ABSTRACT

Title of Dissertation: MEASUREMENTS OF THE CP CONTENT
AND CP VIOLATING ASYMMETRIES
IN NEUTRAL B DECAYS TO TWO D* MESONS
WITH THE BABAR DETECTOR

Vincent A. Lillard, Doctor of Philosophy, 2004

Dissertation directed by: Associate Professor Douglas Roberts
Department of Physics

This dissertation presents a measurement of time-dependent CP-violating asymmetries and a measurement of the CP-odd parity fraction in the decay $B^0 \to D^{*+}D^{*-}$. The measurements are derived from a data sample of $88 \times 10^6 B\bar{B}$ pairs collected by the *BABAR* detector at the PEP-II asymmetric-energy B Factory located at the Stanford Linear Accelerator Center. A one-dimensional angular analysis of the decay products measures the CP-odd fraction to be $0.063 \pm 0.055(\text{stat}) \pm 0.009(\text{syst})$, indicating that the $D^{*+}D^{*-}$ final state is mostly CP-even. The time-dependent CP asymmetry parameters $\mathcal{I}m(\lambda_+)$ and $|\lambda_+|$ are determined from an analysis of the time-dependence of flavor-tagged B decays. One neutral B meson is fully reconstructed in a $D^{*+}D^{*-}$ final state, while the other B is inclusively reconstructed in order to determine its flavor. The Standard Model predicts the CP asymmetry parameters $\mathcal{I}m(\lambda_+)$ and $|\lambda_+|$ to be $-\sin 2\beta$ and 1, respectively, in the absence of penguin diagram contributions. They are determined to be $0.05 \pm 0.29(\text{stat}) \pm 0.10(\text{syst})$ and $0.75 \pm 0.19(\text{stat}) \pm 0.02(\text{syst})$, respectively, which corresponds to a 2.5 sigma deviation from Standard Model predictions with penguin contributions ignored.

MEASUREMENTS OF THE CP CONTENT

AND CP VIOLATING ASYMMETRIES

IN NEUTRAL B DECAYS TO TWO D* MESONS

WITH THE BABAR DETECTOR

by

Vincent A. Lillard

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2004

Advisory Committee:

   Associate Professor Douglas Roberts, Chairman/Advisor
   Associate Professor Andrew Baden
   Professor Abolhassan Jawahery
   Professor Jogesh Pati
   Associate Professor Sylvain Veilleux

DEDICATION

This dissertation is dedicated to my mom, who raised and sacrificed for me, my dad, who instructed and guided me, and my wife who loves and supports me.

ACKNOWLEDGEMENTS

The conclusion of my graduate career and the completion of this dissertation represents even more than the difficult life commitment and academic accomplishment that graduate degrees symbolize. It also signifies the conclusion of a rich season of my life guided and directed by God, who has demonstrated his goodness and faithfulness through the people and experiences of these years.

I am immensely grateful to my advisor, Doug Roberts. His commitment and patience for me has been quite evident over the years. His ability to mentor me has been invaluble for my academic development, and his friendship has made even the difficult times enjoyable.

Working on *BABAR* has been an exciting and rewarding scientific experience. I owe much of my professional development to the many presentations, deliberations, and "reviews" afforded by the collaborative experience. Specifically, I am grateful to the D^*D^* group, who shared in the work and frustrations of publishing these results. Jochen patiently explained concepts to me early on, while Justin, Gloria, Lorenzo and Chung worked hard to get results out "in time". Much of this dissertation represents their commitment and efforts.

There are many special people that support and enable me to succeed in life. While I can not mention everyone, I must thank my close friends, Dusan, Jonathan, Dave, and Amir who walked through the years of graduate school along side me, supporting and encouraging me each in their own way. The Hempels and the Lymans, having made me a part of their families, provided the acceptance and encouragement I will always cherish.

Finally, I thank my parents and my wife, for their unwavering love and support. My mom sacrificed for many years, that I might have a successful life in spite of the many hardships she endured. My dad was patient to teach me even when I was unwilling to listen; his guidance produced in me the focus required in this endeavor. My wife, Sarah, now enables me to finish with confidence and have vision for the future as a new chapter begins.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

1.1   Symmetries in the Physical Laws

Studying and understanding symmetries and symmetry violations in the physical laws is one of the most important pursuits of physicists. A transformation which leaves a physical law unchanged results in an important constraint on the state of a system. In fact, Nöther's theorem [1] states that for each continuous symmetry of a theory, there is an associated conserved quantity. For example, the conservation of energy is a consequence of the invariance of physical laws under time translations. Invariance under spatial translations and rotations result in the conservation of linear momentum and angular momentum.

In addition to continuous transformations, there exists three important discrete transformations[1]:

- *Time reversal*, T, which changes the sign of the time coordinate $(t \rightarrow -t)$ in equations of motion;

- *Parity*, P, which inverts the space coordinates $(\vec{x} \rightarrow -\vec{x})$ of a particle. Axial

---

[1]The existence of a discrete symmetry does not necessitate an associated conserved quantity, but does constrain the physical system by limiting the form of the Lagrangian.

vectors, such as the spin of a particle, are unaffected by this transformation; and

- *Charge conjugation*, C, which transforms a particle in its antiparticle by changing its electrical charge and other *quantum numbers*, but leaves its space-time coordinates unchanged.

A study of classical mechanics and electrodynamics reveals that all interactions are invariant under any of the above transformations. Naively, one would expect all fundamental particle interactions to obey the same symmetries. However, it was demonstrated in 1957 by C. Wu *et al.* that P was violated in the decays of ^{60}Co nuclei [2]. Further, the study of neutrinos reveals that neutrinos are always left-handed (its spin $\vec{s}$ is anti-parallel to its momentum $\vec{p}$), while anti-neutrinos are always right-handed ($\vec{s}$ parallel to $\vec{p}$). The fact that right-handed neutrinos and left-handed anti-neutrinos do not exist in nature is a clear violation of C. A measurement of neutrino helicity was demonstrated first in 1958 by M. Goldhaber *et al.* [3]

Even the combined transformation of C and P, which was thought to be a symmetry of all particle interactions, was demonstrated to be violated in the decays of kaons. In 1964, J. Christenson, J. Cronin, V. Fitch, and R. Turlay proved the existence of CP violation from the discovery of the decay $K_L^0 \rightarrow \pi\pi$ [4]. And recently, measurements of CP asymmetries by the *BABAR* [6] and BELLE [7] collaborations established this effect in the decays of B mesons.

These symmetry violations are, however, clearly explained in the context of the Standard Model of particle interactions. The Standard Model, which describes the hundreds of observed elementary particles and their interactions, is able to account for these symmetry violating effects. The weak interactions between quarks is

regulated by complex coupling constants which can be parameterized by three real parameters and one irreducible complex phase. The source and magnitude of CP violating effects is proportional to this one complex phase. This implies that all CP violating asymmetries observed in nature should be the result of this one parameter, according to the Standard Model. Verifying this fact has become one the most important and interesting tests in particle physics.

The violation of CP symmetry is also very important for modern theories of cosmology. According to these theories, an equal amount of matter and antimatter was present in the early universe after the Big Bang. Presently, however, the universe appears to be composed of only matter. In our galaxy, for example, the primary cosmic-ray nuclei that we observe are composed of particles rather than anti-particles. [10] Large masses of antimatter could be detected through γ-ray emission from annihilation processes with cosmic matter. No such phenomena have been observed.

The existence of CP violation is, according to Sakharov [5], an essential ingredient to explain the abundance of matter in the universe. The small differences in the interactions of matter and anti-matter due to CP-violation could provide a mechanism to generate the observed matter-antimatter asymmetry. However, the predicted size of CP-violating effects in the Standard Model is not able to account for the absence of antimatter. In fact, the effect predicted by the Standard Model is many orders of magnitude too small to produce this asymmetry. This inconsistency motivates physicists to probe the sources and effects of CP-violation in fundamental particle interactions. Theories of the physics beyond the Standard Model generally provide more sources of CP violation, and predict effects which can be clearly distinguished from Standard Model predictions in experiment. Since CP-violating asymmetries are the result of merely one parameter in the Standard

Model, measuring asymmetries in B decays provides an excellent probe for new physics not yet observed.

1.2   Overview of the Contents

In this dissertation, a study of the $B^0 \to D^{*+}D^{*-}$ decay is presented. The expected magnitude of CP-violation in B meson decays is much larger (three orders of magnitude) than asymmetries in kaon decays. This was verified by the recent results of the $BABAR$ and BELLE collaborations, mentioned above. The decay $B^0 \to D^{*+}D^{*-}$ provides an alternate and independent test of the same CP asymmetry amplitude, $\sin 2\beta$. A comparison of measurements of $\sin 2\beta$ from $b \to c\bar{c}s$ modes such as $B^0 \to J/\psi K_S^0$ [8] with that obtained in $B^0 \to D^{*+}D^{*-}$ is an important test of the Standard Model.

A measurement of CP violating asymmetries in the $B^0 \to D^{*+}D^{*-}$ decay is complicated by the fact that the final state is not a CP eigenstate. Here, a pseudoscalar B meson decays to a vector-vector final state, which is composed of three partial waves with different CP parities: even for the S- and D-waves, odd for the P-wave. The relative fraction of CP-parities can be determined from an angular analysis of the decay products and would therefore enable a more accurate measurement of CP-violating asymmetries.

Therefore, in this dissertation a measurement of the CP-odd fraction of the $D^{*+}D^{*-}$ final state is presented. We then present a measurement of the time-dependent CP asymmetry in $B^0 \to D^{*+}D^{*-}$ decays, obtained from a combined analysis of the time dependence of flavor-tagged decays and the one-dimensional angular distribution of the decay products.

The dissertation is organized as follows:

- We begin with an overview of the Standard Model of particle physics and its provisions for CP violation. The necessary formalism for observing CP violation in the time evolution of B mesons is presented as well as the motivation for studying the specific decay $B^0 \rightarrow D^{*+}D^{*-}$.

- The large data sample used for these analyses was produced at the PEP-II collider and recorded by the *BABAR* detector. The primary characteristics of PEP-II and *BABAR* are described in Chapter 3.

- A detailed set of criteria is necessary to distinguish events reconstructed as $B^0 \rightarrow D^{*+}D^{*-}$ from other background events in the data sample. These selection criteria are described in detail in Chapter 4. The resulting distributions of events from the data sample is discussed in Chapter 5.

- The measurement of the CP-odd fraction of the $B^0 \rightarrow D^{*+}D^{*-}$ final state is described in Chapter 6.

- Three primary ingredients are necessary to measure CP-violating time-dependent asymmetries in $B^0 \rightarrow D^{*+}D^{*-}$ decays. They are discussed in Chapter 7.

- The fit procedure for incorporating these ingredients in the analysis of the reconstructed $B^0 \rightarrow D^{*+}D^{*-}$ events is discussed in Chapter 8. The results of the fit are also given.

- Chapter 9 summarizes the two measurements and discusses the implications of the results, as well as future prospects.

Chapter 2

Theoretical Background and Motivation

To understand the context of our study of B mesons (specifically the decay $B^0 \rightarrow D^{*+}D^{*-}$), we will first present a short overview of the Standard Model of particle physics. We explore important aspects of the theory relevant to our discussion of CP violation and B meson physics. We also explore CP violation, its general quantum-mechanical phenomenology, as well as how the Standard Model provides for it. The B meson system and its connection to CP violation will be explained as well. Finally, we conclude this chapter with the motivation for studying the decay $B^0 \rightarrow D^{*+}D^{*-}$, as well as a discussion of the implications of measurements in this mode.

2.1 Overview of the Standard Model

The Standard Model of particle physics describes the fundamental particles and their interactions as the result of three local gauge symmetries. We know there to exist six quarks and six leptons and their corresponding anti-particles. Their relationships and interactions are governed by the combined electro-weak symmetry group of $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ and the quantum chromo-dynamics symmetry group of $\mathrm{SU}(3)$. The $\mathrm{SU}(2)_L$ group describes a symmetry of "weak isospin" and couples to

	Family			Electric		Weak Charge		
Fermion	1	2	3	charge	Color	left-hd.	right-hd.	Spin
Leptons	ν_e	ν_μ	ν_τ	0	0	$\frac{1}{2}$	n/a	$\frac{1}{2}$
	e	μ	τ	-1	0	$-\frac{1}{2}$	0	$\frac{1}{2}$
Quarks	u	c	t	$\frac{2}{3}$	r,b,g	$\frac{1}{2}$	0	$\frac{1}{2}$
	d	s	b	$-\frac{1}{3}$	r,b,g	$-\frac{1}{2}$	0	$\frac{1}{2}$

Table 2.1: The fundamental particles and their quantum numbers.

Force	Coupling Charge	Particle(s) Exchanged	Symmetry
Electro-weak	Electric/weak	Photon (γ), $W^\pm$, Z^0	$U(1) \times SU(2)$
Strong	Color	8 Gluons (g)	$SU(3)$

Table 2.2: The fundamental interactions and their mediating particles.

the left-handed (chirality -1) states of the leptons and quarks. This is combined with a U(1) group symmetry with the "weak hypercharge" Y. The SU(3) group describes the interactions obtained as a result of the color charge that quarks hold and leptons do not. The fundamental particles and their quantum numbers are listed in Table 2.1.

By imposing local gauge invariance to the free-particle Lagrangians, we obtain the appropriate description of interactions between the fundamental particles. Further, with the constraint of local gauge invariance, the theory requires the introduction of gauge fields which act as mediators of the interactions. For the electromagnetic interaction, for example, the photon and it's corresponding gauge field A_μ are the result of imposing local gauge invariance on the Dirac Lagrangian.

The consequence of this method is that the resulting gauge fields are always massless. From experiment, however, we understand that the W^+,W^-, and the Z^0 are the mediators of the weak interaction, and that they are not massless. Thus, to obtain massive gauge fields, but retain the symmetry group of $SU(2)_L \times U(1)_Y$, the Standard Model must include a mechanism for these particles to "acquire" mass. This method is known as "spontaneous symmetry breaking" via the Higgs mechanism and was critical to the model of electroweak symmetry introduced by Weinberg and Salam in 1967-1968.

Table 2.3 displays the Minimal Standard Model's (MSM) three generations of quarks and leptons and the spin-zero Higgs boson. The suggestive notation used to express these fields is indicative of experimental observations. For example, the table lists no right-handed neutrino, reflecting that up to recently, there was no evidence for such a particle.

While the details of gauge theories and spontaneous symmetry breaking are interesting, a full and complete treatment is beyond the scope of this thesis. The reader is referred to any field theory or particle physics text (such as Refs. [12], [14], [16]) for the details of the Standard Model of particle physics. To understand the context of CP violation and B meson decays we will, however, discuss some important aspects of the weak sector of the Standard Model, specifically the weak mixing matrix, as well as the properties of discrete symmetries in nature.

2.1.1 Discrete Symmetries

In both classical and quantum theories of nature, the existence and study of discrete symmetries has been central to a full understanding of physics. Left-right symmetry, or parity symmetry, and time-reversal symmetry are two well-known invariances of classical mechanics and electromagnetism. Here, parity symmetry

Field	$U(1)$	$SU(2)$	$SU(3)$
$u_R^{i,\alpha}$	$2/3$	**1**	**3**
$d_R^{i,\alpha}$	$-1/3$	**1**	**3**
e_R^i	-1	**1**	**1**
$Q_L^i = \begin{pmatrix} u_L^{i,\alpha} \\ d_L^{i,\alpha} \end{pmatrix}$	$1/6$	**2**	**3**
$L_L^i = \begin{pmatrix} \nu_L^i \\ e_L^- \end{pmatrix}$	$-1/2$	**2**	**1**
$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}$	$1/2$	**2**	**1**

Table 2.3: The Minimal Standard Model matter fields. The L and R subscripts indicate left and right-handed fields, respectively. The $i = 1, 2, 3$ index enumerates the generations. The $\alpha = r, g, b$ is used for the $SU(3)$ transformations of the quarks. The $U(1)$ column lists the hyper-charge, while $SU(2)$ and $SU(3)$ columns list the dimension of representation of the fields under the respective gauge transformation.

(P) says that the system behaves exactly the same when the coordinate system is reversed $(\vec{x} \rightarrow -\vec{x})$; while time-reversal symmetry (T) indicates no time direction preference. In quantum field theory these transformations take the form of quantum mechanical operators which act on the particle fields. A third discrete symmetry applies to quantum systems – the charge-conjugation operator (C) changes a particle into its anti-particle.

Before the 1950's it was generally assumed that each of these transformations were fundamental symmetries of nature. However, Lee and Yang [18] determined that this was only an "extrapolated hypothesis" and further experiments showed some nuclear decays violated parity symmetry [2], while other experiments showed that all neutrinos were left-handed (*i.e.* helicity is -1), indicating C violation. Charge-conjugation violation however seems to always occur in conjunction with parity violation in the weak interactions. For example, consider the following related decays:

$$
\begin{aligned}
1: \quad & \pi^+ \rightarrow \mu_R^+ \, \nu_{\mu L}, \\
2: \quad & \pi^- \rightarrow \mu_L^- \, \bar{\nu}_{\mu R}, \\
3: \quad & \pi^+ \rightarrow \mu_L^+ \, \nu_{\mu R}, \\
4: \quad & \pi^- \rightarrow \mu_R^- \, \bar{\nu}_{\mu L}
\end{aligned}
\tag{2.1}
$$

Only the first two decays are observed in nature. We note that the parity operation, since it changes the sign of the helicity of a particle, transforms $(1) \rightarrow (3)$, and $(2) \rightarrow (4)$. But also we see that the C operation takes $(1) \rightarrow (4)$, and $(2) \rightarrow (3)$. So it then follows that the combined operation of C and P *is* a symmetry in nature, because we do see both (1) and (2). In general we see that CP symmetry is conserved for most weak interactions.

However, in 1964, Christensen *et al.* [4] discovered that CP symmetry was

10

violated with the discovery of the decay $K_L^0 \to \pi\pi$. In 2001, the *BABAR* and BELLE experiments discovered CP violation in B meson decays [6, 7]. In the Standard Model, the understanding of the source of CP violation is derived from the CKM matrix, and provides a framework in which to study and understand it. While the scope of this thesis does not cover CP violation in kaon decays, we will, in the following sections, explore this phenomenon in B decays.

To complete our discussion of discrete symmetries in the Standard Model, we must note the combined transformation of CPT. While any of the discrete symmetries C, P, or T, or any combination thereof, may be violated in nature, there is a strong theoretical prejudice against the possibility that CPT is violated. This is the "CPT theorem" which states that any quantum field theory that obeys Lorentz invariance as well as spin statistics must also be CPT invariant. The most basic consequences of CPT symmetry includes the requirement that the masses and decay widths of particle and its antiparticle be exactly equal. In fact, experiments to date are so far in complete agreement with the predictions of the CPT invariant phenomenology. [26]

2.1.2 Weak Flavor Mixing in the Standard Model

One of the primary "ingredients" of the Standard Model is the Higgs mechanism which bestows mass on all of the fundamental particles and breaks the electro-weak gauge symmetry. The theory introduces a scalar field which, by acquiring a vacuum expectation value, v, produces the appropriate mass terms in the Lagrangian. (see, for example, Refs. [16], [12]) The simplest model uses a Higgs doublet scalar field which also belongs to the electroweak gauge group SU(2)×U(1):

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}_L \quad \text{where } \phi^+, \phi^0 \quad \text{are complex fields}$$

We now consider how the Higgs couples to the quarks and leptons to give them mass. The Yukawa couplings arise from the following Lagrangian:

$$\mathcal{L}_Y = - \left(g_d^{ij} \bar{Q}_L^i \, \phi \, d_R^j + g_u^{ij} \bar{Q}_L^i \, \tilde{\phi} \, u_R^j + g_e^{ij} \bar{L}_L^i \, \phi \, e_R^j \right) + h.c. \qquad (2.2)$$

where the indices i, j are over the 3 generations of fermions, and $\tilde{\phi}$ denotes the conjugate SU(2) doublet $(\tilde{\phi} = i\tau_2 \phi^{\dagger T})$. The couplings g_u, g_d, g_e are, in general, 3×3 complex matrices. If we substitute in the above Lagrangian the vacuum expectation value for the Higgs field we obtain the mass terms

$$\mathcal{L}_Y = -\bar{d}_L \, \mathcal{M}_d \, d_R - \bar{u}_L \, \mathcal{M}_u \, u_R - \bar{e}_L \, \mathcal{M}_e \, e_R \qquad (2.3)$$

where $\mathcal{M}_d^{ij} = v \, g_d^{ij}$, $\mathcal{M}_u^{ij} = v \, g_u^{ij}$, and $\mathcal{M}_e^{ij} = v \, g_e^{ij}$ are the mass matrices. These matrices are not necessarily diagonal and therefore introduce mixing between the different generations of quarks. Hence, we observe that the Standard Model Lagrangian is not expressed in terms of mass eigenstates but instead the eigenstates of the interactions.

We can, however, rewrite the fields using a unitary transformation:

$$u_L = U_L^u \, u_L' \quad , \quad u_R = U_R^u \, u_R',$$
$$d_L = U_L^d \, d_L' \quad , \quad d_R = U_R^d \, d_R' \qquad (2.4)$$

so that $\mathcal{M}' = U_L^{\dagger k} \, \mathcal{M}_k \, U_R^k$ is the *diagonal* mass matrix. We note that this redefinition does not affect the kinetic terms of the Lagrangian, nor the Z^0 and A_μ (photon) couplings. Since neutrinos are massless in the Standard Model, the lepton fields may be chosen to be simultaneous mass and weak eigenstates, so the lepton terms are also unaffected. The only change in the Lagrangian is to the (left-handed) quark couplings to the W. The appropriate Lagrangian terms describing the W charged current coupling is written as

$$g \, \bar{u}_L \, \gamma^\mu \, d_L W_\mu^+ \rightarrow g \, \bar{u}_L'^i \, \gamma^\mu \, V_{CKM}^{ij} d_L'^j W_\mu^+ \qquad (2.5)$$

12

where i, j are the quark generation indices and g is the universal weak coupling constant. The above relation shows that by performing the transformation of Eq. 2.4, the matrix V_{ij} encapsulates our understanding of quark mixing in the weak interactions. This matrix was first introduced by Cabibbo, Kobayashi and Maskawa (CKM) [22], is defined as $U_L^{\dagger u} U_L^d$ and is written as

$$V_{\text{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{2.6}$$

where the subscripts denote the appropriate quark transition. For the $d \rightarrow u$ transition, for example, the CKM element corresponds to the well-known Cabibbo angle ($\cos\theta_c \approx 0.97$). The W boson thus couples to the "rotated" quark state $d' = V_{ud}\, d + V_{us}\, s + V_{ub}\, b$; the specific $d \rightarrow u$ transition selects out the appropriate *flavor* eigenstate and the transition is therefore proportional to V_{ud}. Hence, cross-generational quark transitions are obtained via the experimentally determined CKM matrix elements.

A survey of measurements of the magnitudes of CKM elements provide [23]:

$$|V_{\text{CKM}}| = \begin{pmatrix} 0.97504 \pm 0.00049 & 0.2221 \pm 0.0021 & 0.00270 - 0.00371 \\ 0.2220 \pm 0.0021 & 0.97414 \pm 0.00049 & 0.00387 - 0.00432 \\ 0.0072 - 0.0092 & 0.0380 - 0.0427 & 0.99907 - 0.99926 \end{pmatrix}. \tag{2.7}$$

We observe that transitions within each quark-lepton family is much more probable than those between the first and second, which are more probable than second and third, leaving transitions between the first and third generations as the least likely.

2.1.3 *CP* Violation in the Standard Model

As mentioned in Section 2.1.1, we have observed discrete symmetry violations of parity (P) and charge conjugation (C). This is built in to the Standard Model

through the absence of right-handed neutrinos and the lack of $SU(2)$ charge for the right-handed fermions (see Table 2.3). However, finding a source of CP violation requires a closer look at the theory.

We note that pure gauge Lagrangians are necessarily CP-invariant. [24] The QCD Lagrangian is, in fact, manifestly CP-invariant [1] so we are left with the Higgs mechanism as the only possible source of CP-violation. The scalar potential of one Higgs doublet clearly conserves CP [2] and the Yukawa interactions of the fermions with the Higgs do not reveal any source of CP violation. However, the redefinition of the fermion fields as mass eigenstates introduced the CKM matrix in the flavor changing terms of the Lagrangian. We identify this as the only potential source of CP violation.

If we apply the CP operation to the relevant weak interaction terms, we find that

$$(CP)g\,\bar{u}_L^{\prime i}\gamma^\mu (V_{\text{CKM}})^{ij}d_L^{\prime j}W_\mu^+(CP)^\dagger = e^{i\phi}g\,\bar{d}_L^{\prime i}\gamma^\mu (V_{\text{CKM}})^{ij}u_L^{\prime j}W_\mu^-$$

The right hand side above is different from the hermitian conjugate term by only a phase. Thus, CP conservation requires that

$$V_{ij}^* = e^{i\,\phi}V_{ij} \tag{2.8}$$

Here ϕ is an arbitrary phase, which may be chosen to satisfy this condition for one matrix element. This condition, however, is not necessarily satisfied for all elements of the CKM matrix. Hence, if more than one element of the CKM matrix is complex, then CP conservation is violated in the Standard Model.

---

[1] A potential source of CP violation in QCD is through the term $L_\theta = \bar{\theta}\frac{g_s^2}{32\pi^2}F_{\mu\nu}^i F_i^{\mu\nu}$. This is the "strong CP problem" and is beyond the scope of this discussion. The reader is referred to [17], for example. Measurements constrain $\bar{\theta} < 3 \times 10^{-10}$ [19], which is significantly smaller than the parameters of the CKM matrix.

[2] More than one Higgs boson may lead to more sources of CP violation.

The CKM matrix, V, is unitary as a consequence of its definition $V_{CKM} \equiv U_L^{\dagger u} U_L^d$ (see Eq. 2.4); $U_L^{\dagger u}$ and U_L^d are required to be unitary so that $|\bar{u}_L\, u_L| = |\bar{u}'_L\, u'_L|$ and $|\bar{d}_L\, d_L| = |\bar{d}'_L\, d'_L|$. In general, an $N \times N$ unitary matrix may be parameterized by N^2 parameters, but $2N - 1$ phases may be absorbed or changed by re-phasing the quark fields. Any orthogonal matrix is parameterized by $N(N-1)/2$ rotation angles; given that a unitary matrix is the complex extension of an orthogonal matrix, there remains $(N-1)(N-2)/2$ physical phases that cannot be rotated away. For two quark generations, V is defined by one real parameter (the Cabibbo angle) and no physical phases remain – thus, CP violation cannot exist in the Standard Model unless there are 3 or more quark generations. The existence of a third generation was in fact postulated by Kobayashi and Maskawa before the discovery of the b-quark in 1977.

Unitarity of the CKM matrix requires that

$$V^\dagger V = V\, V^\dagger = 1 \quad \Rightarrow \quad \sum_j V_{ji}^* V_{jk} = \sum_j V_{ij} V_{kj}^* = \delta_{ik} \tag{2.9}$$

which results in 9 independent relations. Since each element of the CKM matrix is in principle directly measurable, the consistency of the unitarity conditions may be experimentally confirmed. Any evidence of the failure of the unitarity of V is an indication of new physics beyond the CKM picture of flavor changing processes or evidence for more than three generations of quarks and leptons. The latter possibility is constrained by the number of light neutrino flavors contributing to the width of the Z boson. [21] Therefore, discovering that the CKM matrix elements do not obey the unitarity relations is a good indication of physics beyond the Standard Model.

In order to reflect the possibility of complex elements within the CKM matrix, a convenient parameterization was suggested by Wolfenstein [25] in terms of the

parameter [3]$\lambda \approx 0.22$

$$V = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \qquad (2.10)$$

This parameterization reflects the understanding that transitions within each quark family is much more probable than those between the first and second generations, which are more probable than second and third, and so on. We therefore have 4 real parameters (λ, A, ρ, η) to describe the CKM matrix, and a clear formulation for the existence of CP violating complex phases.

The relations in Eq. 2.9 represent six triangles and three rectangles in the complex plane. In particular, the relation containing all the b quark elements

$$V_{ub}^* V_{ud} + V_{cb}^* V_{cd} + V_{tb}^* V_{td} = 0 \qquad (2.11)$$

is used to pictorially represent the irreducible CP violating phase and is referred to as the Unitarity Triangle. ρ and η then describe the x and y position of the upper vertex of the Unitarity Triangle as shown in Figure 2.1. The angles of the triangle are given by

$$\alpha = \arg\left[-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right], \beta = \arg\left[-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right], \gamma = \arg\left[-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right], \qquad (2.12)$$

and are sensitive to B meson decays to specific final states and are in principle experimentally measurable. Confirming that $\alpha + \beta + \gamma = \pi$ is therefore an important test of the unitarity of the CKM matrix, and has become a primary focus of modern particle physics laboratories.

---

[3]Note this is Cabibbo parameter $\lambda = \theta_c$ and not related to the CP violation parameter λ detailed in Section 2.2.2.
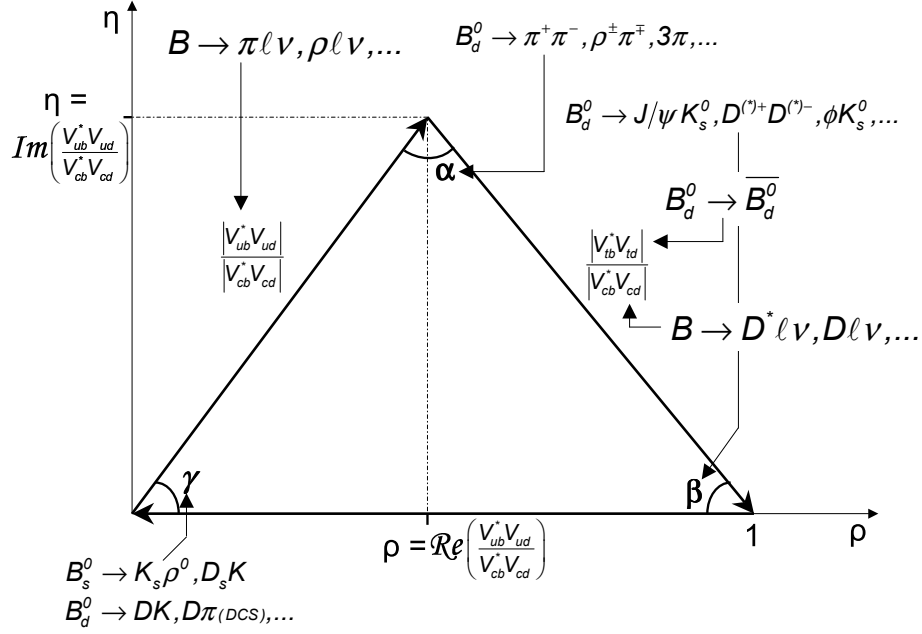
Figure 2.1: The Unitarity Triangle in the $\rho - \eta$ plane. In the B meson system, angles are measurable via time-dependent asymmetries in the modes listed, sides are measurable via semileptonic branching fractions and mixing frequency (with some theoretical error in the CKM extraction).

2.2 CP Violation

As described in Section 2.1.1, the combined discrete transformations of (C) interchanging particles and anti-particles and (P) performing a reflection of the space axes through the origin is a symmetry of nature in most cases. In 1964, Christensen, Cronin, Fitch, and Turlay [4] were able to demonstrate experimentally that CP symmetry was violated in the $K^0 - \overline{K}^0$ meson system. This had far reaching implications, including the prediction of a third generation of quarks and the CKM matrix as the framework for such phenomena. Recently, the *BABAR* and BELLE experiments have confirmed the existence of CP violation in the B meson system, producing results consistent with the Standard Model's framework of CP violation.

However, to date, the theoretical understanding for the source of CP violation is not completely understood.

Besides being a fascinating effect because of its elusiveness at both the experimental and theoretical levels, CP violation might also play an important role in our understanding of cosmology. The observed baryon asymmetry of the Universe could only be generated from an initial situation of equal amounts of matter and antimatter if there is CP violation. However, the CP violation predicted by the Standard Model is not sufficient to account for the entire effect. Hence, study continues in order to understand both the picture the Standard Model provides as well as the search for possibilities beyond the phenomena predicted by the Standard Model ("new physics").

2.2.1 CP Violation Phenomenology

While we have seen a potential source of CP violation in the Standard Model (Section 2.1.3), we now discuss the framework for observing CP violation phenomena in the decays of mesons. We start by considering a simple two-state quantum mechanical system, where a particle X^0 and its antiparticle $\bar{X}^0$ are allowed to decay and to mix (indicating they may not be simultaneous mass and flavor eigenstates – this being the case for K^0 and B^0 mesons). The time-evolution of the state $|\psi(t)\rangle = \alpha(t)|X^0\rangle + \beta(t)|\bar{X}^0\rangle$ is determined by the Schrödinger equation:

$$i\frac{\partial}{\partial t}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \mathcal{H}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \equiv \begin{pmatrix} m_{11} - \frac{1}{2}i\Gamma_{11} & m_{12} - \frac{1}{2}i\Gamma_{12} \\ m_{21} - \frac{1}{2}i\Gamma_{21} & m_{22} - \frac{1}{2}i\Gamma_{22} \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \tag{2.13}$$

The angle in the complex plane of m_{12} represents the phase of the mixing between the flavor eigenstates, and Γ_{12} represents the (complex) coupling to common decay modes of X^0 and $\bar{X}^0$ (for example, $B^0/\bar{B}^0 \rightarrow J/\psi K_S^0$). CPT invariance guarantees that $m_{11} = m_{22}$ and $\Gamma_{11} = \Gamma_{22}$, and that $m_{21} = m_{12}^*$ and $\Gamma_{21} = \Gamma_{12}^*$. The mass

eigenstates are the eigenvectors of the Hamiltonian:

$$|X_L\rangle = p|X^0\rangle + q|\bar{X}^0\rangle$$

$$|X_H\rangle = p|X^0\rangle - q|\bar{X}^0\rangle \tag{2.14}$$

where $|X_L\rangle$ and $|X_H\rangle$ are the lighter and heavier mass eigenstates, respectively, with

$$\frac{q}{p} = \sqrt{\frac{m_{12}^* - \frac{1}{2}i\Gamma_{12}^*}{m_{12} - \frac{1}{2}i\Gamma_{12}}} = \frac{\Delta m - \frac{i}{2}\Delta\Gamma}{2(m_{12} - \frac{1}{2}i\Gamma_{12})} \tag{2.15}$$

and the requirement that $|p|^2 + |q|^2 = 1$. The mass difference $\Delta m = m_H - m_L$ and decay width difference $\Delta\Gamma = \Gamma_H - \Gamma_L$ can be obtained by diagonalizing the "mixing matrix" of Eq. 2.13. An initially pure $|X^0\rangle$ state will, therefore, time evolve as a superposition of the mass eigenstates $|X_L\rangle$ and $|X_H\rangle$. For a B meson, the mixing between states can occur through the "box" diagrams of Figure 2.2.

Following the algebra of the Schrödinger equation, we find the time dependent
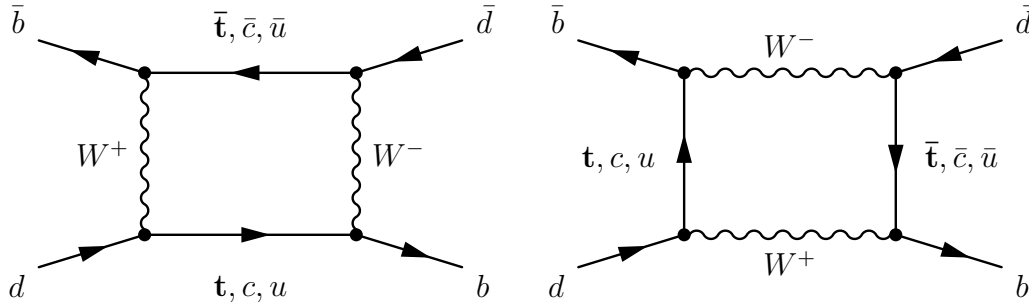
Figure 2.2: The two primary mixing diagrams for neutral B mesons.

mass eigenstates are

$$|X_L(t)\rangle = e^{-im_L t - \Gamma_L t/2}|X_L\rangle$$

$$|X_H(t)\rangle = e^{-im_H t - \Gamma_H t/2}|X_H\rangle \tag{2.16}$$

Using the relationship in Eq. 2.14 the time dependent *flavor* eigenstates are given by

$$|X^0(t)\rangle = e^{-iMt-\Gamma t/2}\left(\cos(\Delta mt/2)|X^0\rangle + i\frac{q}{p}\sin(\Delta mt/2)|\bar{X}^0\rangle\right)$$

$$|\bar{X}^0(t)\rangle = e^{-iMt-\Gamma t/2}\left(\cos(\Delta mt/2)|\bar{X}^0\rangle + i\frac{p}{q}\sin(\Delta mt/2)|X^0\rangle\right) \qquad (2.17)$$

Here the assumption that

$$\Delta\Gamma \ll \Delta m \qquad (2.18)$$

has been made to simplify the expression. For the B mesons, this is a safe approximation, as it is expected that $\Delta\Gamma = \mathcal{O}(10^{-3})\Delta m$. The difference in width is produced by decay channels common to B^0 and $\overline{B}^0$; because the mass of the B is so large, the phase space does not suppress flavor-specific decays (as in the K system), and hence the resulting lifetimes are nearly equivalent.[4]

We now consider the decay of a B meson to a final state f which is an eigenstate of CP:

$$CP|f\rangle = \eta_f|f\rangle$$

where η_f is the CP eigenvalue of f. The decay amplitudes are defined as:

$$A = \langle f|\mathcal{H}|B^0\rangle$$

$$\bar{A} = \langle f|\mathcal{H}|\overline{B}^0\rangle \qquad (2.19)$$

Using Eq. 2.17 the decay rates are given by the absolute square of the amplitudes:

$$
\begin{aligned}
|\langle f|\mathcal{H}|B^0(t)\rangle|^2 &= e^{-\Gamma t}|A|^2\,|\cos(\Delta mt/2) + i\lambda\sin(\Delta mt/2)|^2 \\
&= e^{-\Gamma t}|A|^2\left[\frac{1}{2}(1+|\lambda|^2) + \frac{1}{2}(1-|\lambda|^2)\cos(\Delta mt) - \mathcal{I}m\lambda\sin(\Delta mt)\right]
\end{aligned}
$$

and

$$|\langle f|\mathcal{H}|\overline{B}^0(t)\rangle|^2 = e^{-\Gamma t}|A|^2\left[\frac{1}{2}(1+|\lambda|^2) - \frac{1}{2}(1-|\lambda|^2)\cos(\Delta mt) + \mathcal{I}m\lambda\sin(\Delta mt)\right] \quad (2.20)$$

[4]Note that $\Delta\Gamma_{B_d}$ has not been measured, but this approximation is generally considered to be a model-independent assumption [15].

where λ is given as

$$\lambda = \frac{q}{p}\frac{\bar{A}_f}{A_f} \tag{2.21}$$

The importance of λ and its connection to CP violation will be explained further in Section 2.2.2. However, the difference of the above rates divided by their sum forms a time-dependent CP observable which is sensitive to CP violation:

$$a_{f_{CP}} = \frac{(1-|\lambda|^2)}{(1+|\lambda|^2)}\cos(\Delta mt) - \frac{\mathcal{I}m\lambda}{(1+|\lambda|^2)}\sin(\Delta mt) \tag{2.22}$$

The above formulation illustrates that the time dependence of B decays to CP eigenstates depends on λ and provides an experimental means for measuring different types of CP violation. Before proceeding further, we now examine the three types of CP violation that can be observed in meson systems.

2.2.2   Three Types of CP Violation

The possible manifestations of CP violation can be classified as follows

- CP violation in decay (also called *direct CP* violation), which occurs in both charged and neutral decays, when the amplitude for a decay and its CP conjugate process have different magnitudes;

- CP violation in mixing, which occurs when the two neutral mass eigenstates cannot be chosen to be CP eigenstates

- CP violation in the *interference* between decays with and without mixing, which occurs in decays into final states that are common to B^0 and $\overline{B}^0$.

For any final state f, the ratio of the amplitude A_f to its CP conjugate amplitude $\bar{A}_{\bar{f}}$ is related to direct CP violation. There are two types of phases that may appear in A_f and $\bar{A}_{\bar{f}}$ that warrant discussion. Complex parameters in any

Lagrangian term that contributes to the amplitude will appear in complex conjugate form in the CP conjugate amplitude. Thus, phases that appear in A_f will also appear in $\bar{A}_{\bar{f}}$ with opposite signs; these are called *weak* phases. In the Standard Model they can only occur in the CKM matrix. A second type of phase can appear in decay amplitudes even when the Lagrangian is real. Such phases (designated *strong* phases) do not violate CP, since they appear in A_f and $\bar{A}_{\bar{f}}$ with the same sign. Further, only the relative phases of different terms in an amplitude have physical content; an overall phase rotation of an amplitude will have no physical consequences. It is therefore useful to write contributions to A in three parts: the magnitudes A_i, the weak-phase terms $e^{i\phi_i}$ and the strong-phase terms $e^{i\delta_i}$. If several amplitudes contribute to $B^0 \to f$ we can write:

$$A_f = \sum_i A_i e^{i(\delta_i + \phi_i)}, \quad \bar{A}_{\bar{f}} = \eta_f \sum_i A_i e^{i(\delta_i - \phi_i)} \tag{2.23}$$

where η_f is the CP eigenvalue. The phase convention-independent quantity is then

$$\left| \frac{\bar{A}_{\bar{f}}}{A_f} \right| = \left| \frac{\sum_i A_i e^{i(\delta_i - \phi_i)}}{\sum_i A_i e^{i(\delta_i + \phi_i)}} \right| \tag{2.24}$$

If CP is conserved, the weak phases are all equal. Therefore, if $|\bar{A}_{\bar{f}}/A_f| \neq 1$, then CP violation in *decay* is manifest. It is interesting to note that CP violation of this type will not occur unless at least two terms with different weak phases also have different strong phases. The asymmetry observable is given by:

$$|A|^2 - |\bar{A}|^2 = -2 \sum_{i,j} A_i A_j \sin(\phi_i - \phi_j) \sin(\delta_i - \delta_j) \tag{2.25}$$

This is, in fact, the only way to observe CP violation in charged meson decays.

The second type of CP violation is in the mixing between neutral mesons. Recall, from Eq. 2.15 that the quantity q/p is related to the off-diagonal elements of the mixing matrix.

$$\left| \frac{q}{p} \right|^2 = \left| \frac{m_{12}^* - \frac{1}{2} i \Gamma_{12}^*}{m_{12} - \frac{1}{2} i \Gamma_{12}} \right| \tag{2.26}$$

If CP is conserved, the mass eigenstates $|B_L\rangle$ and $|B_H\rangle$ must be the CP eigenstates. In this case, the relative phase between m_{12} and Γ_{12} vanishes. Therefore, if $|q/p| \neq 1$, then CP violation is manifest in the mixing transitions. For the neutral B system, this can be observed through asymmetries in semileptonic decays. [15]

Finally, the last type of CP violation is observed from the interference between decay with and without mixing. Here, we return to the case where neutral B mesons decay into the same final CP eigenstate, f. The relevant physically meaningful quantity is λ, as shown in Eqs. 2.20, 2.21. We assume there is no CP violation in decay or in mixing; therefore

$$A_f = Ae^{i(\delta + \phi_D)}, \bar{A}_f = \eta_{f_{CP}} Ae^{i(\delta - \phi_D)} \quad \Rightarrow \quad |A_f| = |\bar{A}_f|, \qquad (2.27)$$

$$q/p = e^{2i\phi_M} \quad \Rightarrow \quad |q/p| = 1. \qquad (2.28)$$

We allow for a strong phase δ but denote different decay and mixing weak phases, ϕ_D and ϕ_M. Hence we see that

$$\lambda = \frac{q}{p}\frac{\bar{A}_f}{A_f} = \eta_{f_{CP}} \, e^{2i(\phi_M - \phi_D)} \qquad (2.29)$$

If $|\lambda| \neq 1$, then CP violation is manifest through either mixing or decay, and if $\mathcal{I}m\lambda \neq 0$ then CP violation is manifest through the interference between decays with and without mixing. This is a consequence of the non-vanishing phase between q/p (from mixing) and $\bar{A}_f/A_f$ (from decay). It is clear, then, that the asymmetry shown in Eq. 2.22 provides a very powerful means of probing CP violation in B decays, via the parameter λ. We will see in the next section how λ is directly related to the CKM matrix elements in the Standard Model.

2.2.3 How λ_f Relates to the Unitarity Triangle

In Section 2.1.3 we saw that the unitarity relations can be represented in the complex plane as a triangle whose sides and angles are sensitive to specific B meson

decays. With this understanding of the Unitarity Triangle we can now relate the CP violation parameter λ to specific CKM elements. In general, the amplitudes for B decays often carry contributions from multiple Feynman diagrams, each of which contain different CKM elements. If all of the amplitudes that contribute to A and $\bar{A}$ could be calculated for a given decay, then the relationship of the Unitarity Triangle's angles (and CKM elements) to λ would be simple to identify. Unfortunately calculating amplitudes for hadronic B decays is very difficult (see Section 2.3.6). However, the decay $B^0 \rightarrow J/\psi\,K_S^0$, shown in Figure 2.3, is dominated
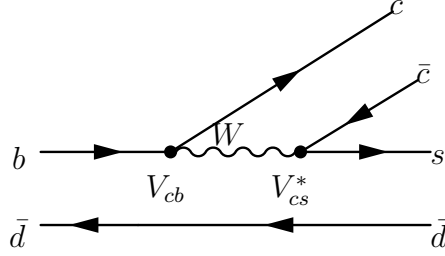
Figure 2.3: The $B^0 \rightarrow J/\psi\,K_S^0$ decay tree diagram.

by a color-suppressed tree diagram and is sensitive to only one weak phase. From Eq. 2.29, we find

$$
\begin{aligned}
\lambda(B^0 \rightarrow J/\psi\,K_S^0) &= -\left(\frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*}\right)\left(\frac{V_{cb} V_{cs}^*}{V_{cb}^* V_{cs}}\right)\left(\frac{V_{cs} V_{cd}^*}{V_{cs}^* V_{cd}}\right) \\
&\Rightarrow \; \mathcal{I}m(\lambda_{J/\psi K_S^0}) = \sin 2\beta
\end{aligned}
\tag{2.30}
$$

where β is one of the angles of the Unitarity Triangle. The first term above is q/p from $B - \bar{B}$ mixing (Figure 2.2), the second term comes from $\bar{A}/A$ and the last term comes from $K^0 - \overline{K}^0$ mixing. A measurement of the asymmetry in Eq. 2.22 with the mode $B^0 \rightarrow J/\psi\,K_S^0$ thus provides a direct measurement of the quantity $\sin 2\beta$. In fact, the BABAR and BELLE collaborations have measured the value of $\sin 2\beta$ precisely in this mode [6, 7] and have confirmed the Standard Model picture

of CP violation for that mode.

The measurements of the other angles and sides of the Unitary Triangle, however, are not complete. There are in fact many analyses devoted to the determination of the other two angles α and γ in addition to efforts to more precisely measure the sides of the Unitarity Triangle. The focus of current B physics experiments is to over-constrain the parameters of the Unitarity Triangle, to determine the consistency of the Standard Model's picture of CP violation and search for physics processes beyond the Standard Model description. Previous experiments have also put constraints on the CKM elements. Figure 2.4 shows the union of constraints from semileptonic branching fractions ratios, CP asymmetries from the K^0 system (ϵ_K), and B_d^0 mixing, as well as the constraint from recent time-dependent CP violating asymmetries in $B^0 \rightarrow J/\psi K_S^0$ at $BABAR$ and BELLE. The measurement of the CP asymmetry in $B^0 \rightarrow J/\psi K_S^0$ is clearly in agreement with expected values.

Nevertheless, other modes which also probe the angle β could possibly show deviations from the measurement of $\sin 2\beta$ in $B^0 \rightarrow J/\psi K_S^0$. The modes $B \rightarrow D^{(*)} \overline{D}^{(*)}$ are an example of these cases; a comparison of $\sin 2\beta$ measurements in these modes with that of $B^0 \rightarrow J/\psi K_S^0$ provides an independent test of the Standard Model picture of CP violation. Any differences could indicate that physics processes beyond the Standard Model are contributing to the known picture. The following section discusses the mode $B^0 \rightarrow D^{*+} D^{*-}$, its sensitivity to $\sin 2\beta$, and its usefulness for understanding other theoretical models.
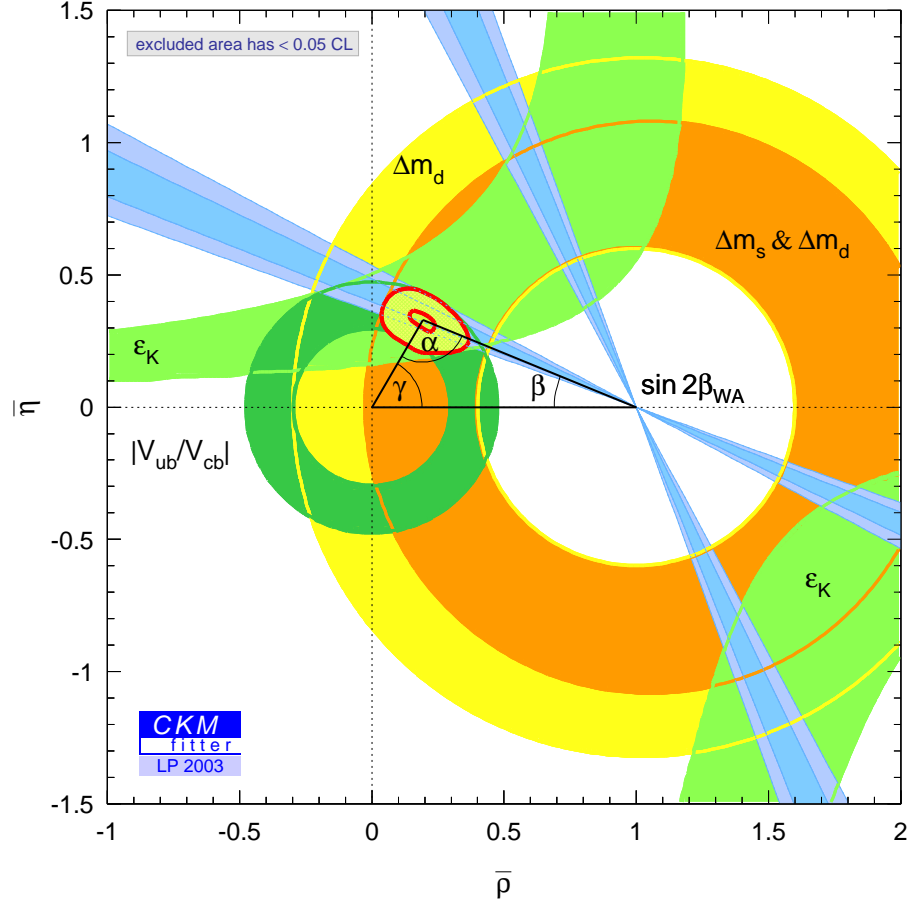
Figure 2.4: Experimental constraints on the Unitarity Triangle from measurements of semileptonic branching fractions ratios ($|V_{ub}/V_{cb}|$), CP asymmetries from the K^0 system (ϵ_K), and B_d^0 and B_s^0 mixing. The $\sin 2\beta$ measurement in $B^0 \to J/\psi K_S^0$ is overlaid to show the accuracy and agreement with other constraints. [23]

2.3 $B \to D^{(*)}\overline{D}^{(*)}$ Physics

2.3.1 Using $B^0 \to D^{(*)+}D^{(*)-}$ to extract $\sin 2\beta$

The modes $B \to D^{(*)}\overline{D}^{(*)}$ provide an alternative means to measuring $\sin 2\beta$ and independently test the validity of the Standard Model picture. However, these modes carry the added complication of multiple decay diagrams contributing to the amplitude. For $B^0 \to D^+ D^-$, for example, the tree diagram is expected to be dominant, but an additional "penguin" diagram is thought to contribute (Figure 2.6). If the tree diagram was truly the only contribution then,

$$\lambda(B^0 \to D^{(*)+}D^{(*)-})_{tree} \;=\; \left(\frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*}\right)\left(\frac{V_{cd}^* V_{cb}}{V_{cd} V_{cb}^*}\right)$$

$$\Rightarrow \;\; \mathcal{I}m(\lambda_{D^{(*)+}D^{(*)-}})_{tree} = -\sin 2\beta \qquad (2.31)$$

A measurement of $\sin 2\beta$ with $B^0 \to D^{(*)+}D^{(*)-}$ modes could then be compared
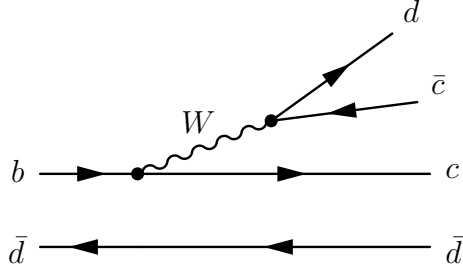
Figure 2.5: The primary ("tree") $B^0 \to D^{(*)+}D^{(*)-}$ decay diagram.
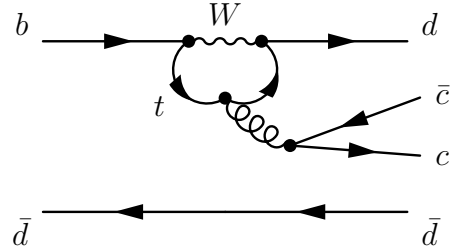
Figure 2.6: The color-suppressed penguin $B^0 \to D^{(*)+}D^{(*)-}$ diagram.

with the measurement in $B^0 \to J/\psi K_S^0$ as a test of the Standard Model, provided the penguin diagram(s) contribution is negligible. Conversely, the measurement of $\lambda(B^0 \to D^{(*)+}D^{(*)-})$ could be used to measure the size of the penguin contribution

in these modes, when the Standard Model picture is assumed to be true (no new physics).

A theoretical estimate of the possible size of the contribution from these penguin diagrams is determined in Ref. [29] (also see [35]). The complexity of calculating decay amplitudes for hadronic decays is discussed further in Section 2.3.6. However, using the technique of effective field theory and the factorization assumption, Ref. [29] predicts the penguin correction to $\sin 2\beta$ in $B^0 \to D^{*+}D^{*-}$ to be about 2%. The validity of the assuptions needed to calculate this number is often debated; hence, the experimental method for measuring $\lambda_{(B^0 \to D^{(*)+}D^{(*)-})}$ is designed to avoid any model-dependent assumption (see Chapter 7).

2.3.2 Penguin Contributions

For completeness, we include here the formulation that allows for a non-zero penguin contribution to the CP asymmetry measurement in $B^0 \to D^{(*)+}D^{(*)-}$. The amplitudes can be written as

$$A_{D^+D^-} = V_{cd}V_{cb}^* T + V_{td}V_{tb}^* P$$

$$\bar{A}_{D^+D^-} = V_{cd}^* V_{cb} T + V_{td}^* V_{tb} P \tag{2.32}$$

where T is the dominant amplitude (tree diagram), and P is the penguin-only amplitude. The CP violation parameter λ_f simplifies to

$$\lambda_f = \eta_f \frac{e^{-i\beta} - |R|e^{i\delta}}{e^{i\beta} - |R|e^{i\delta}} \tag{2.33}$$

where we define

$$R = z\,r \qquad z = \frac{V_{td}^* V_{tb}}{V_{cd}^* V_{cb}} \qquad r = \frac{P}{T} = |r|e^{i\delta} \tag{2.34}$$

z is the ratio of CKM matrix elements, r is the ratio of the penguin-only term P to the tree-dominated term T, and δ is the relative strong phase between the

amplitudes P and T. Since $|R|$ is dependent on the CKM ratio z, it is a function of the weak angles α, β and hence the small $|R|$ limit is not *a priori* justified. Now returning to the formulation in Eq. 2.22, and relating the coefficients of the sine and cosine terms to the Unitarity Triangle angles, we write

$$C_f = \frac{-2|R| \sin\beta \sin\delta}{1 + |R|^2 - 2|R| \cos\beta \cos\delta} \quad S_f = \eta_f \frac{\sin2\beta - 2|R| \sin\beta \cos\delta}{1 + |R|^2 - 2|R| \cos\beta \cos\delta} \quad (2.35)$$

where C_f is the cosine coefficient in Eq. 2.22 and S_f is the sine coefficient, both of which are experimentally measurable quantities. In the limit of small $|R|$,

$$C_f \approx -2|R| \sin\beta \sin\delta \qquad S_f \approx \eta_f \left(\sin2\beta + 2|R| \sin\beta \cos2\beta \cos\delta \right). \qquad (2.36)$$

Lastly, we include the quantitative theoretical estimates on the correction to a $\sin2\beta$ measurement using $B^0 \rightarrow D^{(*)+}D^{(*)-}$ modes. The CP parameter λ can also be written as

$$\lambda_{D^{(*)+}D^{(*)-}} = \eta_f \frac{|\bar{A}|}{|A|} e^{-2i(\beta + \Delta\beta)}$$
$$\Rightarrow \mathcal{I}m(\lambda_{D^{(*)+}D^{(*)-}}) = \sin[2(\beta + \Delta\beta)] \qquad (2.37)$$

where $\Delta\beta$ encapsulates the penguin diagrams' contribution to the measurement of the angle β. Thus when the asymmetry is measured and $\mathcal{I}m\lambda$ determined, the relationship to the angle β and the contributions from penguin diagrams is clearly formulated. Using the factorization assumption and the method of heavy-quark effective theory to calculate the matrix elements, an estimate of the size of $\Delta\beta$ in terms of the parameters ρ and η is found [15]:

$$(\Delta\beta)_{DD} \approx \mathrm{Arg}[1 - 0.088(1 - \rho - i\eta)]$$
$$(\Delta\beta)_{D^*D} \approx \mathrm{Arg}[1 - 0.010(1 - \rho - i\eta)]$$
$$(\Delta\beta)_{D^*D^*} \approx \mathrm{Arg}[1 - 0.029(1 - \rho - i\eta)]. \qquad (2.38)$$
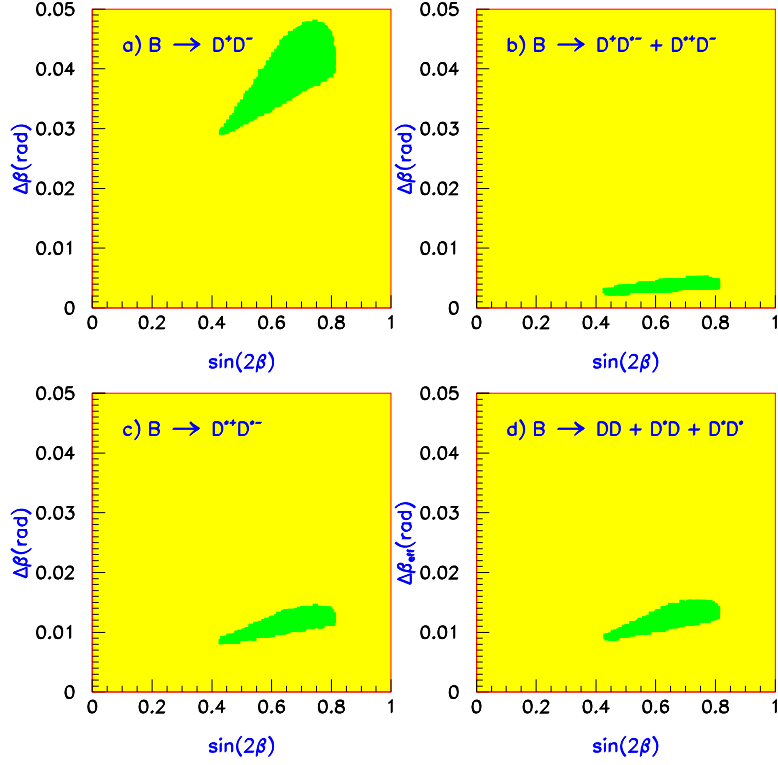
Figure 2.7: Uncertainty $\Delta\beta$ due to penguin diagrams in $B^0 \rightarrow D^{(*)+}D^{(*)-}$ decays, assuming factorization. Spread in values represents the theoretical uncertainties from heavy-quark-effective theory calculations.

Further, using the constraints of $B^0 - \overline{B}^0$ mixing, $|V_{ub}|$ and ϵ_K, the allowed domain of $\Delta\beta$ vs. $\sin 2\beta$ is plotted in Figure 2.7, under the factorization assumption.

2.3.3 Measuring γ via $B^0 \rightarrow D^{(*)+}D^{(*)-}$ Penguin Contributions

A further study of the possible penguin contributions has revealed the possibility of measuring the CKM angle γ using the $B^0 \rightarrow D^{(*)+}D^{(*)-}$ modes. Reference [33] demonstrates the sensitivity of $B^0 \rightarrow D^+D^-$ and $B^0 \rightarrow D^{*+}D^{*-}$ to γ via multiple penguin amplitudes and some theoretical input. While the formulation is not

model-independent and relies on both the $\sin 2\beta$ measurement in $B^0 \to J/\psi K_s^0$ as well as an SU(3) flavor symmetry argument, it may provide an interesting comparison of γ measurements in other modes in the future.

Figure 2.6 shows one possible penguin diagram with the t quark in the loop. This is expected to be the dominant penguin diagram. However, allowing for other quarks in the loop introduces additional penguin diagrams with different CKM elements contributing. We can write Eq. 2.32 more generally as:

$$
\begin{aligned}
A_{D^+D^-} &= (T + P_c)\, V_{cd}\, V_{cb}^* + P_t\, V_{td}\, V_{tb}^* + P_u\, V_{ud}\, V_{ub}^* \\
&= \mathcal{A}_{ct}\, e^{i\delta^{ct}} + \mathcal{A}_{ut}\, e^{i\gamma} e^{i\delta^{ut}}
\end{aligned}
\tag{2.39}
$$

where $\mathcal{A}_{ct} \equiv |(T + P_c - P_t)V_{cd}\, V_{cb}^*|$ and $\mathcal{A}_{ut} \equiv |(P_u - P_t)V_{ud}\, V_{ub}^*|$; the strong phases δ^{ct} and δ^{ut}, as well as the weak phase γ are explicitly written. We have defined here P_i which is the $b \to d$ penguin amplitude with an internal i-quark. The second line is obtained by using the unitarity of the CKM matrix, to eliminate the $V_{td}\, V_{tb}^*$ term. Also, the assumption that the exchange diagram is approximately zero (based on factorization) has also been made.

This formulation provides three physical observables sensitive to β, γ, and the strong phase difference $\delta = \delta^{ct} - \delta^{ut}$. However, even if one takes β from the $B^0 \to J/\psi K_s^0$ measurement, there is still one more theoretical parameter than there are observables. Using SU(3) flavor symmetry, Reference [33] argues that the appropriate information can be obtained with the Cabibbo-favored decay $B^0 \to D_s^+ D^-$ to reasonable accuracy.

The details of this analysis is left to Reference [33]; however, we can see that the $B^0 \to D^{(*)+}D^{(*)-}$ modes provide a means to obtain a measurement of γ in addition to β. Thus, we find that some models predict the penguin contributions in these modes to be small thereby allowing an independent measurement of $\sin 2\beta$;

31

other models use the presence of penguin diagrams as a means to determine γ. Experimentally, the prospect of measuring a deviation from sin2β in $B^0 \to J/\psi K_S^0$ is still somewhat far off statistically. Hence, measuring γ is even further away. However, it is clear that the modes $B \to D^{(*)}\overline{D}^{(*)}$ are capable of many interesting measurements for testing the Standard Model's picture of CP violation as well as uncovering the uncertainties of penguin diagram contributions.

2.3.4   Potential New Physics Effects

Sakharov showed in 1967 that one of the essential conditions for the observed baryon asymmetry in the universe is that CP violation must exist. [5] However, recent calculations show that the CP violation predicted by the Unitarity Triangle constraints is about 12 orders of magnitude below what is required to achieve the observed matter dominance in the universe. It is possible that the CP violation necessary to achieve consistency with observation lies at energies above what can be probed by $BABAR$; however, several well-motivated models, such as supersymmetry, predict extra CP violating phases that can be potentially observed through loop-mediated processes. Hence, it is important to test and constrain the Standard Model predictions for CP violation via tests sensitive to predictions from alternative models.

For example, the supersymmetric contribution to flavor-changing neutral current amplitudes and phases can be large. This will introduce significant differences in the CP asymmetries in $B^0 \to J/\psi K_S^0$ versus $B^0 \to D^{*+}D^{*-}$. A comparison of asymmetries in these modes provides a significant constraint on the parameters of such models.

Secondly, in Multi-Higgs-Doublet Models with n Higgs doublets, there are $2(n-1)$ charged and $2n-1$ neutral scalars that remain after spontaneous symmetry

breaking. These form an extended Higgs sector and can affect the couplings to the weak vector bosons. The dominant effects would be in penguin decays (of which $B^0 \to D^{(*)+}D^{(*)-}$ is potentially sensitive) where rates and CP asymmetries might be affected.

Hence, it has become an important test of the Standard Model to measure CP violation in $B^0 \to D^{(*)+}D^{(*)-}$. Not only will these measurements test the consistency of the Standard Model picture of CP violation, they could also provide a powerful probe into new physics processes not yet observed.

2.3.5 Specific $B^0 \to D^{*+}D^{*-}$ Measurements

Besides the possible penguin contributions in the CP asymmetry measurement, the $B^0 \to D^{*+}D^{*-}$ mode has the added complication that it is not a pure CP eigenstate. The D^* is a vector meson, meaning that it has an intrinsic spin of $S = 1$. Since the B is a pseudoscalar, three partial waves exist for the $D^{*+}D^{*-}$ final state: the $L = 0, 2$ or "S-" and "D-wave" states (which have CP eigenvalue of $+1$), and the $L = 1$ or "P-wave" state (which has $CP = -1$), where L is the relative angular momentum between the two vector particles. Thus, in general, the $B^0 \to D^{*+}D^{*-}$ decay is a mixture of CP-odd and CP-even eigenstates. The CP asymmetry is dependent on the sign of the CP final state, and hence the mixture will "dilute" the overall asymmetry. Equation 2.22 thus becomes

$$a_{f_{CP}} = C_f \cos(\Delta mt) - D\, S_f \sin(\Delta mt) \tag{2.40}$$

where D is given by

$$D = \frac{\Gamma_{CP=+} - \Gamma_{CP=-}}{\Gamma_{total}} = \frac{\Gamma_0 + \Gamma_\| - \Gamma_\perp}{\Gamma_{total}} = 1 - \frac{2\Gamma_\perp}{\Gamma_{total}} \tag{2.41}$$

and $\Gamma_0, \Gamma_\|, \Gamma_\perp$ are partial widths corresponding to an angular momentum decomposition of the decay topology, which is explained below. We can, therefore, perform

33

an angular analysis of the decay products in order to separate out the CP-even and CP-odd components, thereby compensating for the dilution in the CP asymmetry measurement.

For the $B^0 \rightarrow D^{*+}D^{*-} \rightarrow D^0\pi^+ \overline{D}{}^0\pi^-$ decay, it is beneficial to write the differential decay width in terms of the "transversity" basis [35]. This provides a slightly different physical representation of the decay products than the "helicity" basis. In the transversity frame, the following angles are defined: θ_1 is the polar angle of the π^- in the D^{*-} rest frame, θ_{tr} is the polar angle between the normal to the D^{*-} decay plane and the π^+ line of flight, and ϕ_{tr} is the remaining azimuthal angle of the π^+ (see Figure 2.8). The time-dependent amplitudes corresponding
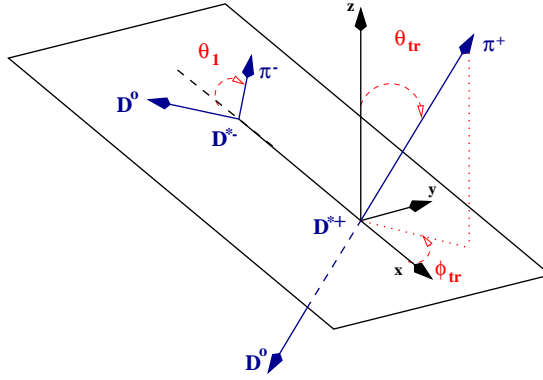
Figure 2.8: The transversity frame for $B^0 \rightarrow D^{*+}D^{*-} \rightarrow D^0\pi^+ \overline{D}{}^0\pi^-$. The decay of the D^{*-} is represented in the B^0 rest frame, while the decay products of the D^{*+} are shown in the D^{*+} rest frame. The x direction is defined by the direction of flight of the D^{*+} in the B^0 rest frame.

to the three transversity states are denoted $A_0, A_\parallel$, and $A_\perp$; the time dependent

angular distribution for $B^0 \to D^{*+}D^{*-}$ is given by:

$$\frac{1}{\Gamma}\frac{\mathrm{d}^4\Gamma}{\mathrm{d}\cos\theta_1 \mathrm{d}\cos\theta_{\mathrm{tr}}\mathrm{d}\phi_{\mathrm{tr}}\mathrm{dt}} = \frac{9}{32\pi}\frac{1}{|A_0|^2 + |A_\||^2 + |A_\perp|^2}$$

$$\{4|A_0|^2 \cos^2\theta_1 \sin^2\theta_{\mathrm{tr}} \cos^2\phi_{\mathrm{tr}}$$

$$+2|A_\||^2 \sin^2\theta_1 \sin^2\theta_{\mathrm{tr}} \sin^2\phi_{\mathrm{tr}}$$

$$+2|A_\perp|^2 \sin^2\theta_1 \cos^2\theta_{\mathrm{tr}}$$

$$+\sqrt{2}\mathcal{R}e(A_\|^* A_0) \sin 2\theta_1 \sin^2\theta_{\mathrm{tr}} \sin 2\phi_{\mathrm{tr}}$$

$$-\sqrt{2}\mathcal{I}m(A_0^* A_\perp) \sin 2\theta_1 \sin 2\theta_{\mathrm{tr}} \cos \phi_{\mathrm{tr}}$$

$$-2\mathcal{I}m(A_\|^* A_\perp) \sin^2\theta_1 \sin 2\theta_{\mathrm{tr}} \sin \phi_{\mathrm{tr}}\} \ . \quad (2.42)$$

For the $\overline{B}^0$ decay, $A_\perp \to -A_\perp$. Integrating the above expression over $\cos\theta_1$, ϕ_{tr} and t provides an interesting result:

$$\frac{1}{\Gamma}\frac{\mathrm{d}\Gamma}{\mathrm{d}\cos\theta_{\mathrm{tr}}} = \frac{3}{4}(1 - R_\perp)\sin^2\theta_{\mathrm{tr}} + \frac{3}{2}R_\perp \cos^2\theta_{\mathrm{tr}} \qquad (2.43)$$

The CP-odd fraction is defined as:

$$R_\perp = \frac{|A_\perp^0|^2}{|A_0^0|^2 + |A_\|^0|^2 + |A_\perp^0|^2} \qquad (2.44)$$

and can be extracted from a one-parameter, one-dimensional analysis of the decay products. We can now see that the magnitude $|A_\perp^0|^2$ corresponds to the CP-odd (P-wave) component of the final state. When $R_\perp$ is close to 1, the cosine term of Equation 2.43 dominates and the angle θ_{tr} is close to 0 or π – hence, the name $A_\perp$ for the transverse polarization perpendicular to the decay plane. When $R_\perp$ is close to 0, the S- and D- waves dominate, leaving a predominately parallel polarization of the decay products; thus, the combination $|A_0^0|^2 + |A_\|^0|^2$ describes the CP-even S- and D-wave component.

Ultimately, it is desirable to measure all of the magnitudes of the amplitudes and their relative phases. Currently, however, the number of reconstructed $B^0 \to D^{*+}D^{*-}$ events is too small to provide statistically meaningful results for all amplitudes, phases, and CP parameters. However, the measurement of $R_\perp$ is very feasible with a sample of $\sim$100 reconstructed events. Further, a measurement of the CP asymmetry parameter(s) λ_f is also possible, provided the final state favors one CP eigenstate over the other. As mentioned before, the existence of penguin diagrams could contribute to the CP asymmetry in $B^0 \to D^{*+}D^{*-}$. Theoretically, then, these diagrams could provide different penguin to tree ratios for each transversity state. To be completely general in the formulation, one should allow for different λ_f's for each of the transversity amplitudes $(\lambda_0, \lambda_\parallel, \lambda_\perp)$. A further discussion of this formulation and the specifics of the analyses performed is given in Chapters 6 and 7.

Besides the measurement of $R_\perp$ and the CP parameters λ_f, it is also of interest to determine the branching ratios for the $B \to D^{(*)}\overline{D}^{(*)}$ modes. While these measurements are interesting in and of themselves, they also provide an important test of possible final state interactions and the theoretical model of factorization. Effective theories are often used to calculate the branching ratios of these hadronic modes; however, these rely on specific models of the hadronization processes which are not easily determined. These methods are discussed briefly in Section 2.3.6.

Nevertheless, the factorization assumption has become an important cornerstone for many theoretical calculations. A few experimental tests exist to help determine the validity of this final-state model for the $B \to D^{(*)}\overline{D}^{(*)}$ decays. The three modes: $B^0 \to D^{*+}D^{*-}$, $B^+ \to D^{*0}D^{*+}$, $B^0 \to D^{*0}\overline{D}^{*0}$ are related through isospin relationships: specifically $B^+ \to D^{*0}D^{*+}$ is the $I = 1$ isospin transition, while the other two decays are the symmetric and antisymmetric combinations of

the $I = 1$ and $I = 0$ amplitudes. Theoretically this implies the relationship [29,35]:

$$|\langle D^{*+}D^{*-}|\mathcal{H}_{eff}|B^0\rangle|^2 + |\langle D^{*0}\overline{D}^{*0}|\mathcal{H}_{eff}|B^0\rangle|^2 = |\langle \overline{D}^{*0}D^{*+}|\mathcal{H}_{eff}|B^+\rangle|^2 \quad (2.45)$$

and further, the Branching Ratios also follow this relationship. The decay of $B^0 \rightarrow D^{*0}\overline{D}^{*0}$ is in fact predicted to be highly suppressed when final state interactions are negligible. Hence, if the $B^0 \rightarrow D^{*0}\overline{D}^{*0}$ branching ratio is too small to be observed (in comparison with $B^0 \rightarrow D^{*+}D^{*-}$) then factorization is presumed to be a good model for these transitions.

While the isospin relationship provides a convenient means of testing factorization, it does not provide any more information about the relative size of penguin diagrams compared to the tree diagrams. In other CP eigenstate modes like $B^0 \rightarrow \pi^+\pi^-$, an isospin analysis of the related modes (like the purely penguin mode $B^0 \rightarrow \pi^0 \pi^0$) provides an estimation of the penguin corrections to the CP asymmetry. Such is not the case with $B \rightarrow D^{(*)}\overline{D}^{(*)}$ modes. Nevertheless, an accurate determination of all $B \rightarrow D^{(*)}\overline{D}^{(*)}$ branching ratios remains a priority for the experimental validation of the theoretical calculations for these modes.

2.3.6 Overview of Theoretical Calculations

Because the quarks are bound by strong dynamics into color neutral hadrons, the perturbative basis for the Feynman calculus is no longer applicable at energy scales near or below Λ_{QCD}. For B meson decays, then, there exists a very complicated interplay between the weak and strong forces in the phenomenology of hadronic decays. As an example, Figure 2.9 shows a more realistic picture of a non-leptonic B decay. The complexity of strong-interaction effects, in fact, increases with the number of quarks appearing in the final state. Bound-state effects in leptonic decays can be lumped into a single parameter (the "decay constant"), while
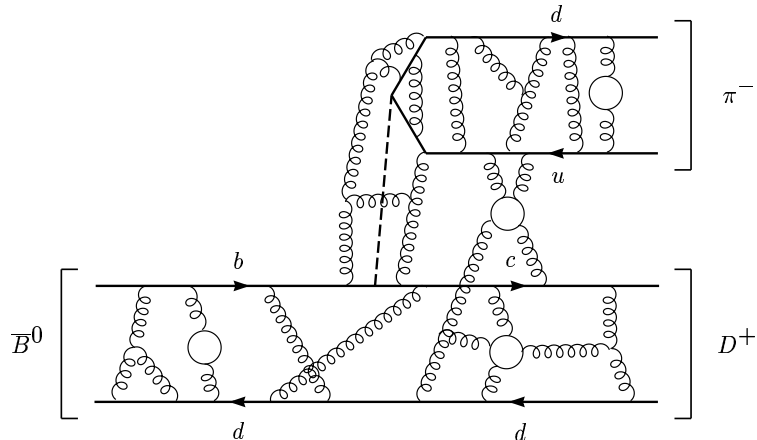
Figure 2.9: A more realistic representation of a non-leptonic B decay.

those in semi-leptonic decays are described by invariant form factors depending on the momentum transfer q^2 between the hadrons. Non-leptonic weak decays (like $B^0 \to D^{*+}D^{*-}$), however, are much more complicated and require the techniques of Heavy Quark Effective Theory (HQET), Chiral Perturbations Theory (ChPT), lattice gauge theory, QCD sum rules, and so on. [34] These theoretical techniques, while often model-dependent and intrinsically *ad hoc*, are able to over-come the some of the difficulties of non-perturbative QCD.

In HQET, for example, one exploits the fact that the heavy quark of the meson system will behave as an external source where the dominant color dynamics result from the light quarks interacting with this color source. [16] Within the limit $m_b, m_c \to \infty$, a "spin-flavor" symmetry yields a variety of predictions for heavy-hadron spectroscopy and decays. For ChPT, the limit is $m_u, m_d, m_s \to 0$; both methods use the perturbative technique to organize deviations from the limiting behavior through a systematic expansion about a small parameter. The technique of effective theories is comparable to Fermi's original theory of weak interactions which are approximated by point-like four-fermion couplings, governed by the cou-

pling constant G_F. In the same way, HQET "integrates out" the degrees of free-
dom of the heavy quark. Then, an effective Lagrangian can be obtained through
an infinite series expansion in powers of m_Q (called the Operator Product Expan-
sion) [15] and can handle the long distance physics at scales around Λ_{QCD}. To
separate short- and long-distance effects, theorists introduce a separation scale μ
such that $\Lambda_{QCD} \ll \mu \ll m_Q$. HQET is constructed in such a way that it is equiv-
alent to QCD in the long-distance region, i.e. for scales below μ. For more detail
on the subject of effective theories for B physics consult Refs. [34], [15], [16].

It is of interest to discuss some of the predictions specific to the decay of
$B^0 \rightarrow D^{*+}D^{*-}$, and give an example of the techniques mentioned above. As
a simple and related (same quark diagram) example is the use of an effective
Hamiltonian for $B^0 \rightarrow D^+D^{--}$ the simplest calculation can be obtained from the
operator expansion method. The tree diagram contribution to $B^0 \rightarrow D^+D^-$ can
be modeled by

$$H_{tree} = \frac{G_F}{\sqrt{2}} \left[V_{ub}V_{ud}^*(c_1 O_1 + c_2 O_2) \right] \tag{2.46}$$

where G_F is the effective Fermi weak coupling constant, and the operators are
defined as

$$O_1 = \bar{d}_j \gamma_\mu (1 - \gamma^5) c^j \bar{c}_i \gamma^\mu (1 - \gamma^5) b^i \quad , \quad O_2 = \bar{d}_j \gamma_\mu (1 - \gamma^5) c^i \bar{c}_i \gamma^\mu (1 - \gamma^5) b^j \tag{2.47}$$

where i, j are color indices. This formulation represents the idea that the interac-
tion with the W boson is strongly local so that the matrix element can be expanded
in powers of q^2/M_W (provided $q^2 \ll M_W$, which for B decays is satisfied). Thus
the "short-distance" effects are lumped into the Wilson coefficients c_k and are able
to be calculated using perturbative QCD. The "long-distance" effects are lumped
into the operators. [20]

The difficultly comes when attempting to calculate the actual hadronic matrix

element for $B^0 \to D^+D^-$. To do so the decay products are often assumed to be *factorizable* – that is, that the hadronization processes of the $\bar{c}d$ and $c\bar{d}$ quark pairs are to first order not contaminated by gluon exchange between them (i.e. the momenta of the quark pairs are large enough to avoid local color interactions). With the above effective interaction coupled with the assumption of factorization, the matrix element for $B^0 \to D^+D^-$ can be written:

$$\mathcal{M} = \frac{G_F}{\sqrt{2}}V_{cd}^*V_{cb}\left(\frac{c_1}{N_c} + c_2\right)\langle D^+|\,\bar{d}\gamma_\mu(1-\gamma^5)c\,|0\rangle\langle D^-|\,\bar{c}\gamma^\mu(1-\gamma^5)b\,|B^0\rangle \quad (2.48)$$

where N_c is the number of colors. The first current term above is simply given by $if_D q^\mu$ where f_D is the decay constant of the D meson and $q = p_B - p_D$. The second current term can be determined from q dependent form factors.

A common means of estimating the branching fraction for $B^0 \to D^+D^-$ is to use the related Cabibbo-favored decay $B^0 \to D_s^+D^-$. The above described factorization method leads to an estimate of $\mathcal{B}(B^0 \to D_s^+D^-) \approx 1.02 \pm 0.55\%$ [30]. If one assumes that the form factor term is approximately the same for $q^2 = m_{D^+}$ versus that for $q^2 = m_{D_s}$, then we can write

$$\frac{\mathcal{B}(B^0 \to D^+D^-)}{\mathcal{B}(B^0 \to D_s^+D^-)} \simeq \tan^2\theta_C \left[\frac{f_D}{f_{D_s}}\right]^2 \quad (2.49)$$

where θ_C is the Cabibbo angle. This relationship together with the theoretical calculation of $\mathcal{B}(B^0 \to D_s^+D^-)$ gives expected branching ratios of $\mathcal{B}(B^0 \to D^+D^-) \simeq 3 \times 10^{-4}$ and $\mathcal{B}(B^0 \to D^{*+}D^{*-}) \simeq 6 \times 10^{-4}$ [30]. Reference [30] also calculates the partial decay widths for the three transversity final states for $B^0 \to D^{*+}D^{*-}$; they estimate the *CP*-odd fraction ($R_\perp$) to be $\approx 6\%$. Given the recent measurements by *BABAR* [40], this theoretical calculation is quite impressive.

Hence, using the techniques of effective theories and factorization, decays such as $B^0 \to D^+D^-$ and the related decay $B^0 \to D^{*+}D^{*-}$ can in fact, be calculated to a reasonable degree of accuracy. Experimental measurements can then be used to

test the reasonableness of these assumptions as well as the possibility of measuring the effects of new physics not yet known.

Chapter 3

The *BABAR* Experiment

As detailed in the previous chapter, the study of B mesons is an exciting and necessary endeavor in order to further our understanding of CP violation as well as other interesting physics topics. In order to produce these exotic particles, and accurately detect and study them, large collaborations of scientists and engineers are formed to design, build, and maintain the accelerators and detectors necessary for such studies. The data studied in this dissertation was collected by the *BABAR* detector which records the e^+e^- collisions inside the PEP-II storage ring at the Stanford Linear Accelerator Center (SLAC).[1]

In 1987, Piermaria Oddone suggested that an asymmetric electron-positron collider would provide an excellent environment for studying CP violation in B decays at the $\Upsilon(4S)$ resonance [44]. Two laboratories ultimately accepted the challenge of developing a program for B meson CP violation studies using an asymmetric high luminosity e^+e^- collider design: KEK in Japan and the Stanford

---

[1]The work presented in this dissertation is representative of the cummulative efforts of the large (~ 600 people) *BABAR* collaboration. The specific D^*D^* analyses were developed by a smaller group ($\sim 5-7$ people) of collaborators. While the author participated in the development of both analyses, his primary contributions were to the studies and results presented in Chapters 4, 5, and 6.

Linear Accelerator Center (SLAC) in California. At SLAC, the Positron Electron Project's (PEP) ring was upgraded to support a high energy electron beam and a new lower energy positron ring was installed. Four and half years after the start of construction, PEP-II's first collisions took place on July 23, 1998. Meanwhile the *BABAR* detector, after a proposal in 1994 and a detailed Technical Design Report in 1995, was assembled in the PEP-II's interaction region 2 (IR-2) hall. *BABAR* has been recording the PEP-II collisions since May 1999.

In order to achieve their physics goals, the machine, detector, software, and personnel were required to operate as a factory. This means producing, recording, and analyzing the decays of tens of millions of B meson pairs a year in order to have the desired sensitivity to the physics the collaboration is interested in probing. In this chapter we'll examine the basics of PEP-II's and *BABAR*'s design, as well as the computing requirements of *BABAR*. The majority of the information in this chapter is derived from reference [45], which documents the *BABAR* hardware and software systems in detail.

3.1   PEP-II

The PEP-II B Factory is an asymmetric e^+e^- collider designed to operate at a luminosity of $3 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and above, at a center-of-mass energy of $10.58\,\mathrm{GeV}$, the mass of the $\Upsilon(4S)$ resonance. This resonance decays exclusively to $B^0\overline{B}^0$ and B^+B^- pairs and thus provides an ideal laboratory for the study of B mesons. We note that, if the $\Upsilon(4S)$ is produced at rest, the small B lifetime ($\tau_B \approx 1.5\mathrm{ps}$) and the small boost provided to the mesons (the $\Upsilon(4S)$ is only slightly more massive than the $B\bar{B}$ pair, so the B's center-of-mass momentum is about $340\,\mathrm{MeV}/c^2$) makes measurements of the separation of the two B meson decay points an experimental

impossibility given today's technologies.[2] Second, the decays of B mesons to CP eigenstates are very infrequent, having branching fractions that are typically less than 10^{-4}. These considerations necessitate that the collider designed for CP-violation studies in B mesons should be energy asymmetric, producing $\Upsilon(4S)$ with sufficient boost so that the decay points of the B mesons are distinguishable, and high luminosity, yielding significant numbers of B decays to CP eigenstates.

3.1.1   Design

SLAC's 2 mile linear accelerator feeds 9.0 GeV electrons into a high energy ring (HER) and 3.1 GeV positrons into a low energy ring (LER) that reside in the 2200 meter PEP-II circular tunnel (see Figure 3.1). These particles, kept in orbit by
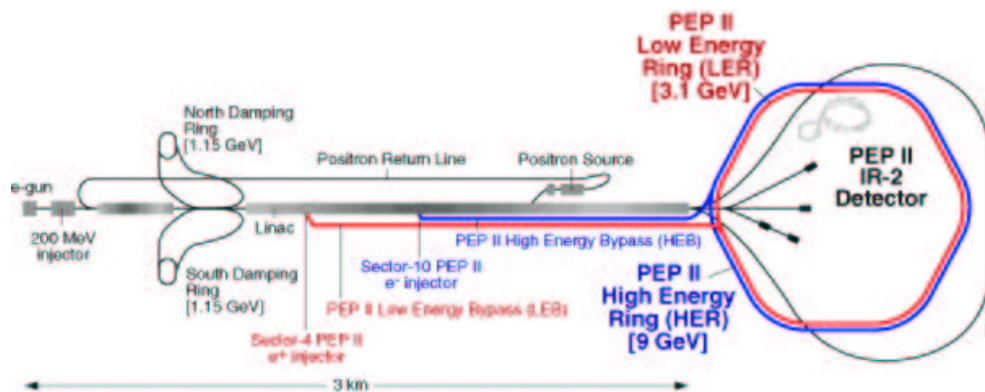
Figure 3.1: The PEP-II asymmetric storage ring and the SLAC linear accelerator. The SLAC linac is the injector for PEP-II. The point of collisions resides at the Interaction Region 2, where *BABAR* is located.

magnets and radio-frequency (RF) acceleration, are collided at an interaction region located at the center of the *BABAR* detector. Here the beams, while being tuned with a series of quadrapole magnets, are brought together and then sepa-

---

[2]Also see Chapter 7 for a thorough motivation of the asymmetric-energy configuration.

$e^+e^- \rightarrow$	Cross-section (nb)
$B\bar{B}$	1.10
$c\bar{c}$	1.30
$s\bar{s}$	0.35
$u\bar{u}$	1.39
$d\bar{d}$	0.35
$\tau^+\tau^-$	0.94
$\mu^+\mu^-$	1.16
e^+e^-	≈ 40

Table 3.1: Approximate production cross sections at PEP-II, including the experimental acceptance of *BABAR*. [15]

rated by a pair of dipole magnets. High luminosity is achieved by simultaneously maintaining several hundred bunches of electrons and positrons in each ring. The two ring design also aids in producing high luminosities by minimizing interactions between the beams.

With a center-of-mass energy at the peak of the $\Upsilon(4S)$ resonance, about 10.58 GeV, the e^+e^- system is Lorentz boosted in the electron direction with $\beta\gamma \approx 0.55$. The resulting B mesons travel an average of 260 microns along the electron beam before decaying. At this resonance, $e^+e^- \rightarrow B\bar{B}$ production accounts for nearly a quarter of the total hadronic cross-section (see table 3.1). The technical design and typical operating parameters of the PEP-II rings are listed in Table 3.2.

Parameters	Design	Typical
Energy HER/LER (GeV)	9.0/3.1	9.0/3.1
Current HER/LER (A)	0.75/2.15	0.7/1.3
Number of Bunches	1658	$553 - 829$
Bunch spacing (ns)	4.2	$6.3 - 10.5$
σ_x (μm)	110	120
σ_y (μm)	3.3	5.6
σ_z (mm)	9	9
Luminosity (10^{33}cm^{-2}s^{-1})	3	2.5
Luminosity (pb^{-1}/day)	135	120

Table 3.2: PEP-II design and typical luminosity operating parameters in the first year of colliding beams. HER and LER refer to the high energy e^- and low energy e^+ ring, respectively. σ_x, σ_y, and σ_z refer to the horizontal, vertical, and longitudinal RMS size of the luminous region. [45]
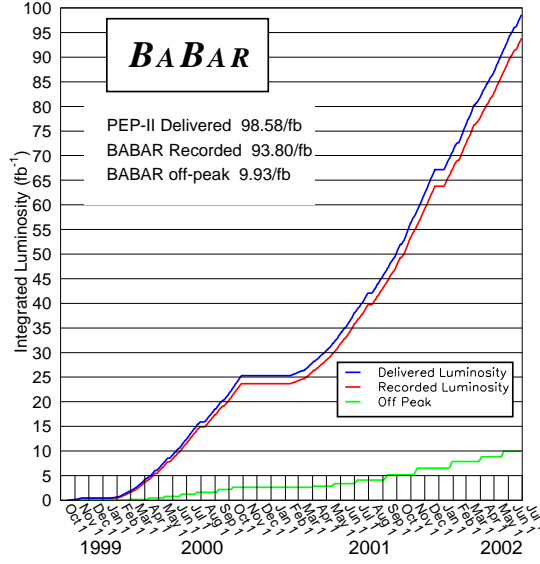
Figure 3.2: The integrated PEP-II luminosities delivered to and recorded by *BABAR*.

3.1.2 Operation and Performance

Though PEP-II operates at the $\Upsilon(4S)$ resonance, a small portion of the running time ($\approx 12\%$) is dedicated to data taking at a CM energy of 40 MeV below the resonance in order to aid in background studies for analyses sensitive to $e^+e^- \rightarrow u\bar{u}, d\bar{d}, s\bar{s}, c\bar{c}$ events, known as *continuum* background.

Figure 3.2 displays the accumulation of total integrated luminosity from October 1999 to June 2002. While the beam energies are calculated by PEP-II from the total magnetic bending strength and the average deviations of the RF accelerating frequencies from their mean values, the absolute luminosity, luminous region size, and beam position and angles are determined by the *BABAR* detector and online computing system. The two $B^0 \rightarrow D^{*+}D^{*-}$ analyses presented in this dissertation use a sample of data corresponding to 82 fb^{-1} (or about 88 million $B\bar{B}$ decays).

3.2   BABAR

The *BABAR* detector was designed to meet the stringent requirements of an ambitious physics program. Given the small branching fractions and high physics backgrounds of many interesting B decay modes and the necessity to determine the flavor of the second B in an event in time-dependent CP analyses (see Chapter 7), the detector had to possess the following properties: large and uniform acceptance, highly efficient and accurate charged and neutral particle reconstruction down to $60\,\mathrm{MeV}/c$, very good momentum (charged particles) and energy (neutrals) resolution and vertexing resolution, and powerful particle identification (especially for electrons and muons in order achieve flavor identification of the decaying B mesons). All of the above is needed while operating reliably for long periods in a possibly high radiation environment and succinctly processing and storing an extremely high volume of data.

It is also important to note the specific physics requirements of the decay $B^0 \to D^{*+}D^{*-}$. D^* mesons often decay to a neutral D meson and a charged π with low momentum (and is therefore called a *slow* pion). The tracking system must be able to reconstruct these slow pions efficiently down to momenta as low as possible. Because of the strong magnetic field, many of these pions never move beyond the inner radius of the drift chamber and thus require excellent reconstruction by the Silicon Vertex Tracker (SVT). The design of the SVT will be discussed in more detail in Section 3.3.

Figures 3.3 and 3.4 display the $y - z$ and $x - y$ cross-section schematics of the *BABAR* detector, where the electron beam direction is defined as the $+z$ direction and $+y$ points upward. Five sub-detectors, built as nested concentric cylinders about the interaction point, are supported by an array of electronics that con-
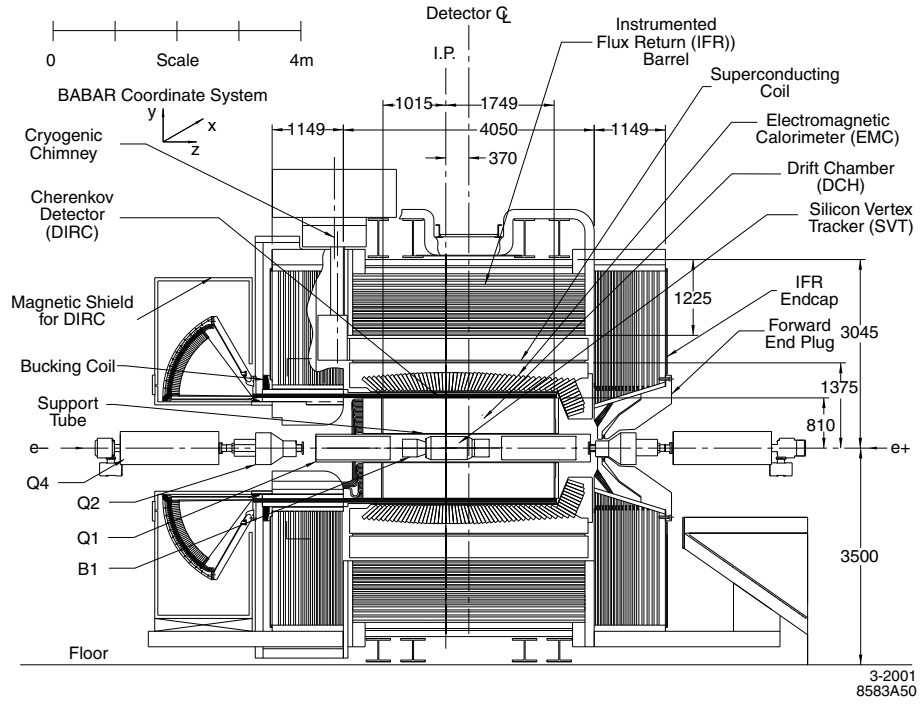
Figure 3.3: $y - z$ cross-section schematic of the *BABAR* Detector.
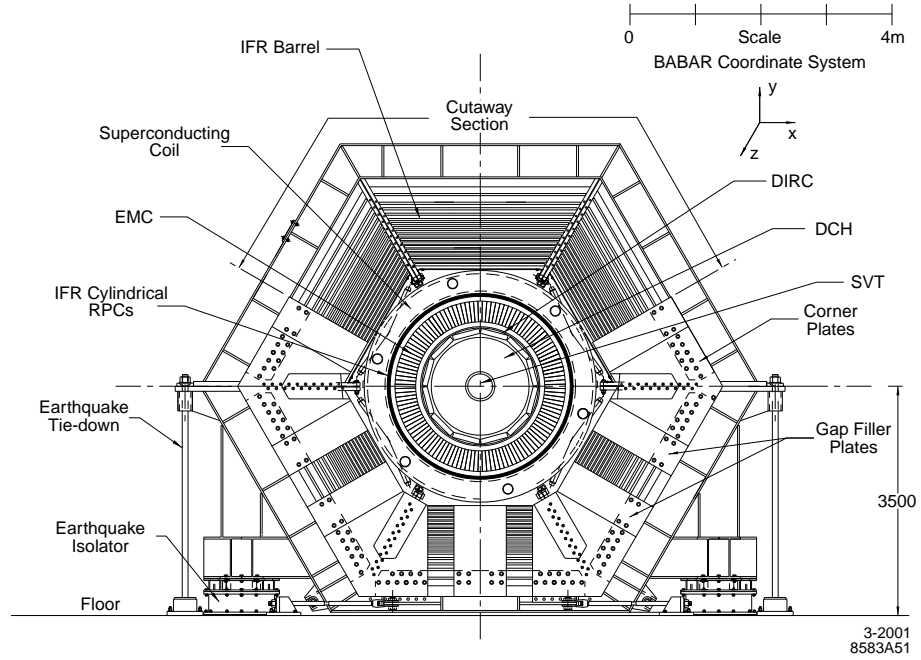
Figure 3.4: $x - y$ cross-section schematic of the *BABAR* Detector.

trol, readout, and monitor every component. Inside a 1.5 Tesla super-conducting solenoid sits the tracking system, composed of the Silicon Vertex Tracker (SVT) and the Drift Chamber (DCH); the Detector of Internally Reflected Cherenkov light (DIRC) for particle identification; and the Electromagnetic Calorimeter (EMC). Outside of the solenoid is the steel Instrumented Flux Return (IFR) which provides muon identification and neutral hadron detection.

3.2.1 Overview of Design Considerations

The detector was designed to be compact; the solenoid radius was chosen by balancing the physics requirements and performance of the drift chamber and calorimeter against the total detector cost. As in many similar detectors, the calorimeter was the most expensive single system and thus considerable effort was made to minimize its total volume without undue impact on the performance of either the tracking system or the calorimeter itself. Since the average momentum of charged particles produced in B-meson decays is less than $1\,\mathrm{GeV}/c$, the precision of the measured track parameters is heavily influenced by multiple Coulomb scattering. Similarly, the detection efficiency and energy resolution of low energy photons are severely impacted by material in front of the calorimeter. Thus, special care was taken to keep material in the active volume of the detector to a minimum.

An overview of the coverage and performance of each of the *BABAR* systems (subsystems) is shown in Table 3.3.

The asymmetry of the machine is also reflected in *BABAR*'s positioning with respect to the interaction point: an offset 37 cm in the direction of the LER maximizes the acceptance of the boosted system. In order to reduce perturbation by the tracking system solenoid, the detector axis is offset 20 mrad relative to the beam axis in the horizontal plane.

Table 3.3: Overview of the coverage and performance of the *BABAR* detector systems. The notation (C), (F), and (B) refers to the central barrel, forward and backward components of the system, respectively. The detector coverage in the laboratory frame is specified in terms of the polar angles θ_1 (forward) and θ_2 (backward). Performance numbers are quoted for $1\,\mathrm{GeV}/c$ particles, except where noted. [45]

System	θ_1 (θ_2)	No. Channels	No. Layers	Segmentation	Performance
SVT	20.1°	150K	5	50-100 μm $r-\phi$	$\sigma_{d_0} = 55\,\mu\mathrm{m}$
	(-29.8°)			100-200 μm z	$\sigma_{z_0} = 65\,\mu\mathrm{m}$
DCH	17.2°	7,104	40	6-8 mm	$\sigma_\phi = 1$ mrad
	(-27.4°)			drift distance	$\sigma_{tan\lambda} = 0.001$
					$\sigma p_T / p_T = 0.47\%$
					$\sigma(dE/dx) = 7.5\%$
DIRC	25.5°	10,752	1	$35 \times 17\,\mathrm{mm}^2$	$\sigma_{\theta_C} = 2.5\,\mathrm{mrad}$
	(-38.6°)			$(r\Delta\phi \times \Delta r)$	per track
				144 bars	
EMC(C)	27.1°	2×5760	1	$47 \times 47\,\mathrm{mm}^2$	$\sigma_E/E = 3.0\%$
	(-39.2°)			5760 crystals	$\sigma_\phi = 3.9\,\mathrm{mrad}$
EMC(F)	15.8°	2×820	1	820 crystals	$\sigma_\theta = 3.9\,\mathrm{mrad}$
	(27.1°)				
IFR(C)	47°	22K+2K	19+2	20-38 mm	90% $\mu^\pm$ eff.
	(-57°)				6-8% $\pi^\pm$ mis-id
IFR(F)	20°	14.5K	18	28-38 mm	(loose selection,
	(47°)				1.5–3.0 GeV/c)
IFR(B)	-57°	14.5K	18	28-38 mm	

The high luminosity environment provided by PEP-II produces various sources of large backgrounds and necessitates radiation-hard detectors and electronics for *BABAR* to protect against damage, aging, and bandwidth limitations from extraneous signals. The majority of the several kW of synchrotron radiation emanating from the quadrapole and dipole magnets is diverted by the design of the beam orbits, the vacuum-pipe apertures, and the synchrotron-radiation masks. Beam-gas backgrounds, caused by bremsstrahlung and Coulomb scattering off of residual gas molecules, are enhanced by vacuum breeches. Thus this sample of energy-degraded particles that reach the interaction region are typically bent by the dipoles and bombard detector components in the horizontal plane. Similarly, interactions of energy degraded electrons or positrons from radiative Bhabha scattering with various apertures produce backgrounds that scale with the instantaneous luminosity. Comparisons of data taken with single HER and LER beam with colliding beam runs show that machine backgrounds constitute a significant portion of all triggers.

The *BABAR* detector was designed to record PEP-II's full luminosity for 10 years. In addition to radiation-hard components, a radiation protection system safe-guards the detector by determining radiation doses from diode leakage currents near the SVT and signals from PIN diodes mounted on small CsI(Tl) crystals in the DCH and EMC. At the same time DCH and IFR high voltage and DIRC and IFR counting rates are monitored to ensure sub-detector safety. *BABAR*-initiated beam aborts are generally induced by instantaneous sources of radiation from dust particles trapped in the beam and non-Gaussian tails from beam-beam interactions.
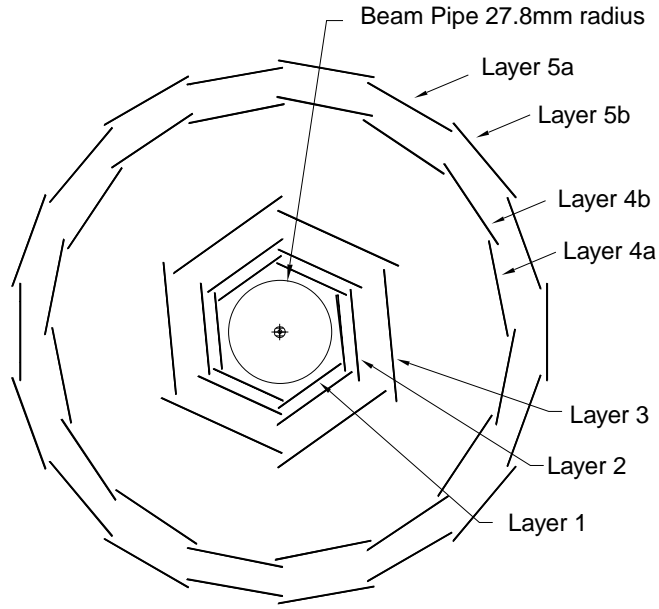
Figure 3.5: $x - y$ cross-section schematic of the *BABAR* SVT Detector.

3.3 The Silicon Vertex Tracker (SVT)

The trajectories of charged particles are determined primarily from the tracking capabilities of the SVT and the Drift Chamber (DCH). The 1.5 T magnetic field of the super-conducting solenoid bends the paths of charged particles thus enabling a measurement of their momenta. The field, produced by currents in two layers of 3060 mm diameter, 3513 mm long niobium-titanium coils, is very uniform in the tracking volume, deviating at most by 2.5% in the direction transverse to the path of high momentum tracks.

The *BABAR* SVT (pictured in figures 3.5 and 3.6) provides the precise vertexing as well as low momentum charge particle reconstruction needed for the *BABAR* physics program. The SVT provides stand-alone tracking capability with accurate position and impact parameter determination. These precision measurements allow for the exclusive reconstruction of *B*- and *D*-meson decays with high resolution
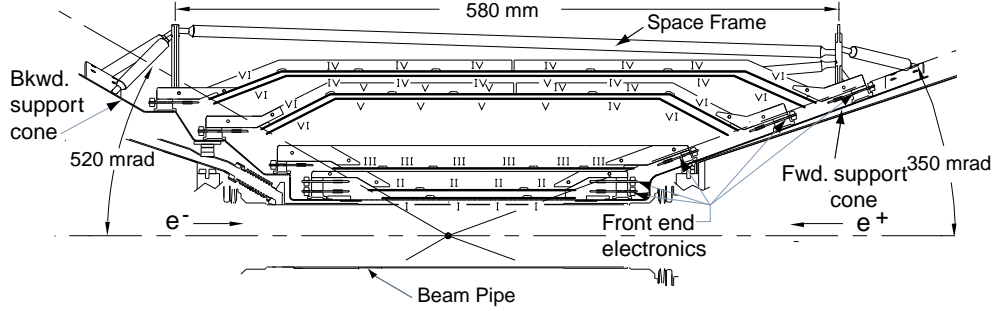
Figure 3.6: $y - z$ cross-section schematic of the *BABAR* SVT Detector.

and thus minimal background. Also, to avoid significant impact of the resolution on the *CP* asymmetry measurements the mean B vertex resolution along the z-axis must be better than $80 \mu m$. Similarly, the SVT must provide vertex resolution on the order $100 \mu m$ in the x-y plane in order to reconstruct decays like $B^0 \to D^+ D^-$.

Mounted on the beam pipe, this detector's 0.96 m^2 of active area is composed of five layers of 340 double-sided silicon wafers mounted on a carbon-fiber frame. On each side of every wafer, strip sensors running orthogonal to ones on the opposite side detect the passage of a charged particle. Thus, the strips running parallel to the beam measure $\phi \equiv \tan^{-1}(x/y)$, while the transversely oriented strips measure z. The wafers are organized in half modules which read out at the two ends of the detector by fanout circuits to custom time-over-threshold (ToT) chips in a total of $150,000$ channels. The inner three layers are placed close to the beam pipe, dominating the determination of track position and angles, while the outer two layers, which are arch-shaped to minimize the silicon use, are placed close to the DCH in order to aid in pattern recognition with the DCH and low p_T tracking. Tilting of modules in the inner layers and dividing the outer layers into two sub-

54

layers produces overlap regions which avoid gaps in the acceptance.

The silicon wafers are composed of n-type substrate with p^+ and n^+ strips on opposite sides of the wafer. The typical depletion (bias) voltages are in the range 25-35V, while operating voltages are about 10V higher. If the deposited charge from a particle is more than 0.95fC, corresponding to 0.25MIP, then this "hit" is processed by the ToT chips, provided a Level 1 (L1) trigger is received for the event. The output information, including the strip address is sent to the VME-based processors called readout modules (ROMs) for further track reconstruction. Event processing will be discussed more later.

The alignment of the silicon wafers is essential for vertexing accuracy and track reconstruction, and is performed in two steps. The first step consists of determining the relative positions of the 340 wafers (local alignment). Once done, the SVT as a whole must be aligned with respect to the global coordinate system defined by the DCH (global alignment). The local alignment procedure is performed using tracks from $e^+e^- \rightarrow \mu^+\mu^-$ events and cosmic rays. The $\mu^+\mu^-$ tracks are simultaneously fit using a Kalman filter technique [47, 48] and the known beam momentum as a constraint. The use of tracks from cosmic rays reduces any systematic distortion that may be introduced due to imprecise knowledge of the beam momenta. Information from the optical survey performed during the assembly of the SVT is also included in the alignment procedure in order to constrain wafers relative to other wafers in the same module.

Using the hit-track residuals from the above described set of tracks and optical survey information, a χ^2 is formed for all the tracks in each wafer. Each χ^2 is minimized with respect to each wafer's six local parameters while holding the parameters of all other wafers fixed. This minimization procedure is iterated because the correlations between the wafers are neglected in a single minimization.
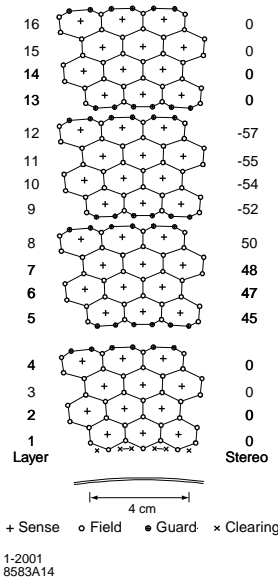
Figure 3.7: Schematic of the 4 inner layers of the *BABAR* DCH.

Because of possible global movement of the SVT, the six global alignment parameters are determined by minimizing the difference between track parameters obtained with the SVT-only and the DCH-only fits. This procedure is performed once per run (about every 2-3 hours) and the parameters are then used in the subsequent run. This procedure, known as *rolling calibration*, ensures that track reconstruction is always performed with up-to-date global alignment constants.

The SVT has performed according to design essentially since its inception. The combined hardware and software hit-finding efficiency is about 97%, excluding the few defective read-out-sections. Single hit resolution for tracks originating from the IP averages $20\mu m$ in both z and ϕ for hits on the inner 3 layers and $40\mu m$ in z and 20 in ϕ for hits in the outer two layers.

3.4 The Drift Chamber (DCH)

The DCH's reconstruction of track trajectories dominates *BABAR*'s measurement of charge particle momenta. It is 3m long and 81cm in diameter and has 7104 hexagonal cells; each cell consists of one grounded tungsten-rhenium sense wire surrounded by six aluminum field wires held at $> +1900$ V.[3] The cells are grouped into 40 layers, which are in turn grouped into 10 superlayers, with the wires in each superlayer oriented as either axial (directly parallel to the z-axis) or "stereo" (at a small angle in ϕ with respect to the z-axis, in order to obtain longitudinal position information). Each cell is able to detect traversing charge particles' ionization of a 80:20 mixture of helium:isobutane gas which is held at 4 mbar above atmospheric pressure. This choice of wire and gas minimizes multiple Coulomb scattering, presenting less than 0.2% of the radiation length (X_0) to tracks.

The readout electronics, mounted on the backward end-plate in order to minimize the material in front of the forward calorimeter end-cap, measure the drift time from ionization and the integrated charge, and provide a single bit to the L1 trigger. The time-to-distance relationship within the cells, as well as calibrations for determining the deposited charge are performed offline. The time-to-distance relationship is determined from two prong events (e^+e^- and $\mu^+\mu^-$ events). The total deposited charge in each cell is used to determine the specific energy loss dE/dx, which in turn is used for particle identification (see Figure 3.8). The dE/dx resolution for Bhabha events is typically 7.5%.

---

[3]The *BABAR* data sample has been collected with the DCH at three different high voltages. Originally the wires were held at 1960 V. After collecting $\approx 10/fb$ of on-resonance data, the voltage was changed to 1900 V for another $\approx 10/fb$ in an effort to extend the lifetime of the chamber. Eventually 1930 V was chosen as optimal for both the chamber's longevity and detection efficiency.
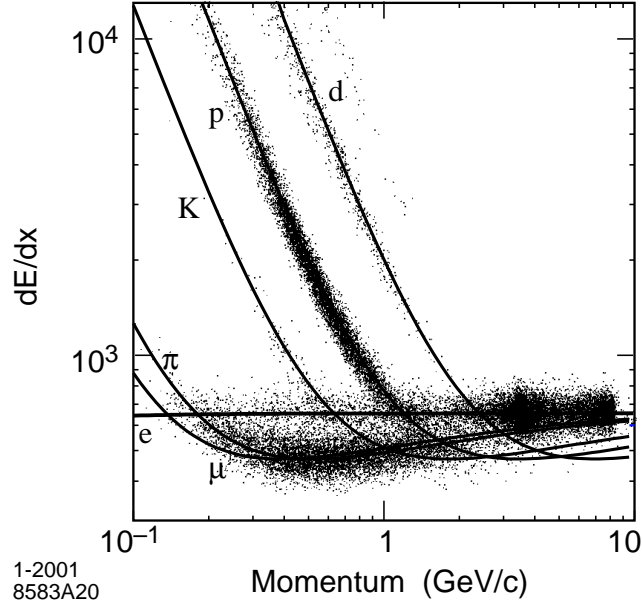
Figure 3.8: DCH particle identification as a function of track momentum using dE/dx.

3.5 Track Reconstruction

Most of the B decay modes are very dependent on precise determination of track parameters. The mode $B^0 \rightarrow D^{*+}D^{*-}$ (and subsequent decays), for example, has on average 7-8 charged tracks and requires precise determination of mass and energy in order to separate these events from combinatoric background; thus, it is very dependent on charged particle tracking.

The reconstruction of charged particle tracks relies on the data obtained from the SVT and DCH. Charged tracks are defined by five parameters $(d_0, \phi_0, \omega, z_0, \tan \lambda)$ and their associated error matrix. These parameters are measured at the point of closest approach to the z-axis; d_0 and z_0 are the distances of this point from the origin in the x-y plane and along the z-axis, respectively. The angle ϕ_0 is the azimuth of the track, λ the dip angle relative to the transverse plane, and $\omega = 1/p_t$

58

is its curvature.

Track reconstruction is done offline and builds tracks from information initially provided by the L3 trigger. L3 provides an estimate of the time at which the e^+e^- interaction occurred, called t_0, as well as four-hit track segments in the DCH superlayers. Tracks are selected by performing a helix fit to these track segments, and a search for additional hits in the DCH that may belong to these tracks is performed; t_0 is improved by using only hits associated with tracks. Next, the L3 DCH tracks are refitted using the more precise time-to-distance calibration, while two other sophisticated pattern-finding algorithms are used to find tracks that do not pass through the entire DCH or do not originate from the IP. At the end of this procedure, tracks are again fit using a Kalman filter method [47,48] that takes into account the detailed distribution of material in the detector and the full map of the magnetic field.

The resulting tracks are extrapolated into the SVT, and SVT track segments are added, provided they are consistent with the expected error through the intervening material and inhomogeneous magnetic field. Among the possible SVT segments, those with the smallest residuals and the largest number of SVT layers are retained and a Kalman fit is performed to the full set of DCH and SVT hits. Any remaining SVT hits are passed to two complementary standalone track finding algorithms. The first of these forms "space-points" from combinations of ϕ and z-hits from opposite sides of a wafer in layers 1, 3 and 5, and then adding consistent space points from the other layers. This algorithm is efficient over a wide range of d_0 and z_0 values. The second algorithm starts with circle trajectories from ϕ hits and then adds z hits to form helices. This algorithm is less sensitive to large combinatorics and to the missing z information in the few defective SVT modules. These two standalone SVT tracking algorithms have a high efficiency for tracks

with low transverse momentum. This feature is important for the detection of the soft pion in D^* decays. In fact, the SVT significantly extends the capability of the charged particle detection down to transverse momenta as low as $50\,\mathrm{MeV}/c$.

3.6 The DIRC and Particle Identification

The *BABAR* physics program has stringent requirements for $\pi - K$ separation over a large momentum range. At low momenta, primarily $< 1\,\mathrm{GeV}/c$, flavor tagging using kaons from cascade decays is an efficient way of determining B flavor. At the high end of the range, reconstructing $B^0 \to \pi^+\pi^-$ and $B^0 \to K^\pm\pi^\mp$ requires separation at momenta up to 4.2 GeV in the lab frame. At intermediate energies, reducing background in charm decays such as $D^0 \to K\pi$ is necessary for $B \to D^{(*)}\overline{D}^{(*)}$ reconstruction.

3.6.1 DIRC

The DIRC (Detector of Internally Reflected Cerenkov light) principle uses internal reflection within quartz bars to propagate Cerenkov light to readout phototubes while preserving the Cerenkov angle. A schematic of the DIRC principle is shown in Figure 3.9. The propagation of photons and the preservation of θ_c requires extremely flat surfaces in order to avoid dispersion of the reflected angles. Fused, synthetic silica quartz is used due to the excellent optical surface it allows through polishing, as well as other favorable properties such as long attenuation length, low chromatic dispersion, small radiation length, and radiation hardness. As shown in figure 3.9, the light is internally reflected to a wedge to reflect photons into a water-filled *standoff box* (SOB). The standoff box is enclosed by an array of 10752 photomultiplier tubes, which are each 29 mm in diameter. The Cerenkov
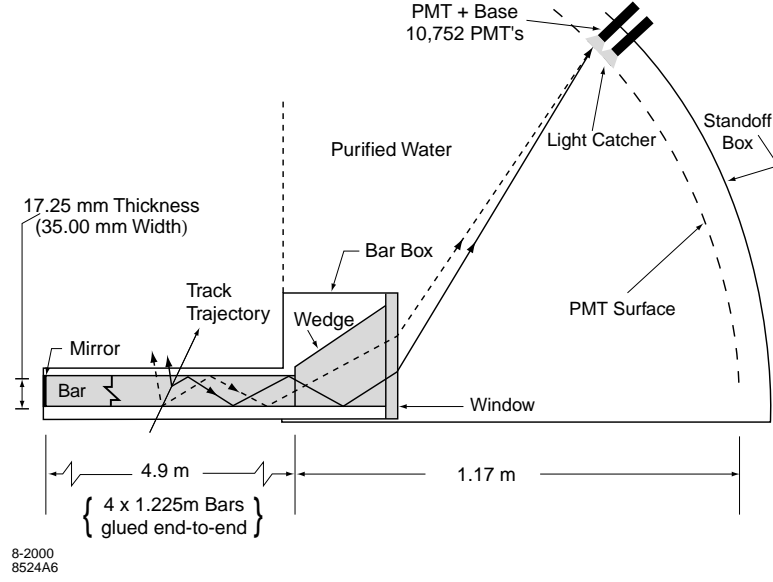
Figure 3.9: Schematic drawing illustrating the detection of Cherenkov photons by *BABAR*'s DIRC.

light from a particle passing through the DIRC forms a ring (essentially a conic section) imaged on the phototubes. The opening angle of this conic section contains information on particle type via the typical relation $\cos\theta_c = 1/n\beta$, where $\beta = v/c$, and n is the mean index of refraction ($= 1.473$ for fused silica).

Data from the phototubes is read out to front-end electronics, which performs the amplification, digitization, and buffering. Reduction of data from out-of-time or noisy PMTs is performed in in the external electronics and reduces the data volume by 50% using rough timing cuts. Online calibration of PMT efficiency, timing response, and electronics delays is provided by a light pulser system which generates precise 1 ns flashes from blue LEDs inside the SOB.

The emission angle and the arrival time of the Cherenkov photons are reconstructed from the observed space-time coordinates of the PMT signals and then transformed into the Cherenkov coordinates (θ_c, ϕ_c, and δt) via a maximum like-

lihood fit. Both efficiency and the timing of the electronics are critical for DIRC performance. Timing accuracy is necessary for background hit rejection, and more importantly, for the exclusion of other tracks in the same event as the source of the photon.

The Cerenkov angle resolution for dimuon events is 2.5 mrad, close to the design goal of 2.2 mrad. This results in $\pi - K$ separation at $3\,\mathrm{GeV}/c$ of 4.2σ. The mean kaon selection efficiency and pion misidentification for a "loose" selection are 96.2% and 2.1% respectively.

3.6.2   Other Particle ID

Measurements of energy loss (dE/dx) by the tracking system also allow charged particle identification at low momenta and allow separation of < 700 MeV pions and kaons. In the SVT, ≈ 10 ToT measurements are converted to dE/dx using a lookup table and a 60% truncated mean is calculated. In the DCH, an 80% truncated mean of the ≈ 40 ionization loss measurements for each track provide a 7.5% dE/dx determination (see figure 3.8).

3.7   The Electromagnetic Calorimeter (EMC)

The primary tasks of *BABAR*'s EMC are the detection of photons, reconstruction of π^0 and η decays, and the identification of electrons for B flavor determination. The requirements for the EMC include energy resolution on the order of $1-2\%$ for π^0's and be able to operate within the 1.5T field of the solenoid for an anticipated ten-year lifetime of the experiment.

The EMC records the energy of the electromagnetic showers from photons and electrons in a finely segmented array of thallium-doped cesium iodide (CsI(Tl))
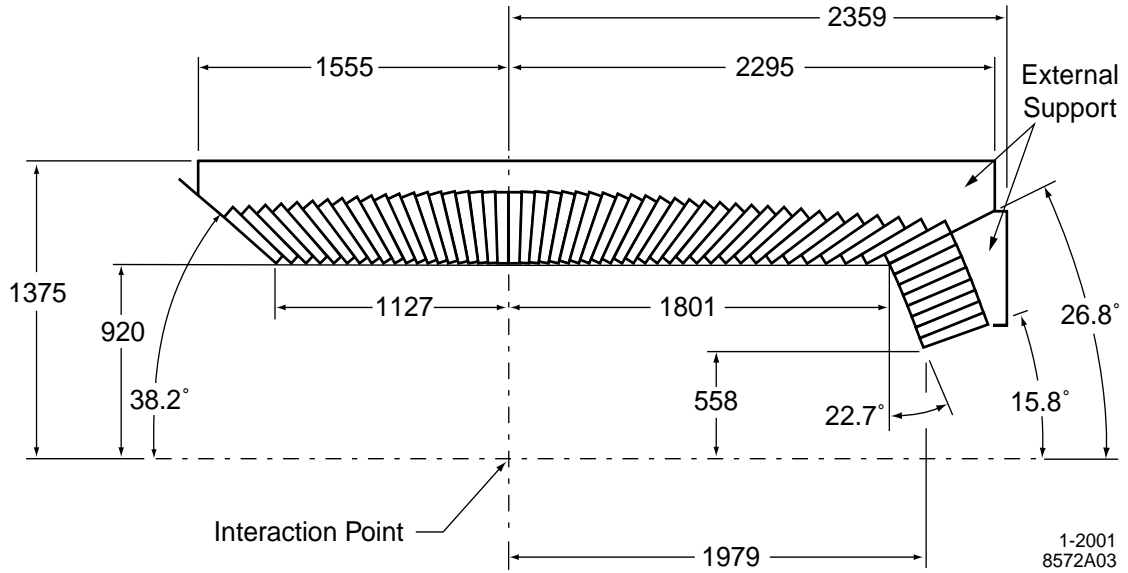
Figure 3.10: $y-z$ Schematic drawing of the top half of *BABAR*'s EMC. The detector is axially symmetric around the z-axis. All dimensions are in mm.

crystals (properties listed in table 3.4). The crystals, with radiation lengths between 16.0 and 17.5 X_0, are each instrumented with a pair of silicon photodiodes. Each diode is connected to a low-noise preamplifier which shapes and amplifies the signal by a factor between 1 and 32. The EMC, in fact, does not buffer the data on front-end electronics; rather it outputs the full digital data stream to the read-out modules in external electronics, which perform, on receipt of a Level 1 trigger, a fit to the digitally filtered data stream to derive energy and time measurements.

The EMC, divided into two sections (a cylindrical barrel and a conical forward end-cap) which externally support each of the 6,580 crystals (see figure 3.10), determines the energy and direction of 90% of the photons emitted in the center of mass system. A typical electromagnetic shower spreads over many adjacent crystals, forming a *cluster* of energy deposits. Pattern recognition algorithms are used to efficiently identify these clusters and to differentiate single clusters with

Parameter	Value
Radiation Length	1.85 cm
Moliere Radius	3.8 cm
Density	4.53 g/cm^3
Light Yield	50,000 γ/MeV

Table 3.4: Properties of CsI(Tl).

one energy maximum from merged clusters with multiple energy maxima.

To determine the energy resolution of the EMC, two methods are used. At low energy a neutron source which produces a mono-energetic 6.13 MeV calibration signal and a xenon flash light pulser system is used and yields $\sigma_E/E = 5.0 \pm 0.8\%$. At high energy, the resolution is derived from Bhabha scattering, where the energy of the detected shower can be predicted from the polar angle of the electron; at 7.5 GeV, $\sigma_E/E = 1.9 \pm 0.07\%$. The reconstructed mass of π^0's is found to be stable to better than 1% over the full photon energy range and the width agrees well with Monte Carlo simulations.

3.8   The IFR and Muon Identification

The Instrumented Flux Return (IFR) is made of 806 Resistive Plate Chambers (RPC) placed inside the steel of the magnet flux return. This detector enables *BABAR* to discriminate muons from hadrons and to detect K_L^0s and other neutral particles. Muon identification is important for determining the flavor of neutral B mesons via semileptonic decays, for the reconstruction of vector mesons like J/ψ, and for the study of semi-leptonic and rare decays.

The RPCs are arranged in 19 barrel and 18 end door layers and separated

64

by steel of increasing thickness from 2 cm to 10 cm. In order to detect particles exiting the EMC, two additional layers with four readout plates are placed inside the magnet cryostat. Planar RPCs contain a 2 mm Bakelite gap with ~ 8 kV across it. Ionizing particles which cross the gap create streamers of ions and electrons in the gas mixture (which is typically 56.7% Argon, 38.8% Freon, and 4.5% isobutane) which in turn creates signals via capacitive coupling on the "x-strips" and "y-strips" on opposite sides of the RPC. The RPC strip segmentation provide measurements of track ϕ/z in the barrel and x/y in the end-cap.

Muon identification uses variables such as number of expected vs. actual interaction lengths transversed and the χ^2 match to the charged track. Muon identification performance may be seen in figure 3.11. K_L^0 efficiency roughly increases linearly with momentum and varies between 20% and 40% between 1 and 4 GeV.

3.9   The Online System

BABAR's electronics, trigger, data acquisition (DAQ), and online computing systems are composed of tightly coupled hardware and software. A brief overview of the trigger system and online monitoring is presented here; the reader is encouraged to consult Reference [45] for more information.

3.9.1   Trigger

The basic requirement for the trigger system is the selection of events of interest (see Table 3.5) with a high, stable, and well-understood efficiency while rejecting background events and keeping the total event rate under 120 Hz in order to satisfy offline computing limitations. Beam-induced background rates are typically about 20 kHz. The trigger is responsible for reducing this rate while accepting over 99%
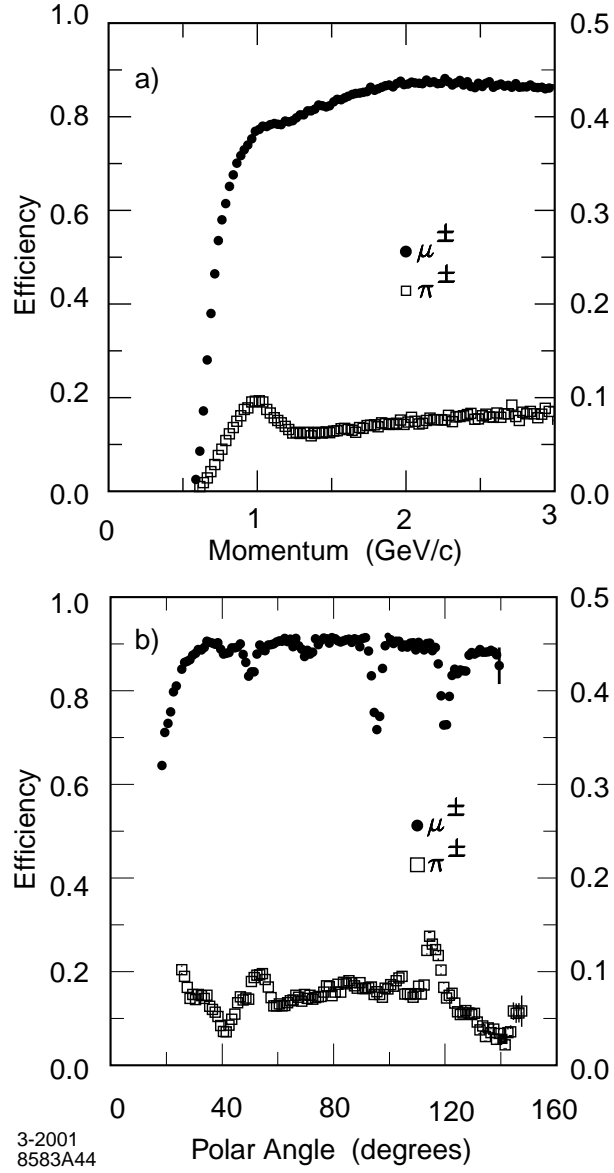
Figure 3.11: Muon efficiency (left scale) and pion misidentification probability (right scale) as a function of a) the laboratory track momentum, and b) the polar angle (for $1.5 < p < 3.0\,\mathrm{GeV}/c$ momentum), obtained with loose selection criteria.

66

of B events, over 95% of hadronic continuum, and over 90% of $\tau^+\tau^-$ events. The *BABAR* trigger is implemented in two levels: a Level 1 hardware trigger (L1), and a Level 3 software trigger (L3). An additional Level 2 trigger was not necessary under the current background and luminosity conditions.

Event type	Cross section (nb)	Production Rate (Hz)	Level 1 Trigger Rate (Hz)
$B\bar{B}$	1.1	3.2	3.2
$u\bar{u} + d\bar{d} + c\bar{c} + s\bar{s}$	3.4	10.2	10.1
e^+e^-	$\sim$53	159	156
$\mu^+\mu^-$	1.2	3.5	3.1
$\tau^+\tau^-$	0.9	2.8	2.4

Table 3.5: Cross sections, production and trigger rates for the principal physics processes at $10.58\,\mathrm{GeV}$ for a luminosity of $3\times10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The e^+e^- cross section refers to events with either the e^+, e^-, or both inside the EMC detection volume.

The Level 1 trigger system consists of four components: the L1 drift chamber trigger (DCT), L1 calorimeter trigger (EMT), and IFR trigger used for calibration (IFT), and the global electronics for producing the final L1 accept signal (GLT). The DCT is algorithms are executed in three types of modules. First, track segments, their ϕ positions and drift time estimates are found using a set of 24 Track Segment Finder (TSF) modules. These data are then passed to the Binary Link Tracker (BLT) module, where segments are linked into complete tracks. In parallel, the ϕ information for segments found in axial superlayers is transmitted to eight transverse momentum discriminator (PTD) modules, which search for tracks above a set transverse momentum (p_T) threshold. All of this information is output

to the GLT.

The calorimeter trigger (EMT) divides the EMC into 280 *towers* of 24 crystals each (22 for the endcap). All crystal energies within a tower which are above a 20 MeV threshold are summed and supplied to the EMT trigger processor boards (TPBs). The TPBs digitally filter the energy deposition and compare neighboring towers to look for clusters which span more than one tower. Again this data corresponding to the energy and placement of found clusters is passed to the GLT.

The GLT receives the trigger line "primitives" (bytes corresponding to trigger type and information) from the EMT and DCT, as well as information from the IFT (which is used for $\mu^+\mu^-$ and cosmic ray triggering), and performs a timing alignment of these input data. The GLT does some rudimentary matching between DCT tracks and EMT clusters, and performs a logical AND of the input trigger primitives. The combined L1 trigger efficiency is $> 99.9\%$ for $B\overline{B}$ events, about 99% for continuum events, and 94.5% for $\tau^+\tau^-$ events.

The L3 trigger software comprises event reconstruction and classifications, a set of event selection filters, and monitoring. This software runs on the online computer farms within the Online Event Processing (OEP) framework. Many events which pass L1 but must be rejected by L3 are beam-induced charged particle background that are produced in material close to the IP. The Level 3 trigger combines DCT tracks (from the TSF system) and EMT clusters with the full DCH and EMC information. The L3 DCH algorithm performs fast pattern recognition and fits L1 tracks to helices and is able to determine the z_0 of tracks, which is important for rejecting the above mentioned background. The L3 EMC based trigger identifies energy clusters with a higher sensitivity than L1 and filters events with either high energy deposits of high cluster multiplicity. The output of both the DCH and EMC L3 filters is dominated by Bhabha events with are mostly rejected
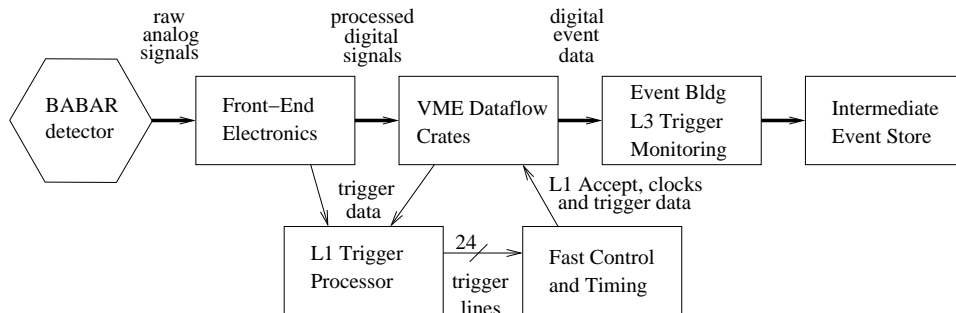
Figure 3.12: Schematic drawing the data path through *BABAR*'s online system.

but are also prescaled in L3 for calibration and luminosity online monitoring and offline measurements.

3.9.2 Data Flow and Detector Monitoring

Online Data-flow (ODF), which provides data transport, buffering, and event building is implemented in a set of VME crates which house 157 Readout Modules (ROMs) running `VxWorks`. Configuration and readout of the FEEs are performed through 1.2 Gbits/s fiber links to the ROMs. These specialized VME-based processors handle Feature Extraction (FEX) of physics signals, perform gain and pedestal corrections, as well as data specification and formatting. The calibration data is stored in a dedicated conditions database. Using a Fast Control and Timing System (FCTS), events are built from data from the individual subsystems and shipped from the ROM modules to the Online Event Processing (OEP) workstations via 100 Mb Ethernet. No dedicated counters are employed to associate events with beam crossings. Instead absolute timing is determined offline using DCH track segment timing, waveforms from EMC, and accelerator timing fiducials. Figure 3.12 presents a schematic diagram of the DAQ system.

Running on a farm of Unix workstations, the OEP software collects and pro-

cesses data from the ODF event builders, partially reconstructing the event in order to apply the Level 3 (L3) trigger algorithms and provide fast monitoring for the data taking personnel on shift. The L3 software examines the complete event information, categorizing and flagging physics, diagnostic, and calibration events for logging into 0.8 TB immediate storage.

Online Prompt Reconstruction (OPR) of the collected events occurs in as little as few hours after logging by farms of several hundred Unix workstations running in parallel. Using the raw detector signals and the partially reconstructed events of the L3 trigger, OPR performs full reconstruction of all physics events and select calibration events. These algorithms categorize potentially interesting events before storing the results into an object database for further analysis. Monitoring and rolling calibration of reconstructed parameters is also applied at this stage.

The Online Detector Control (ODC) system controls and monitors the electronics and environment of the detector and its support systems for safety assurance. Monitoring of machine status, injection inhibition, and beam aborting is achieved through links to PEP-II systems. All collected data is archived into a browsable ambient database.

The online machinery is tied together by the Online Run Control (ORC) system whose logic manages the state of all systems and provides a user interface for calibrations and starting/stopping runs. Detector configurations are stored in a configurations database for reference during reconstruction.

Chapter 4

The Analysis of B Mesons

In Chapter 2 we motivated the study of B mesons and the specific decay of $B^0 \rightarrow D^{*+}D^{*-}$ in the context of understanding CP violation as well as furthering our knowledge of B meson physics. We have also explained the *BABAR* detector in detail to show the capabilities of measuring the decay products of B mesons. This chapter is devoted to explaining the procedures and selection criteria used to obtain a sample of events in the $B^0 \rightarrow D^{*+}D^{*-}$ decay mode from the large data sample collected by *BABAR*. We start, however, with an overview of the two analyses presented in this dissertation.

4.1 Overview of the $B^0 \rightarrow D^{*+}D^{*-}$ Analyses

The decay $B^0 \rightarrow D^{*+}D^{*-}$ is, as mentioned in Section 2.3.5, not a CP eigenstate. The pseudoscalar B^0 meson decays into two vector particles, thus requiring a specific angular momentum configuration of the final state. The partial waves of this final state are the $L = 0, 2$ or "S-" and "D-wave" states (which have CP eigenvalue of $+1$), and the $L = 1$ or "P-wave" state (which has $CP = -1$), where L is the relative angular momentum between the two vector particles. Thus, the $B^0 \rightarrow D^{*+}D^{*-}$ decay has a mixture of CP-odd and CP-even eigenstates.

The amount of CP-odd versus CP-even component in the final state will determine how a time-dependent analysis is performed. It is, therefore, beneficial to 1) measure $R_\perp$, and 2) perform a simultaneous angular analysis with the time-dependent analysis. The latter will provide the most accurate time-dependent CP measurement in this mode, by removing the angular dilution factor $D = 1 - 2R_\perp$ from the CP asymmetry. Further, a time-independent measurement of $R_\perp$ would provide a test of the theoretical calculations in this mode that rely on the factorization approach, in addition to motivating the time-dependent analysis strategy.

The branching ratio of $B^0 \to D^{*+}D^{*-}$ decay is also of interest. While this measurement is not the primary focus of this thesis, the reader is referred to Reference [39] for the first *BABAR* measurement of the branching ratio.

The following sections detail how we analyze the full *BABAR* set (as of 2002) to select specific $B^0 \to D^{*+}D^{*-}$ events. This selection is common to both the time-independent $R_\perp$ measurement and the time-dependent CP asymmetry analysis. Discussion of the data sample itself, as well as the signal extraction is left to Chapter 5.

4.2   Pre-Selection of the Data

Production cross sections for the physics processes at the $\Upsilon(4S)$ energy were listed in Table 3.1. In addition to the $\Upsilon(4S) \to B\bar{B}$ decay, these processes include continuum $q\bar{q}$ and QED events such as $e^+e^- \to e^+e^-, \mu^+\mu^-, \tau^+\tau^-$, and $\gamma\gamma$.

The event topology is significantly different for each type of process. Table 4.1 summarizes the main characteristics of each process.

For the $B^0 \to D^{*+}D^{*-}$ analyses we are only interested in $B\overline{B}$ events. The normalized second Fox-Wolfram moment [51] R_2 is used to reduce background

Event type	Main characteristics
$e^+e^- \to e^+e^-(\gamma)$	Two high-momentum back-to-back tracks, and associated energy deposit in the EMC
$e^+e^- \to \mu^+\mu^-(\gamma)$	Two high-momentum back-to-back tracks
$e^+e^- \to \tau^+\tau^-$	Back-to-back topology with large missing energy, due neutrinos from semileptonic τ decays
$e^+e^- \to \gamma\gamma$	Large missing energy, and small number of tracks due to preferential production of particles along the beam direction
$e^+e^- \to q\bar{q}$ with $q = u, d, s, c$	Large number of hadrons and jet-like topology, due to the hadronization of the quarks which are produced back-to-back.
$e^+e^- \to \Upsilon(4S) \to B\bar{B}$	Large number of hadrons and isotropic topology due to the B decays.

Table 4.1: Main characteristics of the physics processes at the $\Upsilon(4S)$ energy, in the center-of-mass frame.

from continuum $u\bar{u}$, $d\bar{d}$, $s\bar{s}$, and $c\bar{c}$ events.

The ℓth Fox-Wolfram moment $\mathcal{H}_\ell$ is the momentum-weighted sum of Legendre polynomials of ℓth order, computed from the cosine of the angle between all pairs of tracks. Each $\mathcal{H}_\ell$ is essentially a multipole moment of the momentum distribution in an event. The $\mathcal{H}_0$ moment is the analog of the electric charge distribution. The first moment $\mathcal{H}_1$ is zero because the momentum is not a signed quantity (unlike the electric charge) and therefore can not have a dipole moment. The quadrupole moment $\mathcal{H}_2$ can instead discriminate events with a jet-like structure of momentum ($q\bar{q}$ events) from those with a more spherically symmetric topology ($B\bar{B}$ events).

The normalized ratio $R_2 = \mathcal{H}_2/\mathcal{H}_0$ is therefore very close to unity for events with back-to-back tracks such as QED events, and approaches 0 for isotropic events like $B\bar{B}$ events. The distribution of R_2 for the physics processes at the $\Upsilon(4S)$ energy is shown in Figure 4.1. The value of R_2 is computed with both charged tracks and neutral particles and is required to be less than 0.6 for $B^0 \rightarrow D^{*+}D^{*-}$ selection.
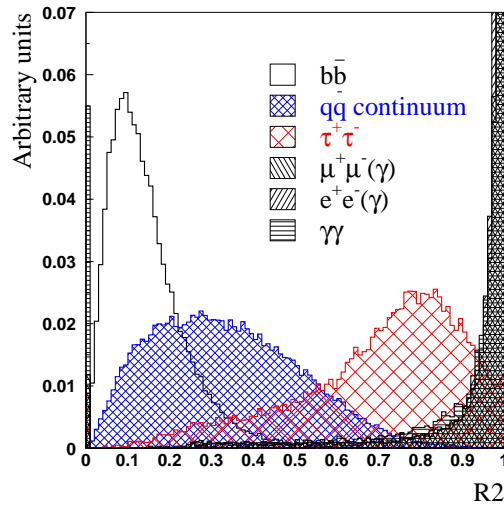
Figure 4.1: Distribution of R_2 for the main physics processes at the $\Upsilon(4S)$ energy. The distributions are normalized to the same area.

Another topological variable is used to help reduce continuum background events. Since the continuum events have a jet-like distribution, the direction of the jets can be used to distinguish events from the more isotropic B decays. The vector $\vec{A}_B$ for a B candidate is found by maximizing the ratio V_T

$$V_T = \frac{\sum_i^{1,N} |\vec{A}_B \cdot \vec{p}_i^{\,*}|}{\sum_i^{1,N} \sqrt{\vec{p}_i^{\,*} \cdot \vec{p}_i^{\,*}}} \ , \tag{4.1}$$

where the sum is over the all charged and neutral particles in the event that were not used to reconstruct the B candidate, and $\vec{p}_i^{\,*}$ is their three-momentum vectors in the $\Upsilon(4S)$ rest frame. The cosine of the thrust angle θ_T between the three-momentum $\vec{p}_B^{\,*}$ of the B candidate and the thrust axis $\vec{A}_B$

$$\cos \theta_T = \frac{\vec{p}_B \cdot \vec{A}_B}{|\vec{p}_B||\vec{A}_B|} \tag{4.2}$$

peaks at ± 1 in the jet-like $q\bar{q}$ continuum events, but is uniformly distributed in the isotropic $B\bar{B}$ events, as shown in Figure 4.2. For the decay $B^0 \to D^{*+}D^{*-}$ we require $\cos \theta_T < 0.9$.

4.3   Components of the Decay Chain

The decay products of the process $B^0 \to D^{*+}D^{*-}$ are fully reconstructed for both time-independent $R_\perp$ and time-dependent CP analyses. The reconstruction of an event involves first the selection of charged tracks and neutral candidates from the procedures outlined in Sections 4.3.1,4.3.2 (also see Sections 3.5,3.7). Next,the tracks are combined to form composite candidates such as K_S^0, D and D^* mesons. The decay modes of the D and D^* that were considered for this analysis are listed in Table 4.2. Finally, two D^* candidates are combined to form a B candidate. The selection variables used to distinguish between signal and background is described in Section 4.5.

Table 4.2: D^*, D^0 and D^+ decay modes and branching fractions [50]. For $B^0 \rightarrow D^{*+}D^{*-}$, the combination where both D^*'s decay to $D^+\pi^0$ is not included in the analysis. The branching fraction for $K_S^0 \rightarrow \pi^+\pi^-$ is included for decays containing a K_S^0.

Decay Mode	Branching Fraction (%)
$D^{*+} \rightarrow D^0\pi^+$	67.7 ± 0.5
$D^{*+} \rightarrow D^+\pi^0$	30.7 ± 0.5
Total Reconstructed D^{*+} Branching Fraction	98.4
Decay Mode	Branching Fraction (%)
$D^0 \rightarrow K^-\pi^+$	3.83 ± 0.09
$D^0 \rightarrow K^-\pi^+\pi^0$	13.9 ± 0.9
$D^0 \rightarrow K^-\pi^+\pi^+\pi^-$	7.49 ± 0.31
$D^0 \rightarrow K_S^0\pi^+\pi^-$	1.85 ± 0.14
Total D^0 Branching Fraction	27.1
Decay Mode	Branching Fraction (%)
$D^+ \rightarrow K^-\pi^+\pi^+$	9.0 ± 0.6
$D^+ \rightarrow K_S^0\pi^+$	0.99 ± 0.09
$D^+ \rightarrow K^-K^+\pi^+$	0.87 ± 0.07
Total D^+ Branching Fraction	10.9
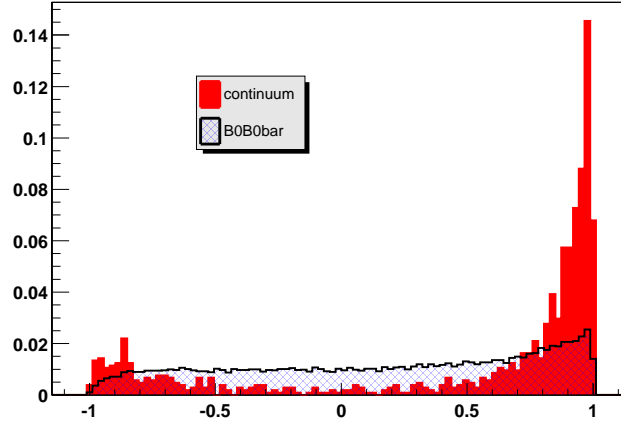
Figure 4.2: Distribution of the cosine of the angle between the thrust axis of the reconstructed B meson and the thrust axis of the remaining tracks in the event. The distributions are normalized to the same area.

Vertex and kinematic fitting techniques are applied to improve the resolution on the measured momentum of reconstructed mesons and further reduce contributions from combinatorial background. A general description of these techniques and their benefits is beyond the scope of this thesis. A comprehensive discussion of the kinematic- and vertex-fitting techniques can be found in a series of lectures by Paul Avery [52].

4.3.1 Selection of Charged Particles

Several quality requirements are applied to the charged tracks reconstructed in the tracking system to define *lists* of tracks for analysis purposes [53]. The lists are hierarchical: starting from a list including all reconstructed charged tracks, tighter requirements are applied to define good quality tracks.

1. `ChargedTracks`: All tracks reconstructed in the drift chamber and/or the silicon vertex tracker with the pion hypothesis.

2. `GoodTracksVeryLoose`: Subset of `ChargedTracks` with additional requirements:

   - center-of-mass momentum p^* less than $10\,\mathrm{GeV}/c$, and

   - distance from the nominal beamspot less than $1.5\,\mathrm{cm}$ in the transverse x-y plane, and less than $10\,\mathrm{cm}$ along the z axis.

3. `GoodTracksLoose`: Subset of `GoodTracksVeryLoose` that satisfy the following requirements:

   - transverse momentum p_t greater than $100\,\mathrm{MeV}/c$, and

   - at least 12 hits in the drift chamber.

4.3.2   Selection of neutral particles

The neutral particles reconstructed in the electromagnetic calorimeter are organized in hierarchical lists, similar to charged tracks, on the basis of the following quality requirements:

1. `CalorNeutral`: All energy bumps in the electromagnetic calorimeter not associated with any charged track.

2. `GoodNeutralLooseAcc`: Subset of `CalorNeutral` with additional requirements:

   - energy E greater than $30\,\mathrm{MeV}/c$,

   - lateral shaper parameter λ_{LAT} [46] less than 1.1, and

- within the *fiducial volume* of the EMC defined as $0.41 < \theta_{\mathrm{LAB}} < 2.409\,\mathrm{rad}$, where θ_{LAB} is the polar angle in the laboratory frame.

3. `GoodPhotonLoose`: Subset of `CalorNeutral` with additional requirements:

   - energy E greater than $30\,\mathrm{MeV}/c$, and

   - lateral shaper parameter λ_{LAT} [46] less than 0.8.

4. `GoodPhotonDefault`: Subset of `GoodPhotonDefault` with minimum energy of $100\,\mathrm{MeV}/c$.

4.3.3 Kaon Selection

Kaon identification is also used to reduce combinatorial background in the reconstruction of B meson candidates. Kaons are distinguished from pions and protons on the basis of specific energy-loss measurements $\mathrm{d}E/\mathrm{d}x$ in SVT and DCH, and the number of Cerenkov photons and the Cerenkov angle in the DIRC (also see Section 3.6).

A likelihood is constructed for each particle hypothesis from the product of two components: the expected number N_γ of Cerenkov photons, assuming a Poisson distribution, and the difference between the measured average Cerenkov angle θ_c (Figure 4.3) and the expected angle θ_c^0, for a given mass hypothesis, assuming a Gaussian distribution. Loose kaon identification is used in exclusive B reconstruction, while the b-flavor tagging (see Section 7.2) is based on tighter criteria.

The `NotAPion` kaon selection is defined by combining individual likelihoods from the SVT and DCH for momenta below $0.5\,\mathrm{GeV}/c$, from the DCH only for momenta between 0.5 and $0.6\,\mathrm{GeV}/c$, and from the DIRC only for momenta above $0.6\,\mathrm{GeV}/c$. Kaon candidates are rejected if the likelihood ratios satisfy $\mathcal{L}_K/\mathcal{L}_\pi < r$
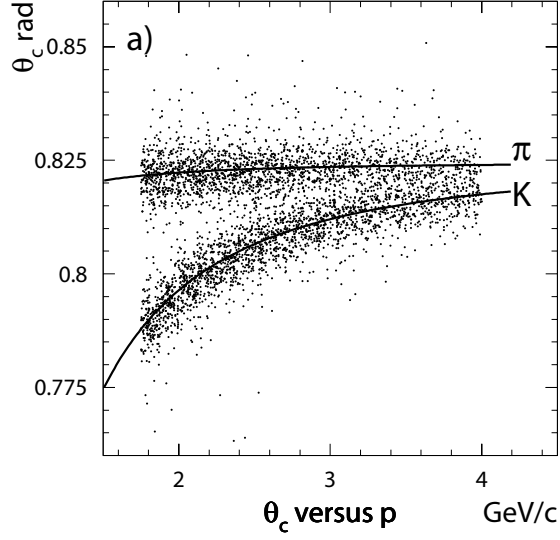
Figure 4.3: Distribution of the measured Cerenkov angle θ_c in a control sample of $D^0 \to K^- \pi^+$ decays.

and $\mathcal{L}_K/\mathcal{L}_p < r$, where $r = 0.1$ for $p < 0.5\,\text{GeV}/c$ and $r = 1$ for $p \geq 0.5\,\text{GeV}/c$. Tracks with no particle information are assumed to be pions.

The `NotAPion` kaon requirement has a nearly constant kaon-identification efficiency of about 96%, and a pion-misidentification probability of not larger than 30%, for tracks in the transverse momentum between 1 and $2.5\,\text{GeV}/c$. Tighter kaon selections require $\mathcal{L}_K/\mathcal{L}_\pi > r$, with r typically greater than one.

For this analysis a list of candidates which satisfies both the `NotAPion` kaon selection as well as the `GoodTrackLoose` track selection is made. This list is referred to as `KMicroNotPionGTL` and is the primary list of kaon candidates used for D meson reconstruction.

4.4 Reconstruction of Composite Particles

The following sections detail the selection and reconstruction of the composite particles involved in the decay of $B^0 \to D^{*+}D^{*-}$.

4.4.1 π^0 Reconstruction

The π^0 meson decays to a pair of photons about 98.8% of the time. Two `GoodPhotonLoose` photon candidates (see Section 4.3.2) are combined to form π^0 candidates. Photon pairs with invariant mass within $\pm 35\,\text{MeV}/c^2$ of the nominal π^0 mass ($135\,\text{MeV}/c^2$ [50]), and a minimum energy of 200 MeV are selected. The invariant mass $m(\gamma\gamma)$ for these candidates is shown in Figure 4.4.
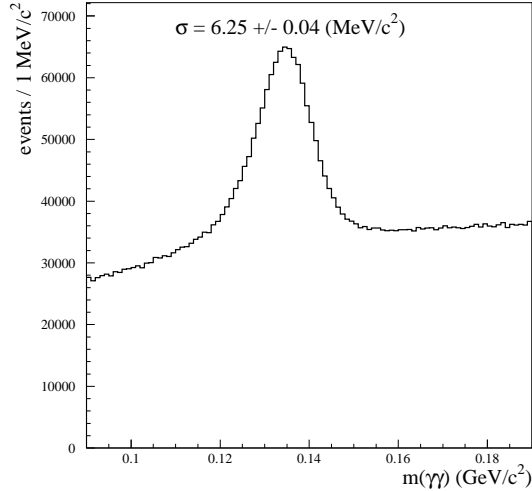
Figure 4.4: Invariant mass $m(\gamma\gamma)$ for selected π^0 candidates with $E_{\pi^0} > 200\,\text{MeV}$.

Selected candidates are subjected to a kinematic fit, with the the $\gamma\gamma$ invariant mass constrained to be the nominal π^0 mass (mass constraint). The mass constraint improves the energy resolution of the selected π^0 candidates.

The π^0 candidates to be used in the reconstruction of $D^{*+} \to D^+ \pi^0$ are also formed from two photons of the `GoodPhotonLoose` selection. Because of the lower momenta of these candidates, we require the composite $\gamma\gamma$ momentum in the center of mass frame to be within the range $70\,\mathrm{MeV}/c < p^* < 450\,\mathrm{MeV}/c$. These candidates are also subjected to a mass-constrained fit.

4.4.2 K_S^0 Reconstruction

The K_S^0 candidates are reconstructed in decay mode $\pi^+ \pi^-$ which has a branching fraction of 68.6% [50]. A pair of oppositely-charged `ChargedTracks` tracks (see Section 4.3.1) are required to originate from a common point (vertex fit). Candidates with invariant mass $m(\pi^+\pi^-)$ within $\pm 15\,\mathrm{MeV}/c^2$ of the nominal K_S^0 mass are selected. In addition, the probability for the tracks to have a common vertex (vertex χ^2 probability) must be greater than 0.1%.

4.4.3 D^0 and D^+ Reconstruction

The decay modes of the charmed D^0 and D^+ mesons reconstructed in this analysis are listed in Table 4.2. Candidates in these modes are formed by combining pion candidates from the `GoodTracksVeryLoose` selection, kaon candidates from the `KMicroNotPionGTL` selection, and K_S^0 and π^0 candidates from the selection described above. The kaon selection used for the mode $D^0 \to K^- \pi^+$ is `GoodTracksLoose`, since the combinatorial background in this mode is lower and does not require additional kaon identification. The invariant mass for D^0 (D^+) candidates are required to be within ± 20 MeV/c^2 of the nominal D^0 (D^+) mass. This cut is used for all D^0 modes except $K^- \pi^+ \pi^0$, which has a looser cut of 35 MeV/c^2 due to the π^0 resolution. A vertex fit is performed on the decay daughters to ensure a common decay point. No requirement is made on the χ^2 probability,

however.

Finally, the momentum p^* of all D^0 and D^+ candidates in the $\Upsilon(4S)$ rest frame is required to be between 1.3 and 2.5 GeV/c. The lower limit corresponds to the minimum momentum of D mesons produced in B decays, and reduces the combinatorial background. The higher limit is needed to reject high-momentum D mesons produced in continuum $c\bar{c}$ events.

4.4.4 D^* Reconstruction

The D^* candidates in this analysis are reconstructed in two decay modes: $D^{*+} \to D^0\pi^+$ and $D^{*+} \to D^+\pi^0$. The two D^* candidates required to construct a B candidate can decay to either mode, however both are not allowed to decay to $D^+\pi^0$. This case contributes too much background because of the poorer π^0 reconstruction and the total branching fraction is too low to observe a significant number of events in the data sample.

The pion in the D^* decay is referred to as the *slow pion* because of its low momentum. Charged candidates are taken from the `GoodTracksVeryLoose` selection and neutral candidates are taken from the π^0 selection described above. The momentum of the charged slow pion is required to be between 70 and 450 MeV/c. The lower limit is the threshold for track reconstruction in the silicon vertex tracker, while the higher limit is the maximum possible momentum, in the laboratory frame, for the soft pion in these decays. The π^0 is required to have momentum in the center-of-mass frame of less than 450 MeV/c.

Before combining the pion candidate with the D candidate, the D^0 or D^+ is constrained to the nominal mass value. Because of this, the mass difference $\Delta m = m(D\pi) - m(D)$ between the $D\pi$ invariant mass and the mass of the D candidate will have a better resolution. Further, for $D^{*+} \to D^0\pi^+$, the beamspot

83

is used as an additional geometric constraint for the soft pion, when the D^* decay vertex is computed. The effective vertical size of the beamspot is increased to $40\mu m$ (from a few microns) to account for the transverse flight of the B mesons, which have a transverse momentum of about $340\,\mathrm{MeV}/c$. The vertex fit is not required to converge and no requirement is made on the χ^2 probability.

The data distribution of Δm for the $D^{*+} \rightarrow D^0\pi^+$ decay mode is shown in Figure 4.5. The resolution of Δm is worse for $D^{*+} \rightarrow D^+\pi^0$ because the energy
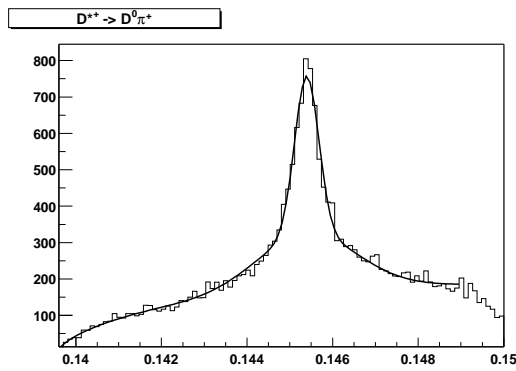
Figure 4.5: Mass difference D^*- D^0 for $D^{*+} \rightarrow D^0\pi^+$ candidates in the data. The D^0 candidate is constrained to the nominal mass before being combined with a soft pion.

resolution for π^0s is worse than that for charged tracks. The value of Δm is required to be within $\pm 11\,\mathrm{MeV}/c^2$ of the nominal Δm value for $D^{*+} \rightarrow D^0\pi^+$; it is required to be within $\pm 30\,\mathrm{MeV}/c^2$ for $D^{*+} \rightarrow D^+\pi^0$. Before being combined to form B^0 candidates, $D^{*+} \rightarrow D^0\pi^+$ candidates are constrained to the nominal D^* mass.

4.5 Reconstruction of B Mesons

B meson candidates are constructed by combining two oppositely charged D^* candidates that have passed the selection criteria described previously. The vari-

ables [54] used to distinguish our signal from background are

1. The difference ΔE, defined as

$$\Delta E = E^* - E^*_{beam} ,\qquad(4.3)$$

is the difference between the energy E^* of the B candidate and the beam energy E^*_{beam} in the $\Upsilon(4S)$ rest frame. In this frame, E^*_{beam} is simply half of the $\Upsilon(4S)$ energy and represents the best estimate for the true energy of the B meson. Hence, ΔE has an expected value of zero for signal candidates. The RMS spread $\sigma(\Delta E)$ is given by the uncertainty σ_E on the measured energy and by the spread σ_B of the true B meson energy

$$\sigma^2(\Delta E) = \sigma_B^2 + \sigma_E^2 .\qquad(4.4)$$

The measured spread in the beam energies result in variations of the $\Upsilon(4S)$ energy and are the main contribution to σ_B. The beam-energy spread is of the order of a few MeV and therefore $\sigma(\Delta E)$ is dominated by σ_E. The selection criteria and data distributions for the ΔE values of B candidates is discussed in detail in Section 5.1.2.

2. The beam-energy–substituted mass m_{ES} is defined as

$$m_{\mathrm{ES}} = \sqrt{E^*_{beam}{}^2 - p^{*2}}\qquad(4.5)$$

where E^*_{beam} is the beam energy and p^* is the measured momentum of the B candidate in the $\Upsilon(4S)$ center-of-mass frame. The RMS spread of m_{ES} is given by [54]

$$\sigma^2(m_{\mathrm{ES}}) \approx \sigma_B^2 + \left(\frac{p}{M_B}\right)^2 \sigma_p^2\qquad(4.6)$$

where σ_p is the uncertainty on the measured momentum. Unlike $\sigma(\Delta E)$, since $p/m_B \approx [325\,\mathrm{MeV}/c]/[5279\,\mathrm{MeV}/c^2] \approx 0.06c$, the uncertainty $\sigma(m_{\mathrm{ES}})$ is dominated by the beam-energy spread σ_B.

3. A likelihood variable, $\mathcal{L}_{Mass}$, is constructed from all candidate D and D^* masses. The usefulness of this variable is two-fold: first, the selection criteria for the reconstructed D and D^* candidates can be optimized by a requirement on this variable, and second, since a given event will have multiple reconstructed B candidates, the $\mathcal{L}_{Mass}$ variable is used to choose one per event.

$\mathcal{L}_{Mass}$ is defined by four multiplicative probability distribution functions (pdf): single Gaussians for each D mass and a double-Gaussian pdf for each Δm term. If $\mathcal{G}(x; \mu, \sigma)$ is a *normalized* Gaussian distribution where x is the dependent variable, μ is the mean, and σ is the resolution, then $\mathcal{L}_{Mass}$ is defined as

$$
\begin{aligned}
\mathcal{L}_{Mass} =\ & \mathcal{G}(m_D; m_{D_{PDG}}, \sigma_{m_D}) \times \mathcal{G}(m_{\bar{D}}; m_{\bar{D}_{PDG}}, \sigma_{m_{\bar{D}}}) \times \\
& \left[f_{core}\mathcal{G}(\Delta m_{D^{*+}}; \Delta m_{D^{*+}_{PDG}}, \sigma_{\Delta m_{core}}) \right. \\
& \left. + (1 - f_{core})\mathcal{G}(\Delta m_{D^{*+}}; \Delta m_{D^{*+}_{PDG}}, \sigma_{\Delta m_{tail}}) \right] \times \\
& \left[f_{core}\mathcal{G}(\Delta m_{D^{*-}}; \Delta m_{D^{*-}_{PDG}}, \sigma_{\Delta m_{core}}) \right. \\
& \left. + (1 - f_{core})\mathcal{G}(\Delta m_{D^{*-}}; \Delta m_{D^{*-}_{PDG}}, \sigma_{\Delta m_{tail}}) \right] \quad (4.7)
\end{aligned}
$$

where PDG refers to the nominal value [50]. For σ_{m_D} the errors calculated candidate-by-candidate from track reconstruction are used. The parameter f_{core} is the ratio of areas for the core and tail Gaussians. This along with $\sigma_{\Delta m_{core}}$ and $\sigma_{\Delta m_{tail}}$ are determined from fitting the Δm distributions in simulated events (Monte Carlo). When reconstructing Monte Carlo events, we use the σ_{m_D} and $\sigma_{\Delta m}$ values as is, however for reconstructing data we scale σ_{m_D} by 1.25 to account for the different resolutions seen in data and Monte Carlo. No additional smearing is applied to the Δm resolutions. Since events contain more than one B candidate, we choose only one candidate per event

by picking the one with the largest value of $\mathcal{L}_{Mass}$. We also establish $\mathcal{L}_{Mass}$ criteria for each final state of the $B^0 \to D^{*+}D^{*-}$ decay. This is discussed fully in Section 4.6.

The $\mathcal{L}_{Mass}$ distribution is displayed in Figure 4.6. In order to make the plot easier to interpret, the variable $-\ln(\mathcal{L}_{Mass}/\mathcal{L}_{Mass}^{MAX})$ is shown, where $\mathcal{L}_{Mass}^{MAX}$ is determined from setting $m_D, m_{D^*} = m_{PDG}$ in the definition of $\mathcal{L}_{Mass}$.

The m_{ES} and ΔE variables are nearly uncorrelated. Thus a signal region can be defined near $m_{\mathrm{ES}} = m_B \, \mathrm{GeV}/c^2$ and $\Delta E = 0 \, \mathrm{MeV}$. For the purpose of determining event yields and purities, four regions are defined in the $(m_{\mathrm{ES}}, \Delta E)$ plane. These regions are illustrated in Figure 4.7. An m_{ES} *projection* plot is used to extract signal yields and is the one-dimensional projection of the data after imposing the ΔE criteria; the composition of the candidates in this region is illustrated in Figure 4.8. The signal component is parameterized with a Gaussian centered at the B meson mass. The background contribution is separated into *combinatorial* and *peaking* components.

The combinatorial background arises from random combinations of charged and neutral particles. The ΔE of these combinations is within the required window, but the m_{ES} is smoothly distributed and does not peak near the B mass. The m_{ES} distribution for these combinations is parameterized with a threshold function

$$\mathcal{A}(m_{\mathrm{ES}}; m_0, \kappa) = N_B \, m_{\mathrm{ES}} \sqrt{1 - (m_{\mathrm{ES}}/m_0)^2} \; \mathrm{e}^{\kappa \left(1 - (m_{\mathrm{ES}}/m_0)^2\right)}, \qquad (4.8)$$

commonly called the ARGUS function [55], where m_0 is the upper kinematic limit fixed at the beam energy E_{beam}, N_B is the normalization factor, and κ controls the slope of the function.

The peaking background is due to mis-reconstructed B candidates which have m_{ES} near the B mass. While $B^0 \to D^{*+}D^{*-}$ reconstruction tends to have little
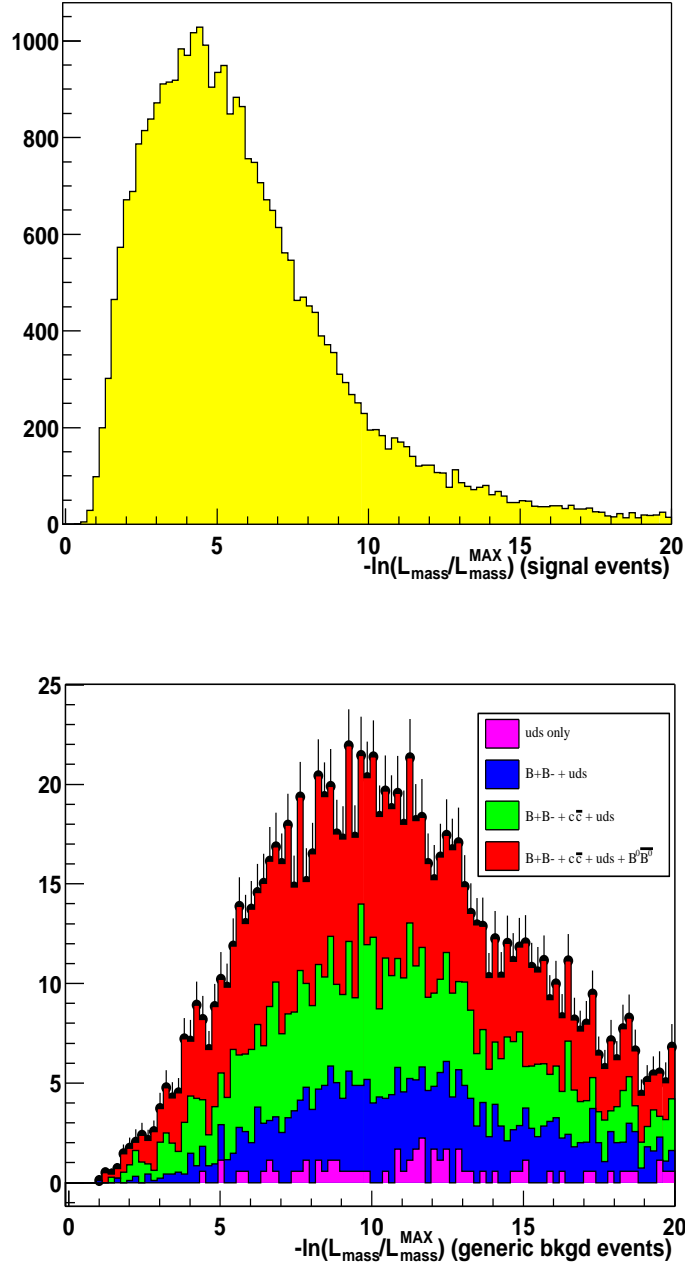
Figure 4.6: Signal and background Monte Carlo distributions of $-\ln(\mathcal{L}_{Mass}/\mathcal{L}_{Mass}^{\mathrm{MAX}})$ with no other selection criteria imposed. All reconstructed submodes of $B^0 \rightarrow D^{*+}D^{*-}$ are shown. The background distribution is a combination of generic $B^0\overline{B}^0, B^+B^-, c\bar{c}$ and uds $(u\bar{u}, d\bar{d}, s\bar{s})$ events scaled to the size of the data sample.
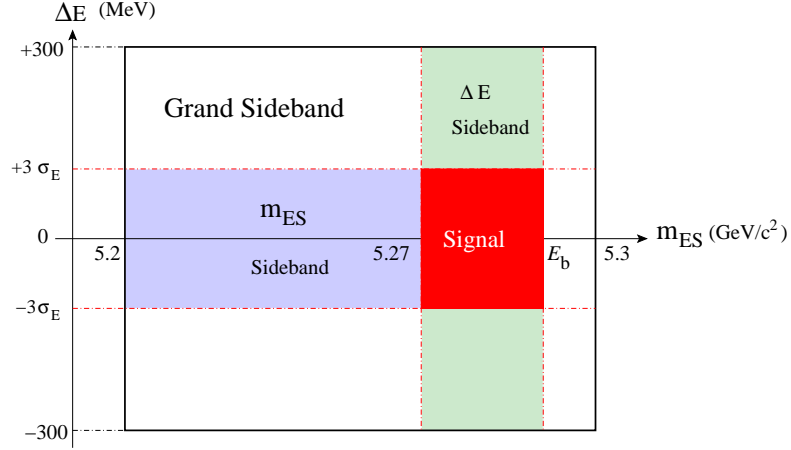
Figure 4.7: Definition of signal and sideband regions in the $(m_{\rm ES}, \Delta E)$ plane. The beam energy E_b is fixed at $5.291\,{\rm GeV}/c^2$.

peaking background contamination, a discussion of the studies done to verify this is in Section 5.2. The Δt distribution of events with a fake B^0 is different from that of the signal, and therefore directly affects the measurement of the time-dependent CP asymmetry in $B^0 \to D^{*+}D^{*-}$. The impact of the peaking background on the measured CP asymmetry is taken into account in the systematic uncertainty, and is discussed in Section 8.4.2.

The purity $\mathcal{P}$ for the selected candidates is defined as

$$\mathcal{P} = \frac{\int_{5.27}^{E_b} dm_{\rm ES}\, \mathcal{G}(m_{\rm ES})}{\int_{5.27}^{E_b} dm_{\rm ES}\, (\mathcal{G}(m_{\rm ES}) + \mathcal{A}(m_{\rm ES}))} \ . \tag{4.9}$$

Each candidate is assigned a per-event signal probability P, on the basis of the measured $m_{\rm ES}$, defined as

$$P(m_{\rm ES}) = \frac{\mathcal{G}(m_{\rm ES})}{\mathcal{G}(m_{\rm ES}) + \mathcal{A}(m_{\rm ES})} \ . \tag{4.10}$$

This probability will be used in the likelihood fits of the CP and $R_\perp$ measurements.

89
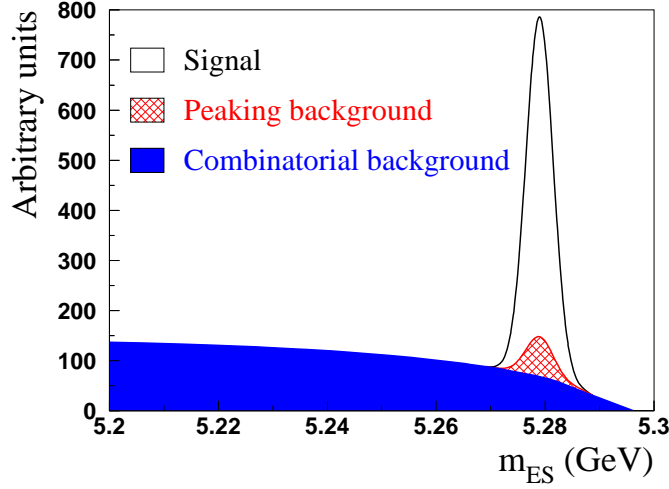
Figure 4.8: The composition of the m_{ES} distribution in the signal region. Note that for $B^0 \rightarrow D^{*+}D^{*-}$ the number of background events which peak in m_{ES} is estimated to be small. The size of the component illustrated here is only an example.

4.6 Selection Optimization

To obtain a sample of $B^0 \rightarrow D^{*+}D^{*-}$ candidates which maximizes signal and minimizes combinatorial background, the requirements on the $\mathcal{L}_{Mass}$ variable was optimized. Many of the requirements on D and D^* masses mentioned in Section 4.4 were not optimized; instead they were chosen to represent somewhat loose criteria in order to reduce the data (or Monte Carlo) sample to a reasonable size. Then, the $\mathcal{L}_{Mass}$ criteria for each decay mode, may then be optimized to select the "best" candidates in a given B candidate reconstruction.

The optimization of the $\mathcal{L}_{Mass}$ requirements was performed by maximizing the value of $S^2/(S+B)$. Here, S is the number of signal events that pass the selection and B is the estimation of background that passes the selection in the signal region. By maximizing this variable, the fractional statistical uncertainty on the

90

branching ratio measurement is minimized.[1] We note that the event selection for a time-dependent CP asymmetry measurement is sensitive to possible CP content in the background events (see Reference [15], for explanation). For this analysis, the difference in selection criteria is expected to be negligible and the more general variable $S^2/(S+B)$ is optimized.

For the signal estimation, exclusive Monte Carlo is generated for each specific $B^0 \to D^{*+}D^{*-}$ sub-decay mode[2] is used. A $B^0 \to D^{*+}D^{*-}$ branching ratio must be assumed to predict the value of S; for $B^0 \to D^{*+}D^{*-}$ 0.1% was used [50]. For the background estimation, appropriately weighted samples of Monte Carlo simulated as *generic* (that is, the mother particle(s) is allowed to decay to any final state to simulate the data) $B^0\overline{B}^0$, B^+B^-, and $c\overline{c}$ events is used. Because the generic $B^0\overline{B}^0$ events contain actual signal events, we remove these events from the sample so as not to overestimate the signal calculation. The number of background events reconstructed in the signal region is often very small. Thus, to decrease the statistical uncertainties incurred from a finite sample, the background distribution of the $\mathcal{L}_{Mass}$ variable is taken from the grand sideband (see Figure 4.7) and scaled to the size of the signal region.

Because the background level and expected signal is different for each sub-decay mode in $B^0 \to D^{*+}D^{*-}$, and because the D and D^* resolutions are potentially different in each submode, we prefer a separate optimized $\mathcal{L}_{Mass}$ requirement for each submode as opposed to one global criterion for all modes. The extracted signal sample used in the $R_\perp$ and CP asymmetry measurements will consist of events in

---

[1]The statistical uncertainty for a branching ratio is proportional to $\sqrt{S+B}$, while the branching ratio itself is proportional to S. Hence, the fractional uncertainty is given by $\sqrt{S+B}/S$.

[2]The term "submode" refers to the final state of the daughters of the B; for example, a "submode" of $B^0 \to D^{*+}D^{*-}$ is $(K\pi, K\pi\pi^0)$.

every sub-decay mode; thus, to optimize the mode-by-mode $\mathcal{L}_{Mass}$ requirements a "global" $S^2/(S+B)$ is constructed. The value of S is given by

$$S = \sum_{i=1}^{\#modes} N_i \times \epsilon_i(\mathcal{L}_{Mass} \text{ cut}) \qquad (4.11)$$

where $\epsilon_i(\mathcal{L}_{Mass} \text{ cut})$ is the efficiency determined from signal Monte Carlo as a function of the $\mathcal{L}_{Mass}$ requirement ("cut value"). The factor N_i includes the appropriate Branching Ratio and luminosity scaling for each mode. The number of background events B expected in the signal region is found by counting the number of events from generic Monte Carlo in the grand sideband and scaling it to the number expected in the signal region. This scale factor is dependent on the ratio of sizes of the signal and background regions.

To simultaneously determine the optimal $\mathcal{L}_{Mass}$ requirements for every submode a minimization program, MINUIT [56], is used. Here, the global $S^2/(S+B)$ value can be maximized as a function of the 22 $\mathcal{L}_{Mass}$ requirements (one for each submode in $B^0 \rightarrow D^{*+}D^{*-}$). Thus, for each mode the value of S and the expected value of B is plotted as a function of the $\mathcal{L}_{Mass}$ "cut value".[3] To streamline the maximization process, each plot is fit to an empirical function which parameterizes the distribution in each submode. These parameterizations were the input to MINUIT and the output parameters were the optimal $\mathcal{L}_{Mass}$ cut values for each mode. The submode requirements for $B^0 \rightarrow D^{*+}D^{*-}$ are listed in Table 4.3.

A technique was also developed to verify the output of the MINUIT optimization as well as to aid the decisions on which submodes to exclude in the analysis. For the mode $B^0 \rightarrow D^{*+}D^{*-}$, we require MINUIT to optimize a 22 parameter space using the input parameterizations of histogrammed data. One method to

---

[3]The variable plotted is actually $-ln(\mathcal{L}_{Mass}/\mathcal{L}_{Mass}^{MAX})$, where $\mathcal{L}_{Mass}^{MAX}$ is determined from setting $m_D, m_{D^*} = m_{PDG}$ in the definition of $\mathcal{L}_{Mass}$. Plotting the log of $\mathcal{L}_{Mass}$ produces a simple χ^2-like distribution that is easier to look at.

Table 4.3: The cuts applied to the $-\ln(\mathcal{L}_{Mass}/\mathcal{L}_{Mass}^{MAX})$ variable used for each decay mode in $B^0 \to D^{*+}D^{*-}$. B candidates reconstructed in the given mode (D decay)(D decay) are required to have a $-\ln(\mathcal{L}_{Mass}/\mathcal{L}_{Mass}^{MAX})$ value that is less than the cut value given below. Entries with a cut value equal to zero indicate the mode was included in the cut optimization process but was excluded from the analysis. The charge conjugate is implied and for clarity is not displayed.

	$(K\pi)$	$(K\pi\pi^0)$	$(K3\pi)$	$(K_S^0\pi\pi)$	$(K\pi\pi)$	$(K_S^0\pi)$	$(KK\pi)$
$(K\pi)$	14.	12.5	10.	11.5	10.	8.5	8.
$(K\pi\pi^0)$		11.5	10.5	9.	7.5	7.	5.5
$(K3\pi)$			8.	7.	6.	6.5	0.
$(K_S^0\pi\pi)$				8.	7.5	20.	0.

verify that MINUIT obtains the "correct" maximum for $S^2/(S+B)$ is to examine slices of that 22 parameters space. To do so, we plot the value of $S^2/(S+B)$ as a function of one submode's cut value after all other $\mathcal{L}_{Mass}$ cuts are applied to all other submodes. We expect the cut value obtained from the MINUIT optimization to be somewhat close to the peak of the above described distribution. A plot of $S^2/(S+B)$ versus the $\mathcal{L}_{Mass}$ cut value for the mode $B^0 \to D^{*+}D^{*-} \to (K\pi\pi^0, K3\pi)$ is shown in Figure 4.9. The histogram is the sum of all histogram data (before the distributions are fitted) and the dotted line shows the $S^2/(S+B)$ value using the fitted parameterizations of the histograms. The vertical line indicates where MINUIT determined the optimal cut to be from the optimization process.

This distribution is also useful for determining which modes are "beneficial" to the analysis and which are not. In Figure 4.10 the same distribution is plotted for the mode $B^0 \to D^{*+}D^{*-} \to (K3\pi, KK\pi)$. It is clear that the global $S^2/(S+B)$
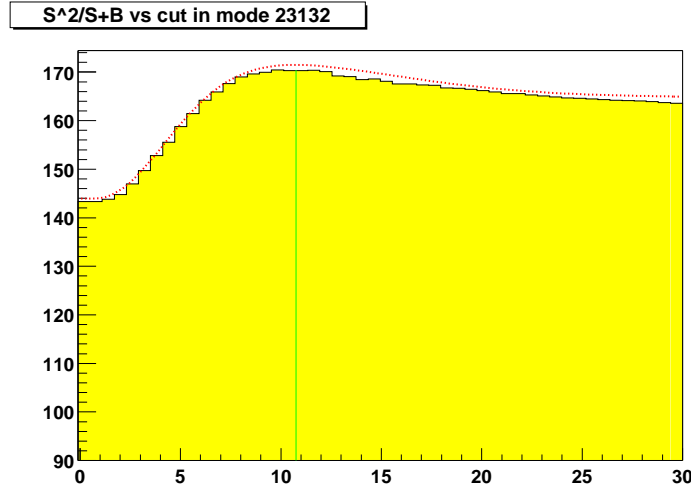
Figure 4.9: The total $S^2/(S + B)$ plotted as a function of the $\mathcal{L}_{Mass}$ cut value for the mode $B^0 \rightarrow D^{*+}D^{*-} \rightarrow (K\pi\pi^0, K3\pi)$ after all other cuts are applied. The histogram is obtained from the histogrammed signal and background Monte Carlo distributions, and the dotted line corresponds to the fits to the mode-by-mode $\mathcal{L}_{Mass}$ and ΔE distributions. The vertical line shows the cut value returned from the MINUIT optimization of $S^2/(S + B)$.

does not increase much when this mode is included, and in fact is detrimental to the analysis if a loose cut is used. Secondly, the "optimal" cut value obtained from MINUIT is considerably tight. A tight cut on the $\mathcal{L}_{Mass}$ variable could result in a large systematic uncertainty because of possible differences between data and Monte Carlo mass resolution differences.

Hence, for modes where the optimal cut is very tight and it is clear that the total $S^2/(S + B)$ does not significantly increase with the inclusion of a particular mode, then this mode will be dropped from the analysis.

Having determined the $\mathcal{L}_{Mass}$ requirements for the selection of $B^0 \rightarrow D^{*+}D^{*-}$ events, we now turn to the final selection variable, ΔE. This, as well as the data

S^2/S+B vs cut in mode 23239
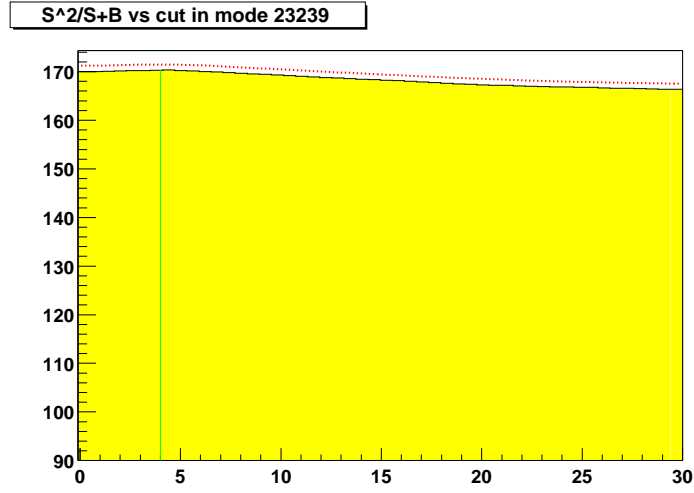
Figure 4.10: The total $S^2/(S + B)$ plotted as a function of the $\mathcal{L}_{Mass}$ cut value for the mode $B^0 \to D^{*+}D^{*-} \to (K3\pi, KK\pi)$ after all other cuts are applied. The histogram is obtained from the histogrammed signal and background Monte Carlo distributions, and the dotted line corresponds to the fits to the mode-by-mode $\mathcal{L}_{Mass}$ and ΔE distributions. The vertical line shows the cut value returned from the MINUIT optimization of $S^2/(S + B)$.

yields resulting from the full selection procedure is discussed in detail in Chapter 5.

Chapter 5

Data Sample and Signal Yields

Having described the methods used to exclusively reconstruct B mesons in the final state D^{*+} D^{*-}, we now present the event sample and the extracted signal yields. We first discuss in detail the final selection variable ΔE. After obtaining a signal distribution, we detail the studies used to determine the background content.

5.1   Event Yields

5.1.1   Data Sample

The data sample used is comprised of the data taken by *BABAR* during the November 1999 - June 2002 time period. The sample is divided into smaller time periods related to the shutdowns of the detector: Run 1 consists of data taken between 1999 and 2000; Run 2 consists of the data taken in 2001 and the first half of 2002. The sample used is data taken at the $\Upsilon(4S)$ resonance energy and corresponds to $81.8\,\mathrm{fb}^{-1}$. The total number of $B\overline{B}$ pairs produced in this sample is found to be $(88.0 \pm 1.0) \times 10^6$. This data set is used for both the time-dependent CP analysis and the time-independent transversity analysis for $B^0 \to D^{*+}D^{*-}$.
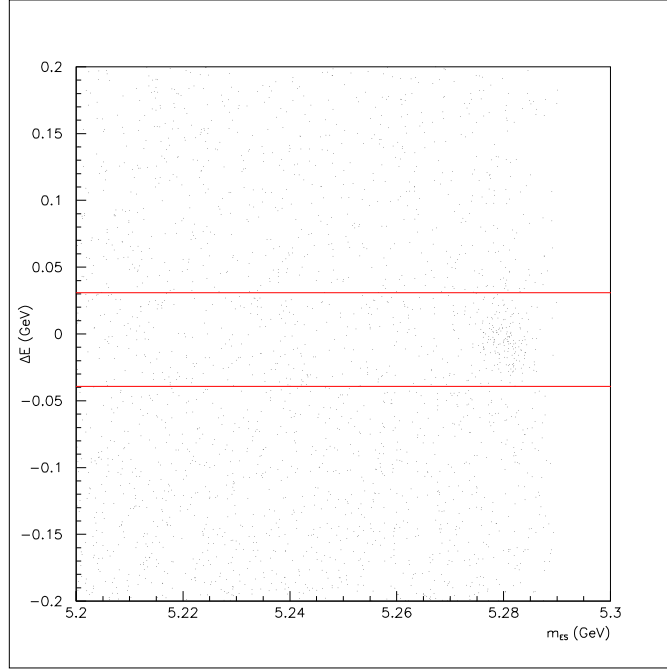
Figure 5.1: The ΔE vs. m_{ES} projection for $B^0 \to D^{*+}D^{*-}$ (full data sample). The lines show where the ΔE cut is applied to obtain the m_{ES} projection.

5.1.2 ΔE Distribution in $B^0 \to D^{*+}D^{*-}$

After all of the selection criteria described in Chapter 4 are applied the resulting sample is displayed in the ΔE versus m_{ES} two-dimensional plane, shown in Figure 5.1. The accumulation of points near $\Delta E = 0$ and $m_{\mathrm{ES}} = m_B$ contains the signal $B^0 \to D^{*+}D^{*-}$ events. A signal probability (see Equation 4.10) is needed in the subsequent CP and $R_\perp$ analyses and is obtained from the m_{ES} projection of the data.

It is of interest to first study the ΔE distribution (projection). Requiring events to have m_{ES} within 3σ of the B mass results in the distribution shown in Figure 5.2.

When the data sample is divided into the time periods described above, large
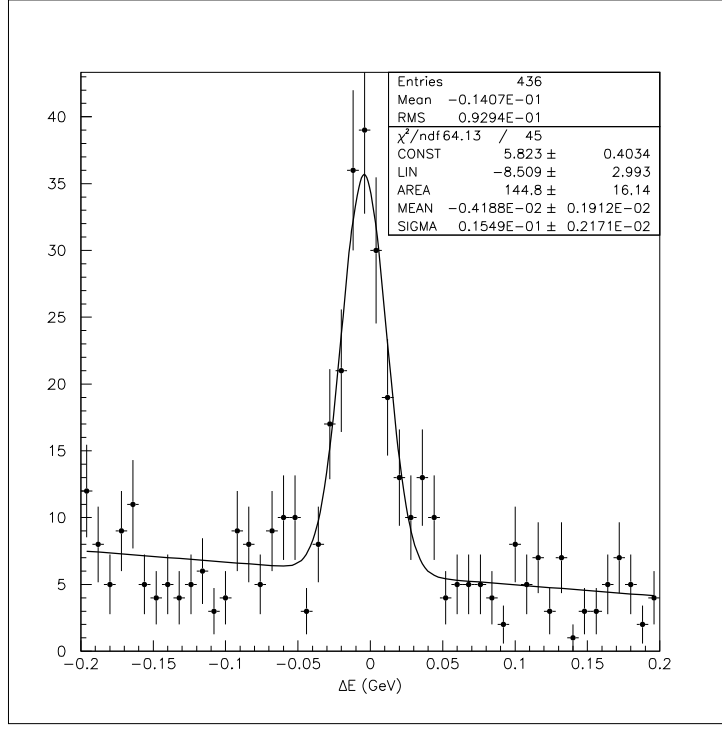
Figure 5.2: The ΔE projection for all $B^0 \to D^{*+}D^{*-}$ modes in the full data sample after applying the signal region cut on m_{ES}. The line is a fit to a Gaussian+linear polynomial shape. While this fit is not used in the final CP analysis, this figure is useful for understanding the ΔE resolution.

Table 5.1: The results of the fits to the ΔE projections for different time periods. The fit is the same as in Figure 5.2 .

		$\sigma_{\Delta E}$ (MeV)	mean of ΔE fit (MeV)
$B^0 \to D^{*+}D^{*-}$	Run1	13.5 ± 3.5	$+1.4 \pm 3.9$
	Run2 (2001)	11.2 ± 2.3	-7.7 ± 2.3
	Run2 (2002)	23.8 ± 3.1	-6.4 ± 4.4
	Total	15.5 ± 2.2	-4.2 ± 1.9
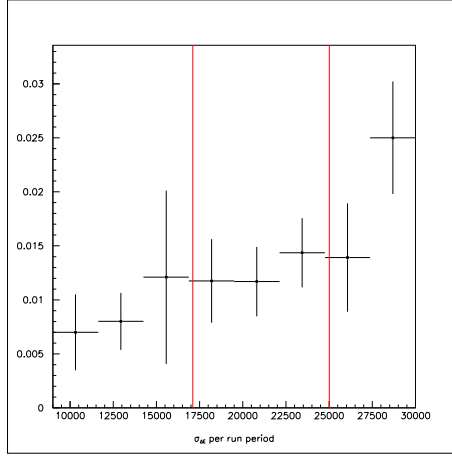
Figure 5.3: The resolution of the ΔE Gaussian+line fit as a function of time period. The data sample is divided into 8 sub-samples based on run number. The vertical lines denote where 2001 and 2002 running periods begin, respectively. Events in this projection are required to pass the m_{ES} signal region cut.
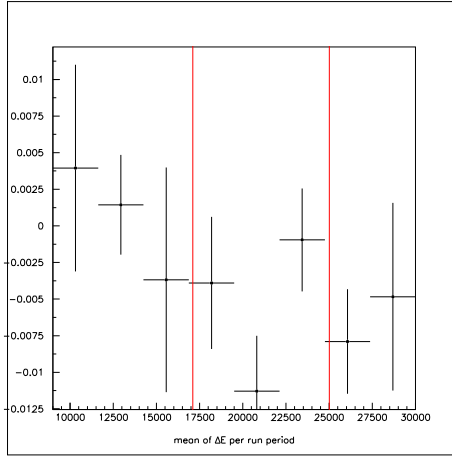
Figure 5.4: The mean of the ΔE Gaussian+line fit as a function of time period. The data sample is divided into 8 sub-samples based on run number. The vertical lines denote where 2001 and 2002 running periods begin, respectively. Events in this projection are required to pass the m_{ES} signal region cut.

fluctuations are observed in the ΔE resolution. This is shown in Table 5.1. In an effort to find possible problems in the data processing, the sample is divided into shorter time periods. Though the total number of $B^0 \to D^{*+}D^{*-}$ events is small, any substantial processing error or data corruption might be seen from these tests. A plot of the ΔE resolution as a function of the run number (representing the *BABAR* data-taking time periods) is shown in Figure 5.3. Though the last period of 2002 data has a very large $\sigma_{\Delta E}$, the corresponding histogram has very few events and a poor fit result. This suggests that a statistical fluctuation in the distribution is an explanation of the wider distribution.

As a second cross check, the data is divided into sets depending on the final state tracks. In Figure 5.5 the first column shows the ΔE distribution for all modes who's final state contains only charged tracks. The second column contains all modes where at least one D^0 decays to $K\pi\pi^0$. The third column contains all modes where one $D^{*\pm}$ decays to $D^{\pm}\pi^0$. While we are limited by low statistics, these plots show that the widened ΔE resolution is not due to any particular problem with track or neutral reconstruction.

Because of the fluctuation in the ΔE resolution, and because the mean is seen to be shifted slightly, the cut value used to make the m_{ES} distribution must be reconsidered. Before the data distribution was seen the ΔE criteria was obtained from an optimization of $S^2/(S+B)$. A single requirement was determined for all $B^0 \to D^{*+}D^{*-}$ submodes using the signal and generic Monte Carlo samples mentioned above. The optimal requirement was found to be $-25 < \Delta E < +25\,\mathrm{MeV}$. However, as seen in Figure 5.2, the data resolution is different than in the Monte Carlo and the $\pm 25\,\mathrm{MeV}$ criteria is not optimal for the sample. Based on the data distribution, the more optimal cut is determined to be a $35\,\mathrm{MeV}$ half-window centered on the mean ($-4.19\,\mathrm{MeV}$); that is, $-39.19 < \Delta E < +30.81\,\mathrm{MeV}$. Thus, for

Figure 5.5: The data sample is divided into three sets based on the D^*D^* final state. Column one shows the full sample, 2001 only, and 2002 only data, respectively, for all modes containing charged tracks only. Column two shows all modes where $D^0 \to K\pi\pi^0$. The third column shows all modes where one $D^{*\pm} \to D^{\pm}\pi^0$. All projections were required to pass the m_{ES} signal region cut.

both CP and $R_\perp$ analyses, the m_{ES} projection is made using this ΔE cut. The m_{ES} distribution is shown in Figure 5.6. The extracted number of $B^0 \to D^{*+}D^{*-}$ signal events is found from a Gaussian+ ARGUS fit to this distribution:

$$N_{B^0 \to D^{*+}D^{*-}} = 156.4 \pm 14.5 \tag{5.1}$$

with a purity (determined from Equation 4.9) of 73.1%. It should also be noted that neither analysis depends on the absolute efficiency determined from Monte Carlo, so a widened ΔE window does not introduce any systematic uncertainty.
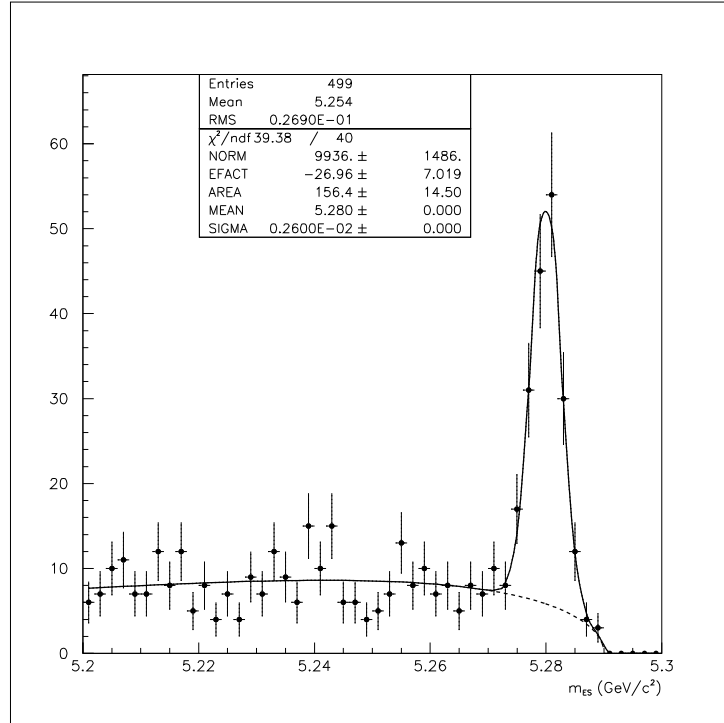
Figure 5.6: The m_{ES} projection for all $B^0 \to D^{*+}D^{*-}$ modes in the full data sample after applying the signal region cut on ΔE. The line is a fit to an AR-GUS+Gaussian shape where the Gaussian resolution is fixed to the value seen in Monte Carlo.

5.2 Peaking Background Studies

As mentioned in Section 4.5, candidates which are not $B^0 \to D^{*+}D^{*-}$ events but peak in the signal m_{ES} region are called "peaking background". Often the sources of these events can be determined from the generic Monte Carlo samples of $B^0\overline{B}^0$, $B^+ B^-$, $c\bar{c}$ events. For $B^0 \to D^{*+}D^{*-}$ reconstruction, there is **no** evidence from these samples of peaking background. However, the sample of generic Monte Carlo is somewhat smaller in size than the full data sample; further, there are potentially a few decays which are not well modeled in the Monte Carlo. Because of this, a more detailed study is performed.

The decay $B^\pm \to D^{*\pm}D^{*0}$ is not included in the Monte Carlo sample that is studied. Here a D^{*0} might be mis-reconstructed, for example, as a $D^{*\pm}$ by exchanging a $\pi^\pm$ for a γ and hence contribute in the m_{ES} signal region of $B^0 \to D^{*+}D^{*-}$ events. About 34,000 generated $B^\pm \to D^{*\pm}D^{*0}$ events were reconstructed as $B^0 \to D^{*+}D^{*-}$; this corresponds to approximately $550\,\mathrm{fb}^{-1}$ of data provided that $\mathcal{B}(B^0 \to D^{*+}D^{*-}) \approx \mathcal{B}(B^\pm \to D^{*\pm}D^{*0})$. Since only 3 events were found in the signal region, this potential source of peaking background was considered negligible.

The decay $B^0 \to D_s^{(*)}D^*$ is also a potential source of peaking background. While this mode is included in the generic Monte Carlo sample studied, the $D_s \to \phi\pi$ branching ratio has a relative error of over 25% (at the time of the study); a mis-reconstructed D_s or D_s^* might enable these events to peak in the $B^0 \to D^{*+}D^{*-}$ m_{ES} signal region. The $B^0 \to D^{*+}D^{*-}$ reconstruction was performed on a sample of 44,000 generated $B^0 \to D_s^{(*)}D^*$ events. This corresponds to an effective luminosity of $20\,\mathrm{fb}^{-1}$; no events passed the selection criteria and hence the mode was not considered to be a source of peaking background.

Other possible sources of background that might peak in m_{ES} but not in ΔE can be investigated by using data sidebands. Here, we are not limited by the size or content of a Monte Carlo sample. Figure 5.7 shows the m_{ES} distribution

MINUIT Likelihood Fit to Plots
mES
File: gloria.rz   4-JUL-2002 16:16
Plot Area Total/Fit   1146.0 / 1146.0   Fit Status 0
Func Area Total/Fit   1148.1 / 1148.1   E.D.M. 1.00
Likelihood =   34.0
χ^2 = 33.8 for 40 - 4 d.o.f.,   C.L.= 57.6%
Errors   Parabolic   Minos
Function 1: Gaussian (sigma)
* AREA   23.344   ± 12.95   - 12.65   + 13.25
* MEAN   5.2801   ± 0.   - 0.   + 0.
* SIGMA   2.60000E-03   ± 0.   - 0.   + 0.
Function 2: ARGUS Background
* NORM   35949.   ± 2793.   - 2705.   + 2882.
* OFFSET   -2.61786E-04   ± 2.4090E-05 - 5.6193E-05 + 0.
* EBEAM   5.2910   ± 0.   - 0.   + 0.
* EFACT   -33.131   ± 3.705   - 3.700   + 3.708

Figure 5.7: The m_{ES} distribution of $B^0 \rightarrow D^{*+}D^{*-}$ in the ΔE sideband fitted with an Argus function plus a Gaussian distribution to evaluate peaking background.

in the ΔE sideband ($50 < |\Delta E| < 200\,\mathrm{MeV}$) fitted with an ARGUS function plus a Gaussian distribution; the sigma and mean of the Gaussian and the end point of the ARGUS were fixed at the fitted values obtained from a fit to the m_{ES} distribution in the ΔE signal region. From this fit we conservatively estimate 3.9 ± 2.2 events of peaking background in the signal region from sources which have a linear distribution in ΔE. Nominally, we will assume, however, that there is no peaking background in the m_{ES} signal region since the estimation is consistent with zero. A systematic error can be determined where appropriate from the above

information (see Section 8.4.2). As an upper limit, which likely overestimates the possible peaking background, a one sigma fluctuation (+2.16 events) plus the estimate from possible $B^{\pm} \to D^{*\pm}D^{*0}$ peaking events (+0.44) is used. Hence, a conservative upper limit for peaking background in $B^0 \to D^{*+}D^{*-}$ is 6.5 events.

5.2.1   Summary

Table 5.2 shows the data sample yields and background estimates for $B^0 \to D^{*+}D^{*-}$ detailed in this chapter.

Table 5.2: Summary of the event sample for $B^0 \to D^{*+}D^{*-}$ in $81.8\,\text{fb}^{-1}$. The number of signal events was determined from a fit to a Gaussian+ARGUS function. The purity is defined in Equation 4.9, and the estimated background is determined from it. The number of peaking background events is assumed to be zero in the data sample; however, a conservative upper limit is given.

	N_{signal}	Purity	N_{bkgd}	N_{peak}
$B^0 \to D^{*+}D^{*-}$	156.4 ± 14.5	73.1%	57.6	0.0 (< 6.5)

Chapter 6

Time-Integrated Transversity Analysis

As mentioned in Section 2.3.5, the $B^0 \to D^{*+}D^{*-}$ final state is not a pure CP eigenstate. Unless we distinguish between the CP-odd (P-wave) and CP-even (S-,D-wave) components of the final state, the time-dependent CP asymmetry is diluted by a factor of $D = 1 - 2R_\perp$ (see Equation 2.40), where $R_\perp$ is defined as:

$$R_\perp = \frac{|A_\perp^0|^2}{|A_0^0|^2 + |A_\parallel^0|^2 + |A_\perp^0|^2} \tag{6.1}$$

(the superscript 0 refers to the magnitudes of the transversity amplitudes at $t = 0$). Theoretical calculations of the value of $R_\perp$ rely on the factorization approximation and predict the CP-odd component to be at the 5% level [29–31]. Since these are model-dependent calculations the value of $R_\perp$ should be determined experimentally. In this chapter we present a one-dimensional angular analysis of the $B^0 \to D^{*+}D^{*-}$ decay, to improve the existing experimental value of $R_\perp$ [39].

6.1 Overview

For any pseudo-scalar decay to two vector mesons, $B \to V_1 V_2$, there are three different angular momentum projections commonly used to describe it: the helicity basis, the transversity basis, and the partial wave decomposition. They are

106

completely equivalent, but quantities defined in these different basis states have different physical interpretations, and thus lead to slightly different physical insights about the underlying process. In the helicity basis, there are three amplitudes, $A_\lambda, (\lambda = 0, \pm 1)$ corresponding to the helicity of V_1 or V_2 in the decay $B \to V_1 V_2$. The "transverse" amplitudes are defined as spin projections for one vector particle parallel and perpendicular to the plane of the decay of the other. The amplitude A_0 remains unchanged, while the other two transversity amplitudes are defined as the following linear combinations of the helicity amplitudes [36, 37]:

$$A_\| = \frac{1}{\sqrt{2}}(A_{+1} + A_{-1}) \qquad\qquad A_\perp = \frac{1}{\sqrt{2}}(A_{+1} - A_{-1}). \qquad (6.2)$$

The helicity formalism gives a straightforward determination of the longitudinal rate, while the transversity formalism is used to determine the CP-odd component of the decay rate, thus allowing a ready interpretation of the CP asymmetry measurement. Lastly, the partial wave decomposition corresponds to the possible S, P and D orbital angular momenta, which refers to the relative angular momenta $L = 0, 1, 2$, respectively, between V_1 and V_2. In terms of the transversity amplitudes, the partial wave amplitudes are

$$S = \frac{1}{\sqrt{3}}(\sqrt{2}A_\| - A_0), \qquad P = A_\perp, \qquad D = \frac{1}{\sqrt{3}}(A_\| + \sqrt{2}A_0). \qquad (6.3)$$

Note that the $(-1)^L$-odd P-wave term is also the transversity amplitude $A_\perp$, while the other two transversity amplitudes are combinations of the $(-1)^L$-even S-wave and D-wave amplitudes.

Given this formalism, it is interesting to observe the physical implications for each basis. In the helicity basis, the angles are defined as follows for $B^0 \to D^{*+}D^{*-} \to D^0\pi^+\overline{D}^0\pi^-$: θ_1 is the polar angle of the π^- in the D^{*-} rest frame, θ_2 is the polar angle of the π^+ in the D^{*+} rest frame, and ϕ is the azimuthal angle between the D^{*+} and D^{*-} decay planes (see Figure 6.1 [38]).
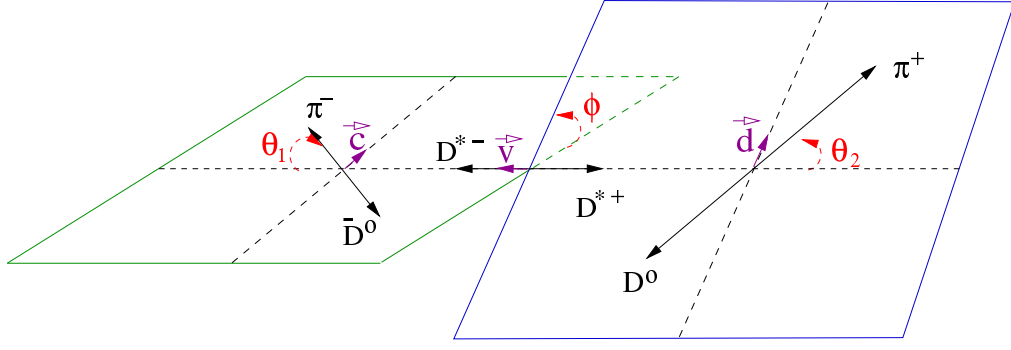
Figure 6.1: The helicity frame for $B^0 \to D^{*+}D^{*-} \to D^0\pi^+ \overline{D}^0\pi^-$. The decay of the D^{*-} is represented in the D^{*-} rest frame, while the decay products of the D^{*+} are shown in the D^{*+} rest frame. The angles are defined in the text.

In the transversity basis, the angles are defined as

- the polar angle θ_1 between the momentum of the π^- in the D^{*-} rest frame, and the direction of flight of the D^{*-} in the B rest frame,

- the polar angle θ_{tr} between the normal z to the D^{*-} decay plane and the π^+ line of flight in the D^{*+} rest frame, and

- the corresponding azimuthal angle ϕ_{tr} (see Figure 6.2).

The time dependent angular distribution of decay products in the transversity
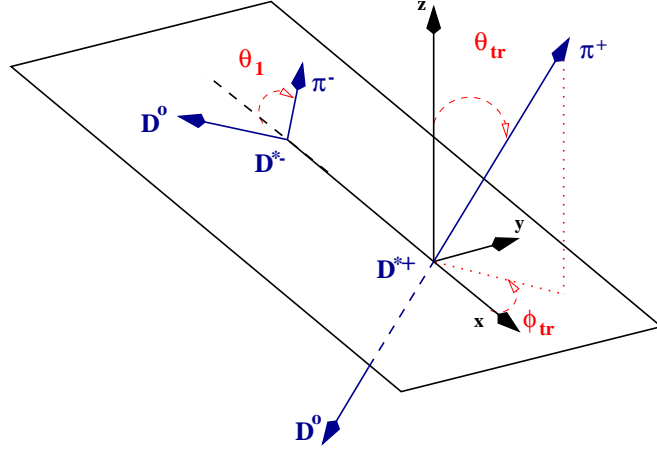
Figure 6.2: The transversity frame for $B^0 \to D^{*+}D^{*-} \to D^0\pi^+ \overline{D}{}^0\pi^-$. The decay of the D^{*-} is represented in the B^0 rest frame, while the decay products of the D^{*+} are shown in the D^{*+} rest frame. The x direction is defined by the direction of flight of the D^{*+} in the B^0 rest frame. The (x, y) plane is defined by the momenta of the D^{*-} decay products in the B^0 rest frame.

frame for the mode $B^0 \to D^{*+}D^{*-}$ is given by [35] (also see Appendix A):

$$
\frac{1}{\Gamma}\frac{\mathrm{d}^4\Gamma}{\mathrm{d}\cos\theta_1 \mathrm{d}\cos\theta_{\mathrm{tr}}\mathrm{d}\phi_{\mathrm{tr}}\mathrm{dt}} = \frac{9}{32\pi}\frac{1}{|A_0|^2 + |A_\parallel|^2 + |A_\perp|^2}
$$

$$
\{4|A_0|^2 \cos^2\theta_1 \sin^2\theta_{\mathrm{tr}} \cos^2\phi_{\mathrm{tr}}
$$

$$
+2|A_\parallel|^2 \sin^2\theta_1 \sin^2\theta_{\mathrm{tr}} \sin^2\phi_{\mathrm{tr}}
$$

$$
+2|A_\perp|^2 \sin^2\theta_1 \cos^2\theta_{\mathrm{tr}}
$$

$$
+\sqrt{2}\mathcal{R}e(A_\parallel^* A_0) \sin 2\theta_1 \sin^2\theta_{\mathrm{tr}} \sin 2\phi_{\mathrm{tr}}
$$

$$
-\sqrt{2}\mathcal{I}m(A_0^* A_\perp) \sin 2\theta_1 \sin 2\theta_{\mathrm{tr}} \cos\phi_{\mathrm{tr}}
$$

$$
-2\mathcal{I}m(A_\parallel^* A_\perp) \sin^2\theta_1 \sin 2\theta_{\mathrm{tr}} \sin\phi_{\mathrm{tr}}\} . \quad (6.4)
$$

where $A_0, A_\parallel, A_\perp$ are the time dependent decay amplitudes in the transversity

basis. For the $\overline{B}^0$ decay $A_\perp \to -A_\perp$. If, however, the flavor of the B is ignored, then the $\mathcal{I}m(A_\parallel^* A_\perp)$ and $\mathcal{I}m(A_0^* A_\perp)$ terms average to zero. Further, we will not consider the time dependence of the amplitudes; that discussion is left to Chapter 7. After integrating out the time dependence, the angular distribution keeps the same form with 4 terms and 4 free parameters remaining.

$$\frac{1}{\Gamma} \frac{\mathrm{d}^3\Gamma}{\mathrm{d}\cos\theta_1 \mathrm{d}\cos\theta_{\mathrm{tr}} \mathrm{d}\phi_{\mathrm{tr}}} = \frac{9}{16\pi} \frac{1}{|A_0^0|^2 + |A_\parallel^0|^2 + |A_\perp^0|^2}$$

$$\left\{ 2\cos^2\theta_1 \sin^2\theta_{\mathrm{tr}} \cos^2\phi_{\mathrm{tr}} |A_0^0|^2 \right.$$

$$+ \sin^2\theta_1 \sin^2\theta_{\mathrm{tr}} \sin^2\phi_{\mathrm{tr}} |A_\parallel^0|^2$$

$$+ \sin^2\theta_1 \cos^2\theta_{\mathrm{tr}} |A_\perp^0|^2$$

$$\left. + \frac{1}{\sqrt{2}} \sin 2\theta_1 \sin^2\theta_{\mathrm{tr}} \sin 2\phi_{\mathrm{tr}} \, \mathcal{R}e(A_0^{0*} A_\parallel^0) \right\} \, . \tag{6.5}$$

Any further integrations over the angles to reduce the number of free parameters is not appropriate without noting the effects of experimental detection efficiency. No detector is able to fully cover the entire space of the decay products. Secondly, it is likely that some regions within the detector have better tracking resolution than others. Together these effects lead to detection efficiencies which are dependent on the angular distribution of the decay products. These effects must be accounted for in the fit method. However, for illustration purposes, we first consider the angular distribution with perfect detection efficiency (acceptance).

If Equation 6.5 is integrated over $\cos\theta_1$ and ϕ_{tr}, assuming a perfect acceptance, then one obtains:

$$\frac{1}{\Gamma} \frac{\mathrm{d}\Gamma}{\mathrm{d}\cos\theta_{\mathrm{tr}}} = \frac{3}{4}(1 - R_\perp)\sin^2\theta_{\mathrm{tr}} + \frac{3}{2}R_\perp \cos^2\theta_{\mathrm{tr}} \tag{6.6}$$

where the proportion of the odd CP component was defined in Equation 6.1. Clearly, a one-parameter fit to the $\cos\theta_{\mathrm{tr}}$ distribution yields a measure of the

CP-odd component in $B^0 \to D^{*+}D^{*-}$. Figure 6.3 illustrates the possibilities of the final state and the expected distribution in the data for a given value of $R_\perp$. If $R_\perp$ is close to zero, then the $\sin^2 \theta_{\text{tr}}$ term of Equation 6.6 will dominate; this
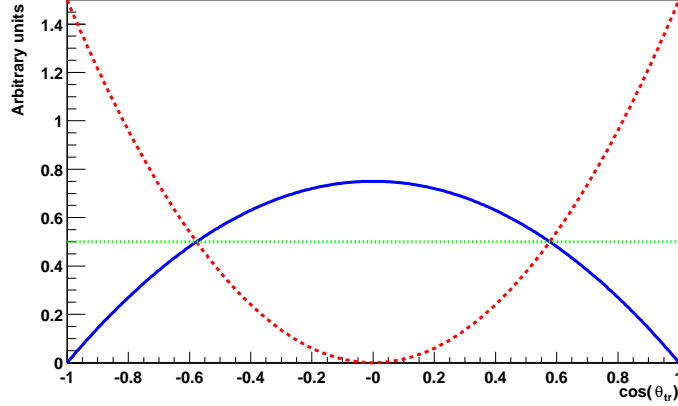
Figure 6.3: Theoretical $\cos \theta_{\text{tr}}$ distributions from Equation 6.6. If $B^0 \to D^{*+}D^{*-}$ has a purely CP-even final state, or $R_\perp = 0.$, then the solid line (blue curve) shows the resulting distribution. If the final state is purely CP-odd, or $R_\perp = 1.$, then the dashed line (red curve) is the corresponding distribution. If $R_\perp = 1/3$, as expected for background events, then the dotted line (green curve) is the expected distribution.

implies that a fully CP-even state will cause the soft pions to decay predominately in the same plane. The opposite case $(R_\perp \approx 1)$ requires the CP-odd final state to have a $\cos^2 \theta_{\text{tr}}$ distribution. We note that a distribution which is flat in $\cos \theta_{\text{tr}}$ demonstrates no directional preference for the decay products, and corresponds to $R_\perp = 1/3$; this is the expected distribution for background events.

Given the above simplified version of the differential decay rate, we now consider the more realistic case where detector acceptance is included. Multiplying Equation 6.5 by a three dimensional efficiency function, $\epsilon(\cos \theta_1, \cos \theta_{\text{tr}}, \phi_{\text{tr}})$, quan-

titatively describes the detector acceptance. After integrating over $\cos\theta_1$ and ϕ_{tr}, as before, we obtain:

$$\frac{1}{\Gamma}\frac{\mathrm{d}\Gamma}{\mathrm{d}\cos\theta_{\text{tr}}} =$$
$$\frac{3}{4}(1-R_\perp)\sin^2\theta_{\text{tr}} \times \left\{\frac{1+\alpha}{2}I_0(\cos\theta_{\text{tr}}) + \frac{1-\alpha}{2}I_\|(\cos\theta_{\text{tr}})\right\}$$
$$+ \quad \frac{3}{2}R_\perp\cos^2\theta_{\text{tr}} \times I_\perp(\cos\theta_{\text{tr}}) \tag{6.7}$$

with:

$$\alpha = \frac{|A_0^0|^2 - |A_\|^0|^2}{|A_0^0|^2 + |A_\|^0|^2}$$

$$\tag{6.8}$$

and the three *acceptance moments*, which are independent of the amplitudes:

$$I_0(\cos\theta_{\text{tr}}) = \frac{3}{2\pi}\int d\cos\theta_1 d\phi_{\text{tr}}\cos^2\theta_1\cos^2\phi_{\text{tr}} \times \epsilon(\cos\theta_1, \cos\theta_{\text{tr}}, \phi_{\text{tr}})$$
$$I_\|(\cos\theta_{\text{tr}}) = \frac{3}{4\pi}\int d\cos\theta_1 d\phi_{\text{tr}}\sin^2\theta_1\sin^2\phi_{\text{tr}} \times \epsilon(\cos\theta_1, \cos\theta_{\text{tr}}, \phi_{\text{tr}})$$
$$I_\perp(\cos\theta_{\text{tr}}) = \frac{3}{8\pi}\int d\cos\theta_1 d\phi_{\text{tr}}\sin^2\theta_1 \times \epsilon(\cos\theta_1, \cos\theta_{\text{tr}}, \phi_{\text{tr}}) \tag{6.9}$$

If these moments are modeled and determined using Monte Carlo, the acceptance effects can be explicitly accounted for in the fit for $R_\perp$. This will be the strategy used in the analysis presented here. The parameter α describes the relative contributions of $|A_0^0|$ and $|A_\|^0|$; the sensitivity of the distribution to this parameter is dependent on the shapes of the acceptance moment integrals. We note that if I_0 and $I_\|$ are of similar shape in $\cos\theta_{\text{tr}}$, then the value of α will not affect the fit result for $R_\perp$.

6.2 Fit Method

The measurement of the *CP*-odd component, $R_\perp$, of the $B^0 \to D^{*+}D^{*-}$ final state is based on an unbinned maximum likelihood fit to the $\cos(\theta_{\text{tr}})$ distribution, with

a simultaneous fit to the m_{ES} distribution. The probability density function for the m_{ES} distribution is given by the sum of normalized ARGUS and Gaussian functions; the relative weight of each function is given by a signal fraction, f_{sig}, which is allowed to float in the likelihood fit. The likelihood is then defined as:

$$\mathcal{L} = \prod_{i=1,n} \mathcal{L}_i = \prod_{i=1,n} \left[f_{sig} \times \mathcal{F}^{sig}_{m_{ES}}(m_{ES,i}; \sigma_{m_{ES}}) \times \mathcal{F}^{sig}_{R_\perp}(\cos(\theta_i); R_\perp) + \right.$$
$$\left. (1 - f_{sig}) \times \mathcal{F}^{bkg}_{m_{ES}}(m_{ES,i}; \kappa) \times \mathcal{F}_{bkg}(\cos(\theta_i); b_2) \right], \tag{6.10}$$

where n is the number of selected events in the m_{ES} distribution, $\mathcal{F}^{sig}_{m_{\mathrm{ES}}}$ is the signal Gaussian for the m_{ES} distribution and $\mathcal{F}^{bkg}_{m_{\mathrm{ES}}}$ is the background ARGUS shape with parameter κ [55] (the threshold parameter for the ARGUS function is fixed to 5.291 GeV). $\mathcal{F}^{sig}_{R_\perp}$ refers to the probability density function (pdf) for signal events (see Equation 6.7), and $\mathcal{F}_{bkg}$ is the background pdf with parameter b_2. The background shape is modeled by a polynomial in $\cos(\theta_{\mathrm{tr}})$:

$$\mathcal{F}_{bkg}(\cos(\theta_{\mathrm{tr}}); b_2) = N \times (1 + b_2 \cos^2(\theta_{\mathrm{tr}})) \tag{6.11}$$

where N is the normalization factor. The background has no *a priori* reason to have a non-even shape in $\cos(\theta_{\mathrm{tr}})$. However, to determine a systematic error on the background parameterization, we allow for all even and odd terms of $\cos(\theta_{\mathrm{tr}})$ up to a fourth order polynomial, with corresponding parameters b_1, b_2, b_3, and b_4.

The parameter α is fixed to zero for the likelihood fit. As mentioned previously, the $R_\perp$ pdf is sensitive to α only if the shapes of the moment integrals, I_0 and $I_\parallel$, are very different. The best method to obtain a value of α is with a full three dimensional angular analysis. The strategy used in this analysis is to fix the value of α and then scan the range (-1 to +1) to determine a systematic error from our lack of knowledge of the *CP*-even amplitudes.

The acceptance moments can be determined from Monte Carlo studies. Once calculated for a given binned distribution in $\cos(\theta_{\mathrm{tr}})$, the distributions can be

parameterized. These "shape functions" are then fixed in the signal pdf (Equation 6.7) to explicitly take into account the detector acceptance. A complete description of the moment integral determination is given in Section 6.3.

Because of possible differences in acceptance between charged and neutral pions, we calculate the acceptance moments for the three soft pion final states. They are denoted as:

- $\pi\pi$ for $D^{*+} \to D^0\pi^+, D^{*-} \to \overline{D}^0\pi^-$

- $\pi\pi^0$ for $D^{*+} \to D^0\pi^+, D^{*-} \to D^-\pi^0$

- $\pi^0\pi$ for $D^{*+} \to D^+\pi^0, D^{*-} \to \overline{D}^0\pi^-$

Because of the three sets of acceptance moments (determined for each combination of soft π type in the final state), three versions of the signal pdf are used and we simultaneously fit the three event types in the data. It also follows that the three event types, $\pi\pi$, $\pi\pi^0$ and $\pi^0\pi$, may also have different purities. Hence, we allow for three different signal fractions f_{sig} in the fit to the m_{ES} distribution.

Thus, the parameters that are floating in the likelihood fit are the following: σm_{ES} of the Gaussian, mean of m_{ES} Gaussian, κ (Argus shape parameter), $f_{sig}^{\pi\pi}$, $f_{sig}^{\pi\pi^0}$, $f_{sig}^{\pi^0\pi}$, b_2, and $R_\perp$.

Lastly, the reconstructed angular resolution of θ_{tr} must also be taken into account in the fit. To do so we modify the signal pdf, $\mathcal{F}_{R_\perp}^{sig}$, by convolving it with a pre-determined resolution function plus an additional component to represent mis-reconstructed events. As described in Section 6.4 the resolution in θ_{tr} is found to contain a fraction of events that are un-correlated with the true value of θ_{tr}. These events are parameterized in the resolution function fit and included as an additional parameterized term in the signal pdf. The resolution function determined from events correlated with θ_{true} is folded into the signal pdf via convolution. Thus,

114

we write:

$$\mathcal{F}^{sig}_{R_\perp}(\cos(\theta_{\mathrm{tr}})) = f_{misReco} \times \mathcal{F}^{sig}_{misReco} + (1 - f_{misReco}) \times \mathcal{F}_{R_\perp} \quad (6.12)$$

where $\mathcal{F}_{R_\perp}$ is detailed in Section 6.4 as the (normalized) convolved form of Equation 6.7 and $\mathcal{F}^{sig}_{misReco}$ is the parameterization of the mis-reconstructed events seen in signal Monte Carlo. The fraction $f_{misReco}$ represents the relative amount of events mis-reconstructed and is determined in the resolution function fit detailed in Section 6.4. The parameters of the resolution function are determined from a fit to signal Monte Carlo and are fixed in the likelihood fit of $R_\perp$.

6.3 Determining the Acceptance Moments

The acceptance moment integrals described by Equation 6.9 are calculated using a Monte Carlo integration technique and then parameterized from the fully reconstructed Monte Carlo samples of signal D^*D^* events. A detailed approach to the derivation and the formulas used for the acceptance moments calculation can be found in Appendix B; here, we describe the general principle of the Monte Carlo Integration estimation and the results determined from the fully reconstructed Monte Carlo.

6.3.1 Monte Carlo Integration Estimation

In general, a function, $h(x)$, has an expectation value defined as:

$$\langle h \rangle = \int_{\mathcal{V}} h(x) f(x) \, dx \quad (6.13)$$

where $f(x)$ is the sampling probability distribution function. Often a discrete set of data is used to estimate $\langle h \rangle$; this is given by the discrete sum over the sample

$\{x_i\}$:

$$\langle h \rangle_N \approx \frac{C}{N} \sum_{i=1}^{N} h(x_i), \tag{6.14}$$

where the N is the total number of events and $f(x)$ is typically normalized such that $C = 1$. The estimated variance in this estimate is given by

$$V(\langle h \rangle_N) = \frac{C^2}{N-1} \left[\frac{1}{N} \sum_{i=1}^{N} h(x_i)^2 - \left(\frac{1}{N} \sum_{i=1}^{N} h(x_i) \right)^2 \right]. \tag{6.15}$$

To apply this technique to the estimation of the acceptance moments, we first note that each moment is of the form:

$$I_\alpha(z) = \int g_\alpha(x, y) \epsilon(x, y, z) \mathrm{d}x \, \mathrm{d}y. \tag{6.16}$$

In this definition $g_\alpha(x, y)$ are the moment distributions, given by

$$
\begin{aligned}
g_0(x, y) &= \frac{3}{2\pi} y^2 \cos^2 x \\
g_\parallel(x, y) &= \frac{3}{4\pi} \sqrt{1 - y^2} \sin^2 x \\
g_\perp(x, y) &= \frac{3}{8\pi} \sqrt{1 - y^2}
\end{aligned}
$$

and $x = \phi_{\mathrm{tr}}$, $y = \cos(\theta_1)$, $z = \cos(\theta_{\mathrm{tr}})$, as seen from Equation 6.9. Since we will bin the Monte Carlo distribution in $\cos \theta_{\mathrm{tr}}$, the average value of the acceptance moment in bin k is defined as:

$$I_\alpha^k = \frac{1}{\Delta z} \int_{z_k}^{z_{k+1}} \mathrm{d}z \, I_\alpha(z) \tag{6.17}$$

Since we will use a sample of Monte Carlo with an underlying non-uniform generation pdf, the events used in the discrete sum estimation must be "de-weighted" by the same pdf. If, for example, the events were distributed evenly throughout the (x, y, z) space, then each event would contribute equally to the sum. At the time of this analysis, such Monte Carlo samples were not available and hence the

acceptance moments' calculation requires this event-weighting technique. Using this formulation the value of an acceptance moment I_α^k in bin k is given as:

$$I_\alpha^k = \frac{1}{N\Delta z} \sum_{\text{events in bin } k} \frac{g_\alpha(x_i, y_i)}{f(x_i, y_i, z_i)} \tag{6.18}$$

with an estimated uncertainty

$$\sigma(I_N^k) = \frac{1}{N\Delta z} \sqrt{\sum \left(\frac{g(x_i, y_i)}{f(x_i, y_i, z_i)}\right)^2 - \frac{1}{N}\left(\sum \frac{g(x_i, y_i)}{f(x_i, y_i, z_i)}\right)^2}. \tag{6.19}$$

Here, N is the total number of events generated in the sample, and $f(x, y, z)$ is the pdf used to generate the Monte Carlo events (see Equation 6.4). The efficiency/acceptance distribution has been taken into account via the sum over *selected* events in each bin (*i.e.* $\epsilon = 1$ for selected events and $\epsilon = 0$ otherwise). This formulation, therefore, provides a simple mechanism to evaluate the acceptance moments in each bin of z using a finite Monte Carlo sample. A more rigorous derivation of the above formulas is found in Appendix B.

6.3.2 Validation of the Acceptance Moments' Calculation

Since the procedure of calculating the acceptance moments involves both an estimation technique as well as a parameterization of the distributions, it is important to validate the results using toy[1] Monte Carlo samples. The dominate acceptance effect is derived from the slow pion efficiency as a function of the track's transverse momentum in the lab frame, p_T. We therefore generate a large sample of events using the probability distribution function of Equation 6.4 with known values of

---

[1]A "toy" Monte Carlo is defined as a sample produced by random number generation where the distributions follow the probability distribution functions of the fit. These samples are completely different than the "fully-reconstructed" Monte Carlo which utilizes the full simulation of the detector and reconstruction techniques.

$R_\perp$ and α and with a theoretical p_T distribution for the slow pion momenta. In this sample, the effects of detector resolution and tracking algorithms are ignored to isolate the effect of the p_T-dependent acceptance. For a given sample, a hard p_T cut[2] is applied to both soft pion tracks. The sample is used to calculate the acceptance moments and the parameterizations are fixed in the signal pdf (Equation 6.7). If the moments calculation is correct then the fitted value of $R_\perp$ should be the same as the generated value. In Figure 6.4 the effect of the acceptance correction is clearly demonstrated. As the p_T cut is increased, the fitted value of $R_\perp$ for the uncorrected signal pdf diverges from the generated value, while the acceptance moments calculation and parameterization corrects for the effect in every case. (Second order even polynomials were used to parameterize the moments' distributions; fourth order even polynomials were also used as a cross check – Figure 6.4 shows that the second order polynomials are sufficient parameterizations for these efficiency distributions.) Thus, it is clear that even for "harsh" estimations of the acceptance, the acceptance moments are able to correctly account for the effect.

6.3.3 Parameterization of the Acceptance Moments

Before the $R_\perp$ fit is performed on the data sample, the parameterization of the acceptance moments is determined from fully-reconstructed signal Monte Carlo samples. The sample used is generated with the amplitudes $(A_0, A_\parallel, A_\perp) = (0.74, 0.62, 0.25)$ and relative phases (Equation 6.4) set to zero (note that the acceptance moments are in principle independent of the amplitudes, see Equation 6.9). This particular set of amplitudes results in a generated value of $R_\perp = 0.0626$ and $\alpha = 0.179$. After requiring events to be in the signal region in m_{ES} and ΔE, about

---

[2]That is, the efficiency is set to zero for $p_T <$ cut value and 100% for $p_T >$ cut value.
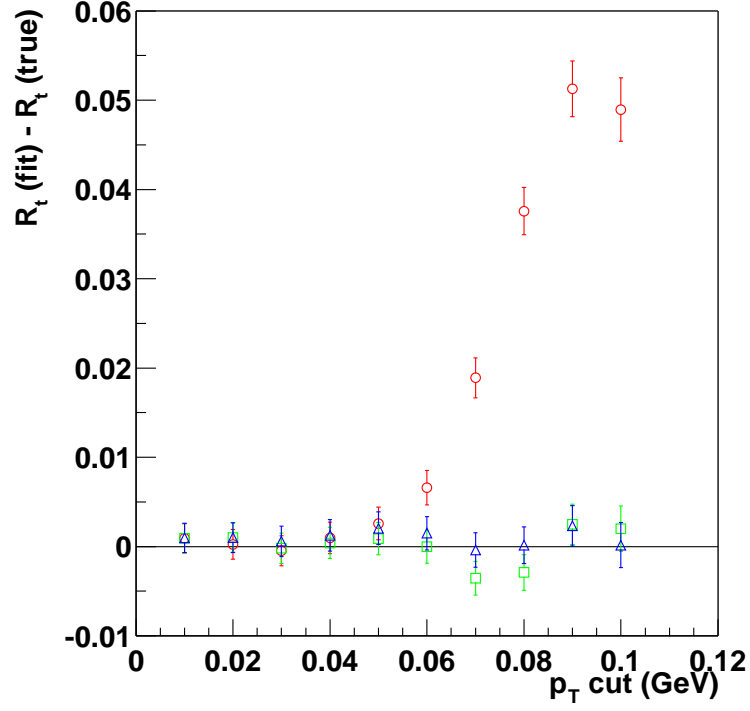
Figure 6.4: Toy Monte Carlo validation for the acceptance moments calculation. The difference between the generated and fitted value of $R_\perp$ is plotted as a function of the p_T cut of the sample. Each point represents one large statistics sample generated using the amplitudes $(A_0, A_\parallel, A_\perp) = (0.74, 0.62, 0.25)$ where the soft pions are required to pass the p_T cut. The circle points represent the fit results when the moments were *not* included in the pdf, and the squares are the fit results when the moment parameterizations are included. The triangles represent a cross check where fourth order even polynomials were used to parameterize the same acceptance moments' calculation.

$30,000$ events remain in the $B^0 \rightarrow D^{*+}D^{*-} \rightarrow D^0\pi^+ \overline{D}^0\pi^-$ modes, and about $23,000$ events remain in $B^0 \rightarrow D^{*+}D^{*-} \rightarrow D^0\pi^+ D^-\pi^0 + c.c.$ modes.

As mentioned in Section 6.2, due to the difference in acceptance for charged and neutral pions, the calculation is performed for each of the three cases of soft pion combinations for the final state. The calculated acceptance moments, using the Monte Carlo true values[3], are shown in Figure 6.5. Each distribution was fit to a second order even polynomial. The odd term in the parameterization is fixed to zero because it was found to be consistent with zero when included in the fit. Further, there is no physical reason why the acceptance should *not* be even in $\cos\theta_{\mathrm{tr}}$ (see Figure 6.2). Therefore, to minimize the number of parameters in the fit, we used even quadratic polynomials for describing the acceptance moments. As a secondary check, the moments were fit to fourth order polynomials (with no odd terms). While these parameterizations were not used in the analysis, they showed that the fourth order term was consistent with zero in every case, validating our choice to use quadratic fits to sufficiently describe the shape.

Once the acceptance moments were calculated, the functional form of the polynomial fits and the parameters obtained in the fits were fixed in the signal PDF (Equation 6.7). A systematic error is incurred from the parameterization of the acceptance moments, and is estimated by "smearing" the moment parameters within their errors. This is detailed in Section 6.7.3.

In Figure 6.5, a clear enhancement or peak is seen near the $\cos(\theta_{\mathrm{tr}}) = 0$ bin. While the errors correctly account for any fluctuation in the distribution, it is of interest to determine the cause of this effect. The primary cause for a large fluctuation in the moments calculation may be due to a few events that have

---

[3]True generated values were used instead of the reconstructed values because the definition of the acceptance moments is based on the true values of ϕ_{tr}, $\cos\theta_1$ and $\cos\theta_{\mathrm{tr}}$.
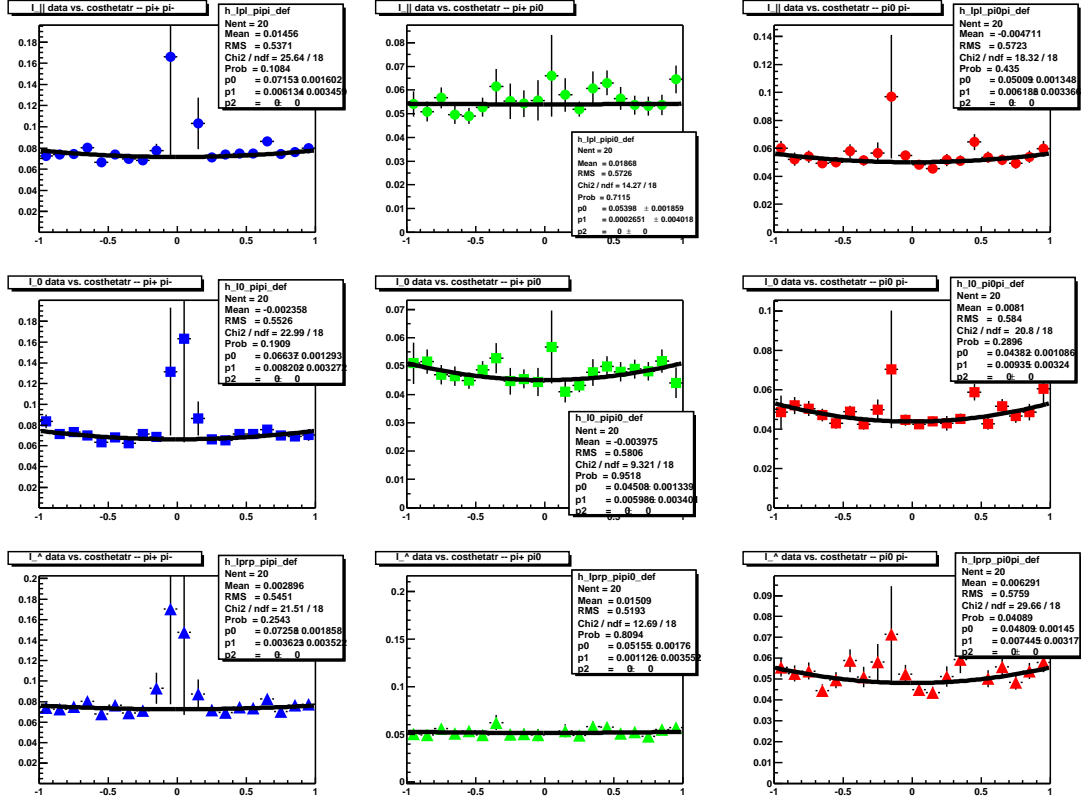
Figure 6.5: $I_{\parallel}$, I_0, $I_{\perp}$ for each pion type of the $B^0 \to D^{*+}D^{*-}$ decay. The solid lines represent fits to second order polynomials with the odd term fixed to zero.

extremely small probability, that is, the value of the pdf is close to zero and hence the corresponding weight in the moments calculation would be extremely large. In Figure 6.6, we can see that the majority (≈ 30000 events) of the Monte Carlo sample used to calculate the moments have weights < 500, while 2-4 events have weights > 2000 and are the reason for the "peaked" structure in Figure 6.5.
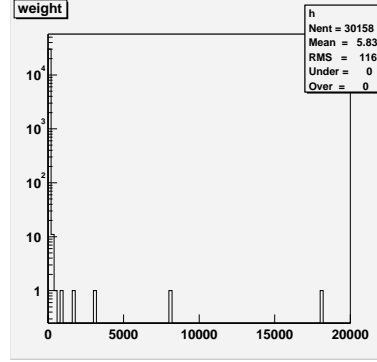
Figure 6.6: Histogram of the weight distribution $(1/f_{gen}(x, y, z))$ calculated from the same Monte Carlo sample used to determine the acceptance moments (shown are all events for modes where $D^{*+}D^{*-} \rightarrow (D^0\pi^+, \overline{D}^0\pi^-)$).

The effect, however, was also seen in a fully-reconstructed Monte Carlo sample where $(A_0, A_\parallel, A_\perp) = (1., 1., 1.)$. It is likely, then, that these few events are probing a region of the phase space that is extremely improbable. Calculating the probability of observing such events in a large sample confirmed that it was not a Monte Carlo generation error. It is likely, therefore, that the Monte Carlo integration technique itself is simply limited in its ability to handle these events. To avoid these effects, a sample of evenly distributed events over the entire phase space would serve as the best Monte Carlo sample for such calculations. In any case, excluding these few events from the sample does not bias the calculation of the acceptance moments. As seen in Figure 6.7, after cutting on the value of the
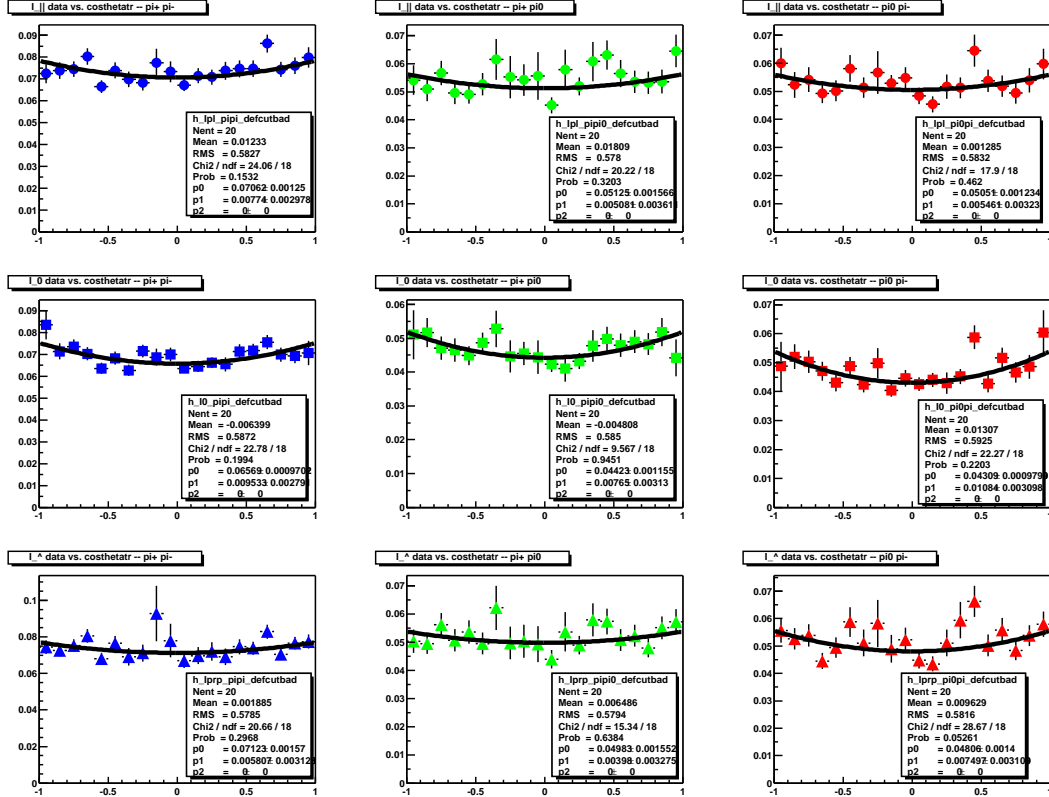
Figure 6.7: $I_\parallel$, I_0, $I_\perp$ for each pion type of the $B^0 \to D^{*+}D^{*-}$ decay. Here the event weight is required to be < 1000 for reasons detailed in the text. The solid lines represent fits to second order polynomials with the odd term fixed to zero. This version of the acceptance moments is used for $R_\perp$ fits to Monte Carlo.

event weight (at 1000) the parameterizations of the acceptance moments are consistent with the previous version (all events included). For the $R_\perp$ measurement, the parameterizations obtained in Figure 6.7 were used.

Finally, it is noted that the full Monte Carlo sample used to determine the acceptance moments does not take into account differences in efficiency between the D-decay sub-modes. To take any possible differences into account we weight events by their sub-mode branching ratios. This version of the acceptance moments is seen in Figure 6.8. It is important to note that the moments found in Figure 6.7 were used to perform Monte Carlo studies, since the sample is not distributed according to the submode branching ratios. When the data is fit, however, the Branching Ratio (BR) weighted version (Figure 6.8) of the acceptance moments is used.

As a cross check, we fit the truth values of the angles in the full Monte Carlo sample (all modes) using the non-BR weighted version of the acceptance moments – we find $R_\perp = 0.0634 \pm 0.0021$. Fitting the same sample using the BR-weighted version of the moments we find $R_\perp = 0.0619 \pm 0.0022$.
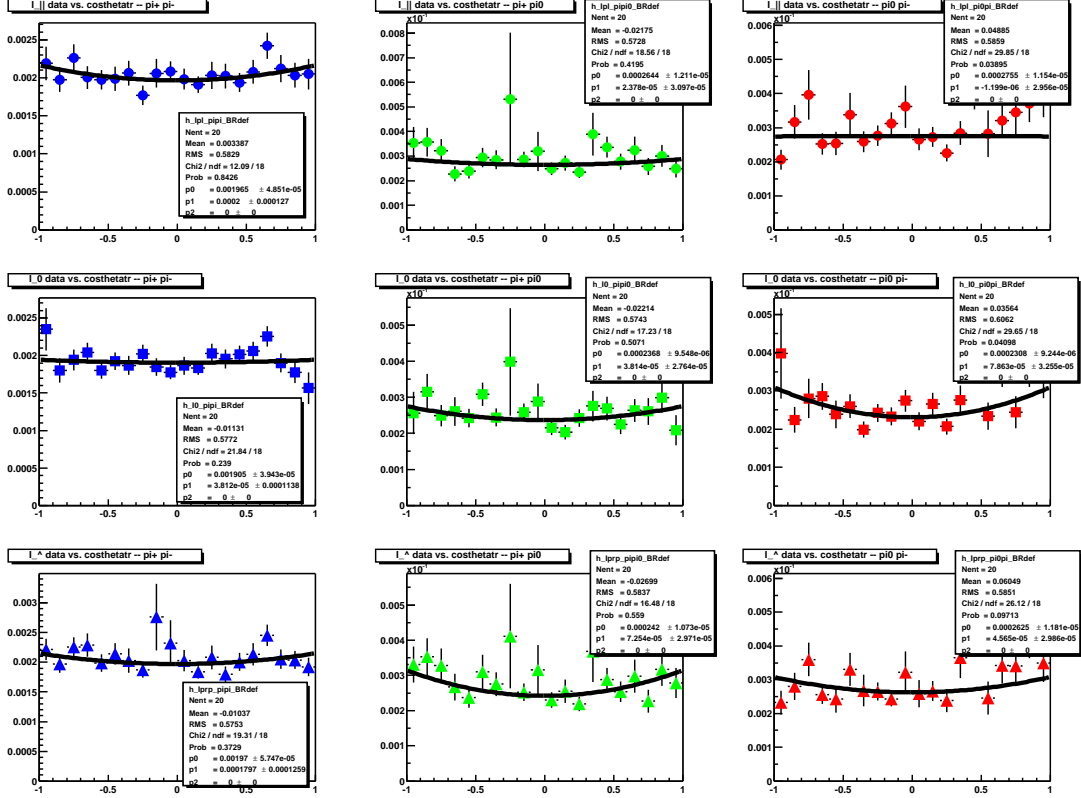
Figure 6.8: $I_\parallel$, I_0, $I_\perp$ for each pion type of decay. Here the event weight is required to be < 1000 for reasons detailed in the text. Events are weighted by their sub-mode Branching Ratios to mimic the data. The solid lines represent fits to second order polynomials with the odd term fixed to zero. This version of the acceptance moments are used only when fitting the data sample.

6.4 Angular Resolution

The angular resolution of θ_{tr} was studied using samples of Monte Carlo with amplitudes $(A_0, A_\parallel, A_\perp) = (0.74, 0.62, 0.25)$ The corresponding value of $R_\perp$ for this Monte Carlo is 0.0626. Shown in Figures 6.9 and 6.10, is the difference between the true value of θ_{tr} and it's reconstructed value. The line in each plot represents a fit to triple-Gaussian functions (which are not used in the analysis but demonstrate the structure of the distributions).

A RooPlot of "thetatr - thetatrMC"

sigma_tail = 0.0970 ± 0.0026
sigma_outl = 0.430 ± 0.010
sigma_core = 0.03501 ± 0.00083
mean_core = 0.00004 ± 0.00043
frac_tail = 0.402 ± 0.014
frac_core = 0.475 ± 0.017
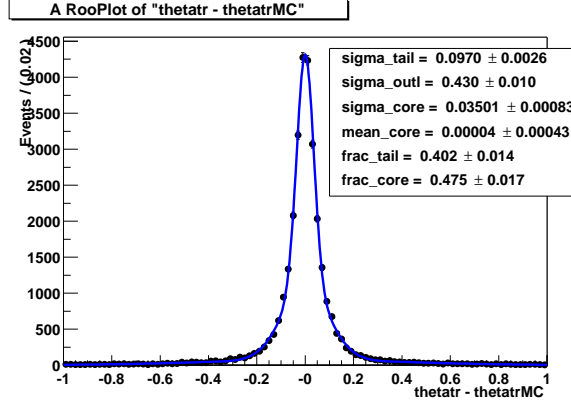
Figure 6.9: The difference between reconstructed and generated values of θ_{tr} for modes where $D^{*+}D^{*-} \to (D^0\pi^+, \overline{D}^0\pi^-)$. The line is a fit to a triple Gaussian where the means of the tail and out-lier Gaussians are fixed to zero. This fit is not used in the analysis, but is shown here for reference.

It is important to note the long tails in the distributions. In early studies, the central (core) Gaussian was used in the convolution with the $R_\perp$ pdf; this resulted in an underestimation of the effect of angular resolution. Also, the triple-Gaussian description was not sufficient and indicated that the events in the tail distributions warranted further study. The following sections detail the nature of the long tails in the resolution distribution and how they are parameterized for the $R_\perp$ fit.
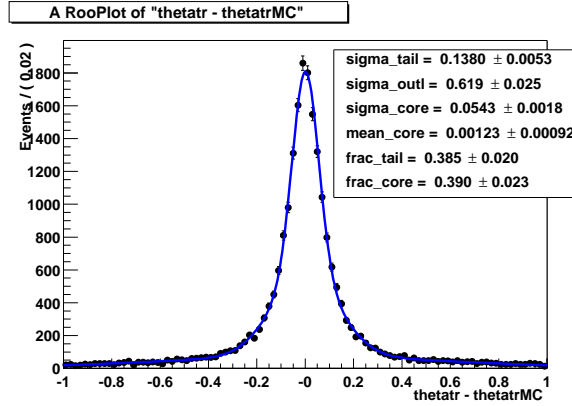
Figure 6.10: The difference between reconstructed and generated values of $\theta_{\rm tr}$ for modes where $D^{*+}D^{*-} \to (D^0\pi^+, D^-\pi^0)$+c.c. The line is a fit to a triple Gaussian where the means of the tail and out-lier Gaussians are fixed to zero. This fit is not used in the analysis, but is shown here for reference.

6.4.1 Mis-reconstruction of Soft Pions

In order to determine the cause of the long tails in the $\Delta\theta_{\rm tr}$ distribution we can, as a simple study, flag mis-reconstructed slow pions in the $D^{*+}D^{*-}$ decay. Using fully-reconstructed signal Monte Carlo, events are flagged when at least one soft pion in the reconstructed event has no Monte Carlo truth partner, or one of the soft pions is "wrong". That is, the truth associated candidate for that soft pion is *not* the correct pion in the truth decay chain.

As indicated in Figures 6.11 and 6.12, the dominant cause for the long tails in $\Delta\theta_{\rm tr}$ are due to mis-reconstructed soft pions. The value of $\theta_{\rm tr}$ for these events will therefore be completely uncorrelated with the true value of $\theta_{\rm tr}$. The projection plots of $\theta_{\rm tr}$ in bins of $\theta_{\rm tr}^{true}$ show that these events peak near $\pi/2$ and are independent of the true value of $\theta_{\rm tr}$. This type of distribution will therefore be parameterized separately in the resolution function used in the $R_\perp$ fit. (see Section 6.4.3).
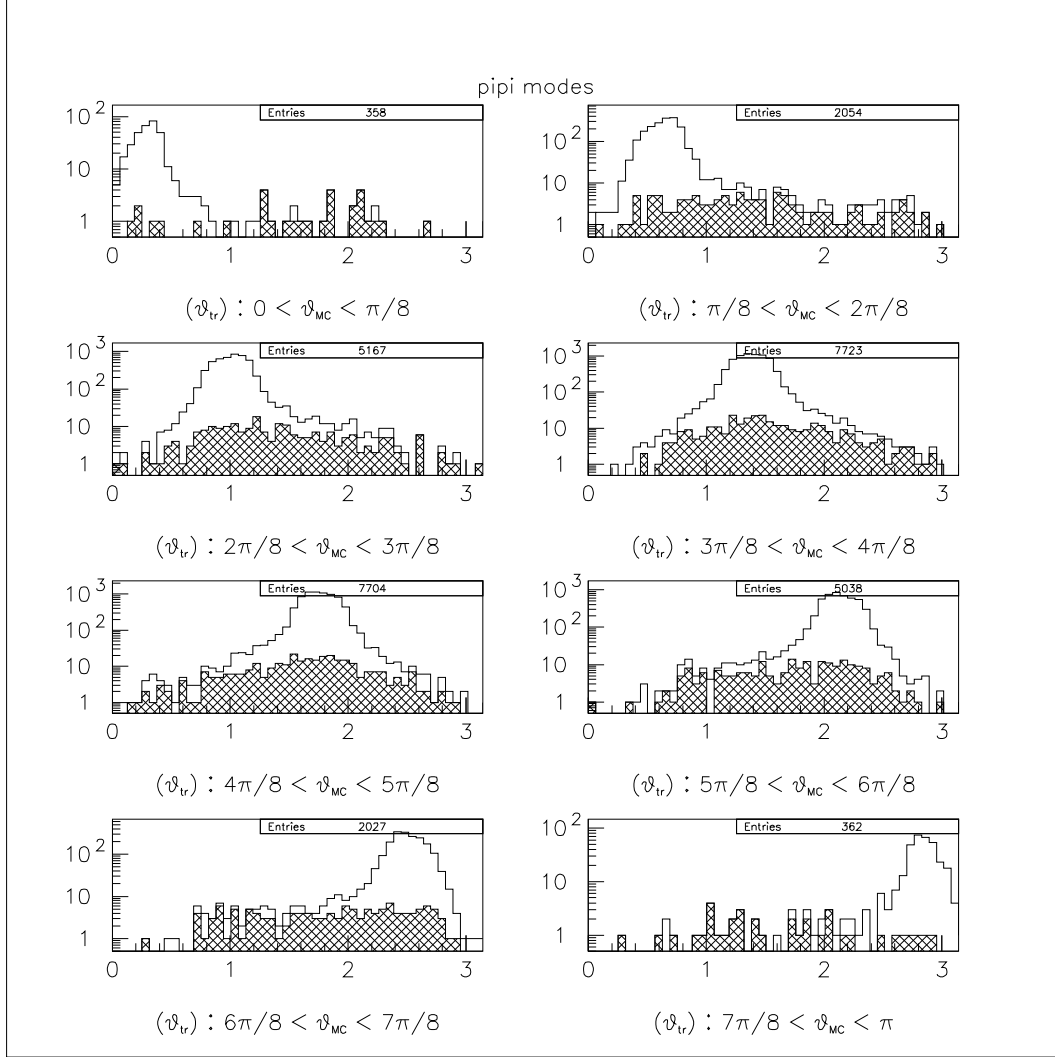
Figure 6.11: Displayed is the signal Monte Carlo distribution of $\theta_{\rm tr}$ in 8 bins of the true value of θ on a log scale. The solid line histogram is the projection of all signal events, while the filled histogram is the projection of events flagged as having a mis-reconstructed soft pion. These histograms contain events reconstructed in all modes where $D^{*+}D^{*-} \to (D^0\pi^+, \overline{D}^0\pi^-)$.

128
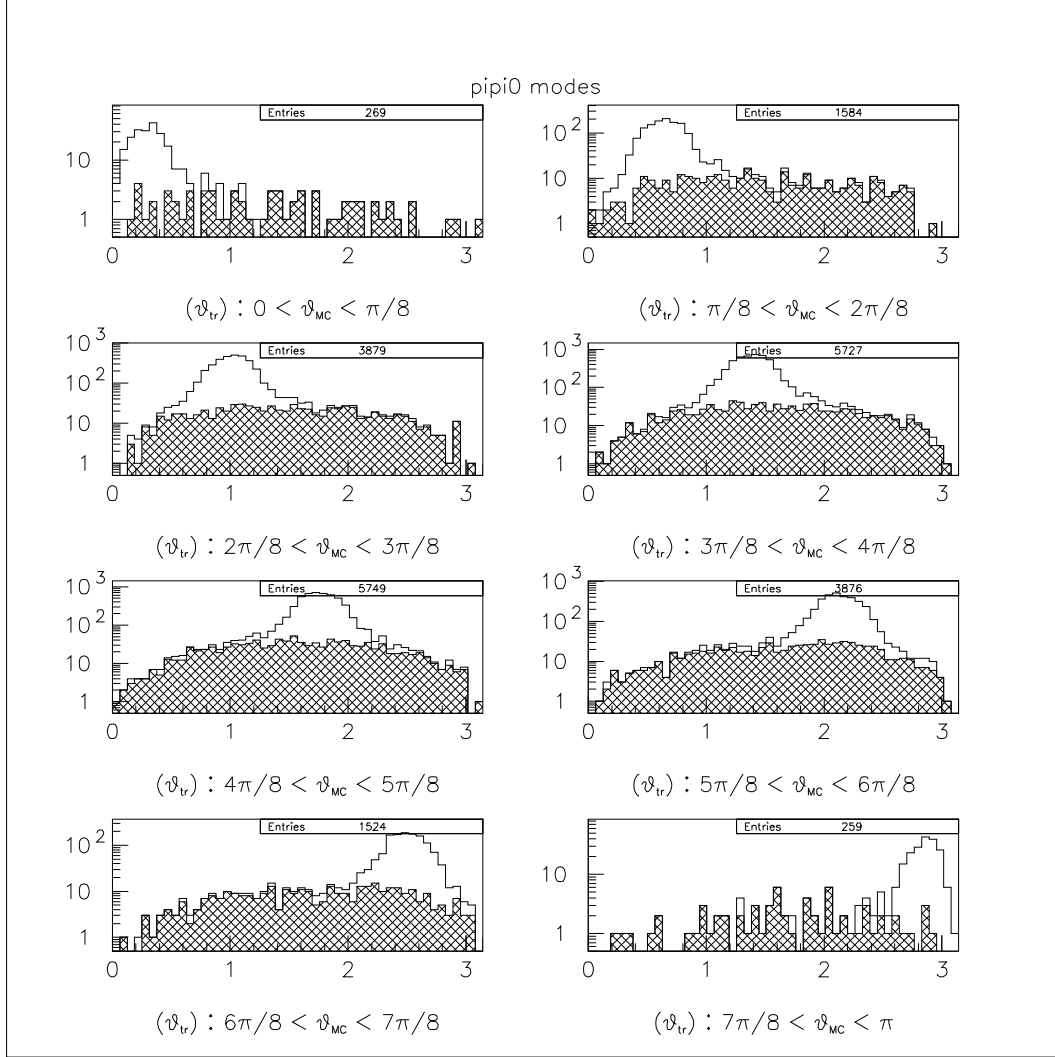
Figure 6.12: Displayed is the signal Monte Carlo distribution of θ_{tr} in 8 bins of the true value of θ on a log scale. The solid line histogram is the projection of all signal events, while the filled histogram is the projection of events flagged as having a mis-reconstructed soft pion. These histograms contain events reconstructed in all modes where $D^{*+}D^{*-} \rightarrow (D^0\pi^+, D^-\pi^0)$+c.c.

6.4.2 Derivation of the Convolved Signal PDF

The signal pdf (Equation 6.7) is a function of $z = \cos(\theta_{\mathrm{tr}})$, while the resolution function is found to be best parameterized as a function of θ_{tr}. Therefore, in order to correctly convolute the signal pdf with the resolution function, we must change the resolution pdf which is defined in terms of θ to one defined in terms of $z = \cos(\theta_{\mathrm{tr}})$. Let $f(\theta; \theta')$ represent the probability density for measuring θ given a known value of θ'. Since θ is defined on $(-\infty, \infty)$ and z is defined on $[-1, 1]$ then,

$$
\begin{aligned}
g(z; z') &= \sum \left| \frac{d\theta}{dz} \right| f(\theta; \theta') \\
&= \sum \left| \frac{1}{\sqrt{1 - z^2}} \right| f(\theta; \theta') \quad\quad (6.20)
\end{aligned}
$$

where the sum is over all values of θ that map on to the same value of z. So $\theta = \cos^{-1} z + 2n\pi$ and $\theta = -\cos^{-1} z + 2n\pi$ where n is any integer on $(-\infty, \infty)$. Let the un-convolved signal pdf be

$$
\frac{1}{\Gamma} \frac{d\Gamma}{d\cos\theta_{\mathrm{tr}}} = P(z)
$$

The convolved pdf is then

$$
\begin{aligned}
\tilde{P}(z) &= \int_{-1}^{1} P(z') g(z; z') dz' \\
&= \frac{1}{\sqrt{1 - z^2}} \sum_{n=-\infty}^{\infty} \int_{-1}^{1} dz' P(z') \left[f(\cos^{-1} z + 2n\pi; \cos^{-1} z') \right. \\
&\quad\quad\quad\quad \left. + f(-\cos^{-1} z + 2n\pi; \cos^{-1} z') \right] \quad (6.21)
\end{aligned}
$$

We now make a change of variables and it can be shown that

$$
\begin{aligned}
\tilde{P}(z) &= \frac{1}{\sqrt{1 - z^2}} \sum_{n=-\infty}^{\infty} \left[\int_{-2n\pi}^{(-2n+1)\pi} d\theta' \sin\theta' P(\cos\theta') f(\theta; \theta') \right. \\
&\quad\quad\quad\quad \left. - \int_{(-2n-1)\pi}^{-2n\pi} d\theta' \sin\theta' P(\cos\theta') f(\theta; \theta') \right] \\
&= \frac{1}{\sqrt{1 - z^2}} \sum_{n=-\infty}^{\infty} (-1)^n \int_{n\pi}^{(n+1)\pi} d\theta' \sin\theta' P(\cos\theta') f(\theta; \theta') \quad (6.22)
\end{aligned}
$$

130

We note here that f is the resolution function in θ and $z = \cos(\theta_{\text{tr}})$; that is, we now have a means of convolving the pdf (which is function of z and limited domain) with a resolution function that is a function of θ and of infinite domain. For this analysis we use a Gaussian (or sum of Gaussians) for the resolution function. If $\sigma_{Gauss} \ll \pi$ (as is the case for the θ_{tr} resolution) then we need only keep the terms in the sum with limits between $[-\pi, 2\pi]$ (ie. $n = -1, 0, 1$ in Equation 6.22). To observe the effect of resolution and the convoluted signal pdf, examples are shown in Figure 6.13; the same $R_\perp$ cases are to be compared with those in Figure 6.3.
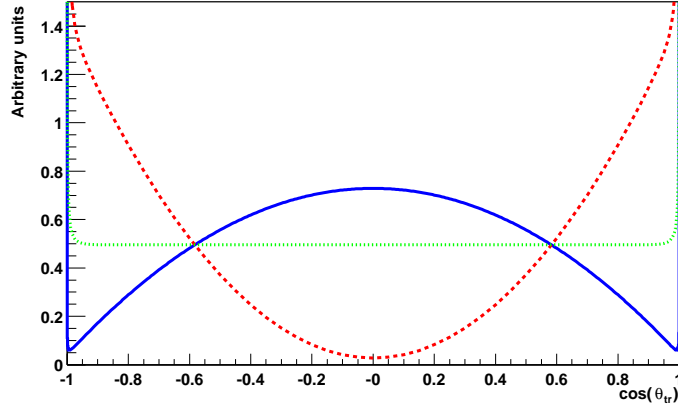
Figure 6.13: Theoretical $\cos \theta_{\text{tr}}$ distributions after Equation 6.6 is convolved with a resolution function described in the text (see Equation 6.22). For $R_\perp = 0.$, the solid line (blue curve) shows the resulting distribution. The dashed line (red curve) is the corresponding distribution for $R_\perp = 1$. The dotted line (green curve) is the expected distribution for $R_\perp = 1/3$.

As mentioned in Section 6.4.1 we have a component of the θ_{tr} resolution that is due to mis-reconstructed signal events, such that θ_{reco} is uncorrelated with θ_{true}. One possible parameterization of this component of the resolution function is a

truncated Gaussian centered on $\theta_{tr} = \pi/2$:

$$f(\theta; \theta') = \frac{1}{\sqrt{2\pi}\sigma} \frac{1}{erf(\frac{\pi}{2\sqrt{2}\sigma})} e^{-\frac{(\theta - \pi/2)^2}{2\sigma^2}} \tag{6.23}$$

Note that this function is independent of θ', and is normalized on the domain $[0, \pi]$. Similar to Equation 6.21, we have:

$$
\begin{aligned}
\tilde{P}(z) &= \frac{1}{\sqrt{1-z^2}} \int_{-1}^{1} dz' P(z') f(\cos^{-1} z; \cos^{-1} z') \\
&= \frac{1}{\sqrt{1-z^2}} \frac{1}{\sqrt{2\pi}\sigma} \frac{1}{erf(\frac{\pi}{2\sqrt{2}\sigma})} e^{-\frac{(\cos^{-1} z - \pi/2)^2}{2\sigma^2}} \int_{-1}^{1} dz' P(z') \\
&= \frac{1}{\sqrt{1-z^2}} \frac{1}{\sqrt{2\pi}\sigma} \frac{1}{erf(\frac{\pi}{2\sqrt{2}\sigma})} e^{-\frac{(\cos^{-1} z - \pi/2)^2}{2\sigma^2}}
\end{aligned}
\tag{6.24}
$$

Since the resolution function does not depend on θ', it comes out of the integral and we are left with the integral of the original pdf, which is equal to 1 (assuming it is normalized). Also note that there is no sum as there is in Equation 6.21 because here the resolution function is defined and normalized on the domain $[0, \pi]$.

6.4.3   Fit Results for the Resolution Function

Using the signal Monte Carlo samples generated with $R_\perp = 0.0626$ we fit for the resolution function in θ_{tr} for the $(D^0\pi^+, \overline{D}^0\pi^-)$ modes and $(D^0\pi, D^-\pi^0)$ and $(D^+\pi^0, \overline{D}^0\pi^-)$ modes separately.[4] We use a double-Gaussian to describe the correlated signal, and the truncated Gaussian centered on $\pi/2$ (Equation 6.23) to describe the uncorrelated, mis-reconstructed signal component. The resolution

---

[4]As with the calculation of the acceptance moments we separate the $(D^0\pi, D^-\pi^0)$ and $(D^+\pi^0, \overline{D}^0\pi^-)$ modes because of the definition of the angle θ_{tr} (defined wrt the D^{*+}). This allows for possible differences in resolution parameters between the two sets of modes, which in fact is seen.

function, f, is written as:

$$f(\theta; \theta') = (1 - f_{misReco})G(\theta; \theta') + f_{misReco}F_{misReco}(\theta; \pi/2, \sigma_{misReco})$$

$$G(\theta; \theta') = (1 - f_{wide})Gauss(\theta; \theta', \sigma_{core}) + f_{wide}Gauss(\theta; \theta', \sigma_{wide}) \quad (6.25)$$

where the function $Gauss(x; x', \sigma)$ is a Gaussian with mean x' and r.m.s σ and is defined for x on $(-\infty, \infty)$ and the function $F_{misReco}$ is the resolution function described in Equation 6.23.

The results of the fit are displayed graphically in Figures 6.15, 6.16, and 6.17. In these plots, the signal Monte Carlo samples are projected into bins of θ_{true} and the projection of the resolution function in the same range of θ_{true} is displayed with the data. To get a general "goodness" of fit for the resulting parameters, the function normalization was allowed to float in each plot, and with all other parameters fixed, a fit was performed to determine a χ^2. The χ^2/ndf can be used as a cross check for how well the curve fits the distribution.

One can also observe the effect of θ_{tr} resolution and mis-reconstructed events on the $\cos\theta_{tr}$ distribution. In Figure 6.14 the theoretical distribution (Equation 6.6) is compared with the more realistic signal pdf where the resolution effects and mis-reconstructed events' parameterizations have been included. For the latter, the parameterizations determined from the Monte Carlo shown in Figure 6.16 are used. In this case, we have shown that for a mostly CP-even ($R_\perp$ close to zero) final state, the resolution and misreconstruction effects will cause events to shift from the central $\cos\theta_{tr}$ bins to the outer bins. Physically, the soft π^+ (and D^0) tracks are more poorly determined; therefore, while they would tend to lie more in the plane ($\cos\theta_{tr} \approx 0$) in the theoretical case, the resolution and misreconstruction cause the tracks to move away from $\cos\theta_{tr} \approx 0$. The opposite would be true for $R_\perp \approx 1$.

133
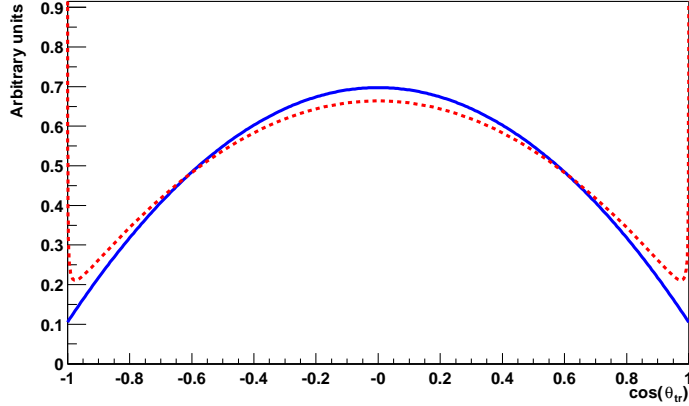
Figure 6.14: The $\cos\theta_{\text{tr}}$ distribution for the theoretical case (Equation 6.6) where $R_\perp = 0.07$ and no resolution, acceptance or misreconstruction effects (solid line). The dashed line shows the more realistic case where both θ_{tr} resolution and misreconstruction of soft pions are taken into account via the parameterizations described in the text.

6.4.4 Validation of the Resolution Function Parameterization

In order to validate the convolution procedure, we generated toy Monte Carlo events with a known resolution in θ_{tr}. For each sample the acceptance was perfect (100% efficiency) and the fit was performed without the acceptance moments in the signal pdf. For a given set of amplitudes, 100k events were generated and the angle θ_{tr} was smeared randomly based on a triple Gaussian parameterization. The same parameters were fixed in the fit for $R_\perp$. The difference between the generated and fitted values of $R_\perp$ are shown in Figure 6.18 and show that the convolution technique is working correctly and no biases are observed.

To test the use of the mis-reconstructed events' parameterization a 200k event sample was generated using the truncated Gaussian distribution (Equation 6.23)

134
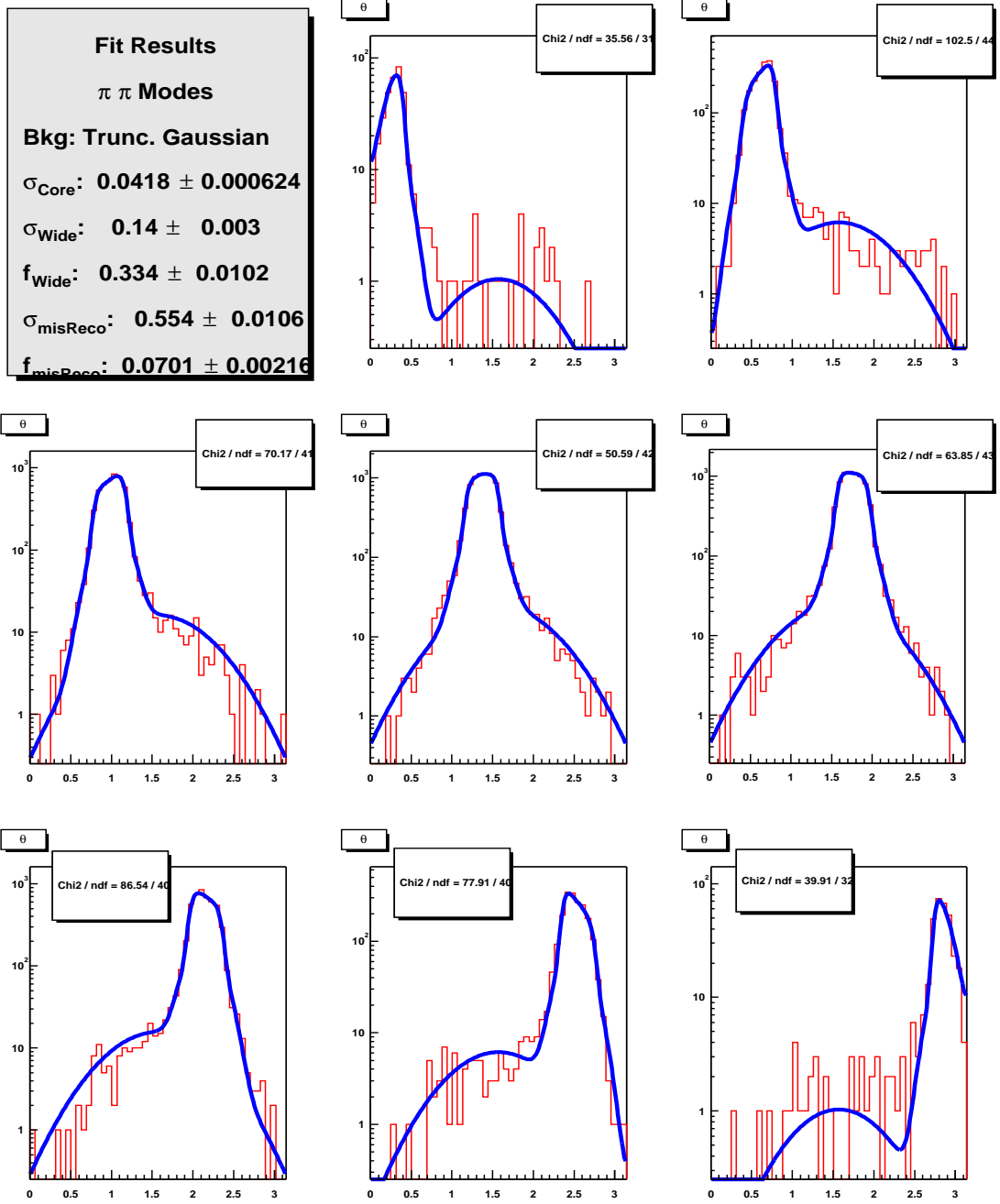
Figure 6.15: Plotted is the result of the resolution fit projected in bins of $\theta_{\mathrm{tr}}^{true}$. The parameters of the fit function (determined from a fit to all events) are fixed and the normalization (only) is allowed to float in each plot as a measure of the "goodness" of the fit. These histograms contain events reconstructed in all modes where $D^*D^* \to (D^0\pi, D^0\pi)$.
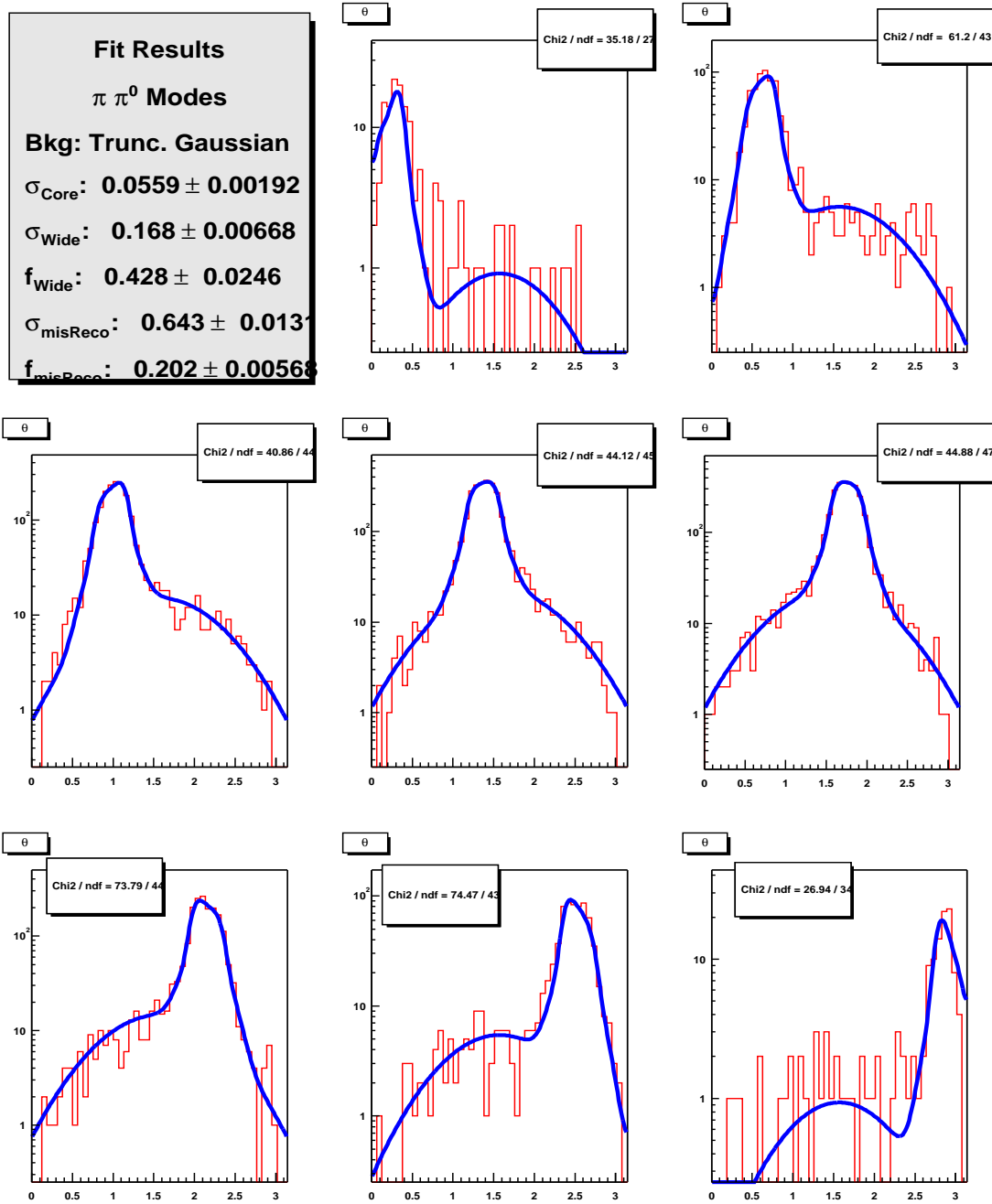
Figure 6.16: Plotted is the result of the resolution fit projected in bins of $\theta_{\mathrm{tr}}^{true}$. The parameters of the fit function (determined from a fit to all events in the modes where $D^*D^* \to (D^0\pi, D^-\pi^0)$) are fixed and the normalization (only) is allowed to float in each plot as a measure of the "goodness" of the fit.
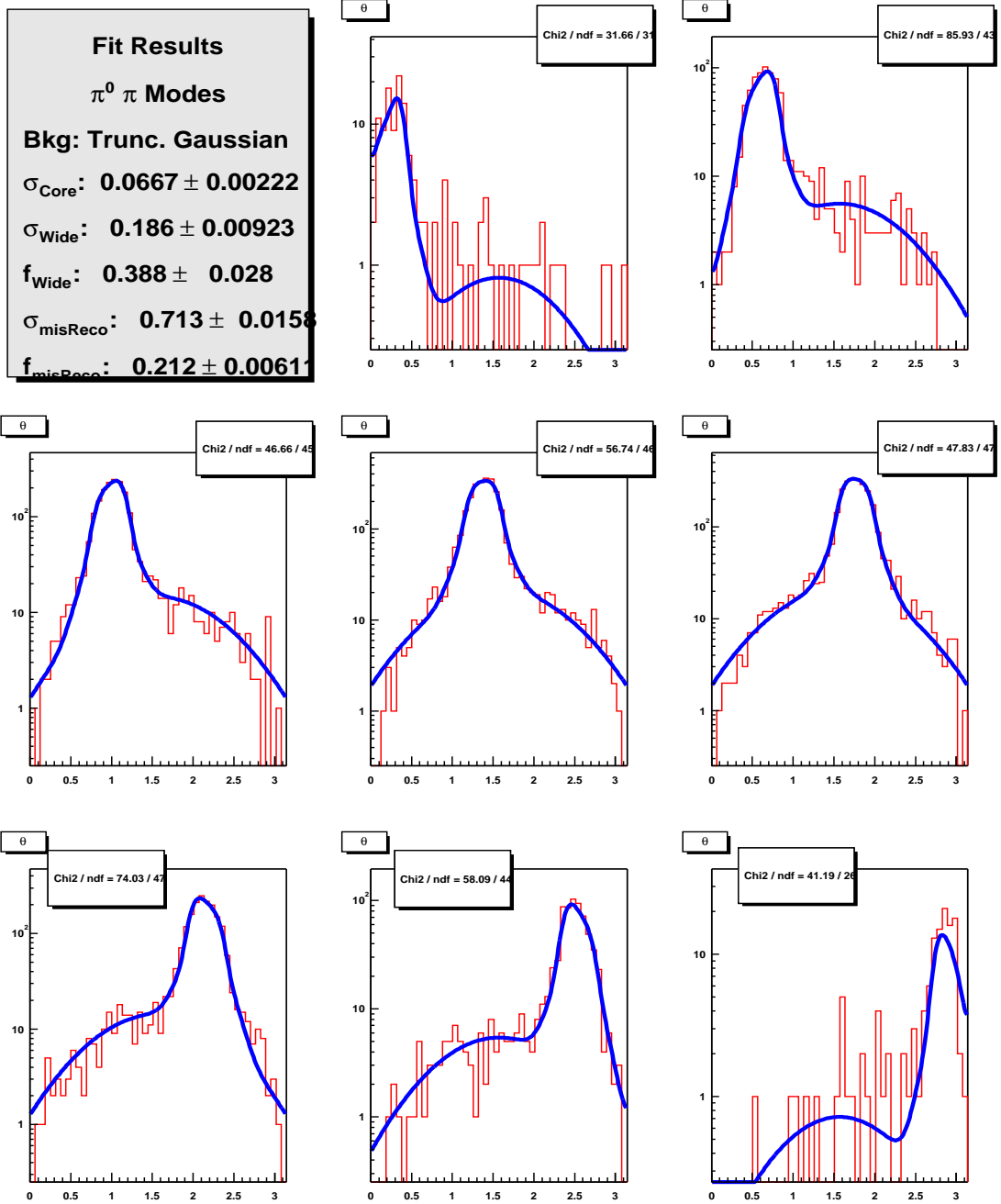
Figure 6.17: Plotted is the result of the resolution fit projected in bins of $\theta_{\mathrm{tr}}^{true}$. The parameters of the fit function (determined from a fit to all events in the modes where $D^*D^* \to (D^+\pi^0, \overline{D}^0\pi^-))$ are fixed and the normalization (only) is allowed to float in each plot as a measure of the "goodness" of the fit.
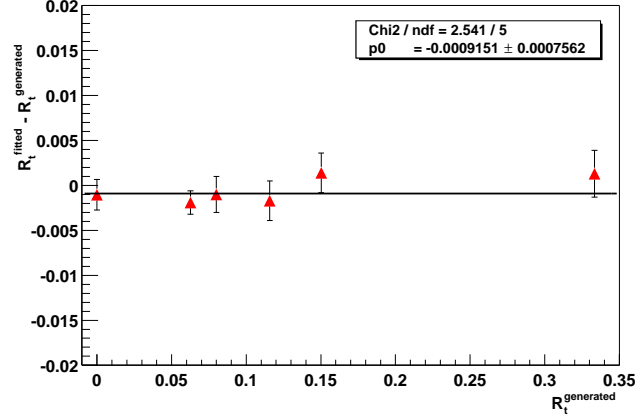
Figure 6.18: Toy Monte Carlo test for convolution technique. The difference between the generated and fitted value of $R_\perp$ is plotted as a function of generated value. Each point represents one large statistics sample where the true value of $\theta_{\rm tr}$ was smeared with a triple Gaussian distribution.

for the fraction of mis-reconstructed events and a double Gaussian smearing of $\theta_{\rm tr}$ for the remainder of the events. This type of smearing would more closely resemble that seen in the full Monte Carlo (Figure 6.16). This toy sample was generated with the same amplitudes as the full Monte Carlo sample ($R_\perp = 0.0626$) and with perfect acceptance. To observe the effect of not correcting for resolution, we first fit the sample without the convolution and mis-reconstruction component of the signal pdf. The fit results in $R_\perp = 0.1063 \pm 0.0013$. To see the effect of the convolution and validate the technique, we fix the parameters of the resolution function to the same values used in the generation of the sample; this resulted in a fit value of $R_\perp = 0.0636 \pm 0.0017$, which is consistent with the generated value.

6.5 Monte Carlo Studies

6.5.1 Likelihood Fit Validation

Besides the toy Monte Carlo tests used to validate the acceptance moments calculation and the convolution technique it is also necessary to validate the performance of the $R_\perp$ likelihood fit. A toy Monte Carlo sample of $150,000$ events was generated with $R_\perp = 0.1$ and $\alpha = 0$. The background was assumed to be flat in $\cos(\theta_{\rm tr})$ and hence b_2 was set to zero. The relative fractions of event types: $\pi\pi$ vs. $\pi\pi^0$ vs. $\pi^0\pi$, were set approximately to that expected in the data. The signal fractions for the three event types $f_{sig}^{\pi\pi}$, $f_{sig}^{\pi\pi^0}$, and $f_{sig}^{\pi^0\pi}$ were also set equal to values expected in data: $0.5, 0.25$, and 0.25 respectively. Also, for the parameters of the signal Gaussian in $m_{\rm ES}$, we set $\sigma m_{\rm ES} = 2.6\,\mathrm{MeV}$, and fixed the mean to $5.2794\,\mathrm{GeV}$; the ARGUS function parameter is set to -30., since that corresponds to the expected shape in data. The parameters of the acceptance moments were set equal to the values determined in Section 6.3.3 and the parameters of the convolution were set to the values determined in Section 6.4.3. The result of the fit is shown in Fig. 6.19, and is clearly consistent with input values. The fitted value of $R_\perp$ was 0.1014 ± 0.0029.

Another validation of the likelihood fit was performed by generating multiple Monte Carlo samples each with the same number of events expected to be seen in the data. Here, a distribution of fitted $R_\perp$ values is expected, the width of which is related to the statistical size of the sample. We generated 1000 samples of 350 events each (approximate number of events expected in the $m_{\rm ES}$ distribution in $82.\,\mathrm{fb}^{-1}$). Each sample was generated with $R_\perp = 0.1, \alpha = 0.$ and flat background parameters. Again, all relative fractions of signal events were distributed as expected in the data. The distribution of fitted values of $R_\perp$ is shown in Figure 6.20.

A RooPlot of "cos(θ_{tr})"

R_t = 0.1014 ± 0.0029
argPar = -29.889 ± 0.45
b2 = 0.010 ± 0.012
mbMean = 5.279410 ± 0.000012
mbSigFrac_pi0pi = 0.2561 ± 0.0034
mbSigFrac_pipi = 0.5005 ± 0.0019
mbSigFrac_pipi0 = 0.2497 ± 0.0034
mbWidth = 0.002612 ± 0.000011
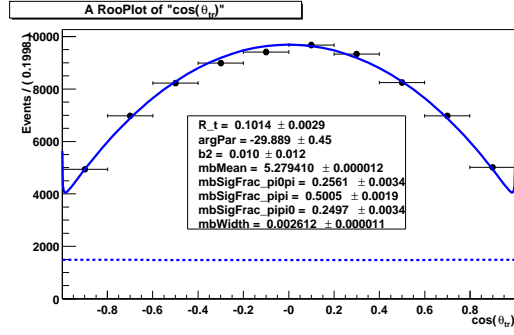
Figure 6.19: Toy Monte Carlo test: one experiment of 150k events generated with the probability density function described in the text. Generated values for the events were $R_\perp = 0.1$, $\alpha = 0$ and flat background parameters. The relative fractions of events as well as signal fractions for the $\pi\pi$, $\pi\pi^0$, $\pi^0\pi$ event types were generated with values close to that expected from data. The data points shown are the events generated projected into the region $m_{ES} > 5.27\,\text{GeV}$; the solid is the projection of the full PDF in the same region. The dashed line represents the contribution to the total pdf from the background.

The far left plot shows the distribution is centered on $R_\perp = 0.1$, as expected. The center plot displays the corresponding distribution of errors from the fits; it implies that our expected error from the data fit should be approximately 0.06. The far right plot shows the pull distribution, and the fitted Gaussian has a width consistent with one and mean consistent with zero, indicating that the fitted values and corresponding errors are appropriate.

A small number of experiments return a fitted value of $R_\perp = 0.$ and the corresponding error is somewhat unrealistic. These cases are seen as a secondary bump or tail in the error distribution and a slightly overpopulated zero bin in the $R_\perp$ distribution of Figure 6.20. These entries correspond to cases where the fit approaches the boundary of the physical limit of $R_\perp$; the minimum of the log(*likelihood*) may

140
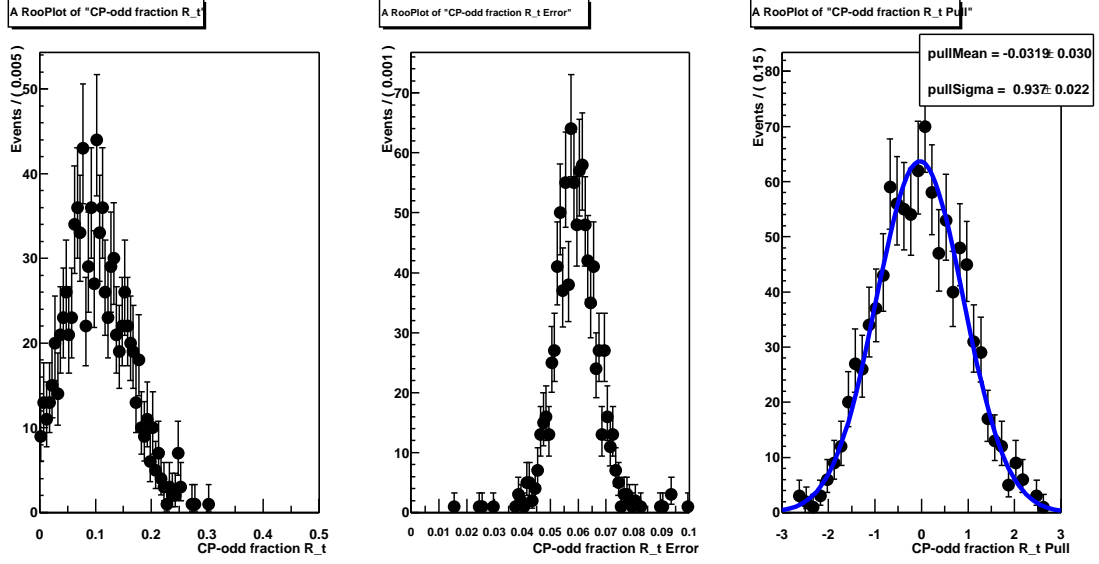
Figure 6.20: Toy Monte Carlo test: 1000 experiments of 350 events each generated with the probability density function described in the text. Generated values for the events were $R_\perp = 0.1$, $\alpha = 0$ and flat background parameters. The relative fractions of events as well as signal fractions for the $\pi\pi$, $\pi\pi^0$, $\pi^0\pi$ event types were generated with values close to that expected from data. The left plot shows the distribution of fitted $R_\perp$ values. The center plot shows the distribution of errors on $R_\perp$, and the far right plot shows the pull distribution with a Gaussian fit.

141

lie below $R_\perp = 0$, but the pdf is undefined for negative values and the returned error is not well determined. If the fit to the data is similar to these cases, then a different technique would be required in order to set an upper bound on $R_\perp$.

As with all likelihood fits, it is important to test the full range of the fit parameters to determine if any bias exists for certain values of the parameter fitted. Shown in Figure 6.21 is the linearity test for the $R_\perp$ fit. The left plot shows the fitted value of $R_\perp$ versus the input generated value; each point represents one 150k event experiment with signal fractions set to 0.50, 0.25, 0.25 for $f_{sig}^{\pi\pi}$, $f_{sig}^{\pi\pi^0}$, $f_{sig}^{\pi^0\pi}$ respectively, and flat background parameters. The plot on the right show the difference between fitted and generated values versus the generated value. No bias is seen.
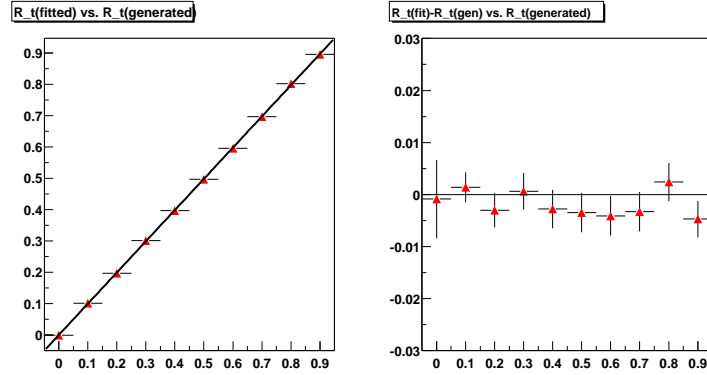
Figure 6.21: Linearity test: The left plot shows the fitted value of $R_\perp$ versus the input generated value for each 150k event experiment of toy Monte Carlo. The plot on the right shows the difference between fitted and generated values versus the generated value of $R_\perp$. The line is drawn at zero to show any possible bias in the results.

6.5.2 Validation from Fully-Reconstructed Monte Carlo

As mentioned in both Section 6.3.3 and Section 6.4 the sample of fully recon-structed Monte Carlo with amplitudes $(A_0, A_{\parallel}, A_{\perp}) = (0.74, 0.62, 0.25)$ was used to calculate the acceptance moments as well as the resolution function parameters. For this sample approximately $40 - 50$k events were generated for each $D^{*+}D^{*-}$ sub-decay mode. As a test of the signal $R_{\perp}$ pdf (with acceptance moments' pa-rameterizations and resolution function parameters fixed appropriately) we fit this sample with the expectation of obtaining the generated value of $R_{\perp}$. Fitting to the modes where $D^*D^* \to (D^0\pi^+, \overline{D}^0\pi^-)$, we find $R_{\perp} = 0.0525 \pm 0.0033$. Fitting to the modes where $D^*D^* \to (D^0\pi^+, D^-\pi^0)$+c.c., we find $R_{\perp} = 0.0508 \pm 0.0050$. Both results are slightly more than 2 standard deviations from the generated value (0.0626).

Also, as an additional test of the acceptance moments calculation, we can fit the truth values of $\theta_{\rm tr}$ in the full Monte Carlo sample; here, the effect of the resolution on $\theta_{\rm tr}$ is removed and the acceptance correction is tested directly. For the $(D^0\pi^+, \overline{D}^0\pi^-)$ modes, $R_{\perp} = 0.0630 \pm 0.0029$; for the $(D^0\pi^+, D^-\pi^0)$+c.c. modes $R_{\perp} = 0.0604 \pm 0.0033$.

A sample of Monte Carlo with amplitudes $(A_{\parallel}, A_0, A_{\perp}) = (1., 1., 1.)$ was also used to validate the fitting procedure. Here, we are using the parameters for the acceptance moments and the resolution function determined from a different sample of Monte Carlo (described above); thus, the fit tests the independence of the acceptance correction and resolution function from the decay amplitudes of the sample. This sample has a corresponding generated value of $R_{\perp} = 1/3$ and is flat in the $\cos(\theta_{\rm tr})$ distribution. The events decay in two modes only: $D^* \to D^0\pi$, $D^0 \to K\pi$ for both D^* mesons or one D^* meson decaying in $D^0\pi$, $D^0 \to K\pi$ and the other $D^* \to D^{\pm}\pi^0$, $D^{\pm} \to K\pi\pi$. Out of the 98k events generated in

each decay chain we observe 28110 reconstructed $D^{*+}D^{*-}$ events. The fit to these events yielded a fitted value of 0.3312 ± 0.0051, which is less than one σ from the generated value. The projection of the pdf onto data points in the region $m_{ES} > 5.27\,\mathrm{GeV}$ is shown in Figure 6.22.
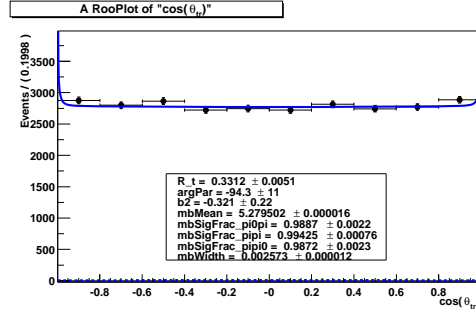
Figure 6.22: Fully reconstructed simulated $D^{*+}D^{*-}$ events used to test the $R_\perp$ fit. The events were generated with $R_\perp = 1/3$. Shown here is the projection of events in the region $m_{\mathrm{ES}} > 5.27\,\mathrm{GeV}$ and the solid line is the projection of the full pdf in the same region. The dotted line represents the background component of the pdf in the same region.

6.6 Fit Results

After all selection criteria are applied to the full data set $(81.8\,\mathrm{fb}^{-1})$, the unbinned maximum likelihood fit is performed on the events in the m_{ES} distribution.
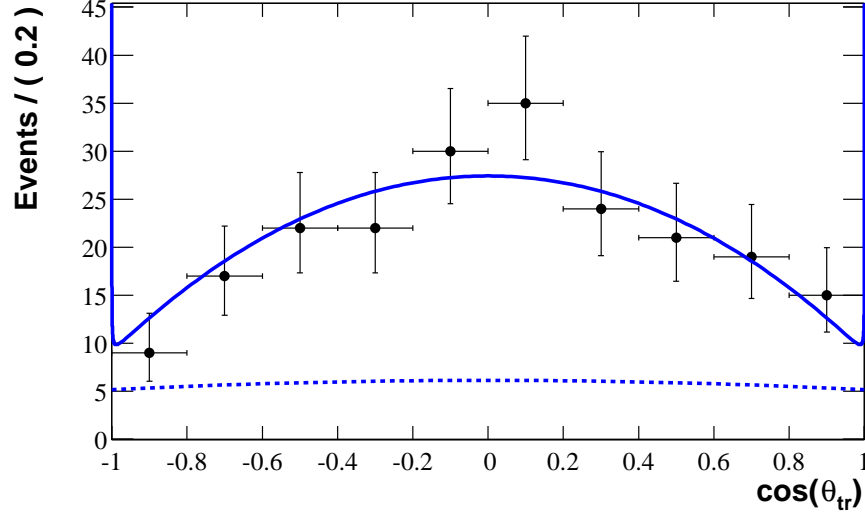
Figure 6.23: Likelihood fit result to the $\cos(\theta_{\mathrm{tr}})$ distribution of the 499 $D^{*+}D^{*-}$ events seen in the full data set. The data points shown are from the region $m_{\mathrm{ES}} > 5.27\,\mathrm{GeV}$ and the solid line is the projection of the fit result in the same region. The dotted line represents the component of the pdf for the background.

The result of the fit is

$$R_{\perp} = 0.063 \pm 0.055(\mathrm{stat}) \tag{6.26}$$

(the corresponding asymmetric errors are +0.0524, -0.0583). The background parameter is consistent with a flat distribution ($b_2 = -0.16 \pm 0.17$). A projection of the data in the region $m_{\mathrm{ES}} > 5.27\,\mathrm{GeV}$ is shown in Figure 6.23; the projection of the likelihood fit in the same region is superimposed. The result of the fit to the m_{ES} distribution is shown in Figure 6.24.
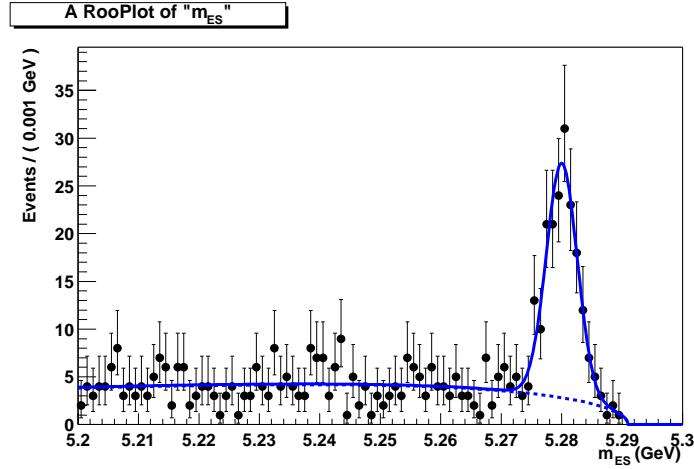
Figure 6.24: Likelihood fit result to the m_{ES} distribution of the 499 $D^{*+}D^{*-}$ events seen in the full data set. The dotted line in the signal region represents the ARGUS component of the pdf for the background in that region.

A projection of the data in the region $m_{\mathrm{ES}} < 5.27\,\mathrm{GeV}$ is shown in Figure 6.25. The line shown is the projection of the pdf in the same region, normalized to the number events in that region. This shows the m_{ES} sideband distribution and confirms the expectation of a flat distribution.

As a measure of the goodness of fit, a set of 500 toy experiments are generated using the same parameters and signal yields seen in the above fit results. The generated value of $R_{\perp}$ was set to 0.08 in each toy experiment. The value of $\log(likelihood)$ is plotted for each experiment in Figure 6.26, and confirms that the value obtained in the fit to the data (-441.8) is reasonable. Figure 6.27 shows the distribution of error values obtained in the fits to the toy data. The value obtained in the fit to the data (0.055) is clearly validated by this distribution.

146
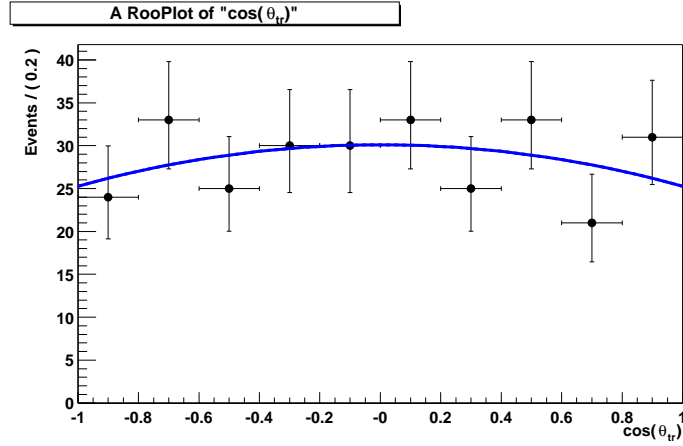
Figure 6.25: Likelihood fit result to the $\cos(\theta_{\mathrm{tr}})$ distribution of the 499 $D^{*+}D^{*-}$ events seen in the data. The data points shown are from the region $m_{\mathrm{ES}} < 5.27\,\mathrm{GeV}$ and the solid line is the projection of the fit result in the same region (primarily the background component).
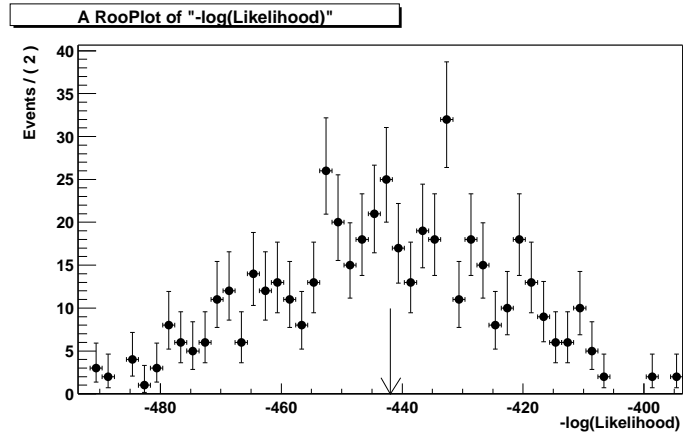
Figure 6.26: Results of the toy Monte Carlo study of the distribution of the $\log(likelihood)$ in 500 simulated experiments with 499 events each. The parameters of the generated data were set to those seen in the data. The result of the fit to data had $\log(likelihood) = -441.8$ and is indicated by the arrow.

147
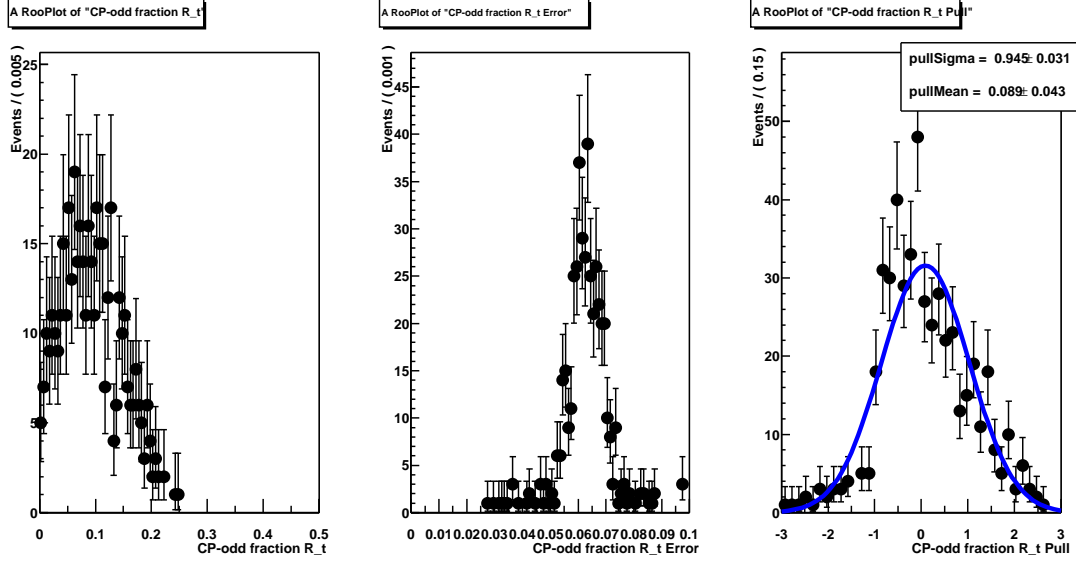
Figure 6.27: Results of the toy Monte Carlo study with 500 experiments, each with the same number of events and parameters seen in the data. The left plot shows the distribution of fitted $R_\perp$ values (input value set to 0.08), the center plot shows the distribution of errors, and the right plot shows the pull of these results. The error obtained in the fit to data was 0.055.

6.7 Systematic Uncertainties on $R_\perp$

Table 6.1 summarizes our estimate of the systematic uncertainty $\delta R_\perp$ on the fitted value of $R_\perp$. In this section we describe the individual contributions to the systematic uncertainty.

6.7.1 Angular Resolution

The systematic uncertainty on the value of $R_\perp$ incurred from the parameterization of the resolution function is determined by changing the fit function and determining the difference in the fitted value of $R_\perp$. As seen in Section 6.4.3 the resolution

148

Systematics source	$\delta R_\perp$
Angular Resolution	0.0054
Acceptance Moments Statistics	0.0049
Moments: soft pion efficiency differences (data/MC)	0.0033
Moments: amplitude dependence	0.0042
α parameter scan	0.0003
Floating background parameters	0.0027
TOTAL	0.0094

Table 6.1: Summary of the systematics uncertainties estimated for the value of $R_\perp$.

function was parameterized as a double Gaussian for the "correlated" signal events and a truncated Gaussian for the "uncorrelated" mis-reconstructed signal events. An alternative parameterization of these mis-reconstructed events is given as the sum of sine functions:

$$F_{misReco}(\theta; \theta') \;\; = \;\; (1 - f_{sin2})\frac{1}{2}\sin\theta + f_{sin2}\frac{2}{\pi}\sin^2\theta \qquad (6.27)$$

Using this parameterization for the mis-reconstructed events (independent of the true value) and the canonical double Gaussian for the correlated events, we again have five resolution function parameters (compare with Equation 6.25): three from the double Gaussian $\sigma_{core}, \sigma_{wide}, f_{wide}$ and $f_{sin2}, f_{misReco}$. Refitting the signal Monte Carlo samples, and using the alternate resolution function parameterization in the likelihood fit for $R_\perp$, the fitted value of $R_\perp$ was found to be 0.0683 ± 0.055. The difference between this fit value and the value obtained using the nominal parameterization of the resolution function is considered the systematic uncertainty due to the angular resolution parameterization (0.0053).

It is also necessary to consider the uncertainty incurred from the finite size of the Monte Carlo sample used to determine the parameters of the resolution function. A simple way to evaluate this is to perform the likelihood fit to the data by raising (or lowering) the value of each parameter in the resolution function by one sigma. Thus, we perform the fit to the data for the 15 resolution function parameters (5 parameters for each of the three resolution functions), each time raising the value of the parameter by one sigma. The differences between fitted and nominal values of $R_\perp$ are added in quadrature; the total uncertainty on $R_\perp$ from this method is found to be 0.0011.

The total systematic uncertainty on $R_\perp$ due to the angular resolution function parameterization is taken as the quadratic sum of the two calculations described above (0.0054).

6.7.2 α Parameter Scan

The parameter α, as described previously, defines the ambiguity in the signal pdf of the even amplitudes, $|A_0|$ and $|A_\parallel|$:

$$\alpha = \frac{|A_0^0|^2 - |A_\parallel^0|^2}{|A_0^0|^2 + |A_\parallel^0|^2}$$

In the likelihood fit to determine $R_\perp$, the parameter α was fixed to zero. If we fix α to other values, and keep all other parameters the same in the fit, the fitted value of $R_\perp$ changes. At $\alpha = -1$, the fitted value of $R_\perp = 0.06357$, and at $\alpha = +1$, $R_\perp = 0.0625$. Given that the probability is the same for observing any value of alpha, the systematic error incurred from a lack of knowledge of α is the full range of $R_\perp/\sqrt{12}$, or 0.0003. The dependence of the value of $R_\perp$ versus the value of α is

shown in Figure 6.28.[5]

R_t(fitted) vs. alpha

p0	= 0.201 ± 0.149
p1	= -0.000534 ± 0.233
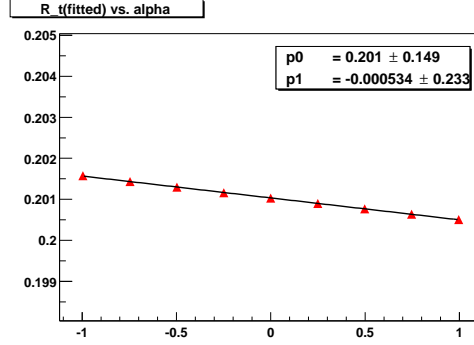
Figure 6.28: The (blinded - see footnote) value of $R_{\perp}$ is plotted as a function of the input value of α used in the fit to the data sample. The line is a linear fit. Each point is a fit to the data sample with a different input value of α. The statistical error bars from the fits are not shown in this plot for clarity.

6.7.3 Moment Integral Parameters

In the definition of the signal $R_{\perp}$ pdf the shapes of the acceptance moments are fixed based on χ^2 fits to 2nd order polynomials. The parameters of these fits, because they are based on limited statistics, will contribute a systematic uncertainty to the fitted value of $R_{\perp}$. In order to estimate this uncertainty, each set of parameters obtained from each acceptance fit are smeared within their errors, taking into account the correlations between parameters. For a given acceptance moment fit, the covariance matrix obtained is used to randomly throw a new set of parameters. These new sets of parameters are then fixed in the signal pdf for $R_{\perp}$ and the data

[5]At the time of the study, the value of $R_{\perp}$ was "blinded" using an arbitrary random number to offset central value. The offset is chosen based on a known seeding string and can be "un-blinded" at the appropriate time. This practice enables the study of systematic effects while avoiding possible biases in the evaluation process.

is refitted. This procedure is repeated 500 times, and the distribution of fitted $R_\perp$ values is then fitted with a single Gaussian. The width of this Gaussian is then taken as the systematic uncertainty in $R_\perp$ based on the errors from the acceptance moments fits. The spread in fitted $R_\perp$ values (blinded – see footnote) for 500 fits to the data is shown in Figure 6.29. The single Gaussian fit returns a $\sigma = 0.0038$.

The mean of this distribution is slightly shifted from the value obtained in the fit to the data (blinded value is 0.2010, the corresponding shift of the mean is 0.0031). It is noted that the smearing of parameters within their errors does not take into account the correlations between the three acceptance moments; this is a likely explanation for the shift in the mean of the distribution. As a conservative estimate of the systematic error we add in quadrature the shift of the mean with the sigma of the fitted Gaussian. Hence, we determine the total error due to the acceptance moments to be 0.0049.
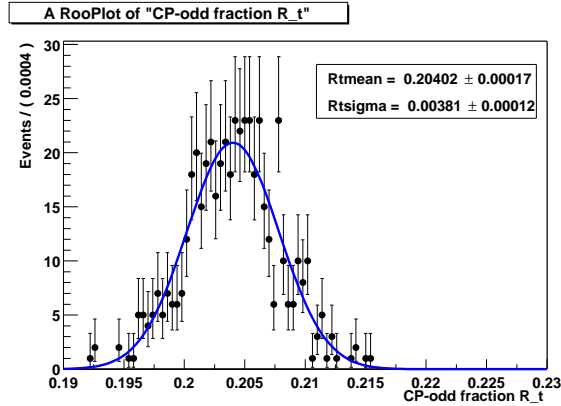
Figure 6.29: The results of smearing the acceptance moments' fit parameters within their errors and refitting the data using the smeared parameters. Shown is the distribution of fitted $R_\perp$ values (blinded – see footnote) for 500 trials of smeared parameters.

6.7.4 Amplitude Dependence of Acceptance Moments

The acceptance moments are derived from the fully-reconstructed Monte Carlo sample. This sample is generated from only one set of amplitudes and phases. While the acceptance moments are in principle independent of the values of the amplitudes, the calculation assumes the efficiency distribution is described entirely as a function of the three decay angles. Since the actual values of α, $R_\perp$, and phase differences are not known in the data, and because different values of these parameters could change the efficiency distribution, we incur a systematic error in the calculation of the acceptance moments from our lack of knowledge of these parameters.

To evaluate our sensitivity to possibly different amplitudes, we first generate a toy sample of 100k events with amplitudes $(A_0, A_\parallel, A_\perp) = (0.74, 0.62, 0.25)$, perfect angular resolution and a soft pion efficiency given by Equation 6.28. The acceptance moments are determined from this sample the fit is performed to determine a "nominal" $R_\perp$ value. Then a second sample is generated with a different value of α and fit using the acceptance moments calculated in the nominal fit. This procedure is repeated for all values of α, the difference in the phases, and for a series of different values of $R_\perp$. The results are seen in Figure 6.30.

To evaluate the systematic error for the amplitude dependence of the acceptance moments, we can take the full range of $\Delta R_\perp$ for each parameter scan seen in Figure 6.30 divided by $\sqrt{12}$. The three cases can then be added in quadrature to obtain a total systematic error from this effect of 0.0042.
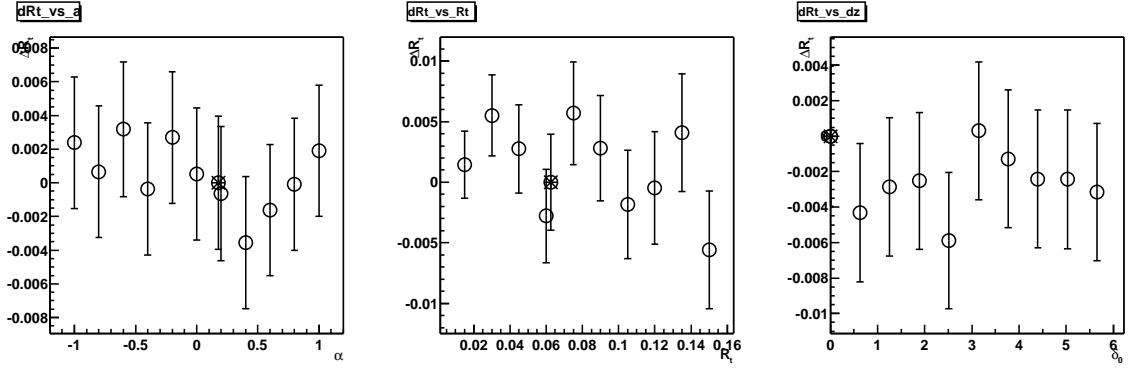
Figure 6.30: Acceptance moments determined from the nominal sample are used in fits to samples where α, $R_\perp$ and the phase difference are changed. Plotted are the difference in fitted values of $R_\perp$ between the nominal sample and the sample where the respective parameter has been changed. Each point represents a different toy Monte Carlo sample of 100k events; the star point is the nominal fit.

6.7.5 Data and Monte Carlo Differences in Soft Pion Efficiency

It has been shown [57] that the Monte Carlo does not model the slow pion efficiency exactly as is seen in data. Since we calculate the acceptance moments from the fully reconstructed Monte Carlo sample and then use them when fitting the data, it is necessary to evaluate how differences between data and Monte Carlo could affect our measurement of $R_\perp$.

We first generate toy Monte Carlo with perfect acceptance (100k events). Reference [57] parameterizes the soft pion efficiency as:

$$\epsilon(p) = \begin{cases} \epsilon_{max}\left(1 - \frac{1}{\beta(p-p_0)+1}\right) & \text{, if } p > p_0 \\ 0 & \text{, if } p \leq p_0 \end{cases} \tag{6.28}$$

where p_0 is the minimum or cutoff value of the soft pion's momentum. Reference [57] also provides the parameters of the above equation from large signal

154

Monte Carlo and data samples of D^* decays. Using the Monte Carlo parameters for this curve, we create a reduced toy sample that accepts/rejects events based on the momentum of the soft pion. This reduced sample now represents the "Monte Carlo -like" version of the acceptance. The acceptance moments are calculated from this sample and the nominal value of $R_\perp$ is determined. Using the same parent sample, we then make a new "data-like" reduced sample using the parameters of [57] corresponding to data. A fit is then performed to this "data-like" sample using the acceptance moments calculated from the "MC-like" sample above. The difference between $R_\perp$ in the "data-like" case and the nominal case is a measure of our sensitivity to the difference between data and Monte Carlo for soft pion efficiency. We find $\Delta R_\perp = -0.0015$.

Reference [57] also provides errors on the parameters derived from the data sample. We can raise (or lower) each parameter by one sigma and refit the corresponding sample. The largest difference in $R_\perp$ for all of the cases is found to be 0.0033. This "worst case" is taken as a conservative estimate of the systematic uncertainty incurred from data/MC differences for soft pion efficiency.

6.7.6 Background Evaluation

As mentioned above, the background in the $\cos(\theta_{\mathrm{tr}})$ distribution is modeled by the shape

$$\mathcal{F}_{bkg}(\cos(\theta_{tr}); b_2) = N \times (1 + b_2 \cos^2(\theta_{tr})) \tag{6.29}$$

and the parameter b_2 is allowed to float in the likelihood fit. To determine the systematic uncertainty incurred from this assumption of the background shape we include all even and odd terms of a fourth order polynomial in $\cos(\theta_{\mathrm{tr}})$ by floating the corresponding parameters b_1, b_2, b_3, and b_4. The fitted value of $R_\perp$ using the full 4th order polynomial in $\cos(\theta_{tr})$ is found to be 0.06564. Thus we assign a

systematic error of 0.0027 for the parameterization of the background shape.

It is also important to validate the result of the fit from the m_{ES} sideband distribution. The highly populated sideband regions defined by

$$-200 < \Delta E < -50\,\mathrm{MeV}$$

$$5.2 < m_{\mathrm{ES}} < 5.3\,\mathrm{GeV}$$

and

$$50 < \Delta E < 200\,\mathrm{MeV}$$

$$5.2 < m_{\mathrm{ES}} < 5.3\,\mathrm{GeV}$$

are statistically strong and provide a means to check the parameterization of the background shape determined in the likelihood fit. Fitting to the upper region in ΔE yields $b_2 = -0.19 \pm 0.12$. In the lower region of ΔE, we see $b_2 = -0.06 \pm 0.11$. In both cases the mean and sigma of the signal Gaussian for the m_{ES} fit were fixed to $5.2794\,\mathrm{GeV}$ and $2.42\,\mathrm{MeV}$, respectively. The fitted values of the signal fractions are also consistent with zero as expected. The projections of these regions and the corresponding fits are shown in Figures 6.31, 6.32. These results show that not only our parameterization of the background is sufficient but also that the parameter b_2 returned in the signal fit are consistent with what is seen in the larger sidebands.
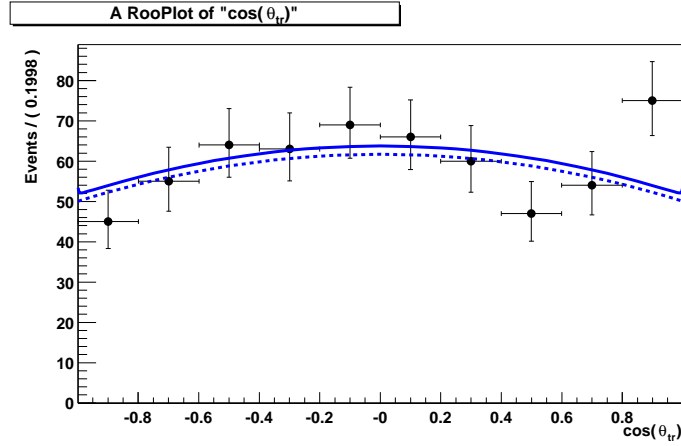
Figure 6.31: The result of the likelihood fit to the region $50 < \Delta E < 200\,\mathrm{MeV}$ and full m_{ES} range is shown. The mean and sigma of the m_{ES} Gaussian fit are fixed to 5.2794 GeV and 2.42 MeV respectively. The parameter b_2, the corresponding parameter of the background parameterization, is found to be -0.19 ± 0.12.
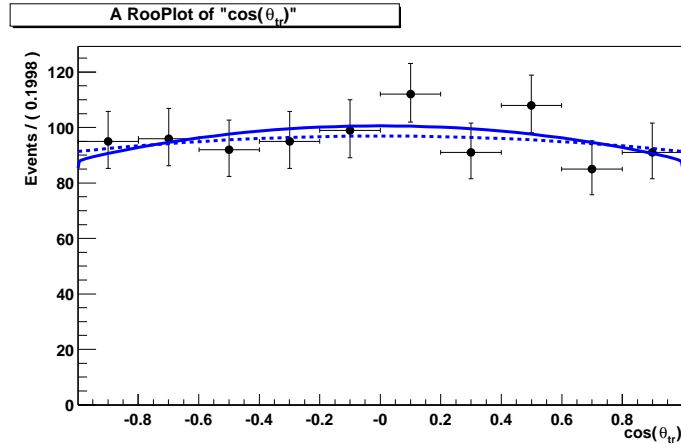
Figure 6.32: The result of the likelihood fit to the region $-200 < \Delta E < -50\,\mathrm{MeV}$ and full m_{ES} range is shown. The mean and sigma of the m_{ES} Gaussian fit are fixed to 5.2794 GeV and 2.42 MeV respectively. The parameter b_2, the corresponding parameter of the background parameterization, is found to be -0.06 ± 0.11.

6.8 Summary

The odd CP parity fraction in the decay $B^0 \rightarrow D^{*+}D^{*-}$, known as $R_\perp$, was determined using the data collected near the $\Upsilon(4S)$ during the period December 1999 - July 2002. We have performed a one-dimensional angular analysis on the 499 fully reconstructed candidate events in the m_{ES} distribution, and estimate the fraction of the CP-odd component of the final state to be

$$R_\perp = 0.063 \pm 0.055(\mathrm{stat}) \pm 0.009(\mathrm{syst}). \tag{6.30}$$

This implies that the $B^0 \rightarrow D^{*+}D^{*-}$ final state is mostly CP-even.

Chapter 7

Measurement of Time-Dependent CP

Asymmetries

The theoretical framework of CP violation and the motivation for measuring CP asymmetries in the decays of B mesons was discussed in Chapter 2. While the $B^0 \to J/\psi K^0_S$ decay provides a direct measure of the quantity $\sin 2\beta$, the decay $B^0 \to D^{*+}D^{*-}$ is also sensitive to the same quantity and provides an independent test of CP violation in the Standard Model. The measurement of time-dependent CP-violating asymmetries with $B^0 \to D^{*+}D^{*-}$ decays requires several experimental ingredients. We begin this chapter with an overview of the analysis and the necessary components of the measurement.

7.1   Analysis Overview

The PEP-II collider, as described in Chapter 3, is a high luminosity e^+e^- storage ring operating at the $\Upsilon(4S)$ resonance energy. The $\Upsilon(4S)$ is a $b\bar{b}$ bound state which decays to a $B^0\overline{B}^0$ or B^+B^- pair. Since the $\Upsilon(4S)$ has spin $S = 1$, and therefore total angular momentum $J = L + S = 1$, the two pseudoscalar mesons must be in an $L = 1$ anti-symmetric state. The time evolution of the $B^0\overline{B}^0$ state is

derived from the relations in Equation 2.17 and represents an example of quantum coherence. The decay of one of the two mesons as a B^0, for example, forces the other meson to necessarily be a $\overline{B}^0$ at the time of the first's decay. If one B, referred to as B_{rec}, is fully reconstructed in a CP eigenstate, then its flavor cannot be determined from its decay products. However, the coherence property of the $\Upsilon(4S)$ decay can be used to infer its flavor at the time of decay of the other B, referred to as B_{tag}. The time-dependent probability distributions of B_{rec} (also see Equation 2.20) are given by

$$
\begin{aligned}
f_{B_{\text{tag}}=B^0}(t_{\text{tag}}, t_{\text{rec}}) \quad \propto \quad & e^{-\Gamma(t_{\text{rec}}-t_{\text{tag}})}\Big\{ 1 + \frac{1-|\lambda_{f_{CP}}|^2}{1+|\lambda_{f_{CP}}|^2}\cos[\Delta m_d(t_{\text{rec}}-t_{\text{tag}})] \\
& - \frac{2\mathcal{I}m\lambda_{f_{CP}}}{1+|\lambda_{f_{CP}}|^2}\sin[\Delta m_d(t_{\text{rec}}-t_{\text{tag}})]\Big\}
\end{aligned}
\tag{7.1}
$$

$$
\begin{aligned}
f_{B_{\text{tag}}=\overline{B}^0}(t_{\text{tag}}, t_{\text{rec}}) \quad \propto \quad & e^{-\Gamma(t_{\text{rec}}-t_{\text{tag}})}\Big\{ 1 - \frac{1-|\lambda_{f_{CP}}|^2}{1+|\lambda_{f_{CP}}|^2}\cos[\Delta m_d(t_{\text{rec}}-t_{\text{tag}})] \\
& + \frac{2\mathcal{I}m\lambda_{f_{CP}}}{1+|\lambda_{f_{CP}}|^2}\sin[\Delta m_d(t_{\text{rec}}-t_{\text{tag}})]\Big\}
\end{aligned}
\tag{7.2}
$$

where t_{rec} is the time of decay of the fully reconstructed CP eigenstate, B_{rec}, and t_{tag} is the time of decay of B_{tag}.[1] A CP-violating asymmetry can be constructed from the time-dependent evolution of B_{rec}. This was introduced in Equation 2.22 as:

$$
a_{f_{CP}} = \frac{(1-|\lambda|^2)}{(1+|\lambda|^2)}\cos(\Delta m_d \Delta t) - \frac{2\mathcal{I}m\lambda}{(1+|\lambda|^2)}\sin(\Delta m_d \Delta t)
\tag{7.3}
$$

where $\Delta t = t_{\text{rec}} - t_{\text{tag}}$ is the appropriate time variable or *clock* for time-dependent measurements in the $\Upsilon(4S) \to B^0\overline{B}^0$ coherent decays. Relations (7.1) and (7.2) are illustrated in Figure 7.1 and are visibly different for events in which B_{tag} is a B^0 and those where B_{tag} is a $\overline{B}^0$. The CP parameter $\lambda = \frac{q}{p}\frac{\bar{A}}{A}$ was discussed

---

[1] A full derivation of the time-dependent probabilities for the $B^0\overline{B}^0$ coherent state is given in Section 1.2.4 of Reference [15].
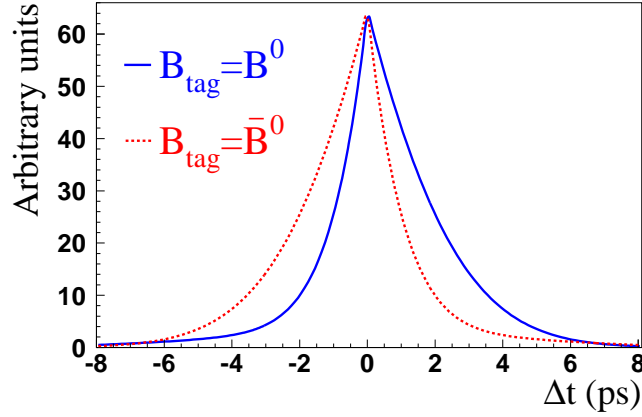
Figure 7.1: Time-dependent rates $f_{B_{\text{tag}}=B^0}$ and $f_{B_{\text{tag}}=\bar{B}^0}$ (see Equations 7.1,7.2). The values of $|\lambda_{f_{CP}}|$ and $\mathcal{I}m\lambda_{f_{CP}}$ are chosen to be 1.0 and 0.75, respectively.

in Section 2.2.3. The $\mathcal{I}m\lambda$ and $|\lambda|$ parameters can be measured experimentally, and are related to different types of CP violation, discussed in Section 2.2.2. In the Standard Model, $|\lambda|$ is expected to be very close to 1. Since the sine term in Equation 7.3 is an odd function of Δt, $\int_{-\infty}^{+\infty} a_{f_{CP}} \mathrm{d}\Delta t = 0$. Therefore, $\mathcal{I}m\lambda$ can only be determined with a time-dependent analysis of the Δt distribution.

Experimentally the value of Δt can be determined from a spatial separation between the decay vertices of the B mesons. In the $\Upsilon(4S)$ rest frame, B mesons are separated by $\sim 30\mu m$ on average. Measuring such a small distance is technologically challenging and the problem is solved by the asymmetric-energy configuration of the PEP-II beams. The $\Upsilon(4S)$ is produced by colliding a 9 GeV electron beam with a 3.1 GeV positron beam, and therefore has a Lorentz boost of $\beta\gamma = 0.55$ (see Section 3.1.1). As a result, the average separation of the two B mesons is about $250\mu m$ along the collision axis (z) in the laboratory frame and can be measured with a precision sufficient for a time-dependent analysis.

The time-dependent measurement of the asymmetry $a_{f_{CP}}$ with $B^0 \to D^{*+}D^{*-}$ decays, therefore, requires three major ingredients:

161

- The flavor of B_{tag} must be determined to establish the flavor of B_{rec} at time $t = t_{\text{tag}}$. In practice, the flavor-tagging procedure will incorrectly assign the flavor of B_{tag} for a the fraction of the events; this must also be measured.

- The decay vertices of B_{rec} and B_{tag} must be determined in order calculate the time difference Δt. The experimental resolution of the Δt measurement must also be determined and parameterized.

- The $D^{*+}D^{*-}$ final state is not a CP eigenstate. An angular analysis of the decay products can separate out the CP-even and CP-odd components and avoid the related dilution on the time-dependent asymmetry measurement.

These ingredients are pictorially represented in Figure 7.2. Each of the above com-
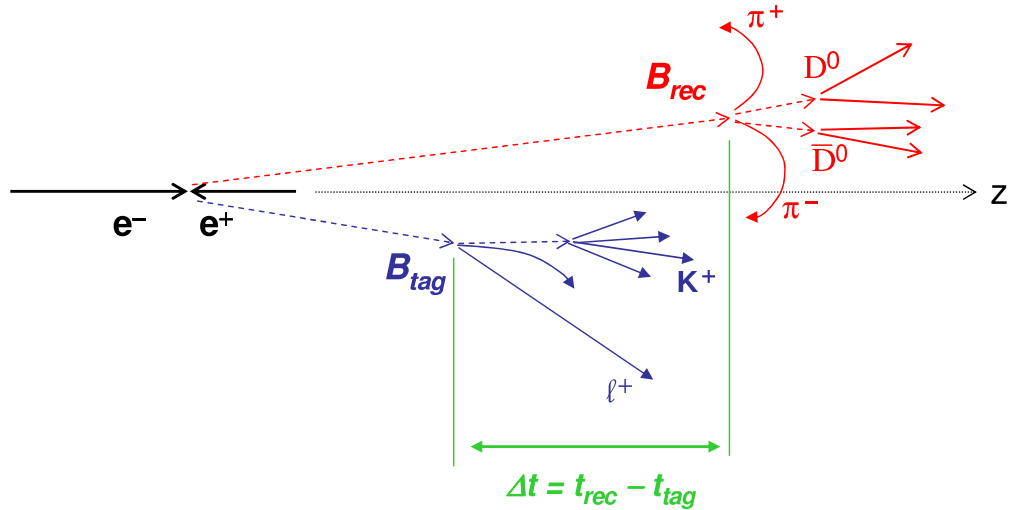
Figure 7.2: An illustration of the $B^0\ \overline{B}^0$ decays used to extract time-dependent CP asymmetries. B_{rec} is fully reconstructed and the remaining particles are analyzed to determine the flavor of B_{tag}. The time-difference between the two B decays is determined from the spacial separation of the decay vertices.

ponents of the analysis warrants further discussion. The following sections provide a brief explanation of the techniques used for flavor-tagging and measurement of Δt in $B^0 \to D^{*+} D^{*-}$ events. The specific $B^0 \to D^{*+} D^{*-}$ angular distribution and fit method is then discussed in detail, followed by the validation and results of the CP asymmetry measurement in Chapter 8.

7.2   Determining the Flavor of B Mesons

One of the primary ingredients in the CP asymmetry measurement is the determination of the flavor of B_{tag}. Since the final state $D^{*+} D^{*-}$ is a mostly CP-even eigenstate, its flavor at time $t = t_{\text{tag}}$ is deduced from the coherence property of the $\Upsilon(4S)$ decay; that is, B_{rec} is a B^0 ($\overline{B}^0$) if B_{tag} is a $\overline{B}^0$ (B^0) at the time of its decay. The flavor of B_{tag} is correlated with the charge of leptons, kaons and pions in its decay chain. It can, therefore, be determined from kinematic properties and particle identification information of a partial reconstruction of the B_{tag} final state.

The final state decays of a B meson offers a variety of characteristic processes that can be recognized by a "flavor-tagging" algorithm. This algorithm has a non-zero probability of choosing the wrong flavor from the final state particles. It is customary, therefore, to define an "effective tagging power" $Q \equiv \epsilon (1-2w)^2$ where ϵ is the efficiency for determining the flavor and w is the fraction of candidates with a wrong flavor assignment. The quantity Q directly affects the statistical uncertainty on $\mathcal{I}m\lambda_{CP}$ (since only flavor-tagged events are used, the statistical power of the sample is dependent on the tagging efficiency and the mis-tag fractions).[2]

The following B processes which give information about the flavor of B_{tag} is

---

[2]The expected statistical uncertainty can be analytically determined [15, 60] from the likelihood function and is proportional to $1/\sqrt{N_{\text{sig}}} \sqrt{\sum_{i=1}^{4} \epsilon_i (1 - 2w_i)^2}$.

summarized below (for more detail, consult Reference [58]).

- The semileptonic decays $B \rightarrow X\ell\nu_\ell$ constitute 20% of all decay modes of the B mesons and represent the primary source of leptons for flavor tagging. In the dominant tree diagram of this process, the lepton is generated from the W^- (W^+) boson emitted by the b ($\bar{b}$) quark; hence, a positively (negatively) charged lepton ℓ^+ (ℓ^-) indicates a parent B^0 ($\overline{B}^0$). While the probability w of assigning a wrong flavor is lowest for this class of decays, there are two sources of wrong flavor assignment: hadrons mis-identified as leptons (fake leptons) and leptons which do not originate from the b quark (wrong-sign leptons from D meson decays, for example). These *secondary* leptons can be distinguished from the *primary* leptons produced by the B using the center-of-mass momentum spectrum of the leptons in the data.

- The quark "cascade" decay chain $b \rightarrow c \rightarrow s$ is the primary source of kaons in B decays. The K^- produced from the hadronization of the s quark is associated with an initial $\overline{B}^0$ state. The W^- emitted by the b quark can also hadronize into a K^- or D_s^+ meson (the latter can be inclusively reconstructed for flavor-tagging information). The W^+ in the decay of the c quark, however, can hadronize into a wrong-sign K^+. Another source of wrong flavor assignment from kaons is the mis-identification of pions as kaons. The majority of kaons in B decays, however, have the right sign and the mis-identification rate is less than 5%.

- The decays $B^0 \rightarrow D^{*-}\pi^+, \rho^+, a_1^+$ can also be used to determine the flavor of neutral B mesons. The charge of the slow pion from the D^* decay is correlated with the flavor of the parent B. Also the high-momentum pion from the W^+ hadronization also provides tagging information from its charge.

Using the above physics-based information an algorithm was developed to exploit these phenomena as well as other information from the kinematic and angular properties of the B_{tag} decay. Neural networks are developed to recognize the characteristic signatures discussed above. The details of the training of the neural networks as well as the structure of the algorithm is beyond the scope of this thesis. [58] However, the output of the algorithm is a sub-division of events into four hierarchical mutually-exclusive tagging categories. The `Lepton` category provides a flavor tag for events with an identified electron or muon and has the lowest estimated mistag probability. If the criteria of the `Lepton` category is not met then events are evaluated for the `Kaon I` and `Kaon II` categories. Here, a probability is determined from the multiple kaons identified in the event; the `Kaon I` category contains events with higher correct-tag probability. Lastly, an `Inclusive` category attempts to determine the flavor tag using the center-of-mass momentum of charged tracks. The intent is to identify fast tracks, *e.g.* fast pions from $B^0 \to D^{*-}\pi^+$ decays, and recover primary leptons not assigned to the `Lepton` category.

7.2.1 Estimating Tagging Performance with the B_{flav} Sample

A large sample of fully reconstructed B-flavor eigenstates is used to estimate the performance of the tagging algorithms as well as estimate the mis-tagging (wrong flavor assignment) rates of the tagging algorithms. The "B_{flav}" sample is composed of the decays $D^{*-}\pi^+/\rho^+/a_1^+$, $D^-\pi^+/\rho^+/a_1^+$, and $J/\psi K^{*0}(K^+\pi^-)$. These decays are used for the precise measurement of the B^0-$\overline{B}^0$ oscillation frequency Δm_d and the B^0 lifetime, and are essential for all the time-dependent CP-violation analyses. Not only is the sample used to determine the performance of the flavor-tagging

algorithm, but is also used to measure the Δt resolution function in data and used as a control sample for the CP-violation asymmetry measurements. The details of selection criteria and sample composition is given in Reference [61]. The m_{ES} distribution for these events is shown in Figure 7.3.
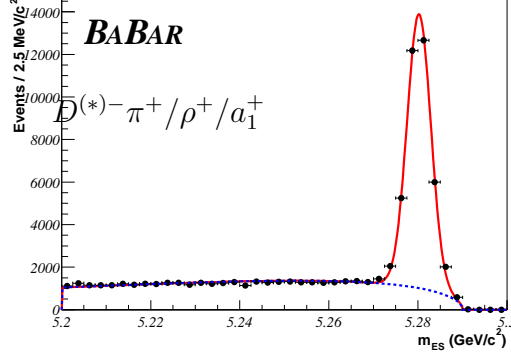
Figure 7.3: Distribution of m_{ES} for selected B^0 candidates in flavor eigenstates in the data sample.

The performance of the b-flavor–tagging algorithm in simulated events is reported in Table 7.1. These results are measured in a large sample of simulated $\Upsilon(4S) \to B\bar{B}$ events, with one B decaying to flavor eigenstates. The fraction w of wrongly tagged B^0 mesons can be different from the fraction $\overline{w}$ of mistagged $\overline{B}^0$ mesons. In order to account for such a difference the average mistag fraction $\langle w \rangle = (w + \overline{w})/2$ and the difference $\Delta w = w - \overline{w}$ are measured. One observes that about 2/3 of all selected B candidates are assigned a flavor tag.

The observed Δt distributions for B^0-tagged and $\overline{B}^0$-tagged events are derived from Equations 7.1 and 7.2 by including the fractions w_i and $\overline{w}_i$, and are given by

$$f'_{B_{\text{tag}}=B^0}(\Delta t; w_i, \overline{w}_i) = (1-w)\, f\left(B_{\text{tag}} \equiv B^0, \Delta t\right) + \overline{w}\, f\left(B_{\text{tag}} \equiv \overline{B}^0, \Delta t\right) \quad (7.4)$$

$$f'_{B_{\text{tag}}=\overline{B}^0}(\Delta t; w_i, \overline{w}_i) = w\, f\left(B_{\text{tag}} \equiv B^0, \Delta t\right) + (1-\overline{w})\, f\left(B_{\text{tag}} \equiv \overline{B}^0, \Delta t\right) \quad (7.5)$$

In order to keep the expression of these distributions simple, it is convenient to

Category	N_{sig}	$\varepsilon(\%)$	$\langle w\rangle(\%)$	$\Delta w(\%)$	$Q(\%)$
lepton	11607 ± 108	10.3 ± 0.1	3.5 ± 0.3	-0.9 ± 0.5	8.9 ± 0.1
Kaon I	19759 ± 141	17.5 ± 0.1	9.0 ± 0.3	-0.2 ± 0.5	11.8 ± 0.2
Kaon II	22557 ± 150	20.0 ± 0.1	21.2 ± 0.4	-2.7 ± 0.6	6.6 ± 0.2
Inclusive	22330 ± 149	19.8 ± 0.1	30.9 ± 0.4	-3.2 ± 0.6	2.9 ± 0.1
Total	113050 ± 336	67.5 ± 0.2			30.2 ± 0.3

Table 7.1: The efficiency ε_i, average mistag fraction $\langle w_i\rangle$, mistag difference Δw_i, and tagging power $Q_i = \epsilon_i \cdot (1 - 2\langle w_i\rangle)^2$ for each tagging category. N_{sig} is the number of simulated signal events.

use two new parameters in place of w and $\overline{w}$. The average dilution $\langle \mathcal{D}\rangle$ and the difference $\Delta \mathcal{D}$ between the dilutions for B^0 and $\overline{B}^0$ are linear functions of w and $\overline{w}$ and are defined as

$$
\begin{aligned}
\langle w\rangle &= \tfrac{1}{2}(w + \overline{w}), & \Delta w &= (w - \overline{w}) \\
\mathcal{D} &= 1 - 2w, & \overline{\mathcal{D}} &= 1 - 2\overline{w} \\
\langle \mathcal{D}\rangle &= \tfrac{1}{2}(\mathcal{D} + \overline{\mathcal{D}}), & \Delta\mathcal{D} &= (\mathcal{D} - \overline{\mathcal{D}}) .
\end{aligned}
$$

After some algebra, the distributions (7.4) and (7.5) can be written as

$$
\begin{aligned}
f'_{B_{\text{tag}}=B^0}(\Delta t; \langle\mathcal{D}\rangle_i, \Delta\mathcal{D}_i) &= \frac{\Gamma}{4}e^{-\Gamma|\Delta t|}\Big\{(1 + \tfrac{1}{2}\Delta\mathcal{D}_i) \\
&\quad -\langle\mathcal{D}\rangle_i\left[S\sin(\Delta m_d\Delta t) - C\cos(\Delta m_d\Delta t)\right]\Big\} \quad (7.7) \\
f'_{B_{\text{tag}}=\overline{B}^0}(\Delta t; \langle\mathcal{D}\rangle_i, \Delta\mathcal{D}_i) &= \frac{\Gamma}{4}e^{-\Gamma|\Delta t|}\Big\{(1 - \tfrac{1}{2}\Delta\mathcal{D}_i) \\
&\quad +\langle\mathcal{D}\rangle_i\left[S\sin(\Delta m_d\Delta t) - C\cos(\Delta m_d\Delta t)\right]\Big\} \quad (7.8)
\end{aligned}
$$

where S and C are the coefficients of the sine and cosine Δt terms, respectively.

In the time-dependent analysis of $B^0 \to D^{*+}D^{*-}$ decays, the values for mistagging rates are determined from the B_{flav} sample. First, the B_{flav} sample in data

Parameter	$D^{*+}D^{*-}$ Truth	B_{flav} Truth	B_{flav} fitted
$\langle\mathcal{D}\rangle$(Lepton)	0.946 ± 0.004	0.931 ± 0.003	0.930 ± 0.006
$\langle\mathcal{D}\rangle$(KaonI)	0.815 ± 0.006	0.806 ± 0.004	0.820 ± 0.006
$\langle\mathcal{D}\rangle$(KaonII)	0.574 ± 0.008	0.566 ± 0.006	0.577 ± 0.008
$\langle\mathcal{D}\rangle$(Incl.)	0.389 ± 0.009	0.370 ± 0.006	0.382 ± 0.008
$\Delta\mathcal{D}$(Lepton)	0.009 ± 0.009	0.012 ± 0.007	0.019 ± 0.011
$\Delta\mathcal{D}$(KaonI)	0.003 ± 0.013	0.0001 ± 0.008	0.003 ± 0.010
$\Delta\mathcal{D}$(KaonII)	0.054 ± 0.017	0.050 ± 0.011	0.053 ± 0.012
$\Delta\mathcal{D}$(Incl.)	0.052 ± 0.018	0.063 ± 0.012	0.064 ± 0.013

Table 7.2: Comparison of mistag dilution parameters in $D^{*+}D^{*-}$ and B_{flav} Monte Carlo samples. "Truth" refers to the dilutions obtained when the true flavor of B_{tag} is known, while "fitted" corresponds to the dilutions obtained from the B_{flav} fit procedure.

is much larger than the $D^{*+}D^{*-}$ event sample so the determination of $\langle\mathcal{D}\rangle_i$ and $\Delta\mathcal{D}_i$ is more precise. Secondly, using a large data sample avoids any dependence on the Monte Carlo simulations which slightly overestimates reconstruction efficiency and relies on known branching fractions for all B decay channels (theoretical assumptions must be used for remaining channels).

Because the B_{flav} sample is used to determine tagging performance, the difference between B_{flav} and $D^{*+}D^{*-}$ samples is studied in simulated samples in order to estimate systematic uncertainties. Table 7.2 shows the values of average dilutions $\langle\mathcal{D}\rangle_i$ and dilution difference $\Delta\mathcal{D}_i$ between B^0 and $\overline{B}^0$ for the four tagging categories, as determined from the tagging algorithm in the Monte Carlo samples.

7.3 Measurement of Δt

The next ingredient of the time-dependent analysis of CP violation is the measurement of the time interval Δt between the decay of the fully reconstructed B meson, $B_{\rm rec}$, and the decay of the tagging B meson, $B_{\rm tag}$. If t_0 is the time when the $\Upsilon(4S)$ decays in two B mesons, there must be always one B^0 and one $\bar{B}^0$ for times $t > t_0$. This conditions holds until the decay of one of the two mesons, $B_{\rm tag}$, in a flavor eigenstate at time $t_{\rm tag} > t_0$. The flavor of the other B meson, $B_{\rm rec}$ which is fully reconstructed, must be opposite to the flavor of $B_{\rm tag}$ at time $t_{\rm tag}$, in order to satisfy the coherence condition.

The time of decay of $B_{\rm rec}$ can be before or after $t_{\rm tag}$. That is, the time evolution is described by the same distribution (see Equation 2.17) with the "boundary condition" that the flavor of $B_{\rm rec}$ is determined at $t = t_{\rm tag}$. Thus, it follows that the time evolution of the B mesons are a function of the interval $\Delta t = t_{\rm rec} - t_{\rm tag}$ (which can be positive or negative) and is not dependent on the $\Upsilon(4S)$ production time, t_0.

The value of Δt can be measured by reconstructing the decay vertices of the B mesons, and measuring the spatial separation between them, without reconstructing the $\Upsilon(4S)$ decay point. The distance is then converted to Δt by using the boost factor $\beta\gamma$ that is known from the beam energies. A naive determination of Δt is given by the relation

$$\Delta z = \beta\gamma c \Delta t \; , \tag{7.9}$$

where $\beta\gamma = 0.55$ is the $\Upsilon(4S)$ Lorentz boost factor, which is known with a precision of 0.1%. Its value is calculated from the beam energies which are monitored every 5 seconds. Equation 7.9 represents the limit where the B mesons are at rest in the $\Upsilon(4S)$ frame, and the boost is exactly along the z axis. In practice, the

detector symmetry axis, and therefore the boost axis, is rotated with respect to the beams by 20 mrad. Furthermore, the B mesons have a momentum of about 340 MeV/c in the $\Upsilon(4S)$ frame. The rotation is incorporated through the Lorentz transformations, and the measured momentum of the B_{rec} candidate can be used to correct (7.9) and account for the B momentum. The exact relationship between Δt and Δz, including these corrections, is discussed in Reference [59].

7.3.1 Determining the B_{tag} and B_{rec} Vertices

The decay vertex of the B_{rec} candidate is reconstructed by using all its decay daughters in the final state. Charged tracks originating from intermediate states, e.g. a D^0 or a K^0_S are replaced by *virtual* composite candidates and appropriate spatial and kinematic constraints are used in the fit to the B_{rec} vertex. The typical resolutions on the position of the vertex along the z axis and in the transverse plane are $\approx$?? μm and $\approx 65\,\mu$m, for the $D^{*+}D^{*-}$ and B_{flav} samples, respectively.

The decay vertex of the B_{tag} candidate is reconstructed with an inclusive technique, using charged tracks not used in the reconstruction of B_{rec}. Charged tracks originating from long-lived particles, K^0_Ss and Λ^0s, are removed and replaced by the reconstructed composite candidates in order to reduce potential biases. These composite candidates and the remaining charged tracks are used as input in a geometrical fit to determine a common decay vertex. Since the three-momentum $\vec{p}_{\mathrm{rec}}$ and the decay vertex of the B_{rec} candidate are measured with good precision, the three-momentum of B_{tag} can be constrained kinematically by using the measured momentum of the $\Upsilon(4S)$ and $\vec{p}_{\mathrm{rec}}$. Since D^0 and D^+ mesons have decay lengths $c\tau$ of about, 125 μm and 315 μm, respectively, the determination of the B_{tag} vertex will typically be biased.[3] Hence, the vertex is determined with an iterative proce-

---

[3]The bias is dependent on the actual flight of the D mesons in the lab frame, which depends

dure, where tracks with a large contribution to the fit χ^2 ($\Delta\chi^2 > 6$) are removed until no track fails the χ^2 requirement or only two tracks remain.

The value of Δz is determined directly in the B_{tag} vertex fit. The fit also provides a correct estimate of the uncertainty $\sigma_{\Delta z}$ by taking into account the correlation between the B_{tag} and B_{rec} vertices (which originates from the use of the three-momentum of B_{rec} as a constraint on the B_{tag} vertex).

7.3.2 Δt Resolution

The measured and true values of Δt differ due to the finite resolution of the detector in the measurement of decay vertices. The detector response for Δt, called the Δt resolution function, is parameterized with a sum of three Gaussian distributions (core, tail, and outliers components) as a function of the residual $\delta_t \equiv \Delta t_{\text{meas}} - \Delta t_{\text{true}}$ as

$$
\mathcal{R}(\delta_t; \hat{a}) = \sum_k^{\text{core,tail}} \frac{f_k}{S_k \sigma_{\Delta t} \sqrt{2\pi}} \exp\left(-\frac{(\delta_t - b_k \sigma_{\Delta t})^2}{2(S_k \sigma_{\Delta t})^2}\right)
$$
$$
+ \frac{f_{\text{outl}}}{\sigma_{\text{outl}} \sqrt{2\pi}} \exp\left(-\frac{\delta_t^2}{2\sigma_{\text{outl}}^2}\right) \tag{7.10}
$$

where f_k is the fraction of events in each component. The width σ of the core and tail components can be written as

$$
\sigma_{\text{core}} = S_{\text{core}} \ \sigma_{\Delta t} \quad , \quad \sigma_{\text{tail}} = S_{\text{tail}} \ \sigma_{\Delta t}
$$

where $\sigma_{\Delta t}$ is the measured uncertainty on Δt, determined for every event, and S_{core} and S_{tail} are scale factor parameters. These factors account for an overall underestimate ($S_k > 1$) or overestimate ($S_k < 1$) of the uncertainty $\sigma_{\Delta t}$ for all events.

---

on both $\beta\gamma$ and the direction of flight.
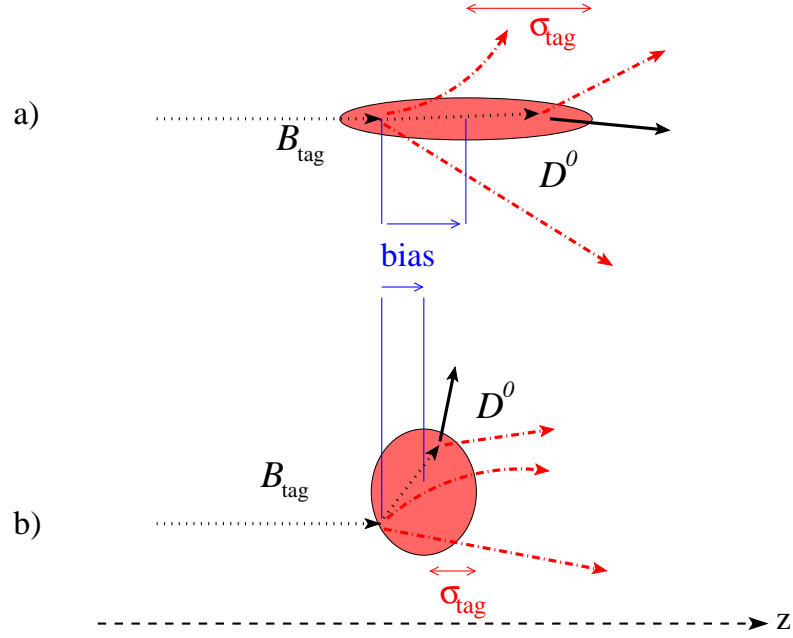
Figure 7.4: Correlation between the bias of the B_{tag} vertex and its uncertainty σ_{tag} when a) the D^0 flies in the direction of B_{tag}, or b) the D^0 is perpendicular to the direction of B_{tag}.

The core and tail Gaussians are allowed to have a non-zero mean offset δ_0 to account for residual charm decay products included in the B_{tag} vertex which tend to bias the Δt determination. These offsets are proportional to the uncertainty $\sigma_{\Delta t}$ and are therefore parameterized as

$$\delta_{\text{core}}^0 = b_{\text{core}} \ \sigma_{\Delta t} \quad , \quad \delta_{\text{tail}}^0 = b_{\text{tail}} \ \sigma_{\Delta t} \ .$$

D mesons in the B_{tag} decay with flight direction perpendicular to the z axis in the laboratory frame have the best z resolution, and introduce the smallest bias in the measured z position of the B_{tag} vertex. D mesons that travel forward in the laboratory, however, have poorer z resolution, and introduce a larger bias in the position of the B_{tag} vertex (see Figure 7.4).

Figure 7.5a,b shows the correlation between the RMS spread of Δt and $\sigma_{\Delta t}$

in simulated events as well as the correlation between the mean of δ_t and the uncertainty $\sigma_{\Delta t}$.
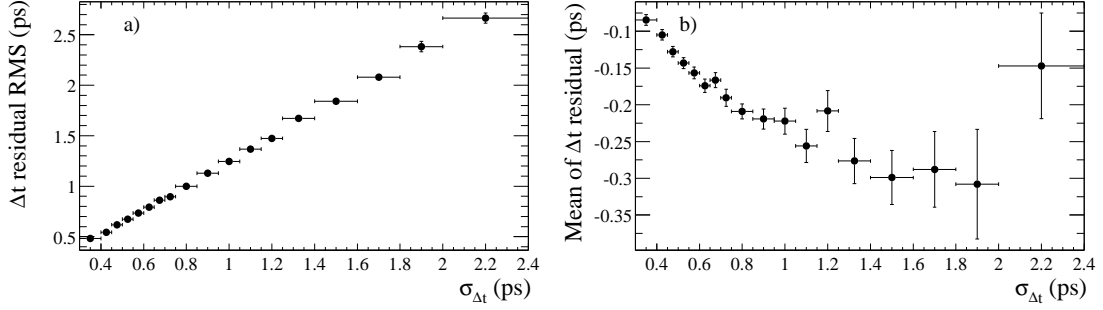
Figure 7.5: Correlation between $\sigma_{\Delta t}$ and a) the RMS spread and b) the mean of the residual $\delta_t = \Delta t_{\text{meas}} - \Delta t_{\text{true}}$, in simulated events.

The outliers component has a fixed width of 8 ps and no offset and accounts for 0.3% of selected events, which have mis-reconstructed vertices. All parameters of the Δt resolution function are measured from the B_{flav} sample in data.

The finite Δt resolution is incorporated by convolving (7.7) and (7.8) with the Δt resolution function $\mathcal{R}(\delta_t; \hat{a})$

$$\mathcal{F}_{B_{\text{tag}}=B^0}\left(\Delta t; \langle \mathcal{D} \rangle_i, \Delta \mathcal{D}_i, \hat{a}_i\right) = f'_{B_{\text{tag}}=B^0}\left(\Delta t_{\text{true}}; \langle \mathcal{D} \rangle_i, \Delta \mathcal{D}_i\right) \otimes \mathcal{R}(\delta_t; \hat{a}_i) \quad (7.11)$$

$$\mathcal{F}_{B_{\text{tag}}=\bar{B}^0}\left(\Delta t; \langle \mathcal{D} \rangle_i, \Delta \mathcal{D}_i, \hat{a}_i\right) = f'_{B_{\text{tag}}=\bar{B}^0}\left(\Delta t_{\text{true}}; \langle \mathcal{D} \rangle_i, \Delta \mathcal{D}_i\right) \otimes \mathcal{R}(\delta_t; \hat{a}_i) \quad (7.12)$$

Figure 7.6 illustrates the distributions (7.11) and (7.12) for realistic choices of mistag fractions and Δt resolution function parameters.

7.3.3   Comparison of Δt Between D^{*+} D^{*-} and B_{flav}

We assume that a common Δt resolution function can be used for events reconstructed in flavor eigenstates (B_{flav}) and in the $B^0 \rightarrow D^{*+}D^{*-}$ final state. Since
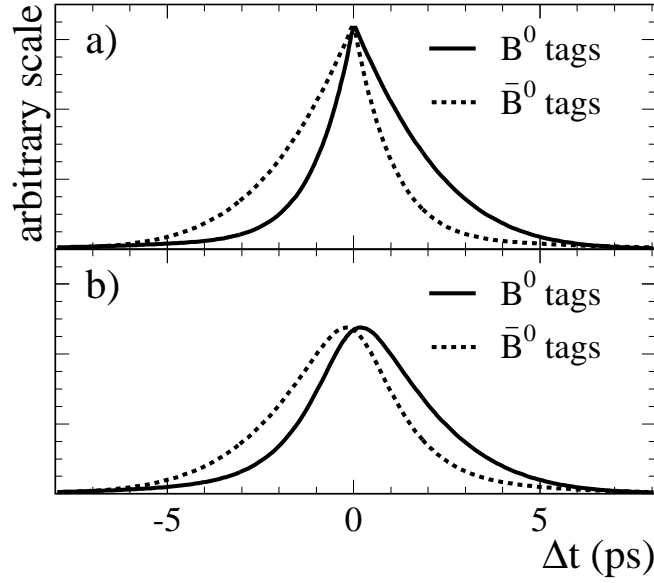
Figure 7.6: Expected Δt distribution for B^0- and $\overline{B}^0$-tagged CP events with a) perfect tagging and Δt resolution, and b) typical mistag fractions and finite Δt resolution. The scale is arbitrary but is the same for the two plots.

the resolution on the B_{tag} vertex is about $190\mu m$, which is much larger than the B_{rec} resolution, the two samples are expected to have similar Δz resolutions. Any difference can be accounted for as a systematic uncertainty.

A comparison of Δt resolution function parameters between $D^{*+}D^{*-}$ signal Monte Carlo and B_{flav} Monte Carlo samples is summarized in Table 7.3.

Figure 7.7a (left) shows Δt residual in $D^{*+}D^{*-}$ signal Monte Carlo events (points). The two curves superimposed are obtained by using the fitted parameters of the signal sample (dashed curve) and on B_{flav} sample (plain curve). Figure 7.7b (right) shows the pull distribution of Δt residual for the two Monte Carlo samples.

174

Parameter	$D^{*+}D^{*-}$ Truth	$D^{*+}D^{*-}$ fitted	B_{flav} Truth	B_{flav} fitted
S_{core}	1.235 ± 0.008	1.175 ± 0.043	1.240 ± 0.005	1.176 ± 0.020
S_{tail}	3.0 (fixed)	3.0 (fixed)	3.0 (fixed)	3.0 (fixed)
$\delta_{\text{core}}^{0 \text{ Lepton}}$	-0.117 ± 0.019	-0.069 ± 0.050	-0.132 ± 0.013	-0.099 ± 0.031
$\delta_{\text{core}}^{0 \text{ KaonI}}$	-0.222 ± 0.016	-0.197 ± 0.041	-0.272 ± 0.010	-0.247 ± 0.026
$\delta_{\text{core}}^{0 \text{ KaonII}}$	-0.267 ± 0.015	-0.231 ± 0.036	-0.276 ± 0.009	-0.246 ± 0.023
$\delta_{\text{core}}^{0 \text{ Incl.}}$	-0.202 ± 0.014	-0.198 ± 0.036	-0.208 ± 0.009	-0.201 ± 0.023
δ_{tail}^{0}	-0.969 ± 0.074	-1.228 ± 0.314	-1.332 ± 0.058	-1.075 ± 0.157
f_{tail}	0.093 ± 0.004	0.113 ± 0.022	0.082 ± 0.003	0.091 ± 0.009
f_{outlier}	0.004 ± 0.001	0.002 ± 0.001	0.005 ± 0.001	0.003 ± 0.001

Table 7.3: Fitted parameters of the Δt resolution function for $D^{*+}D^{*-}$ and B_{flav} Monte Carlo samples.

7.4 Time Dependent Angular Distribution

The last ingredient for measuring time-dependent CP asymmetries in $B^0 \to D^{*+}D^{*-}$ decays is to separate out the CP-even and CP-odd components of the final state. To understand the necessity of this, we return to the time-dependent angular dis-
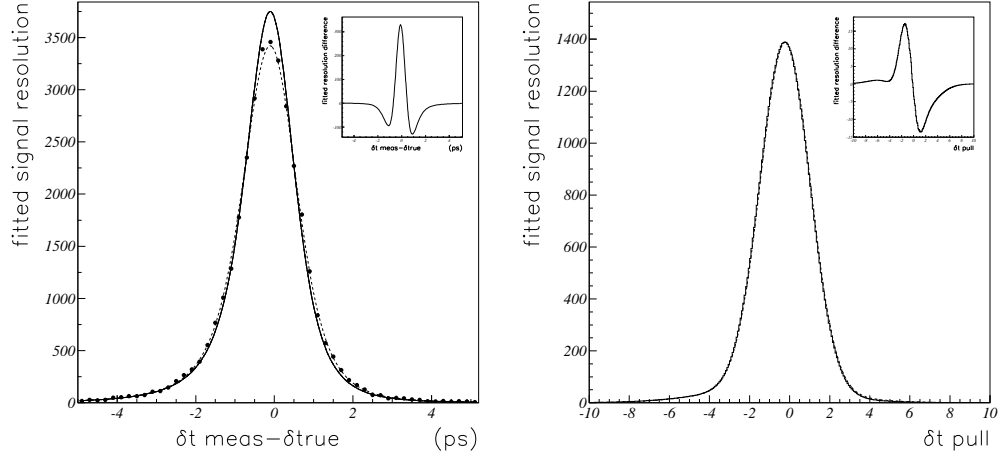
Figure 7.7: Comparison of the fitted δt residual (left) and Δt pull (right) between the B_{flav} and $D^{*+}D^{*-}$ Monte Carlo samples. In the left plot the two curves superimposed are obtained by using the fitted parameters on the signal sample (dashed curve) and on B_{flav} sample (plain curve). The inset distributions show the difference between the two curves.

tribution of the $B^0 \to D^{*+}D^{*-}$ decay in the transversity basis (Equation 2.42):

$$
\frac{1}{\Gamma}\frac{\mathrm{d}^4\Gamma}{\mathrm{d}\cos\theta_1\,\mathrm{d}\cos\theta_{\text{tr}}\,\mathrm{d}\phi_{\text{tr}}\,\mathrm{d}t} = \frac{9}{32\pi}\frac{1}{|A_0|^2 + |A_\parallel|^2 + |A_\perp|^2}
$$

$$
\{4|A_0|^2\cos^2\theta_1\sin^2\theta_{\text{tr}}\cos^2\phi_{\text{tr}}
$$

$$
+2|A_\parallel|^2\sin^2\theta_1\sin^2\theta_{\text{tr}}\sin^2\phi_{\text{tr}}
$$

$$
+2|A_\perp|^2\sin^2\theta_1\cos^2\theta_{\text{tr}}
$$

$$
+\sqrt{2}\mathcal{R}e(A_\parallel^*A_0)\sin2\theta_1\sin^2\theta_{\text{tr}}\sin2\phi_{\text{tr}}
$$

$$
-\sqrt{2}\mathcal{I}m(A_0^*A_\perp)\sin2\theta_1\sin2\theta_{\text{tr}}\cos\phi_{\text{tr}}
$$

$$
-2\mathcal{I}m(A_\parallel^*A_\perp)\sin^2\theta_1\sin2\theta_{\text{tr}}\sin\phi_{\text{tr}}\} . \quad (7.13)
$$

176

Let us now integrate over $\cos\theta_1$ and ϕ_{tr}, keeping the time dependence:

$$\frac{1}{\Gamma}\frac{\mathrm{d}^2\Gamma(\overset{(-)}{B^0}\to D^{*+}D^{*-})}{\mathrm{d}z\mathrm{d}t} = \frac{9}{32\pi}\frac{1}{|A_0|^2+|A_\parallel|^2+|A_\perp|^2}$$
$$\{(|\overset{(-)}{A_\parallel}|^2+|\overset{(-)}{A_0}|^2)G_+(z)+|\overset{(-)}{A_\perp}|^2G_-(z)\} \quad (7.14)$$

where $z=\cos(\theta_{\mathrm{tr}})$ and

$$\begin{aligned}
G_+(z) &= \frac{8\pi}{3}(1-z^2)=\frac{8\pi}{3}\sin^2\theta_{\mathrm{tr}} \\
G_-(z) &= \frac{16\pi}{3}z^2=\frac{16\pi}{3}\cos^2\theta_{\mathrm{tr}}.
\end{aligned} \quad (7.15)$$

These are the CP-even $(+)$ and CP-odd $(-)$ terms which correspond to the extreme values of $R_\perp$ discussed in Section 6.1. We now insert the explicit time-dependence of the transversity amplitudes:

$$\begin{aligned}
A_{0,\parallel}(t) &= A_{0,\parallel}^0\, e^{-imt}\, e^{-\Gamma t/2}\left(\cos\frac{\Delta mt}{2}+i\lambda\sin\frac{\Delta mt}{2}\right) \\
A_\perp(t) &= A_\perp^0\, e^{-imt}\, e^{-\Gamma t/2}\left(\cos\frac{\Delta mt}{2}-i\lambda\sin\frac{\Delta mt}{2}\right)
\end{aligned} \quad (7.16)$$

and:

$$\begin{aligned}
\bar{A}_{0,\parallel}(t) &= \bar{A}_{0,\parallel}^0\, e^{-imt}\, e^{-\Gamma t/2}\left(\cos\frac{\Delta mt}{2}+\frac{i}{\lambda}\sin\frac{\Delta mt}{2}\right) \\
\bar{A}_\perp(t) &= -\bar{A}_\perp^0\, e^{-imt}\, e^{-\Gamma t/2}\left(\cos\frac{\Delta mt}{2}-\frac{i}{\lambda}\sin\frac{\Delta mt}{2}\right)
\end{aligned} \quad (7.17)$$

where the superscript 0 refers to the magnitude of the amplitude at $t=0$. The above expressions also assume that penguin diagram contributions are negligible so that $\lambda=\frac{q}{p}\frac{\bar{A}}{A}$ is the same for the three transversity eigenstates. Therefore, we

find

$$\frac{1}{\Gamma}\frac{\mathrm{d}^2\Gamma(\overset{(-)}{B^0}\to D^{*+}D^{*-})}{\mathrm{d}z\mathrm{d}t} = \frac{9}{32\pi}\frac{e^{-\Gamma t}}{|A_0^0|^2+|A_\parallel^0|^2+|A_\perp^0|^2}\Big\{$$

$$[(|A_\parallel^0|^2+|A_0^0|^2)G_+(z)+|A_\perp^0|^2 G_-(z)]$$

$$\overset{(-)}{+}\frac{(1-|\lambda|^2)}{(1+|\lambda|^2)}\ [(|A_\parallel^0|^2+|A_0^0|^2)G_+(z)+|A_\perp^0|^2 G_-(z)]\cos\Delta mt$$

$$\overset{(+)}{-}\frac{2\mathcal{I}m(\lambda)}{(1+|\lambda|^2)}\ [(|A_\parallel^0|^2+|A_0^0|^2)G_+(z)-|A_\perp^0|^2 G_-(z)]\sin\Delta mt\Big\}.$$

$$(7.18)$$

If we now define the CP dilution factor to be

$$K = \frac{|A_0^0|^2+|A_\parallel^0|^2-|A_\perp^0|^2}{|A_0^0|^2+|A_\parallel^0|^2+|A_\perp^0|^2} = 1-2R_\perp \tag{7.19}$$

and integrate over $z = \cos(\theta_{\mathrm{tr}})$ such that all the angular information is lost then

we have:

$$\frac{1}{\Gamma}\frac{\mathrm{d}\Gamma(\overset{(-)}{B^0}\to D^{*+}D^{*-})}{\mathrm{d}t} = e^{-\Gamma t}\left\{1\overset{(-)}{+}\frac{(1-|\lambda|^2)}{(1+|\lambda|^2)}\cos\Delta mt\overset{(+)}{-}\frac{2\mathcal{I}m(\lambda)}{(1+|\lambda|^2)}K\sin\Delta mt\right\}. \tag{7.20}$$

The time dependent asymmetry is now found to be

$$a_{f_{CP}} = \frac{(1-|\lambda|^2)}{(1+|\lambda|^2)}\cos\Delta mt - \frac{2\mathcal{I}m(\lambda)}{(1+|\lambda|^2)}K\sin\Delta mt = C\cos\Delta mt - S\sin\Delta mt \tag{7.21}$$

where:

$$\begin{aligned}
C &= \frac{(1-|\lambda|^2)}{(1+|\lambda|^2)} \\
S &= K\frac{2\mathcal{I}m(\lambda)}{(1+|\lambda|^2)} = -K\sin(2\beta).
\end{aligned} \tag{7.22}$$

This is to be compared with Equation 7.3 which represents the result for CP

eigenstates like $B^0 \to J/\psi K_S^0$.[4] Because the $B^0 \to D^{*+}D^{*-}$ final state is not

---

[4]Notice that the effective CP eigenvalue for $B^0 \to D^{*+}D^{*-}$ is dependent on the value of K

($R_\perp$) and therefore determines the sign of S. If the final state is mostly CP-even then $\eta_{CP} \approx +1$,

a pure *CP* eigenstate, the measurement of the $\mathcal{I}m\lambda$ will be diluted by a factor $K = 1 - 2R_\perp$. An undiluted measurement requires the information from $\cos\theta_{\text{tr}}$ so that K can be simultaneously determined in the fit for $\mathcal{I}m\lambda$ and $|\lambda|$.

In the above derivation, we have neglected the possibility that penguin diagrams could contribute differently to each transversity amplitude. The more appropriate and model-independent representation of the amplitudes is the following:

$$
\begin{aligned}
A_{0,\|}(t) &= A^0_{0,\|}\, e^{-imt}\, e^{-\Gamma t/2}\left(\cos\frac{\Delta mt}{2} + i\lambda_{0,\|}\sin\frac{\Delta mt}{2}\right) \\
A_\perp(t) &= A^0_\perp\, e^{-imt}\, e^{-\Gamma t/2}\left(\cos\frac{\Delta mt}{2} - i\lambda_\perp\sin\frac{\Delta mt}{2}\right).
\end{aligned}
\tag{7.23}
$$

where we now allow for different λ_f's for each transversity state. The time-dependent decay rate in Equation 7.18 is therefore rewritten as

$$
\begin{aligned}
\frac{1}{\Gamma}\frac{\mathrm{d}^2\Gamma(\overset{(-)}{B^0}\!\to D^{*+}D^{*-})}{\mathrm{d}z\mathrm{d}t} = \;&\frac{e^{-\Gamma t}}{|A^0_0|^2 + |A^0_\||^2 + |A^0_\perp|^2}\Big\{ \\
&[O_+ G_+(z) + O_- G_-(z)] \\
\overset{(-)}{+}\;&[C_+ G_+(z) + C_- G_-(z)]\cos\Delta mt \\
\overset{(+)}{-}\;&[S_+ G_+(z) - S_- G_-(z)]\sin\Delta mt\Big\}
\end{aligned}
\tag{7.24}
$$

The angular dependence of $G_+(z)$ and $G_-(z)$ was defined in Equation (7.15); the

while if it is mostly *CP*-odd then $\eta_{CP}\approx -1$.

six coefficients O_+, O_-, C_+, C_-, S_+ and S_- are given by:

$$
\begin{aligned}
O_+ &= |A_\parallel^0|^2 + |A_0^0|^2 = \frac{1}{2}(1 + K) \\
O_- &= |A_\perp^0|^2 = \frac{1}{2}(1 - K) \\
C_+ &= \frac{(1 - |\lambda_\parallel|^2)}{(1 + |\lambda_\parallel|^2)}|A_\parallel^0|^2 + \frac{(1 - |\lambda_0|^2)}{(1 + |\lambda_0|^2)}|A_0^0|^2 \\
C_- &= \frac{(1 - |\lambda_\perp|^2)}{(1 + |\lambda_\perp|^2)}|A_\perp^0|^2 \\
S_+ &= \left[\frac{2\mathcal{I}m(\lambda_\parallel)}{(1 + |\lambda_\parallel|^2)}|A_\parallel^0|^2 + \frac{2\mathcal{I}m(\lambda_0|)}{(1 + |\lambda_0|^2)}|A_0^0|^2 \right] \\
S_- &= \frac{2\mathcal{I}m(\lambda_\perp)}{(1 + |\lambda_\perp|^2)}|A_\perp^0|^2.
\end{aligned}
\tag{7.25}
$$

We can now define the CP-even parameters $\mathcal{I}m(\lambda_+)$ and $|\lambda_+|$ as

$$
\begin{aligned}
\frac{(1 - |\lambda_+|^2)}{(1 + |\lambda_+|^2)} &= \frac{\frac{(1-|\lambda_\parallel|^2)}{(1+|\lambda_\parallel|^2)}|A_\parallel^0|^2 + \frac{(1-|\lambda_0|^2)}{(1+|\lambda_0|^2)}|A_0^0|^2}{|A_\parallel^0|^2 + |A_0^0|^2} \\
\frac{\mathcal{I}m(\lambda_+)}{(1 + |\lambda_+|^2)} &= \frac{\frac{\mathcal{I}m(\lambda_\parallel)}{(1+|\lambda_\parallel|^2)}|A_\parallel^0|^2 + \frac{\mathcal{I}m(\lambda_0)}{(1+|\lambda_0|^2)}|A_0^0|^2}{|A_\parallel^0|^2 + |A_0^0|^2}.
\end{aligned}
\tag{7.26}
$$

Equation (7.24) can then be rewritten in terms of the five relevant physical quantities $|\lambda_+|$, $\mathcal{I}m(\lambda_+)$, $|\lambda_\perp|$, $\mathcal{I}m(\lambda_\perp)$, and K as:

$$
\begin{aligned}
\frac{1}{\Gamma}\frac{\mathrm{d}^2\Gamma(\overset{(-)}{B^0}\to D^{*+}D^{*-})}{\mathrm{d}z\mathrm{d}t} = \;& \frac{9}{32\pi}e^{-\Gamma t}\Big\{ \\
& \left[\frac{1}{2}(1 + K)G_+(z) + \frac{1}{2}(1 - K)G_-(z) \right] \\
\overset{(-)}{+}\; & \left[\frac{1}{2}(1 + K)\frac{(1 - |\lambda_+|^2)}{(1 + |\lambda_+|^2)}G_+(z) \right. \\
& \left. + \frac{1}{2}(1 - K)\frac{(1 - |\lambda_\perp|^2)}{(1 + |\lambda_\perp|^2)}G_-(z) \right] \cos\Delta mt \\
\overset{(+)}{-}\; & \left[\frac{1}{2}(1 + K)\frac{2\mathcal{I}m(\lambda_+)}{(1 + |\lambda_+|^2)}G_+(z) \right. \\
& \left. - \frac{1}{2}(1 - K)\frac{2\mathcal{I}m(\lambda_\perp)}{(1 + |\lambda_\perp|^2)}G_-(z) \right] \sin\Delta mt \Big\}.
\end{aligned}
\tag{7.27}
$$

180

In the limit of negligible penguin contributions, and taking into account the definition (7.19) of the dilution factor K, one recovers Equation 7.18. Equation 7.27 is the probability density function (pdf) used for the signal distribution in the likelihood fit described in Chapter 8.[5] It explicitly accounts for the CP-odd and CP-even components of the final state. The dilution of $\mathcal{I}m\lambda$ is therefore removed by simultaneously fitting the $z = \cos\theta_{\mathrm{tr}}$ distribution and extracting the value of K. Further, the values of $|\lambda_+|$, $\mathcal{I}m(\lambda_+)$, $|\lambda_\perp|$, $\mathcal{I}m(\lambda_\perp)$ are independent of any model-dependent assumptions about penguin diagram contributions.

7.4.1   Modeling the Angular Resolution

In the time-integrated transversity analysis (Section 6.4) the resolution on θ_{tr} was found to be non-negligible and was parameterized via an angular resolution function. Similarly, the time-dependent CP-asymmetry analysis is also dependent on the θ_{tr} resolution, as seen in Equation 7.27, and is accounted for in a similar way.

Equation 7.27 can be viewed as a linear combination of the CP-even and CP-odd angular terms:

$$
\begin{aligned}
G_+(z) &= \frac{8\pi}{3}\left(1 - z^2\right) = \frac{8\pi}{3}\sin^2\theta_{\mathrm{tr}} \\
G_-(z) &= \frac{16\pi}{3}z^2 = \frac{16\pi}{3}\cos^2\theta_{\mathrm{tr}}.
\end{aligned}
$$

The angular resolution is accounted for by performing a convolution with the two $G(\cos\theta_{\mathrm{tr}})$ functions in the same fashion as the convolution of $P(\cos\theta')$ in

[5]We note that Equation 7.27 does not explicitly show the tagging dilution factors and the Δt resolution function convolution discussed in Sections 7.2.1, 7.3.2. The full pdfs used in the fit include these effects in the same form as in Equations 7.11, 7.12.

Equation 6.22:

$$\tilde{G}(z) = \frac{1}{\sqrt{1-z^2}} \sum_{n=-\infty}^{\infty} (-1)^n \int_{n\pi}^{(n+1)\pi} d\theta' \sin\theta' G(\cos\theta') f(\theta_{\rm tr}; \theta')$$

$$\approx \frac{1}{\sqrt{1-z^2}} \sum_{n=-1}^{1} (-1)^n \int_{n\pi}^{(n+1)\pi} d\theta' \sin\theta' G(\cos\theta') f(\theta_{\rm tr}; \theta') \qquad (7.28)$$

Again we note that $f(\theta_{\rm tr}; \theta')$ is a function of $\theta_{\rm tr}$, while the angular terms G are functions of $z = \cos(\theta_{\rm tr})$. We can, therefore, replace $G_+(z)$ and $G_-(z)$ with the convolved forms $\tilde{G}_+(z), \tilde{G}_-(z)$ (which can be determined analytically) in Equation 7.27. This will properly account for any biases caused by the experimental resolution on $\theta_{\rm tr}$.

The resolution function $f(\theta_{\rm tr}; \theta')$ is the same used in Equation 6.25 for the time-integrated transversity analysis. It is composed of the sum of two Gaussians plus an additional term to model mis-reconstructed signal events (where $\theta_{\rm tr}$ is completely uncorrelated with its true value). The two Gaussians are centered at zero and represent pure angular resolution effects with "core" and "wide" Gaussians. The additional term used to model events with a mis-reconstructed soft pion is a truncated Gaussian centered at $\pi/2$. See Section 6.4.3 for details and the parameterizations determined from Monte Carlo samples.

Measuring the $\cos\theta_{\rm tr}$ distribution and resolution represents the last ingredient to the CP asymmetry measurement. The likelihood fit and validation of this technique is discussed in Chapter 8.

Chapter 8

CP Asymmetry Fit Method and Results

In Chapter 7 the primary ingredients for measuring time-dependent CP asymmetries in $B^0 \rightarrow D^{*+}D^{*-}$ were outlined. Since the $D^{*+}D^{*-}$ final state is not a CP eigenstate, the $\cos\theta_{\rm tr}$ distribution is measured to eliminate the angular dilution on the CP parameters. The time difference $\Delta t = t_{\rm rec} - t_{\rm tag}$ between the two B decays is determined from the spacial separation caused by the asymmetric-energy configuration of the PEP-II beams. The flavor of the $B_{\rm rec}$ candidate at $t = t_{\rm tag}$ is determined from the decay of the other B from the coherence property of the $\Upsilon(4S)$ decay. Mistagging fractions (wrong flavor assignment) can not be estimated from the $B^0 \rightarrow D^{*+}D^{*-}$ sample and must be provided as input to the fit. The fully reconstructed B mesons in flavor eigenstates ($B_{\rm flav}$ sample) can be used to measure the mistag fractions as well as the Δt experimental resolution in the data.

In principle, one can measure the detector parameters with the $B_{\rm flav}$ sample and fix them in the analysis of the $D^{*+}D^{*-}$ sample. But this approach has the disadvantage that correlations between the CP parameters $(\mathcal{I}m\lambda_+, \mathcal{I}m\lambda_\perp)$ and the detector parameters can result in complicated systematic uncertainties. A better approach, which is used in this analysis, is to perform a maximum-likelihood fit to the Δt distributions of the $B_{\rm flav}$ and $D^{*+}D^{*-}$ samples, simultaneously. The

former are used to measure the detector parameters, while the latter constrains the values of $|\lambda_+|$, $\mathcal{I}m(\lambda_+)$, $|\lambda_\perp|$, $\mathcal{I}m(\lambda_\perp)$. The advantage of this approach is that correlations among all parameters are properly taken into account and become part of the statistical uncertainty.

This chapter details the likelihood fit used to measure the CP parameters in $B^0 \to D^{*+}D^{*-}$. The likelihood functions and validation procedures are discussed followed by the results of the fit to data.

8.1   Likelihood Fit Method

The probability distribution functions for the $B^0 \to D^{*+}D^{*-}$ decay are given as

$$
\begin{aligned}
f_\pm(\Delta t_{\mathrm{true}}) &= \frac{\Gamma}{4}\mathrm{e}^{-\Gamma|\Delta t_{\mathrm{true}}|}\left\{(O(1+\tfrac{1}{2}\Delta\mathcal{D})\right. \\
&\qquad \left. \pm\langle\mathcal{D}\rangle\left[C\cos\left(\Delta m_d\Delta t_{\mathrm{true}}\right) - S\sin\left(\Delta m_d\Delta t_{\mathrm{true}}\right)\right]\right\},
\end{aligned}
\tag{8.1}
$$

where the $+$ sign refers to B_{tag} tagged as a B^0 and the $-$ sign is for B_{tag} tagged as a $\overline{B}^0$. The coefficients O, C and S are written as (see Equation 7.27):

$$
\begin{aligned}
O &= \frac{9}{32\pi}\left[\frac{1}{2}(1+K)\tilde{G}_+(z) + \frac{1}{2}(1-K)\tilde{G}_-(z)\right] \\
C &= \frac{9}{32\pi}\left[\frac{1}{2}(1+K)\tilde{G}_+(z)\frac{(1-|\lambda_+|^2)}{(1+|\lambda_+|^2)} + \frac{1}{2}(1-K)\tilde{G}_-(z)\frac{(1-|\lambda_\perp|^2)}{(1+|\lambda_\perp|^2)}\right] \\
S &= \frac{9}{32\pi}\left[\frac{1}{2}(1+K)\tilde{G}_+(z)\frac{2\mathcal{I}m(\lambda_+)}{(1+|\lambda_+|^2)} - \frac{1}{2}(1-K)\tilde{G}_-(z)\frac{2\mathcal{I}m(\lambda_\perp)}{(1+|\lambda_\perp|^2)}\right]
\end{aligned}
\tag{8.2}
$$

where the angular resolution is accounted for in the definitions of $\tilde{G}_+(z)$ and $\tilde{G}_-(z)$ described in Section 7.4.1. The finite resolution of Δt is taken into account by convolving $f_\pm$ with the Δt resolution function $\mathcal{R}(\delta_{\mathrm t} = \Delta t - \Delta t_{\mathrm{true}}; \hat{a})$:

$$
\mathcal{F}_\pm(\Delta t; \Gamma, \Delta m_d, \mathcal{D}_i, O, C, S, \hat{a}) = f_\pm(\Delta t; \Gamma, \Delta m_d, \mathcal{D}_i, O, C, S) \otimes \mathcal{R}(\delta_{\mathrm t}; \hat{a}),
\tag{8.3}
$$

where $\hat{a}$ represents the set of parameters that describe the resolution function (discussed in Section 7.3.2). Events are separated into four different tagging categories,

184

each of which has a different mean mistag fraction, $\langle w \rangle_i$, determined individually for each category.

In addition to signal parameterizations, the likelihood definition must also parameterize backgrounds events. These events are characterized as either peaking or combinatorial background. The probabilities $f^{CP}_{i,\text{sig}}$, $f^{CP}_{i,\text{peak}}$, and $f^{CP}_{i,\text{comb}}$ for an event to be signal or background are estimated from fits to m_{ES} distributions as described in Section 4.5. Fits are performed separately for each tagging category and the probabilities satisfy the constraint $f^{CP}_{i,\text{sig}} + f^{CP}_{i,\text{peak}} + f^{CP}_{i,\text{comb}} = 1$. The probability distribution functions $\mathcal{P}_{\pm,i}$, for the events in the $D^{*+}D^{*-}$ sample are therefore expressed as a sum of three contributions

$$\mathcal{P}_{\pm,i} = f^{CP}_{i,\text{sig}}\mathcal{F}_\pm + f^{CP}_{i,\text{peak}}\mathcal{B}^{CP}_{\pm,i,\text{peak}} + f^{CP}_{i,\text{comb}}\mathcal{B}^{CP}_{\pm,i,\text{comb}} \ , \tag{8.4}$$

where $\mathcal{F}_\pm$ are the signal components (shown in Equation 8.3), $\mathcal{B}^{CP}_{\pm,i,\text{peak}}$ represent the contributions from peaking background, and $\mathcal{B}^{CP}_{\pm,i,\text{comb}}$ are the combinatorial-background components. Though the $D^{*+}D^{*-}$ final state is estimated to have negligible peaking background, the term is included for systematic studies.

The likelihood function for $D^{*+}D^{*-}$ events is then written as

$$\ln\mathcal{L}_{CP} = \sum_i^{N_c}\left[\sum_{B^0 \text{ tag}}\ln\mathcal{P}_{+,i} + \sum_{\overline{B}^0 \text{ tag}}\ln\mathcal{P}_{-,i}\right] + \sum_{\text{untagged}}\ln\mathcal{P}^{\mathcal{D}_i=0}_{\pm,\text{untag}} \tag{8.5}$$

where $N_c = 4$ is the number of tagging categories, and $\mathcal{P}_{+,i}$ ($\mathcal{P}_{-,i}$) is the pdf for events in the ith tagging category for $B_{\text{tag}} \equiv B^0$ ($B_{\text{tag}} \equiv \overline{B}^0$). The four tagging categories are mutually exclusive which means that each event can only belong to one category, and can be tagged as either a B^0 or a $\overline{B}^0$. Events without a flavor tag can not be used to measure the CP parameters ($\mathcal{I}m\lambda_+, \mathcal{I}m\lambda_\perp$), but can be used to constrain K. The last term of Equation 8.5 is only sensitive to K since the tagging dilutions are set to zero for untagged events (leaving only the O coefficient in the signal pdf, Equation 8.1).

185

8.1.1 Modeling Background Events

Each event is assigned a probability to be signal on the basis of its measured energy-constrained mass m_{ES}, as described in Section 4.5. The m_{ES} distribution is described with a single Gaussian distribution $\mathcal{G}(m_{\text{ES}})$ for the signal and an ARGUS parameterization $\mathcal{A}(m_{\text{ES}})$ for the background. The probabilities that appear in Equation (8.4) are defined as

$$
\begin{aligned}
f_{i,\text{sig}}(m_{\text{ES}}) &= \frac{(1 - \delta_{\text{peak}})\mathcal{G}(m_{\text{ES}})}{\mathcal{G}(m_{\text{ES}}) + \mathcal{A}(m_{\text{ES}})} \\
f_{i,\text{peak}}(m_{\text{ES}}) &= \frac{\delta_{\text{peak}}\mathcal{G}(m_{\text{ES}})}{\mathcal{G}(m_{\text{ES}}) + \mathcal{A}(m_{\text{ES}})} \\
f_{i,\text{comb}}(m_{\text{ES}}) &= \frac{\mathcal{A}(m_{\text{ES}})}{\mathcal{G}(m_{\text{ES}}) + \mathcal{A}(m_{\text{ES}})}
\end{aligned}
\tag{8.6}
$$

The fraction δ_{peak} accounts for the peaking-background contribution and is nominally set to zero in the fit.

Backgrounds arise from many different sources. Rather than describing the Δt distribution of each physics process that contributes, an empirical description is used in the fit which allows for different time dependencies. Peaking background is likely to result from real B decays and is parameterized as

$$
\mathcal{B}^{CP}_{\pm,i,\text{peak}} = \frac{\Gamma^{CP}_{\text{peak}}}{4} e^{-\Gamma^{CP}_{\text{peak}}|\Delta t_{\text{true}}|} (1 \pm \langle \mathcal{D} \rangle_i \eta_{\text{peak}} \sin \Delta m_d \Delta t_{\text{true}}) \otimes \mathcal{R}(\delta_{\text{t}}; \hat{a}_i) , \tag{8.7}
$$

where dilutions $\langle \mathcal{D} \rangle_i$ and resolution function parameters $\hat{a}_i$ are the same as those used for the signal, $1/\Gamma^{CP}_{\text{peak}}$ is an empirical lifetime, and η_{peak} is the effective CP eigenvalue (which is varied to estimate the systematic uncertainty).

The Δt spectrum of the combinatorial background is modeled by a sum of two distributions:

$$
\mathcal{B}^{CP}_{\pm,i,comb} = f^{CP}_{i,1} \mathcal{B}_{\pm,i,1}(\Delta t; \hat{b}_i) + (1 - f^{CP}_{i,1})\mathcal{B}_{\pm,i,2}(\Delta t; \hat{b}_i) \tag{8.8}
$$

where the components are defined as

$$\mathcal{B}^{CP}_{\pm,i,1} = \frac{1}{2}\delta(\Delta t_{\text{true}}) \otimes \mathcal{R}(\delta t; \hat{b}_i) \,, \tag{8.9}$$

$$\mathcal{B}^{CP}_{\pm,i,2} = \frac{1}{4}\Gamma^{CP}_2 e^{-\Gamma^{CP}_2 |\Delta t_{\text{true}}|}\left(1 \pm \langle\mathcal{D}\rangle_{i,2}\, \eta^{\text{comb}} \sin \Delta m_d \Delta t_{\text{true}}\right) \otimes \mathcal{R}(\delta t; \hat{b}_i) \tag{8.10}$$

Here, $1/\Gamma^{CP}_2$ is an empirical lifetime, $f^{CP}_{i,1}$ is the fraction of events in the "prompt-lifetime" component, η^{comb} is an effective CP eigenvalue, and $\hat{b}_i$ are the resolution function parameters for the background events which are determined from the background resolution function of the B_{flav} sample. These parameterizations allow for different time-dependencies; background from B decays, for example, will likely follow the $\mathcal{B}^{CP}_{\pm,i,2}$ distribution, while background from continuum events will exhibit no lifetime and correspond to the prompt component.

The parameterization of the combinatorial background (Equation 8.8) is simplified by setting $\eta^{\text{comb}} = 0$ in the nominal fit. This implies that no CP asymmetry is expected in the background events. In addition, a common fraction $f^{CP}_{i,1}$ for the prompt component, and a common empirical lifetime $1/\Gamma^{CP}_2$ for the non-prompt component are used for all tagging categories. The fraction is allowed to float in the fit, while $1/\Gamma^{CP}_2$ is fixed to the B^0 lifetime τ_{B^0}. The systematic uncertainties due to the these assumptions are evaluated in Sections 8.4.3, 8.4.5.

8.1.2 Likelihood Function for Flavor Eigenstates

Events with a fully-reconstructed B meson in flavor eigenstates are used to determine tagging dilutions and Δt resolution function parameters for both signal ($\hat{a}_i$) and background events ($\hat{b}_i$). The flavor of the fully reconstructed B meson (B_{flav}) is known from the reconstructed final state, and the flavor of the other B meson (B_{tag}) is determined by the tagging algorithm discussed in Section 7.2. Since the flavor of both B mesons is known, events can be divided into two categories de-

pending on whether the B_flav changed flavor between the time of its decay and the time of the B_tag decay:

- **Unmixed events**: the B mesons have different flavors, that is $|B_\text{flav}, B_\text{tag}\rangle$ is either $|B^0, \overline{B}^0\rangle$ or $|\overline{B}^0, B^0\rangle$;

- **Mixed events**: the two B mesons have the same flavor, that is $|B_\text{flav}, B_\text{tag}\rangle$ is either $|B^0, B^0\rangle$ or $|\overline{B}^0, \overline{B}^0\rangle$.

After including mistagging rates $(w, \overline{w})$ and simplifying the expressions, the observed Δt distributions are given by [61]

$$h'_\text{unmix}(\Delta t; B_\text{flav} B_\text{tag} \equiv \overline{B}^0 B^0) = \frac{1}{4}\Gamma\, e^{-\Gamma|\Delta t|}\Big(1 + \frac{\Delta\mathcal{D}_i}{2} + \langle\mathcal{D}\rangle_i \cos\Delta m_d\Delta t\Big)\,(8.11)$$

$$h'_\text{unmix}(\Delta t; B_\text{flav} B_\text{tag} \equiv B^0 \overline{B}^0) = \frac{1}{4}\Gamma\, e^{-\Gamma|\Delta t|}\Big(1 - \frac{\Delta\mathcal{D}_i}{2} + \langle\mathcal{D}\rangle_i \cos\Delta m_d\Delta t\Big)\,(8.12)$$

$$h'_\text{mix}(\Delta t; B_\text{flav} B_\text{tag} \equiv B^0 B^0) = \frac{1}{4}\Gamma\, e^{-\Gamma|\Delta t|}\Big(1 + \frac{\Delta\mathcal{D}_i}{2} - \langle\mathcal{D}\rangle_i \cos\Delta m_d\Delta t\Big)\,(8.13)$$

$$h'_\text{mix}(\Delta t; B_\text{flav} B_\text{tag} \equiv \overline{B}^0 \overline{B}^0) = \frac{1}{4}\Gamma\, e^{-\Gamma|\Delta t|}\Big(1 - \frac{\Delta\mathcal{D}_i}{2} - \langle\mathcal{D}\rangle_i \cos\Delta m_d\Delta t\Big)\,(8.14)$$

The probability density functions (PDFs) for the signal component are given by the convolution of (8.11)-(8.14) with the resolution function $\mathcal{R}(\delta t; \hat{a}_i)$ (same as in Section 7.3.2)

$$\mathcal{H}_+\left(\Delta t; \Gamma, \Delta m_d, \langle\mathcal{D}\rangle_i, \Delta\mathcal{D}_i, \hat{a}_i, B^0_\text{tag}\right) = h'_\text{unmix}\left(\Delta t_\text{true}; B^0_\text{tag}\right) \otimes \mathcal{R}(\delta_t; \hat{a}_i)\,(8.15)$$

$$\mathcal{H}_+\left(\Delta t; \Gamma, \Delta m_d, \langle\mathcal{D}\rangle_i, \Delta\mathcal{D}_i, \hat{a}_i, \overline{B}^0_\text{tag}\right) = h'_\text{unmix}\left(\Delta t_\text{true}; \overline{B}^0_\text{tag}\right) \otimes \mathcal{R}(\delta_t; \hat{a}_i)\,(8.16)$$

$$\mathcal{H}_-\left(\Delta t; \Gamma, \Delta m_d, \langle\mathcal{D}\rangle_i, \Delta\mathcal{D}_i, \hat{a}_i, B^0_\text{tag}\right) = h'_\text{mix}\left(\Delta t_\text{true}; B^0_\text{tag}\right) \otimes \mathcal{R}(\delta_t; \hat{a}_i)\quad(8.17)$$

$$\mathcal{H}_-\left(\Delta t; \Gamma, \Delta m_d, \langle\mathcal{D}\rangle_i, \Delta\mathcal{D}_i, \hat{a}_i, \overline{B}^0_\text{tag}\right) = h'_\text{mix}\left(\Delta t_\text{true}; \overline{B}^0_\text{tag}\right) \otimes \mathcal{R}(\delta_t; \hat{a}_i)\quad(8.18)$$

These Δt distributions are illustrated in Figure 8.1.
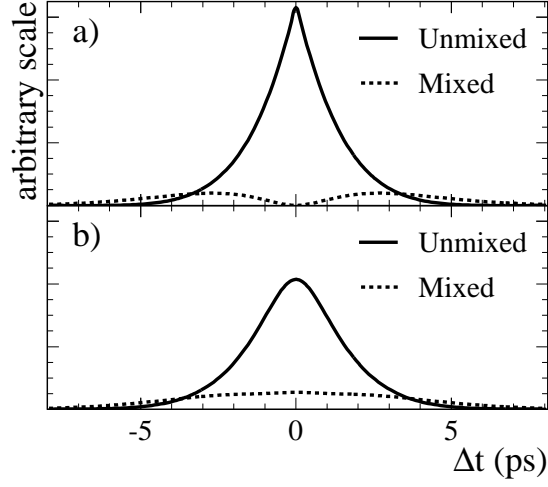
Figure 8.1: The Δt distribution for mixed and unmixed events with a) perfect tagging and Δt resolution, and b) typical mistag fractions and Δt resolution. The scale is arbitrary but is the same for the two plots.

The likelihood function for events in the B_{flav} sample is written, analogous to the $D^{*+}D^{*-}$ sample (see Equation 8.5), as

$$\ln \mathcal{L}_{\text{flav}} = \sum_{i}^{N_c} \left[\sum_{\text{unmixed}} \ln \mathcal{J}_{+,i} + \sum_{\text{mixed}} \ln \mathcal{J}_{-,i} \right], \tag{8.19}$$

where $\mathcal{J}_{+,i}$ and $\mathcal{J}_{-,i}$ are, respectively, the probability density functions for unmixed and mixed events in i^{th} tagging category. The likelihood functions $\mathcal{J}_{\pm,i}$ are defined as a sum of signal, peaking background, and combinatorial background components

$$\mathcal{J}_{\pm,i} = f_{i,\text{sig}}^{\text{flav}} \mathcal{H}_{\pm} + f_{i,\text{peak}}^{\text{flav}} \mathcal{B}_{\pm,i,\text{peak}}^{\text{flav}} + f_{i,\text{comb}}^{\text{flav}} \mathcal{B}_{\pm,i,\text{comb}}^{\text{flav}} . \tag{8.20}$$

The treatment of the Δt distribution of background events in the B_{flav} sample is similar to the method described in Section 8.1.1 for the $D^{*+}D^{*-}$ sample. A similar empirical description (see Equations 8.9, 8.10) is used to model the Δt distributions of the background events. Thus, from the B_{flav} sample, the signal and background tagging dilution parameters $\langle \mathcal{D} \rangle_i, \Delta \mathcal{D}_i$ are extracted as well as the

189

signal and background Δt resolution function parameters ($\hat{a}_i$ and $\hat{b}_i$), all of which are used in the likelihood function definition of the $D^{*+}D^{*-}$ sample. (Reference [61] should be consulted for more detail on the B_{flav} sample fit procedure.)

8.1.3   Simultaneous Fit to B_{flav} and $D^{*+}D^{*-}$ Samples

Using the components described above, we extract the values of the CP parameters $|\lambda_+|$, $\mathcal{I}m(\lambda_+)$, $|\lambda_\perp|$, $\mathcal{I}m(\lambda_\perp)$ from an unbinned likelihood fit to the Δt and $\cos\theta_{\text{tr}}$ distributions of events reconstructed in the $D^{*+}D^{*-}$ final state and events with a fully reconstructed B meson in flavor eigenstates. The function maximized in the fit is given by

$$\ln \mathcal{L}_{\text{tot}} = \ln \mathcal{L}_{CP} + \ln \mathcal{L}_{\text{flav}} \ , \tag{8.21}$$

where $\ln \mathcal{L}_{CP}$ and $\ln \mathcal{L}_{\text{flav}}$ are defined in Equations 8.5 and 8.19.

A total of 38 parameters are varied in the fit, as shown in Table 8.1. A number of

Description	Number of parameters
CP-even asymmetry parameters $\mathcal{I}m\lambda_+, |\lambda_+|$	2
CP-odd asymmetry parameters $\mathcal{I}m\lambda_\perp, |\lambda_\perp|$	fixed
Angular dilution factor K	1
Signal Δt resolution function	9
Signal dilutions for each tagging category	8
Background dilutions for each tagging category	8
Background Δt resolution function	3
Background composition	7
Total	38

Table 8.1: Summary of the floating parameters in the maximum-likelihood fit.

other parameters, e.g. oscillation frequency Δm_d and peaking background fraction, are used as input to the fit and their values are fixed. The signal Δt resolution function parameters floating in the fit are listed in Table 7.3. The signal and background dilution parameters in the $D^{*+}D^{*-}$ likelihood function are determined from the B_{flav} sample likelihood function. The background Δt resolution function is common to both samples and is parameterized by core and outlier Gaussians, only. The fraction $f_{\text{outl}}^{\text{bkgd}}$ between the two components is allowed to float as well as the width and mean-offset parameters of the core Gaussian. The width and offset of the outliers component are fixed, respectively, to 8 ps and 0 ps similar to the signal resolution function. The background composition parameters include the empirical lifetime $1/\Gamma_2^{CP}$, the fraction of prompt background $f_{i,1}^{CP}$ for the $D^{*+}D^{*-}$ sample, and the category-specific prompt-background fractions for the B_{flav} sample.

Because of the results found in Chapter 6, the value of $R_\perp$, and therefore, K, is expected to be small. It follows that the sensitivity to the CP-odd parameters $\mathcal{I}m\lambda_\perp$ and $|\lambda_\perp|$ will be very poor, because of the small number of events in the CP-odd ($G_-(z)$) component of the distribution (also see Equation 7.27). We choose, therefore, to fix the values of $\mathcal{I}m\lambda_\perp$ and $|\lambda_\perp|$ in the fit to the expected values from the Standard Model (-0.75 and 1.0, respectively). The variations on the fitted values of $\mathcal{I}m\lambda_+$ and $|\lambda_+|$ for different input values of $\mathcal{I}m\lambda_\perp$ and $|\lambda_\perp|$ are evaluated as systematic uncertainties.

Lastly, it is noted that no angular acceptance effects are accounted for in the likelihood definition. This differs from the time-integrated transversity analysis method. The value of K extracted is therefore considered to be an *effective* value that is expected to be different than the acceptance-corrected value of $R_\perp$ determined in Chapter 6. A discussion of the comparison of values is left to Section 8.4.8. The effect of detector acceptance on the values of $\mathcal{I}m\lambda_+$ and $|\lambda_+|$ is also studied

and evaluated as a systematic uncertainty.

8.2 Validation Studies

8.2.1 Studies using Toy Monte Carlo Samples

The performance of the likelihood fit was validated using toy[1] Monte Carlo samples. The samples were generated with 1000 times the number of $D^{*+}D^{*-}$ events in the data using the same tagging dilutions, time resolution functions parameters and background fractions expected in the data. Several "experiments" were generated each with different values of $\mathcal{I}m(\lambda_+)$, $|\lambda_+|$ and $\mathcal{I}m(\lambda_\perp)$, $|\lambda_\perp|$. The fit was then performed allowing the four CP parameters and K to float.

Figure 8.2 shows the linearity of the fit of when the input (generated) values of $\mathcal{I}m(\lambda_+)$ and $|\lambda_+|$ were varied in the ranges $[-1., 1.]$ and $[0.7, 1.3]$, respectively.

Figure 8.3 shows the fitted values of $\mathcal{I}m(\lambda_+)$, $|\lambda_+|$, $\mathcal{I}m(\lambda_\perp)$, $|\lambda_\perp|$ and K as a function of the generated value of K. These figures demonstrate the linearity of the fit and it's consistency for all generated values of the parameters. We note that $\mathcal{I}m(\lambda_+)$ $(\mathcal{I}m(\lambda_\perp))$ is poorly determined when K is close to $+1.$ $(-1.)$; this behavior is expected since the value of K determines the distribution of events in the two CP-parity distributions $(G_\pm(z))$.

To validate the convergence properties of the fit for a low statistics sample, multiple toy samples were generated with 310 events each. This corresponds to the data where about 126 signal events (84 tagged) and 184 background events

---

[1]A "toy" Monte Carlo is defined as a sample produced by random number generation where the distributions follow the probability distribution functions of the fit. These samples are completely different than the "fully-reconstructed" Monte Carlo which utilizes the full simulation of the detector and reconstruction techniques.
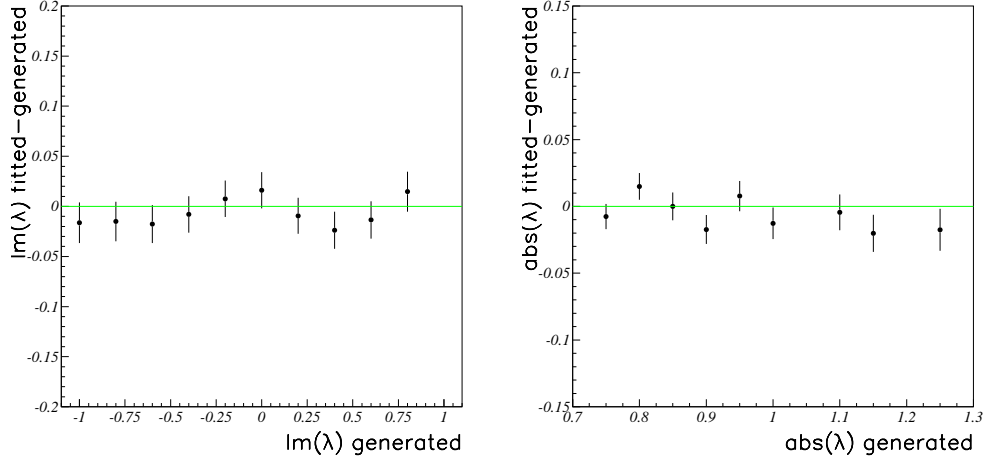
Figure 8.2: Difference between fitted and generated values for the $\mathcal{I}m(\lambda_+)$ (left) and $|\lambda_+|$ (right) parameters as a function of their generated values. In both cases the values are consistent with no bias with a normalized $\chi^2 = 0.56$ and 1.1 respectively.

(described by the Argus+Gaussian distribution of the m_{ES} projection) are expected. The values of the CP-odd parameters $\mathcal{I}m(\lambda_\perp)$ and $|\lambda_\perp|$ were fixed in the one hundred fitted experiments. None of the fits failed to converge.

8.2.2 Angular Acceptance Studies

For the measurement of the CP-odd fraction $R_\perp$ the detector acceptance was studied and the effect was included in the signal distribution used in the likelihood fit (see Section 6.3). For the time-dependent CP analysis, no explicit correction is included in the fit procedure. Therefore, the size of the effect on the fitted values of $\mathcal{I}m(\lambda_+)$, $|\lambda_+|$ is determined using toy Monte Carlo samples with different generated distributions.

The inefficiency in reconstructing the low momentum "soft" pions in the $B^0 \rightarrow$

193
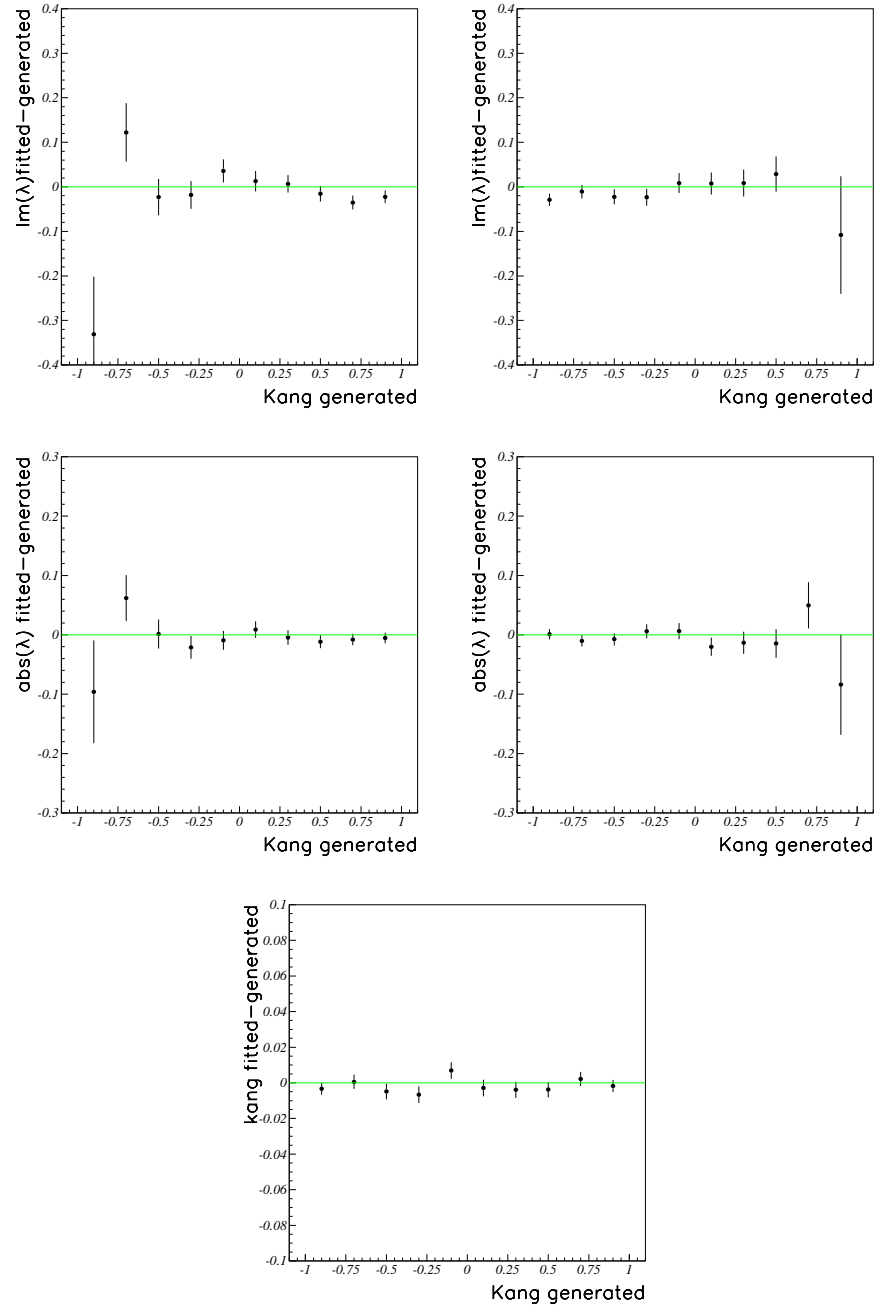
Figure 8.3: Difference between fitted and generated values for $\mathcal{I}m(\lambda_+)$ (top left), $\mathcal{I}m(\lambda_\perp)$ (top right) $|\lambda_\perp|$ (center left) and $|\lambda_\perp|$ (center right) and K (bottom) as a function of the generated value of K.

194

$D^{*+}D^{*-}$ decay is the primary cause of the acceptance effects. Figure 8.4 shows the
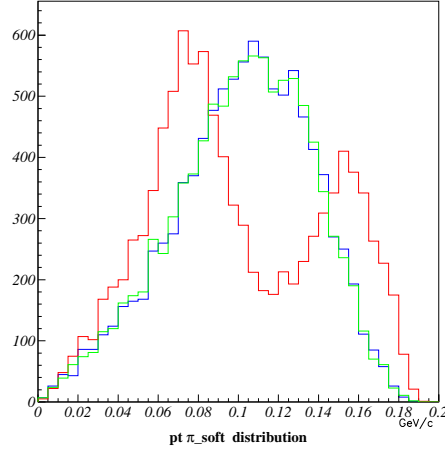
Figure 8.4: Soft pion transverse momentum distribution for pure A_0 final state (red curve) and pure $A_\parallel$ (blue) or pure $A_\perp$ (green).

soft pion transverse momentum (p_T) distributions for the three extreme values of the transversity amplitudes without detector acceptance included: $(A_0, A_\parallel, A_\perp) = (1, 0, 0)$ (red curve), $= (0, 1, 0)$ (blue curve) and $= (0, 0, 1)$ (green curve). The CP-odd amplitude, $A_\perp$ and one of the CP-even amplitudes, $A_\parallel$, show similar soft pion p_T distributions, while the other CP-even amplitude A_0 has a distribution which is more populated at low p_T values, and is therefore more sensitive to the detector acceptance.

The reconstruction efficiency as a function of the soft pion transverse momentum is shown in Figure 8.5 for two $B^0 \rightarrow D^{*+}D^{*-}$ submodes.

In order to study the effect of acceptance we generate large toy Monte Carlo samples with a reconstruction probability that follows the shape of the curves in Figure 8.5 for different input values of K. The results are shown in Figure 8.6. The first four plots show the fitted values of the CP parameters as a function of the gen-

195
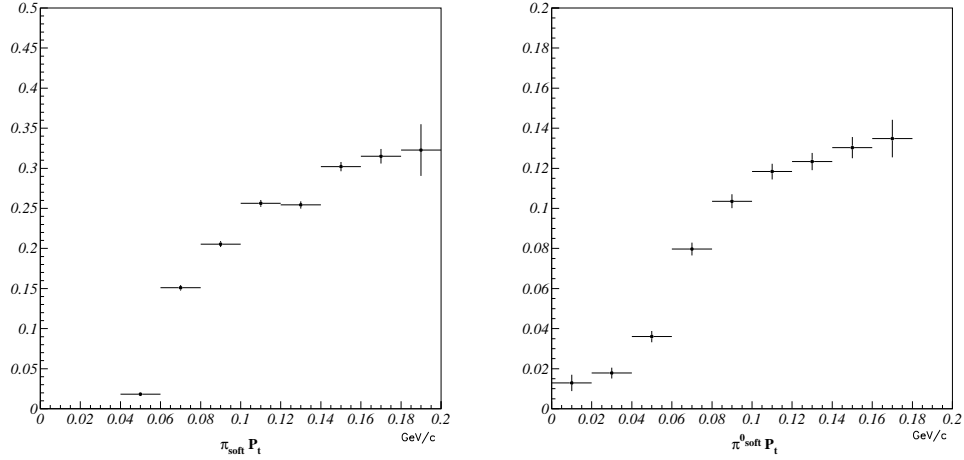
Figure 8.5: Global reconstruction efficiency for $B^0 \rightarrow D^{*+}D^{*-} \rightarrow D^0(K\pi)\pi^+\bar{D}^0(K\pi)\pi^-$ (left) and $B^0 \rightarrow D^{*+}D^{*-} \rightarrow D^0(K\pi)\pi^+D^-(K\pi\pi)\pi^0$ (right) as a function of charged and neutral soft pion transverse momentum.

erated value of K. The green line is a linear fit to the fit results and demonstrates the agreement between fitted and generated values ($\mathcal{I}m(\lambda_+) = \mathcal{I}m(\lambda_\perp) = -0.7033$ and $|\lambda_+| = |\lambda_\perp| = 1$.) The last plot shows the sensitivity of the variable K to the detector acceptance. Without correcting for the effect via acceptance moments (as in the time-integrated measurement of $R_\perp$) the value of K will be biased. The fitted value of K in the data is therefore, considered to be an effective value and is not quoted as a measurement of $R_\perp$.

Because the Monte Carlo may not model the soft pion reconstruction efficiency perfectly we also consider a more extreme case to validate the behavior of the CP fit. The four CP parameters are fit to a sample generated with sharp p_T cuts (similar to the test in Section 6.3.2). The results are shown in Figure 8.7, and demonstrate the values of $\mathcal{I}m(\lambda_+)$ and $|\lambda_+|$ are not biased by soft pion inefficiency for the threshold values expected in the data. No bias correction is therefore used
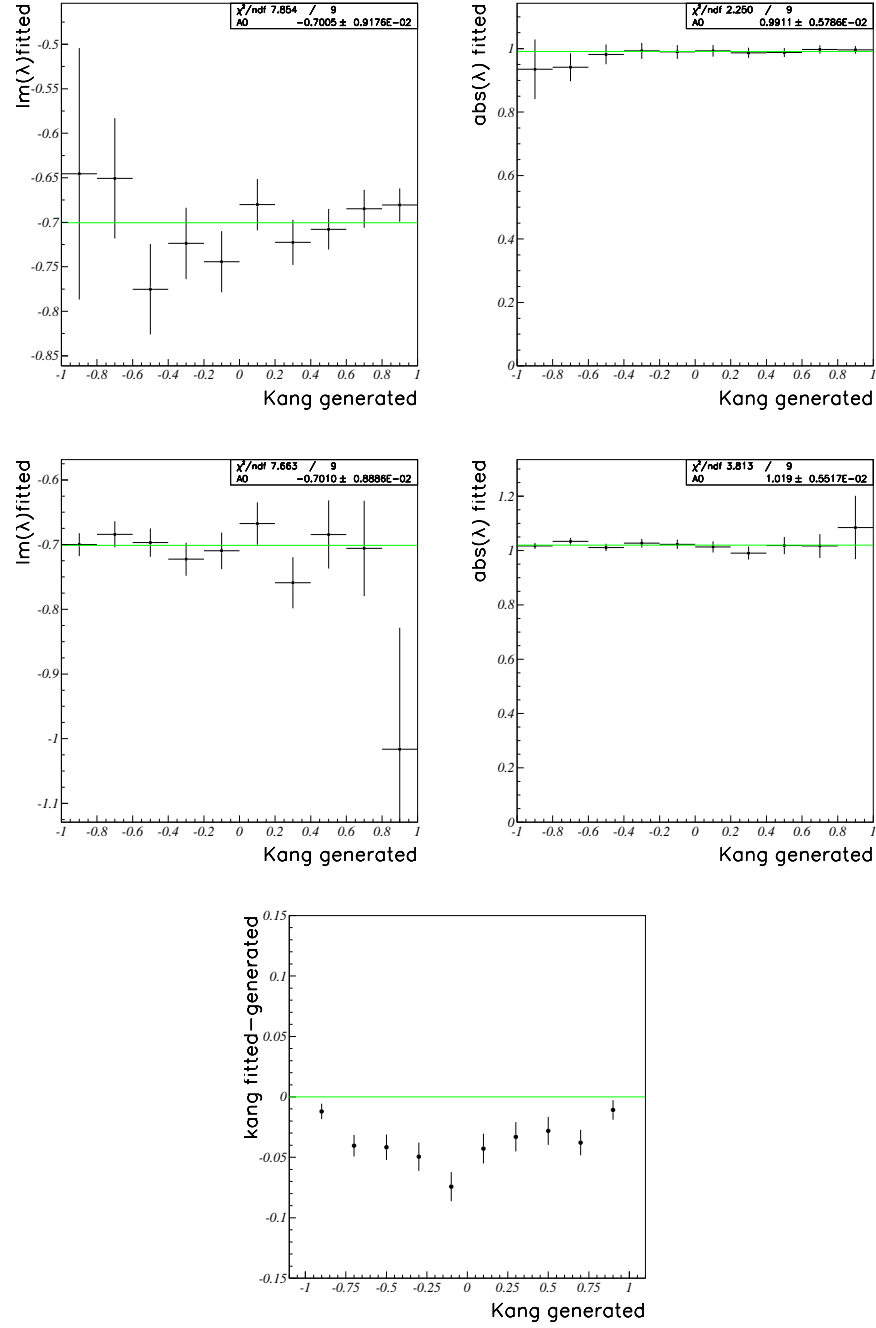
Figure 8.6: Using the soft pion efficiency distributions of Figure 8.5 for the generated toy samples, the fitted values of the CP parameters $\mathcal{I}m(\lambda_+)$ (top left), $|\lambda_+|$ (top right), $\mathcal{I}m(\lambda_\perp)$ (central left) $|\lambda_\perp|$ (central right) are shown as a function of the generated value of K. The bottom plot shows the fitted value of K and demonstrates the effect of detector acceptance.
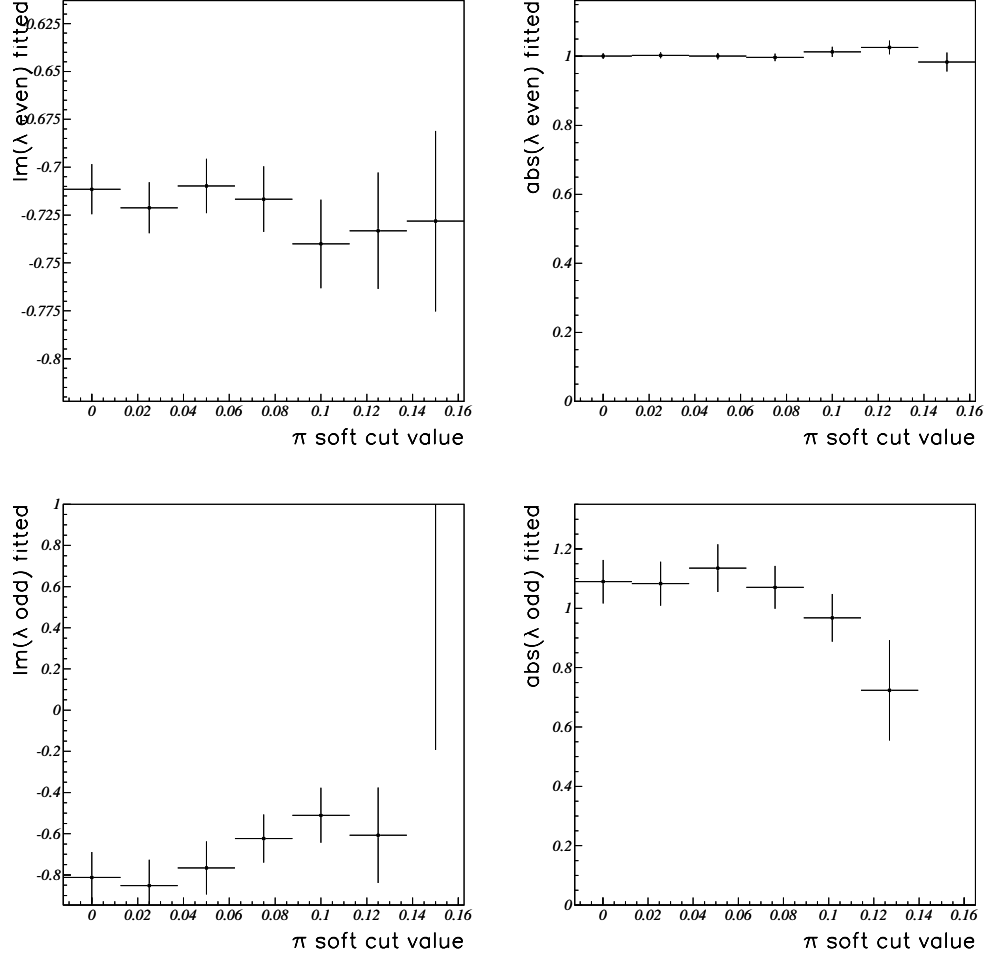
Figure 8.7: Fitted values of the CP parameters $\mathcal{I}m(\lambda_+)$ (top left), $|\lambda_+|$ (top right), $\mathcal{I}m(\lambda_\perp)$ (bottom left) $|\lambda_\perp|$ (bottom right) as a function of a threshold cut on the soft pion transverse momentum, p_T.

198

in the fit to the data, but the effect is evaluated as a systematic uncertainty on the values of $\mathcal{I}m(\lambda_+)$ and $|\lambda_+|$.

8.2.3 Angular Resolution Studies

Section 6.4 described in detail the $\theta_{\rm tr}$ angular resolution of reconstructed $B^0 \to D^{*+}D^{*-}$ events. The reconstructed $\theta_{\rm tr}$ distributions in Figures 6.15 and 6.16 show that about 7% of the signal events in $D^{*+}D^{*-} \to (D^0\pi^+, \overline{D}^0\pi^-)$ modes and about 20% of the events reconstructed in $D^{*+}D^{*-} \to (D^0\pi^+, D^-\pi^0)$ modes are mis-reconstructed ($\theta_{\rm tr}$ is uncorrelated with $\theta_{\rm tr}^{\rm true}$). The primary source of this effect is due to the mis-identification of at least one of the soft pions in the event.

The Δt distribution is much less affected, however. The $B_{\rm rec}$ vertex is primarily determined by the pseudo-tracks of the D mesons and uses little information from the soft pion tracks. Figure 8.8 shows the residual distributions of $\Delta t - \Delta t_{\rm true}$ for mis-reconstructed and well-reconstructed signal events in $(D^0\pi^+, \overline{D}^0\pi^-)$ modes (top) and $(D^0\pi^+, D^-\pi^0)$ modes (bottom) using fully reconstructed Monte Carlo samples. The Gaussian fit is not intended to be a good description of the distributions but instead is used to quantify how the distributions differ. The sigmas of the distributions agree statistically, but the negative bias is larger for the mis-reconstructed events. Because we take into account the average bias of the Δt resolution function in the systematic uncertainties, we do not evaluate an additional systematic uncertainty for this effect.

It is also of interest to study how the values of the CP parameters are impacted by the finite experimental resolution of $\theta_{\rm tr}$. Four toy Monte Carlo samples are generated using different possible "smearings" of the angle $\theta_{\rm tr}$ to mimic the detector's angular resolution. Figure 8.9 shows how the fitted values of CP parameters degrade as the resolution is changed from perfect to the distribution observed in the
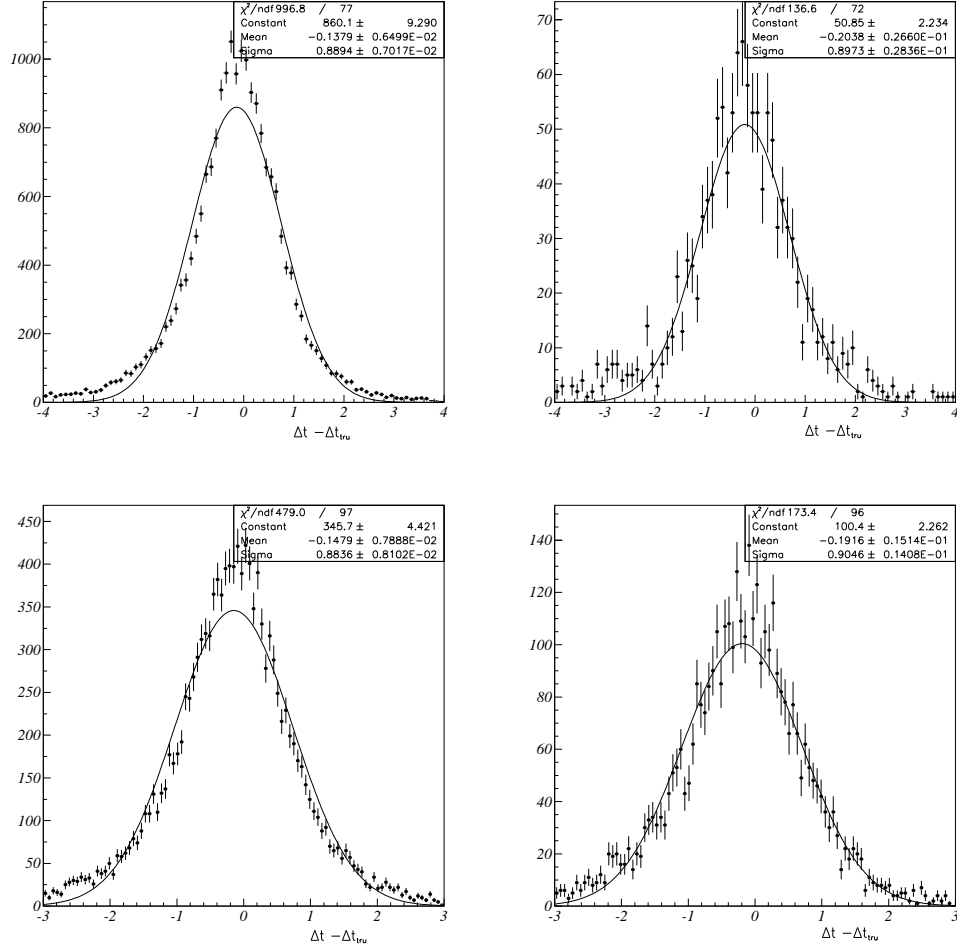
Figure 8.8: Distributions of $\Delta t - \Delta t_{true}$ for mis-reconstructed (right) and well-reconstructed (left) $B^0 \to D^{*+} D^{*-}$ signal events in decay modes with two charged soft pions (top) and in modes with a neutral soft pion and a charged soft pion (bottom).

Monte Carlo samples (see Figures 6.9 and 6.10).

The first toy sample is generated with perfect θ_{tr} resolution (left-most data points in Figure 8.9). The second sample is generated using a single Gaussian distribution with a width equal to σ_{core} from Figure 6.9 to smear the value of θ_{tr}.[2] The third point corresponds to a double-Gaussian smearing. The right-most point corresponds to the full triple-Gaussian smearing observed in Figure 6.9. For all of the fits the unconvolved form of the signal pdf (that is, the $G(z)$ functions do not include the angular resolution convolution described in Section 7.4.1) was used in order to observe the effect.

Figure 8.9 shows that the CP-odd parameters are more sensitive to the angular resolution. As the resolution degrades there is a migration of events from one CP-parity distribution ($G_{\pm}(z)$) to the other CP-parity distribution ($G_{\mp}(z)$). Since K is very close to one for these toy samples, the "leakage" of the CP-even component into the CP-odd distribution causes a larger effect than vice versa. The result is a dilution of the CP-odd asymmetry, while the effect on the CP-even asymmetry parameters is small.

However, the convolution described in Section 7.4.1 is able to correct for these biases. By generating toy Monte Carlo samples where the resolution is simulated by smearing the θ_{tr} distributions, the convolved forms of the signal pdf can be tested. Three different toy samples were generated with different values of K and θ_{tr} resolution similar to that observed in Figures 6.15 and 6.16. The fitted values of $\mathcal{I}m(\lambda_+)$ and $|\lambda_+|$ were found to be consistent with the generated values in every sample.

---

[2]Smearing refers to generating a random number according to the $\theta_{\mathrm{tr}} - \theta_{\mathrm{tr}}^{\mathrm{true}}$ distribution and adding it to the generated value of θ_{tr}.
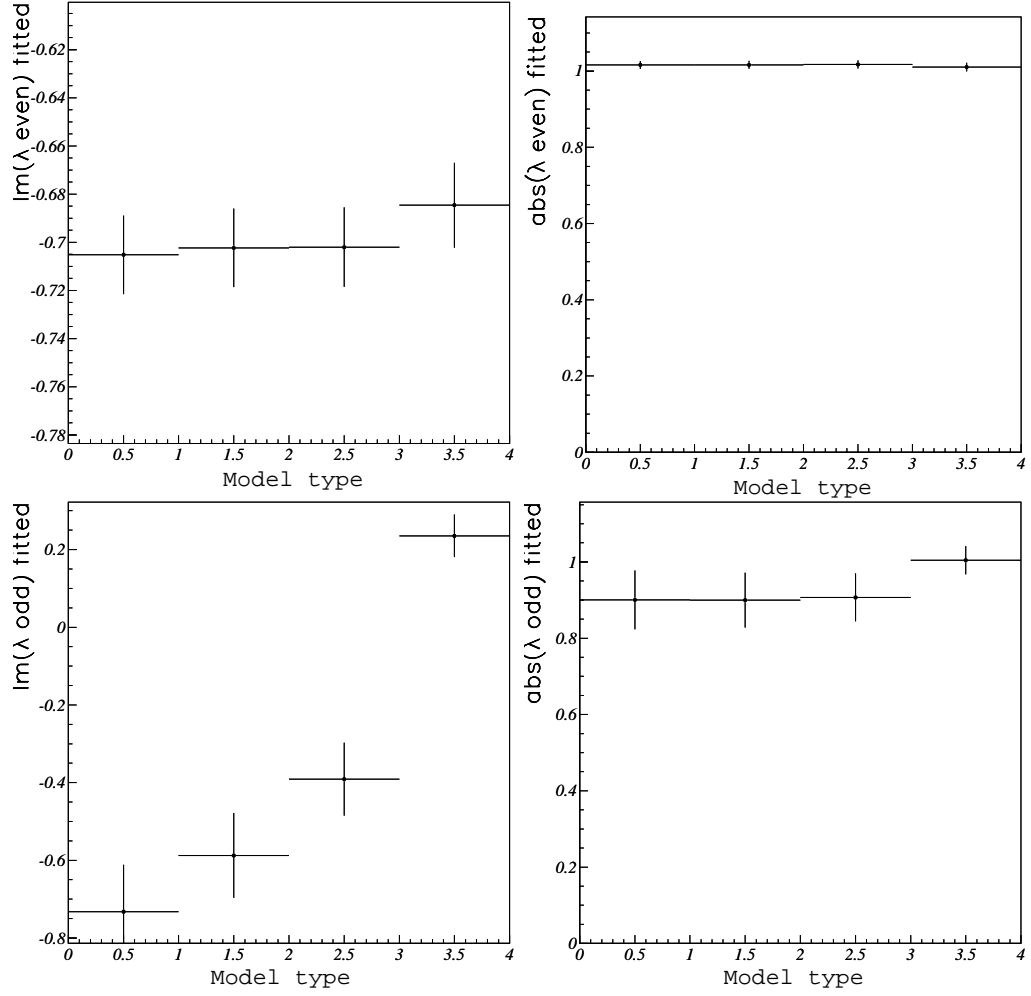
Figure 8.9: Fitted values of the CP parameters $\mathcal{I}m(\lambda_+)$ (top left), $|\lambda_+|$ (top right), $\mathcal{I}m(\lambda_-)$ (bottom left) $|\lambda_-|$ (bottom right) for the four different cases of generated resolution effects described in the text.

8.2.4 Validation from Fully Reconstructed Monte Carlo

Large samples of fully reconstructed signal Monte Carlo events are used to validate the fit procedure. The sample is divided into many smaller samples with the same number of events expected in the data; these multiple "experiments" are used to study the distribution of the mean values and statistical errors on the $\mathcal{I}m(\lambda_+)$ and $|\lambda_+|$ parameters. The generated values of the amplitudes were $(A_{\parallel}, A_0, A_{\perp}) = (0.62, 0.74, 0.25)$ corresponding to $K = 0.875$. The generated values of the CP parameters were $\mathcal{I}m(\lambda_+) = \mathcal{I}m(\lambda_{\perp}) = -0.7033$ and $|\lambda_+| = |\lambda_{\perp}| = 1$.

The fits were performed by fixing the tagging dilutions and Δt resolution function parameters to the fitted values determined in the B_{flav} Monte Carlo sample (this is necessary to decrease the CPU processing time but has no effect on the results). The angular resolution parameters were fixed to the values determined in Section 6.4.3. The CP-even parameters $\mathcal{I}m(\lambda_+)$ and $|\lambda_+|$ as well as K are allowed to float in the fit while the CP-odd parameters are fixed. The results are shown in Figure 8.10.

The fit was also performed on the entire Monte Carlo sample (~ 53000 reconstructed events), again fixing the dilutions and resolution function parameters as described above. The results obtained are: $\mathcal{I}m(\lambda_+) = -0.703 \pm 0.018$, $|\lambda_+| = 0.999 \pm 0.011$ and $K_{eff} = 0.866 \pm 0.006$, which are consistent with the generated values.

The impact of angular resolution is evident if the fit is performed on the same sample without the the angular resolution function in the signal PDF. The result obtained in this case is: $\mathcal{I}m(\lambda_+) = -0.725 \pm 0.018$, $|\lambda_+| = 0.998 \pm 0.011$ and $K_{eff} = 0.810 \pm 0.005$.

We also validate the fit using another sample of $B^0 \rightarrow D^{*+}D^{*-}$ signal Monte Carlo, where the generated amplitudes are $(A_{\parallel}, A_0, A_-) = (1, 1, 1)$. This sample
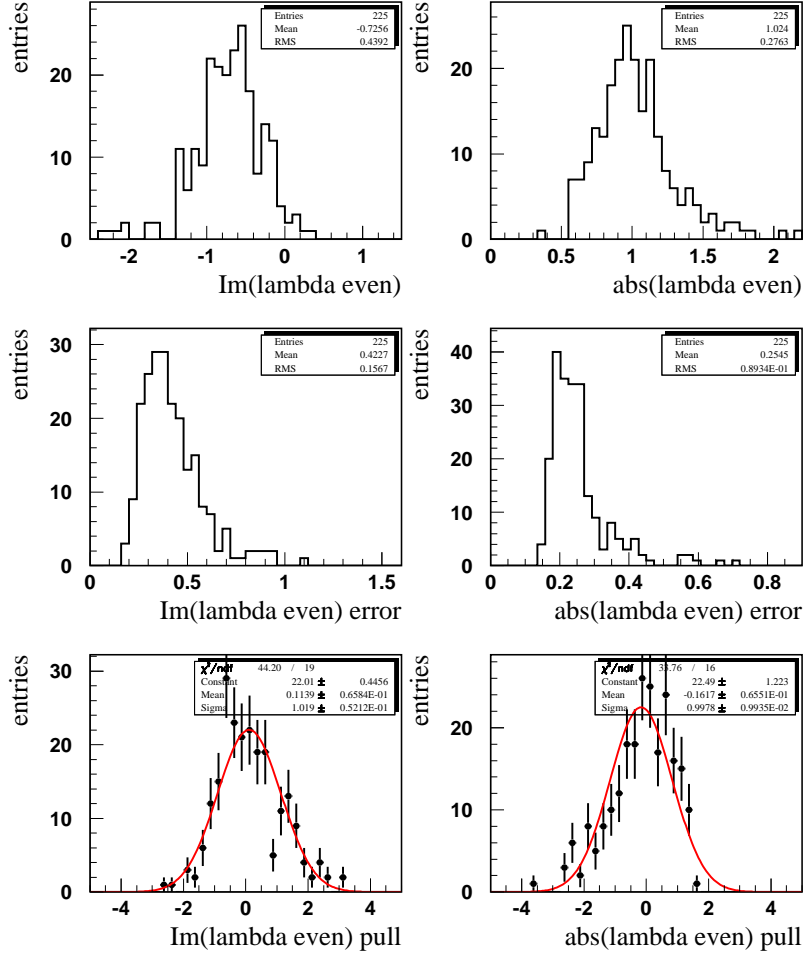
Figure 8.10: Distribution of $\mathcal{I}m(\lambda_+)$ (left) and $|\lambda_+|$ (right) in $B^0 \rightarrow D^{*+}D^{*-}$ for fully reconstructed Monte Carlo samples which are the same size as the data sample. The distribution of the errors (center) and pull distribution (bottom) are also shown.

consists of only two decay chains: $B^0 \to D^{*+}D^{*-} \to (D^0\pi^+, \overline{D}^0\pi^-) \to (K\pi, K\pi)$ and $B^0 \to D^{*+}D^{*-} \to (D^0\pi^+, D^-\pi^0) \to (K\pi, K\pi\pi) + c.c.$ The result of the fit for this sample is $\mathcal{I}m(\lambda_+) = -0.772 \pm 0.035$, $|\lambda_+| = 1.026 \pm 0.021$ and $K_{eff} = 0.276 \pm 0.011$. The K_{eff} fitted value is found to be more than 5σ away from the generated value of K. This effect is due to the acceptance as explained in Section 8.2.2 and is consistent with the effect observed in the bottom plot of Figure 8.6. A similar bias is observed in the time-integrated fit for $R_\perp$ (described in Chapter 6) when the acceptance moments are not included in the fit to the same sample.

8.3   Fit Results

Signal yields for the sample of selected $B^0 \to D^{*+}D^{*-}$ candidates were presented in Section 5.1.1 and summarized in Table 5.2. Additional vertex quality requirements are applied to select events with a well-measured interval Δt:

- the fits performed to determine the decay vertexes of the fully reconstructed and the tagging B mesons must converge;

- the measured Δt must be in the interval $[-20, 20]$ ps. This is a very loose requirement considering the B^0 lifetime of 1.542 ps; and

- the measured uncertainty $\sigma_{\Delta t}$ must be less than 2.5 ps.

These requirements reject almost all events with mis-reconstructed vertexes. The events are divided into five tagging categories: the four from the tagging algorithm (`Lepton`, `Kaon I`, `Kaon II`, and `Inclusive`) and an additional category of untagged events which are used, together with the tagged events, to constrain the value of K.

	No. of events	Lepton	Kaon I	Kaon II	Incl.	unTag
all events	460	18	56	85	112	189
signal	143	10	23	25	37	50
background	317	8	33	60	75	139

Table 8.2: Number of selected events in the $B^0 \to D^{*+}D^{*-}$ m_{ES} distribution for each tagging category. The number of signal events is determined from the area of the Gaussian and the number of background events is determined from the Argus function of the m_{ES} fit.

The number of signal and background events is evaluated from the unbinned likelihood fit to the m_{ES} distribution in two steps. First, all selected events, including untagged events, are fit to determine the mean μm_{ES} and the width σm_{ES} of the Gaussian component for signal, as well as parameter κ of the ARGUS component for background. These values ($\mu = 5.2800\,\mathrm{GeV}/c^2$, $\sigma = 0.0026\,\mathrm{GeV}/c^2$, $\kappa = -36.2$) are fixed in fits to m_{ES} distributions of each tagging category and only the normalizations of the signal and background components are allowed to float. Fixing the parameters is necessary because the amount of background and the number of events in each category are different. For example, the Lepton category has the highest purity, but also the smallest number of events which are not sufficient to determine the shape of the ARGUS function. The number of events in each category is shown in Table 8.2.

The fit to the Δt distributions of the $B^0 \to D^{*+}D^{*-}$ and B_{flav} data samples yields

$$\mathcal{I}m(\lambda_+) = 0.05 \pm 0.29(\mathrm{stat}) \tag{8.22}$$

$$|\lambda_+| = 0.75 \pm 0.19(\mathrm{stat}). \tag{8.23}$$

The value of K is determined to be 0.94 ± 0.12 and is consistent with the time-integrated measurement of $R_\perp$ given that acceptance effects are not considered. The largest correlation between $\mathcal{I}m(\lambda_+)$ and any linear combination of the other free parameters is 10%.

The Δt distributions of $B^0 \to D^{*+}D^{*-}$ events are shown in Figure 8.11. Figure 8.12 illustrates the raw asymmetry defined by

$$\mathcal{A}_{CP}(\Delta t) = \frac{N(\Delta t; B^0_{\text{tag}}) - N(\Delta t; \overline{B}^0_{\text{tag}})}{N(\Delta t; B^0_{\text{tag}}) + N(\Delta t; \overline{B}^0_{\text{tag}})} \ , \tag{8.24}$$

where $N(\Delta t; \overline{B}^0_{\text{tag}})$ and $N(\Delta t; B^0_{\text{tag}})$ are, respectively, the observed number of $\overline{B}^0$-tagged and B^0-tagged events in bins of Δt.
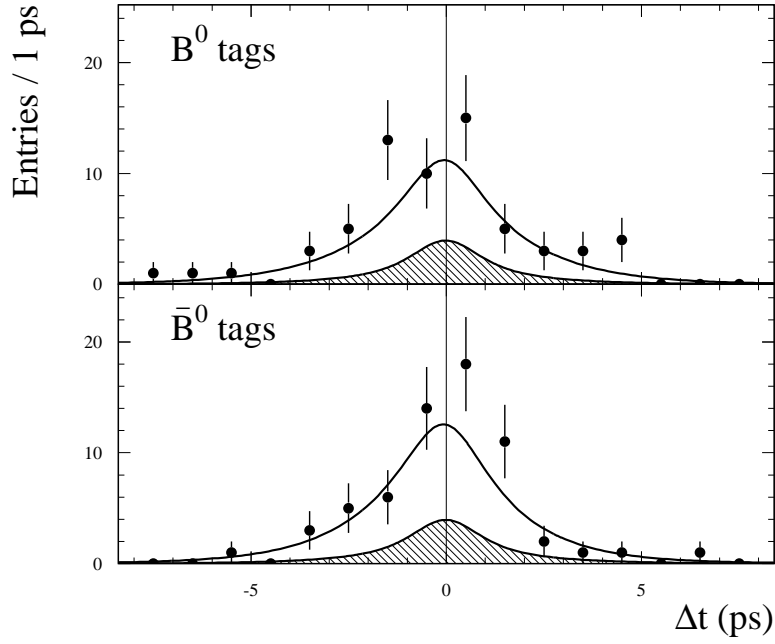
Figure 8.11: Flavor tagged $B^0 \to D^{*+}D^{*-}$ Δt distributions with the likelihood fit result superimposed. The shaded portion represents the background contribution.

207
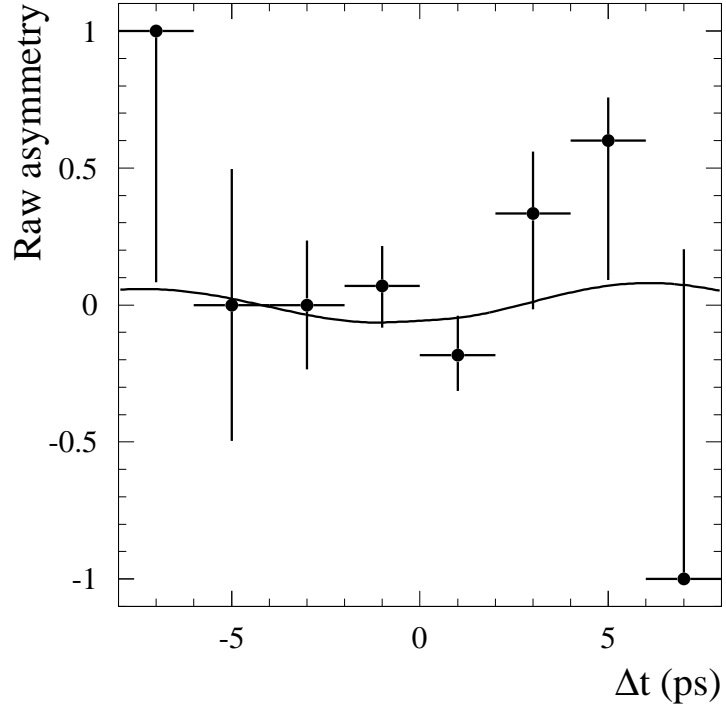
Figure 8.12: Raw $B^0 \rightarrow D^{*+}D^{*-}$ time dependent CP-even asymmetry with the likelihood fit result superimposed.

8.4 Evaluation of Systematic Uncertainties

Systematic uncertainties on $\mathcal{I}m(\lambda_+)$ and $|\lambda_+|$ arise from systematic effects in the measurement of the time-difference Δt, assumptions in the analysis technique, and the parameterization of the Δt distributions for signal and background. These uncertainties are evaluated in data, where possible, or from samples of simulated events.

8.4.1 Common Δt Resolution Function and Tagging Dilutions

The primary assumption in the Δt parameterization for $B^0 \to D^{*+} D^{*-}$ events is that the resolution function and tagging dilutions are the same in the B_{flav} and $D^{*+} D^{*-}$ samples. This assumption is tested using simulated events.

The comparison of tagging dilutions in large samples of B_{flav} and $D^{*+} D^{*-}$ Monte Carlo events was shown in Table 7.2. We assign a systematic uncertainty by observing the difference in $\mathcal{I}m(\lambda_+)$ and $|\lambda_+|$ values for the two sets of tagging dilution parameters. First, the fit is performed using the $\langle \mathcal{D} \rangle_{CP}$ and $\Delta \mathcal{D}_{CP}$ parameters measured in the $D^{*+} D^{*-}$ sample. Second, the same $D^{*+} D^{*-}$ Monte Carlo sample is fit using the $\langle \mathcal{D} \rangle_{\mathrm{flav}}$ and $\Delta \mathcal{D}_{\mathrm{flav}}$ parameters measured in the B_{flav} sample. The difference in the fits correspond to systematic uncertainties of 0.010 for $\mathcal{I}m\lambda_+$ and 0.0002 to $|\lambda_+|$.

The same technique is used to determine the systematic uncertainty for assuming a common Δt resolution function. The resolution function parameters are determined from the large simulated samples of B_{flav} and $D^{*+} D^{*-}$ events and are shown in Table 7.3. For each case, the parameters of the resolution function are fixed in the fit, and the change in $\mathcal{I}m(\lambda_+)$ and $|\lambda_+|$ is determined. The results are shown in Table 8.3. The shift in the values of $\mathcal{I}m(\lambda_+)$ (0.007) and $|\lambda_+|$ (0.007) determined using the truth information is assigned as the systematic error for the common resolution function assumption.

8.4.2 Peaking Background

In Section 5.2 the peaking background in $B^0 \to D^{*+} D^{*-}$ was estimated to be consistent with zero events in the data sample. Thus, in the nominal fit, the fraction

Res.Func. Parameter set	$\mathcal{I}m(\lambda_+)$	$|\lambda_+|$	$\delta\mathcal{I}m(\lambda_+)$	$\delta|\lambda_+|$
$D^{*+}D^{*-}$ MC Truth	-0.736 ± 0.018	1.000 ± 0.011	0	0
$D^{*+}D^{*-}$ MC fitted	-0.734 ± 0.018	0.999 ± 0.011	0.002	0.001
B_{flav} MC Truth	-0.729 ± 0.018	0.993 ± 0.011	0.007	0.007
B_{flav} MC fitted	-0.725 ± 0.018	0.998 ± 0.011	0.011	0.002

Table 8.3: Fitted values of the $\mathcal{I}m(\lambda_+)$ and $|\lambda_+|$ parameters obtained using different Δt resolution function parameters determined from the $D^{*+}D^{*-}$ and B_{flav} Monte Carlo samples (using truth information or directly fitting the reconstructed events).

for the peaking background component, f_{peak}^{CP}, was set to zero. However, the estimated upper limit of peaking events in $82\,\text{fb}^{-1}$ was 6.5 events, which corresponds to $f_{\text{peak}}^{CP} = 0.0505$. Performing the fit to the data using this value for the fraction of peaking background results in a shift in the values of $\mathcal{I}m(\lambda_+)$ (0.009) and $|\lambda_+|$ (0.0001), which are taken as the systematic uncertainties.

Also, the value of η_{peak} (the effective CP eigenvalue of peaking background events) is nominally set to zero in the fit. This parameter is varied between ± 1 for our lack of knowledge of the CP content of the possible peaking component. The full range of the fitted values of $\mathcal{I}m(\lambda_+)$ and $|\lambda_+|$ divided by $\sqrt{12}$ is taken as the systematic uncertainty for this effect. This corresponds to an uncertainty on $\mathcal{I}m(\lambda_+)$ of 0.005 and on $|\lambda_+|$ of 0.0001.

8.4.3 CP Content and Lifetime of Combinatorial Background

In the nominal fit, the value of η^{comb} was assumed to be zero, as no CP asymmetry was expected for combinatorial background events. We determine a systematic uncertainty from this assumption to be the range of the fitted values of $\mathcal{I}m(\lambda_+)$

and $|\lambda_+|$ divided by $\sqrt{12}$ as η^{comb} is varied between ± 1. This corresponds to an uncertainty on $\mathcal{I}m(\lambda_+)$ of 0.075 and on $|\lambda_+|$ of 0.005.

The Δt distribution of the background events is modeled by a prompt-lifetime component and a component with an effective lifetime $\tau_{bkg} = 1/\Gamma_2^{CP}$ that is fixed to the B^0 lifetime (1.542 ± 0.016 [50]) in the nominal fit. We observe a shift in the values of $\mathcal{I}m(\lambda_+)$ and $|\lambda_+|$ as τ_{bkg} is varied between 0.7 and 2.0 ps. The maximum values of the shifts are assigned as the systematic uncertainty on $\mathcal{I}m(\lambda_+)$ (0.018) and $|\lambda_+|$ (0.002).

8.4.4 Background Δt Resolution Function

The background resolution function in the fit was discussed in Section 8.1.3. The parameters of the background resolution function are determined from the B_{flav} sample. A systematic uncertainty is determined to estimate the differences in the background Δt resolution of the $D^{*+}D^{*-}$ and B_{flav} samples. The fit is performed by fixing all parameters except $\mathcal{I}m(\lambda_+)$ and $|\lambda_+|$ and the Δt resolution and lifetime of the background parameters and fitting to the $D^{*+}D^{*-}$ sample only. The systematic uncertainties are determined to be 0.015 for $\mathcal{I}m(\lambda_+)$ and 0.001 for $|\lambda_+|$.

8.4.5 Fixed B^0 lifetime and oscillation frequency Δm_d

In the nominal fit, the B^0 lifetime and the oscillation frequency Δm_d are fixed to the world average values 1.542 ± 0.016 ps and 0.489 ± 0.008 ps^{-1} [50], respectively. The dependency of $\mathcal{I}m(\lambda_+)$ and $|\lambda_+|$ on τ_{B^0} and Δm_d is determined from fits where the fixed values of these parameters are varied $\pm 1\sigma$ (*i.e.* ± 0.016 ps and ± 0.008 MeV, respectively). The variation of the B^0 lifetime results in a change in $\mathcal{I}m\lambda_+$ of 0.001 and a negligible change of $|\lambda_+|$. The variation of the mixing frequency results in a change in $\mathcal{I}m\lambda_+$ of 0.0025 and in $|\lambda_+|$ of 0.0026.

$\mathcal{I}m(\lambda_+)$	$\|\lambda_+\|$	$\mathcal{I}m(\lambda_\perp)$	$\|\lambda_\perp\|$
$-0.7097 +/- 0.3037$	$1.3789 +/- 0.1969$	$1.0(fixed)$	$1.0(fixed)$
$-0.7699 +/- 0.2918$	$1.3749 +/- 0.1915$	$-1.0(fixed)$	$1.0(fixed)$
$-0.7709 +/- 0.2949$	$1.3672 +/- 0.1932$	$-0.75(fixed)$	$0.7(fixed)$
$-0.7623 +/- 0.2910$	$1.3817 +/- 0.1922$	$-0.75(fixed)$	$1.3(fixed)$

Table 8.4: Fitted values of $\mathcal{I}m(\lambda_+)$ and $\|\lambda_+\|$ obtained by varying the values of the fixed parameters $\mathcal{I}m(\lambda_\perp)$ and $\|\lambda_\perp\|$. (The central values of $\mathcal{I}m(\lambda_+)$ and $\|\lambda_+\|$ are displayed with an arbitrary signed offset. Only the difference in the fitted values is used to determine the systematic uncertainty.)

8.4.6 Fixed values of $\mathcal{I}m(\lambda_\perp)$ and $\|\lambda_\perp\|$

In the nominal fit we fix the value of the CP-odd parameters to the values determined from the Standard Model assuming that penguin contributions are negligible: $\mathcal{I}m(\lambda_\perp) = -0.75$ and $\|\lambda_\perp\| = 1.0$. To determine the systematic uncertainty for this assumption, the fit is repeated by fixing $\mathcal{I}m(\lambda_\perp)$ parameters to ± 1 and $\|\lambda_\perp\|$ to 0.7 and 1.3. Table 8.4 shows the results. We therefore assign a systematic uncertainty of 0.056 to $\mathcal{I}m(\lambda_+)$ and 0.008 to $\|\lambda_+\|$.

8.4.7 θ_{tr} Angular Resolution

The effects of angular resolution were studied in Section 8.2.3. Section 7.4.1 described how the effect was accounted for in the signal pdf and fit method. In Section 8.2.4 the angular resolution function parameterization was changed as an estimate of the validity of the parameterization used in the nominal fit. A systematic uncertainty equal to the difference in the fitted values of $\mathcal{I}m(\lambda_+)$ and $\|\lambda_+\|$ for the two parameterizations is assigned (0.005 and 0.000, respectively).

8.4.8   Detector Acceptance

The effect of detector acceptance was studied in Section 8.2.2. The soft pion effi-
ciency distribution in Figure 8.5 was used to generate toy Monte Carlo and estimate
the change in the fitted values of $\mathcal{I}m(\lambda_+)$ and $|\lambda_+|$. The difference between the
average fitted value and the input value is used as an estimation of the systematic
uncertainty from acceptance effects. This corresponds to 0.0028 for $\mathcal{I}m(\lambda_+)$ and
0.0009 for $|\lambda_+|$.

The value of K extracted in the fit was quoted as an "effective" value because
acceptance effects were not included in the fit. The effect was demonstrated with
simulated events in Figure 8.6. As a cross-check, however, we would like to compare
the fitted value of K and the acceptance-corrected value of $R_\perp$ obtained in the
time-integrated fit of Chapter 6.

After obtaining a common set of events (recall the additional selection criteria
of Section 8.3) and configuring the fit without acceptance moments, the time-
integrated fit for $R_\perp$ is repeated. The value obtained from the modified $R_\perp$ fit is
$R_\perp^{eff} = 0.042$, while the corresponding value from the CP fit is $R_\perp^{eff} = (1-K)/2 =$
0.032. (The statistical uncertainties are not quoted since the two results originate
from the same data sample and are highly correlated.) Given that the CP fit
differs in how the signal events are weighted (for each tagging category) and that
it contains correlations between the values of K and $\mathcal{I}m(\lambda_+), |\lambda_+|$, the results are
considered to be consistent. The value of K, however, will not considered to be an
accurate measure of the CP-odd fraction, $R_\perp$, in $B^0 \rightarrow D^{*+}D^{*-}$.

8.4.9 Uncertainty on the Boost

To first approximation, the value of Δt and the measured Δz are related by $\Delta z = \beta\gamma\Delta t$. Hence, variations in $\beta\gamma$ directly impact the value of Δt, which can result in variations in the fitted values of $\mathcal{I}m(\lambda_+)$ and $|\lambda_+|$. The boost factor $\beta\gamma$ is measured from the knowledge of the PEP-II beam energies with a relative uncertainty of 0.6% [45]. The measured value of Δt and its error is changed by this amount and the fit is repeated. The systematic error from the boost uncertainty is determined to be 0.011 for $\mathcal{I}m(\lambda_+)$ and 0.019 for $|\lambda_+|$.

8.4.10 SVT alignment

Reconstruction of the decay vertexes of B mesons relies on the high precision reconstruction of charged-particle trajectories. For particles originating from the interaction point, the measurement of the trajectory parameters is dominated by the silicon vertex tracker. The measured parameters are very sensitive to the relative positions of the silicon wafers and strips that are used to detect the interactions of the charged particles. The knowledge of actual positions of the wafers and strips with respect to their nominal positions is referred to as the SVT *local alignment*, which was described in Section 3.3.

The reconstruction of simulated $D^{*+}D^{*-}$ and B_flav samples uses the correct (true) positions of the SVT wafers, and corresponds to a perfect alignment scenario. Hence, fits to simulated $D^{*+}D^{*-}$ and B_flav samples provide a nominal value for $\mathcal{I}m(\lambda_+)$ and $|\lambda_+|$ with perfect alignment. Possible misalignment scenarios in the data can be expressed in terms of rotations and translations of the wafers, and are used to introduce a *known* misalignment in the reconstruction of simulated events.

A sample of $B^0 \to D^{*+}D^{*-}$ events with amplitudes $(A_0, A_\parallel, A_\perp) = (1, 1, 1)$ was

Monte Carlo sample	$\mathcal{I}m(\lambda_+)$	$\lvert\lambda_+\rvert$	$\delta\mathcal{I}m(\lambda_+)$	$\delta\lvert\lambda_+\rvert$
$D^{*+}D^{*-}$ "aligned SVT"	-0.689 ± 0.084	1.035 ± 0.051	0	0
$D^{*+}D^{*-}$ "misaligned SVT"	-0.718 ± 0.087	1.037 ± 0.052	0.029	0.003

Table 8.5: Fitted values of $\mathcal{I}m(\lambda_+)$ and $\lvert\lambda_+\rvert$, obtained from $D^{*+}D^{*-}$ Monte Carlo samples processed with different SVT alignment sets.

generated using a misalignment scenario similar to that observed in the data. A common set of events is used to compare the fit results for the "perfect alignment" and "misaligned SVT" samples (see Table 8.5). The systematic uncertainties due to SVT misalignment assigned to $\mathcal{I}m(\lambda_+)$ and $\lvert\lambda_+\rvert$ are 0.029 and 0.003 respectively.

8.4.11 Summary of Systematic Uncertainties

The dominant contributions to the systematic uncertainty on $\mathcal{I}m(\lambda_+)$ originate from the unknown CP content of background events (0.075) and the variation of the values of $\mathcal{I}m(\lambda_\perp)$ and $\lvert\lambda_\perp\rvert$ (0.056). The latter will not change very much with additional data as these parameters are poorly determined given the measured value of $R_\perp$.

Table 8.6 summarizes the systematic uncertainties evaluated for $\mathcal{I}m(\lambda_+)$ and $\lvert\lambda_+\rvert$. The total uncertainty is computed by adding in quadrature all the contributions and is found to be 0.10 for $\mathcal{I}m(\lambda_+)$ and 0.02 for $\lvert\lambda_+\rvert$. These uncertainties are about a factor of three smaller than the statistical errors.

215

Systematics source	$\delta\mathcal{I}m(\lambda_+)$	$\delta|\lambda_+|$
Signal parameters		
Signal Tagging Dilutions	0.010	0.0002
Common Δt resolution function	0.007	0.007
Background parameters		
Fraction of peaking background	0.009	0.0001
Background CP content (peaking)	0.005	0.0001
$D^{*+}\ D^{*-}$ background CP content	0.075	0.005
Lifetime of background	0.018	0.002
Background Δt resolution function	0.015	0.001
External parameters		
B^0 lifetime variation	0.001	–
Δm_d variation	0.0025	0.0026
Variation of $\mathcal{I}m(\lambda_\perp)$ and $|\lambda_\perp|$	0.056	0.008
Detector Effects		
Angular resolution parameterization	0.005	–
Acceptance	0.0028	0.0009
Boost uncertainty	0.011	0.019
SVT misalignment	0.029	0.003
TOTAL	0.10	0.02

Table 8.6: Summary of the systematics errors estimated for the $\mathcal{I}m(\lambda_+)$ and $|\lambda_+|$ parameters.

8.5   Summary

Time-dependent CP asymmetries were measured in fully-reconstructed $B^0 \rightarrow D^{*+}D^{*-}$ events using the data collected by the *BABAR* detector from December 1999 - July 2002. A simultaneous fit to events reconstructed in B-flavor eigenstates and as $B^0 \rightarrow D^{*+}D^{*-}$, utilizing tagging information from the other B in the event results in a measurement of the CP-odd parameters:

$$\mathcal{I}m(\lambda_+) = 0.05 \pm 0.29(\text{stat}) \pm 0.10(\text{syst}) \tag{8.25}$$

$$|\lambda_+| = 0.75 \pm 0.19(\text{stat}) \pm 0.02(\text{syst}). \tag{8.26}$$

If the $B^0 \rightarrow D^{*+}D^{*-}$ decay proceeds only through the $b \rightarrow c\bar{c}d$ tree amplitude, then one expects that $\mathcal{I}m(\lambda_+) = -\sin 2\beta$ and $|\lambda_+| = 1.$. To test this hypothesis, we fix $\mathcal{I}m(\lambda_+) = -0.741$ [8] and $|\lambda_+| = 1.$ and repeat the fit. The observed change in the likelihood corresponds to 2.5 standard deviations (statistical uncertainty only). Thus, the time-dependent asymmetry is found to differ slightly from the Standard Model prediction with penguin amplitudes ignored.

Chapter 9

Conclusions and Outlook

9.1 Summary

About 88 million $\Upsilon(4S) \to B\bar{B}$ decays were collected between December 1999 and June 2002 with the *BABAR* detector at the PEP-II asymmetric-energy e^+e^- collider. From this data sample about 150 events were reconstructed as $B^0 \to D^{*+}D^{*-}$. A time-integrated angular analysis was performed to determine the *CP*-odd parity fraction of the final state, called $R_\perp$, and was found to be

$$R_\perp = 0.063 \pm 0.055(\text{stat}) \pm 0.009(\text{syst}). \tag{9.1}$$

This result implies that the $D^{*+}D^{*-}$ final state is mostly *CP*-even. The *CP*-odd fraction is in agreement with a number of theoretical predictions which rely on factorization approximations [30].

Because the *CP*-odd fraction is found to be small, the related dilution in the time-dependent *CP* asymmetry measurement is also expected to be small. A measurement of the *CP*-even asymmetry parameters $\mathcal{I}m(\lambda_+)$ and $|\lambda_+|$ was performed on the $B^0 \to D^{*+}D^{*-}$ event sample. We determine

$$\mathcal{I}m(\lambda_+) \;=\; 0.05 \pm 0.29(\text{stat}) \pm 0.10(\text{syst}) \tag{9.2}$$

$$|\lambda_+| \;=\; 0.75 \pm 0.19(\text{stat}) \pm 0.02(\text{syst}). \tag{9.3}$$

If one assumes that the penguin contributions to the final state are negligible (some models predict the contamination to be at the 2% level [29]), then $\mathcal{I}m(\lambda_+) = -\sin 2\beta$ according to the Standard Model. Our result is equivalent to a 2.5 sigma deviation from the Standard Model prediction which states the result should be the same as in $B^0 \to J/\psi K_S^0$.

Both of the above results appeared in Physical Review Letters [40]. The $R_\perp$ measurement presented here supersedes the previous *BABAR* measurement [39], with a factor of three reduction in the statistical uncertainty. The time-dependent measurement of CP asymmetries in $B^0 \to D^{*+}D^{*-}$ is the first measurement of its kind in that mode. While the statistical errors are currently large, the measurement represents an important test of the Standard Model picture of CP violation.

9.2   Other Measurements of $\sin 2\beta$

The $B^0 \to D^{*+}D^{*-}$ decay represents one of the many $b \to c\bar{c}d$ transitions which are dominated by a tree diagram and have potential additional contributions from penguin diagrams. These types of decays represent one of two additional classes of decays besides the "golden modes" ($J/\psi\ K_S^0$, $J/\psi\ K_L^0$, $\psi(2S)\ K_S^0$, $\chi_{c1}\ K_S^0$, and $\eta_c\ K_S^0$) which are sensitive to the Unitarity Triangle angle β. CP violation asymmetries have also been measured by the *BABAR* collaboration in $B^0 \to D^{*\pm}D^{\mp}$ [41] and $B^0 \to J/\psi\pi^0$ [62] decays.

The other class of decays which are sensitive to β correspond to $b \to s\bar{s}s$ or $b \to d\bar{d}s$ transitions and are dominated by penguin diagrams. The leading contribution to the amplitudes for $B \to \phi^0 K_S^0$ and $\eta' K_S^0$ decays is the penguin diagram illustrated in Figure 9.1. The absence of a second contribution at the order $\mathcal{O}(10^{-2})$ allows another theoretically clean measurement of $\sin 2\beta$.
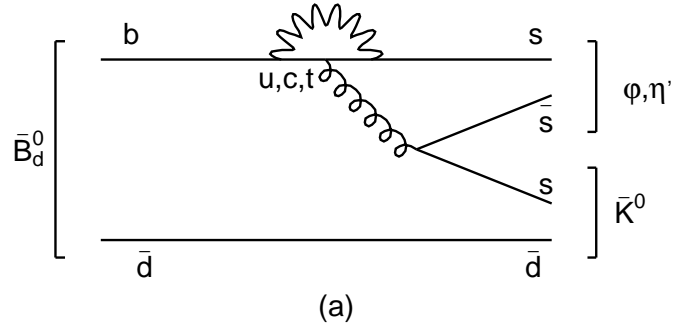
Figure 9.1: Penguin diagram for the $B \to \phi^0 K_S^0, \eta' K_S^0$ decays.

The leading penguin diagram is suppressed in the Standard Model for these decays. Potential contributions from physics beyond the Standard Model could be comparable to the the penguin contribution, and result in discrepancies between the value of $\sin 2\beta$ in these modes and the measurement in the golden modes. In addition to the measurements mentioned above, the *BABAR* collaboration has performed *CP*-violating asymmetry measurements for the modes $B \to \phi^0 K_S^0$ [63], $B \to \eta' K_S^0$ [64], and $B \to \pi^0 K_S^0$ [65].

Some of the above mentioned measurements show a hint of slight discrepancies from Standard Model predictions. They are all limited by the current statistical size of the data sample; however, should these discrepancies become more pronounced as more data is accumulated, these measurements could indicate the effects of New Physics.

A summary of results of $\sin 2\beta$ measurements by the *BABAR* collaboration are shown in Figure 9.2. [1]

---

[1]Note that the S coefficient of the time-dependent asymmetry is shown in the figure. For $B^0 \to D^{*+}D^{*-}$ the relationship between $\mathcal{I}m(\lambda_+)$ and $|\lambda_+|$ and the S and C coefficients is straightforward and is calculated including correlations.

9.3 Future Prospects

The results presented in this dissertation show that the time-dependent CP asymmetry measurement in $B^0 \to D^{*+}D^{*-}$ differs from the Standard Model prediction (with penguin contributions ignored) by 2.5 standard deviations. This discrepancy is statistically inconclusive with the current size of the data sample. The current luminosity model for the PEP-II collider predicts an accumulated data sample of about $500 \, \text{fb}^{-1}$ by the end of 2005, which is six times larger than the sample used for these measurements. If one assumes that the reconstruction efficiency and tagging power of the detector and the analysis remains unchanged, then one should expect the statistical uncertainty on $\mathcal{I}m(\lambda_+)$ to be < 0.12. This estimate approaches the current level of systematic uncertainty of the analysis, and implies that the measurement will be limited by statistics for the next few years. As data is accumulated, the $B^0 \to D^{*+}D^{*-}$ decay will, therefore, provide a statistically interesting measurement of the Unitarity Triangle angle β (provided a measure of the penguin diagrams contribution is accomplished) and begin to test the Standard Model picture of CP violation.
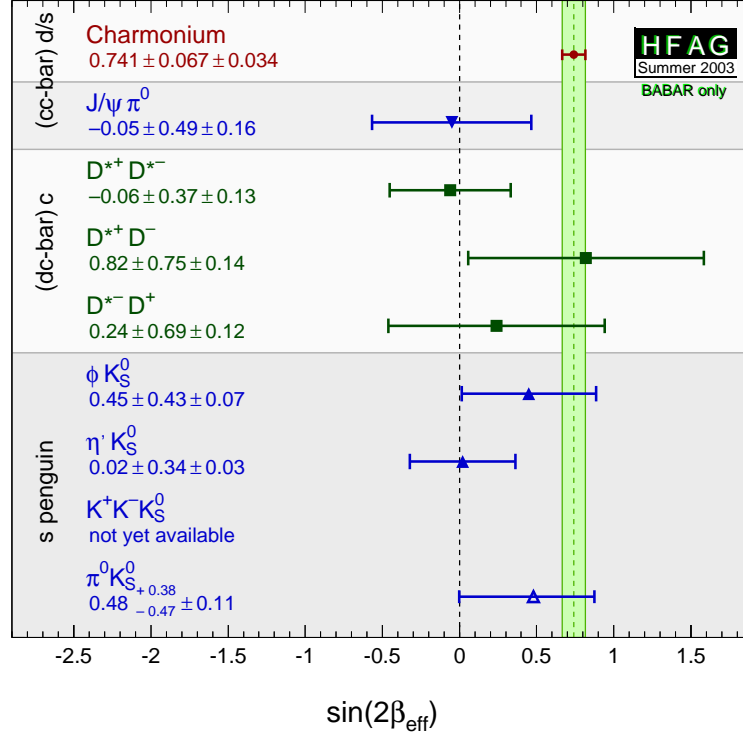
Figure 9.2: A summary of the measurements of CP-violating time-dependent asymmetries by the *BABAR* collaboration as of the summer of 2003. The numbers shown represent the value of the S coefficient in the asymmetry definition (see Equation 2.22) which is interpreted as $\sin 2\beta$ in the Standard Model, provided that only one diagram contributes to the decay amplitude.

Appendix A

Derivation of $B^0 \rightarrow D^{*+}D^{*-}$ Decay Rate

For completeness and pedagogical reasons we derive the differential decay rate for a B decaying to two D^*s. For generality, we consider the D^* to decay to a D meson and a pion (no charge specified). We therefore label the daughters of one of the D^*s D_1 and π_1, and the daughters of the other D_2 and π_2.

We begin by writing the general matrix element for $B^0 \rightarrow D^{*+}D^{*-}$:

$$\mathcal{M} = \langle D^{*+}D^{*-}|\mathcal{H}_{EW}|B^0\rangle \langle j_1 j_2; m_1 m_2|\mathcal{R}|j\ m_B\rangle \tag{A.1}$$

where the first bra-ket denotes the actual physics of a B decaying to two D^*s through the Electro-Weak Hamiltonian and the second bra-ket denotes the total angular momentum ($J = L + S$) information of the decay ($\mathcal{R}$ is the rotation matrix). The j represents the spin of B and m_B is its helicity (or more precisely, the eigenvalue of the J_z operator). j_1 and j_2 correspond to the spins of the two D^*s, and m_1 and m_2 are the eigenvalues of the projected spin states. Since the B is a pseudoscalar it has spin and helicity of zero; a D^* is a vector particle and has spin one. Therefore the two D^*s must be in one of three possible coherent states based on the conservation of angular momentum: $|m_1, m_2\rangle = |1, -1\rangle$ or $|0, 0\rangle$ or $|-1, 1\rangle$. That is,

$$\langle j_1 j_2; m_1 m_2|\mathcal{R}|j\ m_B\rangle = \delta_{m_1, -m_2} \tag{A.2}$$

223

Since there are three possible decays of $B^0 \to D^{*+}D^{*-}$ the $\mathcal{M}$ matrix element must be a sum over the three decay paths (helicities). For the purpose of notation we will use the helicity formalism: here we denote the projection of the spin along the direction of momentum of each D^*. Therefore the above relation requires that λ, the helicity eigenvalue of both D^*s be equal ($\lambda = m_1 = -m_2$).

In order to derive the angular distribution we must also consider the decay of the D^*. We therefore expand $\mathcal{M}$ to include the amplitudes of each $D^* \to D\pi$ and their angular momentum contributions.

$$\mathcal{M} = \sum_\lambda A_{j_1} A_{j_2} \, B_\lambda \, \langle j_1 \; \lambda'_1 | \mathcal{R} | j_{D^*_1} \; \lambda \rangle_\lambda \langle j_2 \; \lambda'_2 | \mathcal{R} | j_{D^*_2} \; \lambda \rangle_\lambda \qquad \text{(A.3)}$$

Here A_{j_1} and A_{j_2} are the D^* decay amplitudes (independent of angular information) and $B_\lambda = \langle D^{*+}D^{*-} | \mathcal{H}_{EW} | B^0 \rangle_\lambda$, the $B^0 \to D^{*+}D^{*-}$ amplitudes for each helicity state. The final two bra-kets denote the angular momentum matrix elements for $D^* \to D\pi$. The sum over λ is the sum over the three possible helicity states of the two D^*s. The D^*s have spin one, so $j_1 = j_2 = 1$. The D and π mesons each have zero spin (pseudoscalars) and therefore must have a total helicity of $\lambda_D - \lambda_\pi = m'_1 + m'_2 = 0$. Therefore we need only determine the Wigner functions

$$\mathcal{D}^j_{m',m} = \langle j \; m' | \mathcal{R} | j \; m \rangle$$

Since the helicity of the $D\pi$ final state is zero, we are interested in

$$\mathcal{D}^1_{\lambda,0}(\theta_1, \phi_1, 0) \;\; \text{and} \;\; \mathcal{D}^1_{\lambda,0}(\theta_2, \phi_2, 0)$$

where $\lambda = m_1 = -m_2$ (the helicity state of the D^*s) and (θ_1, ϕ_1) are the decay angles of the D^*_1 daughters and (θ_2, ϕ_2) are the decay angles of the D^*_2 daughters. (See Figure A.1.)

Note that:

$$\mathcal{D}^1_{\lambda,0}(\theta, \phi, 0) = e^{i\lambda\phi} \; \mathrm{d}^1_{\lambda,0}(\theta)$$
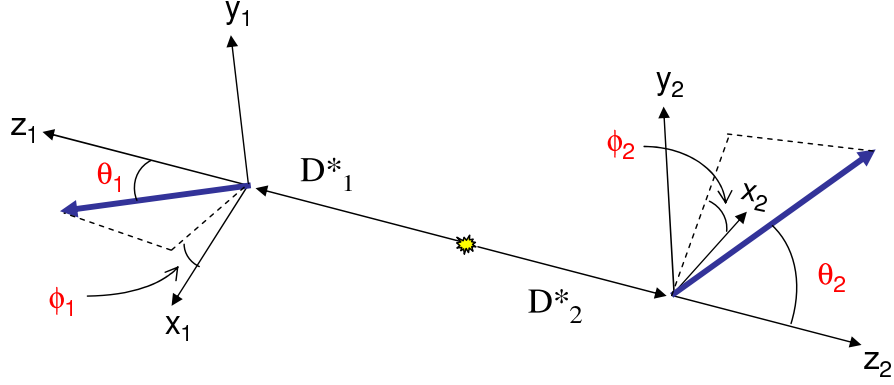
Figure A.1: The decay diagram of $B^0 \rightarrow D^{*+}D^{*-}$ in the "helicity frame". The direction of the z_1 (z_2) axis is defined with respect to the direction of flight of the D_1^* (D_2^*) in the B rest frame. The "helicity" angles are defined in the rest frame of the corresponding D^*.

and

$$d_{1,0}^1 = -\frac{\sin\theta}{\sqrt{2}} \qquad d_{0,0}^1 = \cos\theta \qquad d_{-1,0}^1 = \frac{\sin\theta}{\sqrt{2}} \qquad (A.4)$$

Note also in Eqn. A.3 that the D^* amplitudes are independent of λ and will cancel out when the fractional decay rate $\frac{1}{\Gamma}\frac{d^3\Gamma}{d\Omega}$ is determined. Therefore we have

$$|\mathcal{M}|^2 = \left| B_0 \cos\theta_1 \cos\theta_2 + \frac{1}{2}B_{+1}e^{i(\phi_1+\phi_2)}\sin\theta_1\sin\theta_2 + \frac{1}{2}B_{-1}e^{-i(\phi_1+\phi_2)}\sin\theta_1\sin\theta_2 \right|^2$$

$$(A.5)$$

Simplifying we get

$$
\begin{aligned}
|\mathcal{M}|^2 =\ & |B_0|^2 \cos^2\theta_1 \cos^2\theta_2 \\
& + \frac{1}{4}(|B_1|^2 + |B_{-1}|^2)\sin^2\theta_1\sin^2\theta_2 \\
& + \frac{1}{4}\left[\mathcal{R}e(B_0^*[B_1 + B_{-1}])\cos\chi - \mathcal{I}m(B_0^*[B_1 - B_{-1}])\sin\chi\right]\sin 2\theta_1 \sin 2\theta_2 \\
& + \frac{1}{2}\left[\mathcal{R}e(B_1 B_{-1}^*)\cos 2\chi - \mathcal{I}m(B_1 B_{-1}^*)\sin 2\chi\right]\sin^2\theta_1\sin^2\theta_2 \qquad (A.6)
\end{aligned}
$$

where $\chi = \phi_1 + \phi_2$. While it can be shown how the helicity amplitudes are related to the S-, P-, and D-wave eigenstates, we are more interested in converting the

225

above expression to the transversity basis. To do so, let us first reconsider the coordinate systems shown in Figure A.1.

We choose a new coordinate system (primed) where the angle ϕ_1 is absorbed into the coordinate system of the D_2^*. This clearly defines a "decay plane" where the normal is seen to correspond to y_2' (the new y-axis in the D_2^* coordinate system). Note that θ_1 and the z_1-axis are unchanged; we only require that the D_1^* decay lies completely in the $z_1 - x_1'$ plane. This change make the azimuthal angle in the D_2^* coordinate system equal to $\phi_1 + \phi_2$, which is χ in Eqn. A.6. In the new primed coordinate system, we see that

$$y_2' = z_{\text{tr}}$$
$$x_2' = y_{\text{tr}}$$
$$z_2^{(\prime)} = x_{\text{tr}} \tag{A.7}$$

We now define θ_{tr} as the polar angle in the D_2^* primed frame (or "tr" frame) with respect to the new z_{tr}-axis. The angle ϕ_{tr} is the corresponding azimuthal angle opening from the x_{tr}-axis (originally the z_2-axis). See Figure A.2. It follows that:

$$\cos\theta_2 = \sin\theta_{\text{tr}}\cos\phi_{\text{tr}}$$
$$\sin\theta_2\cos\chi = \sin\theta_{\text{tr}}\sin\phi_{\text{tr}}$$
$$\sin\theta_2\sin\chi = \cos\theta_{\text{tr}} \tag{A.8}$$

Finally, we define the "transversity" amplitudes in terms of the helicity amplitudes:

$$B_0 = T_0$$
$$B_1 = \frac{1}{\sqrt{2}}(T_\| + T_\perp)$$
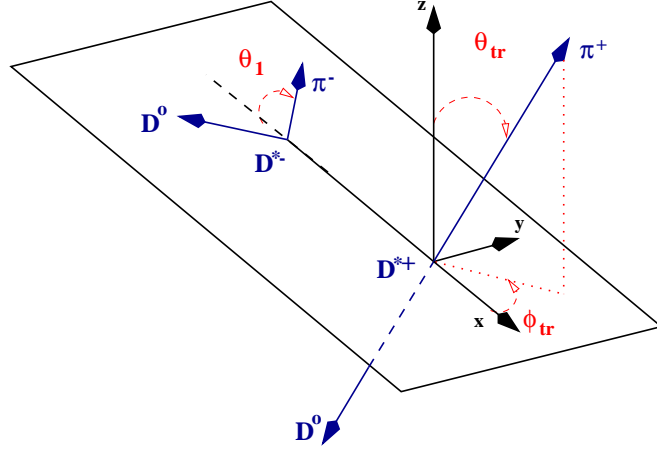$$B_{-1} = \frac{1}{\sqrt{2}}(T_\| - T_\perp) \tag{A.9}$$

Figure A.2: The decay diagram of $B^0 \rightarrow D^{*+}D^{*-}$ in the "transversity frame". Here the decay of the D^{*-} is represented in the B^0 rest frame, while the decay products of the D^{*+} are shown in the D^{*+} rest frame. The x direction is defined by the direction of flight of the D^{*+} in the B^0 rest frame.

and we also note that

$$
\begin{aligned}
|B_1|^2 + |B_{-1}|^2 &= |T_{\|}|^2 + |T_{\perp}|^2 \\
B_1 B_{-1}^* &= \frac{1}{2}(|T_{\|}|^2 - |T_{\perp}|^2) + i\,\mathcal{I}m(T_{\|}^* T_{\perp})
\end{aligned}
\tag{A.10}
$$

Using A.8, A.9, and A.10 we can rewrite Eqn. A.6 as

$$
\begin{aligned}
|\mathcal{M}|^2 =\ & |T_0|^2 \cos^2\theta_1 \sin^2\theta_{\mathrm{tr}} \cos^2\phi_{\mathrm{tr}} \\
& + \frac{1}{2}|T_{\|}|^2 \sin^2\theta_1 \sin^2\theta_{\mathrm{tr}} \sin^2\phi_{\mathrm{tr}} \\
& + \frac{1}{2}|T_{\perp}|^2 \sin^2\theta_1 \cos^2\theta_{\mathrm{tr}} \\
& + \frac{1}{2\sqrt{2}}\mathcal{R}e(T_0^* T_{\|}) \sin 2\theta_1 \sin^2\theta_{\mathrm{tr}} \sin 2\phi_{\mathrm{tr}} \\
& - \frac{1}{2\sqrt{2}}\mathcal{I}m(T_0^* T_{\perp}) \sin 2\theta_1 \sin 2\theta_{\mathrm{tr}} \cos\phi_{\mathrm{tr}} \\
& - \frac{1}{2}\mathcal{I}m(T_{\|}^* T_{\perp}) \sin^2\theta_1 \sin 2\theta_{\mathrm{tr}} \sin\phi_{\mathrm{tr}}.
\end{aligned}
\tag{A.11}
$$

Appendix B

Acceptance Moments' Calculation

The calculation of the acceptance moments required for the $R_\perp$ analysis was outlined in Section 6.3.1. Here we present a more rigorous derivation of the formulas used (Equations 6.18,6.19) to estimate the acceptance moments in a binned representation.

B.1   Monte Carlo Estimation of Integrals

The basic technique of Monte Carlo integration is as follows. Given a sampling probability distribution function (pdf), $f(x)$, the expectation value of any function, $h(x)$, is defined as:

$$\langle h \rangle = \int_{\mathcal{V}} h(x) f(x) \; dx \tag{B.1}$$

The expectation value if h can be estimated by the discrete sum over a given event sample $\{x_i\}$ from

$$\langle h \rangle_N \; \approx \; \frac{C}{N} \sum_{i=1}^{N} h(x_i), \tag{B.2}$$

where the N is the total number of events and $C = \int_{\mathcal{V}} f(x)\mathrm{d}x$. The estimated variance in this estimate is given by

$$V(\langle h \rangle_N) = \frac{C^2}{N-1} \left[\frac{1}{N}\sum_{i=1}^{N} h(x_i)^2 - \left(\frac{1}{N}\sum_{i=1}^{N} h(x_i) \right)^2 \right]. \qquad (B.3)$$

From now on we shall assume that $f(x)$ is normalized, such that $C = 1$. Furthermore, we implicitly assume that all integrals are definite integrals over the full space $\mathcal{V}$.

B.2 Acceptance Functions

The efficiency function $\epsilon(x, y, z)$ represents the probability that an event with configuration (x, y, z) is accepted by the event selection and reconstruction. In general, it is a continuous function taking values between 0.0 and 1.0 as a function of the position in space. However, a more general formulation defines an acceptance function $\epsilon(\mathbf{w})$ to be equal to one of two values: 1 for events which pass all selection criteria and 0 for events which do not. The acceptance of a particular event is in fact dependent on many random variables in the Monte Carlo generation, and is theoretically known. For example, the acceptance of an event can depend on the random number that specifies the energy-loss of a particular track in a particular DCH drift cell. These types of random variables are represented by the set $\mathbf{w}$. This more fundamental acceptance distribution is then denoted by $\epsilon(x, y, z, \mathbf{w})$, where $\mathbf{w}$ represents all degrees of freedom other than (x, y, z). The value of $\epsilon(x, y, z, \mathbf{w})$ is, therefore, defined to be 1. for accepted events and 0. for rejected events.

The efficiency distribution $\epsilon(x, y, z)$ is then formally written as the integral:

$$\epsilon(x, y, z) = \frac{\int \mathrm{d}\mathbf{w}\ \epsilon(x, y, z, \mathbf{w})\ f(x, y, z, \mathbf{w})}{\int \mathrm{d}\mathbf{w}\ f(x, y, z, \mathbf{w})} \qquad (B.4)$$

The Monte Carlo generation pdf $f(x, y, z, \mathbf{w})$ is an *a priori* known, but extremely complicated function. However, its integral over all redundant degrees of freedom

$$f(x, y, z) = \int d\mathbf{w} f(x, y, z, \mathbf{w}) \tag{B.5}$$

is well known and is relatively simple (Equation 6.4). This makes it possible to calculate the acceptance moment integrals using the Monte Carlo events generated with the common *BABAR* Monte Carlo simulation.

B.3 Calculating the Acceptance Moments

The acceptance moments defined by integrals of the form:

$$I_\alpha(z) = \int g_\alpha(x, y) \epsilon(x, y, z) dx\, dy. \tag{B.6}$$

In this definition $g_\alpha(x, y)$ are the moment distributions, given by

$$
\begin{aligned}
g_0(x, y) &= \frac{3}{2\pi} y^2 \cos^2 x \\
g_{||}(x, y) &= \frac{3}{4\pi} \sqrt{1 - y^2} \sin^2 x \\
g_\perp(x, y) &= \frac{3}{8\pi} \sqrt{1 - y^2}
\end{aligned}
$$

and $x = \phi_{\text{tr}}$, $y = \cos(\theta_1)$, $z = \cos(\theta_{\text{tr}})$, as seen from Eqn. 6.9. Since we will bin the Monte Carlo distribution in $\cos \theta_{\text{tr}}$, the average value of the acceptance moment in bin k is defined as:

$$I_\alpha^k = \frac{1}{\Delta z} \int_{z_k}^{z_{k+1}} dz\, I_\alpha(z) \tag{B.7}$$

After inserting the formal definition of the acceptance, this takes the explicit form

$$I^k = \frac{1}{\Delta z} \int_{z_k}^{z_{k+1}} dz \int dx dy d\mathbf{w}\, f(x, y, z, \mathbf{w}) \frac{g(x, y)\epsilon(x, y, z, \mathbf{w})}{f(x, y, z)} \tag{B.8}$$

Using the formalism of Monte Carlo estimation of integrals (Equations B.1,B.2), we can now estimate the value of I_α^k from a finite Monte Carlo event sample:

$$
\begin{aligned}
I_{N_k}^k &= \frac{c_k}{N_k \Delta z} \sum_{\text{all events in bin } k} \frac{g(x_i, y_i)\epsilon(x_i, y_i, z_i, w_i)}{f(x_i, y_i, z_i)} \\
&= \frac{c_k}{N_k \Delta z} \sum_{\text{selected events in bin } k} \frac{g(x_i, y_i)}{f(x_i, y_i, z_i)}
\end{aligned}
\tag{B.9}
$$

where c_k is the integral of $f(x, y, z)$ over one bin and N_k is the number of generated events in bin k. Using Equation B.3, the uncertainty in this estimate is given by

$$
\sigma(I_{N_k}^k) = \frac{c_k}{N_k \Delta z} \sqrt{\sum \left(\frac{g(x_i, y_i)}{f(x_i, y_i, z_i)} \right)^2 - \frac{1}{N_k} \left(\sum \frac{g(x_i, y_i)}{f(x_i, y_i, z_i)} \right)^2}.
\tag{B.10}
$$

Unfortunately, due to the way we have generated our Monte Carlo samples, the exact number of events N_k generated in each bin k is lost and only the total number of generated events N is known. This translates to an additional uncertainty in the estimate of the acceptance moment. An elegant way of calculating this uncertainty requires only to extend the integral over bin k to one over the entire phase space. Let us define a discrete function

$$
\eta_k(z) = = \begin{cases} 1 & \text{for } z_k < z < z_{k+1} \\ 0 & \text{otherwise} \end{cases}
\tag{B.11}
$$

If this function is inserted in the integrand of Equation B.8, we can rewrite it as

$$
I^k = \frac{1}{\Delta z} \int \mathrm{d}z \mathrm{d}x \mathrm{d}y \mathrm{d}\mathbf{w} \; f(x, y, z, \mathbf{w}) \frac{g(x, y)\epsilon(x, y, z, \mathbf{w})\eta_k(z)}{f(x, y, z)}
\tag{B.12}
$$

where the integral is now taken over all values of z. Given a total number of generated events N, the estimate of the moment integral becomes

$$
\begin{aligned}
I_N^k &= \frac{1}{N \Delta z} \sum_{\text{all events}} \frac{g(x_i, y_i)\epsilon(x_i, y_i, z_i, w_i)\eta_k(z)}{f(x_i, y_i, z_i)} \\
&= \frac{1}{N \Delta z} \sum_{\text{selected events in bin } k} \frac{g(x_i, y_i)}{f(x_i, y_i, z_i)}
\end{aligned}
\tag{B.13}
$$

231

with estimated uncertainty

$$\sigma(I_N^k) \;=\; \frac{1}{N\Delta z}\sqrt{\sum\left(\frac{g(x_i,y_i)}{f(x_i,y_i,z_i)}\right)^2 - \frac{1}{N}\left(\sum\frac{g(x_i,y_i)}{f(x_i,y_i,z_i)}\right)^2}. \qquad (\text{B.14})$$

Since the expectation value for N_k is given by Nc_k, there is a remarkable resemblance between the uncertainties in Equation B.14 and Equation B.10.

BIBLIOGRAPHY

[1] See for example Section 2.2 in [14].

[2] Wu, C.S.,Ambler, E., Haywood, R. W., Hoppes, D. D., and Hudson, R. P. (1957). Phys. Rev. **105**, 1413.

[3] M. Goldhaber, L. Grodzins, and A. Sunyar, Phys. Rev. **109**, 1015 (1958).

[4] J.H. Christensen, J.W. Cronin, V.L Fitch, and R. Turlay, Phys. Rev. Lett. **13**, 138 (1964).

[5] A. D. Sakharov, *ZhETF Pis. Red.* **5**, 32 (1967); *JETP Lett.* **5**, 24 (1967).

[6] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **87**, 091801 (2001).

[7] BELLE Collaboration, K. Abe *et al.*, Phys. Rev. Lett. **87**, 091802 (2001).

[8] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **89**, 201802 (2002).

[9] *BABAR* Collaboration, B.Aubert *et al.*, Phys. Rev. D **66**, 032003 (2002).

[10] AMS Collaboration, J. Alcaraz *et at.*, Phys. Lett. **B 461**, 387-396 (1999); M. Aguilar *et al.*, Physics Reports **366** (2002), pp.331-404.

[11] D. Griffiths, *Introduction to Elementary Particles*, John Wiley&Sons, Inc. , New York (1987), ISBN 0-471-60386-4.

[12] F. Halzen and A.D. Martin, *Quarks and Leptons*, Wiley, New York (1984).

[13] D.H. Perkins, *Introduction to High Energy Physics*, Addison-Wesley, New York (1987), ISBN 0-201-12105-0.

[14] M. Peskin and D. Schroeder, *An Introduction to Quantum Field Theory*, Perseus Books, Massachusetts (1995).

[15] P.F. Harrison and H.R. Quinn, eds. "The *BABAR* Physics Book", SLAC-R504 (1998), Chapter 2, and references therein.

[16] A. Manohar and M. Wise *Heavy Quark Physics*, Cambridge University Press (2000).

[17] G. Branco, L. Lavoura, and J. Silva, *CP Violation*, Clarendon Press, Oxford (1999).

[18] Lee, T.D. and Yang, C. N.(1956). Phys. Rev. **104**, 254.

[19] [17] page 335 calculates this limit using Baluni, V. (1979). Phys. Rev. D **19**, 2227, Crewther, R. J., Di Vecchia, P., Veneziano, G., and Witten, E. (1979), Phys. Rev. Lett. **88B**, 123, and Altarev, I.S. *et al.* (1992), Phys. Rev. Lett. **B276**, 242.

[20] See, for example, G. Buchalla, A.J. Buras and M.E. Lautenbacher, Rev. Mod. Phys. **68**, 1125 (1996); A.J. Buras, "Weak Hamiltonian, CP Violation and Rare Decays," in *Probing the Standard Model of Particle Interactions*, ed. F. David and R. Gupta (Elsevier Science B.V. ,1998)

[21] See Figure 39.9 in [50] and references therein.

[22] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963);
M. Kobayashi and T. Maskawa, Prog. Th. Phys. **49**, 652 (1973).

[23] Hocker, A., Lacker, H., Laplace, S., Le Diberder, F. R., *Eur. Phys. J.* **C21** (2001), 225-259.

[24] W. Grimus and M.N. Rebelo, Phys. Rev. B **427**, 172 (1997).

[25] L. Wolfenstein, Phys. Rev. Lett. **51** (1983) 1945.

[26] See for example, Particle Data Group, K. Hagiwara *et al.*, Phys. Rev. D **66**, 010001 (2002) pg.xxx-yyy.

[27] M. Gronau, Phys. Rev. Lett. **63**, 1451 (1989).

[28] A.I. Sanda and Z.Z. Xing, Phys. Rev. D **56**, 341 (1997).

[29] X.Y. Pham and Z.Z. Xing, Phys. Lett. B **458**, 375 (1999).

[30] J. Rosner, Phys. Rev. D **42**, 3732 (1990).

[31] R. Aleksan, *et al.*, Phys. Lett. B **317**, 173 (1993).

[32] Z.Z. Xing, Phys. Rev. D **61**, 014010 (2000).

[33] A. Datta and D. London, hep-ph/0310252

[34] M. Neubert, hep-ph/0001334,...

[35] P.F. Harrison and H.R. Quinn, eds. "The *BABAR* Physics Book", SLAC-R504 (1998), Chapter 5, and references therein. (Eq. 5.44 in this reference appears with wrong signs in the last two terms.)

[36] I. Dunietz *et al.*, Phys. Rev. D **43**, 2193 (1991).

[37] C.P. Jessop *et al.*, (CLEO collaboration), Phys. Rev. Lett. **79**, 4533 (1997).

[38] G. Michelon, "A Theoretical Review of the $B \to D^{(*)}\overline{D}^{(*)}$ Decay Channel", *BABAR* Note #342, (1996).

[39] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **89**, 201802 (2002).

[40] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **91**, 131801 (2003).

[41] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **90**, 221801 (2002).

[42] E. Lipeles, *et al.*, Phys. Rev. D **62** (2000).

[43] K. Abe, *et al.*, Belle-Conf-0104 (2001).

[44] P. Oddone, Proceedings of the UCLA Workshop on Linear Collider B B-Factory Conceptual Design, ed. by D. Stork (1987).

[45] *BABAR* Collaboration, B. Aubert *et al.*, Nucl. Instr. Meth. A **479**, 1 (2002).

[46] A. Drescher *et al.*, Nucl. Instr. and Methods **A237**, 464 (1985).

[47] R. Kalman, Journal of Basic Engineering, 35 (1960).

[48] R. Frühwirth, Nucl. Instrum. Methods Phys. Res. **A262**, 444 (1987).

[49] Throughout this paper, flavor-eigenstate decay modes also imply their charge conjugate.

[50] Particle Data Group, K. Hagiwara *et al.*, Phys. Rev. D **66**, 010001 (2002).

[51] G.C. Fox and S. Wolfram, Phys. Rev. Lett. **41**, 1581 (1978).

[52] Paul Avery, *Applied Fitting Theory*, CLEO Note CBX 91-72 (1991). Available online at
http://www.phys.ufl.edu/∼avery/fitting.html.

[53] For a summary of all reconstructed lists see

http://www.slac.stanford.edu/BFROOT/www/doc/workbook/

nanomicro/newest/Micro/CandLists.html

[54] J. Smith, A. Soffer, and R. Waldi, *Reccomendation for Exclusive B Reconstruction Analysis Variables*, *BABAR* Note 497 (1999).

[55] Defined as $A \sim \sqrt{x} \times \exp(\kappa x)$ where $x = 1 - (m_{\mathrm{ES}}/m_0)^2$ and $m_{\mathrm{ES}} < m_0$. The value of m_0 is fixed to 5.291 GeV/c^2. ARGUS Collaboration, H. Albrecht *et al.*, Z. Phys. C **48**, 543 (1990).

[56] The reference manual for *Minuit, Function Minimization and Error Analysis* is available online at:

http://wwwinfo.cern.ch/asdoc/minuit/minmain.html .

[57] O. Long *et al.* *BABAR* Analysis Document 54 *Measurement of the slow π relative efficiency using helicity distributions* (Version 7)

[58] J. Beringer *et al.* *BABAR* Analysis Document 317 *BTagger – A multivariate Tagging Algorithm with Categories Based on the Physics of the B_{tag} Decay* (2002).

[59] F. Le Diberder, *CP Violation as seen from the Δz Distribution*, *BABAR* Note 42 (1990).

[60] F. Le Diberder, *Precision on CP-Violation Measurements and Requirement on the Vertex Resolution*, *BABAR* Note 34 (1990).

[61] S. Rahatlou *Observation of Matter-Antimatter Asymmetry in the B^0 Meson System*, SLAC-R-677 (2002).

[62] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **91**, 061801 (2003).

[63] *BABAR* Collaboration, *Measurement of sin2beta in $B^0 \rightarrow phi\ KS$* hep-ex/0207070 (2004).

[64] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **91**, 161801 (2003).

[65] *BABAR* Collaboration, B. Aubert *et al.*, hep-ex/0403001 Submitted to Phys. Rev. Lett. (2004).