of Science double-majoring in math and physics. In the summer of 1992, he received a summer undergraduate research fellowship (SURF) at Caltech. He studied under Mike Werner, researching interstellar dust grains using the IRAS satellite data. At Principia, his two undergraduate capstone projects were a study of a mechanical double pendulum and the construction of an astronomical CCD camera. He won departmental awards in both mathematics (The Mathematics Award), and physics (The James Watt Physical Science Award) in 1993.

At the University of Illinois, he joined the FOCUS group in 1994 and participated in the 1996–7 fixed target run. There, he helped to build, install, and calibrate the muon detector arrays. He created several web-based diagnostic tools during the run, including "Oddpack Access" and the "Oddpack Timelines." He has contributed to several aspects of data analysis and the FOCUS simulation. He was the recipient of the 1999 Giulio Ascoli Award for "demonstrating excellence and originality in the study of High Energy Physics" from the University of Illinois.