A SEARCH FOR D^0–$\overline{D}^0$ MIXING IN SEMILEPTONIC DECAYS FROM FOCUS

By

Michael G. Hosack

Dissertation

Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in

Physics

December, 2002

Nashville, Tennessee

Approved:	Date:

ACKNOWLEDGEMENTS

A high energy physics experiment requires the expertise of a large and diverse group of individuals. This includes not only those involved in the design, construction, operation, and analysis of the experiment and accelerators, but also all of the professionals that keep the laboratory running. This enterprise would not be possible without the Department of Energy, the National Science Foundation, and the taxpayers who support these agencies.

The data used in this analysis is the product of the efforts of many, as is the arsenal of software developed to reconstruct and analyze the data. I would like to thank the FOCUS collaboration for all of their efforts as well as for being incredibly friendly and supportive.

I would particularly like to thank Matt Nehring for getting me off to a good start on this analysis. The analysis committee—Jim Wiss, John Cumalat, Matt Nehring, and Ignacio Bediaga—impressed upon me the value of keeping the methodology as simple as possible. This has made the analysis more understandable and a pleasure to describe. I would also like to thank the post-doc's Eric Vaandering and Kevin Stenson. A number of times in this analysis Kevin's suggestions helped me out of an impasse. Eric read and commented on much of this thesis. Both he and Will Johns urged me to present the "what before the how." To the extent that I was able to do this, the thesis is more readable. I would also like to thank Will for offering other perspectives and helping me to realize that a perfect analysis is an obsolete analysis.

My advisor, Paul Sheldon, I thank for suggesting this analysis topic, being a tireless advocate for our research group, and being a congenial colleague and friend.

To Medford Webster, without whose generosity and patience this thesis would be far less convincing, I am forever grateful.

I would like to thank former Vanderbilt graduate student Dan Engh for his work improving the muon simulation as well as his invaluable help with Skim1. I am thankful to former Vanderbilt graduate students Steven Jacobson and Tracy Huard for lending an ear to the tribulations and breakthroughs of my work here.

I am thankful to my thesis committee and my future boss, Tom Coan, for being patient as my analysis ran against one delay after another.

Finally, I would like to thank all the relatives who have supported and encouraged me over the years, especially my parents. They worked hard to provide me with the freedom to pursue my dreams. I am incredibly blessed to have such a loving family.

TABLE OF CONTENTS

Page

Chapter

Appendix

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

I.1 Motivation and Overview

The Standard Model of particle physics has proven to be enormously successful. It describes all known fundamental particles and interactions, with the exception of gravitation (which is not included in the model). The model, in its present form, was completed by 1979. It was formed by joining the Weinberg-Salam electroweak theory[1] with the theory of strong interactions,[2] Quantum Chromodynamics (QCD). All of the fundamental particles in the Standard Model have been experimentally verified except for the Higgs boson.

The model has withstood all experimental tests, except for recent measurements indicating neutrinos can oscillate (mix) between flavors. Neutrino oscillations can be accommodated by allowing neutrinos to have mass; they are massless in the Standard Model. However, this is not a serious conceptual difficulty with the underpinnings of the model. There are a number of extensions to the Standard Model which include neutrinos with mass [1]. Any updated Standard Model, however, will likely include more fundamental constants to account for this.

The large number of fundamental constants is one reason the Standard Model is unsatisfying. There are 19 constants: 3 coupling constants, 9 masses, 3 mixing

---

[1] Electroweak theory combines electromagnetic forces with weak interactions. The most familiar consequence of weak interactions is nuclear beta decays.

[2] Strong interactions bind protons and neutrons in the nucleus. Strong interactions are responsible for most of the mass in the observable universe. Observations indicate the presence of more non-luminous "dark matter" than can be accounted for by current cosmology.

angles, 1 phase, 1 Higgs self coupling, 1 Higgs-related constant with mass units, and 1 QCD vacuum energy parameter which determines the level of CP violation allowed in strong interactions [1]. The model gives no prediction for the values of these constants. They are determined by measurements.

There are a number of approaches for testing a successful theory or model with experiment. One way is to make more precise measurements. These measurements may involve testing null or small predictions. Yet another strategy is to look in untested regions. If a model has a history of success, all of these approaches likely will provide further calculational challenges for theorists (since the easier calculations have likely already been done and tested). This is especially true with the subject of this thesis: D^0–$\overline{D}^0$ mixing.

Early predictions for D^0–$\overline{D}^0$ mixing in the Standard Model indicated an effect far too small to be measured (assuming the Standard Model is correct). Further studies found, however, that non-perturbative calculations are needed to make estimates at the level of the experimental sensitivities. Most of the quantitative predictions made by the Standard Model have involved perturbation theory.[3] Approaches for non-perturbative calculations are not as well developed. At this time, it is not clear whether errors in theoretical calculations can be made smaller than the experimental sensitivities. This will depend in part on measurements not directly related to mixing which are used as inputs to the calculations (see Sec. I.3).

The following chapters present an analysis of mixing with data from the FOCUS experiment at Fermi National Accelerator Laboratory. This chapter provides a brief

---

[3] Perturbation theory assumes the kinetic energies of the interacting particles are much greater than the energy of the interaction.

introduction to D^0–$\overline{D}^0$ mixing and the FOCUS experiment. Chapter II describes how the charm particles were produced. Chapter III describes the FOCUS spectrometer. Chapter IV discusses reconstruction of basic quantities: momentum, position, and particle type as well as initial data reduction. Chapter V discusses the simulation of the experiment. Chapter VI describes the analysis methodology and Chapter VII discusses systematic errors. Chapter VIII summarizes the results from this measurement, results from other measurements, and discusses ongoing and future D^0–$\overline{D}^0$ mixing searches.

I.2 D^0–$\overline{D}^0$ Mixing

A D^0 meson is composed of a charm (c) and an anti-up ($\bar{u}$) quark. D^0's are flavor eigenstates produced in strong interactions. Quarks also interact via the weak force, which causes the eigenstate to evolve into a *mixture* of $c\bar{u}$ and $\bar{c}u$. The weak Hamiltonian, H_{wk}, gives the time evolution of a $D^0/\overline{D}^0$ as[4]

$$\begin{aligned} i\frac{\partial}{\partial t}\begin{pmatrix} D^0 \\ \overline{D}^0 \end{pmatrix} &= H_{\mathrm{wk}}\begin{pmatrix} D^0 \\ \overline{D}^0 \end{pmatrix} \\ &= \begin{pmatrix} M - i\Gamma/2 & M_{12} - i\Gamma_{12}/2 \\ M_{12}^* - i\Gamma_{12}^*/2 & M - i\Gamma/2 \end{pmatrix}\begin{pmatrix} D^0 \\ \overline{D}^0 \end{pmatrix}. \end{aligned} \tag{1}$$

Note H_{wk} is not Hermitian since a $D^0/\overline{D}^0$ has a finite lifetime, $1/\Gamma$. Diagonalizing the weak Hamiltonian, H_{wk}, gives weak eigenstates

$$D_H = pD^0 + q\overline{D}^0 \text{ and } D_L = pD^0 - q\overline{D}^0 \tag{2}$$

---

[4] $(H_{wk})_{11} = (H_{wk})_{22}$ in order to insure CPT invariance.

with masses and lifetimes, $M_{\mathrm{H,L}}$ and $\Gamma_{\mathrm{H,L}}$. The constants, p and q are functions of M, Γ, M_{12}, and Γ_{12}. If H_{wk} conserves CP, $|p| = |q|$, and D_{H} and D_{L} are CP eigenstates. Experimentally, it has been found that $|p|$ and $|q|$ are nearly equal [2].

Starting with a D^0, the probability of having a $\overline{D}^0$ at time t is:

$$r_{\mathrm{mix}}(t) = \frac{1}{4}\left|\frac{q}{p}\right|^2 [e^{-\Gamma_{\mathrm{H}}t} + e^{-\Gamma_{\mathrm{L}}t} - 2e^{-(\Gamma_{\mathrm{H}}+\Gamma_{\mathrm{L}})t/2}\, cos(\Delta M t)] \tag{3}$$

where $\Delta M \equiv (M_{\mathrm{H}} - M_{\mathrm{L}})$ and $\Delta\Gamma \equiv (\Gamma_{\mathrm{H}} - \Gamma_{\mathrm{L}})$. Since experimentally it is known that $\Delta M \ll \Gamma$ and $\Delta\Gamma \ll \Gamma$, to a good approximation

$$r_{\mathrm{mix}}(t) = \frac{1}{4}e^{-\Gamma t}\left|\frac{q}{p}\right|^2 \left(\Delta M^2 + \frac{1}{4}\Delta\Gamma^2\right) t^2. \tag{4}$$

Defining $x \equiv \Delta M/\Gamma$ and $y \equiv \Delta\Gamma/2\Gamma$,

$$r_{\mathrm{mix}}(t) = \frac{1}{4}e^{-\Gamma t}\left|\frac{q}{p}\right|^2 (x^2 + y^2)(\Gamma t)^2. \tag{5}$$

Integrating over all time gives

$$r_{\mathrm{mix}} = \frac{1}{2}\left|\frac{q}{p}\right|^2 (x^2 + y^2). \tag{6}$$

Therefore, r_{mix} is the fraction of D^0's decaying as $\overline{D}^0$'s. The fraction of $\overline{D}^0$'s decaying as D^0's is

$$\bar{r}_{\mathrm{mix}} = \frac{1}{2}\left|\frac{p}{q}\right|^2 (x^2 + y^2). \tag{7}$$

Equation (7) differs from (6) in the interchange of p and q.

I.3 Theoretical Predictions for D^0–$\overline{D}^0$ Mixing

I.3.1 Short Distance

The lowest order contributions to mixing in the Standard Model can be represented by the Feynman diagrams in Fig. 1 (so called "box diagrams"). The

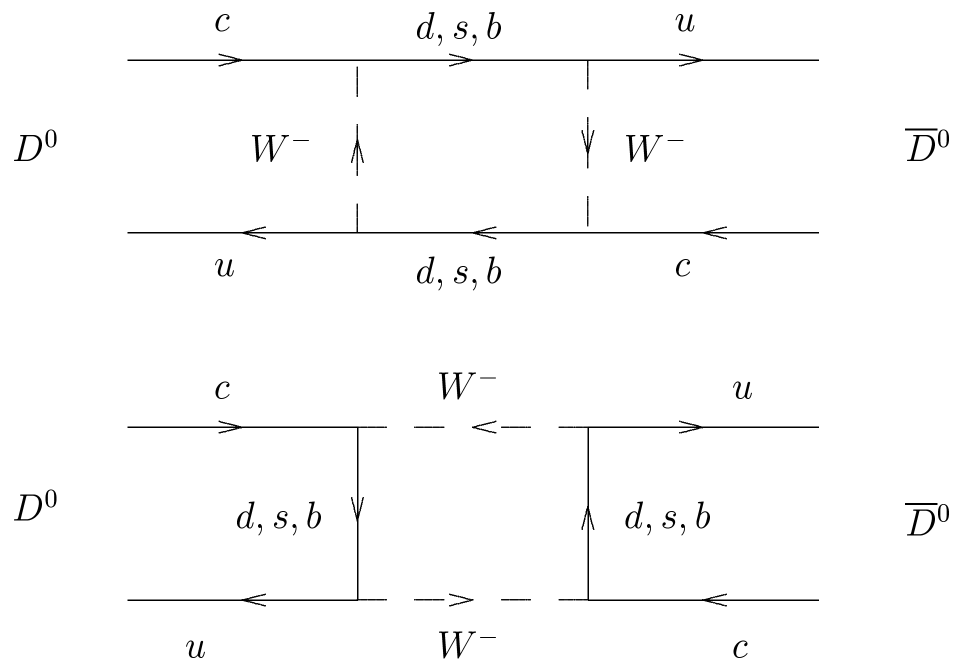


Figure 1: Box diagrams.

contribution from these diagrams is $r_{\text{mix}} \sim 10^{-10}$–$10^{-9}$, which is far below the mixing calculated and observed in either s or b quark mesons. The strength of flavor changing transitions among quarks is parameterized by a 3×3 unitary matrix in the Standard Model. In terms of the elements, V_{ij}, of this "Cabibbo-Kobayashi-Maskawa" (CKM) matrix, the box diagrams give [3]

$$\langle \overline{D}^0 | H_{\text{wk}} | D^0 \rangle = \sum_i \sum_j V_{ci}^* \, V_{cj}^* \, V_{ui} \, V_{uj} \, S(i,j) \tag{8}$$

where $S(i,j)$ is a function of the quark masses (m_i) and i and j run over the $-1/3$ charged quarks: d, s, b. The functions $S(i,j)$ consist of integrals over the internal quark momenta of the box diagrams. Unitarity of the CKM matrix, the small value of $(m_s^2 - m_d^2)/m_W^2$, the small value of $V_{cb}^* \, V_{ub}$, and the large value of m_c relative to m_d and m_s all "conspire" to make this sum very nearly zero [4].

Using unitarity, Eq. (8) becomes

$$\begin{aligned}\langle\overline{D}^0|H_{\rm wk}|D^0\rangle &= V^*_{cd}\,V_{ud}\,V^*_{cs}\,V_{us}[S(d,s)+S(s,d)-S(d,d)-S(s,s)]\\ &\quad+V^*_{cb}\,V_{ub}\,V^*_{cs}\,V_{us}[S(b,s)+S(s,b)-S(b,b)-S(s,s)]\\ &\quad+V^*_{cb}\,V_{ub}\,V^*_{cd}\,V_{ud}[S(d,b)+S(b,d)-S(d,d)-S(b,b)] \qquad (9)\end{aligned}$$

or with $F(i,j)\equiv[S(i,j)+S(j,i)-S(i,i)-S(j,j)]$,

$$\langle\overline{D}^0|H_{\rm wk}|D^0\rangle = V^*_{cd}\,V_{ud}\,V^*_{cs}\,V_{us}\,F(d,s)+V^*_{cb}\,V_{ub}\,V^*_{cs}\,V_{us}\,F(b,s)+V^*_{cb}\,V_{ub}\,V^*_{cd}\,V_{ud}\,F(d,b). \qquad (10)$$

Note $F(i,j)=F(j,i)$ and $F(i,i)=0$. Since $m_d \ll m_b$ and $m_s \ll m_b$, $F(b,s)\doteq F(d,b)\doteq F(b)$ (see [3] for details). Combining the last two terms in Eq. (10) then gives

$$\langle\overline{D}^0|H_{\rm wk}|D^0\rangle \doteq V^*_{cd}\,V_{ud}\,V^*_{cs}\,V_{us}\,F(d,s)+V^*_{cb}\,V_{ub}\,(V^*_{cs}\,V_{us}+V^*_{cd}\,V_{ud})F(b). \qquad (11)$$

The second term can be simplified by using the unitarity condition $V^*_{cs}V_{us}+V^*_{cd}V_{ud}+V^*_{cb}V_{ub}=0$. Thus

$$\langle\overline{D}^0|H_{\rm wk}|D^0\rangle \doteq V^*_{cd}\,V_{ud}\,V^*_{cs}\,V_{us}\,F(d,s)-(V^*_{cb}\,V_{ub})^2\,F(b). \qquad (12)$$

The first term can be simplified further. Using unitarity again:

$$V^*_{cs}\,V_{us}\,V^*_{cd}\,V_{ud}=-V^*_{cs}\,V_{us}\,(V^*_{cs}\,V_{us}+V^*_{cb}\,V_{ub}). \qquad (13)$$

Also $|V^*_{cb}\,V_{ub}|\sim 10^{-3}\,|V^*_{cs}\,V_{us}|$ so

$$V^*_{cs}\,V_{us}\,V^*_{cd}\,V_{ud}\doteq-(V^*_{cs}\,V_{us})^2 \qquad (14)$$

and Eq. (12) becomes

$$\langle\overline{D}^0|H_{\rm wk}|D^0\rangle \doteq -(V^*_{cs}\,V_{us})^2F(d,s)-(V^*_{cb}\,V_{ub})^2\,F(b). \qquad (15)$$

$(V_{cb}^* V_{ub})^2$ is 1.0×10^6 to 5.4×10^6 times smaller in magnitude than $(V_{cs}^* V_{us})^2$. Explicit estimates of the integrals, however, show a factor of $(m_b/m_W)^2$ associated with the b quark term compared to

$$F(d,s) \propto \frac{(m_s^2 - m_d^2)}{m_W^2} \cdot \frac{(m_s^2 - m_d^2)}{m_c^2} \tag{16}$$

for the light quark term [4, 3]. $(m_s^2 - m_d^2)/m_W^2$ reflects the fact that $F(d,s)$ would be zero if $m_s = m_d$ ("GIM" suppression [5]). $(m_s^2 - m_d^2)/m_c^2$ arises from momentum transferred from the charm quark to the internal quarks. (This factor is generally neglected in B^0 and K^0 mixing since the dominant internal quark there is significantly more massive than the external quarks.) Relative to the light quark term, the b quark term is only

$$\frac{m_b^2\, m_c^2}{(m_s^2 - m_d^2)^2} \times \left(\frac{1}{5.4 \times 10^6} \text{ to } \frac{1}{1.0 \times 10^6} \right) \approx 0.015 \text{ to } 0.079 \tag{17}$$

times as large.

For mixing in K^0 $(d\bar{s})$, B^0 $(d\bar{b})$, and B_s^0 $(s\bar{b})$, the internal loop quarks are the charged $+2/3$ quarks: u, c, t.[5] The analogous expression to Eq. (8) is [6]

$$\langle \overline{B}^0 | H_{\text{wk}} | B^0 \rangle = \sum_i \sum_j V_{id}\, V_{ib}^*\, V_{jd}^*\, V_{jb}\, S(i,j). \tag{18}$$

The CKM factors for B^0 mixing are all of similar size, with the t quark $S(t,t)$ term dominating: $\langle \overline{B}^0 | H_{\text{wk}} | B^0 \rangle \approx |V_{td} V_{tb}^*|^2\, S(t,t) \approx |V_{td}|^2\, S(t,t)$. The $S(t,t)$ integrals produce a factor of m_t^2/m_W^2. Comparing this with Eq. (16) for D^0 mixing, it is clear the box diagrams are highly suppressed in D^0 mixing compared to B^0 mixing.

---

[5] D^0 mesons are the only experimentally accessible mesons which mix using the $-1/3$ charged quarks for the internal loops. The two other (Standard Model) mesons which could do this, T^0 $(u\bar{t})$ and T_c^0 $(\bar{t}c)$, do not exist as bound states since the t quark decays immediately.

$S(t,t)$ is also the largest term in B_s^0 mixing: $\langle \overline{B}_s^0 | H_{\rm wk} | B_s^0 \rangle \approx |V_{ts} V_{tb}^*|^2 \, S(t,t) \approx |V_{ts}|^2 \, S(t,t)$. The B_s^0 mixing rate is expected to be roughly $|V_{ts}/V_{td}|^4$ times larger than the B^0 mixing rate (recall the rate is the square of the amplitude). Mixing in B_s^0 mesons has not been observed due to limited statistics. Limits on B_s^0 mixing together with B^0 mixing measurements can be used to place constraints on $|V_{td}|$ [6].

In K^0 mixing, the large $S(t,j)$ terms are suppressed by the relatively small $V_{td}V_{ts}^*$ factors, so mixing is dominated by the $S(c,c)$ term: $\langle \overline{K}^0 | H_{\rm wk} | K^0 \rangle \approx |V_{cd} V_{cs}^*|^2 \, S(c,c) \approx |V_{cd}|^2 \, S(c,c)$. (Although small, the $S(t,c)$ term must be included to account for CP violation [1].) The $S(c,c)$ integrals produce a factor of m_c^2/m_W^2. Again, this is much larger than the corresponding D^0 mixing factor shown in Eq. (16).

I.3.2 Long Distance

Since the internal quarks involved in D^0–$\overline{D}^0$ mixing are light, there are a large number of intermediate hadronic states that can contribute to the amplitude for mixing. Because the contribution from the box diagrams is so small, it is likely that these intermediate "long distance" states dominate mixing. In K^0–$\overline{K}^0$ mixing, the long-distance effects are comparable to the box diagram effects, however, the long-distance effects largely cancel [7]. It is not clear to what degree long-distance effects cancel in D^0–$\overline{D}^0$ mixing.

Long-distance interactions are inherently non-perturbative. Two main approaches have been used to estimate long distance contributions to D^0–$\overline{D}^0$ mixing: "dispersive" approaches and "Heavy Quark Effective Field Theory" (HQEFT). Early dispersive estimates for long-distance mixing indicated $r_{\rm mix}$ could be roughly as large as a percent [8]. More detailed dispersive calculations [9] indicated mixing rates too small to measure but still roughly a 100 times larger than the box calculation. Later calculations using HQEFT indicated effects that are within a factor of ten of the box calculation [10, 11]. More recent dispersive calculations indicate mixing might be close to experimental limits [12, 13].

In dispersive calculations, the long-distance interactions are mediated by virtual mesons. (For example, the D^0 can decay via a weak interaction into a virtual $\pi^-\pi^+$ which then recombines via another weak interaction into a $\overline{D}^0$.) These interactions are categorized as n-particle intermediate states related by $SU(3)$ flavor symmetry (i.e., they consist of mesons with the three closest mass quarks: u, d, and s). For $n > 1$, the larger the symmetry breaking, the larger the mixing contribution. In the limit of exact $SU(3)$ flavor symmetry, the mixing contributions from each set of n particles in the $SU(3)$ group cancel exactly [9].

Consider for example the group of two charged pseudoscalars (spin zero): K^-K^+, $\pi^-\pi^+$, $K^-\pi^+$, $K^+\pi^+$. $SU(3)$ flavor symmetry is known to be badly broken in charm decays. Experimentally [14, 15, 16],

$$\frac{\mathrm{BR}(D^0 \to K^-K^+)}{\mathrm{BR}(D^0 \to \pi^-\pi^+)} = 2.83 \pm 0.15 \tag{19}$$

but this ratio would be 1 for exact $SU(3)$ flavor symmetry and a pure spectator diagram (i.e. weak interactions only) calculation yields a ratio of ~ 1.4 [17]. This gives some indication that mixing from long-distance effects could be large [9]. Dispersive estimates unfortunately depend strongly on experimental uncertainties of the D^0 branching ratios. Because of these uncertainties, Standard Model mixing could be near experimental limits [13].

Other dispersive examples come from 1-body (resonant) intermediate states.

These involve particles with masses near the D^0 mass which can decay to the same final state as a D^0. Examples of such resonances include $K(1830)$, $K(1460)$, $\eta(1760)$, and $\pi(1800)$. These appear to give mixing rate contributions only within a factor of ten of the box contributions but have implications for CP violation [18].

The large effects seen in 2-body dispersive calculations are in contrast to the small values obtained from HQEFT, which are the same order of magnitude as the box diagram estimates. In HQEFT, the mass of the heavy quark (c here) is assumed to be much larger than the scale of strong interactions (usually taken to be near 1 GeV). This assumption is questionable but corrections are possible in principle. Since most of the momentum is carried by the c, non-leptonic D^0 decays are forbidden to leading order in HQEFT [10]. Long-distance effects are accounted for entirely by the running of the renormalized mass below m_c.

There are at least two ways of reconciling the small values for mixing predicted by HQEFT with the large effects seen in dispersive calculations. One is that the dispersive contributions among the individual $SU(3)$ flavor representations cancel [10, 11, 13]. Another interpretation is that the approximation $m_c \gg 1$ GeV$/c^2$ is inadequate or corrections are large. Other leading order HQEFT calculations for charm (not mixing) have had varied success [19, 20]. Next to leading order calculations have been developed [21] as well as an alternative expansion scheme [22] (again in non-mixing contexts). These provide some hope for improved Standard Model D^0–$\overline{D}^0$ mixing calculations in the future.

I.3.3 Other Models and Comparisons of Predictions

Some theories predict extensions to the Standard Model which can give rise to large mixing. Some examples of theories and models predicting mixing close to the experimental bounds include: supersymmetry with quark-squark alignment [23], four generation theories [24], and theories containing leptoquarks [25].

In general, both Standard Model and non-Standard Model predictions of mixing vary by several orders of magnitude, as can be seen from the compilation by Nelson [26] (shown in Fig. 2).

An observation of D^0–$\overline{D}^0$ mixing at present experimental sensitivities would indicate either long-distance interactions (which are difficult to compute), or the presence of non-Standard Model physics.

I.4 Experimental Signatures

To identify mixing one needs to know the flavor content of a neutral meson at some initial time and again at some later time. If the meson has changed into its antiparticle, then mixing has occurred. Since misidentification can mimic a wrong sign signal, another indicator of mixing is useful. The proper lifetime[6] distribution for mixed events is one such indicator. For D^0 mesons, the proper lifetime distribution for mixing has a characteristic $t^2e^{-\Gamma t}$ dependence (see Eqs. (5) and (6)). Events in which mixing does not occur (an initial D^0 decays as a D^0) have a proper lifetime distribution proportional to $e^{-\Gamma t}$, within the approximations discussed in Sec. I.2.

A pure D^0 eigenstate can be produced in strong decays.[7] Traditionally, this is

---

[6] "Proper lifetime" means the decay time in the rest frame of the decaying particle.

[7] At least this can be done in the Standard Model. In some unified field theories this may only be approximately true.

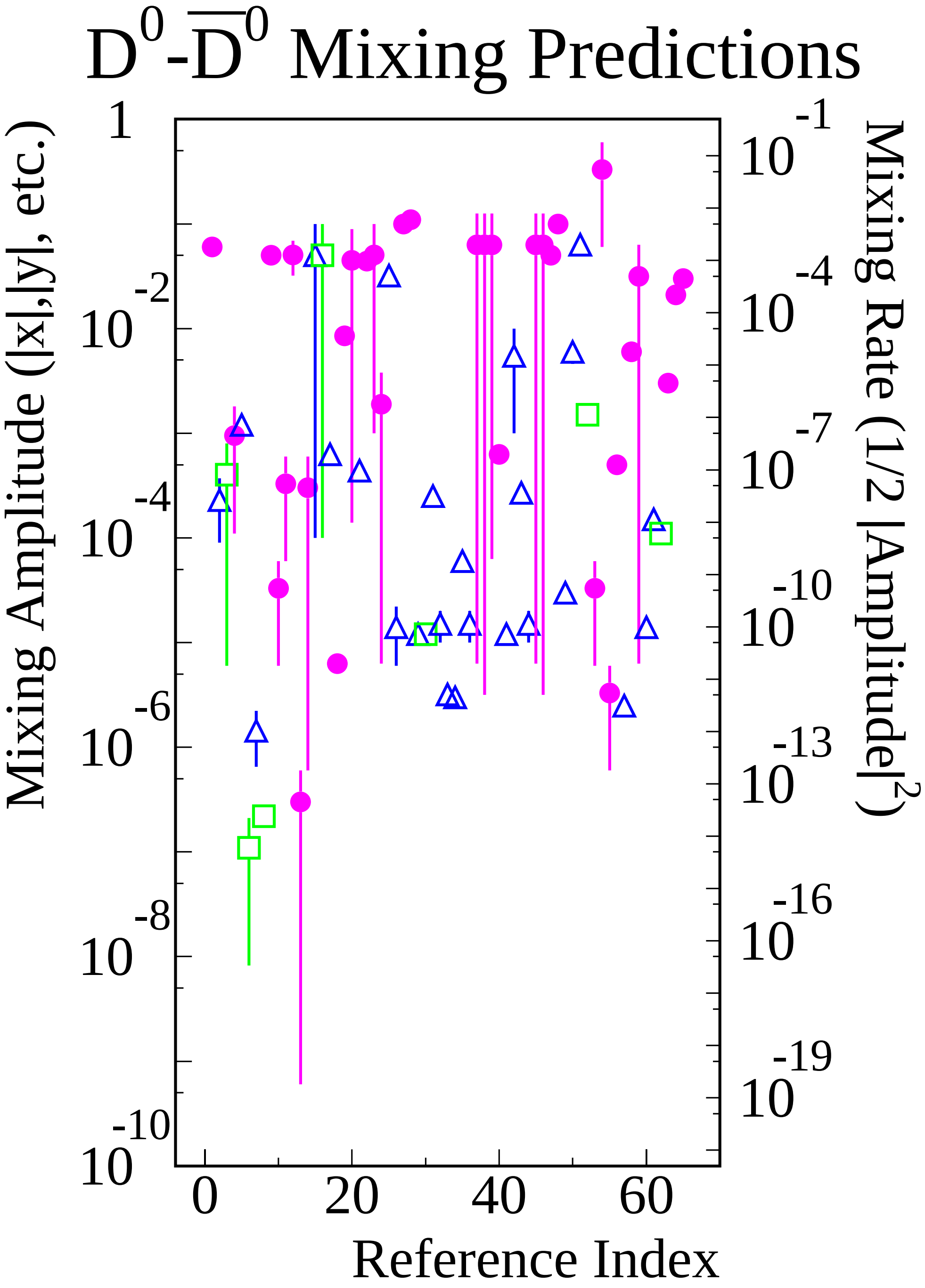


Figure 2: Theoretical “guidance.” The triangles are Standard Model predictions for x, the squares are Standard Model predictions for y, and the circles are non-Standard Model predictions for x. (Figure from H. Nelson, UCSB HEP 99-08, 1999, SLAC SPIRES preprint hep-ex/9908021.)

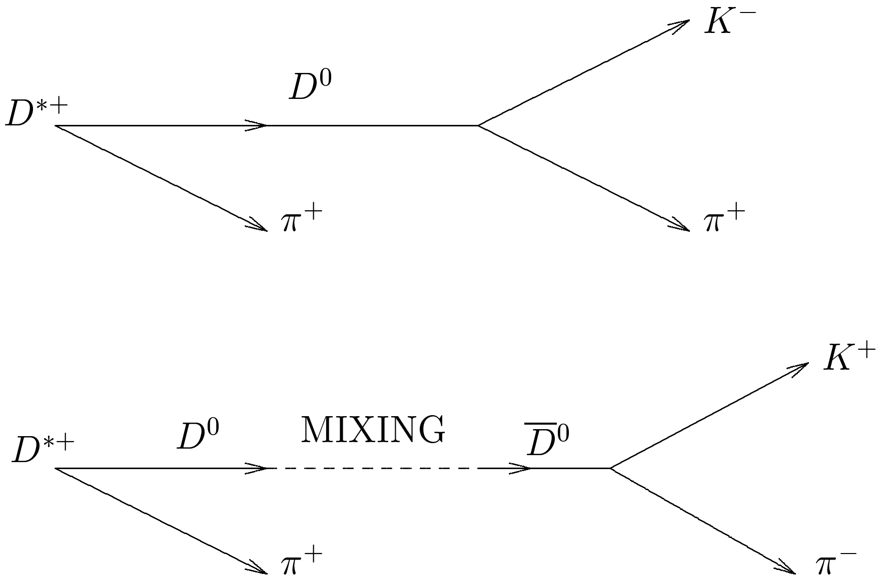


Figure 3: A candidate for mixing arises in $D^0 \to K^-\pi^+$ decays when the pion from the neutral D decay has a charge opposite that of the pion from the D^* decay.

done with the decays $D^{*+} \to D^0\pi^+$ and $D^{*-} \to \overline{D}^0\pi^-$, which is the approach taken in this analysis. The D^{*+} decays essentially instantly after being produced. The charge of the π identifies the initial flavor of the D^0, since flavor is conserved in strong interactions: $D^{*+}(c\bar{d}) \to D^0(c\bar{u})\,\pi^+(u\bar{d})$. The final D^0 flavor can often be determined from its weak decay daughters. A mode frequently used to search for mixing is $D^0 \to K^-\pi^+$ [27, 28]. This mode has the advantage of being relatively easy to reconstruct. Candidates for mixing arise when the pion from the D^{*+} has an opposite charge ("wrong sign") from the pion in the decay of the D^0 (see Fig. 3). The lowest order Feynman diagram for $D^0 \to K^-\pi^+$ is shown in Fig. 4. However there are rare processes in which the daughters of the D^0 can decay into the "wrong sign" without mixing. This occurs with Doubly Cabibbo Suppressed (DCS) decays (Fig. 5). One can derive the proper lifetime distribution for wrong sign decays by

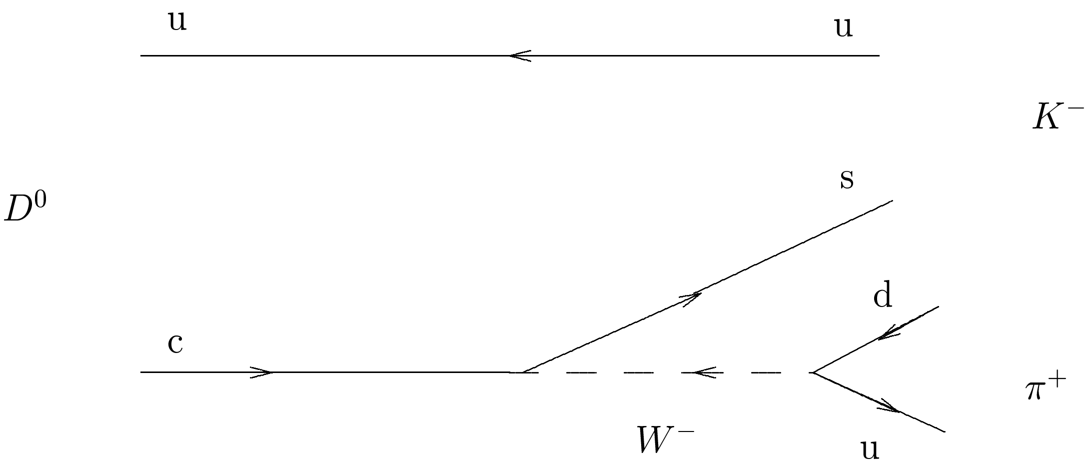


Figure 4: Lowest order Feynman diagram for $D^0 \to K^-\pi^+$.

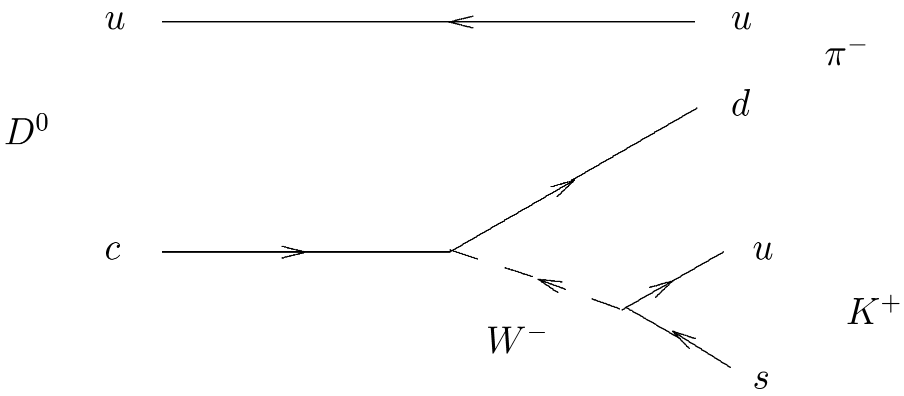


Figure 5: Doubly Cabibbo Suppressed (DCS) decays of D^0 can also give rise to wrong sign events. Both vertices involve cross generational transitions, thus the decay is "doubly suppressed." Although small, DCS decays likely occur at levels exceeding mixing.

combining the diagrams for DCS and mixing. Analysis of $D^0 \to K^+\pi^-$ is interesting in that it provides information about both DCS and mixing, however sensitivity to mixing is somewhat impaired by fitting to a more complicated time distribution.

An approach which provides information about the y part of mixing (see Eq. (6)), without measuring mixing, is to look for an asymmetry between the lifetime of the CP even final state $D^0 \to K^+K^-$, and the CP mixed final state $D^0 \to K^-\pi^+$ [29].

An approach which avoids the DCS complication is to look at semileptonic decays of D^0, such as $D^0 \to K^-\bar{\ell}\nu_\ell$ [30] (Fig. 6). This is the approach taken for this analysis. Although the mixing proper decay time distribution is simpler for semileptonic decays, the neutrino (ν_ℓ) is not reconstructed, so both D^{*+} mass and D^0 decay time resolutions are impaired compared to the fully charged hadronic decays. In semileptonic decays, the charge of the lepton uniquely defines the flavor of the neutral D. When the muon has a charge opposite that of the pion from the D^* ("wrong sign" decays) mixing has occurred (Fig. 7).

Reconstruction, backgrounds, and misidentification of semileptonic candidate events are discussed in Chapters VI and VII.

I.5 The FOCUS Experiment

The data used in this analysis were collected in the FOCUS experiment during the 1996–1997 fixed target run at the Fermi National Accelerator Laboratory (Fermilab). FOCUS/E831 is a charm photoproduction experiment and an upgrade of FNAL-E687. FOCUS has collected over one million fully reconstructed charm hadron decays.

The FOCUS collaboration consists of roughly 110 people from 17 institutions in 5 countries (United States, Italy, Brazil, Mexico, and Korea). Table 1 lists the

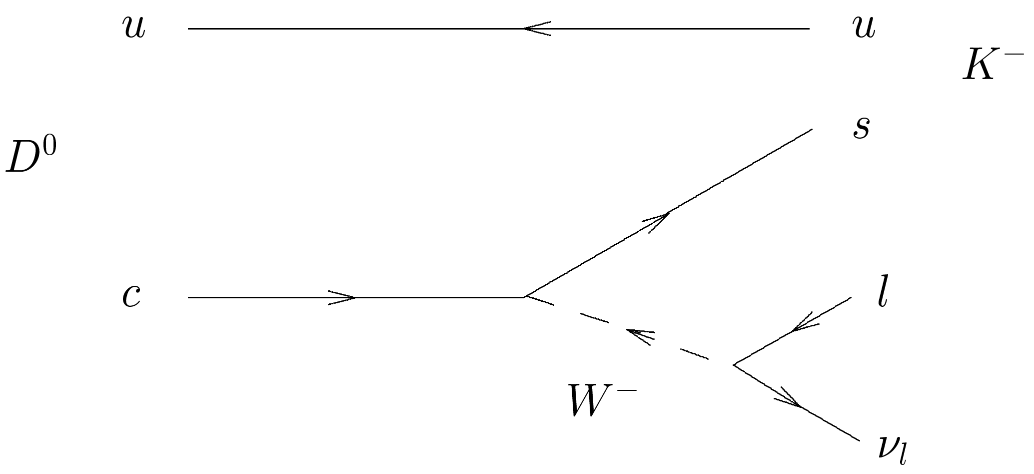


Figure 6: Diagram for $D^0 \to K^-\bar{\ell}\nu_\ell$. The charge of the lepton (or kaon) determines the flavor of the neutral D.

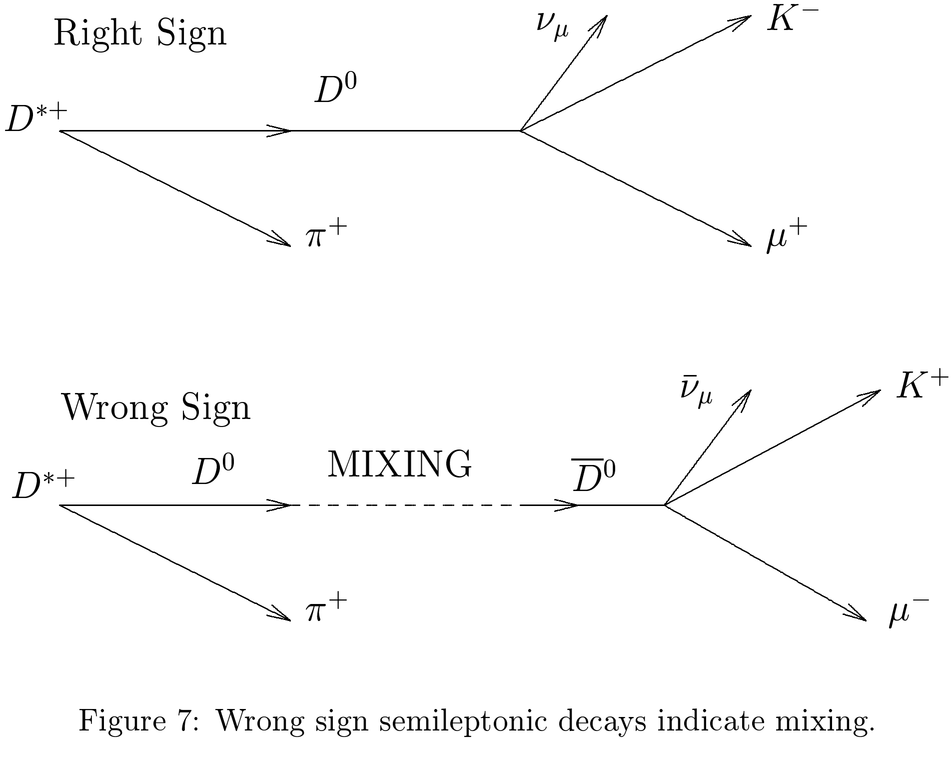


Figure 7: Wrong sign semileptonic decays indicate mixing.

Table 1: Institutions participating in FOCUS. The FOCUS collaboration consists of roughly 110 people from 17 institutions in 5 countries (United States, Italy, Brazil, Mexico, and Korea).

University of California, Davis
CBPF (Brazil)
CINVESTAV (Mexico)
University of Colorado, Boulder
Fermi National Accelerator Laboratory
INFN and Frascati Laboratory
University of Illinois, Urbana-Champaign
Korea University, Seoul
INFN and University of Milano
University of North Carolina, Ashville
INFN and University of Pavia
Universidad Autonoma de Puebla (Mexico)
University of Puerto Rico, Mayaguez
University of South Carolina
University of Tennessee
Vanderbilt University
University of Wisconsin, Madison

institutions participating in FOCUS.

The FOCUS experimental hall is located at the Fermilab Wideband facility (wideband refers to the broad range of photon energies). The photons there were produced from bremsstrahlung of electrons off lead nuclei. These *real* (as opposed to virtual) photons are the highest energy real photons produced by humans with an average triggered energy of $\sim 175\,\mathrm{GeV}$ in FOCUS.[8]

Photons are well suited for producing charm mesons. A meson beam produces a smaller fraction of charm hadrons relative to other hadrons than a photon beam. For a photon beam:

$$\frac{\text{Rate of charm production}}{\text{Rate of total hadronic production}} \simeq 1\%.$$

---

[8] E687 ran with a higher energy but lower rate.

For hadron beams, this ratio is $\sim 0.1\%$. Qualitatively this difference arises from flavor conservation in strong interactions—the *u*-*d* quark content of a meson beam produces more *u*-*d* containing hadrons than charm mesons. The mechanisms for charm photoproduction involve a photon resolving into a charm-anticharm pair within a nucleon. With an electron beam this can be achieved via a virtual photon. In a fixed target experiment, however, an electron beam also produces a large number of unwanted electromagnetic interactions. With a real photon beam, however, electron-positron and muon-antimuon pair production is a significant background. As we will see, measures were taken in the experimental apparatus design to minimize the impact of these backgrounds.

The following chapter discusses how the charm particles are produced. This is followed by a chapter describing the spectrometer. Much of this material is drawn from the extensive article, "Description and Performance of the Fermilab E687 Spectrometer" [31], which the reader can refer to for further details. FOCUS is a major upgrade of E687. Nearly all detector systems changed in some fashion. Differences between FOCUS and E687 will be noted explicitly.

CHAPTER II

THE FOCUS BEAM LINE

II.1 Protons from the Tevatron

The acceleration of protons to near TeV energies proceeds through several types of accelerators. Each accelerator is optimal for a particular range of energies. The five accelerators used at Fermilab for the 1996–97 fixed target run are summarized below and shown in Fig. 8.

Cockcroft-Walton: This accelerator consists of a repeating lattice of diodes and capacitors [32] which builds up a single voltage gap of 750 kV. H^- ions are accelerated across this gap and sent to the LINAC.

LINAC: This is an Alvarez type linear accelerator [33]. It has a coaxial type geometry. The central conductor is split at intervals, forming "drift tubes" separated by gaps. The drift tubes and outer conductor are connected to a radio-frequency (rf) oscillator. This forms a resonant rf cavity with electromagnetic waves propagating in the space between the tubes and outer conductor. The acceleration is provided by roughly parallel electric field lines running across each gap. The field lines reverse direction every half period of the oscillation. During the reversal, the ions are shielded from the fields by traveling within a drift tube. Each successive drift tube is longer than the previous, to account for the increasing velocity of the ions.

The H^- ions are accelerated by the LINAC to a kinetic energy of 400 MeV. They then pass through a thin carbon foil which strips the electrons off, leaving a proton. The proton is sent to the Booster.

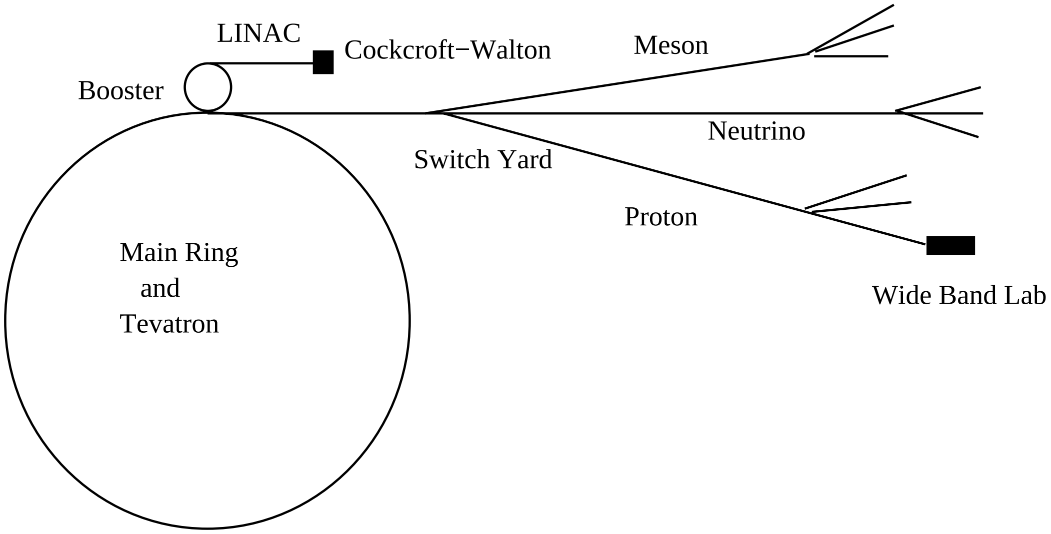


Figure 8: Schematic diagram of the Fermilab accelerators and beam lines.

II.1.1 Synchrotrons

The accelerators after the LINAC are synchrotrons. A synchrotron [33] has a ring geometry, which enables a particle to pass through accelerating rf cavities many times, reducing the number of rf cavities required. The beam is confined to a circular path with magnets. The bending magnetic fields are synchronized to the momentum to maintain a constant radius. Also, the rf frequency is synchronized to the arrival time of the beam particles.

Charged particles moving in a bent path emit substantially more radiation than a particle accelerating in a straight path. For a given momentum, the power emitted by synchrotron radiation is inversely proportional to the fourth power of the mass. This makes synchrotrons substantially less efficient for accelerating electrons than protons.[1]

---

[1] The highest energy electron/positron accelerator is the Large Electron Positron (LEP) synchrotron at CERN (European Organization for Nuclear Research) which reaches 100 GeV. The

The resonant rf cavities [33, 34] in synchrotrons (and many linear accelerators) are designed to operate at higher frequencies than the Alvarez type discussed above. The coaxial design is generally replaced by one or more nearly enclosed cavities. A small opening (iris) at each end of a rf cavity allows passage of the accelerated particles. The size of the iris and shape of the cavity is designed to insure the phase velocity of the enclosed electromagnetic wave is close to the velocity of the accelerated particles (essentially c in the Tevatron).[2] With this condition satisfied, there is no need for shielding with drift tubes.

Three synchrotrons provided the remaining energy boosts for the protons used by FOCUS. These are discussed below.

Booster: The Booster is about 500 feet in diameter and accelerates a group of protons to 8 GeV. Twelve such groups are used to fill the Main Ring. A proton goes around the Booster about 16,000 times to reach 8 GeV. The magnets used in the Booster combine the functions of bending and focusing in a single magnet (the Main Ring and Tevatron use dipole magnets for bending and quadrupole magnets for focusing).

Main Ring: The Main Ring (now decommissioned) and Tevatron share the same tunnel. This ring is about 4 miles in circumference. Conventional copper-coiled steel magnets are used by the Main Ring to guide the protons as their energy is increased to 150 GeV. The protons are then injected into the Tevatron.

Tevatron: In fixed target running, the Tevatron accelerates the protons to 800 GeV. The protons are held to a circular path by dipole magnets wrapped in

---

Large Hadron Collider (LHC) is being built in the same tunnel as the LEP collider. Protons are expected to reach 7 TeV in the LHC.

[2] In general, the phase velocity is also a function of the frequency.

superconducting niobium-titanium alloy wire, cooled to below 4.5 Kelvin. As the beam energy increases from 150 to 800 GeV, the magnets are ramped from 0.66 Tesla to 3.54 Tesla. The rf cavities used for accelerating the protons operate at 53 MHz, producing 18 ns spaced bunches of protons called “buckets” (about 1000 buckets are required to fill the Tevatron).

II.1.2 Proton Extraction

The Tevatron operates on a cycle of beam acceleration and extraction. During the 1996–97 fixed target run, protons were accelerated for 40 sec and then extracted slowly (to keep the intensity low) during a 20 sec “spill.” During extraction, the protons were sent to the “switchyard” where the beam was split among the Proton, Neutrino, and Meson areas. The beam in each of these areas was split further to the various experiments. The Wideband Photon Lab, where FOCUS was located, is in the Proton area.

II.2 Photon Production and Energy Tagging

The photons in FOCUS are produced from bremsstrahlung of electrons and positrons off nuclei in a lead radiator. The production of these electrons, and the removal of other particles from the beam involves several stages, illustrated in Fig. 9.

Magnets are used in various places to deflect or focus charged particles. Dipole sweeping magnets remove charged particles from the beam. Quadrupole magnets are used to focus charged particles. Neutral particles are removed by means of a neutral dump, with the electrons carried around the dump by means of dipole magnets.

The ultimate production of the photon beam can be traced in steps, as follows.

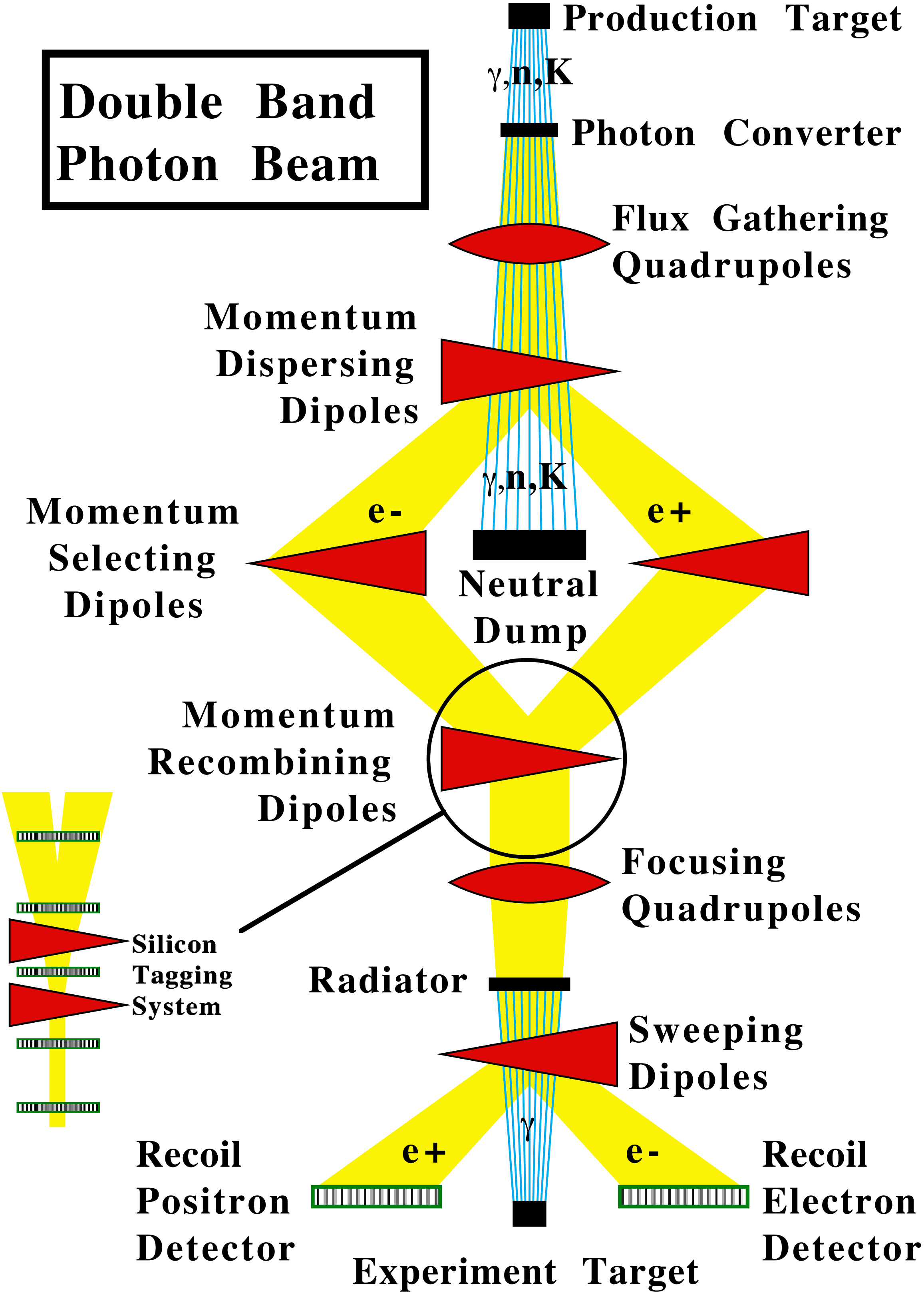


Figure 9: Layout of the tagging system.

A proton beam (800 GeV) from the main accelerator strikes the primary target made of liquid deuterium. The low charge to mass ratio of a deuterium nucleus maximizes hadronic interactions while minimizing electromagnetic interactions. This interaction produces a large number of hadrons (mostly pions), the charged ones being swept aside by magnets.[3] Neutral pions decay instantly and primarily into photon pairs. Some of these photons strike a lead converter producing e^+-e^- pairs in the region near a lead nucleus. The primary target is made of 3.4 meters of liquid deuterium which provides one proton interaction length. A longer target would result in an increase in e^+-e^- pairs produced in the primary target rather than the converter.

The neutral dump collects the remaining neutral particles (except for neutrinos, which rarely interact).[4] The conversion electron passes through two sets of dipole magnets. The electron path is bent with additional dipoles while passing through five planes of silicon microstrip detectors, enabling a determination of the electron's initial momentum ("tagging" the electron). Finally, the electron reaches the radiator (20% of a radiation length[5] of lead) to produce the final desired bremsstrahlung. Passing through the radiator, the electron is swept aside by magnets into a Recoil Electron Shower Hodoscope (RESH). The RESH consists of layers of Lucite scintillator and lead and is split into thirteen segments so that the direction of the deflected electron is measured in addition to its energy. (The central segment uses SiO_2 instead of Lucite, for greater radiation hardness.) The remaining energy is possessed by the photon,

---

[3] Many of these hadrons give rise to "halo" muons in the downstream experiment which must be "vetoed" (see Sec. III.10.1 for further discussion).

[4] Interactions in the neutral dump (largely from K_L^0's and neutrons) also give rise to halo muons (see above footnote).

[5] A "radiation length" is the mean amount of material needed to reduced the energy of an electron to $1/e \doteq 0.368$ of its initial value, where the energy loss comes from bremsstrahlung. Bremsstrahlung is the dominant mechanism for energy loss in high energy electrons.

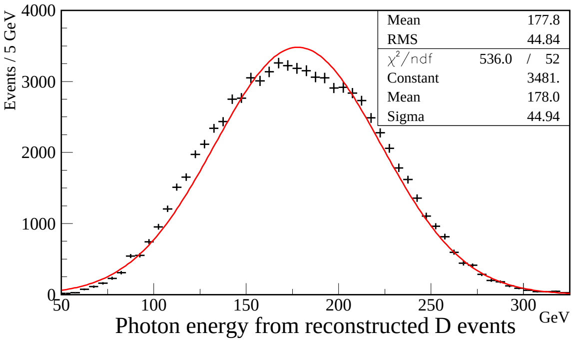


Figure 10: Photon energy from reconstructed $D^0 \to K^-\pi^+$, $D^0 \to K^-\pi^+\pi^+\pi^-$, and $D^+ \to K^-\pi^+\pi^+$ events. The smooth curve is a fit to a Gaussian.

which passes through the magnet undisturbed.[6] These photons continue on to be used by FOCUS. A similar process is followed by the positron branch.

Photons which do not interact in the charm production target are measured in a shower counter, the Beam Gamma Monitor (BGM) located at the end of the spectrometer. The BGM consists of 24 alternating layers of lead and SiO_2. The BGM, like the RESH, has a depth of 24 radiation lengths.

The mean photon energy for reconstructed D events is about 180 GeV (Fig. 10).

II.3 Charm Production Target

The beam photons can interact with one of four BeO target segments. The target material and thickness is chosen to minimize e^+-e^- pair production, multiple Coulomb scattering and re-interactions, while maximizing charm production and the number of charm decays outside of the targets. This requires a material with a small effective Z/A. Be was used for the initial 20% of FOCUS data and all of E687. BeO actually

---

[6] Multiple high energy photons are produced in a small fraction of the bremsstrahlungs. The thickness of the radiator was chosen to minimize the impact of multiple bremsstrahlungs on the performance of the experimental apparatus.

has a slightly higher effective Z/A than Be, but is 63% more dense than Be, so a smaller thickness of BeO is required for the same number of interactions. The smaller thickness allowed for silicon strips to be added within the target, closer to the primary interaction. Further details on this silicon and target layout are in Sec. III.1.

CHAPTER III

SPECTROMETER

The FOCUS spectrometer is illustrated in Fig. 11. The length of the spectrometer is roughly 31 meters from the most upstream target face to the most downstream muon filter face. The beam photons are incident on a segmented BeO target. Silicon microstrip detectors are located between the target segments and also downstream of the target. Five stations of proportional wire chambers and two separated, oppositely polarized, magnets provide momentum measurements for charged particles. These tracking and vertexing detectors provide FOCUS with a proper time resolution $\sim 8\%$ of D^0 lifetime. Three threshold Čerenkov counters are used to discriminate between protons, kaons, pions, and electrons.

For both muon and electromagnetic detectors there are inner and outer regions. The outer regions measure wide angle particles which miss the inner detectors. Inner muons are detected with an array of scintillating paddles with layers of iron serving as absorbers of hadrons and electrons. Outer muons are detected with resistive plate chambers with the downstream analysis magnet serving as an absorber. Two electromagnetic calorimeters identify electrons and photons. The inner electromagnetic calorimeter consists of lead glass scintillator; the outer electromagnetic calorimeter consists of plastic scintillator with lead layers.

The hadron calorimeter consists of iron absorber plates with scintillating tile. The hadron calorimeter is used in the first level trigger.

Each of these detectors are detailed below.

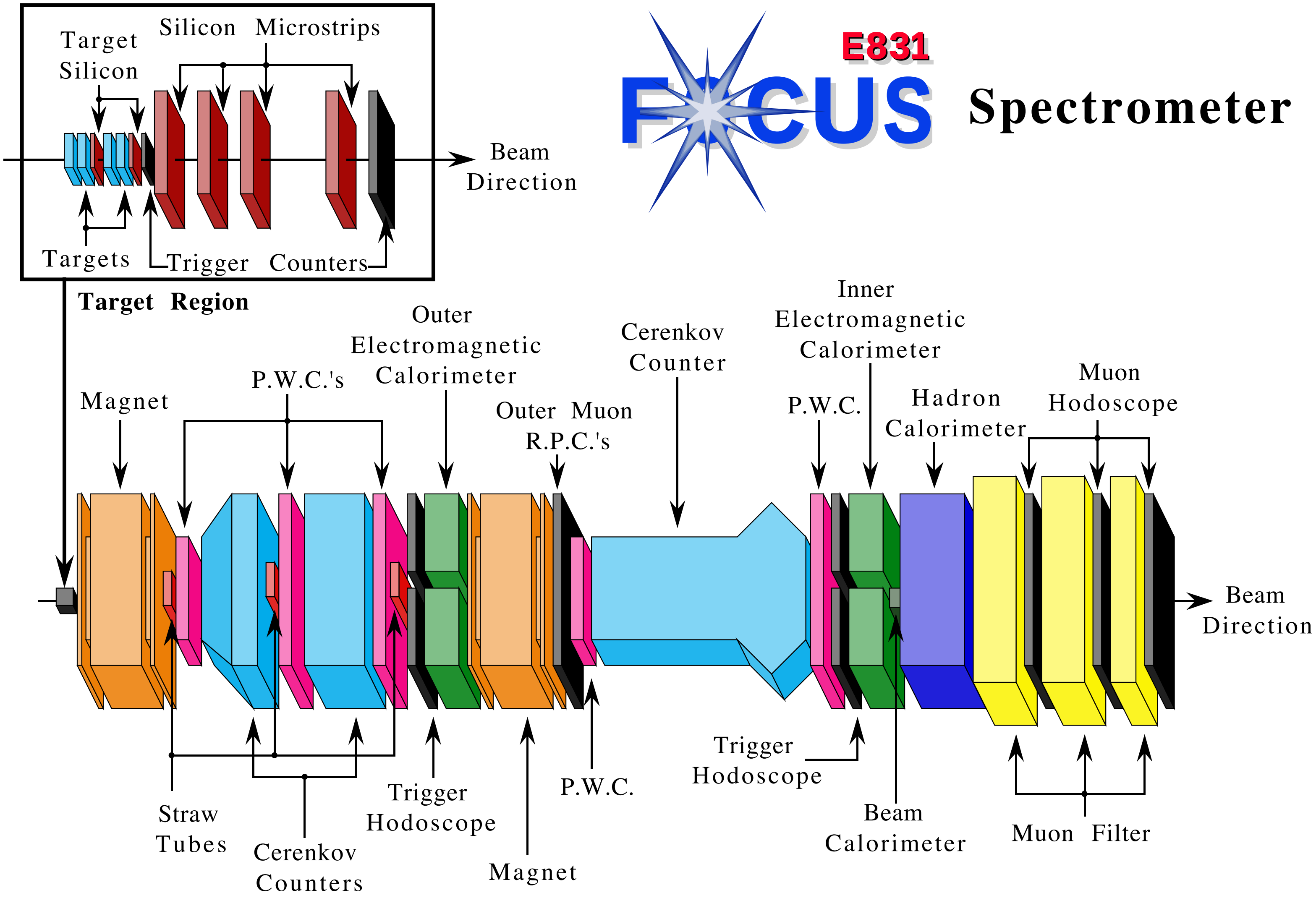


Figure 11: E831 Spectrometer. The length of the spectrometer is roughly 31 meters.

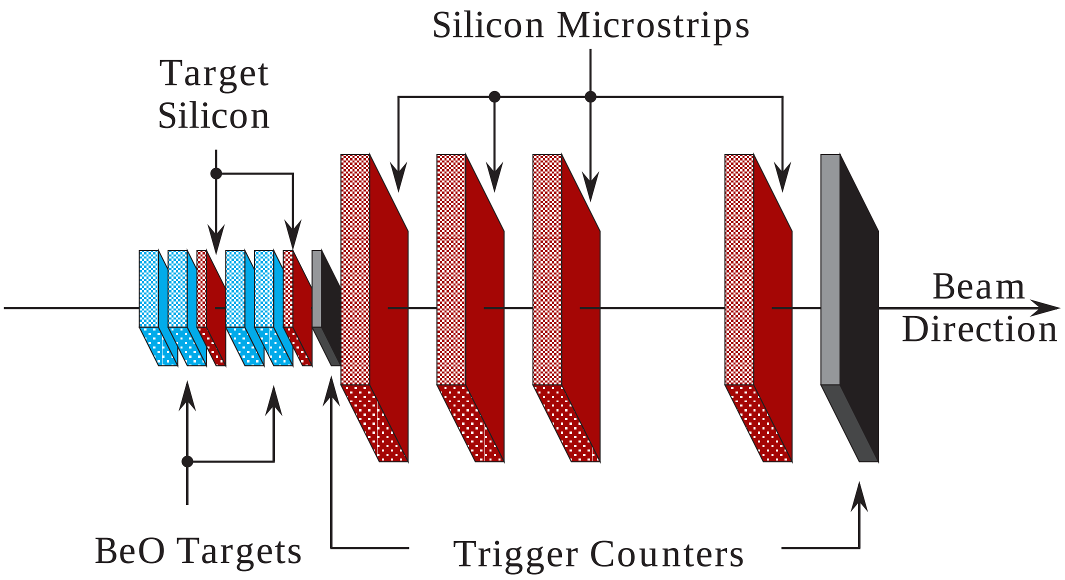


Figure 12: Target segments and silicon microstrip detectors.

III.1 Target and Silicon Microstrips

Tracking and vertexing in the charm production and decay region is accomplished with silicon microstrip detectors. Silicon microstrip detectors consist of parallel strips of reverse biased p-n junctions. The junctions establish an electric field which produces a current from electron-hole pairs liberated by the passage of an ionizing particle.

The target and silicon microstrip layout is shown in Fig. 12. Four BeO segments serve as targets for the photon beam. Two target silicon stations, containing two perpendicular views in each station, are located within and just downstream of the targets. Each view has a 25 micron pitch. The use of a segmented target and the addition of target silicon is an improvement over E687. The use of BeO instead of Be allowed more space for these modifications as well as increasing the cross section for charm production slightly (BeO is denser than the previously used Be). As well as adding measuring stations closer to the primary interaction, the use of a segmented target lessens confusion from hadronic interactions—charm frequently decays outside the target.

Just downstream of the target region are four more stations of silicon microstrip detectors (SSD's) with three views in each station [35]. Each view has a 25 to 100 micron pitch. The vertex resolution in z (along the beam) is about 300 microns; in x and y it is about 6 microns. This provides an average lifetime resolution of 0.028 ps for charged two-body decays, using the combined silicon microstrip and target silicon system.

A coincidence in the trigger counters on either side of the downstream microstrip detectors is one component of the first level trigger. This coincidence selects events originating from the target.

III.2 PWC System

Tracking downstream of the upstream analysis magnet is done with proportional wire chambers (PWC's). PWC's are gas ionizing detectors. As with other gas ionizing detectors (RPC's, proportional tubes, and Geiger-Muller tubes) a charged particle passing through the detector ionizes the gas, electrons migrate towards the anode (high voltage) and positive ions migrate towards the cathode (low voltage or ground). The corresponding current due to this migrating charge is measured. Secondary ionizations due to collisions between the accelerating electrons and ions create tertiary ions and so on, forming an avalanche which amplifies the signal exponentially.[1] The

---

[1] There are actually several ionization mechanisms involved. See Sec. 6.2 of Leo [36] for further details.

anodes and cathodes in PWC's consist of wires. Most of the avalanche is localized to the intense field region near an anode wire.

The proportional wire chamber system consists of 20 planes arranged in five stations with four views per station. The stations are labeled P0, P1, P2, P3, and P4, upstream to downstream. The views in each station measure YVUX, upstream to downstream. The U and V anode wires run 11.3 degrees from the y-measuring anodes (horizontal).

There are two types of stations. Type I (P0 and P3) have an active region of about 30 in. × 50 in. The anode wires are 0.8 mil diameter (1 mil = 1 thousandth of an inch) gold-plated tungsten. The anode wire spacing is 80 mils. Cathode wires are 2.5 mil diameter Cu-Be with a spacing of about 33 mils.

Type II stations (P1, P2, P4) have an active region of 60 in. × 80 in. The anode wires are 1.0 mil diameter gold plated tungsten. The wire spacing is 130 mils. Cathode wires are 3.5 mil diameter Cu-Be with a spacing of about 48 mils.

The gas mixture used was 75% Argon, 25% Ethane. Type I planes ran at 3.3–3.6 kV; type II ran at 2.8–3.2 kV.

III.3 Magnets

Two vertical-bending dipole magnets with opposite polarity were used for momentum determination. The magnets are identical in construction, apart from the shielding plates. The magnets consist of a 66 inch thick steel yoke wound with copper coils. The center of the coils contain a channel for water cooling.

The central field values were roughly 6.6 kGauss and 14.0 kGauss for M1 (upstream) and M2 (downstream), respectively. M1 had a momentum kick, p_t, of

0.400 GeV/c; M2 had a p_t of 0.837 GeV/c.

Tracks spread out by M1 are refocused by M2. The focal point is chosen to be near the inner-electromagnetic calorimeter. Electron-positron pairs are produced profusely in the target. These are swept into a vertical swath by M1. The pairs that pass the geometrical acceptance of M1 and M2 are refocused by M2. Several detectors contain a vertical gap to exclude the pair region. These include: the outer electromagnetic calorimeter (OE), the inner electromagnetic calorimeter (IE), the OH trigger hodoscope (just upstream of the OE), and the H×V trigger hodoscope (located just upstream of the IE).

III.4 Straw Tubes

Straw tubes are gas ionizing detectors (gas ionizing detectors were described briefly in Sec. III.2). For straw tubes, each anode wire is enclosed by a separate small cathode tube (straw).

Three straw tube wire chambers, a new addition for FOCUS, cover the pair region. One is located just upstream of each of the three most upstream PWC's. Each straw chamber contains three views—one x-measuring and two slant views ± 11.33 degrees from vertical. Each view contains three layers of tubes. The straws are 5 mm diameter mylar with an inner coating of copper. The central wire in each straw is 20 micron diameter gold plated tungsten. The gas, a 50% argon, 50% ethane mixture, flowed with a negligible positive pressure. The copper straw coating was held at ground; the central wire at ~ 1.6 kV.

Table 2: Čerenkov counter specifications [37].

Counter	Gas	Threshold (GeV/c)			Number 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

III.5 Čerenkov Counters

When a charged particle traverses a medium faster than the speed of light in the medium, a "shock-wave" of light is emitted in a cone similar to the shock-wave of sound emitted when jet airplane breaks the sound barrier. This shock-wave of light is referred to as Čerenkov radiation. The threshold speed for producing Čerenkov light is $\beta_t = 1/n$ where n is the index of refraction and $\beta = \text{speed}/c$. The half angle of the Čerenkov light cone is given by $\theta_c = \arccos(1/(n\beta))$. The number of photons produced per unit pathlength of the emitting particle is proportional to $\sin^2\theta_c$.

In FOCUS, three threshold Čerenkov counters provide discrimination between protons, kaons, pions, and electrons. Table 2 summarizes the properties of these detectors, including the momentum thresholds for producing Čerenkov light for various species and the average number of photoelectrons detected in a phototube for a Čerenkov cone contained within a cell (there is one cell for each phototube).

The Čerenkov counters in FOCUS are operated in "threshold" mode. The ADC information in each cell is only used to determine if a cell has detected light or not. For each particle hypothesis and counter, θ_c and the expected number of photons is computed. This information, together with the measured distribution of activated cells, is used to determine a likelihood for each hypothesis (Sec. IV.4.1 describes this in more detail).

The design characteristics of the three Čerenkov counters are briefly summarized below. A more detailed discussion can be found in [37] and [31].

III.5.1 C1

C1 is the most upstream Čerenkov counter. The gas mixture is 57% He, 43% N_2 (which differs slightly from the ratio used in E687). The length of the gas volume along the beam direction is 71 inches. There are 90 cells; each cell consists of a mirror and a phototube. Each phototube has a diameter of 2, 3, or 5 inches. The 2 and 3 inch tubes use collection cones to ensure complete light collection. A detailed discussion of C1 can be found in [38].

III.5.2 C2

C2 is located downstream of C1. C2 uses N_2O gas and the length of the gas volume is 74 inches. It has the lowest momentum threshold. There are 54 inner cells and 56 outer cells. The inner cells contain 2 inch phototubes; the outer contain 5 inch phototubes. The faces of the tubes are coated with the wavelength shifter, p-terphenyl. Each cell consists of a light collection cone and a phototube. The cells are actually mounted along the east and west sides of the structure. An array of mirrors direct the Čerenkov light into the cells. C2 contains a vertical gap in the region of high photon and e^+-e^- pair flux.

III.5.3 C3

C3 is the most downstream Čerenkov counter. It has the highest momentum threshold. C3 uses He gas and the length of the gas volume is 277 inches. There are 100 cells. Each cell consists of a focusing mirror, light collection cone, and phototube. The phototubes were coated with waveshifter p-terphenyl. The gaps between the phototubes and gas volume windows were flushed with N_2 to prevent contamination from He diffusion through the windows.

III.6 Inner Muon Detectors

The inner muon system uses three layers of steel to filter out electrons and hadrons. Stations of scintillating paddles are positioned just downstream of each steel layer. E687 used a combination of proportional tubes and coarse grained triggering scintillator. For FOCUS, the proportional tubes were replaced with much faster scintillator to greatly reduced false identification from the muon halo. The most upstream layer of steel/scintillator was also a new addition for FOCUS.

The most upstream layer of steel is 24 inches thick followed by x-measuring and y-measuring scintillator (MH1X and MH1Y). The next layer of steel is just downstream of MH1Y and is 24 inches thick. This is followed by more scintillator (MH2X and MH2Y). Downstream of this is the most downstream layer of steel (27 inches) followed by the final scintillator array (MH3UV) arranged in a slanted "V" configuration 30 degrees from horizontal.

III.7 Outer Muon Detectors

The outer muon (OM) system consists of an array of resistive plate chambers (RPC's) located between the downstream analysis magnet (M2)[2] and the P3 PWC (very little overlap with the RPC's). M2 serves as a filter of hadrons and electrons. Eight towers of chambers are arranged around the M2 aperture. Each tower consists of three chambers—an x-measuring, a y-measuring, and a slant view (45 degrees), for a total of 24 chambers. The 3 cm wide read-out strips were OR'ed into effective 12 cm wide strips in order to reduce electronics and cabling. Due to multiple Coulomb scattering in the magnet iron, 12 cm provided sufficient resolution.

Collaborators from Vanderbilt and Pavia were responsible for testing and installing the outer muon system as well as developing the read-out and trigger electronics. The author of this thesis assisted in this process, and had the primary responsibility for developing the OM reconstruction and simulation software.

The outer muon system in FOCUS completely replaced the proportional tubes and scintillator used in E687. RPC's are insensitive to magnetic fields and there are no phototubes requiring shielding. Since RPC's are operated in “streamer” mode (discussed below), no amplification of the signals was required before discriminating. In addition to quenching gas, the high resistance in an RPC limits the region of the ionization avalanche.

RPC's are gas ionizing detectors (described in general terms in Sec. III.2). The RPC's used in FOCUS were built by General Technica, in Colli, Italy. They contain two gaps with a single readout layer as shown in Fig 13. A double gap provides larger signals, as well as eliminating the slight inefficiency from the gap spacers (the spacers

---

[2] Actually, located downstream of the M2 mirror plate.

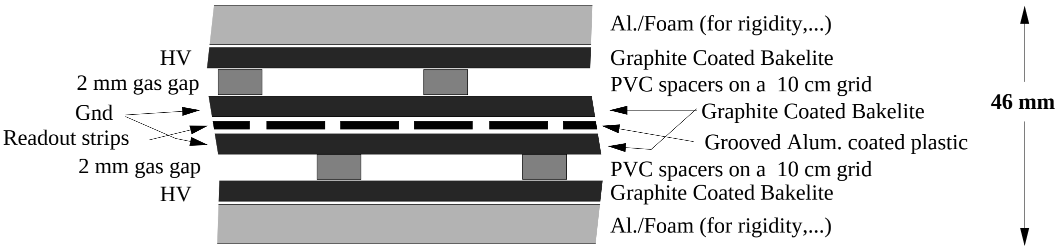


Figure 13: Cross section of a FOCUS RPC module (not to scale).

are staggered to prevent them from overlapping between layers). The high voltage ground planes are made of bakelite with a volume resistivity of about $10^{11}\,\Omega\,\mathrm{cm}$. The readout strips lie between the ground planes and consist of aluminum coated plastic. The outer faces of the high voltage planes are coated with conductive graphite to deliver a uniform voltage. The ground and high voltage planes were treated with linseed oil on the gap facing sides, ostensibly to smooth the surface and improve field uniformity [39]. The planes are encased in foam and aluminum and reinforced with stainless steel “u-channels.” Six of the eight towers used RPC’s which were 1 m across and 1.6 m high. The two towers on either side of the magnet aperture used 1.8 m high RPC’s.

If the applied voltage between the cathode and anode is high enough, a physical limit known as “breakdown” or the “Raether limit” is reached. Because of their high operating voltages ($> 4.5\,\mathrm{kV}$), breakdown is always reached in RPC’s. At these voltages, many avalanches would spread along the length of the detector if they were not prevented by a “quenching” gas. These non-localized avalanches are caused by UV photons which are emitted as electrons recombine with ions [36]. A quenching gas absorbs these photons and disperses their energy.

The measured current needs to be localized to be translated into a position measurement. In a RPC this is achieved by the extremely high resistance ($\sim 10^{11}\ \Omega$ cm) of the ground and high voltage planes. The charge removed from the ground and high voltage planes during a discharge cannot be restored over the duration of the discharge ($\sim$ 10 ns). The relaxation time of the plates is roughly 10^{-2} s. Roughly 0.1 cm^2 remains inactive during the relaxation time [40]. The rapid drop in potential difference in this small area helps in quenching the avalanche. The choice of gas also serves to quench the avalanche—isobutane or CO_2 captures UV photons emitted from electron-ion recombination. Freon, with its high electron affinity, also reduces the region of the discharge [40].

Two gas mixtures were used in FOCUS. The first was chosen for its low flammability and relatively low voltage required (5.8 kV) [41]. This mixture was: 71% Argon, 8% Isobutane, 5% Freon 13B1, and 16% CO_2. Using this mixture in normal running conditions, we found the pulse height distributions to have very long tails. These long pulses caused the RPC's to frequently draw too much current. About a quarter of the way through the run, we switched to using a more conventional gas mixture: 4% Freon, 42% Isobutane, and 54% Argon (8.1 kV operating voltage). This did not have the long pulses and high current draw.

The OM was also used for triggering. The OM produced two types of trigger inputs, indicating either at least one or two active towers (to signal one or two muons). For a tower to be considered active, a signal from at least two views was required within a tower. The two hit trigger required the two active towers be non-adjacent (to eliminate triggering on a single muon passing through two towers). The center top and bottom towers were not used for two hit trigger due to e^+-e^- pairs.

III.8 Electromagnetic Calorimeters

In a calorimeter, the energy of an incident particle is dispersed into a shower of lower energy particles which can be more effectively detected and contained. Near the start of an electromagnetic (EM) shower, bremsstrahlung and e^+e^- pair production are the dominant processes; near the end, ionization and Compton scattering dominate.

The inner and outer electromagnetic calorimeters detect electrons (more precisely EM showers) and photons. They are also used to reconstruct π^0s, which decay into photons.

III.8.1 Inner Electromagnetic Calorimeter

The Inner Electromagnetic Calorimeter (IE) consists of a layer of 802 lead glass blocks 60.2 cm deep (18.8 radiation lengths and 2.2 proton interaction lengths). Čerenkov light (Sec. III.5) produced within the blocks is detected by photomultiplier tubes. Each lead glass block has a 5.8×5.8 cm^2 square face. The detector is split in half by a 5.5 inch wide vertical pair gap. Each block is wrapped in aluminized mylar which reflects produced light into a photomultiplier tube attached to the end.

The glass blocks are of type F-2 manufactured by Schott Glass Technologies, Inc. Type F-2 glass is composed roughly of 45% silica, 45% lead oxide, 5% sodium dioxide, and 5% potassium oxide by weight.

For calibration purposes, light from xenon flash tubes was distributed to each glass block by fiber optics during each inter-spill. This calibration was checked and corrected throughout the run by reconstructing π^0's.

The IE was also used for triggering. Three types of trigger inputs were formed with the IE: a sum of the entire IE energy, a sum of the transverse energy, and a trigger to select $J/\psi \rightarrow e^+e^-$ decays. Energy sums are formed from six nearly square regions (sextants). Three sextants are on each side of the pair gap. The J/ψ trigger required about 20 GeV deposited in two non-adjacent sextants (sextants separated by the pair gap are considered non-adjacent).

III.8.2 Outer Electromagnetic Calorimeter

The Outer Electromagnetic Calorimeter (OE) detects electrons and photons in the outer region of the spectrometer. It consists of alternating layers of aluminum, lead (stiffened with 6% Sb (antimony) by weight), and scintillator. Its outer dimensions are $255 \times 205\,\mathrm{cm}^2$. This corresponds to an angular acceptance of $28 \leq |\theta_x| \leq 142$ mrad and $49 \leq |\theta_y| \leq 114$ mrad. A 9 cm wide vertical gap excludes the e^+e^- pair region.

The scintillator layers are made of strips arranged in horizontal, vertical, and slant views (45° and 135°). Extra layers were added for FOCUS and a scintillating tile "tie breaker" was added. The tie breaker is a single plane of 100 tiles. The x and y hits from two photons give four possible positions. The tie breaker, together with the slant views, help resolve this ambiguity.

The OE also adds 23 cm of equivalent iron absorber for the outer muon system.

III.9 Hadron Calorimeter

The theory of showers produced from hadronic interactions is not as well understood as the theory of EM showers. Descriptions of the shower evolution are largely empirical. In a hadronic shower, multiple hadrons are produced in each interaction within the shower. Many π^0's can be produced which gives rise to a significant EM component to the shower. The EM component can vary considerably. In addition, a variable fraction of the energy is invisible (not measurable). Invisible energy goes into breaking up absorber nuclei or into neutrinos or muons which escape the detector. A more extensive introduction to hadronic and EM showers, with references, can be found in [42].

To enhance selection of hadronic events over e^+e^- photon conversions, an energy requirement of ~ 20 GeV in the Hadron Calorimeter (HC) was included in the first level trigger. For this reason, the HC had to have a fast response, large acceptance ($\sim 100\,\mathrm{mrad}$), and fairly good energy resolution. The sampling gas hadron calorimeter from E687 was replaced with scintillating tile with a fiber readout in FOCUS. The existing HC iron structure was retained.

The HC consists of 28 iron plates interspersed with 0.7 cm thick scintillator planes. The first two plates are 6.4 and 5.1 cm respectively; the remaining plates are each 4.4 cm thick. This corresponds to a total of 7.8 proton interaction lengths and 72.7 radiation lengths. Each scintillator plane is formed from 66 tiles of different size (see Fig. 14).

Blue scintillation light was converted to green by wavelength shifting plastic fibers. The fibers in each tile were thermally fused to clear fibers which transmitted the light to photomultiplier tubes (PMT) via a square mixing block interface. The PMT outputs were sent to 15 CAMAC summer cards. The outputs of each summer were integrated. Four different integrators ran continuously, with two operating during any given clock cycle (the 18 ns accelerator rf).

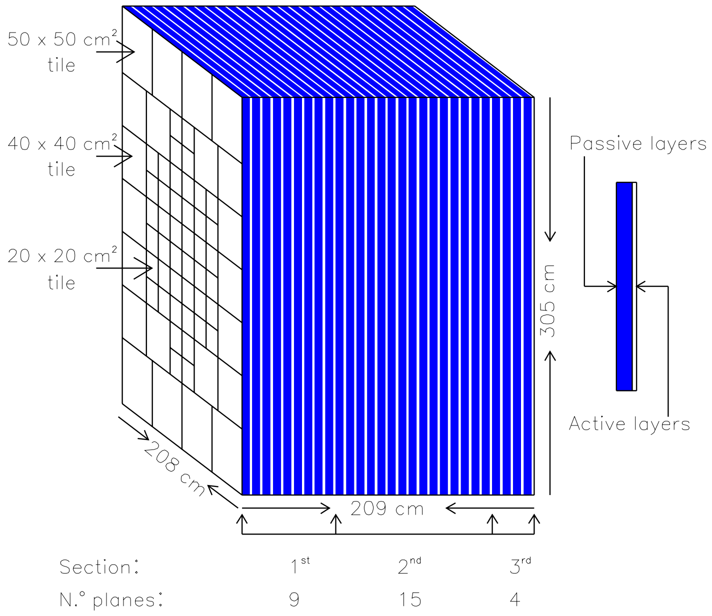


Figure 14: Hadron Calorimeter schematic.

Accounting for all time delays, a trigger signal from the HC was available for the first level trigger $0.34\,\mu$s after an interaction in the target. The HC trigger reduced the first level trigger rate by about a factor of 100 to about 3 kHz.

Two systems were employed to monitor the HC response: a ^{60}Co source and a laser system. Calibrations were performed with muons and pions. The measured energy resolution was $\sigma_E/E = 85\%/\sqrt{E} + 0.86\%$ (a significant improvement over the E687 HC ($\sigma_E/E = 132.6\%/\sqrt{E}$)).

III.10 The Trigger

A trigger is the logic used to select interactions with physics of interest and reject uninteresting background events. FOCUS had roughly 10^8 interactions per spill

(~ 20 seconds) and only triggered on about 3×10^4 per spill. The triggered events were mainly hadronic, the rest were mainly electromagnetic (e^+e^- and $\mu^+\mu^-$ pairs). FOCUS had two levels of triggers: the first level, or Master Gate (MG), and the second level. The MG makes a fast initial selection. If the MG criteria are satisfied, then the second level trigger starts. If the event passes the second level trigger, the event is read out to be eventually written to tape, otherwise the readout electronics are cleared and reset. It takes about 160 ns for signals from the spectrometer to reach the triggering electronics, with a MG decision made about 40 ns later. The second level trigger decision is made 1200 ns after the MG. Reseting takes about 1000 ns. It takes about $60\,\mu$s to readout an event, during which time no other events can be triggered.

III.10.1 Trigger Elements

The following is a discussion of the origin of the various trigger signals.

Triggering Scintillators

Several stations of scintillator arrays were used to trigger on or veto certain classes of events. These are described below.

TR1 is a 1.6 mm thick scintillator lying just downstream of the target silicon (Fig. 12). It is used to indicate an interaction in the target.

TR2 lies just downstream of the silicon microstrip detectors (Fig. 12) and indicates that a charged particle has entered the aperture of the M1 magnet. This counter consists of four separate pieces of scintillator arranged in quadrants with roughly 1 cm overlap at the edges (the intersecting corners at the center are chopped off so that the overlap is at most two layers). The signals from the four are combined with a logical OR. All of the "Master Gate" (first level) triggers require a coincidence of TR1 and TR2.

H×V is an array of 36 counters arranged in two layers with a conversion pair gap, positioned just upstream of the IE. "H" refers to the horizontal pitch layer, "V" refers to the vertical pitch layer. The output of the counters is fed into a fast logic module which produces two possible output signals: $(\mathrm{H}\times\mathrm{V})_1$ or $(\mathrm{H}\times\mathrm{V})_2$, indicating one or two charged particles detected respectively.

OH is a layer of 24 counters positioned just upstream of the OE. It contains a central gap the size of the M2 aperture as well as a vertical pair gap. The signal OH_1 from this array signifies at least one charged particle was measured.

IM1 & IM2 are located close to the MH (muon hodoscope) counters (Sec. III.6). IM1H is just downstream of MH2X; IM1V is just downstream of MH2Y; IM2H is just upstream of MH3UV. The IM counters cover a larger acceptance but have worse resolution than the MH arrays. The IM counters are only used for triggering. Two outputs are derived from the IM arrays—IM_1 and IM_2, signifying one or two muons detected.

AM & AMD are located just upstream of the charm production target, surrounding the beam. A coincidence in these two counters is used to veto muons arising from the primary production target (the target for the proton beam discussed in Sec. II.2) and the neutral beam dump. Since these "halo" muons are produced far upstream, they have very small angles in order to reach the charm production target region (Sec. II.3). A signal from these counters is used to veto certain muon triggers.

Triggers from Other Detectors

Various detectors formed a trigger signal based on an OR or sum of active detector subsystem outputs. Triggers were formed from the HC (Sec. III.9), IE (Sec. III.8.1), and OM (Sec. III.7) detector outputs.

III.10.2 The Master Gate

As discussed above, the Master Gate (first level trigger) makes an initial fast decision to begin reading out an event and start the second level trigger.

Information from the trigger elements listed above is fed to the Master Gate module [43], which outputs eight trigger types (MG1–MG8). Most of the data for this analysis comes through the hadronic MG trigger:

$$\mathrm{MG1} = \mathrm{TR1} \cdot \mathrm{TR2} \cdot \{(\mathrm{H} \times \mathrm{V})_2 + [(\mathrm{H} \times \mathrm{V})_1 \cdot \mathrm{OH}_1]\} \cdot \mathrm{E}_{\mathrm{HI}},$$

where the "·" denotes the logical AND and "+" denotes the logical OR. E_{HI} is summed HC energy greater than 18 GeV.

III.10.3 Second Level Trigger

In the second level trigger, additional requirements are made along with MG1: enough PWC hits for at least four tracks are required (MULT4) and some electromagnetic energy in the IE. The hadronic second level trigger is then:

$$\mathrm{TRIG1} = \mathrm{MG1} \cdot \mathrm{E}_{\mathrm{IE-2}} \cdot \mathrm{MULT4}.$$

III.11 Data Acquisition System

Data passing the second level trigger is saved to 8 mm tape for later analysis. The Data Acquisition System (DAQ) converts the various digitized detector signals into a serial format to be written to tape. It has to deal with a variety of input formats. These include some standard formats such as PCOS (Proportional Chamber Operating System) used by the PWC's and FERA (Fast Encoding and Readout ADC) used by the straw chambers, but generally each detector has a unique format.

Figure 15 illustrates the DAQ schematically. Signals from the various detectors are carried via coax or ribbon cables into modules in CAMAC[3] crates.[4] Various sorts of processing is done there depending on the detector, including time-to-digital conversion (TDC), analog-to-digital conversion (ADC), compression, discrimination, latching, etc. The modules produce ECL[5] outputs.

Each ECL signal is sent to a DYC+[6] [44] which packs the 16 bit words received into 32 bit words. The DYC+ buffers a sub-event[7] of data. When a DYC+ or the FSCC receives a "token," the sub-event is transferred to the DDD via the RS-485 bus.[8] The DYC+ then sends a token to the next DYC+, to allow it to release its buffered sub-event to the RS-485 cable. Tokens are sent along in a circular fashion. DDD is an acronym for three modules which receive, buffer, and sequence the input from the RS-485. The result is sent along a VME bus to a SGI Challenge L workstation. The

---

[3] Computer Automated Measurement and Control—a widely used electronics standard developed by the European Standard of Nuclear Electronics.

[4] Signals from the calorimeters and part of the output of the Čerenkov system bypass the CAMAC and are digitized by Fastbus ADC's.

[5] Emitter-Coupled Logic handles very fast signals, but with close logic levels: -0.9 V and -1.75 V.

[6] Damn Yankee Controller—designed by the Fermilab Physics Department.

[7] A sub-event is the digitized and encoded data generated by a single detector system (or by part of it) in response to a trigger signal.

[8] The RS485 standard allows for up to 10 Mb/s with up to 32 drivers and receivers.

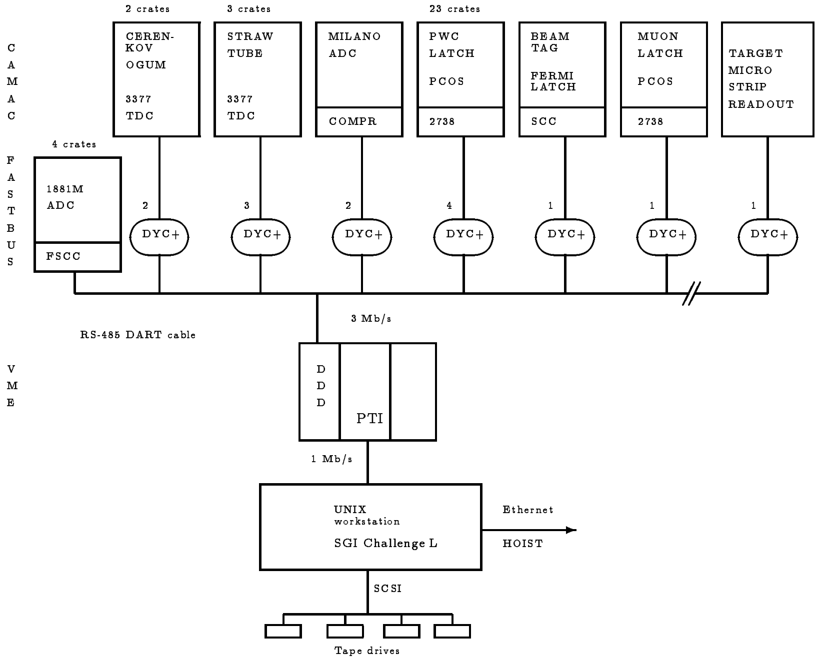


Figure 15: Schematic of the DAQ layout.

workstation buffers the data on disk and later writes it to tape.

The DAQ logged 1500 to 2000 events per second, with a typical event size of 4 kB. The livetime was typically 85–90%. For more details on the DAQ see [45].

CHAPTER IV

DATA RECONSTRUCTION

In this chapter, we discuss general reconstruction and data reduction techniques used in FOCUS. The emphasis will be on techniques that are either typically used in every FOCUS analysis or that are pertinent for semileptonic analyses.

IV.1 Tracking

Tracks in the SSD and PWC systems (Secs. III.1 and III.2) are initially reconstructed separately (although the SSD is used to "seed" the search for tracks in the PWC). Keeping the two systems separate proved useful for identifying and measuring relative position offsets between the SSD and PWC regions.

The tracks from the two systems are then "linked." Linked tracks generally consist of long-lived particles (electrons, muons, pions, protons, kaons). Unlinked tracks arise from particles that either decay to longer lived particles, or which leave the spectrometer acceptance before reaching the PWC's. A typical charm event contains a "primary" vertex where the photon-nucleon interaction produces the charmed particles (typically a charm containing and an anti-charm containing particle) along with a number of other particles (five charged primary tracks is typical), mostly pions. Although they have no momentum measurement, unlinked tracks are included in the primary vertex to further pin down the position of the primary.

IV.1.1 SSD Tracks

The formation of SSD tracks proceeds in three stages. First clusters are formed, then clusters are used to make separate projections in each view (measurement direction), and finally the projections are combined into tracks. Only the four silicon microstrip stations downstream of the target are used for track finding. The target silicon (TSSD) is used later to improve track and vertex parameters.

Groups of up to three active strips are used to form clusters. The integrated charge collected by the clustered strips, as determined from ADC (analog-to-digital converter) counts is required to be consistent with the expected charge for a single minimum ionizing particle (MIP). By weighting the position of each strip in a cluster by its ADC counts, the centroid of the cluster is determined (this approach is referred to as "pulse height sharing"). Projections are created from the cluster positions.

Hits in at least three of the four planes in a given view are required to form a projection. Each projection is required to be consistent with a line with a χ^2/dof (chi-square per degree of freedom) less than 3. A group of projections from the three views is considered a track if they are consistent with a line with a $\chi^2/\mathrm{dof} < 8$. Tracks with shared projections are arbitrated by using the one with the lowest χ^2/dof. Tracks with nearly identical parameters are reduced to a single track.

The efficiency for reconstructing SSD tracks increases with increasing momentum. The resolution also improves with increasing momentum, since there is less multiple Coulomb scattering. For a track traversing the central region of the SSD's, the resolution of the intercept of a SSD track projected back to the center of the target in

E687 was

$$\sigma_x = 11.0\,\mu\text{m}\sqrt{1+\left(\frac{17.5\,\text{GeV}/c}{p}\right)^2} \text{ and } \sigma_y = 7.7\,\mu\text{m}\sqrt{1+\left(\frac{25.0\,\text{GeV}/c}{p}\right)^2}, \quad (20)$$

where p is the track momentum and $11.0\,\mu$m and $7.7\,\mu$m account for the SSD strip granularities. Tracks traversing the outer portion of the SSD have a resolution twice as large. Instead of using the centroid for the the cluster position, E687 used the position of the strip with the most ADC counts. The centroid approach provides a roughly 10% improvement over the errors in Eq. (20). In addition, the target silicon typically improves the resolution by a factor of 1.2–1.5.

IV.1.2 PWC Tracks

The search for PWC tracks is "seeded" with the x-measuring components of the SSD tracks (the magnets bend trajectories predominantly in the y direction). Hits in the PWC x views are matched to the extrapolated SSD tracks. Projections in the y, u, and v PWC views are then combined with the x projection to form tracks. Unused x hits are then used to form more projections for additional tracks.

A number of requirements are made on tracks. The x projections found from SSD tracks must have hits in P0. A track must have hits in at least three chambers. There can be no more than four missing hits, with a maximum of two missing in a single chamber. A least squares fit is performed on tracks to determine track parameters (slopes, intercepts, and M2 magnet bend) and χ^2/dof. Various corrections are made to the least squares fit to account for the full spatial variation of the magnetic fields.

Tracks with hits in all five chambers are confusingly called "tracks" and tracks with hits only in the first three chambers (before M2) are referred to as "stubs."

Tracks terminate in the inner region of the spectrometer; stubs terminate in the outer region.

There are other categories of tracks which do not have to meet the same restrictions as those discussed above. These include four chamber tracks which miss P3 or three chamber tracks missing both P0 and P3 (the smaller chambers). The later generally arise from halo muons, which are useful for various studies. Also, wide angle tracks with hits only in P0 and P1 are seeded with microstrip tracks. Tracks from particles which decay upstream of P2 need not have hits in P0. These include "vees" (vertexed tracks) and "kinks" (charged particles which decay into a charged particle and a neutral particle). A discussion of vees and kinks can be found in Secs. 4.1.5 and. 4.1.6 of Ref. [31].

Events containing more than 30 tracks or 600 PWC hits are discarded since these take a long time to reconstruct. This discards a few percent of the events.

IV.1.3 Linking

In order to associate a momentum with a SSD track, the SSD track must be "linked" to a PWC track. The slopes and intercepts of SSD and PWC tracks are required to be consistent at the M1 bend center. If this is the case, a global least squares fit is made using both PWC and SSD hits. The χ^2/dof resulting from this fit is used to arbitrate links. A maximum of two PWC tracks is allowed to be linked to each SSD track. Excluding double links reduces e^+e^- pair contamination (the initially close pairs are split apart by the M1 magnetic field).

IV.2 Momentum Determination

Momentum is computed from the bend induced by the two analysis magnets. The bend is found by comparing track parameters upstream and downstream of each magnet.

The momentum of five chamber tracks is determined from M2 (which has a higher field and therefore larger bend than M1). The track parameters and magnetic field are input to the fit. The computed momentum and improved track parameters are fit iteratively.[1]

The momentum for linked stubs and four chamber tracks is determined from SSD track parameters and the PWC track parameters between M1 and M2, and the magnetic field of M1. Again, the computed momentum and track parameters are fit iteratively.

Unlinked stubs have no SSD track, so the origin of the stub is approximated to estimate the momentum.[2] Unlinked tracks and unlinked stubs are not used in the analysis in this thesis.

In E687, the momentum resolution for stubs was

$$\frac{\sigma_p}{p} = 0.034 \times \frac{p}{100\,\mathrm{GeV}/c}\sqrt{1+\left(\frac{17\,\mathrm{GeV}/c}{p}\right)^2} \tag{21}$$

and the momentum resolution for tracks measured in M2 was

$$\frac{\sigma_p}{p} = 0.014 \times \frac{p}{100\,\mathrm{GeV}/c}\sqrt{1+\left(\frac{23\,\mathrm{GeV}/c}{p}\right)^2}. \tag{22}$$

Slight changes were made in some of the PWC positions and wire spacings for FOCUS.

---

[1] For five chamber tracks, the momentum is also separately computed from the M1 bend. The energy lost due to bremsstrahlung of electron candidates is less in the weaker field of M1. In this case, M1 can give a more accurate determination of the initial momentum.

[2] To approximate the origin, the x-projection of the stub is extrapolated back to the target region and the nearest vertex is selected as the origin. If there is no vertex, the center of the target material is used as the origin.

IV.3 Vertexing

In FOCUS, the calculation of vertices is made with a software package called `DVERT`. The search for vertices is "candidate driven," meaning tracks which pass some requirements decided by the user are selected as candidates for coming from the same vertex. The candidates are "booked" in `DVERT`. Based on a χ^2 calculation, `DVERT` then computes a vertex position for the selected tracks, and a confidence level for this vertex hypothesis. This "Dee" confidence level (or DCL) is usually required to be greater than 1% (the user decides the level appropriate for his or her particular analysis). Information from the target silicon is incorporated in this calculation.

`DVERT` also contains specialized routines for locating primary (production) vertices as well as various algorithms for determining the degree of isolation of a vertex from other decays. The algorithms used in this analysis are detailed below.

IV.3.1 Primary Vertex

The algorithm most commonly used in FOCUS to locate primary vertices in semileptonic analyses is called `DVFREE`. `DVFREE` is called a "free form" vertex finder to distinguish it from candidate driven vertex finders. When a D^0 decays semileptonically, the full energy of the D^0 cannot be computed from the D^0 decay daughters, so the D^0 momentum direction is not available to point back to the primary. However, the candidate D^0 daughters, which form a "secondary" vertex, are excluded from consideration as primary tracks.

`DVFREE` works as follows. A seed track is selected arbitrarily. Then tracks are combined with the seed track to form a vertex. A track is kept in the vertex if the confidence level of the vertex remains above 1%. The tracks are not accumulated in any particular order. When all tracks have been tried, a new seed track is selected from among the tracks that haven't been included in a primary and the process is repeated to create another candidate primary vertex. Each track can appear in any number of primary vertices, except for a seed track which can only appear in the primary that it spawned. This procedure is repeated until no more seed tracks are available.

From among the list of candidate primaries, some criteria is used to select one. A common choice is to select the primary containing the most tracks. This is what is done for this thesis analysis. Ties are arbitrated by selecting the most upstream of the tied primaries. A new set of primaries is found for each secondary (D^0 candidate), since this changes which tracks are excluded from the primary.

IV.3.2 Vertexing Cuts

Detachment and isolation variables are among the most powerful tools for rejecting non-charm and combinatoric backgrounds. The optimal values of cuts on these variables depends on the particular analysis. Consequently, the most generic data reduction stages ("skims" discussed in Secs. IV.5.2 and IV.5.3) usually have very loose detachment requirements and no isolation requirements.

The detachment between a primary and secondary vertex is quantified by L/σ_L, where L is the distance between the two vertices and σ_L is the error on L. Since charmed particles tend to be long-lived, a cut on this variable is highly effective at rejecting short-lived non-charm backgrounds.

A related quantity is the number of sigma outside of the targets of the secondary vertex position. Most charm produced in FOCUS decays outside the target material, so this is an effective variable for eliminating backgrounds arising from interactions in material. It is not used at the skim level.

Once modest particle identification and detachment requirements are made, the isolation variables are the most effective cut variables in this analysis. They are evaluated as the maximum confidence level (CL) that additional tracks are consistent with a vertex. Therefore, cuts on these variables eliminate high CL's. The three isolation variables used in this analysis are described below.

`Iso1` cuts require the secondary candidates be inconsistent with coming from the primary vertex. The CL of the primary is recomputed with each secondary track included in the primary. The highest resulting CL is the value of `iso1`.

`Iso2` cuts require all tracks except the secondary candidates and tracks from the primary be inconsistent with coming from the secondary vertex.[3] The secondary and primary tracks are added to an exclusion list. Each track not in the exclusion list is separately added to the secondary, and the CL of the secondary (DCL) is recomputed. The highest resulting CL is the value of `iso2`.[4]

`Iso3` is the same as `iso2` except rather than adding all tracks from the primary to the exclusion list, only one of the tracks from the primary is excluded (e.g., the pion candidate in $D^{*+} \to D^0\pi^+$).[5] This is a more restrictive cut than `iso2` and is used in this analysis only to arbitrate among different π^+ candidates for the decay $D^{*+} \to D^0\pi^+$.

---

[3] Tracks from the primary tend to point roughly towards the secondary vertex, so isolating them from the secondary would lead to an overly harsh cut.

[4] This is sometimes labeled `iso2ex`, to distinguish it from the case where primary tracks are not included in the exclusion list.

[5] A note to FOCUS collaborators: this is not the same "`iso3`" in `SEZDEE`.

IV.4 Particle Identification

Information gathered from the detectors described in Secs. III.5–III.8 is used to identify the type of particle (electron, muon, pion, kaon, or proton) associated with each track. The algorithms used to identify these particles are described below.

The Čerenkov detectors are used to distinguish between electrons, pions, kaons and protons. The masses of muons and pions are too close for the Čerenkov to be useful in distinguishing them, so dedicated muon detectors are used for muon identification. The momentum thresholds for the Čerenkov detectors prevents them from distinguishing between electrons and pions at high momentum (above about $8.5\,\mathrm{GeV}/c$ for stubs and $17.4\,\mathrm{GeV}/c$ for tracks). The electromagnetic calorimeters are needed for electron identification at these higher energies.

IV.4.1 Čerenkov

The FOCUS Čerenkov identification algorithm [46] is called `CITADL` (Čerenkov Identification of Tracks by an Algorithm using Digital Likelihood). The algorithm computes a firing probability for all cells within a tracks $\beta = 1$ Čerenkov cone. For an expected number of photoelectrons, μ, produced for a given particle hypothesis and momentum, the likelihood that a cell will fire is given by $(1-\exp(-\mu))$. An accidental firing rate, a, is also incorporated into the likelihood (a is often proportional to the beam intensity). The total likelihood for a cell to be on is then $(1-e^{-\mu})+a-a(1-e^{-\mu})$, while the likelihood for a cell to be off is $1-[(1-e^{-\mu})+a-a(1-e^{-\mu})]$.

In analogy with a χ^2, the sum over cells of the log-likelihoods for a given particle hypothesis, i, is combined into a variable $W(i) \equiv -2\log(\text{likelihood})$. An example of how this is used is with the cut $W(\pi) - W(K) > 1$, which requires that the kaon hypothesis be preferred over the pion hypothesis.

IV.4.2 Electron Calorimetry

The electromagnetic calorimeters contain most of the electromagnetic energy passing through while capturing little hadronic energy. Thus for an electron entering the calorimeter with a momentum of p and a measured deposited energy of E, $E/p \approx 1$. For the IE, a requirement of $0.8 < E/p < 1.2$ is made to identify electrons. This window is slightly wider for the OE. Confirmation from the Čerenkov detectors can be used to enhance electron identification with the IE.

The OE algorithm combines information from E/p, Čerenkov, and the shower evolution as a function of z into a discriminant score.

Electron identification with the IE or OE is not used in this thesis analysis. Semi-electronic modes could potentially double the statistics used in this analysis, however, they are beyond the scope of this thesis.

IV.4.3 Muon Identification

Inner Muons

Muons are identified in the inner muon system (described in Sec. III.6) by computing a confidence level that a track is consistent with hits in the MH (muon hodoscope) planes. The confidence level is based on a χ^2 calculation which incorporates errors due to multiple Coulomb scattering in the steel filters and granularity of the MH paddles. The highest confidence level that a second track shares the same hits is also computed to serve as an isolation variable.

For muon candidates with momentum greater than 10 GeV/c, at least four of the six MH planes are required to have hits. At lower momentum, muons can be absorbed in the steel, so only two planes are required to have hits. These lower momentum inner muons candidates are typically cut out in most analyses in order to eliminate backgrounds.

Outer Muons

Muons are identified in the outer muon system (Sec. III.7) by computing a confidence level that a track is consistent with hits in the RPC's. Like the inner muon confidence level, the outer muon confidence level is based on a χ^2 incorporating errors from multiple Coulomb scattering and detector granularity. The calculation of correlations between the hits in the measuring stations is simplified by having all of the RPC's at the end of the filters. However, the magnetic field within the M2 steel must be accounted for in projecting the trajectories of tracks to the RPC's. Also, a more detailed treatment of energy loss is needed since outer muons are considerably lower momentum. The algorithm treats muons down to 4 GeV/c which is the lowest momentum at which muons will penetrate through the OE and M2.

The outer muon analysis and simulation code was primarily developed by the author of this thesis. A more complete description of this work can be found in Appendix B.

IV.5 Data Processing

FOCUS recorded roughly 6.5 billion photon interactions on 6000 8 mm tapes. This corresponds to about 25 terabytes of information. In order to make this more manageable for individual analyses, the data was reduced and split in three stages. In the first stage, “Pass1,” all the data was reconstructed. In the second stage, “Skim1,” events were split into broad physics topics. In the final stage, “Skim2,” the data was further divided into subtopics.

IV.5.1 Event Reconstruction (Pass1)

Pass1 reconstructed the raw data and wrote the output to another 6000 8 mm tapes. Pass1 ran from January 1998 to October 1998.

Since each event (photon interaction) recorded on a tape is independent, the events need not be processed sequentially. Rather than running a single fast (expensive) process, it is more efficient to have separate CPU’s process different events. For Pass1, the software that coordinated this division of labor was the Fermilab product CPS (Cooperative Process Software).

Each Fermilab computer “farm” which utilizes CPS consists of a server node and about ten worker nodes. The server node runs software for reading and writing data as well as sending events over a high-speed network to worker nodes. FOCUS used up to eight farms, but the availability varied considerably since the farms were shared with other experiments. More details on Pass1 processing can be found in [47].

Pass1 reconstructed data for all the detectors. Pass1 also wrote out raw data so that improvements and corrections to the reconstruction code could be made without

Table 3: Skim1 superstreams and Skim2 processing institutions. Skim1 output tapes served as input to Skim2. Skim2 split the Skim1 superstreams into smaller substreams.

Superstream	Description	Skim2 Institution
1	Semi-leptonic, di-leptonic	Puerto Rico
2	Topological vertexing and K^0_S	Illinois
3	Calibration and rare decays	CBPF, Brazil
4	Baryons	Fermilab
5	Diffractive (light quark states)	UC Davis
6	Hadronic meson decays	UC Davis

going back to the raw data tapes. In order to fit the Pass1 output on the same number of tapes as the raw data, raw ADC information was compressed. Also reconstructed calorimetry was not written out, events with reconstruction errors were discarded, and some very minimal selection cuts were made. About 10% of the events were discarded by Pass1.

IV.5.2 Event Selection (Skim1)

Skim1 divided the files output by Pass1 into six streams (denoted as superstreams to distinguish them from Skim2 output streams). The physics topics targeted in each of these superstreams are summarized in Table 3. Each superstream was written to 200–500 8 mm tapes, for a total of about 2500 tapes. Roughly half of the events in Pass1 survived the more restrictive cuts of Skim1. Summaries were saved rather than raw data. Calorimetry reconstruction was re-run. Improvements were made to the vee and Čerenkov reconstruction algorithms, so they were re-run in Skim1 as well.

In Skim1, computer clusters also processed data in parallel. Each CPU used in Skim1 processed an input file containing about 40,000 events before reading another input file. This differs from Pass1, where a worker node would process about 80 events at a time. In Skim1, input files (from Pass1 tapes) were dumped to disk with tape stackers. Each node (CPU) in the cluster would process one input file and output six files (one for each superstream) to disk. When enough output files were accumulated, they would be written to tape (one tape for each superstream).

Skim1 was primarily processed on computer farms at Vanderbilt and the University of Colorado at Boulder. At each institution, approximately 3000 input Pass1 tapes were processed on 20–30 workstations/PCs. Skim1 ran from March 1998 to April 1999.

IV.5.3 Final Splitting (Skim2)

The hundreds of tapes in each Skim1 superstream are still unwieldy to analyze. Skim2 split each superstream into 5–12 substreams. Tapes from each superstream served as input for processing at Skim2 institutions (indicated in Table 3). Skim2 processing was similar to Skim1. For Skim2, most of the skims used control software called the Generalized Skim Framework.

Cuts applied in Skim2 reduce the number of tapes to a manageable level but are loose enough to optimize particular analyses. The Skim2 substream used in this thesis analysis was the semimuonic substream, which contains 26 tapes. Requiring a primary and secondary vertex with $\mathrm{CL} > 0.01$, $L/\sigma_L > 3$, and $W(\pi) - W(K) > 1$ reduced this down to four tapes.

CHAPTER V

MONTE CARLO SIMULATION

This chapter describes the FOCUS Monte Carlo simulation, MCFOCUS. The shapes used in the fit for r_{mix} in this analysis come entirely from MCFOCUS. Validation of MCFOCUS is largely the topic of Chapters VI and VII.

V.1 MCFOCUS

MCFOCUS consists of three parts: beam simulation, event generation, and spectrometer simulation. (The same reconstruction code used with the data is used with MCFOCUS.) In event generation, particles produced from the photon-nucleon interaction are simulated. These particles are decayed and traced through the spectrometer. Particle interactions in each detector are simulated.

V.1.1 Beam Simulation

The positions and energies of the positrons and electrons incident on the radiator (Sec. II.2) are read from a library produced by a detailed simulation of the beamline. The detailed beamline simulation was generated using the the Fermilab program `TURTLE` (Trace Unlimited Rays Through Lumped Elements). Additional components are added to the beam library trajectories which incorporate position and energy smearing, acceptance, and run dependent positions.

The electron or positron is showered in the radiator. In the radiator, the electron can scatter or emit bremsstrahlung, and the photons may re-interact to produce pairs[1]. Photon energies are selected with a rejection method. The photons are stepped through the target segments where they can convert, shower, or produce charm according to the charm cross-section used by `PYTHIA` (discussed below).

V.1.2 Event Generation

Particles arising from the photon-nucleon interaction are simulated using `PYTHIA` [48] version 6.127. `PYTHIA` has a large number of adjustable parameters. These include (but are not limited to) choices of various baryon production models, amount of energy shared by the remnant baryon, different spin counting rules, parameters describing the string, cross sections for various processes, upper limits on momentum transfers, and charm quark masses. These parameters were tuned to give good agreement with FOCUS data in secondary momentum, secondary transverse momentum, primary vertex multiplicity, and charm-anticharm asymmetries in number and average momentum.

In order to save time generating simulated FOCUS events, `PYTHIA` is set up to only generate the photon-gluon (c-$\bar{c}$) process illustrated in Fig. 16.[2] This is the dominant mechanism for charm photoproduction. A completely generic "minimum bias" generation would be on the order of 100 times slower. The trade-off is that for some analyses reconstructed events have to be subjected to harsher cuts to obtain good agreement between data and Monte Carlo.

The particles generated using `PYTHIA` are then traced through the spectrometer and decayed. The particle trajectories, decays, and detector responses are all

---

[1] Photons that convert to pairs in the radiator are removed from the simulation.

[2] There is also an option for directly producing J/ψ particles, but this is run separately.

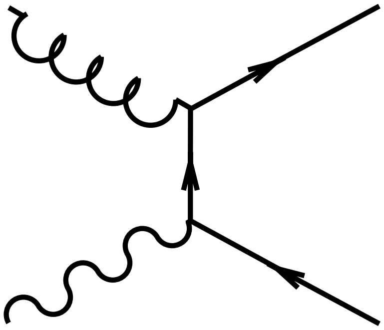


Figure 16: Feynman diagram for the photon-gluon fusion process. The gluon comes from a nearby u or d quark in the nucleon. The solid lines represent a c and $\bar{c}$. This is the dominant mechanism for charm photoproduction.

simulated using code written by the E687 and FOCUS collaborations called `ROGUE`. Particle properties and decay modes are specified in a particle "dictionary" file [49]. The decay modes are grouped into long lived "final states," which are reached by a number of decay paths. If a final state has been studied sufficiently to account for quantum interference between decay paths, the decay paths are replaced by a matrix element. Also if the form of a decay is known, such as pseudoscalar[3] to pseudoscalar lepton neutrino for $D^0 \to K^- \mu^+ \nu$, then a matrix element is used to correctly model the angular distribution of the decay.[4]

V.1.3 Spectrometer Simulation

Multiple Coulomb scattering, elastic, and inelastic scattering are simulated as particles are stepped through the spectrometer. For electrons bremsstrahlung is simulated. Photon conversions are also simulated. For inelastic scattering, the particle

---

[3] "Pseudoscalar" means the particle has a spin of zero with a negative parity quantum number.

[4] MCFOCUS also allows the user to specify particular decay paths be present in every event [49]. This is accomplished by rejecting generated events which do not contain the specified paths. This facility was not used in this analysis.

is destroyed producing a hadron shower in the process.[5]

Multiple Coulomb scattering is simulated with the distributions derived by Molière. This includes corrections to the Gaussian approximation. The author of this thesis developed a fast algorithm to implement Molière scattering in `ROGUE`. Further details of this work are presented in Appendix A.

Each detector has a separate simulation for determining the detector's response to various particles, written by experts on that particular detector. Sometimes this involves analytic approximations. In other cases this involves parameterizing calibration data and/or parameterizing more detailed simulations (e.g., the general purpose but very slow detector simulation program `GEANT`). Showers in calorimeters are frequently modeled by sampling a library of representative events, since detailed simulations are expensive in CPU time.

It is beyond the scope of this thesis to describe the details of each detector simulation. However, the muon simulations will be described in the following section, since this is particularly relevant for this thesis analysis and the author of this thesis was largely responsible for the outer muon simulation.

V.2 Muon Detector Simulations

Essentially all hadrons are absorbed in filters for the inner muon (IM) and outer muon (OM) detectors, so hadronic punch-through is not simulated.[6] The IM system has roughly 21 nuclear interaction lengths from the inner electromagnetic calorimeter,

---

[5] The shower is created by converting a Feynman-Field jet [50] into hadrons. This produces mostly pions and rhos.

[6] Knock-on electrons (also called delta rays) are not simulated either. These arise from electrons kicked out of material by a high energy particle. These can escape if produced near the surface of the filters.

hadron calorimeter, and additional steel filters. The OM system has roughly 18 nuclear interaction lengths arising from the outer electromagnetic calorimeter and the second magnet (M2).[7]

The muon simulations assume all particles incident on the filters are absorbed except for muons. Gaussian approximations are used for multiple Coulomb scattering (MCS) and statistical variations in energy loss (energy "straggling") are ignored. For inner muons, energy loss and MCS are simulated as the muons are stepped through cuboids (rectangular blocks) of material.

A similar procedure is followed for outer muons, except there is the added complication of the magnetic field inside M2. Also, since outer muons have lower energies than inner muons, a more careful accounting of the effect of energy loss on MCS is required.[8] The approach used is to trace the trajectory of a muon through cuboids, accounting for energy loss and the magnetic field, but ignoring MCS. The MCS exit angle and spatial displacement is added to the trajectory after exiting the downstream M2 mirror plate. The MCS scattering widths are computed the same way scattering errors are computed for the OM reconstruction code (described in Appendix A). Decays of muons are not simulated between the OE and RPC's (over this 4.9 m distance 10 GeV/c muons decay about 1% of the time).

The determination of efficiencies for the inner and outer muon detectors is discussed in Sec. VII.2.9. To a good approximation, the inner muon detectors are treated

---

[7] For inner muons, a requirement that the number of muon hodoscope planes missing hits be no more than two insures muon candidates pass though most of the filters. For outer muons, a requirement that a least 150 cm of material be traversed cuts out tracks passing through the lip of M2, which traverse less material.

[8] The mean momentum of detected outer muons in c-$\bar{c}$ Monte Carlo is about 10 GeV/c with about 40% of incident muons absorbed. Muons below 10 GeV/c in the inner muon system are usually discarded to eliminate backgrounds.

as 100% efficient (no efficiency simulated). For the outer muon detectors, separate efficiencies are simulated for each of the 24 RPC planes in six run periods. The efficiency for each plane is assumed to be independent of the other planes.

CHAPTER VI

METHOD

The quantities of interest are r_{mix}, the fraction of D^0's decaying as $\overline{D}^0$'s, and $\bar{r}_{\mathrm{mix}}$, the fraction of $\overline{D}^0$'s decaying as D^0's. If CP is conserved, then $r_{\mathrm{mix}} = \bar{r}_{\mathrm{mix}}$. For the simplicity, CP will be assumed conserved and charge conjugate modes will be implied in this thesis.

Decays of $D^{*\pm}$ produce an initial pure flavor eigenstate, either a D^0 or $\overline{D}^0$. The initial flavor is identified from the charge of the pion: $D^{*+} \to D^0\pi^+$ or $D^{*-} \to \overline{D}^0\pi^-$. The final flavor is identified from the charges and flavors of the D^0 daughters.

The value of r_{mix} can be computed from the number of non-mixed events measured, the number of mixed events measured, and the time dependence for mixing (or equivalently the relative efficiency between mixed and non-mixed events).

VI.1 Calculation of D^{*+} Mass and D^0 Proper Decay Time

Figure 17 illustrates the semileptonic decay mode to be reconstructed.

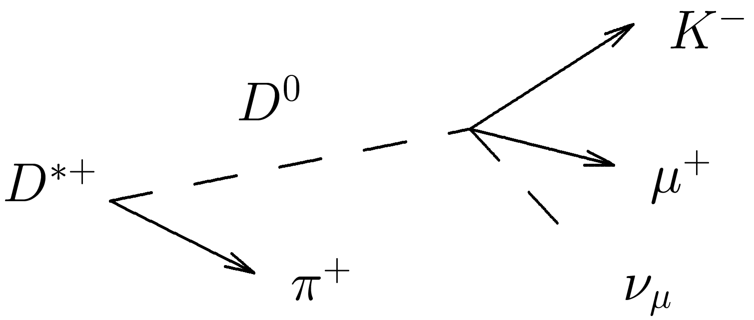


Figure 17: Decay topology for $D^{*+} \to D^0\pi^+ \to (K^-\mu^+\nu_\mu)\pi^+$.

Charges and three-momenta of the π^+, K^-, and μ^+ are measured with the tracking detectors and analysis magnets. The D^{*+} production vertex is determined by selecting the primary with the most tracks. Ties are arbitrated by picking the most upstream primary vertex. The primaries are found using the `DVFREE` algorithm (Sec. IV.3.1). Both primary (D^{*+} production) and secondary (D^0 decay) vertices are required to form a confidence level of at least 1%. The direction and decay length of the D^0 are determined from the position of the primary and secondary vertices. To compute either the proper decay time of the D^0 or the D^{*+} mass, we must determine the energy of the D^0. There are two solutions for the D^0 energy which arise because the neutrino is not measured. Energy-momentum conservation at the D^0 decay vertex can be expressed with four-vectors as:

$$p_0 = p_{K\mu} + p_\nu. \tag{23}$$

where $p_{K\mu} \equiv p_K + p_\mu$ and p_0 is the four-momentum of the D^0. The neutrino momentum can be eliminated as follows

$$\begin{aligned} p_\nu &= p_0 - p_{K\mu} & (24)\\ p_\nu^2 &= (p_0 - p_{K\mu})^2 & (25)\\ p_\nu^2 &= p_0^2 + p_{K\mu}^2 - 2p_0 \cdot p_{K\mu}. & (26) \end{aligned}$$

Define $m_{K\mu}$ as the total energy of the K and μ in the frame where $\vec{p}_{K\mu} + \vec{p}_\nu = \vec{0}$. Since the scalar product of a four-vector is Lorentz invariant, $p_{K\mu}^2 = m_{K\mu}^2$. Likewise, $p_0^2 = m_0^2$ and $p_\nu^2 = m_\nu^2$, where m_0 is the D^0 mass and m_ν is the neutrino mass.

Defining ρ as the three-momentum magnitude of the D^0,

$$m_\nu^2 = m_0^2 + m_{K\mu}^2 - 2E_0E_{K\mu} + 2\rho\hat{p}_0 \cdot \vec{p}_{K\mu} \tag{27}$$

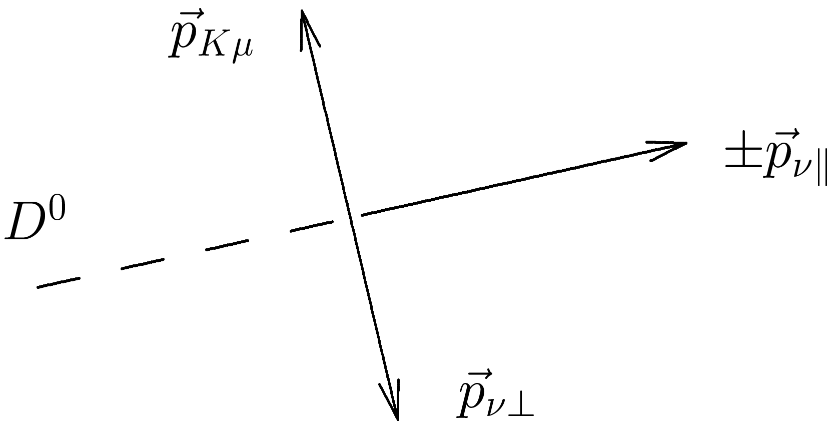


Figure 18: Reference frame with $\vec{p}_0 \cdot \vec{p}_{K\mu} = 0$.

where $\hat{p}_0$ is the unit vector along the D^0 direction. Defining the invariant $M^2 \equiv m_0^2 + m_{K\mu}^2 - m_\nu^2$,

$$\begin{aligned} 2E_0E_{K\mu} &= M^2 + 2\rho\hat{p}_0 \cdot \vec{p}_{K\mu} &(28)\\ \sqrt{\rho^2 + m_0^2} &= \frac{1}{2E_{K\mu}}(M^2 + 2\rho\hat{p}_0 \cdot \vec{p}_{K\mu}). &(29)\end{aligned}$$

Squaring both sides and collecting terms leads to the following quadratic equation for ρ:

$$\left[\left(\frac{\hat{p}_0 \cdot \vec{p}_{K\mu}}{E_{K\mu}}\right)^2 - 1\right]\rho^2 + \left[\frac{M^2\hat{p}_0 \cdot \vec{p}_{K\mu}}{E_{K\mu}^2}\right]\rho + \left[\left(\frac{M^2}{2E_{K\mu}}\right)^2 - m_0^2\right] = 0. \qquad (30)$$

This equation becomes simple in the reference frame with $\hat{p}_0 \cdot \vec{p}_{K\mu} = 0$, where there is only one positive root.[1] This frame is reached by boosting along the direction of the D^0 until the component of $\vec{p}_{K\mu}$ along the boost direction is zero (see Fig. 18). However, in this frame, the D^0 can point either along the boost direction or against the boost direction. This leads to two solutions for p_0, when boosting to any other

---

[1] There is also only one positive solution in any frame where the constant term in the quadratic equation is positive. With $a\rho^2 + b\rho + c = 0$, the product of the two solutions is $\rho_1\rho_2 = c/a$, but a is negative so c is positive for $\rho_1\rho_2 < 0$.

frame. One of the two solutions can be eliminated when, in the lab frame, $\vec{p}_{K\mu} \cdot \hat{p}_0 < 0$. However this never occurs with the beam energy in FOCUS.

Apart from simplifying the quadratic equation, this frame with $\hat{p}_0 \cdot \vec{p}_{K\mu} = 0$ also allows one sometimes to reduce the effect of measurement errors (and increase reconstruction efficiency) by requiring $\rho^2 \geq 0$ in the boosted frame. Due to measurement errors, Eq. (30) often gives a negative value for ρ^2. Rather than discard these events, ρ is set to zero before boosting back to the lab frame. Rare outliers with $\rho^2 < -1\,\mathrm{GeV}^2/c^2$ are discarded. At high enough energies, measurements with $\rho^2 \sim 0$ in the boosted frame become unavoidable. As the D^{*+} momentum in the lab frame increases, the two solutions for the D^0 energy converge. Equivalently, the D^0 is essentially at rest in the boosted frame compared to the boost velocity, for a high energy D^{*+}.

Once the D^0 energy is found, calculation of the D^{*+} mass is straightforward. Applying energy-momentum conservation at the D^{*+} decay vertex:

$$p_* = p_\pi + p_0 \tag{31}$$

$$p_*^2 = p_\pi^2 + p_0^2 + 2p_\pi \cdot p_0 \tag{32}$$

$$m_*^2 = m_\pi^2 + m_0^2 + 2E_\pi E_0 - 2\rho\hat{p}_0 \cdot \vec{p}_\pi . \tag{33}$$

The solution for ρ giving the lowest D^{*+} mass (m_*) is used. Monte Carlo indicates that by choosing this solution, the reconstructed proper time distribution of $D^0 \to K^-\mu^+\nu$ events agrees much better with the generated (true) proper decay time than using either the high or low momentum solution exclusively.

The D^0 proper decay time is found from the D^0 energy and the distance between the two decay vertices.

VI.2 Calculating the Mixing Fraction

r_{mix} is defined as the fraction of D^0's decaying as a $\overline{D}^0$. In terms of the number of mixed signal produced, T_{M}, and the number of non-mixed signal produced, T_{NM}:

$$r_{\text{mix}} = \frac{T_{\text{M}}}{T_{\text{M}} + T_{\text{NM}}}. \tag{34}$$

Since $T_{\text{M}} \ll T_{\text{NM}}$, the following approximation holds and is chosen to simplify error analysis:

$$r_{\text{mix}} \doteq \frac{T_{\text{M}}}{T_{\text{NM}}}. \tag{35}$$

In the remainder of this thesis, r_{mix} will refer to Eq. (35).

In terms of the number of mixed events measured, N_{M}, and the number of non-mixed measured, N_{NM}, Eq. (35) becomes

$$r_{\text{mix}} = \frac{N_{\text{M}}}{N_{\text{NM}}}\alpha \tag{36}$$

where α is the ratio of efficiencies between non-mixed and mixed signal. α is less than one because the spectrometer is more sensitive to long lived decays and mixing extends the life of the particle. By using efficiency corrected Monte Carlo shapes in the fit, r_{mix} is computed directly in the final fit without referring to α. Mixing is not simulated in the Monte Carlo, so the mixing shape is determined by weighting the non-mixed D^0 shapes by the expected mixing functional time dependence. The generated proper time is used in computing the weights. Note since any D^0 can mix, D^0's not from D^{*+}'s are also included in the mixing shape (and included as signal).

The procedure for isolating r_{mix} can be expressed symbolically as follows. The average non-mixed D^0 signal distribution can be expressed as

$$dN_{\text{NM}} = T_{\text{NM}}\,\Gamma e^{-\Gamma t_g} \text{M}(t, m)\, dt\, dm$$

where Γ is the reciprocal of the D^0 lifetime, t is the reconstructed proper decay time, t_g is the generated (or true) proper decay time, and m is the reconstructed D^{*+} mass (whether a D^{*+} is present in the event or not). The function $M(t, m)$ accounts for resolution and efficiency. This is normalized to give the average number of non-mixed D^0 signal measured, N_{NM}

$$N_{\text{NM}} = T_{\text{NM}} \iint \Gamma e^{-\Gamma t_g} \mathrm{M}(t, m)\, dt\, dm. \tag{37}$$

For a mixed signal, this is modified by $\frac{1}{2}\Gamma^2 t_g^2$:

$$N_{\text{M}} = T_{\text{M}} \iint \left(\frac{1}{2}\Gamma^2 t_g^2\right) \Gamma e^{-\Gamma t_g} \mathrm{M}(t, m)\, dt\, dm. \tag{38}$$

From Eqs. (37), (38), and (35),

$$r_{\text{mix}} = \frac{T_{\text{M}}}{T_{\text{NM}}} = \frac{N_{\text{M}}}{N_{\text{NM}}} \left[\frac{T_{\text{NM}} \iint \Gamma e^{-\Gamma t_g} \mathrm{M}(t, m)\, dt\, dm}{T_{\text{NM}} \iint (\frac{1}{2}\Gamma^2 t_g^2)\, \Gamma e^{-\Gamma t_g} \mathrm{M}(t, m)\, dt\, dm}\right]. \tag{39}$$

The quantity in brackets is α. Substituting Eq. (37) into Eq. (39) gives

$$r_{\text{mix}} = \frac{N_{\text{M}}}{T_{\text{NM}} \iint (\frac{1}{2}\Gamma^2 t_g^2)\, \Gamma e^{-\Gamma t_g} \mathrm{M}(t, m)\, dt\, dm} \tag{40}$$

or

$$r_{\text{mix}} \left(\frac{1}{2}\Gamma^2 t_g^2\right) dN_{\text{NM}} = dN_{\text{M}}. \tag{41}$$

This says the mixing signal shape is the non-mixed signal shape (normalized to the number of measured non-mixed signal) scaled event-by-event by $r_{\text{mix}}(\frac{1}{2}\Gamma^2 t_g^2)$.

The mixing shape used in the fit is determined by weighting non-mixed Monte Carlo D^0 events by $\frac{1}{2}\Gamma^2 t_g^2$ and swapping the reconstructed sign (right sign or wrong sign). Further details of this procedure will be discussed in the following sections.

VI.3 Monte Carlo Contributions

The shapes used in the fit for r_{mix} come entirely from the MCFOCUS Monte Carlo (MC) simulation described in Chapter V. The various MC contributions are described and illustrated below. D^{*+} mass distributions are used to contrast $D^{*+} \to D^0\pi^+$ signal with combinatoric (non-peaking) contributions; D^0 proper decay time distributions enable mixed and non-mixed shapes to be distinguished.

VI.3.1 Relative Amounts in Data

Results from fits to data are illustrated in the following figures. The two dominant contributions are tied together in the fit, so there are two parameters fit, representing the dominant and mixing contributions. The two dominant contributions are events containing non-mixed $D^{*+} \to D^0\pi^+$ and all remaining non-mixed events.

Results of a simultaneous right sign (RS), wrong sign (WS), D^{*+} mass, D^0 proper decay time fit to data are shown in Figure 19. The mass distributions are shifted by the fixed D^0 mass. Mixing occurs mainly in WS, so the error on r_{mix} (shown later) is dominated by the combinatoric background in WS.

Figure 20 shows the D^0 semileptonic, D^0 hadronic, and non D^0 contributions predicted from MC and the fit illustrated in Figure 19. Note that even combinatoric (non-peaking) shapes consist mostly of D^0 decays. These are included as signal since any D^0 can mix. The non D^0 contribution is quite low.

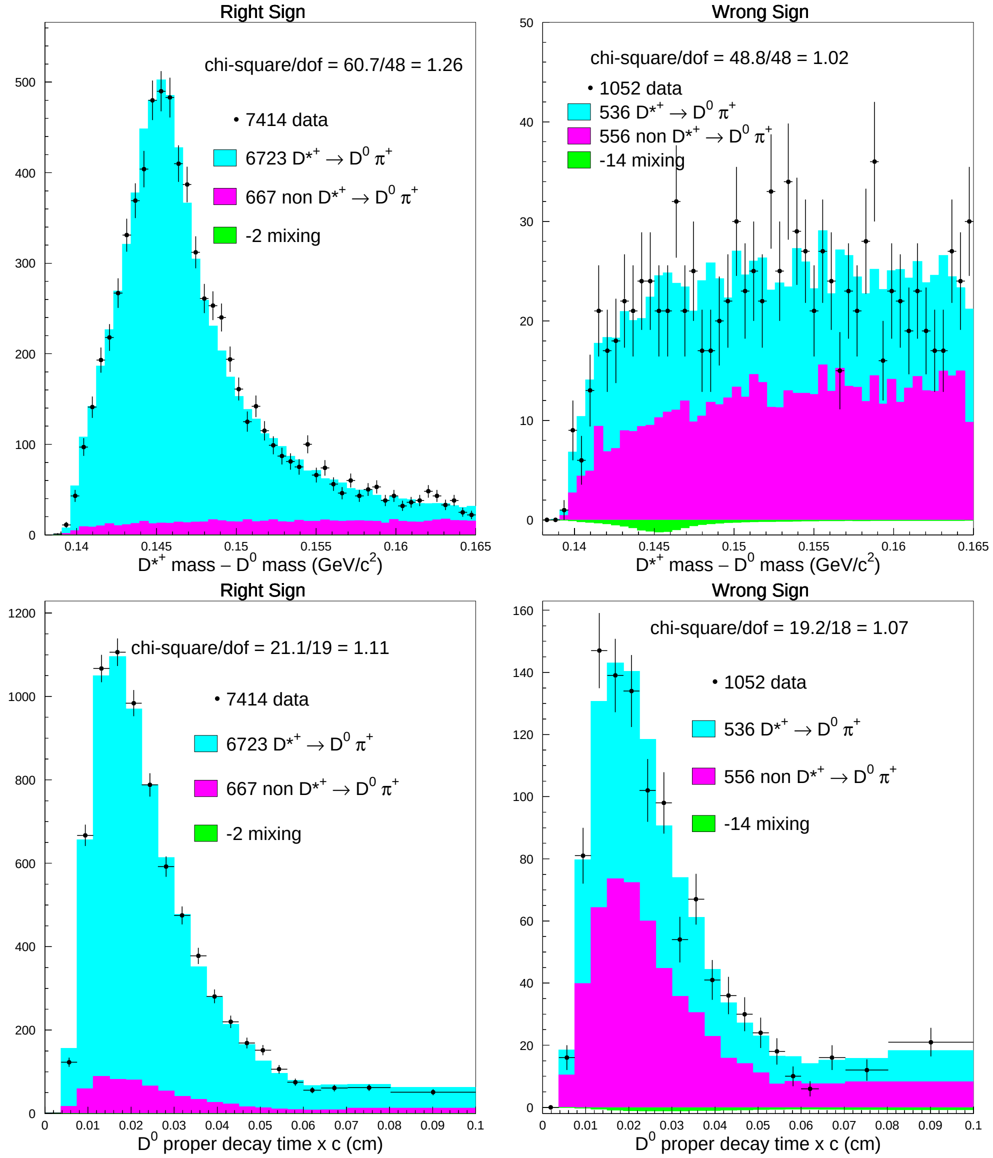


Figure 19: Results of a two parameter, two-dimensional fit to data with Monte Carlo (MC) shapes used in the fit overlaid. The inner and outer muon samples are merged. The low bias bin merging algorithm is used. The first MC shape is from events containing a $D^{*+} \to D^0\pi^+$. The second shape is all events without a $D^{*+} \to D^0\pi^+$. The MC shapes contain slight branching ratio corrections. The ratio of the amounts of the two dominant MC shapes are tied to the MC predicted ratios. ΔM runs from $0.138\,\mathrm{GeV}/c^2$ to $0.165\,\mathrm{GeV}/c^2$ with 50 bins ($0.54\,\mathrm{MeV}/c^2$ wide bins). The proper time runs from 0 cm to 0.1 cm. The first sixteen time bins have widths of 0.00375 cm. The last four time bins have variable widths. In these units, the mean D^0 lifetime is at 0.0124 cm.

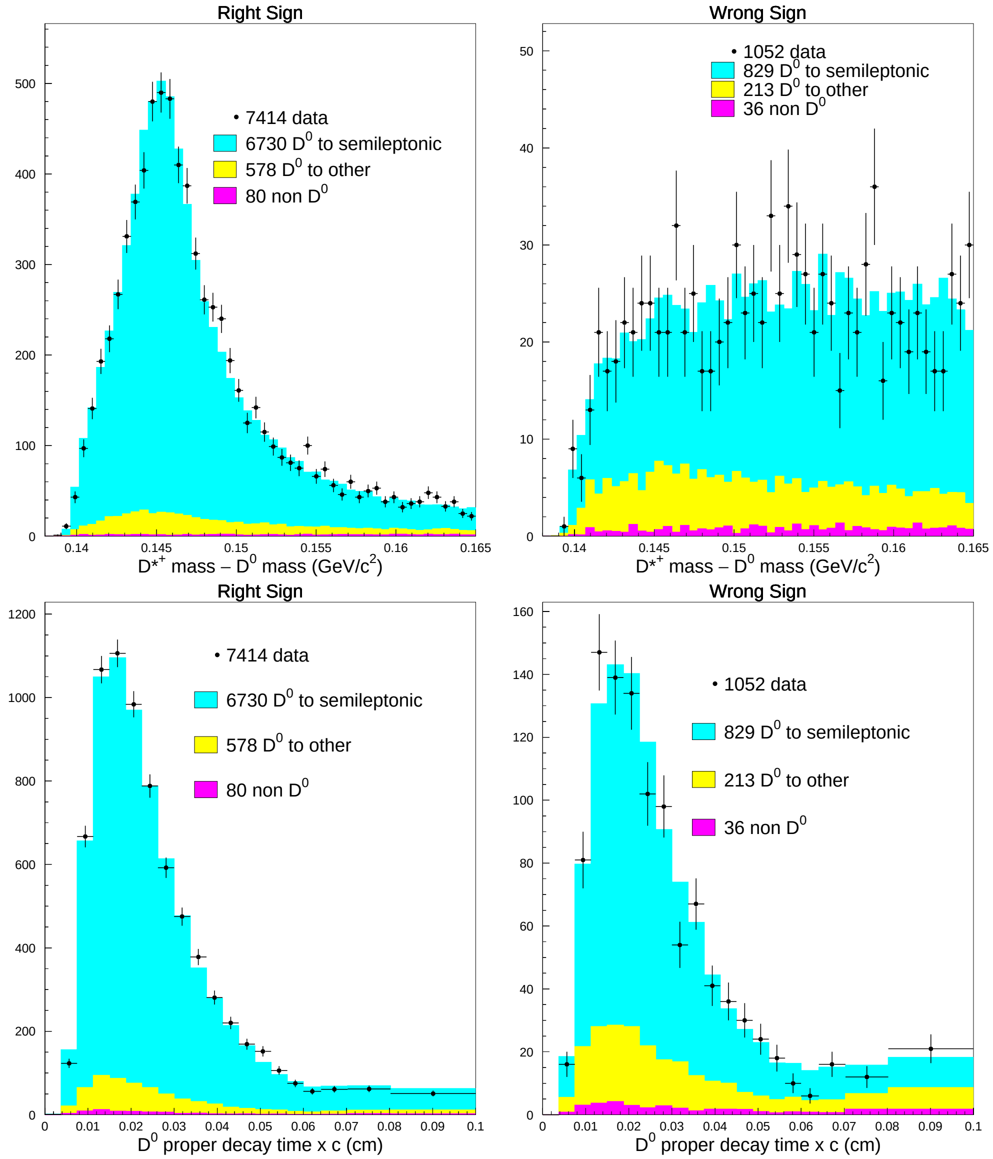


Figure 20: D^0 semileptonic, D^0 hadronic, and non D^0 contributions in the D^{*+} mass and D^0 proper time dimensions. ΔM runs from 0.138 GeV/c^2 to 0.165 GeV/c^2 with 50 bins (0.54 MeV/c^2 wide bins). The proper time runs from 0 cm to 0.1 cm. The first sixteen time bins have widths of 0.00375 cm. The last four time bins have variable widths. In these units, the mean D^0 lifetime is at 0.0124 cm.

VI.3.2 Comparisons of Monte Carlo Shapes

Due to the different scales of the contributions illustrated in the cumulative plots, Figs. 19 and 20, it is difficult to discern some of the shapes. The following plots show these shapes separately.

Figure 21 shows the mass and time distributions for the MC shapes used in the fit. Note the WS mixing mass peak is more narrow than the RS dominant mass peak. This is because mixing emphasizes long-lived decays which have better resolution than short-lived decays. Also note the dramatic rise above $ct = 0.06\,\mathrm{cm}$ due to the t^2 mixing dependence. The proper time binning was chosen to make the RS $D^{*+} \to D^0\pi^+$ ct distribution nearly level for the last four bins, so the mixing plots would be roughly level for $ct > 0.06\,\mathrm{cm}$ without the t^2 enhancement. The drop-off at low ct is due to a $L/\sigma_L > 7$ detachment cut (L/σ_L was defined in Sec. IV.3.2). Figure 22 compares the combined D^0 contributions (scaled to data levels) with the mixing shapes. The mixing shapes in Figure 22 are scaled so that multiplying them by r_{mix} gives the average number of mixed events in the data.

VI.3.3 Determining the Shapes

The "dominant" type is what was directly generated by the MC, which did not simulate mixing.[2] The mixing shape is found by weighting MC events which contain a D^0 by the mixing time dependence $\frac{1}{2}\Gamma^2 t_g^2$, where t_g is the MC generated decay time of the D^0. If there is both a D^0 and a $\overline{D}^0$ in the event, the one that is reconstructed is selected. Which one is reconstructed is determined using matching between microstrip

---

[2] DCS and DCS-mixing interference was not simulated either. This contamination is assessed as a systematic error (discussed in Sec. VII.2.5).

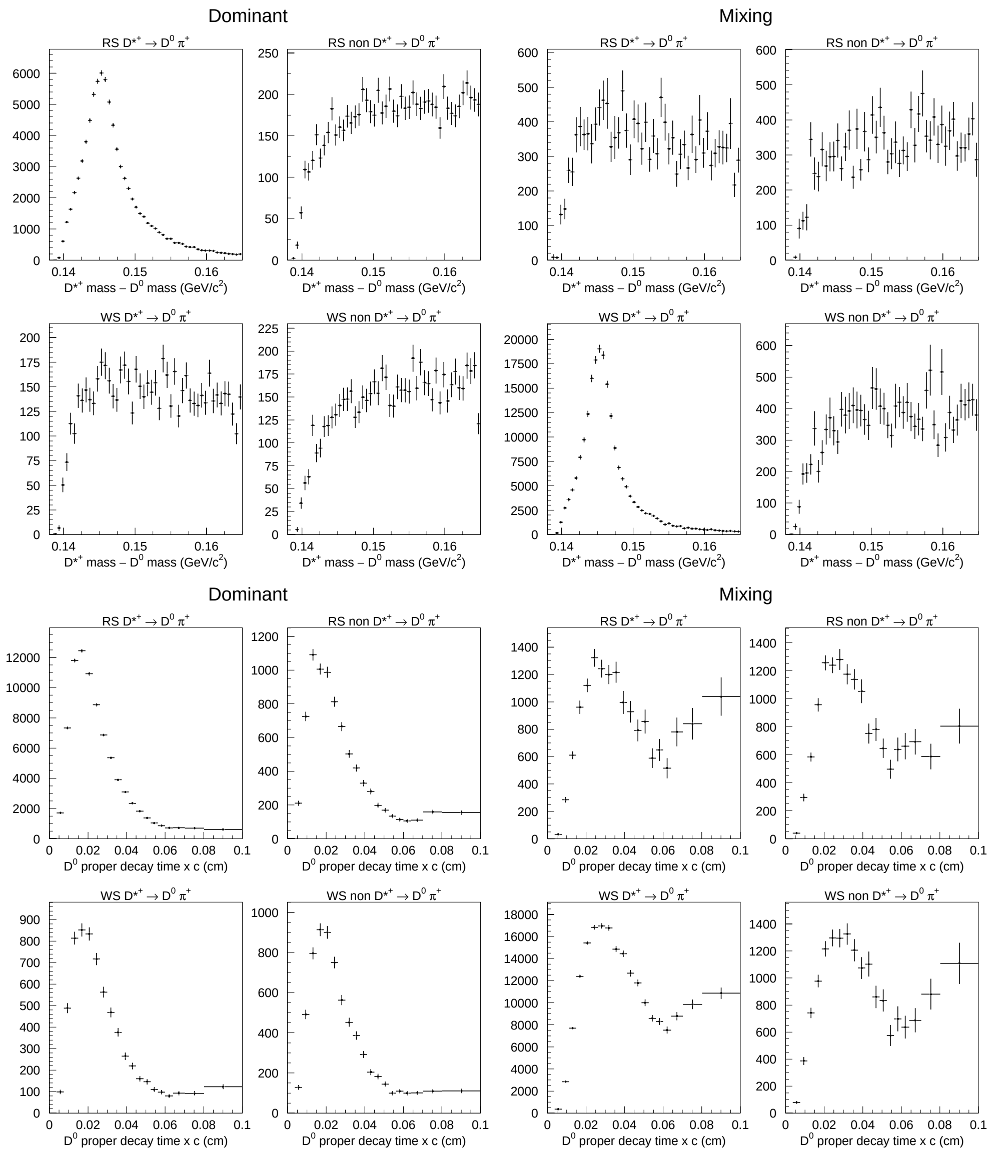


Figure 21: D^{*+} mass and D^0 proper time distributions for the Monte Carlo shapes used in the fit. The top two rows are mass distributions; the bottom two rows are time distributions. The left two columns are non-mixed distributions; the right two columns are mixed distributions.

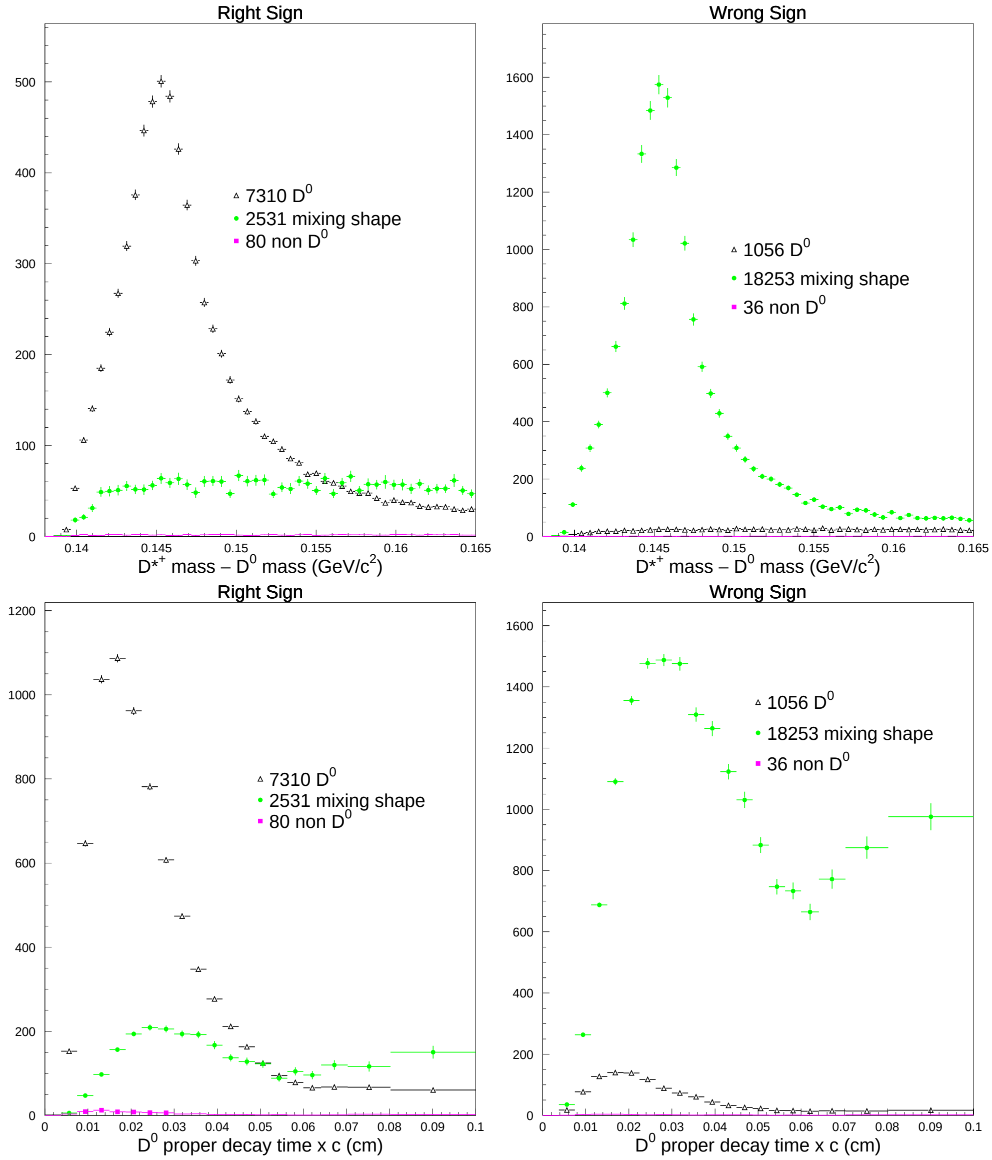


Figure 22: Monte Carlo D^{*+} mass and D^0 proper time distributions comparing the non-mixed D^0, mixing, and non D^0 shapes. The D^0 shapes have hadronic and semileptonic contributions combined. Non-mixed shapes are scaled to data levels, mixing shapes are scaled so that multiplying by r_{mix} gives the average number of mixed events in the data.

tracks and MC particle trajectories. Instead, which D^0 to weight can be selected at random, but this results in more statistical spread in the weighted reconstructed proper time distributions which in turn degrades the error on r_{mix} by about 34%.[3] With more MC, selecting the D^0 at random might be a better approach. Cases where mixing occurs in both the D^0 and $\overline{D}^0$ are ignored since they should occur with negligible frequency.

All of the shapes also contain branching ratio corrections. Except for the $D^0 \to K^-\mu^+\nu$ branching ratio, these have a fairly minor effect on the shapes used in the fit. These corrections are discussed at length in Sec. VII.2.1.

VI.4 Fitting Procedure and Fit Parameters

The amounts of the contributions are determined from a binned Poisson log-likelihood fit to the data, with shapes for each contribution from MC. Errors due to finite MC statistics are incorporated directly in the likelihood function. Data and MC distributions used in the fit are binned in two-dimensional histograms of D^{*+} mass versus D^0 proper decay time. The right sign (RS) and wrong sign (WS) samples are fit jointly.

VI.4.1 Likelihood Function

The log-likelihood function is the same one used by the CERN `HBOOK` [51] library routine `hmclnl` [52].[4] The likelihood to be maximized is the product of Poisson

---

[3] When simulating the mixing time dependence with rejection both the D^0 and $\overline{D}^0$ in an event are independently allowed to mix.

[4] Rather than use `hmclnl`, the author wrote an equivalent routine which does additional error checking and also provides greater flexibility in defining the fit parameters.

distributions in each bin, for both data and Monte Carlo. This likelihood function accounts for errors due to finite MC statistics.

This log-likelihood function is maximized using the CERN software package, Minuit [53]. Parabolic errors are computed with `MIGRAD` and asymmetric errors are found with `MINOS`.[5]

The log of the likelihood is

$$\ln \mathcal{L} = \sum_{i=1}^{n} d_i \ln f_i - f_i + \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ji} \ln A_{ji} - A_{ji} \tag{42}$$

$$f_i = \sum_{j=1}^{m} p_j w_{ji} A_{ji}. \tag{43}$$

f_i is the expected mean number of data events in bin i. A_{ji} is the most probable mean number of MC events in bin i for source j, which is determined bin by bin from a set of **non-coupled** equations, one for each bin (discussed below). a_{ji} is the actual number of MC entries in bin i for source j. d_i is the number of data entries in bin i. w_{ji} is the weight applied to source j in bin i. Weights are used in this analysis to adjust the proper time dependence of non-mixed MC events to simulated mixed events (Sec. VI.3.3), to correct the branching ratios used to generate the MC (Sec. VII.2.1), and to correct the relative efficiency between the inner and outer muon detectors (Sec. VII.2.9). The parameters p_j are found from the fit and are related to the proportions of each distribution in the data, P_j, by $p_j = N_D P_j / N_j$, where $N_j = \sum_i w_{ji} A_{ji}$ and $N_D = \sum_i d_i$. Each p_j is the factor needed to multiply N_j by to obtain the expected amount in data. (With sufficient MC, $N_j = \sum_i w_{ji} A_{ji} \simeq \sum_i w_{ji} a_{ji}$.) If the MC were a perfect representation of reality, the p_j would be independent of cuts

---

[5] `MINOS` computes errors by finding the point where the log-likelihood changes by an amount selected by the user (0.5 for one standard deviation errors). Perhaps a more rigorous approach would be to convert the log-likelihood back into a probability and integrate this probability.

(apart from statistical fluctuations). Determining A_{ji} is facilitated with the change of variables: $t_i \equiv 1 - d_i/f_i$. A_{ji} is computed from

$$A_{ji} = \frac{a_{ji}}{1 + p_j w_{ji} t_i} \tag{44}$$

where t_i is found by numerically solving

$$\frac{d_i}{1 - t_i} = \sum_j \frac{p_j w_{ji} a_{ji}}{1 + p_j w_{ji} t_i}. \tag{45}$$

If one or more a_{ji} are zero, special considerations apply. Further details are found in Ref. [52]. Errors in w_{ji} are neglected in the fit, but can be assessed as a systematic error by fluctuating the weights.

Shapes weighted event-by-event can be incorporated in the fit in the following way. In this case, w_{ji} is just the ratio of the histogram with event-by-event weighting to the histogram of the same events without weighting, and a_{ji} is the histogram without weighting. Weights for bins with no MC entries are computed by averaging over all the mass bins within a given time slice.[6]

Pathologies in the Likelihood Function

The minimizing routine (`MIGRAD`) can search values of p_j which result in negative f_i. This only occurs in bins with $d_i = 0$ and one or more $p_j < 0$. Note with $d_i = 0$, the term $d_i \ln f_i$ vanishes and the only term constraining f_i (in that particular bin) is $-f_i$, which gives a higher $\ln \mathcal{L}$ for lower f_i, including negative values. Negative f_i are only problematic in their statistical interpretation. There is still a unique maximum of $\ln \mathcal{L}$, and automatic normalization is obtained even with negative f_i. However,

---

[6] Averaging only over nearby bins instead of all bins in a slice gives almost exactly the same fitted r_{mix} and errors.

allowing negative f_i leads to a negative bias in r_{mix} of roughly 0.3 times the statistical error.

Skipping bins with $f_i < 0$ can give rise to discontinuities in $\ln \mathcal{L}$ as `MIGRAD` scans different values of p_j. Adding a penalty term for these bins also can produce a sharply changing $\ln \mathcal{L}$, which `MINOS` has trouble handling. Placing constraints on the negative p_j does not resolve this problem, since the values of p_j which maximize $\ln \mathcal{L}$ can occur outside the physical region.

Initially, this pathology was eliminated as follows. Bins with $d_i < 1$ were merged with nearby bins with $d_i \geq 1$. First, time slices with $d_i < 1$ are merged with adjacent time slices with $d_i \geq 1$. Then within each time slice, bins with $d_i < 1$ are merged with the nearest bin along the mass dimension containing $d_i \geq 1$. In this way, as much information about the time dimension is retained at the expense of the mass dimension. Unfortunately this bin merging procedure leads to a positive bias in r_{mix} of about 0.4 times the statistical error. Systematic errors were determined using this high bias algorithm.

An approach which is nearly unbiased is to merge bins based on the expected number of entries, f_i, rather than the data d_i. For the likelihood function to be well defined, the binning must be fixed prior to fitting. Since f_i is a function of the fit parameters, f_i is estimated by the f_i with the mixing parameter fixed to zero (no adjustable binning is needed with mixing absent, since the single remaining fit parameter is well above zero). Denote this estimate by $f_i(0)$. In each bin, i, $f_i(0)$ should be large enough that the addition of a negative mixing contribution would still give $f_i \geq 0$. From the equations in Ref. [52], the maximum mixing contribution in a bin is

$$f_{i,\text{mix}} = r_{\text{min}}\, p_1 \sum_{j,\text{mix}} \frac{w_{ji} a_{ji}}{(1 + r_{\text{min}} p_1 w_{ji})} \tag{46}$$

where r_{min} is the lowest value of r_{mix} `MINUIT` is expected to search, p_1 is the dominant (non-mixing) fit parameter, and the sum is only over terms for the mixing shapes. Bins are then merged so that $\sum_i f_i(0) > -\sum_i f_{i,\text{mix}}$, where the sum denotes the process of merging.

For 2250 fits to MC subsamples, $r_{\text{min}} = -4.5 \times 10^{-3}$ guarantees $f_i \geq 0$ for all bins after merging. A lower value, $r_{\text{min}} = -6.0 \times 10^{-3}$, is used since it eliminates nearly all bias from the fit (discussed further in Sec. VII.2.11).

As with the original approach, time slices are merged before mass bins within each time slice. This insures there are enough entries in the mass bins within each slice to merge. Unlike the original approach, no attempt is made to merge with the closest bins. Instead, bins are merged sequentially in the positive mass direction first, then in the negative direction (to pick up sparse bins on either end of the mass boundaries).

For plots and goodness-of-fit tests the histograms prior to merging are used.[7] The merged bins are only used in the likelihood function.

An alternative way to handle this problem is to compute confidence intervals using simulated data as suggested by Feldman and Cousins. This approach is not implemented for this thesis but is described more fully in Appendix C.

---

[7] The merging discussed here is completely separate from the adaptive binning used for the goodness-of-fit tests discussed in Secs. VI.5 and VI.7.4.

Table 4: Description of the fit parameters. The first two are constrained to be equal.

Fit parameter	Description
q_1	Dominant $D^{*+} \to D^0\pi^+$
q_2	Dominant non $D^{*+} \to D^0\pi^+$
q_3	Mixing

VI.4.2 Fit Parameters

There are only two variable parameters in the fit, but a total of eight MC sources.[8] Equation (43) has a separate parameter, p_j, for each MC source, but these are computed from the fit parameters. RS and WS parameters are constrained to be equal.

Each MC source is described by three categories: reconstructed sign (RS or WS), type of event generated (event contains a $D^{*+} \to D^0\pi^+$ or not), and type of decay (dominant or mixing). These categories produce $2 \times 2 \times 2 = 8$ sources.

RS and WS events occupy separate i bins. WS bins for RS MC sources are empty (for both the a_{ij} and w_{ij} described in Sec. VI.4.1). Likewise, RS bins for WS MC sources are empty. Data occupies both RS and WS bins. In this way, mixing in both the RS and WS sample can be accounted for (D^0's not from a D^{*+} occupy RS and WS roughly equally).

Tables 4 and 5 summarize the MC shapes incorporated in the fits and show how the eight p_j are computed from three fit parameters, q_j. q_1 and q_2 are constrained to be equal for the final fit, but are allowed to be fit separately when selecting cuts.

Separate mixing shapes are determined for the four categories: RS $D^{*+} \to D^0\pi^+$, WS $D^{*+} \to D^0\pi^+$, RS not $D^{*+} \to D^0\pi^+$, WS not $D^{*+} \to D^0\pi^+$.[9]

---

[8] This can be reduced to four MC sources if the two dominant sources are combined. This gives virtually identical results.

[9] Since the mixing shape is determined from non-mixed MC, the non-mixed RS and WS categories are swapped to model mixing.

Table 5: The eight p_j "strengths" for the MC contributions in the likelihood function are determined from the three fit parameters.

"Strength"	p_j value	Description
p_1	q_1	RS dominant $D^{*+} \to D^0\pi^+$
p_2	q_2	RS dominant non $D^{*+} \to D^0\pi^+$
p_3	q_3q_1	RS mixing $D^{*+} \to D^0\pi^+$
p_4	q_3q_2	RS mixing non $D^{*+} \to D^0\pi^+$
p_5	q_1	WS dominant $D^{*+} \to D^0\pi^+$
p_6	q_2	WS dominant non $D^{*+} \to D^0\pi^+$
p_7	q_3q_1	WS mixing $D^{*+} \to D^0\pi^+$
p_8	q_3q_2	WS mixing non $D^{*+} \to D^0\pi^+$

VI.5 Goodness-of-Fit

Standard chi-square tests can be done for the proper time or mass histogram dimensions. For the two dimensional histograms in this analysis, bins must be merged to perform a standard chi-square confidence level test. This is because the bin entries are sparsely populated which results in deviations from Gaussian errors.

To compute a standard chi-square on a two dimensional distribution, a crude adaptive binning procedure is employed. Bins are combined until the expected number of entries in bin i, $\sum_j p_j w_{ji} a_{ji}$, is at least 15. Since the procedure allows variable bin shapes, bins are non-contiguous in a few places. The effect of non-Gaussian errors was investigated using mini-MC (Poisson fluctuating the bin entries with scaled down MC replacing data). For the adaptively binned chi-square, the number of bins are the same, on average, between the original data and fluctuated data and typically varies within three bins from one sample to the next. The adaptively binned chi-square tends to be higher, on average, than the number of degrees of freedom. This means the Gaussian approximation underestimates the confidence level.

Goodness-of-fit confidence levels for fits to data are presented in Sec. VI.7.4.

VI.6 Confidence Intervals

The approach used to compute the limit is a Gaussian approximation which excludes unphysical negative values from the confidence interval and gives an automatic transition between an interval and an upper limit. This method is described in Sec. IV-B of the Feldman-Cousins article [54]. A Gaussian resolution function suffices, since the errors determined from the likelihood function are nearly parabolic. The positive (larger) error is used to compute the limit.

The Feldman-Cousins paper also suggests how to generalize this to a non-parabolic likelihood function using simulated data. See Appendix C for further details.

VI.7 Event Selection and Error Optimization

The plots shown in the preceding sections use cuts selected by an optimization procedure. The MC used to simulate backgrounds in this analysis only simulates events containing charm decays. Even in events containing charm, backgrounds occurring from decays within material are less well understood. In order to reject non-charm backgrounds, it is necessary to apply more restrictive cuts than are optimal for the statistical sensitivity. The problem of optimizing the sensitivity then is to find the loosest set of cuts where the MC is an accurate model of the data.

Most of the cuts were selected with the value of r_{mix} hidden. This was done by adding a constant unknown offset to r_{mix}. After removing the offset, it was necessary to change two cuts to further reduce statistical and systematic errors.

For the final result, the inner and outer muon samples are merged. Separate fits to the inner and outer samples were made when selecting cuts and cross-validating the MC. The shapes used for separate inner and outer muon fits are very similar to the merged sample shapes illustrated above. For the sake of completeness, the separate inner and outer muon shapes are illustrated in Appendix D.

VI.7.1 Procedure for Optimizing Cuts

Since mixing is a tiny contribution, it can be ignored for assessing the match between data and MC. Excluding mixing, with the final cuts the MC is a good match to data in the reconstructed D^{*+} mass and D^0 proper decay time distribution.[10] If the MC is accurately modeling data, the ratio of data to MC should remain constant (within statistical fluctuations) as the cuts are tightened. When fitting, the same should be true for r_{mix} and any p_j parameters (Secs. VI.4.1 and VI.4.2) in the likelihood function.

Since the MC shapes are not reliable representations of the data for arbitrary cuts, each cut variable was scanned individually while leaving other cut variables fixed. Cuts were adjusted and scans repeated iteratively until nearly optimal statistical errors were found with fair agreement between data and MC.

Five types of plots are used to make the cut selections. These are illustrated in Fig. 23 for the inner muon K-μ invariant mass. The top two plots are the fitted r_{mix} and r_{mix} errors as a function of the cut variable. A unknown offset is added to r_{mix}. The middle two plots are the dominant fit parameters as a function of cut. The middle left plot uses the default fit with one dominant parameter. The middle right plot shows what happens when the two dominant contributions are allowed to float separately in the fit. The two dominant contributions are $D^{*+} \to D^0\pi^+$ events

---

[10] In fact, for the mass distribution alone, two adjustable parameters are all that is required to match data with MC even with fairly loose cuts. The proper time is more sensitive to non-charm backgrounds than the mass.

Table 6: Minimal cuts applied prior to optimizing with cut scans.

`DVFREE` primary with highest multiplicity
(most upstream primary as tie-breaker)
$L/\sigma_L > 3$
$\mathrm{CL}_{\mathrm{sec}} > 0.01$
$\mathrm{CL}_{\mu} > 0.01$
No missing planes for inner muons
Material traversed > 150 cm for outer muons
$|\vec{p}_\mu| > 5\,\mathrm{GeV}/c$
$\mathtt{W}(\pi) - \mathtt{W}(K) > 1$ of kaon candidate
$0.8\,\mathrm{GeV}/c^2 < M_{K\mu} < 1.855\,\mathrm{GeV}/c^2$
$30\,\mathrm{GeV}/c < |\vec{p}_{D^{*+}}| < 350\,\mathrm{GeV}/c$ for both D^0 momentum solutions
$0.138 \leq M_{D^{*+}} - M_{D^0} < 0.243\,\mathrm{GeV/c^2}$
$ct \leq 0.1\,\mathrm{cm}$
Exclude double links.

(smaller errors) and all the rest (larger errors). The bottom plot shows the ratio of data to MC as a function of the cut variable. The MC contributions are combined in amounts determined from the (mass-time) fit using the final cuts (shown with the arrow). The amounts also have the same branching ratio corrections as the fitted shapes. The qualitative features of the plots are not very sensitive to the relative amounts of the contributions.

Plots for all the optimized cut variables are shown on separate pages. Zero is suppressed on all of these plots. Outer muon plots are shown after the inner muon plots. Which cut variables to optimize was decided based on experience from preliminary studies (described in Appendix E) as well as previous semileptonic analyses. The remaining cuts were fixed at the minimal values listed in Table 6. The following cut variables were optimized with these scans: K-μ invariant mass, L/σ_L, `iso1`, `iso2`, number of sigma out-of-target for the secondary vertex, muon confidence level,[11] muon

---

[11] The inner muon confidence level was not selected with cut scans. The scans for this variable are

momentum, and pion momentum.

For most of the cut variables illustrated in Figs. 23 through 40, the value of r_{mix} is fairly stable as the cut is tightened (as seen in the upper left plots). For three of the cut variables, r_{mix} does not appear stable as the cut is tightened. In each case, the variation seen in the r_{mix} cut scan is comparable to amount of variation expected from statistics alone in that variable (this will be shown in the next chapter).

The first of these potentially problematic variables is the OM muon momentum (Fig. 38). Above 12 GeV/c, combinatoric backgrounds (non $D^{*+} \to D^0\pi^+$) increase more rapidly over signal ($D^{*+} \to D^0\pi^+$) in data than MC. This is indicated by the center right plot. The statistics also rapidly degrades above 12 GeV/c, as can be seen from the increasing error bars, so the observed split may be a statistical fluke. The ratio plot (bottom) indicates a systematic trend. A similar trend is seen in the inner muon momentum plots (Fig. 29), but the inner muon r_{mix} scan exhibits only a slight drop near 12 GeV/c. If a variable is not accurately simulated and the discrepancy cannot be definitely ascribed to background, then the safest cut to place is the loosest possible one. Outer muons below 4 GeV/c are absorbed in the magnet steel, so a cut of 5 GeV/c is reasonable. The center right plot in Fig. 38 indicates the two dominant (non-mixed) fit components track well with loose cuts, so this is further verification that 5 GeV/c is a reasonable choice.

---

shown in any case.

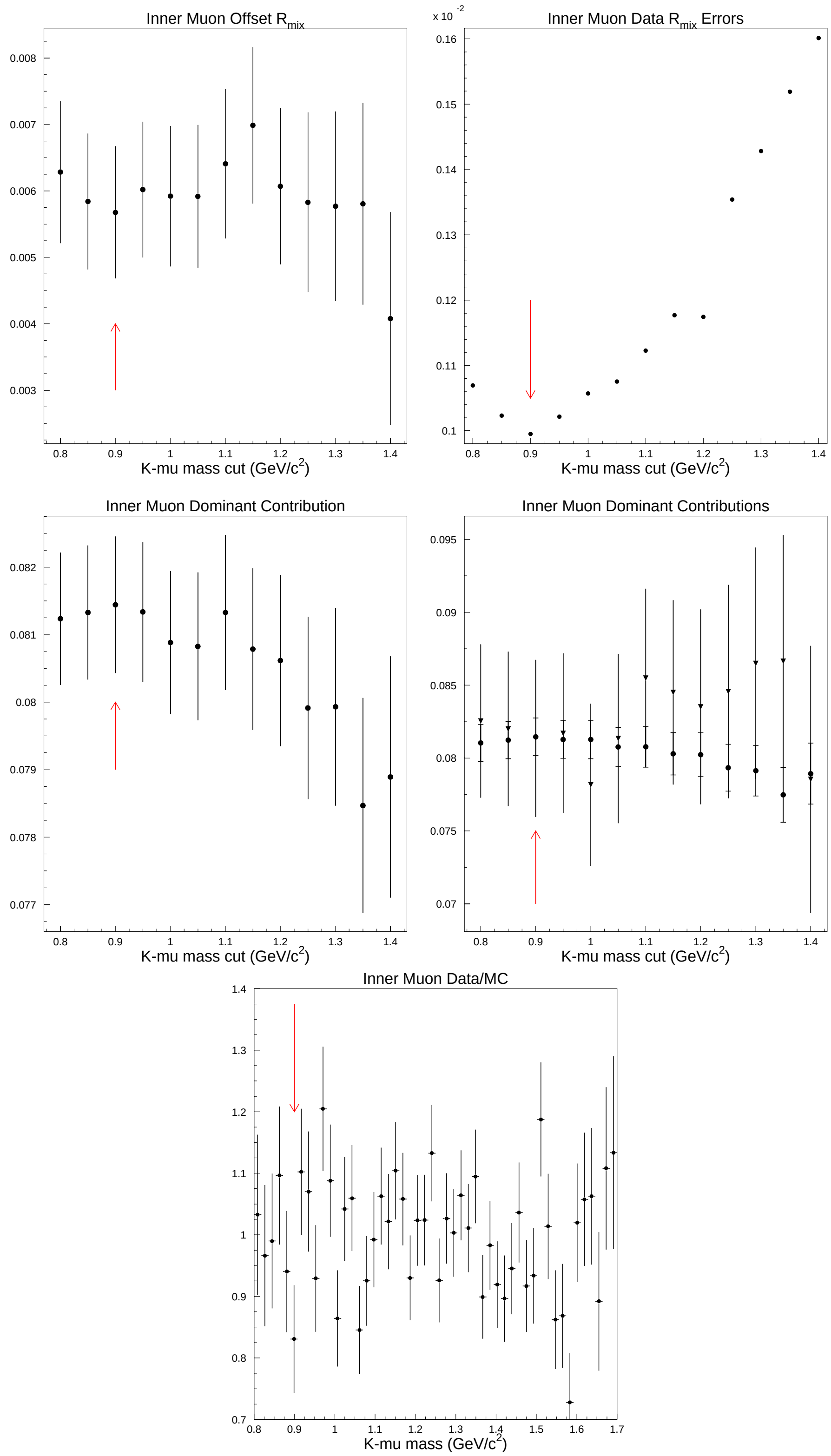


Figure 23: Evolution of fit parameters and data to Monte Carlo ratio as a function of K-μ invariant mass cuts. Arrows indicate where cuts are placed. The inner muon sample is illustrated here.

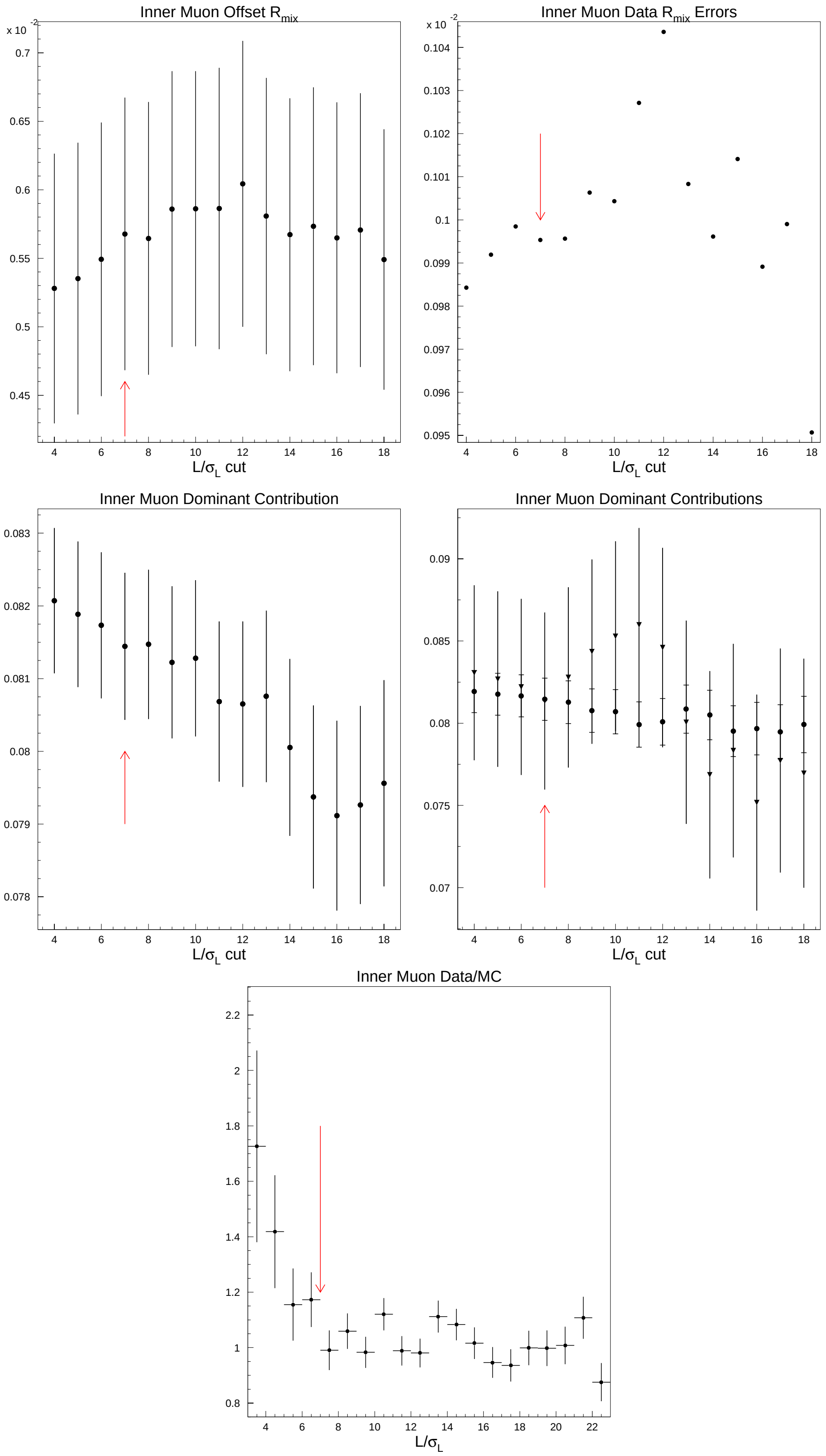


Figure 24: Evolution of fit parameters and data to Monte Carlo ratio as a function of L/σ_L cuts. Arrows indicate where cuts are placed. The inner muon sample is illustrated here.

The second potentially problematic cut variable is the inner muon confidence level (Fig. 28). The ratio plot for this variable suggests a cut of 0.1 is justified. On the other hand, the center right plot indicates that the two dominant contributions agree very well for any cut from 0.01 to 0.2. Cut scans were not made for this variable when selecting cuts. The minimal cut, 0.01, was chosen based only on the ratio plot and the fact that a cut of 0.05 gave a virtually identical r_{mix} and errors as 0.01.

The last potentially problematic cut variable is $W(e) - W(\pi)$ for pion candidates (Fig. 31 for the inner muon sample; Fig. 40 for the outer muon sample). Both the center plots and the ratio plots indicate a strong systematic trend. The fact that the ratio plot is level over a wider range for the outer muon sample than for the inner muon sample is a strong indication that the systematic trend is momentum dependent. This can be seen more clearly by splitting the data and MC in momentum ranges according to the Čerenkov momentum thresholds for pions listed in Table 2. Data and MC $W(e) - W(\pi)$ distributions in the four pion momentum ranges are shown in Fig. 41 for inner muons and Fig. 42 for outer muons. For $W(e) - W(\pi) > 3$, the agreement between data and MC is spectacular in every momentum range except $4.5 \leq |\vec{p}_\pi| < 8.4\,\text{GeV}/c$. From these plots and the center right plots in Figs. 31 and 40 it appears likely that requiring $W(e) - W(\pi) > 3$ eliminates non-simulated backgrounds without seriously distorting the pion momentum distribution. (Ratio plots for the pion momentum distributions are at the bottom of Figs. 30 and 39.)

The $W(e) - W(\pi)$ distributions will be revisited in the next chapter to illustrate a cross-validation procedure.

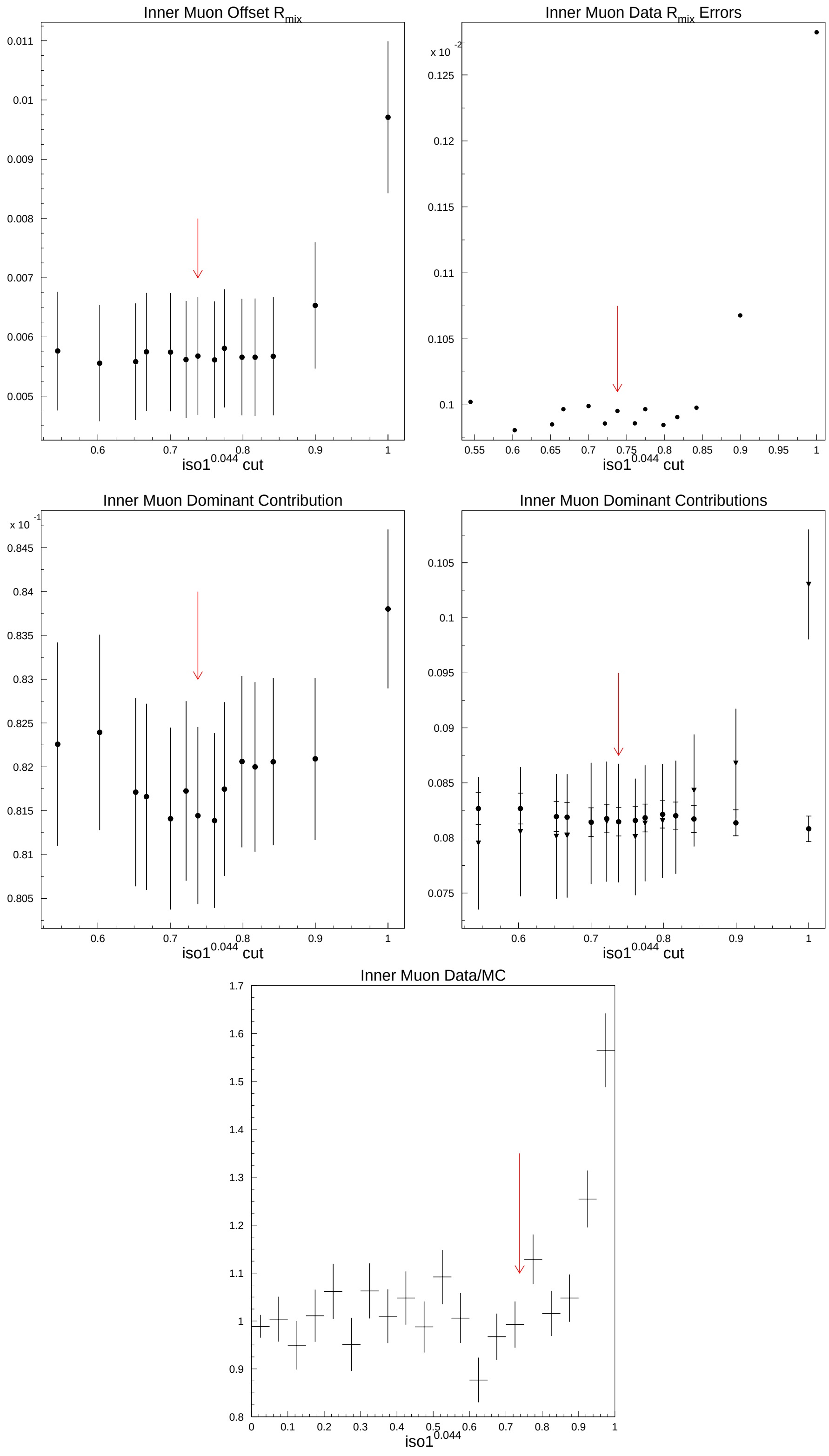


Figure 25: Evolution of fit parameters and data to Monte Carlo ratio as a function of `iso1` cuts. The cut chosen is `iso1` < 0.001. The inner muon sample is illustrated here.

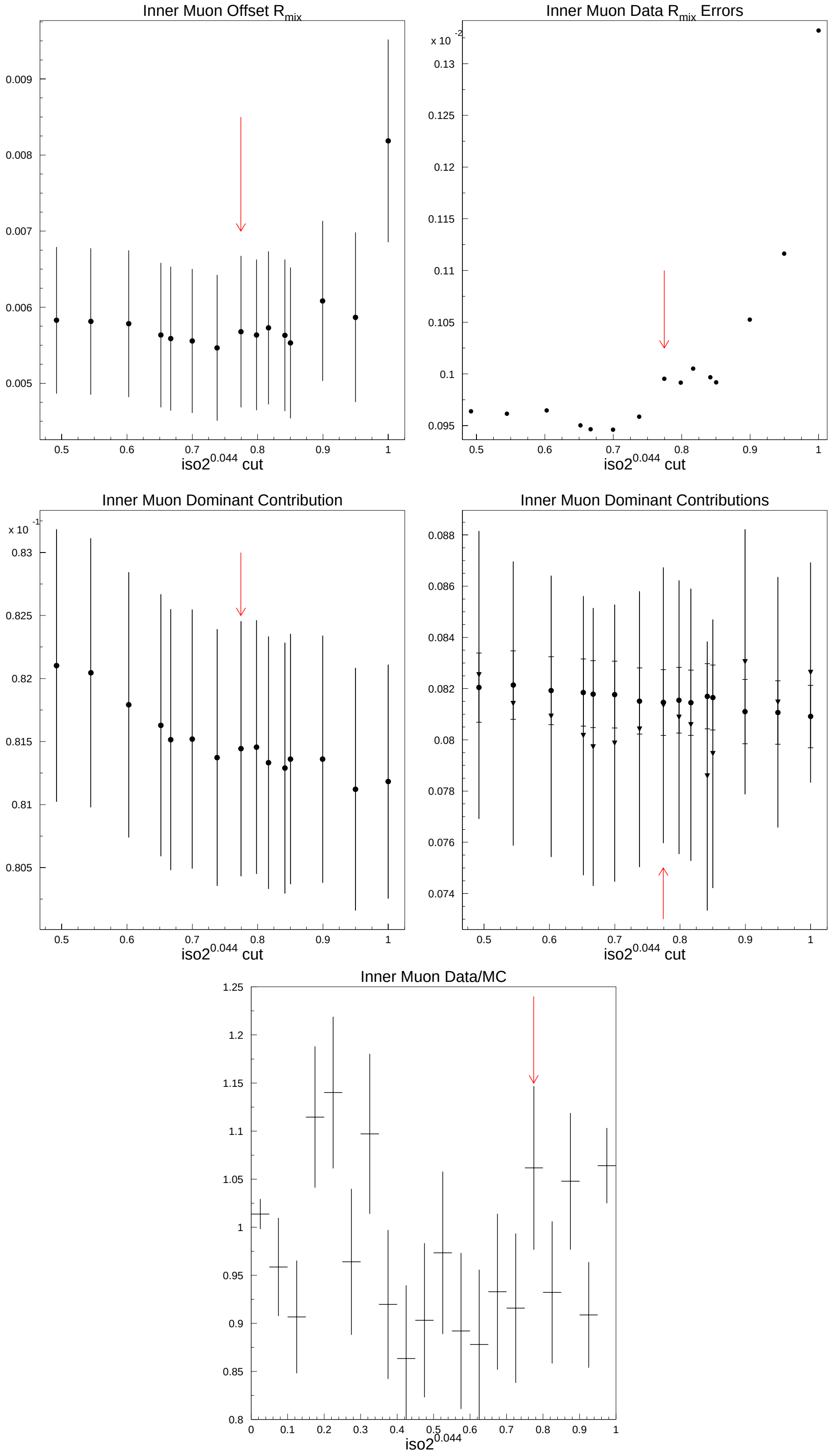


Figure 26: Evolution of fit parameters and data to Monte Carlo ratio as a function of `iso2` cuts. The cut chosen is `iso2` < 0.003. The inner muon sample is illustrated here.

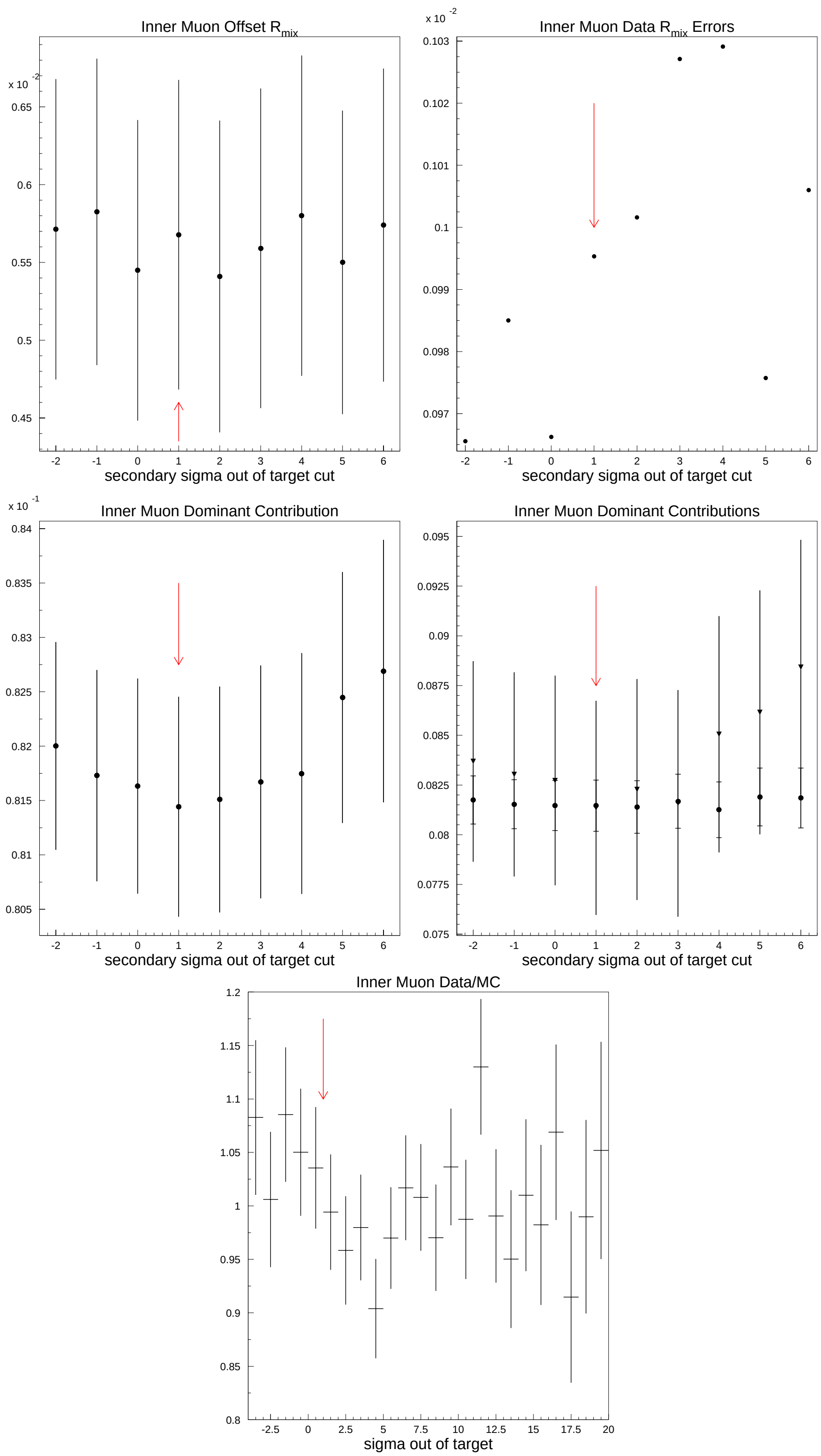


Figure 27: Evolution of fit parameters and data to Monte Carlo ratio as a function of sigma out-of-target cuts. Arrows indicate where cuts are placed. The inner muon sample is illustrated here.

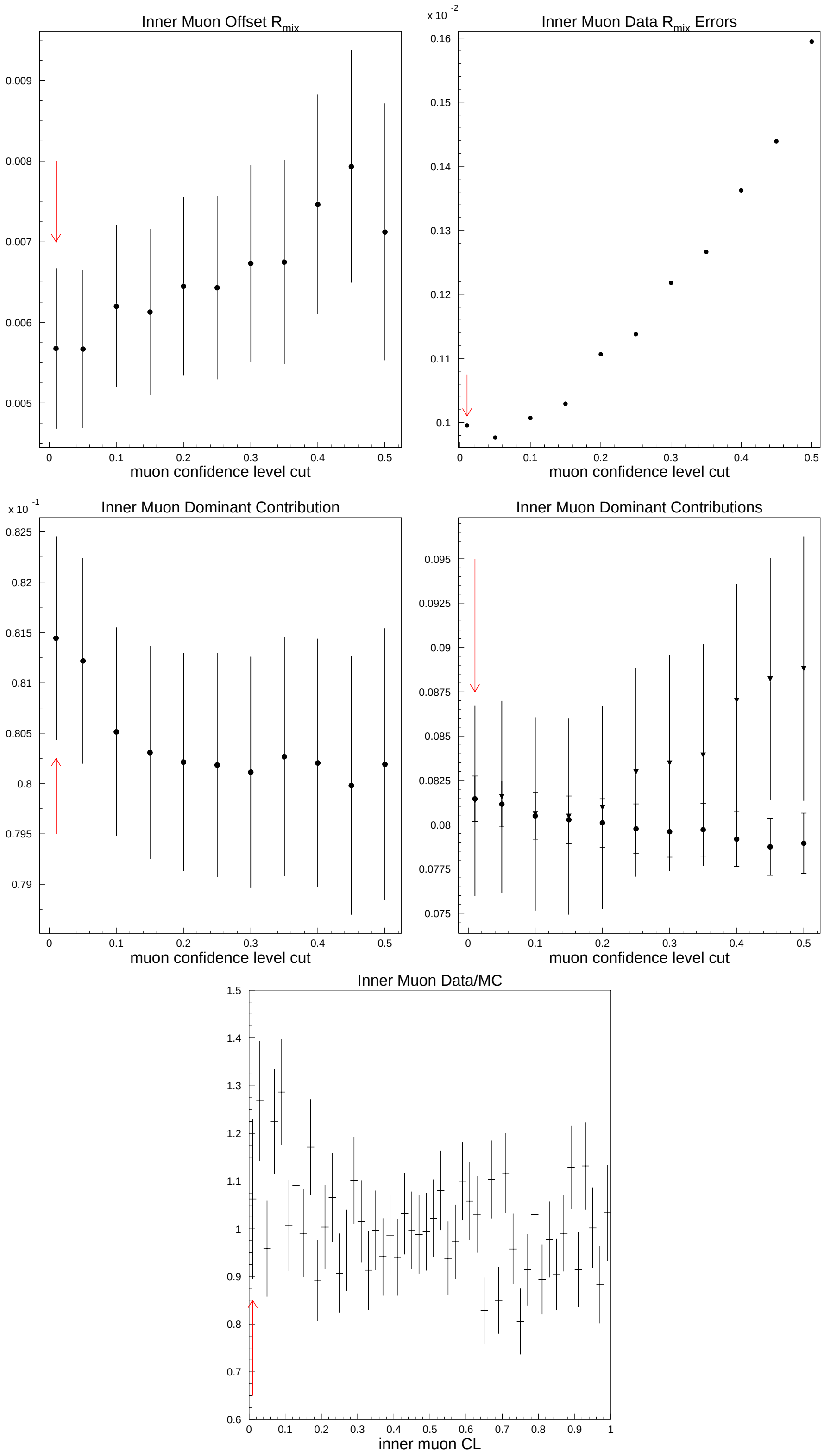


Figure 28: Evolution of fit parameters and data to Monte Carlo ratio as a function of muon confidence level cuts. The inner muon sample is illustrated here.

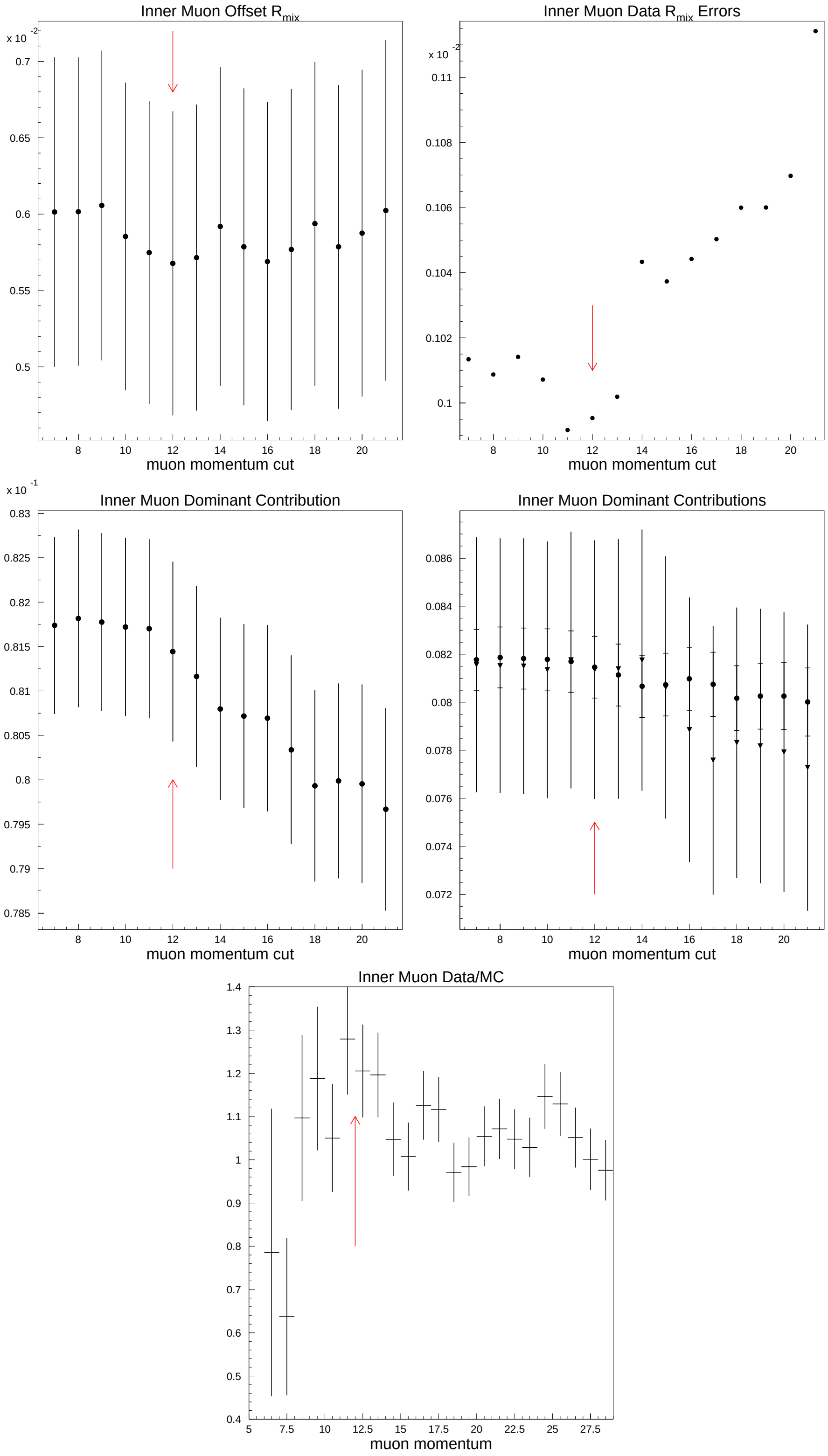
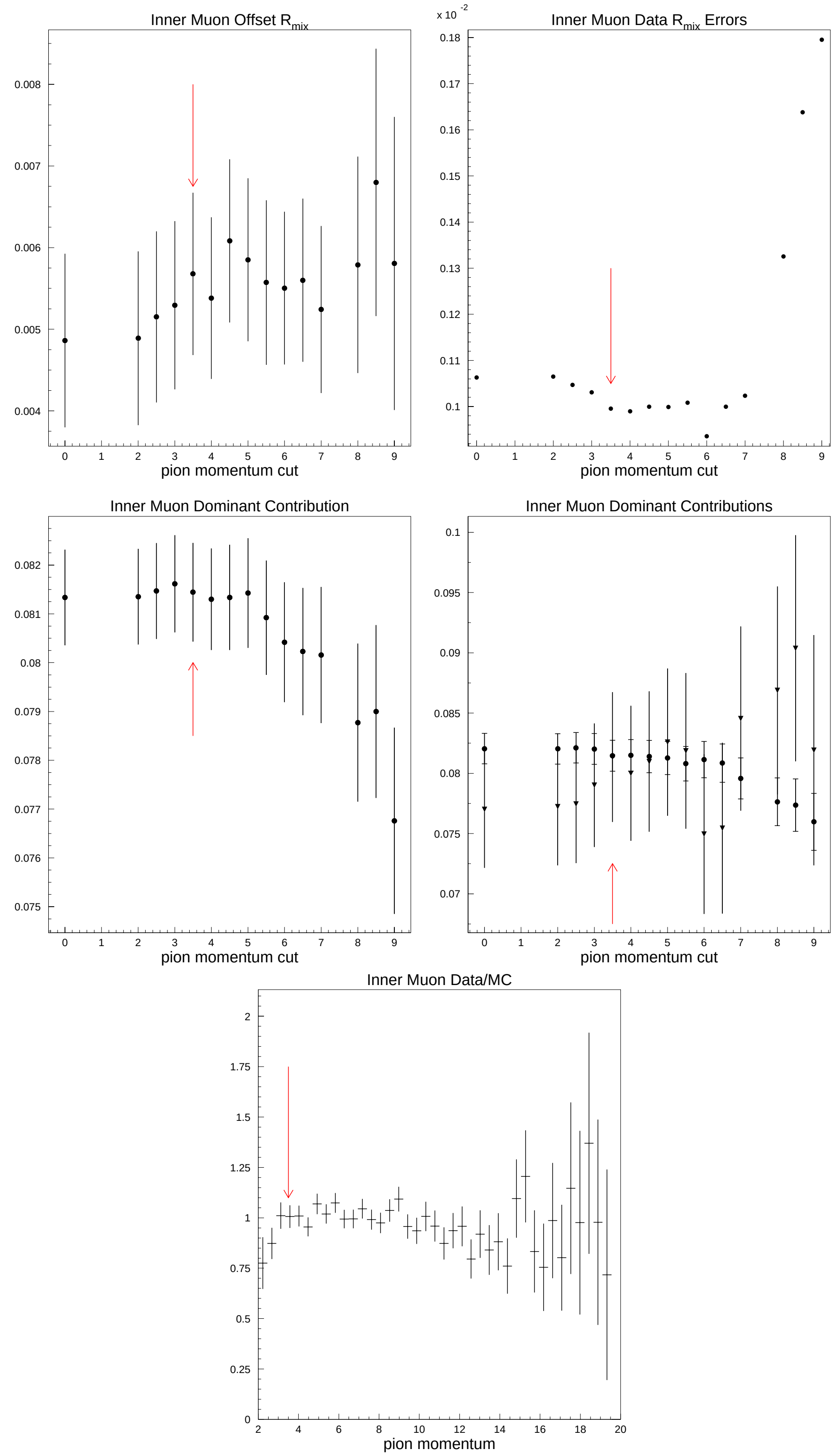


Figure 29: Evolution of fit parameters and data to Monte Carlo ratio as a function of muon momentum cuts. The inner muon sample is illustrated here.



Inner Muon Offset R_{mix}

Inner Muon Data R_{mix} Errors

Inner Muon Dominant Contribution

Inner Muon Dominant Contributions

Inner Muon Data/MC

Figure 30: Evolution of fit parameters and data to Monte Carlo ratio as a function of pion momentum cuts. The inner muon sample is illustrated here.

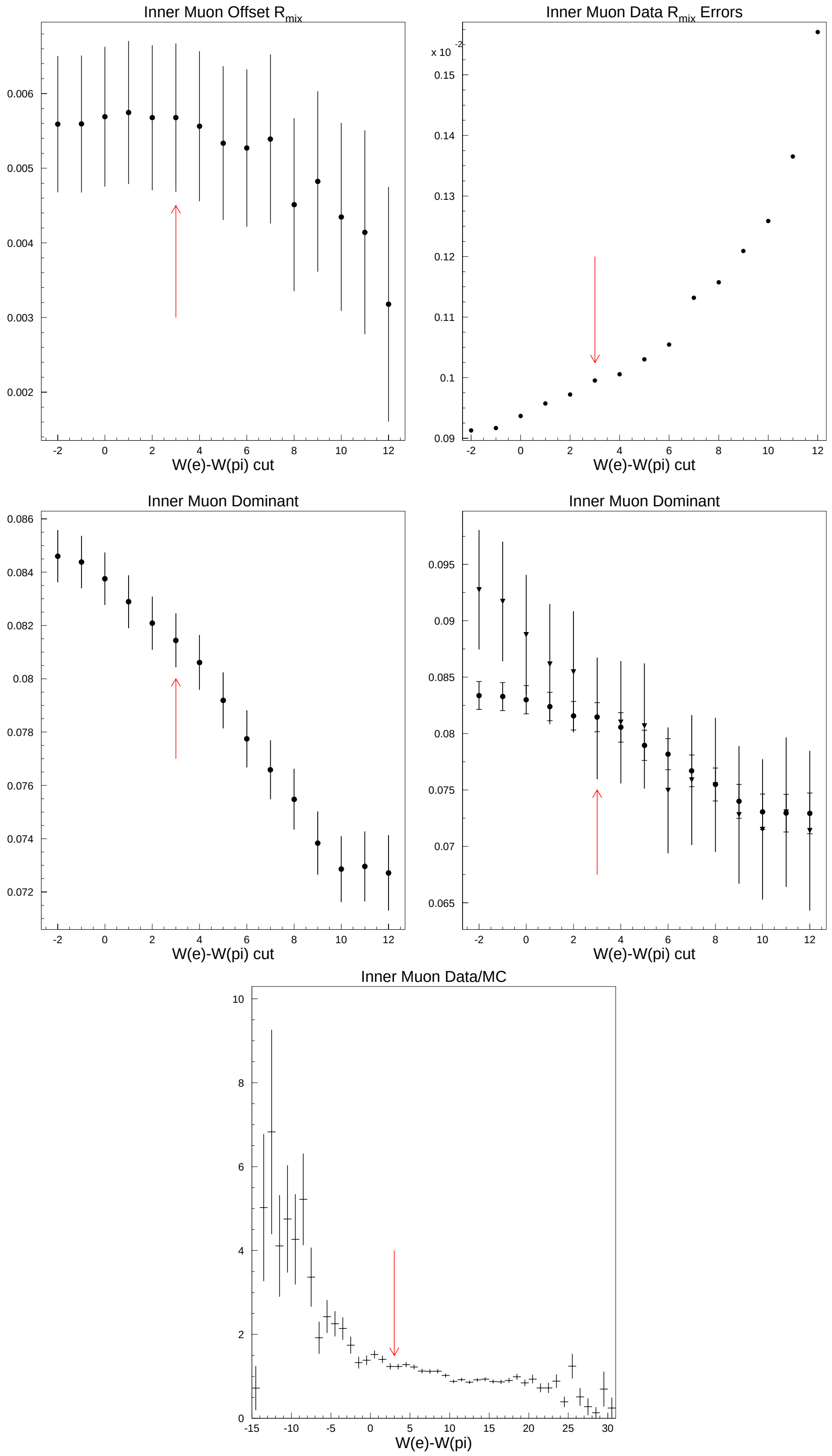


Figure 31: Evolution of fit parameters and data to Monte Carlo ratio as a function of $W(e) - W(\pi)$ cuts. The inner muon sample is illustrated here.

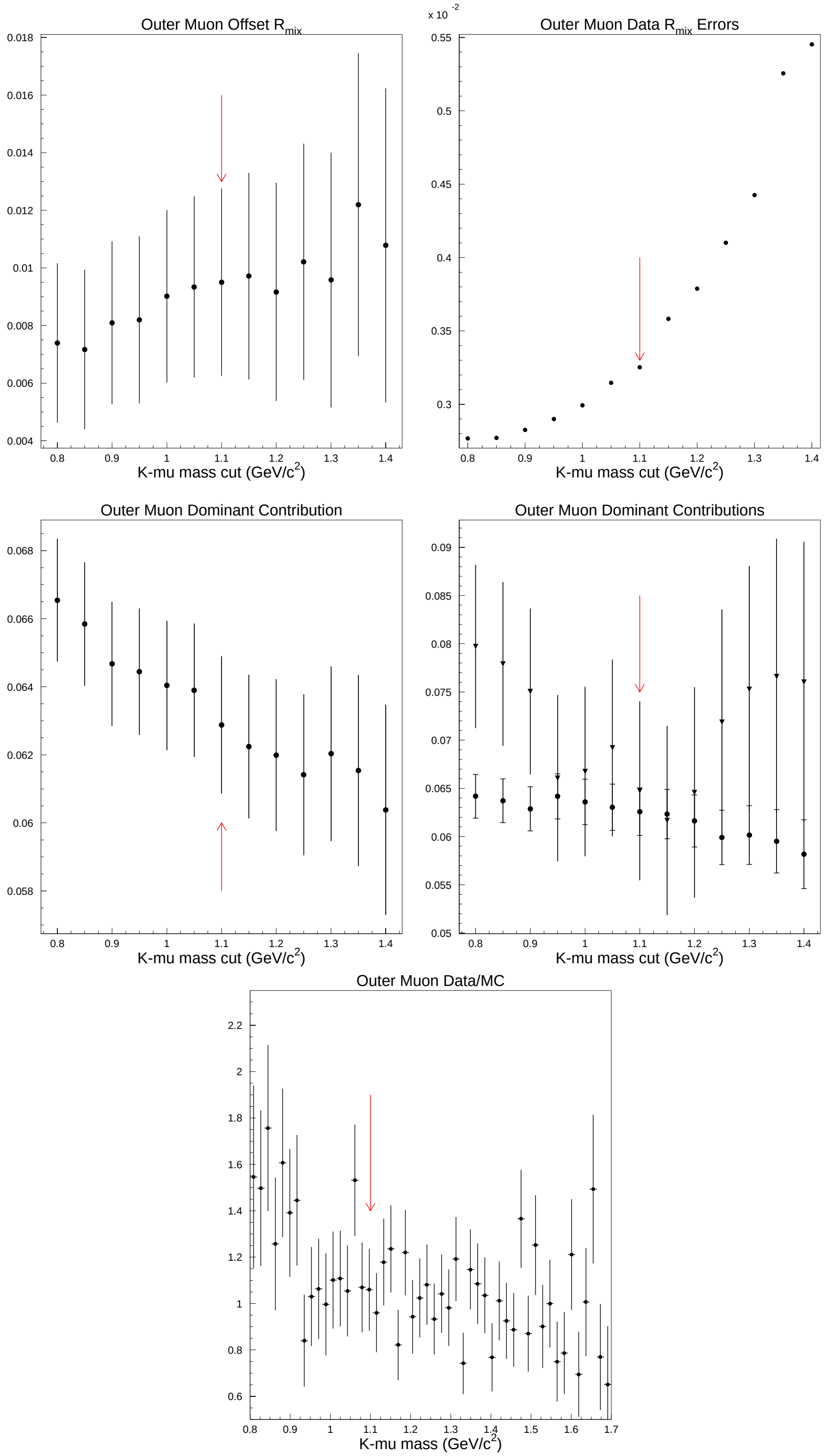


Figure 32: Evolution of fit parameters and data to Monte Carlo ratio as a function of K-μ invariant mass cuts. The outer muon sample is illustrated here.

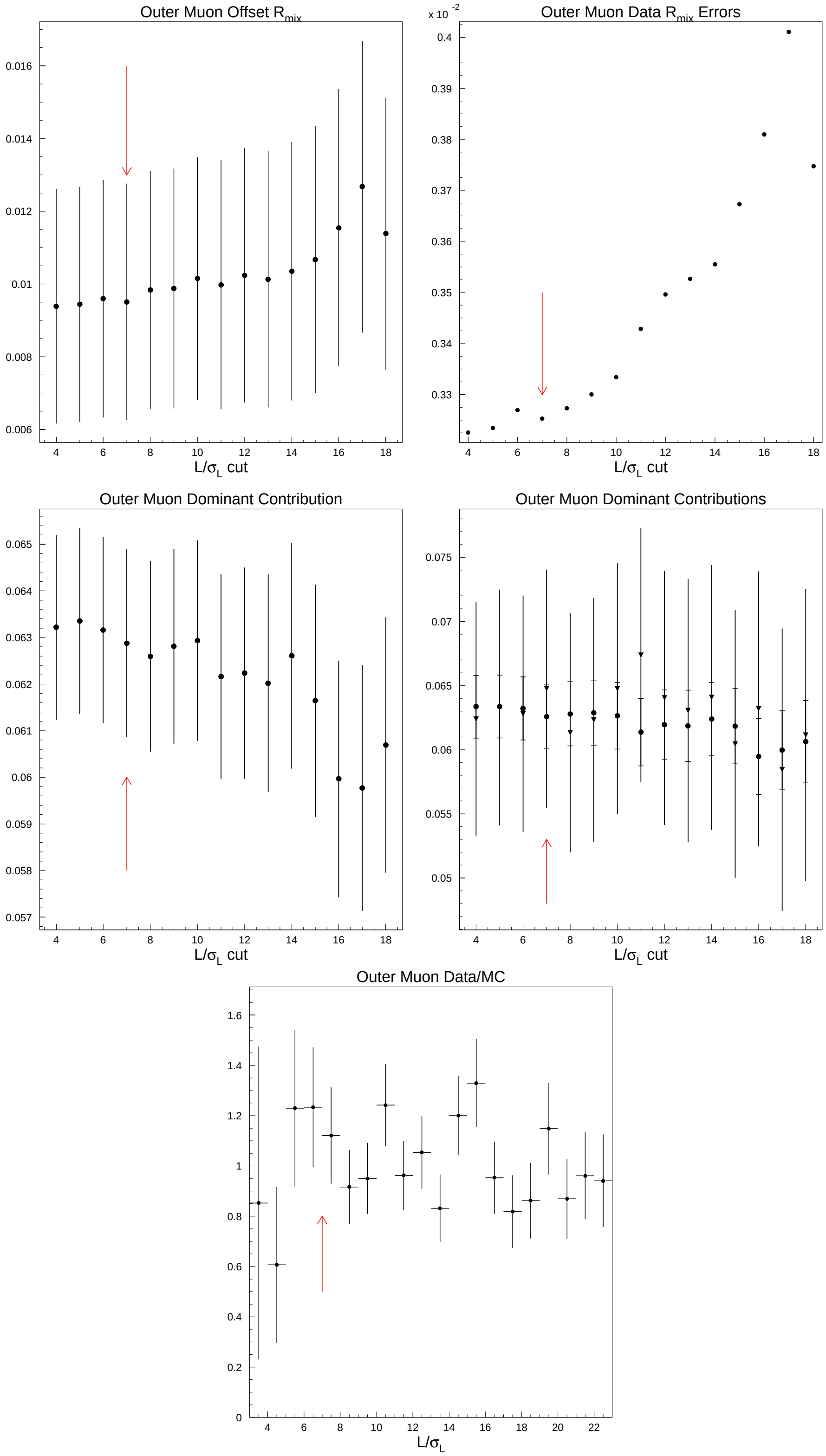


Figure 33: Evolution of fit parameters and data to Monte Carlo ratio as a function of L/σ_L cuts. Arrows indicate where cuts are placed. The outer muon sample is illustrated here.

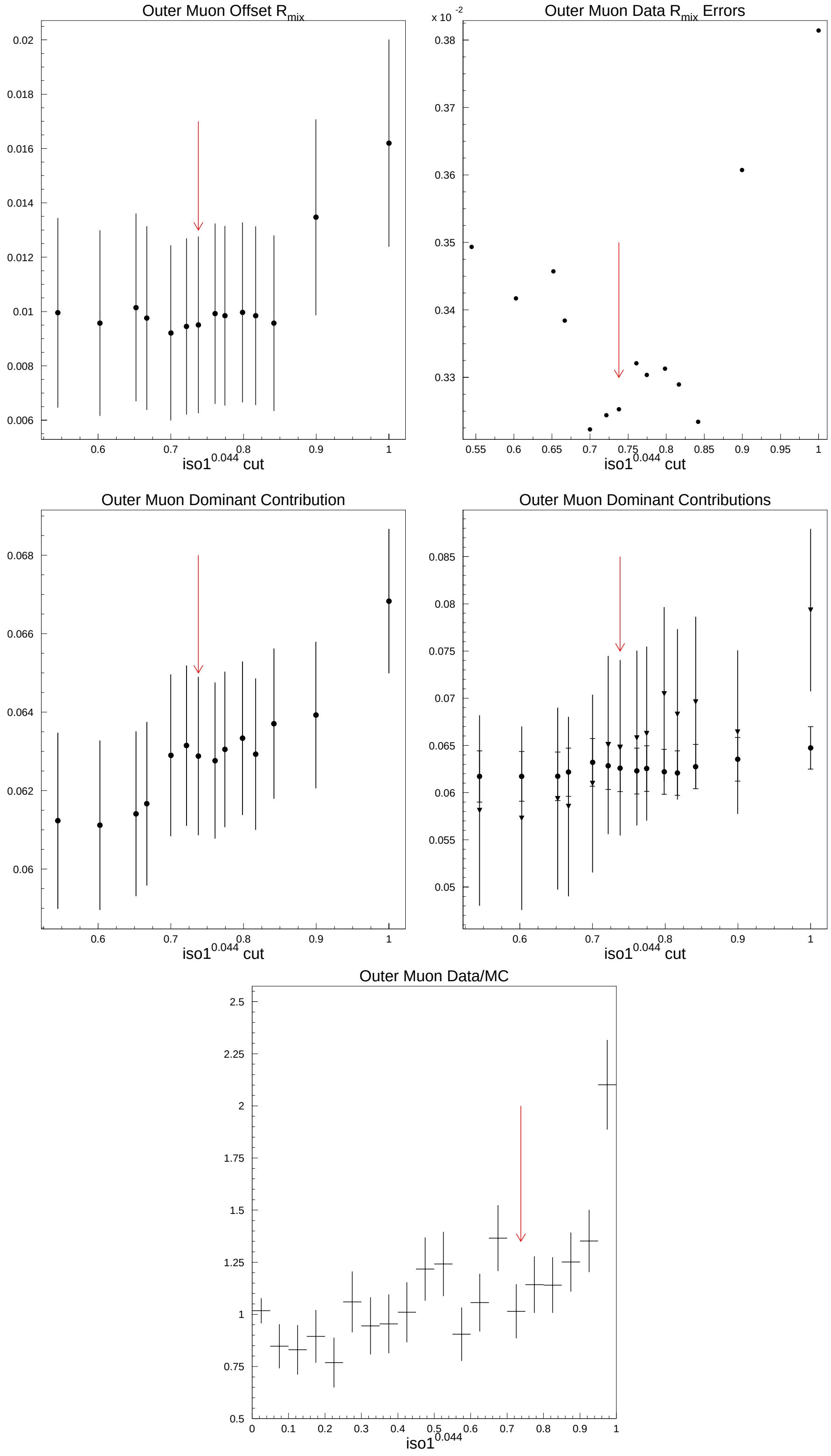


Figure 34: Evolution of fit parameters and data to Monte Carlo ratio as a function of `iso1` cuts. The cut chosen is `iso1` < 0.001. The outer muon sample is illustrated here.

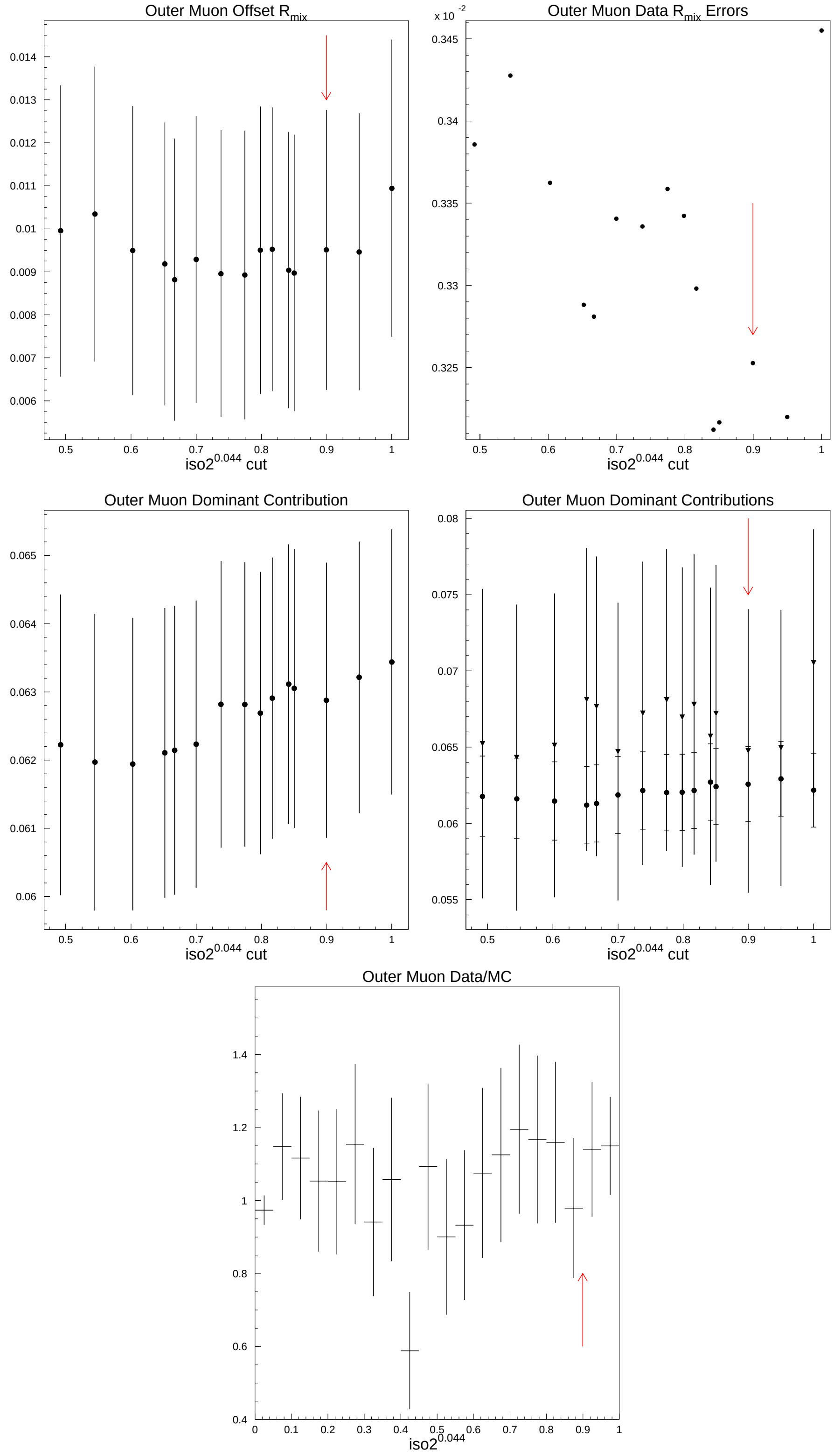


Figure 35: Evolution of fit parameters and data to Monte Carlo ratio as a function of `iso2` cuts. The cut chosen is `iso2` < 0.09. The outer muon sample is illustrated here.

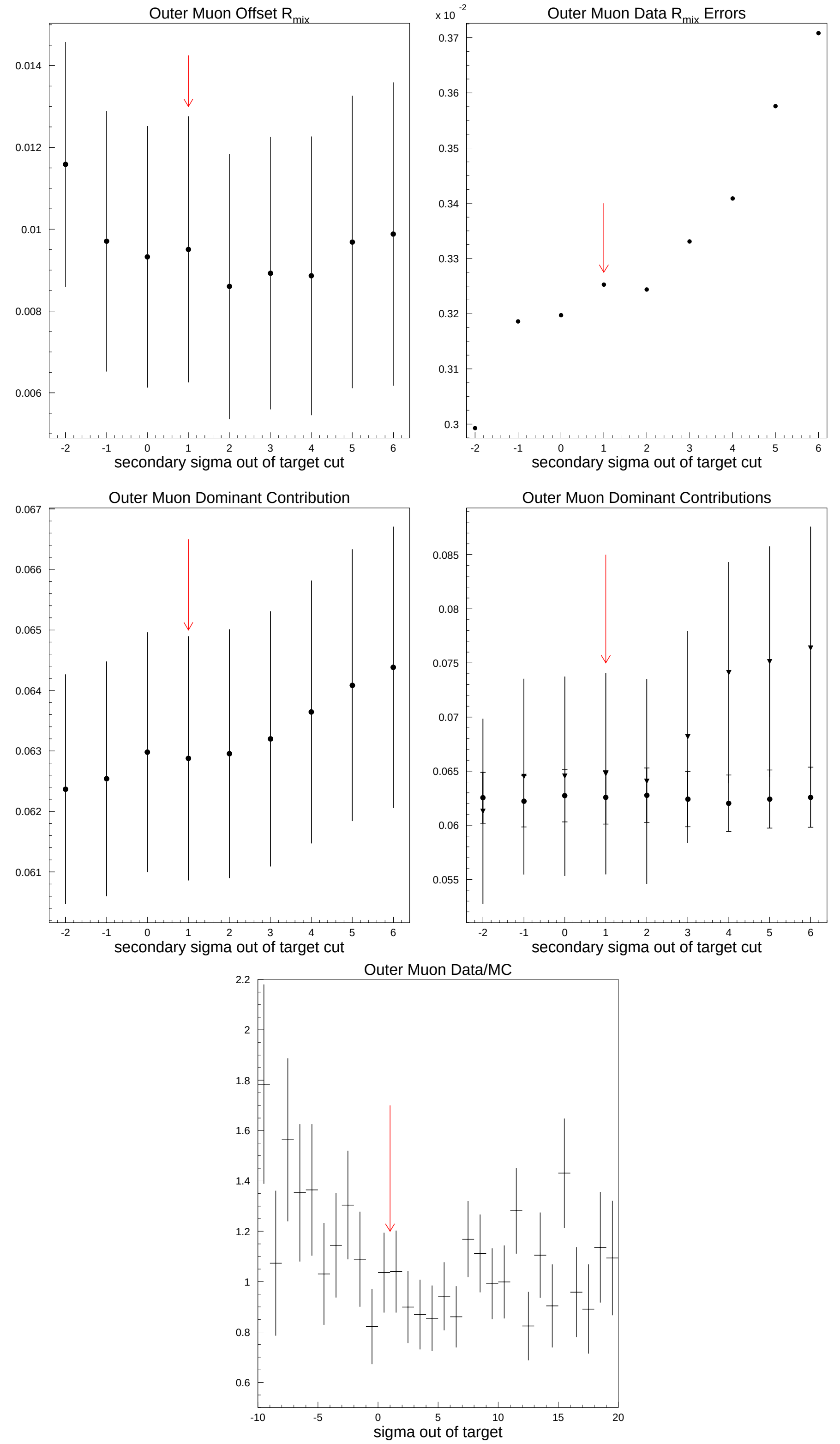


Figure 36: Evolution of fit parameters and data to Monte Carlo ratio as a function of sigma out-of-target cuts. The outer muon sample is illustrated here.

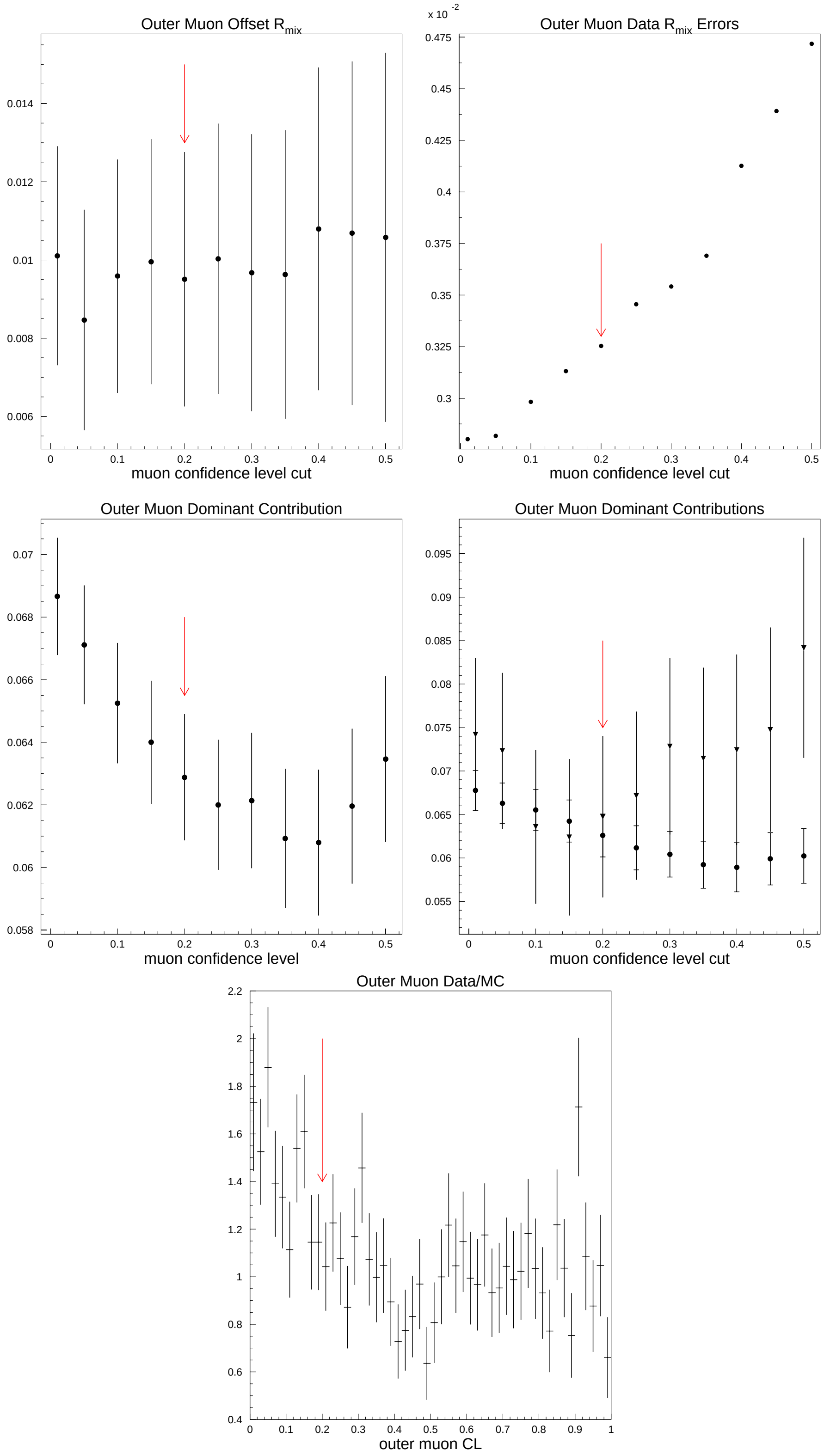
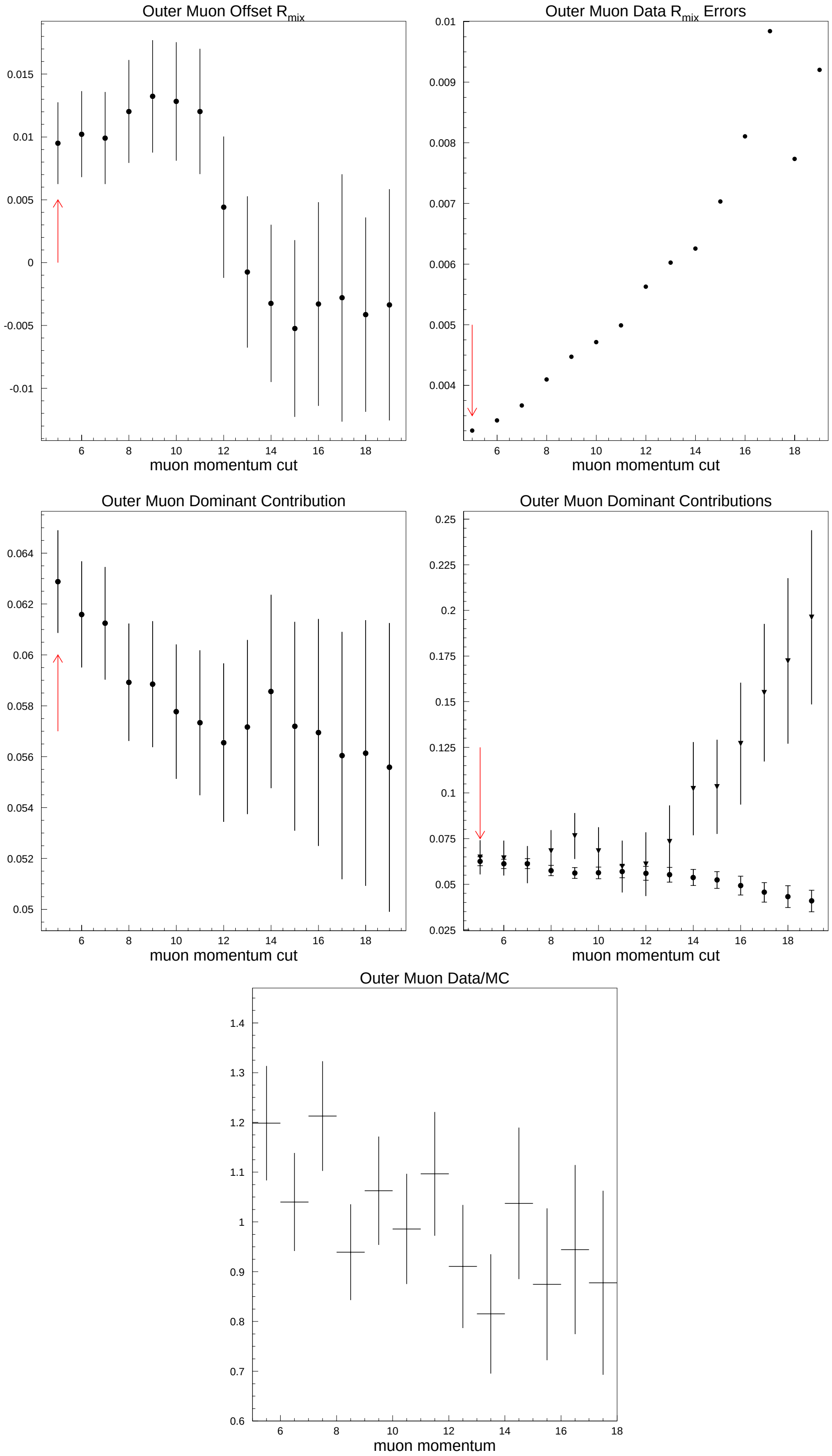


Figure 37: Evolution of fit parameters and data to Monte Carlo ratio as a function of muon confidence level cuts. The outer muon sample is illustrated here.



Outer Muon Offset R_{mix}

Outer Muon Data R_{mix} Errors

Outer Muon Dominant Contribution

Outer Muon Dominant Contributions

Outer Muon Data/MC

Figure 38: Evolution of fit parameters and data to Monte Carlo ratio as a function of muon momentum cuts. The outer muon sample is illustrated here.

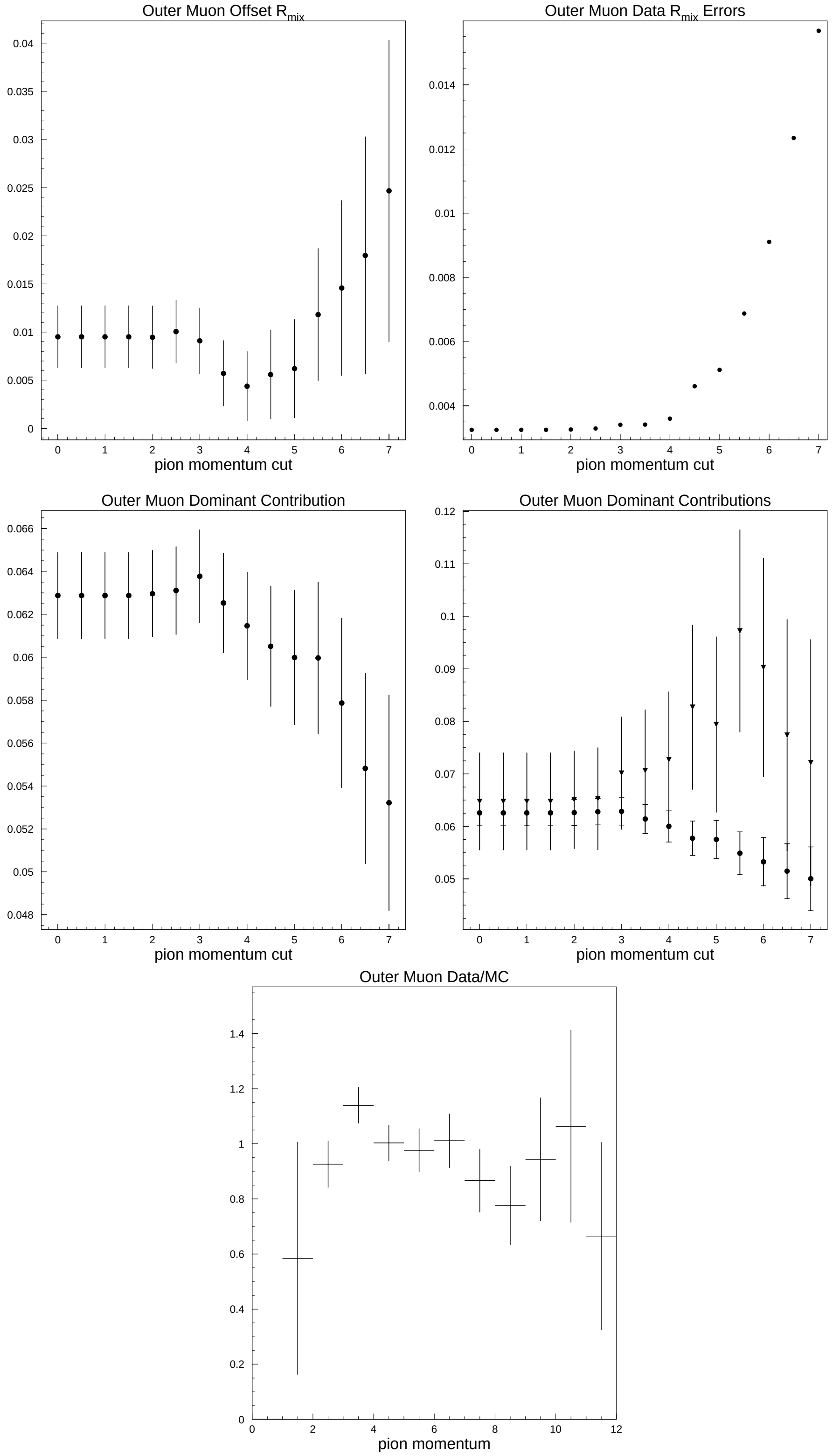


Figure 39: Evolution of fit parameters and data to Monte Carlo ratio as a function of pion momentum cuts. The outer muon sample is illustrated here. No cut is placed on the pion momentum for the outer muon sample.

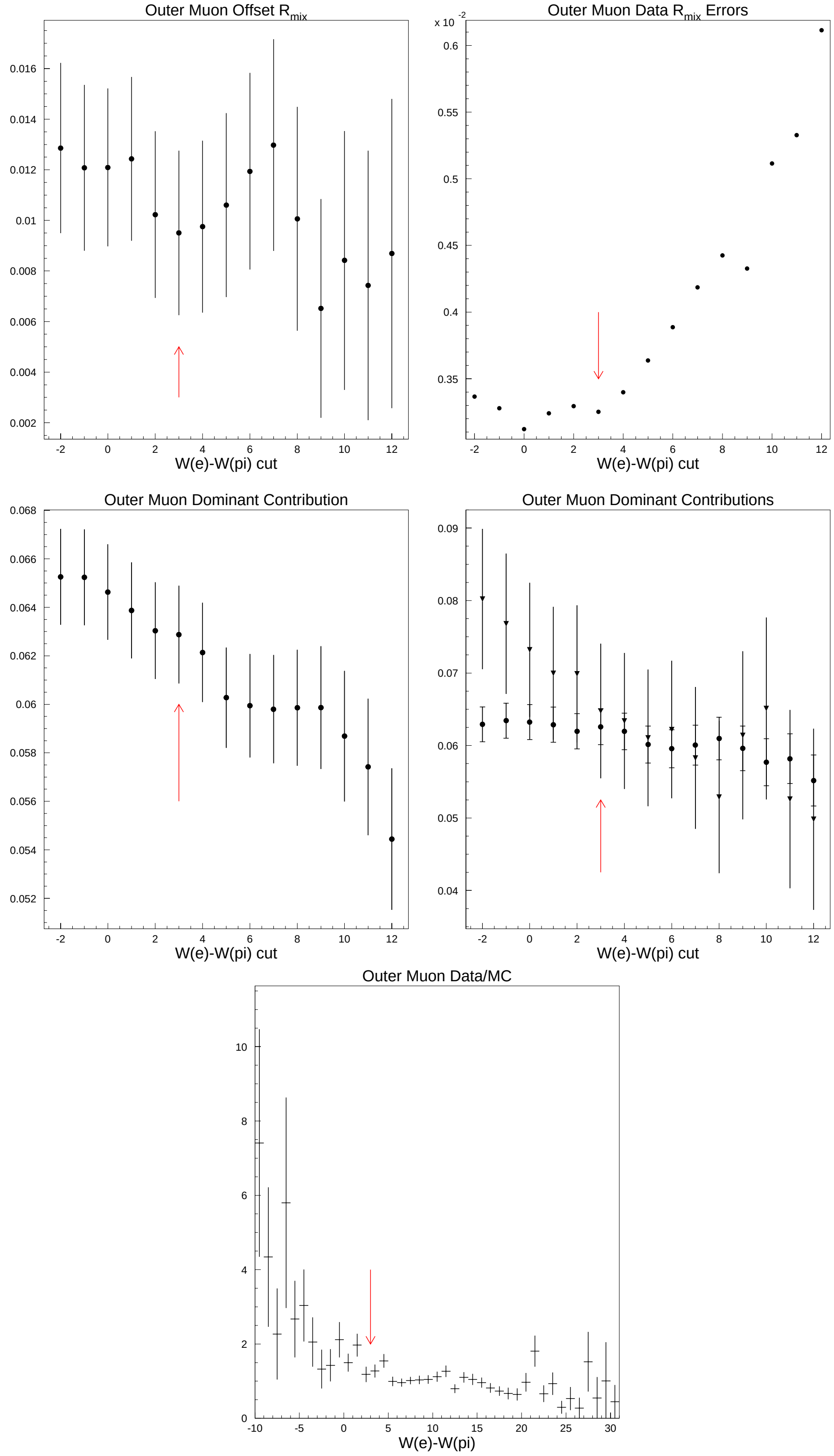


Figure 40: Evolution of fit parameters and data to Monte Carlo ratio as a function of $W(e) - W(\pi)$ cuts. The outer muon sample is illustrated here.

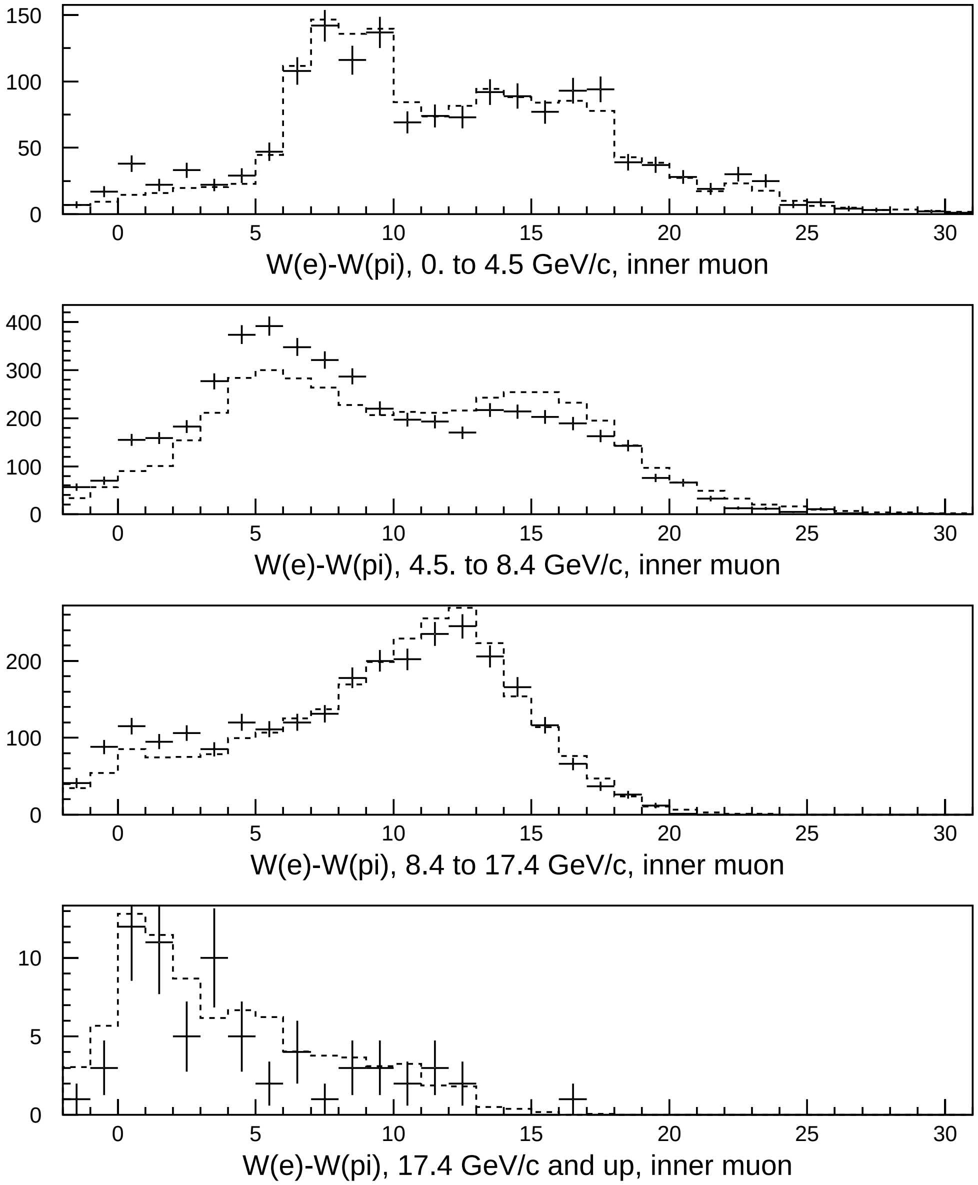


Figure 41: $W(e) - W(\pi)$ distributions for pion candidates split into pion momentum threshold regions. Error bars indicate data and the dashed histograms are Monte Carlo. The inner muon sample is shown here. The pion momentum cut was eliminated for these plots to increase the statistics at low momentum.

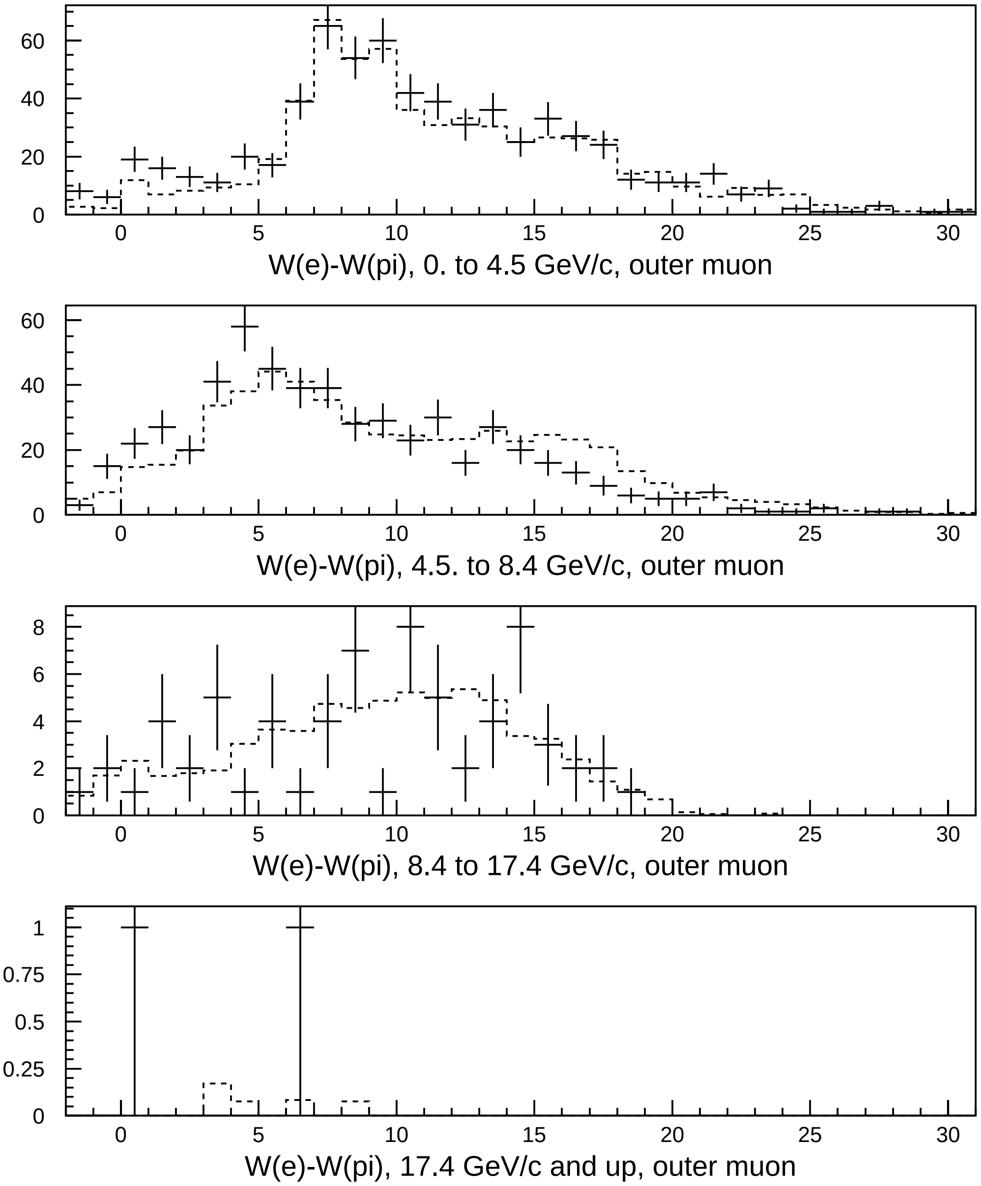


Figure 42: $W(e) - W(\pi)$ distributions for pion candidates split into pion momentum threshold regions. Error bars indicate data and the dashed histograms are Monte Carlo. The outer muon sample is shown here.

VI.7.2 Optimized Cuts

Results from optimizing cuts are presented here. Cuts were selected based on the procedure illustrated above.

Duplicate candidates in an event were eliminated after applying the minimal cuts (Table 6) but before applying additional cuts. (Eliminating duplicates after all cuts is about two percent more efficient in keeping events, but makes trying different cuts complicated.) Duplicate pion candidates for the same secondary are eliminated by selecting the candidate giving the lowest `iso3` value. It is rarely the case that multiple secondaries for the same pion candidate are selected. This rare instance is handled by selecting the secondary with the highest secondary vertex confidence level. Eliminating duplicates increases the signal to background ratio by about 15% (where "signal" here is $D^{*+} \to D^0\pi^+$ with the correct pion selected). Matching SSD tracks to (true) MC particle trajectories indicates this is at least 80% accurate in selecting the correct pion in a $D^{*+} \to D^0\pi^+$ event.

After eliminating duplicates, a narrower mass range cut was applied: $0.138 \leq M_{D^{*+}} - M_{D^0} < 0.165\,\text{GeV/c}^2$. Including more of the "tail" beyond $0.165\,\text{GeV/c}^2$ degrades the agreement between data and MC in the $M_{K\mu}$ distribution. If instead, the narrower mass range is used before eliminating duplicate candidates, an almost identical RS $M_{D^{*+}}$ distribution is obtained in data.

After applying minimal cuts, the isolation variables are the most effective cut variables in this analysis. The ranking discussed in Appendix E indicates `iso3` gives the best discrimination between signal and background in MC. Unfortunately `iso3` cuts eliminate a lot of signal as well. Consequently, `iso3` is only used for eliminating duplicate pion candidates in an event.

Table 7: Additional cuts obtained after optimizing with cut scans.

$0.138 \leq M_{D^{*+}} - M_{D^0} < 0.165\,\mathrm{GeV}/c^2$
$0.9 \leq M_{K\mu} < 1.7\,\mathrm{GeV}/c^2$ for inner muons
$1.1 \leq M_{K\mu} < 1.7\,\mathrm{GeV}/c^2$ for outer muons
$|\vec{p}_\pi| > 3.5\,\mathrm{GeV}/c$ for inner muons
$W(e) - W(\pi) > 3$ for pion candidates
$|\vec{p}_\mu| > 12\,\mathrm{GeV}/c$ for inner muons
$\mathrm{CL}_\mu > 0.2$ for outer muons
`iso2` < 0.003 for inner muons
`iso2` < 0.09 for outer muons
`iso1` < 0.001
$L/\sigma_L > 7$
number of sigma out-of-target of secondary > 1
run number ≥ 6577[a]

---

[a] Very early muon data are not well modeled by the MC.

The three isolation variables were defined in Sec. IV.3.2. Here is a brief summary:

- `Iso1` requires the K and μ candidates be inconsistent with coming from the primary vertex.
- `Iso2` requires all tracks except the K candidate, μ candidate, and tracks from the primary be inconsistent with coming from the secondary vertex.
- `Iso3` requires all tracks except the K, μ, and π candidate be inconsistent with coming from the secondary vertex.

The additional optimal cuts are shown in Table 7.

VI.7.3 Optimized Errors

For optimizing cuts, separate fits were made to the inner and outer muon samples. With the final cuts, the fitted value of r_{mix} for the inner sample is $(-1.17^{+1.00}_{-0.91}) \times 10^{-3}$

Table 8: Adaptively binned goodness-of-fit chi-squares for the merged sample. The high bias algorithm was used in these fits.

	χ^2	N_{dof}	Goodness-of-fit CL
RS mass	58.14	46	0.108
WS mass	45.99	44	0.390
RS *ct*	21.20	18	0.269
WS *ct*	17.08	17	0.449
RS *ct* vs mass	295.1	289	0.390
WS *ct* vs mass	67.82	66	0.415

and the fit to the outer sample gives $(2.66^{+3.25}_{-2.81}) \times 10^{-3}$ (statistical errors only). The fit to the merged sample gives $r_{\mathrm{mix}} = (-0.340^{+0.934}_{-0.867}) \times 10^{-3}$. These results and the results in the next chapter use the high bias bin merging algorithm (Sec. VI.4.1). The low bias algorithm is used for the final results, which are presented in the Conclusions chapter.

VI.7.4 Goodness-of-Fit with Optimized Cuts

The merged sample gives the adaptively binned chi-squares listed in Table 8. The χ^2 for the *ct* vs mass histograms used the adaptive binning described in Sec. VI.5. For the remaining histograms, if a bin has fewer than 15 entries predicted from the combined MC shape, it is merged with an adjacent bin before computing the χ^2 and N_{dof}. The confidence levels shown all assume Gaussian errors. As mentioned in Sec. VI.5, this underestimates the CL for the *ct* vs mass histograms. The number of degrees of freedom, N_{dof}, for the one dimensional histograms was increased by one to account for the fits using both mass and time dimensions.

CHAPTER VII

SYSTEMATIC STUDIES

In the previous chapter, a number of mostly qualitative checks were made of the Monte Carlo (MC) while selecting cuts. In this chapter, more quantitative comparisons are made. In variables where there appears to be statistically significant disagreement between data and MC, the data and MC are split into samples. The fit is repeated with each of the samples left out one at a time. The splitting procedure is repeated using 40 random mixing-free MC subsamples to determine the spread of r_{mix} values expected from statistical variations alone. The frequency of large deviations in the data is consistent with statistical variations alone. Consequently no systematic errors are assessed from this procedure.

Systematic errors are evaluated primarily by varying uncertain quantities in the MC and repeating the fit. Bias in the fitting procedure is also assessed by simulating the data and MC with smoothed MC shapes serving as parent distributions. A positive bias in the fit of roughly 0.4 times the statistical error was discovered after finalizing all cuts and computing systematic errors. The improved adaptive binning algorithm algorithm (described in Sec. VI.4.1) eliminates nearly all the bias. The final results use the improved algorithm, but systematic errors were determined with the old algorithm.

VII.1 Cross Validation

Verification of the MC proceeds in roughly two stages. In the first stage, statistical comparisons are made between data and MC distributions. In the second stage, the effect of deviations between data and MC on the measurement is determined. The first stage is used to determine which variables should be considered for the second stage.

Data and MC distributions were compared in a large number of variables. For each variable, the match was assessed using a Kolmogorov-Smirnov (KS) test [51, 55, 56]. The KS test is intended for unbinned data. Applying it to binned data tends to overestimate the confidence level of the match. The results shown here use 10,000 bins, except for variables that are inherently discrete such as the primary vertex multiplicity. Tables 9 and 10 list the matches from worst to best. Somewhat arbitrarily, only matches that are worse than the proper time, ct, are considered for the second stage.

To assess the effect of deviations on the fitted value of r_{mix}, a "leave-one-out" cross validation procedure is used. First each variable is split into two to eight regions, then each region is left out one at a time and the fit repeated. The regions are determined by dividing data histograms by MC histograms for each variable and looking for roughly level regions. Each region is fit to a flat line and the boundaries of the region are adjusted to give a reasonable χ^2 for the flat line fit.

This selection procedure is illustrated for the $W(e) - W(\pi)$ variable in Fig. 43. Bins containing fewer than 15 data entries are not used in the line fits. The regions selected for the $W(e) - W(\pi)$ variable are more complicated than the regions for other variables, since the regions are divided in pion momentum as well. For $4.5 \leq |\vec{p}_\pi| <$

Table 9: Results of Kolmogorov-Smirnov tests comparing data and Monte Carlo for inner muon candidates.

Variable	CL	Variable	CL
rs no. σ pri. in target	0.00000	ws q^2	0.20508
rs no. PWC tracks	0.00000	ws no. PWC tracks	0.23249
rs $W(e) - W(\pi)$, π can.	0.00000	ws $W(\pi) - W(K)$, K can.	0.27045
ws $W(e) - W(\pi)$, π can.	0.00000	ws `iso3`	0.29420
rs $\lvert\vec{p}_\mu\rvert$	0.00000	rs z pos. sec. vertex	0.30886
rs no. pri. tracks	0.00021	ws no. σ sec. out of target	0.32640
rs angle betw. $\vec{p}_\pi$ and D^0	0.00570	ws $\lvert\vec{p}_\mu\rvert$	0.40302
rs $\lvert\vec{p}_\pi\rvert$	0.00714	ws $\lvert\vec{p}_{K\mu}\rvert$	0.42574
ws z pos. pri. vertex	0.00912	rs L	0.42643
rs luminosity run period	0.02109	ws $\lvert\vec{p}_\pi\rvert$	0.45463
ws IM CL	0.02526	ws no. σ pri. in target	0.46515
ws L/σ_L	0.02696	rs no. σ sec. out of target	0.51403
rs q^2	0.03007	rs `iso1`	0.51678
rs D^* mass	0.03464	ws D^* mass	0.54764
ws z pos. sec. vertex	0.04067	rs CL sec. vertex	0.60675
rs K-μ mass	0.04318	ws K-μ mass	0.60942
rs L/σ_L	0.05546	ws angle betw. $\vec{p}_{K\mu}$ and D^0	0.71102
rs $\lvert\vec{p}_{K\mu}\rvert$	0.05685	ws no. pri. tracks	0.77488
rs angle betw. $\vec{p}_{K\mu}$ and D^0	0.09060	rs ct	0.79450
rs D^* momentum	0.09917	rs $W(\pi) - W(K)$, K can.	0.81562
rs IM CL	0.10361	ws D^* momentum	0.90378
rs z pos. pri. vertex	0.12297	ws angle betw. $\vec{p}_\pi$ and D^0	0.92807
ws luminosity run period	0.12630	ws CL sec. vertex	0.97622
ws ct	0.15656	rs `iso2`	0.99349
rs `iso3`	0.16627	ws `iso1`	0.99726
ws L	0.17020	ws `iso2`	0.99729

Table 10: Results of Kolmogorov-Smirnov tests comparing data and Monte Carlo for outer muon candidates.

Variable	CL	Variable	CL
rs OM thickness	0.00038	rs q^2	0.44114
ws $W(e) - W(\pi)$, π can.	0.00426	rs `iso3`	0.46386
rs $\lvert\vec{p}_\mu\rvert$	0.00875	rs K-μ mass	0.47409
ws luminosity run period	0.01764	rs no. σ sec. out of target	0.47926
rs $W(e) - W(\pi)$, π can.	0.02509	ws q^2	0.48404
ws OM thickness	0.04065	rs $\lvert\vec{p}_\pi\rvert$	0.49612
rs angle betw. $\vec{p}_\pi$ and D^0	0.09282	rs no. pri. tracks	0.53460
rs z pos. pri. vertex	0.11810	rs no. σ pri. in target	0.56036
rs z pos. sec. vertex	0.12321	ws no. PWC tracks	0.59269
rs $W(\pi) - W(K)$, K can.	0.14981	ws D^*	0.64764
ws angle betw. $\vec{p}_{K\mu}$ and D^0	0.18063	rs angle betw. $\vec{p}_{K\mu}$ and D^0	0.65831
rs ct	0.18431	ws D^* momentum	0.66943
rs no. PWC tracks	0.19317	ws $W(\pi) - W(K)$, K can.	0.69476
ws `iso3`	0.19795	rs L	0.72446
ws no. σ sec. out of target	0.23987	ws `iso2`	0.75015
rs D^* momentum	0.24964	rs L/σ_L	0.75085
rs luminosity run period	0.25798	ws ct	0.76651
ws $\lvert\vec{p}_{K\mu}\rvert$	0.25924	ws CL sec. vertex	0.77352
rs `iso1`	0.27005	ws K-μ mass	0.78118
ws $\lvert\vec{p}_\pi\rvert$	0.27055	ws L	0.82036
ws $\lvert\vec{p}_\mu\rvert$	0.27208	rs CL sec. vertex	0.90144
ws no. σ pri. in target	0.28789	rs D^* mass	0.96564
ws angle betw. $\vec{p}_\pi$ and D^0	0.31069	ws z pos. pri. vertex	0.98656
rs $\lvert\vec{p}_{K\mu}\rvert$	0.31665	ws no. pri. tracks	0.99508
ws L/σ_L	0.36955	ws z pos. sec. vertex	0.99748
ws OM CL	0.39764	rs `iso2`	0.99790
rs OM CL	0.40722	ws `iso1`	0.99820

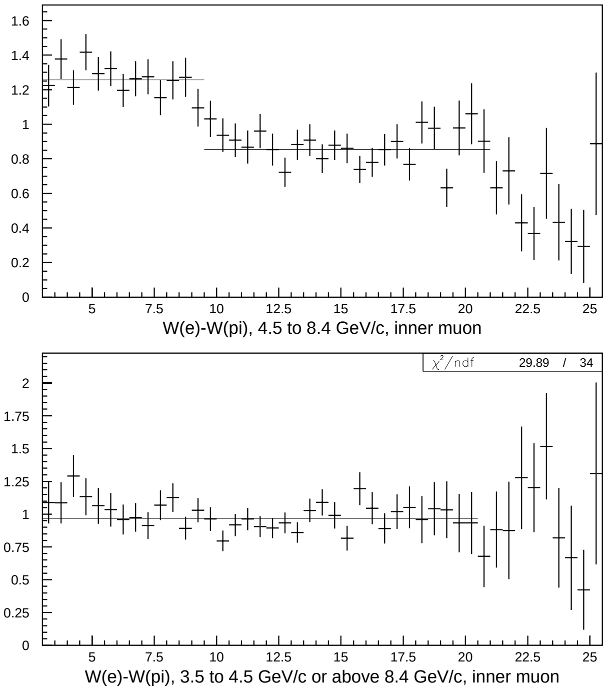


Figure 43: For cross validation, data over Monte Carlo ratio plots are first split into roughly level regions. This is illustrated here for $W(e) - W(\pi)$ for pion candidates in the inner muon sample. Three $W(e) - W(\pi)$ regions are selected from two pion momentum regions. The horizontal lines are results of fits. High $W(e) - W(\pi)$ bins are excluded from the line fits since the statistics are poor there.

$8.4\,\mathrm{GeV}/c$, data and MC are split into two samples at $W(e) - W(\pi) = 9.5$. In the remaining momentum region, no split is made.

To estimate the amount of variation due to statistics, fits were made to 40 random mixing-free MC subsamples (the full MC was still used for the shapes). Each event in a given subsample is independent (sampling without replacement). The leave-one-out procedure was repeated for each of the 40 subsamples, using the same regions as the data. The fraction of MC subsample shifts exceeding data shifts is an approximate confidence level for the shift. The data shifts and MC subsample variation for $W(e) - W(\pi)$ is summarized in Table 11. The shifts in r_{mix} in data could easily be due to

Table 11: Data shifts and RMS MC shifts for $W(e) - W(\pi)$ in the inner muon samples. The fraction of MC subsample shifts exceeding the data shift is denoted as "CL."

Excluded region	Fitted r_{mix}	Data shift	RMS MC shift	CL
$\Delta W \leq 9.5$ and $4.5 \leq \lvert\vec{p}_\pi\rvert < 8.4\,\text{GeV}/c$	$(-1.83^{+1.10}_{-1.01}) \times 10^{-3}$	-0.658×10^{-3}	0.566×10^{-3}	0.25
$\Delta W > 9.5$ and $4.5 \leq \lvert\vec{p}_\pi\rvert < 8.4\,\text{GeV}/c$	$(-0.05^{+1.37}_{-1.25}) \times 10^{-3}$	1.122×10^{-3}	0.837×10^{-3}	0.225
$\lvert\vec{p}_\pi\rvert \geq 8.4\,\text{GeV}/c$ or $3.5 < \lvert\vec{p}_\pi\rvert < 4.5\,\text{GeV}/c$	$(-1.54^{+1.42}_{-1.28}) \times 10^{-3}$	-0.368×10^{-3}	1.008×10^{-3}	0.85
none	$(-1.17^{+1.00}_{-0.91}) \times 10^{-3}$	0	0	—

statistical fluctuations, but the large RMS of the MC shifts make it impossible to be certain. (The "RMS" here is the square root of the mean of the square shifts.)

The shifts in r_{mix}, RMS of MC shifts, and CL for the remaining variables are summarized in Tables 12 and 13 (separate fits are made for inner and outer muons). Ignoring correlations between the variables, the mean and standard deviation of the shift CL's is 0.516 and 0.306 respectively for the inner muon sample (67 regions), and 0.450 and 0.323 respectively for the outer muon sample (29 regions).[1] Even in regions where there may be significant shifts, the RMS is too large to compute an accurate systematic error from this study. How systematic errors are computed is the subject of the next section.

VII.2 Systematic Errors

The methods used to compute the systematic errors vary in details, but the common theme is to shift variables by their error and observe how this shift effects the

---

[1] If the MC and data are statistically consistent (i.e., have the same parent distributions) the CL should follow a uniform distribution, which on average has a mean of 0.5 and standard deviation of $1/\sqrt{12} \doteq 0.289$.

Table 12: Leave-one-out cross validation for inner muons. The fraction of MC subsample shifts exceeding the data shift is denoted as "CL." 40 random subsamples from a 10.8 times FOCUS MC were used.

Variable	Data shift	RMS MC shift	CL	Data shift	RMS MC shift	CL
no. σ pri. in target	-5.39×10^{-5}	9.74×10^{-5}	0.575	0.000100	0.000299	0.775
	-0.000609	0.000576	0.25	-2.63×10^{-5}	0.000191	0.9
	0.001210	0.000885	0.2			
no. PWC tracks	-0.000309	0.000322	0.4	-0.000126	0.000252	0.575
	-0.000125	0.001679	0.85	0.000262	9.65×10^{-5}	0.025
	0.000340	0.000388	0.375			
$W(e) - W(\pi)$, π can.	-0.000658	0.000566	0.25			
	0.001122	0.000837	0.225			
	-0.000368	0.001008	0.85			
$\|\vec{p}_\mu\|$	0.000108	0.000511	0.875			
	0.000568	0.000759	0.65			
	-0.000203	0.000840	0.85			
	-0.000290	0.000205	0.2			
no. pri. tracks	-0.000515	0.000243	0.05	0.000169	0.000354	0.65
	0.000955	0.000666	0.225	-3.97×10^{-6}	0.000243	1.
	-0.000993	0.000638	0.125			
angle betw. $\vec{p}_\pi$ and D^0	-0.001918	0.002747	0.35			
	0.000201	0.000615	0.75			
$\|\vec{p}_\pi\|$	0.000390	0.001584	0.825			
	-0.000249	0.000297	0.55			
z pos. pri. vertex	6.20×10^{-5}	0.000257	0.925	-0.000113	0.000201	0.475
	0.000234	0.000316	0.55	0.000596	0.000535	0.3
	-0.000206	0.000155	0.125	-0.000445	0.000711	0.475
	0.000206	0.000515	0.675	-3.35×10^{-5}	0.000118	0.8
luminosity run period	0.000106	0.000481	0.775			
	-0.000931	0.000494	0.05			
	0.001111	0.000949	0.25			
	0.000444	0.000356	0.15			
IM CL	0.000522	0.000351	0.175			
	0.000244	0.000306	0.375			
	0.000460	0.001056	0.675			
	-0.000369	0.000586	0.6			
L/σ_L	0.000199	0.000645	0.825			
	0.000478	0.000488	0.325			
	0.000662	0.000377	0.05			
	-5.86×10^{-5}	0.000824	0.975			
q^2	-0.000465	0.000509	0.3			
	0.002752	0.002385	0.225			
D^* mass	-0.000602	0.001714	0.725			
	-3.52×10^{-5}	0.000417	0.925			
	-0.000178	0.000120	0.15			
z pos. sec. vertex	-5.07×10^{-5}	0.000223	0.85	-0.000415	0.000177	0.
	0.000251	0.000327	0.375	0.000491	0.000570	0.325
	0.000444	0.000441	0.35	3.72×10^{-5}	0.001036	1.
K-μ mass	0.000127	0.001460	0.9			
	-0.000876	0.000897	0.325			
$\|\vec{p}_{K\mu}\|$	0.000280	0.001151	0.9			
	-0.000740	0.000960	0.5			
angle betw. $\vec{p}_{K\mu}$ and D^0	6.10×10^{-5}	0.000903	0.975			
	0.000180	0.000767	0.9			
	-5.29×10^{-5}	0.000510	0.875			
D^* momentum	0.000793	0.001009	0.45			
	-0.000389	0.000661	0.575			
	-0.000194	0.000114	0.075			

Table 13: Leave-one-out cross validation for outer muons. The fraction of MC subsample shifts exceeding the data shift is denoted as "CL." 40 random subsamples from a 10.8 times FOCUS MC were used.

Variable	Data shift	RMS MC shift	CL
OM thickness	-0.003831	0.001602	0.025
	0.001647	0.001479	0.3
	0.003519	0.001948	0.025
$W(e) - W(\pi)$, π can.	0.001009	0.001023	0.35
	0.000807	0.001388	0.55
	-0.007420	0.003451	0.025
$\|\vec{p}_\mu\|$	-0.004911	0.004085	0.225
	0.002208	0.002053	0.3
luminosity run period	0.000472	0.001448	0.75
	0.000236	0.0012656	0.9
	0.000411	0.001695	0.85
	0.001535	0.001466	0.325
	0.000301	0.000717	0.65
angle betw. $\vec{p}_\pi$ and D^0	0.012518	0.008889	0.15
	-0.004707	0.001778	0.
z pos. pri. vertex	0.000181	0.001176	0.925
	0.000494	0.001477	0.8
	0.000791	0.001262	0.6
	-0.000561	0.001408	0.675
	-0.000352	0.000426	0.475
	-0.000217	0.001883	0.9
z pos. sec. vertex	0.000842	0.000668	0.2
	0.008459	0.009909	0.35
$W(\pi) - W(K)$, K can.	0.001418	0.001192	0.25
	-0.003580	0.006616	0.65
	0.000206	0.000753	0.825
angle betw. $\vec{p}_{K\mu}$ and D^0	0.006036	0.003253	0.05
	-0.004570	0.002115	0.025
	0.000121	0.000949	0.9

Table 14: Systematic error summary.

Source	Value
Branching ratios	$(+1.96, -1.93) \times 10^{-4}$
R_{DCS} and y'	$(+0.30, -1.50) \times 10^{-4}$
Length scale	$\pm 0.07 \times 10^{-4}$
Momentum scale	$\pm 0.37 \times 10^{-4}$
Nominal D^0 mass	$\pm 0.49 \times 10^{-4}$
OM efficiency	$\pm 0.05 \times 10^{-4}$
Fit shape weights	$\pm 0.43 \times 10^{-4}$
Fit bias	$(+0.19, -0.23) \times 10^{-4}$

fitted value of r_{mix}. The change in r_{mix} is the systematic error associated with the shifted variable. Where possible, variables are shifted both in positive and negative directions, so the error can be asymmetric in general. Correlations between the variables are ignored, and the positive and negative errors are separately added in quadrature to obtain the total positive and negative error.

Except where otherwise noted, the systematic errors were assessed for fits made with the high bias bin merging algorithm described in Sec. VI.4.1. Once corrected, the high and low bias results are quite close, so it is sufficient to use systematic errors computed for the high bias fit with the low bias fit. The bias corrected result for the low bias fit is $(-7.46^{+9.89}_{-9.30}) \times 10^{-4}$; the bias corrected result for the high bias fit is $(-7.03^{+9.34}_{-8.67}) \times 10^{-4}$ (the errors shown only include statistical errors from the fit). In each case, the statistical error on the bias correction is about $\pm 0.22 \times 10^{-4}$ (the same fluctuated samples were used for both corrections).

Table 14 summarizes the values of the systematic errors.

VII.2.1 Branching Ratio Corrections

All of the shapes used in the fits are weighted to correct branching ratios used by the MC. The largest systematic errors in this analysis arise from branching ratio uncertainties. Values compiled by the Particle Data Group are used, except where noted otherwise.

Table 15 lists the decays for which branching ratios are adjusted. These include all significant sources which give rise to an excess in the signal region ($M(D^{*+}) - M(D^0) \leq 0.165\,\mathrm{GeV}/c^2$) of a ten times FOCUS MC sample.[2]

All the MC histograms used in the fit are filled with and without event-by-event weighting. Each weighting function, w_{ji}, is found by dividing the weighted histogram by the histogram without weighting. The event-by-event weighting factor is $BR_{\mathrm{PDG}}/BR_{\mathrm{MC}}$. (The event is weighted twice if more than one D in the event decays as one of the modes from Table 15.) Two D^0 decays significant in this analysis are not in the most recent PDG listing. The first is $D^0 \to K^{*-}\mu^+\nu$. Assuming $\Gamma(D^+ \to \bar{K}^{*0}\mu^+\nu) = \Gamma(D^0 \to K^{*-}\mu^+\nu)$,[3] E687 obtained [59]

$$\frac{\Gamma(D^0 \to K^{*-}\mu^+\nu)}{\Gamma(D^0 \to K^-\mu^+\nu)} = 0.62 \pm 0.07 \pm 0.09,$$

which translates to $BR(D^+ \to \bar{K}^{*0}\mu^+\nu) = 2.00\%$. A less restrictive assumption is

$$\frac{BR(D^0 \to K^{*-}e^+\nu)}{BR(D^+ \to \bar{K}^{*0}e^+\nu)} = \frac{BR(D^0 \to K^{*-}\mu^+\nu)}{BR(D^+ \to \bar{K}^{*0}\mu^+\nu)}, \tag{47}$$

from which

$$BR(D^0 \to K^{*-}\mu^+) \quad = \quad \frac{BR(D^0 \to K^{*-}e^+\nu)}{BR(D^+ \to \bar{K}^{*0}e^+\nu)} \times BR(D^+ \to \bar{K}^{*0}\mu^+\nu)$$

---

[2] Some of these sources are more significant for RS than WS and visa versa.

[3] This apparently follows from the semileptonic $\Delta I = 0$ rule [58] inherent in the Glashow-Iliopoulos-Maiani scheme of weak interactions [5].

Table 15: Main Monte Carlo contributions, Monte Carlo branching ratio (MC BR), Particle Data Group (except where noted) branching ratio (PDG BR), and PDG errors on branching ratio (PDG BR error). Modes are listed from most to least significant in the RS sample (the ranking is different for WS). $D^0 \to K^{*-}\mu^+\nu$ is not listed by the PDG.

Mode	MC BR (%)	PDG BR (%)	PDG BR error (%)
$D^0 \to K^-\mu^+\nu$	2.703	3.22	0.17
$D^0 \to K^{*-}\mu^+\nu$	1.585	1.85[a]	0.44
$D^0 \to K^-\pi^+\pi^0$	13.9	13.9	0.9
$D^0 \to K^-\pi^+\pi^0\pi^0$ [b]	13.98	15.4[c]	2.5
$D^0 \to K^-\pi^+$	3.85	3.83	0.09
$D^0 \to \pi^-\mu^+\nu$	0.364	0.37[d]	0.06
$D^0 \to K^{*-}\rho^+$	8.10807	6.1	2.4
$D^0 \to K^-K^+$	0.3820	0.425	0.016
$D^+ \to \bar{K}^{*0}\mu^+\nu$	4.949	4.4[e]	0.6
$D_s^+ \to \phi\mu^+\nu$	2.4711	2.0[f]	0.5

[a] Not listed by the PDG. See text for details.

[b] Not including $D^0 \to K^{*-}\rho^+$ which can decay to this final state.

[c] Not in most recent PDG listing. See text for details.

[d] Only $D^0 \to \pi^-e^+\nu$ is listed by the PDG. $BR(D^0 \to \pi e\nu) \doteq 1.01BR(D^0 \to \pi\mu\nu)$ is used, following E687 [57].

[e] Only peaks in RS when a K^- or π^+ from the $\bar{K}^{*0}$ is misidentified as the soft pion from a D^{*+}. Does not peak in WS.

[f] Actually an average of $\phi\mu\nu$ and $\phi e\nu$.

$$
\begin{aligned}
&= (0.4208 \pm 0.0815) \times (4.4 \pm 0.6)\% \\
&= (1.85 \pm 0.44)\%
\end{aligned}
$$

which is the value in Table 15.

The other important mode missing from the most recent PDG listing[4] is $D^0 \to K^-\pi^+\pi^0\pi^0$ which is omitted because the measurements do not exclude the presence of additional neutral pions. For the purposes of this analysis it is better to include this mode even if it is overestimated. The PDG listed three measurements of $D^0 \to K^-\pi^+\pi^0\pi^0$ in the 2000 edition of the *Review of Particle Physics* [60]. These three measurements are averaged using the PDG method for combining results with asymmetric errors. The three measurements are $0.149 \pm 0.037 \pm 0.030$ [61], 0.177 ± 0.029 [62], and $0.209^{+0.074}_{-0.043} \pm 0.012$ [63]. The average is $0.174^{+0.024}_{-0.022}$. Some of these decays arise from $D^0 \to K^{*-}\rho^+$, which is accounted for separately. Using isospin for the $K^{*-} \to K^-\pi^0$ gives

$$
\begin{aligned}
BR(D^0 \to K^{*-}(K^-\pi^0)\rho^+(\pi^+\pi^0)) &= \frac{1}{3}BR(D^0 \to K^{*-}\rho^+) \\
&= \frac{1}{3}(6.1 \pm 2.4)\% \\
&= (2.03 \pm 0.80)\%
\end{aligned}
$$

which is subtracted from the total $BR(D^0 \to K^-\pi^+\pi^0\pi^0)$:

$$
(17.4 \pm 2.4)\% - (2.03 \pm 0.80)\% = (15.4 \pm 2.5)\%
$$

which is the value in Table 15.

The effect of BR uncertainties was evaluated by shifting the BR corrections applied to the MC and repeating the fit with the modified MC shapes. To compute the

---

[4] URL: `http://pdg.lbl.gov`

Table 16: Shifts in r_{mix} resulting from one standard error shifts in branching ratios.

Mode	Positive BR shift	Negative BR shift
$D^0 \to K^-\mu^+\nu$	0.000138	-0.000151
$D^0 \to K^{*-}\mu^+\nu$	-0.000013	0.000023
$D^0 \to K^-\pi^+\pi^0$	-0.000027	0.000029
$D^0 \to K^-\pi^+\pi^0\pi^0$	-0.000061	0.000070
$D^0 \to K^-\pi^+$	-0.000000	0.000001
$D^0 \to \pi^-\mu^+\nu$	-0.000006	0.000007
$D^0 \to K^{*-}\rho^+$	-0.000060	0.000077
$D^0 \to K^-K^+$	-0.000003	0.000003
$D^+ \to \bar{K}^{*0}\mu^+\nu$	-0.000079	0.000082
$D_s^+ \to \phi\mu^+\nu$	-0.000013	0.000014

systematic error, the BR corrections were individually shifted by one standard error and the fit repeated for each. So there were ten fits (one for each mode) with positively shifted BR's, and another ten fits with negatively shifted BR's. Table 16 lists the shifts obtained from these fits.

VII.2.2 Verification of the Background Model

The backgrounds can be classified into two categories: those that form a peak in the D^{*+} mass distributions or those that do not. Those that do not peak are referred to as "continuum" backgrounds. Events containing no $D^{*+} \to D^0\pi^+$ are usually continuum. Also, $D^{*+} \to D^0\pi^+$ events with the wrong reconstructed π^+ candidate are continuum.

Backgrounds which peak in the D^{*+} mass distributions are almost entirely from misidentification of the D^0 daughters in $D^{*+} \to D^0\pi^+$ decays. The most significant of the peaking backgrounds were listed in Table 15. Misidentification which creates a peak in the RS sample is a relatively small effect since it results in an error in r_{mix} roughly proportional to r_{mix}. The use of the pion flavor tag eliminates most of the peaking background in WS since both D^0 daughters then have to be misidentified. This "double misidentification" is an additive error to r_{mix} and therefore potentially more significant.

Double misidentification in the WS signal can be included in the fit by allowing the WS $D^{*+} \to D^0\pi^+$ contribution to be fit using an additional parameter. This allows the component which peaks in the WS signal region to have a different level, relative to RS, than predicted by MC. Since double misidentification is indistinguishable from DCS, this is also a way to account for DCS contamination (a better way is discussed below in Sec. VII.2.5). The fitted r_{mix} goes from $r_{\text{mix}} = (-0.76^{+0.99}_{-0.93}) \times 10^{-3}$ to $r_{\text{mix}} = (-0.17^{+1.30}_{-1.13}) \times 10^{-3}$ when the WS contribution to $D^{*+} \to D^0\pi^+$ is fit using an additional parameter. (The fits for this comparison used the low bias fit.) This shift is entirely consistent with statistical variations. Fits to 2228 simulated data and Monte Carlo samples[5] indicate that a shift larger than that observed in data occurs 42% of the time.[6]

Since double misidentification is too small to determine with the final set of cuts, a better idea of how well the Monte Carlo models backgrounds can be obtained by loosening some of the cuts. Figures 44 and 45 compares data and MC for samples with a muon confidence level (CL_μ) below 1%. The peak at the high edge of the K-μ mass range comes from $D^0 \to K^-\pi^+$ and some $D^0 \to K^-K^+$. The peak just below that is from $D^0 \to K^-\pi^+\pi^0$. A clear WS peak appears in the D^{*+} mass distribution for the inner sample. The WS proper time distributions indicate that

---

[5] The simulated samples are described in Sec. VII.2.11.

[6] The same conclusion is obtained by adding yet another parameter for the RS $D^{*+} \to D^0\pi^+$ contribution, giving a total of four parameters. In this case, shifts larger than the data occur 48% of the time.

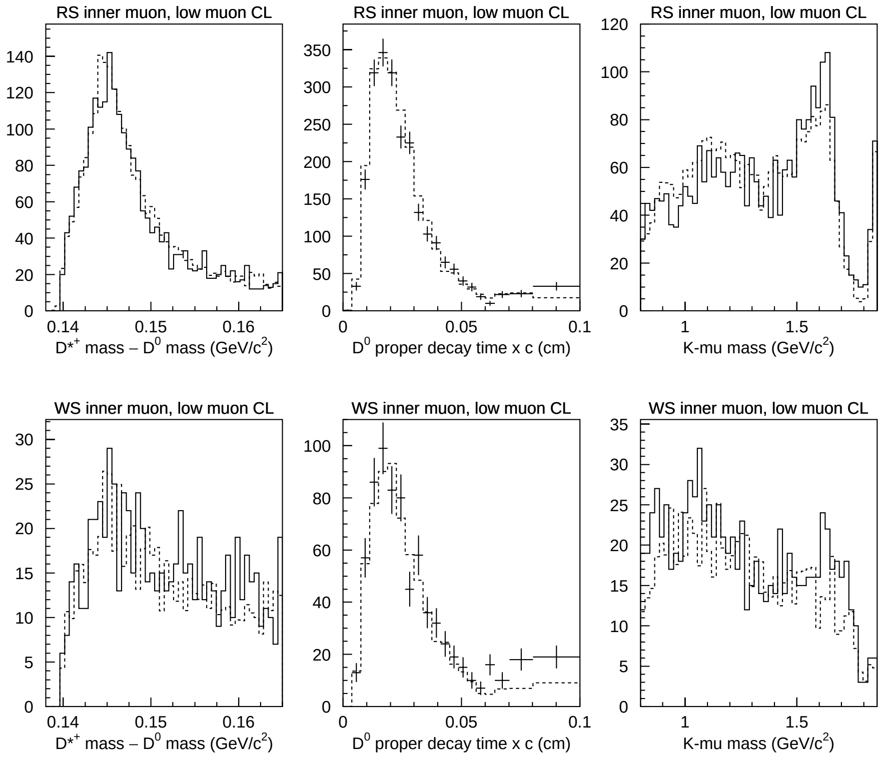


Figure 44: Comparisons of data and Monte Carlo for inner muon candidates with $\mathrm{CL}_\mu \leq 0.01$. Dashed histograms are Monte Carlo. The Monte Carlo level is scaled to the combined RS and WS data level. A wider K-μ mass range is used in the K-μ mass plots for the purposes of illustration.

there is some source of background at long proper time which is underestimated by the Monte Carlo.[7] This long proper time excess gets worse if the lower K-μ mass cut is relaxed. This is a strong indication that some of this excess results from D^0 decays with more than three daughters, which tend to populate the low K-μ mass region due to non-reconstructed energy. The branching ratios for many of these modes are poorly known, particularly those involving multiple neutral pions. Correcting for this excess would lead to a less conservative limit and is not attempted.

This low CL_μ data comes from the "global vertex" skim (also referred to as a

---

[7] The Monte Carlo distributions discussed here do not include mixing.

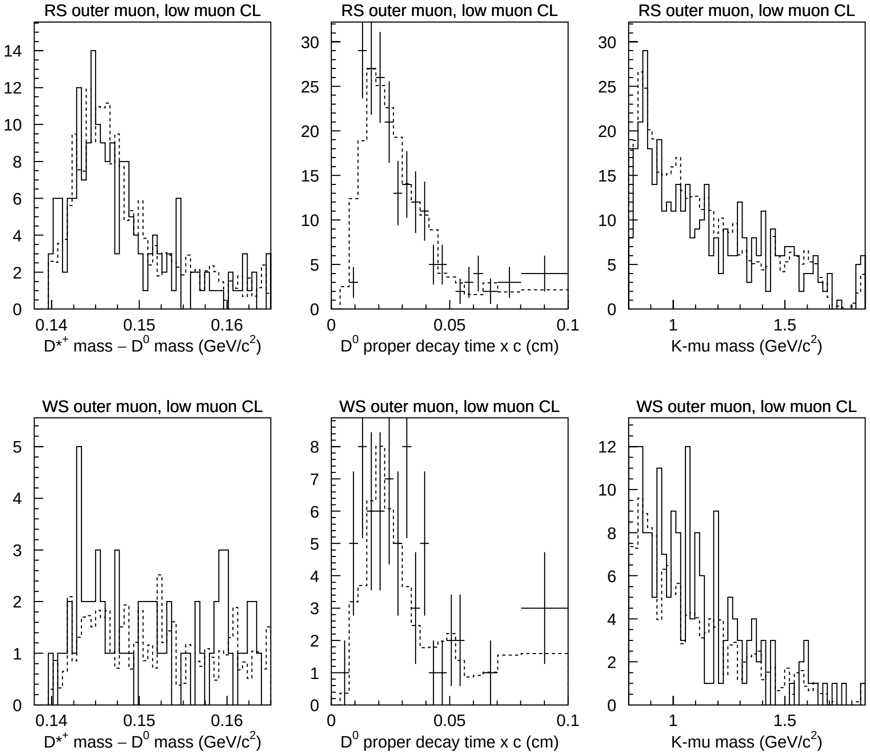


Figure 45: Comparisons of data and Monte Carlo for outer muon candidates with $\mathrm{CL}_\mu \leq 0.01$. Dashed histograms are Monte Carlo. The Monte Carlo level is scaled to the combined RS and WS data level. A wider K-μ mass range is used in the K-μ mass plots for the purposes of illustration.

topological, as opposed to a candidate driven skim). This skim first finds all combinations of two track vertices and then requires at least two such vertices be separated by greater than 4.5 sigma. In addition, the downstream vertex must contain at least one track with $W(\pi) - W(K) > 1$. All the usual cuts (listed in Tables 6 and 7) are applied in addition to the global vertex requirements, except for the CL_μ.

For inner muons, the requirement that there be no missing inner muon plane hits is still retained, although in retrospect it would have been better to have an acceptance requirement instead. Nearly all of the inner muon candidates with CL_μ below 1% are stubs (tracks with PWC hits only upstream of the second analysis magnet). In fact, an inner CL_μ is not evaluated for stubs—it is set to zero for stubs. In addition to $\mathrm{CL}_\mu \leq 1\%$ and no missing inner planes, the predicted thickness of material traversed in the outer muon system is required to be zero.[8]

With these requirements, stubs arise not only from wide angle particles which miss the downstream PWC's, but also from low angle particles which decay in flight. It should be noted that although the shapes and right sign to wrong sign ratios are well predicted by the MC for these tracks, the overall MC level is low by about a factor of 1.8 for inner muons (1.3 for efficiency corrected outer muons) when scaled to the level predicted from high confidence level data. The reason for this discrepancy has not been investigated.

Muon candidates containing missing inner muon planes or a momentum less than 12 Gev/c must pass the modified outer muon requirements. These requirements are $0 < \mathrm{CL}_\mu \leq 0.01$ and the usual thickness > 150 cm and momentum above 5 GeV/c.

Misidentification rates generally decrease with increasing momentum. The low and high CL_μ samples have different momentum distributions as shown in Figs. 46 and 47. The momentum dependence of misidentification is considered further below.

VII.2.3 Muon Misidentification

Particles can be misidentified as muons in a variety of ways: hadrons can "punch through" filters (described in Sec. V.2), a track can accidentally point to a muon hit, or a hadron can decay in flight to a muon. Misidentification rates are on the order of a percent, but vary significantly with momentum [49, 64]. Comparisons of data

---

[8] Due to a coding error, the outer muon thickness variable is not filled correctly when the thickness is zero. However, the outer muon CL is correctly set to zero in these cases. Consequently, the actual requirement is that the outer muon CL is zero, rather than the thickness.

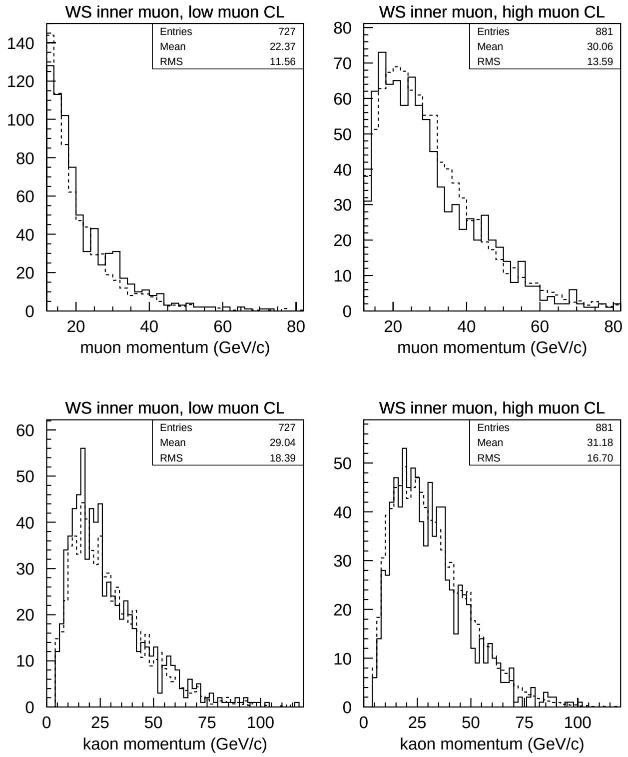


Figure 46: Comparisons of data and Monte Carlo momentum distributions for muon and kaon candidates in the wrong sign inner muon sample. Dashed lines are Monte Carlo. High and low CL_μ samples are shown.

and MC for reconstructed $D^+ \to K^-\pi^+\pi^+$ indicate the total single misidentification rate in MC is 7% lower than data for inner tracks and 30% lower than data for outer tracks [64].

Using the misidentification rates as a function of momentum from Ref. [64], the effect on r_{mix} due to the mismatch between data and MC was assessed. In MC, the particle type for muon candidates and their parents was determined by matching reconstructed SSD tracks to true generated particle trajectories. If a muon candidate was a pion or came from a pion, the event was weighted by the data to MC ratio of the misidentification rate as a function of reconstructed momentum. The fit was repeated with the corrected MC sample. The resulting shift in r_{mix} was $(-0.124\pm0.138)\times10^{-3}$.

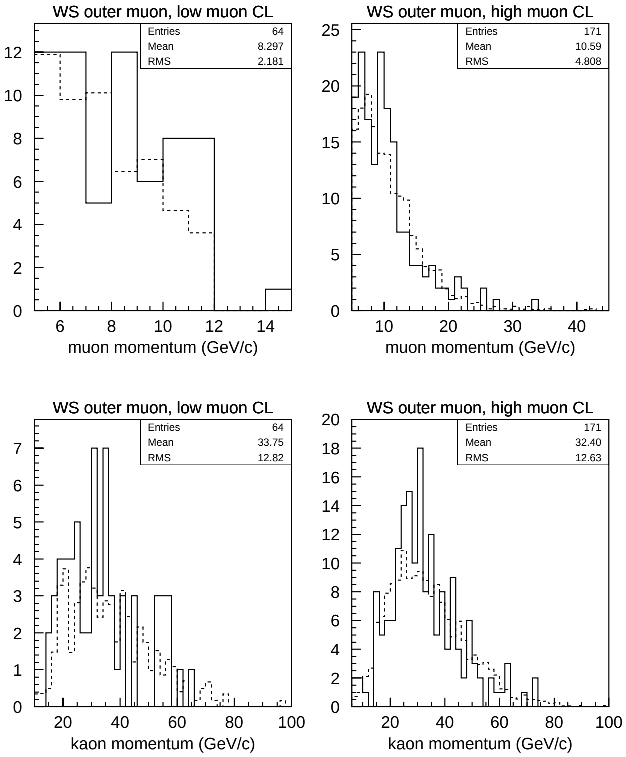


Figure 47: Comparisons of data and Monte Carlo momentum distributions for muon and kaon candidates in the wrong sign outer muon sample. Dashed lines are Monte Carlo. High and low CL$_\mu$ samples are shown.

The error on this number was found by varying the data to MC ratio by its statistical uncertainties. Since the statistical uncertainty in this correction is larger than the correction, and the correction would lead to a less conservative limit, no correction or systematic error is applied for muon misidentification.

VII.2.4 Kaon Misidentification

The kaon momentum distributions shown in Figs. 46 and 47 show good agreement between data and Monte Carlo.[9] Kolmogorov-Smirnov tests comparing data and MC for these distributions give probabilities of 38% (28%) for the inner (outer) muon

---

[9] The Monte Carlo distributions discussed in this section do not include mixing.

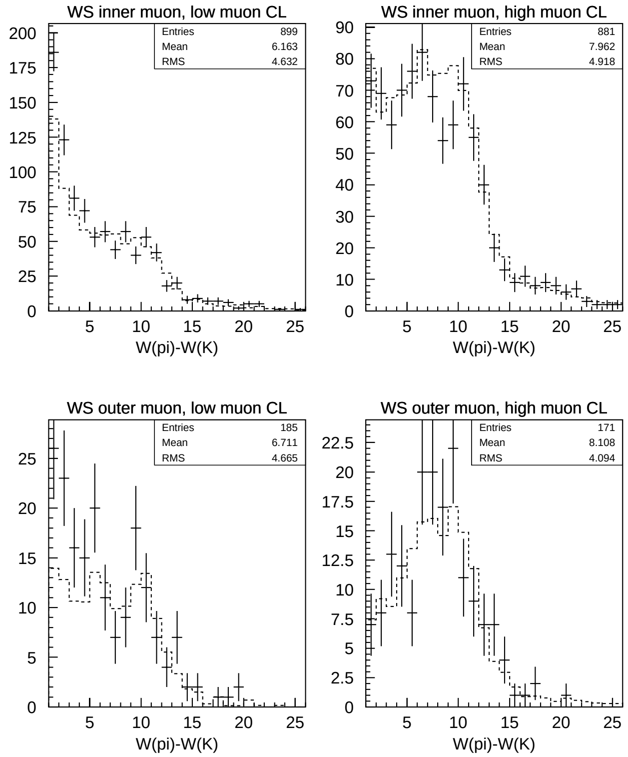


Figure 48: Comparisons of data and Monte Carlo $W(\pi) - W(K)$ distributions for kaon candidates in the wrong sign samples. Dashed lines are Monte Carlo. High and low CL$_\mu$ samples are shown for both inner and outer muons.

high CL$_\mu$ sample and 33% (18%) for the inner (outer) muon low CL$_\mu$ sample. (10,000 bins are used for all Kolmogorov-Smirnov tests in this thesis.) As shown in Fig. 48, the $W(\pi) - W(K)$ distributions for the kaon candidates agree well for the high CL$_\mu$ sample but not the low CL$_\mu$ sample. Kolmogorov-Smirnov tests for the $W(\pi) - W(K)$ distributions give probabilities of 43% (72%) for the inner (outer) muon high CL$_\mu$ sample and 0.23% (1.3%) for the inner (outer) muon low CL$_\mu$ sample. All of these comparisons are between wrong sign data and Monte Carlo. Requiring a higher $W(\pi) - W(K)$ does not improve the low CL$_\mu$ proper time distributions.

No correction or systematic error is assessed for kaon misidentification.

VII.2.5 DCS and DCS-Mixing Interference

Since about 10% of the accepted D^0 decays are non-semileptonic (with the final set of cuts), it is necessary to consider the effect of contamination from DCS and DCS-mixing interference. (DCS was briefly described in Sec. I.4 and Fig. 5.) A contribution from this contamination is included in the fit shapes to assess a systematic error.

As with mixing, DCS and DCS-mixing interference was not simulated in the MC sample used. For DCS, the analogous expression to Eq. (41) is

$$R_{\mathrm{DCS}}\, dN_{\mathrm{NDCS}} = dN_{\mathrm{DCS}}, \tag{48}$$

where N_{DCS} is the expected number of DCS decays measured, N_{NDCS} is the average number of non-DCS non-semileptonic D^0 signal measured and R_{DCS} is the fraction of DCS decays.[10] Likewise, for DCS-mixing interference [28]

$$\sqrt{R_{\mathrm{DCS}}}\, y'\, \Gamma\, t_g\, dN_{\mathrm{NDCS}} = dN_{\mathrm{INT}}, \tag{49}$$

where N_{INT} is the expected contribution of DCS-mixing interference (which can be negative) and the prime in y' denotes an unknown rotation of x and y arising from a possible strong phase between the DCS and Cabibbo favored amplitudes.

The DCS shape comes from all non-semileptonic D^0 MC events. The DCS-mixing interference shape was determined by weighting non-semileptonic D^0 MC events by Γt_g. To evaluate the effect of DCS and DCS-mixing interference, these shapes were included in the fit at fixed amounts. Tables 17 and 18 list the additional fit parameters added to account for DCS and DCS-mixing interference. CLEO's measurement [28], $R_D = (0.48 \pm 0.12 \pm 0.04)\%$ was used and the "or" of the one standard deviation

---

[10] Not all non-semileptonic D^0 decay modes are candidates for DCS, but at least 80% of the non-semileptonic modes that appear in the MC sample are candidates for DCS. To make this analysis simpler, all non-semileptonic D^0 decay modes are allowed in the DCS shapes.

Table 17: Description of the fit parameters including DCS related contamination. The two DCS related parameters are included in the fit at fixed amounts to assess a systematic error, but are otherwise set to zero.

Fit parameter	Description
q_1	Dominant $D^{*+} \to D^0\pi^+$
q_2	Dominant non $D^{*+} \to D^0\pi^+$
q_3	Mixing
q_4	DCS
q_5	DCS-mixing interference

Table 18: The fit is generalized to include contamination from DCS and DCS-mixing interference. Sixteen p_j "strengths" for the MC contributions in the likelihood function are determined from five fit parameters.

"Strength"	p_j value	Description
p_1	q_1	RS dominant $D^{*+} \to D^0\pi^+$
p_2	q_2	RS dominant non $D^{*+} \to D^0\pi^+$
p_3	q_3q_1	RS mixing $D^{*+} \to D^0\pi^+$
p_4	q_3q_2	RS mixing non $D^{*+} \to D^0\pi^+$
p_5	q_4q_1	RS DCS $D^{*+} \to D^0\pi^+$
p_6	q_4q_2	RS DCS non $D^{*+} \to D^0\pi^+$
p_7	q_5q_1	RS DCS-mixing interference $D^{*+} \to D^0\pi^+$
p_8	q_5q_2	RS DCS-mixing interference non $D^{*+} \to D^0\pi^+$
p_9	q_1	WS as above
$\vdots$	$\vdots$	
p_{16}	q_5q_2	

y ranges inferred from the CLEO [28] y' and FOCUS [29] y_{CP} values was used: $y = (-4.13$ to $5.55)\%$. The fit was done using 5.55% and repeated using -4.13%. In each case, $R_D = 0.48\%$ was used. (Hadronic mixing is included as signal and assumed to have the same value of r_{mix} as semileptonic mixing.)

VII.2.6 Length Scale

The length scale uncertainty has been studied extensively for lifetime measurements [65, 66]. Bench measurements determined the downstream microstrip (SSD) positions to better than $3\,\mu\text{m}$ [35]. The upstream microstrip (TSSD) positions were determined from vertices reconstructed from high statistics decay modes with the SSD and also with caliper/depth gauges. Run dependence of the positions was also investigated from high statistics decay modes. The length scale uncertainty is about 0.1%.

Scaling $L, ct, L/\sigma_L$, and number of sigma out of target by 1.001 in the MC sample, then reapplying cuts and repeating the fit results in an increase in r_{mix} of 0.7×10^{-5}. The above variables were modified with very loosely cut MC before applying final cuts. Unfortunately, a cut of $ct < 0.1\,\text{cm}$ was present even in the loosely cut sample. A sample with only $ct < 0.4\,\text{cm}$ is available but would likely take a couple of days to process. Scaling these variables by 0.999 decreases r_{mix} by 4.5×10^{-5}, but this is questionable due to the $ct < 0.1\,\text{cm}$ cut. Presently, the best estimate of the systematic error on r_{mix} due to the length scale uncertainty is $\pm 0.7 \times 10^{-5}$.

VII.2.7 Momentum Scale

The momentum scale uncertainty is inferred from the mass scale uncertainty (since the momentum is generally much higher than the mass, a linear relation between momentum uncertainty and mass uncertainty is a good approximation). The mass scale uncertainty has been studied for mass measurements [67] using the high statistics modes $D^0 \to K^-\pi^+$, $D^0 \to K^-\pi^+\pi^+\pi^-$, and $D^+ \to K^-\pi^+\pi^-$. The mass/momentum scale uncertainty is 0.07%.

Scaling all the momenta used in the analysis by 0.9993 and recomputing the D^* mass, K-μ mass, and ct results in a shift in r_{mix} of -3.7×10^{-5}. A correction of 1.0007 would decrease ct so is not used to assess the positive shift (because of the $ct < 0.1\,\text{cm}$ cut mentioned above in the length scale discussion). Currently, the best estimate of the systematic error on r_{mix} due to the momentum scale uncertainty is $\pm 3.7 \times 10^{-5}$.

VII.2.8 Nominal D^0 Mass

In computing the D^{*+} mass and D^0 proper decay time, a mass of $1.8646\,\text{GeV}/c^2$ is used. The PDG [60] gives an average D^0 mass of $(1.8645 \pm 0.0005)\,\text{GeV}/c^2$. The D^{*+} mass and D^0 proper decay time are recomputed using $1.8640\,\text{GeV}/c^2$ and $1.8650\,\text{GeV}/c^2$ and the fit repeated for each. The recalculation is done using information stored in ntuples rather than the original FOCUS data files, which results in some loss of precision. If the recalculation is done with the original mass, $1.8646\,\text{GeV}/c^2$, r_{mix} goes from -0.340×10^{-3} to -0.336×10^{-3}. The systematic error is computed as the shift from the recomputed value of r_{mix}. The resulting shifts are $+4.9 \times 10^{-5}$ and -17.1×10^{-5}. The negative shift may change with looser cuts on ct, so only the positive shift is used to estimate the error. (Other relevant loose cuts applied before recomputing are $M(D^{*+}) - M(D^0) < 0.243\,\text{GeV}/c^2$, $L/\sigma_L > 3$, and $M_{K\mu} > 0.8\,\text{GeV}/c^2$.)

VII.2.9 Muon Efficiencies

Efficiencies for a particular detector plane are generally estimated by computing the fraction of hits in the plane given that all other overlapping planes have hits. To minimize the effect of accidental hits from detector noise or other tracks, the hits can be required to belong to the same cluster and/or be within a certain radius of a projected track. This calculation assumes that the probabilities for registering a hit in each plane are statistically independent. For the sake of discussion, these will be referred to as "independent efficiencies." This assumption does not hold, for example, if some fraction of the time all the planes are dead. This could arise from deadtime in electronics shared by the planes, or from time needed for planes to recover from the passage of previous particles. The calculation of an independent efficiency only counts events with at most one missing plane so it gives the same value regardless of the correlated efficiency. At present, only independent efficiencies have been accounted for in the muon simulations.

Independent efficiencies for the MH arrays (inner muon) have been measured with halo muons [49] and both $D^+ \to K^{*0}\mu^+\nu \to (K^-\pi^+)\mu^+\nu$ and $J/\psi \to \mu^+\mu^-$ samples [64]. Independent efficiencies for the MH arrays are greater than 98%. Since analyses generally require 4 or 5 hits out of 6, this gives track efficiencies exceeding 99%. For the MH arrays, no inefficiency is simulated in the Monte Carlo.

A thorough study of independent efficiencies for the RPC's (outer muon) is given in Reference [64]. The efficiencies were studied with a variety of methods and samples: with and without clustering, requiring overlap with MH arrays, using $J/\psi \to \mu^+\mu^-$, $D^+ \to K^{*0}\mu^+\nu \to (K^-\pi^+)\mu^+\nu$, and $D^{*+} \to D^0\pi^+ \to (K^-\mu^+\nu)\pi^+$ samples. Position, trigger, and run-dependence were also studied. Separate efficiencies are simulated for each of the 24 RPC planes in six run periods. Efficiencies computed with a variety of methods are available to assess systematic errors. The efficiency varies considerably from plane to plane: 50% to 95%. The fractional systematic uncertainty on the efficiencies is about 5.5% [64].

There appears to be additional sources of inefficiency not accounted for in the Monte Carlo. This could be due to "correlated efficiencies" discussed above, incorrect wide-angle D decay distributions, acceptance problems, tracking inefficiencies, or some other problem. To account for this, the outer muon MC sample is multiplied by 0.772 ± 0.027 so that the number of outer muons relative to inner muons match between data and MC in reconstructed $D^{*+} \to D^0\pi^+ \to (K^-\mu^+\nu)\pi^+$ events.[11]

As we have seen in previous chapters, data and MC match very well for the inner and outer muon samples separately. It should also be possible to assess this additional inefficiency by reconstructing $J/\psi \to \mu^+\mu^-$ data with no outer muon hit requirement (geometric and energy acceptance only), but requiring the other muon be detected in the inner muon system. This has not been done at this time.

To determine the correction, separate fits are made to the inner and outer samples. The ratio of data to MC is returned by the dominant (non-mixing) fit parameter. The correction, ϵ, is the ratio of outer to inner muons for this fit parameter:

$$\epsilon = \frac{0.06287 \pm 0.00204}{0.08144 \pm 0.00102} = 0.7720 \pm 0.0269. \tag{50}$$

For simplicity, the larger (positive) `MINOS` error is used. Systematic errors from this correction arise from the statistical error on the correction. Repeating the fit once with $0.7720 + 0.0269$ and again with $0.7720 - 0.0269$ gives the one sigma range for this systematic error.

---

[11] When a track is identified as a muon by both the inner and outer detectors, it is classified as an inner muon. Only when a track fails the inner muon requirements is it considered a possible outer muon candidate.

VII.2.10 Fit Shape Weights

The MC fit shapes are weighted to account for BR corrections, mixing time dependence, and the OM efficiency correction. The weights are computed from the ratio of MC weighted event-by-event to the same MC without any weighting. The statistical error in a bin for event-by-event weighted MC is computed by the `HBOOK` histogramming package as

$$\text{Error}(i) = \sqrt{\sum_k W_{ki}^2}, \tag{51}$$

where W_{ki} is the weight of event k in bin i. Using the notation defined in Sec. VI.4.1, the extra error (RMS) in a bin due to the error in the weights is

$$\text{Extra error}(i) = \sqrt{\text{Error}(i)^2 - w_{ji}^2 a_{ji}}. \tag{52}$$

(Most of the error comes from error in a_{ji}, but this is already accounted for in the fit.) The effect of this extra error is found by adding Gaussian fluctuations to the event-by-event weighted histograms. The fluctuated event-by-event weighted histograms are then used to recompute the weights used in the fit. 2250 fluctuated MC sets were used to determine the spread in r_{mix} due to statistical errors in the weights.

VII.2.11 Fit Bias

A positive binning related bias of about 0.4 times the statistical error on r_{mix} was discovered after finalizing all cuts and computing systematic errors. Nearly all binning related bias is eliminated by adjusting the r_{min} parameter in the improved bin merging algorithm that was discussed in Sec. VI.4.1. If no bin merging is used, a negative bias of about 0.3 times the statistical error results.

Bias in the fit is determined using simulated data and MC generated from the same parent distributions. The parent distributions are formed from smoothed MC shapes. The smoothed shapes are Poisson fluctuated to generate simulated data and MC samples. The level of bias depends only weakly on the amount of mixing generated so no mixing is generated in computing the bias. 2250 simulated data and MC sets were generated to evaluate bias. For the low bias algorithm, the median of fitted r_{mix} values for the simulated samples is $(-12^{+19}_{-23}) \times 10^{-6}$; the sample mean is $(-8 \pm 23) \times 10^{-6}$. Since the fit errors are slightly asymmetric, the median is a better estimate.

As a cross check, bias was also investigated by randomly separating the full MC sample into simulated data and MC samples. This estimate gives results consistent with the Poisson fluctuated method, however the accuracy of such a sampling procedure is limited by MC statistics.

VII.2.12 Other Tests of the Fit

A mixing signal can be added to the simulated samples described in the previous section. The following table shows the sample means of fitted r_{mix} values for different simulated r_{mix} levels (the low bias bin merging algorithm is used). The two methods used for simulating data and MC are labeled "Poisson fluctuating" and "Random sampling." In MC events where there is both a D^0 and a $\overline{D}^0$ produced, the random sampling approach allows mixing in either D. In the Poisson fluctuating approach, mixing is only simulated in the reconstructed D. The mean positive fit errors are within 3% of the sample standard deviations (not shown).

Table 19: Mean fitted r_{mix} values for different levels of simulated mixing.

	Sample mean	
Input r_{mix}	Poisson fluctuating	Random sampling
0.	$(-8 \pm 23) \times 10^{-6}$	$(6 \pm 33) \times 10^{-5}$
0.001	$(1.031 \pm 0.024) \times 10^{-3}$	$(1.17 \pm 0.34) \times 10^{-3}$
0.01	$(1.0250 \pm 0.0031) \times 10^{-2}$	$(1.147 \pm 0.051) \times 10^{-2}$

VII.3 Improving the Analysis

There are at least three ways the analysis might be improved for journal publication. One way is to increase the 10.8 times FOCUS sample used for this thesis to a 20 times FOCUS sample. This might improve the statistical error by 20%. The other possibility is to include the semielectronic modes. This would give a potential factor of two increase in statistics. This might improve the statistical error by 40%. Finally, a more accurate limit could be computed based on simulated data (similar to what is done in computing the bias), rather than the Gaussian approximation used here. This is described more fully in Appendix C.

CHAPTER VIII

CONCLUSIONS

The results of this analysis are summarized below. Results from other experiments are also summarized and theoretical estimates are compared with experimental limits. The results of this thesis currently represent the world's best limit on r_{mix} in the D^0–$\overline{D}^0$ system. At present, there is no evidence for D^0–$\overline{D}^0$ mixing.

VIII.1 Experimental Results

The value of r_{mix} returned from the low bias fit is $(-7.6^{+9.9}_{-9.3}) \times 10^{-4}$. The fit is illustrated in Fig. 19. Including the bias correction and systematic errors, the result is $(-7.5^{+9.9,+2.1}_{-9.3,-2.6}) \times 10^{-4}$. In computing the limit, systematic and statistical errors are added in quadrature giving $(-7.5^{+10.1}_{-9.6}) \times 10^{-4}$.

The confidence intervals are computed by approximating the resolution as a Gaussian distribution with a standard deviation given by the positive error. The limit is computed with the approach recommended by Feldman and Cousins for a Gaussian measurement near a physical boundary. This approach was discussed in Sec. VI.6. The results are $r_{\mathrm{mix}} < 10.1 \times 10^{-4}$ at 90% confidence level and $r_{\mathrm{mix}} < 13.1 \times 10^{-4}$ at 95% confidence level.

Tables 20 and 21 list the most recently published experimental limits for r_{mix} along with this result. The Feldman-Cousins method (Table 20) for computing a limit is more conservative than the traditionally used method, since the Feldman-Cousins confidence intervals are guaranteed to be non-negative. For comparison, these

Table 20: Most recent limits for published measurements of r_{mix} and this result. The Feldman-Cousins method for a bounded Gaussian is used. The published values, which used a a less conservative method, are shown in Table 21. Only limits for versions of the fits assuming CP conservation are listed (except for the E791 hadronic fit which allows CP violation in the interference term).

Modes	Experiment	90% CL	95% CL
$D^0 \to K^-\mu^+\nu$	FOCUS (this thesis)	1.01×10^{-3}	1.31×10^{-3}
$D^0 \to K^-\mu^+\nu$ $D^0 \to K^-e^+\nu$	E791 [30]	6.0×10^{-3}	7.0×10^{-3}
$D^0 \to K^-\pi^+$ $D^0 \to K^-\pi^-\pi^+\pi^+$	E791 [27]	9.8×10^{-3}	10.9×10^{-3}
$D^0 \to K^-\pi^+$	CLEO [28]	See text.	
$D^0 \to K^-\pi^+$	ALEPH [68]	See text.	

Table 21: Most recently published experimental limits on r_{mix} and this result. The log-likelihood difference method is used: $\Delta \ln \mathcal{L} = 0.82$ for a 90% CL upper limit and $\Delta \ln \mathcal{L} = 1.31$ for a 95% CL upper limit. Only limits for versions of the fits assuming CP conservation are listed (except for the E791 hadronic fit which allows CP violation in the interference term).

Modes	Experiment	90% CL	95% CL
$D^0 \to K^-\mu^+\nu$	FOCUS (this thesis)	0.52×10^{-3}	0.87×10^{-3}
$D^0 \to K^-\mu^+\nu$ $D^0 \to K^-e^+\nu$	E791 [30]	5.0×10^{-3}	$\sim 6.1 \times 10^{-3}$
$D^0 \to K^-\pi^+$ $D^0 \to K^-\pi^-\pi^+\pi^+$	E791 [27]	8.5×10^{-3}	$\sim 9.8 \times 10^{-3}$
$D^0 \to K^-\pi^+$	CLEO [28]	See text.	
$D^0 \to K^-\pi^+$	ALEPH [68]	See text.	

results are also listed using the traditional method for computing a limit (Table 21). The traditional method locates the point were the log-likelihood decreases from the maximum by an amount determined by a corresponding Gaussian limit.[1]

A limit on r_{mix} can also be estimated from CLEO's $D^0 \to K^-\pi^+$ mixing measurement [28]. Their limits are given in terms of x'^2 and y', where the prime denotes an

---

[1] The Gaussian limit implies $\Delta \ln \mathcal{L} = 0.82$ for a 90% CL upper limit, $\Delta \ln \mathcal{L} = 1.31$ for a 95% CL upper limit, $\Delta \ln \mathcal{L} = 1.35$ for a 90% central confidence interval, and $\Delta \ln \mathcal{L} = 1.92$ for a 95% central confidence interval.

unknown rotation of x and y arising from a possible strong phase between the DCS and Cabibbo favored amplitudes. The likelihood contour for x' and y' is complicated (Fig. 49), but the largest distance from the origin gives a maximum 95% confidence level of $r_{\text{mix}} < 1.8 \times 10^{-3}$ for the CP conserved fit. Although this limit is based on a likelihood difference, using $r_{\text{mix}} \doteq \frac{1}{2}(x'^2 + y'^2)$ automatically constrains r_{mix} to be non-negative, so it is most easily compared with the values in Table 20.

The ALEPH collaboration [68] also has limits on r_{mix} using the decay $D^0 \to K^+\pi^-$. It is not apparent from the ALEPH paper whether r_{mix} was constrained to be positive in computing the limits. Assuming CP is conserved they obtain 95% CL upper limits of $r_{\text{mix}} < 0.92\%$ assuming no DCS-mixing interference, $r_{\text{mix}} < 0.96\%$ assuming totally constructive DCS-mixing interference, and $r_{\text{mix}} < 3.6\%$ assuming totally destructive DCS-mixing interference. It is interesting to note that, like CLEO, a wider uncertainty for destructive DCS-mixing interference ($y' < 0$) is obtained.

Table 22 lists published experimental limits on parameters related to r_{mix}. Table 23 lists the y_{CP} measurements used to compute the average in Table 22. DCS modes provide additional information on y through interference between the mixing and DCS matrix elements. Also, since $y \equiv (1/2)\Delta\Gamma/\Gamma$, measurements of the lifetime difference between decay modes with different CP states provides information on y. If CP is conserved, then $\Delta\Gamma = \Gamma(CP \text{ even}) - \Gamma(CP \text{ odd})$ and $y = y_{CP}$.

Figure 49 illustrates these limits in the x-y plane.

VIII.2 Theoretical Predictions

Experimental limits are beginning to constrain theoretical models. Of the 65 predictions compiled in Ref. [26] and illustrated in Fig. 2, 20 are near or above the

Table 22: Most recently published experimental limits on parameters related to r_{mix}. If CP is conserved, $r_{\text{mix}} = \frac{1}{2}(x^2 + y^2) = \frac{1}{2}(x'^2 + y'^2)$.

Modes	Experiment	95% CL
$D^0 \to K^-\pi^+$	CLEO [28]	Most general fit $(1/2)x'^2 < 0.041\%$ $-5.8\% < y' < 1.0\%$
		CP conserving fit $(1/2)x'^2 < 0.038\%$ $-5.2\% < y' < 0.2\%$
See Table 23	Average	$(-1.0 < y_{CP} < 2.5)\%$

Table 23: Most recently published measurements of y_{CP}. If CP is conserved then $y = y_{CP}$. For computing the average, statistical and systematic errors are added in quadrature. The Particle Data Group method for combining measurements with asymmetric errors is used (see Sec. 4.2.1 of [60] for details).

Modes	Experiment	y_{CP}
$D^0 \to K^-\pi^+$ $D^0 \to K^-K^+$ $D^0 \to \pi^-\pi^+$	CLEO [14]	$(-1.2 \pm 2.5 \pm 1.4)\%$
$D^0 \to K^-\pi^+$ $D^0 \to K^-K^+$	Belle [69]	$(-0.5 \pm 1.0^{+0.7}_{-0.8})\%$
$D^0 \to K^-\pi^+$ $D^0 \to K^-K^+$	FOCUS [29]	$(3.42 \pm 1.39 \pm 0.74)\%$
$D^0 \to K^-\pi^+$ $D^0 \to K^-K^+$	E791 [70]	$(0.8 \pm 2.9 \pm 1.0)\%$
	Average	$(0.76^{+0.88}_{-0.90})\%$

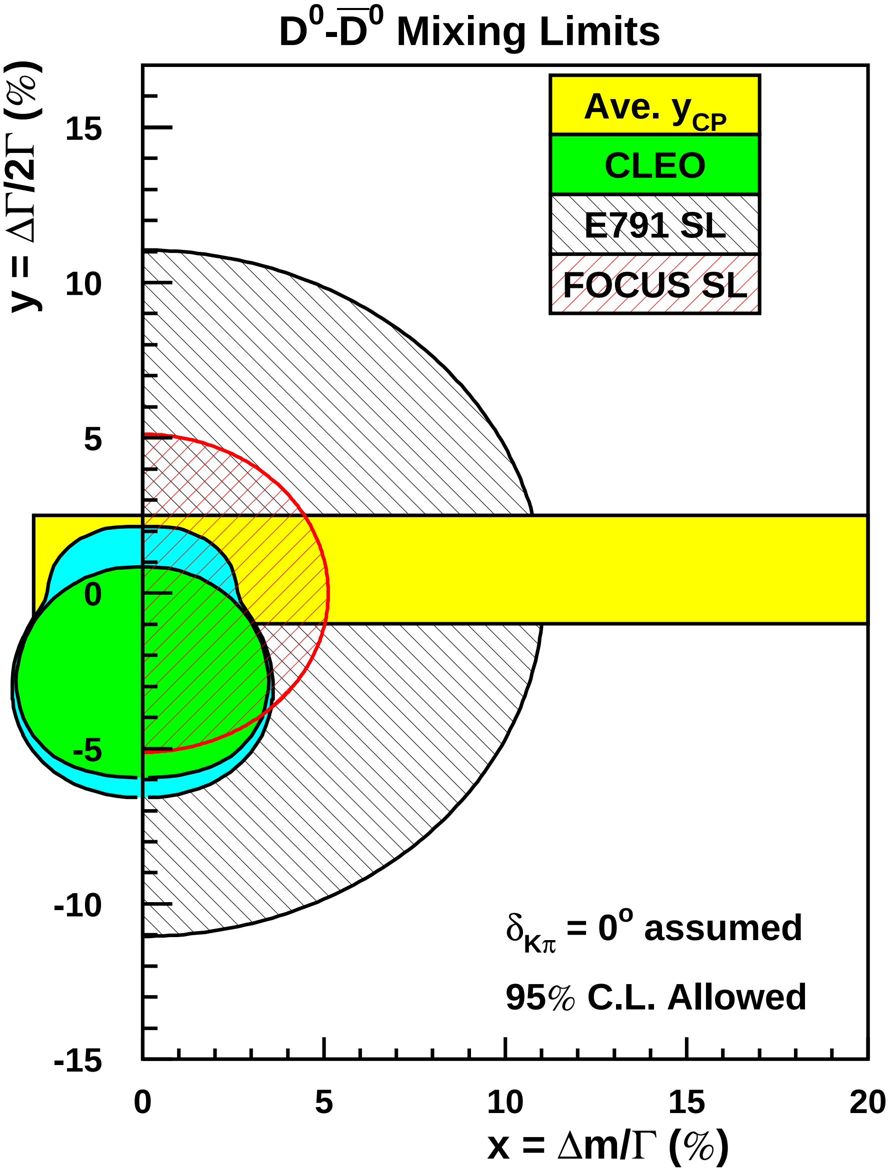


Figure 49: Limit for FOCUS semileptonic mixing superimposed on published limits. The narrow horizontal band is the limit obtained from combining the lifetime asymmetry measurements of E791 (Mar. 1999) [70], FOCUS (June 2000) [29], CLEO (Apr. 2002) [14], and Belle (Apr. 2002) [69]. The pear shaped region is from the CLEO hadronic mixing limit (Dec. 1999) [28] plotted assuming zero for the strong phase angle (the smaller circular region assumes CP is conserved), and the large shaded circle is from the E791 semileptonic mixing limit (August 1996) [30].

Table 24: Theoretical predictions near or above experimental limits. S stands for "Standard Model" and NS stands for "non-Standard Model." The notation "±" does not indicate a $1\,\sigma$ region, but an entire range of predictions, where unknowable parameters govern the variation. This is an abbreviated version of the tables in Ref. [26]. The entries are sorted by year, with the most recent (1998) at the end.

Citation	S/NS	x	$(1/2)x^2$	Comment
[71]	NS	6×10^{-2} [a]	1.8×10^{-3}	Family Symmetry
[72]	NS	5×10^{-2}	1.25×10^{-3}	Higgs Doublet
[72]	NS	$(5.05 \pm 1.85) \times 10^{-2}$	$(0.51 - 2.38) \times 10^{-3}$	Kane-Thun Model [73, 74]
[8]	S	$(0.01 - 10) \times 10^{-2}$ [b]	$5 \times 10^{-9} - 5 \times 10^{-3}$	Long Distance
[75]	NS	$(0.15 - 90) \times 10^{-3}$	$1 \times 10^{-8} - 4.0 \times 10^{-3}$	Superstring-inspired E_6
[76]	NS	4.4×10^{-2}	0.97×10^{-3}	Higgs Doublets, $m_H = 1\,\mathrm{TeV}$
[24]	NS	$(0.1 - 10) \times 10^{-2}$	$5 \times 10^{-7} - 5 \times 10^{-3}$	Fourth Generation
[23]	NS	≈ 0.1	$\approx 5 \times 10^{-3}$	Quark-squark alignment
[77]	NS	≈ 0.11	$\approx 6.0 \times 10^{-3}$	Flavor Changing Scalar Int.
[78]	NS	$(0.006 - 120) \times 10^{-3}$	$1.8 \times 10^{-11} - 7.2 \times 10^{-3}$	Fourth Generation
[78]	NS	$(0.004 - 120) \times 10^{-3}$	$8 \times 10^{-12} - 7.2 \times 10^{-3}$	Higgs Doublet
[78]	NS	$(0.06 - 120) \times 10^{-3}$	$1.8 \times 10^{-9} - 7.2 \times 10^{-3}$	Flavor-Changing Higgs
[79]	NS	$(0.06 - 120) \times 10^{-3}$	$1.8 \times 10^{-9} - 7.2 \times 10^{-3}$	Fourth Generation
[79]	NS	$(0.04 - 120) \times 10^{-3}$	$8 \times 10^{-10} - 7.2 \times 10^{-3}$	Higgs Doublet
[79]	NS	5×10^{-2}	1.25×10^{-3}	Tree Level FCNC
[79]	NS	$\lesssim 0.1$ [c]	$\lesssim 5 \times 10^{-3}$	Squark-gluino box diagrams
[13]	S	6×10^{-2}	1.8×10^{-3}	Upper Limit
[80]	NS	$(0.6 - 6) \times 10^{-1}$	$(1.8 - 180) \times 10^{-3}$	Higgs Doublet
[81]	NS	$(0.06 - 600) \times 10^{-4}$	$1.8 \times 10^{-11} - 1.8 \times 10^{-3}$	Singlet Quarks

[a] It is not obvious how this particular value was extracted from Ref. [71].

[b] Ref. [8] also estimates y with the same range of values as x.

[c] Nelson appears to mistake this for a prediction rather than a limit.

limit from this analysis (Fig. 50). Only two of these predictions are Standard Model estimates. These 20 are listed in Table 24. With the exception of one early estimate, all of these are predictions for x (all the y predictions in Ref. [26] are Standard Model predictions).

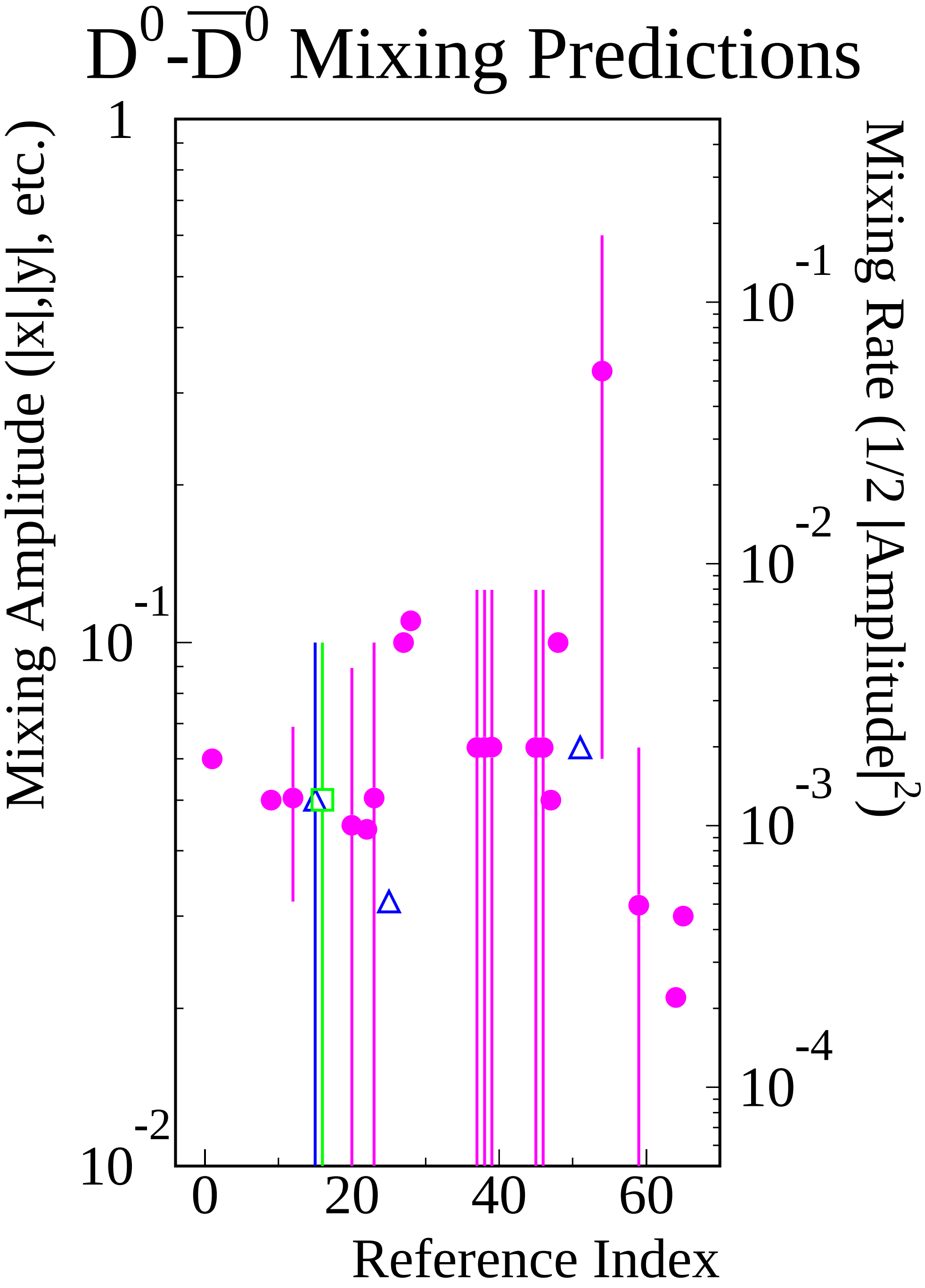


Figure 50: Theoretical model predictions near the experimental limits. The triangles are Standard Model predictions for x, the squares are Standard Model predictions for y, and the circles are non-Standard Model predictions for x. (Figure from H. Nelson at `http://hep.ucsb.edu/people/hnn/wrongd/predictions/`.)

VIII.3 The Future

The BaBar experiment has presented preliminary mixing limits using $D^0 \to K^-\pi^+$ [82] with $57\,\mathrm{fb}^{-1}$ of data. These limits are comparable to the CLEO result [28]. The BaBar experiment should run another five years and accumulate a total of $400\,\mathrm{fb}^{-1}$ of data. The same is true of the Belle experiment, which is already surpassing FOCUS and CLEO in some measurements. Over the next five years, the combined results from BaBar and Belle should lead to roughly a factor of $\sqrt{2 \times 400/57} \approx 4$ improvement[2] in the limits for r_{mix}.

The prospects for D^0–$\overline{D}^0$mixing look even more promising for CLEO-c, which will be dedicated to investigating charm and tau particles. The CLEO experiment is located at the Cornell Electron Storage Ring (CESR). CESR is undergoing modifications to operate at center-of-mass energies between 3 to 5 GeV, including the $\psi(3770)$ and $\psi(4140)$ resonances. The CLEO-c experiment will collect an extremely clean sample of D decays from $\psi(3770) \to D\overline{D}$ and $\psi(4140) \to \gamma D\overline{D}$. The $D^0\overline{D}^0$ pairs will be produced in a coherent state with a definite C quantum number. This will enable the relative strong phase between DCS and Cabibbo favored decays to be measured and x and y possibly separated [83, 84]. CLEO-c is expected to run from 2003 to 2005. With their final data set, CLEO-c anticipates a 95% confidence level sensitivity of $r_{\mathrm{mix}} < 10^{-4}$ [85] which is about a factor of ten better than current experimental limits. This would rule out or constrain some of the larger Standard Model estimates for mixing.

The Fermilab BTeV experiment is expected to start taking data around 2007. In

---

[2] From Eq. 36, the error on r_{mix} scales approximately as $\delta N_{\mathrm{M}}/N_{\mathrm{NM}}$, where δN_{M} is the error on the number of mixing events. δN_{M} scales as the square-root of the number of background events.

a single year, BTeV will collect roughly 1000 times as much charm as the full FOCUS data set. Background levels for charm analyses have yet to be studied in detail for BTeV. Optimistically, a factor of $\sqrt{1000} \approx 30$ improvement over present r_{mix} limits could be obtained with only a single year of running.

VIII.4 Final Remarks

Experimental limits on D^0–$\overline{D}^0$ mixing provide constraints on a large variety of theories beyond the Standard Model as well as providing guidance in computationally challenging areas of the Standard Model. The limits produced from this analysis are among the best in the world. The disadvantages of using semileptonic decays—lower statistics and worse resolution—are offset by the much simpler decay time distribution for mixing in semileptonic decays and the excellent muon identification of FOCUS. D^0–$\overline{D}^0$ mixing will continue to be an active area of research. Experimental sensitivities will likely improve by more than a factor of ten over the next five to ten years, significantly widening the window for discovering new physics.

APPENDIX A

MULTIPLE COULOMB SCATTERING FOR FOCUS MONTE CARLO

This appendix briefly describes multiple Coulomb scattering generated for the FOCUS Monte Carlo by the routine `rgmoliere`.

The theory of small angle multiple Coulomb scattering was worked out by Molière in 1948 [86]. This work was clarified by Bethe [87]. Theoretical calculations for the exit angle have been checked by numerous experiments [88, 89, 90].

A.1 The Scattering Distribution

Molière expresses scattering in terms of two characteristic angles, χ_c and χ_a. Path length dependence is contained in χ_c. Dependence on the cross-section is contained in χ_a. The variable χ_c is defined as

$$\chi_c^2 = 0.157z^2[Z(Z+1)\rho t/A](p^{-2} + (mc)^2 p^{-4}) \tag{53}$$

where t is thickness, ρ is density, and z is the projectile charge. Units used are cm, grams, and MeV. The factor $Z(Z+1)$ was introduced by Bethe [87] and others (replacing Z^2) to approximately account for scattering by atomic electrons. More rigorous, but complicated, corrections exist but the $(Z+1)$ prescription works well for high energy particles [91, 90].

Molière calculated χ_a for a Thomas-Fermi potential (exponentially screened Coulomb charge) and included a correction to the Born approximation. His χ_a is

$$\chi_a^2 = 2.007 \times 10^{-5} Z^{2/3} \frac{1 + 3.34(Zz\alpha)^2}{p^2} \left[1 + \frac{(mc)^2/p^2}{1 + 3.34(Zz\alpha)^2}\right], \tag{54}$$

where the fine structure constant is α.

The quantity

$$\Omega = \chi_c^2/\chi_a^2 \tag{55}$$

is the mean number of collisions. Ω is nearly momentum independent for $\beta \simeq 1$. For $\Omega < 20$, Molière's calculation is not accurate (this region is called "plural scattering"). Molière defined an expansion parameter, B, which satisfies

$$B - \ln B = \ln \Omega + 1 - 2C \tag{56}$$

where $C = 0.5772\ldots$ is Euler's constant.[1] With this, the probability distribution for the polar angle θ is

$$\begin{aligned} f(\theta)\theta\, d\theta &= \vartheta\, d\vartheta[2\exp(-\vartheta^2) + B^{-1}f^{(1)}(\vartheta) + B^{-2}f^{(2)}(\vartheta) + \cdots], && (57) \\ \vartheta &= \theta/(\chi_c B^{\frac{1}{2}}), && (58) \end{aligned}$$

and

$$f^{(n)}(\vartheta) = n!^{-1} \int_0^\infty u du J_0(\vartheta u) \exp\left(-\frac{1}{4}u^2\right) \left[\frac{1}{4}u^2 \ln\left(\frac{1}{4}u^2\right)\right]^n . \tag{59}$$

In practice, one only needs to go up to $n = 2$. $f^{(1)}$ and $f^{(2)}$ are easy to compute with series expansions given in Bethe's paper [87]. The author has reproduced Bethe's table for $f^{(1)}$ and $f^{(2)}$ using his expansions as well as doing the integrals numerically with Runge-Kutta.[2] $f^{(1)}(\vartheta)$ and $f^{(2)}(\vartheta)$ are evaluated by `rgmoliere` using cubic spline interpolation for $\vartheta \leq 10$. For $\vartheta > 10$, the asymptotic formulas Eqs. (57) and (58) from Ref. [92], are used. The idea of using cubic splines was also taken from Ref. [92].

---

[1] B is nearly linear in Ω for materials and thicknesses of interest. In `rgmoliere`, B is first approximated from a linear equation and then improved with one or two iterations of Newton's method.

[2] Equation (29) of Ref. [87] contains a typo—a missing factor of $\frac{1}{4}$. However, the values in Table II of Ref. [87] are still correct.

A.2 Random Number Generation

To generate scattering, a rejection function whose integral can be inverted is used. See Ref. [93], Sec. 7.3 for a description of this method. Knowing that the Molière distribution approaches Rutherford scattering ($f_R(\theta)$) at large ϑ was very useful in finding a rejection function. From Ref. [87],

$$f_R(\theta)\theta d\theta = (2/B)d\vartheta/\vartheta^3.$$

Near $\theta = 0$, Hanson *et al.* [94] suggest the distribution can be approximated by a Gaussian with somewhat narrower width than Molière's leading term. The suggested $1/e$ width is

$$\theta_w = \chi_c(B - 1.2)^{\frac{1}{2}}$$

rather than $\theta_w = \chi_c B^{\frac{1}{2}}$.

After some trial and error, we arrive at the following rejection function:

$$f_{\text{rej}}(\vartheta) = 2\exp(-\vartheta^2/(1 - 1.2/B)) + (c_1 + c_2 B\vartheta^4)^{-1}. \tag{60}$$

The variable c_1 is chosen so f_{rej} agrees exactly with the Molière distribution at $\vartheta = 0$:

$$c_1 = B/(f^{(1)}(0) + f^{(2)}(0)/B).$$

The variable c_2 would be 0.5 to match Rutherford scattering at large ϑ, but with 0.5 the rejection function drops below Molière's distribution at intermediate ϑ. To prevent this, 0.5 is reduced to 0.17, which works for $B \leq 79$ (the largest B in the materials file for FOCUS is about 18). A material dependent c_2 would result in a somewhat more efficient computation.

The two terms in the rejection function can be treated as separate probability distributions. This has the advantage that one needs to invert two simple integrals rather than a single complicated one. The integrals used are

$$\int 2\exp(-\vartheta^2/w)\vartheta d\vartheta = w(1-\exp(-\vartheta^2/w)) \tag{61}$$

$$\int (a+b\vartheta^4)^{-1}\vartheta d\vartheta = \frac{1}{2\sqrt{ab}}\arctan\vartheta^2\sqrt{b/a}. \tag{62}$$

Taking the limits of integration to ∞ gives the area of the two distributions, A_1 and A_2. The fraction of entries taken from the first distribution is $A_1/(A_1+A_2)$.

A.3 χ_a and χ_c for Compounds

Molière generalized χ_a and χ_c to account for energy loss and different layers of material [95]. Energy loss is not accounted for by `rgmoliere`. Energy loss is treated in Appendix B.

The following definitions will be used below:

$$k_c \equiv 0.157[Z(Z+1)\rho/A] \tag{63}$$

$$k_a \equiv 2.007\times10^{-5}Z^{2/3}(1+3.34(Z\alpha)^2) \tag{64}$$

$$a \equiv (1+3.34(Z\alpha)^2)^{-1}. \tag{65}$$

These definitions differ slightly from those in Appendix B. The differences involve dividing out the mass of the projectile, and setting the charge of the projectile, z, to 1. This is done in order to have constants that depend only on the material. The values for these constants are read from an initialization file by `rgmoliere`.

With the above constants and assuming $z=1$, χ_c and χ_a are given by

$$\chi_c^2 = k_c t(p^{-2}+(mc)^2p^{-4}) \tag{66}$$

$$\chi_a^2 = \frac{k_a}{p^2}[1+a(mc)^2/p^2], \tag{67}$$

Effective χ_a and χ_c after traversing layers of different materials are given by summing over the layers:

$$\begin{aligned} \chi_c^2 &= \sum_i (\chi_c^2)_i && (68) \\ \chi_c^2 \ln \chi_a^2 &= \sum_i (\chi_c^2 \ln \chi_a^2)_i . && (69) \end{aligned}$$

After generalizing these equations for compounds, we sum over elements instead:

$$\begin{aligned} k_c &= 0.157\rho \sum_i [wZ(Z+1)/A]_i && (70) \\ k_c \ln k_a &= 0.157\rho \sum_i [w(Z/A)(Z+1)\ln(2.007 \times 10^{-5} Z^{\frac{2}{3}}(1 + 3.34(\alpha Z)^2))]_i && (71) \end{aligned}$$

where w_i is the fraction by weight of the i'th element and ρ is the density of the compound. An extensive list of densities can be found in Ref. [96]. The effective a is found using the effective k_a and solving Eq. (64) for Z numerically and then using Eq. (65). Eq. (71) also assumes $(mc)^2 \ll p^2$.

APPENDIX B

OUTER MUON IDENTIFICATION ALGORITHM

This appendix describes the method used for calculating the confidence level that a track projected to the RPC's is consistent with being a muon. The basic elements of the method include: a covariance matrix calculation which accounts for material effects through different layers, a pattern recognition algorithm which identifies separate clusters of hits, and inclusion of magnetic deflections in the M2 iron. The following sections describe how the the confidence level calculation is done, some Monte Carlo and data results, and ways to improve the calculation.

B.1 Confidence Level Calculation

The confidence level is the probability that a measurement will have a chi-square equal to or greater than some specified value (assuming the measurement errors are approximately Gaussian). If the predicted errors used in the chi-square calculation are correct, then the confidence level distribution should be flat. If the errors are underestimated then the confidence level will peak at less than 0.5; If the errors are overestimated, then the confidence level will peak at greater than 0.5. (Calculation of the confidence level as a function of chi-square and number of degrees of freedom, for Gaussian errors, is illustrated by the function `gammq` in Ref. [93].)

With correlated errors, the chi-square takes the form [97]

$$\chi^2 = \sum_i \sum_j C_{ij}^{-1}(t_i - X_i)(t_j - X_j), \tag{72}$$

where the sums are over the muon planes and C_{ij}^{-1} is the inverse of the covariance matrix. X_i is the measured coordinate and t_i is the extrapolated track coordinate. The following sections discuss the calculation of the covariance matrix.

B.1.1 The Covariance Matrix

The covariance between two variables, x_i and x_j, is

$$C_{ij} = \int\int dx_i\, dy_j\, (x_i - \bar{x}_i)(x_j - \bar{x}_j)P(x_i, x_j) \tag{73}$$

where $\bar{x}_i$ is the mean value of x_i and $P(x_i, x_j)$ is the joint probability distribution for (x_i, x_j). For short, this can be written as $C_{ij} = \langle \Delta x_i \Delta x_j \rangle$. For $i = j$, the covariance is the same as the variance (square of the standard deviation).

The probability distribution for detected muons is determined by multiple scattering and energy loss through the M2 magnet and OE calorimeter. For hadrons, the mean free path for inelastic interactions in iron is about 17 cm, so hadrons typically dissipate their energy in showers which generally do not make it through the OE and M2.[1] Electrons bremsstrahlung their energy away. This leaves mostly muons of above 4 GeV making it to the RPC's.

The displacement, in the x-z plane, of a muon from its extrapolated track intersection point at the exit face of a scattering slab, can be approximated as:

$$x_{\text{dis}} = x_{\text{rms}}((1-\rho^2)^{\frac{1}{2}}\zeta_{1x} + \rho\zeta_{2x}) \tag{74}$$

This is for a particle incident along the z-direction, normal to the slab face. x_{rms} is the x-component of the RMS multiple scattering displacement[2] and ρ is the correlation

---

[1] Since there is a gap in the OE for electron-positron pairs, there are two narrow bands where a hadronic shower could leak through the magnet iron and mirror plates.

[2] The Root-Mean Square scattering here refers to a Gaussian approximation. The RMS of the full distribution is much larger than the Gaussian "core."

coefficient between the exit angle and x_{dis}. The variables ζ_{1x} and ζ_{2x} are Gaussian deviates with a RMS of 1. The exit angle is given by $\zeta_{2x}\phi_0$, indicating that the spatial displacement is correlated to the exit angle. The RMS planar projected exit angle is ϕ_0. Similarly, the displacement in the y-z plane is:

$$y_{\text{dis}} = y_{\text{rms}}((1-\rho^2)^{\frac{1}{2}}\zeta_{1y} + \rho\zeta_{2y}), \tag{75}$$

where (ζ_{1y}, ζ_{2y}) are independent of (ζ_{1x}, ζ_{2x}), and $y_{\text{rms}} = x_{\text{rms}}$.

Eqs. (74) and (75) are somewhat more general than those given in Sec. 22.6 of the Review of Particle Properties [98]. The generalization is needed to account for energy loss and multiple materials. The equations in [98] imply

$$y_{\text{rms}}^2 = \frac{1}{3}(t\phi_0)^2 \tag{76}$$

with pathlength t, which is only valid with a single material and negligible energy loss. Even if multiple materials and energy loss are accounted for in ϕ_0, this relationship gives an incorrect y_{rms}. Calculations of ϕ_0, y_{rms}, and ρ which properly account for multiple materials and energy loss are discussed in Sec. B.1.3.

Eqs. (74) and (75) give the deflection at the exit face of a scatterer, in a coordinate system where the incident track points along the z-axis. However, what is needed is the deflection at a RPC plane. Also, the incident track has a non-zero slope. To account for the slope of a track, t can be taken as the pathlength of the extrapolated track through the scattering slabs, rather than the slab thickness. Then x_{dis} and y_{dis} are displacements perpendicular to the incident track. This is illustrated in Fig. 51.

The projection of y_{dis} in the vertical (y-z) plane is $y_{\text{dis}}\sqrt{1+m_y^2}$. The intersection point of the scattered track with a detector plane is then given by

$$y_i = m_y z_e + y_0 + m_y'(z_i - z_e) + y_{\text{dis}}\sqrt{1+m_y^2}. \tag{77}$$

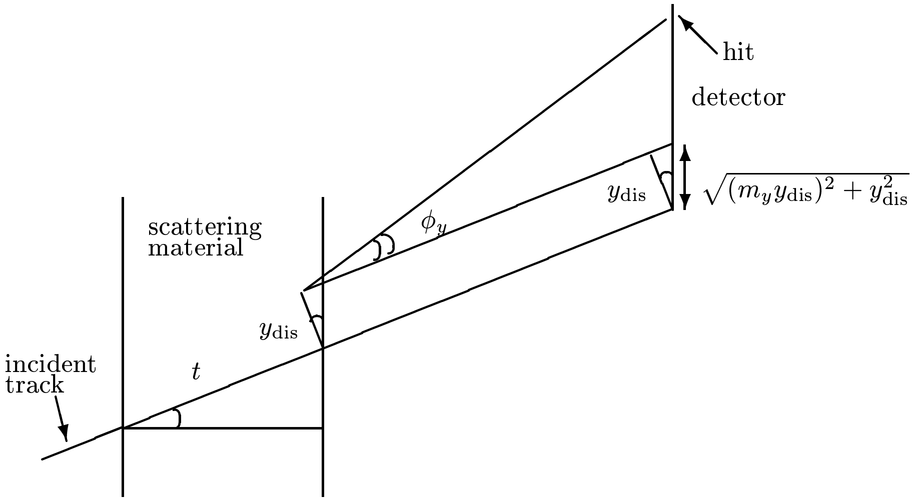


Figure 51: Multiple scattering parameters viewed in the y-z plane. The pathlength through the material is t, ϕ_y is the planar projected multiple scattering angle, $y_{\rm dis}$ is the scattering displacement at the exit from the material, and m_y is the incident track y-slope.

The y-slope of the incident track is m_y [3] and the y-intercept of the line corresponding to the incident track is y_0. The y-slope after scattering is m'_y. The z-position of the detector is z_i, and the z-position of the exit point from the scatterer is z_e. There is a small correction to z_e which has not been included in the above equation.

Using small angle approximations,

$$m'_y \simeq m_y + \phi_y(1 + m_y^2), \tag{78}$$

where the angular deflection is $\phi_y = \zeta_{2y}\phi_0$. To obtain this approximate m'_y, use has been made of the expressions:

$$\sin(A + \phi_y) \simeq \sin A + \phi_y \cos A \tag{79}$$

$$\cos(A + \phi_y) \simeq \cos A - \phi_y \sin A, \tag{80}$$

---

[3] To account for magnetic deflections, m_y, is reinterpreted as the slope that a muon would acquire after magnetic deflection through M2, in the absence of multiple scattering. The variable t is still the pathlength traversed in absence of multiple scattering, but the path is now curved.

with $m_y = \tan A$ and $m'_y = \tan(A + \phi_y)$.

From Eqs. (77) and (78), the deflection in the i'th RPC plane is

$$\Delta y_i = \zeta_{2y}\phi_0(1 + m_y^2)(z_i - z_e) + y_{\text{rms}}((1 - \rho^2)^{\frac{1}{2}}\zeta_{1y} + \rho\zeta_{2y})\sqrt{1 + m_y^2}. \tag{81}$$

An exactly analogous expression holds for Δx_i. The u-view coordinate is given in terms of the x and y coordinate through $u_i = x_i \cos A + y_i \sin A$. For the RPC's, $\sin A = \cos A = 1/\sqrt{2}$. For the u_i coordinate then,

$$\Delta u_i = (\Delta x_i + \Delta y_i)/\sqrt{2} \tag{82}$$

The expressions for Δx, Δy, and Δu can be inserted into Eq. (73) to find the covariance. The integrations are simple:

$$\begin{aligned} \int d\zeta\, P(\zeta) &= 1 & (83)\\ \int d\zeta\, \zeta\, P(\zeta) &= 0 & (84)\\ \int d\zeta\, \zeta^2\, P(\zeta) &= 1. & (85) \end{aligned}$$

Note however, the range of integration is taken from $-\infty$ to $+\infty$. Technically, the range of integration should be limited by the detector acceptance. So, tracks near the edges of the detector acceptance are not treated correctly. By combining each detector in a given view (x, y, or u), into one effective detector, this problem is minimized. Still, there are tracks near the beam aperture that are not treated correctly. Requiring a minimum amount of material traversed cuts out tracks near the aperture.

The elements of the covariance matrix obtained from the above integrations are as follows (with the abbreviation $\Delta z_i \equiv (z_i - z_e)$):

$$\langle \Delta x_i \Delta x_j \rangle = \phi_0(1 + m_x^2)(\phi_0(1 + m_x^2)\Delta z_i \Delta z_j + y_{\text{rms}}\rho(\Delta z_i + \Delta z_j)\sqrt{1 + m_x^2})$$

$$+ y_{\rm rms}^2(1+m_x^2) \tag{86}$$

$$\langle \Delta y_i \Delta y_j \rangle = \phi_0(1+m_y^2)(\phi_0(1+m_y^2)\Delta z_i \Delta z_j + y_{\rm rms}\rho(\Delta z_i + \Delta z_j)\sqrt{1+m_y^2})$$

$$+ y_{\rm rms}^2(1+m_y^2) \tag{87}$$

$$\langle \Delta x_i \Delta y_j \rangle = 0 \tag{88}$$

$$\langle \Delta x_i \Delta u_j \rangle = \langle \Delta x_i \Delta x_j \rangle / \sqrt{2} \tag{89}$$

$$\langle \Delta y_i \Delta u_j \rangle = \langle \Delta y_i \Delta y_j \rangle / \sqrt{2} \tag{90}$$

$$\langle \Delta u_i \Delta u_j \rangle = (\langle \Delta x_i \Delta x_j \rangle + \langle \Delta y_i \Delta y_j \rangle)/2 \tag{91}$$

For equations involving the same view (Eqs. (86), (87), and (91)), it is normally the case that $i = j$. However, it is possible for a single track to hit two planes in the same view since there is some overlap between RPC towers. Hits in separate towers form separate clusters, unless there are less than three views. In this case, hits from an overlapping tower may be added to form a three view (three degree of freedom) cluster.

In order to approximately account for granularity of the detectors, a contribution is added to the diagonal elements of the covariance matrix (the equations above only include the multiple scattering contribution). A uniform distribution is assumed over the width of a readout strip (12.4 cm in the RPC's), and the RMS is calculated as follows:

$$\begin{aligned} \sigma^2 &= \frac{1}{12.4\,\mathrm{cm}} \int_{-6.2\,\mathrm{cm}}^{6.2\,\mathrm{cm}} dx\, x^2 \\ &= 12.813\,\mathrm{cm}^2. \end{aligned} \tag{92}$$

B.1.2 Pattern Recognition

When multiple hits occur in the RPC's in a single event, a decision must be made as to which hits belong to a given track. Taking the hits nearest a track as belonging to the track sometimes results in non-overlapping strips being chosen. Nehring[4] wrote a routine which remedies this problem. The routine identifies clusters of active strips and returns their position. If adjacent strips in a given view (there are three views: x, y, and u) are active, then the cluster coordinate is given as the midpoint between the two strips and the granularity error is halved.[5] The cluster returning the highest confidence level for a given track is treated as a match.

B.1.3 Multiple Coulomb Scattering

The theory of small angle multiple Coulomb scattering was worked out by Molière in 1948 [86]. Molière's basic results were outlined in Appendix A. To a good approximation, the Molière distribution can be replaced by a Gaussian. The distribution, however, has larger tails than a Gaussian. Because of the tails, determining the final width by adding the RMS of successive approximate Gaussian scattering in quadrature is only accurate for a limited range of material thicknesses.

Lynch and Dahl [91] have obtained an expression for the multiple scattering width (ϕ_0 in this thesis), from Gaussian fits to the Molière distribution over a wide range of materials and energies. Previous formulas for ϕ_0 (such as the Rossi or Highland formulas) are also found from fits to the Molière distribution, but are functions of radiation lengths, whereas the Lynch and Dahl formula depends on two characteristic

---

[4] Matthew Nehring, Assistant Professor of Physics at Adams State College, Alamosa, Colorado.

[5] This does not work when adjacent strips are activated by different muons. But this is a rare occurrence.

angles, χ_a and χ_c, from Molière's theory. These angles are functions of momentum, path length, density, atomic weight, atomic charge, and projectile charge. The Lynch and Dahl formula is used for the outer muon simulation.[6]

In order to take into account different materials and changes in momentum, effective χ_a and χ_c need to be calculated through suitable averaging. Formulas for effective χ_a and χ_c are given by Lynch and Dahl in Ref. [91], however they neglect energy loss.

Multiple scattering with energy loss for the Gaussian approximation was derived by Eyges [99], generalizing the work of Fermi [100]. This implies (see corrections and useful comments in [101]):

$$\phi_0^2 = \frac{1}{2}\int_0^t \theta_s^2(x)dx \tag{93}$$

$$y_{\text{rms}}^2 = \frac{1}{2}\int_0^t \theta_s^2(x)(t-x)^2dx \tag{94}$$

$$\langle y\phi\rangle = \frac{1}{2}\int_0^t \theta_s^2(x)(t-x)dx. \tag{95}$$

Here the momentum dependence of $\theta_s^2(x)$ is regarded as a function of position x. The factor $(t-x)^2$ indicates that scattering at the beginning of the path contributes more to y_{rms}^2 than scattering at the end of the path. Thus, the influence of low energy scattering near the end of the path, is less than suggested by the simple formula $y_{\text{rms}}^2 = \frac{1}{3}(t\phi_0)^2$, which neglects energy loss.

Since the Gaussian approximation of Lynch and Dahl [91] is written in terms of the characteristic angles, χ_a and χ_c from Molière's theory, energy loss and multiple layers need to be accounted for using Molière's notation. Most recent references only do this for ϕ_0, but not y_{rms} or ρ. Fortunately, Molière (1955) showed how to express

---

[6] Eq. (7) of the Lynch and Dahl paper [91] contains a typo: σ should be replaced by σ^2.

distributions for y/t and $(y/t+\phi)$ in the same form as equivalent ϕ distributions [95]. This leads to integrals similar to Eqs. (93)–(95). In order to make the integrations simpler, the rate of momentum loss, $\frac{dp}{dx}$, is assumed constant (within 1%) over a given layer. This condition on the rate is checked in the code[7] and the layer is subdivided if needed (there is a minimum cut-off of 8 cm). This assumption enables larger step sizes to be taken than would be possible if the steps were so small that energy loss was negligible over a given layer. Calculations of equivalent χ_a and χ_c are outlined in Sec. B.4.

B.1.4 Ionization Energy Loss

Based on $c\bar{c}$ Monte Carlo, 40% of muons incident on the RPC's are absorbed. It is therefore necessary to treat muons down to a fairly low energy. Because materials are subdivided into layers for which momentum loss rate is constant, there is no need to use a numerical integration of the Bethe-Bloch equation. Eliminating this gives another improvement in speed. Convergence and absorption are checked at several stages of the calculation.

The density correction is included in $\frac{dE}{dx}$, but the shell correction is ignored. Calculation of $\frac{dE}{dx}$ for composite materials is discussed in Refs. [102, 103].

B.1.5 Magnetic Deflections

To approximate the mean muon trajectory, scattering is ignored and the trajectory stepped through the magnet, assuming a constant magnetic field over each step. Since

---

[7] The thickness resulting in a 1% change in $\frac{dp}{dx}$ was fit to a polynomial, giving thickness as a function momentum for iron.

the equations of motion have a p^{-1} dependence, to account for energy loss an effective momentum is used:

$$
\begin{aligned}
1/p_{\text{eff}} &= (p_{\text{exit}} - p_{\text{entrance}})^{-1} \int_{p_{\text{entrance}}}^{p_{\text{exit}}} dp/p & (96) \\
&= (p_{\text{exit}} - p_{\text{entrance}})^{-1} \log(p_{\text{exit}}/p_{\text{entrance}}) & (97)
\end{aligned}
$$

Convergence to within half a millimeter lateral deflection at the RPC's is obtained by limiting the step size to no more than 20 cm. (This assumes muons with incident energies of at least 4 GeV. Most muons with lower energies are absorbed anyway.)

Webster[8] has written a routine which approximates the B field inside the M2 iron. The routine consists of polynomial fits to a two-dimensional field map produced using the Fermilab program, `POISSON`. In calculating the trajectory, the B field in the iron is assumed to end at the edges of the magnet. To first order, only the product of magnet thickness and B field contributes to the deflection, rather than each individually.

B.2 Results

The subroutine which computes the outer muon confidence level for each track is called `omurecon`. Figure 52 shows the confidence level distribution as calculated in `omurecon` from muons in $c\bar{c}$ Monte Carlo. At least one cluster is required to be found in the RPC's in an event and the incident track is within the RPC acceptance. The bottom plot shows the momentum spectrum for these muons. This is merely a consistency check between the confidence level computed in `omurecon` and the simulation of muon hits in `omusim`. As expected, the distribution is fairly uniform.

Figure 53 shows three plots from candidate muons in J/ψ data. One muon in the

---

[8] Medford Webster, Professor of Physics, Vanderbilt University.

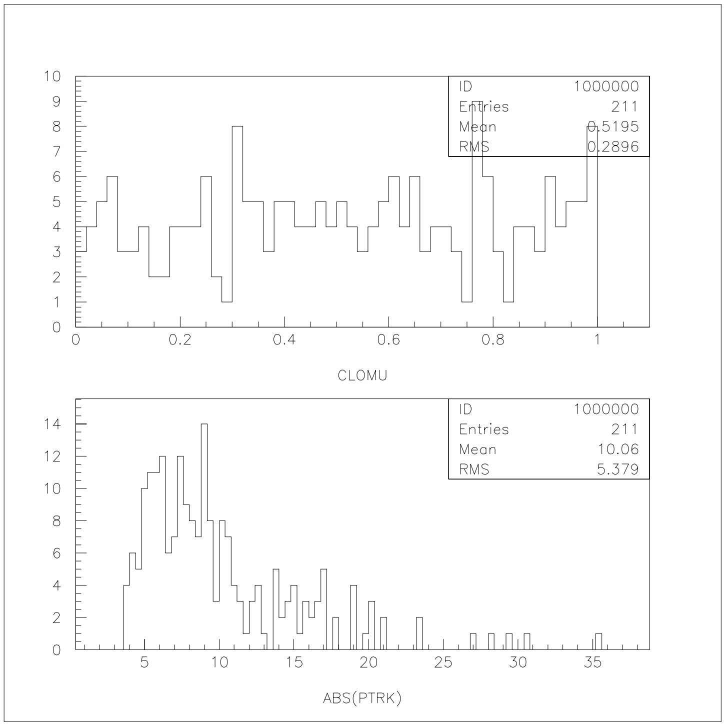


Figure 52: Confidence level and momentum distributions for muons from $c\bar{c}$ Monte Carlo.

pair appears in the plot, the other muon is identified in the inner muon detector. The top plot is the confidence level (entries below 0.001 are not shown), the middle plot is the distribution of projected track minus y-hit position, divided by the predicted RMS (again, entries with confidence level below 0.001 are not shown). The predicted RMS is about 3% larger than the measured RMS. The scattering RMS is about 19% larger than the granularity RMS. The projected track includes magnetic corrections for M2. The bottom plot is the momentum spectrum for entries in the top plot.

B.3 Improving the Algorithm

Based on Monte Carlo, the confidence level distribution for muons is flat, except at high energy (20 GeV and up), where granularity effects become significant. This

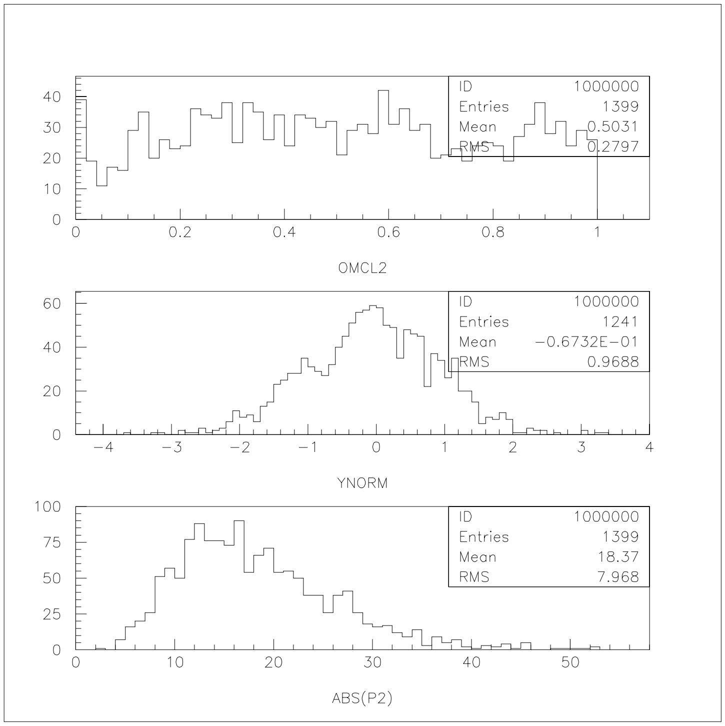


Figure 53: Top: Confidence level distribution for candidate J/ψ muons. Only confidence levels of at least 0.001 are plotted (there are 146 omitted entries from this cut). Middle: Projected track minus y-hit position, divided by predicted RMS. The scattering RMS is about 19% larger than the granularity RMS. Bottom: Momentum spectrum for entries in the top plot.

is because the confidence level is calculated assuming Gaussian errors, but at high momentum the errors are basically rectangular. This tends to artificially push low confidence level tracks to higher confidence level. Thus, efficiency for muons is increased, but more background is accepted. Methods for handling granularity effects in confidence level calculations were developed by Cawlfield, Ruesink, and Wiss [104] which work well for four or more degrees of freedom (hits in four or more layers). Improving the calculation for two or three degrees of freedom might improve background rejection for higher energy muons.

B.4 Calculation of Molière Parameters with Energy Loss

This section outlines the calculation of effective scattering parameters, including energy loss, in Molière's notation [95]. Molière generalized his calculation of scattering distributions to include linear combinations of the form $\phi\cos\lambda + \psi\sin\lambda$, where $\psi = y/t$. Recall y is the y-component of lateral deflection through a thickness, t. The parameters replacing χ_c and χ_a are x_c and x_a. These are given by (Eqs. (3.9a) and (3.9b) of Ref. [95]):

$$x_c^2 = t\int_0^1 da_2\,\chi_c'^2(a_2)(c+sa_2)^2 \tag{98}$$

$$\ln x_a^2 = (t/x_c^2)\int_0^1 da_2\,\chi_c'^2(a_2)(c+sa_2)^2\ln[\chi_a^2(a_2)(c+sa_2)^2] \tag{99}$$

with the abbreviations $\cos\lambda = c$ and $\sin\lambda = s$. $\chi_c'^2$ and χ_a^2 are defined as follows:

$$\chi_c'^2 = 0.157z^2[Z(Z+1)\rho/A](p^{-2}+(mc)^2p^{-4}) \tag{100}$$

$$\chi_a^2 = 2.007\times10^{-5}Z^{2/3}\frac{1+3.34(Zz\alpha)^2}{p^2}\left[1+\frac{(mc)^2/p^2}{1+3.34(Zz\alpha)^2}\right]. \tag{101}$$

Units used in the above are cm, grams, and MeV. The fine structure constant is α and ρ is mass density. The factor $Z(Z+1)$ was introduced by Bethe [87] and others (replacing Z^2) to approximately account for screening of the nucleus by surrounding electrons. The projectile charge is z.

Results of integrating Eqs. (98) and (99) for a single layer, i, are summarized below. The position along the z-axis after exiting the layer is x_i, and the momentum after exiting is $mc\eta_i$. When energy loss is present, $\frac{dp}{dx}$ is treated as constant. Terms higher order than η^{-2} are generally dropped. The following notation is used:

$$k_c \equiv 0.157[Z(Z+1)\rho/A]z^2(mc)^{-2} \tag{102}$$

$$k_a \equiv 2.007\times10^{-5}Z^{2/3}(1+3.34(Zz\alpha)^2)(mc)^{-2} \tag{103}$$

$$
\begin{aligned}
a &\equiv (1+3.34(Zz\alpha)^2)^{-1} && (104)\\
\xi &\equiv \frac{d\eta}{dx} && (105)\\
\Delta x_i &\equiv x_i - x_{i-1} && (106)\\
x_c^2 &= \sum_i (x_c^2)_i && (107)\\
\ln x_a^2 &= \sum_i (\ln x_a^2)_i && (108)
\end{aligned}
$$

For ϕ without energy loss:

$$
\begin{aligned}
(x_c^2)_i &= k_c \frac{\Delta x_i}{\eta_i^2} && (109)\\
x_c^2(\ln x_a^2)_i &= (x_c^2)_i \ln k_a + k_c \Delta x_i (a\eta_i^{-2} - 2\ln \eta_i)\eta_i^{-2} && (110)
\end{aligned}
$$

For ϕ with energy loss:

$$
\begin{aligned}
(x_c^2)_i &= k_c \frac{\Delta x_i}{\eta_i \eta_{i-1}} && (111)\\
x_c^2(\ln x_a^2)_i &= (x_c^2)_i \ln k_a + k_c \xi^{-1}\left[2\eta^{-1}(\ln\eta + 1) - \frac{1}{3}a\eta^{-3}\right]_{\eta_{i-1}}^{\eta_i} && (112)
\end{aligned}
$$

For ψ without energy loss:

$$
\begin{aligned}
(x_c^2)_i &= -k_c t[(1-x_i/t)^3 - (1-x_{i-1}/t)^3]/(3\eta_i^2) && (113)\\
x_c^2(\ln x_a^2)_i &= (x_c^2)_i(\ln k_a - 2\ln\eta_i + a\eta_i^{-2})\\
&\quad -\frac{2}{3}tk_c\eta_i^{-2}\left[\left(1-\frac{x}{t}\right)^3\left(\ln\left(1-\frac{x}{t}\right) - \frac{1}{3}\right)\right]_{x_{i-1}}^{x_i} && (114)
\end{aligned}
$$

For ψ with energy loss:

$$
\begin{aligned}
(x_c^2)_i &= k_c \xi^{-1}[-A^2\eta^{-1} + 2AB\ln\eta + B^2\eta]_{\eta_{i-1}}^{\eta_i} && (115)\\
A &\equiv 1 + (t\xi)^{-1}\eta_{i-1} - x_{i-1}/t && (116)\\
B &\equiv -(t\xi)^{-1} && (117)
\end{aligned}
$$

Unfortunately, the author is unable to find an analytic expression for x_a^2 with energy loss for ψ. The 8 point Newton-Cotes formula (see Abramowitz and Stegun [105] Eq. (25.4.17)) works well enough for this purpose (at least for the materials, geometry, and energies in the outer muon system).

The correlation coefficient, ρ, can be found from $\langle y\phi\rangle$, which in turn can be found from $\langle(\psi+\phi)^2\rangle$. The integrals used for $(\psi+\phi)/\sqrt{2}$ are similar to ψ and can be found by simple substitutions. The substitutions can be found from Eqs. (98) and (99) and a little algebra. For ψ, $c=0$ and $s=1$; For $(\psi+\phi)/\sqrt{2}$, $c=1/\sqrt{2}$ and $s=1/\sqrt{2}$. Aside from some constant factors, the integrals are the same for both ψ and $(\psi+\phi)/\sqrt{2}$.

APPENDIX C

CONFIDENCE INTERVALS FOR NON-GAUSSIAN RESOLUTIONS

The Feldman-Cousins paper suggests how to construct confidence intervals for non-parabolic log-likelihood functions with measurements near a physical boundary [54]. Their paper considers a somewhat more complicated two-parameter example than needed here. Only the one parameter case will be considered here.

With simulated data, "acceptance intervals" can be constructed using an ordering principle based on the log-likelihood difference

$$\Delta\chi^2 = -2\ln\mathcal{L}_{\text{gen}} + 2\ln\mathcal{L}_{\text{best}} \tag{118}$$

where $\mathcal{L}_{\text{gen}}$ is the likelihood evaluated at the generated r_{mix}, r_{gen}, and $\mathcal{L}_{\text{best}}$ is the the likelihood evaluated at the most likely non-negative r_{mix} (either zero or the maximum likelihood from a fit).[1] Based on many simulated experiments, a distribution for $\Delta\chi^2$ is found for the r_{gen}. From this distribution, a $\Delta\chi^2_c$ is found such that α of the simulated experiments have $\Delta\chi^2 < \Delta\chi^2_c$ (each r_{gen} has a single $\Delta\chi^2_c$). Then $\Delta\chi^2(r_{\text{mix}})$ for the real data is compared to $\Delta\chi^2_c(r_{\text{mix}})$ and the confidence interval for r_{mix} is all points such that

$$\Delta\chi^2(r_{\text{mix}}) > \Delta\chi^2_c(r_{\text{mix}}). \tag{119}$$

---

[1] Feldman and Cousins actually recommend a particular form for $\mathcal{L}$. Presumably other forms for $\mathcal{L}$ would work also.

APPENDIX D

INNER AND OUTER MUON DISTRIBUTIONS

In this appendix, shapes for the inner and outer muon samples are shown separately. The final results used the merged sample but inner and outer muon samples were separated to select cuts (Sec. VI.7) and cross-validate the Monte Carlo (Sec. VII.1). Fits made using the separate samples used the high bias bin merging algorithm (Sec. VI.4.1).

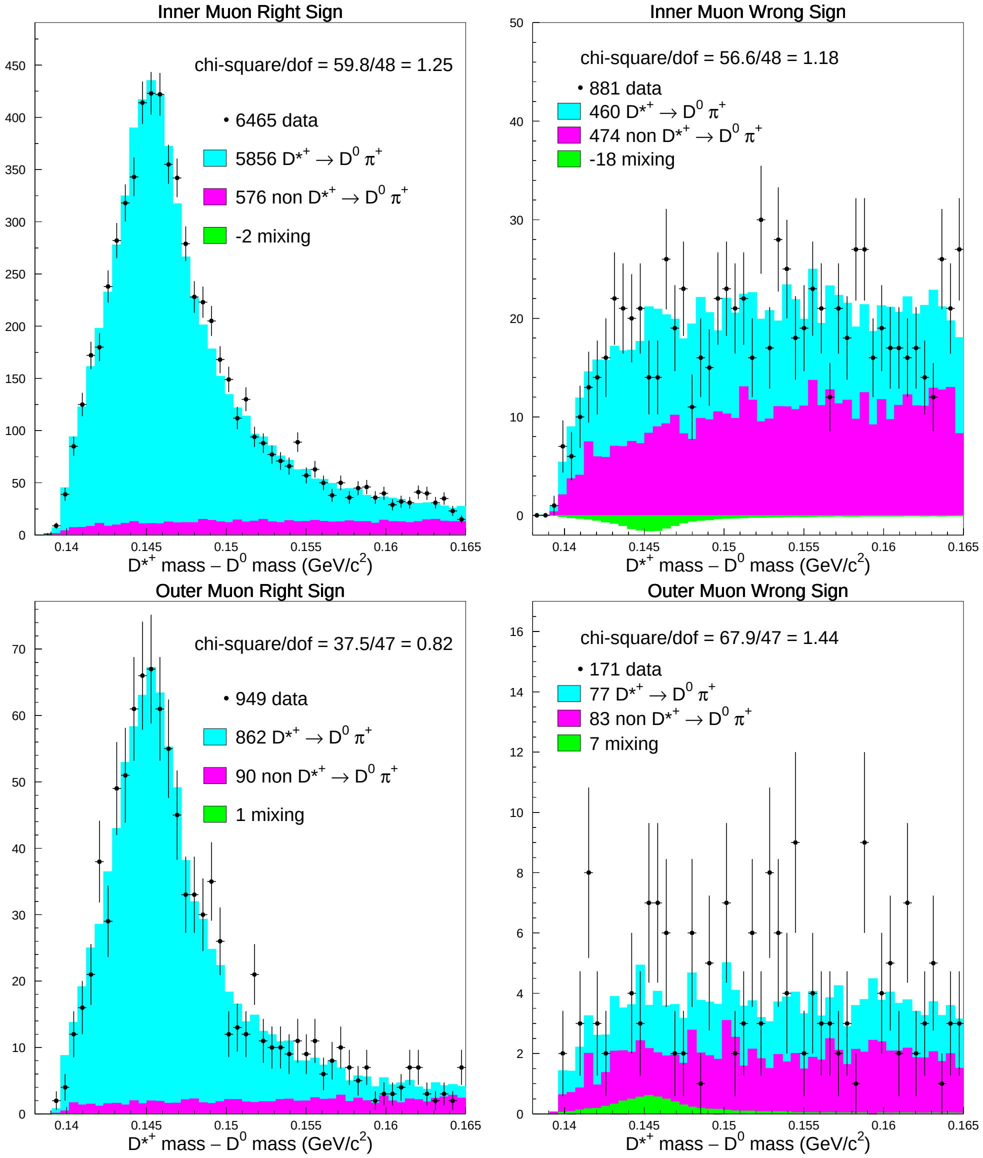


Figure 54: Results of a two parameter, two-dimensional fit to data with Monte Carlo (MC) shapes used in the fit overlaid. Only the D^{*+} mass dimension is shown. The first MC shape is from events containing a $D^{*+} \rightarrow D^0\pi^+$. The second shape is all events without a $D^{*+} \rightarrow D^0\pi^+$. The MC shapes contain slight branching ratio corrections. The ratio of the amounts of the two dominant MC shapes are tied to the MC predicted ratios. ΔM runs from $0.138\,\mathrm{GeV}/c^2$ to $0.165\,\mathrm{GeV}/c^2$ with 50 bins ($0.54\,\mathrm{MeV}/c^2$ wide bins). The proper time runs from 0 cm to 0.1 cm.

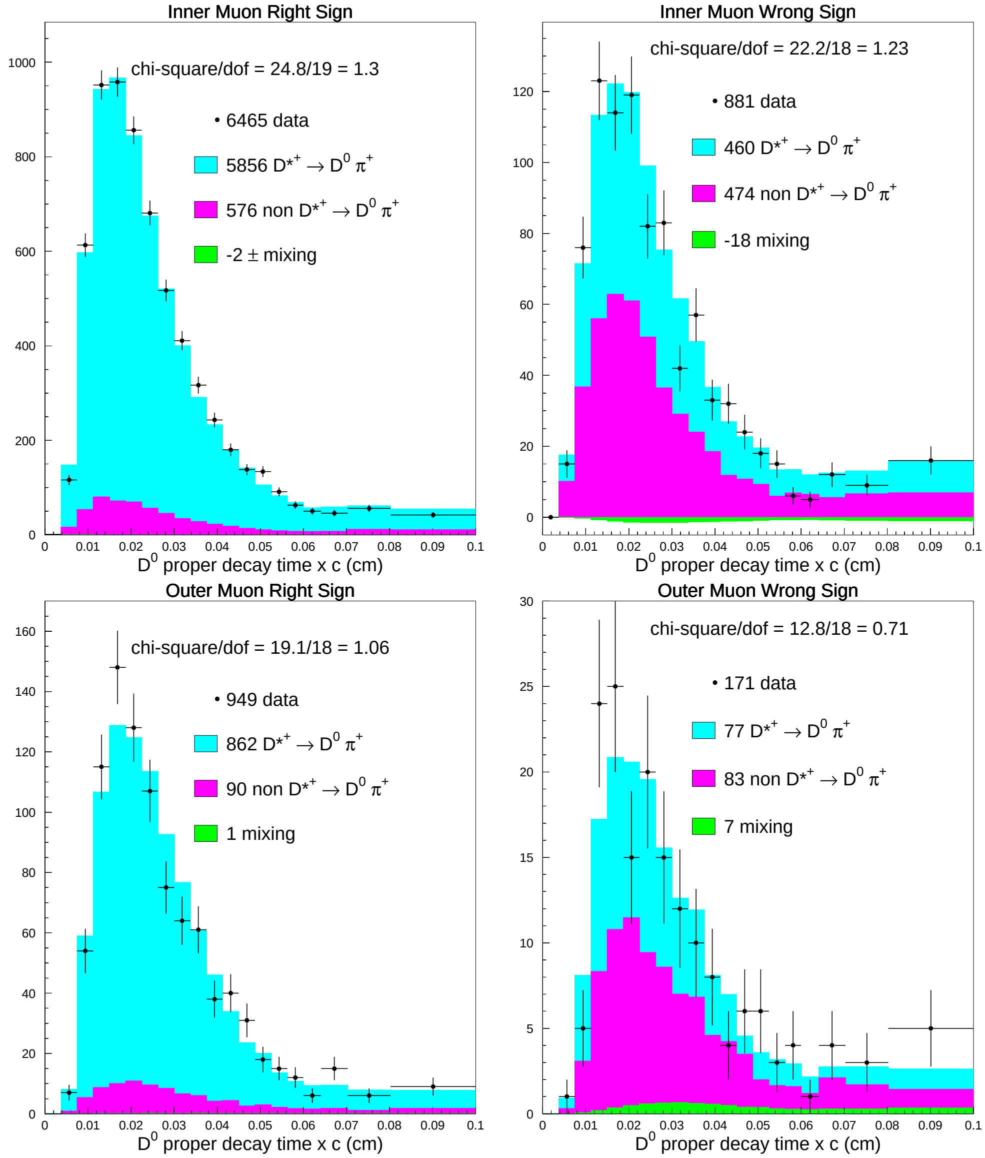


Figure 55: Results of a two parameter, two-dimensional fit to data with Monte Carlo (MC) shapes used in the fit overlaid. Only the D^0 proper decay time dimension is shown. The first MC shape is from events containing a $D^{*+} \rightarrow D^0\pi^+$. The second shape is all events without a $D^{*+} \rightarrow D^0\pi^+$. The MC shapes contain slight branching ratio corrections. The ratio of the amounts of the two dominant MC shapes are tied to the MC predicted ratios. ΔM runs from $0.138\,\mathrm{GeV}/c^2$ to $0.165\,\mathrm{GeV}/c^2$. The proper time runs from 0 cm to 0.1 cm. The first sixteen time bins have a widths of 0.00375 cm. The last four time bins have variable widths. In these units, the mean D^0 lifetime is at 0.0124 cm.

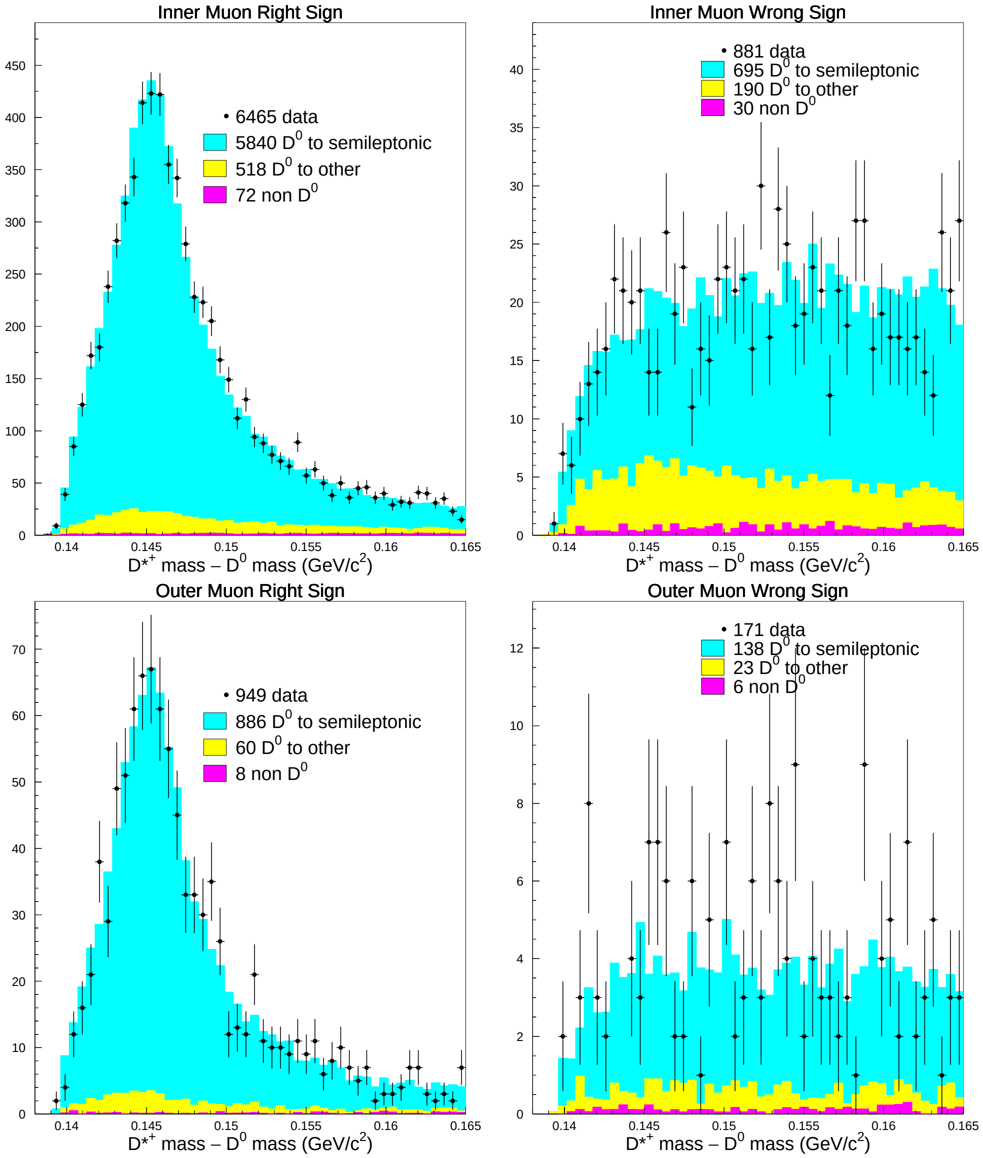
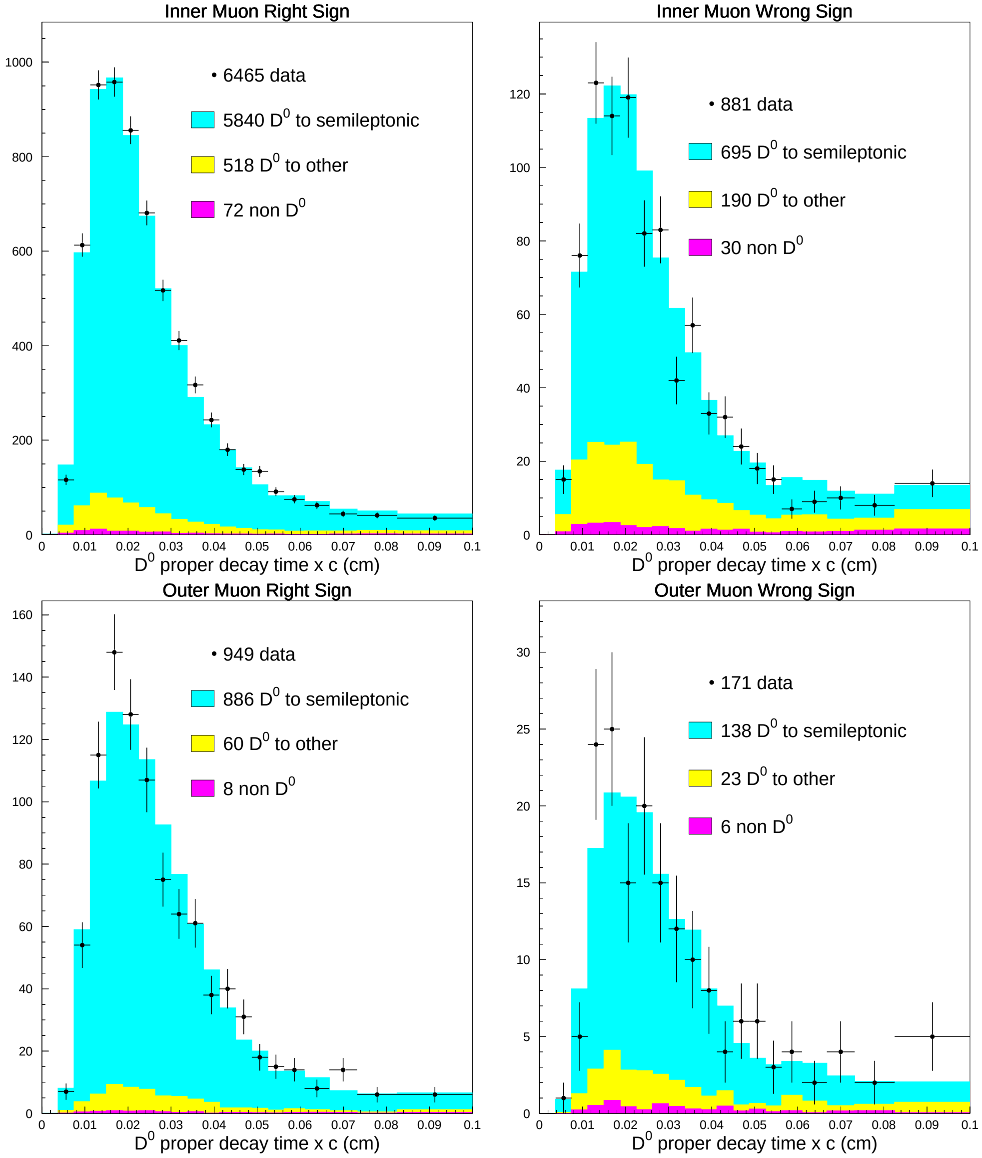


Figure 56: D^0 semileptonic, D^0 hadronic, and non D^0 contributions in the D^{*+} mass dimension. ΔM runs from $0.138\,\mathrm{GeV}/c^2$ to $0.165\,\mathrm{GeV}/c^2$ with 50 bins ($0.54\,\mathrm{MeV}/c^2$ wide bins). The proper time runs from $0\,\mathrm{cm}$ to $0.1\,\mathrm{cm}$.


Inner Muon Right Sign

Inner Muon Wrong Sign

Outer Muon Right Sign

Outer Muon Wrong Sign

Figure 57: D^0 semileptonic, D^0 hadronic, and non D^0 contributions in the D^0 proper time dimension. ΔM runs from $0.138\,\mathrm{GeV}/c^2$ to $0.165\,\mathrm{GeV}/c^2$. The proper time runs from 0 cm to 0.1 cm. The first sixteen time bins have a widths of 0.00375 cm. The last four time bins have variable widths. In these units, the mean D^0 lifetime is at 0.0124 cm.

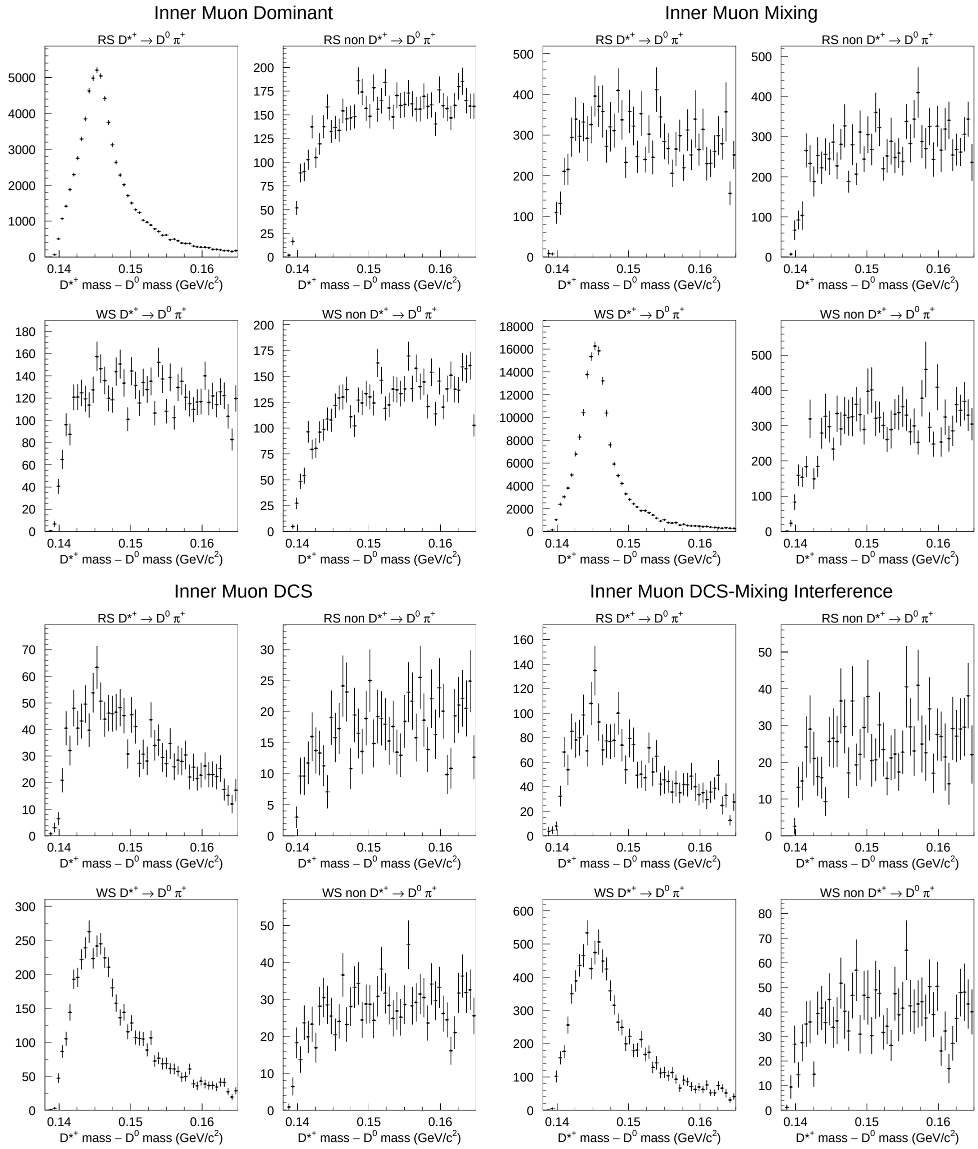


Figure 58: Inner muon D^{*+} mass distributions for the Monte Carlo shapes used in the fit.

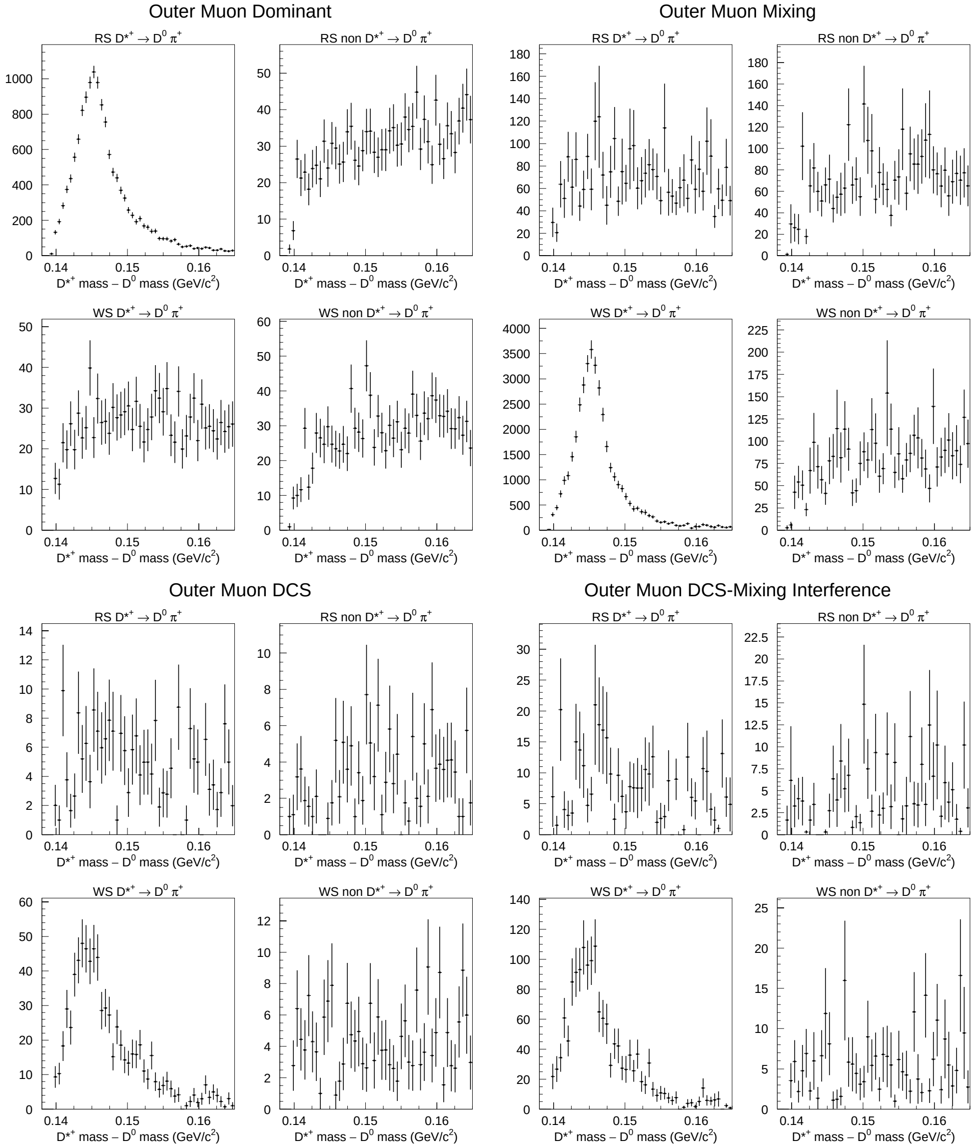


Figure 59: Outer muon D^{*+} mass distributions for the Monte Carlo shapes used in the fit.

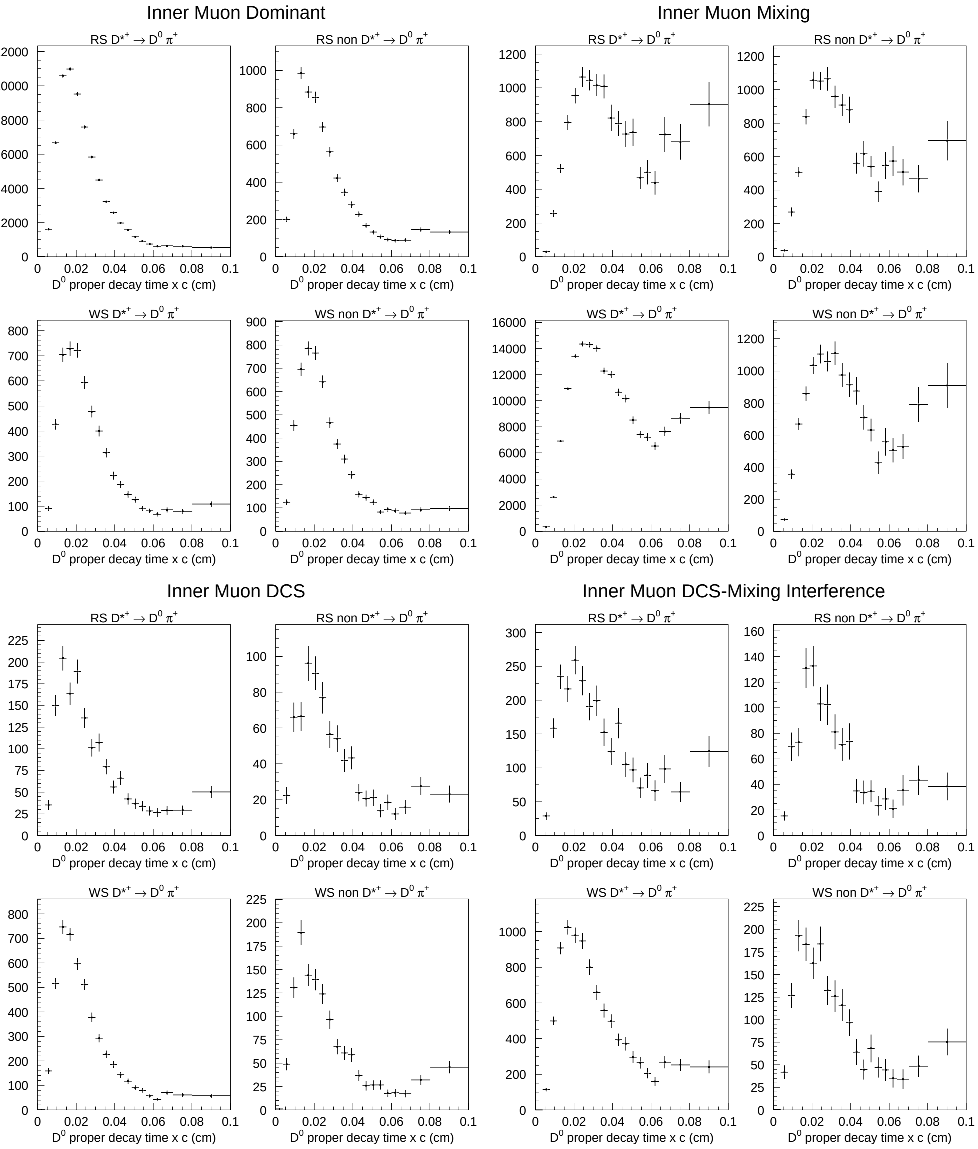


Figure 60: Inner muon D^0 proper time distributions for the Monte Carlo shapes used in the fit.

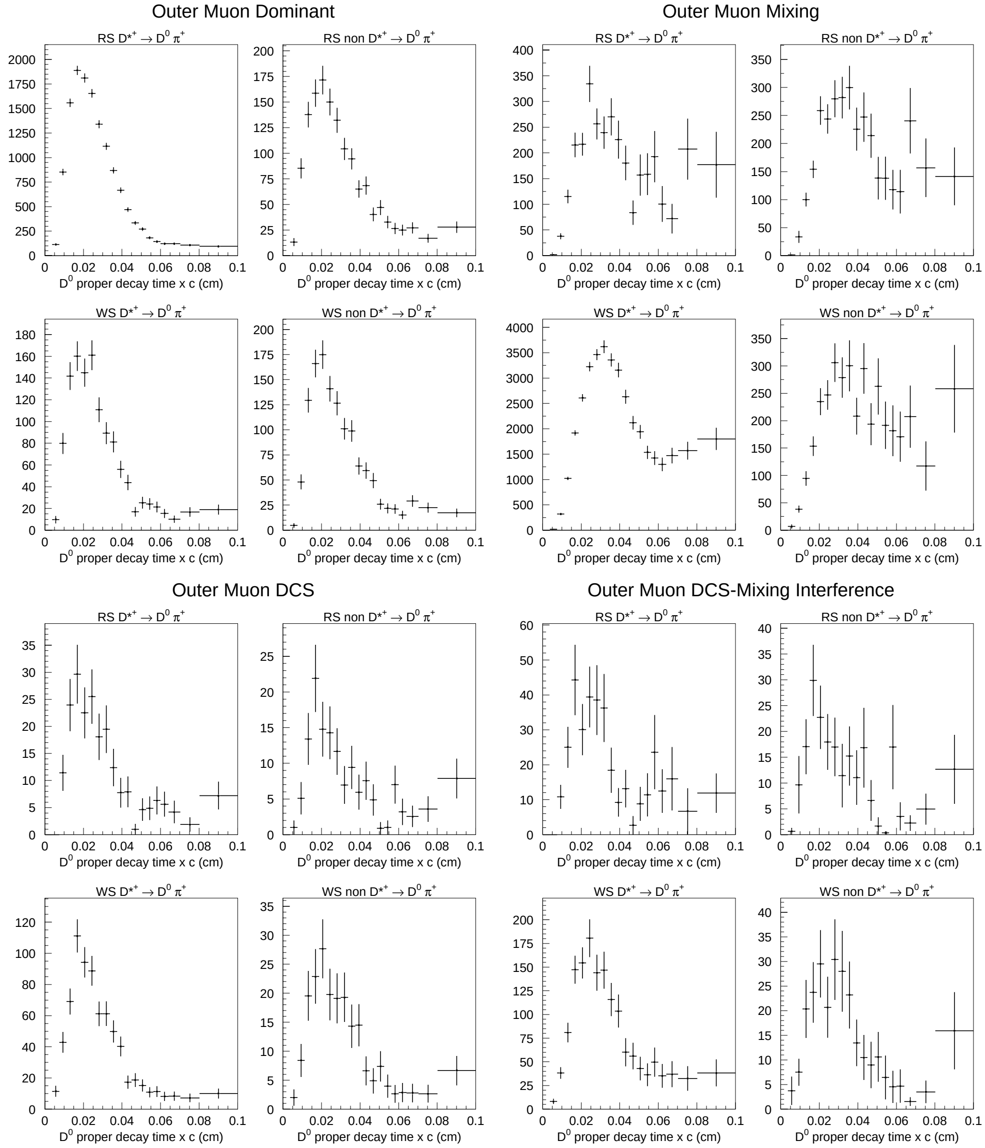


Figure 61: Outer muon D^0 proper time distributions for the Monte Carlo shapes used in the fit.

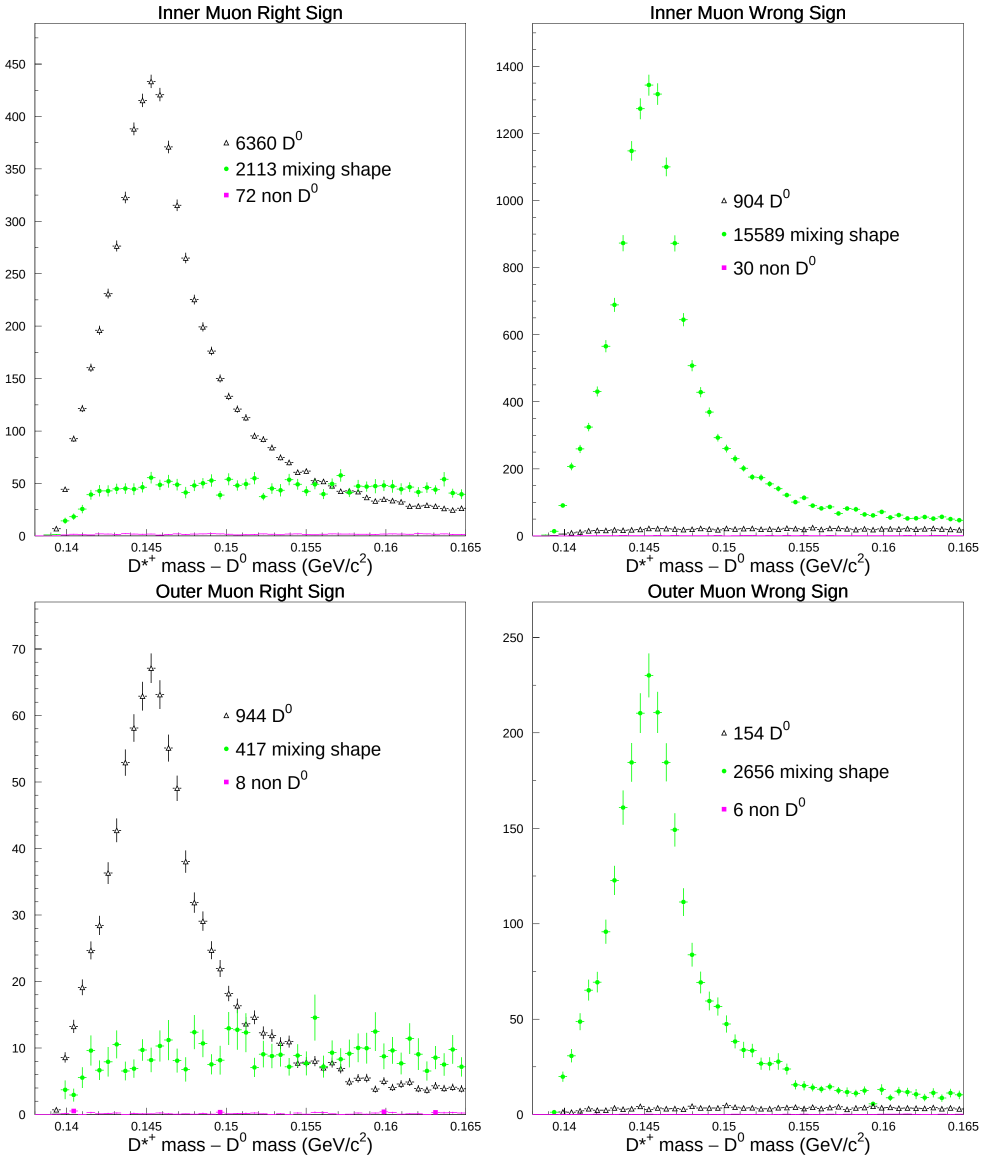


Figure 62: Monte Carlo D^{*+} mass shapes comparing the non mixed D^0, mixing, and non D^0 shapes. The D^0 shapes have hadronic and semileptonic contributions combined. Non mixed shapes are scaled to data levels, mixing shapes are scaled so that multiplying by r_{mix} gives the average number of mixed events in the data.

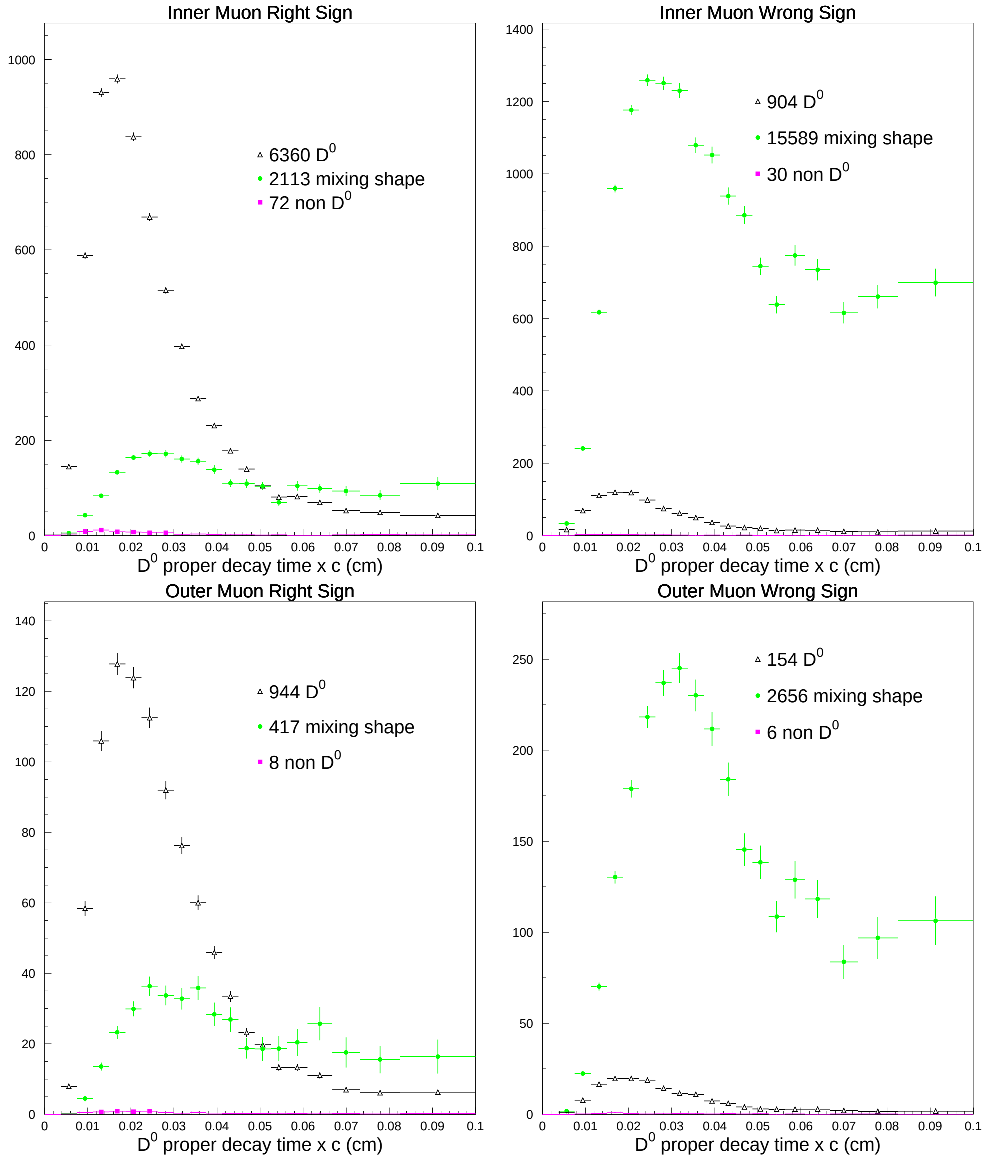


Figure 63: Monte Carlo D^{*+} mass shapes comparing the non mixed D^0, mixing, and non D^0 shapes. The D^0 shapes have hadronic and semileptonic contributions combined. Non mixed shapes are scaled to data levels, mixing shapes are scaled so that multiplying by r_{mix} gives the average number of mixed events in the data.

APPENDIX E

RANKING CUT VARIABLES

Since a large number of variables were considered in preliminary optimizations, some way of ranking the variables according to their ability to distinguish between signal and background was useful in determining the order in which to try various cuts. Once a smaller effective set of variables was found, this ranking was basically ignored.

We can define a variable called `signal` which is equal to 1 for signal events and 0 for background events. Variables are ranked according to their correlation with `signal`. For Gaussian variables, the linear correlation coefficient is a handy measure of correlation. Since all of the variables are non-Gaussian, Kendall's tau [93] is used instead as the measure of correlation between `signal` and the cut variable. Kendall's tau is ideally suited for comparing variables with different functional forms. However, computing Kendall's tau, can be very slow for large samples. The algorithm can be coded to run more efficiently when one of the variables is binary, as is `signal`. 3000 data points are generally sufficient.

The signal and background samples used in computing tau are from Monte Carlo. The signal is RS $D^{*+} \to D^0\pi^+ \to (K^-\mu^+\nu)\pi^+$ and $D^{*+} \to D^0\pi^+ \to (K^{*-}\mu^+\nu)\pi^+$. The soft pion tracks and D^0 daughter tracks are required to match the Monte Carlo particle trajectories for these decays. Mixing proper time dependence is simulated using event rejection on this sample. The background is all WS events.

Table 25 shows the initial ranking of variables considered using Kendall's tau. Table 26 ranks these same variables using linear correlation coefficients. Some powerful cut variables have low ranking because they have already been applied (e.g., muon confidence level, CL_μ). Also, some variables which are very effective for eliminating backgrounds in data have a low ranking. In particular, the number of sigma out of material of the secondary is essential for reducing non-charm backgrounds. Another such variable is $\mathtt{W}(e) - \mathtt{W}(\pi)$ of the π candidate.

Minimal cuts were listed in Sec. VI.7.1. The rankings shown are for inner muons. Outer muons have similar lists. This study used an older, less accurate, Monte Carlo sample than used in the final analysis.

Table 25: Variables ranked according to their correlation with `signal`. Kendall's tau is used as the measure of correlation.

Rank	Variable	Kendall's tau
1	`iso3`	-0.2283
2	L/σ_L	0.1944
3	L	0.1813
4	`iso1`	-0.1752
5	`iso2`	-0.1737
6	$\texttt{W}(K) - \texttt{W}(\pi)$ of π candidate	0.1211
7	K-μ invariant mass	0.1113
8	$\lvert\vec{p}_\pi\rvert$	0.1018
9	$\lvert\vec{p}_K + \vec{p}_\mu\rvert$	0.1002
10	D^{*+} momentum	0.0936
11	q^2	-0.0810
12	$\lvert\vec{p}_\mu\rvert$	0.0803
13	$\lvert\vec{p}_K\rvert$	0.0618
14	Angle between $\vec{p}_K + \vec{p}_\mu$ and $\vec{p}_{D^0}$	-0.0513
15	No. of sigma out of material of secondary	0.0465
16	$\lvert(\vec{p}_K + \vec{p}_\mu) \times \hat{p}_{D^0}\rvert$	-0.0390
17	$\texttt{W}(e) - \texttt{W}(\pi)$ of μ candidate	-0.0311
18	$\texttt{W}(\pi) - \texttt{W}(K)$ of K candidate	0.0309
19	$\texttt{W}(e) - \texttt{W}(\pi)$ of π candidate	0.0280
20	CL_μ	0.0257
21	$c\lvert\Delta t\rvert$ (two solutions)	-0.0121
22	$\mathrm{CL}_{\mathrm{sec}}$	0.0113
23	No. of sigma out of material of primary	0.0098

Table 26: Variables ranked according to their correlation with `signal`. Linear correlation coefficients are used here as the measure of correlation.

Rank	Variable	Linear correlation
1	L/σ_L	0.2709
2	L	0.2473
3	`iso3`	-0.2009
4	$\lvert\vec{p}_K + \vec{p}_\mu\rvert$	0.1655
5	K-μ invariant mass	0.1562
6	$\mathtt{W}(K) - \mathtt{W}(\pi)$ of π candidate	0.1314
7	`iso2`	-0.1265
8	$\lvert\vec{p}_\mu\rvert$	0.1200
9	$\lvert\vec{p}_K\rvert$	0.1145
10	q^2	-0.1093
11	Angle between $\vec{p}_K + \vec{p}_\mu$ and $\vec{p}_{D^0}$	-0.0927
12	`iso1`	-0.0837
13	D^{*+} momentum	0.0809
14	No. of sigma out of material of secondary	0.0768
15	$\lvert\vec{p}_\pi\rvert$	0.0762
16	$\mathtt{W}(e) - \mathtt{W}(\pi)$ of μ candidate	-0.0503
17	$\lvert(\vec{p}_K + \vec{p}_\mu) \times \hat{p}_{D^0}\rvert$	-0.0384
18	$\mathrm{CL}_{\mathrm{sec}}$	0.0348
19	CL_μ	0.0305
20	$\mathtt{W}(\pi) - \mathtt{W}(K)$ of K candidate	0.0150
21	$\mathtt{W}(e) - \mathtt{W}(\pi)$ of π candidate	0.0119
22	$c\lvert\Delta t\rvert$ (two solutions)	-0.0071
23	No. of sigma out of material of primary	-0.0017

APPENDIX F

APPROXIMATE ERROR ON A SMALL PARAMETER

It is possible to obtain a rough estimate of the error on a small parameter without doing a full fit. For this purpose, we can ignore the error due to finite Monte Carlo statistics. Maximizing the likelihood is then equivalent to minimizing

$$-\ln\mathcal{L} = \sum_{i}^{n}(f_i - d_i \ln f_i), \tag{120}$$

where f_i is the expected mean number of entries. If p_1 denotes the small parameter (e.g., r_{mix})

$$f_i = p_1 w_{1i} a_{1i} + \sum_{j\neq 1}^{m} p_j w_{ji} a_{ji}. \tag{121}$$

a_{ji} are the Monte Carlo shapes, w_{ji} are weights, and p_j are parameters determined from preliminary fits (these would be independent of cuts if the Monte Carlo was a perfect representation of data). The above equation can be expressed as

$$\begin{aligned} f_i &= \left(\sum_{j\neq 1}^{m} p_j w_{ji} a_{ji}\right)\left[1 + \frac{p_1 w_{1i} a_{1i}}{\sum_{j\neq 1}^{m} p_j w_{ji} a_{ji}}\right] && (122) \\ x_i &\equiv \frac{w_{1i} a_{1i}}{\sum_{j\neq 1}^{m} p_j w_{ji} a_{ji}}. && (123) \end{aligned}$$

If the second term in the [ ] is small compared with 1, then $\ln f_i$ can be expanded in a series:

$$\ln f_i \simeq \ln\left(\sum_{j\neq 1}^{m} p_j w_{ji} a_{ji}\right) + x_i p_1 - \frac{1}{2}x_i^2 p_1^2. \tag{124}$$

Finding the point where the log-likelihood changes by 1/2 gives the error on p_1, δp_1:

$$-\ln\mathcal{L} - [-\ln\mathcal{L}_0] = \sum_{i}^{n}\left[w_{1i}a_{1i}\delta p_1 - d_i\left(x_i\delta p_1 - \frac{1}{2}x_i^2(\delta p_1)^2\right)\right] = \frac{1}{2}, \tag{125}$$

where $\ln \mathcal{L}_0$ is evaluated at $p_1 = 0$. Collecting these terms:

$$(\delta p_1)^2 \left[\sum_i^n \frac{d_i}{2} \left(\frac{w_{1i} a_{1i}}{\sum_{j \neq 1}^m p_j w_{ji} a_{ji}} \right)^2 \right] + \delta p_1 \left[\sum_i^n \left(p_1 w_{1i} a_{1i} - \frac{d_i w_{1i} a_{1i}}{\sum_{j \neq 1}^m p_j w_{ji} a_{ji}} \right) \right] - \frac{1}{2} = 0. \tag{126}$$

With $d_i \simeq f_{i0} = \sum_{j \neq 1}^m p_j w_{ji} a_{ji}$ (i.e., we model the data with the Monte Carlo predicted shape, excluding p_1) this simplifies to

$$(\delta p_1)^2 \left[\sum_i^n (w_{1i} a_{1i})^2 f_{i0}^{-1} \right] \simeq 1. \tag{127}$$

Finally,

$$\delta p_1 \simeq \left[\sum_i^n (w_{1i} a_{1i})^2 f_{i0}^{-1} \right]^{-\frac{1}{2}} \tag{128}$$

which is easily evaluated. $w_{1i} a_{1i}$ is, for example, the predicted mixing shape and f_{i0} is the expected shape for mixing free data. This approximation assumes the other parameters are known and error free. In the limit of a single bin, this is just the familiar result $\delta p_1 \simeq \sqrt{N_{\text{background}}}/N_{\text{signal}}$.

REFERENCES

[1] Pierre Ramond. *Journeys Beyond the Standard Model.* Perseus Books, 1999. Chapter 8 discusses massive neutrinos. Chapter 9 discusses the QCD vacuum.

[2] J. M. Link et al. Search for CP violation in D^0 and D^+ decays. *Phys. Lett.*, B491:232–239, 2000.

[3] A. Datta and D. Kumbhakar. D^0–$\overline{D}^0$ mixing: A possible test of physics beyond the Standard Model. *Zeitschrift für Physik*, C27:515–522, 1985.

[4] G. Burdman. Charm mixing and CP violation in the Standard Model. In *The Future of High-Sensitivity Charm Experiments*, page 414016. Fermilab, 1994. FERMILAB-Conf-94-200. SLAC SPIRES preprint hep-ph/9407378.

[5] S. L. Glashow, J. Iliopoulos, and L. Maiani. Weak interactions with lepton-hadron symmetry. *Phys. Rev.*, D2:1285–1292, 1970.

[6] A. Pich. Weak decays, quark mixing and CP violation: Theory overview. *Nucl. Phys. Proc. Suppl.*, 66:456–465, 1998.

[7] P. Cea and G. Nardulli. An estimate of the long-distance dispersive contributions to the K_{L}–K_{S} mass difference. *Phys. Lett.*, B152:251–255, 1985.

[8] L. Wolfenstein. D^0–$\overline{D}^0$ mixing. *Phys. Lett.*, B164:170–172, 1985.

[9] J. F. Donoghue et al. Dispersive effects in D^0–$\overline{D}^0$ mixing. *Phys. Rev.*, D33:179–183, 1986.

[10] H. Georgi. D–$\overline{D}$ mixing in Heavy Quark Effective Field Theory. *Phys. Lett.*, B297:353–357, 1992.

[11] T. Ohl, G. Ricciardi, and E. H. Simmons. D–$\overline{D}$ mixing in Heavy Quark Effective Field Theory: The sequel. *Nucl. Phys.*, B403:605–632, 1993.

[12] P. Colangelo, G. Nardulli, and N. Paver. On D^0–$\overline{D}^0$ mixing in the Standard Model. *Phys. Lett.*, B242:71–76, 1990.

[13] T. A. Kaeding. D-meson mixing in broken $SU(3)$. *Phys. Lett.*, B357:151, 1995.

[14] S. E. Csorna et al. Lifetime differences, direct CP violation, and partial widths in D^0 meson decays to K^+K^- and $\pi^+\pi^-$. *Phys. Rev.*, D65:092001, 2002.

[15] E. M. Aitala et al. Branching fractions for $D^0 \to K^+K^-$ and $D^0 \to \pi^+\pi^-$ and a search for CP violation in D^0 decays. *Phys. Lett.*, B421:405–411, 1998.

[16] P. L. Frabetti et al. A measurement of the Cabibbo suppressed decays $D^0 \to \pi^-\pi^+$ and $D^0 \to K^-K^+$. *Phys. Lett.*, B321:295–302, 1994.

[17] I. Bigi. D^0–$\overline{D}^0$ mixing and CP violation in D decays—can there be high impact physics in charm decays? In Lydia V. Beers, editor, *Proceedings: Study of tau, charm and J/ψ physics, development of high luminosity e^+e^- colliders*, pages 169–195. SLAC, 1989. Tau-Charm Factory Workshop, Stanford, CA, May 23–27, 1989.

[18] E. Golowich and A. A. Petrov. Can nearby resonances enhance D^0–$\overline{D}^0$ mixing? *Phys. Lett.*, B427:172, 1998.

[19] J. P. Ma and J. S. Xu. Soft gluon approach for diffractive photoproduction of J/ψ. *Nucl. Phys.*, B640:283–308, 2002.

[20] J. P. Ma. Revisiting spin alignment of heavy meson in its inclusive production. *Nucl. Phys.*, B622:416–428, 2002.

[21] W. Y. Wang and Y. L. Wu. A consistent calculation of heavy meson decay constants and transition wave functions in the complete HQEFT. *Intl. J. of Mod. Phys.*, A16:377–408, 2001.

[22] Z. A. Baccouche et al. Excited heavy baryons and their symmetries III: Phenomenology. *Nucl. Phys.*, A696:638–666, 2001.

[23] Y. Nir and N. Seiberg. Should squarks be degenerate? *Phys. Lett.*, B309:337–343, 1993.

[24] K. S. Babu et al. Fourth generation signatures in D^0–$\overline{D}^0$ mixing and rare D decays. *Phys. Lett.*, B205:540, 1988.

[25] M. Leurer. A comprehensive study of leptoquark bounds. *Phys. Rev.*, D49:333–342, 1994.

[26] H. Nelson. Compilation of D^0–$\overline{D}^0$ mixing predictions. *UCSB HEP 99-08*, 1999. SLAC SPIRES preprint hep-ex/9908021.

[27] E. M. Aitala et al. A search for D^0–$\overline{D}^0$ mixing and doubly-Cabibbo-suppressed decays of the D^0 in hadronic final states. *Phys. Rev.*, D57:13, 1998.

[28] R. Godang et al. Search for D^0–$\overline{D}^0$ mixing. *Phys. Rev. Lett.*, 84:5038–5042, 2000.

[29] J. M. Link et al. A measurement of lifetime differences in the neutral D-meson system. *Phys. Lett.*, B485:62–70, 2000.

[30] E. M. Aitala et al. Search for D^0–$\overline{D}^0$ mixing in semileptonic decay modes. *Phys. Rev. Lett.*, 77:2384, 1996.

[31] P. L. Frabetti et al. Description and performance of the Fermilab E687 spectrometer. *Nucl. Inst. Meth.*, A320:519, 1992. Also in Fermilab publication FERMILAB-Pub-90/258-E.

[32] J. Takács. *Energy Stabilization of Electrostatic Accelerators.* John Wiley and Sons Ltd., 1997. Pages 30–33 discusses the Cockcroft-Walton.

[33] H. Wiedemann. *Particle Accelerator Physics I—Basic Principles and Linear Beam Dynamics.* Springer-Verlag, 1999. Pages 31–33 discusses LINACs, pages 33–46 discusses waveguides , and pages 69–71 discusses synchrotrons.

[34] S. Y. Lee. *Accelerator Physics.* World Scientific Publishing Co., 1999. Pages 314–332 gives a nice introduction to RF systems.

[35] G. Bellini et al. The microstrip vertex detector for the E687 experiment at Tevatron. *Nucl. Inst. Meth.*, A252:366–372, 1986.

[36] William R. Leo. *Techniques for Nuclear and Particle Physics Experiments: A How-to Approach.* Springer-Verlag, second edition, 1994. Lots of useful information.

[37] Amir M. Rahimi. *Amplitude Analysis of the* D^+, $D_s^+ \to K^-K^+\pi^+$ *and* $D^0 \to K_S^0K^+K^-$ *Final States.* PhD thesis, University of Illinois at Urbana-Champaign, 2000. The Čerenkov system is discussed on pages 36–43.

[38] Glen Richard Jaross. *Photoproduction of* $D^{*\pm}$ *and* D^0 *Mesons.* PhD thesis, University of Illinois at Urbana-Champaign, 1991. The Čerenkov system is discussed on pages 25–33.

[39] Ludovico Pontecorvo. Study of the characteristics of single and double gap resistive plate chambers for a muon trigger at LHC. *Nucl. Inst. Meth.*, A344:105–109, 1994.

[40] R. Cardarelli et al. Progress in resistive plate counters. *Nucl. Inst. Meth.*, A263:20–25, 1988.

[41] V. Arena et al. Performance of resistive plate counters with low percentage of isobutane. *Nucl. Inst. Meth.*, A356:264–269, 1995.

[42] R. Frühwirth et al. *Data Analysis Techniques for High-Energy Physics.* Cambridge University Press, 2000.

[43] H. W. Cheung. Proposal for a programmable MG module for FOCUS/E831. *FOCUS Internal Memo*, 1995.

[44] Steve Bracker and Sten Hansen. Description of the Damn Yankee Controller (DYC). SLAC SPIRES preprint hep-ex/0210034, 2002.

[45] Art Kreymer and Francesco Prelz. DAQ architecture for E831. *FOCUS Internal Memo*, 1997.

[46] J. M. Link et al. Čerenkov particle identification in Focus. *Nucl. Inst. Meth.*, A484:270–286, 2002. Also in Fermilab publication FERMILAB-Pub-01-243-E.

[47] Eric W. Vaandering. *Mass and Width Measurements of Σ_c Baryons.* PhD thesis, University of Colorado, 2000.

[48] Torbjörn Sjöstrand et al. High-energy-physics event generation with PYTHIA 6.1. *Comp. Phys. Comm.*, 135:238, 2001. This is available at `http://www.thep.lu.se/~torbjorn/Phythia.html` which also contains additional information.

[49] Christopher Alan Cawlfield. *Studies of the Semileptonic Decays of D Mesons.* PhD thesis, University of Illinois at Urbana-Champaign, 2001.

[50] R. D. Field and R. P. Feynman. A parameterization of the properties of quark jets. *Nucl. Phys.*, B136:1–76, 1978.

[51] René Brun et al. *HBOOK—Statistical Analysis and Histogramming.* CERN, 1995. This is available at `http://wwwinfo.cern.ch/asdoc/`.

[52] Roger Barlow and Christine Beeston. Fitting using finite Monte Carlo samples. *Comp. Phys. Comm.*, 72:219–228, 1993.

[53] F. James. *MINUIT—Function Minimization and Error Analysis.* CERN, 1994–1998. This is available at `http://wwwinfo.cern.ch/asdoc/`.

[54] G. J. Feldman and R. D. Cousins. Unified approach to the classical statistical analysis of small signals. *Phys. Rev.*, D57:3873–3889, 1998.

[55] A. Kolmogorov. Sulla determinazione empirica di una legge disitribuzione. *G. Ist. Ital. Attuari*, 4:83, 1933.

[56] N. V. Smirnov. Table for estimating the goodness of fit of empirical distributions. *Ann. Math. Statist.*, 19:279, 1948.

[57] P. L. Frabetti et al. Analysis of the Cabibbo suppressed decay $D^0 \to \pi^- l^+ \nu$. *Phys. Lett.*, B382:312–322, 1996.

[58] A. Pais and S. B. Treiman. Charmed-meson lifetime ratios and production in e^+-e^- collisions. *Phys. Rev.*, D15:2529–2532, 1977.

[59] P. L. Frabetti et al. Analysis of the decay mode $D^0 \to K^- \mu^+ \nu_\mu$. *Phys. Lett.*, B364:127–136, 1995.

[60] Particle Data Group. *Review of Particle Physics.* Springer, 2000. Published in The European Physical Journal C, vol. 15, number 1–4.

[61] J. Adler et al. Reanalysis of charmed–D-meson branching fractions. *Phys. Rev. Lett.*, 60(2):89–92, 1988.

[62] S. Barlag et al. A reanalysis of branching fractions of charmed mesons D^0, D^+ and D_s^+. *Zeitschrift für Physik*, C55:383–390, 1992.

[63] M. Aguilar-Benitez et al. Exclusive hadron branching ratios of the D meson. *Zeitschrift für Physik*, C36:559, 1987. Also erratum in 1988 ZPHY C38 520.

[64] Daniel J. Engh. *A search for physics beyond the Standard Model through the three-body rare and forbidden charm decays* $D^+, D_s^+ \rightarrow K^+\mu^+\mu^-, K^-\mu^+\mu^+, \pi^+\mu^+\mu^-, \pi^-\mu^+\mu^+, \mu^+\mu^+\mu^-$. PhD thesis, Vanderbilt University, 2002.

[65] J. M. Link et al. A high statistics measurement of the Λ_c^+ lifetime. *Phys. Rev. Lett.*, 88:161801, 2002.

[66] H. W. Cheung. Systematics is not all voodoo science: Λ_c^+ lifetime as a case study. *FOCUS Internal Memo*, 2001.

[67] J. M. Link et al. Measurements of the Σ_c^0 and Σ_c^{++} mass splittings. *Phys. Lett.*, B488:218, 2000.

[68] M. Aleppo et al. Charm decays and D^0–$\overline{D}^0$ mixing. *Nucl. Phys.*, A675:291C–296C, 1996.

[69] K. Abe et al. Measurement of the lifetime difference in D^0 meson decays. *Phys. Rev. Lett.*, 88:162001, 2002.

[70] E. M. Aitala et al. Measurements of lifetimes and a limit on the lifetime difference in the neutral D-meson system. *Phys. Rev. Lett.*, 83:32, 1999.

[71] G. Volkov, V. A. Monich, and B. V. Struminski. On CP violation and mixing mechanisms in neutral K and D meson systems. *Yad. Fiz.*, 34:435–450, 1981. English translation in Sov. J. Nucl. Phys.

[72] A. Datta. D^0–$\overline{D}^0$ mixing: Standard versus nonstandard scenarios. *Phys. Lett.*, B154:287, 1985.

[73] G. L. Kane and R. Thun. Searches for effects of flavor changing neutral currents. *Phys. Lett.*, B94:513, 1980.

[74] R. N. Cahn and H. Harari. Bounds on the masses of neutral generation changing gauge bosons. *Nucl. Phys.*, B176:135, 1980.

[75] B. Mukhopadhyaya, A. Raychaudhuri, and A. Ray-Mukhopadhyaya. D^0–$\overline{D}^0$ mixing in superstring inspired $E(6)$ models. *Phys. Lett.*, B190:93, 1987.

[76] T. P. Cheng and M. Sher. Mass matrix ansatz and flavor nonconservation in models with multiple Higgs doublets. *Phys. Rev.*, D35:3484, 1987.

[77] L. Hall and S. Weinberg. Flavor changing scalar interactions. *Phys. Rev.*, D48:979–983, 1993.

[78] J. L. Hewett. *The Albuquerque Meeting.* World Scientific, River Edge, N.J., 1995. Edited by Sally Seidel. pp. 951–955.

[79] G. Burdman. *Workshop on the Tau/Charm Factory.* AIP Press, Woodbury, N.Y., 1995. Edited by Jose Repond. pp. 409–424.

[80] D. Atwood, L. Reina, and A. Soni. Phenomenology of two Higgs doublet models with flavor changing neutral currents. *Phys. Rev.*, D55:3156–3176, 1997.

[81] I. Kakebe and K. Yamamoto. Flavor nonconservation and CP violation from quark mixings with singlet quarks. *Phys. Lett.*, B416:184–191, 1998.

[82] U. Egede. Mixing in the D^0–$\overline{D}^0$ system at BaBar. Talk presented at the First International Workshop on Frontier Science, Oct. 6–11, 2002, INFN Laboratory, Frascati, Italy.

[83] M. Gronau, Y. Grossman, and J. L. Rosner. Measuring D^0–$\overline{D}^0$ mixing and relative strong phases at a charm factory. *Phys. Lett.*, B508:37–43, 2001.

[84] Z. Z. Xing. New time distributions of D^0–$\overline{D}^0$ or B^0–$\overline{B}^0$ mixing and CP violation. *Phys. Lett.*, B463:323–329, 1999.

[85] M. Artuso. Beyond the Standard Model studies at CLEO-c—introductory considerations. Talk presented at the CLEO-c Workshop. `http://http://www.lns.cornell.edu/public/CLEO/CLEO-C/sciprog.html`, May 2001.

[86] G. Molière. *Z. Naturforsch*, 3a:78, Jul 1948.

[87] H. A. Bethe. Moliere's theory of multiple scattering. *Phys. Rev.*, 89(6):1256–1266, Mar 1953.

[88] G. Shen et al. Measurement of multiple scattering at 50 to 200 GeV/c. *Phys. Rev.*, D20(7):1584–1588, Oct 1979.

[89] S. A. Akimenko et al. Multiple Coulomb scattering of 7.3 to 11.7 GeV/c muons on a Cu target. *Nucl. Ins. Meth.*, A243:518–522, 1986.

[90] B. Gottschalk et al. Multiple Coulomb scattering of 160 MeV protons. *Nucl. Ins. Meth.*, B74:467–490, 1993.

[91] G. R. Lynch and O. I. Dahl. Approximations to multiple Coulomb scattering. *Nucl. Ins. Meth.*, B58:6–10, 1991.

[92] J. M. Fernández-Varea et al. On the theory and simulation of multiple elastic scattering of electrons. *Nucl. Ins. Meth.*, B73:447–473, 1993.

[93] William H. Press et al. *Numerical Recipes: The Art of Scientific Computing.* Cambridge University Press, 1986.

[94] Hanson, Lanzl, Lyman, and Scott. *Phys. Rev.*, 84:634, 1951.

[95] G. Molière. Theorie der Streuung schneller geladener Teilchen III. *Z. Naturforsch*, 10a:177, 1955.

[96] R. M. Sternheimer, M. J. Berger, and S. M. Seltzer. Density effect for ionization loss. *At. Data and Nucl. Data Tables*, 30:262, 1984.

[97] J. Wiss. Thoughts on muon identification algorithms for E831. *FOCUS Internal Memo*, Sep 1994.

[98] Particle Data Group. Review of particle properties. *Phys. Rev.*, D54(1):134–135, Jul 1996.

[99] L. Eyges. Multiple scattering with energy loss. *Phys. Rev.*, 74:1534–1535, 1948.

[100] B. Rossi and K. Greisen. *Rev. Mod. Phys.*, 13:267, 1941.

[101] H. Bichsel, K. M. Hanson, and M. E. Schillaci. Lateral displacement in small angle multiple scattering. *Phys. Med. Biol.*, 27(7):959–961, 1982.

[102] S. M. Seltzer and M. J. Berger. *Int. J. Applied Rad.*, 33:1193, 1982.

[103] R. M. Sternheimer and R. F. Peierls. General expression for the density effect for the ionization loss of charged particles. *Phys. Rev.*, B3(11):3681–3692, 1971.

[104] C. Cawlfield, M. Ruesink, and J. Wiss. Muon identification χ^2 confidence levels. *FOCUS Internal Memo*, Dec 1994.

[105] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions.* Dover Publications, Inc., 1972.