MIXING AND DOUBLY CABIBBO-SUPPRESSED DECAYS OF NEUTRAL CHARMED MESONS

BY

MICHAEL ALAN MARSH

A.B., Cornell University, 1994
M.S., University of Illinois at Urbana-Champaign, 1995

THESIS

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

Urbana, Illinois

MIXING AND DOUBLY CABIBBO-SUPPRESSED DECAYS OF NEUTRAL CHARMED MESONS

Michael Alan Marsh, Ph.D.
Department of Physics
University of Illinois at Urbana-Champaign, 2001
Prof. Gary Gladding, Advisor

A study of the decay $D^0 \rightarrow K^+\pi^-\pi^+\pi^-$ is presented. This "wrong-sign" decay can proceed through either a doubly Cabibbo-suppressed decay or mixing to a $\overline{D}^0$ followed by the Cabibbo-favored (or "right-sign") decay $D^0 \rightarrow K^-\pi^+\pi^-\pi^+$. The Cabibbo-favored decay is used as a normalization for the wrong-sign decay. A detailed description of the fitting procedure, including extensive tests, is also included.

Acknowledgements

In my seven years of graduate study, numerous people have helped to make my research possible, bearable, and (frequently) fun. An exhaustive list would be far too long, but a number of them bear explicit mention.

My advisor, Gary Gladding, certainly needs to be mentioned. Without his guidance and knowledge, this work would not have been possible. He has always maintained a positive attitude even through my most muddled attempts to explain what I was doing and why.

It has been a privilege and a pleasure to work with all of the Illinois CLEO professors and researchers. I am especially endebted to Professors Mats Selen and Jon Thaler for their contributions. I have also benefitted greatly from the former and current Illinois CLEO postdocs Mark Palmer, Jesse Ernst, Randal Hans, and Kamal Benslama. Randal, in particular, has been an invaluable asset as a consultant regarding my analysis, as coordinator of the CLEO charm mixing working group, and as a friend. I would also like to thank the former and current Illinois CLEO graduate students Tim Bergfeld, Ed Johnson, Chris Sedlack, Charles Plager, Jeremy Williams, and Norm Lowrey for advice, assistance, and camaraderie.

I would like to thank the entire CLEO charm mixing group, past and present, for their input and help. Dave Asner and Alex Smith's work on similar decay modes greatly smoothed out many of the bumps in the road for this analysis. Harry Nelson, Dave Cinabro, Craig Prescott, and Sean McGee provided valuable insights. Until his departure (and somewhat slightly beyond), Tony Hill provided a constant impetus and source of enthusiasm for the entire group.

Mats Selen, Alex Smith, and Steve Pappas served on my paper committee, for which I am grateful. Mats, Randal, Giancarlo Monetti, and Dave Besson also deserve mention as chairs of the Charm PTA at CLEO.

As for my software work, I would like to thank the software coordinators who facilitated my job: John O'Neill, Rich Baker, Lawrence Gibbons, and Dave Kreinick. I would also like to thank Craig Prescott, my predecessor as software librarian, for training me and leaving a stable system, and Xin Zhao, my successor, for taking over. Deserving of special mention are Lawrence Gibbons, who was instrumental in having me relieved of my CLEO II/II.V librarian duties; Mark Palmer, who was instrumental in having me relieved of my CLEO III software development duties (supervised by Lawrence); and Gary Gladding, who was instrumental in having me relieved of my inner tracker development duties (supervised by Mark). I would especially like to thank Simon Patton, Chris Jones, and Martin Lohner both for their development of an easy-to-use software infrastructure and for taking the time to explain enough of it to me that I could (mostly) understand how it works.

I would like to thank the friends who made the Misspent Youth of my graduate career enjoyable. My high school friends Marco Cavagna, Tom Mon, and Chris Shaw have been a constant link to simpler times. Brynnen Owen and Clark Wagner showed me how to brew beer, which Amy Oldenburg, Paul Kennedy, and William Tucker helped us consume, occasionally while trying to think of humorous (and often tasteless) ideas for the UIUC Physics department's annual "Physical Revue." Martin and Raghida Lohner, Samina Jafri, Peter Gaidarev and Tatyana Pyntikova, Tony Rubiera, Basit Athar, Chris Sedlack, Casey Westerman, Norm Lowrey, Kamal Benslama, Adi Bornheim and Steve Pappas helped keep the Friday night Chapter House tradition going. Casey in particular, as a non-physicist, was subjected to a large number of mind-numbingly boring conversations. I would also like to thank Julie Elefante for keeping me smiling from afar.

Finally, I would like to thank my parents Howard and Lynn, and my sister Ivy

for their constant support throughout my life. My father never wanted me to follow in his footsteps in studying Physics, but he hasn't seemed to hold that against me.

This material is based upon work supported by the National Science Foundation under Grant PHY 95-53157 FFW, the Department of Energy under Grant DE FG02 91ER40677, and the Department of Education under the GAANN Fellowship Program.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Mixing is a basic consequence of quantum mechanics. Quantum mechanical operators define eigenstates, fixed points of their interactions, which provide bases that can be used to describe the space. A simple example of this is the classic Stern-Gerlach experiment for separating particles based on the projections of their spins onto a particular axis. It is frequently the case that there will be two or more states which are orthogonal under an interaction, but otherwise identical, such as two electrons in the same atomic orbital but having opposite spin. If there is another operator for which these states are not fixed points, the action of this second operator will tend to mix the initial states. Successive Stern-Gerlach devices oriented along different axes is an example of this.

In elementary particle Physics, mesons form as bound eigenstates of the strong interaction Hamiltonian. Ground-state meson decay is goverened by the total Hamiltonian, however, which incorporates the electroweak interaction as well. For mesons which are not their own antiparticles, the situation is analogous to the paired Stern-Gerlach devices. The strong interaction first separates states into meson and anti-meson, and the slower electroweak interaction then rotates them into the basis that describes their decays [1].

Mixing was first observed in 1956 in the decays of $K^0(\bar{s}d)$ mesons[2]. The neutral K^0 meson was known to decay to two pions, and it was predicted that a longer-lived partner of approximately the same mass would decay to three particles. These two particles were labelled K_1^0 and K_2^0, later changed to K_S^0 and K_L^0 to emphasise their "short" and "long" lifetimes. In 1987, mixing was observed in $B^0(\bar{b}d)$ mesons, but this time there was a larger split in the masses of the eigenstates B_1^0 and B_2^0 than in their lifetimes[3].

One thing that the $K^0 - \overline{K}^0$ and $B^0 - \overline{B}^0$ systems have in common is that both sets of mesons contain a light quark (in both cases a d) and a heavier quark (s or b). If the heavy quark is instead a charmed quark (c), it can be paired with a light u quark to form a D^0 or $\overline{D}^0$ meson. These too should mix, though this has never been observed. The Standard Model predicts that the rate of mixing (R_{mix}) is at most 10^{-6} [4]. That means at least one million D^0 or $\overline{D}^0$ mesons would have to be produced for every one $D^0 - \overline{D}^0$ mixing event. The mixing rates for K and B mesons, in comparison, are approximately 0.5 and 0.2, respectively.

In order to search for mixing, we must choose a specific final state to reconstruct. These fall into three general categories: states accessible only to either a D^0 or $\overline{D}^0$; states accessible equally to both; and states accessible to both, but at different levels.

The first category is the simplest to search for, in some sense. If you can determine that the charmed meson was created as a D^0 but decayed as a $\overline{D}^0$, then it is an unambiguous signal of mixing. These are semileptonic modes, where the charge of the lepton is constrained by the charge of the virtual W boson emitted by the c (or $\bar{c}$) quark. Until recently, this has been the most popular way to search for mixing, and produces excellent limits. The best semileptonic result comes from a search for $D^0 \rightarrow K^+ l^- \nu_l$, which limits the mixing rate to less than 0.5% at the 90% confidence level [5].

The second category is also simple. It does not require knowing whether the initial meson was a D^0 or a $\overline{D}^0$, since both produce (approximately) equal amounts of the final state. These are CP eigenstates, where the final state particles are either themselves CP eigenstates or paired with their antiparticles, such as K^+K^- and $K^+K^-\pi^0$. In this case, we look at the lifetimes, and try to constrain the amount of mixing related to lifetime splitting. This is not sensitive to mixing related to mass splitting, however. If we do know what the initial meson is, we can also compare the relative amounts of the final state from D^0 and $\overline{D}^0$. This gives us an asymmetry which can be related to the CP violation[1] in the D^0 system. The lifetime-splitting part of the mixing amplitude, which is referred to as y, has recently been measured by FOCUS [7] and E791 [8] in the mode $D^0 \to K^+K^-$. The average of their results is $y = (2.9 \pm 1.4)\%$, and new results are expected soon which will further constrain y.

The third category is more complicated, and contains the fully hadronic decay modes that are not CP eigenstates. As with the semileptonic modes, we need to determine whether we initially had a D^0 or $\overline{D}^0$ meson. There is no clear mixed final state, however, so we will see both mixing and the lower-rate process. The lifetime provides additional information, however. The mixed and unmixed events have different lifetime distributions. This, in itself, would only help if mixing were already large enough to be seen in semileptonic modes. Where this method gains sensitivity is in the interference between mixed and unmixed decays to the same final state. This interference can be comparatively large, even if the actual rate of mixing is small. Searches of this type have only gained popularity since the advent of high-precision detectors, such as silicon pixel or microstrip detectors, which allow the reconstruction of high-precision lifetimes. The current best measurement is in the decay mode $D^0 \to K^+\pi^-$ and comes from CLEO II.V [9]. This analysis measured two different parameters, x' and y', which are components of the mixing amplitude.[2] The limits obtained were $\frac{1}{2}x'^2 < 0.041\%$ and $-5.8\% < y' < 1.0\%$, both of which are at the 95% confidence level.

[1] CP violation will not be discussed here. A good explanation of it can be found in [6].
[2] Section 2.3.1 has a detailed discussion of this notation.

Chapter 2

Theory of $D^0 - \overline{D}^0$ Mixing

2.1 The Standard Model

2.1.1 Fundamental Particles and Interactions

The fundamental particles in the Standard Model are fermions, which form the building blocks of matter, and bosons, which carry the forces between particles. The fundamental fermions have the structure

$$\begin{matrix} \begin{pmatrix} u \\ d \end{pmatrix} & \begin{pmatrix} c \\ s \end{pmatrix} & \begin{pmatrix} t \\ b \end{pmatrix} \\ \begin{pmatrix} e \\ \nu_e \end{pmatrix} & \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix} & \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix} \end{matrix} \tag{2.1}$$

where each of the three columns is a generation of particles with similar properties but increasing mass. Each generation has the charge structure

$$\begin{matrix} \begin{pmatrix} +\frac{2}{3} \\ -\frac{1}{3} \end{pmatrix} \\ \begin{pmatrix} -1 \\ 0 \end{pmatrix} \end{matrix} \tag{2.2}$$

The top pairings are *quarks* (with *flavors* 'up', 'down', 'charmed', 'strange', 'top', and 'bottom') and the bottom *leptons*. The quarks are distinguished from the leptons primarily by their fractional charges and that they carry a property called *color* ('red', 'green', or 'blue'). Each fermion also has an associated antiparticle with opposite quantum numbers (commonly indicated as $\bar{p}$ for a particle p).

The gauge bosons, which govern interactions, are the photon (γ), the charged W's, the Z, and the gluons (g). Photons, massless and possessing no charge, couple charged particles via the electromagnetic force. W and Z bosons carry the weak force, which couples to left-handed particles and right-handed antiparticles. The W is charged, and so comes in two varieties, W^+ and W^-. This means it couples leptons with neutrinos and up-type quarks with down-type quarks. This is noteworthy, as no other interaction couples particles of different types. The Z is neutral, and couples all left-handed particles with themselves, providing (among other things) a creation and annihilation mechanism for $\nu\bar{\nu}$ pairs. The W and Z bosons are massive, and so can only be produced at sufficient energies. Otherwise, they appear as virtual particles, with lifetimes bound by the Uncertainty Principle based on how far "off-shell" their energies are ($\Delta E = M_{W,Z}c^2 - E$). Finally, the gluons are massless and neutral particles which couple via the strong force to particles with color. They themselves possess color and anticolor (the antiparticle version of color), and so couple to themselves. The strong force preserves color, as the electromagnetic force preserves charge (though there are three "charges" to the strong force). It binds quarks together into bound states, and dominates bound-state decays to lower-energy states.

There are two varieties of quark bound states. The first, called a *meson*, contains a quark and an antiquark. These must form a colorless combination, such as red-antired ($r\bar{r}$). Any two flavors can be combined, as long as the quarks themselves have complimentary colors, but heavy-quark combinations will generally decay strongly to bound states containing one heavy quark and light quarks (u or d), unless doing so would violate some conservation law (*ie.*, a D_s meson has insufficient mass to decay strongly to a D and K). Because it is made of two fermions, the meson will be a boson. The second type of bound state is a *baryon*, and contains three quarks (or antiquarks), again in a colorless combination. Because there is no anticolor present, the quarks are constrained so that one must be red, another green, and the third blue. Baryons have the same behavior, in terms of heavy-quark state decays, as mesons. They are fermions, however, since they contain an odd number of fermions. It is also possible to have higher-energy states of both mesons and baryons due to relative angular momentum between the constituent quarks, and some are linear combinations of differing quark states. This produces a rich spectroscopic structure with various selection rules for decays analogous to the atomic system.

2.1.2 Quantum Numbers

As was alluded to before, there are quantum numbers which characterize the fundamental particles and hence their combinations as well. Most familiar is the electric charge Q. This is conserved in all interactions. As is standard in quantum mechanics, we also consider the quantization of angular momentum, given by the quantum numbers J and J_3, which have their usual meanings.

Specific to the Standard Model are *lepton number* and flavor. Each species of lepton; electron, muon, and tau; has its own lepton number, with the following convention:

$$L_\ell = \begin{cases} +1 & ; \; \ell^-, \nu_\ell \\ -1 & ; \; \ell^+, \bar{\nu_\ell} \end{cases} \tag{2.3}$$

Note that for charged leptons we use the charge as a superscript to distinguish particles (negative) from antiparticles (positive). The subscript ℓ denotes the appropriate generation (e, μ, or τ). These lepton numbers are also conserved by all interactions.

As lepton number applies only to leptons, flavor applies only to quarks. The flavor quantum numbers are treated slightly differently from the flavors listed in Section 2.1.1. Because the up and down quarks are very light, well below the characteristic energy scale of the strong force, they obey an approximate symmetry under strong interactions with spin-like characteristics, known as *isospin*. We write the total isospin as I, and its $\hat{z}$ component as I_3. Only u and d quarks (and their antiquarks) carry non-zero isospin. The remaining four quarks are each assigned their own flavor quantum number, which is $+1$ or -1 according to the quark (or antiquark) charge. Expressed as $|I, I_3\rangle$, the up quark is $|1/2, 1/2\rangle$ and the down quark is $|1/2, -1/2\rangle$.

The $\bar{u}$ and $\bar{d}$ have opposite I_3. The strong force preserves isospin, and all forces except the charged weak force preserve flavor in general (that is, they may violate isospin, but will not change an up quark into a down quark).

There are three more standard quantum numbers: charge-conjugation (C), parity (P), and time reversal (T). C is only applicable to particles which are their own antiparticles (such as the γ and certain bound states), and is the eigenvalue of the wavefunction transform taking particle to antiparticle. P is the quantum number of spatial inversion ($\vec{x} \to -\vec{x}$) applied to a particle's wavefunction, and T is the quantum number of time reversal ($t \to -t$) applied to the wavefunction. All three are multiplicative. In general, these are not conserved, though the product CPT is believed to be.

2.1.3 Decays

The Standard Model is a quantum field theory, and as such is Lorentz invariant. The kinematics of particle interactions must therefore conserve four-momentum. Because $E^2 = p^2c^2 + m^2c^4$, this tells us that a particle can only decay into particles with a combined mass less than or equal to its own. This and angular momentum form the dominant effects in terms of experimentally reconstructing and identifying decays.

The customary way to express decays (and other interactions) is through Feynman diagrams. These consist of fermion and boson lines with interactions represented as vertices. A simple electromagnetic interaction is shown in Figure 2.1, and corresponds to an electron radiating a photon.

The quark generations, which are eigenstates of the strong interaction, are not eigenstates of the weak interaction. For this reason, processes involving a $W^\pm$ can couple quarks from different generations. This is effectively a change of basis, and is

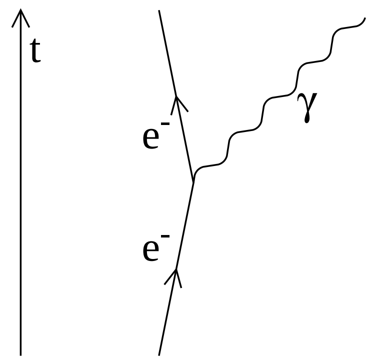


Figure 2.1: Feynman diagram for the interaction $e \to e\gamma$.

represented by the Cabibbo-Kobayashi-Maskawa (CKM) matrix:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{2.4}$$

We then have weak coupling between the u and d', the c and s', and the t and b'. The diagonal elements of the CKM matrix are all near one, and the further from the diagonal an element is, the smaller. This is what allows the lightest mesons and baryons involving a heavier quark to decay further into lighter bound states, as well as annihilating and producing photons or leptons.

It now becomes useful to discuss the particular decay $D^0 \to K^-\pi^+\pi^-\pi^+$. This is a complicated process, and can proceed either directly in a non-resonant fashion or by creating one of several intermediate bound states, all with extremely short lifetimes. In any of these cases, the decays occur via one weak gauge boson and a number of gluon interactions. The wave functions for these various decays overlap to some extent, which causes interference among the different decay channels. Each intermediate resonance is a mesonic bound state which decays to two or three lighter mesons (K or π), which are then observed in the detector. As such, we can attempt to reconstruct these intermediate resonances, and potentially gain some useful information about the phase space structure of the decay. To simplify our notation, we label the final-state particles in the following way: $K^- = 1, \pi^+_{slow} = 2, \pi^- = 3, \pi^+_{fast} = 4$. Here "fast" and "slow" are used to distiguish the two like-charge pions in an unambiguous way based on their total momenta in the lab frame.

For each particle i in the final state, we have a four-momentum p_i^μ which describes its kinematics. Spatial position can be ignored, since all four share a common production point, and the underlying physics is invariant under spatial transformations. All invariants of the system can be expressed in terms of the invariants $p_i^\mu p_{\mu j} (\equiv p_i \cdot p_j)$, of which there are ten $(p_j \cdot p_i \equiv p_i \cdot p_j)$. Four of these are trivial $(p_i \cdot p_i = m_i^2)$, and we assume that our final products have sufficiently well-measured and narrow mass distributions that they can be taken as constants.

We now define masses of the various two- and three-particle combinations (which in some cases will be exactly the masses of intermediate resonances) as $m_{ij}^2 \equiv (p_{i\mu} + p_{j\mu})(p_i^\mu + p_j^\mu)$, and $m_{ijk}^2 \equiv (p_{i\mu} + p_{j\mu} + p_{k\mu})(p_i^\mu + p_j^\mu + p_k^\mu)$. It is obvious that these masses can be expressed as linear combinations of the invariants mentioned above. These, in turn, can be inverted to give the six non-trivial invariants in terms of the masses. This inversion is degenerate, so we must choose a particular "basis" of invariant masses, one of which is:

$$\begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 \end{pmatrix} \begin{pmatrix} p_1 \cdot p_2 \\ p_1 \cdot p_3 \\ p_1 \cdot p_4 \\ p_2 \cdot p_3 \\ p_2 \cdot p_4 \\ p_3 \cdot p_4 \end{pmatrix} = \begin{pmatrix} m_{12}^2 - m_1^2 - m_2^2 \\ m_{14}^2 - m_1^2 - m_4^2 \\ m_{23}^2 - m_2^2 - m_3^2 \\ m_{34}^2 - m_3^2 - m_4^2 \\ m_{234}^2 - m_2^2 - m_3^2 - m_4^2 \\ m_D^2 - m_1^2 - m_2^2 - m_3^2 - m_4^2 \end{pmatrix} \tag{2.5}$$

What is important to note is that many of these masses are what we have taken to be constants, leaving us with four two-particle combinations, one three-particle combination, and the total final state. The combinations have been selected to be physically meaningful. That is, all two-particle combinations are neutral (since we do not expect doubly charged intermediate resonances), and the three-particle combination consists of the three pions (the most likely three-particle resonance combination).

It is furthermore possible to kinematically constrain the final state to have the central value of the measured D^0 mass, which reduces the number of independent phase space variables from six to five.

2.1.4 Oscillations through Common Decay States

Because processes are reversable (time-reversal T, while not a perfect symmetry, is approximately conserved), the matrix element $\langle f \,|H|\, D^0\rangle$ is equal to the matrix element $\langle D^0 \,|H|\, f\rangle$ for some final state f. This means the reaction $D^0 \to f$ could be followed by $f \to D^0$. If f is a final state accessible to both D^0 and $\overline{D}^0$, we can furthermore have the reaction $D^0 \to f \to \overline{D}^0$. In other words, it is possible for a D^0 to become its antiparticle $\overline{D}^0$. It should be noted that f is not merely a collection of final-state particles, but includes their relative kinematics and quantum numbers. Charge conservation, for example, prevents D^+ and D^- mesons from decaying to identical final states, while conservation of momentum prevents D^0 and K^0 mesons, both of which can decay to $\pi^+\pi^-$, from mixing.

2.2 Mixing of Quantum Mechanical States

Most of the general features of mesonic mixing can be seen in a purely quantum-mechanical formulation. We begin by defining the basis of our system, which is constructed of the eigenstates of the strong interaction:

$$\begin{aligned} |D^0\rangle &\equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ |\overline{D}^0\rangle &\equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{aligned} \tag{2.6}$$

We also define the mixing Hamiltonian to be $\mathcal{H}_{mix}$, which has the mass eigenstates $|D_1^0\rangle$ and $|D_2^0\rangle$(following the naming convention of [1]). The time evolution of the system is

$$\begin{pmatrix} |D^0(\tau)\rangle \\ |\overline{D}^0(\tau)\rangle \end{pmatrix} = e^{-i\mathcal{H}_{mix}\tau} \begin{pmatrix} |D^0(0)\rangle \\ |\overline{D}^0(0)\rangle \end{pmatrix} \tag{2.7}$$

while $|D_1^0(\tau)\rangle = |D_1^0(0)\rangle$ and $|D_2^0(\tau)\rangle = |D_2^0(0)\rangle$.

Specifically, we can write $\mathcal{H}_{mix}$ as[10]

$$\mathcal{H}_{mix} = \mathbf{M} - i\frac{\mathbf{\Gamma}}{2} = \begin{pmatrix} M_{11} - i\frac{\Gamma_{11}}{2} & M_{12} - i\frac{\Gamma_{12}}{2} \\ M_{12}^* - i\frac{\Gamma_{12}^*}{2} & M_{11} - i\frac{\Gamma_{11}}{2} \end{pmatrix}. \tag{2.8}$$

$\mathbf{M}$ and $\mathbf{\Gamma}$ are the Hermitian matrices whose eigenvalues give the masses and widths of the mixed states $|D_1^0\rangle$ and $|D_2^0\rangle$(Hermitian because these eigenvalues must be real, as physical observables). The diagonal elements are identical, because the D^0 and $\overline{D}^0$ are antiparticles, and so have the same mass and lifetime. The off-diagonal elements represent the amount of mixing between the two. The eigenvalues of $\mathbf{M}$ are

$$M_{1,2} \equiv M_{11} \pm |M_{12}| \tag{2.9}$$

and the eigenvalues of $\mathbf{\Gamma}$ are similarly defined. For notational convenience, we also define

$$\begin{aligned} M_D &\equiv M_{11} \\ \Gamma_D &\equiv \Gamma_{11} \end{aligned} \tag{2.10}$$

as the mass and width of the strong eigenstates. We can now define the quantities x and y as

$$\begin{aligned} x &\equiv \frac{2M_{12}}{\Gamma_D} \\ y &\equiv \frac{\Gamma_{12}}{\Gamma_D} \end{aligned} \tag{2.11}$$

The eigenvalues of $\mathcal{H}_{mix}$ are

$$\lambda_{1,2} = M_{11} - i\frac{\Gamma_{11}}{2} \pm \sqrt{(M_{12} - i\frac{\Gamma_{12}}{2})(M_{12}^* - i\frac{\Gamma_{12}^*}{2})}. \tag{2.12}$$

The corresponding states are the eigenstates of the weak interaction, and we can decompose the eigenvalues into the masses and widths of these states:

$$\lambda_{1,2} = M_{1,2}^W - i\frac{\Gamma_{1,2}^W}{2} \tag{2.13}$$

which means the (real) physical observables of the weak eigenstates are

$$
\begin{aligned}
M^W_{1,2} &= M_{11} \pm \Re\sqrt{\left(M_{12} - i\frac{\Gamma_{12}}{2}\right)\left(M^*_{12} - i\frac{\Gamma^*_{12}}{2}\right)} \\
\Gamma^W_{1,2} &= \Gamma_{11} \mp 2\Im\sqrt{\left(M_{12} - i\frac{\Gamma_{12}}{2}\right)\left(M^*_{12} - i\frac{\Gamma^*_{12}}{2}\right)}
\end{aligned}
\tag{2.14}
$$

We now separate $\mathcal{H}_{mix}$ into flavor-preserving and flavor-mixing pieces:

$$
\mathcal{H}_{mix} \equiv \mathcal{H}_0 + \tilde{\mathcal{H}} = \begin{pmatrix} M_D - i\frac{\Gamma_D}{2} & 0 \\ 0 & M_D - i\frac{\Gamma_D}{2} \end{pmatrix} + \begin{pmatrix} 0 & M_{12} - i\frac{\Gamma_{12}}{2} \\ M^*_{12} - i\frac{\Gamma^*_{12}}{2} & 0 \end{pmatrix} \tag{2.15}
$$

If mixing is a very small effect, as is the case in the $D^0 - \overline{D}^0$ system, we can make the approximation

$$
e^{-i\mathcal{H}_{mix}\tau} \approx e^{-i\mathcal{H}_0\tau}(\mathbf{1} - i\tilde{\mathcal{H}}\tau) \tag{2.16}
$$

From (2.15) and (2.16), we see that the mixing amplitudes are

$$
\begin{aligned}
\left\langle \overline{D}^0(\tau) \,\middle|\, D^0(0) \right\rangle &= e^{-iM_D\tau - \frac{\Gamma_D}{2}\tau}\left(-iM^*_{12} - \frac{\Gamma^*_{12}}{2}\right)\tau \\
\left\langle D^0(\tau) \,\middle|\, \overline{D}^0(0) \right\rangle &= e^{-iM_D\tau - \frac{\Gamma_D}{2}\tau}\left(-iM_{12} - \frac{\Gamma_{12}}{2}\right)\tau
\end{aligned}
\tag{2.17}
$$

Using (2.11) we can write the mixing amplitudes as

$$
\begin{aligned}
\left\langle \overline{D}^0(\tau) \,\middle|\, D^0(0) \right\rangle &= e^{-iM_D\tau - \frac{\Gamma_D}{2}\tau}\left(-ix^* - y^*\right)\frac{\tau\Gamma_D}{2} \\
\left\langle D^0(\tau) \,\middle|\, \overline{D}^0(0) \right\rangle &= e^{-iM_D\tau - \frac{\Gamma_D}{2}\tau}\left(-ix - y\right)\frac{\tau\Gamma_D}{2}
\end{aligned}
\tag{2.18}
$$

In the absence of CP violation (*ie.*, $\left\langle \overline{D}^0(\tau) \,\middle|\, D^0(0) \right\rangle = \left\langle D^0(\tau) \,\middle|\, \overline{D}^0(0) \right\rangle$), it can be shown that M_{12} and Γ_{12} must be real. Our notation for x and y then becomes consistent with that of [11], and from (2.9) we can express x and y in terms of the mass and width splittings of the mixed states:

$$
\begin{aligned}
x &= \frac{\Delta M}{\Gamma_D} \\
y &= \frac{\Delta\Gamma}{2\Gamma_D}.
\end{aligned}
\tag{2.19}
$$

This also allows us to write a single amplitude for mixing:

$$\left\langle \overline{D}^0(\tau) \mid D^0(0) \right\rangle = e^{-iM_D\tau - \frac{\Gamma_D}{2}\tau} \left(-ix - y\right) \frac{\tau\Gamma_D}{2} \tag{2.20}$$

At this point we will note that the width Γ arises from interactions with real states, and so the width splitting $\Delta\Gamma$ (and therefore y) comes from oscillations through real intermediate states. The mass splitting (and x), on the other hand, comes from virtual intermediate states.

As an aside, even in the general case where CP violation is allowed, we are free to choose a gauge where y is real, and the complex phase is completely contained within x. If we express M_{12} as $|M_{12}|e^{im}$ and Γ_{12} as $|\Gamma_{12}|e^{i\gamma}$, our choice of gauge need only preserve the quantity $\cos(m - \gamma)$ for the eigenvalues of $\mathbf{M}$, $\mathbf{\Gamma}$, and $\mathcal{H}_{mix}$ (*ie.*, the physical observables) to be preserved.

We can simplify (2.20) by a change of variable $\tau = t \cdot \tau_D$, where τ_D is the mean lifetime of the D^0 , so t is dimensionless. τ_D is furthermore just $1/\Gamma_D$. This then gives us

$$\left\langle \overline{D}^0(t) \mid D^0(0) \right\rangle = -\frac{1}{2}(ix + y) \cdot t \cdot e^{-\frac{t}{2}} e^{-i\frac{M_D}{\Gamma_D}t} \tag{2.21}$$

The final exponential is a phase rotation that all D^0 decays will share, and so it can be neglected.

Another simplification that we can make is in the expressions for the masses and widths of the weak eigenstates (2.14). Because M_{12} and Γ_{12} are real, this simplifies to

$$
\begin{aligned}
M^W_{1,2} &= M_{11} \pm \Re\sqrt{\left(M_{12} - i\frac{\Gamma_{12}}{2}\right)^2} \\
&= M_{11} \pm M_{12} \\
&= M_{1,2} \\
\Gamma^W_{1,2} &= \Gamma_{11} \mp 2\Im\sqrt{\left(M_{12} - i\frac{\Gamma_{12}}{2}\right)^2} \\
&= \Gamma_{11} \pm \Gamma_{12} \\
&= \Gamma_{1,2},
\end{aligned}
\tag{2.22}
$$

so the weak and mass eigenstates are identical.

2.3 $D^0 - \overline{D}^0$ Mixing in the Standard Model

2.3.1 Cabibbo-Favored and Doubly Cabibbo-Suppressed Decay Amplitudes

Because the c quark in a D^0 has a lighter partner s in the same generation, it has a Cabibbo-favored decay (CF: all weak couplings on-diagonal in the CKM matrix) to $K^-\pi^+\pi^-\pi^+$ (Figure 2.2). We will call the amplitude for this decay A. Similarly, there is a doubly Cabibbo-suppressed decay (DCS: both weak interactions involve off-diagonal CKM matrix elements) to $K^+\pi^-\pi^+\pi^-$ (Figure 2.3). We will call the amplitude for this decay B. We will, for the moment, only consider the non-resonant decays, and return to the matter of intermediate resonances in Section 2.3.2.

The final state we are interested in is one arising from mixing, and so we are much more likely to see a Cabibbo-favored state after mixing than a doubly Cabibbo-suppressed state. However, when a D^0 mixes to a $\overline{D}^0$ and then decays through the Cabibbo-favored mode, the final state is $K^+\pi^-\pi^+\pi^-$, just as it is for the DCS decay. It is important, therefore, that we can disentangle these two processes. The

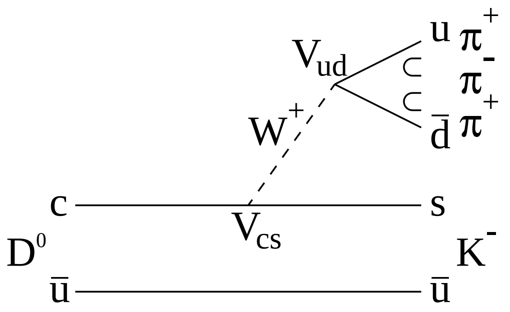


Figure 2.2: Example of a tree-level diagram for the Cabibbo-favored (right-sign) decay $D^0 \to K^-\pi^+\pi^-\pi^+$.

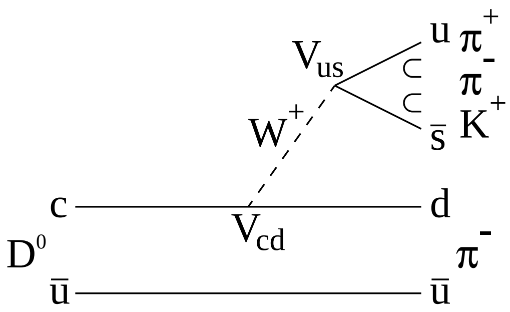


Figure 2.3: Example of a tree-level diagram for the doubly Cabibbo-suppressed (wrong-sign) decay $D^0 \to K^+\pi^-\pi^+\pi^-$.

amplitudes for both decays are

$$\begin{aligned} \mathcal{A}(D^0 \to K^+\pi^-\pi^+\pi^- \,, \mathrm{DCS}) &= Be^{-\frac{t}{2}} \\ \mathcal{A}(D^0 \to K^+\pi^-\pi^+\pi^- \,, \mathrm{mix}) &= -\frac{1}{2}\overline{A}(ix+y)\cdot te^{-\frac{t}{2}} \end{aligned} \tag{2.23}$$

These have different time dependences, and when added and squared give us

$$\begin{aligned} \Gamma(D^0 \to K^+\pi^-\pi^+\pi^-) \propto \Big[&|B|^2 - \left\{\Re(\overline{A}B^*)(y - \Im(x)) - \Im(\overline{A}B^*)\Re(x)\right\} \cdot t \\ &+ \frac{1}{4}|\overline{A}|^2(|x|^2 + |y|^2 - 2y\Im(x)) \cdot t^2 \Big] e^{-t} \end{aligned} \tag{2.24}$$

from which we can divide out a factor of $|\overline{A}|^2$ to get

$$\begin{aligned}\Gamma(D^0 \to K^+\pi^-\pi^+\pi^-) \propto |\overline{A}|^2 \Bigg[R_{\mathrm{DCS}} - t\frac{\Re(\overline{A}B^*)(y - \Im(x)) - \Im(\overline{A}B^*)\Re(x)}{|\overline{A}|^2} \\ + \frac{1}{4}(|x|^2 + |y|^2 - 2y\Im(x)) \cdot t^2 \Bigg] e^{-t}\end{aligned} \tag{2.25}$$

Here we have defined $R_{\mathrm{DCS}} \equiv |B|^2/|\overline{A}|^2$ as the relative rate of DCS $D^0 \to K^+\pi^-\pi^+\pi^-$ to the right-sign $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$. To simplify this, we first write the amplitudes as

$$\begin{aligned}\overline{A} &= |\overline{A}| \cdot e^{ia} \\ B &= -|B| \cdot e^{ib}\end{aligned} \tag{2.26}$$

where the minus sign in B accounts for the sign of the real part of the ratio of CKM matrix elements $V_{cd}V_{us}^*/V_{cs}V_{ud}^*$. The combination $\overline{A}B^*$ is then

$$\overline{A}B^* = -|\overline{A}| \cdot |B| \cdot e^{i(a-b)} \tag{2.27}$$

We can then define the phase difference between the DCS and CF amplitudes $\delta \equiv a-b$, and write the rate as

$$\begin{aligned}\Gamma(D^0 \to K^+\pi^-\pi^+\pi^-) \propto |\overline{A}|^2 \Bigg[R_{\mathrm{DCS}} - t\frac{|B|}{|\overline{A}|} \left\{(y - \Im(x))\cos\delta - \Re(x)\sin\delta\right\} \\ + \frac{1}{4}(|x|^2 + |y|^2 - 2y\Im(x)) \cdot t^2 \Bigg] e^{-t}\end{aligned} \tag{2.28}$$

If we now make the assumption of CP conservation, so that x and y are real, the second term of (2.28) can be further simplified by rotating x and y:

$$\begin{aligned}x' &= x\cos\delta + y\sin\delta \\ y' &= y\cos\delta - x\sin\delta\end{aligned} \tag{2.29}$$

This gives us

$$\Gamma(D^0 \to K^+\pi^-\pi^+\pi^-) \propto |\overline{A}|^2 \left[R_{\mathrm{DCS}} + \sqrt{R_{\mathrm{DCS}}} \cdot y' \cdot t + \frac{1}{4}(x'^2 + y'^2) \cdot t^2 \right] e^{-t} \tag{2.30}$$

which is a simple parameterization to fit.

2.3.2 Resonant Sub-Structure

In both the Cabibbo-favored (and mixed) and doubly Cabibbo-suppressed decays, there are a number of strongly decaying intermediate resonances which can appear. The most prominent of these are the ρ^0, the $\bar{K}^*(892)^0$, the $a_1(1260)^+$, and the $K_1(1270)^-$. Each resonant sub-mode, as well as the non-resonant contribution, has a different amplitude and phase-space structure. These can interfere with one another, producing a total amplitude for the decay that reflects not only the relative strengths of the resonances, but also different reconstruction efficiencies due to where they populate phase space. In addition, since both the mixed and doubly Cabibbo-suppressed decays have these intermediate resonances, and furthermore can interfere, it follows that each individual sub-mode from one decay will interfere with some or all of the sub-modes from the other decay.

It is worth noting, however, that regardless of the particulars of the interference, we can still express the total amplitude of the right-sign and wrong-sign decays as $\overline{A}$ and B. The phase δ, unfortunately, is not easily interpreted. A full fit to the phase space, plus proper lifetime, can separate the various modes and interferences. This requires high statistics in both the right-sign and wrong-sign decays, however. The current best measurements of the right-sign sub-structure come from Mark III[12] (with 1281 events) and E691[13] (with 1745 events). While CLEO II.V observes many more right-sign events, the number of wrong-sign events is far fewer. Given that only the inclusive resonant decay $D^0 \to K^-\pi^+\rho^0$ has been measured to better than 10%, and no specific resonance has been measured to better than 15%, we would not expect to be able to distinguish any significant features in the wrong-sign phase space. Moreover, the wrong-sign backgrounds are large compared with the backgrounds from either Mark III or E691, further complicating a fit.

2.3.3 Real and Virtual Intermediate States

Mixing proceeds through intermediate states accessible to both the particles, as described in Section 2.1.4. In the case of $K^0 - \overline{K}^0$ mixing, this takes place primarily through two- and three-pion states. Both of these are low enough in mass that they can be produced by a neutral kaon, and so are *on-shell* reactions involving real intermediate states. The differing volumes of phase space for the allowed decays of the mixed eigenstates result in different total widths for the K_1^0 and K_2^0, commonly referred to as *short* and *long*. In contrast, the $B^0 - \overline{B}^0$ system mixes through intermediate states with higher mass than is available, and so are *off-shell* reactions involving virtual intermediate states. In this case the splitting involves self-energy, producing a mass difference.

In the $D^0 - \overline{D}^0$ system, both on- and off-shell reactions are greatly suppressed in comparison to the previous two systems. Due to the unitarity of the CKM matrix, if all quark masses were the same there would be complete cancellation of the amplitude, and hence no mixing. This was first described in a paper by Glashow, Iliopoulos, and Maiani[14], and is referred to as the *GIM mechanism* or *GIM cancellation*. In the B system, the presence of the heavy top quark breaks the symmetry of the GIM mechanism, allowing fairly substantial mixing to occur. In the D system, however, all of the quark masses are light enough in comparison to the mass of the propagator ($m_W \approx 80\text{GeV}/c^2$) that the GIM cancellation is almost complete. Long-range interactions between final states can yield predictions for both $|x|$ and $|y|$ on the order of 0.1%[4]. The current best limit on y comes from the FOCUS analysis of $D^0 \to K^+K^-$ [7], which reports $y = (3.42 \pm 1.39 \pm 0.74)\%$. CLEO II.V reports limits on y' of $-5.8\% < y' < 1.0\%$ [9]. The central values of these measurements have opposite sign, and if these are correct then the strong phase difference between $D^0 \to K^-\pi^+$ and $D^0 \to K^+\pi^-$ is very large [4, 16].

2.4 $D^0 - \overline{D}^0$ Mixing Beyond the Standard Model

Perhaps the most compelling reason to search for $D^0 - \overline{D}^0$ mixing is the potential window on new Physics beyond the Standard Model. Previous searches for mixing have pushed the limits on y to the point where a discovery would be consistent with Standard Model predictions. Mixing in x, on the other hand, remains a fertile ground for new Physics. This requires the ability to disentangle mixing due to real and virtual intermediate states, which is considerably more complicated than disentangling mixing from doubly Cabibbo-suppressed decays.

Various proposed models provide enhancements to x, primarily through the introduction of higher-mass particles which enter the box diagrams, breaking the GIM cancellation. A review of theoretical predictions from various extensions to the Standard Model (as well as a number of Standard Model calculations) has been compiled by Harry Nelson[15], and it demonstrates how effective setting new mixing limits can be in ruling out new theories.

Chapter 3

The CLEO II.V Detector

The CLEO II.V detector (Figure 3.1) is located in Ithaca, New York at the Cornell Electron Storage Ring. It surrounds the center-of-mass interaction point of the storage ring, providing angular coverage equal to roughly 90% of 4π.

The detector is an upgrade to the CLEO II detector[17], and is described in detail elsewhere[18, 19]. Surrounding the beampipe is a three-layer silicon microstrip subdetector for precision position information. Outside of this are two wire chambers for momentum measurement and particle identification. The outer wire chamber is surrounded by a time-of-flight system for additional particle identification, around which are crystal calorimeters for charged- and neutral-particle electromagnetic energy measurements. This is enclosed in an iron solenoid, which generates a 1.5 tesla magnetic field oriented along the beam direction, and which also absorbs the bulk of the remaining particles except for muons. A large number of the muons can reach the outermost set of concentric wire chambers, each separated by additional iron.

The most natural way to define a coordinate system is with respect to the detector and the colliding beams. The $\hat{z}$ direction is defined to be along the axis through the center of the detector in the direction of positron flow. $\hat{y}$ is defined to be up. This leads to $\hat{x}$ horizontal and outwards from the center of the storage ring.

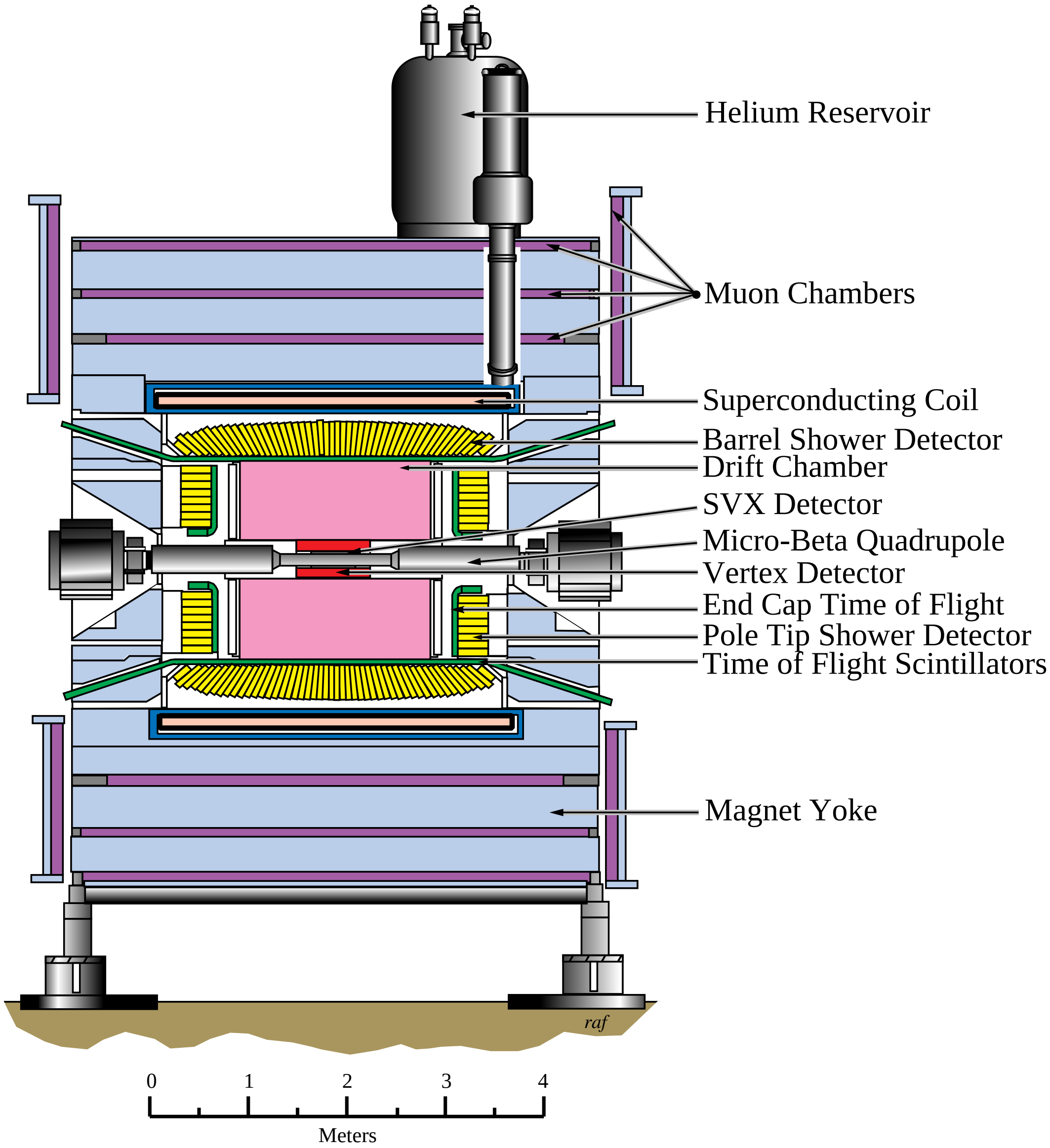


Figure 3.1: The CLEO II.V Detector

3.1 Tracking Chambers

At the heart of the detector are the tracking chambers, which record a particle's trajectory as it passes through. In each sub-detector this trajectory registers as a sequence of discrete points, which can then be put together to re-create the full trajectory.

Because the detector is in a solenoidal field, charged particles will follow helical paths, where the curvature κ is inversely proportional to the momentum transverse to the field, which we call p_T. This allows us to measure the particle's momentum and charge, since positive and negative particles will bend in opposite directions.

Tracking resolutions in CLEO II.V are excellent, as can be seen in the impact parameters from $e^+e^- \to \mu^+\mu^-$ events (Figures 3.2 and 3.3).

3.1.1 Silicon Microstrip Detector

The silicon microstrip detector, called the Silicon Vertex Detector (SVD), consists of three concentric sets of silicon wafers. Each wafer has strips on either side, measuring along orthogonal axes. One side measures positions along the $\hat{z}$ axis, which (as previously mentioned) is the direction of the positron beam. The other side measures in the $r-\phi$ plane, along the wafers as shown in Figure 3.4.

Sensitive strips are produced via either p $(r-\phi)$ or n (z) implanting of the silicon. On the $r-\phi$ sides of the wafers, sensitive strips are closely spaced. Every fourth strip is instrumented, with the charge deposition on the other three determined by capacitive coupling. In z, every strip is instrumented, but due to geometric constraints each strip is "ganged" in the readout with strips on other wafers in its layer. Determination of which strip in the gang was actually hit is based on tracking. Implant and readout pitches are summarized in Table 3.1.

Since every strip in one direction crosses a number of strips in the orthogonal direction, it is necessary to pair up hits from both sides. We can do this by looking for charge-deposition correlations, since the particles tend to ionize small regions of

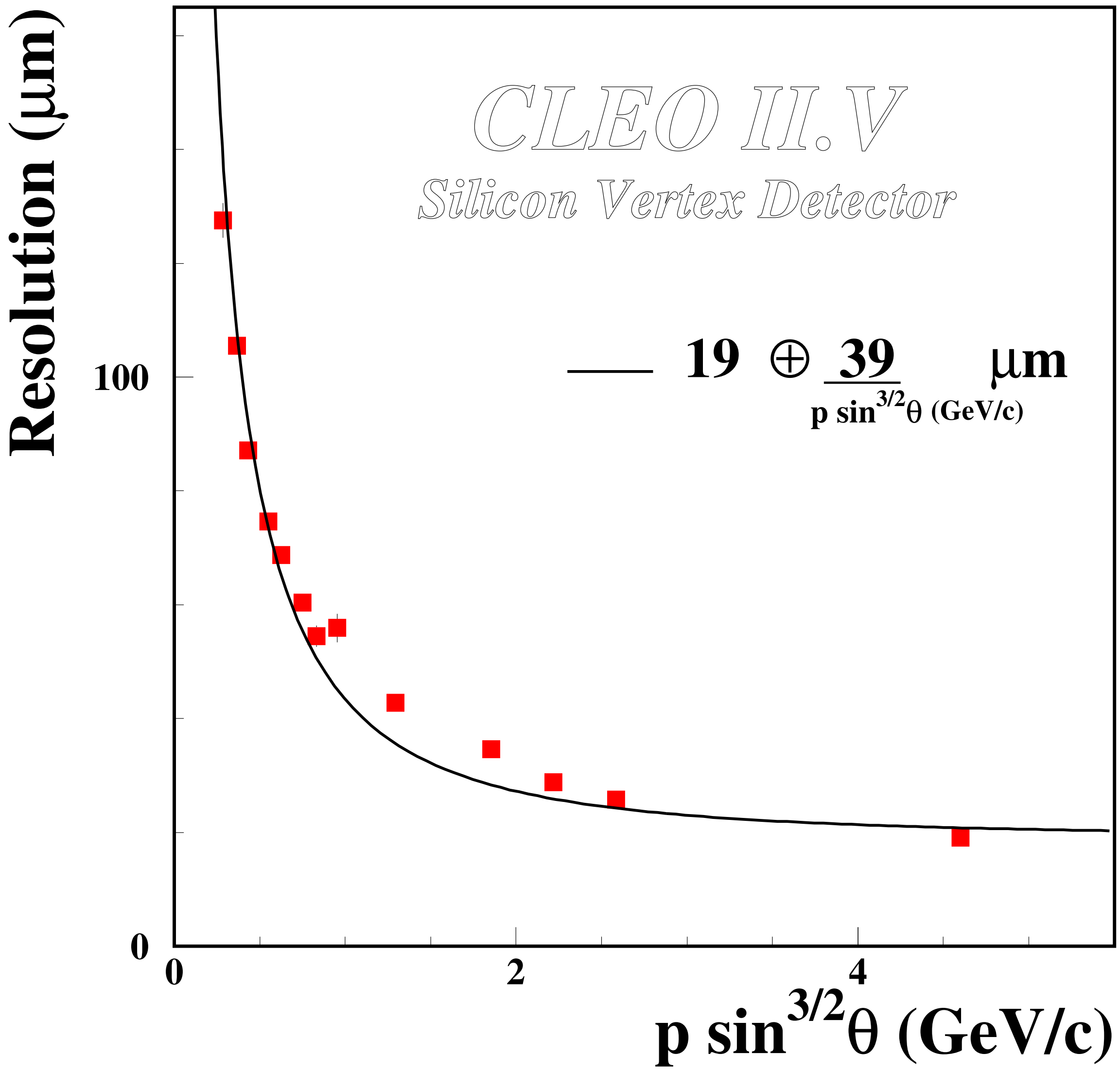


Figure 3.2: Impact parameter resolution in the $r-\phi$ plane for the CLEO II.V detector as a function of particle momentum and incidence angle.

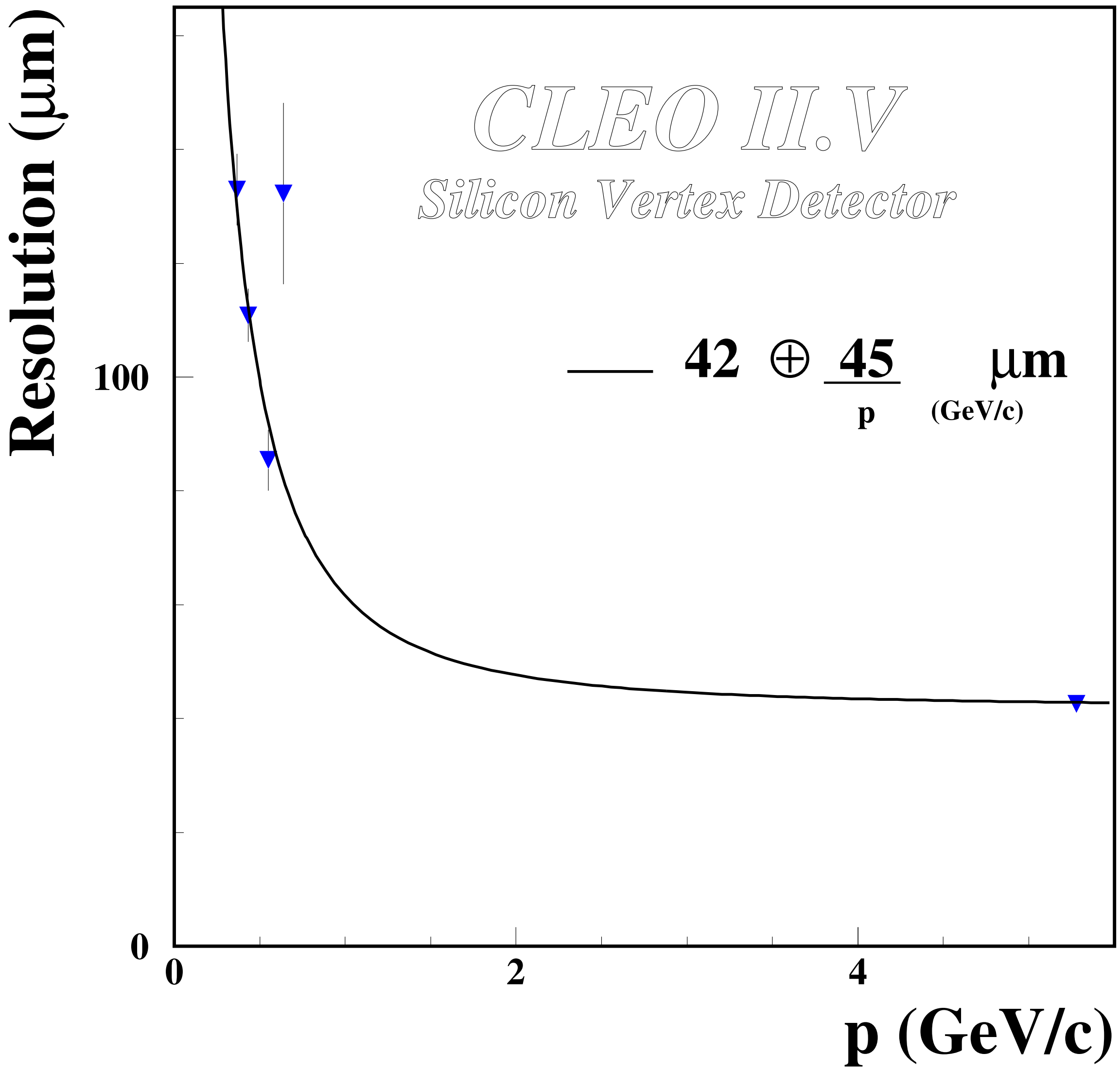


Figure 3.3: Impact parameter resolution in z for the CLEO II.V detector as a function of particle momentum.

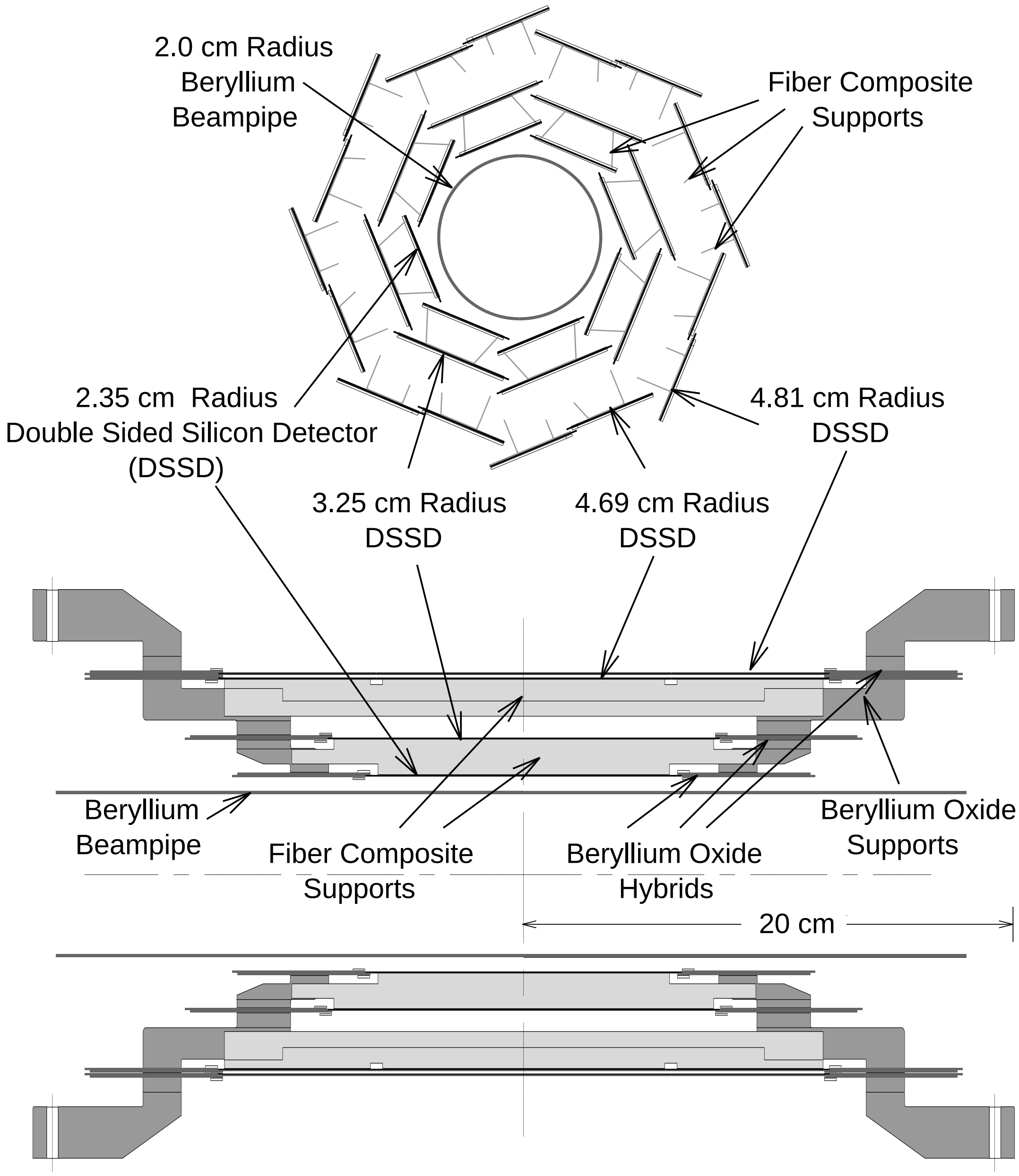


Figure 3.4: The CLEO II.V Silicon Vertex Detector

Table 3.1: SVD Wafer Specifics

	Layer 1	Layer 2	Layer 3
Wafer Radius (mm)	23.5	32.5	46.9, 48.1
$r-\phi$ Implant Pitch (μm)	28.0	28.5	28.0
$r-\phi$ Readout Pitch (μm)	112	114	112
z Implant Pitch (μm)	105	101	105

the silicon, with the electrons drifting to one side and the holes to the other.

Due to the small number of points provided by the SVD for a given particle, it does not do as well as the wire chambers in measuring momentum. However, it gives us our best measurements of position (Figure 3.5).

With one hit on each orientation of the silicon, we have enough information to greatly improve the position measurement for a particle. In order to measure the direction accurately, we need a second set of high-precision points. This is referred to as having "two hits per view," and is a common requirement for the reconstructed tracks used in silicon-dependent analyses. The efficiency for finding two hits per view on electron tracks coming from the initial-state radiation process $\gamma\gamma \to e^+e^-$ is shown in Figure 3.6 as a function of transverse momentum, with an average efficiency of 91%. For tracks from the decay $D^{*+} \to D^0\pi^+, D^0 \to K^-\pi^+$, the average efficiency is 89% (Figure 3.7).

3.1.2 Wire Chambers

There are two wire chambers in the CLEO II.V detector. The inner chamber is known as the Vertex Detector (VD), and the outer is the Central Drift Chamber (DR). Both are filled with Helium-Propane gas (60% He, 40% Propane), which improves ionization-based particle identification over the CLEO II Argon-Ethane mixture. This will be discussed in more detail in section 3.2.1. The wires collect electrons released as a passing charged particle ionizes the gas, recording both the amount of charge and the time at which it is collected. Using properties of the gas and the electric

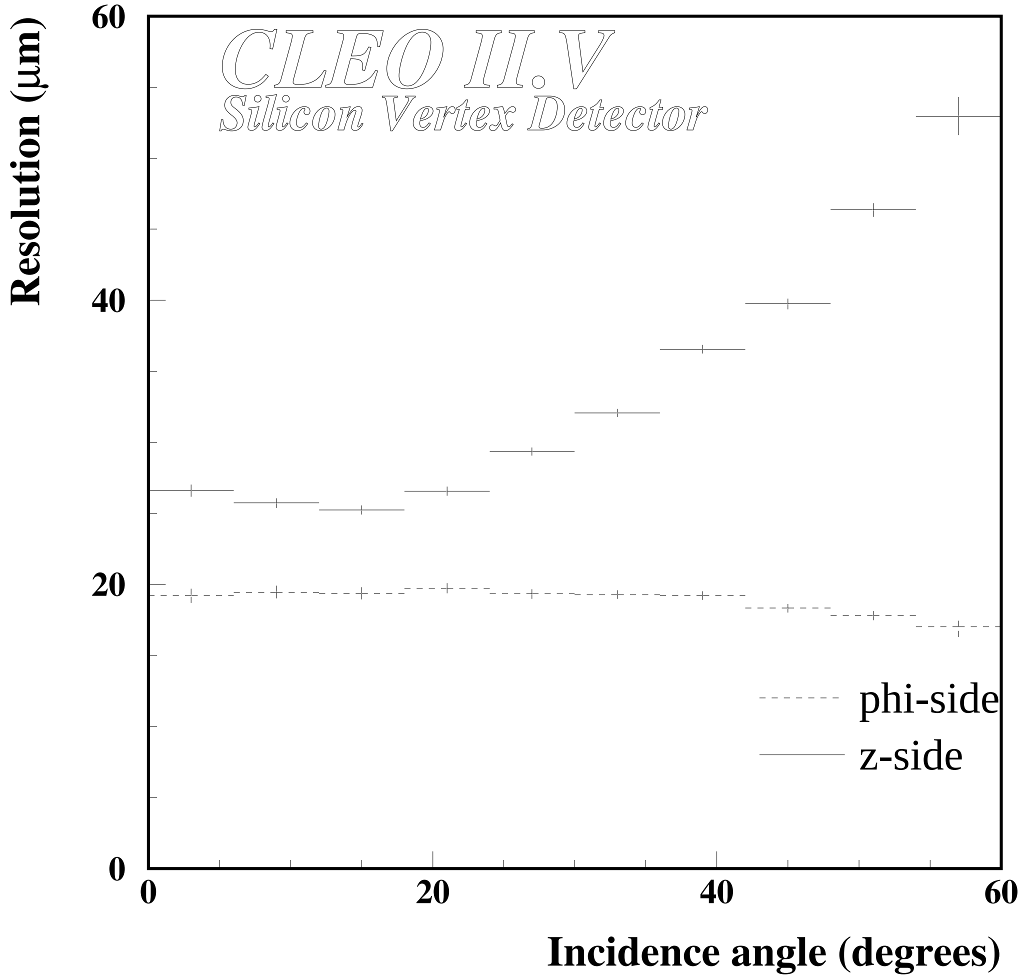


Figure 3.5: Position resolution of the CLEO II.V Silicon Vertex Detector as a function of particle incidence angle.

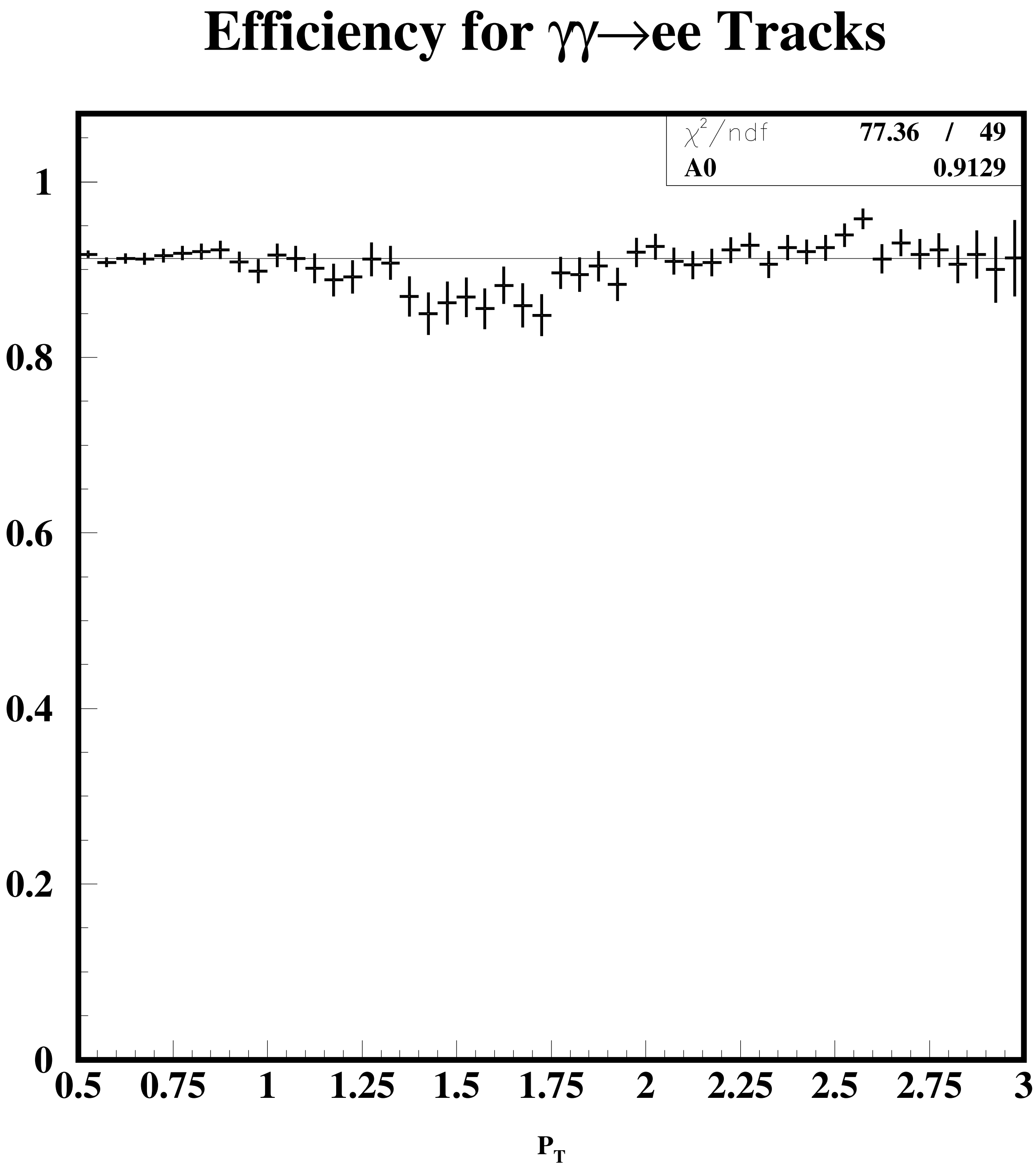


Figure 3.6: Efficiency of finding two silicon hits per view for tracks coming from the process $\gamma\gamma \to e^+e^-$, as a function of transverse momentum (in units of GeV/c).

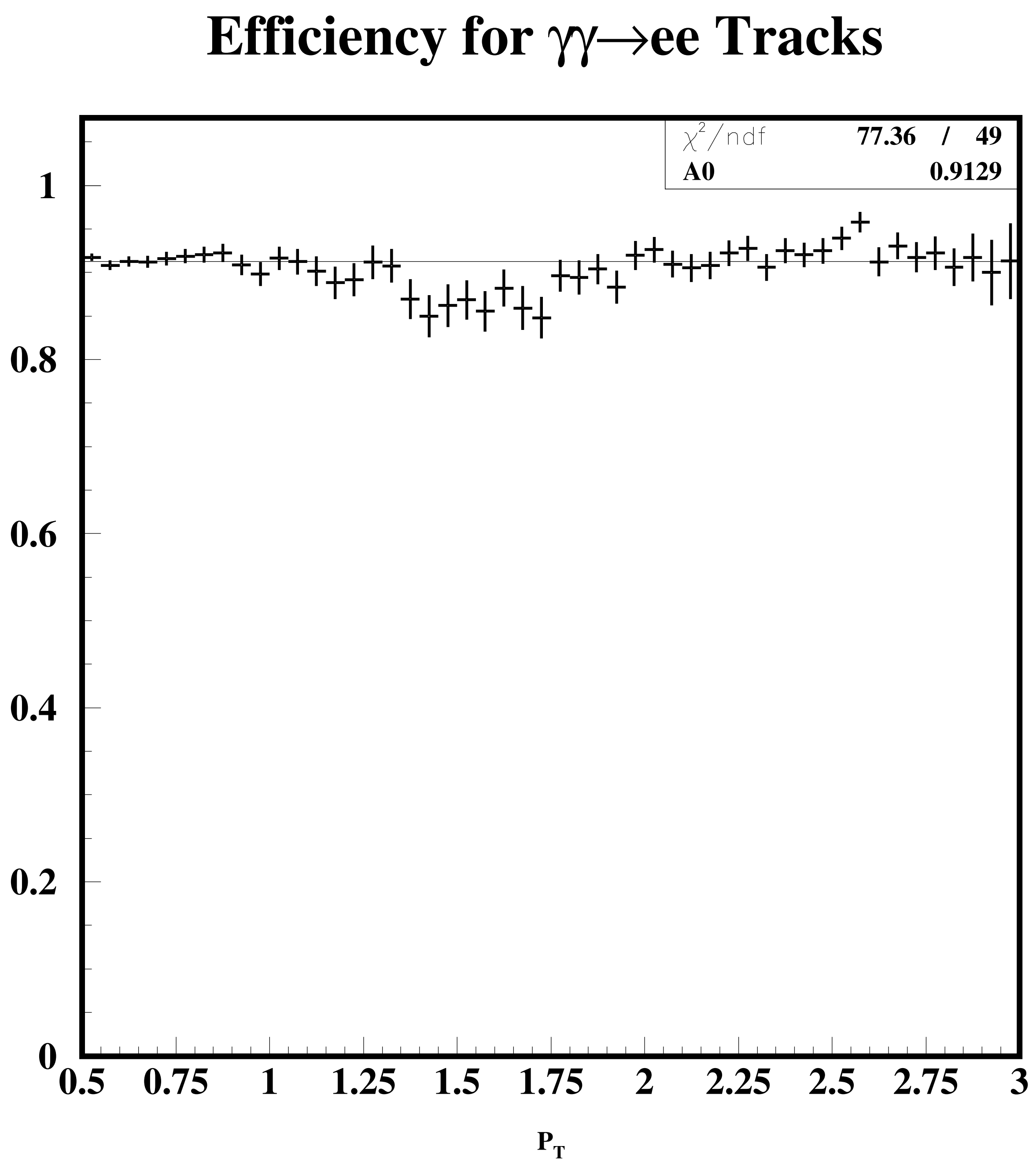


Figure 3.7: Efficiency of finding two silicon hits per view for tracks coming from the decay $D^{*+} \to D^0\pi^+, D^0 \to K^-\pi^+$, as a function of transverse momentum (in units of GeV/c).

field, we can calculate a drift distance from the time. The electric field is produced by another set of "field" wires which are oriented parallel to the sense wires.

The Vertex Detector is a cylindrical tube containing 10 layers (groups of equally-spaced wires at a constant radius) of sense wires oriented axially (along $\hat{z}$). The first five layers contain 64 sense wires each, and the last five contain 96 sense wires each.

The Central Drift Chamber contains another 40 layers of axial wires, as well as 11 layers with stereo angles ($\Delta\phi/\Delta z$) ranging from $3.77°\mathrm{m}^{-1}$ to $6.89°\mathrm{m}^{-1}$ interspersed at every fourth layer. These stereo angles in turn alternate between positive and negative. As a particle travels in the $\hat{z}$ direction, its distance from a particular stereo wire, or even which wire in a layer it will pass nearest, changes. Thus the stereo layers, when a sufficient number have been hit, allow us to determine p_z/p_T for a particle. The innermost layer of the drift chamber contains 96 sense wires, and the outermost 384.

Position is determined by the drift distance. The drift distance resolution is 101μm in the axial layers, and 131μm in the stereo layers. For 5.3GeV muons, this gives us a 47MeV resolution on the momentum. Momentum resolution tends to be degraded somewhat for slower particles, since there are fewer hits available to make a measurement. Increased "dip" angle (towards the beam direction) also degrades resolution due to distortions in the electrostatic fields.

Both wire chambers have layers of cathode panels at their inner and outer radii. These cathodes measure image charges from the nearest sense wire layers, and provide measurements along the $\hat{z}$ direction. The individual pads (sensitive strips) on the cathodes range in width from 5.85mm on the innermost to about 1cm on the outermost. The pad widths are chosen to spread the image of the collected charge from one wire across three pads.

3.2 Particle Identification

For many analyses, it is essential to identify decay products as particular types of particles in order to reduce backgrounds. Because a large number of pions are produced in a typical hadronic event, it is helpful to be able to separate pions from kaons. There are two sources of particle-identity information, the drift chambers and the time-of-flight counters.

3.2.1 Drift Chamber Energy Loss, $\frac{dE}{dx}$

As a charged particle passes through the drift chamber, it loses energy due to the collisions which produce the ionization seen at the sense wires. The ionization is proportional to the particle's energy, so we can, by recording the amount of charge collected by a particle's passing, obtain a series of snapshots of the particle's decreasing energy. These can then be compared to theoretical expectations, based principally on the mass of the particle, to determine a measure of consistency between a particle and a specific mass hypothesis. This is usually expressed in terms of the width of a distribution obtained from a clean sample of each type of particle. The expected ionizations for kaons and pions, as well as the 1σ deviations, are shown in Figure 3.8. There is a region of overlap between the two curves, which unfortunately corresponds rather closely to the momentum spectrum of the kaons produced in $D^0 \to K3\pi$.

3.2.2 Time-of-Flight Detector

Another way to determine a particle's mass is by comparing its momentum and velocity, since $p = mc\beta\gamma$. The momentum is determined by the tracking chambers. The velocity is determined by comparing the time of the primary interaction with the time a particle hits a scintillator mounted between the wire chambers and the crystal calorimeter. This time, along with the distance between the interaction point and the scintillator (which is momentum-dependent, due to the helical motion), gives us the particle's velocity.

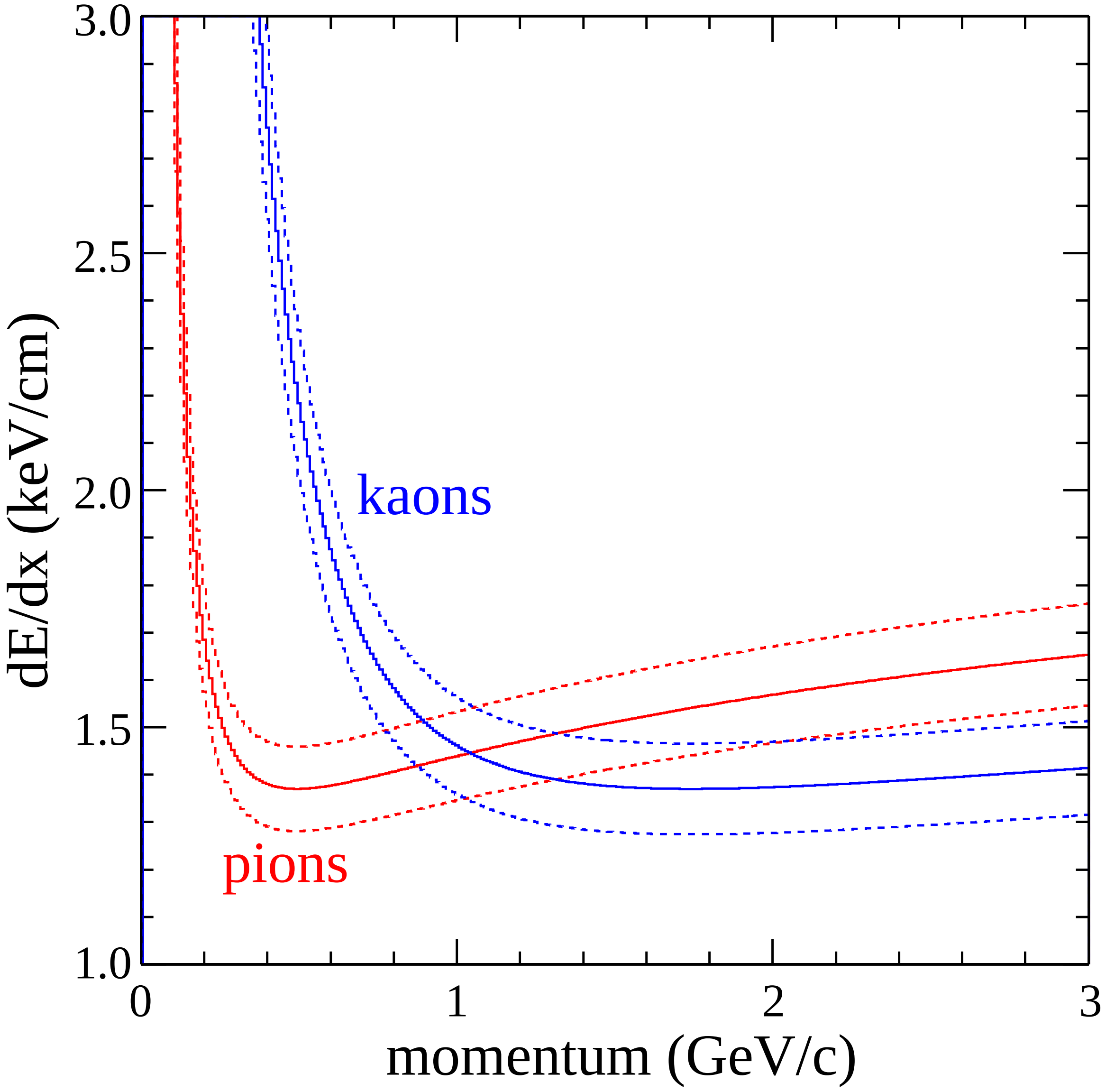


Figure 3.8: Specific ionization curves for kaons and pions as a function of momentum. 1σ bands are shown.

There are two sections of the time-of-flight system. The cylindrical barrel region surrounds the tracking chambers, and has a time resolution of $0.2-0.3$ns. The endcap region has a time resolution of $0.4-0.7$ns. When time-of-flight information is added to the analysis procedure, the improvement is negligible, so time-of-flight is not included in the final analysis.

3.3 Crystal Calorimeter

Outside of the tracking chambers are 7800 thallium-doped cesium iodide crystal scintillators. These react to both charged and neutral particles (excluding neutrinos) with an electromagnetic shower. Electrons and photons deposit all of their energy in the calorimeter, while hadrons deposit less of their energy but spread among more crystals. Its primary purpose is to detect photons and help distinguish electrons from other charged particles.

At normal incidence, a particle traverses about sixteen radiation lengths in passing through a crystal scintillator. The energy of the resultant shower is spread in the transverse direction with a width given approximately by the Molière radius, which is 3.8 cm. The crystals are 5 cm square in this plane, so an electromagnetic shower will generally be confined to nine or fewer crystals, depending on how close to the center or edge of a crystal the impact is.

The energy resolution is less than 4% of the measured energy for much of the range detected (Figure 3.9), and the angular resolutions are also quite good, on the order of a few to a few tens of milliradians.

The crystal calorimeter is not used in this analysis, as all of the final-state particles are charged, and hence sufficiently well measured by the tracking chambers.

3.4 Muon Chambers

The outermost detectors are the muon chambers. These are separated from the other detector elements and each other by thick layers of iron, which act as both a flux return for the superconducting magnetic coil and a barrier for particles produced in the collision. Because muons interact very weakly with ordinary matter, many of them penetrate through several of the layers of iron.

The muon chambers are wire chambers, similar to the tracking wire chambers. There is, however, no magnetic field present, and the wire are oriented both along $\hat{z}$ and perpendicular to it. Because of this, the chambers cannot be cylindrical, so

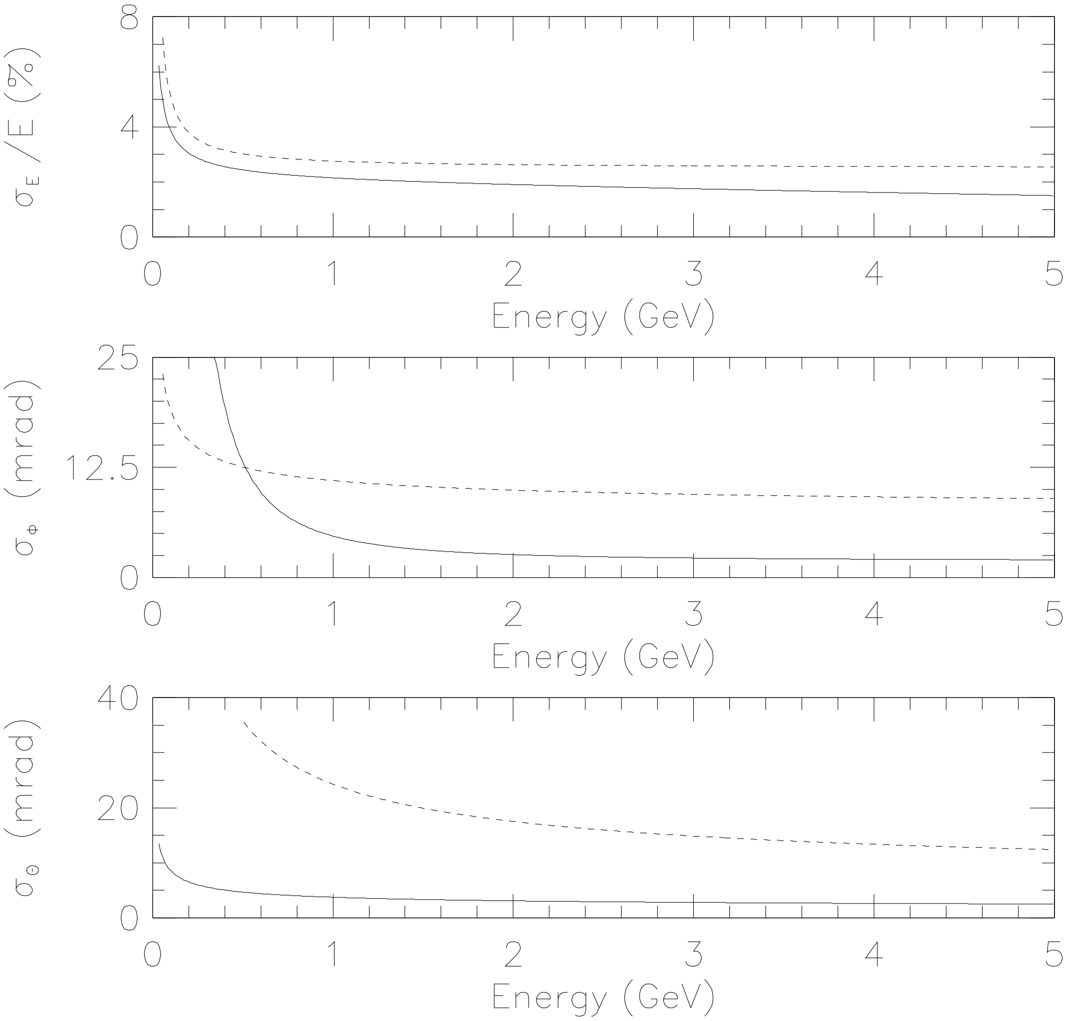


Figure 3.9: Energy and angular resolutions for the cesium iodide electromagnetic calorimeter. The solid lines show the resolutions for the barrel crystals, and the dashed lines show the resolutions for the endcap crystals. The barrel θ resolution is proportional to $\sin\theta$ and is shown for the worst case ($\theta = 90°$).

instead each layer of the muon chambers is split into eight separate chambers, forming a large octagon. The resulting charge deposits are then matched up to an existing drift chamber track to identify that track as a muon.

This analysis does not use information from the muon detectors.

Chapter 4

Search for the Decay $D^0 \to K^+\pi^-\pi^+\pi^-$

In the decay of a D^0 meson to a K and three charged π's, $D^0 \to K^-\pi^+\pi^-\pi^+$ is the Cabibbo-favored (CF) decay, and is illustrated in Figure 4.1. The 'wrong-sign' decay $D^0 \to K^+\pi^-\pi^+\pi^-$ can proceed through either mixing or a doubly Cabibbo-suppressed decay (DCSD), as shown in Figure 4.2. While it has never been observed, the relative wrong-sign rates for the decays $D^0 \to K^+\pi^-$ and $D^0 \to K^+\pi^-\pi^0$, the results of a previous search[21], and the expected sensitivity of the CLEO II.V detector motivate a search for $D^0 \to K^+\pi^-\pi^+\pi^-$.

Mixing of D^0 and $\bar{D}^0$ states can proceed through either real (Figure 4.3) or virtual (Figure 4.4) intermediate states. It is in the latter, which in the Standard Model are

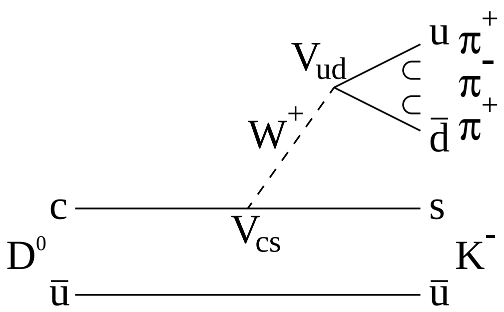
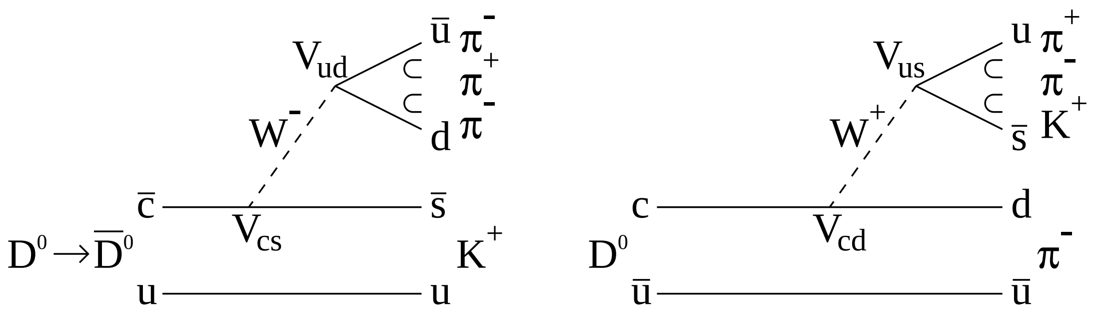


Figure 4.1: Tree-level diagram for the decay $D^0 \to K^-\pi^+\pi^-\pi^+$. Production of additional quark-antiquark pairs after the weak decay vertices corresponds to the intermediate resonance expected to dominate.



(a) Mixed

(b) DCSD

Figure 4.2: Tree-level diagrams for the decay $D^0 \to K^+\pi^-\pi^+\pi^-$ for (a) mixing and (b) DCSD. Production of additional quark-antiquark pairs after the weak decay vertices corresponds to the intermediate resonance expected to dominate.

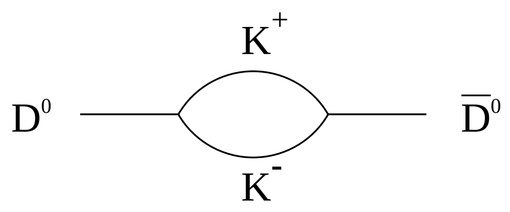


Figure 4.3: D^0 mixing through real intermediate states.

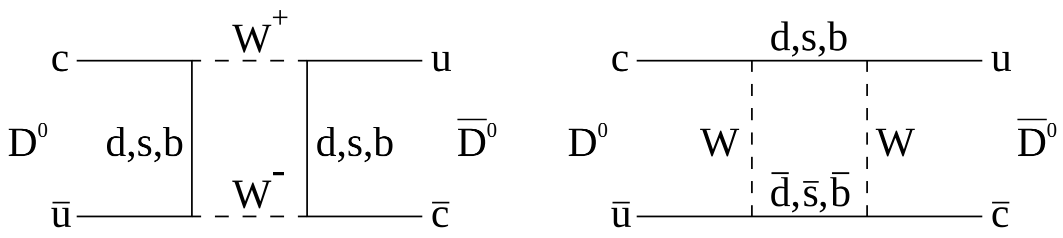


Figure 4.4: D^0 mixing through virtual intermediate states.

suppressed by the GIM mechanism[14], that we would expect to see enhancements due to new Physics. The decay $D^0 \to K^+\pi^-\pi^+\pi^-$ has a complicated resonant substructure which, when combined with the strong phases between the CF and DCS amplitudes for each sub-mode, makes disentangling mixing and DCS effects nearly impossible without a much larger (and cleaner) dataset. Nonetheless, we can still measure a total wrong-sign rate,

$$R_{WS} \equiv \frac{\mathcal{B}(D^0 \to K^+\pi^-\pi^+\pi^-)}{\mathcal{B}(D^0 \to K^-\pi^+\pi^-\pi^+)}, \tag{4.1}$$

which should be comparable to the rates from $D^0 \to K^+\pi^-$ and $D^0 \to K^+\pi^-\pi^0$, since they are all expected to be about $\tan^4\theta_C \approx 3 \times 10^{-3}$.

4.1 Data Sample

For this analysis, we used the full CLEO II.V dataset, with a total integrated luminosity of 9.0fb^{-1}. Of this, 6.2fb^{-1}was collected on the $\Upsilon(4S)$ resonance, and the remainder was collected at a center-of-mass energy slightly below this resonance. We use both on- and off-resonance data, since both contain the $e^+e^- \to c\bar{c}$ events which interest us.

4.2 Monte Carlo Simulation Sample

In order to model the backgrounds, we produced a large sample of Monte Carlo (MC) simulated data. This was pre-selected to contain events with a high likelihood of being selected by the loose analysis requirements, which greatly speeds up the generation time. This sample was generated over the full range of running conditions present in taking the data, and so accounts for variations in detector performance due to calibration, aging of components, and environmental factors.

Not specifically included in the pre-selection are backgrounds from τ decays and interactions between particles and the detector. These are expected to be both small and flat in all variables of interest. From a dead-reckoning comparison of right-sign signal statistics in the data and simulation, the total effective luminosity simulated corresponds to six times the data, or $54\mathrm{fb}^{-1}$.

An additional sample of generic Monte Carlo equal to roughly five times the data statistics is also included. This is a fully generic sample, with no pre-selection to increase the fraction which will be accepted by this analysis.

None of the generic Monte Carlo sample contains the wrong-sign signal decay. The right-sign, however, is present in approximately the correct ratio of signal to background.

4.3 Event Reconstruction

We fully reconstruct the decay chain $D^{*+} \to D^0\pi_s^+, D^0 \to K3\pi$ using the tracks left by the final-state particles as they passed through the silicon and drift chambers. The D^0 candidate is reconstructed first, and then paired with a suitable pion to form a D^{*+} candidate.

For each track, we have a Lorentz momentum calculated assuming the particle to have either the pion or kaon mass. As this is output from a Kalman [22] fit, we also have a reliable covariance matrix which gives us both the uncertainties on the components (both momentum and position) and the correlations between them. Using this, we can impose a common-vertex constraint on the four tracks in a D^0 candidate to get the best possible measurements of the combined position and momentum. The mass distribution for the right-sign candidates in the Monte Carlo is shown in Figure 4.5.

Given a $K3\pi$ combination, we can project it back to the interaction point (IP), where the initial electron and positron collided. Rather than attempt to reconstruct an event-by-event IP, we take advantage of the typical dimensions of the run-averaged luminous region. Because this region is very narrow in the vertical direction (σ_y is about $10\mu m$[23]), we can simply project the $K3\pi$ candidate to the average y of the luminous region, and use that as the IP. This also gives us the vertical displacement of the D^0 from the IP, from which we can calculate a proper lifetime for the candidate

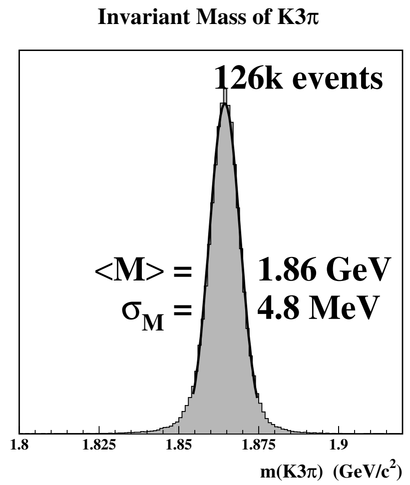
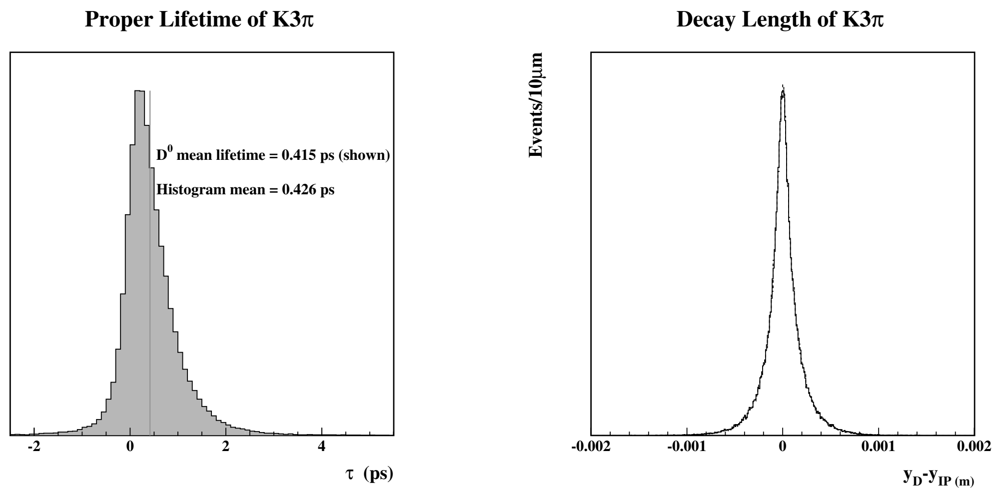


Figure 4.5: Invariant mass of $K3\pi$ combinations in right-sign Monte Carlo.

(both are shown in Figure 4.6). For comparison, Figure 4.6(b) also shows the vertical displacement using the actual IP from the event as a dashed line. The two distributions are indistinguishable, which indicates that the vertical position of the center of the run-averaged luminous region is an excellent approximation to the event-by-event vertical position of the IP.

To reconstruct the D^{*+} candidate, we look for an additional pion and once again subject it to a vertex constraint. However, while the $K3\pi$ vertex was allowed to vary (since we were interested in determining the D^0 decay point), the pion is forced through the projected interaction point. The momentum of the pion from the D^{*+} is constrained kinematically to be less than about $450\mathrm{MeV}/c$, with the mean value being significantly lower. In this momentum region, the dominant effect contributing to mis-measurement is multiple Coulomb scattering from the beam pipe and silicon wafers. By re-fitting the pion through the IP, we can remove the bulk of this effect.

Now that we have both a $K3\pi$ and π_s Lorentz momentum, we can add them together to get the Lorentz momentum of the D^{*+} candidate. From this we calculate an invariant mass for the $K3\pi\pi_s$ combination. This invariant mass correlates strongly



(a) Proper Lifetime

(b) $y_D - y_{IP}$

Figure 4.6: $K3\pi$ combinations in right-sign Monte Carlo: (a) proper lifetime, and (b) $y_D - y_{IP}$. Overlaid on (b) as a dashed line is the difference between y_D and the true interaction point (IP). That the two distributions are indistinguishable is a confirmation of our reconstruction technique.

with the $K3\pi$ mass, so by subtracting the two, we are left with a much narrower distribution. If we also subtract the mass of the pion, what remains is the energy released by the decay of the D^{*+} candidate, Q ($= m(K3\pi\pi_s)\cdot c^2 - m(K3\pi)\cdot c^2 - m_\pi\cdot c^2$), which peaks at 5.827MeV[20]. The width of this distribution is dominated by the momentum of the slow pion (Figure 4.7).

4.4 Selection Criteria

4.4.1 Track Selection

The tracks used to form a D^0 candidate are subjected to a number of quality requirements. The pattern-recognition algorithm allows tracks with no $\hat{z}$ information and tracks put together in a final "dredge" pass of as-yet unused drift chamber hits. These are needed for some analyses which have to account for all of the energy ob-

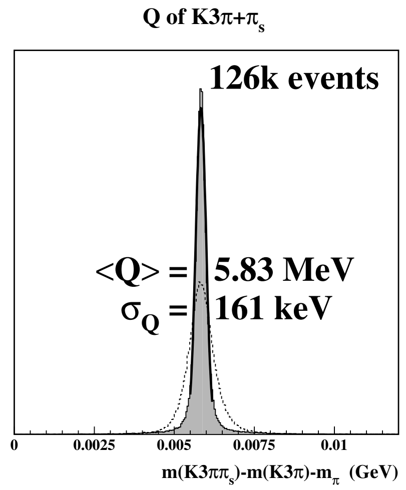


Figure 4.7: Energy release Q of $K3\pi\pi_s$ combinations in RS Monte Carlo. The dashed line shows the Q distribution without the slow-pion refit.

served by the detector. For this analysis these tracks must be rejected. Because a D^0 will not typically travel more than $100 - 200\mu$m before decaying, we also require that all tracks be consistent with the interaction point. This amounts to requiring that they originate inside of the beam pipe. We impose an additional requirement of good silicon information, meaning at least two hits on each orientation of the silicon wafers, for at least two of the tracks, in order to improve the spatial vertex resolution for the candidate. The entire combination must also have no net charge.

For the π coming from the decay of the D^{*+}, we apply the same criteria as for the other four tracks, but good silicon information is always required.

4.4.2 D^0 Candidate Selection

Identifying the kaon among the four particles in a D^0 candidate is crucial, since this will determine whether the candidate is right-sign or wrong-sign. Because the rate for the right-sign decay is so much higher than the rate for the wrong-sign decay, feed-through from the right-sign is one of the most significant sources of background in the wrong-sign measurement. We have two basic handles on this: kinematics and particle identification.

Kinematic Separation

We calculate the invariant mass of a particle from the Lorentz momenta of its decay products. When two of these final-state particles have approximately equal total momentum, their masses can be exchanged without affecting the invariant mass very much. We refer to this exchange as a "mass-flip." In the case of the CLEO II.V $D^0 \to K^+\pi^-$ analysis[9], there was sufficient momentum asymmetry between the kaon and pion that requiring a large separation between flipped masses was a very efficient method for eliminating feed-through. We would expect a similar requirement in this analysis to be less efficient, since we have more particles in the final state, which reduces the amount of energy available for the particles' momenta. This requirement is crucial, however, since we cannot allow right-sign feed-through into the wrong-sign distribution. We create four D^0 candidates by subjecting the tracks to a common-vertex constraint (changing which track is assigned the kaon mass). For each of these candidates, we can compute a mass and measurement uncertainty. From these and the world-average D^0 mass (1.865 GeV/c^2)[20], we can calculate a normalized mass:

$$\tilde{m}_i = \frac{m_i - m_{D^0}}{\sigma_i} \quad (i = 1...4) \tag{4.2}$$

A candidate i passes the kinematic mass-flip separation requirement if the alternate candidates satisfy $|\tilde{m}_i| \geq 3.85$. The uncertainties in the reconstruction are well-measured, so the normalized mass distribution for real signal has a Gaussian shape, and a selection criterion can be reliably interpreted in terms of σ_m. Figure 4.8 shows this distribution for right-sign data, which has a small amount of background that is fairly flat. Since our main concern is correctly tagging the strangeness of the final state, this is only applied to alternates with oppositely charged kaons to the one being considered. In addition, we only consider candidates with masses between 1.8 and 1.92 GeV/c^2. We can see in Figure 4.9 that this removes a very large amount of background that peaks in Q (D^0 momentum nearly correct, though possibly not

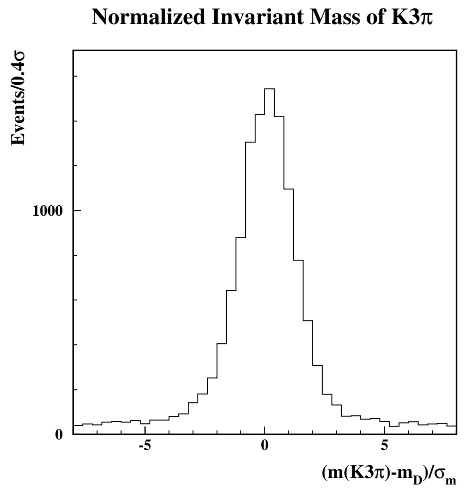


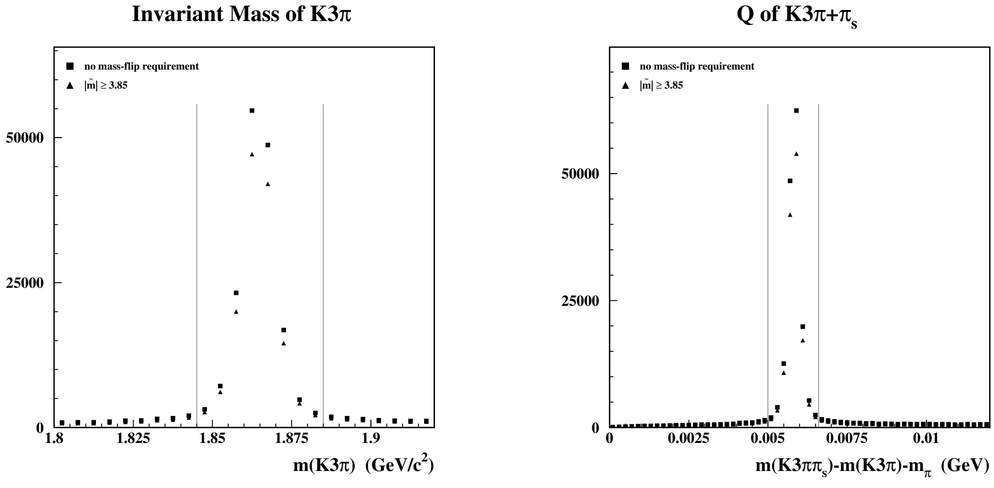
Figure 4.8: Normalized mass of right-sign $K3\pi$ candidates from data.

the mass). This is 87% efficient for signal, while reducing the wrong-sign background level in the signal region by a factor of 6.

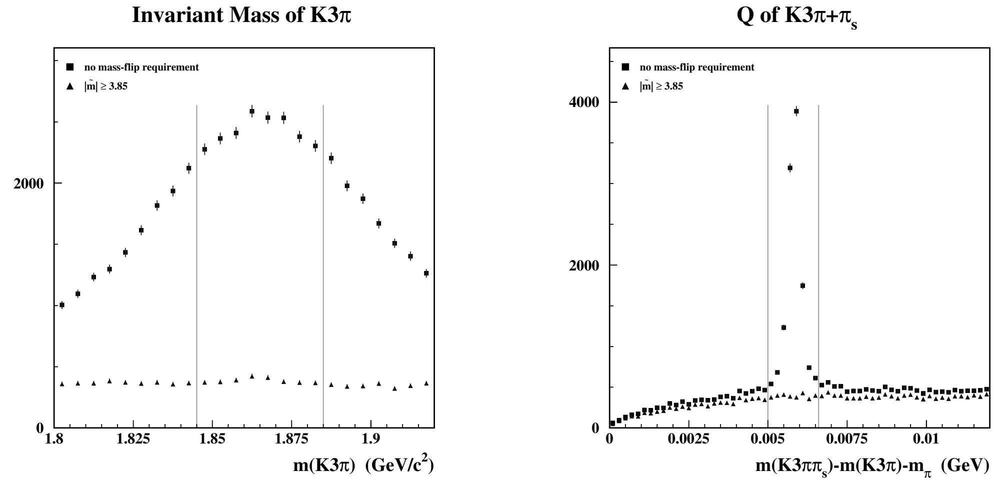
There is an important effect due to treatment of errors. In general, the physical processes which cause particles to scatter have non-Gaussian distributions. We are unable, however, to treat errors in any way other than as Gaussian distributions, and so there are cases where the uncertainty in the reconstructed invariant mass is either under- or over-estimated. Both of these affect the ability to properly reject feed-through. This will typically only happen in cases where a particle has been badly mis-measured, though, so we would not expect these combinations to form a peak in either invariant mass or Q.

Particle Identification

As an additional method of correctly assigning particle masses, we can use the particle identification (PID) information available track-by-track. Specifically, we consider the ionization of the gas in the drift chamber per unit length of the particle's trajectory as it traveled through it (dE/dx). This is dependent on $\beta\gamma$ for the particle, which will differ for particles of the same momentum but different masses. The curves of



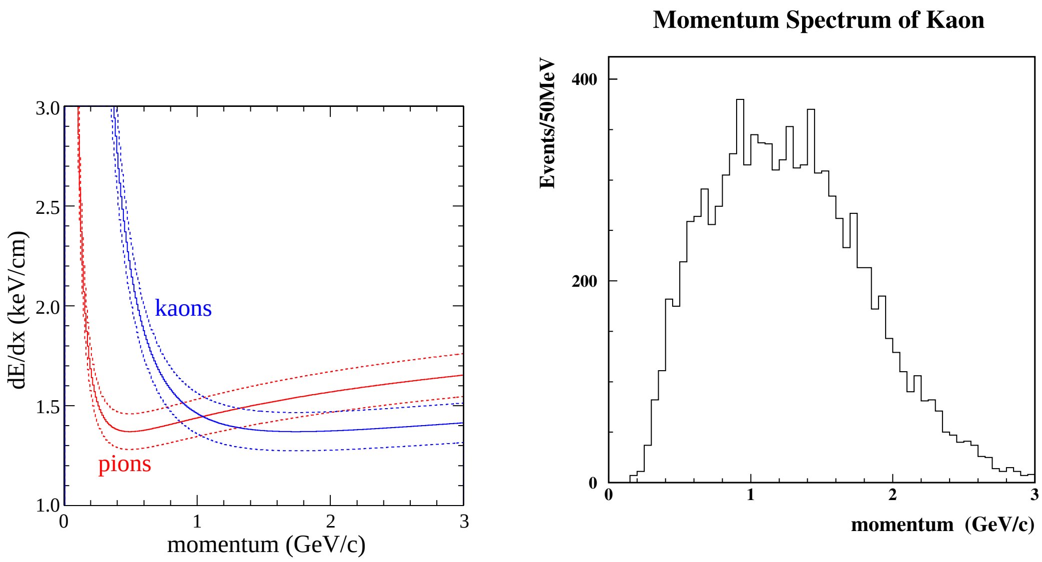
(a) Right-Sign from Generic Monte Carlo



(b) Wrong-Sign from Generic Monte Carlo

Figure 4.9: Right-sign and wrong-sign mass and Q distributions with and without the mass-flip veto. The alternate mass assignments peak broadly in mass, but narrowly in Q. There is a slight residual peak in wrong-sign invariant mass from correctly assigned $\bar{D}^0$ combinations matched with a random slow π^+, which does not peak in Q. All other selection criteria are applied.

most-likely ionization for kaons and pions cross, however, and in a fairly broad range around this crossing the separation between the two mass hypotheses is not very large (Figure 4.10). Nonetheless, it can provide considerable additional ambiguity resolution. The ionization is translated into a number of (Gaussian-distributed) standard deviations from the most-likely value for each mass hypothesis, which we can label σ_K and σ_π.



(a) Ionization Curves

(b) Kaon Momentum Spectrum

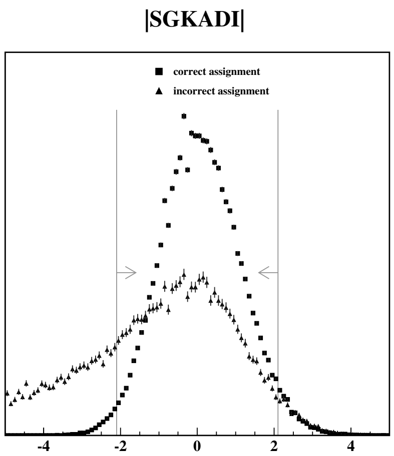
Figure 4.10: Ionization curves for particles of varying masses passing through the drift chamber. One-sigma bands are shown. Shown for comparison is the momentum spectrum of kaons from right-sign signal in data.

We can, in principle, also include information from the time-of-flight scintillators. Given the momentum of a particle and the time it takes to travel from the interaction point to a fixed detector, we can determine its mass. Because of resolution issues both in time and the location of the IP, a more practical approach is to consider expected flight time for a particle with a certain mass and momentum, and compare the measured value. When properly calibrated, this gives us an additional

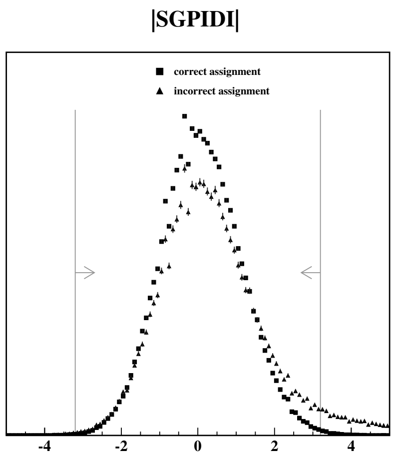
Gaussian-distributed σ for each track and mass hypothesis. This information is available somewhat less frequently than dE/dx particle identification. However, in the case where a particle's mass assignment is unresolved by dE/dx alone, either because the particle has insufficient information or is in the "cross-over" region for dE/dx, it can occasionally provide enough additional information to determine the particle's correct mass assignment. Time-of-flight was not included for initial data processing, but was considered afterwards for possible improvement in K/π discrimination. As with dE/dx, time-of-flight has an overlap region where the expected measurements for kaons and pions are very close. Unfortunately, this region is very similar to that for dE/dx. Simple consistency requirements from time-of-flight, comparable to those from dE/dx, were placed on the data (Monte Carlo processing takes considerably longer). Comparing the acceptance in the signal region of the right-sign data with that in the sideband (non-signal) regions of the wrong-sign data, the additional background rejection was almost entirely offset by the loss of signal. For this reason, time-of-flight is not used in this analysis.

We use particle identification to check consistency with a mass hypothesis. For kaon candidates, we require $|\sigma_{K,dE/dx}| \leq 2.1$, which is a fairly tight requirement because the backgrounds are dominated by pions masquerading as kaons. For the pion candidates, we make the looser requirement of $|\sigma_{\pi,dE/dx}| \leq 3.2$, since kaons appearing to be pions is a much rarer problem, making signal efficiency more important than background rejection in this case. Distributions for both quantities are shown in Figure 4.11, and the mass and Q distributions are in Figure 4.12. These requirements reduce the background by a factor of three with 87% efficiency for the signal.

It is worth noting that we can, in principle, impose more complicated requirements on particle identification information. For example, we could attempt to use the available information to positively identify the species of a particle. The advantage of using a simple consistency requirement rather than a complicated rejection requirement is that the acceptance of the consistency requirement is almost flat as a function of momentum, while any attempt to use rejection will eliminate more parti-



(a) σ_K



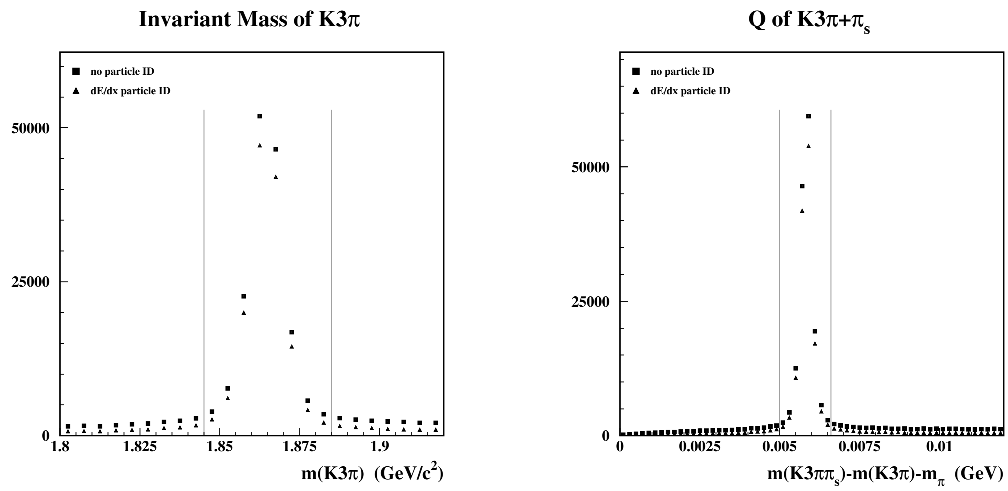
(b) σ_π

Figure 4.11: Distributions of σ_K and σ_π for correct and incorrect assignments. The distributions are normalized to unit area, including underflows and overflows.

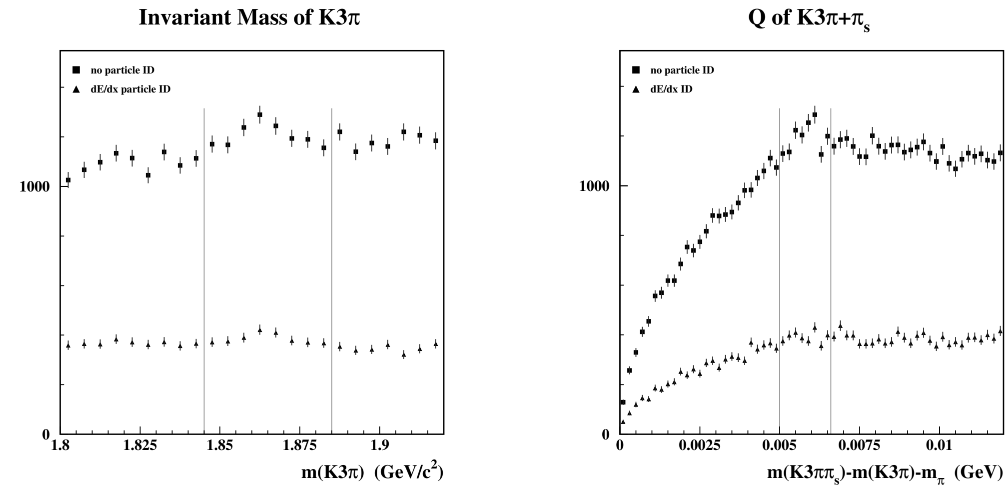
cles in the cross-over region, which is where the bulk of the kaons appear. This also makes our method more robust against mis-modeling of the initial D^{*+} momentum spectrum.

Additional Requirements

Some additional requirements are made, in order to ensure that the candidate is well reconstructed. The measurement uncertainty on the invariant mass is required to be no larger than 10 MeV/c^2 (Figure 4.13), and the probability of the χ^2 for the vertex-constraint fit is required to be at least 10^{-4} (Figure 4.14). The latter, especially, eliminates combinations where the particles originated from spatially distinct points (Figure 4.15). As with the kinematic separation (Section 4.4.2), non-Gaussian errors cause a deviation in the shape of the probability distribution from the expected (flat) shape. Over-estimation of errors causes a slight increase at high probability, while under-estimation (large non-Gaussian scatters) causes a sizeable peak at zero. We lose



(a) Right-Sign from Generic Monte Carlo



(b) Wrong-Sign from Generic Monte Carlo

Figure 4.12: Right-sign and wrong-sign mass and Q distributions with and without dE/dx particle identification. All other selection criteria are applied.

close to one third of our signal candidates because of this, since a large scatter in any one of the four tracks will prevent them from forming a decent vertex. Additionally,

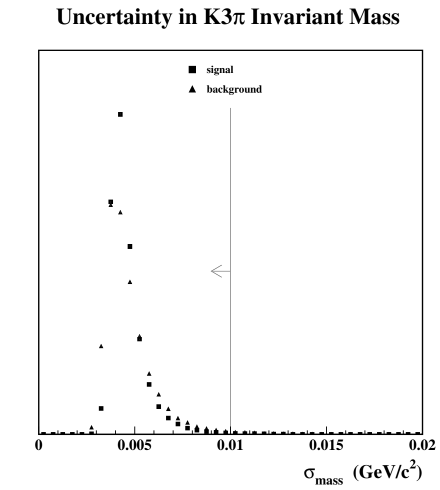


Figure 4.13: Measurement uncertainty on $K3\pi$ invariant mass for signal and background. The background shown comes from continuum Monte Carlo events not containing the decay $D^0 \to K3\pi$. The distributions are normalized to unit area, including overflows.

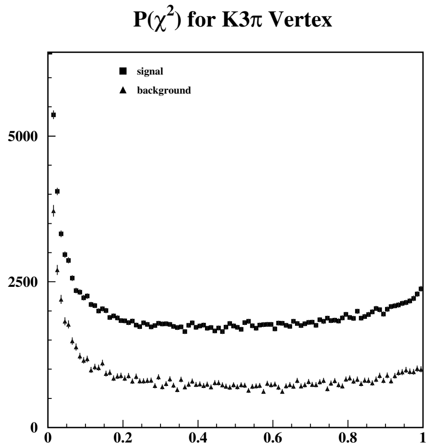
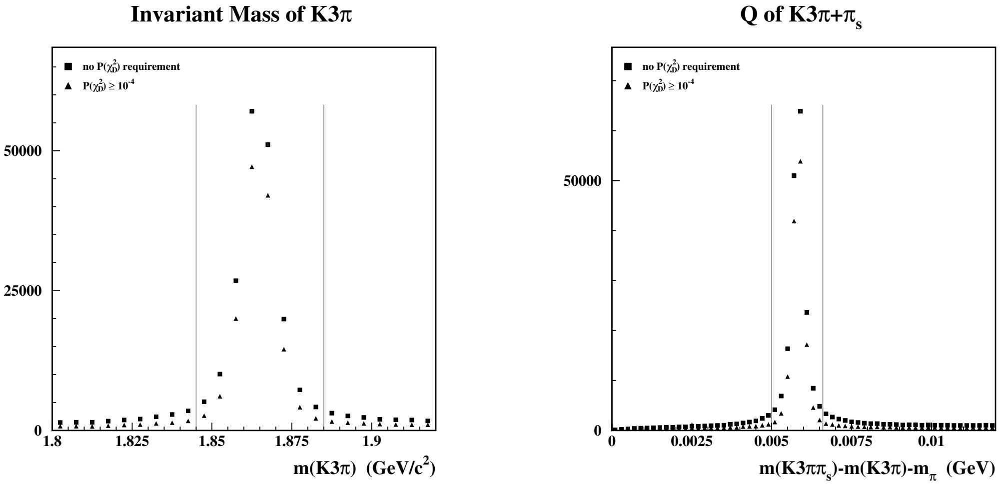
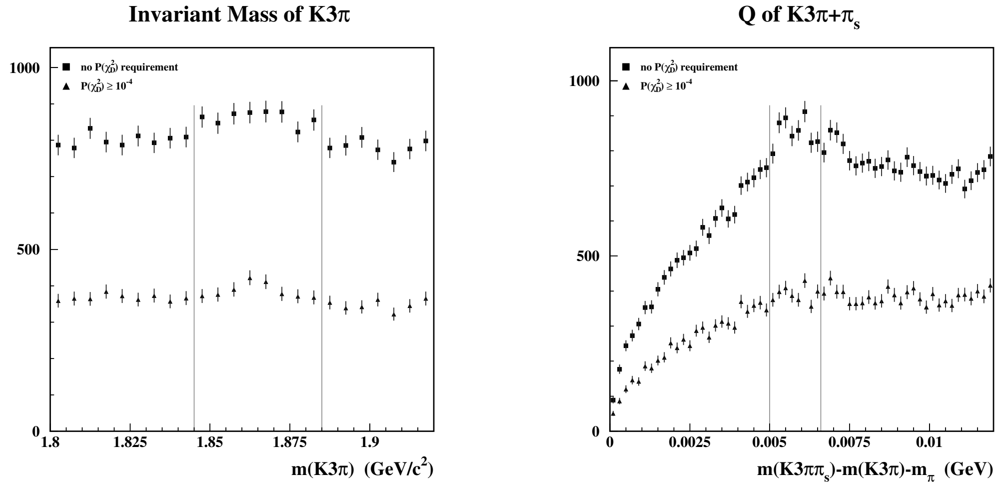


Figure 4.14: $P(\chi^2)$ of $K3\pi$ vertex-constrained fits for signal and background. The background shown comes from continuum Monte Carlo events not containing the decay $D^0 \to K3\pi$. The distributions are normalized to unit area. The bin between 0 and 0.01 is suppressed so that the rest of the shape can be seen. About twice as much of the background populates the first bin as the signal, primarily at zero.



(a) Right-Sign from Generic Monte Carlo



(b) Wrong-Sign from Generic Monte Carlo

Figure 4.15: Right-sign and wrong-sign mass and Q distributions with and without the $P(\chi^2)$ requirement on the $K3\pi$ vertex. All other selection criteria are applied.

the uncertainty in the proper lifetime is required to be less than or equal to 1.1cm (effectively a non-cut, see Figure 4.16).

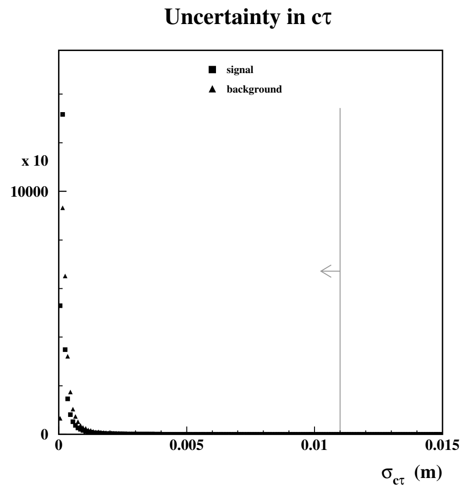


Figure 4.16: Measurement uncertainty on $c\tau$ of the $K3\pi$ candidate for signal and background. The background shown comes from continuum Monte Carlo events not containing the decay $D^0 \to K3\pi$. The distributions are normalized to unit area, including overflows.

The D^0 candidate selection criteria are summarized in Table 4.1. Acceptance is measured with all other analysis requirements in place, and comes from a signal region approximately 2σ to either side of the nominal mass and Q. The right-sign data and Monte Carlo are used to calculate the signal acceptances, and the wrong-sign Monte Carlo is used to calculate the background rejection factor.

Table 4.1: Summary of D^0 selection criteria and acceptances.

Quantity	**Requirement**	**Signal Acceptance (%)**		**Background Rejection Factor**
		RS(Data)	**RS(MC)**	**WS(MC)**
invariant mass	$m \in [1.8, 1.92]\mathrm{GeV}/c^2$	—	—	—
norm. flipped mass	$\lvert\tilde{m}_i\rvert \geq 3.85$	87.2	86.3	6.3
kaon PID	$\lvert\sigma_{K,dE/dx}\rvert \leq 2.1$	87.2	89.8	3.1
pion PID	$\lvert\sigma_{\pi,dE/dx}\rvert \leq 3.2$	*incl. in kaon acceptance*		
invar. mass uncert.	$\sigma_m \leq 10\mathrm{MeV}/c^2$	99.7	99.6	1.0
vert. constraint qual.	$P(\chi^2_{K3\pi}) \geq 10^{-4}$	73.6	79.9	2.2
lifetime uncert.	$\sigma_{c\tau} \leq 1.1\mathrm{cm}$	99.9	99.9	1.0

4.4.3 D^{*+} Candidate Selection

We require Q to be less than or equal to 12MeV, and its error no larger than 500keV(Figure 4.17). We further require that the D^{*+} candidate's momentum be

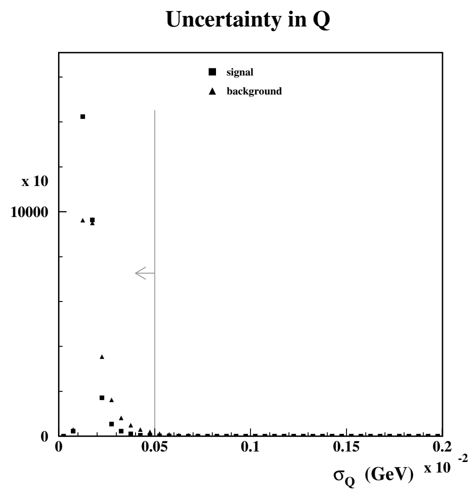


Figure 4.17: Measurement uncertainty on Q for signal and background. the background shown comes from continuum Monte Carlo events not containing the decay $D^0 \to K3\pi$. The distributions are normalized to unit area, including overflows.

at least 2.68GeV/c (Figure 4.18). This predominantly eliminates combinatoric backgrounds, and also eliminates D^{*+} mesons produced in B meson decays. The latter are not especially problematic, but their backgrounds are different, which would require additional modeling to properly handle. The momentum requirement also increases the spatial separation between the D^0 production and decay points. Figure 4.19 shows the right-sign and wrong-sign mass and Q distributions with and without the D^{*+} momentum requirement, and demonstrates the effectiveness of this requirement. The background is reduced by a factor of two, while the signal efficiency is 85%. The probability of χ^2 for the slow pion refit is required to be at least 10^{-4} (Figures 4.20 and 4.21), as with the $K3\pi$ vertex. Unlike the displaced vertex from the D^0, most of the slow pions come from the primary interaction, so the vertex requirement does not eliminate combinatorics. It has a dramatic effect on the width of the Q distribution,

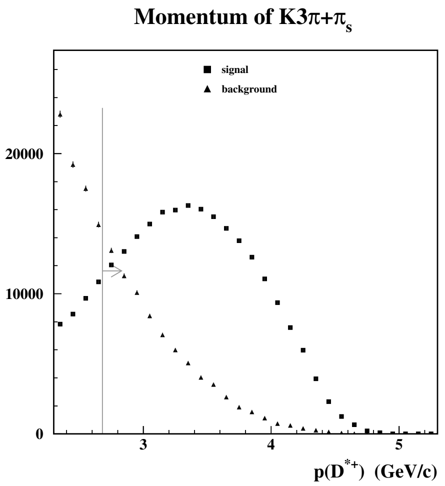


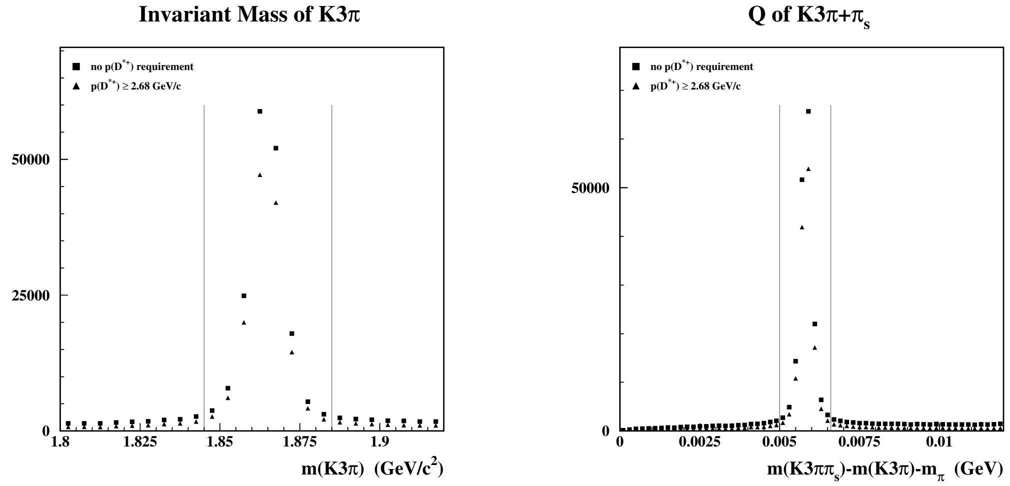
Figure 4.18: Momentum of the $K3\pi\pi_s$ candidate for signal and background. The background shown comes from continuum Monte Carlo events not containing the decay $D^0 \to K3\pi$. Normalization is affected by a hard cut-off of 2.3GeV$/c$ (the lower limit of the plot) used in generating the high-statistics Monte Carlo sample.

however, by reducing the effects of multiple scattering, as mentioned in Section 4.3.

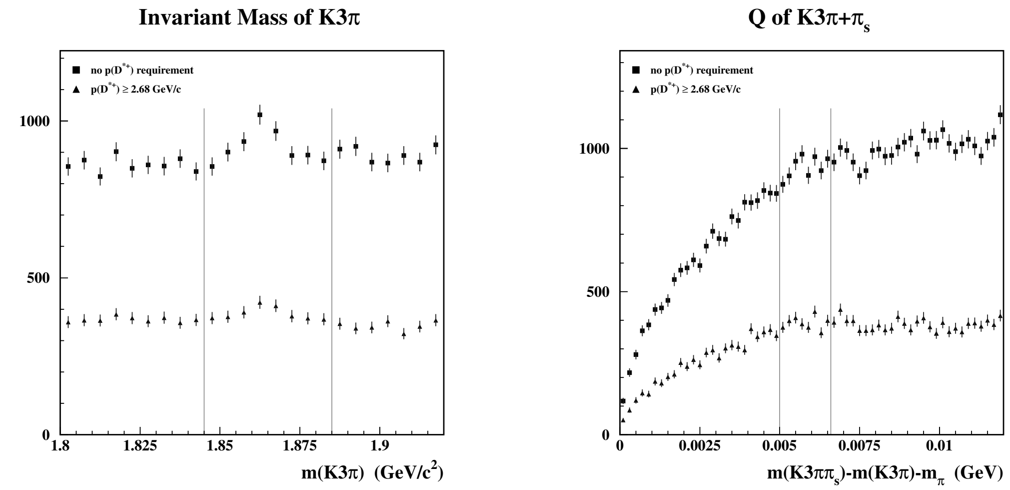
The D^{*+} candidate selection criteria are summarized in Table 4.2. Acceptance is measured with all other analysis requirements in place, and comes from a signal region approximately 2σ to either side of the nominal mass and Q. The right-sign data and Monte Carlo are used to calculate the signal acceptances, and the wrong-sign Monte Carlo is used to calculate the background rejection factor. The acceptance of the momentum requirement is greatly affected by the lower limit used in the initial reconstruction and Monte Carlo generation. For this reason, a common minimum momentum of 2.3GeV$/c$ is imposed for the calculation.

4.4.4 $\bar{K}^0K^+\pi^-$ Suppression

If in the decay $D^0 \to \bar{K}^0K^+\pi^-$ (Figure 4.22), which is singly Cabibbo-suppressed, the $\bar{K}^0$ decays as a K^0_S to two charged pions, we end up with a final state identical to our signal, but with an additional weak decay. Many of the $\bar{K}^0$ mesons will decay



(a) Right-Sign from Generic Monte Carlo



(b) Wrong-Sign from Generic Monte Carlo

Figure 4.19: Right-sign and wrong-sign mass and Q distributions with and without the $K3\pi\pi_s$ momentum requirement. All other selection criteria are applied.

far enough from the D^0 decay point that the final-state particles will fail our analysis procedure, but those that remain will be very similar to our signal kinematically.

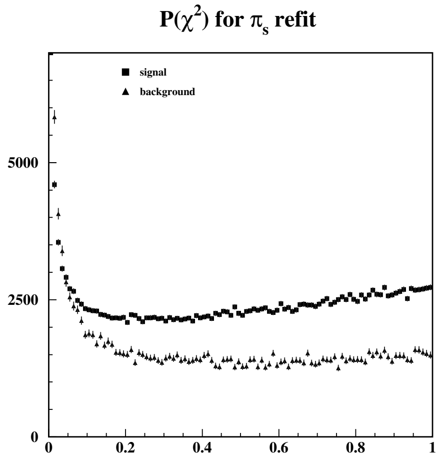
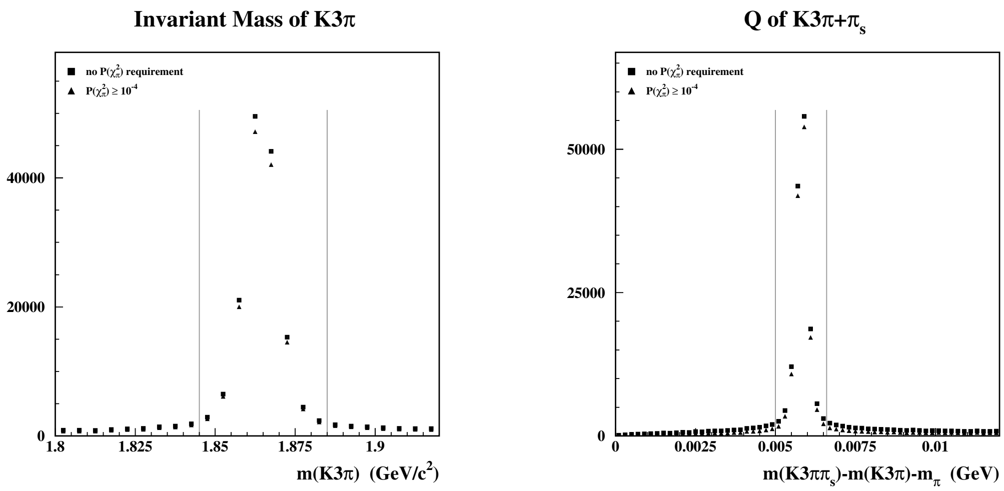


Figure 4.20: $P(\chi^2)$ of the slow-pion refit for signal and background. The background shown comes from continumm Monte Carlo events not containing the decay $D^0 \to K3\pi$. The distributions are normalized to unit area. The bin between 0 and 0.01 is suppressed so that the rest of the shape can be seen. Almost three times as much of the background populates the first bin as the signal, primarily at zero.

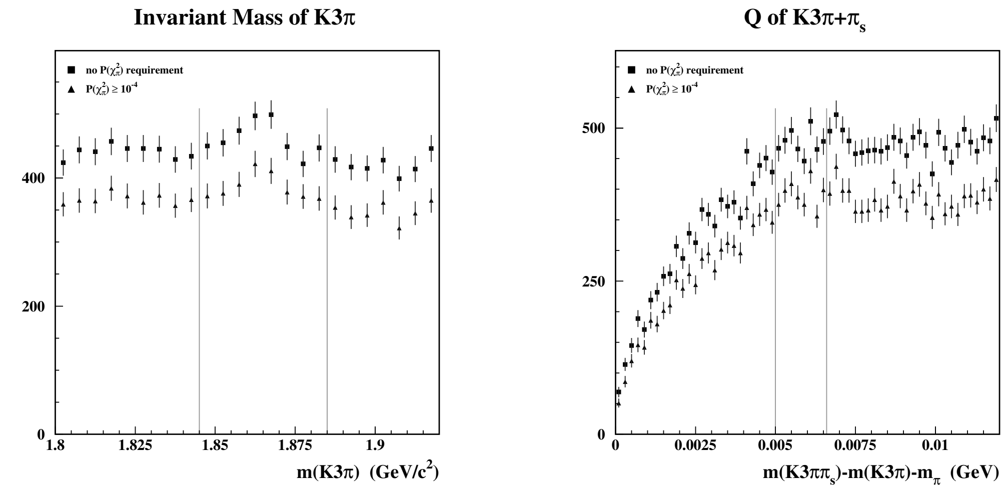
Table 4.2: Summary of D^{*+} selection criteria and acceptances.

Quantity	**Requirement**	**Signal Acceptance (%)**		**Background Rejection Factor**
		RS(Data)	**RS(MC)**	**WS(MC)**
energy release	$Q \leq 12\text{MeV}$	—	—	—
energy release uncert.	$\sigma_Q \leq 500\text{keV}$	99.9	99.9	1.0
momentum	$p \geq 2.68\text{GeV}/c$	84.2	85.6	1.6
slow-pion refit qual.	$P(\chi^2_{\pi_s}) \geq 10^{-4}$	94.3	95.2	1.2

We can select these decays explicitly in the Monte Carlo, and Figure 4.23 shows their mass and Q distributions after passing all of the other analysis requirements. The widths of these distributions are the same as for the signal, so it is impossible to separate this background from the signal in the fit. The branching fraction for $D^0 \to \bar{K}^0 K^+ \pi^- \to \pi^+\pi^- K^+\pi^-$ is $(1.7 \pm 0.3) \times 10^{-3}$[24, 25], compared with the right-



(a) Right-Sign from Generic Monte Carlo



(b) Wrong-Sign from Generic Monte Carlo

Figure 4.21: Right-sign and wrong-sign mass and Q distributions with and without the $P(\chi^2)$ requirement on the slow pion refit. This is more useful for removing multiple-scattering effects than suppressing backgrounds. All other selection criteria are applied.

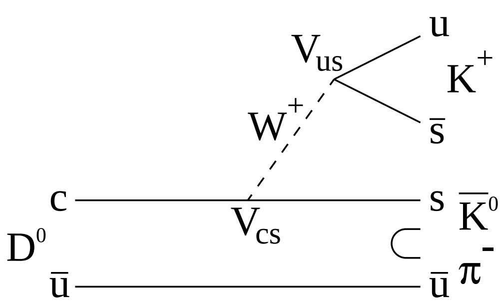
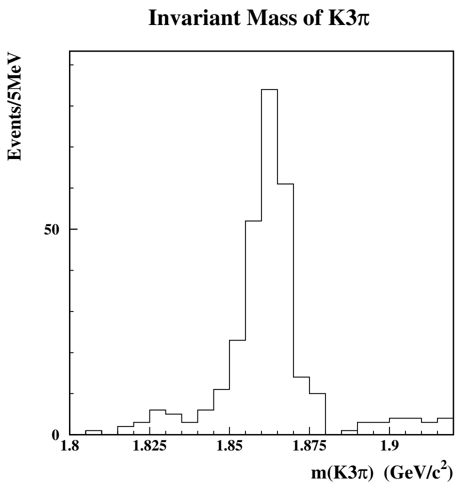
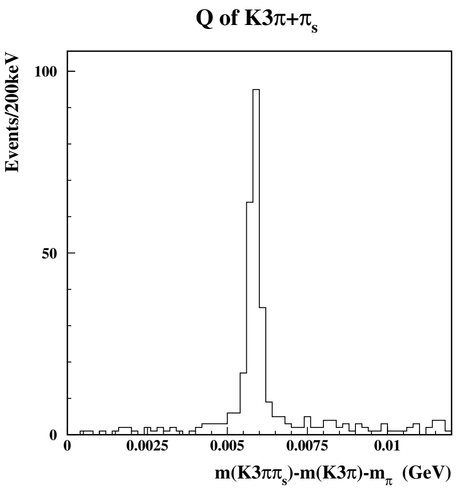


Figure 4.22: Tree-level diagram for the decay $D^0 \to \bar{K}^0 K^+ \pi^-$.



(a) Mass



(b) Q

Figure 4.23: Mass and Q for MC events containing $D^0 \to K^+\pi^-\pi^+\pi^-$ via a K^0_S . The widths of the distributions are comparable to signal.

sign branching fraction of 7.6×10^{-2}. We are not efficient for reconstructing this mode, with the common-vertex constraint rejecting much of the background from $\bar{K}^0$, so we will not see the full enhancement to R_{WS} of $(2.2 \pm 0.4) \times 10^{-2}$. From Monte Carlo, we predict an actual enhancement to R_{WS} of $(0.07 \pm 0.02) \times 10^{-2}$.

We can form a $\pi^+\pi^-$ invariant mass at the four-particle vertex. This mass is shifted slightly downward, because the K_S^0 tends to move at least some distance from its creation point, and forcing its decay products back to the D^0 vertex decreases the opening angle between the tracks from its true value. From Monte Carlo simulation (Figure 4.24), this shift is slightly over $5\text{MeV}/c^2$, to $492\text{MeV}/c^2$. The width of the

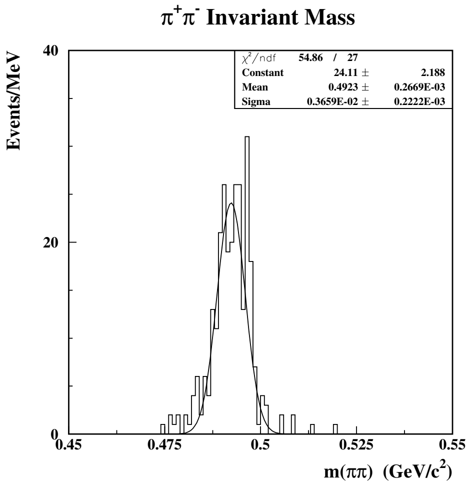


Figure 4.24: Invariant mass of $\pi^+\pi^-$ combinations for tagged K_S^0's reconstructed at the four-track vertex.

distribution is about $4\text{MeV}/c^2$, so rejecting $\pi^+\pi^-$ combinations within $20\text{MeV}/c^2$ of the shifted K_S^0 mass completely eliminates this background while rejecting only 10% of the signal.

The CP-conjugate final state $D^0 \to K^0 K^- \pi^+$ also appears in the right-sign, at a level much below the Cabibbo-favored $D^0 \to K^-\pi^+\pi^-\pi^+$ mode. It is also suppressed by the above requirement.

4.5 Determination of R_{WS}

4.5.1 Backgrounds

There are three general categories of backgrounds in this analysis. The first consists of properly reconstructed $\bar{D}^0$ mesons combined with a random slow pion. This background has the correct shape in mass, since the $K3\pi$ combination has been reconstructed properly with the correct mass assignments, but in Q it has a threshold shape with a turn-on at 0MeV. Because of its shape in the plane of Q vs. mass, we commonly refer to this background as the "ridge". The ridge also has the same lifetime distribution and resonant sub-structure as the right-sign signal.

The second background category consists of all other contributions from charmed mesons and baryons. Monte Carlo simulation shows that D^0 mesons dominate this background.

The final background is everything else. Having removed the contribution from $\Upsilon(4S) \to B\bar{B}$ with the D^{*+} momentum requirement, we are left with events of the type $e^+e^- \to u\bar{u}, d\bar{d}, s\bar{s}$. As with the "other charm" background, this does not peak, and has no resemblance to the signal in phase-space. It is further distinguished by having a mean lifetime of 0, since all decays to stable particles occur via strong interactions.

4.5.2 Fitting Procedure

Because we have a signal that peaks in two variables, $m(K3\pi)$ and Q, and backgrounds that peak in neither or one, we benefit from performing a two-dimensional fit. Rather than fit to complicated analytical forms, we instead use histograms for each fit component. This procedure also simplifies the number of parameters in the fit, as we only have one normalization for each component. This is essentially the same technique that has been used in previous analyses[9, 26]. By extending the fitting region well past the signal region in both variables, we get both a larger lever arm for fitting and a way to cross-check some of the results and assumptions. While the backgrounds are fit for absolute normalizations, the signal is fit directly for R_{WS}, since that is the quantity that interests us.

For the signal, we have a sufficiently clean sample of 13735 events from the right-sign. Because of the identical kinematics between the right- and wrong-sign signals, this will perfectly model the wrong-sign signal shape, within statistical fluctuations of the right-sign distribution, which should be small.

The background shapes come from Monte Carlo simulation with roughly seven times the statistics of the data. The simulation should properly represent the shapes arising from various sources of background, and comparisons between the predicted shapes and data in the sideband regions support this. However, it may not have the correct relative abundances of different particles and their decay modes. Many of the requirements described in Section 4.4 are specifically designed to eliminate backgrounds that have small-scale structure, so that all of the remaining backgrounds (which cannot be removed without removing most or all of the signal) only exhibit broad-scale variation. These variations are sufficiently similar for each general category of backgrounds that one shape can account for different relative amounts of the specific backgrounds in that category.

The fitter was tested extensively,[1] and found to produce unbiased estimators on toy Monte Carlo where the same smooth parent distributions were used to generate the simulated data and the shapes used to fit it. A somewhat pessimistic conclusion was reached regarding the fitter's ability to distinguish broad shapes in general, but the shapes used here do not appear to suffer from this problem.

4.5.3 Results

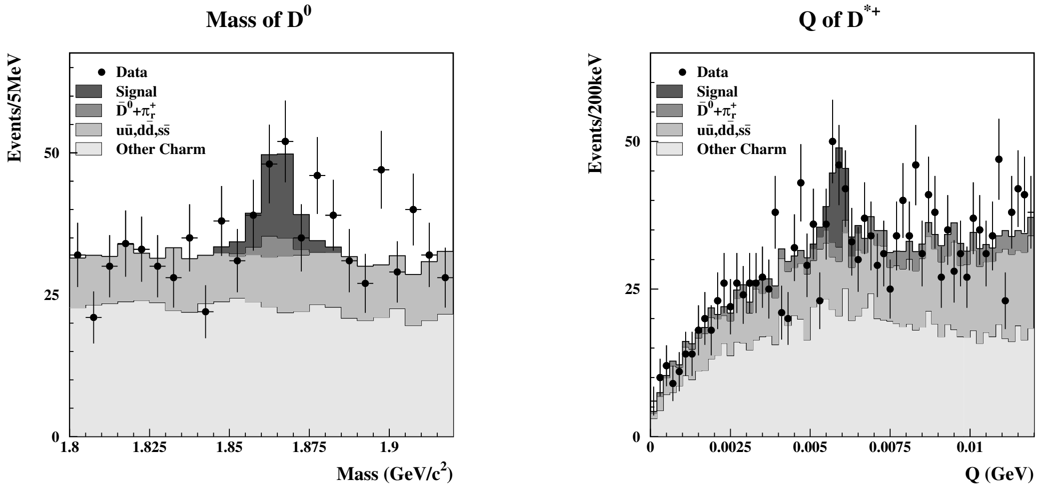
The final result of the fit is $R_{WS} = (0.408^{+0.121}_{-0.114}(\text{stat.}))\%$. If we fix R_{WS} to 0.0 and fit again, the increase in $-\log\mathcal{L}$ is 7.6, which gives us a statistical significance of 3.9σ. The fit results for all parameters are given in Table 4.3. Using the observed 13735 right-sign events and a one-dimensional fit in mass to a constant plus a double

[1]See Appendix A for complete details on testing the fitter.

Table 4.3: Parameters from the fit to wrong-sign data.

Component	Fit Result
R_{WS}	$(4.08^{+1.21}_{-1.14}) \times 10^{-3}$
total number of ridge events	$140.^{+43}_{-43}$
total number of other charm events	6127^{+304}_{-303}
ratio tot. # $u\bar{u}, d\bar{d}, s\bar{s}$: tot. # oth. c events	$0.548^{+0.079}_{-0.072}$

Gaussian in order to estimate the right-sign background, we obtain a total of 54 wrong-sign signal events. While the fit is performed on two-dimensional distributions, it is useful to look at projections of these distributions in mass and Q. This allows us to easily compare the fit results with the data. Figure 4.25 shows the mass and Q projections for the $\pm 2\sigma$ signal region in the other variable. The data is shown as



(a) Mass

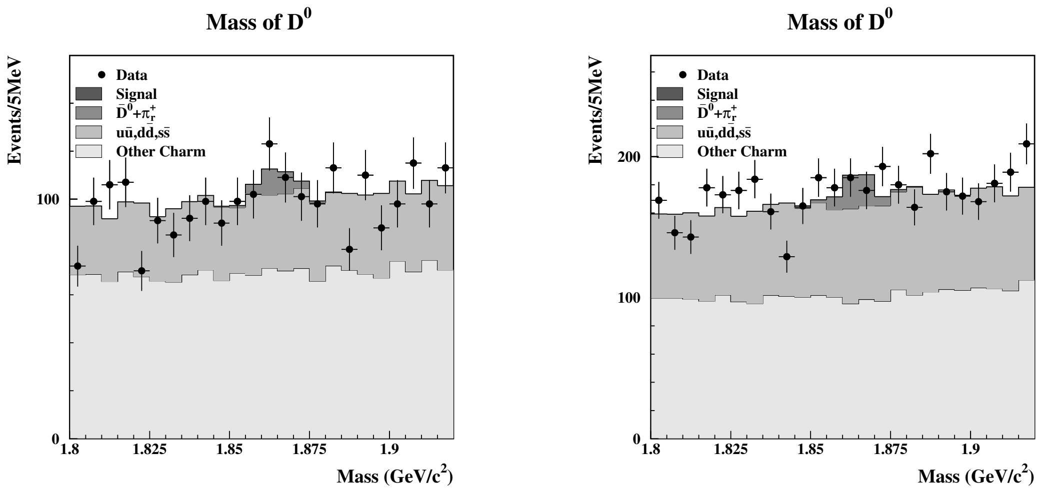
(b) Q

Figure 4.25: Wrong-sign fit results. Mass is plotted for $5.4\text{MeV} \leq Q \leq 6.2\text{MeV}$. Q is plotted for $1.855\text{GeV} \leq m(K3\pi) \cdot c^2 \leq 1.875\text{GeV}$. These are each approximately $\pm 2\sigma$ slices.

points with error bars, and the fit components are shown as histograms, normalized to their fit-result values. The error bars on the data points do not correspond to anything used in the fit, as the points are the results of summing over several bins, but are illustrative of the level of statistics present.

The dominant source of background in the fit is "other charm" events, which accounts for nearly two-thirds of the background. The eye is drawn to what appears to be a slight excess in the mass projection at just under 1.9GeV/c^2, but this feature is much too narrow to represent more than statistical fluctuation. A similar feature appears at around 8MeV in Q, but the high bins are staggered with bins very close to the predicted shapes. The four highest bins in Q show a possible effect, though not a statistically significant one.

Similar mass projections are shown in Figure 4.26 for values of Q below and above the signal region. The "ridge" formed of correctly reconstructed $\overline{D}^0$ mesons paired with uncorrelated slow pions is visible in the low-Q region especially. Projections in

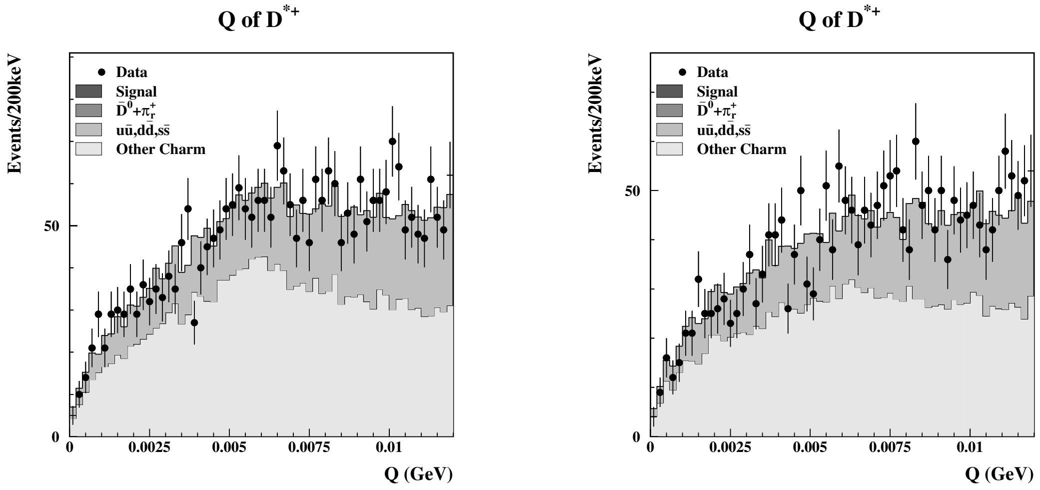


(a) Low Q Sideband

(b) High Q Sideband

Figure 4.26: Wrong-sign fit results in $K3\pi$ invariant mass for Q sidebands. The low Q sideband is $0.5\text{MeV} \leq Q \leq 4.5\text{MeV}$, and the high Q sideband is $7.1\text{MeV} \leq Q \leq 11.5\text{MeV}$.

Q for low- and high-mass regions are shown in Figure 4.27. Another excess, again not significant, appears for large values of Q in the high-mass sideband. It is possible that there is some small mis-modeling in the Monte Carlo which is responsible for this, and increased Monte Carlo statistics might resolve this issue.



(a) Low Mass Sideband

(b) High Mass Sideband

Figure 4.27: Wrong-sign fit results in Q for mass sidebands. The low mass sideband is $1.805\text{GeV} \leq m(K3\pi) \cdot c^2 \leq 1.841\text{GeV}$, and the high mass sideband is $1.889\text{GeV} \leq m(K3\pi) \cdot c^2 \leq 1.917\text{GeV}$.

It is worth noting that some number of right-sign candidates can be correctly matched with an appropriate slow pion as well as with an uncorrelated slow pion of the opposite charge. These events therefore will show up in both the right-sign and wrong-sign distributions, and will be double-counted. In the wrong-sign, these candidates will show up in the "ridge," and we can estimate their contribution to the total background. In the $\pm 2\sigma$ signal region of the $Q - M$ plane, the ridge shape from Monte Carlo contains 152 entries. Of these, 41 are really right-sign events from the decay $D^{*+} \rightarrow D^0 \pi_s^+$ (as opposed to $\bar{D}^0$ mesons from other sources). Multiple counting of events therefore represents about 27% of the ridge. This has no impact on the results of the fit; it merely affects the prominence of this particular background.

Confidence Level

Because we are doing a likelihood fit rather than a χ^2 fit, there is no clearly defined confidence level based on the resultant likelihood. We can, however, attempt to form an ensemble of likelihoods from a toy Monte Carlo[27, 28]. If the Monte Carlo shapes that we are using are a fair representation of the data, then we can start with a "parent" distribution which is the result of the fit. By fluctuating each bin according to Poisson statistics and fitting the new distribution as data, we can determine a typical expected likelihood. Many such fluctuations and fits can then be performed, and the fraction for which $-\log \mathcal{L}$ exceeds that of the fit to the actual data is the confidence level of the data fit. In ten thousand trials, one of the fits produce a higher $-\log \mathcal{L}$ than the data, so the confidence level from this technique is 1×10^{-4} (Figure 4.28).

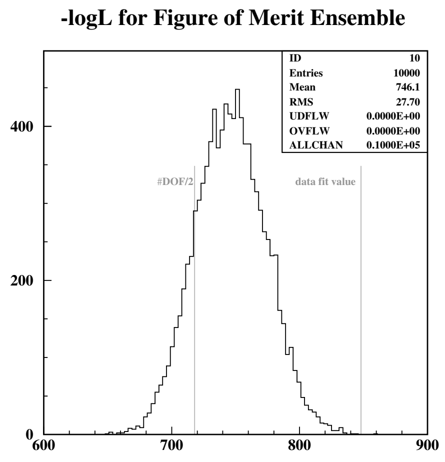


Figure 4.28: Distribution of $-\log \mathcal{L}$ for fits of simulated data produced by fluctuations of the scaled shapes from Monte Carlo. The actual data fit result and the χ^2-limit expected mean are also shown.

This technique is not completely reliable, as it is highly susceptible to statistical fluctuations in the fitting shapes. This can lead to fluctuations about a local minimum which are distributed more narrowly than they might be if the initial shapes were smooth. The fit to the data looks good to the eye when viewed in projections, and the $-\log \mathcal{L}$ distribution is noticeably non-Gaussian, so the reasonable statement would be that the likelihood we obtain from the fit to the data is close to what we would expect based on our Monte Carlo sample.

Comparison of Monte Carlo Predictions and Fit Results

The statistics of the fitting shapes from Monte Carlo are given in Table 4.4, along with

Table 4.4: Monte Carlo background statistics, predicted background fractions, and fit background fractions.

Background	**Number of Entries**	**Fraction**	
		Predicted	**Fit**
ridge	2306	0.0214	0.0145
other charm	76349	0.7084	0.6368
$u\bar{u}, d\bar{d}, s\bar{s}$	29121	0.2702	0.3487
total	107776	—	—

the predicted and fit fractions of the backgrounds. This shows that the simulation does a reasonable job of predicting the background composition. From the right-sign signal, we can estimate the effective size of the Monte Carlo sample. The Monte Carlo has sufficient information that we can count the number of true right-sign decays that we reconstruct. The data, on the other hand, requires that we determine the background level. The background in mass is very flat, so we can easily fit the right-sign data to a double Gaussian plus a constant. When this constant is subtracted from the mass distribution, we are left with our best estimate of the total right-sign signal yield in data. Dividing the Monte Carlo signal yield by this number gives us the estimate that there is 11.1 times as much Monte Carlo as data, which is our right-sign MC-to-data scaling factor. We can make dead-reckoning estimates of the background by dividing the total Monte Carlo predictions by the right-sign scaling factor of 11.1. Table 4.5 summarizes these estimates in two sideband regions. Shown

Table 4.5: Monte Carlo dead-reckoning background predictions.

Region	**Predicted Background using N_{RS} to determine MC Luminosity**	**Data Results**	
		Fit Values	**Total Results**
signal mass, sideband Q	1492	1442.1	1474
sideband mass, signal Q	516	505.3	499

are the dead-reckoned predictions, the fit results, and the data yield. Signal regions are approximately $\pm 2\sigma$, and sidebands consist of everything outside of approximately $\pm 4\sigma$. The Monte Carlo predictions and actual background levels agree to within the statistical uncertainty of the data, so the total background is also well-predicted by the Monte Carlo. This level of agreement is excellent, and shows that the Monte Carlo simulates both the production mechanisms and detector interactions much better than we would expect.

Efficiency Ratio of Right-Sign vs. Wrong-Sign

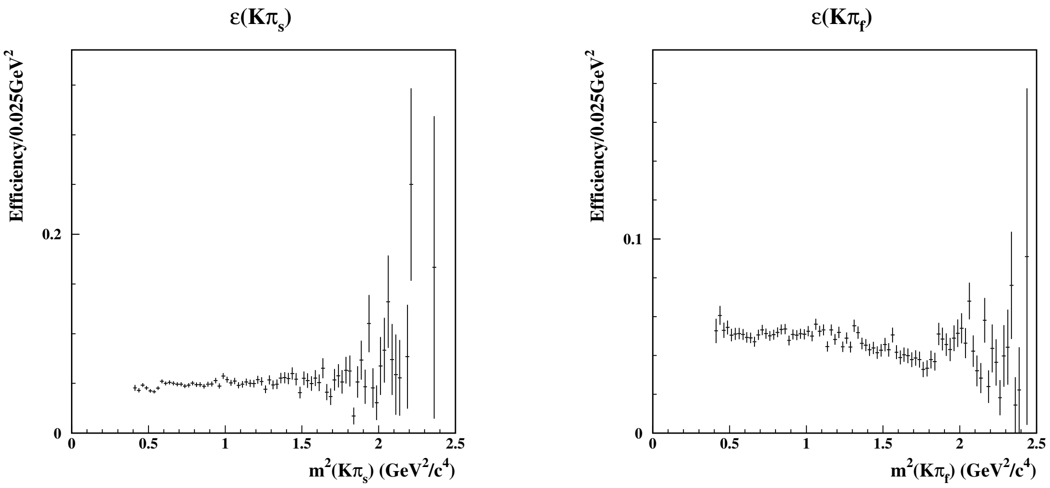
There is, in principle, an efficiency correction that must be made to R_{WS} due to the differing phase-space distributions of the right-sign and wrong-sign. This difference arises because while the mixed contribution has the same structure, the doubly Cabibbo-suppressed contribution can have different amplitudes and phases for the intermediate resonances. This is not an issue for the $D^0 \to K^+\pi^-$ analysis, since there are no intermediate resonances, but it was studied extensively for the $D^0 \to K^+\pi^-\pi^0$ analysis[26]. The efficiency correction was found to be 1.00 ± 0.02, which supports the hypothesis that the resonant sub-structure is dominated by final-state interactions, which will be identical between mixing and DCSD. If this hypothesis is true, or the efficiency is flat across all of phase space, then an efficiency correction is not needed.

Because of the high dimensionality of the phase space and the relatively low statistics, we cannot examine the efficiency in the typical manner whereby we form bins of phase space and calculate the efficiency of each bin. Instead we use a method[27] that relies on the distribution of the data and small regions of phase space surrounding them ("bubbles"). From a large sample of signal Monte Carlo generated with flat phase space, we can count the number of events in these bubbles, as well as the number of these events which would have been reconstructed successfully. These then give us the local efficiencies, or their inverses, event weights. The sum of these weights is the number of "actual" events present in the data, which when divided into the number observed gives us the average efficiency for the decay mode.

This is sufficient in the right-sign, where the backgrounds are low enough to be safely neglected. In the wrong-sign, however, the backgrounds are comparably large, so some care must be taken to remove their effects. In order to do this, we take advantage of the events which are close to the nominal mass of the D^0, but not to the nominal Q of the D^{*+}. Since the phase space distribution of the four-track combination does not depend on the slow pion from the D^{*+}, these events will have the same phase space distribution as the backgrounds under the signal, to the extent that the relative background abundances are consistent. We can then calculate the total event weight of the sideband region, scale it by the ratio of background areas in the signal and sideband regions, and subtract the result from the wrong-sign weight, which yields the wrong-sign signal weight. For the number of observed signal events, we take the right-sign area and scale by the fit result for R_{WS}. What remains of the data in the wrong-sign signal region is taken as the number of background events. This then is independent of further reliance on Monte Carlo simulation of the backgrounds.

Following this procedure, we obtain an average right-sign efficiency of 0.05207 ± 0.00007, and an average wrong-sign efficiency of 0.04889 ± 0.00476, for an efficiency ratio of 1.07 ± 0.10. The size of the "bubbles" used in this is $0.2\mathrm{GeV}^2/c^4$, which corresponds to a volume of under 1% of the total phase space.

An efficiency correction around unity is not unexpected. A simple study comparing expected yield to observed wrong-sign candidates (signal and background) using variable-sized bubbles shows consistency, though not quantifiably convincing, between right-sign and wrong-sign. Moreover, projections of the efficiency onto the eight meaningful two- and three-particle combinations (neutral and singly charged, Figures 4.29-4.33), with the exception of the small pieces of phase space cut out by



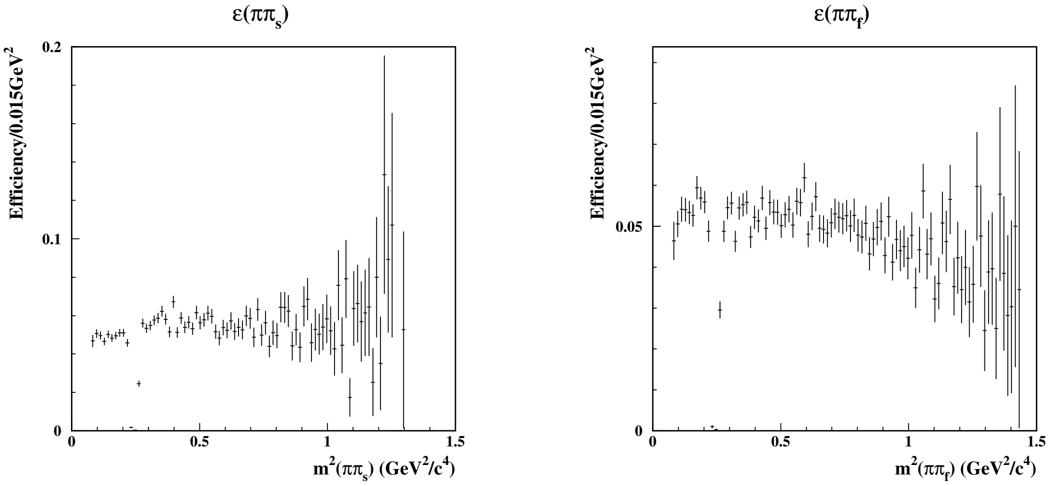
(a) Slower π

(b) Faster π

Figure 4.29: Efficiency for signal Monte Carlo generated with flat phase space projected onto the two neutral $K\pi$ axes.

the K_S veto, vary smoothly by no more than 10-20%. There does not appear to be any significant resonant contribution near the K_S mass, and up to the vetoed region the efficiency again appears flat, so this does not give us cause for concern. Given the near equivalence of the four daughter particles of the D^0, it is not surprising that we see no large variations in efficiency across phase space.

We have no reason to expect the efficiency ratio to differ from one, and given the great degree of observed consistency between right- and wrong-sign efficiencies, we take the ratio to be exactly one as a theoretical assumption, and so there is no



(a) Slower π (b) Faster π

Figure 4.30: Efficiency for signal Monte Carlo generated with flat phase space projected onto the two neutral $\pi\pi$ axes. The K_S veto is easily seen around $0.25\mathrm{GeV}^2/c^4$.

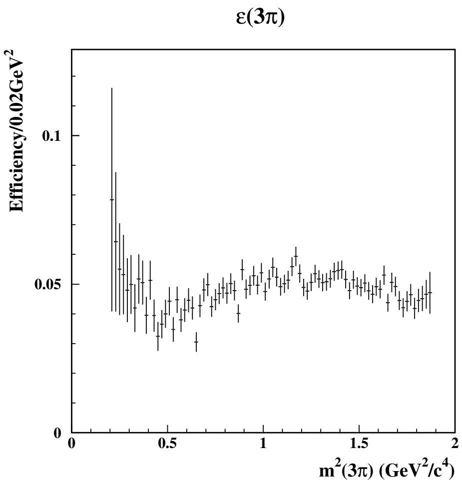
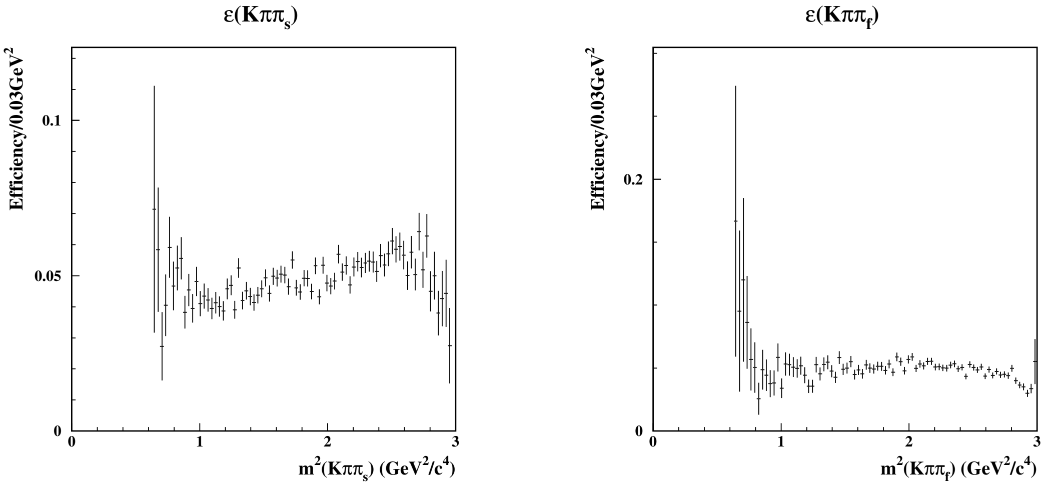


Figure 4.31: Efficiency for signal Monte Carlo generated with flat phase space projected onto the singly charged 3π axis.



(a) Slower π (b) Faster π

Figure 4.32: Efficiency for signal Monte Carlo generated with flat phase space projected onto the two singly charged $K\pi\pi$ axes with one of the pions having the same charge as the kaon.

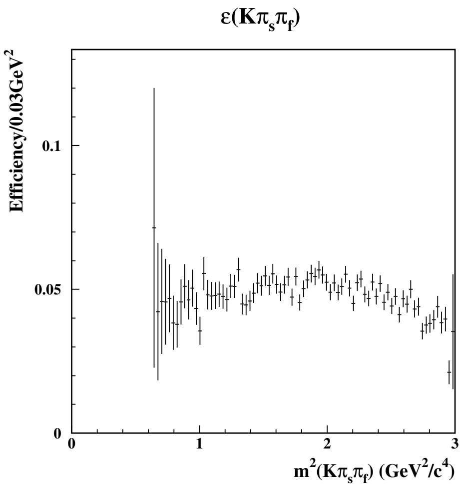


Figure 4.33: Efficiency for signal Monte Carlo generated with flat phase space projected onto the singly charged $K\pi\pi$ axis with both of the pions having the opposite charge as the kaon.

multiplicative correction to R_{WS}. This is consistent with the findings of the previous search for this decay[21].

4.5.4 Variation of R_{WS} with Selection Criteria

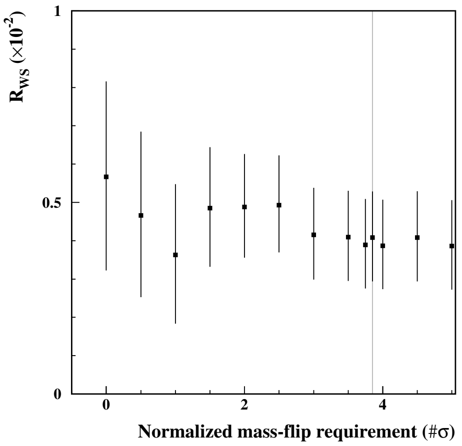
As a check, it is useful to observe the fit value of R_{WS} as we vary some of the selection criteria. We choose the quantities most likely to affect our result: kinematic reflections ($\tilde{m}$), particle identification, $K3\pi + \pi_s$ momentum, and $\pi^+\pi^-$ invariant mass.

As the mass-flip requirement is changed, we will allow different amounts of feed-through to appear in the wrong-sign distributions. Feed-through events are doubly mis-identified, so they will be tagged in the Monte Carlo as mis-reconstructed D^0 decays, which are classified as part of the "other charm" background. This means that as the requirement changes, the "other charm" fitting shape will change to accomodate the feed-through, and we would expect R_{WS} to be insensitive to the variations, as is shown in Figure 4.34(a). We can, alternatively, prevent the feed-through events from being included in any of the fitting shapes, which allows us to fit for the level of feed-through observed as we vary the mass-flip criterion. The mass distribution of feed-through events is very broad compared with the wrong-sign signal, however, so our usual fitting procedure does not work. Instead we perform one-dimensional fits of projections onto the Q axis. For the right-sign, we only consider candidates within 2σ of the D^0 mass, and the projection is fit with a single Gaussian. For the wrong-sign, we integrate over all masses in the fit region, and fit to a threshold function plus two Gaussians. The ratios of wrong-sign to right-sign area are then the feed-through rate, and they are shown in Figure 4.34(b).

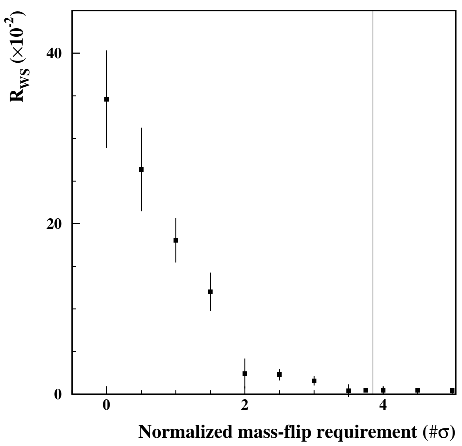
The particle identification requirements are primarily a matter of signal cleanliness. Hence, we would not expect to see a systematic effect in our fit value of R_{WS} as these requirements are varied. This is indeed the case, as is shown in Figure 4.35.

Similarly, varying the D^{*+} momentum requirement should not have a significant effect on R_{WS} . Figure 4.36 confirms this.

The $\overline{K}^0$ veto specifically eliminates a background which is nearly identical to the



(a) Normal Fitting Procedure



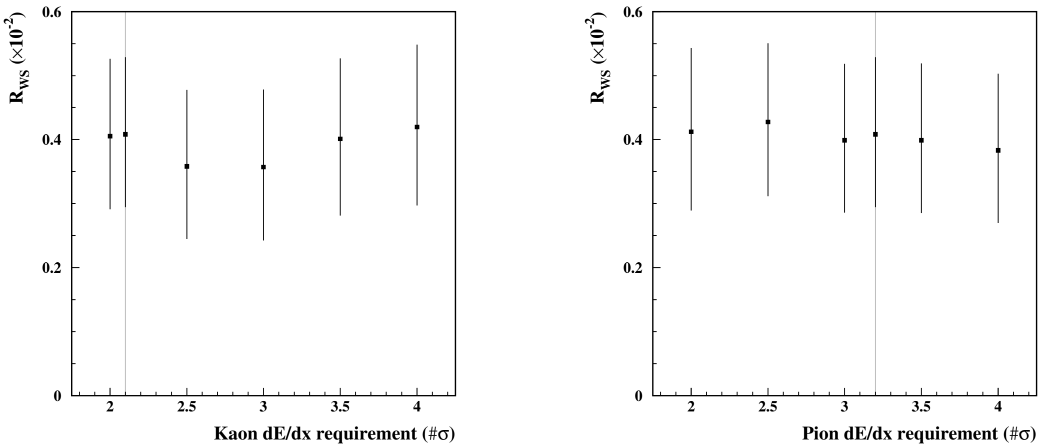
(b) Fitting Specifically for Feed-Through

Figure 4.34: Variation in R_{WS} as a function of "mass-flip" rejection. The vertical line indicates the nominal requirement. The normal fitting procedure places feed-through from the Monte Carlo in the "other charm" distribution, so it does not appear significantly in the fit value for R_{WS} . Fitting specifically for feed-through requires that we not label these events as "other charm."

signal. Consequently, loosening this veto should enhance the observed value of R_{WS} . From Section 4.4.4, we predict an excess of $(0.07 \pm 0.02)\%$. Figure 4.37 shows the variation of R_{WS} as the K_S veto is changed, and the observed excess is approximately 0.095%, which is consistent with our prediction.

4.6 Study of Systematic Errors

One advantage of measuring a ratio is that many of the sources of systematic uncertainty cancel. The remaining sources to consider are the mis-modeling of the quark fragmentation in Monte Carlo simulations, the statistics of the Monte Carlo simulation, and the Monte Carlo-derived shape distributions. A summary of all sources of systematic uncertainty is given in Table 4.6.



(a) Kaon

(b) Pion

Figure 4.35: Variation in R_{WS} as a function of dE/dx requirements for (a) kaons and (b) pions. The vertical lines indicate the nominal requirements.

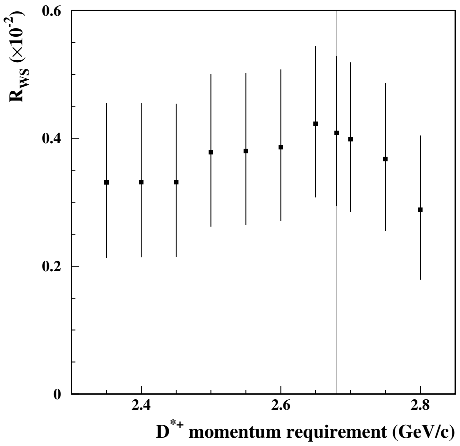


Figure 4.36: Variation in R_{WS} as a function of the D^{*+} momentum requirement. The vertical line indicates the nominal requirement.

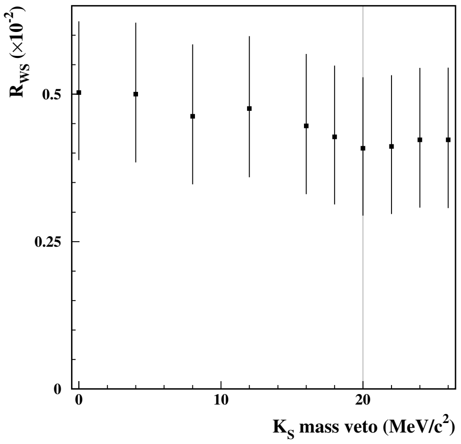


Figure 4.37: Variation in R_{WS} as a function of the K_S^0 rejection. The vertical line indicates the nominal requirement.

Table 4.6: Sources of Systematic Error

Source	**Systematic on R_{WS}**
fragmentation in MC	8×10^{-6}
MC statistics	3.2×10^{-4}
MC background distributions	3.1×10^{-4}
Total	0.044×10^{-2}

4.6.1 Fragmentation in Monte Carlo

It is known that the Monte Carlo does not produce the correct momentum spectrum for prompt D^{*+} mesons, though the difference is not especially large (Figure 4.38). Ideally, we would produce Monte Carlo with the correct fragmentation, but this is computationally prohibitive. Even producing targeted Monte Carlo for the specific decay modes would take a long time.

Because the differences are small, we can approximate the fragmentation model effects by separating the backgrounds containing charm into bins of transverse momentum. Each bin is assigned a weighting that will produce the correct final spectrum,

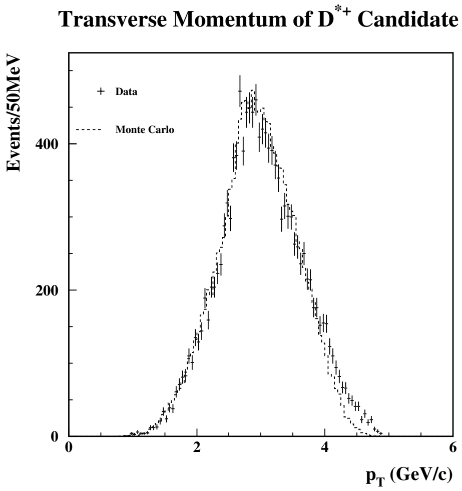


Figure 4.38: Comparison of momentum spectra of D^{*+} mesons before correcting for mis-modeling of fragmentation in the Monte Carlo.

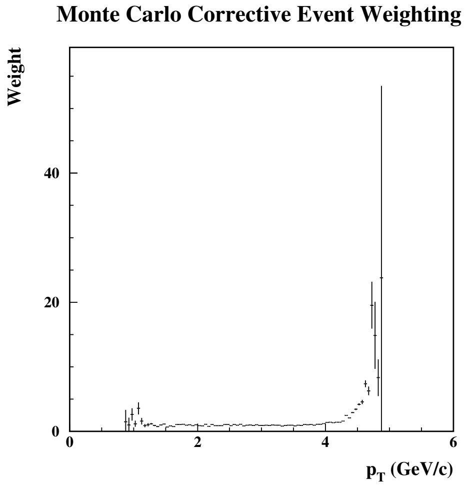
which is used when constructing the $Q - M$ histograms.

Specifically, if a given momentum bin p contains m_p Monte Carlo signal candidates and d_p data signal candidates (in right-sign, they are all signal to first order), the weight for Monte Carlo candidates in that bin is d_pM/m_pD, where M and D are the total number of Monte Carlo and data candidates (Figure 4.39).

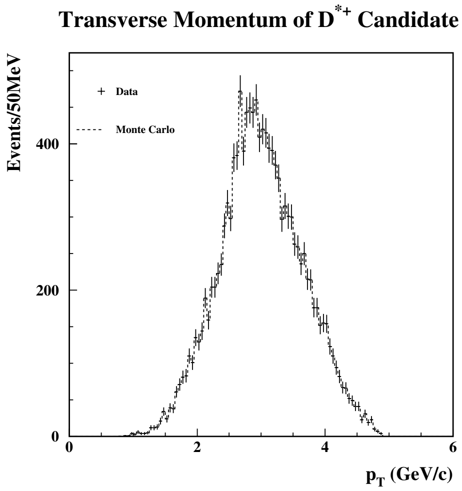
With the fragmentation-corrected shapes, the fit result for R_{WS} is 0.409%. For the systematic uncertainty we use the standard deviation of this result and the nominal fit result, which is $1/\sqrt{2}$ times their difference. We assign a systematic uncertainty on R_{WS} of 8×10^{-6}due to this effect, which is negligible.

4.6.2 Monte Carlo Statistics

The effects of finite Monte Carlo statistics can be investigated by performing Poisson-fluctuations on the bins of the various MC shapes. Treating the current bin contents as the "true" distribution is not, strictly speaking, correct, but should be a good approximation under the assumption that the statistical variation from the true mean values is not very large. This is, to a certain extent, circular, but if it turns out to be



(a) Correction Factors



(b) Corrected Momentum Spectrum

Figure 4.39: Correction factors and momentum spectra of D^{*+} mesons after applying the corrections. The post-correction distributions are identical by construction, but serve as a check that we are applying the scaling correctly.

a poor approximation, then there is a serious deficiency in the Monte Carlo statistics, which does not appear to be the case when looking at distributions and fits by eye.

Ten thousand fluctuations were performed, and the distribution obtained for the signal variation is shown in Figure 4.40. The width of this distribution, and the systematic on R_{WS} assigned, is 3.2×10^{-4}.

4.6.3 Monte Carlo Background Distributions

Estimating the effects of Monte Carlo mis-modeling on R_{WS} is a non-trivial exercise, as the Monte Carlo shapes are the only real handle we have to understand the backgrounds. The excellent agreement between predicted and observed background contributions, along with the general lack of features in the $Q - m(K3\pi)$ plane, indicate that this mis-modeling effect cannot be large.

The approach used in other wrong-sign analyses is to mask out various pieces of

Distribution of R_{WS} for MC Statistical Fluctuations

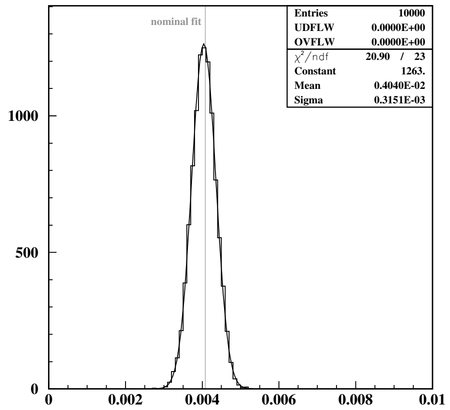


Figure 4.40: Distribution of signal parameter for fits to fluctuations of the shapes from Monte Carlo. The un-fluctuated (nominal) fit result is also shown.

the fitting plane, and observe the variation in the central value of R_{WS}. This suffers from the additional complication that we are restricting the statistics of our fit, which will accentuate statistical fluctuations in the data. The regions used for the various fits are shown in Figure 4.41, and the corresponding fit results for R_{WS} are shown in Table 4.7. The sample standard deviation of these restricted fits (excluding the nominal fit) is 3.1×10^{-4}, which is what we take as a conservative systematic on R_{WS}.

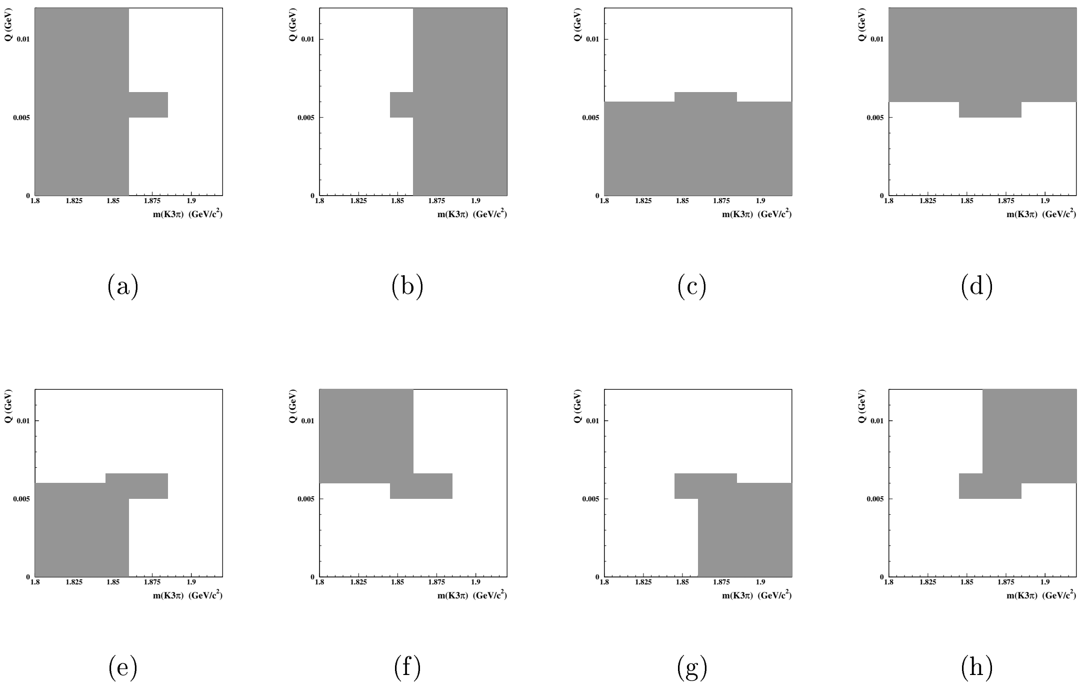


Figure 4.41: Masked fit regions. The shading indicates the areas used in the fit.

Table 4.7: Fit results with masks. Regions fit correspond to those shown in Figure 4.41.

Fit Region	**Fit Result for $R_{WS}(\times 10^{-2})$**
All	$0.408^{+0.121}_{-0.114}$
a)	$0.436^{+0.140}_{-0.134}$
b)	$0.412^{+0.122}_{-0.115}$
c)	$0.423^{+0.127}_{-0.121}$
d)	$0.410^{+0.124}_{-0.118}$
e)	$0.498^{+0.155}_{-0.150}$
f)	$0.444^{+0.147}_{-0.141}$
g)	$0.431^{+0.128}_{-0.122}$
h)	$0.399^{+0.126}_{-0.120}$

Chapter 5

Conclusions

We have searched for the decay $D^0 \to K^+\pi^-\pi^+\pi^-$ in the CLEO II.V dataset, with a total integrated luminosity of 9.0fb^{-1}. Our result is $R_{WS} = (0.408^{+0.121}_{-0.114}(\text{stat.}) \pm 0.044(\text{sys.}))\%$, which has a statistical significance of just under 4σ. While insufficient to claim the first observation of this mode, it constitutes very strong evidence, and is a considerable improvement on the result of the previous search, $(0.25^{+0.36}_{-0.34} \pm 0.03)\%$ from Fermilab E791[21].

Mixing in the $D^0 - \overline{D}^0$ system remains elusive. A number of measurements have been made recently comparing "wrong-sign" hadronic modes to their corresponding "right-sign" modes[9, 26, 30], all with relative rates comparable to each other as well as the result of this analysis. These analyses present a challenge in extracting limits on mixing, due to the interference between mixing and doubly Cabibbo-suppressed decays. In the two body modes this results in limits on rotated quantities, with the phase rotation still unknown. In the three and four body modes, it is difficult to get even this far. Higher backgrounds make any extraction of lifetime information difficult, and the interference terms contain resonance-by-resonance effects, rather than a single physically understandable phase.

Considerable progress has been made in the hadronic CP eigenstates[29, 31, 32, 33], placing tighter constraints on the lifetime-splitting y. Reconciling these results with the CP-averaged wrong-sign modes indicates a potentially large phase difference between the Cabibbo-favored and doubly Cabbibo-suppressed decays.

While no longer the most sensitive probes of mixing, several semileptonic analyses from CLEO and FOCUS are expected to shed additional insight on mixing in the charm sector via direct measurements of R_{mix}. Other techniques sensitive to most or all of the mixing parameters are also being developed.

A summary of the current published limits on x and y is depicted graphically in Figure 5.1. The semileptonic results are from E791 [5]. For the band defined by the $D^0 \to K^+K^-$ analyses, the results from FOCUS [7] and E791 [8] are combined with a weighted average. Two sets of limits are shown from the CLEO II.V $D^0 \to K^+\pi^-$ analysis [9]. The more restrictive limits are from a fit which does not allow CP violation. The looser limits are from a fit where CP violation is allowed. In both cases, the strong phase that rotates x and y into x' and y' is assumed to be zero when displaying the results.

With their larger datasets, the B factories Belle and BaBar should be able to further constrain, if not measure, charm mixing. The possibility of running a dedicated charm factory at CESR presents another unique opportunity, both from greatly reduced backgrounds and possibly exploitable coherence effects due to particular initial states such as the ψ''. It is likely that this next generation of experiments will either observe mixing or confine it within the bounds of the Standard Model.

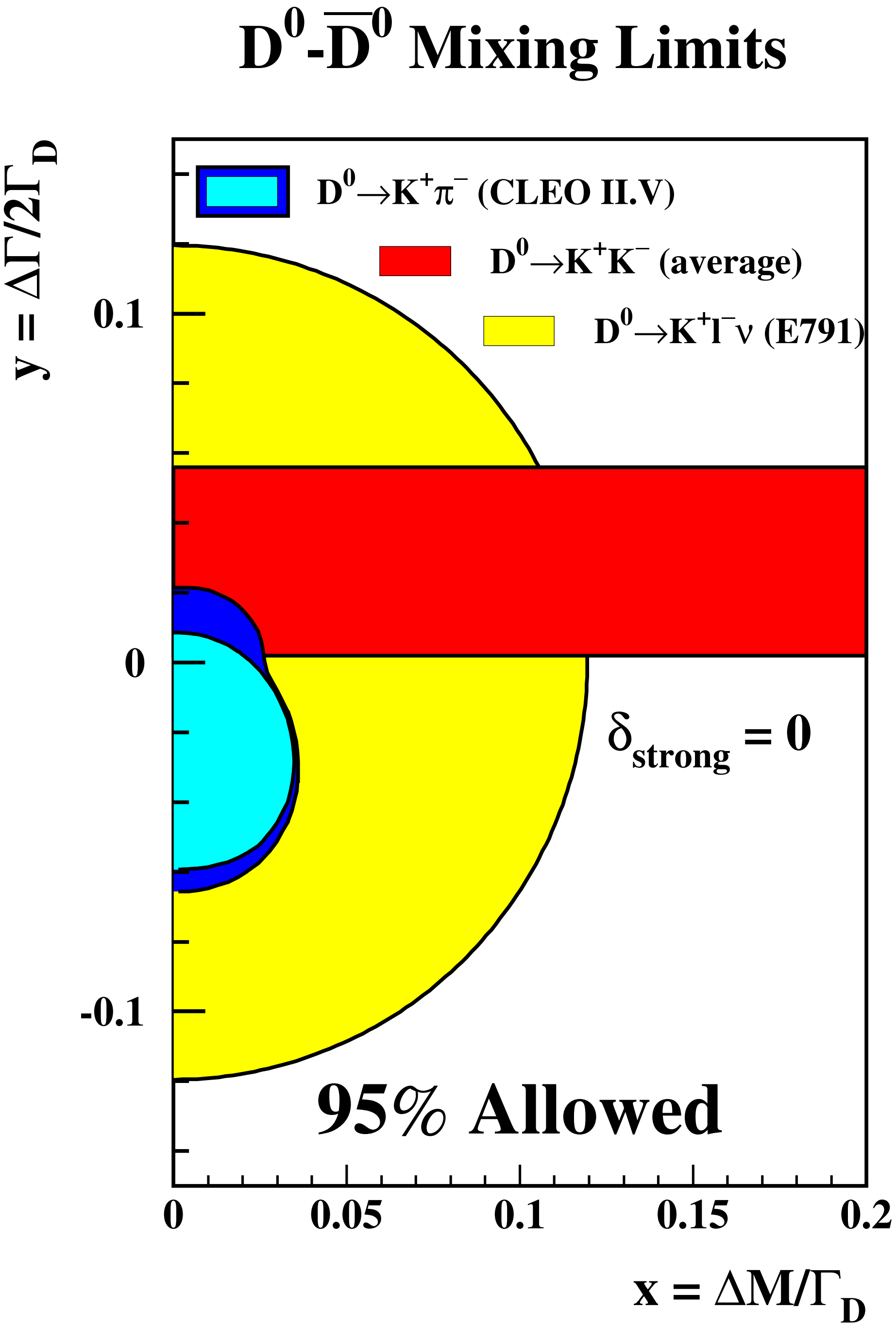


Figure 5.1: Limits on the mixing parameters x and y. A strong phase of zero is assumed for comparison between the $K^+\pi^-$ and the K^+K^- results.

Appendix A

Testing the 2D Fitter for $K3\pi\pi_s - K3\pi$ Energy Release vs. $K3\pi$ Invariant Mass

A.1 Introduction

A central piece of the wrong-sign hadronic searches for charm mixing is a two-dimensional fit in the $m(D^0) - Q(D^{*+})$ plane. This fit uses shape histograms obtained from the right-sign data (for the signal) and a large sample of Monte Carlo (for the backgrounds).

In general, we separate the background into three samples: good $\bar{D}^0$ candidates combined with a random π^+ (which form a "ridge" narrow in mass and broad in Q), other events originating from a $c\bar{c}$-pair, and events originating from lighter quarks. Each of these samples provides a shape histogram which is scaled in the fit to match a portion of the background in data.

While this fitting procedure has already been used in other wrong-sign D^0 hadronic analyses[9, 26]. the tests of its efficacy were performed by varying the data with Poisson fluctuations and making "pull" distributions assuming that the fit's central value on the un-modified data is the correct value. These pulls have tended to look reasonable, but there is always the possibility that the "correct" central values are biased, and the fluctuations are moving the fit results around in a local minimum, rather than around the true value.

In order to have greater certainty that the results of the fit can be trusted, it is necessary to perform tests approximating the conditions of the data, but where the true central values are known.

A.2 General Procedure

There are two main characteristics that a fitting package must have in order to be a useful tool: good estimators of the fit parameters and reasonable error estimates. We test these by running repeated "toy Monte Carlo" experiments with shapes rougly approximating what we expect to see in the data. Using the notation of [20], we have a set of parameters $\boldsymbol{\alpha}$ with estimators $\hat{\boldsymbol{\alpha}}$ obtained from a maximum likelihood fit, and their errors $\mathrm{Var}(\hat{\boldsymbol{\alpha}})$ (henceforth referred to as $\hat{\boldsymbol{\sigma}}_{\alpha}^2$).

Given the nature of the fit (discrete measurements in bins with differing probability distribution functions), the Central Limit Theorem does not guarantee a Gaussian p.d.f. for the estimators $\hat{\boldsymbol{\alpha}}$. The estimators will, in fact, be distributed according to Poisson statistics, since a binned maximum likelihood fit for Poisson-distributed data conserves the area of the histogram being fit.

A.2.1 Likelihood Function

Basic Likelihood Function

The basic likelihood function that we use is a bin-by-bin Poisson distribution of the data d, where the mean μ is the sum of the bin contents of the shape histograms, scaled appropriately by the current fitting parameters:

$$\mathcal{L} = \prod_{j=1}^{N_{mass}} \prod_{k=1}^{N_Q} \frac{(\mu_{jk})^{d_{jk}} e^{-\mu_{jk}}}{d_{jk}!} \tag{A.1}$$

where $\mu_{jk} = \sum_{i=1}^{N_{shapes}} c_i N_{ijk}$ is the prediction for the bin, and c_i is the (scaled) fit parameter.

Extended Likelihood Function

The basic likelihood function is sufficient when the fitting histograms have large statistics. For lower-statistics shapes, we have to modify the likelihood function with additional terms that account for the Monte Carlo bin contents themselves being observed quantities subject to fluctuations. The full likelihood function is:

$$\mathcal{L} = \prod_{j=1}^{N_{mass}} \prod_{k=1}^{N_Q} \frac{(\mu_{jk})^{d_{jk}} e^{-\mu_{jk}}}{d_{jk}!} \prod_{i=1}^{N_{shapes}} \frac{(P_{ijk})^{N_{ijk}} e^{-P_{ijk}}}{N_{ijk}!} \tag{A.2}$$

and μ_{jk} is composed of the P_{ijk} rather than the N_{ijk}.

Finding the Parent Distributions P_{ijk}

In (A.2) we have additional unknown quantities P_{ijk} which must be determined. To do this, we need to maximize the likelihood for each P_{ijk}. In practice, it is simpler to minimize $-\ln \mathcal{L}$:

$$\begin{aligned} -\ln \mathcal{L} = \sum_{j=1}^{N_{mass}} \sum_{k=1}^{N_Q} \Bigg\{ & -d_{jk} \ln \left[\sum_{i=1}^{N_{shapes}} c_i P_{ijk} \right] \\ & + \sum_{i=1}^{N_{shapes}} c_i P_{ijk} \\ & + \sum_{i=1}^{N_{shapes}} \left[-N_{ijk} \ln P_{ijk} + P_{ijk} \right] \Bigg\} \\ & + \text{constant} \end{aligned} \tag{A.3}$$

It is then simple to compute $\partial(-\ln \mathcal{L})/\partial P_{mjk}$ and set it to zero, giving us

$$\left[1 - \frac{d_{jk}}{\sum_{i=1}^{N_{shapes}} c_i P_{ijk}} \right] c_m - \frac{N_{mjk}}{P_{mjk}} + 1 = 0 \tag{A.4}$$

which can be further simplified to

$$0 = [\mu_{jk}(c_m + 1) - d_{jk}c_m]\, P_{mjk} - \mu_{jk}N_{mjk} \tag{A.5}$$

where μ_{jk} is defined in section A.2.1.

This set of non-linear coupled equations, which must be solved for each bin and set of $\{c_i\}$, can be solved numerically by a routine such as CERNLIB's DSNLEQ[36].

A.2.2 Estimators and Biases

Total Area from Simple Likelihood

We can, in principle, choose different sets of estimators for the fits. The fit parameters $\{c_i\}$ are one set of estimators, but we could equally well choose a set $\{a, f_i\}$ where a is the overall normalization, and f_i is the fractional contribution to the total area of shape i (with $\sum f_i = 1$). In this latter case, a is an unbiased estimator for the simple likelihood function given in (A.1). We can see this as follows:

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial a} &= \sum_{j=1}^{N_{mass}} \sum_{k=1}^{N_Q} \frac{\mathcal{L}\; d_{jk}!}{(\mu_{jk})^{d_{jk}} e^{-\mu_{jk}}} \left[-(\mu_{jk})^{d_{jk}} + d_{jk}(\mu_{jk})^{d_{jk}-1}\right] \frac{e^{-\mu_{jk}}}{d_{jk}!} \frac{\partial \mu_{jk}}{\partial a} \\
&= \mathcal{L} \sum \sum \frac{\partial \mu_{jk}}{\partial a} \left(-1 + \frac{d_{jk}}{\mu_{jk}}\right) \\
&= \mathcal{L} \sum \sum \frac{\mu_{jk}}{a} \left(\frac{d_{jk}}{\mu_{jk}} - 1\right) \\
&= \mathcal{L} \sum \sum \frac{1}{a} (d_{jk} - \mu_{jk}) \\
&= 0
\end{aligned} \tag{A.6}$$

Since we're maximizing the likelihood, we are not interested in the case $\mathcal{L} = 0$, and the normalization estimator a is constant over all the bins, so this becomes

$$\sum_{j=1}^{N_{mass}} \sum_{k=1}^{N_Q} (d_{jk} - \mu_{jk}) = 0 \tag{A.7}$$

which, when the individual terms in the sums are split up and performed, tells us that the total area (which is a) is equal to the area of the data histogram. In other words, a is unbiased so long as the data is itself unbiased, as is the case for Poisson processes, and there are no bins for which $\mu_{jk} = 0$ and $d_{jk} \neq 0$, since including those bins causes the likelihood to vanish.

Estimators $\{c_i\}$ from Simple Likelihood

Considering the estimators we are using in the fit, we can ask if a particular c_i is biased. By replacing ∂a with ∂c_i everywhere in (A.6), we get

$$0 = \sum_{j=1}^{N_{mass}} \sum_{k=1}^{N_Q} \frac{N_{ijk}}{\mu_{jk}} (d_{jk} - \mu_{jk})$$

or, multiplying each term by the constant c_i,

$$0 = \sum_{j=1}^{N_{mass}} \sum_{k=1}^{N_Q} \frac{c_i N_{ijk}}{\mu_{jk}} (d_{jk} - \mu_{jk}) \tag{A.8}$$

We can sum (A.8) over i, and because $\mu_{jk} \equiv \sum_i c_i N_{ijk}$ we again see that area is conserved between the data and fit. However, we have no real information about the $\{c_i\}$, so we must conclude that they are, in principle, biased so long as there are at least two parameters. (One parameter c_1 is trivially the same as a.)

Total Area from Extended Likelihood

For the extended likelihood given in (A.2), it is simplest to test area conservation by considering the estimators $\{c_i\}$. $\partial\mathcal{L}/\partial c_i$ is almost identical, except that because μ_{jk} is the sum of $c_i P_{ijk}$ rather than $c_i N_{ijk}$, the final equation we get is

$$0 = \sum_{j=1}^{N_{mass}} \sum_{k=1}^{N_Q} \frac{c_i P_{ijk}}{\mu_{jk}} (d_{jk} - \mu_{jk}) \tag{A.9}$$

or

$$\sum_{j=1}^{N_{mass}} \sum_{k=1}^{N_Q} d_{jk} = \sum_{j=1}^{N_{mass}} \sum_{k=1}^{N_Q} \mu_{jk} \tag{A.10}$$

where the left-hand side is still the total area of the data histogram, but the right-hand side is the total area of the "parent" shape distributions. Since each c_i is normalized by $\sum_j \sum_k N_{ijk}$, this is not *a priori* the same as the total fit area.

This means that we must determine if each shape is also area-conserved when fit for the parent distribution. Summing the constraint equation (A.5) over the fitting shapes, we have

$$0 = \sum_m \left[\left(\mu_{jk}(c_m + 1) - d_{jk} c_m\right) P_{mjk} - \mu_{jk} N_{mjk} \right]$$

We can expand this and collect like terms as follows:

$$\begin{aligned}
0 &= \mu_{jk}\left[\sum_m c_m P_{mjk} + \sum_m P_{mjk}\right] - d_{jk}\sum_m c_m P_{mjk} - \mu_{jk}\sum_m N_{mjk} \\
&= \mu_{jk}^2 + \mu_{jk}\left[\sum_m P_{mjk} - \sum_m N_{mjk}\right] - d_{jk}\mu_{jk}
\end{aligned}$$

Because we are only concerned with bins where the total contribution from all shapes is non-zero (the others give us no information), we can further simplify this to

$$d_{jk} = \mu_{jk} + \sum_{m=1}^{N_{shapes}} (P_{mjk} - N_{mjk}) \tag{A.11}$$

We can see that if we sum (A.11) over the bins, we can apply (A.10) to get

$$\sum_{j=1}^{N_{mass}} \sum_{k=1}^{N_Q} \sum_{m=1}^{N_{shapes}} P_{mjk} = \sum_{j=1}^{N_{mass}} \sum_{k=1}^{N_Q} \sum_{m=1}^{N_{shapes}} N_{mjk} \tag{A.12}$$

This is an interesting result, because it says that the sum of the areas of the observed fitting histograms and the sum of the areas of their parent distributions are the same, which is *not* the same as saying that each fitting histogram and its individual parent distribution have equal area. Since the total area we need to consider is a weighted sum of these histograms, it tells us that, in general, the extended likelihood is *not* an area-conserving fit, and so *the total area of the fit is no longer an unbiased estimator.*

There is another important result of this. If we only have a single fitting shape, then its area will be conserved in the parent distribution, once again making area an unbiased estimator. This also makes the associated c_i unbiased. However, if the other fitting shapes are present (as empty histograms) and allowed to fluctuate, it is possible for them to fluctuate to non-zero parent distributions, again introducing a potential bias in the total area. Assuming that the routine used to solve the constraint equations (A.5) is well-behaved in these circumstances, any bias from this will be very small if present at all.

A.2.3 Goodness of Fit

Pull Distributions

Pull distributions are a popular way of verifying a fitting package's performance. A pull distribution is the distribution of normalized residuals $r_i = (y_i - F(x_i;\alpha))/\sigma_i$ where y_i is a measurement at the point x_i, σ_i is its error, and $F(x_i;\alpha)$ is the expected mean at x_i given parameter α. This distribution should be Gaussian with a mean of zero and a width of one if y_i has a Gaussian p.d.f. about $F(x_i;\alpha)$. The appeal of this sort of distribution is that it tells you both if the result is biased and whether the calculated errors reflect true variations.

How do these quantities correspond to what the fitter is doing? x_i is any individual toy MC experiment, and $F(x_i;\alpha)$ is independent of x_i, since the mean is just the true value about which we are fluctuating. If we consider y_i to be $\hat{\alpha}$, then $F(\alpha)$ is just α. σ_i is then the error from the fit.

Now for appropriateness. We have already seen in Section A.2.2 that the total area is not conserved in the fit, so even in the limit where the Poisson distribution is a Gaussian, we fail the test of y_i's p.d.f. having a mean of $F(\alpha)$. However, pulls still provide an easy way to see biases and verify that the errors correspond to the observed variances. Of course, since we are dominated by asymmetric errors, we have to define an asymmetric pull as:

$$pull = \begin{cases} \frac{\hat{n}-n}{\hat{\sigma}^+} & ,(\hat{n} < n) \\ \frac{\hat{n}-n}{\hat{\sigma}^-} & ,(\hat{n} \geq n) \end{cases} \tag{A.13}$$

Meaningful Test Quantities

What ultimately matters is that the parameter estimators have biases which are small in comparison with their uncertainties, though of course the smaller the bias, the better. The most natural definition for the bias of a measurement is $b_i = \hat{n}_i - n_i$. As usual, $\hat{n}_i$ is the estimator for the parameter n_i, the number of events contributed by shape i.

The significance of the bias is $b_i/\sqrt{\sigma^2/M}$, where M is the number of experiments and σ^2 is the variance in n_i, which is just n_i for a Poisson or Gaussian distribution. Assuming the bias is the only problem with the estimator, this will approximately equal the RMS-squared of the bias distribution, which should be $\langle\hat{n}_i\rangle$.

In practice, we will mostly concern ourselves with pull distributions, though for the single-shape fits the bias plots will be included as well.

A.2.4 Fitting Shapes

Each experiment must be statistically independent of the others, as well as independent of the shape histograms used in the fit. When fitting the wrong-sign data, statistical independence of the shapes comes from using Monte Carlo simulation for the backgrounds, and the right-sign data for the signal. In the toy MC, it comes from generating shape histograms separately from the toy data. The histograms contain 55 bins in Q and 32 in D^0 mass, for a total of 1760 bins.

The toy fitting shapes are generated to contain some number N_i (for fitting shape i) of observed events. This is done using an acceptance-rejection technique[20] where a triplet of random numbers is generated, corresponding to a mass, Q, and height. If the height is less than the functional form for the given mass and Q, the event is accepted, otherwise a new triplet is thrown. This continues until the desired number of events have been accepted.

For the toy data histogram we cannot follow this same procedure, because we are fitting for n_{tot}, the true yield, and so must allow the data histogram to fluctuate about this value. What we do instead is generate one parent distribution and fluctuate the contents of each bin according to a Poisson distribution. We begin by generating perfect shape histograms for each fit component, which are shown in Figure A.1. We can then scale each distribution to the desired total area n_i and add them, obtaining an ideal data distribution, an example of which is shown in Figure A.2. It is this shape which is then subjected to a bin-by-bin fluctuation for each toy experiment. One such fluctuation is shown in Figure A.3.

A.3 Tools Used in the Tests

The basic fitting package used is MINUIT[35], with DSNLEQ[36] used to solve the non-linear parent-distribution equations (A.5). Both routines are part of the software package CERNLIB[34].

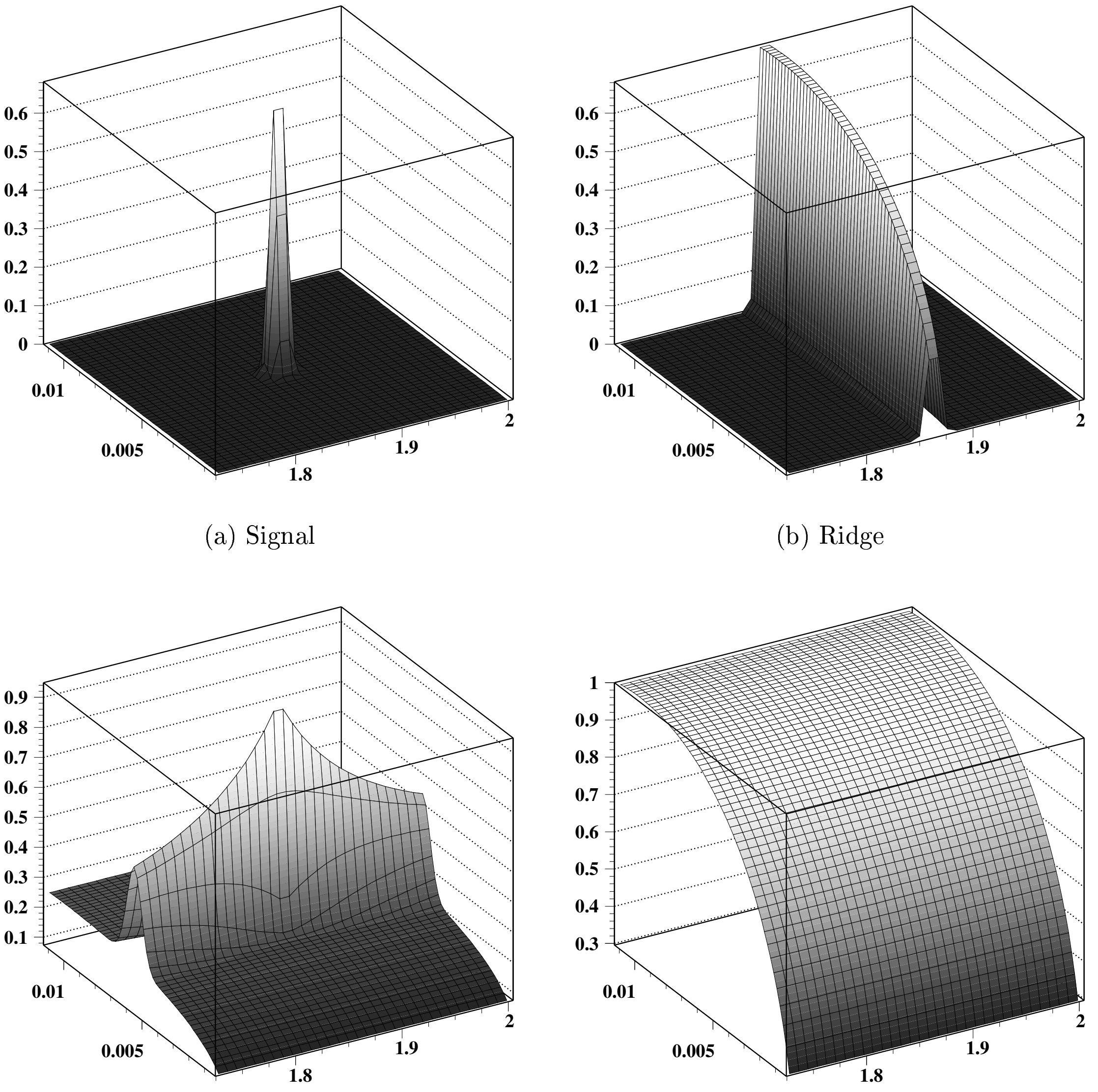
Random numbers are generated using RANLUX[37], also from CERNLIB. For Poisson distributions, the method described in [20] is used. Figure A.4 shows a distribution generated for $\mu = 4$ using this method. The error on the mean is the RMS divided by the square-root of the number of entries, which is 6×10^{-3}, considerably larger than the deviation of the mean from μ.

A.4 Single-Shape Fits

The simplest tests we can perform are to look at distributions where only one shape is present (the rest are fixed to be zero both in generation and in the fit). This is done for various data and fitting statistics for each of the four shapes.

A.4.1 Basic Likelihood Fits

We begin by looking at fits with only the basic form of the likelihood function given by equation (A.1). The stability of the fit is tested for each shape first by using a large-statistics fitting shape with decreasing amounts of data. For the background shapes, we use 8 million events in the fitting shape, and vary the data statistics from 2000 to 500. The results are shown in Figures A.5-A.10. For the signal shape, we maintain a constant data-to-fitting-shape ratio of 1:300, which reflects the expected yield in the actual data. We use data sample sizes from 100 to 20 to 10, the results of which are shown in Figures A.11-A.13.



(a) Signal

(b) Ridge

(c) Other Charm

(d) $u\bar{u}, d\bar{d}, s\bar{s}$

Figure A.1: Perfect Shape Histograms

With a single shape and high statistics in the fitting shape, the fitter gives good results even down to an average of less than one-third of an event per bin. We also see that the signal shape is stable even with relatively low fitting shape statistics, or at least as low as we are ever likely to see. The pulls for the signal, however, do not accurately reflect the fitter's performance, as the plots of fit results show.

We can now reduce the fitting shape statistics to 80,000 and again look at the three background shapes. Because we already know that the fits are stable down to 500

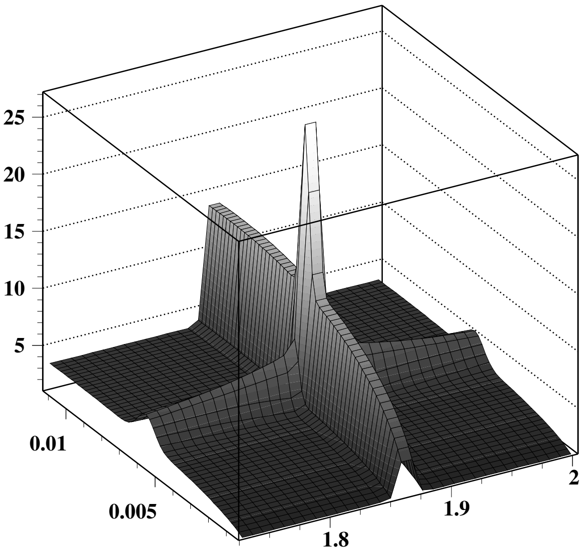


Figure A.2: Perfect Data Histogram

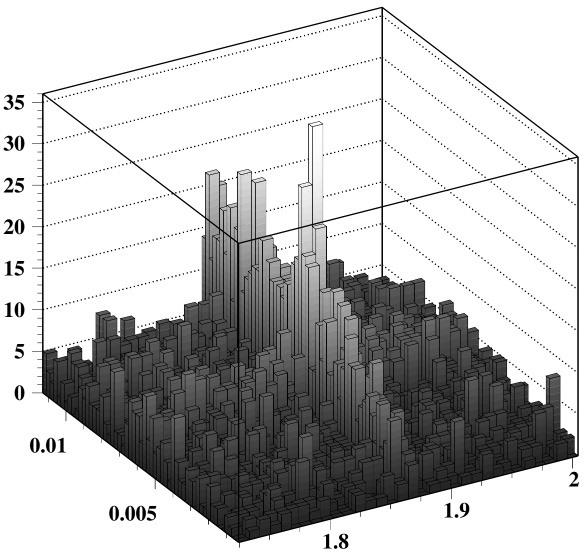


Figure A.3: Example of a Data Fluctuation

data events, we will only look at these results, which are shown in Figures A.14-A.16.

Again, the results are very reasonable. We can further reduce the statistics of the fitting shapes to 5000 events, which is shown in Figures A.17-A.19. This corresponds to a Monte Carlo sample with ten times the data statistics, which is comparable to what we will have when fitting the data.

The fits for the broad shapes, which have about three events per bin in the fitting histograms, are badly biased. The Poisson likelihood fit is supposed to be area

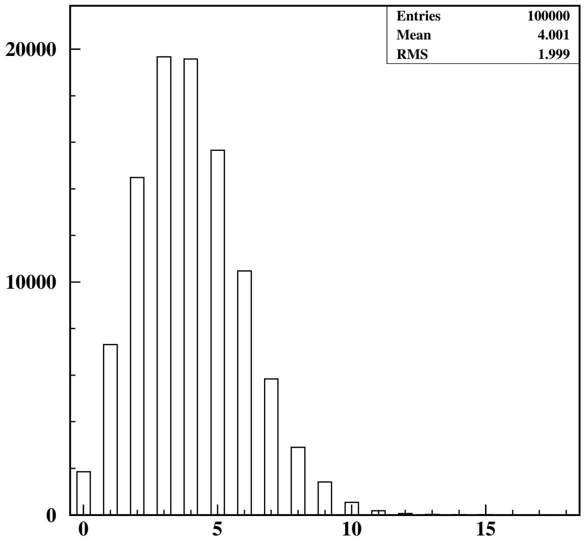


Figure A.4: Randomly-Generated Poisson Distribution for $\mu = 4$

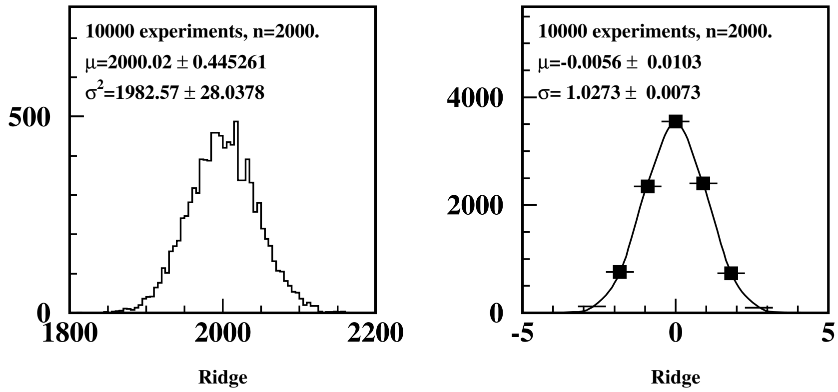


Figure A.5: $N_{ridge} = 8,000,000$; $n_{ridge} = 2000$ (no shape fluctuations)

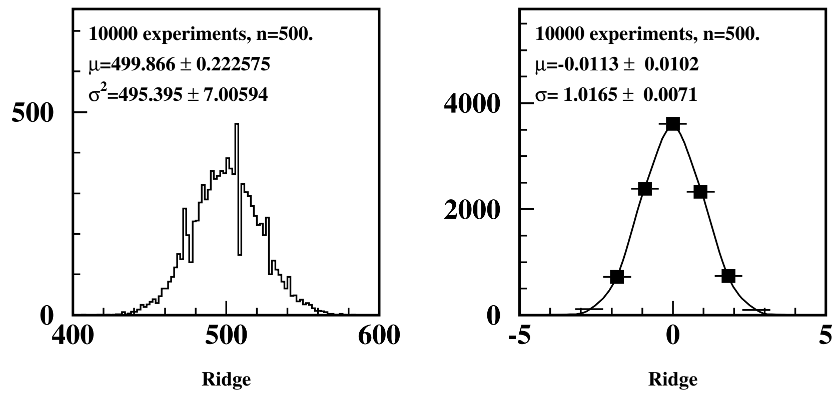


Figure A.6: $N_{ridge} = 8,000,000$; $n_{ridge} = 500$ (no shape fluctuations)

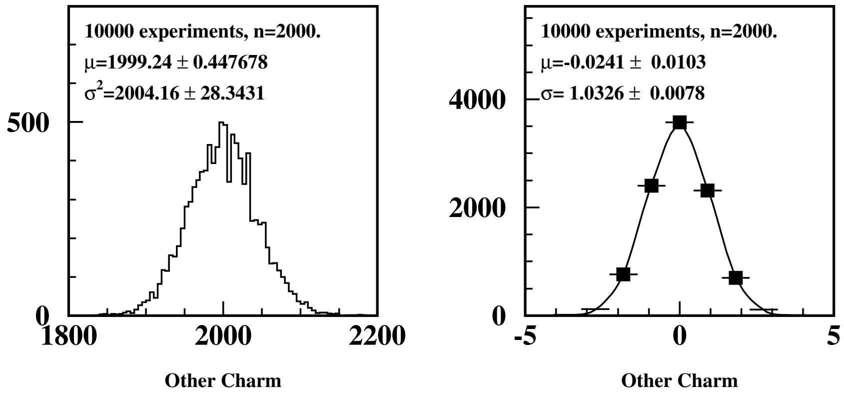


Figure A.7: $N_{charm} = 8,000,000$; $n_{charm} = 2000$ (no shape fluctuations)

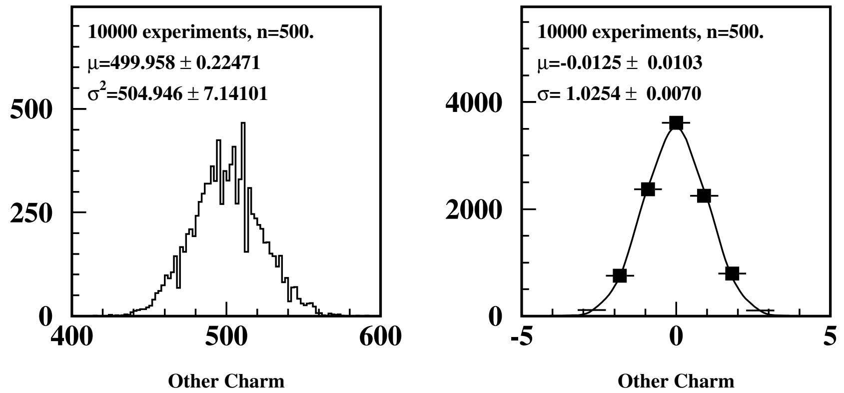


Figure A.8: $N_{charm} = 8,000,000$; $n_{charm} = 500$ (no shape fluctuations)

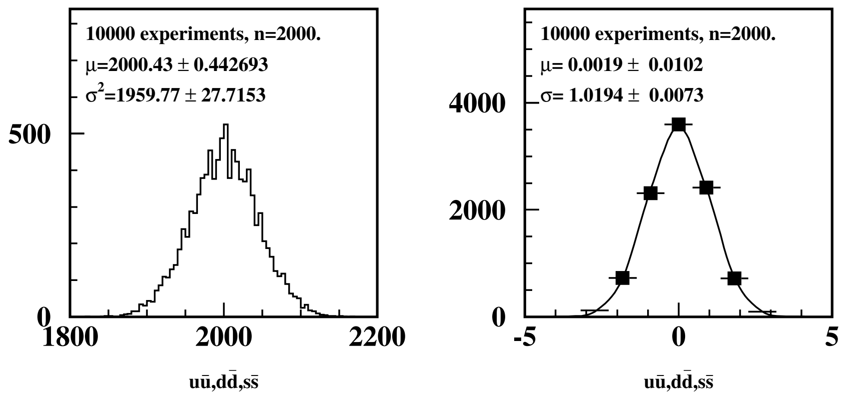


Figure A.9: $N_{u\bar{u},d\bar{d},s\bar{s}} = 8,000,000$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 2000$ (no shape fluctuations)

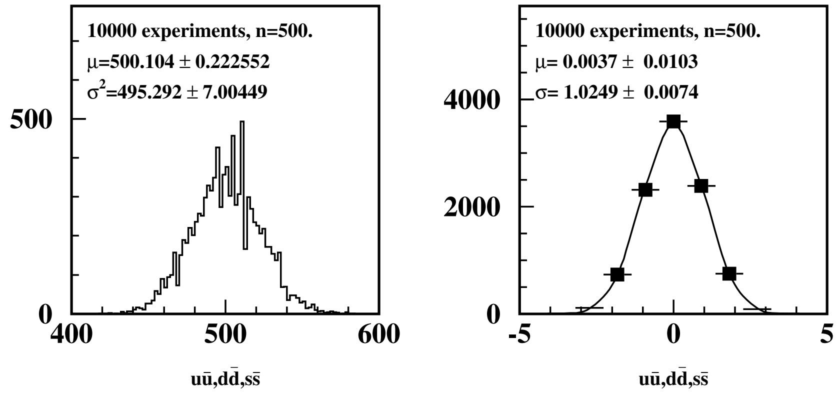


Figure A.10: $N_{u\bar{u},d\bar{d},s\bar{s}} = 8,000,000$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 500$ (no shape fluctuations)

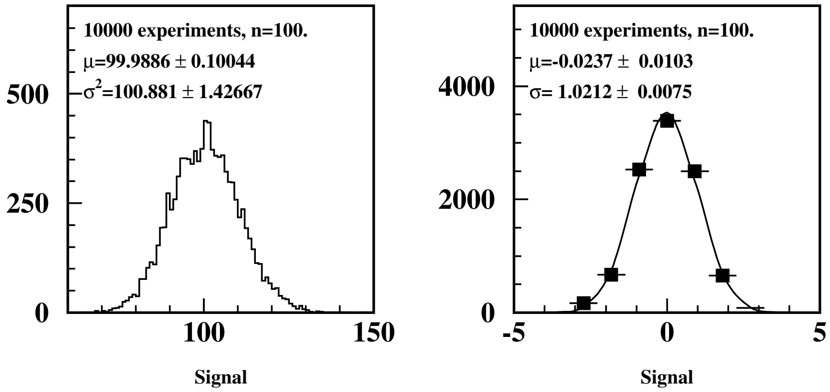


Figure A.11: $N_{signal} = 30,000$; $n_{signal} = 100$ (no shape fluctuations)

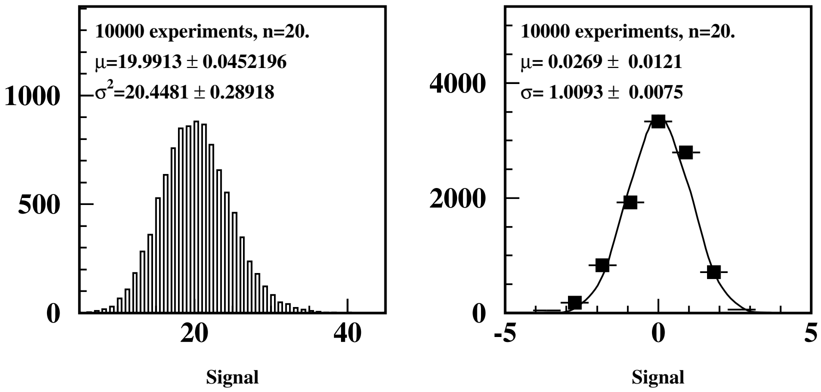


Figure A.12: $N_{signal} = 6000$; $n_{signal} = 20$ (no shape fluctuations)

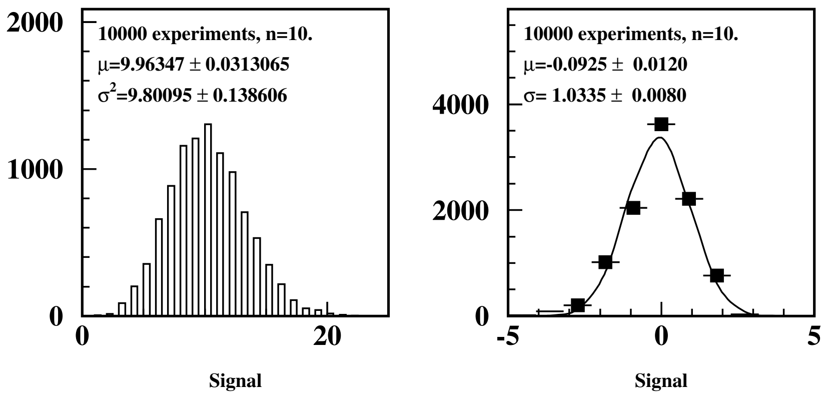


Figure A.13: $N_{signal} = 3000$; $n_{signal} = 10$ (no shape fluctuations)

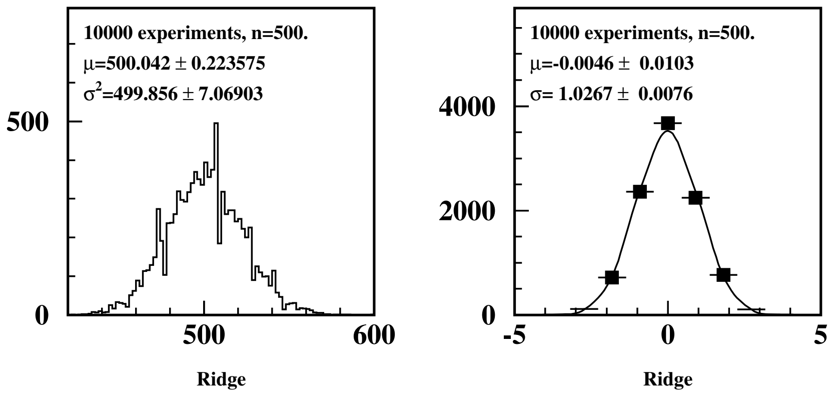


Figure A.14: $N_{ridge} = 80,000$; $n_{ridge} = 500$ (no shape fluctuations)

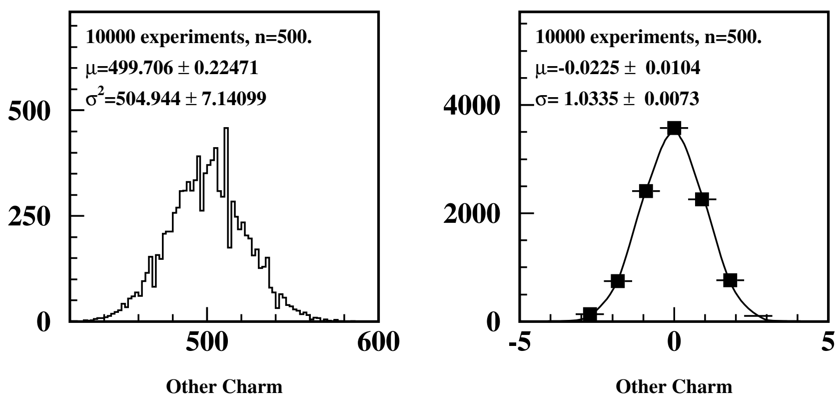


Figure A.15: $N_{charm} = 80,000$; $n_{charm} = 500$ (no shape fluctuations)

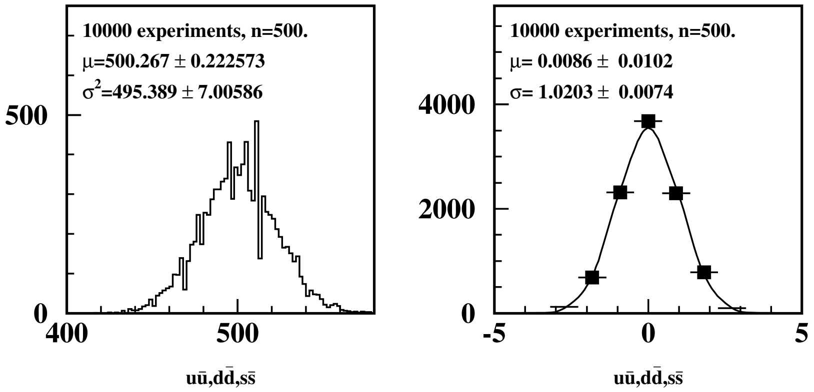


Figure A.16: $N_{u\bar{u},d\bar{d},s\bar{s}} = 80,000$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 500$ (no shape fluctuations)

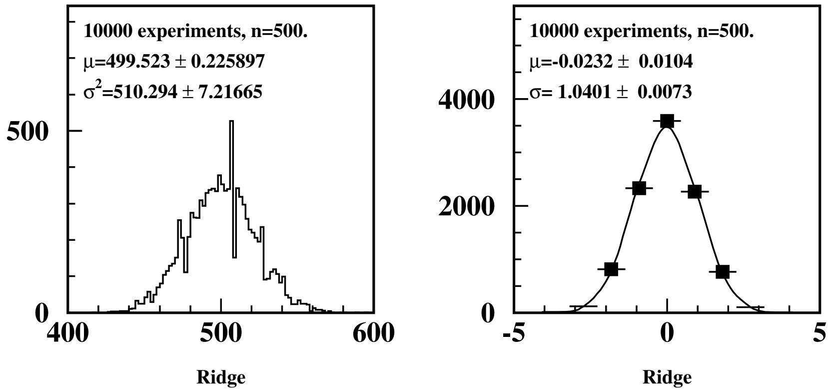


Figure A.17: $N_{ridge} = 5000$; $n_{ridge} = 500$ (no shape fluctuations)

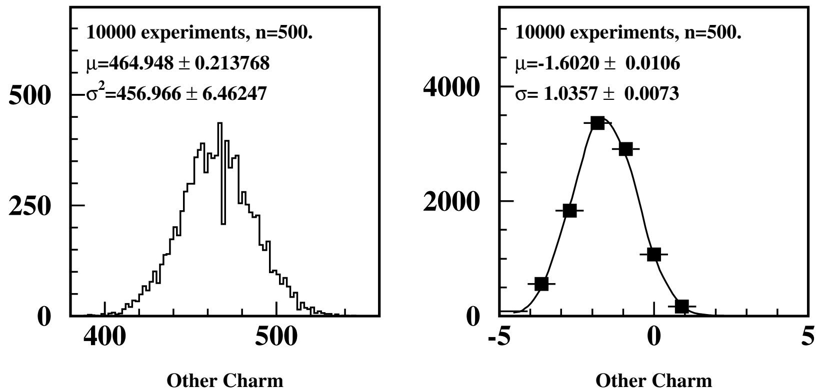


Figure A.18: $N_{charm} = 5000$; $n_{charm} = 500$ (no shape fluctuations)

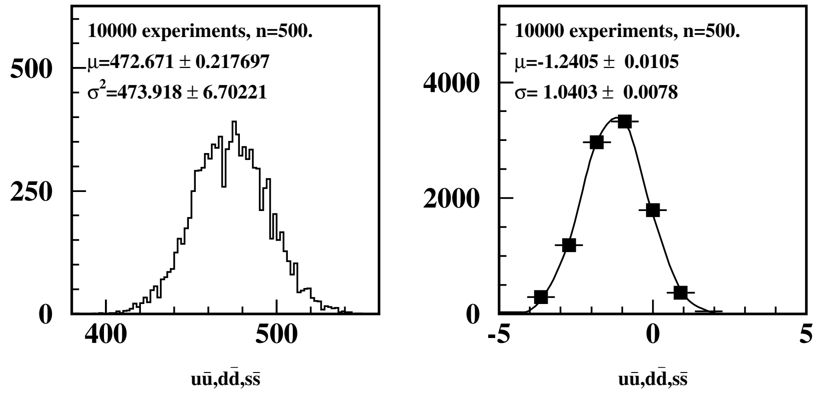


Figure A.19: $N_{u\bar{u},d\bar{d},s\bar{s}} = 5000$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 500$ (no shape fluctuations)

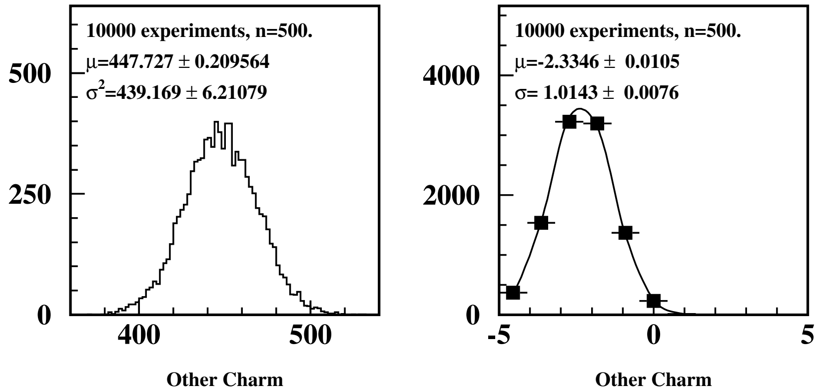


Figure A.20: $N_{charm} = 5000$; $n_{charm} = 500$ (allowing shape fluctuations)

conserving, but this is only true for the bins with non-zero fitting shape statistics, because the likelihood is not well-defined otherwise. With an average of three events per bin, there is a probability of e^{-3} to measure no events in a single bin, which means that about 1/20 of the fitting shape's bins will be zero. This then becomes an under-estimation of the total area of about 5%, which is roughly what we see. This should, in principle, be corrected by using the extended likelihood (A.2).

A.4.2 Extended Likelihood Fits

As mentioned in the previous section, low fitting shape statistics can cause a deficit in the total fit area. If we allow for the finite statistics of the shapes in the likelihood function, a bin with no events in the fitting shape can fluctuate up to agree with the data. We can see the effect of this on the two biased shapes in Figures A.20 and A.21.

The biases are actually worse, indicating that the fitter is finding parent distributions with more zero bins where the data is non-zero than the original distribution.

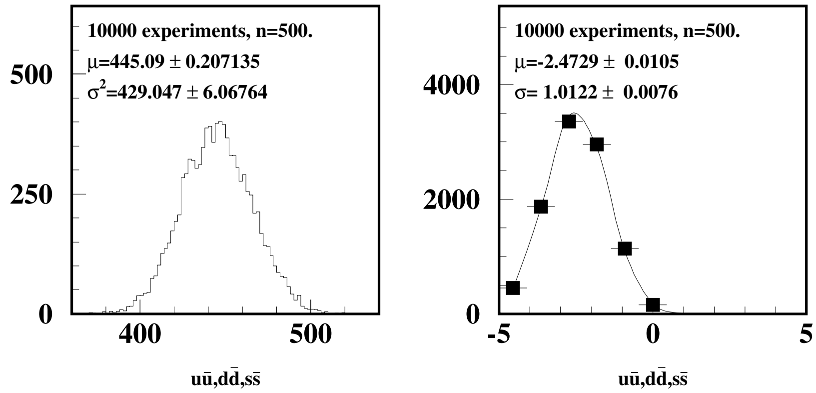


Figure A.21: $N_{u\bar{u},d\bar{d},s\bar{s}} = 5000$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 500$ (allowing shape fluctuations)

We can understand why this is so by looking at equation (A.5). Because we have only one non-zero shape, and the other shapes are fixed to zero contributions, the quantity R_{mjk} is identically zero, which means that only the terms proportional to P_{mjk} and P^2_{mjk} are present, so $P_{mjk} = 0$ is a fixed point. When $N_{mjk} = 0$ and we begin with $P_{mjk} = 0$, there is no other value of P_{mjk} which will be considered. Furthermore, any time DSNLEQ tries zero as a value, the fit will converge. So, not only will we not gain back the area lost by the zero bins in the fitting shape, but we will also lose additional area by the fit hitting the bad fixed point.

There is, however, a way around this problem. If we consider the derivative of $-\ln\mathcal{L}$ given by equation (A.4), we see that is inconsistent that we have either $P_{mjk} = 0$ or $\sum_i c_i P_{ijk} = 0$. Since we have shapes which really do have finite extent, we are somewhat lax about the former requirement. The latter, on the other hand, is a real problem. We are perfectly justified in pushing the fit away from this condition. The way we do this is by checking R_{mjk}. If it is zero (or very small), then the third term of (A.5) is zero and we can force P_{mjk} non-zero by dividing it out of the equation, leaving us with

$$0 = c_m(c_m+1)P_{mjk} + ((c_m+1)R_{mjk} - c_m d_{jk} - c_m N_{mjk}) \qquad \text{(A.14)}$$

This should also work in cases where there are multiple fitting shapes which have extent in a bin but happen to be zero. In this case, the first step will push both away from zero, and they can then converge back towards zero if that is their favored value.

We can now see how modifying the equations for the parent distributions affects the fits in Figures A.22 and A.23.

Clearly, these are much better. There are still appreciable biases, and the pull widths are not quite right, but given that these tests are pushing the statistics farther than they are likely ever to have to go, we can consider these tests successful. Now we need to verify that we have not impacted the higher-statistics fits. We can test this by once again running with eight million fitting shape events, the results of which are

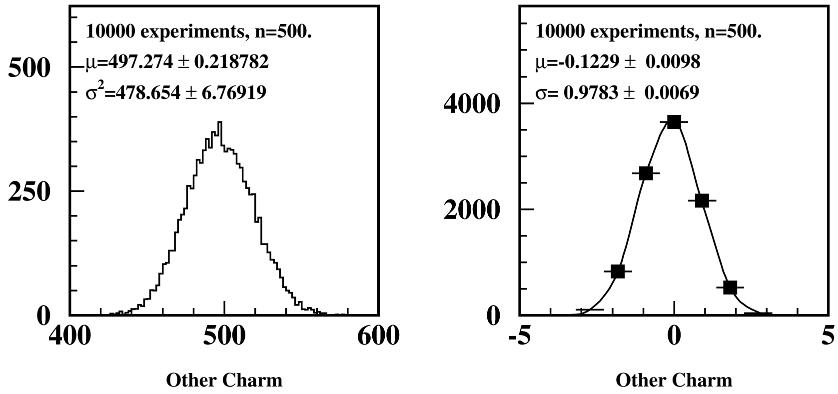


Figure A.22: $N_{charm} = 5000$; $n_{charm} = 500$ (modified shape fluctuations)

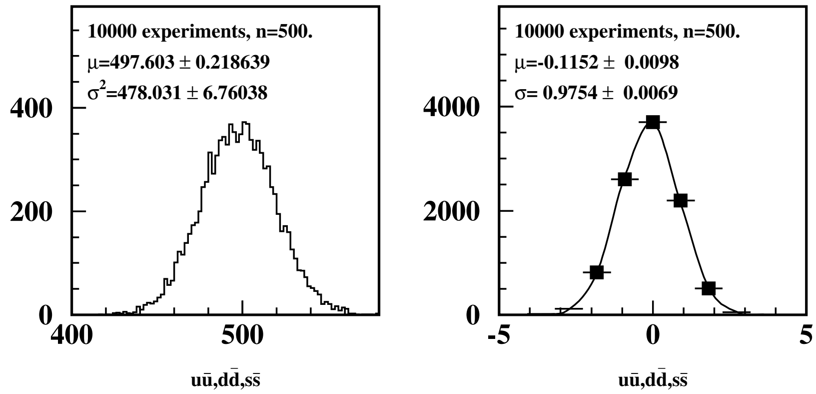


Figure A.23: $N_{u\bar{u},d\bar{d},s\bar{s}} = 5000$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 500$ (modified shape fluctuations)

shown in Figures A.24 and A.25.

These shapes look good, though the log files indicate that for almost every toy experiment, most of the bins were unable to converge in DSNLEQ. When this happens, the un-fluctuated values are used. It is possible that this is an issue of the relative scalings of the variables. To test this (since we already know that, DSNLEQ problems aside, eight million shape events works), we can run with 80,000 events in the shapes, which are Figures A.26 and A.27.

Once again, everything looks just fine, and the logfiles confirm that DSNLEQ converged on every toy experiment. This means the implementation of the extended likelihood (A.2) and the solution of the equations (A.5) is working. All plots from this point on will include the extra terms, and the reduced equations (A.14) when $R_{mjk} < 10^{-10}$, unless otherwise noted.

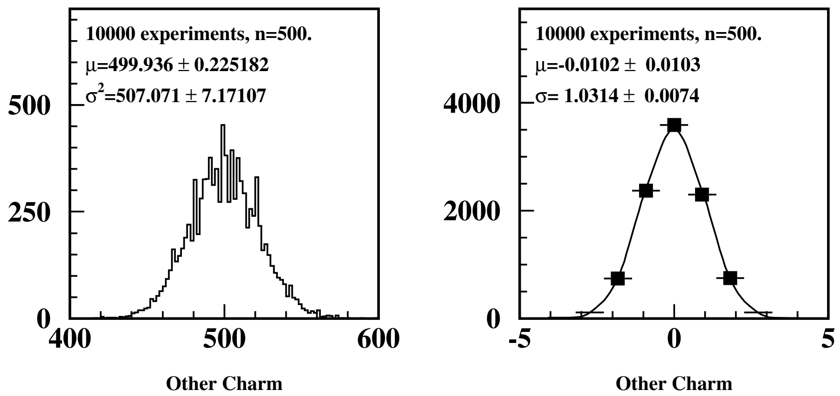


Figure A.24: $N_{charm} = 8,000,000$; $n_{charm} = 500$ (modified shape fluctuations)

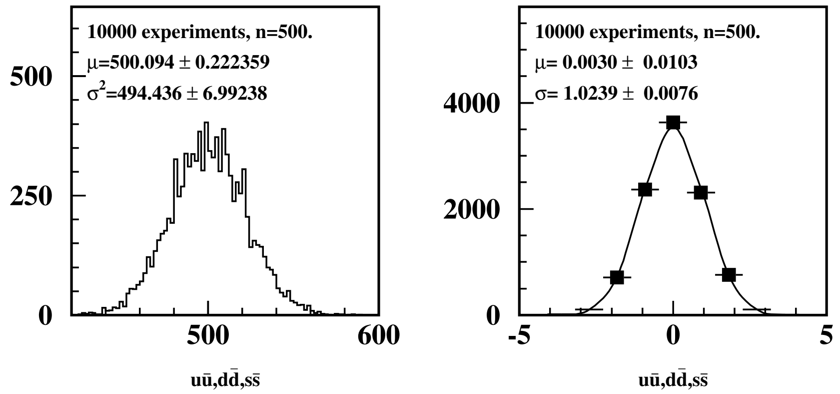


Figure A.25: $N_{u\bar{u},d\bar{d},s\bar{s}} = 8,000,000$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 500$ (modified shape fluctuations)

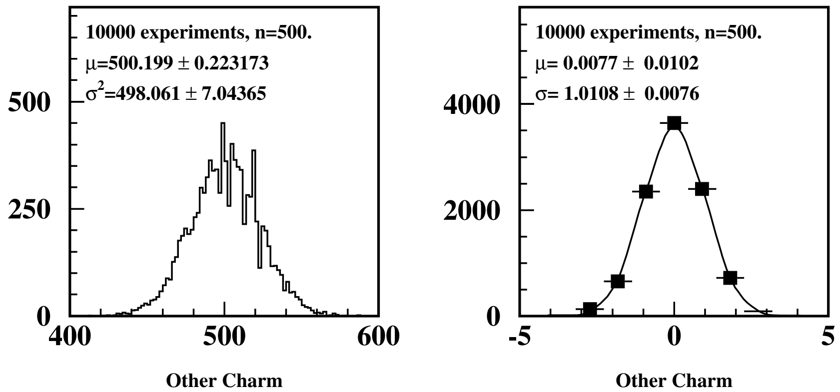


Figure A.26: $N_{charm} = 80,000$; $n_{charm} = 500$ (modified shape fluctuations)

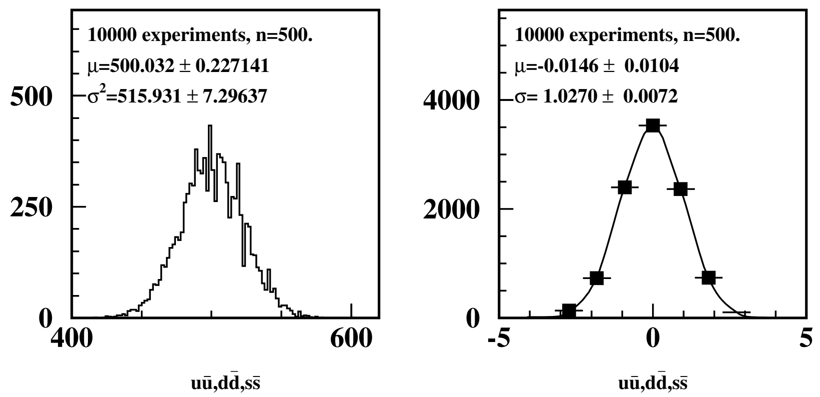


Figure A.27: $N_{u\bar{u},d\bar{d},s\bar{s}} = 80,000$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 500$ (modified shape fluctuations)

A.5 Two-Shape Fits

Now that we have shown the single-shape fits to work, we must start looking at combinations of shapes. To begin with, we will test the signal shape with each of the backgrounds, in order to verify that the fitter can at least distinguish these.

Initial tests showed serious convergence problems. The fitter either was not converging, or it converged on negative values. This also affected the error calculations. Several modifications were made to attempt to fix this.

When the total fit contribution to a bin μ_{jk} becomes small or negative, the likelihood for that bin is set to one, removing it from that iteration of the fit. When μ_{jk} is negative for one bin, however, it will be negative for most or all of the bins. This means the likelihood function has no discriminating power, and MINUIT will move the fit parameters around more-or-less at random. To prevent this, for negative μ_{jk} we set $-\ln \mathcal{L}_{jk} = \mu_{jk}^2$ in order to push the parameters back towards zero.

Another change made is a preliminary scan of all free fit parameters to find starting points that are perhaps closer to the minima than the values fed to the fitter. MINUIT has a routine to do this within specified limits. For the signal parameter, we scan from 0 to 0.01. For the background parameters, we scan from 1 to $0.2N_i$. After scanning, a HESSE error estimation is performed, which calculates the approximate second derivatives of the fitting function, giving MINUIT a better "feel" for the function.

The final change made was to release the parameter limits on the signal after the first minimization call. The limits were added to keep the signal from going negative, but limits can cause problems in error calculations. This was not seen in the single-shape fits, but with multiple shapes it is more likely for the signal parameter to be close to the lower limit in a particular fit.

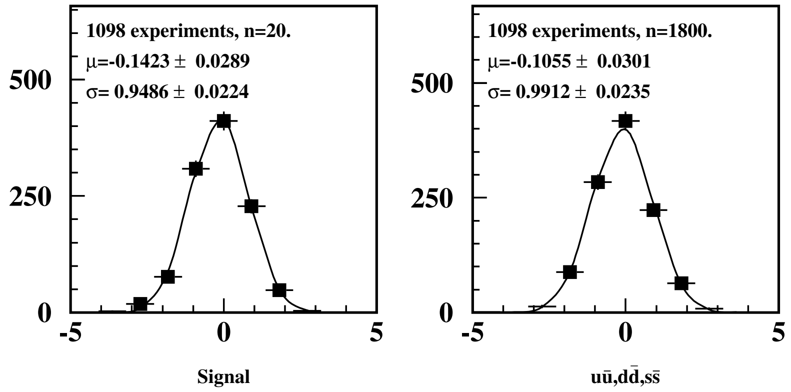


Figure A.28: $n_{signal} = 20$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 1800$ (initial two-shape modifications)

A.5.1 Signal plus $u\bar{u}, d\bar{d}, s\bar{s}$

Given the length of time each fit takes with these changes, as well as from having multiple shapes to deal with, tests were initially performed only on the signal plus $u\bar{u}, d\bar{d}, s\bar{s}$ sample, which showed the most problems in the first set of two-shape tests. The results for 2000 toy experiments are shown in Figure A.28.

These pulls do not look too bad, except that the number of experiments successfully fit is only 1098, or about 55% of the trials. The fact that both are biased low is a reflection of the generated distributions. The mean number of events after fluctuating the "perfect" data distribution is only 1815, rather than the expected central value of 1820 events. The expected error on this is roughly 1.3 events, so this is a significant systematic error. The poisson fluctuations are believed to be correct (see section A.3), and the area of the parent distribution has been verified to be 1820 events.

Further modifications to the fitting procedure were then made. Rather than relying solely on the MIGRAD algorithm, we allow MINUIT to use the slower and more robust SIMPLEX algorithm. This, however, does not produce reasonable error estimates. Given that we calculate asymmetric errors in another step, however, this should not be a great problem.

This does not improve things greatly, though it does a bit, as can be seen by looking at a scan of the likelihood function over the $u\bar{u}, d\bar{d}, s\bar{s}$ parameter, shown in Figure A.29, and the signal parameter, shown in Figure A.30. The $u\bar{u}, d\bar{d}, s\bar{s}$ scan looks fine, but the signal scan shows a lot of jitter in the function. There is basically no way that a function with this behavior can be successfully minimized.

Clearly, something else is needed. Since the likelihood function's jitter is most likely related to problems solving (A.5), we can change DSNLEQ's convergence criteria to something more restrictive. This does not completely fix the problem, but it greatly reduces the number of intractable problems. It also gives us a method of more gracefully dealing with the data: We can increase the demanded precision

```
 **********
 **  370 **SCAN    5.000       1.000       3000.
 **********
0SCAN OF PARAMETER NO.  5,  uds
   10800.00     ....................................
                  . *                 .            .
   10400.00     ...                   .            .
                  .                   .            .
   10000.00     ...                   .            .
                  .                   .            .
   9600.000     ... *                 .            .
                  .                   .            .
   9200.000     ...                   .            .
                  .                   .            .
   8800.000     ...                   .            .
                  .                   .            .
   8400.000     ...  *                .            .
                  .                   .            .
   8000.000     ...                   .            .
                  .   *               .            .
   7600.000     ...                   .            .
                  .    *              .            .
   7200.000     ...    *              .            .
                  .     *             .            .
   6800.000     ...      *            .            .
                  .       *           .            .
   6400.000     ...       **          .            .
                  .         **        .          **.
   6000.000     ...           *****   .  ********  .
                  .................***&***..........
   5600.000     ....................................
                  /         /         /         /
            -100.0     900.0     1900.     2900.
            ONE COLUMN= 100.0000       Overprint character is &
```

Figure A.29: Scan of $u\bar{u}, d\bar{d}, s\bar{s}$ Parameter

```
 **********
 **  371 **SCAN    2.000      0.0000E+00  0.2000E-01
 **********
0SCAN OF PARAMETER NO.  2,  WSrate
   6006.000     .............................................
                 .               .                          *.
   6002.000     ...              .                         * .
                 .               .                        *  .
   5998.000     ...              .                       *   .
                 .               .                      *    .
   5994.000     ...              .                     *     .
                 .               .                    *      .
   5990.000     ...              .             *    *        .
                 .               .                 *         .
   5986.000     ...              .                *          .
                 .               .               *           .
   5982.000     ... *       *    .        ** ** *            .
                 .             *  .       *  *               .
   5978.000     ...  *           .                           .
                 .               . * **                      .
   5974.000     ...   *          .  *                        .
                 .               .                           .
   5970.000     ...    *         .     **                    .
                 .      **       .                           .
   5966.000     ...       **    **.                          .
                 .           **   .*                         .
   5962.000     .................&...........................
   5960.000     .............................................
                  /         /         /         /         /
            -.5000E-030.4500E-020.9500E-020.1450E-010.1950E-01
            ONE COLUMN=0.5000000E-03   Overprint character is &
```

Figure A.30: Scan of Signal Parameter

until we have a well-behaved function. In practice, however, it is better to have a precision already set which we believe to have a high degree of success. The most common problem now seen is illustrated in Figure A.31, where MINUIT selects as its minimum a discontinuous point that is clearly not what we are looking for. This is easy to deal with interactively, since we can set the parameter appropriately and re-minimize. The reason for the discontinuity is not immediately obvious—it could be related to some bins being removed from the fit after DSNLEQ fluctuations. In any event, it is straight-forward to deal with.

If we now run 1000 toy experiments, simply ignoring the ones that do not trivially converge, we get the pull distributions shown in Figure A.32.

The yield for successful minimization is still very low. 54% of the toy experiments failed to converge properly (that is, with good asymmetric errors). Of these, over 95% converged in MIGRAD, but failed in the MINOS error calculation. This indicates that the discontinuity shown in Figure A.31 is very prevalent. There is another problem when negative fit results are obtained. Again, these are successful minimizations from MIGRAD's standpoint, but obvious garbage to us.

The problem boils down to one of discontinuities caused by the fact that logarithms can only be taken for positive numbers. This means that bins occasionally have to be thrown out, which can (and does) destroy the continuity of the likelihood function. We can change the likelihood function once again to use a special function $\mathcal{F}(x;a)$ which is defined as

$$\mathcal{F}(x;a) = \begin{cases} \ln(x) & ,(x \geq a) \\ \ln(a) + \frac{1}{a}(x-a) - \frac{1}{2a^2}(x-a)^2 & ,(x < a) \end{cases} \tag{A.15}$$

This is well-defined for all x, and is twice-continuously-differentiable at $x = a$, which is important for error calculations, which use second derivatives. We can see what this looks like for $a = 0.1$ and $a = 0.05$ in Figure A.33.

The likelihood function (A.2) now becomes

$$\mathcal{L} = \prod_{j=1}^{N_{mass}} \prod_{k=1}^{N_Q} \frac{\exp(d_{jk}\mathcal{F}(\mu_{jk};a) - \mu_{jk})}{d_{jk}!} \prod_{i=1}^{N_{shapes}} \frac{\exp(N_{ijk}\mathcal{F}(P_{ijk};a) - P_{ijk})}{N_{ijk}!} \tag{A.16}$$

With these modifications, and a set to 0.05, we can again run 1000 toy experiments. The fraction of "good" experiments is much higher, about 93%, with only 1.5% not "easily" recoverable. The pulls are shown in Figure A.34. The pull for the $u\bar{u}, d\bar{d}, s\bar{s}$ parameter is excellent. The pull for the signal parameter is not nearly as good, but it is reasonable.

A.5.2 Signal plus Ridge

We now look at the signal and ridge fits. Once again we use 20 signal events in the parent distribution. For the ridge, we use 100 events. The pulls for 1000 toy

```
 **********
 **  725 **SCAN    2.000      0.0000E+00  0.1000E-01
 **********
0SCAN OF PARAMETER NO.  2,  WSrate
5958.500 ......................................................
           .                                  .               .
5957.500 ...                                  .               .
           .                                  .               .
5956.500 ...   *                              .               .
           .                                  .               .
5955.500 ...                                  .               .
           .    *                             .               .
5954.500 ...                                  .              *.
           .                                  .             * .
5953.500 ...      *                           .            *  .
           .                                  .               .
5952.500 ...                                  .               .
           .       *                          .               .
5951.500 ...                                  .               .
           .        *                         .               .
5950.500 ...                                  .          *    .
           .         *                        .         *     .
5949.500 ...           *                      .        *      .
           .            *                     .       *       .
5948.500 ...                                **.     *         .
           .             **                *  .    *          .
5947.500 ...                *           **    .   *           .
           .                 *** **** **      . *             .
5946.500 .....................................&................
5946.000 ......................................................
           /         /         /         /         /
     -.2000E-030.1800E-020.3800E-020.5800E-020.7800E-02
     ONE COLUMN=0.2000000E-03   Overprint character is &
```

Figure A.31: Discontinuity in Scan of Signal Parameter (trimmed for display)

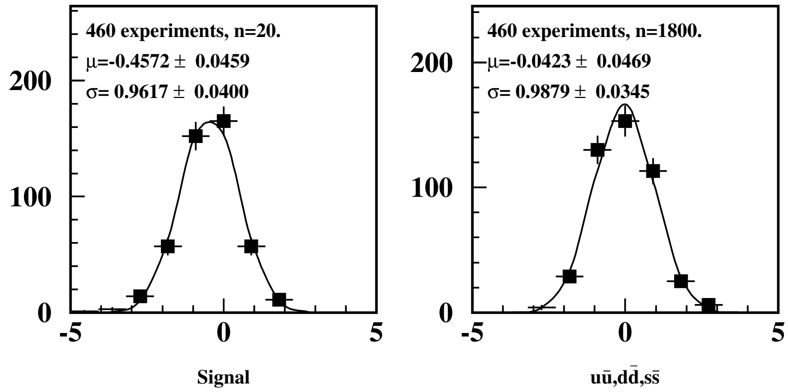
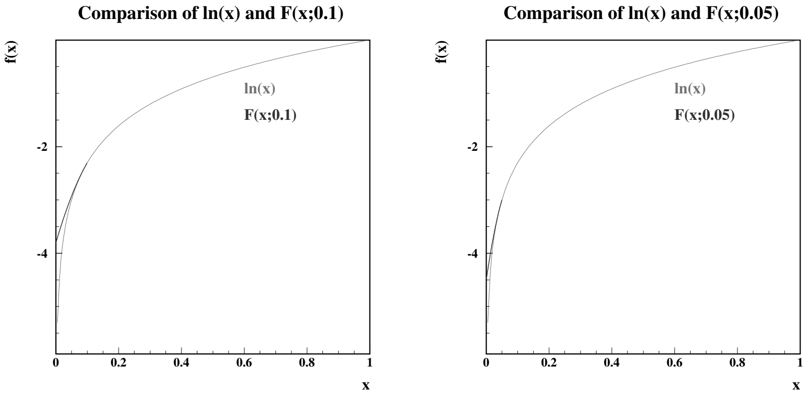


Figure A.32: $n_{signal} = 20$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 1800$ (increased DSNLEQ precision)



(a) $a = 0.1$

(b) $a = 0.05$

Figure A.33: Two examples of $\mathcal{F}(x; a)$

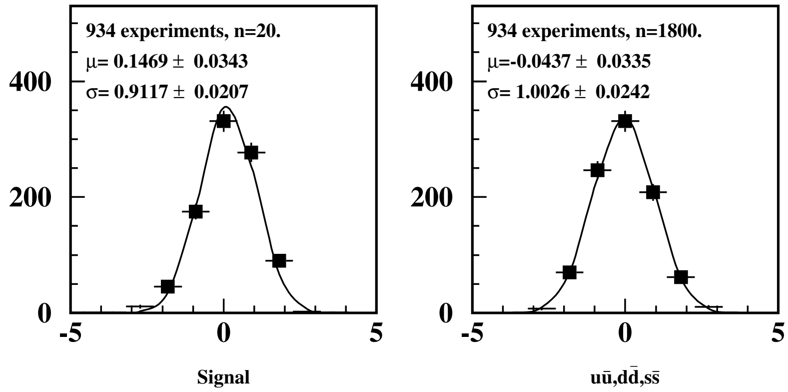


Figure A.34: $n_{signal} = 20$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 1800$ ($\ln(x) \rightarrow \mathcal{F}(x; 0.05)$)

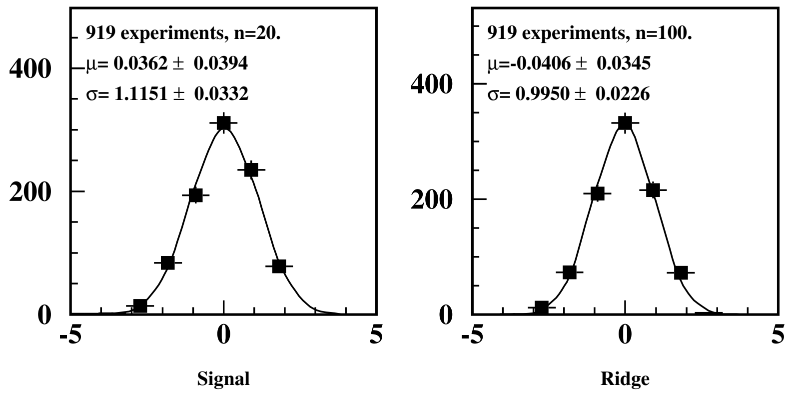


Figure A.35: $n_{signal} = 20$; $n_{ridge} = 100$ ($\ln(x) \rightarrow \mathcal{F}(x; 0.05)$)

experiments can be seen in Figure A.35. The pulls are unbiased, and as with the $u\bar{u}, d\bar{d}, s\bar{s}$ fits the background shape has the correct width while the signal shape does not. In this case, the errors are under-estimated by about 10%. For comparison with signal plus $u\bar{u}, d\bar{d}, s\bar{s}$, 92% of the toy experiments are good, and 1.8% are not easily recoverable.

A.5.3 Signal plus Other Charm

Moving on to signal plus other charm, the situation is rather different. While the pulls (Figure A.36) are reasonable by the standards of the other fits, only 37% of the toy experiments converge and yield asymmetric errors properly. Furthermore, 7.7% do not fit into the category that is expected to be easy to fit interactively. This is, of course, worth a closer look.

First running interactively, and then increasing once again the tolerance used in DSNLEQ, it appears that in many cases the likelihood function is badly-behaved, including the cases which were previously believed to be recoverable. This may be due to the similarity of the shapes in the one region where the signal is present. The question then becomes whether the shape used for the other charm component is a fair representation of what we expect to see. It is possible, for example, that the piece with structure is over-emphasized compared with the phase-space shape. It is also possible that the shape is just generally unrealistic. In either case, we know that the other charm and $u\bar{u}, d\bar{d}, s\bar{s}$ shapes are similar over much of the fitting plane, which means we may be fighting a similar effect when both are present.

A.6 Full Fits

At this point we know that fits which succeed are behaving correctly. It now becomes interesting to consider the full four-shape fits, both to look at pull distributions and

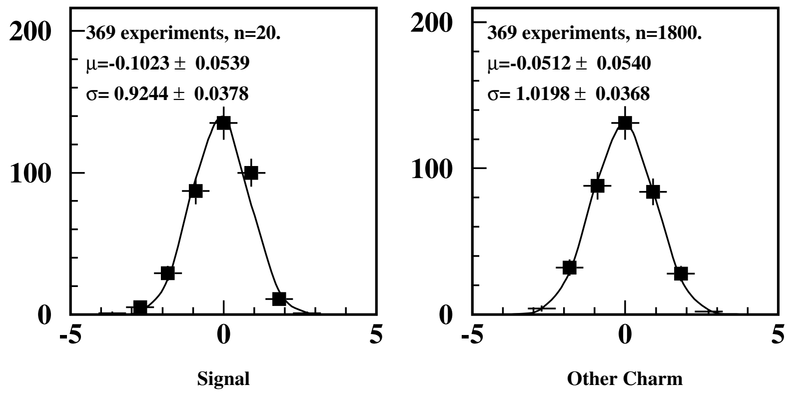


Figure A.36: $n_{signal} = 20$; $n_{charm} = 1800$ ($\ln(x) \to \mathcal{F}(x; 0.05)$)

to see how frequently we encounter problems. As before, the signal shape-to-data statistics ratio is 300:1, and for the backgrounds the ratio is 10:1.

To begin with, we will use the statistics of the two-shape fits. That is, 20 signal events, 100 ridge events, and 1800 events each of other charm and $u\bar{u}, d\bar{d}, s\bar{s}$. The pulls from 1000 toy experiments are shown in Figure A.37.

Only 66% of the fits succeeded, with 21% of the failures converging either with negative fit results or no asymmetric errors. Of the four components, only the ridge appears to have a reasonable pull distribution. The signal is not terrible, certainly, though it is biased more than we would expect for the mean of the total events (3719.5, compared with a nominal 3720). The two broad shapes are also badly biased, and their expanded widths are likely due to anti-correlation between them. This is understandable to some extent, since they both have broad, relatively flat regions which make it difficult to distinguish between them. The structure of the other charm shape does not appear to help matters much, probably because its structure can be accomodated by fluctuations from a smaller contribution, as well as fluctuations in the $u\bar{u}, d\bar{d}, s\bar{s}$ shape. Additionally, we have already seen in Section A.5.3 that the fitter can easily become confused between the signal shape and the other charm shape in the signal region.

What this tells us is that if we actually end up with one shape that is either very close to another, or looks like a composite of two or more of the other shapes, we will not be able to fit reliably. We can switch to a new shape for the "other charm" contribution, which resembles somewhat what is seen in $D^0 \to K\pi\pi^0$, and is shown in Figure A.38. Running 1000 toy experiments with this shape yields the pulls in Figure A.39.

The pulls certainly look better, and none of the contributions show significant bias. There are still problems, however. Only 35% of the fits succeed, and the broad shapes have errors underestimated by about 50%. Another 41% converge with some problems. Given the occasional irregularities in the likelihood function, it is

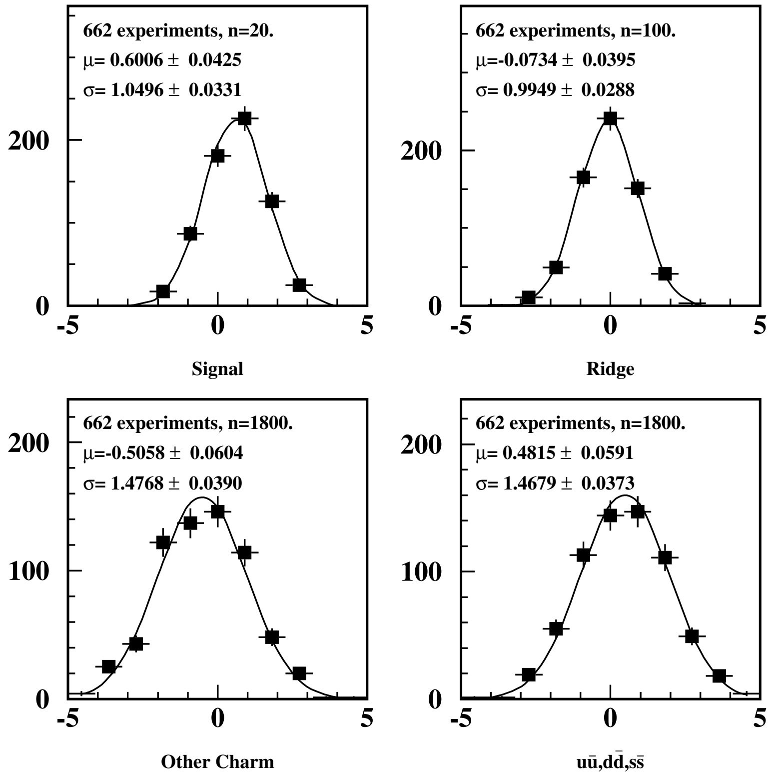


Figure A.37: $n_{signal} = 20$; $n_{ridge} = 100$; $n_{charm} = 1800$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 1800$

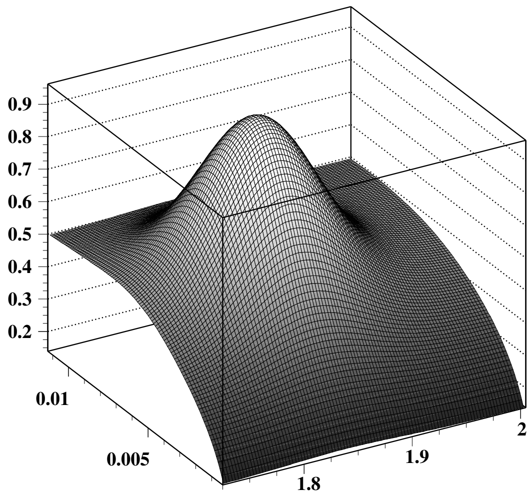


Figure A.38: New “Other Charm” Shape

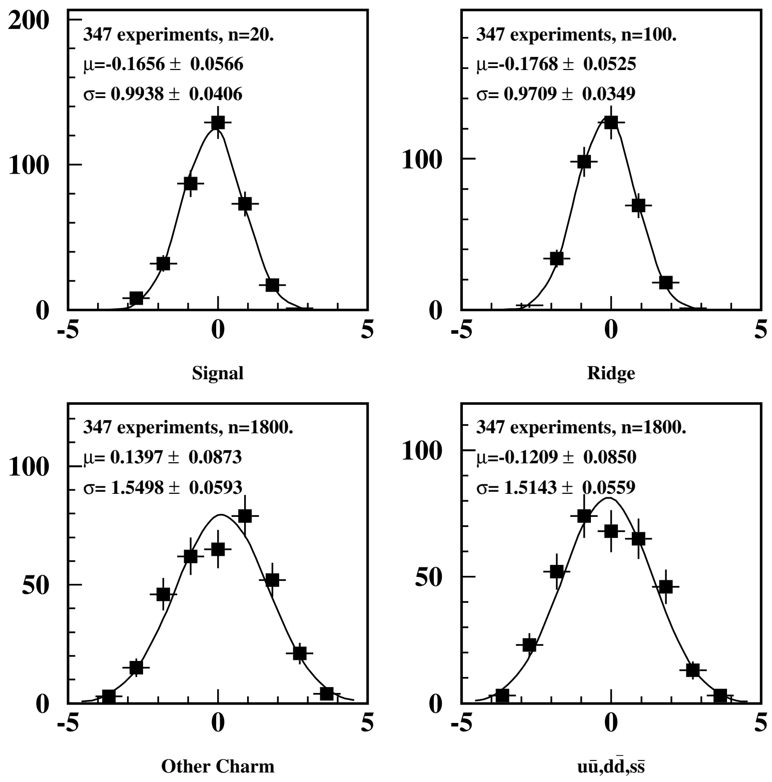


Figure A.39: New Shape

possible that many of these are error-calculation problems, where MINOS cannot travel far enough along the function before hitting a discontinuity. In these cases, it is possible that the symmetric errors are good enough. We can perform another 1000 toy experiments, the results of which are shown in Figure A.40.

This is certainly better in the signal, and the yield for good fits is up to 84% when we only consider convergence. The ridge is slightly more biased, but not really horribly so. The broad shapes are worse, however.

A.6.1 More Two-Shape Fits

Clearly, there is still a problem with the way the two broad shapes interact, so we will now look at another two-shape fit with only other charm and $u\bar{u}, d\bar{d}, s\bar{s}$ present. The results are shown in Figure A.41.

As we would expect, this shows similar problems to what we observed in the full fit. If we modify the $u\bar{u}, d\bar{d}, s\bar{s}$ shape to give it some structure, the fitter can hopefully begin to distinguish them more easily. The shape we use is very much like the new other charm shape, but with the Gaussian in a different location. The separation between the shapes is 1.1σ in mass and 0.2σ in Q. The results of this are

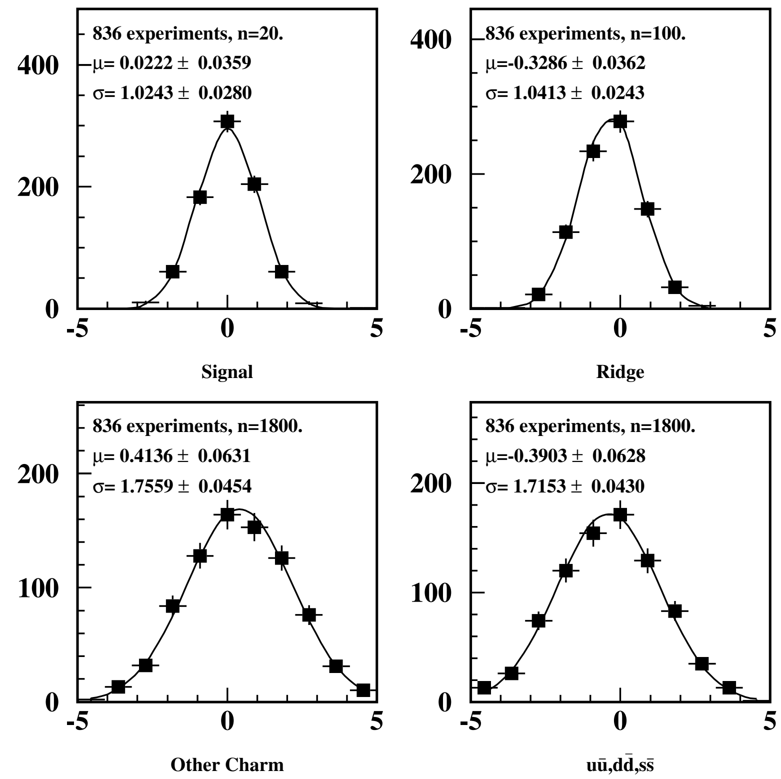


Figure A.40: New Shape, Allow Symmetric Errors

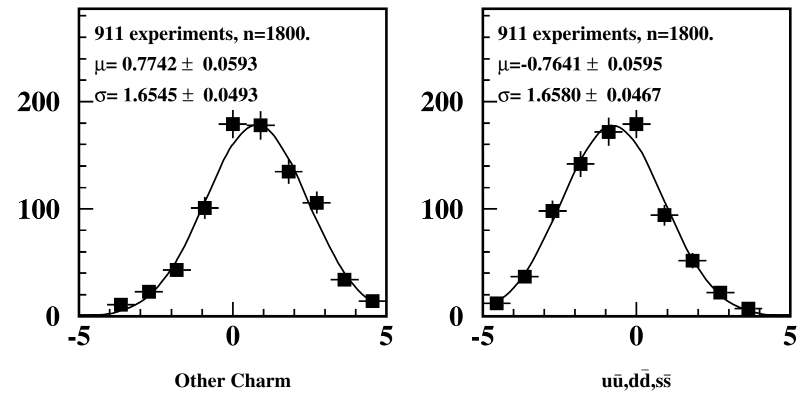


Figure A.41: $n_{charm} = 1800$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 1800$

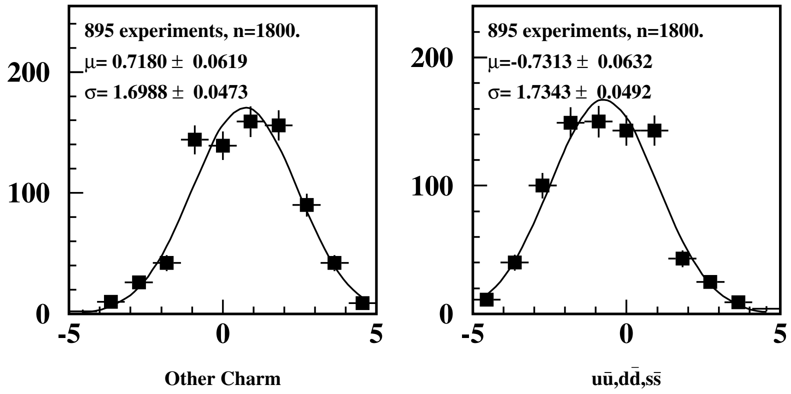


Figure A.42: $n_{charm} = 1800$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 1800$ (new $u\bar{u}, d\bar{d}, s\bar{s}$ shape)

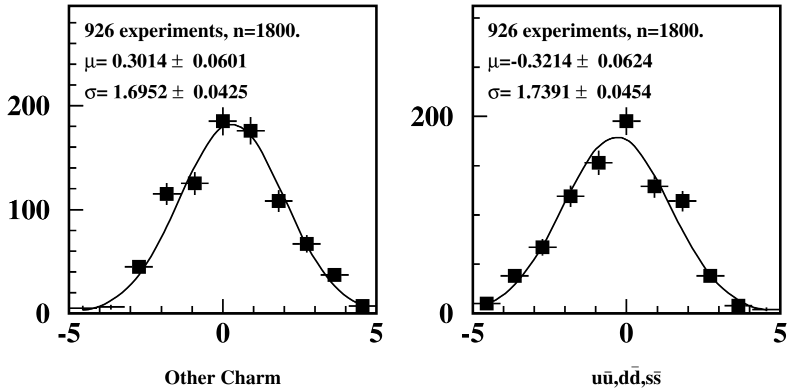


Figure A.43: $n_{charm} = 1800$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 1800$ (increased separation)

in Figure A.42.

This is really no improvement, but the Gaussians are so close that the fitter is most likely confusing them. We can increase the separation in Q to 0.8σ, which is shown in Figure A.43.

This is a little better, though not enough. We can make the Gaussians narrower in Q by a factor of 2, making a 1.6σ separation. This brings the total separation to 2σ. The results are in Figure A.44.

The widths are a little better. Three successive attempts to run 1000 experiments, however, each failed at the six or seven hundred level. What is shown is the last attempt, which is also the most successfully-run experiments before encountering a floating exception. The success rate is about 87%. It is, perhaps, worthwhile at this point to look at pulls without the additional terms for the finite shape statistics. The results of this are in Figure A.45.

This looks much better. Both pulls are unbiased and have unit width. The shape statistics are still reasonably high, though: The $D^0 \to K\pi\pi^0$ analysis sees about half

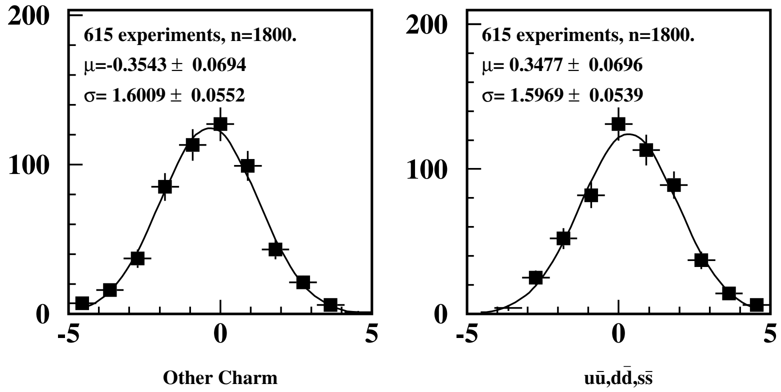


Figure A.44: $n_{charm} = 1800$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 1800$ (2σ separation, 702 experiments)

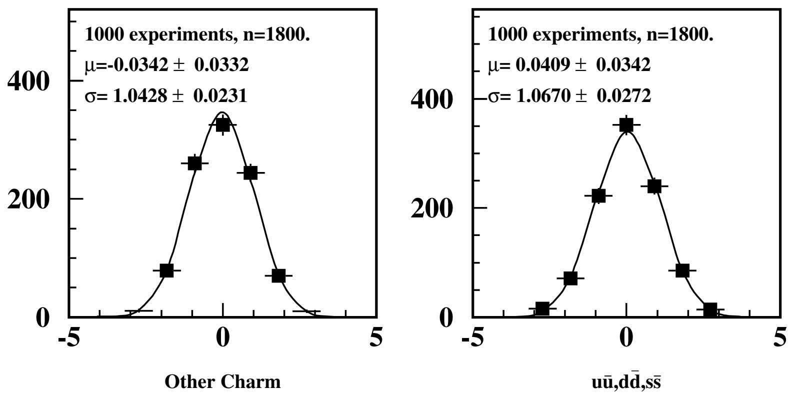


Figure A.45: $n_{charm} = 1800$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 1800$ (no finite shape statistics terms)

of this. We can easily run the half-statistics test with the basic likelihood function, and the results are shown in Figure A.46.

Again, everything looks fine. If we again halve the statistics, we get the pulls in Figure A.47.

Aside from some slight feed-across between the two shapes, leading to small biases, the pulls still look good. However, we have equal contributions from each shape, so on average each bin has the same probability of fluctuating to zero in either shape. This means that the two shapes are likely trading off zero bins, and absorbing each others' contributions to those bins. We can test this by making the contributions more asymmetric. We can leave the fitting shape statistics the same and skew the contributions to the data, which is shown in Figure A.48. We can also skew the fitting shape statistics and leave the contributions equal, which is shown in Figure A.49.

These plots are very interesting (and no, the central values are not reversed). We see that the fit results are tracking not changes in the contributions to data, but the statistics of the fitting shapes. A test using the extended likelihood had problems

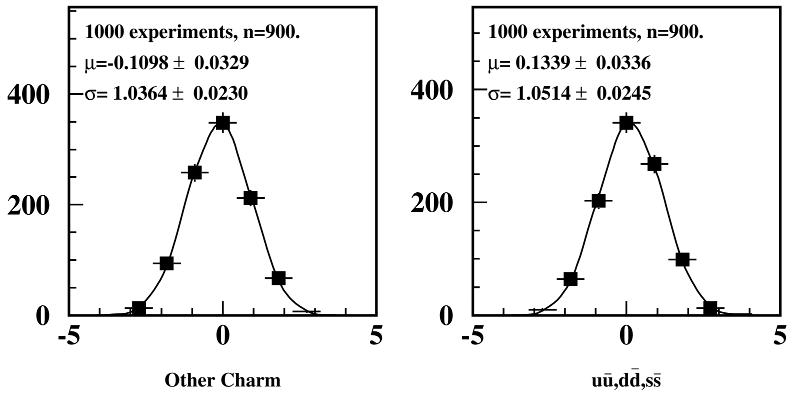


Figure A.46: $n_{charm} = 900$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 900$ (no finite shape statistics terms)

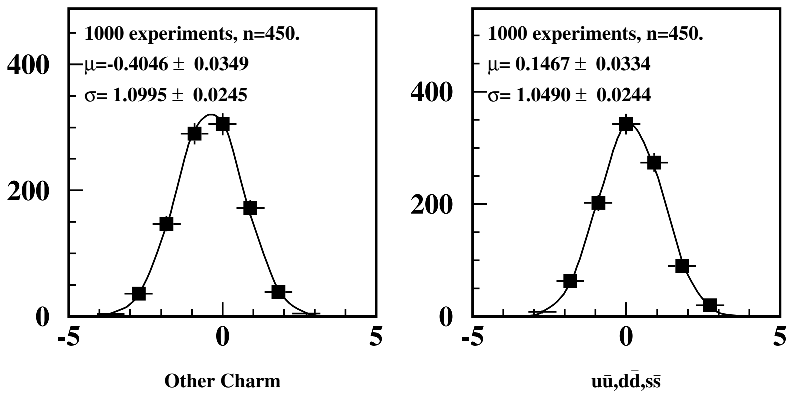


Figure A.47: $n_{charm} = 450$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 450$ (no finite shape statistics terms)

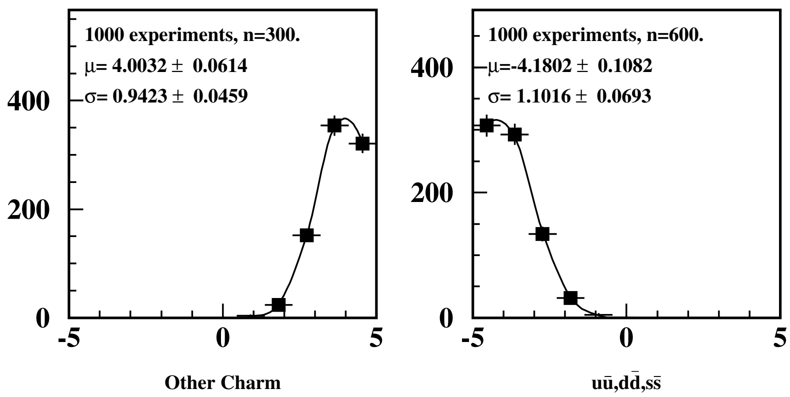


Figure A.48: $N_{charm} = 4500$; $N_{u\bar{u},d\bar{d},s\bar{s}} = 4500$; $n_{charm} = 300$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 600$ (no finite shape statistics terms)

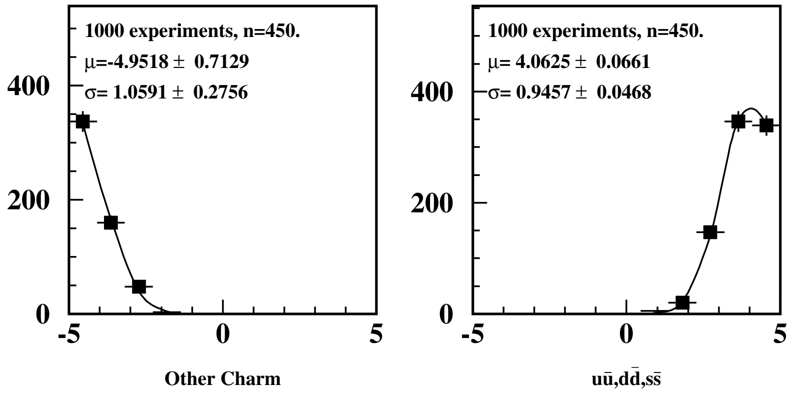


Figure A.49: $N_{charm} = 3000$; $N_{u\bar{u},d\bar{d},s\bar{s}} = 6000$; $n_{charm} = 450$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 450$ (no finite shape statistics terms)

running, including a fatal error in MINUIT.

A.7 Constrained Fits

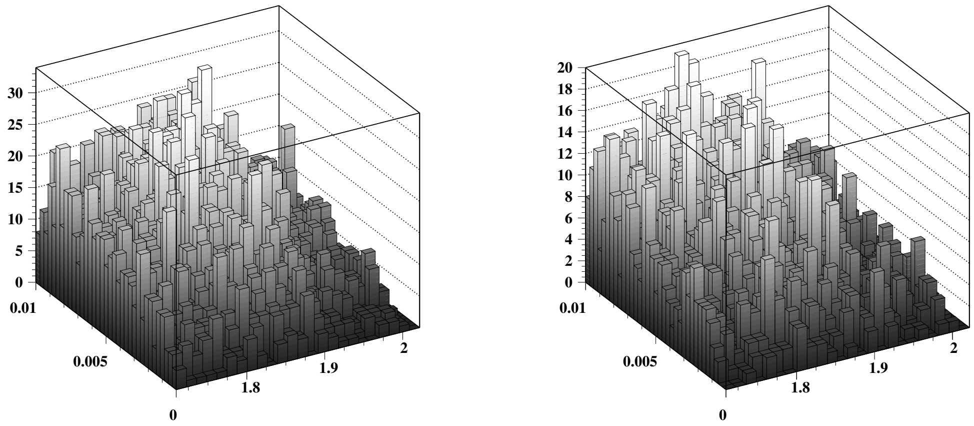
It seems fairly obvious that we are not going to successfully distinguish two similar broad shapes. We now amend our procedure as follows:

1. Re-define the broad shapes to correspond more closely to observed shapes in $D^0 \to K\pi\pi^0$ (Section A.7.1)
2. Verify that three-shape fits (only one broad shape present) work properly.
3. Change the $u\bar{u}, d\bar{d}, s\bar{s}$ fit parameter to the ratio $n_{u\bar{u},d\bar{d},s\bar{s}} : n_{charm}$
4. Fix the ratio parameter to correct and incorrect values, and observe the effect on the signal and ridge pull distributions.

A.7.1 Fitting Shapes from $D^0 \to K\pi\pi^0$

The $K\pi\pi^0$ analysis[26] has the Monte Carlo shapes shown in Figure A.50. These are clearly subject to fluctuations from finite Monte Carlo, but we can apply a smoothing algorithm to them in order to more closely approximate the parent distributions (Figure A.51).

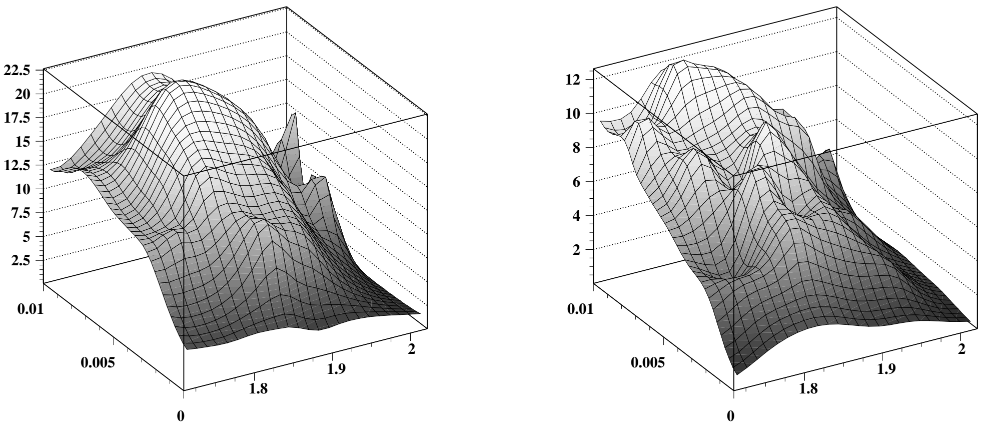
These appear to follow the phase-space turn-on shape in Q and a polynomial in mass, with a broad two-dimensional gaussian added in. The analytical approximations made based on this are shown in Figure A.52, and our subsequent tests will use these shapes.



(a) Other Charm

(b) $u\bar{u}, d\bar{d}, s\bar{s}$

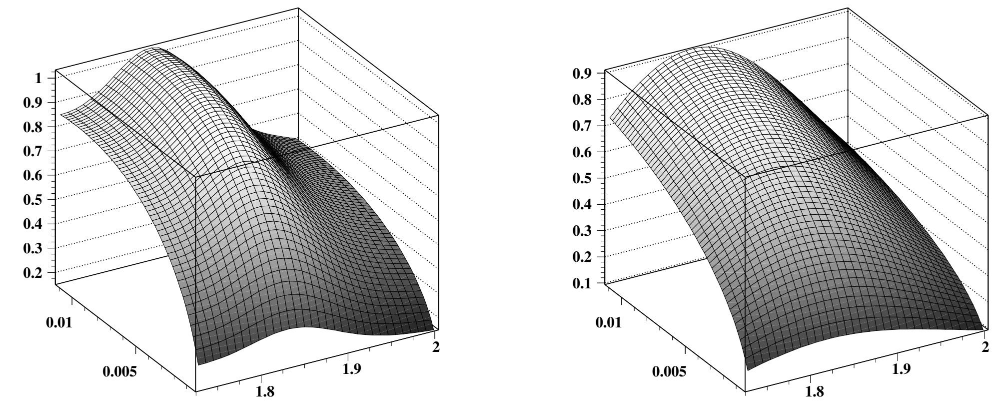
Figure A.50: Monte Carlo histograms from the $D^0 \to K\pi\pi^0$ analysis



(a) Other Charm

(b) $u\bar{u}, d\bar{d}, s\bar{s}$

Figure A.51: Smoothed versions of Figure A.50



(a) Other Charm (b) $u\bar{u}, d\bar{d}, s\bar{s}$

Figure A.52: Analytic shapes determined from shapes in Figure A.51

A.7.2 Three-Shape Fits

If we exclude the $u\bar{u}, d\bar{d}, s\bar{s}$ shape from the fits, we should be able to distinguish all of the remaining shapes. Of course, this needs to be verified before we can proceed to the constrained fits. To begin with, we perform the fits with only the basic likelihood function. These are shown in Figure A.53.

These pulls certainly are not bad, though the signal is slightly biased. Whether we need it for this case or not, the extended likelihood fit should also be tested, since we may find it necessary when performing the four-shape constrained fit. The results are shown in Figure A.54.

One last test before the constrained fits is that the fitter works when treating one broad shape as the other (the evidence for which prompted the constraint). We generate a sample with signal, ridge, and other charm, but we use the $u\bar{u}, d\bar{d}, s\bar{s}$ shape instead of other charm in the fit. The results are shown in Figure A.55. This seems ample confirmation: The pulls are not perfect, but the signal at least looks just fine.

A.7.3 Full Constrained Fits

Now that we know that we can distinguish the signal and ridge shapes from a broad background, we can re-introduce the $u\bar{u}, d\bar{d}, s\bar{s}$ shape in a fixed ratio to the other charm contribution. What we are most interested in seeing is how the signal (and to a

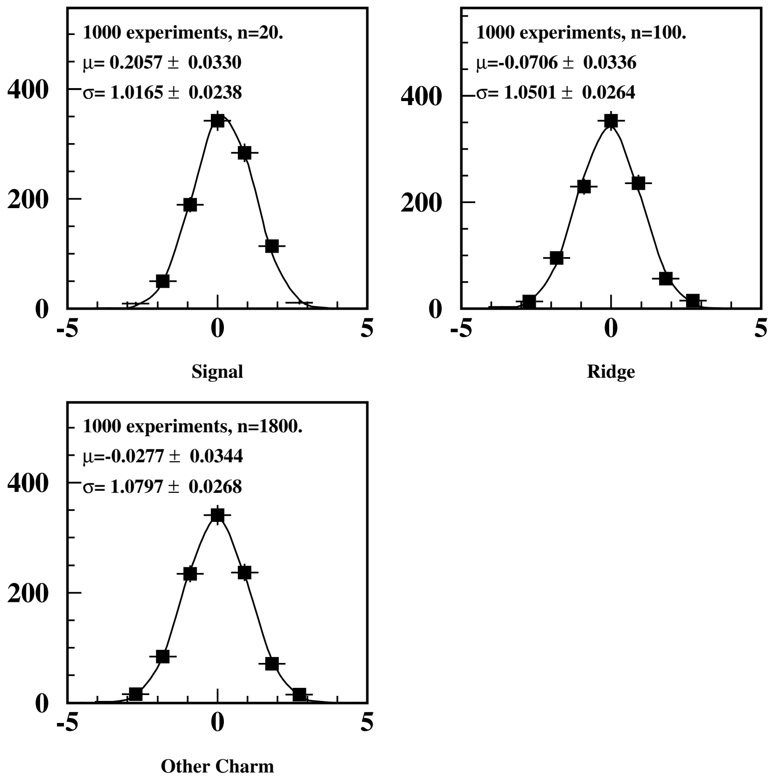


Figure A.53: $n_{signal} = 20$; $n_{ridge} = 100$; $n_{charm} = 1800$ (no finite shape statistics terms)

lesser extent the ridge) pull distributions behave as we vary the ratio from its correct value to various other values. This will give us some indication of the systematic errors we can expect in the data. For all of these tests, unless otherwise indicated the simple likelihood was used.

To begin with, we use a sample with equal amounts of both broad shapes. Figure A.56 shows the pulls with $n_{u\bar{u},d\bar{d},s\bar{s}} : n_{charm}$ fixed to the correct value of 1. Figure A.57 uses a fixed ratio of 0.5, and Figure A.58 uses a ratio of 2. The pulls look best for the correct ratio, but for the two incorrect ratios the pulls also look fine. The signal becomes slightly biased, but the bias is small compared with the statistical errors.

Now we can test a highly-asymmetric case, where the true ratio of $u\bar{u}, d\bar{d}, s\bar{s}$ to charm is $1:9$. The pulls are shown in Figures A.59-A.63.

The pulls in Figure A.63 are certainly not good, but the signal is not terribly-biased, and the ratio was fixed to a value 18 times what was generated. Summary plots of the effects on the signal for the varying ratios are shown in Figures A.64 and A.65. Of course, it is useful to do a test with the extended likelihood to see whether this improves things. If it does, then the data fit should probably use the extended

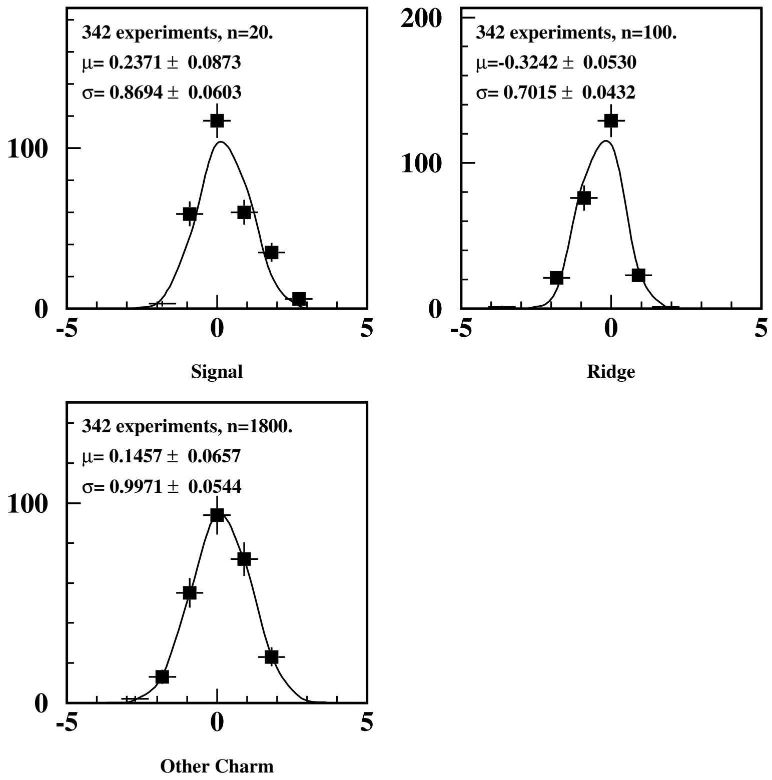


Figure A.54: $n_{signal} = 20$; $n_{ridge} = 100$; $n_{charm} = 1800$

likelihood (if possible) to reduce systematics in the signal measurement. The results of the extended-likelihood test are shown in Figure A.66 as means, because all of the fit results are far from the correct values. The incomplete set (811 of 1000 trials) is due to a hardware glitch.

A.8 Using Lifetime to Distinguish Shapes

One final thing to consider is that the two types of background leading to the broad shapes have very different average lifetimes. For the "other charm" events, the average lifetime is that of the D^0. For the $u\bar{u}, d\bar{d}, s\bar{s}$ events, the average lifetime is zero. We can use this information to create another χ^2 term (with a factor of $\frac{1}{2}$ because $\chi^2 = -2\ln\mathcal{L}$, and we use $-\ln\mathcal{L}$) of the form

$$\frac{1}{2}\chi^2 = \frac{\left(\langle\tau_{fit}\rangle - \langle\tau_{data}\rangle\right)^2}{2\sigma_{data}^2} \tag{A.17}$$

where $\langle\tau_{fit}\rangle$ is the average lifetime computed from the events in the fitting distributions and their relative weights, and $\langle\tau_{data}\rangle$ is the average lifetime of the events in the data sample (with error σ_{data}). In both averages, the signal region is excluded.

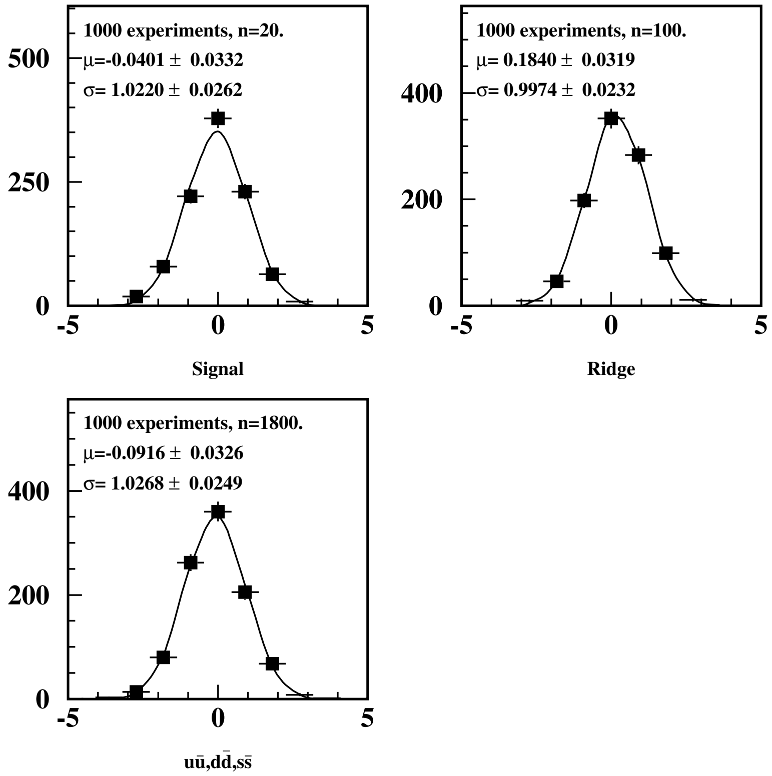


Figure A.55: $n_{signal} = 20$; $n_{ridge} = 100$; $n_{charm} = 1800$ (no finite shape statistics, other charm treated as $u\bar{u}, d\bar{d}, s\bar{s}$)

The first comparison to make is between fits with and without lifetime information using the simple likelihood function. These are shown in Figures A.67 and A.68. The pull widths for the two broad shapes have decreased to around 0.8, and the biases are *larger*, so clearly the lifetime information does not help in this case.

Still, it is possible that the extended likelihood will yield different results, necessitating the additional tests. The results are shown in Figures A.69 and A.70. Again, the biases increase, though the widths are better with the lifetime information included. This may be due to two problems cancelling one another out. Regardless, the biases are still too large.

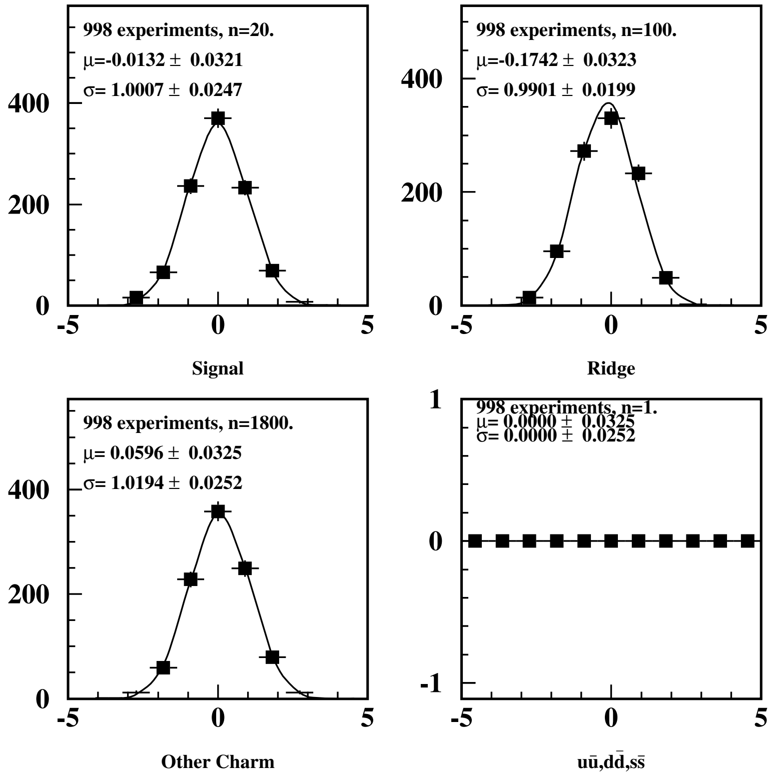


Figure A.56: $n_{signal} = 20$; $n_{ridge} = 100$; $n_{charm} = 1800$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 1800$; $u\bar{u}, d\bar{d}, s\bar{s}\ : charm \equiv 1$

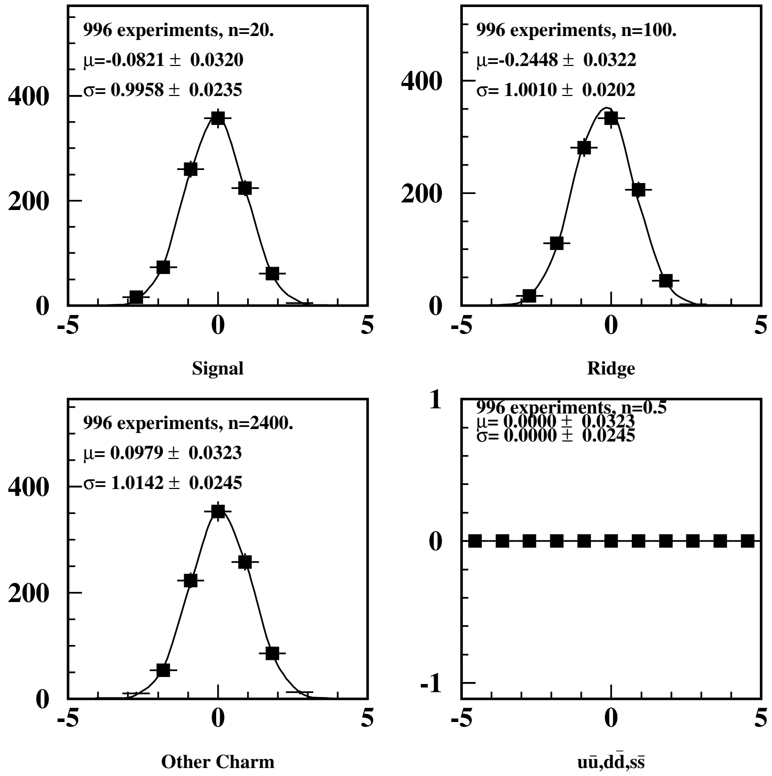


Figure A.57: $n_{signal} = 20$; $n_{ridge} = 100$; $n_{charm} = 1800$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 1800$; $u\bar{u}, d\bar{d}, s\bar{s}\ : charm \equiv 0.5$

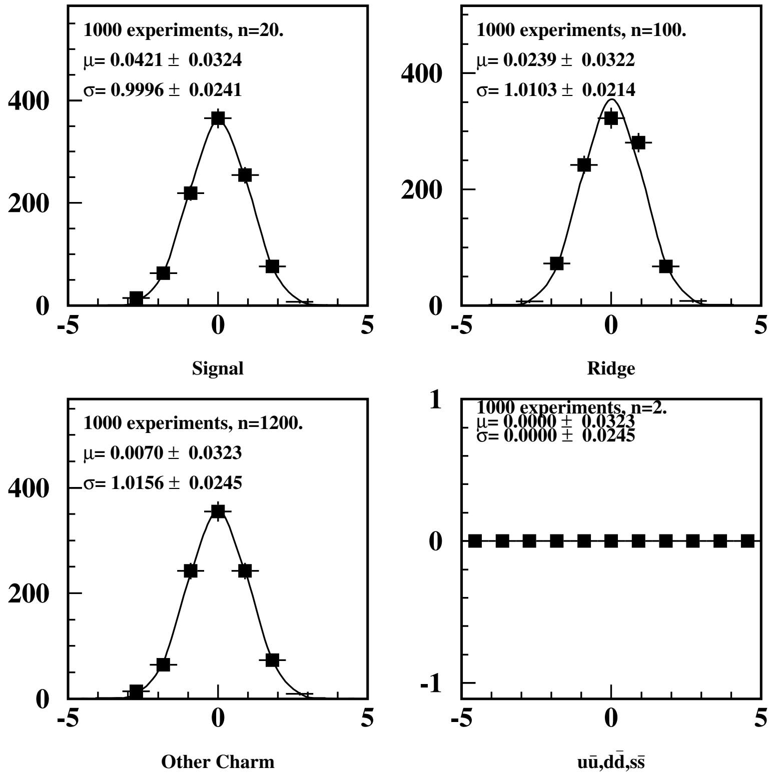


Figure A.58: $n_{signal} = 20$; $n_{ridge} = 100$; $n_{charm} = 1800$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 1800$; $u\bar{u}, d\bar{d}, s\bar{s}\ : charm \equiv 2$

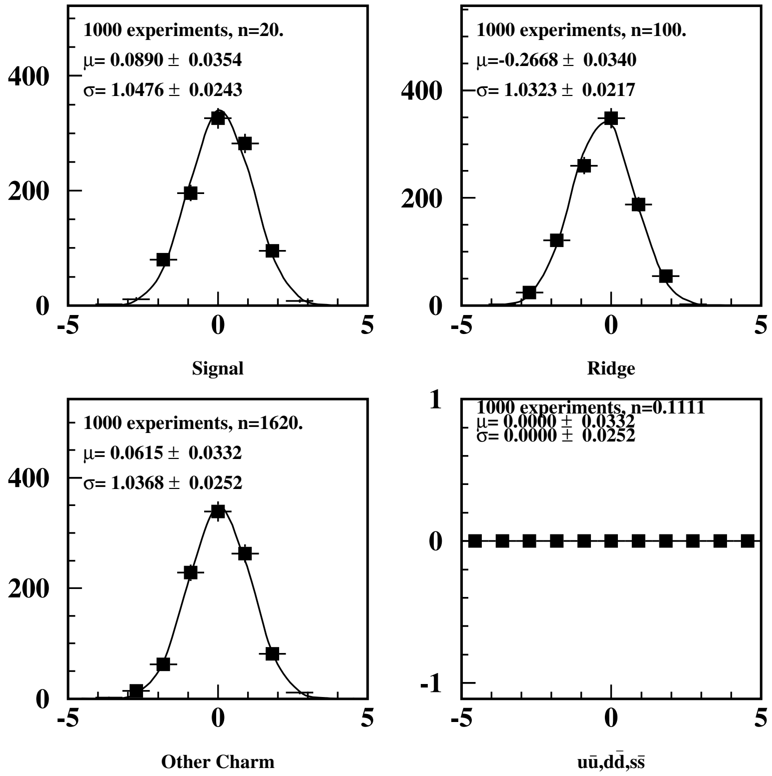


Figure A.59: $n_{signal} = 20$; $n_{ridge} = 100$; $n_{charm} = 1620$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 180$; $u\bar{u}, d\bar{d}, s\bar{s}$: $charm \equiv 0.1111$

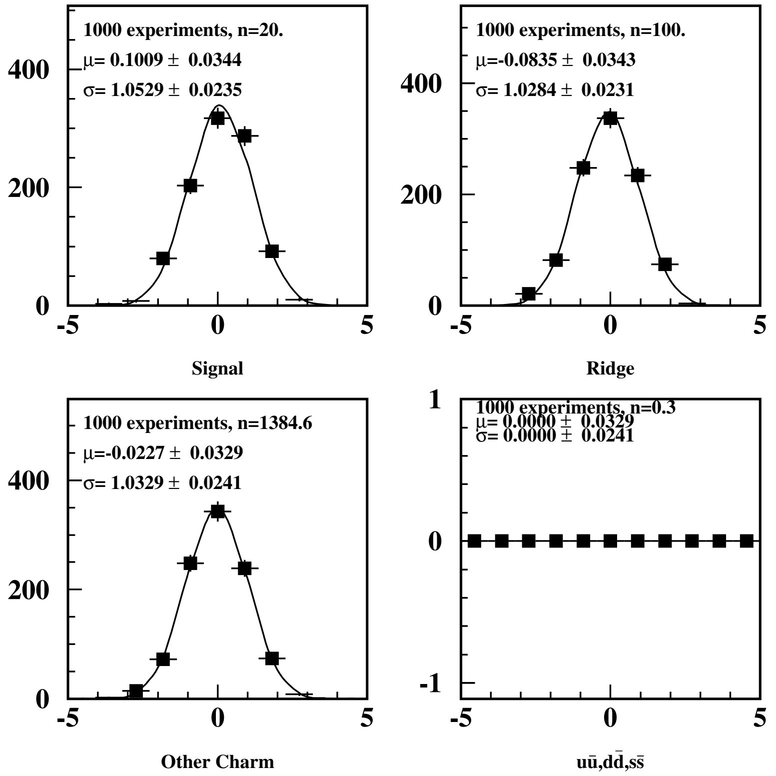


Figure A.60: $n_{signal} = 20$; $n_{ridge} = 100$; $n_{charm} = 1620$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 180$; $u\bar{u}, d\bar{d}, s\bar{s}$: $charm \equiv 0.3$

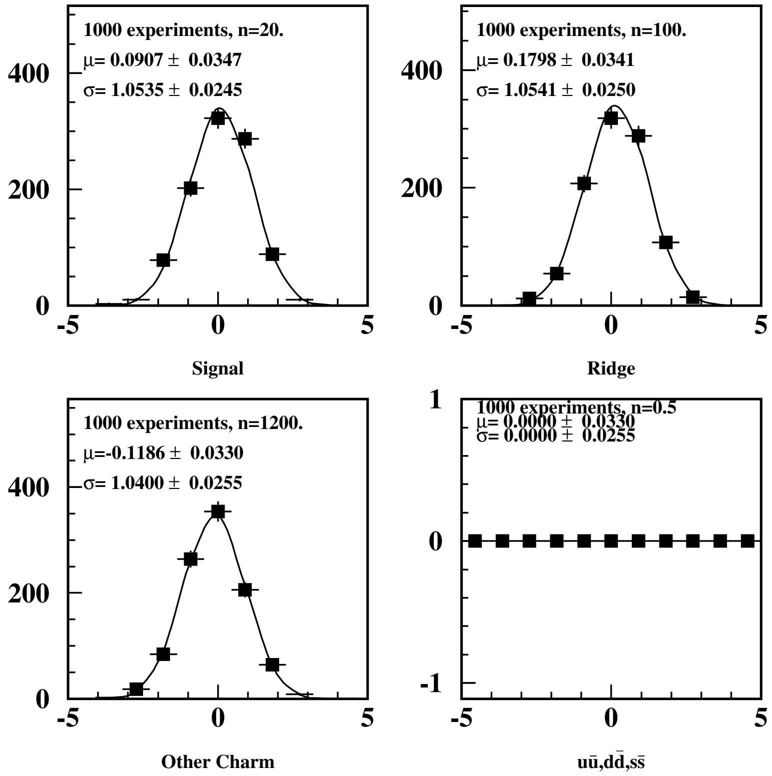


Figure A.61: $n_{signal} = 20$; $n_{ridge} = 100$; $n_{charm} = 1620$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 180$; $u\bar{u}, d\bar{d}, s\bar{s}$ $: charm \equiv 0.5$

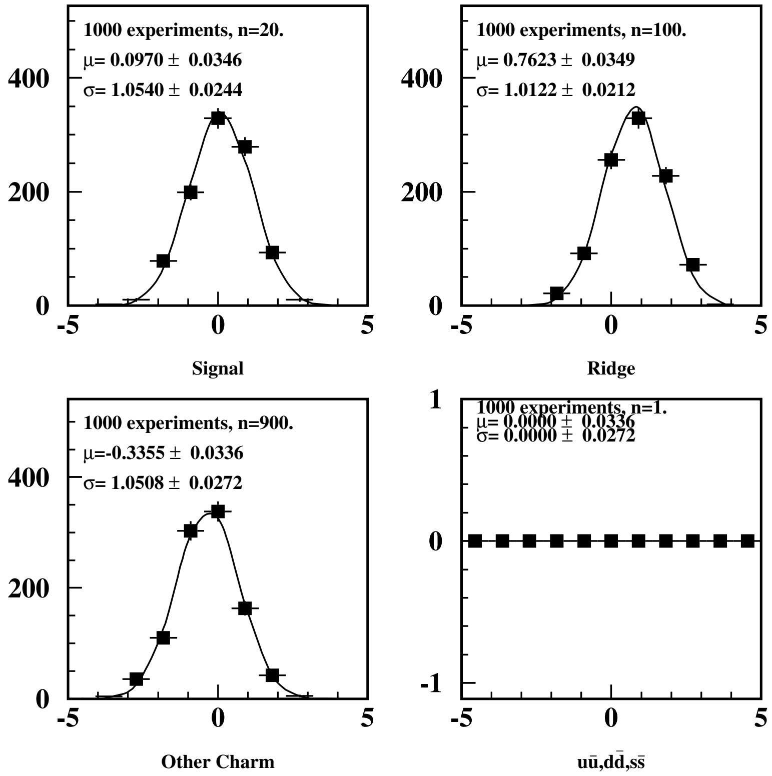


Figure A.62: $n_{signal} = 20$; $n_{ridge} = 100$; $n_{charm} = 1620$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 180$; $u\bar{u}, d\bar{d}, s\bar{s}$: $charm \equiv 1$

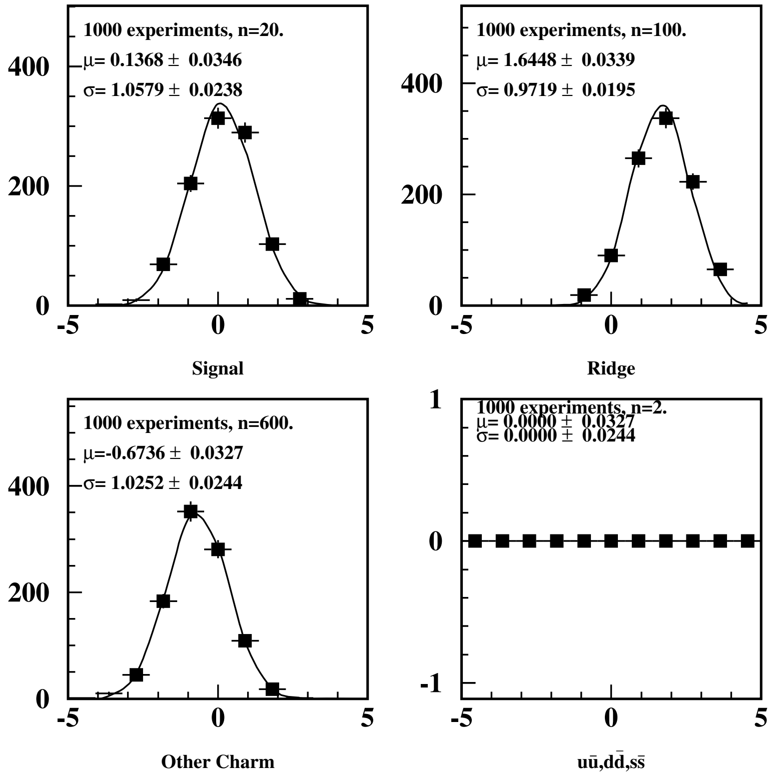
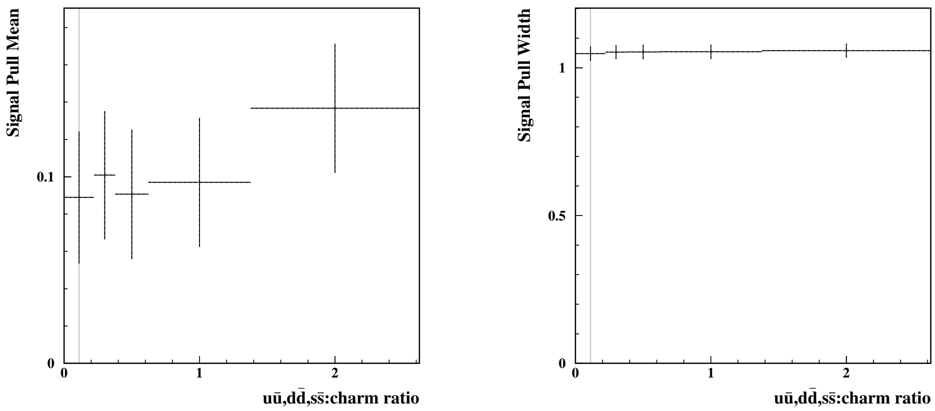


Figure A.63: $n_{signal} = 20$; $n_{ridge} = 100$; $n_{charm} = 1620$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 180$; $u\bar{u}, d\bar{d}, s\bar{s}$: $charm \equiv 2$



(a) Signal Pull Mean

(b) Signal Pull Width

Figure A.64: Signal pull mean and width as a function of the fixed $u\bar{u}, d\bar{d}, s\bar{s}\ : charm$ ratio. The grey line shows the true value of the ratio.

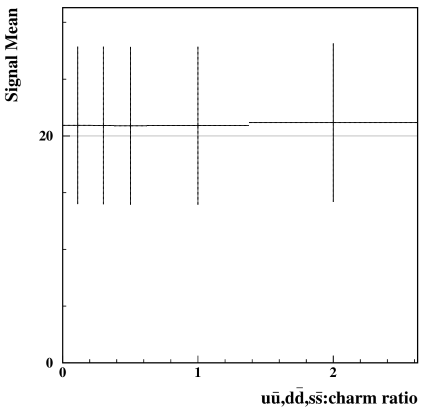


Figure A.65: Signal mean as a function of the fixed $u\bar{u}, d\bar{d}, s\bar{s}\ : charm$ ratio. The grey line shows the true signal level.

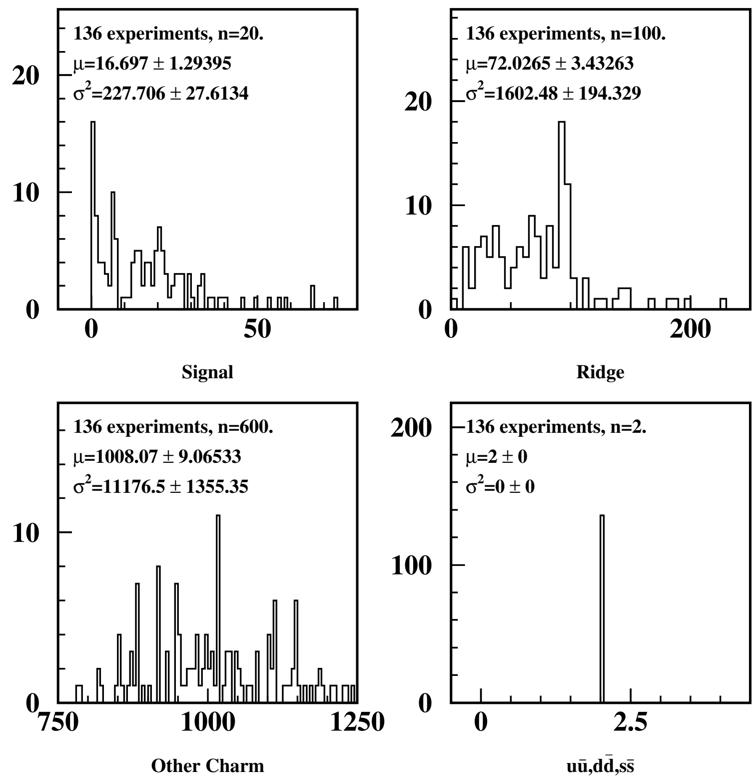


Figure A.66: $n_{signal} = 20$; $n_{ridge} = 100$; $n_{charm} = 1620$; $n_{u\bar{u},d\bar{d},s\bar{s}} = 180$; $u\bar{u}, d\bar{d}, s\bar{s}$: $charm \equiv 2$ (extended likelihood, 811 events total)

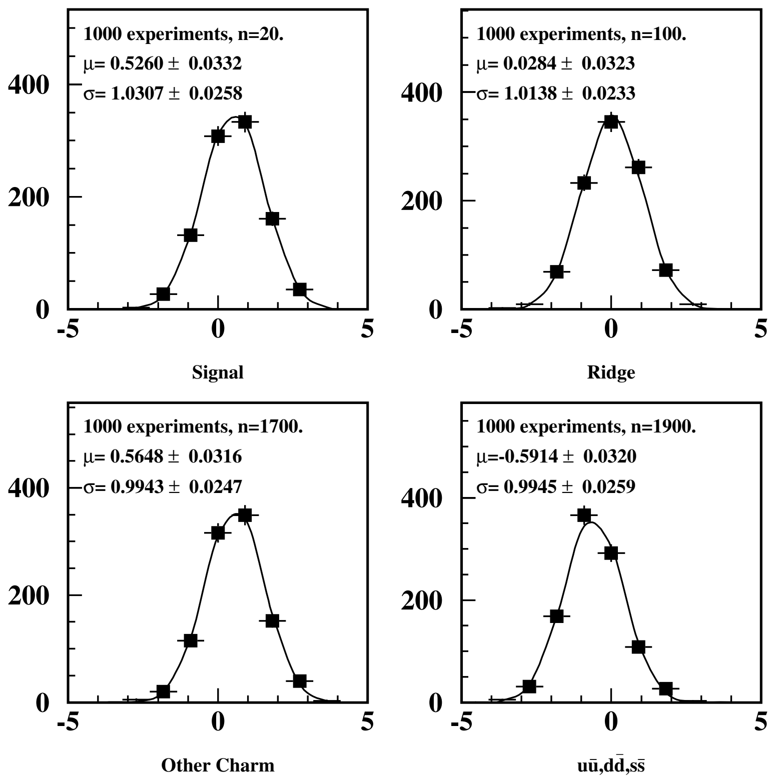


Figure A.67: No lifetime information in fit. No finite MC statistics terms in fit.

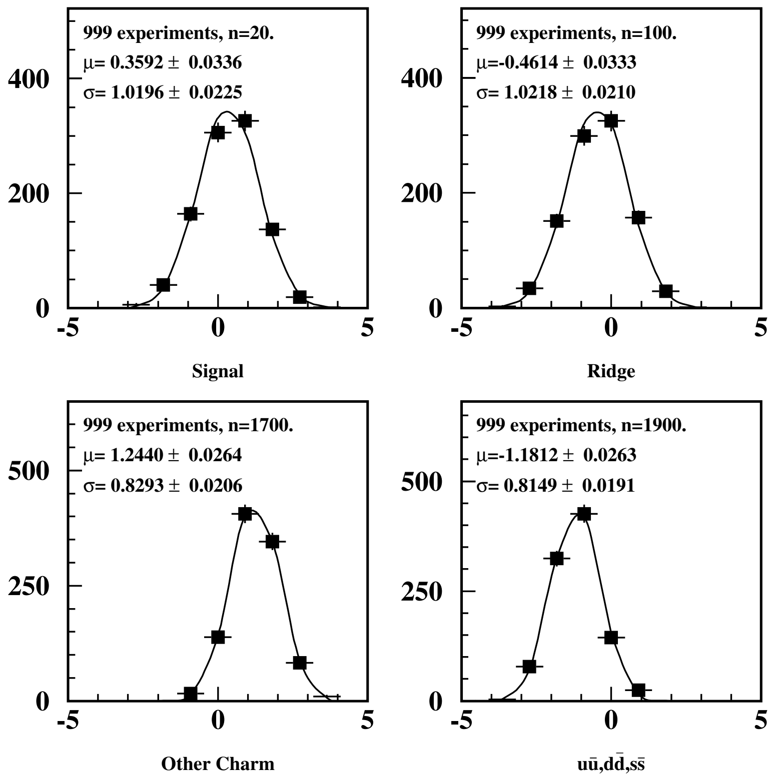


Figure A.68: Lifetime information included in fit. No finite MC statistics terms in fit.

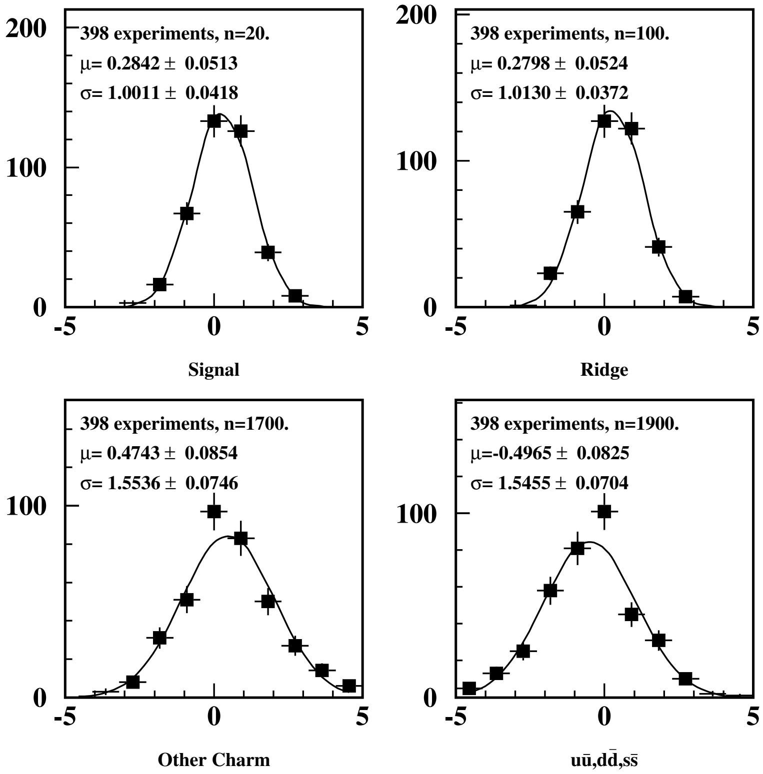


Figure A.69: No lifetime information in fit. Finite MC statistics terms included in fit.

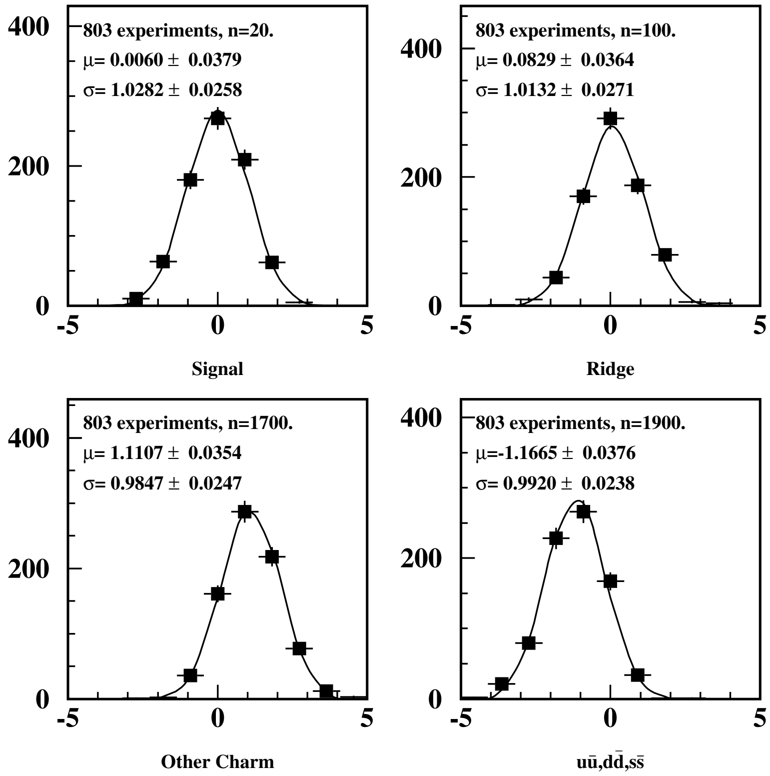


Figure A.70: Lifetime information and finite MC statistics included in fit.

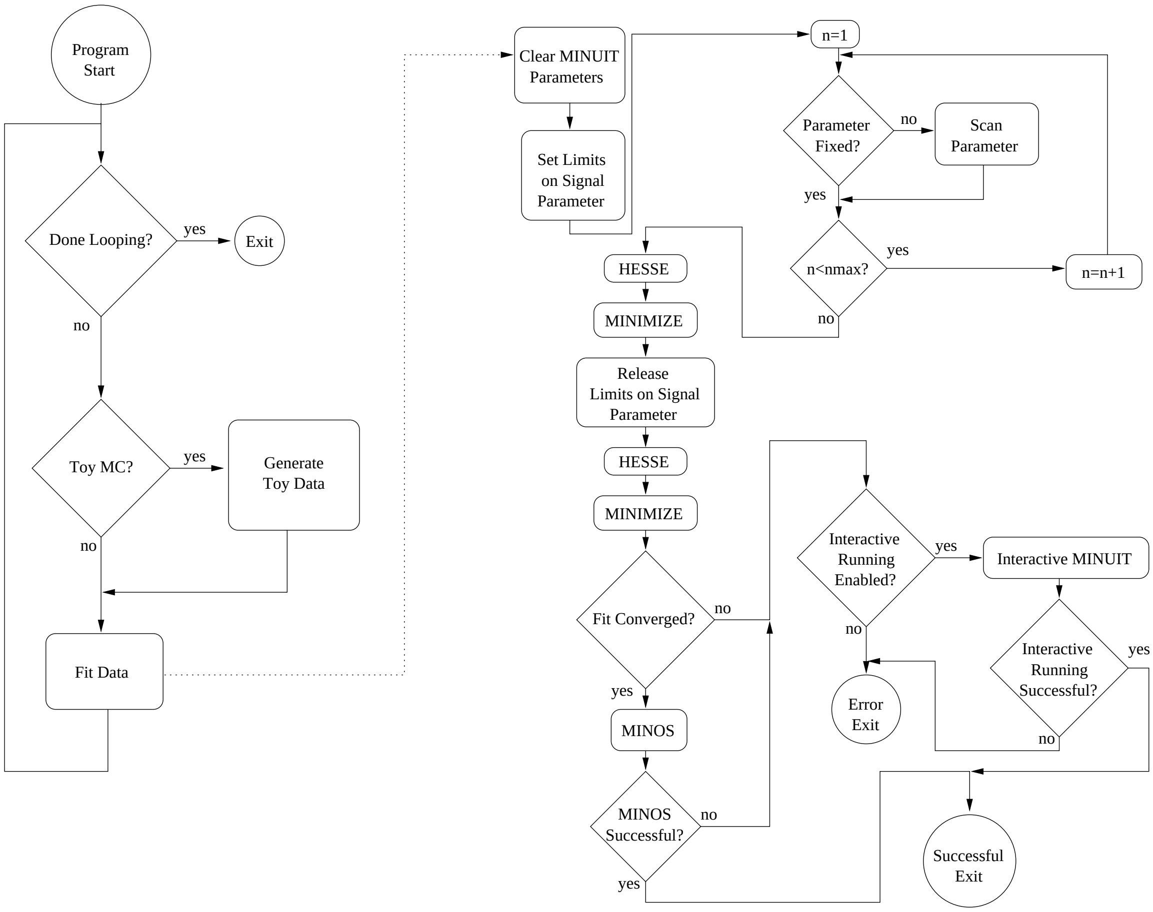


Figure A.71: Flowchart of Fitting Program

A.9 Summary

A flowchart of the general fitting procedure is shown in Figure A.71. The right-hand chart details the actual fitting step, from setting the parameters to deciding whether the fit succeeded or failed. The FFREAD "cards" used to control the program's behavior are given in Table A.1.

It seems safe to conclude that the fitter works. There are, however, certain limitations. The two broad components appear to be basically indistinguishable from the fitter's perspective, so the trivial approach of allowing both to float is insufficient. Instead, the ratio of the two fit areas should be fixed to a "reasonable" value (such as the ratio of Monte Carlo yields). As long as this ratio is not too different from the true value, the results should be reliable. This introduces a systematic error, though, which can be estimated by re-fitting with different values of the ratio. It is probably not worth fitting for the ratio to get a starting value, since the fitter has already been demonstrated to give very unreliable central values when both broad shapes are

Card	Type	Default Value	Meaning
NMBN	Integer	32	# of mass bins
NQBN	Integer	20	# of Q bins
DMAS	Real	1.8646	"True" D^0 mass
MHWP	Real	0.14	Half-width of plot in mass
MHWS	Real	0.0214	Half-width of signal region in mass
MDIF	Real	0.0058	"True" D^* Q
QHWP	Real	0.0055	Half-width of plot in Q
QHWS	Real	0.000566	Half-width of signal region in Q
LTOY	Logical	T	Use toy Monte Carlo?
NSIG, NRID, NOTC, NUDS	Integer	10k, 20k, 20k, 20k	Toy fitting shape statistics
SSIG, SRID, SOTC, SUDS	Integer	50, 1000, 3000, 2500	Shape contributions to toy data
LLUP	Logical	F	Run in looping mode (toy MC only)?
NLUP	Integer	0	Number of times through loop (0=forever)
NSKP	Integer	0	Number of loop iterations to skip
POIS	Logical	F	Use extended Poisson terms?
NSPB, NRPB, NCPB, NLPB	Integer	10, 10, 10, 10	# of bins for pull plots
SIGS, SIGR, SIGC, SIGL	Real	5., 5., 5., 5.	Half-width for pull plots
LTME	Logical	T	Use time to seed RANLUX?
SEED	Integer	0	Seed for RANLUX (if LTIME is false)
LUXL	Integer	3	RANLUX luxury level
RLK1	Integer	0	RANLUX parameter K1
RLK2	Integer	0	RANLUX parameter K2
FSIG, FRID, FOTH, FUDS	Logical	F, F, F, F	Fix shape parameter?
FTOT	Logical	T	Fluctuate toy data by total bin contents (rather than component-by-component)?
VSIG, VRID, VOTH, VUDS	Real	0.005, 1000., 3000., 2500.	Starting parameter values
INTM	Logical	F	Run interactive MINUIT on failed fits?
VRBS	Logical	F	Verbose MINUIT output in loop mode?
LRUC	Logical	F	Make $u\bar{u}, d\bar{d}, s\bar{s}$ parameter $u\bar{u}, d\bar{d}, s\bar{s}$ $/charm$?
LTAU	Logical	F	Include lifetime information in fit?

Table A.1: FFREAD "cards" used by the fitter

allowed to float, and also faces much greater chances of failing.

As for the use of the extended likelihood, the initially-observed problem with missing area can be corrected by ensuring that the histogram binning is sufficiently coarse. If the Monte Carlo statistics are large compared to the data statistics, there is no observed benefit in using the extended likelihood, but it increases the fitting time and probability of fit failure—a definite net loss. There is also no improvement in the pulls for grossly-incorrect $u\bar{u}, d\bar{d}, s\bar{s}$ -to-other-charm ratios. This leads us to the conclusion that the extended likelihood is not worth using in the fits to data.

Finally, the addition of lifetime information to the fit does not help distinguish between the two broad shapes, even though they have very different lifetimes. This is true both with and without the extended likelihood terms.

References

[1] M. Gell-Mann and A. Pais, Phys. Rev. **97**, 1387(1955).

[2] K. Lande *et al.*, Phys. Rev. **103**, 1901(1956).

[3] ARGUS Collaboration, H. Albrecht *et al.*, Phys. Lett. B **192**, 245(1987).

[4] A.A. Petrov, "Width Difference in the $D^0 - \overline{D^0}$ System," `hep-ph/0009160`. To be published in the proceedings of 4th Workshop on Continuous Advances in QCD, Minneapolis, Minnesota, 12-14 May 2000.

[5] E791 Collaboration, E.M. Aitala *et al.*, Phys. Rev. Lett. **77**, 2384 (1996).

[6] J. F. Donoghue, E. Golowich, and B. R. Holstein, *Dynamics of the Standard Model*. Cambridge: Cambridge University Press, 1992.

[7] FOCUS Collaboration, J.M. Link *et al.*, Phys. Lett. B **485**, 62 (2000).

[8] E791 Collaboration, E.M. Aitala *et al.*, Phys. Rev. Lett. **83**, 32 (1999).

[9] CLEO Collaboration, R. Godang *et al.*, Phys. Rev. Lett. **84**, 5038(2000).

[10] T.D. Lee and C.N. Yang, Phys. Rev. **106**, 340(1957).

[11] A. Pais and S.B. Treiman, Phys. Rev. D **12**, 2744(1975).

[12] Mark III Collaboration, D. Coffman *et al.*, Phys. Rev. D **45**, 2196(1992).

[13] E691 Collaboration, J.C. Anjos *et al.*, Phys. Rev. D **46**, 1941(1992).

[14] S.L. Glashow, J. Iliopoulos, and L. Maiani, Phys. Rev. D **2**, 1285(1970).

[15] H.N. Nelson, "Compilation of $D^0 - \overline{D}^0$ Mixing Predictions," submitted to the 1999 Lepton-Photon Symposium, `hep-ex/9908021` (unpublished).

[16] S. Bergmann, Y. Grossman, Z. Ligeti, Y. Nir, and A.A. Petrov, Phys. Lett. B **486**, 418(2000).

[17] CLEO Collaboration, Y. Kubota *et al.*, Nucl. Instrum. Methods Phys. Res., Sec. A **320**, 66(1992).

[18] T.S. Hill, Nucl. Instrum. Methods Phys. Res., Sec. A **418**, 32(1998).

[19] R. A. Briere, "Tracking in Helium-Based Gases: Present and Future B Factories," in *Heavy Flavor Physics*, ed. C. Campagnari, World Scientific (Singapore, 1999).

[20] C. Caso *et al.*, The European Physical Journal **C3**, (1998) pp.1-794.

[21] E791 Collaboration, E.M. Aitala *et al.*, Phys. Rev. D **57**, 13(1998).

[22] Look up reference for Kalman filtering.

[23] D. Cinabro *et al.*, `physics/0011075`. Submitted to Nucl. Instrum. Meth. A. CLNS 00/1706.

[24] E691 Collaboration, J.C. Anjos *et al.*, Phys. Rev. D **43**, 635 (1991).

[25] CLEO Collaboration, R. Ammar *et al.*, Phys. Rev. D **44**, 3383 (1991).

[26] CLEO Collaboration, G. Brandenburg *et al.*, Phys. Rev. Lett. **87**, 071802 (2001).

[27] H. Nelson and T. Nelson, "A Guide to Unbinned Maximum Likelihood Fits," CLEO internal note CBX 98-61.

[28] H. Nelson, "Goodness of Fit for Poisson Distributed Data," CLEO internal note CBX 98-73.

[29] D. Asner, T. Hill, and H. Nelson, "Cabibbo Suppressed $D^0 \to P^+P^-$ Decays," CLEO internal note CBX 01-28.

[30] FOCUS Collaboration, J.M. Link *et al.*, Phys. Rev. Lett. **86**, 2955(2001).

[31] CLEO Collaboration, G. Bonvicini *et al.*, Phys. Rev. D **63**, 071101(R) (2001).

[32] FOCUS Collaboration, J.M. Link *et al.*, Phys. Lett. B **491**, 232(2000).

[33] E791 Collaboration, E.M. Aitala *et al.*, Phys. Rev. Lett. **83**, 32(1999).

[34] `http://wwwinfo.cern.ch/asd/cernlib.html`

[35] F. James, "MINUIT — Function Minimization and Error Analysis", CERN Program Library entry **D506**, 1998.

[36] J.J. Moré and M.Y. Cosnard, "RSNLEQ — Numerical Solutions of Systems of Nonlinear Equations", CERN Program Library entry **C201**, 1994.

[37] F. James, "RANLUX — Uniform Random Numbers of Guaranteed Quality", CERN Program Library entry **V115**, 1994.

Vita

Michael Alan Marsh was born in Lowell, Massachusetts on June 24, 1972. He lived in Nashua, New Hampshire; Gaithersburg, Maryland; Waterloo, Belgium; and Potomac, Maryland before beginning his university education at Cornell University in 1990. He earned a Bachelor of Arts degree in Physics and another in Mathematics in May of 1994. During his senior year he was inducted into the Phi Beta Kappa and Golden Key honor societies, and participated in an undergraduate research project with his Physics advisor, Professor Albert J. Sievers.

After graduating from Cornell, he continued his studies as a graduate student in the Physics department of the University of Illinois at Urbana-Champaign. He earned a Master of Science degree in May of 1995. Under the direction of Professor Gary Gladding, he researched mixing of light charmed neutral mesons, earning a Doctor of Philosophy degree in October of 2001, and was very active in software development and maintenance both for the CLEO II/II.V and CLEO III experiments. From 1997 to 2001 he was a Visiting Fellow at Cornell University's Laboratory of Nuclear Studies.