A SEARCH FOR PHYSICS BEYOND THE STANDARD MODEL THROUGH THE THREE–BODY RARE AND FORBIDDEN CHARM DECAYS $D^+, D_s^+ \to K^+\mu^+\mu^-,\ K^-\mu^+\mu^+,\ \pi^+\mu^+\mu^-,\ \pi^-\mu^+\mu^+,\ \mu^+\mu^+\mu^-$

By

Daniel J. Engh

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

May 2002

Nashville, Tennessee

Approved: Date:

Physics

A SEARCH FOR PHYSICS BEYOND THE STANDARD MODEL THROUGH THE THREE–BODY RARE AND FORBIDDEN CHARM DECAYS $D^+, D_s^+ \rightarrow K^+\mu^+\mu^-,\ K^-\mu^+\mu^+,\ \pi^+\mu^+\mu^-,\ \pi^-\mu^+\mu^+,\ \mu^+\mu^+\mu^-$

DANIEL J. ENGH

Dissertation under the direction of Professor Paul Sheldon

The Standard Model forbids lepton number violating (LNV) decays and predicts branching ratios for flavor changing neutral current (FCNC) charm decays below present experimental sensitivities. An observation of these rare or forbidden decays in current experiments would indicate new physics beyond the Standard Model. Fermilab experiment E831 (FOCUS) is a high statistics, fixed target, charm photoproduction experiment with the largest sample of fully reconstructed charm particles to date.

This dissertation presents an analysis of the ten rare and forbidden decays D^+, $D_s^+ \rightarrow K^+\mu^+\mu^-,\ K^-\mu^+\mu^+,\ \pi^+\mu^+\mu^-,\ \pi^-\mu^+\mu^+,\ \mu^+\mu^+\mu^-$. A novel analysis method was adopted using bootstrap data re-sampling techniques to reduce bias. No decays violating the Standard Model were observed and new upper limits on their branching ratios are reported. These limits represent an improvement on previous limits by factors of 1.5–10 and illustrate a new technique for analyses of rare signals.

Approved______________________________ Date__________

ACKNOWLEDGEMENTS

The results presented in this thesis are the handiwork of hundreds of people. The FOCUS collaboration, Fermilab, and the U.S. Department of Energy all form a backdrop of support without which a high-energy physics experiment of this type would not be possible. I would like to thank the members of the FOCUS collaboration for their long efforts in the design and execution of the FOCUS experiment. The analysis in this dissertation is based upon their years of efforts developing the tools to analyze FOCUS results. I would like to thank all the members of FOCUS, and especially Angel Lopez and Wolfgang Rolke for their efforts in improving methods for rare decay analyses.

I would like to thank the people at Vanderbilt for all their help in my graduate studies. Paul Sheldon, Medford Webster, and Will Johns form a first-rate research group with excellent resources and experience. Thanks to each one of them for their assistance and support. Thanks also to Mike Hosack for help with the outer muon system, to Eric Vaandering for his help running Monte Carlo, and Kevin Stenson for his extensive help in the details of this analysis.

Thanks to the students at Vanderbilt for their personal support, especially: Scott Adams, Lionel Crews, Tracy Huard, Szabi and Zsuzsa Marka, Mike Papantonakis, Katrina Wagner, and D.J. Wagner. I enjoy all the good memories and our friendships that will last a lifetime.

Finally I would like to thank my family. Thanks to my sister, brothers, nieces, and nephews for all our letters, phone calls, visits and vacations together, all of which really uplifted and sustained me. And a special thanks go to Mom and Dad, for their undying love and support in all that I do.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

PDG	Particle Data Group
PWC	Proportional Wire Chamber
RPC	Resistive Plate Chamber
FCNC	Flavor Changing Neutral Current
M1	Analysis Magnet 1
M2	Analysis Magnet 2
MH	Muon Hodoscope
LPS1	Scaled 2nd level Hadronic Trigger
LNV	Lepton Number Violating Decay Mode
E831	Fermilab Experiment Number 831 (aka FOCUS)
FOCUS	Fermilab Experiment Number 831 (aka E831)
SSD	Silicon Microstrip Detector
EZDEE	Easy D reconstruction package for FOCUS
DVERT	Vertex reconstruction package for FOCUS
PASS1	First "pass" thru the data converting raw detector data to reconstructed events
MCFOCUS	FOCUS Monte Carlo software package
PYTHIA	Standard-Model simulation software from University of Lund
SUSY	Supersymmetry extension to the Standard Model
CCBAR	Charm, charm-bar Monte Carlo

CHAPTER I

INTRODUCTION

It is an ancient belief that the universe is made of simple materials governed by a set of universal laws. The quest for this knowledge has driven the rise of science into the present age. Our current understanding of the structure of the universe is summarized in a set of theories and parameters known as the Standard Model. With great precision, this model describes the composition and behavior of matter — from macroscopic materials to subatomic particles. Despite the success of the Standard Model, many basic questions remain unanswered. Laboratories worldwide are conducting experiments on immense scales to probe further into the nature of matter.

1.1 The Standard Model

The Standard Model is a set of ideas that describes the structure of matter and many of its interactions [1]. Twelve fundamental particles provide the building blocks, and four forces govern their interactions. Fundamental particles are defined here to be the smallest constituents of matter – localized units that contain the smallest amounts of energy, momentum, charge, spin, and other measurable quantities. The particles are split into two groups: leptons and quarks; and the forces are currently summarized in two quantum field theories, quantum chromodynamics (QCD) and quantum electrodynamics (QED). The major omission from the Standard Model is gravity. Although gravity is the dominant force on the largest scales, gravity is the weakest of the four forces. For experimental particle physics, gravitational effects are negligible and are ignored.

The twelve fundamental particles exhibit symmetries that group them into three doublets of leptons and quarks as shown in Equation (1). The doublets are ordered by increasing mass, with the lightest doublets on the left. The higher mass (or energy) states are unstable and these particles quickly decay to the lowest mass states. Since the lowest mass states are the most stable, most of the matter of the universe consists of particles in the lowest mass states. In addition, each particle has an anti-particle with identical properties except for opposite charge (and opposite particle number). All twelve particles have been experimentally detected.

$$\begin{array}{ccc} \begin{pmatrix} \nu e \\ e \end{pmatrix} & \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix} & \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix} \\ \begin{pmatrix} u \\ d \end{pmatrix} & \begin{pmatrix} c \\ s \end{pmatrix} & \begin{pmatrix} t \\ b \end{pmatrix} \end{array} \tag{1}$$

The familiar electron, discovered in 1897, is the lightest of the charged leptons and defines the basic unit of electric charge, $|e|$. The charged leptons (e, μ, τ), called the electron, muon and tau, all have the same electric charge $-1|e|$. Each of these leptons is grouped with an uncharged lepton, or neutrino $(\nu_e, \nu_\mu, \nu_\tau)$, which has very low (or possibly no) mass. The quarks (u, d, c, s, t, b) are called up, down, charm, strange, top, and bottom, respectively. The u, c, t quarks have charge $+\frac{2}{3}|e|$ and the d, s, b quarks have charge $-\frac{1}{3}|e|$. Quarks have never been detected in isolation, but are deduced to be the fundamental units of all hadronic matter. Quarks are always detected in groups of three quarks or three anti-quarks, which are called baryons, or quark anti-quark pairs, which are called mesons. The particles discussed in this thesis are predominantly mesons. The familiar neutron and proton are the lowest mass states of baryons, composed of u and $\overline{d}$ quarks. Pions are the lowest mass states of mesons, also composed of u and $\overline{d}$ quarks.

Table 1: Properties of quarks and leptons [2].

Quarks			Leptons		
Flavor	Mass (MeV/c^2)	Charge	Flavor	Mass (MeV/c^2)	Charge
d	3–9	$-\frac{1}{3}$	e	0.5110	-1
u	1.5–5	$+\frac{2}{3}$	μ	105.66	-1
s	50–200	$-\frac{1}{3}$	τ	1777.1	-1
c	1200–1800	$+\frac{2}{3}$	ν_e	< 2.8 eV/c^2	0
b	4000–5000	$-\frac{1}{3}$	ν_μ	< 0.17 MeV/c^2	0
t	$(173.8 \pm 5.2) \times 10^3$	$+\frac{2}{3}$	ν_τ	< 18.2 MeV/c^2	0

Four forces govern the interactions of particles: the gravitational, electromagnetic, weak nuclear, and strong nuclear forces. The gravitational and electromagnetic forces have long been recognized as forces of nature. In the 20th century it was discovered that a new force, termed the weak force, is responsible for radioactivity and that a fourth force, the strong force, holds protons and neutrons together in the atomic nucleus.

Particles called gauge bosons, listed in Table 2, mediate these forces. All particles with mass interact via the gravitational force, but this force is not included in the Standard Model. The graviton is postulated to be the gravitational force carrier, but has not been detected experimentally. Particles with electric charge interact via the electromagnetic force, mediated by photons (γ). Quarks and leptons can directly change flavors (decay into other quarks and leptons) only via weak interactions, and these interactions are mediated by the massive $W^{\pm}$ and Z^0 bosons. The fourth force is the strong nuclear force, mediated by gluons (g), which binds quarks together to form mesons and baryons. This force is responsible for quark confinement − the fact that quarks are never seen isolated.

According to the standard model, neutrinos interact only weakly with other particles (they generally pass through the earth undetected); charged leptons interact with other particles primarily via the weak and electromagnetic interaction; and quarks are governed by weak, electromagnetic, and strong forces.

Table 2: Properties of the gauge bosons [2].

Boson	Force	Mass (GeV/c^2)	Spin
graviton	Gravity	0	2
γ (photon)	Electromagnetic	0	1
$W^{\pm}, Z^0$	Weak	$80.4, 91.2$	1
g (gluon)	Strong	0	1

1.2 Charm Physics and Weak Decays

Particle decays that involve quark flavor changes are the result of weak force interactions. Experimentally these weak decays are measured in terms of the particle masses and decay rates, or expressed somewhat differently, in terms of particle mass resonances (or "eigenstates") and weak decay resonances. The quark doublets in Equation (1) are expressed in the basis of mass eigenstates. The following section discusses the relationship of mass and weak eigenstates.

In the 1960s, only the u, d, and s quarks were known. Weak decays had been observed that implied $s \to u$ transitions. The eigenstates of weak decays were modelled [1] as a single isospin doublet

$$\mathcal{Q} = \begin{pmatrix} u \\ d \end{pmatrix} \tag{2}$$

using a modified version of Fermi's V-A theory.[1] For a weak particle interaction the amplitude of the weak interaction, $\mathcal{M}$, is represented as a function of the weak current:

$$\mathcal{M} = \frac{4G}{\sqrt{2}} J^{\mu} J^{\dagger}_{\mu} \tag{3}$$

where G represents the Fermi weak coupling constant, and J^{μ} is the weak current 4-vector. The charged currents for a $u \to d$ transition are

$$J^{+}_{\mu} = \overline{u}_L \gamma_{\mu} d_L \quad , \qquad J^{-}_{\mu} = \overline{d}_L \gamma_{\mu} u_L \tag{4}$$

and the neutral currents can be approximated[2] by

$$J^{3}_{\mu} = \overline{u}_L \gamma_{\mu} u_L + \overline{d}_L \gamma_{\mu} d_L \tag{5}$$

and $J^{+}_{\mu}, J^{-}_{\mu}, J^{3}_{\mu}$ can be combined into a triplet of weak currents

$$J^{i}_{\mu} = \overline{\mathcal{Q}_L} \gamma_{\mu} \frac{\tau^i}{2} \mathcal{Q}_L \ , \qquad\qquad i = 1, 2, 3 \tag{6}$$

where (since weak isospin current couples only left handed fermions) $\mathcal{Q}_L$ is the left handed isospin doublet for d and u, and τ^i are the Pauli spin matrices.

[1]This is a low energy approximation which ignores the boson propagator terms, for example.

[2]This simplified example serves merely to illustrate the suppression flavor changing neutral currents (FCNC) by the GIM mechanism.

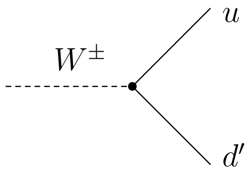


Figure 1: Weak Current Diagram

This formalism holds for both leptons and quarks, but the quark sector is complicated by the fact that the weak eigenstates do not correspond exactly with the mass eigenstates. In the 1960s, only u, d, and s quarks were known, so $d \to u$ as well as $s \to u$ transitions led Cabibbo to construct a weak isospin doublet

$$\mathcal{U} = \begin{pmatrix} u \\ d' \end{pmatrix} \tag{7}$$

where $d' = d\cos\theta_c + s\sin\theta_c$ is a mixture of d and s mass eigenstates, parameterized by the Cabibbo angle θ_c [3]. Since the weak isospin current has the form

$$J^i_\mu = \overline{\mathcal{U}_L}\gamma_\mu\frac{\tau^i}{2}\mathcal{U}_L \qquad\qquad i = 1, 2, 3 \tag{8}$$

a problem arises. The resulting amplitude includes a flavor changing neutral current (FCNC) term

$$\overline{d}_L\gamma_\mu s_L\cos\theta_c\sin\theta_c \tag{9}$$

with a probability similar to (flavor changing) charged currents. These FCNC decays (e.g. $K^+ \to \pi^+\mu^+\mu^-$) were within experimental sensitivities of the time, but were not observed.

1.2.1 GIM Mechanism

A solution to the problem was proposed in 1970 by Glashow, Iliopoulos, and Maiana [4] with the introduction of a new quark, charm, and a second weak isospin doublet

$$\mathcal{U} = \begin{pmatrix} u \\ d' \end{pmatrix} , \quad \mathcal{C} = \begin{pmatrix} c \\ s' \end{pmatrix} \tag{10}$$

where $s' = -d \sin\theta_c + s\cos\theta_c$. The mixing between the mass and the weak eigenstates can be expressed as the unitary Cabibbo rotation matrix

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix} . \tag{11}$$

The weak isospin current for weak quark interactions then becomes

$$J^i_\mu = \overline{\mathcal{U}_L}\gamma_\mu \frac{\tau^i}{2}\mathcal{U}_L \; + \; \overline{\mathcal{C}_L}\gamma_\mu \frac{\tau^i}{2}\mathcal{C}_L \qquad\qquad i = 1, 2, 3 \tag{12}$$

and the flavor changing neutral current (FCNC) terms exactly cancel.

Charm was introduced to remove unwanted terms in the formalism describing $u, s,$ and d quarks. This solution, known as the GIM mechanism, was initially thought to be artificial, since no particles containing charm quarks had been detected. But the GIM mechanism introduced satisfying symmetries into weak quark interactions, and further developments extended the quark sector with a third family, (with the top and bottom, or truth and beauty, quarks). The Cabibbo rotation matrix is then replaced with the 3×3 unitary CKM [5] matrix

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} . \tag{13}$$

The discovery of a $c\bar{c}$ resonance, named the J/ψ particle in 1974 and the dramatic explosion of new particles discovered in the next few years confirmed the GIM mechanism and the existence of three quark families. The development of the electroweak theory combined these developments into a renormalizable theory which is the current theoretical framework for weak interactions in the Standard Model.

1.3 Rare and Forbidden Decays

The term "rare decays" here refers to FCNC decays forbidden at tree level and suppressed at higher orders by the GIM mechanism. Forbidden decays are strictly forbidden at all orders and have branching ratios of exactly zero according to the Standard Model. The forbidden decays considered here fall under the categories of lepton number violating (LNV) and lepton family number violating (LFNV) modes. Any observation of a forbidden decay suggests new physics. Since the branching ratios of the the rare decays considered here are below FOCUS sensitivities, any observation of these rare decays in FOCUS data suggests new physics.

1.3.1 Flavor Changing Neutral Currents

The FCNC decay channels in this analysis have the form

$$D \to X\mu^+\mu^- \tag{14}$$

where $D = D_s^+$ or D^+ and $X = K^+$ or π^+. The Standard Model branching ratios for these decays are summarized as follows.

The rare decay $D^+ \to \pi^+\mu^+\mu^-$ has a branching ratio close to, but below, FOCUS sensitivities. The $D^+ \to \pi^+\phi \to \pi^+l^+l^-$ transition has a branching ratio of

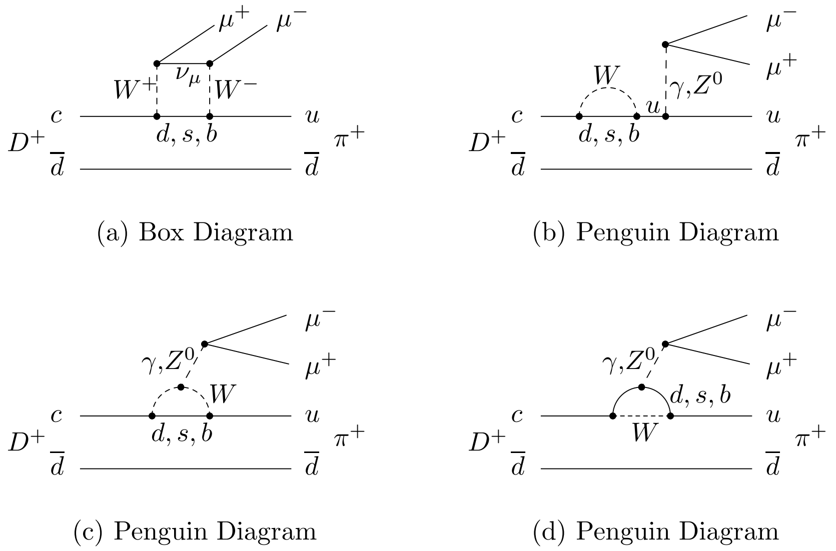


Figure 2: Example lowest order short range contributions to the FCNC decay $D^+ \to \pi^+\mu^+\mu^-$. The branching ratio for these contributions is estimated to be $\sim 2 \times 10^{-8}$.

$\sim 1.5{\times}10^{-6}$. Final state interactions for this decay have been calculated in 1998 by Singer and Zhang [6], using a vector meson dominance (VMD) mechanism to calculate long distance contributions away from the ϕ resonance. Including these contributions, the branching ratio becomes $\sim 1.9{\times}10^{-6}$.

Short-range contributions to this decay were calculated by Schwartz [7], adapting the method of Inami and Lim [8] (used for rare K decays). Electromagnetic and Z^0 penguin diagrams as well as the W-box diagram represent the lowest order contributions, as shown in Figure 1.3.1. The branching ratio of the inclusive $c \to u\ l^+l^-$ process is 1.8×10^{-8}. The exclusive process $D^+ \to \pi^+l^+l^-$ is expected to be $\sim 10\%$ of the inclusive rate leading to branching ratios $\sim 10^{-9}$ for the short range contributions.

Short range contributions to the FCNC decay $D^+ \to K^+\mu^+\mu^-$ are estimated to be $\sim 10^{-15}$ [7]. There are no Standard Model long range interactions for this decay

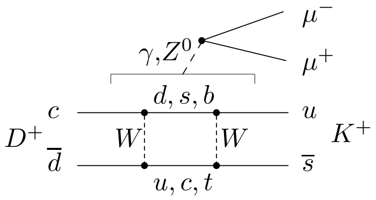
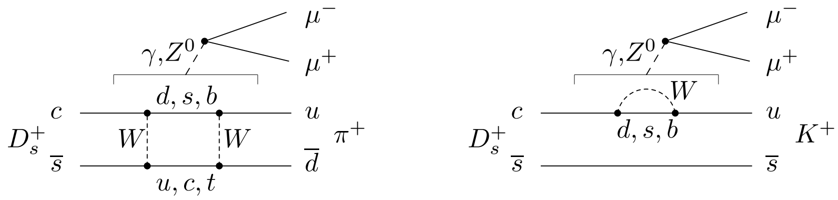


Figure 3: Lowest order short range contributions to $D^+ \to K^+\mu^+\mu^-$. (The bracket indicates that the γ or Z^0 line can emanate from any of the quark or W lines.)



(a) $D_s^+ \to \pi^+\mu^+\mu^-$ (b) $D_s^+ \to K^+\mu^+\mu^-$

Figure 4: Lowest order short range contributions to D_s^+ FCNC decays. The $D_s^+ \to K^+\mu^+\mu^-$ diagrams are the same as for $D^+ \to \pi^+\mu^+\mu^-$ (Figure 1.3.1) except for the spectator quark. (The bracket indicates that the γ or Z^0 line can emanate from any of the (non-spectator) quark or W lines.)

mode and the branching ratios are far below experimental sensitivities. The branching ratios are summarized in Table 3.

Table 3: Predicted FCNC Branching Ratios

Decay Mode	$BR_{\phi-resonance}$	BR_{short}	BR_{long-D^*}	$BR_{long-\phi}$
$D^+ \to \pi^+\mu\mu$	1.9×10^{-6}	2×10^{-8}	$\sim 10^{-8}$	6×10^{-7}
$D^+ \to K^+\mu\mu$	$< 5 \times 10^{-8}$		$\sim 10^{-15}$	
$D_s^+ \to \pi^+\mu\mu$	1.3×10^{-5}			
$D_s^+ \to K^+\mu\mu$	$< 2 \times 10^{-7}$	1×10^{-8}		

1.3.2 Lepton number violating.

For all particle decays, the number of leptons in the decay products has always been observed to equal the number of leptons in the original state. We are motivated to search for violations of this rule since lepton number conservation has no theoretical justification in the Standard Model. The three-body lepton number violating (LNV) modes in D decays have the form

$$D \to X\mu^+\mu^+ \tag{15}$$

where $D = D_s^+$ or D^+ and $X = K^-$ or π^-.

1.3.3 Lepton family number violating

Similarly, lepton family number has always been conserved in hadronic decays. As shown in Equation (1), the leptons are paired in three generations or families. Thus for a decay with a parent lepton family number of zero for all three lepton families, the lepton and antilepton pairs in the decay products must be in the same family. The lepton family violating modes have the form

$$D \to X\mu^\pm e^\mp \tag{16}$$

where $D = D_s^+$ or D^+ and $X = K^+$ or π^+.

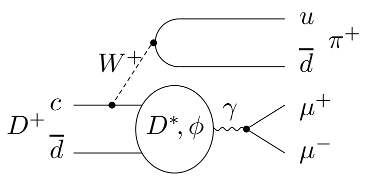
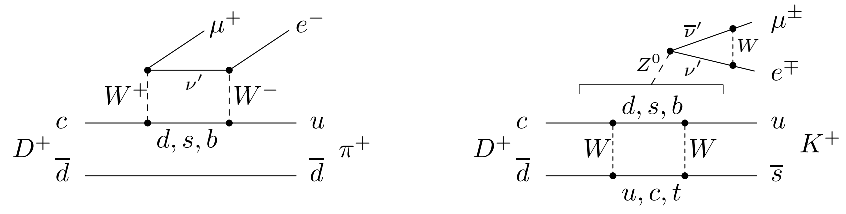


Figure 5: The Standard Model predicts branching ratios $< 10^{-6}$ for long range contributions to $D^+ \to \pi^+\mu^+\mu^-$



(a) $D^+ \to \pi^+\mu^+e^-$

(b) $D^+ \to K^+\mu^\pm e^\mp$

Figure 6: Standard Model Lepton Family Changing modes via a neutrino oscillation mechanism. These decays are allowed (but are *extremely* rare) if neutrinos mix (i.e. $\nu' = \nu_e, \nu_\mu, \nu_\tau$) in a fashion similar to quarks.

Recent neutrino oscillation observations indicate that neutrinos may oscillate between families. This implies that neutrinos have mass and that the weak eigenstates are mixtures of the mass eigenstates. If we add to the Standard Model a mixing matrix for leptons analogous to the CKM matrix, we can draw a lowest-order diagram for lepton family changing (LFC) decays (Figure 6). With this extension to the Standard Model, LFC decays are not strictly forbidden but are expected to be *extremely* rare. We can estimate [9] the branching ratio for the decay $D^+ \to \pi^+\mu^+e^-$ (shown in Figure 6) with

$$BR \sim g^4 \left[\frac{(m_{\nu_2}^2 - m_{\nu_1}^2)}{m_W^2} \cdot \frac{(m_b^2 - m_u^2)}{m_W^2} \right]^2 . \tag{17}$$

Using PDG limits on individual neutrino masses (Table 1) this branching ratio is $< 10^{-20}$. Using mass difference limits from neutrino oscillation experiments this branching ratio becomes $< 10^{-40}$.

1.4 **Beyond the Standard Model**

The Standard Model is a useful set of ideas that successfully describes the behavior of matter over a very wide range of energies. However, the model contains many unanswered questions and many arbitrary or incomplete theories. The quest to simplify or complete the Standard Model motivates the search for new behavior of matter that would violate the standard model. In response to some of its arbitrary features, many extensions to the Standard Model have been proposed that would allow decays not currently allowed. Among these are horizontal gauge bosons, leptoquark models, or supersymmetry. Many of these theories relate directly to measurements of rare and forbidden decays, which can constrain standard model extensions.

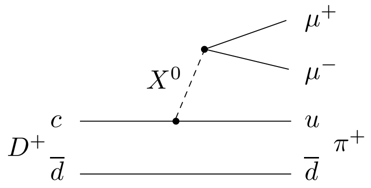
1.4.1 Horizontal Gauge Bosons

Weak interactions in the Standard Model occur "vertically" within a weak eigenstate generation. Generation changes result from the mixing of the mass and weak eigenstates, but still occur vertically. A direct transition "horizontally" across generations could result from a non-Standard Model neutral generation-changing gauge boson [10] as illustrated in Figure 7. The mass scale M_H of such a horizontal boson can be calculated by a comparison to the topologically similar semileptonic decay mediated by the W gauge boson. Assuming the coupling constants are similar, this comparison becomes, for the flavor-changing neutral current case,

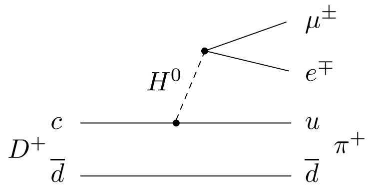
$$M_X \sim M_W \left[\frac{BR(D^+ \to \overline{K}^0 \mu^+ \nu)}{BR(D^+ \to FCNC)}\right]^{1/4} \tag{18}$$

or the lepton number violating case,

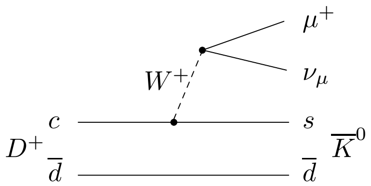
$$M_H \sim M_W \left[\frac{BR(D^+ \to \overline{K}^0 \mu^+ \nu)}{BR(D^+ \to LNV)}\right]^{1/4} . \tag{19}$$



(a) FCNC Horizontal Gauge Boson



(b) LFNV Horizontal Gauge Boson



(c) Semileptonic Decay

Figure 7: A comparison of Standard Model semileptonic decay vs non Standard Model LNV or FCNC decays can be used to establish an upper limit on the horizontal gauge boson masses.

For example, the upper limit on the branching ratios for the rare decay $D^+ \rightarrow \pi^+\mu^+e^-$ set by E687 [11] set a lower mass bound on the corresponding horizontal gauge boson, $m_H > 445GeV/c^2$. The improvement on this branching ratio by E791 [12] constrains this limit even further to $m_H > 565GeV/c^2$.

1.4.2 Leptoquark Extensions

The Standard Model allowed transitions from the quark to the lepton sector are mediated by the weak or electromagnetic gauge bosons. A direct coupling of a quark to a lepton would require a mediating particle that contains quantum numbers of both leptons and quarks [10]. A decay involving such a particle, termed a "leptoquark," is illustrated in Figure 8.

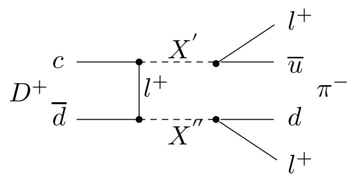


Figure 8: Leptoquark Diagram for LNV decays

1.4.3 Supersymmetry R-Parity Violation

Supersymmetry [13] (SUSY) generalizes the space-time symmetries of quantum field theory to allow the coupling of bosons and fermions. SUSY extensions to the Standard Model provide a theoretical framework for particle decays from the quark sector to the lepton sector mediated by the exchange of "squarks." The minimal supersymmetric extension to the Standard Model (MSSM) is made by taking the particles of the Standard Model and adding the corresponding supersymmetric partners. Burdman, et. al., [10] have considered supersymmetry parameters that may be measured with rare charm decays, specifically parameters that test the conservation of R-parity, defined as

$$R = (-1)^{3(B-L)+2S} \tag{20}$$

where B,L, and S represent the baryon number, lepton number, and spin of the respective fields.

There is no *a priori* justification for R-parity conservation. It is a useful assumption, however, since it avoids baryon and lepton number violating terms in the MSSM superpotential. Additional symmetries such as baryon-parity and lepton-parity would preserve baryon and lepton number while allowing R-parity violation. From the analysis of Burdman, et. al., the most general R-parity violating superpotential can be

Table 4: Constraints on R-parity violating MSSM [10]. FOCUS experimental sensitivities are able to constrain parameters in the R-parity violating MSSM. $\not{R}_p$ indicates maximum the branching ratio allowed using current constraints.

Decay Mode	S.M Limit	$\not{R}_p$	Exp. Limit.
$D^+ \to \pi^+ e^+ e^-$	2.0×10^{-6}	2.3×10^{-6}	5.2×10^{-5}
$D^+ \to \pi^+ \mu^+ \mu^-$	1.9×10^{-6}	1.5×10^{-5}	1.5×10^{-5}
$D^+ \to \pi^+ \mu^+ e^-$	0	3.0×10^{-5}	6.6×10^{-5}

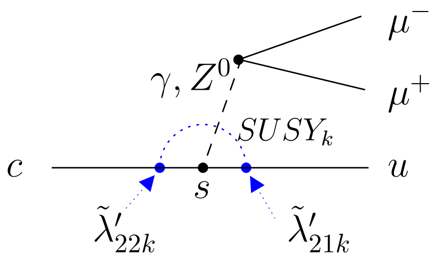


Figure 9: SUSY R-parity violating FCNC diagram [14]

written as

$$\mathcal{W}_{R_p} = \epsilon_{ab} \left\{ \frac{1}{2} \lambda_{ijk} L_i^a L_j^b \bar{E}_k + \lambda'_{ijk} L_i^a Q_j^b \bar{D}_k + \frac{1}{2} \epsilon_{\alpha\beta\gamma} \lambda''_{ijk} \bar{U}_i^\alpha \bar{D}_j^\beta \bar{D}_k^\gamma \right\} \tag{21}$$

where L,Q,$\bar{E}$,$\bar{U}$,and $\bar{D}$ are the chiral superfields. The $SU(3)$ color indices are denoted by $\alpha, \beta, \gamma = 1, 2, 3$, the $SU(2)_L$ indices by $a, b = 1, 2$, and the generation indices by $i, j, k = 1, 2, 3$. The term of interest to rare charm decays is the coupling parameter, λ'_{ijk}, which, through the exchange of squarks, can give rise to tree-level contributions to the decay modes $D \to X \ell^+ \ell^-$ or $D \to X \mu^+ e^-$. After rotating the fields into the mass basis, a new coupling, $\tilde{\lambda}'_{ijk}$, is defined

$$\tilde{\lambda}'_{ijk} \equiv \lambda'_{irs} \mathcal{U}_{rj}^L \mathcal{D}_{sk}^{*R} \tag{22}$$

where $\mathcal{U}^L$ and $\mathcal{D}^R$ are the matrices used to rotate the up and down quark fields to the mass basis. The current PDG limit $\mathcal{B}r^{exp}_{D^+\to\pi^+\mu^+\mu^-} < 1.5 \times 10^{-5}$ bounds the product

$$\tilde{\lambda}'_{22k}\ \tilde{\lambda}'_{21k} < 0.0004 \tag{23}$$

This upper bound will be reduce approximately as the square root of the improvement of $\mathcal{B}r^{exp}_{D^+\to\pi^+\mu^+\mu^-}$. FOCUS experimental sensitivities can improve upon this branching ratio by up to a factor of two, reducing the upper bound on these parameters:

$$\tilde{\lambda}'_{22k}\ \tilde{\lambda}'_{21k} < 0.0003 \tag{24}$$

Table 4 shows the branching ratio, $R\!\!\!/_p$, using current limits on R-parity violation parameters in MSSM.

1.4.4 Into the Future

Physics in the near future is presented with an exciting array of new questions and the potential for new physics which may change the Standard Model as we know it. The questions of neutrino mass and the mixing of mass and weak eigenstates in the Lepton sector may soon be well understood. The search for the origins of mass and the testing of the Higgs mechanism may soon open new horizons and shape much of future high-energy physics experimentation. The quest for simplifications to the Standard Model and a deeper understanding of its many parameters may lead to a discovery of supersymmetry, or perhaps something unexpected. Physicists still cannot satisfying answer the questions why are are there three generations of matter, what is the origin of mass, and what are the ultimate symmetries in matter? Probing the limits of the Standard Model will constrain these ideas further until eventually new physics will emerge. The prospects for new physics beyond the Standard Model in the coming decades are bright.

1.5 The FOCUS Experiment

At Fermilab (Fermi National Accelerator Laboratory, Batavia, Illinois), experiment E831 is conducting a broad research program into particle interactions involving charm [15]. E831 (also known as FOCUS) is a continuation of Fermilab experiment E687 and is the collected effort of more than 150 people at 20 universities in the United States, Italy, Korea, Brazil, and Mexico. Experiment E687, which collected data during several data runs from 1987 to 1992, represented the culmination of a long effort in the Fermilab fixed target program to produce charm using a photon-gluon fusion process as shown in Figure 10. E687 established many important measurements on charm decays. FOCUS provided a substantial upgrade to E687 yielding more than a factor of 10 increase in statistics and what is currently the world's large sample of charm decays. FOCUS collected data during a 1996-1997 data run and is currently publishing many new results on charm physics.

1.6 Scope of the Thesis

Since the FOCUS charm data represents the world's largest sample of charm decays, FOCUS has the best sensitivity to date on several rare and forbidden charm decays. In this thesis I search for three-body rare and forbidden charm decays and set limits on these decays. The detector is summarized as well as a general overview of the reconstruction algorithms. Special attention is paid to the calibration of Monte Carlo muon efficiencies and muon misidentification which significantly impact this

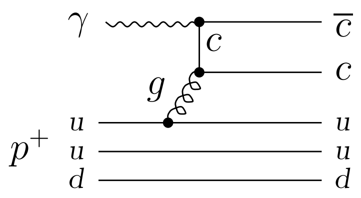


Figure 10: Charm Photoproduction Diagram

analysis. Novel methods for optimizing small signals and reducing bias comprise a major aspect of this analysis and are explained in detail. The 90% confidence upper limit results are presented in a way that shows the value of these results in relation to our expected sensitivities and in relation to previous measurements.

CHAPTER II

THE FERMILAB FOCUS EXPERIMENT

This chapter provides a general description of the FOCUS experiment with an emphasis on the FOCUS detector elements which are important to the analysis in this thesis. Photoproduction of charm requires converting the Fermilab Tevatron proton beam into a photon beam with the energies necessary for the production of charm with a fixed target. The multi-step process used in the production of the proton beam, its conversion into a photon beam, the production of charm particles, and the detection of charm decays is outlined in the following sections.

The FOCUS spectrometer [16] was built to detect charm particles and their decays. The spectrometer is an upgrade to the E687 detector [17] and is shown in Figure 14. This spectrometer features several silicon microstrip detectors and proportional wire chambers for vertexing and tracking, two analysis magnets for momentum measurement, three Čerenkov detectors, two muon detectors, two electromagnetic calorimeters, a hadronic calorimeter, several trigger hodoscopes, and calorimeters for beam tagging.

The FOCUS experiment produced the world's largest sample of fully reconstructed charm decays with low electromagnetic backgrounds, high vertexing resolution, and excellent particle identification.

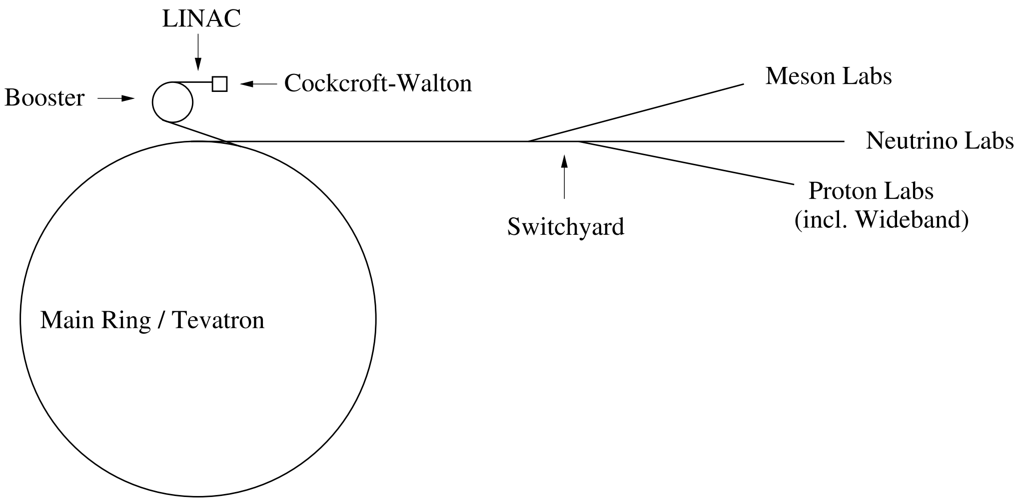


Figure 11: Layout of the beam acceleration stages

2.1 Tevatron Proton Beam

Fermilab uses a multi-step process to deliver 800 GeV/c protons to the Wideband laboratory as illustrated in Figure 11.

Cockcroft-Walton: Initially, a puff of hydrogen gas is negatively ionized. The H^- ions are electrostatically accelerated in a 750,000 V Cockcroft-Walton voltage gap and then injected into the LINAC (with an energy of 750 keV).

LINAC: The LINAC is a 500 ft. long linear accelerator, in which alternating electromagnetic fields (in radio-frequency cavities) create accelerating electromagnetic wavefronts. The H^- ions are accelerated by the wavefronts to an energy of 400 MeV. The ions then pass through a thin carbon foil which strips off the electrons. The H^+ ions (protons) are injected into the Booster.

2.1.1 Synchrotrons

In linear accelerators, a particle is accelerated with only one pass through the accelerator. In a synchrotron, particles are cycled through RF cavities thousands of times and can be accelerated to very high energies. A synchrotron is composed of alternating stations of beam accelerating RF cavities and beam shaping magnets. Dipole magnets are used to bend the proton beam along a circular path. Energy is lost as "synchrotron radiation" when particles are bent in a magnetic field, however this loss is relatively small for protons. Quadrapole (and higher order) magnets are used to focus the beam and reduce beam aberrations.

Booster: The Booster is a 500 ft. diameter rapid cycling synchrotron. The protons travel through the Booster about 12,000 times and are accelerated to 8 GeV/c. The resulting bunch, or pulse, of protons is then injected into the Main Ring. Normally a rapid sequence of twelve proton pulses is used to fill the Main Ring.

Main Ring: Two synchrotrons, the "Main Ring" and the "Tevatron," occupy the largest Fermilab synchrotron tunnel, about four miles in circumference. The Main Ring (the original synchrotron at Fermilab – now decommissioned) used conventional copper-coiled steel dipole magnets to accelerate protons to energies of 150 GeV. The protons were then injected into the Tevatron.

Tevatron: The Tevatron uses liquid helium cooled, superconducting, dipole magnets to contain the proton beam more efficiently than conventional magnets. This allows the Tevatron to reach near TeV energies. The Tevatron operates with a repeating cycle of beam acceleration and beam extraction. The beam is accelerated for approximately 40 seconds followed by a 20 second “spill” (beam extraction) of 800 GeV/c protons. The beam extraction is 20 s long (when it could be much faster) to deliver a long-duration beam of uniform intensity to the fixed-target laboratories. The RF cavities that accelerate the beam operate at 53 MHz which corresponds to 18 ns spaced bunches or “buckets” of protons. The proton beam is then sent to the fixed-target laboratories.

Fixed-Target Laboratories: The fixed-target laboratories are located about a mile from the Tevatron and are grouped in three areas known as “Meson,” “Neutrino,” and “Proton.” A switchyard splits the proton beam into several less intense beams and sends them to the separate laboratories so that multiple experiments can run simultaneously. FOCUS occupied the Wideband laboratory in the “proton” section. Approximately 4.5×10^{12} protons/spill were delivered to Wideband.

After 1999, several configuration changes were made at Fermilab. The Main Ring was decommissioned and replaced with the Main Injector, a new synchrotron (complete with its own tunnel) which now feeds the Tevatron. In addition, the fixed-target program was retired bringing an end to the series of experiments at Wideband that culminated in FOCUS.

2.2 Wideband Photon Beam

At the Wideband laboratory (named after its wide-band photon beam), a multistep process is used to convert the 800 GeV/c protons from the Tevatron into a usable photon beam [18], with an average energy of about 200 GeV, as illustrated in Figure 12. A hadronic beam incident on the experiment target could have been used

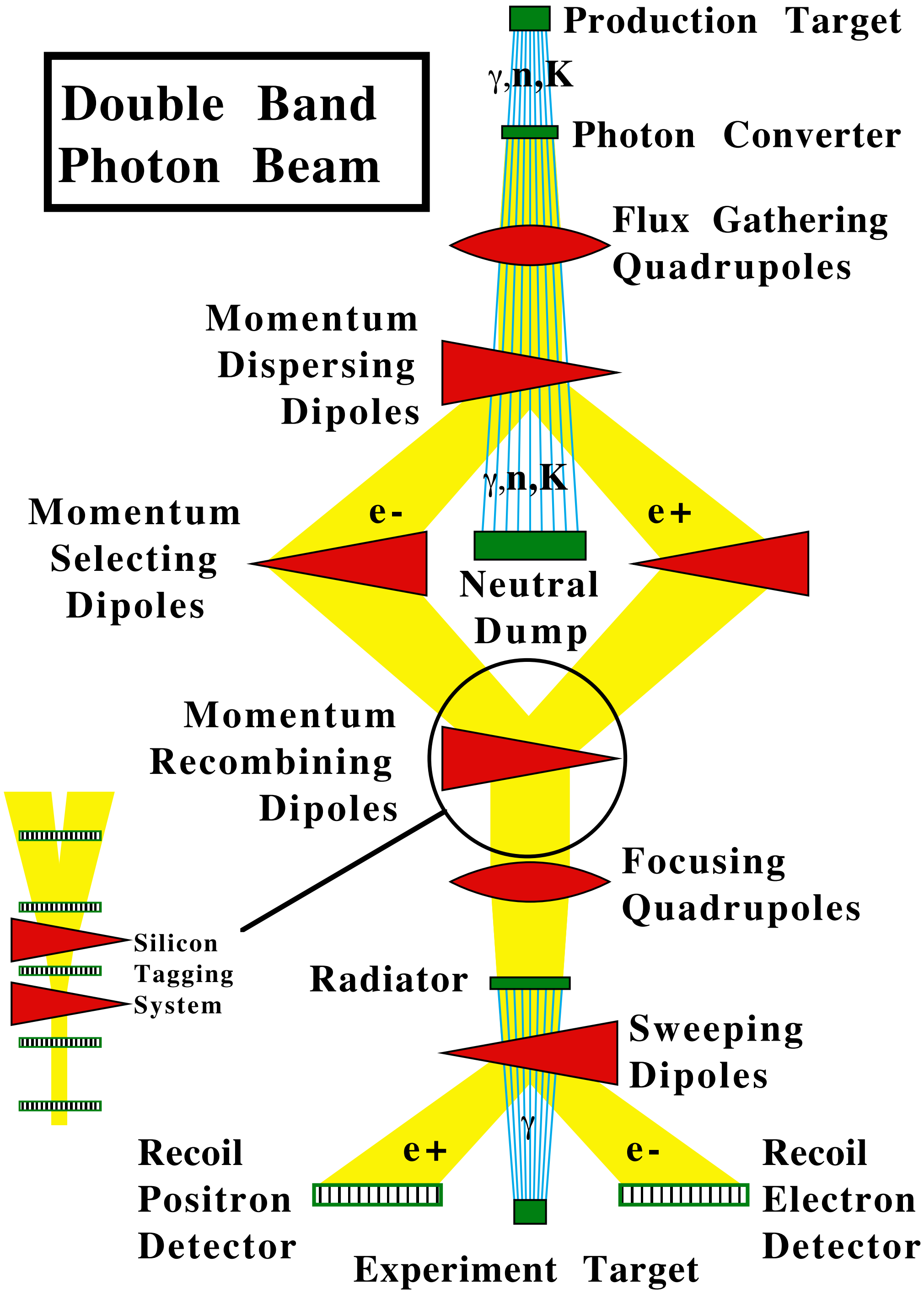


Figure 12: Schematic of the FOCUS Proton to Photon Conversion

to produce charm (as done in E791), however there are advantages to using a photon beam. A high-energy photon beam produces charm with less hadronic contamination but with higher electromagnetic contamination. With triggering it is easier to separate hadronic from electromagnetic events than it is to separate charm from other hadronic events. So even though the conversion to a photon beam reduces the intensity of experimental target interactions, the charm statistics are improved by a cleaner charm sample.

Production Target: The 800 GeV/c protons from the Tevatron interact with the Production Target to produce hadrons, including π^0's which quickly decay ($\tau \sim 10^{-16}$ s) into two photons. The Production Target is a tank of cryogenically cooled liquid deuterium 1.6 m in length. The high A/Z^2 ratio of liquid deuterium provides a large cross section for hadronic interactions, proportional to the number of nucleons, A, while providing low photon absorption, which scales as the square of nuclear charge, Z.

Electron/Positron Beam: As illustrated in Figure 12, dipole magnets sweep away charged particles produced in the deuterium target. The photons (along with other neutral particles) then strike the photon converter, a 50-60% radiation length lead sheet. Electron/positron pairs are created in the converter while neutral hadrons pass through. The e^+e^- pairs are focused into a beam with "flux gathering" quadrapole magnets and then directed around the neutral beam dump with dipole magnets. The beam dump absorbs the remaining neutral hadrons which would produce higher backgrounds if allowed to interact with the experimental target.

The separate electron and positron beams (the "double bands") are recombined and focused on the experimental target. The positron beam was an addition over E687 designed to double the target interaction rate. However, size constraints in the beam tunnel created an occlusion to the positron beam resulting in a 30% loss of the positron beam relative to the electron beam. The electron beam had negligible hadronic contamination but the positron beam adds a small hadronic contamination (due to Λ decays after the neutral sweeper) so that 3.8% of the triggered events are produced by hadrons in the beam. The beam is also contaminated with a small number of muons from the production target. These muons are described in detail in Section 4.2.1. Beam collimators select a nominal lepton beam momentum of 300 GeV/c and accept a momentum bite of $\pm 15\%$, a relatively wide band.

Photon Beam: Bremstrahlung photons are produced by scattering the electrons and positrons with the radiator, composed of a 20% radiation length lead sheet. The lepton beam is focused on the experiment target so that the bremstrahlung photons are focused on the experiment target as well. The electrons and positrons are then swept toward the recoil electron/positron calorimeters with dipole magnets. What remains incident on the experiment target are photons. The mean energy of the photons which produce reconstructible charm is about 190 GeV. The incoming 4.5×10^{12} protons/spill result in 10^{9} photons/spill.

2.2.1 Photon Beam Tagging System

A beam tagging system [19] is used in analyses that require information on the energy of the incident photon for a particular event. Beam tagging is not used in this

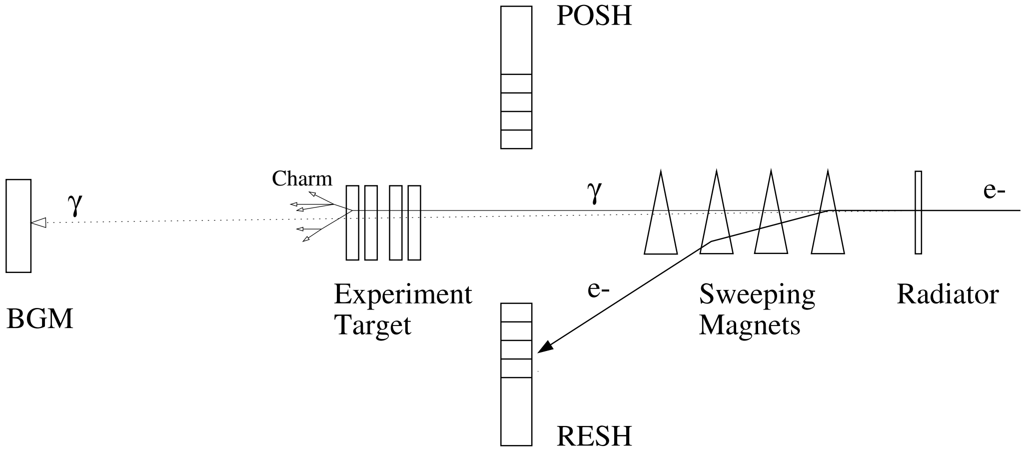


Figure 13: Diagram of the Beam Tagging System

analysis but a description is included for completeness. The energy of the photon created in the radiator is the difference between the energy of the incident electron, E_i, and the energy of the electron after the radiator, E_r, plus the energy of any additional photons that do not interact with the experimental target, E_{non}.

$$E_\gamma = E_i - E_r - E_{\text{non}} \tag{25}$$

Silicon Tagging System The Silicon Tagging System determines E_i, the incident electron energy, by measuring the horizontal bend of the electron as it passes through the two momentum recombining dipole magnets. The system consists of five planes of silicon strip detectors positioned before, between, and after the two magnets as shown in Figure 12. This system is the same as in E687.

RESH and POSH Calorimeters The Recoil Electron/Positron Shower Calorimeters measure E_r, the energy of electrons or positrons after passing through the radiator. These sampling calorimeters are segmented horizontally to provide a measure of the horizontal position (and thus the bend) of the electron in addition to the energy from the electron shower in the calorimeter.

BGM Calorimeter The Beam Gamma Monitor Calorimeter measures the energy of photons that do not interact hadronically with the experimental target, E_{non}. The BGM Calorimeter or Beam Calorimeter is located past the middle of the spectrometer, just downstream of the Inner Electromagnetic Calorimeter. The RESH/POSH Calorimeters consist of alternating layers of lead and Lucite. The BGM and innermost segment of RESH/POSH use Si0_2 instead of Lucite for radiation hardness.

2.3 Experiment Target

The high energy photon beam interacts with the experiment target to produce charm through a photon-gluon fusion process. Other hadrons, primarily pions, are produced with the charm, and these extra particles are seen to come from a common point. This point locates where the charm is produced and is called the primary vertex. The charm particles will decay quickly into their daughter particles. These daughter particles will be seen to come from a common point downstream of the primary vertex. The location of the charm decay is called the secondary, or decay, vertex.

Non-charm particles from the primary vertex can re-interact with the target, producing a vertex of particles. The vertex from secondary interactions can be confused

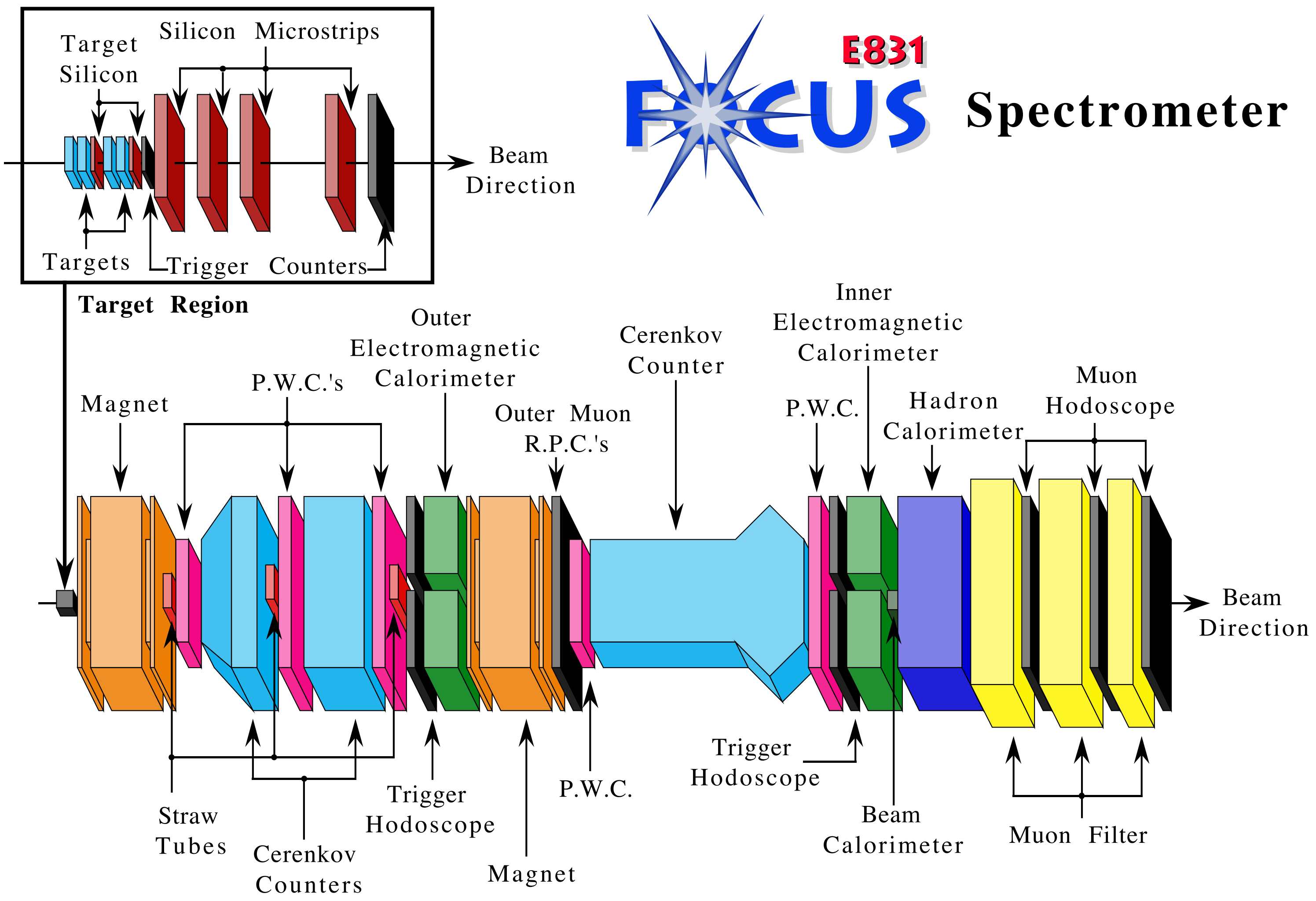
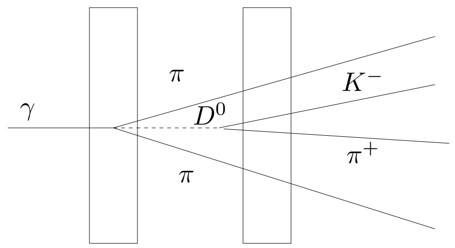
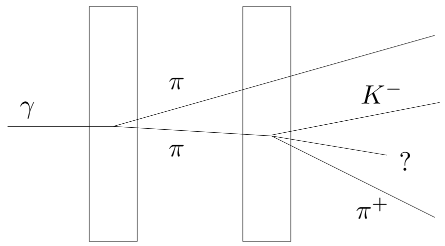


Figure 14: Schematic of the FOCUS Spectrometer



(a) Out-of-Target Charm Decay



(b) Secondary Interaction

Figure 15: Advantage of a segmented target. The use of a segmented target increases the number of decays that can occur outside of the target.

with the vertex from charm decays and will produce a background to the charm signal. The experiment target for FOCUS was originally the single beryllium target used in E687. Early in the FOCUS data run, the target was replaced with a segmented target to increase the number of charm particles that decay outside of the target material. This reduces the chance of confusing a secondary target interaction with the charm decays and makes the backgrounds easier to model.

The first segmented target was made from beryllium but this was changed to beryllium oxide for most of the FOCUS data run. The high A/Z^2 ratio of beryllium oxide maximizes strong interactions for charm photoproduction while minimizing the production of conversion pairs, and minimizing multiple scattering. The high density of beryllium oxide allows the target to be thin, increasing out-of-target decays. Each of the four targets is 30 mm square. by 6.75 mm thick, which corresponds to a 15% hadronic interaction length and a 5% electromagnetic radiation length.

2.4 **Tracking**

Particles are tracked in FOCUS using a silicion microstrip detector (SSD), located immediately downstream of the target, and proportional wire chambers (PWC), located between the first analysis magnet (M1) and the inner electromagnetic calorimeter (IE). The SSD is a high-precision tracking system used to resolve particles that are very closely spaced since they have just left the experiment target. By the time the particles reach the PWC's, they have separated significantly so a lower spatial resolution is required for the PWC's. PWC's are widely separated to provide accurate slopes for momentum determination. SSD's are high-precision and near the target to provide good vertexing resolution for lifetime and vertex separation measurements.

Between the SSD's and the first PWC chamber (P0), charged tracks are bent by the magnetic field of M1. Between the third and fourth PWC chamber (P2 and P3), charged tracks are bent by the magnetic field of the second analysis magnet (M2). The reconstruction algorithms connect together (or "link") SSD and PWC tracks and calculate the momentum from the track bends induced by the analysis magnets.

2.4.1 Silicon Microstrip Detector

Charged particles passing through silicon semiconductor strips induce a current in the semiconductor. Silicon microstrips are located both immediately downstream of the target and embedded within the target segments [20]. The precision of the SSD's

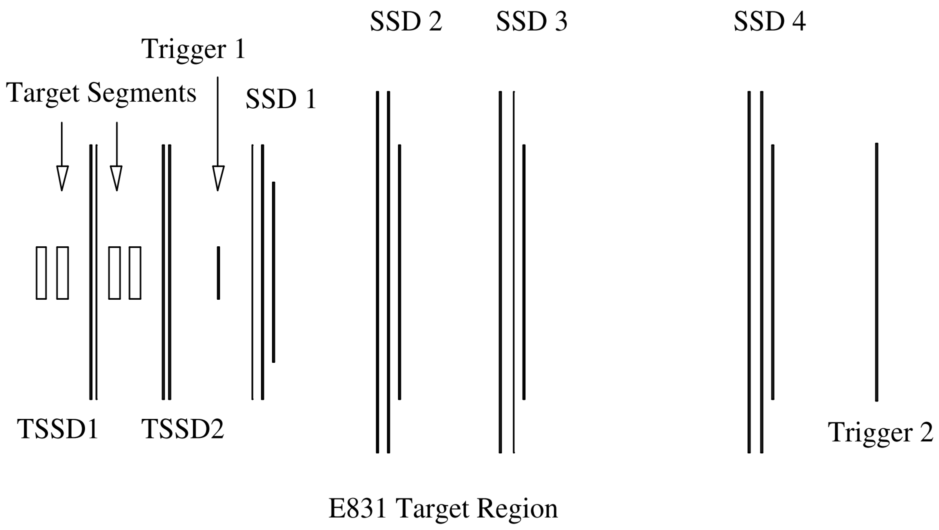


Figure 16: Relative Positions of the Target, Target Silicon, and SSD's.

provides excellent vertexing and tracking information with an asymptotic lifetime resolution of 12 fs for two-body charm meson decays.

The first two microstrip stations (TSSD's) are embedded in the target. The TSSD's are configured into two stations ±45° from horizontal. Each view consists of 1024 strips 25 μm wide and 50 mm long making a total area $25 \times 50\,\mathrm{mm}^2$ square. The TSSD's were installed in 1997 and were operational for about two-thirds of the FOCUS data collected.

Downstream of the first trigger counter (TR1) are four more microstrip stations (SSD's). Each station has three views -135°,-45°, and -90°, relative to the horizontal, which are separated by 5 mm. The first three stations are separated by 6 cm, and the last separated by 12 cm. Each view has two regions with different strip densities. The first SSD station has an inner strip spacing of 25 μm and an outer strip spacing of 50 μm. The last three SSD stations have 50 μm and 100 μm strip densities.

2.4.2 Proportional Wire Chambers

Proportional Wire Chambers (PWC's) are built from alternating planes of high voltage wires and grounded sense wires. The chambers are filled with a gas selected for its ionizing properties. The planes are up to $60 \times 90\,\mathrm{in}^2$ and are separated by a gap of only a few millimeters. The voltage differences between sense and high voltage wires is roughly 3 kV. Electrons from atoms in the PWC gas are liberated by charged particles passing through the chamber. These electrons are accelerated toward the grounded sense wires. The electrons ionize the gas further, producing a cascade of electrons that reach the sense wires and ions that move away from the sense wires. This current is amplified and, if it is above a certain threshold, produces a hit.

The FOCUS spectrometer has five PWC chambers (labelled P0-P4) with four planes of sense wires in each chamber. As illustrated in Figure 17, each chamber has views for x (vertical wires) and y (horizontal wires) plus u and v wires at $\pm 11.3^\circ$ from horizontal. The stereo (angled) views are used to resolve ambiguities. 11.3° was used instead of 45° to provide more information in the vertical (bend) view, since the analysis magnets both bend in the vertical direction. Outer angle tracks that miss P3 and P4 are called 3-chamber tracks or "stubs." Inner angle tracks that leave hits in all PWC's are referred to as 5-chamber tracks.

All chambers use a gas mix of 75% argon and 25% ethane bubbled through methanol to remove water contamination. There are two types of chambers: Type I chambers (P0 and P3) are positioned just downstream of the analysis magnets and are sized to match the magnet apertures. These chambers have shorter wires and

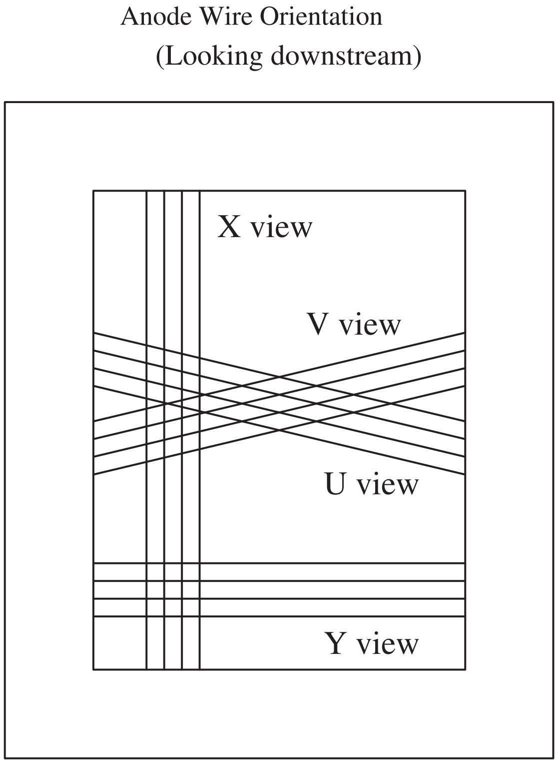


Figure 17: PWC Views

smaller diameter gold plated wires so higher voltages can be applied. Type II chambers (P1, P2, and P4) are larger chambers with looser wire spacing, copper wires and lower voltages.

2.4.3 Straw Tube Chambers

The analysis magnets bend in the vertical direction forming a vertical stripe with a high e^+e^- flux. To prepare for the possibility that the conventional PWC's might need to be deadened in the high flux pair region, three straw tube wire chambers were installed to provide additional tracking in this region. The PWC's ultimately did not need to be deadened, and this system was not used.

2.5 Analysis Magnets

Two analysis magnets are used to determine the momentum of charged particles. The first magnet, M1, is located just downstream of the SSD's. The second magnet, M2, is located near the middle of the spectrometer. The magnets deflect vertically and with opposite polarities so that e^+e^- pairs produced in the experiment target are roughly focused back onto the Beam Calorimeter (BGM). M1 operates at 1020 amps and provides a p_T kick of 0.400 GeV/c. M2 operates at 2000 amps and provides a p_T kick of 0.836 GeV/c.

2.6 Cerenkov Counters

Čerenkov light is emitted when particles passing through a material travel faster than the speed of light in that material, that is,

$$\beta = \frac{p}{E} = \frac{p}{\sqrt{p^2 + m^2}} \geq \frac{1}{n} \tag{26}$$

where n is the index of refraction of the material (using units where $c \equiv 1$). Solving for p at the threshold yields

$$p_{\text{thresh}} = \frac{m}{\sqrt{n^2 - 1}} . \tag{27}$$

For our threshold Čerenkov detectors, the spatial resolution of the detector is low compared to a ring imaging detector, so that the light from a particular Čerenkov light cone is detected by one or only a few photodetectors. This resolution is sufficient to link separate Čerenkov signals to individual PWC tracks. The intensity of the Čerenkov light for a particle (which will only emit Čerenkov light if it is above threshold) along with momentum measurements for that particle are used to calculate the probability of a hypothesis for the particle type. The FOCUS spectrometer has

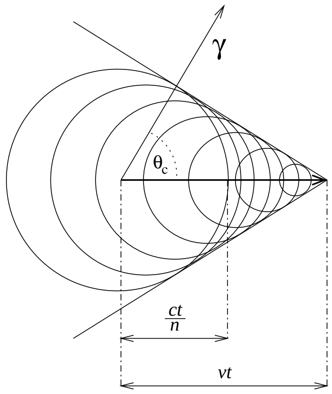


Figure 18: Čerenkov light. As a charged particle travels through a medium faster than the speed of light for that medium, electromagnetic shock waves produce Čerenkov light. The direction of the Čerenkov photons is perpendicular to the electromagnetic shock wave. The Čerenkov cone is not the cone of the shock wave, but rather the cone of light defined by the Čerenkov angle at the point the particle enters the Čerenkov detector. Figure from Reference 21.

three threshold Čerenkov detectors, C1, C2, and C3, to distinguish electrons, pions, kaons, and protons [22].

C1: Čerenkov detector C1 is located between proportional wire chambers P0 and P1 with xy dimensions[1] of $50 \times 80\,\mathrm{in}^2$. C1 consists of 90 cells, split between inner and outer portions. The outer portion has spherical mirrors which focus the Čerenkov light onto 40 photomultiplier tubes (PMT's). The inner portion uses planar mirrors that reflect the Čerenkov light onto Winston cones which concentrate the light onto 50 PMT's. C1 uses a mixture of 58% helium and 42% nitrogen and has a pion threshold of 8.5 GeV/c.

C2: Čerenkov detector C2 is located between P1 and P2 with xy dimensions of $64 \times 100\,\mathrm{in}^2$. C2 consists of 110 cells, split between inner and outer portions with a

[1] all dimensions are ordered $x \times y$ unless otherwise indicated.

Table 5: Properties of the three Čerenkov Detectors.

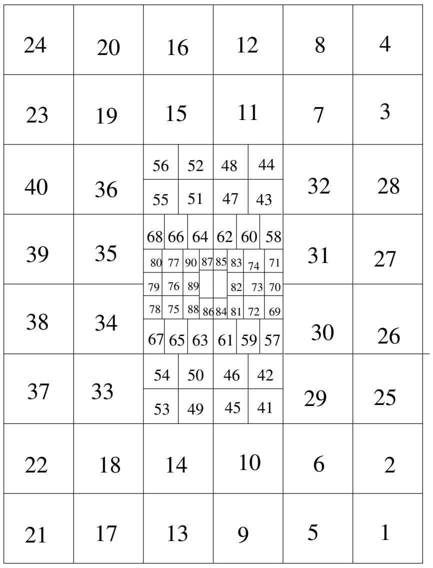
Detector	Gas	Threshold (GeV/c) π	K	p	No. of Cells
C1	He/N_2	8.5	29.9	56.8	90
C2	N_2O	4.5	16.2	30.9	110
C3	He	17.0	61.0	116.2	100

design similar to C1. C2 uses nitrous oxide gas (N_2O) and has a pion threshold of 4.5 GeV/c.

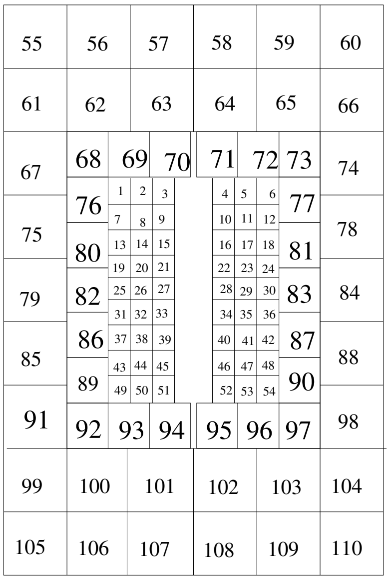
C3: Čerenkov detector C3 is located between P3 and P4 with xy dimensions of $60 \times 93\,\text{in}^2$. C3 consists of 100 cells which all use spherical mirrors to focus onto PMT's. C3 uses helium gas and has a pion threshold of 17.0 GeV/c.

2.7 Calorimeters

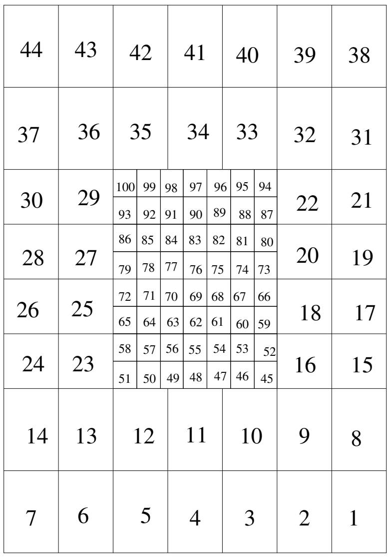
Calorimeters determine the energy of a particle by measuring the particle showers produced in the calorimeter. Sampling calorimeters are made with alternating layers of absorber and scintillator. The incoming particles interact with the absorbing material (lead or steel) producing more particles, which in turn interact with the next layers of absorbing material, resulting in a shower of particles. In a lead glass calorimeter, Čerenkov light from the showering particles is measured. The lead enhances the electromagnetic interactions and the glass transmits the light. In other calorimeters, plastic scintillator detects the showers. For electromagnetic showers, bremstrahlung and pair conversion produce the showers, for hadronic showers, strong interactions produce (initially) the showers. The amount of light produced by the



(a) C1. Outer cells 1–40 use spherical mirrors. Inner cells 41–90 use planar mirrors.



(b) C2. Outer cells 55–110 use spherical mirrors. Inner Cells 1–54 use planar mirrors.



(c) C3. All cells use spherical mirrors.

Figure 19: The arrangement of the light gathering cells for the three Čerenkov counters, C1, C2, and C3.

showers indicates the number of particles produced in the shower and is proportional to the incident particle energy. Since calorimeters can detect strongly interacting particles, whether charged or uncharged, calorimetry is a useful way to detect uncharged hadrons (as well as photons).

2.7.1 Inner Electron Calorimeter

The inner electromagnetic calorimeter (IE) [23] was built from 802 lead glass blocks each with a face $5.8 \times 5.8\,\mathrm{cm}^2$ and a depth of 60.2 cm, equivalent to 18.75 radiation lengths or 2.2 proton interaction lengths. Each block is wrapped with aluminized mylar to reflect light back into the block and arranged as shown in Figure 20. The light is detected with a photomultiplier tube at the back of each block. Summing modules extracted a small portion of the signal from each block to create the IE triggers, which included: sums of the total energy E_{IE} (used in the second level J/ψ trigger) and $E_{IE\cdot 2}$ (used in the second level hadronic trigger); the sum of the total transverse energy $E_{T(IE)}$; and a two-body trigger, IE_2 (used in the J/ψ level-one trigger).

2.7.2 Outer Electron Calorimeter

The Outer Electron Calorimeter (OE) [24] is a sampling calorimeter which measures the energies of outer angle electrons and photons. The OE is positioned before M2 with an opening of $55 \times 88\,\mathrm{cm}^2$ coinciding with the M2 aperture and an overall size of $255 \times 205\,\mathrm{cm}^2$ with a vertical gap to avoid conversion pairs. The OE is built from alternating layers of lead and scintillator. There are 23 layers of 3.1 in. wide scintillator paddles in x, y, u, or v orientations. The u and v planes are oriented $\pm 45°$

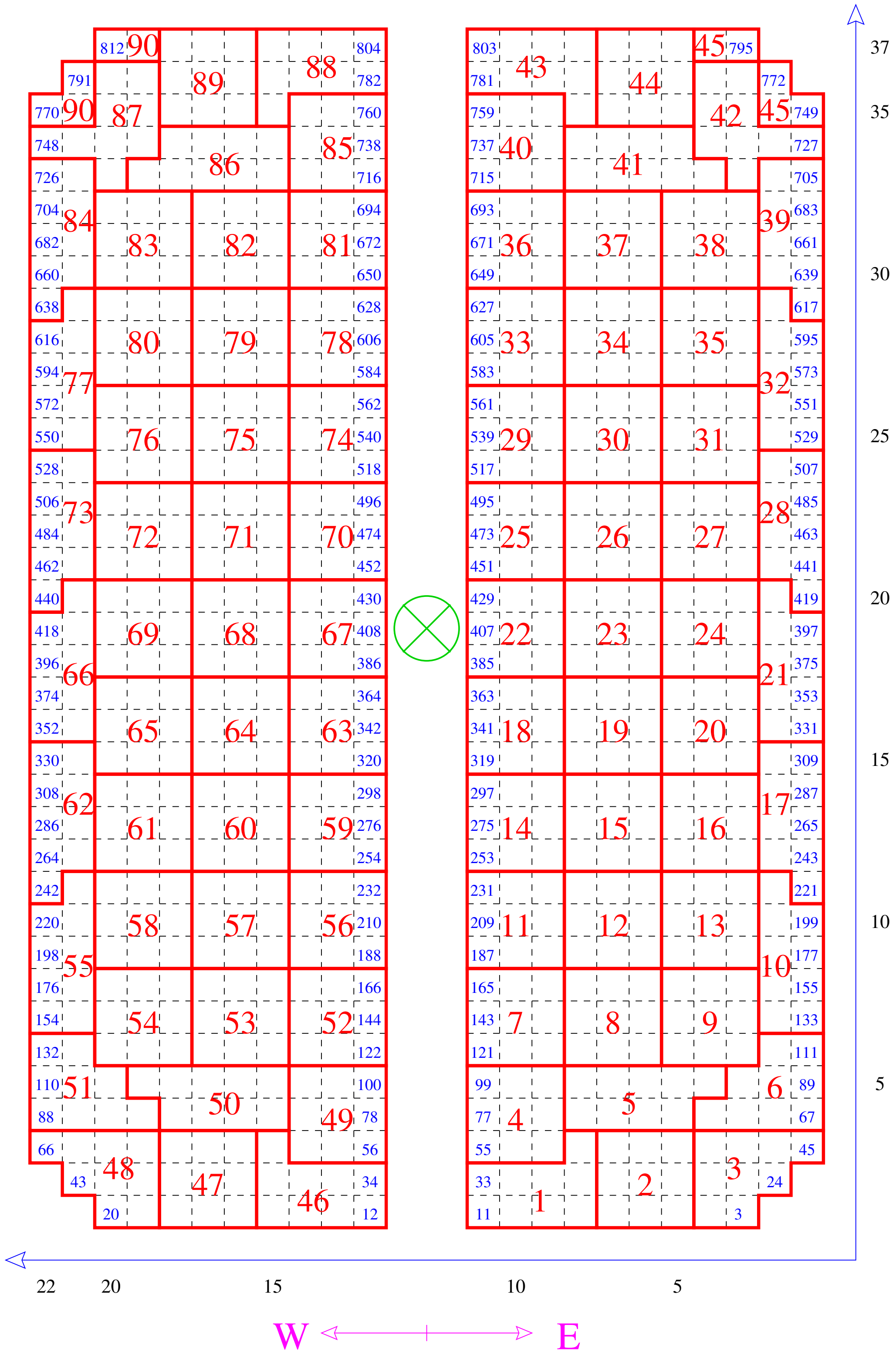
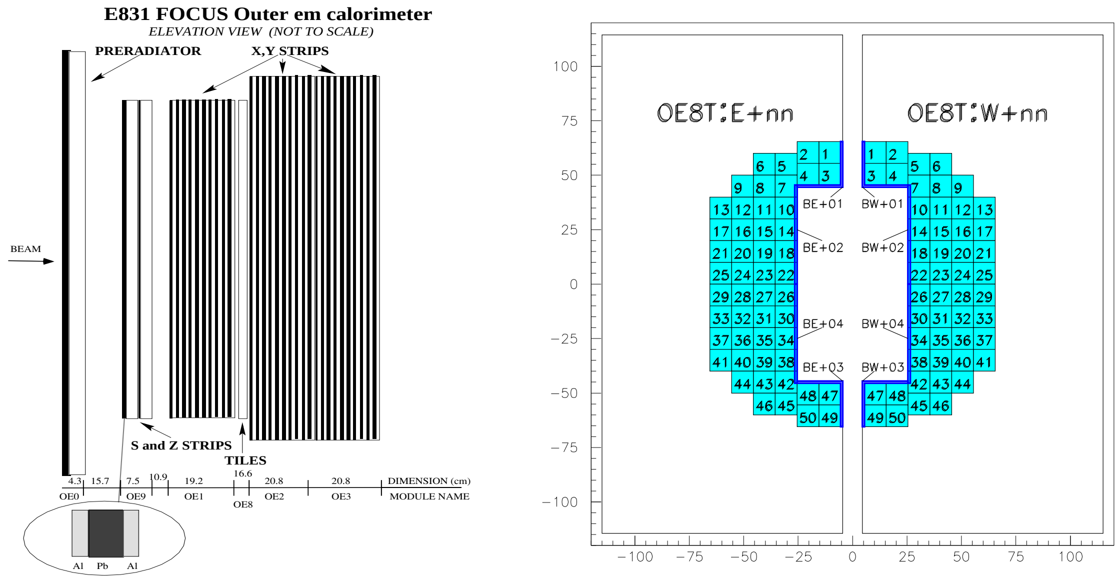


Figure 20: Inner Electromagnetic Calorimeter Layout.



(a) OE Side View

(b) OE Tie Breaker

Figure 21: OE Layout

from horizontal. A single layer of 100 scintillator tiles, located in the higher rate inner region of the detector, is used as a tiebreaker for two simultaneous hits. Side and front views of the OE are shown in Figure 21. OE information is not used in this analysis, but is described for completeness.

2.7.3 Hadron Calorimeter

The Hadron Calorimeter (HC) [25] [26] is located downstream of the IE. It is a sampling calorimeter built with 28 alternating layers of scintillator and steel. The face of the HC measures $200 \times 300\,\mathrm{cm}^2$. The HC is 209 cm deep, corresponding to 7.8 hadronic interaction lengths. Each layer is composed of 4.4 cm thick steel and scintillator which is subdivided into 66 tiles (or pads) as illustrated in Figure 22. The inner region is built with 20×20 cm tiles, the outer region with 40×40 cm and

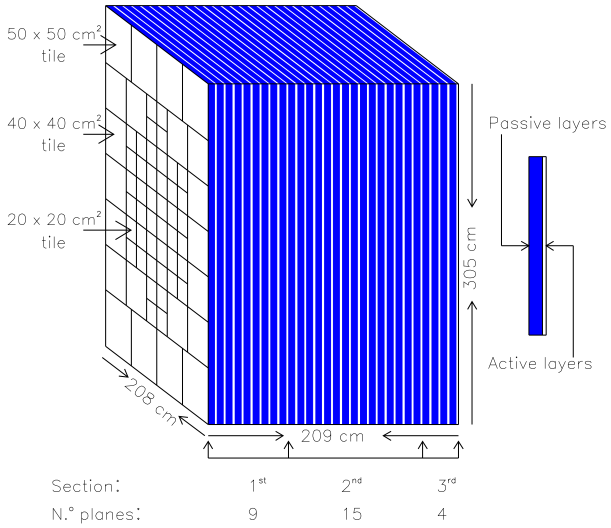


Figure 22: HC Layout

$50 \times 50\,\text{cm}$ tiles. The layers are grouped into three sections in which corresponding tiles from each layer are optically combined (with optical fiber). This reduces the readout to 66 channels for each section; a total of 198 channels for the entire HC. The first 9 layers form the first section, the next 15 layers form the second section and the last 4 layers form the final section. The HC signals are summed to form two HC triggers, at a higher threshold (about 20 GeV) E_{HI}, used in the main level-one trigger, and a lower threshold, E_{LO}, which was not used.

2.8 Muon Detectors

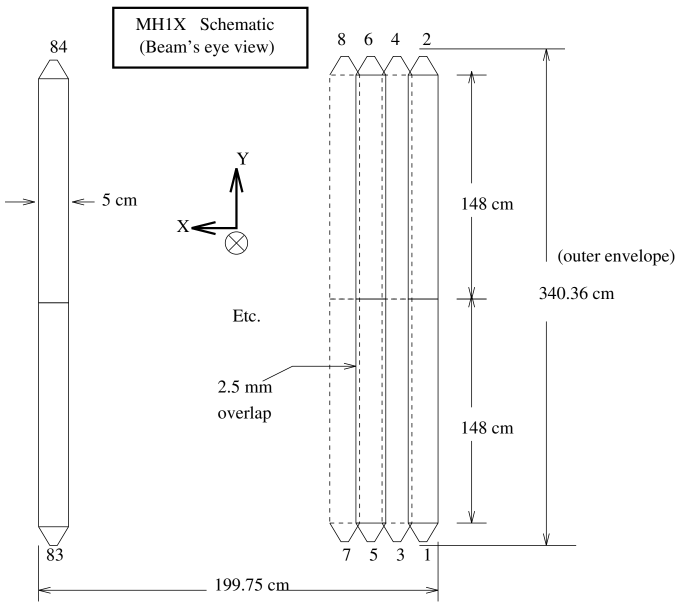
Since muons can penetrate much more material than electrons or hadrons, muon detection is achieved by looking for the passage of charged particles through a length of absorbing material large enough to absorb the electrons and hadrons. The FOCUS spectrometer has two muon detection systems. Outer angle muons are detected in the outer muon detector located near the center of the spectrometer just downstream of M2. The Inner Muon Detector is located at the end of the spectrometer, downstream of the Hadron Calorimeter.

2.8.1 Inner Muon Detector

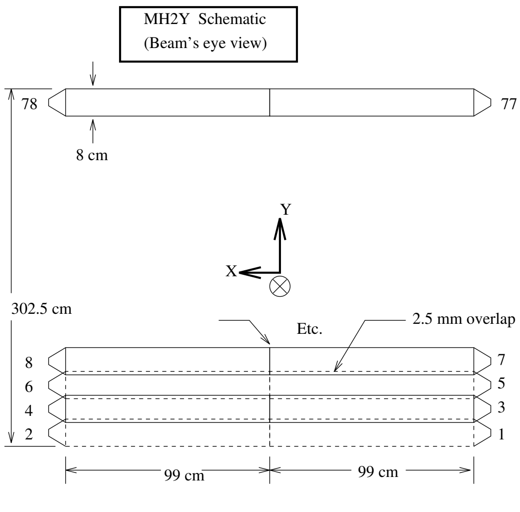
The inner muon system detects muons with arrays of scintillator, also called muon hodoscopes [27]. The muon hodoscopes (MH) are arranged in three stations named MH1, MH2, and MH3. The MH stations are separated by steel muon filters as shown in Figure 14. The three steel filters are 61 cm, 129 cm, and 68 cm thick. The HC, just upstream of the inner muon system, provides an additional 126 cm of steel. Each MH station has two views. MH1 and MH2 have x and y views, while MH3 has u and v views at $\pm 30^\circ$ from horizontal. As muons pass through more iron, the extrapolation error from the tracking chambers increases (due to multiple Coulomb scattering). Therefore, the scintillator strip widths are increased for stations further downstream. The strip width is 5 cm for MH1, 8 cm for MH2, and 10 cm for MH3. Inner Muon triggering is accomplished with a separate set of hodoscopes described in Section 2.9.1.

2.8.2 Outer Muon Detector

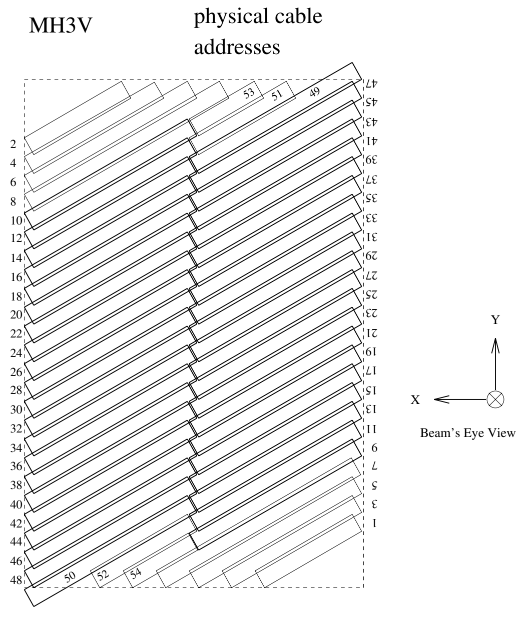
The Outer Muon Detector [28] is located just downstream of M2. It is positioned to detect outer angle charged particles that have passed through the material of M2 and the OE (about 4 m of iron). Since the outer muon system is positioned close to M2, in magnetic fields up to 1 kgauss, and in a confined space, resistive plate chambers



(a) Layout of MH1X and MH2X.



(b) Layout of MH2Y and MH1Y.



(c) MH3V. MH3U is the mirror of MH3V.

Figure 23: Views showing the counter arrangements of the Inner Muon Hodoscope arrays.

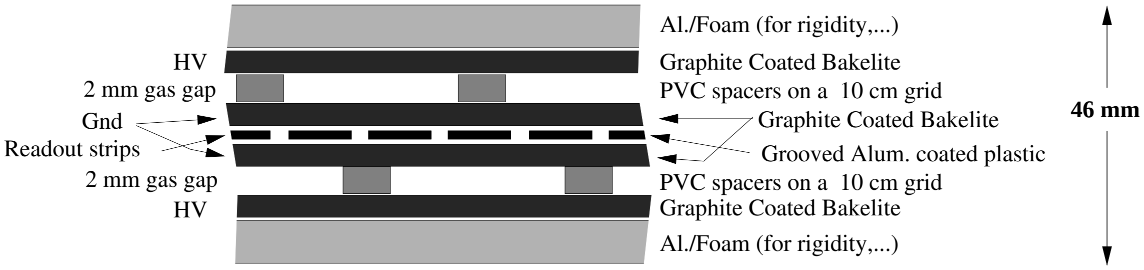


Figure 24: RPC module cross section

(RPC's) were used. RPC's work in higher magnetic fields than other systems and have high efficiency without several layers of detector.

The RPC's are double gap modules as shown in Figure 24. High voltage (5.8 kV) is applied to graphite coated bakelite across a gap filled with a gas mixture of 5% freon, 8% isobutane, 16% CO_2, and 71% argon. The gap is maintained with spacers that result in dead regions. Using a double gap with the spacers staggered on adjacent gaps provides complete coverage and some redundancy against gap failure.

The 24 RPC modules were assembled into three views with eight RPC modules in each view. Each module is $1 \times 1.6\,\mathrm{m}^2$ or $1 \times 1.8\,\mathrm{m}^2$. Readout strips are located between the bakelite modules. Each strip is 2.9 cm wide with a 2 mm gap between strips. The x view plane has two sets of vertical strips that each cover half of the RPC module. The y view has one set of horizontal strips that cover the full width of the module. The u view has one set of strips at 45° that cover the full module. Every four adjacent readout strips were OR'd together creating an effective strip width of 3.1 cm. The outer muon system efficiencies are discussed in Section 4.1.

The modules from the three views that are aligned in the z direction define a "tower." A muon candidate is defined by two of the three modules in a tower having

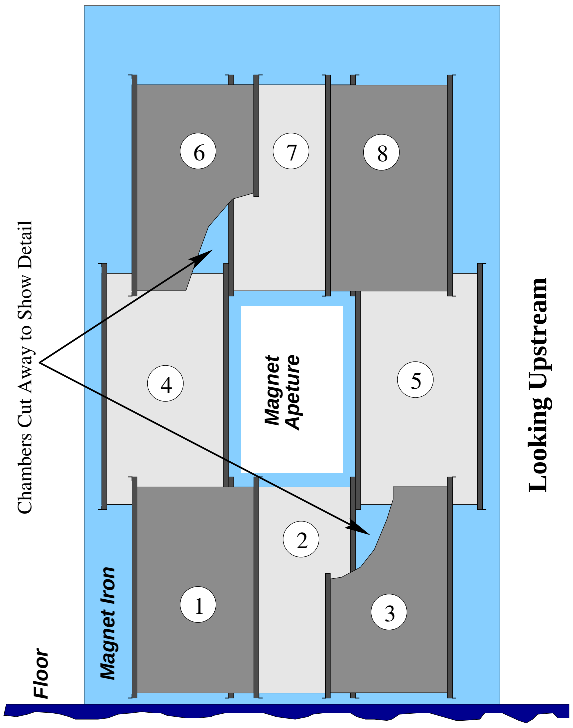


Figure 25: Illustration showing the placement of 8 RPC's into one view. The outer muon detector has three such views, x, y and u.

a hit. The single muon trigger, OM1, is defined by a single muon candidate. The dimuon trigger, OM2, is defined by two muon candidates in separate, non-adjacent towers.

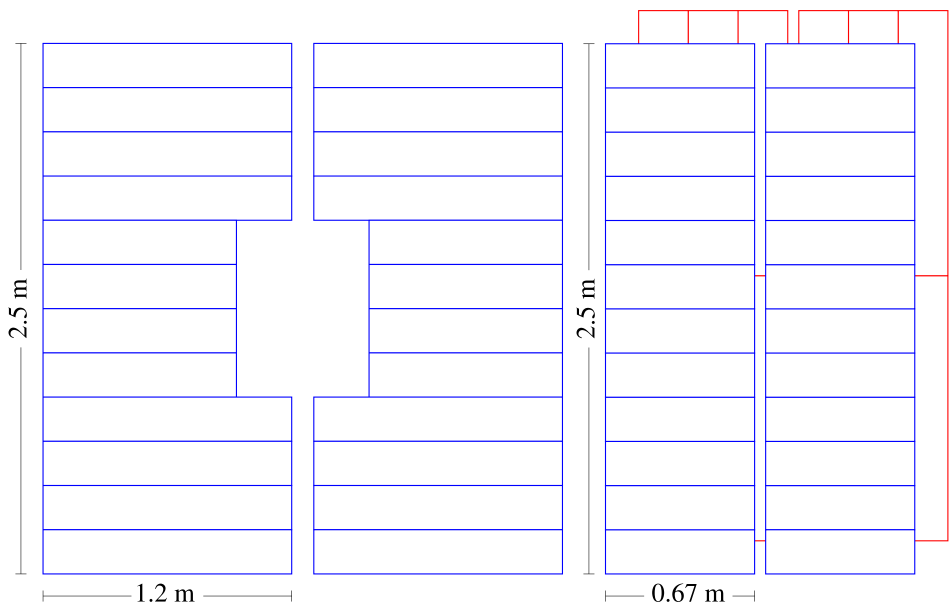
2.9 Triggering and Data Acquisition

To improve the charm signal it is desirable to reject as much electromagnetic background as possible. In FOCUS a typical (20 second) spill contained $\sim 10^8$ interactions, mostly electromagnetic. FOCUS used a hadronic trigger to select charm and reject electromagnetic background. The hadronic trigger is the initial requirement in the selection of events in this analysis. FOCUS triggered on about 30,000 interactions per spill (mostly hadronic). A summary of the separate trigger signals is given in Table 6. The signals combined to form the hadronic trigger are summarized in Table 7 and Table 8.

The triggering is accomplished in two stages. The first level trigger, or master gate (MG), is a fast (200 ns) trigger on a candidate event which initiates the data readout process for most detector elements. The second level trigger decides $1.2\,\mu$s later whether the data readout should be completed. If the second level trigger rejects the event, a $1\,\mu$s clear cycle resets the readout electronics. If the event is accepted the readout continues and lasts approximately $100\,\mu$s, depending on the event.

Table 6: Summary of FOCUS triggers.

Trigger	Description
TR1	A particle from the target
TR2	A particle from the silicon
OH_1	At least one outer particle
$(H \times V)_1$	At least one inner particle
$(H \times V)_2$	At least two inner particles
IE_2	At least two hits in the IE
IM_1	At least one hit in the IM
IM_2	At least two hits in the IM
OM_1	At least one hit in the OM
OM_2	At least two hits in the OM
E_{HI}	Hadronic energy sum over a high threshold
E_{LO}	Hadronic energy sum over a low threshold
E_{IE}	Electromagnetic energy sum over threshold
$E_{\mathrm{IE\text{-}2}}$	Improved electromagnetic energy sum
MULTn	Enough PWC hits for at least n tracks
AM·AMD	Halo muons (veto)
IM(E+W)	Hits in both halves of IM triggers (veto)



(a) OH (b) H × V

Figure 26: The OH and H × V Hodoscope Arrays.

2.9.1 Scintillating Hodoscopes

TR1 is a single scintillator plane situated between the target silicon and the SSD's. A hit in TR1 ensures that there was an interaction in the experiment target. TR2 is a set of four scintillator planes located just after the SSD's. TR2 ensures the tracks went through the SSD's for vertexing and provides a check on TR1.

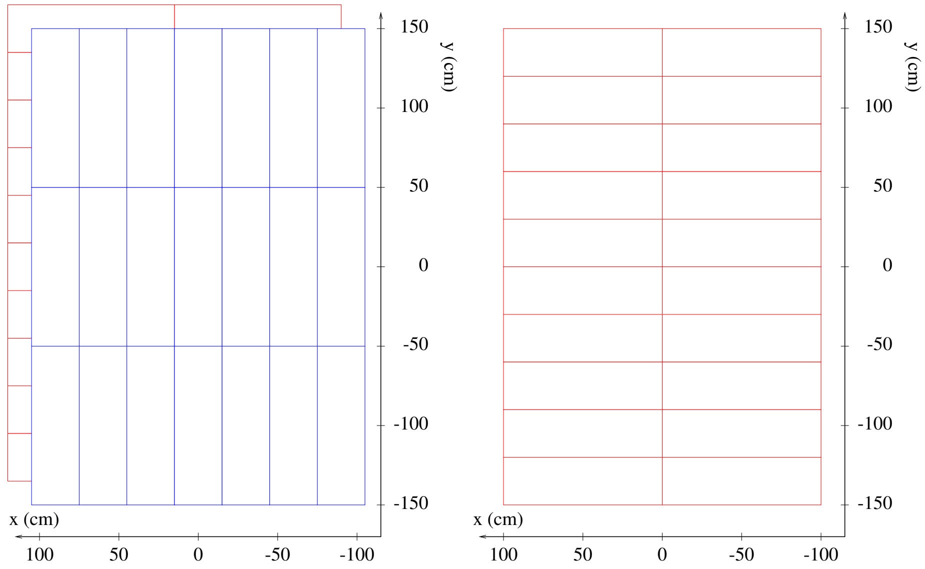
The OH is a set of horizontal scintillators located just upstream of the OE. The OH array has an aperture sized to match the M2 aperture and a pair region gap. The OH checks for a single particle in the outer (wide angle) region of the detector. The analysis magnets bend in the vertical direction so conversion pairs spread out in a narrow vertical band throughout the detector. Hadronic events have greater transverse momentum than electromagnetic, so FOCUS has a wide angle requirement in the trigger. The HxV array is a set of horizontal and vertical scintillators located after the last PWC (P4) just before the IE calorimeter. HxV has a vertical pair region gap and is used to test for wide angle tracks. HxV produces two triggers, one charged particle, $(\mathrm{H}\times\mathrm{V})_1$, or two charged particles, $(\mathrm{H}\times\mathrm{V})_2$, passing through the scintillator array. The OH and HxV arrays, illustrated in Figure 26 are combined to make the two-body requirement:

$$2\mathrm{B} \equiv (\mathrm{H}\times\mathrm{V})_2 + [(\mathrm{H}\times\mathrm{V})_1 \cdot \mathrm{OH}_1] \tag{28}$$

where + represents a logical "OR" and · represents a logical "AND." This two-body requirement checks for two charged tracks in the inner region or a charged track in the inner region and a charged track in the outer region on the opposite side, to avoid a halo muon which travels at low angles (nearly horizontal) across the length of the spectrometer.

IM1 is a set of two planes of scintillators located just downstream of MH1, IM2 is a single plane of scintillators located just upstream of MH2. These arrays are pictured in Figure 27. IM1 and IM2 use logic similar to HxV to produce two triggers for a single inner muon, IM_1, and at least two inner muons, IM_2. These triggers are incorporated into the J/ψ triggers.

AM and AMD are two planes of scintillator upstream of the experiment target and arranged around the beam. These are used to detect muons produced upstream. A coincident signal in both planes is used as a trigger on halo muons, which can be used as a veto in some muon triggers.



(a) IM1V and IM1H

(b) IM2H

Figure 27: The IM trigger counter arrays. a) IM1 has both horizontal and vertical scintillator arrays, b) IM2 has a single horizontal array

Table 7: FOCUS Master Gates.

Trigger	Description	Physics signal
MG1	TR1·TR2·2B·E_{HI}	Hadronic trigger
MG2	TR1·TR2·2B·IE$_2$	$J/\psi \to e^+e^-$
MG3	TR1·TR2·$[\mathrm{IM}_1 + \mathrm{OM}_1] \cdot E_{\mathrm{LO}}$	Semi-muonic decays
MG4	TR1·TR2·2B·$[\mathrm{IM}_2 + \mathrm{OM}_2 + \mathrm{IM}_1 \cdot \mathrm{OM}_1]$	$J/\psi \to \mu^+\mu^-$
MG5	TR1·TR2	e^+e^- pairs (PS)
MG6	TR1·TR2·2B	two-body events (PS)
MG7	TR1·TR2·$[\mathrm{IM}_1 + \mathrm{OM}_1]$	One-muon events (PS)

The FOCUS level-one triggers, or "master gates" are summarized in Table 7. The hadronic trigger requires hits in TR1 and TR2, which are located just upstream and downstream of the SSD's, a two-body requirement and a minimum HC energy.

Table 8: FOCUS Second Level Triggers.

Trigger	Description	Physics signal
TRIG1	MG1$\cdot E_{\text{IE-2}}\cdot$MULT4	Hadronic trigger
TRIG2	MG2$\cdot$(H $\times$ V)$_2$ $\cdot E_{IE}$	$J/\psi \to e^+e^-$
TRIG4	MG4$\cdot$IM$_2$ $\cdot$ (H $\times$ V)$_2\cdot$ (AM$\cdot$AMD)	J/ψ, inner only
TRIG5	MG5	Prescaled MG5
TRIG6	MG6	Prescaled MG6
TRIG8	MG1	Prescaled MG1
TRIG9	MG4$\cdot$OH$\cdot$OM$_2\cdot$MULT2$\cdot$!(AM$\cdot$AMD)	J/ψ, outer only
TRIG11	MG4$\cdot$IM$_1\cdot$OM$_1\cdot$MULT1$\cdot$(H $\times$ V)$_1$ $\cdot$ IM(E+W)	J/ψ, inner/outer

The FOCUS level-two triggers begin with a level-one trigger and apply additional requirements as summarized in Table 8. The prescaled triggers are level one triggers that are used without additional requirements, but trigger at a prescaled fraction of the level-one trigger rate. The prescaled triggers are used for detector calibration. All data used in this analysis requires the hadronic trigger, TRIG1, except for J/ψ data used in muon efficiency studies.

2.9.2 Data Acquisition

The FOCUS Data Acquisition system (DAQ) [29] is outlined in Figure 3.2. The data acquisition system digitized the analog signals from each of the detector elements and merged the results onto magnetic tape. All detector elements digitized their signals and were connected to a single digital RS-485 DAQ bus. The information on the bus is stored in a Dual Ported Memory system. This data is staged on an SGI workstation and written to tape. Approximately 30,000 events were saved per 20 second spill. The readout time was about 100μs per event and about 4 kbytes per

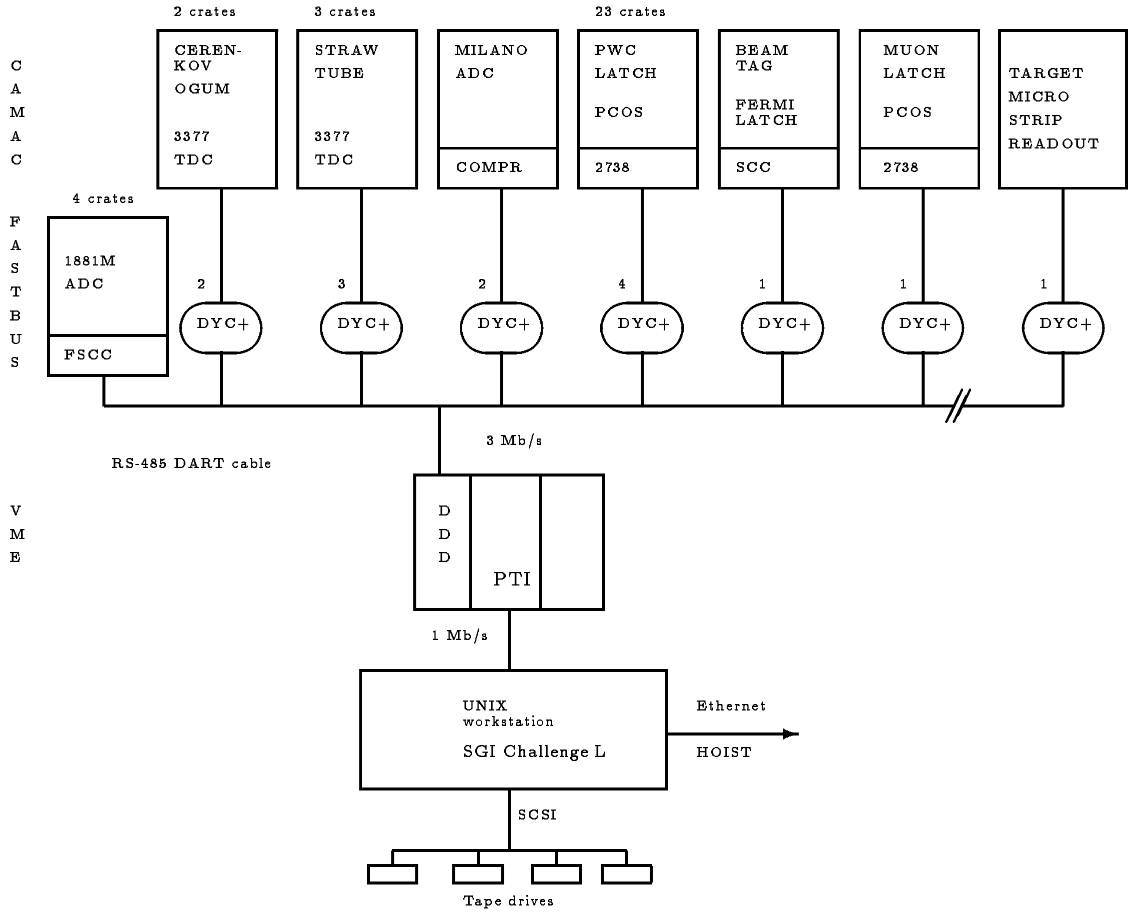


Figure 28: Diagram of FOCUS Data Acquisition

event was written to tape. The time waiting for an event (livetime) for the DAQ was about 85-90%.

The information from all the detector elements is saved in self-contained records for each event. Storing events in complete, independent units allows the data reconstruction and analysis to be done in parallel.

CHAPTER III

DATA RECONSTRUCTION

3.1 **Tracking**

As described in Section 2.4, the FOCUS tracking system is comprised of two systems, the silicon microstrip detector (SSD), located immediately downstream of the experiment target, and the proportional wire chambers (PWC's), located between the first analysis magnet (M1) and the inner electromagnetic calorimeter (IE). The track reconstruction constructs tracks from SSD and PWC hits, links SSD and PWC tracks together, and determines track momenta by measuring the track bends caused by the two analysis magnets.

3.1.1 Silicon Microstrip Tracks

The Silicon microStrip Detector (SSD) described in Section 2.4.1, is located immediately downstream of the target. Each of the four SSD stations contains three views of microstrips. An ionizing particle produces signal in the microstrips and the ADC information resulting from this signal is saved for track reconstruction. SSD track reconstruction proceeds in three steps. The first step is to look at hit clusters to identify the center of the hit. An ionizing particle can deposit charge on several microstrips within a view producing a cluster of hits. The center of the cluster is determined using ADC pulse height information to weight the hits. The next step is to compare hits from corresponding views in the four stations. For each view, hits from three out of four planes are required to be consistent with a line (or a "projection") with a $\chi^2/\text{DOF} < 3$. Hit clusters in the middle stations can be shared for multiple projections but hit clusters from the first and last station cannot be shared in the case of multiple projections. The final step is to combine the projections found for each view into tracks, with the requirement that the track have a $\chi^2/\text{DOF} < 8.0$. When multiple projections are consistent with a track, the projection with the lowest χ^2/DOF is chosen.

The reconstruction efficiency is proportional to momentum. The resolution is also a function of momentum. E687 located the center of SSD hit clusters by averaging the hits with a uniform weighting. For the central region of the SSD's, the track resolution in E687 improved with momentum according to the function

$$\sigma_x = 11.0\ \mu\text{m}\sqrt{1 + \left(\frac{17.5\ \text{GeV}/c}{p}\right)^2} \tag{29}$$

$$\sigma_y = 7.7\ \mu\text{m}\sqrt{1 + \left(\frac{25.0\ \text{GeV}/c}{p}\right)^2}, \tag{30}$$

where p is the track momentum, and the constants 11.0μm and 7.7μm indicate the SSD granularity. For a track in the outer region of the SSD's the resolution is twice as large. In FOCUS, the track resolution was improved by weighting the SSD hits with the ADC pulse height information.

3.1.2 PWC Tracks

The Proportional Wire Chambers (PWC's) are positioned throughout the detector between the first analysis magnet (M1) and the inner electron calorimeter (IE) (cf. Section 2.4.2). The entire PWC system consists of five chambers with four views in each chamber.

Similar to the SSD method, PWC hits are formed into projections for the x, y, u, and v views separately. Since the first analysis magnet will bend the tracks vertically, SSD tracks are used to seed PWC projection finding in the x-view, where the other three views use only PWC hits for projection finding. The x-view projections are not restricted to the SSD seed projections; additional hits in the x-view can be used to make additional x-projections.

Several PWC hit requirements are applied. A track must have PWC hits in the x-view for the first PWC chamber (P0). All other chambers can have a maximum of two missing hits. The total number of missing hits for all views in all chambers cannot be greater than four. These requirements are applied to the two main categories of tracks: 5-chamber tracks and 3-chamber tracks (or "stubs") which are located in the outer angle region of the detector and so have hits only in the first three chambers. The main tracking driver also includes tracks, labeled "type-7000," which are 4-chamber tracks that miss P3. The same stringent tracking requirements were not applied to special categories of events such as kinks (a charged track decaying into a charged and uncharged track) and vees (two tracks that form a vertex anywhere from the target to P0).

Other tracking algorithms, not included in the main driver, were made for additional studies. For halo muon studies, "P1P2P4" tracks were defined with the requirement that tracks only need hits in the large PWC chambers. This allowed for tracking of roughly horizontal-trajectory particles that did not necessarily come from the target region.

A χ^2 fit is done to obtain the track slopes and intercepts. The 3-chamber tracks only leave hits in the chambers located upstream of M2 and so are fit to a single line.

The 5-chamber tracks are bent by M2 and are fit to two lines. In addition, since 5-chamber tracks are bent by M2, momentum determination is possible with only PWC information for these tracks while 3-chamber tracks can have that momentum determined only with the M1 bend information and require the addition of SSD information to determine the track bend in M1. Finally, a maximum of 30 PWC tracks and 600 PWC hits were allowed per event.

3.1.3 Linking

The SSD's are all located upstream of M1 and the PWC's are all located downstream of M1. The linking of SSD tracks and PWC tracks is done by extrapolating both tracks to the M1 bend center and requiring a match. A global χ^2/DOF is calculated for the linked track using both SSD and PWC hits. This χ^2/DOF is used to determine the best match when two PWC tracks match to a single SSD track, although both links are saved. If more than two PWC tracks match well to an SSD track, no link is recorded.

A requirement that SSD tracks are singly linked reduces backgrounds such as e^+e^- pairs. Linking also enables the calculation of the track bend in M1 for the determination of stub momentum. SSD tracks linked to 5-chamber tracks are assigned the M2 momentum.

3.2 **Vertexing**

The point where the photon beam interacts with the target to produce a charm particle is termed the primary vertex. The beam interaction produces other particles at the primary vertex, mainly pions, along with the charm particle. The point at

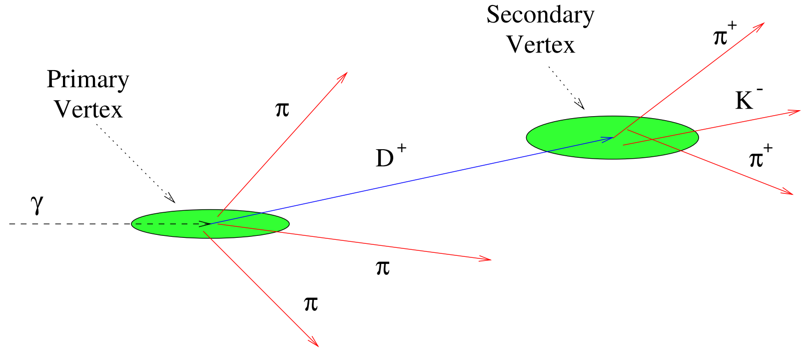


Figure 29: Vertexing Diagram. The uncertainty in vertex positions is represented by the error ellipses which are elongated in z. Not drawn to scale.

which the charm particle decays into its daughter particles is termed the secondary vertex. Charm particles such as a D^+ or D_s^+ decay, on average, a few millimeters from the primary vertex. This means that the charm particle rarely produces a hit in the silicon detectors. The secondary and the primary vertex are found by extrapolating the daughter tracks back to the target region.

To find a vertex with n tracks, χ^2 is minimized:

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{x - (x_i + x_i' z)}{\sigma_{x,i}} \right)^2 + \left(\frac{y - (y_i + y_i' z)}{\sigma_{y,i}} \right)^2 \tag{31}$$

where (x, y, z) are the vertex coordinates, and x_i, x_i', y_i, y_i' are the SSD track slopes and intercepts. The search for secondary charm vertices is straightforward, combinations of tracks are tested and the χ^2 and confidence level is obtained. The confidence level of the secondary vertex is generally required to be greater than 1%.

3.2.1 Primary Vertexing Methods

The search for the primary vertex is complicated by methods of correlating the secondary vertex to the primary vertex and by methods of combining tracks into secondary and primary vertices. DVERT [30] is the FOCUS vertexing package which incorporates two methods for finding the primary vertex.

DVFREE uses a "free-form" method for finding the primary vertex. This method first locates the secondary vertex and then uses the remaining tracks to find the primary vertex (which is required to be upstream of the secondary vertex). Multiple primary vertex candidates are generally selected with criteria such as vertex multiplicity, vertex separation, and which vertex is farthest upstream. For a $D^+ \to K^-\pi^+\pi^+$ decay, for example, a minimum of five tracks is required in free-form vertexing: three tracks to locate the secondary vertex, and at least two tracks to locate the primary vertex.

DVNUCL uses a "candidate-driven" approach. The momenta of the secondary daughter tracks are used to construct the charm track. The reconstructed charm track is included in the search for the primary vertex along with other (non-secondary) tracks. DVERT initialization determines the primary vertex position using SSD hits. DVNUCL then refines the primary vertex position by re-fitting the SSD tracks from the primary vertex using SSD and target silicon (TSSD) hits. Including the candidate charm track and target silicon in the primary vertex determination makes DVNUCL more efficient at fully reconstructing charm events than free-form vertexing, and is used in this analysis.

3.2.2 Vertexing Cuts

The detachment between the primary and secondary vertex, termed ℓ/σ_ℓ, is defined as the distance between the vertices divided by the error. Typical ℓ/σ_ℓ cuts for D

meson analyses range from 3 to 30. Three additional vertex isolation variables are defined as well.

The primary vertex isolation variable is defined as the best vertexing confidence level achieved when each of the secondary daughter tracks is included in the primary vertex. This variable is required to be less than the specified cut value.

Secondary vertex isolation cuts measure the possibility that other tracks in the event originate in the secondary vertex. The secondary isolation variables are defined in two ways: the first method calculates the best secondary vertex confidence level achieved with all unused tracks (excluding primary tracks) tested in the secondary vertex; the second secondary isolation variable repeats the procedure of the first, but considers all non-secondary tracks (including primary tracks). The secondary isolation variables are required to be less than the specified cut value.

Since the charm particle is produced by photon beam interactions with the beryllium target, it is reasonable to restrict the primary vertex (within its error) to the target region. A secondary target interaction may result when charm particles decay within the target. In FOCUS, we sought to reduce this with target separation, and in many analyses it is advantageous to restrict the secondary vertex to be outside the target region. These requirements are termed "out-of-target" cuts and are defined as the distance of the vertex from the nearest target divided by the vertexing error. Similar requirements, termed "out-of-material" cuts, include the targets and the target silicon planes.

3.3 Momentum Calculation

The momentum of a track is determined by measuring how much the track bends in a magnetic field. The momentum of a 5-chamber track is determined with the bend induced by the second analysis magnet M2. The fit of the track parameters is then improved by incorporating the momentum into the track fit. This is done by tracing the particle, using the measured momentum and trajectory, through a model of the M2 magnetic field. The new track is fit to the PWC hits, and a new momentum is calculated. This process is repeated until the global fit to PWC hits and momentum has converged.

The momentum for linked 3-chamber tracks is calculated in a similar way, but using the bend induced by the first analysis magnet M1. SSD hits define the track before M1 and PWC hits define the track after M1. The resolution of the momentum determined with M1 is

$$\frac{\sigma_p}{p} = 0.034 \times \frac{p}{100 \text{ GeV}/c}\sqrt{1 + \left(\frac{17 \text{ GeV}/c}{p}\right)^2} \tag{32}$$

the resolution of the momentum determined with M2 is

$$\frac{\sigma_p}{p} = 0.014 \times \frac{p}{100 \text{ GeV}/c}\sqrt{1 + \left(\frac{23 \text{ GeV}/c}{p}\right)^2}. \tag{33}$$

TRKFIT is an algorithm which improves the track momentum resolution by refitting the tracks with additional constraints. Stubs are constrained to meet the microstrip track in M1. TRKFIT constrains 5-chamber tracks the same way and, in addition, uses the bend angles from both M1 and M2 to calculate a new momentum and confidence level for the fit. The momentum from TRKFIT improves the mass resolution for $D^+ \to K^-\pi^+\pi^+$ by $\sim 5\%$. The TRKFIT confidence level can be used to reduce muon misidentification. Particles that have decayed in-flight between M1 and M2 will cause a discrepancy between the momentum calculated using M1 or M2. Applying a TRKFIT confidence level cut for 5-chamber tracks reduces inner muon misidentification caused by particle decays in-flight.

3.4 Particle Identification

Several systems are used in FOCUS to identify the type of particle for a particular track. Three threshold Čerenkov detectors are designed to distinguish electrons, pions, kaons, and protons. Inner and outer electron calorimeters are used to identify electrons. Finally, inner and outer muon detectors provide excellent muon identification.

3.4.1 Čerenkov

The FOCUS spectrometer contains three threshold Čerenkov detectors (described in Section 2.6) which are used to evaluate the likelihoods of an electron, pion, kaon, and proton hypothesis for a track. The FOCUS Čerenkov algorithm, CITADL, compares the observed number of photons in each of the three Čerenkov detectors with the predicted number of photons that would result from a hypothetical particle with the momentum of the track. The Čerenkov data reconstruction methods are discussed at length in Reference 22 and Reference 31.

Photons from a particle above threshold are emitted at the Čerenkov angle creating a Čerenkov cone (with the "apex" of the cone at the point the particle enters the chamber and the "open" end of the cone at the detector end of the chamber, cf. Figure 18). The Čerenkov angle is different for the different gasses in the 3 different

Table 9: Čerenkov Particle ID Cuts

Čerenkov cut variables			
particle		definition	typical cut
kaonicity	(kaon not pion)	$\Delta W_K \equiv W_\pi - W_K$	> 0.5
pion id 1	(pion not kaon)	$\Delta W_\pi \equiv W_K - W_\pi$	> 5
pion id 2	(pion consistency)	$\pi_{con} \equiv \min\,(W_e, W_K, W_p) - W_\pi$	> -6
proton id 1	(proton not pion)	$\Delta W_{p1} \equiv W_\pi - W_p$	> 5
proton id 2	(proton not kaon)	$\Delta W_{p2} \equiv W_K - W_p$	> 1

Čerenkov chambers. A longer detector will have a larger cone and more photons to detect, the chamber length, cone size, and detector size are all matched to provide efficient detection.

Only the detector cells within the particle's Čerenkov cone are considered in the comparison of the observed signal with the predicted signal. A likelihood, $\mathcal{L}$, is constructed with the product of the firing probabilities for the relevant detector cells for all three Čerenkov detectors. CITADL returns the variable $W_{obs} = -2\ln\mathcal{L}$ for all four particle hypotheses. A common FOCUS variable used to distinguish kaons from pions is the degree of "*kaonicity*" defined as:

$$kaonicity = W_{obs}(\pi) - W_{obs}(K) \tag{34}$$

Since W_{obs} is a negative log likelihood, $W_{obs}(\pi)$ is expected to be higher than $W_{obs}(K)$ for a kaon.

3.4.2 Electron Calorimetry

The threshhold Čerenkov detectors can distinguish electrons and pions at low momentum, but above the hadronic Čerenkov threshholds calorimetry is needed. The electron calorimeters, described in Section 2.7, are designed so that most of the electromagnetic energy, E, is contained within the calorimeter, while most of the hadronic energy passes through. (Muons leave neglible amounts of energy behind.) Since the electron rest mass is only a small component of the total electron energy, electrons are selected by requiring that E/p be close to 1.0.

For the inner electron calorimeter (IE), electrons were required to have $0.8 < E/p < 1.2$. Tracks passing this cut are assigned an IEID value of 10. The IEID is then incremented by 0-3 based on a Čerenkov analysis, with an IEID score of 13 being the best electron ID score. The outer electron calorimeter identifies electrons in a similar way. Using an E/p cut, Čerenkov information, and characteristics of the shower evolution, an OESCORE between 300 and 399 is set for electrons. The outer electron calorimeter was not used in this analysis. A minimum IE energy is required in the second level hadronic trigger, but the inner electron calorimeter was otherwise not used in this analysis. IE reconstruction methods are discussed at length in Reference 32.

3.4.3 Muon Identification

As described in Section 2.8.1 and Section 2.8.2 the inner muon detector is composed of six planes of muon hodoscopes (MH) and the outer muon detector is composed of three planes of resistive plate chambers (RPC). The detectors are placed downstream of a thick layer of steel which muons can penetrate rather easily but hadrons cannot, so that charged particles passing through the muon detection planes are predominantly muons.

Muon candidates are constructed by projecting PWC tracks into the muon detectors and looking for hits [33]. For the inner (outer) muon system, at least 4 (2) out of the 6 (3) planes must have hits associated with the track. After hit finding, the tracks are fit to the muon hits and a confidence level for the fit is obtained. A muon confidence level greater than 0.01% was required for muon identification in the initial data selection (SKIM1).

For inner muons [34], the set of hits used to calculate the confidence level for a track is also used to calculate the isolation of that track from other muon candidates. Using this same set of hits, all the other tracks in the event have a confidence level calculated. The highest and second highest confidence level that result are saved.

Low momentum tracks have a higher search radius for muon hits than high momentum tracks. This makes it easier to obtain a good confidence level for low momentum tracks. Therefore a minimum momentum cut is necessary for the muon confidence level to be believed. Also, if a muon candidate is accompanied by a nearby low momentum track, the muon candidate will have a lowered probability of being isolated.

Outer muon identification [35] is complicated by the presence of the internal magnetic field of M2. The outer muon algorithm incorporates the bending due to the magnetic field in addition to the smearing due to multiple Coulomb scattering.

3.5 **Data Processing**

At the completion of the FOCUS data run, nearly 6.5 billion events had been written to $\sim$6000 8 mm tapes amounting to 25 terabytes of information. The individual events are to be independent and the DAQ information for each event was written to tape in a single self-contained record. This meant events could be analyzed individually, in parallel, on large clusters of computers at many different institutions. The DAQ saved "raw" detector information (signals from each PWC wire, Čerenkov cell, RPC strip, etc.) for each event. From the raw data, events were reconstructed into tracks and vertex objects with momentum, energy, and particle identification information.

3.5.1 Event Reconstruction

The first step of data reconstruction, Pass-1, applied only the most general event selection criteria and resulted in an output of an additional 6000 tapes. To reduce the data sets into sizes managable for individual users (generally on the order of 50 tapes) it was necessary to separate events into different categories and apply selection cuts. The cuts needed to be hard enough to make the data sizes manageable, but loose enough so that optimizations can be performed for each analysis.

3.5.2 Event Selection

During Skim-1, reconstructed events were separated into different physics topics (or "data streams"). The six data streams were written on separate tapes totalling approximately 2500 tapes. Table 10 lists the six different Skim-1 data streams.

Skim-2 further separated the Skim-1 streams into 40 data streams using an additional 2500 data tapes. From stream 6, the rare and forbidden meson decays, as well as the appropriate normalizing modes, were written to the data stream labeled "FSFE" which totaled 16 tapes. The cuts applied to the FSFE data are summarized in Table 13.

Table 10: Skim-1 Superstreams

Stream	Description
1	Semi-leptonic, di-leptonic
2	Topological vertexing and K_s^0
3	Calibration
4	Baryons
5	Diffractive (light quark states)
6	SEZDEE (Mesons)

CHAPTER IV

MONTE CARLO SIMULATION

The FOCUS Monte Carlo simulation, "ROGUE," utilizes PYTHIA (version 6.127) to simulate charm production. PYTHIA simulates the production of particles from a fixed-target neutron or proton. PYTHIA is input a random photon energy corresponding to FOCUS beam energies and outputs a list of particles and their momenta. The particles are produced according to the standard model simulation selected; in our case, a photon-gluon fusion process.

The output particles are selected from a list of particles relevant to FOCUS data. If PYTHIA fails to produce an event that would likely appear in FOCUS data, PYTHIA is given a new random photon energy. When PYTHIA produces an event of interest, ROGUE simulates the subsequent decays (if any) of the particles and steps the particles through a simulation of the FOCUS detector. ROGUE incorporates the decay matrix elements when known (e.g. $D^+ \to K^-\pi^+\pi^+$). For decays with unknown matrix elements, the matrix elements are defined to be uniform in phase space.

The events are then reconstructed in a manner as close as possible to the data reconstruction, with many simulation parameters available to tune the particle production, detector response, and background simulations. The FOCUS collaboration has extensively tuned Pythia and ROGUE to match these properties. The most significant aspect of the Monte Carlo simulation for this analysis is the calibration of muon efficiencies and muon misidentification rates, which are discussed at length.

4.1 Muon Efficiencies

As described in Section 2.8.2 and Section 2.8.1 the inner muon detector is composed of six planes of muon hodoscopes (MH) and the outer muon detector is composed of three planes of resistive plate chambers (RPC). The detectors are placed downstream of a thick layer of steel which muons can penetrate rather easily but hadrons cannot, so that charged particles passing through the muon detection planes are predominantly muons.

The variables for selecting muons are described in Section 3.4.3. For this analysis muon tracks were identified with a muon confidence level greater than 1%. Inner muon candidates were required to have hits in at least 4 of 6 planes. Outer muon candidates were required to have hits in at least 2 of 3 planes and were required to have a track trajectory that traverses at least 150 cm of material (mostly iron).

However, hadrons can "punch-through" the absorbing material, so a small percentage of hits in the muon detectors may be from particles other than muons. Hadrons can decay into muons to produce a muon hit. Detector noise and halo muons produce random hits which can lie in the path of a tracked particle that never reaches the muon detector. The resultant backgrounds affect both the calibration of muon detector efficiency and particle misidentification.

To calibrate the efficiency of muon detection, a large sample of muons is available from J/ψ and semileptonic events. However, J/ψ events are not representative of typical FOCUS charm events. In J/ψ's, charm is "hidden" in a particle with a charm quantum number of zero. For typical FOCUS charm events, such as D^+ decays, charm is openly manifested in particles with a non-zero charm quantum number. J/ψ events are a very clean source of muons. For example, dimuon triggered J/ψ events have lower multiplicity and less background than hadronic triggered events. J/ψ muons also have a higher momentum than open-charm muons. Therefore muon detection efficiencies determined by J/ψ muons need to be compared to efficiencies measured for open-charm muons.

To calculate the muon identification efficiency with the requirement that at least y out of n planes recorded a muon hit, recall the binomial probability distribution:

$$P(y) = \begin{pmatrix} n \\ y \end{pmatrix} p^y q^{n-y} \tag{35}$$

where

$$\begin{pmatrix} n \\ y \end{pmatrix} \equiv \frac{n!}{y!(n-y)!} \tag{36}$$

and p is the probability of a hit (the efficiency of a plane) and $q = 1 - p$. For example, requiring hits (correlated with the track) in at least 4 of 6 planes, with an average MH plane efficiency of 95%, results in a muon identification efficiency greater than 99.9%.

The efficiency of an individual plane is determined by the probability that the plane fired given that all the other planes fired. For hadronic triggered data this method is unbiased. For dimuon triggered data, this method is unbiased as long as the muon hits recorded are the same as the muon hits that trigger and the trigger requirements are the same or less than the muon identification requirements. The outer muon electronics that record the data (the data latches) and the outer muon trigger electronics may have different efficiencies. For example, if an RPC has inefficient trigger hardware, a hit in that RPC would often fail to trigger the dimuon trigger but might still be recorded by a trigger from another RPC in that tower. The

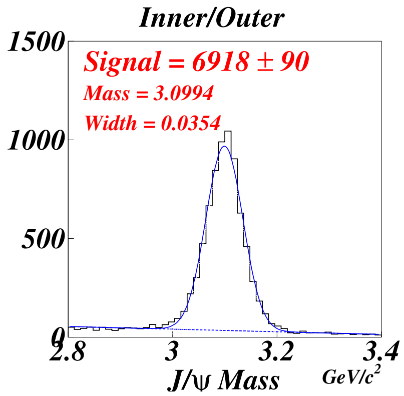


Figure 30: $J/\psi \to \mu^+\mu^-$ mass plot with one inner and one outer muon.

relative number of hits in that RPC would be high relative to the other RPC's in that tower. The discrepancy between the outer muon data latches and muon triggers has been determined to be small for the dimuon trigger.

The J/ψ sample, shown in Figure 30, is obtained by selecting events where an inner muon track and an outer muon track form a vertex with the J/ψ mass. The inner/outer requirement is made to ensure there is only one muon incident on the inner or outer detector. The semileptonic decays $D^+ \to K^{0*}\mu^+\nu \to K^-\pi^+\mu^+\nu$ and $D^{*+} \to \pi^+D^0 \to K^-\pi^+\mu^+\nu$ provide the largest available sample of muons ("D^*-tagged" and "K^*-tagged") from open-charm events.

The $D^+ \to K^{*0}\mu^+\nu \to K^-\pi^+\mu^+\nu$ sample, shown in Figure 31, is obtained from events where the K, π, and μ tracks form a vertex with an invariant mass less than the D^+ mass and the K and π, tracks form a vertex within 50 MeV/c^2 of the K^* mass. A vertex detachment requirement $\ell/\sigma_\ell > 10\%$ was applied (The ℓ/σ_ℓ for K^* calculated using the FOCUS routine "pvrt"). The K^*-tagged muon sample yield is obtained by fitting the K^* mass to a Breit-Wigner signal and a linear background.

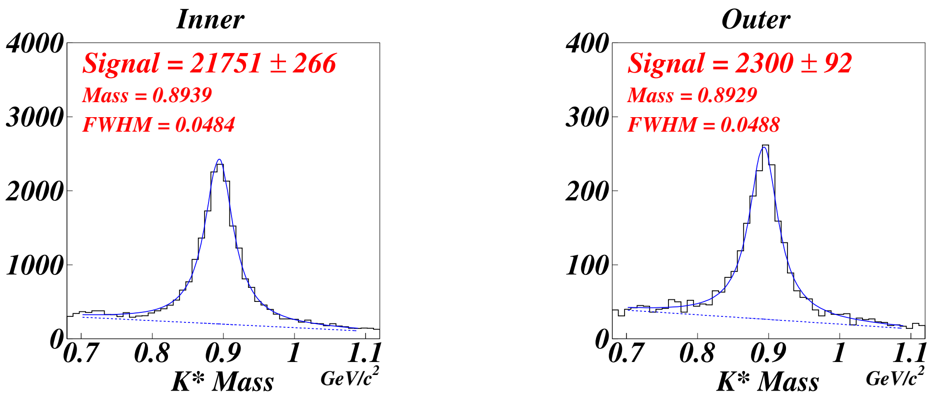


Figure 31: K^* mass plots for $D^+ \rightarrow K^{0*}\mu^+\nu \rightarrow K^-\pi^+\mu^+\nu$ events with an inner or outer muon.

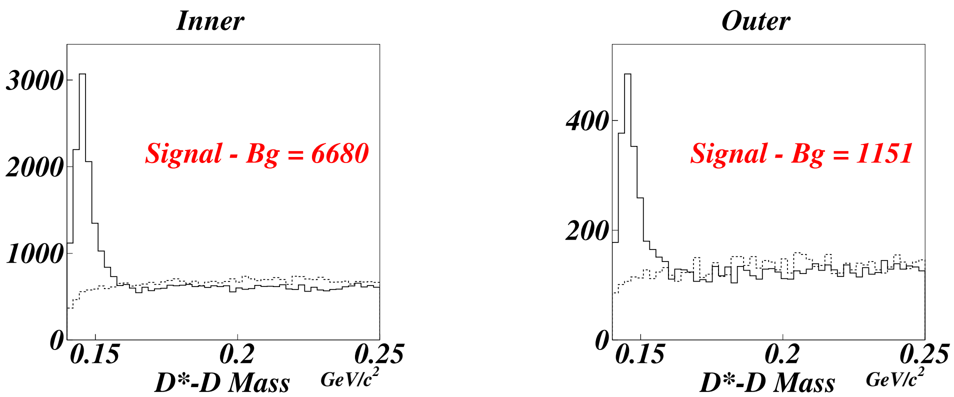


Figure 32: $D^* - D$ mass difference plots of $D^{*+} \rightarrow \pi^+ D^0 \rightarrow K^-\pi^+\mu^+\nu$ events with an inner or outer muon.

The $D^{*+} \rightarrow \pi^+ D^0 \rightarrow K^-\pi^+\mu^+\nu$ sample is obtained by selecting a secondary vertex with K and μ daughters and a π from the primary vertex. The neutrino momentum is determined by assuming a D^0 mass and calculating the neutrino momentum up to a quadratic ambiguity (the solution with the lowest $D^* - D^0$ mass is selected). The $D^* - D^0$ mass is plotted in Figure 32 and the D^*-tagged muon yield is calculated as the right-sign ($D^0 \rightarrow K^-\mu^+\nu$) data minus the wrong sign ($D^0 \rightarrow K^+\mu^-\nu$) data

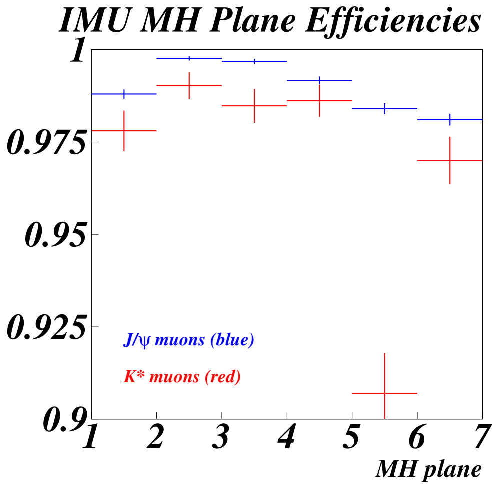


Figure 33: MH plane efficiencies calculated as the probability that a plane had a hit given that the other 5 planes had hits.

4.1.1 Inner Muon Efficiencies

The inner muon MH plane efficiencies determined with K^*-tagged and J/ψ muons are shown in Figure 33. The inner muon hodoscopes are modeled with 100% efficiency for the Monte Carlo. Since the average MH plane efficiency is greater than 98%, the Monte Carlo assumption of 100% MH plane efficiency is reasonable. The muon identification inefficiency becomes negligible when the 4 of 6 plane requirement is applied (cf. Section 4.1).

4.1.2 Outer Muon Efficiencies

As shown in Figure 34, the outer muon system is composed of three planes of RPC's. Each plane is built with eight RPC's arranged outside the aperture of M2. Three planes aligned in z define a tower.

During data acquisition and "Pass1" data reconstruction, two methods were used to monitor the outer muon RPC efficiencies. The RPC efficiency was calculated as

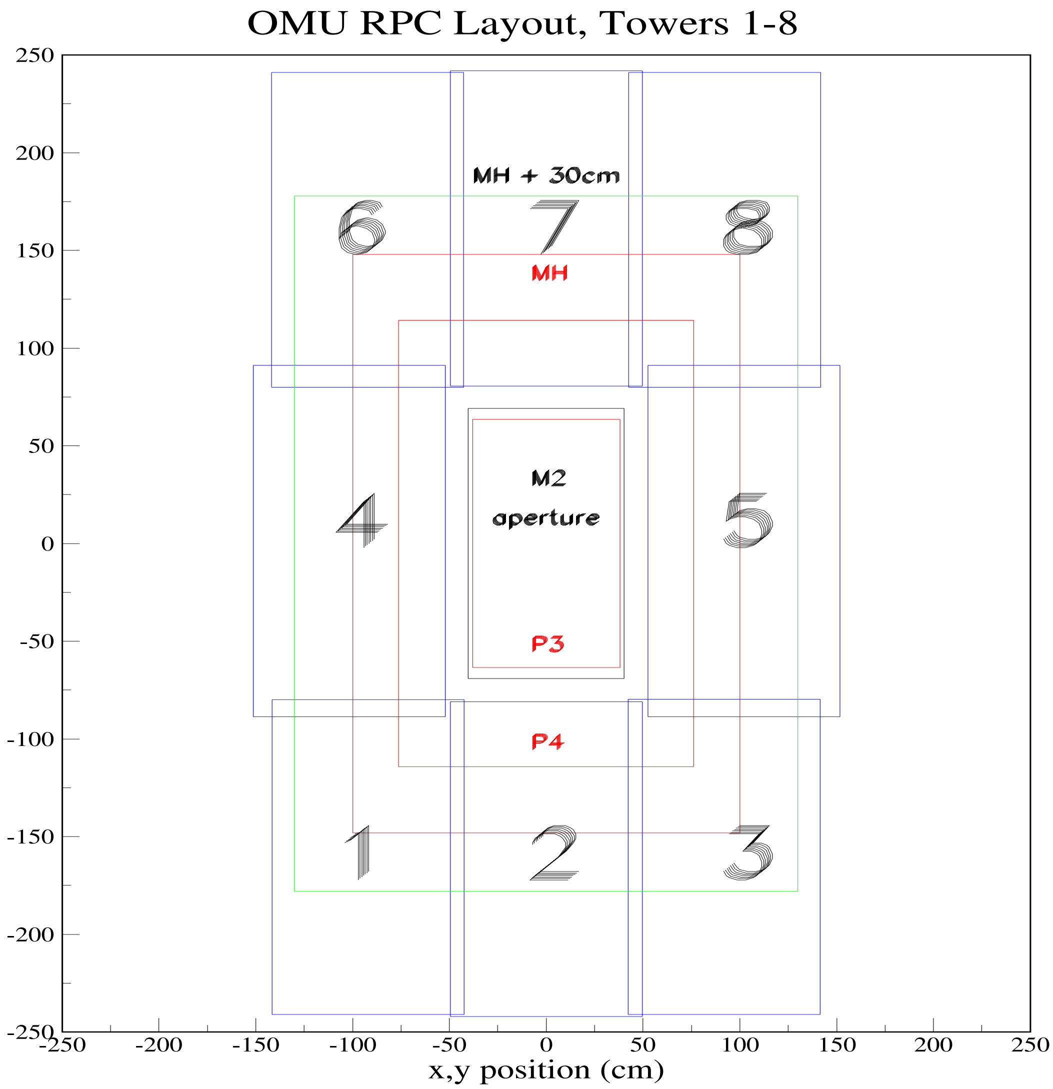


Figure 34: The outer muon RPC's showing the x,y position of each RPC. Each plane consists of eight partially overlapping RPC's labeled 1-8. Comparisons with other detector sizes are shown (from innermost to outermost): (i) PWC P3, (ii) M2 aperture, (iii) PWC P4, (iv) the MH planes, (v) 30 cm outside the MH planes.

the number of times a RPC fired given that both the other RPC's in that tower fired. The first online method simply required a hadronic trigger. The second method had no trigger requirement, but incorporated a crude horizontal trajectory requirement using both muon systems. Outer muon hit positions were determined by comparing the overlap ("cluster") of RPC strips that recorded hits from 2 or more views. Inner muon hit positions were determined by comparing the overlap of MH planks that recorded hits from 4 or more views. The inner muon system is located ~1300 cm downstream of the outer muon system. The two detectors overlap 45 cm in x and 70 cm in y (Figure 34). A horizontal trajectory was determined by requiring that an outer muon cluster be within 30 cm (in x and y) of an inner muon cluster. This crude tracking requirement selects "halo" muons with a trajectory originating outside the target region. The RPC efficiencies from Pass1 method-2 are ~10% higher than those from Pass1 method-1.

In addition to the online methods, RPC efficiencies were tested with reconstructed data from J/ψ muons (Figure 30) and open-charm muons (Figure 31 and Figure 32). The RPC efficiencies measured with the relatively clean J/ψ events are ~10% higher than the efficiencies measured with the semileptonic data. The efficiencies from the two reconstructed data samples closely match the efficiencies from the two Pass1 calibration methods.

Figure 35 compares the different efficiencies measured for all 24 RPC's, for all four methods. The hatched region shows the range of RPC efficiencies measured by the two Pass1 methods on unreconstructed data (whose statistics are very high so the statistical errors are insignificant here). The points with error bars show the RPC efficiencies for J/ψ muons and open-charm muons. The values are listed in RPC order as described in Figure 34. The RPC's in the corners (1,3,6,8,etc.) have very few hits. The center RPC's on each side (4,5,12,13,20,21) have the highest rate of hits and the center RPC's top and bottom (2,7,10,15,18,23) have a significant number of hits as well. The outer muon CL requires at least 2 of 3 planes to fire. The average of the

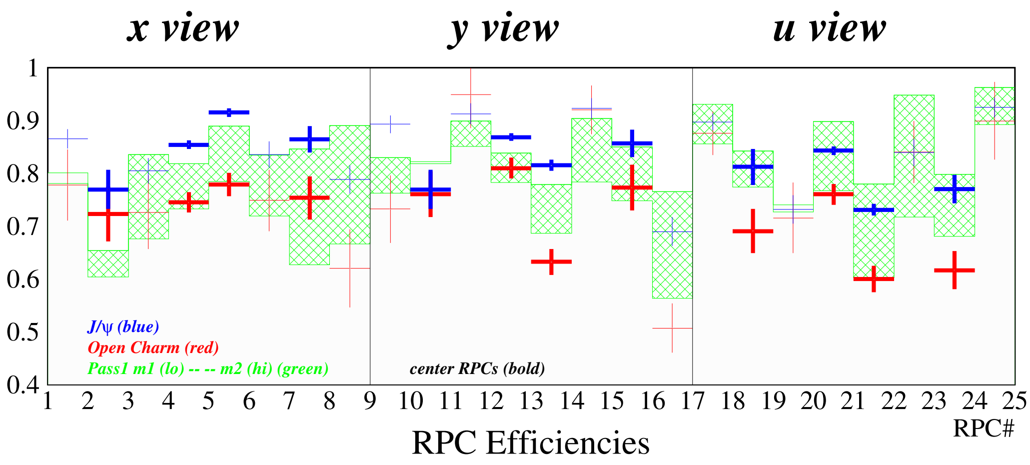
Pass1 online methods shows a 77% average RPC efficiency, which translates to an 87% outer muon ID efficiency.

The J/ψ events in the upper plot of Figure 35 include both hadronic triggered and dimuon triggered events. To more closely match the character of hadronic events, J/ψ events with a hadronic trigger requirement are shown in the lower plot of Figure 35. The ratio of J/ψ muon efficiencies to open-charm muon efficiencies is shown in Figure 36. With the hadronic trigger requirement, the efficiencies from J/ψ and open-charm muons are consistent within $\sim$10% for the center RPC's.

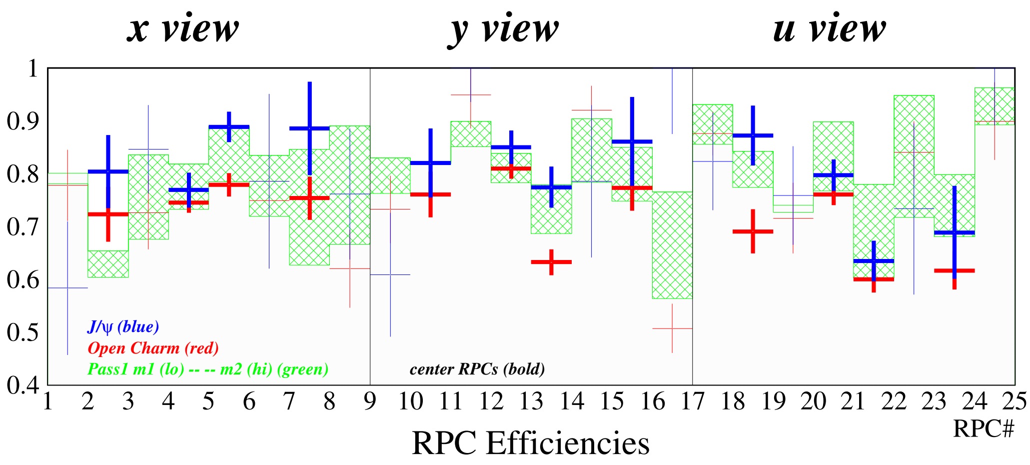
4.1.3 Outer Muon Position-Dependent Efficiencies

Muons from J/ψ events have a different radial distribution than muons from open-charm events as shown in Figure 37. The J/ψ radial distribution is also biased by the outer dimuon trigger which excludes the center towers 2 and 7. Radial RPC efficiency variations, combined with differences in radial distributions for two data samples will result in different average efficiencies for the two samples.

RPC efficiencies as a function of x (for towers 4,5) and as a function of y (for towers 2,7) are shown in Figure 38. The J/ψ muon efficiencies are slightly higher than open charm muon efficiencies, but the spatial variation is significant. There is no consistent pattern for the spatial variation in efficiency of the corner RPC towers, as shown in Figure 39, Figure 40, and Figure 41. Therefore, some of the efficiency differences due to spatial variations in the muon hits can be removed by weighting the J/ψ muons by the radial distribution (Figure 42) of open-charm muons for the center towers.

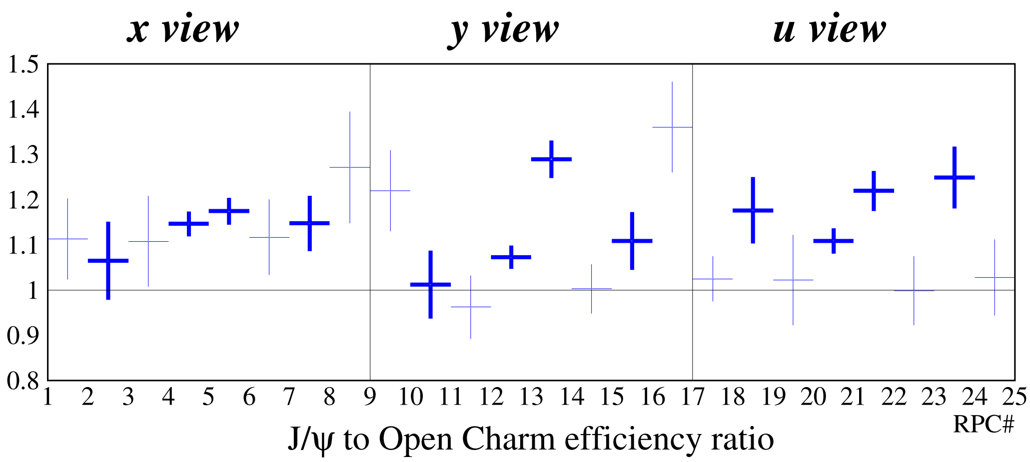


(a) All J/ψ events included

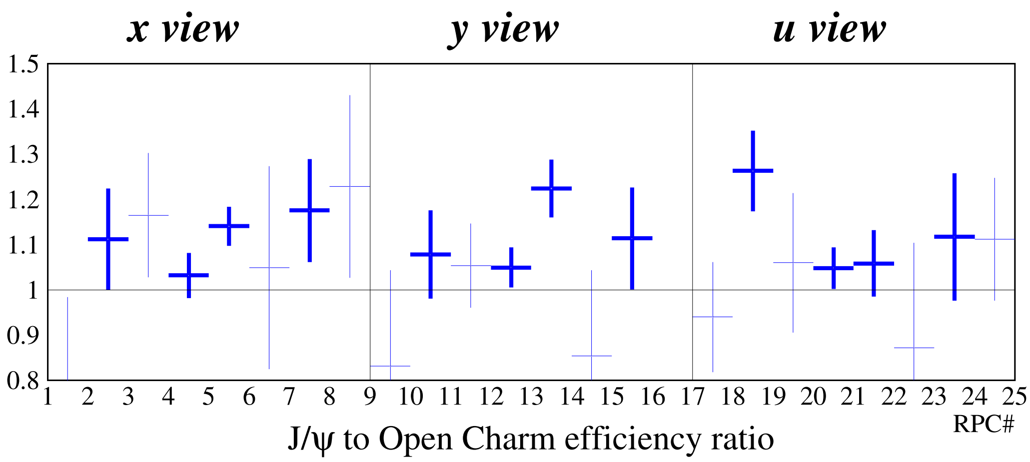


(b) Hadronic trigger required for J/ψ

Figure 35: Efficiencies for all 24 RPC's (3 planes of 8 chambers) are shown for the various samples. The hatched area gives the range of efficiencies measured with the 2 Pass1 methods. The red points give the efficiencies measured with muons from the combined open-charm samples. The blue points give efficiencies measured with J/ψ muons. The points in bold face indicate the middle towers which contain most of the hits.



(a) All J/ψ events included

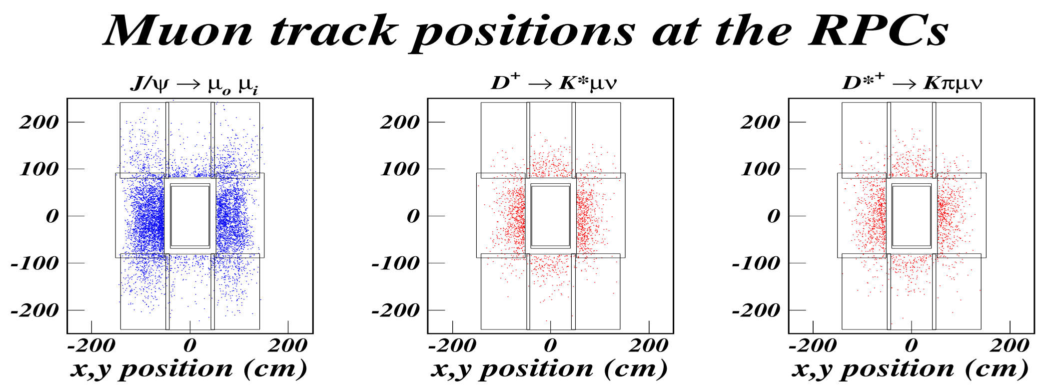


(b) Hadronic trigger required for J/ψ

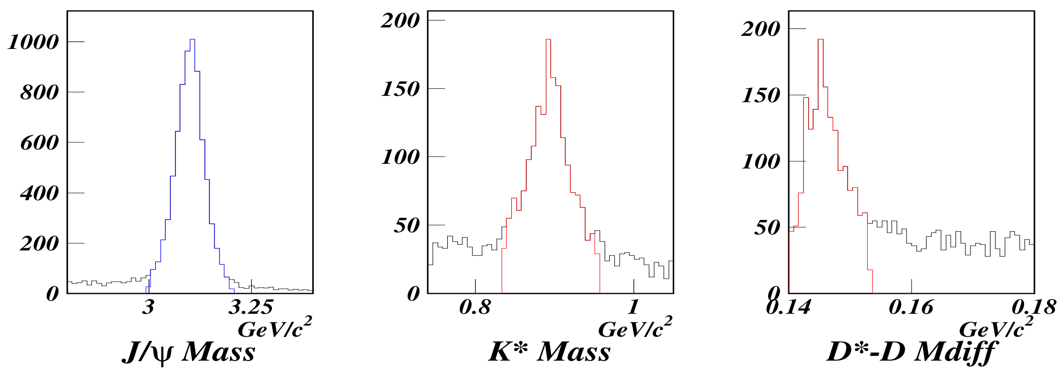
Figure 36: The ratio of efficiencies measured with J/ψ muons vs. open-charm muons shows more clearly the spread in efficiency uncertainty. The points in bold indicate the center towers which contain most of the hits.

4.1.4 Outer Muon Efficiency Monte Carlo

RPC efficiencies were incorporated into the Monte Carlo simulation for each individual RPC and for 6 run periods. Several efficiency parameter files were created



(a) muon hit distribution relative to the eight RPC towers

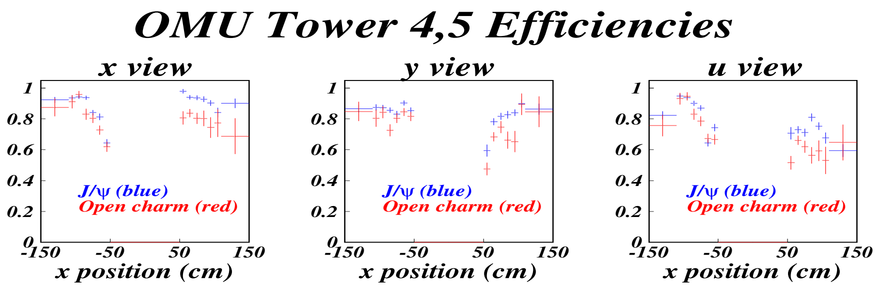


(b) invariant mass of each sample

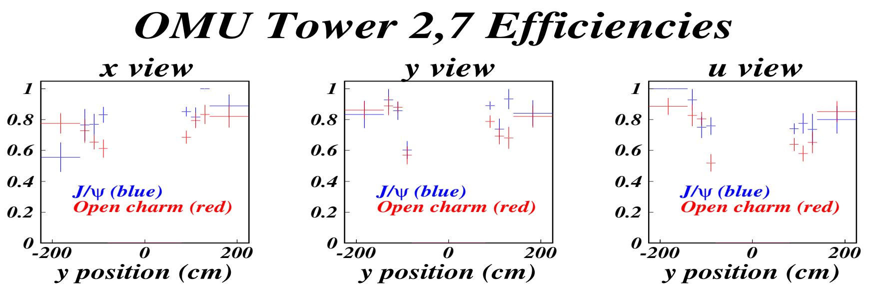
Figure 37: Muons from 3 separate sources are shown. The muon track xy-positions at the RPC's are plotted above with the corresponding mass plots shown below.

and the efficiencies can be chosen to match the desired method. The RPC efficiency code in Monte Carlo was checked by setting all RPC efficiencies to 85%, generating a K^*-tagged and J/ψ muon sample, and comparing RPC hits in the three views. The reproduction of input efficiencies is shown in Figure 43. Since most muon hits appear in towers 4 and 5, the Monte Carlo code is accurate to $\sim$1%.

The effect of the OMU efficiency uncertainty is shown in Figure 44. K^*-tagged muons were generated under seven different efficiency scenarios: (1) Pass1 method-1,



(a) efficiencies vs x-position

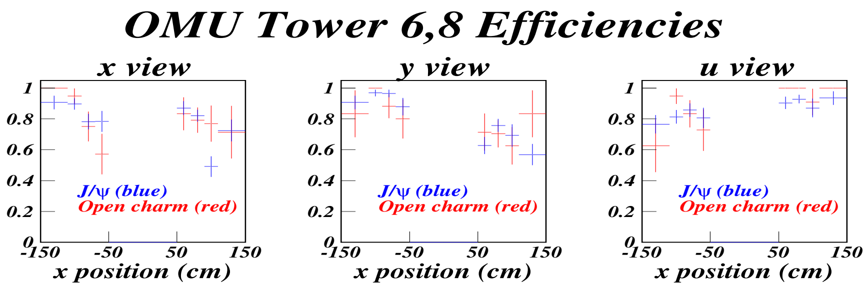


(b) efficiencies vs y-position

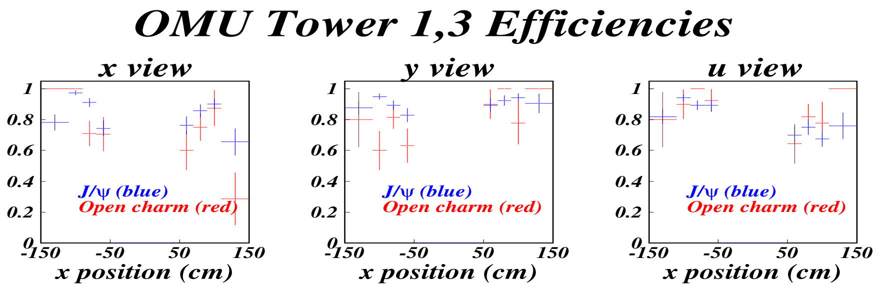
Figure 38: RPC efficiencies as a function of x and y for the central towers. The most central positions have the highest hit rate and in some cases have a much lower efficiency. Efficiencies from J/ψ muons are shown in blue, open-charm muons are shown in red.

(2) Pass1 method-2, (3) the average of the Pass1 methods, (4) open-charm muons and finally (5) J/ψ muons, selected three ways. The three methods for selecting J/ψ muons are: (i) use all events including dimuon-triggered events, (ii) use all events with re-weighting to the open-charm radial distribution, and (iii) as in (ii) but require the hadronic trigger.

A systematic error from the outer muon efficiency uncertainty can be calculated by comparing the Monte Carlo yields from different efficiency settings and calculating



(a) upper corner towers



(b) lower corner towers

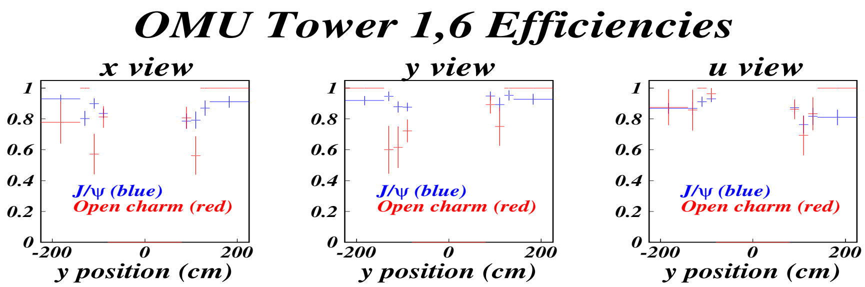
Figure 39: No clear pattern is apparent for efficiencies as a function of x for the corner RPC's.

the resulting rms of the yields [36]:

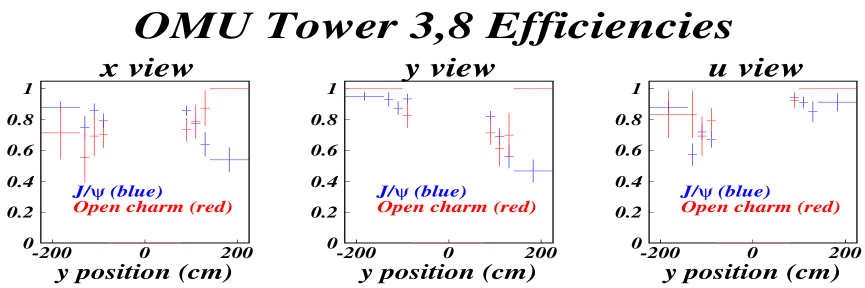
$$\sigma_{sys} = \sqrt{\frac{\left(\sum Y_i^2\right) \; - \; N < Y >^2}{N \; - \; 1}}. \tag{37}$$

Four outer muon systematic errors for open-charm events can defined as follows: (1) include Pass1 method-1, Pass1 method-2, the default setting (Pass1 average) and the open charm setting, but ignore any J/ψ number; (2) repeat (1) and include efficiencies for all J/ψ events; (3) repeat (1) and include efficiencies for all J/ψ events with radial re-weighing; (3) repeat (1) and include efficiencies for hadronic triggerred J/ψ events.

The Monte Carlo sample is 10$\times$ the data sample, and this decay channel provides the largest sample of muons found in FOCUS data. For this data sample, the relative



(a) upper corner towers



(b) lower corner towers

Figure 40: No clear pattern is apparent for efficiencies as a function of y for the corner RPC's.

Table 11: Outer Muon Efficiency Systematic Errors

method	description	sys. error
1	w/o any J/ψ efficiencies	5.6%
2	incl all J/ψ events	6.3%
3	incl all J/ψ w/ radial adj.	5.8%
4	incl hadronic triggered J/ψ events	5.2%
	data relative statistical error	3.9%

statistical error is 3.9% compared to a relative systematic error of ∼5.5%. For all other muon samples, the statistical error will be much greater than the outer muon efficiency systematic error. The Pass1 method-1 and open charm efficiencies are consistent and

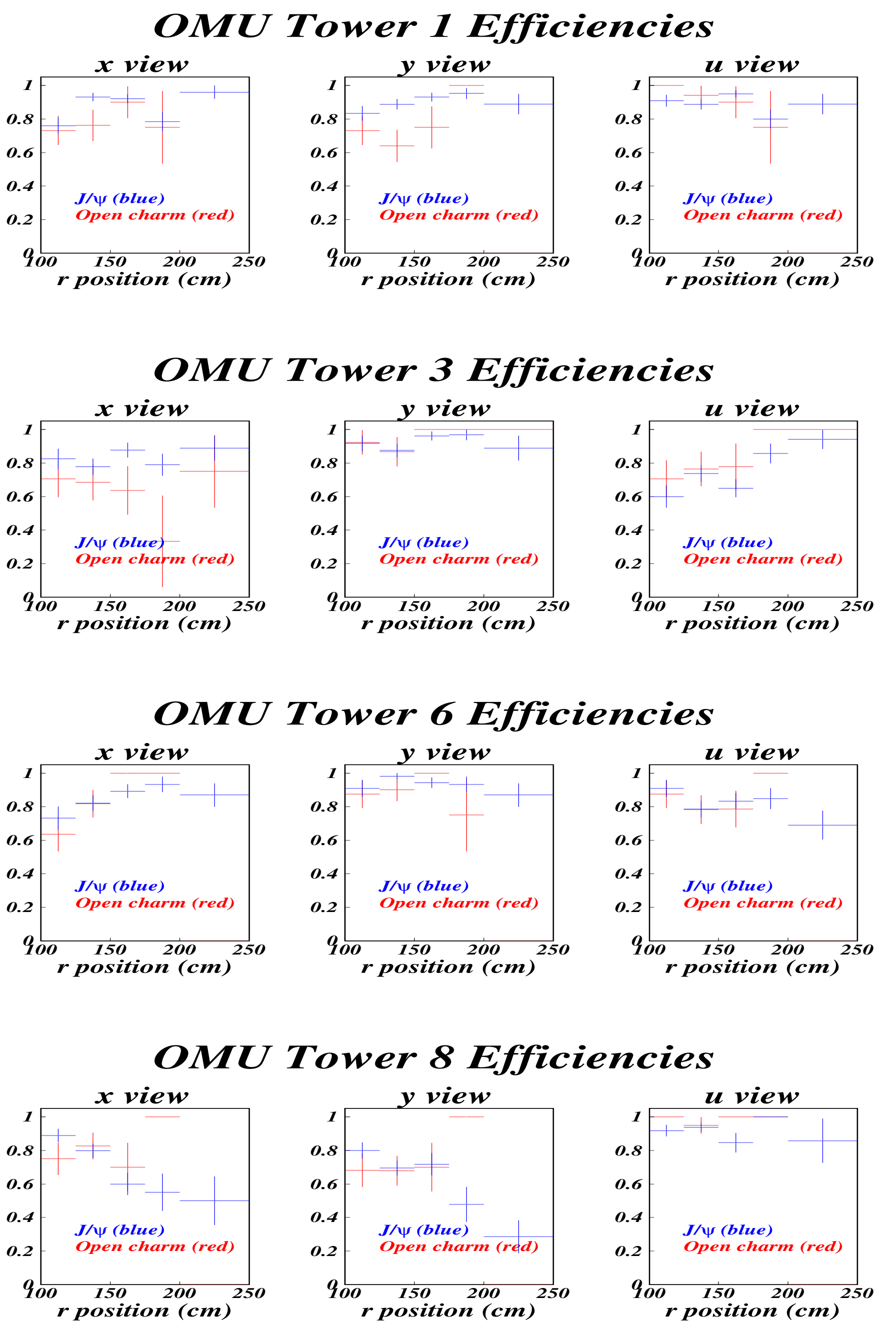


Figure 41: No clear pattern is apparent for efficiencies as a function of radial position for the corner RPC's.

are low probably due to backgrounds. The Pass1 method-2 and J/ψ efficiencies are consistent and are high probably due to the inclusion of dimuon triggered events.

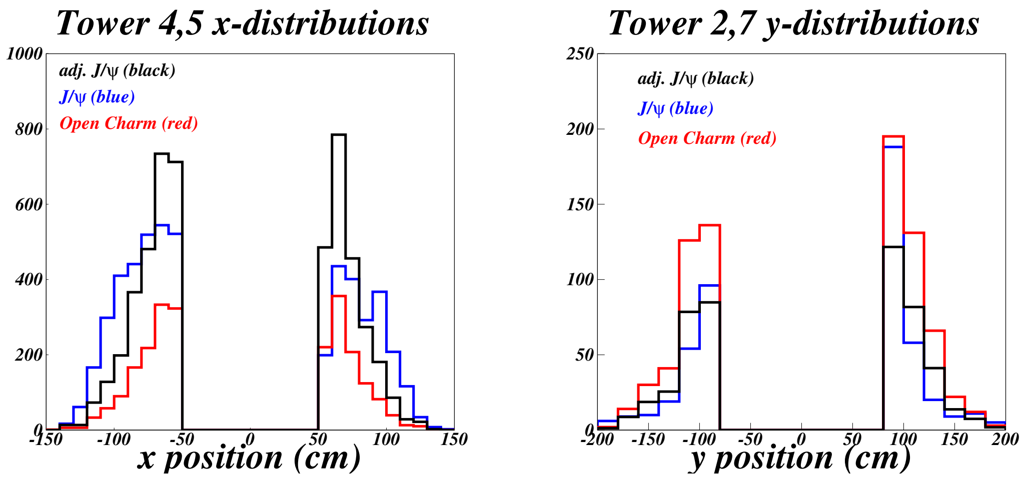
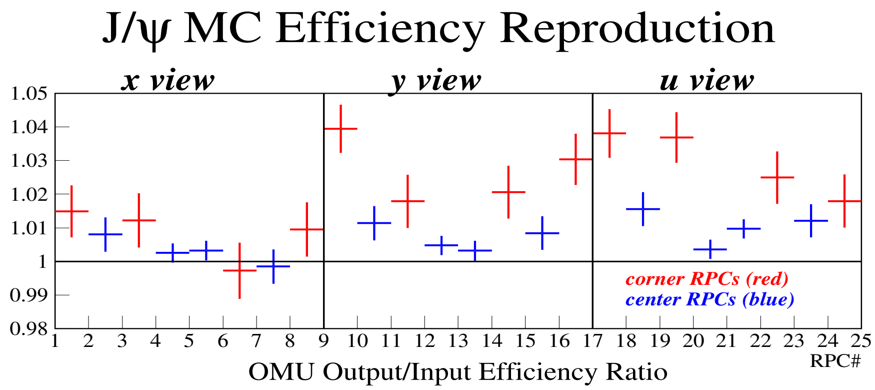


Figure 42: Radial distributions for J/ψ and open-charm muon tracks at the RPC's. Since the central tower efficiencies have a radial dependence, the differing radial distributions of open charm and J/ψ muons explain part of the efficiency difference for the two samples.

It follows that the average of the Pass1 methods represents the best measurement of RPC efficiencies for typical hadronic events. Systematic errors are obtained by comparing results with the range of efficiency settings.

4.2 Muon Misidentification

Particles other than muons are misidentified as muons in three ways. Firstly, a non-muon particle can "punch through" the thick wall of absorbing material (iron or steel) positioned upstream of the muon detector planes and register hits. Secondly, a particle that never reaches the detector can have a trajectory that points to a muon hit leading to a false match between the track and the muon hit. Thirdly, a hadron can decay in flight producing a muon. Inner muon misidentification from punch-through is reduced by requiring hits in at least 4 (out of 6) MH planes. A



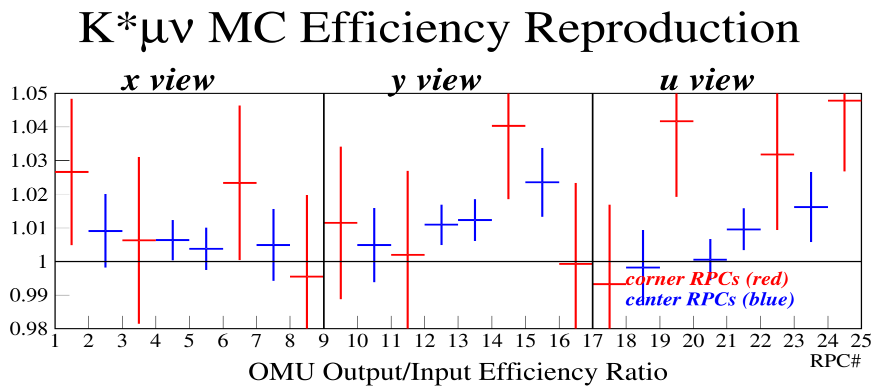
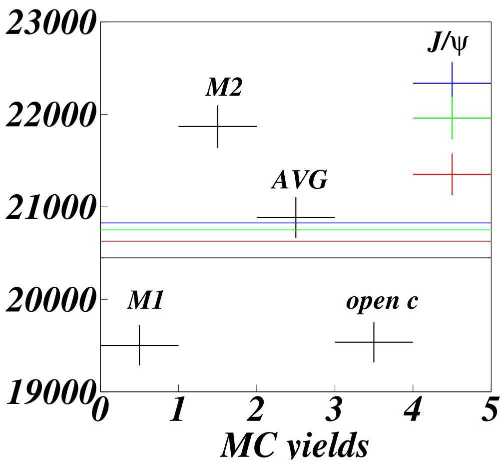


Figure 43: A test of the Monte Carlo efficiency simulation shows that the central RPC's reproduce the input efficiencies to $\sim$1% for both J/ψ and open-charm muons.

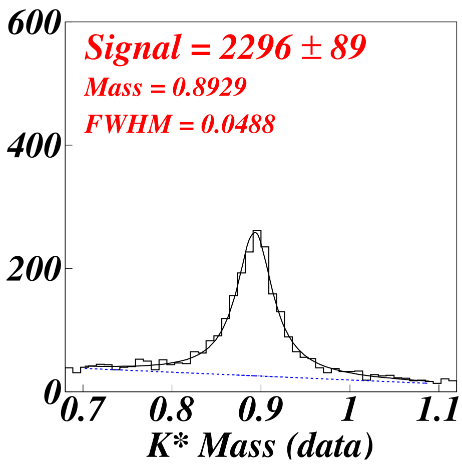
requirement that outer muons traverse more than 150 cm of steel reduces the outer muon punch through which occurs primarily along the edges of the M2 aperture.

4.2.1 Muon Halo

Detector noise and halo muons produce hits which can be randomly aligned with trajectories of particles absorbed by the steel. Halo muons originate from beam line interactions upstream of the experiment target and form a nearly uniform wall or "halo" of horizontal muons spread over the entire acceptance of the spectrometer.



(a) Monte Carlo yield variation



(b) Data yield

Figure 44: The effects of outer muon efficiency uncertainty is summarized here with K^*-tagged data vs a 10× Monte Carlo sample. Shown are the Monte Carlo yields resulting from efficiencies for: (1) Pass1 method-1, (2) Pass1 method-2, (3) the average of the Pass1 methods, (4) open-charm muons and finally (5) the from J/ψ muons for three methods. Listed in order of decreasing efficiency these three methods are: (i) including dimuon-triggered muons, (ii) including dimuon-triggered events with re-weighting to the open-charm radial distribution, and (iii) excluding dimuon-triggered muons. The average of the first four methods is shown in black. The average of the first four methods with one of the J/ψ methods is shown in the color corresponding to the J/ψ method. The data yield is shown for comparison of the statistical error to the systematic error resulting from uncertainty in outer muon efficiencies.

These muons will be scattered by the materials and magnetic fields of the beamline and spectrometer and so will have a wide distribution of slopes.

For this study, halo muons were identified in two ways. PWC tracks associated with a muon hit with a good confidence level provide one sample of muons. These "tracked" muons are limited in x and y to the acceptance of the PWC tracking chambers, and were required to be not linked with a microstrip track. This data sample was used to determine the halo muon momenta and slopes which were then incorporated into the simulation.

Muons identified with the clustering method (cf. Section 4.1.2) provide a second sample of muons. Since the outer muon detector acceptance is much larger than the PWC acceptance for horizontal tracks (Figure 34), the simulated halo muon production rate and spatial distributions were modeled using the muon cluster data. The incorporation of an accurate halo muon simulation into the Monte Carlo improves the muon background simulation and the muon misidentification level.

A random-trigger calibration run with the same configuration as a data run (beam stop out, analysis magnets on, etc.) was used to obtain a generic muon sample unbiased by the hadronic trigger, PWC aperture, or event selection criteria. Figure 45 shows the distribution of halo muon clusters. The upper plots show the distribution of random-trigger muons which follows a broad Gaussian distribution peaking 100 cm below the beam center. The initial Monte Carlo simulation, shown in the central plots, was modeled with a Gaussian which was too narrow in x and y. The lower plots show the Monte Carlo distributions matched to random-trigger data. The x and y distributions of muon clusters are shown separately in Figure 46 to more clearly illustrate the match of Monte Carlo to data.

Figure 47 to Figure 49 show the two different halo muon components. The first component is the flat "halo" of muons which originate far upstream of the production target. The second component consists of 260 GeV/c momentum muons that originated from the production target region and follow the the e^+e^- beam line. This second component of muons is probably the result of decays from hadrons produced in the production target or possibly Bethe-Heitler production of muons in the radiator.

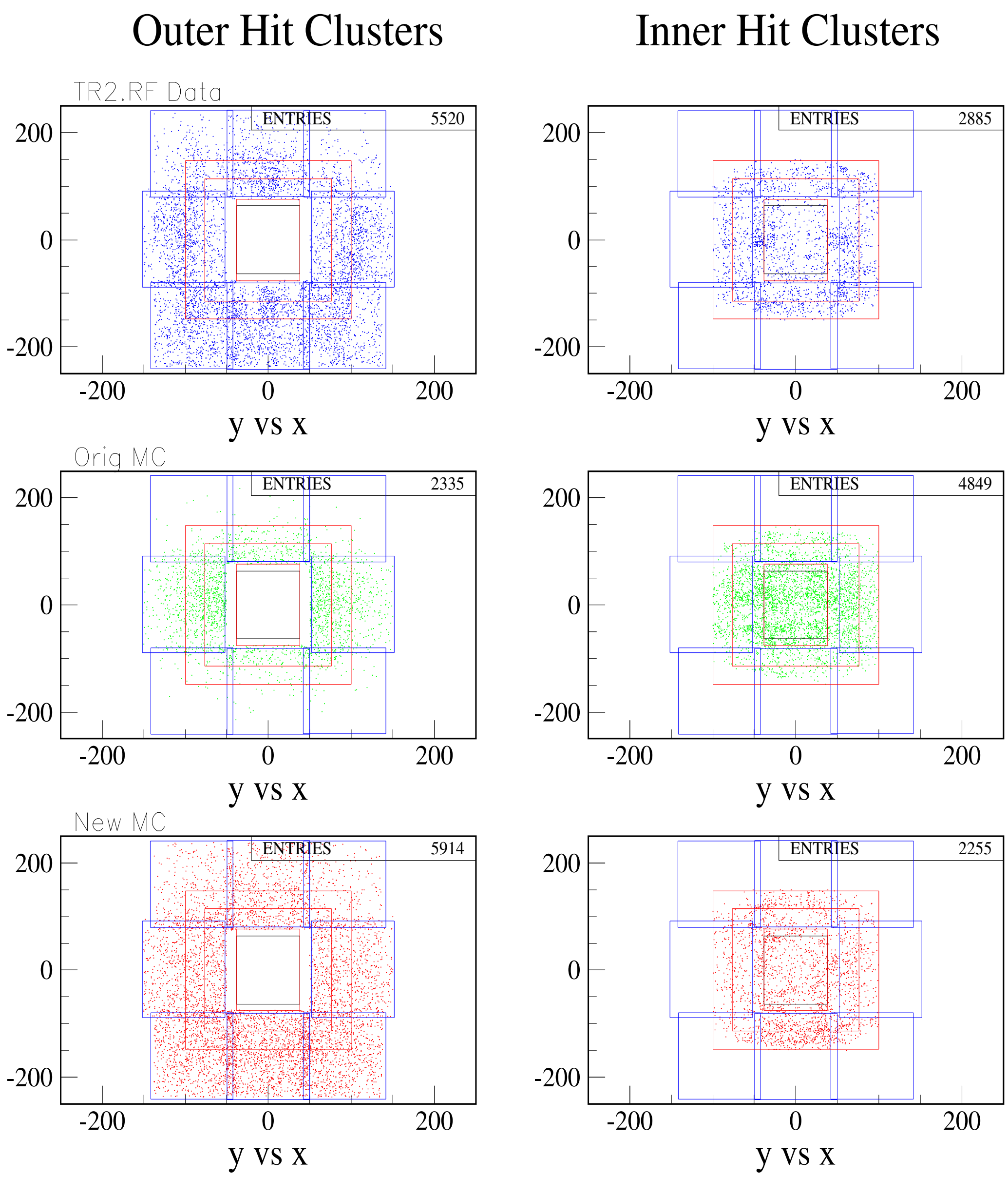


Figure 45: Muon hit clusters for inner and outer muons. Clusters are determined by looking at coinciding hits for all muon planes. Muon hit clusters from randomly triggered data are shown in the upper plots. The lower plots show clusters from muon Monte Carlo: from the simple original Monte Carlo model (middle) and from a model tuned to match the data (bottom).

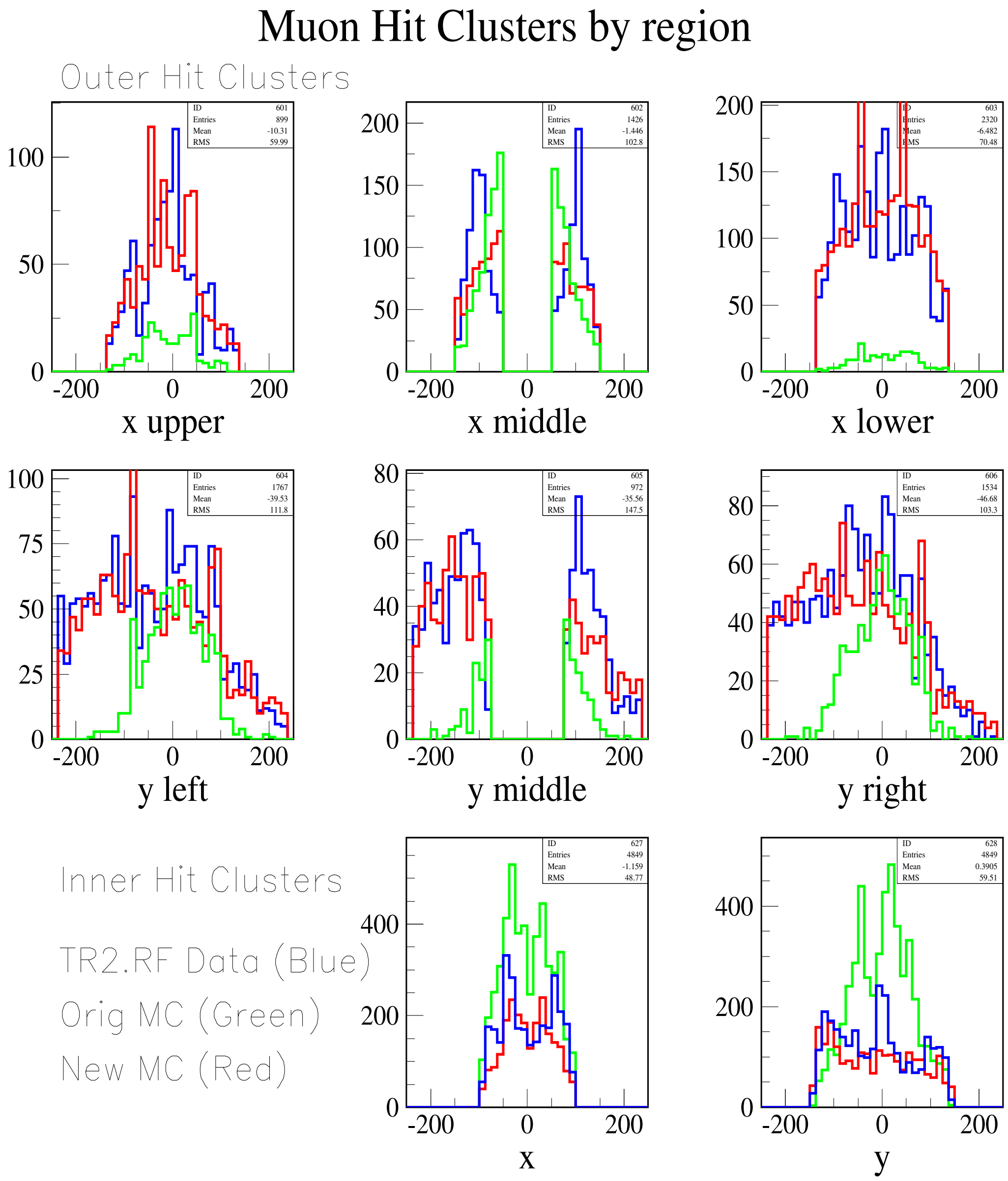


Figure 46: Muon hit clusters are shown as histograms in x or y (compare to the xy scatter plot of Figure 45). The plots clearly show the fidelity of the new muon halo simulation for various regions of the detector. The upper 3 plots show the distributions as a function of x for 3 swaths of the outer muon detector: an upper, central (spanning the M2 aperture), and lower swath. The middle 3 plots show the distributions as a function of y for 3 swaths of the outer muon detector. The lower 2 plots show the cluster distribution for the inner muon detector as a function of x and y (units in cm). Blue indicates randomly triggered data, green indicates the original Monte Carlo simulation, red indicates the new Monte Carlo simulation.

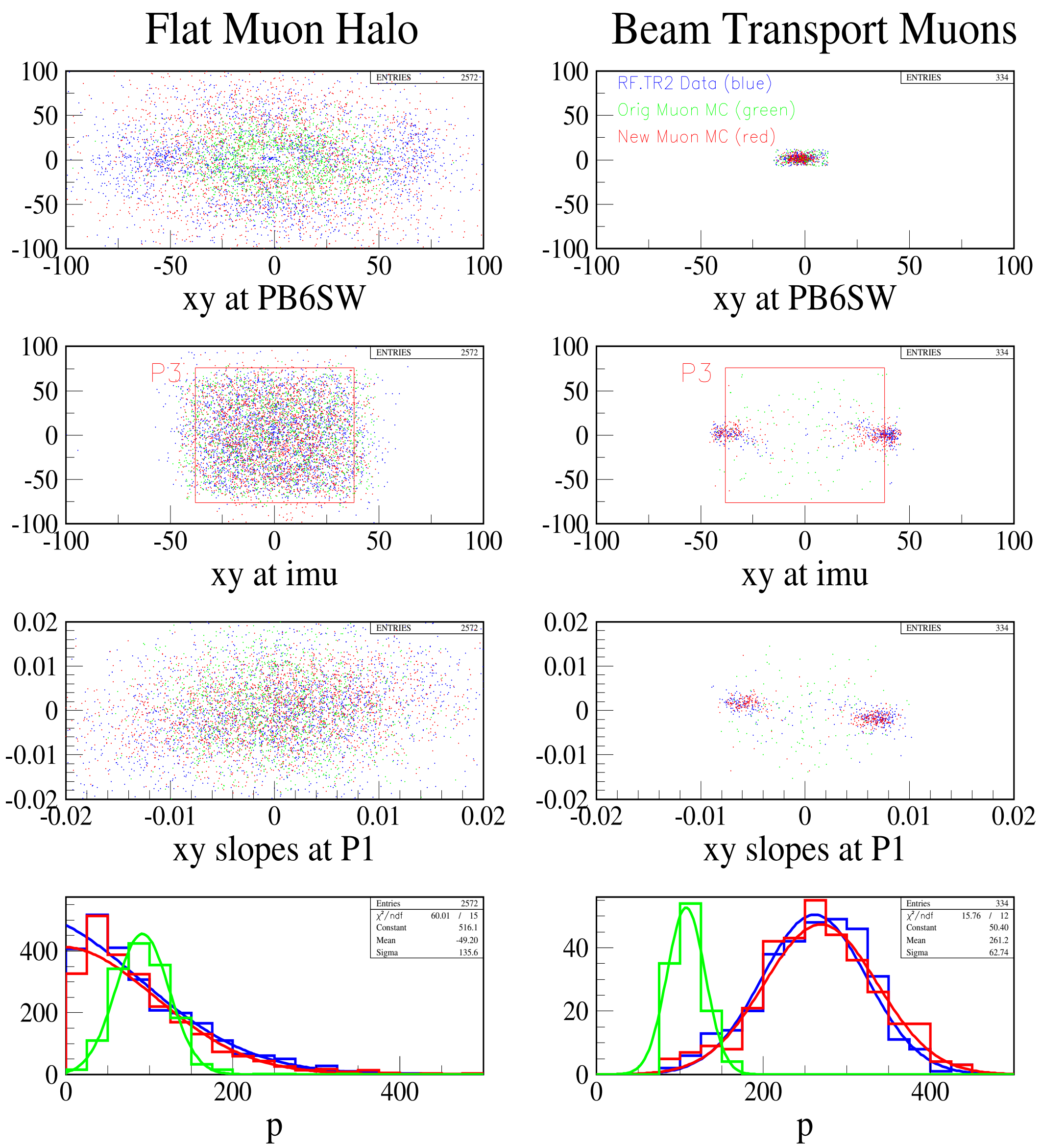


Figure 47: Scatter plots of muon halo tracks. The halo muons are present as a flat distribution and as two spots from muons transported along with the e^+e^- beam. The beam transport muons can be seen to originate from the beam sweeper magnet (PB6SW), have an average momentum of 260 GeV/c, and graze the edges of P3. Blue indicates random triggered data, green indicates original Monte Carlo, and red indicates the new Monte Carlo which incorporates both sources of halo muons.

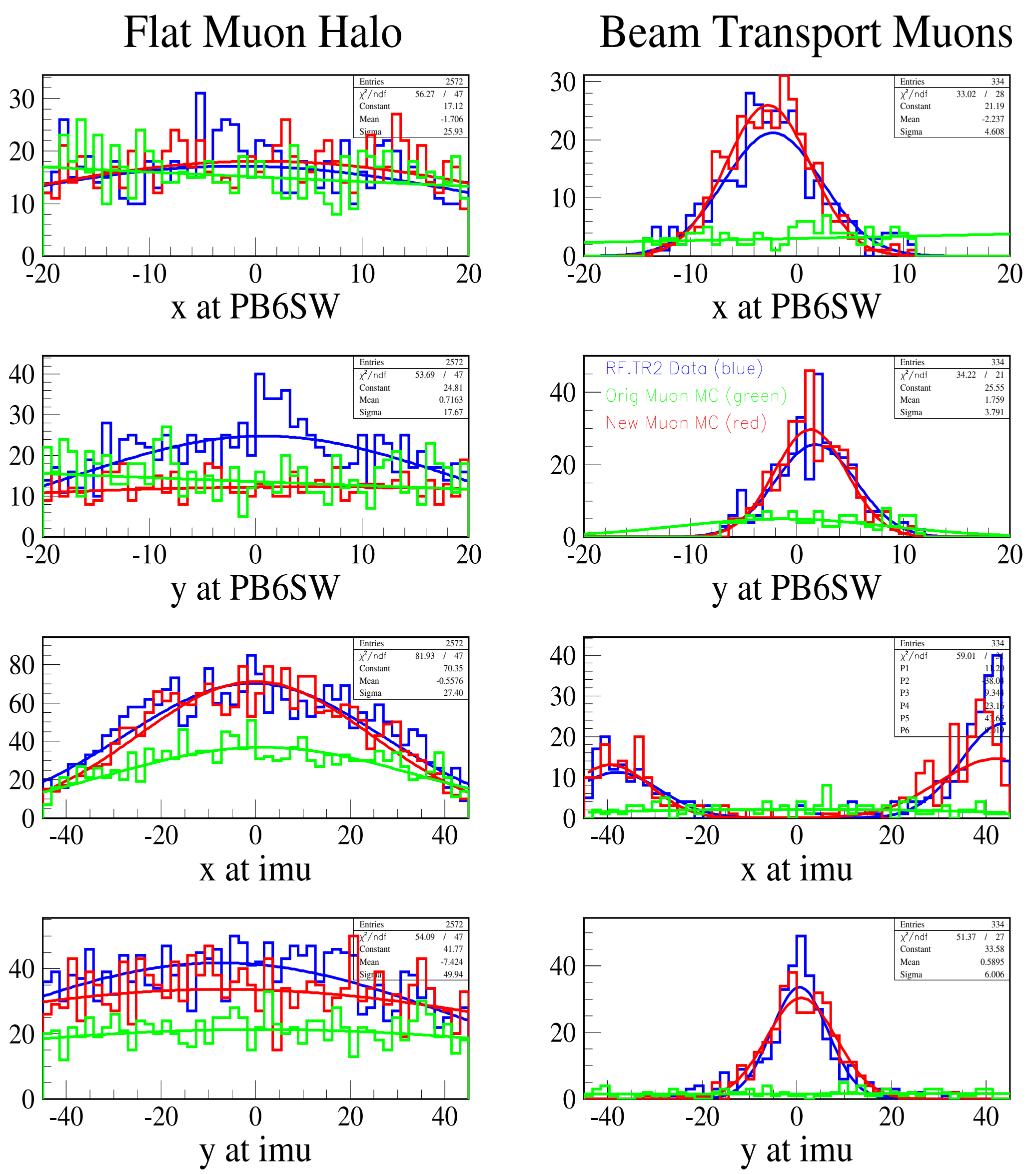


Figure 48: Muon halo tracks are shown as histograms in x and y (compare to the xy scatter plots of Figure 47 at the beam sweeper magnet (PB6SW) and at the face of the inner muon detector). Tracking parameters for halo muons (red) are matched to data (blue), with green showing the original Monte Carlo settings.

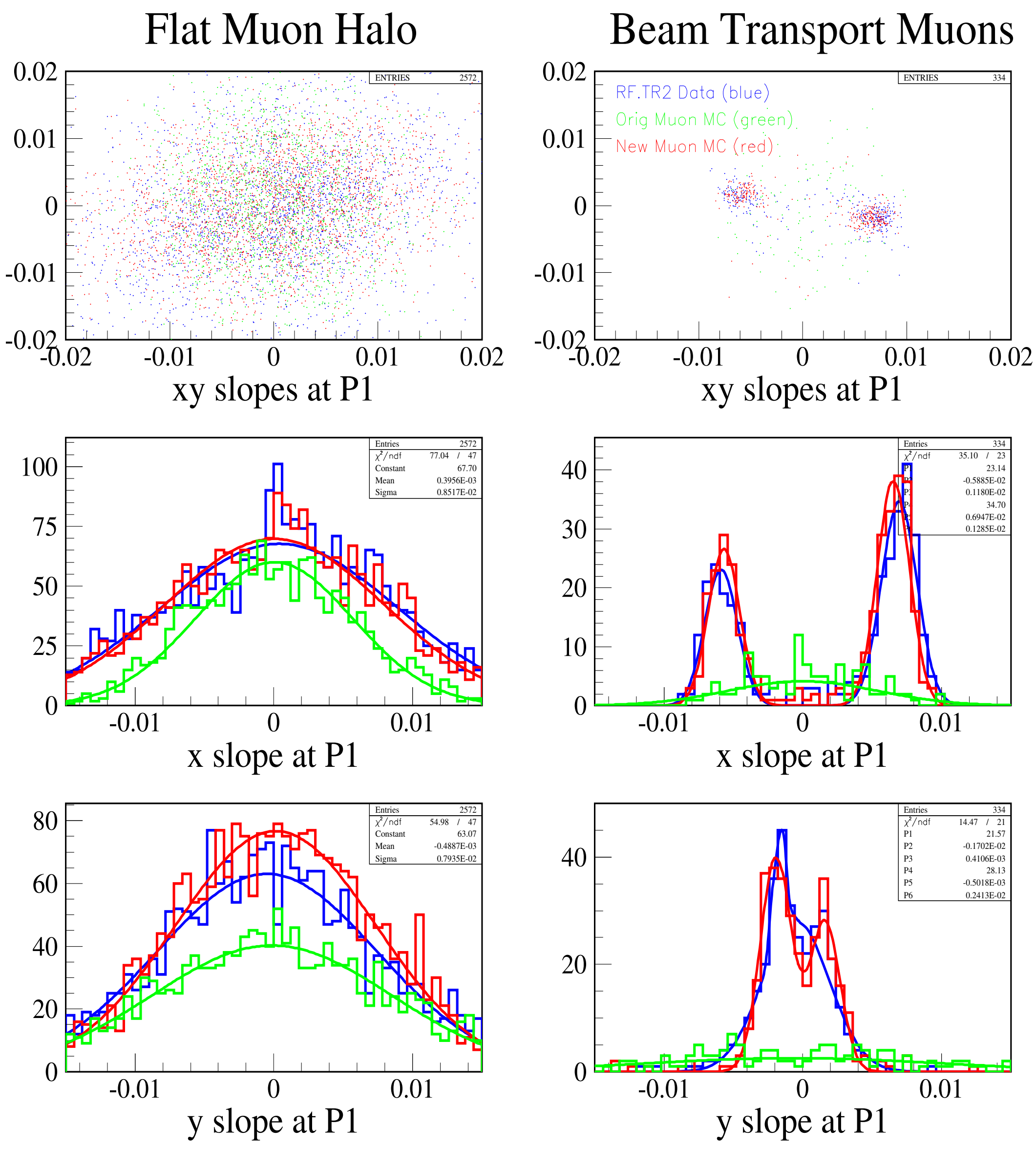


Figure 49: Muon halo slopes are shown as scatter plots and then in x and y separately. Monte Carlo (red) is matched to data (blue), with the original Monte Carlo settings in green.

Approximately 10% of the halo muons within the inner muon detector acceptance are the beam transport muons. Both components of the muon halo were incorporated into the muon halo simulation.

The upper plots of Figure 47 show the xy-distributions of halo muon tracks at the center of the sweeper magnets and at the face of the inner muon system. The track slopes and momenta are shown in the lower plots. The x and y distributions are plotted separately in Figure 48 and Figure 49 to show more clearly the match of the Monte Carlo simulation to data.

The halo muon distribution was also studied for reconstructed charm and K_s events by requiring the event have no tracked muons coming from the target region (i.e. not linked to an SSD microstrip track). With this event selection, the muon clustering and tracking was studied as for the random-trigger data. The data and Monte Carlo match well in all three cases.

To set the rate of halo muon production in Monte Carlo, muon cluster rates in data for three cases are considered: randomly triggered data, halo muons from charm events, and halo muons from K_s events. The number of muon hit clusters per event for data and Monte Carlo for various muon pile up settings (halo muon rate per RF bucket) are listed in Table 12. A muon pileup setting of 15% matches data.

4.2.2 Muon Misidentification Rates

The single muon misidentification rate was tested for this analysis using golden-mode charm decays. Charm decay modes with all daughter tracks being charged kaons or charged pions are the easiest to reconstruct and are termed "golden" decay modes. The combined golden-mode charm decays $D^0 \to K^-\pi^+$, $D^+ \to K^-\pi^+\pi^+$,

Table 12: Halo Muon Hit Rates

Random Data, Clusters/event		
Pilemu Setting	inner clusters per event	outer/inner clusters
Data	.03	1.9
0.15	.04	5.4
0.30	.09	2.9
0.50	.14	3.0
$K_s \to \pi\pi$, Clusters/event		
Pilemu Setting	inner clusters per event	outer/inner clusters
data	.07	1.9
0.15	.08	1.6
0.30	.12	2.0
0.50	.16	2.3
$D^+ \to K\pi\pi$, Clusters/event		
Pilemu Setting	inner clusters per event	outer/inner clusters
data	.08	1.8
0.15	.07	1.5
0.30	.11	2.1
0.50	.16	2.4

and $D^0 \to K^-\pi^+\pi^+\pi^-$ provides a large sample of pions from charm. A subset of the golden mode sample, $D^+ \to K^-\pi^+\pi^+$ ($K\pi\pi$), provide the largest sample of pions most representative for the charm decays in this analysis.

Particle momentum is determined by measuring the track bend (kick) induced by the analysis magnets. Therefore a pion decaying after M2 will create a hit in the muon detector, but have the correct momentum for reconstructing the parent mass. Muons from pion decays before M2 will have lower momentum (the missing momentum carried by the neutrino) and will then be bent more by M2. Under a

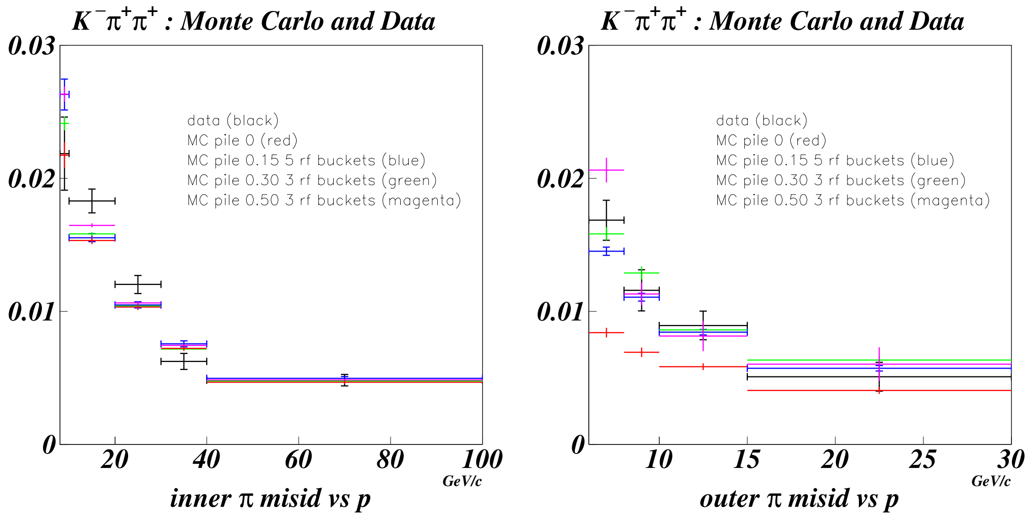


Figure 50: $K\pi\pi$ single muon misid as a function momentum. The effect of different PILEMU and RFOMU settings is compared.

pion hypothesis, this missing momentum will make the parent mass appear lower, distorting the low mass end of the signal shape.

To test the level of misidentification, the $K\pi\pi$ signal was used with single misid including decays in flight. The signal and background shapes were determined from CCBAR Monte Carlo and fit to data. Mode-specific Monte Carlo was used to test the misidentification rates for different PILEMU and RFOMU settings. The PILEMU setting was set by an independent factor — the muon clustering rates discussed in the previous section — but the misidentification sensitivity to the different settings was tested as well.

The outer muon misidentification rate was inflated by adding four out of time RF buckets to the outer muon Monte Carlo simulation. This increases the length of time for the OMU simulation for each event thereby increasing the halo muon hit rate for the outer muon detector and increasing the outer muon misidentification rate. Figure 50 shows the misidentification rate vs momentum. The total single misidentification in Monte Carlo is 7% lower than data for inner tracks and 30% lower than data for outer tracks. For $D^+ \to K^-\pi^+\pi^+$, the total single misidentification rate in Monte Carlo is 14% lower than data.

4.2.3 Muon Double Misidentification

The double muon misidentification rate may not simply be the square of the single misidentification rate. Closely spaced tracks can match to the same hits with a high confidence level, increasing the double misidentification rate. $K^-\pi^+\pi^+$ events can be used for single misidentification, but $K^-\pi^+\pi^+$ double misidentification events are nearly indistinguishable from $K^-\mu^+\mu^+$ rare events, so $K\pi\pi$ cannot be used for double misidentification calibration.

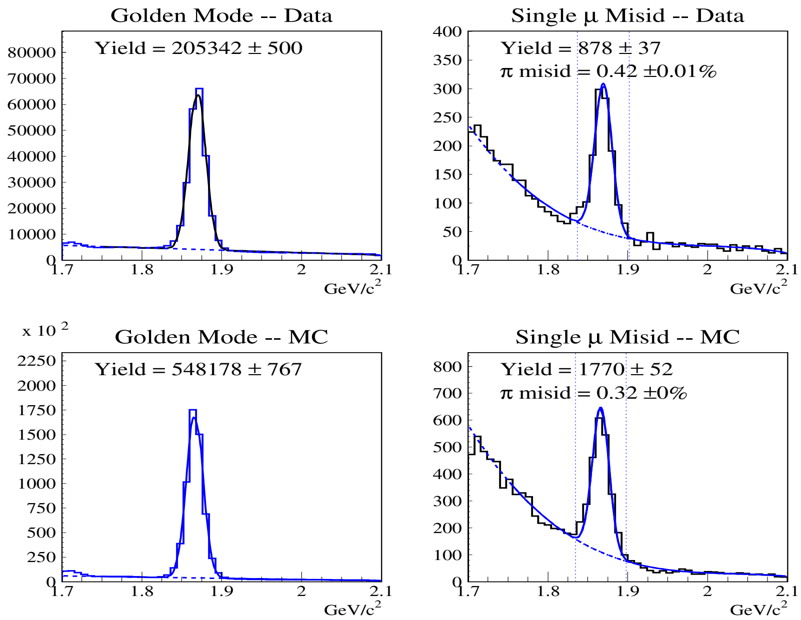
Since double misid effects can arise from a variety of factors, it is desireable to reduce the number of factors that can vary the level of misid. The primary complication in double misid events are the correlations that may arise from closely spaced tracks. This problem is reduced by reducing the effect of decays in flight and by imposing a muon separation requirement. A requirement the muons from two-muon events be clearly separate is employed. This separation requirement is achieved by requiring the two muon tracks to have separate muon hits.

In principal, removing the decay-in-flight events requiring that muon tracks be clearly separate from each other greatly simplifies the number of factors the Monte Carlo must model correctly for the double misid effect to behave in a manner similar to the single misid effects. With these muon event simplifications we can estimate the effect of double misid uncertainty using single misid information.

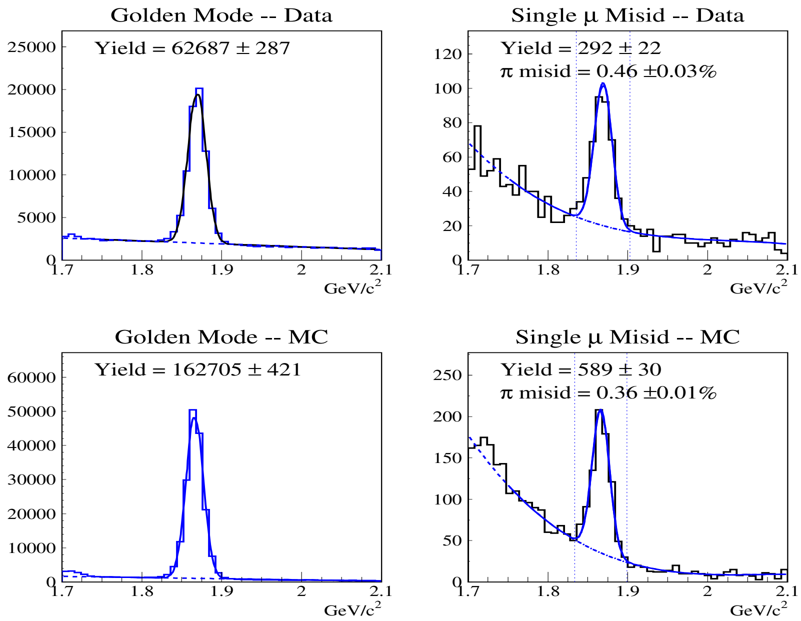
TRKFIT momentum and confidence levels were used to reduce decays in flight. The muon hit separation cuts significantly reduced overall level of misid. These additional cuts significantly reduce backgrounds, but we need to re-measure the match of single misid to CCBAR Monte Carlo.

Figure 52 shows the single misid behavior for inner and outer tracks with the new cuts for $K\pi\pi$. Figure 51 shows the single misid behavior for inner and outer tracks with the new cuts for combined golden-mode decays. The $D^+ \to K^-\pi^+\pi^+$ mass plot is fit to a linear background and a Gaussian signal. The $D^+ \to K^-\mu^+\pi^+$ mass plot is fit to a third-order polynomial background and a Gaussian signal with the mean and width fixed to the values of the previous plot. This ratio of yields for the two plots gives the rate of a single pion misidentified.

The overall single misidentification rate in Monte Carlo is ~20% low which implies a double misidentification rate ~40% low. The muon misidentification discrepancy in Monte Carlo can be tested by inflating the $D^+ \to K^-\pi^+\pi^+$ component of ccbar Monte Carlo 40%. This Monte Carlo sample is used to test the uncertainty in the misid rate. Results with the default Monte Carlo sample are compared with the $K\pi\pi$-inflated Monte Carlo sample and the most conservative result is used.

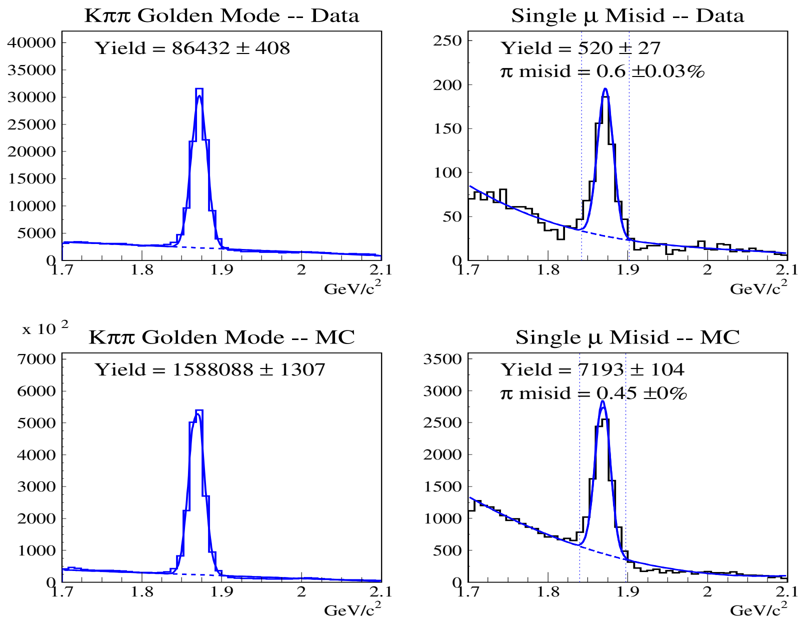


(a) Inner Muon Misidentification

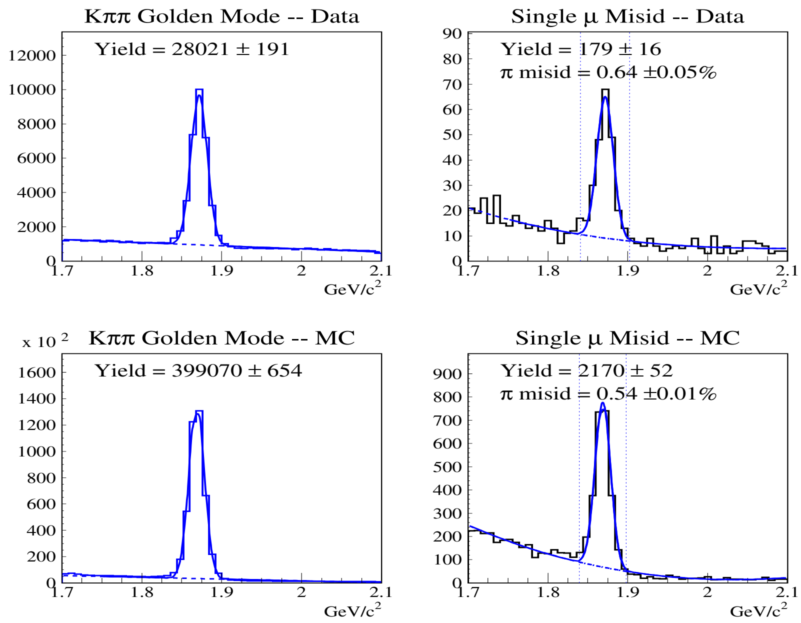


(b) Outer Muon Misidentification

Figure 51: Combined Golden Mode Muon Misidentification. Comparisons of data to CCBAR Monte Carlo Show single misidentification to be approximately 23% lower in Monte Carlo than data.



(a) Inner Muon Misidentification



(b) Outer Muon Misidentification

Figure 52: $K\pi\pi$ Golden Mode Muon Misidentification. Comparisons of data to CCBAR Monte Carlo Show single misidentification to be approximately 23% lower in Monte Carlo than data.

CHAPTER V

SELECTION OF RARE AND FORBIDDEN CANDIDATES

5.1 **Rare Decay Data Skim**

From the large sample of charm decays collected by FOCUS, events that pass the requisite conditions for rare decays are "skimmed" from the sample as a whole and collected into a smaller sample of candidate events. The general data reconstruction methods are described in Chapter III. This chapter details the analysis method used specifically for the selection of rare and forbidden decays.

5.1.1 EZDEE

EZDEE (Easy-D) is the FOCUS software package used to select charm events. EZDEE uses DVERT vertexing to find D candidates with the correct topology. EZDEE then applies the user-defined selection criteria on tracking, vertexing, particle ID, mass, etc. and calculates a standard set of analysis variables for each event. In the data skims, EZDEE variables were used to make skim decisions, but the standard reconstructed data blocks (not the EZDEE variables) were saved in the output streams. In the final user analysis, EZDEE was used to identify events, and the EZDEE analysis variables, along with additional user defined variables, were saved in an HBOOK ntuple format [37].

Table 13: Skim Cuts

Variable	Skim1 Sezskm cuts
muons	$CL \geq 0.0001$
kaons	$\Delta W_K = W_\pi - W_K > 0.5$
vertexing	DCL > 1%
mass (GeV/c^2)	1.7 < mass < 2.1

5.1.2 Skim Results

As described in Section 3.5, the FOCUS data were reconstructed and separated into several data streams. The FSFE tapes contained the rare-decay candidates and related normalizing modes. The cuts applied FSFE rare and forbidden decays are summarized in Table 13.

5.2 **Final Event Selection**

For the final event selection, data were skimmed from the FSFE tapes, again using EZDEE. Seven secondary vertex topologies were specified: the normalizing modes $K^-\pi^+\pi^+$, $K^-K^+\pi^+$, and the rare decay modes $K^-\mu^+\mu^+$, $K^+\mu^+\mu^-$, $\pi^-\mu^+\mu^+$, $\pi^+\mu^+\mu^-$, and $\mu^-\mu^+\mu^+$.

5.2.1 Cut Variables

The rare decay channel is determined relative to a well-known decay channel (normalizing mode). Ideally the analysis of the rare decay channel should be done exactly as for the normalizing mode. A comparison of the two signals then normalizes the rare signal so that an absolute branching ratio can be determined.

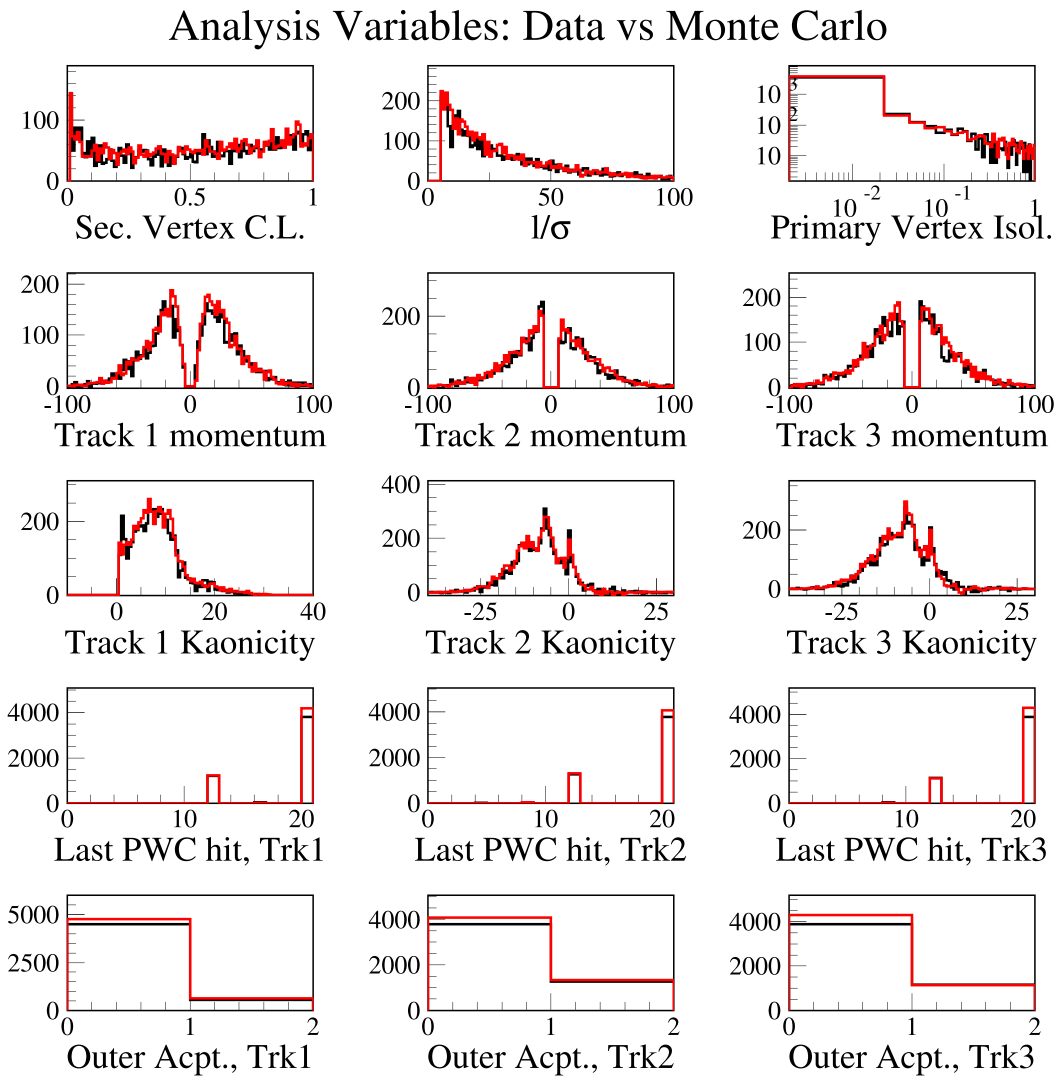


Figure 53: Sideband subtracted analysis variables: Golden mode ($K\pi\pi$) from Data (black) and CCBAR Monte Carlo (red).

An analysis cut used to separate signal from background will cut out signal as well as background. The amount of signal that is preserved after a cut is the efficiency for that cut. Identical vertexing cuts are applied to the rare decay modes and the normalizing modes to make the efficiencies nearly equal. The efficiencies corrections for these cuts will cancel in the ratio of rare signal to normalizing signal so they do not need to be determined. These cuts included acceptance, momentum, tracking, and vertexing cuts. Particle ID cuts cannot be applied symmetrically to the normalizing modes so the absolute efficiencies of those cuts need to be studied. A summary of the analysis variables follows:

Acceptance Cuts. The muon confidence level cut has an implicit muon detector acceptance cut which does not necessarily apply to the normalizing mode. For tracks in the normalizing mode that correspond to tracks in the rare decay mode, an explicit inner or outer muon acceptance requirement was applied.

Momentum Cuts. Since muon misidentification increases at low momentum, minimum momentum cuts were applied to muon tracks for rare decays and for the corresponding tracks in the normalizing modes.

Vertexing cuts. The detachment between the primary and secondary vertex is defined as the distance between the vertices divided by the error. For this variable, termed ℓ/σ_ℓ, a minimum cut of 5 was applied in Skim-2. The confidence levels of the primary and secondary vertex fits were required to be greater than 1%.

The primary vertex was required to be isolated from the secondary vertex daughter tracks using the primary isolation variable, defined as the best vertexing confidence level achieved when each of the secondary daughter tracks is included, one at a time, in the primary vertex. A primary isolation confidence level (ISO1) less than 1% removes the D^* backgrounds.

Secondary vertex isolation cuts measure the possibility that other tracks in the event originate in the secondary vertex. The secondary isolation variables were defined in two ways. The first method (ISO2) calculated the best secondary vertex confidence level achieved with each unused track (excluding primary tracks) included, one at a

Table 14: Analysis Cuts – Cut Grid

Variable	Loosest to Hardest
ℓ/σ_ℓ	>5, 7, ..., 21
Primary Isol.	< 0.1, 0.03, 0.01, 0.003, 0.001
Secondary CL	> 1%, 2%, 4%
Kaonicity	> 0.5, 1.0, 2.0
Pionicity	> -15, -3, -1
Muon CL	> 1%, 5%, 10%
inner mom. (GeV)	> 8, 9, 10, 11, 12, 14
outer mom. (GeV)	> 6, 7, 8, 9, 10

time, in the secondary vertex; the second method (ISO3) repeated the procedure of the first, but considered *all* non-secondary tracks (including primary tracks). The secondary isolation variable (excluding primary tracks) is required to be less than the specified confidence level.

The routine TRKFIT re-fits tracks as described in Section 3.3. TRKFIT calculates improved track momentum and returns a confidence level for the track. A TRKFIT confidence level greater than 0.01% was required for all tracks. This reduces the number of hadrons that decayed in flight included in the final sample.

Particle ID Cuts. The threshold Čerenkov detectors (cf. Section 2.6) are used to distinguish kaons from pions. A "kaonicity" variable was defined as the negative log likelihood of a particle being a pion minus the negative log likelihood of a particle being a kaon. A minimum kaonicity of 0.5 is required for kaon candidates ensuring the kaon hypothesis is favored over the pion hypothesis. "Pionicity" is defined to be the negative of Kaonicity. A pionicity cut is applied to the pion in $\pi\mu\mu$ modes.

A minimum 1% muon confidence level cut is applied to muon candidates along with the requirement that the muon is predicted to have traversed enough material. For outer muons this cut requires the muon traverse more than 150 cm of material and have hits in at least 2 of 3 planes. For inner muons this cut requires at least 4 of 6 planes have a hit correlated with the muon track.

The simultaneous identification of two muons is often the result of two tracks pointing to the same hits within a detector. To eliminate these events, a simple muon separation variable was defined for this analysis. Muons were defined to be separated if they did not share too many hits. The first step in outer muon identification is grouping RPC hits into clusters, and then matching tracks to clusters. The FOCUS outer muon identification algorithm recorded the muon confidence level for each track and the RPC cluster matched to that track. Outer muon separation was achieved by requiring that two outer muons be matched to separate RPC clusters. The FOCUS inner muon identification algorithm recorded the muon confidence level for each track as well as the hits matched to that track. The hit for each MH plane that was matched to the track was recorded. Inner muon separation was achieved by requiring that no more than two hits were shared between two muon tracks. The two-hit requirement allows parallel muons that are well separated.

In addition to the muon separation variables just described, muon isolation variables also were used in the analysis. Muon isolation for a given track is determined by considering the muon confidence level of other tracks. This is done by looking at the set of muon hits for a given muon candidate. All other tracks that are correlated with this same set of muon hits are considered. The highest muon confidence level of these tracks defines the muon isolation. Muon isolation is applied by requiring the value to be less than a specified cut.

Reflection Cuts. For the D_s signal, a cut was applied to reduce the reflection occurring when one of the pions from a $D^+ \to K^-\pi^+\pi^+$ decay is misidentified as a kaon. First, the kaon track (in $K^+K^-\pi^+$) with the same charge as the pion is switched to a pion hypothesis. The parent mass for the three tracks is then recalculated under the new hypothesis. If the recalculated parent mass is within 3σ of the D^+ mass, then a very strong kaonicity cut (kaonicity > 6) was applied to the kaon in question.

Cut Grid. A fixed set of minimum cuts was applied in all cases. Additional cuts were applied with varying strengths to find the cut values for optimal signal to noise (optimal experimental sensitivity). The range of cut values for these variable cuts are summarized in Table 14. The variable cuts are applied in all possible combinations producing a cut "grid" of thousands of cut combinations.

5.2.2 Backgrounds

The invariant mass (using TRKFIT momentum) for each of the rare decays is shown in Figure 54 with the loosest cuts and in Figure 55 with more optimal cuts. All of the backgrounds are approximately linear with the exception of $K^-\mu^+\mu^+$. The plots show the behavior of the data with the range of cuts applied. The data points are shown with error bars and the signal region is masked out.

As shown in Figure 54, a Gaussian signal plus linear background is fit to the normalizing modes while a ccbar Monte Carlo shape is fit to the rare decay sidebands. The normalizing mode backgrounds are linear for both loose and tight cuts, and sideband subtraction is used to obtain the normalizing mode data yield. The rare

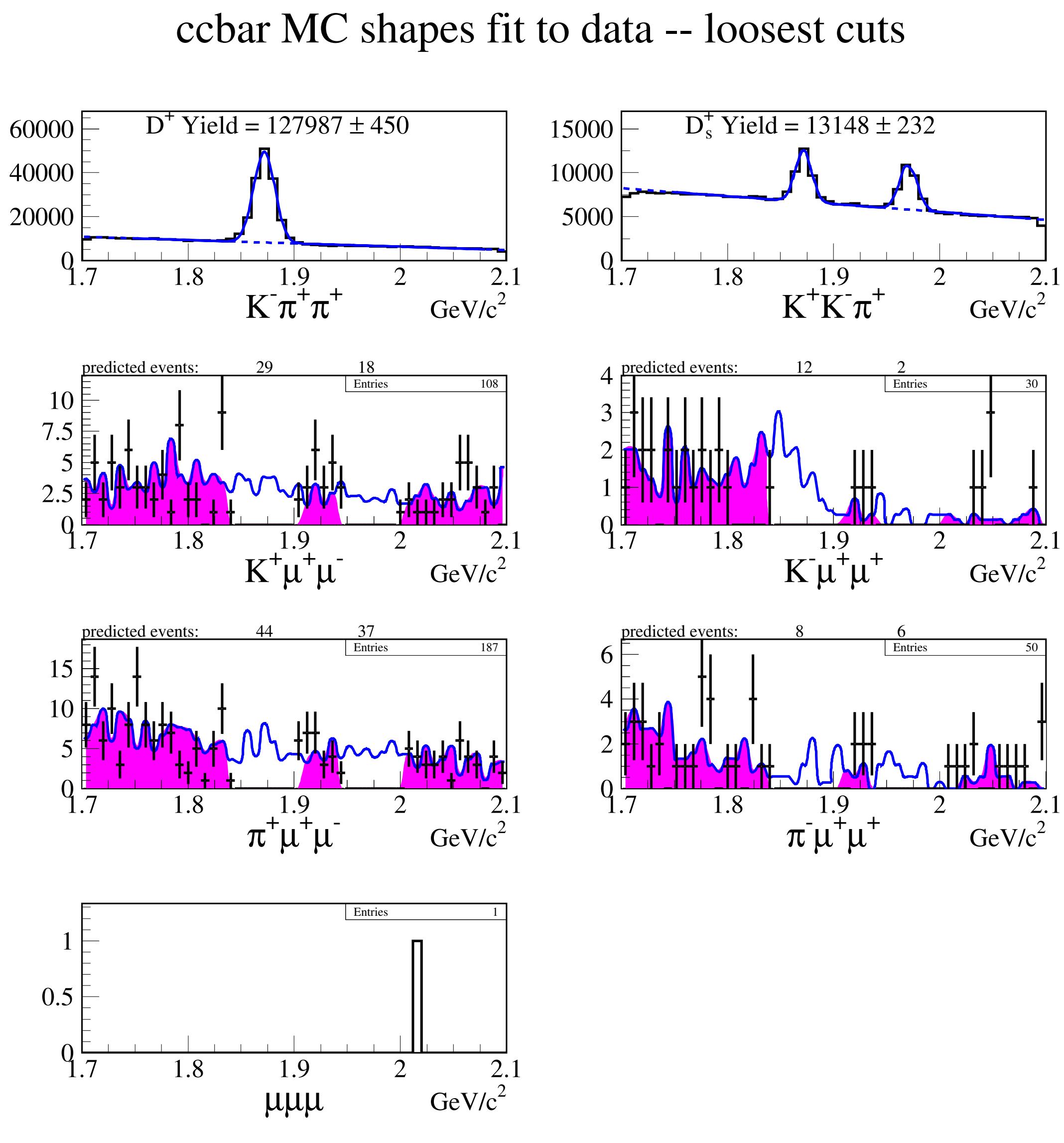


Figure 54: Mass plots of normalizing modes and rare and forbidden decay modes with the loosest cuts applied. The normalizing modes are shown with a Gaussian signal plus linear background. The rare modes are shown with events in the signal regions masked for both the D^+ and D_s^+ decay modes. The rare mode data is fit with the ccbar Monte Carlo shapes to data in the shaded regions. The unshaded regions show the signal predicted by Monte Carlo.

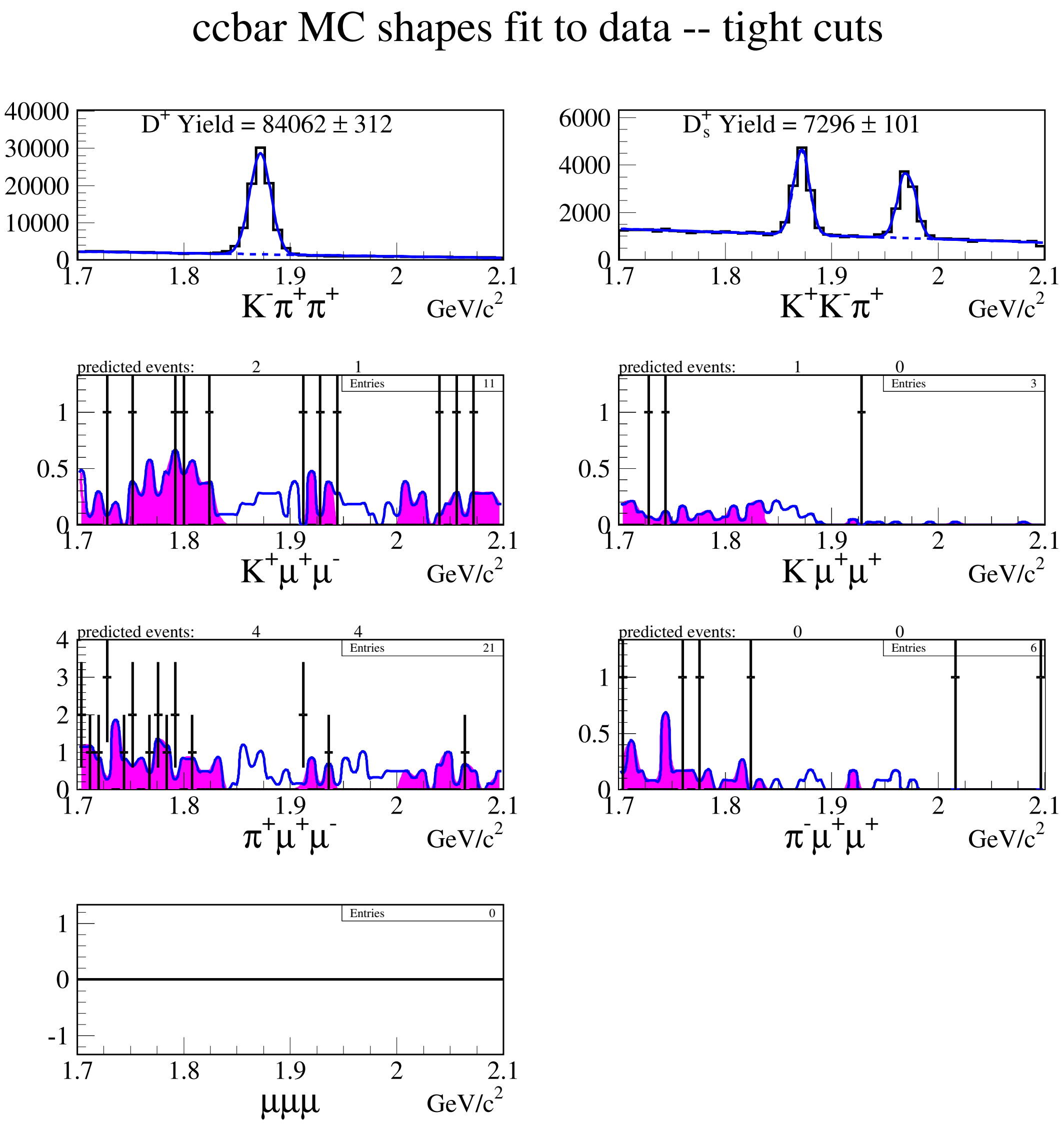


Figure 55: Mass plots of normalizing modes and rare and forbidden decay modes with tight cuts applied. The plots are presented as in Figure 54. This and the preceding figure illustrate the effect of the range of cuts used in the cut grid. Thousands of cut combinations are applied with the cut grid, and the branching ratio sensitivities are determined by counting signal and sideband events without any fitting (as done here in this example for illustration purposes only.)

decay events are cut out completely with the very hardest cut combinations (not shown). These plots illustrate changes in the data that occur during the sensitivity optimization process, discussed in the following sections.

5.3 Sensitivity Optimization

The calculation of the rare decay branching ratio is straightforward,

$$\frac{\mathcal{B}r_{rare}}{\mathcal{B}r_{norm}} = \frac{Y_{rare}}{Y_{norm}} \times \epsilon \tag{38}$$

where ϵ is the relative efficiency, $\mathcal{B}r_{rare}$ is the 90% confidence interval upper limit on the rare decay branching fraction, $\mathcal{B}r_{norm}$ is the current PDG branching fraction value for the normalizing decay mode, Y_{norm}, the normalizing signal yield, comes solely from sideband subtracted data, and Y_{rare} is the 90% confidence interval upper limit on the rare signal yield (described in the next section). The relative efficiency of the normalizing mode and decay mode is the ratio of mode-specific Monte Carlo reconstructed yields (sideband subtracted) for the two modes (each run with the same number of events). (The branching ratio for each mode was set to zero in mode-specific Monte Carlo. This ensured that the specified decay came from only one of the charm decays and was excluded from the opposite sign charm decays.) So five different event samples are used to calculate the branching ratio: (i) rare data, (ii) normalizing data, (iii) ccbar Monte Carlo, (iv) normalizing mode Monte Carlo, and (v) rare decay Monte Carlo. The cuts are applied equally to all five samples.

The experimental sensitivity is calculated the same way as the branching ratio, except that a background prediction for the rare signal is used instead of the actual signal, which is described in the following section.

5.3.1 Confidence Intervals for small signals

For Poisson processes with background, a 90% confidence level upper limit on the result generally is determined. Feldman and Cousins [38] have developed a method, based on the Neyman [39] construction, with an additional likelihood ratio test, that has become the current standard. However, several variations on this method have subsequently been proposed [40, 41]. These methods typically neglect the effects of background uncertainty. When using these methods, background uncertainty can be incorporated into an upper limit as a *systematic* error using a method proposed by Cousins and Highland [42]. However, Rolke and Lopez have developed an improved method for constructing confidence intervals which incorporates the *statistical* uncertainty of the background into the construction of Poisson confidence intervals [43].

The discrete nature of Poisson statistics means that Poisson confidence intervals cannot be determined exactly. A 90% confidence interval is constructed to cover the result *at least* 90% of the time. With this criterion, confidence intervals generally over-cover (i.e. are conservative). It is desirable to minimize this over-coverage. The Rolke-Lopez method provides more accurate coverage of Poisson confidence intervals than the Feldman-Cousins method.

For Poisson processes with background, the probability of observing a signal x for a signal mean μ and a known background b is given by:

$$P_{\mu,b}(x) = \frac{(\mu + b)^x}{x!}\, e^{-(\mu+b)} \tag{39}$$

The procedure for the Neyman construction of confidence intervals is as follows: (i) For a given μ, b, $P_{\mu,b}(x)$ is calculated for $x = 0, 1, 2, ...$ (ii) each x is then ordered — and the ordering criteria varies for different methods. (iii) the confidence interval for

μ is constructed by summing $P_{\mu,b}(x)$ for each value of x in decreasing order of rank until the total probability is greater than or equal to the desired confidence level α (i.e. $\alpha = 90\%$ is typically the standard used).

The Feldman-Cousins method uses the Neyman construction with an ordering principle based on likelihood ratios. The rank $R_{\mu,b}(x)$ for each x is defined as the the ratio of the likelihood $P_{\mu,b}(x)$ to the maximum likelihood for that value of x,

$$R_{\mu,b}(x) = \frac{P_{\mu,b}(x)}{max\,\{\,P_{0,b}(x),\ P_{1,b}(x),\ P_{2,b}(x),\ ...\,\}}\ . \tag{40}$$

The probabilities for each x are summed, in decreasing order of rank, R, until the total probability is greater than or equal to the desired confidence level, α. The α confidence interval for μ is obtained from the μ values that correspond to endpoints of the range of x used in summation of probabilities.

The Feldman-Cousins method assumes that the signal background is known without error. In practice, the signal background is derived from sideband data with either a linear model or a Monte Carlo background shape. However, sideband data for small signals usually has significant statistical uncertainty even when the sidebands are much larger than the signal region. The Rolke-Lopez method incorporates the statistical uncertainty of the background rate into its confidence intervals. The background rate of the signal region, b, is extrapolated from the data sidebands. τ, defined as the ratio of sideband to signal data:

$$\tau \equiv \frac{Monte\ Carlo\ sideband}{Monte\ Carlo\ signal}. \tag{41}$$

is used to determine the mean background rate,. The uncertainty in τ is assumed to be small in the Rolke-Lopez formulation; τ can be defined using a linear assumption or a high-statistics Monte Carlo shape.

The statistical uncertainty in both the signal and sideband regions results in a two-dimensional Poisson function. The probability of observing x events in the signal region and y events in the sidebands, for a signal mean μ and a signal background mean, b, and a sideband to signal background ratio τ is given by:

$$P_{\mu,b,\tau}(x,y) = \frac{(\mu+b)^x}{x!}\, e^{-(\mu+b)} \cdot \frac{(\tau b)^y}{y!}\, e^{-(\tau b)} \tag{42}$$

The determination of confidence intervals for the average signal rate μ is done in two ways. The first method uses a two dimensional confidence region in x, b, constructed with an ordering procedure which uses simply $P_{\mu,b,\tau}(x,y)$ as the ordering principle (and not a likelihood ratio test). The second method uses a large sample approximation to the likelihood ratio test statistic. This yields a continuous profile likelihood function approximating a chi-square distribution, and this function can be used to construct confidence intervals for μ. For the final determination of the confidence intervals for μ, the profile likelihood method is used except when the number of observations in the signal region is small compared to the number of background events, in which case the confidence region method is used. The method is described in detail in Reference 43.

For a given τ with y sideband events, the experimental sensitivity, $S_\tau(y)$, for the 90% confidence upper limit of the signal mean is

$$S_\tau(y) = \sum_{x=0}^{\infty} U_\tau(x,y) \cdot P_y(x) \tag{43}$$

where $U_\tau(x,y)$ is the Rolke-Lopez 90% confidence interval upper limit of the signal mean and $P_y(x)$ is the Poisson probability of x given a mean y. Using code provided by Wolfgang Rolke, the confidence intervals and sensitivities were calculated for values of $y = 1$ to 30, $x = 1$ to 130, and $\tau = 0.5$ to 20. The observed sideband events y, and the observed signal events, x, are integer values, but τ may be a fractional value and is calculated in steps of 0.5 for the tables. Since τ can be any fractional value, the values calculated from the Rolke-Lopez tables are linearly interpolated.

5.3.2 Blind Analysis

The search for rare signals involves selecting optimal cuts on the data. Cuts that optimize the rare decay signal relative to the normalizing signal while maximizing efficiency are considered to be optimal. However, the selection of optimal cut values for low statistics data could be biased by statistical fluctuations or the user's desire for a good result. The blind analysis method proposed by Schwartz [44] is designed to remove potential user bias by disallowing the user any information on the signal region until the final cuts have been selected. Low statistics data may fluctuate widely and the user selecting specific cuts is prone to influence the cut selection. With the Schwartz method, cuts are optimized, selected by the user and frozen. For this analysis a different method was chosen to reduce bias: the bootstrap method discussed in Section 5.4.1.

The rare decay result is determined from the actual signal given a predicted background. The signal background is predicted from the sideband data and Monte Carlo shapes. The sideband regions are chosen to be more statistically significant (larger) than the signal region so the background prediction has a lower statistical fluctuation than the actual signal.

For a blind analysis, the selection of optimal cuts is "blind" to the data signal region. Cut variables and optimal cut values are chosen by using sideband data and Monte Carlo shapes to maximize the experimental sensitivity. While a blind analysis is unbiased by statistical fluctuations in the signal data, it is still susceptible to fluctuations in the sideband data. The Monte Carlo shapes should not introduce statistical problems so they are determined from a sample which is 15× larger than the data sample.

The human element could introduce bias by selecting cuts on signal data. To protect from this bias, all cut variables were chosen in a blind manner, and an automated procedure is used to select the optimal cut values.

5.3.3 Cut Grid Results

With a cut grid, one does not "search" a parameter surface for a global minimum (e.g. a minimum chi-square to a fit or a minimum sensitivity). Instead, the entire parameter space (i.e. all cut combinations) is calculated. Although the cut grid method does not find the absolute global minimum, it has several advantages. A cut grid does not get trapped in local minima and produces results for a parameter surface that has a poorly defined or very broad global minimum. The cut grid for this analysis uses the cuts defined in Section 5.2.1. The results of the cut grid are displayed as a histogram of all cut combinations. Figure 56 and Figure 57 show the histogram of sensitivities for each decay mode.

5.3.4 Efficiencies

The efficiencies for each cut combination are calculated using Monte Carlo generated for the rare and normalizing mode. Equal amounts of mode specific Monte Carlo are generated for both the normalizing mode and the rare mode. The Monte Carlo simulates the trigger efficiency as well as the cut efficiencies. The ratio of normalizing

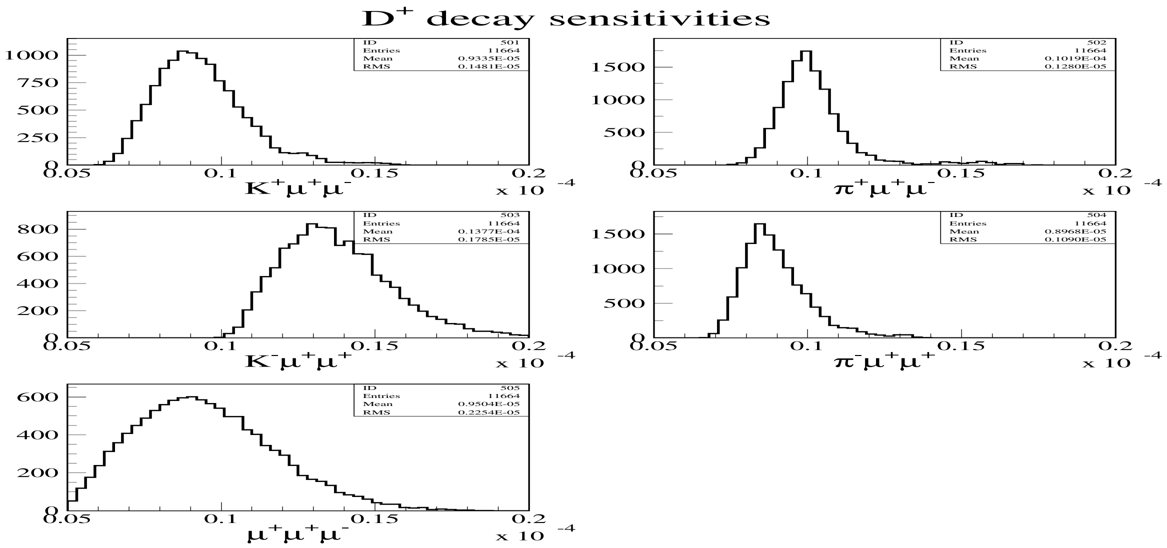


Figure 56: The experiment sensitivity for D^+ decay modes is shown as a histogram of all cut combinations.

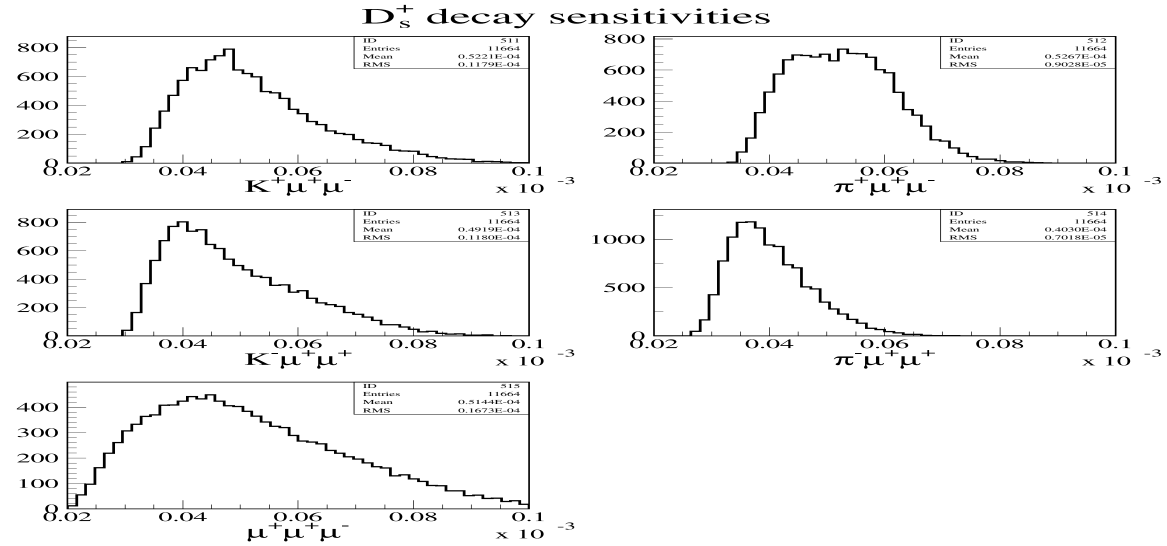


Figure 57: The experiment sensitivity for D_s^+ decay modes is shown as a histogram of all cut combinations.

mode signal to rare mode signal gives the relative efficiency (since the branching ratio is set to zero in particle2.dat). By comparing the ratio of a three-body normalizing

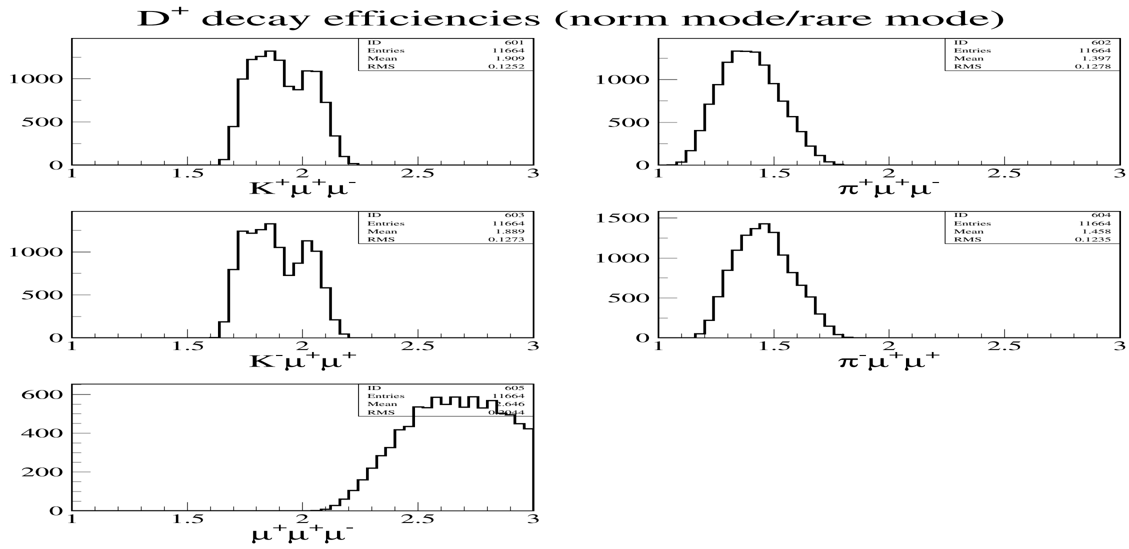


Figure 58: Relative efficiencies for the D^+ decay modes resulting from all cut combinations are shown. Each plot is a histogram of sensitivities resulting from all cut combinations in the cut grid.

mode to a three-body body rare mode, the efficiencies of the cuts that are applied equally to both are assumed to nearly cancel. Acceptance, vertexing, and momentum cuts are applied equally to the normalizing and rare modes. Muon ID efficiencies were studied in Section 4.1. Histograms of the ratio of normalizing mode efficiency to rare mode efficiency resulting from all cut combinations are shown in Figure 58 and Figure 59.

5.3.5 Signal Region Optimization

The invariant mass constructed from a fully reconstructed final state with correct particle identification appears as a Gaussian for a high statistics signal. The invariant mass plots of the normalizing mode for data and Monte Carlo were fit to a Gaussian and the means and widths determined. These values are used to define the signal

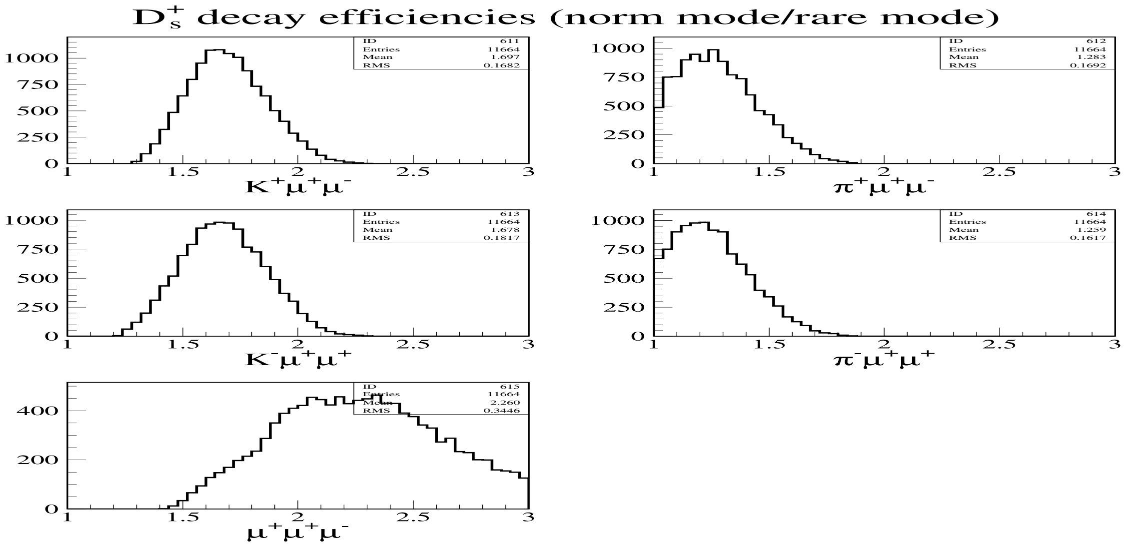


Figure 59: Relative efficiencies for the D_s^+ decay modes resulting from all cut combinations are shown. Each plot is a histogram of sensitivities resulting from all cut combinations in the cut grid.

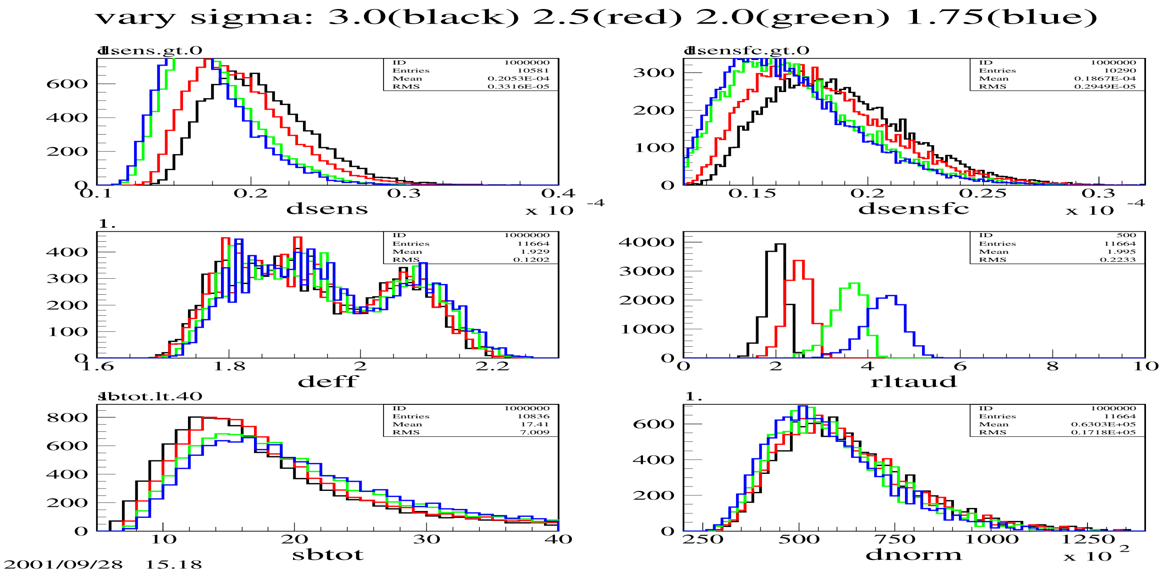


Figure 60: The sensitivity of the mode $D^+ \to K^+\mu^+\mu^-$ was tested with the signal region set to four different sizes. The upper plots show the differences in sensitivity for the Rolke-Lopez method and for the Feldman-Cousins method. The lower plots show the separate variables used in the sensitivity calculation: efficiency, τ, sideband total, and normalizing signal. A $\pm 2\sigma$ signal region is optimal.

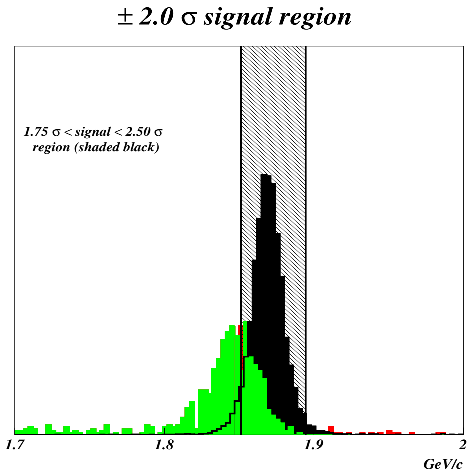


Figure 61: For the $D^+ \rightarrow K^-\mu^+\mu^+$ decay channel, the $D^+ \rightarrow K^-\pi^+\pi^+$ Monte-Carlo reflection shape is compared to the $K^-\mu^+\mu^+$ Monte-Carlo signal shape. A change in the "double misid" rate of two pions misidentified as two muons will have an effect on the calculation of τ. Reducing the signal region to 1.75σ makes ratio of signal to sideband for the reflection more closely equal the ratio of sideband to signal for the rest of the background. This reduces the effect the uncertainty in the K2pi reflection size has on the calculation of the ratio of the sideband and signal events.

and sideband regions. A $\pm 3\sigma$ signal region will contain 99.7% of a Gaussian signal, however a better signal region size can be set that maximizes signal to background. Decreasing the signal region from $\pm 3\sigma$ to $\pm 2\sigma$ will reduce a Gaussian signal by 5% and a linear background by 33%. The optimal value depends on the relative sizes of the signal and background. As shown in Figure 60, the sensitivity for different signal sizes was tested and a signal region of 2σ was chosen as the optimal signal region size.

The $D^+ \rightarrow K^-\mu^+\mu^+$ decay channel is complicated by the presence of the $D^+ \rightarrow K^-\pi^+\pi^+$ ($K\pi\pi$) reflection arising from the simultaneous misidentification ("double misid") of the two pions as two muons. The shape of this reflection is shown in Figure 61. As described in Section 4.2.3, there is uncertainty in the double misid rate which can affect the τ used in the Rolke-Lopez confidence tables. A τ value which is too low predicts too little background in the signal region and can lead to a poor upper limit or a false signal. A τ value which is too high predicts too much signal and can lead to an unjustifiably good limit with a "negative" signal (i.e. fewer entries than expected from background).

Since the $K\pi\pi$ reflection is present in both the signal and sideband regions, the effect of double muon misidentification uncertainty can be reduced by setting the signal and sideband regions so that the $K\pi\pi$ reflection "cancels out." If the $K\pi\pi$ reflection is the only background this would mean setting the signal region and the sideband regions to have equal amounts of $K\pi\pi$ reflection in them. In reality there is a linear background in the signal region which is about as large as the reflection, so the relative amounts of $K\pi\pi$ reflection in the sideband and signal regions should be equal to what τ would be without the reflection, something difficult to do since the reflection and combinatoric backgrounds change with different cuts. However, setting the signal region to be large and the sidebands to be small for this decay mode minimizes the impact that a large variation in the $K\pi\pi$ reflection has on the calculation of τ. For the $D^+ \to K^-\mu^+\mu^+$ decay channel the signal region was set to 3σ and the sidebands were decreased. This reduces the sensitivity but also reduces the systematic error for this mode.

5.3.6 Rare Decay Monte Carlo Backgrounds

To look up the confidence intervals in the Rolke-Lopez tables, one needs to use the sideband data, the signal data, and τ – the known relationship between sideband and signal data. τ can be determined in various ways: with a Monte Carlo data sample, with an appropriate independent data sample, or with an assumed background shape.

A large Monte Carlo data sample is used in this analysis to calculate τ. When the statistics of the Monte Carlo backgrounds becomes small, the resulting τ values produce a spread, and in some cases a deviation, from the average τ value. the tightest cuts in the cut grid can produce very few events in the Monte Carlo sample, even with a Monte Carlo sample many times larger than the data sample. The values of τ over the cut grid is shown in Figure 62. The average τ value determined with higher statistics was used to define a tau representative of all cut combinations. This was done by requiring that the number of events in the (Monte Carlo) signal region be greater than ten, and using the average of the τ distribution.

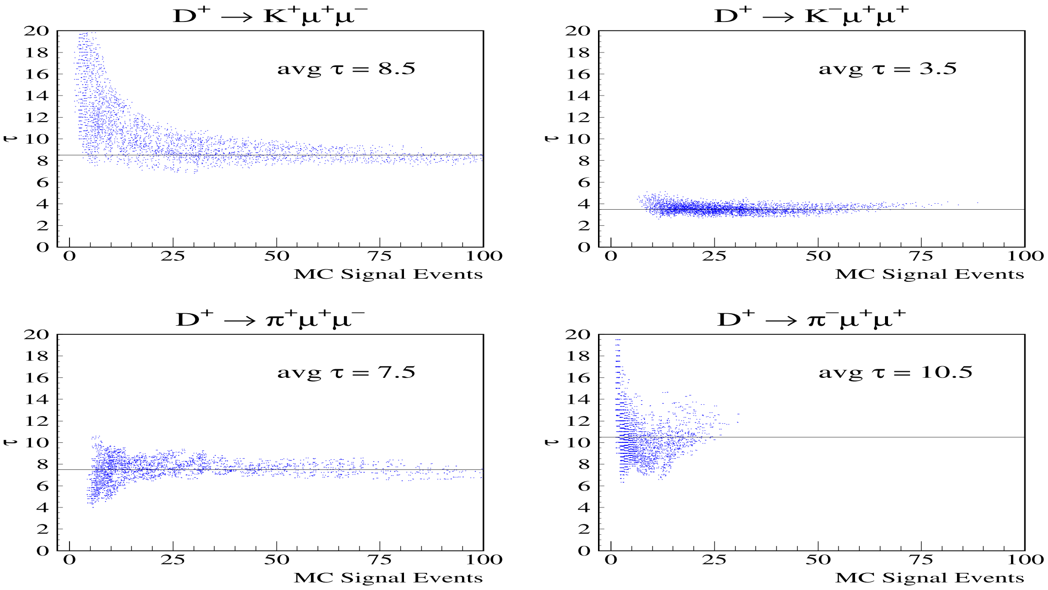
5.4 **Experimental Sensitivity and Bias Reduction**

For the present analysis we did not follow the Schwartz method for blindly determining an optimal set of analysis cuts. We used a bootstrap method that automates the selection of cuts in a blind manner and incorporates a random re-sampling of the data to reduce bias from statistical fluctuations.

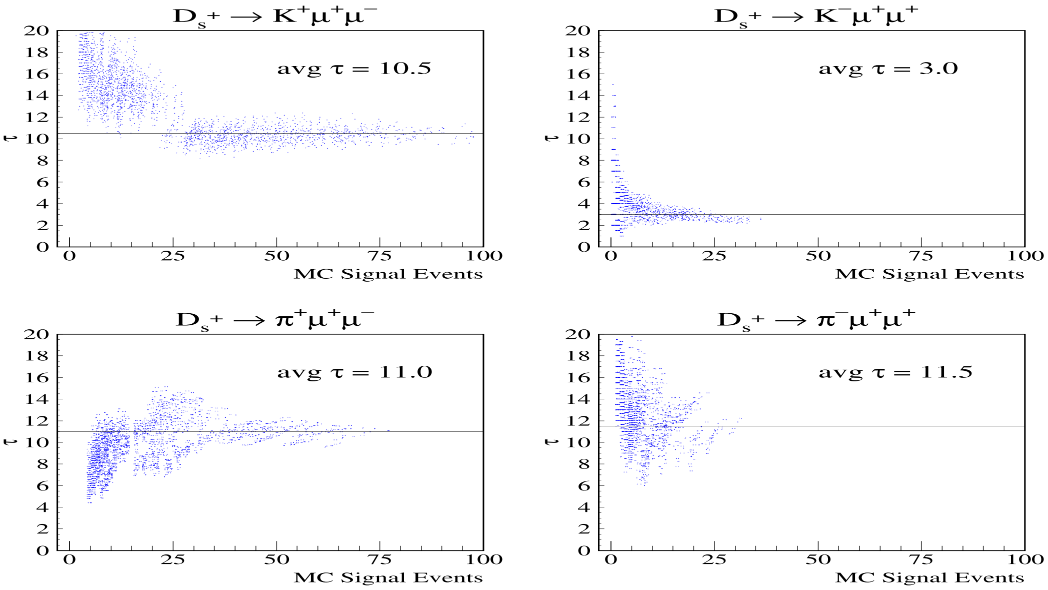
5.4.1 Bootstrap Statistical Methods

The bootstrap[1] is a method to reduce bias from results sensitive to statistical fluctuations. Split sample studies are a familiar form of bias testing, but splitting reduces the statistics of the sample and is not useful in the search for upper limits. Other examples are "jackknife" methods in which data points are removed from the original sample. While jackknife methods simply remove data points, bootstrap methods randomly remove and duplicate data points. The inclusion of data points

[1]as defined in statistics, not the S-matrix bootstrap used in quantum field theory



(a) τ averages for D^+ decays



(b) τ averages for D_s^+ decays

Figure 62: τ values calculated over the cut grid. The average value for τ is determined for cuts where the number of Monte Carlo signal events are greater than 10. Fluctuations at low statistics (i.e. with tight cuts) could result in spurious values for τ. Therefore, the high statistics average value for τ is used to calculate the sensitivities and branching ratios for all cuts.

more than once, and the ability to create a bootstrap sample the same size as the original data set gives bootstrap methods more power to test for bias.

The procedure is shown in Figure 63 and outlined in the following paragraphs. For the present analysis, data points are randomly chosen until the bootstrap sample is the same size as the original sample. The bootstrap sample will then have the statistical sensitivity of the original sample but the population of statistical outliers will vary measurably with different bootstrap samples. If the bootstrap sample is too small (i.e. less than 20 events), the results could vary too much for the method to effectively reduce bias. If the bootstrap sample is too large (i.e. greater than 1000 events), the ability to remove bias becomes less significant.

Wolfgang Rolke and Angel Lopez [45] have developed a bootstrap method for rare decay analysis, which they call the "double bootstrap," to decouple the optimization procedure from the limit determination. The procedure as applied to this analysis is as follows: (i) create a root sample of rare decay candidate events by using the loosest cuts in the cut grid; (ii) create a bootstrap sample from the root sample; (iii) calculate the sensitivity of the bootstrap sample for all the cuts in the cut grid, and find the cut combination which produces the best sensitivity; (iv) create a second bootstrap sample from the root sample; (v) calculate the sensitivity and branching ratio from the second bootstrap sample using the best cut combination determined in step (iii).

Steps (ii) to (v) are repeated 1000 times and the results from step (v) produce a distribution of sensitivities. The average of the distribution of double-bootstrap sensitivities is the bias-reduced result. The creation of the bootstrap samples in steps (ii) and (iv) do sample from both the signal and sideband regions. But step (iii)

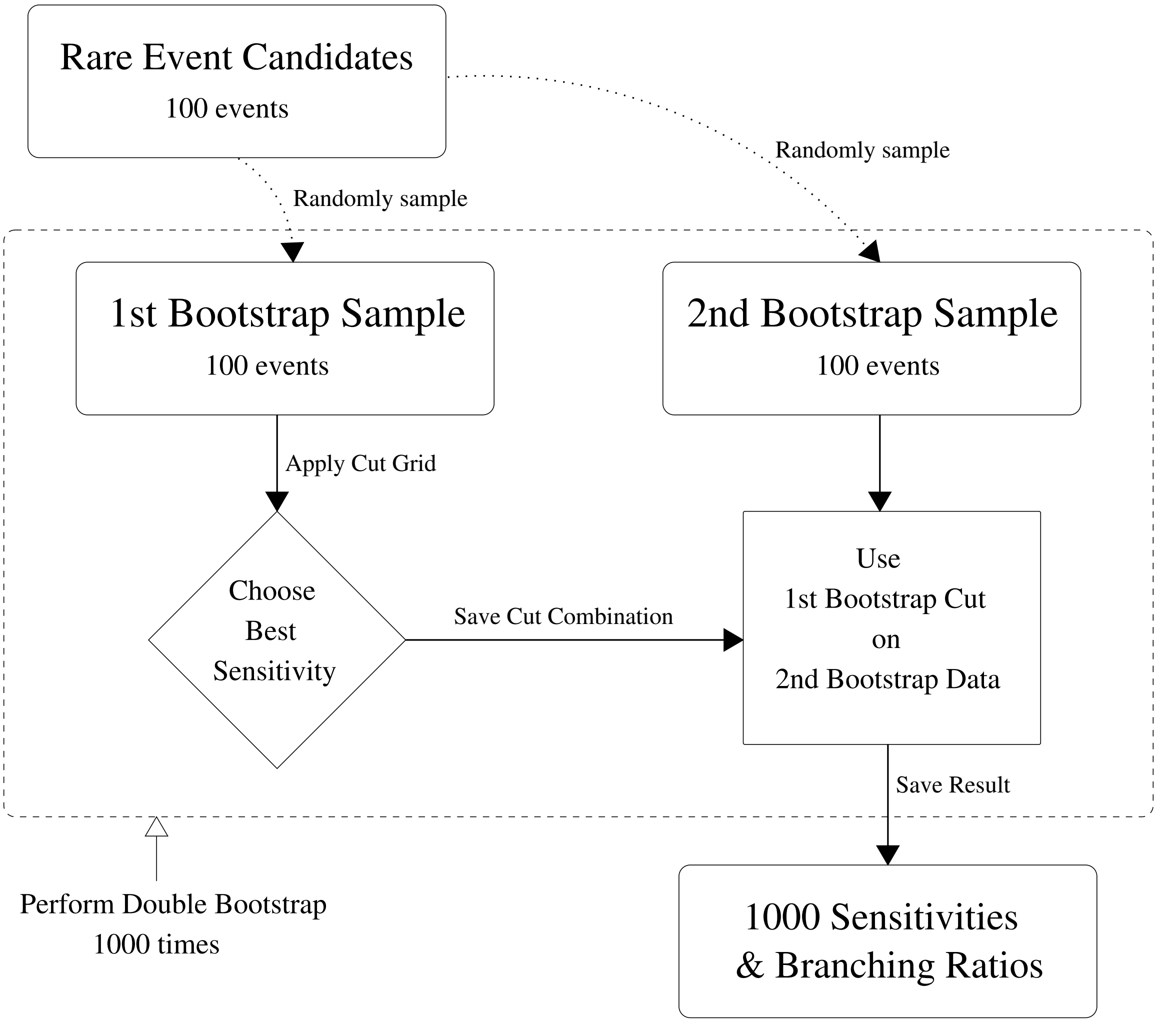


Figure 63: Double Bootstrap Flow Chart.

optimizes the sensitivity with the signal region masked, so the analysis is still blind. The second bootstrap sample is used to determine a bias-reduced measure of the sensitivity and the branching ratio. The double bootstrap procedure is outlined in Figure **??**.

The behavior of the bootstrap method is shown in Figure 64. The large histogram shows the sensitivities determined for the entire cut grid. The two smaller histograms are the results of the double bootstrap. The leftmost histogram is the bootstrapped

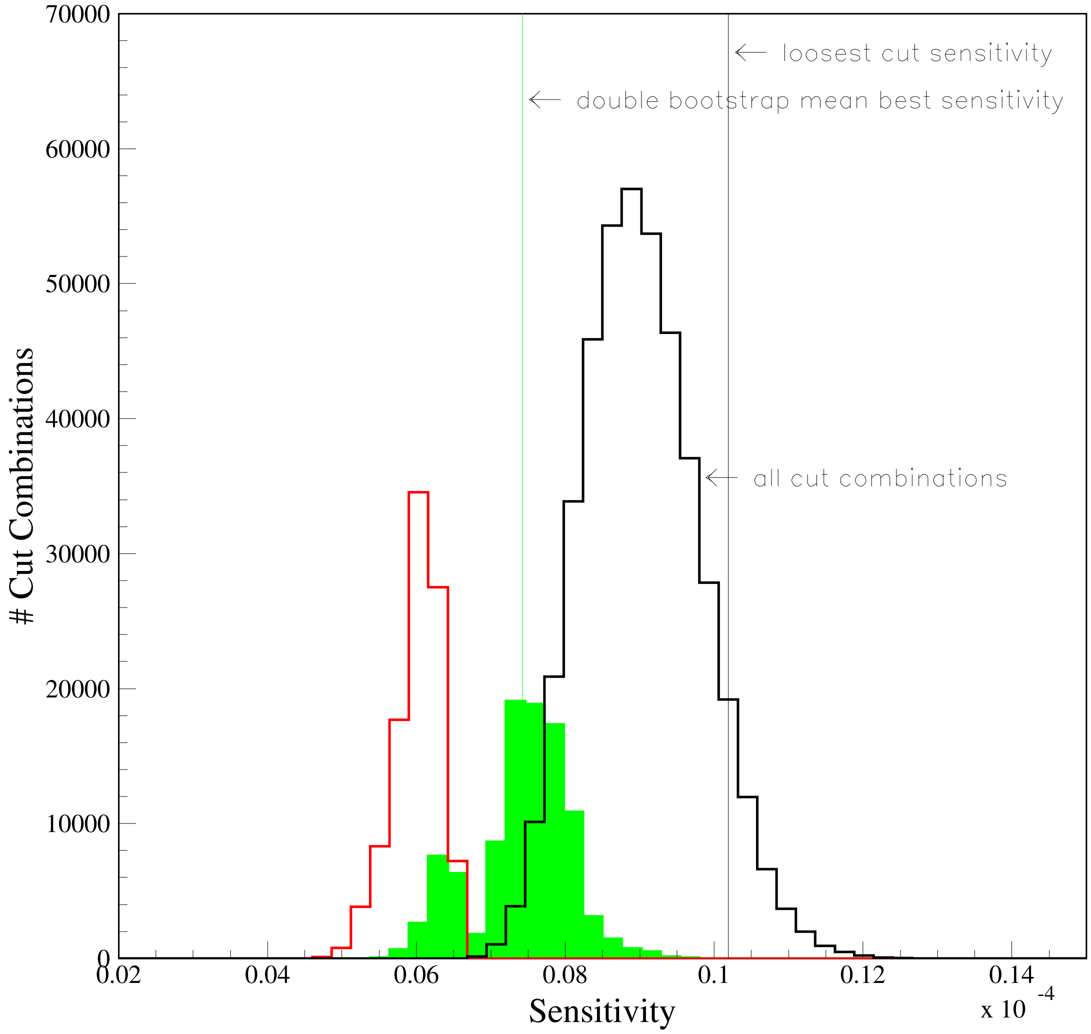


Figure 64: Histogram of $D^+ \rightarrow K^+\mu^+\mu^-$ sensitivities with the double bootstrap. The larger histogram is the sensitivity resulting from all cut combinations. The smaller line histogram is the best sensitivity from the first bootstrap sample. The shaded histogram is the sensitivity from the second bootstrap sample using the best cut combination derived from the first bootstrap sample.

best sensitivity. This value is biased toward selecting larger statistical fluctuations.

The second bootstrap histogram is the bias-reduced result.

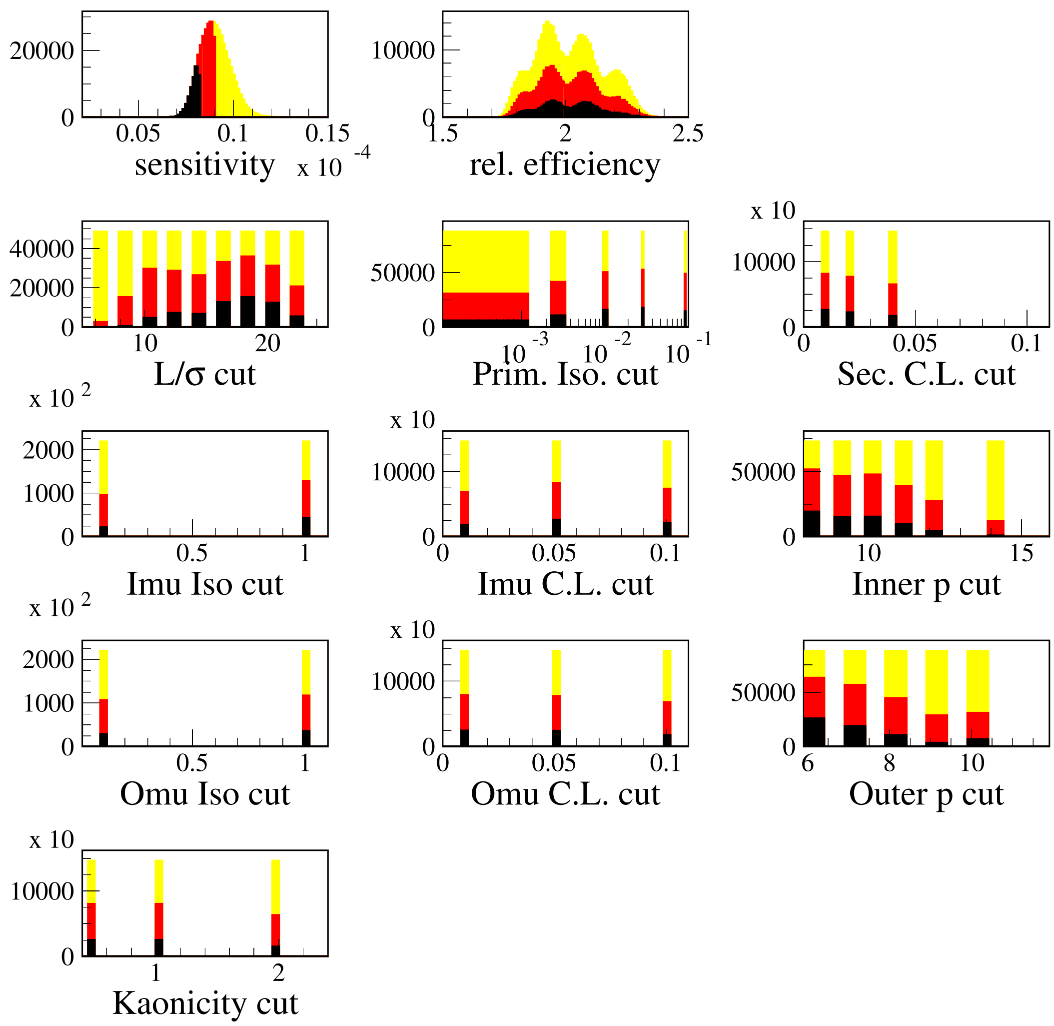


Figure 65: Histogram of individual cut effects for $D^+ \to K^+\mu^+\mu^-$. These histograms illustrate which cuts are producing the best sensitivities. The lightly shaded data corresponds to the worst sensitivities. The dark-shaded data corresponds to the best sensitivities. Ideally the cut ranges and step sizes should be defined so that the optimal cut values are bracketed. The behavior of ℓ/σ_ℓ is an example of a good cut range definition. Some cut variables may have a less dramatic effect on the sensitivity making their cut ranges and step sizes harder to define.

The behavior of the individual cut variables is show in Figure 65. Each individual plot is shaded so that events with best sensitivity are shown in the darker color. ℓ/σ_ℓ

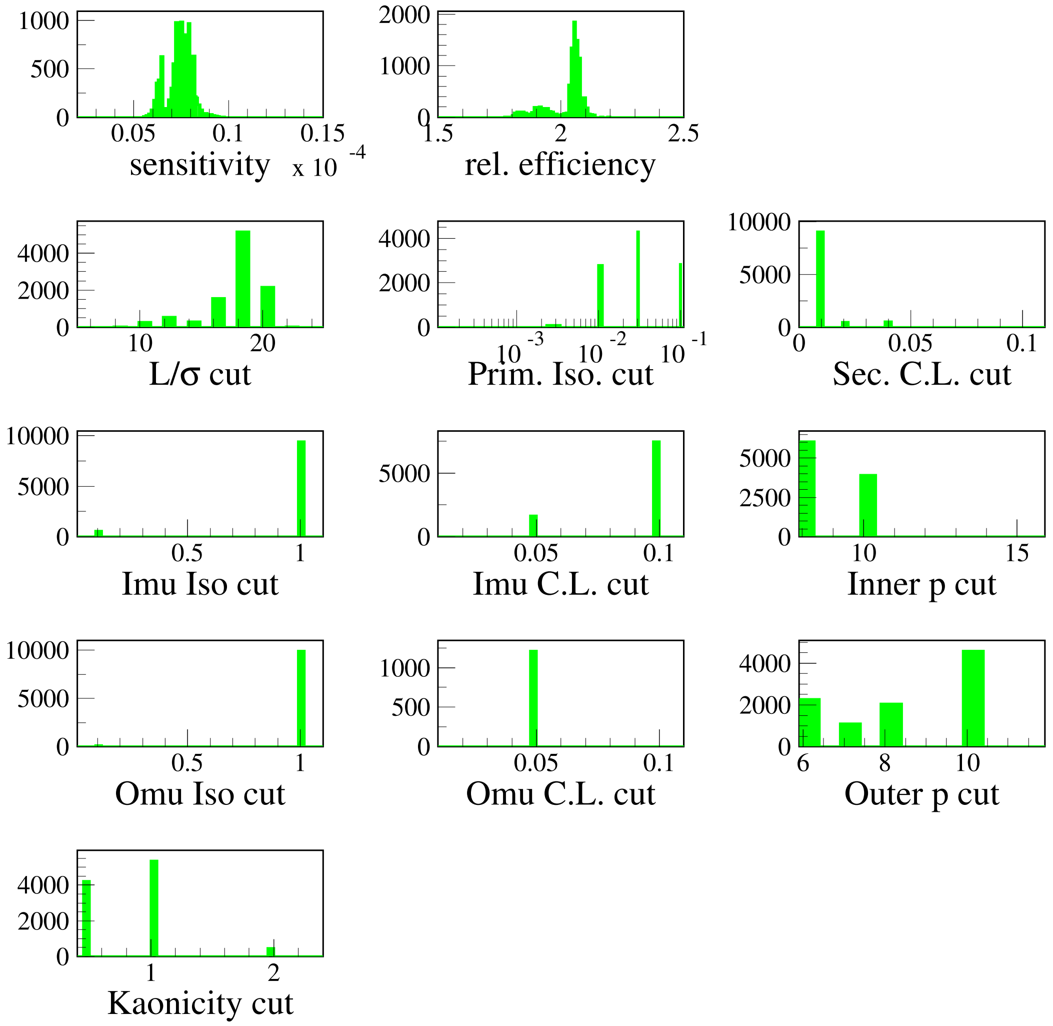


Figure 66: Bootstrap Best Cuts for $D^+ \to K^+\mu^+\mu^-$. The best cut for each of the 1000 bootstrap samples is shown for each cut variable in the cut grid.

illustrates the method for providing the correct range for the cut grid. The bootstrap should be allowed to find the optimal value for a specific cut variable. An ℓ/σ_ℓ cut of 15 ($\ell/\sigma_\ell > 15$) is chosen most often for the optimal sensitivity. The range for a cut in the cut grid should bracket the optimal value. The other cuts show a much broader peak (or no clear peak). Flat can be included for completeness, but should not affect the result. The cut ranges are restricted by the minimum skim criteria, such as muon confidence level cut 1% (to reduce muon misidentification) and a primary isolation cut of 10% (to reduce D^* backgrounds).

For the Schwartz method for a blind analysis, the user picks a single set of cuts without using signal region information. For the double bootstrap method the sensitivity calculation is blind to the signal region, and the selection of specific cuts is automated and repeated 1000 times. The definition of optimal experimental sensitivity is often biased by the subjective nature of the definition. One objective definition of sensitivity is to choose the cut combination giving the very best sensitivity. Optimizing cuts on this principle is biased toward selecting downward fluctuations in the sidebands. If one backs off from the hardest cuts somewhat, the choice of cuts becomes somewhat subjective and could be biased by the user. The double bootstrap objectively and accurately defines a sensitivity that is based on the measured sensitivity for each decay mode and that reduces the bias that arises during the sensitivity optimization procedure. The features of the double bootstrap are discussed in greater detail in the final chapter with the incorporation of the branching ratio results.

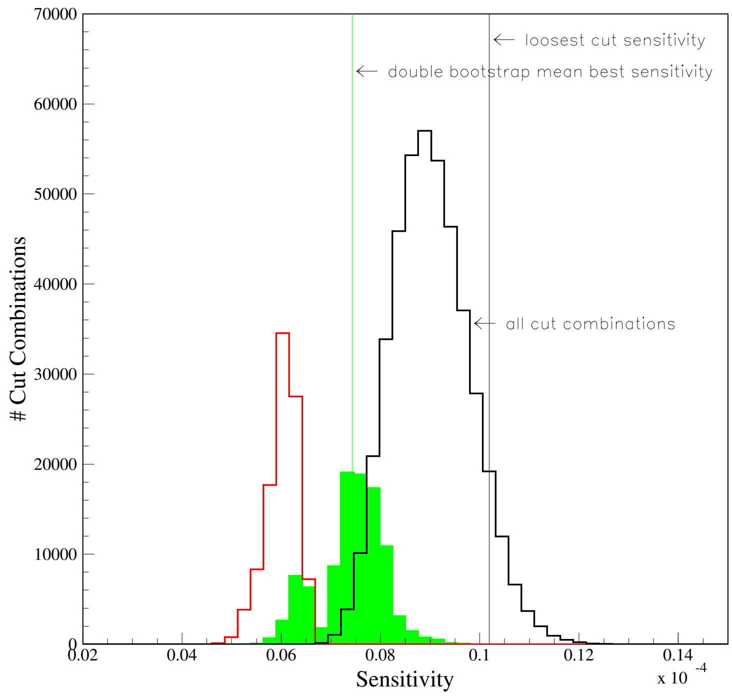
5.4.2 Bootstrap Sensitivities

The bootstrap sensitivities for all decay modes (except the $\mu\mu\mu$ modes) are shown in detail in Figure 67 to Figure 74. Each figure shows the bootstrap sensitivity accompanied be the cut grid analysis for that decay mode. As can be seen in each of the sensitivity histograms, the double bootstrap behaves in a very similar fashion for all decay modes. The optimal sensitivity selected by the double bootstrap selects a value close to, but slightly greater than the tail of the histogram.

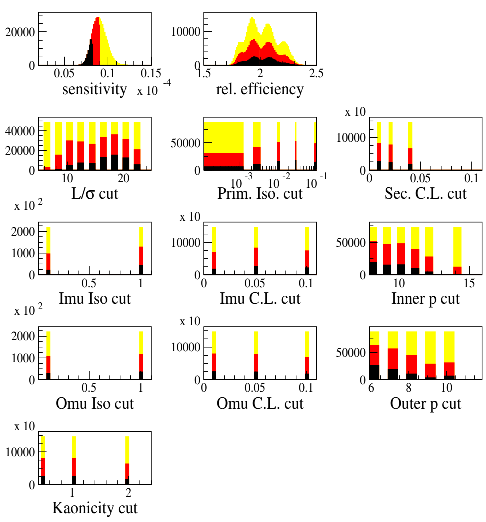
Table 15: Bootstrapped Sensitivities

Mode Mode	E831 Sensitivity	E791 Result	E687 Result
$D^+ \to K^+\mu^+\mu^-$	7.4×10^{-6}	4.4×10^{-5}	9.7×10^{-5}
$D^+ \to K^-\mu^+\mu^+$	5.0×10^{-6}	—	12.0×10^{-5}
$D^+ \to \pi^+\mu^+\mu^-$	7.6×10^{-6}	1.5×10^{-5}	8.9×10^{-5}
$D^+ \to \pi^-\mu^+\mu^+$	5.3×10^{-6}	1.7×10^{-5}	8.7×10^{-5}
$D^+ \to \mu^+\mu^+\mu^-$	—	—	—
$D_s^+ \to K^+\mu^+\mu^-$	3.3×10^{-5}	1.4×10^{-4}	—
$D_s^+ \to K^-\mu^+\mu^+$	2.1×10^{-5}	1.8×10^{-4}	—
$D_s^+ \to \pi^+\mu^+\mu^-$	3.1×10^{-5}	1.4×10^{-4}	—
$D_s^+ \to \pi^-\mu^+\mu^+$	2.3×10^{-5}	0.8×10^{-4}	—
$D_s^+ \to \mu^+\mu^+\mu^-$	—	—	—

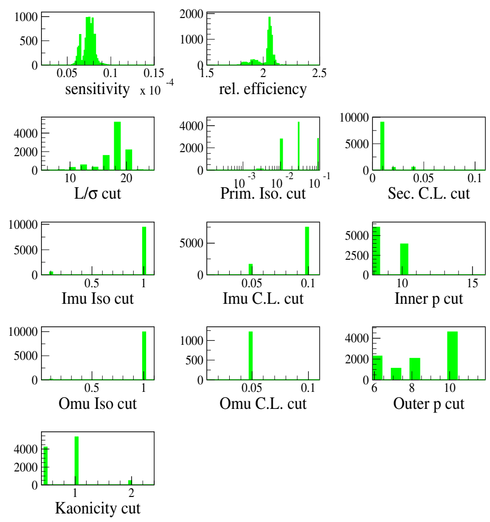
The histograms of the cut grids show which cuts tend to be selected for optimal sensitivities. The ℓ/σ_ℓ variable has the most clearly defined optimal cut values. The other cut variables have less pronounced optimal values. Even so, the fact that a range cut values can provide optimal sensitivity demonstrates the usefulness of all of the cuts variables. These variables, in many different combinations, can provide effective sensitivity optimization. The results of the double-bootstrapped sensitivities are listed in Table 15.



(a) Double Bootstrap Sensitivities

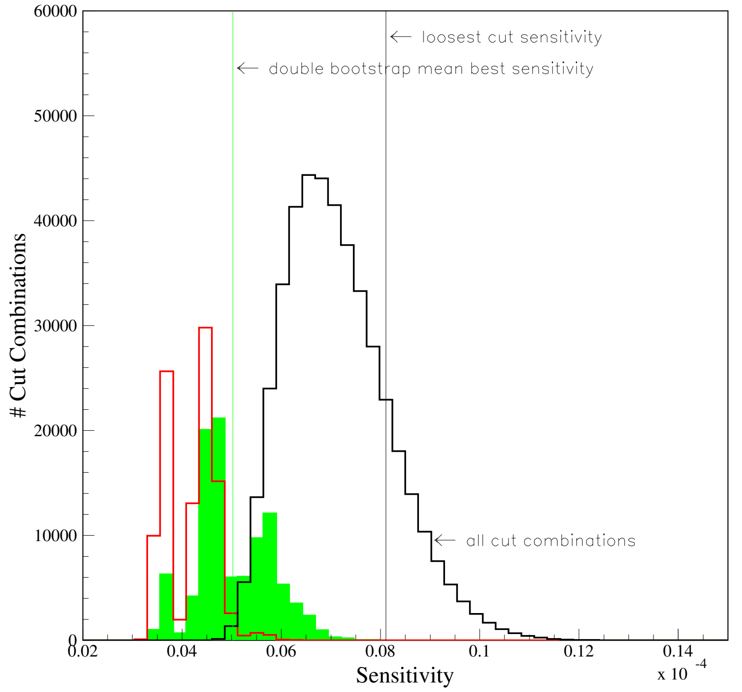


(b) Cut Grid Analysis

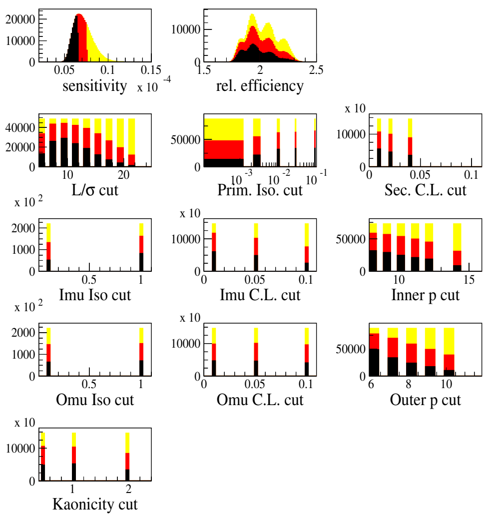


(c) Bootstrap Best Cut

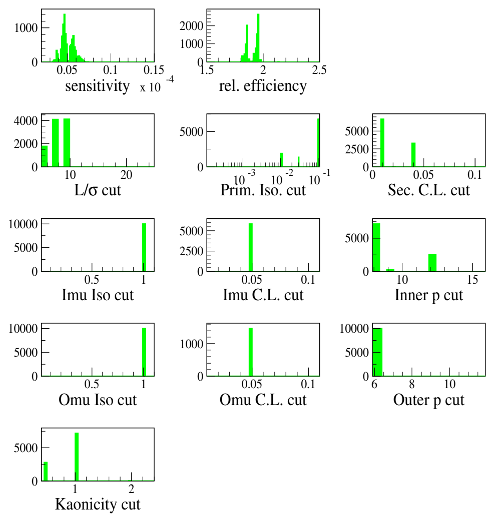
Figure 67: $D^+ \rightarrow K^+\mu^+\mu^-$ Sensitivity Analysis



(a) Double Bootstrap Sensitivities

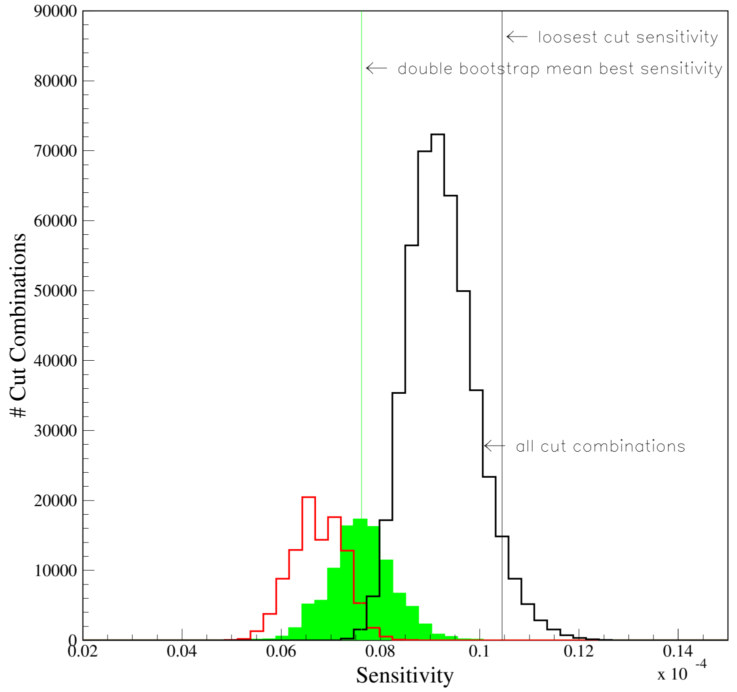


(b) Cut Grid Analysis

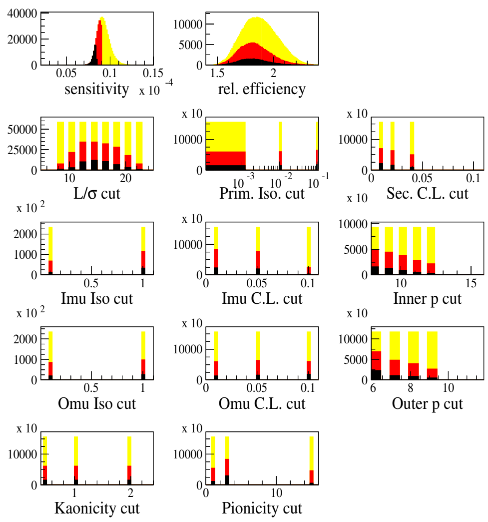


(c) Bootstrap Best Cut

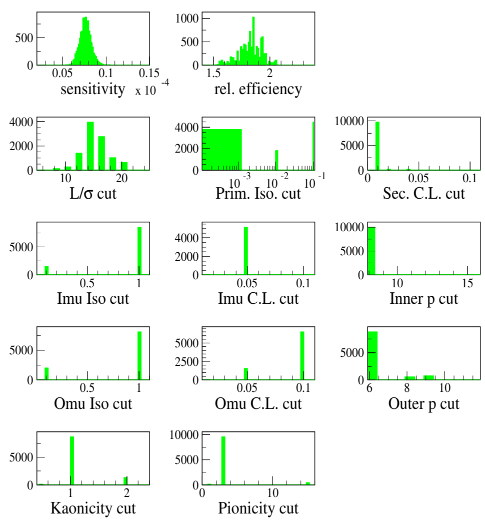
Figure 68: $D^+ \to K^-\mu^+\mu^+$ Sensitivity Analysis



(a) Double Bootstrap Sensitivities

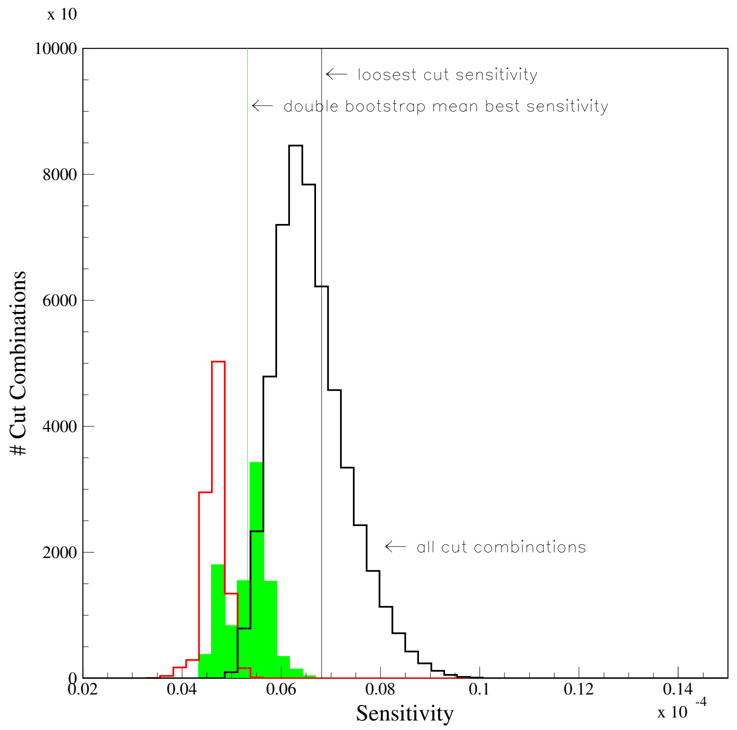


(b) Cut Grid Analysis

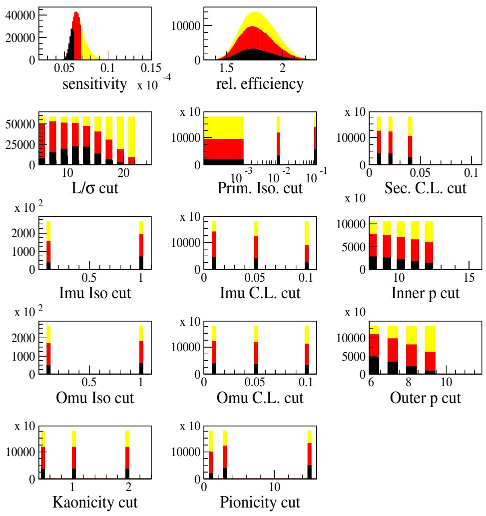


(c) Bootstrap Best Cut

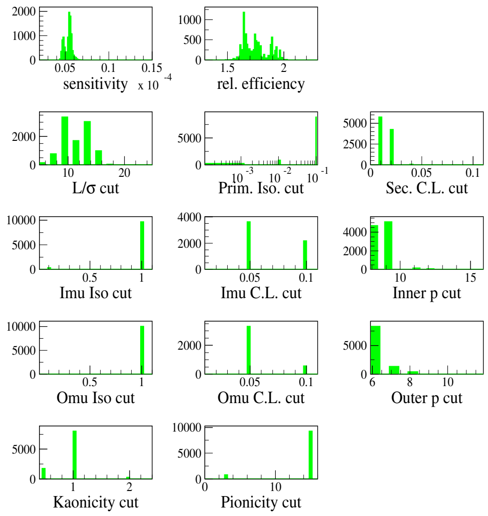
Figure 69: $D^+ \rightarrow \pi^+\mu^+\mu^-$ Sensitivity Analysis



(a) Double Bootstrap Sensitivities

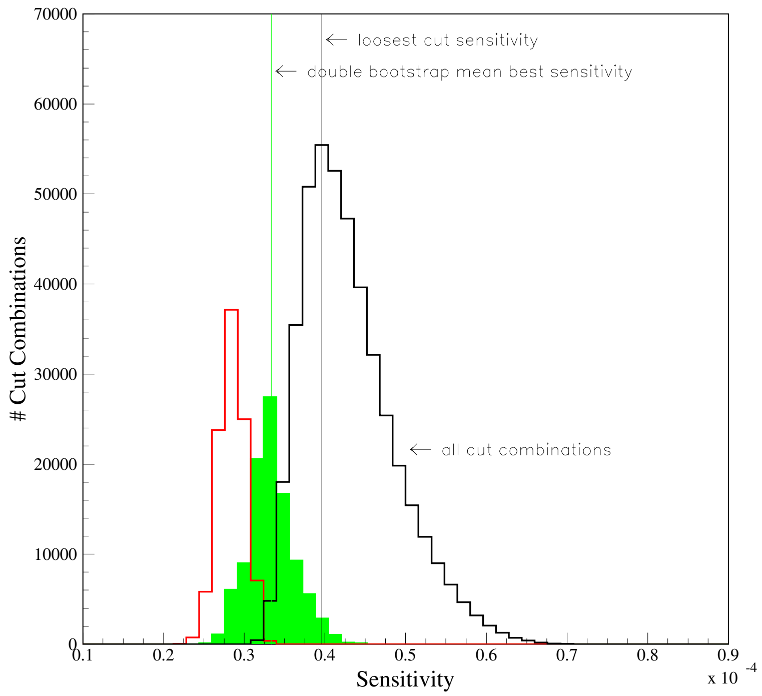


(b) Cut Grid Analysis

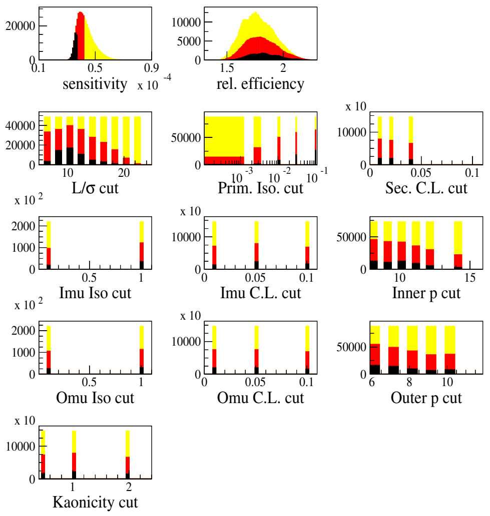


(c) Bootstrap Best Cut

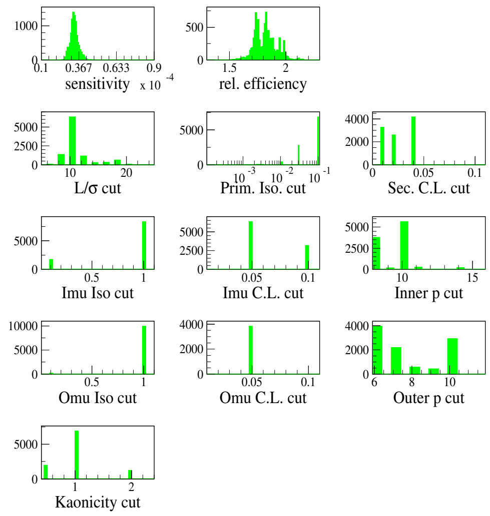
Figure 70: $D^+ \to \pi^- \mu^+ \mu^+$ Sensitivity Analysis



(a) Double Bootstrap Sensitivities

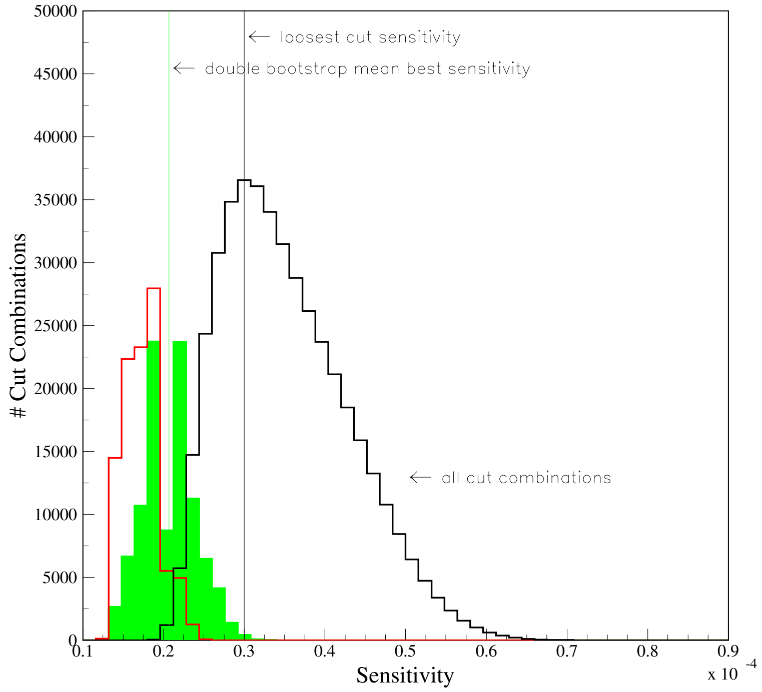


(b) Cut Grid Analysis

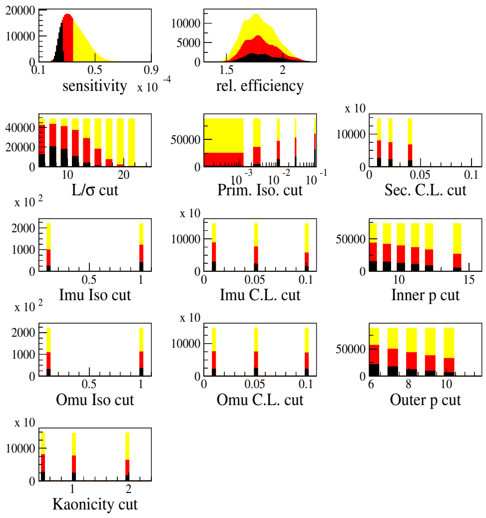


(c) Bootstrap Best Cut

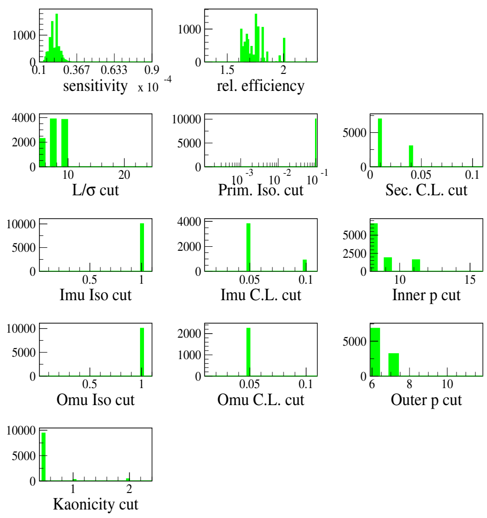
Figure 71: $D_s^+ \to K^+\mu^+\mu^-$ Sensitivity Analysis



(a) Double Bootstrap Sensitivities

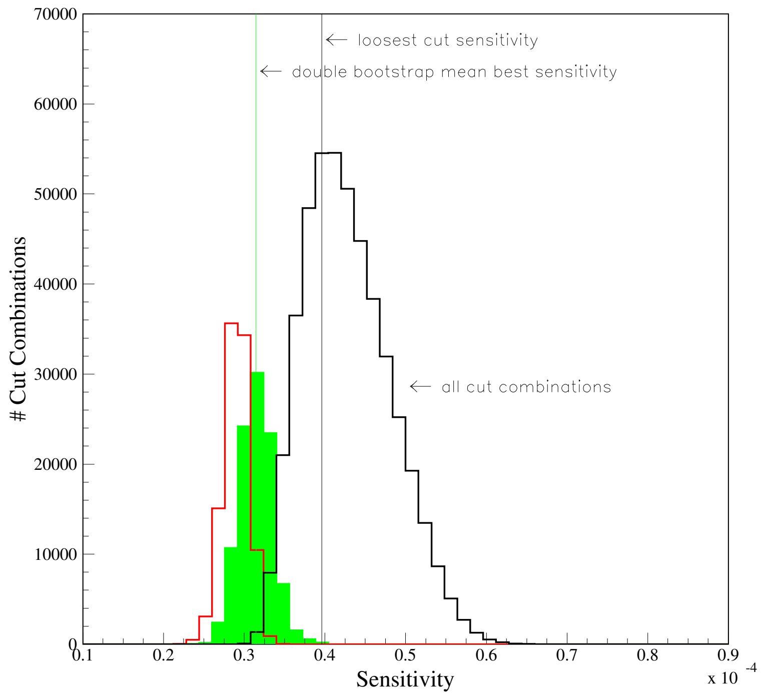


(b) Cut Grid Analysis

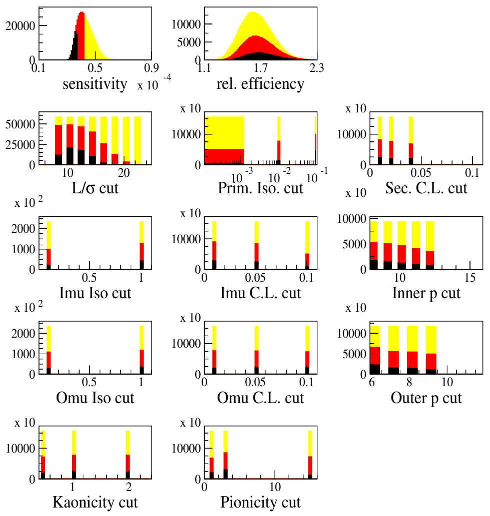


(c) Bootstrap Best Cut

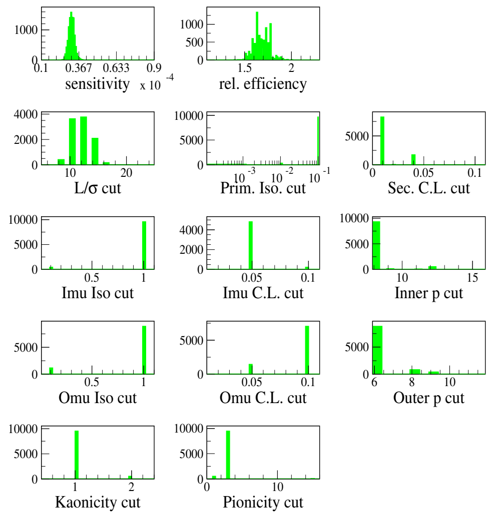
Figure 72: $D_s^+ \to K^-\mu^+\mu^+$ Sensitivity Analysis



(a) Double Bootstrap Sensitivities

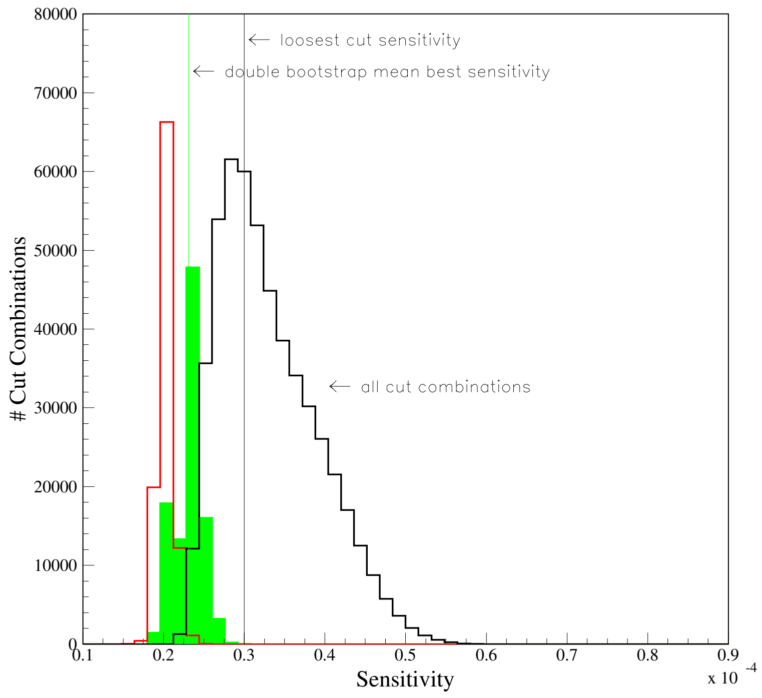


(b) Cut Grid Analysis

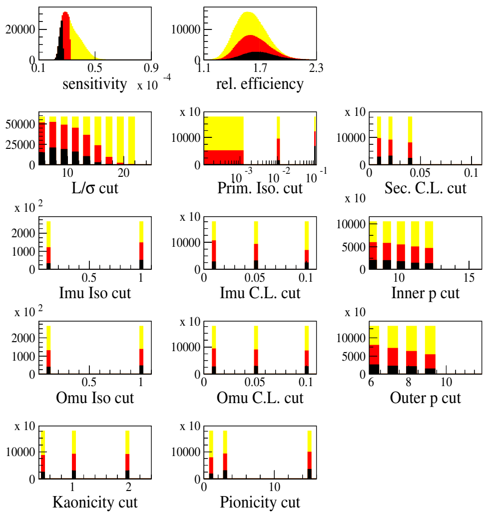


(c) Bootstrap Best Cut

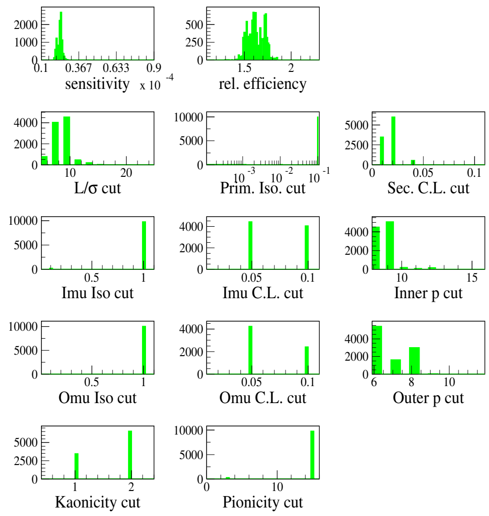
Figure 73: $D_s^+ \to \pi^+\mu^+\mu^-$ Sensitivity Analysis



(a) Double Bootstrap Sensitivities



(b) Cut Grid Analysis



(c) Bootstrap Best Cut

Figure 74: $D_s^+ \to \pi^-\mu^+\mu^+$ Sensitivity Analysis

CHAPTER VI

RESULTS

In this chapter we will discuss the incorporation of systematic errors into the upper limits, the impact of matrix element uncertainties, and the limits on branching ratios for the 10 decay modes.

6.1 Systematic Errors

Systematic errors that are independent of the level of data statistics can arise from uncertainties inherent in the analysis method. The most significant sources of systematic error in this analysis are uncertainties in the Monte Carlo simulation and uncertainty in the branching ratio of the normalizing mode, σ_{BR}. The dominant Monte Carl systematic errors for this analysis are the uncertainty in muon id efficiency, $\sigma_{\mu-id}$, and triggering uncertainties arising from HC simulation uncertainties, σ_{HC}. These errors, which will be discussed in the following sections, are determined individually and combined, in quadrature, to produce a total relative systematic error σ_r^2,

$$\sigma_r^2 = \sigma_{HC}^2 + \sigma_{\mu-id}^2 + \sigma_{BR}^2 \,. \tag{44}$$

Since the final result of this analysis is a 90% confidence interval upper limit, the statistical and systematic errors are not stated separately from the results but are incorporated into the 90% confidence interval upper limit.

6.1.1 Normalizing Mode Uncertainties

The branching ratio of the rare decay is determined relative to a normalizing mode. Thus, the uncertainty in the branching ratio of the normalizing mode must be included in the uncertainty of the rare decay. The 2001 Particle Data Group (PDG) branching ratio for $D^+ \to K^-\pi^+\pi^+$ is 9.0 ± 0.6 which represents a 6.7% systematic error. The 2001 PDG branching ratio for $D_s^+ \to K^+K^-\pi^+$ is 4.4 ± 1.2 which represents a 27.3% systematic error.

6.1.2 Hadronic Calorimeter simulation uncertainties

The Hadronic Calorimeter (HC) determines a particle's energy by measuring the size of the hadronic shower produced by that particle as it interacts with the iron plates of the calorimeter (cf. Section 2.7.3). These showers were modeled with the software package GEANT. Since the GEANT simulation is slow, FOCUS Monte Carlo provides two faster options for HC shower simulation. The default method is to use a pre-stored shower library which incorporates the distribution of showers generated with GEANT for different particles at different energies. The third, and fastest, method termed "crude," uses a simple parameterization of the showers as a function of the particle energy and type.

The hadronic trigger is required for all events in this analysis. As discussed in Section 2.9, one of the components in the Hadronic Trigger is a minimum energy deposited in the HC. In charm events, the energy deposited in the HC comes from a combination of charm daughters and other hadrons. Since muons deposit much less energy in the HC than hadrons, 3-body decays with muon daughter tracks will deposit less energy in the HC than 3-body decays with three hadronic daughters. The triggering efficiency for the normalizing modes will therefore be higher than the rare decay modes.

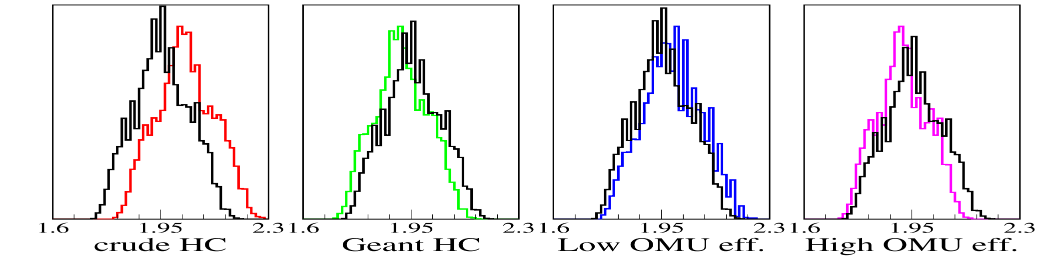
The relative triggering efficiency of the normalizing mode and the rare decay mode is included in the efficiencies calculated with the ratios of rare and normalizing mode Monte Carlo. The uncertainty in the relative triggering efficiency can be estimated by comparing the Monte Carlo results with the three different HC simulations.

Figure 75 shows how the relative efficiency (of normalizing mode Monte Carlo to rare decay mode Monte Carlo) differs from the default HC simulation (using the HC shower library) when different HC simulation options are chosen. The upper plots are histograms of the relative efficiency for each cut combination in the cut grid. The lower plots show, for the most optimal cut combinations, the relative efficiencies for the different HC simulation options divided by the relative efficiency of the default HC simulation for individual cut combinations.

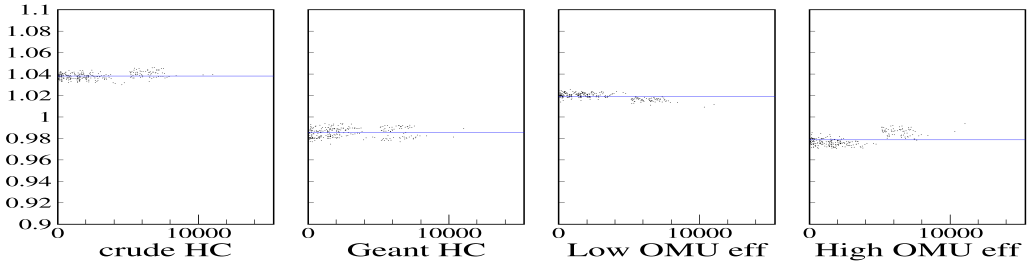
6.1.3 Muon ID uncertainties

It was concluded in the outer muon efficiency study (cf. Section 4.1.2) that the outer muon RPC efficiencies are best modeled with the average efficiency of the two Pass-1 methods. The inner muon efficiency is near 100% and so will not contribute significantly to the overall systematic error. Therefore, a systematic error of the muon ID efficiency can be determined using the upper and lower range of Pass-1 efficiencies for the outer muon RPCs.

Figure 75 shows how the overall relative efficiency (of normalizing mode Monte Carlo to rare decay mode Monte Carl) differs from the default when the range of outer muon RPC efficiencies is used. The upper plots are histograms of the relative



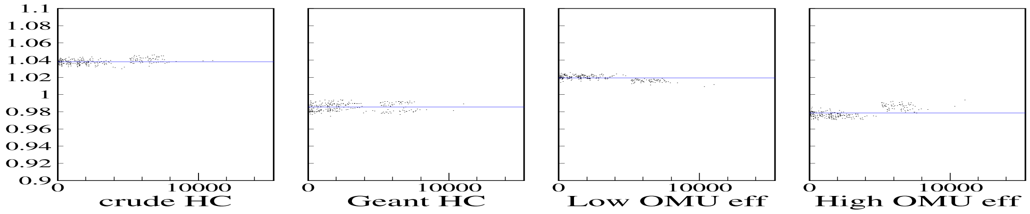
(a) Efficiency Histograms



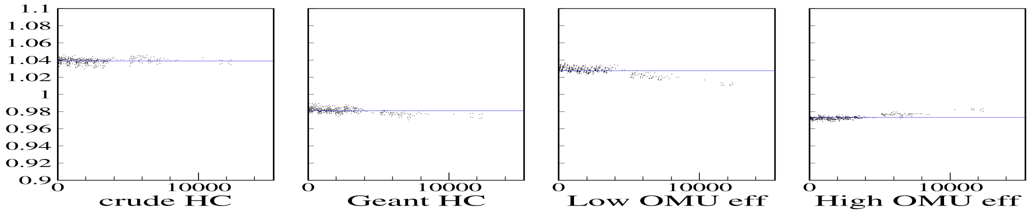
(b) Efficiency Ratios

Figure 75: Histograms of relative efficiencies. These plots show the relative efficiencies defined as normalizing mode MC yield over rare decay mode MC yield. For example Outer Muon Pass-1 method 1 is low efficiency so rare signal is lower and norm/rare ratio is higher.

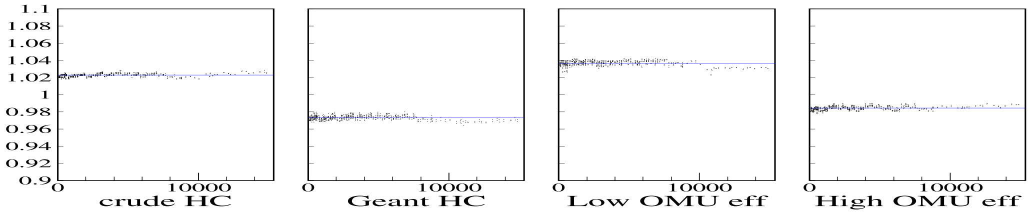
efficiency for each cut combination in the cut grid. The lower plots show, for the most optimal cut combinations, the relative efficiencies generated with Monte Carlo using the different RPC efficiency options divided by the the relative efficiency generated with the default Monte Carlo method.



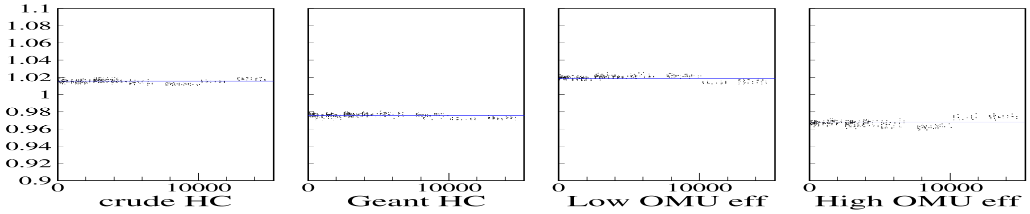
(a) $D^+ \to K^+\mu^+\mu^-$ Efficiency Ratios



(b) $D^+ \to K^-\mu^+\mu^+$ Efficiency Ratios

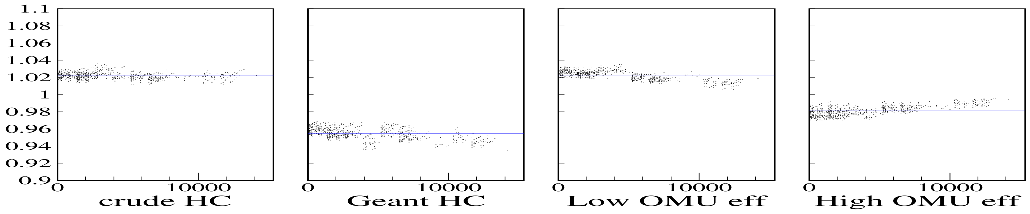


(c) $D^+ \to \pi^+\mu^+\mu^-$ Efficiency Ratios

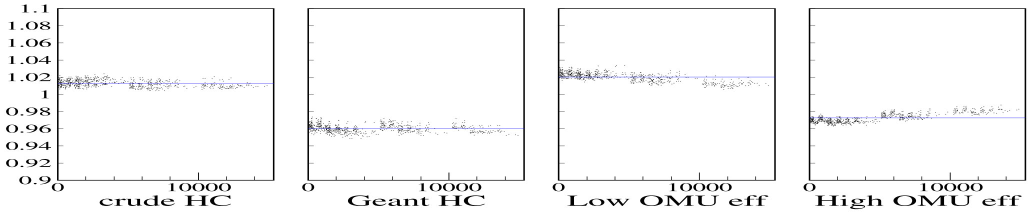


(d) $D^+ \to \pi^-\mu^+\mu^+$ Efficiency Ratios

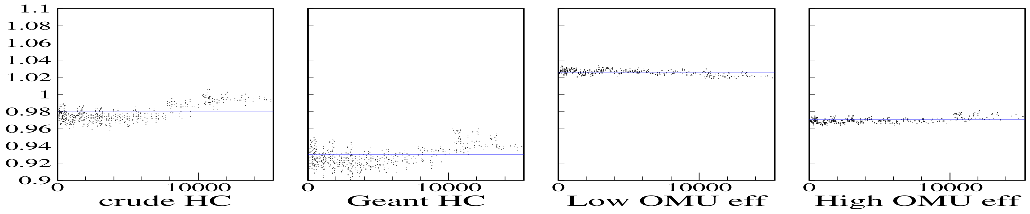
Figure 76: D^+ Monte Carlo Systematics



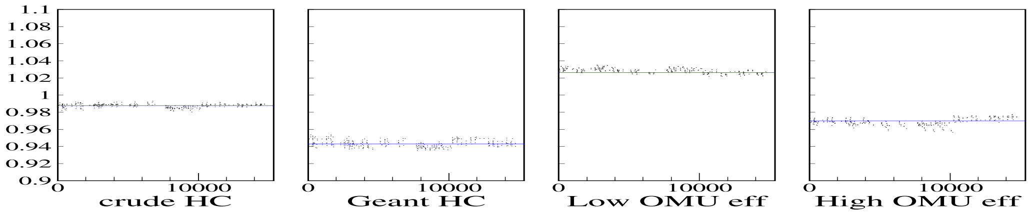
(a) $D_s^+ \to K^+\mu^+\mu^-$ Efficiency Ratios



(b) $D_s^+ \to K^-\mu^+\mu^+$ Efficiency Ratios



(c) $D_s^+ \to \pi^+\mu^+\mu^-$ Efficiency Ratios



(d) $D_s^+ \to \pi^-\mu^+\mu^+$ Efficiency Ratios

Figure 77: D_s^+ Monte Carlo Systematics

6.1.4 Combined Systematic Errors

The separate systematic errors are combined to produce a total systematic error for each decay mode. The three methods for simulating HC showers (crude, shower-library, and GEANT) are combined using Equation (37) to produce a single value, σ_{HC}, for each decay mode. Similarly, the three methods for simulating outer muon efficiencies (default, high, and low) are combined to produce a single value, $\sigma_{\mu-id}$, for each decay mode. These two numbers are added in quadrature (Equation (44)) to the branching ratio uncertainty giving a total relative systematic error σ_r for each decay mode. These errors are summarized in Table 16.

Table 16: Systematic Error Summary (as percentage).

Decay Mode	σ_{HC}	$\sigma_{\mu-id}$	σ_{PDG}	σ_r
$D^+ \to K^+\mu^+\mu^-$	2.8	1.9	6.7	7.5
$D^+ \to K^-\mu^+\mu^+$	2.7	2.6	6.7	7.7
$D^+ \to \pi^+\mu^+\mu^-$	2.5	2.7	6.7	7.6
$D^+ \to \pi^-\mu^+\mu^+$	2.0	2.6	6.7	7.5
$D^+ \to \mu^+\mu^+\mu^-$	8.4	4.0	6.7	11.5
$D_s^+ \to K^+\mu^+\mu^-$	3.0	1.9	27.3	27.5
$D_s^+ \to K^-\mu^+\mu^+$	2.3	2.5	27.3	27.5
$D_s^+ \to \pi^+\mu^+\mu^-$	3.6	2.7	27.3	27.7
$D_s^+ \to \pi^-\mu^+\mu^+$	3.0	2.8	27.3	27.6
$D_s^+ \to \mu^+\mu^+\mu^-$	8.5	3.2	27.3	28.8

6.1.5 Cousins Highland Method

The systematic error is incorporated into the 90% confidence interval upper limit for each decay mode, rather than being quoted separately from the final branching ratio. Cousins and Highland have developed a method for incorporating systematic uncertainties into an upper limit [42]. Using this method, the increase in the Poisson upper limits, ΔU_n, is

$$\Delta U_n = \frac{1}{2}\, U_{_{RL}}^2 \;\, \sigma_r^2 \;\frac{U_{_{RL}} + b - n}{U_{_{RL}} + b} \tag{45}$$

for a predicted background b, a signal n, the Rolke-Lopez limit on the signal $U_{_{RL}}$, and a total relative systematic error, σ_r. The change in the 90% confidence interval upper limit becomes much smaller than than the systematic error using this method.

6.2 Forbidden Decay Matrix Elements

The simulation of a 3-body decay requires the calculation of the partial decay rate

$d\Gamma$

$$d\Gamma = \frac{(2\pi)^4}{2m}|M|^2 d\Phi(P; p1, p2, p3) \tag{46}$$

by integrating the Lorentz-invariant matrix M over 3-body phase space

$d\Phi(P; p1, p2, p3)$

$$d\Phi(P; p1, p2, p3) = \delta^4\left(P - \sum_{i=1}^{3} p_i\right)\prod_{i=1}^{3}\frac{d^3 p_i}{(2\pi)^3 2E_i}. \tag{47}$$

The matrix elements are undefined for decays forbidden by the standard model, so M is assumed to be uniform in phase space for these decays. The standard model predicts that the rare decays in this analysis are below FOCUS sensitivities, so rare decays are modeled with matrix elements uniform in phase space.

Extensions to the standard model could provide the matrix elements for rare or forbidden decays. To set a limit for a specific model, the phase space efficiency for that model should be used. The efficiencies across the Dalitz plot are shown for each of the decay channels for this analysis in Figure 78 and Figure 79. The efficiencies are calculated by comparing the Dalitz plot for events accepted with optimal cuts to the Dalitz plot for generated events. The tracks are labelled (1,2,3) in the order specified in the title for each plot. Thus for $K^+\mu^+\mu^-$, K^+ is track 1, μ^+ is track 2, μ^- is track 3. For the modes where tracks 2 and 3 have the same sign, these tracks are explicitly randomized by the FOCUS simulation and may be interchanged with no bias introduced to the Dalitz plot.

Failure to adjust the limit for a specific model will introduce a systematic error, but this error should be small if the resonances are widely spread across the Dalitz plot. Since we do not know the matrix elements for the forbidden decays, and we are not testing any particular Standard Model extensions, uncertainties arising from unknown matrix elements are not included in the systematic errors. For the future possibility of adjusting results using a specific model, the Dalitz plot efficiencies are included for completeness.

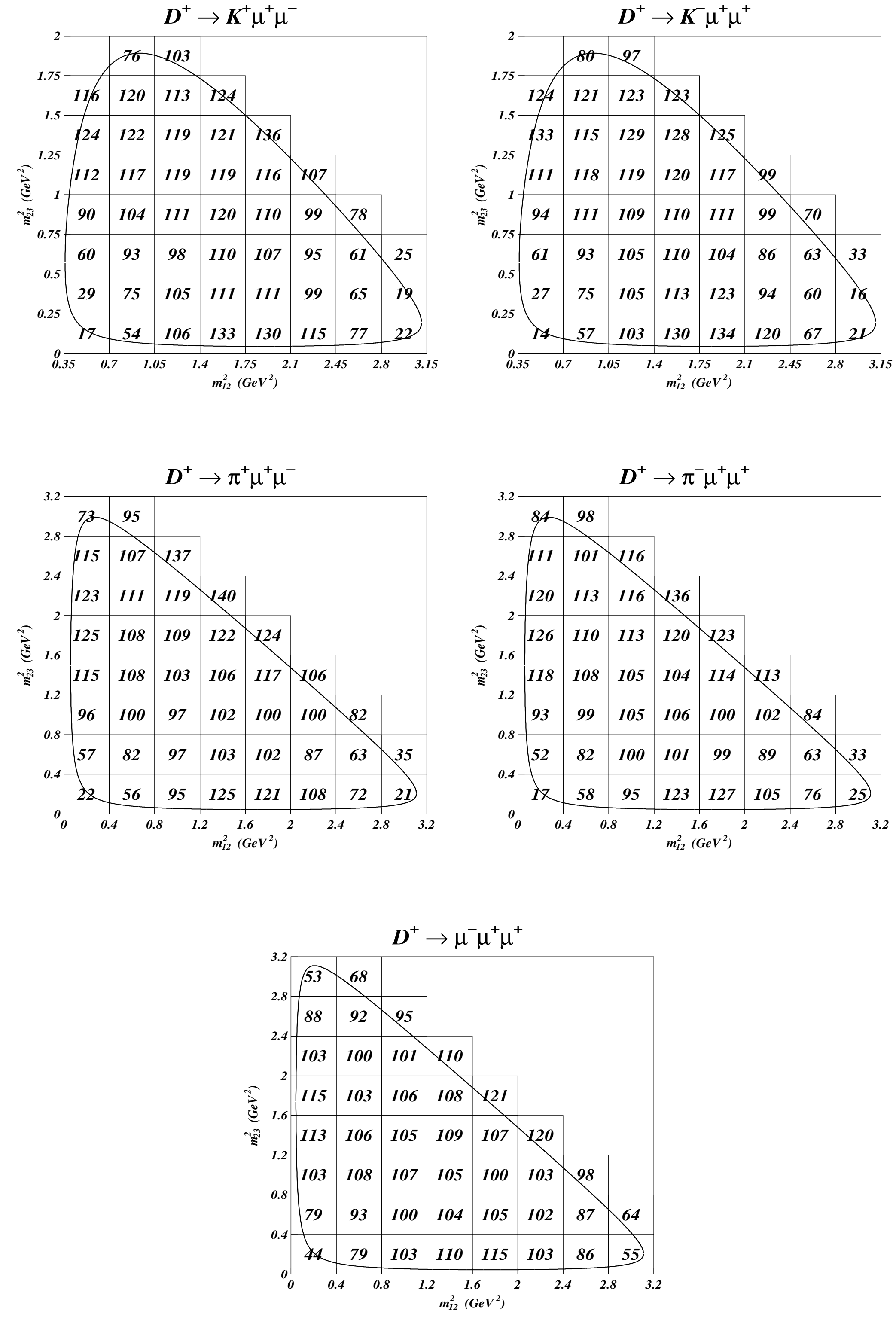


Figure 78: Variation in Monte Carlo efficiencies (for optimal cuts) across the Dalitz plot are shown for each D^+ decay mode. Bins are labelled with efficiencies (as a percent of the mean efficiency).

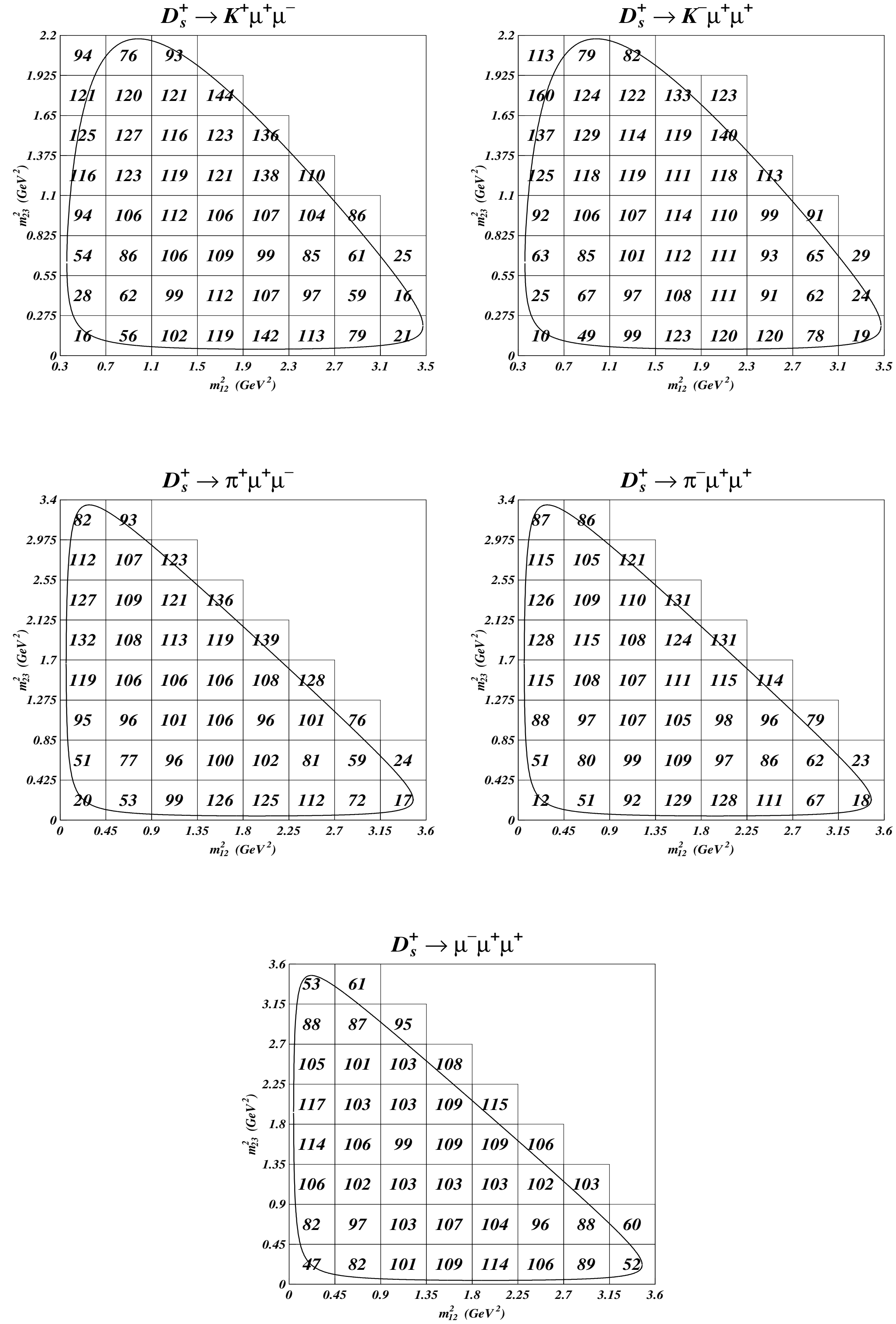


Figure 79: Variation in Monte Carlo efficiencies (for optimal cuts) across the Dalitz plot are shown for each D_s^+ decay mode. Bins are labelled with efficiencies (as a percent of the mean efficiency).

6.3 **Branching Ratios**

Now that we have defined a mechanism for determining analysis cuts, and a method for incorporating systematic errors into the confidence intervals, we will look at the data in the signal region and analyze the branching ratios. The branching ratios are determined in two ways. The principal method is to use the double bootstrap samples to get an average branching ratio of the re-sampled data. A second method is used to define a single cut combination for all decay modes. This method provides a comparison to other experiments that quote results based on a single cut combination. The $\mu\mu\mu$ decay modes have only one sideband event and are treated separately.

6.3.1 Double Bootstrap Results

The use of the double bootstrap algorithm provides two major features to the analysis. The first feature provided is a definition of sensitivity and branching ratio that is an objective and accurate representation of the experimental data. The second feature provided is the reduction of bias arising from the optimization process.

In general, the sensitivity should be optimal when the predicted background is reduced to nearly zero events. Therefore, analysis cuts are tightened until only very small signals remain. The Poisson fluctuations of small signals create a situation in which many very different but equally justifiable results may occur. The variation in these results can be significant which in turn could make cut selection bias a significant factor.

Figure 80 shows the experimental sensitivity and branching ratio as a function of the cut grid and the bootstrap. The two large histograms give the sensitivity and

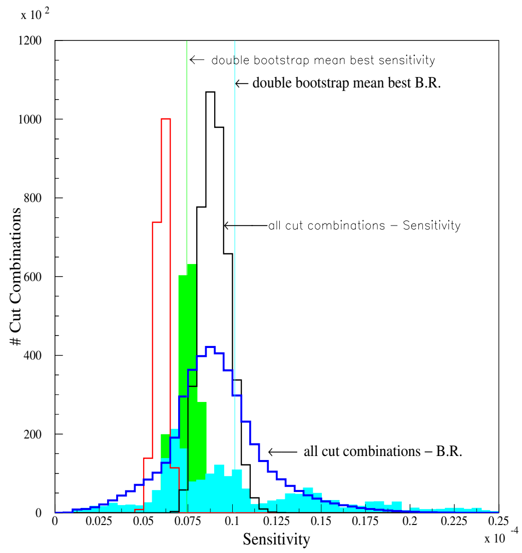


Figure 80: Double Bootstrap Results. The histograms are described in detail in the text. The labelled unshaded histograms are the non-bootstrapped sensitivities and branching ratios for all cut combinations. The shaded histograms are the double bootstrap sensitivities and branching ratios.

branching ratio as a function of all cut combinations of the cut grid. The double bootstrap sensitivities are shown in red and green as in Figure 64. The cut grid contains approximately 450,000 cut combinations. The bootstrap re-sampling was performed 10,000 times, and only one cut combination is plotted for each bootstrap sample. The bootstrap histograms are scaled up a factor of 10 for visual comparison to the non-bootstrap histograms.

The first bootstrap sample (the red histogram) is used to find the best cut combination and the second bootstrap sample uses the optimal cut combination found by the first sample. The second bootstrap sample is used to find the bias reduced sensitivity and branching ratio (the shaded histograms). The mean of the second bootstrap sensitivity and branching ratio are labeled.

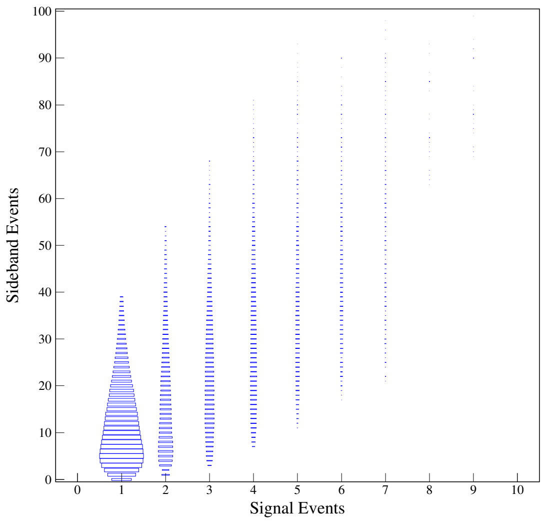
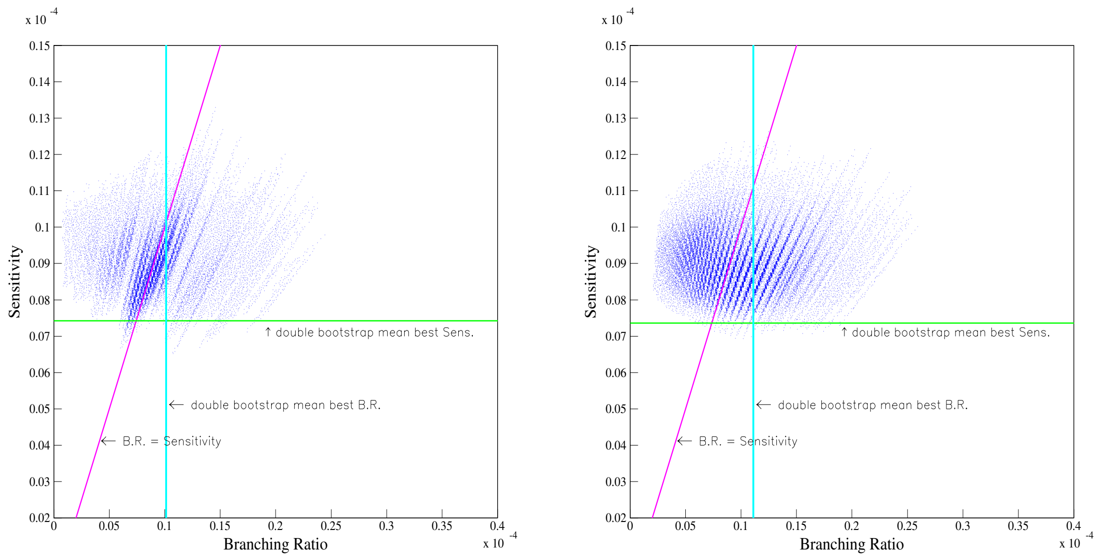


Figure 81: Sideband vs. Signal Spread

As is to be expected, the branching ratio has a wider spread in results than the sensitivity. The sensitivity is calculated using the Monte Carlo prediction for the signal region. The actual signal fluctuates in comparison with the sidebands as cuts vary giving rise to a spread in results. Due to the Poisson nature of these results, the spread can be large.

Also, the bootstrap re-sampling increases the number of available results introducing an additional spread to the bootstrap results. As can be seen in the histogram of the first bootstrap sample best sensitivities (red), the mean appears to be outside the range of all sensitivities for the non-bootstrapped sample. However, it should be noted that the first bootstrap histogram results from selecting the very leftmost tail of the cut combination histogram for each (first) bootstrap sample. The combined broadening effects of the branching ratio and bootstrap results in a very broad distribution for the bootstrapped branching ratio.



(a) Rolke-Lopez Sens. vs. B.R.

(b) Feldman-Cousins Sens. vs. B.R.

Figure 82: $D^+ \rightarrow K^+\mu^+\mu^-$ Branching Ratio Analysis

To understand the spread in branching ratio results, we can inspect Figure 81 which shows the number of signal events and sideband events as a function of the cut grid. For each signal event, the number of sideband events can vary dramatically. To get the final branching ratio the efficiency of the cut must be incorporated, but the changes in efficiency are small compared to the changes in Rolke-Lopez upper limits that result from a large spread in the signal to sideband events.

Figure 82 shows the sensitivity on the vertical axis vs. the branching ratio on the horizontal axis. Each point in the plot represents a different cut combination. This plot clearly shows the spread of results arising from the Poisson fluctuation of the small signals and the large availability of cuts.

For results on small signals, many problems arise when attempting to produce a statistically meaningful result that accurately reflects the full discriminating power of the experimental data. Among these problems is the construction of confidence intervals for small signals. Additionally, problems arise in the methods for selecting cuts that accurately probe the sensitivity limits of the data. If cuts are chosen too tightly and the data disappears, the sensitivity limits of the experimental data have not been tested. Similarly, if cuts are chosen too loosely, too much data may be present to effectively analyze the limits of the data. A method for cut selection that accurately probes the limits of the experimental data needs to be used.

As mentioned earlier, there has been much recent work developing methods for producing confidence intervals with accurate coverage for Poisson statistics. The Feldman-Cousins confidence intervals are currently the most commonly used. The Rolke-Lopez method uses a similar method, but has improved coverage [43]. As seen in Figure 82 results using the Feldman-Cousins limits have an apparent under-coverage whereas the Rolke-Lopez results have better apparent coverage. Since we are measuring the signal, this does not demonstrate the improved coverage of the Rolke-Lopez method but is consistent with it.

It is interesting to consider the spread of branching ratio vs. sensitivity for the tightest regions of the cut grid. Cuts that are tight enough to remove all background and signal events result in the observed signal usually matching the predicted signal (zero events). In this case the branching ratio should equal the sensitivity. In general, sensitivity improves with tighter cuts until the background is cut to zero or nearly zero events. Cuts tighter than the optimal cuts will loose efficiency, but the sensitivity will still be equal to the branching ratio. The branching ratio spread narrows around the "BR=Sensitivity" line with tighter cuts, but beyond the optimal sensitivity, the efficiency suffers and the sensitivity vs branching ratio plot doubles back on itself. This explains the concentration of results near the "BR=Sensitivity" line.

A common problem of the confidence interval methods is the question of how to interpret results when there is a statistical "hole" in the signal region. If the signal region is fluctuated downward relative to the sidebands, then a result significantly lower than the experimental sensitivity is possible. In practice, this problem may be less a problem with the confidence intervals themselves than with the method of dealing with a wide range of possible results.

One possible method for choosing a single result is to select a cut combination that produces the absolute best sensitivity and choose the branching ratio from this single cut combination as the result. This method clearly has the bias of selecting downward statistical fluctuations in the sidebands, which would tend to produce a worse limit. A response to this bias may be to find a cut combination that is "consistent" with several decay modes and declare the branching ratios from this new cut combination to be a fair and unbiased representation of the data. However, a single cut combination consistent with several decay modes will likely not be optimal for any particular decay mode.

As can be seen in Figure 82, there is a wide degree of uncertainty in results when selecting a single cut combination. Analysis methods that utilize a single cut combination require that the cut choice be done in a blind manner, to avoid potential bias arising from a subjective selection of cuts. A blind analysis will remove this user bias, but it still unsatisfying since there is a large (and random) uncertainty in the result. It is clearly not true to say that a single sampling of a wide distribution provides a useful characterization of that distribution.

Since there is a wide distribution of results in this analysis, a method for assessing the distribution of results is necessary to provide an accurate determination of the experimental data. An assessment of this distribution of results could be achieved by determining the spread of results for a good range of sensitivities — say the best 5% of cut combinations. The mean of this distribution of branching ratios could define a final result. However, the subjective nature of determining the best range of cuts (why 5%?) introduces a potential bias.

The double bootstrap method provides an *objective* method for determining the experimental sensitivity and branching ratio. The objective nature of the cut selection is a primary justification for using this method. The double bootstrap method thoroughly probes the sensitivity limits of the data, and reduces the bias that results from optimizing on small signals. The averaging of branching ratios over many bootstrapped cut combinations, provides a robust answer that is representative of the full "discriminating power" of the data rather than a single random sampling of possible answers.

For the double bootstrap method to be able to effectively probe the limits of data enough data points are needed to provide variable bootstrap samples. This is why minimum cuts were selected for each decay mode to try to provide between 50 and 100 events. Fewer events reduces the effectiveness of the method.

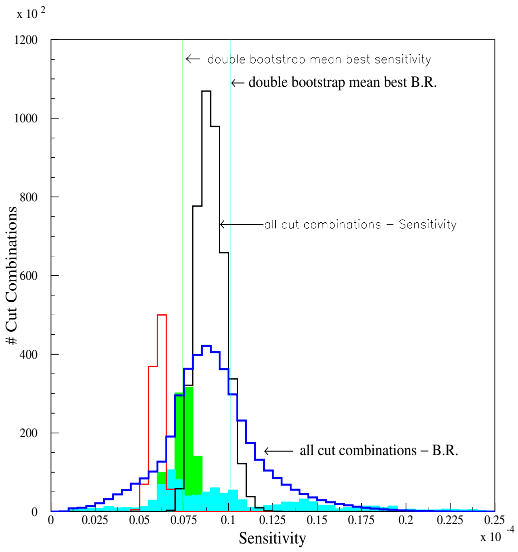
It is clear in Figure 82 that the spread in branching ratios below the double bootstrap sensitivity is much greater than the difference between the mean double bootstrap branching ratio and the best sensitivity branching ratio. It follows that the primary effect of the double bootstrap method is to reduce the uncertainty in branching ratios that would result from a single cut selection. A smaller effect is the reduction of bias that would appear if the single most optimal cut defined the answer.

The results for all decay modes show that the double bootstrap branching ratio is generally near the result for the optimal sensitivity. There is no apparent tendency to be either higher or lower than the optimal sensitivity. The $D^+ \to K^-\mu^+\mu^+$ result indicates a slight signal. This is mostly likely due to the muon misidentification uncertainty seen in data. The $D_s^+ \to K^-\mu^+\mu^+$ and the $\mu\mu\mu$ modes do not contain enough data points to apply the cut grid or the bootstrap, so the results are determined by the minimum skim cuts used for all modes.

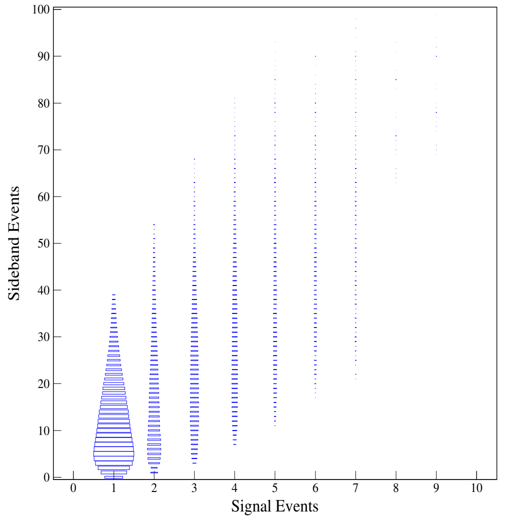
The systematic error is incorporated into the bootstrap limit by applying the Cousins-Highland method to each individual bootstrap sample. The difference in the limits is shown in Table 17.

Table 17: Bootstrapped Rare Decay Results

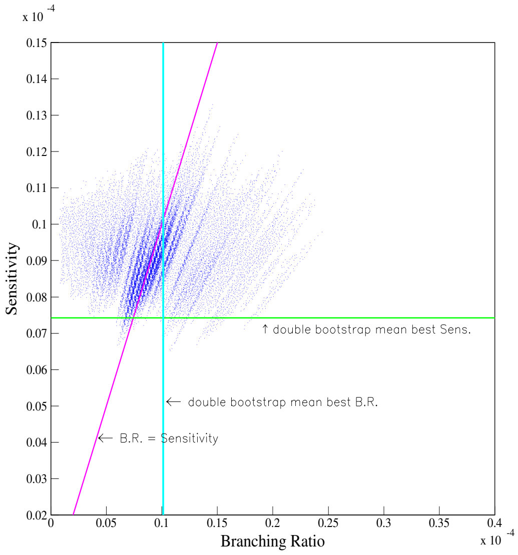
Decay Mode	Bootstrap Result	σ_r	BS result incl σ_r
$D^+ \to K^+\mu^+\mu^-$	10.4×10^{-6}	7.5%	10.4×10^{-6}
$D^+ \to K^-\mu^+\mu^+$	12.9×10^{-6}	7.7%	13.1×10^{-6}
$D^+ \to \pi^+\mu^+\mu^-$	9.6×10^{-6}	7.6%	9.7×10^{-6}
$D^+ \to \pi^-\mu^+\mu^+$	5.0×10^{-6}	7.5%	5.1×10^{-6}
$D_s^+ \to K^+\mu^+\mu^-$	3.5×10^{-5}	27.5%	3.8×10^{-5}
$D_s^+ \to K^-\mu^+\mu^+$	1.2×10^{-5}	27.5%	1.3×10^{-5}
$D_s^+ \to \pi^+\mu^+\mu^-$	2.5×10^{-5}	27.7%	2.7×10^{-5}
$D_s^+ \to \pi^-\mu^+\mu^+$	3.1×10^{-5}	27.6%	3.4×10^{-5}



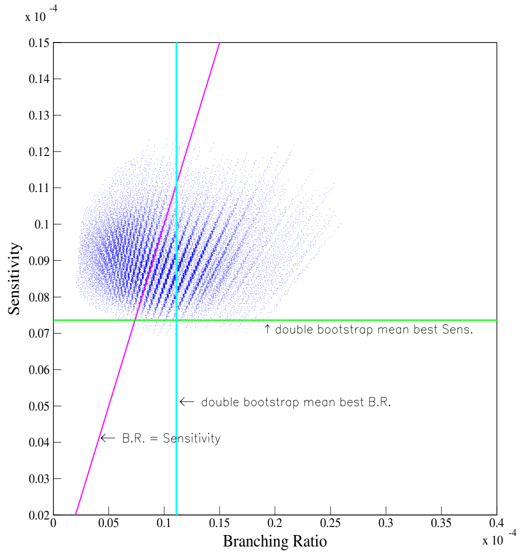
(a) Double Bootstrap Results



(b) Sideband vs. Signal Spread

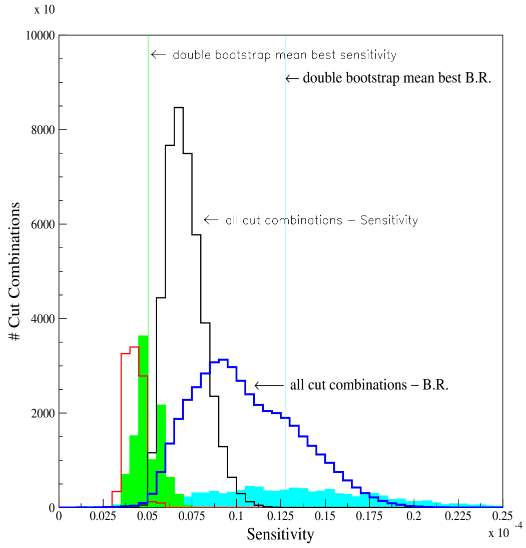


(c) Rolke-Lopez Sens. vs. B.R.

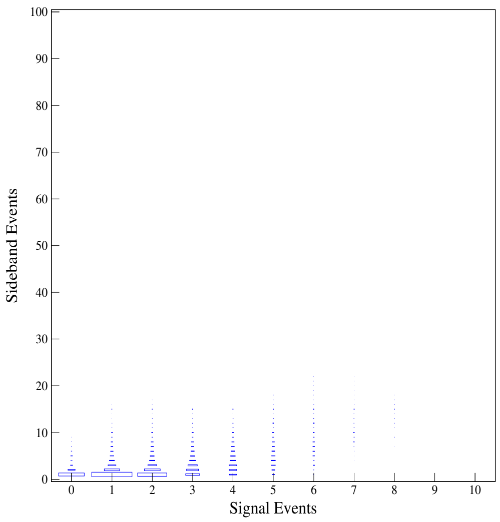


(d) Feldman-Cousins Sens. vs. B.R.

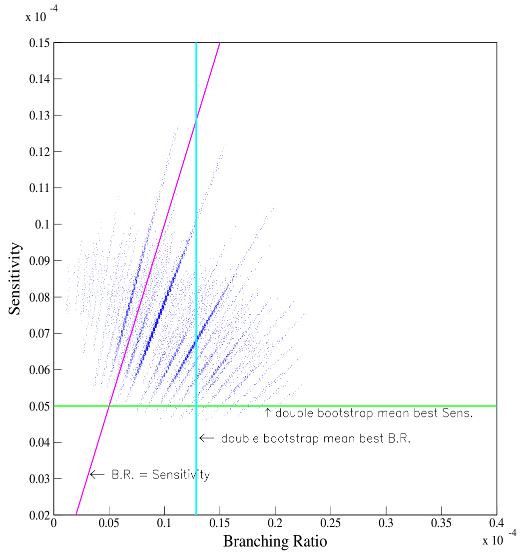
Figure 83: $D^+ \to K^+\mu^+\mu^-$ Branching Ratio Analysis



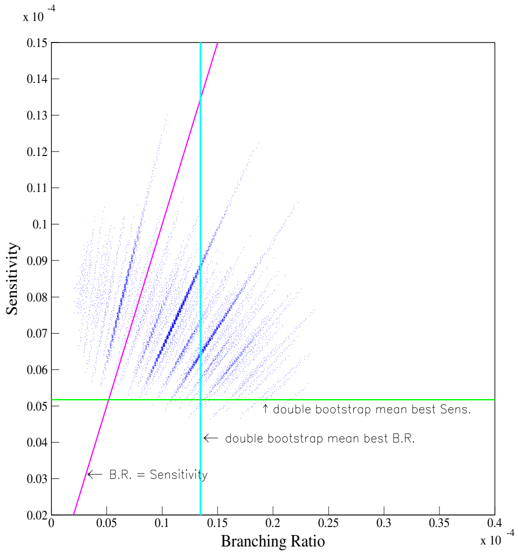
(a) Double Bootstrap Results



(b) Sideband vs. Signal Spread

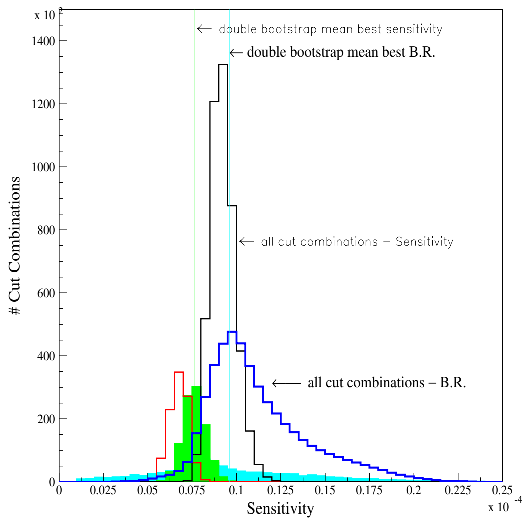


(c) Rolke-Lopez Sens. vs. B.R.

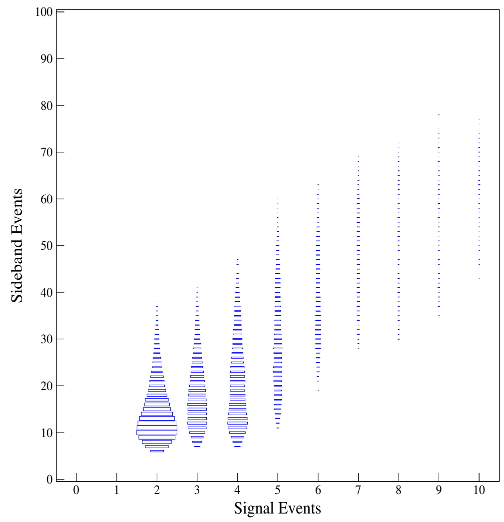


(d) Feldman-Cousins Sens. vs. B.R.

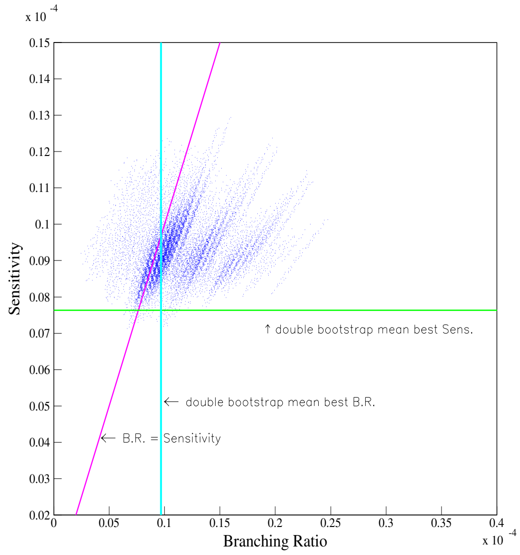
Figure 84: $D^+ \to K^-\mu^+\mu^+$ Branching Ratio Analysis



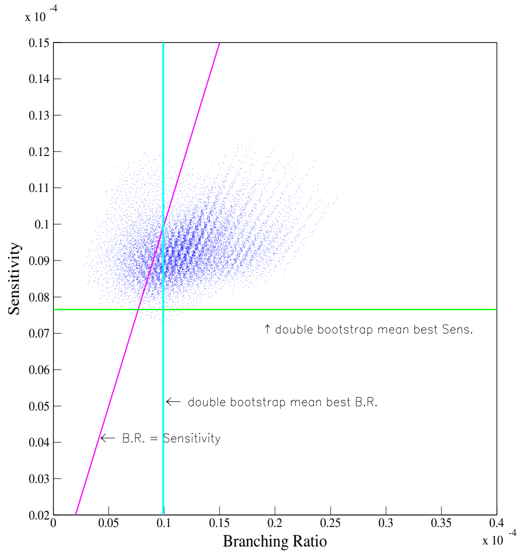
(a) Double Bootstrap Results



(b) Sideband vs. Signal Spread

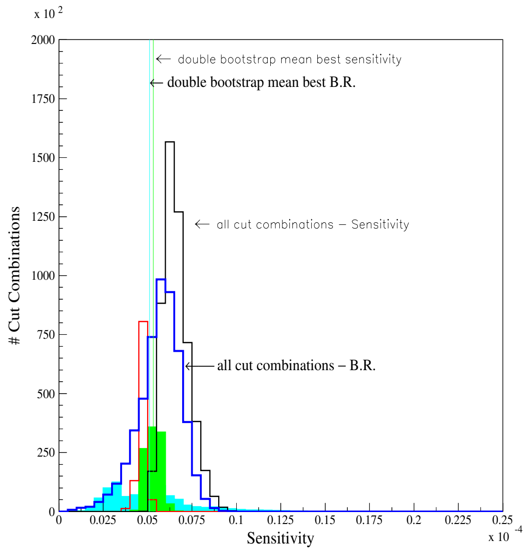


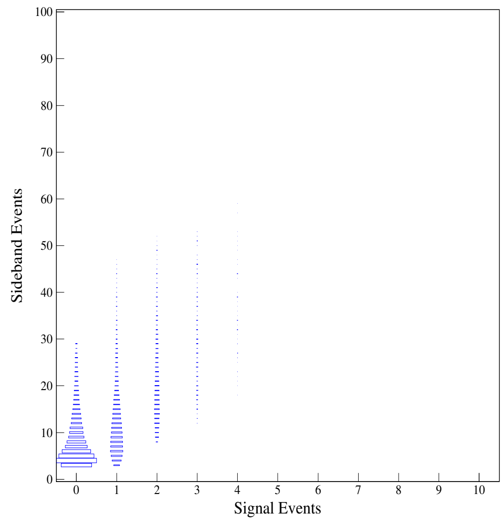
(c) Rolke-Lopez Sens. vs. B.R.



(d) Feldman-Cousins Sens. vs. B.R.

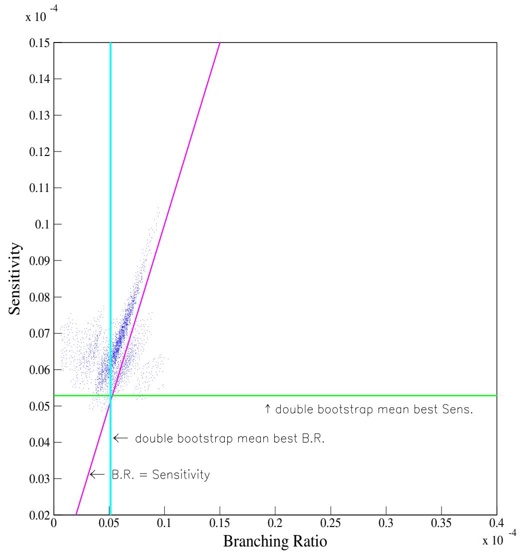
Figure 85: $D^+ \to \pi^+\mu^+\mu^-$ Branching Ratio Analysis

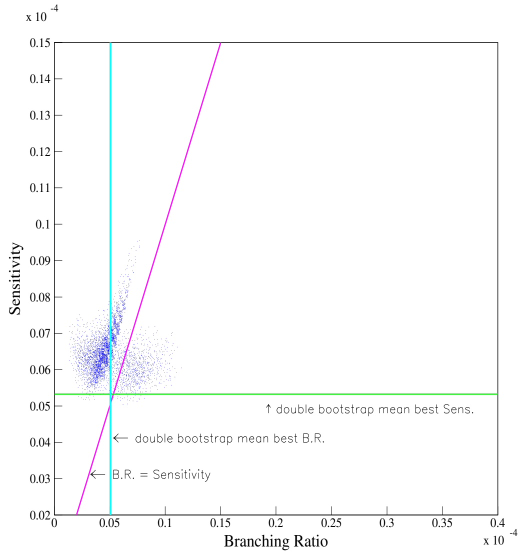




(a) Double Bootstrap Results

(b) Sideband vs. Signal Spread

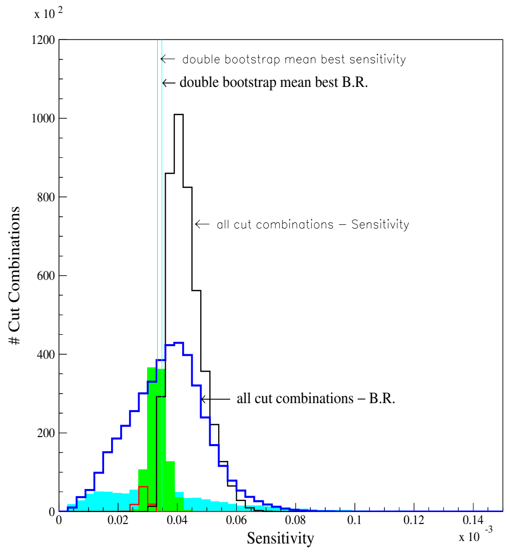




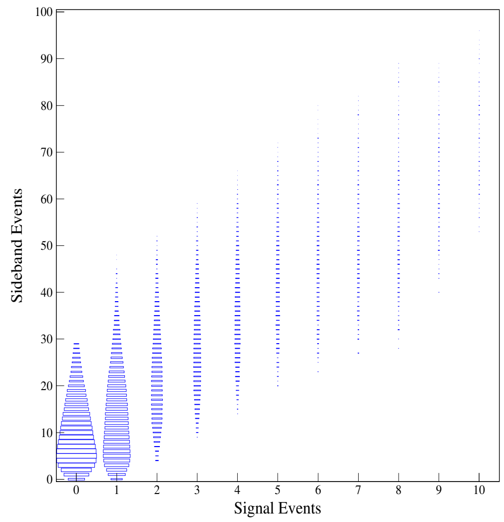
(c) Rolke-Lopez Sens. vs. B.R.

(d) Feldman-Cousins Sens. vs. B.R.

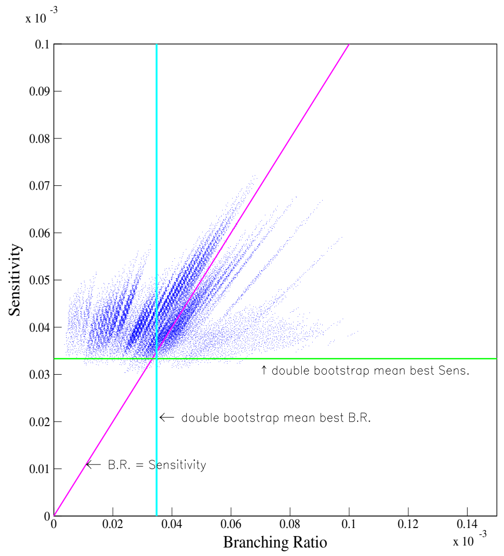
Figure 86: $D^+ \to \pi^-\mu^+\mu^+$ Branching Ratio Analysis



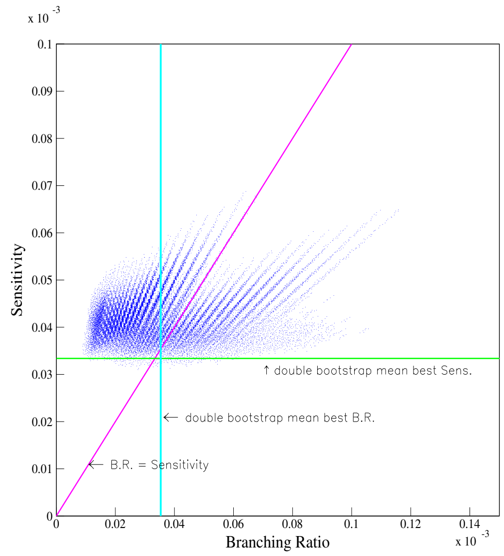
(a) Double Bootstrap Results



(b) Sideband vs. Signal Spread

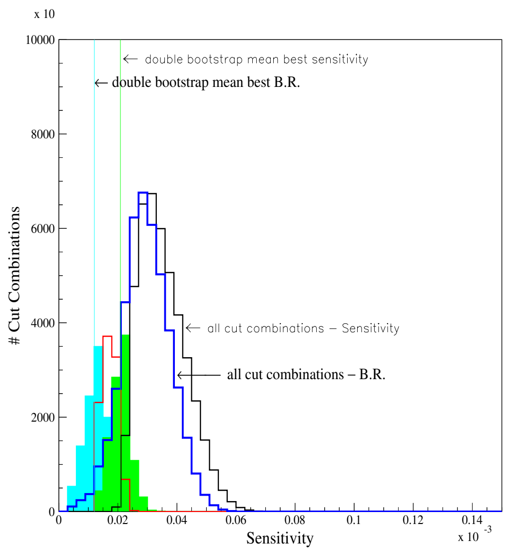


(c) Rolke-Lopez Sens. vs. B.R.

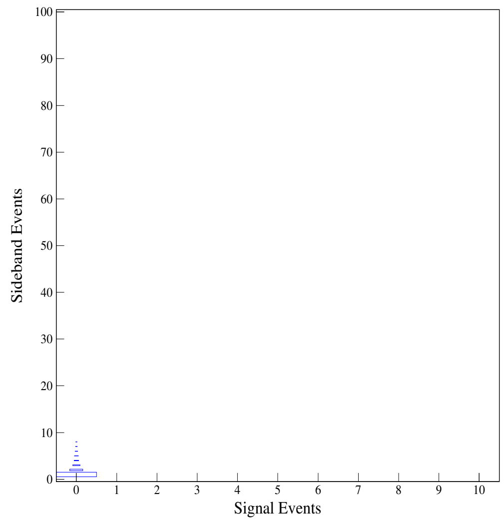


(d) Feldman-Cousins Sens. vs. B.R.

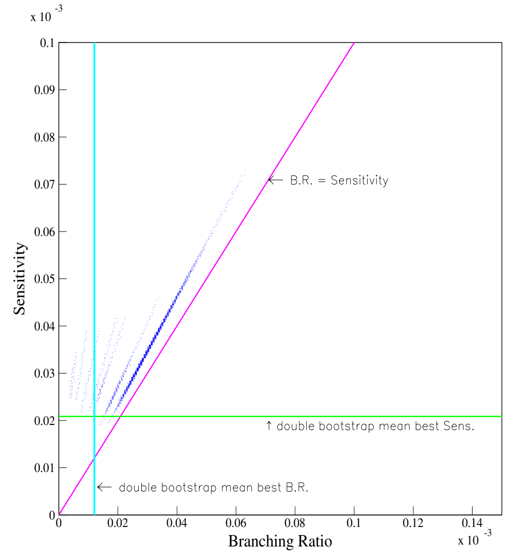
Figure 87: $D_s^+ \to K^+\mu^+\mu^-$ Branching Ratio Analysis



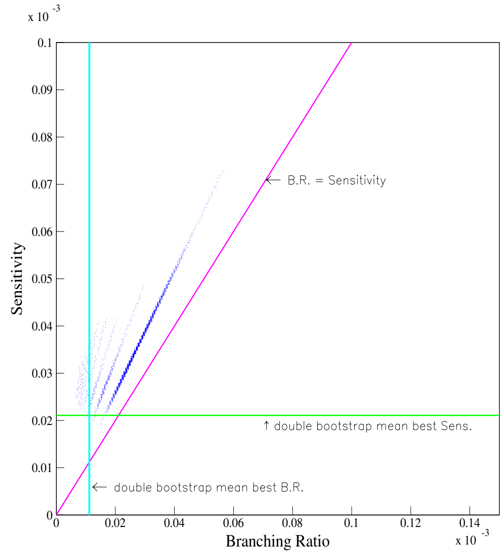
(a) Double Bootstrap Results



(b) Sideband vs. Signal Spread

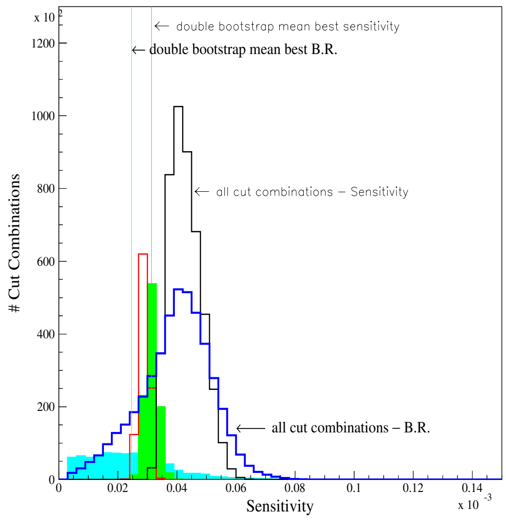


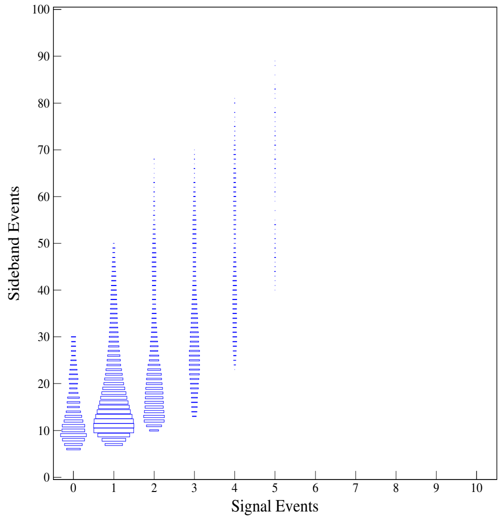
(c) Rolke-Lopez Sens. vs. B.R.



(d) Feldman-Cousins Sens. vs. B.R.

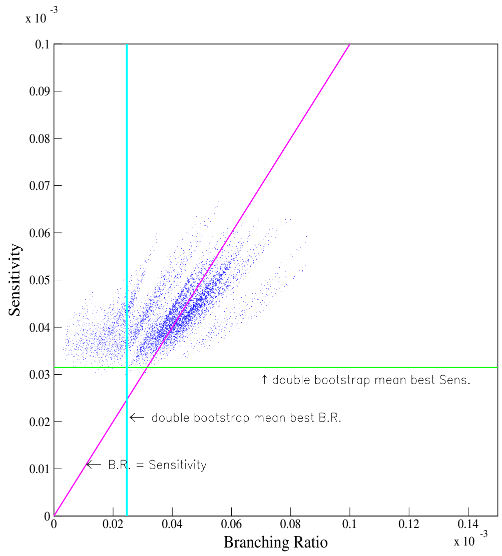
Figure 88: $D_s^+ \to K^-\mu^+\mu^+$ Branching Ratio Analysis

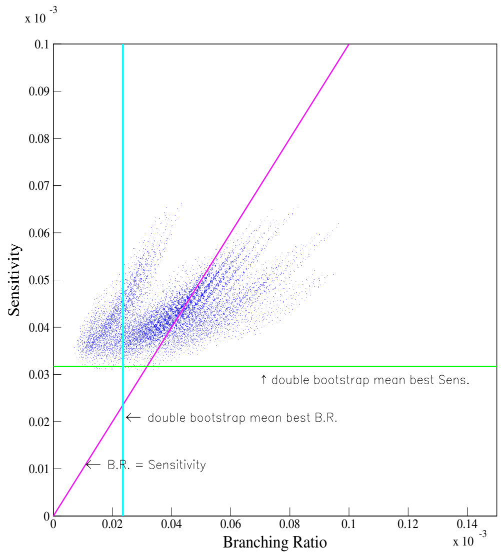




(a) Double Bootstrap Results

(b) Sideband vs. Signal Spread

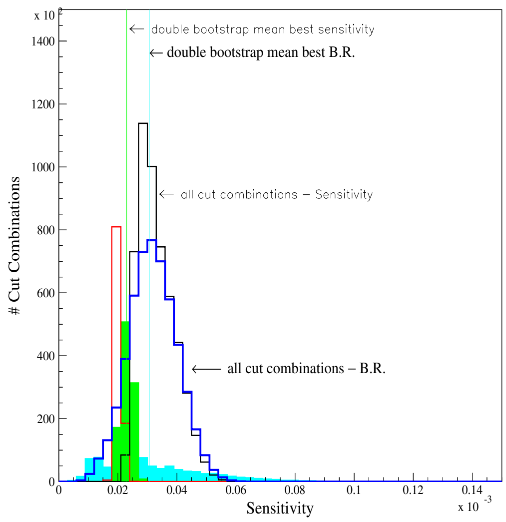




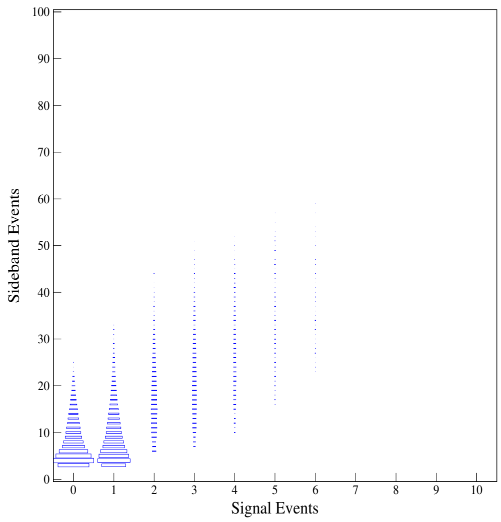
(c) Rolke-Lopez Sens. vs. B.R.

(d) Feldman-Cousins Sens. vs. B.R.

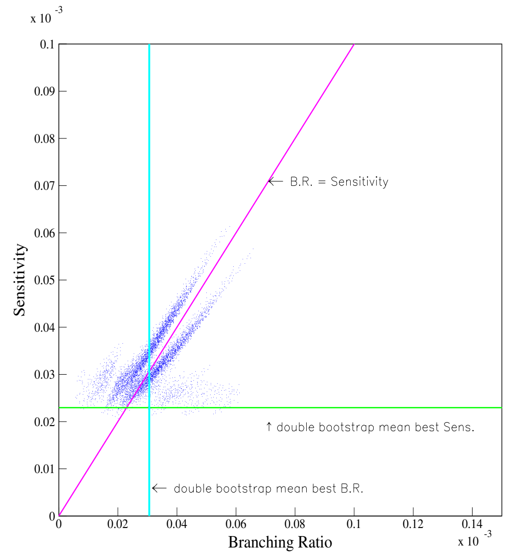
Figure 89: $D_s^+ \to \pi^+\mu^+\mu^-$ Branching Ratio Analysis



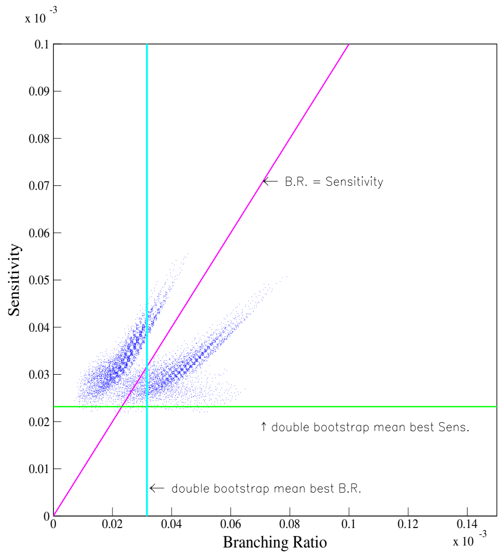
(a) Double Bootstrap Results



(b) Sideband vs. Signal Spread

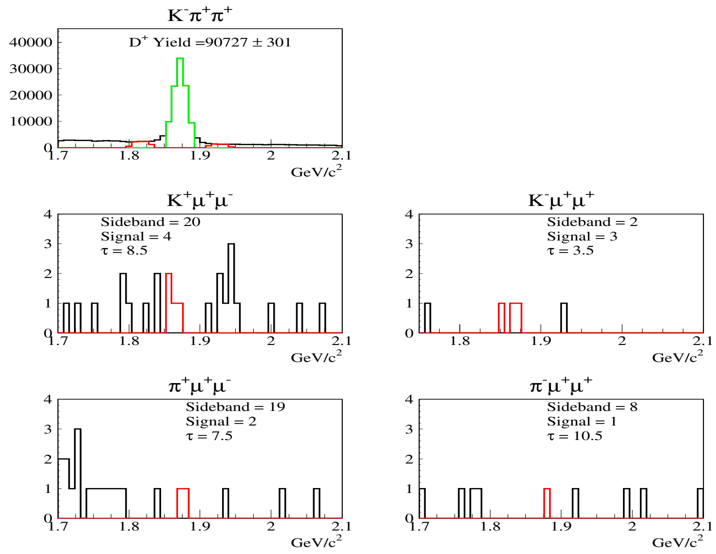


(c) Rolke-Lopez Sens. vs. B.R.

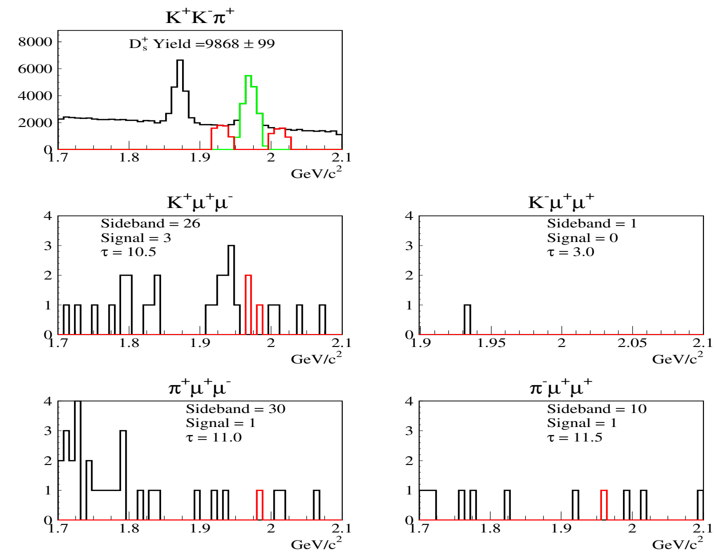


(d) Feldman-Cousins Sens. vs. B.R.

Figure 90: $D_s^+ \to \pi^-\mu^+\mu^+$ Branching Ratio Analysis



(a) D^+ Decay Modes



(b) D_s^+ Decay Modes

Figure 91: Results with the average optimal bootstrap cut. The bootstrap best cuts are combined for all D decay modes and separately for all D_s decay modes. This provides a comparison to other experiments which may have used a single set of cuts for all decay modes, but does not have the reduced bias of the double bootstrap results.

Table 18: Average Bootstrap Cuts for D^+ (D_s^+)

Variable	Cut
ℓ/σ_ℓ	> 13 (10)
Primary Isol.	< 0.1
Secondary CL	$> 1\%$ (2%)
Kaonicity	> 1.0
Inner Muon CL	$> 5\%$
Outer Muon CL	$> 5\%$
inner mom. (GeV)	> 9
outer mom. (GeV)	> 7

6.3.2 Average Bootstrap Cuts

A result can be derived from a single set of cuts for a comparison with other experiments that quote single cut combination results. A single set of cuts also provides an illustration of results with mass plots. As discussed earlier, the problem of selecting a single set of cuts has the potential for bias arising from subjective cut selection criteria. One objective method for choosing a single set of cuts can be defined as follows: combine the histograms of the bootstrap best cuts (cf. Figure 66) for all the decay modes (except the $\mu\mu\mu$ modes) to get the average bootstrap best cut for all decay modes. Selecting the cut closest to the average cut provides an objective definition for a cut combination that is appropriate for all decay modes. This cut selection also has the added feature of using the bootstrap test for bias. In the spirit of a blind analysis, this selection of cuts was done once, in a blind manner.

For all decay modes several cuts are clearly preferred by the bootstrap test. The unambiguous cut selections are track momentum, muon confidence level, kaonicity, and primary vertex isolation. Two cuts are more variable: a secondary isolation level of 1% or 2% is selected with the bootstrap, and ℓ/σ_ℓ varies with the bootstrap. The modes with same sign muons have less data and individually tend to select looser ℓ/σ_ℓ cuts than modes with opposite sign muons. An average over all decay modes results in a relatively loose ℓ/σ_ℓ cut for opposite sign muons and a relatively tight ℓ/σ_ℓ cut for same sign muons as can be seen in Figure 91. The results for the average bootstrap cuts are shown in Table 19.

As can be seen in the mass plots of Figure 91, the decay modes with opposite sign muons contain more events than the decay modes with same sign muons. The averaging of cuts is done separately for D^+ and D_s^+ decays. Since the cuts are averaged over many decay modes, the single cut combination is not optimal for any of the decay modes. The cuts are loose for the decay modes with opposite sign muons and tight for the decay modes with same sign muons. In general, loose cuts result in a wide uncertainty in the branching ratio.

The average cut result for the decay mode $D^+ \to \pi^+\mu^+\mu^-$ is an improvement branching ratio over the bootstrapped branching ratio. This result could be compared to results from E791 which use a similar method for determining a single cut combination. But it should be recognized that this improvement is essentially random and that a single cut result does not accurately reflect the full "discriminating power" of the data.

6.3.3 The $\mu\mu\mu$ Decay Modes

The $\mu\mu\mu$ decay modes are treated separately since there is too little data to use a cut grid or bootstrap analysis. The triple muon id requirement is a very tight cut that nearly eliminates all background, leaving only one sideband event after the minimum

Table 19: Average Cut Results

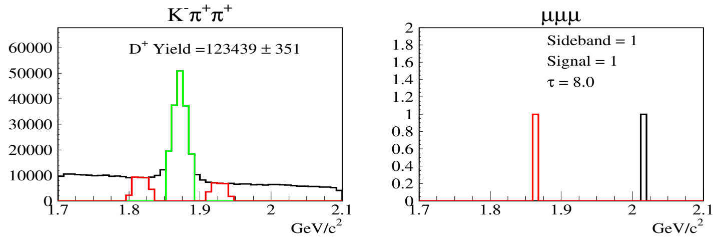
Decay Mode	SB	Sig	τ	R.L. 90%	σ_r	C.H. 90%	Rel Eff.	BR
$D^+ \to K^-\pi^+\pi^+$			90,727 events					
$K^+\mu^+\mu^-$	20	4	8.5	5.96	7.5%	6.01	1.970	11.8×10^{-6}
$K^-\mu^+\mu^+$	2	3	3.5	6.28	7.7%	6.35	1.972	12.4×10^{-6}
$\pi^+\mu^+\mu^-$	19	2	7.5	4.00	7.6%	4.03	1.865	7.5×10^{-6}
$\pi^-\mu^+\mu^+$	8	1	10.5	2.91	7.5%	2.93	1.778	5.2×10^{-6}
$D_s^+ \to K^+K^-\pi^+$			9,868 events					
$K^+\mu^+\mu^-$	26	3	10.5	4.40	27.5%	4.81	1.778	3.8×10^{-5}
$K^-\mu^+\mu^+$	1	0	3.0	2.35	27.5%	2.56	1.783	2.0×10^{-5}
$\pi^+\mu^+\mu^-$	30	1	11.0	2.18	27.7%	2.32	1.737	1.8×10^{-5}
$\pi^-\mu^+\mu^+$	10	1	11.5	2.80	27.6%	3.02	1.653	2.2×10^{-5}

Table 20: Cuts for $\mu\mu\mu$ Decay Modes

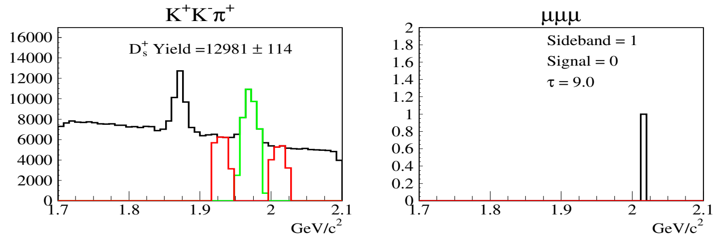
Variable	Cut
ℓ/σ_ℓ	> 5
Primary Isol.	< 0.1
Secondary CL	$> 1\%$
Kaonicity	> 0.5
Inner Muon CL	$> 1\%$
Outer Muon CL	$> 1\%$
inner mom. (GeV)	> 8
outer mom. (GeV)	> 6

skim cuts. Given this situation the results are simply determined with the minimum skim cuts as summarized in Table 20. The branching ratios that result are plotted in Figure 92 and summarized in Table 21.

As can be seen in the mass plots shown in Figure 92, there is one sideband event. The is an event in the D^+ signal region but no events in the D_s^+ signal region.



(a) D^+ Decay Modes



(b) D_s^+ Decay Modes

Figure 92: 3mu Results

Table 21: $\mu\mu\mu$ Results

Decay Mode	SB	Sig	τ	R.L. 90%	σ_r	C.H. 90%	Rel Eff.	BR
$D^+ \to K^-\pi^+\pi^+$	123,439 events							
$\mu^+\mu^+\mu^-$	1	1	8.0	3.53	11.5%	3.59	3.1	8.2×10^{-6}
$D_s^+ \to K^+K^-\pi^+$	12,981 events							
$\mu^+\mu^+\mu^-$	1	0	9.0	2.50	28.8%	2.76	2.9	2.7×10^{-5}

6.4 Conclusion

The combined results are shown in Table 22. All decay channels set new limits, by a factors of approximately 1.5 to 10 for all decay channels. No new physics signals are seen, and limits consistent with the standard model are set.

The broad distribution of branching ratios seen for cuts at optimal sensitivities demonstrates the need for using a method for using this distribution of results to generate an answer rather than a single random sampling of possible results. The double bootstrap mechanism generates results that are conservative but also that objectively test the limits of the data, rather than cutting too tightly, cutting too loosely, or cutting randomly. The analysis method illustrates the possible application of bootstrap methods to other high energy physics analyses with small signals.

Table 22: Results Summary

Decay Mode	Single Cut Result	Bootstrap Result	Previous Limit
$D^+ \to K^+\mu^+\mu^-$	11.8×10^{-6}	10.4×10^{-6}	4.4×10^{-5}
$D^+ \to K^-\mu^+\mu^+$	12.4×10^{-6}	13.1×10^{-6}	12.0×10^{-5}
$D^+ \to \pi^+\mu^+\mu^-$	7.5×10^{-6}	9.7×10^{-6}	1.5×10^{-5}
$D^+ \to \pi^-\mu^+\mu^+$	5.2×10^{-6}	5.1×10^{-6}	1.7×10^{-5}
$D^+ \to \mu^+\mu^+\mu^-$	8.2×10^{-6}	—	—
$D_s^+ \to K^+\mu^+\mu^-$	3.8×10^{-5}	3.8×10^{-6}	1.4×10^{-4}
$D_s^+ \to K^-\mu^+\mu^+$	2.0×10^{-5}	1.3×10^{-6}	1.8×10^{-4}
$D_s^+ \to \pi^+\mu^+\mu^-$	1.8×10^{-5}	2.7×10^{-6}	1.4×10^{-4}
$D_s^+ \to \pi^-\mu^+\mu^+$	2.2×10^{-5}	3.4×10^{-6}	0.8×10^{-4}
$D_s^+ \to \mu^+\mu^+\mu^-$	2.7×10^{-5}	—	—

REFERENCES

[1] F. Halzen and A. D. Martin, *Quarks & Leptons: An Introductory Course in Modern Particle Physics*, John Wiley & Sons, Inc., New York, 1984.

[2] C. Caso et al., Eur. Phys. J. **C3**, 1 (1998).

[3] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963).

[4] J. I. S. L. Glashow and L. Maiani, Phys. Rev. **D2**, 1285 (1970).

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

[6] P. Singer and D.-X. Zhang, Phys. Rev. **D55**, 1127 (1996).

[7] A. J. Schwartz, Modern Phys. Lett. **A8**, 967 (1993).

[8] T. Inami and C. S. Lim, Prog. Theor. Phys. **65**, 297 (1981).

[9] T. J. Weiler, 2001, Private communication.

[10] J. H. G. Burdman, E. Golowich and S. Pakvasa, (2001).

[11] J. Cao, *Search for Physics Beyond the Standard Model Through Rare and forbidden Decays of the Charmed Meson D^+*, PhD thesis, Vanderbilt University, 1997.

[12] e. a. E. M. Aitala, Phys. Lett. **B462**, 401 (1999).

[13] S. P. Martin, (1999), This article, *A Supersymmetry Primer*, provides a good introduction to SUSY.

[14] T. J. Weiler, 2002, Private communication.

[15] FOCUS Collaboration, `http://www-focus.fnal.gov` is the Fermilab FOCUS web site.

[16] FOCUS Collaboration, Description and performance of the fermilab focus spectrometer, 1999, `http://www-focus.fnal.gov/nim/focus-nim/nim.ps`, In preparation for publication in *Nucl. Instrum. Meth.*

[17] P. L. Frabetti et al., Nucl. Instrum. Meth. **A320**, 519 (1992).

[18] P. L. Frabetti et al., Nuclear Instruments and Methods **A329**, 62 (1993).

[19] G. Alimonti et al., Nucl. Instrum. Meth. **A314**, 411 (1992).

[20] G. Bellini et al., Nucl. Instrum. Meth. **A252**, 366 (1986).

[21] J. D. Jackson, *Classical Electrodynamics*, John Wiley & Sons, Inc., New York, second edition, 1975.

[22] FOCUS Collaboration, Čerenkov particle indentification in FOCUS, 2001, hep-ex/0108011, submitted to *Nucl. Instrum. Meth.*

[23] L. Cinquini, J. Cumalat, E. Erdos, B. O'Reilly, and E. Vaandering, A description of the Inner Electromagnetic Calorimeter for FOCUS, FOCUS memo, `http://www-hep.colorado.edu/~e687/focus_doc/iecal.ps`.

[24] S. Bianco et al., The upgraded outer EM calorimeter of FOCUS at Fermilab, 1999, hep-ex/9912066, presented at the *VIII International Conference on Calorimetry in HEP, June 13-19, 1999, Lisbon (Portugal).*

[25] G. Bonomi, An iron/scintillator tile calorimeter for the FOCUS experiment at Fermilab, in *Calorimetry in high energy physics. Proceedings, 7th International Conference, Tucson, USA, November 9-14, 1997*, edited by E. Cheu, T. Embry, J. Rutherfoord, and R. Wigmans, pages 276–280, Singapore, Singapore: World Scientific, 1998.

[26] G. Boca et al., Nucl. Instrum. Meth. **A409**, 561 (1998).

[27] J. Wiss, FOCUS MH homepage, `http://www.hep.uiuc.edu/e687/muon/muon.html`.

[28] P. Liguori, P. Sheldon, P. Vitulo, M. Webster, and J. Wilson, OM design and considerations, 1994, FOCUS memo, `http://www-focus.fnal.gov/papers/outermu_1.ps`.

[29] A. Kreymer and F. Prelz, DAQ architecture for E831, 1997, FOCUS memo, `http://www-focus.fnal.gov/papers/daq_main/daq_main/daq_main.html`.

[30] R. L. Culbertson, *Four-Body Semileptonic Decays of D Mesons*, PhD thesis, University of Illinois at Urbana-Champaign, 1993.

[31] A. M. Rahimi, *Amplitude Analyses of the* $D^+, D_s^+ \rightarrow K^-K^+\pi^+$ *and* $D^0 \rightarrow K_s^0 K^+ K^-$ *Final States*, PhD thesis, University of Illinois at Urbana-Champaign, 2000.

[32] E. Vaandering, *Mass and Width Measurements of* Σ_c *Baryons*, PhD thesis, University of Colorado at Boulder, 2000.

[33] J. Wiss, Thoughts on muon identification algorithms for E831, 1994, FOCUS memo, `http://web.hep.uiuc.edu/e687/muon/mu_id_thoughts.ps`.

[34] C. Cawlfield, M. Ruesnink, and J. Wiss, Muon identification χ^2 confidence levels, 1994, FOCUS memo, `http://web.hep.uiuc.edu/e687/muon/chisq.ps`.

[35] M. G. Hosack, Confidence level for outer muons (2nd ed.), 1998, FOCUS memo, `http://www-focus.fnal.gov/papers/omu_conf2.ps`.

[36] J. Wiss and R. Gardner, Estimating systematic errors, 1994, E687 memo: E687-94-030, `http://web.hep.uiuc.edu/e687/memos/DALITZ_SYS.PS`.

[37] CERN Application Software Group, A summary of the HBOOK software package can be found at `http://wwwinfo.cern.ch/asdoc/hbook/HBOOKMAIN.html`.

[38] G. J. Feldman and R. D. Cousins, Phys. Rev. **D57**, 3873 (1998).

[39] J. Neyman, Trans. R. Soc. London **A236**, 333 (1937).

[40] C. Giunti, Phys. Rev. **D59**, 053001 (1999).

[41] B. P. Roe and M. B. Woodroofe, Phys. Rev. **D60**, 053009 (1999).

[42] R. D. Cousins and V. L. Highland, Nucl. Instrum. Meth. **A320**, 331 (1992).

[43] W. Rolke and A. Lopez, Nucl.Instrum.Meth. **A458**, 745 (2001).

[44] A. J. Schwartz, (1997), From PRINCETON/HEP/97-6, OFFLINE_DOC_263.

[45] W. A. Rolke and A. M. Lopez, (2002), *Correcting the Minimization Bias in Searches for Rare Decays* — document in preparation.