Measurement of the Leptonic Branching Fraction of the D_s Meson and Determination of its Decay Constant

Thesis by

Michael H. Kelsey

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

1996
(Submitted May 21, 1996)

Acknowledgements

Doctoral thesis research is supposed to be an individual endeavor, an "original contribution to human knowledge," but physics, especially experimental high energy physics, is a cooperative effort. When accelerators cost billions of dollars, and detectors a tenth of that, it is inconceivable that a single researcher, especially a graduate student, could accomplish anything without relying on the work and efforts of many colleagues, friends, and fellow scientists.

First and foremost, I must thank my research advisor David Hitlin, who provided me with the support and encouragement I needed to pursue this project. He is one of the best advisors, as well as one of the best physicists, I have ever met. Despite his constant work in designing, developing, and promoting the SLAC B Factory, he was always available to me for questions, discussions, or just as a sounding board for my ideas. He gave me the freedom to pursue projects on my own, and to learn how to be a physicist by figuring out how to solve problems, not by being told the solution. He shielded me from the mundane issues of campus politics and the search for funding, so that I could fully develop my research skills. Although at times I thought he was spending more time on administration than on physics, his grasp of the details of current experiments, and his ability to explain clearly those details, and the bigger picture of research, was a constant source of knowledge, learning, and understanding for me. I am most grateful that I have had the opportunity to learn from David how to be a real physicist, and perhaps to pass on what I have learned to a new generation of students in years to come.

I could not hope to mention everyone who contributed to my learning and understanding these last eight years—it is only for economy of space that I limit my thanks to a few. My former office mates Lawrence Jones and Ernest Prabhakar (both of whom have left the field for greener pastures, and bank accounts), and my fellow Caltech graduate student Jim Panetta, were for years my colleagues, debating oppon-

ents, and friends. The energy and excitement of this field was apparent whenever we were together. My senior colleagues in BES, Walter Toki, Joe Izen, Bill Dunwoodie, and Zhang Changchun have given me advice, assisted me in solving problems, and supported me in the decision-making processes of our collaboration. Their suggestions at many junctures helped me to pursue my research, and to produce the best results I possibly could. My fellow graduate students at other universities, especially the "Gingerman Collaboration" of John Standifird, Pat Gratton, and others at the University of Texas at Dallas, made many suggestions and asked questions which often resolved my own misconceptions, and helped me to be sure that my work was as good as it could be.

I also express my thanks to the staff and faculty of the Chinese Institute for High Energy Physics, without whom this thesis would not even exist. In running the accelerator and the detector, and in managing the BES collaboration, they facilitated many measurements in tau-charm physics which have probed details of the Standard Model in unique ways, and at unprecedented levels of precision. Finally, the United States Department of Energy made my entire graduate career possible, by funding the Caltech High Energy Physics group through Contract No. DE–FG03–92–ER40701.

Abstract

Between February 1992 and June 1994, the Beijing Spectrometer (BES) experiment collected 24.5 pb^{-1} of data in e^+e^- collisions at a center of mass energy of 4.03 GeV, above threshold for production of $D_s^+D_s^-$ meson pairs. In this thesis, these data are searched for $D_s^+D_s^-$ events by identifying sets of charged tracks consistent with the decay channels $D_s \to \phi\pi, \phi \to K^+K^-$; $D_s \to K^{*0}K, K^{*0} \to K^+\pi^-$; or $D_s \to K_S^0K, K_S^0 \to \pi^+\pi^-$. A global maximum-likelihood technique is used, which combines particle identification detector data and a kinematic fit to each decay hypothesis, for maximum sensitivity. A sample of 190 candidates is found in the mass range 1958.5–1978.5 MeV/c^2, of which 76.9 $\pm$ 13.8 are estimated (from a fit to 1097 combinations with $1800 < m < 2015$ MeV/c^2) to be true $e^+e^- \to D_s^+D_s^-$ events.

This event sample is searched for leptonic decays of the D_s meson in the channels $D_s \to \mu\nu_\mu$ and $D_s \to \tau\nu_\tau$ (with $\tau \to \mu\overline{\nu}_\mu\nu_\tau$, $\tau \to e\overline{\nu}_e\nu_\tau$, or $\tau \to \pi\nu_\tau$). Five candidate decays are found. A likelihood function is formed as a product over the candidate events, where each contributes a sum over all of the leptonic decay channels, weighted according to the likelihoods of the particle identifications and the kinematics of the decay. This likelihood function is maximized with respect to the branching fraction $\mathcal{B}(D_s \to \mu\nu_\mu)$, or the D_s pseudoscalar decay constant f_{D_s}, with the results

$$\mathcal{B}(D_s \to \mu\nu_\mu) = .75^{+1.23}_{-.47}\,(stat) \pm .06\,(syst)\ \%\ ,$$
$$f_{D_s} = 308^{+164}_{-189}\,(stat) \pm 12\,(syst)\ \text{MeV}\ .$$

Comparison of the measured value of f_{D_s} with other experimental determinations and with a variety of theoretical models is discussed, with implications for constraining those models.

Contents

List of Figures

List of Tables

For Elizabeth

Part I

Introduction

Chapter 1 Overview

As a taxonomic enterprise, high energy physics is competitive with the most ardent zoological survey. The 1994 edition of the *Review of Particle Properties* [1] lists more than 180 identified particles, and thousands of particle decay channels or interactions. Particle names are a wondrous cacophony of Latin and Greek letters, subscripts, superscripts, symbols and numbers. All these measurements are empirical, the fruits of hundreds of person-years of labor and billions of dollars spent worldwide on accelerators, detectors, and computers.

Yet high energy physics is not merely "stamp collecting." Over the past thirty years, two models have developed to systematize and explain the ever-growing volume of experimental data. The electromagnetic and weak nuclear forces are unified as components of the spontaneously broken symmetry group $SU(2)_L \times U(1)_Y$ [2,3]. The strong nuclear force is explained as the interaction of particles in the $SU(3)$ symmetry group, via the theory of quantum chromodynamics [4]. Together, these form the Standard Model of particle physics, a theory which has met with success at every test, despite its empirical basis and its relatively large number of experimentally-determined parameters.

Of the vast number of experimental measurements, a few directly probe the underlying assumptions of the Standard Model. High energy scattering of electrons on nuclei (deep inelastic scattering) reveal pointlike quarks as the constituents of baryons [5]. Inclusive measurement of event rates in e^+e^- collisions verifies the fractional charge (and mass hierarchy) of quarks [6]. Observation of multiple jets of particles in high-energy collisions demonstrates the existence of gluons in quark-quark interactions [7,8]. Observation of mesons decaying exclusively to leptons probes the meson wavefunctions and the potential which holds quarks in bound states [9].

In this thesis, I present an absolute measurement of the decay rate of the pseudoscalar D_s meson (the lowest energy $c\bar{s}$ quark bound state) into leptons, $D_s \to \mu\nu_\mu$

and $D_s \to \tau\nu_\tau$. From these decays, I determine the decay constant of the D_s meson, f_{D_s} which can be related directly to the amplitude of the D_s wavefunction at zero distance. The data from which I make these measurements were taken with the Beijing Spectrometer (BES) at the Beijing Electron-Positron Collider (BEPC), a facility of the Institute for High Energy Physics of the People's Republic of China. BES is an international collaboration, involving approximately 150 researchers from more than a dozen institutions in the United States and China: China's Institute for High Energy Physics, the Chinese University of Science and Technology (Hefei), Shandong University (Jinan), Boston University, the California Institute of Technology, Colorado State University, the Massachusetts Institute of Technology, the Stanford Linear Accelerator Center, the former Superconducting Supercollider Laboratory, the University of California at Irvine, the University of Hawaii, the University of Texas at Dallas, and Washington University.

In the next chapter, I present an overview of the Standard Model of electroweak interactions, and theoretical models of meson structure and dynamics. In Part II, I present the experimental basis for my analysis. Chapter 3 covers the BEPC accelerator and the BES detector, emphasizing those components central to the identification of D_s leptonic decay events. Chapter 4 describes the detector triggers and filtering criteria used to select events for further analysis. Chapter 5 describes the various Monte Carlo data sets that were generated for analysis of event selection efficiencies and sources of background. Part III covers my analysis methodology (Chapter 6), the specific techniques I apply to particle identification (Chapter 7), selection of events containing possible D_s mesons (Chapter 8) and candidates for D_s leptonic decays (Chapter 9). In Part IV, I derive physics results from this analysis using the maximum likelihood technique (Chapter 10), including leptonic decay rates and the D_s decay constant (Chapter 11), and discuss the implications of these results for both theoretical models and other experiments (Chapter 12). Some of the software packages I developed or used in my analysis are presented, for reference, in the Appendices (Part V).

Chapter 2 Theoretical Underpinnings

2.1 The Standard Model

The Standard Model of high-energy physics is the most successful description we have of the fundamental interactions of matter. Despite having approximately 20 empirical parameters[1] there is a certain simplicity and elegance in its conventional formulation as a Lorentz-covariant Lagrangian quantum field theory, symmetric under transformations of the group $SU(3)_c \times SU(2)_L \times U(1)_Y$. The unified electroweak Lagrangian ($SU(2)_L \times U(1)_Y$) is [1]

$$\begin{aligned}\mathcal{L}_F &= \sum_i^{\text{fermions}} \overline{\psi}_i(i\,\partial\!\!\!/ - m_i - \frac{g m_i H}{2M_W})\psi_i \\ &\quad - \frac{g}{2\sqrt{2}}\overline{\psi}_i\gamma^\mu(1-\gamma^5)(T^+W_\mu^+ + T^-W_\mu^-)\psi_i \\ &\quad - eq_i\overline{\psi}_i\gamma^\mu\psi_i A_\mu - \frac{g}{2\cos\theta_W}\overline{\psi}_i\gamma^\mu(g_{V,i} - g_{A,i}\gamma^5)\psi_i Z_\mu \qquad (2.1)\end{aligned}$$

where $\theta_W = \tan^{-1}(g'/g)$ is the mixing angle between the $SU(2)$ (g coupling) and $U(1)$ (g' coupling) symmetries; $e = g\sin\theta_W$ is the positron charge; and A_μ is the vector potential (photon) electromagnetic field. The $T^\pm$ are the weak isospin raising and lowering operators, which couple weak isodoublet partners to the $W_\mu^\pm$ charged weak field[2]. The weak vector and axial couplings to the Z^0 (the Z_μ field) are

$$g_{V,i} \equiv t_{3,i} - 2q_i\sin^2\theta_W$$

[1]There are 19 basic parameters: the masses of the charged leptons and the quarks, the coupling strengths of each symmetry group, the mass of the Higgs boson and its coupling to the fermions, the mixing between the $U(1)_Y$ and $SU(2)_L$ symmetries, and the CKM quark mixing matrix. Some authors include neutrino masses and a hypothetical leptonic mixing matrix within the context of the Standard Model, which adds seven parameters; other authors consider the photon identically masses, which constrains the relationship between g, g', and θ_W.

[2]For quarks, the $W^\pm$ does not couple to the mass eigenstates q_i, but to the weak eigenstates $q_i' = \sum_j V_{ij}q_j$, where V_{ij} is the Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix between the ith and jth down-type quarks [10].

$$g_{A,i} \equiv t_{3,i} , \tag{2.2}$$

with $t_{3,i}$ the weak isospin of fermion i ($+\frac{1}{2}$ for up-type quarks and neutrinos; $-\frac{1}{2}$ for down-type quarks and electrons, muons, and taus). The q_i are the fermion electric charges in units of e. The Higgs mechanism spontaneously breaks the electroweak symmetry, giving mass to the weak vector bosons ($W^{\pm}$ and Z^0) and making the photon massless, and introducing a coupling of the fermions to the scalar Higgs field H, proportional to the fermion mass.

The strong interaction (QCD) Lagrangian [1] is symmetric under $SU(3)$ transformations of the quark color charge,

$$\mathcal{L}_{\rm QCD} = -\frac{1}{4}F^{(a)}_{\mu\nu}F^{(a)\mu\nu} + \sum_{q}^{\rm quarks} i\overline{\psi}^i_q\gamma^\mu(D_\mu)_{ij}\psi^j_q - m_q\overline{\psi}^i_q\psi^j_q , \tag{2.3}$$

$$F^{(a)}_{\mu\nu} \equiv \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g_s f_{abc}A^b_\mu A^c_\nu , \tag{2.4}$$

$$(D_\mu)_{ij} \equiv \delta_{ij}\partial_\mu - ig_s\sum_{a}^{\rm gluons}\frac{\lambda^a_{i,j}}{2}A^a_\mu , \tag{2.5}$$

where $g_s = \sqrt{4\pi\alpha_s}$ is the QCD coupling constant, and the f_{abc} are $SU(3)$ structure constants. The f_{abc} and λ matrices are discussed in detail in Ref. [1]. The (eight) gluon fields A^a_μ couple both to quark spinors ψ_q (with color indices i and j) and to one another.

In the Standard Model matter appears as fermions (spin-$\frac{1}{2}$ particles) which couple to one another via spin-1 gauge bosons. Left-handed fermions are doublets under the weak $SU(2)$ symmetry, coupling to the $W^{\pm}$ and Z^0 bosons; right-handed fermions are $SU(2)$ singlets. The $U(1)$ electromagnetic interaction, mediated by the photon (A^μ), couples to electrically charged fermions only, with equal left- and right-handed couplings. The two symmetry groups $SU(2)_L$ and $U(1)_Y$ are mixed, and the mass symmetry of their couplings broken, by the Higgs interaction. The neutral fields W^μ_3 ($SU(2)$) and B^μ ($U(1)$) are transformed into the physical Z^μ and A^μ via the weak

mixing angle θ_W,

$$\begin{aligned} A^\mu &= W_3^\mu \sin\theta_W + B^\mu \cos\theta_W \\ Z^\mu &= W_3^\mu \cos\theta_W - B^\mu \sin\theta_W \ . \end{aligned}$$

Finally, the quarks couple via their strong (color) charge to eight gluons (g) in $SU(3)$ color triplets, which are mixtures of the electroweak-coupling quark spinors. The mixing is described by the Cabibbo-Kobayashi-Maskawa matrix [10], operating on the down-type quarks, $d_i' \equiv \sum_j V_{ij} d_j$, where the d_j are the $SU(3)$ (strong coupling and mass) eigenstates, and the mixed d_i' are the electroweak eigenstates.

The fermions appear in three "families," or "generations," which differ only in mass; current measurements [1, 11–14] rule out the existence of more than three light neutrinos, a result which is conventionally interpreted as a limit on the number of quark and lepton generations. Figure 2.1 summarizes the "Periodic Table" of fundamental particles in the Standard Model [15].

The success of the Standard Model is both qualitative and quantitative: its symmetries and interactions are consistent with the observed interactions and decays of particles; and the current world-average results are consistent with it to within two standard deviations [1]. Where some experiments do not agree (for example, the solar neutrino problem [16]) there are several additional assumptions, which may be inadequate, built into the predictive calculations or the experimental interpretations.

2.2 Leptonic Decays of Mesons

Despite the success of the Standard Model, extracting results is still somewhat problematic, especially in quantum chromodynamics (QCD). The strong coupling constant, α_s, is a function of energy, becoming larger as the energy (or momentum transfer q) becomes small. At a scale below 1 GeV, α_s becomes unity, rendering perturbative (Feynman diagram) calculations, which rely on expansions in powers of α_s, useless. Non-perturbative techniques, such as lattice QCD [17], are promising, but do not

Leptons spin=1/2		
Flavor	Mass GeV/c^2	Charge
ν_e	$< 5 \times 10^{-9}$	0
e	0.000511	-1
ν_μ	< 0.0003	0
μ	0.106	-1
ν_τ	< 0.03	0
τ	1.7771	-1

Quarks spin=1/2		
Flavor	Approx Mass GeV/c^2	Charge
u	0.005	2/3
d	0.01	-1/3
c	1.5	2/3
s	0.2	-1/3
t	170	2/3
b	4.7	-1/3

Bosons spin=1		
Flavor	Mass GeV/c^2	Charge
γ (photon)	0	0
$W^\pm$	80.22	± 1
Z^0	91.187	0
g (gluon)	0	0

Figure 2.1: The Standard Model of high-energy physics. Adapted from Ref. [15].

yet provide precise, detailed results. Instead, calculations of hadronic physics rely on phenomenology and effective field theories, which convert the unknown effects of non-perturbative QCD into arbitrary parameters or functions, to be determined by experiment.

The leptonic decay of mesons is a good example of this phenomenological approach [18], and is the topic of my thesis. Figure 2.2 shows the lowest order Feynman diagram for the leptonic decay of a meson M with quark content $(Q\overline{q})$ to a lepton ℓ^- and its associated antineutrino $\overline{\nu}_\ell$.

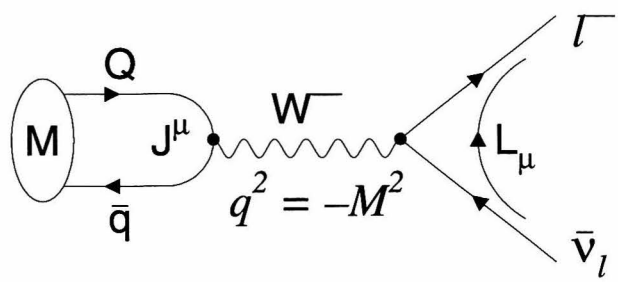

Figure 2.2: Lowest order Feynman diagram for the leptonic decay of meson M with quark content $(Q\overline{q})$, showing the quark current in the meson (*left*), the W propagator (*middle*), and the final state lepton current (*right*).

The currents L_μ and H^μ in Figure 2.2 are written

$$L_\mu = \overline{\ell}\gamma_\mu \frac{1}{2}(1-\gamma_5)\nu_\ell \tag{2.6}$$

$$J^\mu = \sum_{i,j} \overline{Q}_i \gamma^\mu \frac{1}{2}(1-\gamma_5) V_{ij} q_j \tag{2.7}$$

(assuming that Q is a down-type quark). Naïvely, one would expect to write the matrix element for Figure 2.2 as

$$\mathcal{M}(M_{Qq} \to \ell^- \overline{\nu}_\ell) = -i\frac{G_F}{\sqrt{2}} H^\mu L_\mu \ , \tag{2.8}$$

where $H^\mu = \langle 0|J^\mu|M\rangle$. The quark current is not free, and the non-perturbative effects that embed the quarks in the mesonic bound state cannot be evaluated directly

from QCD. We parameterize these unknown effects using functions (form factors) of the available four-vectors and polarization states in the problem. For leptonic decays of pseudoscalar mesons, these complications simplify tremendously: there is no polarization of the initial state, and the four-momentum transfer q^μ is constrained to $q^2 = -m_M^2$. What would otherwise be a form factor $f(q^2)$ becomes a constant f_M, the *decay constant* of the meson M, and we can write

$$\begin{aligned} \mathcal{M}(M_{Qq} \to \ell^- \overline{\nu}_\ell) &= i\frac{G_F}{\sqrt{2}} V_{Qq} f_M q^\mu L_\mu \\ \Gamma(M \to \ell\nu) &= \frac{G_F^2 |V_{Qq}|^2}{8\pi} f_M^2 m_M m_\ell^2 \left(1 - \frac{m_\ell^2}{m_M^2}\right)^2 . \end{aligned}$$

Specializing to the case of the D_s meson ($Q = c$, $\overline{q} = \overline{s}$), the leptonic partial width and branching fraction are

$$\begin{aligned} \Gamma(D_s \to \ell\nu_\ell) &= \frac{G_F^2 |V_{cs}|^2}{8\pi} f_{D_s}^2 m_{D_s} m_\ell^2 \left(1 - \frac{m_\ell^2}{m_{D_s}^2}\right)^2 \\ \mathcal{B}(D_s \to \ell\nu_\ell) &= \tau_{D_s} \Gamma(D_s \to \ell\nu_\ell) . \end{aligned} \tag{2.9}$$

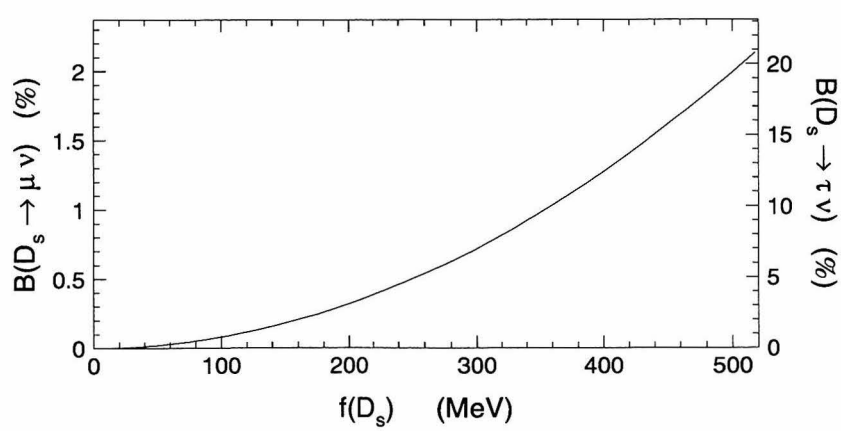

Figure 2.3: Dependence of $\mathcal{B}(D_s \to \ell\nu_\ell)$ on f_{D_s}, assuming lepton universality. The leptonic branching fraction is quadratic in the decay constant f_{D_s} for all lepton families. The left scale shows $\mathcal{B}(D_s \to \mu\nu_\mu)$ in percent, and the right scale shows $\mathcal{B}(D_s \to \tau\nu_\tau) = 9.75 \cdot \mathcal{B}(D_s \to \mu\nu_\mu)$.

Using the world-average values [1] for the parameters of equation 2.9, the branching

fractions depend only on f_{D_s} and the lepton mass,

$$B_\ell \equiv \mathcal{B}(D_s \to \ell\nu_\ell) = \left(7.179 \times 10^{-12}\mathrm{MeV}^{-4}\right) m_\ell^2 \left(1 - \frac{m_\ell^2}{m_{D_s}^2}\right)^2 f_{D_s}^2 \tag{2.10}$$

(with f_{D_s} and m_ℓ expressed in MeV), or

$$\frac{B_\ell}{B_\mu} = \left(\frac{m_\ell}{m_\mu}\right)^2 \left(\frac{m_{D_s}^2 - m_\ell^2}{m_{D_s}^2 - m_\mu^2}\right)^2 . \tag{2.11}$$

This gives the relation $B_e : B_\mu : B_\tau = 2.35 \times 10^{-5} : 1 : 9.75$, independently of f_{D_s}. In Figure 2.3, I plot equation 2.10 for B_μ and B_τ. For later reference, I write equation 2.10 as $\mathcal{B}(D_s \to \ell\nu_\ell) = A_\ell f_{D_s}^2$, where

$$\begin{aligned} A_\ell &\equiv \left(7.179 \times 10^{-12}\mathrm{MeV}^{-4}\right) m_\ell^2 \left(1 - \frac{m_\ell^2}{m_{D_s}^2}\right)^2 , \\ A_\tau &= 7.771 \times 10^{-7}\ \mathrm{MeV}^{-2} , \\ A_\mu &= 7.969 \times 10^{-8}\ \mathrm{MeV}^{-2} , \end{aligned} \tag{2.12}$$

using world-average values [1] for the lepton and D_s masses.

In systems containing at least one heavy quark, such as the charm and bottom hadron systems, one of the earliest models created to explain the observed masses involves a non-relativistic Schrödinger equation [9,19], in which the quarks are bound in a Coulombic potential with a linear confining term at large distances, *e.g.* [20],

$$V(r) = \kappa r + c - 4\alpha_s(r)/3r .$$

In this model, the pseudoscalar decay constant is related to the amplitude of the meson wavefunction at the origin (or equivalently, the overlap of the two quark wavefunctions), since the annihilation of Figure 2.2 is a contact interaction. The result is [9]

$$f_M^2 = \frac{12}{M}|\psi(0)|^2 . \tag{2.13}$$

Part II

Experimental Data

Chapter 3 Apparatus

3.1 Beijing Electron-Positron Collider (BEPC)

The Beijing Electron-Positron Collider (BEPC), located in southwestern Beijing, P.R.C., is similar in structure to the SPEAR storage ring at SLAC [21, 22]. The accelerator and storage ring are shown schematically in Figure 3.1. The complex supports e^+e^- collisions at center-of-mass energies from 3 to 5 GeV, as well as single-beam operation for generating synchrotron radiation. It has two interaction regions available, of which one is currently occupied by the BES detector (the other is not used). The nominal operating parameters of the accelerator are listed in Table 3.1 [22]. The coordinate system of the accelerator and detector are defined by the beams: the electron beam travels along the z axis from $-\infty$ to $+\infty$, and the $x-y$ plane is normal to the beam line, with the x axis horizontal.

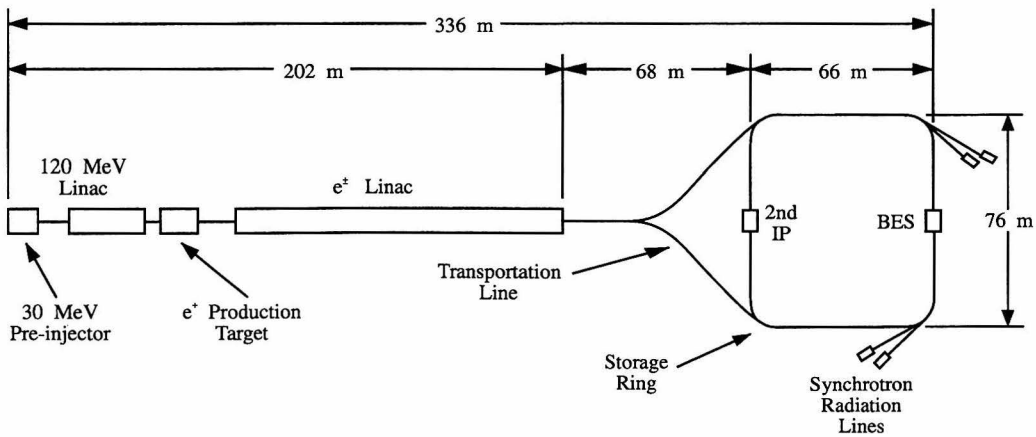

Figure 3.1: The BEPC Accelerator and Storage Ring, IHEP, Beijing, P.R.C.

Table 3.1: Operating Parameters of BEPC (Adapted from Refs. [21] and [22]).

Parameter	Symbol	Nominal Value
Center of mass energy	E_{cm}	3–5 GeV
Storage ring circumference	L	240.400 m
Bunch spacing	t_b	801.888 ns
Number of bunches	N	1
Natural bunch length	σ_l	5.20 cm
Particles per bunch at injection	N_b	6.8×10^{10}
Revolution frequency	f_0	1247.057 kHz
Horizontal beta function at IP	β_x^*	130 cm
Vertical beta function at IP	β_y^*	10 cm
Horizontal spot size at IP	σ_x^*	592 μm
Vertical spot size at IP	σ_x^*	38.7 μm
Horizontal linear tune shift	$\Delta\nu_x$	0.035
Vertical linear tune shift	$\Delta\nu_y$	0.035
Beam Pipe inner radius at IP	r^*	7.5 cm

The expected luminosity of BEPC is

$$L_{\rm nom} = \frac{N N_b^2 f_0}{(4\pi\sigma_x^*\sigma_y^*)} = 2.4 \times 10^{30}\ {\rm cm}^{-2}\,{\rm s}^{-1}\ .$$

The actual luminosity measured by BES (Section 3.2.2, below) during the 1992–1994 runs was typically a few 10^{30} cm^{-2} s^{-1}, matching this expectation well.

3.2 Beijing Spectrometer (BES)

The Beijing Spectrometer (BES) is a general purpose detector optimized for physics in the few GeV region. It has been described in detail elsewhere [23]. Figure 3.2 shows schematic views of the detector in the axial and transverse directions. The major components of the detector, from the center (beam line) outward, are described briefly in the following sections.

3.2.1 Beam Pipe

The beam pipe in the vicinity of the BES interaction point (IP) is 0.3 mm thick aluminum reinforced with 2 mm of carbon fiber, with an outer diameter (OD) of 15.4 cm. This corresponds to 0.0223 radiation lengths of material transverse to the beam line. The beam pipe encloses a vacuum of 2×10^{-10} torr.

3.2.2 Luminosity Monitor

The BES luminosity monitor is a set of four combination plastic scintillator/shower counter telescopes, arranged in pairs at opposite ends of the detector, as shown in Figure 3.3. Small pieces of plastic scintillator define a fiducial area for coincidence of Bhabha-scattered electrons, while a sampling calorimeter of tungsten and scintillator provides energy measurements for identification of luminosity (Bhabha scattering) events.

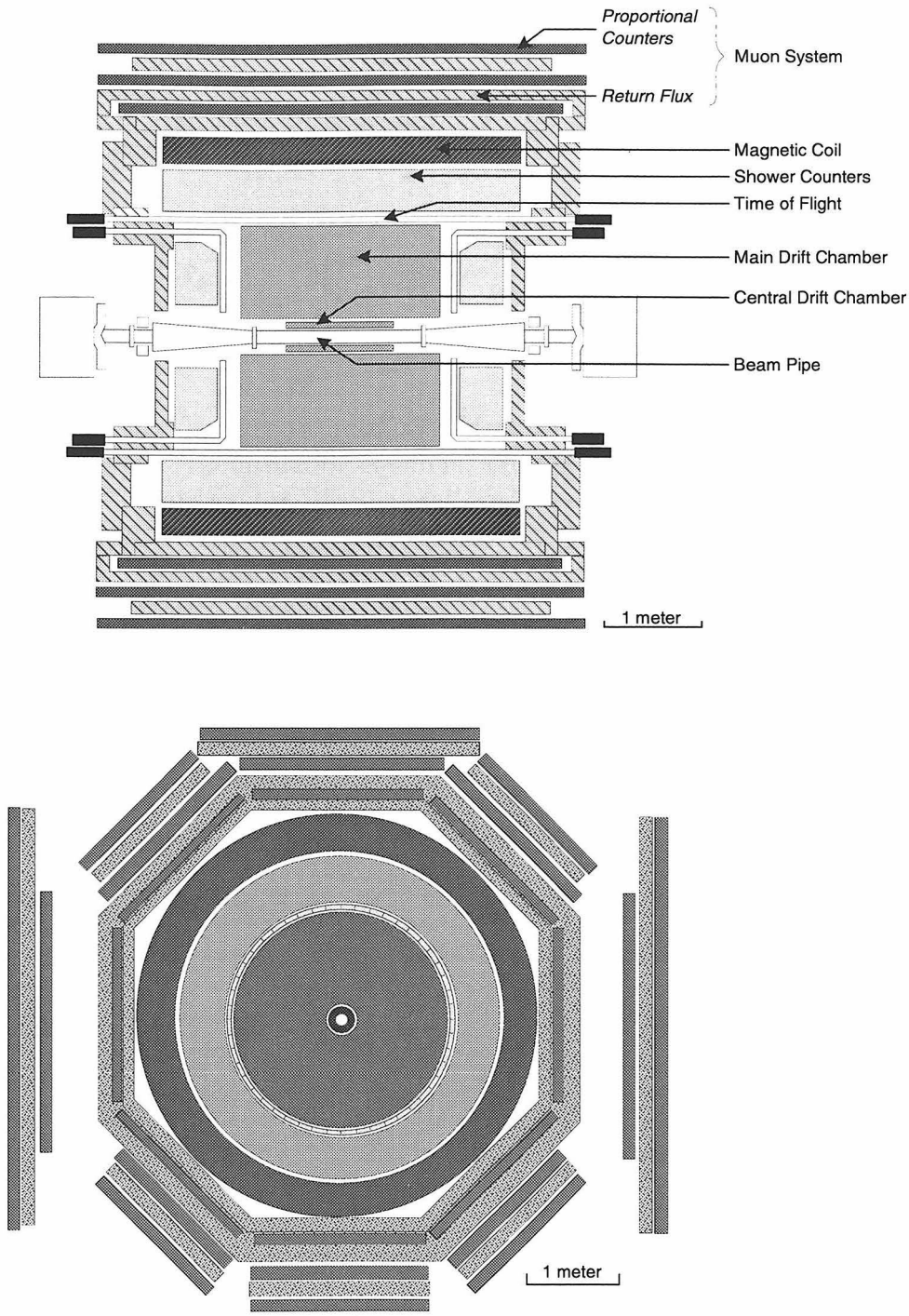

Figure 3.2: The BES detector. *(Top)* Transverse view; *(Bottom)* axial view.

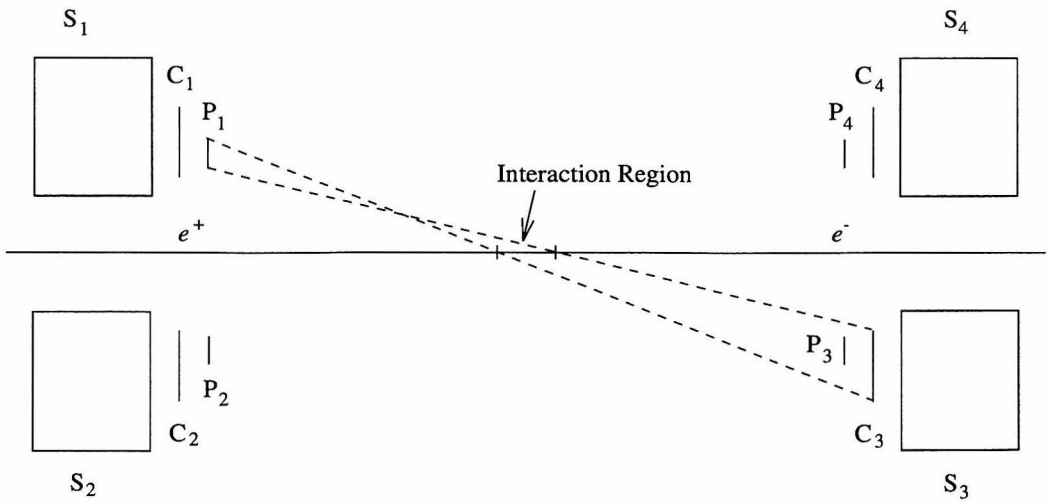

Figure 3.3: The BES Luminosity Monitor, showing the position (P) and coincidence (C) scintillation counters, and the sampling calorimeters (S). The angular acceptance for Bhabha-scattered electrons is also shown for one set of counters.

3.2.3 Central Drift Chamber

A four-layer central drift chamber (CDC, ID 18.4 cm, OD 30.2 cm, length 110 cm) is designed to provide tracking of charged particles close to the interaction point, for precise determination of primary and secondary event vertices. In practice, it was used only for triggering of events, and was not involved in reconstruction or data analyses. The CDC is a multiwire proportional chamber, using HRS gas (89% argon, 10% CO_2, 1% CH_4) as the active medium, and contains 48 $\times$ 1-sense wire cells per layer, with resolution of $\sigma_x = 150$ μm, $\sigma_z = 1$ cm (by charge division), and solid angle coverage of 98% of 4π.

3.2.4 Main Drift Chamber

The main drift chamber (MDC, ID 31.0 cm, OD 230.0 cm, length 220.0 cm) is the primary component for identifying charged particles in the BES detector. It is a multiwire proportional chamber also using HRS gas as the active medium, with 10 layers of 4-sense-wire cells, 5 axial and 5 stereo, as shown in Figure 3.4. There are

48 to 108 cells per layer. Tracking and momentum measurement of charged particles is done with pulse-time readouts of the sense wires; the momentum resolution is $\sigma_p/p = 2.1\%\sqrt{1+p^2}$ (p in GeV/c). Particle identification is done by measuring energy loss in traversing the chamber, via pulse-height measurements from each sense wire; dE/dx resolution is 8.5% FWHM for Bhabha-scattered electrons. The solid angle coverage of the MDC is 96% of 4π at the second layer, and 75% of 4π at the last (tenth) layer.

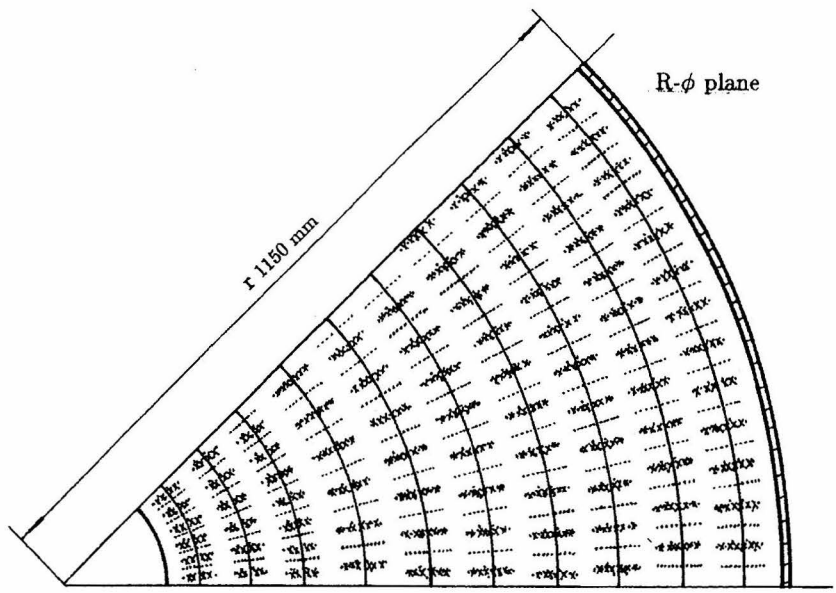

Figure 3.4: One sector of the BES main drift drift chamber (MDC) showing the ten layers of cells. Guard wires are indicated by the lines of dots ($\cdots$), the field shaping and signal wires are indicated by the lines of '$\times$'s.

3.2.5 Time of Flight Counters

Besides dE/dx, particle identification is also provided by time of flight counters in the barrel and end caps. The barrel TOF contains 48 15 cm $\times$ 5 cm $\times$ 284 cm scintillation counters arranged around the main drift chamber. The end cap TOFs are each 24 pieces of 2.5 cm thick scintillator forming a disk (ID 75 cm, OD 211 cm). Time resolution for electrons was approximately 300 psin 1990, and about 400 psin 1994, with combined coverage of 96% of 4π (75% of 4π for the barrel counters only).

3.2.6 Shower Counter

Identification of electrons and neutral particles (photons and π^0 mesons) is provided by electromagnetic calorimeters consisting of shower counters on both the barrel (BSC, ID 247.0 cm, OD 338.2 cm, length 385.0 cm) and end caps (ESC, ID 74.6 cm, OD 192.0 cm, thickness 41.0 cm), as shown in Figures 3.5 and 3.6. The chambers consist of 12 radiation lengths of lead/gas-counter sandwich, with the gas counters running in streamer mode to count shower particles. Resolution is $\sigma_E/E = 22\%/\sqrt{E}$ (E in GeV), $\sigma_\phi = 4.5$ mrad, $\sigma_z = 2$ cm in the BSC, and $\sigma_x = 0.7$ cm, $\sigma_y = 13\%L$ (tube length) in the ESC. The coverage of the BSC is 70% of 4π; combined coverage is 94% of 4π.

3.2.7 Magnet Coil

The BES detector is operated with a 4 kG axial magnetic field for momentum measurement. The field is provided by a conventional aluminum coil magnet surrounding the barrel shower counter (ID 3.48 m, OD 4.14 m, length 3.60 m), and in turn surrounded by several layers of steel-plate flux return.

3.2.8 Muon Tracking Counters

The magnetic flux return is instrumented with three double layers of proportional counter tubes sandwiched between layers of iron absorber, which are read out at both ends via charge division. This *muon counter* (MUC) system is most sensitive to muons, which can penetrate the material of the shower counter, magnet coil, and flux return much more readily than hadrons. Position resolution is $\sigma_z = 5$ cm, $\sigma_{r\phi} = 3$ cm; solid angle coverage is 68% of 4π.

3.3 BES Trigger System

The BES trigger system is used to decide whether the data acquired by the detector should be accepted and written to tape for reconstruction. The decision is made by

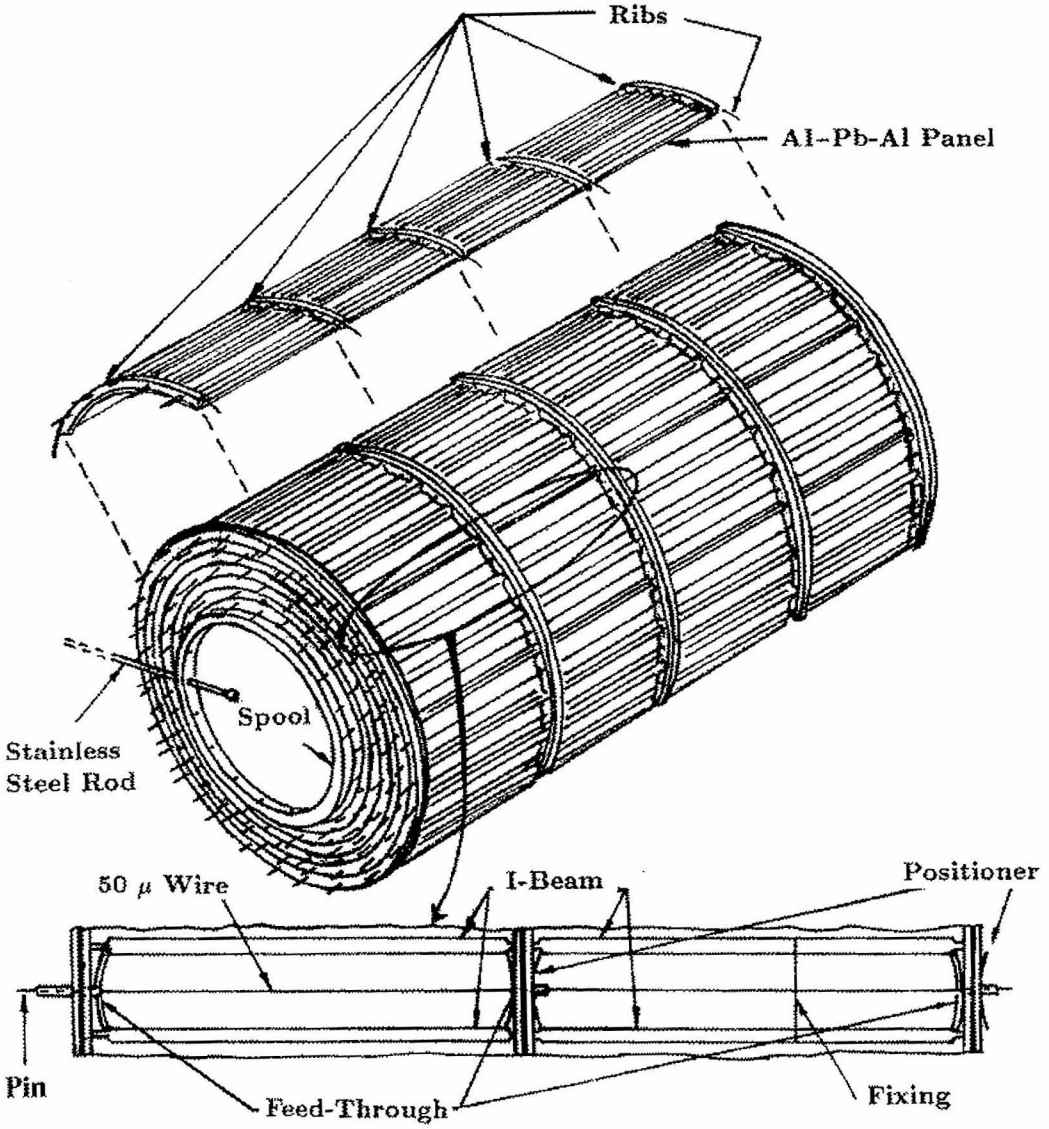

Figure 3.5: The BES barrel shower counter (BSC) showing overall construction *(top)* and an individual gas counter *(bottom)*.

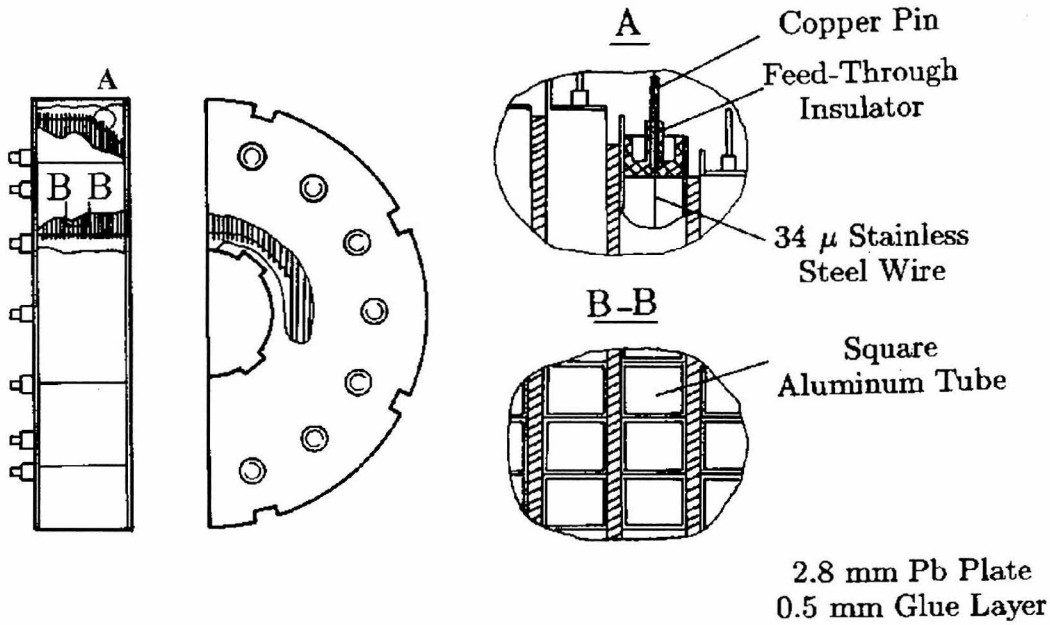

Figure 3.6: One of two BES end cap shower counters (ESC) showing overall construction in cross section *(left)* and plan *(middle)* and details of individual gas counters *(right)*.

combining up to sixteen trigger *conditions*, each corresponding to data from a single detector component, into a *trigger*. For a given data-taking period, up to eight triggers may be defined. The trigger conditions are listed in Table 3.2; specific triggers are defined for different running energies and physics selections. See Chapter 4 for D_s trigger definitions.

Table 3.2: BES Trigger Conditions.

Name	System	Description of Condition
TOF B-B	Barrel TOF	Paired hits in back-to-back counters
Ntof$\geq$1	Barrel TOF	At least one counter with paired hits
Ntof$\geq$2	Barrel TOF	At least two counters with paired hits
N-Veto	BSC+MDC	Non-zero energy and at least one track
End B-B	End cap TOF	Hits in back-to-back counters
Ntrk$\geq$1	MDC	At least one charged track
Ntrk$\geq$2	MDC	At least two charged tracks
Ntrk$\geq$4	MDC	At least four charged tracks
Muon-OR	MUC	At least one hit in any chamber
CDC-OR	CDC	At least one hit in layers 3 or 4
ESC-Et	ESC	Total energy above threshold
Etot-L	BSC	Total energy above low threshold
Etot-H	BSC	Total energy above high threshold

Chapter 4 Data Acquisition and Filtering

The BEPC accelerator and BES detector operated at a center-of-mass energy $E_{cm} = 4.03$ GeV for three running periods, from 1992 through 1994. An integrated luminosity of approximately 24.5 pb^{-1} was collected on tape during this time.[1] The global properties of this data, and the hadronic (charm) event selection procedures, are described below.

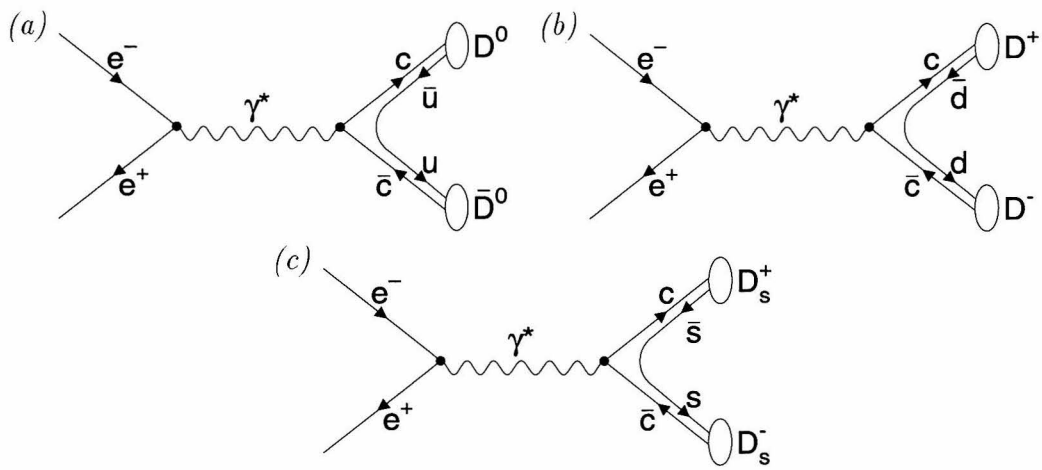

Figure 4.1: Lowest-order Feynman diagrams for production of charmed mesons in e^+e^- collisions: *(a)* neutral D production; *(b)* charged D production; *(c)* D_s production. Near threshold, the relative amplitudes of these diagrams, and the possible excitations of the mesonic bound states, must be computed using phenomenological models (c.f. Appendix A).

Figure 4.1 shows the lowest order Feynman diagrams for production of charmed mesons at BES. At $E_{cm} = 4.03$ GeV, the accelerator is above threshold for production of $D\overline{D}$, D^*D, and $D^*\overline{D}^*$ in both charged and neutral channels, as well as $D_s^+ D_s^-$. In addition to the "signal" charm states, a number of other physics channels, including light quark production, lepton pair production, and two-photon interactions, contrib-

[1] Luminosity is determined from a count of small-angle e^+e^- (Bhabha) scattering events and cross-checked against $e^+e^- \to e^+e^-$ and $e^+e^- \to \mu^+\mu^-$ events reconstructed in the BES detector.

ute to the total cross section (Figure 4.2). Both initial and final state radiation (e.g., $e^+e^- \rightarrow e^+e^-\gamma$) can modify the diagrams shown, as well as opening new physics channels, such as $e^+e^- \rightarrow \gamma\psi'$ [24] (Figure 4.3).

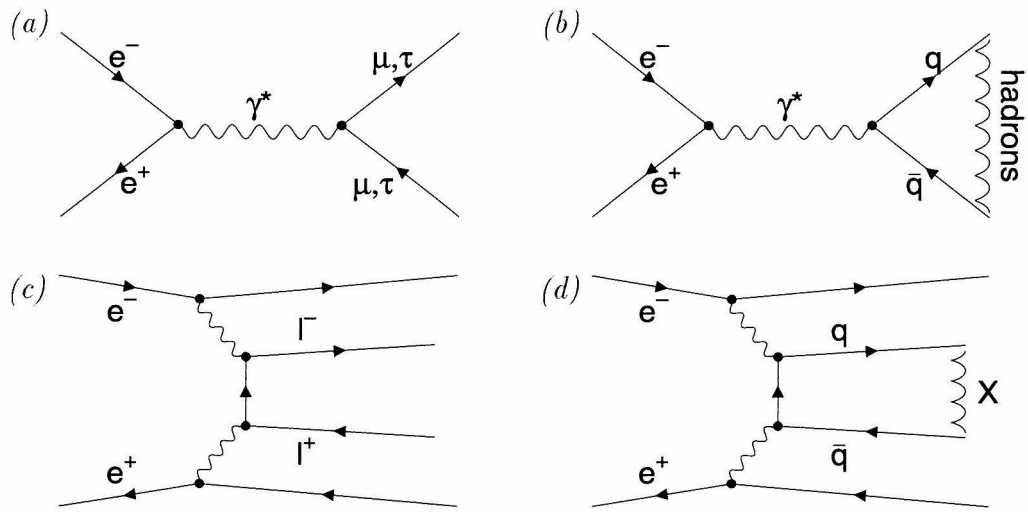

Figure 4.2: Additional physics processes accessible in e^+e^- collisions at $E_{cm} = 4.03$ GeV: Direct production of leptons *(a)* and hadrons *(b)*, and two-photon production of leptons *(c)* and hadrons *(d)* (the scattered beam particles tend to be lost at small angles).

4.1 Raw Data Sample

The BES trigger system (Section 3.3) was configured for the D_s running periods to provide maximum acceptance for the hadronic events expected from charm production (Figure 4.1). The combination of triggers used during the running is summarized in Table 4.1. With these triggers, the event rate was typically 5–10 Hz.

Table 4.2 lists the integrated luminosity for each D_s running period. During the 1993 runs, a data acquisition failure suppressed the readout of dE/dx information for approximately half the data. As a result, two different luminosities are reported: the total $\int L\,dt$ determined from small-angle Bhabha events during the run, and a "good dE/dx" luminosity, computed from reconstructed large angle e^+e^- scattering events for which dE/dx data was available for both tracks.

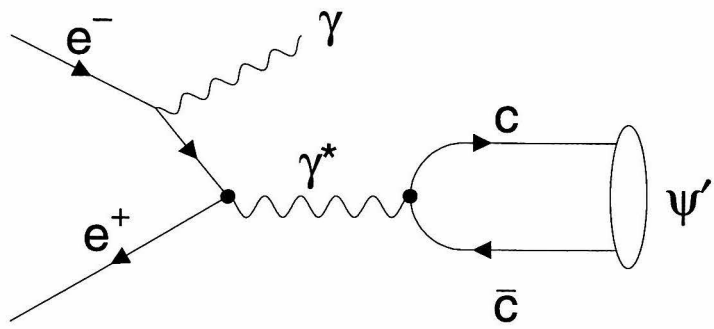

Figure 4.3: Initial-state radiation diagram (lowest-order) for charmonium production at $E_{cm} = 4.03$ GeV.

Table 4.1: Trigger selections defined and used ($\star$) during BES D_s data taking periods, 1992–1994. For each trigger, '+' indicates a required condition; '—' indicates a veto (required absence) condition. See section 3.3 for the details of trigger conditions.

	$\star$			$\star$			$\star$	
Conditions	CHARGED	COSMIC	DS1	DS2	DS3	DS4	ESC	2-MU
TOF B-B								
Ntof$\geq$1	+	+			+			
Ntof$\geq$2			+	+				+
N-VETO								
END B-B						+	+	
Ntrk$\geq$1	+	+	+			+		+
Ntrk$\geq$2				+	+			
Ntrk$\geq$4		—						—
MUON-OR		+						+
CDC	+		+	+	+	+	+	+
ESC-Et							+	
Etot-L				+				
Etot-H	+							

My analysis utilizes only the "good dE/dx" data, since not only is particle identification essentially impossible without dE/dx measurements, but the electronics failure invalidated the drift chamber timing information for those events, making the track reconstruction useless.

Table 4.2: Integrated luminosity for BES D_s running periods (*yymmdd*). Run numbers are used internally for reference by the BES collaboration; run numbers and periods not listed refer to data taken at different energies for calibration or for other physics studies. Both the total integrated luminosity ($\int L\,dt$) and the luminosity for "good dE/dx" [good] events are listed; see text for details. N_{had} is the number of events after the hadronic event selection (Section 4.2) is applied.

Period	Run Numbers	$\int L\,dt$ [good] (nb^{-1})		N_{had} (events)
920124–920319	3097–3578	2296.257	*all*	117 359
920402–920421	3739–3966	982.820	*all*	58 700
921206–921221	4363–4423	253.531	[124.845]	0
930214–930511	4489–5304	5012.759	[2517.876]	432 833
930511–930615	5305–5698	2020.537	*all*	131 329
940106–940131	6060–6413	3251.615	*all*	211 726
940131–940331	6414–6946	6519.888	[6519.622]	352 520
940401–940514	6947–7426	4163.488	*all*	271 102
Total		24500.895	[21877.052]	1 575 569

4.2 Hadronic Event Selection

At 4.03 GeV, the total charm production cross section is predicted to be $\sigma(\mathrm{charm}) = 12.8$ nb, according to the coupled-channel model of Eichten, *et al.* (Appendix A). With a typical luminosity of a few 10^{30} $\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, we expect a charm ($D$, D^*, or D_s) event rate $\sim$ 0.03 Hz, or one or two events per minute. A vast majority of the triggered events in the raw data samples are not going to be important for D_s or other charm physics analyses. As part of the reconstruction process, only events consistent with hadronic production were selected for further analysis (separate selections of two-prong

Bhabha-scattering or dimuon production were also made).

The hadronic event selection [25] consists of three general criteria which reasonably describe the physics of multihadron production in e^+e^- collisions: multiple charged and neutral tracks, all originating from a common vertex (the interaction point), with summed momenta consistent with the e^+e^- collision. The detailed selection criteria are

1. $N_{chg} < 12$ OR $\{N_{chg} \geq 12$ AND $P_{sum} \leq 10$ GeV$\}$,
2. $N_{good} \geq 3$ OR $\{N_{good} < 2$ AND $N_\gamma \geq 2\}$,
3. $V_{xy} < 2$ cm AND $|V_z| < 20$ cm, and
4. $P_{sum} > 1.5$ GeV

N_{chg} is the total number of reconstructed charged tracks, and N_{good} the reconstructed tracks with converged fits ($\chi^2 < 50$). N_γ is the number of isolated photons in the event: neutral energy clusters in the shower counter with $E > 50$ MeV and at least 14° separation from the nearest charged track ($\cos\theta_{\gamma c} \equiv \max(\vec{r}_\gamma \cdot \vec{p}_c) < 0.97$). V_{xy} and V_z are estimates of the primary vertex position (relative to the center of the detector),

$$V_{xy} = \frac{1}{N}\left[\left(\sum_i w_i x_i^{\rm DCA}\right)^2 + \left(\sum_i w_i y_i^{\rm DCA}\right)^2\right]^{1/2} \tag{4.1}$$

$$V_z = \frac{1}{N}\sum_i w_i z_i^{\rm DCA} \tag{4.2}$$

$$w_i \equiv 1 - \frac{r_i^{\rm DCA}}{10\ {\rm cm}}$$

averaged over all good charged tracks with reasonable momentum ($0.1 < p < 4$ GeV) and a distance of closest approach (DCA) to the beam axis of less than 10 cm. The scalar momentum $P_{sum} = \sum_i |\vec{p}_i|$, is summed over all charged tracks with reasonable momenta, and all isolated photons (see above).

Criterion (1) eliminates detector "noise" corresponding to large sectors of the main drift chamber (apparently) firing simultaneously. The reconstruction software attempts to form tracks from all the contiguous "hits," resulting in many dozen tracks,

or tracks with physically impossible momenta (tens or hundreds of GeV). This criterion generally limits the data to a maximum of 12 prong events, while allowing higher multiplicities if the momentum sum is sensible. With an average charm decay multiplicity of ~ 2.6 [26], less than 0.45% of charm events could be rejected; with the momentum-sum cross-check, that fraction is reduced still further.

Criterion (2) selects $e^+e^- \to q\overline{q}$ physics. It requires more than two charged tracks (eliminating Bhabha-scattering, dimuon production, and cosmic rays), or at least two isolated photons if only two charged tracks are present.

Criterion (3) and (4) eliminate beam-gas and beam-wall interactions, by requiring all of the charged tracks to originate from the interaction point in both the xy plane and along the beam (z) axis, and by requiring that the charged tracks in the event have significant momentum, corresponding to at least (approximately) the beam energy.

Chapter 5 Monte Carlo Generated Data

At the BES center-of-mass energy E_{cm} = 4.03 GeV, a number of different charmed mesons may be produced, along with tau lepton pairs, and multi-hadron events from light quark ($u\overline{u}$, $d\overline{d}$, $s\overline{s}$) production; see Chapter 4. The BES collaboration has generated Monte Carlo simulations of many of the physics channels available at 4.03 GeV, in groups of 100,000 events. These simulations include the expected angular distributions of the initial physics processes, decays of generated hadrons according to the current world average branching fractions [1], and a realistic detector response based on inclusive analyses of the 1992 data sets. Since not all possible decay channels have been observed experimentally, a variety of methods are used to estimate the "missing" branching fractions: theoretical models such as that of Bauer, Stech, and Wirbel [27]; isospin and $SU(3)$ symmetries to relate measured and unmeasured ratios; and approximate phase space relationships for modes which differ only in the number of final state pions.

For my analysis, I use these Monte Carlo data sets to make estimates of the signal efficiency and various sources of background for both my D_s event selection and my identification of candidate D_s leptonic decays. The coupled-channel model of Eichten, *et al.* [28], makes predictions for the absolute production rates of charmed mesons in e^+e^- collisions, as described in Appendix A; I use these predictions to normalize the results from each Monte Carlo data set, to obtain realistic estimates for the relative background levels from different sources.

Part III

Data Analysis

Chapter 6 The Maximum Likelihood Method

A likelihood function is analogous to a probability density function (p.d.f.), where the *parameters* of the likelihood function may be treated as random variables. Although this may not be formally true, it is sometimes a convenient fiction when constructing a likelihood function, as well as when considering the properties and interpretation of likelihood results. Likelihoods are multiplicative: a likelihood function representing a complex situation may often be written as a product of simpler terms, just as a joint p.d.f. may be constructed as a product of statistically independent p.d.f.'s. For a given set of measurements, the "most probable" values of the likelihood parameters are those for which the likelihood function is maximized, just as the most probable value of a random variable is that for which the p.d.f. is maximized.

In this chapter, I will discuss general techniques for constructing a likelihood function for an analysis, the application of the maximum likelihood technique to measurements and hypothesis testing, and the use of integrated likelihoods for testing the consistency of results, and for constructing confidence intervals.

6.1 Constructing A Likelihood Function

There are two broad classes of likelihood functions, both of which appear as part of the analysis I describe in this thesis. The first class consist of parameterized functions of experimental data, in which one or more of the parameters are unspecified. Usually, the unspecified parameter(s) represent definite, but unknown, physical quantities such as a branching fraction or the peak position and width of a resonance. If the experimental situation is well-understood, we may construct p.d.f.'s for the measured data inter terms of underlying physical quantities: for example, the probability of observing

some number of decays in a fixed time is given by a Poisson distribution, in which the expected number depends on detection efficiency and on the true decay rate in a sample,

$$\begin{aligned} p(N;\mu) &= \frac{e^{-\mu}\mu^N}{N!} \\ \mu &= N_0 \varepsilon e^{-\Delta t/\tau} \ . \end{aligned}$$

The true decay rate τ is an unspecified, but definite, parameter of this p.d.f., while our experimental measurement N is the random variable. We may turn this into a likelihood function by considering N to be fixed by our measurement, and τ as the unknown. We assume that the measurement we got (the only information we have) is the most probable measurement, and solve for the value of τ which maximizes p, as I will discuss further below.

The second class of likelihood functions are not functions of some unknown parameter, but are sets of discrete values, each corresponding to one hypothesis among several (the first class described above is a limiting case of this, where the "set of hypotheses" are all of the unknown values of the parameter). A simple example is my particle identification technique, which I describe in Chapter 7. We know how a given detector should respond to each of several types (mass, charge, momentum) of particle. We can construct a likelihood function which measures the "distance" between a measured detector output and each of the expected responses, for example with a chi-squared that takes into account resolutions of detector components. The measured output is a random variable; if it is distributed normally about the expected response for a given particle h, then we can write

$$\mathcal{L}_h \equiv p_h(x_{\rm obs}) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{1}{2}\left(\frac{x_{\rm obs} - x_h}{\sigma_x}\right)^2\right) \ .$$

This likelihood "function" has no unspecified parameters; it is just the probability p_h that a given particle h will cause our detector to respond as we observed it to respond. As a likelihood we can compare several different possible h's, and interpret the set

$\{\mathcal{L}_h\}$ as the relative probabilities that a particle of such type actually traversed our detector to cause the observed measurement.

6.2 Hypothesis Selection by Maximum Likelihood

Given a likelihood function $\mathcal{L}(p)$, we conventionally define the "measured" value p_0 of the parameter as that for which $\mathcal{L}(p_0) > \mathcal{L}(p)$, $\forall p \neq p_0$, as I alluded to above. When $\mathcal{L}(p)$ is a continuous function of the parameter, we find p_0 formally by differentiating the likelihood with respect to p, setting

$$\frac{\partial \mathcal{L}}{\partial p} = 0 \ ,$$

and solving for the value p_0 which satisfies this condition. In realistic situations, for example in my analysis of D_s leptonic decays (Part IV), the likelihood function is too complex to find a solution analytically. There are a variety of programs available which can perform numerical maximization (or minimization of $-\ln \mathcal{L}$); the program MINUIT [29], distributed as part of the CERN Program Library, is widely used in the high energy physics community. When the likelihood function consists a of discrete set of values, hypothesis selection is much simpler. The set of values is compared, and that with the largest value of $\mathcal{L}_h$ is considered correct.

6.3 Consistency Checking with Confidence Levels

In addition to selecting the best hypothesis, it is important to ensure that the measured data are consistent with that hypothesis. Not all measurements are necessarily drawn from the hypothetical distribution of interest: noise, detector problems, or other effects can generate data which are not physically reasonable for any charged particle hypothesis, for example. Nonetheless, a formal likelihood expression may still be evaluated, and some hypothesis h_0 or parameter value p_0 will turn out to have a likelihood greater than the others. To evaluate the "reasonableness" of that hypothesis,

the confidence level or probability is a useful tool [30]. Typically, this is the upper tail integral of the probability density or likelihood function is used,

$$Q(y) = \int_{-\infty}^{y} q(y')\,dy' \ , \tag{6.1}$$

where $q(y)$ is the normalized likelihood function,

$$q(y) = \frac{\mathcal{L}(y)}{\int_{-\infty}^{\infty} \mathcal{L}(y')\,dy'} \ . \tag{6.2}$$

In my analyses, I define my likelihood functions such that condition 6.2 holds naturally, $q(y) \equiv \mathcal{L}(y)$; this will be assumed for the rest of the discussion below.

A joint likelihood $\mathcal{L}(y_1, ..., y_n)$ has a confidence level defined as the multidimensional integral

$$Q(y_1, ..., y_n) = \int_{-\infty}^{y_1} dy_1' \cdots \int_{-\infty}^{y_n} dy_n' \mathcal{L}(y_1', ..., y_n')$$

which, for uncorrelated variables, factors into

$$\begin{aligned} Q(y_1, ..., y_n) &= \int_{-\infty}^{y_1} dy_1' \cdots \int_{-\infty}^{y_n} dy_n' \mathcal{L}(y_1') \cdots \mathcal{L}(y_n') \\ &= \int_{-\infty}^{y_1} dy_1' \mathcal{L}(y_1') \cdots \int_{-\infty}^{y_n} dy_n' \mathcal{L}(y_n') \\ &= Q(y_1) \cdots Q(y_n) \ . \end{aligned} \tag{6.3}$$

For uncorrelated variables, the total confidence level is thus the product of the separate confidence level expressions.[1]

These general techniques of maximum likelihood hypothesis selection and minimum probability (confidence level) hypothesis testing are used throughout my analysis for particle identification, for selection of hadronic D_s decays, and for the deter-

[1]Computationally, it is often simpler to work with chi-squared variables than with confidence levels, since for uncorrelated distributions $P(\chi_a^2; n_a)P(\chi_b^2; n_b) = P(\chi_a^2 + \chi_b^2; n_a + n_b)$. In my analysis software, I convert non-Gaussian confidence levels into "equivalent chi-squareds" by inverting

$$P(\tilde{\chi}^2; n_{df}) = Q(y) = \int_{-\infty}^{y} \mathcal{L}(y')\,dy'$$

to obtain $\tilde{\chi}^2$, the χ^2 equivalent to y.

mination of $\mathcal{B}(D_s \to \ell\nu_\ell)$ and f_{D_s}. I believe that this technique, applied consistently, ensures sensitivity to all available data, and accounts for statistical fluctuations and backgrounds in a natural, quantitative way. I discuss this aspect of my analysis below, with a qualitative comparison of the likelihood technique with more common selection, or "cut," methods.

6.4 Comparison with Other Techniques

In a "typical" analysis in high-energy physics, myriad detector measurements are distilled down to a relatively small number of physically useful quantities—momenta, energies, particle identification parameters, etc. Those quantities are then combined and *cut* to select among various hypotheses. While this is often an intuitively appealing exercise, it is not particularly efficient. To separate overlapping hypotheses, or to isolate a signal from surrounding background, each cut necessarily rejects some true events. When many cuts, each with $\varepsilon < 1$, are combined, the final efficiency can be miniscule. In a low statistics environment, the loss of precision that results can mean the difference between a measurement and a limit, or even preclude any measurement.

Consider the selection of $K^0_S \to \pi^+\pi^-$ decays as a simple example. A traditional analysis might proceed by selecting pairs of oppositely charged tracks, and requiring them to be identified as pions. Considering the overlap between π, μ, and K identification in a typical detector, this selection might have $\varepsilon = 0.7$ per track, or 0.49 for the pair. Next, the invariant mass of the pair is computed, and compared to the expected mass of a K^0_S. To eliminate combinatoric background, a $\pm 2\sigma$ selection might be done, with an efficiency of $\varepsilon = 0.95$. To reduce background further, it might be necessary to impose a requirement that the pion tracks cross at a point separated from the event production vertex; depending on the position resolution of the detector, this could eliminate as many as half the true K^0_S's in the event. Overall, then, this typical selection method is less than 25% efficient.

The technique of maximum likelihood selection can compensate for the loss of efficiency in a cut cascade. Instead of proceeding systematically through a logical

chain, selecting each component of a physics process, we can treat each of those components as having some probability distribution, or likelihood function. The entire process of interest is then described by a joint likelihood which is the product of all the components, and the final selection is made by evaluating and cutting on that single function. Different aspects of the selection process can compensate one another, leading to a net efficiency which is greater than that of the cascade of separate cuts. For the $K^0_S \to \pi^+\pi^-$ example above, we can use equation 6.3 to form a "joint probability" combining the particle identification measurements for both tracks with position and mass resolution functions. This joint probability might be required to be greater than 0.1 or 0.2 for good identification. If it has been well-constructed, the probability should be uniform for true K^0_S decays, implying that the efficiency is just the complement of the selection value ($\varepsilon = 0.9$ for $P(K^0_S) > 0.1$), which can be significantly greater than the efficiency of a traditional analysis.

Chapter 7 Particle Identification

An essential component of my analysis is efficient, background-free identification of all of the charged particles in each event. Identification and separation of pions and kaons is critical to select the D_s hadronic decays in candidate $e^+e^- \to D_s^+ D_s^-$ events. Good identification of muon and electron tracks, and rejection of hadrons, is central to identifying D_s leptonic decay candidates. In this chapter, I present an application of the maximum likelihood method, which makes optimal use of all available detector information to do particle identification and separation over the full momentum range of interest.

7.1 Kinematic Samples

The first step in constructing a particle identification function is knowing the response of the detector to different particles. For some detector components, this could be done with test beams produced at accelerator facilities; a powerful simulation program, such as GEANT [31], could be used to model the response of detector components provided the detailed structure of the detector, including such complex variables as gas temperature, pressure, and composition, are sufficiently known. For a variety of reasons, neither of these options were available for the BES detector. Instead, we have used actual particle tracks, selected according to kinematic constraints, to measure the response of the BES detector components.

In the sections below, I present the physics process and selection criteria used to produce each sample of particles and show distributions of measured quantities in the major detector components.

7.1.1 Electrons

A sample of electrons is selected by "tagging" radiative Bhabha events, $e^+e^- \rightarrow e^+e^-\gamma$. Events are selected by requiring exactly two oppositely charged tracks, both with good helix fits in the main drift chamber, and which are acollinear,

$$\frac{\vec{p}_1 \cdot \vec{p}_2}{p_1 p_2} > -0.999$$

(*i.e.*, the momentum are less than 177° apart. The event must also contain at least one neutral cluster with energy $E > 100$ MeV. One electron is tagged by imposing stringent identification requirements: $p > 0.7E_{\rm beam}$, $E_{\rm BSC} > 0.7p$, $|\texttt{XSE}| < 1.5$, and no hits in the muon tracking system. The other (recoil) electron has no selection criteria imposed.

Out of 3,370 candidate events in the J/ψ sample, 3,004 are single tag events and 366 are double tags (both tracks satisfied the electron tag criteria), giving a total of 3,736 tracks between 50 MeV/c and 2 GeV/c.

As Figure 7.1 shows, the radiative Bhabha sample has an excess of several hundred tracks at low (nearly minimum ionizing) shower counter energy, and momenta near the beam momentum. These tracks are concentrated in the detector end caps, $|\cos\theta| > 0.7$, (Figure 7.2). When the end cap data are excluded, the energy distribution is much cleaner (Figure 7.3).

7.1.2 Muons

A sample of muons was produced at IHEP by selecting cosmic ray candidates from the 1992 and 1993 D_s data runs [32]. Two-prong events were selected with one track having $p < 3$ GeV/c and hits in the muon tracking system; these candidates were then specially reconstructed with modifications to the BES reconstruction software to correct for t_0 distortions in the incoming "track" of each cosmic ray. Interaction events ($e^+e^- \rightarrow \mu^+\mu^-$) are removed from the sample by requiring $|t_{\rm TOF,1} - t_{\rm TOF,2}| > 4$ ns, because the software modifications would disrupt reconstruction of normal e^+e^- event

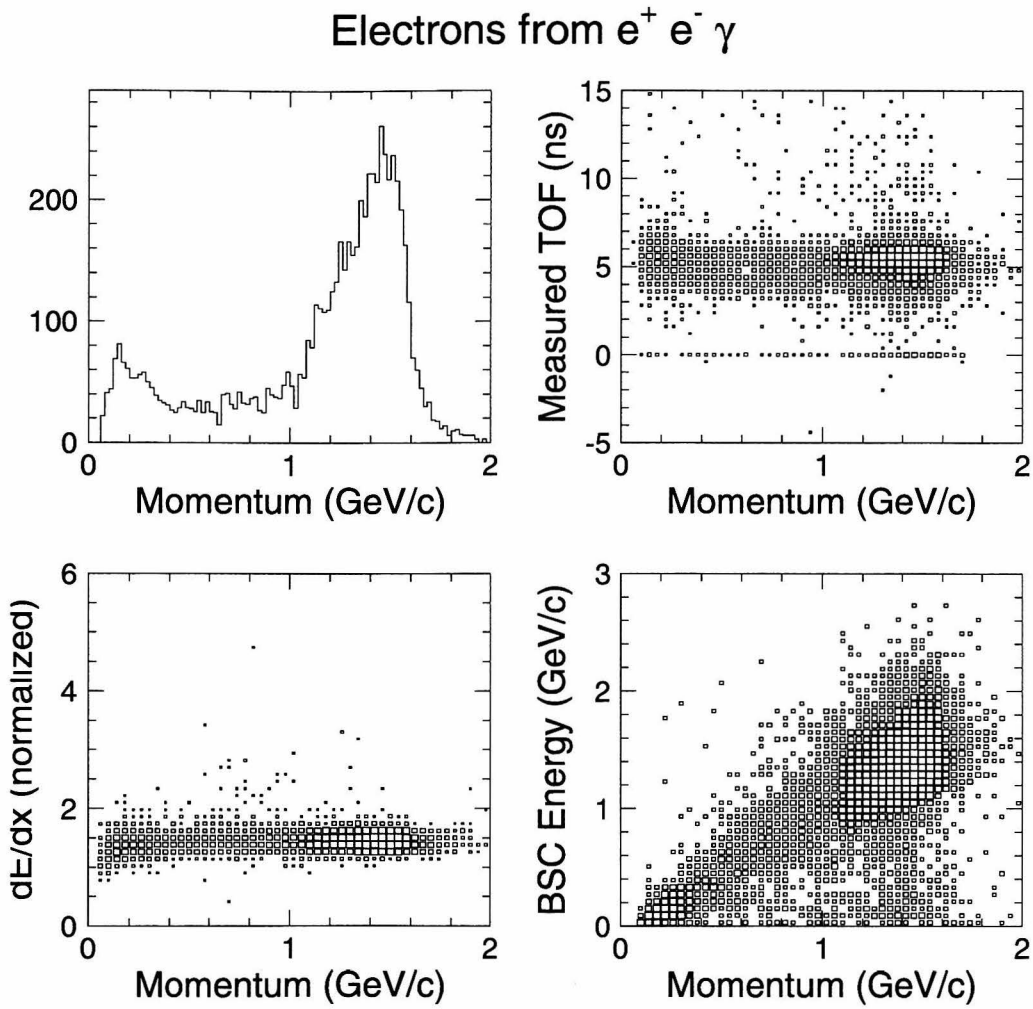

Figure 7.1: Reconstructed detector data from recoil electrons in radiative Bhabha selection.

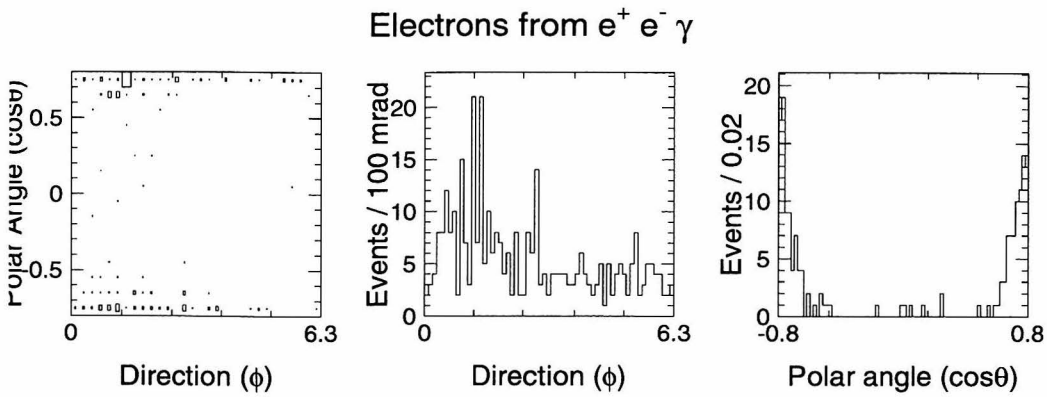

Figure 7.2: Distributions in ϕ and $\cos\theta$ of radiative Bhabha electrons with anomalous energy.

tracks. The final sample consists of 13 groups of runs, containing 21,409 events. The plot of dE/dx *vs.* momentum (Figure 7.4) indicates that a small background of protons from beam gas events is also present in this sample.

7.1.3 Pions

A sample of pions is selected using the decay chain

$$\begin{aligned} J/\psi \;&\rightarrow\; \omega(782)\,\pi^+\,\pi^- \quad \mathcal{B} = (7.0 \pm 0.7) \times 10^{-3} \\ &\quad\;\; \hookrightarrow \pi^+\,\pi^-\,\pi^0 \quad \mathcal{B} = (88.8 \pm 0.6)\,\% \end{aligned}$$

which has a net branching fraction of $(6.4 \pm 0.9) \times 10^{-3}$ [1]. The selection criteria used require only the existence of good charged tracks, or neutral calorimeter clusters; no data from any other detector system is used. A candidate event must contain exactly four charged tracks, two positive and two negative, and all with good helix fits in the main drift chamber. The event must also contain a minimum of two neutral energy clusters, but may contain more.

The neutral energy clusters are tested pair-wise for consistency with the decay $\pi^0 \rightarrow \gamma\gamma$, requiring $|m_{\gamma\gamma} - m_{\pi^0}| \leq 60$ MeV/c^2. Each π^0 candidate is matched with a π^+ and π^-, and the triplet is compared to an $\omega(782)$, requiring $|m_{\pi\pi\pi} - m_{\omega}| \leq 50$ MeV/c^2.

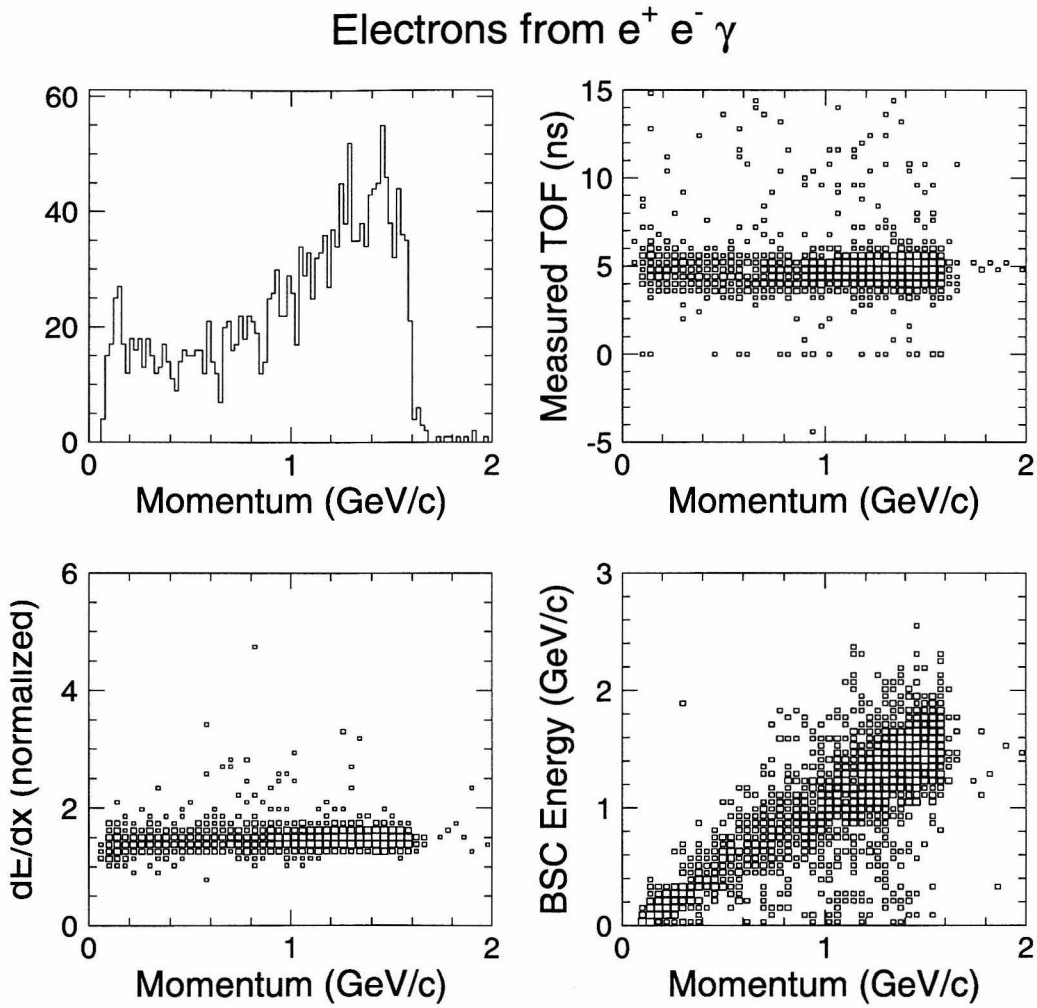

Figure 7.3: Recoil electrons from radiative Bhabha events, excluding detector end caps (*i.e.*, $|\cos\theta| < 0.7$).

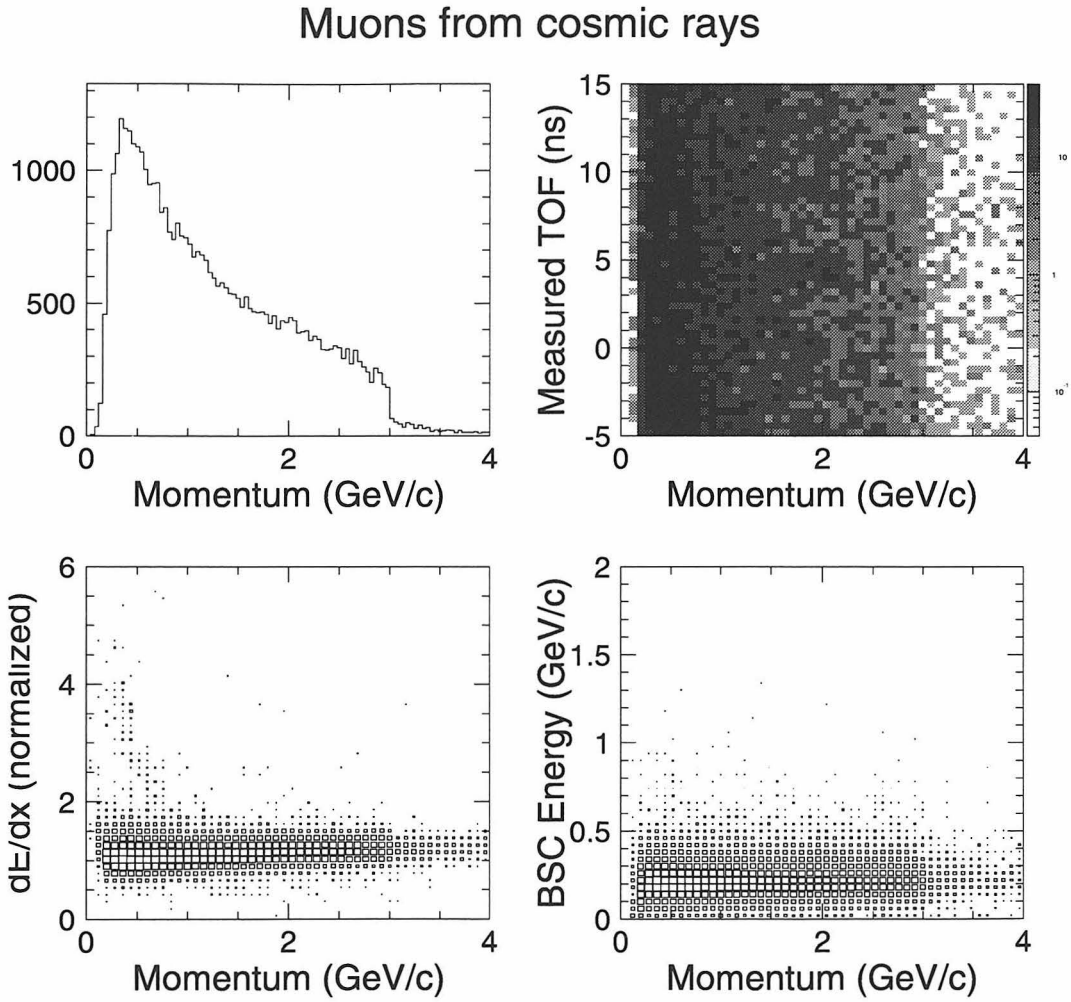

Figure 7.4: Cosmic ray muon sample from 1992 and 1993 D_s data.

Each successful candidate combination is kinematically fit, using the SQUID package, to the event topology above [33]. An event is accepted if at least one combination has $\chi^2 < 50$.

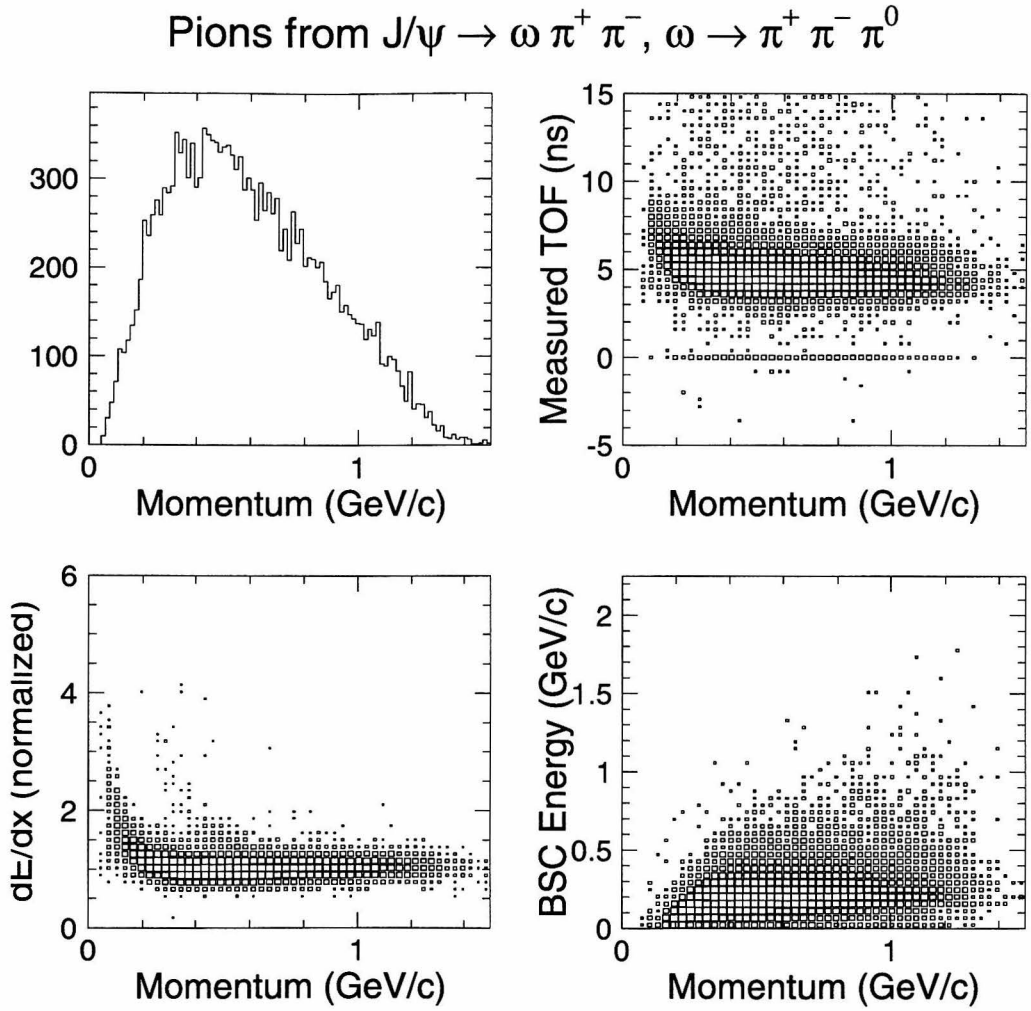

Figure 7.5: Pion sample selected kinematically from $J/\psi \to \omega(782)\pi^+\pi^-$.

In the sample of J/ψ data used, we find 4,200 candidates, or $(1.017\pm0.016)\times10^{-3}$ of the data. Considering the geometric acceptance of the detector, 85% for charged tracks and about 70% for neutral clusters in the barrel calorimeter, this sample is consistent with 100% efficiency for the decay channel. The distributions in Figure 7.5, especially dE/dx *vs.* momentum, reveal a small contamination from kaons in the sample, but

this is not significant enough to affect my analysis.

7.1.4 Kaons

A sample of kaons is selected using the decay chain

$$\begin{aligned} J/\psi \;\rightarrow\; & \phi(1020)\, K^+ K^- \quad \mathcal{B} = (7.4 \pm 1.1) \times 10^{-4} \\ & \hookrightarrow K^+ K^- \qquad\quad \mathcal{B} = (49.1 \pm 0.9)\,\% \end{aligned}$$

which has a net branching fraction of $(3.63 \pm 0.54) \times 10^{-4}$ [1]. This decay channel has a significant background from $J/\psi \rightarrow p\bar{p}\pi^+\pi^-$. Candidate events are selected by requiring exactly four well-fit charged tracks with total charge zero. No restriction on neutral energy is required.

Each candidate event is kinematically fit with the TELESIS package [34, 35] to the candidate signal and to the background final state, using all possible track combinations. For each combination, two tracks are constrained to the $\phi(1020)$ resonance, $m_\phi = 1019.4$ MeV/c^2 [1], and the full event is constrained to center-of-mass production with no missing momentum or energy. An event is accepted for the kaon sample if the kinematic fit of at least one combination has a probability $P(\chi^2_{\phi KK}) > 10^{-3}$, and no combination has $P(\chi^2_{p\bar{p}\pi\pi}) > 10^{-3}$.

From the J/ψ sample data, 96 events are selected which meet the criteria. Of these, 17 events contain at least one track which is inconsistent (from TOF and dE/dx information) with identification as a kaon at 99% CL. Eleven of those events contain two or more "non-kaon" tracks, implying a possible decay channel other than $\phi(1020)KK$. Two events contain a pair of tracks consistent with p and $\bar{p}$. Overall, the sample of 384 candidate tracks is 93% pure from inspection of TOF and dE/dx information.

Assuming 94 signal events, with a geometric acceptance of 85% for charged tracks, the net efficiency for ϕK^+K^- is about 12%. The momentum spectrum of kaons in $J/\psi \rightarrow \phi KK$ ranges from zero to about 1200 MeV/c; about 1% of that is below the 150 MeV/c threshold to enter the MDC. Measurements indicate that about 40% of

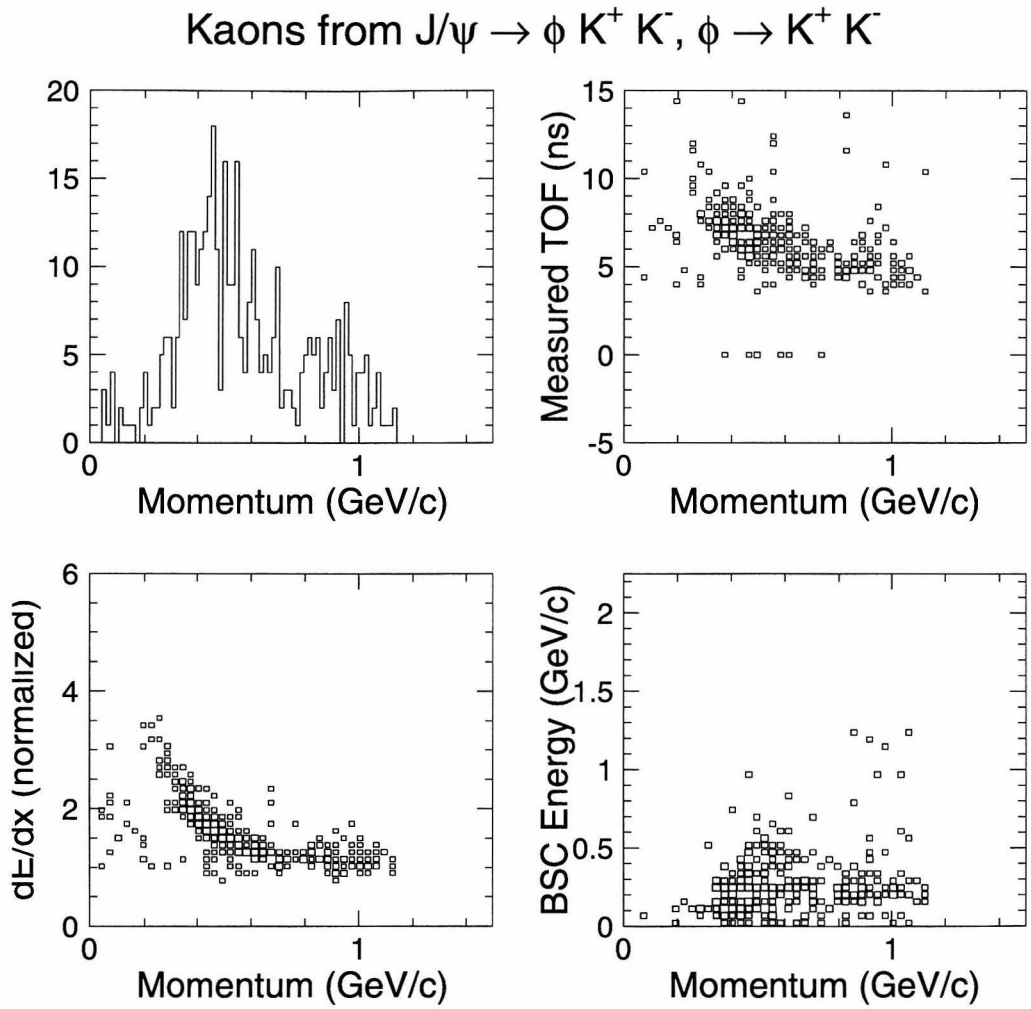

Figure 7.6: Kaon sample selected kinematically from $J/\psi \to \phi(1020)\, K^+\, K^-$.

kaons decay in flight. These two sources are sufficient (at 37%) to account for the observed efficiency.

7.2 The Likelihood Function

For particle identification, we may construct an overall likelihood as a product of terms for each detector system—the assumption is that a measurement by one set of electronics is independent of, and unaffected by, measurements made by other electronics. This assumption, while valid for major detector components, would not be valid for measurements of individual wires in the main drift chamber, for example, which are subject to crosstalk.

For the BES detector, I write a likelihood function for particle identification, involving the four main detector components, as

$$\begin{aligned} \mathcal{L}^h &= \mathcal{L}_{\mathrm{MDC}}(\chi^2, N_{\mathrm{hits}}) \cdot \mathcal{L}^h_{dE/dx}(dE/dx; p) \cdot \mathcal{L}^h_{\mathrm{TOF}}(t; p, L) \\ &\quad \times \mathcal{L}^h_{\mathrm{BSC}}(E; p) \cdot \mathcal{L}^h_{\mathrm{MUC}}(\mathrm{layers}; p, L) \end{aligned} \tag{7.1}$$

where $h \in \{e, \mu, \pi, K, p\}$ is the charged particle identification hypothesis being tested. The "best" hypothesis for h_0 for a given track is given by that with the maximum value of the likelihood function, equation 7.1,

$$\mathcal{L}^{h_0} > \mathcal{L}^h, \forall h \neq h_0 \in \{e, \mu, \pi, K, p\} \ .$$

To separate between two identification hypotheses, the method I shall use is the likelihood ratio, or normalized weight. Considering hypotheses a and b, let

$$W_{ab} = \frac{\mathcal{L}_a}{\mathcal{L}_a + \mathcal{L}_b} \tag{7.2}$$

which has the convenient property that $W_{ab} \to 1$ if hypothesis a is correct, and $W_{ab} \to 0$ if hypothesis b is correct. If the two hypotheses are indistinguishable, then $W_{ab} \sim 0.5$.

Hypotheses selected according to either maximum likelihood or normalized weights must be verified as consistent with the data. This is done by computing the confidence level for each likelihood using the prescription given in Chapter 6 (equation 6.1), and requiring that $P > P_{min}$, typically 0.01 (99% confidence level).

In the sections which follow, I derive expressions for each of the likelihood factors in equation 7.1, including empirical corrections where the observed detector distributions do not match theoretical expectations. For non-Gaussian distributions, I will also derive confidence level distributions useful for evaluating the consistency of the data with the identification hypotheses.

7.2.1 Charged Track Fitting

The BES track helix fitting routine, `DCFIND`, returns χ^2 and number of degrees of freedom for a five parameter fit. Besides being the best estimate for the selection of hits used for the track (and hence for the dE/dx calculation), the MDC fit χ^2 is also a good figure of merit for the consistency of track extrapolation to other subsystems with the track. Rather than attempting to calculate some other expression, the simplest solution is to use $\mathcal{L}_{\rm MDC} = \exp(\chi^2/2)$ as the likelihood, and $P(\chi^2_{\rm MDC}, n_{\rm hits} - 5)$ as a consistency check.

7.2.2 dE/dx in Main Drift Chamber

The BES main drift chamber (MDC) reports pulse heights for each wire hit in a track. These pulse heights are averaged (with an 80% truncated mean) to give an estimate for the ionization energy loss in the MDC gas. With 40 wire layers, and a "good" track required to have at least 24 hits, the BES dE/dx system has a resolution of 8.5% for ultrarelativistic particles (Bhabha electrons), and approximately 12% for low momentum hadrons. The dE/dx reconstruction routine calculates a normalized residual of the measured pulse height for each of the five charged particle hypotheses

$$\texttt{XS}h = \frac{dE/dx - dE/dx_h(p)}{\sigma_{dE/dx}}, \quad h \in \{e, \mu, \pi, K, p\} \; . \tag{7.3}$$

This quantity is used directly in the likelihood expression,

$$\mathcal{L}^h_{dE/dx}(dE/dx;p) = \exp\left(-\frac{\texttt{XS}h^2}{2}\right) . \tag{7.4}$$

During the 1993 D_s data run, a number of BES collaborators found that the model used for the expected dE/dx_h had systematic deviations. Efforts were made to re-calculate the normalized residuals for already reconstructed data, and to correct the model in the reconstruction software. These efforts have been largely successful, and it is expected that future reconstruction will produce well-distributed values of $\texttt{XS}h$.

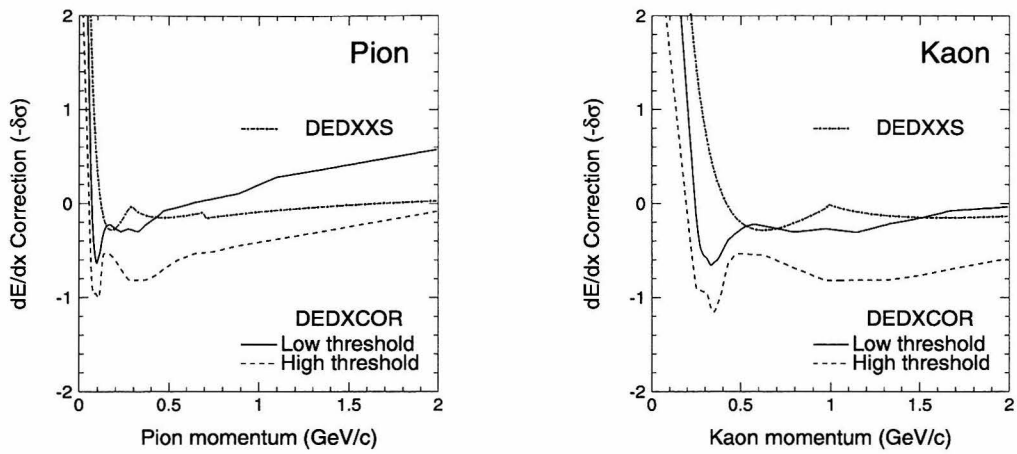

Figure 7.7: Empirical corrections to dE/dx reconstruction in D_s data. These corrections are applied as subtractions to the reported $\texttt{XS}h$ values. Corrections are shown for pions *(left)* and kaons *(right)* as a function of momentum, for both "high threshold" (dashed curves, runs 3097–3450) and "low threshold" (solid curves, runs 3451*ff*) data. Also shown (beaded curves) is the effect of recalculation based on the 1991 tau data.

For analysis of older data, routines are available to recompute the five normalized residuals. Corrections for data from the tau mass run [36] use a ten-parameter fit to the measured dE/dx distributions, including adjustments to both momentum ($\beta\gamma$) and pulse height (dE/dx). For the 1992, 1993, and 1994 D_s data sets, the corrections [37] are binned in $\beta\gamma$, and directly compute an offset in the $\texttt{XS}h$ variables for hadronic identification hypotheses. These corrections are shown in Figure 7.7.

7.2.3 Time of Flight

The time of flight counters have an essentially Gaussian distribution of measured time about the true time-of-flight of a given particle. The likelihood factor is therefore quite simple,

$$\mathcal{L}^h_{\mathrm{TOF}}(t;p,L) = \exp\left(-\frac{1}{2}\left(\frac{t_{\mathrm{TOF}} - t_h(p,L)}{\sigma_t}\right)^2\right), \quad \sigma_t \approx 400 \text{ ps} . \tag{7.5}$$

For any particle, the expected time of flight t_h from production to detection should be given by a simple expression involving the path length of the particle's trajectory and its momentum:

$$t_h(p,L) = \frac{L}{pc}\sqrt{p^2 + m_h^2} . \tag{7.6}$$

In the BES detector, Bhabha electrons are used to calibrate the time of flight detector system. A comparison of TOF measurements with the expected $t_h(p,L)$ values for hadrons and muons reveals a systematic deviation of unknown origin. A set of empirical corrections [38] have been developed to offset the expected flight time as a function of momentum for each particle identification hypothesis. The effect of these corrections is shown in Figure 7.8.

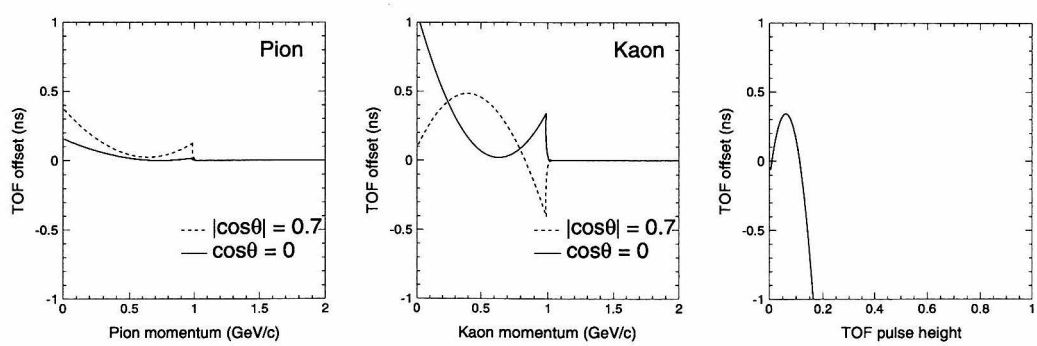

Figure 7.8: Time of flight deviations from nominal values for pions *(left)* and kaons *(center)* as a function of momentum and polar angle. The curves indicate the deviation at $\cos\theta = 0$ (solid) and $|\cos\theta| = 0.7$ (dashed). The right plot shows an additional deviation with respect to the measured pulse height for pions.

Additionally, measurements of the time of flight resolution for each counter have been made for the 1992 and 1993 data sets. These resolutions, shown in Figure 7.9, are often significantly different than the "resolution" reported by the reconstruction software. The resolution is parameterized as a function of polar angle, $\sigma = A + B\cos^2\theta$, with a counter-dependent scale. An overall scaling factor of 0.9 is applied to the 1992 D_s data.

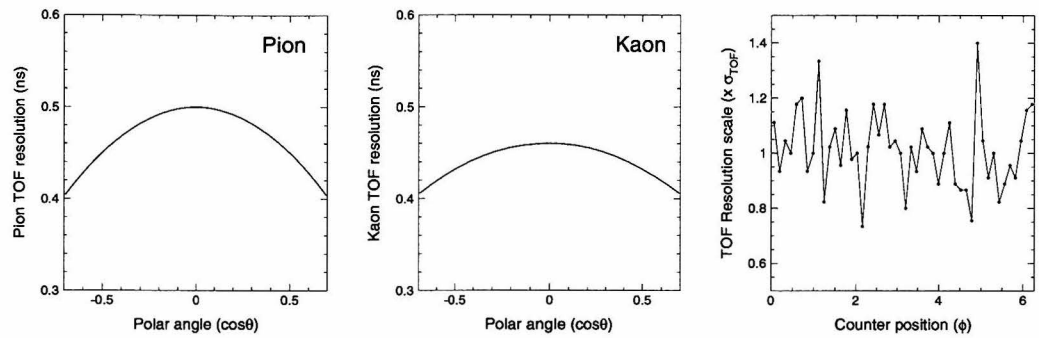

Figure 7.9: Time of flight measured resolutions for pions *(left)* and kaons *(center)*, as a function of polar angle. At right is shown the counter-by-counter scaling factors for the barrel TOF, plotted versus the ϕ position of the counter. A scaling factor of 1 corresponds to $\sigma = 450$ ps.

Using both sets of empirical adjustments, equation 7.5 becomes

$$\mathcal{L}^h_{\rm TOF}(p,L) = \exp\left(-\frac{1}{2}\left(\frac{t - t_h(p,L) - t_{{\rm corr},h}(p,L;{\rm run})}{\sigma_{t,h}(p;\phi_{\rm TOF},{\rm run})}\right)^2\right) . \qquad (7.7)$$

7.2.4 Barrel Shower Counter

The BES shower counter (BSC and ESC) is a sampling calorimeter, optimized for detecting electromagnetic showers from electrons and photons. The detector calibration and reconstruction software are designed with this in mind, with the result that the shower counter response to hadrons is fundamentally different from that to electrons. Electrons (and photons) deposit their energy in the BSC through electromagnetic showers in the lead between proportional tube layers, and the shower secondaries produce ionization in the proportional tubes. This response is Gaussian, with $E_{\rm SC} = p$

and $\sigma_{EM}(E) \approx 0.23\sqrt{E}$.

Most muons do not interact in the BSC except by depositing ionization energy in the gas of the proportional tubes. Since the tubes are operated in limited streamer mode, no pulse height information is available; muons produce the same response at all momenta, corresponding to the energy of one hit per layer (about 200 MeV). This produces a Landau response with approximately fixed peak position and width (Figure 7.10).

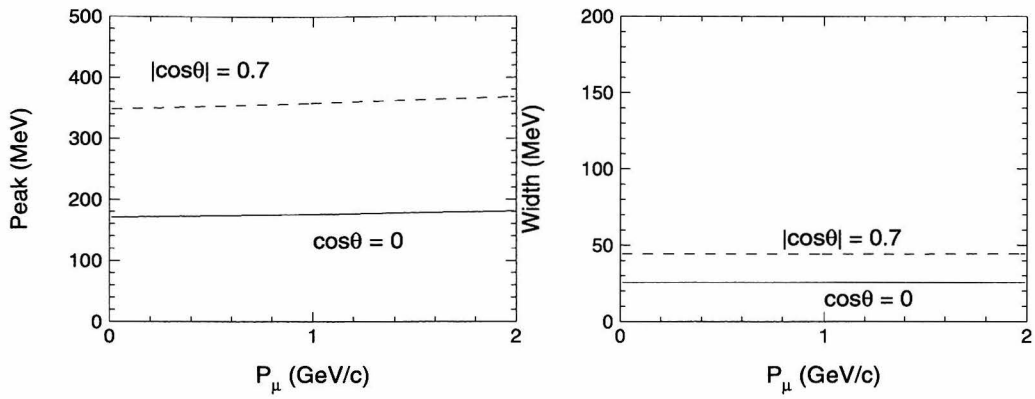

Figure 7.10: Parameterization of Landau response of BSC to muons, as functions of momentum. *(Left)* Most probable energy response. *(Right)* Width of Landau peak.

Hadrons interact in the BSC quite differently, with a strong dependence on momentum, charge and hadron species, especially at lower momenta ($\lesssim$ 500 MeV/c). The dominant interaction is ionization in the proportional tubes, giving an apparent energy deposition distributed according to the Landau function [39, 40]. This is modified by the large cross section for hadronic interactions at low momentum, which cause range-out in the BSC and a lower reconstructed energy (fewer tubes hit); for kaons, this effect also leads to a significant broadening of the Landau response around 500 MeV/c. Figures 7.11–7.13 show the peak energy and width parameterizations determined from data.

These parameterizations are empirical, piecewise-smooth functions which approximate the observed data distributions reasonably well. I have not attempted to in-

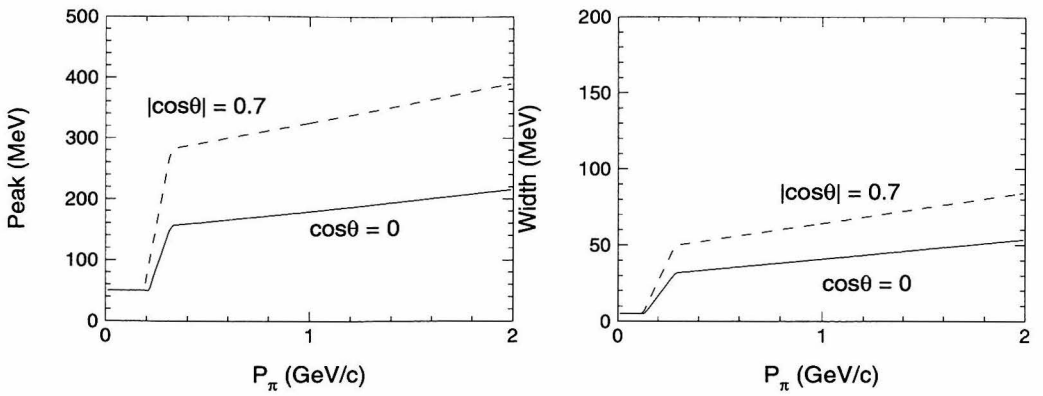

Figure 7.11: Parameterization of Landau response of BSC to pions, as functions of momentum. *(Left)* Most probable energy response. *(Right)* Width of Landau peak.

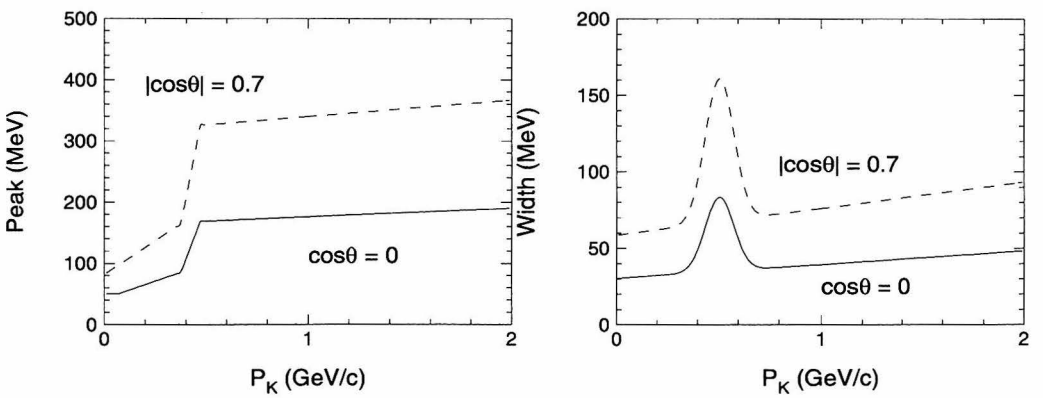

Figure 7.12: Parameterization of Landau response of BSC to kaons, as functions of momentum. *(Left)* Most probable energy response. *(Right)* Width of Landau peak.

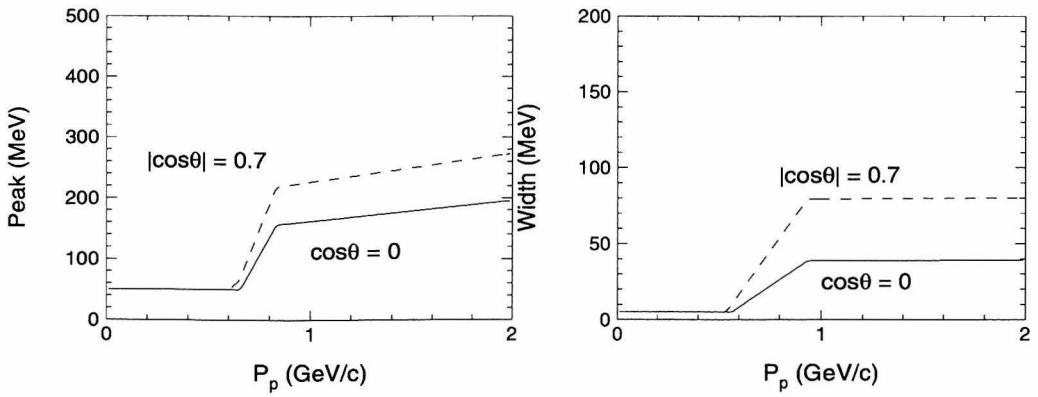

Figure 7.13: Parameterization of Landau response of BSC to protons, as functions of momentum. *(Left)* Most probable energy response. *(Right)* Width of Landau peak.

corporate the large variety of physical effects which contribute to these distributions; that task is better suited to Monte Carlo development.

Combing the different responses of the shower counter, I write the likelihood factor as

$$\mathcal{L}^h_{\rm BSC}(E;p) = \begin{cases} \exp\left(-\frac{1}{2}\left(\frac{E-p}{\sigma_{EM}(p)}\right)^2\right) & h = e \\ {\rm Lan}\left(\frac{E-E_{0,h}(p)}{\sigma_{E,h}(p)}\right) & h \in \{\mu\,,\,\pi\,,\,K\,,\,p\} \end{cases} \tag{7.8}$$

where ${\rm Lan}(x)$ is the Landau distribution.

7.2.5 Muon Tracking

The most natural likelihood expression for the muon tracking system is the probability of observing or not observing a hit at each layer along the track's projection from the drift chamber through the layers of steel flux return.

$$\mathcal{L}^h_{\rm MUC}({\rm layers};p) = \prod_i^{\rm hits} P^h_{\rm range}(\lambda_i|p) \cdot \prod_j^{\rm nohits} \left(1 - \varepsilon P^h_{\rm range}(\lambda_j|p)\right)\,, \tag{7.9}$$

where p is the particle's momentum as measured in the main drift chamber, and λ_i is the total amount of material, expressed in equivalent g cm^{-2} of steel, traversed by the

particle from the IP up to the ith layer. The hit probability, P^h_{range}, is expected to be a simple sigmoid, with the range (50% probability) determined by integrating ionization energy loss along the particle's path through material [41]. Fluctuations in the actual energy loss at each interaction, multiple Coulomb scattering, and other effects smear the endpoint of what would otherwise be a step function at a depth $\lambda = R_h(p)$,

$$P^h_{\text{range}}(\lambda; p) = \frac{1}{1 + \exp \frac{\lambda - R_h(p)}{S_h(p)}} \ . \tag{7.10}$$

P^h_{range}, scaled by the muon tracking system efficiency ε, is shown schematically in Figure 7.14.

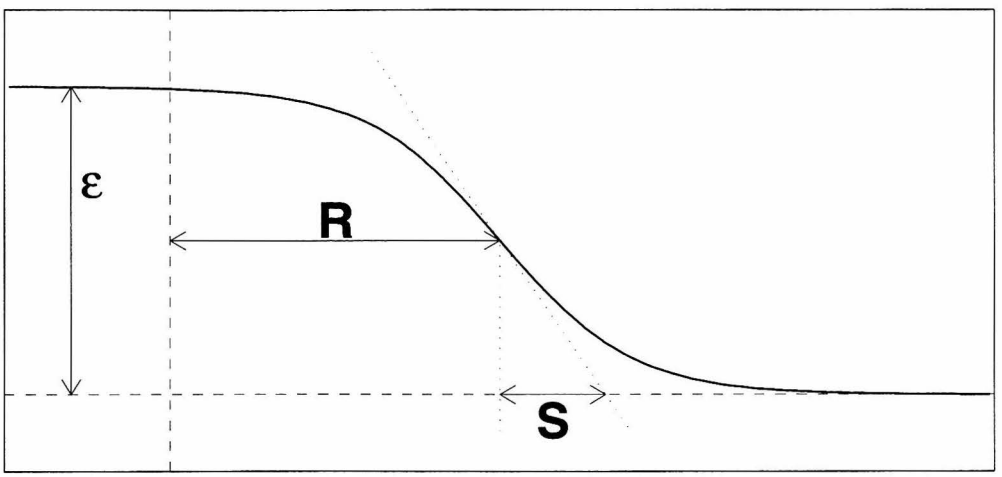

Figure 7.14: General functional form of sigmoid curve used to parameterize penetration probability in muon tracking system, $P(\lambda) = \varepsilon/[1 + \exp(\frac{\lambda - R}{S})]$, where ε is efficiency.

I determine the range $R_h(p)$ and straggling parameters $S_h(p)$ experimentally using pure samples of cosmic ray muons and pions from $J/\psi \to \omega\pi\pi$, by fitting the observed rates of hits at a given momentum to equation 7.10. The results are shown in Figures 7.15 and 7.16.

The muon tracking system is unique in my particle identification system, in that the consistency check (confidence level) variable is *not* the integral of the likelihood

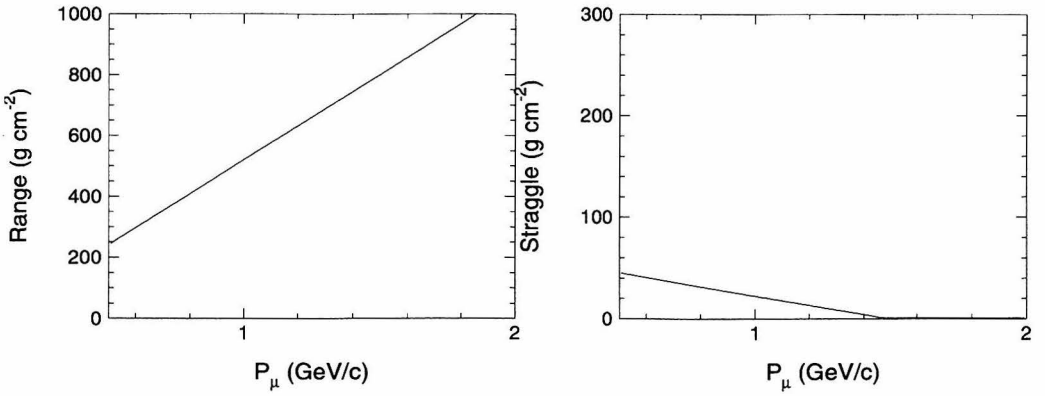

Figure 7.15: Range *(left)* and straggling *(right)* as functions of initial (drift chamber) momentum for muons in the BES muon tracking system.

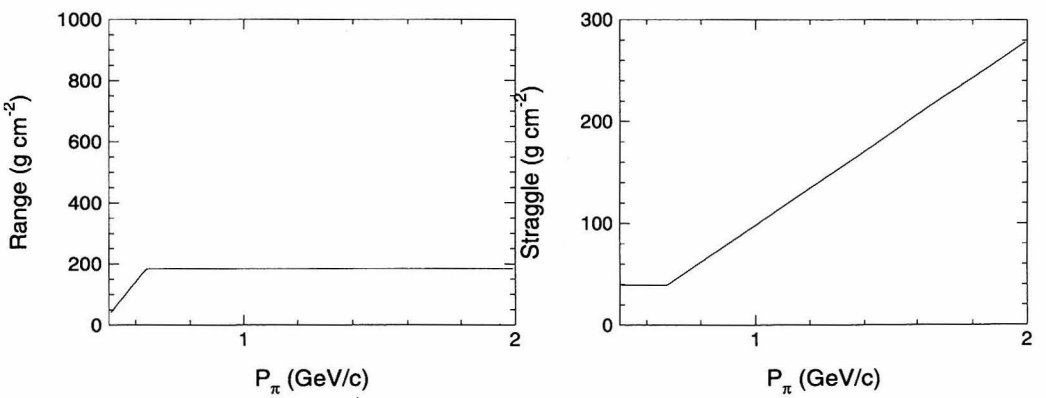

Figure 7.16: Range *(left)* and straggling *(right)* as functions of initial (drift chamber) momentum for pions in the BES muon tracking system.

expression, equation 7.9. Instead, I determine consistency from the alignment of the observed hits in each layer, compared to the projection of the reconstructed charged track through the layers. Since the muon tracking system consists of flat steel plates, the x and y coordinates are necessarily correlated; I find it convenient to convert to pseudo-cylindrical coordinates (l, ϕ, z), with l the layer index for the hit and $\phi = \tan^{-1}(y/x)$.

From cosmic ray muon data, the residuals

$$\Delta\phi = \phi_{\text{hit}} - \phi_{\text{proj}} \quad ; \quad \Delta z = z_{\text{hit}} - z_{\text{proj}} \tag{7.11}$$

are essentially Gaussian, with non-Gaussian tails distributed symmetrically about each peak (Figure 7.17). The broadness of the Δz distributions in each layer ($\gtrsim$ 10 cm) and the higher tails make them unsuitable for evaluating hit association with high efficiency. The $\Delta\phi$ distributions are quite narrow, 20 mrad corresponding to $\sigma_{xy} \sim 5$ cm position alignment. The Δz distributions are extremely broad ($\sigma \sim$ 10 cm), and offset significantly from the expected average zero; for these reasons, and because many reported hits do not include z information at all, I do not include the Δz alignment as part of the muon tracking system consistency check.

I fit the distributions of $\Delta\phi$ in each layer to a family of Gaussians as a function of momentum, and determine the peak offset ($\langle\Delta\phi\rangle$) and width ($\sigma_{\Delta\phi}$) (Figure 7.18) as functions of momentum. The systematic offsets are an indication that the algorithm used to project tracks from the main drift chamber through the muon system may not be accurate, but I can correct for this with the empirical parameters determined here. The distributions of the parameters are fit to simple functions of momentum,

$$\langle\Delta\phi\rangle(p) = A + Bp \quad ; \quad \sigma_{\Delta\phi}(p) = C + \frac{D}{p - E}$$

(Figure 7.18) which results in the functions shown in Table 7.1. From these paramet-

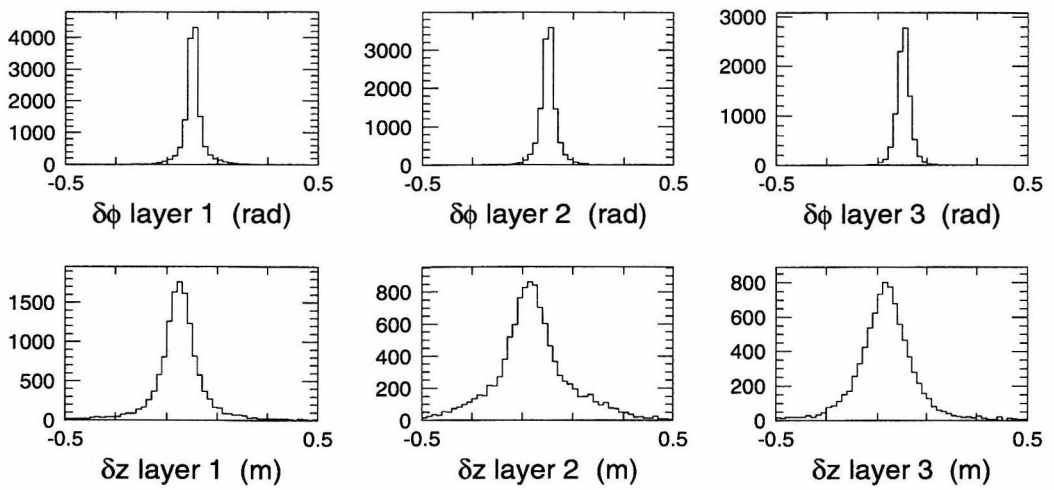

Figure 7.17: Muon tracking residuals in ϕ *(top)* and z *(bottom)* for each layer of the muon tracking system.

ers, I normalize the residuals $\Delta\phi_i$ as

$$\widehat{\delta\phi_i} = \frac{\Delta\phi_i - \langle\Delta\phi\rangle_i(p)}{\sigma_i(\Delta\phi)(p)} \ .$$

	$\langle\Delta\phi\rangle(p) =$ (mrad)	$\sigma_{\Delta\phi}(p) =$ (mrad)
Layer 1	$1.390 - 0.252\,p$	$4.253 + 17.66/(p - 0.370)$
Layer 2	$0.529 + 0.078\,p$	$3.192 + 25.38/(p - 0.332)$
Layer 3	$4.707 - 0.445\,p$	$4.426 + 25.03/(p - 0.409)$

Table 7.1: Parameterizations of offset and width for ϕ residuals in the muon tracking system, as functions of momentum.

Since the $\widehat{\delta\phi}$ are normalized Gaussian variables, the natural expression to evaluate the hit association would be $\chi^2 = \sum \widehat{\delta\phi_i}^2$. However, the muon tracking residuals are strongly correlated between layers, and a proper χ^2 must account for the correlations. For a set $\vec{x}$ of n correlated Gaussian variables, the χ^2 such that $P(\chi^2)$ is the probability that $\vec{x}$ lies outside an n-ellipsoid centered on the true value $\vec{x}_0$, may be computed via a transformation involving the covariance matrix $V = \langle x_i x_j \rangle$ [1]. Specifically,

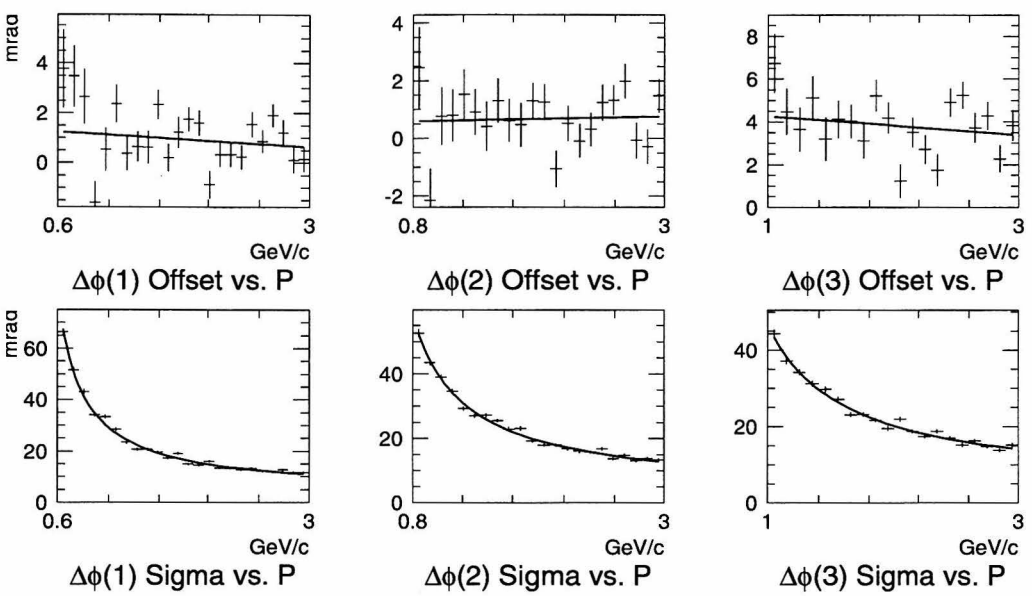

Figure 7.18: Distribution of hit association offsets and widths *vs.* momentum for each layer of the muon tracking system. *(Top)* Best fit offsets, with linear parameterization; *(bottom)* best fit widths, with $1/p$ parameterization. Fit results for each layer are listed in Table 7.1.

$\chi^2 = \vec{x}^T V^{-1} \vec{x}$ has the desired statistical properties for n degrees of freedom.

From cosmic ray muon data, the covariance matrix, $V_{\mathrm{MUC},ij} = \langle \widehat{\delta\phi}_i \cdot \widehat{\delta\phi}_j \rangle$ (where i and j index the layers), is

$$V_{\mathrm{MUC}} = \begin{pmatrix} 0.89968 & 0.78028 & 0.75376 \\ 0.78028 & 0.80974 & 0.76471 \\ 0.75376 & 0.76471 & 0.87911 \end{pmatrix} . \tag{7.12}$$

I compute the inverse matrix as well as the three 2×2 submatrices (for the cases where one layer does not have a hit) from equation 7.12, and define

$$\chi^2_{\mathrm{MUC}} = \widehat{\delta\phi}^{\,i}\, V^{-1}_{\mathrm{MUC},i}{}^{j}\, \widehat{\delta\phi}_j \ , \tag{7.13}$$

summed over the observed hits, and $P(\chi^2_{\mathrm{MUC}}; n_{hits})$ gives the probability that the hits are correctly associated with the projected track.

7.2.6 Summary

With all of the detector information described above, my particle identification likelihood function, equation 7.1, may be written out in full as

$$\begin{aligned} \mathcal{L}^h \;=\; & \exp\left[-\frac{1}{2}\left(\chi^2_{\mathrm{MDC}} + \left(\tfrac{dE/dx - dE/dx_h(p)}{\sigma_{dE/dx}}\right)^2 + \left(\tfrac{t - t_h(p,L) - t_{\mathrm{corr},h}(p,L;\mathrm{run})}{\sigma_{t,h}(p;\phi_{\mathrm{TOF}},\mathrm{run})}\right)^2\right)\right] \\ & \times \begin{cases} \exp\left(-\frac{1}{2}\left(\frac{E-p}{\sigma_{EM}(p)}\right)^2\right) & h = e \\ \mathrm{Lan}\left(\frac{E - E_{0,h}(p)}{\sigma_{E,h}(p)}\right) & h \in \{\mu\,,\,\pi\,,\,K\,,\,p\} \end{cases} \\ & \times \prod_i^{\text{MUC hits}} \frac{1}{1 + \exp[(\lambda_i - R_h(p))/S_h(p)]} \\ & \times \prod_j^{\text{MUC no-hits}} \left(1 - \frac{\varepsilon_{\mathrm{MUC}}}{1 + \exp[(\lambda_j - R_h(p))/S_h(p)]}\right) \end{aligned} \tag{7.14}$$

where the various parameterized functions are described in the preceding sections. Using equation 6.1 I can convert this likelihood to a probability, or confidence level,

to check the consistency of a given track with each particle identification hypothesis. If the likelihood functions are properly constructed, then the confidence level distribution should be uniform between 0 and 1 for the true hypothesis. Figure 7.19 shows the confidence level distributions obtained from each of my kinematically selected particle samples. They are all reasonably flat, within the statistical fluctuations, and show that this method of identifying particles is useful.

7.3 Efficiency and Misidentification

There are several ways to apply the likelihood function of equation 7.1 for particle identification: simple consistency using a minimum confidence level (*e.g.*, 90% or 99%); maximum likelihood value among all or several identification hypotheses; a likelihood ratio to distinguish between two hypotheses. In my analyses for this thesis (see Chapters 8 and 9), I implicitly use a consistency requirement to obtain maximum efficiency. This method has the worst background rejection, since it takes no account of the relative probabilities of the different hypotheses. Because I incorporate the particle identification likelihood function as one factor among several in a global analysis, I do not need to preselect identified tracks with high purity.

In this section, I present efficiency and misidentification rates for the particle ID likelihood function as a way of evaluating the utility of the technique. The identity of each particle is assigned as the maximum likelihood value combining dE/dx, time of flight, and the muon tracking system. I also require 99% confidence level consistency with the selected hypothesis. Figures 7.20 through 7.23 show the fraction of tracks in each momentum bin which are identified as indicated, in each of the kinematically selected track samples. For the sample of cosmic ray muons, the likelihood function does not include information from the time-of-flight (TOF) detectors; as Figure 7.4 shows, the time of flight is uniformly distributed across the entire time window ($\sim \pm 50$ ns) of the trigger system.

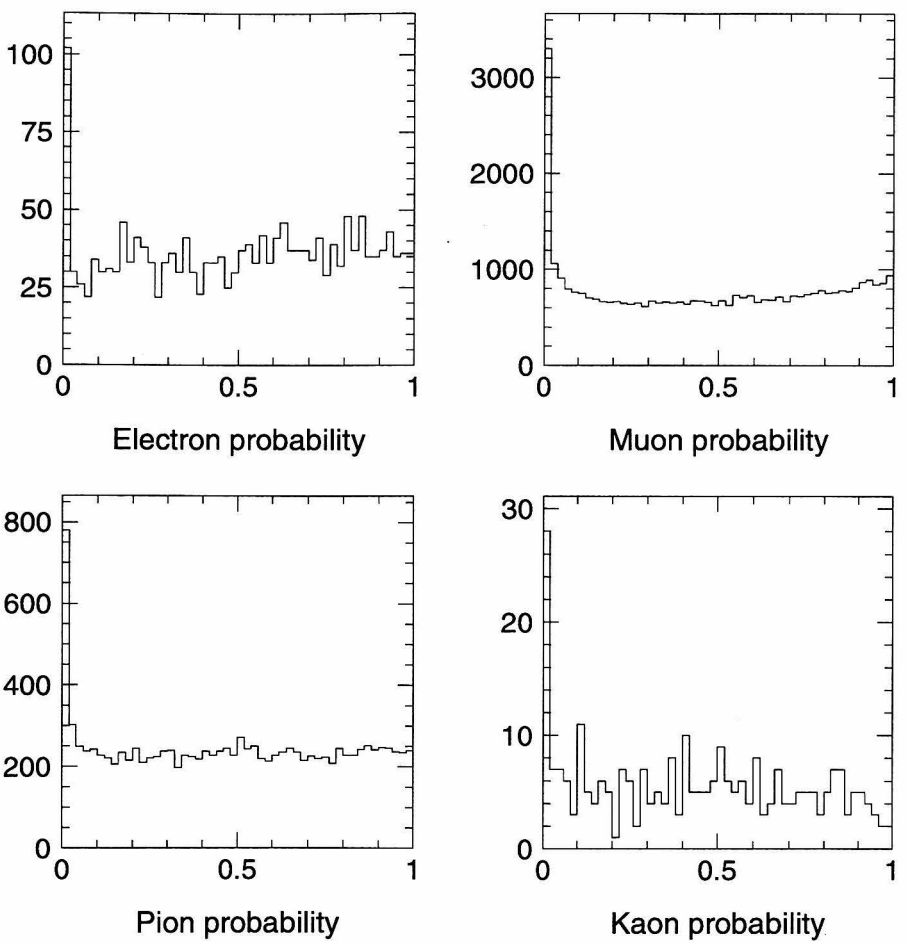

Figure 7.19: Probability distributions for identification of particles. Each distribution is obtained using the kinematically selected sample of the indicated particle type, evaluated for the "correct" hypothesis (e, μ, π, K). The spike at zero is due to particles for which the detector subsystems did not report valid information (*e.g.*, an end cap track with $|\cos\theta| > 0.7$, bad time of flight matching, etc.).

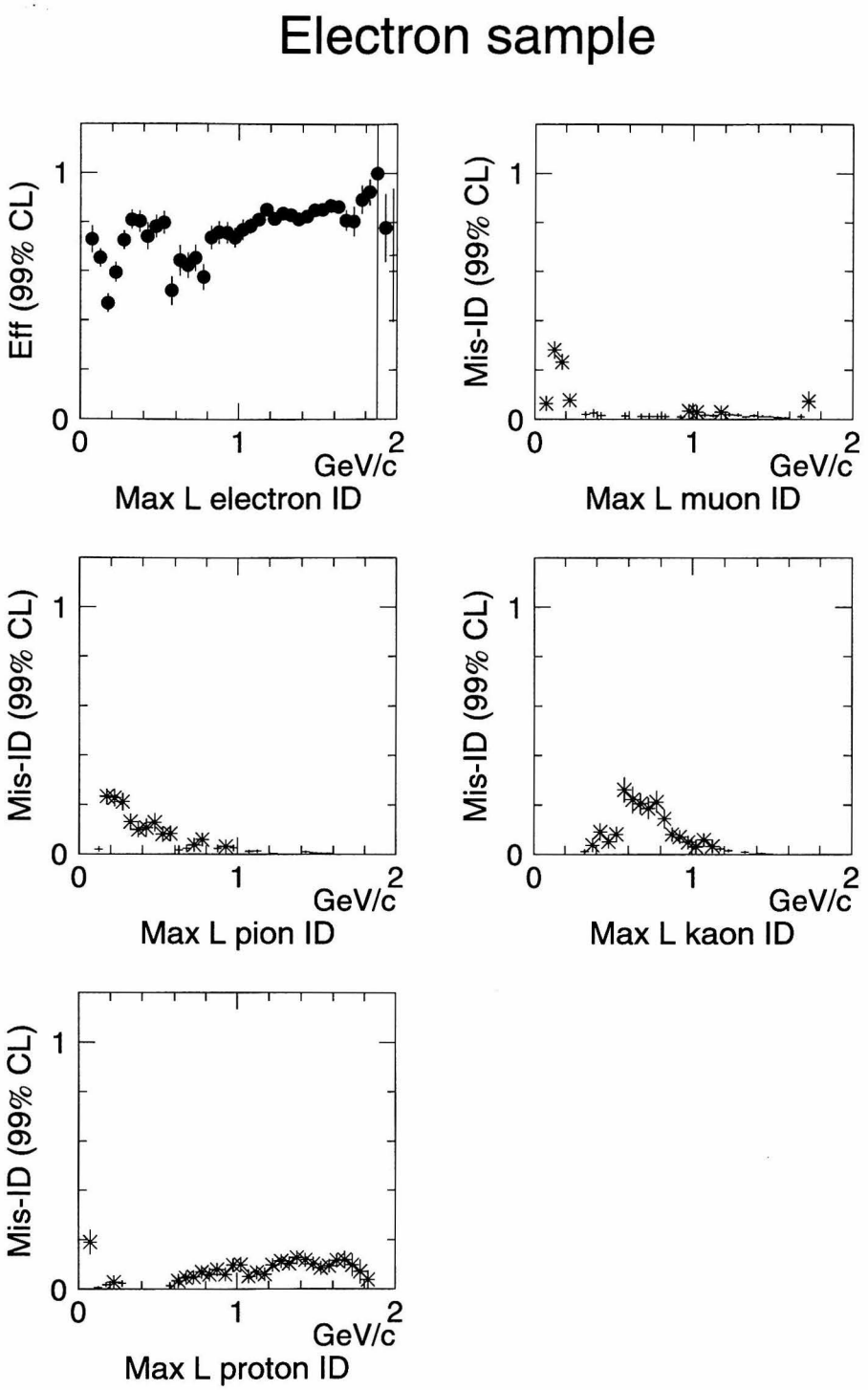

Figure 7.20: Efficiencies and misidentification rates of electrons using the likelihood technique described in the text. The five plots show the fraction of known (radiative Bhabha) electron tracks identified as electrons, muons, pions, kaons, and protons, respectively.

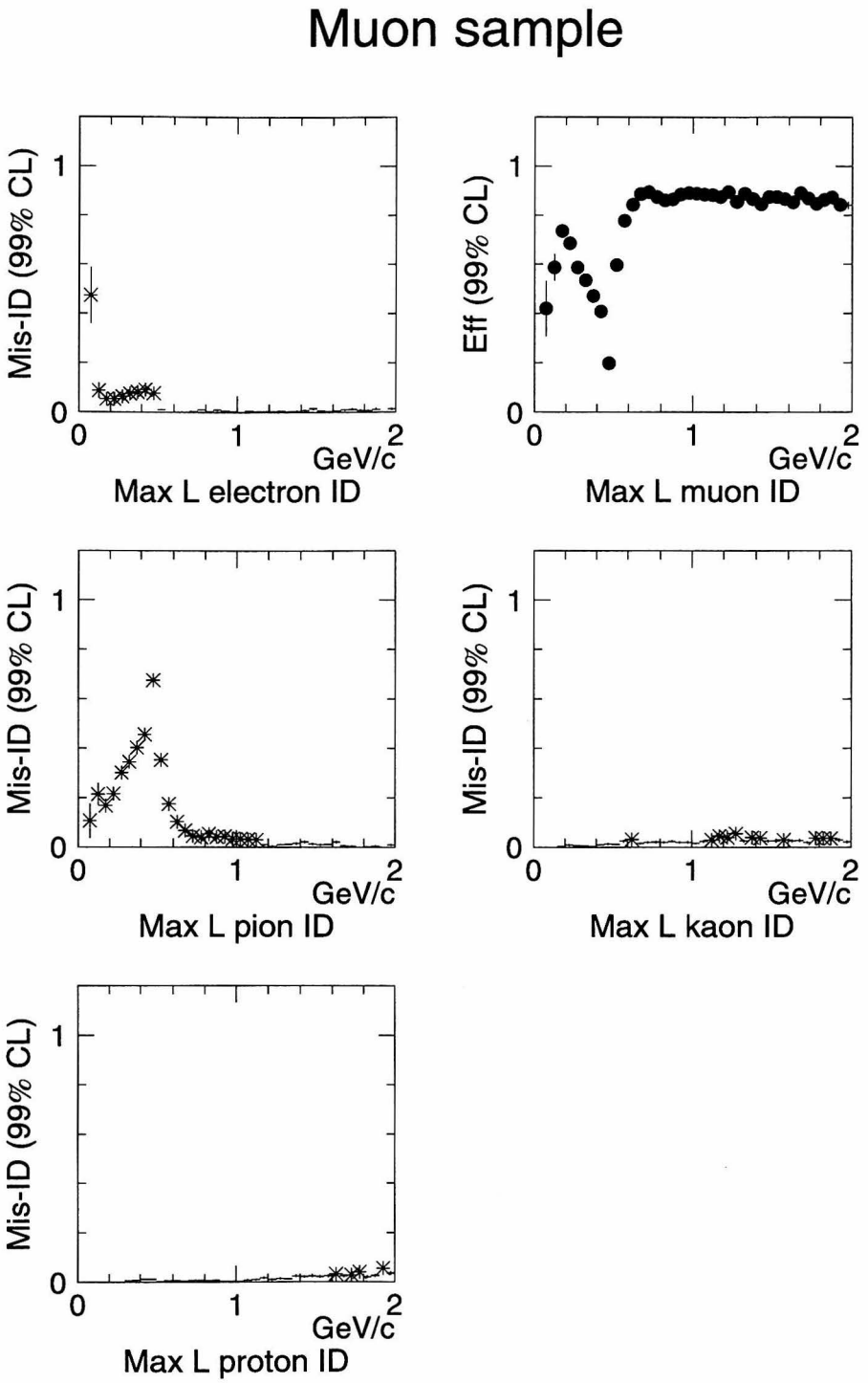

Figure 7.21: Efficiencies and misidentification rates of muons using the likelihood technique described in the text, excluding the time of flight (TOF) detector. The five plots show the fraction of known (cosmic ray) muon tracks identified as electrons, muons, pions, kaons, and protons, respectively.

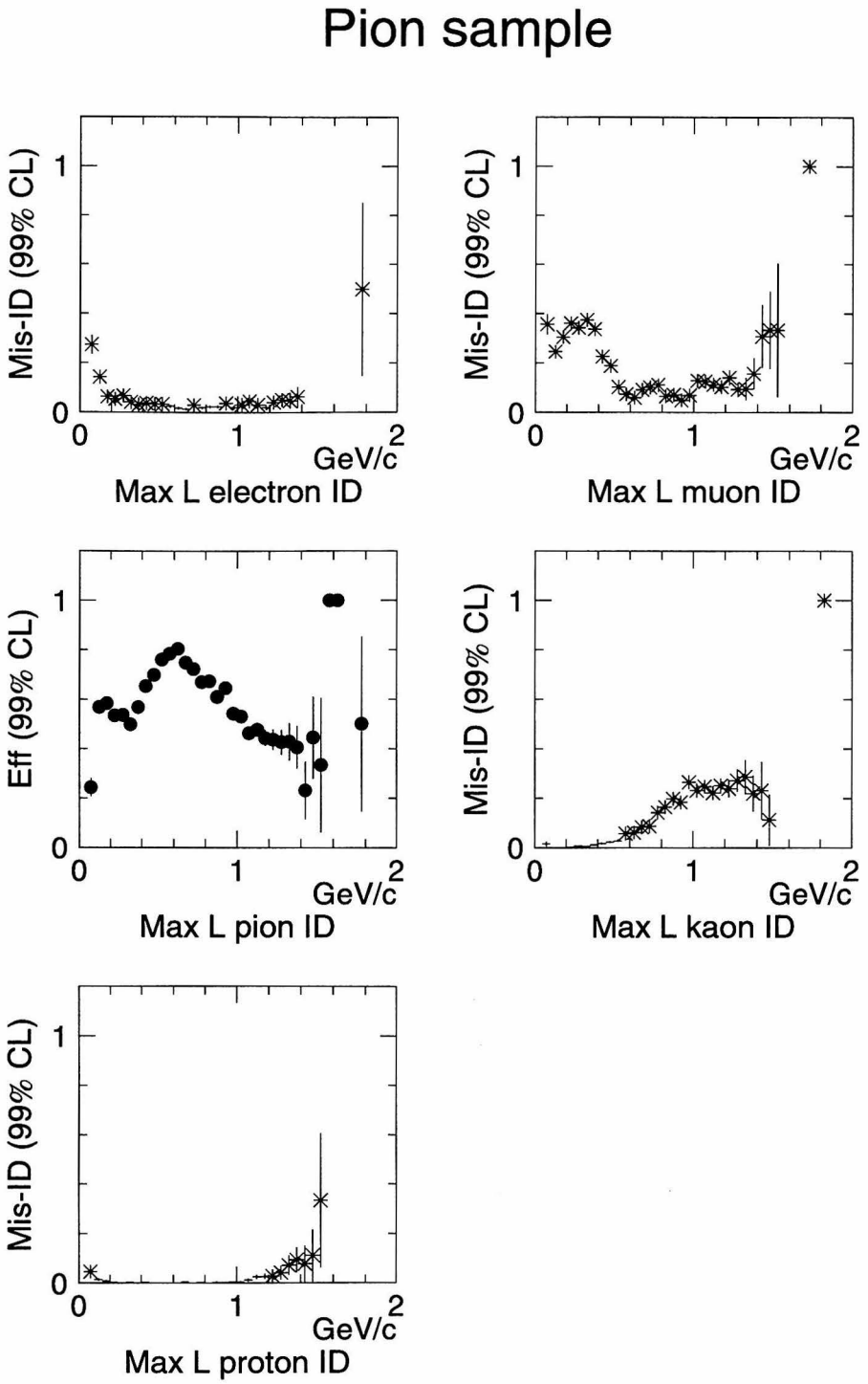

Figure 7.22: Efficiencies and misidentification rates of pions using the likelihood technique described in the text. The five plots show the fraction of known ($J/\psi \rightarrow \omega\pi^+\pi^-$) pion tracks identified as electrons, muons, pions, kaons, and protons, respectively.

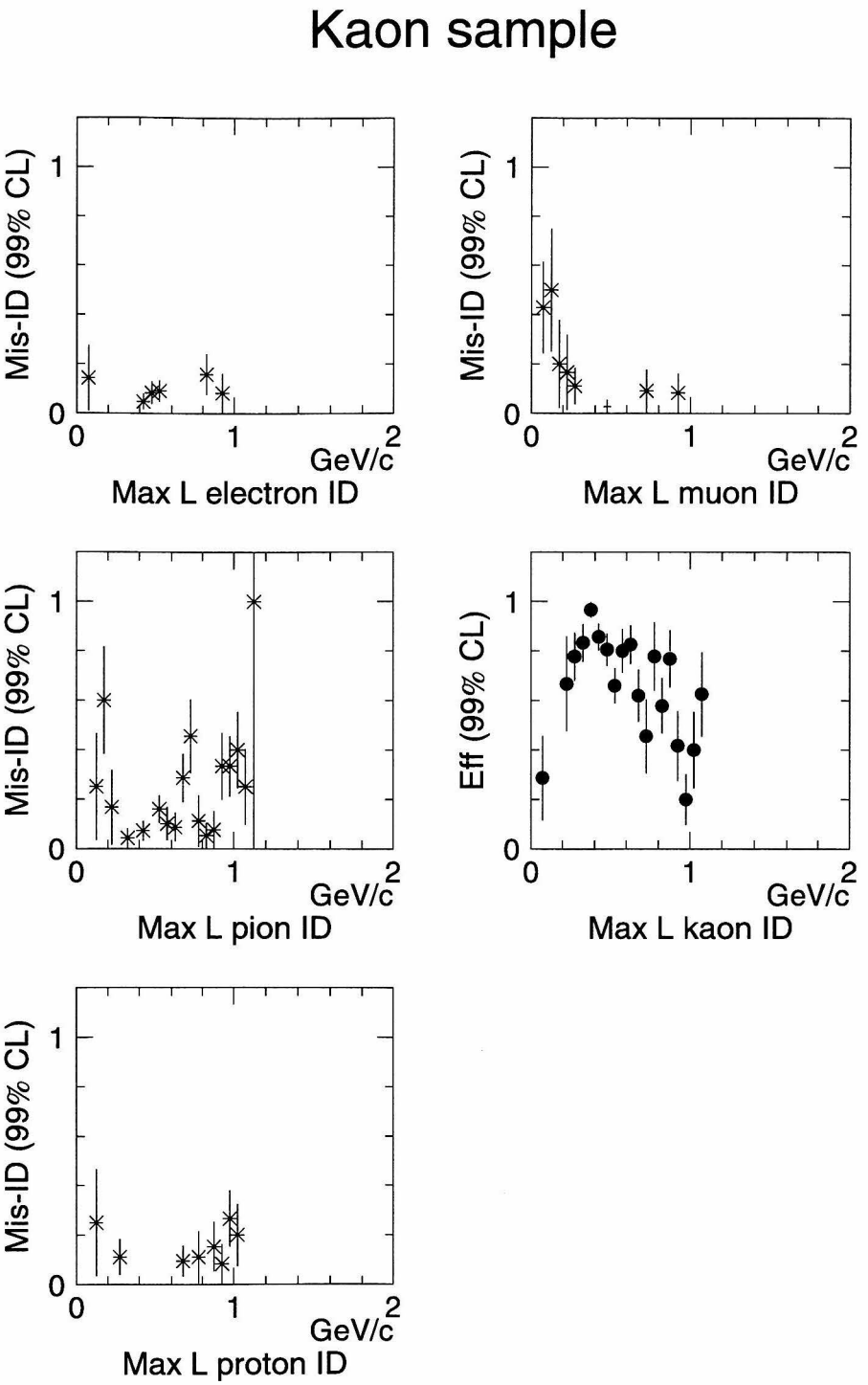

Figure 7.23: Efficiencies and misidentification rates of kaons using the likelihood technique described in the text. The five plots show the fraction of known ($J/\psi \to \phi K^+K^-$) kaon tracks identified as electrons, muons, pions, kaons, and protons, respectively.

Chapter 8 Selection of D_s Production Events

At 4.03 GeV, a variety of charmed meson states are produced. Their decays all have generally similar topologies, though there are characteristics (*e.g.*, slow pions from $D^* \to \pi D$) by which particular initial states can be selected. For the case of $e^+e^- \to D_s^+ D_s^-$, the two main features are (i) back-to-back production in the laboratory reference frame, and (ii) multiple strangeness in the final states, from the $c \to sW^+$ quark transition. In my analysis, I exploit both of these features to select candidate D_s events (*tags*) over other charm "background" using reconstruction of exclusive decay channels. Back-to-back production implies that a kinematic fit to the hypothesis $e^+e^- \to X\overline{X}$ (a one-constraint fit) can select $D_s^+ D_s^-$ events (along with $D\overline{D}$ and $D^*\overline{D}^*$) while rejecting other topologies, and the spectator strange quark in D_s mesons implies that Cabibbo-suppressed charm decays should be Cabibbo-favored in D_s decays, and *vice versa*.

The use of exclusive channels to identify candidate D_s events is also essential for the other half of my analysis, the identification of candidate leptonic decays. By fully reconstructing the hadronic tags, I guarantee that all remaining tracks in the event belong to the decay of the other D_s meson in the event. I can then identify candidate leptonic decays by requiring a single *recoil* track, as I discuss in detail in Chapter 9.

8.1 Hadronic Tag Channel Selection

I have chosen three decay channels for my tag selection, each involving three charged tracks and no neutrals (photons or π^0's): $D_s^+ \to \phi(1020)\pi^+, \phi \to K^+K^-$ (which I will refer to as $\phi(K^+K^-)\,\pi^+$); $D_s^+ \to \overline{K}^{*0}(892)K^+, \overline{K}^{*0} \to K^-\pi^+$ ($\overline{K}^{*0}(K^+\pi^-)\,K^+$); and $D_s^+ \to K_S^0 K^+, K_S^0 \to \pi^+\pi^-$ ($K_S^0(\pi^+\pi^-)\,K^+$). Although the total branching fraction

for D_s decays to these final states is small (Table 8.1), they are expected to be relatively "clean:" they are all Cabibbo-favored for D_s decays, and Cabibbo-suppressed for D^+ decays, and they do not contain neutral final-state particles which can drastically increase backgrounds. Throughout this discussion, I imply both charge conjugation states for the decays, unless I explicitly state otherwise, for example, $D_s^+ \to \overline{K}^{*0}K^+$ and $D_s^- \to K^{*0}K^-$.

Table 8.1: Branching fractions for tag channels $D_s \to A(ab)c$ ($D_s \to Ac, A \to ab$). $B_{D_s} = \mathcal{B}(D_s \to Ac)$ is the branching fraction for D_s decay to the primary state Ac; $B_A = \mathcal{B}(A \to ab)$ is the branching fraction for the secondary meson ($\phi(1020)$, $K^{*0}(892)$, or K_S^0) to decay into the indicated final state, with appropriate Clebsch-Gordan coefficients. Branching fractions are taken from the 1994 *Review of Particle Properties* [1].

$D_s \to A(ab)c$	B_{D_s} (%)	B_A (%)	B_{total} (%)
$\phi(K^+K^-)\,\pi^+$	3.5 ± 0.4	49.1 ± 0.9	1.72 ± 0.2
$\overline{K}^{*0}(K^+\pi^-)\,K^+$	3.3 ± 0.5	$2/3 \times (99.770 \pm 0.020)$	2.2 ± 0.3
$K_S^0(\pi^+\pi^-)\,K^+$	3.5 ± 0.7	$1/2 \times (68.61 \pm 0.28)$	1.20 ± 0.24
Total branching fraction (%) to final states			5.1 ± 0.4

In each event, I search for hadronic tag candidates by choosing groups of three charged tracks $\{i, j, k\}$ with the following general characteristics:

1. all three tracks must be reconstructed with $\chi^2_{\text{MDC}} < 50$;

2. all three tracks must project into the central portion of the BES detector ($|\cos\theta| < 0.75$);

3. the total charge of the tracks must be ± 1;

4. the invariant mass of the three tracks (using the particle identifications for the hypothesized final state) must be within 250 MeV of the D_s mass (1968.5 MeV) [1];

5. the sum of the momenta of the three tracks must be less than the beam energy.

For each acceptable group of tracks, I compare every permutation to the $D_s \to A(ab)c$ decay channels, requiring identification of the intermediate resonance, a good kinematic fit, and overall consistency with the decay hypothesis combining particle identification, the resonance mass, and the kinematic fit. I construct a global likelihood for these criteria, and identify the candidate tag as that track permutation with the maximum likelihood value of those surviving the selection cuts. Figure 8.1 shows the distributions of kinematic fit mass (Section 8.2) for every track permutation tested, with no maximum likelihood or other selection criteria.

8.2 Kinematic Fitting

For each track permutation, I perform a kinematic fit with the TELESIS software package [34,35], assigning the appropriate mass hypotheses to each one. The fit hypothesis is to an event of the form $e^+e^- \to X\overline{X}$, with the three tracks forming resonance X recoiling against an unidentified but equal-mass partner. For the $\phi(K^+K^-)\pi^+$ and $\overline{K}^{*0}(K^+\pi^-)K^+$ channels, the X resonance is only the final state (e.g., $K^+K^-\pi^+$) and not the full $A(bc)c$ decay chain. In these cases the fit has one constraint, that the fitted energy of X is equal to the beam energy.[1]

For the $K^0_S(\pi^+\pi^-)K^+$ channel, I require that the two pions from $K^0_S \to \pi^+\pi^-$ originate from a secondary vertex, separated from the IP. I do this using routines in the TELESIS package to "swim" the pion momentum vectors along their helical trajectories to the point where those trajectories cross. I then do a kinematic fit to

[1]The unidentified $\overline{X}$ introduces four free parameters, $p^\mu_{\overline{X}}$. The event fit constrains $p^\mu_X + p^\mu_{\overline{X}} \equiv (\vec{0}, E_{cm})$. The equal-mass requirement constrains $p^\mu_X p_{X,\mu} \equiv p^\mu_{\overline{X}} p_{\overline{X},\mu}$. Expanding these we obtain

$$\vec{p}_X + \vec{p}_{\overline{X}} = \vec{0} \;\to\; \vec{p}_X = -\vec{p}_{\overline{X}} \tag{8.1}$$

$$E_X + E_{\overline{X}} = E_{cm} \tag{8.2}$$

$$|\vec{p}_X|^2 - E_X^2 = |\vec{p}_{\overline{X}}|^2 - E_{\overline{X}}^2 \,. \tag{8.3}$$

Equations 8.1 and 8.3 imply $E_X^2 = E_{\overline{X}}^2$, which reduces equation 8.2 to

$$E_X = E_{cm}/2 = E_{beam}$$

as the only constraint.

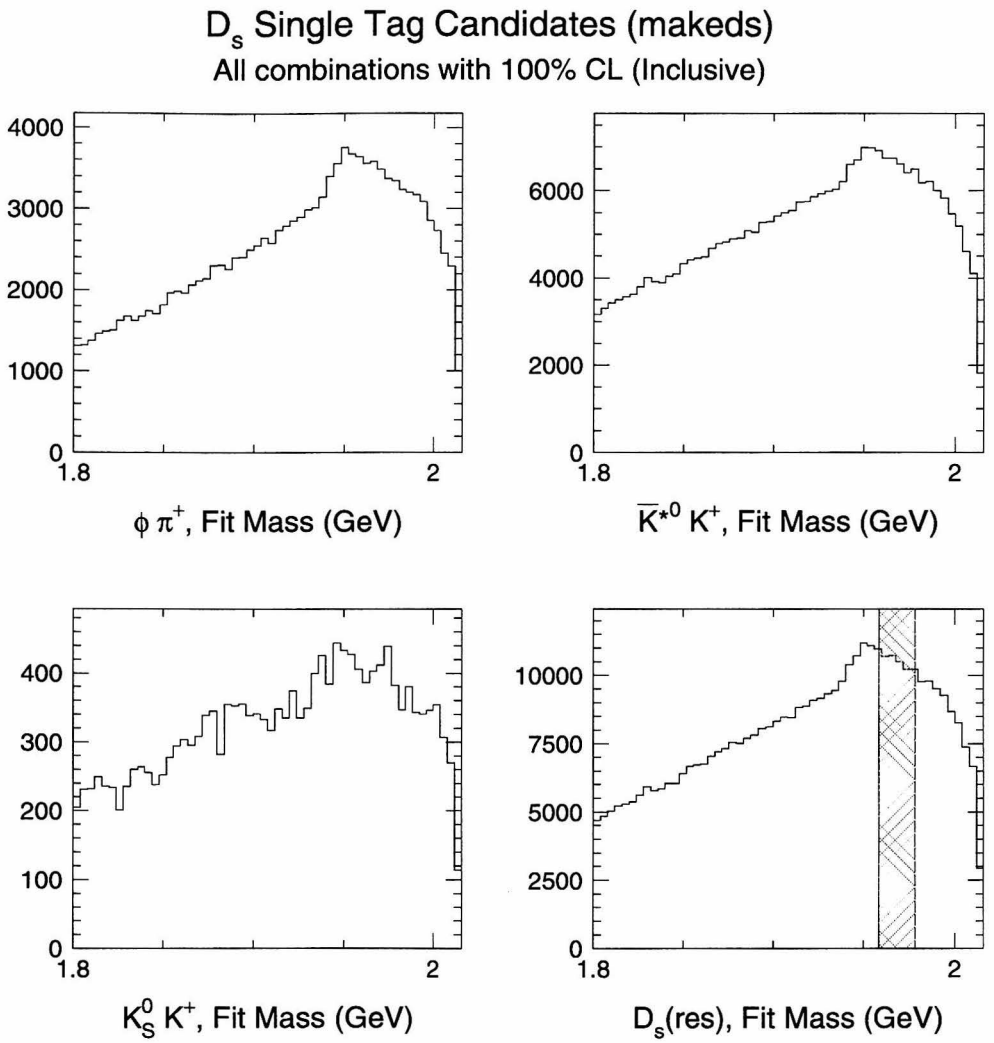

Figure 8.1: Inclusive distributions of fitted D_s "tag" mass for all three-track permutations. Only the broad candidate selection criteria described in Section 8.1 have been applied to the D_s data sets. The hatched area on the lower right plot shows the D_s peak region, 1958.5–1978.5 MeV/c^2.

the final state $K^+\pi^+\pi^-$, requiring the invariant mass of the two-pion system to equal the K^0_S mass (497.67 MeV/c^2) [1].

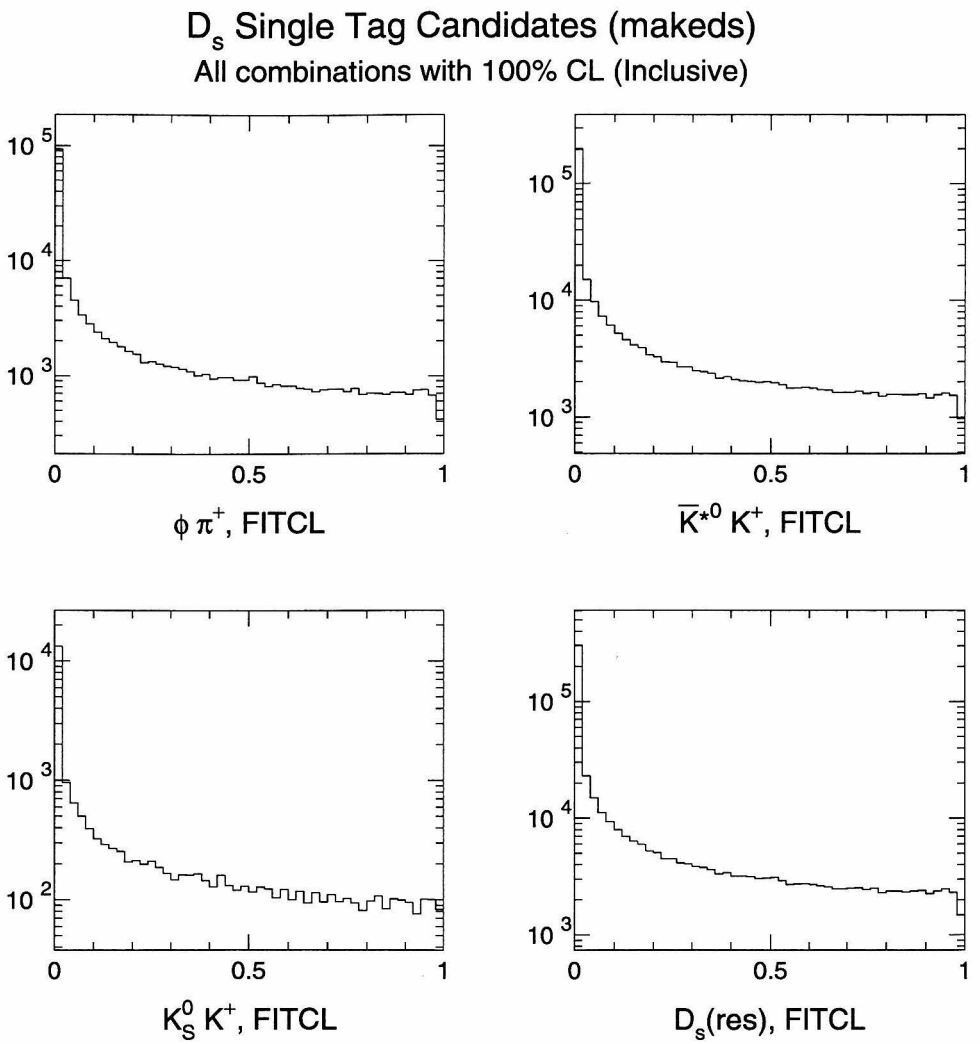

Figure 8.2: Inclusive kinematic fit confidence levels for D_s hadronic decay candidates, for each of the channels shown. The plot at lower right is the sum of the other three.

Figure 8.2 shows the distribution of kinematic fit confidence levels, $P(\chi^2_{fit}; ndf)$ (ndf = 1 or 2), for my inclusive sample of candidate track permutations. It is clear from the plots (as in Figure 8.1) that the inclusive sample is dominated by background, since the kinematic fit confidence levels are not flat, as they should be for true signal events. Some additional selection criteria must be applied in order to reduce the

background levels, before using the confidence levels to select candidate signal events; these additional criteria are discussed below.

8.3 Secondary Resonance Selection

In order to reduce the overwhelming background surrounding the assumed D_s mass peak (as Figure 8.1 shows), I require each track permutation to have a pair of tracks consistent (at 95% confidence level) with the resonance expected for the given decay channel (ϕ, K^{*0}, or K^0_S), including both invariant mass and particle identification. To determine the resonance parameters (mass, width, and shape), I have reconstructed inclusive two-track combinations using a small subset of the BES D_s data.[2] By doing this fit to real data, I naturally incorporate all of the effects of detector resolution, scattering, and energy loss on the track momenta.

In Figure 8.3 I show the inclusive invariant mass distributions of the resonance candidates $\phi(1020) \to K^+K^-$, $K^{*0}(892) \to K^+\pi^-$, and $K^0_S \to \pi^+\pi^-$. I form these distributions by taking every two-charged-track combination in each event, assigning the mass (and energy) of each track based on the hypothetical final state, and computing the invariant mass of the system. It is clear from the distributions that some selection is required to reduce the overwhelming inclusive backgrounds in order to fit the resonance shapes accurately. For this purpose, restrictive selection criteria are useful, provided there are sufficient candidates in the final sample.

The criteria used to select inclusive samples for fitting are different for each resonance. For the $\phi(1020) \to K^+K^-$ candidates, the kinematics are fairly clean, except for a peculiarity in the BES track reconstruction code: A slow particle (usually an electron) which curls up in the drift chamber ($50 \lesssim p \lesssim 80$ MeV/c) will be reconstructed as a pair of exactly back-to-back oppositely charged tracks. It turns out that if such tracks are assigned the mass of kaons, and their momenta are "corrected" for the energy loss expected of kaons in the beam pipe [42], the result is an apparent

[2]The subset consists of 100 runs from each of the 1992, 1993, and 1994 running periods, a total of 160,506 events.

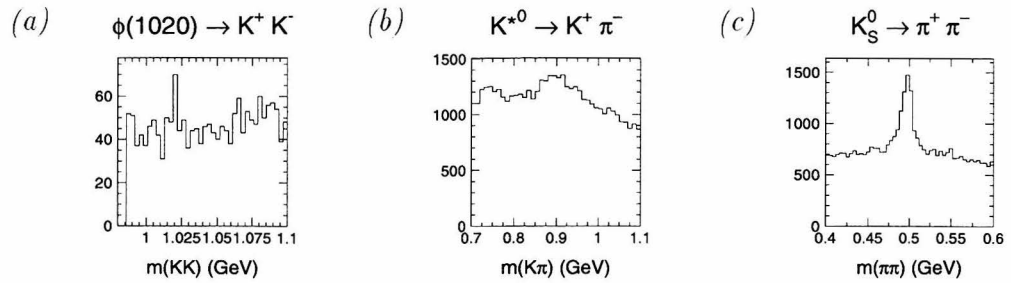

Figure 8.3: Inclusive invariant mass distributions of two-body resonances, from 300 runs of BES 4.03 GeV data. *(a)* Hypothetical $\phi(1020) \to K^+K^-$ candidates; *(b)* hypothetical $K^{*0}(892) \to K^+\pi^-$ candidates; and *(c)* hypothetical $K^0_S \to \pi^+\pi^-$ candidates. No particle identification or other selection has been applied to the two-track combinations.

"peak" in invariant mass around 1030 MeV/c^2. Since real ϕ's are never at rest in the lab, restricting the candidates to noncollinear pairs ($\vec{p}_1 \cdot \vec{p}_2 > -0.99$) eliminates this false peak. Figure 8.4 shows my fit to the resonance parameters of the $\phi(1020)$. The results are[3]

$$M(\phi) = 1019.4 \pm 0.2 \text{ MeV}/c^2, \quad \sigma(\phi) = 1.1 \pm 0.2 \text{ MeV}/c^2.$$

For the $K^{*0}(892) \to K^+\pi^-$ candidates, the background is dominant. I require that both tracks be well identified (maximum likelihood) and that the two tracks originate within two centimeters of one another along the z axis. The restrictive particle ID is necessary because the K^{*0} reflection (*i.e.*, reversing the identification assignments of the tracks from a real K^{*0}) is very close to the actual K^{*0} resonance, leading to an apparent shift and broadening of the observed peak. Figure 8.5 shows my fit to the K^{*0} resonance parameters, giving

$$M(K^{*0}) = 894.4 \pm 2.6 \text{ MeV}/c^2, \quad \sigma(K^{*0}) = 26.3 \pm 3.1 \text{ MeV}/c^2.$$

The K^0_S has a macroscopic lifetime ($\tau_{K^0_S} = 89.26 \pm 0.12$ ps, or $c\tau = 2.676$ cm [1]).

[3]For all of these resonances, I report a fitted mass and a Gaussian width σ, not the Breit-Wigner resonance width Γ. Since these measurements are convoluted with the response and resolution of the BES detector, the Gaussian approximation is reasonable, and is much simpler computationally.

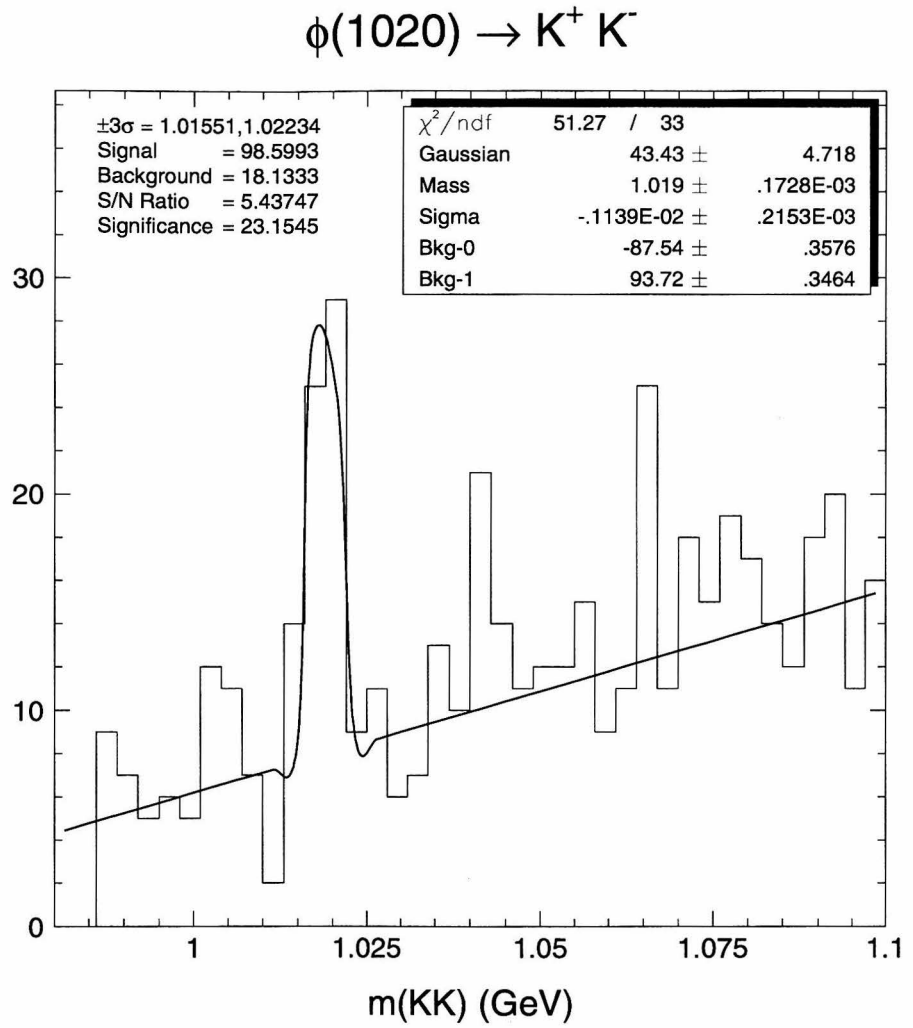

Figure 8.4: Invariant mass distribution of $\phi(1020) \rightarrow K^+K^-$ candidates, fit to a Gaussian peak and linear background function. The background shape away from the peak is not consequential, except to normalize the number of events.

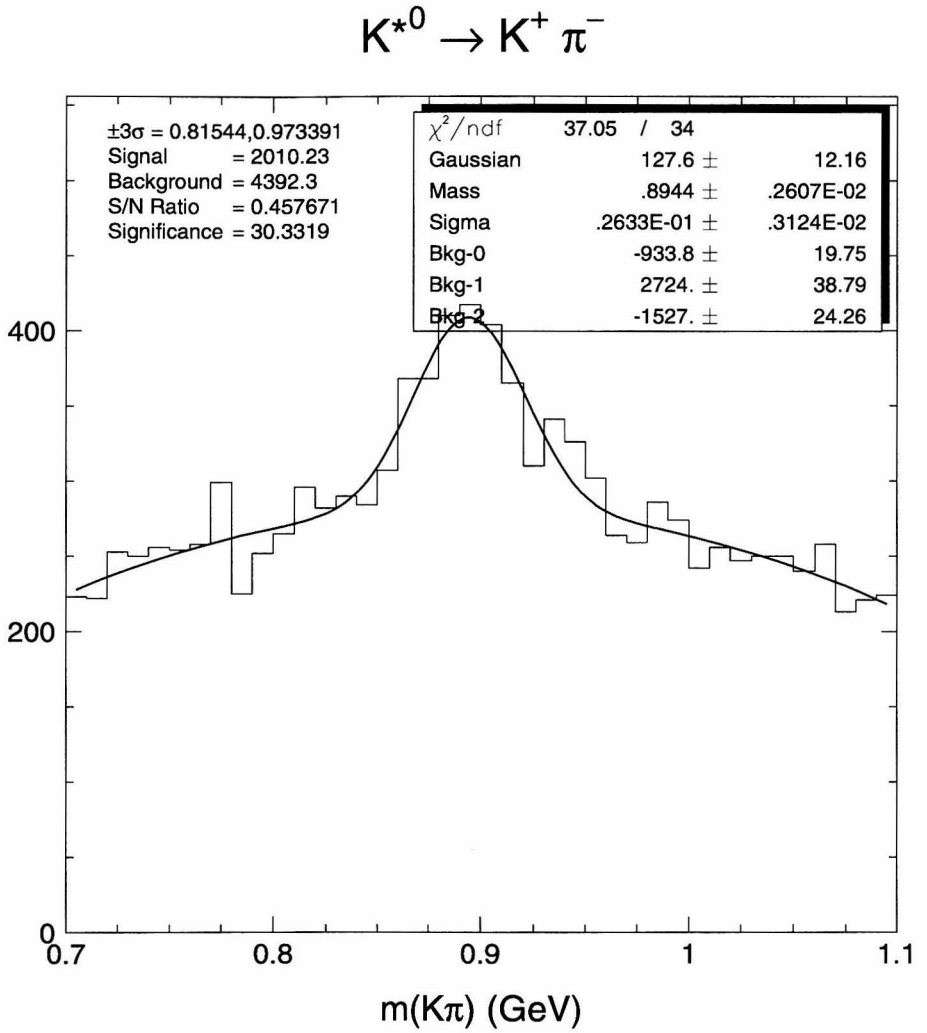

Figure 8.5: Invariant mass distribution of $K^{*0}(892) \to K^+\pi^-$ candidates, fit to a Gaussian peak and quadratic background function.

In a charmed meson decay, a typical K^0_S (with $p \sim 1$ GeV/c, $\beta\gamma \sim 2$) will travel a few centimeters before decaying. I identify K^0_S candidates from among the inclusive pion pairs by requiring that the helical tracks intersect at a point offset from the event IP by at least 1 cm. Since the momentum vectors follow the helical path, I calculate the invariant mass of the candidate at the secondary vertex, which improves the mass resolution for real K^0_S decays. Figure 8.6 shows the distribution for these candidates, and my fit to the resonance parameters,

$$M(K^0_S) = 497.3 \pm 0.2 \text{ MeV}/c^2, \quad \sigma(K^0_S) = 7.63 \pm 0.31 \text{ MeV}/c^2.$$

I use the resonance parameters as part of my likelihood function for evaluating candidate D_s decays, as I show in Section 8.4.

8.4 Likelihood and Consistency Function

To evaluate the consistency of each tag hypothesis with the given set of tracks, I construct a global likelihood function,

$$\begin{aligned} \mathcal{L}(A(ab)c) &= \mathcal{L}_{\text{pid}}(i = a) \cdot \mathcal{L}_{\text{pid}}(j = b) \cdot \mathcal{L}_{\text{pid}}(k = c) \\ &\quad \times \mathcal{L}_{\text{res}}(m_{ij}; M_A, \Gamma_A) \cdot \mathcal{L}_{\text{fit}}(\chi^2, ndf) \ , \end{aligned} \tag{8.4}$$

where $\mathcal{L}_{\text{pid}}$ is the BES particle identification likelihood function (equation 7.1) which I discuss in detail in Chapter 7; $\mathcal{L}_{\text{res}}$ is a Gaussian distribution for the two-body invariant mass m_{ij} with respect to the resonance (all three resonances are dominated by detector resolution, so the use of a Breit-Wigner distribution is unnecessary); and $\mathcal{L}_{\text{fit}}$ is the likelihood (ndf-dimensional Gaussian) for the result of the kinematic fit. For the $K^0_S(\pi^+\pi^-)\,K^+$ hypothesis, I identify the K^0_S by fitting the two pions to a secondary vertex (Section 8.3) with the invariant mass of the pair constrained to the nominal K^0_S mass, 497.67 MeV/c^2 [1]; this is included as a second constraint in the kinematic fit, in place of $\mathcal{L}_{\text{res}}(K^0_S)$.

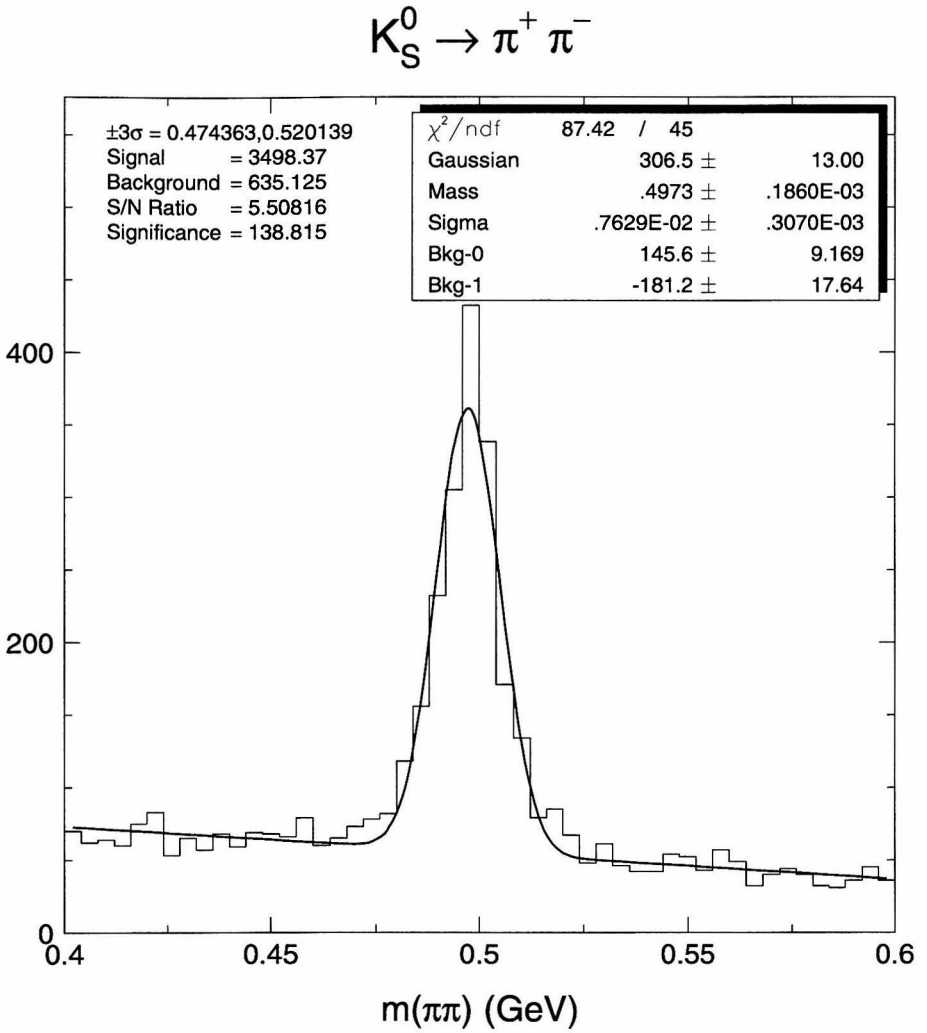

Figure 8.6: Invariant mass distribution of $K^0_S \to \pi^+\pi^-$ candidates, using track momenta swum to secondary vertex. Fit is to a Gaussian peak and linear background function.

As outlined in Chapter 6, I define a probability or confidence level, $P(A(ab)c)$, for each hypothesis by integrating the tail of the likelihood function (in practice, I sum the equivalent χ^2 and degrees of freedom of the factors in equation 8.4). I require candidates to be consistent with the given hypothesis by requiring $P(A(ab)c) > P_{min} = 0.05$ (95% confidence level). If the set of tracks has multiple consistent hypotheses, I choose the one with maximum likelihood value as "best." The efficiency of this selection method, and the rate of wrong identifications, are presented in Section 8.8. Figure 8.7 shows the fitted mass distributions of my tag candidates after this procedure. Compared to Figure 8.1, the signal-to-background ratio is much improved, but the signal is not prominent enough for me to use it for my recoil analysis.

8.5 Background Channel Suppression

I tag my sample of $D_s^+ D_s^-$ events with three hadronic decay channels, $D_s \to \phi(K^+K^-)\pi^+$, $\overline{K}^{*0}(K^+\pi^-)K^+$, and $K_S^0(\pi^+\pi^-)K^+$. To suppress backgrounds from other D_s and $D^\pm$ decay channels, I apply my analysis technique to the additional final states $\overline{K}^{*0}(K^-\pi^+)\pi^+$ and $K_S^0(\pi^+\pi^-)\pi^+$. I evaluate the likelihoods for these channels in the same way as for my signal channels (equation 8.4), and include the values to select the maximum-likelihood hypothesis for each three-track combination. This helps significantly to suppress background from $D^\pm$ decays, especially the Cabibbo-favored $K_S^0(\pi^+\pi^-)\pi^+$ mode.

8.6 Additional Selection Criteria

I have designed the candidate event selection, described above, to be efficient as possible. However, as Figure 8.7 shows, it is not especially pure. The background levels would render my determination of the leptonic branching fractions useless. In order to obtain a reasonable signal to background ratio, I impose additional selection criteria on each D_s decay channel. For every channel (Figures 8.8 through 8.10), I require that the total energy of the candidate ($E_{fit} \equiv \sqrt{p_{fit}^2 + m_{D_s}^2}$) be within 50 MeV

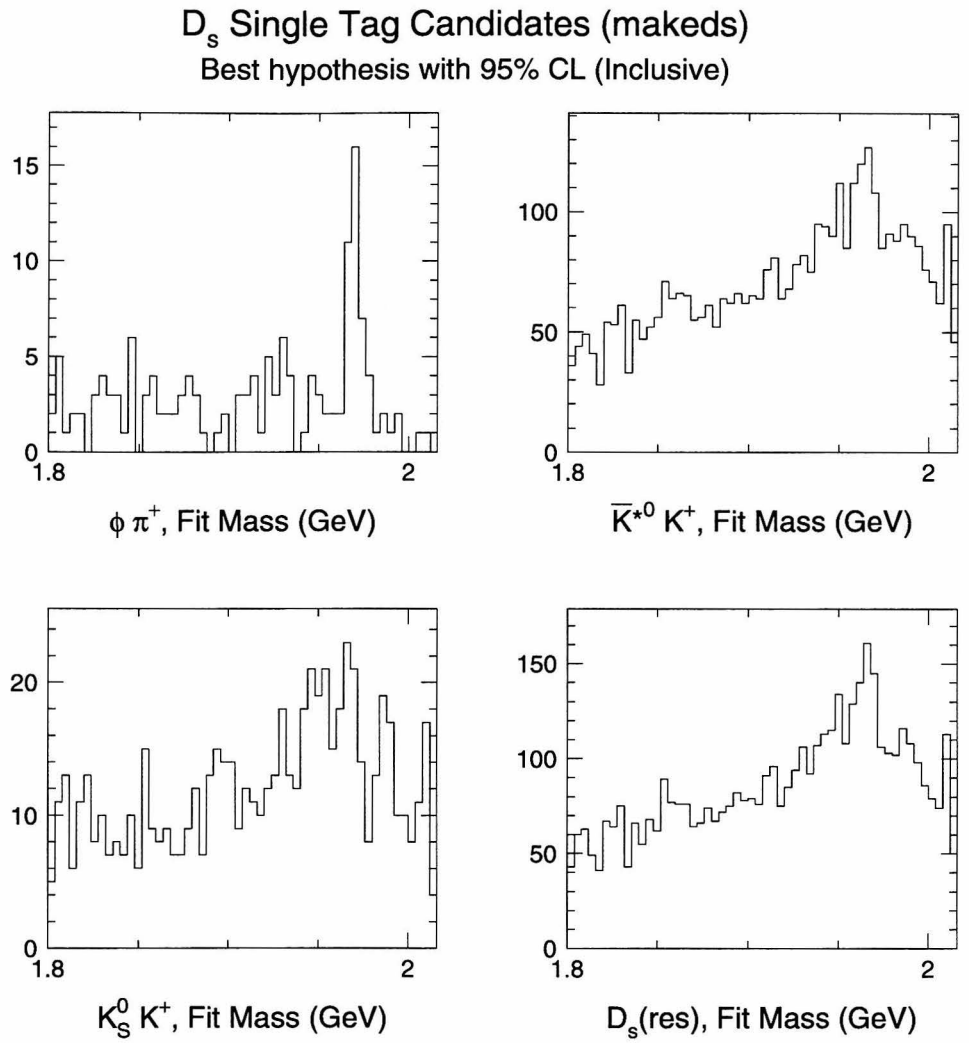

Figure 8.7: Mass distributions of D_s tag candidates after maximum likelihood event selection. Candidate tags in each event are required to have at least a 95% confidence level ($P_{tag} > 0.05$). If multiple permutations pass this selection, the one with the maximum likelihood value is identified as the "true" or "best" tag, to eliminate double counting.

of the beam energy (2.015 GeV), and that the fitted momentum of the candidate, $\vec{p}_{fit}$, must point into the central region of the BES detector ($|\cos\theta_{fit}| < 0.7$).

The $\phi(K^+K^-)\,\pi^+$ channel requires few additional selection criteria, since the kinematics of the $\phi \to K^+K^-$ decay eliminate nearly all of the background (as Figure 8.7, top left, shows). Figure 8.8 shows the distributions of the two selection variables (E_{fit} and $\cos\theta_{fit}$) with respect to the fitted mass of each candidate. The hatched areas indicate the selection criteria (rejected candidates) I have imposed in addition to the global likelihood.

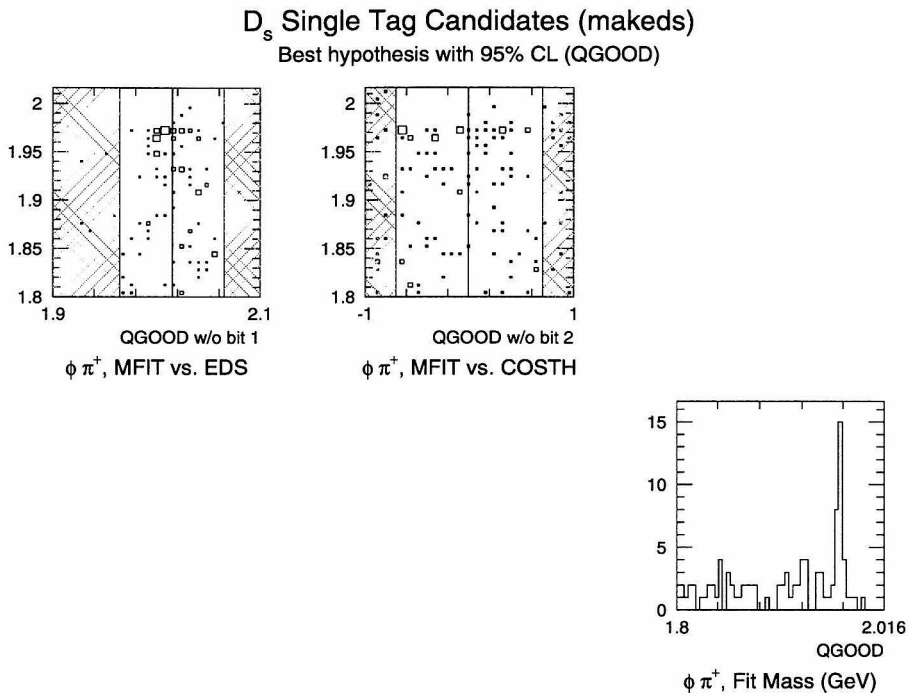

Figure 8.8: Distributions of variables for final selection of candidate $D_s \to \phi(K^+K^-)\,\pi^+$ decays. Each plot shows the distribution of one selection variable (x axes) with respect to the fitted mass of the candidate D_s decays (y axes). The hatched areas, as described in the text, are excluded from the final sample, as shown in the histogram in the lower right corner.

The $\overline{K}^{*0}(K^+\pi^-)\,K^+$ channel has much higher background than $\phi(K^+K^-)\,\pi^+$; to reduce this background, I require that the K^{*0} resonance be well identified, using selection criteria similar to those I described in Section 8.3,

1. the two tracks from the hypothetical K^{*0} must reconstruct to within $\Delta z < 2$ cm at the IP;

2. the helicity angle[4] of the K^{*0}, θ_π, must have $|\cos\theta_\pi| > 0.4$.

In Figure 8.9 I show how these selection variables vary with the candidate fitted mass. The hatched areas in each plot show the excluded candidates for the indicated criterion.

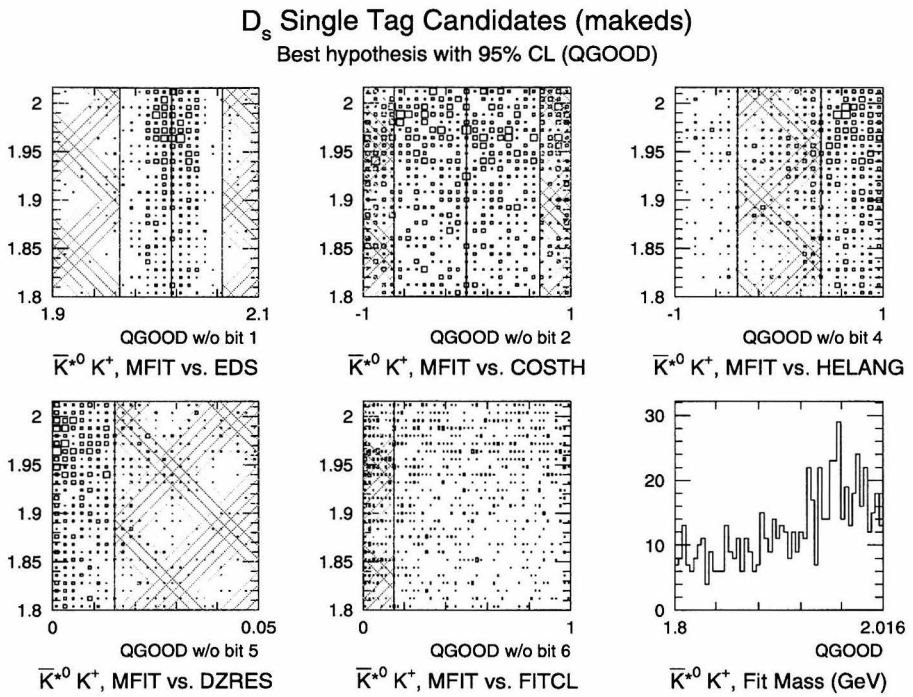

Figure 8.9: Distributions of variables for final selection of candidate $D_s \to \overline{K}^{*0}(K^+\pi^-)\,K^+$ decays. Each plot shows the distribution of one selection variable (x axes) with respect to the fitted mass of the candidate D_s decays (y axes). The hatched areas, as described in the text, are excluded from the final sample, as shown in the histogram in the lower right corner.

The $K^0_S(\pi^+\pi^-)\,K^+$ channel has significant background in the mass region 1920–1980 MeV/c^2, originating mainly in charged and neutral D^*D events (see Section 8.8).

[4]The *helicity angle* for $K^{*0} \to K^+\pi^-$ decays is defined as the angle between the pion momentum in the K^{*0} rest frame, and the momentum of the K^{*0}in the lab frame.

I have studied a number of selection variables, including K^0_S kinematics and identification, track alignment (in both the $x-y$ plane and along the z axis), and the kinematic fit. None of these variables are particularly useful at reducing this background, at least not without impacting the selection efficiency for true $K^0_S(\pi^+\pi^-)\,K^+$ events. As I discuss in Section 8.8, I model the background using Monte Carlo data, and include that background in my determination of the number of tags. Thus, to maximize my event selection efficiency, I do not impose any selection criteria beyond the two listed above, and 2-C kinematic fit of $D_s \to K^0_S(\pi^+\pi^-)\,K^+$ with the pions constrained to the K^0_S mass (497.6 MeV/c^2) at an identified vertex. The distributions of the selection variables E_{fit} and $\cos\theta_{fit}$ are shown in Figure 8.10.

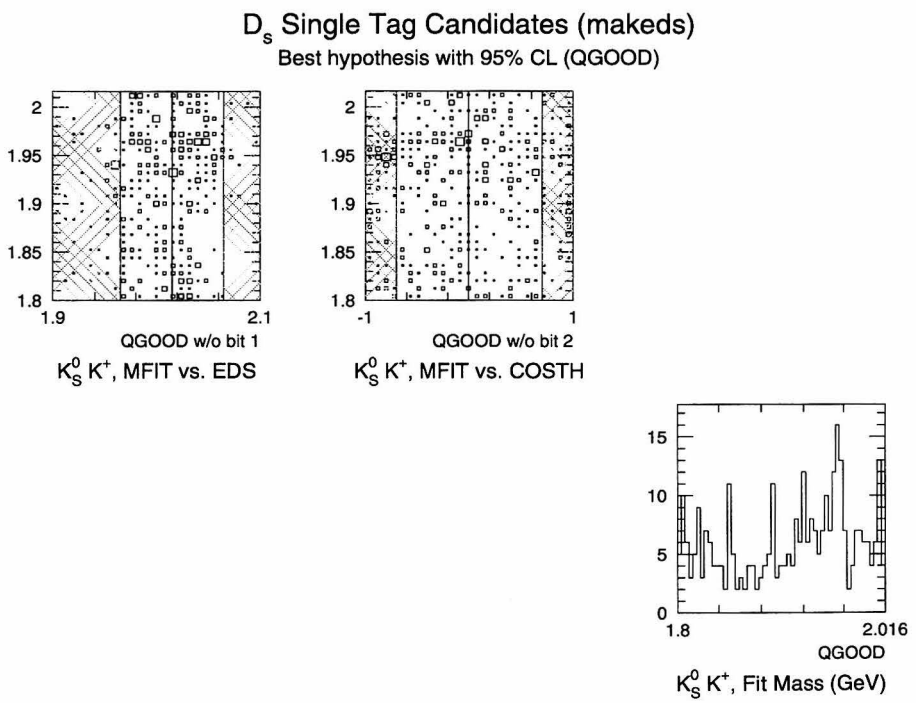

Figure 8.10: Distributions of variables for final selection of candidate $D_s \to K^0_S(\pi^+\pi^-)\,K^+$ decays. Each plot shows the distribution of one selection variable (x axes) with respect to the fitted mass of the candidate D_s decays (y axes). The hatched areas, as described in the text, are excluded from the final sample, as shown in the histogram in the lower right corner.

8.7 Tagged Event Sample

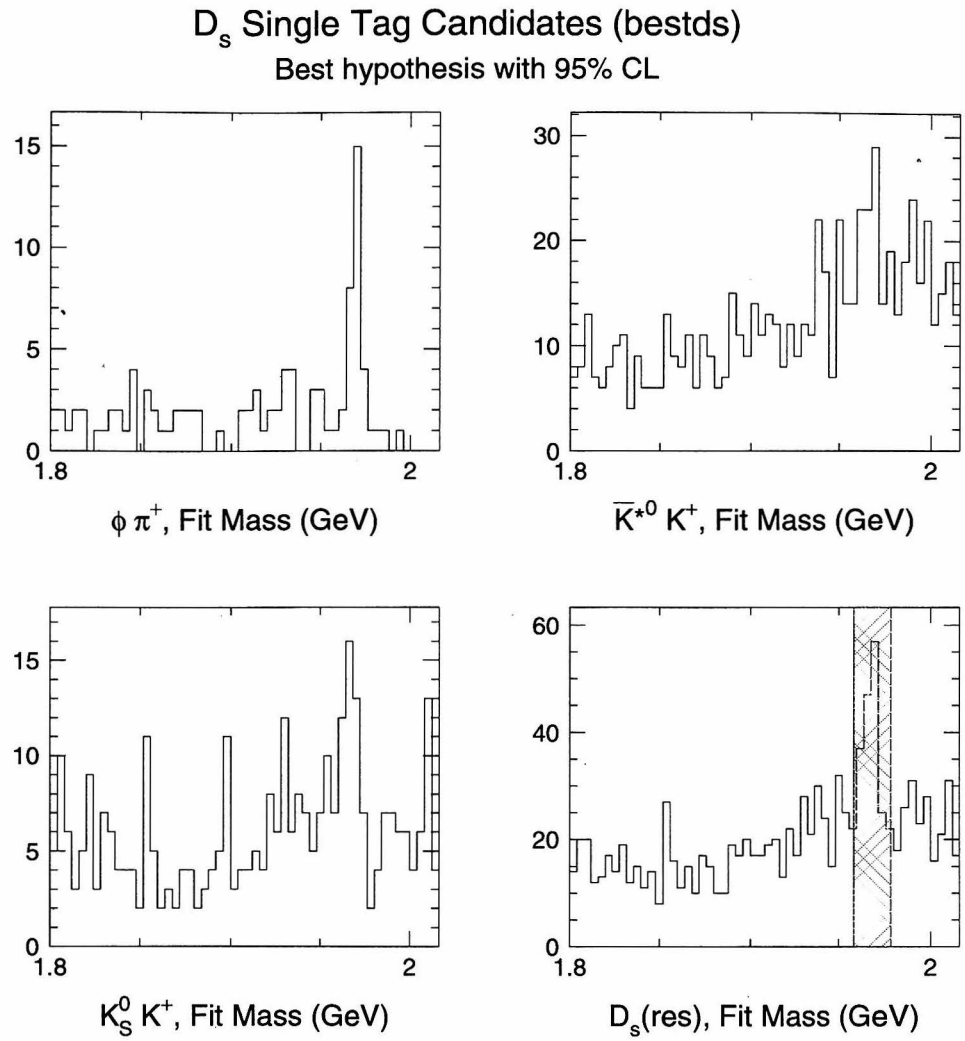

Figure 8.11: Kinematically fitted mass distributions of final sample of hadronic D_s events. All event selection criteria described in the text have been applied to the BES data. The event sample consists of those events with a fitted mass within 10 MeV/$c^2$of the D_s peak at 1968.5 MeV/c^2 (hatched).

Combining all of the information in this chapter—the kinematic fit, resonance selection, maximum likelihood selection among candidate permutations, and additional cuts—I obtain the final distribution of D_s hadronic decay candidates masses shown in Figure 8.11. To determine the number of "true" $D_s^+ D_s^-$ events in this set, I perform an unbinned maximum-likelihood fit [43] of the candidate masses to a Gaussian signal

(the natural width of the $D_s$is essentially zero) plus a background function chosen to model the shape of the data,

$$f_{\rm bkg}(m) = (m_0 - m)^a \exp\left(-b(m_0 - m)\right) \qquad (m < m_0) \quad (8.5)$$

$$= \exp\left(\frac{(m_0 - m_{max})\ln(m_0 - m) - (m_0 - m)}{\lambda}\right) , \quad (8.6)$$

where m_0 is the maximum allowed mass ($f_{\rm bkg} \equiv 0$ for $m > m_0$), and a and b are parameters which give the background a slow rise at low mass, and a rapid fall to zero near the maximum. In equation 8.6, I have rewritten this function in a form which relates more intuitively to the shape of the background distribution: $m_{max} \equiv m_0 - a/b$ is the point where the background is maximum, and $\lambda \equiv 1/b$ is the scale height of the low-mass tail, in units of mass. I show the result of this fit in Figure 8.12. There are 190 candidate events within ± 10 MeV/c^2 of the D_s mass (1958.5 to 1978.5 MeV/c^2); the signal contains 74.7 ± 14.6 events. The fitted D_s mass is 1968.5 ± 0.8 MeV/c^2, in excellent agreement with the world average [1]; the peak has a Gaussian width (σ) of 3.0 ± 0.8 MeV/c^2.

8.8 Single Tag Selection Studies

Technically, the sample of tagged D_s events stands on its own, as the "given" input to my search for D_s leptonic decays. It is an important and useful exercise, however, to determine the efficiency of the selection procedure, as well as the levels and sources of backgrounds to the single tag sample. In particular, I will use the results of these selection studies to make a more detailed fit of the single-tag mass distribution for signal and background numbers.

With a sufficiently large data sample, I could use "self-consistency" of the data itself to determine efficiency and background levels. For example, given a sample of tagged D_s events, I could apply the same selection procedure to the recoil side of each

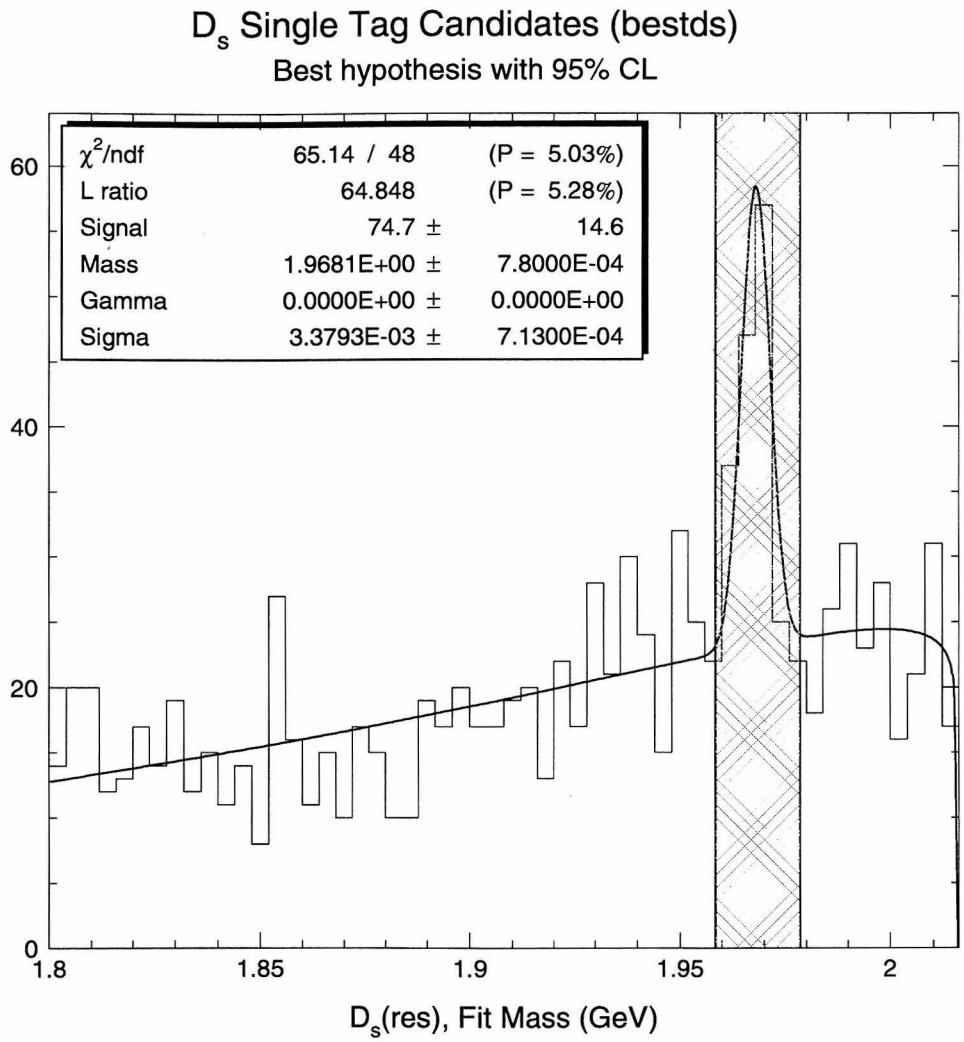

Figure 8.12: Unbinned maximum-likelihood fit of final D_s candidate sample to a Gaussian signal plus the background function $f_{\rm bkg}$ described in the text. The hatched region shows the final candidate event selection, $1958.5 < m_{fit} < 1978.5$ MeV/c^2.

event (*double-tagging*), and solving the system of equations

$$\begin{aligned} N_{single} &= N_{data}(\varepsilon_{tag} + \eta_{bkg}) \\ N_{double} &= N_{data}(\varepsilon_{tag}^2 + \eta_{bkg}^2) \end{aligned}$$

for the tagging efficiency ε_{tag} and the background η_{bkg}. Unfortunately, this technique requires a statistically significant sample of double-tagged events, much larger than that available from the BES experiment. A similar double-tag study [44] found only two candidate events for the $D_s \to \phi(K^+K^-)\,\pi^+$ recoil channel, which is not sufficient to determine either ε_{tag} or η_{bkg} with any significance.

As a result, my efficiency and background studies (as well as those for the leptonic event selection described in Chapter 9) are done using Monte Carlo simulations. In order for these studies to be successful—to provide accurate estimates for efficiencies within each signal channel, for the overall efficiency of tag selection, and for the levels background—the physics incorporated into the simulation, and the detector response, must both be accurate, as discussed in Chapter 5. For my background studies, I use events generated in all of the available charmed meson channels ($D\overline{D}$, $D^*\overline{D}$, $D^*\overline{D}^*$ and $D_s^+D_s^-$), using world average branching fractions for D and D^* decays; the relative proportions of each charm channel (with charged and neutral mesons considered separately) are specified using the coupled-channel model of Eichten *et al.* [28] (Appendix A).

8.8.1 Background Distributions

The backgrounds present in my data (about half the total number of events in my sample) come from a variety of sources: non-D_s charm production, $\tau^+\tau^-$ semi-hadronic decays, light-quark production (both direct and two-photon), and wrong combinations of tracks from true $D_s^+D_s^-$ production. As mentioned in Chapter 5, Monte Carlo data sets exist for all of these sources except the two-photon physics. Although I will show (Section 8.8.2) that none of these sources has a high rate of misidentification as a D_s event, the approximately 1:1 signal-to-background ratio ob-

served (Figure 8.12) arises because the $D_s^+ D_s^-$ production rate is small compared to any of the other sources. In Table 8.2, I show the cross sections for various e^+e^- processes, based on first order QED calculations (for the $\tau^+\tau^-$ and $q\bar{q}$ processes) and the coupled-channel model (for charm production) [28].

Table 8.2: Production rates for sources of D_s candidate backgrounds at 4.03 GeV. The $\tau^+\tau^-$ and $q\bar{q}$ rates are first order QED calculations; the charm rates are predicted from the coupled-channel model, normalized to $\sigma(\mu^+\mu^-) = 5.348$ nb.

$e^+e^- \to$	σ (nb)
$D_s^+ D_s^-$	0.808
$q\bar{q}$	10.696
$\tau^+\tau^-$	3.316

$e^+e^- \to$	σ (nb)
$D^0\overline{D}^0$	0.051
$D^0\overline{D}^{*0}$	3.341
$D^{*0}\overline{D}^{*0}$	3.088

$e^+e^- \to$	σ (nb)
$D^+\overline{D}^-$	0.021
$D^{\pm}\overline{D}^{*\mp}$	3.605
$D^{*+}\overline{D}^{*-}$	1.901

In Figure 8.13, I show the fitted mass distributions for each of the background sources in Table 8.2. Each plot is produced by applying my full selection technique to a sample of 100,000 Monte Carlo events of the indicated source. Thus, except for normalization, these plots should be compared to Figure 8.12 to determine the contribution each one makes to my background.

I have also used the plots in Figure 8.13 in an alternative method of fitting my final data sample. By smoothing the background plots with a 10-point cubic spline, I can reduce the statistical fluctuations and approximate the "true" shape of each background source's candidate mass distribution. I can then form the weighted sum of all of the background plots, normalized to the size of each Monte Carlo sample and weighted by the expected production cross section (Table 8.2), as shown in Figure 8.14. If I use this as my background shape, in place of equation 8.6, I should get a better approximation to my data than the latter function provides. The result of this procedure is shown in Figure 8.15, and the difference is clear. The number of candidate signal events is increased to 76.9 ± 13.8 (which is consistent with my previous results), and the background shape no longer systematically overestimates my data.

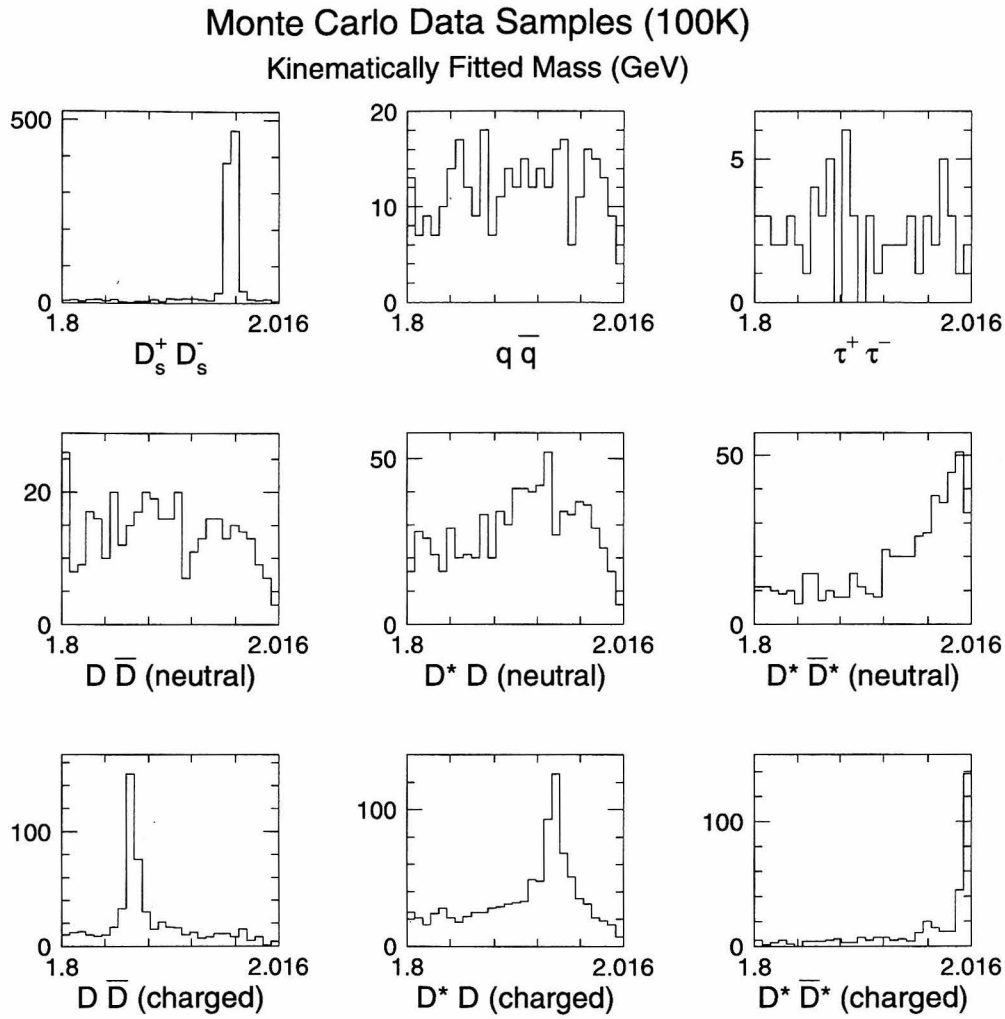

Figure 8.13: Fitted mass distributions of three-track combinations identified as candidate D_s decays, from Monte Carlo data sets. The upper left plot shows the "ideal" signal, from a Monte Carlo sample of $D_s \rightarrow \phi(K^+K^-)\pi^+$, $\overline{K}^{*0}(K^+\pi^-)K^+$, and $K_S^0(\pi^+\pi^-)K^+$ decays. All of the other plots are from the background sources indicated, with input data sets of 100,000 events each.

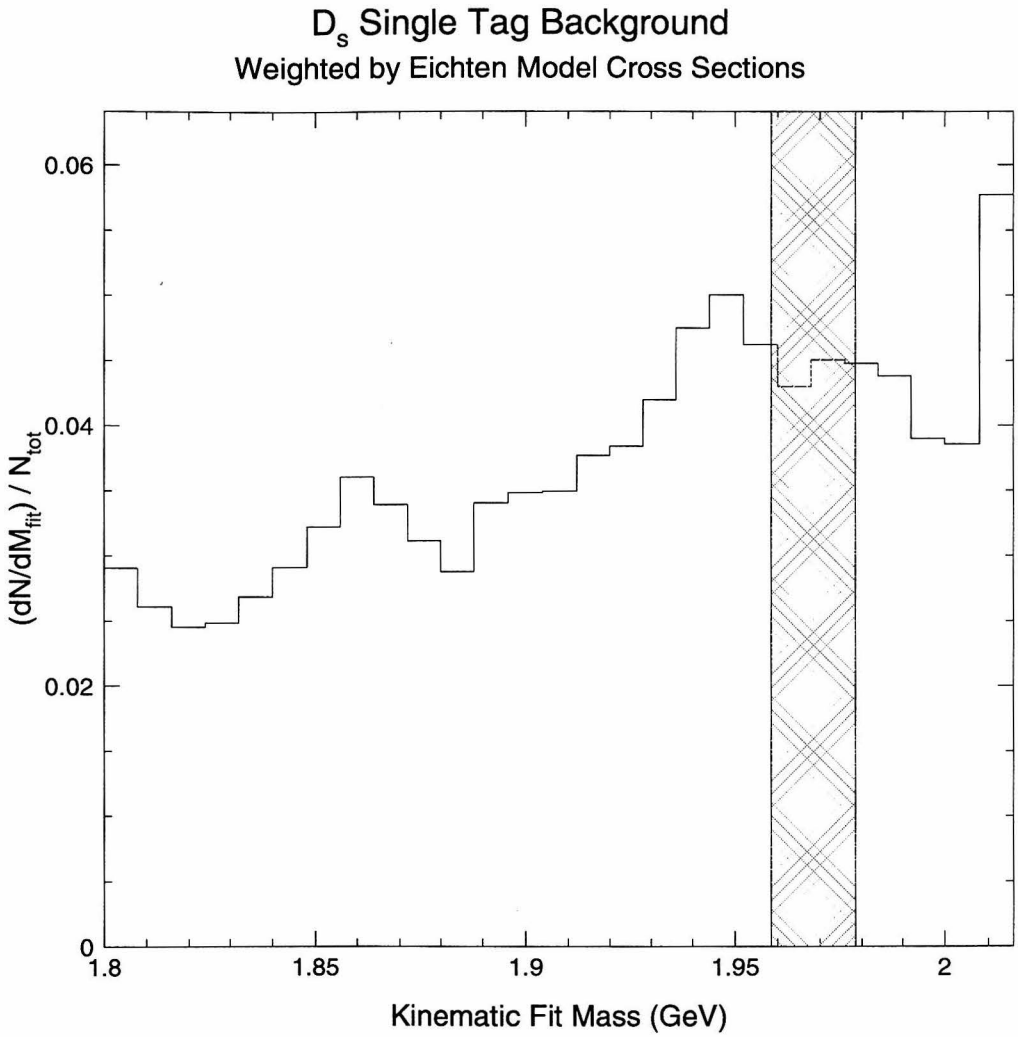

Figure 8.14: Background shape predicted for D_s candidate events from Monte Carlo data. The shape is the sum of the distributions shown in Figure 8.13, normalized to the Monte Carlo sample size and weighted according to the charm production cross sections of the coupled-channel model (Table 8.2). The hatched region is the mass range for the final candidate event selection, $1958.5 < m_{fit} < 1978.5$ MeV/c^2.

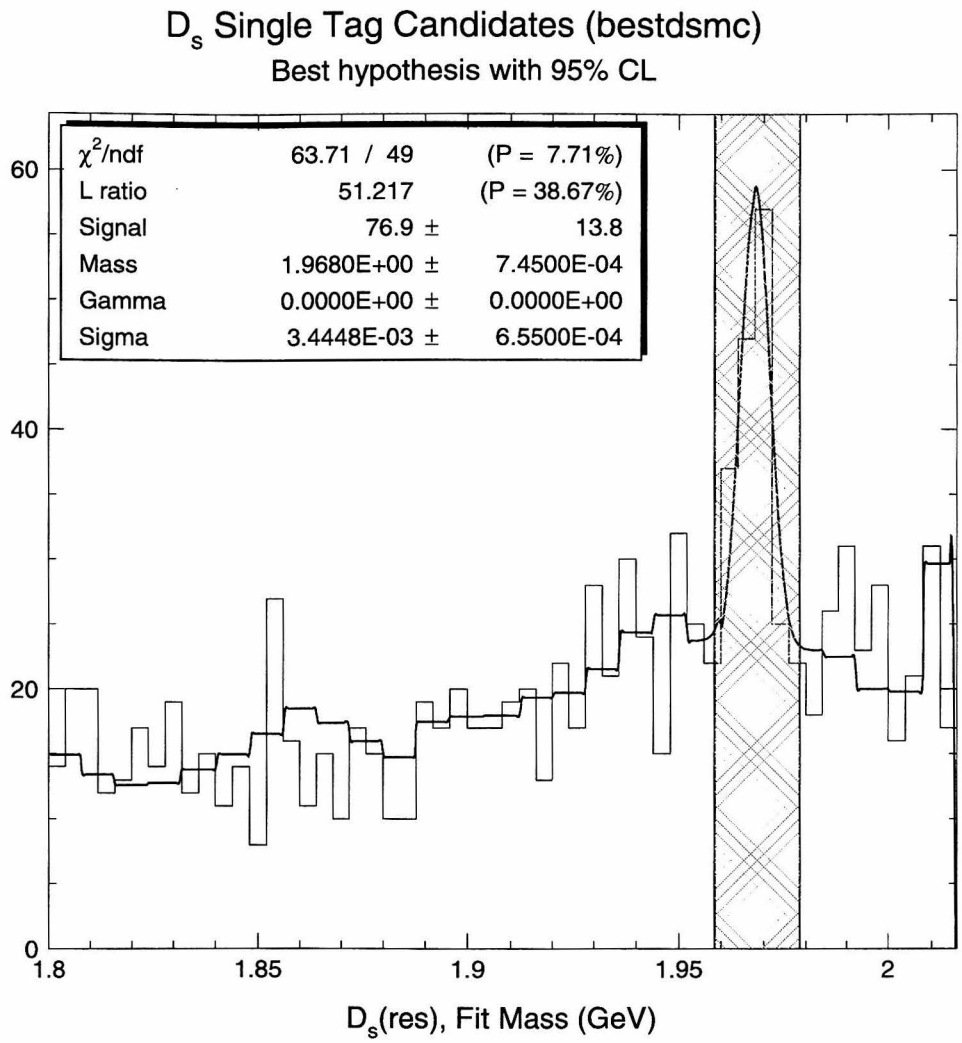

Figure 8.15: Unbinned fit of D_s candidate events to the weighted-sum background shape shown in Figure 8.14. The D_s signal is fit to a Breit-Wigner line shape convoluted with a Gaussian, and the overall background level is the only other free parameter. The hatched region is the mass range for the final candidate event selection, $1958.5 < m_{fit} < 1978.5$ MeV/c^2.

8.8.2 The Fold Matrix

A general technique for determining efficiencies and background levels in a system with multiple signals is the *fold matrix*. The fold matrix maps the rates of generated channel mapping to observed (reconstructed) events. I follow the simulated data through reconstruction and analysis, on an event-by-event basis. By keeping track of which state was generated and which state was finally assigned by my analysis, I can determine unambiguously the efficiency, the rates of background contributions, and the rates of cross-talk for each signal state. The fold matrix for my D_s tagging analysis is shown in Table 8.3. Note that the matrix is normalized to the number of events generated in each channel. I can determine the actual number of signal and background events by multiplying the fold matrix by a column vector of production rates (absolute or relative) of each channel. The product vector will give the predicted number of events (or fractions) for each final state.

Table 8.3: Fold matrix for D_s event tagging. Rows indicate generated (Monte Carlo) sources of events; columns indicate reconstructed signal channels. The values shown are percentages (%); round-off errors may affect the last digit. Matrix elements are normalized to the number of generated events in each source, so that they are true rates of efficiency or misidentification.

$\phi\pi$	$K^{*0}K$	K^0_SK	Combined	(in %)
7.685	.000	.000	7.685	$\phi\pi$
.000	5.306	.000	5.306	$K^{*0}K$
.000	.000	2.519	2.519	K^0_SK
.007	.021	.010	.039	$D^0\overline{D}^0$
.002	.019	.008	.030	$D^{*0}D^0$
.001	.029	.004	.034	$D^{*0}\overline{D}^{*0}$
.006	.018	.006	.029	D^+D^-
.002	.028	.009	.039	$D^{*\pm}D^{\mp}$
.004	.031	.003	.038	$D^{*+}D^{*-}$
.001	.017	.009	.028	$q\overline{q}$
.000	.003	.002	.006	$\tau^+\tau^-$
.018	.072	.006	.095	D_s bkg

8.8.3 Selection Efficiency

The fold matrix (Table 8.3) shows that there is actually no cross-talk between the three D_s decay channels $\phi(K^+K^-)\pi^+$, $\overline{K}^{*0}(K^+\pi^-)K^+$, and $K^0_S(\pi^+\pi^-)K^+$. Thus, the selection efficiencies may be read from the diagonal of the first three rows of the fold matrix,

$$\begin{aligned} \varepsilon(\phi\pi) &= 21.60\% \\ \varepsilon(K^{*0}K) &= 10.17\% \\ \varepsilon(K^0_S K) &= 7.66\% \end{aligned}$$

where the secondary branching fractions ($\mathcal{B}(A)$ in Table 8.1 are factored out. Combining these values with the branching fractions in Table 8.1, I find that the overall efficiency for selecting D_s events is $\varepsilon_{\mathrm{tag}} = 0.27\%$. With the 21.9 pb^{-1} of "good dE/dx" used in my analysis, my fitted result of 74.7 ± 14.6 events corresponds to a production cross section $\sigma(e^+e^- \to D_s^+ D_s^-) = 1.08 \pm 0.21$ nb, in good agreement with the prediction of the coupled-channel model [28] (Appendix A).

Chapter 9 D_s Leptonic Decay Selection

Given the sample of tagged $e^+e^- \to D_s^+ D_s^-$ events, I search the *recoil* (those tracks not used for the tag) for events containing a single charged track and no significant neutral energy. The events which pass this selection are identified as potential D_s leptonic decay events in one of the channels $D_s \to \mu\nu_\mu$ or $D_s \to \tau\nu_\tau$ ($\tau \to \mu\overline{\nu}_\mu\nu_\tau$, $\tau \to e\overline{\nu}_e\nu_\tau$ or $\tau \to \pi\nu_\tau$) which are shown in Figure 9.1. The two channels $D_s \to e\nu_e$ and $D_s \to \tau\nu_\tau, \tau \to K\nu_\tau$ are not considered, as their branching fractions (10^{-7} and 10^{-3}, respectively) imply that we should observe much less than one event in the BES data sample, even with 100% efficiency.

As discussed in Chapter 2, we expect the branching fraction $\mathcal{B}(D_s \to \mu\nu_\mu)$ to be a few 10^{-3}, and the total $\mathcal{B}(D_s \to \tau\nu_\tau)$ a few percent. If I use a "reasonable" value for f_{D_s} of 300 MeV (c.f. Chapter 12), I can estimate the number of events in my candidate final states, given the size of my $D_s^+ D_s^-$ event sample, as shown in Table 9.1.

Table 9.1: Estimated number of events in each leptonic decay final state, using a theoretical value of $f_{D_s} = 300$ MeV, and 76.9±13.8 identified $D_s^+ D_s^-$ events. These estimates do not include detector acceptance or efficiency issues.

Final State	$\mathcal{B}(D_s \to \ell\nu_\ell)$ (%)	$\mathcal{B}(\tau \to X)$ (%)	$\mathcal{B}(tot)$ (%)	$N_{\rm exp}$ (of 76.9 ± 13.8)
$\mu\nu_\mu$	0.72	—	0.72	0.55 ± 0.10
$e\nu_e\nu_\tau$	6.99	17.88	1.25	0.96 ± 0.17
$\mu\nu_\mu\nu_\tau$		17.46	1.22	0.94 ± 0.17
$\pi\nu_\tau$		11.3	0.79	0.61 ± 0.11
Total			3.98	3.06 ± 0.55

A sample size of three events (assuming 100% efficiency!) is not sufficient for a high precision measurement. Even with an unreasonably high value of f_{D_s}, say 500 MeV, I could still only expect to find seven or eight events. In order to achieve maximum

sensitivity from these limited statistics, I will not attempt to uniquely identify each candidate event as coming from one or another of the leptonic decay channels. Instead, as I will show in Chapter 10, I compute for each event a likelihood that it represents each of the decay channels, or background. I then extract the branching fractions, and the D_s decay constant, by maximizing the total likelihood for the set of events as a function of those branching fractions. In the remainder of this chapter, I describe the physics characteristics and experimental signals of D_s leptonic decays and show the candidate event distributions after event selection.

9.1 D_s Leptonic Decay Kinematics

D_s leptonic decays are characterized by low multiplicity (few charged tracks), and missing four-momentum carried by one or more final-state neutrinos, as shown in Figure 9.1. The kinematic variable I use to identify leptonic decays is the missing mass squared, $m^2_{\rm miss}$, defined by

$$\begin{aligned} m^2_{\rm miss} &\equiv (\sum p_\nu)^\alpha (\sum p_\nu)_\alpha \\ &= (p_{D_s} - p_h)^\alpha (p_{D_s} - p_h)_\alpha \\ &= (E_{beam} - E_h)^2 - (\vec{p}_{D_s} - \vec{p}_h)^2 \end{aligned} \tag{9.1}$$

where h is the observed charged track from the leptonic decay.

In Figure 9.2, I show the expected missing mass distributions from each of the leptonic decay final states, derived from special Monte Carlo data sets of 30,000 $D_s \to \mu\nu_\mu$ and 30,000 $D_s \to \tau\nu_\tau$ decays. In each event, one of the D_s mesons decays into one of the single-tag hadronic final states (10,000 of each in each data set), while the other D_s decays leptonically. For muonic decays, $D_s \to \mu\nu_\mu$ (Figure 9.1*(a)*), there is a single, isolated muon track (which is monochromatic in the D_s rest frame), and the missing mass is consistent with zero (up to the resolution of the BES detector), as shown in Figure 9.2.

The leptonic decays involving tau leptons are more complex, since the τ decay

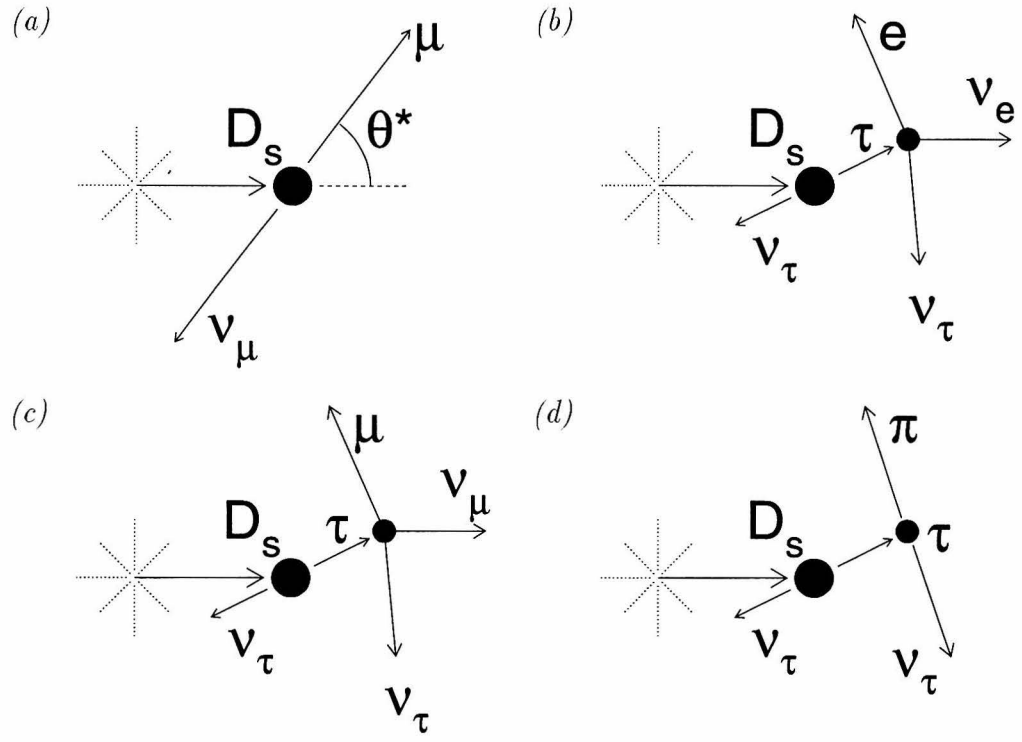

Figure 9.1: Kinematics of $D_s \to \ell\nu_\ell$ decays to single-prong final states. $D_s \to \mu\nu_\mu$ *(a)* has a monochromatic muon in the D_s rest frame, and zero missing mass. All $D_s \to \tau\nu_\tau$ channels (*(b)-(d)*) have non-zero missing mass, since at least two neutrinos are produced in the final state. $D_s \to \tau\nu_\tau$, $\tau \to e\overline{\nu}_e\nu_\tau$ *(b)* and $D_s \to \tau\nu_\tau$, $\tau \to \mu\overline{\nu}_\mu\nu_\tau$ *(c)* have three neutrinos, and a broad distribution in missing mass. $D_s \to \tau\nu_\tau$, $\tau \to \pi\nu_\tau$ *(d)* has only two neutrinos, and the pion is monochromatic in the τ rest frame; the missing mass distribution is an asymmetric peak at about 800 MeV/c^2.

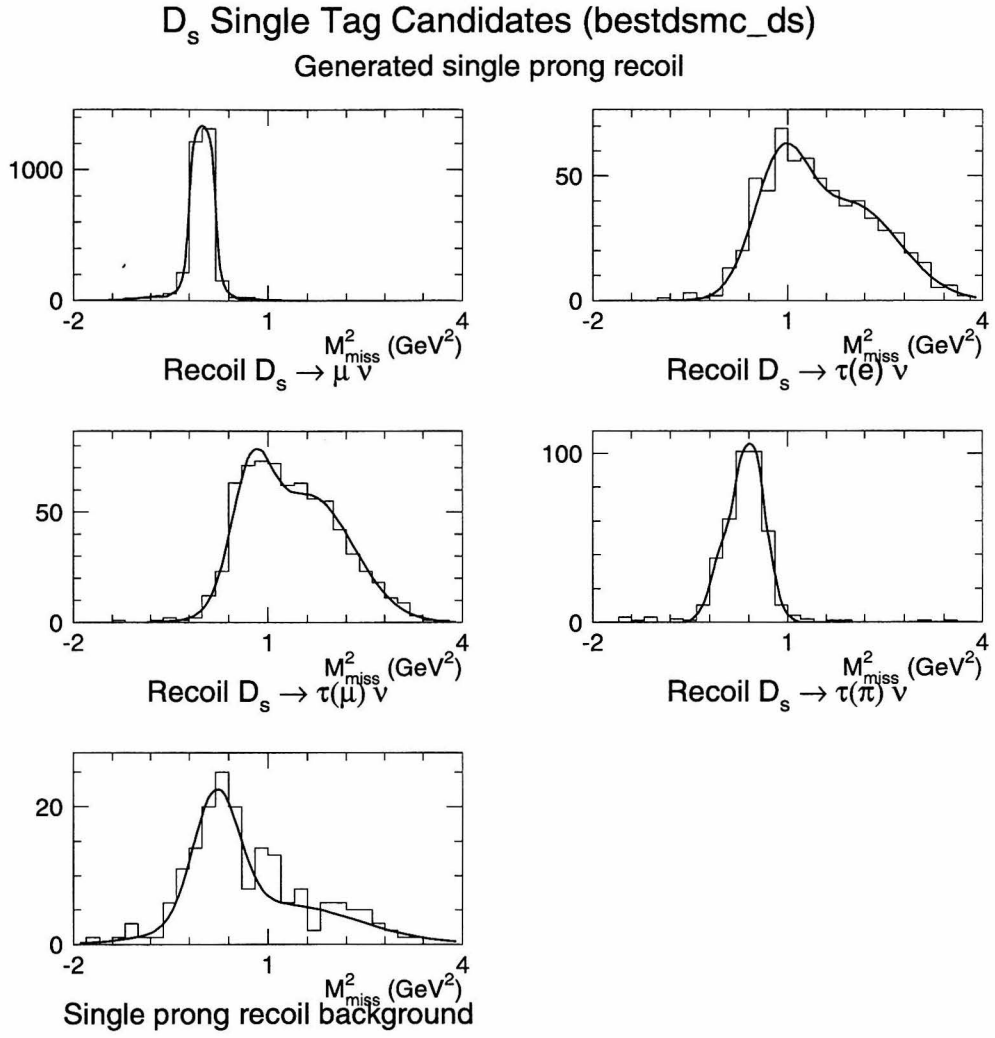

Figure 9.2: Distributions of missing mass squared ($m^2_{\rm miss}$) in D_s leptonic decay channels and expected backgrounds. In each figure, the histogram indicates the number of events selected from Monte Carlo simulated decays; the smooth curve is a six-parameter fit of the data to a pair of Gaussians, as described in Chapter 10.

produces additional neutrinos in the final state. For the leptonic final states, $D_s \to \tau\nu_\tau$ with $\tau \to e\overline{\nu}_e\nu_\tau$ or $\tau \to \mu\overline{\nu}_\mu\nu_\tau$ (Figures 9.1*(b)* and *(c)*), the charged leptons are produced over the full kinematic range, and the missing mass distributions span the available phase space, from 0 to 4 $\mathrm{GeV}^2/\mathrm{c}^4$.

The pionic final state, $D_s \to \tau\nu_\tau$, $\tau \to \pi\nu_\tau$ (Figure 9.1*(d)*) has only two neutrinos, and the pion is monochromatic in the tau rest frame. The missing mass distribution for this final state is narrow, like the muonic decay, but is significantly non-zero.

The background distribution in Figure 9.2 is the inclusive distribution of tagged events in a Monte Carlo simulation of D_s decays, with a single recoil track. For this inclusive distribution, I did not identify or reject isolated photons (see the discussion below, Section 9.2). For comparison with the Monte Carlo distributions, I show in Figure 9.3, with the same binning, the distribution of missing mass in the sample of candidate D_s events.

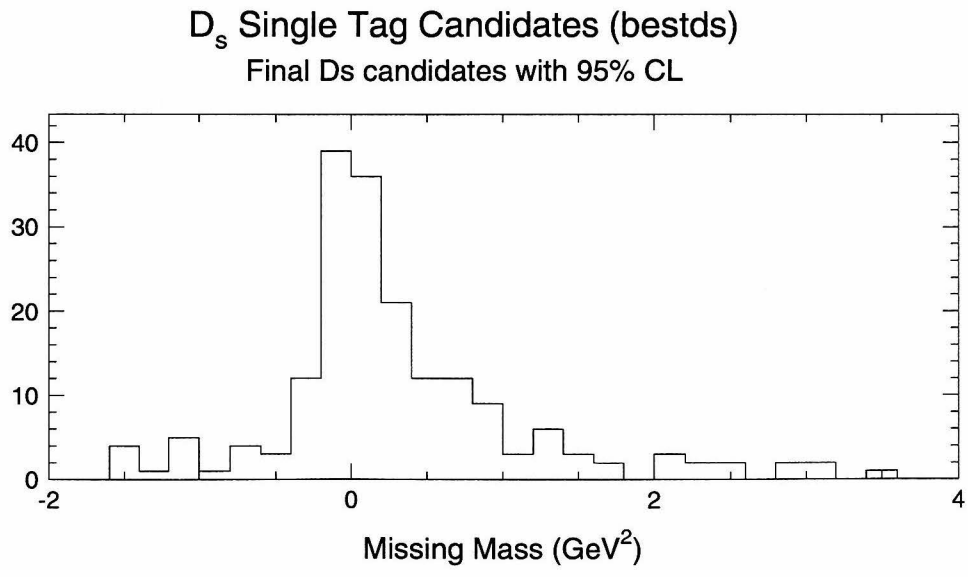

Figure 9.3: Distribution of missing mass squared (m^2_{miss}) in the D_s event sample, with the same binning as the Monte Carlo distributions (Figure 9.2).

9.2 Recoil Event Selection

I select candidate events for D_s leptonic decays from the recoils of the identified tagged D_s production events. My selection criteria are designed to select all of the candidate channels with equal efficiency, while rejecting non-leptonic background sources.

1. The recoil contains exactly one reconstructed charged track, in the central region of the detector ($|\cos\theta| \leq 0.75$).

2. The event contains no isolated neutral tracks (photons, π^0's, K_L^0's). I define an "isolated neutral" as a reconstructed cluster of hits in the shower counter (BSC) which is separated from the nearest charged track by at least 18 degrees ($\cos\theta_{\gamma q} < 0.95$), and from other neutral clusters by at least the same amount.

3. The missing momentum in the event projects into the central region of the detector ($|\cos\theta_{miss}| \leq 0.75$). This reduces dramatically the number of non-leptonic decay events with "lost" tracks which could feed into my candidate sample.

I find six events consistent with these criteria in my D_s event sample. The properties of these events (tag channel, $m^2_{\rm miss}$, and particle ID likelihoods, including the "best" recoil track ID) are listed in Table 9.2, and the "distribution" of missing mass is shown in Figure 9.4. The individual events are depicted in Figures 9.5 through 9.9.

9.3 Leptonic Event Selection Studies

I determine the efficiency for selection of leptonic decay events using special Monte Carlo simulations of $D_s^+ D_s^-$ events, with one of the D_s mesons decaying to the three hadronic tag states (10,000 events each), and the other D_s decaying to the leptonic final states; for the $D_s \to \tau\nu_\tau$ samples, the τ is allowed to decay according to the world average branching fractions [1], and I select the channels of interest afterward.

Table 9.2: Characteristics of the candidate $D_s \to \ell\nu_\ell$ events selected from the D_s event sample. The tagged D_s decay channel and recoil track identification are determined using the methods described in Chapters 8 and 7, and are provided for reference. The other properties, m^2_{miss} and the recoil track ID likelihoods, will be used in the likelihood function to determine $\mathcal{B}(D_s \to \ell\nu_\ell)$ and f_{D_s}, in Chapters 10 and 11.

BES Run Event	Tag Tracks	m^2_{miss} (GeV^2)	$\cos\theta_{\text{miss}}$	Recoil Track	Recoil ID $\mathcal{L}(e,\mu,\pi,K)$	
3273	$K^{*0}K$	-.093	-.338	μ	.0000	.9656
14523	2, 3, 4			1	.0037	.0063
3382	$K^0_S K$	3.082	.680	π	.0004	.0032
24240	1, 3, 4			2	.9964	.0000
6150	$K^{*0}K$	-.053	.175	K	.0000	.2586
13927	1, 2, 4			3	.2089	.5325
6233	$K^0_S K$	.381	.418	μ	.0000	.5639
14185	1, 3, 4			2	.4298	.0063
6274	$K^0_S K$	1.570	.053	e	.9921	.0042
17978	1, 3, 4			2	.0029	.0008
6543	$K^{*0}K$	1.109	-.650	μ	.0651	.4482
14314	2, 3, 4			1	.2992	.1876

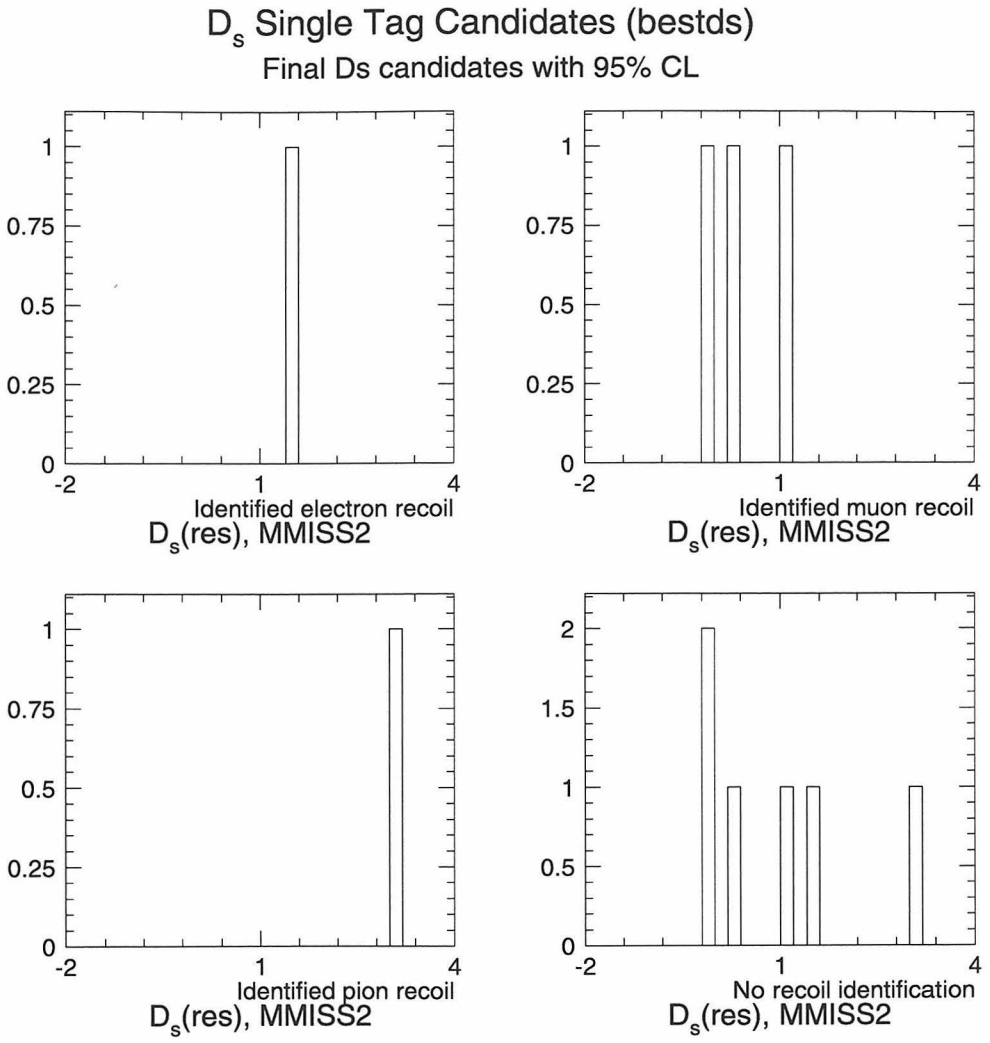

Figure 9.4: Distribution of missing mass squared ($m^2_{\rm miss}$) for the six candidate $D_s \to \ell\nu_\ell$ events described in the text, with the recoil track identified by maximum likelihood as an electron *(upper left)*, muon *(upper right)*, or pion *(lower left)*. The plot at lower right shows the combined distribution for all the candidate events. The binning of these plots are the same as that used for the Monte Carlo distributions (Figure 9.2).

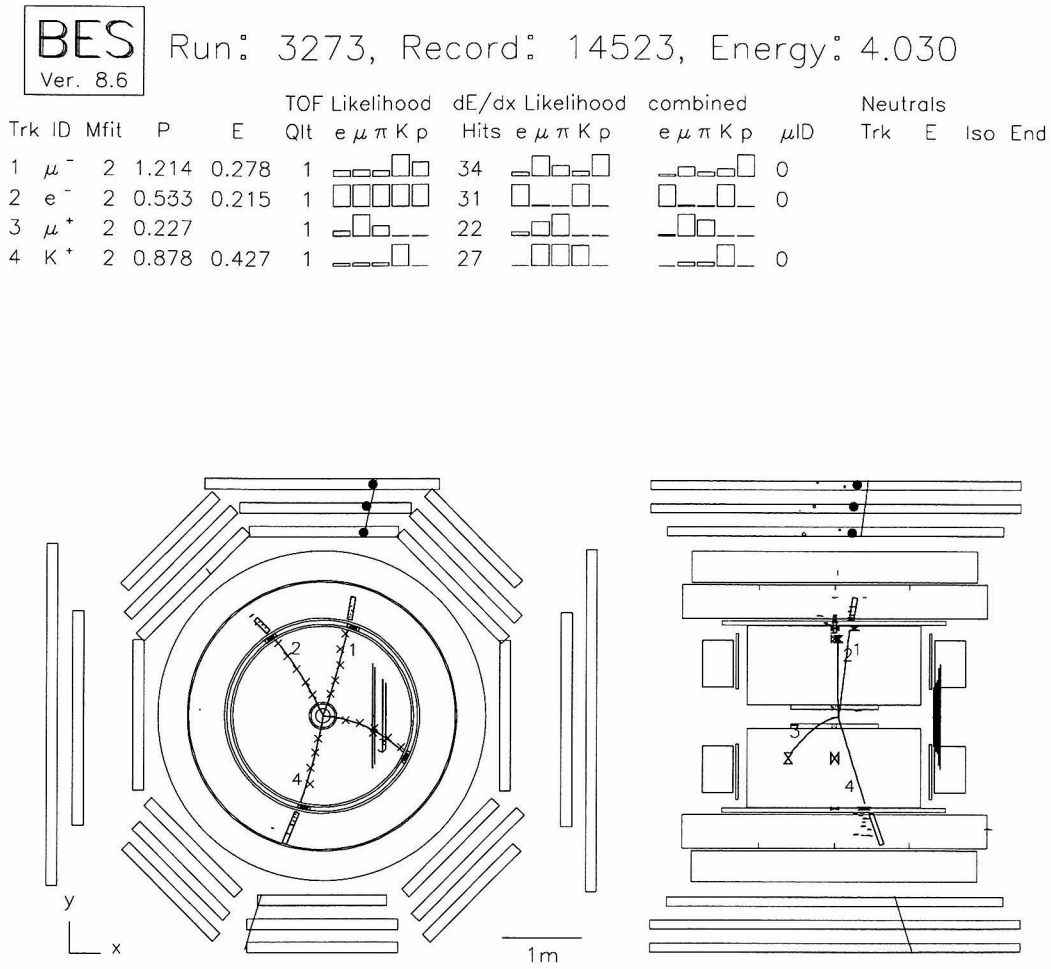

Figure 9.5: BES event display depicting candidate leptonic decay event from run 3273, record 14523. Tracks 2, 3, and 4 are identified as a $D_s \to \overline{K}^{*0}(K^+\pi^-)\,K^+$ tag, and track 1 is identified as a muon by maximum likelihood.

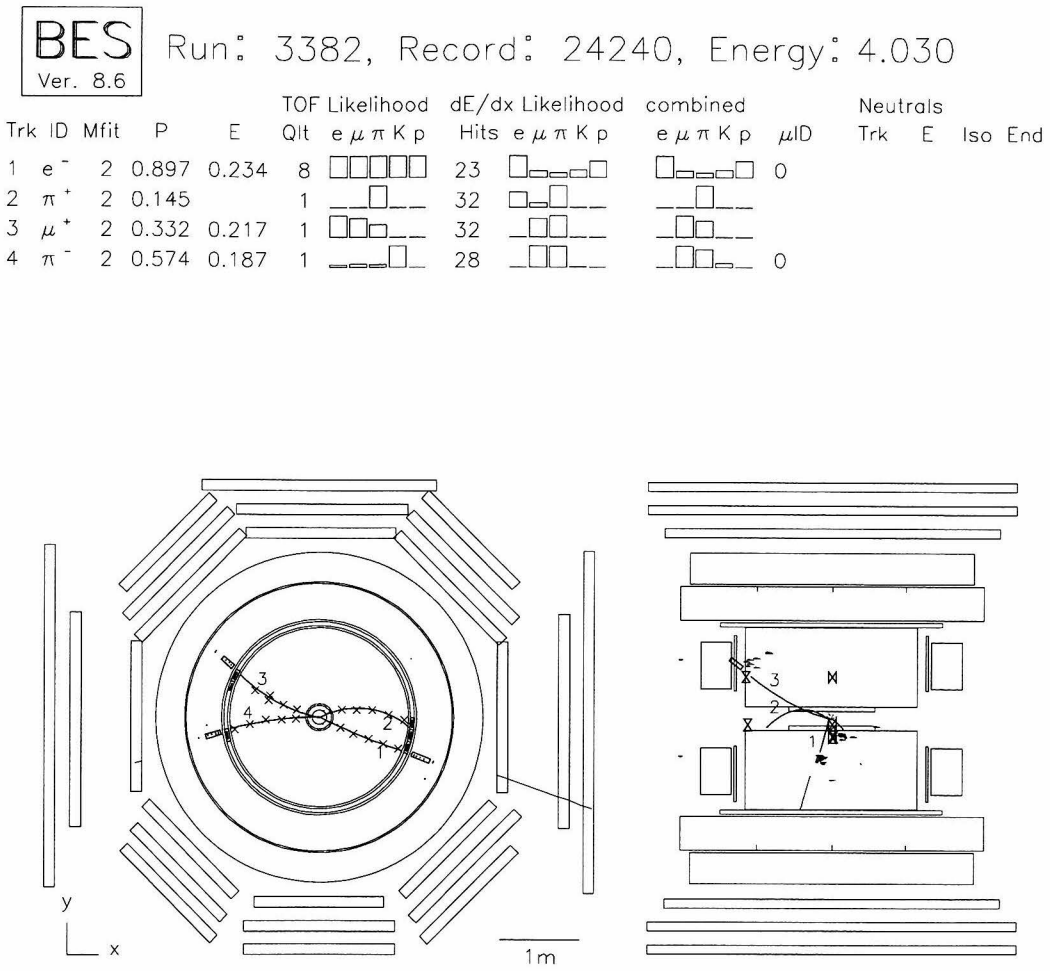

Figure 9.6: BES event display depicting candidate leptonic decay event from run 3382, record 24240. Tracks 1, 3, and 4 are identified as a $D_s \to K^0_S(\pi^+\pi^-)\,K^+$ tag, and track 2 is identified as a pion by maximum likelihood.

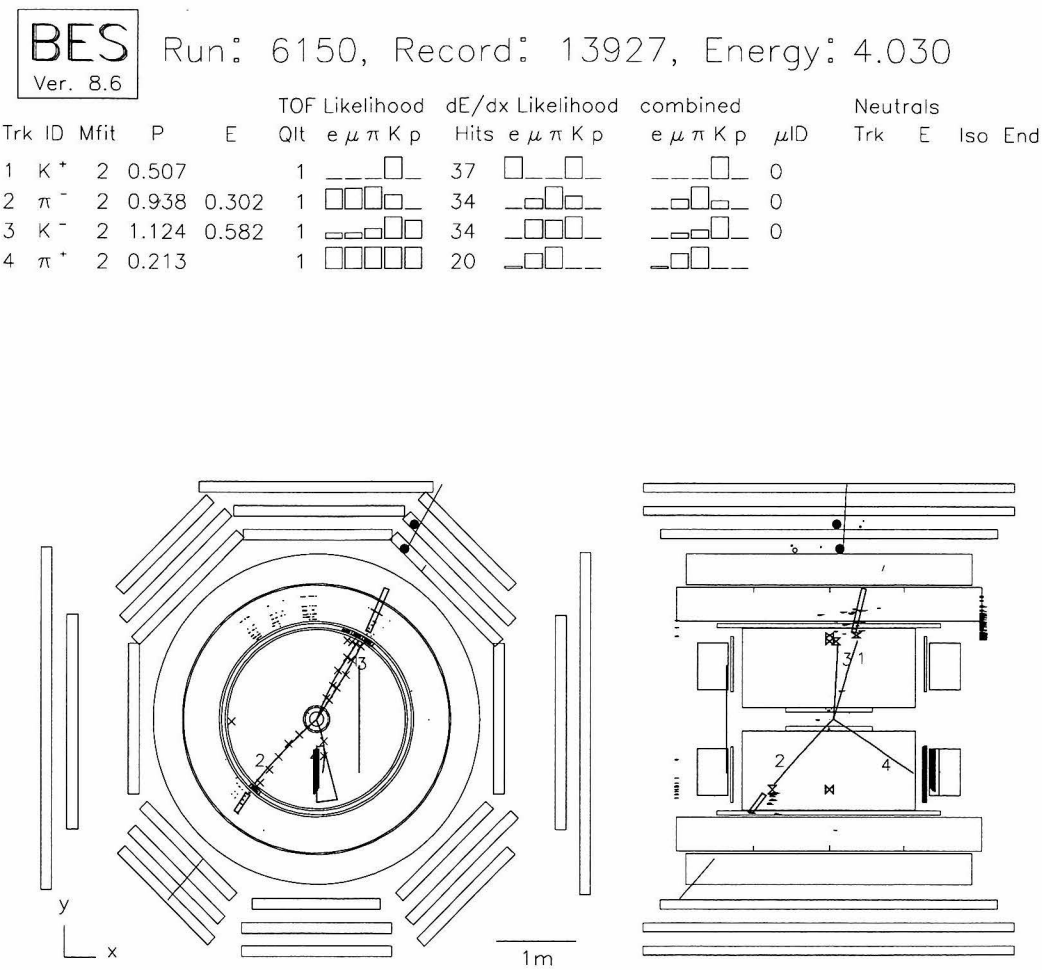

Figure 9.7: BES event display depicting candidate leptonic decay event from run 6150, record 13927. Tracks 1, 2, and 4 are identified as a $D_s \to \overline{K}^{*0}(K^+\pi^-)\,K^+$ tag, and track 3 is identified as a kaon by maximum likelihood.

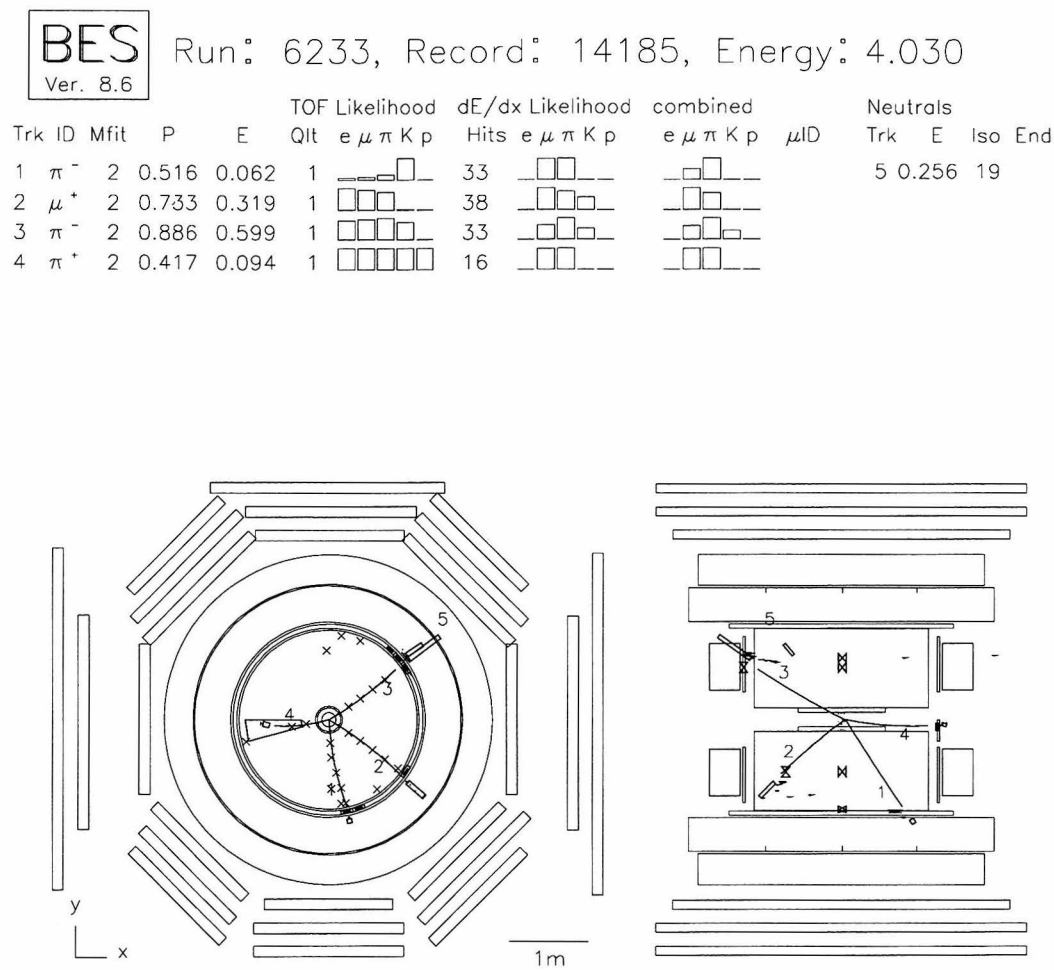

Figure 9.8: BES event display depicting candidate leptonic decay event from run 6233, record 14185. Tracks 1, 3, and 4 are identified as a $D_s \to K^0_S(\pi^+\pi^-)\,K^+$ tag, and track 2 is identified as a muon by maximum likelihood.

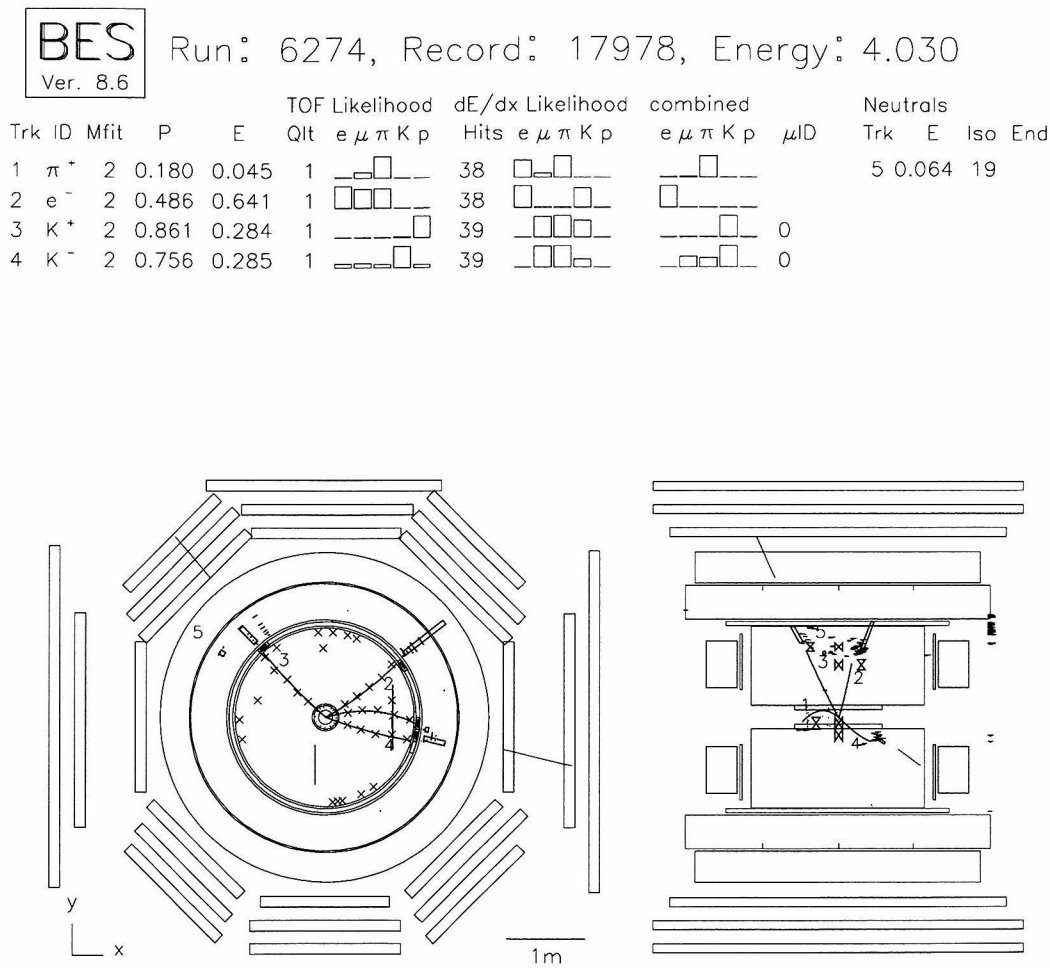

Figure 9.9: BES event display depicting candidate leptonic decay event from run 6274, record 17978. Tracks 1, 3, and 4 are identified as a $D_s \to K^0_S(\pi^+\pi^-)\,K^+$ tag, and track 2 is identified as an electron by maximum likelihood.

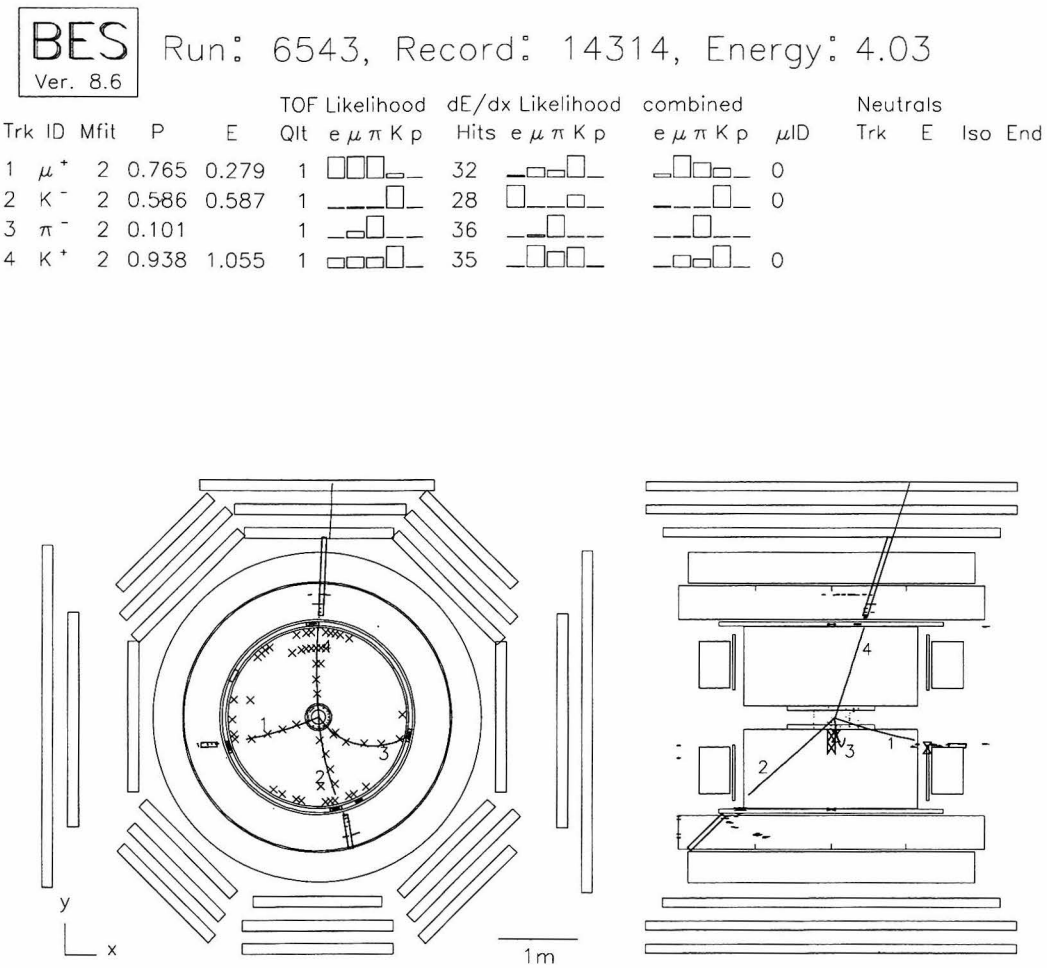

Figure 9.10: BES event display depicting candidate leptonic decay event from run 6543, record 14314. Tracks 2, 3, and 4 are identified as a $D_s \to \overline{K}^{*0}(K^+\pi^-)\,K^+$ tag, and track 1 is identified as a muon by maximum likelihood.

I apply my full analysis (Chapter 8 and this one) to the simulated events, and "identify" the recoil track in each event using the maximum likelihood value of the particle ID function (Chapter 7). The selection efficiencies are just the fraction of events in each final state for which my identification of the recoil track matches that generated,

$$\begin{aligned} \varepsilon(\mu\nu) &= 58.20\% \\ \varepsilon(\tau \to e) &= 58.59\% \\ \varepsilon(\tau \to \mu) &= 54.09\% \\ \varepsilon(\tau \to \pi) &= 46.75\% \end{aligned}$$

with statistical errors of about $\pm 1\%$, based on my Monte Carlo statistics.

I determine the background levels in a similar way, using the other Monte Carlo data sets described in Chapter 5. For the non-D_s initial states, I obtain the background rates directly as the fraction of D_s-tagged events (themselves fairly small) for which I find a single recoil track with the appropriate maximum-likelihood identity. For the D_s "cocktail" data set, I measure the backgrounds in the same way, but I cross-check the generated recoil state, and do not count events (roughly 100 in the sample) which are correctly identified as $D_s \to \mu\nu_\mu$ or $D_s \to \tau\nu_\tau$.

Table 9.3 shows the background rates for each recoil identity from the various sources of tagged events (the zero entries indicate that no tagged event satisfied my recoil selection criteria, listed in Section 9.2). Weighting these values by the production cross sections (Table 8.2), I find the inclusive background rates for my leptonic decay event selection to be

$$\begin{aligned} \eta(e) &= .05\% \\ \eta(\mu) &= .12\% \\ \eta(\pi) &= .53\% \end{aligned}$$

$$\eta(tot) \;=\; .79\%$$

with statistical errors of approximately $\pm 10\%$, based on the number of tagged events in the Monte Carlo data sets. With 190 candidate $D_s^+ D_s^-$ events, this corresponds to approximately 1.51 background events out of my sample of six leptonic candidates.

Table 9.3: Recoil track identification background rates, obtained from Monte Carlo data sets. Zeros indicate that no tagged events satisfied my recoil selection criteria (Section 9.2). Rates are normalized to the number of tagged events in each data set. Inclusive rates for each track identification are found by weighting these values by the production cross sections of the initial states.

e ID	μ ID	π ID	Inclusive	(%)
.00	.00	.00	.20	$D^0\overline{D}^0$
.00	.11	.11	.43	$D^{*0}D^0$
.00	.17	.00	.17	$D^{*0}\overline{D}^{*0}$
.00	.00	.16	.16	D^+D^-
.09	.27	.37	.73	$D^{*\pm}D^{\mp}$
.27	.27	.00	.81	$D^{*+}D^{*-}$
.00	.00	.24	.24	$q\overline{q}$
.00	.00	2.60	2.60	$\tau^+\tau^-$
.53	1.15	1.86	4.96	D_s bkg

These background rates include both "topological" backgrounds, track combinations incorrectly identified as D_s hadronic decays, which have isolated single recoil tracks; and "identification" backgrounds, true $D_s^+ D_s^-$ events, with single charged tracks in the recoil, but for which the identification assignment is not correct. I incorporate both the efficiencies and the background rates into the final likelihood function in the next chapter, as factors in the weights assigned to each event.

Part IV

Results

Chapter 10 Likelihood Function for Leptonic Decays

In this thesis, I measure the branching fractions of the D_s meson to the leptonic final states $\mu\nu$ and $\tau\nu$, and extract the D_s decay constant f_{D_s} from these measurements. To obtain these results with maximum sensitivity, I define a likelihood function which depends on these parameters, and evaluate that function for my sample of candidate leptonic decays, maximizing the function with respect to the branching fractions (or equivalently, to f_{D_s}).

10.1 Likelihood Expression

For my sample of candidate D_s leptonic decay events, I define the likelihood to obtain those particular events to be

$$\mathcal{L} = P(N_{\text{events}}; N_{\text{exp}}) \prod_{i}^{\text{events}} \mathcal{L}_i^{\text{lept}}(m^2_{\text{miss},i}, \vec{\mathcal{L}}_i^{PID}) \tag{10.1}$$

where N_{exp} is the number of leptonic decay events expected, and

$$\mathcal{L}_i^{\text{lept}}(m^2_{\text{miss},i}, \vec{\mathcal{L}}_i^{PID}) = \sum_{m}^{\text{channels}} W_m \mathcal{L}_{m,i}^{PID} \mathcal{L}_m(m^2_{\text{miss},i}) \tag{10.2}$$

is the likelihood that event i is due to a leptonic D_s decay, or is background. The subscript $m \in \{D_s \to \mu\nu_\mu; D_s \to \tau\nu_\tau, \tau \to e\overline{\nu}_e\nu_\tau; D_s \to \tau\nu_\tau, \tau \to \mu\overline{\nu}_\mu\nu_\tau; D_s \to \tau\nu_\tau, \tau \to \pi\nu_\tau$ or background$\}$ covers both the final state leptonic decay channels, as well as the inclusive background rate. In the case of background, the particle identification likelihood is unity, since the background sums over all possible identifications.

I equation 10.1, $P(N_{\text{events}}; N_{\text{exp}})$ is the Poisson probability of the number of observed candidates N_{events}, given a number N_{exp} of events expected. The number of

expected events depends on the total number of candidate D_s events, the branching fractions to each final state, and the background rate estimated from Monte Carlo data,

$$N_{\rm exp} = N_{D_s} \sum_m \eta_m \tag{10.3}$$

$$\eta_m = \begin{cases} \varepsilon_m \mathcal{B}(m) & m \in \{\mu\nu, \tau e, \tau\mu, \tau\pi\} \\ \eta_{\rm bkg} & m \text{ background} , \end{cases} \tag{10.4}$$

and $\eta \equiv \sum_m \eta_m$ for convenience.

In Chapter 9 I presented the missing mass distributions, $\mathcal{L}_m(m^2_{\rm miss})$ (Figures 9.2), where I derived them from analyses of Monte Carlo-generated samples of leptonic decays. To simplify the use of these distributions in my analyses, I parameterize them as a pair of overlapping Gaussians,

$$\mathcal{L}_m(m^2_{\rm miss}) = H_{1m} \exp\left(-\frac{1}{2}\left(\frac{m^2_{\rm miss} - M^2_{1m}}{\sigma_{1m}}\right)^2\right) + H_{2m} \exp\left(-\frac{1}{2}\left(\frac{m^2_{\rm miss} - M^2_{2m}}{\sigma_{2m}}\right)^2\right) \tag{10.5}$$

normalized to unit area. The parameters H_{im}, M_{im}, and σ_{im} are listed in Table 10.1.

In the likelihood function (equation 10.2), I use a weighted sum of the missing-mass likelihoods for all of the possible final states. Rather than assigning each candidate event to one or another final state, with the concomitant issues of efficiency and cross-talk between channels, I let each event contribute to all possible topologies, and weight those contributions according to the number of expected events in that topology, and the probability that the charged track observed matches the particle in the final state. The weight is the the fraction of expected events in channel m, η_m/η, with η_m defined in equation 10.4. These weights ensure that each possible final state contributes proportionally to the overall likelihood for each event; otherwise the likelihood would be biased towards the $D_s \to \mu\nu_\mu$ and $D_s \to \tau\nu_\tau$,$\tau \to \pi\nu_\tau$ channels, which have narrow and high missing mass distributions (see Figure 9.2 in Chapter 9).

An important feature of my analysis is my use of the particle identification like-

Table 10.1: Parameters of the missing mass likelihood function, equation 10.5.

$D_s \rightarrow$	H_1	M_1	σ_1
$\mu\nu$	$6.51 \pm .16$	$.00 \pm .00$	$.13 \pm .00$
$\tau \rightarrow e$	$1.05 \pm .22$	$.89 \pm .09$	$.42 \pm .06$
$\tau \rightarrow \mu$	$.94 \pm .14$	$.74 \pm .05$	$.29 \pm .04$
$\tau \rightarrow \pi$	$3.64 \pm .42$	$.41 \pm .05$	$.23 \pm .03$
Bkg	$1.40 \pm .24$	$.21 \pm .06$	$.34 \pm .07$

	H_2	M_2	σ_2
$\mu\nu$	$.22 \pm .03$	$-.15 \pm .03$	$.57 \pm .03$
$\tau \rightarrow e$	$.78 \pm .11$	$2.01 \pm .22$	$.71 \pm .11$
$\tau \rightarrow \mu$	$1.05 \pm .09$	$1.62 \pm .07$	$.69 \pm .03$
$\tau \rightarrow \pi$	$.92 \pm .56$	$-.06 \pm .10$	$.16 \pm .05$
Bkg	$.44 \pm .10$	$1.17 \pm .26$	$1.18 \pm .17$

lihoods (Chapter 7) as part of my determination of f_{D_s}. The particle identification vector, $\vec{\mathcal{L}}^{PID}$, is the set of normalized weights returned by my `TRACKWT` function (see Appendix B). The component of $\vec{\mathcal{L}}^{PID}$ used in equation 10.2 is just the charged particle species in the leptonic final state m.

In Figures 10.1 and 10.2, I show the form of this weighted sum, equation 10.2, evaluated for each of the six leptonic events in my sample, as functions of $\mathcal{B}(D_s \rightarrow \mu\nu_\mu)$ and f_{D_s}, respectively. The total likelihood function for $\mathcal{B}(\mu\nu)$ or f_{D_s} is the product of these shapes, multiplied by the Poisson likelihood of the six observed events *vs.* the number of events expected. I show the final likelihood results in the next chapter (Figures 11.2 and 11.3).

10.2 Statistical Uncertainties

The likelihood function, equation 10.1, is non-Gaussian by construction, and as I will show in Chapter 11, is quite asymmetric. Clearly, I should not quote symmetric

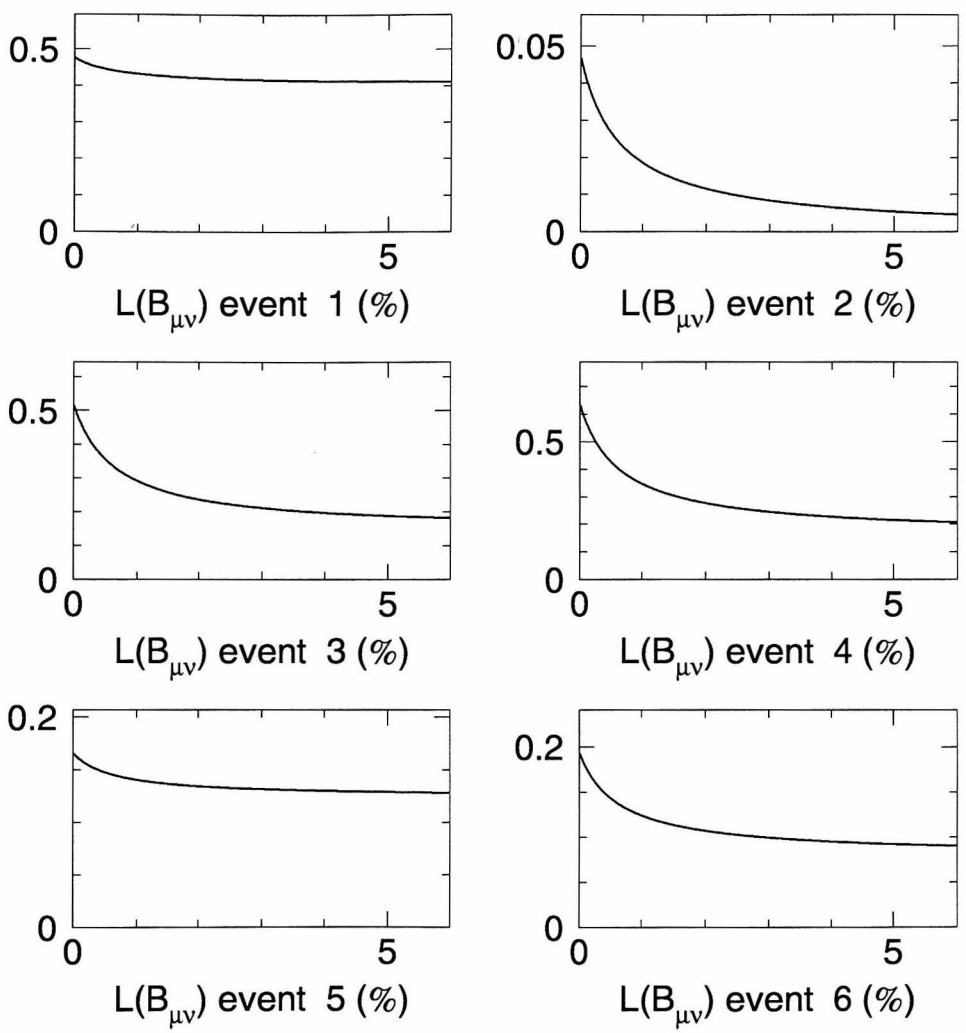

Figure 10.1: Likelihood functions for the candidate leptonic events, with respect to the branching fraction $\mathcal{B}(D_s \to \mu\nu_\mu)$. The total likelihood, equation 10.1 is the product of these six functions and the Poisson likelihood for observing the candidate events.

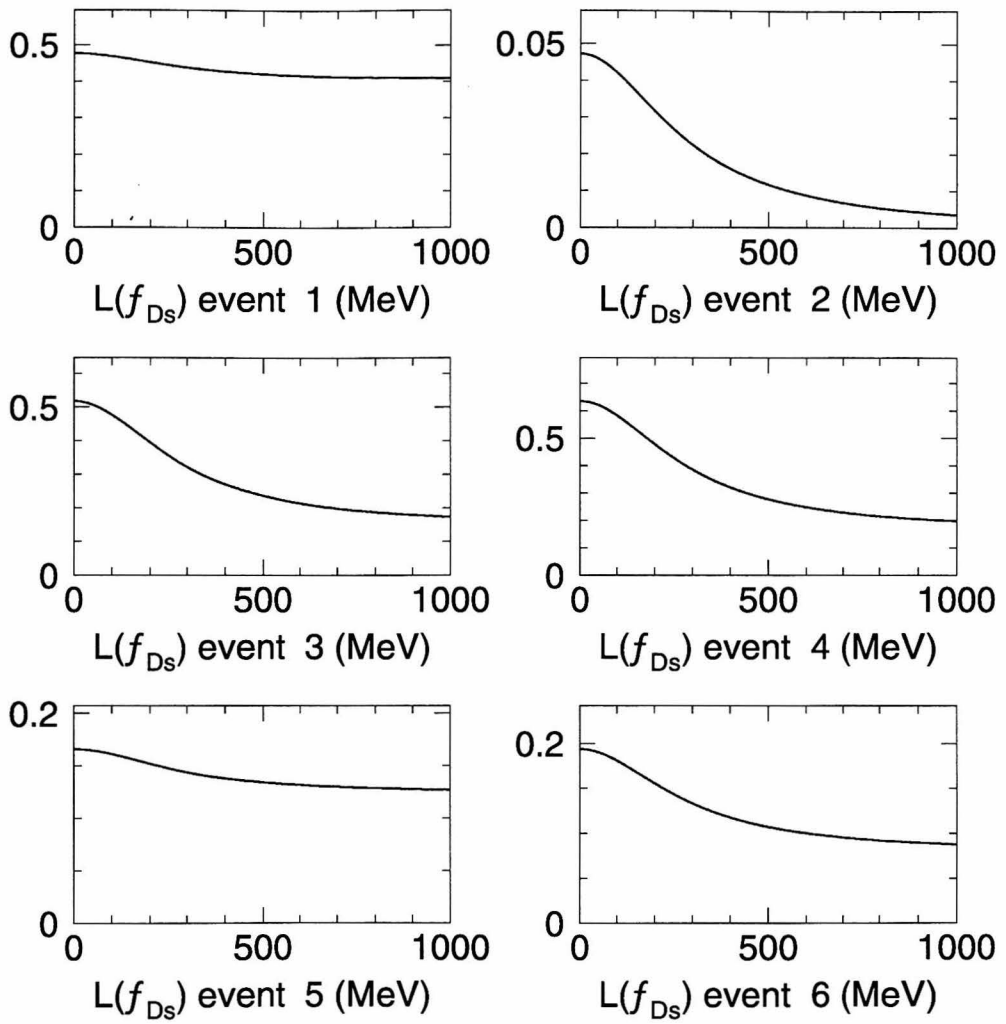

Figure 10.2: Likelihood functions for the candidate leptonic events, with respect to the decay constant f_{D_s}. The total likelihood, equation 10.1 is the product of these six functions and the Poisson likelihood for observing the candidate events.

error bars for my results, since this obscures the essentially non-Gaussian nature of the problem. For the most reasonable comparison with other, "Gaussian," results, I would like to quote a 68% confidence interval about the maximum likelihood value. Unfortunately, there is no consensus in the high energy physics community about how to define such an interval[1] [45].

For my analysis, I interpret my likelihood function as a probability density, treating my measurement as a random variable. Although there exists a "true" value of f_{D_s}, for example, which is fixed and certain, my experimental technique introduces randomness into the value I *measure*; that randomness is presumably reflected in the shape of my likelihood function, provided that function is well-constructed. Accordingly, I define a confidence interval in terms of the area A of the likelihood function, such that

1. a quoted "p% confidence interval" contains an area $Ap/100$;
2. the maximum likelihood value is contained within the quoted interval;
3. the asymmetry of the likelihood function is qualitatively reflected in the quoted positive and negative uncertainties.

Given a likelihood function $\mathcal{L}(x)$, with a single maximum-likelihood value x_0 ($\mathcal{L}(x_0) > \mathcal{L}(x)\ \forall x \neq x_0$, I will quote a result

$$x_0 {}^{+\delta x_+}_{-\delta x_-},$$

where the uncertainties δx_- and δx_+ are defined by

$$\int_{x_0-\delta x_-}^{x_0} \mathcal{L}(x)dx = p\int_{-\infty}^{x_0} \mathcal{L}(x)dx \tag{10.6}$$

$$\int_{x_0}^{x_0+\delta x_+} \mathcal{L}(x)dx = p\int_{x_0}^{+\infty} \mathcal{L}(x)dx \tag{10.7}$$

for a $(100p)$% confidence interval. For "one sigma" errors, $p = 0.6827$, which is what

[1]The BES experiment contacted several members of the Particle Data Group regarding this issue, and received several different, and sometimes conflicting, recommendations.

I will use in computing the statistical uncertainties on my results for $\mathcal{B}(D_s \to \mu\nu_\mu)$ and f_{D_s} in Chapter 11.

10.3 Systematic Uncertainties

The likelihood function, equation 10.1, is a relatively simple function of the branching fractions or f_{D_s}. They appear only in the weights applied to each event, and in the overall Poisson normalization, and so (in principle) $\mathcal{L}$ can be differentiated and solved for them algebraically. The maximum with respect to the variable X occurs at

$$\frac{\partial \mathcal{L}}{\partial X} = 0 \ .$$

If $\mathcal{L}$ is being maximized with respect to several variables (cf. Chapter 11), this becomes a system of equations to be solved simultaneously; I will not consider this case here.

Let $B \in \{\mu\nu, \tau\nu\}$ be the channel to be determined. Differentiating equation 10.1 with respect to $\mathcal{B}(B)$,

$$\frac{\partial \mathcal{L}}{\partial \mathcal{B}(B)} = \frac{\partial P}{\partial \mathcal{B}(B)} \prod_i \frac{\sum_m C_{mi}\eta_m}{\eta} + P(N_{\text{events}}; N_{\text{exp}}) \frac{\partial}{\partial \mathcal{B}(B)} \prod_i \frac{\sum_m C_{mi}\eta_m}{\eta} \ , \quad (10.8)$$

where $C_{mi} = \mathcal{L}^{PID}_{m,i}\mathcal{L}_m(m^2_{\text{miss},i})$. Again, I use i to indicate the candidate leptonic decay events, and m the signal and background channels. For the tau decay channels, $\mathcal{B}(m) \equiv \mathcal{B}(D_s \to \tau\nu_\tau)\,\mathcal{B}(\tau \to m)$. The derivative of the Poisson distribution has a simple expression,

$$\begin{aligned}
\frac{\partial}{\partial \mathcal{B}(B)} P(N;\mu) &= \frac{\partial \mu}{\partial \mathcal{B}(B)} \frac{\partial}{\partial \mu} P(N;\mu) = \mu' \frac{\partial}{\partial \mu}\left(\frac{e^{-\mu}\mu^N}{N!}\right) \\
&= \frac{\mu'}{N!}\left(e^{-\mu}N\mu^{N-1} - e^{-\mu}\mu^N\right) \\
&= \mu'\left(P(N;\mu)\frac{N}{\mu} - P(N;\mu)\right)
\end{aligned}$$

$$= P(N;\mu)\,\mu'\left(\frac{N}{\mu}-1\right) ,$$

where $\mu' \equiv \partial\mu/\partial\mathcal{B}(B)$. Substituting $N \to N_{\rm events}$ and $\mu \to N_{\rm exp} = N_{D_s}\eta$,

$$\frac{\partial}{\partial\mathcal{B}(B)}P(N_{\rm events};N_{\rm exp}) = P(N_{\rm events};N_{\rm exp})N_{D_s}\,\eta'\left(\frac{N_{\rm events}}{N_{\rm exp}}-1\right) . \tag{10.9}$$

The second term in equation 10.8 expands to

$$\begin{aligned}
\frac{\partial}{\partial\mathcal{B}(B)}\prod_i\frac{\sum_m C_{mi}\eta_m}{\eta} &= \sum_i\left[\frac{\prod_j\sum_m C_{mj}\eta_m/\eta}{\sum_m C_{mi}\eta_m/\eta}\sum_m C_{mi}\frac{\partial}{\partial\mathcal{B}(B)}\frac{\eta_m}{\eta}\right] \\
&= \sum_i\left[\frac{\prod_j\sum_m C_{mj}\eta_m/\eta}{\sum_m C_{mi}\eta_m/\eta}\sum_m C_{mi}\left(\frac{\eta'_m}{\eta}-\frac{\eta_m\eta'}{\eta^2}\right)\right] \\
&= \left(\prod_i\frac{\sum_m C_{mi}\eta_m}{\eta}\right)\left[\sum_i\frac{\sum_m C_{mi}\eta'_m}{\sum_m C_{mi}\eta_m}-N_{\rm events}\frac{\eta'}{\eta}\right] . \qquad (10.10)
\end{aligned}$$

Substituting equations 10.9 and 10.10 into equation 10.8, and using $N_{\rm exp} = N_{D_s}\eta$, I obtain

$$\begin{aligned}
\frac{\partial\mathcal{L}}{\partial\mathcal{B}(B)} &= N_{D_s}\,\eta'\left(\frac{N_{\rm events}}{N_{\rm exp}}-1\right)P(N_{\rm events};N_{\rm exp})\prod_i\frac{\sum_m C_{mi}\eta_m}{\eta} \\
&\quad + P(N_{\rm events};N_{\rm exp})\left(\prod_i\frac{\sum_m C_{mi}\eta_m}{\eta}\right)\left[\sum_i\frac{\sum_m C_{mi}\eta'_m}{\sum_m C_{mi}\eta_m}-N_{D_s}N_{\rm events}\frac{\eta'}{N_{\rm exp}}\right] \\
&= \mathcal{L}\times\left\{N_{D_s}\,\eta'\left(\frac{N_{\rm events}}{N_{\rm exp}}-1-\frac{N_{\rm events}}{N_{\rm exp}}\right)+\sum_i\frac{\sum_m C_{mi}\eta'_m}{\sum_m C_{mi}\eta_m}\right\} \\
&= \mathcal{L}\times\left(\sum_i\frac{\sum_m C_{mi}\eta'_m}{\sum_m C_{mi}\eta_m}-N_{D_s}\,\eta'\right) . \qquad (10.11)
\end{aligned}$$

Recall that $\mathcal{B}(D_s\to\ell\nu_\ell) \propto f_{D_s}^2$. It is equivalent, then, to maximize the likelihood with respect to $\mathcal{B}(D_s\to\ell\nu_\ell)$, and change variables to get the maximization with respect to f_{D_s}. Setting $\partial\mathcal{L}/\partial\mathcal{B}(B) = 0$, the likelihood is maximized when

$$\sum_i\frac{\sum_m C_{mi}\eta'_m}{\sum_m C_{mi}\eta_m} = N_{D_s}\,\eta' , \tag{10.12}$$

since $\mathcal{L} = 0$ cannot be a maximum. This expression should be solved for the ap-

propriate parameter, but in general cannot be, since the right-hand side is a rational polynomial of $\mathcal{O}(N_{\text{events}})$ in η_B. For f_{D_s}, I should formally apply the chain rule and require, *e.g.*,

$$0 = \frac{\partial \mathcal{L}}{\partial f_{D_s}} = \frac{\partial \mathcal{L}}{\partial \mathcal{B}(B)} \frac{\partial \mathcal{B}(B)}{\partial f_{D_s}} = 2A_B f_{D_s} \frac{\partial \mathcal{L}}{\partial \mathcal{B}(B)} \ ,$$

where $A_B = A_\mu$ or A_τ is the expression given in equation 2.12 involving the D_s and lepton masses, couplings, etc. This change of variables adds an extra factor of $2A_B f_{D_s}$ to both sides of equation 10.12, which will simply cancel out.

In fact, I do not actually need to solve equation 10.12 analytically to evaluate the systematic errors: I can find the total differential of both sides, and solve for $d\mathcal{B}(B)$ in terms of the differentials of all of the input parameters,

$$d\left[N_{D_s}\,\eta'\right] = d\left[\sum_i \frac{\sum_m C_{mi}\eta'_m}{\sum_m C_{mi}\eta_m}\right] \ , \tag{10.13}$$

where

$$d\left[N_{D_s}\eta'\right] = \eta'\, dN_{D_s} + N_{D_s} \sum_P \frac{\partial \eta'}{\partial P}\, dP \tag{10.14}$$

and

$$\begin{aligned} d\left[\sum_i \frac{\sum_m C_{mi}\eta'_m}{\sum_m C_{mi}\eta_m}\right] &= \sum_{i,P} dP\, \frac{\partial}{\partial P}\left[\frac{\sum_m C_{mi}\eta'_m}{\sum_m C_{mi}\eta_m}\right] \\ &= \sum_{i,P}\left[\frac{1}{\sum_m C_{mi}\eta_m} dP\, \frac{\partial}{\partial P}\sum_m C_{mi}\eta'_m - \frac{\sum_m C_{mi}\eta'_m}{(\sum_m C_{mi}\eta_m)^2} dP\, \frac{\partial}{\partial P}\sum_m C_{mi}\eta_m\right] \\ &= \sum_{i,P} \frac{1}{\sum_m C_{mi}\eta_m} \sum_m C_{mi}\left[\frac{\partial \eta'_m}{\partial P} - \frac{\sum_m C_{mi}\eta'_m}{\sum_m C_{mi}\eta_m}\frac{\partial \eta_m}{\partial P}\right] dP \ , \end{aligned} \tag{10.15}$$

so that

$$\eta'\, dN_{D_s} + N_{D_s} \sum_P \frac{\partial \eta'}{\partial P}\, dP = \sum_{i,P} \frac{1}{\sum_m C_{mi}\eta_m} \sum_m C_{mi}\left[\frac{\partial \eta'_m}{\partial P} - \frac{\sum_m C_{mi}\eta'_m}{\sum_m C_{mi}\eta_m}\frac{\partial \eta_m}{\partial P}\right] dP \ . \tag{10.16}$$

This is a simple linear equation in the differentials dP, with the partial derivatives $\partial \mathcal{B}(B)/\partial P$ given implicitly as the coefficients of the dP. If I specialize to the case of finding $\mathcal{B}(D_s \to \mu\nu_\mu)$ assuming lepton universality, using the definitions of η_m

(equation 10.4),

$$\begin{aligned} \eta'_m &= \{\, \varepsilon_{\mu\nu},\ 9.75\varepsilon_{\tau e}B_{\tau e},\ 9.75\varepsilon_{\tau\mu}B_{\tau\mu},\ 9.75\varepsilon_{\tau\pi}B_{\tau\pi}, \\ &= \quad 0 \text{ (background)} \,\} \end{aligned} \tag{10.17}$$

$$\sum_P \frac{\partial \eta'_m}{\partial P} dP = \begin{cases} d\varepsilon_{\mu\nu} \\ 9.75(\varepsilon_{\tau e} dB_{\tau e} + B_{\tau e} d\varepsilon_{\tau e}) \\ 9.75(\varepsilon_{\tau\mu} dB_{\tau\mu} + B_{\tau\mu} d\varepsilon_{\tau\mu}) \\ 9.75(\varepsilon_{\tau\pi} dB_{\tau\pi} + B_{\tau\pi} d\varepsilon_{\tau\pi}) \\ 0 \qquad \text{(background)} \end{cases} \tag{10.18}$$

$$\begin{aligned} \sum_P \frac{\partial \eta'}{\partial P} dP &= \sum_{P,m} \frac{\partial \eta'_m}{\partial P} dP \\ &= d\varepsilon_{\mu\nu} + 9.75(\varepsilon_{\tau e} dB_{\tau e} + B_{\tau e} d\varepsilon_{\tau e} + \varepsilon_{\tau\mu} dB_{\tau\mu} \\ &\qquad + B_{\tau\mu} d\varepsilon_{\tau\mu} + \varepsilon_{\tau\pi} dB_{\tau\pi} + B_{\tau\pi} d\varepsilon_{\tau\pi}) \end{aligned} \tag{10.19}$$

$$\sum_P \frac{\partial \eta_m}{\partial P} dP = \begin{cases} \eta'_m d\mathcal{B}(B) + \mathcal{B}(B) \sum_P \partial_P \eta'_m dP & \text{(signal)} \\ d\eta_m & \text{(background)} \,. \end{cases} \tag{10.20}$$

Substituting these expressions (equations 10.17 through 10.20) into equation 10.16,

$$\begin{aligned} \eta' dN_{D_s} + N_{D_s} \sum_P \frac{\partial \eta'}{\partial P} dP &= \sum_i \Bigg[\frac{1}{\sum_m C_{mi}\eta_m} \sum_m C_{mi} \sum_P \frac{\partial \eta'_m}{\partial P} dP \\ &\quad - \frac{\sum_m C_{mi}\eta'_m}{(\sum_m C_{mi}\eta_m)^2} \sum_m^{\text{sig}} C_{mi} \left(\eta'_m d\mathcal{B}(B) + \mathcal{B}(B) \sum_P \frac{\partial \eta'_m}{\partial P} dP \right) \\ &\quad - \frac{\sum_m C_{mi}\eta'_m}{(\sum_m C_{mi}\eta_m)^2} C_{\text{bkg},i} d\eta_{\text{bkg}} \Bigg] \end{aligned}$$

which I can solve for $d\mathcal{B}(B)$,

$$\begin{aligned}
&\sum_i \left(\frac{\sum_m C_{mi}\eta'_m}{\sum_m C_{mi}\eta_m}\right)^2 d\mathcal{B}(B) \\
&\quad = \sum_i \Bigg[\frac{1}{\sum_m C_{mi}\eta_m}\sum_m C_{mi}\sum_P \frac{\partial\eta'_m}{\partial P}\,dP - \frac{\mathcal{B}(B)\sum_m C_{mi}\eta'_m}{(\sum_m C_{mi}\eta_m)^2}\sum_m^{\text{sig}} C_{mi}\sum_P \frac{\partial\eta'_m}{\partial P}\,dP \\
&\qquad - \frac{\sum_m C_{mi}\eta'_m}{(\sum_m C_{mi}\eta_m)^2} C_{\text{bkg},i}\,d\eta_{\text{bkg}}\Bigg] - N_{D_s}\sum_P \frac{\partial\eta'}{\partial P}\,dP - \eta' dN_{D_s} \\
&\quad = \sum_i \frac{\sum_m C_{mi}\eta_m - \mathcal{B}(B)\sum_m C_{mi}\eta'_m}{(\sum_m C_{mi}\eta_m)^2}\sum_m^{\text{sig}} C_{mi}\sum_P \frac{\partial\eta'_m}{\partial P}\,dP \\
&\qquad - \sum_i \frac{\sum_m C_{mi}\eta'_m}{(\sum_m C_{mi}\eta_m)^2} C_{\text{bkg},i}\,d\eta_{\text{bkg}} - N_{D_s}\sum_P \frac{\partial\eta'}{\partial P}\,dP - \eta' dN_{D_s}\ . \qquad (10.21)
\end{aligned}$$

Defining

$$A_1 = \left[\sum_i \left(\frac{\sum_m C_{mi}\eta'_m}{\sum_m C_{mi}\eta_m}\right)^2\right]^{-1} \qquad (10.22)$$

$$A_{2i} = \frac{\sum_m C_{mi}\eta'_m}{(\sum_m C_{mi}\eta_m)^2} \qquad (10.23)$$

$$A_{3i} = \frac{1}{\sum_m C_{mi}\eta_m} - \mathcal{B}(B)\,A_{2i}\ , \qquad (10.24)$$

equation 10.21 becomes

$$d\mathcal{B}(B) = A_1\left\{\sum_m^{\text{sig}}\left(\sum_i A_{3i}C_{mi} - N_{D_s}\right)\sum_P \frac{\partial\eta'_m}{\partial P}\,dP - \sum_i A_{2i}C_{\text{bkg},i}\,d\eta_{\text{bkg}} - \eta' dN_{D_s}\right\}$$

and the partial derivatives can be written explicitly as

$$\frac{\partial\mathcal{B}(B)}{\partial N_{D_s}} = -A_1\,\eta'\ , \qquad (10.25)$$

$$\frac{\partial\mathcal{B}(B)}{\partial B_{\tau e}} = 9.75\varepsilon_{\tau e}A_1\left[\sum_i A_{3i}C_{\tau(e),i} - N_{D_s}\right]\ , \qquad (10.26)$$

$$\frac{\partial\mathcal{B}(B)}{\partial B_{\tau\mu}} = 9.75\varepsilon_{\tau\mu}A_1\left[\sum_i A_{3i}C_{\tau(\mu),i} - N_{D_s}\right]\ , \qquad (10.27)$$

$$\frac{\partial \mathcal{B}(B)}{\partial B_{\tau\pi}} = 9.75\varepsilon_{\tau\pi}A_1\left[\sum_i A_{3i}C_{\tau(\pi),i} - N_{D_s}\right] , \tag{10.28}$$

$$\frac{\partial \mathcal{B}(B)}{\partial \varepsilon_{\mu\nu}} = A_1\left[\sum_i A_{3i}C_{\mu\nu,i} - N_{D_s}\right] , \tag{10.29}$$

$$\frac{\partial \mathcal{B}(B)}{\partial \varepsilon_{\tau e}} = 9.75B_{\tau e}A_1\left[\sum_i A_{3i}C_{\tau(e),i} - N_{D_s}\right] , \tag{10.30}$$

$$\frac{\partial \mathcal{B}(B)}{\partial \varepsilon_{\tau\mu}} = 9.75B_{\tau\mu}A_1\left[\sum_i A_{3i}C_{\tau(\mu),i} - N_{D_s}\right] , \tag{10.31}$$

$$\frac{\partial \mathcal{B}(B)}{\partial \varepsilon_{\tau\pi}} = 9.75B_{\tau\pi}A_1\left[\sum_i A_{3i}C_{\tau(\pi),i} - N_{D_s}\right] , \tag{10.32}$$

$$\frac{\partial \mathcal{B}(B)}{\partial \eta_{\text{bkg}}} = -A_1\sum_i A_{2i}C_{\text{bkg},i} . \tag{10.33}$$

The systematic uncertainty due to each parameter is given by

$$\delta_P\mathcal{B}(B) = \left|\frac{\partial \mathcal{B}(B)}{\partial P}\right|\delta P \tag{10.34}$$

and

$$\begin{aligned}\delta_P f_{D_s} &= \left|\frac{\partial f_{D_s}}{\partial \mathcal{B}(B)}\right|\delta_P\mathcal{B}(B) \\ &= \frac{1}{2\sqrt{A_B\mathcal{B}(B)}}\delta_P\mathcal{B}(B)\end{aligned} \tag{10.35}$$

where $A_B = 7.9694 \times 10^{-8}$ MeV^{-2} for $D_s \to \mu\nu_\mu$ (see equation 2.12).

Chapter 11 Results of Maximum Likelihood Analysis

In this chapter, I present the results of evaluating the leptonic decay likelihood function, equation 10.1, using the six candidate leptonic decay events from Table 9.2 and the other parameters derived throughout this thesis, summarized in Table 11.1. The three quantities which I derive are the branching fractions $\mathcal{B}(D_s \to \mu\nu_\mu)$ and $\mathcal{B}(D_s \to \tau\nu_\tau)$ and the D_s decay constant, f_{D_s}. I can rewrite equation 10.1 with some of the terms expanded to show the explicit dependence on the branching fractions:

$$\begin{aligned}\mathcal{L} &= P\left(N_{\text{events}}; N_{D_s}\cdot\left(\sum_m \varepsilon_m \boxed{\mathcal{B}(m)} + \eta_{\text{bkg}}\right)\right)\times \\ &\times \prod^{\text{events}} \frac{\sum_m^{\text{sig}} \mathcal{L}_m^{PID}\mathcal{L}_m(m^2_{\text{miss}})\varepsilon_m \boxed{\mathcal{B}(m)} + \mathcal{L}_{\text{bkg}}(m^2_{\text{miss}})\eta_{\text{bkg}}}{\sum_m \varepsilon_m \boxed{\mathcal{B}(m)} + \eta_{\text{bkg}}}\end{aligned} \tag{11.1}$$

where η_{bkg} is the inclusive background rate with no particle identification.. The D_s leptonic branching fractions, or the decay constant f_{D_s}, appear in the boxed terms as components of the weights for each possible final state, as well as in the expectation of the Poisson distribution. In this chapter I present my results maximizing the likelihood function [29] with respect to $\mathcal{B}(D_s \to \tau\nu_\tau)$, $\mathcal{B}(D_s \to \mu\nu_\mu)$, and f_{D_s}, along with my evaluation of the statistical (Section 10.2) and systematic (Section 10.3) uncertainties on those results. At the end, I will discuss additional sources of uncertainty which cannot be fully quantified, but which should be kept in mind when evaluating the significance of my results.

11.1 $\mathcal{B}(D_s \to \mu\nu_\mu)$ and $\mathcal{B}(D_s \to \tau\nu_\tau)$

The muonic and tauonic branching fractions, $\mathcal{B}(D_s \to \mu\nu_\mu)$ and $\mathcal{B}(D_s \to \tau\nu_\tau)$, may be treated as independent parameters of the likelihood function, equation 11.1. Figure 11.1 shows the resulting likelihood surface, normalized to the maximum value, which occurs at

$$\begin{aligned} \mathcal{B}(D_s \to \mu\nu_\mu) &= 3.20^{+6.39}_{-3.15}\ \% \\ \mathcal{B}(D_s \to \tau\nu_\tau) &= 1.63^{+15.88}_{-1.63}\ \% \end{aligned}$$

where the errors indicate points on the 68% confidence interval ellipse. The correlation coefficient r between the two parameters is 0.3.

11.2 $\mathcal{B}(D_s \to \ell\nu_\ell)$ with Lepton Universality

In equation 2.11, I showed the relationship between the various leptonic branching fractions, assuming that the $W-\ell-\nu$ coupling is independent of the lepton flavor ("lepton universality"). Using it, I can write the likelihood function in terms of $\mathcal{B}(D_s \to \mu\nu_\mu)$ alone, and maximize the resulting one-dimensional function. In Figure 11.2, I show this function, normalized to the maximum value, along with the 68% confidence interval around the maximum calculated according to equation 10.7. The result for the branching fraction is

$$\mathcal{B}(D_s \to \mu\nu_\mu) = 0.75^{+1.23}_{-0.47}\ \% \ .$$

Figure 11.2 also shows, for reference, the positions where $\log(\mathcal{L}/\mathcal{L}_{\max}) = -1/2$, which are often quoted as errors.

11.3 The Decay Constant f_{D_s}

Using the expression for the partial width of the D_s to decay leptonically, equation 2.9, and assuming lepton universality, I can express the likelihood function directly in terms

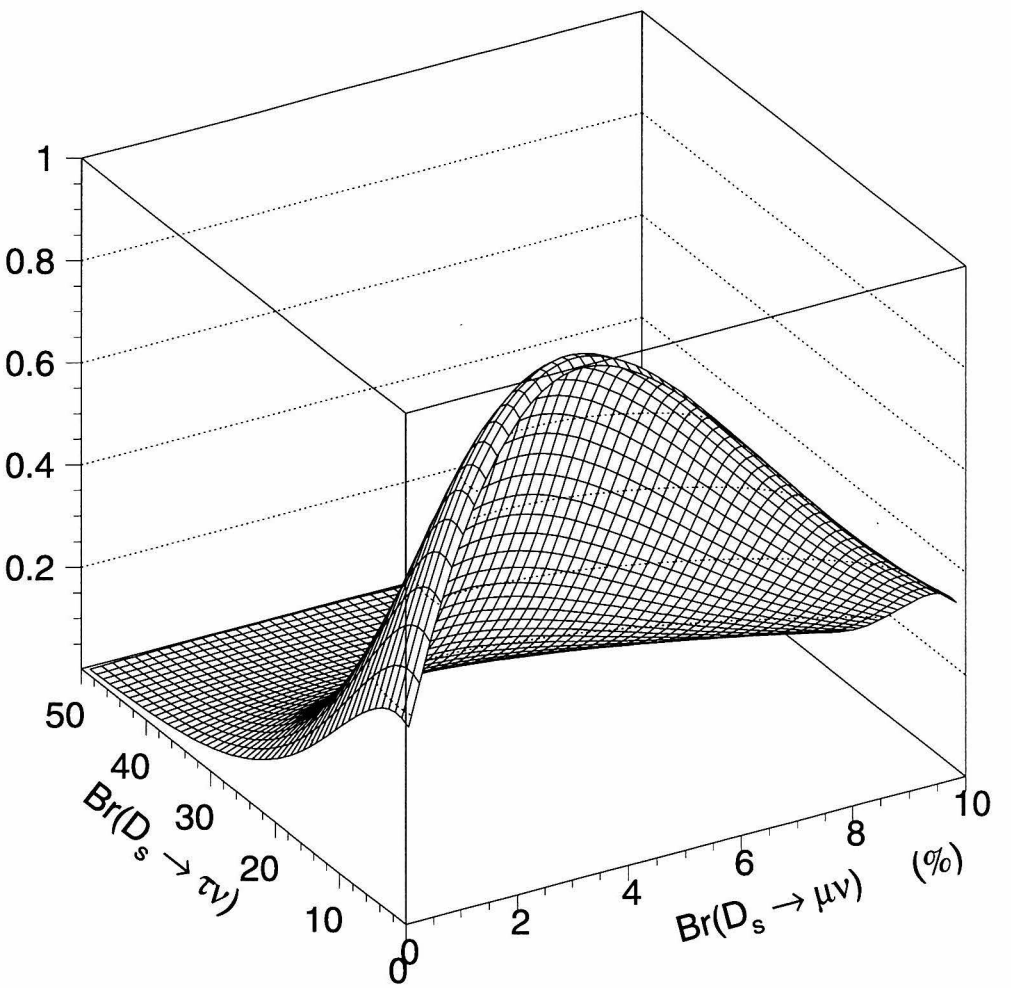

Figure 11.1: Likelihood function for D_s leptonic decay plotted with respect to $\mathcal{B}(D_s \to \mu\nu_\mu)$ and $\mathcal{B}(D_s \to \tau\nu_\tau)$ simultaneously. The likelihood function has been normalized to have a maximum value of 1.

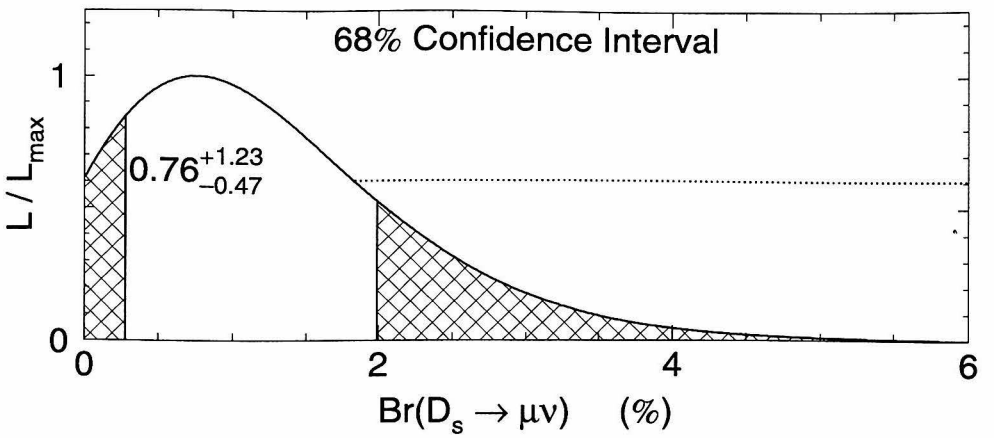

Figure 11.2: Likelihood function for D_s leptonic decay plotted with respect to $\mathcal{B}(D_s \to \mu\nu_\mu)$, assuming lepton universality: $\mathcal{B}(D_s \to \tau\nu_\tau) = 9.75\mathcal{B}(D_s \to \mu\nu_\mu)$, as discussed in Chapter 2. The hatched regions are excluded at 68% confidence level ("one sigma"). The value indicated is the maximum likelihood result and 68% CL errors. The dotted line indicates $\log(\mathcal{L}/\mathcal{L}_{\max}) = -1/2$, for reference.

of f_{D_s}, and maximize it to determine the best value for f_{D_s} based on my candidate event sample. I show the likelihood function in Figure 11.3, along with the 68% confidence interval statistical uncertainties (equation 10.7), with the result

$$f_{D_s} = 308^{+164}_{-189}\ \text{MeV} .$$

Figure 11.3 also shows, for reference, the positions where $\log(\mathcal{L}/\mathcal{L}_{\max}) = -1/2$.

11.4 Parameter Ranges and Systematic Errors

Although my results are clearly dominated by statistical uncertainties (since there are only six candidate leptonic decays in a sample of about 100 tagged $D_s^+ D_s^-$ events), an evaluation of the systematic errors is important. If nothing else, it provides insight into the validity of my technique for extracting $\mathcal{B}(D_s \to \ell\nu_\ell)$ and f_{D_s}, since the systematic errors measure the "intrinsic" uncertainty of my procedure, rather than the random fluctuations of signal and background statistics. In Chapter 10 (Section 10.3)

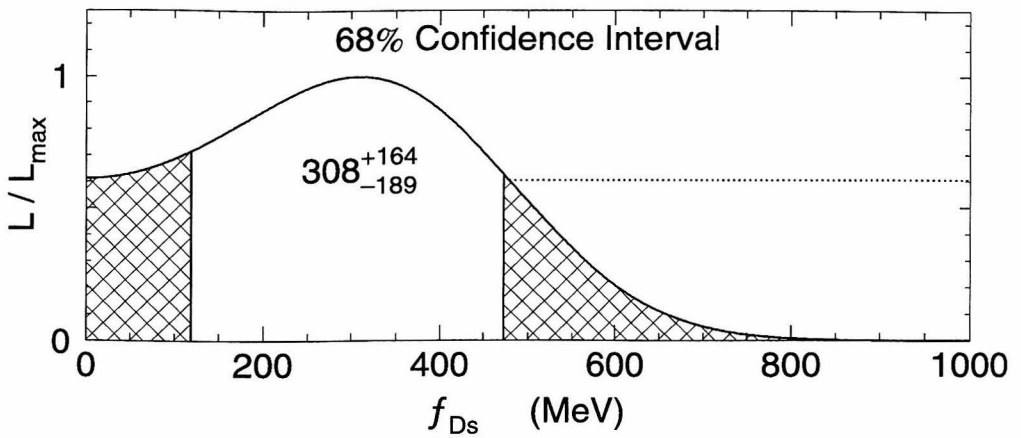

Figure 11.3: Likelihood function for D_s leptonic decay plotted with respect to f_{D_s}. This assumes the Standard Model process for $D_s \to \ell\nu_\ell$, and lepton universality, as discussed in Chapter 2. The hatched regions are excluded at 68% confidence level ("one sigma"). The value indicated is the maximum likelihood result and 68% CL errors. The dotted line indicates $\log(\mathcal{L}/\mathcal{L}_{\max}) = -1/2$, for reference.

I presented analytic expressions for the systematic uncertainty deriving from each of the "fixed" input parameters of my likelihood function, listed in Table 11.1.

When I evaluate equations 10.25 through 10.33 at the nominal values of these parameters and the kinematics of my candidate events, and substitute them into equations 10.34 and 10.35, I obtain the systematic uncertainties listed in Table 11.2. By construction, these systematic uncertainties are a first-order estimate of the 68% (one sigma) confidence interval about the measured results, and may be added in quadrature with the statistical uncertainties to obtain overall error bars for my results. The dominant systematic error arises from my the selection efficiency for $D_s \to \tau\nu_\tau, \tau \to e\overline{\nu}_e\nu_\tau$ events. This source is nearly three times as large as the next significant source, the fit uncertainty on the number of single-tag $D_s^+ D_s^-$ events.

Table 11.1: Input parameters of the leptonic decay likelihood function.

Input Parameter	Symbol	Range
Candidate D_s events	N_{D_s}	76.9 ± 13.8
$\mathcal{B}(\tau \to e\overline{\nu}_e\nu_\tau)$	$B_{\tau e}$	$.1788 \pm .0018$
$\mathcal{B}(\tau \to \mu\overline{\nu}_\mu\nu_\tau)$	$B_{\tau\mu}$	$.1746 \pm .0025$
$\mathcal{B}(\tau \to \pi\nu_\tau)$	$B_{\tau\pi}$	$.113 \pm .004$
Efficiencies (%)	$\varepsilon_{\mu\nu}$	58.2 ± 5.8
	$\varepsilon_{\tau e}$	58.6 ± 5.9
	$\varepsilon_{\tau\mu}$	54.1 ± 5.4
	$\varepsilon_{\tau\pi}$	46.8 ± 4.7
Background rate (%)	$\eta_{\rm bkg}$	$.79 \pm .08$

Table 11.2: Contributions to the systematic uncertainty in $\mathcal{B}(D_s \to \mu\nu_\mu)$ and f_{D_s} from each of the likelihood function input parameters. $\delta f_{D_s} = \delta\mathcal{B}(\mu\nu)\,/2\sqrt{A_B\mathcal{B}(\mu\nu)}$ as discussed in the text. The signs indicate the "direction" of systematic variation, *e.g.*, $\delta f_{D_s} < 0$ indicates that f_{D_s} would decrease if the indicated parameter increased. All uncertainties represent half of a 68% confidence interval about the nominal (maximum likelihood) value.

Input Parameter	δP	$\delta\mathcal{B}(\mu\nu)$ ($\mathcal{B}$ in %)	δf_{D_s} (MeV)
N_{D_s} from fit	13.8	-.016	-3.2
$\mathcal{B}(\tau \to e\overline{\nu}_e\nu_\tau)$	.0018	-.002	-.3
$\mathcal{B}(\tau \to \mu\overline{\nu}_\mu\nu_\tau)$	.0025	-.002	-.3
$\mathcal{B}(\tau \to \pi\nu_\tau)$	.0040	-.002	-.3
$\varepsilon_{\mu\nu}$	.0582	.004	.9
$\varepsilon_{\tau e}$	.0586	-.016	-3.2
$\varepsilon_{\tau\mu}$	.0541	-.011	-2.3
$\varepsilon_{\tau\pi}$	.0468	-.005	-1.0
$\eta_{\rm bkg}$	.0008	-.053	-10.9
Total Systematic Error		.059	12.1

11.5 Other Sources of Uncertainty

In addition to the variation of the likelihood parameters presented in Section 11.4, there are additional factors which must be considered in evaluating the uncertainties of my results for $\mathcal{B}(D_s \to \ell\nu_\ell)$ and f_{D_s}. Unfortunately, some of these factors are computationally difficult to quantify, while others have no quantitative character at all. In contrast to Section 11.4, my presentation here will be mainly qualitative, with perhaps some order-of-magnitude estimates of additional systematic uncertainties in my results.

Among the quantitative but computationally difficult sources of error are the parameterizations of the embedded likelihood functions $\mathcal{L}(m^2_{\rm miss})$ and $\vec{\mathcal{L}}^{PID}$. As I noted in equation 10.5, I approximate the missing mass distributions by overlapping Gaussians, which are fit to samples of Monte Carlo generated events. These fit parameters have approximately Gaussian uncertainties associated with them, arising from the limited statistics and the fitting process. In principle, I should follow the procedure described in Section 11.4, expanding the differentials of equation 10.16 to include $\partial C_{mi}/\partial H_{1m}$, $\partial C_{mi}/\partial \sigma_{2m}$ and so on. With six parameters each for five different functions, this would add 30 terms to the derivation, making an already difficult problem essentially intractable. Given that the uncertainties on these parameters are small (see Table 10.1), I expect that their contribution to the overall systematic error would also be small, of order $\delta f_{D_s} \sim 10$ MeV.

The particle identification likelihoods are similar. I described the functional form of these terms in Chapter 7. Here it is the detector response functions and resolutions which should be treated as parameters with uncertainties. Again, in principle each of the particle ID likelihood functions should be differentiated with respect to all of its parameters, and the resulting terms incorporated into the differential expression. Such an exercise would lead to terms which depend on the underlying detector data, rather than on kinematical quantities, and would also lead to an unacceptable and intractable multiplicity of terms. For purposes of this thesis, I believe that any systematic uncertainty resulting from the particle identification system is adequately accounted

for in the uncertainties on my identification efficiencies and backgrounds, since those values depend directly on my maximum-likelihood particle identification method.

Other sources of systematic uncertainty are not amenable to the quantitative, confidence-interval analysis I have used up to now. For example, my definition of an "isolated photon" includes fixed limits on energy deposition and proximity to charged tracks. If I vary these limits, it is possible that my sample of candidate leptonic decays would change, but I cannot sensibly assign a "confidence interval" to such variation. Without such an interval, I cannot derive systematic uncertainties with the same interpretation as those in Table 11.2. Similar arguments apply, for example, to the initial hadronic event selection applied to the BES data, and to my definition of the "central region" of the detector, where I restrict my event reconstruction. However, these various criteria ultimately affect my efficiencies and background levels, so I believe they are accounted for by my evaluations of $\delta\varepsilon$ and $\delta\eta$.

Chapter 12 Discussion

The pseudoscalar decay constants are fundamental parameters of meson structure. They appear in expressions for meson leptonic decays, for the mixing of neutral mesons (K^0–$\overline{K}^0$, B^0–$\overline{B}^0$), and in detailed balance expressions which relate hadronic or semi-leptonic branching fractions of different mesons. A precision measurement of f_{D_s}, therefore, could be used to constrain a variety of theoretical models and calculations, as well as to make accurate numerical predictions of physics in the B meson sector.

Unfortunately, the measurements I have discussed in this thesis are not sufficiently precise to serve these purposes. With only six observed candidate decays, out of a sample of less than 100 $e^+e^- \to D_s^+ D_s^-$ events, my statistical and systematic uncertainties are too large to do more than verify that my results,

$$\mathcal{B}(D_s \to \mu\nu_\mu) = .75^{+1.23}_{-.47}\,(stat) \pm .06\,(syst)\ \%$$

and

$$f_{D_s} = 308^{+164}_{-189}\,(stat) \pm 12\,(syst)\ \text{MeV}$$

are consistent with current theoretical expectations. On the other hand, my technique of constructing a global likelihood function provides the maximum possible sensitivity for the given data, and I believe it could be used to great advantage for similar measurements at future experiments, such as a B meson factory [46].

In this chapter I will compare my result with the few other experimental determinations of f_{D_s} [47,48], and describe some of the significant differences between those results and my analysis. I will also present a representative collection of theoretical calculations, and compare my result to them. Finally, I will discuss some of the implications of my results with respect to tests of the Standard Model in both the leptonic and quark sectors.

12.1 Other Experimental Determinations of f_{D_s}

In 1992, the WA75 collaboration at CERN presented the first observations of $D_s \to \mu\nu_\mu$ decays [47]. WA75 was a fixed target experiment, in which a π^- beam is incident on an emulsion target to produce D_s and other charmed hadrons by photoproduction,

$$\pi^- \to \pi^-\gamma^*;\ \gamma^*(Z,A) \to X_c + \text{hadrons}\ (X_c \sim D_s)\ ,$$

and a downstream hodoscope identifies the muon from $D_s \to \mu\nu_\mu$ decay. After selecting events consistent with a two-body charged decay to a single charged particle, or a two-body neutral decay to two charged particles, they measured the momentum of the muon transfers to the line of flight of the decaying charmed hadron ($p_\perp^\mu$). After fitting these distributions to semileptonic decays determined from Monte Carlo simulations, WA75 observed six charged events (out of 144) above the kinematic limit of 880 MeV/c for semileptonic decays. Normalizing to their observed $D^0 \to \mu^+\nu_\mu X$ signal, they extracted a branching fraction

$$\begin{aligned} \mathcal{B}(D_s \to \mu\nu_\mu) &= (3.9^{+1.8+0.8}_{-1.4-0.6} \pm 1.4) \times 10^{-3} \\ f_{D_s} &= 225 \pm 45 \pm 20 \pm 41\ \text{MeV}\ , \end{aligned}$$

where the last systematic error is due to their normalization of the D_s production rate, based on measurements of D^0 and D_s cross sections and branching fractions from several experiments.

In 1994, the CLEO experiment at Cornell used D_s^* production in charm jets to observe a few dozen candidate $D_s \to \mu\nu_\mu$ decays [48]. The CLEO detector, at the CESR e^+e^- storage ring at Cornell University, operates in the Υ resonance region ($E_{cm} \sim 10$ GeV) to produce $B\overline{B}$ meson pairs, as well as continuum quarks. For this analysis, they identified direct D_s^* production by tagging the photon in the decay $D_s^* \to D_s\gamma$, and rejecting events consistent with $B\overline{B}$ production. Selecting $D_s \to \mu\nu_\mu$ decays from their sample of D_s tags, and measuring the $\gamma-\mu$ correlations, they extract

the $D_s \to \mu\nu_\mu$ branching fraction relative to $\mathcal{B}(D_s \to \phi\pi)$. Their result,

$$\frac{\mathcal{B}(D_s \to \mu\nu_\mu)}{\mathcal{B}(D_s \to \phi\pi)} = 0.245 \pm 0.052 \pm 0.074 \ ,$$

must be normalized to the $\phi\pi$ branching fraction ($3.5 \pm 0.4\,\%$) in order to extract f_{D_s}; this introduces an additional systematic error in their published result, $f_{D_s} = 344 \pm 37 \pm 52 \pm 23$ MeV.[1]

Both of these experimental results, the only published values for f_{D_s} before BES, suffer from the problem of normalization. In order to convert a count of observed events to a branching fraction (and thence to f_{D_s}), WA75 must make assumptions about the rate of D_s production in πN interactions; CLEO must either know the absolute D_s^* production rate at the $\Upsilon(4S)$, or they must normalize their $\mathcal{B}(D_s \to \mu\nu_\mu)$ to some other D_s decay channel, such as $D_s \to \phi\pi$. In either case, the normalizations are determined from other experiments, so the analysis is ultimately not "self-contained."

In contrast, the BES data described in this thesis were obtained at $E_{cm} = 4.03$ GeV, which is just above threshold for $D_s^+ D_s^-$ production. As I have noted in Part II of this thesis, we can identify an entire event as $e^+e^- \to D_s^+ D_s^-$ by reconstructing only one of the D_s mesons. We do not have to apply any selection criteria to the other (recoil) half of the event. A sample of such tagged events, used as input to various recoil analyses, allows us to make absolute determinations of branching fractions, independent of $D_s^+ D_s^-$ cross sections, our tagging efficiency, or other D_s branching fractions. BES published a search for $D_s \to \ell\nu_\ell$ decays in such a "double-tag" analysis in 1995 [50]; that analysis was independent of mine, using a different tagging method, different criteria for recoil event selection, and a different likelihood function with which to extract $\mathcal{B}(D_s \to \mu\nu_\mu)$ and f_{D_s}. The published result, based on 94.3 ± 13.2 tagged $D_s^+ D_s^-$ events and three candidate leptonic decays (with electron or muon recoil tracks) was $f_{D_s} = 430\,^{+150}_{-130} \pm 40$ MeV.

[1] In 1995, CLEO reported [49] an updated calculation, based on 50% additional data (3.06 fb^{-1} *vs.* 2.13 fb^{-1}). Their new result, $\mathcal{B}(\mu\nu)/\mathcal{B}(\phi\pi) = 0.184 \pm 0.03 \pm 0.038$, corresponds to $f_{D_s} = 284 \pm 30 \pm 30 \pm 16$ MeV.

12.2 Theoretical Calculations of f_{D_s}

Because of their relationship to meson wavefunctions, decay constants are often calculated as a test of theoretical models of different types. Conversely, experimental determination of decay constants can directly validate or exclude those same theoretical models. In this section, I will briefly review several theoretical models which can make numerical predictions of f_{D_s}. Many of these models also make predictions, sometimes with smaller uncertainties, for such ratios as f_D/f_{D_s} or f_{D_s}/f_{B_s} [51,52].

12.2.1 Lattice QCD

The most popular calculational method is lattice QCD, where spacetime is modeled as a discrete lattice, with typical spacing $\lesssim \mathcal{O}(1\text{ fm})$. This method is not a model, but is rather an implementation of the exact theory of QCD in a form suitable for computation. The quarks occupy the coordinates of the four-dimensional lattice, and the gluons occupy the "edges" connecting pairs of coordinates, exchanging four-momentum between the quarks at each end. Observables are measured by sampling the configuration space of possible quark-gluon arrangements, and taking statistical averages. The calculations of lattice QCD are comparable to experimental results in the limit where the lattice spacing becomes zero, and where the number of lattice points (typically $\mathcal{O}(16^4)$) becomes infinite. This requires extreme computing resources, and so such calculations have historically been of limited value. Modern massively-parallel processors are making the calculation of lattice QCD results much more reliable and mutually consistent. Bernard and Soni have done several such calculations in the last decade, reporting $f_{D_s} = 234 \pm 64.35$ MeV [53] and 230 ± 11 MeV [54]. In 1988, DeGrand, *et al.* reported $f_{D_s} = 157 \pm 17$ MeV [55].

12.2.2 QCD Sum Rules

Derivation of QCD sum rules, like lattice QCD, should provide model-independent predictions for f_{D_s}, based on exact, non-perturbative QCD. Sum rules are derived

by constructing dispersion relations (integral equations over the energy transfer Q) to relate low energy parameters such as hadron masses and structure functions to current correlators [56]. These correlators (two-point integrals over spacetime) are expanded in short-distance operator products, representing perturbative and nonperturbative effects at a particular order in the derivatives. The operator products themselves are expressed in terms of vacuum expectation values of the quark and gluon fields. If sufficient sum rules can be determined, then the self-consistency of QCD implies that experimental input from one sector (such as the charmonium mass spectrum or inclusive e^+e^- cross sections) can be used to compute, or at least to constrain, unmeasured parameters in other sectors such as f_{D_s}. In 1987, Dominguez and Paver used Hilbert transform power-moment sum rules to compute pseudoscalar decay constants at $Q^2 = 0$, obtaining $f_{D_s} = 194 \pm 12$ MeV, while Narison used QCD duality to obtain $f_{D_s} = 154 \pm 14$ MeV (from his results $f_D = 122 \pm 11$ MeV and $f_{D_s} \simeq 1.26\, f_D$). Bian and Juang in 1993 reported their calculation of $f_{D_s} = 332$ MeV. While some of the experimental data used to convert sum rules to numerical predictions changed between the 1980's and 1990's, the broad range of results is characteristic, and far larger than the reported uncertainties of any individual calculation.

12.2.3 Non-relativistic Potentials

The non-relativistic quark model was discussed in Chapter 2, where I showed equation 2.13, relating a meson's decay constant to its mass and wavefunction,

$$f_P^2 = \frac{12}{M_P} |\psi_{Qq}(0)|^2 \ .$$

The calculation of $\psi(0)$ depends on the precise form of the binding potential, the value of α_s used, and on the constituent masses of the two quarks. Different authors make different choices for all of these parameters, which lead to the variety of predictions for f_{D_s} in the literature, an which are reflected in Figure 12.1.

12.2.4 Heavy Quark Effective Theory

The heavy quark effective theory (HQET), originated by Isgur, Wise, and others [57], is an approximation of QCD in which the mass of the heavy quarks (bottom, top, and sometimes charm) are taken to be infinite. It is similar to the non-relativistic potential models, but is an exact field theory, rather than a parameterized model. In this limit, a number of useful symmetries appear in the QCD Lagrangian, including an $SU(2N)$ symmetry among the heavy quark flavors and spins, which can be exploited to relate experimental measurements in different sectors. Because HQET is a field theory of QCD, rather than a phenomenological model, the relations it predicts can be evaluated as operator product expansions similar to QCD sum rules, as Hamiltonians for potential models, or using any of the other quark models taken in the appropriate ($m_Q \to \infty$) limit. Generally, HQET predictions of decay constants are expressed as ratios, f_D/f_B or f_{D_s}/f_D, rather than absolute values. One of the problems with using HQET to extract charm parameters, such as f_{D_s}, is that the charm quark mass of $\lesssim 2$ GeV is only slightly larger than the QCD scale $\Lambda \sim 1$ GeV; it is not clear that the infinite mass limit assumed by HQET is correct in this case, although some of the results are consistent with experimental data.

Recently [51], Amundson has used HQET to compare various quark potential models, and to extract a model-independent expression for the separate pseudoscalar decay constants to first order in $1/m_Q$. His result for $f_{D_s}$is 376 ± 87.3 MeV, where I quote the uncertainty as the difference between his calculations for non-relativistic and semirelativistic models.

12.2.5 Relativistic Quark Model

In the literature, the term "relativistic quark model" covers a fairly wide variety of models and calculational methods. Some authors have used a "relativized Schrödinger equation," where the $p^2/2m$ operator is replaced by $\sqrt{p^2+m^2}$ [58]; others construct a Dirac wavefunction for the bound quarks in a non-relativistic confining potential [59]; others use "trial" wavefunctions of simple form (Gaussian or hydrogenic) to construct

Lorentz-covariant currents [60]. These different approaches, and the latitude of input parameters available, lead to a wide range of calculations of decay constants, varying over a factor of two, as Figure 12.1 indicates.

12.2.6 The Factorization Hypothesis

The factorization hypothesis is a simplification of the problem of computing matrix elements for heavy hadron weak decays. It presumes that the matrix element for the decay $M \to XY$,

$$\mathcal{M} = \langle XY|H|M\rangle \ ,$$

can be factored into separate transitions, for example the initial heavy hadron to a heavy final-state hadron X, and production of light hadrons Y from vacuum,

$$\mathcal{M} = \langle Y|A^{\mu}|0\rangle\langle X|H_{\mu}|M\rangle \ .$$

In terms of tree-level Feynman diagrams, this factorization assumes that the creation and decay of the W boson propagator are independent, and that there are no interference or final-state interaction effects contributing to the XY final state.

With respect to meson decay constants, the factorization hypothesis allows, for example, f_{D_s} to be related to charmed semileptonic decays (for which factorization is necessarily valid) and nonleptonic decays to strange (quark) final states, since f_{D_s} parameterized all of the QCD interactions at the $c{-}s{-}W$ vertex. Calculations by Bortoletto *et al.* [61], and by Rosner [19] give values for f_{D_s} of 276 ± 62.94 and 259 ± 74, respectively.

12.3 Summary of Comparisons

In Figure 12.1, I summarize the current status of experimental and theoretical measurements of f_{D_s}, in comparison with both the recently published BES result [50] and my result from this thesis. Clearly, the statistical uncertainties on my result are

consistent with everything in the current literature. Nevertheless, it is interesting that most of the theoretical calculations (both with and without uncertainties) are systematically lower than my value for f_{D_s}, and lower than the other experimental determinations.

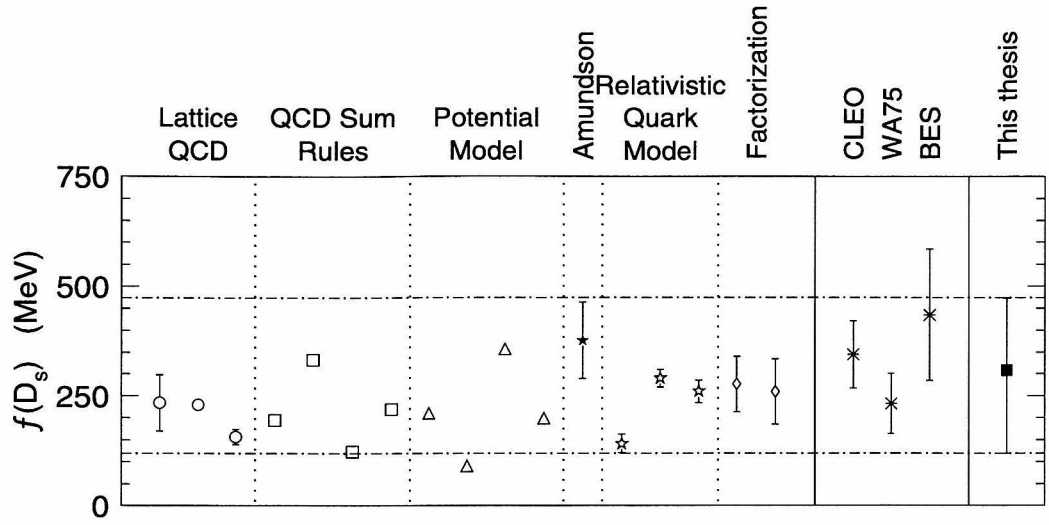

Figure 12.1: Comparison of this measurement of f_{D_s} with other experimental and theoretical calculations. The theoretical calculations shown cover lattice QCD [53–55], QCD sum rules [56,62–64], quark-quark potential models [65–68] including Amundson's application of heavy quark effective theory [51], relativistic wavefunction calculations [58,60,69], and factorization models [19, 61]. The details of the experimental results [47,48] are discussed in the text. The BES analysis shown [50,70] was independent of this thesis. The figure is adapted from Ref. [70].

12.4 Unitarity of D_s Decays

My determination of $\mathcal{B}(D_s \to \mu\nu_\mu)$ and $\mathcal{B}(D_s \to \tau\nu_\tau)$ as independent results (Section 11.1) could be used as a test of lepton universality. Equation 2.11 makes a clear prediction for the ratio of the branching fractions, assuming that both muons and taus couple identically to the W boson. Any non-standard coupling, or non-Standard Model processes which contribute to the final states I have studied, would lead to a discrepancy with this prediction. In fact, the statistical uncertainty on my result

make this a somewhat futile exercise. Using my results,

$$\begin{aligned} \mathcal{B}(D_s \to \mu\nu_\mu) &= 3.20^{+6.39}_{-3.15}\ \% \\ \mathcal{B}(D_s \to \tau\nu_\tau) &= 1.63^{+15.88}_{-1.63}\ \% \end{aligned}$$

I obtain the ratio

$$\frac{\mathcal{B}(D_s \to \tau\nu_\tau)}{\mathcal{B}(D_s \to \mu\nu_\mu)} = .51^{+17.11}_{-3.55}\ , \tag{12.1}$$

which, while not close to the expected ratio of 9.75, is consistent with lepton universality within the quoted errors, as well as being consistent with zero.

Chapter 13 Summary of Analysis

In this thesis I have presented my determination of the branching fraction of the D_s meson to leptons,

$$\mathcal{B}(D_s \to \mu\nu_\mu) = .75^{+1.23}_{-.47}\,(stat) \pm .06\,(syst)\ \%$$

and the pseudoscalar decay constant,

$$f_{D_s} = 308^{+164}_{-189}\,(stat) \pm 12\,(syst)\ \text{MeV}\ .$$

My analysis is based on 21.9 pb^{-1} of e^+e^- collision data taken with the BES detector at a center of mass energy of 4.03 GeV. From an initial sample of 1.6 million hadronic ($\geq$ 3 charged track) events, I extracted 133,000 possible $e^+e^- \to D_s^+ D_s^-$ production events, which I then analyzed in detail using a maximum likelihood technique. The BES data sample, at $E_{cm} = 4.03$ GeV, is below threshold for production of D_s^* mesons; thus, D_s decays may be studied using a "double-tag" technique, where identification of one D_s decay in the event constrains the other, recoil, side to also be a D_s, in an unbiased way. Information about, or uncertainties arising from, the absolute size of the data sample or D_s production cross sections are eliminated naturally, leading to absolute branching fraction determinations.

I presented a maximum-likelihood based method of single-track particle identification, which combines information from all of the BES detector systems. This method includes response functions for each charged particle type ($e, \mu, \pi^\pm, K^\pm, p$) determined from kinematically selected (non-particle ID based) samples of each type, and correctly integrates both Gaussian and non-Gaussian probability distributions. The software package with which I implemented this method is documented in this thesis (see Appendix B), and has been used for other published BES analyses [44, 50].

To obtain a sample of $D_s^+ D_s^-$ events, I constructed a likelihood function (equation 8.4) to describe the decay of a D_s meson into a fully-reconstructed final state, $\phi(K^+K^-)\,\pi^+$, $\overline{K}^{*0}(K^+\pi^-)\,K^+$, or $K_S^0(\pi^+\pi^-)\,K^+$. This function,

$$\begin{aligned}\mathcal{L}(D_s \to A(ab)c) \;&=\; \mathcal{L}_{PID}(i=a)\cdot\mathcal{L}_{PID}(j=b)\cdot\mathcal{L}_{PID}(k=c) \\ &\quad\times \mathcal{L}_{\rm res}(m_{ij};M_A,\Gamma_A)\cdot\mathcal{L}_{\rm fit}(\chi^2,ndf)\ ,\end{aligned}$$

includes my particle identification function for each track ($\mathcal{L}_{PID}$), correspondence of the two-body mass with the expected resonance ($\phi(1020)$, K^{*0}, or K_S^0) in the D_s hadronic decay, and a kinematic fit of the three-body final state to one side of an $e^+e^- \to X\overline{X}$ event, where the mass of X is left as a free parameter of the fit. By selecting the maximum likelihood permutation of each set of three tracks, and requiring that that permutation have a probability (cumulative likelihood) $P > 0.05$, I identified a total of 190 events with fitted masses between 1958.5 and 1978.5 MeV/c^2 (*i.e.*, within 10 MeV/c^2 of the world-average mass of the D_s meson, 1968.5 MeV [1]). Of these events, I determined that 76.9 ± 13.8 were true $D_s^+ D_s^-$ events, based on an unbinned maximum-likelihood fit of the masses of 1097 candidates to a D_s signal peak plus background (see Figure 8.15).

Finally, I searched my sample of 190 candidate $D_s^+ D_s^-$ events for recoils containing a single charged track in the central region of the detector ($|\cos\theta| < 0.75$), no isolated photons (neutral energy in the barrel calorimeter more than 18 degrees from the nearest charged track), and missing momentum pointing into the central detector. This latter condition is characteristic of events containing neutrinos, such as leptonic and semi-leptonic D_s decays. My criteria were designed to select D_s leptonic decays in one of four final states, $D_s \to \mu\nu_\mu$ or $D_s \to \tau\nu_\tau$ with $\tau \to e\overline{\nu}_e\nu_\tau$, $\tau \to \mu\overline{\nu}_\mu\nu_\tau$, or $\tau \to \pi\nu_\tau$ (see Figure 9.1). I identified six candidates, shown graphically in Figures 9.5 through 9.10, of which 1.51 are estimated from Monte Carlo studies to be background.

Based on these six events, I have extracted measurements of the branching fraction $\mathcal{B}(D_s \to \mu\nu_\mu)$ and the D_s decay constant f_{D_s}, using a maximum likelihood technique. My likelihood function (equation 10.2) is a product of likelihoods for each event,

constructed as a sum of likelihoods for the missing mass in the event compared to that expected in each of the four leptonic-decay final states, or background,

$$\begin{aligned}\mathcal{L} &= P(N_{\rm events}; N_{\rm exp}) \prod_i^{\rm events} \left\{ \sum_m^{\rm channels} W_m \mathcal{L}_m^{PID} \mathcal{L}_m(m^2_{\rm miss}) \right. \\ &\quad \left. + W_{\rm bkg}\mathcal{L}_{\rm bkg}(m^2_{\rm miss}) \right\} .\end{aligned}$$

Each of the likelihood factors is weighted by the likelihood that the recoil track is the particle expected in the final state, and by the fraction of events expected in the final state.

I can maximize this likelihood function with respect to the leptonic branching fractions $\mathcal{B}(D_s \to \mu\nu_\mu)$ and $\mathcal{B}(D_s \to \tau\nu_\tau)$, which appear in the weights W_m, or with respect to the decay constant f_{D_s} ($\mathcal{B}(D_s \to \ell\nu_\ell) \propto f^2_{D_s}$) assuming universal couplings of the leptons. My results for all of these cases were shown graphically in Chapter 11. For the two branching fractions, my likelihood function is maximized at

$$\begin{aligned}\mathcal{B}(D_s \to \mu\nu_\mu) &= 3.20^{+6.39}_{-3.15}\ \% \\ \mathcal{B}(D_s \to \tau\nu_\tau) &= 1.63^{+15.88}_{-1.63}\ \%\end{aligned}$$

when the two are allowed to vary independently. Assuming lepton universality, $\mathcal{B}(D_s \to \tau\nu_\tau) = 9.75\mathcal{B}(D_s \to \mu\nu_\mu)$, and my likelihood function is maximized at

$$\mathcal{B}(D_s \to \mu\nu_\mu) = 0.75^{+1.23}_{-0.47}\ \%$$

corresponding to a decay constant

$$f_{D_s} = 308^{+164}_{-189}\ \text{MeV}$$

where the quoted uncertainties are statistical, and correspond to a 68% confidence interval about the central (maximum likelihood) value.

Since my likelihood function is fully analytical (the missing mass distributions are parameterized as overlapping Gaussians, fit to Monte Carlo data), I have determined

the systematic uncertainties on my analysis by formally differentiating the maximum likelihood condition

$$\frac{\partial \mathcal{L}}{\partial \mathcal{B}(D_s \to \mu\nu_\mu)} = 0 \ ,$$

to obtain the total differential $d\mathcal{B}(\mu\nu)$ in terms of the various parameters of the likelihood function, and the differentials of those parameters. I described this procedure in detail, and presented explicit expressions for all of the derivatives $\partial\mathcal{B}(\mu\nu)/\partial P$, in Section 11.4. Adding the individual contributions in quadrature, I found that my total systematic uncertainty is $\delta\mathcal{B}(D_s \to \mu\nu_\mu) = 1.5 \times 10^{-3}$ (*i.e.*, 0.15%), or $\delta f_{D_s} = 27$ MeV. In the following section (11.5) I discussed qualitatively sources of systematic error which arise outside my likelihood function, and hence cannot be defined rigorously in terms of a confidence interval.

The decay constant f_{D_s} is a fairly fundamental parameter of the Standard Model. Its measurement probes several aspects of QCD, including the effective quark-quark potential in mesons, the accuracy of lattice calculations, and symmetry relationships among the hadrons. Although my measurement of f_{D_s} is unique in being independent of production cross sections or hadron decay fractions (such as $\mathcal{B}(B \to D_s X)$), the limited statistics of the BES data set make it less useful for validating or excluding the many theoretical calculations in the literature. In Chapter 12 I compared my result with existing theoretical calculations, and with experimental results from fixed target [47] and high-energy charm production [48]; within my statistical and systematic uncertainties, my result is consistent with all of them. In addition, my independent determination of the muonic and tauonic branching fractions can be used to test leptonic universality: the ratio of branching fractions,

$$\frac{\mathcal{B}(D_s \to \tau\nu_\tau)}{\mathcal{B}(D_s \to \mu\nu_\mu)} = .51^{+17.11}_{-3.55}$$

is consistent, within large errors, to the expected value of 9.75.

Part V

Appendices

Appendix A The Coupled-Channel Model of Charm Production

In 1980, Eichten *et al.* presented [28] calculations, based on a non-relativistic quark-model potential, for the production of charmed mesons in e^+e^- collisions in the 3–5 GeV center-of-mass energy range. Specifically, the model predicts values for R_{charm} as a function of E_{cm}, for the reactions

$$\begin{array}{lll} e^+e^- \to D^0\overline{D}^0 & e^+e^- \to D^+D^- & e^+e^- \to D_s^+D_s^- \\ e^+e^- \to D^0\overline{D}^{*0} & e^+e^- \to D^\pm D^{*\mp} & e^+e^- \to D_s^\pm D_s^{*\mp} \\ e^+e^- \to D^{*0}\overline{D}^{*0} & e^+e^- \to D^{*+}D^{*-} & e^+e^- \to D_s^{*+}D_s^{*-} \end{array}$$

as well as to higher charmed resonance states. The Mark III experiment used this model [71] to investigate possible running energies for charmed meson studies, including D_s production.

I have adapted the Mark III software [72] for use by the BES experiment, and have used it to generate updated predictions, based on the latest [1] charmed meson masses and coupling constants (α_s). Figures A.1 and A.2 show these predictions for the energy range 3.7–4.5 GeV. Of particular interest are the production cross sections predicted at $E_{cm} = 4.03$ GeV, where there is expected to be an enhancement in the $e^+e^- \to D_s^+D_s^-$ rate, listed in Table A.1.

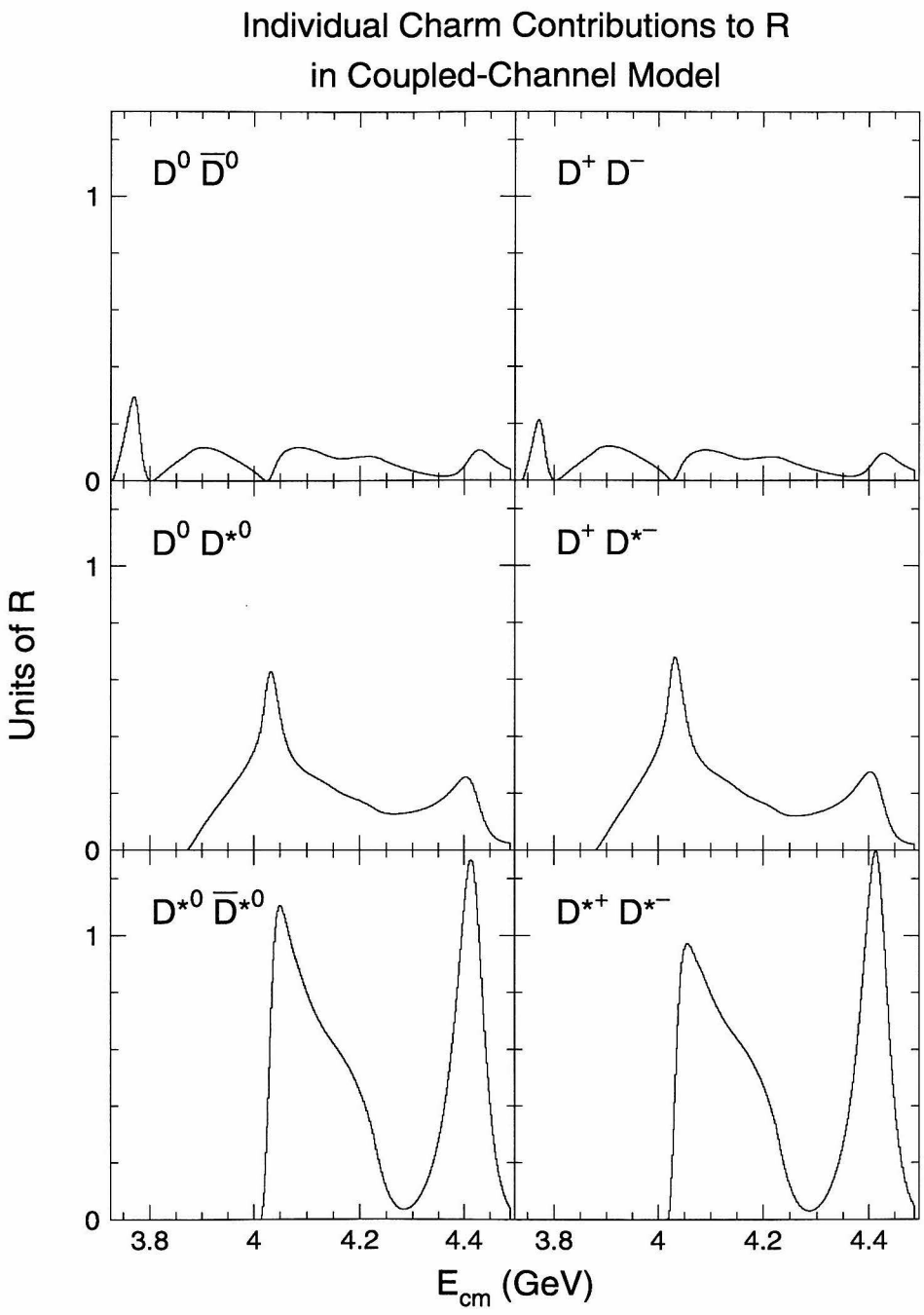

Figure A.1: Predictions for production of charmed mesons in e^+e^-collisions, using the coupled-channel model. Predictions are expressed in units of $R = \sigma(e^+e^- \to X)/\sigma(e^+e^- \to \mu^+\mu^-)$.

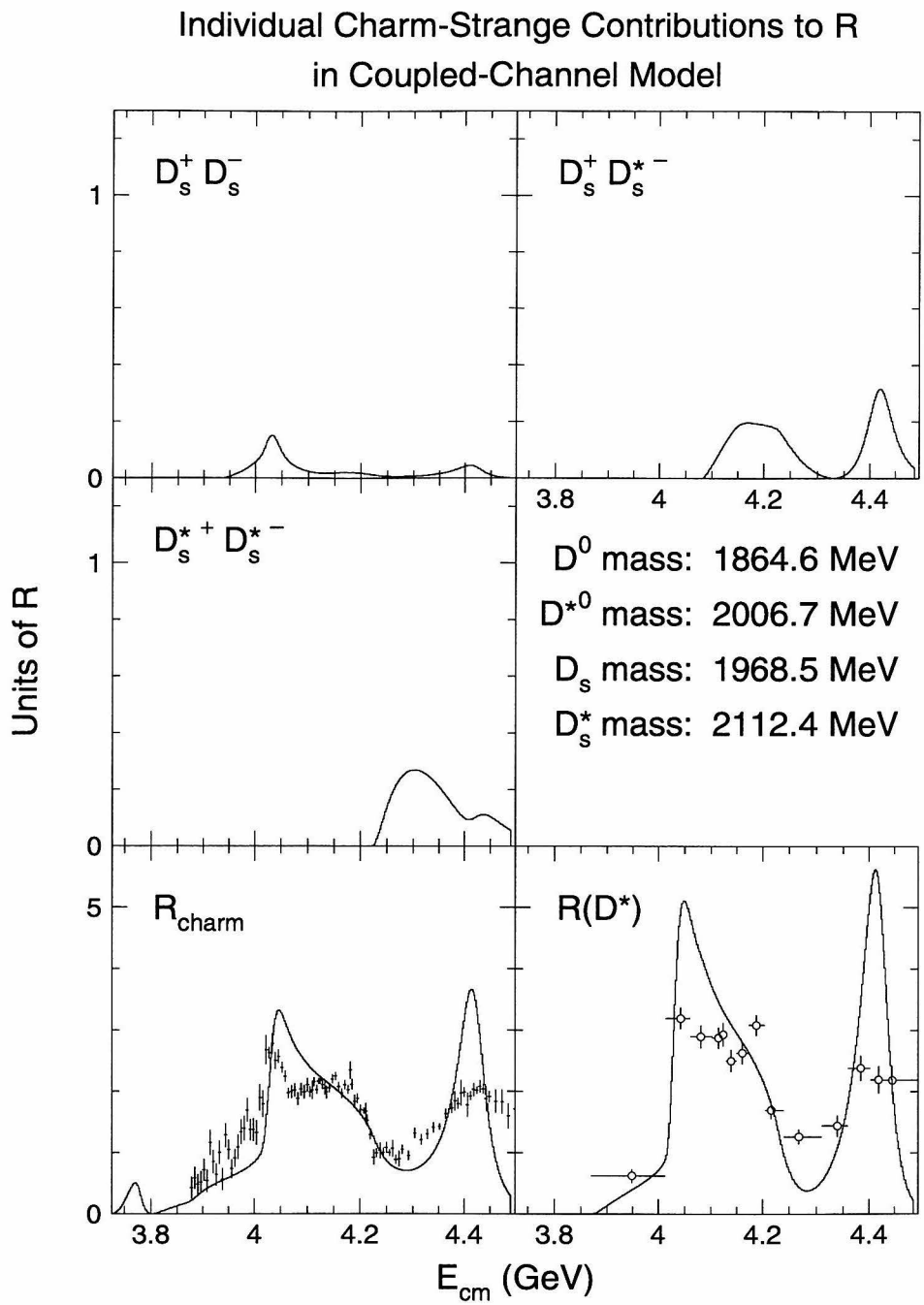

Figure A.2: Predictions for production of D_s mesons in e^+e^-collisions, using the coupled-channel model. Predictions are expressed in units of $R = \sigma(e^+e^- \to X)/\sigma(e^+e^- \to \mu^+\mu^-)$.

Table A.1: Predictions for production of charmed meson states at center-of-mass energy $E_{cm} = 4.03$ GeV, using the coupled-channel model.

	R_{mode}			σ(Total)
$e^+e^- \rightarrow$	Charged	Neutral	Total	(nb)
$D\overline{D}$	3.9×10^{-3}	0.01	0.01	0.07
D^*D	0.67	0.62	1.30	6.95
$D^*\overline{D}^*$	0.36	0.58	0.93	4.99
$D_s^+D_s^-$	0.15		0.15	0.81
Total			2.40	12.8

Appendix B Particle Identification with TRACKID

I have implemented the BES particle identification algorithm described in Chapter 7 as a package of routines called `TRACKID`. This package calculates the χ^2 of a reconstructed track for any charged particle hypothesis $\{e, \mu, \pi, K, p\}$, or for the neutral hypothesis of an isolated photon. Additionally, routines are provided for computing a normalized weight between any two identification hypotheses, and for filling a vector (N-tuple) with the reconstructed track parameters and identification confidence levels.

`TRACKID` is available as part of the BES software library. No special compilation or linking is required. All of the routines are written in portable Fortran 77, and may be modified or improved by users as necessary.

The basic structure of `TRACKID` is shown in Figure B.1. there are three "top-level" routines, intended to be called by the user, and a modular hierarchy of routines which first fill a vector with data from each detector system, then use that data to compute particle identification likelihoods and χ^2's. At each stage, the number of degrees of freedom contributed by the detector system is accumulated, so that a final confidence level for each hypothesis may be calculated.

The following sections discuss the user interface in detail, and also describe how the program flow in Figure B.1, and the calculations, may be controlled by the user.

B.1 User Interface

The user interface to `TRACKID` consists of four program routines, and one include file, which is only necessary if the user wants access to the N-tuple version of the track data. All of the routines take a track index into `TRKLST` as the initial argument. This argument is an `INTEGER` and must be in the valid range for tracks in the event,

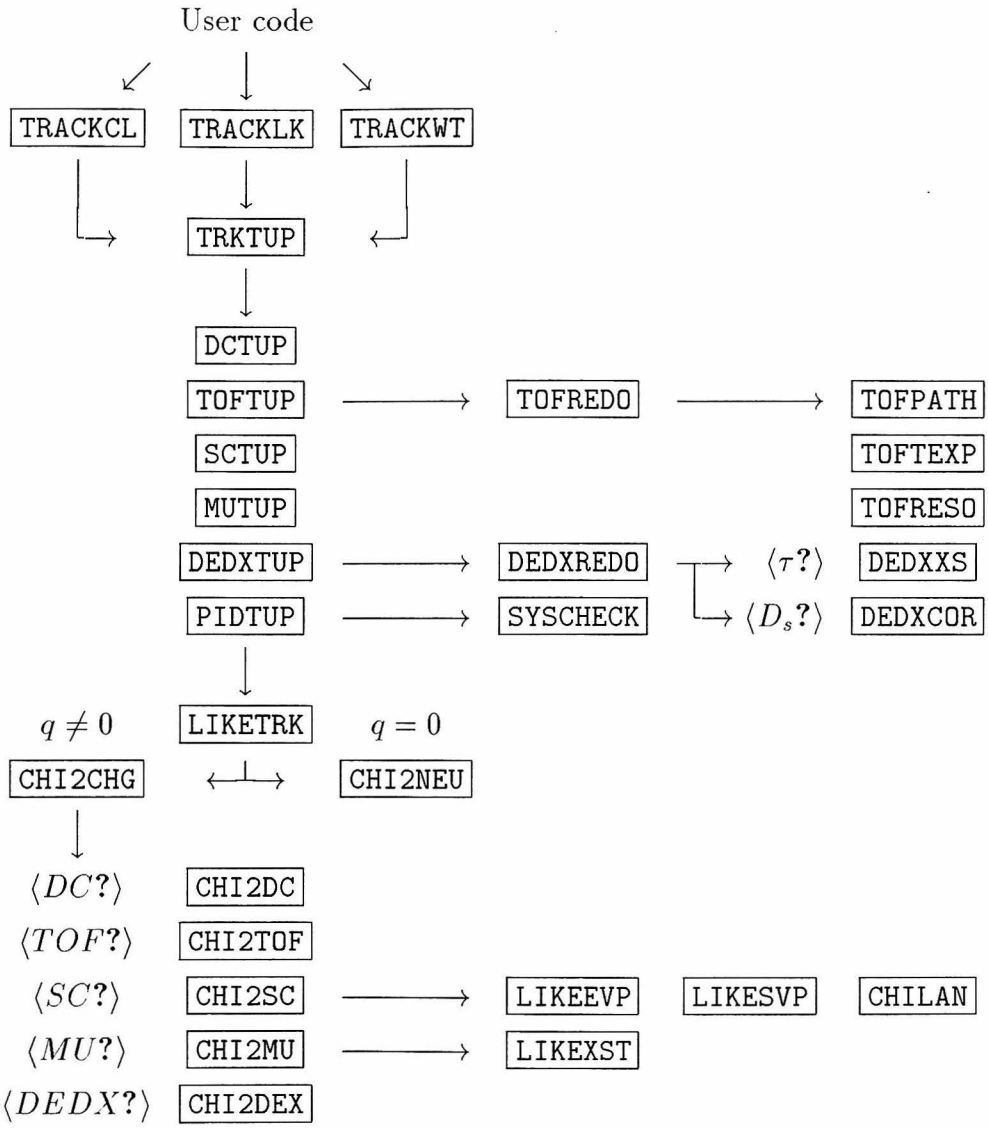

Figure B.1: Structure of TRACKID software package.

$1 \leq$ i $\leq$ TOTTKS.

The particle identification hypothesis h is coded using an INTEGER value, taken from the BES Monte Carlo particle numbering scheme (see XMCHTY.INC):

$$\gamma = 1,\ e = 2,\ \mu = 3,\ \pi = 4,\ K = 5,\ p = 6\ . \tag{B.1}$$

B.2 Universal Particle Identification

```
EXTERNAL TRACKID
INTEGER  TRACKID, IDBEST
INTEGER  i, hypgood, hypbad
REAL clmin
IDBEST = TRACKID(i,hypgood,hypbad,clmin)
```

The function TRACKID provides a universal interface to the TRACKID package. Given appropriate values of hypgood, hypbad, and clmin, it will return the best available identification (using the BES Monte Carlo particle numbering scheme (equation B.1). If no identification can be made, IDBEST = 0.

Normalized Weight. With hypgood $\neq 0$ and hypbad $\neq 0$, a normalized weight is computed to compare the two hypotheses, as is done with TRACKWT (Section B.5). The additional argument clmin in this case sets the threshold for identification with the normalized weight. If clmin = 0, the threshold is set to 0.5.

Maximum Likelihood. With hypgood = hypbad = 0, a maximum likelihood selection is done, among all six identification hypotheses (equation B.1). If clmin $\neq 0$, it is used as a threshold for a global consistency check, to determine if the maximum likelihood identity is reasonable.

Consistency Check. With hypgood $\neq 0$ but hypbad = 0, a simple consistency check for the specified hypothesis is done, using clmin as the minimum acceptable confidence level. If clmin = 0, then TRKCUT (Section B.8) is used as the minimum confidence level.

B.3 Filling a Vector of Likelihoods

```
INTEGER i
REAL    PIDLIKE(6)
CALL TRACKLK(i,PIDLIKE)
```

The subroutine TRACKLK returns an array of normalized likelihoods for all possible particle identification hypotheses for the selected track. Note that the internal N-tuple vector TKVAR is *not* available to the calling user code.

The array PIDLIKE will be filled on return with the likelihoods for all particle identification hypotheses, normalized to their sum. The array subscripts correspond to the BES Monte Carlo particle numbering scheme (equation B.1). If the track has no data available for identification, then all the likelihoods will be zero.

B.4 Filling a Vector of Confidence Levels

```
INTEGER i
REAL    PIDCL(6)
INTEGER NDFPID
CALL TRACKCL(i,PIDCL,NDFPID)
```

The subroutine TRACKCL returns an array of global confidence levels for all possible particle identification hypotheses for the selected track. Note that the internal N-tuple vector TKVAR is *not* available to the calling user code.

The array PIDCL will be filled on return with the confidence levels, combined from all available detector systems for all particle identification hypotheses. NDFPID will indicate the number of degrees of freedom used to compute the confidence levels, typically one for each detector system. The array subscripts correspond to the BES Monte Carlo particle numbering scheme (equation B.1). If the track has no data available for identification, then NDFPID = 0.

B.5 Computing a Normalized Weight

```
EXTERNAL TRACKWT
REAL     TRACKWT, WEIGHT
INTEGER  i, hypgood, hypbad
WEIGHT = TRACKWT(i,hypgood,hypbad)
```

The function TRACKWT computes a normalized weight, or likelihood ratio, comparing two selected hypotheses. The two hypothesis codes hypgood and hypbad correspond to the subscripts a and b, respectively, in equation 7.2, so that the returned weight is close to 1 if the track matches hypgood, and close to 0 if the track matches hypbad.

The track is required to be consistent with at least one of the hypotheses. If $CL(\texttt{hypgood}) < \texttt{TRKCUT}$ and $CL(\texttt{hypbad}) < \texttt{TRKCUT}$, or if any of the calling arguments i, hypgood, or hypbad are invalid, the function returns a weight of -1.

A specific use of TRACKWT is in π/K separation. One BES particle identification routine [73] calculates a normalized weight XPIK which is close to 1 for a good kaon, and close to 0 for a good pion; the track is rejected if neither hypothesis is above 0.01. Using the TRACKID package, XPIK = TRACKWT(i, 5, 4).

B.6 Filling a Row-wise HBOOK N-tuple

```
INCLUDE 'BESINCN:TRKTUP.INC'
INTEGER i
CALL TRKTUP(i,TKVAR)
```

The one-dimensional array (N-tuple) TKVAR is defined in TRKTUP.INC. TRKTUP fills this array with the reconstructed data for the ith track in the event, with some quantities recomputed using the various empirical corrections discussed earlier. The N-tuple also contains the confidence levels and likelihoods for all of the appropriate identification hypotheses (isolated photon for neutral tracks or $\{e, \mu, \pi, K, p\}$ for charged tracks) and the number of particle ID degrees of freedom.

Symbolic parameters are defined in TRKTUP.INC to reference each of the elements of the N-tuple (users should reference TRKTUP.INC on their local system for a full list). Base indices are also defined for groups of related data. Thus,

TKVAR(TKV_PTRK)	=	Reconstructed track momentum
TKVAR(TKV_XS + 5)	=	Number of dE/dx sigma from kaon hypothesis.

The N-tuple array TKVAR is defined as a *local* variable only. It is not included in any common block. To share the data in it between several routines, the array must be defined with

```
INCLUDE 'BESINCN:TRKTUP.INC'
```

in the top-level (parent) routine, where TRKTUP is called, and passed as an argument to all the underlying (child) routines. In the latter routines, the argument should also be declared by using the INCLUDE statement, *not* by declaring "REAL TKVAR(*)" explicitly. See the TRACKWT source code for an example of such a "parent" routine, and TRKTUP.FOR itself for a "child" routine receiving TKVAR as an argument.

B.7 Passing a Filled N-tuple

All of the TRACKID user-interface routines described above have an alternate calling sequence, where a previously filled TKVAR N-tuple is passed in place of the track index i.

```
INCLUDE 'BESINCN:TRKTUP.INC'
EXTERNAL TKVARID, TKVARLK, TKVARCL, TKVARWT
INTEGER  TKVARID, IDBEST
REAL     TKVARWT, WEIGHT
INTEGER  i, hypgood, hypbad
REAL     clmin
REAL     PIDLIKE(6), PIDCL(6)
INTEGER  NDFPID
```

```
CALL TRKTUP(i,TKVAR)
IDBEST = TKVARID(TKVAR,hypgood,hypbad,clmin)
WEIGHT = TKVARWT(TKVAR,hypgood,hypbad)
CALL TKVARLK(TKVAR,PIDLIKE(6))
CALL TKVARCL(TKVAR,PIDCL(6),NDFPID)
```

The valid argument requirements, return values, and error conditions for the TKVAR*xx* routines are the same as those described previously for the TRACK*xx* versions.

B.8 Configuration

The TRACKID package has a number of configuration parameters, all stored in the common block /TRKCUTS/, in the TRKCUTS.INC include file. There are parameters for error protection (to prevent over- and underflows during calculations), for data quality checks, and for selection of detector systems.

The parameters are initialized in a BLOCK DATA subprogram loaded automatically if any TRACKID routine is called by the user. The default values can be overridden simply by assigning a new value to the appropriate variable in user code (*not* by using DATA statements).

CHIMAX (REAL*4) specifies the maximum value of χ in expressions of the form $\exp(-\chi^2/2)$. The default value, CHIMAX = 13, corresponds to 2×10^{-37}, which is nearly the single-precision (REAL*4) limit on most platforms.

TRKCUT (REAL*4) specifies the minimum confidence level required for a track to be consistent with a given hypothesis. It is used in the TRACKWT function to verify that the track is consistent with one of the two hypotheses requested by the user. The default value is TRKCUT = 0.01.

The total χ^2 for data consistency checking may include contributions from every detector system. The parameter TRKSYS (INTEGER*4) is a bit mask which specifies which detector systems should be used in the actual calculation in routine CHI2CHG. TRKCUTS.INC defines symbolic parameters for each of the bit positions, so the user does not need to use numeric values:

```
INTEGER TRKSYS_DC, TRKSYS_TOF, TRKSYS_SC
INTEGER TRKSYS_DEDX, TRKSYS_MU, TRKSYS_ECTOF
INTEGER TRKSYS_ESC


PARAMETER (TRKSYS_DC    =  1)
PARAMETER (TRKSYS_TOF   =  2)
PARAMETER (TRKSYS_SC    =  4)
PARAMETER (TRKSYS_DEDX  =  8)
PARAMETER (TRKSYS_MU    = 16)
PARAMETER (TRKSYS_ECTOF = 32)
PARAMETER (TRKSYS_ESC   = 64)
```

The default value is TRKSYS = 10 = TRKSYS_TOF + TRKSYS_DEDX, so that only the TOF and dE/dx systems contribute to particle identification.

Preselection of the data in each subsystem is done for charged track fits, TOF, and dE/dx data. All of the data quality variables are INTEGER*4 values or arrays.

For drift chamber track fitting, MFITOK is an array of three acceptable values of the MFIT quality code. NMFITOK indicates how many MFITOK values have been set: NMFITOK = 0 accepts all charged tracks. The default values are NFMITOK = 2 and MFITOK(i) = 2, −19.

For the time of flight data, TQMASK is a bit mask indicating acceptable values of the TQUAL quality code. The position (0 to 31) of each set bit is used as a TQUAL value which is acceptable. The default value is TQMASK = $0x0086$, corresponding to TQUAL = 1, 2 or 7.

For dE/dx energy loss reconstruction, MINDEDX selects the minimum number of hits for a reasonable dE/dx value. The default value is MINDEDX = 15.

An additional data quality check is available, for analyses which have a known and limited number of *true* track hypotheses (e.g., a kinematically pure sample of $\pi^{\pm}$ and $K^{\pm}$ tracks). The parameter HYPSYS (INTEGER*4) is a bit mask which specifies a list of hypotheses with which each detector subsystem (as selected in TRKSYS) must be consistent. TRKCUTS.INC defines symbolic parameters for each of the bit positions in HYPSYS, so the user does not need to use numeric values:

```
INTEGER HYPSYS_G, HYPSYS_E, HYPSYS_MU
INTEGER HYPSYS_PI, HYPSYS_K, HYPSYS_P


PARAMETER (HYPSYS_G  =  1)
PARAMETER (HYPSYS_E  =  2)
PARAMETER (HYPSYS_MU =  4)
PARAMETER (HYPSYS_PI =  8)
PARAMETER (HYPSYS_K  = 16)
PARAMETER (HYPSYS_P  = 32)
```

The default value is HYPSYS = 0. This disables system consistency checking. All data which pass the quality checks above will be used for particle identification.

The consistency requirement is imposed as a hard-wired $\chi^2_{\text{system},h} < 9$ for each hypothesis h in HYPSYS, in the routine SYSCHECK (INTEGER*4). If a detector system fails this consistency check for every hypothesis listed in HYPSYS, the data is presumed to be invalid and is not used for the track. SYSCHECK returns an integer value which may be substituted for TRKSYS before particle identification is done.

Bibliography

[1] L. Montanet *et al.*, Phys. Rev. **D50**, 1173 (1994), and 1995 off-year partial update for the 1996 edition available on the PDG WWW pages ⟨URL: `http://pdg.lbl.gov/`⟩.

[2] S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967).

[3] A. Salam, in *Elementary Particle Theory*, edited by N. Svaratholm (Almquist and Forlag, Stockholm, 1968).

[4] M. Gell-Mann and Y. Ne'eman, *The Eightfold Way* (W. Benjamin, Reading, Massachussetts, 1964).

[5] S. Drell and T. Lee, Phys. Rev. **D5**, 1738 (1972).

[6] M. Dine and J. Sapirstein, Phys. Rev. Lett. **43**, 688 (1979).

[7] G. Altarelli, in *New Phenomena in Lepton-Hadron Physics*, Vol. 49 of *NATO ASI B: Physics* (Addison Wesley, Menlo Park, CA, 1978), .

[8] G. Altarelli, Phys. Rep. **81C**, 1 (1982).

[9] R. van Royen and V. Weisskopf, Nuovo Cimento **50A**, 617 (1967), **51A**, 583(E) (1967).

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

[11] P. Abreu *et al.*, Nucl. Phys. **B418**, 403 (1994).

[12] M. Acciarri *et al.*, Z. Phys. **C62**, 551 (1994).

[13] R. Akers *et al.*, Z. Phys. **C61**, 19 (1994).

[14] D. Buskulic *et al.*, Z. Phys. **C62**, 539 (1994).

[15] C. P. E. Project, The Standard Model of Fundamental Particles and Interactions, 1995, available electronically from ⟨URL: `http://pdg.lbl.gov/cpep.html`⟩.

[16] T. Bowles and V. Gavrin, Annu. Rev. Nucl. Part. Sci. **43**, 117 (1994).

[17] A. S. Kronfeld and P. B. Mackenzie, Annu. Rev. Nucl. Part. Sci. **43**, 793 (1993).

[18] J. D. Richman and P. R. Burchat, submitted to Reviews of Modern Physics (unpublished).

[19] J. L. Rosner, Phys. Rev. **D42**, 3732 (1990).

[20] P. J. O'Donnell, Phys. Lett. **261B**, 136 (1991).

[21] M. H. Ye and Z. P. Zheng, Int. J. Mod. Phys. A **2**, 1707 (1987).

[22] M. H. Ye and Z. P. Zheng, in *Proceedings of the 1989 International Symposium on Lepton and Photon Interactions at High Energies*, Stanford University, edited by M. Riordan (World Scientific, Stanford, California, 1989), p. 122.

[23] J. Z. Bai *et al.*, Nucl. Instrum. Meth. **A344**, 319 (1994).

[24] X. Lou, in *Proceedings of the 17th International Symposium on Lepton and Photon Interactions*, Chinese Institute for High Energy Physics (World Scientific, Beijing, 1995), contributed paper.

[25] Hu Xiaoqing, `STRIP.FOR`, BES internal software.

[26] C. G. Matthews, Ph.D. thesis, California Institute of Technology, Pasadena, California, 1993.

[27] M. Bauer, B. Stech, and M. Wirbel, Z. Phys. **C34**, 103 (1987).

[28] E. Eichten *et al.*, Phys. Rev. **D21**, 203 (1980).

[29] F. James, MINUIT *Minimization Package*, CERN (European Particle Physics Laboratory), Geneva, Switzerland, 1994.

[30] F. Porter, Nucl. Instrum. Meth. **A368**, 793 (1996).

[31] S. Giani *et al.*, GEANT *Detector Description and Simulation Tool*, CERN (European Particle Physics Laboratory), Geneva, Switzerland, 1994.

[32] A. Breakstone, Lan Huibin's Mu Pair Tapes, posted to `bes.physics` newsgroup, message ID ⟨`93293.093525ALANBJ@SLACVM.SLAC.STANFORD.EDU`⟩.

[33] Chen Shaomin, `WPIPI.FOR`, BES internal software.

[34] S. Wasserbach, *Telesis*, Mark III Collaboration, Stanford, California, 1988.

[35] Lan Huibin, *Telesis*, BES Collaboration, IHEP, Beijing, China, 1993, adapted from Mark III version.

[36] E. Soderstrom, `SLACDX.FOR`, BES internal software.

[37] Li Weiguo, `DEDXCOR.FOR`, BES internal software.

[38] Li Weiguo, `TOFFSIG.FOR`, BES internal software.

[39] *Techniques and Concepts of High-Energy Physics IV*, Vol. 164 of *NATO ASI B: Physics*, edited by T. Ferbel (Addison Wesley, Menlo Park, CA, 1987).

[40] L. Landau, J. Phys. USSR **8**, 201 (1944).

[41] B. Rossi, *High-Energy Particles* (Prentice-Hall, New York, 1952).

[42] M. Kelsey, Energy Loss in the BES Beampipe, 1993, BES internal memorandum (BES Note 93–005).

[43] Chen Shenjian, BWSPFIT *Resonance Fitting Package*, BES Collaboration, IHEP, Beijing, China, 1994.

[44] J. Bai *et al.*, Phys. Rev. **D52**, 3781 (1995).

[45] J. Izen, Questions to PDG about asymmetric errors, Private communication.

[46] D. Boutigny *et al.*, Technical Report No. SLAC-R-95–457, (*BABAR* Collaboration), SLAC, Stanford, California (unpublished).

[47] S. Aoki *et al.*, Progress of Theor. Phys. **89**, 131 (1993).

[48] D. Acosta *et al.*, Phys. Rev. **D49**, 5690 (1994).

[49] D. Gibaut *et al.* (unpublished).

[50] J. Bai *et al.*, Phys. Rev. Lett. **74**, 4599 (1995).

[51] J. F. Amundson, Phys. Rev. **D52**, 2926 (1995).

[52] D. S. Hwang and G.-H. Kim, Phys. Lett. **B367**, 353 (1996).

[53] C. Bernard, T. Draper, G. Hockney, and A. Soni, Phys. Rev. **D38**, 3540 (1988).

[54] C. W. Bernard, J. N. Labrenz, and A. Soni, Phys. Rev. **D49**, 2536 (1994).

[55] T. A. DeGrand and R. D. Loft, Phys. Rev. **D38**, 954 (1988).

[56] C. Dominguez and N. Paver, Phys. Lett. **197B**, 423 (1987).

[57] N. Isgur and M. B. Wise, in *International Conference on Hadron Spectroscopy (Hadron '91)*, edited by S. Oneda and D. C. Peaslee (World Scientific, College Park, MD, 1991), pp. 549–572, invited talk.

[58] P. Colangelo, G. Nardulli, and M. Pietroni, Phys. Rev. **D43**, 3002 (1991).

[59] R. R. Mendel and H. D. Trottier, Phys. Lett. **B231**, 312 (1989).

[60] S. Capstick and S. Godfrey, Phys. Rev. **D41**, 2856 (1990).

[61] D. Bortoletto and S. Stone, Phys. Rev. Lett. **65**, 2951 (1990).

[62] J.-G. Bian and T. Huang, Modern Phys. Lett. **A8**, 635 (1993).

[63] S. Narison, Phys. Lett. **198B**, 104 (1987).

[64] E. Shuryak, Nucl. Phys. **B198**, 83 (1982).

[65] H. Krasemann, Phys. Lett. **96B**, 397 (1980).

[66] M. Suzuki, Phys. Lett. **162B**, 392 (1985).

[67] S. Sinha, Phys. Lett. **178B**, 110 (1986).

[68] P. Cea, P. Colangelo, L. Cosmai, and G. Nardulli, Phys. Lett. **206B**, 691 (1988).

[69] V. Galkin, A. Mishurov, and R. Faustov, Sov. J. Nucl. Phys **53**, 1026 (1991), in translation.

[70] M. Kelsey, in *8th Meeting of the Division of Particles and Fields*, American Physical Society (World Scientific, Albuquerque, New Mexico, 1994).

[71] W. S. Lockman, D and D_s Production in the Range $3.8 < \sqrt{s} < 4.5$ GeV, Mark III internal memorandum.

[72] W. S. Lockman, `FCALC FORTRAN`, Mark III internal software.

[73] Li Weiguo, `PARTID1.FOR`, BES internal software.