SLAC-R-645

A Measurement of the Lifetime and Mixing Frequency of Neutral B Mesons with Semileptonic Decays in the BABAR Detector*

Chih-Hsiang Cheng

Stanford Linear Accelerator Center
Stanford University
Stanford, CA 94309

SLAC-Report-645

Prepared for the Department of Energy
under contract number DE-AC03-76SF00515

Printed in the United States of America. Available from the National Technical Information Service, U.S. Department of Commerce, 5285 Port Royal Road, Springfield, VA 22161.

* Ph.D. thesis, Stanford University, Stanford, CA 94305.

A MEASUREMENT OF THE LIFETIME AND MIXING FREQUENCY OF NEUTRAL B MESONS WITH SEMILEPTONIC DECAYS IN THE BABAR DETECTOR

A DISSERTATION

SUBMITTED TO THE DEPARTMENT OF PHYSICS

AND THE COMMITTEE ON GRADUATE STUDIES

OF STANFORD UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Chih-Hsiang Cheng

August 2002

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Patricia Burchat
(Principal Adviser)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Aaron Roodman

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Lance Dixon

Approved for the University Committee on Graduate Studies:

Abstract

The neutral B meson, consisting of a b quark and an anti-d quark, can mix (oscillate) to its own anti-particle through second-order weak interactions. The measurement of the mixing frequency can constrain the quark mixing matrix in the Standard Model of particle physics. The PEP-II B-factory at the Stanford Linear Accelerator Center provides a very large data sample that enables us to make measurements with much higher precisions than previous measurements, and to probe physics beyond the Standard Model.

The lifetime of the neutral B meson τ_{B^0} and the B^0-$\overline{B}^0$ mixing frequency Δm_d are measured with a sample of approximately 14,000 exclusively reconstructed $B^0 \to D^{*-}\ell^+\nu_\ell$ signal events, selected from 23 million $B\overline{B}$ pairs recorded at the $\Upsilon(4S)$ resonance with the *BABAR* detector at the asymmetric-energy e^+e^- collider, PEP-II. The decay position of the exclusively reconstructed B is determined by the charged tracks in the final state, and its b-quark flavor at the time of decay is known unambiguously from the charge of the lepton. The decay position of the other B is determined inclusively, and its b-quark flavor at the time of decay is determined (tagged) with the charge of tracks in the final state, where identified leptons or kaons give the most information. The decay time difference of two B mesons in the event is calculated from the distance between their decay vertices and the Lorentz boost of the center of mass. Additional samples of approximately 50,000 events are selected for studies of background events.

The lifetime and mixing frequency, along with wrong-tag probabilities and the time-difference resolution function, are measured simultaneously with an unbinned maximum-likelihood fit that uses, for each event, the measured difference in B decay times (Δt), the calculated uncertainty on Δt, the signal and background probabilities, and b-quark tagging for the other B meson. The results are

$$\tau_{B^0} = (1.523\,^{+0.024}_{-0.023} \pm 0.022)\ \text{ps}$$

and

$$\Delta m_d = (0.492 \pm 0.018 \pm 0.013)\ \text{ps}^{-1}\ ,$$

where the first error is statistical and the second is systematic. The statistical correlation coefficient between τ_{B^0} and Δm_d is -0.22.

This result is consistent with the current world average values, the the total errors are comparable with other most-precise measurements.

Acknowledgement

As I put a period to my journey through graduate school, I would like to acknowledge the guidance and support of my adviser Pat Burchat. I would also like to thank David Kirkby for his always-ingenious suggestions and insight, Tim Meyer, and Chris Roat for their camaraderie in finishing this ambitious analysis and making my graduate life so enjoyable, and Chris LeClerc at U. C. Berkeley, who, with David and Tim, has contributed a great deal to this analysis. I would also like to thank Wouter Verkerke for providing an amazing fitting package and answering numerous questions. This work cannot be accomplished without the dedication of all *BABAR* and PEP-II members. I appreciate the careful reading and comments by Pat Burchat and Lance Dixon in preparing this manuscript.

Finally, I would like to thank my family for the support and encouragement, especially my wife, Joyce, who has added the pleasure to my life. Without her, this thesis would not have been possible.

Contents

List of Tables

List of Figures

Chapter 1

Introduction and Overview

The time evolution of neutral B mesons is governed by the overall decay rate $\Gamma \equiv 1/\tau_{B^0}$, and the B^0-$\overline{B}^0$ oscillation frequency Δm_d. [1] The phenomenon of particle-anti-particle oscillation or mixing has been observed in neutral mesons containing a down quark and a strange quark (K mesons), or a bottom quark (B mesons) [1, 2]. By interchanging the quark content of B^0 ($\overline{b}d$) with that of $\overline{B}^0$ ($b\overline{d}$), B^0-$\overline{B}^0$ mixing changes the bottom quantum number by two units; i.e., $|\Delta B| = 2$. In the Standard Model of particle physics, such a process is the result of second-order charged weak interactions involving box diagrams containing virtual quarks with charge 2/3, illustrated in Fig. 1.1. In B^0-$\overline{B}^0$ mixing, the diagram containing the top quark dominates. Therefore, the mixing frequency Δm_d is sensitive to the Cabibbo-Kobayashi-Maskawa quark-mixing matrix element V_{td} [3, 4]. Because of this interaction, the mass eigenstates of this two-state system in B^0-$\overline{B}^0$ space are no longer degenerate. The mixing frequency Δm_d is equal to the difference between the two mass eigenvalues of neutral B meson states. In the neutral K meson system, mixing also has contributions from real intermediate states accessible to both K^0 and $\overline{K}^0$. As a

[1] Through out this thesis, the fundamental constant c and $\hbar$ are ignored, that is, units are changed so that mass and momentum are in the unit of energy, e.g., MeV, or equivalently, in the unit of inverse time, e.g., ps^{-1}.

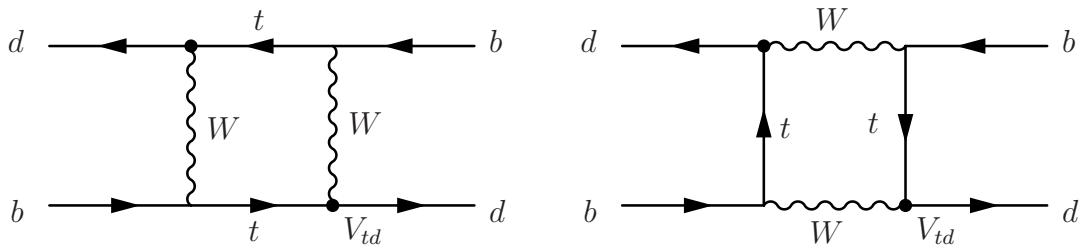

Figure 1.1: Second-order charged weak interaction that interchanges $b\overline{d}$ with $\overline{b}d$. The virtual quark in the loop can be u, c, and t, but the diagram containing the top quark dominates.

result, the two mass eigenstates of the neutral K system have a large decay rates difference $\Delta\Gamma$. For the B^0 system, $\Delta\Gamma$ is much smaller than Δm_d. Therefore, in this analysis, $\Delta\Gamma$ is assumed to be negligible.

In the naive spectator model [5], the decay rate of hadrons containing a b quark and one or two light quarks (u, d, or s) is determined by the decay rate of the b quark, and therefore their lifetimes should be the same. However, some b-hadrons have significantly different lifetimes from others[2]. A precise measurement of the B^0 lifetime, along with other b-hadron (e.g., B^+, B_s, Λ_b) lifetimes, can improve our understanding of the effect of the light quarks in the b-hadron.

In *BABAR* [7], $b\overline{b}$ bound states called $\Upsilon(4S)$ are created by the asymmetric-energy e^+e^- collider, PEP-II [8], at the Stanford Linear Accelerator Center. More than 96% of the $\Upsilon(4S)$ states decay to a pair of $B^0\overline{B}^0$ or B^+B^- [6]. The two B mesons evolve coherently, until one decays. The proper decay-time difference Δt between two neutral B mesons is governed by the following probabilities to observe an *unmixed* event, i.e., two B mesons decaying to opposite flavors,

$$P(B^0\overline{B}^0 \to B^0\overline{B}^0) \propto e^{-|\Delta t|/\tau_{B^0}}(1+\cos\Delta m_d\Delta t)\ , \tag{1.1}$$

or a *mixed* event; i.e., two B mesons decaying to the same flavor because one B^0 has

[2]For example, the lifetime of Λ_b is about 20% smaller than that of B^0 [6].

oscillated to its anti-particle $\overline{B}^0$, or vice versa,

$$P(B^0\overline{B}^0 \to B^0B^0 \text{ or } \overline{B}^0\overline{B}^0) \propto e^{-|\Delta t|/\tau_{B^0}}(1-\cos\Delta m_d\Delta t)\ . \tag{1.2}$$

Therefore, if Δt is measured and the b-quark flavor of both B mesons at their time of decay is measured, one can extract the B^0 lifetime τ_{B^0} and the mixing frequency Δm_d.

In this analysis, one B^0 (referred to as $B_{\rm rec}$) is reconstructed in the mode $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_\ell$,[3] where ℓ is an electron or a muon. The charges of the final state particles identify the B as a B^0 or $\overline{B}^0$. The remaining charged particles in the event, which originate from the other B meson (referred to as $B_{\rm tag}$), are used to reconstructed the $B_{\rm tag}$ decay vertex, and to determine, or *tag*, its b-flavor content.

Since the mass of two B mesons is only slightly less than the mass of the $\Upsilon(4S)$, the B mesons move very slowly ($v \simeq 0.06c$) in the center-of-mass frame. In *BABAR*, because the e^+ and e^- in the PEP-II storage rings have different energy, the center-of-mass frame, i.e., the $\Upsilon(4S)$ rest frame, is boosted along the beam axis ($\simeq z$-axis) with a known Lorentz factor $\beta\gamma \simeq 0.55$, so that the decay vertices of the B meson pair can be separated. The decay-time difference Δt can be calculated from the separation of the two decay vertices along the z-axis, $\Delta t \equiv t_{\rm rec} - t_{\rm tag} \simeq \Delta z/\beta\gamma c$, where $t_{\rm rec}$ and $t_{\rm tag}$ are the proper decay times of the $B_{\rm rec}$ and $B_{\rm tag}$ mesons, respectively. The average separation is about 250 μm.

The theoretical Δt distributions of unmixed and mixed events, given in Eq. 1.1 and 1.2, and the time-dependent asymmetry, defined as

$$A(\Delta t) = \frac{N_{\rm unmixed}(\Delta t) - N_{\rm mixed}(\Delta t)}{N_{\rm unmixed}(\Delta t) + N_{\rm mixed}(\Delta t)}\ , \tag{1.3}$$

[3]Through out this thesis, the charge conjugate modes are implied, unless specified otherwise.

are shown in Fig. 1.2, in which the world average values [9] of the lifetime and mixing frequency are used. The time-dependent asymmetry is equal to $\cos(\Delta m_d \Delta t)$, with the oscillation amplitude equal to unity. In reality, there are three major experimental complications that affect the Δt distributions shown in Fig. 1.2. The tagging algorithms that determine the b-flavor of $B_{\rm tag}$ at its time of decay have a finite probability of making a mistake, called the *mistag* probability, ω. As a result, some mixed events will be tagged as unmixed events, and vice versa. An example of this effect with $\omega = 0.1$, which represents a quite good tagging power in reality, is shown in the left plot in Fig. 1.3. The Δt distributions become

$$P(\Delta t) = \frac{e^{-|\Delta t|/\tau_{B^0}}}{4\tau_{B^0}} \left[1 \pm (1-2\omega)\cos\Delta m_d \Delta t\right] \ ,$$

for the unmixed $(+)$ and mixed $(-)$ events, and the asymmetry amplitude becomes $1-2\omega$. Second, the resolution for Δt is comparable to the lifetime and must be well understood. The effect of non-perfect resolution is shown in the right plot in Fig. 1.3. A final complication is that the sample of selected $B^0 \to D^{*-}\ell^+\nu_\ell$ candidates is not pure signal. Various background levels and their time evolution properties must be understood, in order to extract the lifetime and mixing frequency.

The $B_{\rm rec}$ is reconstructed in the decay chain $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_\ell$, $D^{*+} \to D^0\pi^+$, and $D^0 \to K^-\pi^+$, $K^-\pi^+\pi^0$, $K^-\pi^+\pi^-\pi^+$, or $K_S\pi^+\pi^-$. Although the neutrino cannot be detected, the requirement of a reconstructed $D^{*-} \to \overline{D}^0\pi^-$ and an identified high-momentum lepton that satisfies the kinematic constraints consistent with a $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_\ell$ decay allows the isolation of a signal sample with (65 – 89)% purity, depending on the D^0 decay mode and whether the lepton candidate is an electron or a muon. To characterize the backgrounds, a set of control samples enriched in each type of background is selected along with the signal sample to determine the signal and background probabilities of each events in all samples.

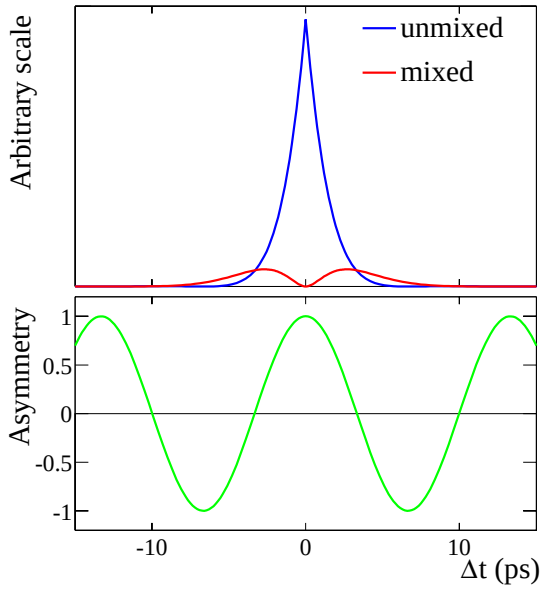

Figure 1.2: Theoretical expectation of B^0 decay Δt distributions and the asymmetry between unmixed and mixed events.

The mixing frequency Δm_d and the average lifetime of the neutral B meson, τ_{B^0}, are measured simultaneously with an unbinned maximum-likelihood fit to measured Δt distributions for events that are classified as mixed and unmixed. This is in contrast to published measurements in which only Δm_d is measured with τ_{B^0} fixed to the world average value, or only τ_{B^0} is measured. There are several reasons to measure the lifetime and mixing frequency simultaneously. The statistical precision for both τ_{B^0} and Δm_d is comparable to the uncertainty on the world average. Therefore, it is appropriate to measure both quantities rather than fixing the lifetime to the world average. Since mixed and unmixed events have different Δt distributions, the mixing information for each event gives greater sensitivity to the Δt resolution function and a smaller statistical uncertainty on τ_{B^0}. Also, since B^+B^- events do not mix, the Δt distributions for mixed and unmixed events can help discriminate between $B^0\overline{B}^0$ signal events and B^+B^- background events in the lifetime and mixing measurement.

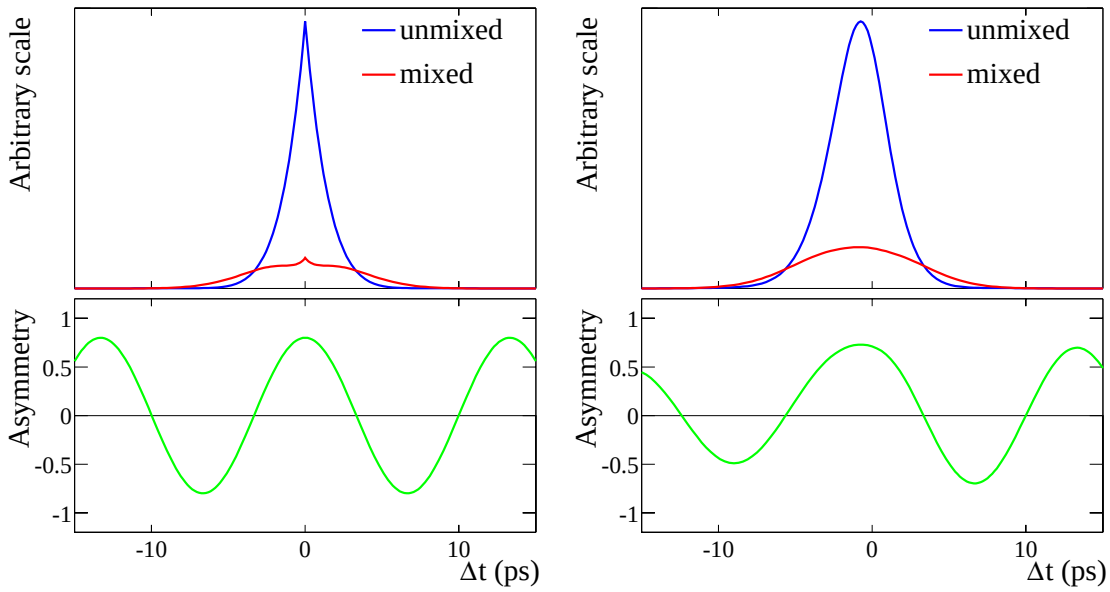

Figure 1.3: Experimental effects on the expected Δt distributions and the asymmetry between unmixed and mixed events. (a) Effect due to non-perfect flavor tagging. (b) Effect due to a non-perfect Δt determination and flavor tagging. A resolution function with a realistic width $\sigma \simeq 1$ ps and a somewhat exaggerated negative tail is used.

This thesis is organized in the following way. The physics that is related to the neutral B meson time evolution is discussed in Ch. 2. A review of the current experimental status of B lifetime and mixing measurements is given in Ch. 3. The detector facility, and the particle reconstruction and identification are described in Ch. 4 and Ch. 5. The measurements of B decay vertices and decay-time difference are discussed in Ch. 6. The algorithms that identify the b-flavor of $B_{\rm rec}$ are given in Ch. 7. The event selection for signal and background control samples, and the calculation of signal and background probabilities for each event are described in Ch. 8 and Ch. 9. In Ch. 10, the full analysis of the measured decay time difference is presented, including physics models and Δt resolution functions used to describe the measured Δt distribution for signal and background, the unbinned maximum-likelihood fit and

the fit results. The validation exercises and systematic uncertainty studies are shown in Ch. 11. Finally the conclusions and future prospect are presented in Ch. 12.

Chapter 2

Theory and Phenomenology

Overview

In this chapter, I review the theoretical aspects of the time evolution of the neutral B meson system. The phenomenon of B^0-$\overline{B}^0$ oscillation is governed by the *electroweak interaction* in the *Standard Model* of particle physics, which provides an elegant theoretical framework and represents our current understanding of the fundamental building blocks of nature and their interactions, except for gravity. In this chapter, I start with a brief overview of the Standard Model and focus on the electroweak interaction that eventually leads to the quark mixing matrix (the CKM matrix). I then demonstrate the formalisms for the time evolution of the neutral B meson and the coherent $\Upsilon(4S) \to B\overline{B}$ state, and their implication for C, P and T symmetries. The phenomenology for b-hadron lifetimes is also discussed. Finally, the kinematics of semileptonic decays of the B meson are presented.

2.1 Standard Model Physics

Particle physics theory is usually formulated by a Lagrangian that is a function of matter fields and their first derivatives. The Lagrange density $\mathcal{L}$ is related to the classical action through

$$S = \int d^4x \mathcal{L}(\phi, \partial_\mu \phi) \, ,$$

where $\phi \equiv \phi(x)$ is a generalized coordinate as a function of space-time, and $\partial_\mu \phi \equiv \partial\phi/\partial x^\mu$, in which μ is a space-time index. Through the principle of least action, one can deduce the Euler-Lagrangian equations for the Lagrange density,

$$\frac{\partial \mathcal{L}}{\partial \phi(x)} = \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi(x))} \, ,$$

which lead to the explicit form of equations of motion for the fields. A Lagrangian ("density" is dropped for convenience) must be a scalar and be invariant under Lorentz transformations, so that all predictions of the theory are Lorentz invariant.

The existence of symmetries plays an important role in the development of the Standard Model of particle physics. A physical system is said to have a particular symmetry if the Lagrangian $\mathcal{L}$ that describes this system is invariant under a certain transformation. A set of transformations that leave a system invariant can have the algebraic structure of a group, in which case there is said to be a symmetry group. The types of symmetry transformation can be categorized in various ways; e.g., whether the parameters of the transformation are discrete or continuous, whether the transformation is acting on space-time (geometrical) or on the internal space, and whether the transformation varies at different space-time points (local) or not (global). Continuous symmetries of the Lagrangian and constants of motion are closely related. This connection is embodied in Noether's theorem [10], which states that for a system described by a Lagrangian, any continuous symmetry that leaves

invariant the action leads to the existence of a conserved current S_μ with $\partial^\mu S_\mu = 0$.

In the Lagrangian formalism, the interaction arises naturally when local gauge invariance is imposed on the system. In nature, matter consists of spin-1/2 fermions; i.e., quarks and leptons. The Lagrangian for a massless spin-1/2 fermion can be expressed as

$$\mathcal{L} = i\overline{\psi}\gamma^\mu\partial_\mu\psi\,,$$

where ψ is a Dirac spinor. In order to maintain invariance under the local phase transformation (called gauge transformation for historical reasons) $\psi(x) \rightarrow e^{i\alpha(x)}\psi(x)$, one needs to introduce a vector field A^μ and replace the space-time derivative ∂^μ with a covariant derivative $\mathcal{D}^\mu = \partial^\mu - igA^\mu$, where g represents the coupling strength and A^μ is the field of a massless spin-1 gauge boson, which mediates this particular interaction. This formalism does not restrict the symmetry group of the local gauge transformation.

The Standard Model is based on an $SU(3) \times SU(2) \times U(1)$ gauge theory. This is basically saying that the particles in this theory have an internal space that is invariant under local gauge transformations that belong to $SU(3)$, $SU(2)$ and $U(1)$ symmetry groups. The $U(1)$ group is a simple (one-dimensional) phase rotation. For an $SU(n)$ group, the physical state is represented by an n-dimensional multiplet, and the gauge transformation and the corresponding covariant derivative are

$$\psi \rightarrow \exp(i\sum_{a=1}^{n^2-1}\theta_a T^a)\psi\,,$$
$$\mathcal{D}^\mu = \partial^\mu - ig_n\sum_{a=1}^{n^2-1}T^a G_a^\mu\,,$$

where θ_a is a rotation angle, T^a represents n^2-1 rotation generators of this group, and G_a^μ are gauge boson fields. $SU(3)$ is the symmetry group of the strong interaction,

which is described by the theory of *quantum chromodynamics*, QCD (the charge of the strong interaction is called *color* charge). The electroweak part ($SU(2) \times U(1)$) is referred to as the Glashow-Salam-Weinberg model [11, 12, 13].

The full covariant derivative can be written as

$$\mathcal{D}^\mu = \partial^\mu - ig_1 Y B^\mu - ig_2 T^i W_i^\mu - ig_3 \Lambda^a G_a^\mu \,,$$

where $i = 1 \ldots 3$, $a = 1 \ldots 8$, and Y, T^i, and Λ^a are generators of $U(1)$, $SU(2)$ and $SU(3)$, respectively, and B^μ, W_i^μ and G_a^μ are the corresponding gauge boson fields.

So far in the Lagrangian the gauge bosons are all massless. However, the fact that the weak gauge bosons are massive particles shows that electroweak $SU(2) \times U(1)$ is not a symmetry of the vacuum. The massive weak bosons, $W^\pm$ and Z acquire mass through the spontaneous breakdown of $SU(2) \times U(1)$, called the *Higgs mechanism* [14], which requires a single Higgs doublet[1] to acquire a vacuum expectation value. One linear combination of the $SU(2) \times U(1)$ generators is left unbroken, the electric charge generator $Q = T^3 + Y/2$. The field sources of the electroweak interaction include left-handed quark and lepton iso-doublets, right-handed quark and lepton iso-singlets and a scalar Higgs iso-doublet. These are summarized in Table 2.1, including their $SU(2)$ quantum numbers, $|T|$, T^3, $U(1)$ hypercharge Y, and electric charge Q.

$SU(3) \times SU(2) \times U(1)$ gauge invariance also prevents bare mass terms for quarks and leptons from appearing in the Lagrangian. The quarks and leptons get mass from the Yukawa couplings to the Higgs doublet,

$$\mathcal{L}_{\text{Yukawa}} = g_u^{ij} \overline{u}_R^i H^T \epsilon Q_L^j - g_d^{ij} \overline{d}_R^i H^\dagger Q_L^j - g_e^{ij} \overline{e}_R^i H^\dagger L_L^j + \text{h.c.} \,, \tag{2.1}$$

[1] A single Higgs doublet is the simplest way to achieve the observed spontaneous symmetry breaking, but a more complicated scalar sector is possible.

Table 2.1: The three families of leptons and quarks, the single Higgs doublet, and their electroweak quantum numbers. The index i runs from 1 to 3, representing three families.

fields				$\|T\|$	T^3	Y	Q
$L_L^i =$	$\begin{pmatrix}\nu_e\\ e\end{pmatrix}_L$	$\begin{pmatrix}\nu_\mu\\ \mu\end{pmatrix}_L$	$\begin{pmatrix}\nu_\tau\\ \tau\end{pmatrix}_L$	$1/2$	$+1/2$	-1	0
				$1/2$	$-1/2$	-1	-1
$e_R^i =$	e_R	μ_R	τ_R	0	0	-2	-1
$Q_L^i =$	$\begin{pmatrix}u\\ d\end{pmatrix}_L$	$\begin{pmatrix}c\\ s\end{pmatrix}_L$	$\begin{pmatrix}t\\ b\end{pmatrix}_L$	$1/2$	$+1/2$	$1/3$	$+2/3$
				$1/2$	$-1/2$	$1/3$	$-1/3$
$u_R^i =$	u_R	c_R	t_R	0	0	$+4/3$	$+2/3$
$d_R^i =$	d_R	s_R	b_R	0	0	$-2/3$	$-1/3$
$H =$	$\begin{pmatrix}H^+\\ H^0\end{pmatrix}$			$1/2$	$+1/2$	1	$+1$
				$1/2$	$-1/2$	1	0

where i and j are quark-lepton family indices and are summed if repeated, and h.c. denotes Hermitian conjugate. The antisymmetric matrix ϵ is given by

$$\epsilon = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} .$$

The Higgs field has a vacuum expectation value v at the minimum of the Higgs potential and a charge-neutral field h^0 along the symmetry-breaking direction,

$$H = \begin{pmatrix} 0 \\ (v + h^0)/\sqrt{2} \end{pmatrix} .$$

The terms in Eq. 2.1 become

$$g_x^{ij} \frac{v}{\sqrt{2}} \overline{x}_R^i x_L^j + g_x^{ij} \frac{h^{0(\dagger)}}{\sqrt{2}} \overline{x}_R^i x_L^j \,, \tag{2.2}$$

for $x = u\,, d\,, \text{or}\, e$. The first term gives rise to fermion masses, and the second term

represents the interaction between fermions and the Higgs boson.

The 3×3 quark and lepton mass matrices, $M_x^{ij} = g_x^{ij} v/\sqrt{2}$, can be transformed into diagonal matrices by separate unitary matrices,

$$V_{uR} M_u V_{uL}^\dagger = D_u \,, \quad V_{dR} M_d V_{dL}^\dagger = D_d \,, \quad \text{and} \quad V_{eR} M_e V_{eL}^\dagger = D_e \,,$$

where V_{xL} and V_{xR} are unitary matrices and D_x are diagonal and real. The quark and lepton fields can be transformed accordingly,

$$\begin{aligned} u_L^m &= V_{uL} u_L \,, \quad u_R^m = V_{uR} u_R \,, \\ d_L^m &= V_{dL} d_L \,, \quad d_R^m = V_{dR} d_R \,, \\ e_L^m &= V_{eL} e_L \,, \quad e_R^m = V_{eR} e_R \,, \end{aligned}$$

where mass eigenstates are labeled by the superscript m, so that the mass terms in Eq. 2.2 become diagonal, e.g., $\overline{u}_R^m D_u u_L^m$. The weak iso-doublet of left-handed quarks can be rewritten as

$$\begin{pmatrix} u_L \\ d_L \end{pmatrix} = \begin{pmatrix} V_{uL}^\dagger u_L^m \\ V_{dL}^\dagger d_L^m \end{pmatrix} = V_{uL}^\dagger \begin{pmatrix} u_L^m \\ V d_L^m \end{pmatrix} \,,$$

where $V = V_{uL} V_{dL}^\dagger$ is called the Cabibbo-Kobayashi-Maskawa (CKM) [3, 4] mixing matrix. The overall unitary transformation by $V_{uL}^\dagger$ has no effect, but the unitary matrix V changes the form of the weak charged-current interactions; for example,

$$\frac{g_2}{\sqrt{2}} W_\mu^+ \overline{u}_L \gamma^\mu d_L = \frac{g_2}{\sqrt{2}} W_\mu^+ \overline{u}_L^m \gamma^\mu V d_L^m \,. \tag{2.3}$$

The quark kinematic energy terms and the weak neutral current and electromagnetic interactions are unaffected. The right-hand side of Eq. 2.3 shows that the mass

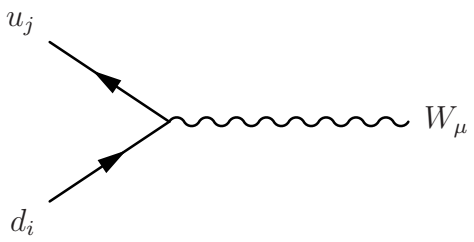

Figure 2.1: Diagram for quark flavor changing charged current.

eigenstate of an up-type quark (charge $+2/3$) can interact with a mixture of the mass eigenstates of down-type quarks (charge $-1/3$) of different families at the tree level through weak charged currents. Conventionally, the mass eigenstates of quarks of three families are labeled as (u, c, t) and (d, s, b). The weak eigenstates shown in the weak iso-doublet are labeled in the same way for up-type quarks and with a prime for down-type quarks,

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

The Feynman rule that describes the flavor-changing charged current between a down-type quark of family i and an up-type quark of family j (Fig. 2.1) is then

$$-i \left(\frac{G_F M_W^2}{\sqrt{2}} \right)^{1/2} \bar{u}_j \gamma_\mu (1 - \gamma_5) d_i \cdot V_{ij} \, ;$$

i.e., the coupling strength is proportional to the CKM matrix element V_{ij}.

The unitary CKM matrix for three families can in general be parameterized by nine real numbers. However, one can rotate the phases of the six quarks independently (while keeping the relative phase of left- and right-handed quarks of the same flavor unchanged). An overall phase leaves the CKM matrix unchanged. The remaining rotations can eliminate five phases and leave three Euler angles and one

phase to parameterize the CKM matrix. It has been found that the CKM matrix has a hierarchical structure; i.e.,

$$|V_{ud}|^2 \simeq |V_{cs}|^2 \simeq |V_{tb}|^2 \simeq 1 \quad \text{and} \quad |V_{ub}|^2 \ll |V_{cb}|^2 \ll |V_{us}|^2 \ll 1\,.$$

It is best illustrated by the Wolfenstein parameterization [15], which is a power-series expansion in $\lambda = \sin\theta_C \simeq 0.22$, where θ_C is called the Cabibbo angle [3]. Up to $\mathcal{O}(\lambda^3)$, it is expressed as

$$V = \begin{pmatrix} 1-\frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}\left(\lambda^4\right)\,. \tag{2.4}$$

The three Euler angles and one phase in the matrix are now parameterized by the four real numbers, λ, A, ρ, and η. Their absolute values, except for λ, are all of order unity.

Since the CKM matrix is a unitary matrix, the products of the elements in any two columns satisfy

$$\sum_{k=1}^{3} V_{ki}V_{kj}^* = \delta_{ij}\,, \tag{2.5}$$

where $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise. Therefore the three terms in Eq. 2.5 from two different columns ($i \neq j$) can be visualized as a triangle in the complex plane, called a unitary triangle. There are three such triangles [2]. Two of them have one side of the triangle much smaller than the other two sides by $\mathcal{O}(\lambda^4)$ (columns 1 and 2) or $\mathcal{O}(\lambda^2)$ (columns 2 and 3). The remaining triangle, which is formed by columns 1 and 3, has three sides of the same order $\mathcal{O}(\lambda^3)$. Therefore the three angles of this triangle are more likely to be significantly different from 0 or π than the other

[2]There are actually three more triangles formed by the products of two different rows of the matrix.

two triangles. This triangle formed by columns 1 and 3 is most relevant to this thesis, It is usually presented after each side is divided by $V_{cd}V_{cb}^*$ so that one side is unit length and lies along the real axis. The three angles are denoted by α, β and γ. This unitary triangle is shown in Fig. 2.1, and the equation is given explicitly by

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0\,. \tag{2.6}$$

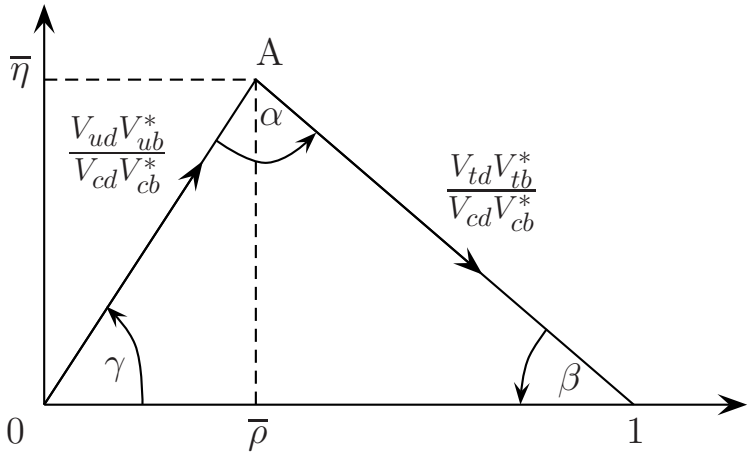

Figure 2.2: The normalized unitary triangle formed by columns 1 and 3 of the CKM matrix, where $\overline{\rho} = (1 - \lambda^2/2)\rho$ and $\overline{\eta} = (1 - \lambda^2/2)\eta$.

2.2 A Few Words on C, P and T Symmetry

There are three important finite (discrete) symmetry groups in particle physics: space-inversion (or parity, P), particle-antiparticle conjugation (C) and time-reversal (T) transformations. Each of them contains only two elements, the identity e and an element g, which satisfies $g^2 = e$. The P and C invariances require that g is represented by a unitary operator. The T invariance is the only symmetry that requires an antiunitary operator [16].

The week interaction violates the C and P symmetries, separately [17]. This is modeled by the $V-A$ structure of the electroweak theory in the Standard Model. The combined symmetry of CP was thought to be preserved until the discovery of the CP violation in the neutral kaon system [18]. The neutral kaon system remained the sole system that exhibits CP violation until the recent observations of CP violation in the neutral B system [19, 20]. The conservation of the combined symmetry of CPT is a consequence of the Lorentz symmetry of local quantum theories (see, for example, the text book [21]). So far no observation indicates that CPT is not a perfect symmetry in nature.

The phenomenon of CP violation can be accommodated in the three-family Standard Model. The magnitude of CP asymmetry is related to the area of the unitary triangle[3] [22]. The evidence of observed CP violation in the neutral B system was a non-zero β angle in Fig. 2.1.

There are three types of CP asymmetry: (a) *CP violation in decay* is due to the differences between the magnitudes of the decay amplitude $B \to f$ and its CP conjugate amplitude $\overline{B} \to \overline{f}$, which can occur for both neutral and charged mesons; (b) *CP violation in mixing* occurs when the two neutral mass eigenstate are not CP eigenstates; (c) *CP violation in the interference between decays with and without mixing*, caused by an extra phase in the decay amplitude with mixing, occurs in decays of neutral B into final states that are common to B^0 and $\overline{B}^0$. All these three types can coexist with each other. To a good approximation, the observed CP violation in B-factories [19, 20] is solely due to the type (c). The one that is most relevant to the analysis in this thesis is the CP violation in mixing. It will be discussed in Sec. 2.3.2.

[3] All six unitary triangles have the same area.

2.3 Neutral B Meson System

In this section, I present the phenomenology calculation that leads to a general description of the time evolution of a neutral B meson. I also show the summary of the Standard Model calculations of the mass and decay rate differences of the two mass eigenstates in the B^0-$\overline{B}^0$ system.

2.3.1 Two-state phenomenology

Considering a simplest case in which only two degenerate states $|P^0\rangle$ and $|\overline{P}^0\rangle$ exist and they can neither decay nor transform into each other. An arbitrary state can be represented by

$$\psi(t) = a(t)|P^0\rangle + b(t)|\overline{P}^0\rangle \equiv \begin{pmatrix} a(t) \\ b(t) \end{pmatrix} .$$

The Hamiltonian in the Schrodinger equation can simply be written as a diagonal matrix with real elements

$$\mathcal{H} = \begin{pmatrix} m_P & 0 \\ 0 & m_P \end{pmatrix} ,$$

and $|P^0\rangle$ and $|\overline{P}^0\rangle$ are mass eigenstates with an eigenvalue m_P.

If an interaction is turned on so that these two states can turn into each other, directly or through intermediate states, in a simplest case without expanding the Hilbert space, the Hamiltonian can be written as

$$\mathcal{H} = \begin{pmatrix} m_P & \Delta \\ \Delta & m_P \end{pmatrix} ,$$

where $\Delta = \langle P^0|\mathcal{H}|\overline{P}^0\rangle = \langle \overline{P}^0|\mathcal{H}|P^0\rangle$. No matter how small Δ is, the mass eigenstates

of this Hamiltonian become

$$|P^0_{1,2}\rangle = \frac{1}{\sqrt{2}}(|P^0\rangle \pm |\overline{P}^0\rangle)\,,$$

with eigenvalues $m_P \pm \Delta$.

In the real B^0-$\overline{B}^0$ system, we have to consider the full Hilbert space including all final and intermediate states that B^0 and $\overline{B}^0$ can access. A general state is then given by

$$\psi(t) = a(t)|B^0\rangle + b(t)|\overline{B}^0\rangle + \sum_n c_n(t)|f_n\rangle\,.$$

The solution of the Schrodinger equation for the full Hamiltonian

$$i\hbar\frac{\partial}{\partial t}\Psi = \mathcal{H}\Psi\,,$$

where $\mathcal{H}$ is an infinite-dimensional Hermitian matrix, is difficult to obtain due to complex strong dynamics. However one can simplify the situation by restricting oneself to times that are much larger than a typical strong interaction scale and assuming there is no interaction between final states[4].

We now consider a Hamiltonian

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_w\,, \tag{2.7}$$

such that $|B^0\rangle$, $|\overline{B}^0\rangle$ and $|f_n\rangle$ are eigenstates of $\mathcal{H}_0$; i.e.,

$$\mathcal{H}_0|B^0\rangle = m_{B^0}|B^0\rangle\,, \quad \mathcal{H}_0|\overline{B}^0\rangle = m_{\overline{B}^0}|\overline{B}^0\rangle\,, \quad \text{and} \quad \mathcal{H}_0|f_n\rangle = E_n|f_n\rangle\,,$$

and $\mathcal{H}_w$ is a perturbative interaction that induces $|B^0\rangle$ or $|\overline{B}^0\rangle \to |f_n\rangle$, a $|\Delta B| = 1$ transition, and/or $|B^0\rangle \leftrightarrow |\overline{B}^0\rangle$, a $|\Delta B| = 2$ transition. Considering an initial state

[4]This is called the Weisskopf-Wigner approximation [23].

only containing $|B^0\rangle$ and $|\overline{B}^0\rangle$, and a time scale much longer than the reciprocal of the energy difference between the initial state and any intermediate states, one can obtain the matrix element of $\mathcal{H}_w$ between initial state $|i\rangle$ and final state $|f\rangle$ using the second-order time-dependent perturbation calculation,

$$\begin{aligned}(\mathcal{H}_w)_{fi} = \langle f|\mathcal{H}_w|i\rangle + \mathcal{P}\sum_n \frac{\langle f|\mathcal{H}_w|f_n\rangle\langle f_n|\mathcal{H}_w|i\rangle}{m_B - E_n} \\ -i\pi\sum_n \langle f|\mathcal{H}_w|f_n\rangle\langle f_n|\mathcal{H}_w|i\rangle \cdot \delta(m_B - E_f)\end{aligned} \tag{2.8}$$

where $m_B = (m_{B^0} - m_{\overline{B}^0})/2$ and $\mathcal{P}$ stands for the principal value. The summation here is in a general sense. If the final or intermediate states have continuous spectra, the summation should read an integral.

One can then write down a time-dependent equation on the B^0-$\overline{B}^0$ subspace,

$$i\hbar\frac{\partial}{\partial t}\psi(t) = \mathbf{H}\psi(t)\,, \qquad \psi(t) = a(t)|B^0\rangle + b(t)|\overline{B}^0\rangle = \begin{pmatrix} a(t) \\ b(t)\end{pmatrix}\,. \tag{2.9}$$

The 2×2 matrix $\mathbf{H}$ can be expressed in terms of two Hermitian matrices $\mathbf{M}$ and $\mathbf{\Gamma}$,

$$\mathbf{H} = \mathbf{M} - \frac{i}{2}\mathbf{\Gamma} = \begin{pmatrix} M_{11} - \frac{i}{2}\Gamma_{11} & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{12}^* - \frac{i}{2}\Gamma_{12}^* & M_{22} - \frac{i}{2}\Gamma_{22}\end{pmatrix}\,. \tag{2.10}$$

$\mathbf{H}$ is generally not Hermitian since it only describes a subspace of the infinite-dimensional basis space.

Comparing Eq. 2.10 and Eq. 2.8 with $|i\rangle = |B^0\rangle$ and $|f\rangle = |\overline{B}^0\rangle$, one can obtain

$$M_{11} = m_{B^0} + \langle B^0|\mathcal{H}_w|B^0\rangle + \mathcal{P}\sum_n \frac{|\langle f_n|\mathcal{H}_w|B^0\rangle|^2}{m_B - E_n} \tag{2.11}$$

$$M_{22} = m_{\overline{B}^0} + \langle \overline{B}^0|\mathcal{H}_w|\overline{B}^0\rangle + \mathcal{P}\sum_n \frac{|\langle f_n|\mathcal{H}_w|\overline{B}^0\rangle|^2}{m_B - E_n} \tag{2.12}$$

$$M_{12} = M_{21}^* = \langle B^0|\mathcal{H}_w|\overline{B}^0\rangle + \mathcal{P}\sum_n \frac{\langle B^0|\mathcal{H}_w|f_n\rangle\langle f_n|\mathcal{H}_w|\overline{B}^0\rangle}{m_B - E_n} \tag{2.13}$$

$$\Gamma_{11} = 2\pi\sum_n |\langle f_n|\mathcal{H}_w|B^0\rangle|^2 \cdot \delta(m_B - E_f) \tag{2.14}$$

$$\Gamma_{22} = 2\pi\sum_n |\langle f_n|\mathcal{H}_w|\overline{B}^0\rangle|^2 \cdot \delta(m_B - E_f) \tag{2.15}$$

$$\Gamma_{12} = \Gamma_{21}^* = 2\pi\sum_n \langle B^0|\mathcal{H}_w|f_n\rangle\langle f_n|\mathcal{H}_w|\overline{B}^0\rangle \cdot \delta(m_B - E_f) \tag{2.16}$$

The matrix elements of $\mathcal{H}_w$ in the B^0, $\overline{B}^0$ subspace; i.e., $\langle B^0 \text{ or } \overline{B}^0|\ \mathcal{H}_w\ |\overline{B}^0 \text{ or } B^0\rangle$ represent the direct "superweak" transitions [24], and are not expected in the Standard Model.

If the CPT symmetry is conserved, which is the consequence of the Lorentz symmetry of local quantum theories [21], the diagonal elements of the 2×2 matrix in Eq. 2.10 are identical; i.e., $M_{11} = M_{22}$ (that is, $m_{B^0} = m_{\overline{B}^0}$) and $\Gamma_{11} = \Gamma_{22}$. The eigenstates of Eq. 2.9 can be written as

$$|B_\pm\rangle = p|B^0\rangle \pm q|\overline{B}^0\rangle\,, \tag{2.17}$$

with $|p|^2 + |q|^2 = 1$. The eigenvalues $\mu_\pm$ and p, q can be solved for,

$$\begin{aligned}\omega_\pm &= M - \frac{i}{2}\Gamma \pm \sqrt{H_{12}H_{21}} \\ &= M - \frac{i}{2}\Gamma \pm \sqrt{\left(M_{12} - \frac{i}{2}\Gamma_{12}\right)\left(M_{12}^* - \frac{i}{2}\Gamma_{12}^*\right)}\end{aligned} \tag{2.18}$$

where $M = M_{11} = M_{22}$ and $\Gamma = \Gamma_{11} = \Gamma_{22}$, and

$$\frac{q}{p} = \left(\frac{H_{21}}{H_{12}}\right)^{1/2} = \left(\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}\right)^{1/2} . \tag{2.19}$$

Since the eigenvalues show up in the exponent in a time-dependent quantum state, $e^{-i\omega t}$, the real part of the eigenvalue is the mass term and the imaginary part is half the negative decay rate. Therefore one can write

$$\begin{aligned} M_\pm &= \Re(\omega_\pm) = M \pm \Re[(H_{12}H_{21})^{1/2}] , \\ \Gamma_\pm &= -2 \cdot \Im(\omega_\pm) = \Gamma \mp 2 \cdot \Im[(H_{12}H_{21})^{1/2}] . \end{aligned} \tag{2.20}$$

Since one mass eigenstate is heavier than the other, we can call $|B_+\rangle$ the *B-heavy* ($|B_H\rangle$) and $|B_-\rangle$ the *B-light* ($|B_L\rangle$), and define $\Delta M \equiv M_H - M_L = M_+ - M_-$ and $\Delta\Gamma \equiv \Gamma_H - \Gamma_L = \Gamma_+ - \Gamma_-$. ΔM and $\Delta\Gamma$ satisfy

$$\begin{aligned} \Delta M^2 - \frac{1}{4}\Delta\Gamma^2 &= 4\left(|M_{12}|^2 - \frac{1}{4}|\Gamma_{12}|^2\right) \\ \Delta M \Delta\Gamma &= 4\Re(M_{12}\Gamma_{12}^*) . \end{aligned}$$

Eq. 2.19 can be expressed as

$$\frac{q}{p} = \frac{\Delta M - \frac{i}{2}\Delta\Gamma}{2\left(M_{12} - \frac{i}{2}\Gamma_{12}\right)} = \frac{2\left(M_{12}^* - \frac{i}{2}\Gamma_{12}^*\right)}{\Delta M - \frac{i}{2}\Delta\Gamma} .$$

As shown in Eq. 2.11–2.16, the intermediate states that contribute to the matrix elements of $\mathbf{M}$ are virtual; i.e., their energy can be different from the B meson mass, while those contribute to $\mathbf{\Gamma}$ are real (on-shell); i.e., they have the same energies as the B meson mass and thus represent real decays. The major contribution to the matrix element M_{12} is given by the box diagrams (Fig. 1.1) that contain top quarks in the loop, due to the large mass of the top quark, and the result shows that $M_{12} \propto m_t^2$.

More details will be given in the following paragraphs. On the other hand, the box diagrams can also provide a good approximation for Γ_{12}, but the internal quarks can only be up and charm quarks because the top quark is too heavy so that it cannot contribute to on-shell intermediate states. Because the charm and up quarks are considerably lighter than the bottom quark, the mass of the latter sets the scale, and it leads to $\Gamma_{12} \propto m_b^2$. Therefore,

$$\left|\frac{\Gamma_{12}}{M_{12}}\right| = \mathcal{O}\left(\frac{m_b^2}{m_t^2}\right) \ll 1\,. \tag{2.21}$$

With this approximation, the mass and decay rate differences of two mass eigenstates and q/p become simply

$$\Delta M = 2|M_{12}|\,, \qquad \Delta\Gamma = 2\Re(M_{12}\Gamma_{12}^*)/|M_{12}|\,. \tag{2.22}$$

and

$$\frac{q}{p} = \frac{M_{12}^*}{|M_{12}|}\left(1 - \frac{1}{2}\Im\left(\frac{\Gamma_{12}}{M_{12}}\right)\right) = \frac{M_{12}^*}{|M_{12}|}\left(1 + \mathcal{O}\left(10^{-3}\right)\right)\,. \tag{2.23}$$

From Eq. 2.21 and 2.22, we can see that $\Delta\Gamma \ll \Delta M$. Currently $\Delta\Gamma$ has not been measured. On the other hand, ΔM for B^0 has been measured quite precisely [6],

$$x_d \equiv \Delta m_d/\Gamma_d \simeq 0.77\,.$$

This also means $\Delta\Gamma \ll \Gamma$.

2.3.2 Neutral B decay time evolution

Any neutral B states can be written as a superposition of two mass eigenstates,

$$|B_H(t)\rangle = e^{-\Gamma_H t/2} e^{-iM_H t} |B_H(0)\rangle \tag{2.24}$$

$$|B_L(t)\rangle = e^{-\Gamma_L t/2} e^{-iM_L t} |B_L(0)\rangle \,. \tag{2.25}$$

The state that is purely $|B^0\rangle$ at $t = 0$, denoted as $|B^0_{\rm phys}(t)\rangle$, evolves into a mixed state of $|B^0\rangle$ and $|\overline{B}^0\rangle$,

$$|B^0_{\rm phys}(t)\rangle = \frac{1}{2p}(|B_L(t)\rangle + |B_H(t)\rangle) = g_+(t)|B^0\rangle + (q/p)g_-(t)|\overline{B}^0\rangle \,. \tag{2.26}$$

Likewise

$$|\overline{B}^0_{\rm phys}(t)\rangle = (p/q)g_-(t)|B^0\rangle + g_+(t)|\overline{B}^0\rangle \,. \tag{2.27}$$

The time dependent functions $g_\pm(t)$ are given by

$$g_+(t) = e^{-\Gamma t/2} e^{-iMt} \cos\left[\left(\Delta M - i\frac{\Delta\Gamma}{2}\right) t/2\right] \tag{2.28}$$

$$g_-(t) = e^{-\Gamma t/2} e^{-iMt} i \sin\left[\left(\Delta M - i\frac{\Delta\Gamma}{2}\right) t/2\right] \,, \tag{2.29}$$

where $M = (M_H + M_L)/2$, $\Gamma = (\Gamma_H + \Gamma_L)/2$, and $\Delta M = M_H - M_L$, $\Delta\Gamma = \Gamma_H - \Gamma_L$.

In order to study the decay probability, the amplitudes of a B^0 or $\overline{B}^0$ decaying to an arbitrary final state f are introduced and denoted as $A_f \equiv \langle f|H|B^0\rangle$, and $\overline{A}_f \equiv \langle f|H|\overline{B}^0\rangle$. The decay rate of a B^0 or $\overline{B}^0$ created at $t = 0$ to the final state f at time t is then

$$|\langle f|H|B^0_{\rm phys}(t)\rangle|^2 = |g_+(t)A_f + (q/p)g_-(t)\overline{A}_f|^2 \tag{2.30}$$

$$|\langle f|H|\overline{B}^0_{\rm phys}(t)\rangle|^2 = |(p/q)g_-(t)A_f + g_+(t)\overline{A}_f|^2 \tag{2.31}$$

In reality, the branching ratio of a B^0 decaying to a final state that a $\overline{B}^0$ can also decay into is very small ($\mathcal{O}(10^{-3})$). So to a good approximation, the sum of branching ratios of those final states with $\overline{A}_f \neq 0$ and $A_f = 0$ is unity. Therefore, the probability of observing a state that was created as a B^0 decaying to a final state that can only be originated from $\overline{B}^0$ is

$$\begin{aligned}\sum_{f,\, A_f=0} |\langle f|H|B^0_{\rm phys}(t)\rangle|^2 &= \left|\frac{q}{p} g_-(t)\right|^2 \\ &= \frac{1}{2}\left|\frac{q}{p}\right|^2 e^{-\Gamma t}\left[\cosh(\Delta\Gamma t/2) - \cos(\Delta M t)\right] .\end{aligned} \tag{2.32}$$

We call this a *mixed* event as it decays as a different flavor of b from what was created. Similarly the probability of $|B^0_{\rm phys}(t)\rangle$ decaying to a final state that can only be originated from B^0 is

$$\begin{aligned}\sum_{f,\, \overline{A}_f=0} |\langle f|H|B^0_{\rm phys}(t)\rangle|^2 &= |g_+(t)|^2 \\ &= \frac{1}{2} e^{-\Gamma t}\left[\cosh(\Delta\Gamma t/2) + \cos(\Delta M t)\right] .\end{aligned} \tag{2.33}$$

We call this an *unmixed* event. Exchanging p and q in Eq. 2.32, one gets the probability of observing a mixed event from $|\overline{B}^0_{\rm phys}(t)\rangle$. The unmixed probability for $|\overline{B}^0_{\rm phys}(t)\rangle$ is the same as $|B^0_{\rm phys}(t)\rangle$.

If CP is conserved, the mass eigenstates must also be CP eigenstates. We can write down the CP transformation on $|B^0\rangle$ and $|\overline{B}^0\rangle$ with a certain phase convention,

$$CP|B^0\rangle = e^{-i\xi}|\overline{B}^0\rangle\,, \qquad CP|\overline{B}^0\rangle = e^{+i\xi}|B^0\rangle\,.$$

If we let $CP|B_H\rangle = \eta_H|B_H\rangle$ and $CP|B_L\rangle = \eta_L|B_L\rangle$, we find

$$\frac{q}{p} = \frac{p}{q}e^{-2i\xi}\,,$$

which means q/p is a pure phase ($|q/p| = 1$). Therefore if $|q/p| \neq 1$, we have CP violation. This type of CP violation is called "CP violation in mixing", which results from the mass eigenstates being different from the CP eigenstates. In the neutral K system, this type of CP violation has been observed unambiguously [18]. CP violation in mixing can be tested with semileptonic decays by examining the asymmetry

$$\frac{\Gamma(|\overline{B}^0_{\rm phys}(t)\rangle \to \ell^+\nu_\ell X) - \Gamma(|B^0_{\rm phys}(t)\rangle \to \ell^-\overline{\nu}_\ell X)}{\Gamma(|\overline{B}^0_{\rm phys}(t)\rangle \to \ell^+\nu_\ell X) + \Gamma(|B^0_{\rm phys}(t)\rangle \to \ell^-\overline{\nu}_\ell X)} = \frac{1-|q/p|^4}{1+|q/p|^4}\,. \tag{2.34}$$

This equation basically probes the asymmetry between $|B^0\rangle \to |\overline{B}^0\rangle \to \overline{f}$ and $|\overline{B}^0\rangle \to |B^0\rangle \to f$.

In the analysis described in this thesis, $\Delta\Gamma$ is assumed to be negligible and CP violation in mixing is ignored. After proper normalization, the probability of observing a unmixed ("+" sign) or mixed ("−" sign) event from a state that was created as a pure $|B^0\rangle$ or $|\overline{B}^0\rangle$ is

$$P_\pm(t) = \frac{e^{-\Gamma|\Delta t|}}{4\tau}\left[1 \pm \cos(\Delta M \Delta t)\right] \tag{2.35}$$

where $\tau = 1/\Gamma$ is the average lifetime, and the decay time t is replaced by the decay time difference Δt, which is a signed quantity. An extra factor of $1/2$ is due to the fact that the domain of Δt is now $(-\infty,+\infty)$, instead of $(0,+\infty)$.

The time integrated mixing probability can be obtained by comparing the integrals of Eq. 2.32 and 2.33,

$$\begin{aligned}\chi^{B^0\to\overline{B}^0} &= \frac{\int_0^\infty dt|\langle\overline{B}^0|B^0_{\rm phys}(t)\rangle|^2}{\int_0^\infty dt|\langle\overline{B}^0|B^0_{\rm phys}(t)\rangle|^2+\int_0^\infty dt|\langle B^0|B^0_{\rm phys}(t)\rangle|^2} \\ &= \frac{|q/p|^2(x^2+y^2)}{|q/p|^2(x^2+y^2)+2+x^2-y^2}\,,\end{aligned} \tag{2.36}$$

where $x=\Delta M/\Gamma$ and $y=\Delta\Gamma/(2\Gamma)$. Again, exchanging q and p, one can get $\chi^{\overline{B}^0\to B^0}$.

2.3.3 Standard Model calculation of ΔM

As mentioned before, the matrix element M_{12} can be calculated from box diagrams. The amplitude is roughly proportional to the masses of the two quarks in the loop. Therefore, the contributions of the diagrams with one or two quark lines being up or charm quark are negligible compared to the one with two top quarks.

With a theoretical tool called *operator product expansion* (OPE) [25], one can develop a low-energy effective Hamiltonian $H^{|\Delta B|=2}$ that uses a local $|\Delta B|=2$ four quark operator, $Q=\bar{d}\gamma_\nu(1-\gamma_5)b\bar{d}\gamma_\nu(1-\gamma_5)b$. M_{12} is given by

$$M_{12}=\frac{\langle B^0|H^{|\Delta B|=2}|\overline{B}^0\rangle}{2m_B}\left(1+\mathcal{O}\left(\frac{m_b^2}{M_W}\right)\right)\,. \tag{2.37}$$

The result is of the form

$$H^{|\Delta B|=2}=\left(\frac{G_F}{4\pi}\right)^2(V_{tb}V_{td}^*)^2M_W^2S\left(\frac{m_t^2}{M_W^2}\right)\eta_B b_B(\mu)Q(\mu)+\text{h.c.}\,, \tag{2.38}$$

where the function $S(x)$ is the box diagram calculation,

$$S(x)=x\left(\frac{1}{4}+\frac{9}{4(1-x)}-\frac{3}{2(1-x)^2}\right)-\frac{3}{2}\left(\frac{x}{1-x}\right)^3\ln x\,,$$

and η_B and $b_B(\mu)$ contain the short-distance QCD corrections. μ is the renormalization scale. The matrix element of the local operator Q between two hadronic states $\langle B^0|Q|\overline{B}^0\rangle$ depends on μ as well, and can be parameterized as

$$\langle B^0|Q(\mu)|\overline{B}^0\rangle = \frac{8}{3}B_B(\mu)f_B^2 m_B^2 = \frac{8}{3}\frac{\hat{B}_B}{b_B(\mu)}f_B^2 m_B^2\,, \tag{2.39}$$

where f_B is the decay constant and $\hat{B}_B$ the bag constant. The physical quantity should not depend on the unphysical parameter μ. Therefore, $b_B(\mu)$ should be canceled in Eq. 2.38, and $\hat{B}_B$ does not depend on μ. However, in practice, such calculations involves approximations, so the dependence still exists, and therefore a "reasonable" range of μ is considered to estimate the theoretical uncertainties. Combining Eq. 2.39, 2.38 and 2.37, one can obtain the mass difference of two mass eigenstates

$$\Delta m_d = 2|M_{12}| = \frac{G_F^2}{6\pi^2}\eta_B m_B \hat{B}_B f_B^2 M_W^2 S\left(\frac{m_t^2}{M_W^2}\right)|V_{tb}V_{td}^*|^2\,. \tag{2.40}$$

The measurement of Δm_d can constrain the length of one side of the unitary triangle shown in Fig. 2.1, from (0,1) to $(\overline{\rho},\overline{\eta})$, i.e., $|V_{td}V_{tb}^*|/|V_{cd}V_{cb}^*|$. Among these CKM matrix elements, $|V_{cd}|$, $|V_{cb}|$ and $|V_{tb}|$ are much better known than $|V_{td}|$. Therefore Δm_d can in principle provide a constraint on $|V_{td}|$. However, currently the best estimates of f_B and $\hat{B}_B$, which are from lattice QCD calculations [26, 27], have much larger uncertainties ($\sigma(f_B\sqrt{\hat{B}_B}) \sim 15\%$) than the experimental uncertainty on Δm_d (2%).

The same formalism can be applies to B_s mixing. The only difference is that B_s contains an s quark instead of a d quark. The mass difference Δm_s takes the same formula as Eq. 2.40, with $|V_{td}|^2 m_B \hat{B}_B f_B^2$ replaced by $|V_{ts}|^2 m_{B_s}\hat{B}_{B_s} f_{B_s}^2$. The lattice

calculation of the ratio

$$\xi \equiv \left(f_{B_s} \sqrt{\hat{B}_{B_s}} \right) / \left(f_{B_d} \sqrt{\hat{B}_{B_d}} \right) \tag{2.41}$$

is more precise than the individual quantities ($\sigma(\xi) \simeq 4$–8% [26, 27, 28]). Therefore, the ratio of Δm_s to Δm_d,

$$\frac{\Delta m_s}{\Delta m_d} = \frac{m_{B_s}}{m_{B_d}} \xi^2 \frac{|V_{ts}|}{|V_{td}|} = \frac{1}{\lambda} \frac{m_{B_s}}{m_{B_d}} \xi^2 \frac{1}{(1-\overline{\rho})^2 + \overline{\eta}^2} \,, \tag{2.42}$$

can provide a much better constraint on $|V_{td}|$. Presently only the lower limit of Δm_s is available [6].

2.3.4 Coherent $B\overline{B}$ state

At the PEP-II e^+e^- collider, the B^0 and $\overline{B}^0$ mesons are produced in pairs from $\Upsilon(4S)$ decays. A general state consisting of two neutral B mesons can be expressed in terms of products of mass eigenstates,

$$|BB_{\rm phys}(t_1, t_2)\rangle = c_1 e^{-i(\omega_+ t_1 + \omega_- t_2)} |B_+^{(1)} B_-^{(2)}\rangle + c_2 e^{-i(\omega_- t_1 + \omega_+ t_2)} |B_-^{(1)} B_+^{(2)}\rangle \,,$$

where t_1 and t_2 are the proper time of $|B^{(1)}\rangle$ and $|B^{(2)}\rangle$, respectively. When spin-0 B mesons are produced from a spin-1 $\Upsilon(4S)$ decay, they must be in an antisymmetric p-wave state ($L = 1$). If they were to decay *simultaneously* ($t_1 = t_2$) into the same bosonic final state f, there would be two identical bosonic states in an antisymmetric state. It is forbidden according to Bose statistics [29]. Therefore $\langle ff|BB_{\rm phys}(t,t)\rangle = 0$ in this case, which requires $c_1 = -c_2$, that is,

$$\begin{aligned} |BB_{L=1}(t_1, t_2)\rangle &= e^{-i(\omega_+ t_1 + \omega_- t_2)} |B_+^{(1)} B_-^{(2)}\rangle - e^{-i(\omega_- t_1 + \omega_+ t_2)} |B_-^{(1)} B_+^{(2)}\rangle \\ &= e^{-i2\omega t} \left[e^{+i\Delta\omega\Delta t} |B_+^{(1)} B_-^{(2)}\rangle - e^{-i\Delta\omega\Delta t} |B_-^{(1)} B_+^{(2)}\rangle \right] \,, \end{aligned} \tag{2.43}$$

in which the variables in the second line have been changed to

$$t \equiv \frac{t_1 + t_2}{2}\,, \qquad \Delta t \equiv t_2 - t_1\,,$$

and

$$\omega \equiv \frac{\omega_+ + \omega_-}{2} = M - \frac{i}{2}\Gamma\,, \qquad \Delta\omega \equiv \frac{\omega_+ - \omega_-}{2} = \frac{1}{2}\left(\Delta M - \frac{i}{2}\Delta\Gamma\right)\,. \tag{2.44}$$

If one B decays to a final state f_1 at time t_1, the resulting partially projected state is

$$\langle f_1|BB_{L=1}(t_1,t_2)\rangle = e^{-i2\omega t}\left[e^{+i\Delta\omega\Delta t}\langle f_1|B_+^{(1)}\rangle\cdot|B_-^{(2)}\rangle - e^{-i\Delta\omega\Delta t}\langle f_1|B_-^{(1)}\rangle\cdot|B_+^{(2)}\rangle\right]\,, \tag{2.45}$$

We define $A_1 \equiv \langle f_1|B^0\rangle$ and $\overline{A}_1 \equiv \langle f_1|\overline{B}^0\rangle$, then

$$\langle f_1|B_\pm\rangle = pA_1 \pm q\overline{A}_1\,.$$

Eq. 2.45 can be written in terms of A and $\overline{A}$,

$$\begin{aligned}\langle f_1|BB_{L=1}(t_1,t_2)\rangle = e^{-i2\omega t}\Big\{ & q\left[e^{-i\Delta\omega\Delta t}|B_+^{(2)}\rangle + e^{+i\Delta\omega\Delta t}|B_-^{(2)}\rangle\right]\overline{A}_1 \\ & - p\left[e^{-i\Delta\omega\Delta t}|B_+^{(2)}\rangle - e^{+i\Delta\omega\Delta t}|B_-^{(2)}\rangle\right]A_1\Big\}\,,\end{aligned} \tag{2.46}$$

Using Eq. 2.26–2.29, Eq. 2.44, and $t = t_1 + \Delta t/2$, one can get

$$\langle f_1|BB_{L=1}(t_1,\Delta t)\rangle = 2e^{-i2\omega t_1}pq\left[|B^0_{\rm phys}(\Delta t)\rangle\cdot\overline{A}_1 - |\overline{B}^0_{\rm phys}(\Delta t)\rangle\cdot A_1\right] \tag{2.47}$$

If the state f_1 can only be originated from B^0 not $\overline{B}^0$, then f_1 can "tag" the flavor of $|B^{(1)}\rangle$ as a $\overline{B}^0$, and the amplitude $\overline{A}_1 = 0$. Therefore Eq. 2.47 indicates that the state at $\Delta t = 0$ is a pure $\overline{B}^0$. This describes the picture that when a B in this $L = 1$

coherent state decays to a flavor eigenstate, the wave function "collapses" to a B with the opposite flavor.

2.4 Lifetime

2.4.1 Naive expectation

The lifetime of b hadrons, of order picosecond was much longer than anticipated when it was first measured in 1983 [30, 31]. The b hadron decay width is dominated by the total width of a b quark. Naively, one can scale the expression of the total width of a muon decay to estimate the b quark decay width, which is dominated by the $b \to c$ transition,

$$\Gamma_\mu = \frac{G_F}{192\pi^3} m_\mu^5 \quad \Longrightarrow \quad \Gamma_b \sim \frac{G_F}{192\pi^3} m_b^5 |V_{cb}|^2 \times (2 \times 3 + 3)\,, \tag{2.48}$$

where the factor of 2×3 comes from two hadronic channels at quark level (the virtual $W^{-*} \to \bar{c}\, s$ and $\bar{u}\, d$), each with three colors, and the additional factor of 3 from three leptonic channels (the virtual $W^{-*} \to e^-\, \overline{\nu}_e$, $\mu^-\, \overline{\nu}_\mu$, and $\tau^-\, \overline{\nu}_\tau$). All these particles in the final states are assumed to be much lighter than m_b for this estimation. The b quark lifetime is then

$$\tau_b \sim \frac{\tau_\mu (m_\mu/m_b)^5}{9|V_{cb}|^2} \simeq \frac{10^{-15}}{|V_{cb}|^2}\ \mathrm{s}\,.$$

This means that $|V_{cb}|$ is about $1/30$, which is of the order of $\sin^2\theta_C$. $|V_{cb}|$ is more suppressed than the $\sin\theta_C$ suppression expected for inter-family transitions in the early 1980's.

2.4.2 OPE and spectator model

The total width of a hadron H_b containing a b quark can be written in terms of the imaginary part of the forward matrix element of the transition operator by the optical theorem,

$$\Gamma(H_b \to X) = \frac{1}{m_{H_b}} \Im \left[\langle H_b | \hat{T} | H_b \rangle \right] ,$$

where the transition operator $\hat{T}$ is given by the time-ordered product T of the relevant effective weak Hamiltonian $\mathcal{H}_{\text{eff}}$,

$$\hat{T} = i \int d^4x T \left[\mathcal{H}_{\text{eff}}(x) \mathcal{H}_{\text{eff}}(0) \right] . \tag{2.49}$$

The dominant space-time separation x is related to the inverse of the energy release, which is close to the mass of the b quark. One can construct an operator product expansion (OPE) [25] for the bilocal transition operator (Eq. 2.49) and expand it as a series of local operators, whose coefficients contain inverse powers of the b quark mass. The result of the inclusive decay rate can be written [32, 33, 34],

$$\Gamma(H_b \to X) = |V_{\text{CKM}}|^2 \frac{G_F^2 m_b^5}{192\pi^3} \left\{ c_3 \left(1 - \frac{\mu_\pi(H_b)}{2m_b^2} \right) + c_5 \frac{\mu_G(H_b)}{2m_b^2} + \mathcal{O}\left(\frac{1}{m_b^3} \right) \right\} , \tag{2.50}$$

where $\mu_\pi(H_b)$ and $\mu_G(H_b)$ parameterize the matrix elements of the kinetic-energy and the chromo-magnetic operators, respectively.

The leading corrections of order $1/m_b^3$ are called *spectator effects* [5], which is the effect of a light constituent quark participating in the weak process. This effect contributes directly to the differences in the lifetimes of different b-hadrons. It is suppressed by $1/m_b^3$ because the light quark has to be very close to the b quark, and the b quark inside a hadron only occupies a volume of $\mathcal{O}\left((\Lambda_{\text{QCD}}/m_b)^3\right)$.

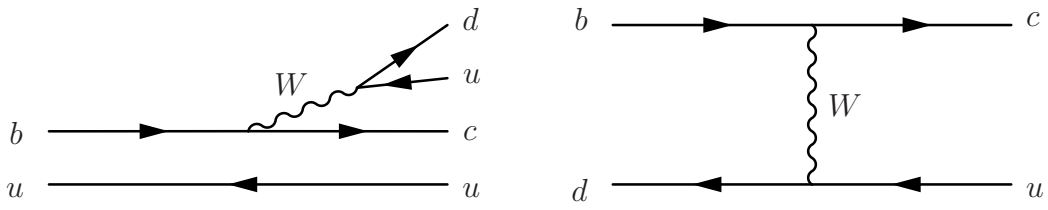

Figure 2.3: Diagrams for spectator effect. Left, Pauli-interference diagram (for B^+). Right, W-exchange diagram (for B^0).

The spectator effects arise from diagrams shown in Fig. 2.3, referred to as Pauli-interference and W-exchange diagrams. They contribute differently for B^0 and B^+ mesons. For the Pauli-interference diagram, the two u quarks in the final state interfere with each other, and thus have a contribution in addition to the simple tree diagram.

Current experimental result shows that the lifetime of Λ_b is shorter that that of B^0 by about 20%, which *cannot* be accounted for by theoretical calculations. Other lifetime ratios, such as $\tau(B^+)/\tau(B^0)$ and $\tau(B_s)/\tau(B^0)$ are consistent with the calculations, however the experiment precisions are not high enough to probe in detail the model dependent calculations.

2.5 B Semileptonic Decays

The signal for this analysis is defined as $B^0\overline{B}^0$ events in which at least one B decays into a final state satisfying the following criteria: The final state contains a $D^{*\pm}$ and either an electron or a muon of the opposite charge produced by $W^{\mp}$ decay; the D^* and lepton originate from a common vertex, within the detector resolution; and those decays that have the same time-dependent decay structure as the mode $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_\ell$.

With this definition, the D^* and lepton must either be the direct decay products

of the parent B

$$\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_\ell\ (X)\ ,$$

or else be the secondary decay products of a short-lived intermediate resonance. The only types of resonances expected to contribute are the radially and orbitally excited D states ($Y_c = {D_1}^+, {D_1'}^+, {D_2^*}^+$):

$$\overline{B}^0 \to Y_c\ \ell^-\overline{\nu}_\ell\ (X) \quad , \quad Y_c \to D^{*+}\ X'\ .$$

Table 2.2 summarizes the signal branching ratios which are used in the Monte Carlo simulations, and defines the signal categories used in this thesis.

Table 2.2: Signal decay modes and their branching fractions in the default generic Monte Carlo. The branching ratios for the excited D resonances (${D_1}^+$, ${D_1'}^+$, ${D_2^*}^+$) include their branching ratio into $D^{*+}\pi^0$. The branching ratio for radiative decays $B^0 \to D^{*+}e^-\nu_e\ \gamma(\gamma\ldots)$ is not an explicit model parameter but depends on cutoff parameters in the PHOTOS [35] decay model. The fractions are relative to the total signal branching ratio of 10.88%.

Type	Decay Mode	Branching Ratio	Fraction
1	$\overline{B}^0 \to D^{*+}e^-\nu_e$	4.04%	37%
2	$\overline{B}^0 \to D^{*+}\mu^-\nu_\mu$	4.90%	45%
3	$\overline{B}^0 \to D^{*+}e^-\nu_e\ \gamma(\gamma\ldots)$	0.860%	7.9%
4	$\overline{B}^0 \to D^{*+}\ell^-\nu_\ell\ \pi^0\ (\gamma\ldots)$	0.200%	1.8%
5	$\overline{B}^0 \to {D_1'}^+\ell^-\nu_\ell\ (\gamma\ldots)$	0.467%	4.3%
6	$\overline{B}^0 \to {D_1}^+\ell^-\nu_\ell\ (\gamma\ldots)$	0.267%	2.5%
7	$\overline{B}^0 \to {D_2^*}^+\ell^-\nu_\ell\ (\gamma\ldots)$	0.144%	1.3%
	TOTAL	10.88%	100%

The phase space of semileptonic decays can be parametrized as

$$d\Phi\left(\overline{B}^0 \to X^+\ell^-\overline{\nu}_\ell\right) \propto d\hat{E}_\ell\, d\hat{Q}^2\ ,$$

where X represents any hadronic state, $\hat{E}_\ell$ is the lepton energy in the B rest frame and $\hat{Q}^2$ the $W^\pm$ invariant mass, calculated as $m^2_{B^0} + m^2_X - 2m_{B^0}\hat{E}_X$, where $\hat{E}_X$ is the hadron energy in the B rest frame. The variables $\hat{E}_X$ and $\hat{Q}^2$ are useful to disentangle the effects of phase space and dynamics. Figure 2.4 compares the distribution of these variables ("Dalitz plot") for three-body phase space and for $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_\ell$ decays with a heavy quark effective theory [36] model of form factors [5]

The magnitude of the hadron momentum, $\hat{p}_X$, and the angle between the direction of the hadron and lepton momenta, $\hat{\theta}_{X,\ell}$, (both defined in the B rest frame) are functions of the Dalitz variables,

$$\begin{aligned} \hat{p}_X &= \frac{1}{2m_B}\sqrt{\left(m_B^2 - (\hat{Q}+m_X)^2\right)\left(m_B^2 - (\hat{Q}-m_X)^2\right)} \\ \cos\hat{\theta}_{X,\ell} &= \frac{\hat{Q}^2 + m_\ell^2 - \hat{E}_\ell(m_B^2 + m_X^2 - \hat{Q}^2)/m_B}{2\hat{p}_X\hat{p}_\ell}\,, \end{aligned}$$

and so selection criteria applied to more than two of the following variables are not independent: $\hat{Q}^2$, $\hat{E}_\ell$, $\hat{p}_X$, and $\cos\hat{\theta}_{X,\ell}$. Figure 2.5 shows the distribution of $\hat{p}_X$ and $\cos\hat{\theta}_{X,\ell}$ for $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_\ell$ decays with known form factors and generic three-body decays whose probability is uniform on the phase space. Figure 2.5(a) shows that the hadron momentum has a sharp edge at the upper kinematic limit, although the spectrum is significantly softened by the form factors and helicity correlations. Figure 2.5(b) shows that the D^* and lepton tend to be back-to-back in the B decay frame, and that this effect is mostly due to the phase space constraints.

All kinematic quantities discussed in this section so far are defined in the B rest frame. Since the mass of Υ(4S) state is just above the threshold of two open B mesons, the B momentum in the Υ(4S) rest frame is very small, $\sim$ 340MeV. To a good approximation, the B rest frame and the Υ(4S) rest frame are equivalent. Since in general the B flight direction in the Υ(4S) frame is not known, the Υ(4S)

[5] Form factor values: $\rho^2 = 0.92$, $R_1 = 1.18$, $R_2 = 0.72$.

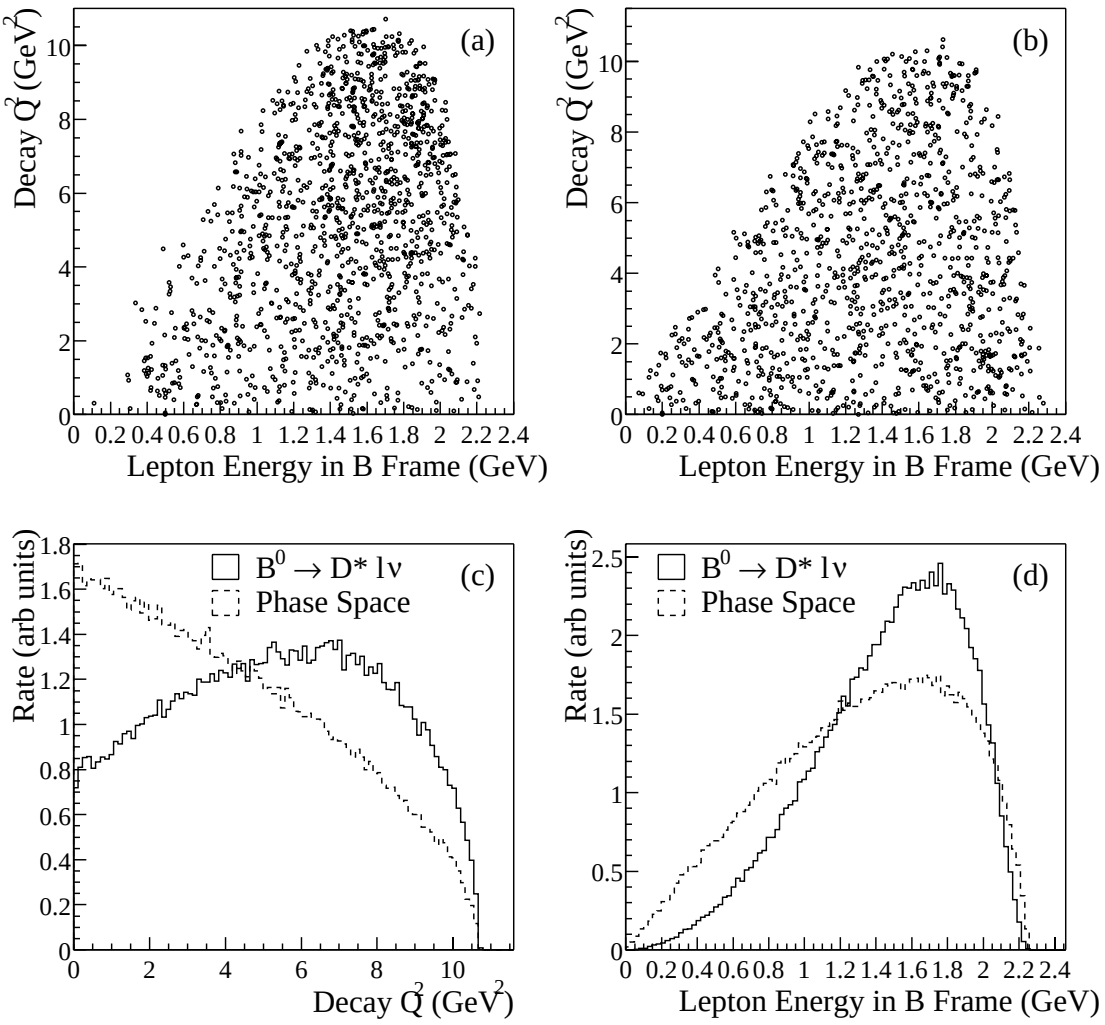

Figure 2.4: Dalitz distribution of $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_\ell$ using (a) the form factors and helicity correlations, or (b) three-body phase space. The lower plots show the normalized projections of these variables: (c) $\hat{Q}^2$, and (d) $\hat{E}_\ell$.

frame is often used for calculating variables that should have been evaluated in the B rest frame.

Since the neutrino cannot be detected directly in *BABAR*, its information must be inferred from other kinematic quantities. The mass of the neutrino can be calculated

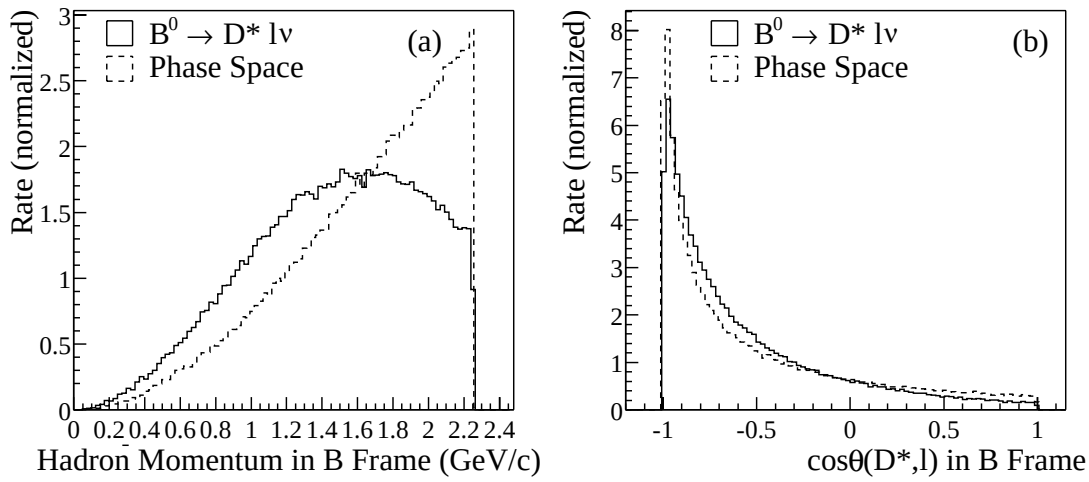

Figure 2.5: Normalized distributions of (a) the hadron momentum, $\hat{p}_X$, and (b) the angle between the hadron and lepton momenta, $\cos\hat{\theta}_{X,\ell}$, measured in the B^0 decay frame. The plots compare the distributions calculated using either three-body phase space, or else including the full decay dynamics.

as

$$\begin{aligned} m_\nu^2 = p_\nu^2 &= (p_{B^0} - p_{X\ell})^2 \\ &= m_{B^0}^2 + m_{X\ell}^2 - 2E_{B^0}E_{X\ell} + 2|\vec{p}_{B^0}||\vec{p}_{X\ell}|\cos\theta_{B^0,X\ell} \ . \end{aligned} \tag{2.51}$$

All quantities on the right hand side of the equal sign in the second line of Eq. 2.51 can be measured or calculated in the Υ(4S) rest frame, except $\cos\theta_{B^0,X\ell}$, which has a range of (−1,1). If the missing mass is calculated as $M_{\rm miss} = m_{B^0}^2 + m_{X\ell}^2 - 2E_{B^0}E_{X\ell}$, its distribution has a peak around zero and an RMS of the order of 1 GeV^2. Apparently, the RMS of $M_{\rm miss}$ distribution depends on the coefficient of $\cos\theta_{B^0,X\ell}$, i.e., $|\vec{p}_{B^0}||\vec{p}_{X\ell}|$. Some analyses (for example [37]) select events with a variable $M_{\rm miss}$ range according to $|\vec{p}_{B^0}||\vec{p}_{X\ell}|$) to increase sensitivity.

In the analysis described in this thesis, an alternative but equivalent approach is used. The expression in Eq. 2.51 is equal to zero if the only missing particle in the

event is a massless neutrino. Therefore the quantity

$$\frac{-m_{B^0}^2 - m_{X\ell}^2 + 2E_{B^0}E_{X\ell}}{2|\vec{p}_{B^0}||\vec{p}_{X\ell}|} = \cos\theta_{B^0,X\ell} \tag{2.52}$$

is in the range (−1,+1) for signal, while for background the distribution can extend beyond this range. Eq. 2.52 does not need the knowledge of B direction in the $\Upsilon(4S)$ rest frame, which can not be measured in the case of semileptonic decays. Only the magnitude of the B momentum is needed, which can be calculated from the center-of-mass energy and the B mass. The events outside the physical range are removed to enhance signal to background ratio.

Chapter 3

Current Experimental Status

In this chapter I summarize the current experimental status and analysis techniques of B^0 lifetime and B^0-$\overline{B}^0$ mixing frequency measurements.

3.1 Mixing Measurements

The phenomenon of particle-antiparitcle mixing in the neutral B meson system was first observed in 1987 by the UA1 [1] and ARGUS [2] collaborations. Both experiments are based on detecting same-sign high (transverse) momentum lepton pairs, which are predominately from B semileptonic decays. The time-dependent B^0 oscillation frequency was first reported by the ALEPH collaboration [38] in 1993. Since then both time-integrated and time-dependent techniques [9] have been used to measure the B^0 $\overline{B}^0$ mixing frequency. These experiments are operated at the $\Upsilon(4S)$ resonance or the Z resonance at e^+e^- colliders, or at higher center-of-mass energies at hadron colliders. In this section, I summarize the techniques and results from major experiments and the current world average value.

3.1.1 Time-integrated measurements

The basic idea of the time-integrated technique is to directly measure the overall mixing probability χ_d, which is the overall probability of a meson that is produced as a B^0 but decays to final states that can only originate from $\overline{B}^0$, or vice versa (see Eq. 2.36), by counting the fraction of mixed events; that is, the fraction of events in which b quarks in both B mesons that are created as opposite flavor decay as the same flavor (one B going through a $|\Delta B| = 2$ process). Although B^0 mixing was discovered in a $p\overline{p}$ collider by the UA1 Collaboration at CERN (where B mesons are created through $q\overline{q} \to g^* \to b\overline{b}$ and $gg \to b\overline{b}$ process), by observing a non-zero χ_d, the time-integrated method is mostly used in symmetric e^+e^- colliders whose center-of-mass (c.m.) energy is equal to the $\Upsilon(4S)$ mass. At these colliders, B meson pairs are created almost at rest; their time evolution is not measurable. Only time-integrated results can be obtained.

Most time-integrated measurements have been made by the ARGUS [39, 40] and CLEO [41, 42] Collaborations. Both use B mesons from $e^+e^- \to \Upsilon(4S) \to B^0\overline{B}^0$ process. One simple method is to measure the fraction of events with same-sign lepton pairs among all dilepton events whose leptons are directly coming from B mesons through semileptonic decays [41];

$$\chi_d = \frac{N(\ell^\pm\ell^\pm)}{N(\ell^\pm\ell^\pm) + N(\ell^\pm\ell^\mp)} \,.$$

The sign of the lepton indicates the sign of the virtual W in the b decay, which directly identifies the flavor of the b quark. Since the B meson pairs from $\Upsilon(4S)$ decays have opposite flavors, a pair of same-sign direct leptons indicates a mixed event. The advantage of this method is that the branching fraction and the reconstruction efficiency are high. The major source of systematic uncertainty is the estimation of the $B^0\overline{B}^0$ and B^+B^- fractions in $\Upsilon(4S)$ decays, since charged B mesons do not mix.

Other sources of systematic uncertainty include $b \to c \to \ell$ contamination, since the lepton from the cascade charm decay has opposite sign to that of the primary lepton.

Assuming $\Delta\Gamma = 0$ and $|q/p| = 1$, the mixing probability χ_d can be expressed as $\chi_d = \frac{1}{2}\frac{x_d^2}{x_d^2+1}$ (see Eq. 2.36), where $x_d = \Delta m_d/\Gamma = \Delta m_d \times \tau_{B^0}$. Obviously, in order to extract Δm_d from χ_d, one needs an independent measurement of τ_{B^0}. However, because the time-integrated mixing probability approaches 1/2 asymptotically as Δm increases, the time-integrated measurement for B_s is not sensitive to Δm_s due to its large value ($x_s \equiv \Delta m_s \times \tau_{B_s} > 19.0$ at 95% CL [6]). The mixing of B_s has not been measured. It is virtually impossible to measure it with time-integrated methods. The best hope lies in the time-dependent measurements.

Several other techniques are used to improve the B^0 purity in dilepton samples. The most commonly used method is to reconstruct the soft pion from $\overline{B}^0 \to D^{*+}\ell^-\nu$, $D^{*+} \to D^0\pi^+$ decay chain [41, 40]. In addition to dilepton events, partially reconstructed hadronic events have also been used. CLEO [42] uses the sign of the π or ρ in partially reconstructed $\overline{B}^0 \to D^{*+} \to \pi^-, \rho^-$, $D^{*+} \to D^0\pi^+$ events (only direct π (ρ) and soft π from D^{*+} are reconstructed) to determine the flavor of the B and a high-momentum lepton to tag the other B. These measurements greatly reduce the systematic error due to charged B events in the data sample.

The most precise time-integrated measurement of χ_d has been made by CLEO [42]. They use 9.6×10^6 $B\overline{B}$ pairs and find $\chi_d = 0.198 \pm 0.013 \pm 0.014$. This measurement dominates the world average of χ_d using the time-integrated technique, which is $\chi_d = 0.182 \pm 0.015$ [6]. This number corresponds to the oscillation frequency $\Delta m_d = (0.491 \pm 0.033)$ ps^{-1} if we assume $y_d = 0$ and use the world average B^0 lifetime $\tau_{B^0} = (1.542 \pm 0.016)$ ps [6] ($\chi_d = \frac{x^2}{2(1+x^2)}$, where $x = \Delta m_d \tau_{B^0}$). However, as will be shown in Sec. 3.1.2, Δm_d is better measured with time-dependent analyses.

3.1.2 Time-dependent measurements

The time-dependent analyses study the oscillatory behavior of B^0 mesons and extract the frequency Δm_d directly. This method is used in experiments operated at the Z mass (LEP [43, 44, 45] and SLD [46]), asymmetric B-factories at the $\Upsilon(4S)$ resonance (*BABAR* and Belle [47]), and higher center-of-mass energy hadronic colliders (Tevatron [48]). The basic ingredients are the proper decay time of the B^0 and the flavors of the B^0 at the times of its creation and decay. The probability of observing a B^0 meson decaying in the originally produced state (unmixed, "+" sign) and in the opposite state (mixed, "−" sign), at time t after its production is proportional to $1 \pm \cos(\Delta m t)$.

The B^0 is either fully or partially reconstructed. Fully (or exclusively) reconstructed analyses usually use hadronic final states or semileptonic decays that contain an open charm meson. Partially reconstructed analyses usually use the lepton from semileptonic decays and/or the soft pion from D^* or D^* itself from $B^0 \to D^* X$, or the topological information such as the detached secondary vertices.

The proper decay time of the B meson at LEP, SLD and Tevatron is measured by $t = \frac{m_B}{p} L$, where L is the distance between the production vertex and the decay vertex, and p is the momentum of the B. At these higher energy experiments, the reconstruction efficiencies are usually much lower than those operated at $\Upsilon(4S)$, therefore the fully reconstructed hadronic final states are rarely used because their branching ratios are an order of magnitude or more smaller than the semileptonic decays. As a result, the momentum of the B cannot be measured solely from the detected final particles from the B. The momentum of the B can be calculated with $p = \sqrt{E_B^2 - m_B^2}$. The energy of the B (E_B) can be measured from the c.m. energy and/or jet energy, and the energies of objects that are not associated with the reconstructed B. For a partial reconstruction analysis, extra care should be taken to reject or reduce the weights of particles that belong to the B final state.

At asymmetric energy B-factories (PEP-II and KEKB [47]), the $\Upsilon(4S)$ decays to only a pair of B mesons. Both B meson move away from the production vertex by a few hundred microns before they decay due to a boost in the beam direction, $\langle\beta\gamma\rangle \simeq 0.55$ for *BABAR* and $\langle\beta\gamma\rangle \simeq 0.43$ for Belle. Therefore, only the decay time difference between the two B mesons is measurable, instead of the total time from production to decay. The decay time difference Δt is calculated with $\Delta t = \Delta z/(\beta\gamma c)$, where Δz is the separation between two decay vertices of the B mesons along the beam direction. Because of the coherence between two B^0 mesons from $\Upsilon(4S)$ decays (see Sec. 2.3), when one B^0 decays to a flavor eigenstate, the other one will have the opposite flavor without a component of the other flavor. So one can start the timer at the time of the first decay, and replace the decay time t with Δt and maintain the formalism for observing an unmixed or mixed event.

The flavors of the B^0 meson at its production and decay are determined through charge correlations between the b quark and the final states. The most commonly used method is to use the charge of the lepton from $b \to \ell$ decays. Alternative methods are to use the charge of charged kaons from $b \to c \to s$ decays or the charged pion from a D^* produced from a B. These are what B-factories use to tag the B that is not being reconstructed. Other methods such as jet-charge [49] and charge dipole [50] are also used. The jet-charge method sums over charges of relevant tracks weighted by a function of their momenta. the charge dipole method finds tracks associated with primary, secondary and tertiary vertices, respectively, and use the fact that the secondary vertex is more likely to be a B decay and the tertiary vertex is more likely to be a charm decay. At a high-energy collider, one can tag the initial flavor of the b quark that the reconstructed B candidate contains (same-side tag), as well as the recoiling b quark (opposite-side tag). On the same side, the charge of a track from the primary vertex is correlated with the production state of the B if that track is a decay product of a B^{**} or the first particle in the fragmentation chain. Jet-charge

works on both side, while the charge dipole method only works on the opposite side.

The statistical significance of an oscillation signal can be approximated as [51]

$$\mathcal{S} \simeq \sqrt{N/2} f_{\rm sig}(1-2\omega) e^{-(\Delta m \sigma_t)^2/2},$$

where N and $f_{\rm sig}$ are the total number of candidates and signal fraction, ω is the total mistag probability, and σ_t is the resolution on proper time or proper time difference. At high-energy colliders, the proper time resolution has two components — one is the resolution of the decay length, typically about 0.05–0.3 ps; the other is the relative momentum resolution σ_p/p, typically about 10–20% for partially reconstructed decays. For B-factories, the decay time difference resolution is dominated by the uncertainty of the vertex location, which is typically 1–1.5 ps because of the much smaller Lorentz boost. The mistag probability at LEP can reach $\sim 26\%$. For SLD it can reach $\sim 22\%$ because the polarized beam produces a sizeable forward-backward asymmetry in the $Z \to b\bar{b}$ decays, which adds more information for b-tagging. At *BABAR* and Belle, effective mistag probabilities of $\sim 24\%$ for full efficiency are reached.

The current published world average of Δm_d is 0.489 ± 0.008 ps^{-1} [6], which is dominated by the measurements at *BABAR* and Belle. The combined precision measured by these two B-factories is ± 0.009 ps^{-1}, while other high-energy experiments contribute to the world average with a combined uncertainty of about ± 0.016 ps^{-1}.

3.2 Lifetime Measurements

The first measurement of the b hadron lifetime was by MAC [30, 31] in 1983. Since then, like time-dependent mixing measurements, experiments at LEP [52, 53], SLD [54], Tevatron and B-factories [55, 56] have significantly improved the precision.

The techniques used for lifetime measurements are very similar to those used in time-dependent mixing measurements, except the flavor tagging of the B meson at its production is not necessary.

The current published world average of τ_{B^0} is 1.542 ± 0.016 ps. The world average precision is greatly improved recently by the measurement at B-factories, which have a combined uncertainty of ± 0.020 ps. The experiments at high-energy colliders over the last decade contribute to the world average with a combined uncertainty of about ± 0.025 ps.

Chapter 4

The *BABAR* Experiment

The analysis described in this thesis is based on data collected at the *BABAR* detector [57] at the PEP-II asymmetric-energy e^+e^- B-factory [8] at the Stanford Linear Accelerator Center (SLAC) in California. The *BABAR* Collaboration consists of approximately 600 physicists and engineers from 72 institutions in 9 countries. In this chapter I describe the overall design of the *BABAR* detector, the PEP-II storage rings [8, 58], the components of the *BABAR* detector, the trigger and data-acquisition system.

4.1 Introduction

The *BABAR* experiment is designed for the systematic study of CP-violating asymmetries in the decay of neutral B mesons to CP eigenstates. The detector is also well suited for precision measurements of decays of bottom and charm mesons and τ lepton, and for searches of rare processes that become accessible with the high luminosity provided by the PEP-II B-factory.

The PEP-II B-factory is an asymmetric-energy e^+e^- collider operating at a center-of-mass (c.m.) energy of 10.58 GeV, the peak of the $\Upsilon(4S)$ resonance, with

a design luminosity of $3 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and above. In the PEP-II storage rings, a 9 GeV electron beam collides head on with a 3.1 GeV positron beam, resulting in a Lorentz boost of $\beta\gamma \simeq 0.55$ between the center-of-mass and laboratory frames. This boost enables us to study the decay time difference of the two B mesons from decays of the $\Upsilon(4S)$.

Decays of B mesons to CP eigenstates have very small branching ratios, typically 10^{-4} or less, and involve two or more charged particles and often one or more π^0 mesons. To fully reconstructing these final states with high efficiency, the *BABAR* detector has a large and uniform acceptance, and excellent momentum and energy resolution and reconstruction efficiencies for both charged particles and photons over a wide range of momentum. The detector also has very good vertex resolution, both transverse and parallel to the beam direction.

Another crucial aspect of many studies in the *BABAR* experiment, including CP violation and neutral B meson mixing, is the determination of the flavor of the second B meson at the time of decay, which utilizes several flavor sensitive features, including lepton charge, kaon charge and the soft pion from D^* decay. Therefore, excellent electron and muon identifications and kaon/pion separation are crucial features of the *BABAR* detector.

Figure 4.1 shows the longitudinal cross section of the *BABAR* detector through its center. Figure 4.2 shows the end view of the detector. A superconducting coil provides a 1.5-T magnetic field. Inside the solenoid, charged particles are detected and their momenta measured with a silicon vertex tracker (SVT) and a drift chamber (DCH). The ionization energy loss (dE/dx) in both detectors is also used for particle identification. Beyond the outer radius of the DCH is a detector of internally reflected Cherenkov light (DIRC), which is used primarily for charged-hadron identification. A CsI(Tl) electromagnetic calorimeter (EMC) located outside the DIRC is used to detect photons and neutral hadrons, and also to identify electrons. Outside the

solenoid is the instrumented flux return (IFR), which is used for the identification of muons and neutral hadrons.

The right-hand coordinate system is anchored on the main tracking system, the drift chamber, with the z-axis coinciding with its principal axis. The positive y-axis points upward and the x-axis points away from the center of the PEP-II storage rings. The beam axis near the interaction point (IP) is rotated relative to the z-axis by about 20 mrad in the horizontal plane. The positive z direction is defined by the direction of the high-energy beam (electron), except for the 20 mrad rotation. To maximize the acceptance, the center of the whole detector is shifted by 0.37 m relative to the IP along the direction of the high energy beam. The polar angle coverage extends to within 350 mrad of the beam in the forward ($+z$) direction and 400 mrad in the backward direction. The coverage is limited by components of PEP-II, a pair of dipole magnets (B1).

4.2 The PEP-II Storage Rings

PEP-II is an e^+e^- storage ring designed to operate at a center-of-mass energy of 10.58 GeV, which corresponds to the peak of the $\Upsilon(4S)$ resonance, with high luminosity. PEP-II utilizes the original PEP (Positron-Electron Project) beam pipe and magnets for the high-energy electron storage ring (HER). The circumference of the ring is 2200 m. The low-energy positron ring (LER) is built on top of the HER for most of the circumference except for about ± 60 m from the interaction point, where the LER is brought to the same horizontal plane as the HER.

The layout of the PEP-II storage ring and the two mile long accelerator is shown in Fig. 4.3. The linear accelerator is capable of accelerating electrons and positions to more than 90 GeV. For the operation of the B-factory, only a fraction of the linacs are used. The parameters of electron and positron beams are shown in Table 4.1.

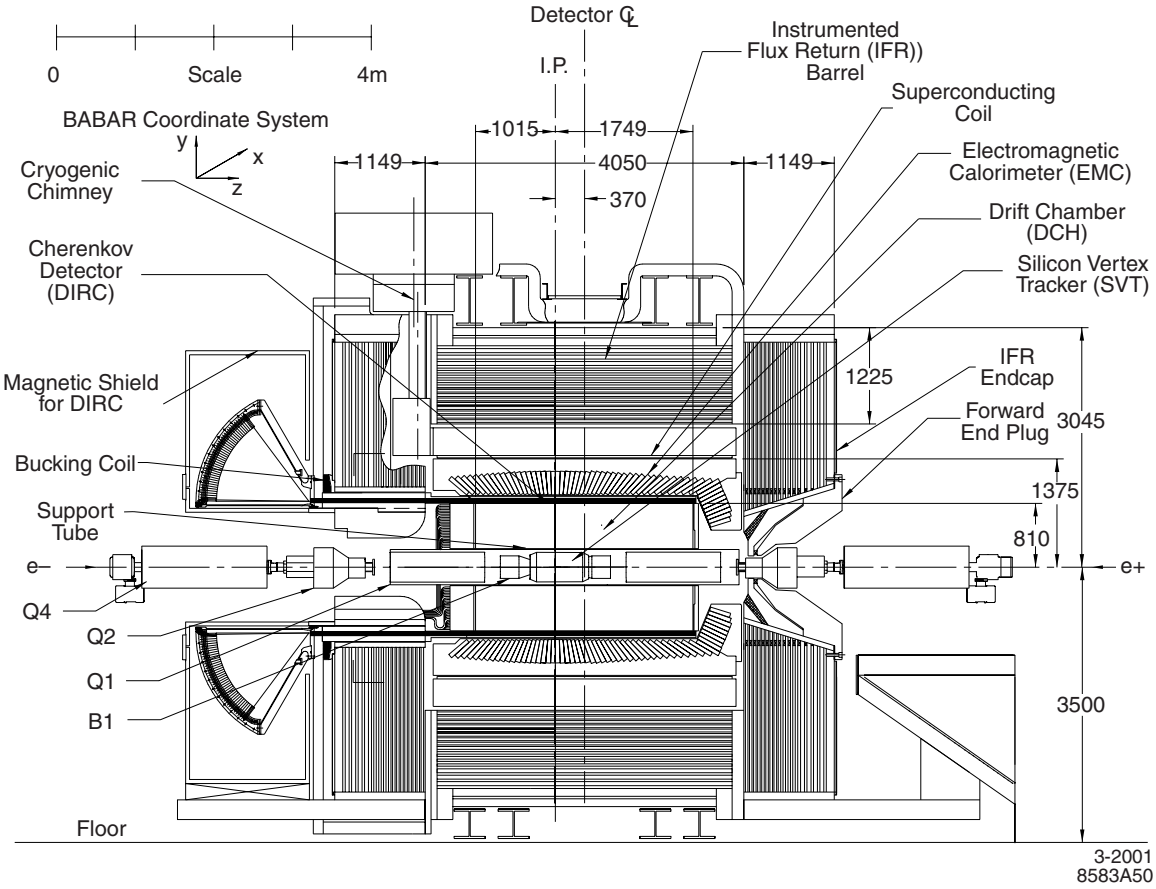

Figure 4.1: *BABAR* detector longitudinal section.

The most challenging part of the the PEP-II design is the interaction region (IR) [59]. As can be seen in Fig. 4.4, because of the zero crossing angle design for the colliding beams, a pair of bending magnets (B1) have to be placed very close to the IP to avoid parasitic collisions. This compact design has great impacts on the design of the *BABAR* detector. For example, the detector coverage is limited by the B1 magnets, and the beam-induced backgrounds increase because the final bending of the beams, which produces strong synchrotron radiations, and the materials near the IP, which generate electromagnetic showers.

A set of radiation protection and monitoring systems are installed for the SVT,

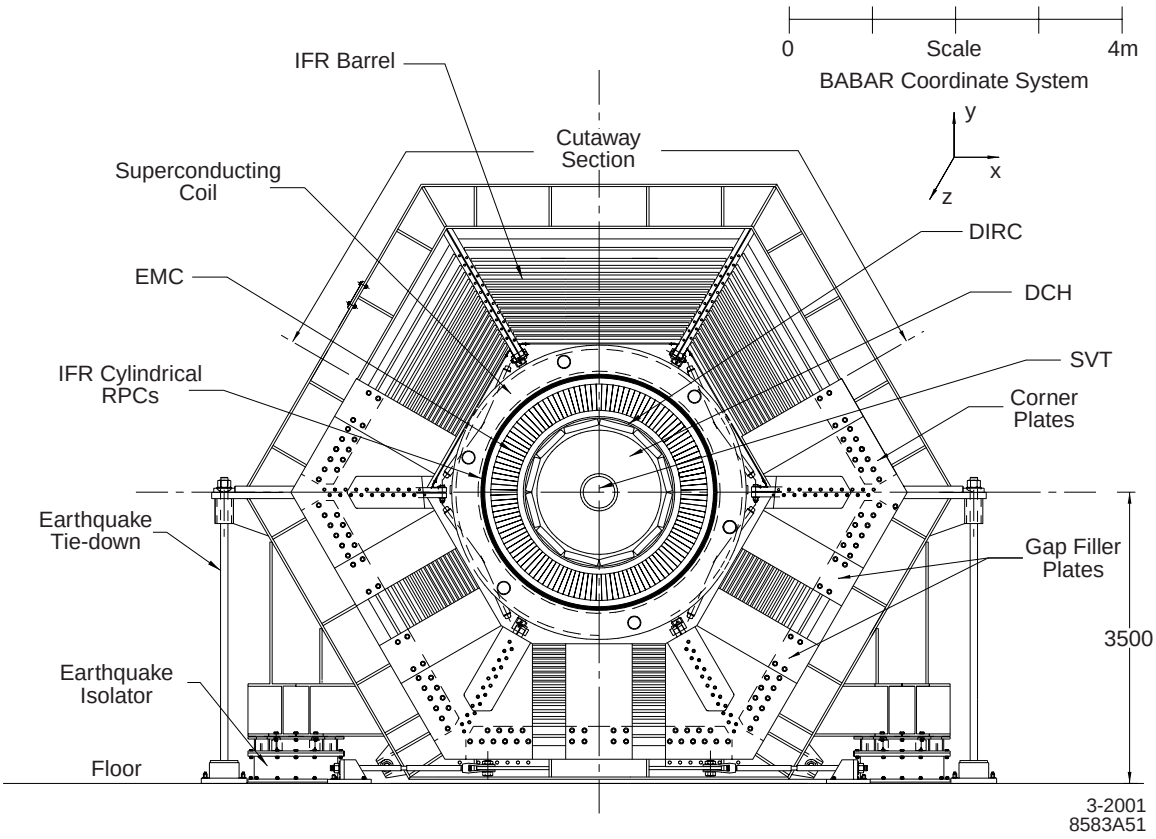

Figure 4.2: *BABAR* detector end view.

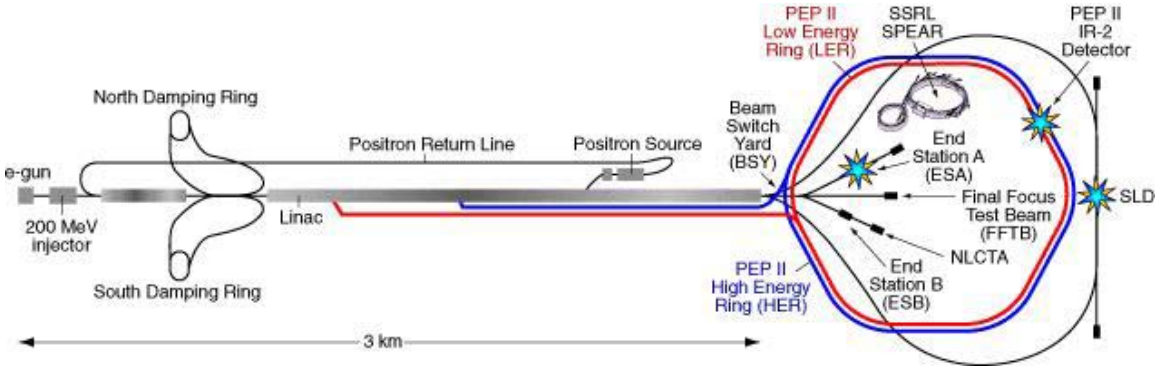

Figure 4.3: PEP-II and accelerator layout.

Table 4.1: PEP-II beam parameters. Values are given both for the design and for typical colliding beam operation in the first year. HER and LER refer to the high energy electron and low energy positron ring, respectively. σ_{L_x}, σ_{L_y} and σ_{L_z} refer to the horizontal, vertical and longitudinal RMS size of the luminous region.

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_{L_x} (μm)	110	120
σ_{L_y} (μm)	3.3	5.6
σ_{L_z} (mm)	9	9
Luminosity ($10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$)	3	2.5
Luminosity (pb^{-1}/d)	135	120

the DCH electronics and the EMC, to reduce the impact of the beam-induced background on the experiment and to prolong the lifetimes of the detector components. The radiation doses are measured with silicon photodiodes. For the SVT, 12 $1\,\mathrm{cm} \times 1\,\mathrm{cm} \times 300\,\mu$m diodes are arranged in three horizontal planes, at, above, and below the beam level, with four diodes in each plane, placed at $z = +12.1\,\mathrm{cm}$ and $z = -8.5\,\mathrm{cm}$ and at a radial distance of 3 cm from the beam line, beneath the SVT readout electronics [60]. The increase of diode leakage current, due to the electron-hole pairs generated by the passing of photons or charged particles, is proportional to the dose rate. For the DCH and the EMC, the PIN diodes are mounted on small CsI(Tl) crystals. The radiation dose is approximately proportional to the photons generated in the crystal. These diodes are installed in sets of four, three on the front face of the endcap calorimeter and one set on the backward end plate of the DCH.

It takes approximately 10–15 minutes to fill the beams. PEP-II typically operates on a 40–50 minute fill cycle. At the end of each fill, it takes about three minutes to refill the beams. While most of the data are recorded at the peak of the $\Upsilon(4S)$

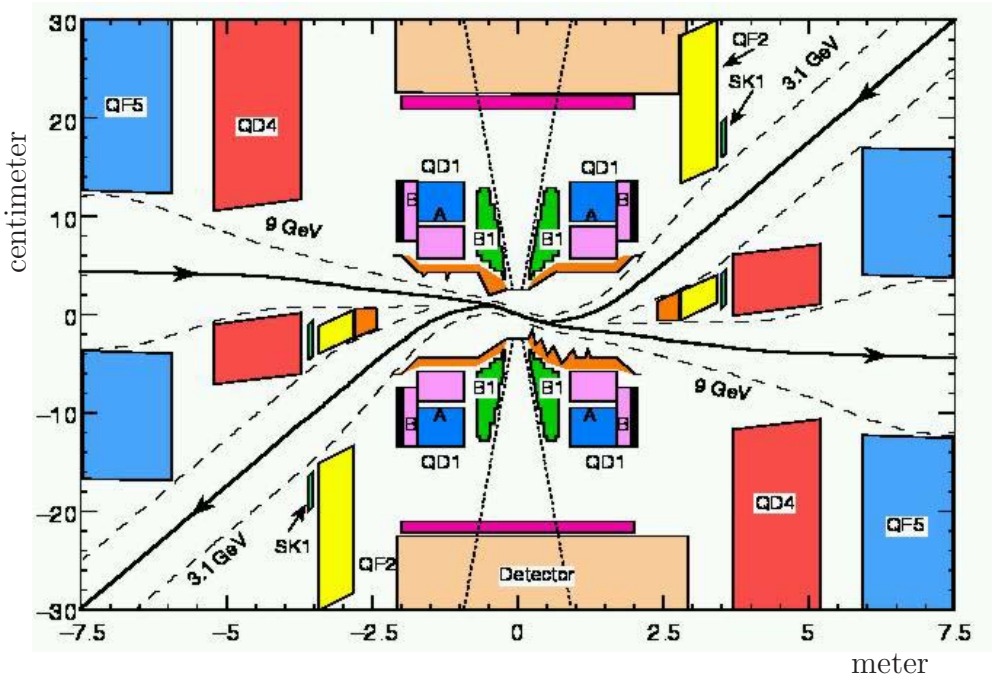

Figure 4.4: The components near the interaction region, including the final bending magnets (B1), focusing magnets, synchrotron radiation masks, and the HER/LER trajectories.

resonance, about 12% are taken at a center-of-mass energy 40 MeV lower to allow for studies of non-$B\overline{B}$ background. Energy scans are performed to find the peak energy of the Υ(4S) by calculating the ratio of the detected multi-hadron event to the di-muon events (see Fig. 4.5).

The relative luminosity is monitored by detecting radiative Bhabha scattering. The absolute luminosity is calculated from data collected by *BABAR* with other QED processes, such as e^+e^- and $\mu^+\mu^-$ pairs. For a data sample of 1 fb^{-1}, the statistical error is less than 1%. The total systematic errors for the absolute luminosity is about 1.5%.

The total number of $B\overline{B}$ pairs detected by *BABAR* is estimated from the selected multi-hadron and $\mu^+\mu^-$ events in both on- and off-resonance data [61] with the

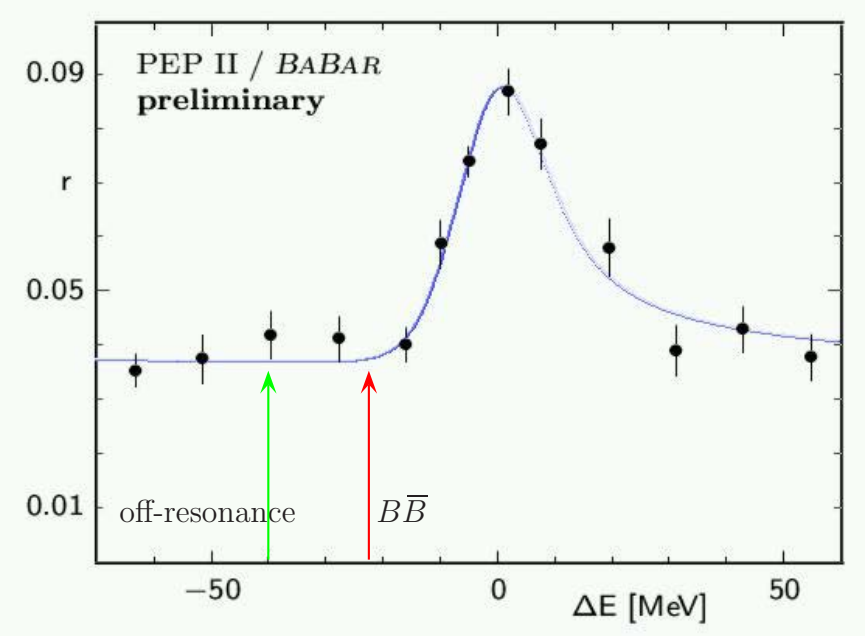

Figure 4.5: $\Upsilon(4S)$ energy scan. The PEP-II is normally operated at the peak, $\Delta E = 0$. About 12% of data are taken at 40MeV below the peak, indicated by the green arrow on the left. The $B\overline{B}$ threshold is indicated by the red arrow on the right.

equation,

$$N_{\Upsilon}(4S) = (N_{\rm MH} - N_{\mu^+\mu^-} \cdot R_{\rm off} \cdot \kappa)/\epsilon_{B\overline{B}}\,,$$

where $N_{\rm MH}$ and $N_{\mu^+\mu^-}$ are numbers of selected multi-hadron and $\mu^+\mu^-$ events, $R_{\rm off}$ is the multi-hadron to $\mu^+\mu^-$ events ratio for off-resonance data, κ is correction factor of $\mathcal{O}(1)$, and $\epsilon_{B\overline{B}}$ is the efficiency of $B\overline{B}$ events to satisfy the selection criteria. The result of the total number of $\Upsilon(4S)$ mesons in data taken in years 1999–2000 is $(22.74 \pm 0.36) \times 10^6$.

4.3 Silicon Vertex Tracker

The innermost component of *BABAR* is the silicon vertex tracker (SVT), located just outside the beryllium beam pipe, and inside a support tube with a radius of about 22 cm, whose central section is fabricated from a carbon-fiber epoxy composite. The SVT provides a vertex resolution of 80 μm or better for a fully reconstructed B decay. It also serves as a stand-alone tracking system for particles with transverse momentum less than 120 MeV, the minimum that can be measured reliably in the DCH alone.

The SVT consists of five layers of 300 μm thick, double-sided silicon strip sensors organized in 6, 6, 6, 16, and 18 modules, respectively. Each module consists of four (layer 1 and 2) to eight (layer 5) sensors. The strips on the opposite sides of the module are orthogonal to each other. The strips that are parallel to the beam axis measure ϕ of a hit (ϕ strips). The ones that are transverse to the beam axis measure z position (z strips). The modules of the inner three layers are straight, while the modules of layer 4 and 5 are arch-shape. The schematic views of SVT are shown in Fig. 4.6 and 4.7,

The modules are divided electrically into two half-modules, which are read out at the ends. The sensors have five different shapes, sizes ranging from $43 \times 42\,\mathrm{mm}^2$ ($z \times \phi$) to $68 \times 53\,\mathrm{mm}^2$. Two identical trapezoidal sensors are added to form the arch modules. The ϕ strips of sensors in the same half-module are electrically connected with wire bonds to form a single readout strip. This results in a total strip length up to 140 mm (240 mm) in the inner (outer) layers. The length of the z strip is about 50 mm in inner layers. For layer 4 and 5 the number of z strips exceeds the number of electronics channels available, requiring two z strips on different sensors be electrically connected (ganged) to a single electronics channel. The total length of these strips is about 100 mm. The readout pitch for z strips is 100 μm (210 μm)

for layer 1–3 (layer 4, 5) and for ϕ 100 μm (110 μm) for layer 1, 4, 5 (layer 2, 3) with one floating strip between readout strips. Parts of the ϕ sides of layer 1 (2) are bounded at 50 μm (55 μm) pitch with no floating strip. The total number of readout channels is approximately 150,000.

The total active silicon area is 0.96 m^2 and the material traversed by particles is about 4% of a radiation length. The geometrical acceptance of SVT is 90% of the solid angle in the c.m. system, typically 86% are used in charged particle tracking.

The combined hardware and software efficiencies are about 97%, excluding defective readout sections (9 out of 208) that were damaged during the installation. As shown in Fig. 4.8, the hit resolution for tracks at normal incidence is about 15 μm for layer 1–3 and about 35 μm for outer layers in z coordinate, and about 10 μm for layer 1 and 2 and about 20 μm for layer 3–5 in ϕ coordinate.

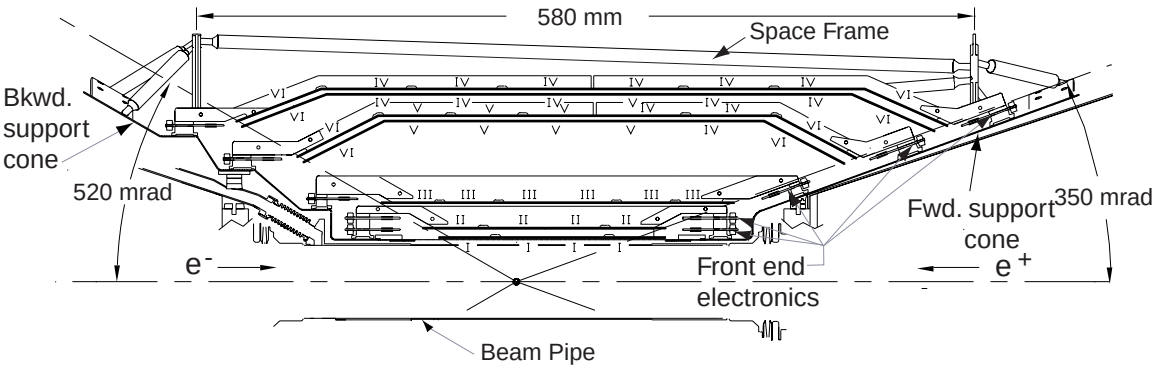

Figure 4.6: Schematic view of SVT: longitudinal section

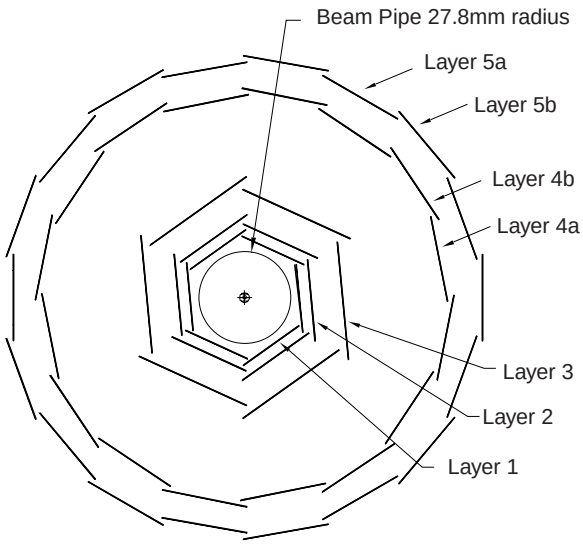

Figure 4.7: Schematic view of SVT: transverse section

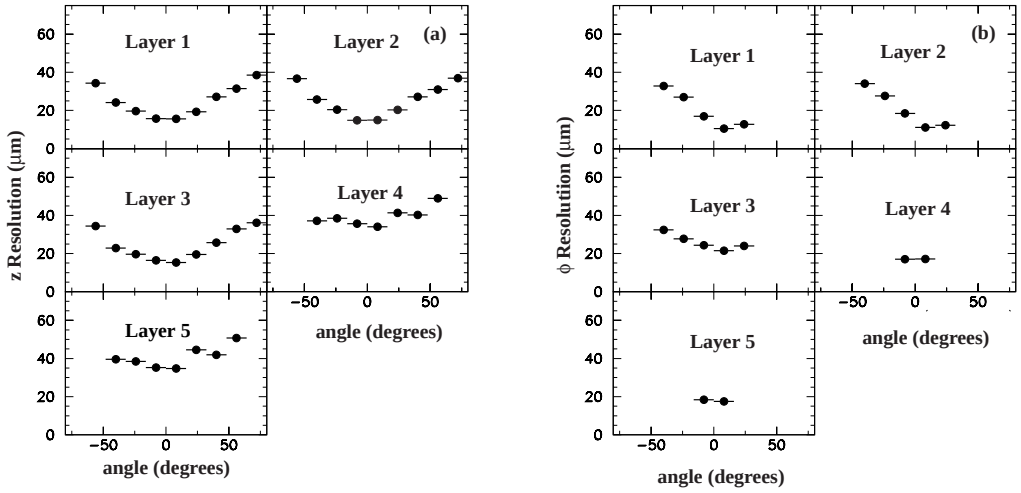

Figure 4.8: SVT hit resolution in the z and ϕ coordinate in microns, plotted as a function of track incident angle in degrees.

4.4 Drift Chamber

Outside the support tube is the drift chamber (DCH). The principal purpose of the DCH is the efficient detection of charged particles and the measurement of their

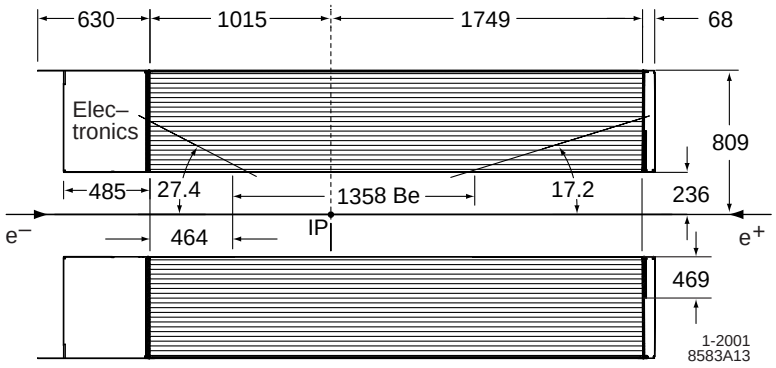

Figure 4.9: Longitudinal section of the DCH with principal dimensions; the chamber center is offset by 370 mm from the interaction point (IP).

momenta and angles with high precision. The DCH also supply information for the charged particle trigger with a maximum time jitter of 0.5 μs (Sec. 4.9). For low momentum particles, the DCH also provide particle identification by measurement of ionization energy loss (dE/dx).

Figure 4.9 shows the longitudinal cross section of the DCH with principal dimensions. The DCH has 40 layers of small hexagonal cells providing up to 40 spatial and ionization loss measurements for charged particles with transverse momentum greater then 180 MeV. Longitudinal position information is obtained by placing the wires in 24 of the 40 layers at small angles with respect to the z-axis. Particles emitted at polar angles between 17.2° and 152.6° traverse at least half of the layers of the chamber before exiting through the endplates. The gas used in the DCH is a 80:20 mixture of helium:isobutane.

The DCH consists of a total of 7,104 small hexagonal drift cells of 11.9 mm by approximately 19.0 mm, arranged in 40 cylindrical layers. The layers are grouped by four into ten superlayers. The stereo angles of the superlayers alternate between axial (A) and stereo (U,V) pairs, in the order AUVAUVAUVA, as shown in Fig. 4.10. Each cell consists of one sense wire surrounded by six field wires. The sense wires

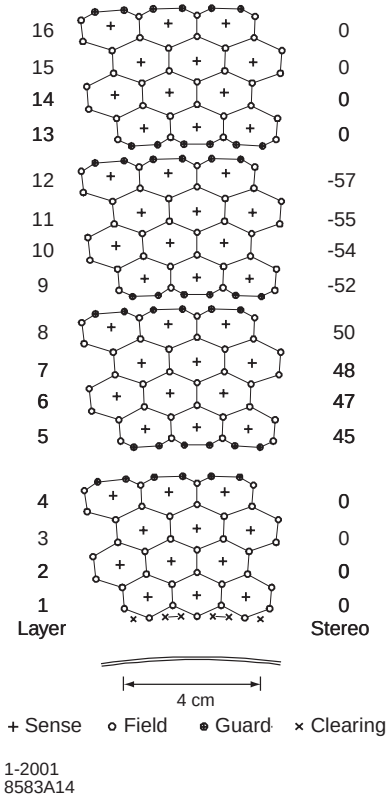

Figure 4.10: Schematic layout of drift cells for the four innermost superlayers. Lines have been added between field wires to aid in visualization of the cell boundaries. The numbers on the right side give the stereo angles (mrad) of sense wires in each layer. The 1 mm-thick beryllium inner wall is shown inside of the first layer.

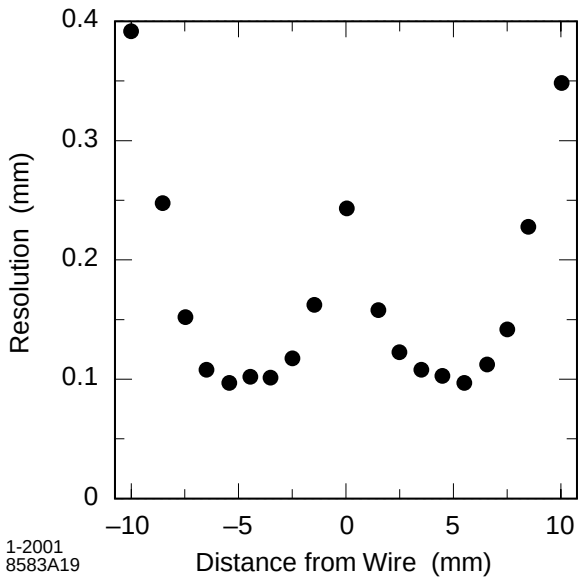

Figure 4.11: DCH position resolution as a function of the drift distance in layer 18, for tracks on the left and right side of the sense wire. The data are averaged over all cells in the layer.

are made of tungsten-rhenium with 20 μm diameter. The field wires are made of aluminum with a diameter of 120 μm. All wires are coated with gold. A positive high voltage (nominal value is 1960 V) is applied to the sense wires, and the field wires are at ground potential.

The typical position resolution of charged tracks as a function of the drift distance is shown in Fig. 4.11. The results are based on multi-hadron events for data averaged over all cells in layer 18.

The dE/dx, for charged particles traversing the DCH is derived from measurement of the total charge collected in each drift cell through an extraction algorithm with various corrections. The RMS resolution of the measured dE/dx is typically 7.5%.

4.5 Cherenkov Light Detector

The detector of internally reflected Cherenkov light (DIRC) is a novel ring-imaging Cherenkov detector, located outside the outer shell of the DCH. The principal purpose of the DIRC is to provide good π/K separation from 0.7 to 4 GeV/c.

Figure 4.12 shows a schematic of the DIRC geometry that illustrates the principles of light production, transport, and imaging. The radiator is a set of long, thin bars made of synthetic, fused silica, with rectangular cross section. These bars also serve as light guides for the light trapped in the radiator by total internal reflection. The magnitudes of light angles are preserved by the parallel flat surfaces of the bars.

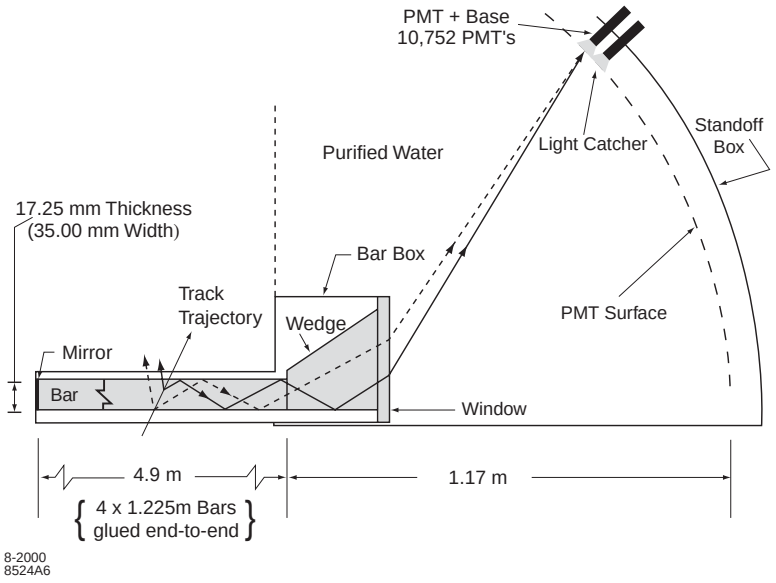

Figure 4.12: Schematics of the DIRC fused silica radiator bar and imaging region. Not shown is a 6 mrad angle on the bottom surface of the wedge (see text).

The mean index of refraction (n) of fused silica is 1.473. The Cherenkov angle (θ_C) is related to n and the speed of charged particle (v) through the familiar relation $\cos\theta_C = 1/n\beta$, where $\beta = v/c$, and c = velocity of light. Therefore, by measuring θ_C and the momentum of the track, one can identify the mass of the particle.

For particles with $\beta \simeq 1$, some photons will always lie within the limits of total internal reflection, and will be transported to either or both ends of the bar, depending on the incident angle. A mirror is placed at the forward end, perpendicular to the bar axis, to reflect incident photons to the backward end, where is instrumented. Most photons that arrive at the instrumented end will emerge into a water-filled expansion region, call the standoff box. The photons are detected by an array of densely packed photo-multiplier tubes (PMTs), which are placed at a distance of about 1.2 m form the bar end.

The bars are placed into 12 hermetically sealed containers called bar box. The bar boxes are arranged in a 12-sided polygonal barrel. Each bar box contains 12 bars, for a total of 144 bars. Within a bar box the 12 bars are optically isolated by a 150 μm air gap between neighboring bars. The bars are 17 mm-thick, 35 mm-wide, and 4.9 m-long. Each bar is assembled from four 1.225 m pieces that are glued end-to-end.

The standoff box is made of stainless steel, consisting of a cone, cylinder, and 12 sectors of PMTs. It contains about 6,000 liters of purified water, with an average index of refraction of about 1.346. Each of the 12 PMT sectors contains 896 PMTs with 20 mm-diameter, in a closely packed array inside the water volume. A hexagonal light catcher cone is mounted in front of the photocathode of each PMT, which results in an effective active surface area light collection fraction of about 90%. The support structure and geometry of the DIRC are shown in Fig. 4.13 and 4.14.

The radiator bars subtend a solid angle corresponding to about 94% of the azimuth and 83% of the c.m. polar angle. The geometric contribution to the single photon Cherenkov angle resolution due to the sizes of the bars and PMTs is about 7 mrad. This value is slightly larger than the rms spread of the photon production and transmission dispersions. The overall single photon resolution is estimated to be about 10 mrad.

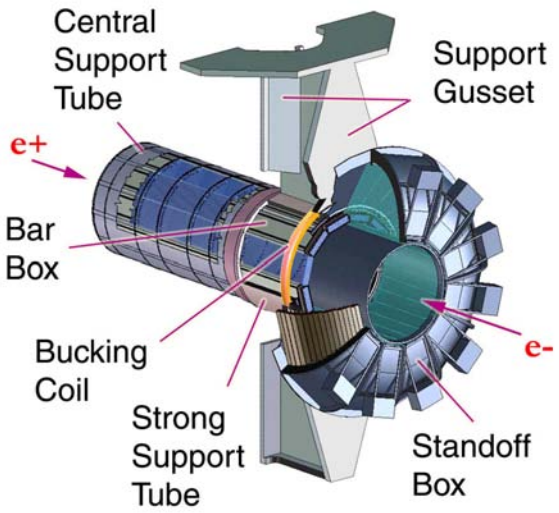

Figure 4.13: Exploded view of the DIRC mechanical support structure.

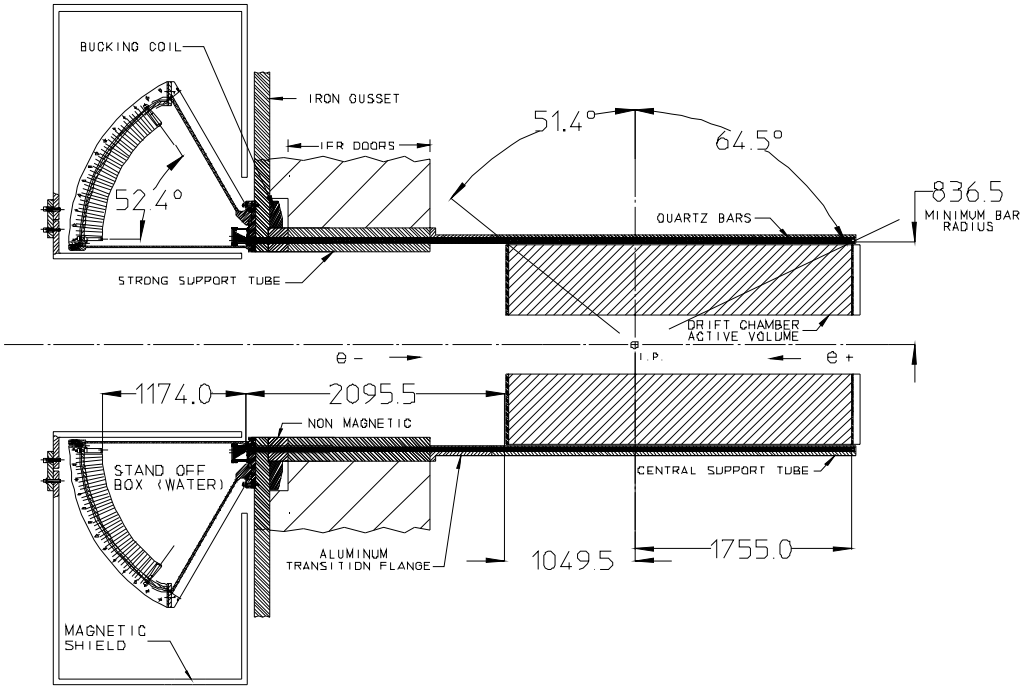

Figure 4.14: Elevation view of the nominal DIRC system geometry.

About 80% of the light is maintained after a few hundreds of bounces along the bars. The overall detection efficiency is dominated by the quantum efficiency of PMTs ($\mathcal{O}(20\%)$). The expected number of photoelectrons (N_{pe}) is about 28 for a $\beta = 1$ particle entering normal to the surface at the center of a bar, and increases by over a factor of two in the forward and backward direction.

An unbinned maximum likelihood formalism is used to incorporate all information provided by the space and time measurements form the DIRC. Currently a likelihood value of each of the five stable particle types (e, μ, π, K, p) is calculated if the track passes through the active volume of the DIRC. The expected separation between kaon and pions is about 4.2 σ at 3 GeV/c and about 2.5 σ at 4.2 GeV/c.

4.6 Electromagnetic Calorimeter

The electromagnetic calorimeter (EMC) is designed to measure electromagnetic showers with excellent efficiency, energy and angular resolution over the energy range from 20 MeV to 9 GeV. This capability allows the detection of photons from π^0 and η decays as well as from electromagnetic and radiative processes. The EMC is also a major component for electron identification.

The EMC consists of 6,580 Thallium-doped CsI crystals, of which 5,760 crystals are contained in a cylindrical barrel support structure arranged in 48 distinct rings, and 820 crystals arranged in eight rings. The EMC has full coverage in azimuth and extends in polar angle from 15.8° to 141.8° corresponding to a solid-angle coverage of 90% in the c.m. system (see Fig. 4.15).

CsI crystals are doped with 0.1% thallium. They are machined into tapered trapezoids. The transverse dimensions of the crystals vary. The typical area of the front face is $4.7 \times 4.7\,\mathrm{cm}^2$ and the back face area is $6.1 \times 6.0\,\mathrm{cm}^2$. The length of the crystals increases from 29.6 cm in the backward direction to 32.4 cm in the forward

direction. The surface of the crystal is polished and wrapped with two layers of diffuse white reflector. A schematic of the crystal is shown in Fig. 4.16. The light yield is required to be uniform to within $\pm 2\%$ in the front half of the crystal; the limit increases linearly to $\pm 5\%$ at the rear face. The photons of electromagnetic showers in a crystal are detected by two $2 \times 1\,\mathrm{cm}^2$ silicon PIN diodes. The diodes have a quantum efficiency of 85% for the CsI(Tl) scintillation light. The depletion voltage is about 70 V, at which voltage the typical leakage current is 4 nA and capacitance is 85 pF; the diodes are operated at 50 V. For each crystal-diode assembly, the light yield is measured with the 1.836 MeV photon line from a ^{88}Y radioactive source. The resulting signal distribution has a mean and RMS of 7300 and 890 photoelectrons/MeV. The electronics system has an equivalent noise energy of less than 250 keV, which has negligible impact on the energy resolution of electromagnetic showers from 20 MeV to 9GeV.

A typical electromagnetic shower spreads over several crystals, called a cluster. The reconstruction algorithm requires that at least one crystal in a cluster exceeds 10 MeV and the surrounding crystals pass certain thresholds. A cluster can contain multiple local energy maxima, called bumps. An iterative algorithm is used to determine the energy of the bumps, by calculating the weight of each crystal associated with a certain bump according to the distances between crystals and the centroid of the bump, and calculating the bump centroid according to the weights. The position of a bump is calculated using a center-of-gravity method with logarithmic, rather than linear weights.

To determine whether a bump is associated with a charged particle, the track is projected onto the inner surface of the EMC. If the centroid of the bump is consistent with the angle and momentum of the track, the bump is associated with this charged particle, otherwise, it is assumed to originate from a neutral particle.

The energy resolution can be measured with or inferred from several sources,

including radioactive sources, mass resolutions of π^0 and η mesons decaying to two photons, the decay of $\chi_{c1} \to J/\psi\gamma$, and electrons from Bhabha scattering. A fit to the energy dependence with an empirical function results in

$$\frac{\sigma_E}{E} = \frac{(2.32 \pm 0.30)\%}{\sqrt[4]{E(\mathrm{GeV})}} \oplus (1.85 \pm 0.12)\% \,.$$

The measurement of the angular resolution is based on the analysis of π^0 and η decays to two photons of approximately equal energy. The resolution varies between about 12 mrad at low energies and 3 mrad at high energies. A fit with an empirical parameterization of energy dependence results in

$$\sigma_\theta = \sigma_\phi = \left(\frac{3.87 \pm 0.07}{\sqrt{E(\mathrm{GeV})}} + 0.00 \pm 0.04 \right) \mathrm{mrad} \,.$$

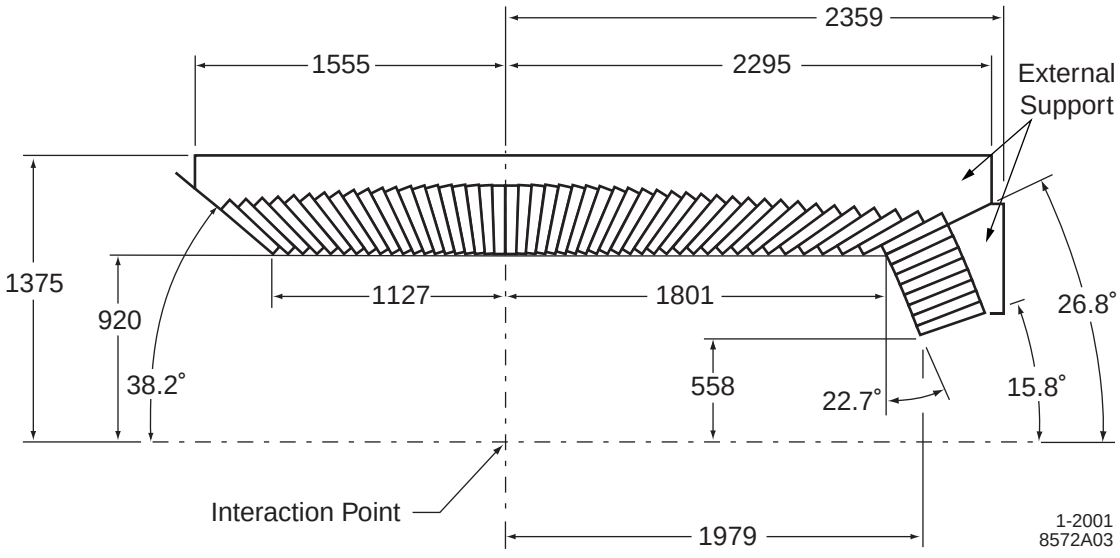

Figure 4.15: A longitudinal cross section of the EMC (only the top half is shown) indicating the arrangement of the 56 crystal rings. The detector is axially symmetric around the z-axis. All dimensions are given in mm.

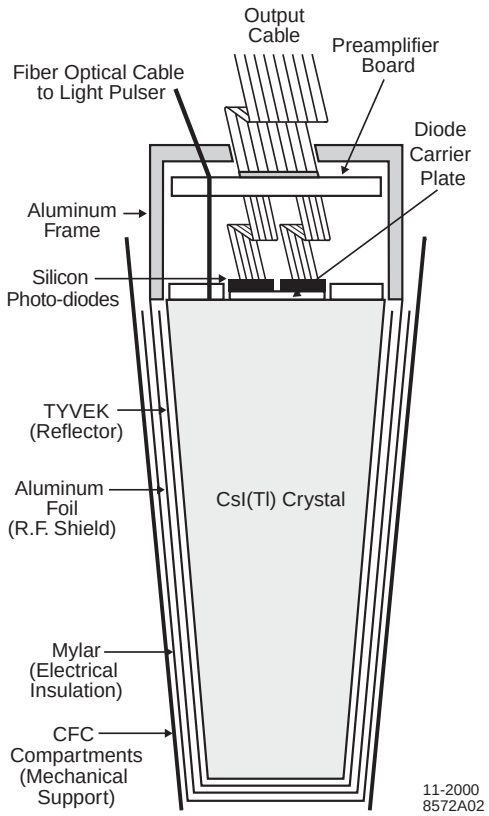

Figure 4.16: A schematic of the wrapped CsI(Tl) crystal and the front-end readout package mounted on the rear face. Also indicated is the tapered, trapezoidal CFC compartment, which is open at the front. This drawing is not to scale.

The π^0 mass resolution reconstructed from two photons over the full photon energy range in $B\overline{B}$ events is 6.9 MeV.

The particle identification techniques with the EMC are described in Chapter 5.

4.7 Instrumented Flux Return

The main purpose of the Instrumented Flux Return (IFR) is to identify muons with high efficiency and good purity, and to detect neutral hadrons (primarily K_L and neutrons) over a wide range of momenta and angles.

The active detectors in the IFR are single gap resistive plate chambers (RPCs) with two-coordinate readout. They are installed in the gaps of 18 steel plates in the barrel and the end doors of the flux return, as illustrated in Fig. 4.17. There are 19 RPC layers with 342 modules in the barrel and 18 RPC layers with 432 modules in two end doors. The thickness of the steel plates ranges from 2 cm for the inner nine layers to 10 cm for the outermost ones. The gap between the steel plates is 3.5 cm in the inner layers of the barrel and 3.2 cm elsewhere. In addition, two layers of cylindrical RPCs with 32 modules are installed between the EMC and the magnet cryostat to detect particles exiting the EMC.

RPCs detect streamers from ionizing particles via capacitive readout strips. The RPCs consist of two 2-mm-thick bakelite (phenolic polymer) sheets, separated by a gap of 2 mm, in which a gas mixture, typically 56.7% Argon, 38.8% Freon 134a, and 4.5% isobutane is filled. The external surfaces of bakelite sheets are coated with graphite and are connected to high voltage ($\sim$8 kV) and ground. The bakelite surfaces facing the gap are treated with linseed oil except on the cylindrical RPCs. A cross section of an RPC is shown schematically in Fig. 4.18. The widths of the strips are between 16 mm and 38 mm.

To calculate the efficiency in a given chamber, nearby hits in a given layer and hits in different layers are combined to form clusters. The residual distributions from straight line fits to two-dimensional clusters typically have an RMS width of less than 1 cm. An RPC is considered efficient if a signal is detected at a distance of less than 10 cm from the fitted straight line to either of the two readout planes. Of the active RPC modules, 75% exceed an efficiency of 90%.

Muons can penetrate many layers of steel plates while other particles (except for non-detectable neutrinos) can't. This property enable the IFR to separate muons from other particles. Details of muon identification is described in Chapter 5.

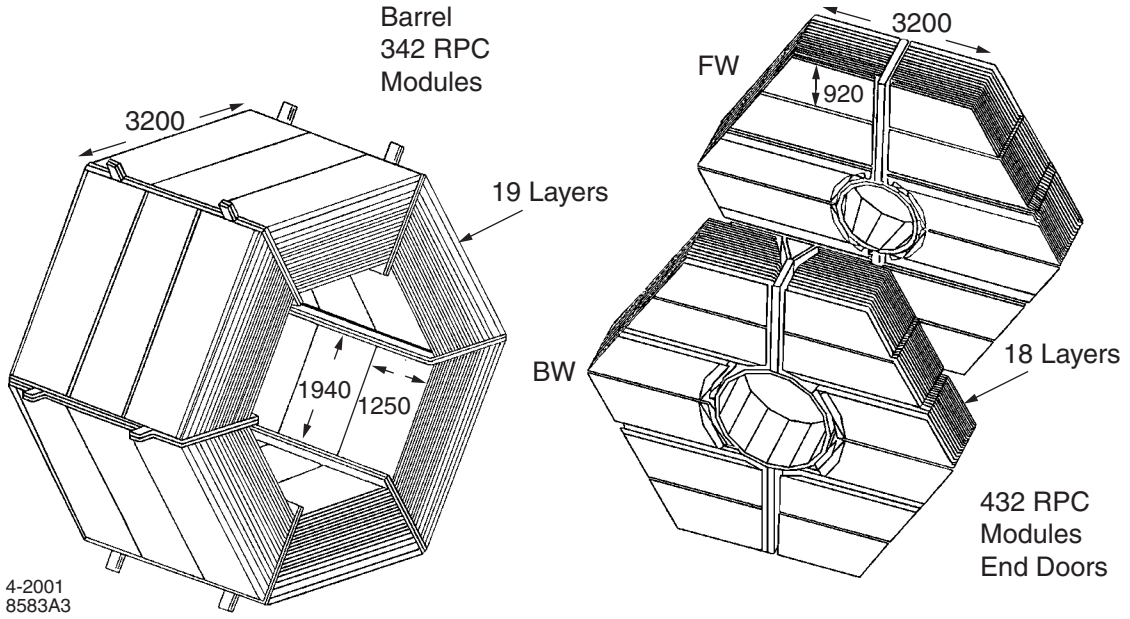

Figure 4.17: Overview of the IFR: Barrel sectors and forward (FW) and backward (BW) end doors; the shapes of the RPC modules and their dimensions are indicated (in mm).

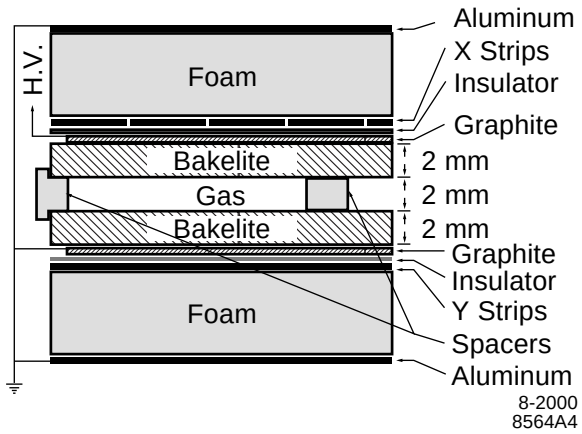

Figure 4.18: Cross section of a planar RPC with the schematics of the high voltage (HV) connection.

4.8 Solenoid and Magnets

The *BABAR* magnet system consists of a superconducting solenoid, a segmented flux return and a field compensating coil. Figure 4.1 shows major components of the magnets. The main purposes of the magnet system are to provide magnetic field for measuring charged particle momentum, serve as hadron absorber for separating muons and hadrons, and provide the overall structure and support for *BABAR* detector elements.

The flux return consists of a hexagonal barrel and forward and backward doors. The coil cryostat is mounted inside the barrel. The coil is made of 16 strand niobium-titanium Rutherford cable coextruded with pure aluminum stabilizer. Each strand has a diameter of 0.8 mm and is wound from thin filaments, each less than 40 μm in diameter. The coil is cooled to 4.5 K by liquid helium and the operating current is 4596 A. The structure of the flux return is designed to sustain the forces produced by the magnetic field, and earthquake.

The central field produced by the coil is 1.5 T. The samarium-cobalt B1 and Q1 magnets are located inside the solenoid as shown in Fig. 4.1. They cannot sustain high radial magnetic field. The radial component at Q1 and $r = 200$mm is kept below 0.25 T. The stray field leaking into the conventional iron quadrupole Q2, Q4 and Q5 is less than 0.01 T averaged over their apertures.

The magnetic field is mapped by five sets of B_r and B_z, and two B_ϕ movable Hall probes. One NMR probe located at $r = 89$mm provides a very precise field reference near the z-axis. In the tracking volume, B_r and B_z fields vary by less than 0.05 T, and B_ϕ is less than 1 mT. The variation of bending field for high momentum tracks is less than 2.5% along the path.

4.9 Trigger

The trigger system provides a fast filtering before events are recorded. The system is required to select events of interests (see Table 4.2) with a high, stable and well-understood efficiency while rejecting background events to keep the event rate under 120 Hz. The trigger should contributes no more than 1% to dead time. The total trigger efficiency is required to exceed 99% for $B\overline{B}$ events and at least 95% for continuum events. Less stringent requirements apply to other types of events.

Table 4.2: Cross sections, production and trigger rates for the principal physics processes at 10.58 GeV for a luminosity of 3×10^{33} cm^{-2}s^{-1}. The e^+e^- cross section refers to events with either one, or both e^+e^- inside the EMC detection volume.

Event type	Cross section (nb)	Production rates (Hz)	Level 1 Trigger rates (Hz)
$b\overline{b}$	1.1	3.2	3.2
other $q\overline{q}$	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

There are two levels in the trigger system. Level 1 (L1) is implemented in hardware, followed by Level 3 (L3)[1] in software. L1 is configured to have an output rate of typically 1 kHz during normal operation. Triggers are produced within a fixed latency window of 11–12 μs after the e^+e^- collision, and delivered to the Fast Control and Timing System. The L1 trigger is based on the track segments in the DCH, energy deposit in the EMC *towers*, each of which consists of 19–24 crystals, and the hits on the triggers in each of the ten IFR trigger sectors. The IFR trigger is used for triggering on $\mu^+\mu^-$ events and cosmic rays.

For a typical L1 rate of 1 kHz, Bhabha and other physics events contribute

[1]There is no Level 2 due to a historic reason.

*sim*130 Hz. There are also about 100 Hz of cosmic ray and 20 Hz of random beam crossing triggers. The remaining triggers are due to the lost particles (off-energy/orbit particles from e^+e^- beams) interacting with the beam pipe and other components near the interaction region.

The L3 receives the output from L1, performs a further rate reduction, and identifies and flags the special categories of events needed for luminosity determination, diagnostic, and calibration purposes. The L3 software runs on the online computer farm. The L3 filters have access to the complete event data for making their decisions. L3 operates by refining methods used in L1. For example, better DCH tracking and vertex resolution, and EMC clustering filters allow for greater rejection of beam backgrounds and Bhabha events. For a typical run on the Υ(4S) peak with an beverage luminosity of 2.6×10^{33} cm^{-2}s^{-1}, the desired physics events contribute 13% of the total output while the calibration and diagnostic samples comprise 40%.

4.10 Data Acquisition

The *BABAR* data acquisition system is a chain from the common front-end electronics through the embedded computing processors, to the logging of event data. The data acquisition chain supports an L1 trigger accept rate of up to 2 kHz, with an average event size of ∼32 kbytes, and a maximum output rate of 120 Hz. It should contribute less than a time-average 3% to dead time during normal data acquisition.

The data acquisition system consists of the following major subsystems: *Online Dataflow*, for communication with and control of the detector systems' front-end electronics, and the acquisition and building of event data from them; *Online Event Processing*, for processing of complete events, including L3 triggering, data quality monitoring, and the final stages of calibration; and *Logging Manager*, for receiving selected events sent from the Online Event Processing and writing them to disk files

for use as input to the *Online Prompt Reconstruction* processing. The entire system is coded primarily in the C++ language, with some Java for graphical user interfaces.

The data from the front-end electronics are routed via optical fiber links to a set of 157 custom VME readout modules (ROMs), which are then grouped and housed in 23 data acquisition VME crates. One ROM in each crate aggregates the data and forwards them for event building to 32 commercial Unix workstations [62]. The crates and farm computers communicate via full-duplex 100 Mbits/s Ethernet, linked by a network switch. The 32 online farm computers host the Online Event Processing and L3 trigger software. The events accepted by the trigger are logged via TCP/IP to a logging server and written to a disk buffer for later reconstruction and archival storage.

Chapter 5

Reconstruction and Particle Identification

5.1 Charged Track Reconstruction

5.1.1 Algorithm

Charged particle tracks are reconstructed by DCH and SVT. The track finding and the fitting procedures utilize a Kalman filter algorithm [63], which takes into account the detailed distribution of material and magnetic field that tracks pass through. The L3 trigger and tracking algorithm provide information for the offline charged particle track reconstruction. It first estimates the event start time t_0 from a fit to the distance of closest approach to the z-axis, the azimuth angle of the track and t_0 based on the four-hit track segments in the DCH superlayers. Tracks are then selected by performing helix fits to the hits found by the L3 track finding algorithm. Additional DCH hits that may belong on these tracks are searched, and t_0 is further improved by using only hits associated with tracks; the noise and background hits are reduced due to improved t_0. Two more tracking procedures designed to find

tracks that either do not pass through the entire DCH or do not originate from the IP are performed. These algorithms primarily use tracks that have not been assigned to other tracks. At the end of this process, tracks are again fit with a Kalman filter method.

The resulting tracks from the DCH are extrapolated into the SVT. The SVT track segments that are consistent with the expected error in the extrapolation are added to the tracks. The SVT segments with 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 standalone track finding algorithms. The first reconstructs tracks starting with triplets of space points (matched ϕ and z hits) in layers 1, 3, and 5 of the SVT, and adding consistent space points from the other layers. A minimum of four space points are required to from a good track. The second algorithm starts with circle trajectories from ϕ hits and then adds z hits to from helices.

Finally, an attempt is made to combine tracks that are only found by one of the two tracking systems and thus recover tracks scattered in the material of the support tube.

5.1.2 Tracking efficiency

The efficiency of reconstructing tracks in the DCH is determined as the ratio of the number of tracks reconstructed in the DCH to the number of tracks detected in the SVT. Figure 5.1 shows the DCH track reconstruction efficiencies as functions of the transverse momentum and polar angle, for two operating voltage settings. At design voltage of 1960 V, the efficiency averages $98 \pm 1\%$ per track above 200 MeV.

The reconstruction efficiency of the standalone SVT tracking algorithm is estimated with a detailed Monte Carlo study using soft pions (π_s) from events of the

type $\overline{B} \to D^{*+}X$ followed by $D^{*+} \to D^0\pi_s^+ \to K^-\pi^+\pi_s^+$. A comparison of the detected slow pion spectrum with the Monte Carlo prediction, and the reconstruction efficiency based on Monte Carlo study are shown in Fig. 5.2. The spectrum between data and Monte Carlo prediction are in good agreement. The slow pion reconstruction efficiency shown in Fig. 5.2 demonstrates that the SVT significantly extends the capability of the charged particle detection down to transverse momenta of ~ 50MeV; this is especially important for analyses involving charged D^*.

5.1.3 Track parameter resolution

Tracks are parameterized with five parameters, $(d_0,\ z_0,\ \phi_0,\ \omega,\ \tan\lambda)$. These are measured at the point of closest approach to the z-axis. d_0 and z_0 are the distance of this point to the origin of the coordinate system in the x-y plane and along the z-axis, respectively, ϕ_0 the azimuthal angle, ω the curvature ($\omega = 1/p_T$), and λ the dip angle relative to the transverse plane.

The track parameter resolutions are measured with cosmic ray muons that pass close to the interaction point by comparing impact parameters of upper and lower halves of the track, and with multi-hadron events by comparing parameters of a single track with the vertex reconstructed from the remaining tracks. The resolutions of tracks with momenta above $p_T > 3\,\mathrm{GeV}$ measured with cosmic rays are:

$$\sigma_{d_0} = 23\,\mu\mathrm{m} \qquad \sigma_{\phi_0} = 0.43\,\mathrm{mrad}$$
$$\sigma_{z_0} = 29\,\mu\mathrm{m} \qquad \sigma_{\tan\lambda} = 0.53 \times 10^{-3}\,.$$

The resolutions of parameters d_0 and z_0 as a function of the transverse momentum measured with multi-hadron events are shown in Fig. 5.3.

While the position and angle measurements near the IP are dominated by the SVT measurements, the DCH contributes primarily to the p_T measurement. The

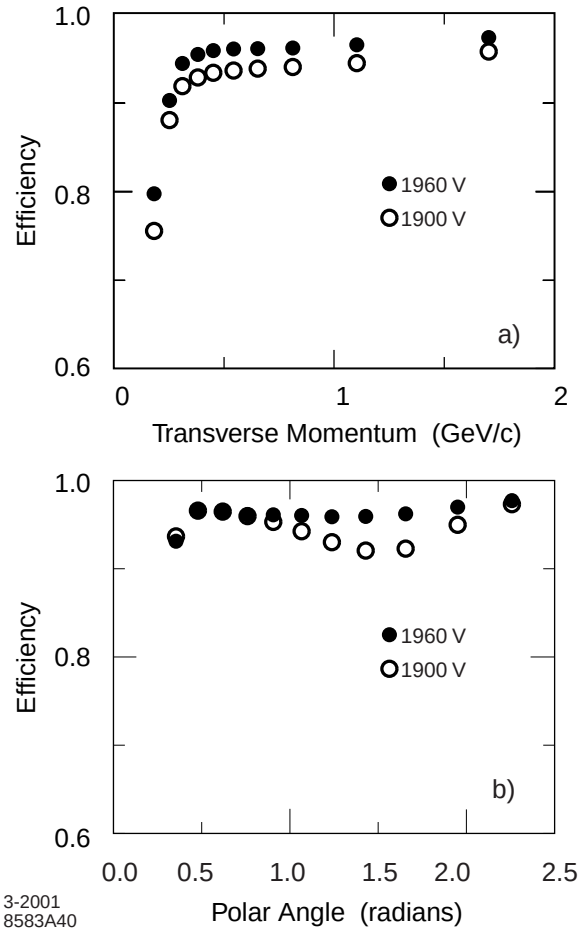

Figure 5.1: The track reconstruction efficiency in the DCH at operating voltages of 1900 V and 1960 V, as a function of a) transverse momentum and b) polar angle. The efficiency is measured in multi-hadron events as the fraction of all tracks detected in the SVT for which the DCH portion is also reconstructed.

resolution of p_T is measured as

$$\sigma_{p_T}/p_T = (0.13 \pm 0.01)\% \cdot p_T + (0.45 \pm 0.03)\% \,,$$

where p_T is measured in GeV.

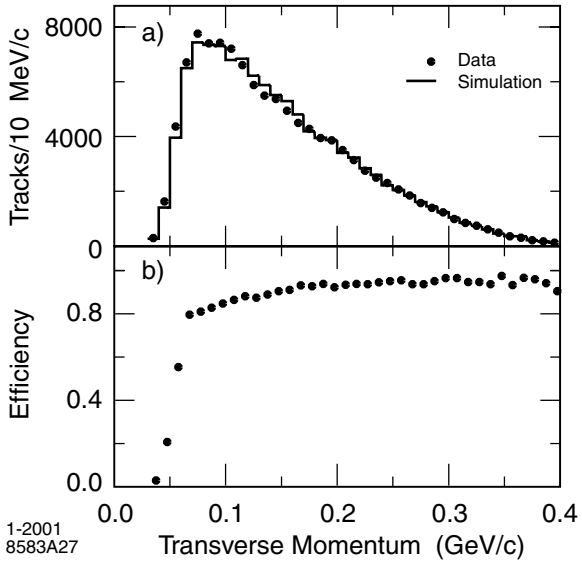

Figure 5.2: Monte Carlo studies of low momentum tracks in the SVT: a) comparison of data (contributions from combinatoric background and non-$B\overline{B}$ events have been subtracted) with simulation of the transverse momentum spectrum of pions from $D^{*+} \to D^0\pi^+$ in $B\overline{B}$ events, and b) efficiency for slow pion detection derived from simulated events.

5.2 Particle Identification

5.2.1 Electron identification

An electron is separated from other charged particles according primarily to the shower energy and lateral shape in the EMC, and its momentum. The Cherenkov radiation angle in DIRC and dE/dx in DCH are required to be consistent with the particle being an electron [64]. The most important quantity is the ratio of energy deposited in the EMC to the momentum of the track (E/p). An electron is very likely to deposit all its energy into the EMC, therefore its E/p distribution is peaked at one, while other types of charged particles usually deposit only minimum ionizing energy, or some random amount of energy if nuclear interactions occur between hadrons and nucleons.

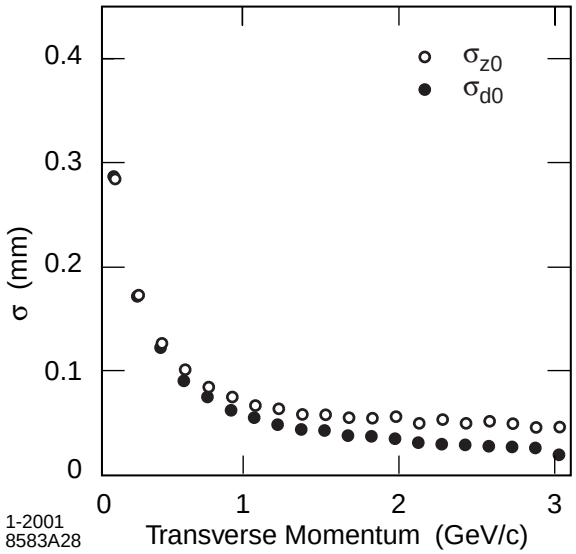

Figure 5.3: Resolution in the parameters d_0 and z_0 for tracks in multi-hadron events as a function of the transverse momentum. The data are corrected for the effects of particle decays and vertexing errors.

The shower shape also provides important information. The shower shape created by an electron is usually more regular and narrower, while the shower shape of an hadron can be irregular and wide when nuclear interaction occurs. Two shower shape variables are used, lateral energy distribution (LAT) [65] and Zernike moments (A_{mn}) [66]. LAT measures the spread of the shower:

$$LAT = \frac{\sum_{i=3}^{n} E_i r_i^2}{\sum_{i=3}^{n} E_i r_i^2 + E_1 r_0^2 + E_2 r_0^2}, \qquad E_1 \geq E_2 \geq \cdots \geq E_n \,,$$

where the sum extends over all crystals in a shower, $r_0 = 5$cm, the average distance between two crystal front faces, and r_i is the distance between crystal i and the shower center. Electromagnetic showers in the EMC have most of their energy deposited in one or two crystals, hence LAT is smaller for electromagnetic showers than for hadronic showers.

The Zernike moments measures the irregularity of the shower shape:

$$A_{nm} = \sum_{r_i \leq R_0}^{n} \frac{E_i}{E} f_{nm}\left(\frac{r_t}{R_0}\right) e^{-im\phi_i}, \qquad R_0 = 15\text{cm}\,,$$

where ϕ is the orientation of the lateral coordinate system and

$$f_{nm}(\rho) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)! \rho^{n-2s}}{s!((n+m)/2-s)!((n-m)/2-s)!}\,,$$

with $n, m \geq 0$ integers, $n - m$ even and $m \leq n$. In *BABAR*, only $|A_{42}|$ is used. Shower shapes and E/p are complementary. If a hadron deposits a large amount of energy, the shower shape variables tend to have large values as well.

The distance between the centroid of the cluster and the extrapolation of the corresponding track at the surface of the EMC also help to reject hadrons. The distribution of this distance (measured by $\Delta\phi$) for electrons is narrower than for hadrons, typically less than 0.05, depending on the track's transverse momentum.

There are five levels of electron selections used in *BABAR* analyses: *noCal, very loose, loose, tight* and *very tight*. The *noCal* is for tracks in the region outside the EMC acceptance and uses only dE/dx information. The selections relevant to the analysis described in this thesis are *loose* and *very tight*. Both selections require at least three crystals in the cluster. The loose selection requires that $E/p > 0.65$ and $500 < dE/dx < 1000$; the very tight selection requires that $0.89 < E/p < 1.3$, $540 < dE/dx < 860$, LAT < 0.6, $|A_{42}| < 0.11$, and the DIRC cherenkov angle and $\Delta\phi$ consistent with an electron.

The efficiency for the loose (very tight) selection is higher than 98% (90%) for tracks more than 1 GeV of momentum, and the average misidentification probabilities are lower than 0.2% for very tight selection.

5.2.2 Muon identification

The primary detector for muon identification [67] is the IFR. A muon is expected to penetrate many layers of steel and leave only minimum ionizing energy in the EMC. Several variables are used to discriminate muons from other charged particles. They include (1) the total interaction lengths (λ) from the IP to the last layer of RPC associated to the track and the difference between the expected and observed interaction lengths ($\Delta\lambda \equiv \lambda_{\rm exp} - \lambda$), (2) the average ($\overline{m}$) and standard deviation (σ_m) of the multiplicity of hit strips per layer, (3) the χ^2 per degree of freedom of the IFR hit strips in the cluster with respect to the track extrapolation ($\chi^2_{\rm trk}$), and that with respect to a polynomial fit of the two-dimensional IFR cluster ($\chi^2_{\rm fit}$), and (4) the continuity of the hits in the IFR (T_c), which is defined as the fraction of IFR hit layers out of total number of layers from the first hit layer to the last one.

The difference between the expected and observed interaction length ($\Delta\lambda$) takes into account the history of RPC efficiencies, which reduces the dependence on the status of the IFR detector, and thus increases the overall efficiency. $\Delta\lambda$ for muons is distributed below 1, while for hadrons, the distribution extends beyond 3 without a clear peak. The interaction length for muons is generally larger than 2 and peaked around 4. For pions, this distribution is peaked around 2, depending on the momentum.

Five selection criteria are used in *BABAR* muon identification: *minimum ionizing, very loose, loose, tight* and *very tight*. The minimum ionizing selection only requires the energy released in the EMC to be less than 0.5 GeV. All other selections have the same or tighter cut on EMC energy and always require at least two IFR layers with hit strips. The selections relevant to this analysis are *loose* and *very tight*. The additional requirements for loose (very tight) selections are $E_{\rm EMC} < 0.5$ ($0.05 < E_{\rm EMC} < 0.4$) if applicable, $\Delta\lambda < 2$ (0.8), $\lambda > 2$ (2.2), $\chi^2_{\rm trk} < 7$ (5), $\chi^2_{\rm fit} < 4$ (3), $T_c > 0.2$ (0.34) if applicable, $\overline{m} < 10$ (8), and $\sigma_m < 6$ (4).

The efficiency for loose (very tight) selection is around 80% (50%) at a momentum of 1 GeV, and increases to higher than 85% (70%) for momentum higher than 2 GeV. The pion misidentification probability for loose (very tight) selection is about 8% (2%). The kaon misidentification probability is similar to the pion on average, but has higher dependence on momentum: it is smaller than pion misidentification at low momentum and larger at high momentum.

5.2.3 Kaon identification

The *BABAR* experiment uses a likelihood based kaon selection [68], which uses the ionization energy loss dE/dx in the SVT and the DCH, and the Cherenkov light emission angle θ_C with respect to the direction of the track entering the quartz bar of the DIRC.

The central dE/dx values for long-lived particles (e.g., e, μ, π, K, and p) in both the SVT and the DCH are parameterized based on the Bethe-Bloch functions as functions of momentum. The Bethe-Bloch parameters and the dE/dx resolution are calibrated by measuring the average charge deposited by a minimum-ionizing particle for each cell. The relative dE/dx error for the SVT is about 10% and for the DCH about 8%. The probability density functions for both detectors are modeled by a Gaussian distribution.

The likelihood for one of the five particle hypotheses (e, μ, π, K, and p) consists of two parts, which are assumed to be uncorrelated. The first part is a Gaussian probability based on the measured Cherenkov angle θ_C, the error on the measurement ($\sim$2.5 mrad), and the expected Cherenkov angle for each of the five hypotheses, based on the equation $\cos\theta_C = 1/(n\beta)$, where n is the index of reflection and $\beta = v/c$. This probability is normalized so that the sum of the five hypotheses equals one. If the momentum is below the threshold for certain particle hypotheses, their probability is assigned to be 1/5. The second part is a Poisson probability based on the number

of photons; i.e., the number of detected photons $N_{\rm sig}$, the number of background photons $N_{\rm bg}$ (fixed at 3) and the expected number of photons $N^i_{\rm exp}$ for particle type i. The Poisson probability is then $p^i = {\rm Poisson}(N_{\rm sig} + N_{\rm bg}, N^i_{\rm exp} + N_{\rm bg})$. Again, this probability is then normalized so that the sum of the five probabilities is 1. The total probability ℓ_i for each particle type is the product of these two parts. The same normalization scheme is applied for the total probabilities.

Five selection criteria are used in *BABAR* kaon identification: *not-a-pion*, *very loose*, *loose*, *tight*, and *very tight*. The selection criteria relevant to the analysis in this thesis is *not-a-pion* and *tight*. Whether or not a sub-detector is used for kaon identification depends on the particle's momentum. For the tight criterion, the used detectors are: SVT+DCH for $p < 0.6$ GeV, SVT+DCH+DIRC for $(0.6 < p < 0.7)$ GeV, and DIRC only for $p > 0.7$ GeV. For the not-a-pion criterion, the used detectors are: SVT+DCH for $p < 0.5$ GeV, DCH for $p < 0.6$ GeV, and DIRC for $p > 0.6$ GeV. The tight criterion accepts particles satisfying both $\ell_K > r_\pi \ell_\pi$ and $\ell_K > \ell_p$, where $r_\pi = 1$ for $(0.7 < p < 2.7)$ GeV, $r_\pi = 80$ for $p > 2.7$ GeV, and $r_\pi = 15$ for $(0.5 < p < 0.7)$ GeV. The not-a-pion criterion *rejects* particles satisfying $\ell_K < r_\pi \ell_\pi$ and $r_\pi \ell_\pi > \ell_p$, where $r_\pi = 0.1$ for $p \leq 0.5$ GeV and $r_\pi = 1.0$ for $p > 0.5$ GeV.

Chapter 6

Vertex and Decay Time Difference

This analysis relies on good vertex resolution to separate the decay points of two B mesons. The distance between two decay vertices in the z direction, Δz, to a good approximation, is proportional to the decay time difference of the two B mesons, Δt.

In this chapter, the vertexing algorithm and the measurement of Δz are briefly summarized in Sec. 6.1, issues about the conversion from Δz to Δt are discussed in Sec. 6.2.1, the behavior of Δt resolution is discussed in Sec. 6.3, and the measurement of the beam spot is described in Sec. 6.4.

6.1 Decay Vertices Measurement

The algorithm used to reconstruct decay vertex in *BABAR* is described in detail in [69]. This section highlights the most important aspects of vertexing that relate to the decay vertices separation for this analysis.

In this analysis, the decay chain, $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_\ell$, $D^{*+} \to D^0\pi^+$, and $D^0 \to K^-\pi^+$, $K^-\pi^+\pi^-\pi^+$, $K^-\pi^+\pi^0$, or $K_S\pi^+\pi^-$ is reconstructed (see Ch. 8). The vertex and momentum of the K_S candidate are calculated by fitting two charged tracks to a common vertex, and the result forms a virtual track for the D^0 reconstruction. The

π^0 candidate is form by combining two neutral clusters and is fit with a constraint so that the invariant mass of the π^0 candidate is equal to the true value [9]. The momentum of the π^0 candidate contributes to the calculation of the D^0 invariant mass, but does not contribute to the information for D^0 vertex. The daughter particles of the D^0 candidate are then fit to a common vertex, with a constraint so that the invariant mass of the D^0 candidate is equal to the world average [9]. The resulting momentum and position are combined with the slow pion and lepton candidate in a vertex fit that assumes the B meson decays within the beam colliding region (beam spot) on the plane transverse to the beam axis. The beam spot is assumed to be an ellipsoid with a Gaussian profile. As will be discussed in Sec. 6.4, the real beam spot size is about 1 cm in the z direction, 100 μm in the x direction, and 5 μm in the vertical (y) direction. The vertical size provides the most stringent constraint. However, to account for the flight length of the B meson in the direction transverse to the beam direction, the vertical size of the beam spot is artificially set to 30 μm. By including the beam spot information, the RMS of the distribution of the $D^*\ell$ vertex residual ($z_{\rm measured} - z_{\rm true}$) in Monte Carlo simulation improves from about 90 μm to about 77 μm. The mean of the residual distribution is consistent with zero (see Fig. 6.1).

The vertex of the second B meson (called $B_{\rm tag}$, see Ch. 7) is constructed from all good quality tracks in the event except the daughter particles of the D^* ℓ candidate. In order to reduce bias and tails due to long-lived particles, K_s and Λ^0 candidates are reconstructed by combining two charged tracks to a common vertex, and are used as inputs for the fit in place of their daughters. In addition, tracks that are consistent with photon conversions are excluded. The vertex is also constrained to the beam spot region in the transverse plane. To reduce the bias contributed by the charm decay products, the track with the largest vertex χ^2 contribution is removed if the contribution is greater than 6. The fitting procedure is repeated until no track

contributes more than 6 to the χ^2, or there is only one track left. (Because the beam spot information is used, it is possible to determine a vertex position from one track.) In this analysis, only the events that use two or more tracks in the $B_{\rm tag}$ vertex are retained. The RMS of the z-vertex residual distribution in simulation data ($\sim$155 μm) is much larger than that for $B_{\rm rec}$, and the mean of the distribution is significant shifted by $\sim$34 μm (see Fig. 6.1).

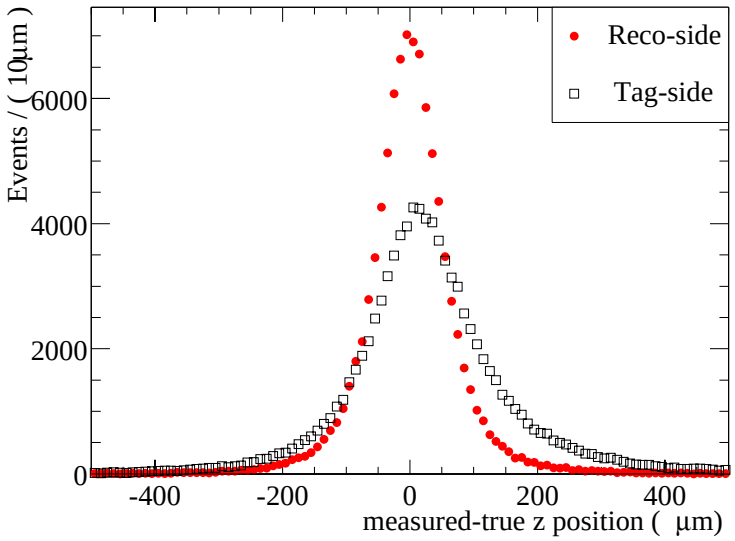

Figure 6.1: z-vertex residual distributions for $B_{\rm rec}$ and $B_{\rm tag}$ in Monte Carlo samples, where the $B_{\rm rec}$ is correctly reconstructed.

6.2 Decay Time Difference

6.2.1 Conversion from Δz to Δt

The decay time difference between B mesons can be calculated from the difference in decay positions along the *BABAR* z-axis with the following equation:

$$\Delta t = t_1 - t_2 = m_B \left(\frac{z_1}{p_1^z} - \frac{z_2}{p_2^z} \right) ,$$

where we define the $\Upsilon(4S)$ decay point to be at $z = 0$. The decay times t_1 and t_2 are defined in the respective B rest frames, which are different due to the small momentum of the B meson in the $\Upsilon(4S)$ frame, $\langle p_B^* \rangle = \frac{1}{2}\sqrt{s - 4m_{B^0}^2} \simeq 320\,\mathrm{MeV}$.[1] In this analysis, the direction of B mesons in the $\Upsilon(4S)$ frame is not known, therefore this small momentum is ignored and the approximation for decay time difference is

$$\Delta t = \frac{\Delta z}{\gamma \beta_z c} , \tag{6.1}$$

where γ is the boost factor for the $\Upsilon(4S)$ in the lab frame and β_z its velocity projected on the *BABAR* z-axis. In *BABAR*, $\beta_z \gamma \simeq 0.55$.[2]

Equation 6.1 ignores two effects. The most important is the finite momentum of the B meson in the $\Upsilon(4S)$ rest frame, mentioned in the previous paragraph and quantified below. Second, the $\Upsilon(4S)$ momentum is smeared due to the momentum spread of the colliding beams. This smearing is about 6 MeV, which corresponds to 0.1% smearing of $\beta\gamma$. The factor of β_z accounts for the fact that the direction of the colliding beams is tilted by 20 mrad with respect to the detector z-axis. The boost (momentum of the $\Upsilon(4S)$) along the z-axis is reduced by 1.2 MeV (0.02%) due to this

[1] We use values used in Monte Carlo simulation: $E_{e^+} = 3.11175\,\mathrm{GeV}$, $E_{e^-} = 8.99\,\mathrm{GeV}$, $\sqrt{s} = 10.5782\,\mathrm{GeV}$ and $m_{B^0} = 5.2794\,\mathrm{GeV}$.

[2] The value used in Monte Carlo simulation is $\beta_z \gamma = 0.555596$.

effect. There is also an effective boost of about 118 MeV in the x direction, which generates an azimuthal asymmetry, but has no direct impact on the estimation of decay time difference.

For a B meson that decays with polar angle θ^* with respect to the boost direction in the $\Upsilon(4S)$ frame, the distance Δz can be written as [69]

$$\Delta z = c\gamma\beta\gamma^*(t_1 - t_2) + c\gamma\gamma^*\beta^* \cos\theta^*(t_1 + t_2)\,, \tag{6.2}$$

where $\gamma^*\beta^*$ is the boost of the B meson in the $\Upsilon(4S)$ frame. If the acceptance does not depend on θ^*, the transformation Eq. 6.1 is only biased by a factor γ^*, which is approximately 1.002, since $\langle\cos\theta^*\rangle = 0$.

In general we don't know $t_1 + t_2$ event by event, but if the B decay angle θ^* is known (as is the case for fully reconstructed B's), the estimate of event-by-event Δt can be improved using the expectation value of $t_1 + t_2$. Since both t_1 and t_2 are positive, the minimum value of $t_1 + t_2$ is $|\Delta t|$. By integrating $t_1 + t_2$ from $|\Delta t|$ to infinity, we get $\langle t_1 + t_2\rangle = \tau_B + |\Delta t|$. Substituting this correction in Eq. 6.2, we get

$$\frac{\Delta z}{\gamma\beta\gamma^* c} = \Delta t + \frac{\beta^*\cos\theta^*}{\beta}(\tau_B + |\Delta t|)\,.$$

In this analysis, we do not measure the B decay direction. The effect of the second term in Eq. 6.2 is to smear the exponential distribution of Δz. The RMS of this term is

$$c\gamma\gamma^*\beta^*\sqrt{\langle\cos^2\theta^*\rangle\,\langle(t_1 + t_2)^2\rangle} \simeq 21\cdot\sqrt{6}\,\tau_B\sqrt{\langle\cos^2\theta^*\rangle}\ (\mu\mathrm{m}) \tag{6.3}$$

where τ_{B^0} is in ps, and $\beta^* \simeq 0.0606$, $\gamma \simeq 1.144$ and the expectation value of $\langle(t_1 + t_2)^2\rangle = 6\tau_B^2$ are used. The angular distribution for $\Upsilon(4S) \to B\overline{B}$ is given by $(1 - \cos^2\theta^*)d(\cos\theta^*)$, which gives $\langle\cos^2\theta^*\rangle = 1/5$. The RMS of the Δz residual

due to the boost approximation is then about 35 μm, which is about 0.21 ps for Δt. This result is for integration over the entire (t_1, t_2) space. For a given Δt, the expectation value of $(t_1 + t_2)^2$ depends on $|\Delta t|$. We can integrate $(t_1 + t_2)^2$ from $|\Delta t|$ to infinity and get

$$\left\langle (t_1 + t_2)^2 \right\rangle |_{\Delta t} = 2\tau_B^2 + 2\tau_B |\Delta t| + |\Delta t|^2 \,.$$

If we further integrate $|\Delta t|$ from zero to infinity, we get the factor of $6\tau_B^2$ that appears in Eq. 6.3. Therefore the smearing of Δt introduced by using the approximation $\Delta z/(\gamma\beta\gamma^* c)$ can be written as a function of $|\Delta t|$:

$$\frac{\beta^*}{\sqrt{5}\beta}\sqrt{\tau_B^2 + (\tau_B + |\Delta t|)^2} \simeq 0.0557\sqrt{\tau_B^2 + (\tau_B + |\Delta t|)^2}\,. \tag{6.4}$$

We can demonstrate this effect by plotting the RMS of $\Delta z_{\rm truth}/(\gamma\beta\gamma^* c) - \Delta t_{\rm truth}$ in bins of $|\Delta t_{\rm truth}|$. Figure 6.2 shows this plot for 80k signal simulation events that pass all our final signal cuts and 100k events directly from the event generator, and a fit to a function of the form given in Eq. 6.4: $p_0 \cdot \sqrt{p_1^2 + (p_1 + |\Delta t|)^2}$.

This smearing effect is visible at the level of reconstructed Δt. The top plot in Fig. 6.3 shows the RMS of the pull of reconstructed Δt in bins of true Δt, where the pull is defined as the residual $\delta\Delta t \equiv \Delta t_{\rm meas} - \Delta t_{\rm true}$ divided by the calculated error, $\sigma_{\Delta t}$ (see Sec. 6.2.2). The dependence of the RMS on $\Delta t_{\rm true}$ is clearly seen. The middle plot shows the same sort of distribution except that $\Delta t_{\rm meas}$ is replaced by $\Delta z_{\rm true}/\gamma\beta\gamma^* c$ in the calculation of the RMS of the pull. The bottom plot is the bin-by-bin difference in quadrature between the top two plots. The fact that the dependence of the pull RMS on $\Delta t_{\rm true}$ almost disappears after removing the effect due to boost in the conversion from Δz to Δt suggests that the boost approximation is the major source of the resolution dependence on $\Delta t_{\rm true}$. Since this dependence can be fully understood, in principle we can modify our model to accommodate this

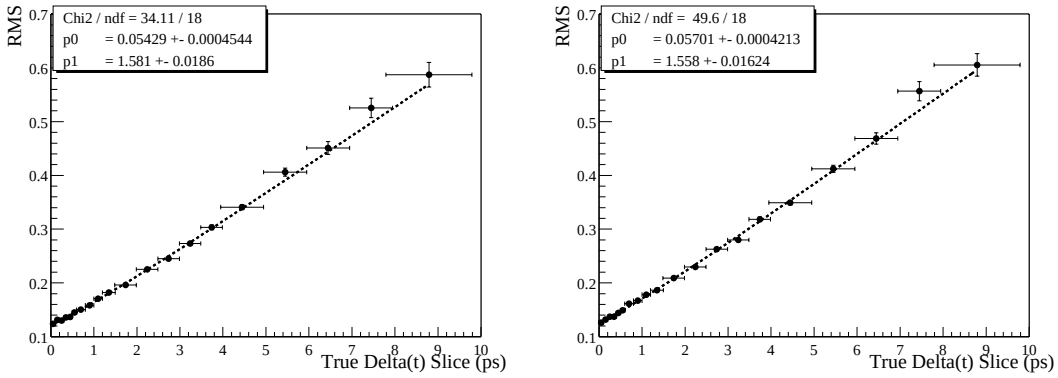

Figure 6.2: RMS of the distribution of residuals $\Delta z_{\rm truth}/(\gamma\beta\gamma^* c) - \Delta t_{\rm truth}$ in bins of $|\Delta t_{\rm truth}|$ for 80k signal simulation events that pass all event selection criteria (left) and for 100k events from the event generator (right), and a fit to a function $p_0 \cdot \sqrt{p_1^2 + (p_1 + |\Delta t|)^2}$. The $\Delta z \to \Delta t$ conversion factor used here is 166.87μm/ps.

effect.

6.2.2 Calculation of error on Δt

The calculated uncertainties on the vertex position of the $B^0 \to D^{*-}\ell^+\nu_\ell$ candidate and the tagging B account for uncertainties on the track parameters due to SVT and DCH hit resolution and multiple scattering, our knowledge of the beam spot size, and the average B flight length in the vertical direction. The calculated uncertainty does not account for errors in pattern recognition in tracking or errors in associating tracks with the B vertex. The calculated uncertainties will also be inaccurate if the assumptions for the amount of material in the detector or the beam spot size or position are incorrect. The errors on the two vertices directly translate to the error ($\sigma_{\Delta t}$) on Δt. The distribution of $\sigma_{\Delta t}$ for Monte Carlo events is shown in the left plots in Fig. 6.4. We use parameters in the resolution model, measured with data, to account for uncertainties and biases introduced by these effects.

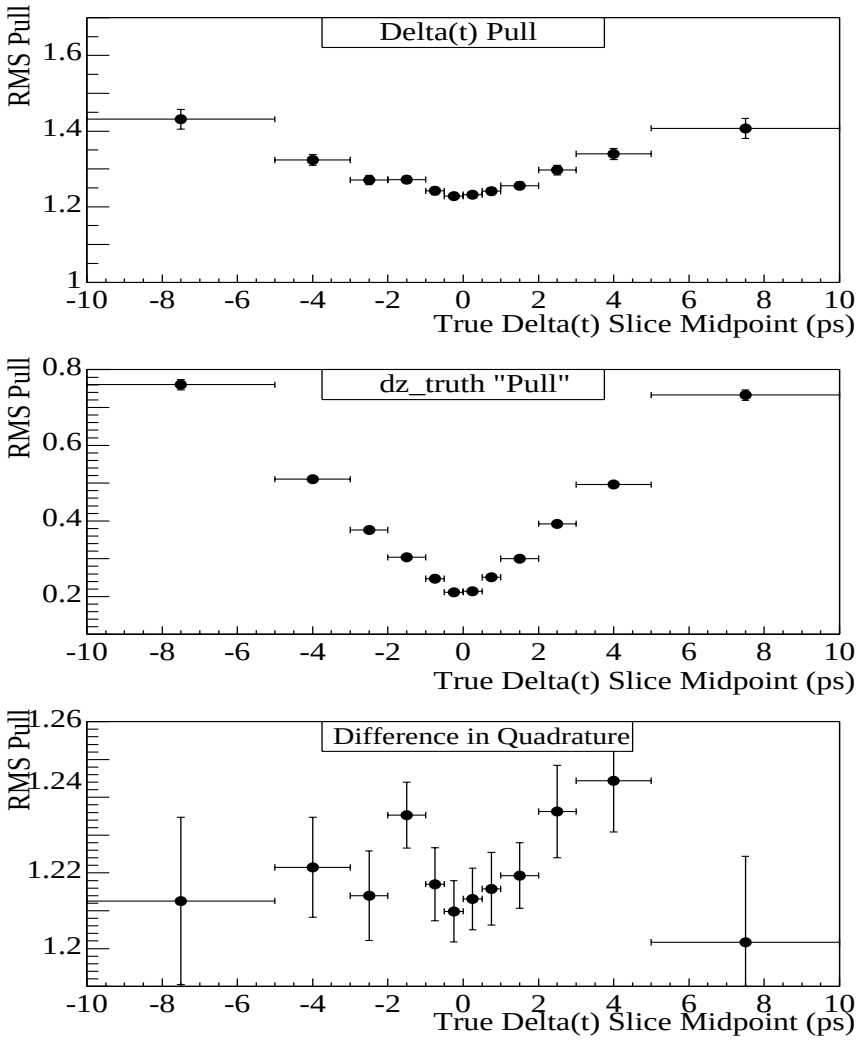

Figure 6.3: Top: the RMS of the pull ($\delta\Delta t/\sigma_{\Delta t}$) in bins of $\Delta t_{\rm true}$ for Monte Carlo simulation. Middle: same plot with Δt replaced by $\Delta z_{\rm true}/\gamma\beta\gamma^* c$ in the calculation of the RMS of the pull. Bottom: bin-by-bin difference in quadrature between the top two histograms. Note the factor of 10 difference in the vertical scales of the top and bottom histograms.

6.3 Decay-Time Difference Resolution

The behavior of the Δt residual ($\delta\Delta t \equiv \Delta t_{\rm meas} - \Delta t_{\rm true}$) will be modeled for the final fit (see Sec. 10.1.1). In this section, the residual of measured decay time difference

Δt and its correlation with the calculated error, $\sigma_{\Delta t}$ are studied.

The left-hand plot in Fig. 6.4 shows the distribution of $\sigma_{\Delta t}$ values observed in simulation. If the calculated errors are accurate and the Δt reconstruction is unbiased, then the distribution of pulls $\delta\Delta t/\sigma_{\Delta t}$ is a unit Gaussian with a mean equal zero, and the resulting resolution function has no free parameters:

$$\mathcal{R}(\delta\Delta t, \sigma_{\Delta t}) = G(\delta\Delta t;\, 0, \sigma_{\Delta t}) \;,$$

where we define the Gaussian function

$$G(x;\, x_0, \sigma) \equiv \frac{1}{\sqrt{2\pi}\,\sigma} \cdot \exp\!\left(-(x - x_0)^2/(2\sigma)^2\right) . \tag{6.5}$$

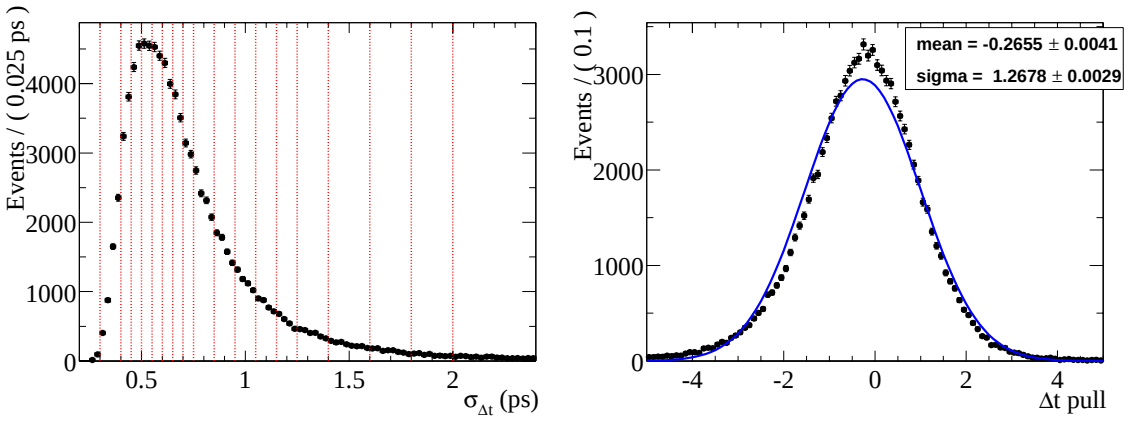

Figure 6.4: Distribution of calculated vertex errors $\sigma_{\Delta t}$ (left-hand side) and pulls (right-hand side) obtained from correctly reconstructed signal events in the simulation sample. Vertical dashed lines in the $\sigma_{\Delta t}$ distribution indicate the slices used to select subsamples with approximately equal calculated errors for further study. The solid curve superimposed on the pull distribution is the result of a single Gaussian fit.

The right-hand plot in Fig. 6.4 shows that the pull distribution observed in signal

simulation events is approximately Gaussian, but is biased towards negative residuals and has an RMS of about 1.3, which indicates that the calculated errors are underestimated on average by about 30%.

In order to study these deviations from the ideal case in more detail, we divide the sample of correctly reconstructed signal Monte Carlo events into slices of $\sigma_{\Delta t}$. Fig. 6.5 shows the mean and RMS calculated directly from the residuals in each slice, as a function of each slice's midpoint $\sigma_{\Delta t}$ value. The RMS of the residuals scales linearly with the calculated error as expected. It is clear from the plot (b) in Fig. 6.5 that the bias also scales with the calculated error.

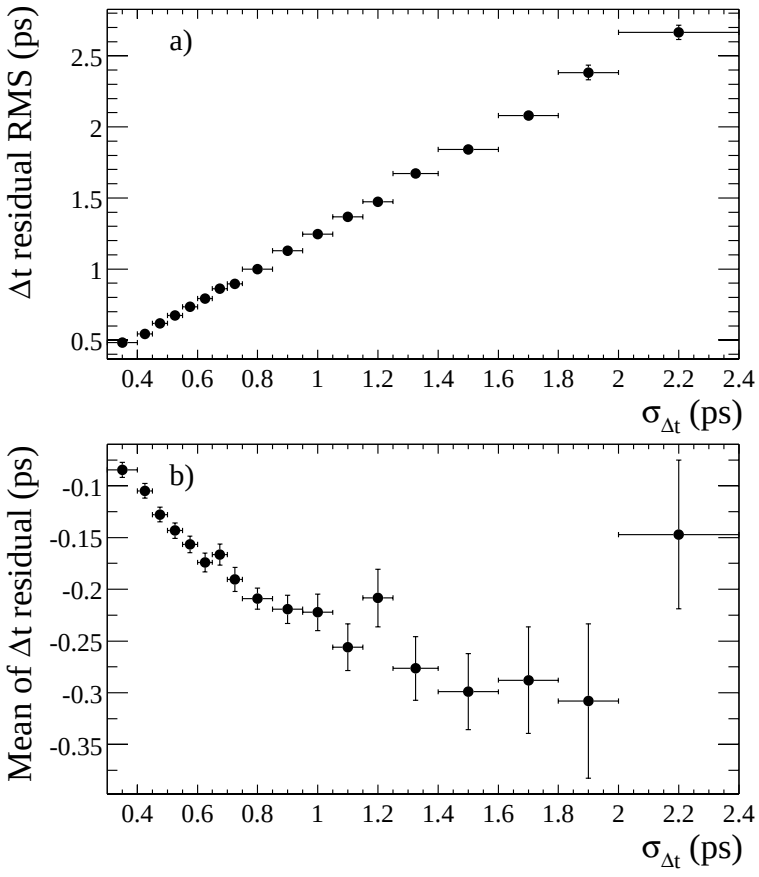

Figure 6.5: (a) RMS and (b) mean of the residual distributions obtained from correctly reconstructed signal Monte Carlo simulation, in slices of $\sigma_{\Delta t}$.

The explanation for this is the fact that the vertex error ellipse for the D decay products is oriented with its major axis along the D flight direction, leading to a correlation between the D flight direction and the per-event D^0 error. D mesons that have a flight direction perpendicular to the z direction in the laboratory frame will have the best resolution and will introduce the least bias in a measurement of the z position of the parent B meson. D mesons that travel forward in the laboratory will have poorer resolution and will introduce a larger bias in the measurement of $z_{\rm tag}$. This effect is illustrated in Fig. 6.6, and is first demonstrated in [70].

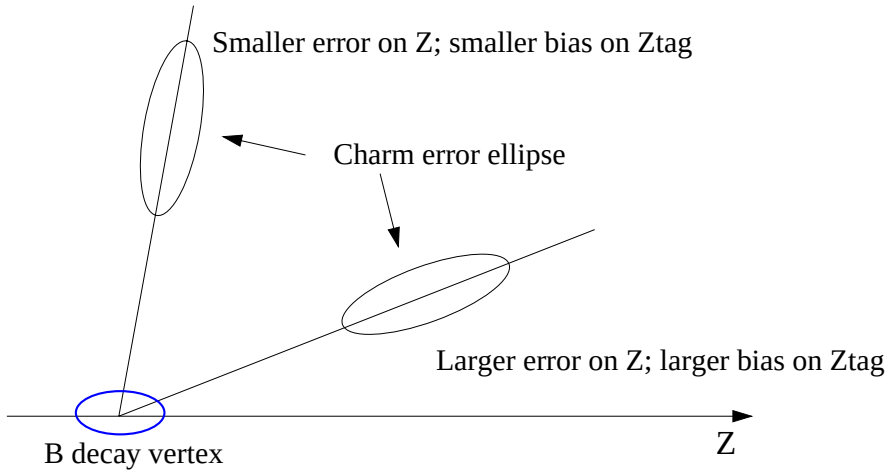

Figure 6.6: Schematic graph that illustrate the correlation between vertex bias and calculated error on tag-side z-vertex through charm flight direction. The measured tag-side vertex in the z direction is somewhat between the B vertex and the charm vertex if tracks from both vertices are used. The projected vertex error on the z-axis is smaller when the charm flight polar angle is closer to 90°

To verify that this is the correct explanation for the scaling of the bias with per-event error, we study the correlations among D meson direction, $\sigma_{\Delta t}$ and Δt bias with signal Monte Carlo simulation samples.

From the signal Monte Carlo sample, correctly reconstructed candidates are selected according to Monte Carlo true information. We trace the decay tree for all

tracks used for determining the tag-side vertex and look for long-lived charm (D^+, D^0, ${D_s}^+$ and Λ_c) and τ. There can be up to three such long-lived particles in the decay tree. For the simplicity of this study, for each event, only the information of the first charm-tau found in the event is recorded. Events with no charm-tau are also selected. Since the residual and error on Δt are dominated by the tag-side, we only focus on the tag-side.

The correlation between the mean of Δt residual and $\sigma_{\Delta t}$ as seen in Fig. 6.5 can be reproduced at the tag-side. Figure 6.7 shows the mean tag-side vertex residual (scaled by $\gamma\beta c$) in slices of $\sigma_{\Delta t}$, for events with one and zero charm-tau particle. Clearly, the correlation in question is due to charm-tau on tag-side, since the events without charm-tau have no bias.

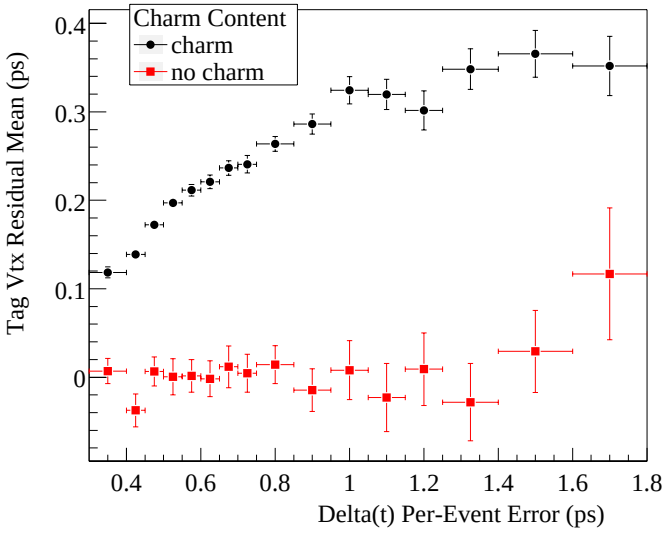

Figure 6.7: Mean residual of tag-side vertex (scaled by $\gamma\beta c$ of Υ(4S) in lab frame) in slices of $\sigma_{\Delta t}$ for simulation events in which at least one track used for tag-vertexing is produced from charm-tau (black) and for events in which no track used for tag-vertexing is produced from charm-tau (red)

As shown in Fig. 6.8, the mean residual and the mean $\sigma_{\Delta t}$ are highly correlated with charm-tau flight direction. However, the minimum of $\sigma_{\Delta t}$ is not at $\cos\theta_D = 0$

as one would have expected by charm vertex topology. The fact that the minimum shifts towards positive $\cos\theta_D$ is largely due to the correlation between the tracking error and the momentum of that track. The charm-tau momentum in the lab frame is highly correlated with its flight direction due to the boost of the center-of-mass frame. In general the higher momentum a track has, the smaller its error is, therefore, the error ellipse of the charm-tau vertex is smaller for high momentum (forward direction) charm-tau, since they tend to produce higher momentum tracks.

These two effects (topology and momentum correlation) compete with each other. But since most events are populated in the forward direction, the topology effect dominates most of the $\sigma_{\Delta t}$ range. At the region of large $\sigma_{\Delta t}$, however, the sign of the correlation between the mean of Δt residual and $\sigma_{\Delta t}$ changes, as shown in plot (b) in Fig. 6.5 for $\sigma_{\Delta t} > 1.8$ ps. To maintain the sign of this correlation so that the resolution model discussed in the next paragraphs is valid for the full range of $\sigma_{\Delta t}$ in the event sample, the events with $\sigma_{\Delta t} > 1.8$ is removed. The events that are removed represent only about 1.5 % of the sample and contribute even less to the statistical sensitivity, since their vertices are poorly measured.

6.4 Beam Spot Determination

The nominal luminous region (beam spot) of e^+e^- collision in PEP-II is a thin flat ribbon-like shape. It can be approximated by a ellipsoid with a Gaussian distribution along three axes, with $\sigma_z \simeq 1$ cm, $\sigma_x \simeq 100$ μm, and $\sigma_y \simeq 5$ μm. The vertical size, σ_y, is much smaller than the tracking and vertexing resolution in *BABAR*. Therefore, the knowledge of the y position of the beam spot provides a constraint that can improve the sensitivity of many measurements, such as the vertex resolution and the slow pion direction.

The average beam spot position and apparent size are measured on a run-by-run

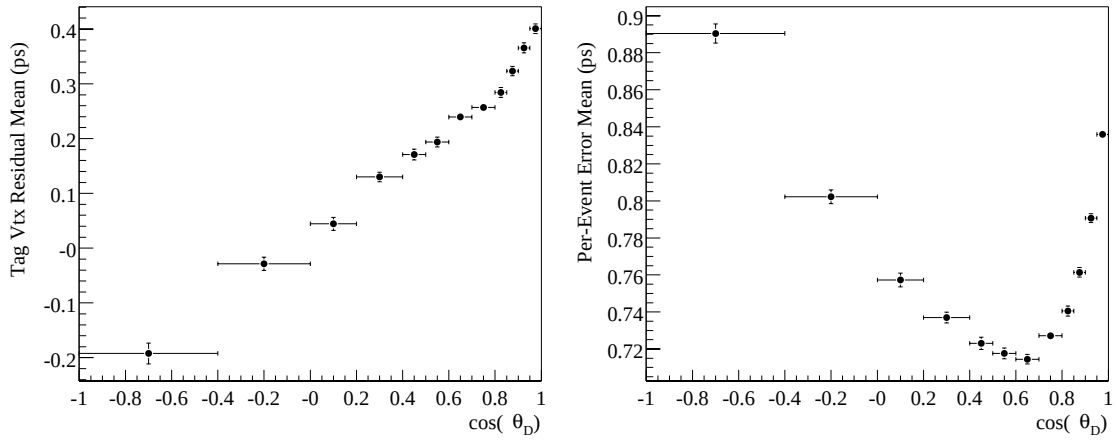

Figure 6.8: Tag-side residual (left) and $\sigma_{\Delta t}$ (right) in slices of $\cos\theta_D$ where θ_D is the polar angle of charm-tau (from which at least one track used for tag-vertexing is produced from) for signal simulation events.

basis. The length of each run is typically two hours. Two different methods are used to measure the beam spot parameters. They are described in the following paragraphs.

Measure beam spot parameters with two-prong tracks

This method uses the high momentum tracks from two-prong events (e^+e^- and $\mu^+\mu^-$ events). The distance of closest approach (doca) of each track with respect to the *BABAR* origin on the transverse (x-y) plane is calculated. The doca can be expressed as doca $= -x\sin\phi + y\cos\phi$, where (x, y) is the coordinate of a point on the track, and ϕ is the azimuthal angle. If all tracks originate from the same point (x_0, y_0), doca is a sinusoidal function of ϕ. The doca-ϕ distribution is shown in Fig. 6.9 for a typical run, where the doca for horizontal tracks ($\phi = 0$) is equal to y_0, for vertical tracks ($\phi = \pi/2$) equal to x_0, and spread of the distribution represents the apparent size of the beam spot. Therefore, one can fit the doca-ϕ distribution to the sinusoidal function to extract the centroid and the apparent size of the beam spot. This method

can be expanded to include the z information to account for the $\simeq -20$ mrad tilt angle of beam spot on z-x plane. The z information can be calculated from the interception of two lines projected by the two tracks on the plane that is defined by the two direction vectors at their points of closest approach.

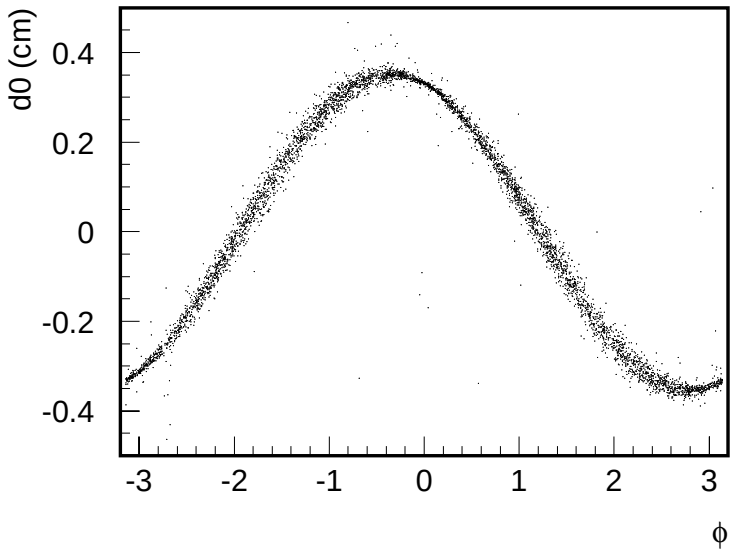

Figure 6.9: The doca-ϕ distribution for two-prong events in a typical run.

The beam spot determination is implemented in the prompt-reconstruction rolling calibration. The prompt-reconstruction is running on multiple computers in parallel and the accumulated data on all computers are merged and calculated in a "*finalize*" stage. The structure of the data that hold quantities used for calibration does not allow the accumulation of data with variable sizes (e.g., lists). Therefore accumulating all information of the selected tracks, iterating over all tracks, and performing a fit is not possible under this framework. The solution to this problem is to linearize the question; that is, to make the quantities needed in the finalize stage be summations of some quantities for individual tracks. Information of individual track is not recorded, but only the summation is. This way, the size of the accumulated

information is fixed.

A likelihood function for each event is constructed as a nine-parameter function, assuming the beam spot is a ellipsoid with a Gaussian profile,

$$\mathcal{L} = \sum_{k=1}^{2} \left[b + \frac{1-b}{\sqrt{2\pi}\sigma_d^k} \exp\left(-\frac{(d_k - d_0^k)^2}{2(\sigma_d^k)^2} \right) \right] \frac{1}{\sqrt{2\pi}\sigma_z} \exp\left(-\frac{(z_v^k - z_0)^2}{2\sigma_z^2} \right). \qquad (6.6)$$

where the sum is over the two tracks and

$$\begin{aligned} d_k &= -x_{poca}^k \sin\phi_k + y_{poca}^k \cos\phi_k, \\ (\sigma_d^k)^2 &= \sigma_x^2 \sin^2(\phi_k - \phi_{xy}) + \sigma_y^2 \cos^2(\phi_k - \phi_{xy}), \\ d_0^k &= [y_0 + m_{zy}(z_v^k - z_0)] \cos\phi_k - [x_0 + m_{zx}(z_v^k - z_0)] \sin\phi_k. \end{aligned}$$

with ϕ_k the track direction and z_v^k the z vertex. The parameters to be fit are x_0, y_0, z_0, σ_x, σ_y, σ_z, the beam spot slopes on z-x and z-y planes, m_{zx}, m_{zy}, and the beam spot tilt angle on x-y plane, ϕ_{xy}. The σ_i are measured along the beam axis that is close to the i-th axis of *BABAR* coordinates, not the projection on the *BABAR* axis. The uniform background term b is kept constant at the value 0.01 in the fit.

The sum of $\log\mathcal{L}$ over all selected events is Taylor-expanded to the second order. The maximum can be solved by setting the first derivatives to zero, if the function is a purely quadratic function, and the errors can be estimated from the second derivatives. This method satisfies the prompt reconstruction restriction, but needs the initial parameters to be close to the true values.

The measured σ_y is about 50 μm for e^+e^- events and about 25 μm for $\mu^+\mu^-$ events, which are dominated by tracking resolution. Both are well above the true σ_y. The statistical error on the mean of the y coordinate is of the order of 1 μm, since thousands of tracks contribute to the beam spot calculation. However, the vertical position of the beam can move up to 10 μm/hour during data taking due to machine

operation and diurnal motion correlated with temperature. The true collision point of an event can be above or below the calculated mean position by more than a few μm.

The the beam spot measurements as function of time in the year 2000 are shown in Fig. 6.10–6.13, including mean positions, error on the means, measured sizes and tilt angles.

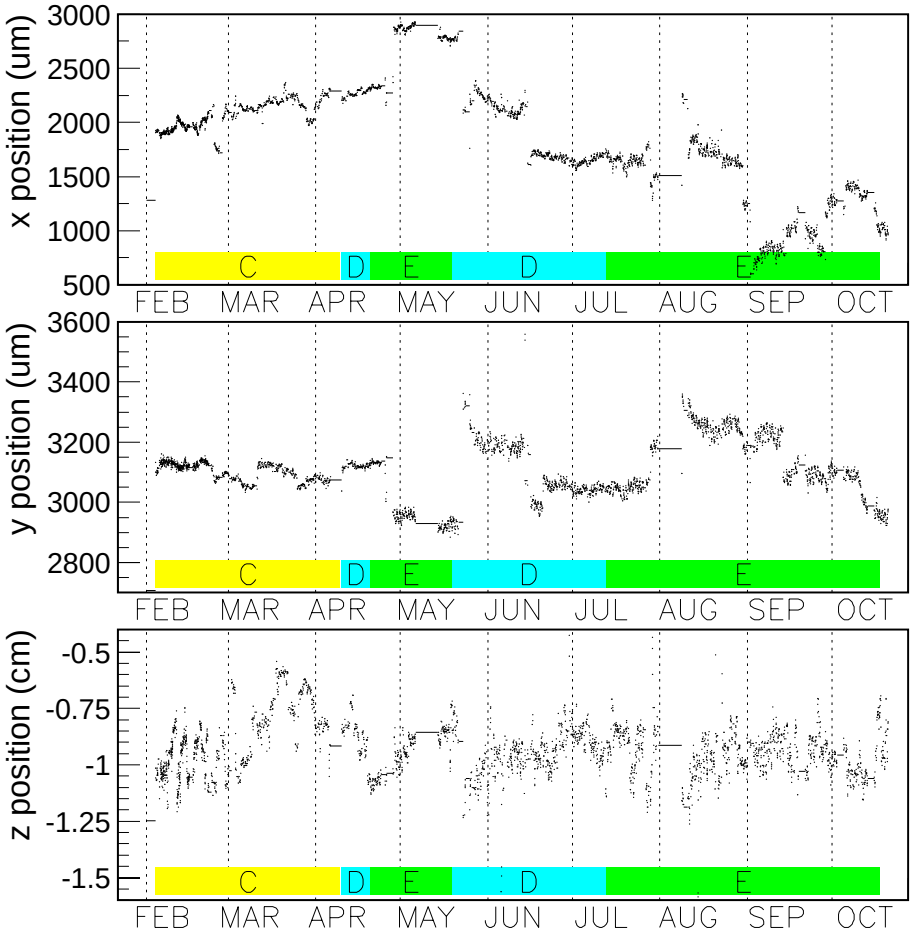

Figure 6.10: Mean position of the beam spot in year 2000.

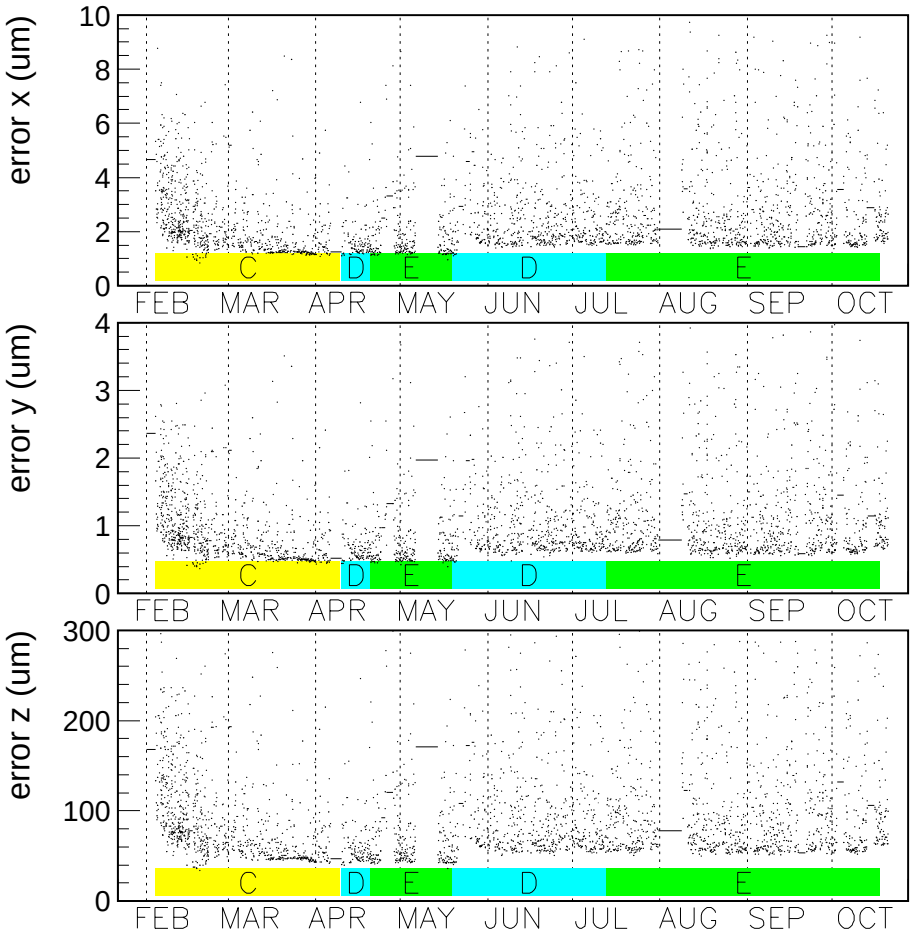

Figure 6.11: Error on the mean position of the beam spot in year 2000.

Measure beam spot parameters with event vertices

Another method is to simply to select a two-prong or hadronic events, perform a vertexing algorithm, and accumulate the first and second moment of x, y, and z coordinates of each vertex. This method is simpler than the method described in the previous paragraphs, but it depends on the vertexing algorithm and has worse resolution, especially for hadronic events, which can be larger than 70 μm.

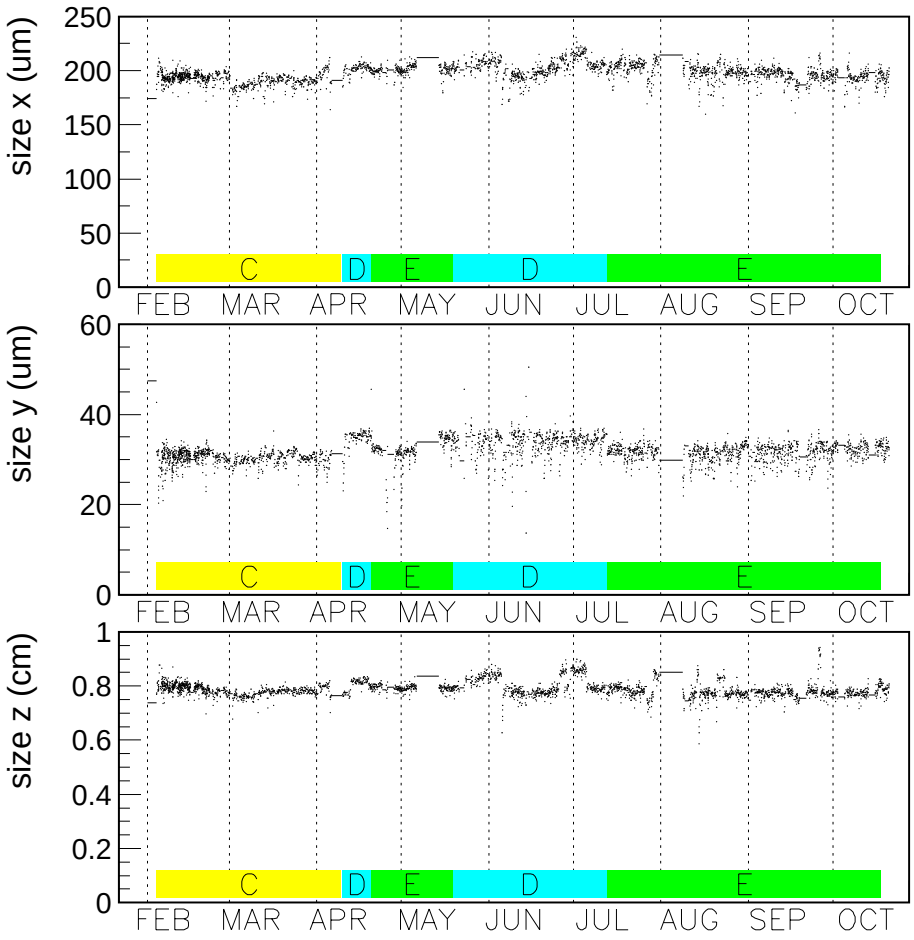

Figure 6.12: Measured size of the beam spot in year 2000.

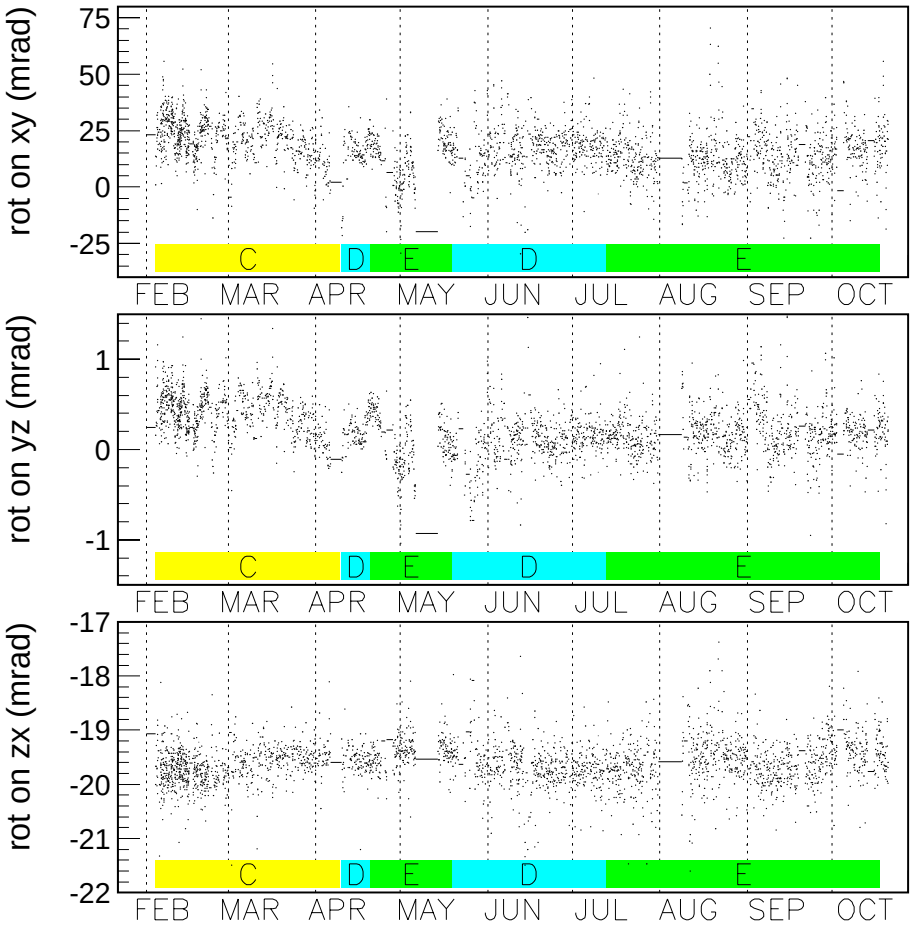

Figure 6.13: Tilt angle of the beam spot in year 2000.

Chapter 7

Flavor Tagging

In this analysis, one B^0 from $\Upsilon(4S)$ decay is reconstructed exclusively (except for the neutrino). After all tracks originating from this reconstructed B^0 are removed, for signal events, the remaining charged tracks are from the other B meson. In this chapter, the technique that identifies (tags) the flavor of the second B meson using these remaining tracks at the time of its decay is described. High efficiency and low mistag rate of tagging techniques are essential for $B^0\overline{B}^0$ mixing measurements, because the statistical uncertainty for mixing measurement is roughly proportional to $1/\sqrt{Q} \equiv 1/\sqrt{\epsilon(1-2\omega)^2}$, where ϵ is the tagging efficiency; i.e., the probability that a certain tagging technique can be used to separate B^0 decays from $\overline{B}^0$, and ω is the mistag rate; i.e., the probability that this technique makes a incorrect decision. The quantity Q is often called quality factor or effective tagging efficiency.

In *BABAR*, five different types of flavor tag, or tagging categories are used. The first two tagging categories rely on the presence of a prompt lepton, or one or more charged kaons, in the event. They are described in Sec. 7.1.1. The other three tagging categories exploit a variety of inputs with a neural-network algorithm, which is described in Sec. 7.1.2. Finally, the correlation between flavor tagging and decay vertexing is discussed in Sec. 7.2.

7.1 Algorithm

7.1.1 Single variable-based tagging — `lepton` tag and `kaon` tag

The `lepton` and `kaon` tagging categories use the correlation between the flavor of the decaying b quark and the charge of a primary lepton from a semileptonic decay or the charge of a kaon from the decay chain $b \to c \to s$.

The `lepton` category uses both electrons and muons, which are required to satisfy the "very tight" and "tight" selections, respectively (see Sec. 5.2). A minimum requirement of 1.0 (1.1) GeV on the electron (muon) center-of-mass momentum is applied to reduce the contamination from softer, opposite-sign lepton coming from cascade semileptonic decays of charm mesons. In each event, the electron or muon with the highest momentum in the center-of-mass frame is used for flavor tagging. For a tiny fraction of events, where both electron and muon are present, the electron is used for flavor tagging. The existence of a positively (negatively) charged primary lepton indicates a B^0 ($\overline{B}^0$) decay.

The kaon content in an event is evaluated from the sum of the charges of the kaons identified by a neural network algorithm, called `K` subnet, which is described in Sec. 7.1.2. This kaon identification algorithm has be optimized to maximize the quality factor, Q. A positive (negative) charged kaon is more likely to come from a B^0 ($\overline{B}^0$) decay. On average there are 0.8 charged kaons per B decay, and roughly 15% of these have the wrong sign. The wrong-sign kaons are primarily from B decays to a charm-anti-charm pair of mesons, or from right-sign D^+ decays to a wrong-sign kaon (see Sec.7.2).

An event is assigned to the `lepton` category if a high momentum lepton is identified, and no identified kaon has the opposite sign. Otherwise, an event is assign

to the `kaon` category if no high momentum lepton is identified and the sum of the charges of the identified kaons is non-zero. The flavor of the B is assigned according to the charge of the lepton, or the sum of the charges of kaons.

The `lepton` tagging category has the lowest mistag probability among the five tagging categories. The `kaon` tagging category has a higher mistag rate than the `lepton` category, but the efficiency is also higher. About 45% of events belong to these two categories. The remaining events are passed to the neural-network based algorithm.

7.1.2 Neural network tagging

A multi-variate analysis based on a neural network is carried out for the events that do not belong to `lepton` and `kaon` tagging categories. Three different neural networks, called "subnet", have been trained to identify primary leptons (`L` subnet), kaons (`K` subnet) and soft pions from D^* decays (`SoftPi` subnet). More details are described in [71].

The `L` subnet uses the output of the electron and muon identification algorithm, the center-of-mass momentum of the input track, and two kinematic variables that use the information of missing momentum and are sensitive to the primary lepton from semileptonic decays. The `K` subnet uses the input track momentum in the laboratory frame, together with the three relative likelihoods $\mathcal{L}_\mathcal{K}/(\mathcal{L}_\pi + \mathcal{L}_\mathcal{K})$ for the SVT, the DCH and the DIRC. The `SoftPi` subnet uses the center-of-mass momentum of the input track, the angle of the input track with the thrust axis calculated with all charged tracks and neutral clusters in the $B_{\rm tag}$, and the center-of-mass momentum of the minimum momentum track. The direction of a soft π from a D^* decay is likely to be aligned with the D^* direction, which is correlated with the trust angle. Each of these subnets is applied to each charged track in the event.

The final neural network that is trained to distinguish between B^0 and $\overline{B}^0$ uses

the following inputs: the maximal values over each track in the event of the `L` and `SoftPi` subnet outputs and the highest and the second-highest values of the `K` subnet output, all subnet output values multiplied by the charge of the corresponding input track; the center-of-mass momentum of the maximum momentum track multiplied by its charge; and the number of tracks with significant impact parameters, which is an indicator of the presence of K_s mesons.

The output from the final neural network, x_{NT}, is mapped onto the interval $(-1, 1)$. The assigned flavor is B^0 if the x_{NT} is positive, and $\overline{B}^0$ otherwise. Events with $|x_{NT}| > 0.5$ are assigned to `NT1` category, events with $0.2 < |x_{NT}| < 0.5$ are assigned to `NT2` category, and the remaining events are assigned to `NT3` category. The distribution of the neural network output can be seen in Fig. 7.1. About 30% of the total events are in `NT3` category which carries very little tagging power; i.e., their mistag probability is close to 50%, but they increase the sensitivity to the lifetime measurement

7.2 Mistag and Vertexing Correlation

A strong correlation between the mistag fraction and calculated error on B decay time difference, $\sigma_{\Delta t}$ for kaon tagging category (`KTag`) is found. Figure 7.2 shows the mistag fractions for five different tagging categories in slices of $\sigma_{\Delta t}$ in Monte Carlo samples. The mistag fraction for `KTag` increases almost linearly for small $\sigma_{\Delta t}$ and levels off at larger $\sigma_{\Delta t}$, while mistag fractions for other categories are basically independent of $\sigma_{\Delta t}$.

Apart from detector effects (e.g. kaon mis-id), wrong-sign kaons can originate from many physics processes, e.g., double charm decays, fragmentation, wrong-sign kaon from singly/doubly Cabibbo-suppressed charm decays.

In this section I study the sources of this ω–$\sigma_{\Delta t}$ correlation using $B^0 \to D^* \ell \nu$

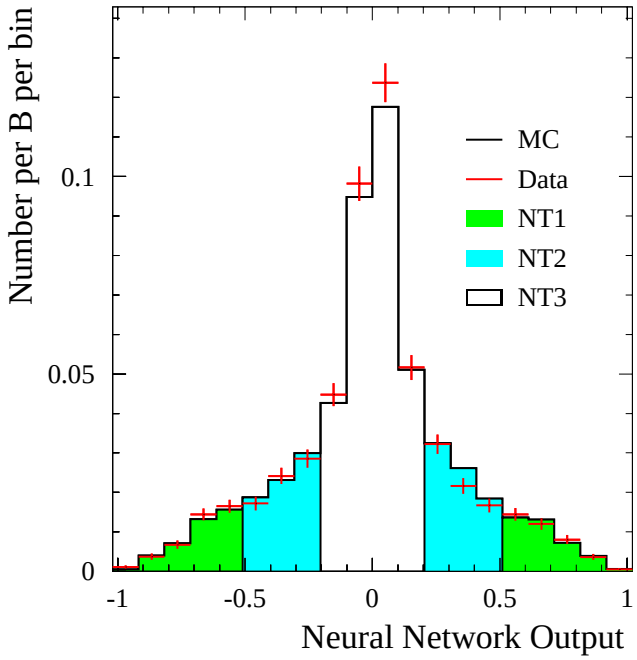

Figure 7.1: The distribution of neural network output for the events that are not assigned to `lepton` and `kaon` categories. The points are from an analysis of fully reconstructed hadronic B events [72], and the histogram from simulation.

signal Monte Carlo sample by studying the tag-side B^0 decays in the events where the $B_{\rm rec}$ candidate is correctly reconstructed.

7.2.1 Transverse momenta of used tracks

It was first confirmed by Art Snyder [73] that the ω–$\sigma_{\Delta t}$ correlation is (almost) solely due to the correlations between ω, $\sigma_{\Delta t}$ and transverse momenta of used charged tracks. Here I demonstrate the effect with $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_\ell$ simulation sample.

Roughly speaking the error of the tracking parameters is proportional to $1/p$, where p is the magnitude of the momentum. A single track i contributes to the error

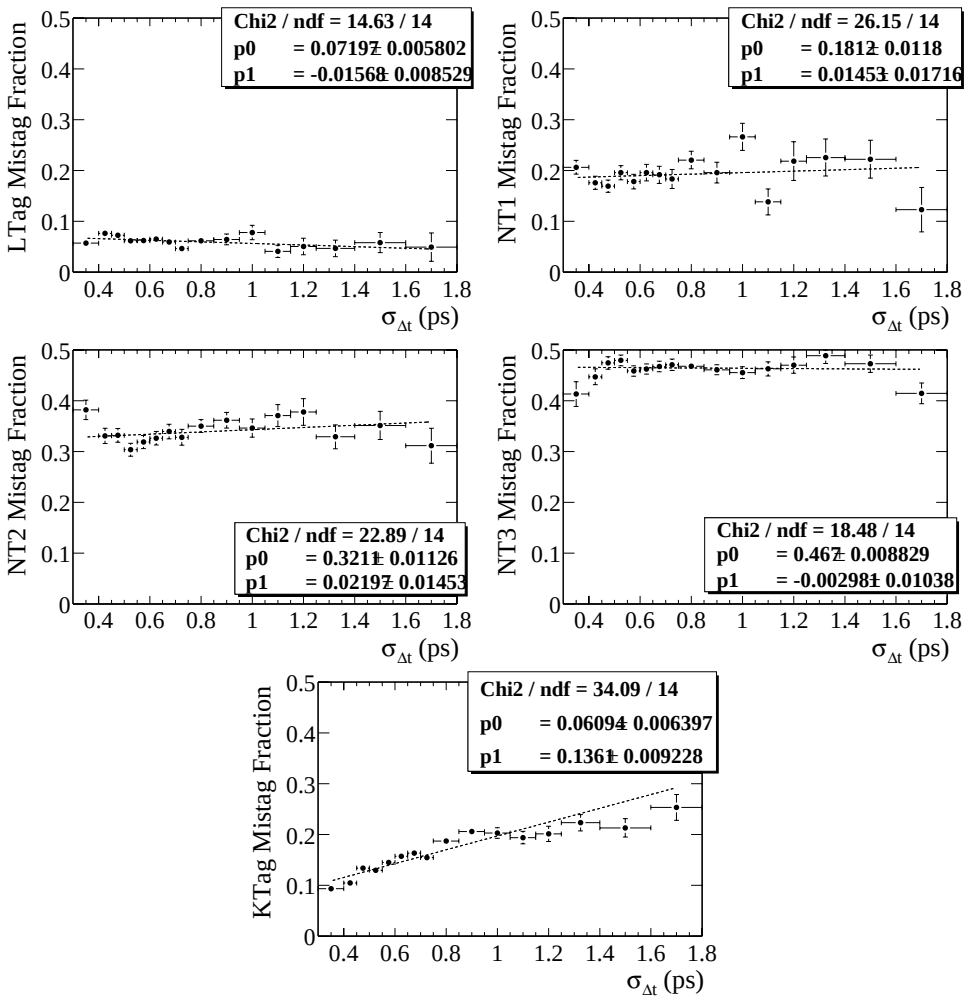

Figure 7.2: Mistag fractions in slices of $\sigma_{\Delta t}$ for five tagging categories for true signal events in $B^0 \to D^*\ell\nu$ signal Monte Carlo sample. The values are calculated by counting the numbers of incorrectly tagged events by comparing the observed and true mixing status. The plots are fitted with a straight line. Only `KTag` shows a significant slope.

of the vertex position, projected on the z-axis is,

$$\sigma_z^i \propto 1/(p^i \sin\theta^i) = 1/p_t^i \,.$$

The error on a vertex in the z direction, after summing contributions from all used

tracks, is then given by

$$\frac{1}{(\sigma_z)^2} = \sum_i \frac{1}{(\sigma_z^i)^2} \propto \sum_i (p_t^i)^2 \,.$$

So $\sigma_{\Delta t}$ (dominated by $\sigma_{z_{\rm tag}}$) should be a linear function of $1/\sqrt{\sum_i (p_t^i)^2}$. The left plot in Fig. 7.3 shows this relation along with a linear fit. The mistag fraction is also strongly correlated with $1/\sqrt{\sum_i (p_t^i)^2}$, as shown in the right plot. To understand whether the correlation between the mistag probability and $\sigma_{\Delta t}$ is due to p_t correlation, or due to other effect, we can "predict" the mistag probability using p_t and plot the ratio of measured to predicted mistag probability as a function of $\sigma_{\Delta t}$. According to a linear fit to mistag versus $1/\sqrt{\sum_i (p_t^i)^2}$, the expected mistag is $\omega' = 0.151/\sqrt{\sum_i (p_t^i)^2} + 0.009$. The relation ω/ω' versus $\sigma_{\Delta t}$ is shown in Fig. 7.4. The correlation is basically eliminated by this transformation. This confirms that the ω–$\sigma_{\Delta t}$ correlation is basically due to the correlations between ω, $\sigma_{\Delta t}$ and p_t.

This conclusion suggests that the events with wrong tagging information have a softer p_t spectrum. The p_t spectra of correctly and incorrectly tagged events by `KTag` are plotted explicitly in Fig. 7.5.

To understand the sources of p_t differences, I use the *BABAR* event generator to examine the types of events from B^0 decays that produce at least one charged kaon in the final state. The charges of the final charged kaons are summed. If the sum is positive (negative) the B meson is tagged as B^0 ($\overline{B}^0$). If the sum is zero, there is no tagging information.

We first look at the tagging efficiencies for events with 0, 1, and 2 charm mesons (D^0, D^+, D_s), and also for events that have at least one fragmentation kaon, $B^0 \to K^{(*)}X$. The efficiencies, separated by tagging information, are summarized in Table 7.1. The wrong tag fraction is about 10% for all tagged events.

We further break down the wrong tag events by the sources of the wrong-sign

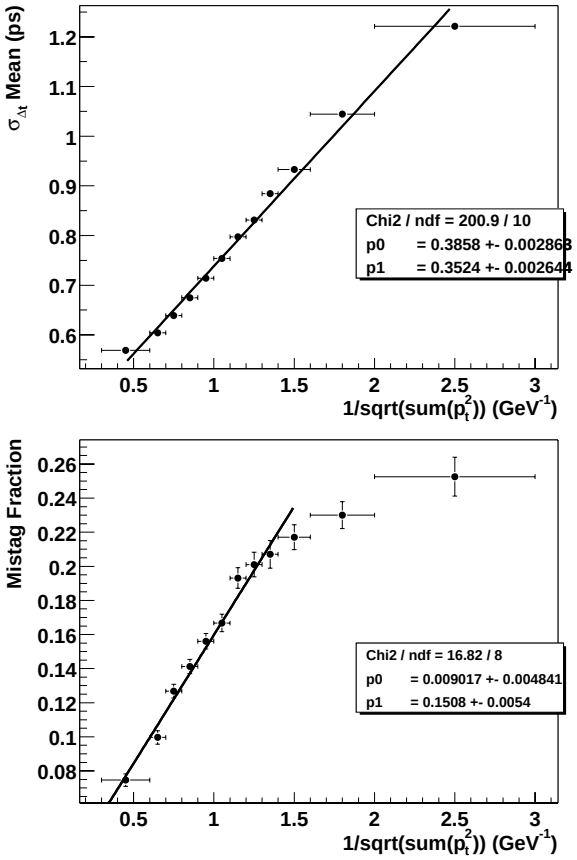

Figure 7.3: Mean of $\sigma_{\Delta t}$ (left) and mistag (right) as functions of $1/\sqrt{\sum_i (p_t^i)^2}$ in signal Monte Carlo simulation.

kaons. The percentages for different sources are listed and ranked in Table 7.2. The major sources of wrong-sign kaons are: two charm yielding wrong-sign D decaying to its right-sign K, one right-sign charged D decaying to wrong-sign K, and kaons from fragmentation.

I compare the $\sqrt{\sum_i p_t^2}$ spectrum among no-, one-, and two-charm events, and events that have at least one fragmentation kaon. The comparison is shown in Fig. 7.6. For the double charm and B fragmentation events, the spectra are softer. The primary reasons are that two charm events have lower momentum due to the

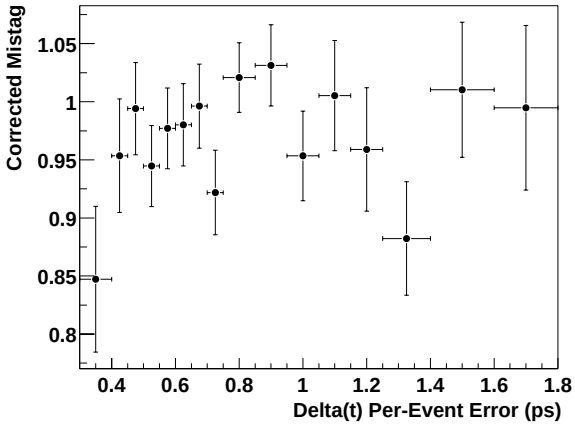

Figure 7.4: Mistag ratio $\omega/(0.151/\sqrt{\sum_i (p_t^i)^2} + 0.009)$ versus $\sigma_{\Delta t}$.

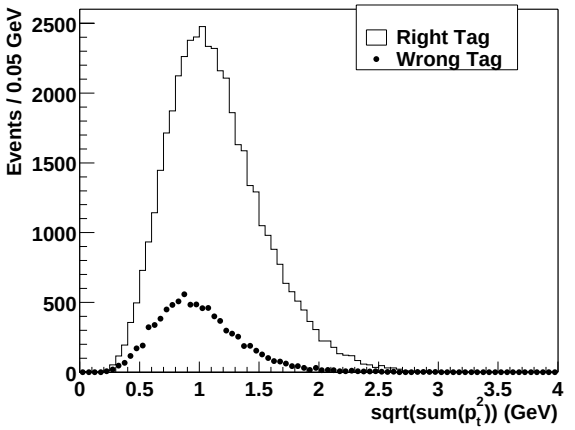

Figure 7.5: The $\sqrt{\sum_i p_t^2}$ spectra for correctly and incorrectly tagged events by `KTag` in signal Monte Carlo simulation. The summation is over all charged tracks used for tag-side vertexing.

production of two heavy quarks, and the fragmentation decays lose momentum to "spin energy" [74], etc.

We also compare right and wrong tag events for 0–2 charm events. As shown in Fig. 7.7, the wrong tag spectra are always softer than the right tag, driving the ω–$\sigma_{\Delta t}$ correlation. The differences are most noticeable in one-charm events, which

Table 7.1: The kaon tagging efficiencies at generator level, for events with 0, 1, and 2 charm mesons (D^0, D^+, D_s), and also for events that have at least one fragmentation kaon, $B^0 \to K^{(*)}X$. The fragmentation kaon sample is not a proper subset of the "No charm" sample because X can contain charm.

	Tag info			
Event type	Right	Wrong	No	Total
No charm	2.8%	0.6%	7.0%	10.4%
One charm	29.1%	1.9%	36.6%	67.6%
Two charm	9.8%	2.2%	10.0%	22.0%
Total	41.6%	4.7%	53.6%	
Frag. K	8.8%	1.0%	3.1%	12.8%

Table 7.2: The percentages of wrong sign kaons from different sources for wrong tag events estimated from Monte Carlo generator. The values are weighted by number of wrong sign kaons.

Source	%
Wrong-sign D^0, two charm	25.1
Right-sign D^+, one charm	21.5
Fragmentation $B^0 \to K^{(*)}$X	17.8
Right-sign D^+, two charm	7.6
Wrong-sign D^+, two charm	7.1
Wrong-sign ${D_s}^+$, two charm	5.8
$B^0 \to$ ResX, Res $\to K$Y	4.4
Other K mother (Λ_c, etc)	4.1
Right-sign D^0, one charm	1.9
Wrong-sign D^0, one charm	1.6
Right-sign D^0, two charm	1.5
Right-sign ${D_s}^+$, one charm	1.0
Wrong-sign D^+, one charm	0.4
Right-sign ${D_s}^+$, two charm	0.3

is dominated by D^+. The reason for the dominance of D^+ in the wrong-sign sample is that the branching ratio for a D^0 decaying to a right sign K^- (53%) is twice as

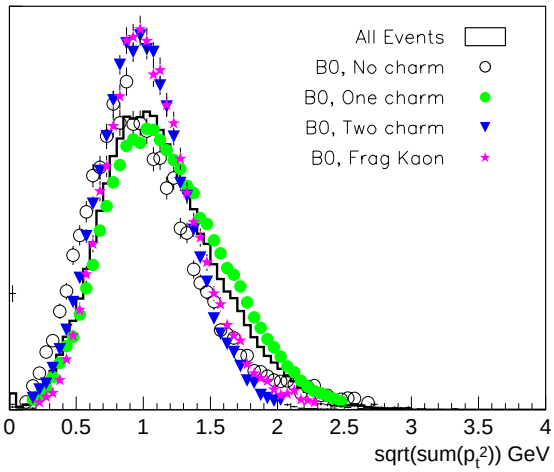

Figure 7.6: $\sqrt{\sum_i p_t^2}$ spectra for all, 0–2 charm, and fragmentation kaon events. Histograms are normalized to the same area.

large as a D^+ to K^- (24%), while for wrong sign K^+, D^0 is only about half (3.4%) of D^+ (5.8%). The right versus wrong tag spectra for events with only one right sign D^+ are shown in Fig. 7.8. This case is particularly easy to understand. Since the right sign D^+ decay produces a K^-, charge conservation requires at least two more charged particles in the final state. It results in a higher multiplicity of charged tracks for right-sign decay than the wrong-sign decay. Figure 7.10 shows the B and D multiplicities for right and wrong sign K for events with exactly one charged K in the final state. The minimum number of charged tracks for D^+ to right sign K ($D^+ \to K^-$) is three, because two extra positively charged tracks are needed to conserve charge. But for the wrong sign K ($D^+ \to K^+$), the dominant channels are $K^+\overline{K}^{(*)}(\pi^0)$, which have only one charged track. Since a lot of momentum is carried away by neutrals and fewer charged tracks are left, the wrong sign decay from charged D has softer $\sqrt{\sum_i p_t^2}$ spectrum than right sign decay.

The same spectra are compared for events with at least one kaon from B fragmentation, as shown in Fig. 7.9.

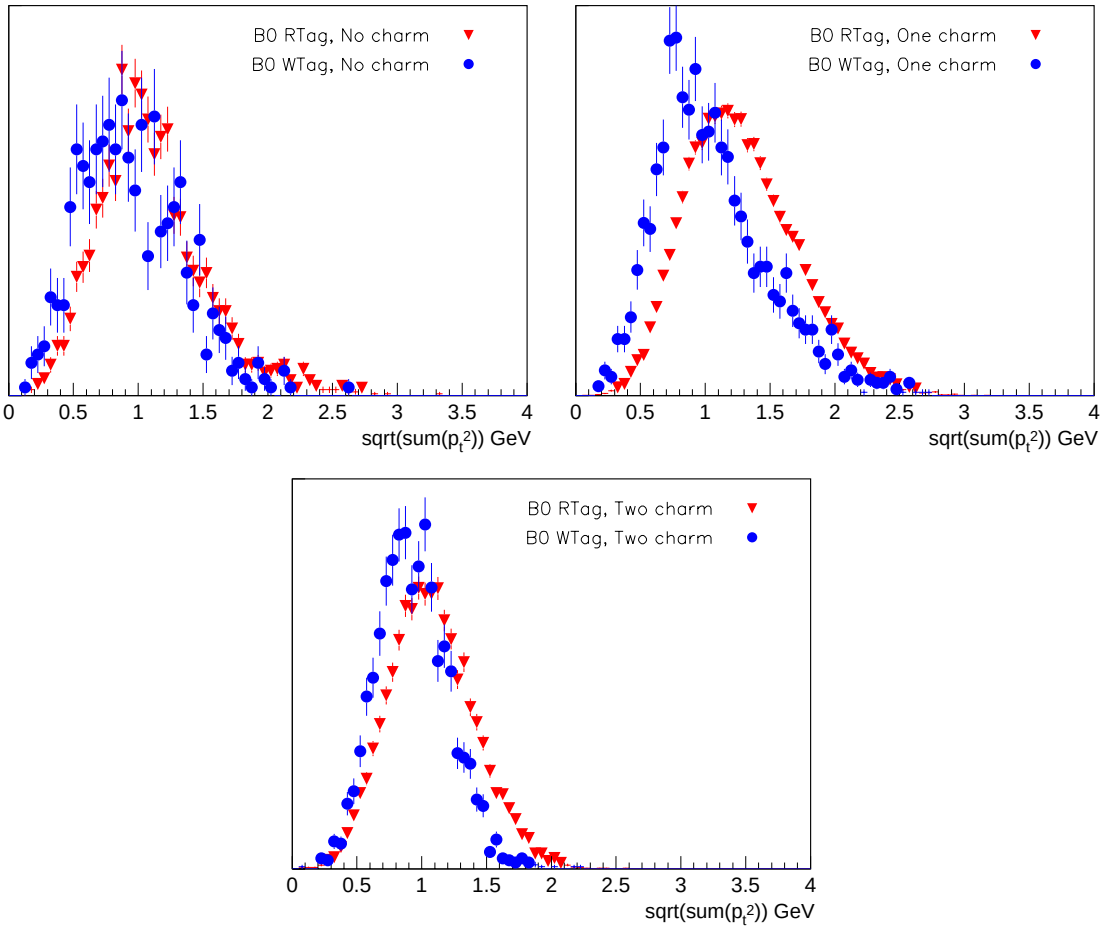

Figure 7.7: Right and wrong tag event spectra for 0–2 charm events. Histograms are normalized to the same area.

All these spectrum comparisons show the same pattern, that ultimately contributes to the increase of mistag fraction as $\sigma_{\Delta t}$ increases.

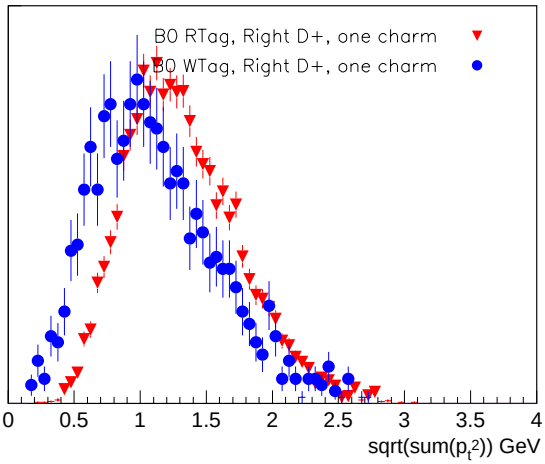

Figure 7.8: Right and wrong tag event spectra for events with only one right sign D^+. Histograms are normalized to the same area.

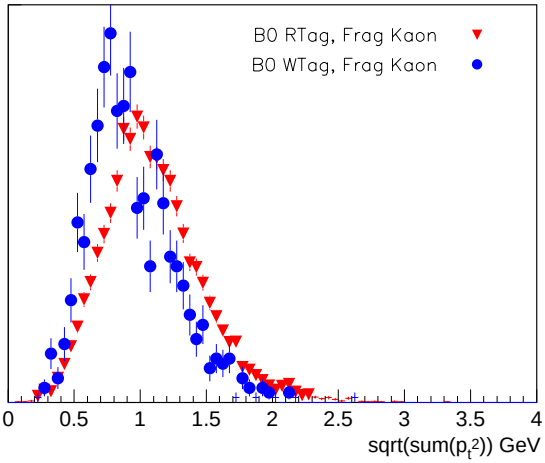

Figure 7.9: Right and wrong tag event spectra for events with at least one kaon from B fragmentation. Histograms are normalized to the same area.

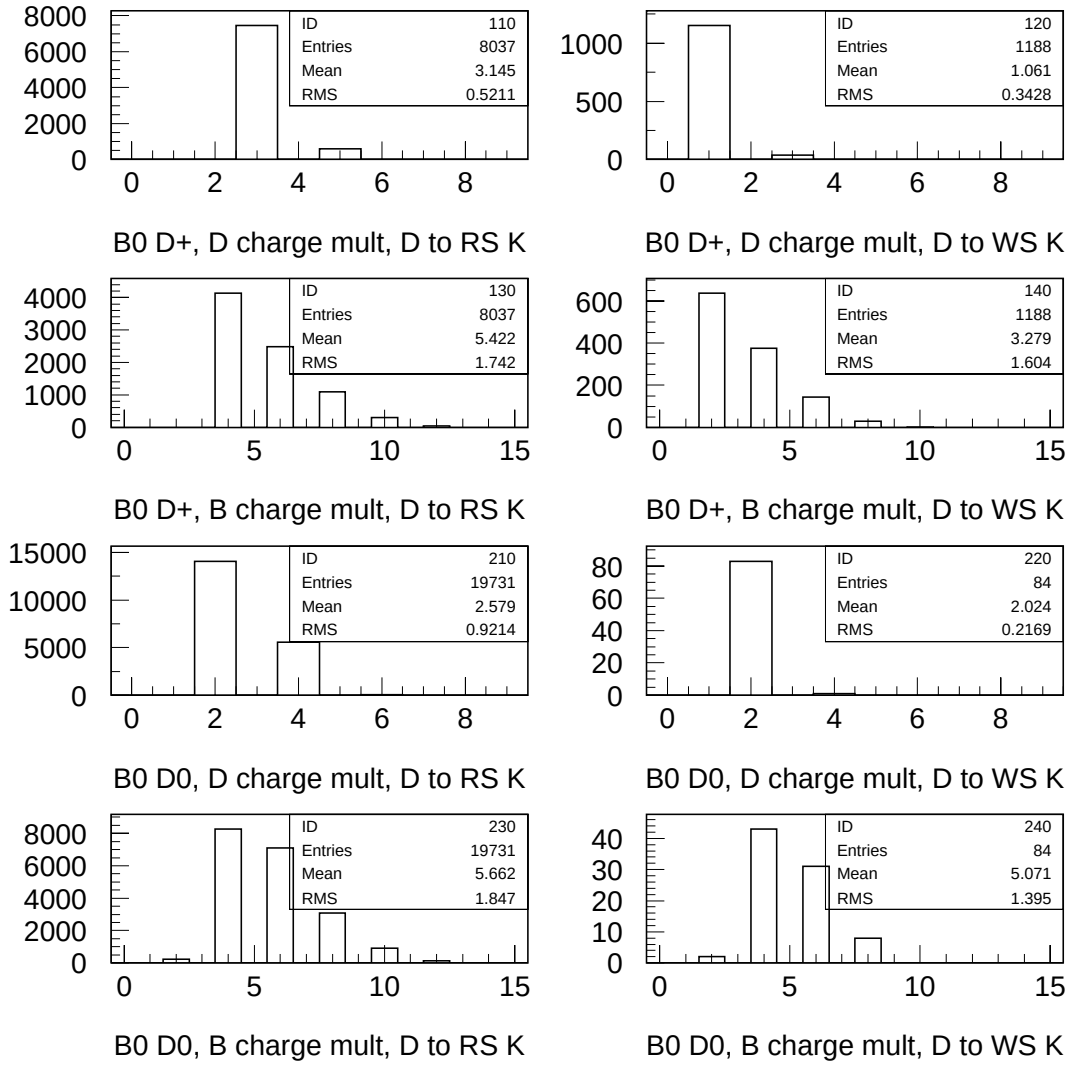

Figure 7.10: Charge multiplicities for B and D.

Chapter 8

Event Selection

8.1 Introduction

The events in this analysis are selected from the semileptonic decay of neutral B mesons, which contains a charged D^* and an electron or muon. The main source of signal events is $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_\ell$ decays[1]. A few percent originates from $\overline{B}^0 \to D^{*+}X\ell^-\overline{\nu}_\ell$ decays, where D^*X can be from higher charm resonances or simply from fragmentation. In both cases the decay time structures are identical to that of the main decay mode.

The D^* is reconstructed through the $D^{*+} \to D^0\pi^+$ channel, and the D^0 through four decay modes, $D^0 \to K^-\pi^+$, $K^-\pi^+\pi^0$, $K^-\pi^+\pi^-\pi^+$, or $K_S\pi^+\pi^-$. A charged D^* is fully reconstructed and is combined with an identified oppositely-charged electron or muon. This $D^*\ell$ pair is then required to pass kinematic cuts that enhance the contribution of semileptonic $B^0 \to D^{*-}\ell^+\nu_\ell$ decays. In addition to the signal sample, several background-enriched control samples that are used to characterize the main sources of background are also selected. These backgrounds can be categorized

[1]Be reminded that throughout this thesis, unless explicitly specified otherwise, charged conjugate modes are implied.

according to the D^* reconstruction status and the source of the lepton candidate:

1. events with a correctly reconstructed D^{*+} candidate,
   (a) events originating from $B\overline{B}$ events,
      i. events with a correctly identified lepton candidate,
         A. **signal**: $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_\ell X$ decays,
         B. or, **charged B background**: $B^- \to D^{*+}\ell^-\overline{\nu}_\ell X^-$,
         C. or, **correlated background**: $\overline{B}^0$ or $B^- \to D^{*+}X$, $X \to \ell^- Y$,
         D. or, **uncorrelated background**: $\overline{B}^0$ or $B^- \to D^{*+}X$, and the other B through $B \to \ell^- Y$,
      ii. or, **fake-lepton background**: events with a misidentified lepton candidate.
   (b) or, **continuum background**: $c\overline{c} \to D^{*+}X$.
2. or, **combinatoric background**: events with a mis-reconstructed D^* candidate, either from $B\overline{B}$ events or continuum.

The data analyzed in this thesis are collected during 1999–2000 of *BABAR* operation, of which 20.6 fb^{-1} are on peak and 2.6 fb^{-1} are at 40 MeV below the peak. The total number of $B\overline{B}$ pairs collected by *BABAR* during this period of time is estimated to be about 23 million (see Sec. 4.2). A total amount of 68,000 events are selected, of which 27,000 events are in signal samples, including combinatoric background, and the remainder are in background control samples.

The common selection criteria for signal and background control samples are described in Secs. 8.2–8.8. Sec. 8.9 describes the selections specific to signal and different control samples.

8.2 Track Selection

All charged tracks reconstructed in the DCH and/or SVT are categorized hierarchically according to the selection requirements.

The pions for K_S reconstruction are selected from all charged tracks without additional requirements.

The soft pion candidate from the D^* decay is selected from a list that uses very loose requirements, which are (1) the momentum measured in the lab frame is less than 10 GeV and (2) the distance of closest approach (doca) to the nominal beam spot centroid calibrated for each run is less than 1.5 cm on the transverse plane and less than 10 cm along the z-axis.

The remaining charged tracks, except the lepton candidates, are selected from a list that imposes additional requirements: (1) the transverse momentum is larger than 100 MeV, and (2) at least 20 hits are recorded in the drift chamber, out of a possible maximum of 40 hits for tracks perpendicular to the beam pipe.

The lepton candidates are selected from a tight list that requires, in addition to those requirements shown above, the distance of closest approach to be less than 1 cm on the transverse plane and less than 3 cm along the z-axis.

At the final stage of event selection, the same doca requirement for lepton candidates is imposed on all charged tracks in the final state except those used for K_S reconstruction.

8.3 Event Shape

To reduce continuum background, we calculate the angle $\theta_{\rm thrust}$ between the thrust of the $D^*\ell$ candidate and that of the rest of the event in the e^+e^- center-of-mass frame. Events from continuum tend to be jetty, therefore the angle between two

thrusts tends to be close to zero or 180 degrees, so that the angle is peaked near $|\cos\theta_{\rm thrust}| = 1$, while the decay topology of $B\overline{B}$ events is more spherical since two B mesons decay almost at rest, therefore the directions of the two trusts are not correlated and the angular distribution is flat. Figure 8.1 shows the $|\cos\theta_{\rm thrust}|$ distributions of $B\overline{B}$ and continuum events. We select events with $|\cos\theta_{\rm thrust}| < 0.85$. The efficiencies for signal and continuum events to satisfy this requirement are about 85% and 46%, respectively.

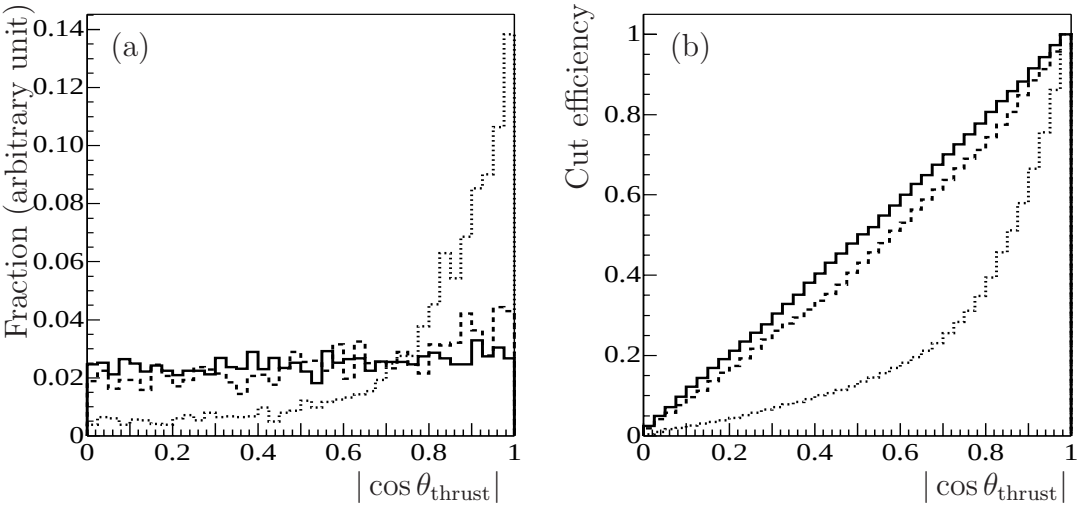

Figure 8.1: (a) Distributions of the angle between the thrusts of the $D^*\ell$ candidate and the rest of the event for signal events (solid histogram), D^*-hadron events (dashed histogram) from $B\overline{B}$ Monte Carlo, and events that are selected from off-resonance data (dotted histogram). (b) Efficiencies of making cut on $\cos\theta_{\rm thrust}$ for these types of events.

8.4 π^0 Reconstruction

Neutral pion candidates are formed from pairs of EMC bumps with energy greater than 30 MeV, assumed to be photons originating from the interaction point. The

invariant mass of the photon pair is required to be between 119.2 MeV and 150.0 MeV, which is about $\pm 2.5\sigma_{\gamma\gamma}$ of the nominal π^0 mass, where $\sigma_{\gamma\gamma}$ is the $\gamma\gamma$ mass resolution, with a minimum total energy of 200 MeV. Selected γ candidates are kinematically fitted so that $\gamma\gamma$ invariant mass is equal to the nominal π^0 mass. The χ^2 probability of the fit is required to be greater than 1%. Within the acceptance of the EMC, efficiencies for this selection vary from about 55% to 65% for π^0 energies from 0.3 to 2.5 GeV, the range of π^0 energies in B decays [75].

The $\gamma\gamma$ mass distribution and the χ^2 probability of mass-constraint are shown in Fig. 8.2.

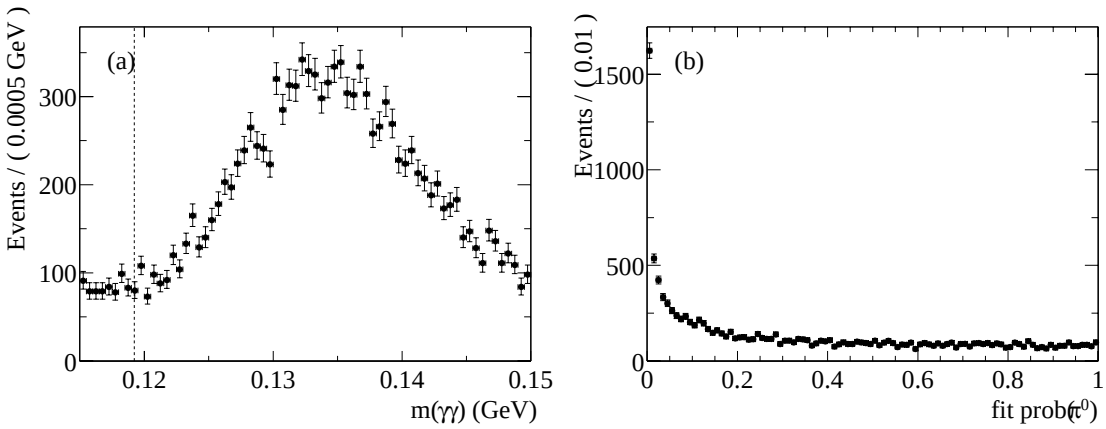

Figure 8.2: (a) $\gamma\gamma$ invariant mass invariant mass distribution for π^0 candidates. (b) χ^2 probability of the kinematic fit for π^0 candidates. The plotted distributions are from data within $m(D^*) - m(D^0)$ peak (see Sec. 8.8), with selection criteria close to the final set except that no D^0 mass cut (see Sec. 8.6) and no χ^2 probabilities for any vertex fits (see Sec. 8.8) are applied.

8.5 K_S Reconstruction

K_S candidates are formed from pairs of charged tracks, assumed to be pions. These two tracks are fitted to a common vertex. The fit χ^2 probability is required to be

greater than 1%. The $\pi\pi$ invariant mass, calculated at the vertex of the two tracks, is required to be within ± 15 MeV of the nominal K_S mass [9].

The $\pi\pi$ mass distribution evaluated at the vertex, and the χ^2 probability of mass-constraint fit are shown in Fig. 8.3.

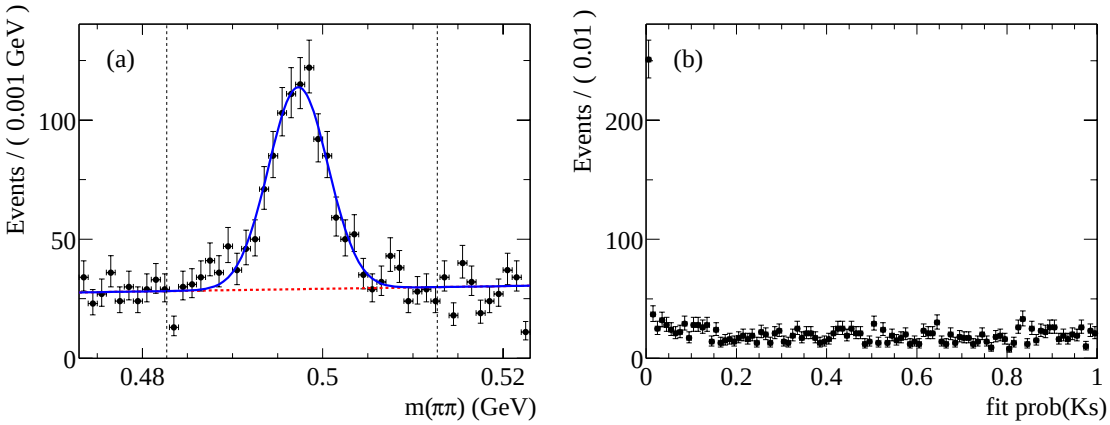

Figure 8.3: (a) π π invariant mass for K_S candidates. Curve is a Gaussian plus a flat distribution. The width of the Gaussian is about 3.3 MeV. (b) χ^2 probability of π π vertexing for K_S candidate. The plotted distributions are from data within $m(D^*) - m(D^0)$ peak (see Sec. 8.8),, with selection criteria close to the final set except that no D^0 mass cut (see Sec. 8.6) and no χ^2 probabilities for any vertex fits (see Sec. 8.8) are applied.

8.6 D^0 Reconstruction

The D^0 candidates are obtained from reconstructing $D^0 \to K^-\pi^+$, $D^0 \to K^-\pi^+\pi^-\pi^+$, $D^0 \to K^-\pi^+\pi^0$ and $D^0 \to K_S\pi^+\pi^-$ decay modes. Charged kaon of $D^0 \to K^-\pi^+$ ($D^0 \to K^-\pi^+\pi^-\pi^+$, $D^0 \to K^-\pi^+\pi^0$) is required to pass *not-a-pion (tight)* criteria (see Sec. 5.2.3 for definitions). No particle identification is applied for $D^0 \to K_S\pi^+\pi^-$ modes.

The raw invariant mass distributions of D^0 candidates are shown in Fig. 8.4. The widths of the peak Gaussian are about 14 MeV for $K^-\pi^+\pi^0$ mode and about

7 MeV for other three modes, which do not have a π^0 in their final states. The selection windows are chosen to be ± 34 and ± 17 MeV from nominal D^0 mass [9] for $K^-\pi^+\pi^0$ and other three modes, respectively, which are about 2.5 sigmas. The D^0 candidates are then reconstructed with the invariant mass constrained to be the nominal D^0 mass. The χ^2 probability of the fit is required to be larger than 0.1%. This probability distributions is shown in Fig. 8.5

The $D^0 \to K^-\pi^+\pi^0$ and $D^0 \to K_S\pi^+\pi^-$ modes are known to have resonance substructure [76, 77]. The major resonances are $K^-\rho^+$, $K^{*-}\pi^+$ and $K^{*0}\pi^0$ for $D^0 \to K^-\pi^+\pi^0$ mode and $K^{*-}\pi^+$,$K^0\rho^0$, $K^0 f_0$, $K^0 f_2$, etc., for $D^0 \to K_S\pi^+\pi^-$ mode. One can take advantage of these structures by selecting D^0 candidates near these resonances to increase signal to background ratios. Figure 8.6 shows the Dalitz plots [78] for these two modes, in which the density is calculated based on the decay amplitudes and phases found in [77] (Details are discussed in Appendix. C.). The event density (Dalitz weight) distributions and the cut efficiencies for the signal and flat background are also shown in Fig. 8.6.

For each event, its expected density on the Dalitz plot is calculated. The event is kept if this value is larger than 10 % of the maximum value (within the red contours in plots (a) and (d) in Fig. 8.6). This cut is very effective against flat background.

It is worth noticing that the Dalitz plot for $D^0 \to K_S\pi^+\pi^-$ is not symmetric under the exchange of π^+ and π^-. It is due to the fact that a D^0 can produce the resonance $K^{*-}\pi^+$ but not $K^{*+}\pi^-$, and vice versa for $\overline{D}^0$. Therefore, we need to know whether the flavor of the mother particle is a D^0 or $\overline{D}^0$, (which cannot be determined from the final state $K_S\pi^+\pi^-$ alone), before the Dalitz density can be calculated correctly. We need to rely on the charge association between the D^* and soft π, i.e., $D^{*+} \to D^0\pi^+$, and $D^{*-} \to \overline{D}^0\pi^-$ to determine the flavor of D^0.

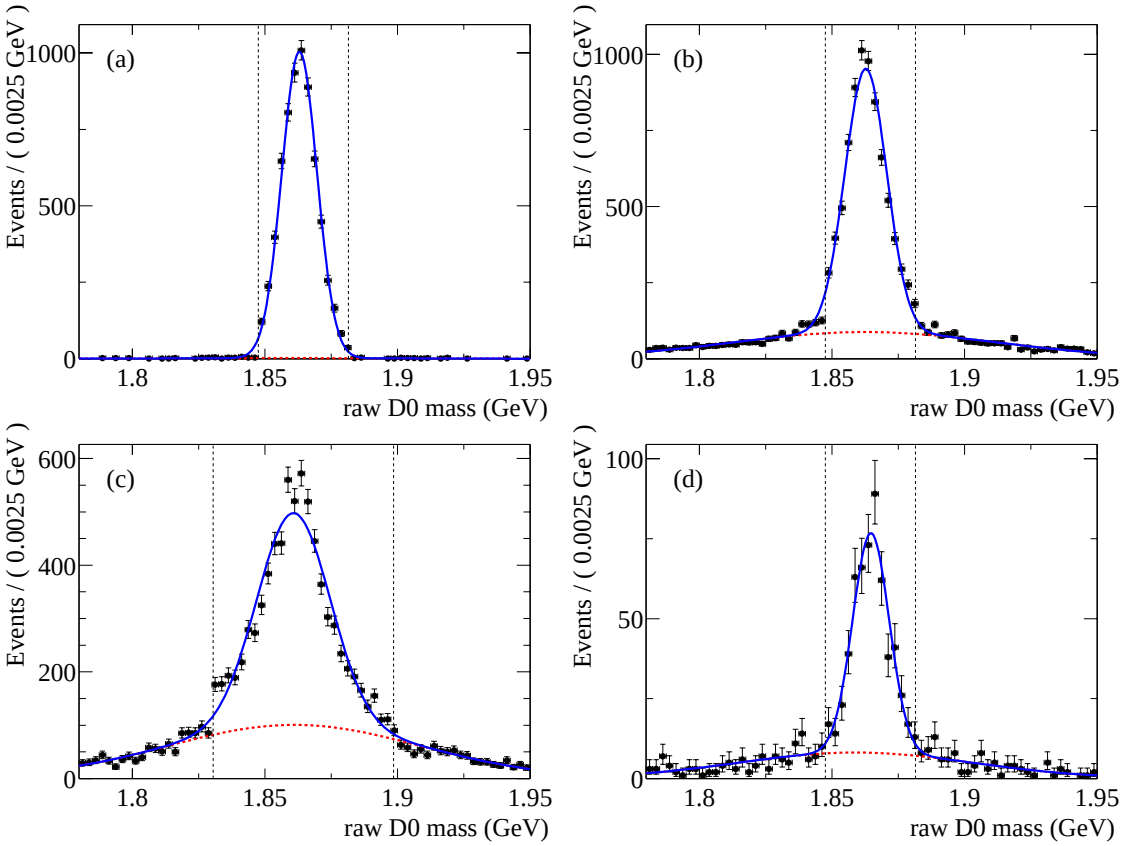

Figure 8.4: Raw D^0 candidate mass distributions for (a) $K^-\pi^+$, (b) $K^-\pi^+\pi^-\pi^+$, (c) $K^-\pi^+\pi^0$, and (d) $K_S\pi^+\pi^-$ modes. The plotted distributions are from data within $m(D^*) - m(D^0)$ peak (see Sec. 8.8), with selection criteria close to the final set except that no χ^2 probabilities for any vertex fits (see Sec. 8.8) are applied, and the mass windows for π^0 and K_S are tightened to 125–145 MeV and 495–505 MeV, respectively. Fitted curves are double-Gaussian. Backgrounds are not flat because the events are pre-selected such that at least one D^0 candidate in each event has mass within $\pm$35 MeV of nominal D^0 mass for $K^-\pi^+\pi^0$ mode and $\pm$20 MeV for other three modes.

8.7 D^* Reconstruction

D^* candidates are reconstructed by combining the D^0 candidate with a charged π (called the soft pion or π_s^+) that has transverse momentum greater than 50 MeV and charge opposite to the charge of the kaon candidate. For the $K_S\pi^+\pi^-$ mode, the

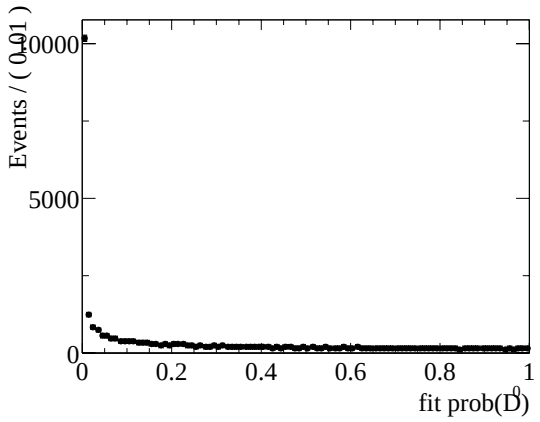

Figure 8.5: χ^2 probability of D^0 mass-constraint fit.

charge of the soft pion is used to identify the flavor of D^0 candidates, so that the Dalitz density can be calculated correctly (see Sec. 8.6.).

Since we are selecting D^* candidates from a B decay, D^* candidates are required to have momentum less than 2.5 GeV in the $\Upsilon(4S)$ rest frame. As shown in Fig. 8.7, this cut preserves all signal events but removes a significant amount of D^* from continuum events.

The spectrum that is used to determine whether a D^* is reconstructed correctly is the difference between the candidate $D^0\pi_s^+$ mass with the candidate D^0 mass constrained to the true D^0 mass, and the true D^0 mass ($m(D^*) - m(D^0)$), after the π_s^+ is refitted with D^*-ℓ vertex and beams-pot constraint (See Ch. 6.). The spectrum is peaked around 145.5 MeV with a width less than 1 MeV, depending on the soft pion reconstruction status (See Ch. 10 for more details.). Figure 8.8 shows the $m(D^*) - m(D^0)$ distribution for signal sample in data. All events with $m(D^*) - m(D^0) < 165$ MeV will be used in the final fit. There is no explicit cut on this quantity for signal, but this distribution is used for determining the probability of a given event being a certain type of background in the final fit.

8.8 D^* ℓ Vertex and Tag-side B Requirement

D^* ℓ candidates are retained if the χ^2 probability that the daughter tracks form a common vertex is greater than 1%. Events are selected if the fit that determines the distance Δz converges, the decay time difference Δt calculated from Δz is between ± 18 ps, the error on Δt estimated from the tracking resolution ($\sigma_{\Delta t}$) is less than 1.8 ps, and at least two tracks are used to determine the vertex of the other B (tagging B). The motivation of the last two criteria to reduce outlier events on Δt residual space and to keep the linearity between $\sigma_{\Delta t}$ and the mean of Δt residual valid (See Sec. 10.1.1 for details.).

8.9 Background Control Samples

As discussed in Sec. 2.5, the cosine of the angle between the D^* and ℓ momenta in the $\Upsilon(4S)$ rest frame ($\cos\theta^*_{D^*\ell}$) peaks at -1, and the cosine of the angle between the B^0 and $D^*\ell$ candidate momenta ($\cos\theta^*_{B^0,D^*\ell}$, see Eq. 2.52.) is between ± 1. Figure 8.10 shows these two distributions for signal and background events.

For the events in the signal sample, the lepton satisfies electron or muon selection and is on the *"opposite-side"* of the D^* candidate, where opposite-side is defined as $\cos\theta^*_{D^*\ell} < 0$ and $|\cos\theta^*_{B^0,D^*\ell}| < 1.1$. The range ± 1.1 is used instead of ± 1 to include some signal events outside of ± 1 due to detector resolution. A background control sample, enhanced in uncorrelated-lepton background and called the *"same-side"* sample is composed of $D^*\ell$ candidates with $\cos\theta^*_{D^*\ell} > 0$ and $|\cos\theta^*_{B^0,D^*(-\ell)}| < 1.1$, which is calculated by Eq. 2.52 in the $\Upsilon(4s)$ rest frame after the lepton candidate is flipped with respect to its origin by 180° ($\vec{p}_\ell \to -\vec{p}_\ell$, denoted as $-\ell$). The quantity $\cos\theta^*_{B^0,D^*(-\ell)}$ is used instead of $\cos\theta^*_{B^0,D^*\ell}$ because the distribution of $\cos\theta^*_{B^0,D^*(-\ell)}$ in the background control sample is similar to the distribution of $\cos\theta^*_{B^0,D^*\ell}$ for

uncorrelated-lepton background in the signal sample, based on simulation, whereas the distribution of $\cos\theta^*_{B^0,D^*\ell}$ in the background control sample is systematically different.

Events in the signal sample must be selected from on-resonance events. The lepton candidates for the signal sample must satisfy a "very-tight" particle identification criteria for either electrons or muons. A lepton candidate is also selected if it fails "loose" criteria for both electron and muon identification. This sample is called *"fake-lepton"* control sample. We define three categories for lepton identification, satisfying electron or muon selection, or failing both selections, two for the angular relation between lepton and D^* candidates, and two for on or off $\Upsilon(4S)$ resonance at which events are recorded. Out of the 12 possible combinations of these three attributes, two represent the signal sample and the remaining ten are background control samples (see Table 8.2). Of the ten control samples, five are primary samples that we use to directly study backgrounds that contribute to the signal sample, and the remaining five are secondary samples that we use to study backgrounds in the primary control samples. The description of these three attributes is listed in the upper half of Table 8.1.

Each of the 12 samples is further divided into 30 subsamples according to three additional reconstruction characteristics listed in the lower half of Table 8.1: DCH hits of the soft pion (`svtDch`, two choices), D^0 reconstruction modes (`Dmode`, three choices), and tagging categories (`tagCat`, five choices, detail described in Ch.7). The `svtDch` category is useful to isolate the subsample (of about 20%) of candidates for which the D^* mass is more precisely determined because the slow pion has higher momentum and DCH hits are available. `Dmode` and `tagCat`, are used to account for the different amounts and properties of the combinatoric background that we observe in these subsamples. The sample is also subdivided by `tagCat` in order to separate events with different flavor tagging performances and Δt resolution to improve the

Table 8.1: Attributes used to subdivide the signal and control samples for the purposes of background characterization and extraction of lifetime and mixing parameters.

Attribute	N_{cat}	Description
`angCut`	2	What are the angular correlations between the D^* and ℓ? (1) candidates satisfy the "opposite-side" (OS) criteria (2) candidates satisfy the "same-side" (SS) criteria
`onOffRes`	2	At what center-of-mass energy was the event recorded at? (1) near the Υ(4S) resonance (2) below the Υ(4S) resonance
`leptID`	3	Is the kinematically-selected ℓ candidate lepton-like? (1) candidate satisfies "very-tight" electron selection (2) candidate satisfies "very-tight" muon selection (3) candidate fails "loose" electron and muon selections
`svtDch`	2	How was soft-pion candidate track reconstructed? (1) in the SVT only (2) in the SVT and the DCH
`Dmode`	3	How was the D^0 candidate reconstructed? (1) $D^0 \to K\pi$ (2) $D^0 \to K\pi\pi^0$ (3) $D^0 \to K\pi\pi\pi$ or $D^0 \to K_S\pi\pi$
`tagCat`	5	How was the flavor of the non-$D^*\ell$ B determined? (1) with the `lepton` category algorithm, (2) with the `kaon` category algorithm, (3) with the `NT1` category algorithm (4) with the `NT2` category algorithm (4) with the `NT3` category algorithm

sensitivity to mixing oscillations in decay-time fits.

Table 8.2: Summary of the signal and control samples selected for lifetime and mixing analysis. Samples 1–2 are signal, 3–7 are primary background control samples, and 8–12 are secondary background control samples.

Sample	`angCut`	`onOffRes`	`leptID`	Sample enriched in...
1	OS	On	pass e	electron signal
2	OS	On	pass μ	muon signal
3	OS	On	fail e & μ	fake BG to samples 1 and 2
4	OS	Off	pass e	continuum BG to sample 1
5	OS	Off	pass μ	continuum BG to sample 2
6	SS	On	pass e	uncorrelated BG to sample 1
7	SS	On	pass μ	uncorrelated BG to sample 2
8	OS	Off	fail e & μ	fake BG to samples 4 and 5
9	SS	On	fail e & μ	fake BG to samples 6 and 7
10	SS	Off	pass e	continuum BG to sample 6
11	SS	Off	pass μ	continuum BG to sample 7
12	SS	Off	fail e & μ	fake BG to samples 10 and 11

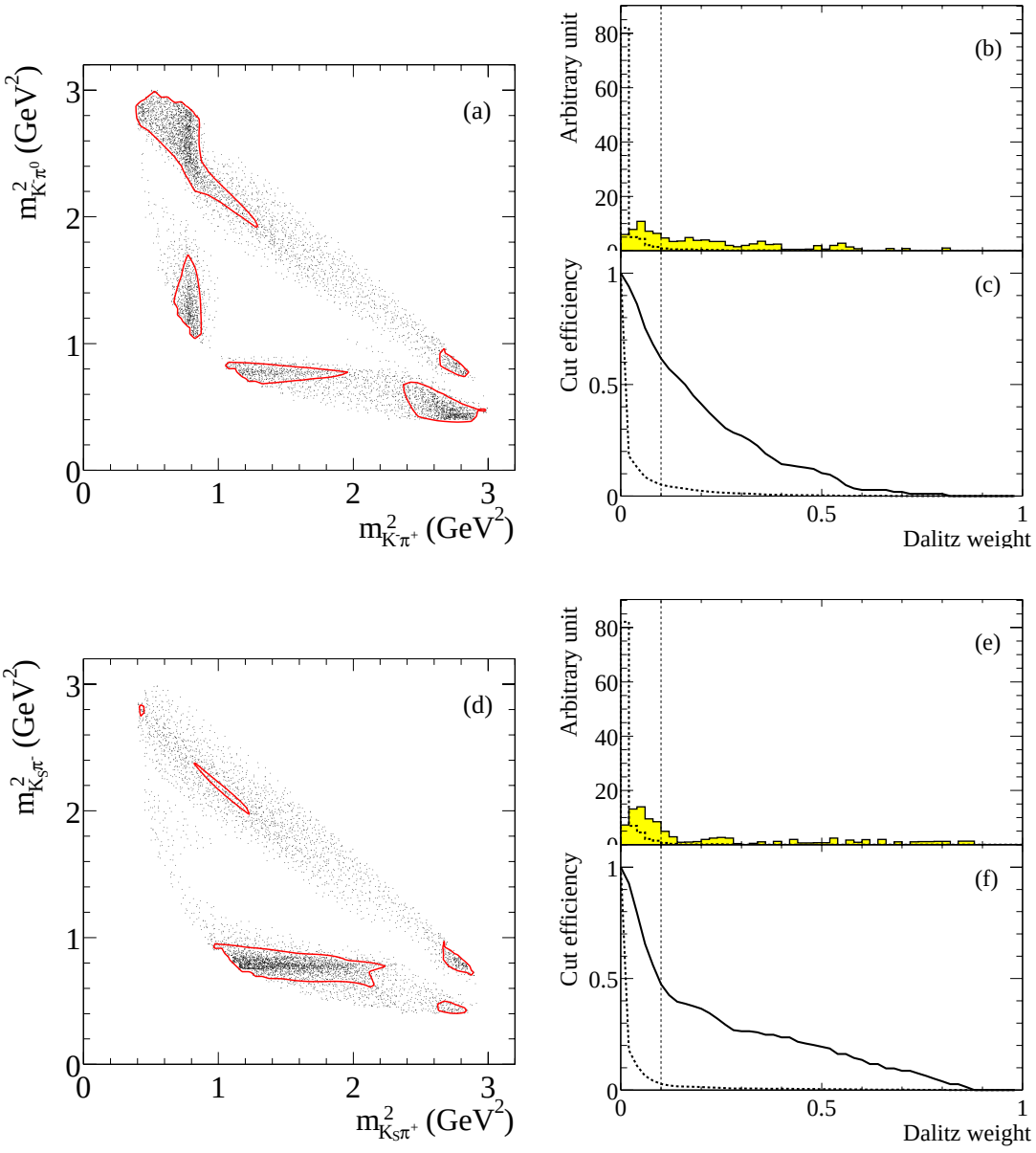

Figure 8.6: (a) Dalitz plots calculated with the amplitudes and phases measured by [77] for $D^0 \to K^-\pi^+\pi^0$. Red contours represent the regions where the density is higher than 10 % of the maximum value. (b) Distributions of density on the Dalitz plot for signal (solid histogram) and flat background (dashed histogram). (c) Cut efficiencies against signal (solid line) and flat background (dashed line). (d)–(f) Same plots as (a)–(c) for $D^0 \to K_S\pi^+\pi^-$.

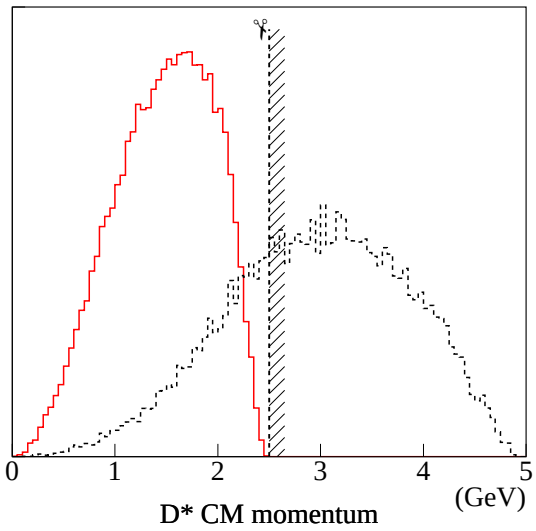

Figure 8.7: D^* momentum distribution in $\Upsilon(4S)$ rest frame from *BABAR* event generator, for signal (red histogram) and for continuum events (black dashed histogram).

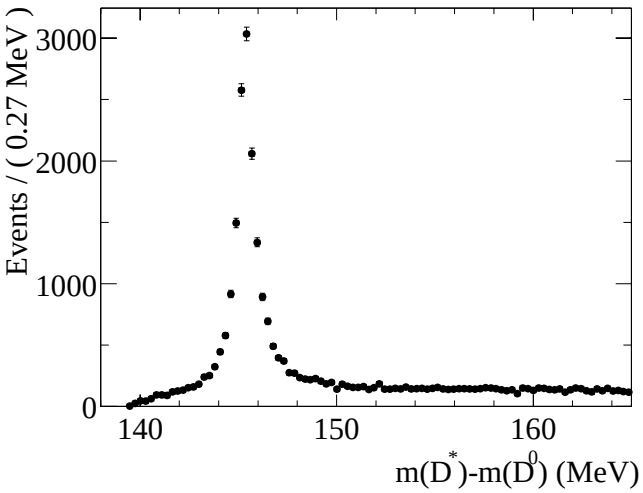

Figure 8.8: The distribution of $m(D^*) - m(D^0)$ for signal samples in data.

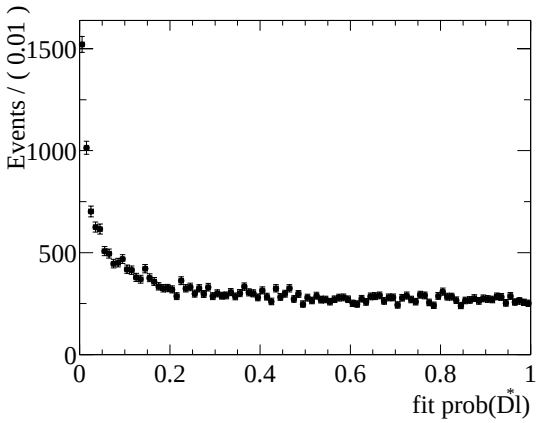

Figure 8.9: χ^2 probability of vertex fit for D^* ℓ candidate. The plotted distribution is from data within $m(D^*) - m(D^0)$ peak, with selection criteria close to the final set except that no D^0 mass cut and χ^2 probabilities for any vertex fits are applied.

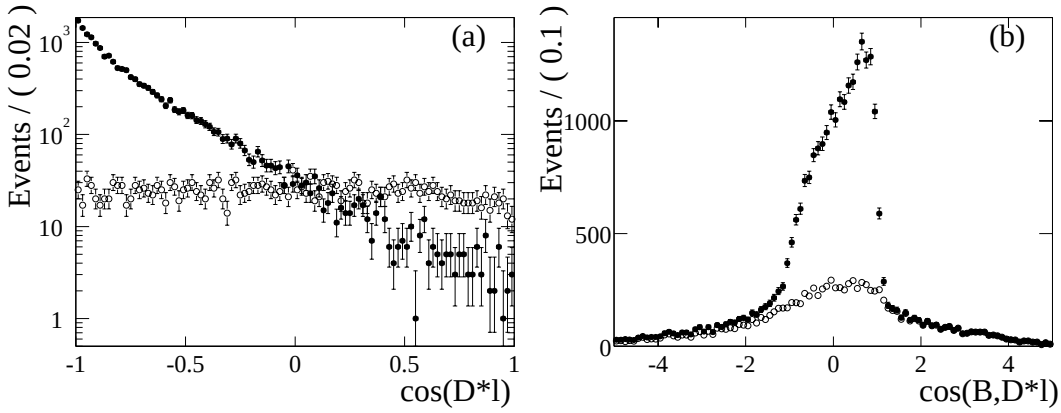

Figure 8.10: (a) Distributions of $\cos\theta^*_{D^*\ell}$ for signal (solid dots) events and for background events (open circles) whose D^* and ℓ candidates are from different B mesons; (b) distributions of $\cos\theta^*_{B^0,D^*\ell}$ calculated by Eq. 2.52 for all selected $B\overline{B}$ events (solid dots) and for all background events (open circles). Both angles are evaluated in the Υ(4S) rest frame. The plotted distributions are from signal Monte Carlo simulation.

Chapter 9

Background Characterization

In this section, the procedure of determining the components of major backgrounds is described. Since the neutrino in a event cannot be detected, the energy and mass of a B candidate are not available for discriminating signal and background. The distribution of the $D^* - D^0$ mass difference ($\delta m \equiv m(D^*) - m(D^0)$) is therefore used as one of the discriminating variables. It determines whether the D^* in an event is correctly reconstructed. As described in Sec. 8.9, several background control samples are selected for estimating levels of major background in the signal sample. The off-resonance data are used to estimate the continuum contribution in on-resonance data. Events whose lepton candidate fails the lepton identification ("fake-lepton" sample) are used to calculate the contamination due to lepton mis-identification. Finally, the events whose lepton and D^* candidates are in the same hemisphere ("same-side" control sample) are used to project the contribution from uncorrelated leptons.

9.1 Lepton Identification

As will be shown in Sec. 9.6, the momentum- and angle-weighted particle identification efficiencies are used to disentangle the hadron contribution in each control

sample. In this section, the procedure of calculating these efficiencies is discussed.

We use the "very tight" electron and muon selectors (see Sec. 5.2) to select the e and μ signal samples. The lepton candidates that fail both the "loose" electron and "loose" muon selectors are defined as the fake sample. In Fig. 9.1, we show the number of events that pass a particular selector and fail the tighter selectors, for electrons and muons in signal Monte Carlo simulation. In order to calculate the efficiency of the lepton selection and the contribution of non-lepton backgrounds, we use angle- and momentum-weighted selector efficiencies $\epsilon(x \to y)$ measured in data, weighted according to the angular and momentum distributions in the lab frame for the relevant particles in $B^0\overline{B}^0$ and B^+B^- generic Monte Carlo samples. The notation $x \to y$ means the efficiency for selecting a particle of type x with the selector for particle of type y. We define the angle- and momentum-averaged efficiencies as

$$\epsilon(x \to y) \equiv \frac{1}{N} \int dp\, d\cos\theta\, d\phi \, \frac{dN}{dp\, d\cos\theta\, d\phi} \cdot \omega_{x\to y}(p, \cos\theta, \phi) \ , \tag{9.1}$$

where $\omega_{x\to y}(p, \cos\theta, \phi)$ is the selector efficiency as a function of particle momentum and direction measured in the lab frame, and the function

$$\frac{1}{N} \frac{dN}{dp\, d\cos\theta\, d\phi}$$

describes the distribution of these variables in the relevant physics samples. The lab momentum distributions of lepton and hadron tracks (according to Monte Carlo truth) from generic Monte Carlo simulation are shown in Fig. 9.2.

The lepton candidate momentum distributions for opposite-side (OS) and same-side (SS) (defined in Table 8.1) samples are significantly different. As can be seen in Fig. 9.3, the SS sample has a lower peak at 1 GeV, while OS sample peaks at 2 GeV. Therefore, we calculate two sets of PID efficiencies for OS and SS samples separately.

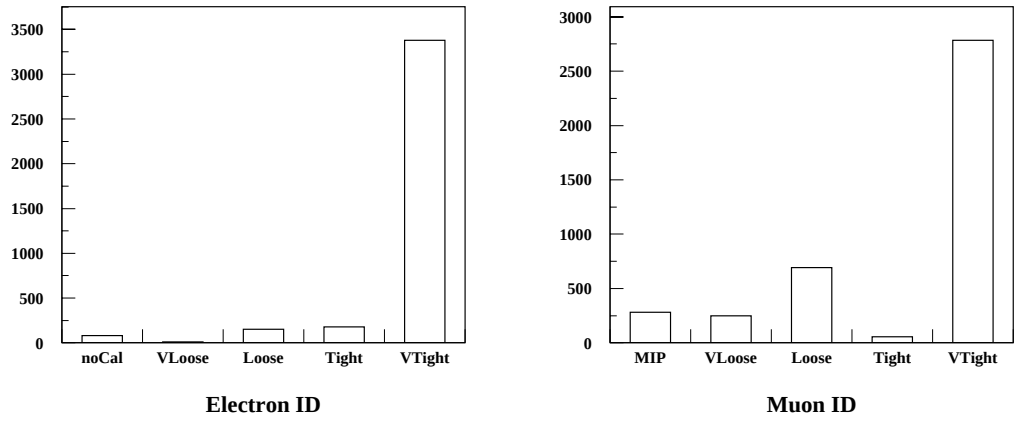

Figure 9.1: Number of events that pass one of the standard lepton selectors and fail the tighter selectors, for electrons (left plot) and muons (right plot) in signal Monte Carlo.

The electron and muon selector efficiencies $\omega_{x\to y}(p, \cos\theta, \phi)$ are estimated with particle identification efficiency and misidentification tables calculated from electron, muon, pion, kaon and proton control samples in data [64, 67]. The efficiencies and misidentification rates are tabulated in bins of (p, θ, ϕ), measured in the lab frame. For each particle type determined from the Monte Carlo truth (electron, muon, π, K, or p), a list of particle momenta and angles, $\{(p, \theta, \phi)_\ell\}_{\ell=1}^{N}$, from generic $B^0\overline{B}^0$ and B^+B^- Monte Carlo events are generated to calculate the momentum- and angle-weighted selector efficiency:

$$\epsilon(x \to y) \simeq \frac{1}{\sum_s f_s N_s} \sum_{n=1}^{m} (\sum_s f_s N_{n;s})\, \omega_{x\to y}(n) \ ,$$

where m is the total number of bins and $s = B^0\overline{B}^0, B^+B^-$. N_s is the total number of events in sample s and $N_{n;s}$ is the number of events in sample s that fall in bin n. The scaling factor f_s is the ratio of the equivalent luminosity of the $B^0\overline{B}^0$ Monte Carlo sample to that of sample s. The uncertainty on $\epsilon(x \to y)$ is calculated with

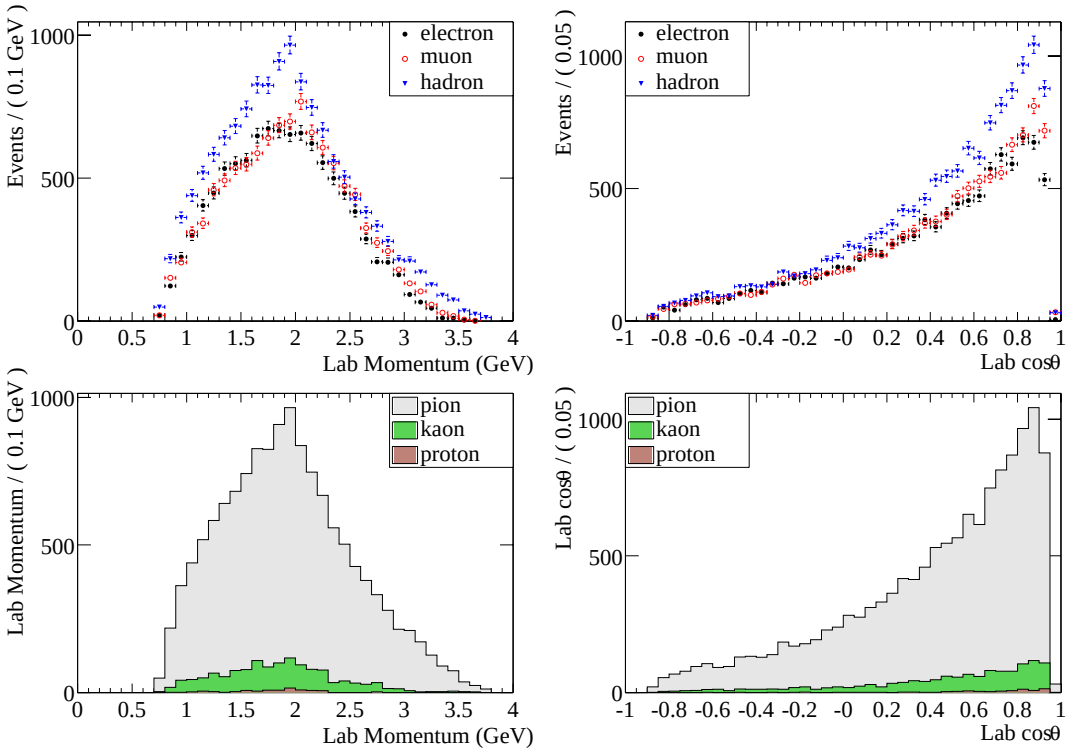

Figure 9.2: Upper plots: distributions of momentum (left) and polar angle (right) in the lab frame of real leptons and real hadrons (according to Monte Carlo information) for the events that pass all the cuts except lepton identification, and the D^* is reconstructed correctly, in generic $B^0\overline{B}^0$ and B^+B^- Monte Carlo samples, without reweighting. Lower plots: The relative π, K and proton contributions to the hadron sample.

the following formula:

$$\delta\epsilon(x \to y) =$$

$$\frac{1}{\sum_s f_s N_s} \sqrt{\sum_{n=1}^{m} \left[\left(\sum_s f_s N_{n;s} \right)^2 \delta^2_{x\to y}(n) + \left(\sum_s f_s^2 N_{n;s} \right) \left(\omega^2_{x\to y}(n) - \epsilon^2(x \to y) \right) \right]} .$$

The first term in the sum over bins ($n = 1$ to m) corresponds to the uncertainty due to the uncorrelated statistical errors $\delta_{x\to y}(n)$ on the efficiency calculated from

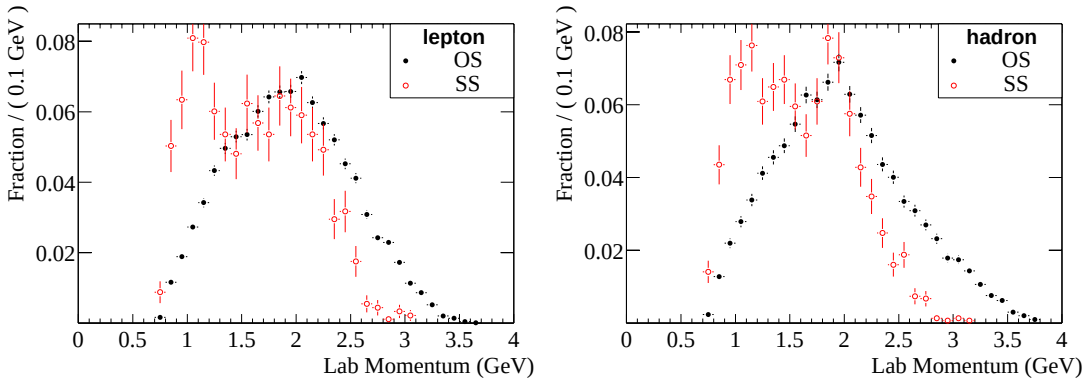

Figure 9.3: Distributions of lab momentum for real lepton (left) and real hadron (right) in generic $B^0\overline{B}^0$ and B^+B^- Monte Carlo samples for opposite-side (OS) and same-side (SS) samples. The areas of the histograms are normalized to one. Event are required to satisfy all selections except lepton identification.

the data control samples. The second term corresponds to the statistical uncertainty due to the Monte Carlo sample used for the angle and momentum weighting. To demonstrate the source of each term, we consider a simple example with only two bins and calculate the error on the efficiency ϵ. Assume the total number of tracks in the Monte Carlo sample is N. The number of events that pass the selection is

$$N_p = N\epsilon = N_1\epsilon_1 + N_2\epsilon_2\,,$$

where N_i is the number of tracks that fall in bin i, and ϵ_i is the efficiency for bin i. The error on N_p will be

$$(\delta N_p)^2 = (\delta N)^2\epsilon^2 + N^2\delta\epsilon^2 = (\delta N_1)^2\epsilon_1^2 + N_1^2\delta\epsilon_1^2 + (\delta N_2)^2\epsilon_2^2 + N_2^2\delta\epsilon_2^2\,.$$

Using $\delta N_i = \sqrt{N_i}$ and $N = N_1 + N_2$, we find that the error on the efficiency is

$$\delta\epsilon^2 = \frac{1}{N^2}(N_1^2\delta\epsilon_1^2 + N_2^2\delta\epsilon_2^2 + N_1\epsilon_1^2 + N_2\epsilon_2^2 - (N_1+N_2)\epsilon^2)\,.$$

When scaling factors f are included, N_i is replaced by fN_i and δN_i is replaced by $f\sqrt{N_i}$.

The candidate leptons and hadrons used here are those with a true reconstructed D^* in the event, according to the Monte Carlo information. In addition, the event is required to pass all D^* ℓ (real or fake lepton) selection criteria, except the lepton identification requirements. The momentum- and angle-weighted efficiencies are shown in Table 9.1.

Table 9.1: Momentum- and angle-weighted lepton identification efficiencies and fake rates for very tight and loose selectors, for leptons and hadrons from $B^0\overline{B}^0$ and B^+B^- simulation events that pass all cuts except lepton identification, and the D^* is reconstructed correctly.

	OS		SS	
Eff.	very tight (%)	loose (%)	very tight (%)	loose (%)
$e \to e$	90.948 ± 0.044	98.587 ± 0.019	90.99 ± 0.13	98.418 ± 0.057
$\pi \to e$	0.1671 ± 0.0087	5.715 ± 0.042	0.1409 ± 0.0071	5.622 ± 0.051
$K \to e$	0.124 ± 0.037	2.59 ± 0.11	0.115 ± 0.046	3.34 ± 0.22
$\mathrm{p} \to e$	0.82 ± 0.19	15.8 ± 1.0	0.53 ± 0.15	14.9 ± 1.1
$\mu \to \mu$	69.60 ± 0.12	86.726 ± 0.095	64.54 ± 0.85	83.59 ± 0.77
$\pi \to \mu$	2.37 ± 0.11	7.26 ± 0.11	2.27 ± 0.34	7.09 ± 0.34
$K \to \mu$	3.052 ± 0.093	8.54 ± 0.14	2.61 ± 0.15	7.67 ± 0.25

For the overall rate of misidentifying a hadron as an electron or muon, $\epsilon(h \to \ell)$, we use the average of the misidentification rates for π, K, and p, weighted by the relative fraction of true π, K, and p candidates predicted by generic Monte Carlo for $B^0\overline{B}^0$ and B^+B^-:

$$\epsilon(h \to \ell) = f_\pi\, \epsilon(\pi \to \ell) + f_K\, \epsilon(K \to \ell) + f_{\mathrm{p}}\, \epsilon(\mathrm{p} \to \ell)\,.$$

The relative fractions f_π, f_K, and f_{p} are given in Table 9.2.

In the final selection in the lifetime and mixing analysis, the electron (muon)

Table 9.2: Relative contributions of π, K and protons to hadron sample used for calculating misidentification probabilities.

Sample	π	K	p	total
$B^0\overline{B}^0 + B^+B^-$	12138	1226	83	13447
Relative Fraction	90.3%	9.1%	0.62%	

sample is required to pass the very tight electron (muon) selector and fail the very tight muon (electron) selector. The fake sample is required to fail both the loose electron and loose muon selectors. We denote the efficiency for species x to satisfy two criteria y and z based on the very tight criteria as $\epsilon_x^{\rm VT}(y;z)$ and loose criteria as $\epsilon_x^{\rm LS}(y;z)$. The relations between these efficiencies and the individual selector efficiencies are

$$\begin{aligned}
\epsilon_e^{\rm VT}(e;!\mu) &= \epsilon^{\rm VT}(e\to e)(1-\epsilon^{\rm VT}(e\to\mu)) \simeq \epsilon^{\rm VT}(e\to e)\\
\epsilon_\mu^{\rm VT}(e;!\mu) &= \epsilon^{\rm VT}(\mu\to e)(1-\epsilon^{\rm VT}(\mu\to\mu)) \simeq 0\\
\epsilon_h^{\rm VT}(e;!\mu) &= \epsilon^{\rm VT}(h\to e)(1-\epsilon^{\rm VT}(h\to\mu))\\
\epsilon_e^{\rm VT}(!e;\mu) &= \epsilon^{\rm VT}(e\to\mu)(1-\epsilon^{\rm VT}(e\to e)) \simeq 0\\
\epsilon_\mu^{\rm VT}(!e;\mu) &= \epsilon^{\rm VT}(\mu\to\mu)(1-\epsilon^{\rm VT}(\mu\to e)) \simeq \epsilon^{\rm VT}(\mu\to\mu)\\
\epsilon_h^{\rm VT}(!e;\mu) &= \epsilon^{\rm VT}(h\to\mu)(1-\epsilon^{\rm VT}(h\to e))\\
\epsilon_e^{\rm LS}(!e;!\mu) &= (1-\epsilon^{\rm LS}(e\to e))(1-\epsilon^{\rm LS}(e\to\mu)) \simeq 1-\epsilon^{\rm LS}(e\to e)\\
\epsilon_\mu^{\rm LS}(!e;!\mu) &= (1-\epsilon^{\rm LS}(\mu\to e))(1-\epsilon^{\rm LS}(\mu\to\mu)) \simeq 1-\epsilon^{\rm LS}(\mu\to\mu)\\
\epsilon_h^{\rm LS}(!e;!\mu) &= (1-\epsilon^{\rm LS}(h\to e))(1-\epsilon^{\rm LS}(h\to\mu))
\end{aligned}$$

Here we assume that $\epsilon(e\to\mu)\simeq 0$ and $\epsilon(\mu\to e)\simeq 0$ for both the very tight and loose selectors.

9.2 Angular Variables

As discussed in Sec. 8.9, the opposite-side and same-side samples are selected, according to angular variables, $\cos\theta^*_{D^*\ell}$, $\cos\theta^*_{B^0,D^*\ell}$, and $\cos\theta^*_{B^0,D^*(-\ell)}$, to disentangle the uncorrelated-lepton background events in the signal sample. Distributions of these variables and correlations between them are shown in Figs. 9.4–9.6 for signal events, D^*-hadron events, and D^*-lepton events in $B^0\overline{B}^0$ Monte Carlo simulation. The regions defining the opposite- and same-side control samples are also shown in the figures.

The upper-right plot in Fig. 9.6 demonstrates why we use the flipped variable $\cos\theta_{B^0,D^*(-\ell)}$ to select events in the uncorrelated-lepton sample: the distribution of this flipped variable in the control sample is similar to the distribution of the unflipped variable $\cos\theta_{B^0,D^*\ell}$ for uncorrelated background in the signal sample, whereas the distribution of $\cos\theta_{B^0,D^*\ell}$ in the control sample is systematically different. In other words, we use the flipped variable in order to select a control sample that is more representative of the background we would like to study.

To determine the contribution from uncorrelated-lepton background, the relative efficiencies for a true signal or a true uncorrelated event to pass the opposite-side or same-side selection criteria are needed. They are obtained from generic Monte Carlo simulation. The result is tabulated in Table 9.3.

The efficiencies for SVT-only events and SVT+DCH events are significantly different so we treat them separately. The efficiency difference between SVT-only and SVT+DCH events originates from the correlation between the $\cos\theta_{D^*\ell}$ distribution and D^* momentum in the semileptonic decay process: the higher the D^* momentum in the $\Upsilon(4S)$ rest frame is, the stronger the $\cos\theta_{D^*\ell}$ is peaked at -1. Since the available energy in the D^* decay is very small (6 MeV), the soft π p_t is highly correlated with the D^* momentum. As a result, $\cos\theta_{D^*\ell}$ is more strongly peaked at -1 for high

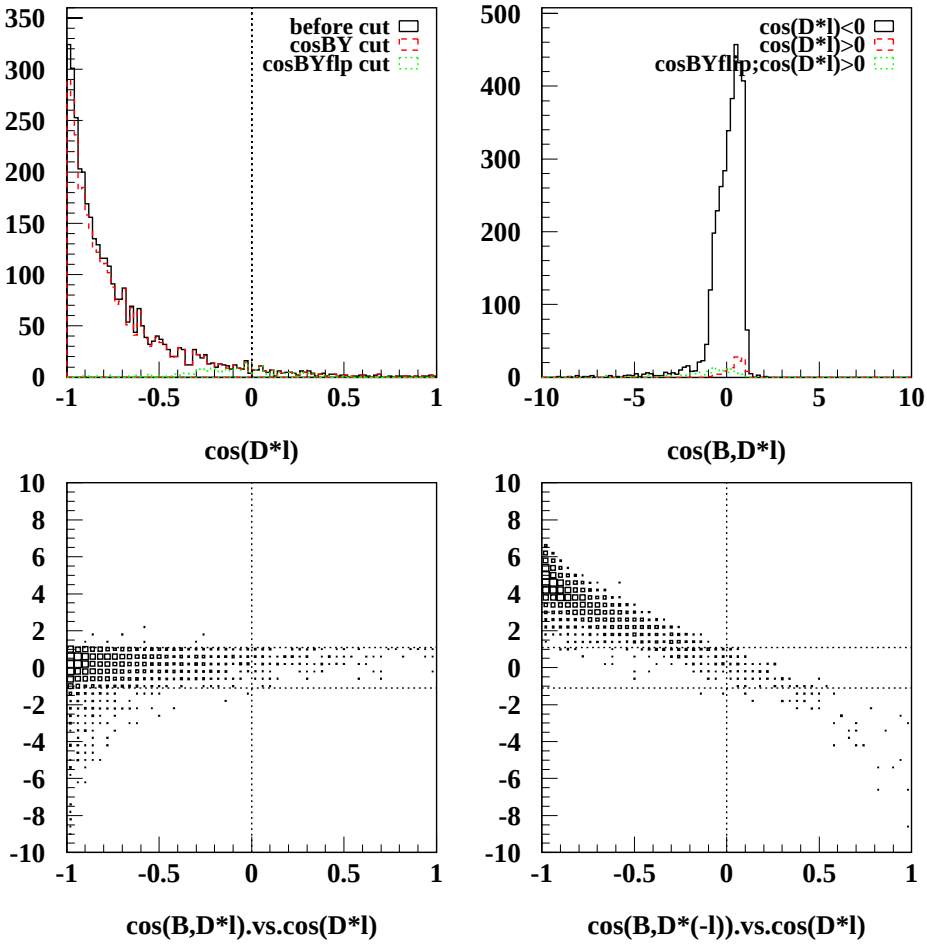

Figure 9.4: Angular variable distributions for signal events. cosBY and cosBYflp cuts mean $|\cos\theta_{B^0,D^*\ell}| < 1.1$ and $|\cos\theta_{B^0,D^*(-\ell)}| < 1.1$, respectively. Selected events for the opposite-side and same-side samples are in the left-middle rectangle in the lower-left and lower-right plots, respectively.

p_t soft π events, which leads to a higher opposite-side efficiency for events whose soft π has DCH hits.

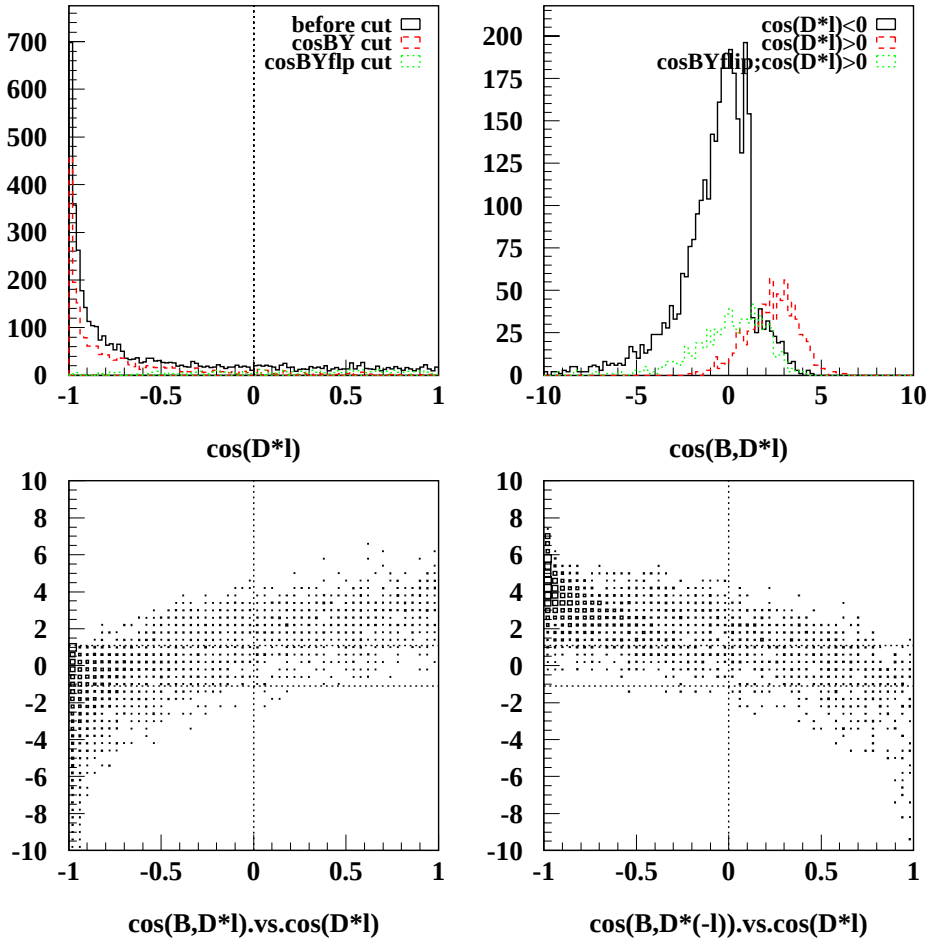

Figure 9.5: Angular variable distributions for D^*-hadron events. cosBY and cosBYflp cuts mean $|\cos\theta_{B^0,D^*\ell}| < 1.1$ and $|\cos\theta_{B^0,D^*(-\ell)}| < 1.1$, respectively. Selected events for the opposite-side and same-side samples are in the left-middle rectangle in the lower-left and lower-right plots, respectively.

9.3 D^*-D^0 Mass Difference

We use the variable δm to calculate signal probability on an event by event basis. In this calculation, the daughter momenta are those from a simultaneous constrained

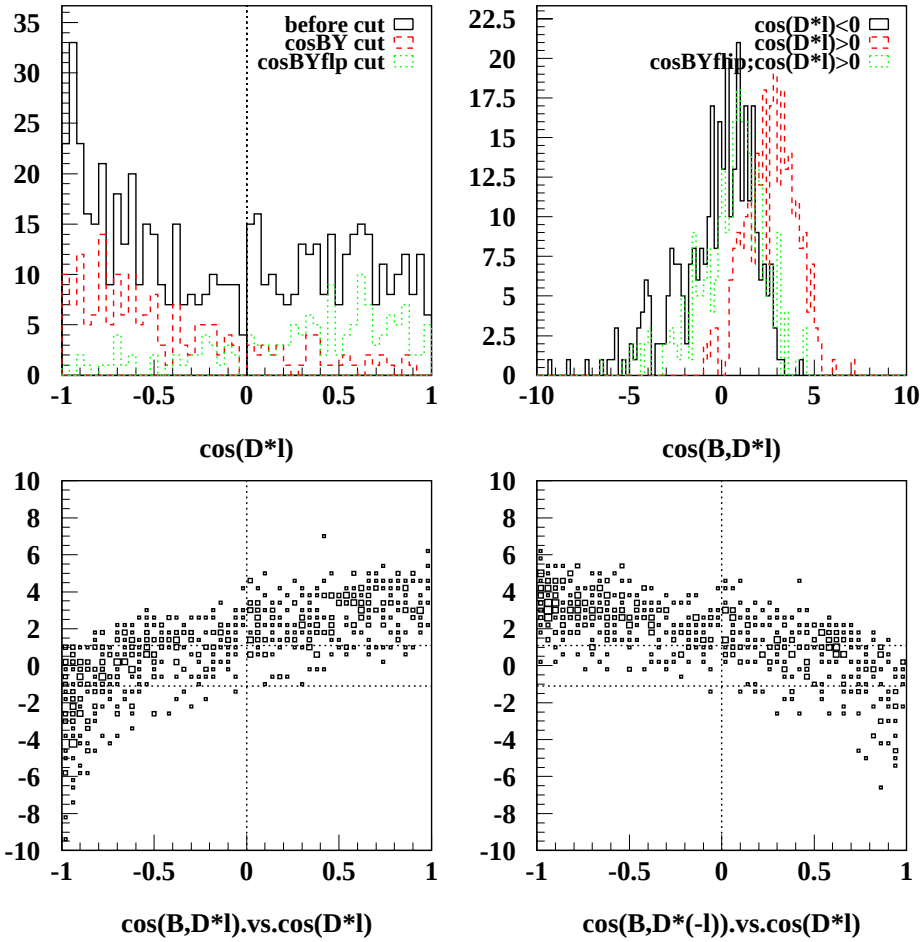

Figure 9.6: Angular variable distributions for uncorrelated lepton events. cosBY and cosBYflp cuts mean $|\cos\theta_{B^0,D^*\ell}| < 1.1$ and $|\cos\theta_{B^0,D^*(-\ell)}| < 1.1$, respectively. Selected events for the opposite-side and same-side samples are in the left-middle rectangle in the lower-left and lower-right plots, respectively.

vertex fit for the $D^*\ell$ candidate that properly accounts for correlations between the D^0 and D^*. [79, 80] The algorithm and its performance are described further in [69, 81]. The D^0 and D^* vertices are fit simultaneously to properly handle correlations between them. The D^0 daughters are constrained to a common D^0 vertex, which

Table 9.3: Relative efficiencies for passing opposite-side or same-side selection criteria, calculated from Monte Carlo simulation events.

Event	opposite-side	same-side
SVT-only; signal	0.9761 ± 0.0012	0.0239 ± 0.0012
SVT+DCH; signal	0.9958 ± 0.0011	0.0042 ± 0.0011
SVT-only; uncorr	0.553 ± 0.015	0.447 ± 0.015
SVT+DCH; uncorr	0.623 ± 0.047	0.377 ± 0.047

is separated explicitly from the $D^*\ell$ candidate by including a constraint for the D^0 flight direction and allowing a D^0 decay length. This constrained fit is used only for the calculation of δm and not for the calculation of the distance between two B decay vertices.

A fit to the δm distribution can determine a) a model for the shape of the peak due to candidates with a correctly reconstructed D^*, b) a model for the shape of the combinatoric background due to candidates with an incorrectly reconstructed D^*, and c) a parameterization of the yields of peak and combinatoric components in each sample. The yields are then used to calculate the fractions of different sources of peaking background, i.e., backgrounds that contains a real D^*. Finally, we will use the calculated amounts and fitted shapes of each background source to estimate the probability of each candidate to be due to signal or each type of background we consider (combinatoric, continuum, fake-lepton, or uncorrelated) in our fits to obtain the lifetime and mixing parameters.

The function used to model the the shape of the peak in a δm distribution is a double Gaussian, which has five free parameters:

$$\mathcal{F}_{\rm peak}(\delta m) = \frac{f}{\sqrt{2\pi}\sigma_1}\exp\left(-\frac{(\delta m-\delta m_1)^2}{2\sigma_1^2}\right)+\frac{1-f}{\sqrt{2\pi}\sigma_2}\exp\left(-\frac{(\delta m-\delta m_2)^2}{2\sigma_2^2}\right)\ . \quad (9.2)$$

The shape of the combinatoric background is modeled with a function

$$\mathcal{F}_{\rm comb}(\delta m) = \frac{1}{N}\left[1 - \exp\left(-\frac{\delta m - \delta m_0}{c_1}\right)\right]\left(\frac{\delta m}{\delta m_0}\right)^{c_2}, \tag{9.3}$$

where N is a normalization constant and δm_0 is the kinematic threshold equal to m_π.

If we are only interested in determining shape parameters, and not absolute yields, we can simply combine these models with one additional parameter $f_{\rm peak}$:

$$\mathcal{F}(\delta m) = f_{\rm peak}\cdot\mathcal{F}_{\rm peak}(\delta m) + (1 - f_{\rm peak})\cdot\mathcal{F}_{\rm comb}(\delta m)\ . \tag{9.4}$$

The value of $f_{\rm peak}$ corresponds to the fraction of peak events in the entire sample being fit (including the sideband).

Once the model $\mathcal{F}(\delta m)$ has been specified, we next need to identify subsamples that will be fit with a single set of shape parameters. At one extreme, we could fit the entire sample (which includes background control samples) at once to a single set of shape parameters. We choose not to do this because we find that different subsamples have significantly different shapes (see Figures 9.7–9.12). Instead, we account for these differences to improve our statistical sensitivity, and more importantly, to reduce possible systematic biases in lifetime and mixing fits. At the other extreme, we could independently fit each of the 360 subsamples we obtain with the full breakdown by the attributes listed in Table 8.1. This approach provides the maximum statistical sensitivity, in principle, but suffers from problems with fitting small samples.

The compromise we have chosen is to divide the sample into six groups that each share a common set of 5 peak shape parameters, `peakGroup` (6) = `svtDch` (2) × `Dmode` (3). We select this breakdown through a process of trial and error to determine which subdivisions of the sample result in statistically significant differences in peak

shape. In particular, we find that the peak shape does not depend on whether an event is recorded on or off the peak (`onOffRes`), whether a candidate passes or fails the angular cuts (`angCut`), or whether the lepton candidate passes or fails lepton ID requirements (`leptID`). Using the same peak shape (within a peak group) for the signal and its control samples will simplify the characterization of peaking backgrounds later.

Within each of the six peak groups, we further subdivide the sample into 12 subgroups that each share a common set of 2 combinatoric background shape parameters. We again use the shape variations observed in data to determine an appropriate breakdown. The on-resonance, opposite-side sample is divided by tagging category (5 choices) and whether the lepton candidate belongs to signal sample or fake lepton sample (2 choices; e and μ samples combined). The remaining two subgroups are all off-resonance events and all same-side events in the on-resonance sample, both of which are smaller than each of the ten on-resonance, opposite-side subgroups.

In order to determine the 5 peak + 12×2 background shape parameters for each peak group, we perform a simultaneous fit of each group using a likelihood

$$\mathcal{L} \equiv \prod_{k=1}^{12} \prod_{\text{evts } j} \left\{ f_{\text{peak}}^{(k)} \cdot \mathcal{F}_{\text{peak}}(\delta m_j; \vec{q}) + (1 - f_{\text{peak}}^{(k)}) \cdot \mathcal{F}_{\text{comb}}(\delta m_j;\, c_1^{(k)}, c_2^{(k)}) \right\} ,$$

$$\vec{q} = (m_1, \sigma_1, m_2, \sigma_2, f_1) ,$$

where each candidate's contribution is calculated with the PDF in this product corresponding to its background group k. With this approach we introduce 12 additional relative normalization parameters $f_{\text{peak}}^{(k)}$ for a total of 5 (peak) + 24 (bkg) + 12 (fractions) = 41 parameters in each simultaneous fit.

Figures 9.7–9.12 show the results of the simultaneous fits to each of the 6 peak groups. Note that these plots only list values for $c_1^{(k)}$, $c_2^{(k)}$, and $f_{\text{peak}}^{(k)}$ even though

a common set of peak shape parameters was also allowed to vary in each fit. The combined data and the sum of the 12 models in each of the six peak groups are shown in Fig. 9.13, in which the peak shape parameters are shown.

Once we have determined the peak and combinatoric background shape distributions expected for all candidates in our sample, we calculate the absolute yields of peak and combinatoric events within each of the 360 subsamples resulting from our most detailed decomposition by `svtDch` (2) × `Dmode` (3) × `onOffRes` (2) × `angCut` (2) × `leptID` (3) × `tagCat` (5). The main advantage of such a detailed breakdown is that it offers the maximum flexibility for later regrouping subsamples. The potential disadvantage is that our procedure involves fits to some categories with very few events (even zero events in some cases).

In fact, we find that these fits are stable and give reasonable results once the shape parameters are fixed. This is not surprising since, to a good approximation, the fits could simply be replaced by counting the number of events inside and outside of the peak region (with some assumption about the extrapolation of the sideband under the peak). The refinement that the fits provide is to account for the variations in the expected peak width and sideband extrapolation based on the actual composition of events in each of the 360 subsamples. For example, consider a sample consisting of a single candidate with $\delta m = 146$ MeV. By assuming an average shape for the peak and combinatoric BG, we would calculate that this event is more likely to be in the peak than the background. With our approach, if in fact the candidate's soft pion is measured in the SVT and DCH, we correctly identify it as being background.

The peak and combinatoric models are combined into an unbinned extended

maximum-likelihood fit[1] with

$$\mathcal{L}(N'_{\rm peak}, N'_{\rm comb}) = \prod_{\rm evts\ j} \left\{ \frac{N_{\rm peak}}{N_{\rm tot}} \cdot f_{\rm peak}(\delta m_j) + \frac{N_{\rm comb}}{N_{\rm tot}} \cdot f_{\rm comb}(\delta m_j) \right\} \cdot P(N_{\rm tot}) \ ,$$

where $N_{\rm tot} \equiv N_{\rm peak} + N_{\rm comb}$ and $P(N_{\rm tot})$ is the Poisson probability of the observed number of events when $N_{\rm tot}$ are expected. The actual fit parameters in this likelihood, $N'_{\rm peak}$ and $N'_{\rm comb}$, are yields calculated within a signal window $(\delta m_1, \delta m_2)$

$$N'_{\rm peak} \equiv N_{\rm peak} \cdot \int_{\delta m_1}^{\delta m_2} f_{\rm peak}(\delta m) d(\delta m) \quad , \quad N'_{\rm comb} \equiv N_{\rm comb} \cdot \int_{\delta m_1}^{\delta m_2} f_{\rm comb}(\delta m) d(\delta m) \ .$$

The choice of signal window does not affect the subsequent analysis since we calculate per-event signal probabilities over the full range of δm. We use 143–148 MeV for all categories of events.

9.4 Combinatoric Background

We define combinatoric background as events in which the selected D^* candidate is either not a real D^* or else is misreconstructed. This background then corresponds to the extrapolation of the upper δm sideband to the region under the peak, and we determine its contribution to each subsample directly from the combinatoric yield parameter, $N'_{\rm comb}$, in the δm fits described in the previous section.

[1] All modeling and fitting in this analysis are using the **RooFit** packages [82, 83, 84], which are developed by D. Kirkby and W. Verkerke in the ROOT [85] environment. The core fitting engine is MINUIT [86].

9.5 Continuum Background

The continuum background contribution ($c\overline{c} \to D^*X$) can be estimated using off-resonance data. We scale the peak contribution by the ratio of on-resonance luminosity to off-resonance luminosity, which is approximately 7.9. The fractions of events in the δm peak that are due to continuum background are shown as $f_{\rm cont}$ in Table 9.4.

9.6 Fake Lepton Background

After subtracting continuum contributions from off-resonance data, the remaining portion in the peak region of the δm distributions will contains events with a real D^* and a real or misidentified lepton from $B\overline{B}$ events, The observed numbers of candidates in the D^* peak, after continuum subtraction, in the lepton signal and fake control samples (N_e, N_μ, N_f) are related to the numbers of D^* ℓ candidates where the lepton candidate is a true lepton or a true hadron $(\eta_e, \eta_\mu, \eta_h)$ through a set of linear equations:

$$\begin{pmatrix} N_e \\ N_\mu \\ N_f \end{pmatrix} = \begin{pmatrix} \epsilon_e(e;!\mu) & \epsilon_\mu(e;!\mu) & \epsilon_h(e;!\mu) \\ \epsilon_e(!e;\mu) & \epsilon_\mu(!e;\mu) & \epsilon_h(!e;\mu) \\ \epsilon_e(!e;!\mu) & \epsilon_\mu(!e;!\mu) & \epsilon_h(!e;!\mu) \end{pmatrix} \begin{pmatrix} \eta_e \\ \eta_\mu \\ \eta_h \end{pmatrix}. \tag{9.5}$$

The true numbers of events $(\eta_e, \eta_\mu, \eta_h)$ can easily be found by inverting the matrix. The fake fractions in the signal and control samples are therefore

$$\begin{aligned} f_e &= \eta_h \epsilon_h(e; !\mu)/N_e \\ f_\mu &= \eta_h \epsilon_h(!e; \mu)/N_\mu \\ f_f &= \eta_h \epsilon_h(!e; !\mu)/N_f \end{aligned}$$

This calculation of fake fractions is done with the lepton and fake yields extracted from each of the 60 separate subdivisions of the on-resonance data: `angCut` (2) $\otimes$ `svtDch` (2) $\otimes$ `Dmode` (3) $\otimes$ `tagCat` (5). We use two different sets of efficiencies $\epsilon_x(y; z)$ for opposite-side (OS) and same-side (SS) sample, and assume they are valid for all samples with the same `angCut`.

The fractions of events in the δm peak that are due to fake leptons are shown as $f_{\rm fake}$ in Table 9.4.

9.7 Uncorrelated Lepton Background

After subtracting the fake and continuum peaking background contributions, the remaining events in the δm peak in on-resonance data come from $B^0\overline{B}^0$ and B^+B^- events. To estimate the uncorrelated background, defined as a real D^* from one B and a real lepton from the other B, we use the same-side control samples and the relative efficiencies for a true signal or a true uncorrelated event to pass the opposite-side or same-side selection criteria from generic Monte Carlo simulation (see Sec. 9.2).

The measured yields of opposite-side ($N_{\rm OS}$) and same-side ($N_{\rm SS}$) events, after subtracting fake and continuum fractions, are related to the true number of signal

($\eta_{\rm sig}$) and uncorrelated background ($\eta_{\rm unc}$) events by

$$\begin{pmatrix} N_{\rm OS} \\ N_{\rm SS} \end{pmatrix} = \begin{pmatrix} \epsilon_{\rm OS}^{\rm sig} & \epsilon_{\rm OS}^{\rm unc} \\ \epsilon_{\rm SS}^{\rm sig} & \epsilon_{\rm SS}^{\rm unc} \end{pmatrix} \begin{pmatrix} \eta_{\rm sig} \\ \eta_{\rm unc} \end{pmatrix} . \tag{9.6}$$

After solving for ($\eta_{\rm sig}$, $\eta_{\rm unc}$), the uncorrelated background fractions can be expressed as

$$\begin{aligned} f_{\rm OS} &= \epsilon_{\rm OS}^{\rm unc}\eta_{\rm unc}/N_{\rm OS} \\ f_{\rm SS} &= \epsilon_{\rm SS}^{\rm unc}\eta_{\rm unc}/N_{\rm SS} \end{aligned}$$

Note that although we normalize the relative efficiencies for passing the OS or SS selections so that the sum of the efficiencies is 1, the resulting background fractions do not depend on this normalization. In other words, we are depending on the Monte Carlo for the relative efficiencies within each class of events (signal or uncorrelated lepton) but not the absolute efficiencies or the relative efficiencies between the two classes of events.

This calculation of uncorrelated background fractions is done with the OS and SS yields extracted from each of the 90 separate subdivisions of the data: `leptID` (3) $\otimes$ `svtDch` (2) $\otimes$ `Dmode` (3) $\otimes$ `tagCat` (5). We have assumed that one efficiency matrix is valid for all SVT-only samples and one for all SVT+DCH samples.

The fractions of events in the δm peak that are due to uncorrelated leptons are shown as $f_{\rm uncorr}$ in Table 9.4.

9.8 Sample Composition

In order to summarize the results of the procedures described here, we regroup the 360 subsamples used for the final 2-parameter yield fits into a smaller number of "supersamples" for the purposes of display. The curves in the resulting plots are obtained by summing the results of the individual fits that contribute to each supersample.

Figures 9.14 and 9.15 show the combined results when the sample is divided according to `angCut`, `onOffRes` and `leptID`. Table 9.4 summarizes the background fractions calculated in each of these 12 supersamples. Figures A.1–A.3 summarize the results of the background studies when the sample is divided according to `svtDch`, `Dmode`, and `tagCat`, respectively. The supersamples used in Figures A.1–A.3 are signal samples only (i.e., they do not include any background control samples) and show electron and muon results separately.

We also study the background contribution for unmixed, mixed and untagged events separately. The uncorrelated background for mixed events is significantly higher than for others. The reason is that for mixed events the probability of getting a lepton with correct charge from different B is higher since two B decay as the same flavor. The combinatoric background for mixed events is also higher for the similar reason. The mixed events tend to have more D^* with the same sign so that there are more combinations to form a D^* candidate from their daughter particles. We will take into account the differences of background levels among events with different tagging information for lifetime and mixing fits.

Table 9.4: Peak yields and fake, continuum and uncorrelated fractions of the peak yield, and the combinatoric fraction of total events in δm window for the various subsamples shown in Figs. 9.14 and 9.15. Peak yields include the peaking backgrounds.

Category			Peak Yield	$f_{\text{cont.}}(\%)$	$f_{\text{fake}}(\%)$	$f_{\text{uncorr.}}(\%)$	$f_{\text{comb.}}(\%)$
On	OS	e	7008 ± 91	1.53 ± 0.42	0.1678 ± 0.0042	3.14 ± 0.39	17.89 ± 0.24
		μ	6569 ± 88	2.27 ± 0.57	2.669 ± 0.067	2.85 ± 0.48	18.36 ± 0.25
		f	8770 ± 108	12.8 ± 1.3	72.4 ± 1.8	0.7 ± 1.6	31.40 ± 0.24
	SS	e	306 ± 21	0.0 ± 0.0	0.533 ± 0.039	56.9 ± 7.0	34.0 ± 1.3
		μ	299 ± 20	5.1 ± 3.6	8.89 ± 0.64	48.9 ± 8.0	34.4 ± 1.3
		f	1350 ± 45	20.4 ± 4.1	74.4 ± 5.4	3.6 ± 7.8	42.59 ± 0.61
Off	OS	e	13.6 ± 3.8	100	0	0	21.0 ± 4.5
		μ	18.9 ± 4.7	100	0	0	25.1 ± 4.3
		f	142 ± 14	100	0	0	44.2 ± 1.8
	SS	e	0.0 ± 0.0	100	0	0	100 ± 100
		μ	1.9 ± 1.4	100	0	0	38 ± 14
		f	34.9 ± 7.1	100	0	0	52.5 ± 3.6

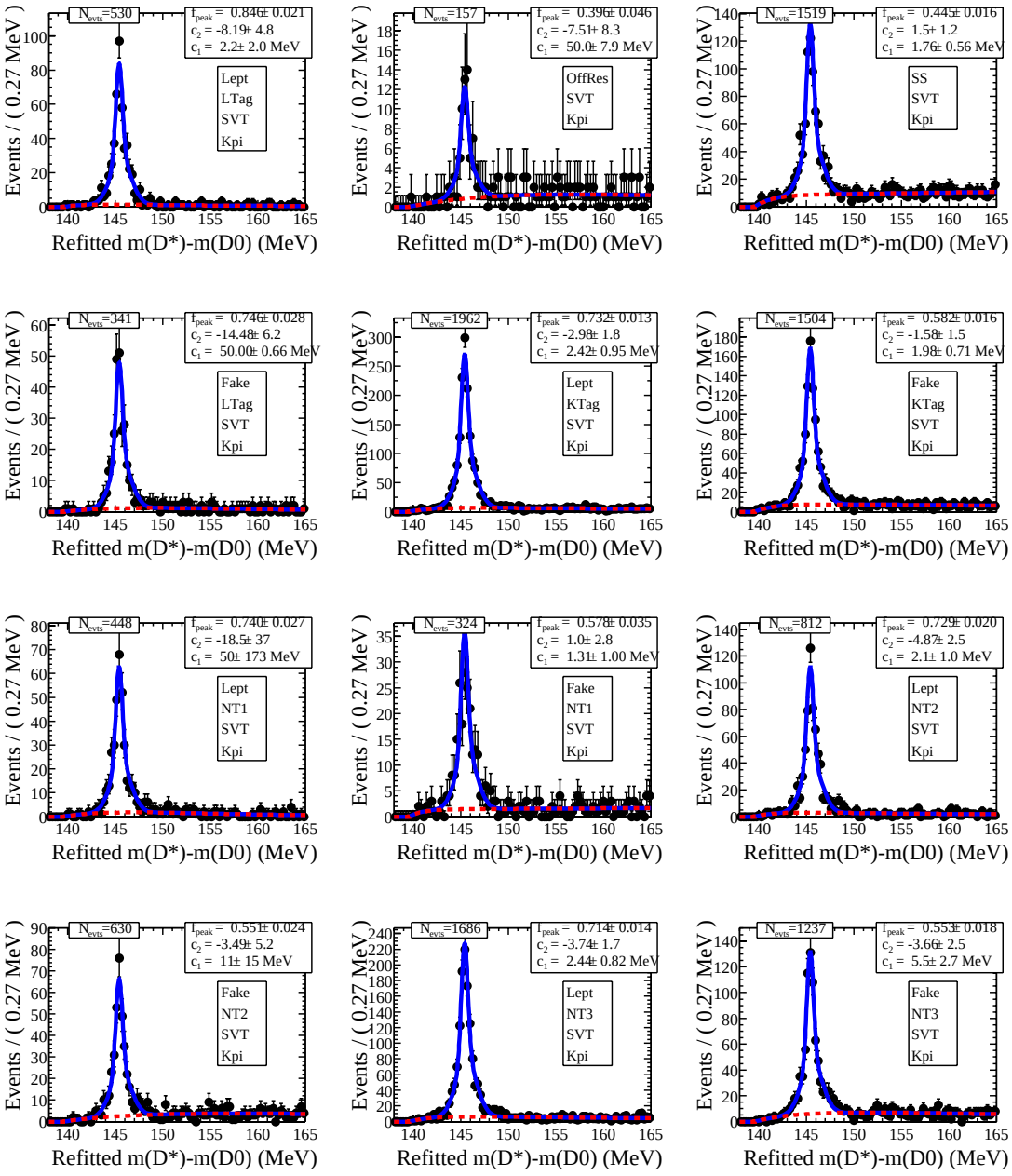

Figure 9.7: Results of the simultaneous fit described in the text to the peak group consisting of candidates reconstructed as $D^0 \to K\pi$ and their soft pion reconstructed in the SVT only. The dashed curves show the fitted contributions of combinatoric background.

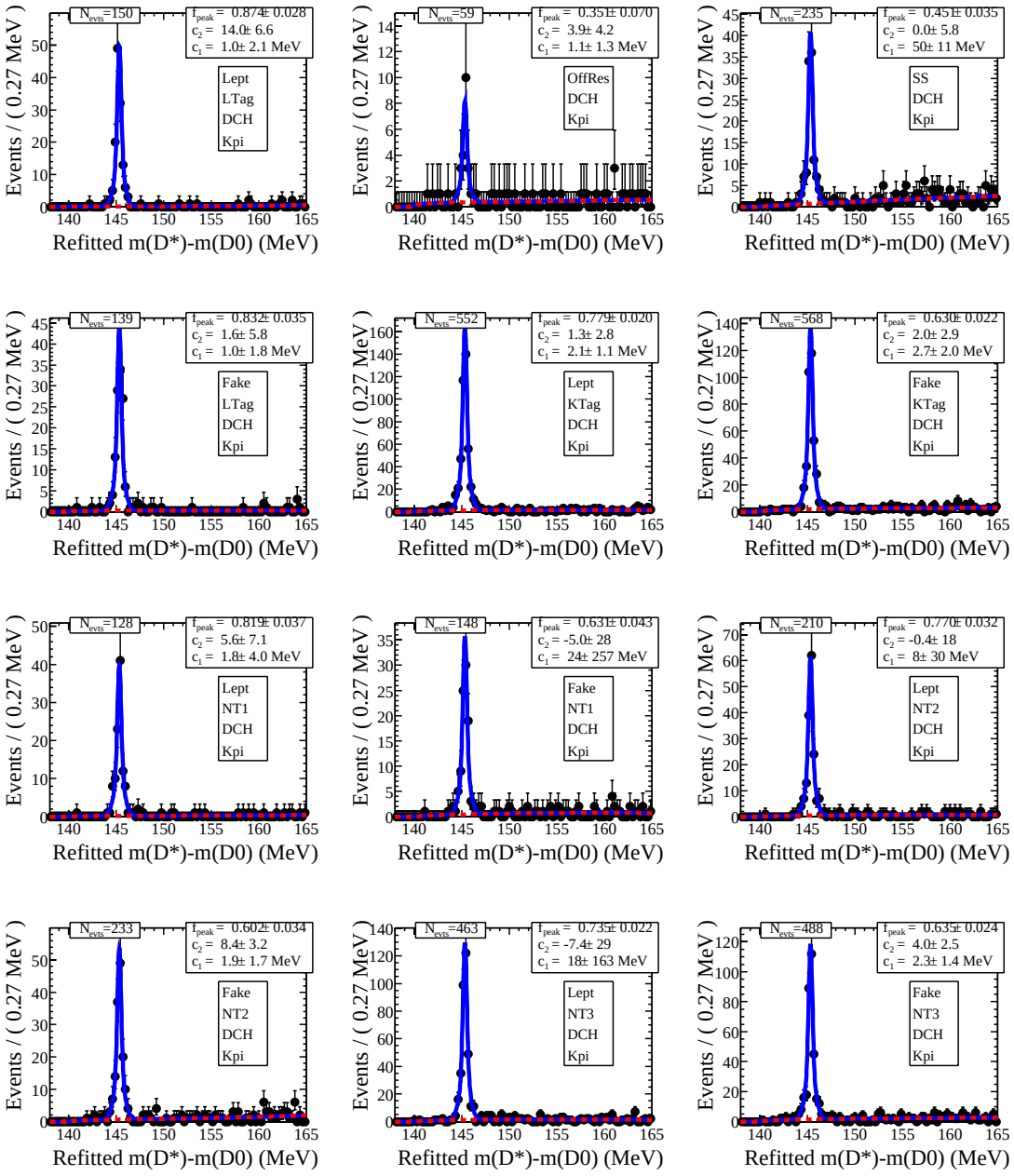

Figure 9.8: Results of the simultaneous fit described in the text to the peak group consisting of candidates reconstructed as $D^0 \to K\pi$ and their soft pion reconstructed in the SVT and DCH. The dashed curves show the fitted contributions of combinatoric background.

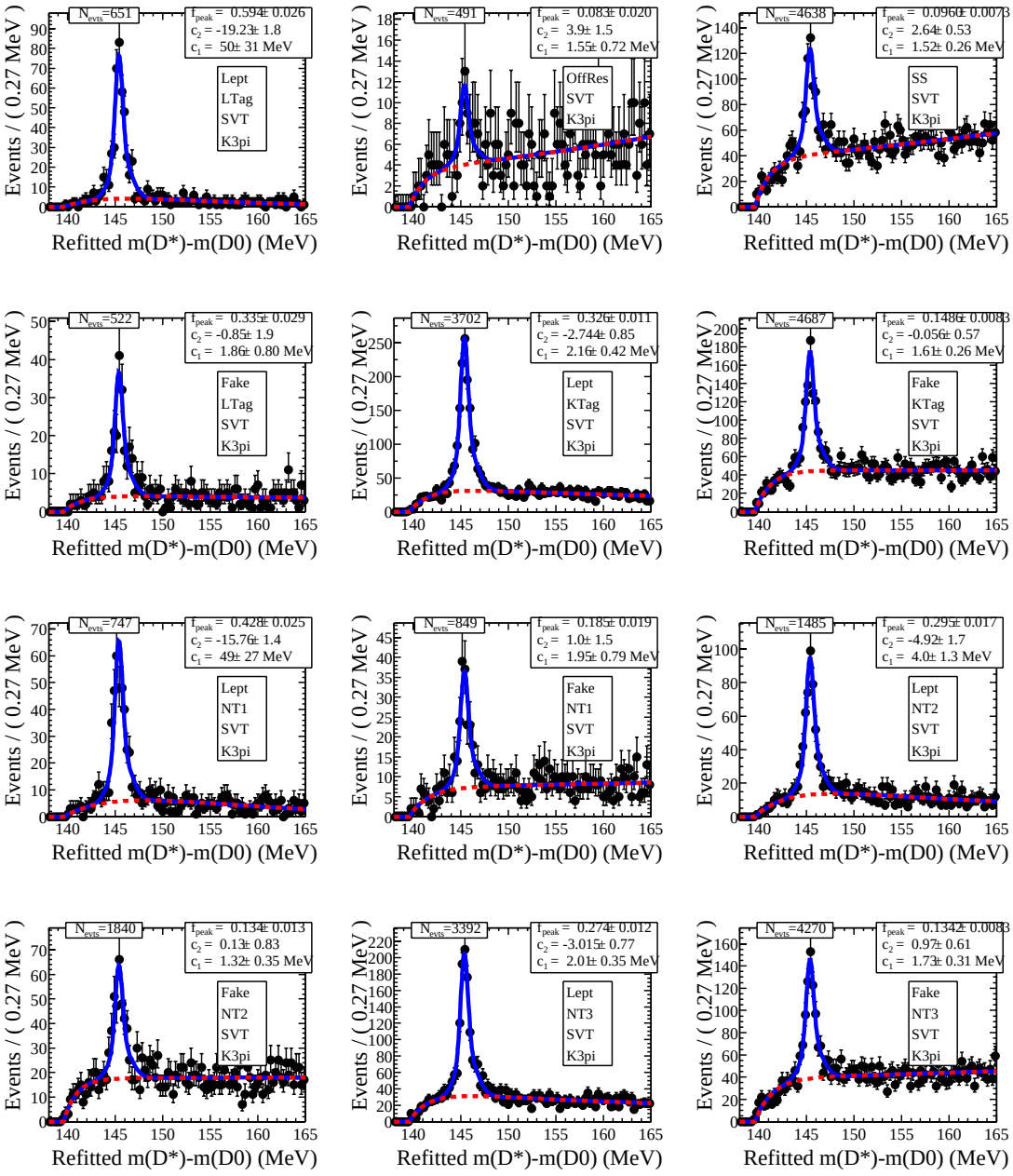

Figure 9.9: Results of the simultaneous fit described in the text to the peak group consisting of candidates reconstructed as $D^0 \to K3\pi$ or $D^0 \to K_S\pi\pi$ and their soft pion reconstructed in the SVT only. The dashed curves show the fitted contributions of combinatoric background.

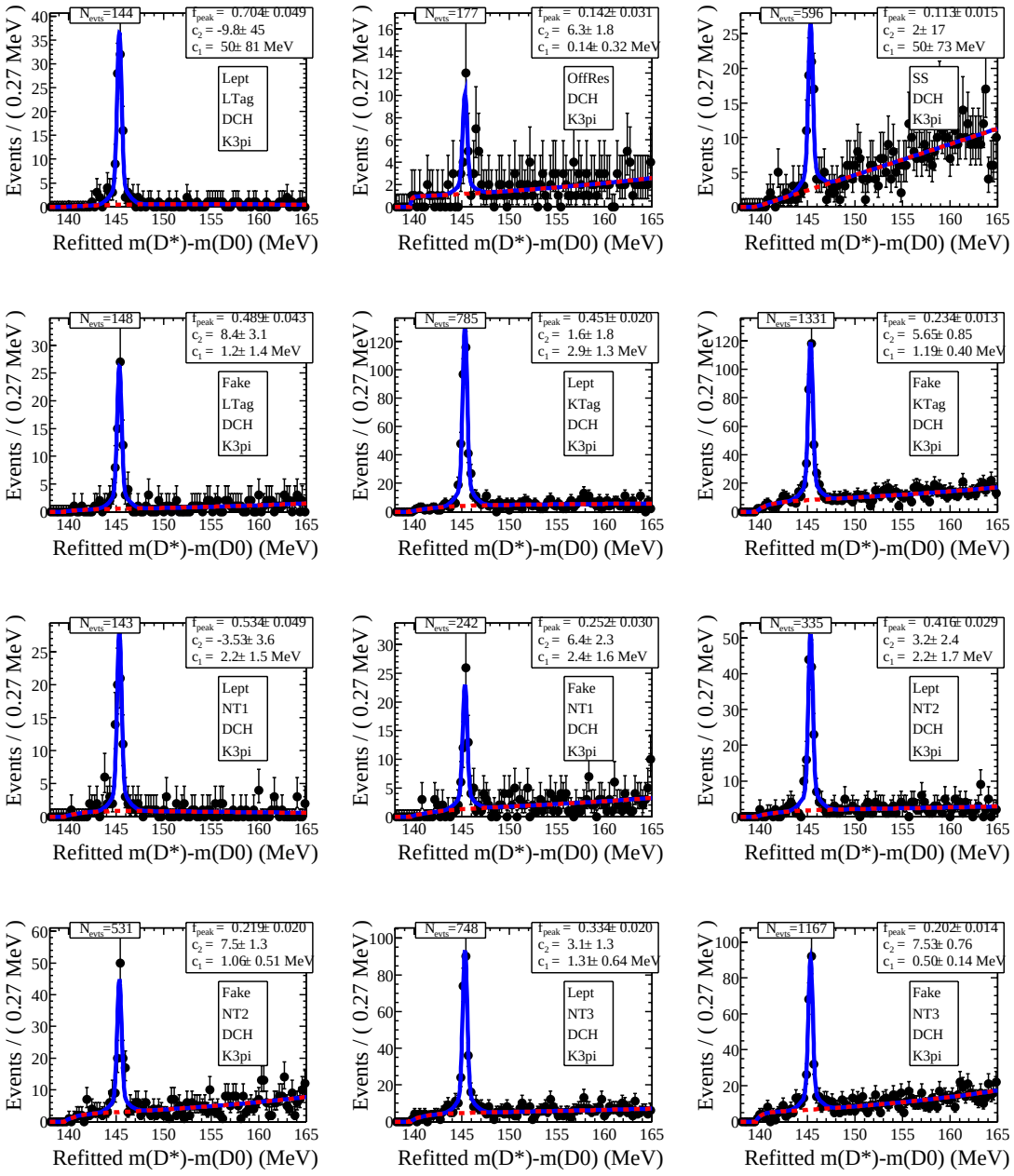

Figure 9.10: Results of the simultaneous fit described in the text to the peak group consisting of candidates reconstructed as $D^0 \to K3\pi$ or $D^0 \to K_S\pi\pi$ and their soft pion reconstructed in the SVT and DCH. The dashed curves show the fitted contributions of combinatoric background.

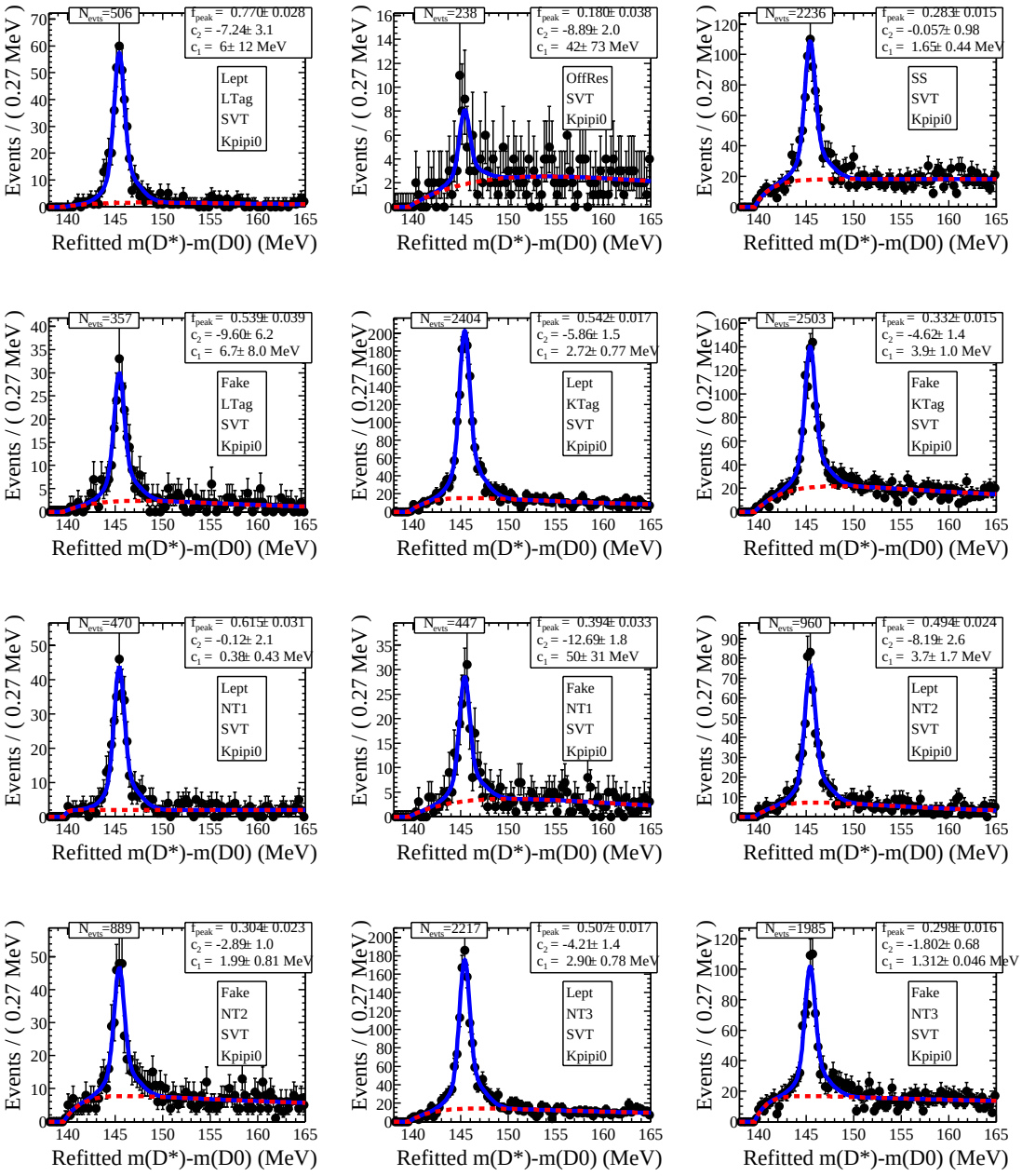

Figure 9.11: Results of the simultaneous fit described in the text to the peak group consisting of candidates reconstructed as $D^0 \to K\pi\pi^0$ and their soft pion reconstructed in the SVT only. The dashed curves show the fitted contributions of combinatoric background.

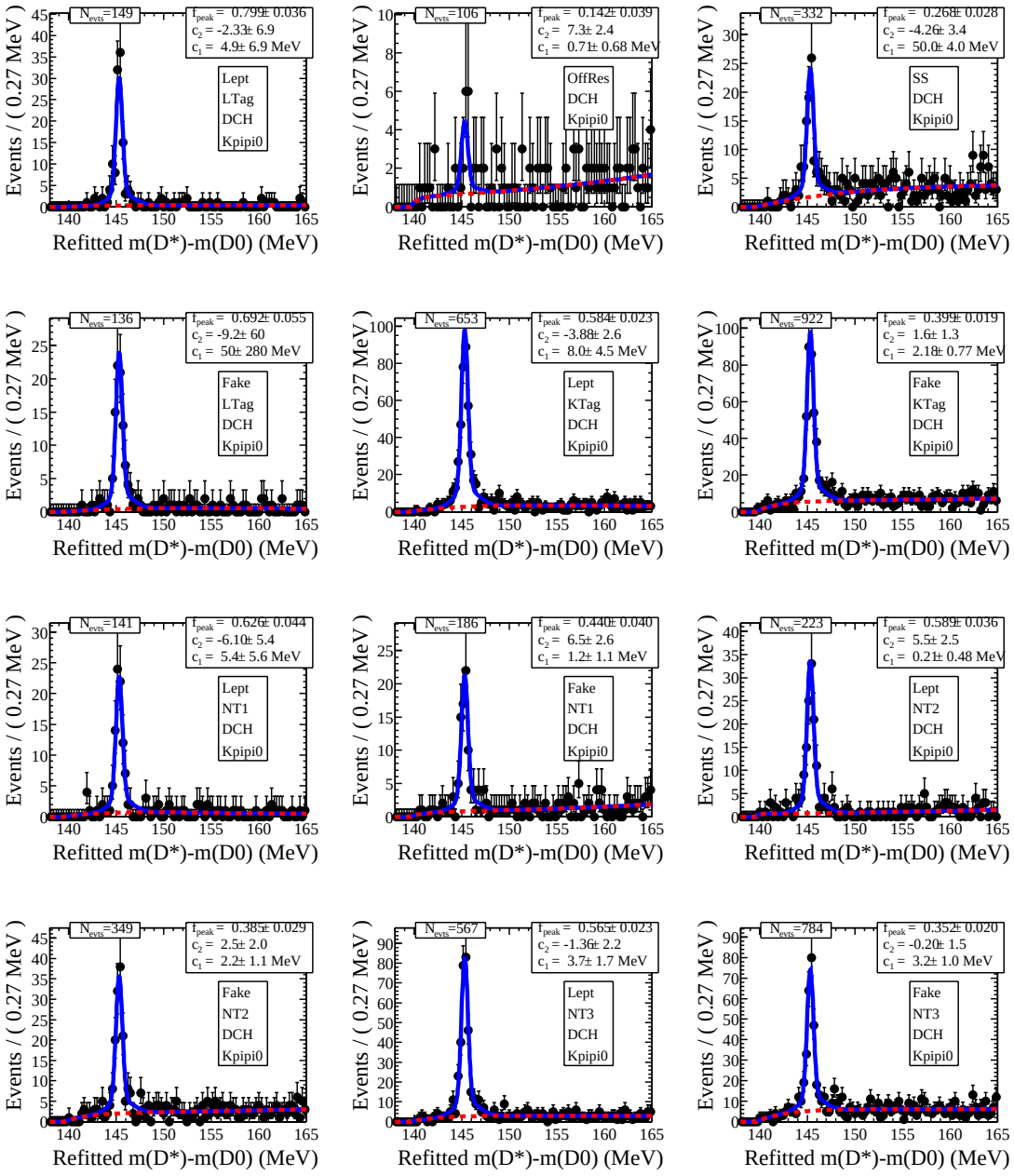

Figure 9.12: Results of the simultaneous fit described in the text to the peak group consisting of candidates reconstructed as $D^0 \to K\pi\pi^0$ and their soft pion reconstructed in the SVT and DCH. The dashed curves show the fitted contributions of combinatoric background.

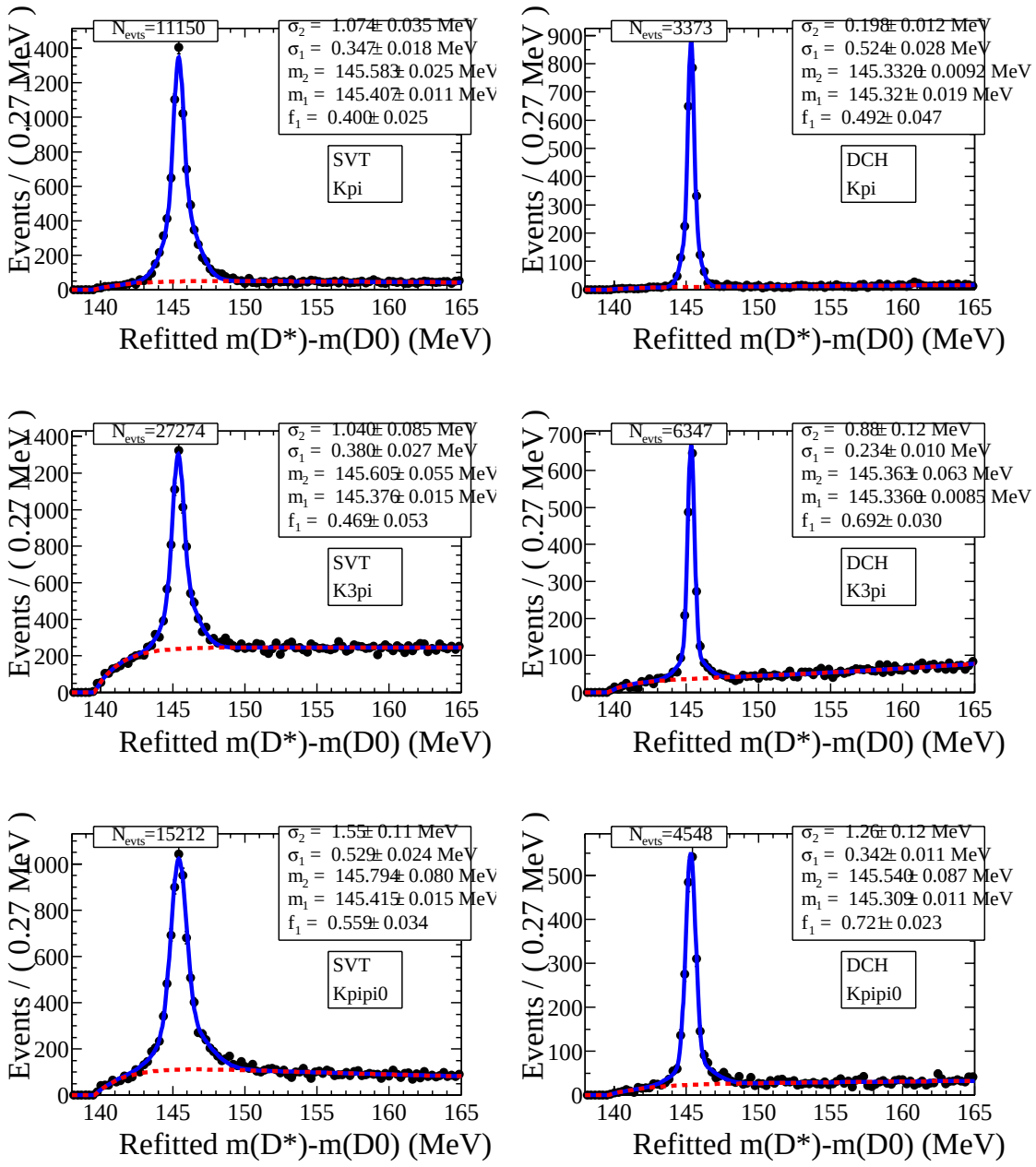

Figure 9.13: Results of the simultaneous fit described in the text to each of the 6 peak groups over the peak region 143–149 MeV. The dashed curves show the fitted contributions of combinatoric background.

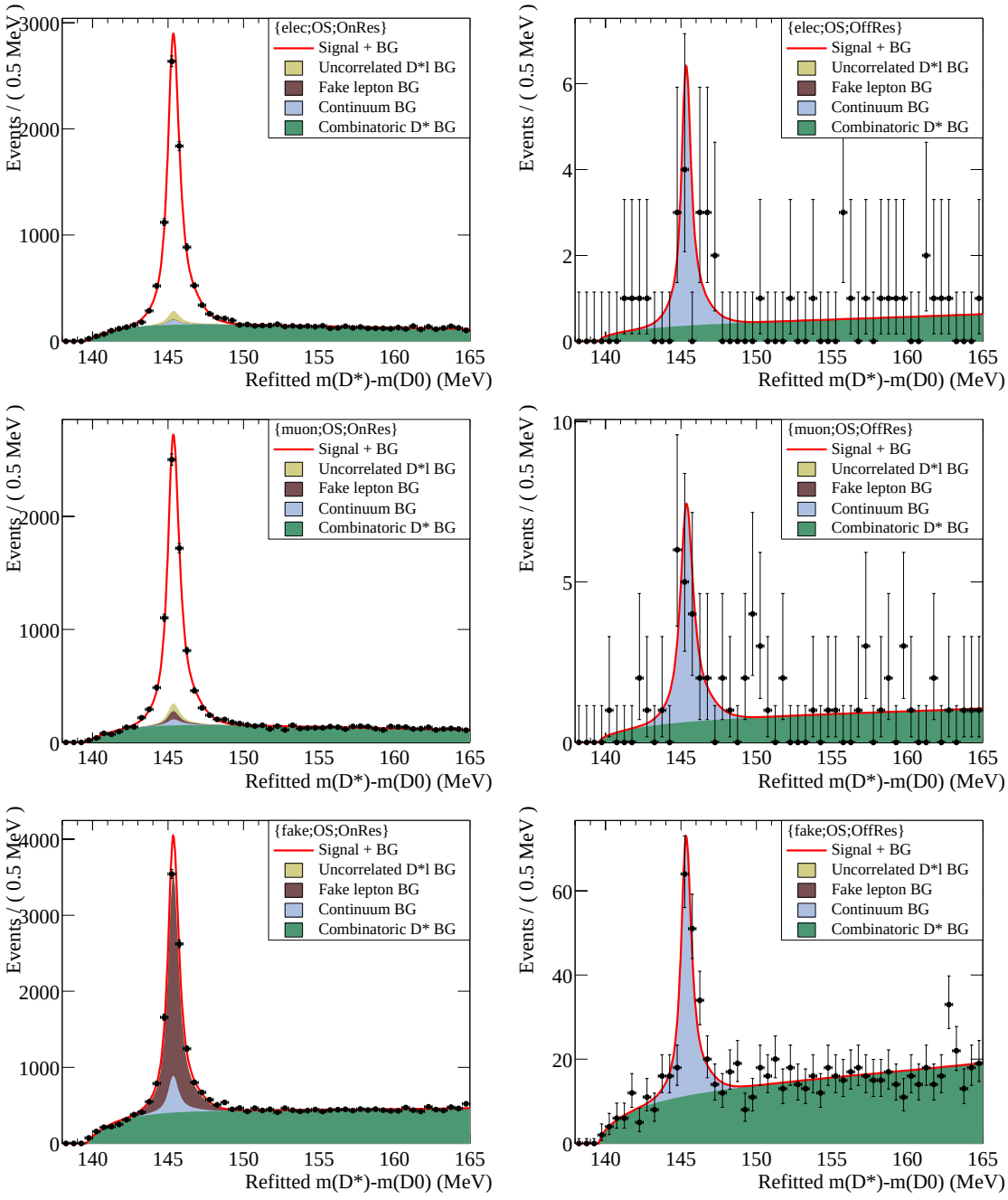

Figure 9.14: Uncorrelated D^* ℓ, fake ℓ and continuum D^* X, and fake D^* contributions for Opposite-Side samples. From top to bottom: electron, muon and fake control samples; left column: on-resonance and right column: off-resonance.

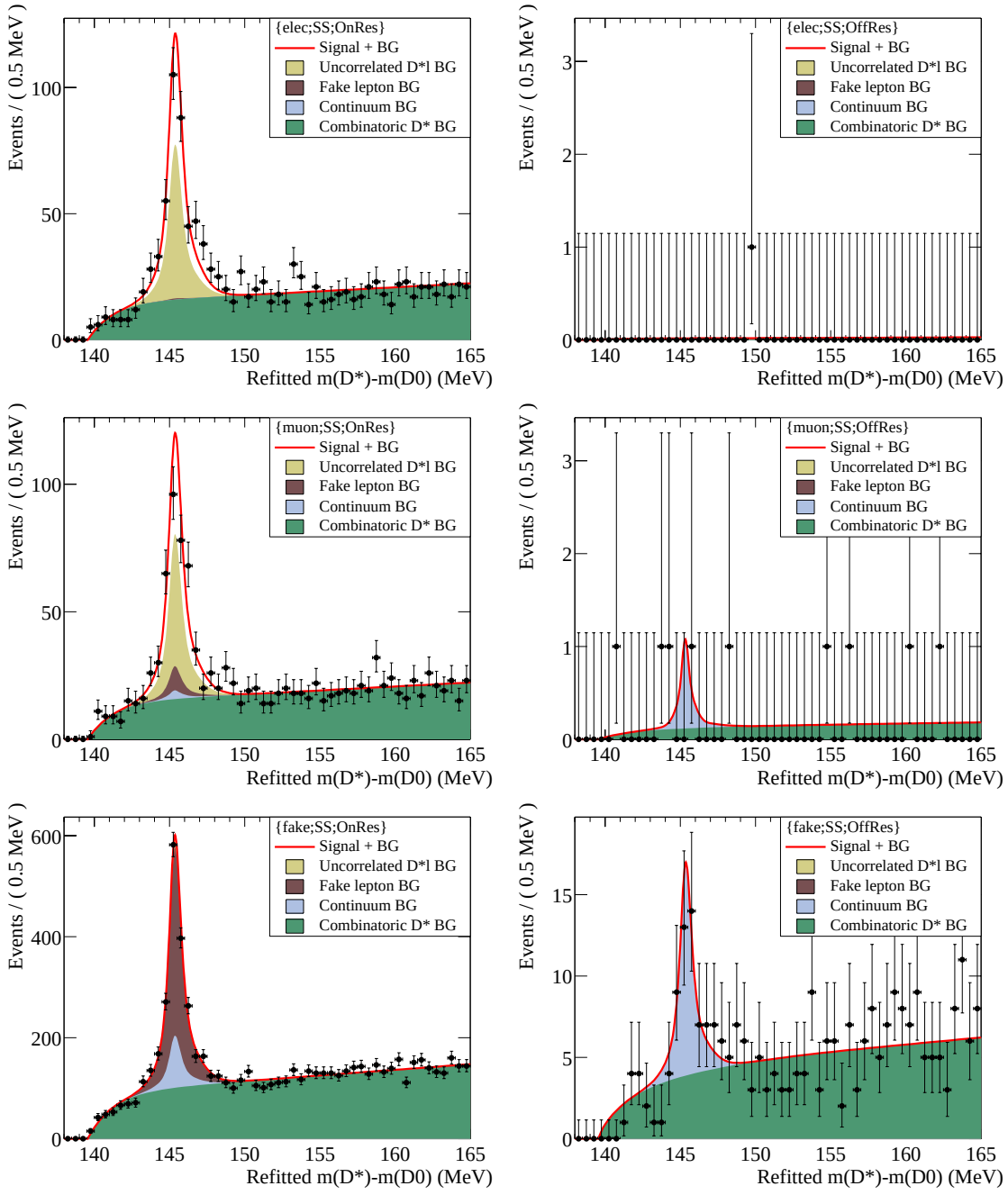

Figure 9.15: Uncorrelated D^* ℓ, fake ℓ, continuum D^* X, and fake D^* contributions for Same-Side samples. From top to bottom: electron, muon and fake control samples; left column: on-resonance and right column: off-resonance.

Chapter 10

Analysis of B^0 Decay Time Evolution

In this chapter, the analysis procedure leading to final fitting is described, including the background characterization, fit model building and fit results.

After the signal and background control samples are selected, combinatoric and peaking background probabilities of each event are determined in a procedure described in Ch. 9. The lifetime τ_{B^0} and oscillation frequency Δm_d are measured simultaneously with an unbinned maximum likelihood fit [82, 83, 84] that use, for each event, the measured decay time difference Δt, the calculated uncertainty on Δt, the combinatoric and peaking background probabilities, and tagging information from the rest of the event. The resulting statistical uncertainties are approximately 0.02 ps for the lifetime measurement and about 0.02 ps^{-1} for the oscillation frequency. For comparison, the uncertainty on the world average in the year 2000 [9] is 0.032 ps for τ_{B^0} and 0.017 ps^{-1} for Δm_d. The challenge in this analysis is to reduce the systematic uncertainties so that they do not dominate the total uncertainty. In order to do so, the types and sizes of backgrounds, the Δt resolution function for signal and the Δt distributions for combinatoric and peaking backgrounds have to be understood.

There are several advantages to fitting the B^0 lifetime and oscillation frequency simultaneously.

1. The expected precision for both τ_{B^0} and Δm_d is not much larger than that of world average [9] in the year 2000. Therefore, there is little or no benefit from fixing either values to the world average, depending on their correlation that is caused by the resolution on Δt.

2. Mixing information improves the measurement of τ_{B^0} in two distinct ways:

   (a) Since mixed and unmixed events have different Δt distributions, the mixing information for each event gives greater sensitivity to the Δt resolution function and a smaller statistical uncertainty on τ_{B^0}. The improvement is about 15%.

   (b) Since B^+B^- and $B^0\overline{B}^0$ events have different mixing behavior, we can use the Δt distributions for mixed and unmixed events to help discriminate between $B^0\overline{B}^0$ signal events and peaking B^+B^- background events.

10.1 Resolution Models

10.1.1 Function forms

A resolution model $\mathcal{R}$ specifies the expected distribution of residuals $\delta\Delta t \equiv \Delta t_{\rm meas} - \Delta t_{\rm true}$ for an event with measured $\Delta t_{\rm meas}$ and calculated vertex error $\sigma_{\Delta t}$. We assume that the resolution is independent of $\Delta t_{\rm meas}$, but expect a strong correlation with $\sigma_{\Delta t}$; i.e., $\mathcal{R} = \mathcal{R}(\delta\Delta t, \sigma_{\Delta t})$.

As discussed in Sec. 6.3, the mean and RMS of the $\delta\Delta t$ distribution strongly depend on $\sigma_{\Delta t}$. In order to accommodate these effects, we consider two different models for the "core" of the resolution function (loosely defined as $|\delta\Delta t| < 5 \cdot \sigma_{\Delta t}$).

The first core model consists of a double Gaussian whose biases and widths are scaled by the measured $\sigma_{\Delta t}$:

$$\mathcal{R}_{\mathrm{G+G}}(\delta\Delta t, \sigma_{\Delta t};\, b_1, b_2, s_1, s_2, f) \equiv f \cdot G(\delta\Delta t;\, b_1\sigma_{\Delta t}, s_1\sigma_{\Delta t}) + (1-f)\cdot G(\delta\Delta t;\, b_2\sigma_{\Delta t}, s_2\sigma_{\Delta t})\ .$$

The second model is the sum of a single Gaussian and the same Gaussian convoluted with a one-sided exponential $E(\delta\Delta t;\, \kappa)$:

$$\mathcal{R}_{\mathrm{GExp}}(\delta\Delta t, \sigma_{\Delta t};\, \kappa, s, f) \equiv f \cdot G(\delta\Delta t;\, 0, s\,\sigma_{\Delta t}) + (1-f)\cdot \frac{1}{2\,\kappa\sigma_{\Delta t}} \int_{-\infty}^{0} du\ \exp\bigl(+u/(\kappa\,\sigma_{\Delta t})\bigr)\cdot G(u-\delta\Delta t;\, 0, s\,\sigma_{\Delta t})\ . \quad (10.1)$$

The decay constant κ is introduced to account for residual charm decay products included in the B_{tag} vertex. The integral in this definition can be performed analytically to give

$$\mathcal{R}_{\mathrm{GExp}}(\delta\Delta t, \sigma_{\Delta t};\, \kappa, s, f) = f\, G(\delta\Delta t;\, 0, s\,\sigma_{\Delta t}) + (1-f)\,\frac{1}{2\,\kappa\sigma_{\Delta t}}\ \exp\left(\frac{s^2 + 2\,\delta\Delta t\cdot\kappa/\sigma_{\Delta t}}{2\kappa^2}\right)\ \mathrm{erfc}\left(\frac{s^2+\delta\Delta t\cdot\kappa/\sigma_{\Delta t}}{\sqrt{2}\,s\,\kappa}\right)\ .$$

Note that the widths of Gaussian functions in G+G model and GExp models, and the decay constant in GExp model are scaled by $\sigma_{\Delta t}$.

Once the parameters of a resolution model are known, we can calculate the corresponding moments of the residual distribution directly in terms of the moments of the per-event error distribution. For the G+G model, the first and second moments

are

$$\langle \delta\Delta t \rangle_{\rm G+G} = \left(f b_1 + (1-f) b_2 \right) \cdot \langle \sigma_{\Delta t} \rangle \tag{10.2}$$

$$\langle (\delta\Delta t)^2 \rangle_{\rm G+G} = \left(f(s_1^2 + b_1^2) + (1-f)(s_2^2 + b_2^2) \right) \cdot \langle (\sigma_{\Delta t})^2 \rangle \ , \tag{10.3}$$

from which we calculate the G+G RMS squared

$$\begin{aligned} (\delta\Delta t)^2_{\rm RMS} = & \left(f s_1^2 + (1-f) s_2^2 \right) \cdot \langle (\sigma_{\Delta t})^2 \rangle + f(1-f) \left(b_1 - b_2 \right)^2 \cdot \langle \sigma_{\Delta t} \rangle^2 \\ & + \left(f b_1^2 + (1-f) b_2^2 \right) \cdot (\sigma_{\Delta t})^2_{\rm RMS} \ . \end{aligned}$$

For comparison, the first and second moments of the GExp model are

$$\langle \delta\Delta t \rangle_{\rm GExp} = -(1-f)\, \kappa \cdot \langle \sigma_{\Delta t} \rangle \tag{10.4}$$

$$\langle (\delta\Delta t)^2 \rangle_{\rm GExp} = \left(s^2 + 2\,(1-f)\, \kappa \right) \cdot \langle (\sigma_{\Delta t})^2 \rangle \ , \tag{10.5}$$

from which we calculate the GExp RMS squared

$$\begin{aligned} (\delta\Delta t)^2_{\rm RMS} = & \, s^2 \cdot \langle (\sigma_{\Delta t})^2 \rangle + (1-f^2)\, \kappa \cdot \langle \sigma_{\Delta t} \rangle^2 \\ & + 2\,(1-f)\, \kappa \cdot (\sigma_{\Delta t})^2_{\rm RMS} \ . \end{aligned}$$

Both of these models predict a mean and RMS residual that scale linearly with $\sigma_{\Delta t}$ and vanish in the limit of zero measured error. The predicted slopes for the dependence of the mean residual with the measured error are

$$f\, b_1 + (1-f)\, b_2 \quad \text{for G+G}, \qquad -(1-f)\, \kappa \quad \text{for GExp} \ ,$$

and the predicted slopes for the RMS are

$$\sqrt{f\,s_1^2 + (1-f)\,s_2^2 + f(1-f)\,(b_1-b_2)^2} \quad \text{for G+G}, \qquad \sqrt{s^2 + (1-f^2)\,\kappa} \quad \text{for GExp}\ .$$

Both models perform similarly on the fit to Δt residual.

In addition to the core of the resolution function, we model the contribution of "outliers": candidates whose residual $\delta\Delta t$ is (on average) at least 5 times larger than their calculated error $\sigma_{\Delta t}$ (and which are therefore poorly vertexed.) Outliers represent less than 1% of the candidates selected in the signal Monte Carlo sample. We model outliers with a single Gaussian whose width and bias are *not* scaled by the calculated error

$$\mathcal{R}_{\rm out}(\delta\Delta t;\, b_{\rm out}, s_{\rm out}) \equiv G(\delta\Delta t;\, b_{\rm out}, s_{\rm out})\ .$$

The reason for not scaling the outlier model by $\sigma_{\Delta t}$ is that most outlier events are from those where the vertexing algorithm for tag-side B mistakenly finds a secondary vertex, instead of the primary vertex, and $\sigma_{\Delta t}$ does not contain any information on the distance between the primary and second vertices.

We combine the two core and outlier models using an extra parameter $f_{\rm out}$ that specifies the relative fraction of events in the outlier contribution:

$$\begin{aligned}\mathcal{R}_{\rm G+G+G}(\delta\Delta t, \sigma_{\Delta t};\, b_1, b_2, b_{\rm out}, s_1, s_2, s_{\rm out}, f, f_{\rm out}) \equiv \\ f\cdot G(\delta\Delta t;\, b_1\sigma_{\Delta t}, s_1\sigma_{\Delta t}) + (1-f-f_{\rm out})\cdot G(\delta\Delta t;\, b_2\sigma_{\Delta t}, s_2\sigma_{\Delta t}) + f_{\rm out}\cdot\mathcal{R}_{\rm out}\ ,\end{aligned} \tag{10.6}$$

and

$$\mathcal{R}_{\mathrm{GExp+G}}(\delta\Delta t, \sigma_{\Delta t};\, \kappa, s, f, b_{\mathrm{out}}, s_{\mathrm{out}}, f_{\mathrm{out}}) \equiv f \cdot G(\delta\Delta t;\, 0, s\,\sigma_{\Delta t}) + (1 - f - f_{\mathrm{out}}) \cdot \frac{1}{2\,\kappa\sigma_{\Delta t}} \int_{-\infty}^{0} du\, \exp\big(+u/(\kappa\,\sigma_{\Delta t})\big) \cdot G(u - \delta\Delta t;\, 0, s\,\sigma_{\Delta t}) + f_{\mathrm{out}} \cdot \mathcal{R}_{\mathrm{out}}(\delta\Delta t, \sigma_{\Delta t};\, b_{\mathrm{out}}, s_{\mathrm{out}})\,. \quad (10.7)$$

10.1.2 Model test on lifetime

In this subsection, a series of fits to the signal Monte Carlo sample is performed. The model used here is a simple lifetime model ($\exp(-|\Delta t|/\tau)$) convoluted with a resolution model. All of the candidates used in these studies have been selected by the signal sample selection and are correctly reconstructed according to the truth information. Three kinds of fits are carried out:

- fit the true Δt and Δz distributions for the lifetime,
- fit the distribution of residuals $\delta\Delta t$ and $\sigma_{\Delta t}$, to obtain the parameters of the two resolution models G+G+G and GExp+G, and
- fit the reconstructed Δt distribution to obtain the lifetime, using the two resolution models, and with the resolution parameters either fixed (to values obtained in the previous step) or floating in the fit.

The resulting lifetimes derived from these fits is shown in Table 10.1.

The fit to Δt_{true} reveals any bias introduced by the event selection. By comparing the lifetimes extracted from the fits to Δt_{true} and Δz_{true}, we can estimate any additional bias due to the boost approximation (Sec. 6.2.1). However, the difference between the lifetime from these two fits is probably an overestimate of the boost effects since some of the bias will be absorbed into the resolution model. When we

Table 10.1: Results of fitting lifetime compared with reference values in various stages. $\tau_{\rm gen} = 1.548$ ps is the value used in the Monte Carlo generator; $\tau(\Delta t_{\rm true})$ is the result from a fit to true Δt; $\tau(\Delta z_{\rm true})$ is the result from a fit to true Δz; $\tau(\Delta t)$ is the result from a fit to measured Δt.

Type	$\mathcal{R}$ fixed	$\mathcal{R}$ floating
$\tau(\Delta t_{\rm true}) - \tau_{\rm gen}$ (fs)	-3.4 ± 5.0	-
$\tau(\Delta z_{\rm true}) - \tau_{\rm gen}$ (fs)	1.3 ± 5.0	-
GExp+G, $\tau(\Delta t) - \tau(\Delta t_{\rm true})$ (fs)	-11.8 ± 5.9	3.4 ± 10.2
G+G+G, $\tau(\Delta t) - \tau(\Delta t_{\rm true})$ (fs)	-9.8 ± 5.0	2.0 ± 15.7

fit to the $\Delta t_{\rm true}$ distribution, we measure a lifetime of 1.5446 ± 0.0051 ps, which is (3.4 ± 5.1) fs below the generated value. When we fit to the $\Delta z_{\rm true}$ distribution and scale the fit parameter by the average value of $\beta\gamma\gamma^* c = 166.87$ μm/ps for the generated simulation sample, we measure a lifetime of 1.5494 ± 0.0054 ps. This value is 4.8 fs higher than that extracted from $\Delta t_{\rm true}$, which is significant, given that the events are identical.

This bias, however, is expected. The reason is the following. Fitting to a pure exponential $e^{-|x|/\tau}$ is mathematically identical to calculating the average of the value $|x|$. According to Eq. 6.2, for large $t_1 - t_2$ and small $t_1 + t_2$, Δz is always positive, independent of $\cos\theta^*$. The average of $|\Delta t|$ is equal to that of $|\Delta z|/\beta\gamma\gamma^* c$ for these events since $\langle \cos\theta^* \rangle = 0$. However, for small $t_1 - t_2$ and large $t_1 + t_2$, Δz can be negative even if $t_1 - t_2 > 0$. After taking the absolute value, these events will contribute in positive values, the average over $\cos\theta^*$ will not cancel, which will therefore bias the average high.

To estimate this bias, we generate 100 toy experiments of ≈81k events simply using $e^{-t/\tau}$ and $(1 - \cos^2\theta^*)$ angular distribution of the B mesons in the Υ(4S) frame, and using Eq. 6.2 to calculate Δz. We then calculate the average of $|\Delta t|$ and $|\Delta z|/\beta\gamma\gamma^* c$. The histogram of the shifts is shown in Fig. 10.1. The blue line shows

the value found in the fit to the signal simulation. The mean bias is 4.7 fs with an RMS of 0.76 fs. Signal simulation fit is consistent with the toy experiments.

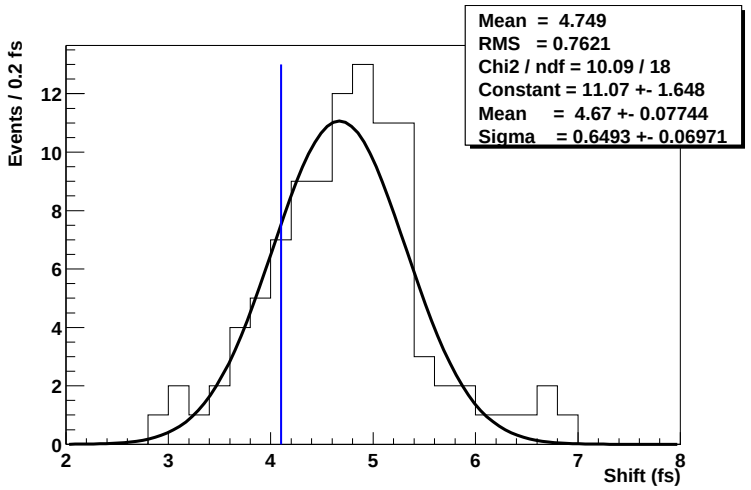

Figure 10.1: Shift of $\langle|\Delta z|/\beta\gamma\gamma^* c\rangle$ from $\langle|\Delta t|\rangle$ with 100 toy experiments. Blue line indicates the value found in signal simulation fit.

Table 10.1 also shows the difference between the fit to the measured Δt with resolution floating and fixed at the values from fits to residual. When the resolution function is fixed at the result from a fit to Δt residual, the lifetime result is significantly lower then the fit to $\Delta t_{\rm true}$ and the simultaneous fit for lifetime and resolution function. This implies that the resolution model may not be totally accurate. We already know that the RMS of the residual $\Delta z_{\rm true} - \Delta t_{\rm true}$ depends on $|\Delta t_{\rm true}|$. We can check the significance of this dependence at the level of reconstructed Δt. The top plot in Fig. 6.3 shows the RMS of the pull of reconstructed Δt in bins of true Δt. We can clearly see the dependence on $\Delta t_{\rm true}$. The middle plot shows the same sort of distribution except that $\Delta t_{\rm meas}$ is replaced by $\Delta z_{\rm true}/\gamma\beta\gamma^* c$ in the calculation of the RMS of the pull. The bottom plot is the bin-by-bin difference in quadrature between the top two histograms. The fact that the dependence of the pull RMS on $\Delta t_{\rm true}$ almost disappears after removing the effect due to boost in the conversion from Δz

to Δt suggests that the boost approximation is the major source of the dependence of resolution on $\Delta t_{\rm true}$. Since this dependence can be fully understood, in principle the model can be modified to accommodate this effect.

The effect of not modeling this $\Delta t_{\rm true}$ dependence of resolution is studied with toy Monte Carlo simulations. A set of datasets are generated with the same number of events as in the full signal Monte Carlo. The pull of Δt is smeared by double-Gaussian or GExp resolution models, whose RMS varies as a function of $\Delta t_{\rm true}$ according to the shape found in the signal Monte Carlo simulation (first plot in Fig. 6.2). A simple double-Gaussian or GExp model without $\Delta t_{\rm true}$ dependence is then used to fit the residual, τ_{B^0}, Δm_d, and mistag fraction with and without fixing the resolution to the fits to the residual. The mean values of bias on lifetime are listed in Table 10.2. In the table the biases in fits to the full signal simulation data are also compared [87].

Table 10.2: Lifetime fit biases in the toy study described in the text and in the fits to full signal simulation data.

	GExp toy	G+G toy	GExp+G full MC	G+G+G full MC
Fix R	-4.6	-5.5	-7.9 ± 6.3	-5.8 ± 6.3
Float R	4.2	-2.4	7.5 ± 9.1	3.2 ± 9.4

This effect, however, does not exhibit statistically significant bias when the resolution is floating in the fit. The resolution function is able to absorb this small systematic effect and leave the physical parameters unbiased.

By fitting the mixing frequency simultaneously, the improvement of the statistical uncertainty on lifetime is observed. With this simulation sample, fitting to the lifetime and the resolution function only results in a statistical error of 11 fs on lifetime. While fitting to the mixing frequency at the same time, this error reduces to better than 9.5 fs, which represents a 15% improvement.

10.2 Fit Method

We perform an unbinned fit simultaneously to events in each of the 12 signal and control samples (indexed by s) that are further subdivided into 30 subsamples (indexed by c) using a likelihood

$$\mathcal{L} = \prod_{s=1}^{12} \prod_{c=1}^{30} \prod_{k=1}^{N(s,c)} P_{s,c}(\vec{x}_k \,;\vec{p}) \;, \tag{10.8}$$

where k indexes the $N(s,c)$ events $\vec{x}_k$ in each of the 360 subsamples. The probability $P_{s,c}(\vec{x}_k \,;\vec{p})$ of observing an event $\vec{x}_k = (\delta m, \Delta t, \sigma_{\Delta t}, g)$ is calculated as a function of the parameters $\vec{p} = (f_{s,c}^{\rm comb}, \vec{p}_{s,c}^{\rm comb}, \vec{p}_c^{\rm peak}, \vec{q}_{s,c}^{\rm comb}, f_{s,c,1}^{\rm pkg}, f_{s,c,2}^{\rm pkg}, f_{s,c,3}^{\rm pkg}, \vec{q}_{s,c,1}^{\rm pkg}, \vec{q}_{s,c,2}^{\rm pkg}, \vec{q}_{s,c,3}^{\rm pkg}, \vec{q}_c^{\rm sig})$ as

$$\begin{aligned} &P_{s,c}(\delta m, \Delta t, \sigma_{\Delta t}, g \,;\vec{p}) = \\ &f_{s,c}^{\rm comb} \cdot \mathcal{F}^{\rm comb}(\delta m \,;\vec{p}_{s,c}^{\rm comb}) \cdot \mathcal{G}^{\rm comb}(\Delta t, \sigma_{\Delta t}, g \,;\vec{q}_{s,c}^{\rm comb}) + \left(1 - f_{s,c}^{\rm comb}\right) \cdot \mathcal{F}^{\rm peak}(\delta m \,;\vec{p}_c^{\rm peak}) \cdot \\ &\quad \left[\sum_{j=1}^{3} f_{s,c,j}^{\rm pkg} \cdot \mathcal{G}_j^{\rm pkg}(\Delta t, \sigma_{\Delta t}, g \,;\vec{q}_{s,c,j}^{\rm pkg}) + \left(1 - \sum_{j=1}^{3} f_{s,c,j}^{\rm pkg}\right) \cdot \mathcal{G}^{\rm sig}(\Delta t, \sigma_{\Delta t}, g \,;\vec{q}_c^{\rm sig}) \right] \;, \end{aligned} \tag{10.9}$$

where j indexes the three sources of peaking background and $\delta m = m(D^*) - m(D^0)$. The index g is $+1$ (-1) for unmixed (mixed) events. Functions labeled with $\mathcal{F}$ are peak and combinatoric components of δm distributions, defined in Eqs. 9.2 and 9.3, and thus describe the probability of observing a particular value of δm. Functions labeled with $\mathcal{G}$ give probabilities for values of Δt and $\sigma_{\Delta t}$ in category g, which are discussed in detail in Secs. 10.2.1 and 10.2.2. By allowing different effective mistag rates for apparently mixed or unmixed events in the background functions $\mathcal{G}^{\rm comb}$ and $\mathcal{G}^{\rm pkg}$, we accommodate the different levels of backgrounds observed in mixed and unmixed samples. Parameters labeled with f describe the relative contributions

of different types of events. Parameters labeled with $\vec{p}$ describe the shape of a δm distribution, and those labeled with $\vec{q}$ describe a (Δt, $\sigma_{\Delta t}$) shape.

Note that we make explicit assumptions that the δm peak shape, parameterized by $\vec{p}_c^{\,\mathrm{peak}}$, and the signal ($\Delta t$, $\sigma_{\Delta t}$) shape, parameterized by $\vec{q}_c^{\,\mathrm{sig}}$, depend only on the subsample index c. The first of these assumptions is supported by data, and simplifies the analysis of peaking background contributions. The second assumption reflects our expectation that the Δt distribution of signal events does not depend on whether they are selected in the signal sample or appear as a background in a control sample.

The ultimate aim of the fit is to obtain the B^0 lifetime and mixing frequency, which by construction are common to all sets of signal parameters $\vec{q}_c^{\,\mathrm{sig}}$. Most of the statistical power for determining these parameters comes from the signal sample, although the fake and uncorrelated background control samples also contribute due to their signal content (see Table 9.4).

We bootstrap the full fit with a sequence of initial fits using reduced likelihood functions to a partial set of samples, to determine the appropriate parameterization of the signal resolution function and the background Δt models, and to determine starting values for each parameter in the full fit.

1. We first find a model that describes the Δt distribution for each type of event: signal, combinatoric background, and the three types of backgrounds that peak in the $m(D^*) - m(D^0)$ distribution. To establish a model, we use Monte Carlo samples that have been selected to correspond to only one type of signal or background event based on Monte Carlo truth information. These samples are used to determine the Δt model and the categories of events (*e.g.,* tagging category, fake or real lepton) that can share each of the parameters in the model. Any subset of parameters can be shared among any subset of the 360 subsamples. We choose parameterizations and sharing of parameters that

minimize the number of different parameters while still providing an adequate description of the Δt distributions.

2. We then find the starting values for the background parameters by fitting to each of the background-enhanced control samples in data, using the model (and sharing of parameters) determined in the previous step. Since these background control samples are not pure, we start with the purest control sample (combinatoric background events from the $m(D^*) - m(D^0)$ sideband) and move on to less pure control samples, always using the models established from earlier steps to describe the Δt distribution of the contamination from other backgrounds.

The result of the above two steps is a Δt model for each type of event and a set of starting values for all parameters in the fit. When we do the final fit, we fit all signal and control samples simultaneously (≈68k events), leaving essentially all parameters free in the fit (72 free parameters). The physics parameters τ_{B^0} and Δm_d were kept hidden until all analysis details and the systematic errors were finalized, to eliminate experimenters' bias. However, statistical errors on the parameters and changes in the physics parameters due to changes in the analysis were not hidden.

10.2.1 Signal Δt model

For signal events in a given tagging category c, the probability density function (PDF) for Δt consists of a physics model convolved with a Δt resolution function:

$$\mathcal{G}^{\rm sig}(\Delta t, \sigma_{\Delta t}, g\,; \vec{q}_c^{\rm sig}) = \left\{ \frac{1}{4\tau_{B^0}} e^{-|\Delta t_{\rm true}|/\tau_{B^0}} \left(1 + g(1-2\omega_c)\cos(\Delta m_d \Delta t_{\rm true})\right) \right\} \otimes \mathcal{R}(\delta\Delta t, \sigma_{\Delta t}; \vec{q_c})\,,$$

where $\mathcal{R}$ is a resolution function, which can be different for different event categories, g is $+1$ (-1) for unmixed (mixed) events, and $\delta\Delta t$ is the residual $\Delta t - \Delta t_{\rm true}$. The

GExp+G resolution model (Eq. 10.7) is used in the final fit. Be reminded that width of the core Gaussian and the decay constant of the exponential in the resolution are scaled by $\sigma_{\Delta t}$, while the width and bias for the outlier Gaussian are not. The G+G+G model (Eq. 10.6) is used for systematic studies. The physics model shown in the above equation has seven parameters: Δm_d, τ_{B^0}, and mistag fractions ω_c for each of the five tagging categories. To account for an observed correlation between the mistag rate and $\sigma_{\Delta t}$ in the kaon category (described in Sec. 7.2), we allow the mistag rate in the kaon category to vary as a linear function of $\sigma_{\Delta t}$:

$$\omega_{\rm kaon} = m_{\rm kaon} \cdot \sigma_{\Delta t} + \omega_{\rm kaon}^{\rm offset} \,. \tag{10.10}$$

In addition, we allow the mistag fractions for B^0 tags and $\overline{B}^0$ tags to be different. We define $\Delta\omega = \omega_{B^0} - \omega_{\overline{B}^0}$ and $\omega = (\omega_{B^0} + \omega_{\overline{B}^0})/2$, so that

$$\omega_{B^0/\overline{B}^0} = \omega \pm \frac{1}{2}\Delta\omega \,.$$

The mean and RMS spread of Δt residual distributions in Monte Carlo simulation vary significantly among tagging categories. For example, the mean of kaon-tagged events is -0.233 ± 0.006 ps and that of lepton-tagged events is -0.105 ± 008 ps. This is due to the fact that a kaon-tagged event is more likely to contain a charm decay, which causes a bias on $B_{\rm tag}$ vertex. We study several ways to accommodate this effect by allowing some parameters in the resolution function to be different for each tagging category. We find that by allowing the core Gaussian fraction f in GExp+G model, or the core Gaussian bias b_1 in G+G+G model to be different for each tagging category, the resolution function can accommodate this effect. In addition, we find that the correlations among the three parameters describing the outlier Gaussian ($b^{\rm out}$, $s^{\rm out}$, $f^{\rm out}$) are large and that the outlier parameters are highly correlated with other resolution parameters. Therefore, we fix the outlier bias $b^{\rm out}$

Table 10.3: All 13+10 parameters of default fit to signal event in signal Monte Carlo sample with GExp+G model, in which the core Gaussian fraction is split by tagging category. For this particular set of results, we do not hold either of the outlier shape parameters to be constant, since the statistical power of the full signal simulation sample is sufficient for direct measurement.

GExp+G			
par.	value	par.	value
Δm_d	0.4714 ± 0.0043 ps^{-1}	$f^1_{\rm LT}$	0.841 ± 0.035
τ_{B^0}	1.5523 ± 0.0096 ps	$f^1_{\rm KT}$	0.686 ± 0.024
$\omega_{\rm LT}$	0.0561 ± 0.0035	$f^1_{\rm N1}$	0.853 ± 0.029
$\omega_{\rm N1}$	0.1840 ± 0.0061	$f^1_{\rm N2}$	0.740 ± 0.026
$\omega_{\rm N2}$	0.3392 ± 0.0054	$f^1_{\rm N3}$	0.719 ± 0.022
$\omega_{\rm N3}$	0.4617 ± 0.0040	κ	1.16 ± 0.07
$\omega^{\rm offset}_{\rm KT}$	0.0672 ± 0.0085	s^1	1.001 ± 0.021
$m_{\rm KT}$	0.128 ± 0.012 ps^{-1}	$b^{\rm out}$	-1.680 ± 1.05 ps
$\Delta\omega_{\rm LT}$	0.0044 ± 0.0058	$f^{\rm out}$	0.0044 ± 0.0022
$\Delta\omega_{\rm KT}$	-0.01403 ± 0.0045	$s^{\rm out}$	5.44 ± 0.86 ps
$\Delta\omega_{\rm N1}$	0.0120 ± 0.0098		
$\Delta\omega_{\rm N2}$	-0.04155 ± 0.0083		
$\Delta\omega_{\rm N3}$	-0.07418 ± 0.0061		

and scale factor $s^{\rm out}$, and vary them over a wide range to evaluate the systematic uncertainty on the physics parameters due to fixing these parameters (see Sec. 11.2). The resolution model then has 8 free parameters: s, κ, $f^{\rm out}$, and five fractions f_c (one for each tagging category c).

This signal model is tested on the correctly reconstructed signal Monte Carlo events. The result is shown in Table 10.3, and the simulation data and signal Δt model projection are shown in Fig. 10.2. Both τ_{B^0} and Δm_d are consistent with generated values and the results from a fit to true Δt and tagging information.

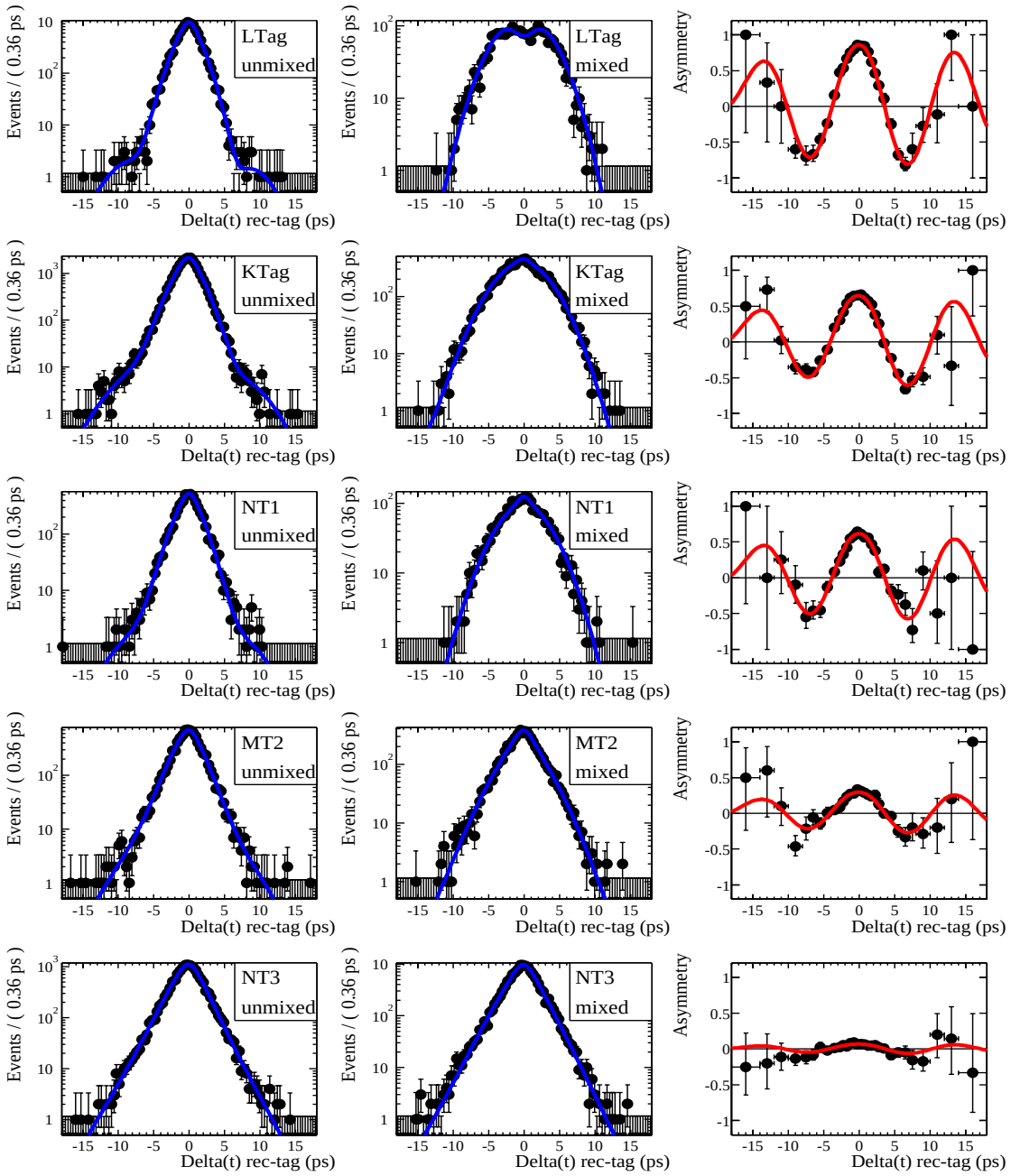

Figure 10.2: Δt distributions for unmixed and mixed events on a log scale, and the asymmetry plots for the signal Monte Carlo sample for the five tagging categories. The curves are projections of the signal model from the GExp+G fit to signal Monte Carlo sample.

10.2.2 Background Δt model

Although the true Δt and resolution on Δt are not well-defined for background events, we still describe the total Δt model as a "physics model" convolved with a "resolution function".

The background Δt physics models we use in this analysis are each a linear combination of one or more of the following terms, corresponding to prompt (zero lifetime), exponential lifetime, and oscillatory distributions:

$$\begin{aligned}
\mathcal{G}^{\rm pmt}_{\rm phys}(\Delta t_{\rm true}, g) &= (1/2)\cdot\delta(\Delta t_{\rm true})\cdot\left(1+g\cdot(1-2\omega^{\rm pmt})\right), \\
\mathcal{G}^{\rm life}_{\rm phys}(\Delta t_{\rm true}, g) &= (1/4)\cdot\exp(-|\Delta t_{\rm true}|/\tau^{\rm bg})\cdot\left(1+g\cdot(1-2\omega^{\rm life})\right), \\
\mathcal{G}^{\rm osc}_{\rm phys}(\Delta t_{\rm true}, g) &= (1/4)\cdot\exp(-|\Delta t_{\rm true}|/\tau^{\rm bg})\cdot\left(1+g\cdot(1-2\omega^{\rm osc})\cos\Delta m^{\rm bg}\Delta t_{\rm true}\right),
\end{aligned}$$

where $\delta(\Delta t)$ is a δ-function, $g = +1$ for unmixed and -1 for mixed events, and $\tau^{\rm bg}$ and $\Delta m^{\rm bg}$ are the effective lifetime and mixing frequency for the particular background.

For backgrounds, we use a resolution function that is the sum of a narrow and a wide Gaussian distribution:

$$\begin{aligned}
&\mathcal{R}_{\rm G+G}(\delta\Delta t, \sigma_{\Delta t}; b, s, f, b^w, s^w) \\
&= f\cdot G(\delta\Delta t; b\sigma_{\Delta t}, s\sigma_{\Delta t}) + (1-f)\cdot G(\delta\Delta t; b^w\sigma_{\Delta t}, s^w\sigma_{\Delta t}) .
\end{aligned}$$

Combinatoric background

Events in which the D^* candidate corresponds to a random combination of charged tracks (called combinatoric background) constitute the largest background in the

signal sample. We use two sets of events to determine the appropriate parameterization of the Δt model for combinatoric background: events in data that are in the upper $m(D^*) - m(D^0)$ sideband (above the peak due to real D^* decays); and events in Monte Carlo simulation that are identified as combinatoric background, based on the true information for the event, in both the $m(D^*) - m(D^0)$ sideband and peak region. The data and Monte Carlo Δt distributions are described well by a prompt plus oscillatory term convolved with a double-Gaussian resolution function:

$$\begin{aligned}\mathcal{G}^{\rm comb} = &\left[f^{\rm osc}\mathcal{G}^{\rm osc}_{\rm phys}(\Delta t_{\rm true}, g; \tau^{\rm comb}, \Delta m^{\rm comb}, \omega^{\rm osc}) + (1 - f^{\rm osc})\mathcal{G}^{\rm pmt}_{\rm phys}(\Delta t_{\rm true}, g; \omega^{\rm pmt})\right] \\ &\otimes \mathcal{R}_{\rm G+G}(\delta\Delta t, \sigma_{\Delta t}; b, s, f, b^w, s^w) \ . \end{aligned} \tag{10.11}$$

The parameters $\omega^{\rm pmt}$, $\Delta m^{\rm comb}$, $\tau^{\rm comb}$, f, b^w, and s^w are shared among all control samples. The parameters $\omega^{\rm osc}$, $f^{\rm osc}$, b, and s are allowed to be different depending on criteria such as tagging category, whether the data was recorded on- or off-resonance, whether the candidate lepton passes real- or fake-lepton criteria, whether the event passes the criteria for same-side or opposite-side D^* and ℓ, and how many identified leptons are in the event. The total number of free parameters in the combinatoric background Δt model is 24.

The bias and width of the resolution are scaled by $\sigma_{\Delta t}$. We justify this model by fitting to the sideband data in slices of $\sigma_{\Delta t}$ with combinatoric Δt model. We still maintain the scaling in the model. If the bias and width of the resolution do *not* scale by $\sigma_{\Delta t}$, we should see the scale factors decrease as $\sigma_{\Delta t}$ increases. If the model is adequate, the scale factors should be consistent to each other for different $\sigma_{\Delta t}$ slices. The scale factor for width changes by only less than 20%, although the mean of $\sigma_{\Delta t}$ changes by about factor of three. This study shows that scaling the bias and width is a adequate approach for combinatoric model.

The relative fraction of $B^0\overline{B}^0$ and B^+B^- events in the combinatoric background

depends slightly on $m(D^*) - m(D^0)$. However, no significant dependence of the parameters of the Δt model on $m(D^*) - m(D^0)$ is observed in data or Monte Carlo simulation. The sample of events in the $m(D^*)-m(D^0)$ sideband is used to determine the starting values for the parameters in the final full fit to all data samples.

To reduce the total number of free parameters in the fit, parameters that describe the shape of the wide Gaussian (bias and width) are shared between combinatoric background and the three types of peaking background: continuum, fake-lepton, and uncorrelated-lepton. The outlier fraction is allowed to be different for each type of background.

Continuum background

All $c\overline{c}$ events that have a correctly reconstructed D^* are defined as continuum peaking background, independent of whether the associated lepton candidate is a real lepton or a fake lepton. The $c\overline{c}$ Monte Carlo sample and off-resonance data are used to identify the appropriate Δt model and sharing of parameters among subsamples. The combinatoric-background Δt model and parameters described in the previous section are used to model the combinatoric-background contribution in the off-resonance Δt distribution in data.

Events with a real D^* from continuum $c\overline{c}$ production should have vanishing Δt in the case of perfect reconstruction. Therefore, we use the following model for the Δt distribution of these events:

$$\mathcal{G}^{\rm cont} = \mathcal{G}^{\rm pmt}_{\rm phys}(\Delta t_{\rm true}, g; \omega^{\rm pmt}) \otimes \mathcal{R}_{\rm G+G}(\delta\Delta t, \sigma_{\Delta t}; b, s, f, b^w, s^w) \; .$$

Dependence on the flavor tagging information is included to accommodate any differences in the amount of background events classified as mixed and unmixed.

By fitting to the data and Monte Carlo control samples with different sharing of

parameters across subsets of the data, we find that the apparent "mistag fraction" for events in the `kaon` category is significantly different from the mistag fraction for other tagging categories. We also find that the core Gaussian bias is significantly different for opposite-side and same-side events. We introduce separate parameters to accommodate these effects.

The total number of parameters used to describe the Δt distribution of continuum peaking background is six. The off-resonance control samples in data are used to determine starting values for the final full fit to all data samples.

Fake lepton background

To determine the Δt model and sharing of parameters for the fake-lepton peaking backgrounds, we use $B^0\overline{B}^0$ and B^+B^- Monte Carlo events in which the D^* is correctly reconstructed but the lepton candidate is misidentified. In addition, we use the fake-lepton control sample in data. The combinatoric and continuum peaking background Δt models and parameters described in the previous two sections are used to model their contribution to the fake-lepton Δt distribution in data. For this study, the contribution of the signal is described by the signal parameters found for signal events in the Monte Carlo simulation.

Since the fake-lepton peaking background is due to B decays in which the fake lepton and the D^* candidate can originate from the same B or different B mesons, we include both prompt and oscillatory terms in the Δt model:

$$\mathcal{G}^{\rm fake} = \left[f^{\rm osc} \cdot \mathcal{G}^{\rm osc}_{\rm phys} + (1 - f^{\rm osc}) \cdot \mathcal{G}^{\rm pmt}_{\rm phys} \right] \otimes \mathcal{R}_{\rm G+G}(\delta\Delta t, \sigma_{\Delta t}; b, s, f, b^w, s^w) \; .$$

We find that the apparent mistag rates for both the prompt and mixing terms, and the bias of the core Gaussian of the resolution function, are different between some tagging categories. The total number of parameters used to describe the fake-lepton

background is 14. The fake-lepton control samples in data are used to determine starting values for the final full fit to all data samples.

Uncorrelated background

To determine the Δt model and sharing of parameters for the uncorrelated-lepton peaking backgrounds, we use $B^0\overline{B}^0$ and B^+B^- Monte Carlo events in which the D^* is correctly reconstructed but the lepton candidate is from the other B in the event or from a secondary decay of the same B. In addition, we use the same-side control sample in data, which is only about 30% uncorrelated-lepton background in the $m(D^*) - m(D^0)$ peak region due to significant contributions from combinatoric background and signal. The combinatoric and other peaking background Δt models and parameters described in the previous two sections are used to model their contribution to the same-side Δt distribution in data. For this initial study, the contribution of the signal is described by the signal parameters found for signal events in the Monte Carlo simulation.

Physics and vertex reconstruction considerations suggest several features of the Δt distribution for the uncorrelated-lepton sample. First, we expect the reconstructed Δt to be systematically smaller than the true Δt value since using a lepton and a D^* from different B decays will generally reduce the separation between the reconstructed $B_{\rm rec}$ and $B_{\rm tag}$ vertices. We also expect that events with small true Δt will have a higher probability of being misreconstructed as an uncorrelated lepton candidate because it is more likely that the fit of the D^* and ℓ to a common vertex will converge for these events. Finally, we expect truly mixed events to have a higher fraction of uncorrelated-lepton events because in mixed events the charge of the D^* is opposite that of primary leptons on the tagging side. These expectations are confirmed in the Monte Carlo simulation.

We do not expect the uncorrelated-lepton background to exhibit any mixing

behavior and none is observed in the data or Monte Carlo control samples. We describe the Δt distribution with the sum of a lifetime term and a prompt term, convolved with a double-Gaussian resolution function:

$$\mathcal{G}^{\rm uncor} = \left[f^{\rm life} \cdot \mathcal{G}^{\rm life}_{\rm phys}(\Delta t_{\rm true}, g; \tau^{\rm uncor}, \omega^{\rm life}) + (1 - f^{\rm life}) \cdot \mathcal{G}^{\rm pmt}_{\rm phys}(\Delta t_{\rm true}, g; \omega^{\rm pmt}) \right] \otimes \mathcal{R}_{\rm G+G}(\delta\Delta t, \sigma_{\Delta t}; b, s, f, b^w, s^w) \; . \tag{10.12}$$

The effective mistag rates $\omega^{\rm pmt}$ and $\omega^{\rm life}$ accommodate different fractions of uncorrelated-lepton backgrounds in events classified as mixed and unmixed. We find that the apparent mistag rate for the lifetime term is different between some tagging categories. All other parameters are consistent among the different subsamples. The total number of parameters used to describe the uncorrelated-lepton background is six. The uncorrelated-lepton control samples in data are used to determine starting values for the final full fit to all data samples.

10.2.3 Charged B model

The charged-B peaking background is due to decays of the type $B^{\pm} \to D^* \ell \nu_\ell X$. Since charged B's do not exhibit mixing behavior, our strategy is to use the Δt and tagging information to discriminate charged-B peaking background events from neutral-B signal events, in the simultaneous fit to all samples. We use the same resolution model and parameters as for the neutral-B signal since the decay dynamics are very similar. The signal model, with the charged B background term, becomes

$$\mathcal{G}^{\rm sig}(\Delta t, \sigma_{\Delta t}, g\,; \vec{q}_c^{\,\rm sig}) = [\frac{1 - f_{B^+}}{4\tau_{B^0}} e^{-|\Delta t_{\rm true}|/\tau_{B^0}} \, (1 + g(1 - 2\omega^c_{B^0}) \cos(\Delta m_d \Delta t_{\rm true})) + \frac{f_{B^+}}{4\tau_{B^+}} e^{-|\Delta t_{\rm true}|/\tau_{B^+}} \, (1 + g(1 - 2\omega^c_{B^+}))] \otimes \mathcal{R}(\delta\Delta t, \sigma_{\Delta t}; \vec{q}_c) \, ,$$

where $\omega^c_{B^0}$ ($\omega^c_{B^+}$) is the mistag fraction for neutral (charged) B mesons for tagging category c.

Given that the ratio of the charged B to neutral B lifetime is close to one and the fraction of charged B mesons in the peaking sample is only a few percent [71], sensitivity of this analysis is too low to distinguish the B^0 and B^+ lifetimes in the fit. We parameterize the physics model for the B^+ in terms of the lifetime ratio τ_{B^+}/τ_{B^0}, and fix this ratio to the world average value in 2002 [6]. We vary the ratio by the error on the world average to estimate the corresponding systematic uncertainties on τ_{B^0} and Δm_d (see Sec. 11.2).

The fit is sensitive to only two parameters among ω_{B^+}, ω_{B^0} and the charged B fraction (f_{B^+}). The reason is that, given the fact that the two lifetimes are very close, the time evolution of mixed and unmixed events can be expressed approximately as [88]

$$\begin{aligned} U &\propto (1 + f_{B^\pm} D_{B^\pm})/2 + ((1 - f_{B^\pm}) D_{B^0} \cos \Delta m_d \Delta t)/2 \\ M &\propto (1 - f_{B^\pm} D_{B^\pm})/2 - ((1 - f_{B^\pm}) D_{B^0} \cos \Delta m_d \Delta t)/2 \,, \end{aligned}$$

where $D \equiv 1 - 2\omega$. Only two quantities, $f_{B^\pm} D^\pm_B$ and $(1 - f_{B^\pm}) D^0_B$ are measurable. Therefore we fix the ratio of mistag rates, $\omega_{B^+}/\omega_{B^0}$, to the value of the ratio measured with fully reconstructed charged and neutral hadronic B decays in data, for each tagging category.

The ability of this model to distinguish charged B background from neutral B signal is tested with fits to Monte Carlo simulation, in which 17290 events are real signal and 952 are $B^+ \to D^{*-} \ell^+ \mathrm{X}$ events, and there are no other backgrounds. The charged B fraction in this mixture is 5.2%. The lifetime ratio τ_{B^+}/τ_{B^0} and mistag ratio $\omega_{B^+}/\omega_{B^0}$ are set to be 1.069 and 0.97, respectively. The lifetime ratio is based on the world average in 2000, which is the value used in the event generator of

the Monte Carlo samples. The mistag ratio is calculated from the overall mistag fractions for $B^0\overline{B}^0$ and B^+B^- samples by counting event in the sample based on the true Monte Carlo information. The fit results for Δm_d, τ_{B^0} and the charged B fraction are shown in Table 10.4, along with the results of the fit to the B^0 signal only. In all fits, the resolution function and mistag fractions are free parameters. The fitted charged B fraction is consistent with the true value, and the τ_{B^0} and Δm_d are consistent with the true values and the fit to B^0 events only.

Table 10.4: Signal model fit to B^0 signal with or without B^+ peaking background in signal and generic Monte Carlo simulation. In all fits, resolution function and mistag fractions are floating, except for the two outlier model parameters. *Note that the fit to signal Monte Carlo events differs slightly from that shown in Table 10.3 because of the fixed outlier model shape.

parameter	Sig MC*	Gen $B^0\overline{B}^0$	Gen $B^0\overline{B}^0$, B^+B^-
Δm_d (ps^{-1})	0.4715 ± 0.00427	0.4640 ± 0.0101	0.4617 ± 0.0132
τ_{B^0} (ps)	1.5523 ± 0.00915	1.5749 ± 0.0227	1.5656 ± 0.0227
f_{B^+}	-	-	0.047 ± 0.030

10.3 Fit Result

The total number of free parameters in the final fit is 72: 22 of them in the signal model, 24 in the combinatoric background model, and 26 in peaking background models. The ratio of B^+ to B^0 lifetimes is fixed at the world average in 2002 [6], 1.083 ± 0.017. The ratios of B^+ to B^0 mistag fractions are fixed at the central values extracted from tagging performance analysis with hadronic events [72]. They are summarized in Table 10.5. The outlier Gaussian bias and width are fixed at -5 ps and 6 ps (see Sec. 11.2 for details.). The fit is done "blind", that is, the central values of τ_{B^0} and Δm_d are randomize by Gaussian distributions with width 0.06 ps

and 0.04 ps^{-1}, respectively. The actual shifts are kept hidden to avoid human bias due to experimenters' expectation. These hidden shifts are consistent through out the whole analysis so that the effects of changing analysis parameters or strategies can be studied. The hidden values are revealed after all analyses are finalized.

Table 10.5: B^+ to B^0 mistag fraction ratios extracted from hadronic events analyses [72], except the NT3 category, which is not available. We calculate the ratio for NT3 from generic Monte Carlo sample.

Category	ratio
`LTag`	0.54 ± 0.10
`KTag`	0.68 ± 0.05
`NT1`	0.99 ± 0.12
`NT2`	1.05 ± 0.07
`NT3`	$1.12 \pm 0.12^*$

The fitted signal Δt model parameters are shown in Table 10.6, and the combinatoric and peaking background model parameters are shown in Table 10.7 and Table 10.8, respectively. The error on τ_{B^0} is slightly asymmetric.

The statistical correlation coefficient between τ_{B^0} and Δm_d is $\rho(\Delta m_d, \tau_{B^0}) = -0.22$. The global correlation coefficients for τ_{B^0} and Δm_d, and some of the correlation coefficients between τ_{B^0} or Δm_d and other parameters, are shown in Table 10.9.

Figure 10.3 shows the Δt distributions for unmixed and mixed events in the signal sample (opposite-side D^*-lepton candidates in on-resonance data). The points correspond to data. The curves correspond to the sum of the projections of the appropriate relative amounts of signal and background Δt models for the signal sample in the δm range between 143 and 148 MeV. Figure 10.4 shows the asymmetry

$$A = \frac{N_{\rm unmixed}(\Delta t) - N_{\rm mixed}(\Delta t)}{N_{\rm unmixed}(\Delta t) + N_{\rm mixed}(\Delta t)} \, .$$

Table 10.6: Results of full fit to data — signal model and resolution function parameters. A small correction, described in Sec. 11.1, has been applied to τ_{B^0} and Δm_d. The parameters without errors are kept constant in the fit.

Signal Model and Δt Resolution Function Parameters					
parameter	value	parameter	value	parameter	value
Δm_d (ps^{-1})	0.492 ± 0.018	f_{B^+}	0.082 ± 0.029	s	1.201 ± 0.063
τ_{B^0} (ps)	$1.523\ ^{+0.024}_{-0.023}$	$\omega_{\rm lepton}$	0.071 ± 0.015	κ	0.86 ± 0.17
-	-	$\omega^{\rm offset}_{\rm kaon}$	0.002 ± 0.024	$f_{\rm lepton}$	0.72 ± 0.10
-	-	$m_{\rm kaon}$	0.229 ± 0.036	$f_{\rm kaon}$	0.609 ± 0.088
-	-	$\omega_{\rm NT1}$	0.212 ± 0.020	$f_{\rm NT1}$	0.69 ± 0.13
-	-	$\omega_{\rm NT2}$	0.384 ± 0.018	$f_{\rm NT2}$	0.70 ± 0.10
-	-	$\omega_{\rm NT3}$	0.456 ± 0.012	$f_{\rm NT3}$	0.723 ± 0.078
-	-	$\Delta\omega_{\rm lepton}$	-0.001 ± 0.022	$f^{\rm out}$ (%)	0.27 ± 0.17
-	-	$\Delta\omega_{\rm kaon}$	-0.024 ± 0.015	$b^{\rm out}$ (ps)	-5.000
-	-	$\Delta\omega_{\rm NT1}$	-0.098 ± 0.032	$s^{\rm out}$ (ps)	6.000
-	-	$\Delta\omega_{\rm NT2}$	-0.112 ± 0.028	-	-
-	-	$\Delta\omega_{\rm NT3}$	-0.023 ± 0.019	-	-

The unit amplitude for the cosine dependence of A is diluted by the mistag probability, the experimental Δt resolution, and backgrounds.

Additional plots that compare data and the fit result for background-enriched samples are shown in Appendix B.

Since we float many parameters in the model, it is interesting to see how the errors on τ_{B^0} and Δm_d, and their correlation change when different parameters are free in the fit, or fixed to their best value from the full fit. We perform a series of fits, fixing all parameters at the values obtained from the default fit, except (a) Δm_d and τ_{B^0}, (b) Δm_d, τ_{B^0}, and all mistag fractions in the signal model, (c) Δm_d, τ_{B^0}, and f_{B^+}, (d) Δm_d, τ_{B^0}, f_{B^+}, and all mistag fractions in the signal model, (e) all parameters in the signal Δt model. The one-sigma error ellipses for these fits and for the default fit are shown in Fig. 10.5.

Table 10.7: Result of the full fit to data — Δt model parameters for combinatoric background, $G^{(\mathrm{comb})}$. The parameters without errors are kept constant in the fit.

Final Fit to Data, Combinatoric Δt Model Parameters			
parameter	value	parameter	value
Δm^{comb} (ps^{-1})	0.422 ± 0.020	${b^1}_{\mathrm{LNT12}}$	-0.056 ± 0.021
τ^{comb} (ps)	1.234 ± 0.024 ps	${b^1}_{\mathrm{KNT3}}$	-0.104 ± 0.012
$\omega^{\mathrm{osc}}_{\mathrm{Fake;LTag}}$	0.434 ± 0.026	$f^{\mathrm{osc}}_{\mathrm{Off}}$	0.099 ± 0.035
$\omega^{\mathrm{osc}}_{\mathrm{Fake;KTag}}$	0.128 ± 0.017	$f^{\mathrm{osc}}_{\mathrm{SS}}$	0.385 ± 0.023
$\omega^{\mathrm{osc}}_{\mathrm{Fake;NT1}}$	0.443 ± 0.032	$f^{\mathrm{osc}}_{\mathrm{NoLept}}$	0.526 ± 0.019
$\omega^{\mathrm{osc}}_{\mathrm{Fake;NT2}}$	0.507 ± 0.024	$f^{\mathrm{osc}}_{\mathrm{OneLept}}$	0.894 ± 0.022
$\omega^{\mathrm{osc}}_{\mathrm{Fake;NT3}}$	0.589 ± 0.020	$f^{\mathrm{osc}}_{\mathrm{TwoLept}}$	1.000
$\omega^{\mathrm{osc}}_{\mathrm{Lept;LTag}}$	0.156 ± 0.023	${s^1}_{\mathrm{OffSS}}$	1.337 ± 0.024
$\omega^{\mathrm{osc}}_{\mathrm{Lept;KTag}}$	0.297 ± 0.010	${s^1}_{\mathrm{On;OS}}$	1.259 ± 0.021
$\omega^{\mathrm{osc}}_{\mathrm{Lept;NT1}}$	0.377 ± 0.025	f^1	0.9666 ± 0.0071
$\omega^{\mathrm{osc}}_{\mathrm{Lept;NT2}}$	0.401 ± 0.016	b^{out}	-0.98 ± 0.24
$\omega^{\mathrm{osc}}_{\mathrm{Lept;NT3}}$	0.491 ± 0.011	s^{out}	4.60 ± 0.42
ω^{pmt}	0.3960 ± 0.0093	-	-

We can see that the error on τ_{B^0} changes very little until we float the signal resolution function. Floating the background parameters adds a very small contribution to the error. The contribution from the charged B fraction and mistag fractions to the τ_{B^0} error is negligible. On the other hand, the charged B fraction changes the error on Δm_d the most. The contributions from floating the mistag fractions, resolution functions, and background Δt models are relatively small.

Table 10.8: Result of the full fit to data — Δt model parameters for peaking backgrounds. The outlier bias ($\mu_{\rm out}$) and scale factor ($\sigma_{\rm out}$) for these peaking background are shared with combinatoric background parameters. The parameters without errors are kept constant in the fit.

Final Fit to Data, Peaking BG Model Parameters					
cont. par.	value	fake par.	value	unco. par.	value
$\omega_{\rm KT}$	0.083 ± 0.028	$\Delta m^{\rm fake}$	0.444 ± 0.031	$\tau^{\rm unco}$	1.07 ± 0.17
$\omega_{\rm LNT}$	0.457 ± 0.040	$\tau^{\rm fake}$	1.416 ± 0.060	$\omega^{\rm life}_{\rm KNT}$	0.71 ± 0.13
${b^1}_{\rm OS}$	0.04 ± 0.11	$\omega^{\rm osc}_{\rm LT}$	0.248 ± 0.038	$\omega^{\rm life}_{\rm LT}$	0.9998
${b^1}_{\rm SS}$	-0.08 ± 0.13	$\omega^{\rm osc}_{\rm KT}$	0.142 ± 0.029	$\omega^{\rm pmt}$	0.0 ± 0.00064
s^1	1.300 ± 0.090	$\omega^{\rm osc}_{\rm NT1}$	0.342 ± 0.044	$f^{\rm life}$	0.877 ± 0.085
f^1	0.909 ± 0.040	$\omega^{\rm osc}_{\rm NT2}$	0.377 ± 0.039	b^1	-0.30 ± 0.18
-	-	$\omega^{\rm osc}_{\rm NT3}$	0.512 ± 0.042	f^1	1.000
-	-	$\omega^{\rm pmt}_{\rm KNT3}$	0.46 ± 0.14	s^1	1.34 ± 0.24
-	-	$\omega^{\rm pmt}_{\rm LNT12}$	0.64 ± 0.15	-	-
-	-	$f^{\rm osc}$	0.852 ± 0.046	-	-
-	-	${b^1}_{\rm KNT3}$	-0.208 ± 0.062	-	-
-	-	${b^1}_{\rm LNT12}$	-0.078 ± 0.065	-	-
-	-	f^1	0.944 ± 0.023	-	-
-	-	s^1	1.159 ± 0.091	-	-

Table 10.9: Global correlation coefficients for Δm_d and τ_{B^0} from the full fit to data and other correlation coefficients for pairs of key parameters in the fit.

Δm_d global correlation	0.74
τ_{B^0} global correlation	0.69
$\rho(\Delta m_d, \tau_{B^0})$	-0.22
$\rho(\Delta m_d, f_{B^+})$	0.58
$\rho(\tau_{B^0}, \sigma^1_{\rm sig})$	-0.49
$\rho(\tau_{B^0}, f^{\rm out}_{\rm sig})$	-0.26

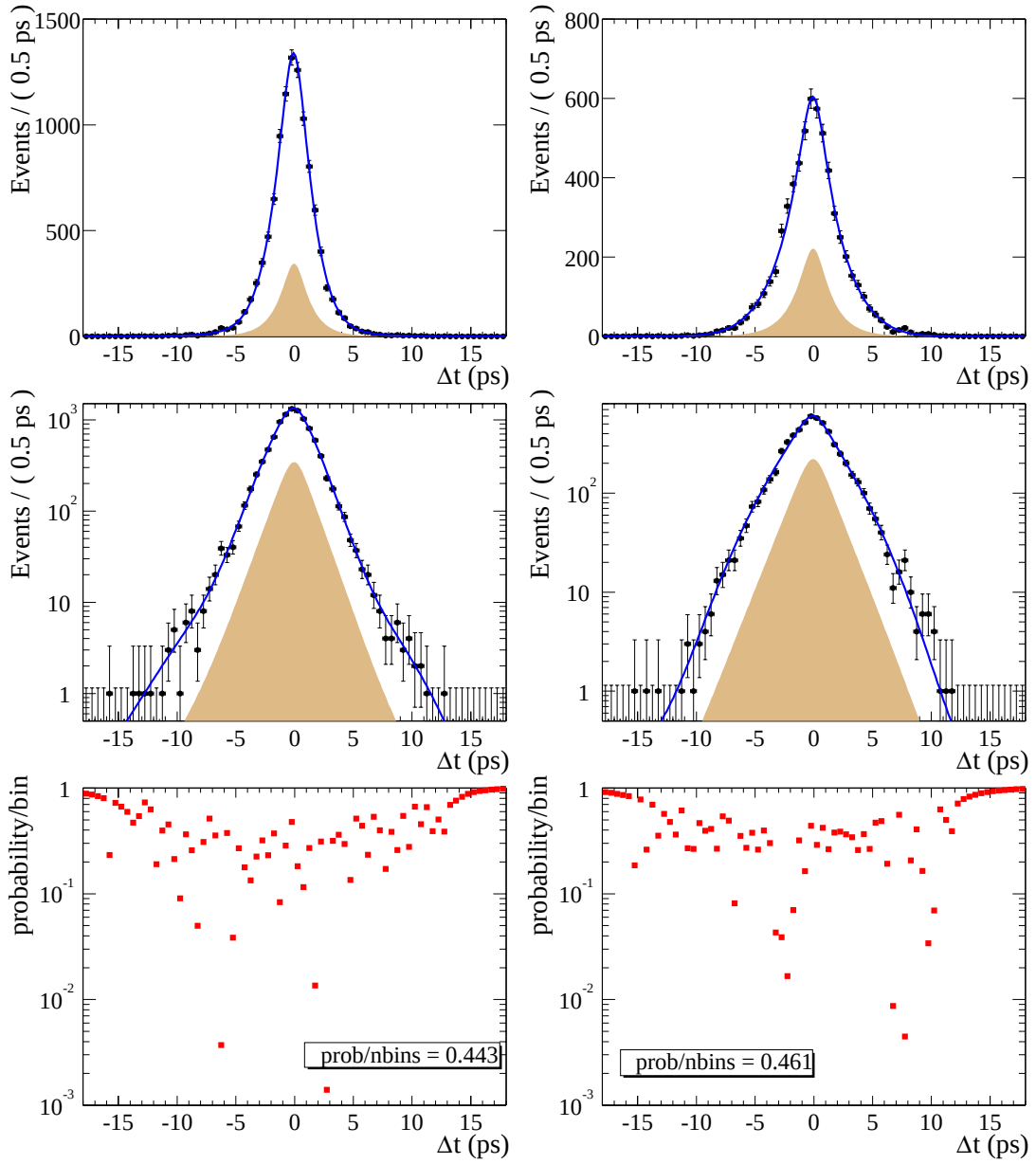

Figure 10.3: The Δt distribution for unmixed and mixed events in the signal sample (opposite-side D^*-lepton candidates in on-resonance data) and the projection of the fit results. The left-hand plots are for unmixed events; the right-hand plots for mixed events. The shaded area shows the background contribution to the distributions. The bottom row is the Poisson probability of observing n events or smaller (larger) if n is smaller (larger) than the expected value; i.e., $y = \sum_{i=0}^{n} P(i;\mu)$ for $n < \mu$ and $y = \sum_{i=n}^{\infty} P(i;\mu)$ for $n > \mu$.

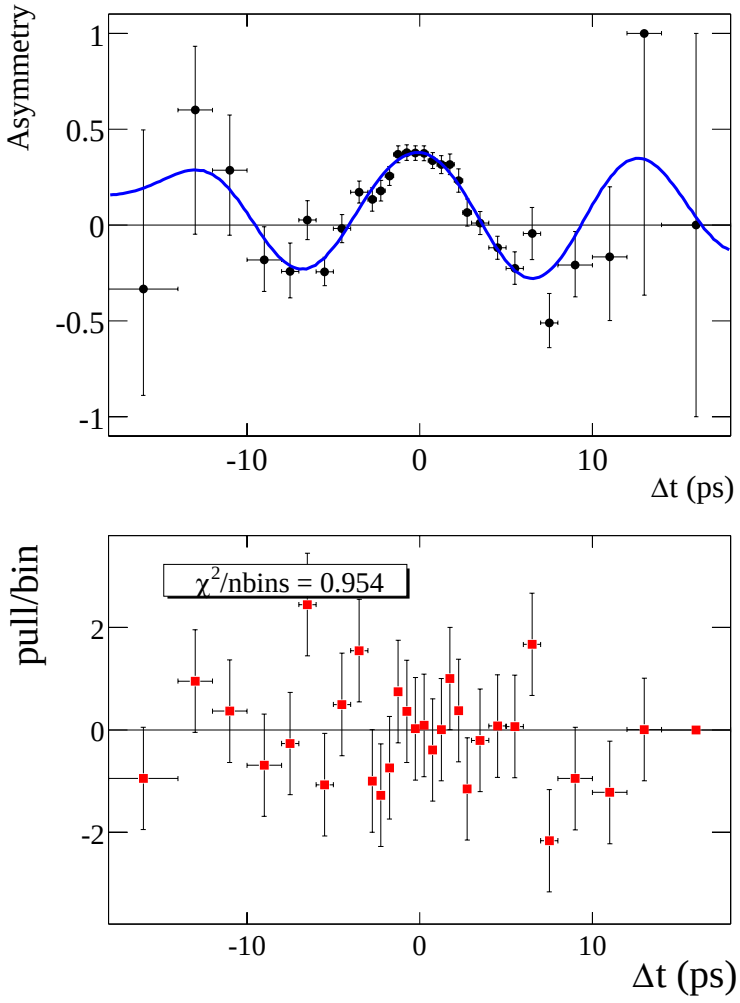

Figure 10.4: The asymmetry plot for mixed and unmixed events in the signal sample (opposite-side D^*-lepton candidates in on-resonance data) in the δm range from 143 to 148 MeV, and the projection of the fit results. The lower plot shows the bin-by-bin difference normalized by the error estimated from data on each bin.

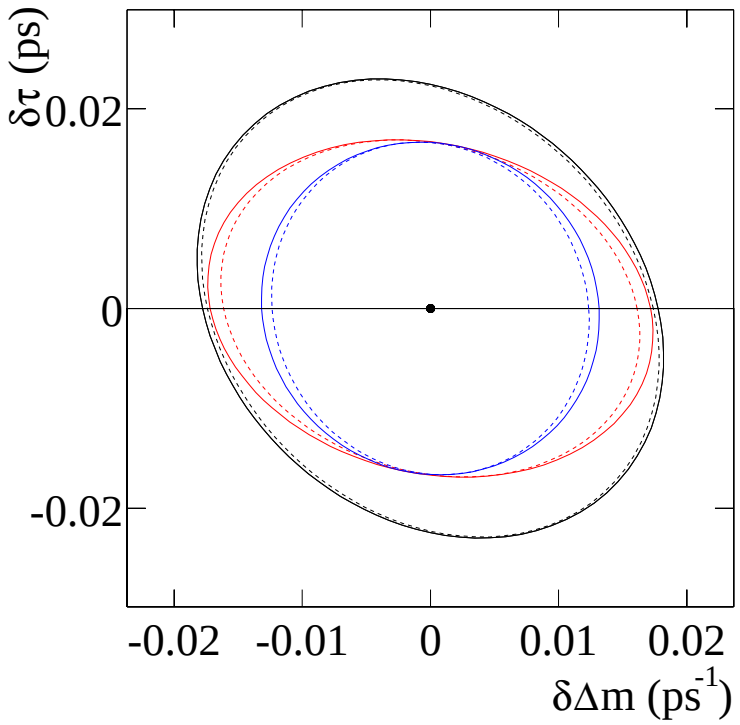

Figure 10.5: Comparison of one-sigma error ellipses in the Δm_d-τ_{B^0} plane for fits in which different sets of parameters are free. From the innermost to the outermost ellipse, the floating parameters are (Δm_d, τ_{B^0}), (Δm_d, τ_{B^0}, mistag fractions), (Δm_d, τ_{B^0}, f_{B^+}), (Δm_d, τ_{B^0}, f_{B^+}, mistag fractions), all signal Δt parameters, and the default fit (72 floating parameters).

Chapter 11

Validation and Systematic Study

In this chapter, validation exercises on both data and simulation, and systematic uncertainty studies are described. A correction that is comparable to a full statistical error in data is applied for both τ_{B^0} and Δm_d. Detail is described in Sec. 11.1.1.

11.1 Validation and Consistency Checks

In Sec. 11.1.1, I describe several tests of the fitting procedure that are performed with both fast parameterized Monte Carlo simulations and full detector simulations. In Sec. 11.1.2, the results of consistency checks on data are given, including fitting to different subsamples of the data and fitting with variations to the standard fit.

A bias on the fit values of τ_{B^0} and Δm_d is observed when fitting a combination of signal and background Monte Carlo events. A correction (approximately equivalent to the statistical uncertainty on τ_{B^0} and Δm_d in data) for this bias is applied to the final fit results in data and apply a systematic uncertainty equivalent to the statistical uncertainty on the Monte Carlo fit result.

11.1.1 Tests of fitting procedure with Monte Carlo simulations

Fast parameterized (toy) Monte Carlo simulation

A test of the fitting procedure is performed with fast parameterized Monte Carlo simulations, where 87 experiments are generated with signal and background control sample sizes and compositions corresponding to the data. The Δt, mixing status, and the flavor of the tagging B are generated with a random generator with probabilities according to the full model with parameters obtained from the full likelihood fit to data. Because the fitting function does not model the $\sigma_{\Delta t}$ and δm distribution and category variables, for each event $\sigma_{\Delta t}$, δm and category variables are taken from the corresponding events in data. In other words, the random generator goes through each event in data, chooses the parameters according to this event's $\sigma_{\Delta t}$, δm and category variable values, and generates Δt, mixing status and $B_{\rm tag}$ flavor according to the full model. The full fit is then performed on each of these experiments. The distributions of the fit values of Δm_d, τ_{B^0} and $f_{B^\pm}$ are shown in Fig. 11.1. The mean and RMSs of Δm_d, τ_{B^0} and $f_{B^\pm}$ are:

$$\begin{array}{lll} \Delta m & \text{mean} = 0.5280 \pm 0.0018 \text{ ps}^{-1} & \text{RMS} = 0.0171 \pm 0.0013 \text{ ps}^{-1} \\ \tau_{B^0} & \text{mean} = 1.5390 \pm 0.0024 \text{ ps} & \text{RMS} = 0.0223 \pm 0.0017 \text{ ps} \\ f_{B^\pm} & \text{mean} = 0.076 \pm 0.004 & \text{RMS} = 0.034 \pm 0.003 \,, \end{array}$$

and should be compared with the central value and statistical error on the (blind) fit values from data:

$$\Delta m_d = 0.527 \pm 0.018 \text{ ps}^{-1} \text{ (blind)} \,,$$

$$\tau_{B^0} = 1.539 \pm 0.023 \text{ ps (blind)} ,$$

and

$$f_{B^\pm} = 0.082 \pm 0.029 .$$

The central values of Δm_d, τ_{B^0} and $f_{B^\pm}$ from the fit to data (used to generate the toy samples) are shown as vertical dashed lines in Fig. 11.1. Both means and RMS's are consistent between fits to toy sample and data. The correlation coefficient calculated from the Δm_d-τ_{B^0} distribution is -0.21, which is in a good agreement with the fit to data, -0.22.

The distribution of minimized negative log-likelihood (minNll) is shown in Fig. 11.2. Seventeen of the fits to the 87 experiments result in a value of the negative log likelihood that is smaller (better) than that found in data.

We also check the statistical errors on data by measuring the increase in negative log likelihood in data in the two-dimensional (τ_{B^0}, D^-) space in the vicinity of the minimum of the negative log likelihood. We found that the positive error on τ_{B^0} is about 6% larger than that predicted by the fitting program, whereas the other errors are the same as predicted. We increased the positive statistical error on τ_{B^0} by 6%.

Full simulation

We also fit two types of Monte Carlo samples that include full detector simulation: pure signal and signal plus background. To check whether the selection criteria introduce any bias in the lifetime or mixing frequency, we fit the signal physics model to the true lifetime distribution, using true tagging information, for a large sample of signal Monte Carlo events that pass all selection criteria. We also fit the measured Δt distribution, using measured tagging information, with the complete signal Δt model described in Sec. 10.2.1. We find no statistically significant bias in the values of τ_{B^0} or Δm_d extracted in these fits.

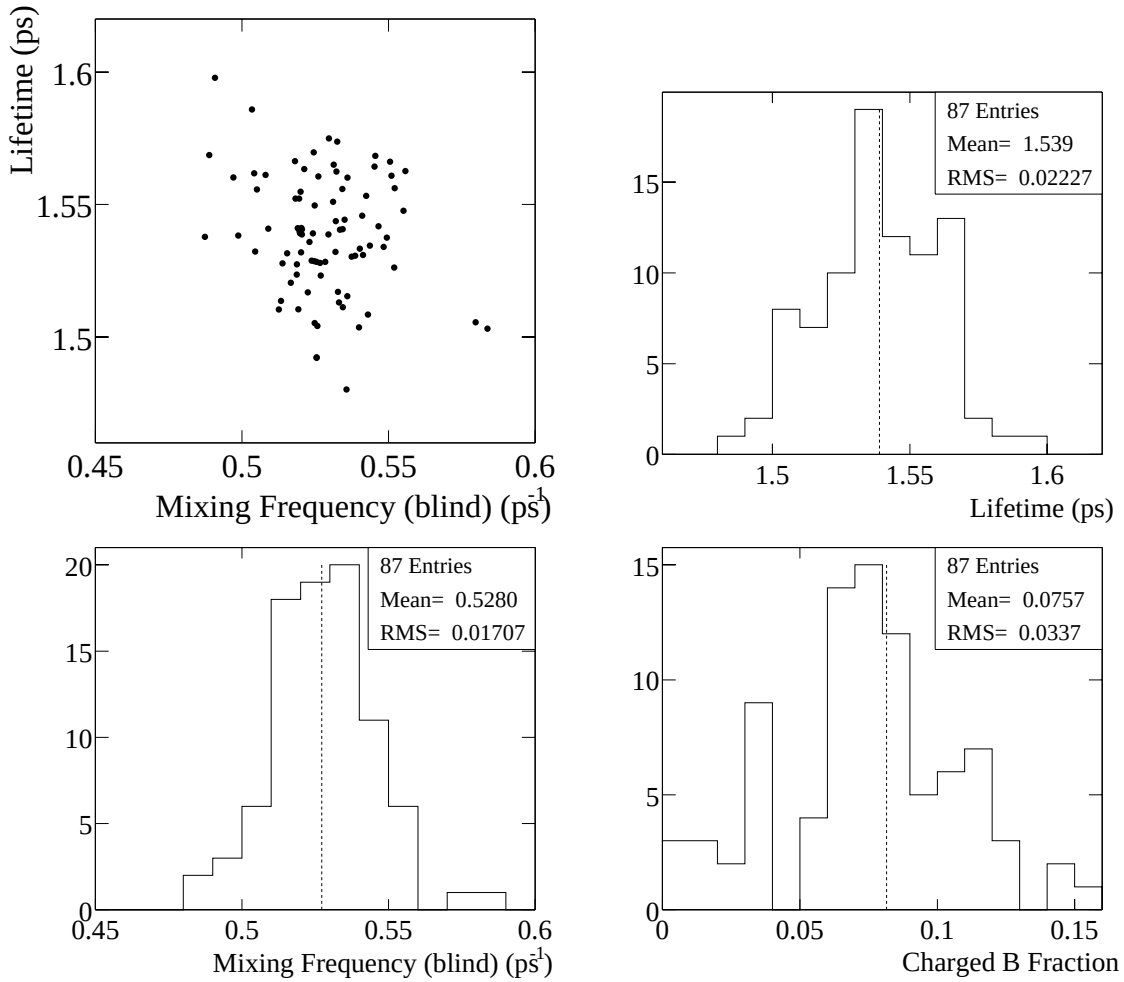

Figure 11.1: The distributions of τ_{B^0}, Δm_d and $f_{B^\pm}$ from 87 toy Monte Carlo sample fits. The values from the default fit to data, which were used to generate the toy samples, are indicated by vertical dashed lines. Note that the values shown here are before the hidden values for τ_{B^0} and Δm_d are revealed. Therefore, the mean values *cannot* be compared with the result shown in Table 10.6.

The $B^0\overline{B}^0$, B^+B^-, and $c\overline{c}$ Monte Carlo samples that provide simulated background events along with signal events are much smaller than the pure signal Monte Carlo samples. In addition, they are not much larger than the data samples. In order to increase the statistical sensitivity to any bias introduced when the background

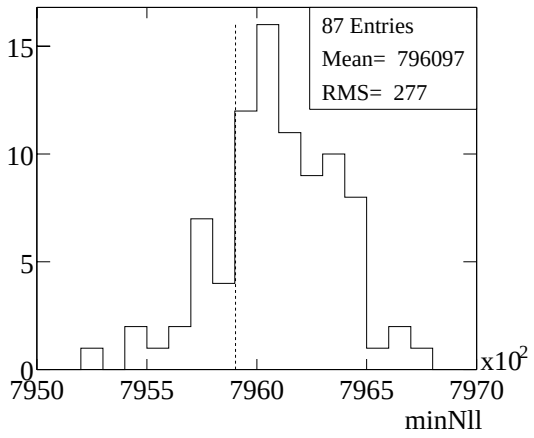

Figure 11.2: The distribution of minimized negative log-likelihood (minNll) from 87 toy Monte Carlo sample fits. The value from default fit to data is indicated by the vertical dashed lines. The number of fits that result in a smaller (better) value of minNll is 17.

samples are added to the fit, we compare the values of τ_{B^0} and Δm_d from the fit to signal plus background events, and pure signal events from the same sample. We find that when background is added, the value of τ_{B^0} increases by (0.022 ± 0.009) ps and the value of Δm_d increases by (0.020 ± 0.005) ps^{-1}, where the error is the difference in quadrature between the statistical errors from the fit with and without background. We correct our final results in data for these biases, which are each roughly the same size as the statistical error on the results in data. We conservatively apply a systematic uncertainty on this bias equal to the full statistical error on the measured result in Monte Carlo simulation with background: ± 0.0178 ps for τ_{B^0} and ± 0.0123 ps^{-1} for Δm_d.

11.1.2 Data consistency checks

We perform the full maximum-likelihood fit on different subsets of the data and find no statistically significant difference in the results for different subsets. The fit is

performed on datasets divided according to tagging category, b-quark flavor of the D^* ℓ candidate, b-quark flavor of the tagging B, and D^0 decay mode. We also vary the range of Δt over which we perform the fit (maximum value of $|\Delta t|$ equal to 10, 14, and 18 ps), and decrease the maximum allowed value of $\sigma_{\Delta t}$ from 1.8 ps to 1.4 ps. These checks are summarized in Figs. 11.3 and 11.4. Again, no statistically significant changes in the values of τ_{B^0} or Δm_d are found.

11.1.3 Change in kaon mistag slope

The kaon mistag is modeled as a linear function of $\sigma_{\Delta t}$ in signal model. The slope from the fit to signal only in $B^0\overline{B}^0$ Monte Carlo simulation is 12%, while the slope from the full fit to signal and background events is $20 \pm 4\%$. The change is due to the fact that we ignore a similar effect in background events; i.e., the kaon mistag in the background model is independent of $\sigma_{\Delta t}$. We can confirm that the kaon mistag in background events do increase as $\sigma_{\Delta t}$ increases by calculating the observed *mixed* events in background in slices of $\sigma_{\Delta t}$. Kaon tagged events have a slope similar to kaon tagged signal events, while events in other tagging categories do not have a non-zero slope.

The correlation coefficients between this slope and (Δm_d, τ_{B^0}) are very small, so that fixing this slope at 12% does not change τ_{B^0} and Δm_d by more than 0.001 ps or ps^{-1}. Therefore, we can safely ignore the slope for backgrounds.

11.1.4 Sensitivity to NT3 tagging information

Using the tagging information for events in the NT3 category allows us to treat all events on an equal footing for the lifetime and mixing measurement. Since the Q value (effective statistical power of an event for mixing measurement, defined as $\epsilon(1-2\omega)^2$, where ϵ is the tagging efficiency and ω is the mistag fraction) for this

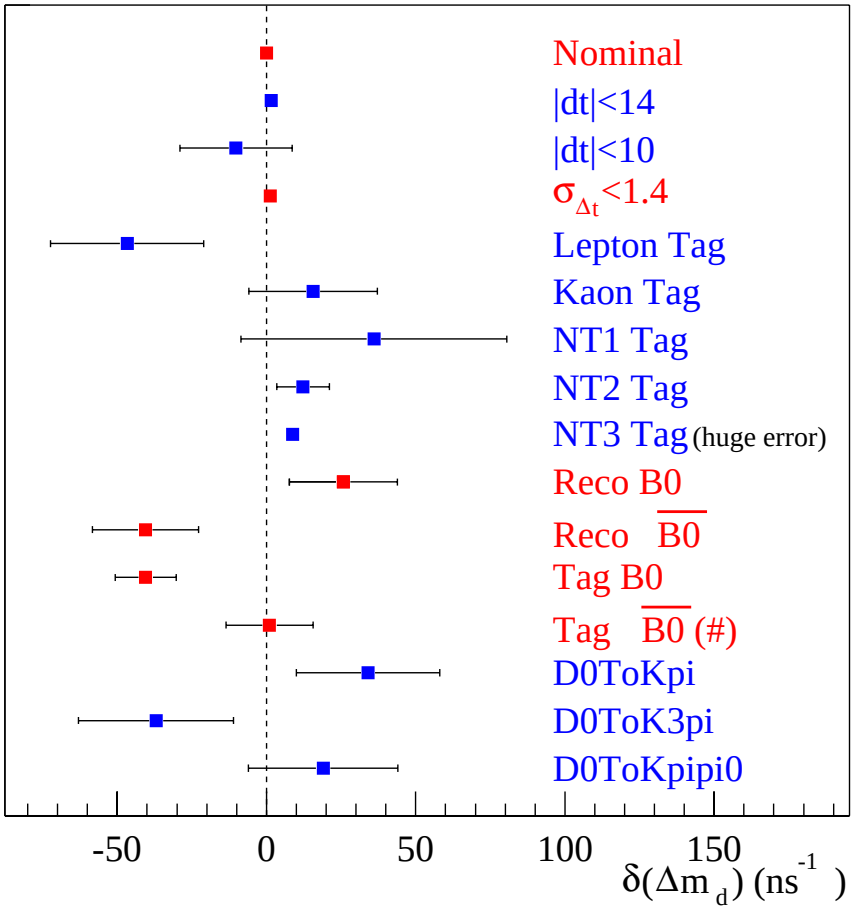

Figure 11.3: These plots show the Δm_d differences from the standard fit for various fits to the Data sample. The units of the plots are ns^{-1}. The errors shown on the plots are the sample and subsample errors subtracted in quadrature. The line marked with (#) represents a fit that only converged when f_{B^+} was fixed to the nominal value.

category is estimated to be about 0.3 %, it has only a small impact on the mixing measurement. This analysis is currently the only analysis in *BABAR* that uses tagging information from the `NT3` category. Since the `NT3` tagging performance is not well verified in *BABAR*, we also perform a fit to data in which we assume the tagging power

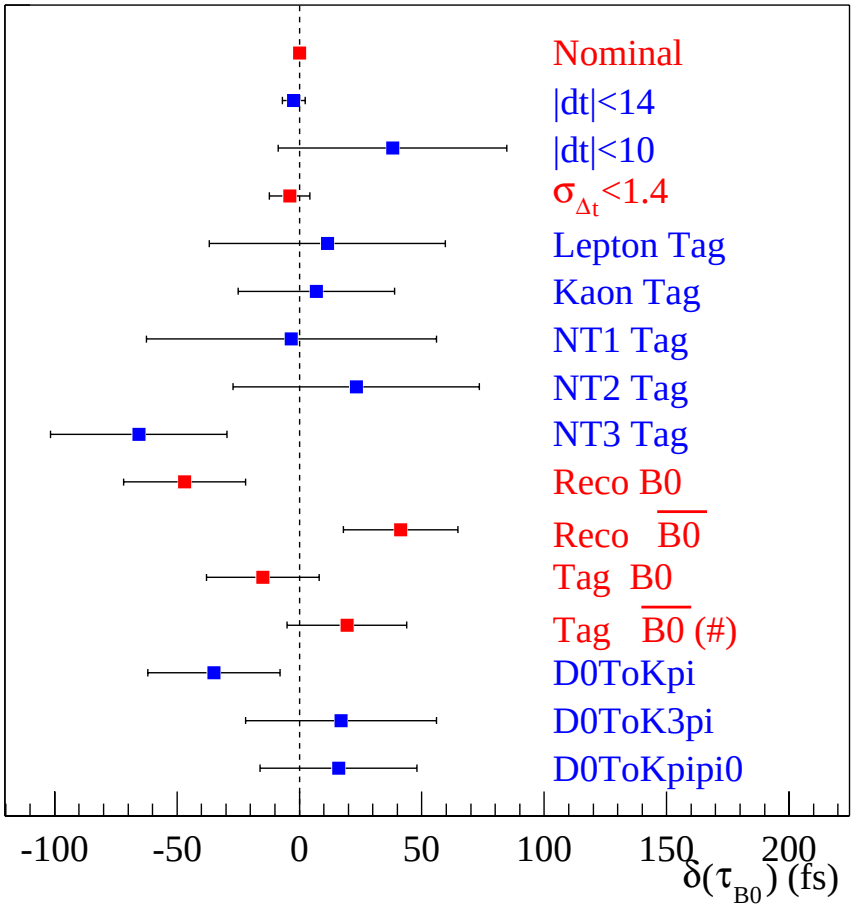

Figure 11.4: These plots show the τ_{B^0} differences from the standard fit for various fits to the Data sample. The units of the plots are fs. The errors shown on the plots are the sample and subsample errors subtracted in quadrature.

of NT3 is 0. It is equivalent to assuming that all events in this tagging category have a mistag rate of 50%.

We fix the NT3 mistag rates (ω_{NT3}) in signal and background Δt models to be 0.5. In a few parameters in background models, NT3 is grouped with other categories, we maintain the original parameter grouping, without separating NT3 from them, and

still let these parameters float.

The change in parameters are all very minor. When the NT3 tagging information is *not* used, the lifetime increases by 0.0022 ps and the statistical error increases by 0.002 ps (in quadrature); the change in Δm is $+0.0008$ ps^{-1} with an error that is larger by 0.0036 ps^{-1} (in quadrature). We therefore conclude that extracting NT3 mistag rates from data is warranted, and is slightly beneficial.

11.2 Systematic Uncertainties

Outlier in signal model

The outlier portion of the resolution function is intended to describe the Δt distribution for candidates whose residual is not consistent with the calculated error. These outliers are typically due to errors in tag-side vertex reconstruction, when a secondary vertex (e.g., charm decays) is identified instead of the primary vertex. Since there is little reason to expect the mean and width of the signal outlier Gaussian to scale with the Δt error, the outlier resolution function is parameterized as a Gaussian as a function of Δt residual.

We found that the fit to data is not very robust when all three parameters for the outlier Gaussian of the resolution function are allowed to float in the fit. In particular, the best fit value for the bias of the outlier function is always at the lower end of the allowed range, even when the lower limit is set at -18 ps. This is likely due to the negative tail in the Δt distribution for events in the signal sample; see Fig. 10.3. In addition, when all outlier parameters are free, many fits fail to converge or have an error matrix that is not positive definite. Therefore, we repeated the full fit with 36 different *fixed* values of the outlier bias and outlier width, covering a wide range for each parameter (-1 to -10 ps in bias, 4 to 11 ps in width).

Figure 11.5 shows the negative log-likelihood surface as a function of outlier bias

and width. The negative log-likelihood has been offset so that the minimum value over the plot is zero. For fixed width, the minimum in negative log-likelihood always occurs at the most negative bias. For fixed bias, the likelihood curve as a function of resolution becomes broader as the bias becomes more negative. Note that there is a local minimum at small bias. Also note that the change in negative log-likelihood is less than 0.5 for most of the values of bias and width scanned in this study.

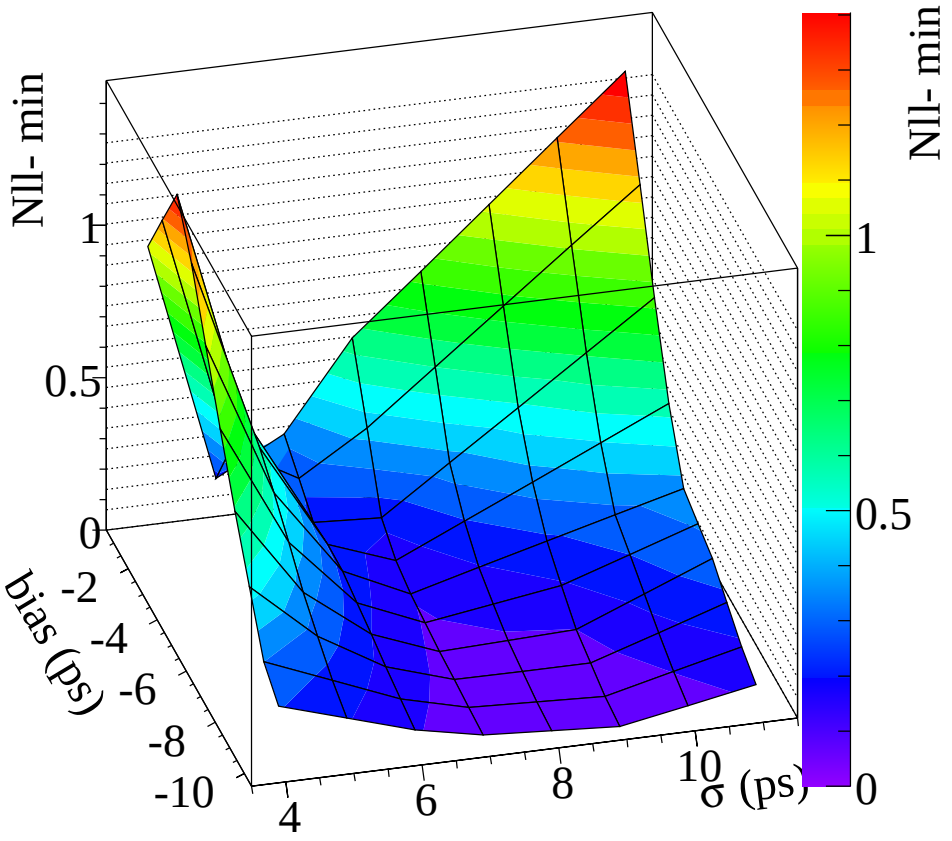

Figure 11.5: The negative log-likelihood surface in the space of outlier bias and outlier width for the outlier Gaussian of the signal resolution model in Data. The vertical scale has been offset so that minimum is at 0. The surface is constructed from 36 fits to the Data with different fixed values of the signal outlier parameters.

A scatter plot of the resulting values of Δm_d and τ_{B^0} is shown in Fig. 11.6. Note that the spread in values of τ_{B^0} and Δm_d is small compared to the statistical uncertainty on each quantity. This is consistent with the fact that the change in ln-likelihood is small as the outlier bias and width are varied (see Fig. 11.5).

In order to choose an outlier bias and width for the default fit to data, we select a point near the middle of the cluster of points in the τ_{B^0} versus Δm_d scatter plot (Fig. 11.6) and use the corresponding bias and width. We choose a bias of -5 ps and a width of 6 ps.

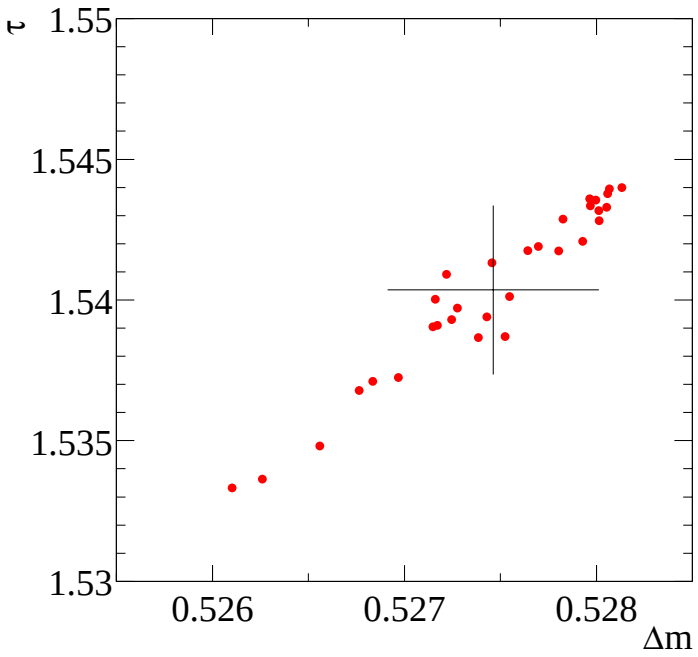

Figure 11.6: Scatter plot of the set of (blinded) Δm_d, τ_{B^0} values obtained from data for 36 fits with different fixed values of the outlier bias and width in the range -1 to -10 ps for bias and 4 to 11 ps for width. The cross indicates the mean and RMS of the Δm_d, τ_{B^0} distribution.

We assign half the full spread of lifetime and mixing results as the systematic uncertainty due to the outlier parameters, 0.005 ps for τ_{B^0} and 0.001 ps^{-1} for Δm_d.

Event selection bias and Monte Carlo Correction

The largest source of systematic uncertainty on both parameters is the limited statistical precision for determining the bias due to the fit procedure (in particular, the background modeling) with Monte Carlo events. We assign the statistical errors of a full fit to Monte Carlo samples including background to estimate this systematic uncertainty: ± 0.018 ps for τ_{B^0} and ± 0.012 ps^{-1} for Δm_d. See Sec. 11.1.1 for more details.

PEP-II boost and z scale uncertainty

Since Δt is directly proportional to the measured average PEP-II boost, and the measured Δz, the errors on τ_{B^0} and Δm_d are directly related to the uncertainty on the boost and the z scale. The PEP-II boost is found to be accurate at 0.1% level [57] based on the knowledge of beam energies. The z scale uncertainty is determined by reconstructing protons scattered from the beam pipe and comparing the measured beam pipe shape with the optical survey data that are taken before the beam pipe is covered [89]. The z scale uncertainty is less than 0.4%

The systematic uncertainties due to these effects are ± 0.0015 and ± 0.0005 ps^{-1} for Δm_d, and ± 0.0060 and 0.0015 ps for τ_{B^0}.

B^+ mistag rate and lifetime

The B^+ to B^0 lifetime ratio is fixed at the world average [6], i.e., 1.083 ± 0.017 ps We vary this ratio up and down by one sigma and repeat the full fit. The mixing and lifetime results are shown in Fig. 11.1. We assign the systematic errors to be half of the observed difference.

The B^+ to B^0 mistag ratios for all five tagging categories are also fixed. We move all five mistag ratios up and down by one sigma according to Table 10.5 and

Table 11.1: Results of full fits to data for the two physics parameters for different values of the fixed B^+ to B^0 lifetime ratio. All other fit parameters were nearly identical.

Parameter	$(\tau_{B^+}/\tau_{B^0}) + 1\sigma$	$(\tau_{B^+}/\tau_{B^0}) - 1\sigma$	diff./2
$\delta(\Delta m)$ (ps^{-1})	-0.00002	0.0006	0.0003
$\delta(\tau_{B^0})$ (ps)	-0.0016	0.0021	0.0019

repeat the full fit. The results are shown in Table 11.2. Again we assign half of the difference as the systematic error due to fixing mistag ratios.

Table 11.2: Results of fits to data for the two physics parameters for different values of the fixed B^+ to B^0 mistag rate ratio.

	$(\omega_{B^+}/\omega_{B^0}) + 1\sigma$	$(\omega_{B^+}/\omega_{B^0}) - 1\sigma$	diff./2
$\delta(\Delta m)$ (ps^{-1})	-0.00007	-0.00023	0.00008
$\delta(\tau_{B^0})$ (ps)	-0.0002	0.0004	0.0003

SVT alignment

The vertex precision is sensitive to the knowledge of the positions of charged track hits on the SVT, which in turn relies on the knowledge of the position of each silicon wafers. Several sets of constants describing the deviation of these wafers from their nominal location are produced by studying the hit residual from dimuon events and cosmic ray data with the SVT alone [57]. The optical survey information taken during the assembly of the SVT is also used to improve the precision. These constants are then used to re-process charged tracks.

The sets of constants used to process data events for this analysis are called alignment set "D" and "E". However, a much accurate set of constants (called set "L") is calculated after the data used in this analysis have been processed.

The SVT wafer location in Monte Carlo simulation is known perfectly. To estimate the effect of using less accurate alignment constants, Monte Carlo events are reprocessed with the SVT wafers moved by the amount that is described by the difference between alignment constant set D or E and L (called "diffDL" and "diffEL"), and repeat the fit to signal events with the signal model. The signal event sample used here is a mixture of four D^0 decay modes according to their branching ratios, and it size is equivalent to 20 fb^{-1}.

By comparing the best-fit values for τ_{B^0} and Δm_d between the nominal "perfect" alignment and the degraded models of real detector alignment, we estimate the uncertainty in the result from data due to the uncertainty about the true internal alignment of the SVT. Table 11.3 summarizes the change in physics and resolution parameters in the two (mis)alignment scenarios diffDL and diffEL with respect to the default case. We assign the average change of these two cases as the systematic uncertainties: 0.0056 ps for τ_{B^0} and 0.0030 ps^{-1} for Δm_d.

Table 11.3: Parameter values and their change from nominal values for a full signal Δt fit (with Monte Carlo truth tagging information) to the signal Monte Carlo cocktail in two different SVT alignment scenarios.

Parameter	Nominal Value	Δ diffDL	Δ diffEL
Δm	0.477 ± 0.00516	-0.00316	-0.00292
τ_{B^0}	1.53 ± 0.0155	0.0057	0.0055
κ	0.812 ± 0.131	0.0452	0.00591
s^1	0.969 ± 0.0423	0.0569	0.108
$f^1_{\rm KT}$	0.507 ± 0.0767	-0.0213	0.0159
$f^1_{\rm LT}$	0.687 ± 0.0812	-0.0254	0.00904
$f^1_{\rm N1}$	0.831 ± 0.0839	-0.0997	0.0589
$f^1_{\rm N2}$	0.555 ± 0.084	0.0059	-0.0105
$f^1_{\rm N3}$	0.566 ± 0.069	-0.05	0.00131
$b^{\rm out}$	-0.440 ± 0.551	0.112	0.368
$f^{\rm out}$	0.0291 ± 0.00498	0.000435	0.00319

Beam spot

The beams pot is used as a constraint in both the tag-side and reconstructed-side vertices. We investigate the dependence of the fitted values of τ_{B^0} and Δm_d on the beam-spot parameters by varying the beam spot position in y direction, which provides most constraining power. The Monte Carlo signal event vertices are recalculated after the beam spot position is shifted consistently for every event by up to 80 μm, or randomly for each event according to a Gaussian distribution with widths up to 80 μm, and refitted with signal Δt model. The results are listed in Table. 11.4.

Based on these observations, and knowing that the difference in estimated beams pot position between different algorithm is less than 10 μm, and that average resolution of the beam spot size in y direction is about 30 μm, we estimate the systematic error due to uncertainties in the beams pot position to be ± 0.005 ps for τ_{B^0} and ± 0.001 ps^{-1} for Δm_d.

Table 11.4: Difference in final parameter values from the full signal fit to the signal Monte Carlo samples for different variations applied to the beam spot parameters.

	Beams pot Variation [μm]			
Parameter	shift 10	shift 30	shift 80	shift/random 20
Δm	-0.0005	-0.0013	-0.0020	0.0001
τ_{B^0}	-0.0042	-0.0001	0.0012	-0.0024
Parameter	random 20	random 40	random 80	shift/random 80
Δm	0.0005	-0.0002	-0.0030	-0.0056
τ_{B^0}	-0.0050	-0.0036	-0.0029	0.0024

Background fractions

In this analysis, the signal and background probabilities are calculated based on the fitted $m(D^*)-m(D^0)$ shapes, and peak and combinatoric background yields of small

subsamples. The shape parameters and yields are kept constant in the final Δt fit. The uncertainty on τ_{B^0} and Δm_d due to the statistical fluctuation of the signal and background probability calculation is estimated in the following way. The δm shape parameters and peak and background yields are perturbed randomly according to the errors from the fit results, assuming Gaussian distribution and taking into account all correlations, and then the Δt fit is carried out to estimate this contribution to uncertainties.

A faster test, where only τ_{B^0} and Δm_d are floating in the full Δt model, is performed on 111 different perturbations of the δm parameters. A full test with all 72 parameters floating is also performed on 20 different random perturbations to verify that the 2-parameter fit results are not an inappropriate estimate. The results are shown in Fig. 11.7 and Tab. 11.5. A systematic error of ± 0.0029 ps^{-1} on Δm_d and ± 0.0032 ps on τ_{B^0} is assigned, according to the 2-parameter fit distributions.

Table 11.5: Characteristics of the distribution of fitted results for Δm_d and τ_{B^0} (with respect to the baseline) from the 111 fits with perturbed background fraction when floating only the two physics parameters.

	$\delta(\Delta m_d)$ (ps^{-1})	$\delta(\tau_{B^0})$ (ps)
Mean	-0.0013 ± 0.0003	-0.0006 ± 0.0003
RMS	0.0029 ± 0.0002	0.0032 ± 0.0002

Resolution models

The data is fitted with an alternative resolution function for signal Δt model, a triple-Gaussian (G+G+G). The outlier Gaussian shape is kept the same; i.e., $\sigma_{\rm out} = 6$ ps and $b_{\rm out} = -5$ ps. The biases and widths of the other Gaussians are scaled by $\sigma_{\Delta t}$. The number of free parameters is 74, two more than that of default fit. In comparison with the default fit, the lifetime shifts by $+0.0034$ ps and Δm_d shifts by

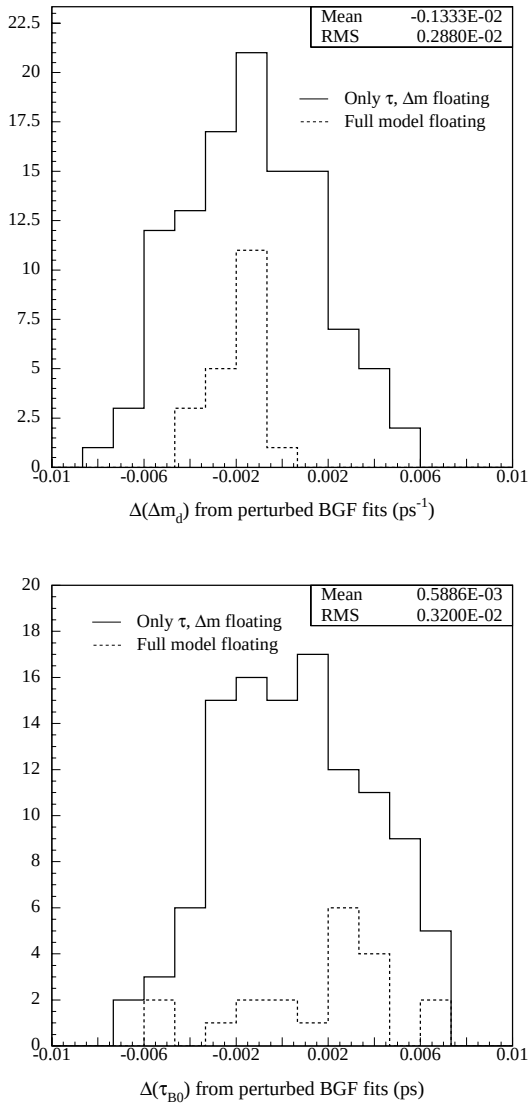

Figure 11.7: Histograms of the spread in fitted results with respect to the default fit result from Data with 111 different statistical variations of the input $m(D^*)-m(D^0)$ parameters and floating only τ_{B^0} and Δm_d in the final Δt fits. The superimposed shaded histograms show the distribution of fit results for the 20 fits where the entire Δt model was floating in the fit.

-0.0009 ps^{-1}. The minimized negative log-likelihood is 795904.4, (for reference, the

value for the default fit is 795903.8.).

Although most of the parameters in the Δt distributions for the backgrounds are floating, there could be a bias in the result due to the choice of models. Since the largest background contribution is from combinatoric events, only the combinatoric model is studied for the systematic error. An important feature of the background Δt model is its effort to accommodate the mixing structure of the combinatoric background. As a suitable variation, the model is changed with a pure lifetime-based model. Comparing the results from a full fit to data gives us an estimate of our systematic error due to the choice of model for the largest background.

The fit result with this combinatoric model shows that the lifetime shifts by $+0.0063$ ps and Δm_d shifts by -0.0012 ps^{-1}. It is worth noticing that the minimized negative log-likelihood is 795996.3, which is much larger than the default fit value, 795903.8. The charged B fraction, $f_{B^\pm}$, also changes from 8% to 6%. Although this is an extreme test for exposing our sensitivity to assumptions about the functional form of the background Δt distributions, the size of this shift is used as an estimate of the systematic uncertainty.

Summary

The results of this section are summarized in Table 11.6, which lists each source of systematic uncertainty and our evaluation of its contribution to the error on τ_{B^0} and Δm_d, and the sum in quadrature.

Table 11.6: Summary of systematic uncertainties of τ_{B^0} and Δm_d.

Source	$\delta(\Delta m)$ (ps^{-1})	$\delta(\tau_{B^0})$ (ps)	Note
Selection and fit correction	± 0.0123	± 0.0178	Stat. error on generic MC fit
z scale	± 0.0020	± 0.0060	0.4%
PEP-II boost	± 0.0005	± 0.0015	0.1%
SVT alignment	± 0.0030	± 0.0056	test on modified alignment constants with sig MC
Beam spot position	± 0.0010	± 0.0050	Move/smear BS in simulation
Bkg / signal prob.	± 0.0029	± 0.0032	Vary $m(D^*) - m(D^0)$ fit result
Fixed B^+/B^0 lifetime ratio	∓ 0.0003	± 0.0019	Vary lifetime ratio by $\pm 1\sigma$ [6]
Fixed B^+/B^0 mistag ratio	∓ 0.0001	∓ 0.0003	Vary mistag ratios by $\pm 1\sigma$ [72]
Fixed signal outlier shape	± 0.0010	± 0.0054	Vary outlier width & bias
Signal resolution model	± 0.0009	± 0.0034	G+G+G vs GExp+G
Bkg Δt models	± 0.0012	± 0.0063	Vary combinatoric bkg model
Total syst. error	± 0.013	± 0.022	

Chapter 12

Conclusion and Outlook

With approximately 14,000 exclusively reconstructed $B^0 \to D^{*-}\ell^+\nu_\ell$ signal events, the B^0 lifetime τ_{B^0} and B^0-$\overline{B}^0$ oscillation frequency Δm_d have been measured simultaneously, with an unbinned maximum-likelihood fit. The results are

$$\tau_{B^0} = (1.523\,^{+0.024}_{-0.023} \pm 0.022)\ \text{ps}\,,$$

and

$$\Delta m_d = (0.492 \pm 0.018 \pm 0.013)\ \text{ps}^{-1}\,,$$

where the first uncertainty is statistical and the second is systematic. The statistical correlation coefficient between τ_{B^0} and Δm_d is -0.22. Both the lifetime and mixing frequency have combined statistical and systematic uncertainties that are comparable to those of most precise previously-published experimental measurements [6]. The results are consistent with the world average measurements of $\tau_{B^0} = (1.542 \pm 0.016)$ ps and $\Delta m_d = (0.489 \pm 0.008)\ \text{ps}^{-1}$ [6].

The comparisons of the results in this thesis and other measurement, along with the improved world average including the results in this thesis are shown in Figs. 12.3 and 12.4. Figures 12.1 and 12.2 show the comparisons among the most precise

measurements.

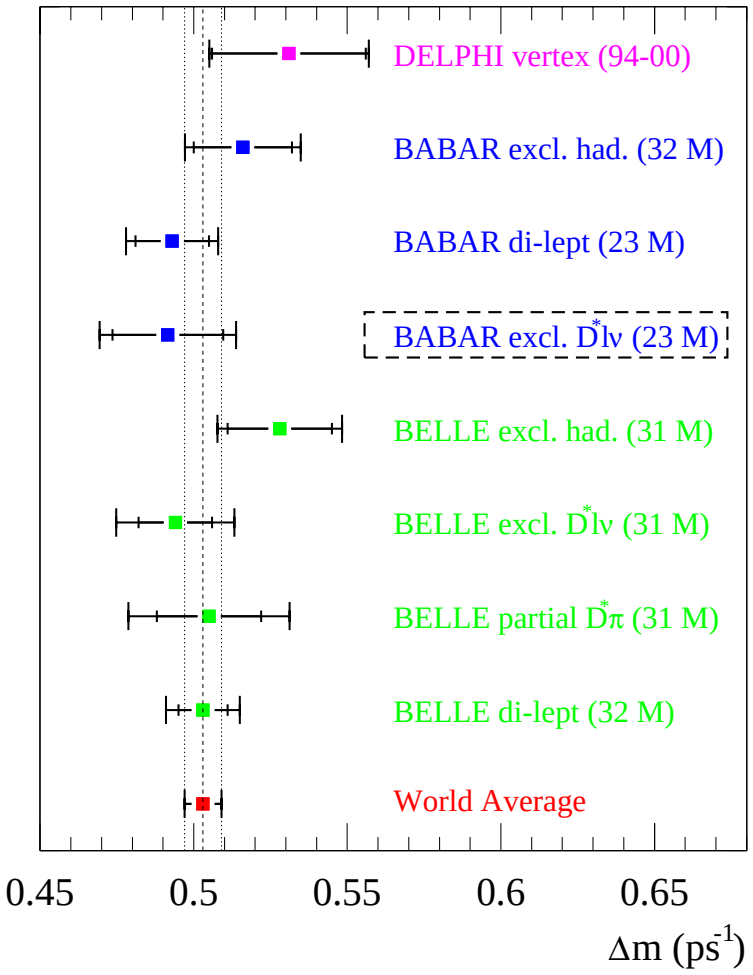

Figure 12.1: Mixing frequency comparison with most precise measurements [90, 91, 92, 93, 94, 95, 96, 97]. The numbers in the parentheses are either the number of $B\overline{B}$ that particular experiment uses (e.g., 23 M), or the years the data are collected in (e.g., 91–00). The dashed box indicates the measurement in this thesis.

This analysis is still statistically limited. In addition to τ_{B^0} and Δm_d, a large number of parameters are floating in the fitting process. The statistical uncertainty is increased because of the correlations between τ_{B^0}, Δm_d and other floating parameters. However, the advantage of this approach is that one can avoid making

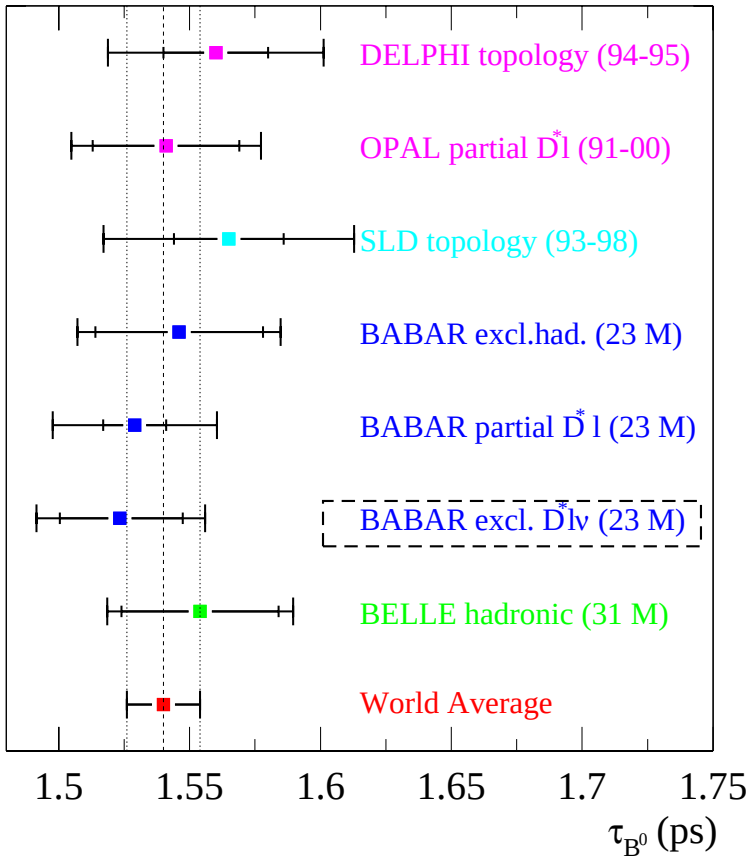

Figure 12.2: Lifetime comparison with most precise measurements [98, 52, 53, 54, 55, 99, 93, 56]. The numbers in the parentheses are either the number of $B\overline{B}$ that particular experiment uses (e.g., 23 M), or the years the data are collected in (e.g., 91–00). The dashed box indicates the measurement in this thesis.

many assumptions, such as the charged B fraction, resolution functions and background behaviors, so that complex systematic uncertainty studies can be reduced. This analysis tries to use data as much as possible, so that the dependence on the simulation can be minimized, thus the systematic uncertainties due to the possible inconsistency between data and simulation can be reduced.

Although the systematic uncertainties are not very small compared to the statistical uncertainties, the dominant systematic uncertainties for both lifetime and mixing

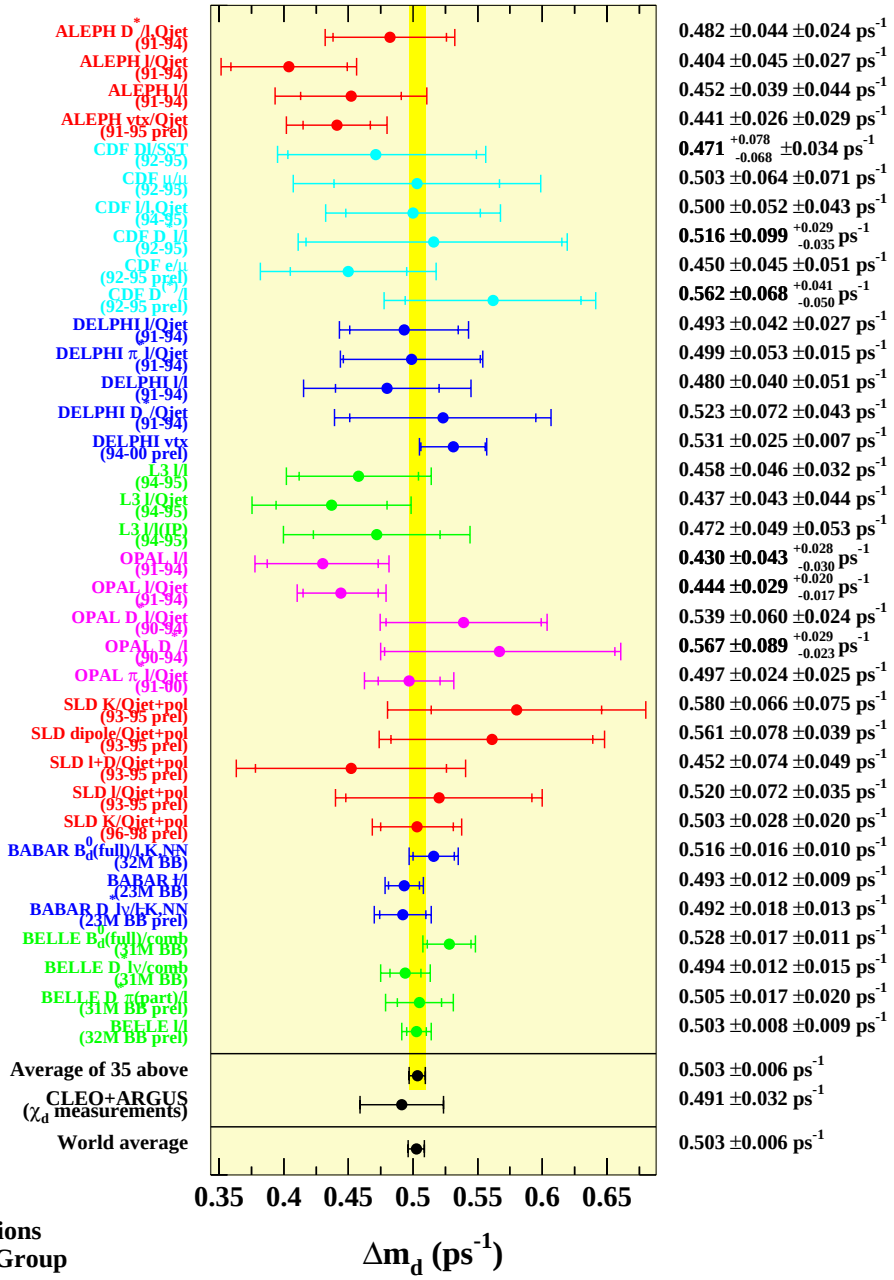

Figure 12.3: Mixing comparison with world average and other experiments [98].

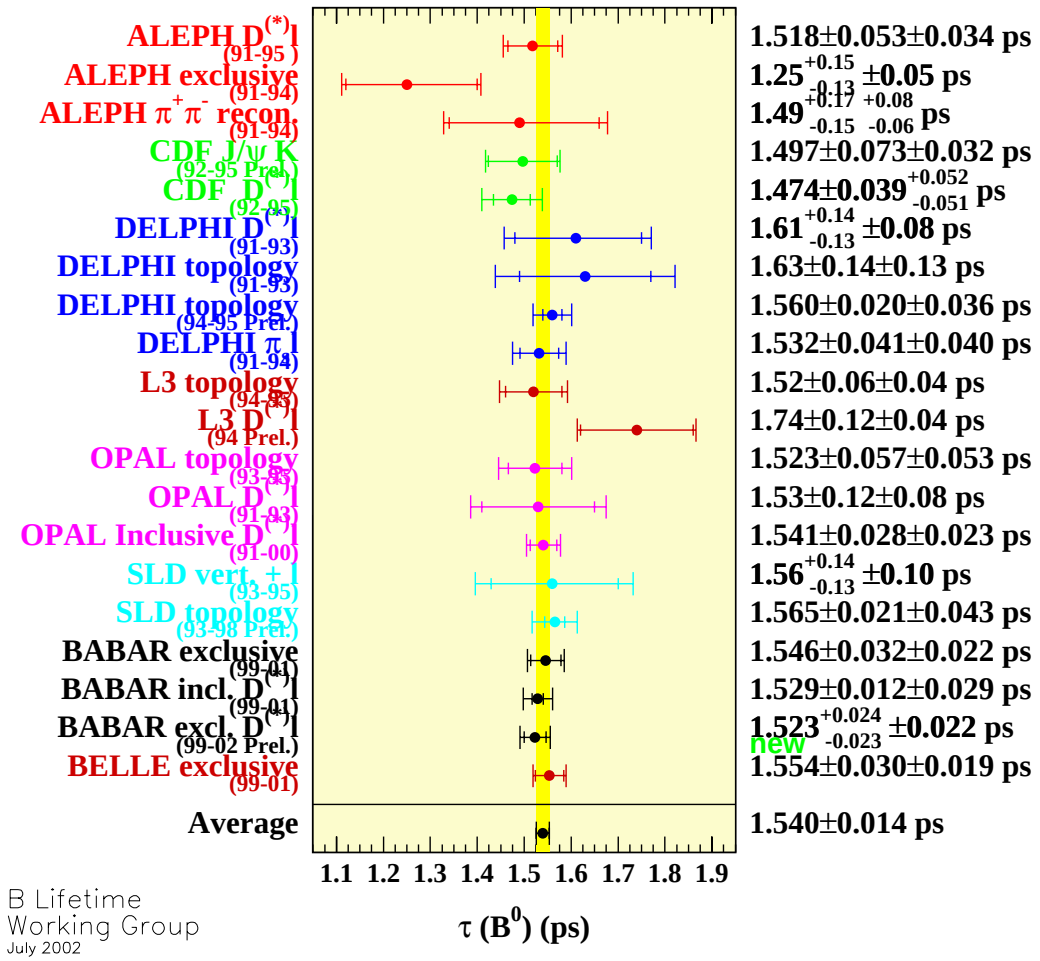

Figure 12.4: Lifetime comparison with world average and other experiments [98].

frequency are due to the small size of simulation data, which can be improved in the near future. Up to summer of 2002, *BABAR* already accumulated about four times as much data as those used in this analysis. With existing data, and assuming that ten times as much simulation data will be available soon, the naive projection of this analysis shows that the total uncertainty for Δm_d can reach 2% and for τ_{B^0} 1.3%. This precision is approaching the Standard Model expectations of the decay

width difference $\Delta\Gamma$ between B_H and B_L, and CP violation in mixing. The physics probability functions for observing an unmixed or mixed events have to be modified to incorporate $\Delta\Gamma$ and q/p,

$$P_+ = \frac{e^{-|\Delta t|/\tau}}{4\tau}\left[\cosh(\Delta\Gamma\Delta t/2) + \cos(\Delta m\Delta t)\right] \tag{12.1}$$

$$P_- = \frac{e^{-|\Delta t|/\tau}}{4\tau}\left[\cosh(\Delta\Gamma\Delta t/2) - \cos(\Delta m\Delta t)\right] \cdot \frac{1}{2}\left(\left|\frac{p}{q}\right|^2 + \left|\frac{q}{p}\right|^2\right) . \tag{12.2}$$

The $\cosh(\Delta\Gamma\Delta t/2)$ term will make the overall probability larger than the pure exponential since it is always greater than one, and thus bias the measured lifetime higher than the true value if this term is ignored. The factor of $\frac{1}{2}(|p/q|^2 + |q/p|^2)$ will have a very small impact on the overall mixing probability since

$$\frac{1}{2}\left(\left|\frac{p}{q}\right|^2 + \left|\frac{q}{p}\right|^2\right) - 1 \simeq 2\left(\left|\frac{q}{p}\right| - 1\right)^2 = \mathcal{O}\left(10^{-4}\text{–}10^{-6}\right)$$

if we take $|q/p| = 1 + \mathcal{O}\left(10^{-2}\text{–}10^{-3}\right)$. The decay rate asymmetry between $|B^0_{\rm phys}(t)\rangle$ and $|\overline{B}^0_{\rm phys}(t)\rangle$ shown in Eq. 2.34 has a better sensitivity to $|q/p| - 1$, but a better understanding of charge asymmetry in the detector is needed.

As the statistical uncertainties quickly decreases due to large amount of data, to carry this analysis forward, many cares should be taken to reduce the systematic uncertainties:

- First of all, the bias observed in the simulation data must be understood. A very large amount of Monte Carlo simulation must be generated, and a better tuning on simulation to match real data is needed.

- The resolution functions and background time dependent models must be revisited to describe the behaviors in data with higher precisions. Especially it is already known (Sec. 6.2.1) that the resolution is correlated with the true decay

time difference. This effect must be modeled properly. In addition, the resolution function for the charged B component should be separated from the signal resolution function since an extra soft charged particle from the B^+ decay is in principle included in the tagging side vertex.

- With a larger data size, the outlier component of resolution model should be easier to handle. Nevertheless, better understanding of the origin and behavior of these outliers, and reducing them are essential.

- The distribution of $\cos\theta_{B^0,D^*\ell}$ contains information of charged B background. It will increase the sensitivity to the estimation of charged B fraction.

- Better models for mistag-$\sigma_{\Delta t}$ correlation, including all tagging categories besides kaon tag, for both signal and background are necessary.

- The mistag rate for charged B in this analysis needs more study because it can be different from the mistag rate for pure $B^\pm$ events since an extra soft charged track can be included in the tagging side.

- It is better to understand the more fundamental cause of the $m(D^*) - m(D^0)$ resolution difference between whether or not the soft pion has DCH hits, so that a more sensible separation of data sample can be applied to increase sensitivities.

This list is certainly not complete. The underline true Δt distribution can be complicated by many interference terms [100] from effects such as CPT violation and CP violation in decay, and doubly-CKM suppressed decays in $B_{\rm tag}$ that will fake tagging flavors.

With future upgrades of PEP-II and *BABAR*, the accumulative data size will reach $\sim$500 fb^{-1} in four years.*BABAR* will have a good chance to observe the CP violation

and make a strong constraint on CPT symmetry, or discover new phenomena that are forbidden or highly suppressed in Standard Model. The analysis and the technique in this thesis can be carried forward and contribute to this process.

Appendix A

Sample Composition

This appendix summarizes the peak yields and background fractions for different "supersamples" other than those shown in Sec. 9.8 to demonstrate the fact that the background levels are significantly different for some event categories.

Table A.1: Peak yields and fake, continuum and uncorrelated fractions of the peak yield, and the combinatoric fraction of total events in $m(D^*) - m(D^0)$ window for signal samples (e/μ, opposite-side and on-Resonance) for the soft pion in the SVT only or SVT+DCH (Fig. A.1).

Category	Peak Yield	$f_{\text{cont.}}(\%)$	$f_{\text{fake}}(\%)$	$f_{\text{uncorr.}}(\%)$	$f_{\text{comb.}}(\%)$
e;SVT	5427 ± 81	1.26 ± 0.44	0.1468 ± 0.0045	3.71 ± 0.48	19.50 ± 0.29
e;DCH	1581 ± 41	2.5 ± 1.1	0.240 ± 0.010	1.19 ± 0.47	11.83 ± 0.37
μ;SVT	5053 ± 78	2.62 ± 0.69	2.352 ± 0.073	3.20 ± 0.53	20.31 ± 0.31
μ;DCH	1517 ± 41	1.08 ± 0.83	3.73 ± 0.16	1.7 ± 1.1	11.14 ± 0.37

Table A.2: Peak yields and fake, continuum and uncorrelated fractions of the peak yield, and the combinatoric fraction of total events in $m(D^*)-m(D^0)$ window for signal samples (e/μ, opposite-side and on-Resonance) for different D^0 modes (Fig. A.2).

Category	Peak Yield	$f_{\rm cont.}(\%)$	$f_{\rm fake}(\%)$	$f_{\rm uncorr.}(\%)$	$f_{\rm comb.}(\%)$
e;$K\pi$	2623 ± 53	0.89 ± 0.51	0.1692 ± 0.0063	3.19 ± 0.57	7.02 ± 0.27
e;$K\pi\pi\pi$	2219 ± 54	2.47 ± 0.93	0.1531 ± 0.0083	3.34 ± 0.77	28.63 ± 0.50
e;$K\pi\pi^0$	2166 ± 51	1.36 ± 0.75	0.1811 ± 0.0073	2.89 ± 0.68	16.85 ± 0.45
μ;$K\pi$	2491 ± 52	1.53 ± 0.70	2.657 ± 0.098	2.81 ± 0.80	7.38 ± 0.29
μ;$K\pi\pi\pi$	1939 ± 51	3.0 ± 1.3	2.61 ± 0.14	2.71 ± 0.81	30.93 ± 0.54
μ;$K\pi\pi^0$	2139 ± 50	2.5 ± 1.0	2.73 ± 0.11	3.01 ± 0.86	16.10 ± 0.44

Table A.3: Peak yields and fake, continuum and uncorrelated fractions of the peak yield, and the combinatoric fraction of total events in $m(D^*) - m(D^0)$ window for signal samples (e/μ, opposite-side and on-Resonance) for different tagging categories (Fig. A.3).

Category	Peak Yield	$f_{\rm cont.}(\%)$	$f_{\rm fake}(\%)$	$f_{\rm uncorr.}(\%)$	$f_{\rm comb.}(\%)$
e;LTag	783 ± 29	0.0 ± 0.0	0.1711 ± 0.0079	2.41 ± 0.99	8.20 ± 0.57
e;KTag	2565 ± 55	3.00 ± 0.95	0.1693 ± 0.0075	3.61 ± 0.66	17.86 ± 0.40
e;NT1	630 ± 27	1.2 ± 1.2	0.153 ± 0.013	3.7 ± 1.3	14.26 ± 0.76
e;NT2	921 ± 33	1.6 ± 1.1	0.193 ± 0.011	3.1 ± 1.1	20.87 ± 0.72
e;NT3	2108 ± 51	0.39 ± 0.51	0.1580 ± 0.0081	2.68 ± 0.71	20.74 ± 0.46
μ;LTag	746 ± 28	0.0 ± 0.0	2.68 ± 0.12	3.2 ± 1.2	8.34 ± 0.60
μ;KTag	2393 ± 53	2.9 ± 1.1	2.71 ± 0.12	2.95 ± 0.80	18.59 ± 0.42
μ;NT1	545 ± 25	1.3 ± 1.4	2.65 ± 0.23	4.5 ± 1.5	15.09 ± 0.83
μ;NT2	958 ± 34	3.1 ± 1.5	2.77 ± 0.15	3.0 ± 1.7	19.37 ± 0.69
μ;NT3	1928 ± 49	2.2 ± 1.1	2.58 ± 0.13	2.02 ± 0.73	21.76 ± 0.49

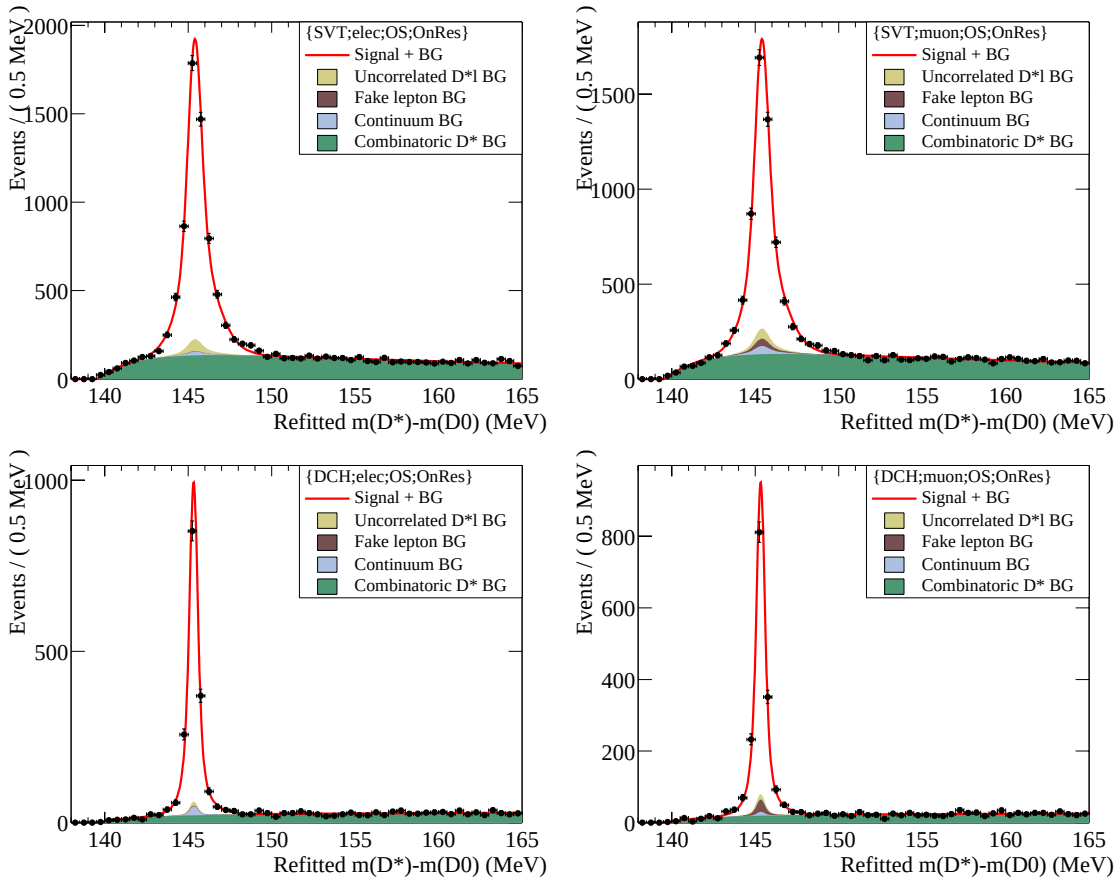

Figure A.1: Signal, and combinatoric and peaking background contributions for the signal samples. Left: electron sample; right: muon sample. Up: SVT-only; bottom: SVT+DCH.

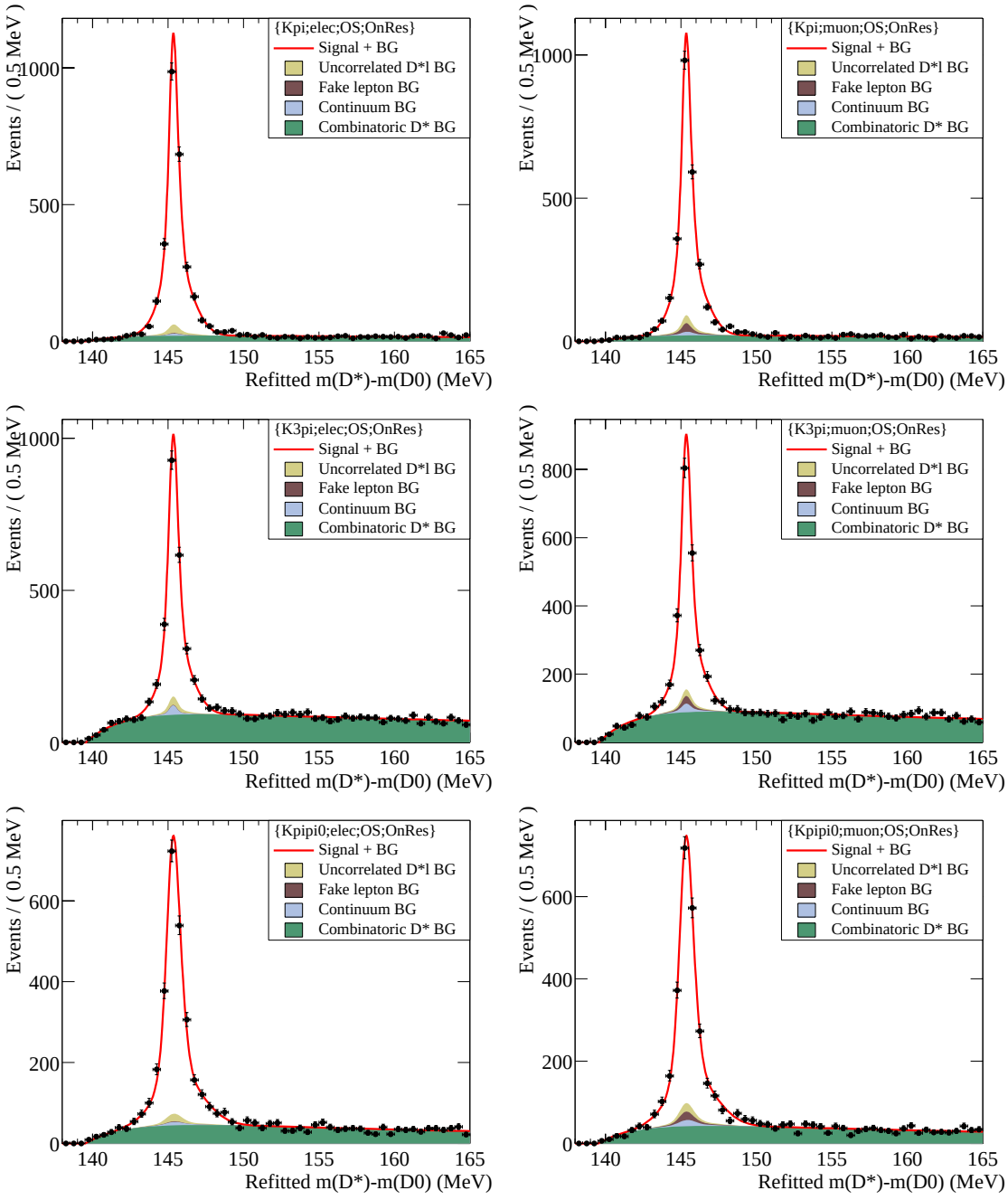

Figure A.2: Signal, and combinatoric and peaking background contributions for the signal samples. Left: electron sample; right: muon sample. From top to bottom, $D^0 \to K\pi$,$D^0 \to K\pi\pi\pi$, and $D^0 \to K\pi\pi^0$ modes.

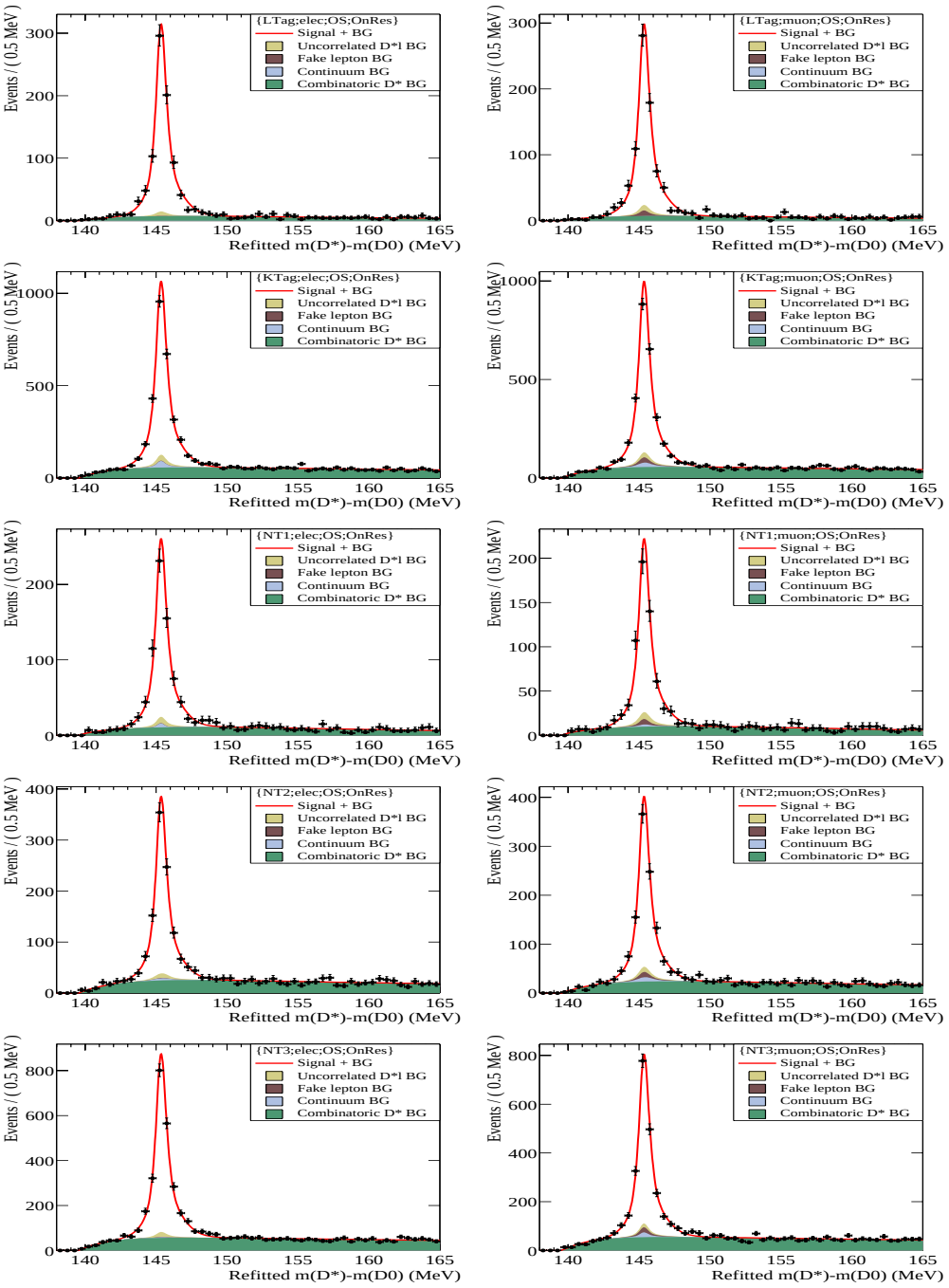

Figure A.3: Signal, and combinatoric and peaking background contributions for the signal samples. Left: electron sample; right: muon sample. From top to bottom, `LTag`, `KTag`, `NT1`, `NT2`, and `NT3` tagging categories.

Appendix B

Additional Δt-Projection Plots

In this appendix, additional plots that illustrate the performance of the full fit result to data, especially the background components. Events are selected according to the signal and background probabilities based on $m(D^*) - m(D^0)$ fits. Figure B.1 shows the Δt distribution and mixed-unmixed asymmetry for a subsample that are 99.5% pure combinatoric background events. Figure B.2 shows the same distributions for a subsample that are 60% pure fake-lepton background events.

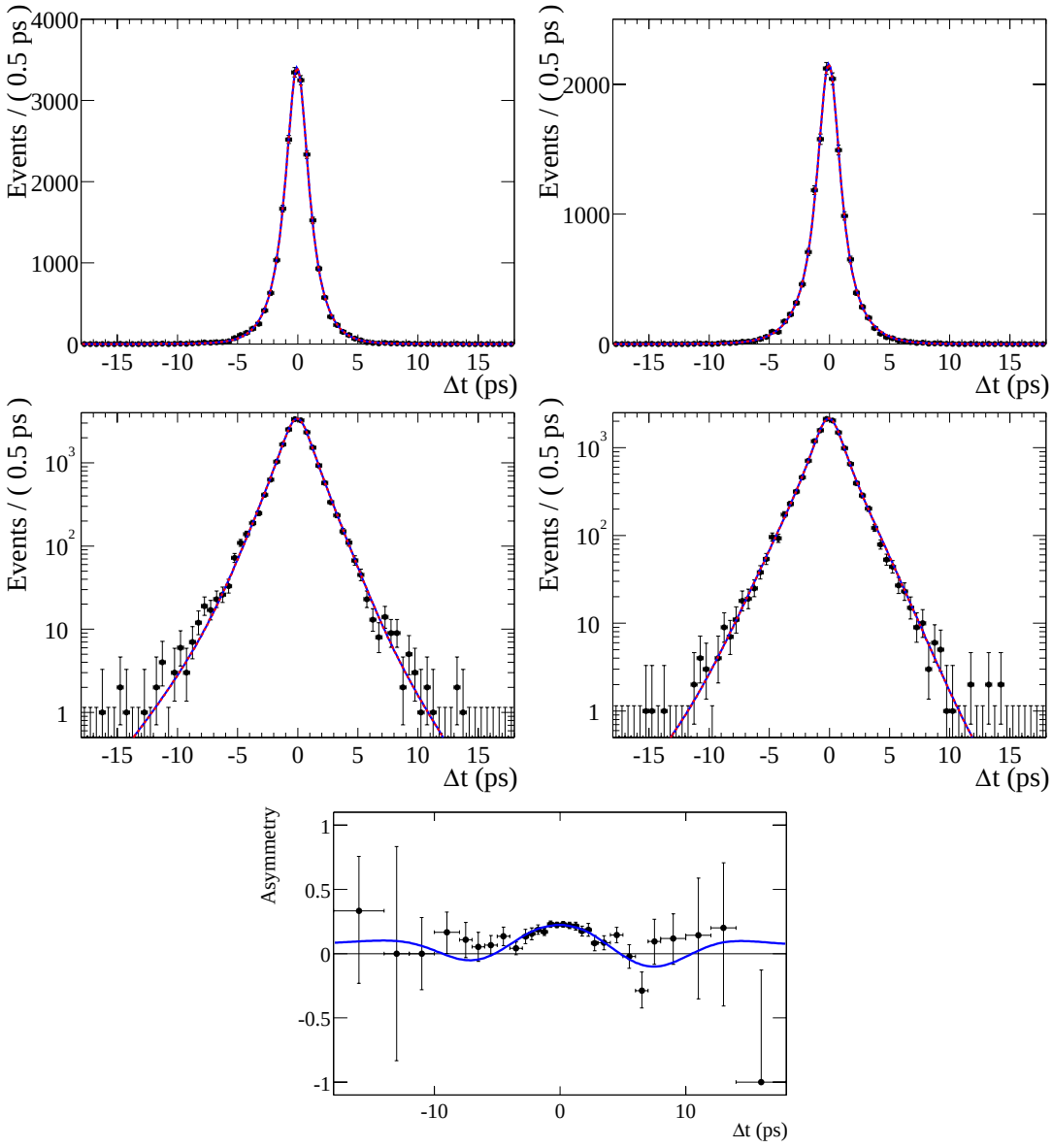

Figure B.1: The Δt distribution and mixed-unmixed asymmetry for combinatoric background and the projection of the fit results. The left-hand plots are for unmixed events and the right-hand plots for mixed events. The bottom plot is the mixed-unmixed asymmetry. The events plotted here are from a sample of 99.5% pure combinatoric background events.

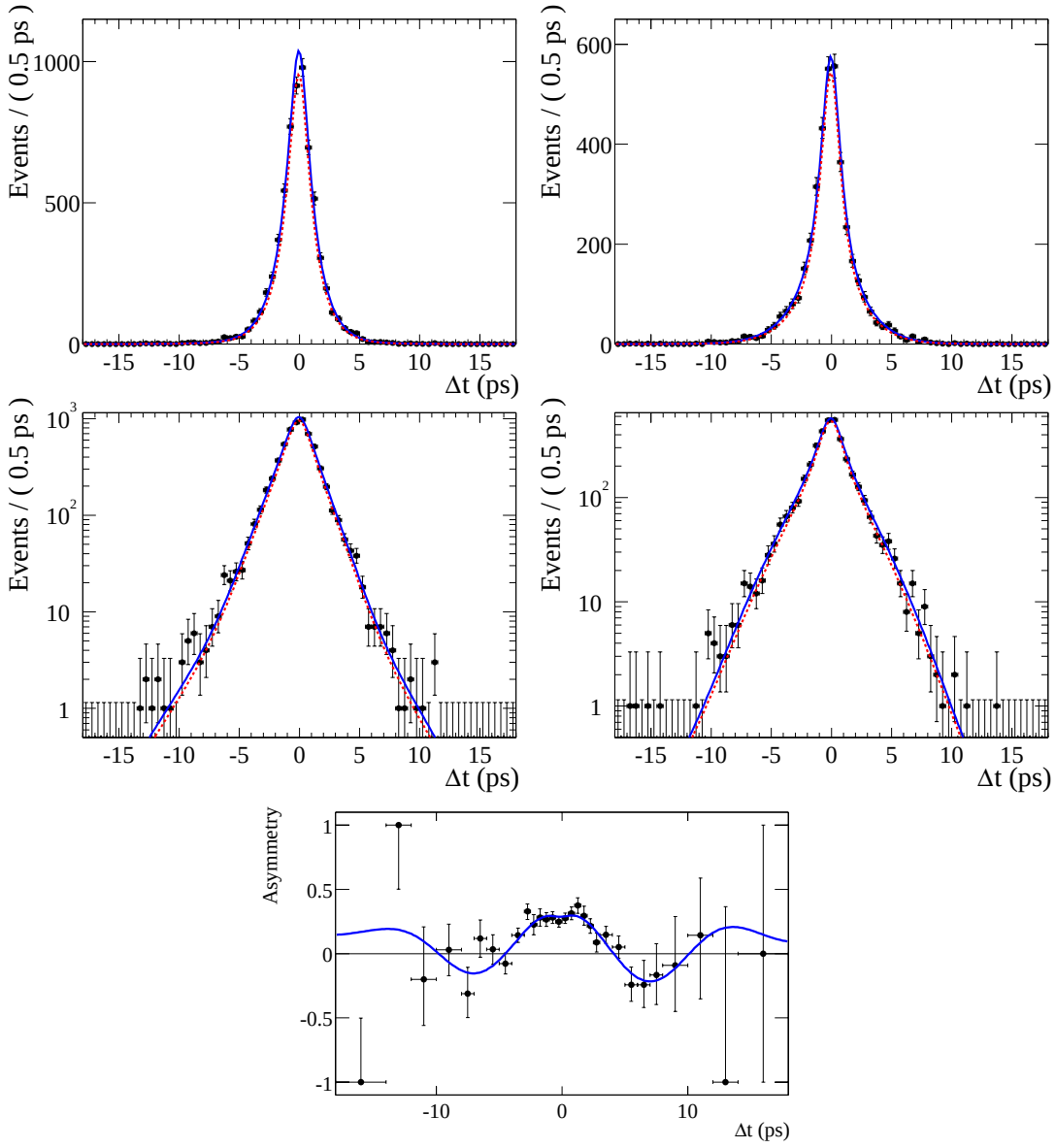

Figure B.2: The Δt distribution and mixed-unmixed asymmetry for fake-lepton background and the projection of the fit results. The left-hand plots are for unmixed events and the right-hand plots for mixed events. The bottom plot is the mixed-unmixed asymmetry. The events plotted here have an average purity of 60% being fake-lepton events. The red-dashed curves are the contribution from events that are not signal (including combinatoric background).

Appendix C

Dalitz Plot Calculation

For a three-body decay $M \to a + b + c$, the full kinematics can be described by two parameters; e.g., the invariant mass of a+b and a+c systems (m_{ab}, m_{ac}). See, for example, a review in [9]. The phase space factor is uniform on (m_{ij}, m_{ik}) plane within the kinematically allowed range, where $i \neq j \neq k$, $i \neq k$. The event distribution in this plane is called the Dalitz plot. If a fraction of this three-body decay amplitude comes from a resonance that decays to particles 1 and 2; i.e., $M \to (ab)+c \to a+b+c$, a peak will show up in $a + b$ invariant mass distribution and a higher-density band perpendicular to m_{ab} axis will show up in the Dalitz plot. By studying the Dalitz plot, one can extract the fractions of decay amplitudes and phases through resonances.

For the analysis descirbed in this thesis, the known resonance structure on the Dalitz plot is used to select events within the high density regions in order to increase signal to background ratios (see Sec. 8.6). In this appendix, I document the formulas and parameters (based on the E687 experiment at the Fermilab [77]) used to calculate the expected Dalitz plot density.

The total amplitude of a three-body decay can be expressed as

$$A_{\text{total}} = A_0 e^{i\phi_0} + \sum_i A_i e^{i\phi_i} B_{\text{bw}} S \,,$$

where the first term represents the non-resonant component, A_i and ϕ_i are the amplitude and phase for resonance i, respectively, B_{bw} is the Breit-Wigner function [101], and S is an angular correlation function due to the spin of the resonance. In particular, for a resonance $r \to a + b$, B_{bw} can be written as

$$B_{\text{bw}}(a, b|r) = \frac{F_D F_r}{m_r^2 - m_{ab}^2 - i\Gamma m_r} \,,$$

where m_r is the peak of the resonance r, Γ an energy-dependent width, and F_D and F_r are Blatt-Weisskopf peneration factors [102]. The angular functions S are different for different resonance spin:

$$\begin{aligned}
S &= 1 && \text{for spin= } 0 \,, \\
S &= 2 p_b p_c \cos\theta_{bc} && \text{for spin= } 1 \,, \\
S &= 2 p_b^2 p_c^2 (3\cos^2\theta_{bc} - 1) && \text{for spin= } 2 \,,
\end{aligned}$$

where p_b and p_c are the magnitudes of three-momentum of b and c and θ_{bc} is the angle between $\vec{p}_b$ and $\vec{p}_c$, all measured in the a+b rest frame.

The factors F_D and F_r are chosen to be

$$\begin{aligned}
F &\sim 1 && \text{for spin= } 0 \,, \\
F &\sim \frac{1}{\sqrt{R^2 p_b^2}} && \text{for spin= } 1 \,, \\
F &\sim \frac{1}{\sqrt{9 + 3R^2 p_b^2 + R^4 p_b^4}} && \text{for spin= } 2 \,,
\end{aligned}$$

where $R = 1.5$ GeV^{-1} is the Blatt and Weiskopf meson radius[1].

The energy-dependent width for spin=s is

$$\Gamma = \Gamma_0 \frac{m_r}{m_{ab}} \left(\frac{p_b}{p_b^R} \right)^{2s+1} \frac{F^2}{F_R^2},$$

where p_b^R is the b momentum in the a+b rest frame when $m_{ab} = m_r$, and F_R is the F factor when p_b is substituted by p_b^R.

The parameters used to calculate the Dalitz plot density are taken from the E687 result [77], and are shown in Table C.1. The results presented [77] are in terms of

Table C.1: Parameters used for Dalitz plot calculations.

$D^0 \to K^-\pi^+\pi^0$					
resonance	A	ϕ	Γ_0 (GeV)	m_r (GeV)	spin
K^{*0}	0.0862	$-2°$	0.0507	0.8961	1
ρ^+	0.1942	$0°$	0.1502	0.7693	1
K^{*-}	0.0813	$162°$	0.0508	0.8917	1
non-res	0.1784	$-122°$	-	-	-
$D^0 \to \overline{K^0}\ \pi^+\ \pi^-$					
resonance	A	ϕ	Γ_0 (GeV)	m_r (GeV)	spin
K^{*-}	0.1703	$0°$	0.0508	0.8917	1
K^{*0} (1430)	0.1222	$-166°$	0.294	1.412	0
ρ^0	0.1324	$-136°$	0.1502	0.7693	1
$f_0(975)$	0.0283	$38°$	0.07	0.98	0
$f_2(1270)$	0.6032	$-174°$	0.1851	1.2754	2
$f_0(1400)$	0.1622	$-45°$	0.35	1.4	0

decay fractions, not amplitude A. Here I have normalized the amplitudes so that the ratio of the integration of the signal intensity for a certain mode to the integration with all modes combined matches the decay fractions in [77].

[1] In principle this value should be different for different resonances, but the final result is not sensitive to the choice of R.

Bibliography

[1] C. Albajar *et al.*, [UA1 Collaboration], "Search for B^0 $\overline{B}^0$ oscillations at the CERN proton-antiproton collider," Phys. Lett. B **186**, 247–254 (1991), [Erratum: *ibid* B197, 565 (1987)].

[2] H. Albrecht *et al.*, [ARGUS Collaboration], "Observation of B^0-$\overline{B}^0$ mixing," Phys. Lett. B **192**, 245 (1987).

[3] N. Cabibbo, "Unitary symmetry and leptonic decays," Phys. Rev. Lett. **10**, 531–532 (1963).

[4] M. Kobayashi and T. Maskawa, "CP violation in the renormalizable theory of weak interaction," Prog. Th. Phys. **49**, 652–657 (1973).

[5] M. Neubert and C. T. Sachrajda, "Spectator effects in inclusive decays of beauty hadrons," Nucl. Phys. B **483**, 339 (1997).

[6] K. Hagiwara *et al.*, [Particle Data Group Collaboration], "Review of particle physics," Phys. Rev. D **66**, 010001+ (2002).

[7] "The *BABAR* collaboration, letter of intent for the study of CP violation and heavy flavor physics at PEP-II," Technical Report (1994). SLAC-443.

[8] "PEP-II: An asymmetric B factory, conceptual design report," Technical Report (1993). SLAC-418, LBL-5379.

[9] D. Groom *et al.*, [Particle Data Group Collaboration], "Review of particle physics," Eur. Phys. Jour. C **15**, 1+ (2000).

[10] E. Noether, Nachr. Kgl. Ges. Wiss. Gottingen page 235 (1918).

[11] Glashow, "Partial symmetries of weak interactions," Nucl. Phys. **22**, 579 (1961).

[12] S. Weinberg, "A model of leptons," Phys. Rev. Lett. **19**, 1264 (1967).

[13] A. Salam, "Elementary particle theory: relativistic groups and analyticity," Nobel Symposium **8**, 367 (1968).

[14] P. W. Higgs, "Broken symmetries, massless particles and gauge fields," Phys. Lett. **12**, 132 (1964).

[15] L. Wolfenstein, "Parameterization of the Kobayashi-Maskawa matrix," Phys. Rev. Lett. **51**, 1945 (1983).

[16] L. Schiff, *Quantum Mechanics*, McGraw-Hill, New York, 3rd edition (1955).

[17] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, "Experimental test of parity conservation in beta decay," Phys. Rev. **105**, 1413–1414 (1957).

[18] J. H. Christenson *et al.*, "Evidence for the 2 π decay of the K_2^0 meson," Phys. Rev. Lett. **13**, 138 (1964).

[19] B. Aubert *et al.*, [*BABAR* Collaboration], "Observation of CP violation in the B^0 meson system," Phys. Rev. Lett. **87**, 091801 (2001).

[20] K. Abe *et al.*, [Belle Collaboration], "Observation of large CP violation in the neutral B meson system," Phys. Rev. Lett. **87**, 091802 (2001).

[21] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, Addison-Wesley (1995). Reading, USA. 842 p.

[22] C. Jarlskog, "Phenomenology of CP violation," (1988). Based on lectures given at Int. Meeting on Fundamental Physics, CP Nonconservation and B Physics, Peniscola, Spain, Apr. 25–29.

[23] V. F. Weisskopf and E. P. Wigner, "Calculation of the natural brightness of spectral lines on the basis of Dirac's theory," Z. Phys. **63**, 54 (1930).

[24] L. Wolfenstein, "Violation of CP invariance and the possibility of very weak interactions," Phys. Rev. Lett. **13**, 562 (1964).

[25] G. Buchalla, A. J. Buras, and M. E. Lautenbacher, "Weak decays beyond leading logarithms," Rev. Mod. Phys. **68**, 1125 (1996).

[26] L. Lellouch and C. J. D. Lin, `[UKQCD Collaboration]`, "Standard Model matrix elements for neutral B meson mixing and associated decay constants," Phys. Rev. D **64**, 094501 (2001), `hep-ph/0011086`, Revised Feb 2002.

[27] A. Ali Khan *et al.*, "B meson decay constant from two-flavor lattice QCD with non-relativistic heavy quarks," Phys. Rev. D **64**, 054504 (2001).

[28] A. S. Kronfeld and S. M. Ryan, "Remark on the theoretical uncertainty in B^0-$\overline{B}^0$ mixing," submitted to Phys. Rev. Lett. (2002), `hep-ph/0206058`.

[29] W. Pauli, "The connection between spin and statistics," Phys. Rev. **58**, 716–722 (1940).

[30] E. Fernandez *et al.*, `[MAC Collaboration]`, "Lifetime of particles containing b quarks," Phys. Rev. Lett. **51**, 1022 (1983).

[31] N. Lockyer *et al.*, [Mark II Collaboration], "Meausrement of the lifetime of bottom hadrons," Phys. Rev. Lett. **51**, 1316 (1983).

[32] I. I. Bigi, N. G. Uraltsev, and A. I. Vainshtein, "Nonperturbative corrections to inclusive beauty and charm decays: QCD versus phenomenological models," Phys. Lett. B **297**, 477 (1993).

[33] A. V. Manohar and M. B. Wise, "Inclusive semileptonic B and polarized Λ_b decays from QCD," Phys. Rev. D **49**, 1310 (1994).

[34] M. Neubert, "A fresh look at the B semileptonic branching ratio and beauty lifetimes," (1996), hep-ph/9605256.

[35] E. Richter-Was, "QED bremsstrahlung in semileptonic B and leptonic tau decays," Phys. Lett. B **303**, 163 (1993).

[36] M. Neubert, "Heavy quark symmetry," Phys. Rep. **245**, 259 (1994), and references there in.

[37] B. Barish *et al.*, [CLEO Collaboration], "Measurement of the $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_\ell$ branching fractions and $|V_{cb}|$," Phys. Rev. D **51**, 1014 (1995).

[38] D. Buskulic *et al.*, [ALEPH Collaboration], "Observation of the time dependence of B^0-$\overline{B}^0$ mixing," Phys. Lett. B **313**, 498 (1993).

[39] H. Albrecht *et al.*, [ARGUS Collaboration], "A New determination of the B^0-$\overline{B}^0$ oscillation strength," Z. Phys. C **55**, 357 (1992).

[40] H. Albrecht *et al.*, [ARGUS Collaboration], "A Study of $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_\ell$ and B^0-$\overline{B}^0$ mixing using partial D^{*+} reconstruction," Phys. Lett. B **324**, 249 (1994).

[41] J. Bartelt *et al.*, [CLEO Collaboration], "Two measurements of B^0-$\overline{B}^0$ mixing," Phys. Rev. Lett. **71**, 1680 (1993).

[42] B. H. Behrens *et al.*, [CLEO Collaboration], "Precise measurement of B^0-$\overline{B}^0$ mixing parameters at the Υ(4S)," Phys. Lett. B **490**, 36 (2000).

[43] D. Decamp *et al.*, [ALEPH Collaboration], "ALEPH: a detector for electron-positron annihilations at LEP," Nucl. Instr. Meth. A **303**, 393 (1991).

[44] K. Ahmet *et al.*, [OPAL Collaboration], "The OPAL detector at LEP," Nucl. Instr. Meth. A **305**, 275 (1991).

[45] P. A. Aarnio *et al.*, [DELPHI Collaboration], "The DELPHI detector at LEP," Nucl. Instr. Meth. A **303**, 233 (1991).

[46] A. K *et al.*, [SLD Collaboration], "Precise measurement of the left-right cross-section asymmetry in Z boson production by e^+e^- collisions," Phys. Rev. Lett. **73**, 25 (1994).

[47] A. E. Bondar *et al.*, [Belle Collaboration], "KEKB performance," Nucl. Instr. Meth. A **462**, 139 (2001).

[48] F. Abe *et al.*, [CDF Collaboration], "The CDF detector: an overview," Nucl. Instr. Meth. A **271**, 387 (1988).

[49] K. Ackerstaff *et al.*, [OPAL Collaboration], "Investigation of CP violation in $B^0 \to J/\psi {K_s}^0$ decays at LEP," Eur. Phys. Jour. C **5**, 379 (1998).

[50] J. Thom, [SLD Collaboration], "Search for B_s-$\overline{B_s}$ oscillations with a charge dipole technique at SLD," (2002), SLAC-R-585.

[51] H. G. Moser and A. Roussarie, "Mathematical methods for B^0-$\overline{B}^0$ oscillation analyses," Nucl. Instr. Meth. A **384**, 491 (1997).

[52] [DELPHI Collaboration], "A precise measurement of the B^+, B^0 mesons and mean b-hadron with the DELPHI detector at LEP I," July 2002, paper contributed to the International Conference ICHEP 2002 July 24–31, Amsterdam, Holland.

[53] G. Abbiendi *et al.*, [OPAL Collaboration], "Measurement of the B^0 lifetime and oscillation frequency using $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_\ell$ decays," Phys. Lett. B **493**, 266 (2000).

[54] [SLD Collaboration], "Measurements of the B^+ and B^0 lifetimes using topological vertexing at SLD," July 1999, paper contributed to the international conference, EPS-HEP 99, July 15–21, 1999, Tampere, Finland, SLAC-PUB-8206.

[55] B. Aubert *et al.*, [*BABAR* Collaboration], "Measurement of the B^0 and B^+ meson lifetimes with fully reconstructed hadronic final states," Phys. Rev. Lett. **87**, 201803 (2001), hep-ex/0107019.

[56] K. Abe *et al.*, [BELLE Collaboration], "Precise measurement of B meson lifetimes with hadronic decay final states," Phys. Rev. Lett. **88**, 171801 (2002), hep-ex/0202009.

[57] B. Aubert *et al.*, [*BABAR* Collaboration], "The *BABAR* detector," Nucl. Instr. Meth. A **479**, 1–116 (2002), hep-ex/0105044.

[58] J. Seeman *et al.*, "Status report on PEP-II performance," (2000), *Proceedings of the IEEE Particle Accelerator Conference (EPAC 2000)*, Vienna, Austria.

[59] M. Sullivan, "B-factory interaction region designs," (1997), *Proceedings of the IEEE Particle Accelerator Conference (PAC 1997)*, Vancouver, B.C., Canada, SLAC-PUB-7563.

[60] T. I. Meyer, "PIN photodiodes for radiation monitoring and protection in the *BABAR* Silicon Vertex Tracker," (2000), *Proceedings of the Meetings of the Division of Particles and Fields of the American Physical Society (DPF 2000)*, Columbus, OH, USA (2000), SLAC-PUB-8651.

[61] C. Hearty, [*BABAR* Collaboration], "Measurement of the number of $\Upsilon(4S)$ mesons produced in run 1 (B counting)," *BABAR* Analysis Document **134**, v.01 (2001).

[62] Sun Ultra 5, with single 333 MHz UltraSPARC-IIi CPUs and 512 Mbytes of RAM, Sun Microsystems, Inc. Palo Alto, CA, USA.

[63] P. Billoir, "Track fitting with multiple scattering: a new method," Nucl. Instr. Meth. A **225**, 352 (1984).

[64] U. Langenegger *et al.*, [*BABAR* Collaboration], "Cut-based electron identification," *BABAR* Analysis Document **90**, v.03 (2000).

[65] A. Drescher *et al.*, "Calibration and monitoring of the Argus shower counters," Nucl. Instr. Meth. A **237**, 464 (1985).

[66] R. Sinkus and T. Voss, "Particle identification with neural networks using a rotational invariant moment representation," Nucl. Instr. Meth. A **391**, 360 (1997).

[67] F. Fabozzi *et al.*, [*BABAR* Collaboration], "Muon identification in the *BABAR* experiment," *BABAR* Analysis Document **60**, v.01 (2000).

[68] S. M. Spanier and G. Mancinelli, [*BABAR* Collaboration], "Kaon selection at *BABAR*," *BABAR* Analysis Document **116** (2001).

[69] F. Martinez-Vidal and M. Carpinelli, [*BABAR* Collaboration], "The *BABAR* vertexing," *BABAR* Analysis Document **102**, v.06 (2001).

[70] F. Martinez-Vidal *et al.*, [*BABAR* Collaboration], "Vertexing and data quality studies for analyses using fully reconstructed B events are described in this supporting document for summer 2001 conferences," *BABAR* Analysis Document **254**, v.01 (2001).

[71] B. Aubert *et al.*, [*BABAR* Collaboration], "A study of time dependent CP-violating asymmetries and flavor oscillations in neutral B decays at the Υ(4S)," (2002), hep-ex/0201020, Submitted to Phys. Rev. D .

[72] J. Beringer and D. Payne, [*BABAR* Collaboration], "B tagging in *BABAR*: status for the $\sin 2\beta$ winter conference results," *BABAR* Analysis Document **119**, v.04 (2001).

[73] A. Snyder, "Dilution Resolution Correlation Explained," Vertexing and Composition Tools HyperNews September 2001, http://babar-hn.slac.stanford.edu:5090/HyperNews/get/VertexTools/227.html, (and others in same thread).

[74] C. Boros and Z. Liang, "Spin content of Lambda and its longitudinal polarization in e^+e^- annihilation at high energies," Phys. Rev. D **57**, 4491 (1998).

[75] A. L. and other, [*BABAR* Collaboration], "Studies on π^0 reconstruction: status report," *BABAR* Analysis Document **20**, v.03 (2000).

[76] J. C. Anjos *et al.*, [E691 Collaboration], "A Dalitz plot analysis of $D \to K\pi\pi$ decays," Phys. Rev. D **48**, 56 (1993).

[77] P. L. Frabetti *et al.*, [E687 Collaboration], "Analysis of three $D \to K\pi\pi$ Dalitz plots," Phys. Lett. B **331**, 217–226 (1994).

[78] R. H. Dalitz, Phil. Mag. **44**, 1068 (1953).

[79] G. Raven, "D^*-D^0 vertexing in $B \to D^*\ell\nu_\ell$ events," September 2000, `http://www.slac.stanford.edu/BFROOT/www/Organization/CollabMtgs/2000/detSep2000/Thu3a/raven.ps`, Presented at the *BABAR* Collaboration Meeting.

[80] P. Avery, "Directly determining lifetime using a 3-D Fit," (1999), `http://www.phys.ufl.edu/~avery/fitting/lifetime.ps`.

[81] F. Martinez-Vidal and M. Carpinelli, [*BABAR* Collaboration], "Vertexing performances and systematic checks with fully reconstructed B events," *BABAR* Analysis Document **130**, v.04 (2001).

[82] W. Verkerke and D. Kirkby, "RooFit toolkit for data modelling," (2002), `http://www.slac.stanford.edu/BFROOT/www/Computing/Offline/ROOT/RooFit/`.

[83] D. Kirkby *et al.*, [*BABAR* Collaboration], "A user's guide to the RooFitTools package for unbinned maximum likelihood fitting," *BABAR* Analysis Document **18** (2000).

[84] D. Kirkby and W. Verkerke, [*BABAR* Collaboration], "Core design of analysis modeling tools for ROOT," .

[85] R. Brun and F. Rademakers, "ROOT: An object-oriented data analysis framework," Nucl. Instr. Meth. A **389**, 81 (1997).

[86] F. James *et al.*, "MINUIT: Function minimization and error analysis," (1998), `http://wwwinfo.cern.ch/asdoc/minuit/`, CERN Program Library Long Writeup D506.

[87] C.-H. Cheng, "Status of lifetime and mixing with the semileptonic sample," June 2001, http://www.slac.stanford.edu/BFROOT/www/Organization/CollabMtgs/2001/detJun2001/Tues4a/cheng.pdf, Presentation at the *BABAR* Collaboration meeting.

[88] R. Cahn, B^0-$\overline{B}^0$ Mixing Analysis HyperNews January 2002, http://babar-hn.slac.stanford.edu:5090/HyperNews/get/Bmixing/330/2/1.html.

[89] W. Dunwoodie *et al.*, [*BABAR* Collaboration], "Study of Material Interactions with Gamma Conversions and Protons," .

[90] O. Leroy *et al.*, [LEP B Lifetimes Working Group Collaboration], "Averages for Amsterdam, ICHEP 2002," July 2002, http://lepbosc.web.cern.ch/LEPBOSC/, and references therein.

[91] B. Aubert *et al.*, [*BABAR* Collaboration], "Measurement of B^0-$\overline{B}^0$ flavor oscillations in hadronic B^0 decays," Phys. Rev. Lett. **88**, 221802 (2002), hep-ex/0112044.

[92] B. Aubert *et al.*, [*BABAR* Collaboration], "Measurement of the B^0-$\overline{B}^0$ oscillation frequency with inclusive dilepton events," Phys. Rev. Lett. **88**, 221803 (2002), hep-ex/0112045.

[93] B. Aubert *et al.*, [BABAR Collaboration], "Simultaneous measurement of B^0 meson lifetime and mixing frequency with $B^0 \to D^{*-}\ell^+\nu_\ell$ decays," (2002), hep-ex/0207071, Contributed to the 31^{st} International Conference on High Energy Physics, Amsterdam, The Netherlands.

[94] T. Tomura *et al.*, [Belle Collaboration], "Measurement of the oscillation frequency for B^0-$\overline{B}^0$ mixing using hadronic B^0 decays," hep-ex/0207022.

[95] K. Hara *et al.*, [Belle Collaboration], "Mesurement of the B^0-$\overline{B}^0$ mixing parameter Δm_d using semileptonic B^0 decays," hep-ex/0207045.

[96] K. Abe *et al.*, [Belle Collaboration], "Measurement of B^0-$\overline{B}^0$ mixing rate with $B^0(\overline{B}^0) \to D^{*\mp}\pi^{\pm}$ partial reconstruction," BELLE-CONF-0204.

[97] K. Abe *et al.*, [Belle Collaboration], "Measurement of B/d0 - anti-B/d0 mixing rate from the time evolution of dilepton events at the Upsilon(4S)," Phys. Rev. Lett. **86**, 3228–3232 (2001), hep-ex/0011090.

[98] L. Di Ciaccio *et al.*, [LEP B Oscillations Working Group Collaboration], "Averages for Summer 2002," July 2002, http://lepbosc.web.cern.ch/LEPBOSC/lifetimes/lepblife.html, and references therein.

[99] B. Aubert *et al.*, [*BABAR* Collaboration], "Measurement of the B^0 lifetime with partially reconstructed $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_\ell$ decays," Phys. Rev. Lett. **89**, 011802 (2002), hep-ex/0202005.

[100] D. Kirkby and R. Cahn, [*BABAR* Collaboration], "A general model for neutral B decay time distributions," *BABAR* Analysis Document **188**, v.05 (2002).

[101] J. D. Jackson, "Remarks on the phenomenological analysis of resonances," Nuovo Cim. **34**, 1644 (1964).

[102] B. J. M. and W. V.F., *Theoretical Nuclear Physics*, John Wiley & Sons, New York (1952).